2018 21st International Conference on Intelligent Transportation Systems (ITSC)
Maui, Hawaii, USA, November 4-7, 2018

A Literature Review on the Prediction of Pedestrian Behavior in Urban Scenarios

Daniela Ridel
ICMC/USP
Sao Carlos, Brazil
danielaridel@usp.br

Eike Rehder
Daimler AG
Sindelfingen, Germany
eike.rehder@daimler.com

Martin Lauer
MRT/KIT
Karlsruhe, Germany
martin.lauer@kit.edu

Christoph Stiller
MRT/KIT
Karlsruhe, Germany
stiller@kit.edu

Denis Wolf
ICMC/USP
Sao Carlos, Brazil
denis@icmc.usp.br

***Abstract*—The ability to anticipate pedestrian actions on streets is a safety issue for intelligent cars and has increasingly drawn the attention of the automotive industry. Estimating when pedestrians will cross streets has proved a challenging task, since they can move in many different directions, suddenly change motion, be occluded by a variety of obstacles and distracted while talking to other pedestrians or typing on a mobile phone. Moreover, their decisions can also be affected by several factors. This paper explores the ways pedestrians' intention estimation has been studied, evaluated, and evolved. It provides a literature review on pedestrian behavior prediction, addresses available solutions, state-of-the-art developments, and hurdles to be overcome towards reaching a solution that is closer to the human ability to predict and interpret such scenarios. Although many studies can precisely estimate pedestrians' positioning one second before they cross a street, most of them cannot precisely predict when they will stop at a curb.**

***Index Terms*—pedestrian, road user, prediction, estimation, recognition, behavior, intention, trajectory, crossing**

I. Introduction

Pedestrians account for 22% of the worldwide 1.24 million deaths caused by traffic accidents every year [1]. Most of such deaths occur when pedestrians are crossing a street [2] at sunset [1] and may be caused by poor visibility and drivers' fatigue.

Many researchers have focused on the development of algorithms that estimate pedestrians' intentions of crossing streets [3], [4], [5], [6], [7]. However, the problem is still challenging, since pedestrians can move in any direction and suddenly change motion [8], [9]. Such a prediction is a prerequisite for the safe operation of automated vehicles. Moreover, cars can travel at high speeds, which results in a short time-window for a prediction and a decision to be made.

The main contribution of this manuscript is a detailed analysis of the literature on prediction of pedestrians' behaviors, evolution of research over the years, state-of-the-art solutions, and challenges and gaps to be considered. It also addresses the advantages and drawbacks of the currently available datasets and potential future trends.

II. Pedestrian Behavior Estimation

Several studies have focused on the construction of systems that increase the safety of pedestrians on streets. Keller et al. [10] developed a system using stereo vision which enabled autonomous cars to perform evasive steering maneuvers. Roth et al. [11] evaluated pedestrians' and drivers' awareness for estimating a potential collision risk. Li et al. [7] designed a system for concurrently detecting pedestrians and cyclists. Applications regarding pedestrian detection have already been employed in the industry, e. g. Lexus RX 2017 whose system warns the driver on the detection of a pedestrian. In other cases, the system is granted the control of the car if no action is performed by the human driver, e. g. Ford Fusion 2017 and BMW 3 Series Sedan. The detection of pedestrians has been investigated in several studies [12], [13], [14], [15]. Most of them use images [16], [17], 3D point clouds [18], [19], or even the fusion of both sets of information [20], [21].

The *estimation* of pedestrians' intention is even more challenging due to uncertainties regarding their impending motion [8]. In a fraction of a second, pedestrians can decide to move in one of many different possible directions, stop walking abruptly [9], [8], [22], have their image/point cloud occluded by a variety of obstacles, and be distracted talking to other pedestrians or even on a mobile phone. According to [23], the analysis of non-critical situations has not received considerable attention and Quintero et al., [24], observed the difference between an effective and a non-effective intervention can depend merely on a few centimeters or a fraction of a second.

Pedestrian intention can be estimated jointly with path prediction, as proposed in [3] and [24]. A comparative study of pedestrian path prediction was conducted by [9], whereas [25] provides a survey on types of interactions between autonomous vehicles and humans. Shirazi and Morris [26] reviewed pedestrian, driver, and vehicle behaviors at intersections and analyzed features that distinguish different pedestrian motion patterns. A related approach can be found in [4].

According to the literature, research initiatives are concerned with both short-term and long-term predictions. Studies that comprise long-term predictions usually draw information from static cameras and aim at predicting either the pedestrians' final destination, or the path followed [27], [28], [29]. Karasev et al., [27], modeled pedestrians' intention in a Markov decision-process framework and inferred their state using a Rao-Backwellized filter. They focused on each pedestrian individually and neglected their interactions with other traffic participants. Kitani et al. [28] predicted future actions of pedestrians using noisy visual data and the effects of the physical environment on pedestrians' behavioral choices combining ideas from Control Theory and static semantic scene understanding. Pedestrians can decide to quickly change direction, which makes their long-term predictions a challenging task [8], [22].

