SPECTRAL UNMIXING USING THE BETA COMPOSITIONAL MODEL

Alina Zare[*]*, Paul Gader*[†]*, Dmitri Dranishnikov*[†]*, Taylor Glenn*[†]

[*]Electrical and Computer Engineering, University of Missouri
[†]Computer and Information Science and Engineering, University of Florida

ABSTRACT

This paper introduces a beta compositional model as a mixing model for hyperspectral images. Endmembers are represented via beta distributions, hereafter referred to as betas, to constrain endmembers to a physically-meaningful range. Two associated spectral unmixing algorithms are described and applied to simulated and real hyperspectral imagery.

1. THE BETA COMPOSITIONAL MODEL

The *Linear Mixing Model* (LMM) models pixels as convex combinations of endmembers [1, 2]. The LMM represents endmembers as single spectra. However, spectral signatures of materials will vary for many reasons including illumination, atmospheric or environmental conditions, and inherent material variability. Several methods have been investigated to represent variability including grouping endmembers into sets (i.e., "endmember bundles") [3] or using Gaussians [4]. Unmixing and endmember estimation methods using these representations have been developed [5–9].

Here, betas are proposed to represent endmembers. Spectral pixels are often converted to reflectance, constrained to the interval $[0, 1]$. The domain of the beta is also $[0, 1]$. Therefore, using betas, endmembers are constrained to physically meaningful values. In contrast, the domain of the Gaussian is $[-\infty, \infty]$, which allows samples outside of $[0, 1]$.

The *Beta Compositional Model* (BCM) represents endmembers as random vectors where each band is distributed according to a univariate beta distribution,

$$e_{md} \sim \mathcal{B}(\cdot|\alpha_{md}, \beta_{md}) \tag{1}$$

where e_{md} is the d^{th} band of the m^{th} endmember, $\mathcal{B}$ is the beta distribution shown in (2), and $\alpha_{md} \geq 0$ and $\beta_{md} \geq 0$ are the beta parameters for the d^{th} band of the m^{th} endmember.

$$\mathcal{B}(e|\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta)^{-1}e^{\alpha-1}(1-e)^{\beta-1} \tag{2}$$

The model presented here does not represent statistical relationships between bands. However, the appropriate approach is under investigation, (see Section 4).

The spectra in a BCM are random vectors distributed as a convex combination of beta endmember distributions,

$$x_{id} \sim \mathcal{F}(\cdot|\mathbf{p}_i, \mathbf{a}_d, \mathbf{b}_d) \tag{3}$$

where x_{id} is the d^{th} band of the i^{th} data point, $\mathbf{p}_i$ an $M \times 1$ vector of abundances associated with data point i satisfying (4), $\mathbf{a}_d$ and $\mathbf{b}_d$ are $M \times 1$ vectors containing the α and β parameters for the d^{th} band of the M endmembers, and $\mathcal{F}$ is the distribution for a convex combinations of betas.

$$p_{ik} \geq 0 \quad \forall k = 1, \ldots, M; \quad \sum_{k=1}^{M} p_{ik} = 1 \tag{4}$$

For example, in the case of two beta endmembers, the distribution of the resulting convex combination is $\mathcal{F}_2$,

$$\mathcal{F}_2(x|\mathbf{p}, \mathbf{a}, \mathbf{b}) = \begin{cases} f_1(x|\mathbf{p}, \mathbf{a}, \mathbf{b}) & \text{if } 0 \leq x \leq p_1 \\ f_2(x|\mathbf{p}, \mathbf{a}, \mathbf{b}) & \text{if } p_1 \leq x \leq p_2 \\ f_3(x|\mathbf{p}, \mathbf{a}, \mathbf{b}) & \text{if } p_2 \leq x \leq 1 \end{cases} \tag{5}$$

