Multi-Pedestrian Tracking and Map-Based Intention Estimation for Autonomous Driving Scenario

Ali Dehghani and Lucila Patino Studencki
Coburg University of Applied Sciences and Arts, Faculty of Mechanical and Automotive Engineering, Coburg, Germany

Keywords: Pedestrian Intention Estimation, Multiple Pedestrian Tracking, Situational Awareness, Autonomous Driving, Autonomous Shuttle.

Abstract: Pedestrian intentions estimation and tracking have become essential for the development of autonomous vehicles (AVs). The vehicles need to be aware of pedestrians to avoid fatalities even in complex urban traffic. This requires understanding the most probable trajectory of pedestrians to accordingly plan the vehicle's maneuvers. This complex task requires modeling how multiple pedestrians interact with each other and move depending on their environment. This paper employs a Gaussian Mixture Probability Hypothesis Density Filter, enhanced by the Generalized Potential Field Approach (GMPHD-GPFA), to simultaneously track multiple pedestrians and determine and predict their behavior seconds ahead. The model used considers the static environment of the pedestrians to estimate their intentions and improve prediction accuracy. The paper evaluates both the tracking efficiency of the algorithm and its capability to predict the intentions of multiple pedestrians.

1 INTRODUCTION

Intention estimation and tracking of pedestrians is a fundamental aspect of Vehicle Environment Perception, a crucial component in the advancing field of Intelligent Vehicle Technologies. An intelligent vehicle should safely maneuver through complex environments, including vulnerable road users (VRUs), such as pedestrians. This capability is essential for protecting VRUs and contributes to improving the overall travel experience for passengers. By accurately understanding and predicting pedestrian behaviors, intelligent vehicles can seamlessly integrate into urban traffic and adjust their navigation strategies accordingly. Motivated by experiences with the Shuttle Modellregion Oberfranken (SMO) project in Kronach, Germany (SMO, 2022), this study addresses the critical need for advanced pedestrian intention estimation in autonomous shuttle operations. As stated in (Dehghani et al., 2023), the challenges encountered, particularly those involving unforeseen pedestrian intentions that often cause shuttle abrupt halts, illuminate the necessity for precise prediction of pedestrian goals. Pedestrian movements depend on a multitude of factors, including different customs and informal regulations (social norms) related to each country that significantly impact how people behave in traffic and how they communicate their intentions (Färber, 2016). Furthermore, factors such as the street's width and the presence of traffic signals impact pedestrian behavior. In narrower or signalized areas, pedestrians may become less cautious, often crossing without checking for traffic (Rasouli et al., 2017). Considering all these factors, predicting pedestrian intentions requires an accurate model. Nonetheless, this task is complex due to the variability in the number of pedestrians and their reactions to environmental factors such as traffic density, road conditions, regulations, social influences, and other circumstances (Rasouli et al., 2017). Figure 1 illustrates an urban traffic scenario observed from the viewpoint of an autonomous vehicle, highlighting the complex interaction of different dynamic elements in challenging environmental circumstances. The scene involves multiple pedestrians, each potentially following separate routes and having different objectives, various moving vehicles, plenty of traffic signs, and adverse weather conditions, which raises critical questions about the final objectives of pedestrians. What are all pedestrians' final intentions, and which pedestrian can cause a collision?

Rudenko et al. have devised a taxonomy that organizes current solutions based on their motion modeling techniques and the degree of contextual information utilized (Rudenko et al., 2020). They divided the modeling approach for predicting pedestrian motion

Dehghani, A. and Studencki, L.
Multi-Pedestrian Tracking and Map-Based Intention Estimation for Autonomous Driving Scenario.
DOI: 10.5220/0012691700003702

In *Proceedings of the 10th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2024)*, pages 386-393
ISBN: 978-989-758-703-0; ISSN: 2184-495X

Figure 1: A complex autonomous driving scenario in Kronach, Germany under challenging conditions.

