Directional Primitives for Uncertainty-Aware Motion Estimation in Urban Environments

Ransalu Senanayake, Maneekwan Toyungyernsub, Mingyu Wang, Mykel J. Kochenderfer, and Mac Schwager

Abstract— We can use driving data collected over a long period of time to extract rich information about how vehicles behave in different areas of the roads. In this paper, we introduce the concept of *directional primitives*, which is a representation of prior information of road networks. Specifically, we represent the uncertainty of directions using a mixture of von Mises distributions and associated speeds using gamma distributions. These location-dependent primitives can be combined with motion information of surrounding vehicles to predict their future behavior in the form of probability distributions. Experiments conducted on highways, intersections, and roundabouts in the Carla simulator, as well as real-world urban driving datasets, indicate that primitives lead to better uncertainty-aware motion estimation.

I. INTRODUCTION

Autonomous vehicles will not only have to interact with themselves but also will have to coexist with other vehicles operated by humans. Hence, it is important for autonomous cars to learn the behavior of both human and robotic agents to safely maneuver in busy urban environments. Similar to a human driver, we expect autonomous cars to learn from experience. For instance, when we, as human drivers, approach an intersection, it is possible for us to anticipate how other vehicles in the surrounding would plausibly act. This prediction depends not only on the current behavior but also heavily on our previous experience of observing vehicle interactions in intersections.

Modern vehicles can be used to collect large amounts of data to gain insight into how drivers typically behave in various parts of the road networks. Going beyond the capacities of human drivers, driverless cars can build lifelong models and can learn from one another. The advancement of vehicle-to-network (V2N) and vehicle-to-everything (V2X) technologies would support the development of such communication needs. The behavior of vehicles, both autonomous and human-driven, is highly variable and can be difficult to predict. Therefore, in order to build robust models about the road networks and driver behavior, it is important to consider multimodal probabilistic formulations. For instance, as shown in Figure 1, when a car approaches a four-way intersection, there are three possible directions it could turn. It is important that our models capture the uncertainty in the future directions. If an autonomous car observes a vehicle ahead, it should be able to infer which directions the vehicle

R. Senanayake, M. Toyungyernsub, and M.J. Kochenderfer are with the Stanford Intelligent Systems Laboratory (SISL); and M. Wang and M. Schwager are with the Stanford Multi-Robot Systems Laboratory (MSL) in the Department of Aeronautics and Astronautics at Stanford University, Stanford, CA 94305, USA (email: {ransalu, maneekwt, mingyuw, mykel,schwager}@stanford.edu}).

could turn at each position of its possible future trajectories. Such an understanding about the environment can be used to enhance safe decision-making.

In this paper, we propose *directional primitives*, a representation of prior multimodal directional uncertainties that are local to different segments of the road networks. Unlike previous approaches, our focus is to:

1) model the *uncertainty* in speed as well as directions,
2) associate prior knowledge with current knowledge to make accurate predictions, and
3) use prior information to generate possible trajectories.
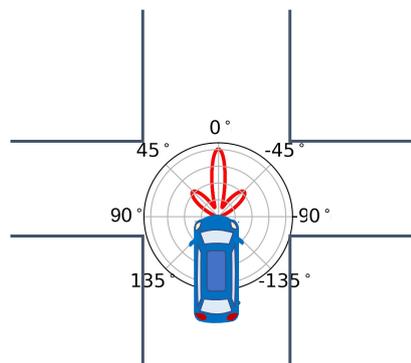
(a) A tri-modal directional distribution obtained from prior vehicle trajectories.
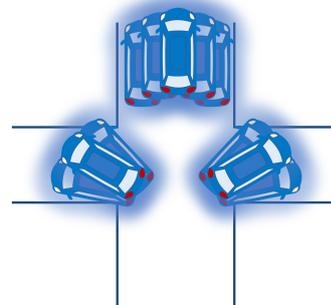
(b) Hallucinated vehicle positions in the next time step.

