NON-LINEAR MIXTURE MODEL AND APPLICATION FOR ENHANCED RESOLUTION MULTISPECTRAL CLASSIFICATION

Mark J. Carlotto, PSR Corporation, 1400 Key Blvd. Suite 700, Arlington, VA 22209 USA (mjcarlotto@ma.psrw.com)

Abstract -- A new mixture model based on a probabilistic formulation is described. The model is non-linear, and unlike previous mixture models, does not need to be inverted to compute mixtures from spectra. Instead, mixtures are estimated using an interpolation approach. There are no constraints between the number of bands and endmembers. The mixture model is used to enhance the spatial resolution of classification images - increasing the classification accuracy by about 20% in the example presented. The enhanced resolution classifier is better able to detect features near the resolution of the sensor, to identify subtle patterns not fully resolved in the original data, and to more precisely delineate the boundaries of areal features.

BACKGROUND

The application of mixture models to multispectral sensors was first suggested by Horwitz el al [1] and Detchmendy and Pace [2] in order to estimate the fraction of materials present within a pixel. A linear model is generally used to relate the spectral response of a mixed pixel to the spectral responses of possible material types (endmembers) weighted by their relative fraction within the pixel:

$$x_n = \sum_m r_{mn} y_m \tag{1}$$

or $\mathbf{x} = \mathbf{R}\mathbf{y}$ where x_n is the spectral response of a mixed pixel in the n-th band, r_{mn} is the spectral response of the m-th material in the n-th spectral band and y_m is the fraction of the m-th surface material present in the pixel where $0 \le y_m \le 1$ and $\Sigma y_m = 1$. In order to compute $\mathbf{y}$ from $\mathbf{x}$, the mixture model must be inverted. However, standard techniques for solving linear equations cannot be used since they do not insure that the mixture proportions computed from arbitrary spectra will be non-negative and sum to one. A variety of techniques have thus been developed to address this problem [1-4]. This paper describes a new approach to the problem based on ideas previously introduced in [5].

NON-LINEAR MIXTURE MODEL

Given a training set consisting of K spectral N-vectors $\mathbf{x}_k$ and corresponding mixture M-vectors $\mathbf{y}_k$, our goal is to estimate the mixture vector $\hat{\mathbf{y}}$ for any spectral vector $\mathbf{x}$ in the image. If one assumes the existence of an underlying joint density $p(\mathbf{x},\mathbf{y}) = p(\mathbf{u})$, it can be shown that the optimal estimate for $\mathbf{y}$ given $\mathbf{x}$ is the conditional mean [6]

$$\hat{\mathbf{y}} = \int \mathbf{y} p(\mathbf{y}|\mathbf{x}) d\mathbf{y} = \frac{\int \mathbf{y} p(\mathbf{y},\mathbf{x}) d\mathbf{y}}{\int p(\mathbf{y},\mathbf{x}) d\mathbf{y}} \tag{2}$$

The histogram $\frac{1}{K}\delta(\mathbf{u}-\mathbf{u}_k)$ is a poor estimate of the joint density since it cannot be used to determine mixture vectors for spectra outside the training set. Instead we approximate the joint density by the histogram convolved by a Gaussian smoothing function (Parzen window):

$$\hat{p}(\mathbf{u}) = c\sum_{k=1}^{K} \exp\left[-\tfrac{1}{2}(\mathbf{u}-\mathbf{u}_k)^T \mathbf{D}^{-1}(\mathbf{u}-\mathbf{u}_k)\right] \tag{3}$$

where c is a normalizing constant and $\mathbf{D}$ is a diagonal matrix. The estimated mixture $\hat{\mathbf{y}}$ can then be expressed as a sum of the $\mathbf{y}_k$ mixtures weighted by a (non-linear) function of the distances between the associated spectral vectors $\mathbf{x}_k$ and $\mathbf{x}$

