Computers & Geosciences 83 (2015) 1–16

Contents lists available at ScienceDirect

Computers & Geosciences

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

Semi-supervising Interval Type-2 Fuzzy C-Means clustering with spatial information for multi-spectral satellite image classification and change detection

Long Thanh Ngo[a,*], Dinh Sinh Mai[a], Witold Pedrycz[b,c,d]

[a] Department of Information Systems, Faculty of Information Technology, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, Vietnam
[b] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6R 2V4
[c] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University Jeddah, 21589, Saudi Arabia
[d] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

ARTICLE INFO

Article history:
Received 18 December 2014
Received in revised form
4 May 2015
Accepted 16 June 2015
Available online 23 June 2015

Keywords:
Interval type-2 fuzzy sets
Fuzzy C-Means
Type-2 fuzzy clustering
Satellite image analysis
Land cover classification
Change detection

ABSTRACT

Data clustering has been widely applied to numerous real-world problems such as natural resource management, urban planning, and satellite image analysis. Especially, fuzzy clustering with its ability of handling uncertainty has been developed for image segmentation or image analysis e.g. in health image analysis, satellite image classification. Normally, image segmentation algorithms like fuzzy clustering use spatial information along with the color information to improve the cluster quality. This paper introduces an approach, which exploits local spatial information between the pixel and its neighbors to compute the membership degree by using an interval type-2 fuzzy clustering algorithm, called IIT2-FCM. Besides, a Semi-supervising Interval Type-2 Fuzzy C-Means algorithm using spatial information, called SIIT2-FCM, is proposed to move the prototype of clusters to the expected centroids which are pre-defined on a basis of available samples. The proposed algorithms are applied to the problems of satellite image analysis consisting of land cover classification and change detection. Experimental results are reported for various datasets of the LandSat7 imagery at multi-temporal points and compared with the results produced by some existing algorithms and obtained from some survey data. The clustering results assessed with regard to some validity indexes demonstrate that the proposed algorithms form clusters of better quality and higher accuracy in problems of land cover classification and change detection.

1. Introduction

In clustering-based image segmentation approaches, the most important problem is to establish a method to determine whether or not the considered pixel belongs to a certain cluster. The "conventional" clustering algorithms like k-Means, Fuzzy C-Means (FCM), and interval type-2 Fuzzy C-Means (IT2-FCM) exhibit the same strategy based on the Euclidean distance to compute the degree of similarity between objects to be assigned to clusters with the corresponding membership degree. Not only color based similarity of the pixels but also the spatial relationship between pixels and their neighbors certainly impact the clustering results.

Satellite image analysis methods based on statistical parameters have been widely used because of their easy implementation and accuracy. However, these methods are often quite expensive, time consuming and only applicable to small areas. Currently, there are several approaches to classify satellite imagery in which fuzzy logic have been widely applied because of its advantages in handling ambiguous data. Normally, satellite imagery are affected by noise because of weather and errors associated with the photographic equipment and in this case fuzzy clustering becomes of interest.

Type-2 fuzzy sets form an extension of original fuzzy sets of type-1. They have been developed and applied to various problems (Karnik et al., 1999; Karnik and Mendel, 2001; Liang and Mendel, 2000; Mendel and John, 2002; Mendel et al., 2006; Liu, 2008; Ngo and Nguyen, 2012; Nguyen et al., 2015; Hwang and Rhee, 2007; Fisher, 2010) including data clustering. Fuzzy C-Means (FCM) clustering (Bezdek et al., 1984) and its variants have been widely applied to many problems including satellite image analysis. The drawback of the FCM algorithms is the limitation in handling uncertainty. Hence, the use of interval type-2 fuzzy sets in data clustering as the interval Type-2 Fuzzy C-Means clustering algorithm (IT2FCM) was studied in Hwang and Rhee (2007). In this method, the FOU (footprint of uncertainty) of the type-2 fuzzy set

* Corresponding author.
E-mail addresses: ngotlong@mta.edu.vn (L.T. Ngo), maidinhsinh@gmail.com (D.S. Mai), wpedrycz@ualberta.ca (W. Pedrycz).

http://dx.doi.org/10.1016/j.cageo.2015.06.011
0098-3004/

is built by using two values of the fuzzifier (fuzzification coefficient) being one of the essential parameters of the FCM algorithm. Besides, in order to improve the quality of clustering in image segmentation including satellite images, various ways of using spatial information together with color information have been developed. Remote sensing image analysis is studied using various approaches exploiting fuzzy logic (Fisher, 2010; Stavrakoudis et al., 2011; Shankar et al., 2011; Liu et al., 2013; Ghaffarian and Ghaffarian, 2014; Martinez and Martinez, 2014).

The paper deals with a novel type-2 fuzzy clustering approach to the problems of land cover classification and change detection. Three essential issues are addressed in this paper:

(1) the use of local spatial information between the pixel and its neighbors to compute the membership degree in interval type-2 fuzzy clustering algorithm, called IIT2-FCM;
(2) development of the semi-supervised interval type-2 Fuzzy C-Means algorithm using spatial information, called SIIT2-FCM, with intent to navigate the prototype of clusters to the expected centroids which are pre-defined on the basis of existing samples;
(3) application of IIT2-FCM and SIIT2-FCM to the problems of land cover classification and change detection of multi-temporal points in multi-spectral remote sensing imagery.

