Perspectives in Science (2016) 8, 722—726

ScienceDirect

journal homepage: www.elsevier.com/pisc

Performance evaluation of hyperspectral classification algorithms on AVIRIS mineral data☆

Ramesh Nityanand Adep*, Aswin P. Vijayan, Amba Shetty, H. Ramesh

Department of Applied Mechanics and Hydraulics National Institute of Technology Karnataka, Surathkal 575025, Karnataka, India

Received 20 February 2016; accepted 9 June 2016
Available online 5 July 2016

KEYWORDS
Spectral discriminatory probability;
Spectral discriminatory entropy;
Spectral discriminatory power;
Hyperspectral data

Summary Hyperspectral sensors enable the acquisition of data with increased number of spectral bands at a higher spectral resolution. Data acquired through such sensors has been widely utilized in many remote sensing applications including mineral mapping. Development of classification algorithms is also continued along with hyperspectral sensors. Choice of an algorithm is extremely important for proper classification. To the best of our knowledge, there exists limited knowledge on the relative performance of similarity measures on a mineral classification. This study uses three statistics namely spectral discriminatory probability (SDP), spectral discriminatory entropy (SDE) and spectral discriminatory power (SDPW) to assess the performance of various similarity measures. Similarity measures chosen are Spectral Angle Mapper (SAM), Spectral Information Divergence (SID), Jeffries—Matusita distance (JM) and their hybrid combinations of SID—SAM, SID—SCA, and JM—SAM. All the similarity measures and statistics were developed on MATLAB® platform and evaluated the same using freely available AVIRIS mineral data from U.S. Geological Survey spectral library. Analysis of statistical results collectively revealed that among the chosen algorithms SID—SAM and SID—SCA outperform the other similarity measures when tested on mineral data. This result has an important implication on choosing of appropriate similarity measure for mineral classification.

Introduction

Hyperspectral data has more than 50 bands with narrower bandwidth and continuous spectral information. The emergence of hyper spectral sensors enabled the acquisition of

☆ This article belongs to the special issue on Engineering and Material Sciences.

* Corresponding author. Tel.: +91 7276756774.
E-mail address: ramesh2civil@gmail.com (R.N. Adep).

http://dx.doi.org/10.1016/j.pisc.2016.06.070

data with increased number of spectral bands and higher spectral resolution has certainly given significant impacts on our ability to map. It has been widely utilized in many remote sensing applications such as mineral mapping, soil salinity mapping and forestry applications, etc. The challenge is to process the data. There have been several classification algorithms developed to process this data. Algorithms which measure similarity between targets are Euclidean distance (ED) (Gower, 1985), Spectral Angle Mapper (SAM) (Kruse et al., 1993), Spectral Information Divergence (SID) (Chang, 1999), Jeffries—Matusita distance (JM) (Richards, 1999) and Spectral Correlation Angle (SCA) (De Carvalho and Meneses, 2000). Some researchers tried to improve the accuracy of mapping by combing two similarity measures such as SID—SAM (Du et al., 2004), SID—SCA (Naresh Kumar et al., 2011) and JM—SAM (Padma and Sanjeevi, 2014). The details of all the algorithms are provided in their respective articles.

Hybrid similarity measures algorithms

Du et al. (2004) proposed SID—SAM which combines deterministic SAM and stochastic SID in two ways SID $\times$ *tan* (SAM) and SID $\times$ *sin* (SAM). To make two similar spectra signatures more similar, while two dissimilar spectral signatures to more distinct. Naresh Kumar et al. (2011) replaced SAM by SCA and proposed SID—SCA to measure spectral similarity for discrimination among Vigna genus species, which gives higher relative discriminatory power in the 400—700 nm spectral region when compared with SID—SAM. Padma and Sanjeevi (2014) proposed another combination of deterministic Spectral Angle Mapper and stochastic Jeffries—Matusita (JM) measure, where the qualitative distance measure (JM) was replaced by SID to increase spectral discrimination ability. The developed JM—SAM algorithm was implemented to discriminate the mangrove species and the land cover classes using the Hyperion image datasets. The objective of this study is to know the performance of all the similarity measures on AVIRIS mineral data and to select the suitable similarity measure for further analysis on the basis of three performance measures. It was found from the literature *tan* version of hybrid measures showing superior performance than *sin* version of hybrid measures. In this study only tan version of hybrid measures were considered.

