Transportation Research Part C 56 (2015) 161–176

Contents lists available at ScienceDirect

Transportation Research Part C

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

Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application

Sohail Zangenehpour [a], Luis F. Miranda-Moreno [b,*], Nicolas Saunier [c,1]

[a] Department of Civil Engineering and Applied Mechanics, McGill University, Room 165, Macdonald Engineering Building, 817 Sherbrooke Street West, Montréal, Québec H3A 2K6, Canada

[b] Department of Civil Engineering and Applied Mechanics, McGill University, Room 268, Macdonald Engineering Building, 817 Sherbrooke Street West, Montréal, Québec H3A 2K6, Canada

[c] Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, C.P. 6079, succ. Centre-Ville, Montréal, Québec H3C 3A7, Canada

A R T I C L E I N F O

Article history:
Received 8 April 2014
Received in revised form 2 April 2015
Accepted 2 April 2015

Keywords:
Traffic data collection
Video analysis
Object classification
Object tracking
Road safety

A B S T R A C T

Pedestrians and cyclists are amongst the most vulnerable road users. Pedestrian and cyclist collisions involving motor-vehicles result in high injury and fatality rates for these two modes. Data for pedestrian and cyclist activity at intersections such as volumes, speeds, and space–time trajectories are essential in the field of transportation in general, and road safety in particular. However, automated data collection for these two road user types remains a challenge. Due to the constant change of orientation and appearance of pedestrians and cyclists, detecting and tracking them using video sensors is a difficult task. This is perhaps one of the main reasons why automated data collection methods are more advanced for motorized traffic. This paper presents a method based on Histogram of Oriented Gradients to extract features of an image box containing the tracked object and Support Vector Machine to classify moving objects in crowded traffic scenes. Moving objects are classified into three categories: pedestrians, cyclists, and motor vehicles. The proposed methodology is composed of three steps: (i) detecting and tracking each moving object in video data, (ii) classifying each object according to its appearance in each frame, and (iii) computing the probability of belonging to each class based on both object appearance and speed. For the last step, Bayes' rule is used to fuse appearance and speed in order to predict the object class. Using video datasets collected in different intersections, the methodology was built and tested. The developed methodology achieved an overall classification accuracy of greater than 88%. However, the classification accuracy varies across modes and is highest for vehicles and lower for pedestrians and cyclists. The applicability of the proposed methodology is illustrated using a simple case study to analyze cyclist–vehicle conflicts at intersections with and without bicycle facilities.

* Corresponding author. Tel.: +1 (514) 398 6589; fax: +1 (514) 398 7361.
E-mail addresses: sohail.zangenehpour@mail.mcgill.ca (S. Zangenehpour), luis.miranda-moreno@mcgill.ca (L.F. Miranda-Moreno), nicolas.saunier@polymtl.ca (N. Saunier).
1 Tel.: +1 (514) 340 4711x4962.

http://dx.doi.org/10.1016/j.trc.2015.04.003
0968-090X/

1. Introduction

With the increase in computing power and capacity of sensors coupled with their decreasing economical cost, the field of Intelligent Transportation Systems (ITS) has seen considerable improvements in automated traffic monitoring systems. The aim is not only to collect traffic data, e.g. flow, density and average speed at specific locations in the road network, but also detailed microscopic information about each road user (position and speed) continuously and over large areas of the network. A great amount of the workload of traffic monitoring will thus shift from human operators to emerging automated systems with improved performance and the possibility to perform new tasks such as road safety monitoring (Anon, 2001), leading to the development of more accurate and practical methods for data collection, safety diagnostics and evaluation of traffic engineering countermeasures at critical road facilities such as intersections.

Intersections are critical elements of the road network for safety, given that a high concentration of conflicts, crashes and injuries occurs at these locations. In cities like Montreal, 60% of pedestrian and cyclist injuries occur at intersections (Strauss et al., 2013). Given the importance of this topic, in research and practice, several recent studies have looked at different safety issues at intersections using traditional approaches based on historical crash data (Miranda-Moreno et al., 2011) and surrogate approaches such as conflict analysis (Ismail et al., 2009b). Independent of the method for road safety diagnosis, obtaining macroscopic and microscopic traffic data is fundamental. In the traditional safety approach, exposure measures are often developed based on traffic counts of each user type (e.g., vehicle, pedestrian and bicycle volumes over a given time period). In the surrogate approach, road user trajectories are necessary to compute measures such as Time To Collision (TTC), Post Encroachment Time (PET), and gap time (Saunier et al., 2010).

Road users can be detected and classified using a variety of sensors like inductive-loops, magnetic sensors, microwave and laser radars, infrared and ultrasonic sensors (Klein et al., 2006). However, it seems that the most convenient way to obtain spatial data such as road user trajectories over a certain area, if not the only, is through video sensors. These sensors have several advantages, in particular the ability to capture naturalistic movements of road users with a small risk of catching their attention, the relative ease of installation, the richness of extracted data and the relatively low cost (Saunier et al., 2011).

The main challenge with video sensors is developing an automated process to obtain trajectories by user type (e.g. for pedestrians, cyclists and vehicles) in order to avoid manual processing that is costly and time-consuming. Automated video processing is even more complex at urban signalized intersections which have a high mix of traffic conditions where all three main road user types (pedestrians, cyclists and motorized vehicles) are present. This issue has attracted some attention in research. The need for classification algorithms has been highlighted and addressed in Zaki and Sayed (2013) and Ismail et al. (2010). Tracking and collecting observational data for cyclists and pedestrians is more difficult than for vehicles because of their non-rigidity, more varied appearance, and less organized movements. In addition, they often move in groups close to each other which make them even harder to detect and track.

Accordingly, this research aims to develop an automated road user classification methodology combining a tracking algorithm and a HOG approach to obtain object appearance and speeds. Different classifiers are proposed to determine object class. Moving objects are classified into three categories: pedestrians, cyclists, and motor vehicles. The proposed methodology includes different tasks: detection and tracking of each moving object and classification according to object appearance in each frame and its speed throughout the video. Although this methodology may not be entirely novel in the field of computer science, this work is the first to combine and use this method in the field of transportation. An existing open-source tracking tool called *Traffic Intelligence* is used (Saunier, n.d.). As part of this research, the accuracy of different classifiers is evaluated to show the advantages of classifiers using two sources of information for classification: the appearance and speed of an object. Finally, the proposed method is demonstrated through an example which investigates the safety effectiveness of a specific treatment at an intersection, in this case a bicycle facility. This case study aims to illustrate one of the potential applications of the developed methodology. This methodology, however, is not limited to analysing the effectiveness of different safety treatments.

The following section provides a literature review on studies looking at object classification in traffic video, limitations and applications.

2. Background

The literature on automated video data analysis for traffic operations and safety is very extensive. In this paper, the literature review is concentrated on studies looking at object tracking and classification as well as their applications in road safety. Some of the gaps in this field are also discussed.

The first step in extracting microscopic data from a video is tracking objects (finding the position of each object in time). In Yilmaz et al. (2006), the authors provide a general survey on object tracking. As reported in Ismail et al. (2009a) and according to Forsyth et al. (2005), the different approaches for the detection and tracking of road users are categorized into tracking by detection, using flow, or with probability.

- *Tracking by detection:* object detection is done using background detection and subtraction with the current image (Antonini et al., 2006) or a model of image appearance using color distribution or edge characteristics (Gavrila and Munder, 2006). This approach works well in situations where the objects are well separated.