Augmented Vehicle Tracking under Occlusions for Decision-Making in Autonomous Driving

Enric Galceran, Edwin Olson, and Ryan M. Eustice

***Abstract*— This paper reports on an algorithm to support autonomous vehicles in reasoning about occluded regions of their environment to make safe, reliable decisions. In autonomous driving scenarios, other traffic participants are often occluded from sensor measurements by buildings or large vehicles like buses or trucks, which makes tracking dynamic objects challenging. We present a method to augment standard dynamic object trackers with means to 1) estimate the occluded state of other traffic agents and 2) robustly associate the occluded estimates with new observations after the tracked object re-enters the visible region of the sensor horizon. We perform occluded state estimation using a dynamics model that accounts for the driving behavior of traffic agents and a hybrid Gaussian mixture model (hGMM) to capture multiple hypotheses over discrete behavior, such as driving along different lanes or turning left or right at an intersection. Upon new observations, we associate them to existing estimates in terms of the Kullback-Leibler divergence (KLD). We evaluate the proposed method in simulation and using a real-world traffic-tracking dataset from an autonomous vehicle platform. Results show that our method can handle significantly prolonged occlusions when compared to a standard dynamic object tracking system.**

I. Introduction

Reasoning about unknown or occluded regions of a robot's surroundings is crucial for robots operating in a multi-agent setting. In the particular case of autonomous driving, autonomous vehicles must reason about the state of occluded dynamic objects to safely handle complex situations like passing, merging, or intersection handling. However, most decision-making approaches for autonomous vehicles to date assume complete knowledge of the states of dynamic objects in the environment [1–6], even if such objects transition through completely occluded regions of the vehicle's sensor horizon. Unfortunately, this assumption can lead to dangerous situations, as shown by the simulation in Fig. 1.

While occlusions can be mitigated by using sensors like some radar variants that allow sensing through solid objects [7], the geometric information provided by such sensors is sparser and less accurate when compared to LiDAR or cameras, on which we focus in this paper. Another alternative to mitigate sensor occlusions is to rely on external road sensing infrastructure [8]. However, such infrastructure is currently not available on most roads and it is hugely expensive to deploy.

This work was supported in part by a grant from Ford Motor Company via the Ford-UM Alliance under award N015392 and in part from DARPA under award D13AP00059.

E. Galceran, E. Olson, and R. Eustice are with the University of Michigan, Ann Arbor, Michigan 48109, USA
{egalcera, ebolson, eustice}@umich.edu.

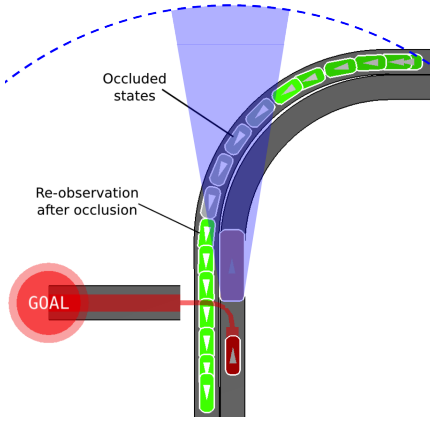


Fig. 1: Demonstration of our method in a simulated T-intersection navigation scenario. The self-driving vehicle (dark red) wishes to take a left turn while its observations are completely occluded for a few seconds by a large vehicle upfront (light red). The self-driving vehicle's occluded region is shaded in blue, and its sensor horizon is marked by the dashed blue line. A naïve object tracker would only observe the oncoming vehicle (green states) when it is visible within the sensor horizon, giving the self-driving vehicle a very short reaction time. In contrast, our method allows the self-driving vehicle to track the oncoming vehicle *even when the oncoming vehicle is occluded* (light gray states), thus enabling the self-driving vehicle to make a safer decision.

Therefore, a dynamic object tracking system that can handle prolonged, complete occlusions is highly desirable. We propose a model-based method to augment standard dynamic object trackers to track objects passing through prolonged occlusions. Thus, we are able to capture the state of previously-observed dynamic objects even when they lie completely in occluded regions of the current sensor horizon.

