Linear Mixture Models, Full and Partial Unmixing in Multi- and Hyperspectral Image Data*

Allan Aasbjerg Nielsen
IMM, Department of Mathematical Modelling
Technical University of Denmark, Building 321
DK-2800 Lyngby, Denmark
E-mail: aa@imm.dtu.dk, Internet http://www.imm.dtu.dk/∼aa

Abstract

As a supplement or an alternative to classification of hyperspectral image data the linear mixture model is considered in order to obtain estimates of abundance of each class or end-member in pixels with mixed membership. Full unmixing and the partial unmixing methods orthogonal subspace projection (OSP), constrained energy minimization (CEM) and an eigenvalue formulation alternative are dealt with. The solution to the eigenvalue formulation alternative proves to be identical to the CEM solution. Also, spectral angle mapping (SAM) is described. The matrix inversion involved in CEM can be avoided by working on (a subset of) orthogonally transformed data such as signal maximum autocorrelation factors, MAFs, or signal minimum noise fractions, MNFs. This will also cause the noise isolated in the MAF/MNFs not included in the analysis not to influence the partial unmixing result. CEM and the eigenvalue formulation alternative enable us to perform partial unmixing when we know the desired end-member spectra only and not the full set of end-member spectra. This is an advantage over full unmixing and OSP. An example with a simple generated 2-band image shows the ability of the CEM method to isolate the desired signal. A case study with a 30-band subset of AVIRIS data from the Mojave Desert, California, USA, shows the utility of the methods to more realistic data.

1 Introduction

In ordinary discriminant analysis which is often used to classify for instance multi- or hyperspectral remote sensing image data it is assumed that each observation (or pixel) is a member of one and only one of a number of pre-determined classes. Linear mixture models allow us to estimate the abundance of each class in pixels with mixed class membership, [20, 13, 30, 23].

This article gives an overview of several methods for full and partial unmixing described in the literature. Also, some new ideas are presented, namely 1) the inclusion of a constant in the linear mixture model (α_0 below), 2) an eigenvalue formulation alternative to constrained energy minimization, and 3) partial unmixing in MAF/MNF space to avoid matrix inversion and to exclude noise isolated in MAF/MNFs not included in the analysis.

*Submitted to SCIA'99, Kangerlussuaq, Greenland, 7-11 June 1999, http://www.diku.dk/scia99.

2 Linear Mixing

We assume that the signal measured at each pixel consists of a linear combination of p so-called end-members. End-members are pure pre-determined classes with 100% abundance of one element and with no mixtures. We think of our l-dimensional signal for end-member i as a vector $\boldsymbol{m}_i = [m_{i1} \ \dots \ m_{il}]^T, \ i = 1, \dots, p$ and represent the end-members by a matrix

$$\boldsymbol{M} = [\boldsymbol{m}_1 \dots \boldsymbol{m}_p] = \begin{bmatrix} m_{11} & \cdots & m_{p1} \\ \vdots & \ddots & \vdots \\ m_{1l} & \cdots & m_{pl} \end{bmatrix} \quad (1)$$

with one column for each end-member. We write each observation $\boldsymbol{r}(x,y) = [r_1(x,y) \ \dots \ r_l(x,y)]^T$ as a linear combination of the end-members $\boldsymbol{M}$; the abundances $\boldsymbol{\alpha}(x,y) = [\alpha_1(x,y) \ \dots \ \alpha_p(x,y)]^T$ are the coefficients

$$\boldsymbol{r}(x,y) = \boldsymbol{M}\boldsymbol{\alpha}(x,y) + \boldsymbol{n}(x,y) \quad (2)$$

where $\boldsymbol{n}(x,y) = [n_1(x,y) \ \dots \ n_l(x,y)]^T$ is the residual or the noise, i.e. the variation in $\boldsymbol{r}(x,y)$ not explained by the model. This is the linear mixture model. The term linear means linear in the coefficients. The expected value of the noise $\mathrm{E}\{\boldsymbol{n}\} = \boldsymbol{0}$. In linear models a constant term α_0 is often introduced (from now on we omit (x,y) from the notation). Here, α_0 represents effects not explained by the chosen end-members. If we introduce α_0 we get

$$\boldsymbol{M} = \begin{bmatrix} 1 & m_{11} & \cdots & m_{p1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & m_{1l} & \cdots & m_{pl} \end{bmatrix} \quad (3)$$

$$\boldsymbol{\alpha} = [\alpha_0 \ \alpha_1 \ \dots \ \alpha_p]^T. \quad (4)$$

Sometimes the column of ones is replaced by a column of zeros. This represents the end-member "total shade."

To solve the system of equations involved we minimize the sum of squared residuals $\boldsymbol{n}^T\boldsymbol{n}$ or more generally $\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n}$ where $\boldsymbol{\Sigma}_n$ is the dispersion or covariance matrix of the residuals. This is done by setting the partial derivative $\partial(\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/\partial\boldsymbol{\alpha} = \boldsymbol{0}$. The result is

$$\boldsymbol{\alpha} = (\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{r}. \quad (5)$$

The estimator $\boldsymbol{\alpha}$ is central with dispersion $(\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{M})^{-1}$. When $\boldsymbol{\Sigma}_n = \sigma^2\boldsymbol{I}$ where $\boldsymbol{I}$ is the $l \times l$ unit matrix and σ^2 is the variance of all residuals, $\mathrm{V}\{n_i\} = \sigma^2$

$$\boldsymbol{\alpha} = (\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T\boldsymbol{r} \tag{6}$$

with dispersion $\sigma^2(\boldsymbol{M}^T\boldsymbol{M})^{-1}$.

