2016 IEEE Intelligent Vehicles Symposium (IV)
Gothenburg, Sweden, June 19-22, 2016

Driver and Pedestrian Awareness-based Collision Risk Analysis

Markus Roth[†,1], Fabian Flohr[‡,1] and Dariu M. Gavrila[‡,1,2]

***Abstract*— We present a novel approach for vehicle-pedestrian collision risk analysis that incorporates mutual situational awareness, a degree of potential motion coupling and the spatial layout of the environment. The approach uses a Dynamic Bayesian Network (DBN) for modeling the individual object paths; collision risk is subsequently computed by an intersection operation.**

More specifically, the proposed DBN consists of two sub-graphs for modeling pedestrian and vehicle path, respectively. They consist of latent states on top of Switching Linear Dynamical Systems (SLDSs) to anticipate changes in object dynamics. The pedestrian- and vehicle-related sub-graphs contain latent states to model whether the pedestrian has seen the oncoming vehicle, and conversely, whether the driver has seen the pedestrian (associated measurements involve the respective head orientations). The pedestrian-related sub-graph furthermore contains a latent state modeling whether the pedestrian is at the curbside or not. Finally, a latent state is shared by the two sub-graphs, which models the potential motion coupling (i.e. at full awareness of the other traffic participant).

We consider the scenario of a crossing pedestrian, who might stop or continue walking at the curb, in combination with an approaching vehicle, that might stop or continue driving. In experiments we illustrate that with the proposed approach, a more anticipatory driver warning and/or vehicle control strategy can be implemented.

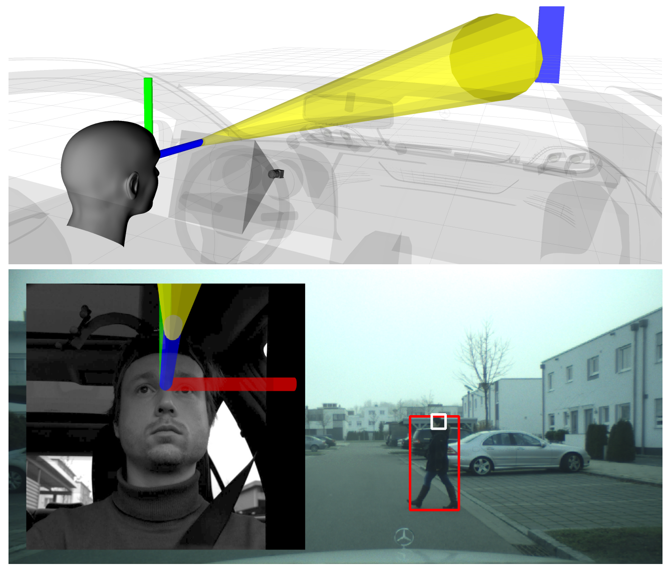

Fig. 1. Looking-in and looking-out experimental vehicle set-up

I. Introduction

The past decade has seen a significant progress on active pedestrian safety leading to systems being deployed in most mid- to premium-range vehicles on the market. An area that holds major potential for further improvement is situation assessment.

Current active pedestrian systems are designed conservatively in their warning/control strategy in order to avoid false system activations. This is understandable due to the difficulty to model the highly dynamic pedestrian motion and due to the limited processing power available. Indeed, the basic Kalman filters that are predominant in production vehicles are efficient and adequate for (position) state estimation, and thus for a reactive warning/control strategy. However, they are less suited for (path) prediction and for implementing a more effective *anticipatory* warning/control strategy utilizing additional context features. An additional limitation of most active pedestrian systems is that they incorporate no information about the driver state which could improve vehicle path prediction. Consequently, systems need to defer action in a potentially dangerous situation till the last moment, just in case the driver might act after all.

In this paper, we present a novel approach for vehicle-pedestrian collision risk analysis that incorporates mutual situational awareness, a degree of potential motion coupling and the spatial layout of the environment. See Fig. 1 for an illustration. The approach uses a Dynamic Bayesian Network (DBN) for modeling the individual object paths and their dependencies.

II. Previous Work

We first cover approaches for pedestrian state estimation, path prediction and action classification (for a survey on vision-based pedestrian detection, see e.g. [1]). The various approaches can be distinguished by the type of features they use: position, low-level features derived from a local region of interest (e.g. optical flow, intensity profiles inside a bounding box provided by a pedestrian detector), and semantic features (i.e. related to body pose, scene context, or goals).

