

SUPERPIXEL BASED FACTOR ANALYSIS AND TARGET TRANSFORMATION METHOD FOR MARTIAN MINERALS DETECTION

Wu Xing^{a,b}, Zhang Xia^{a*}, Lin Honglei^{a,b}

^a. Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, 100101, China
(wuxing, zhangxia, linhl)^a@radi.ac.cn

^b. University of Chinese Academy of Science, Beijing, 100049, China

Commission III, ICWG III/II

KEY WORDS: Superpixel, Factor Analysis and Target Transformation, Minerals detection, Mars

ABSTRACT:

The Factor analysis and target transformation (FATT) is an effective method to test for the presence of particular mineral on Martian surface. It has been used both in thermal infrared (Thermal Emission Spectrometer, TES) and near-infrared (Compact Reconnaissance Imaging Spectrometer for Mars, CRISM) hyperspectral data. FATT derived a set of orthogonal eigenvectors from a mixed system and typically selected first 10 eigenvectors to least square fit the library mineral spectra. However, minerals present only in a limited pixels will be ignored because its weak spectral features compared with full image signatures. Here, we proposed a superpixel based FATT method to detect the mineral distributions on Mars. The simple linear iterative clustering (SLIC) algorithm was used to partition the CRISM image into multiple connected image regions with spectral homogeneous to enhance the weak signatures by increasing their proportion in a mixed system. A least square fitting was used in target transformation and performed to each region iteratively. Finally, the distribution of the specific minerals in image was obtained, where fitting residual less than a threshold represent presence and otherwise absence. We validate our method by identifying carbonates in a well analysed CRISM image in Nili Fossae on Mars. Our experimental results indicate that the proposed method work well both in simulated and real data sets.

1. INTRODUCTION

A record of the evolution of Mars is preserved in the rocks and sediments exposed at its surface. Minerals can fingerprint many processes that build the Martian rock record (Ehlmann et al., 2014). Spectroscopy allows for the analysis of surface minerals via remote sensing observations (Thomas et al., 2017). As Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) had been sent to Mars, a huge quantity of data is getting available for research, which provide an improved understanding of Martian mineralogy. However, due to insufficient spatial resolution and spatial complexity, pixels in images are likely to be a mixture of pure spectral constituents rather than a single substance (Bioucas-Dias et al., 2012). Besides, owing to the instrumental or observational biases, which further complicate the extraction of interesting but subtle spectral features (such as hydrated silicates) (Carter et al., 2013).

There are many efforts have been made to identify minerals based on CRISM data. Spectral parameters provide an analysis tool for rapid assessment of the vast amounts of data (Pelkey et al., 2007, Viviano-Beck et al., 2014), which are widely used to identify a diverse range of minerals on Martian surface (Ehlmann et al., 2008, Mustard et al., 2008, Ehlmann et al., 2009). However, a spectral parameter may account for multiple minerals (For example, D2300 represents Fe/Mg phyllosilicates (Viviano-Beck et al., 2014)), it's still difficult to identify the unambiguous mineral. The Factor analysis and target transformation (FATT) is an effective method to test for the presence of particular mineral on Martian surface. It has been used both in thermal infrared (Thermal Emission Spectrometer, TES) (Bandfield et al., 2000) and near infrared (CRISM) (Thomas et al., 2017) hyperspectral data. Factor analysis derived a set of orthogonal eigenvectors from a mixed system and typically selected first 10 eigenvectors to least square fit the library mineral spectra. If the

trial spectrum fits well, this spectrum is a component of the system and a possible spectral endmember (Bandfield et al., 2000). However, there are some challenges in FATT. First of all, when an image is analysed using factor analysis, higher order eigenvectors (corresponding to noise in statistics) will be discarded. Thus, minerals present only in limited pixels might be ignored because of its weak spectral features compared with full image signatures. There is no systematic discussion on the effect of different number of eigenvectors on fitting accuracy. Furthermore, FATT can only tell the presence of specific mineral, but can't give the potential locations.

The purpose of this paper is to address the above problems. We first analysed the performance of different number of selected eigenvectors upon different signal to noise ratio (SNR) data. Besides, we proposed a novel superpixel based target transformation method for Martian minerals detection. The rest of this paper is organized as follows. Section 2 presents the proposed method. Section 3 and section 4 describe our experimental results with simulated datasets and CRISM data, respectively. The conclusions are drawn in section 5.

