ARTICLE IN PRESS

ISPRS Journal of Photogrammetry and Remote Sensing xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Learning a constrained conditional random field for enhanced segmentation of fallen trees in ALS point clouds

Przemyslaw Polewski [a,b,*], Wei Yao [a], Marco Heurich [c], Peter Krzystek [a], Uwe Stilla [b]

[a] *Dept. of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany*
[b] *Photogrammetry and Remote Sensing, Technische Universitaet Muenchen, 80333 Munich, Germany*
[c] *Dept. of Conservation and Research, Bavarian Forest National Park, 94481 Grafenau, Germany*

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 1 March 2017
Accepted 3 April 2017
Available online xxxx

Keywords:
CRF
Segmentation
Fallen tree detection
LIDAR

ABSTRACT

In this study, we present a method for improving the quality of automatic single fallen tree stem segmentation in ALS data by applying a specialized constrained conditional random field (CRF). The entire processing pipeline is composed of two steps. First, short stem segments of equal length are detected and a subset of them is selected for further processing, while in the second step the chosen segments are merged to form entire trees. The first step is accomplished using the specialized CRF defined on the space of segment labelings, capable of finding segment candidates which are easier to merge subsequently. To achieve this, the CRF considers not only the features of every candidate individually, but incorporates pairwise spatial interactions between adjacent segments into the model. In particular, pairwise interactions include a collinearity/angular deviation probability which is learned from training data as well as the ratio of spatial overlap, whereas unary potentials encode a learned probabilistic model of the laser point distribution around each segment. Each of these components enters the CRF energy with its own balance factor. To process previously unseen data, we first calculate the subset of segments for merging on a grid of balance factors by minimizing the CRF energy. Then, we perform the merging and rank the balance configurations according to the quality of their resulting merged trees, obtained from a learned tree appearance model. The final result is derived from the top-ranked configuration. We tested our approach on 5 plots from the Bavarian Forest National Park using reference data acquired in a field inventory. Compared to our previous segment selection method without pairwise interactions, an increase in detection correctness and completeness of up to 7 and 9 percentage points, respectively, was observed.

1. Introduction

Probabilistic graphical models are a class of methods for compactly describing complex, structured probability distributions (Koller and Friedman, 2009). Although these techniques originated in the field of statistical physics, over the last several decades they have been gradually adopted and further developed within the machine learning and computer vision communities. Particularly the class of undirected (acausal) graphical models, such as Markov Random Fields (MRF) has been successfully applied to classical computer vision problems including image denoising (Besag, 1986), depth map estimation (Tseng and Lai, 2011), classification (Berthod et al., 1996), semantic segmentation (Márquez-Neila et al., 2014) and others. Graphical models use a graph-based representation of the target probability distribution to encode the independence structure among the random variables. The joint probability distribution can be expressed compactly as a product over factors associated with cliques existing within the graph, leading to a model which is much more computationally tractable compared to the naive approach of explicitly modelling the cooccurrence of *all* variables simultaneously. Regarding classification tasks, a special kind of undirected graphical model, *Conditional Random Fields* (CRF) (Lafferty et al., 2001), has attained prominence. In contrast to standard MRFs, the CRF describes the *conditional* distribution of the dependent variable (class labels of classified objects) given the values of independent variables (the objects' features). This obviates the need for modeling possibly complex interactions between independent variables (e.g. image pixel colors) and allows to directly characterize the discriminative

* Corresponding author at: Dept. of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany.

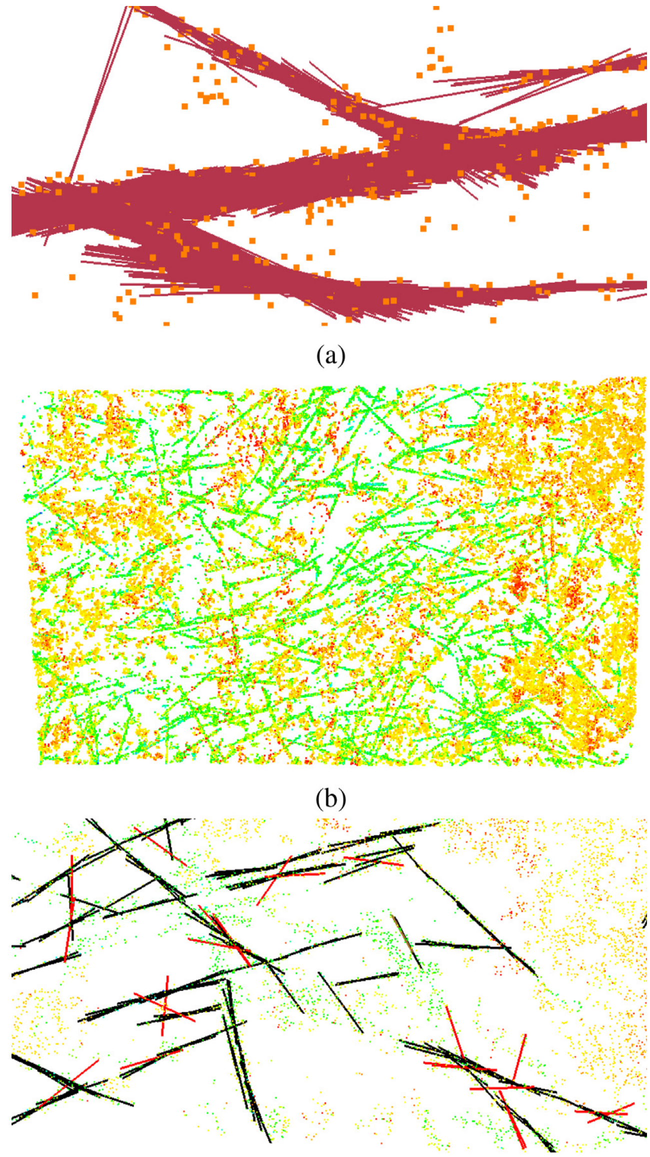
E-mail addresses: polewski@hm.edu (P. Polewski), yao@hm.edu (W. Yao), marco.heurich@npv-bw.bayern.de (M. Heurich), krzystek@hm.edu (P. Krzystek), stilla@tum.de (U. Stilla).