To evaluate the goodness of the model we use $R^2 = (\tilde{\boldsymbol{r}}^T\tilde{\boldsymbol{r}} - \boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/\tilde{\boldsymbol{r}}^T\tilde{\boldsymbol{r}}$, the coefficient of determination (if $\boldsymbol{M}$ contains the column of ones $\tilde{\boldsymbol{r}}$ is $\boldsymbol{r}$ centered, if not $\tilde{\boldsymbol{r}} = \boldsymbol{r}$), and the estimate of residual variance, $s^2 = (\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/(l-p-1)$. s is called the root mean square error, RMSE. $l-p-1$, the degree of freedom, must be positive. If an extra column is added to $\boldsymbol{M}$ p is replaced by $p+1$.

3 Full Unmixing

To perform a full unmixing one needs to know the spectra for all end-members present in the scene. If we interpret $\boldsymbol{\alpha}$ as abundances of all end-members for each observation we demand that the α_is add to 100%, $\boldsymbol{1}^T\boldsymbol{\alpha} = 1$, where $\boldsymbol{1}$ is a column vector of ones, and that $\alpha_i \geq 0$. The first constraint can be dealt with by introducing a Lagrange multiplier -2λ and minimizing $F = \boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n} + 2\lambda(\boldsymbol{1}^T\boldsymbol{\alpha} - 1)$ without constraints. The solution obtained by setting the partial derivatives $\partial F/\partial\boldsymbol{\alpha} = \boldsymbol{0}$ and $\partial F/\partial\lambda = 0$ is

$$\begin{bmatrix} \boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{M} & \boldsymbol{1} \\ \boldsymbol{1}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{r} \\ 1 \end{bmatrix}. \tag{7}$$

The latter constraint can be dealt with by means of methods from convex quadratic programming.

Often, knowledge of all end-member spectra is not available. Therefor partial unmixing methods where we estimate the presence of one or a few desired spectra only are important. α_0 with a column of ones in $\boldsymbol{M}$ above will to some extent reflect the presence of end-members not accounted for in $\boldsymbol{M}$ and so will R^2 and RMSE.

4 Partial Unmixing

Partial unmixing builds on the usual linear mixture model in equation 2. We split the $\boldsymbol{M}\boldsymbol{\alpha}$ term into two terms, one which is the desired end-member $\boldsymbol{d}$ with a corresponding abundance α_p (without loss of generality we place $\boldsymbol{d}$ in the last column of $\boldsymbol{M}$), and one which consists of the undesired end-members $\boldsymbol{U}$ with a corresponding $(p-1)\times 1$ vector, $\boldsymbol{\gamma}$, of abundances. $\boldsymbol{U}$ contains the first $p-1$ columns of $\boldsymbol{M}$ and $\boldsymbol{\gamma}$ contains the first $p-1$ elements of $\boldsymbol{\alpha}$. Hence

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{n} \\ &= \boldsymbol{d}\alpha_p + \boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{n}. \end{aligned} \tag{8}$$

$\boldsymbol{U}\boldsymbol{\gamma}$ is often termed the interference. In partial unmixing we want to develop methods to eliminate or minimize the effect of $\boldsymbol{U}$ and $\boldsymbol{\gamma}$. Often the term matched filtering is applied to such methods.

4.1 Orthogonal Subspace Projection, OSP

The idea in OSP, [21, 10], is to project $\boldsymbol{r}$ onto a subspace where $\boldsymbol{\gamma}$ is removed from the linear mixture model, equation 8. Applying the $l \times l$ matrix $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{U}(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T$ we obtain

$$\begin{aligned} \boldsymbol{P}\boldsymbol{r} &= \boldsymbol{P}\boldsymbol{d}\alpha_p + \boldsymbol{U}\boldsymbol{\gamma} - \boldsymbol{U}(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{P}\boldsymbol{n} \\ &= \boldsymbol{P}\boldsymbol{d}\alpha_p + \boldsymbol{P}\boldsymbol{n}. \end{aligned} \tag{9}$$

We have indeed removed $\boldsymbol{\gamma}$ from the linear mixture model but as with full unmixing we need $\boldsymbol{U}$, i.e. we need all the end-member spectra, both desired and undesired.

In [29] it is shown that full linear unmixing and OSP as described in [10] are identical (except that OSP is computationally slightly more expensive).

4.2 Constrained Energy Minimization, CEM

Constrained energy minimization, CEM, [28, 31, 11], builds on the linear mixture model in equation 8. In CEM we project $\boldsymbol{r}$ onto $\boldsymbol{w}$ with the intent to highlight presence of the desired end-member, and to suppress the presence of the undesired end-members and noise. We do this by requesting the following

1. we want the output (the projected value) to be one when we project the desired spectrum, $\boldsymbol{d}$, i.e. we want $\boldsymbol{w}^T\boldsymbol{d} = 1$;
2. in general, we want the output to be close to zero, we want its expected value to be 0, $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$;
3. also we want to minimize the expected value of the squared difference between the output, $\boldsymbol{w}^T\boldsymbol{r}$, and the desired output, 0, i.e. we want to minimize $\mathrm{E}\{(\boldsymbol{w}^T\boldsymbol{r} - 0)^2\}$.

Since $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$ we get $\mathrm{E}\{(\boldsymbol{w}^T\boldsymbol{r} - 0)^2\} = \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\}$. Hence the job is to minimize $\mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\} = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$ with the constraint $\boldsymbol{w}^T\boldsymbol{d} = 1$. $\boldsymbol{\Sigma}$ is the dispersion matrix of $\boldsymbol{r}$. To do this we introduce a Lagrange multiplier -2λ and minimize $F = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} + 2\lambda(\boldsymbol{w}^T\boldsymbol{d} - 1)$ without constraints. This is done by setting the partial derivatives $\partial F/\partial\boldsymbol{w} = \boldsymbol{0}$ and $\partial F/\partial\lambda = 0$. This leads to

$$\begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{d} \\ \boldsymbol{d}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix} \tag{10}$$

or

$$\boldsymbol{w} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{d}}{\boldsymbol{d}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{d}} \tag{11}$$

with $\lambda = -1/\boldsymbol{d}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{d}$ (which is constant). $(\boldsymbol{w}^T\boldsymbol{r})^2$ the expectation of which is minimized is often termed "energy," hence the name CEM.

