Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

Safety assessment of unsignalized pedestrian crossings by means of advanced movement tracking – The OBSERVE project

Christian Stefan[a]*, Claus Aichinger[a], Rainer Stütz[a] and Oliver Sidla[b]

[a] *AIT – Austrian Institute of Technology, Giefinggasse 2, 1210 Vienna, Austria*
[b]*SLR Engineering GmbH, Gartengasse 19, 8010 Graz, Austria*

Abstract

Pedestrians are among the most vulnerable victims of road traffic accidents. Establishing an unsignalized pedestrian crossing at intersections occasionally results in a high crash risk due to the fact that many vehicle drivers do not heed the legitimate right of way of pedestrians, either deliberately or because of some kind of distraction, speeding or deficiencies in the traffic environment. The primary objective of the OBSERVE project was to develop a novel approach for evaluating crosswalks based on data from observed pedestrian-vehicle driver interactions and local site conditions. Within the project, 85 unsignalized pedestrian crossings in the cities of Graz and Vienna were investigated by means of video observation. The trajectories of different road user categories were analysed to obtain information on driving and walking speeds, traffic behaviour, time gaps etc. That information was subsequently used to model driving behaviour. For the modelling process, data from 54 zebra crossings were used. A beta-regression model identified the parameters 'pedestrian crossing type' and 'pedestrian crossing width' having the highest influence on the stopping probability.

Keywords: Pedestrian crossing, safety, decision-support system, stopping probability

* Corresponding author. Tel.: +43 (0)50550 6329; fax: +43 (0)50550 6439.
E-mail address: christian.stefan@ait.ac.at

1. Introduction

1.1. Pedestrian crashes and Safe System Approach

Pedestrians are among the most vulnerable victims of road traffic accidents. Unprotected by vehicle body, safety belts or helmets, they are especially exposed to risk of serious injury and have a smaller chance of surviving a crash with a motorized vehicle. There continues to be a problem in Austria related to the safety of pedestrians who attempt to cross streets, particularly at unsignalized pedestrian crossings. According to Stefan et al. (2017), every 6th (16.8%) road injury crash involving at least one pedestrian occurs at an uncontrolled pedestrian crossing. Zeeger et al. (2017) investigated different types of pedestrian treatments at crosswalks and concludes that both the road environment (lane width, number of lanes, traffic volume etc.) and the installed safety equipment play a critical role in determining pedestrian safety levels.

The Safe System Approach has been adopted by road agencies around the world. Sweden's 'Vision Zero' and the Dutch 'Sustainable Safety' concepts are the underlying basis for the Safe System approach. 'Vision Zero' suggests that it is not acceptable for fatal or serious injuries to occur on the road system, and that account must be taken of human tolerances when designing road infrastructure. 'Sustainable Safety' is also underpinned by human-centered principles such as predictability of road course by a recognizable road design, homogeneity of mass and/or speed and forgivingness of the environment (Steinmetz et al., 2015). The safe system concept merges those principles and comprises of the following elements: a) safe roads and roadsides, b) safe speeds c) safe vehicles and d) safe road users.

National road administrations and road operators play a crucial role for the provision of road safety – it is both in their power and responsibility to affect the design and layout of the first element of a safe system approach, i.e. safe roads and roadsides. In order to do so, stakeholders need comprehensible and scientific-based information concerning the cause-effect relationship of crashes on dedicated pedestrian facilities.

1.2. Decision-support tool for safety assessment

The primary objective of the OBSERVE project was to develop a novel approach for evaluating unsignalized pedestrian crossings based on data from observed pedestrian-vehicle driver interactions and site conditions. According to Stefan et al. (2007), the likelihood of a crash occurring at an unsignalized pedestrian crossing highly negatively correlates with the willingness of vehicle drivers to stop in front of the crosswalk. Hence, (infrastructural) factors influencing the motivation of drivers to heed the right of way of pedestrians at such facilities play a crucial role for improving safety levels for vulnerable road users. Within the OBSERVE project, a decision-support tool was developed to evaluate the effectiveness of different infrastructural mitigating measures such as speed bumps or road islands and also to give an indication on the importance of well-established traffic parameters such as vehicle or pedestrian flow.

