Pattern Recognition Letters 48 (2014) 24–33

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A rule-based classification methodology to handle uncertainty in habitat mapping employing evidential reasoning and fuzzy logic ☆

Zisis I. Petrou [a,b,*], Vasiliki Kosmidou [a], Ioannis Manakos [a], Tania Stathaki [b], Maria Adamo [c], Cristina Tarantino [c], Valeria Tomaselli [d], Palma Blonda [c], Maria Petrou [a]

[a] *Information Technologies Institute, Centre for Research and Technology Hellas, P.O. Box 60361, 6th km Xarilaou – Thermi, 57001 Thessaloniki, Greece*
[b] *Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*
[c] *Institute for Studies on Intelligent System for Automation (ISSIA), National Research Council (CNR), Via Amendola 122/D-O, 70126 Bari, Italy*
[d] *Institute of Plant Genetics (IGV), National Research Council (CNR), via G. Amendola 165/A, 70126 Bari, Italy*

ARTICLE INFO

Article history:
Available online 15 November 2013

Keywords:
Dempster–Shafer theory
Fuzzy rule-based object-oriented classification
Remote sensing
Habitat mapping
Uncertainty handling
Biodiversity

ABSTRACT

Habitat mapping is a core element in numerous tasks related to sustainability management, conservation planning and biodiversity monitoring. Land cover classifications, extracted in a timely and area-extensive manner through remote sensing data, can be employed to derive habitat maps, through the use of domain expert knowledge and ancillary information. However, complete information to fully discriminate habitat classes is rarely available, while expert knowledge may suffer from uncertainty and inaccuracies. In this study, a rule-based classification methodology for habitat mapping through the use of a pre-existing land cover map and remote sensing data is proposed to deal with uncertainty, missing information, noise afflicted data and inaccurate rule thresholds. The use of the Dempster–Shafer theory of evidence is introduced in land cover to habitat mapping, in combination with fuzzy logic. The framework is able to handle lack of information, by considering composite classes, when necessary data for the discrimination of the constituting single classes is missing, and deal with uncertainty expressed in domain expert knowledge. In addition, a number of fuzzification schemes are proposed to be incorporated in the methodology in order to increase its performance and robustness towards noise afflicted data or inaccurate rule thresholds. Comparison with reference data reveals the improved performance of the methodology and the efficient handling of uncertainty in expert rules. The further scope is to provide a robust methodology readily transferable and applicable to similar sites in different geographic regions and environments. Although developed for habitat mapping, the proposed rule-based methodology is flexible and generic and may be well extended and applied in various classification tasks, aiming at handling uncertainty, missing information and inaccuracies in data or expert rules.

1. Introduction

Habitat mapping is mainly performed through either in situ or remote sensing observations, the latter being increasingly popular due to their advantages in large area coverage, time and cost efficiency (Nagendra, 2001). Land cover (LC) maps extracted from remote sensing data in a more straightforward way, are often used as proxies for habitat map extraction, since they describe observable characteristics of a landscape, through the use of ancillary information. Habitat changes constitute significant indicators for biodiversity monitoring, ecosystem preservation and sustainability management, thus their mapping attracts the interest of various organizations and authorities worldwide (Bunce et al., 2013; Schmeller, 2008). Based on the generalized trend of international, national and regional authorities in producing LC maps, partially due to legal obligations (Tomaselli et al., 2013), an efficient framework for the conversion of LC into habitat classes is largely beneficial for sustainability management, conservation planning and biodiversity monitoring.

A large variety of classification approaches has been employed and evaluated to perform habitat or LC mapping using remote sensing data. They include supervised (Chan et al., 2012; Walker et al., 2010; Féret and Asner, 2012; Longépé et al., 2011; Vyas et al., 2011) or unsupervised (Muad and Foody, 2012; Mwita et al., 2013) classification techniques. In cases where prior expert knowledge is available in the form of explicit rules, rule-based approaches may be employed to incorporate such information in the classification process (Kumar and Patnaik, 2013; Lucas et al., 2011a; Evans et al., 2010). Although efficient in a number of classification tasks, such methods may prove inadequate in handling uncertainty, due

☆ This paper has been recommended for acceptance by Edwin Hancock
* Corresponding author at: Information Technologies Institute, P.O. Box 60361, 6th km Xarilaou – Thermi, 57001, Thessaloniki, Greece. Tel: +30 2311 257711; fax: +30 2310 474128.

E-mail addresses: z.petrou10@imperial.ac.uk, zpetrou@iti.gr (Z.I. Petrou).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.11.002

to missing information, and inaccuracies, caused by noisy data or vague rules. Such problems are common in remote sensing applications, where, on the one hand, necessary sensor or ancillary data may be unavailable for certain landscapes, and, on the other hand, noise affliction may be introduced to the data, during acquisition and processing, such as registration, quantization and topographic and atmospheric correction. Despite various attempts in LC to habitat conversion (Adamo et al., 2013; Tomaselli et al., 2013), no previous framework has been suggested employing evidential reasoning for handling uncertainty and missing information.

Dempster–Shafer (DS) theory, a mathematical theory of evidence, has been broadly used for information fusion and handling uncertainty and missing data (Saffiotti, 1994; Yager, 1987). In pattern recognition, DS theory has been used, principally, to combine results generated from multiple classifiers, even resulting in different classes (Ahmadzadeh and Petrou, 2003), and, less frequently, as the core of individual rule-based inference engines.