For hyperspectral data these operations can be performed on a subset of orthogonally transformed data such as signal maximum autocorrelation factors, MAFs, or signal minimum noise fractions, MNFs, [33, 3, 9, 5, 19, 2, 27, 26, 22, 32, 18, 24]. In this case matrix inversion is not needed since $\boldsymbol{\Sigma}_M = \boldsymbol{I}$. Hence $\boldsymbol{w}_M = \boldsymbol{d}_M/\boldsymbol{d}_M^T\boldsymbol{d}_M$ where the subscript M denotes dispersion, projection vector and desired spectrum after the MAF or MNF transformation.

Note, that nothing in the above ensures the ideal: $0 \leq \boldsymbol{w}^T\boldsymbol{r} \leq 1$. On the contrary, we have requested that $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$ which means that some projections must necessarily be negative.

As opposed to OSP and full linear unmixing, CEM does not require knowledge of all end-member spectra. Only the desired spectrum is needed.

4.3 An Eigenvalue Formulation Alternative to CEM

As a new alternative approach to CEM consider equation 8 and the projection $\boldsymbol{w}^T\boldsymbol{r}$ again

$$\boldsymbol{w}^T\boldsymbol{r} \quad = \quad \boldsymbol{w}^T\boldsymbol{d}\alpha_p + \boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{w}^T\boldsymbol{n}. \tag{12}$$

Consider the variance of $\boldsymbol{w}^T\boldsymbol{r}$

$$\begin{aligned} \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\} \quad &= \quad \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{d}\alpha_p\} + \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma}\} + \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{n}\} + \\ & \quad\quad 2\mathrm{Cov}\{\boldsymbol{w}^T\boldsymbol{d}\alpha_p, \boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma}\} \end{aligned} \tag{13}$$

where we assume no covariation between the abundance of the desired spectrum and noise, and the abundances of the undesired spectra and noise. $\mathrm{Cov}\{\cdots\}$ denotes covariance. This can be written as

$$\begin{aligned} \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} \quad &= \quad \mathrm{V}\{\alpha_p\}\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} + \boldsymbol{w}^T\boldsymbol{U}\,\mathrm{D}\{\boldsymbol{\gamma}\}\boldsymbol{U}^T\boldsymbol{w} + \\ & \quad\quad \boldsymbol{w}^T\boldsymbol{\Sigma}_n\boldsymbol{w} + 2\boldsymbol{w}^T\boldsymbol{d}\,\mathrm{Cov}\{\alpha_p, \boldsymbol{\gamma}\}\boldsymbol{U}^T\boldsymbol{w} \\ &= \quad \mathrm{V}\{\alpha_p\}\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} + \boldsymbol{w}^T\boldsymbol{E}\boldsymbol{w} \end{aligned} \tag{14}$$

where $\boldsymbol{E}$ represents all undesired effects, namely dispersions of interference and noise, and covariance between abundances of desired and undesired spectra. $\boldsymbol{E}$ is unknown. $\mathrm{D}\{\cdots\}$ denotes dispersion. From this we get

$$1 \quad = \quad \mathrm{V}\{\alpha_p\}\frac{\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}} + \frac{\boldsymbol{w}^T\boldsymbol{E}\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}}. \tag{15}$$

To minimize the variance of all the undesired effects we must minimize the last term on the right-hand-side of equation 15 and therefore since the sum is constant we must maximize the Rayleigh coefficient in the first term. Since $\boldsymbol{d}\boldsymbol{d}^T$ is rank 1 we get one solution only, namely the $\boldsymbol{w}$ that satisfies the generalized eigenvalue problem

$$\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} \quad = \quad \rho\boldsymbol{\Sigma}\boldsymbol{w}. \tag{16}$$

If we insert the solution for $\boldsymbol{d}$ found in equation 11 we see that these two solutions are identical with $\rho = -1/\lambda$.

5 Spectral Angle Mapping, SAM

As a supposedly more physically oriented way of establishing a measure of closeness to a desired spectrum, spectral angle mapping, SAM, has been suggested, [14]. In SAM the angle between the desired spectrum, $\boldsymbol{d}$, and the spectrum in each pixel, $\boldsymbol{r}$, is measured. The angle is the inverse cosine of the normalized inner product $\boldsymbol{d}^T\boldsymbol{r}/(||\boldsymbol{d}||\;||\boldsymbol{r}||)$. Apart from a scaling constant this inner product corresponds to CEM with $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$. The result from SAM is ideally insensitive to illumination effects.

6 Computer Programs

Five computer programs developed at IMM are useful in this type of analysis, `unmix`, `maf`, `project`, `seed` and `spam`.

`unmix` performs full unmixing either without constraints or with the natural constraints that the non-negative abundances sum to one, and partial unmixing with the natural constraints that the non-negative abundances sum to a quantity not greater than one. The unconstrained problem is solved by LINPACK routines, [4], the constrained problems by a linearly constrained least squares algorithm, LSSOL, which solves the problem: minimize $\frac{1}{2}||\boldsymbol{r} - \boldsymbol{M}\boldsymbol{\alpha}||^2$ over $\boldsymbol{\alpha}$ in this case with $\alpha_i \geq 0$ and $\boldsymbol{1}^T\boldsymbol{\alpha} = 1$ respectively $\boldsymbol{1}^T\boldsymbol{\alpha} \leq 1$, [8].