Our approach augments any dynamic object tracker that is capable of yielding Gaussian[1] state estimates of the tracked objects. It does so by providing model-based estimation for the duration of sensor occlusions and a data association mechanism to match estimated occluded states to new sensor observations of the target object.

To estimate occluded states, we employ a dynamics model that captures typical driving behavior of tracked dynamic objects on a road network as per [9]. We then can simulate forward in time the execution of the control actions prescribed by the model for the tracked objects, obtaining an estimate of their occluded future state. Furthermore, we employ a

[1]Requiring Gaussian estimates allows us to use a Gaussian mixture belief representation and to compute the Kullback-Leibler divergence (KLD) in closed form for data association, as we describe below.

hybrid Gaussian mixture model (hGMM) to capture multiple hypotheses over the state of the occluded object induced by road network topology elements like multiple lanes on a highway, bifurcations, or intersections. Finally, we handle re-observation of occluded dynamic objects by matching the estimated occluded states to actual sensor observations in terms of the KLD. We evaluate the performance of our method using simulations and real-world traffic-tracking data from an autonomous vehicle platform. The central contributions of this work are:

1) A method for tracking vehicles passing through prolonged occlusions, including a model-based estimation of the occluded state and a data association mechanism to associate estimated states to re-observations of the target object, and
2) Evaluation of our method in simulations and using a real-world traffic-tracking dataset from an autonomous vehicle platform.

II. Related Work

To date, most motion planning and decision-making approaches for autonomous driving assume good estimates from a dynamic object detection and tracking system [1–6]. A notable exception is the approach of Richter et al. [10], which considers occluded space by commanding a vehicle based on learned probabilities of collision, although their approach does not reason about the current state of dynamic objects based on prior observations. However, most trackers from prior work do not explicitly handle prolonged sensor occlusions such as that shown in Fig. 1.

Dynamic object trackers from the literature are typically based either on camera imaging or on LiDAR range-finders. Several early visual trackers were based on the Kalman filter, such as the approach by Marcenaro et al. [11]. More recent visual tracking approaches seek to explicitly handle occlusions in video sequences involving pedestrians [12, 13]. These approaches leverage motion planning algorithms to model possible paths taken by occluded tracks, however, these methods are specifically tailored for pedestrians and have been only shown to work over short occlusion horizons.

Early LiDAR-based approaches focused on detection, rather than tracking, of moving objects [14, 15]. More recent LiDAR-based trackers include those by Leonard et al. [16], Vu and Aycard [17], Petrovskaya and Thrun [18], and Choi et al. [19]. A notable state-of-the-art tracker by Held et al. [20] integrates a motion model, LiDAR, and color from vision. All of these trackers, however, do not to handle complete occlusions for prolonged periods of time, which can be problematic for decision-making in autonomous driving.

Estimating the occluded state of a tracked object is inherently a probabilistic problem. Several approaches in the literature seek to anticipate future states of dynamic objects using standard filtering techniques such as the extended Kalman filter [21, 22]. However, these methods often perform poorly when dealing with nonlinear dynamics and multimodalities induced by discrete decisions like continuing straight, merging or passing.

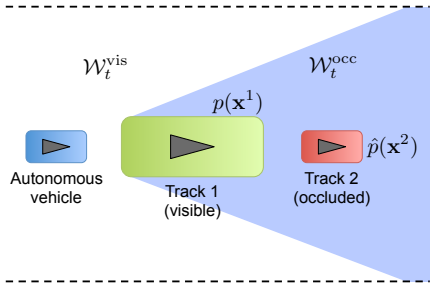


Fig. 2: Problem statement visualization. A self-driving vehicle (blue icon) must estimate the state of occluded dynamic objects, like track 2 in this example, that transition from the visible region of the workspace, $\mathcal{W}_t^{\text{vis}}$, into the occluded region of the workspace, $\mathcal{W}_t^{\text{occ}}$. In this work we propose to use prior dynamic object estimates, $p(\mathbf{x}^i)$, to achieve estimates of the future states of occluded dynamic objects, $\hat{p}(\mathbf{x}^i)$.