Short-term approaches [6], [30], [31], [5] predict pedestrians' position up to the next 2.5 seconds. According to Rehder and Kloeden [32], information, such as head orientation and body movement is relevant in short-time predictions, while long-term predictions are more goal-oriented.

The subsequent sections address our expansion of the taxonomy proposed by Brouwer et al. [33], which categorized related studies according to input data.

A. Body Pose

The human movement analysis based on vision has been a topic of research in many fields and applications, as games, character animation, surveillance systems, traffic analysis, social interfaces, sign-language translation, and dance choreography [34].

Some studies explore the use of pedestrians' contour [35], [4], posture [36], [37] and body language [38], [24] to predict their intentions. Hariyono and Jo [39] used pose recognition, lateral speed, orientation, and scene comprehension as input to a neural network to predict actions, as walking, starting off, bending in and stopping.

The approach of Quintero et. al. [38] used a Gaussian Process (GP) model to reduce the dimensionality of the 3D coordinates of pedestrian body pose and learned to predict pedestrian's dynamics related to walking, starting off, stopping and standing. Quintero et. al. [24] also used Gaussian Process Dynamical Models (GPDM) for estimating pedestrian intention regarding a pose prediction of a one-second horizon. However, more than four seconds were necessary for the model to perform the prediction. Furuhashi and Yamada [36] predicted if a pedestrian standing on the crosswalk would cross the street based on the analysis of postural change throughout consecutive frames captured by a static camera. The model proposed by Köhler et. al. ([35] and [4]) used a HOG-like descriptor for motion contour pedestrian detection along with a Support Vector Machine (SVM) to estimate pedestrians' intentions.

B. Inter-pedestrian/Social-related Behavior

Some researchers considered decisions made by pedestrians based on social norms commonly followed within a shared common space [40], [41], [42], [43]. They observed the patterns used by pedestrians in such interactions and identified several norms, e.g. pedestrians maintain some distance from each other, pedestrians avoid others coming towards them, pedestrians can follow the flow of other pedestrians on the scene, etc. Alahi et. al [44], proposed a method that applied a social layer over the Long Short-Term Memory (LSTM) network for each pedestrian and implicitly learned such interactions.

C. Dynamics-Based Prediction

Attempts towards predictions of pedestrians' positions originated from tracking, which is naturally the second step after the detection of a pedestrian. Many studies predicted pedestrians' positions through Kalman Filter (KF) and Particle Filter (PF) [45], [46]. In [47], as in similar research initiatives, the direction of a pedestrian walking is estimated according to his/her position within multiple consecutive image frames regarding the distance from the vehicle. Kataoka et. al. [48] inferred pedestrian intention by recognizing a walking activity through pedestrian localization and activity analysis.

Schneider and Gavrila [9] compared Extended Kalman Filter (EKF) based on single dynamic models and Interacting Multiple Model (IMM) algorithm, combining Constant Velocity (CV), Constant Acceleration (CA), and Constant Turn (CT). They also proposed a dataset composed of 4 pedestrian motion categories, namely crossing, stopping, bending in and starting off, which was significantly used in subsequent work.

Keller and Gavrila [49] designed a method and compared it with different approaches related to pedestrian path prediction, namely GP, Probabilistic Hierarchical Trajectory Matching (PHTM), KF, and IMM at various time horizons. The results showed a similar prediction performance when pedestrians' locations are precise, however, the approaches that used GPDM and PHTM performed better in stopping scenarios. For comparative reasons, they also employed human subjects tasked with predictions of pedestrians possibly stopping or crossing a street.

Switching dynamics (Linear Dynamical System (LDS)) was used by Kooij at. al. [50] towards more accurate path predictions. They established certain actions more likely to occur in the future depending on previous movements and current locations. Best and Fitch [51] estimated the goal destination and trajectory to be achieved. An environmental map was incorporated into the model and the probability distribution that represents the estimated succeeding position was computed by a Bayesian framework.

Nevertheless, Schmidt and Färber [30] observed the use of only dynamics would not be sufficient, e.g. a KF tracking a pedestrian walking parallel to the ego-vehicle would always predict pedestrian's future positions to be set further. However, pedestrian constantly turning his/her head towards the autonomous vehicle and the road is an indication of where he/she intends to go (e.g. the other side of the street). Therefore, an approach that solely relies on pedestrians' dynamics will never predict their intention of crossing a street.

D. Dynamics and Awareness

Information on head orientation has been incorporated in estimation methods towards improving the estimation of pedestrians' intentions [52], [3], which has given rise to research on perfecting classification of head orientation [32].

Several studies [3], [53], [54], [55], [37] use information on pedestrian dynamics coupled with the awareness of the situation, i.e., the possible pedestrian's visualization of a vehicle and a critical situation.

Goldhammer et al. [52] focused on trajectory prediction of pedestrians on crosswalks and estimated gait initiation through a piecewise linear model and a sigmoid model for calculating velocity and inferring a trajectory. They designed an approach [56] that uses a Multilayer Perceptron (MLP) network based on head orientation information to predict a continuous trajectory for a 2.5 second future time horizon and motion types (starting and stopping). Kwak et al. [57] also aimed at estimating the intention of a pedestrian crossing a street or stopping, however, they used dynamic fuzzy automata in low-light images from a thermal camera.