Approaches that use position-only features can apply existing state estimation theory. State estimation in dynamical systems often involves the assumption that the underlying model is linear and that the noise is Gaussian, and allows the use of the Kalman filter (KF) [2] as an efficient inference algorithm for such Linear Dynamical Systems (LDS). In the intelligent vehicle domain, the KF is the most popular

† Markus.R.Roth@daimler.com
‡ Fabian.Flohr@daimler.com
‡ Dariu.Gavrila@daimler.com
1 Environment Perception, Daimler R&D, Ulm, Germany
2 Informatics Institute, Univ. of Amsterdam, The Netherlands

choice for pedestrian tracking (see [3] for an overview, and a public dataset). The state distribution of a LDS can be propagated into the future without incorporating new observations to account for missing measurements, or to perform path prediction. Switching LDS (SLDS) are needed for maneuvering targets that alternate various motion types. SLDSs use a top-level discrete Markov chain to select per time step the system dynamics of the underlying LDS.

Within the class of approaches that use position and low-level features, [4] proposed two non-linear, higher order Markov models to estimate whether a crossing pedestrian will stop at the curbside, one using Gaussian Process Dynamical Models (GPDM), and one using Probabilistic Hierarchical Trajectory Matching (PHTM). Both models use dense optical flow features in the bounding box obtained by a stereo vision-based pedestrian detector. [5] also uses GPDM but regresses on 3D articulated body pose, rather than on optical flow images (like [4]). [6] detects a pedestrian's action (starting, stopping, bending-in) based on the motion history of silhouettes.

In terms of approaches that use semantic features, pedestrian head/torso orientation is a valuable cue for inferring pedestrian awareness and most likely path. For pedestrian head/body orientation estimation from a moving vehicle, see e.g. [7], [8]. [9] presents a Dynamic Bayesian Network (DBN) which incorporates the pedestrian situational awareness, the potential motion coupling ("situation criticality") and spatial layout of the environment as latent states on top of a SLDS to anticipate changes in the pedestrian dynamics. [10] use multiple features capturing the pedestrian dynamics and the awareness of the nearby traffic situation in order to learn a Latent-dynamic Conditional Random Field model.

Influences on pedestrian behavior can be captured on an individual level using agent models. These have been used to reason about pedestrian intent [11], [12], [13] (i.e. where does observed agent want to go), account for preferences to move around certain regions of a static scene [12], [13], and to avoid collision with static obstacles and other agents, as is done in social force models [14], [15]. [16] enhanced social force towards group behavior by introducing sub-goals such as "following a person". The related Linear Trajectory Avoidance model [17] for short-term path prediction uses the expected point of closest approach to foreshadow and avoid possible collisions. These agent-based models assume that pedestrians are fully aware of their environment [12], [13], [17]. However, this assumption does not hold when dealing with inattentive pedestrians in the intelligent vehicle context. [18] presented a study on head turning behaviors at pedestrian crosswalks regarding the best point of warning for inattentive pedestrians.

In a different line of research, there has been work on the so-called looking-in and looking-out framework (LiLo) (cf. Mohan Trivedi lab [19]), where sensors simultaneous capture the surrounding environment of a vehicle, while monitoring the driver and measuring the vehicle dynamic state. Some work along this line specifically addresses the driver awareness of a pedestrian [20], [21], [22].

III. Proposed Approach

We are interested in an accurate path prediction of both vehicle and pedestrian for calculating collision risk. In [9], pedestrian motion dynamics is modeled by a DBN with a structure of a context-based SLDS. The transition matrix of the SLDS switching state is conditioned on latent factors that are likely going to influence the pedestrians motion type. In a scenario of a lateral crossing pedestrian, [9] argues that the pedestrians decision to continue walking or to stop is largely influenced by the existence of an approaching vehicle on collision course, the pedestrians awareness thereof, and the position of the pedestrian with respect to the curbside.

In this paper, we extend the context-based SLDS for the pedestrian of [9] with one context-based SLDS for the vehicle (which we describe in more detail below). Analogously, we argue that the vehicle's dynamics will be largely influenced by whether the driver has seen the pedestrian or not, and whether the vehicle is on collision course with the pedestrian. We combine the two context-based SLDS into a joint DBN, where the collision course latent state (termed SC "Situation Critical") is shared. Taking into account pedestrian and driver awareness combined, sets this work apart from previous pedestrian path prediction works.

