Transportation Research Part C 25 (2012) 18–33

Contents lists available at SciVerse ScienceDirect

Transportation Research Part C

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

On selecting an optimal wavelet for detecting singularities in traffic and vehicular data

Zuduo Zheng *, Simon Washington

Queensland University of Technology, Brisbane, QLD 4001, Australia

ARTICLE INFO

Article history:
Received 6 August 2011
Received in revised form 2 March 2012
Accepted 2 March 2012

Keywords:
Wavelet transform
Singularity detection
Traffic data analysis
The Mexican hat wavelet
Short-time Fourier transform
Oblique cumulative curve

ABSTRACT

Serving as a powerful tool for extracting localized variations in non-stationary signals, applications of wavelet transforms (WTs) in traffic engineering have been introduced; however, lacking in some important theoretical fundamentals. In particular, there is little guidance provided on selecting an appropriate WT across potential transport applications. This research described in this paper contributes uniquely to the literature by first describing a numerical experiment to demonstrate the shortcomings of commonly-used data processing techniques in traffic engineering (i.e., averaging, moving averaging, second-order difference, oblique cumulative curve, and short-time Fourier transform). It then mathematically describes WT's ability to detect singularities in traffic data. Next, selecting a suitable WT for a particular research topic in traffic engineering is discussed in detail by objectively and quantitatively comparing candidate wavelets' performances using a numerical experiment. Finally, based on several case studies using both loop detector data and vehicle trajectories, it is shown that selecting a suitable wavelet largely depends on the specific research topic, and that the Mexican hat wavelet generally gives a satisfactory performance in detecting singularities in traffic and vehicular data.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Irregular structures and transient phenomena (singularities hereafter) of a signal (curve, graph) often contain considerably rich information regarding the phenomenon being studied, as examples edge detection in image processing (Mallat and Zhong, 1992), vibration analysis in machine health monitoring (Peng and Chu, 2004), and irregular structure detection in physics (Argoul et al., 1988), etc. Considering applications in transport, singularities in traffic data may indicate activation or deactivation of a bottleneck, state changes of traffic flow, and occurrences of abnormal events; these in vehicular data may indicate the start of a driver accelerating or decelerating, or changing lanes. Thus, to correctly understand traffic flow characteristics and individual driving behavior, singularities in traffic data and/or vehicular data need to be accurately and reliably detected. Despite these rather straightforward examples, however, random and systematic noise contained in transport related data often makes detection of singularities extremely challenging.

Obviously, there is a tradeoff between dampening noise's impact on the underlying signal and preserving authentic singularities in the data. Unfortunately, commonly-used data processing techniques in traffic engineering (e.g., average, moving average, second-order difference of cumulative summation, and oblique cumulative curve[1]) attenuate the impact of noise at the cost of distorting or totally losing the original information, and potentially important singularities in particular. Such

* Corresponding author. Address: 2 George St., Brisbane, QLD 4001, Australia. Tel.: +61 7 3138 9989; fax: +61 7 3138 1170.
E-mail address: zuduo.zheng@qut.edu.au (Z. Zheng).

[1] Although these methods are data denoising and trend analysis techniques, they are often used essentially as singularity detection tools directly or indirectly in traffic flow characteristic analysis, see e.g., Mauch and Cassidy (2002) and Munoz and Daganzo (2003) among others.

0968-090X/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2012.03.006

practice has significant consequences, especially when the underlying process is subtle, e.g., a sudden reduction in roadway capacity. Findings using inferior methods might identify artifacts created by the methods themselves, leading researchers perhaps to conclusions contradictory to what the underlying data really suggest. Importantly, this matter has not received sufficient attention in the literature, and thus serves as a primary research motivation here.

