

Will the Pedestrian Cross?

Probabilistic Path Prediction Based on Learned Motion Features

Christoph G. Keller¹, Christoph Hermes², and Dariu M. Gavrilă^{3,4}

¹ Image & Pattern Analysis Group, Univ. of Heidelberg, Germany

² Applied Informatics Group, Univ. of Bielefeld, Germany

³ Environment Perception, Group Research, Daimler AG, Ulm, Germany

⁴ Intelligent Systems Lab, Fac. of Science, Univ. of Amsterdam, The Netherlands

Abstract. Future vehicle systems for active pedestrian safety will not only require a high recognition performance, but also an accurate analysis of the developing traffic situation. In this paper, we present a system for pedestrian action classification (walking vs. stopping) and path prediction at short, sub-second time intervals. Apart from the use of positional cues, obtained by a pedestrian detector, we extract motion features from dense optical flow. These augmented features are used in a probabilistic trajectory matching and filtering framework.

The vehicle-based system was tested in various traffic scenes. We compare its performance to that of a state-of-the-art IMM Kalman filter (IMM-KF), and for the action classification task, to that of human observers, as well. Results show that human performance is best, followed by that of the proposed system, which outperforms the IMM-KF and the simpler system variants.

1 Introduction

Strong gains have been made over the years in improving pedestrian recognition performance. However, the initiation of an emergency vehicle manoeuvre requires a precise estimation of the current and future position of the pedestrian with respect to the moving vehicle. A deviation of, say, 25 cm in the estimated lateral position of the pedestrian can make all the difference between a “correct” and an “incorrect” manoeuvre initiation. One major challenge is the highly dynamic behaviour of pedestrians, which can change their walking direction in an instance, or start/stop walking abruptly. As a consequence, prediction horizons for active pedestrian systems are typical short; even so, small performance improvements can produce tangible benefits. For example, accident analysis [15] shows that being able to initiate emergency braking 0.16 s (4 frames @ 25 Hz) earlier, at a Time-to-Collision of 0.66 s, reduces the chance of incurring injury requiring hospital stay from 50% to 35%, given an initial vehicle speed of 50 km/h.

The paper studies the case of a pedestrian walking towards the road curbside, and poses the question whether the pedestrian will cross or not. See Figure 1. This setting is inspired by an earlier human factors study [16], where participants were asked the same question, upon viewing similar video content. This study varied the amount of visual information provided to the participants and examined its effect on human classification performance; in the baseline case, the pedestrian was fully visible, whereas in other cases, parts of the images were masked out. Masking the complete pedestrian, and leaving only positional information (bounding box), turned out to decrease human accuracy markedly.

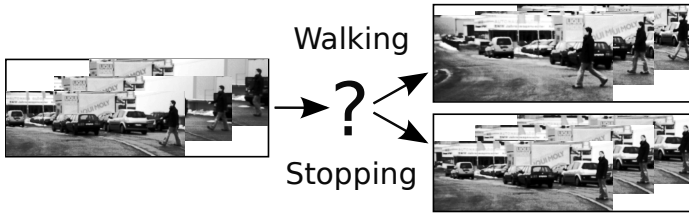


Fig. 1. Everyday problems vehicle drivers are faced with: Will the pedestrian cross?

Besides action classification, the aim of this paper is to accurately predict the pedestrian path. For this, an input pedestrian trajectory fragment is matched against previously stored pedestrian trajectories in a database. The retrieved trajectories in the database are used for extrapolation and path prediction. The intuition is that more complex dynamics (e.g. the process of a pedestrian stopping) are more accurately captured by such learning-based approaches than with generic modeling approaches relying on constant motion assumptions.

2 Previous Work

In this section, we focus on previous work on pedestrian motion models and path prediction. The reader is referred to [9,10] for surveys on pedestrian detection. There is furthermore a wealth of work on human activity recognition [18], including some on detecting unusual pedestrian motion patterns (e.g. [14]).