2. METHOD

There are two steps in FATT. First, R-mode factor analysis derives a set of orthogonal eigenvectors from mixed spectral data, and associated eigenvalues indicate the relative importance of the eigenvectors. Geminal et al. (2015) used Principal Component Analysis (PCA) to obtain eigenvectors. In this work, the eigenvalues and eigenvectors of covariance matrix of the mean-removed spectral data was calculated. Second, a linear least squares (lsq) fitting of the eigenvectors onto a test mineral spectrum is performed to determine whether it is one of the endmember in the scene. Then we analysed the fitting residuals as a function of the number of eigenvectors in different noise

level corrupted data. Next, we utilized a state-of-art superpixel algorithm called Simple Linear Iterative Clustering (SLIC)(Achanta et al., 2012) to partition the image into multiple spectral homogeneous connected image regions, this process can highlights weak signatures in a small region by elevating their pixels' ratio. There are two main parameters affect the results. One is (approximate) number of output segment, the other is compactness, which balances spectral proximity and space proximity, higher value give more weight to space proximity, and vice versa. We compared the fitting results of different segment parameter settings. Finally, we validate our method by identifying carbonates in a well analysed CRISM image in Nili Fossae on Mars.

3. EXPERIMENTS WITH SIMULATED DATA

In this section, we illustrate the fitting performance of the proposed method using two simulated hyperspectral data sets. Data sets 1 are used to study the fitting results of different numbers of eigenvectors upon different signal to noise ratio (SNR) data. Data sets 2 are used to study the performance of our proposed superpixel based target transformation method. For quantitative analysis, the root mean square error (rmse) is used to evaluate the fitting accuracy. Let y be the true spectrum with n bands, y_i be the model spectrum, rmse can be computed as follows:

$$rmse = \sqrt{\sum_{i=1}^n (y - y_i)^2 / n} \quad (1)$$

3.1 Simulated Data Sets 1

The spectral data we used in this experiment is provided by CRISM spectral library, where 4 endmember signatures are used to generate an 8×8 pixels synthetic data, including Hematite BKR1JB041, Clinopyroxene C1XP20, Orthopyroxene CBSB52, Olivine C1OL01. The endmembers cover the wavelength range of 1.03~2.60μm with 238 bands. In each pixel, the fractional abundances of the endmembers follow a dirichlet distribution and abundance nonnegativity constraint (ANC) and sum-to-one constraint (ASC). We generated the data according to the linear mixing model(Bioucas-Dias et al., 2012). Then we randomly implanted a serpentine spectrum LASR06 into the pixel as (2):

$$R = f \times ref_{serp} + (1 - f) \times ref \quad (2)$$

where ref_{serp} is reflectance of serpentine, ref is reflectance of original mixed pixel, f represents abundance, and $0 \leq f \leq 1$. The obtained data was then contaminated with i.i.d. Gaussian noise, for four levels of SNR: 20, 30, 40, 50dB.

Simulated Data Cube 1 (DC1): We randomly chose 10 pixels to insert serpentine spectrum, and the serpentine abundance is 10% in each pixel. Fig. 1 shows that rmse decreases rapidly as the number of eigenvectors K increase to 4 when SNR larger than 20dB, then decreases slowly. That's because the data is mainly mixed by 4 endmembers. Intuitively, fitting residuals decrease by an order of magnitude when SNR increase from 20dB to 30dB.

Simulated Data Cube 2 (DC2): We randomly chose 1 pixel to insert serpentine spectrum, and the serpentine abundance is 100%. From Fig. 2, we can get a conclusion similar to Fig.1. Besides, the rmse of DC2 decrease by an order of magnitude when

compared with DC1. Fig. 3 shows the spectrum fitting with first 4 eigenvectors of DC1 and DC2 when SNR is 30dB. From the above two experiments, we can draw conclusions: 1) The number of eigenvectors K can be determined according to intrinsic dimension of data when data quality is good. 2) Although the total abundance of serpentine is the same as DC1, rmse of DC2 is much lower than DC1, this suggests that the limited high abundance target pixel is insensitivity to noise and easier identified than distributed more widely but with low abundance pixel.