Fig. 1: A vehicle approaches an intersection. For demonstration purposes, we assume one-way roads. The vertical road is the main road and the horizontal roads are collector lanes. Based on prior knowledge, indicated by the polar plot in (a), the vehicle is more likely to continue straight than turn. Some predicted locations are shown in (b).

(a) Linear projection   (b) Polar projection
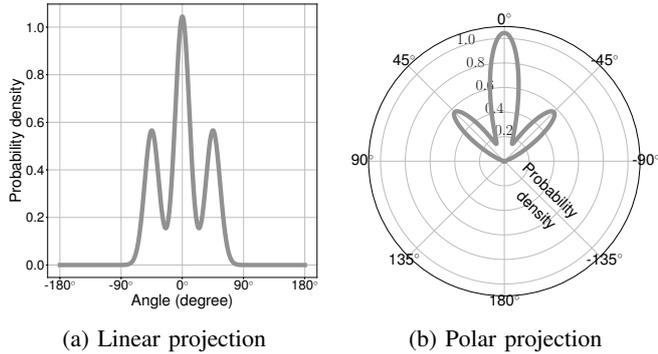
Fig. 2: Two different representations (projections) for a mixture of von Mises distributions with parameters $M = 3, \boldsymbol{\mu} = [-45°, 0°, 45°], \boldsymbol{\kappa} = [20, 20, 20],$ and $\mathbf{w} = [0.25, 0.5, 0.25]$.

II. BACKGROUND

Various methods have been used to model the motion of vehicles and pedestrians [1], predicting occupancy levels [2], and predicting trajectories [1]. Much of this prior research framed these problems as a supervised learning problem and typically do not take into account prior knowledge about location-specific agent behavior for estimation.

In contrast with the objective of motion prediction research, our focus is incorporating high-level information about the behavior of observed vehicles into decision-making. For instance, if a vehicle in front of us is signaling to turn right in the intersection shown in Figure 1, we can specify this current notion of signaling as a directional probability distribution. This distribution can then be used in conjunction with the prior distributions that specify the probabilities of directions and speeds of vehicles in the past in order to infer the possible future locations of the vehicle.

Estimating the current direction of a moving object has been studied by Senanayake and Ramos [3]. This work has been extended to the temporal domain using LSTMs [4]. While our proposed model can plausibly be extended to the temporal domain [5], [6], this paper focuses only on the spatial domain. None of these previous methods consider the variation of speed, which is also crucial for accurate motion prediction.

III. DIRECTIONAL PRIMITIVES

A. Location-dependent direction-speed priors

We model the uncertainty of directions at every location in a road network. The uncertainty of speeds associated with each mode of direction is also modeled.

1) Uncertainty of directions: In order to effectively model uncertainty in real-world robotic applications such as autonomous driving, the stochasticity of directions must be modeled. Unlike many other physical quantities, there are several challenges when modeling directional variability using a probability density function. A density function that represents directions must have a limited support $\theta \in [0, 2\pi)$ with the two limits, 0 and 2π, being the same. We consider the von Mises distribution [7] used for directional grid maps

(DGM) [3]. Intuitively, a von Mises distribution can be thought of as a normal distribution wrapped around a circle.

In order to understand in which directions vehicles typically move in different segments of the road network, we begin by discretizing the environment into a grid with P cells. A mixture of von Mises distributions is assigned to each cell. The mixture is to handle the multiple possible directions [3], [8]. This mixture,

$$p(\theta \mid \mathbf{x}) = \sum_{m=1}^{M} w_m \mathcal{VM}(\theta; \mu_m, \kappa_m), \quad (1)$$

models the probability of possible directions $\theta \in [0, 2\pi)$ at anywhere in the longitude-latitude space $\mathbf{x} \in \mathbb{R}^2$. With weights $\sum_{m=1}^{M} w_m = 1$, the distribution is composed of M von Mises distributions,

$$\mathcal{VM}(\theta; \mu_m, \kappa_m) = \frac{1}{2\pi J_0(\kappa_m)} \exp\left(\kappa_m \cos(\theta - \mu_m)\right), \quad (2)$$

where μ_m is the mean direction and κ_m is the concentration.[1] The higher the concentration parameter, the lower the dispersion of angles is. $J_r(\kappa)$ is the rth order and 1st kind modified Bessel function. The mean and variance of the mth von Mises distribution are μ_m and $\left(1 - J_1(\kappa_m)/J_0(\kappa_m)\right)$, respectively. Figure 2 shows a mixture of von Mises distributions with three components, also known as modes.