One way to perform path prediction relies on closed-form solutions for Bayesian filtering; in the Kalman Filter (KF) [3], the current state can be propagated to the future by means of the underlying linear dynamical model, without the incorporation of new measurements. The same idea can be applied to filter extensions, which involve multiple linear dynamical models, i.e. to the Interacting Multiple Model Kalman Filter (IMM-KF) [3]. An alternative approach for path prediction involves stochastic models. Possible trajectories are generated by Monte Carlo simulations, taking into account the respective dynamical models. For example [1], use Bayesian filtering by means of particle filtering and a constant motion model. [8] distinguishes lateral and longitudinal pedestrian velocity and models these independently by a random walk. In [19] pedestrian motion is modeled by means of four states of a Markov chain, corresponding to standing still, walking, jogging and running. Each state has associated probability distributions magnitude and direction of pedestrian velocity; the state changes are controlled by various transition probabilities. More recently, complex pedestrian motion models also account for group behaviour and spatial lay-out (e.g. entry/exit points) [2].

The limited amount of available training data precludes the use of modeling approaches which compute joint probability distributions over time intervals. Indeed, most pedestrian motion models consists of states that correspond to single time steps, and are first-order Markovian. This potentially limits their expressiveness and precision. In contrast [4] describe an extension of particle filtering to incrementally match trajectory models to input data. It is used for motion classification of 2D gestures and expression. A further development [17] adds an efficient tree search in the context of

articulated 3D human pose recovery. In [13] this technique is applied to vehicle motion prediction, utilizing the quaternion-based rotationally invariant longest common sub-sequence (QRLCS) metric for trajectory matching. [5] propose a multi-level prediction model, in which the higher levels are long-term predictions based on trajectory clustering matching, whereas the low level uses an Auto-Regressive model to predict the next time step.

The main contribution of this paper is a system for pedestrian action classification (walking vs. stopping) and accurate path prediction from a moving vehicle, at short time intervals. We borrow from the efficient trajectory matching approaches of [13,17] but augment the underlying features to include motion cues. The latter is motivated by the before-mentioned study [16]. For our use in trajectory matching, however, we need to derive a low-dimensional optical flow representation. We use real video data, which results in a realistic modeling of sensor uncertainty.

3 Proposed System

An overview of the proposed system is given in Figure 2. The feature extraction step uses dense stereo [11] and dense optical flow [20] data computed over the bounding boxes returned by a pedestrian detector. Lateral and longitudinal position of the pedestrian is obtained with the center of the detector box and median disparity values inside the box, respectively. Vehicle ego motion is compensated by rotation and translation of positions to a global reference point using velocity and yaw-rate measurements from on-board sensor data. Motion features involve a low-dimensional histogram representation of optical flow. Measured pedestrian positions and motion features are subsequently used in a trajectory matching and filtering framework. From the filter state and information about the class labels of the matched trajectories, a future pedestrian position and action is derived.

Motion Features. We propose a low dimensional feature that captures flow variations on the pedestrian legs and upper body. In order to operate from a moving vehicle additional invariance to pedestrian distance and vehicle motion is important. Features are designed to allow bounding box localization errors from a pedestrian detection system. Figure 3 illustrates the feature extraction steps. Vehicle velocity and yaw-rate measurements from on-board sensor data in combination with stereo measurements are used to compute the ego-motion compensated optical flow field. Flow vectors are normalized

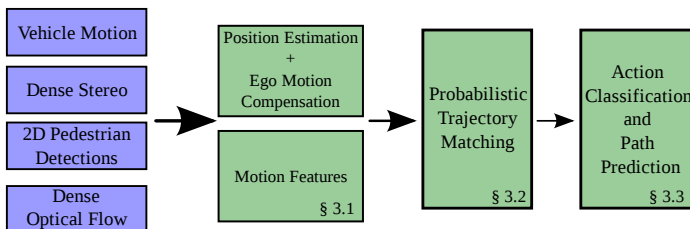


Fig. 2. Overview of the proposed system for pedestrian action classification and path prediction