$$\hat{\mathbf{y}} = \frac{\sum_{k=1}^{K} \mathbf{y}_k \exp\left[-\tfrac{1}{2}(\mathbf{x}-\mathbf{x}_k)^T \mathbf{D}_x^{-1}(\mathbf{x}-\mathbf{x}_k)\right]}{\sum_{k=1}^{K} \exp\left[-\tfrac{1}{2}(\mathbf{x}-\mathbf{x}_k)^T \mathbf{D}_x^{-1}(\mathbf{x}-\mathbf{x}_k)\right]} \tag{4}$$

where $\mathbf{D}_x$ is an NxN diagonal matrix with elements σ_n^2. It can be shown that if the elements of the $\mathbf{y}_k$ mixture vectors are non-negative and sum to one, the elements of $\hat{\mathbf{y}}$ will also be non-negative and sum to one for any value of $\mathbf{x}$.

In (4) the σ_n^2 variances control the amount of smoothing along each of the N spectral dimensions. Assume that $\sigma_n^2 = \gamma\,\beta_n^2$ where β_n^2 is the variance in the n-th spectral band and γ is a parameter. Let

$$\mathbf{y}_{(k)}(\gamma) = \frac{\sum_{k' \ne k} \mathbf{y}_k \exp\left[-\tfrac{1}{2}(\mathbf{x}_k-\mathbf{x}_{k'})^T \mathbf{D}_\gamma^{-1}(\mathbf{x}_k-\mathbf{x}_{k'})\right]}{\sum_{k' \ne k} \exp\left[-\tfrac{1}{2}(\mathbf{x}_k-\mathbf{x}_{k'})^T \mathbf{D}_\gamma^{-1}(\mathbf{x}_k-\mathbf{x}_{k'})\right]} \tag{5}$$

be the estimate of the k-th mixture vector using all $\mathbf{y}_{k'}$ except for $\mathbf{y}_k$ where $\mathbf{D}_\gamma$ is an NxN diagonal matrix with elements $\sigma_n^2 = \gamma\,\beta_n^2$. These estimates can be used, in turn, to estimate the average rms error over the training set as a function of γ

$$\varepsilon(\gamma)=\frac{1}{K}\sum_k\sqrt{\left[\mathbf{y}_k-\mathbf{y}_{(k)}(\gamma)\right]^T\left[\mathbf{y}_k-\mathbf{y}_{(k)}(\gamma)\right]} \qquad (6)$$

Optimal values for the σ_n^2 are determined by sweeping the parameter γ over a given range and finding the value with the minimum error. This value γ_0 is then used to compute the variances $\sigma_n^2=\gamma_0\,\beta_n^2$.

ENHANCED RESOLUTION CLASSIFIER

Mixture models are typically used to detect the presence of particular materials within a pixel (sub-pixel detection) [7]. Here we use the above mixture model to enhance the spatial resolution of classification images. Since the mixture vector gives the fraction of each class present in a pixel, it can be used as a basis for sub-dividing pixels.

Assume a 3x3 neighborhood (Fig. 1). Let $Q=\{0,1,2,3,4,5,6,7,8\}$ be the set of indices of the pixels in the neighborhood. We shall use the index q to denote the location of a pixel within a 3x3 block of pixels as well as the location of a sub-pixel within the central 3x3 block of sub-pixels. $y_{m,q}$ is the fraction of the m-th class in the q-th pixel. The mixture vector $\mathbf{y}_0=\{y_{m,0}\}$ is used to divide the central pixel into a block of 9 sub-pixels denoted z_0-z_8 where each sub-pixel is assigned one of the original M classes. The number of sub-pixels in each class is proportional to the fraction for that class.