http://dx.doi.org/10.1016/j.isprsjprs.2017.04.001

Please cite this article in press as: Polewski, P., et al. Learning a constrained conditional random field for enhanced segmentation of fallen trees in ALS point clouds. ISPRS J. Photogram. Remote Sensing (2017), http://dx.doi.org/10.1016/j.isprsjprs.2017.04.001

class label probability of interest. The strength of graphical models lies in their ability to express prior information about the spatial relationships between parts of the scene in images, *e.g.* 'the sky is *above* the sea' or 'the book lies *on* the table'. This is a consequence of the fact that within a CRF, the class labels of adjacent objects influence each other, resulting in *contextual classification*, as opposed to the usual classification paradigm where every object's label is assigned independently. Due to their attractive properties, CRFs have attracted the attention of researchers from the remote sensing community. Concerning raster images, Wegner et al. (2011) used a CRF binary with a regular neighborhood structure for enhanced building detection from InSAR data combined with optical imagery. Also, some authors propose to build graphical models on arbitrarily shaped superpixels, which results in irregular neighborhoods. Yang and Förstner (2011) define a multi-scale CRF with pairwise interactions on superpixels resulting from mean-shift segmentation for classifying man-made objects in optical images. In a similar setting of urban area classification, Volpi and Ferrari (2015) construct a multiple-ring neighborhood on the superpixels based on the distances between their centers. They then utilize a structured SVM to learn the binary CRF's interaction potential weights, yielding a prior on the localized pairwise class interactions. In the context of road network extraction, Wegner et al. (2015) depart from the purely binary model in favor of a higher-order CRF. They show that the higher-order potentials enforcing within-clique label consistency are suitable for detecting roads which appear as thin, elongated structures. Some effort has also been dedicated to transfer CRF-based methods to the 3D point cloud domain. This was done mainly with respect to urban area classification in ALS (Niemeyer et al., 2014) and MLS (Weinmann et al., 2015) point clouds. Note that due to the characteristics of the data acquisition process, it is no longer possible to define a regular neighborhood as was the case with raster images. Instead, a k-nearest neighbor or fixed radius based neighborhood of a cylindrical or spherical shape is usually employed.

In this work, we propose a constrained conditional random field as a preprocessing step before semantic segmentation of objects from ALS point clouds. Specifically, we target the task of segmenting individual fallen tree stems in forest ALS scenes. Detection of dead wood using remote sensing techniques is an active research topic due to its importance in forest nutrient cycles and its contribution to carbon stock (Nyström et al., 2014; Polewski et al., 2015b). Moreover, up to one third of plant and animal species in temperate forests depend on fallen and standing dead wood for their survival (Stokland et al., 2012). Since forests make up some of Earth's most rich and diverse habitats, this highlights the importance of dead wood for biodiversity of species. Due to the high costs of conducting field surveys and their intrinsic low spatial coverage, remote sensing methods for detecting dead trees are highly desirable. From a pattern recognition perspective, this task is quite challenging owing to the presence of ground vegetation, as well as the fact that the fallen trees may lie in close proximity and nontrivial spatial configurations (see Fig. 1b). We extend our original approach for fallen tree segmentation described in Polewski et al. (2015a), where an in-depth discussion of competing methods is available. Our approach is bottom-up: we first detect short, equal-length primitives (stem segments), and subsequently merge them to form entire objects (fallen trees). We generate a large number of highly overlapping segment candidates. These candidates are then classified to find true stem segments. Due to the overwhelming number of candidates (see Fig. 1a for an example), only a small percentage of them may be selected for later merging to ensure computational feasibility. Therefore, a method for jointly classifying the segment candidates and selecting an appropriate subset for merging is necessary. In our prior work, this was carried out in two separate steps: the segments were classified independently (without considering neighbors), and then the representative items were chosen from the positively labeled candidates based only on the laser points which they cover. In particular, no influence was given to the spatial relationships of adjacent items, which at times led to the formation of divergent segment groups with visually implausible conformations (see Fig. 1c). This observation can be quantitatively confirmed by our previous findings, which indicate that improving precisely the segment classification step in the processing pipeline can yield the greatest increase in detection accuracy (Polewski et al., 2015a). We hypothesize that by using a graphical model, we can incorporate *context* into the selection process, providing candidates for merging which better match each other (higher collinearity of adjacent segments) and contain less redundant information (spatial overlap of neighbors), thus making the merging process simpler. The main contribution



(a)

(b)

(c)

Fig. 1. (a) Point cloud with generated stem segment candidates. Even a small point grouping results in thousands of highly overlapping segments. (b) Raw input cloud (after DTM filtering) with visible overlapping stems and ground vegetation, colored by reflection intensity (green is high, red is low). (c) Stem segments selected without considering pairwise interactions, erroneously chosen items marked with red color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Polewski, P., et al. Learning a constrained conditional random field for enhanced segmentation of fallen trees in ALS point clouds. ISPRS J. Photogram. Remote Sensing (2017), http://dx.doi.org/10.1016/j.isprsjprs.2017.04.001