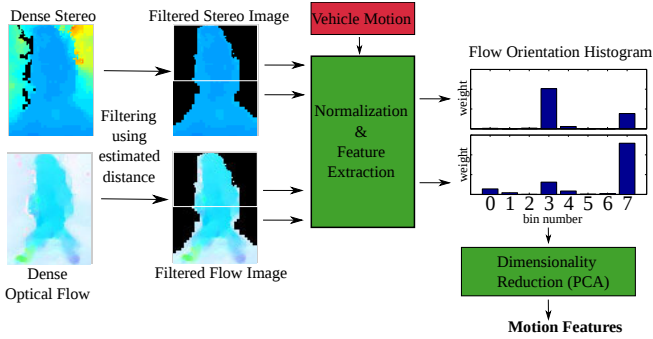


Fig. 3. Overview of the motion feature extraction procedure

with the camera cycle time and measurements from dense stereo for robustness to frame drops and invariance to different pedestrian distances to the camera. This normalized motion field is used to extract features given a bounding box detection and distance estimation z_{ped} from a pedestrian detection system. To ensure that the pedestrian is located in the box for all possible limb extensions and slight localization errors a bounding box aspect ratio of 4:3 is used. Motion vectors not belonging to the pedestrian body are suppressed by using only values at a depth similar to the estimated pedestrian distance. Remaining values in the motion field are used to compute the median object motion and extract orientation histograms. To capture motion differences between torso and legs the bounding box is split into an upper and lower sub-box. For each sub-box the median motion is removed to compensate the pedestrian ego motion. Resulting orientation vectors $v = [v_x, v_y]^T$ are assigned to bins $b \in [0, 7]$ using their 360^0 orientation $\theta = atan2(v_y, v_x)$ and bin index $b = \lfloor \frac{\theta}{\pi/4} \rfloor$. Bin contributions are weighted by their magnitude and resulting histograms are normalized with the number of contributions. A feature vector is formed by concatenating the histogram values and the median flow for the lower and upper box. Dimensionality reduction of the feature vector is archived by applying principal component analysis (PCA). The first three PCA dimensions with the largest eigenvalue are used as final histograms of orientation motion (HoM) features.

Probabilistic Trajectory Matching. A pedestrian trajectory X is represented using the ordered tuples $X = ((x_1, t_1), \dots, (x_N, t_N))$. For every timestamp t_i the state x_i consists of the lateral and longitudinal position of the pedestrian and additional features extracted from optical flow (Figure 4(a)). For motion prediction retrieval, it is possible to compare each trajectory in a motion database with an observed history using a similarity measure. With the Quaternion-based Rotationally Invariant Longest Common Subsequence (QRLCS) metric [13] the optimal translation and rotation parameters to superimpose two trajectories are derived. The distance $dist_{QRLCS}(A, B) \in [0, 1]$ between two trajectories is given by the number of possible assignments determined by an ϵ area around each trajectory state, normalised by the number of trajectory states. Figure 4(a) illustrates this matching process. We replaced this greedy search by a probabilistic search framework [13,17] where the search time depends on the number of sampling points.

