Robust Unmixing Algorithms for Hyperspectral Imagery

Abderrahim Halimi(1), Yoann Altmann(1), Gerald S. Buller(1), Steve McLaughlin(1), William Oxford(2), Damien Clarke(2), and Jonathan Piper(2)

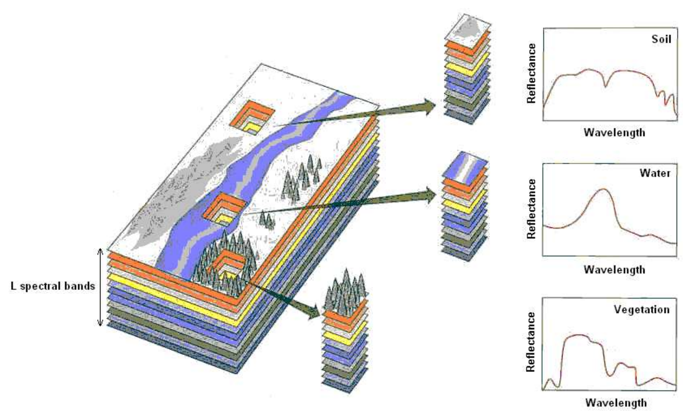
1 School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K
2 Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire U.K.
Corresponding author: a.halimi@hw.ac.uk

This work was supported by the EPSRC Grants EP/J015180/1, EP/N003446/1, and EP/K015338/1

1. Introduction

Hyperspectral imagery

- same scene observed at different wavelengths
- pixel represented by a vector of hundreds of measurements



Hyperspectral imaging concept.

⇒ Analysis of robust algorithms in presence of different effects (nonlinearity NL, endmember variability EV, outliers, ...)

2. Models and algorithms

General formulation

- Residual component analysis model (RCA)

$$\boldsymbol{y}_n = f\left[\boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n\right],$$

- $\boldsymbol{M} = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_R]$: matrix of endmembers
- $\boldsymbol{a}_n = [a_{1,n}, \ldots, a_{R,n}]^T$: abundance vector
- $\boldsymbol{\phi}_n$: residual component
- $f(.)$: function introducing noise (additive, Gaussian, Poisson, ...)

Robust nonnegative matrix factorization (RNMF) [1]

- **Observation model:** $\boldsymbol{y}_n^{\text{RNMF}} = \sum_{r=1}^{R} a_{r,n}\boldsymbol{m}_r + \boldsymbol{r}_n$
- **Assumptions:**
  - $\boldsymbol{r}_n, \forall n$ denote positive residuals that are spatially sparse
  - Two similarity measures $D(\boldsymbol{Y}|\boldsymbol{M}\boldsymbol{A}+\boldsymbol{R})$:
    squared Euclidean distance (SED) ⟹ Gaussian noise
    Kullback-Leibler divergence (KLD) ⟹ Poisson statistics
- **Estimation algorithm for** $(\boldsymbol{A}, \boldsymbol{R})$
  - An optimization approach using the cost function:

$$\min_{\boldsymbol{A},\boldsymbol{R}} \left\{ D(\boldsymbol{Y}|\boldsymbol{M}\boldsymbol{A}+\boldsymbol{R}) + \lambda ||\boldsymbol{R}||_{2,1} \right\}$$

  where $||\boldsymbol{R}||_{2,1} = \sum_{n=1}^{N} \sqrt{\boldsymbol{r}_n^T \boldsymbol{r}_n}$
  - Coordinate descent algorithm (CDA).

Robust Bayesian linear unmixing (RBLU) [2]

- **Observation model:** $\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n + \boldsymbol{e}_n$, with $\boldsymbol{e}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ and a diagonal covariance $\boldsymbol{\Sigma} = \text{diag}\{\boldsymbol{\sigma}^2\}$
- **Assumptions:**
  - $\boldsymbol{\phi}_n = \boldsymbol{z}_n \odot \boldsymbol{x}_n, \forall n$: where $\odot$ is a term-wise product
  - The support $\boldsymbol{z}_n$ is spatially-spectrally sparse and correlated
  - The values $\boldsymbol{x}_n, \forall n$ are independent
- **Estimation algorithm for** $\Theta_1 = (\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{Z}, s^2, \boldsymbol{\beta})$ where $(s^2, \boldsymbol{\beta})$ are hyperparameters
  - A Bayesian approach using the likelihood $f(\boldsymbol{Y}|\Theta_1)$ and the parameter prior distributions $f(\Theta_1)$ to build the posterior
  - Posterior distribution:

$$f(\Theta_1|\boldsymbol{Y}) \propto f(\boldsymbol{Y}|\Theta_1) f(\Theta_1),$$

  - Evaluate the MMSE or MAP estimators using a Markov chain Monte-Carlo (MCMC) algorithm.

