PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Geological applications of machine learning on hyperspectral remote sensing data

C. H. Tse, Yi-liang Li, Edmund Y. Lam

C. H. Tse, Yi-liang Li, Edmund Y. Lam, "Geological applications of machine learning on hyperspectral remote sensing data," Proc. SPIE 9405, Image Processing: Machine Vision Applications VIII, 940512 (27 February 2015); doi: 10.1117/12.2178400

Event: SPIE/IS&T Electronic Imaging, 2015, San Francisco, California, United States

SPIE.

Geological Applications of Machine Learning in Hyperspectral Remote Sensing Data

C. H. Tse*[a], Yi-liang Li[a], Edmund Y. Lam[a,b]
[a]Department of Earth Sciences, The University of Hong Kong, Pokfulam, Hong Kong, China
[b]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China

ABSTRACT

The CRISM imaging spectrometer orbiting Mars has been producing a vast amount of data in the visible to infrared wavelengths in the form of hyperspectral data cubes. These data, compared with those obtained from previous remote sensing techniques, yield an unprecedented level of detailed spectral resolution in additional to an ever increasing level of spatial information. A major challenge brought about by the data is the burden of processing and interpreting these datasets and extract the relevant information from it. This research aims at approaching the challenge by exploring machine learning methods especially unsupervised learning to achieve cluster density estimation and classification, and ultimately devising an efficient means leading to identification of minerals.

A set of software tools have been constructed by Python to access and experiment with CRISM hyperspectral cubes selected from two specific Mars locations. A machine learning pipeline is proposed and unsupervised learning methods were implemented onto pre-processed datasets. The resulting data clusters are compared with the published ASTER spectral library and browse data products from the Planetary Data System (PDS). The result demonstrated that this approach is capable of processing the huge amount of hyperspectral data and potentially providing guidance to scientists for more detailed studies.

Keywords: Hyperspectral, unsupervised learning, cluster classification, CRISM

1. INTRODUCTION

Since the launch of human's first planetary probe, imaging and spectrometry have been the key tools for scientists to study planetary objects remotely [1]. While conventional photographic imaging provides the visual context of the objects of interest, spectrometry allows worker access the composition of materials on the surface [2, 3, 4].

Mars, as our closest planetary neighbor, has been intensively studied for its present and past conditions on its surface and subsurface, and a lot of hints can be obtained by studying its mineralogy [5]. When multispectral sensors were first deployed to Mars about two decades ago, spectral parameters such as band ratios and spectral slopes were utilized to determine the global mineralogy of the Martian surface [6]. While such a robotic rendezvous can undoubtedly bring about datasets with a spatial resolution higher than ground-based observations, better spectral resolution must be achieved in order to characterize the surface mineralogy in greater details, so as to answer key questions such as whether aquatic environments ever existed on Mars [5].

Higher spectral resolution is now accomplished by hyperspectral imaging instruments capable of providing unprecedented spectral details to its observation targets [7]. The Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) aboard the Mars Reconnaissance Orbiter (MRO) is the most recent and advanced hyperspectral imager to arrive at Mars, covering visible to near-infrared (VNIR) wavelengths (0.37–3.92um) at 6.55 nm per channel. Two observation modes are provided by this instrument. The survey mode can achieve global coverage of the planet at 200 m/pixel in 72 selected channels, while targeted mode conducts observations covering less than 1% of the planet's surface at an unprecedented spatial resolution of 20 m/pixel in 545 channels for hyperspectral data. [8]

*tsekevin@hku.hk; phone 852 6110-1154; fax 852 3106-8039

Image Processing: Machine Vision Applications VIII, edited by Edmund Y. Lam, Kurt S. Niel,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9405, 940512 · © 2015 SPIE-IS&T
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.000000

While hyperspectral instruments open up a new era of planetary remote sensing, compared with multi-spectral imaging, new challenges have accompanied with its new strengths. One key challenge is the need to analyze and interpret the massive amount of hyperspectral data. Manual investigation of entire hyperspectral datasets is not feasible (e.g. CRISM high-resolution observations contain over 10^9 of spectra). Machine learning techniques have proven to be able to increase the efficiency of scientific discovery within such a vast amount of data and unsupervised machine learning provides a practical means for scientists to extract useful and meaningful portions in the sea of data for further study [9,10].

2. STUDY AREA AND DATA

2.1 Study Area

Two 10km x 10km (approximate) areas with channel / depression topography were chosen along the Martian equator in the study with a greater chance of mineral diversity especially hydrated species representing a past environment of liquid water. The terrain data was obtained from Mars Orbiter Laser Altimeter (MOLA) on Mars Global Surveyor (MGS) since 1996.