A. Graphical Model

The proposed DBN is shown in Fig. 2. We distinguish two sets of variables: those relating to a pedestrian/vehicle SLDS (consisting of switching state M^P/M^V, latent position state X^P/X^V and associated observation Y^P/Y^V) and those related to the pedestrian/vehicle context, i.e. spatial layout, situation criticality, the pedestrian's and the driver's awareness. More precisely the modeled context consists of discrete latent variables $Z = \{SP, HSP, SV, HSV, SC, AC\}$ that influence the SLDS switching state, and associated observables $E = \{DHO, HO, D^{min}, DTC\}$. Details on the computation of observables are given towards the end of this section. Below, we only discuss the vehicle context-based SLDS $\{SP, HSP, SC\}$. For the pedestrian context-based SLDS $\{SV, HSV, SC, AC\}$, and for DBN inference, the reader is referred to [9].

We will assume fixed time-intervals, and from here on set $\Delta t = 1$. The SLDS filtering has the form

$$X_t = A^{(M_t)} X_{t-1} + \begin{bmatrix} \epsilon_t \\ 0 \end{bmatrix} \qquad \epsilon_t \sim \mathcal{N}(0, Q^{(M)}) \quad (1)$$

$$Y_t = C X_t + \eta_t \qquad \eta_t \sim \mathcal{N}(0, R) \quad (2)$$

where the switching state M_t selects the appropriate linear state transformation $A^{(m)}$, Y_t is the observable, X_t the latent state and C the observation matrix.

The zero-mean process noises $\epsilon_t \sim \mathcal{N}(0, Q^{(M)})$ and $\eta_t \sim \mathcal{N}(0, R)$ allow for deviations of the used motion and the observation model. From the definition of the SLDS, we obtain the following conditional probability distributions for the graphical model: $P(X_t|X_{t-1}, M_t) = \mathcal{N}(X_t|A^{(M_t)}X_{t-1}, Q^{(M_t)})$ and $P(Y_t|X_t) = \mathcal{N}(Y_t|CX_t, R)$.

For the vehicle SLDS, $Y_t^V \in \mathbb{R}$ denotes observed longitudinal vehicle position with observation matrix $C^V = [1\ 0]$.

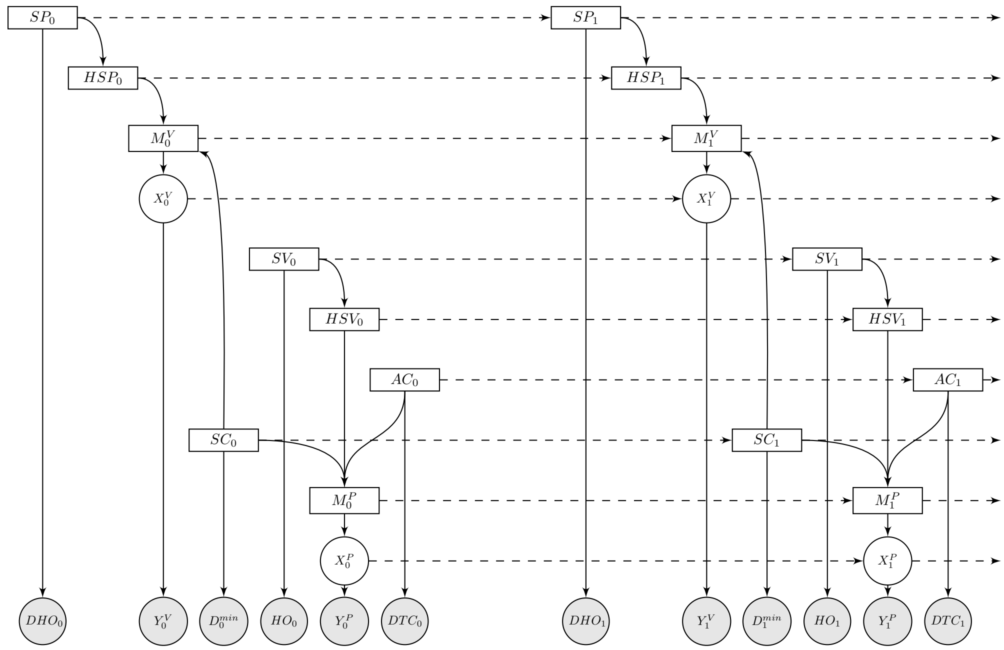

Fig. 2. Directed graphical model representation of the Dynamic Bayesian Network (DBN), unrolled for two time slices. Discrete nodes are rectangular, continuous nodes are circular, observed nodes are shaded. Dashed arcs depict temporal connections between nodes in subsequent time instances.

