ISPRS Journal of Photogrammetry and Remote Sensing 132 (2017) 102–118

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds

Loïc Landrieu [a,*], Hugo Raguet [b], Bruno Vallet [a], Clément Mallet [a], Martin Weinmann [a,c,*]

[a] Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mande, France
[b] Institut de Mathématiques de Marseille, Aix-Marseille Université, 39 rue Frédéric Joliot-Curie, F-13453 Marseille, France
[c] Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Englerstraße 7, D-76131 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 7 April 2017
Received in revised form 25 August 2017
Accepted 26 August 2017

Keywords:
Semantic labeling, Point cloud
Classification
Regularization
Scene interpretation
Structured optimization
Graphical models

ABSTRACT

In this paper, we introduce a mathematical framework for obtaining spatially smooth semantic labelings of 3D point clouds from a pointwise classification. We argue that structured regularization offers a more versatile alternative to the standard graphical model approach. Indeed, our framework allows us to choose between a wide range of fidelity functions and regularizers, influencing the properties of the solution. In particular, we investigate the conditions under which the smoothed labeling remains probabilistic in nature, allowing us to measure the uncertainty associated with each label. Finally, we present efficient algorithms to solve the corresponding optimization problems.

To demonstrate the performance of our approach, we present classification results derived for standard benchmark datasets. We demonstrate that the structured regularization framework offers higher accuracy at a lighter computational cost in comparison to the classic graphical model approach.

1. Introduction

Due to the steadily increasing availability of geospatial information, their automated analysis has become a topic of major interest in photogrammetry, remote sensing, robotics, and computer vision. In particular, the representation of a scene in the form of a 3D point cloud and a subsequent semantic interpretation of this point cloud serve as the basis for many applications, such as scene modeling, autonomous navigation, or object detection. For instance, the analysis of 3D point cloud data acquired within urban environments benefits from a semantic labeling since the latter can be exploited for the creation of large-scale city models (Lafarge and Mallet, 2012) or urban accessibility diagnosis (Serna and Marcotegui, 2013).

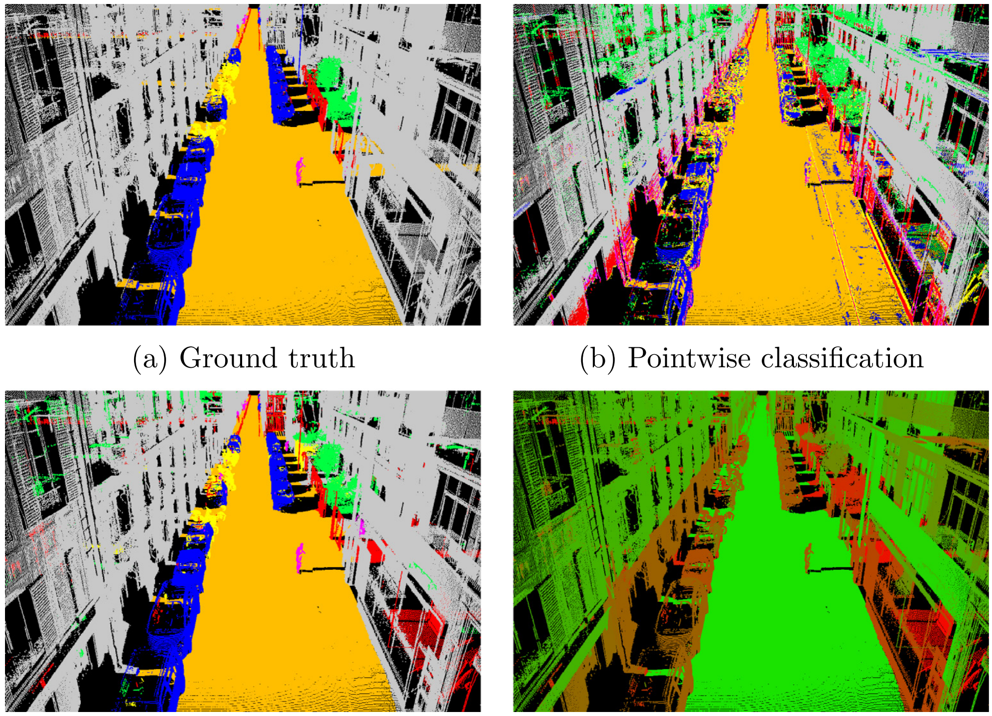
The semantic interpretation typically consists in assigning a semantic label (e.g. *building*, *ground* or *vegetation*) to each point of the considered 3D point cloud, as shown in Fig. 1a–c. This assignment can be accompanied by an estimation of the confidence of the labeling of each point in the form of a probability distribution over the labels, as illustrated in Fig. 1d. Such a certainty assessment can prove useful when either the precision or the recall of the classification is more crucial for a given application. In the case of autonomous navigation for example, merely the possibility of an obstacle can be enough to alter course, and a probabilistic occupancy map is preferred to a binary one (Moravec and Elfes, 1985; Hornung et al., 2013). In the case of reconstruction tasks which necessitate the removal of a specific semantic class beforehand (Clode et al., 2004), precision is the focus in order to not accidentally remove relevant information. In a context of active learning, an assessment of the labeling certainty can guide an operator to the areas of the point cloud in which the classification is least certain, as they are more prone to be labeled incorrectly and might require manual re-labeling (Jing et al., 2004). The nature of the assignment, either a probability or a label, depends on the choice of the method used for inference.