into three types: *physics-based*, *pattern-based*, and *planning-based*. In *physics-based* approaches, many motion prediction techniques model human movement using basic kinematic principles (using Newton's laws to model movement), capturing position, velocity, and acceleration for simplicity and effectiveness in stable conditions with short-term forecasts, for instance, (Elnagar, 2001) use of a Kalman Filter (KF) for tracking dynamic obstacles. However, the aforementioned work only makes predictions that are one step ahead and ignore contextual cues from the environment. In *pattern-based* approaches, utilizing the data collected from the environment or previous observed trajectories to predict motion patterns also demonstrates enhanced accuracy (Chen et al., 2016). Razali et al. (Razali et al., 2021) present a vision-based system that integrates pedestrian localization, body pose estimation, and intention prediction using a multi-task convolutional neural network, offering enhanced precision in intention prediction. However, the effectiveness of data-driven prediction methodologies largely depends on the quantity, quality, and variety of data, including various factors such as age, gender, geographical landscapes, weather conditions, lighting conditions, specific traffic scenarios, cultural norms, legal norms and social norms. Consequently, acquiring and processing such a substantial volume of labeled training data poses a challenge in real-world scenarios due to the computational intensity required (Keller and Gavrila, 2013). Moreover, they mostly do not consider the interaction of multiple pedestrians in the scenario. *Planning-based* approaches to pedestrian motion prediction try to understand the intentions behind a pedestrian's movement by following a sense-reason-predict scheme about the likely goals and possible path to reach the goal. They typically focus on using optimization techniques by applying predefined cost functions (forward planning) or learning these functions from observed behavior (inverse planning). A number of approaches model the probabilities of the future motion based on cost-to-go value estimates. They propose a probabilistic goal-directed motion model that accounts for several goals in the environment (Best and Fitch, 2015)(Vasquez, 2016). While these approaches are suitable for scenarios where understanding the underlying intent or goal is crucial, they are not as effective in dynamically changing environments where objects frequently appear and disappear or when dealing with a large number of objects. These methods can be expanded to consider different contextual cues (map-based, and dynamic environment cues) that impact pedestrian behavior. This combination facilitates the creation of more accurate and contextually sensitive forecasts by considering factors such as societal norms, traffic signals, environmental layout, and psychological conditions.

We propose a comprehensive solution for predicting pedestrian motion using a technique that inherits physics-based and planning-based characteristics that can simultaneously handle multiple pedestrians in a complex automotive driving scenario. Such an approach would leverage the accuracy of physics-based models that adhere to Newton's laws for movement and the insight of planning-based models that infer intentions and goals to forecast future paths. This hybrid method would not only model the immediate physical interactions but also incorporate a deeper understanding of pedestrian behavior, making predictions more robust in complex environments where anticipating future movements is crucial. Our algorithm enhances the Gaussian Mixture Probability Hypothesis Density (GMPHD) Filter (Clark et al., 2006) with the Generalized Potential Field Approach (GPFA) (Particke et al., 2017). This hybrid prediction approach creates a dynamic pedestrian motion model, which integrates a broader range of influences, including environmental layouts and individual pedestrian goals, into a unified framework to find the most probable objective (intention) of all pedestrians. This paper is structured as follows: In Section II the proposed method including environmental data modeling as a potential field, a dynamic model for pedestrians and Probability Hypothesis Density Filter(PHD) is presented. The experiments for demonstrating the effectiveness of our algorithm comes in Section III, and the paper's conclusion and suggestions for further research are presented in Section IV.

2 MULTI-PEDESTRIAN TRACKING

2.1 Tracking Algorithm

In general, as shown in Figure 2, tracking multiple pedestrians requires that they are first detected by some sensory input. Advanced algorithms are then applied to interpret the raw data, distinguish pedestrians from other objects, and predict the intentions. In this paper, we assume that the pedestrian tracking had already been performed and positions in a 3D coordinate system were available. We focus mainly on the Map to Potential Field, the tracking algorithm, and the Pedestrian Trajectory Prediction parts.

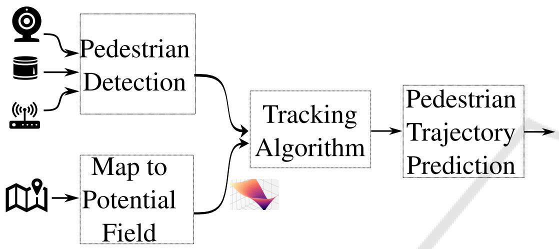

Figure 2: General Architecture of the Multi-Pedestrian Tracking System.