Recent work in traffic analysis applications and autonomous driving [23–25] uses Gaussian process (GP) regression to capture typical motion patterns of traffic participants and predict their future trajectories. Nonetheless, these methods require large amounts of training data in order to reflect the many possible motion patterns of a target vehicle, which can be time-consuming to collect.

A common anticipation strategy in autonomous driving (see, e.g., [6, 15, 26]) consists in determining the possible goals of a target vehicle by planning from its standpoint. Havlak and Campbell [27] propose a hGMM approach that formally captures probabilistic hypotheses over multiple discrete decisions. Here, we similarly adopt a hGMM representation for multiple hypotheses over the state of occluded tracks.

III. Problem Statement

As illustrated in Fig. 2, let $\mathcal{W}_t \subset \mathbb{R}^3 = \mathcal{W}_t^{\text{vis}} \cup \mathcal{W}_t^{\text{occ}}$ be the local workspace of our autonomous vehicle at time t, where $\mathcal{W}_t^{\text{vis}}$ is the region of the workspace directly observable with the vehicle's sensors (e.g., cameras, range sensors) and $\mathcal{W}_t^{\text{occ}}$ is the workspace region either beyond the sensor horizon or occluded by the presence of objects in the environment.

Current estimates $p(\mathbf{x}^i)$ of the state of dynamic objects within $\mathcal{W}_t^{\text{vis}}$, for the i^{th} of N dynamic objects, are assumed to be provided by a dynamic object detector and tracker (based, for instance, on one of the approaches described in [16–20]).

Hereby, we are interested in:

1) Estimating the occluded state $\mathbf{x}^i$ of each tracked dynamic object while the object lies completely within $\mathcal{W}_t^{\text{occ}}$, and
2) Correctly associating the estimated occluded state of each track $\hat{p}(\mathbf{x}^i)$ with a potentially re-observed object estimate $p(\mathbf{x}^i)$ when it re-enters $\mathcal{W}_t^{\text{vis}}$.

IV. Method

The flow diagram in Fig. 3 provides an overview of our proposed method for solving the problem stated in §III. For each tracked object passing through an occlusion, our approach iteratively propagates the last known state estimate provided by the tracker through a dynamics model that accounts for road network topology and driving behavior.

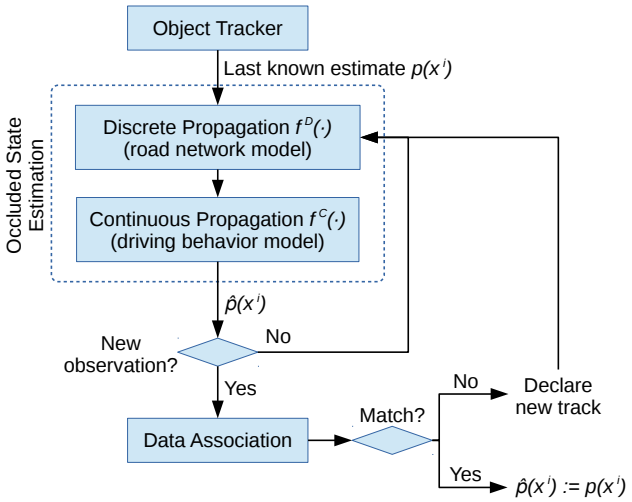


Fig. 3: Flow diagram of our proposed augmented tracking method for a single tracked object. Time subindices are omitted for clarity.

This propagation iterates until a new observation matching the estimated occluded state is received from the object tracker. We next elaborate on our driving behavior model, our occluded state representation and propagation strategy, and our data association method.

A. Dynamics Model

Building upon our prior work [9], we capture the potential behavior of dynamic objects in the environment as a discrete set of policies $\pi \in \Pi$, where each policy captures a typical high-level behavior, such as driving along a lane, doing a lane change, or turning at an intersection. The set of applicable policies in each driving scenario is dictated by a prior map of the environment [9]. Here, a policy is a mapping $\pi : \{1 \dots N\} \times \mathcal{B} \to \mathcal{A}$ that yields a prescribed action $a \in \mathcal{A}$ for dynamic object $i \in \{1 \dots N\}$ *only*, given the current belief $b \in \mathcal{B}$ over the state of *all* dynamic objects, i.e., $b = (p(\mathbf{x}^1), \dots, p(\mathbf{x}^N))$. Each action a prescribes steering-wheel-ange and forward-speed commands.

