9-2004

MAP Estimation for Hyperspectral Image Resolution Enhancement Using an Auxiliary Sensor

Russell C. Hardie
University of Dayton, rhardie1@udayton.edu

Michael T. Eismann
Air Force Research Laboratory

Gregory L. Wilson
Mission Research Corp.

MAP Estimation for Hyperspectral Image Resolution Enhancement Using an Auxiliary Sensor

Russell C. Hardie[†], *Senior Member, IEEE,* Michael T. Eismann[‡], *Member, IEEE,* and Gregory L. Wilson [§]

[†]Department of Electrical and Computer Engineering and Electro-Optics Program

University of Dayton, Dayton, Ohio, 45459–0226 USA

Telephone: (937) 229–3178, Fax: (937) 229–2097

Email: rhardie@udayton.edu

[‡]Sensors Directorate, Air Force Research Laboratory

2241 Avionics Circle, Wright-Patterson AFB OH, 45433-7700 USA

Telephone: (937) 255–9902 $\times$4405, Fax: (937) 255–2959

Email: Michael.Eismann@wpafb.af.mil

[§] Mission Research Corporation

3975 Research Blvd., Dayton OH, 45430 USA

Telephone: (937) 429–9261, Fax: (937) 429–2625

Email: gwilson@mrcday.com

Abstract

This paper presents a novel maximum *a posteriori* (MAP) estimator for enhancing the spatial resolution of an image using co-registered high spatial-resolution imagery from an auxiliary sensor. Here we focus on the use of high-resolution panchomatic data to enhance hyperspectral imagery. However, the estimation framework developed allows for any number of spectral bands in the primary and auxiliary image. The proposed technique is suitable for applications where some correlation, either localized or global, exists between the auxiliary image and the image being enhanced. To exploit localized correlations, a spatially varying statistical model, based on vector quantization, is used. Another important aspect of the proposed algorithm is that it allows for the use of an accurate observation model relating the "true" scene with the low-resolutions observations. Experimental results with hyperspectral data derived from the Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) are presented to demonstrate the efficacy of the proposed estimator.

Index Terms

Hyperspectral, Multispectral, MAP estimation, Resolution Enhancement, Multisensor, Panchromatic Sharpening.

EDICS Category: 2-COLO, 2-MLTF

MAP Estimation for Hyperspectral Image Resolution Enhancement Using an Auxiliary Sensor

I. Introduction

Recently there has been significant interest in hyperspectral sensing technology to support a variety of civilian, commercial, and military remote sensing applications. The wealth of information that resides in the spectral domain provides significant advantages over traditional panchromatic and multispectral imagery, particularly for ground cover and material identification and classification. However, the design and development of practical hyperspectral sensors often result in a significant trade-off in spatial resolution. Therefore, important spatial features such as shape and texture can be lost and the spatial fidelity of the resulting hyperspectral products is reduced.

The inherent trade-off between spatial and spectral resolution has resulted in the development of remote sensing systems that include low-resolution hyperspectral coupled with high-resolution panchromatic and/or multispectral imaging subsystems. An example is the NASA Earth Observer 1 satellite, which includes a 30 m hyperspectral sensor and a 10 m panchromatic imager. Commercial panchromatic satellite imagery approaching 1 m spatial resolution is also available. This provides the opportunity to jointly process the hyperspectral and higher resolution panchromatic imagery to potentially achieve improved detection and/or classification performance.

A variety of techniques have been presented in the literature for merging imagery of different spatial and spectral resolution [1]–[15]. Many of these techniques have been designed to sharpen multispectral imagery for human interpretation using broadband panchromatic data. Component substitution methods transform the multispectral imagery and replace one component with the broadband high-resolution imagery [3], [5], [16]. Commonly used transformations are intensity-hue-saturation (IHS), where the intensity is replaced, or principal component analysis (PCA), where the first principal component is replaced. Clearly, information in the lower components, that may be critical in classification and detection, is not enhanced with such an approach.

High pass techniques add high spatial frequency content from the high-resolution image to the bands of the low-resolution data [1], [4], [7], [8], [16]. The technique in [11] uses a statistical approach that adds a linear combination of the high-resolution data to a pixel replicated version of the low-resolution imagery. This technique was designed to enhance the spatial resolution of the Landsat Thematic Mapper thermal band from the remaining bands. The method described in [2] and [6] models the relationship between the high-resolution image and the image to be estimated. Regression techniques are used at the available lower resolution to obtain the model parameters.

Another approach to hyperspectral resolution enhancement is based on spectral mixture analysis. A method based on the linear mixing model has been investigated that employs constrained nonlinear optimization techniques to obtain high resolution endmember fractions [12]–[14]. An alternative approach appends the high resolution image to the hyperspectral data and computes a mixture model based on the joint data set [15]. A high resolution hyperspectral image, however, is not explicitly estimated.

In this paper, we describe a novel maximum *a posteriori* (MAP) estimation framework [17] for enhancing the spatial resolution of an image using co-registered high spatial-resolution imagery from an auxiliary sensor. Here we focus on the use of high-resolution panchomatic data to enhance hyperspectral imagery. However, the estimation framework developed allows for any number of spectral bands in the primary and auxiliary sensor. The proposed technique is suitable for applications where some correlation, either localized or global, exists between the auxiliary image and the image being enhanced. A spatially varying statistical model is used to help exploit localized correlation between the primary and auxiliary image. Another important aspect of the proposed algorithm is that it allows for the use of an accurate observation model relating the “true” scene with the low-resolutions observations. This means that a potentially wavelength-dependent spatially-varying system point spread function (PSF) can be incorporated into the estimator.

The remainder of this paper is organized as follows. In Section II, the observation model relating the true high-resolution hyperspectral image to the observed low-resolution data is formulated. In Section III, the MAP framework and estimator is presented. Section IV examines methods for estimating the statistical parameters required in forming the MAP estimate. The relationship of the proposed method to some prior methods is discussed in Section V. Experimental results are presented in Section VI. Finally, conclusions are given in Section VII.

II. Observation Model

Consider the desired high-resolution hyperspectral image having P bands with N pixels per band. This hyperspectral data-cube in band-interleaved-by-pixel lexicographical notation will be denoted

$$\mathbf{z} = [z_{1,1}, ..., z_{P,1}, z_{1,2}, ..., z_{P,2}, ..., z_{1,N}, ..., z_{P,N}]^T. \tag{1}$$

These variables can be expressed as $\mathbf{z} = \left[\mathbf{z}_1^T, \mathbf{z}_2^T, ..., \mathbf{z}_N^T\right]^T$, where $\mathbf{z}_n = [z_{1,n}, z_{2,n}, ..., z_{P,n}]^T$ represents the spectral response pattern at each spatial position n, for $n = 1, 2, ..., N$. The vectors, $\mathbf{z}_n$, will be referred to as "hyperpixels."

In practice, the available co-registered imagery is generally either panchromatic or multispectral. Let the number of bands in this associated data set be Q. In lexicographical form this data set is denoted

$$\mathbf{x} = [x_{1,1}, ..., x_{Q,1}, x_{1,2}, ..., x_{Q,2}, \ldots, x_{1,N}, ..., x_{Q,N}]^T. \tag{2}$$

Expressing these data according to hyperpixel index, we get $\mathbf{x} = \left[\mathbf{x}_1^T, \mathbf{x}_2^T, ..., \mathbf{x}_N^T\right]^T$, where $\mathbf{x}_n = [x_{1,n}, x_{2,n}, ..., x_{Q,n}]^T$, for $n = 1, 2, ..., N$. The observed low-resolution hyperspectral data cube, having $M < N$ pixels per band, is denoted

$$\mathbf{y} = [y_{1,1}, ..., y_{P,1}, y_{1,2}, ..., y_{P,2}, ..., y_{1,M}, ..., y_{P,M}]^T. \tag{3}$$

Expressing $\mathbf{y}$ in terms of low-resolution hyperpixels, we obtain $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_M^T\right]^T$, where $\mathbf{y}_m = [y_{1,m}, y_{2,m}, ..., y_{P,m}]^T$, for $m = 1, 2, ..., M$.

We model the observed hyperspectral image with the following linear observation model

$$\mathbf{y} = \mathbf{W}\mathbf{z} + \mathbf{n}. \tag{4}$$

Here $\mathbf{W}$ is an $MP \times NP$ matrix which, when operating on $\mathbf{z}$, performs the PSF blurring and down-sampling. In many cases the PSF can be determined theoretically by taking into account the spatial integration of the detectors and modeling the optics [18]. The vector $\mathbf{n}$ is a random vector of noise samples introduced during image acquisition. We will assume that $\mathbf{n}$ is a zero-mean Gaussian random vector that is independent of $\mathbf{z}$ and $\mathbf{x}$. Specifically, the

probability density function (pdf) for the noise vector is given by

$$p(\mathbf{n}) = \frac{1}{\sqrt{(2\pi)^{MP}|\mathbf{C_n}|}} \exp\left\{-\frac{1}{2}\mathbf{n}^T\mathbf{C_n}^{-1}\mathbf{n}\right\}, \tag{5}$$

where $\mathbf{C_n}$ is the $PM \times PM$ covariance matrix for the noise. For notational simplicity, all the pdfs defined here will be distinguished only by their arguments (in this case $\mathbf{n}$). In many applications it is reasonable to model the noise samples as independent and identically distributed from pixel-to-pixel and band-to-band. In that case, the noise covariance matrix is given by $\mathbf{C_n} = \sigma_n^2\mathbf{I}$. Note that the total number of independent linear equations defined by our observation model is PM and the number of unknowns in $\mathbf{z}$ is PN. Thus, without the additional information provided by $\mathbf{x}$ and *a priori* knowledge about $\mathbf{z}$, the problem of estimating $\mathbf{z}$ is ill-posed.