Performance measures

To evaluate the performance of the classification algorithms which measure the similarity between targets, three statistics were introduced by (Chang, 1999) are spectral discriminatory power (SDPW), spectral discriminatory probability (SDP) and spectral discriminatory entropy (SDE).

SDPW measures the spectral discrimination power of the algorithm. It helps to find the effective similarity measure when two spectral similarity measures are given. It is designed based on the power of discriminating one pixel from another relative to a reference pixel. Let '$\boldsymbol{d}$' be the spectral signature of a reference spectrum and $\boldsymbol{s}_i$, $\boldsymbol{s}_j$ be the spectral signatures of two other similar spectra. Spectral discriminatory power is given by

$$\mathrm{PW}^m(\boldsymbol{s}_i, \boldsymbol{s}_j : \boldsymbol{d}) = \max\left\{\frac{m(\boldsymbol{d},\ \boldsymbol{s}_i)}{m(\boldsymbol{d},\ \boldsymbol{s}_j)}, \frac{m(\boldsymbol{d},\ \boldsymbol{s}_j)}{m(\boldsymbol{d},\ \boldsymbol{s}_i)}\right\} \quad (1)$$

The PW^m ($\boldsymbol{s}_i$, $\boldsymbol{s}_j$:$\boldsymbol{d}$) defined by Eq. (1) provides an index of spectral discrimination capability of a specific similarity measure $m(,)$ between any two spectral signature $\boldsymbol{s}_i$, $\boldsymbol{s}_j$ relative to '$\boldsymbol{d}$'. Higher the PW^m ($\boldsymbol{s}_i$, $\boldsymbol{s}_j$:$\boldsymbol{d}$) value, better the discriminatory power of measure.

SDP measures the probability of algorithm to identify the target from the database or image. Let $\{\boldsymbol{s}_j\}_{j=1}^{J}$ be $\boldsymbol{J}$ spectral signatures in the database (Δ), $\boldsymbol{t}$ be a target to be identified using Δ and $m(,)$ be any similarity measure value between two spectra. Definition of spectral discriminatory probabilities of all $\boldsymbol{s}_j$'s in Δ with respect to $\boldsymbol{t}$ is as follows:

$$P^m_{t,\Delta^i} = \frac{m(\boldsymbol{t},\ \boldsymbol{s}_j)}{\sum_{J}^{j=1} m(\boldsymbol{t},\ \boldsymbol{s}_j)} \quad \text{for } i = 1,2,\ \ldots,\ J, \quad (2)$$

where $\sum_{j=1}^{J} m(\boldsymbol{t},\ \boldsymbol{s}_j)$ is a normalization constant determined by $\boldsymbol{t}$ and Δ.

$P^m_{t,\Delta} = (P^m_{t,\Delta(1)}, P^m_{t,\Delta(2)}, \ldots, P^m_{t,\Delta(J)})^T$ is called the spectral discriminatory probability vector of Δ with respect to $\boldsymbol{t}$. Using the smallest spectral discriminatory probability value obtained from Eq. (2), one can identify the target $\boldsymbol{t}$ because the target and the selected one have the minimum spectral discrimination. SDE measures the uncertainty of identifying the target by the algorithm in the database or image. Since $P^m_{t,\Delta} = (P^m_{t,\Delta(1)}, P^m_{t,\Delta(2)}, \ldots, P^m_{t,\Delta(J)})^T$ given by Eq. (2) is the spectral discriminatory probability vector of $\boldsymbol{t}$ using a spectral library Δ. The spectral discriminatory entropy of Δ with respect to $\boldsymbol{t}$ is given by

$$H^m(t : \Delta) = -\sum_{j=1}^{J} P^m_{t,\Delta(j)} \times \log_2 P^m_{t,\Delta(j)} \quad (3)$$

Eq. (3) provides the uncertainty measure of identifying $\boldsymbol{t}$ using the spectral signature in Δ. A smaller entropy value indicates a better chance to identify $\boldsymbol{t}$.