The key in this multipolicy model is that policies account for the state estimates of the commanded object *and* of other objects, thus yielding an action that reacts in closed loop to the actions of other traffic participants. For example, a lane-following policy can adapt the commanded vehicle speed to match the speed of the vehicle upfront, achieving an adaptive-cruise-control-type behavior. While it is possible that this multipolicy factorization could include behavior of pedestrians and other types of objects, in this paper we focus on vehicle-like dynamic objects.

This dynamics model allows us to effectively sample from the likely actions of traffic participants, providing likely estimates of the future occluded states of tracked objects.

B. Occluded State Representation

We represent the estimates of the occluded states of dynamic objects using a hGMM[2], which allows us to jointly denote the belief over the continuous state and multiple discrete hypotheses over the policies the object might be executing. For example, an occluded object might go through a bifurcation in the road network, after which we need to consider two discrete policy hypotheses, one for each fork.

Thus, similarly to Havlak and Campbell [27], we consider the underlying state $\mathbf{x}_t^i$ of each dynamic object i at time t to be partitioned into continuous components (such as pose, velocities, etc.) ${}^C\mathbf{x}_t^i$ and a vector of discrete hypotheses over policies ${}^D\mathbf{x}_t^i$:

$$\mathbf{x}_t^i = \begin{pmatrix} {}^C\mathbf{x}_t^i \\ {}^D\mathbf{x}_t^i \end{pmatrix}. \tag{1}$$

The occluded state estimate is given by the hGMM as

$$\hat{p}(\mathbf{x}_t^i) = \sum_{j=1}^{M_t} w_t^j \cdot \hat{p}^j(\mathbf{x}_t^i), \tag{2}$$

where M_t is the number of mixture components at time t and each w_t^j is the weight of each component such that

$$\sum_j w_t^j = 1.$$

Given that during an occlusion there are no observations available to infer each component weight w_t^j, in this work we use normalized weights $w_t^j = \frac{1}{M_t}$ throughout the augmented tracking process, assigning equal weight to all hypotheses. However, we note that prior knowledge could be incorporated into the discrete transition model $f^D(\cdot)$ to account for the relative frequency of discrete driving events. For example, at a highway off ramp, the relative frequency of a vehicle taking the ramp versus that of it continuing along the lane could be used to establish the component weights according to recorded data.

Each mixture component is defined as a Gaussian distribution over the continuous components of the state and a hypothesis over the policies the object is executing:

$$\hat{p}^j(\mathbf{x}_t^i) = \delta({}^D\mathbf{x}_t^i - \alpha_t^j) \cdot \mathcal{N}({}^C\mathbf{x}_t^i; \boldsymbol{\mu}_t^j, \Sigma_t^j), \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta function, α_t^j is the discrete hypothesis over the current policies, and $\boldsymbol{\mu}_t^j$ and Σ_t^j are the mean and covariance of the Gaussian over the continuous component of the state.

C. Occluded State Estimation

We estimate the occluded state of each dynamic object by propagating, in a two-step process, the hGMM through a probabilistic model accordingly partitioned into discrete and continuous components $f^D(\cdot)$ and $f^C(\cdot)$. First, we use the discrete component of the model to update the hypothesis over the current policies according to a prior road network

[2]Here hybrid refers to jointly capturing continuous and discrete components of the belief.

map:

$$f^D\begin{pmatrix}(w_t^1,\alpha_t^1,\boldsymbol{\mu}_t^1,\Sigma_t^1)\\ \vdots \\ (w_t^{M_t},\alpha_t^{M_t},\boldsymbol{\mu}_t^{M_t},\Sigma_t^{M_t})\end{pmatrix}= \begin{pmatrix}(w_{t^-}^1,\alpha_{t+1}^1,\boldsymbol{\mu}_{t^-}^1,\Sigma_{t^-}^1)\\ \vdots \\ (w_{t^-}^{M_{t+1}},\alpha_{t+1}^{M_{t+1}},\boldsymbol{\mu}_{t^-}^{M_{t+1}},\Sigma_{t^-}^{M_{t+1}})\end{pmatrix}. \quad (4)$$

Note that the number of components at the next timestep M_{t+1} might change after the propagation according to the updated hypothesis over policies.

