Computers and Electronics in Agriculture 83 (2012) 92–101

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

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

A hybrid method combining SOM-based clustering and object-based analysis for identifying land in good agricultural condition

Kadim Taşdemir *, Pavel Milenov, Brooke Tapsall

Monitoring Agricultural Resources Unit, Institute for Environment and Sustainability, European Commission Joint Research Centre, Via E. Fermi 2749, 21027 Ispra (VA), Italy

ARTICLE INFO

Article history:
Received 19 October 2011
Received in revised form 17 January 2012
Accepted 26 January 2012

Keywords:
Land cover identification
Unsupervised clustering
Self-organizing maps
CONN linkage
Object based image analysis
Good agricultural condition

ABSTRACT

Remotely sensed imagery is currently used as an efficient tool for agricultural management and monitoring. In addition, the use of remotely sensed imagery in Europe has been extended towards determination of the areas potentially eligible for the farmer subsidies under the Common Agricultural Policy (CAP), through interactive or automatic land cover identification. For accurate quantification and fast identification of agricultural land cover areas from the imagery, a hybrid method, which combines automated clustering of self-organizing maps with object based image analysis, and called SOM + OBIA, is proposed. Performance analysis on three test zones (using multi-temporal Rapideye imagery) indicates that for the basic land cover categories (forest, water, vegetated areas, bare areas and sealed surfaces), unsupervised classification with the proposed SOM + OBIA method achieves an identification accuracy comparable to the accuracy of the traditional interactive object oriented analysis, with considerably less user interaction.

1. Introduction

Land cover identification from remote sensing images has been essential for agricultural management and monitoring. Particularly in Europe, in order to support the European farmers, to regulate the impact of agricultural practices to the environment, and to ensure economic sustainability in rural areas, the European Union (EU) established the Common Agricultural Policy (CAP) followed later by an annual Control with Remote Sensing (CwRS) program where a large amount of remote sensing imagery is employed on a regular basis. A specific payment scheme under the CAP applied in numerous member states is the Single Area Payment Scheme (SAPS) which manages payments per eligible hectare of agricultural land. According to the SAPS, agricultural area eligible for payments is the utilized agricultural area, maintained in good agricultural condition (GAC) at a given reference date. Two exceptions for the "reference date" condition are Bulgaria and Romania where any utilized agricultural area, maintained in GAC at the time of the farmer declaration, regardless of its past status (without a historical reference to a given year), can be considered eligible for payment in these countries. Even though it provides flexibility to these countries in terms of managing the EU subsidies, it also creates a substantial administrative challenge for them, since agricultural land eligible for payment should be assessed annually in the years following their EU accession, requiring identification of the potentially eligible agricultural land (for the entire country).

Land cover classification from remotely sensed imagery is often performed at pixel level using a multi-class setting, by supervised or unsupervised methods (see Wilkinson (2005), and references therein). On the one hand, supervised classification algorithms necessitate labeled samples (pixels with known class membership for each class) for their training; and thus their performance depends on availability of labeled samples that can only be obtained purely from field inspection and through manual expert labeling (Mitra et al., 2004). On the other hand, unsupervised classification, achieved by clustering methods, depends on pixel similarity determined according to some criteria (Xu and Wunsch, 2005) (rather than on labeled training samples), requiring very limited supervision (i.e. labeling the final extracted clusters). A detailed review for unsupervised classification methods is given in Goncalves et al. (2008), indicating that clustering based on the self-organizing map (SOM, an artificial neural network based on observed properties of neural maps, introduced by Kohonen (1997)) is a promising tool due to SOM's important properties, such as the input space approximation, topological ordering, density matching, and the ease of its implementation. These properties facilitate extraction of clusters through interactive SOM visualization (see Vesanto (1999), Taşdemir and Merényi (2009), and references therein) and automated SOM clustering, which often use hierarchical agglomerative clustering methods as reviewed in Vesanto and Alhoniemi (2000) and Taşdemir et al. (2011). Therefore, SOMs are widely used for cluster extraction from remotely sensed images

* Corresponding author. Tel.: +39 0332785040; fax: +39 0332789029.
E-mail address: kadim.tasdemir@jrc.ec.europa.eu (K. Taşdemir).

0168-1699/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2012.01.017

with complicated data structures (Ji, 2000; Villmann et al., 2003; Merényi et al., 2009). SOM based clustering is also shown more successful than ISODATA, a commonly used unsupervised clustering method, in remote sensing applications (Merényi et al., 2007). A recent method by Taşdemir et al. (2011) utilizes the SOM properties by introducing a hierarchical clustering based on detailed local density distribution and achieves an accuracy higher than the accuracy obtained by distance based clustering approaches.