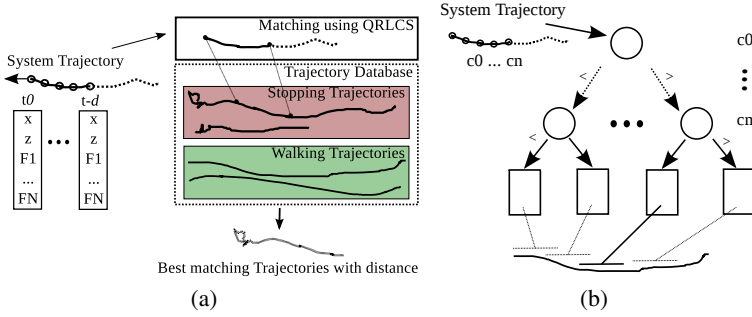


Fig. 4. a) System trajectory with history of length d containing position and feature information for every entry is matched to the training database. Resulting matching position and similarity distance to trajectories in the training database describes a possible trajectory course and class label. b) Tree representation of the trajectory training database. Leaf nodes represent trajectory snippets of fixed length. Similar trajectories are search by traversing the tree using the trajectory descriptors for every level.

Given a motion history $M_{1:t}$ up to the current time step t , the probability that a future pedestrian state ϕ_T occurs is computed by

$$p(\phi_T|M_{1:t}) = \eta p(M_{1:t}|\Psi_t) \int p(\Psi_t|\Psi_{t-1}) p(\Psi_{t-1}|M_{1:t-1}) d\Psi_{t-1} \quad (1)$$

with a normalisation constant η and the current state Ψ_t which represents a sequence of trajectory points including position, optical flow features and its history over a temporal sliding window with a manually defined number of time steps d . $p(\phi_T|\Psi_t)$ is the probability of observing a future state ϕ_T , and is determined from the motion database. This distribution $p(\phi_T|M_{1:t})$ is represented by a set of samples or particles $\{\Psi_t^{(s)}\}_S$, which are propagated in time using a particle filter [4]. Therefore, each particle $\Psi_t^{(s)}$ represents a sub-trajectory for the current state.

A set of overlapping sub-trajectories (snippets, e.g. [12]) with fixed number of trajectory points d is created from the motion database. By piling the features for each trajectory point in a snippet into a description vector and applying the PCA method to these vectors, their principal dimensions can be ordered according to the largest eigenvalue. The resulting transformed description vector \mathbf{v} is used to build a binary tree. For each level l the snippet is assigned to the left or right sub-tree depending on the sign of v_l . Figure 4(b) illustrates this search tree. Particle prediction is performed by a probabilistic search in the constructed binary tree and a lookup for the next state in the motion database. The distribution $p(M_{1:t}|\Psi_t)$ represents the likelihood that the measurement trajectory $M_{1:t}$ can be observed when the model trajectory is given. In the context of particle filters, this value corresponds to the weight of a particle and is approximated using $w^{(s)} = 1 - \text{dist}_{\text{QRLCS}}$ for each particle $\Psi_t^{(s)}$. Each particle is a representation of the assumed current pedestrian state with an assigned likelihood.

Action Classification and Path Prediction. The distribution of the predicted state $p(\Phi_t|M_{1:t})$ is approximated by means of the particle filter. An estimated state $\Phi_T^{(s)}$

Table 1. Number sequences and frames for each class with moving and non-moving ego vehicle

Sequences / Frames	vehicle standing	vehicle moving
ped. stopping	12 / 1526	5 / 587
ped. walking	9 / 1686	4 / 750

representing the pedestrian state in the future $T = t + \Delta T$ can be derived by looking ahead on the associated trajectories for the current state $\Psi_t^{(s)}$. This results in many hypotheses which are compensated using a weighted mean-shift algorithm [6] with a Gaussian kernel and weights $w^{(s)} \sim p(\Phi_T^{(s)} | M_{1:t})$.