`maf` finds principal components, (rotated) principal factors, maximum autocorrelation factors, minimum noise fractions, canonical discriminant functions, (multiset) canonical variates and linear combinations that give maximal multivariate differences of two sets of variables, [24]. The eigenvalue problems associated with the analysis are solved by means of LAPACK routines, [6]. For a fuller description, see [22].

`project` projects data in feature space onto a unit vector representing a desired end-member spectrum.

`seed` grows a training area from one or a few pixels by requesting spatial as well as spectral closeness. Spatial closeness is ensured by requesting 8-neighbor connectivity. Spectral closeness is ensured by requesting low Euclidean or Mahalanobis distance in feature space. For a fuller description, see [7, 25].

`spam` performs spectral angle mapping.

All these programs are written to comply with the HIPS standard, [17, 16, 15].

7 Case Studies

7.1 Simple, Generated Data

The data used consist of two bands, one with a centered horizontal bar and one with a centered vertical bar. The bars and the backgrounds have graylevel values 1 and 0 respectively. Both bands have Gaussian noise with standard deviation 0.5 added. Figure 1 shows the two bands without noise in the first column, the two bands with noise in the second column, the CEM results for end-members $[1\ 0]^T$ (horizontal bar) and $[0\ 1]^T$ (vertical bar) stretched linearly from minimum to maximum in the third column, and the CEM results stretched linearly from 0 to 1 in the fourth column. Figure 2 shows results from a full unmixing without constraints. Column one is abundances of end-member $[1\ 0]^T$, column two is abundances of end-member $[0\ 1]^T$, column three is R^2, and column four is RMSE. In the top row all quantities are stretched linearly between minimum and maximum, in the bottom row all quantities are stretched linearly between 0 and 1. Figure 3 shows the same results for the full, constrained unmixing. Both full unmixings are carried out with three variables, namely bands 1 and 2 and their product.

7.2 AVIRIS data over the Mojave Desert, California, USA

The data used here is the 30 bands subset of AVIRIS data, [34, 1], over a small part of the Mojave Desert, California, USA, that come with the LinkWinds software, [12]. These bands cover the spectral range 0.52–2.33 μm. The images have 180 rows and 360 columns. Figure 4 shows every other of the 30 bands (row-wise).

To establish which regions in the image contain extreme values and therefore are potential end-members we look for the minimum and maximum values in the MAFs. The first 14 MAFs are shown in Figure 5 (row-wise). Since we don't

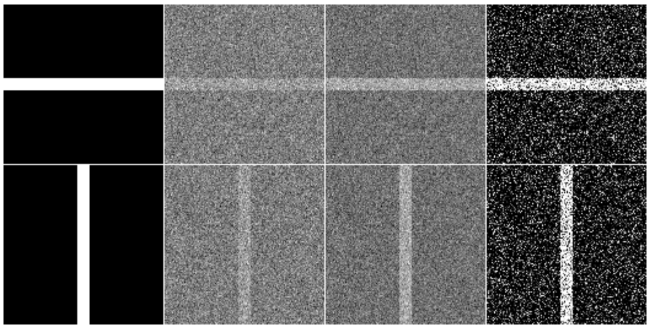

Figure 1: Two bands simulated (without noise in column one, with noise in column two) and the CEM result (columns three and four)

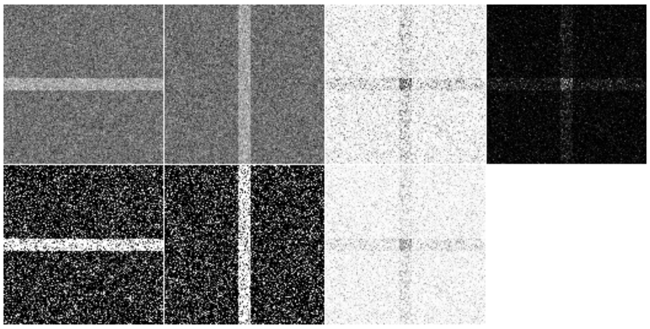

Figure 2: Two bands simulated, full unmixing without constraints

want our partial unmixing results to be based on noise spectra we use training areas grown from the pixels with extreme values as seeds to calculate average spectra instead of using the spectra from the extreme pixels themselves directly.

This is a "true remote sensing situation," we don't know what is on the ground. Our aim here is to illustrate the CEM and SAM methods and not to classify or identify material on the ground. We arbitrarily choose six potential end-members corresponding to the extreme values of MAFs 1-3.

Figure 6 shows the six training areas grown from these extremes by `seed`. Figures 7 and 8 show the resulting abundance images as estimated from the first 9 MAFs by CEM. In Figure 7 the abundance images are stretched linearly from minimum to maximum, in Figure 8 they are stretched linearly from 0 to 1. Figures 11 and 12 show the corresponding histograms.

Figures 9 and 10 show the results from the SAM analyses. Figure 9 shows the normalized inner products, Figure 10 the spectral angles. Obviously, the spectral angle is small where the abundance is high. Figures 13 and 14 show the corresponding histograms.