where $p_1 \leq p_2$ and F_1 is the first Appell function [10],

$$\begin{aligned} f_1(x) &= \frac{x^{\alpha_1+\alpha_2-1}(p_1-x)^{\beta_1-1}B(\alpha_1,\alpha_2)}{p_1^{\alpha_1+\beta_1-1}p_2^{\alpha_2+\beta_2-1}B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)} \qquad (6)\\ &\quad F_1\left(\frac{-x}{p_1-x}, \frac{x}{p_2}, \alpha_2, 1-\beta_1, 1-\beta_2, \alpha_1+\alpha_2\right) \\ f_2(x) &= \frac{(x-p_1)^{\alpha_2-1}(1-x)^{\beta_2-1}}{p_2^{\alpha_2+\beta_2-1}B(\alpha_2,\beta_2)} \qquad (7)\\ &\quad F_1\left(\frac{-p_1}{x-p_1}, \frac{p_1}{1-x}, \beta_1, 1-\alpha_2, 1-\beta_2, \alpha_1+\beta_1\right) \\ f_3(x) &= \frac{1-x^{\beta_1+\beta_2-1}(x-p_1)^{\beta_1-1}B(\beta_1,\beta_2)}{p_1^{\beta_1}p_2^{\alpha_2-\beta_2-1}B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)} \qquad (8)\\ &\quad F_1\left(\frac{1-x}{p_1}, \frac{x-1}{x-p_1}, \beta_1, 1-\alpha_1, 1-\alpha_2, \beta_1+\beta_2\right). \end{aligned}$$

Section 2 describes unmixing algorithms using the BCM. Section 3 presents results on simulated and real hyperspectral data. Section 4 summarizes and discusses future work.

2. SPECTRAL UNMIXING WITH THE BCM

Two unmixing algorithms based on the BCM are presented here. Both algorithms use the approximation in Section 2.1.

2.1. Approximating a Convex Combination of Betas

The distribution of a convex combination of betas is complicated, as seen in (5). To reduce complexity of BCM-based unmixing, an approximation is considered. In [11], a single beta, $\mathcal{B}(\cdot|e, f)$ with f and e given by (9) and (10), is used to approximate the distribution governing the random variable $X_i = \sum_{m=1}^{M} p_{im} E_{md}$,

$$f = \left(F - S(1+F)^2\right) S^{-1} (1+F)^{-3} \quad (9)$$
$$e = Ff \quad (10)$$

such that, letting $C_{md} = \alpha_{md} + \beta_{md}$ and $g = e + f$,

$$F := E_i (1-E)^{-1} \quad (11)$$
$$E := E(X_{id}) = \sum_{m=1}^{M} p_m E(E_{md}) \quad (12)$$
$$= eg^{-1} = \sum_{m=1}^{M} p_m \alpha_{md} C_{md}^{-1}$$
$$S := Var(X_{id}) = \sum_{m=1}^{M} p_{im}^2 Var(E_{md}) \quad (13)$$
$$= efC_{mn}^{-2}(C_{mn}+1)^{-1}$$
$$= \sum_{m=1}^{M} p_{im}^2 \alpha_{md}\beta_{md} C_{md}^{-2}(C_{md}+1)^{-1}.$$

The α_{md} and β_{md} values are assumed to be known as these parameters define the known endmembers. If many pure examples of spectra are provided for each endmember material, these α_{md} and β_{md} could be estimate, for example, using a maximum likelihood estimation approach.

Twelve experiments were performed to explore the approximation accuracy. In the n^{th} experiment, the following process was repeated 10 times: $K = 50,000$ pairs of proportions were drawn from $\mathcal{B}(d_1^{(n)}, d_2^{(n)})$ and used to construct K linear combinations of two betas, $\mathcal{B}(\alpha_1^{(n)}, \beta_1^{(n)})$ and $\mathcal{B}(\alpha_2^{(n)}, \beta_2^{(n)})$, resulting in K normalized histograms. In addition, K normalized histograms were produced using the approximation $\mathcal{B}(e^{(n)}, f^{(n)})$. The parameters $d_1^{(n)}, d_2^{(n)}, \alpha_1^{(n)}, \beta_1^{(n)}, \alpha_2^{(n)}$, and $\beta_2^{(n)}$ were pre-selected. The histograms used a bin width of 0.01. The normalized histograms from the linear combinations were compared to those from the approximation using the Symmetric K-L Divergence (SKLD). Results are in Table 1. Figure 1 compares approximated pdfs.