2. Video observation and trajectory extraction

2.1. Hardware

Detection of the pedestrian-vehicle driver interactions was accomplished by means of a high-resolution camera (see Fig. 1) which was reassembled specifically for this project in order to continuously monitor pedestrian and vehicular movement over a time period of several days. The weatherproof camera body was mounted at lamp posts or similar objects (traffic signs, trees) near the crosswalk (15-20m) at a height of 4m in order to provide an unobstructed view on the crossing and the most relevant road pathway.

The project-specific requirements concerning the camera equipment were as follows:

- Autonomous operation for more than 72 hours
- Quick and easy-to-accomplish installation in order not to disturb traffic flow
- Non-destructive installation of the camera unit on different objects in order to guarantee comparable monitoring conditions

Furthermore, a mobile measuring vehicle (BMW X3) with stereoscopic cameras and GPS-tracking equipment (Applanix) installed on the roof top was used to collect data concerning the road infrastructure (road width, number of lanes, traffic signs, road markings etc.) of each pedestrian crossing under investigation.

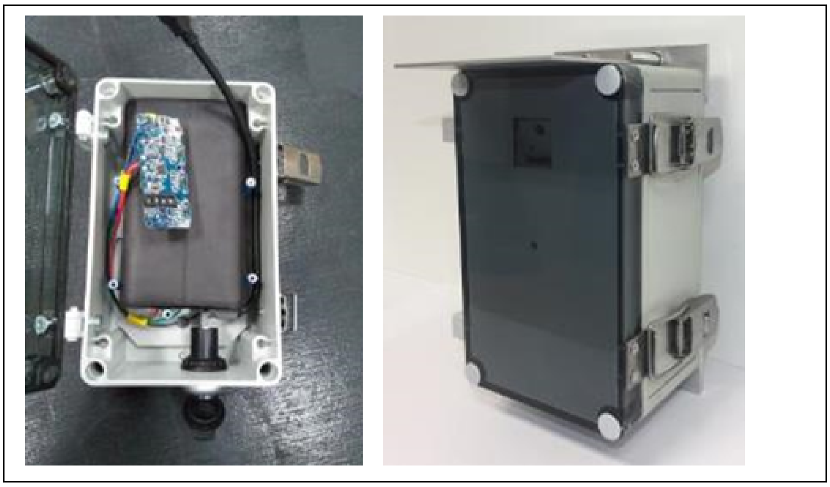

Fig. 1 Mobility Observation Box for detection of pedestrian-vehicle driver interactions

2.2. Computer vision

One of the main challenges within this project originated from the need to evaluate hundreds of hours of outdoor videos, which were used as data source for the automated traffic analysis. To this end, a visual pedestrian detector and tracker was developed, which should be reasonably i) efficient, so that processing time remains manageable, and ii) accurate enough to allow for solid tracking data in a real-world outdoor scenario.

According to both Dalal and Triggs (2005) and Benenson et al. (2014), cascaded HOG (Histogram of Oriented Gradients) detectors and the recently developed Deep Convolutional Networks (DCNN) were suitable approaches for this task (Zhang et al., 2016). For the work presented in this paper, the speed of HOG with its fairly good detection accuracy and a specifically trained DNN were combined into a novel five stage cascaded detector with overall high performance, high detection rate, and low false positive rate (see Fig. 2). The resulting detection cascade was fast and robust enough to provide the automatic evaluation of a large amount of traffic video data.