The concept of potential fields has been extensively applied in various research areas, including flocking behavior, trajectory planning, and pedestrian crowd analysis. However, existing methods like the social force model face limitations when dealing with individual pedestrians or small groups, as they are optimized for pedestrian crowds. Moreover, the number of parameters to be set is huge. To overcome these challenges, the Generalized Potential Field Approach (GPFA) was developed. It combines a potential field with a kinematic motion model, ensures applicability to single pedestrians and small groups, and simplifies parametrization (Particke, 2020). To calculate the potential field, every pedestrian is regarded as a test particle in several different potential fields. Each field (ϕ_N^k) stands for a different information source, such as a map of the surrounding area. Each potential field is made up of a variety of potential sources (ϕ_i^k), such as obstacles. The potential at the pedestrian's position is calculated using the following equation:

$$\phi_N^k = \sum_{i=1}^{n^k} p^k\left(d_{iN}^k\right)\phi_i^k \tag{1}$$

In this equation, $p^k(d_{iN}^k)$ represents the weight of each potential source, which depends on the Euclidean distance d_{iN}^k between the pedestrian and the potential source but is independent of time.

As demonstrated in (Particke, 2020) the influence of the potential field on the pedestrian can be modeled as an acceleration vector $\vec{a}_N^k$ of source k at position PN, which affects directly the pedestrian's dynamics.

The dynamic model of the pedestrian considers both the gradient of the potential field $\vec{\nabla}\phi_N^k$ and the flow resistance ($F_W = c_w v_N^2 \vec{e}vN$):

$$\vec{a}_N^k = \frac{-\vec{\nabla}\phi_N^k - c_w v_N^2 \vec{e}vN}{m_p} \tag{2}$$

where the pseudo mass m_p and the drag coefficient c_w parameters must be configured appropriately to represent the expected dynamics of the pedestrian.

Similar to (Particke et al., 2017), a constant velocity model as a dynamic model in the Kalman Filter for the pedestrian movement was used.

The PHD filter is a well-known method for multi-target tracking based on the ideas of random finite sets and was first introduced by Mahler and Ronald (Mahler, 2003). Later, Clark and et al. (Clark et al., 2006) proposed the Gaussian Mixture PHD (GM-PHD) filter, a computationally effective implementation of the PHD filter. The PHD filter is exceptionally well suited for handling an unknown and time-varying number of targets (Gao et al., 2021), which is a frequent challenge when attempting to follow numerous pedestrians with various intentions. Each object in a GM-PHD filter is presumed to follow a linear Gaussian model. However, the multiple target posterior distribution need not have the same covariance matrices so that it will be a Gaussian mixture (GM). Given a state $p(\mathbf{x}_{k-1})$ at time $k-1$, the probability density of a transition to the state $p(\mathbf{x}_k)$ at time k at time k is Transition Density and given by:

$$f_{k|k-1}(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \tag{3}$$

In the context of the GM-PHD filter, the Kalman Filter is utilized for the state prediction of each target, considering their acceleration. The state prediction equation in the Kalman Filter is given by:

$$\mathbf{x}_k = F_k\mathbf{x}_{k-1} + B_k\mathbf{u}_k \tag{4}$$

$\mathbf{x}_k$ is the state vector at time k, which typically includes position and velocity. F_k is the state transition matrix, mapping the previous state $\mathbf{x}_{k-1}$ to the current state $\mathbf{x}_k$. B_k is the control input model. $\mathbf{u}_k$ is the control vector, incorporating the acceleration ($\vec{a}$) of the pedestrian, obtained from the potential field.

In addition, assuming a state $\mathbf{x}_k$ at time k, the probability density of receiving the detection $\mathbf{z}_k$ gives the Likelihood Function as following:

$$g_k(\mathbf{z}_k \mid \mathbf{x}_k) \tag{5}$$

The probability density of state $\mathbf{x}_k$ given all the prior observations is represented by the notation $p_k(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ for the posterior density. Applying Bayes' recursion, we can demonstrate that the posterior density is actually as following using an initial density of $p_0(\cdot)$:

$$p_k(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{g_k(\mathbf{z}_k \mid \mathbf{x}_k)\, p_{k|k-1}(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})}{\int g_k(\mathbf{z}_k \mid \mathbf{x})\, p_{k|k-1}(\mathbf{x} \mid \mathbf{z}_{1:k-1})\, d\mathbf{x}} \quad (6)$$

In the GM-PHD filter, each target is treated as independent from the others regarding the generation of observations and its evolution. The two equations clearly demonstrate that the PHD filter effectively eliminates the combinatorial calculations resulting from the unassigned association of measurements to specific targets.