Combining the results of different classifiers, DS has been applied in numerous fields, including three-dimensional object reconstruction (Díaz-Más et al., 2010), industrial parts inspection (Osman et al., 2011; Basir and Yuan, 2007; Kaftandjian et al., 2003), desertification risk and water quality assessment (Ahmadzadeh and Petrou, 2001; Aminravan et al., 2011), medical imaging (Bloch, 1996), road extraction from satellite images (Cleynenbreugel and Osinga, 1991), speaker identification (Chen et al., 1997) and optical character recognition (Rogova, 1994), providing results outperforming those derived from the individual classifiers. Fuzzy logic has been incorporated in various classifier fusion tasks using DS theory, in order to deal with data and rule vagueness (Deng et al., 2011; Zhang et al., 2011; Deng et al., 2010; Wu, 2009). In landscape characterization, fuzzy DS frameworks have been used to combine sensor data (Sarkar et al., 2005; Pinz et al., 1996) or contextual information (Laha et al., 2006) to improve classification.

Individual rule-based classifiers have also been designed based on the combination of DS theory and fuzzy logic (Liu et al., 2004; Parikh et al., 2001; Yager, 1992). In landscape monitoring, DS theory with fuzzy sets has been used for LC classification in agricultural (Lein, 2003) and complex landscapes (Cayuela et al., 2006). The ability of the theory to incorporate multiple sources of information has been demonstrated by Franklin et al. (2002), where a classifier based on DS was compared with a conventional maximum likelihood classifier unable to incorporate all available ancillary information, thus resulting in significantly lower accuracy in discriminating habitat classes, compared with the DS-based classifier.

The contribution of this study lies in the proposal of a robust methodology based on DS theory and incorporating fuzzy logic for habitat mapping using remote sensing data. The proposed methodology builds on a pre-existing LC map and converts it into a habitat map, incorporating domain expert rules and additional information. Different fuzzification approaches are proposed and introduced to the DS theory framework to deal with inaccurate rules provided by domain experts or noise afflicted data. The objective is to increase the framework robustness and make it readily applicable and transferable to similar landscapes in different locations. The flexibility of the framework in handling composite classes when adequate information for the discrimination of single classes is missing is also studied.

2. Application field and methods

2.1. Land cover to habitat mapping

The application field of the proposed fuzzy evidential reasoning classification approach lies in the area of ecological monitoring, biodiversity assessment and ecosystem preservation. In particular, the developed classification framework deals with the fusion of diverse information and rules provided by domain experts for habitat mapping, based on a LC map and through the use of remote sensing data.

The employed LC map is expressed in the Land Cover Classification System (LCCS) taxonomy, proposed by the United Nations Food and Agriculture Organization (FAO) (di Gregorio and Jansen, 2005). LCCS classes are organized in eight main categories, depending on whether the area element of interest is vegetated or not, aquatic or terrestrial and managed or artificial or (semi-) natural. Classes are further refined with the inclusion of additional information, such as life form (e.g., woody or herbaceous vegetation), vegetation coverage, leaf type and phenology (e.g., broadleaved, evergreen, deciduous), canopy height, soil type and lithology. LCCS taxonomy has been proposed as a generic framework able to describe adequately any LC class globally, while has been recently recognized as the most appropriate LC taxonomy to serve as basis for habitat mapping (Tomaselli et al., 2013).

Habitat classification in this study is expressed in a recently developed taxonomy, the General Habitat Categories (GHC), based on life and non-life forms (Bunce et al., 2008). GHC classes are organized in five main categories, namely: (i) urban, (ii) cultivated, (iii) sparsely vegetated, (iv) trees and shrubs and (v) herbaceous vegetation. Various classes belong in each category, based on life or non-life forms present in a studied area element, leaf properties, height of canopy, etc. The classes were initially defined to link in situ and remote sensing observations, thus facilitating their extraction through data derived from satellite or airborne sensors.

Based on a LCCS map, information from remote sensing and ancillary data is combined using evidential reasoning to perform habitat classification of a study area, using expert decision rules (Kosmidou et al., 2014; Adamo et al., 2013). DS theory is employed to handle uncertainty and multiple classes, when adequate information for the discrimination among single classes is unavailable, and to provide a framework for embedding fuzzy logic to counteract for noisy data and increase framework robustness and transferability.

2.2. Dempster–Shafer theory principles

DS theory, introduced by Dempster (1967) and Shafer (1976), is a mathematical theory of evidence, considered as a generalised form of the Bayesian theory of subjective probability. It is popular in rule-based expert systems, mainly because of its ability in handling uncertainty, lack of information and vague rules leading to composite events (Ahmadzadeh and Petrou, 2003).

To each individual event, or set of events, *belief* and *plausibility* values are assigned, defining a *belief interval*. Belief on an event expresses the degree of confidence that the event holds, based on supporting evidence. Its plausibility value reflects the highest confidence on an event if all missing information were to support its validity. The difference between the plausibility and belief of a single or composite event expresses its uncertainty. When no uncertainty exists, plausibility and belief values coincide.

One of the principal concepts in DS theory is the basic probability assignment function, m, describing the degree an event, A, from the set of all possible events, or *frame of discernment* Θ, is supported with evidence. A can be a single event or a set of two or more single events; m values assigned to the latter indicate lack of adequate evidence to distinguish among the single events. The m values assigned to all subsets of Θ sum up to 1.

The belief function, $bel : 2^{\Theta} \rightarrow [0,1]$, of a set $A \subseteq \Theta$, is defined as the summation of the m values of all subsets of A, i.e.,

$$bel(A) = \sum_{X \subseteq A} m(X), \quad \text{for all } A \subseteq \Theta. \tag{1}$$