ISPRS Journal of Photogrammetry and Remote Sensing 124 (2017) 106–118

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

EXhype: A tool for mineral classification using hyperspectral data

Ramesh Nityanand Adep *[,1], Amba shetty, H. Ramesh

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

ARTICLE INFO

Article history:
Received 10 February 2016
Received in revised form 10 July 2016
Accepted 27 December 2016

Keywords:
Artificial neural network
Expert system
Human expertise
Knowledge base
Segmented upper hull method
Spectral angle mapper

ABSTRACT

Various supervised classification algorithms have been developed to classify earth surface features using hyperspectral data. Each algorithm is modelled based on different human expertises. However, the performance of conventional algorithms is not satisfactory to map especially the minerals in view of their typical spectral responses. This study introduces a new expert system named 'EXhype (Expert system for hyperspectral data classification)' to map minerals. The system incorporates human expertise at several stages of it's implementation: (i) to deal with intra-class variation; (ii) to identify absorption features; (iii) to discriminate spectra by considering absorption features, non-absorption features and by full spectra comparison; and (iv) finally takes a decision based on learning and by emphasizing most important features. It is developed using a knowledge base consisting of an Optimal Spectral Library, Segmented Upper Hull method, Spectral Angle Mapper (SAM) and Artificial Neural Network. The performance of the EXhype is compared with a traditional, most commonly used SAM algorithm using Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data acquired over Cuprite, Nevada, USA. A virtual verification method is used to collect samples information for accuracy assessment. Further, a modified accuracy assessment method is used to get a real users accuracies in cases where only limited or desired classes are considered for classification. With the modified accuracy assessment method, SAM and EXhype yields an overall accuracy of 60.35% and 90.75% and the kappa coefficient of 0.51 and 0.89 respectively. It was also found that the virtual verification method allows to use most desired stratified random sampling method and eliminates all the difficulties associated with it. The experimental results show that EXhype is not only producing better accuracy compared to traditional SAM but, can also rightly classify the minerals. It is proficient in avoiding misclassification between target classes when applied on minerals.

1. Introduction

The emergence of hyperspectral sensors enabled the acquisition of data with increased number of spectral bands and higher spectral resolution has certainly given significant impacts on our ability to map. High spectral information i.e. narrower bandwidth and contiguous spectral information facilitate to distinguish even the spectrally similar features and enabled to identify sub-class features such as different types of vegetation, mineral, soil and water, etc. To identify minerals, mostly absorption features are considered as a diagnostic characteristic. Absorption features indicate the presence of a particular mineral and the depth of absorption is a function of its grain size, absorption coefficient, and abundance (Clark et al., 2003). Available hyperspectral sensors can capture Electromagnetic Radiation in the region of visible to shortwave infrared. Thereby, in the reflectance curve, only two types of absorption features are visible electronic and vibrational absorption features (Clark, 1999).

Advances in data capturing have posed a challenge to researchers to classify the data more accurately. Advancement in image classification techniques has been happening in three different directions such as supervised, unsupervised and hybrid (i.e. combination of supervised and unsupervised). Recent years have witnessed an extensive use of supervised classification algorithms to classify minerals. Classification algorithms were improved by researchers from traditional algorithms such as Spectral Angle Mapper (SAM) (Kruse et al., 1993a); Spectral Information Divergence (SID) (Chang, 1999); Artificial Neural Network (ANN) (Paya et al., 1997) and Support Vector Machines (SVM) (Melgani and Bruzzone, 2004), etc., to hybrid combinations such as hybrid

* Corresponding author.
E-mail addresses: ramesh2civil@gmail.com (R.N. Adep), amba_shetty@yahoo.co.in (A. shetty), ramesh.hgowda@gmail.com (H. Ramesh).
[1] Residential address: 276/1, Padmanagar, Bhiwandi, Dist: Thane, Maharashtra 421305, India.

http://dx.doi.org/10.1016/j.isprsjprs.2016.12.012
0924-2716/

similarity measures (Du et al., 2004; Naresh Kumar et al., 2011; Padma and Sanjeevi, 2014); Boosted Genetic Fuzzy Classifier (BGFC) (Stavrakoudis et al., 2011); Extended Spectral Angle Mapper (ESAM) (Li et al., 2014) and Support Vector Neural Networks (SVNN) (Lokman and Yilmaz, 2015), and went up to advance absorption feature oriented (Clark et al., 2003; Koerting et al., 2015; Mielke et al., 2016); spectral unmixing techniques (Tits et al., 2012; Zhu et al., 2014; Zhong et al., 2016) and machine learning concepts (Chen et al., 2014, 2015; Romero et al., 2016).

A close look into the developments has revealed that algorithms were incorporated with domain experts knowledge. The best examples are study of vegetation by indices (Sakamoto et al., 2011; Yu et al., 2014; Marshall and Thenkabail, 2015), feature matching techniques for mineral identification by Clark et al. (1990, 2003), Koerting et al. (2015), Mielke et al. (2016) and Optimal Spectral Library to deal with intra-class variation by Luc et al. (2005). In all the three cases, experts understood the behaviour of reflectance curve to their respective objective and then modelled the algorithms to classify features. It was also observed that researchers not only extracting more information from the reflectance curve but also from neighbourhood pixels to classify pixels accurately (Segl et al., 2003; Liu et al., 2015). Developed expert knowledge is translated into a computer-usable format and stored in the knowledge base (Huang and Jensen, 1997). Using such knowledge base, expert systems have been developed from time to time (Kruse et al., 1993b; Clark et al., 2003; Koerting et al., 2015; Mielke et al., 2016; Brossard et al., 2016). These expert systems have shown superior performance than the conventional algorithms. But with advancement in data i.e., in terms of spectral and spatial resolution, user's desire more accuracy in mapping. Also, it has seen that expert knowledge increased with time. The present study is mainly focused on to develop a new expert system by incorporating available expert knowledge to classify minerals.