In a GMPHD-based GPFA, the tracking system initializes with predefined system parameters and an "Intention Map" that outlines the pedestrians' probable objectives. The GPFA calculates the movement acceleration of pedestrians toward each intended destination, serving as the control input for the PHD prediction process. At each time step, the system measures actual pedestrian movements, and during the association phase, the system generates multiple hypotheses based on the prior predictions and new measurements. Finally, these hypotheses undergo refinement in the update process.

Figure 3 shows the general workflow of the proposed PHD-GPFA algorithm. It is assumed, that the vehicle knows where it is and loads the topological map of its environment. Based on this, the potential field map for each intention is calculated according to (2.1) and finally the acceleration for each hypothetical intention is derived. Using this information the prediction for time point k is calculated. Subsequently the measurements are acquired and the association for each of the hypothesis based on the mahalanobis distance is performed. Later, the measurement update for each hypothesis is performed and finally, the selection of the most probable track ('Confirmed Tracks') according to the likelihood is performed.

2.2 Map to Potential Field

The primary aim of the GPFA is to improve the efficiency of models utilized in classic Kalman filtering or Monte Carlo techniques. For this purpose, an acceleration vector should be generated from information sources like attractive components (intentions) and repulsive components. In simpler terms, the pedestrians can be likened to test particles moving within a potential field, where their movements are influenced by both attractions towards certain goals and repulsions from obstacles.

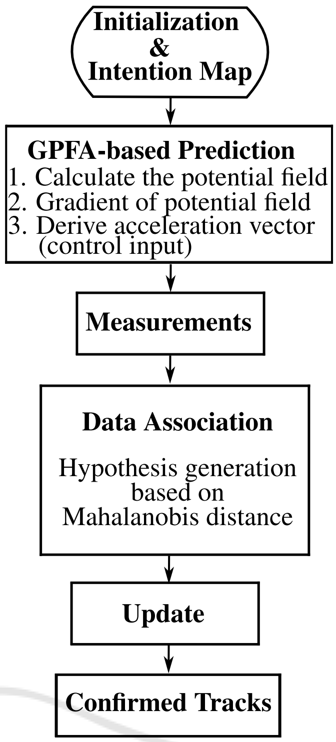

Figure 3: PHD-GPFA Flow diagram.

The explained potential field is generated using mainly the topological description of the area around the car. However some assumptions regarding social force models and pedestrian behaviour are being made:

- Pedestrians have specific intentions (destinations) and they aim to achieve them
- Pedestrians tend to go on sidewalks
- Pedestrians tend to cross road on zebras or corners
- Pedestrians tend to avoid collisions
- Pedestrians tend to keep their travel direction

Following this assumption, map elements such as zebras or sidewalks are described as attraction areas. Map elements like buildings are described as repulsion areas. Moreover, accounting for the direction of walk is necessary to define plausible destinations.

The potential field generation is a key function of the component termed map to potential field. For illustration, Figure 4 demonstrates the concept within a street scenario. Additionally, an example of a potential field for a driving scenario is depicted in Figure 5."

Using the defined potential field, the tracking and posterior prediction of pedestrian trajectories are performed.

How can we improve pedestrian trajectory prediction using map data? To answer this question, we

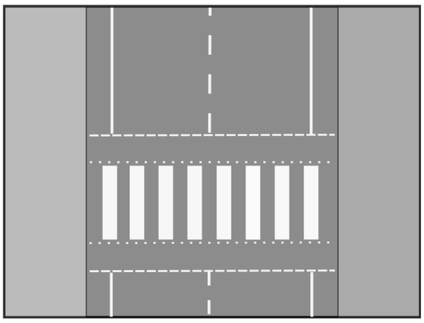

Figure 4: Street scenario.

