Detection of nonlinear mixtures using Gaussian processes: Application to hyperspectral imaging

Tales Imbiriba, José Carlos Bermudez, Jean-Yves Tourneret, Cédric Richard

DETECTION OF NONLINEAR MIXTURES USING GAUSSIAN PROCESSES: APPLICATION TO HYPERSPECTRAL IMAGING

T. Imbiriba[⋆] *J. C. M. Bermudez*[⋆] *J.-Y. Tourneret*[†] *C. Richard*[‡]

[⋆] Federal University of Santa Catarina, Florianópolis, SC, Brazil
[†]University of Toulouse, IRIT-ENSEEIHT, CNRS, Toulouse, France
[‡]Université de Nice Sophia-Antipolis, CNRS, Nice, France

ABSTRACT

This paper investigates the use of Gaussian processes to detect nonlinearly mixed pixels in hyperspectral images. The proposed technique is independent of nonlinear mixing mechanism, and therefore is not restricted to any prescribed nonlinear mixing model. The observed reflectances are estimated using both the least squares method and a Gaussian process. The fitting errors of the two approaches are combined in a test statistics for which it is possible to estimate a detection threshold given a required probability of false alarm. The proposed detector is compared to a robust nonlinearity detector recently proposed using synthetic data and is shown to provide a better detection performance. The new detector is also tested on a real hyperspectral image.

Index Terms— Nonlinearity detection, Hyperspectral images, Gaussian processes.

1. INTRODUCTION

The analysis of hyperspectral images has been recognized as an important tool to infer about the materials present in a scene and about their relative contribution to the scene [1–3]. Such analysis aims at unmixing the spectral information present in the hyperspectral image to identify the composing materials (endmembers) and their abundances in the region from which the data has been acquired. Most unmixing techniques rely on a parametric mixing model, from which the parameters must be estimated [4]. The simplest of these models assumes linear mixing of the endmember contributions [3] (Linear Mixing Model - LMM). However, it has been recognized that the mixing in some pixels of a region is actually nonlinear [3–5]. This finding has triggered a plethora of techniques for analyzing nonlinearly mixed pixels (see for instance [4–15] and references therein). Though nonlinear unmixing permits a better understanding of the endmember contributions, the corresponding analysis techniques are necessarily more complex than linear unmixing. Hence, it makes sense to detect the nonlinearly mixed pixels in an image prior to the analysis. Doing that allows the utilization of the simplest possible unmixing technique to analyze each pixel.

A possible approach for detecting nonlinearly mixed pixels assumes a parametric model for the nonlinearity. The parameters controlling the nonlinearity are then estimated and hypothesis tests are developed based on such estimator. For instance, a single parameter polynomial post-nonlinear mixing model (PPNMM) is assumed in [16]. The main question about a parametric modeling of the nonlinear mixing is whether the chosen model is capable of capturing the actual nonlinear effects present in the hyperspectral image. This is a difficult question to answer, as the actual nonlinear mixing taking place in hyperspectral images is usually unknown. When little or nothing is known about the nonlinear mixing mechanism, an interesting approach is to use nonparametric techniques to obtain information about the nonlinearity directly from the observed data. Nonparametric nonlinearity modeling is not new. For instance, Chen *et al.* [12] introduced a nonlinear unmixing technique using kernel-based expansions. However, the work in [12] is not concerned about nonlinearity detection. Recently, Altmann *et al.* [17] proposed a robust nonlinear mixture detector that does not use a parametric model for the nonlinear mixture. The detector is based on the fact that linear mixing confines the noiseless data to a low-dimensional hyperplane. The hypothesis test thus uses the distance between the observed pixel and that hyperplane. The alternative hypothesis is characterized by an extra deterministic contribution to the mean value of the observations. Though the test proposed in [17] is robust to the actual nonlinear mixing mechanism, it conveys too little information about the nonlinearity as a tradeoff to guarantee simplicity.