The latent state $X_t^V = [x_t^V, v_t^V]$ represents longitudinal vehicle position x_t^V and driving speed v_t^V. The initial distribution on the state X_0^V expresses our prior beliefs about a vehicle's longitudinal position and driving speed, as learned from the training data (see Sec. IV-B).

We assume that the vehicle can exhibit the two motion types, *braking* ($M_t^V = m_b^V$) and *driving* ($M_t^V = m_d^V$), which are implemented by constant position (CP) and constant velocity (CV) motion models, respectively. The motion dynamics are then defined by $x_t = x_{t-1} + v_t + \epsilon_t$ with velocity v_t^V:

$$v_t^V = \begin{cases} 0 & \text{iff } M_t = m_b^V \\ v^{m_d^V} & \text{iff } M_t = m_d^V. \end{cases} \tag{3}$$

The linear state transformation $A^{(m)}$ is set model dependent by:

$$A^{(m_b^V)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad A^{(m_d^V)} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{4}$$

Driver Context. The transition probability of the SLDS switching state is conditioned on the Boolean latent driver context variables Z^V where the temporal transition is factorized by the probability tables

$$P(Z_t^V|Z_{t-1}^V) = \mathcal{P}(HSP_t|HSP_{t-1}, SP_t) \times \mathcal{P}(SC_t|SC_{t-1}).$$

The latent variable *Sees-Pedestrian* (SP) indicates whether the driver is currently seeing the pedestrian. *Has-Seen-Pedestrian* (HSP) indicates whether the driver is aware of the pedestrian, i.e. whether $SP_{t'} = \text{true}$ for some $t' \leq t$. The transition probability of HSP_t encodes simply a logical OR between the Boolean HSP_{t-1} and SP_t nodes:

$$\mathcal{P}(HSP_t|HSP_{t-1}, SP_t) = \begin{cases} 1 & \text{iff } HSP_t = (HSP_{t-1} \vee SP_t) \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Next we discuss observations E_t^V which provide evidence for the latent context Z_t^V,

$$P(E_t^V|Z_t^V) = P(\theta_t|SP_t = sp). \tag{6}$$

The *Driver-Head-Orientation* observable DHO serves as evidence for the *Sees-Pedestrian* (SP) variable.

IV. Experiments

A. Dataset and Observations

Our dataset consists of 48 sequences captured by a vehicle-mounted hardware setup. Exterior images are captured by a stereo camera (baseline 26 cm, 16 fps, 1920×1024 pixels). Additionally, images of the driver are collected for verification and demonstration purposes by an interior camera mounted in front of the speedometer pointing towards the driver.

All sequences involve scenarios of a single pedestrian with the intention to cross the street in front of the approaching vehicle. A scenario captures different instantiations of pedestrian movement, pedestrian-vehicle-awareness, driver-pedestrian-awareness and vehicle motion patterns in different situation criticalities (cf. [9]).[1] An overview of scenarios can be found in Table I. Pedestrian Focus-of-Attention (FOA) corresponds to *Sees-Vehicle* (*SV*), Driver FOA corresponds to *Sees-Pedestrian* (*SP*).

For scenarios 19 and 27, the strong brake maneuver data is altered to evoke an artificial collision of vehicle and pedestrian. Therefore vehicle position and velocity are

[1] Note that no dangerous situations were performed. As in [9], "critical situation" refers to a theoretic outcome where both the approaching vehicle and pedestrian would not stop.

TABLE I

SCENARIOS CAPTURED IN THE DATASET

Scenario	Pedestrian motion	Pedestrian FOA	Driver FOA	Vehicle motion	SC
3	Cross	Away	Away	Drive	0
7	Cross	Away	Pedestrian	Drive	0
11	Cross	Vehicle	Away	Drive	0
15	Cross	Vehicle	Pedestrian	Drive	0
19	Cross	Away	Away	Drive (collision)	1
23	Cross	Away	Pedestrian	Strong brake	1
25	Stop	Vehicle	Away	Drive	1
27	Cross	Vehicle	Away	Drive (collision)	1
29	Stop	Vehicle	Pedestrian	Drive	1

replaced by interpolated values, starting from the first brake actuation.