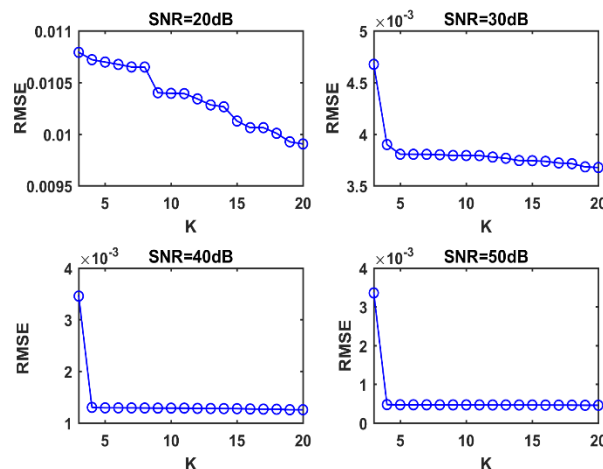


Figure 1. Plot of root mean square error of simulated data cube 1 as a function of the number of eigenvectors.

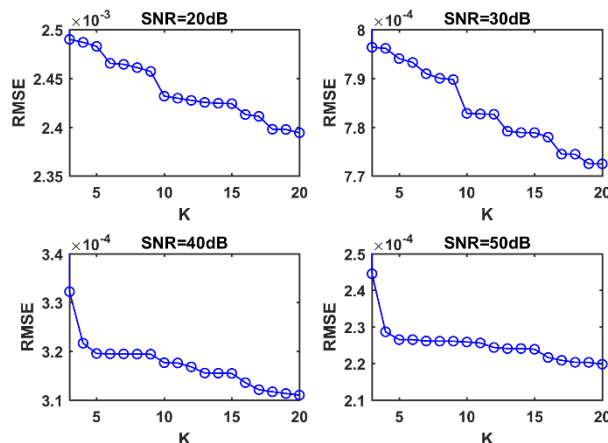


Figure 2. Plot of root mean square error of simulated data cube 2 as a function of the number of eigenvectors in 30dB.

3.2 Simulated Data Sets 2

The spectral data we used in this experiment is provided by the USGS spectral library, where 15 endmember signatures are used. The above 15 spectra are collected in 224 bands uniformly spanning from 0.4 to 2.5 μm. Simulated data 2 with 64×64 pixels was generated as(Zou et al., 2015), and then corrupted by a Gaussian white noise with 30dB SNR. In this experiment, Labradorite HS17.3B is used as the target spectrum. Fig. 4(a) shows the first band of simulated data 2 and Fig. 4(b) shows the groundtruth.

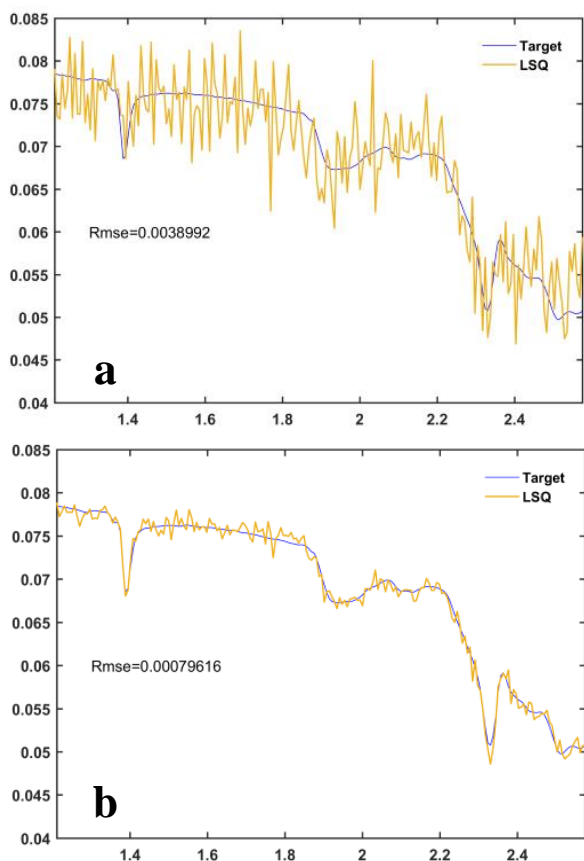


Figure 3. Spectrum reconstruction with first 4 eigenvectors of DC1(a) and DC2(b). Both data cube are contaminated with 30dB Gaussian noise.