For the entire road network with P cells and M_p mixture components for the pth cell, the set of parameters is $\{\{(\mu_{pm}, \kappa_{pm}, w_{pm})\}_{m=1}^{M_p}\}_{p=1}^{P}$. With angles extracted from vehicle trajectory data, these parameters can be learned using P individual expectation-maximization routines [3]. M_p is determined for each cell using a *density-based clustering* [9].

2) Uncertainty of speeds: In Section III-A.1, we obtained a multimodal distribution of directions. In addition to the directions, we also want to model the distribution of speed for each directional modality. Having such speed priors is important because the speeds are different in various segments of the road network. For instance, the speed of vehicles is relatively low near intersections, crosswalks, and roundabouts compared to a highway. Since speed is strictly a non-negative quantity, it is modeled using a gamma distribution given by,

$$p(s \mid \mathbf{x}) = \frac{\beta^\alpha}{\Gamma(\alpha)} s^{\alpha-1} \exp\left(-\beta s\right), \quad (3)$$

where α and β are the shape and rate parameters, respectively. Γ is the gamma function. The mean and variance of a gamma distribution can be computed as $\alpha\beta^{-1}$ and $\alpha\beta^{-2}$, respectively. Gamma distributions are estimated by maximizing the likelihood with data that lie within twice the standard deviation of each von Mises component.

B. Hallucinating future locations

If an autonomous car observes another vehicle at a particular location, then the corresponding cell the vehicle belongs to can be used to predict where the observed vehicle could

---

[1]Analogous to a Gaussian distribution, these can be thought as the mean and inverse variance parameters.

(a) Sampling the directional distribution   (b) Combining the samples with speeds   (c) Projected positions
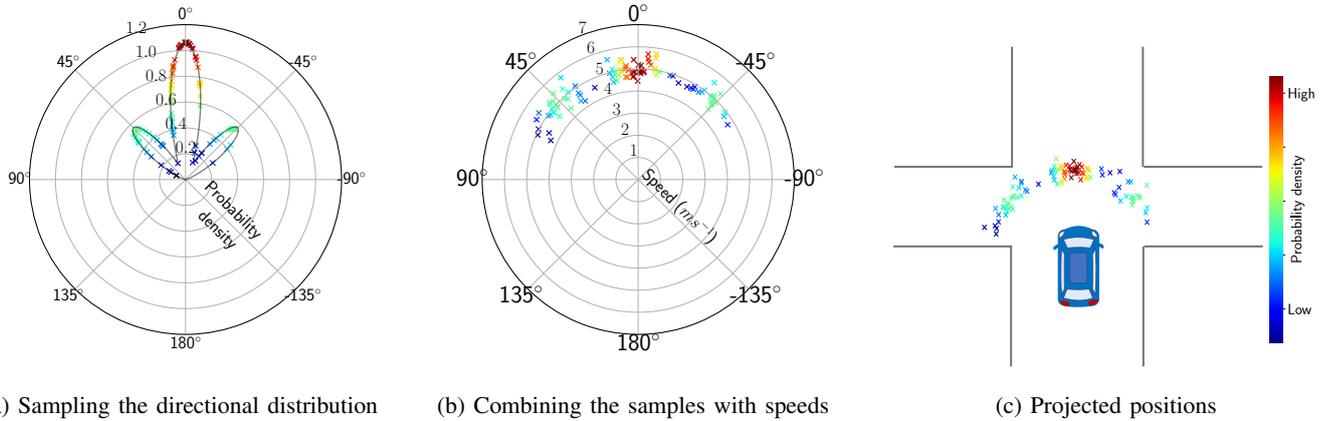
Fig. 3: (a) A mixture of von Mises distribution and 100 samples drawn from it. (b) With different speeds in different directions, samples are projected into the future. (c) Projected locations of the vehicle.

move. As depicted in Figure 1, finitely many such hallucinations can be obtained by sampling the possible directions from the mixture of von Mises distribution and projecting the vehicle towards the sampled directions of motion.