Data encountered in transport often exhibit variations across time. As a time–frequency decomposition tool, the wavelet transform (WT) is particularly effective for extracting local information from non-stationary time-series by moving the wavelet location and squeezing or dilating the wavelet window. Such a time-scale representation of the original time-series data, which are often noisy and aperiodic, finds plentiful applications in fluid mechanics, engineering testing and monitoring, medicine, finance, geophysics, and network operations to name just a few (Addison, 2002). The WT has also been introduced to traffic engineering and intelligent transportation engineering. Combined with data mining techniques such as clustering, fuzzy logic, and neural networks, the WT has been adopted to investigate various traffic-related issues, such as automatic detection of freeway incidents (Adeli and Samant, 2000; Ghosh-Dastidar and Adeli, 2003; Karim and Adeli, 2002, 2003; Samant and Adeli, 2000, 2001), traffic features around freeway work zones (Adeli and Ghosh-Dastidar, 2004; Ghosh-Dastidar and Adeli, 2006), traffic flow forecasting (Boto-Giralda et al., 2010; Jiang and Adeli, 2005; Vlahogianni et al., 2007; Xie et al., 2007), and traffic pattern recognition (Jiang and Adeli, 2004; Vlahogianni et al., 2008). These pioneering studies have demonstrated the potential of WTs in analyzing non-stationary or noisy traffic data. Recently, Zheng et al. (2011a) demonstrated WT's capabilities of analyzing important features related to bottlenecks and traffic oscillations in a systematic manner. Furthermore, using WTs enabled the identification of the origins of stop-and-go driving and the measurement of microscopic features of wave propagation (Zheng et al., 2011b).

In applying WTs to solve practical problems, however, there are hundreds of different wavelets, and it is also quite straightforward to design a new wavelet for a particular research or practical question. As a consequence of the wide range of readily available and newly derived wavelets, researchers are free to select a WT without a reasoned justification or explanation. As a general rule, most WTs perform well if visual verification is satisfactory for the research purposes at hand. However, if more exacting precision is needed, such as microscopic features of traffic flow, selecting a different wavelet may produce significantly different results, suggesting that justification of wavelet selection is needed. Zheng et al. (2011a,b) neither showed mathematically why WT is a powerful tool in detecting singularities in traffic and vehicular data nor discussed how to select a suitable wavelet. These two important topics are closely related. Without a sound understanding of the mathematical features of wavelets, selecting a good wavelet for a particular research application is problematic at best.

This paper aims to fill this gap and provide sufficient and clear guidance on how and why to prefer a particular wavelet to others. Towards this end, the remainder of this paper is organized as follows. Based on a numerical experiment, Section 2 demonstrates shortcomings of several data processing techniques that are widely used in traffic engineering to detect singularities. Section 3 provides a theoretical background of WT essential for understanding its capability for detecting singularities of non-stationary signals. Section 4 first presents criteria of selecting a suitable wavelet, then uses a numerical experiment to demonstrate how to select a suitable wavelet in the context of traffic engineering. Finally performance of the selected wavelet is further verified using several case studies. Section 5 discusses conclusions and future research.

2. Traffic data processing techniques

In this section the performances of popular techniques for analyzing traffic data in both the time and frequency domains are discussed and compared with that of WT using numerical simulation, with the intent to uniquely compare these techniques' performances and to underscore their advantages and disadvantages. Through such a comparative analysis, undesirable consequences of using these popular data processing techniques are demonstrated and superiority of WT is confirmed.

The use of numerical simulation enables the objective comparison of candidates' performances with the ground truth that is unattainable using field data (Cheng and Washington, 2005; Washington and Cheng, 2008). Using the same motivation, a numerical experiment is also employed in selecting a suitable wavelet, as is discussed later.

2.1. Numerical experiment I (NE I)

To mimic vehicle counts in rush hours collected at a loop detector downstream of a bottleneck, a time series is randomly generated, representing a sample of traffic data with a mean of 12 vehicles and a standard deviation of five vehicles. The simulation period is 3 h with a time resolution of 20 s (so the average flow is 2160 vehicles per hour). In total, 540 data points are generated. Note that the simulated data are stochastic and exhibit white noise properties except for a non-zero mean, as shown in Fig. 1.

2.2. Averaging

To dampen or attenuate statistical noise in the traffic data, the simplest and one of the most commonly used techniques in the time domain is to aggregate data over a certain time interval (e.g., 5 or 15 min), which may be sufficiently effective to reveal long-term trends in traffic patterns. However, the shortcoming of this technique should be apparent: fine resolution information will be smoothed out and inaccurate or distorted information may be obtained because of the decreased time

Download English Version:

https://daneshyari.com/en/article/525209

Download Persian Version:

https://daneshyari.com/article/525209

Daneshyari.com