Schulz and Stiefelhagen [3] developed a Latent Dynamic Conditional Random Fields (LDCRF) system that controls an IMM using pedestrian dynamics and head orientation as input. The lateral position error, with 1-second prediction horizon, is used for evaluation purposes. IMM provides pedestrian tracking and path prediction and LDCRF provides the intention estimation. Through a similar approach [53], the authors used LDCRF to recognize pedestrian's intention. The model automatically learned correlated intrinsic structure and dynamics among different actions.

However, relying solely on head orientation may not be the best alternative, since pedestrians may be looking at an advertisement or searching for someone; in such moments, their head might not indicate their current direction.

E. Dynamics, Awareness and Scene Understanding

Some studies [58], [6] have evaluated the influence of the environment on the behavior of pedestrians. Cloutier et al. [58] observed different crossing surface materials and one-way streets were significantly associated with fewer interactions with vehicles, whereas streets with parked vehicles and main streets were associated with more interactions.

The next generation of the state-of-the-art methods uses information from the environment, therefore, relations among environment, autonomous car, and pedestrians are structured [59], [39], [6], [60], [3], [54], [55], [5], [37], [61], [62].

Bock et al. [63] proposed a method based on Recurrent Neural Networks through which pedestrians' trajectories are learned from data collected by cameras on both vehicle and infrastructure. Other approaches focused on infrastructure cameras [60], [52], [35], [63].

Ferguson et al. [8] predicted pedestrian orientations and trajectories associated with each decision through GP towards incorporating uncertainty. Other research initiatives also used GP [8], [38] and [24].

Völz et al. [64] inferred the pedestrians' decision to cross a street using an SVM, however, they did not predict the exact trajectory. They aimed to determine the necessary information for detecting pedestrians' intention and its possible changes caused by other road users. Völz et al. [62] used data on pedestrians, as velocity, distance traveled, distance from the curb, and distance from the crosswalk, jointly with data on car movement and position. However, the authors evaluated only cases in which the pedestrian would invariably cross a street. Köhler et al. [35] also employed an SVM, however, they used data extracted from gait analysis.

Bonnin et al. [5] predicted whether a pedestrian would cross a street creating relations among pedestrian, crosswalk, and ego vehicle, and combining two models, namely a standard inner-city model, which is always activated, and a model activated only in crosswalks.

Dynamic Bayesian Network (DBN) was used in [55], [6], [54]. The latter, [54], considered external surroundings context, pedestrian behavior, physical movement, and information on a pedestrian being in a group or alone [55]. Kooij et al. [6] proposed a DBN that captures some factors as latent states that affect a Switching Linear Dynamical System (SLDS). A pedestrian that always intends to cross a street is the subject of the test sequences. The authors used 3 types of information on top of SLDS, namely 1) minimum distance between pedestrian and ego-vehicle if both keep the same velocity (indicating criticality of situation), 2) pedestrian's head orientation (awareness), and 3) distance from the pedestrian to the curbside.

Bonnin et al. [5] and Kooij et al. [6] employed almost the same observable features, i.e. distance to curb, distance to ego-vehicle, and head orientation. They did not use information from other pedestrians and cars and focused on short-term predictions.

According to Hashimoto et al. [54], the detection of pedestrians' decision to stop is a harder task than their decision to cross a street. Although the approach in [6] achieved better results for scenarios where the pedestrian stopped, most of the studies [65], [49] have yielded worse results on predictions of a pedestrian deciding to stop at a curb. Table I shows some approaches on prediction of pedestrians' intentions.

III. Discussion

According to Shirazi and Morris [26], some external factors can influence the pedestrian's decision-making and motion patterns, such as street width, being alone or in a group, and presence of other cars. Some studies investigated the shift in the speed of the pedestrians' stride, the reason for abrupt stops [70], and the speed variation in gender and age groups.

Schmidt and Färber [30] performed experiments for discovering the parameters that most influence pedestrians' decision to cross a street and the factors that human drivers take into consideration when predicting if a pedestrian will cross a street or not. They observed tighter gaps were chosen on streets of a narrower width and both older people and young adults choose insufficiently large gaps.

Iacoboni et al. [71] analyzed the cerebral activity of people watching others performing some actions and observed some neural cells were activated as soon as the intention had been inferred (before an action was actually performed). In other words, humans observing other humans' actions can implicitly understand their intentions. According to Keller et al. [72], algorithms still do not predict pedestrian intentions as well as humans, and pedestrians' behaviors in urban scenarios are not random [73], which raises the question: What information is neglected by current algorithms?

Our focus is on pedestrians moving in inner-city scenarios. Most of the prior work considered pedestrians a generic entity. However, characteristics regarding different types of pedestrians can be useful information for predictions of their actions. For instance, drivers are usually more careful when they see a child on the curbside, since children may exhibit unpredictable behaviors and inadvertently run into the street. Likewise, when drivers spot a pedestrian with a cane, they assume he/she may be blind or have difficulties in moving. Elderly people, on the other hand, walk slower than an average adult and may face difficulties due to the deterioration of their perceptual and cognitive abilities.