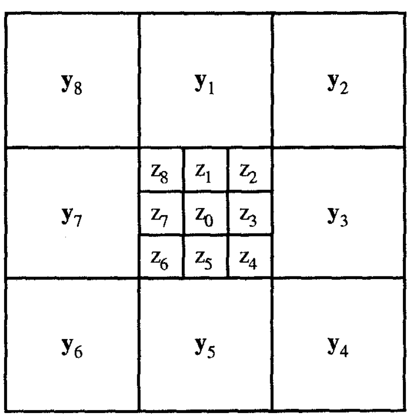


Fig.1 3x3 neighborhood scheme (the central pixel is $\mathbf{y}_0$)

With the exception of features near the resolution limit of the sensor, one typically observes in remotely-sensed imagery that the probability neighboring pixels are the same class is greater than the probability they are different classes. We thus attempt to position the sub-pixels within the central block so that they will likely be near other sub-pixels of the same class in neighboring blocks.

Because of the large number of possible configurations, we have developed the following heuristic algorithm to position sub-pixels, one block at a time:

Do { Assign class $m'=\arg\max_m\{y_{m,0}\}$ to a sub-pixel;
Find neighboring pixel $q'=\arg\max_{q\in Q}\{y_{m',q}\}$ with largest fraction in class m';
Move sub-pixel to corresponding position in block $z_{q'}=m'$;
Decrement allocated fraction $y_{m',0}=y_{m',0}-\frac{1}{9}$;
Delete neighbor q' from Q;
} While $Q\neq\varnothing$

EXPERIMENTAL RESULTS

The accuracy of the non-linear mixture model and enhanced resolution classifier were evaluated using simulated Landsat TM data sets at 5 and 25 meter/pixel. The data sets were derived from 16-channel M7 data set over Gordonsville, VA by averaging and sub-sampling. 192 ground truth samples containing at least 9 contiguous 5 meter pixels were extracted and assigned one of 12 classes. These data were used to classify the 5m simulated TM by assigning the class with the highest fraction on a pixel-by-pixel basis. The 25m simulated TM was similarly classified using the same ground truth resampled to 25m.

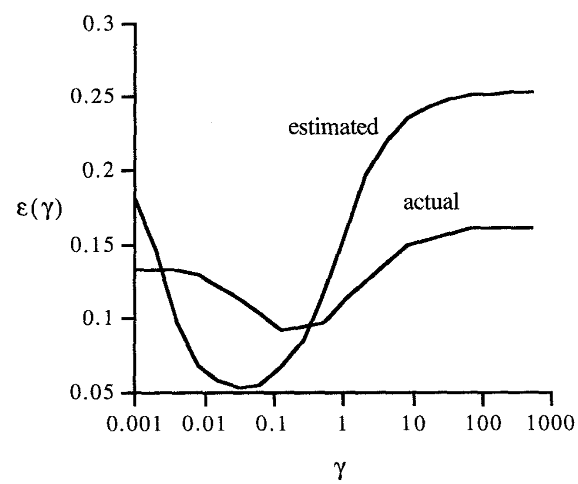


Fig. 2 Rms errors for different γ smoothing factors

A series of 8.33m enhanced classifications were then generated from the 25m imagery using different smoothing factors. Fig. 2 shows the estimated rms error over the training set (6), and the actual rms error between the resultant 8.33m enhanced resolution classifications and the 5m max-fraction classification which was used as the reference. We compared the proportion of classes within 5x5 blocks in the 5m max-fraction classification to 3x3 blocks in the 8.33m enhanced resolution classification. The minimum rms error was 0.093. For comparison, the rms error between the 25m and 5m max-fraction classifications was 0.12. By sub-dividing 25m pixels the rms error was reduced by about 20% in this example. The minimum error estimated over the training set is lower than the actual error and occurs at a lower γ because the training set contains only pure pixels.

Fig. 3 presents black-and-white renditions of classification results over a portion of the Gordonsville data set. The region shown contains a stream and forested wetland running from the upper left to the lower right. Patterns emanating from the stream can be seen in the 5m classification a) but are less evident in the 25m classification b). However, the enhanced resolution classification result in c) shows definite indications of the patterns seen in a).