Experimental data and implementation

Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) mineral spectra were used in this study, which is freely available in the U.S. Geological survey (USGS) website. AVIRIS data has 224 continuous spectral bands of wavelength range 0.38—2.5 μm with a spectral interval of approximately 10 nm. For using all three statistics to know the performance of similarity measures on AVIRIS data same experimental design was adopted as followed by Chang (1999). He has taken a total of five spectra in two groups based on spectral similarity. One group has three spectra and another group has two spectra. For this study similarly, five mineral spectra dickite, halloysite, muscovite, kaolinite and chlorite + muscovite were selected. From, Fig. 1 it can be seen that halloysite, kaolinite and dickite have similar spectral pattern and muscovite and chlorite + muscovite have similar spectral signatures. These five minerals have diagnostic characteristic absorption feature only in the vibrational region (2—2.5 μm), thereby all minerals were spectrally subset to 2—2.5 μm wavelength region (52 bands were

Table 1 Spectral similarity value among five signatures produced by the similarity measures.

	SAM	JM	SID	SCA	JM—SAM	SID—SAM	SID—SCA
T1—T2	0.126	0.224	0.023	0.303	0.029	0.003	0.007
T1—T3	0.156	0.519	0.037	0.392	0.082	0.006	0.015
T1—T4	0.077	0.135	0.010	0.176	0.010	0.001	0.002
T1—T5	0.141	0.989	0.031	0.436	0.141	0.004	0.014
T2—T3	0.080	0.575	0.012	0.334	0.046	0.001	0.004
T2—T4	0.066	0.099	0.006	0.209	0.007	0.000	0.001
T2—T5	0.111	1.113	0.020	0.338	0.125	0.002	0.007
T3—T4	0.114	0.569	0.021	0.393	0.065	0.002	0.009
T3—T5	0.108	1.294	0.019	0.260	0.141	0.002	0.005
T4—T5	0.122	1.042	0.023	0.409	0.127	0.003	0.010

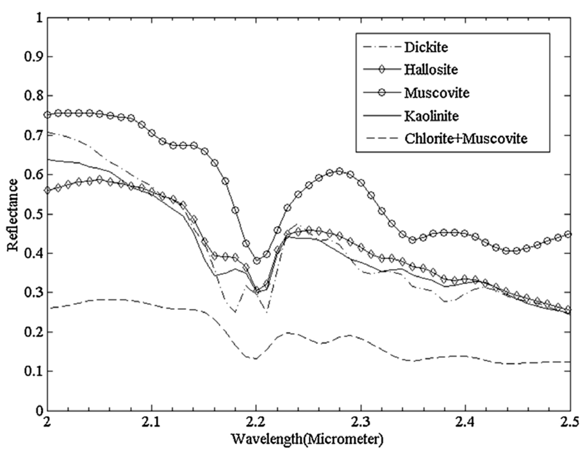

Figure 1 Reflectance spectra of minerals from AVIRIS dataset.

selected). All the similarity measures and statistics were implemented in MATLAB® platform version 2007.

Results and discussion

In this study performance of SAM, SCA, SID, JM and hybrid measures of SID × *tan* (SAM), SID × *tan* (SCA) and JM × *sin* (SAM) were compared through three statistical measures. For the sake of simplicity and ease of understanding in further discussion, minerals dickite, halloysite, muscovite, kaolinite and chlorite + muscovite are represented as T1, T2, T3, T4 and T5 respectively. All similarity measures SID × *tan* (SAM), SID × *tan* (SCA) and JM × *tan* (SAM) are also represented as SID—SAM, SID—SCA and JM—SAM respectively. Table 1 shows the spectral similarity values among five signatures obtained by different similarity measures.These similarity values cannot be used to compare their performance. In order to see which similarity measure is effective in terms of spectral discrimination power, in the first trial, three similar spectra of minerals such as dickite as '*d*', halloysite as 'S_i' and kaolinite as 'S_j' were chosen. The computed spectral discriminatory power of all the algorithms using Eq. (1) are shown in Table 2. From Table 2, in first the trial, it can be inferred that the performance of hybrid measures SID—SCA (3.99) followed by SID—SAM (3.75) was better as compared to other similarity measures. Spectral discrimination power of SID—SCA to distinguish dickite from halloysite is approximately 4 times better than to distinguish dickite from kaolinite. But in trial 2 and trial 3 performance of SID—SAM was better than SID—SCA. On the basis of spectral discrimination power statistics, one can say in any case hybrid measures have more discrimination than non—hybrid measure. Among hybrid measures, SID—SAM has better discrimination power for discriminating similar minerals.