Factors, as age, gender, number of people walking together (when alone, people usually walk faster), constraints posed by the environment (e. g. number of lanes and width of the street) significantly contribute to a pedestrians' speed [74]. Sobhani et al. [75] evaluated the difference in crossing patterns between ordinary pedestrians and pedestrians using a mobile phone. Additionally, a mobile phone in the hands of a pedestrian may provide information regarding his/her level of awareness of the situation.

We believe an individual analysis of pedestrians enables intelligent vehicles to make better predictions and adaptable decisions, as such information may enhance the understanding of pedestrians' behavior patterns and speed of motion in street scenarios.

The estimation of pedestrians' intentions still requires approaches that perceive real-world scenarios, since most of the studies perform predictions using datasets in which pedestrians are actors with predetermined actions. Therefore, they do not represent a real scenario where several factors can influence a decision. A pedestrian may prefer to wait because, for instance, the street is too wide, he/she stops to answer a text message, or is pushing a stroller. All such factors demand more time, therefore, the pedestrian decides to wait for the next gap between vehicles.

We believe a further step for improvements in the state-of-the-art of predictions of pedestrians' intentions will be the use of pedestrian-specific characteristics jointly with dynamics and contextual scene information. The use of contextual information was proposed in [59], [39], [6], [60], [3], [54], [55], [5], [37], [61], however, to the best of our knowledge, the inclusion of different types of pedestrians (i .e. specific pedestrian characteristics) concerning mainly mobility, other indicators of awareness (apart from head orientation information), and wariness has not been addressed. A dataset with information as gender, age and size of pedestrian groups was recently proposed by Rasouli et al. [76], however, the authors did not evaluate in depth the influence of those aspects on the estimation of pedestrians' intentions.

Pedestrian dynamics, i .e., position and velocity, can be accurately assessed through stereoscopic camera images along with detection and tracking algorithms. Pedestrian awareness can be estimated through the observation of head orientation (as proposed in [32]) and also detection of possible objects that may interfere with their attention (e.g mobile phone, cane, and stroller). Scene information can be assessed through the identification of presence of signaling devices (traffic lights and crosswalks), curb location and number of lanes. Pedestrians' characteristics can be indexed according to age groups (child, adult, and senior), detection of objects they may be packing (cane, mobile phone, and stroller) and number of people walking together (alone, pair or family). Such factors may interfere with the pedestrian's awareness, mobility, and wariness towards their surroundings.

The identification of pedestrians' characteristics would provide intelligent cars with information on the different pedestrians' types, enable cars to use adaptable systems that infer actions and understand

TABLE I: Studies of pedestrian intention estimation

Study/author	*Sensor/input*	*Method*	*Objective/output*	*Evaluation/dataset*
[9] (Schneider and Gavrila, 2013)	Stereo-cameras	Comparative study on recursive Bayesian filters IMM and EKF	[crossing, stopping, bending in, starting]	Accuracy of position estimation and path prediction
[66] (Köhler et al., 2013)	Lidar and two cameras	Interacting Multiple Model Extended Kalman Filter (IMM-EKF), Motion Contour image based Histogram Of Gradients (MCHOG) and SVM	[crossing, not crossing]	Receiver Operating Characteristic (ROC)
[5] (Bonnin et al., 2014)	Image, CAN	MLP	[crossing, not crossing]	False Positive/ True Positive Rate
[6] (Kooij et al., 2014)	Image	DBN	Trajectory prediction	Comparison with PHTM through Log Likelihood
[8] (Ferguson et al., 2015)	Lidar	GP	Trajectory prediction	Probability of correct motion pattern and RMS error
[64] (Völz et al., 2015)	Lidar	SVM	[crossing, not crossing]	False Positive/ True Positive Rate/Classification Accuracy
[56] (Goldhammer et al., 2015)	Camera	MLP	Trajectory prediction	Mean square deviation (RMSD2D) from the predicted position to Ground Truth (GT). Comparison with KF CV
[3] (Schulz and Stiefelhagen, 2015)	Stereo-cameras	IMM and LDCRF	Trajectory prediction	Lateral position error
[24] (Quintero et al., 2015)	3D pedestrian body Pose	Balanced Gaussian Process Dynamical Models (B-GPDM) and naive-Bayes	Trajectory prediction and classification of behavior [walking, starting, stopping and standing]	Confusion Matrix and mean error for position
[54] (Hashimoto et al., 2015)	Three monocular cameras	DBN and PF	[crossing, not crossing]	Mean of the estimated probability of the pedestrian crossing/waiting
[57] (Kwak et al., 2017)	Night images from thermal camera	Dynamic fuzzy	[Standing Sidewalk, Walking Sidewalk, Walking Crossing and Running-Crossing]	Precision and recall rate. Comparison with Markovian model-based method, DBN and Fuzzy Finite Automata (FFA)
[63] (Bock et al., 2017)	Pedestrian with sensors	LSTM	Trajectory prediction	Mean displacement error and final displacement error. Comparison with KF. Intersection dataset based on infrastructure sensors and information on pedestrian localization. Single pedestrian
[67] (Li et al., 2017)	Indoor positioning wifi data	LSTM	Trajectory prediction	Comparison with Gated Recurrent Unit (GRU) and vanilla Recurrent Neural Network (RNN)
[68] (Dominguez-Sanchez et al., 2017)	Stereo-camera	Convolutional Neural Network (CNN)	Pedestrian moving direction	Accuracy among different CNNs
[69] (Rehder et al., 2018)	Images	CNN, LSTM and path planning	Trajectory and goal prediction	Predicted probability distribution and comparison with a Constant Position (CP)-IMM and CV-IMM

