Ordering on the probability simplex of endmembers for hyperspectral morphological image processing

Gianni Franchi, Jesus Angulo

Ordering on the probability simplex of endmembers for hyperspectral morphological image processing

Gianni Franchi and Jesús Angulo

MINES ParisTech, PSL-Research University,
CMM-Centre de Morphologie Mathématique; France
gianni.franchi@mines-paristech.fr,jesus.angulo@mines-paristech.fr

Abstract. A hyperspectral image can be represented as a set of materials called endmembers, where each pixel corresponds to a mixture of several of these materials. More precisely pixels are described by the quantity of each material, this quantity is often called abundance and is positive and of sum equal to one. This leads to the characterization of a hyperspectral image as a set of points in a probability simplex. The geometry of the simplex has been particularly studied in the theory of quantum information, giving rise to different notions of distances and interesting preorders. In this paper, we present total orders based on theory of the ordering on the simplex. Thanks to this theory, we can give a physical interpretation of our orders.

Keywords: Hyperspectral image, mathematical morphology, learning an order, quantum information

1 Introduction

Hyperspectral images, which represent a natural extension of conventional optical images, can reconstruct the spectral profiles through the acquisition of hundreds of narrow spectral bands, generally covering the entire optical spectral range. Thus, at each pixel, there is a vector which corresponds to the spectrum of reflected light. For a long time the processes associated with this type of images were limited to treatments where each pixel was considered just as a vector independently of its location on the image domain. Subsequently techniques to account for spatial information were developed [10] [12]. Between these techniques, mathematical morphology has been also used [13][19]. Adding spatial information in the treatment of hyperspectral images greatly improve tasks such as classification.

Mathematical morphology is a non-linear methodology for image processing based on a pair of adjoint and dual operators, dilation and erosion, used to compute sup/inf-convolutions in local neighborhoods. Therefore the extension of morphological operators to multivariate images, and in particular to hyperspectral images, requires the introduction of appropriate vector ordering strategies.

Lots of research done has aimed at developing orders in multivariate images. First, we found techniques based on marginal order [15]: one applies a morphological operator at each channel of the hyperspectral image (or each eigenimage after PCA)[12], without taking care of the vector nature of spectral information. Then there are other approaches based on learning a vector order using classification techniques [17] or nonparametric statistical tools [18]. Supervised learning an order [17] involves to introduce a prior spectral background and foreground. This happens if, for example, one knows that the vegetation areas are the object of interest, and does not care of image structures from other materials: then prior vegetation spectra would be the foreground, and the spectra of other materials the background. But, sometimes, this prior configuration is unknown. There might be two possible solutions. One approach proposed in [18] consists in computing an order based on the "distance to central value", obtained by the statistical depth function. Another solution builds an order based on distances to pure materials on the image [2]: more a spectrum is a mixture of materials less its position in the order is. In hyperspectral image processing, these pure materials are called endmembers. The goal of this paper is to investigate order techniques in a physical representation similar to the one proposed in [2]. From a geometric viewpoint, the space where the spectral values are represented is not an Euclidean space, but a simplex, which is also useful to manipulate discrete probability distributions. In our case, we will consider several alternatives, which can be related to notions as majorization, stochastic dominance, distances and divergence between probability distributions, etc.

2 Hyperspectral image linear model

From a mathematical viewpoint, a hyperspectral image is considered as a function f defined as

$$f : \begin{cases} E \to \mathbb{R}^D \\ x \mapsto v_i \end{cases}$$

where D is the number of spectral bands and E is the image domain (support space of pixels. This multivariate image can be also seen as set of D grey-scale images which correspond to the spectral bands, i.e., $f(x) = (f_1(x), \cdots, f_D(x))$, $f_j : E \to \mathbb{R}$.

Unmixing on hyperspectral images. Spectral unmixing [6–8] is an ill-posed inverse problem. The idea is to find both the physical pure materials, also known as endmembers, present in the scene and their corresponding abundances. By abundance we mean the quantity of each material at each pixel. However, in hyperspectral imaging, because of the low resolution of these images, a pixel may correspond to a large surface, where several different physical materials are present: these materials can be mixed and may affect the unmixing algorithms. A model often used by its simplicity is the linear one. In this model, ground is considered flat such that when the sun's rays reach the study area, they are directly reflected to the sensor. It receives, from a convex area, a linear combination of all the materials present in the area. A strong and extreme assumption which is

often made is that one can find the endmembers from the hyperspectral image itself, because they are already present in the image, or because the geometric problem is not very complex.