In order to compute the statistic (SDP), a mixed spectral signature was generated for use as a target signature (*t*). Two trials were conducted, in the first trial (Table 3) target signature was taken 0.75 of muscovite and remaining 0.25 of other four spectra in equal proportion. Similarly, for second trial (Table 4) target signature was taken 0.75 of kaolinite and remaining 0.25 of other four spectra in equal proportion. From Tables 3 and 4, it can be seen that, in both trials, SID—SCA and SID—SAM have same and lowest spectral discriminatory probability value.

Thereby SID—SCA and SID—SAM can identify the target better than other similarity measures. Entropy values from Table 3 and Table 4 suggests that JM—SAM, SID—SAM and SID—SCA can have less uncertainty in identifying the target than other measures. Non-hybrid measure SID has low entropy value than SAM, JM and SCA measures.

As we can see from Fig. 1, the spectrum of muscovite and chlorite + muscovite are very similar. From Table 3, the ratio

Table 2 Spectral discriminatory power produced by the similarity measures.

Trial No.		SAM	JM	SID	SCA	JM-SAM	SID-SAM	SID-SCA
		(SDPW values)						
1	**T1-d, T2-s_i and T4-s_j**	1.65	1.66	2.26	1.72	2.74	3.75	3.99
2	**T2-d, T1-s_i and T4-s_j**	1.90	2.27	4.16	1.45	4.33	7.95	6.14
3	**T4-*d* T1-s_i and T2-s_j**	1.15	1.37	1.84	1.19	1.58	2.12	1.54

Table 3 Spectral discriminatory probability vectors produced by the measures with ***t*** chosen to be a mixture of 0.75 T3, 0.0625 T1, 0.0625 T2, 0.0625 T4 and 0.0625.

	SAM	JM	SID	SCA	JM-SAM	SID-SAM	SID-SCA
T1-*t*	0.333	0.132	0.429	0.269	0.204	0.535	0.473
T2- *t*	0.162	0.128	0.120	0.226	0.095	0.072	0.109
T3- *t*	0.035	0.113	0.005	0.045	0.018	0.001	0.001
T4- *t*	0.234	0.134	0.219	0.267	0.144	0.191	0.239
T5- *t*	0.236	0.494	0.228	0.194	0.538	0.201	0.177
Entropy	2.10	2.01	1.90	2.16	1.78	1.69	1.81

Table 4 Spectral discriminatory probability vectors produced by the measures with ***t*** chosen to be a mixture of 0.75 T4, 0.0625 T1, 0.0625 T2, 0.0625 T3 and 0.0625.

	SAM	JM	SID	SCA	JM-SAM	SID-SAM	SID-SCA
T1- *t*	0.214	0.098	0.201	0.147	0.079	0.160	0.104
T2- *t*	0.161	0.072	0.085	0.166	0.044	0.051	0.050
T3- *t*	0.283	0.313	0.326	0.316	0.333	0.346	0.375
T4- *t*	0.033	0.053	0.005	0.036	0.007	0.001	0.001
T5- *t*	0.309	0.464	0.384	0.334	0.538	0.443	0.470
Entropy	2.10	1.86	1.86	2.06	1.55	1.70	1.61

Table 5 A number of times effective in identifying ***t*** as a muscovite (target) than the other spectrum by the measures.