The recordings were performed by two trained drivers and five different pedestrians in five different urban locations. Each sequence lasts several seconds (min / max / mean = 1.69s / 5.12s / 3.01s). Pedestrians are typically tracked only a couple of frames before arriving at the curbside, due to parked vehicles. Thereafter, pedestrians are unoccluded. The time for a pedestrian to reach the curb is (min / max / mean = 0.07s / 2.16s / 0.91s).

Positional ground truth of the pedestrian is obtained by manual labeling the pedestrian bounding boxes and 3D information is obtained from the median disparity [23] over the upper pedestrian body. Ego-compensated lateral pedestrian position is then used as positional observation Y^P. Despite using ground truth bounding boxes with added noise here, our system can be used with any state-of-the-art pedestrian detector.

We observe pedestrian's head orientation HO in the angular domain $[0, 360)$ using the method of [8] to infer pedestrian's focus-of-attention as done in [9]. Scenario metainformation of Table I is used for training and evaluating the context variable states. All other parameters concerning the pedestrian model are estimated like in [9] from training data based on the labeled ground truth.

The observations for *Driver-Head-Orientation* (DHO) originate from a simulated sensor for driver head orientation estimation. We sample the 3D orientation DHO_t from a 3D von Mises distribution $\mathcal{VM}(DHO^{gt}, \kappa_{DHO})$ with concentration parameter κ_{DHO}. Based on DHO_t we calculate the angular distance θ_t to the imaginary 3D line between drivers' head and the 3D pedestrian detection. The likelihood $P(\theta_t|SP_t = sp)$ is then modeled by a Gamma distribution over θ_t given SP, parametrized by shape a_{sp} and scale b_{sp},

$$P(\theta_t|SP_t = sp) = \Gamma(\theta_t|a_{sp}, b_{sp}). \quad (7)$$

The expected minimum distance D^{min} for situation criticality observation is calculated as in [17] for each time step based on current position and velocity of both pedestrian and vehicle. Vehicle position and velocity is provided by ego-motion-estimation and on-board sensors.

The curbside is detected using Hough transform [24].

B. Parameter Estimation

All DBN distribution parameters and baseline method parameters are estimated from annotated training data. The DBN subgraph consisting of the latent states SV, HSV, AC, SC, M, X^P and the visible states D^{min}, HO, Y^P and DTC are congruent to the DBN modeled in [9]. We estimate the pedestrian related parameters in the same way, i.e. (a) pedestrian motion model switching states between standing and walking by TTE and scenario information, (b) process and observation noise of pedestrian by difference between GT and measured position, (c) pedestrian head observation by thresholding the orientation obtained by a detector array, (d) distance-to-curb from curb positions and labels, and (e) prior and transition probability tables of the discrete states by enumeration of their GT labels.

The parameters of the driver-related nodes in our DBN are estimated as follows. For the distribution $P(\theta_t|SP_t = sp) = \Gamma(\theta_t|a_{sp}, b_{sp})$, we define per trajectory one value for all SP_t labels ($\forall t SP_t = true$ for trajectories where the driver is aware of the pedestrian, $\forall t SP_t = false$ otherwise), and estimate the distributions $\Gamma(\theta_t|a_{sp}, b_{sp})$ for the cases $SP_t = true$ and $SP_t = false$ from sampled angular distances θ_t between driver's head orientation and the imaginary 3D line between driver's head and the 3D pedestrian position.

The transition probability of HSP is a logical OR, thus memorizing whether the driver has seen the pedestrian at any point in the past.

C. Evaluation

We use the proposed model for joint vehicle and pedestrian path prediction. To show the benefit of using information about the driver focus of attention in an intelligent forward collision warning ADAS, we compare collision warning based on predicted vehicle and pedestrian trajectories against a baseline vehicle motion model which is not aware of the driver. We choose a LDS as a vehicle motion model baseline which is modeled by Kalman filter using a CV model. For the pedestrian path prediction, we utilize the full context-based SLDS as described in [9].

Our evaluation targets the following questions: (a) are there false alarms in terms of forward collision warning? (b) when do forward collision warnings occur? *Leave-one-out* cross validation is performed to split training and test sequences.

