Kernel Fully Constrained Least Squares
Abundance Estimates

1,2Joshua Broadwater and 1Rama Chellappa
The University of Maryland
Dept. of Electrical and Computer Engineering
Center for Automation Research
College Park, MD USA
Joshua.Broadwater@jhuapl.edu
rama@cfar.umd.edu

2Amit Banerjee and 2Philippe Burlina
The Johns Hopkins University
Applied Physics Laboratory
Laurel, MD USA
Amit.Banerjee@jhuapl.edu
Phil.Burlina@jhuapl.edu

Abstract—A critical step for fitting a linear mixing model to hyperspectral imagery is the estimation of the abundances. The abundances are the percentage of each endmember within a given pixel; therefore, they should be non-negative and sum to one. With the advent of kernel based algorithms for hyperspectral imagery, kernel based abundance estimates have become necessary. This paper presents such an algorithm that estimates the abundances in the kernel feature space while maintaining the non-negativity and sum-to-one constraints. The usefulness of the algorithm is shown using the AVIRIS Cuprite, Nevada image.

Keywords—kernel functions, hyperspectral imagery, spectral unmixing, abundance estimates

I. INTRODUCTION

Hyperspectral image (HSI) data is typically modeled as a linear combination of endmembers and abundances. Endmembers are the spectra representing the unique materials in a given image such as water, soil, and vegetation. Abundances are the percentage of each endmember within a given pixel. This linear mixing model can be expressed as

\[ \mathbf{x} = \mathbf{Ea}, \quad a_i \geq 0, \quad \sum_{i=1}^{M} a_i = 1 \]  

where \( \mathbf{x} \) is an \( L \times 1 \) vector that represents the spectral signature of the current pixel, \( M \) is the number of endmembers within the image, \( \mathbf{E} \) is an \( L \times M \) matrix where each column represents the \( i^{th} \) endmember, and \( \mathbf{a} \) is an \( M \times 1 \) vector where the \( i^{th} \) entry represents the abundance value \( a_i \). From this linear mixing model, a number of hyperspectral analysis algorithms have been developed for classification and detection applications.

Recently, new algorithms based on kernel methods have been introduced in the literature [1][2]. These algorithms generalize the linear mixing model by introducing non-linearities through kernel functions. Kernel functions assume a function \( \Phi(\mathbf{x}) \) exists that maps the HSI data to a high-dimensional space called the feature space. The trick to the method is that the function \( \Phi(\mathbf{x}) \) does not have to be known. Instead, inner products are replaced with kernel functions that implicitly calculate \( \Phi(\mathbf{x}) \).

The drawback with current kernel algorithms for classification and detection is that they 1) use non-physical endmembers and/or 2) produce abundance estimates that do not meet the non-negativity and sum-to-one constraints. In [3], we address the first issue by introducing the support vector data description (SVDD) algorithm to extract physical endmembers in the feature space. To address the second issue, we develop a kernel based abundance estimate to meet the physical abundance constraints.

The SVDD algorithm generates endmembers that do not necessarily form a convex boundary in the original input space. This is a desirable property because the endmembers are allowed to model the data without imposing any assumptions about its underlying distribution. Current algorithms such as [4][5][6][7] cannot be used in this case as they will ignore those endmembers not on the convex boundary.

To overcome this problem and estimate abundance estimates in the correct feature space, we propose a kernel fully constrained least squares (KFCLS) algorithm. This algorithm generalizes previous abundance algorithms by using a kernel function. If we choose a linear kernel function (i.e., an inner product), the KFCLS algorithm is simply the FCLS algorithm [7]. However, if we choose a more complex kernel like a radial basis function, the KFCLS algorithm can calculate abundances for those endmembers that form a non-convex boundary.

This paper is organized as follows. Section II provides a derivation of the KFCLS algorithm. Section III provides experimental results using the well documented AVIRIS Cuprite, Nevada imagery. Section IV presents a summary of our work and future directions of research.

II. KFCLS DERIVATION

The KFCLS derivation begins with the non-negativity constraints. As with the other algorithms, we begin by minimizing the Least Squares Error (LSE) to estimate the non-negative abundance values. Mathematically this is expressed as

\[ \min_{\mathbf{a}} (\mathbf{x} - \mathbf{Ea})^T (\mathbf{x} - \mathbf{Ea}), \ s.t. \ a_i \geq 0 \ \forall i \]  

\[ (2) \]
where \( E \) is a matrix of \( n \) background endmembers. Knowing that the abundances have to be non-negative, we write a Lagrangian function \( J \) such that

\[
J = \frac{1}{2} (\langle x, x \rangle - 2 \sum_i a_i < x, e_i > ) + \sum_{i,j} a_i a_j < e_i, e_j > + \sum_i \lambda_i (a_i - c_i)
\]

where \( e_i \) is the \( i \)th column from \( E \), \( a_i = c_i \), and each \( c_i \) is non-negative to enforce the non-negativity constraint. By replacing the inner products with their kernel counterparts, the new Lagrangian is

\[
J = \frac{1}{2} (K(x, x) - 2 \sum a_i K(x, e_i) + \sum_{i,j} a_i a_j K(e_i, e_j) + \sum_i \lambda_i (a_i - c_i)
\]

To calculate the estimate of \( a \), we take the partial derivative of \( J \) with respect to each \( a_i \) to obtain

\[
\frac{\partial J}{\partial a_i} = -K(x, e_i) + \sum_j a_j K(e_j, e_i) + \lambda_i.
\]

The equation above contains a number of unknowns for each \( a_i \) and each \( \lambda_i \). We can solve for each of these unknowns to produce the following equations.

\[
\lambda_i = \sum_j a_j K(e_j, e_j) - K(x, e_i)
\]

\[
a_i = K(e_i, e_i)^{-1} \left[ \sum_j a_j K(e_j, e_j) - K(x, e_i) + \lambda_i \right]
\]

To solve for the non-negative abundances, we set all the Lagrange multipliers to zero and calculate the abundance using (7). From this solution, we identify those abundance values that are greater than zero and place them in the passive set \( P \). The remaining non-positive abundance values are placed in the active set \( R \). Equations (6) and (7) are iterated until all Lagrange multipliers in the active set are zero and all Lagrange multipliers in the passive set are either zero or negative. At this point, the Kuhn-Tucker conditions have been met and an optimal solution for the non-negative abundance values has been found.