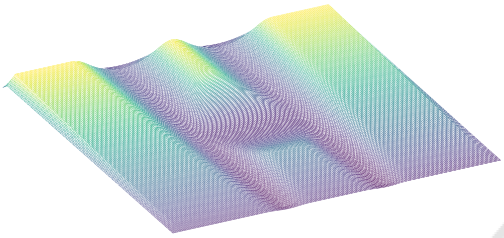

Figure 5: Potential field example for a street scenario.

need to model the map as a potential field by utilizing the map information. The potential field is computed by evaluating the influence of conic components across a grid of points that extends the domain of interest. While the influence of repulsive components typically exhibits an inversely exponential relationship with distance, modulated by a scaling factor, the influence of conic components in the case of "intention," in this instance, exhibits a linear relationship with distance, also modulated by a scaling factor (Particke et al., 2017). This influence decreases linearly with increasing distance. Figure 6 depicts the resulting field in both 3D surface and 2D heatmap formats. These visualizations depict the strength of the field at a given point as the Z-value in the 3D surface plot and the color intensity in the 2D heatmap, respectively.

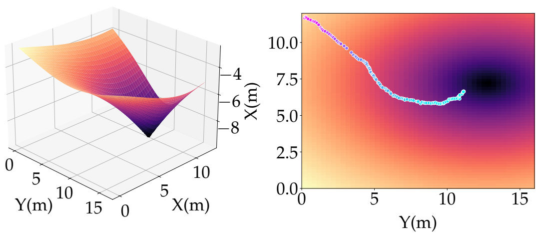

Figure 6: Pedestrian in potential field map.

3 EXPERIMENTAL RESULTS

In this study, we utilized a publicly available pedestrian trajectory dataset from the University of Edinburgh's School of Informatics (Majecka, 2009), chosen for its overhead camera system that captures clear, minimally noisy pedestrian paths in a public space. The dataset's precise, real-world coordinate trajectories and the variety of pedestrian movements provide an ideal base for adding controlled noise (specifically additive white Gaussian noise) for analysis. Its diverse path patterns, originating from different points but diverging towards multiple destinations, accurately represent the dynamic nature of pedestrian traffic in real environments.

For the experiment, trajectories of three pedestrians going to two target regions (two intentions) were selected. The extracted data were used as ground truth, and some artificial measurement noise ($\sigma = 0.1$) was added to assess the performance of our approach.

The approach was first evaluated using the best-case scenario, where one pedestrian with a known intention is tracked. Subsequently, the algorithm's capability to identify the real intention of a pedestrian was evaluated by tracking one pedestrian with three possible intentions (defined as hypotheses for the PHD-GPFA). Finally, the algorithm was tested on scenarios involving multiple pedestrians each with multiple unknown intentions.

3.1 One Pedestrian with One Known Intention

The first experiment conducted involves tracking a single pedestrian, with the pedestrian's intention assumed to be known at the position $x = 6.61m$ and $y = 11.11m$. The acceleration vector calculated by the GPFA is then used as a control input for the Kalman filter's state prediction stage. The KF and KF-GPFA methods were used to estimate the next state based on the previous measurement and the motion model. In addition, predictions for future time points ranging from 1 to 10 seconds were made using both methods.

In addition, predictions for future time points, ranging from 1 to 10 seconds, were made using both methods. These predictions represent the time an autonomous vehicle would need to react in case of a collision and the position uncertainty related to this time-frame. The results of the estimation and prediction processes are presented in Figure 7.

On closer inspection of the data, it is evident that the Kalman Filter provides a reasonably accurate estimate of the pedestrian's position, with only minor deviations compared to the ground truth data. Si-

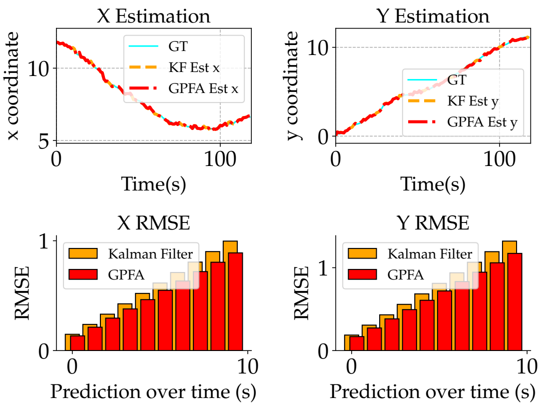

Figure 7: Trajectory estimation and prediction using KF (state-of-the-art) and GPFA compared to Ground Truth (GT).

multaneously, the KF-GPFA technique demonstrates remarkable performance. This demonstrates the potential benefits of using an enhanced method such as KF-GPFA for more precise pedestrian position estimation. On the lower part of Figure 7, the prediction capabilities of a vehicle, based on the estimated trajectory and intention for different time intervals, are shown. As expected, increasing the time interval for making a prediction, the error increases. In terms of forecasting pedestrian positions, the KF-GPFA displays a noticeable enhancement over the standard Kalman Filter.