scenarios that strongly resemble the way humans perform such tasks. Additionally, the use of the features previously discussed would improve both the development of traffic infrastructures based on reports on most common user types and construction of human/machine interfaces and advance the Assistive Intelligent Transportation System [77] field.

IV. DATASETS

Datasets created for the detection and tracking of pedestrians in images usually focus on surveillance and inner-city scenarios. EPFL Multi-camera pedestrian videos [78], [79] and PETS 2009 [80] used static cameras in indoor and outdoor scenarios, respectively. ETHZ Sequences: Inner City Sequences from Mobile Platforms [81] and TUD Stadtmitte [82] used moving cameras in inner-city scenarios. KITTI detection and tracking benchmark [83] is a well-established benchmark that significantly contributes to the autonomous vehicles field by enabling comparisons of results of different approaches under the same input. The KITTI pedestrian detection benchmark comprehends more than one hundred ranked methods and promotes a considerable evolution in the autonomous vehicle field by providing a rich infrastructure through which tests and comparisons of different approaches for solving e.g., pedestrian detection and tracking. However, no benchmark is available for comparisons of multiple methods regarding prediction of pedestrians' behaviors/intentions. In general, each study uses a different dataset and provides their own results and possible comparisons with one or two other methods.

Some of the currently available datasets for pedestrian path prediction are Daimler Pedestrian Path Prediction Benchmark Dataset ECCV14 [6] and Daimler Pedestrian Path Prediction Benchmark Dataset GCPR13 [9]. The problem with such datasets is the available data are related to only one pedestrian who takes predetermined actions. The former dataset does not provide the images captured. Another available dataset for crosswalk behavior classification [76] provides labels with information on weather conditions, presence of traffic elements (traffic light, crosswalk) and gender; however, it does not capture the coordinates of the pedestrians' positions, but only the action associated with each pedestrian. In general, the available datasets fall into two possible categories, i.e. scenarios where the pedestrian is instructed to perform some predetermined actions and more realistic scenarios with real pedestrians. In both cases, the number of data collected is usually very limited.

Two major problems have been identified in the available datasets. The pedestrian is instructed to perform some predetermined actions, i.e. subjects impersonate pedestrians, but carry out predetermined actions (crossing, stopping, utilizing the crosswalk, etc). Such a scenario does not comprise all information, which may encourage the pedestrian to perform an action, and the variability behind real pedestrians' decisions. In more realistic scenarios that use real inner-city data (i.e. where pedestrians are not actors performing some predetermined actions), the problem is the loop between pedestrian/ego-vehicle actions. According to a report available in [84], 80 percent of the pedestrians search for eye contact with the driver before crossing a street. Since the datasets were collected by sensors in a car driven by a human driver, we believe the behavior of pedestrians may have been affected by such an interaction even in realistic scenarios. In other words, the communication established between the driver and the pedestrian may interfere with the data collected, since such an occurrence is not in the stored data.

The evaluation procedures depend on the output. Keller et al. [49] conducted evaluations in different time horizons and compared the results with humans' answers (crossing or not crossing). The methods based on classification usually measure the True Positive Rate (TPR), False Negative Rate (FNR), and ROC curve. Studies that predict the future exact position of a pedestrian usually measure the lateral position error [9], [3] calculated by the mean squared error between the prediction and the GT, or the log likelihood.

V. Final Considerations

This paper has provided a literature review on prediction of pedestrians' intentions. Intelligent cars' applications require high accuracy and precision and low response time, in which a prediction must be made. Despite the efforts devoted towards developing reliable solutions, a lot more effort still needs to be put forth to create a system that can ensure pedestrian's safety on streets.

We believe the next step for the evolution of the state-of-the-art on the estimation of pedestrians' intentions will involve the development of robust methods that approximate to the human analytical capabilities of identifying an intention prior to an action. Information on each pedestrian's profile can also provide useful data (and we just need to look at them once). Finally, a comprehensive benchmark for the ranking of the currently available methods and those to be designed must be developed, so that proper data and tools can be used for comparisons of algorithms.

Acknowledgement

The authors acknowledge CAPES and Erasmus Mundus for the financial support provided to this research under PROEX-8999660/D and BM16DM1046 processes.

References

[1] W. H. O. WHO, "Global status report on road safety," 2015, p. 8.

[2] M. M. M., O. M., G. D., M. Marc, E. Morris, R. Töns, and M. Lettelier, "Strategies in terms of vulnerable road user protection," in *EU Project SAVE-U*, 2003.

[3] A. T. Schulz and R. Stiefelhagen, "A controlled interactive multiple model filter for combined pedestrian intention recognition and path prediction," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 173–178.