In this paper we propose a new robust nonlinearity mixing test that captures more information about the nonlinearity. We propose to model the contribution of the endmembers to the observations using a Gaussian process (GP). The nonlinearity is estimated from the GP and compared to the linear least squares (LS) LMM estimator for the same data. A hypothesis test is then proposed based on the linear and nonlinear estimation errors. As in [17] we assume that the endmembers are available or have been estimated by an appropriate endmember extraction algorithm.

2. MIXING MODELS

The LMM [3] is described as

$$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{a} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{y}$ is the L-dimensional observed pixel, L being the number of spectral bands, $\boldsymbol{M}$ is the $L \times R$ endmember matrix, R is the number of endmembers, $\boldsymbol{a}$ is the R-dimensional abundance vector, and $\boldsymbol{n}$ is the additive noise assumed to be Gaussian with zero mean and covariance matrix $\sigma_n^2\boldsymbol{I}$, that is, $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}_L, \sigma_n^2\boldsymbol{I})$, where $\boldsymbol{I}$ is the identity matrix. The abundances must also obey the following constraints

$$\sum_{r=1}^{R} a_r = 1, \qquad a_r \geq 0, \quad \forall r \in \{1, \ldots, R\}. \tag{2}$$

This work was partly supported by CNPq under grants Nos 305377/2009-4, 400566/2013-3 and 141094/2012-5, and by the Agence Nationale pour la Recherche, France, (Hypanema project, ANR-12- BS03-003), and by ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02.

For better understanding of (4) below, it is important to notice that, under the LMM (1), the ith component y_i of $\boldsymbol{y}$ is given by $y_i = \boldsymbol{a}^\top \boldsymbol{m}_i + n_i$ where $\boldsymbol{m}_i$ is the transpose of the ith row of $\boldsymbol{M}$. In the case of a general mixing, we represent the observation vector $\boldsymbol{y}$ as

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{M}) + \boldsymbol{n}. \tag{3}$$

We suggest to model $\boldsymbol{g}(\boldsymbol{M})$ as a realization of a GP that describes a distribution over functions.

3. GP REGRESSION

By analogy with the LMM, which writes $y_i = \boldsymbol{a}^\top \boldsymbol{m}_i + n_i$, consider modeling the ith row of (3) in a nonlinear way as

$$y_i = f(\boldsymbol{m}_i) + n_i, \tag{4}$$

where $n_i \sim \mathcal{N}(0, \sigma_n^2)$ and $f(\cdot)$ is a smooth latent function. In the context of GPs, we define a Gaussian prior for $f(\cdot)$ with zero mean[1] and covariance function $k(\cdot,\cdot)$. Following [18] and considering the training set $\{\boldsymbol{y}, \boldsymbol{X}\}$, with inputs $\boldsymbol{X} = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_L]$, and outputs or observations $\boldsymbol{y} = [y_1, \ldots, y_L]^\top$, the GP prior distribution for $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right), \tag{5}$$

with $\boldsymbol{K}$ the Gram matrix whose entries $\boldsymbol{K}_{ij} = k(\boldsymbol{m}_i, \boldsymbol{m}_j)$ are the kernel (covariance) functions [19] of the inputs $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ (rows of $\boldsymbol{M}$), and $\boldsymbol{I}$ is the $L \times L$ identity matrix.
GP regression aims at inferring the latent function distribution of f_* for a new (or test) input $\boldsymbol{m}_*$. Using the *marginalization property* [18], (5) can be rewritten as

$$\begin{bmatrix} \boldsymbol{y} \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{K} + \sigma_n^2 \boldsymbol{I} & \boldsymbol{k}_* \\ \boldsymbol{k}_*^\top & k_{**} \end{bmatrix}\right) \tag{6}$$

with $\boldsymbol{k}_*^\top = [k(\boldsymbol{m}_*, \boldsymbol{m}_1), \ldots, k(\boldsymbol{m}_*, \boldsymbol{m}_L)]$ and $k_{**} = k(\boldsymbol{m}_*, \boldsymbol{m}_*)$. The predictive distribution of f_*, or posterior of f_*, can be obtained by conditioning (6) on the data as

$$\begin{aligned} f_* | \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{m}_* \sim \mathcal{N}\Big(& \boldsymbol{k}_*^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{y}\,, \\ & k_{**} - \boldsymbol{k}_*^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{k}_* \Big). \end{aligned} \tag{7}$$

