Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 100 (2015) 460 – 465

Procedia Engineering

www.elsevier.com/locate/procedia

25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014

A Virtual Dimensionality Method for Hyperspectral Imagery

Daniela Baran*, Nicolae Apostolescu

National Institute for Aertospace Research "Elie Carafoli", 220 Iuliu Maniu, Bucharest, Bucharest 061126, Romania

---

Abstract

Hyperspectral sensors capture images in hundreds of narrow spectral channels. The spectral signatures for each spatial location provide rich information about an image scene, leading to better "separation" between physical materials and objects. Hyperspectral data are spectrally overestimated and the useful signals usually occupy lower dimensional subspace which needs to be inferred. The signal information is usually concentrated in lower dimensional subspaces. For estimating the number of spectrally distinct signatures present in a hyperspectral data the concept of Virtual Dimensionality (VD), give us the minimum number of spectrally distinct signal sources that characterize this data from a perspective view of target detection and classification.

Considering these facts it is important to reduce the volume of data with minimum loss of information and this is the main idea of our algorithm. In this paper, a new VD modified method for estimating the number of spectral distinct signatures has been proposed. To demonstrate the applicability of the developed software tools, we used a few well-known hyperspectral image-data sets.

Peer-review under responsibility of DAAAM International Vienna

Keywords: hyperspectral images; virtual dimension; spectral signatures; endmembers

---

1. Introduction

Hyperspectral sensors used in Hyperspectral Imagery collect information of earth surfaces as a set of images that correspond to the same spatial scene, but are acquired at many different spectral bands with high resolution. These images contain abundant spatial, spectral, and radiometric information, which makes earth observation and information acquisition much more efficient for material applications. In terms of spectral properties, the high resolution has the capability of uncovering unknown sources, which cannot be identified by visual inspection.

---

* Corresponding author. Tel.: 40-21-4340083; fax: 40-21-4340082.
E-mail address: baran.daniela@incas.ro

Peer-review under responsibility of DAAAM International Vienna
doi:10.1016/j.proeng.2015.01.391

The large dimensionality of a hyperspectral dataset often requires data transformation which can effectively reduce noise in data sets with minimum loss of information. Several methods have been implemented for determining of pure substances from images. This paper refers to estimating the minimum number of spectrally distinct signal sources that characterize the hyperspectral data. This is an important step in the classification process. Also, it is used to highlight spectral properties of many different types of data. Different transformations are intended to find the minimum number of parameters required to represent the observed properties of the data. This number is so-called "intrinsic dimension" estimation (ID) according to the definition given in Fukunaga [3] and[10]. Other authors define their own intrinsic dimension as "the smallest number of parameters needed to contain all of the variability in the data through a mapping function" or the dimension of signal subspace [8]. Therefore, because the pixels in an image can be considered vectors, the intrinsic dimension of a random vector X is usually defined as the number of "independent" parameters needed to represent X. A general approach to ID estimation is the linear technique PCA that use the eigenvalue distribution, but it can be difficult to implement if it is applied to hyperspectral imagery. Chein-I Chang appreciate that ID may not be suitable for use with applications and image data characterized by significant spatial correlations (like hyperspectral data). Reference [9] provides a summary of methods to determine the intrinsic dimension and number of endmembers of a hyperspectral image.

In the hyperspectral image, reflectance information depends only of the materials spectral responses in the scene. A mixed pixel is either linear or nonlinear combination of pure pixels signatures weighted by the correspondent abundance fraction. Many techniques of unmixing in hyperspectral image analysis are based on geometric approach where each pixel is seen as a spectral vector of p (number of spectral bands). The linear model is assuming that the number of substances and their spectra are known, but in reality these are not known and, then, hyperspectral unmixing falls into the blindly classes. When the mixture between materials is macroscopic, the linear mixing model of spectra is generally admitted because this model assumes no interaction between materials. Spectral unmixing is one of the most important techniques for analyzing hyperspectral images, it consists in endmember extraction. Today, the researchers are concerned to answer the questions: How many spectral signatures are required to unmix data? How many pure spectral signatures, referred to as endmembers, are supposed to be present in the data to be procesed? It is common practice to assume that the number of signatures used for spectral unmixing is the same number of endmembers, contradicted by reality.

For estimating the number of spectrally distinct signatures present in a hyperspectral data was introduced the concept of Virtual Dimensionality (VD), which is the "minimum number of spectrally distinct signal sources that characterize the hyperspectral data from a perspective view of target detection and classification" [2]. Since a spectrally distinct signature is determined by different applications such as endmember extraction, anomaly detection, etc., the VD also varies with how a spectrally distinct signature is interpreted. VD provides an effective alternative because provides a good estimate of the number of dimensions after dimensionality reduction, therefore it can be used to predict how many spectrally distinct signal sources are present in the data. Also, the concept of VD helps us to know that how many bands need to be clustered.

Harsanyi, Farrand, and Chang developed a new method, referred to as HFC method, to VD estimation or to determine the number of spectral endmembers in Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data.

The HFC method, derived from the concept of the Neyman–Pearson detection theory, is based on the observed data properties specified by signatures that can be discriminated spectrally band-by-band. HFC implements a binary hypothesis test for each spectral dimension to test if each of spectral dimensions can be used to accommodate one signal source as opposed to Principal Component Analysis (PCA) that uses eigenvalues to determine the total number of signal sources instead of a signal source in an individual spectral dimension [12].

Wei Xiong and Chein-I Chang develops a new approach to VD estimation based on data representation in a certain form, specifically based on linear spectral mixture analysis (LSMA)[4]. With the LSMA in mind the proposed approach interprets a spectrally distinct signature as an image endmember that can be used to specify a particular spectral class. All the methods use the eigenvalues of the observation correlation and covariance matrices. In all cases, not needs to be known about basis vectors, which is an advantage over supervised methods [17,18].

2. The proposed VD Algorithm

A hyperspectral image can be illustrated as an image cube with the two dimensions of the face of the cube represents the spatial information and the third dimensional representing the spectral information. The information