To show the benefit of incorporating a SLDS which uses the context information of the driver, we analyze the risk of a collision by creating predictive distributions for vehicle and pedestrian locations for future time steps. Therefore, we predict the context-based SLDS of the pedestrian and the vehicle $\delta = 16$ time steps (1s) ahead without incorporating further observations. This results in four one-dimensional Gaussian distributions (two for each model) for the vehicle (longitudinal components only) and the pedestrian (lateral component only).

To estimate a collision probability the combined predicted mean for the pedestrian $\mu_{x^P_{t+\delta}}$ and the vehicle $\mu_{x^V_{t+\delta}}$ and the corresponding combined standard deviations $\sigma_{x^P_{t+\delta}}$ and $\sigma_{x^V_{t+\delta}}$ are found by moment matching.

Let the bivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}^C, \mathbf{C}^C)$ with

$$\mu^C = \begin{bmatrix} \mu^P \\ \mu^V \end{bmatrix} \qquad C = \begin{bmatrix} \sigma^P & 0 \\ 0 & \sigma^V \end{bmatrix}. \tag{8}$$

represent the probability density for a collision. We calculate a probability for a collision by taking the integral over a collision area, which is defined by all possible intersections between vehicle and pedestrian locations. A collision warning is emitted if the collision probability is above a threshold which we empirically set to 0.8 for both SLDS and LDS.

TABLE II

NUMBER OF SEQUENCES WITH CORRECT ACTION AS PART OF TOTAL NUMBER OF SEQUENCES. BETWEEN BRACKETS, AVERAGE AND STANDARD DEVIATION OF FRAMES TILL COLLISION.

scenario	LDS	SLDS
3	5/5	5/5
7	5/5	5/5
11	4/4	4/4
15	5/5	5/5
19	5/5 (μ=14.60, σ=0.89)	5/5 (μ=13.20, σ=1.30)
23	0/5	5/5
25	5/5	5/5
27	5/5 (μ=9.25, σ=3.50)	5/5 (μ=9.00, σ=3.16)
29	5/5	5/5

We exemplarily discuss three scenarios (see Table II). Scenario 19 (pedestrian crosses, driver does not see pedestrian, vehicle continues driving) represents a collision between vehicle and pedestrian[2]. Since the driver is not looking at the pedestrian, vehicle SLDS infers that the most likely vehicle motion model is CV. Collision warnings are consistently emitted as a result of the predicted position distributions of vehicle and pedestrian. As expected, the vehicle LDS baseline model behaves the same, since it always assumes CV. See Figure 5.

In scenario 23 (pedestrian crosses, driver sees pedestrian, vehicle stops) the driver is aware of the pedestrian crossing the street. Our SLDS based vehicle model predicts the deceleration induced by the driver and gives more weight to the CP model. No collision warning is emitted. See Figures 3 and 4 for latent variable states of the DBN during prediction and spatial layout of predicted collision probability. The vehicle LDS baseline, however, emits false alarms in 5 of 5 sequence instances of scenario 23. Through the CV model, a collision is predicted. It cannot adapt fast enough to the decelerating vehicle, induced by the aware driver.

In scenario 29 (pedestrian stops, driver sees pedestrian, vehicle continues driving) the pedestrian sees the driver and stops at the curb. Neither in case of vehicle SLDS nor LDS a warning is emitted, due to the path predicted by the pedestrian SLDS. See Figure 6.

[2]Note that, as stated above, no real collision took place. Instead, vehicle position and velocity have been altered to evoke a collision scenario.

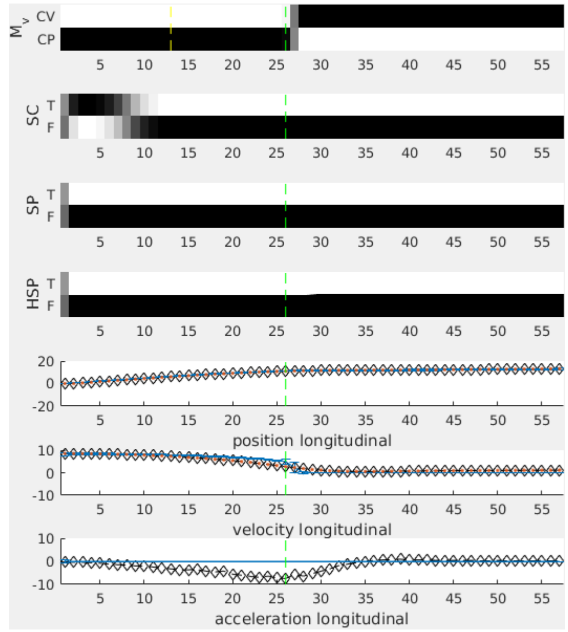

Fig. 3. Scenario 23. DBN model variable states (top) and the camera image (bottom) for a vehicle motion model switch. The DBN switches into CP vehicle motion model at frame 26.

