ROBUST UNMIXING OF HYPERSPECTRAL IMAGES: APPLICATION TO MARS

Mario Parente, John F. Mustard

Brown University,
Dept. of Geological Sciences,
Providence, RI 02912

Scott Murchie, Frank Seelos

Applied Physics Laboratory,
Johns Hopkins University
Laurel, MD 20723

ABSTRACT

Planetary missions such as the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) can benefit from the use of automatic approaches and statistical learning techniques due to the amount of data involved. Thanks to its high sensor resolution, CRISM data volumes overwhelm scientists capacity for exhaustive manual analysis. Planetary investigations would benefit from an automated process that could identify the unique spectral signatures present in a CRISM scene and store them for further examination or interpretation. If installed aboard an orbital system, such a tool could relieve transmission constraints for high-bandwidth hyperspectral datasets by giving priority to the most informative data products. This paper introduces an algorithm that extracts image endmembers of a CRISM scene, which can be used as the scene concise mineralogical representation for cataloging purposes, in addition to existing browse products and parameter maps. The approach uses robust techniques, resilient to CRISM noise. This work benefits from the results of previous efforts [6] and it is currently being extended to other hyperspectral datasets.

Index Terms— One, two, three, four, five

1. INTRODUCTION

Hyperspectral sensors such as the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) [1, 2] can benefit from the use of automatic approaches and statistical learning techniques due to the amount of data involved. Thanks to its high sensor resolution, CRISM data volumes overwhelm scientist capacity for exhaustive manual analysis. Automated summary analysis could identify the unique spectral signatures present in a CRISM scene and store them for further examination or interpretation. If installed aboard an orbital system in a future mission, such a tool could even relieve transmission constraints for high-bandwidth hyperspectral datasets by selecting only the most informative data products for transmission.

This paper introduces an algorithm that extracts image endmembers of a hyperspectral scene, which can be used as the scene concise ground compositional representation for cataloging purposes and demonstrates the proposed technique's benefits in an application to a CRISM dataset. Automatic endmember extraction algorithms applied to CRISM datasets face particular challenges due to CRISM's peculiar noise characteristics [3, 4]. The proposed algorithm uses robust techniques, resilient to CRISM noise. This work benefits from the results of previous efforts [5] and could be a promising addition to already available spectral browse products and spectral summary parameter maps [6].

2. ENDMEMBER EXTRACTION PIPELINE

The unmixing strategy followed in this work can be summarized in the processing pipeline illustrated in Figure 1. The endmember extraction procedure [7] has several stages. CRISM images are first reduced by standard calibration and atmospheric reduction [1, 2]. An initial signal denoising is performed [3, 4]. Then the algorithm focuses on smaller areas of interest based on where available spectral parameter maps show higher spectral variability and higher concentrations of minerals of interest. Otherwise, the algorithm divides the image in smaller sub-images and operates separately on each one.

The steps that follow leverage the representation of the hyperspectral image as a cloud of pixel vectors in a space of dimension equal to the number of spectral channels. A nonlinear dimensionality reduction technique follows that is a novel variation over the t-Distributed Stochastic Neighbor Embedding [8], called similarity-based SNE (SSNE), where high-dimensional similarity measures between data-points x_i are converted into Gaussian probabilities $\mathbf{p}_{ij}$. The same representation is considered in the lower-dimensional space except that the probability distributions $\mathbf{q}_{ij}$ between datapoints y_i in the low dimensional space have longer tails to compensate for the mismatch in dimensionality. The similarity measures adopted try to capture local differences in the spectral shapes, which are the main diagnostic information for spectral matching. If the low-dimensional datapoints correctly represent the high-dimensional ones, their joint probabilities $\mathbf{q}_{ij}$ and $\mathbf{p}_{ij}$ should be equal. Therefore, the aim of SSNE is to minimize the mismatch between $\mathbf{q}_{ij}$ and $\mathbf{p}_{ij}$ in the form of