The semantic labeling of 3D point clouds has been addressed by numerous investigations (Munoz et al., 2009; Shapovalov et al., 2010; Mallet et al., 2011; Niemeyer et al., 2014; Xu et al., 2014; Guo et al., 2015; Weinmann et al., 2015a; Weinmann, 2016; Hackel et al., 2016b). However, this problem remains challenging due to the irregular point sampling, occlusions, and the complexity of the considered scenes, which induce a loose yet meaningful structure to the data. Furthermore, the consideration of larger scenes typically results in a huge amount of data and efficient

* Corresponding authors at: Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mande, France (L. Landrieu and M. Weinmann).

E-mail addresses: loic.landrieu@ign.fr (L. Landrieu), hugo.raguet@gmail.com (H. Raguet), bruno.vallet@ign.fr (B. Vallet), clement.mallet@ign.fr (C. Mallet), martin.weinmann@kit.edu (M. Weinmann).

http://dx.doi.org/10.1016/j.isprsjprs.2017.08.010



(a) Ground truth (b) Pointwise classification

(c) Regularized classification (d) Confidence map

Fig. 1. Visualization of a 3D point cloud labeling for a part of the Paris-rue-Cassette Database (Vallet et al., 2014). In (a), (b), and (c), the color encoding addresses the classes *Façade* (gray), *Ground* (orange), *Cars* (blue), *2-Wheelers* (yellow), *Road Inventory* (red), *Pedestrians* (magenta) and *Vegetation* (green). In (d), the confidence is represented from green to red: confident uncertain. Remark that misclassifications in (c) correspond to the least confident areas in (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

techniques for semantic labeling are therefore desirable. To foster research regarding semantic labeling of 3D point cloud data, a variety of benchmark datasets acquired within urban environments have been released (Munoz et al., 2009; Serna et al., 2014; Vallet et al., 2015; Hackel et al., 2016b).

The straightforward approach for semantically labeling a considered 3D point cloud consists in extracting a variety of features for all points, concatenating these features to a feature vector, which is then classified with a classifier trained on representative training examples. This strategy has for instance been followed in the framework introduced by Weinmann et al. (2015a), in which a diversity of distinctive geometric low-level features serve as input for a standard supervised classification scheme. While this rather simple approach already yields good classification rates due to the use of distinctive features, the visualization of the classified 3D point cloud reveals a noisy behavior as each point is treated individually by only considering the respective feature vector for classification. To illustrate this effect, we provide a ground truth labeling for a considered 3D point cloud in Fig. 1a and the result of a pointwise classification based on distinctive geometric low-level features in Fig. 1b. Considering the ground truth labeling, one can observe a high spatial regularity of the labeling. Indeed, as the number of 3D points far exceeds the number of objects in the scene, it is reasonable to assume that most 3D points are surrounded by points of the same label.

To impose spatial smoothness on this classification result, contextual information among neighboring 3D points is typically taken into account. For this purpose, the spatial structure of a 3D point cloud can be captured by a graph encoding the adjacency relationship between 3D points. Thereby, the adjacency relationship can be derived from the local neighborhood of each 3D point (Weinmann et al., 2015b), pre-segmentations (Niemeyer et al., 2016), or super-voxel structures (Lim and Suter, 2009). Based on the defined adjacency relationship, a context model is typically derived in the form of a graphical model, e.g. a Markov random field (MRF) (Munoz et al., 2009; Shapovalov et al., 2010; Lu and Rasmussen, 2012; Najafi et al., 2014) or its discriminative counterpart, the conditional random field (CRF) (Niemeyer et al., 2011; Niemeyer et al., 2014; Schmidt et al., 2014; Weinmann et al., 2015b). As a consequence of imposing spatial smoothness on the derived labeling, the corresponding classification results are often significantly improved as can be observed in Fig. 1c. However, the choice of the inference strategy (marginal, maximum-a-posteriori) will have a profound impact on the precision and nature of the solution (probabilistic or labeling), as well as the computation times.

In this paper, we propose to consider the problem of spatially smoothing semantic labelings of 3D point clouds from a structured regularization perspective. While using such a model results in a loss of interpretability compared to a probabilistic approach, it offers several advantages. In particular, the structured regularization approach allows:

- the choice from a wide range of fidelity functions and regularizers,[1]
- the choice to retain or not the probabilistic aspect of the input labeling, and
- the use of fast solving algorithms, compared to slow and memory-intensive message-passing algorithms.

After briefly introducing the used notation and the formal description of the considered problem, we summarize related work

[1] Notably, this framework allows us to express the graphical model approach as a special instance.