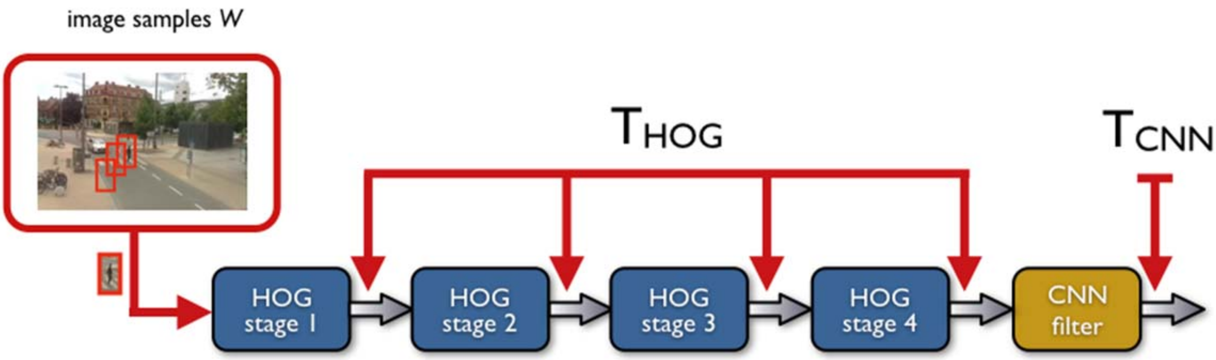

Fig. 2 Structure of the proposed HOG/CNN detector cascade

The first step within the developed framework was to localize object candidates by means of the implemented HOG cascade and then verify each of the candidates with a DCNN. The threshold for the HOG detector was set at a low value so that basically all objects of interest within an image were detected. The parameters of each HOG cascade stage (blocks and cells positioning) were optimized using a genetic optimization algorithm and the selected features were then trained and classified by means of a linear Support Vector Machine (SVM)

(Cristianini and Shawe-Taylor, 2000). The resulting cascaded detector was particularly suited for localizing objects within an image – even the highly overlapped full scanning of an entire HD image (sometimes up to 200,000 of HOG detections and SVM classifications) could be done closely to real-time on a standard CPU (without the use of a GPU).

Even though a fine tuned HOG detector can provide good detection capabilities, it certainly has its limits. Preliminary results showed that the useful number of training samples for HOG was about 50,000 samples. On the contrary, deep DCNNs allow unlimited training potential with the sample numbers currently in use. In almost each modern detection benchmark, deep network based algorithms resulted in superior detection results. Hence, the task of the last detector stage, the DCNN, was to weed out the remaining false positives. In this way, only a very small number of detection candidates needed to be processed, thus spending little time on the demanding DCNN inference process.

To render the detection algorithm practically useful, scale invariance was achieved by testing sample windows of different sizes at every image location. To support the detection process and to be able to create useful measurements from the detection results, a geometric calibration of the camera view was undertaken. By referring to an expected size range for each object class (e.g. the known size range of pedestrians) in an image, we were able to limit the detector size window to reasonable values and thus save computing time and enhance the detection accuracy.

3. Trajectory analysis

3.1. Trajectory (pre)processing and filtering

Trajectories were (pre)processed by projecting the video coordinates into a plane world coordinate system and applying smoothing filters to reduce irregular movements induced by (small) errors in the tracking algorithm. By providing additional information about the pedestrian crossing, trajectories were filtered and analysed/interpreted with respect to local conditions. A wide range of parameters describing the traffic situation were derived as well and used in the modelling process described in Chapter 4.

3.1.1. Projection

Camera calibration was needed for the analytics after detection, e.g. the automatic measurement of distances and velocities. A defined camera calibration allowed for projecting every image point to its corresponding world point on the ground plane and vice versa.

The calibration process was based on Tsai's method of lens camera calibration (Tsai, 1986), which uses a set of corresponding points from the camera to ground plane and basic information about the sensor geometry in order to establish the following camera parameters: the focal length f, a radial lens distortion coefficient k, the translation Tx, Ty, Tz and rotation Rx, Ry, Rz of the camera in world coordinates. Due to the monocular setup, a calibration could only be established for a preset elevation above the ground plane, which is typically the street level. Depending on the camera resolution and the distance of the object to the camera, measurement accuracies of a few centimeters were possible using the proposed calibration method.