III. MAP Estimation

A. MAP Estimation Framework

Treating $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ as random vectors, the MAP estimate of $\mathbf{z}$, given we observe $\mathbf{x}$ and $\mathbf{y}$, is defined as

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} p(\mathbf{z}|\mathbf{x},\mathbf{y}), \tag{6}$$

where $p(\mathbf{z}|\mathbf{x},\mathbf{y})$ represents the conditional probability density function (pdf) of $\mathbf{z}$ given $\mathbf{x}$ and $\mathbf{y}$. In other words, the MAP estimate is given by the vector $\mathbf{z}$ that maximizes $p(\mathbf{z}|\mathbf{x},\mathbf{y})$. Using Bayes rule we can write

$$p(\mathbf{z}|\mathbf{x},\mathbf{y}) = \frac{p(\mathbf{x},\mathbf{y}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{x},\mathbf{y})}. \tag{7}$$

Based on the model in (4) it can be seen that $\mathbf{y}$, given knowledge of $\mathbf{z}$, is simply $\mathbf{n}$ plus a constant vector. In our observation model we have assumed that the primary sensor noise, $\mathbf{n}$, is independent of both $\mathbf{x}$ and $\mathbf{z}$. Thus, it follows that $\mathbf{x}$ and $\mathbf{y}$, conditioned on knowledge of $\mathbf{z}$, are independent. This allows us to express (7) as

$$p(\mathbf{z}|\mathbf{x},\mathbf{y}) = \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{y}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{x},\mathbf{y})}. \tag{8}$$

Using Bayes rule on $p(\mathbf{x}|\mathbf{z})$ yields

$$p(\mathbf{z}|\mathbf{x},\mathbf{y}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})p(\mathbf{y}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{z})p(\mathbf{x},\mathbf{y})}. \tag{9}$$

Lastly, cancelling $p(\mathbf{z})$ and noting that $p(\mathbf{x})$ and $p(\mathbf{x},\mathbf{y})$ are not functions of $\mathbf{z}$, yields an alternative and equivalent expression for the MAP estimator given by

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} p(\mathbf{y}|\mathbf{z})p(\mathbf{z}|\mathbf{x}). \tag{10}$$

Specifying the required pdfs is addressed in the following section.

B. Specifying the Probability Density Functions

The pdf, $p(\mathbf{y}|\mathbf{z})$, comes from the observation model in (4). In particular, given $\mathbf{z}$, the random vector $\mathbf{y}$ has a mean of $\mathbf{Wz}$ and variation due only to the noise $\mathbf{n}$. Thus, using (5) we obtain

$$p(\mathbf{y}|\mathbf{z}) = \frac{1}{\sqrt{(2\pi)^{MP}|\mathbf{C_n}|}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{Wz})^T\mathbf{C}_\mathbf{n}^{-1}(\mathbf{y}-\mathbf{Wz})\right\}. \tag{11}$$

The role that this pdf plays in the final algorithm will be to constrain any estimate of $\mathbf{z}$ to closely (depending on the level of noise) satisfy the observation model equation $\mathbf{y} = \mathbf{Wz}$.

Let us now consider the conditional pdf, $p(\mathbf{z}|\mathbf{x})$. If we assume that $\mathbf{x}$ and $\mathbf{z}$ are jointly Gaussian, then by Theorem 10.2 of [17], the conditional pdf is also Gaussian with the following form

$$p(\mathbf{z}|\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{NP}|\mathbf{C}_{\mathbf{z}|\mathbf{x}}|}} \exp\left\{-\frac{1}{2}(\mathbf{z}-\mu_{\mathbf{z}|\mathbf{x}})^T\mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}(\mathbf{z}-\mu_{\mathbf{z}|\mathbf{x}})\right\}, \tag{12}$$

where $\mu_{\mathbf{z}|\mathbf{x}} = \mathrm{E}\{\mathbf{z}|\mathbf{x}\}$ is the expected value of $\mathbf{z}$ given $\mathbf{x}$ and $\mathbf{C}_{\mathbf{z}|\mathbf{x}}$ is the $NP \times NP$ covariance matrix of $\mathbf{z}$ given $\mathbf{x}$. These conditional parameters can be related to the joint statistics, which are generally easier to estimate [17]. In particular, the conditional mean is given by

$$\mu_{\mathbf{z}|\mathbf{x}} = \mathrm{E}\{\mathbf{z}\} + \mathbf{C}_{\mathbf{z},\mathbf{x}}\mathbf{C}_{\mathbf{x},\mathbf{x}}^{-1}\left[\mathbf{x} - \mathrm{E}\{\mathbf{x}\}\right] \tag{13}$$

and the conditional covariance is given by

$$\mathbf{C}_{\mathbf{z}|\mathbf{x}} = \mathbf{C}_{\mathbf{z},\mathbf{z}} - \mathbf{C}_{\mathbf{z},\mathbf{x}}\mathbf{C}_{\mathbf{x},\mathbf{x}}^{-1}\mathbf{C}_{\mathbf{z},\mathbf{x}}^{T}, \tag{14}$$

where $\mathrm{E}\{\mathbf{x}\}$ represents the mean of $\mathbf{x}$, $\mathrm{E}\{\mathbf{z}\}$ represents the mean of $\mathbf{z}$. The variables $\mathbf{C}_{\mathbf{z},\mathbf{z}}$, $\mathbf{C}_{\mathbf{z},\mathbf{x}}$, and $\mathbf{C}_{\mathbf{x},\mathbf{x}}$ are

cross-covariance matrices of the form

$$\mathbf{C}_{\mathbf{u},\mathbf{v}} = \mathrm{E}\left\{ (\mathbf{u} - \mathrm{E}\{\mathbf{u}\}) (\mathbf{v} - \mathrm{E}\{\mathbf{v}\})^T \right\}, \tag{15}$$

where $\mathbf{u}$ and $\mathbf{v}$ are arbitrary multivariate random vectors.

C. The MAP Cost Function

To maximize the expression in (10), we can alternatively minimize the sum of the exponential terms in (11) and (12). Thus, the MAP estimator can be expressed as

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} C(\mathbf{z}), \tag{16}$$

where $C(\mathbf{z})$ is a cost function given by

$$C(\mathbf{z}) = \frac{1}{2}(\mathbf{y} - \mathbf{W}\mathbf{z})^T \mathbf{C}_{\mathbf{n}}^{-1} (\mathbf{y} - \mathbf{W}\mathbf{z}) + \frac{1}{2}(\mathbf{z} - \mu_{\mathbf{z}|\mathbf{x}})^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1} (\mathbf{z} - \mu_{\mathbf{z}|\mathbf{x}}). \tag{17}$$

Note that first term is the result of the observation model constraint and the second term comes from the correlation between the auxiliary image and the desired image. The first term is minimized when $\mathbf{z}$ is selected so that when it is degraded, according to the observation model, the result matches the observed low-resolution hyperspectral image data. The second term is minimized when $\mathbf{z} = \mu_{\mathbf{z}|\mathbf{x}}$. These two competing constraints are intuitive. We want our estimate to exploit the correlation with $\mathbf{x}$, and still be consistent with the observation model.

To derive the closed form solution, we begin by computing the gradient of the cost in (17) with respect to $\mathbf{z}$. This is given by

$$\nabla_{\mathbf{z}} C(\mathbf{z}) = \mathbf{W}^T \mathbf{C}_{\mathbf{n}}^{-1} \mathbf{W}\mathbf{z} - \mathbf{W}^T \mathbf{C}_{\mathbf{n}}^{-1} \mathbf{y} + \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1} (\mathbf{z} - \mu_{\mathbf{z}|\mathbf{x}}). \tag{18}$$

Setting the gradient to zero and solving for $\mathbf{z}$ yields

$$\hat{\mathbf{z}} = \left[\mathbf{W}^T \mathbf{C}_{\mathbf{n}}^{-1} \mathbf{W} + \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1} \right]^{-1} \left[\mathbf{W}^T \mathbf{C}_{\mathbf{n}}^{-1} \mathbf{y} + \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1} \mu_{\mathbf{z}|\mathbf{x}} \right]. \tag{19}$$

Note also that if the noise covariance goes to zero, this form is invalid. Applying the matrix inversion lemma [17]

and simplifying the result yields

$$\hat{\mathbf{z}} = \mu_{\mathbf{z}|\mathbf{x}} + \mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T \left[\mathbf{W}\mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T + \mathbf{C}_{\mathbf{n}}\right]^{-1} \left[\mathbf{y} - \mathbf{W}\mu_{\mathbf{z}|\mathbf{x}}\right]. \tag{20}$$

This is, perhaps, a more useful result as it requires the inversion of an $MP \times MP$ matrix $\mathbf{W}\mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T + \mathbf{C}_{\mathbf{n}}$ rather than the $NP \times NP$ matrix $\mathbf{C}_{\mathbf{z}|\mathbf{x}}$ in the previous form. Furthermore, when the noise in the observation model is set to zero, (20) reduces to a useful result that is equivalent to that obtained with constrained optimization, as shown in Appendix A.