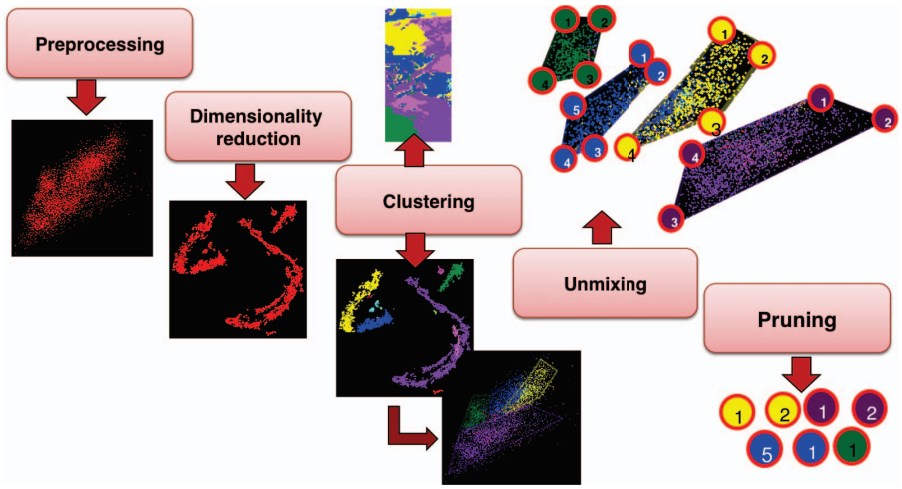


Fig. 1: Pipeline for current approach to unmixing

their Kullback-Leibler divergence.

SSNE tends to produce a lower dimensional representation of the data with well separated components. However the technique tends to create clusters that have non-convex shapes. A nice way of representing the data-points $\{y_1 \ldots, y_n\}$ is in form of the similarity graph $G = (V, E)$. Each vertex v_i in this graph represents a data point y_i. Two vertices are connected if the similarity $\mathbf{q}_{ij}$ between the corresponding datapoints y_i and y_j is positive or larger than a certain threshold, and the edge is weighted by $\mathbf{q}_{ij}$. The problem of identifying and labeling clusters in the data can now be reformulated as a graph partitioning. A variation on the spectral clustering algorithm proposed by Ng, Jordan, and Weiss [9] is use for partitioning the graph. The technique successfully identifies and labels the separated components produced by the SSNE as natural clusters.

The region in the original space corresponding to each cluster is further analyzed by a geometric unmixing algorithm that approximates the geometry of the region by a convex polyhedron whose vertices represent local endmembers. This algorithm assumes that the data-points of such a region can be approximated by convex combinations of the local endmembers. The technique is based on a robust variation of a Non-negative Matrix Factorization algorithm[10], originally applied to Mars exploration Rovers (MER) Panoramic camera multispectral images and is effective in combating the effects of CRISM noise. Briefly, if the data-points in a region are organized as the columns of a matrix $Z \in \mathbb{R}^{mxn}$, the k estimates of the endmembers as the columns of W. Each column of a matrix H can hold the mixing coefficients (or abundances) for each point. Under this setup linear spectral unmixing can be obtained by solving the following problem:

$$\begin{aligned} \text{minimize} \quad & \varphi(Y - WH) + \lambda ||DW||_F^2 \\ \text{subject to} \quad & W \geq 0, H \geq 0, \ \mathbf{1}^T H = \mathbf{1}^T, \end{aligned}$$

where $\varphi(x)$ is the $L1 - L2$ estimator, $||U||_F$ is the Frobenius norm of the matrix U and where in the quadratic constraint $||DW||_F^2$, D is a matrix that imparts smoothness to the endmember spectra while trying to preserve the shape of local features.