As the final predicted state Φ_T^* the cluster center with the highest accumulated weight is selected. The trajectory database contains two classes of trajectory snippets, the class C_s in which the pedestrian is stopping and the class C_w where the pedestrian continues walking. For the predicted object state Φ_T^* derived using cluster members $L = \{\Phi_t^{(l)}\}$ and the corresponding weight $w^{(l)}$ the stopping probability can be approximated using:

$$p(C_s | L) \approx \frac{\sum_{\Phi_t^{(l)} \in C_s} w^{(l)}}{\sum_{\Phi_t^{(l)} \in C_s} w^{(l)} + \sum_{\Phi_t^{(l)} \in C_w} w^{(l)}}. \quad (2)$$

4 Experiments

Video data of two scenario types was recorded using a stereo camera (22 fps) on-board a vehicle. The first scenario features the stopping of a pedestrian at the curbstone. In the second scenario, the pedestrian crosses the street. Recorded data consisted of runs where the vehicle is stationary at a distance of 15 m to 17 m to the pedestrian and runs where the vehicle is moving at speeds of 20 – 30 km/h. Tables 1 summarizes the recorded data with four different pedestrians.

Pedestrians were shape labeled to derive the ground-truth position in the world. The median disparity value on the pedestrian upper body and the center of gravity of the shape is used to project the 3D position. For each trajectory where the pedestrian is stopping the moment of the last placement of the foot is labeled as the stopping moment. By definition, all frames earlier to this event will have a time-to-stop value (TTS in frames) greater than zero. Frames after the stopping moment have a TTS value smaller than zero. In sequences where the pedestrian continues walking the closest point to the curbstone (with closed legs) is labeled. Analogous to the TTS definition, it is called time-to-curb value (TTC). Since the focus of the paper is not a particular pedestrian detection system, we first provide as input to the evaluated methods the ground truth 2D bounding boxes perturbed by noise. Artificially generated uniform noise is added to the height and center of the 2D bounding boxes up to 10% of the original height value.

Analyzing walking trajectories shows an average gait cycle of 10 to 14 frames for different pedestrians. A trajectory database as described in previous Section is generated in a sliding window fashion to contain sub-trajectories with a fixed length of ten frames. For test trajectories a history of 14 frames is used to capture gait cycle variations. Trajectories in which the pedestrian did not stop as well as trajectory slices with a $TTS > 20$ (twice the gait cycle) are member of the class C_w . The stopping class C_s consists of the remaining slices.

Evaluated Models and Parameter Settings. The particle filter approximates the current probability density with $S = 400$ particles and uses a search deviation parameter of $\beta = 0.05$ for the tree search (see [13]). The mean shift position procedure operates with a kernel width value $h = 0.1$. Two feature combinations for the PF are compared: The base configuration uses the lateral and longitudinal object position (Pos) for the trajectory matching. The second combination uses the median horizontal object flow ($MFlowU$) in the image plane and additional motion features (HoM) described in previous Section. For the evaluation the training and testing data has been processed using *leave-one-out* cross validation. In order to ensure the optimal search result by the PF, the proposed system is compared against a full-search (i.e. brute force, BF) method over all training trajectories. For each trajectory slice from the training set, the BF method selects the best hypothesis based on the weight value, equal to the particle weights in the PF system.

As a second model of comparison to the proposed PF system, we use the state-of-the-art Integrated Multiple Model Kalman Filter (IMM-KF) [3]. Two process models describing pedestrian motion are used: A steady walking pace can be represented with the constant velocity (CV) model with process noise parameter q_{CV} . For non-moving pedestrians the constant position (CP) model with q_{CP} applies. For the following evaluation $q_{CV} = 0.009$ and $q_{CP} = 0.01$ has been derived from the set of training trajectories, with respect to the positions minimum mean-RMS error. Each model is provided with the same 3D position data as for the PF as a measurement input. In preliminary experiments a constant acceleration model (CA) was tested as well, but did not produce acceptable results. As mentioned in [3], this can be explained by the relatively short time span of the deceleration process, a comparatively large measurement noise and the use of the second derivative of the position data to estimate the deceleration. Measurement noise for each test trajectory is computed from the training trajectories and yields a standard deviation of $\sigma_z = 0.25$ m in longitudinal and $\sigma_x = 0.15$ m in lateral direction. The Markov matrix P describing the transition probabilities between the CV and CP model was derived from the available training data $P := [0.999, 0.001; 0.001, 0.999]$. Choosing larger values for the model transitions result in more frequent, undesired switches, especially with noisy measurements.