To sample from the mixture, we first sample $\tilde{m}_p \in \{1, 2, \ldots, M_p\}$ from a categorical distribution $\mathcal{CAT}([w_{p1}, w_{p2}, \ldots, w_{pM_p}])$. Then, we sample an angle $\tilde{\theta}_p$ from the $\tilde{m}_p$th von Mises mixture component. This process can be repeated to obtain N samples. If a sample has a speed s, in a unit time, the vehicle will be projected into the positions $\{[s\cos(\tilde{\theta}_p^{(n)}), s\sin(\tilde{\theta}_p^{(n)})]\}_{n=1}^{N}$ in the vehicle's coordinate system. If the distribution of speeds corresponding to the particular location is also known, it can also be incorporated into computations by replacing s with $s(\tilde{\theta}_p^{(n)})$ and sampling from the gamma distribution introduced in Section III-A.2. As illustrated in Figure 3, uncertainty in directions and speeds would project the vehicle not only into various angles, but also into different positions.

C. Incorporating current knowledge

Note that the estimation in Section III-B is purely based on previous experiences of observing the behavior of thousands of vehicles in the past. Nonetheless, past information alone is not sufficient for accurate decision-making. Within our framework, we can also effectively use high-level information about the environment to bias the predictions so as to make informed decisions.

If we have some indications of how an observed car would behave, then we incorporate this current high-level information with the prior information to make better predictions. For instance, consider the distributions in Figure 4. According to the prior information $p_0(\theta)$, shown in blue, there is a very high chance that the vehicle would continue straight, and there is less chance that it would turn either left or right. However, if we observe the vehicle leans towards the right lane or it signals the right blinkers, we have a current probability distribution $p_t(\theta)$ with the mean direction towards the right and very low concentration indicating high
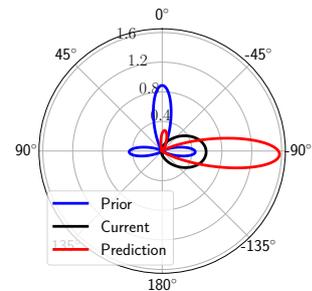

Fig. 4: The prior distribution $p_0(\theta)$ obtained from past vehicle trajectories is shown in blue. The distribution in black $p_t(\theta) = \mathcal{VM}(-90°, 2.5)$ indicates the belief about a vehicle. For instance, if the vehicle signals to turn right, we can set the mean parameter to $-90°$ and concentration to a very low value (i.e. high uncertainty). Combining the prior and current information, we can obtain a new predictive distribution shown in red.

uncertainty. Based on the prior and current information, we can compute the new directional distribution as,

$$p_*(\theta) \propto p_t(\theta)p_0(\theta). \tag{4}$$

Because of current information, the new distribution $p_*(\theta)$, shown in red in Figure 4, has a higher probability towards right and negligible probability towards straight and almost zero probability towards left.

In order to obtain samples from the "product of mixtures" $p_*(\theta)$, a Monte Carlo technique can be used. In this paper, as illustrated in Figure 5, we propose a *rejection sampling* scheme. Firstly, we select an arbitrarily large constant k and a unimodal von Mises distribution with a small concentration as the *proposal distribution* $q(\theta)$. This distribution should be broad enough to enclose the underlying true distribution. We then draw a sample $\tilde{\theta}$ from $q(\theta)$ and evaluate $k \times q(\tilde{\theta})$ and $p_*(\tilde{\theta})$. Independent to this process, another sample $\tilde{u}$ is drawn from the uniform distribution $\mathcal{U}[0, k*q(\tilde{\theta})]$. Samples

$\tilde{\theta}$ are accepted as samples from $p_*(\theta)$, if $\tilde{u} \le p_*(\tilde{\theta})$. The entire process is repeated until it satisfactorily converges. All accepted samples are kept as an approximation to $p_*(\theta)$. In order to further improve the sampling efficiency, similar to sampling from a product of mixture of Gaussians, Gibbs sampling can be performed using a KD-tree.