After the discrete update, we update the continuous component of the estimate based on the multipolicy dynamics model presented above:

$$f^C\begin{pmatrix}(w_{t^-}^1,\alpha_{t+1}^1,\boldsymbol{\mu}_{t^-}^1,\Sigma_{t^-}^1)\\ \vdots \\ (w_{t^-}^{M_{t+1}},\alpha_{t+1}^{M_{t+1}},\boldsymbol{\mu}_{t^-}^{M_{t+1}},\Sigma_{t^-}^{M_{t+1}})\end{pmatrix}= \begin{pmatrix}(w_{t+1}^1,\alpha_{t+1}^1,\boldsymbol{\mu}_{t+1}^1,\Sigma_{t+1}^1)\\ \vdots \\ (w_{t+1}^{M_{t+1}},\alpha_{t+1}^{M_{t+1}},\boldsymbol{\mu}_{t+1}^{M_{t+1}},\Sigma_{t+1}^{M_{t+1}})\end{pmatrix}. \quad (5)$$

Here, each component is updated independently by simulating forward in time, for the duration of the timestep, the execution of action

$$a = \pi_k(i,(p(\mathbf{x}_t^1),\ldots,p(\mathbf{x}_t^N))), \quad (6)$$

where the index k into Π is determined by each policy hypothesis α_{t+1}^j of each track. The forward simulation of a is a non-linear function due to the underlying vehicle dynamics. Therefore, we use the sigma-points transform (also known as unscented transform) [28] to propagate each pair $(\boldsymbol{\mu}_t^j,\Sigma_t^j)$ through the corresponding policy and obtain $(\boldsymbol{\mu}_{t+1}^j,\Sigma_{t+1}^j)$.

D. Data Association

Upon receiving a new dynamic object track in $\mathcal{W}_t^{\text{vis}}$, we must decide whether the new observation $p(\mathbf{x}^{N+1})$ corresponds to an existing occluded track $\hat{p}(\mathbf{x}^i)$ or it is rather an actual new observation. Our solution to this problem is a nearest-neighbor data association in terms of the KLD between the new observation and each Gaussian component j of each existing (occluded) track i:

$$(i^*,j^*) = \operatorname*{argmin}_{i,j} D_{\text{KL}}(p(\mathbf{x}^{N+1})\|\hat{p}^j(\mathbf{x}_t^i)). \quad (7)$$

The KLD measures the relative entropy between two distributions, and can be computed in closed form for two multivariate Gaussians $\mathcal{N}_0(\boldsymbol{\mu}_0,\Sigma_0)$ and $\mathcal{N}_1(\boldsymbol{\mu}_1,\Sigma_1)$ of dimension d as

$$D_{\text{KL}}(\mathcal{N}_0\|\mathcal{N}_1) = \frac{1}{2}\left(\operatorname{tr}\left(\Sigma_1^{-1}\Sigma_0\right)+(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)^\top\Sigma_1^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)- d+\ln\left(\frac{\det\Sigma_1}{\det\Sigma_0}\right)\right). \quad (8)$$

Fig. 4: Autonomous vehicle sensor platform used for experimental results. This platform is a Ford Escape equipped with a TORC ByWire XGV drive-by-wire system, four Velodyne HDL-32E LiDAR scanners, and an Applanix POS-LV 420 INS. A dynamic object detection and tracking system similar to that of Leonard et al. [16] running onboard allows tracking of other traffic participants.

This applies to our case, since the tracker state estimates $p(\mathbf{x}^{N+1})$ and each component $\hat{p}^j(\mathbf{x}_t^i)$ are Gaussian.