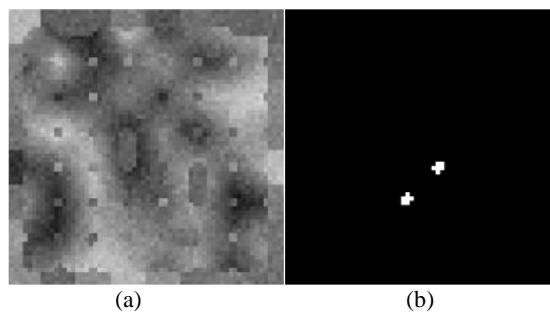


Figure 4. (a) First band of simulated data 2, (b) Groundtruth of targets.

We set a series of initial number of segments 20, 50, 80, 100, 150, 250, 300 to run SLIC algorithm, the results are shown in Fig. 5. We conducted experiments to analyse the effects of image segmentation on detection performance. We manually set K equals 6 in each segmentation result. Target transformation was then performed to each region iteratively. For better illustration, we used the reciprocal of rmse to display the final detection results. The smaller the rmse, the larger its reciprocal, and the more likely it's to be a target. Finally, we set a threshold to determine target pixels. Fig. 6 shows detection maps in different segmentation results. The target pixels are shown in white. From a qualitative point of view, our method reaches its best detection performance at Fig. 6(g).

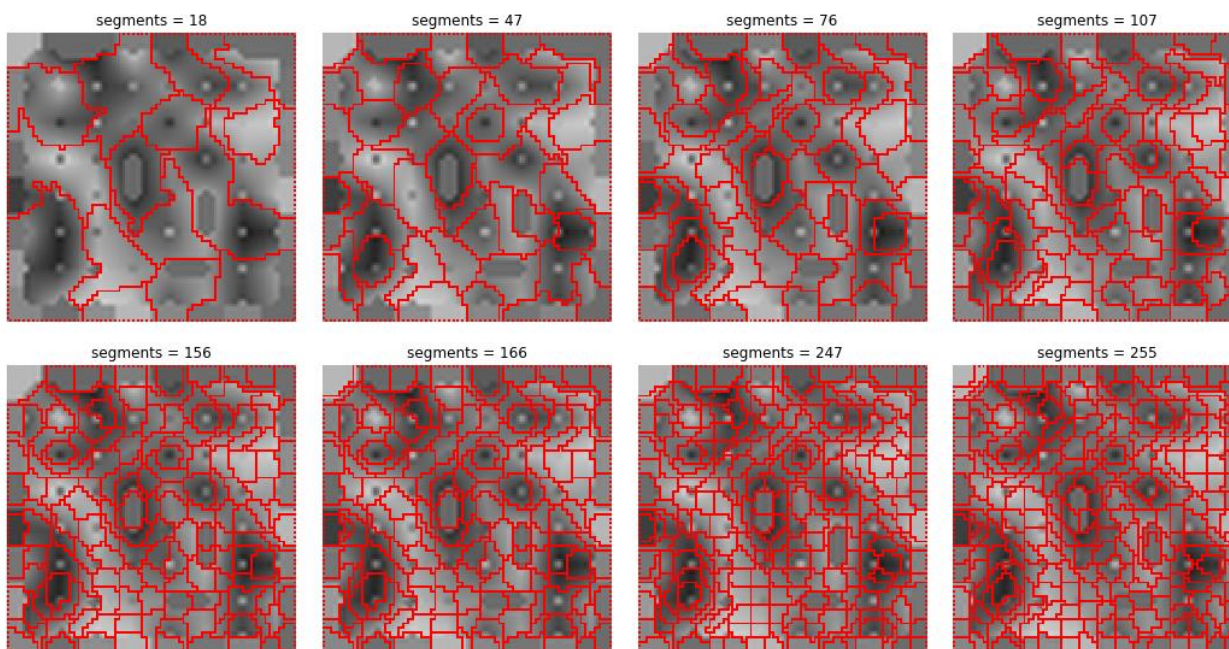


Figure 5. Superpixel segmentation results by SLIC with different input parameter (segments), the titles show the actual number of segments in each figure, the boundaries are represented by red lines.