Experimental results are reported on various datasets of Landsat7 images at multi-temporal points. A comparative analysis is provided involving results produced by some existing algorithms available in the literature. Our interest is to quantify the quality of clustering results in terms of well-known validity indexes and look at the evaluation of the clusters in terms of accuracy in land cover classification and change detection.

The paper is organized as follows: Section 2 presents a literature review on interval type-2 fuzzy clustering, spatial information and satellite analysis. Section 3 covers a background material by looking at type-2 fuzzy sets and interval type-2 C-Means clustering; Section 4 introduces spatial information, IT2-FCM with spatial information, semi-supervising IT2FCM with spatial information; while Section 5 demonstrates how to apply IIT2FCM, SIIT2FCM to land-cover classification and change detection. Section 6 covers conclusions and identifies some future works.

2. Literature review

Let us note that the main issues studied in this paper are as follows: (1) how to use interval type-2 fuzzy sets and exploit their ability of handling uncertainty better in satellite image analysis; (2) how to use spatial information together with color information in clustering algorithms; and (3) how to apply the method to satellite image analysis (land cover classification and change detection).

In what follows, we briefly review some related studies.

Some linkages with the use of type-2 fuzzy sets to clustering algorithms, applications have been identified. Ji et al. (2014) proposed interval-valued possibilistic FCM clustering to incorporate interval type-2 sets into the possibilistic FCM to better handle and manage the uncertainty implied by data. Qiu et al. (2013) introduced the modified interval type-2 FCM using spatial information to handle uncertainty in MR images. Torshizi and Zarandi (2014a, 2014b) proposed an algorithm of general type-2 fuzzy clustering for analyzing gene expression data with newly developed general type-2 cluster validity index. Zarinbal et al. (2014) proposed Interval Type-2 Relative Entropy FCM in which the uncertainty associated with membership functions is the main concern and an application to MR image segmentation was discussed. The combination of IT2FCM and other techniques such as multiple kernel technique was also proposed (Nguyen et al., 2015). These methods handle uncertainties and deal with the input features coming from multiple sources.

With regard to satellite image classification, various applications of type-2 fuzzy sets related to satellite image analysis were introduced, consisting of land cover classification (Ngo and Nguyen, 2012; Fisher, 2010). Fisher (2010) proposed remote sensing of land cover classes given that type 2 fuzzy sets reveal uncertainty, and allows the analysis to be specific about minimum, maximum, and average degree of land cover types. Multi-spectral satellite images contain uncertainty. For instance, each pixel is captured from a square depending on the resolution of image (maybe 30×30 m) which may exhibit several land covers (water, soil or vegetable). The other uncertainty could come as a result of error of equipment or software used of processing data. Hence, clustering algorithms using type-2 fuzzy sets are suitable to deal with satellite image classification.

In order to improve the quality of clustering in image segmentation including satellite images, various ways of using spatial information together with color information in FCM algorithm were proposed (Despotovic et al., 2010; Wang et al., 2009, 2013; Zhao et al., 2011, 2013; Liu and Pham, 2012; Liu et al., 2012; Zhao, 2013; Vargas et al., 2013; Benaichouche et al., 2013). Despotovic et al. (2010) used a mask whose center is the considered pixel, while relationships between pixels in the mask and the center are used to determine the degree of similarity being used to estimate the membership values. Wang et al. (2009) proposed adaptive spatial information-theoretic clustering (ASIC) algorithm with the modified objective function which exhibits a new dissimilarity measure and the weighting factor for neighborhood effect becomes fully adaptive to the image content. The ASIC enhances the smoothness towards piecewise-homogeneous segmentation and reduces the edge blurring effect. Zhao et al. (2011) and Zhao (2013) proposed two novel fuzzy clustering algorithms using the self-tuning non-local spatial information. In the first algorithm, the self-tuning non-local spatial information for each pixel is defined and then introduced into the objective function of FCM. In the second one, a novel gray level histogram is constructed by using the self-tuning non-local spatial information for each pixel, and then clustering is performed on a basis of this gray level histogram. Zhao et al. (2013) also included spatial information in the objective function of a certain generalized Fuzzy C-Means clustering algorithm, and then the kernel induced distance is adopted to substitute the Euclidean distance in the new objective function. Liu and Pham (2012) presented a fuzzy clustering algorithm which can handle spatial constraints, which is based on the notions of hyperplanes, Fuzzy C-Means, and spatial constraints. By adding a spatial regularizer into the fuzzy hyperplane based objective function, the proposed method can take into account additional important information of inherently spatial data. Liu et al. (2012) have come up with a novel fuzzy spectral clustering algorithm with robust spatial information for image segmentation (FSCRS). The similarity matrix was obtained by using a robust gray-based fuzzy similarity measure. The spectral graph partitioning method can be applied to the similarity matrix to group the gray values of the new generated image and then the corresponding pixels in the image are reclassified to obtain the final segmentation result. Vargas et al. (2013) introduced two enhanced Fuzzy C-Means clustering algorithms with spatial constraints for noisy color image segmentation. The Rank M-type L (RM-L) and L-estimators of spatial information of the pixels are involved in the FCM algorithm to provide robustness to the segmentation schemes. Wang et al. (2013) proposed an adaptive spatial information-theoretic fuzzy clustering algorithm to improve the robustness of the conventional Fuzzy C-Means (FCM) clustering algorithms for image