Working on this paradigm, many different algorithms have been proposed to compute the endmembers [8]. We have used in our examples N-FINDR [21]. Once the endmembers are obtained, the corresponding abundance should be computed at each pixel, which consequently yields an abundance image for each endmember, i.e., $\alpha_r(x) : E \to \mathbb{R}_+$, $1 \leq r \leq R$, where $R \in \mathbb{N}$ is the number of endmembers. We have used a fully constrained least squared on each pixel [8], where one imposes that each pixel is a linear nonnegative and convex combination of endmembers. We just discuss below the geometric interpretation of such constrained regression.

Due to the fact that in general the number of endmembers is significatively lower than the number of spectral bands, i.e., $R \ll D$, working on the abundance maps $\{\alpha_j(x)\}_{1\leq j\leq R}$ instead the original spectral bands $\{f_j(x)\}_{1\leq j\leq D}$ is a way to counterpart the curse of dimensionality [4], which designates the phenomena that occur when data are studied in a large dimensionality space. In summary, thanks to unmixing techniques, the hyperspectral image is embedded into a lower dimension space.

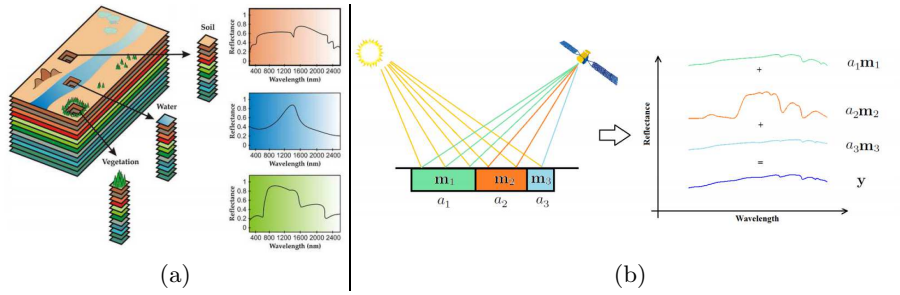

Fig. 1. (a) Representation of a hyperspectral image, (b) Zoom in the linear mixing model. Figures borrowed from [14].

Geometrical viewpoint of unmixing. Let us consider in Fig. 1(a) the typical diagram to illustrate spectral unmixing. To fix the ideas, we can consider that the image is composed of water, soil, and vegetation. Finding the endmembers on this image is therefore a way to find the spectra of these three pure materials. Under the linear model, each pixel of the image can be written as a positive combination of the different endmembers m_1, m_2 and m_3, see Figure 1(b), with a_1, a_2 and a_3 being the abundances. Thus, if we consider the spectrum at a pixel as the vector $v_i \in \mathbb{R}^D$, then, it can be written as

$$v_i = \sum_{r=1}^{R} a_{r,i} m_r + n_i, \tag{1}$$

where $\{m_r\}_{r=1}^{R}$ represent the set of R endmembers, $a_{r,i}$ the abundance at vector i of each endmember r, and n_i an additive noise. This last term can be neglected.

Let us consider a geometric interpretation of the representation by endmembers [11]. Spectra correspond to a set of points $v_i \in \mathbb{R}^D$, $1 \leq i \leq \mid E \mid$. Since each v_i represents a physical spectrum, the v_i are nonnegative data, thus every point lies in the positive orthant $\mathcal{P}^D$ of $\mathbb{R}^D$. Therefore the endmembers should be also nonnegative vectors. The endmembers basis $\Phi = \{m_r\}_{1 \leq r \leq R}$, $m_r \in \mathbb{R}^D$ generates a simplicial cone Γ_Φ containing the data and which lies in $\mathcal{P}^D$:

$$\Gamma_\Phi = \{v \ : \ v = \sum_{r=1}^{R} a_r m_r, \ a_r \geq 0\}. \tag{2}$$

Hence, the extraction of endmembers can be seen as finding the simplicial cone containing the data. In general, for a given set of vectors there are many possible simplicial cones containing the vectors. A way to reduce the number of possible representations consist in restrict the nonnegative coefficients a_r to be a convex combination such that $\sum_{r=1}^{R} a_r = 1$. By using this additional constrain, it is guaranteed to work on a $(R-1)$-simplex. We remind that the unit n-simplex is the subset of $\mathbb{R}^{n+1}$ given by

$$\Delta^n = \{(s_1, \cdots, s_{n+1})^t \in \mathbb{R}^{n+1} : \ \sum_{k=1}^{n+1} s_k = 1 \text{ and } s_k \geq 0\}. \tag{3}$$

where the $n+1$ vertices of a regular tetrahedron. In our case, the vertices are the endmembers $\{m_r\}_{1 \leq r \leq R}$, but we can work geometrically on the canonical simplex Δ^{R-1}. In summary, the abundances $(a_{1,i}, \cdots, a_{R,i})^t$ are just the barycentric coordinates of vector i in the Δ^{R-1}. In the case of the hyperspectral model studied in Fig. 1 with just 3 endmembers, its abundances lie in the triangle represented in Fig. 2(a).

3 Ordering on the simplex of endmembers

Let us consider that the hyperspectral image of spectral pixels $v_i \in \mathbb{R}^D$ is represented by their abundances, together with the set of endmembers $\{m_r\}_{1 \leq r \leq R}$. Let us consider that we have two spectral vectors v_i and v_j whose coordinates in the simplex are $(a_{1,i}, \cdots, a_{R,i})^t$ and $(a_{1,j}, \cdots, a_{R,j})^t$. We address in this section the question of how can we order this pair of vectors.

3.1 Lexicographic abundance order

An easy way to define a total order between v_i and v_j is based on a lexicographic order of their barycentric coordinates. To avoid an arbitrary order between the endmembers, we propose to order them according to their norm $\|m_r\|$, i.e., $m_r \preceq m_s \Leftrightarrow \|m_r\| \leq \|m_s\|$. Then a permutation τ according to this order

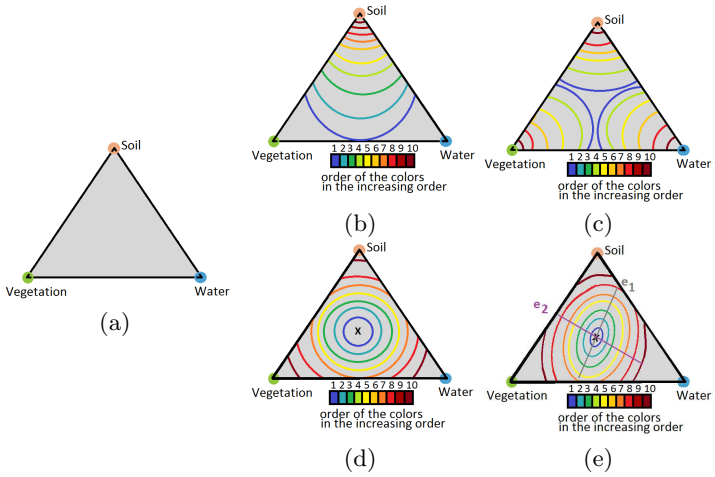

Fig. 2. (a) Simplex representation of an hyperspectral image with three endmembers. (b) First iso-order lines for $\leq_{\text{lex}}$. (c) First iso-order lines for $\leq_{\text{maj}}$. (d) Iso-order lines for $\leq_{d_{L^2}}$. (e) Iso-order lines for $\leq_{d_M}$.

of endmembers is applied on the abundance coordinates, so v_i and v_j are represented by $(a_{\tau^{-1}(1),i}, \ldots, a_{\tau^{-1}(R),i})^t$ and $(a_{\tau^{-1}(1),j}, \ldots, a_{\tau^{-1}(R),j})^t$. Finally, we define the lexicographic abundance order $\leq_{\text{lex}}$ as

$$v_i \leq_{\text{lex}} v_j \Leftrightarrow \begin{cases} a_{\tau^{-1}(1),i} < a_{\tau^{-1}(1),j} \text{ or} \\ a_{\tau^{-1}(1),i} = a_{\tau^{-1}(1),j} \text{ and } a_{\tau^{-1}(2),i} < a_{\tau^{-1}(2),j} \text{ or} \\ \vdots \\ a_{\tau^{-1}(1),i} = a_{\tau^{-1}(1),j} \text{ and } \ldots \text{ and } a_{\tau^{-1}(R),i} \leq a_{\tau^{-1}(R),j} \end{cases} \tag{4}$$