RCA with mismodelling effects (RCA-ME) [3]

- **Observation model:** $\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{d}_n + \boldsymbol{e}_n$, with $\boldsymbol{e}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ and a diagonal covariance $\boldsymbol{\Sigma} = \text{diag}\{\boldsymbol{\sigma}^2\}$
- **Assumptions:**
  - $\boldsymbol{d}_n, \forall n$ are spectrally smooth
  - The energies of $\boldsymbol{d}_n, \forall n$ are spatially correlated
- **Estimation algorithm for** $\boldsymbol{\Theta}_2 = (\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{D}, \boldsymbol{\epsilon}, \boldsymbol{w})$ where $(\boldsymbol{\epsilon}, \boldsymbol{w})$ are hyperparameters
  - A Bayesian approach using the likelihood $f(\boldsymbol{Y}|\boldsymbol{\Theta}_2)$ and the prior distributions $f(\boldsymbol{\Theta}_2)$ to build the posterior distribution:

$$f(\boldsymbol{\Theta}_2|\boldsymbol{Y}) \propto f(\boldsymbol{Y}|\boldsymbol{\Theta}_2) f(\boldsymbol{\Theta}_2),$$

  - An optimization approach using the cost function:

$$\min_{\boldsymbol{\Theta}_2} \left\{ -\log\left[f(\boldsymbol{\Theta}_2|\boldsymbol{Y}) \right] \right\}$$

  - CDA to approximate the MAP estimator of $\boldsymbol{\Theta}_2$.

Characteristics of the robust models.

	Residuals			Statistics
	Positivity	Spatial	Spectral	
RNMF	✓	Sparse energies	None	-SED -KLD
RBLU	X	Correlated and sparse support	Correlated support	$\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$
ME	X	Correlated energies	Correlated values	$\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$

3. Results on a real image

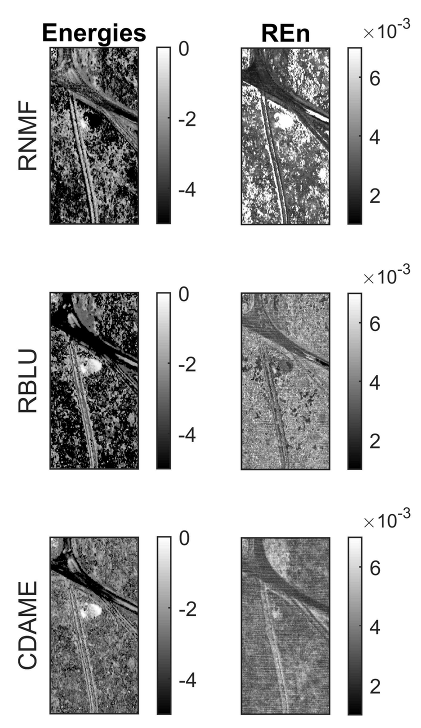
Data description and evaluation criteria

- A real image acquired by the Defence Science and Technology Laboratory (DSTL) in 2014 over Porton Down, U.K
- It contains 400×200 pixels, $L = 140$ bands in $[415, 990]$ nm, $R = 5$ components with other man-made outliers.

$$\text{RE} = \sqrt{\frac{1}{N\,L}\sum_{n=1}^{N} ||\hat{\boldsymbol{y}}_n - \boldsymbol{y}_n||^2}$$

