Classification of Remote Sensing Imagery Based on Density and Fuzzy c-Means Algorithm

Trinh Le Hung, Le Quy Don Technical University, Hanoi, Viet Nam
Mai Dinh Sinh, Le Quy Don Technical University, Hanoi, Viet Nam

ABSTRACT

The goal of data clustering is to divide a set of data into different clusters, so that the data in the same cluster show some similar characteristics. There are many clustering methods for satellite image segmentation, such as k-means, c-means, iso-data, minimum distance algorithms. Each method has certain advantages and disadvantages, but generally they are based on brightness value to divide the pixels of the image into clusters. Actually, the probability of occurrence of frequency of appearance of pixel has certain effects on clustering results. In this article, the authors propose a method for clustering satellite imagery based on density. It consists of two main steps: find cluster centroid using density and data clustering using fuzzy c-Means algorithm (DFCM). The results obtained in this study can be used to potentially improve classification accuracy of satellite image.

KEYWORDS

Density, Fuzzy Clustering, Fuzzy C-Means, Multi-Spectral Image, Remote Sensing

INTRODUCTION

Remote sensing data clustering is an extremely important part of satellite image processing (Torahi & Chai, 2011; Mai, Trinh & Ngo, 2016; Mai & Ngo, 2015; Mai & Ngo, 2018). The results of satellite image classification can be used for a variety of purposes, such as natural resource research and environmental monitoring, urban planning and ensure national defense and security. Meanwhile, optical remote sensing data sources are often affected by weather conditions and the accuracy of the receiver, this make the image classification more complicate (Ngo, Mai & Pedrycz, 2015). In fact, uncertainty is inherently present in decision making. As such, it is increasingly imperative to research and develop new theories and methods based on fuzzy clustering (Li, 2017).

There are many satellite image classification methods (Han, Chi & Yeon, 2005; Gordo, Martinez, Gonzalo & Arquero, 2013), such as manual thresholds methods (Yang et al., 2016), unsupervised classification methods (Genitha & Vani, 2013), supervised classification methods (Jog and Dixit, 2016), fuzzy clustering method (Rauf, Valentijn & Leonid, 2009) and method based on intuitionistic...
fuzzy sets (Li, 2004; Li & Cheng, 2002). These methods often use some common algorithms, such as K-means, c-Means, Iso-data, minimum distance and Fuzzy c-means. These clustering algorithms are essentially using the same strategy based on brightness to split into clusters (Jog & Dixit, 2016; Rauf et al., 2009) without regard to the density of the pixels, while high density pixels are usually located near the centroid of the cluster (Peherstorfer, Pflüger & Bungartz, 2012; Chen, Yan & Wang, 2014; Benmouiza & Cheknane, 2016).

Many scientists in the field of remote sensing data processing have proposed clustering methods based on density of pixels, in which density based spatial clustering of applications with noise (DBSCAN) is commonly used for satellite image classification (Khan, Rehman, Aziz, Fong & Saravady, 2014; Benmouiza & Cheknane, 2016). This algorithm requires only one input parameter and supports the users in determining an appropriate value for it. It discovers clusters of arbitrary shape and divides high density areas into cluster without depend on the size of data. In terms of implementation, this algorithm is also difficult to find the optimal radius of the density function around each pixel. In addition, the execution time of this algorithm is quite slow, especially when tested on large datasets, such as satellite imagery (Ngo, Mai & Nguyen, 2012). To overcome these limitations, many scientists are interested in improving this algorithm. Peherstorfer et al. (2012) presented a grid-based density estimation method to improve the speed of clustering. Chen et al. (2014) improved the DBSCAN algorithm by expanding the clusters which uses the margins of the objects, such as a pixel, to reduce the computation time. These improvements significantly reduce clustering time; however, affect the accuracy of clustering results.

To solve the above problem, this study proposed a method for approximating the centroid of cluster based on the density of pixels. Next step, the authors use approximation centroids for classification satellite imagery using the fuzzy c-means algorithm.

**PROPOSAL METHODOLOGY**

**Scientific Basic**

**Density**

The concept of density can be understood as the quantity representing the amount of matter in unit of measure (length, area, volume). So, the pixel density is the frequency of the pixel per unit of measure. Usually, the centroid of cluster is the average value of the pixels, so if the pixel has high frequency of appearance, that pixel is closer to the centroid of cluster (Ngo et al., 2012).

**Fuzzy C-Means Algorithm**

In general, fuzzy memberships in fuzzy c-means clustering algorithm (FCM) achieved by computing the relative distance among the patterns and cluster centroids (Bezdek & Ehrlich, 1984). Hence, to define the primary membership for a pattern, FCM algorithm defines the membership using value of $m$. The use of fuzzifier gives different objective function as follows:

$$J_m(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{C} (u_{ik})^m d_{ik}^2$$  \hspace{1cm} (1)

where $d_{ik} = \| x_{ik} - v_i \|$ is Euclidean distance between the pattern $x_{ik}$ and the centroid $v_i$, $C$ is number of clusters and $N$ is number of patterns, $m$ is any real number greater than 1.

Degree of membership $u_{ik}$ is determined as follow:
Related Content

Adaptive Awareness of Hospital Patient Information through Multiple Sentient Displays
Jesus Favela, Mónica Tentori, Daniela Segura and Gustavo Berzunza (2011). Ubiquitous Developments in Ambient Computing and Intelligence: Human-Centered Applications (pp. 31-42).
www.igi-global.com/chapter/adaptive-awareness-hospital-patient-information/53323?camid=4v1a

Design and Usage of a Process-Centric Collaboration Methodology for Virtual Organizations in Hybrid Environments
Thorsten J. Dollmann, Peter Loos, Michael Fellmann, Oliver Thomas, Andreas Hoheisel, Peter Katranuschkov and Raimar Scherer (2013). Organizational Efficiency through Intelligent Information Technologies (pp. 45-64).
www.igi-global.com/chapter/design-usage-process-centric-collaboration/71960?camid=4v1a
Fuzzy Rule Based Environment Monitoring System for Weather Controlled Laboratories using Arduino

Cloud-Based Big Data Analytics in Smart Educational System
[www.igi-global.com/chapter/cloud-based-big-data-analytics-in-smart-educational-system/186477?camid=4v1a](www.igi-global.com/chapter/cloud-based-big-data-analytics-in-smart-educational-system/186477?camid=4v1a)