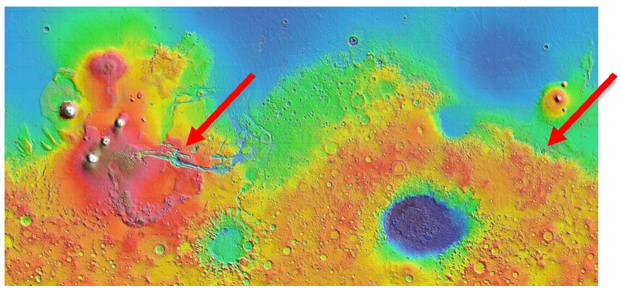

Figure 1. Red arrows indicated the two hyperspectral data cubes extracted in this study, namely the North of Melas Chasma Wall (left) and South of Gale Crater (right).

2.2 Data

Two instruments on CRISM with one collecting light from wavelength of 360 to 1060 nm (sensor ID ‘S’) and another one collecting light from 1000 to 3920 nm (sensor ID ‘L’) constituting the dataset. The data product available from the PDS Geosciences received an approximate correction for photometric effects (differences in local solar time and incidence solar angle), and converted from radiance to I/F [8].

Table 1. Hyperspectral data cubes details.

Observation ID	000132E6	00018285
Center Latitude	-7.81046	-5.569
Center Longitude	-75.78116	138.28294
Year/Day of Year	2009/164	2010/111
Location	North of Melas Chasma Wall	South of Gale Crater

2.3 Spectral Library

The ASTER library provides a comprehensive collection of over 2300 spectra comprising contributions from the Jet Propulsion Laboratory (JPL), Johns Hopkins University (JHU) and the United States Geological Survey (USGS). The library includes spectra of rocks, minerals, lunar soils, terrestrial soils, manmade materials, meteorites, vegetation, snow and ice covering the visible through thermal infrared wavelength region (0.4-15.4 μm). The purity and composition of each mineral sample was determined using standard X-ray Diffraction analysis. [11]

.

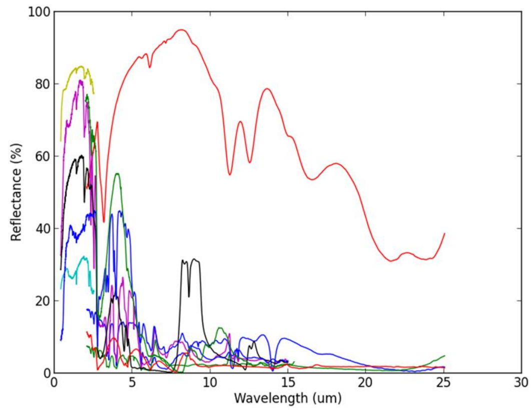

Figure 2. Selected 20 reflectance spectra for igneous (granite, quartzite, silicate), sedimentary (limestone, sandstone), hydrated (carbonate, fluorite, goethite, phosphate, sulfate) minerals and lithologies.

3. METHODS

3.1 Data Pipeline

A set of software tools is constructed for the study and the scientific programming language of Python is adopted for the following reasons:

i) cross-platform portability

ii) ease of implementation

iii) abundance of supporting modules and libraries.

Many successful Python packages (e.g. AstroML) [12] are available and one goal of this study is to develop a module to be used by other scientists working on similar hyperspectral datasets from the PDS of Geosciences of NASA [13].

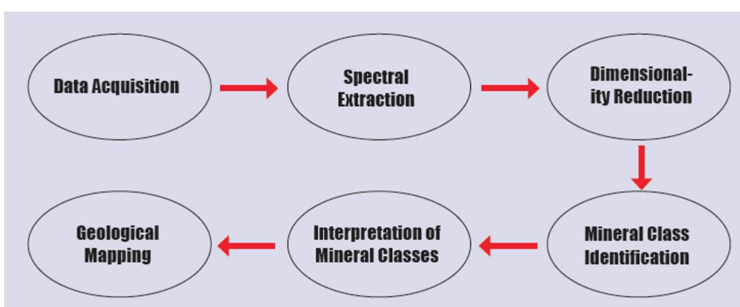

Figure 3. The proposed pipeline from data acquisition to geological mapping on hyperspectral data.

Spectral are extracted for each pixel in the hyperspectral datasets and in order to augment the S/N ratio and improve the efficiency of the classification algorithm, a group of up to five nearby pixels are adopted and “binned” to obtain the median spectra for each pixel in the observation dataset.