SUMMARY

A new mixture model based on a probabilistic formulation was presented. The model is non-linear, and unlike previous mixture models, does not need to be inverted to compute mixtures from spectra. Instead the model uses spectral vectors and associated mixture vectors from a training set to estimate mixtures from spectra by interpolation. There are no constraints between the number of bands and endmembers. The mixture model was then used to enhance the spatial resolution of classification images. In addition to improving classification accuracy the enhanced resolution classifier is better able to detect features such as roads near or below the resolution of the sensor, to identify subtle patterns not fully resolved in the original data, and to more precisely delineate the boundaries of areal features.

REFERENCES

[1] H.M. Horwitz, R.F. Nalepka, P.D. Hyde, and J.P. Morgenstern, "Estimating the proportions of objects within a single resolution element of a multispectral scanner," Proc. 7-th Int. Symp. Remote Sens. Environ., Ann Arbor, MI, 1971.

[2] D.M. Detchmendy and W.H. Pace, "A model for spectral signature variability for mixtures," Earth Resources Observation and Information Analysis Systems Conf., Tullahoma, TN, 1972.

[3] C.R. Hallum, "On a model for optimal proportions estimation for category mixtures," Proc. 8-th Int. Symp. Remote Sens. Environ., Ann Arbor, MI, 1972.

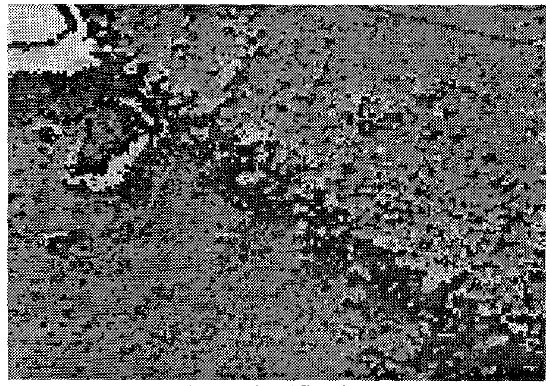
[4] Y.E. Shimabukuro and J.A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," IEEE Trans. Geo. and Remote Sensing, Vol. 29, No. 1, 1991.

[5] M.J. Carlotto and D. Izraelevitz, "System realization using associative memory building blocks," SPIE Vol. 1192, 1989.

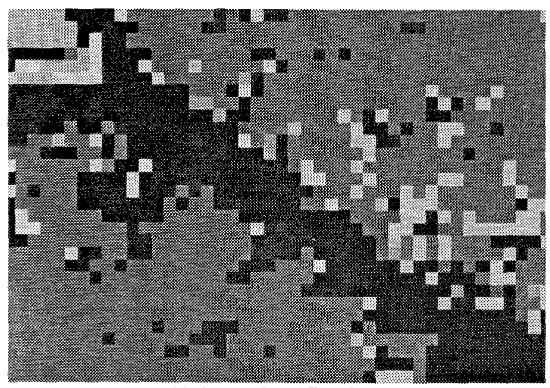
[6] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw Hill, 1965.

[7] R.L. Huguenin, "Subpixel analysis process improves accuracy of multispectral classifications," Earth Observation Magazine, July 1994.

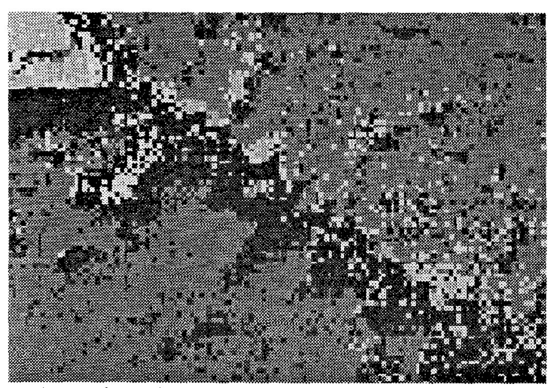
The author would also like to thank Keith Hartt and Michael Stein for early contributions to this work.



a) 5m classification



b) 25m classification



c) Enhanced resolution (8.33m) classification derived from 25m imagery

Fig. 3 Comparison of classification results