D. Generating multi-modal trajectories

If a vehicle is observed in cell $\mathcal{C}_p$ at time t, it is possible to sample an angle $\tilde{\theta}_p^{[t]}$ and move in that direction for one time step. Then we can sample from the directional (and speed) distribution in the new cell. We can repeat this for T time steps. This process is summarized in Algorithm 1.

Algorithm 1 Trajectory generation

1: **input** Initial vehicle position $\mathbf{x}_0$
2: **for** $k \in \{0, 1, \dots, K\}$ trajectories
3:      **for** $t \in \{0, 1, \dots, T_k\}$ time steps
4:          $\mathcal{C}_p \leftarrow$ Determine the cell corresponding to $\mathbf{x}_t^{[k]}$
5:          $\tilde{\theta}_p^{[t,k]} \leftarrow$ Sample from the distribution over $\mathcal{C}_p$
6:          $\mathbf{x}_{t+1}^{[k]} \leftarrow$ Move in the direction $\tilde{\theta}_p^{[t,k]}$
7: **return** Trajectories $\{[\mathbf{x}_0^{[k]}, \mathbf{x}_1^{[k]}, \dots, \mathbf{x}_{T_k}^{[k]}]\}_{k=1}^K$

IV. EXPERIMENTS

The Carla simulator was used to collect simulation data. Town 3 of the simulator was used with 20 cars running in the autopilot mode for around 2 hours. To increase diversity, the system was restarted by spawning new cars every 10000 timesteps. Spurious trajectories such as rare collisions that are generated due to the limitations of the driver model of the simulator were removed based on visual inspection. To demonstrate the idea of directional primitives, we considered two important areas of the environment: a roundabout and an intersection (Figure 6). Stanford Drone dataset and the Lankershim segment of the NGSIM dataset were used as real-world datasets. The former is an aerial dataset that contains trajectories of pedestrians, cyclists, etc. whereas the latter contains trajectories of highways in the US.
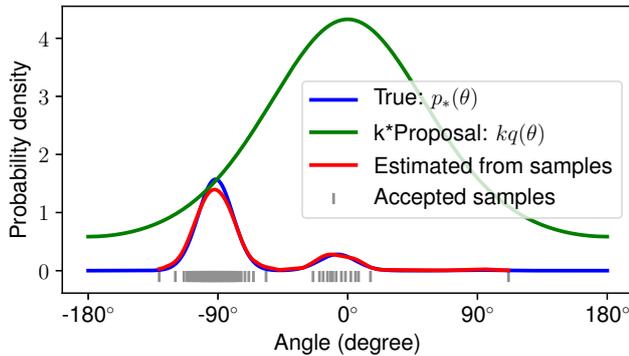

Fig. 5: Rejection sampling for a mixture of von Mises distribution. Kernel density estimate on the accepted samples was used to plot the red curve.

Roads in each environment were divided into cells depending on the road geometry. Von Mises and gamma distributions were then fitted to each cell as described in Sections III-A.1 and III-A.2, respectively. As shown in Figure 8, this results in a collection of von Misses distributions spread throughout the environment. For each directional mode in each cell, we also have an associated speed distribution. An example of a pair of speed distributions corresponding to a given cell with a bimodal von Mises directional distribution is shown in Figure 7.

On average, it takes 82 milliseconds for the EM algorithm to learn the parameters of each mixture of von Mises distributions. Using 10% of the dataset as the test set, to assess how well the distribution is fitted, the average probability density was computed (Table I). Since the test dataset was prepared randomly, it is expected that a higher number of data points are under the higher density area. Hence, the average probability density should be higher than a benchmark uniform distribution. The lower the concentrations of the fitted distribution are, the higher the average density values would be. However, since this is a density estimation problem, designing a metric for evaluating the dispersion is challenging. We compared our results with DGM, Gaussian process directional estimation (GP) [10], and an uninformative von Mises distribution ($\kappa \simeq 0$). We considered the sparse subset of data approximation to GPs [11] as the datasets we consider are extremely large to fit a full GP in a reasonable time. The last column in Table I also reports average density for the gamma speed distributions.