2.2. Method 1: Quadratic Programming Approach

In the proposed unmixing methods, single betas are fit to groups of spectrally similar data points. Then, using the approximation in (9) to (13) and known beta endmember distributions, proportion values are determined by minimizing the difference between the mean and/or variance of the fitted single betas to the linear combination of beta endmember distributions. In the first approach, only the difference in the mean values between the convex combination of beta endmembers

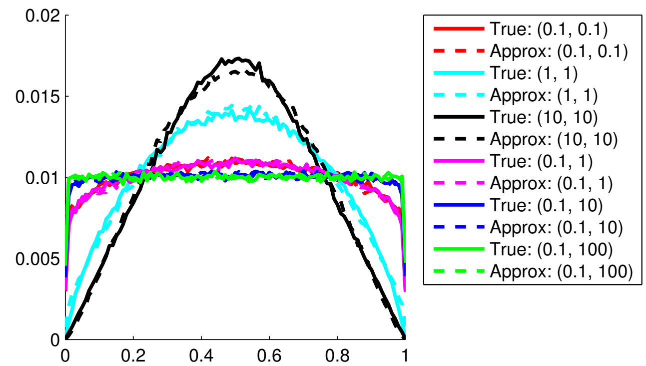


Fig. 1. Comparison of pdfs generated by sampling from a linear combination of betas and by sampling from the corresponding approximation for the linear combination of betas. Legend lists the Dirichlet parameters used to sample proportion values. Parameters $\alpha_1, \beta_1, \alpha_2, \beta_2$ are set to 1.

Table 1. **SKLD between BCM and approximation.**

$d_1^{(n)}$	$d_2^{(n)}$	$\alpha_1^{(n)}$	$\beta_1^{(n)}$	$\alpha_2^{(n)}$	$\beta_2^{(n)}$	$SKLD \times 1000$
0.1	0.1	1.0	1.0	1.0	1.0	5 ± 0.5
1.0	1.0	1.0	1.0	1.0	1.0	9 ± 0.6
0.1	1.0	1.0	1.0	1.0	1.0	4 ± 0.4
0.1	10.0	1.0	1.0	1.0	1.0	4 ± 0.6
1.0	10.0	1.0	1.0	1.0	1.0	2 ± 2.0
2.0	5.0	1.0	1.0	1.0	1.0	8 ± 0.8
1.0	1.0	0.1	1.0	1.0	1.0	8 ± 0.9
1.0	1.0	10	1.0	1.0	1.0	9 ± 0.7
1.0	1.0	0.1	1.0	1.0	0.1	9 ± 1.0
1.0	1.0	10	1.0	1.0	10	8 ± 0.9
1.0	1.0	0.1	1.0	0.1	1.0	8 ± 1.2
1.0	1.0	2.0	5.0	2.0	5.0	8 ± 0.9

Algorithm 1 Beta Compositional Model Spectral Unmixing - Quadratic Programming Approach

Inputs: $\mathbf{x_i}$, for $i = 1, \ldots, N$ and α_{md} and β_{md} for $m = 1, \ldots, M; d = 1, \ldots, D$

FOR $n \leftarrow 1, \ldots, N$

  Identify K nearest neighbors for pixel $\mathbf{x}_i$

  Estimate e_d and f_d (e.g. using MATLAB's betafit) to fit a beta distribution to each band of the K neighbors

  Return proportion values found by using quadratic programming to minimize (14) subject to the constraints in (4).

ENDFOR

and the single beta fit to to the data are minimized. In the subsequent Metropolis-Hastings approach, both the approximated mean and variance are used to estimate proportions.

First, the K-nearest neighbors of each spectral pixel are found. A beta is fit to each dimension of the K-nearest neighbors of each pixel. This estimates the parameters e_d and f_d for $d = 1 \ldots D$ of the beta approximating the linear combination of endmembers. The values e_d and f_d and the known α_{md} and β_{md} for each beta endmember, can be use to estimate the proportions for the i^{th} pixel can be found by minimizing the following objective function subject to the constraints in (4),

$$J = \left(\frac{e_d}{e_d + f_d} - \sum_{m=1}^{M} p_{im} \frac{\alpha_{md}}{\alpha_{md} + \beta_{md}} \right)^2 \qquad (14)$$