With stretching from 0 to 1 the CEM abundance estimates

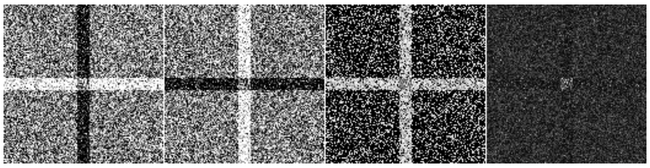

Figure 3: Two bands simulated, full unmixing with constraints

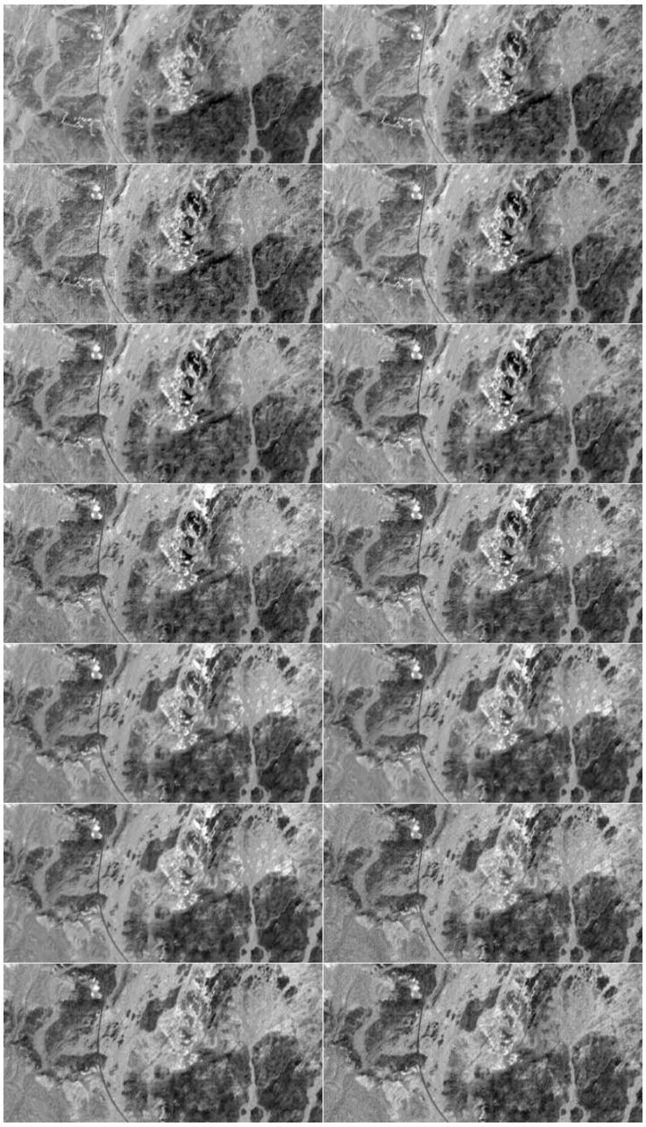

Figure 4: Every other of the 30 bands subset of AVIRIS data over the Mojave Dessert, California, USA

give a more distinct and visually pleasing impression of the spatial distribution of spectrally similar material than does spectral angles. SAM seems to give information on what is very different from the desired spectrum whereas CEM seems to give information on what is very similar to the desired spectrum.

8 Conclusions

CEM and an eigenvalue formulation alternative enable us to perform partial unmixing when we know the desired end-member spectra only and not the full set of end-member spectra. This is an advantage over full unmixing and OSP. When applying the CEM method or the eigenvalue formulation alternative to MAFs or MNFs, matrix inversion is not needed and also the noise isolated in the MAF/MNFs not included in the analysis does not influence the matched filtering performed. Apart from a scaling constant the inner product in the simpler SAM method corresponds to CEM without allowing for covariance between the variables.

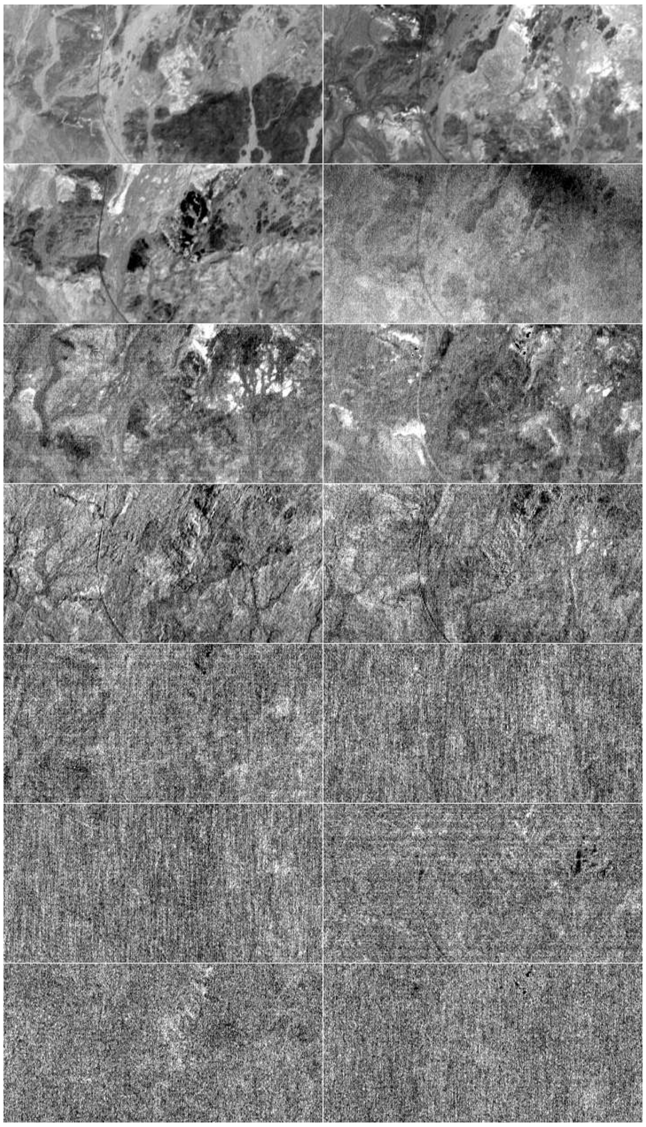

Figure 5: The first 14 MAFs of the AVIRIS data

Acknowledgment

Professor Knut Conradsen and Dr. Bjarne K. Ersbøll, both IMM, and Dr. Rasmus Larsen, formerly IMM, now Novo Nordisk A/S, provided strong backing over the years. Rasmus wrote the `maf` and `seed` programs in close cooperation with the author. Johan Dóre, IMM, wrote part of the `seed` program.