Fig. 2(b) gives an example of first iso-order lines, where the soil is the dominant material (i.e., largest endmember in the sense of its norm). By first iso-order lines we mean those which correspond to the first condition in (4): $a_{\tau^{-1}(1),i} \leq a_{\tau^{-1}(1),j}$.

We note that this order has two important weaknesses. First, all the information of the hyperspectral images is not taking into account, since the first coordinate dominates all the others. Second, we need to order the materials. Nevertheless it is easy to use and to apply.

3.2 Foreground abundance order: use of majorization

We consider now a partial ordering based on the position of vectors in the simplex with respect to the so-called foreground. More precisely the foreground here corresponds to the set of endmembers.

The proposed approach is based on the notion of majorization [9], which is a technique for ordering vectors of same sum. Let us consider two vectors

$\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ and $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{R}^n$, then we say that C weakly majorizes D, written $\mathbf{c} \succ_w \mathbf{d}$, if and only if

$$\begin{cases} \sum_{i=1}^{k} c_i^{\downarrow} \geq \sum_{i=1}^{k} d_i^{\downarrow} \ , \forall k \in [1, n] \\ \sum_{i=1}^{n} c_i = \sum_{i=1}^{n} d_i \end{cases} \tag{5}$$

where $c_i^{\downarrow}$ and $d_i^{\downarrow}$ represent respectively the coordinates of C and D sorted in descending order. Majorization is not a partial order, since $\mathbf{c} \succ \mathbf{d}$ and $\mathbf{d} \succ \mathbf{c}$ do not imply $\mathbf{c} = \mathbf{d}$, it only implies that the components of each vector are equal, but not necessarily in the same order.

We propose a majorization-like partial order adapted to the abundances. Similarly to the majorization, a permutation τ_i of the coordinates of the vectors v_i in the simplex is applied such that they are sorted in descending order. The majorization-like order $\leq_{\text{maj}}$ is given defined as

$$v_i \leq_{\text{maj}} v_j \Leftrightarrow \begin{cases} a_{\tau_i^{-1}(1),i} < a_{\tau_j^{-1}(1),j} \text{ or} \\ a_{\tau_i^{-1}(1),i} = a_{\tau_j^{-1}(1),j} \text{ and } a_{\tau_i^{-1}(2),i} < a_{\tau_j^{-1}(2),j} \text{ or} \\ \vdots \\ a_{\tau_i^{-1}(1),i} = a_{\tau_j^{-1}(1),j} \text{ and } \ldots \text{ and } a_{\tau_i^{-1}(R),i} \leq a_{\tau_j^{-1}(R),j} \end{cases} \tag{6}$$

See in Fig. 2(c), the corresponding first iso-order lines. An important advantage of this order is the fact that there is no need for an order between the different materials. However, as one may notice, it is not possible to find a order-related metric that would follow the geometry of the simplex.

Partial order based on distances to reference vectors is classical in color morphology [1]. The partial order $\leq_{\text{maj}}$ is also related to the approach of hyperspectral order proposed in [2]. The fundamental difference is that the present order is based on abundances (barycentric coordinates in the $R-1$-simplex) whereas in [2], the partial order is based on distances in $\mathbb{R}^D$ between v_i and each endmember m_r (we note that computing distances in a high dimensional space is a tricky issue).