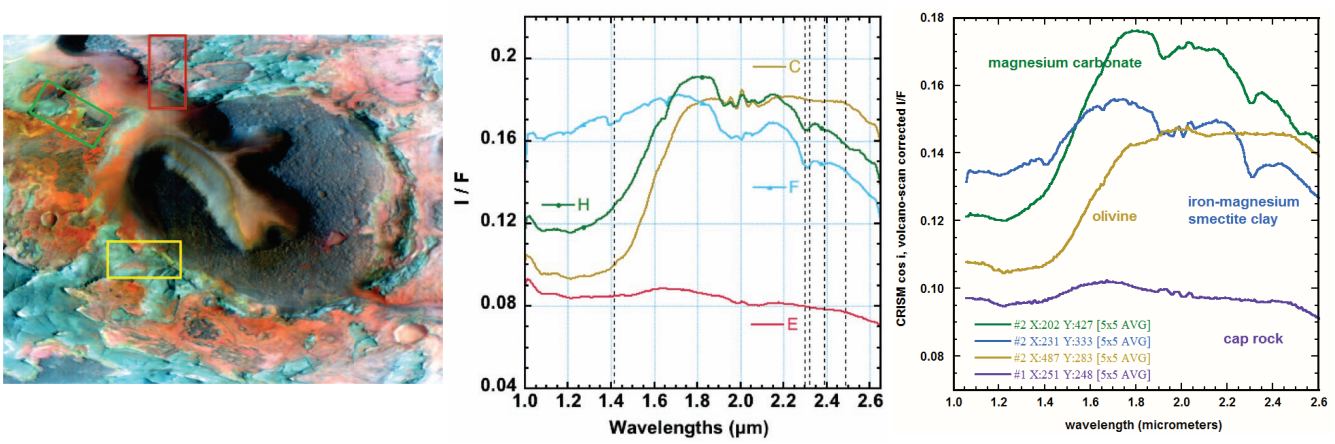
The list of all local endmembers has to be screened since the endmembers locally to a cluster might not be global endmembers for the whole data cloud. This happens, for example when a corner of a cluster lies in the interior of the data cloud. Screening is necessary in order to remove redundant spectra and endmembers due to instrument or calibration artifacts such as spectra affected by residual atmospheric contributions. The procedure screens the candidate spectra based on a spectral similarity score. Two spectra are considered similar if the score between them is smaller than a certain threshold.

3. RESULTS

We performed the procedure described above for the CRISM image FRT00003E12. The algorithm selected 3 areas of interest based on the parameter maps for band depths at 1 , 2.3 and 2.5 μm and extracted endmembers from each of them.

In order to evaluate the endmember detection method we first note that a good automated system should detect the same features that would be found from an exhaustive manual search. For the image under study FRT00003E12 we collected the endmembers found by the algorithm over interesting areas throughout the image and compare them with expert manual selections in Figure 2. Figure 2(a) represent the image FRT00003E12 with interesting areas in colored boxes. The manually picked cyan, gold and green spectra in Figure 2(c) divided by the 'cap-rock' spectrum have been published in several sources (e.g. [11]).

The proposed algorithm selected the spectra labeled H, C and F and E in Figure 2(b), which present similar features



(a) original image (b) Proposed algorithm (c) Manual selection from an expert

Fig. 2: Comparison of selected endmembers with expert assessment from the whole image

with respect to the spectra labeled 'magnesium carbonate', 'olivine', 'iron-magnesium smectite clay' and 'cap-rock' in Figure 2(c). The spectrum labeled H is identified as magnesium carbonate by virtue of the 2 absorptions – the dips or minima in the spectrum – around 2.3 and 2.5 μm. The spectrum labeled C is identified as olivine due to the broad absorption around 1 μ. The smectite clay label for spectrum F comes from the absorption around 2.29 μm (and a smaller absorption around 2.4 μm).

As a further test of performance, we compared the results of the proposed unmixing algorithm with 4 competitors on a case study – the red box in Figure 2(a) – in order to visually assess the quality of each approach. The competitor algorithms are the vertex component analysis (VCA) [12], the sequential maximum angle convex cone (SMACC) [13], the pixel purity index (PPI) algorithm [14] and the N-FINDR [15] method. Since all algorithms required the number of endmembers to be found as an input we set that parameter to 5, the same number extracted by the proposed algorithm. we ran all four algorithms on the case study. VCA, PPI and NFINDR did not require additional parameters. SMACC has a pruning features that eliminates one of the endmembers from a pair of candidates if their angle is smaller then a user defined threshold. We used the default value of 0.1 for the parameter. The results of the comparison are shown in Figure 3.

