ISPRS Journal of Photogrammetry and Remote Sensing 130 (2017) 63–80

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A higher order conditional random field model for simultaneous classification of land cover and land use

Lena Albert, Franz Rottensteiner *, Christian Heipke

Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, D-30167 Hannover, Germany

ARTICLE INFO

Article history:
Received 10 August 2016
Received in revised form 2 April 2017
Accepted 11 April 2017

Keywords:
Contextual classification
Land cover
Land use
Conditional random fields
Higher order potential
Aerial imagery

ABSTRACT

We propose a new approach for the simultaneous classification of land cover and land use considering spatial as well as semantic context. We apply a Conditional Random Fields (CRF) consisting of a land cover and a land use layer. In the land cover layer of the CRF, the nodes represent super-pixels; in the land use layer, the nodes correspond to objects from a geospatial database. Intra-layer edges of the CRF model spatial dependencies between neighbouring image sites. All spatially overlapping sites in both layers are connected by inter-layer edges, which leads to higher order cliques modelling the semantic relation between all land cover and land use sites in the clique. A generic formulation of the higher order potential is proposed. In order to enable efficient inference in the two-layer higher order CRF, we propose an iterative inference procedure in which the two classification tasks mutually influence each other. We integrate contextual relations between land cover and land use in the classification process by using contextual features describing the complex dependencies of all nodes in a higher order clique. These features are incorporated in a discriminative classifier, which approximates the higher order potentials during the inference procedure. The approach is designed for input data based on aerial images. Experiments are carried out on two test sites to evaluate the performance of the proposed method. The experiments show that the classification results are improved compared to the results of a non-contextual classifier. For land cover classification, the result is much more homogeneous and the delineation of land cover segments is improved. For the land use classification, an improvement is mainly achieved for land use objects showing non-typical characteristics or similarities to other land use classes. Furthermore, we have shown that the size of the super-pixels has an influence on the level of detail of the classification result, but also on the degree of smoothing induced by the segmentation method, which is especially beneficial for land cover classes covering large, homogeneous areas.

1. Introduction

1.1. Motivation and goals

Geospatial land use databases contain important information with high benefit for several users, especially in the field of urban management and planning. The number of possible applications of such data increases with a higher level of detail, both in terms of the size of the geometrical entities as well as the diversity of land use classes. Because of fast changes of the land use due to urban growth and land use conversion, such geospatial databases become outdated quickly. The update process requires enormous efforts (Champion, 2007), so that the automation of this task on the basis of remote sensing data is desirable.

In contrast to *land cover*, which describes the physical material of the earth's surface (e.g. *grass*, *asphalt*), *land use* reveals the socio-economic function of a piece of land (e.g. *residential*, *agricultural*). Land use objects are typically composed of different land cover elements, which may be arranged in complex structures. On the other hand, land cover of a specific type can be a part of different land use objects. Thus, land cover and land use classification based on remote sensing data are tasks that pursue different objectives (Barnsley and Barr, 2000). Whereas land cover classification focuses on the assignment of class labels to (frequently small) image sites, the goal of land use classification is to assign a land use label to larger spatial entities which form a functional unit and which are typically represented by polygonal objects in a land use database. Land cover can be derived directly from spectral

* Corresponding author.
E-mail addresses: albert@ipi.uni-hannover.de (L. Albert), rottensteiner@ipi.uni-hannover.de (F. Rottensteiner), heipke@ipi.uni-hannover.de (C. Heipke).

http://dx.doi.org/10.1016/j.isprsjprs.2017.04.006
0924-2716/© 2017 Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS).

characteristics of remote sensing data, but this is more difficult for land use; the classification of land use additionally requires information about the composition and arrangement of different land cover elements within a land use object. As a consequence, a procedure consisting of a sequence of two classification steps is frequently applied to obtain land use information from remote sensing imagery, e.g. (Albert et al., 2014a; Hermosilla et al., 2012). In a first step, land cover information is derived by a classification of remote sensing data. The second step consists of a land use classification, often based on segments or objects from a geospatial database, in which some of the features are derived from the results of the first step. It is the main drawback of this approach that wrong decisions taken during land cover classification cannot be reversed at later stages, which can easily lead to misclassifications of land use.