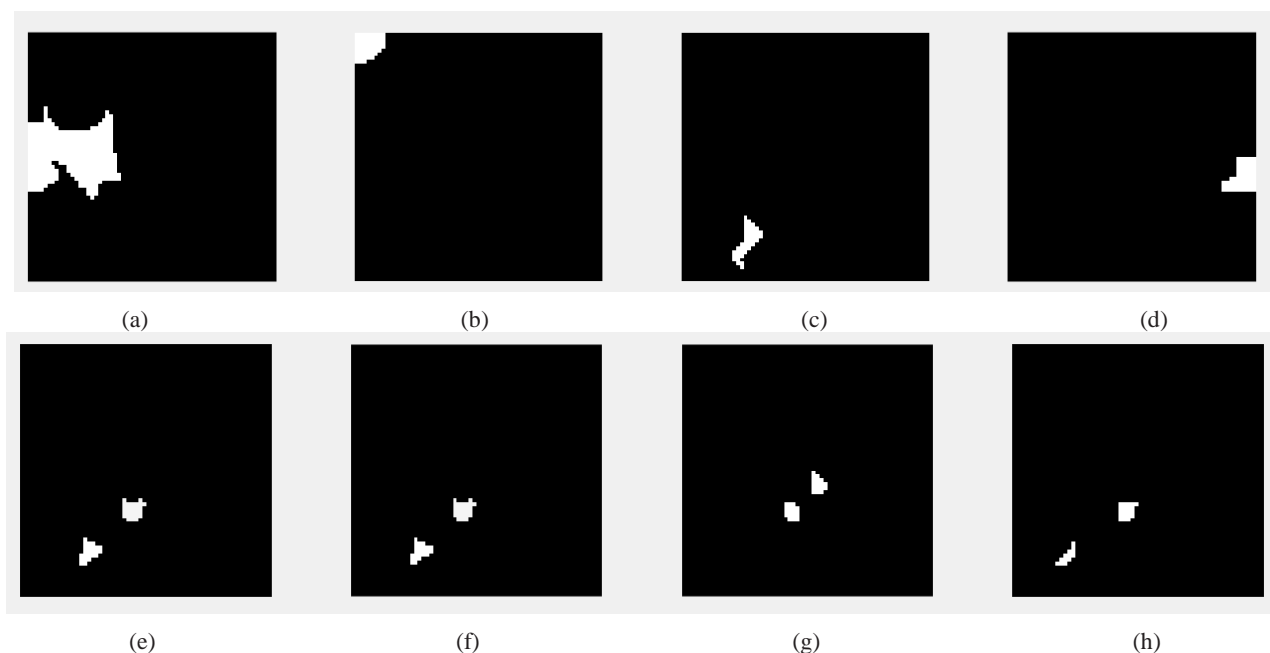


Figure 6. Detection maps of the simulated data 2 (with gaussian white noise 30dB SNR) in different segmentation results corresponding to Fig. 5.

The receiver-operating-characteristic (ROC) curve is employed to quantitatively evaluate the detection ability. ROC curve used a detection probability P_d and a false alarm rate P_f to provide an unbiased, quantitative, and threshold-free performance comparison. P_d and P_f are defined as:

$$P_d = \frac{N_d}{N_t} \quad (3)$$

$$P_f = \frac{N_{miss}}{N_{all}} \quad (4)$$

where N_d represents the number of detected target pixels at a certain threshold, N_t represents the number of target pixels in the image, N_{miss} represents the number of background pixels mistaken as targets, and N_{all} represents all the pixels in the image.

A good detection ROC curve should lie near to the top left. Fig. 7 shows the ROC curves of the detection results of above segmentation schemes, the curves prove that segment 7 is best, which is consistent with the conclusion in Fig.6.