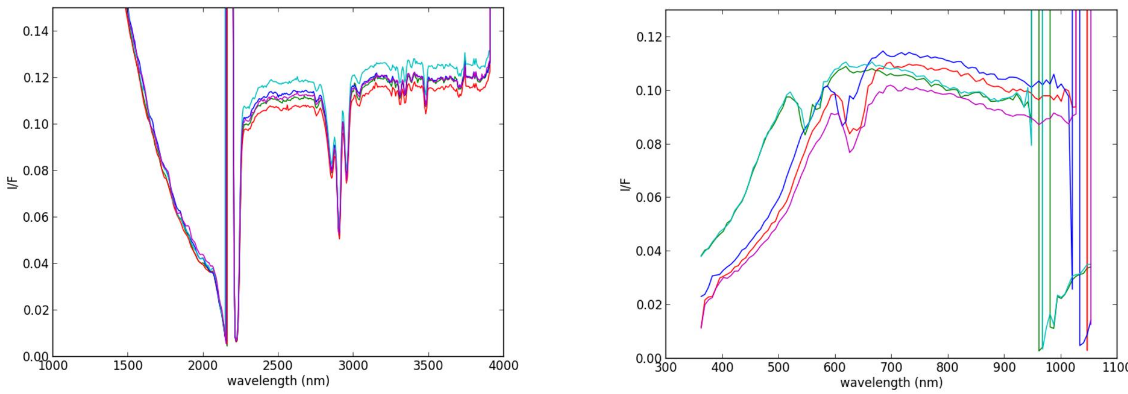

Figure 4. Sample pixel spectra for (Obs #000132E6L) on the left and (Obs #00018285S) on the right.

3.2 Data Implementation

$$U \sum V^T = \frac{1}{\sqrt{N-1}} X \tag{1}$$

$$\widehat{f}_N(x) = \frac{1}{Nh^D} \sum_{i=1}^{N} K\left(\frac{d(x, x_i)}{h}\right) \tag{2}$$

Progressive Principal Component Analysis (PG-PCA) from equation 1 [14] is adopted to achieve dimensional reduction to the TRDR (Targeted Reduced Data Record) from the CRISM hyperspectral data, where U and V are orthogonal basis where the data is to be projected onto.

Non-parametric Kernel Density Estimator (nKDE) from equation 2 [15] is the unsupervised learning method used in this study, where a uniform kernel was implemented and a range of bandwidths was adopted upon a number of different trials.

Slight modifications based on the packages AstroML [15] and scikit-learn [16] were adopted in the implementation of the above mathematical equations on the data.

4. RESULTS

Unsupervised classification based on nKDE was performed on hyperspectral cubes (Obs# FRT000132E6_07_IF154L and FRT00018285_07_F185S). Classification results were compared with browse products from Enhanced Visible Color RGB (592, 533, 492, 1330, BD530, SH600, BD1000, OLINDEX, LCPINDEX, HCPINDEX, BD2300, BD2210, BD1900, SINDEX, BD2100, BD1900 set out in [6]). Classification map is produced with the classes representing potential mineral groups but the interpretation of these classes are to be verified with other published data and future rover missions when a close-up geological survey is available.

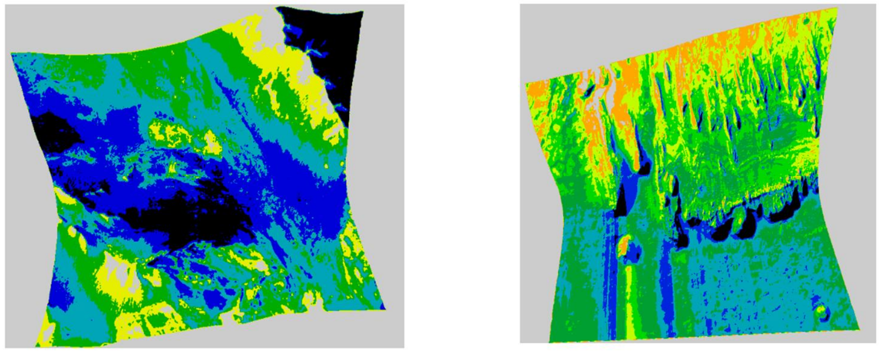

Figure 5. Cluster classification for (Obs #000132E6L) on the left and (Obs #00018285S) on the right.

5. CONCLUSIONS

Considering the sheer volume of data from planetary missions, unsupervised machine learning will be the only practical avenue in future space exploration in the identification of mineral classes. With the advent of software and algorithm design, the processing and visualization of hyperspectral data will help scientists to understand more about the geological past of Mars.

ACKNOWLEDGMENTS

I am hugely indebted to my family for their patience and support during my study. Special thanks also to Dr. Queenie Chan of NASA Johnson Space Center for her assistance in the hyperspectral data.