In addition to the pixel based classification approaches, object based image analysis (OBIA) methods, which segment the image into objects, then classify the resulting objects rather than the pixels, are becoming popular for land cover identification as well (Aplin et al., 1999; Blaschke et al., 2000; Shackelford and Davis, 2003; Walter, 2004). This is mainly due to the capacity of the OBIA to accommodate and consider various types of information, resulting in interpretation of the spatial context of land cover features, integrating and applying thematic information from different spatial datasets, such as the Land Parcel Identification System (LPIS).[1] Therefore, compared to pixel based methods, object based classification may produce more recognizable maps that can also be integrated with the geographical information system environments (Benz et al., 2004; Walter, 2004). A detailed review of object oriented methods is given in Blaschke et al. (2000). Some studies such as Laliberte et al. (2007) and Ruiz et al. (2009) perform object based classification by a decision tree classifier, due to the facts that decision trees, which are semi-automatic approaches, can drastically reduce time spent during the classification process; and the decision rules in the tree are non-parametric statistical techniques. In addition, decision trees allow the use of ancillary data such as the Land Parcel Identification System (LPIS) or other thematic variables (Ruiz et al., 2009). The combination of object oriented approach with decision trees produced low error rates and reduced object features in urban mapping (Thomas et al., 2003) and in rural areas (Laliberte et al., 2007), while allowing the flexibility to strengthen the rules by selection of features of interest and omitting features irrelevant to analysis. Definiens Enterprise Image Intelligence Suite (DEIIS), a commercial object based image analysis system, introduced this approach to a major community by providing an easy implementation and producing reliable classification performance (Flanders et al., 2003; Aplin and Smith, 2008). However, this approach may potentially suffer from the inherent ill posed problem of the image segmentation, which may produce both omission and commission segmentation errors (Baraldi et al., 2010). In addition, a heavy (expert) user interaction is often necessary to determine the decision rules and corresponding parameters and their thresholds (which are often needed to be reset for different imagery and application), resulting in a huge processing time.

To utilize the strong points of pixel-based and object-based approaches for identification of lands in good agricultural condition, a new method, SOM + OBIA, is proposed. The SOM + OBIA combines the initial unsupervised land cover identification, based solely on spectral information from the multi-temporal image, obtained by automated SOM clustering (Taşdemir et al., 2011) (the strong point of SOMs), with further object based analysis for interactive interpretation of several ancillary data layers (the strong point of the OBIA). Therefore the SOM + OBIA requires limited human intervention (and thus much less processing time), due to its construction based on a self-learned classification algorithm, and yet achieves similar accuracies of the traditional object based image analysis (OBIA) by decision trees.

The paper is organized as follows: Section 2 describes the study area and the remote sensing imagery used in this study; Section 3 describes the proposed SOM + OBIA method in detail; Section 4 presents and discusses the classification performances of the SOM + OBIA for three test zones; and Section 5 provides conclusion and future directions.

2. Study area and the imagery

2.1. Good agricultural condition

In order to ensure a correct assessment of the agricultural land suitable for the SAPS payments, a necessary preliminary step is to determine what is 'good agricultural condition' (GAC) in the national context, as there is no common legal definition of GAC at EU level. GAC (for Bulgaria and Romania) was defined in Tapsall et al. (2010), using two important criteria: *agricultural potential* and *accessibility*. Namely, the land shall have the potential to produce certain type of biomass naturally or using certain standard agricultural practices; and there should be no obstacles, neither natural nor man-made, preventing the access and use of the land for agricultural activities (cropping, grazing, etc.). At the country level, the regions in GAC can be of great diversity in terms of land cover dynamics, which can only be evaluated by monitoring vegetation development during the year (phenological cycle) and considering the detailed local crop calendar; whereas the regions in non-GAC are lands which are permanently bare or non-vegetated during the (cultivation) year (for example sealed surfaces, natural bare areas, urban regions and water bodies) or have features preventing the agricultural activity even though they are vegetated (for example closed forest, woodland, wetland, etc.). Therefore, detection and qualification of the permanently non-vegetated areas and the areas not accessible for agriculture are the primary targets. Generally, it is possible to efficiently extract bare, non-vegetated areas or woodlands using a single Rapideye imagery (if acquired in the correct period of the year). However, in order to capture and correctly interpret temporary bare areas, such as harvested agricultural fields, the analysis necessitates several time series due to varying calendars of agricultural crops. Additionally, accessibility of the lands having agricultural potential can be evaluated using textural properties and spatial content. For GAC analysis, GAC (including fallow lands) is grouped in one cluster whereas non-GAC is grouped into (i) permanently bare areas (including artificial surfaces such as urban regions, roads, etc.), (ii) water bodies, (iii) forests and woodlands (including hedges), (iv) vegetation in urban regions, and (v) vegetation enclosed by forests.

2.2. Study area

Bulgaria joined the European Union in 2007, and adopted legislation of the European Community for the management and monitoring of its agricultural land and benefit payments. The country is about 111.000 km^2 in size, extending from the western boundaries of the Black Sea on the East to Serbia and FYROM on the West. In the North, its boundary follows closely along the Danube River and Romania, whereas in the South, Turkey and Greece are neighboring countries. Land cover of Bulgaria is diverse throughout the country. For this reason, to provide sufficient representation of the different landscapes within Bulgaria, three test zones, shown in

[1] The Land Parcel Identification System (LPIS) is the fundamental component of the Integrated Administration and Control System (IACS), which is part of the technical implementation of the Common Agricultural Policy stipulated since the early 1990s. LPIS is based on spatial objects (so-called "reference parcels") – geographically referenced land units, as cadastral parcels or production blocks – within a GIS environment, to allow the identification, location and administrative crosschecks of the agricultural parcels declared by European farmers. Any LPIS has spatial (e.g. boundary coordinates and areas) and alphanumerical attributes (e.g. unique identification, maximum eligible hectares value). Each EU Member State has implemented a Land Parcel Identification System (LPIS) to administrate and control agricultural land and payments. Detailed information can be found in Devos (2011).