3.1.2. Smoothing

Varying object-detection-accuracy may lead to discontinuities in the trajectories. To smooth out such irregularities, a centered moving average over 10 consecutive frames was applied to the trajectory data. At a sampling rate of 25Hz, this amounts to a window duration of 0.4s. Given the time scale of an average pedestrian-vehicle interaction and the procedure developed below, no loss of potential information due to smoothing was to be expected. The smoothed trajectories (see Fig. 3) provided more reliable information on pedestrian and vehicle driving dynamics and were subsequently used for the calculation of the parameters for the pedestrian-driver interactions.

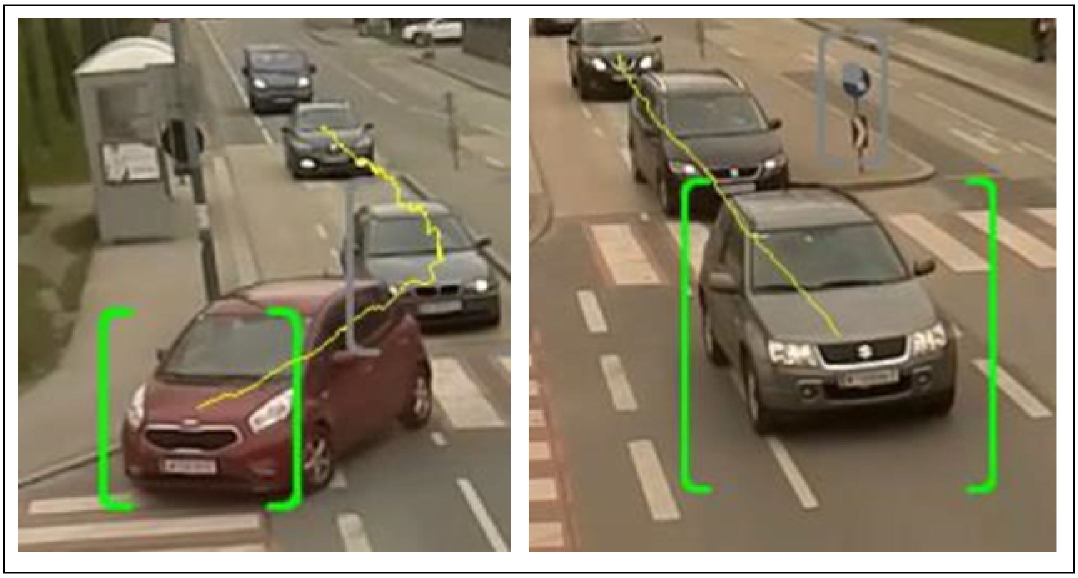

Fig. 3 (a) Raw vehicle trajectory (b) smoothed trajectory

3.1.3. Context and Filtering

In order to interpret the trajectories in the given local environmental context of the pedestrian crossing, specific markers as well as regions encapsulating this information were defined. The aerial shot in Fig. 4 gives an example of the markers and regions outlined below.