To assess how much accuracy is gained by combining prior and current belief distributions, with reference to the example given in Figure 4, the RMSE value between the true ($-90°$) and 100 angles sampled from each distribution were computed. The RMSE values of angles dropped from approximately $60°$ to $40°$ after combining prior and current estimates. As reported in Table II, we further evaluated the effectiveness of combining prior information with observations on the Carla roundabout and intersection datasets using the improvement in likelihood as a metric. Firstly, we randomly picked several locations in the environment. We then computed the prior likelihood $\mathcal{L}_0$ on the test set. As a proxy, we considered the direction of observation as the most probable direction indicated by the directional distribution in the cell right before the prior cell. Its likelihood $\mathcal{L}_t$ is also calculated. We then computed the prior-observation

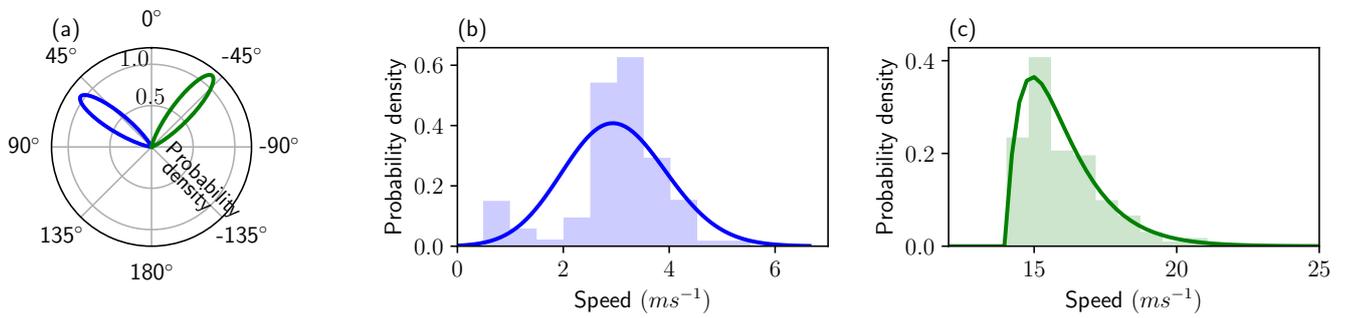Fig. 6: Carla roundabout and intersection.

Fig. 7: (a) A bimodal directional distribution. Blue and green indicates the two modes. (b) and (c) show the speed distributions, gamma, corresponding to the blue and red lobes of the directional distribution.