The extension to a multivariate predictive distribution with test data $\boldsymbol{X}_* = [\boldsymbol{m}_{*1}, \ldots, \boldsymbol{m}_{*L}]$ is straightforward and yields

$$\begin{aligned} \boldsymbol{f}_* | \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{X}_* \sim \mathcal{N}\Big(& \boldsymbol{K}_*^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{y}\,, \\ & \boldsymbol{K}_{**} - \boldsymbol{K}_*^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{K}_* \Big) \end{aligned} \tag{8}$$

where $[\boldsymbol{K}_*]_{ij} = k(\boldsymbol{m}_{*i}, \boldsymbol{m}_j)$ and $[\boldsymbol{K}_{**}]_{ij} = k(\boldsymbol{m}_{*i}, \boldsymbol{m}_{*j})$.

Different kernels can be used in (8) [18]. Here we use the Gaussian kernel

$$k(\boldsymbol{m}_p, \boldsymbol{m}_q) = \sigma_f^2 \exp\left\{-\frac{1}{2s^2}\|\boldsymbol{m}_p - \boldsymbol{m}_q\|^2\right\} \tag{9}$$

for its smoothness and non-informativeness, as we lack any knowledge about the unknown function $f(\cdot)$. Hence, the function estimation is done in a *reproducing kernel Hilbert space* (RKHS) with universal approximating capability [20, p. 35].

[1]The zero mean can be considered even for hyperspectral signatures since we can first subtract the pixel by its mean.

We estimate the noise variance and the kernel hyperparameters in $\boldsymbol{\theta} = \{\sigma_f^2, s^2, \sigma_n^2\}$ by maximizing the marginal likelihood function $p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta})$. Hence,

$$\hat{\boldsymbol{\theta}} = \arg_{\boldsymbol{\theta}} \max \log p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}) \tag{10}$$

where

$$\begin{aligned} \log p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}) = & -\frac{1}{2}\boldsymbol{y}^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{y} - \frac{1}{2}\log|\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}| \\ & - \frac{L}{2}\log(2\pi). \end{aligned}$$

Using the minimum mean squared error (MMSE) criterion, the predictor $\hat{\boldsymbol{y}}_g$ of $\boldsymbol{f}$ is defined as the mean of the predictive distribution in (8). Hence,

$$\hat{\boldsymbol{y}}_g = \widehat{\boldsymbol{f}}_*^{\mathrm{MMSE}} = \boldsymbol{K}_*^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1} \boldsymbol{y}. \tag{11}$$

4. NONLINEAR MIXTURE DETECTOR

Given an observation vector $\boldsymbol{y}$, we formulate the nonlinear mixture detector as the following binary hypothesis test problem

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{M}\boldsymbol{a} + \boldsymbol{n} \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{M}) + \boldsymbol{n} \end{cases} \tag{12}$$

where we assume that the endmember matrix $\boldsymbol{M}$ is available or has been estimated from the image using an endmember extraction technique [5].

We propose to compare the fitting errors resulting from estimating $\boldsymbol{y}$ using an LS estimator and the GP-based estimator (11). Under hypothesis $\mathcal{H}_0$, both the LS and the GP-based estimators should provide good estimates, while under $\mathcal{H}_1$ the LS estimation error should be significantly larger than that resulting from the GP-based estimation. Next, we describe the two estimation errors.

4.1. LS fitting error

The LS estimation error is given by

$$\boldsymbol{e}_\ell = \boldsymbol{y} - \hat{\boldsymbol{y}}_\ell \tag{13}$$

where $\hat{\boldsymbol{y}}_\ell = \boldsymbol{M}\hat{\boldsymbol{a}}$ is the LS estimator of $\boldsymbol{f}$, with

$$\hat{\boldsymbol{a}} = (\boldsymbol{M}^\top \boldsymbol{M})^{-1} \boldsymbol{M}^\top \boldsymbol{y}. \tag{14}$$

Then, simple calculation yields

$$\boldsymbol{e}_\ell = \boldsymbol{P}\boldsymbol{y} \tag{15}$$

where $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{M}(\boldsymbol{M}^\top \boldsymbol{M})^{-1}\boldsymbol{M}^\top$ is an $L \times L$ projection matrix of rank $\rho = L - R$.