2. Procuring of knowledge for proposed Expert System

This section presents the review of various supervised classification algorithms to identify the qualities for proposed expert system and to construct a knowledge base for it.

Most commonly used supervised classification algorithms for classification of hyperspectral data which measure similarity between spectra are Euclidean Distance (ED) (Gower, 1985), Spectral Angle Mapper (SAM) (Kruse et al., 1993a), Mahalanobis Distance (MD) (Palacios-Orueta and Ustin, 1996), Spectral Information Divergence (SID) (Chang, 1999), Jeffries-Matusita Distance (JMD) (Richards and Richards, 1999), Spectral Correlation Angle (SCA) (De Carvalho and Meneses, 2000), etc. and their hybrid combinations SAM-SID (Du et al., 2003), SCA-SID (Naresh Kumar et al., 2011) and JMD-SAM (Padma and Sanjeevi, 2014). Each algorithm employs a different approach to discriminate the target and image spectra. For instance, ED computes the distance between two spectra, MD computes the statistical distance between a reference spectral vector and multivariate distribution of points, SAM treats spectra as two vectors and measures the angle between them and the SID measures the probability of spectral discrepancy (Shanmugam and SrinivasaPerumal, 2014). The limitation of these algorithms is that only one discrimination value is obtained, which is an average fit over the entire spectral range or a subset of the dataset that is used in classification. This single discrimination value may not be sufficient to discriminate targets especially when there is a subtle difference between their spectra.

For correct identification of the target, one should consider each salient feature of the spectrum in addition to overall spectrum. The salient feature can be an absorption feature and/or any other trend of the spectrum. Spectral Feature Fitting (SFF) proposed by Clark et al. (1990) considers only absorption features. To discriminate spectra, depth and shape of the absorption feature in target spectrum and image spectrum are compared. Only one absorption feature is considered and the user has to give a range of wavelengths within which a unique absorption feature exists for the chosen target spectrum. This technique was later modified as Multi-Range Spectral Feature Fitting (MRSFF) (Clark et al., 2003; Pan et al., 2013), where the absorption features at various wavelength ranges are considered for matching. Optional weights to each spectral range were also given to emphasize more important features. Both SFF and MRSFF use continuum removal method to separate absorption features from entire spectra and only absorption features were taken into account to find the discrimination between the target spectrum and image spectrum.

He and He (2011) introduced Weight Spectral Angle Mapper (WSAM) to improve the differentiability between similar minerals by setting a weight in the 'difference range' to reduce the similarity and to increase the discriminability. This method uses both absorption features and non-absorption features to compare spectra, but the adjustment of weights should be done manually.

The reflectance spectra of any land cover features vary within the class due to factors such as weather conditions, bidirectional reflectance distribution function effects, soil conditions, shadows, and phenological stage (Luc et al., 2005). Fig. 1 shows the variation of alunite mineral spectra in the Cuprite region. Even though such a variation exists in alunite, one can easily recognize all spectra as alunite. This variation of spectra is called intra-class variability (Luc et al., 2005). An Optimized Spectral Angle Mapper (OSAM) was proposed (Luc et al., 2005) to capture intra-class variability. Usually, a reference or target spectrum is taken as an average spectrum of each Region Of Interest (ROI), This implies a reference spectrum is unable to represent the spectral variability present within each ROI. OSAM contains an Optimal Spectral Library (OSL) which preserves the spectral variability present within each ROI. This library is called optimal because it contains all the spectra that can classify all pixels of a certain class correctly. All the image pixel spectra are classified using the reference spectra stored in the optimal spectral library to avoid misclassification due to intra-class variation. Finally, it has the same limitation as the traditional algorithms, being only one overall discrimination value.

Human ability to think, remember, and solve problems inspired many researchers to develop artificial models whose architecture are based on the way that the human brain performs computations (Hagan et al., 1996). A sub-class of artificial models called Artificial Neural Networks (ANN) were developed, which were simplified based on the biological learning process of the human brain (Paya et al., 1997). Many authors have used ANN to classify satellite images. Applications and limitations of ANN in remote sensing context were explained in several review papers such as (Paola and Schowengerdt, 1995; Kanellopoulos and Wilkinson, 1997; Kavzoglu and Mather, 2003; Mas and Flores, 2008). Use of ANN is not appreciable because most of the time bands were given as an input which increases the computational load and complexity in recognizing the pattern, in turn, reduces the performance. Present scenario has changed, a pre-processed information was fed to the neural network to classify satellite images (Chen et al., 2014, 2015; Romero et al., 2016).

The literature review has revealed that there is not a single algorithm:

1. To deal with intra-class variation;
2. To identify absorption features;
3. To discriminate spectra by considering absorption features as well as non-absorption features and also by full spectra comparison;
4. And finally to take a decision based upon the learning and by emphasizing most important features.