

QUANTIFYING MINERAL ABUNDANCES IN MIXTURES USING RAMAN SPECTROSCOPY: TOWARD A METHOD FOR SPECTRAL UNMIXING. CJ Carey¹, L. B. Breitenfeld², M. D. Dyar², M. C. Crowley², C. Leight², and E. Watts². ¹College of Information and Computer Science, Univ. of Massachusetts, Amherst, MA 01003, ccarey@cs.umass.edu, ²Mount Holyoke College, Dept. of Astronomy, South Hadley, MA, 01075.

Introduction: Planetary scientists who use reflectance and thermal emission spectroscopy are spoiled by the availability, breadth, and sophistication of unmixing software and spectral libraries available for commonly used techniques, like the Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER) library at JPL [1], the U.S.G.S. library [2], and Reflectance Experiment Library (RELAB) at Brown University [3]. Many techniques for extracting quantitative abundances of minerals from spectra of mineral mixtures are available for these wavelengths. For example, radiative transfer (RT) theory can be used to tackle the problem of nonlinear spectral mixing and extract quantitative mineral abundance in the VNIR [4-8]. Multiple scattering models can provide approximate numerical [9] or analytical solutions to the RT in a particulate medium by considering a small elementary volume of the scattering medium in which the individual particles have some average albedo and phase function [10,11]. However, there is no analogous methodology for mineral unmixing for the technique of Raman spectroscopy because most development work has been focused on micro-Raman techniques [12] that probe pure minerals. Impending use of Raman spectroscopy on Mars by SuperCam [13], which will have a cm-scale sampling size, will however require development of techniques for unmixing.

There is hope for theoretical approaches to unmixing of Raman spectra because of synergies between the two different types of spectroscopy. Raman features arise from scattering of energy while FTIR spectra have absorption features. Raman provides information on covalency of molecular bonds, while FTIR indicates ionic character. Raman peaks reflect changes in bond polarization, while FTIR peaks record dipole changes. Moving forward, unmixing approaches that build on techniques currently used for FTIR should at a minimum provide a starting point for theoretical unmixing models in Raman.

However, the current best approach to unmixing of Raman signals comes from the machine learning community. This abstract provides a new method for automatically unmixing Raman data.

Data: All Raman spectra used are pairwise mixtures of 24 minerals in varying proportions designed to provide measurable peak areas in each spectrum; the resultant ratios and combinations as described in [14,15]. The 24 minerals species used include alunite, anhydrite, augite (clinopyroxene), bytownite (feld-

spar), calcite, clinocllore (chlorite), diamond, diopside (clinopyroxene), enstatite (orthopyroxene), forsterite (olivine), gypsum, hematite, ilmenite, jarosite, rozenite, labradorite (feldspar), magnesite, magnetite, montmorillonite, nontronite, saponite, siderite, tremolite, and chabazite (zeolite) [16]. Chemical analyses were obtained using electron microprobe, and all samples were crushed, handpicked for purity, and sieved to grain sizes <64 μm . X-ray diffraction was used to confirm the homogeneity of each pure phase. Samples were then weighed into binary mixtures in ratios of 50:50, 20:80, and 95:5 volume percent depending on the Raman cross sections of the paired phases. These data permit comparisons to be made between mineral-diamond mixtures, spectra from the RRUFF (<http://rruff.info/>) Raman library (using only spectra from samples for which XRD confirmation is provided), mineral-mineral mixtures excluding diamond, and the pure mineral end-members used in all the mixtures.

Spectra were acquired on a Bruker Optics, Inc. BRAVO Raman spectrometer, which uses dual laser (758 and 852 nm) excitation and a patented fluorescence mitigation strategy involving successive heating of the laser. Each sample was run three times with a scan time of 10s and wave-number range of 300-3350 cm^{-1} , then the runs were averaged to produce the final spectrum. The BRAVO spectrometer produces baseline subtracted data.

Methods: All spectra were preprocessed with maximum intensity normalization, and pairwise comparisons between spectra were made using whole spectrum matching (WSM) techniques [16]. The chosen WSM algorithm used the generalized match score formulation [17]:

$$MS(y_A, y_B) = (1 - w)y_A y_B - w|y_A - y_B|,$$

where w is a blending parameter between 0 and 1. The overall distance between two spectra is computed by summing $MS(y_A, y_B)$ for all corresponding intensities, y , in the pair [16].

The procedure for mineral unmixing is as follows: (1) Preprocess the query spectrum of an unknown mixture of minerals, as well as the target spectral library of known pure minerals. (2) Using WSM, output a list of the query's top k matches in the pure-phase target library. (3) Using the original query, Q , and each top- k matching spectrum, M_i , create a new query spectrum, Q' , by zeroing the intensities where $MS(y_Q, y_{M_i})$ is below some threshold. In our experiments, we defined

this threshold as the first percentile of all $MS(y_Q, y_{M_i})$. (4) Repeat step 2 using Q' to find the second mixture component, multiplying the distances from each match to produce a final, composite distance measure. (5) Order the resulting k^2 mixtures by their composite distances.