2.3. Method 2: Metropolis Hastings Sampler

The stochastic model for the d^{th} band of a pixel is

$$E = \sum_{m=1}^{M} p_m \alpha_{md} C_{md}^{-1} \qquad (15)$$

$$S = \sum_{m=1}^{M} p_m^2 \alpha_{md} \beta_{md} C_{md}^{-2} (C_{md} + 1)^{-1} \qquad (16)$$

which leads to the problem of maximizing the likelihood in (17), where $\Theta = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_m, \sigma_v)$.

$$E, S | \Theta \sim \mathcal{N}\left(E \middle| \sum_{m=1}^{M} p_m \alpha_{md} C_{md}^{-1}, \sigma_m \right) \times \qquad (17)$$
$$\mathcal{N}\left(S \middle| \sum_{m=1}^{M} p_m^2 \alpha_{md} \beta_{md} C_{md}^{-2} (C_{md}+1)^{-1}, \sigma_v \right)$$

The proposal distribution[12] used here to sample proportions is $p(\mathbf{p}_i^{new} | \mathbf{p}_i^{old}) = Dir(\cdot | \mathbf{1}_M)$ where $\mathbf{1}_M$ is a $1 \times M$ vector of ones. Given this proposal distribution, the acceptance ratio used to evaluate $\mathbf{p}^{new}$ is

$$a = \frac{\Pi(\mathbf{p}^{new} | \mathbf{X}, \alpha, \beta, \sigma_m, \sigma_v)}{\Pi(\mathbf{p}^{old} | \mathbf{X}, \alpha, \beta, \sigma_m, \sigma_v)}$$

where $\Pi(\mathbf{p} | \mathbf{X}, \alpha, \beta, \sigma_m, \sigma_v)$ is proportional to (17). The σ_m and σ_v parameters weight the relative importance of the mean and variance factors in the likelihood. In all experiments here, $\sigma_m = 0.001$ and $\sigma_v = 100$.

Algorithm 2 BCM Spectral Unmixing - Metropolis-Hastings Sampler

Inputs: $\mathbf{x_i}$, for $i = 1, \ldots, N$, and α_{md} and β_{md} for $m = 1, \ldots, M; d = 1, \ldots, D$

FOR $n \leftarrow 1, \ldots, N$

  Identify K nearest neighbors for pixel $\mathbf{x}_i$

  Estimate e_d and f_d (e.g. using MATLAB's betafit) to fit a beta distribution to each band of the K neighbors

  FOR $i \leftarrow 1, \ldots,$ Number of Iterations

    Sample a proportion vector from the proposal Dirichlet

    Accept or reject the sample using (18)

  ENDFOR

  Set $\mathbf{p}_i$ to sample with the largest likelihood (17)

ENDFOR

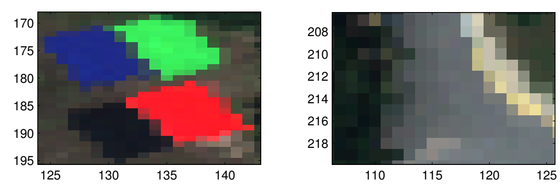


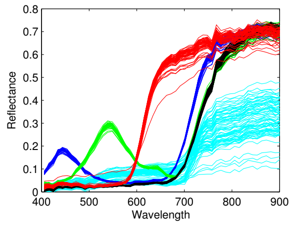
Fig. 2. RGB images of the two scenes over Gulfport, MS.

3. EXPERIMENTAL RESULTS

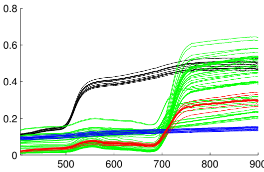
Experimental results are shown for two hyperspectral images collected over Gulfport, MS, using a CASI-1500 with spatial resolution of $1m^2$ and wavelengths ranging from 400 to 900 nm. Figure 2 shows RGB images of the two scenes. Since the beta endmember distributions are assumed to be known, for scene 1, betas were fit to spectra for each material pulled from the image (using maximum likelihood estimation) and, for scene 2, beta endmembers were estimated from ground spectra collected with a handheld ASD-spectrometer. The spectra used for each endmember in the two images are shown in Figure 3. Both of the proposed spectral unmixing methods were applied the images. Figure 4 shows the estimated proportion maps.

4. SUMMARY AND FUTURE WORK

This paper introduces the beta compositional model for hyperspectral unmixing and endmember estimate. Two spectral unmixing algorithms based on this model are presented.