All algorithms with the exception of the proposed one are sensitive to different degrees to CRISM spiking noise. It is interesting to notice that 3 out of 4 competitor algorithms identify a spectrum with a large spike as an endmember, as reflected by the red spectrum in Figure 3(c) (for VCA), the green and magenta ones in Figure 3(b) (for SMACC) and the green curve in Figure 3(d) (for NFINDR). PPI is less sensitive to noise, with the green and red spectra in Figure 3(e) presenting some spikes. This results are not unexpected since only the proposed algorithm implements a strategy for robustness against CRISM noise.

The method presented in this paper seems to capture the main spectral signatures present in the case study image (spectra T, O and J in Figure 3(a) are similar to spectra H, F and C of Figure 2(b), spectrum E in Figure 3(a) could be a possible mafic signature). The spectral variability identified by the 4 competitors is relatively limited. Only 2 types of spectra can be identified in Figures 3(b)-(e) based on the absorption band around 2.3 μm and the large absorption at 1μm. This result could be explained by the fact that these algorithms optimize a global objective that can miss local variations in the data cloud.

4. CONCLUSIONS

We presented an endmember extraction technique that performed well on a CRISM scene when compared to manual selection of endmembers and 4 currently available endmember selection competitors. The results are confirmed on several other images and we will report the findings on an extended version of this paper. Furthermore, we will propose method as a possible data summarization tool for the CRISM team.