An important source of information to be tapped for land cover and land use classification is *context*. First, there are *spatial* dependencies between neigbouring sites. Both for land cover and for land use, certain classes are more likely to occur next to each other than others. For instance, a *residential* land use object is usually located next to a land use object of the class *street*. On the other hand, neighbouring land cover sites are typically small and, thus, likely to belong to the same class; to give an example, in a forest, a land cover site belonging to the class *tree* is usually surrounded by other trees rather than buildings or sealed area. Second, there is a mutual *semantic* dependency between land cover and land use. The classification of land use requires information about land use, but the knowledge about the current land use facilitates the determination of land cover, because, for instance, the land cover *tree* is more likely to occur inside a land use object of class *forest* than inside a *traffic* object. This additional hint becomes particularly beneficial for high-resolution aerial image data, where the appearance of many land cover classes is heterogeneous, making land cover classification more challenging.

In this paper, we present an approach for the simultaneous classification of land cover and land use that considers semantic as well as spatial context. Land cover classification is carried out at the level of super-pixels (Achanta et al., 2012), whereas land use is determined at the level of objects from a geospatial database, represented by polygons that are assumed to be correct. The rationale for this assumption, which also corresponds to the application scenario in our experiments, is that in regions such as Central Europe, the object boundaries in geospatial data bases are usually related to property boundaries which are kept up-to-date by governmental surveying authorities, whereas property owners are not obliged to report changes in land use, so that the update of the semantic information contained in such a database lags behind. Thus, our approach can be seen as a contribution to the automatic derivation of this semantic information from aerial imagery. If higher semantic accuracy is required, it can be the first step of a semi-automatic scheme for database updating, in which land use objects for which the classification result is incompatible with the information contained in the database are highlighted to a human operator for visual inspection, e.g. (Helmholz et al., 2014).

The method presented in this paper is supervised, i.e. it requires training data for both land cover and land use objects. The statistical framework for contextual classification is given by Conditional Random Fields (CRF), which in the standard formulation require unary potentials, typically the output of a local classifier, and pairwise potentials for the context model (Kumar and Hebert, 2006). Our method requires aerial images and derived products such as digital surface models (DSM), as well as the polygons from a geospatial database as input. Whereas the standard CRF framework is applied for modelling spatial context, the dependencies between the land cover and the land use layers of the CRF require the formulation of a high-order potential and a specific inference method for obtaining the optimal configuration of class labels in both layers. The scientific contribution of our method can be summarized as follows:

- *Simultaneous contextual classification of land cover and land use:* Unlike existing two-step approaches (Hermosilla et al., 2012; Walde et al., 2014; Albert et al., 2014a), where the semantic dependency between land cover and land use is only considered in one direction (i.e. from land cover to land use), our method integrates land cover and land use classification in a unified approach, allowing both tasks to interact in the classification procedure. This avoids early decisions and is expected to improve the classification accuracy, in particular for the land cover layer.
- *CRF with higher order potentials:* Standard CRF-based models only consider interactions between pairs of objects to be classified (Kumar and Hebert, 2006). However, such a model is not appropriate in our case, where one (large) land use object may interact with many (small) land cover objects. Thus, we propose a new type of higher order potential in our CRF-based approach which we expect to provide a better model of the complex statistical dependencies between land cover and land use and which, unlike other models for higher order potentials (Kohli et al., 2009), can also deal with different class structures.
- *Inference for CRF with higher order potentials:* Unlike existing techniques placing restrictions on the formulation of these potentials (Kohli et al., 2009), our method can cope with a generic model for the complex interactions between land cover and land use. Inference is carried out in an iterative way. In each iteration, we determine a partial solution for one of the two layers (land cover, land use), keeping current optimal state in the other layer fixed. As a consequence, the higher order potential is simplified to a simple unary potential based on the output of a local classifier, allowing the application of standard inference techniques for CRF. This procedure is inspired by Roig et al. (2011), but we use a different and more complex model for the higher order potential.
- *Semantic context features:* The local classifier used to represent the higher order potential in the approximate inference technique is based on a comprehensive set of contextual features that expresses the local arrangement of land cover primitives in a land use object and vice versa. These features are computed on the basis of the partial solution that is an intermediate result in the inference procedure, considering the current beliefs for all classes as weights.

The remainder of this paper is structured as follows. We start with a review of related work in Section 1.2. Our approach for land cover and land use classification is presented in Section 2, whereas Section 3 describes the experimental evaluation of our approach based on two test sites in Germany. Conclusions and an outlook are given in Section 4.

1.2. Related work

We start this review of related work by a discussion of papers that deal with different strategies for land use classification from high-resolution remote sensing data, focussing on the overall strategy and the way in which land cover is integrated into the process. In the second part of this review, we will discuss methodological aspects related to context-based classification, the use of multilayer CRF, and CRF with higher order potentials.

High-resolution satellite or aerial images are the data source most frequently used for deriving land use information (Walde et al., 2014; Hermosilla et al., 2012; Albert et al., 2014a). The definition of the entities to be classified depends on the application.