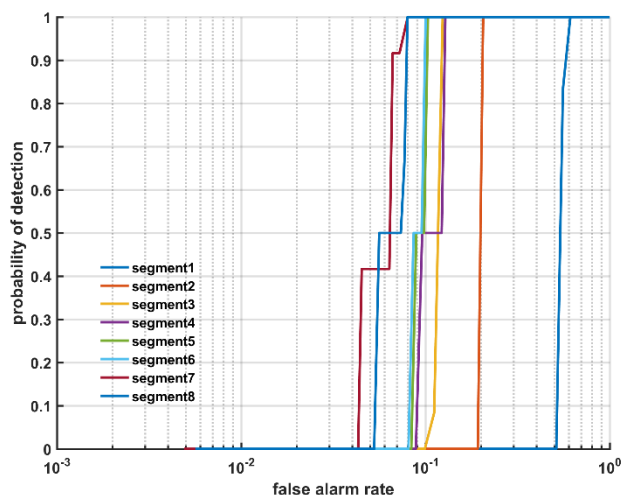


Figure 7 The ROC curves of our method with different segmentation results.

4. EXPERIMENTS WITH REAL DATA

Carbonates are key minerals for understanding ancient Martian environments because they are indicators of potentially habitable, and may be an important reservoir for paleoatmospheric CO₂(Wray et al., 2016). We conducted an experiment to identify carbonate (magnesite) in a well studied CRISM FRT00003e12 to validate our method.

CRISM is a VNIR imaging spectrometer onboard the Mars Reconnaissance Orbiter (MRO) that covers the wavelength range of 0.36~3.94 μ m. In this work, we used one targeted mode observation frt00003e12, which has a full spatial resolution (FRT) of 18 m/pixel(Murchie et al., 2007), and selected a range of 133 spectral bands from 1.7–2.6 μ m as(Thomas et al., 2017) suggested. All preprocessing were performed by CRISM Analysis Toolkit. We manually set a series of segments (500, 1000, 1500, 2000, 2500, 3000) to run SLIC algorithm, the results are shown in Fig. 8. FATT was then performed to fit magnesite spectrum CACB06 in each region iteratively. The number of

eigenvectors was determined according to the data's intrinsic dimension, which was estimated by HySime algorithm(Bioucas-Dias et al., 2008) Finally, we used a decision fusion strategy(Li et al., 2015), where a detection map is produced in each segmentation and the final distribution map is generated with a voting strategy.

Fig. 9 (a) shows CRISM spectral parameters (R(MIN2295_2480), G(MIN2345_2537), B(CINDEX)). Red/magenta colors indicate Mg carbonates, blue color indicates carbonates, while green/cyan

colors indicate Fe/Ca carbonates(Viviano-Beck et al., 2014). Fig. 9 (b). shows the final detection map. Our detection results are to a certain extent related to red and blue colors distribution in Fig. 9 (a), that's because we used magnesite as a target. Naturally, we didn't detect Fe/Ca carbonates (red/magenta colors in Fig. 9 (a)). The modeled spectrum together with magnesite spectrum are shown in Fig. 9 (c). However, there are still some discrepancies between our detection result and spectral parameters map, this will be solved in the future.