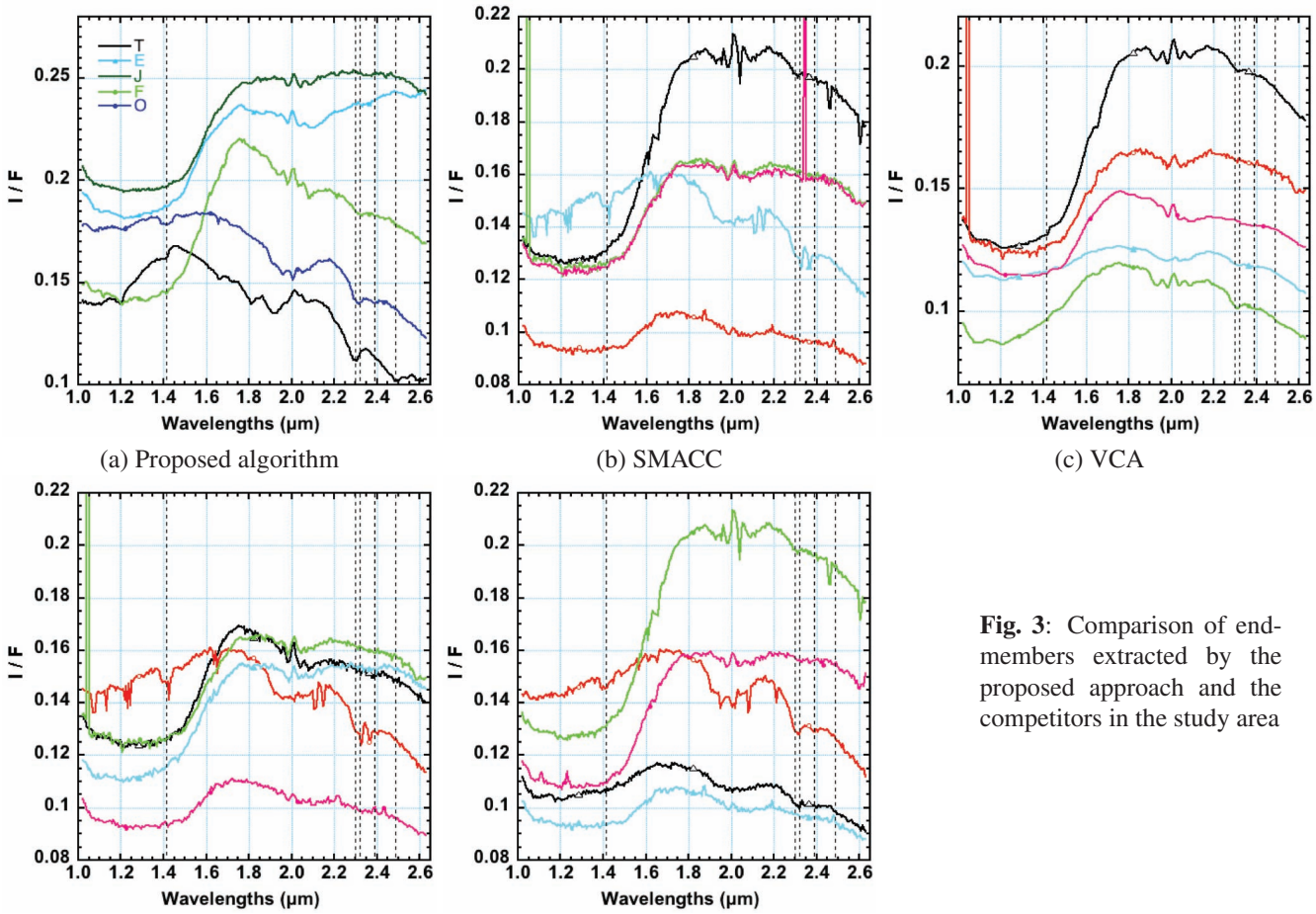
REFERENCES

[1] S.L. Murchie et al., "Compact reconnaissance imaging spectrometer for mars (crism) on mars reconnaissance orbiter (mro)," *J. Geophys. Res*, vol. 112-5, 2007.

[2] S.L Murchie et al., "Compact reconnaissance imaging spectrometer for mars investigation and data set from the mars reconnaissance orbiter's primary science phase," *J. Geophys. Res.*, vol. 114, no. E00D07, 2009.

[3] M. Parente, "A New Approach to Denoising CRISM Images," in *Lunar and Plan. Sci. Conf. XXXIX*, 2008, number 2528.

[4] M Parente, J.T Clark, and J.L Bishop, "End-to-end simulation and analytical model of remote-sensing systems:



(a) Proposed algorithm (b) SMACC (c) VCA

(d) NFINDR (e) PPI

Fig. 3: Comparison of endmembers extracted by the proposed approach and the competitors in the study area

Application to crism," *IEEE Trans. Geosc. and Rem. Sens.*, vol. 48, no. 11, pp. 3877–3888, 2010.

[5] M. Parente and J.L. Bishop, "Extracting endmember spectra from crism images: comparison of new direx image transform technique with mnf," in *Lunar and Plan. Sci. Conf. XXXX*, 2010, number 2633.

[6] S. Pelkey et al., "Crism multispectral summary products: Parameterizing mineral diversity on mars from reflectance," *J. Geophys. Res.*, vol. 112, no. E08S14, 2007.

[7] M. Parente et al., "Unsupervised unmixing of hyperspectral images: Imaging the martian surface," *IEEE Trans. Geosci. Remote Sens. (submitted)*, 2010.

[8] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *J. of Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008.

[9] AY Ng, MI Jordan, and Y Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 2, pp. 849–856, 2002.

[10] M Parente, JL Bishop, and JF Bell, "Spectral unmixing for mineral identification in pancam images of soils in gusev crater, mars," *Icarus*, vol. 203, no. 2, pp. 421–436, 2009.

[11] J.F. Mustard et al., "Hydrated silicate minerals on Mars observed by the CRISM instrument on MRO," *Nature*, vol. 454, pp. 305–309, 2008.

[12] J.M.P. Nascimento and J.M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosc. and Rem. Sens.*, vol. 43, no. 4, pp. 898–910, 2005.

[13] J. Gruninger, A. Ratkowski, and M. Hoke, "The sequential maximum angle convex cone (smacc) endmember model," *Proceedings of SPIE*, vol. 5425, 2004.

[14] J. Boardman, F.A. Kruse, and R.O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Summaries of the JPL Airborne Earth Science Workshop*. JPL Publication, 1995, pp. 23–26.

[15] M.E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proceedings of SPIE*, 1999, vol. 3753, pp. 266–270.