It is interesting to note that if $\mathbf{C}_{\mathbf{z}|\mathbf{x}}$ is positive definite, then $\mathbf{W}\mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T + \mathbf{C}_{\mathbf{n}}$ is positive definite for any $\mathbf{W}$, provided that $\mathbf{C}_{\mathbf{n}}$ is positive definite (e.g., $\mathbf{C}_{\mathbf{n}} = \sigma_n^2\mathbf{I}$, where $\sigma_n^2 > 0$). Another important point is that these closed form solutions may require the inversion of a very large matrix. However, by making assumptions about the structure of the covariance matrices and the system PSF, the problem can be decomposed spatially into independent subsets. This greatly reduces the computational and memory demands of the closed form solutions. In other cases where such decomposition is not possible, iterative minimization techniques can be employed such as those described in [18], [19]. These methods may not require a matrix inverse and they can more readily exploit the sparse nature of $\mathbf{W}$ inherent in most practical image formation models.

IV. Estimating the Required Statistical Parameters

Estimating the full $NP \times NP$ covariance matrix $\mathbf{C}_{\mathbf{z}|\mathbf{x}}$, used in (12), is impractical for a typical size hyperspectral image. To make the problem manageable, constraints on the form of this covariance are required to bring down the number of statistical parameters to be estimated. While there may be numerous ways to accomplish this, we believe that a reasonable approach is to model the unknown high-resolution hyperpixels as conditionally independent spatially, yielding

$$p(\mathbf{z}|\mathbf{x}) = \prod_{n=1}^{N} p(\mathbf{z}_n|\mathbf{x}_n). \tag{21}$$

Writing out the individual conditional pdfs yields

$$p(\mathbf{z}|\mathbf{x}) = \prod_{n=1}^{N} \frac{1}{\sqrt{(2\pi)^P |\mathbf{C}_{\mathbf{z}_n|\mathbf{x}_n}|}} \exp\left\{-\frac{1}{2}(\mathbf{z}_n - \mu_{\mathbf{z}_n|\mathbf{x}_n})^T \mathbf{C}^{-1}_{\mathbf{z}_n|\mathbf{x}_n}(\mathbf{z}_n - \mu_{\mathbf{z}_n|\mathbf{x}_n})\right\}, \tag{22}$$

where $\mu_{\mathbf{z}_n|\mathbf{x}_n} = \mathrm{E}\{\mathbf{z}_n|\mathbf{x}_n\}$ and $\mathbf{C}_{\mathbf{z}_n|\mathbf{x}_n}$ is the $P \times P$ covariance matrix for $\mathbf{z}_n$ given $\mathbf{x}_n$. The relationship between the individual hyperpixel statistical parameters and the global statistical parameters in (12) is given by

$$\mathbf{C}_{\mathbf{z}|\mathbf{x}} = \begin{bmatrix} \mathbf{C}_{\mathbf{z}_1|\mathbf{x}_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\mathbf{z}_2|\mathbf{x}_2} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{C}_{\mathbf{z}_N|\mathbf{x}_N} \end{bmatrix} = \bigoplus_{n=1}^{N} \mathbf{C}_{\mathbf{z}_n|\mathbf{x}_n}, \tag{23}$$

where $\bigoplus$ represents a direct sum and

$$\mu_{\mathbf{z}|\mathbf{x}} = \left[\mu^T_{\mathbf{z}_1|\mathbf{x}_1}, \mu^T_{\mathbf{z}_2|\mathbf{x}_2}, \ldots, \mu^T_{\mathbf{z}_N|\mathbf{x}_N}\right]^T. \tag{24}$$

Furthermore, applying the results in (13) and (14) on the hyperpixels yields

$$\mu_{\mathbf{z}_n|\mathbf{x}_n} = \mathrm{E}\{\mathbf{z}_n\} + \mathbf{C}_{\mathbf{z}_n,\mathbf{x}_n}\mathbf{C}^{-1}_{\mathbf{x}_n,\mathbf{x}_n}\left[\mathbf{x}_n - \mathrm{E}\{\mathbf{x}_n\}\right] \tag{25}$$

and

$$\mathbf{C}_{\mathbf{z}_n|\mathbf{x}_n} = \mathbf{C}_{\mathbf{z}_n,\mathbf{z}_n} - \mathbf{C}_{\mathbf{z}_n,\mathbf{x}_n}\mathbf{C}^{-1}_{\mathbf{x}_n,\mathbf{x}_n}\mathbf{C}^T_{\mathbf{z}_n,\mathbf{x}_n}. \tag{26}$$

Note that the covariance matrix of the joint random vector, $\mathbf{s}_n = \left[\mathbf{x}_n^T, \mathbf{z}_n^T\right]^T$, is related to the cross-covariance matrices in (25) and (26) as follows

$$\mathbf{C}_{\mathbf{s}_n} = \begin{bmatrix} \mathbf{C}_{\mathbf{x}_n,\mathbf{x}_n} & \mathbf{C}^T_{\mathbf{z}_n,\mathbf{x}_n} \\ \mathbf{C}_{\mathbf{z}_n,\mathbf{x}_n} & \mathbf{C}_{\mathbf{z}_n,\mathbf{z}_n} \end{bmatrix}. \tag{27}$$

With the simplifying assumption of conditional independence, $\mathbf{C}_{\mathbf{z}|\mathbf{x}}$ is block diagonal, reducing the number of statistical parameters from $(NP)^2$ to NP^2. That is, the problem reduces to estimating the hyperpixel conditional statistics in (25) and (26). We may wish to further reduce the number of statistical parameters to be estimated by treating some or all hyperpixels as having the same conditional covariance. If, for example, we assume that all hyperpixels have the same covariance, then we have only P^2 statistical parameters in the conditional covariance matrix to estimate. One area of future work might be to explore other forms for the conditional covariance matrix (implying other assumptions regarding the nature of the high-resolution hyperpixels).

To estimate $\mathrm{E}\{\mathbf{z}_n\}$, we propose using the spatially interpolated observed hyperspectral imagery, denoted $\hat{\mu}_{\mathbf{z}_n}$.

We have found that spline interpolation tends to yield the best results here. To estimate $\mathrm{E}\{\mathbf{x}_n\}$, we use a spatially smoothed version of the band or bands in $\mathbf{x}$. We have observed that the best results are obtained when the smoothing is done to mimic the way in which $\hat{\mu}_{\mathbf{z}_n}$ relates to $\mathbf{z}_n$. That is, we degrade $\mathbf{x}_n$ with a PSF and down-sampling factor similar to that defined by $\mathbf{W}$. This degraded image is interpolated using spline interpolation to produce the estimate of $\mathrm{E}\{\mathbf{x}_n\}$, which we will denote $\hat{\mu}_{\mathbf{x}_n}$, for all n. In this fashion it may be said that $\hat{\mu}_{\mathbf{x}_n}$ relates to $\mathbf{x}_n$ as $\hat{\mu}_{\mathbf{z}_n}$ relates to $\mathbf{z}_n$. These estimates tend to track the non-stationarity in the mean exhibited by most natural images [20].

The joint covariance in (27) must also be estimated in order to get the conditional statistics in (25) and (26). In most cases it will not be possible to obtain statistically similar data at the high resolution required. Thus, we attempt to estimate the required joint covariance from the observed imagery. To do so, we will estimate a joint covariance at the lower resolution of the observed hyperspectral imagery and apply it at the higher resolution. While the joint statistics may differ at different resolutions, it is hoped that there is sufficient symmetry of spatial scale in the statistical parameters to provide a useful result. In particular, we artificially degrade the spatial resolution and size of $\mathbf{x}$ to match that of $\mathbf{y}$. Let this degraded image be denoted $\tilde{\mathbf{x}} = \left[\tilde{\mathbf{x}}_1^T, \tilde{\mathbf{x}}_2^T, ..., \tilde{\mathbf{x}}_M^T\right]^T$. The local means at this resolution, obtained using the same method applied to the original resolution, are removed from $\tilde{\mathbf{x}}$ and $\mathbf{y}$. Now, the joint covariance information is estimated. One relatively simple approach seeks a single global covariance using a sample covariance estimate. This covariance is used as an estimate of $\mathbf{C}_{\mathbf{s}_n}$ for all n.

However, in order to more fully exploit the information in $\mathbf{x}$, we wish to capture the changing joint covariance as the spectral content in the scene varies spatially. Since it is impractical to estimate a joint covariance for each hyperpixel, we use a simple clustering approach based on vector quantization. To begin, we form joint vector $\tilde{\mathbf{s}}_m = \left[\tilde{\mathbf{x}}_m^T, \mathbf{y}_m^T\right]^T \in \mathcal{R}^{\mathcal{Q}+\mathcal{P}}$ for $m = 1, 2, \ldots, M$, where $\mathcal{R}^{\mathcal{Q}+\mathcal{P}}$ represents the $Q + P$ dimensional real space. These vectors are grouped into K clusters (or classes) using the Linde-Buzo-Gray (LBG) algorithm [21]. The cluster centroids define the Voronoi partitions of the spectral space. That is, a given vector in $\mathcal{R}^{\mathcal{Q}+\mathcal{P}}$ space is assigned to class k if it lies closest, in a Euclidean sense, to the centroid of cluster k. For each cluster, the sample joint covariance is computed using $\tilde{\mathbf{s}}_m$ for $m \in \Omega_k$, where Ω_k is the set of all indices corresponding to Voronoi partition k. To estimate the joint covariance, $\mathbf{C}_{\mathbf{s}_n}$, we assign $\hat{\mathbf{s}}_n = \left[\mathbf{x}_n^T, \hat{\mu}_{\mathbf{z}_n}^T\right]^T$ to a partition and let the covariance for this high-resolution spatial position n be the corresponding cluster covariance. With these covariance estimates in hand, the final MAP estimate can be formed. In some cases we have observed improvement in performance if we

form an estimate of the conditional mean in (25), $\hat{\mu}_{\mathbf{z}_n|\mathbf{x}_n}$, and then re-classify the image using $\hat{\mathbf{s}}_n = [\mathbf{x}_n^T, \hat{\mu}_{\mathbf{z}_n|\mathbf{x}_n}^T]^T$ for the purpose of assigning cluster covariances to each high-resolution position.