(a) Carla Roundabout
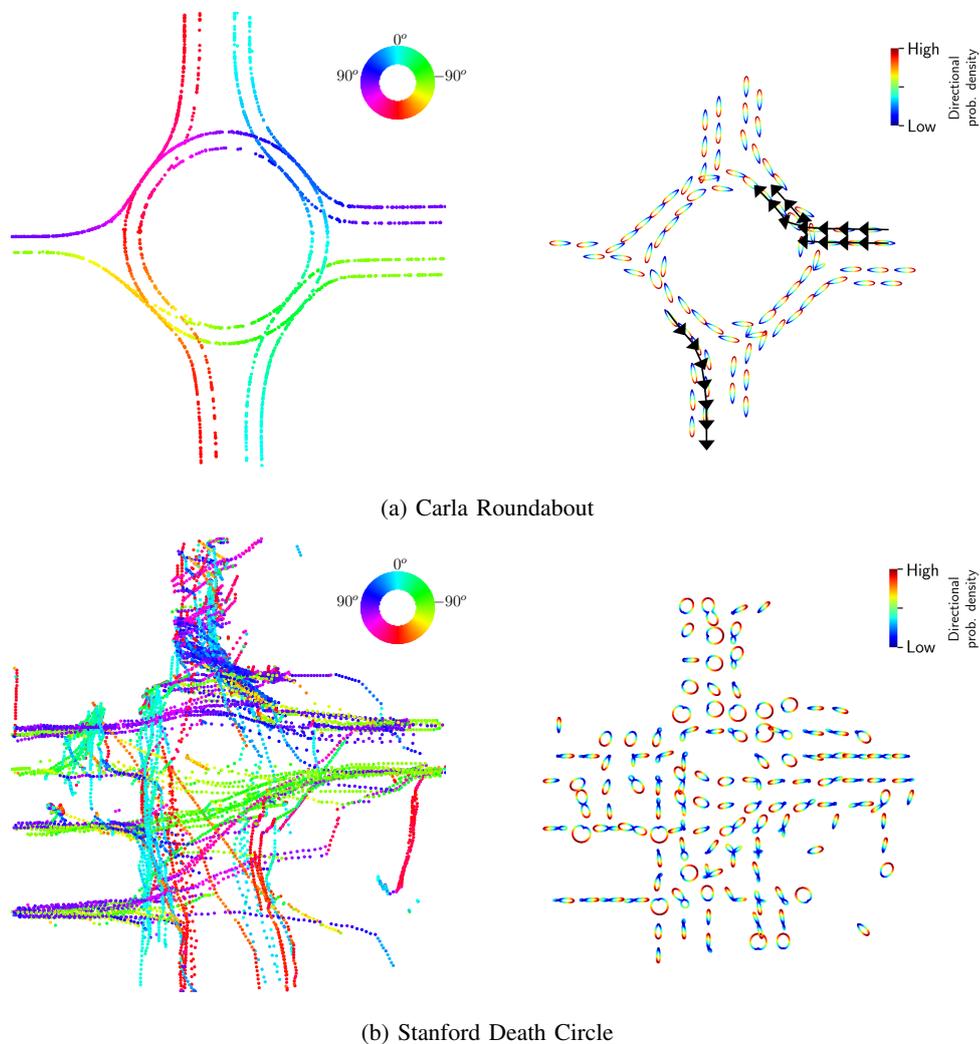

(b) Stanford Death Circle

Fig. 8: Directional polar plots for various locations in the (a) Carla roundabout and (b) Stanford Death Circle. The dataset in (a) only contains trajectories of cars while dataset (b) contains trajectories of pedestrians, bicyclists, skateboarders, buses, etc. This plot is obtained by overlaying many individual multimodal directional distributions. Each distribution takes the form of Figure 1a. Blue to red indicates lower to higher directional density. Therefore, directions with red are the most probable directions of motion. Black arrows indicate some generated paths starting at three different locations. It is possible to generate finitely many trajectories that originate from anywhere in the environment.

TABLE I: Average probability density for different models. The higher, the better.

Dataset	Direction				Speed
	Primitive	DGM	GP	Uninformative	Primitive
Carla roundabout	1.891 ± 0.471	1.480 ± 0.727	1.484 ± 1.029	0.159 ± 0.000	0.304 ± 0.126
Carla intersection	1.893 ± 0.468	1.588 ± 0.758	0.657 ± 0.505	0.159 ± 0.000	0.269 ± 0.110
NGSIM Lankershim	1.726 ± 0.552	1.381 ± 0.742	0.114 ± 0.051	0.159 ± 0.000	0.324 ± 0.273
Stanford Death Circle	0.453 ± 0.409	0.387 ± 0.352	0.119 ± 0.044	0.159 ± 0.000	2.344 ± 5.615

TABLE II: Percentage average improvement of likelihood after combining observations with the prior.

Quantity	Roundabout	Intersection
Prior likelihood	267.072	1288.951
Observation likelihood	187.271	1051.696
Posterior likelihood	280.905	1619.642
Percentage improvement	54.119%	48.993%

combined distribution using eq. (4) and the posterior likelihood $\mathcal{L}_*$ on the test dataset. The percentage improvement in likelihood is $(\mathcal{L}_* - \mathcal{L}_t)/\mathcal{L}_t \times 100\%$. We also simulated future trajectories of vehicles. With the fitted distributions, in Figure 8a, three possible trajectories starting at three different locations around the roundabout are shown in black arrows.