References

[1] AVIRIS. Internet http://makalu.jpl.nasa.gov/aviris.html.

[2] Knut Conradsen, Bjarne Kjær Ersbøll, and Allan Aasbjerg Nielsen. Noise removal in multichannel image data by a parametric maximum noise fractions estimator. In ERIM, editor, *Proceedings of the 24th International Symposium on Remote Sensing of Environment*, pages 403–416, Rio de Janeiro, Brazil, May 1991.

[3] Knut Conradsen, Bjarne Kjær Nielsen, and Tage Thyrsted. A comparison of min/max autocorrelation factor analysis and ordinary factor analysis. In *Proceedings from Symposium in Applied Statistics*, pages 47–56, Lyngby, Denmark, January 1985.

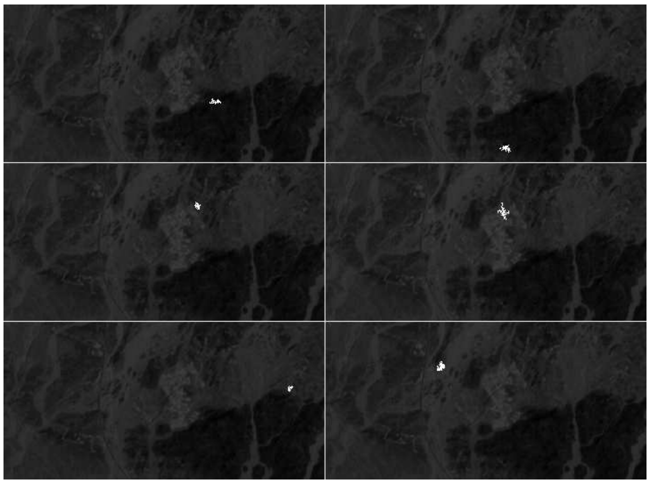

Figure 6: Seed grown training areas under which average spectra are calculated shown in white on top of MAF1

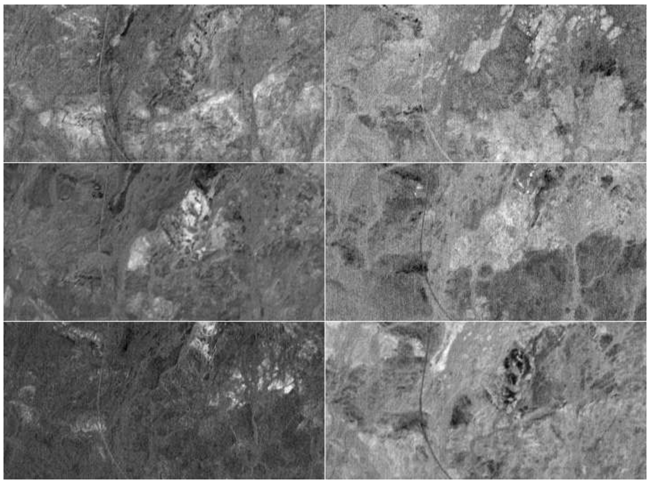

Figure 7: CEM abundance estimates stretched linearly from minimum to maximum

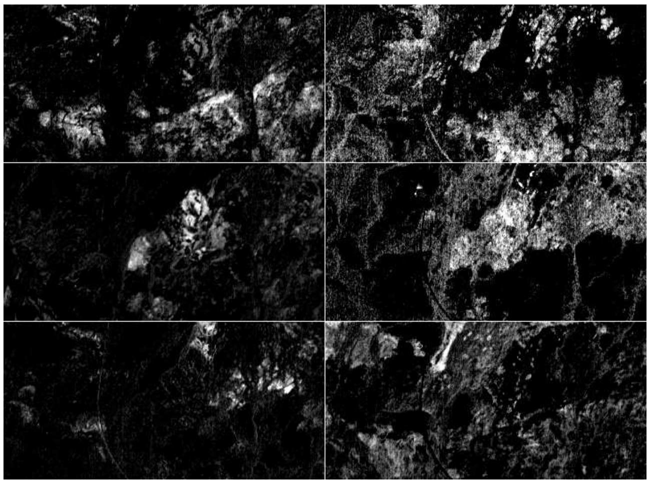

Figure 8: CEM abundance estimates stretched linearly from 0 to 1

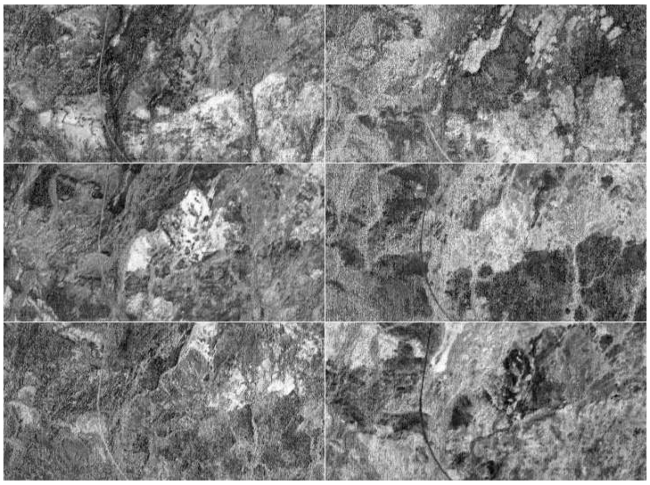

Figure 9: Cosines of spectral angles

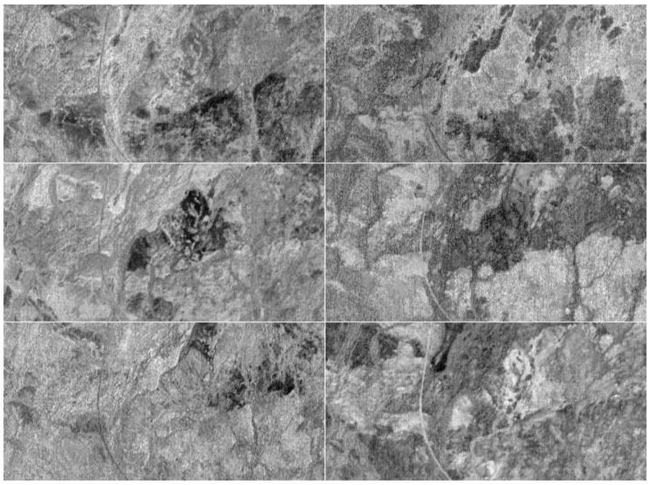

Figure 10: Spectral angles

[4] J. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart. *LINPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, 1979.