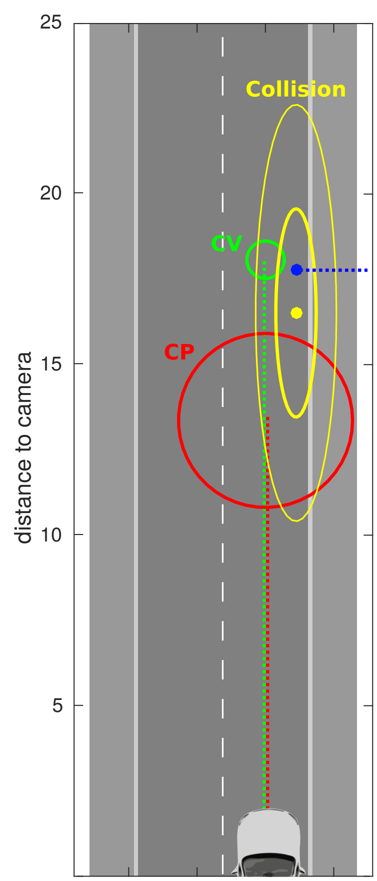

Fig. 4. Scenario 23. Spatial layout of predicted probability distributions for both vehicle motion models (red: CP, green: CV). Yellow ellipse denotes collision probability distribution. Blue dot denotes predicted pedestrian position. Dashed lines represent trajectories. Prediction of frame 28, with last observations from frame 12 ($\delta = 16 = 1$s).

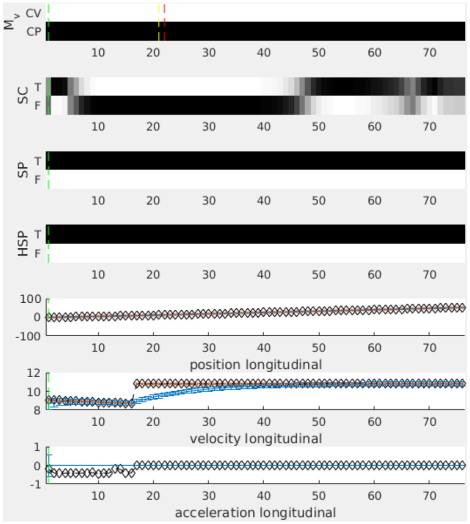

Fig. 5. Scenario 19. DBN model variable states.

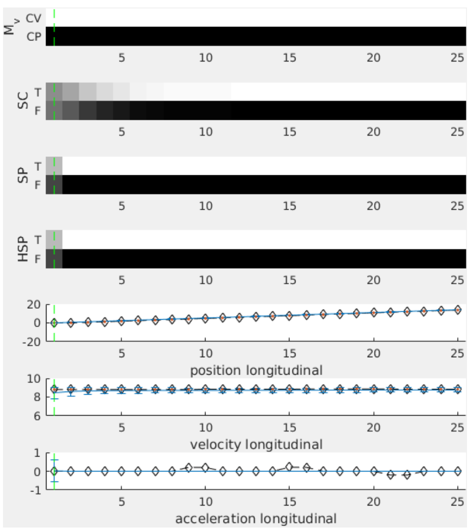

Fig. 6. Scenario 29. DBN model variable states.

V. Conclusions

We presented a novel approach for vehicle-pedestrian collision risk analysis that incorporated mutual situational awareness between driver and pedestrian, a degree of potential motion coupling and the spatial layout of the environment. The approach used a Dynamic Bayesian Network (DBN) for modeling the individual object paths; collision risk was subsequently computed by an intersection operation. The proposed DBN performed correctly on 9 different scenarios (48 recorded sequences) representing a variety of possible interactions between driver and pedestrian. Among others it was able to (a) correctly predict a collision in case of an unaware driver and (b) to suppress collision risk in case of an aware driver, compared to simpler, and context-free LDS vehicle motion model. Further work includes testing the proposed approach on larger datasets and improving video sensing.