3.3 Foreground abundance order: use of stochastic dominance

The stochastic dominance is a partial ordering between probability distributions [16]. The term is used in decision theory to say if a distribution can be considered "superior" to another one. We first introduce the second order stochastic dominance [16]. Let us consider two random variables X and Y of respective cumulative distribution F_X and F_Y, then we say that X is second-order stochastically dominant over Y if and only if

$$\int_{-\infty}^{c} F_X(w)dw \leq \int_{-\infty}^{c} F_Y(w)dw \ , \forall c \in \mathbb{R} \tag{7}$$

In our case, the abundance coordinates $(a_{1,i}, \ldots, a_{R,i})^t$ of each vector v_i can be seen as a discrete probability distribution, since the sum of the coordinates is

equal to one. However, to be able to calculate the cumulative distribution of the vector, we need an order between the coordinates. We use the order of the endmembers and the corresponding permutation τ introduced for the lexicographic abundance order. So v_i is represented by $(a_{\tau^{-1}(1),i}, \ldots, a_{\tau^{-1}(R),i})^t$, then we calculate the discrete cumulative distribution from v_i, noted by $(b_{1,i}, \ldots, b_{R,i})^t$:

$$(a_{1,i}, \ldots, a_{R,i})^t \mapsto (b_{1,i}, \ldots, b_{R,i})^t, \text{ where } b_{j,i} = \sum_{k=1}^{j} a_{\tau^{-1}(k),i}.$$

Stochastic dominance gives rise to a partial order, which induces a supremum $\vee_{dom}$ between vectors v_i and v_j whose barycentric coordinates are $(a_{\tau(1),\vee}, \ldots, a_{\tau(R),\vee})^t$ such that

$$(a_{1,\vee}, \ldots, a_{R,\vee})^t \to v_1 \vee_{dom} v_2,$$

where

$$\begin{aligned} a_{1,\vee} &= b_{1,i} \vee b_{1,j}, \\ a_{2,\vee} &= b_{2,i} \vee b_{2,j} - b_{1,i} \vee b_{1,j}, \\ &\cdots \\ a_{R,\vee} &= b_{R,i} \vee b_{R,j} - b_{R-1,i} \vee b_{R-1,j}. \end{aligned}$$

Supremum $\vee_{dom}$ introduces "false abundances", in the sense that they are not present in the original set of vectors, however, the corresponding spectra are obtained from the endmembers and therefore the spectra are physically plausible. The infimum is obtained dually.

3.4 Background abundance order: use of distance/divergence

Previous orders focuss mainly on the relationship between the vectors in the simplex and its vertices, which correspond in a certain way to the foreground of the hyperspectral image (the pure pixels).

We propose now to think in a dual paradigm. Basically, the idea is to have an order given by how far the vectors are from the background. The background correspond to the situation where the material are totally mixed: the vector which has the same quantity of materials or in geometric terms, the center of the simplex $c \in \mathbb{R}^R$, i.e., in barycentric coordinates $c = (1/R, \ldots, 1/R)^t$. Then, the partial order between two vectors v_i and v_j will depend on the distance between c and each vector. We adopted the convention that further a point is from the center c, higher its position in the order is. Given a distance $d(\cdot, \cdot)$, we have the corresponding background partial order $\leq_{d_{L^2}}$:

$$v_i \leq_d v_j \Leftrightarrow d\left(c, (a_{1,i}, \ldots, a_{R,i})^t\right) \leq d\left(c, (a_{1,j}, \ldots, a_{R,j})^t\right). \tag{8}$$

In case of equality, we just use the lexicographic abundance order to have a total order, similarly to [1]. We have considered several distance metrics, for instance the L^1 and L^2 Minkowski norm (see in Fig. 2(d) the iso-order lines for $\leq_{d_{L^2}}$. We

have also considered the interest of the Mahalanobis distance $d_M(\cdot,\cdot)$, i.e., given two random vectors $\mathbf{c}$ and $\mathbf{d}$ following the same distribution with covariance matrix Σ, it is given by

$$d_M(\mathbf{c},\mathbf{d}) = \sqrt{(\mathbf{c}-\mathbf{d})^t \Sigma^{-1} (\mathbf{c}-\mathbf{d})}.$$

In our case, we have defined Σ as the empirical covariance matrix from the endmembers $\{m_r\}_{1\leq r\leq R}$. We represented the first iso-order lines for $\leq_{d_M}$ in Fig. 2(e). As one may notice, Mahalanobis distance involves elliptical lines whose directions of principal axes e_1 and e_2 are the two eigenvectors of Σ. Hence the geometry is deformed according to the correlations between the endmembers.