[4] S. Köhler, M. Goldhammer, K. Zindler, K. Doll, and K. Dietmeyer, "Stereo-vision-based pedestrian's intention detection in a moving vehicle," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 2317–2322.

[5] S. Bonnin, T. H. Weisswange, F. Kummert, and J. Schmuedderich, "Pedestrian crossing prediction using multiple context-based models," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2014, pp. 378–385.

[6] J. F. P. Kooij, N. Schneider, F. Flohr, and D. M. Gavrila, "Context-based pedestrian path prediction," in *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 618–633.

[7] X. Li, L. Li, F. Flohr, J. Wang, H. Xiong, M. Bernhard, S. Pan, D. M. Gavrila, and K. Li, "A unified framework for concurrent pedestrian and cyclist detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 2, pp. 269–281, 2017.

[8] S. Ferguson, B. Luders, R. C. Grande, and J. P. How, "Real-time predictive modeling and robust avoidance of pedestrians with uncertain, changing intentions," in *Algorithmic Foundations of Robotics XI: Selected Contributions of the Eleventh International Workshop on the Algorithmic Foundations of Robotics*, H. L. Akin, N. M. Amato, V. Isler, and A. F. van der Stappen, Eds. Cham: Springer International Publishing, 2015, pp. 161–177.

[9] N. Schneider and D. M. Gavrila, "Pedestrian path prediction with recursive bayesian filters: A comparative study," in *Pattern Recognition: 35th German Conference, GCPR 2013, Saarbrücken, Germany, September 3-6, 2013. Proceedings*, J. Weickert, M. Hein, and B. Schiele, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 174–183.

[10] C. G. Keller, T. Dang, H. Fritz, A. Joos, C. Rabe, and D. M. Gavrila, "Active pedestrian safety by automatic braking and evasive steering," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1292–1304, 2011.
[11] M. Roth, F. Flohr, and D. M. Gavrila, "Driver and pedestrian awareness-based collision risk analysis," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 454–459.
[12] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 12, pp. 2179–2195, 2009.
[13] H. J. Benenson R., Omran M. and B. Schiele, "Ten years of pedestrian detection, what have we learned?" in *Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part II*, B. M. M. Agapito L. and R. C., Eds. Cham: Springer International Publishing, 2015, pp. 613–627.
[14] M. O. J. H. S. Zhang, R. Benenson and B. Schiele, "How far are we from solving pedestrian detection?" in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 1259–1267.
[15] D. M. Gavrila and S. Munder, "Multi-cue pedestrian detection and tracking from a moving vehicle," *International Journal of Computer Vision*, vol. 73, no. 1, pp. 41–59, Jun 2007.
[16] M. Bertozzi, L. Castangia, S. Cattani, A. Prioletti, and P. Versari, "360 degree; detection and tracking algorithm of both pedestrian and vehicle using fisheye images," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, June 2015, pp. 132–137.
[17] Q. Ye, J. Liang, and J. Jiao, "Pedestrian detection in video images via error correcting output code classification of manifold subclasses," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 193–202, March 2012.
[18] S. R. D. Meissner and K. Dietmayer, "Real-time detection and tracking of pedestrians at intersections using a network of laserscanners," in *2012 IEEE Intelligent Vehicles Symposium*, June 2012, pp. 630–635.
[19] L. Jin, Q. Niu, H. Hou, H. Shunxi, and W. Fangrong, "Study on vehicle front pedestrian detection based on 3d laser scanner," in *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, Dec 2011, pp. 735–738.
[20] S. F. Lin and C. H. Lee, "Pedestrians and vehicles recognition based on image recognition and laser distance detection," in *2016 16th International Conference on Control, Automation and Systems (ICCAS)*, 2016, pp. 1232–1237.
[21] J. Schlosser, C. K. Chow, and Z. Kira, "Fusing lidar and images for pedestrian detection using convolutional neural networks," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 2198–2205.
[22] T. Gandhi and M. M. Trivedi, "Computer vision and machine learning for enhancing pedestrian safety," in *Computational Intelligence in Automotive Applications*, D. Prokhorov, Ed. Berlin: Springer Berlin Heidelberg, 2008.
[23] B. Völz, K. Behrendt, H. Mielenz, I. Gilitschenski, R. Siegwart, and J. Nieto, "A data-driven approach for pedestrian intention estimation," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 2607–2612.
[24] R. Quintero, I. Parra, D. Llorca, and M. Sotelo, "Pedestrian intention and pose prediction through dynamical models and behaviour classification," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 83–88.
[25] E. Ohn-Bar and M. M. Trivedi, "Looking at humans in the age of self-driving and highly automated vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 90–104, 2016.
[26] M. S. Shirazi and B. Morris, "Observing behaviors at intersections: A review of recent studies amp; developments," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, 2015, pp. 1258–1263.
[27] V. Karasev, A. Ayvaci, B. Heisele, and S. Soatto, "Intent-aware long-term prediction of pedestrian motion," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 2543–2549.
[28] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part IV*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 201–214.
[29] N. Deo and M. M. Trivedi, "Learning and predicting on-road pedestrian behavior around vehicles," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–6.
[30] S. Schmidt and B. Färber, "Pedestrians at the kerb – recognising the action intentions of humans," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 4, pp. 300 – 310, 2009.
[31] B. Färber, *Communication and Communication Problems Between Autonomous Vehicles and Human Drivers.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 125–144.
[32] E. Rehder and H. Kloeden, "Goal-directed pedestrian prediction," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, 2015, pp. 139–147.
[33] N. Brouwer, H. Kloeden, and C. Stiller, "Comparison and evaluation of pedestrian motion models for vehicle safety systems," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2016, pp. 2207–2212.
[34] D. M. Gavrila, "The visual analysis of human movement: A survey," *Computer Vision and Image Understanding*, vol. 73, pp. 82–98, 1999.
[35] S. Köhler, M. Goldhammer, S. Bauer, K. Doll, U. Brunsmann, and K. Dietmayer, "Early detection of the pedestrian's intention to cross the street," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, 2012, pp. 1759–1764.
[36] R. Furuhashi and K. Yamada, "Estimation of street crossing intention from a pedestrian's posture on a sidewalk using multiple image frames," in *The First Asian Conference on Pattern Recognition*, Nov 2011, pp. 17–21.
[37] J. Hariyono and K. Jo, "Detection of pedestrian crossing road: A study on pedestrian pose recognition," *Neurocomputing*, vol. 234, no. Supplement C, pp. 144 – 153, 2017.
[38] R. Quintero, I. Parra, D. F. Llorca, and M. A. Sotelo, "Pedestrian path prediction based on body language and action classification," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2014, pp. 679–684.
[39] J. Hariyono and K.-H. Jo, "Pedestrian action recognition using motion type classification," in *2015 IEEE 2nd International Conference on Cybernetics (CYBCONF)*, 2015, pp. 129–132.
[40] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E*, vol. 51, pp. 4282–4286, May 1995.
[41] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 261–268.
[42] Y. Tamura, P. D. Le, K. Hitomi, N. P. Chandrasiri, T. Bando, A. Yamashita, and H. Asama, "Development of pedestrian behavior model taking account of intention," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2012, pp. 382–387.
[43] W. Zeng, P. Chen, H. Nakamura, and M. Iryo-Asano, "Application of social force model to pedestrian behavior analysis at signalized crosswalk," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 143 – 159, 2014.
[44] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 961–971.
[45] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking system with Kalman filtering," in *IEEE Intelligent Vehicles Symposium, 2004*, June 2004, pp. 584–589.
[46] A. Møgelmose, M. M. Trivedi, and T. B. Moeslund, "Trajectory analysis and prediction for improved pedestrian safety: Integrated framework and evaluations," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, June 2015, pp. 330–335.
[47] J. Hariyono, A. Shahbaz, and K. H. Jo, "Estimation of walking direction for pedestrian path prediction from moving vehicle," in *2015 IEEE/SICE International Symposium on System Integration (SII)*, Dec 2015, pp. 750–753.
[48] H. Kataoka, Y. Aoki, Y. Satoh, S. Oikawa, and Y. Matsui, "Fine-grained walking activity recognition via driving recorder