	SAM	JM	SID	SCA	JM-SAM	SID-SAM	SID-SCA
T1-*t* / T3-*t*	9.5	1.2	85.8	6.0	11.3	535.0	473
T2-*t* / T3- *t*	4.6	1.1	54.0	5.0	5.3	72.0	109
T4-*t* / T3-*t*	6.7	1.2	43.8	5.9	8.0	191.0	39
T5-*t* / T3-*t*	6.7	4.4	45.6	4.3	29.9	201.0	177

Table 6 A number of times effective in identifying ***t*** as a kaolinite (target) than the other spectrum by the measures.

	SAM	JM	SID	SCA	JM-SAM	SID-SAM	SID-SCA
T1-*t* / T4-*t*	6.5	1.8	40.2	4.1	11.3	160.0	104.0
T2-*t* / T4-*t*	1.9	1.4	17.0	4.6	6.3	51.0	50.0
T3-*t* / T4-*t*	8.6	5.9	65.2	8.8	47.6	345.0	375.0
T5-*t* / T4-*t*	9.4	8.8	7.8	9.3	76.9	443.0	470.0

of SID—SAM between ***t*** and muscovite to SID—SAM between ***t*** and chlorite + muscovite is 0.201:0.01 ≈ 201. In comparison, SID—SCA yielded 0.177:0.01 ≈ 177. This infers, that SID—SAM nearly 1.14 times more effective in identifying the ***t*** as muscovite than SID—SCA. In the second trial (Table 4), SID—SAM identified approximately (0.160/0.01) 160 times more effectively target ***t*** as kaolinite than dickite. It can also be deduced from Tables 5 and 6 that spectral discrimination capability of hybrid measures SID—SAM and SID—SCA was far better than other measures. In case of non-hybrid measures the performance of SID is better than the others. From this, it can be concluded that hybrid measures have less uncertainty in identifying the target and have higher discrimination capability of target from other spectra.

Conclusion

To know the performance of various similarity measures on AVIRIS mineral data, this study used three statistics. Which provided insight about measures without performing traditional classification process. On the basis of spectral discrimination power statistics, one can say in any case hybrid measures have more discrimination power than non-hybrid measure. Among hybrid measures, SID—SAM has better discrimination power for discriminating similar minerals. From SDE statistical results it was concluded that SID—SAM, JM—SAM and SID—SCA hybrid measures have less uncertainty in identifying the target. Also from SDP statistical results, SID—SAM and SID—SCA have higher discrimination capability of target from other spectra. In case of non-hybrid similarity measure SID has shown superior performance than the other measures in all the three statistics. Analysis of statistical results collectively inferred that among the chosen algorithms SID—SAM and SID—SCA outperform the other similarity measures. This result has an important implication on choosing of an appropriate similarity measure for mineral classification.

References

Chang, C.I., 1999. Spectral information divergence for hyperspectral image analysis. In: Geoscience and Remote Sensing

Symposium, IGARSS'99 Proceedings. IEEE 1999 International, vol. 1, pp. 509—511.

De Carvalho, O.A., Meneses, P.R.,2000. Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM). In: Summaries of the 9th JPL Airborne Earth Science Workshop. JPL Publication. 9, no. 00-18, Pasadena, CA.

Du, Y., Chang, C.I., Ren, H., Chang, C.C., Jensen, J.O., 2004. New hyperspectral discrimination measure for spectral characterization. Opt. Eng. 43 (8), 1777—1786.

Gower, J.C., 1985. Properties of Euclidean and non-Euclidean distance matrices. Linear Algebra Appl. 67, 81—97.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote Sens. Environ. 44 (2), 145—163.

Naresh Kumar, M., Seshasai, M.V.R., Vara Prasad, K.S., Kamala, V., Ramana, K.V., Dwivedi, R.S., Roy, P.S., 2011. A new hybrid spectral similarity measure for discrimination among Vigna species. Int. J. Remote Sens. 32 (14), 4041—4053.

Padma, S., Sanjeevi, S., 2014. Jeffries—Matusita-Spectral Angle Mapper (JM—SAM) spectral matching for species level mapping at Bhitarkanika, Muthupet and Pichavaram mangroves. Int. Archiv. Photogramm. Remote Sens. Spat. Inf. Sci. 40 (8), 1403.

Richards, J.A., 1999. Remote Sensing Digital Image Analysis, vol. 3. Springer, Berlin.