3.2 One Pedestrian with Three Possible Intentions

In the prior experiment, one of the limitations was the assumption that the pedestrian's intention or goal was known, whereas, in autonomous scenarios, this information is typically unavailable. Motivated by this discrepancy, the current experiment focuses on integrating an unknown intention into pedestrian tracking and prediction, aiming to address the question of how pedestrians' intentions can be estimated in advance.

As mentioned before, the proposed solution involves estimating the probability of each possible intention. This is achieved by modeling the intentions as conic well components in the potential field, which is derived from the map of the immediate environment. These components are then evaluated at each estimated pedestrian position. Figure 8 illustrates a potential field map, highlighting three plausible intentions (hypotheses) for a pedestrian.

The estimation result for the most probable hypothesis is presented in the upper subplot of Figure 9. The estimated positions closely align with the actual

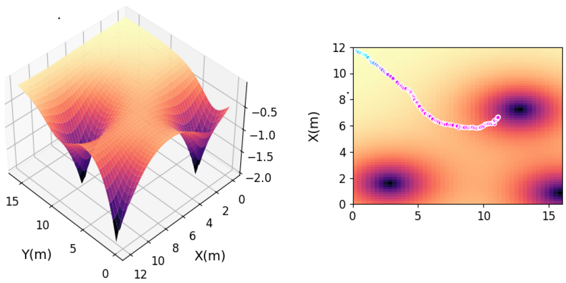

Figure 8: Potential field map for the unknown intention, where the three wells represent the intentions.

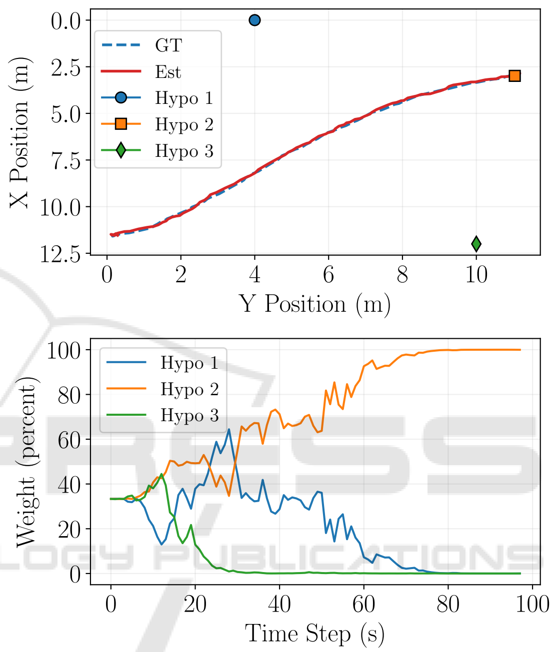

Figure 9: Tracking and probabilistic intention estimation of a pedestrian over time, compared to ground truth (GT), with an analysis of hypothesis (Hypo) weights.

positions, indicating accurate estimation. The pedestrian's intentions are represented by polygons of various hues, denoting various hypotheses. The second subplot depicts the likelihood; the larger the weight, the more probable the hypothesis. This graph illustrates the efficacy of the GPFA algorithm in finding the most probable intention, estimating positions, and tracking the evolution of hypothesis probabilities.

3.3 Multiple Pedestrian with Multiple Possible Intentions

While the previously mentioned experiment, focused on a single pedestrian with multiple intentions, real automotive driving scenarios necessitate tracking multiple intentions of several pedestrians simultaneously. In this experiment the goal is to validate whether our algorithm could effectively handle multiple pedestrians. For this, we examined three pedestrians, each potentially having one of two distinct intentions, labeled as intention one and intention two. In order to evaluate the algorithm's efficiency, we now consider three pedestrians walking in the same area, with two possible intention hypotheses given for each pedestrian. The results of the tracking are presented in Figure 10 as a 2D plot and in Figure 11 as root mean square error (RMSE) for each of the six hypotheses.