V. Relationship to Other Approaches

It is interesting to explore the relationship between the proposed MAP estimate and some previously proposed approaches. Consider that if one neglects the observation model for $\mathbf{y}$ in the MAP formulation, the resulting cost function would simply be the second term in (17). This would lead to an estimate that is the conditional mean in (25). This result is essentially the same as that derived by Nishii *et al.* [11] for Landsat Thematic Mapper thermal band estimation, given specific choices for the estimates of $\mathrm{E}\{\mathbf{z}_n\}$ and $\mathrm{E}\{\mathbf{x}_n\}$. In particular, one must use zero-order-hold (ZOH) interpolation (i.e., pixel replication) on each band of the observed hyperspectral image to estimate $\mathrm{E}\{\mathbf{z}_n\}$ and average the auxiliary image pixels within the span of each low-resolution hyperpixel to estimate $\mathrm{E}\{\mathbf{x}_n\}$. Using these estimates for $\mathrm{E}\{\mathbf{z}_n\}$ and $\mathrm{E}\{\mathbf{x}_n\}$ guarantees that the average of the estimated hyperpixels within the span of a low-resolution hyperpixel will be equal to the low-resolution hyperpixel. Nishii *et al.* [11] explore the use of both local and global covariances. Thus, in comparison to their method, the proposed MAP framework is novel in how it explicitly incorporates an arbitrary system PSF and in how it allows for various statistical models and estimates of the statistical parameters. We will use the method of Nishii *et al.* [11] as one of our performance benchmarks for comparison in Section VI.

Another method used as a benchmark is the estimator proposed by Price [2], [6]. This method was designed to combine multispectral imagery with a panchromatic auxiliary image (i.e., $Q = 1$). For bands strongly correlated with the panchromatic sensor, the estimate is based on a linear mapping of the panchromatic image, yielding

$$\hat{z}_{p,n} = a_p x_{1,n} + b_p, \tag{28}$$

for $n = 1, 2, ..., N$ and $p = 1, 2, ..., P$. This estimate is then scaled so that the average of the high-resolution hyperpixels within the span of a low-resolution hyperpixel is equal to the low-resolution hyperpixel. The coefficients, a_p and b_p are estimated with least-squares regression using low-resolution hyperspectral band p and a degraded version of the panchromatic image (degraded to match the spatial resolution of the low-resolution hyperspectral band). For weakly correlated bands, a look-up-table (LUT) method is employed [2], [6]. The LUT is created based

on the relationship between the low-resolution pixel values in a given band and the corresponding pixel values in the degraded panchromatic image. Once the LUT is generated it is applied to the full-resolution panchromatic image to form a high resolution estimate of the desired band. As before, this estimate is scaled so that the average of the high-resolution hyperpixels within the span of a low-resolution hyperpixel is equal to the low-resolution hyperpixel.

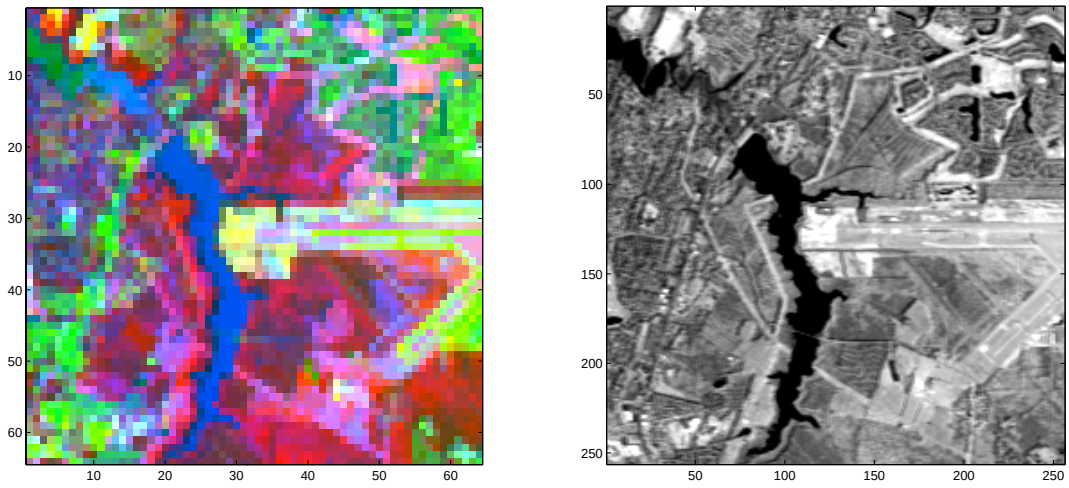
VI. Experimental Results

In this section we present a number of experimental results in order to demonstrate the efficacy of the proposed estimator in comparison to the benchmark techniques. Simulated data are used here to allow for quantitative performance analysis. The details of the data set are provided in Section VI-A. In Sections VI-B and VI-C, quantitative error analysis is presented in the spectral band space and in the principal component space, respectively. Finally, noise analysis is presented in Section VI-D.

A. Simulated Data

The simulated data sets are derived from a hyperspectral image collected by the Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) sensor [22]. AVIRIS is a scanning dispersive hyperspectral imaging sensor that flies on the NASA ER-2 aircraft at approximately 20 km above sea level with a spatial resolution of approximately 6 m per pixel. The sensor collects 224 contiguous spectral bands in the range of 0.4 to 2.5 μm. The specific scene used has been collected over Yorktown Virginia (Flight F980703T01, Run 02, ID 1828000ST23).

A 256×256 portion of the scene is used as the true high spatial resolution hyperspectral image $\mathbf{z}$. This is artificially degraded to form $\mathbf{y}$. A simple rectangular detector model is used for the system PSF [18]. In particular, the PSF is a 4×4 kernel with equal weights of $1/16$. The image is subsampled by a factor of 4 in both spatial dimensions. This PSF model leads to a simple structure in the matrix $\mathbf{W}$. This structure, combined with the spatial conditional independence assumption, allows us to process each low-resolution hyperpixel to form a corresponding 4×4 set of hyperpixels independently (after the mean estimates are formed using interpolation). For imagery in the mid- and long-wave infrared, diffraction effects tend to become more pronounced and can be added to the observation model [18]. The associated high resolution sensor in this case is modeled as a panchromatic broadband imager ($Q = 1$). These data are formed by averaging the 224 AVIRIS bands at the original resolution. A false



(a) (b)

Fig. 1. Simulated observed images derived from AVIRIS data. (a) False color image of principal components one, two, and three of the low spatial resolution hyperspectral data. (b) High spatial resolution panchromatic image.

color image, formed by mapping the first three principal components of the low-resolution hyerpspectral data to red, green and blue, respectively, is shown in Fig. 1(a). The simulated broadband image is shown in Fig. 1(b).

The eigenvalue (variance) associated with each principal component of the low-resolution hyperspectral data is plotted in Fig. 2. This clearly indicates that the vast majority of signal power is contained in the leading components. For example, after 20 components, the eigenvalue has dropped by approximately five orders of magnitude from the top component. In order to reduce the computational burden, we process the imagery in the PCA space in the top twenty dimensions. The lower 204 dimensions are processed using spline interpolation. The processed components are then transformed back to the original spectral space. Note that due to the nature of the PCA transformation, the estimation algorithm applies identically in the principal component space as shown in Appendix B.

B. Spectral Space Performance Analysis

To quantitatively assess the performance of the MAP estimator, we compare the estimates with the "true" hyperspectral image (the original resolution AVIRIS image). Our image fidelity metric is signal-to-noise ratio (SNR), where "noise" here refers to estimation error. This metric is computed as the sample variance of the "desired" image divided by the mean squared error (MSE). Scaling the reciprocal of the MSE by the variance of the desired image is helpful in allowing one to compare performance between bands with significantly different signal powers. This

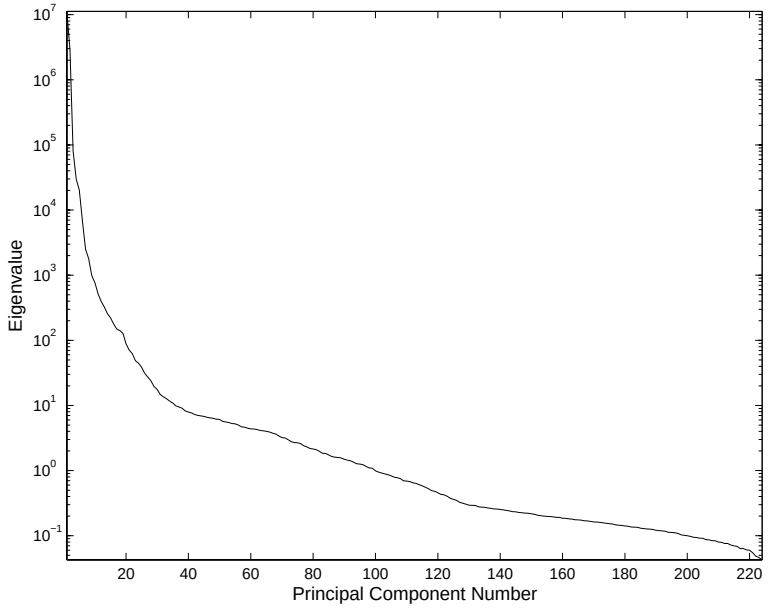


Fig. 2. Eigenvalue versus component number for the low-resolution hyperspectral image.

is particularly useful in principal component space, where power in the bands can vary by orders of magnitude.