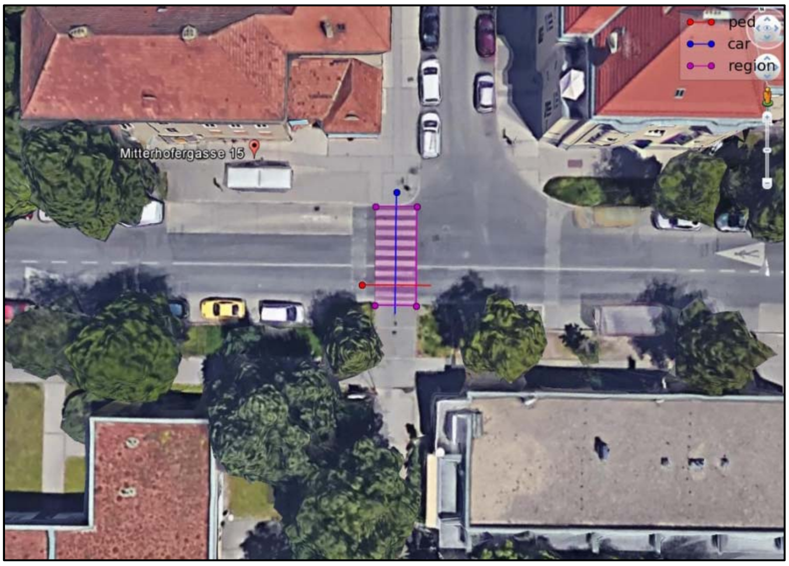

Fig. 4 Pedestrian crossing markers for trajectory interpretation

Crosswalk-Region: a (purple) polygonal region covering the crosswalk covers the crossing itself. An additional buffer (+10% offset) defines a proximity area. A pedestrian inside this region is interpreted as an individual who wants to cross the road.

Pedestrian-Marker: a (red) line perpendicular to the crosswalk differentiates between the before- and after-crossing situations. A pedestrian needs to cross this line to be considered relevant for the analysis. Furthermore, the direction of movement of the pedestrian can be inferred.

Vehicle-Marker: a (blue) line perpendicular to the vehicle lane relevant for the pedestrian crossing under consideration allows to differentiate different vehicle trajectories in a similar fashion as for pedestrians. A vehicle needs to cross this line to be considered relevant.

Trajectories with a length below 3m were excluded from the analysis. A vehicle was considered to be in proximity of the pedestrian crossing if the distance between the vehicle and the vehicle-marker was below 15m.

By computing the time at which a given trajectory crosses the corresponding line, the course of events, i.e. a sequence stating who crosses when, could be derived automatically.

3.1.4. Parameters

From the smoothed trajectories in the world coordinate system, an object's speed could be estimated by computing the distance covered between consecutive frames (the distance between 10 frames turned out to provide a satisfying approximation of the actual speed). Based on above markers and regions, additional information answering below questions were attached to each trajectory:

- Does the object cross the corresponding marker and if so when?
- From which direction does the object approach the crosswalk?
- For each point in (frame), is the object in proximity of the crosswalk?
- Vehicle: for each point in time (frame), is the vehicle able to stop before passing the crosswalk? Here constant deceleration with a reaction time of 1s and a deceleration rate of $8m/s^2$ is assumed.

Besides per-object information, summary statistics describing the traffic situation (number of vehicles/pedestrians, average (vehicle) velocities) were also calculated.

As outlined in the introduction, the willingness to stop was used as a proxy to describe the traffic safety level at a pedestrian crossing. In the context of the present work, the willingness to stop was defined as follows:

Willingness to stop = Amount of vehicles that behave correctly / Amount of all relevant vehicles

<u>Here</u>:
"all vehicles that behave correctly" was a subset of "all relevant vehicles". To apply this reasoning, a precise definition of "correct/incorrect" and "relevance" was needed.

3.1.5. Trajectory/Object Selection and Iteration

Out of all observed trajectories, those objects and time spans are selected that satisfy the following criteria:

<u>Vehicle:</u>
- approaches crosswalk from relevant direction
- is close to the crosswalk, i.e. distance is below 15m
- is next to cross the crosswalk (vehicle-marker)

<u>Pedestrian:</u>
- is at least once inside the buffer region
- is close to the crosswalk, i.e. inside the buffer region
- is next to cross the crosswalk (pedestrian-marker)

As already pointed out, both the vehicle and the pedestrian trajectories have to cross the crosswalk and be sufficiently long, i.e. more than 3m. The importance of the last bullet point needs to be highlighted, since it allows to assign to each vehicle a unique sequence of pedestrians thereby defining the relevant actors at each moment (frame). In an iterative manner, for each vehicle all pedestrians which are observed within the same time period are identified. For each of those vehicle-pedestrian pairs, the events in the common time interval in which both are going to be the next objects to cross the crosswalk are further analysed. During this overlapping time interval the so called pedestrian-vehicle-interaction takes place.