Results: Matching accuracies for minerals in mixtures are given in **Table 1**, broken down by Dana class, type, group, and species (broadest to most specific mineral names). The top set of numbers shows how well our algorithm matches both minerals in a mixture in its top-scoring pair. The lower set of data indicates whether a successful exact match was found among any of the top five predictions. In each case, three different combinations are used: 1) using minerals in mixtures with diamond and matching them against the RRUFF library, 2) matching mixtures of minerals other than diamond vs. the RRUFF library, and 3) matching mixtures of non-diamond phases against a library including only their own pure mineral spectra.

Table 1. %Accuracy of Binary Unmixing

Dana:	class	type	group	species
Diamond mixtures vs. RRUFF library (top match):				
At least one match	91.25	91.25	90.00	90.00
Perfect match	33.75	30.00	30.00	23.75
Non-diamond mixtures vs. RRUFF library (top match):				
At least one match	45.24	42.06	39.68	31.35
Perfect matches	9.52	6.35	6.35	2.78
Non-diamond mixtures vs. pure samples (top match):				
At least one match	51.59	50.00	42.86	38.10
Perfect match	7.54	7.54	6.75	3.57
Diamond mixtures vs. RRUFF library (top 5 matches):				
At least one match	93.75	92.50	91.25	91.25
Perfect match	45.00	41.25	40.00	35.00
Non-diamond mixtures vs. RRUFF library (top 5 matches):				
At least one match	55.56	53.18	50.00	42.06
Perfect match	11.91	7.94	7.54	4.76
Non-diamond mixtures vs. pure samples (top 5 matches):				
At least one match	83.33	82.94	75.34	73.02
Perfect match	22.62	19.84	17.46	11.91

These results much be viewed in the context of the fact that mineral identification based on Raman data of pure mineral phases is far from perfect: pure mineral spectra can be matched against the RRUFF database only ~88% of the time [17].

As shown, these results show some interesting trends. Most importantly, they demonstrate that whole spectrum matching is flexible enough to perform well on certain classes of impure spectra, identifying at least one of the mixture components in 90% of the diamond mixtures. The best matching accuracy for non-diamond pairs resulted when the pure phases used in the actual mixtures comprised the target library. This setting eliminates most of the variation caused by sample acquisition differences, while narrowing the pool of po-

tential mixture components. The overall accuracy of perfect matching is disappointingly low, reminding us that this is a non-trivial problem and we are just beginning to understand it. For example, to what extent do grain size and Raman cross-section cause complex interactions to occur among Raman peaks, affecting peak intensities?

Future work: This preliminary study suggests to us several avenues for improvement in unmixing of Raman data. We will undertake customized baseline removal from all spectral data (both in libraries and unknowns), as this has been shown to improve matching accuracy [18]. We also suspect we can improve matching results by more carefully removing the first endmember's contribution before attempting to match successive minerals. Mixtures with overlapping peaks will also need special treatment. Work to pursue each of these avenues is in progress.

Acknowledgements: We are grateful to Yan Wang and Bruker Optics, Inc. for the loan of a BRAVO Raman unit. This work was supported by NASA SSERVI grant NNA14AB04A (RIS⁴E) and the Massachusetts Space Grant Consortium.

References: [1] Rull F. et al. (2014) *8th Intl. Mars Conf.*, Abstract #1277. [2] Beegle L. W. et al. (2014) *11th Intl. GeoRaman Conf.*, Abstract #5101. [3] Maurice S. et al. (2015) *LPS LXXVI*, Abstract #2818. [4] Clark R. N. and Roush T. L. (1984), *JGR*, 89, 6329-6340. [5] Mustard J. F. and Pieters, C.M. (1987) *Icarus*, 125, 145-163. [6] Cahill J. T. and Lucey, P. G. (2007) *JGR*, 112, 2156-2202. [7] Li S. and Li L. (2011) *JGR*, 116, 2156-2202. [8] Mishenko M. et al. (1999) *J. Quant. Spectrosc. Radiat. Trans.*, 63, 409-432. [9] Hapke, B. (2012) *Theory of Reflectance and Emittance Spectroscopy*, 2nd edition. Cambridge University Press, NY. [10] Shkuratov Y. et al. (1999) *Icarus*, 137, 235-246. [11] Haskin L. A. et al. (1997) *JGR*, 102, 19293-19306. [12] Clegg S. M. et al. (2015) *LPS LXXVI*, Abstract #2781. [13] Breitenfeld, L. et al. (2016) *LPS LXXVII*, Abstract #2430. [14] Breitenfeld, L. et al. (2016) *LPS LXXVII*, Abstract #2186. [15] Carey CJ et al. (2015) *J. Raman Spectrosc.*, 46, 894-903. [16] Carey CJ et al. (2015) *IJCAI15*, Buenos Aires. [17] Carey CJ et al. (2015) *LPS LXXVI*, Abstract #2464.