However, neither the L^p metric nor the Mahalanobis distance do not follow the intrinsic geometry of the simplex. For this purpose, we have studied the interest of the Kullback-Leibler divergence and its generalization, the Rényi divergence [3, 5]. Given two discrete probability measures P and Q, their similarity can be computed by the (non-symmetric) Kullback-Leibler divergence defined as

$$D_{\mathrm{KL}}(P\|Q) = \sum_i P(i)\, \log \frac{P(i)}{Q(i)}. \tag{9}$$

Rényi divergence of order q, $q > 0$ of a distribution P from a distribution Q is given

$$D_q(P\|Q) = \frac{1}{q-1} \log \sum_i P(i)^q Q(i)^{1-q}. \tag{10}$$

Parameter q allows the introduction to a new family of divergences, for instance $D_{q\to 1}(P\|Q) = D_{\mathrm{KL}}(P\|Q)$. The corresponding partial order is obtained as:

$$v_i \leq_{D_q} v_j \Leftrightarrow D_q\left(c,(a_{1,i},\ldots,a_{R,i})^t\right) \leq D_q\left(c,(a_{1,j},\ldots,a_{R,j})^t\right). \tag{11}$$

Fig. 3 illustrates the iso-order lines in the simplex for different values of q. We can notice in particular that the case $q = 1/2$ follows rather well the geometry of the simplex. Finally, we note that Rényi divergence and majorization are related [20].

4 Application to hyperspectral image processing

Given a hyperspectral image f, represented by its endmembers and the abundance maps:

$$f(x) = \sum_{r=1}^{R} \alpha_r(x) m_r,$$

the previous partial orders can be used to compute supremum and infimum needed for dilation and erosion of the abundance images. That is the vector order on the simplex is used on multivariate image $\{\alpha_r(x)\}_{1\leq r\leq R}$, to compute

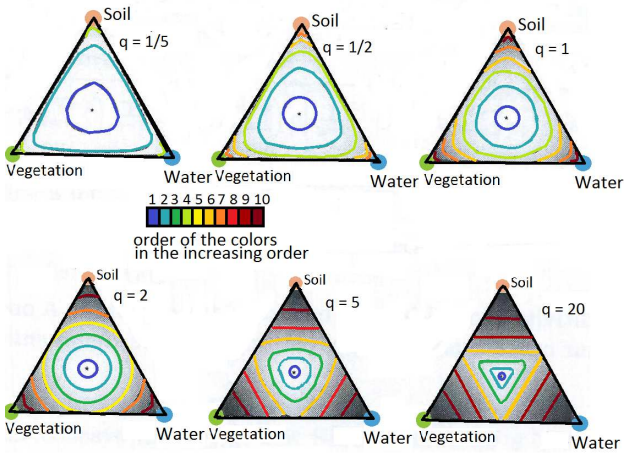

Fig. 3. Iso-order lines for partial order $\leq_{D_q}$ based on for Rényi divergence of parameter q.

for instance the operator $\phi : E \times \Delta^{R-1} \to E \times \Delta^{R-1}$. By abuse of notation, we write $\phi(\alpha_r)$ the r-abundance map obtained by operator ϕ. Then, the processed hyperspectral image by operator ϕ, noted $\phi(f)$ is obtained as

$$\phi(f)(x) = \sum_{r=1}^{R} \phi(\alpha_r)(x) m_r.$$

Fig. 4 gives a comparative example of the opening $\gamma_B(f)$ of a hyperspectral image f using various of the discussed orders.

To judge the effectiveness of a particular morphological operator is not easy. It depends mainly on the application. We have decided to evaluate the alternative orders for hyperspectral images in the context of image regularization and its interest to improve spectral classification. Basically, we compare the result of the supervised spectral classification obtained on the original hyperspectral image or the hyperspectral image filtered by a sequential filter according to one of the presented orders in the simplex.