[5] Bjarne K. Ersbøll. *Transformations and Classifications of Remotely Sensed Data: Theory and Geological Cases*. PhD thesis, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, 1989. 297 pp.

[6] E. Anderson et al. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, second edition, 1995.

[7] Harald Flesche, Allan Aasbjerg Nielsen, and Rasmus Larsen. Supervised mineral classification with semi-automatic training and validation set generation in scanning electron microscope energy dispersive spectroscopy images of thin sections. *Submitted*, 1998.

[8] P. E. Gill, S. J. Hammarling, W. Murray, M. A. Saunders, and M. H. Wright. User's guide for LSSOL (version 1.0): A FORTRAN package for constrained linear least-squares and convex quadratic programming. Technical Report 86-1, Department of Operations Research, Stanford University, California, USA, 1986.

[9] A. A. Green, M. Berman, P. Switzer, and M. D. Craig. Transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, 1988.

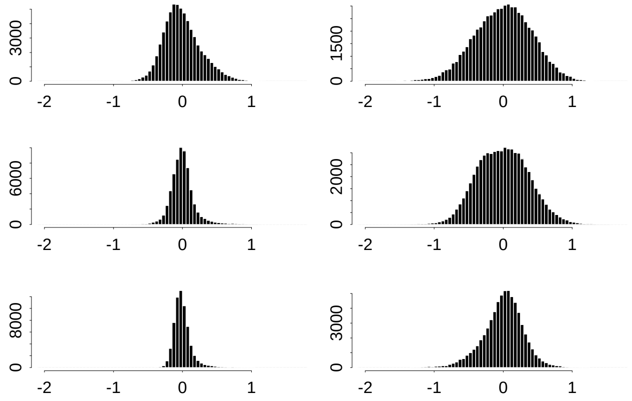

Figure 11: Histograms for CEM abundances

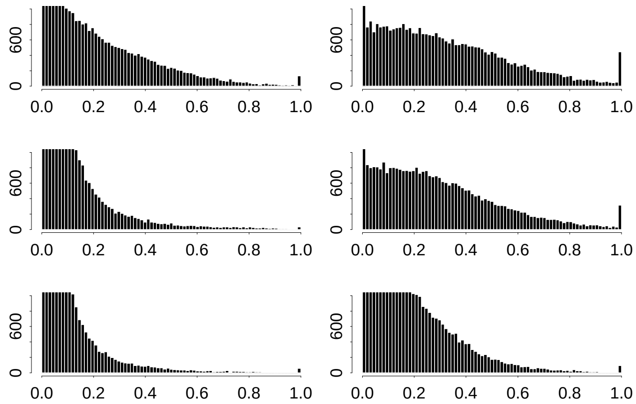

Figure 12: Histograms for CEM abundances saturated at 0 and 1

[10] Joseph C. Harsanyi and Chein-I Chang. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Transactions on Geoscience and Remote Sensing*, 32(4):779–785, 1994.

[11] Anne Jacobsen, Kathleen B. Heidebrecht, and Allan Aasbjerg Nielsen. Monitoring grasslands using convex geometry and partial unmixing - a case study. In M. Schaepman, D. Schläpfer, and K. Itten, editors, *Proceedings of the First EARSeL Workshop on Imaging Spectroscopy*, Zürich, Switzerland, 6-8 October 1998.

[12] A. S. Jacobson, A. L. Berkin, and M. N. Orton. Linkwinds: Interactive scientific data analysis and visualization. *Communications of the ACM*, 37(4):42–52, 1994. Internet http://linkwinds.jpl.nasa.gov/.

[13] J. T. Kent and K. V. Mardia. Spatial classification using fuzzy membership models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):659–671, 1988.

[14] F. A. Kruse, A. B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz. The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44:145–163, 1993.

[15] M. S. Landy. HIPS-2 software for image processing: Goals and directions. In K. L. Boyer and L. Stark, editors, *Proceedings of the SPIE 1964 Applications of Artificial Intelligence 1993: Machine Vision and Robotics*, pages 382–391, 1993. Internet http://www.cns.nyu.edu/home/msl/hipsdescr.cgi/.