4.2. GPM fitting error

The GP-based estimation error is given by

$$\boldsymbol{e}_g = \boldsymbol{y} - \hat{\boldsymbol{y}}_g \tag{16}$$

where $\hat{\boldsymbol{y}}_g$ is determined using (11) with $\boldsymbol{X}_* = \boldsymbol{X}$. This is because our interest is to evaluate the fitting between the model and the available data, and not to make predictions for new data. Hence, the fitting error from (11) becomes

$$\boldsymbol{e}_g = \boldsymbol{y} - \widehat{\boldsymbol{f}}_*^{\mathrm{MMSE}}\Big|_{\boldsymbol{X}_* = \boldsymbol{X}} = \boldsymbol{H}\boldsymbol{y} \tag{17}$$

where $\boldsymbol{H} = \boldsymbol{I}_L - \boldsymbol{K}^\top \left[\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}\right]^{-1}$ is a real-valued symmetric matrix of rank L.

4.3. The test statistics

To decide between the two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ we propose to compare the squared norms of the two fitting error vectors. In doing that, we also need a test statistics whose distribution is known or at least can be approximated, so that a test threshold can be adjusted from a given probability of false alarm (PFA) and the detector can be designed. Given these objectives, we propose the test

$$T = \frac{2\|\boldsymbol{e}_g\|^2}{\|\boldsymbol{e}_g\|^2 + \|\boldsymbol{e}_\ell\|^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\lessgtr}} \tau, \tag{18}$$

where τ is the detection threshold. The reasoning behind the choice of T is as follows. First, as $\boldsymbol{e}_\ell$ and $\boldsymbol{e}_g$ have zero-mean Gaussian distributions, both $\|\boldsymbol{e}_g\|^2$ and $\|\boldsymbol{e}_\ell\|^2$ are chi-square random variables. Now, we write $\boldsymbol{e}_\ell$ as $\boldsymbol{e}_g + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is assumed to be also Gaussian and neglect the cross-term $2\boldsymbol{e}_g^\top \boldsymbol{\epsilon}$, compared to $\|\boldsymbol{\epsilon}\|^2$, when evaluating $\|\boldsymbol{e}_\ell\|^2$ under $\mathcal{H}_0$. The latter approximation is due to the lack of correlation between $\boldsymbol{e}_g$ and $\boldsymbol{\epsilon}$, which can be largely attributed to mismatches resulting from the numerical optimization required to solve (10). Under these considerations and defining $Z = \|\boldsymbol{e}_g\|^2$, (18) can be written as $T = 2Z/(Z + \|\boldsymbol{\epsilon}\|^2)$ with both Z and $\|\boldsymbol{\epsilon}\|^2$ independent and chi-square distributed. Such a statistics is known to follow a Beta distribution [21].

As the GP-based estimator tends to fit better a nonlinearly mixed data, T should be less than 1 under hypothesis $\mathcal{H}_1$. Conversely, T should be close to one for linearly mixed pixels, as $\|\boldsymbol{\epsilon}\|^2$ tends to be much less than $2\|\boldsymbol{e}_g\|^2$. Hence, as per (18), we accept $\mathcal{H}_0$ if $T > \tau$ and we conclude for the nonlinear mixing (of $\mathcal{H}_1$) if $T < \tau$.

5. EXPERIMENTS

This section presents experiments using synthetic and real data.