This approach implies (and is therefore based on this assumption) that at each point in time (frame) the situation at a crosswalk can be sufficiently described by a single vehicle and a single pedestrian. Since only single lane roads are considered and an assessment of the average vehicle-pedestrian interaction is sought, these assumptions appear justified.

Extensive manual review of annotated video material indicates that the methodology outlined in this paper provides a stable and reasonable assessment of pedestrian- vehicle interactions.

The flowchart in Fig. 5 describes the procedure to assess a vehicle-pedestrian-interaction. A more elaborate description of the decision steps is summarized in Table 1.

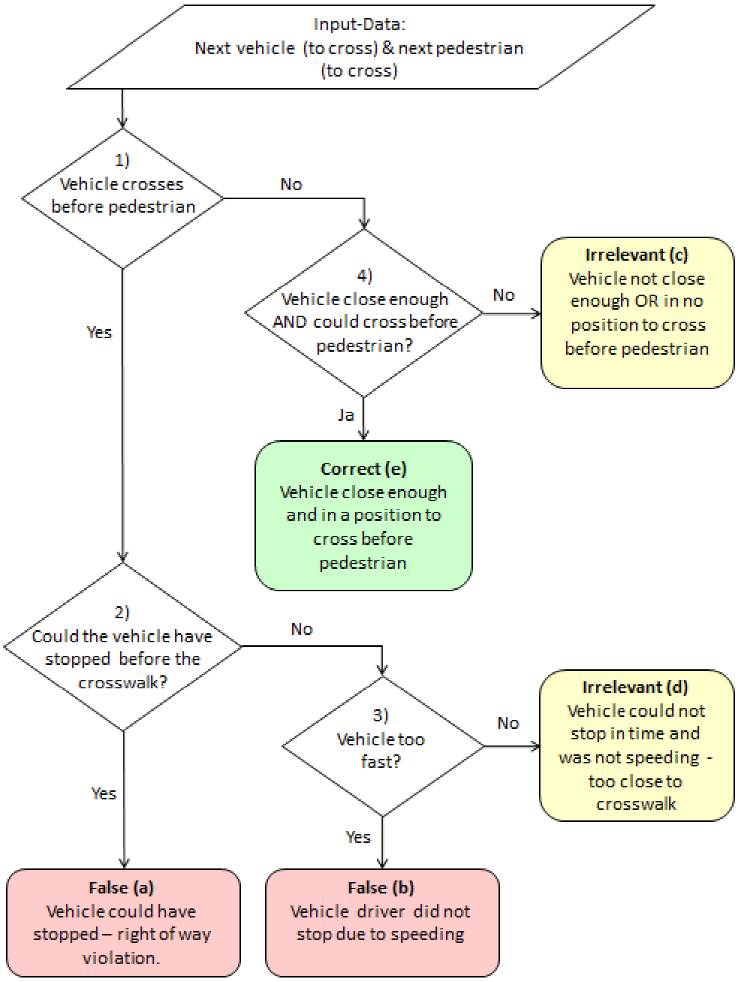

Fig. 5 Interaction analysis for pedestrian crossing

Table 1 Decision criteria for evaluation of pedestrian-vehicle interactions

Code	Decision criteria	Definition/Translation
(1)	Does pedestrian cross before vehicle?	The vehicle is observed to cross before the pedestrian does.
(2)	Is the vehicle, at any point in time, able to stop before the pedestrian passes?	Is at some point the stopping distance smaller than the actual distance to the crosswalk? (Then the vehicle would be able to stop safely)
(3)	Is the vehicle driving too fast?	At some point in time, the speed of the vehicle is greater than the allowed speed limit + 5km/h.
(4)	Is the vehicle in close proximity AND (logical or) could it pass the intersection before the pedestrian crosses?	The vehicle is closer than 15m OR, assuming constant velocity, the vehicle would pass before the pedestrian.