Pedestrian Action Classification. We first tested the ability of various systems to classify pedestrian actions, i.e. whether the pedestrian will cross or not. Figure 5 illustrates the mean performance of each system with different feature sets on stopping and walking test trajectories; depicted is the estimated probability of stopping, as a function of TTS or TTC. For the PF and BF approaches the stopping probability is approximated as described in previous Section. For the IMM-KF filter the probability of the CP model is used. To put the performance of the systems in context, we also evaluated human performance. Video data was presented to several human observers using a graphical user interfaces, where playback was automatically stopped at different TTC or TTS moments (20, 11, 8, 5, 3). For each test, the observers had to decide whether the pedestrian will stop at the curbstone or cross the street and provide a confidence using a slider ranging from 0 to 1. Sequence and playback stopping point were randomly selected before being presented, to avoid the effect of re-identification.

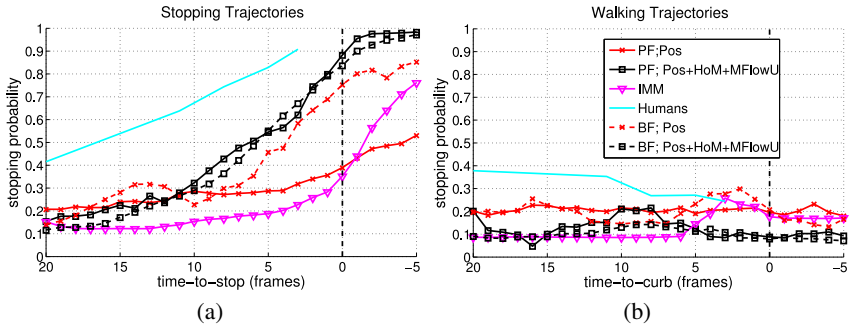


Fig. 5. Estimated probability of stopping over time for (a) stopping or (b) walking test trajectory (averaged over all respective sequences)

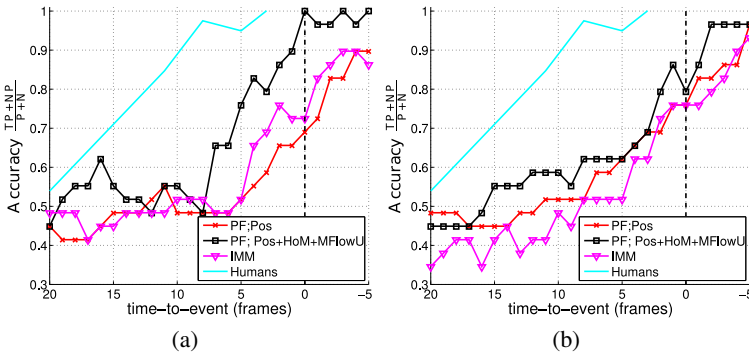


Fig. 6. Classification accuracy of the different systems over time. (a) Results for the jittered ground-truth bounding boxes. (b) Results for the bounding boxes of a HOG detector.

On trajectories where the pedestrian continues walking, all systems show a low and relatively constant stopping probability. BF and PF results for different feature combinations show a similar performance. The CV model for the IMM-KF filter remains more likely all time. On trajectories where the pedestrian is stopping, all systems initially start with a low stopping probability, since stopping is preceded by walking. But about 11 frames before the stopping event the confidence increases more markedly. Class membership of an input trajectory for each time instant is assigned by thresholding the observed stopping probability (cf. Figure 5). Based on the training set, we selected for each system a threshold that minimizes its classification error (i.e. stopping classified as walking and vice versa) over all sequences and time instants. Figure 6 illustrates the classification accuracy using these “optimal” thresholds. Human estimates of the pedestrian action class outperforms current methods. The proposed system using the feature combination Pos + HoM + MFlowU dominate the classification accuracy compared to other evaluated competitive models at all times. An accuracy of 0.8 in classifying the correct pedestrian’s action is reached 570 ms and 180 ms before a possible standstill by the human and proposed system; it is only reached *after* the possible standstill with the other methods, due to sensor noise.