The SNR versus wavelength is shown in Fig. 3 for the MAP estimator ($K = 16$), the method of Nishii *et al.* (with global covariance statistics) [11], Price's method [2], and spline interpolation. For Price's method, both the linear model and LUT approach are used and the best of the two SNRs for each band is reported. Here no noise is introduced to either the low-resolution hyperspectral imagery or the panchromatic imagery. The effects of noise are studied in Section VI-D. Note that significant improvement over spline interpolation is obtained in many bands with all the techniques. Not surprisingly, the bands with the highest correlation with the panchromatic image tend to have the highest SNRs. The spectral band estimates with very low SNR are a result of the original data having very low signal power due to atmospheric absorption. It can be seen from Fig. 3 that the MAP estimate with ($K = 16$) provides the highest SNRs for these data.

C. Principal Component Space Performance Analysis

The spectral domain error analysis indicates that many spectral bands can be significantly enhanced with the use of the panchromatic imagery. However, it is insightful to examine the performance in the principal component space. Table I shows the SNR in the first 5 principal components for spline interpolation, the method of Nishii *et al.*

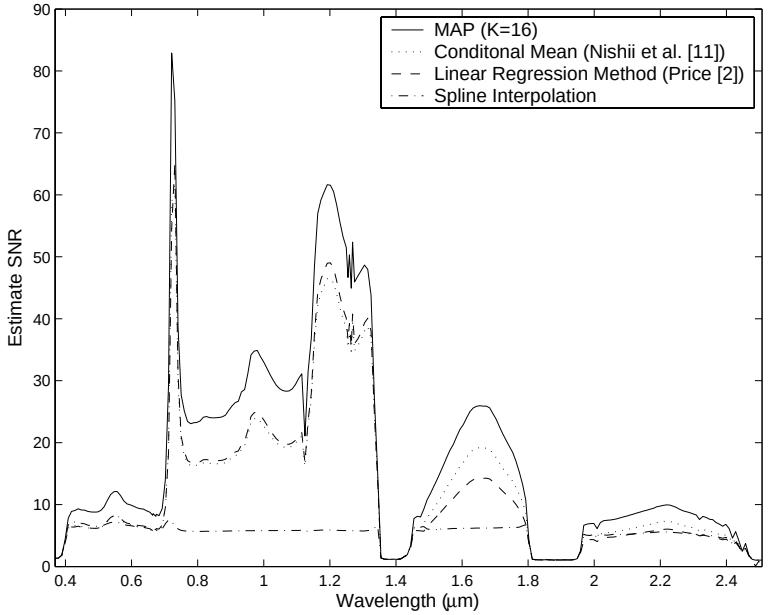


Fig. 3. SNR versus AVIRIS band wavelength for the MAP estimator ($K = 16$), the method of Nishii *et al.* (with global covariance statistics) [11], Price's method [2], and spline interpolation.

(with global covariance statistics) [11], Price's method [2], and the MAP estimator for $K = 1$ and $K = 16$. Clearly the top principal component is dramatically improved by incorporating information from the panchromatic image. However, the lower components are far more difficult to enhance due to the weak correlation with the broadband image. The MAP estimator does provide a modest increase in SNR over spline interpolation for some of the lower components, while the benchmark techniques have lower SNRs than that obtained with spline interpolation. Note also that principal component substitution methods typically seek to enhance only the principal component, and do not enhance the lower components at all. Thus, we believe that any enhancement in these lower components is a promising result. The use of multispectral high-resolution imagery (rather than panchromatic) could provide the means to better improve these lower components. To focus on these lower components, Fig. 4 shows the percentage SNR improvement over spline interpolation for the various estimators in components 2 through 20. Note that the MAP estimator with $K = 16$ generally shows the most improvement. We believe that the improvement seen with $K = 16$ versus $K = 1$ is because more correlation is present in the individual classes than exists globally.

False color images formed with the top three principal components mapped to red, green, and blue are shown in Fig. 5. In particular, the true high resolution hyperspectral image components are shown in Fig. 5(a). Spline

TABLE I
SNRS FOR ESTIMATES OF THE TOP 5 PRINCIPAL COMPONENT IMAGES

Method	PC 1	PC 2	PC 3	PC 4	PC 5
Spline Interpolation	5.87	6.36	2.86	4.86	3.90
Conditional Mean (Nishii *et al.* [11])	26.92	5.69	2.48	4.21	3.62
Linear Regression Method (Price [2])	27.92	5.81	2.48	4.24	3.36
MAP ($K = 1$)	34.74	7.42	2.99	5.04	4.44
MAP ($K = 16$)	38.97	8.37	3.05	5.24	4.48

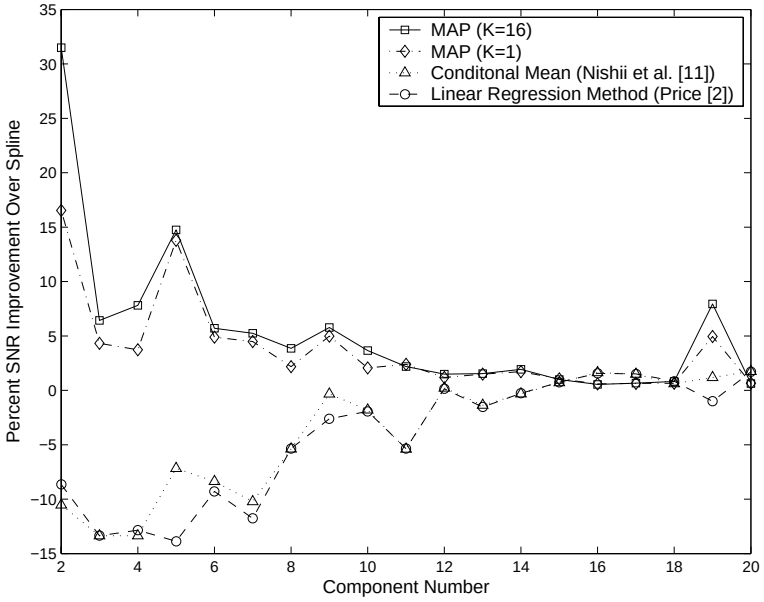


Fig. 4. Percentage improvement in SNR over straight spline interpolation for estimates of principal components 2 through 20.

interpolated components are shown in Fig. 5(b). The estimate using Price's method is shown in Fig. 5(c). Finally, the MAP estimate for $K = 16$ is shown in Fig. 5(d). An enlargement of the upper right corner of the image is shown in Fig. 6 for principal components two, three, and four. We believe that the MAP estimates generally appear sharper than the spline interpolated images and exhibit less prominent block artifacts than the estimates using Price's method (an observation consistent with the quantitative analysis).