Having estimated the willingness to stop as well as additional parameters describing the traffic situation for all pedestrian crossings under investigation, a safety assessment as described in the following chapter was carried out.

4. Safety assessment

To assess the safety of unsignalized crosswalks, a model describing the relationship between the willingness to stop and the adjacent road environmental was developed. Video observations were then conducted in the two largest cities of Austria, Vienna and Graz. To find appropriate locations for the surveillance, several open spatial data sets were used to obtain an overview of the whole population – open government data from the municipality of Vienna and OpenStreetMap data in Graz. Vienna provided a detailed digital map with a relative accuracy of a few centimeters, showing the delimitations of public street areas and a polygonal vector layer which contained detailed information on buildings, roads, buildings, sidewalks, parking lots, traffic islands, speed humps etc.

Video observations and trajectory computations were conducted for a stratified sample of 85 different locations. All other relevant infrastructure parameters (e.g. traffic signs, road width) were collected on-site. After a manual quality check n=54locations were considered for statistical modelling.

Since the willingness to stop at unsignalized crosswalks as defined in Section 3.2 is given by proportion data, the response/dependent variable was defined from the closed unit interval [0,1]. Various methods are proposed in the literature to model proportion data. A common strategy for most of them is to transform the response, e.g. using the logit transformation, and apply ordinary least squares (OLS) regression. This has some major drawbacks, as the assumptions of OLS regression are often not met (e.g. homoscedasticity). An alternative approach is given by regression models which are based on the binomial distribution, i.e. logistic regression. However, in practice proportion data are often non-binomial, e.g. in case of overdispersion, where the observed variance is greater than the expected variance given by the underlying distribution. To overcome the aforementioned problems and limitations, a so-called beta regression model (Ferrari and Cribari-Neto (2004), Cribari-Neto and Zeileis (2010)) was used in this study. The model is based on the assumption that the response is beta-distributed. Beta regression is similar to a binomial generalized linear model (GLM) but provides much more flexibility due to the beta law.

The class of beta regression models assumes the continuous variables y to be restricted on the open unit interval (0,1), therefore the transformation $y \rightarrow \frac{(y(n-1)+0.5)}{n}$ (Smithson and Verkuilen, 2006) was applied to the response in a first step. The following set of explanatory variables was considered in the beta regression model:

Categorical predictors:
- posted speed limit (<50 km/h (24) and 50 km/h (30))
- crossing type (zebra crossing (17), zebra crossing with traffic island (22) and zebra crossing with speed humps (15))

Continuous covariables:
- pedestrian crossing width
- number of vehicles
- number of pedestrians

Note that a beta regression model is linear in the coefficients. To find an appropriate functional form of the predictors, generalized additive models were employed. Variable selection was performed using boosting methods (Mayr et al., 2012). The regression parameters of the beta regression model are interpretable in terms of the mean of the response and are shown for a final model in Table 2. Bootstrapping methods and boosting also showed that the crossing type and pedestrian crossing width have the highest influence on the response variable (i.e. the readiness to stop).

Table 2 Model coefficients and 95% confidence interval of the Beta regression model

Beta regression model	
centre island (c_isl)	0.313
	(-0.124, 0.750)
speed bump (s_bump)	0.715^{**}
	(0.104, 1.325)
traffic volume (n_veh)	-0.001^{*}
	(-0.003, 0.0002)
Pedestrian volume (n_ped)	-0.020
	(-0.045, 0.004)
I(n_veh^2)	0.0003^{**}
	(0.00002, 0.0005)
road width (r_width)	-0.118^{***}
	(-0.195, -0.040)
speed limit <50 km/h (s_limit)	-0.504^{**}
	(-0.918, -0.089)
constant	3.363^{***}
	(2.439, 4.287)
observations	54
pseudo-R^2	0.379
log Likelihood	68.900
note	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

