Future Generation Computer Systems 27 (2011) 419–426

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

A simulated annealing feature extraction approach for hyperspectral images

Yang-Lang Chang *

Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 15 March 2010
Received in revised form
15 July 2010
Accepted 11 August 2010
Available online 17 September 2010

Keywords:
Hyperspectral images
SAR images
Simulated annealing feature extraction
Dimensionality reduction

ABSTRACT

In this paper, a novel study of the *simulated annealing feature extraction* (SAFE) for high-dimensional remote sensing images is proposed. The approach is based on the *greedy modular eigenspace* (GME) scheme. GME was developed by clustering highly correlated bands into a smaller subset based on the *greedy* algorithm. Unfortunately, GME doesn't guarantee to reach a global optimal solution by the *greedy* algorithm except by the *exhaustive search* method. Accordingly, finding an optimal (or near-optimal) solution is very expensive. In order to overcome this disadvantage, the SAFE scheme is introduced to improve the performance of GME feature extraction optimally by modifying the *correlation coefficient* operations and taking sets of non-correlated bands for hyperspectral images based on a *heuristic optimization algorithm*. It presents a framework, which consists of two algorithms, referred to as SAFE and the *feature scale uniformity transformation* (FSUT). SAFE is designed to extract features by a new defined three-dimensional *simulated annealing modular eigenspace* (SAME) to optimize the modular eigenspace, while FSUT is performed to fuse most correlated features from different spectrums associated with different data sources. The performance of the proposed method is evaluated by applying it to hyperspectral and airborne *synthetic aperture radar* (SAR) images. The experimental results demonstrated that SAFE is not only an effective scheme of *feature extraction* but also an alternative to the existing *dimensionality reduction* methods.

1. Introduction

For the most contemporary remote sensing advances, hyperspectral images are emerging fields with an abundance of applications, which cover a wide range from satellite imaging, monitoring systems to medical imaging and industrial product inspection. A lot of attention has been focused on the developing of hyperspectral feature extraction devoted to earth remote sensing. The increment of such high-dimensional data volumes greatly enhances the information content, but provides a challenge to the current techniques for analyzing such high volume data sets. In this paper, a novel technique of *simulated annealing feature extraction* (SAFE) is proposed for hyperspectral feature extraction of earth remote sensing. It utilizes the inherent separability of different classes in hyperspectral images to reduce dimensionality and further to generate a unique *simulated annealing modular eigenspace* (SAME) feature which is developed based on the *greedy modular eigenspace* (GME) scheme [1].

A general issue in hyperspectral images of remote sensing is how to improve class-separability without incurring the curse of dimensionality [2]. Researchers of various research communities, including statistics, pattern recognition, and data mining, all describe the difficulties associated with the feasibility of distribution estimation associated with this problem. Accordingly, extracting the most valuable and meaningful information has become ever more important. Numerous techniques were developed for *feature extraction/selection* to reduce dimensionality without loss of class separability for dealing with high-dimensional data sets [3–7]. The most widely used approach is the *principal components analysis* (PCA) which reorganizes the data coordinates in accordance with data variances so that features are extracted based on the magnitudes of their corresponding eigenvalues [8]. *Fisher discriminant analysis* uses the between-class and within-class variances to extract desired features and reduce dimensionality [9]. Another well-known approach is *orthogonal subspace projection* [10–12], which projects all undesired pixels into a space orthogonal to the space generated by the desired pixels to achieve high accuracy of classification in dealing with high-dimensionality classifications. They focus on the estimation of statistics at full dimensionality to extract classification features. The potential differences between class covariances are not explored.

Chang et al. [1] proposed a GME *feature extraction* (GMEFE) approach to solve this problem. GMEFE was developed by clustering highly correlated bands into a smaller subset based on the *greedy* algorithm and was shown effective in hyperspectral feature extraction. It was developed by grouping highly correlated hyperspectral bands of different classes into a smaller subset of band modular to overcome the dependency on global statistics, while preserving the inherent separability of different classes. Most classifiers seek only one set of features that discriminates all

* Tel.: +886 2 2771 2171; fax: +886 2 2731 7187.
E-mail address: ylchang@ntut.edu.tw.