5.1. Synthetic Data

To test the performance of the detection method proposed in the previous section, we generated synthetic data that contain both linearly and nonlinearly mixed pixels. The amount of nonlinearity is characterized by a degree of nonlinearity. The linearly mixed pixels were generated using the LMM (2) with a known matrix $\boldsymbol{M}$. The nonlinearly mixed pixels were generated using the simplified generalized bilinear model (GBM) used in [17], with a new scaling that permits the control of the degree of nonlinearity for each nonlinear pixel generated. More precisely, the nonlinearly mixed pixels were generated using the following model

$$\boldsymbol{y} = \kappa \boldsymbol{M}\boldsymbol{a} + \boldsymbol{\mu} + \boldsymbol{n} \tag{19}$$

where $0 \leq \kappa \leq 1$, $\boldsymbol{\mu} = \gamma \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} a_i a_j \boldsymbol{m}_i \odot \boldsymbol{m}_j$ is the nonlinear term, γ is the parameter that governs the amount of nonlinear contribution, and $\odot$ is the Hadamard product. Given the parameters $\boldsymbol{M}$, $\boldsymbol{a}$, γ and σ_n^2, this model generates samples with same energy and SNR as the LMM if

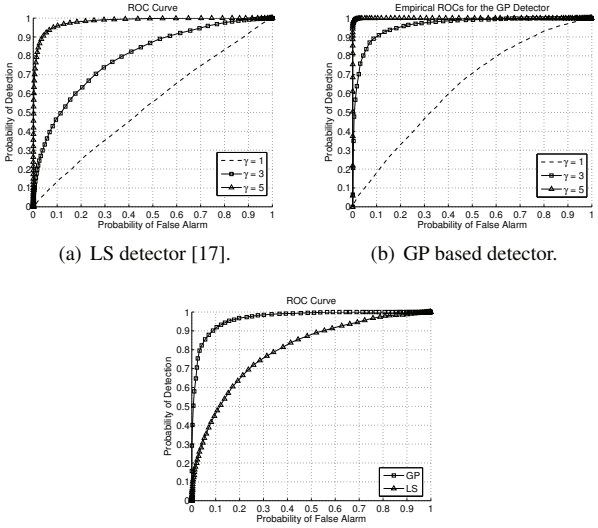
$$\kappa = \left[-2E_{\ell\mu} + \sqrt{4E_{\ell\mu}^2 - 4E_\ell(E_\mu - E_\ell)}\right] / 2E_\ell \tag{20}$$

where $E_\ell = \|\boldsymbol{y}_\ell\|^2$ is the energy of a noiseless linear pixel (i.e., $\boldsymbol{a}^\top \boldsymbol{M}^\top \boldsymbol{M}\boldsymbol{a}$), $E_{\ell\mu} = \boldsymbol{y}_\ell^\top \boldsymbol{\mu}$ is the "cross-energy" of the linear and nonlinear parts, and $E_\mu = \|\boldsymbol{\mu}\|^2$ is the energy of the nonlinear contribution. The degree of nonlinearity of a pixel is then defined as the ratio of the nonlinear portion to the total pixel energy

$$\eta_d = \frac{2\kappa E_{\ell\mu} + E_\mu}{\kappa^2 E_\ell + 2\kappa E_{\ell\mu} + E_\mu}. \tag{21}$$

For the simulations presented here, the endmember matrix $\boldsymbol{M}$ was composed of $R = 3$ materials (green grass, olive green paint and galvanized steel metal) extracted from the spectral library of the software ENVI$^{\text{TM}}$ [22]. Each endmember $\boldsymbol{m}_r$ has $L = 826$ bands that were uniformly decimated to $L = 83$ bands for simplicity. The abundance vector $\boldsymbol{a} = [0.3,\ 0.6,\ 0,\ 1]^\top$ was arbitrarily fixed, and $\sigma_n^2 = 0.0011$ was chosen to produce an SNR of 21dB for both linear and nonlinear samples.

Figure 1 presents the empirical Receiver Operating Characteristic (ROC) curves for both the LS-based detector presented in [17] and the GP detector (18) for $\gamma = 1$ ($\eta_d = 0.21$), $\gamma = 3$ ($\eta_d = 0.55$) and $\gamma = 5$ ($\eta_d = 0.80$). It can be verified that the GP detector presents an improved performance in all three cases. As an example, for $\gamma = 3$ and $\text{PFA} = 0.1$ the LS detector has a probability of detection (PD) in the order of 0.45, while the GP detector has $\text{PD} \approx 0.9$. These results indicate that the extra computational complexity required by the GP detector is justified for detecting nonlinearly mixed pixels independently of the nonlinear mixing model.