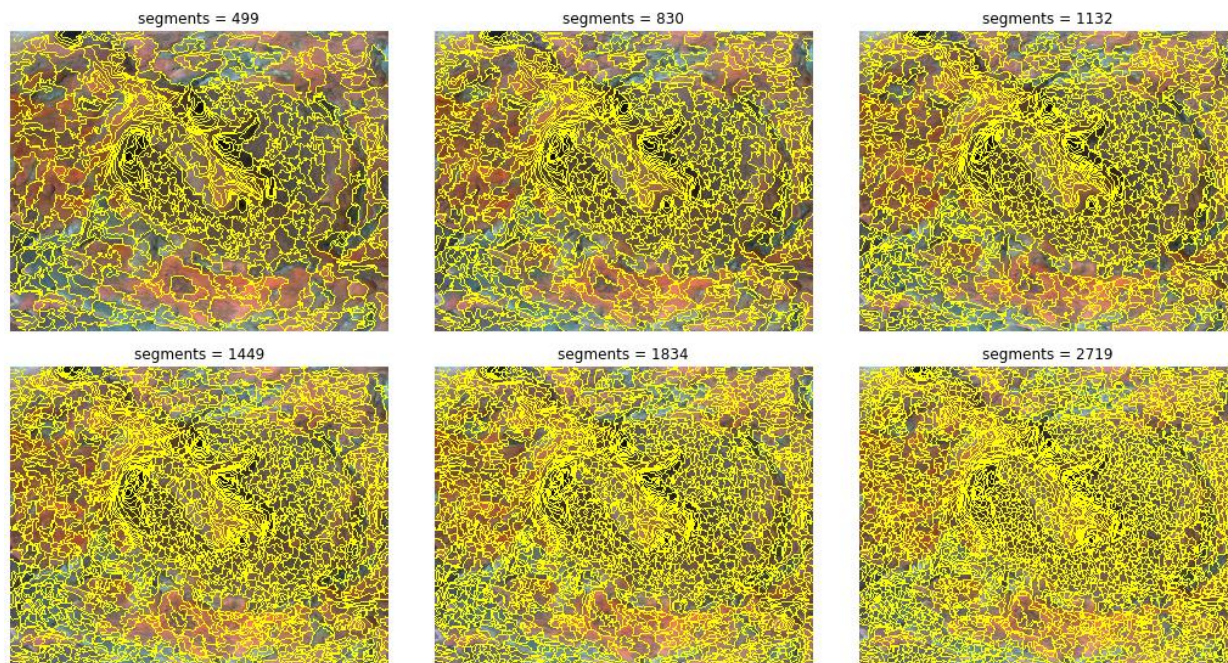


Figure 8. Superpixel segmentation results of FRT00003e12 (unprojected) by SLIC with different input parameter (segments), the titles show the actual number of segments in each figure. The base map is the default false color of 3e12 (R:2.529 μ m, G:1.506 μ m, B: 1.080 μ m), the boundaries are represented by yellow lines. The denser the boundaries, the finer the segmentation.

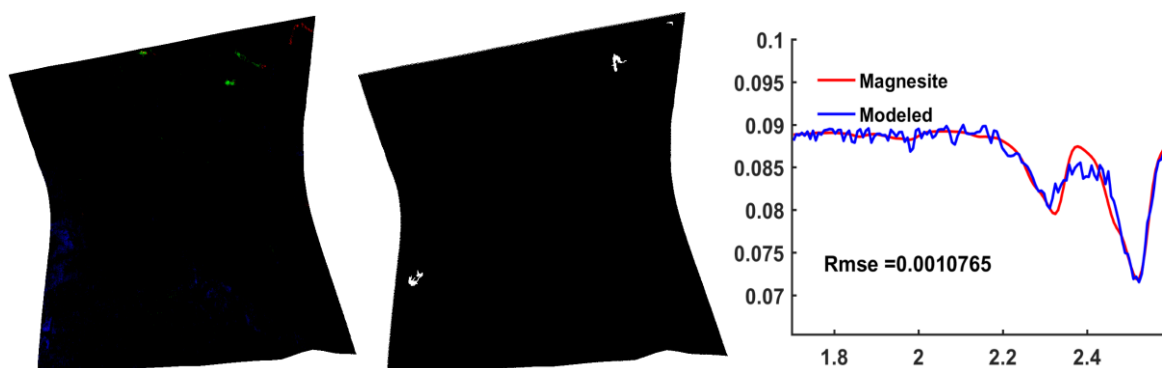


Figure 9. (a) CRISM spectral parameters map (R(MIN2295_2480), G(MIN2345_2537), B(CINDEX)), (b) Our detection map, (c) Modeled spectrum by FATT and magnesite spectrum

5. CONCLUSION AND FUTURE WORK

In this work, we first systematic investigated the performance of different number of eigenvectors upon different signal to noise ratio (SNR) data. Then we proposed a novel target transformation method, this method introduces the superpixel into the popular FATT technology. Finally, we used a decision fusion strategy to produce the final mineral map. Our experimental results, conducted using simulated data sets 1 indicate that the limited high abundance target pixel is insensitivity to noise and easier

identified than distributed more widely but with low abundance target pixel. What's more, both simulated data set 2 and real hyperspectral data CRISM, illustrate the proposed method can identify target minerals in the scene effectively. Although the detection results are encouraging, we have to take such problems into consideration: (1) how to set the threshold of fitting residual of different minerals adaptively, (2) the targets are narrowed to a small segments, how to determine its location in pixel scale.

ACKNOWLEDGEMENTS

The authors are very grateful for the financial support provided by the National Natural Science Foundation of China (grant no.41671360).