This empirical study has been made on two images conventionally used in hyperspectral image processing domain: i) the Pavia image, which represents the campus of Pavia university (urban scene), of size 610×340 pixels and $D = 103$ spectral bands; ii) the Indian Pines image, test site in North-western Indiana composed for two thirds of agriculture, and one-third of forest, of 145×145 pixels and $D = 224$ spectral bands. We fixed the number of endmembers: $R = 9$ for Pavia image and $R = 10$ for the Indian Pines image. The endmembers are computed using N-FINDR [21] and the abondances by fully-constrained linear regression. Finally, for the different orders, a sequential filter $\gamma_B \varphi_B \gamma_B(f)$ is computed, where the structuring element B is a square size 3. The supervised clas-

sification is done using the least square SVM [22] algorithm with a RBF kernel as learning technique. Result of classification for both images are summarized in Tables 1 and 2.

Order	**Overall Accuracy with RBF kernel**	**Kappa statistic with RBF kernel**
without morpho. processing	60	0.57
lexicographic abundance order	61	0.59
majorization-like order	60	0.57
stochastic dominance-like order	71	0.69
Background order L^1	66	0.64
Background order L^2	61	0.59
Background order L^∞	58	0.55
Background order Mahalanobis distance,	64	0.62
Background order Rényi divergence $q = 0.2$**,**	68	0.66
Background order Rényi divergence $q = 1$**,**	69	0.67
Background order Rényi divergence $q^* = 4.65$**,**	74	0.72

Table 1. Comparison of result of classification on the Indian Pine hyperspectral image.

Order	**Overall Accuracy with RBF kernel**	**Kappa statistic with RBF kernel**
without morpho. processing	88	0.87
lexicographic abundance order	87	0.86
majorization like-order	88	0.86
stochastic dominance like-order	90	0.89
Background order L^1	84	0.82
Background order L^2	85	0.83
Background order L^∞	87	0.85
Background order Mahalanobis distance,	88	0.86
Background order Rényi divergence $q = 0.2$**,**	85	0.83
Background order Rényi divergence $q = 1$**,**	81	0.79
Background order Rényi divergence $q^* = 1.74$**,**	93	0.92

Table 2. Comparison of result of classification on the Pavia hyperspectral image.

We do not claim that this is the best way to improve spectral classification, however this study highlights the impact of the order used in the morphological operators, which involves in certain case on the choice of the metric. We observe that, by optimizing the parameter q of Rényi divergence, noted q^*, it is possible to significantly improve the classification score.

5 Conclusions

We have proposed different kinds of partial orders based on the an endmember representation of the hyperspectral images. They are just useful to compute morphological operators for images represented by the linear mixing model. From

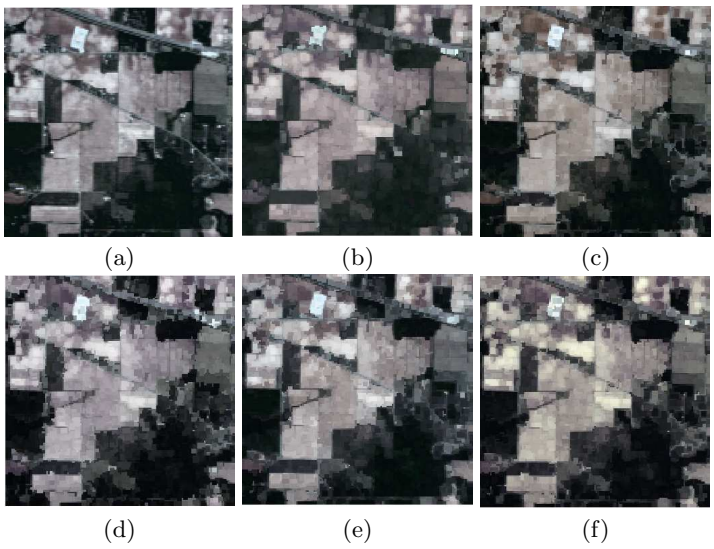

Fig. 4. (a) False RGB color image (using three spectral bands) of a hyperspectral image Indian Pines f, (b) opening $\gamma_B(f)$ using lexicographic abundance order, (c) opening $\gamma_B(f)$ using majorization-like order, (d) opening $\gamma_B(f)$ using stochastic dominance-like order, (e) opening $\gamma_B(f)$ using Mahalanobis distance-based order, (f) opening $\gamma_B(f)$ using L^1-based order, (g) opening $\gamma_B(f)$ using Rényi divergence-based order, $q = 1$.

a mathematical morphology viewpoint, orders considered here can be used for other data lying in the simplex. This is the case for instance of images where at each pixel a discrete probability distribution is given.