(a) LS detector [17]. (b) GP based detector.

(c) GP vs. LS detectors.

Fig. 1: (a) Empirical ROCs for the LS-based detector for 20,000 synthetic samples (10,000 for each hypothesis). (b) Empirical ROCs for the GP detector for 4,000 samples (2,000 for each hypothesis). The data was generated using (2) and (19), and the nonlinear pixels were generated with $\gamma = [1,\ 3,\ 5]$ ($\eta_d = [0.22,\ 0.55,\ 0.80]$). (c) Comparison of the empirical ROCs for GP and LS detectors for $\gamma = 3$ and 4,000 samples. The noise power σ_n^2 was chosen in the three tests to obtain a SNR of 21dB.

5.2. Unknown $\boldsymbol{M}$

In this section we illustrate the sensitivity of the detection performance to the endmember estimation as a function of the degree of nonlinearity. These results are for an endmember extraction using the well-known vertex component analysis (VCA) [23]. Figure 2 presents the results of 4 experiments using synthetic data with 5,000 samples, SNR of 21dB, random abundances, and proportion of nonlinearly mixed pixels in the image varying from 10% to 50%. For every experiment, the endmember matrix was extracted using VCA. These results show how the detection performance can degrade as

the number of nonlinear pixels increases and VCA looses accuracy in extracting the endmembers from the image. Thus, alternatives to VCA must be sought for nonlinearly-mixed pixels.

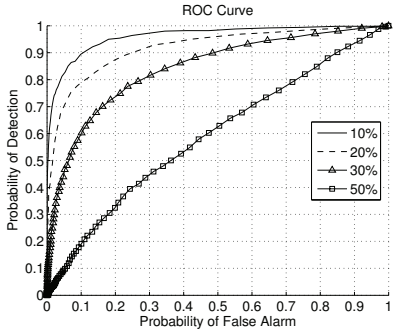


Fig. 2: ROCs for different proportions of nonlinearly mixed pixels.

5.3. Real Data

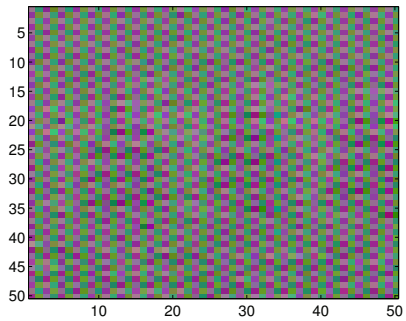
To test the GP detector using real images, we used the well-known data set available at the Indian Pines test site in North-western Indiana [24]. This image was captured by the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer), and has 145×145 samples. Each sample has 220 contiguous bands with wavelengths ranging from 366 to 2497 nm. Prior to analysis, noisy and water absorption bands were removed resulting in a total of 200 bands that were decimated to 50 to speed up simulations. The data set has a ground truth map that divides the samples in 17 mutually exclusive classes. Figure 3(a) displays an image with 50×50 pixels from the Indian Pines region. Figure 3(b) presents the ground truth map for this image, where each class is represented by a different color. The detection was performed in small 10×10 windows of the original 50×50 image at a time. For each window, the endmembers were locally extracted as follows:

1. a principal component analysis (PCA) was performed on the pixels belonging to a given class included in this window. The vector associated with the largest eigenvalue was selected for each class;
2. the data belonging to a given class of this window were projected onto the corresponding eigenvector, and the residual error between each pixel and its projection was computed;
3. we computed one endmember per class by averaging the 50% of the pixels having the smallest residual error.

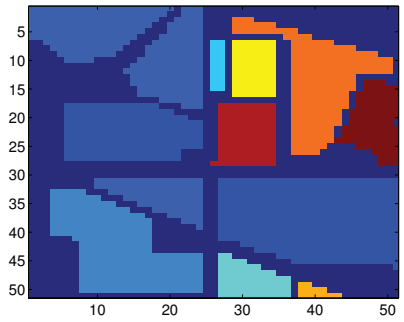
Once the endmember matrix had been estimated as described above, we computed a detection threshold for each window as follows:

1. we created an image with the linear model $\boldsymbol{Y}_s = \hat{\boldsymbol{M}}\hat{\boldsymbol{A}}$, using the endmember matrix $\hat{\boldsymbol{M}}$ extracted as described above. Matrix $\hat{\boldsymbol{A}}$ was computed using LS;
2. we then computed the detection test statistics $T|\mathcal{H}_0$ under $\mathcal{H}_0$ defined in (18);
3. finally, we computed the test threshold from the given PFA and the inverse cumulative distribution of the beta distribution.