dataset," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Sept 2015, pp. 620–625.

[49] C. G. Keller and D. M. Gavrila, "Will the pedestrian cross? a study on pedestrian path prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 494–506, 2014.

[50] J. F. P. Kooij, N. Schneider, and D. M. Gavrila, "Analysis of pedestrian dynamics from a vehicle perspective," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 1445–1450.

[51] G. Best and R. Fitch, "Bayesian intention inference for trajectory prediction with an unknown goal destination," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2015, pp. 5817–5823.

[52] M. Goldhammer, M. Gerhard, S. Zernetsch, K. Doll, and U. Brunsmann, "Early prediction of a pedestrian's trajectory at intersections," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2013, pp. 237–242.

[53] Schulz and R. Stiefelhagen, "Pedestrian intention recognition using latent-dynamic conditional random fields," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, 2015, pp. 622–627.

[54] Y. Hashimoto, G. Yanlei, L. T. Hsu, and K. Shunsuke, "A probabilistic model for the estimation of pedestrian crossing behavior at signalized intersections," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 1520–1526.

[55] Y. Hashimoto, Y. Gu, L. T. Hsu, and S. Kamijo, "Probability estimation for pedestrian crossing intention at signalized crosswalks," in *2015 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 2015, pp. 114–119.

[56] M. Goldhammer, S. Köhler, K. Doll, and B. Sick, "Camera based pedestrian path prediction by means of polynomial least-squares approximation and multilayer perceptron neural networks," in *2015 SAI Intelligent Systems Conference (IntelliSys)*, 2015, pp. 390–399.

[57] J.-Y. Kwak, B. C. Ko, and J.-Y. Nam, "Pedestrian intention prediction based on dynamic fuzzy automata for vehicle driving at nighttime," *Infrared Physics and Technology*, vol. 81, no. Supplement C, pp. 41 – 51, 2017.