References

[1] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *IEEE Trans. on PAMI* , vol. 31, no. 12, pp. 2179–2195, 2009.

[2] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House Norwood, 1999.

[3] N. Schneider and D. M. Gavrila, "Pedestrian path prediction with recursive Bayesian filters: A comparative study," in *Lecture Notes in Computer Science (LNCS)*. Springer, 2013, vol. 8142, pp. 174–183.

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

[5] D. F. L. R. Quintero, I. Parra and M. A. Sotelo, "Pedestrian intention and pose prediction through dynamical models and behaviour classification," in *Proc of the IEEE ITSC* , 2015.

[6] S. Köhler, M. Goldhammer, K. Zindler, K. Doll, and K. Dietmayer, "Stereo-vision-based pedestrians intention detection in a moving vehicle," in *Proc of the IEEE ITSC* , 2015.

[7] M. Enzweiler and D. M. Gavrila, "Integrated pedestrian classification and orientation estimation," in *Proc. of the IEEE CVPR* . IEEE, 2010, pp. 982–989.

[8] F. Flohr, M. Dumitru-Guzu, J. F. P. Kooij, and D. M. Gavrila, "A probabilistic framework for joint pedestrian head and body orientation estimation," *IEEE Trans. on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1872–1882, 2015.

[9] J. F. P. Kooij, N. Schneider, F. Flohr, and D. M. Gavrila, "Context-based pedestrian path prediction," in *Proc. of the ECCV*, vol. 8694. Springer, 2014, pp. 618–633.

[10] A. Schulz and R. Stiefelhagen, "Pedestrian intention recognition using latent-dynamic conditional random fields," in *Proc. of the IEEE Intelligent Vehicles Symposium* , 2015.

[11] T. Bandyopadhyay, K. Won, E. Frazzoli, D. Hsu, W. Lee, and D. Rus, "Intention-aware motion planning," in *Algorithmic Foundations of Robotics X*. Springer, 2013, pp. 475–491.

[12] K. Kitani, B. Ziebart, J. Bagnell, and M. Hebert, "Activity forecasting," in *Proc. of the ECCV*. Springer, 2012, pp. 201–214.

[13] E. Rehder and H. Kloeden, "Goal-directed pedestrian prediction," in *IEEE ICCV Workshop*, 2015.

[14] G. Antonini, S. V. Martinez, M. Bierlaire, and J. P. Thiran, "Behavioral priors for detection and tracking of pedestrians in video sequences," *IJCV* , vol. 69, no. 2, pp. 159–180, 2006.

[15] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E*, vol. 51, no. 5, p. 4282, 1995.

[16] Y. Tamura, P. D. Le, K. Hitomi, N. Chandrasiri, T. Bando, A. Yamashita, and H. Asama, "Development of pedestrian behavior model taking account of intention," in *IEEE IROS*, 2012, pp. 382–387.

[17] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *Proc. of the IEEE ICCV*, 2009, pp. 261–268.

[18] H. Hamaoka, T. Hagiwara, M. Tada, and K. Munehiro, "A study on the behavior of pedestrians when confirming approach of right/left-turning vehicle while crossing a crosswalk," in *Proc. of the IEEE Intelligent Vehicles Symposium*, 2013, pp. 106–110.

[19] A. Tawari, S. Martin, and M. Trivedi, "Continuous head movement estimator for driver assistance: Issues, algorithms, and on-road evaluations," *IEEE Trans. on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 818–830, 2014.

[20] M. Phan, V. Fremont, I. Thouvenin, M. Sallak, and V. Cherfaoui, "Recognizing driver awareness of pedestrian," in *Proc of the IEEE ITSC* , 2014.

[21] A. Tawari, A. Mogelmose, S. Martin, T. Moeslund, and M. Trivedi, "Attention estimation by simultaneous analysis of viewer and view," in *Proc of the IEEE ITSC* , 2014.

[22] K. Tateiwa and K. Yamada, "Estimating driver awareness of pedestrians in crosswalk in the path of right or left turns at an intersection from vehicle behavior," in *Proc. of the IEEE Intelligent Vehicles Symposium* , 2015.

[23] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. on PAMI* , vol. 30, no. 2, pp. 328–341, 2008.

[24] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, 1972.