Applied Soft Computing 12 (2012) 2683–2692

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

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

Fuzzy clustering algorithms incorporating local information for change detection in remotely sensed images

Niladri Shekhar Mishra [a], Susmita Ghosh [b], Ashish Ghosh [c,*]

[a] Department of Electronics and Communication Engineering, Netaji Subhash Engineering College, Kolkata 700152, India
[b] Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, India
[c] Center for Soft Computing Research, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India

ARTICLE INFO

Article history:
Received 13 April 2011
Received in revised form 23 January 2012
Accepted 5 March 2012
Available online 24 April 2012

Keywords:
Remote sensing
Change detection
Multitemporal images
Local information
Fuzzy *c*-means clustering
Gustafson–Kessel clustering
Genetic algorithms
Simulated annealing
Xie–Beni and fuzzy hypervolume validity measures

ABSTRACT

In this paper we have used two fuzzy clustering algorithms, namely fuzzy *c*-means (FCM) and Gustafson–Kessel clustering (GKC) along with local information for unsupervised change detection in multitemporal remote sensing images. In conventional FCM and GKC no spatio-contextual information is taken into account and thus the result is not so much robust to small changes. Since the pixels are highly correlated with their neighbors in image space (spatial domain), incorporation of local information enhances the performance of the algorithms. In this work we have introduced a new technique for incorporation of local information. Change detection maps are obtained by separating the pixel-patterns of the difference image into two groups. Hybridization of FCM and GKC with two other optimization techniques, genetic algorithm (GA) and simulated annealing (SA), is made to further enhance the performance. To show the effectiveness of the proposed technique, experiments are conducted on two multispectral and multitemporal remote sensing images. Two fuzzy cluster validity measures (Xie–Beni and fuzzy hypervolume) have been used to quantitatively evaluate the performance. Results are compared with those of existing state of the art Markov random field (MRF) and neural network based algorithms and found to be superior. The proposed technique is less time consuming and unlike MRF does not require any a priori knowledge of distributions of changed and unchanged pixels.

1. Introduction

In remote sensing applications, *change detection* is a process aimed at identifying the differences in the state of a land cover by analyzing a pair of images acquired on the same geographical area at different times [1,2]. Such a problem plays an important role in different domains like studies on land use/land cover dynamic [3], monitoring shifting cultivations, burned area identification [4], analysis of deforestation processes [5,6], assessment of vegetation changes [7], monitoring of urban growth [8] and oceanography [9]. Since all these applications usually require an analysis of large areas, development of completely automatic and unsupervised change detection techniques is of high relevance to reduce the time (effort) required by manual image analysis.

Change detection in remotely sensed data may be done either in supervised or in unsupervised manner [4,5,7,8,10–16]. In supervised techniques, a set of *training patterns* is required for learning the classifier. In real-life, it is difficult to have data containing spectral signatures of changes from which *training patterns* can be generated. In unsupervised techniques, there is no need of *training data*. Thus the usefulness of unsupervised techniques is more than supervised ones for this problem. We may think of unsupervised change detection problem as a clustering one where the task is to partition the data into two groups *changed* and *unchanged*.

Before performing *change detection* between two multitemporal images, a certain degree of (pre)processing is needed [11] because of co-registration error [11,17,18], radiometric and geometric errors [19]. Thus, in literature [1,2], three steps are suggested to be performed sequentially for unsupervised *change detection*. They are (i) pre-processing, (ii) image comparison and (iii) image analysis. In step (i), operations like coregistration, radiometric and geometric corrections and noise reduction are done to make the two multitemporal images compatible. To remove the effects of sensor errors and environmental factors, radiometric corrections are needed [11,19].

In step (ii), the two pre-processed images are taken as input and compared pixel by pixel and thereafter another image is generated, called the difference image (*DI*). To generate the *DI*, we may consider: (a) only one spectral band (i.e. Univariate Image Differencing – UID) [2], (b) multiple spectral bands (i.e. Change Vector Analysis – CVA) [2], (c) vegetation indices (Vegetation Index

* Corresponding author. Tel.: +91 33 2575 3110/3100; fax: +91 33 2578 3357.
E-mail addresses: niladrimishra@gmail.com (N.S. Mishra), susmitaghoshju@gmail.com (S. Ghosh), ash@isical.ac.in (A. Ghosh).

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2012.03.060

Differencing – VID) [2,20], etc. Tasseled Cap Transformation [21] is also a popular method. The most popular of these is the CVA and is used in our study. We have chosen CVA because by using this technique reflectance properties of various land cover types can be combined.

After performing the above two steps (pre-processing and image comparison and thereby generating the *DI*) *change detection* (step (iii)) is done on the *DI*. Either context-insensitive or context-sensitive procedure is adopted [4] for this. Histogram thresholding [16] is of the first kind. The threshold value may be detected by manual trial-and-error (MTET) process or by automatic techniques by analyzing the statistical distribution of the *DI*. In these cases spatial correlation between the neighboring pixels is not taken into account. Most of the context-sensitive techniques [10,15] are based on MRF, require the selection or estimation of a model for the statistical distributions of *changed* and *unchanged* classes, and can overcome the drawbacks of context-insensitive approaches mentioned earlier. Algorithms (like Expectation-Maximization (EM) [22]) are required for estimating the class distributions assuming different standard distributions e.g. Gaussian, generalized Gaussian [10] and mixture of Gaussians. A few context-sensitive techniques using neural networks are also suggested recently [4,5,23,24].