Table 2. Mean and standard deviation of the RMSE (in m) for *walking* and *stopping* trajectories with different systems and prediction horizons (#frames). **left:** Results using the jittered ground truth bounding boxes **right:** Results using recognitions from a pedestrian detector.

		Walking					Stopping					Walking		Stopping	
		0	3	5	11	17	0	3	5	11	17	0	17	0	17
IMM-KF	Mean	0.62	0.73	0.80	1.06	1.33	0.51	0.66	0.77	1.14	1.54	0.64	2.37	0.38	1.52
	\pm Std	0.51	0.57	0.61	0.73	0.87	0.48	0.57	0.65	0.93	1.23	0.69	2.52	0.28	1.26
Pos	Mean	0.32	0.42	0.48	0.70	0.89	0.25	0.35	0.43	0.63	0.86	0.28	0.68	0.25	0.83
	\pm Std	0.09	0.14	0.18	0.29	0.34	0.07	0.11	0.16	0.22	0.26	0.25	0.49	0.17	0.42
Pos+HoM +MFlowU	Mean	0.37	0.49	0.56	0.79	1.07	0.28	0.04	0.46	0.66	0.88	0.43	0.99	0.31	0.89
	\pm Std	0.17	0.21	0.25	0.31	0.39	0.11	0.18	0.22	0.31	0.43	0.45	0.79	0.20	0.61

Path Prediction Accuracy. The second set of experiments evaluates the world localization accuracy of path prediction for different prediction horizons for every time step (i.e. frame). The predicted object position for the PF is computed as described in previous Section by a look-up on the subsequent parts of the matched trajectories. Position predictions of the IMM-KF are derived by predicting the current filter state without additional measurement updates. Given a prediction time step inside the range $[20, -5]$ frames, where frame 0 denotes the manually labeled TTS/TTC moment, the predicted localization error is evaluated for different prediction horizons.

Localization error for the different prediction horizons are summarized in Table 2. As can be seen, the IMM-KF has a higher localization error for any prediction horizon than PF feature combinations. Because the IMM-KF uses the filtered velocity for path prediction, the increasing localization error can be explained by erroneous velocity estimations, without new measurements the current velocity estimates are propagated unchanged. The proposed system outperforms the IMM-KF by a factor of about 1.7 and 1.2 in stopping and walking situations, respectively, in terms of the reduction of mean RMS. The addition of motion features does not result in improved accuracy, likely because the right trajectory snippets in the database are already found with position-only features; accuracy gains can be expected by a larger training set.

Evaluation using a Real Pedestrian Detector. We also evaluated performance using the bounding boxes provided by a state-of-the-art HOG pedestrian detector [7]. Missing detections were filled in using a basic correlation tracker (one sequence had to be excluded from the evaluation because of 11 successive missing detections). For IMM-KF systems the measurements noise derived from training sequences has been adjusted to $\sigma_z = 0.48$ m and $\sigma_x = 0.37$ m. Process noise was left unchanged. We found that the position and height error of the detector boxes compared to the ground truth is normally distributed with a standard deviation of 5% of the box height but with small number of outliers with larger errors. These outliers affect both position prediction and classification performance of the IMM-KF method (cf. Figure 6(b) and Table 2). For the IMM-KF the outliers lead to frequent model switches resulting in a less accurate action class decision and velocity estimation. Classification performance of all methods decreases compared to the previous experiments. With the particle filter approach no significant change in the localization accuracy is observed compared to the results obtained previously; we attribute this to the robustness of the QRLCS metric to outliers.