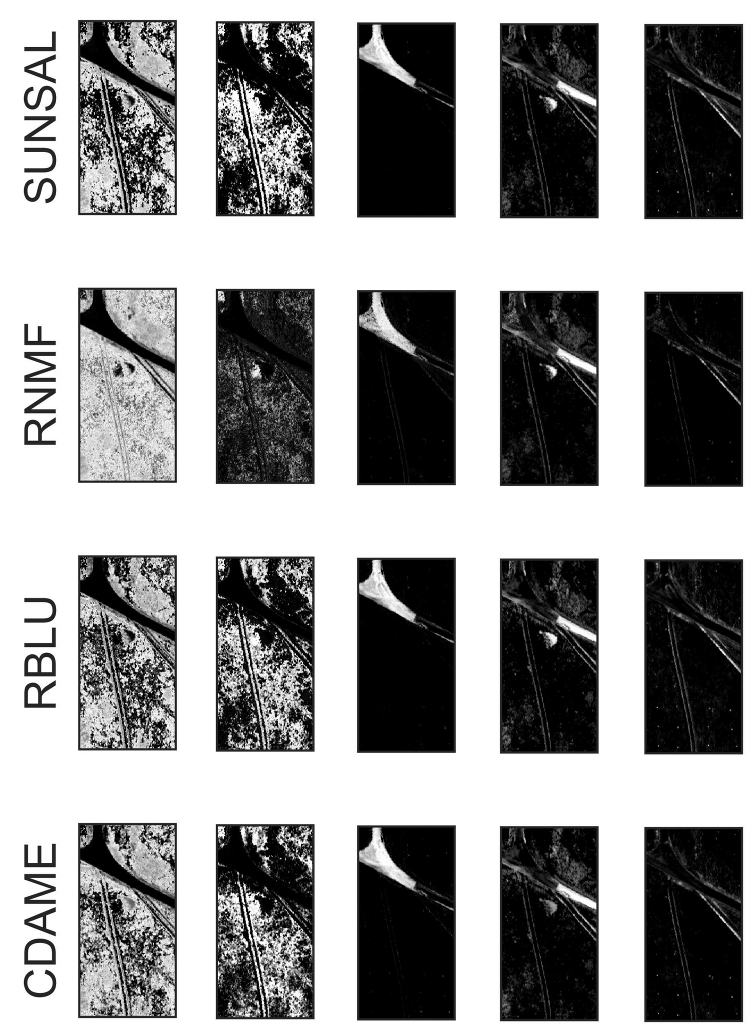
$$\text{SAM} = \frac{1}{N}\sum_{n=1}^{N} \arccos\left(\frac{\hat{\boldsymbol{y}}_n^T \boldsymbol{y}_n}{||\boldsymbol{y}_n||\ ||\hat{\boldsymbol{y}}_n||}\right)$$

Real Porton Down image.



(Left) square root of the energies of the residuals obtained with $||\boldsymbol{\phi}_n||$. (Right) reconstruction error obtained with $\text{RE}_n = \frac{1}{\sqrt{L}}||\hat{\boldsymbol{y}}_n - \boldsymbol{y}_n||$.

Abundance maps

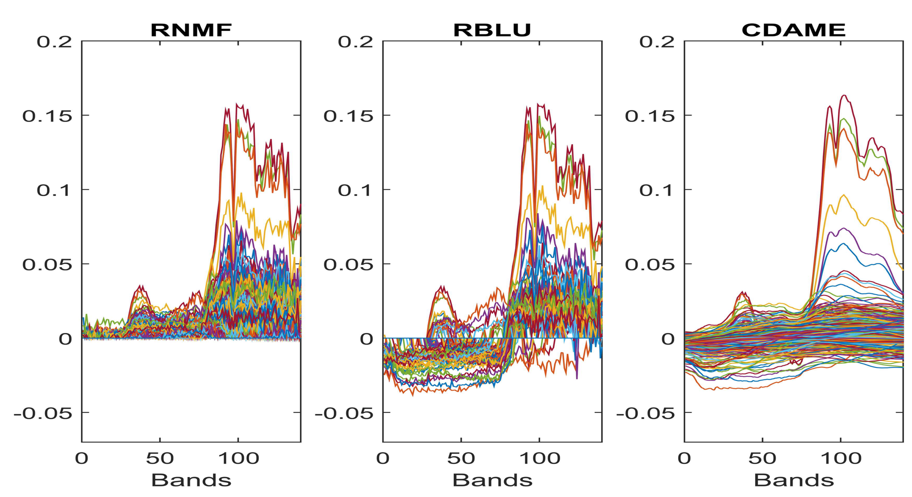


Estimated abundance maps with different algorithms (the gray scale ranges between 0 and 1). From left to right: grass, tree, soil 1, road, and soil 2.

Quantitative results

	RE ($\times 10^{-3}$)	SAM ($\times 10^{-2}$)	Time (min)
SUNSAL [4]	9.32	3.34	0.03
RNMF	7.17	2.05	23.92
RBLU	4.28	2.32	1440
CDA-ME	3.87	2.21	6.98

Unmixing performance on a real images (400×200 pixels).



Example of outlier spectra.

4. Discussion

- The robust algorithms detect man-made outliers
- Complementary results in terms of the detected physical effects, computational time and estimation quality

	Effects	Time	RE	Noise	Endmember estimation
SUNSAL [4]	LMM	+++	-	-	-
RNMF	LMM + NL	+	+	++	+
RBLU	LMM + NL EV + Shadow	-	++	++	+
CDAME	LMM + NL EV + Shadow	++	+++	++	-

Evaluation of the algorithms. (+++) best results, (-) fair results.

References

[1] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, June 2015.

[2] Y. Altmann, S. McLaughlin, and A. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEE Trans. Comput. Imaging*, vol. 1, no. 2, pp. 74–85, June 2015.

[3] A. Halimi, P. Honeine, and J. M. Bioucas-Dias, "Hyperspectral unmixing in presence of endmember variability, nonlinearity or mismodelling effects," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4565–4579, Oct 2016.

[4] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE GRSS WHISPERS*, June 2010, pp. 1–4.