[58] M.-S. Cloutier, U. Lachapelle, A.-A. d'Amours Ouellet, J. Bergeron, S. Lord, and J. Torres, ""outta my way!" individual and environmental correlates of interactions between pedestrians and vehicles during street crossings," vol. 104, no. Supplement C, 2017, pp. 36 – 45.

[59] J. Hariyono and K. H. Jo, "Detection of pedestrian crossing road using action classification model," in *2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, July 2015, pp. 21–24.

[60] K. Kim, Y. Owechko, and S. Medasani, "Active safety and collision alerts using contextual visual dataspace," in *2010 IEEE Intelligent Vehicles Symposium*, June 2010, pp. 424–430.

[61] Y. Gu, Y. Hashimoto, L. T. Hsu, and S. Kamijo, "Motion planning based on learning models of pedestrian and driver behaviors," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2016, pp. 808–813.

[62] B. Völz, H. Mielenz, R. Siegwart, and J. Nieto, "Predicting pedestrian crossing using quantile regression forests," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 426–432.

[63] J. Bock, J. Kotte, T. Beemelmanns, and M. Klösges, "Self-learning trajectory prediction with recurrent neural networks at intelligent intersections," in *Proceedings of the 3rd International Conference on Vehicle Technology and Intelligent Transport Systems*, 2017, pp. 346–351.

[64] B. Völz, H. Mielenz, G. Agamennoni, and R. Siegwart, "Feature relevance estimation for learning pedestrian behavior at crosswalks," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Sept 2015, pp. 854–860.

[65] R. Quintero, I. Parra, J. Lorenzo, D. Fernández-Llorca, and M. A. Sotelo, "Pedestrian intention recognition by means of a hidden markov model and body language," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–7.

[66] S. Koehler, M. Goldhammer, S. Bauer, S. Zecha, K. Doll, U. Brunsmann, and K. Dietmayer, "Stationary detection of the pedestrian's intention at intersections," *IEEE Intelligent Transportation Systems Magazine*, vol. 5, no. 4, pp. 87–99, 2013.

[67] J. Li, Q. Li, N. Chen, and Y. Wang, "Indoor pedestrian trajectory detection with lstm network," in *Computational Science and Engineering (CSE) and Embedded and Ubiquitous Computing (EUC), 2017 IEEE International Conference on*, vol. 1. IEEE, 2017, pp. 651–654.

[68] A. Dominguez-Sanchez, M. Cazorla, and S. Orts-Escolano, "Pedestrian movement direction recognition using convolutional neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3540–3548, 2017.

[69] E. Rehder, F. Wirth, M. Lauer, and C. Stiller, "Pedestrian prediction by planning using deep neural networks," in *IEEE Int. Conf. International Conference on Robotics and Automation*, 05 2018.

[70] L. Mantecchini and F. Paganelli, "Empirical analysis of pedestrian delay models at urban intersections," vol. 8, pp. 981–990, 11 2015.

[71] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta, and G. Rizzolatti, "Grasping the intentions of others with one's own mirror neuron system," *PLOS Biology*, vol. 3, 02 2005.

[72] C. G. Keller, C. Hermes, and D. M. Gavrila, "Will the pedestrian cross? probabilistic path prediction based on learned motion features," in *Pattern Recognition: 33rd DAGM Symposium, Frankfurt/Main, Germany, August 31 – September 2, 2011. Proceedings*, R. Mester and M. Felsberg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 386–395.

[73] P. Vasishta, D. Vaufreydaz, and A. Spalanzani, "Natural vision based method for predicting pedestrian behaviour in urban environments," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–6.

[74] M. Tarawneh, "Evaluation of pedestrian speed in jordan with investigation of some contributing factors," *Journal of safety research*, vol. 32, no. 2, pp. 229–236, 2001.

[75] A. Sobhani, B. Farooq, and Z. Zhong, "Distracted pedestrians crossing behaviour: Application of immersive head mounted virtual reality," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–6.

[76] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Are they going to cross? a benchmark dataset and baseline for pedestrian crosswalk behavior," in *The IEEE International Conference on Computer Vision (ICCV)*.

[77] D. Fernandez-Llorca, R. Q. Minguez, I. P. Alonso, C. F. Lopez, I. G. Daza, M. A. Sotelo, and C. A. Cordero, "Assistive intelligent transportation systems: The need for user localization and anonymous disability identification," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 2, pp. 25–40, 2017.

[78] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua, "Multiple Object Tracking using K-Shortest Paths Optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.

[79] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multi-Camera People Tracking with a Probabilistic Occupancy Map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 267–282, February 2008.

[80] J. Ferryman and A. Shahrokni, "Pets2009: Dataset and challenge," in *Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, 2009, pp. 1–6.

[81] A. Ess, B. Leibe, K. Schindler, , and L. van Gool, "A mobile vision system for robust multi-person tracking," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08)*. IEEE Press, June 2008.

[82] M. Andriluka, S. Roth, and B. Schiele, "Monocular 3d pose estimation and tracking by detection," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 623–630.

[83] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[84] Semcon, "Self-driving car smiles at pedestrians," in *Attitudes to self-driving cars*, 2016.