0167-739X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2010.08.008

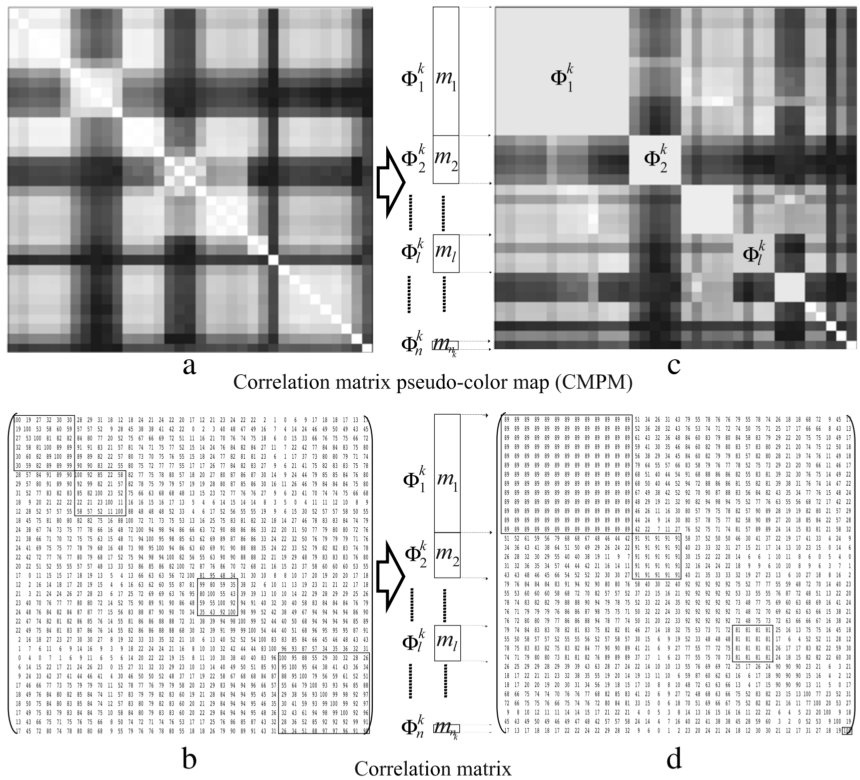

Fig. 1. An example illustrating (a) an original GME-CMPM, (b) its corresponding correlation matrix for class $\boldsymbol{\omega}_k$, (c) the GME set for class $\boldsymbol{\omega}_k$, and (d) its corresponding correlation matrix after GME feature extraction.

classes simultaneously. This not only requires a large number of features, but also increases the complexity of the potential decision boundary. The GMEFE method solves this problem and speeds up the feature extraction processes significantly. Although GMEFE can provide acceptable results for feature selection and dimensionality reduction, it doesn't guarantee to reach a global optimal or near-optimal solution by the *greedy* algorithm except by the *exhaustive search* method. Accordingly, finding the optimal or near-optimal solution is very expensive.

Correspondingly, finding an efficient alternative has become necessary to overcome the aforementioned drawback of GMEFE. One consequence is the development of a technique known as *simulated annealing* (SA) [13,14] for feature extraction of high-dimensional data sets. SA optimization has been widely adopted in fields such as electronics design automation [15,16] and graph partitioning problems [17]. Instead of adopting the band-subset-selection paradigm underlying the *greedy optimization* approach of GMEFE, the proposed SAFE uses the *heuristic optimization algorithm* to group the subsets of non-correlated bands for hyperspectral images. SAFE can readily extract each band and sort different classes into the most common band subset. It can not only speed up the procedure to simultaneously extract the most significant features according to the SA *optimization scheme*, but also makes use of the hyperspatial characteristics embedded in SAME features.

The performance of the proposed SAFE is evaluated by fusing the *MODIS/ ASTER airborne simulator* (MASTER), a hyperspectral sensor, and airborne *synthetic aperture radar* (SAR) images for land cover classification during the Pacrim II campaign. Experimental results demonstrate that the proposed SAFE approach is an effective method for feature selection and dimensionality reduction. Compared to GMEFE, SAFE can not only effectively group highly correlated bands, but also further improve the discriminatory properties which are crucial to the subsequent classification process. The rest of this paper is organized as follows. In the next section, the proposed SAFE method is described in detail. In Section 3, a set of experiments is conducted to demonstrate the feasibility and utility of the proposed approach. Finally, in the last section, several conclusions are presented.

2. Methodology

2.1. Review of GMEFE

A visual *correlation matrix pseudo-color map* (CMPM) proposed by Lee and Landgrebe [18] is used in Fig. 1(a) and (c) to emphasize the second-order statistics in hyperspectral data and to illustrate the magnitude of correlation matrices in the proposed SAFE method. Also shown in Fig. 1(c) is a GME-CMPM set $\boldsymbol{\Phi}^k$, $\boldsymbol{\Phi}^k = (\Phi_1^k, \ldots, \Phi_l^k, \ldots, \Phi_{n_k}^k)$ for class $\boldsymbol{\omega}_k$ which was reported by Chang et al. [1]. It illustrates the original CMPM and the reordered one after SAFE. Each modular eigenspace Φ_l^k, a subset of GME, includes a subset of highly correlated bands. Each ground cover type or material class has a distinct set of GME-generated *feature eigenspaces*.

GMEFE is a spectral-based technique that explores the correlation among bands. It utilizes the inherent separability of different classes for the high-dimensional data sets to reduce dimensionality and formulate a unique GME feature. GMEFE performs a *greedy* iteration searching algorithm which reorders the correlation coefficients in the data correlation matrix row by row and column by column simultaneously, and groups highly correlated bands as GME *feature eigenspaces* that can be further used for *feature extraction* and *selection*. Reordering the bands in terms of wavelengths in high-dimensional data sets, without regard for the original order, is an important characteristic of GMEFE. Fig. 2 shows the graphical mechanism of GMEFE spectral band reordering.

GMEFE defines a correlation submatrix $\mathbf{c}_{\Phi_l^k}[m_l][m_l]$ which belongs to the lth modular eigenspace Φ_l^k of GME $\boldsymbol{\Phi}^k$, $\boldsymbol{\Phi}^k = (\Phi_1^k, \ldots, \Phi_l^k, \ldots, \Phi_{n_k}^k)$, for a land cover class $\boldsymbol{\omega}_k$ in the dataset,