In the experimental section, we have illustrate the potential interest of corresponding morphological operators for a spatial regularization before the classification. However, any image processing task tackled with morphological operators (scale-space decomposition, image enhancement, etc.) can be extended to hyperspectral images using the present framework.

References

1. Angulo, J. (2007). Morphological colour operators in totally ordered lattices based on distances: Application to image filtering, enhancement and analysis. *Computer Vision and Image Understanding,* 107(1):56–73.
2. Aptoula, E., Courty, N., & Lefevre, S. (2014). An end-member based ordering relation for the morphological description of hyperspectral images. In *Proc. of the IEEE ICIP'2014*.
3. Bengtsson, I., & Zyczkowski, K. (2006). *Geometry of quantum states: an introduction to quantum entanglement,* Cambridge University Press.
4. Bellman, R.E. (1961). *Adaptive control processes,* Princeton University Press, New Jersey, 1961.

5. Bromiley, P.A., Thacker, N.A., & Bouhova-Thacker, E. (2004). Shannon entropy, Renyi entropy, and information. *Statistics and Inf. Series (2004-004).*
6. Chang, C.I. (1998). Further results on relationship between spectral unmixing and subspace projection. *IEEE Trans. on Geoscience and Remote Sensing,* 36(3):1030–1032.
7. Chang, C.I. (Ed.). (2007). *Hyperspectral data exploitation: theory and applications,* John Wiley & Sons.
8. Keshava, N. (2003). A survey of spectral unmixing algorithms. *Lincoln Laboratory Journal,* 14(1): 55-78.
9. Marshall, A.W., Olkin, I., & Arnold, B.C. (2010). *Inequalities: Theory of Majorization and Its Applications: Theory of Majorization and Its Applications,* Springer.
10. Mohan, A., Sapiro, G. & Bosch, E. (2007). Spatially coherent nonlinear dimensionality reduction and segmentation of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters,* 4(2): 206–210.
11. Donoho, D., & Stodden, V. (2003). When does non-negative matrix factorization give a correct decomposition into parts?. In *Proc. of Advances in neural information processing systems 16,* MIT Press.
12. Fauvel, M., Benediktsson, J.A., Chanussot, J., & Sveinsson, J.R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. IEEE Trans. on Geoscience and Remote Sensing, 46(11):3804–3814.
13. Fauvel, M., Tarabalka, Y., Benediktsson, J.A., Chanussot, J., & Tilton, J.C. (2013). Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE,* 101(3), 652–675.
14. Altmann, Y. (2013) *Nonlinear unmixing of hyperspectral images,* Ph.D. thesis manuscript, Université de Toulouse (France), HAL, tel-00945513.
15. Serra, J. (1993). Anamorphoses and Function Lattices (Multivalued morphology). Chapter 13, in *(Dougherty, E., Ed.) Mathematical Morphology in Image Processing,* Marcel Dekker.
16. Shaked, M., & Shanthikumar, J.G. (1994). *Stochastic Orders and their Applications,* Associated Press.
17. Velasco-Forero, S., & Angulo, J. (2011). Supervised ordering in $\mathbb{R}^n$: Application to morphological processing of hyperspectral images. *IEEE Transactions on Image Processing,* 20(11): 3301–3308.
18. Velasco-Forero, S., & Angulo, J. (2012). Random projection depth for multivariate mathematical morphology. *IEEE Journal of Selected Topics in Signal Processing*, 6(7):753–763.
19. Velasco-Forero, S., & Angulo, J. (2013). Classification of hyperspectral images by tensor modeling and additive morphological decomposition. *Pattern Recognition*, 46(2):566–577.
20. van Erven, T. & Harremoës, P. (2010) Rényi divergence and majorization. In Proc of the *IEEE International Symposium on Information Theory (ISIT'10).*
21. Winter, M.E. (1999). N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data. In *Proc. of SPIE Image Spectrometry V,* SPIE Vol. 3753, 266–277.
22. Camps-Valls, G., & Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. Geoscience and Remote Sensing, IEEE Transactions on, 43(6), 1351-1362.