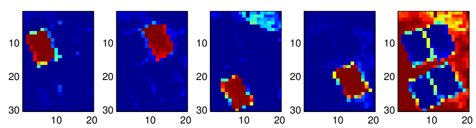


(a) Blue: Blue Cloth, Green: Green Cloth, Red: Red Cloth, Black: Black Cloth, Cyan: Grass

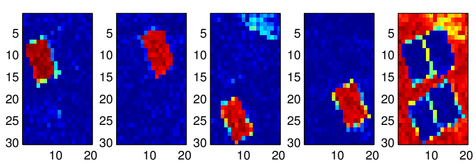


(b) Blue: Asphalt, Black: Yellow Curb, Green: Oak Leaves, Red: Grass

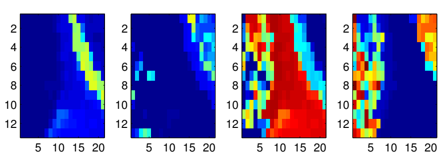
Fig. 3. (a) Image spectra used to estimate beta endmember distributions. Materials included four sheets (colored red, green, blue and black) and grass (b) Ground spectra collected using an ASD-spectrometer (i.e., from a spectral library). Materials included yellow curb, asphalt, oak leaves and grass. Note the grass and oak leaves have similar mean spectral shape but different variance.



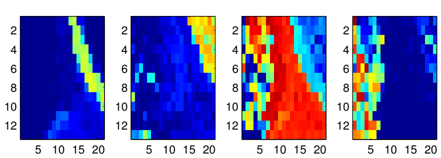
(a) QP: Blue Cloth, Green Cloth, Black Cloth, Red Cloth, Grass



(b) MH: Blue Cloth, Green Cloth, Black Cloth, Red Cloth, Grass



(c) QP: Yellow Curb, Grass, Asphalt, Oak



(d) MH: Yellow Curb, Grass, Asphalt, Oak

Fig. 4. Proportion maps estimated for scene 1 using the (a) QP and (b) MH methods and for scene 2 using the (c) QP and (b) MH methods. Note for scene 2, the inclusion of the variance term in the MH approach provided the ability to distinguish between grass and oak leaves where as the QP method could do not since it does not take into account variance during proportion estimation.

There are many avenues for future investigation. These include developing methods to estimate beta endmember distributions given input hyperspectral data, expanding the beta compositional model to represent covariance between spectral bands, and develop additional spectral unmixing algorithms (such as a gradient descent approach) that balance computational efficiency with proportion accuracy.

References

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, pp. 44–57, 2002.

[2] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[3] C.A. Bateson, G.P. Asner, and C.A. Wessman, "Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 2, pp. 1083 –1094, mar 2000.

[4] M. T. Eismann and D. W. J. Stein, *"Stochastic mixture modeling," in Hyperspectral Data Exploitation: theory and applications*, vol. 148, John Wiley & Sons, 2007.

[5] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603–1616, 2011.

[6] G. P. Asner and D. B. Lobell, "A biogeophysical approach for automated swir unmixing of soils and vegetation," *Remote Sensing of Environment*, vol. 74, pp. 99–112, 2000.

[7] P. E. Dennison and D. A. Roberts, "Multiple endmember spectral mixture analysis using endmember average rsme," *Remote Sensing of Environment*, vol. 87, no. 2-3, pp. 123–135, 2003.

[8] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model: Application to hyperspectral imagery," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1403–1413, June 2010.

[9] A. Zare, P. Gader, and G. Casella, "Sampling piecewise convex unmixing and endmember extraction," *IEEE Transactions on Geoscience and Remote Sensing*, 2013.

[10] T. Pham-Gia and N. Turkkan, "Distribution of the linear combination of two general beta variables and applications," *Communications in Statistics - Theory and Methods*, vol. 27, no. 7, pp. 1851–1869, 1998.

[11] B. Johannesson and N Giri, "On approximations involving the beta distribution," *Communications in Statistics - Simulation and Computation*, vol. 24, no. 2, pp. 489–503, 1995.

[12] Chib S. and Greenberg E., "Understanding the metropolis-hastings algorithm," *The American Statistician*, vol. 49, no. 4, pp. 327–335, april 1995.