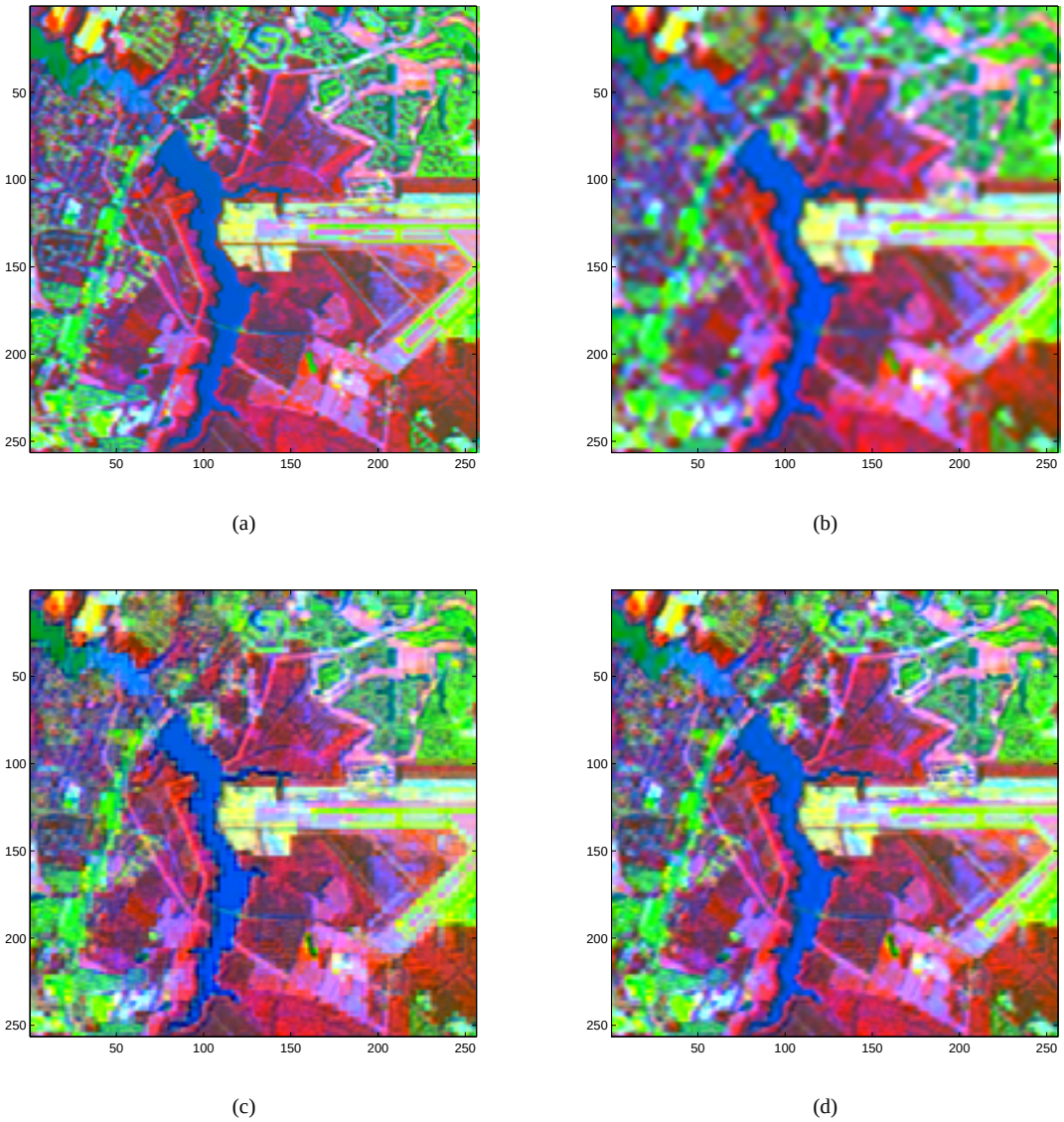


Fig. 5. False color images showing the top three principal components for (a) the true high-resolution hyperspectral image (b) spline interpolated components (c) linear regression method (Price [2]) (d) the MAP estimate with $K = 16$.

D. Noise Analysis

In this section we consider how noise impacts the performance of the MAP estimator. First we consider the impact of noise in the observed hyperspectral imagery and then we consider noise in the panchromatic imagery. The SNRs of the estimates of principal component two as a function of the average SNR of the observed low-resolution hyperspectral bands are shown in Fig. 7. Here no noise in the panchromatic image is introduced. One curve shows the SNRs for the MAP estimates with $K = 16$ when zero noise variance is assumed. Another curve

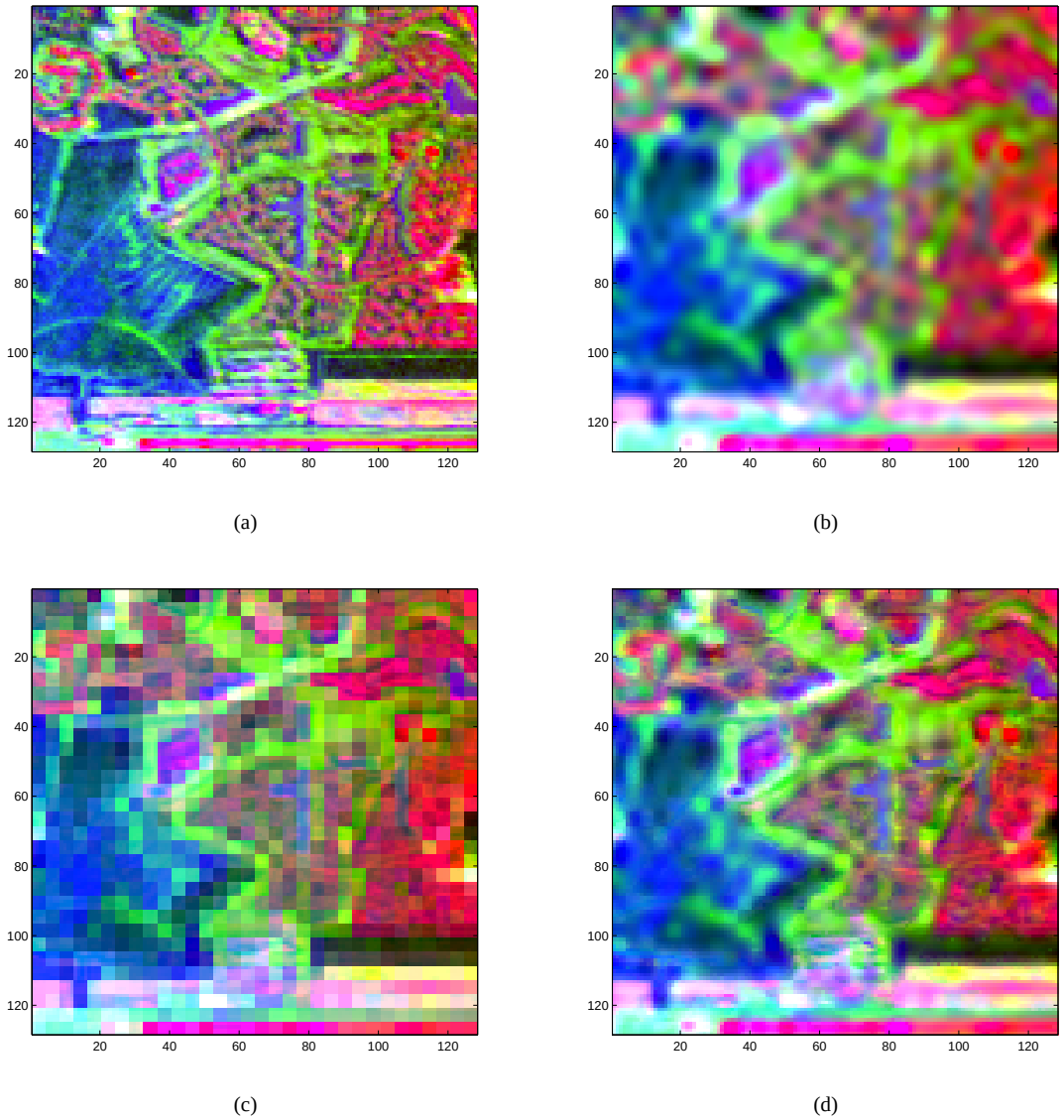


Fig. 6. False color images showing principal components two, three, and four for (a) the true high-resolution hyperspectral image (b) spline interpolated components (c) linear regression method (Price [2]) (d) the MAP estimate with $K = 16$.

shows the SNRs when the correct noise variance is known and used. Note that when significant levels of noise are present, it can be important to specify the appropriate noise variance.

The SNRs of the estimates of principal component two as a function of the panchromatic image SNR are shown in Fig. 8. Here no noise in the hyperspectral image is introduced. Note that the spline interpolator does not depend on $\mathbf{x}$, and thus, is not affected by noise in the panchromatic image. Also note that when the SNR of the panchromatic image is low, little improvement is possible as correlation with the hyperspectral bands is reduced. With very low

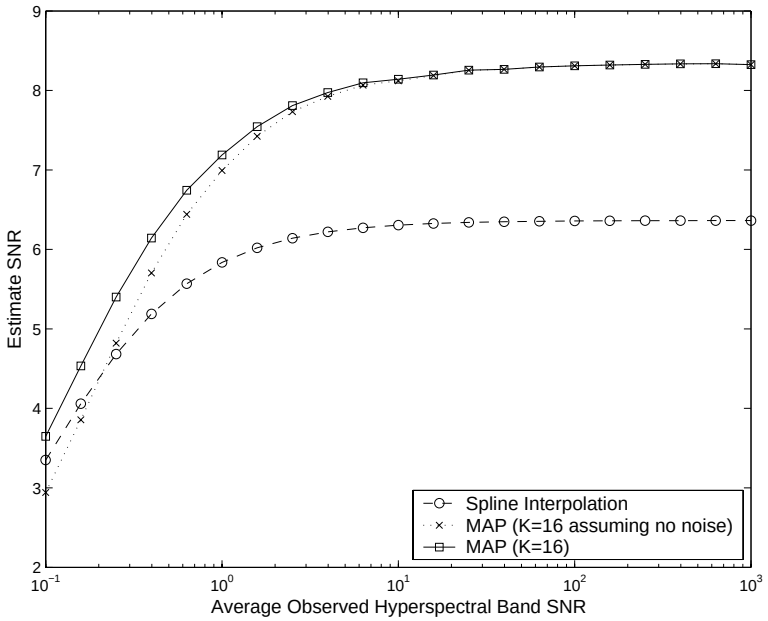


Fig. 7. The SNRs of the estimates of principal component two as a function of the average SNR of the observed low-resolution hyperspectral bands.

SNR panchromatic imagery, the correlation at the lower resolution is not indicative of the correlation at the higher resolution. This is because the noise at the lower resolution is reduced as the panchromatic image is artificially degraded to the lower resolution. The use of pre-filters for noise reduction could help to mitigate this effect.

