Spatial Statistics 14 (2015) 532–545

Contents lists available at ScienceDirect

Spatial Statistics

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

Linear street extraction using a Conditional Random Field model

Padraig Corcoran [a,*], Peter Mooney [b], Michela Bertolotto [c]

[a] School of Computer Science and Informatics, Cardiff University, Cardiff, Wales, UK

[b] Department of Computer Science, Maynooth University, Ireland

[c] School of Computer Science and Informatics, University College Dublin, Dublin, Ireland

ARTICLE INFO

Article history:
Received 11 May 2015
Accepted 14 October 2015
Available online 24 October 2015

Keywords:
Conditional Random Field (CRF)
Street network
Pattern extraction

ABSTRACT

A novel method for extracting linear streets from a street network is proposed where a linear street is defined as a sequence of connected street segments having a shape similar to a straight line segment. Specifically a given street network is modeled as a Conditional Random Field (CRF) where the task of extracting linear streets corresponds to performing learning and inference with respect to this model. The energy function of the proposed CRF model is *submodular* and consequently exact inference can be performed in polynomial time. This contrasts with traditional solutions to the problem of extracting linear streets which employ heuristic search procedures and cannot guarantee that the optimal solution will be found. The performance of the proposed method is quantified in terms of identifying those types or classes of streets which generally exhibit the characteristic of being linear. Results achieved on a large evaluation dataset demonstrate that the proposed method greatly outperforms the aforementioned traditional solutions.

1. Introduction

The automated extraction of geometrical patterns from street networks represents an important component in many geo-spatial applications (Yang et al., 2010). For example, Weiss and Weibel (2014)

* Corresponding author.
E-mail address: corcoranp@cardiff.ac.uk (P. Corcoran).

http://dx.doi.org/10.1016/j.spasta.2015.10.003
2211-6753/

used a measure of street network centrality to determine the most significant subset of a street network to be represented when reforming map generalization. Porta et al. (2006) extracted connected groups of street segments and used this as a platform for performing network analysis. In order to align two different street network representations of a common area, Koukoletsos et al. (2012) first extracted geometrical patterns from both representations and subsequently aligned these patterns.

A street network can exhibit a wide spectrum of geometrical patterns, a taxonomy of which was proposed by Marshall (2004). In this work we focus on the extraction of the geometrical pattern of linear streets which may be defined as sequences of street segments which have a shape *similar* to that of a straight line segment. Linear streets are also commonly referred to as strokes (Thomson, 2006; Heinzle et al., 2005; Touya, 2010). Some authors define strokes to be sequences of street segments which exhibit good continiuty, that is where the turning angle between adjacent segments is small (Yang et al., 2011), as opposed to sequences of streets segments which have shape *similar* to that of a line segment. For the purposes of this paper we consider linear and stroke patterns to be equivalent and adopt the latter, as opposed to the former, definition. The extraction of linear streets represents a challenging task for the following reason. Being a linear street is a characteristic of a sequence of street segments. The set of all street segments in a given street network are however interconnected. Therefore in order to extract linear streets, the tasks of segmenting and labelling streets with respect to the characteristic of being linear must be solved simultaneously. This is commonly refereed to as the task of semantic segmentation in the domains of computer vision and robotics (Anand et al., 2012).

Existing methods for extracting linear streets generally employ heuristic search procedures where individual street segments are iteratively expanded or grown to form linear streets. These methods do not guarantee that the optimal solution will be found and no statements regarding the distance from the solution obtained to the optimal solution can be made. In this paper we propose a novel method for extracting linear streets which overcomes these limitations. Specifically we formulate the problem in terms of performing learning and inference with respect to a Conditional Random Field (CRF) which is a type of undirected probabilistic graphical model (Koller and Friedman, 2009). We demonstrate that performing inference with respect to this model corresponds to minimizing a submodular energy function. As a consequence of this fact, the optimal solution can be computed in polynomial time.

The remainder of this paper is structured as follows. In Section 2 we review related works on the extraction of geometrical patterns from spatial data. In Section 3 the proposed model is described. Finally, in Sections 4 and 5 we present results and draw conclusions respectively.

2. Related works

Existing techniques for extracting geometrical patterns from spatial data generally fall into two broad categories corresponding to those which extract patterns relating to buildings and those which extract patterns relating to street networks. We only considered those techniques in the latter category; the interested reader is directed to the following works in the former category (Lüscher et al., 2009; Zhang et al., 2013). A street network can exhibit a wide spectrum of patterns, a taxonomy of which was proposed by Marshall (2004). Ultimately the actual patterns one attempts to extract depends on the intended application. Some of the most common applications are network analysis (Porta et al., 2006), map generalization (Zhou and Li, 2012) and network matching (Koukoletsos et al., 2012).

Existing approaches to the extraction of linear streets employ heuristic search procedures. We now review these methods. Luan and Yang (2010) and Thomson (2006) describe a simple search procedures which iteratively expands a linear street to contain adjacent street segments. Liu et al. (2010) describes a similar search procedure which is initialized using a single street segment and terminates when the linear street in question cannot be expanded any further. The procedure is initialized using every segment in the street network and a set of criterion is used to evaluate if and how an expansion is performed. Yang et al. (2011) and Luan and Yang (2010) propose search procedures for extracting linear streets which are integrated with methods for extracting dual carriageways and complex junctions. This integrations allows the extraction of linear streets across such features. Zhou and Li (2012) present a comparison of different search procedures which differ in