To handle the sum-to-one constraints, an easy modification of the aforementioned algorithm was developed to retain the optimality guaranteed under the Kuhn-Tucker conditions for numerical optimization on a finite computing machine [8]. In the modification, the endmember matrix and pixel signatures are extended such that

\[
\tilde{E} = \begin{bmatrix} \delta E \\ 1^T \end{bmatrix}
\]

is the new endmember matrix and

\[
\tilde{x} = \begin{bmatrix} \delta x \\ 1 \end{bmatrix}
\]

is the new pixel signature where \( \delta \) is a small number (typically \( 1 \times 10^{-5} \)). The \( \delta \) variable controls how tightly the solution will sum to one so that smaller values provide a better solution, but may need longer convergence time. The new endmember matrix and pixel signature are then used in (6) and (7) to obtain an abundance solution that meets both the non-negativity and sum-to-one constraints.

The original purpose for creating the KFCLS algorithm was to guarantee the abundance estimates are being calculated in the same feature space as the endmembers. This is a necessary condition because different kernel functions and even different kernel parameters generate different feature spaces. Therefore, the endmembers and abundances need to be calculated in the same feature space otherwise we may obtain results that are erroneous. Because of this reason, the KFCLS algorithm uses the same kernel and kernel parameters as the SVDD algorithm.

III. EXPERIMENTAL RESULTS

A. Data Description

The data used for our analysis comes from the Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) developed at the California Institute of Technology Jet Propulsion Laboratory (JPL) [9]. The AVIRIS sensor is a hyperspectral sensor that collects 224 contiguous spectral bands from 400 to 2500 nm. The ground sampling distance is typically 17 m per side. Imagery collected by AVIRIS can be received in the original raw radiance units or corrected reflectance units. In this analysis, we chose the reflectance units so we could compare the endmembers and abundances calculated with known laboratory spectra.

The AVIRIS image we used for our analyses is the well documented Cuprite, Nevada image set. This imagery has been studied in previous papers and has documented ground truth information. The Cuprite image is broken into five parts for easier processing. We chose the fourth image of the set which is used by most other researchers. As was done in [10][11], we also limited the spectrum to the short-wave infrared bands from 2.0 to 2.5 microns. This gave us a final image size of 512x614x57. Fig. 1 shows the broadband short-wave infrared Cuprite AVIRIS image.
B. Endmember Extraction

As mentioned in the Section I, the KFCLS algorithm is designed to work with endmember extraction algorithms that use kernel methods. Therefore, we used our SVDD technique described in [3] to extract the endmembers. The SVDD technique provides two critical other parameters besides the endmembers. First, the SVDD algorithm provides the kernel function that should be used by the KFCLS algorithm. For this experiment, the kernel function is radial basis function described by

\[ K(x, y) = \exp(-\|x - y\|^2 / \sigma^2) \]  

Second, the SVDD algorithm provides the values for the kernel parameter \( \sigma^2 \). The kernel parameter controls the number of endmembers extracted and is found using a rate-distortion algorithm that automatically chooses the “correct” number of endmembers for a scene. For our experiments, the SVDD algorithm found 21 endmembers using the RBF kernel function with \( \sigma^2 = 115 \). Figure 2 shows the endmembers found for the Cuprite image using the SVDD algorithm.

C. Abundance Results

Using the endmembers and kernel parameters from [3], we calculated 21 abundance maps using the KFCLS algorithm. Abundance maps are images of the abundance values shown across all pixels in the image. The images are mapped such that dark pixel represent abundances of zero while white pixels represent abundances of one.

Three abundance maps are shown in Figures 3 through 5 for the minerals alunite, calcite, and kaolinite. We chose these three mineral endmembers because they are well studied and have been mapped in previous research [10][11][12]. Comparing our results to the previous research, the KFCLS abundance maps match exceptionally well. This result shows the ability of the KFCLS algorithm to generate physically meaningful abundances even when calculated in feature space.
D. Timing Results

A valid concern with an algorithm like KFCLS is the time required to estimate the abundances. KFCLS is basically a constrained quadratic programming problem that has to be implemented for each pixel in the image. The algorithm we present in Section II however is optimized to reduce processing time. For the Cuprite image, KFCLS ran in 179 seconds in the Matlab® R2006a environment on a Windows XP laptop with an 800 MHz processor and 1 GByte of memory. Thus, a pixel was being processed every $5.7 \cdot 10^{-4}$ seconds. Given the age of the computer and the fact the processing was done within Matlab using unoptimized code, the time to estimate abundances using KFCLS is very good.

IV. SUMMARY

We present a new kernel based abundance estimate for hyperspectral data analysis. The KFCLS algorithm generalizes previous approaches by the use of a kernel function. A simple linear kernel function allows KFCLS to estimate the same values as the FCLS algorithm. By using more complex kernels such as radial basis functions, the KFCLS algorithm can also find abundances for endmembers that do not lie on convex boundaries. The results show the abundances estimated for such a case are still physically meaningful. This type of algorithm allows the expansion of current kernel based algorithms to include physically meaningful estimates and does so in the same feature space.

REFERENCES