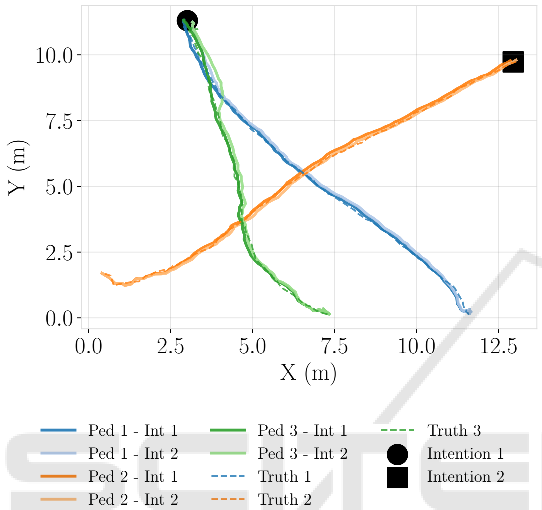

Figure 10: Track estimation for multi-pedestrian scenarios utilizing GMPHD-GPFA; involving three pedestrians each with two potential intentions, yielding a total of six hypothesis scenarios.

Our observations reveal that the hypotheses corresponding to the correct intentions—*Pedestrian 1* with *Intention 1*, *Pedestrian 2* with *Intention 2*, and *Pedestrian 3* with *Intention 1*—demonstrate lower RMSE errors, aligning with our expectations. This suggests that the algorithm can effectively differentiate between the correct and incorrect intentions for each trajectory.

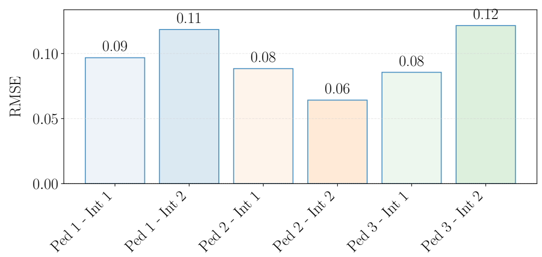

Figure 11: RMSE of multiple trajectories.

Analyzing the weight of different intentions also illustrates the finding of the correct intention according to Figure 12. After some time steps (between 40-60) the real intention can be clearly distinguished, corresponding to around 4 til 6 seconds. Although it is a significant improvement with respect to the state-of-the-art algorithms, this is long for a successful awareness of an autonomous vehicle. This is due to the fact that the algorithm makes its inferences mainly based on position data, and only if the trajectories take notable different ways can the intention be clearly determined.

Analyzing the weight of different intentions also illustrates the finding of the correct intention according to Figure 12. After some time steps (between 40–60) the real intention can be clearly distinguished, corresponding to around 4 to 6 seconds. Although it is a significant improvement with respect to the state-of-the-art algorithms, this duration is long for a successful awareness of an autonomous vehicle. This is because the algorithm makes its inferences mainly based on position data, and only if the trajectories take notably different paths can the intention be clearly determined.

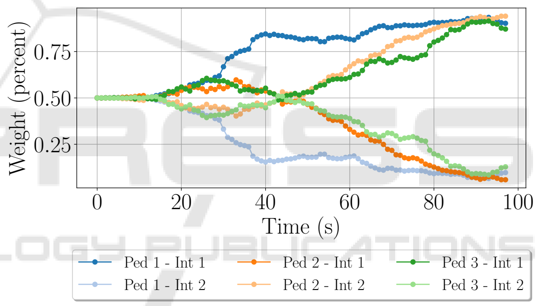

Figure 12: Assessing Diverse Hypothesis Weights for Pedestrian Intention Detection.

The Table 1 shows the RMSE results for both estimations and predictions:

Table 1: RMSE Comparison of the prediction capabilities in each experiment.

Trajectory	Estimation		Prediction		
	KF	GMPHD-GPFA	1s	2s	3s
Pedestrian 1	0.11	0.09	0.42	1.26	2.43
Pedestrian 2	0.07	0.06	0.49	1.42	2.50
Pedestrian 3	0.10	0.08	0.32	0.98	1.91

The first column corresponds to state of the art algorithm based on KF and PHD, the second column coincides with the results in Figure 11, for the prediction of the RMSE observed at a 1–3 second interval. While this may seem high compared to results from individual pedestrians, it is crucial to note the differences between the experiments. The intentions in the experiment involving multiple pedestrians are unknown, unlike in the single-pedestrian scenario. Additionally, a higher number of hypotheses influences the estimation quality, adding complexity that the algorithm must manage. Although the algorithm successfully tracks objects and infers intentions, it does not directly consider changes in those intentions.