Relevance of fuzzy set theoretic methods in pattern recognition and image analysis problems has adequately been addressed in the literature [25–35]. Fuzzy clustering incorporating local information for change detection in remotely sensed images has not been reported in the literature. So, in order to overcome the limitations imposed by the need of selecting or estimating a statistical model for *changed* and *unchanged* class distributions, we propose unsupervised, distribution free and context-sensitive change detection techniques based on fuzzy clustering [26] approach. Normally the pixels of the *DI* belonging to two clusters *changed* and *unchanged* are not separable by sharp boundaries (as they are highly overlapped). As fuzzy clustering techniques are more appropriate and realistic to separate overlapping clusters [19], we have chosen fuzzy clustering techniques to have a better judgement of the two groups. In this regard we have used two fuzzy clustering algorithms namely fuzzy *c*-means [26] and Gustafson–Kessel [36]. There are several fuzzy cluster validity indexes available in the literature to evaluate fuzzy clustering results. We have used two of them. The first one is proposed by Xie and Beni [37] and the second one is by Gath and Geva [38]. They consider both intra cluster compactness and inter cluster separation. While evaluating the outcome of GK-type clustering using Xie–Beni validity index we have used Mahalanobis norm as in [39].

In image clustering applications FCM or GKC can be treated differently from data clustering. Pixels are normally highly correlated to their neighbors in the image space. This should be exploited for more efficiency. Also the homogeneous and nonhomogeneous regions (in context with gray values of pixels) in one image do not bear the same information. So instinct suggests that the amount of local information should vary from zone to zone, and better if varied from pixel to pixel. A pre-computation is done to incorporate the local neighborhood information in a *variable* fashion to the pixels of the *DI*. After generating the patterns they are subjected to clustering for identifying their class labels (*changed* or *unchanged*). Local information is incorporated here in such a way that its amount can vary from pixel to pixel automatically depending on the degree of homogeneity of its surrounding pixels (over a fixed window). It makes this method more robust for small changes and experimental results show that this technique is very efficient than the existing ones.

Two well known other optimization techniques namely genetic algorithms (GAs) [40] and simulated annealing (SA) [41] have been used to minimize the objective functions of the above mentioned clustering techniques to yield better results.

To assess the proposed technique, experiments are carried out on two real world data sets and compared the results with those obtained by already published techniques [4] for solving the same problem of change detection on the same data sets. We also compared the result of hard clusterings to show the effectiveness of the fuzzy ones. The proposed techniques have an edge with respect to both error and time requirements.

This paper is organized as follows: Section 2 provides a brief description of a few crisp and fuzzy clustering algorithms. The next section is about validity measures for fuzzy clustering. The proposed *change detection* technique has been described in Section 4. The data sets used in the experiments and the results obtained are described in Sections 5 and 6, respectively. Finally, in Section 7, conclusions are drawn.

2. Clustering

The clustering algorithms [42] (both fuzzy and non-fuzzy) used in the present investigation are described here in brief.

2.1. Hard c-means (HCM) clustering

The HCM [42] algorithm minimizes the following objective function to divide the data set into *c* clusters.

$$J(\mathbf{X};\mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{n} D_{ik}, \tag{1}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ is the set of n patterns, $\mathbf{x}_k$ is the kth pattern $\in \mathbf{X}$ and $D_{ik} = ||\mathbf{x}_k - \mathbf{v}_i||^2$ (Euclidean norm) is the dissimilarity measure between the sample $\mathbf{x}_k$ and the ith cluster center $\mathbf{v}_i$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c]$.

2.2. Fuzzy clustering

2.2.1. Fuzzy c-means (FCM) clustering

FCM [26] attempts to find fuzzy partitioning of a given data set by minimizing the objective functional

$$J_m(\mathbf{X}; U, \mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{n} (\mu_{ik})^m D_{ik}, \tag{2}$$

where $U = [\mu_{ik}] \in M_{fcn}$, fuzzy partition matrix of $\mathbf{X}$, and

$$\mathbf{v}_i = \frac{\sum_{k=1}^{n}(\mu_{ik})^m \mathbf{x}_k}{\sum_{k=1}^{n}(\mu_{ik})^m} \tag{3}$$

with μ_{ik} (degree of belonging of pattern $\mathbf{x}_k$ to the ith cluster) is expressed as

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c}(d_{ik}/d_{jk})^{2/(m-1)}}, \tag{4}$$

where $d_{ik} = \sqrt{D_{ik}}$ and m (>1) is a parameter, called fuzzifier, which controls the fuzziness of the patterns. During optimization of the functional $J_m(\mathbf{X}; U, \mathbf{V})$, following two constraints must be satisfied: (i) $\sum_{i=1}^{c}\mu_{ik} = 1$ and (ii) $\mu_{ik} \in [0, 1]$.

2.2.2. Gustafson–Kessel clustering (GKC)

Gustafson and Kessel introduced [36] adaptive distance norm to measure the distance between clusters using fuzzy covariance matrix (F_i). Each cluster has its own norm-inducing matrix A_i, a positive definite symmetric one, for automatically adapting its shape. F_i for the ith cluster is expressed as

$$F_i = \frac{\sum_{k=1}^{n}(\mu_{ik})^m(\mathbf{x}_k - \mathbf{v}_i)(\mathbf{x}_k - \mathbf{v}_i)^T}{\sum_{k=1}^{n}(\mu_{ik})^m}. \tag{5}$$