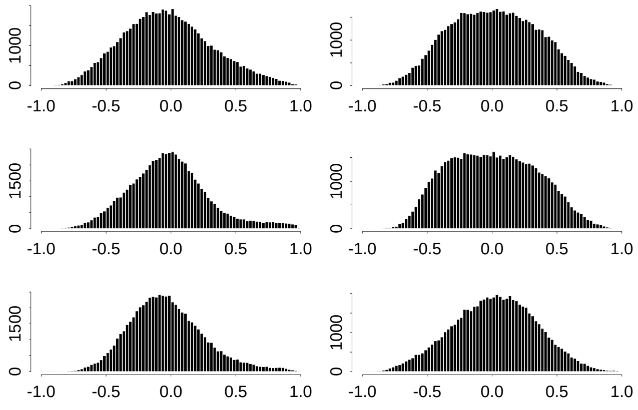

Figure 13: Histograms for cosines of spectral angles

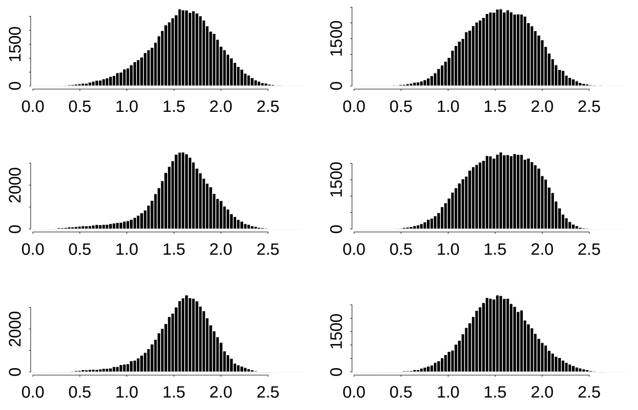

Figure 14: Histograms for spectral angles

[16] M. S. Landy, Y. Cohen, and G. Sperling. HIPS: A UNIX based image processing system. *Computer Vision, Graphics and Image Processing*, 25(3):331–347, 1984. Internet http://www.cns.nyu.edu/home/msl/hipsdescr.cgi/.

[17] M. S. Landy, Y. Cohen, and G. Sperling. HIPS: Image processing under UNIX. Software and applications. *Behavior Research Methods, Instrumentation, and Computers*, 16(2):199–216, 1984. Internet http://www.cns.nyu.edu/home/msl/hipsdescr.cgi/.

[18] Rasmus Larsen, Allan Aasbjerg Nielsen, and Knut Conradsen. Restoration of hyperspectral push-broom scanner data. In Preben Gudmandsen, editor, *Proceedings of the 17th EARSeL Symposium on Future Trends in Remote Sensing*, pages 157–162, Lyngby, Denmark, June 1997.

[19] J. B. Lee, A. S. Woodyatt, and M. Berman. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing*, 28(3):295–304, 1990.

[20] S. E. Marsh, P. Switzer, W. S. Kovalik, and R. J. P. Lyon. Resolving the percentage of component terrains within single resolution elements. *Photogrammetric Engineering and Remote Sensing*, 46:1079–1086, 1980.

[21] John W. V. Miller, James B. Farison, and Youngin Shin. Spatially invariant image sequences. *IEEE Transactions on Image Processing*, 1(2):148–161, 1992.

[22] Allan Aasbjerg Nielsen. *Analysis of Regularly and Irregularly Sampled Spatial, Multivariate, and Multi-temporal Data*. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, 1994. Internet http://www.imm.dtu.dk/∼aa/phd/.

[23] Allan Aasbjerg Nielsen. Linear mixture models and partial unmixing in multi- and hyperspectral image data. In M. Schaepman, D. Schläpfer, and K. Itten, editors, *Proceedings of the First EARSeL Workshop on Imaging Spectroscopy*, Zürich, Switzerland, 6-8 October 1998.

[24] Allan Aasbjerg Nielsen, Knut Conradsen, and James J. Simpson. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bi-temporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, 64:1–19, 1998.

[25] Allan Aasbjerg Nielsen, Harald Flesche, Rasmus Larsen, Johannes M. Rykkje, and Mogens Ramm. Semi-automatic supervised classification of minerals from x-ray mapping images. In A. Buccianti, G. Nardi, and R. Potenza, editors, *Proceedings of the Fourth Annual Conference of the International Association of Mathematical Geology (IAMG'98)*, pages 473–478, Ischia, Italy, 6–8 October 1998.

[26] Allan Aasbjerg Nielsen and Rasmus Larsen. Restoration of GERIS data using the maximum noise fractions transform. In ERIM, editor, *Proceedings from the First International Airborne Remote Sensing Conference and Exhibition, Volume II*, pages 557–568, Strasbourg, France, 1994.

[27] Neil Pendock and Allan Aasbjerg Nielsen. Multispectral image enhancement neural networks and the maximum noise fraction transform. In Paul Cilliers, editor, *Proceedings of the Fourth South African Workshop on Pattern Recognition*, pages 2–13, Simon's Town, South Africa, 25–26 November 1993.

[28] R. G. Resmini, M. E. Kappus, W. S. Aldrich, J. C. Harsanyi, and M. Anderson. Mineral mapping with HYperspectral Digital Imagery Collection Experiment (HYDICE) sensor data at Cuprite, Nevada, U.S.A. *International Journal of Remote Sensing*, 18(7):1553–1570, 1997.

[29] J. J. Settle. On the relationship between spectral unmixing and subspace projection. *IEEE Transactions on Geoscience and Remote Sensing*, 34(4):1045–1046, 1996.

[30] J. J. Settle and N. A. Drake. Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14(6):1159–1177, 1993.

[31] Sorel S. Stan. *Mineral Identification and Mapping by Imaging Spectroscopy Data Analysis*. PhD thesis, Computational and Applied Mathematics Department, University of the Witwatersrand, Johannesburg, South Africa, 1997.

[32] P. Strobl, R. Richter, A. Müller, F. Lehmann, D. Oertel, S. Tischler, and A. A. Nielsen. DAIS system performance, first results from the 1995 evaluation campaigns. In ERIM, editor, *Proceedings from the Second International Airborne Remote Sensing Conference and Exhibition, Volume II*, pages 325–334, San Francisco, California, USA, June 1996.

[33] Paul Switzer and Andrew A. Green. Min/max autocorrelation factors for multivariate spatial imagery. Technical Report 6, Stanford University, 1984.

[34] G. Vane and A. F. H. Goetz. Terrestrial imaging spectroscopy. *Remote Sensing of Environment*, 24:1–29, 1988.