Despite the promising results, the prediction capability falls short of meeting the timing and accuracy requirements for autonomous vehicles operating in urban environments. There is a need for further development to enhance the model with faster intention detection techniques. Such improvements could involve using gestures or other indicators, extending beyond reliance solely on trajectory data.

4 CONCLUSION

This paper introduces an approach that combines physical-based and planning-based modeling for tracking and predicting the positions and intentions of multiple pedestrians around an autonomous vehicle. Utilizing a Probability Hypothesis Density Filter (PHD) integrated with a Generalized Potential Field Approach (GPFA), the proposed algorithm generates multiple hypotheses and continuously tracks them, effectively identifying pedestrians' actual intentions. This enables autonomous vehicles to accurately forecast pedestrian movements and re-planing maneuvers accordingly. However, accelerating the detection of intentions remains a challenge that requires further development. The study also highlights the critical role of incorporating map information in defining tracking hypotheses, significantly enhancing the model's precision and reliability.

ACKNOWLEDGEMENTS

The SMO project is supported by the Federal Ministry of Transport and Digital Infrastructure of Germany. For more information about the project, please see: www.shuttle-modellregion-oberfranken.de

REFERENCES

Best, G. and Fitch, R. (2015). Bayesian intention inference for trajectory prediction with an unknown goal destination. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5817–5823. IEEE.

Chen, Y., Liu, M., Liu, S.-Y., Miller, J., and How, J. P. (2016). Predictive modeling of pedestrian motion patterns with bayesian nonparametrics. In *AIAA guidance, navigation, and control conference*, page 1861.

Clark, D. E., Panta, K., and Vo, B.-N. (2006). The gm-phd filter multiple target tracker. In *2006 9th International Conference on Information Fusion*, pages 1–8. IEEE.

Dehghani, A., Salar, H., Srinivasan, S., Zhou, L., Arbeiter, G., Lindner, A., and Patino-Studencki, L. (2023). Enhancing availability of autonomous shuttle services: A conceptual approach towards challenges and opportunities. Manuscript under review.

Elnagar, A. (2001). Prediction of moving objects in dynamic environments using kalman filters. In *Proceedings 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation (Cat. No. 01EX515)*, pages 414–419. IEEE.

Färber, B. (2016). Communication and communication problems between autonomous vehicles and human drivers. *Autonomous driving: Technical, legal and social aspects*, pages 125–144.

Gao, Y., Jiang, D., Zhang, C., and Guo, S. (2021). A labeled gm-phd filter for explicitly tracking multiple targets. *Sensors*, 21(11):3932.

Keller, C. G. and Gavrila, D. M. (2013). Will the pedestrian cross? a study on pedestrian path prediction. *IEEE Transactions on Intelligent Transportation Systems*, 15(2):494–506.

Mahler, R. P. (2003). Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic systems*, 39(4):1152–1178.

Majecka, B. (2009). Statistical models of pedestrian behaviour in the forum. *Master's thesis, School of Informatics, University of Edinburgh*.

Particke, F. (2020). *Predictive Pedestrian Awareness with Intention Uncertainties for Autonomous Driving*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

Particke, F., Patino-Studencki, L., Thielecke, J., and Feist, C. (2017). Pedestrian tracking using a generalized potential field approach. In *VISIGRAPP (6: VISAPP)*, pages 509–514.

Rasouli, A., Kotseruba, I., and Tsotsos, J. K. (2017). Understanding pedestrian behavior in complex traffic scenes. *IEEE Transactions on Intelligent Vehicles*, 3(1):61–70.

Razali, H., Mordan, T., and Alahi, A. (2021). Pedestrian intention prediction: A convolutional bottom-up multi-task approach. *Transportation research part C: emerging technologies*, 130:103259.

Rudenko, A., Palmieri, L., Herman, M., Kitani, K. M., Gavrila, D. M., and Arras, K. O. (2020). Human motion trajectory prediction: A survey. *The International Journal of Robotics Research*, 39(8):895–935.

SMO (2022). Shuttle modellregion oberfranken (smo) project. https://www.shuttle-modellregion-oberfranken.de/.

Vasquez, D. (2016). Novel planning-based algorithms for human motion prediction. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3317–3322. IEEE.