VII. Conclusions

This paper has developed a MAP estimation framework for estimating an enhanced resolution image using co-registered high-resolution imagery from another sensor. Here we have focused on the enhancement of a hyperspectral image using high-resolution panchomatic data. However, the estimation framework developed allows for any number of spectral bands in the primary and auxiliary image. We believe that the proposed technique is suitable for applications where some correlation exists between the auxiliary image and the image being enhanced. The results with AVIRIS imagery indicate that a number of methods do well enhancing the top principal component image, where strong global correlation exists with the panchromatic image. Much more difficult to enhance are the lower component images. Notwithstanding this, we have demonstrated that the proposed estimator is capable of providing modest improvement in some of these lower components (something not seen with the benchmark techniques). The spatially varying statistical model (i.e., $K = 16$), using vector quantization, does provide some additional

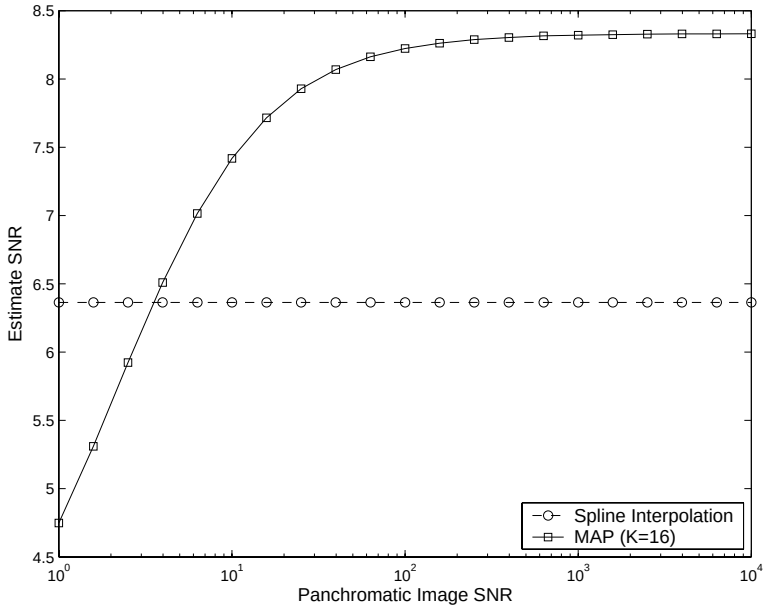


Fig. 8. The SNRs of the estimates of principal component two as a function of the panchromatic image SNR.

performance gain over global statistics (i.e., $K = 1$). We believe this is a result of exploiting the higher correlations present within the spectral classes (correlations that get "washed out" in the global statistics).

We believe that one of the merits of the proposed estimation framework is that it allows for an arbitrary linear observation model. Furthermore, the estimation framework opens up opportunities to improve upon these results with the use of more sophisticated statistical models and methods for estimating the statistical parameters for those models. For example, improved performance may be possible using a spatial-spectral model for the desired high-resolution image (i.e., not assuming the desired hyperpixels are conditionally independent, given the auxiliary image). Future work is focusing on incorporating spectral mixing models in order to improve the estimate of the local statistics. We are also exploring the use of high-resolution multispectral imagery to enhance the hyperspectral data.

REFERENCES

[1] R. A. Schowengerdt, "Reconstruction of multispatial, multispectral image data using spatial frequency content," *Photogrammetric Engineering & Remote Sensing*, vol. 46, no. 10, pp. 1325–1334, 1980.

[2] J. C. Price, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sens. Environ.*, vol. 21, pp. 119–128, 1987.

[3] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data," *Photogrammetric Engineering & Remote Sensing*, vol. 56, 1990.

[4] P. S. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogrammetric Engineering & Remote Sensing*, vol. 57, no. 3, pp. 295–303, 1991.

[5] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogrammetric Engineering & Remote Sensing*, vol. 58, pp. 561–567, 1992.

[6] C. K. Munechika, J. S. Warnick, C. Salvaggio, and J. R. Schott, "Resolution enhancement of multispectral image data to improve classification accuracy," *Photogrammetric Engineering & Remote Sensing*, vol. 59, pp. 67–72, Sept. 1993.

[7] E. Iverson and J. R. Lersch, "Adpative image sharpening using multiresolution representations," *Proceedings of SPIE*, vol. 2231, pp. 72–83, 1994.

[8] D. P. Filberti, S. E. Marsh, and R. A. Schowengardt, "Synthesis of imagery from high spatial and spectral resolution from multiple image sources," *Optical Engineering*, vol. 33, no. 8, pp. 49–56, Aug. 1994.

[9] B. Zhukov, M. Berger, F. Lanzl, and H. Kaufman, "A new technique for merging multispectral and panchromatic images revealing sub-pixel spectral variation," in *Proc IEEE IGARSS*, Firenze, Italy, July 1995, pp. 2154–2156.

[10] B. Zhukov, D. Oertel, and F. Lanzl, "A multiresolution, multisensor technique for satellite remote sensing," in *Proc IEEE IGARSS*, Firenze, Italy, July 1995, pp. 51–53.

[11] R. Nishii, S. Kusanobu, and S. Tanaka, "Enhancement of low resolution image based on high resolution bands," *IEEE Tranactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1151–1158, Sept. 1996.

[12] H. N. Gross and J. R. Schott, "Application of spatial resolution enhancement and spectral mixture analysis to hyperspectral images," *Proceedings of SPIE*, vol. 2821, pp. 30–41, 1996.

[13] ——, "Application of spectral mixture analysis and image fusion techniques for image sharpening," *Remote Sens. Environ.*, vol. 63, pp. 85–94, 1998.

[14] G. D. Robinson, H. N. Gross, and J. R. Schott, "Evaluation of two applications of spectral mixing models to image fusion," *Remote Sens. Environ.*, vol. 71, pp. 272–281, 2000.

[15] M. E. Winter and E. M. Winter, "Physics-based resolution enhancement of hyperspectral data," *Proceedings of SPIE*, vol. 4725, pp. 580–587, 2002.

[16] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation," *Photogrammetric Engineering & Remote Sensing*, vol. 66, pp. 49–56, 2000.

[17] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewoods Cliffs, New Jersey: Prentice Hall, 1993.

[18] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering*, vol. 37, no. 1, pp. 247–260, Jan. 1998.

[19] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.

[20] B. R. Hunt and T. M. Cannon, "Nonstationary assumptions for Gaussian models of images," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, pp. 876–881, Dec. 1976.

[21] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantization," *IEEE Transactions on Communication Theory*, vol. 28, no. 1, pp. 84–95, Jan. 1980.

[22] W. M. Porter, "Evolution of the airborne visible/infrared imaging spectrometer (AVIRIS) flight and ground data processing system," *Proceedings of the SPIE*, vol. 1298, pp. 11–17, 1990.

APPENDIX A

CONSTRAINED OPTIMIZATION FOR NEGLIGIBLE NOISE CASE

In the case where the noise in the observed data is negligible, it may be desirable to strictly enforce the observation model constraints. In particular, instead of the unconstrained minimization of (17), we wish to minimize

$$C(\mathbf{z}) = \frac{1}{2}(\mathbf{z} - \mu_{\mathbf{z}|\mathbf{x}})^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}(\mathbf{z} - \mu_{\mathbf{z}|\mathbf{x}}) \tag{29}$$

subject to the constraint that $\mathbf{y} = \mathbf{W}\mathbf{z}$. Multiplying out the terms and keeping only those that are functions of $\mathbf{z}$, we arrive at an equivalent cost function

$$C(\mathbf{z}) = \frac{1}{2}\mathbf{z}^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}\mathbf{z} - \left(\mu_{\mathbf{z}|\mathbf{x}}^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}\right)\mathbf{z}. \tag{30}$$

Thus, we are faced with a minimization of a quadratic form with linear constraints. We can solve the problem using the method of Lagrangian multipliers [17]. The Lagrangian is formed as

$$L(\mathbf{z}, \lambda) = \frac{1}{2}\mathbf{z}^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}\mathbf{z} - \left(\mu_{\mathbf{z}|\mathbf{x}}^T \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}\right)\mathbf{z} + \lambda^T(\mathbf{W}\mathbf{z} - \mathbf{y}), \tag{31}$$

where λ is an $MP \times 1$ vector of Lagrangian multipliers. The gradient with respect to $\mathbf{z}$ is given by

$$\nabla_{\mathbf{z}} L(\mathbf{z}, \lambda) = \mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1}\mathbf{z} - (\mathbf{C}_{\mathbf{z}|\mathbf{x}}^{-1})^T \mu_{\mathbf{z}|\mathbf{x}} + \mathbf{W}^T\lambda. \tag{32}$$

Setting this equal to zero and solving for $\mathbf{z}$ yields

$$\hat{\mathbf{z}} = \mu_{\mathbf{z}|\mathbf{x}} - \mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T\lambda. \tag{33}$$

Now let us find the λ that allows our solution to meet the linear constraint. To do so we impose the linear constraint on (33) and solve for λ. This gives us

$$\mathbf{y} = \mathbf{W}\hat{\mathbf{z}} = \mathbf{W}\left(\mu_{\mathbf{z}|\mathbf{x}} - \mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T\lambda\right). \tag{34}$$

Solving for λ we get

$$\lambda = \left(\mathbf{W}\mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T\right)^{-1}\left(\mathbf{W}\mu_{\mathbf{z}|\mathbf{x}} - \mathbf{y}\right). \tag{35}$$

Plugging this into (33) yields our final solution

$$\hat{\mathbf{z}} = \mu_{\mathbf{z}|\mathbf{x}} - \mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T\left[\left(\mathbf{W}\mathbf{C}_{\mathbf{z}|\mathbf{x}}\mathbf{W}^T\right)^{-1}\left(\mathbf{W}\mu_{\mathbf{z}|\mathbf{x}} - \mathbf{y}\right)\right]. \tag{36}$$

Note that this result matches (20) when the covariance of the noise is set to zero.