5 Conclusion

We presented a system for short-term pedestrian action classification and path prediction, that makes use of learned, labeled trajectory data. On the task of classifying whether a pedestrian nearing the curb side will stop, human performance was best, followed by the proposed system, and on third spot the state-of-the-art IMM-KF and simpler system variants, without an augmented motion feature set. Regarding the path prediction accuracy, our system leads to a significant lower position error, especially for large prediction horizons. We would like to thank C. Wöhler and M. Enzweiler for helpful discussions.

References

1. Abramson, Y., Steux, B.: Hardware-friendly pedestrian detection and impact prediction. In: *IEEE Intell. Veh.*, pp. 590–595 (2004)
2. Antonini, G., Martinez, S.V., Bierlaire, M., Thiran, J.: Behavioral priors for detection and tracking of pedestrians in video sequences. *IJCV* 69(2) (2006)
3. Bar-Shalom, Y., Li, X., Kirubarajan, T.: *Estimation with applications to tracking and navigation*. Wiley-Interscience, Hoboken (2001)
4. Black, M.J., Jepson, A.D.: A probabilistic framework for matching temporal trajectories: CONDENSATION-based recognition of gestures and expressions. In: Burkhardt, H.-J., Neumann, B. (eds.) *ECCV 1998*. LNCS, vol. 1406, pp. 909–924. Springer, Heidelberg (1998)
5. Chen, Z., Ngai, D., Yung, N.: Pedestrian behavior prediction based on motion patterns for vehicle-to-pedestrian collision avoidance. In: *Proc. of IEEE ITSC*, pp. 316–321 (2008)
6. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. *IEEE PAMI* 24(5), 603–619 (2002)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *Proc. CVPR*, pp. 886–893 (2005)
8. De Nicolao, G., Ferrara, A., Giacomini, L.: A collision risk assessment approach as a basis for the on-board warning generation in cars. In: *IEEE Intell. Veh.* (2002)
9. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: A benchmark. In: *Proc. CVPR* (2009)
10. Enzweiler, M., Gavrila, D.M.: Monocular pedestrian detection: Survey and experiments. *IEEE PAMI* 31, 2179–2195 (2009)
11. Hirschmüller, H.: Stereo processing by semiglobal matching and mutual information. *IEEE PAMI* 30(2), 328–341 (2008)
12. Howe, N., Leventon, M., Freeman, W.: Bayesian reconstruction of 3d human motion from single-camera video. In: *Proc. NIPS*, pp. 820–826 (2000)
13. Käfer, E., Hermes, C., Wöhler, C., Ritter, H., Kummert, F.: Recognition of situation classes at road intersections. In: *Proc. ICRA*, pp. 3960–3965 (2010)
14. Makris, D., Ellis, T.: Spatial and probabilistic modelling of pedestrian behaviour. In: *Proc. BMVC*, pp. 557–566 (2002)
15. Meinecke, M.M., et al.: Strategies in terms of vulnerable road user protection. EU Project SAVE-U, Deliverable D6 (2003), <http://www.save-u.org>
16. Schmidt, S., Färber, B.: Pedestrians at the kerb - recognising the action intentions of humans. *Transportation Research Part F* 12(4), 300–310 (2009)
17. Sidenbladh, H., Black, M.J., Sigal, L.: Implicit probabilistic models of human motion for synthesis and tracking. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) *ECCV 2002*. LNCS, vol. 2350, pp. 784–800. Springer, Heidelberg (2002)
18. Turaga, P.K., Chellappa, R., Subrahmanian, V.S., Udrea, O.: Machine recognition of human activities: A survey. *CirSysVideo* 18(11), 1473–1488 (2008)
19. Wakim, C., Capperon, S., Oksman, J.: A Markovian model of pedestrian behavior. In: *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, pp. 4028–4033 (2004)
20. Wedel, A., et al.: Duality tv-l1 flow with fundamental matrix prior. In: *Proc. IVCNZ* (2008)