5. Conclusions

A literature review conducted at the beginning of the project revealed that previous model approaches mainly focused on the influence of vehicle speed on willingness to stop but did not include several influencing factors of the road environment. Within this project, the methodological approach was widened in order to include important risk factors such as pedestrian crossing type, traffic and pedestrian volume or road width. The established beta regression model gives planers and decision makers a valuable tool to evaluate the safety related consequences of different mitigating measures up front. The modelling results prove that speed bumps and the available road width have the largest influence on vehicle driver’s willingness to stop. Installing a speed bump at unsignalized pedestrian crossings increases the stopping probability by more than 8 percent. Road islands in conjunction with pedestrian bulbouts nearly have the same effect.

After conducting several test runs at unsignalized pedestrian crossings, the correct setup for the observational study also became evident. Sites need to be selected with great care in order to prevent selection bias blemishing the results of the modelling process. A reasonable straight road segment leading to the crosswalk under investigation was found to be paramount for reasonable trajectory detection. Furthermore, the distance of the mounted camera unit to the pedestrian crossing should not exceed 20m. Otherwise, movement tracking becomes increasingly instable and the number of analyzable trajectories decreases significantly.

Acknowledgements

This paper is based on work carried out as part of the OBSERVE research project which latest from 2015-2017. The authors are grateful for the funding received from the Austrian Traffic Safety Funds (VSF).

6. References

Benenson, R., Omran, M., Hosang, J. and Schiele, B., 2014. Ten years of pedestrian detection, what have we learned? European Conference on Computer Vision, CVRSUAD workshop.

Cristianini, N. and Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines and other kernel-based Learning methods", Cambridge University Press.

Cribari-Neto, F., Zeileis, A., 2010. Beta Regression in R. *Journal of Statistical Software*, vol. 34, no. 2, pp. 1–24.

Ferrari, S. and Cribari-Neto, F., 2004. Beta Regression for Modelling Rates and Proportions. Journal of Applied Statistics, vol. 31, no. 7, pp. 799–815.

Dalal, N. and Triggs, B., 2005. Histograms of Oriented Gradients for Human Detection, Conference on Computer Vision and Pattern Recognition, pp. 886-893, San Diego.

Mayr, A., Fenske, N., Hofner, B., Kneib, T. and Schmid, M., (2012). Generalized additive models for location, scale and shape for high-dimensional data - a flexible approach based on boosting. *Journal of the Royal Statistical Society, Series C - Applied Statistics*, vol. 61, no. 3, pp. 403–427.

Smithson, M. and Verkuilen, J., 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, vol. 11, no. 1, pp. 54–71.

Stefan, C., Aichinger, C., Aleksa, M., Stütz, R., Schwieger, K., Harnisch, M., Sidla, O. and Lipetski, Y., 2017. Evaluierung von Querungsstellen mittels risikobasierter Bewertungsverfahren. Österreichischer Verkehrssicherheitsfonds (VSF), Nr. 61, Vienna.

Stefan, C., Smuc, M., Schreder, G., Gatscha, M., Reichenauer, S. and Riccabona-Zecha, C., 2007. Verkehrssicherheitsvergleich unterschiedlicher Varianten der Schutzwegausstattung, Österreichischer Verkehrssicherheitsfonds (VSF), Forschungsarbeiten aus dem Verkehrswesen, Band 161, Vienna.

Steinmetz, L., Jurewicz, C. and Taylor, S., 2015. Safe system in the planning process. Austroads Research Project AP-R488-15, Sydney.

Tsai, R.Y., 1986. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach.

Zeeger, C., Srinivasan, R., Lan, B., Carter, D. and Smith, S., 2017. Development of Crash Modification Factors for Uncontrolled Pedestrian Crossing Treatments. NCHRP Research Report 841, Transportation Research Board, Washington.

Zhang, S., Benenson, R., Omran, M., Hosang, J. and Schiele, B., 2016. How far are we from solving pedestrian detection? Conference on Computer Vision and Pattern Recognition.