The detection threshold was determined as described above for a PFA = 0.001. Figure 4 shows the pixels detected as non-linearly mixed (indicated by black circles) superimposed to the ground truth. The nonlinear mixtures are mainly detected close to the the class boundaries and in the background.



(a) Indian Pines.



(b) Ground Truth.

Fig. 3: (a) Part of the Indian Pines image with 50×50 pixels. (b) Ground truth (each color corresponds to a given material).

Fig. 4: Detection map: black circles indicate the pixels detected as non-linearly mixed.

6. CONCLUSIONS

A GP-based nonlinearity detection strategy was introduced to detect nonlinearly mixed pixels in hyperspectral images. The proposed detector does not require the use of a parametric model for the underlying nonlinear mixing function. Simulations using synthetic data indicate that the proposed detector outperforms a robust method previously presented in the literature. The detector was also tested on the Indian Pines image showing that pixels close to the class boundaries and in the background seem to be non-linearly mixed. Future work includes joint detection of nonlinear mixtures and unmixing.

7. REFERENCES

[1] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, 2002.

[2] D. Manolakis, D. Marden, and Shaw G. A., "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 79–116, 2003.

[3] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[4] N. Dobigeon, J.-Y. Tourneret, C. Richard, J.-C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: models and algorithms," *IEEE Signal Processing Magazine*, 2014.

[5] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.

[6] T. W. Ray and B. C. Murray, "Nonlinear spectral mixing in desert vegetation," *Remote Sensing of Environment*, vol. 55, no. 1, pp. 59–64, 1996.

[7] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," *Proceedings of SPIE*, vol. 7477, 2009.

[8] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. Van der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1183–1193, 2009.

[9] K. J. Guilfoyle, M. L. Althouse, and C.-I. Chang, "A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 10, pp. 2314–2318, 2001.

[10] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulatedforest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.

[11] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a polynomial post nonlinear model for hyperspectral imagery," in *Proc. IEEE ICASSP*, 2011, pp. 1009–1012.

[12] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 480–492, 2013.

[13] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral images with multi-kernel learning," in *Proc. IEEE WHISPERS*, 2012, pp. 1–4.

[14] J. Chen, C. Richard, and P. Honeine, "Estimating abundance fractions of materials in hyperspectral images by fitting a post-nonlinear mixing model," in *Proc. IEEE WHISPERS*, 2013, pp. 1–4.

[15] J. Chen, C. Richard, and P. Honeine, "Nonlinear estimation of material abundances in hyperspectral images with L1-norm spatial regularization," *IEEE Transactions on Geoscience and Remote Sensing*, 2013 (to appear).

[16] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinearity detection in hyperspectral images using a polynomial post-nonlinear mixing model.," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1267–1276, 2013.

[17] Y. Altmann, N. Dobigeon, J.-Y. Tourneret, and J. C. M. Bermudez, "A robust test for nonlinear mixture detection in hyperspectral images," in *Proc. IEEE ICASSP*, 2013.

[18] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, 2006.

[19] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, The MIT Press, 2001.

[20] W. Liu, J. C. Prncipe, and S. Haykin, *Kernel Adaptive Filtering: A Comprehensive Introduction*, Wiley, 2010.

[21] N. J. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, Wiley-Interscience, 1994.

[22] RSI (Research Systems Inc., "Envi user's guide version 4.0," Sept. 2013.

[23] J. M. P. Nascimento and J. M. Bioucas Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.

[24] I. Dopido, M. Zortea, A. Villa, A. Plaza, and P. Gamba, "Unmixing prior to supervised classification of remotely sensed hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 760–764, 2011.