REFERENCES

[1] Pilorget, C. and Bibring, J.P., "NIR reflectance hyperspectral microscopy for planetary science: Application to the MicrOmega instrument", Planetary and Space Science 76,42-53 (2013)

[2] Izenberg, N.R., Klima, R.L., Murchie, S.L., Blewett, D.T., Holsclaw, G.M., McClintock, W.E., Malaret, E., Mauceri, C., Villas, F., Sprague, A.L., Helbert, J., Domingue, D.L., Head III, J.W., Goudge, T.A., Solomon, S.C., Hibbitts, C., Dyar and Dyar, M.D., "The low-iron, reduced surface of Mercury as seen in spectral reflectance by MESSENGER", Icarus 228, 364-374 (2014)

[3] Stephan, K., Hibbitts, C.A., Hoffmann, H. and Jaumann, R., "Reduction of instrument-dependent noise in hyperspectral image data using the principal component analysis: Applications to Galileo NIMS data", Planetary and Space Science 56, 406-419 (2008)

[4] Bell III, J. F., McSween Jr., Crisp, J.A., Morris, R.V., Murchie, S.L., Bridges, N.T., Johnson, J.R., Britt, D.T., Golombek, M.P., Moore, H.J., Ghosh, A., Bishop, J.L., Anderson, R.C., Brfickner, J., Economou, T., Greenwood, J.P., Gunnlaugsson, H.P., Hargraves, R.M., Hviid, S., Knudsen, J.M., Madsen, M.B., Reid, R., Rieder, R. and Soderblo, L., "Mineralogic and compositional properties of Martian soil and dust: Results from Mars Pathfinder", Journal of Geophysical Research 105, 1721-1755 (2000)

[5] Bandfield, J. L., "Global mineral distributions on Mars", Journal of Geophysical Research 107,5042 (2002)
[6] Pelkey, S.M., Mustard, J.F., Murchie, S., Clancy R.T., Wolff, M., Smith, M., Milliken, R., Bibring J.P., Gendrin, A., Poulet, F., Langevin Y. and Gondet, B., "CRISM multispectral summary products: Parameterizing mineral diversity on Mars from reflectance", Journal of Geophysical Research 112 (2007)
[7] Van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., Hecker, C.A., Bakker, W.H., Noomen, M.F., van der Meijde, M., Carranza, E.J.M., Boudewijn de Smeth, J. and Woldai T., "Multi- and hyperspectral geologic remote sensing: A review", Int. Journal of Applied Earth Observation and Geoinformation 14,112-128 (2012)
[8] Murchie, S., Arvidson, R., Bedini, P., Beisser, K., Bibring, J.-P., Bishop, J., Boldt, J., Cavender, P., Choo, T., Clancy, R. T., Darlington, E. H., Des Marais, D., Espiritu, R., Fort, D., Green, R., Guinness, E., Hayes, J., Hash, C., Heffernan, K., Hemmler, J., Heyler, G., Humm, D., Hutcheson, J., Izenberg, N., Lee, R., Lees, J., Lohr, D., Malaret, E., Martin, T., McGovern, J. A., McGuire, P., Morris, R., Mustard, J., Pelkey, S., Rhodes, E., Robinson, M., Roush, T., Schaefer, E., Seagrave, G., Seelos, F., Silverglate, P., Slavney, S., Smith, M., Shyong, W.-J., Strohbehn, K., Taylor, H., Thompson, P., Tossman, B., Wirzburger, M. and Wolff M., "Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on Mars Reconnaissance Orbiter (MRO)", J. Geophysics. Res., 112, E05S03 (2007)
[9] Carter, J., Poulet, F., Murchie, S. and Bibring, J.P., "Automated processing of planetary hyperspectral datasets for the extraction of weak mineral signatures and applications to CRISM observations of hydrated silicates on Mars", Planetary and Space Science 76,53-67 (2013)
[10] Cracknell, M.J. and Reading, A.M., "Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information", Computers & Geosciences 63,22-33 (2014)
[11] Baldridge, A. M., Hook, S.J., Grove, C.I. and Rivera, G., "The ASTER Spectral Library Version 2.0" (2008)
[12] Ivezić, Z., Connolly, A.J., VanderPlas, J.P. and Gray, A. [Statistics, Data Mining and Machine Learning in Astronomy], Princeton (2014)
[13] Brown, A.J. and Storrie-Lombardi, M., "MR PRISM: A Spectral Analysis tool for the CRISM," SPIE Optics and Photonics, 6309-28 (2006).
[14] Chang, C.I. [Hyperspectral Data Processing – Algorithm Design and Analysis], Wiley (2013)
[15] Murphy, K. [Machine Learning - A Probabilistic Perspective], MIT, Cambridge, Massachuett (2012)
[16] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., "Scikit-learn: Machine Learning in {P}ython", 12, 2825-2830 (2011)
[17] Christensen, P. R., Bandfield, J.L., Hamilton, V.E., Howard, D.A., Piatek, J.L., Ruff, S.W., Stefanov, W.L., "A thermal emission spectral library of rock-forming minerals", Journal of Geophysical Research 105, 9735-9739 (2000)