REFERENCES

- Achanta R, Shaji A, Smith K, Lucchi A, Fua P and Susstrunk S., 2012. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 34(11), pp. 2274-2282.
- Bandfield J L, Christensen P R and Smith M D., 2000. Spectral data set factor analysis and end-member recovery: Application to analysis of Martian atmospheric particulates. *Journal of Geophysical Research Planets*, 105(E4), pp. 9573–9587.
- Bioucas-Dias J M and Nascimento J M P., 2008. Hyperspectral Subspace Identification. *IEEE Transactions on Geoscience & Remote Sensing*, 46(8), pp. 2435-2445.
- Bioucas-Dias J M, Plaza A, Dobigeon N, Parente M, Du Q, Gader P and Chanussot J., 2012. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 5(2), pp. 354-379.
- Carter J, Poulet F, Murchie S and Bibring J P., 2013. Automated processing of planetary hyperspectral datasets for the extraction of weak mineral signatures and applications to CRISM observations of hydrated silicates on Mars. *Planetary & Space Science*, 76(1), pp. 53-67.
- Ehlmann B L and Edwards C S., 2014. Mineralogy of the Martian Surface. *Earth and Planetary Sciences*, 42(42), pp. 291-315.
- Ehlmann B L, Mustard J F, Murchie S L, Poulet F, Bishop J L, Brown A J, Calvin W M, Clark R N, Marais D J, Milliken R E, Roach L H, Roush T L, Swayze G A and Wray J J., 2008. Orbital identification of carbonate-bearing rocks on Mars. *Science*, 322(5909), pp. 1828-1832.
- Ehlmann B L, Mustard J F, Swayze G A, Clark R N, Bishop J L, Poulet F, Des Marais D J, Roach L H, Milliken R E, Wray J J, Barnouin-Jha O and Murchie S L., 2009. Identification of hydrated silicate minerals on Mars using MRO-CRISM: Geologic context near Nili Fossae and implications for aqueous alteration. *Journal of Geophysical Research*, 114(E2), pp. 538-549.
- Geminale A, Grassi D, Altieri F, Serventi G, Carli C, Carrozzo F G, Sgavetti M, Orosei R, D'Aversa E, Bellucci G and Frigeri A., 2015. Removal of atmospheric features in near infrared spectra by means of principal component analysis and target transformation on Mars: I. Method. *Icarus*, 253, pp. 51-65.
- Li W and Du Q., 2015. Decision fusion for dual-window-based hyperspectral anomaly detector. *Journal of Applied Remote Sensing*, 9(1), pp. 097297.
- Murchie S, Arvidson R, Bedini P, Beisser K, Bibring J P, Bishop J, Boldt J, Cavender P, Choo T and Clancy R T., 2007. Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on Mars Reconnaissance Orbiter (MRO). *Journal of Geophysical Research Atmospheres*, 112(E5), pp. 431-433.
- Mustard J F, Murchie S L, Pelkey S M, Ehlmann B L, Milliken R E, Grant J A, Bibring J P, Poulet F, Bishop J and Dobrea E N., 2008. Hydrated silicate minerals on Mars observed by the Mars Reconnaissance Orbiter CRISM instrument. *Nature*, 454(7202), pp. 305-309.
- 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., 2007. CRISM multispectral summary products: Parameterizing mineral diversity on Mars from reflectance. *Journal of Geophysical Research*, 112(E8), pp. 1-18
- Thomas N H and Bandfield J L., 2017. Identification and refinement of martian surface mineralogy using factor analysis and target transformation of near-infrared spectroscopic data. *Icarus*, 291, pp. 124-135.
- Viviano-Beck C E, Seelos F P, Murchie S L, Kahn E G, Seelos K D, Taylor H W, Taylor K, Ehlmann B L, Wiseman S M, Mustard J F and Morgan M F., 2014. Revised CRISM spectral parameters and summary products based on the currently detected mineral diversity on Mars. *Journal of Geophysical Research: Planets*, 119(6), pp. 1403-1431.
- Wray J J, Murchie S L, Bishop J L, Ehlmann B L, Milliken R E, Wilhelm M B, Seelos K D and Chojnacki M., 2016. Orbital evidence for more widespread carbonate-bearing rocks on Mars. *Journal of Geophysical Research: Planets*, 121(4), pp. 652-677.
- Zou Z and Shi Z., 2015. Hierarchical Suppression Method for Hyperspectral Target Detection. *IEEE Transactions on Geoscience & Remote Sensing*, 54(1), pp. 330-342.