In Section III, we divided the environment into cells. In future work, as in the context of decentralized collision avoidance [12], it is possible to use Voronoi cells. In 2D space, a Voronoi diagram is a partition of the space based on the positions of a group of N generators $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. The Voronoi cell associated with the ith generator consists of all the points that are closer to $\mathbf{x}_i$ than any other generators. If Voronoi cell i and j are adjacent, generator i and j are called Voronoi neighbors. For the directional map, we can jointly learn the cell boundaries and directional parameters by solving a variant of $\min \sum_{c \in \mathcal{C}} \sum_{\mathbf{x} \in \mathcal{A}} - \log p(\mathbf{x}_c \mid \theta_c)$ for $\mathcal{C}$ cells and $\mathcal{A}$ all possible assignments of data points to cells.

V. CONCLUSION

We introduced directional primitives as a framework to represent uncertainty in directions and speeds of a road network. We showed that this prior information can be combined with current information about the vehicles observed in the environment to infer the possible directions a vehicle could turn. Future work will focus on optimizing road tessellation, incorporating spatiotemporal state dependencies, and employing computer vision models for estimating the distribution of observed vehicles. These efforts will lead to developing safe-decision-making algorithms for autonomous vehicles operating in urban environments [13].

ACKNOWLEDGMENTS

REFERENCES

[1] S. H. Park, B. Kim, C. M. Kang, C. C. Chung, and J. W. Choi, "Sequence-to-sequence prediction of vehicle trajectory via LSTM encoder-decoder architecture," in *IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1672–1678.

[2] R. Senanayake, L. Ott, S. O'Callaghan, and F. Ramos, "Spatio-temporal hilbert maps for continuous occupancy representation in dynamic environments," in *Advances in Neural Information Processing Systems (NIPS)*, 2016.

[3] R. Senanayake and F. Ramos, "Directional grid maps: Modeling multimodal angular uncertainty in dynamic environments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 3241–3248.

[4] W. Zhi, R. Senanayake, L. Ott, and F. Ramos, "Spatiotemporal learning of directional uncertainty in urban environments with kernel recurrent mixture density networks," *Robotics and Automation Letters (RA-L)*, vol. 4, no. 4, pp. 4306 –4313, 2019.

[5] T. Vintr, S. Molina, R. Senanayake, G. Broughton, Z. Yan, J. Ulrich, T. P. Kucner, C. S. Swaminathan, F. Majer, M. Stachová, *et al.*, "Time-varying pedestrian flow models for service robots," in *2019 European Conference on Mobile Robots (ECMR)*, IEEE, 2019, pp. 1–7.

[6] T. Vintr, Z. Yan, T. Duckett, and T. Krajník, "Spatio-temporal representation for long-term anticipation of human presence in service robotics," in *IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2019, pp. 2620–2626.

[7] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Wiley, 2006.

[8] B. Ivanovic, E. Schmerling, K. Leung, and M. Pavone, "Generative modeling of multimodal multi-human behavior," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.

[9] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, vol. 96, 1996, pp. 226–231.

[10] S. T. O'Callaghan, S. P. Singh, A. Alempijevic, and F. T. Ramos, "Learning navigational maps by observing human motion patterns," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 4333–4340.

[11] R. Herbrich, N. D. Lawrence, and M. Seeger, "Fast sparse gaussian process methods: The informative vector machine," in *Advances in Neural Information Processing Systems (NIPS)*, 2003, pp. 625–632.

[12] M. Wang, Z. Wang, S. Paudel, and M. Schwager, "Safe distributed lane change maneuvers for multiple autonomous vehicles using buffered input cells," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 4678–4684.

[13] L. Sun, W. Zhan, C.-Y. Chan, and M. Tomizuka, "Behavior planning of autonomous cars with social perception," in *IEEE Intelligent Vehicles Symposium (IV)*, 2019, pp. 207–213.