APPENDIX B

PRINCIPAL COMPONENT SPACE ESTIMATION

The PCA technique starts with the construction of a $P \times P$ "spectral" covariance matrix, typically estimated from the available low-resolution hyperpixels. Using eigenvalue decomposition, an orthonormal set of eigenvectors of the covariance is obtained. These are used to construct an orthogonal matrix $\mathbf{E}$ whose columns are the eigenvectors. The unknown high-resolution hyperpixels $\mathbf{z}_n$ are related to the PCA space hyperpixels, $\bar{\mathbf{z}}_n$, through the transformation $\mathbf{z}_n = \mathbf{E}\bar{\mathbf{z}}_n$, for $n = 1, 2, \ldots, N$. Let us define $\mathbf{E}_N = \bigoplus_{n=1}^{N}\mathbf{E}$ and $\mathbf{E}_M = \bigoplus_{m=1}^{M}\mathbf{E}$. Note that since $\mathbf{E}^{-1} = \mathbf{E}^T$, the same is true of $\mathbf{E}_N$ and $\mathbf{E}_M$. Therefore, we can relate the PCA space and spectral space variables as $\bar{\mathbf{z}} = \mathbf{E}_N^T\mathbf{z}$, $\bar{\mathbf{y}} = \mathbf{E}_M^T\mathbf{y}$, and $\bar{\mathbf{n}} = \mathbf{E}_M^T\mathbf{n}$. The estimation problem in the PCA space now becomes one of estimating $\bar{\mathbf{z}}$ from $\bar{\mathbf{y}}$ and $\mathbf{x}$.

The observation model defined in (4) in the spectral space can be transformed into the PCA space as $\bar{\mathbf{y}} = \bar{\mathbf{W}}\bar{\mathbf{z}} + \bar{\mathbf{n}}$ where $\bar{\mathbf{W}} = \mathbf{E}_M^T\mathbf{W}\mathbf{E}_N$ and the transformed noise is a zero-mean Gaussian random vector with covariance $\mathbf{C}_{\bar{\mathbf{n}}} = \mathbf{E}_M^T\mathbf{C}_{\mathbf{n}}\mathbf{E}_M$. Also, if $\mathbf{z}$ and $\mathbf{x}$ are jointly Gaussian, then so are $\bar{\mathbf{z}}$ and $\mathbf{x}$. Thus, the MAP estimate in PCA space has the same form as that in (20), but with the variables replaced with their PCA counterparts, yielding

$$\hat{\bar{\mathbf{z}}} = \mu_{\bar{\mathbf{z}}|\mathbf{x}} + \mathbf{C}_{\bar{\mathbf{z}}|\mathbf{x}}\bar{\mathbf{W}}^T\left[\bar{\mathbf{W}}\mathbf{C}_{\bar{\mathbf{z}}|\mathbf{x}}\bar{\mathbf{W}}^T + \mathbf{C}_{\bar{\mathbf{n}}}\right]^{-1}\left[\bar{\mathbf{y}} - \bar{\mathbf{W}}\mu_{\bar{\mathbf{z}}|\mathbf{x}}\right]. \tag{37}$$

Here $\mu_{\bar{\mathbf{z}}|\mathbf{x}} = \mathrm{E}\{\bar{\mathbf{z}}|\mathbf{x}\}$ and $\mathbf{C}_{\bar{\mathbf{z}}|\mathbf{x}}$ is the conditional covariance of $\bar{\mathbf{z}}$ given $\mathbf{x}$. These are estimated in a fashion similar to their spectral domain counterparts. The estimate can be mapped back to the spectral space using $\hat{\mathbf{z}} = \mathbf{E}_N \hat{\bar{\mathbf{z}}}$, if desired.

Note that in the special case where $\mathbf{C}_{\mathbf{n}} = \sigma_n^2 \mathbf{I}$, then $\mathbf{C}_{\bar{\mathbf{n}}} = \mathbf{E}_M^T \sigma_n^2 \mathbf{I} \mathbf{E}_M = \sigma_n^2 \mathbf{I}$. Also, if $\mathbf{W}$ is formed using a system PSF which is identical in each spectral band, then it can be shown that $\bar{\mathbf{W}} = \mathbf{W}$. To see this, consider partitioning $\mathbf{W}$ into $P \times P$ blocks $\mathbf{W}_{i,j}$, where $1 \leq i \leq M$ and $1 \leq j \leq N$, such that each block operates on one hyperpixel. The corresponding (i,j) block of $\bar{\mathbf{W}}$ is given by $\mathbf{E}^T \mathbf{W}_{i,j} \mathbf{E}$. The assumption that $\mathbf{W}$ does not vary spectrally is equivalent to $\mathbf{W}_{i,j}$ being a multiple of a $P \times P$ identity matrix, $\mathbf{I}_P$, for all (i,j). Therefore, $\mathbf{W}_{i,j}$ commutes with $\mathbf{E}$. Since $\mathbf{E}$ is orthogonal, $\mathbf{E}^T \mathbf{E} = \mathbf{I}_P$. Hence, the (i,j) block of $\bar{\mathbf{W}}$ is $\mathbf{W}_{i,j}$, implying that $\bar{\mathbf{W}} = \mathbf{W}$.

The advantage of PCA space processing lies in the fact that most of the signal energy lies in the PCA subspace corresponding to the top eigenvalues. The lower principal component dimensions can be treated with a simpler algorithm, such as interpolation, or even set to zero if they are deemed to be primarily noise. Processing in a lower dimension PCA subspace reduces the size of the matrix inverse required in the closed form solution and requires the estimation of fewer statistical parameters. In addition to the benefit in terms of processing time, this tends to lead to covariance estimates that are better conditioned, making the resulting MAP algorithm more robust.

ACKNOWLEDGMENTS

This work was supported under Air Force Research Laboratory contract F33615-02-C-1169. The authors would like to thank Capt. Byron Formwalt at Air Force Research Laboratory for his valuable input and for supporting the project. We would also like to gratefully acknowledge the technical and project management support of Thomas Fitzgerald at Mission Research Corporation, in Dayton, OH.

Russell C. Hardie graduated *Magna Cum Laude* from Loyola College in Baltimore Maryland in 1988 with a B.S. degree in Engineering Science. He obtained an M.S. and Ph.D. degree in Electrical Engineering from the University of Delaware in 1990 and 1992, respectively. Dr. Hardie served as a Senior Scientist at Earth Satellite Corporation in Maryland prior to his appointment at the University of Dayton in 1993. He is currently an Associate Professor in the Department of Electrical and Computer Engineering with a joint appointment in the Electro-Optics Program. He also serves as a consultant to iCAD, Inc. in the area of medical image processing.

In 1998, Dr. Hardie was the co-recipient of the Rudolf Kingslake Medal and Prize from SPIE for work on multi-frame image resolution enhancement algorithms. Dr. Hardie received the School of Engineering Award of Excellence in Teaching at the University of Dayton in 1999 and received the first University of Dayton IEEE Professor of the Year Award in 2002 (presented by the student chapter of the IEEE). His current research interests are in the areas of signal and image restoration, medical image processing, nonlinear filters, and remote sensing.

Michael T. Eismann is the Electro-Optical Targeting Branch technical advisor at the Sensors Directorate of the Air Force Research Laboratory (AFRL), Wright-Patterson Air Force Base, Ohio. He is responsible for overseeing the development of passive electro-optical and infrared sensor technology, and transition into operational airborne targeting and reconnaissance systems. His recent research emphasis has been the development of hyperspectral imaging sensors and image processing. Prior to joining AFRL in 1996, he was employed by the Environmental Research Institute of Michigan (ERIM). Mr. Eismann joined ERIM in 1991, where he was involved in research concerning active and passive optical and infrared targeting and reconnaissance, optical information processing, and holographic optics. He has a B.S. in Physics from Thomas More College, an M.S. in Electrical Engineering from the Georgia Institute of Technology, and is pursuing a Ph.D. in Electro-Optics at the University of Dayton.

Gregory L. Wilson graduated *Summa Cum Laude* from California State University Hayward earning a B.S. degree in mathematics in 1977. In 1979 he graduated with an M.S. degree in mathematics from the same institution with highest honors while maintaining a 4.0 GPA. In 1983, he received his Ph.D. in mathematics from the University of California at Santa Barbara, working in the field of Modular Function Theory.

Since joining Mission Research in 1983, Dr. Wilson has been exposed to a wide variety of physics and engineering problems. As a result, he has developed several areas of interest resulting in well over 30 reports or publications in diverse areas, including one voted best technical paper for the Antenna Measurements Techniques Association (AMTA) Symposium of 2001/2002. Interests include mathematical modeling and parallel programming in the context of hyperspectral image enhancement. He is currently working with a multi-computer development workstation, developing C programs to run in a real time parallel environment.

LIST OF TABLES

LIST OF FIGURES