If $D_{\text{KL}}(p(\mathbf{x}^{N+1})\|\hat{p}^{j^*}(\mathbf{x}_t^{i^*}))$ is below a user-provided threshold θ, we declare the new observation to match the corresponding existing occluded track. Otherwise, we declare the new observation to correspond to a new tracked object.

V. Experiments

We now explore the performance of our proposed method in simulation and using real-world traffic-tracking data acquired from our autonomous vehicle platform, shown in Fig. 4. We first illustrate the advantage of our augmented tracking approach versus occlusion-unaware tracking in an intersection navigation scenario, and then we show the capability of our method to capture multiple hypotheses over occluded vehicle states and correctly associate them to re-observations. Next, we quantitatively evaluate the accuracy of our occluded state estimation approach on real-world traffic tracking data from our autonomous vehicle platform using a tracking system similar to that of Leonard et al. [16]. Finally, we perform Monte Carlo simulations to evaluate the performance of our data association strategy as a function of dynamics model error in a congested traffic scenario. For these experiments, we use C implementations of our augmented tracking algorithm, a traffic simulation engine, and the dynamic object tracker. The experiments were performed using a standard PC with a 2.8GHz Intel i7 processor.

In all experiments, we consider the continuous state of occluded vehicles to be given by

$${}^C\mathbf{x}_t^i = (x,y,\psi,v)^\top,$$

where (x,y) is the 2-dimensional position of the vehicle on the plane, ψ is its orientation, and v is its forward speed. Similarly, we use a data association threshold $\theta = 55$ nats in all experiments (in simulation and using real-world data). A single lane-following policy [9] is used in these experiments to model the continuous behavior of all discrete hypotheses.

A. Decision-Making Scenario in Simulation

Our first experiment illustrates the importance of estimating the occluded states of other traffic participants using the simulated autonomous driving decision-making scenario introduced in Fig. 1. As shown in the comparison in Fig. 5, without using our augmented tracking method (e.g., as per [16]) our vehicle sees the oncoming vehicle disappear as it is occluded. As a result, it declares clearance and proceeds turning left only to detect the oncoming vehicle at an extremely risky close distance after it re-enters its field of view. The oncoming vehicle is forced to brake abruptly to avoid a collision. In contrast, using our proposed method our vehicle is able to track the oncoming vehicle throughout the occlusion, and hence account for its presence and yield to it before proceeding safely through the intersection.

B. Multi-hypothesis Scenario in Simulation

The following experiment, shown in Fig. 6, illustrates the capability of our method to account for multiple hypotheses over the occluded state of other vehicles via a hGMM. In this simulated highway driving scenario, our vehicle instantiates a hGMM with three hypotheses (determined by three possible lanes of travel) over the occluded state of a tracked vehicle. Despite the tracked vehicle executing a lane change maneuver during the occlusion, our method is able to correctly associate the re-observed state with the appropriate component of the multi-hypothesis occluded state estimate.

C. Real-world Traffic Dataset

To evaluate the accuracy of the occluded state estimates provided by our state propagation model, we recorded 32 dynamic object trajectories on University of Michigan's North Campus. Each recorded trajectory lasts between 13s and 38s. Since ground truth positioning data from the tracked objects is not available, we run our augmented tracking method and artificially inject an occlusion (that is, we drop tracking updates) for 60% of each trajectory, with the occlusion period centered at halfway through the trajectory. A sample augmented tracking sequence from the dataset is shown in Fig. 7. We note that this dataset includes trajectories involving a single lane-following behavior only. Thus, the focus of this experiment is to evaluate the accuracy of the continuous propagation model $f^C(\cdot)$.

We then compute the root mean squared error (RMSE) of the estimated occluded trajectory provided by our proposed method using the actual tracking data as ground truth. As shown in Fig. 8, on average our dynamics model keeps the estimation error under 6m after 20s of occlusion time, which we find to be sufficient to obtain correct data association in typical occlusions in traffic scenarios. We note, however, that a more accurate dynamics model that better captures longitudinal behavior would help further reduce this estimation error. Nonetheless, our method is able to handle occlusions of over 20s, while we have observed that the standard tracking system running on the vehicle loses track of objects in less than 1s after an occlusion.

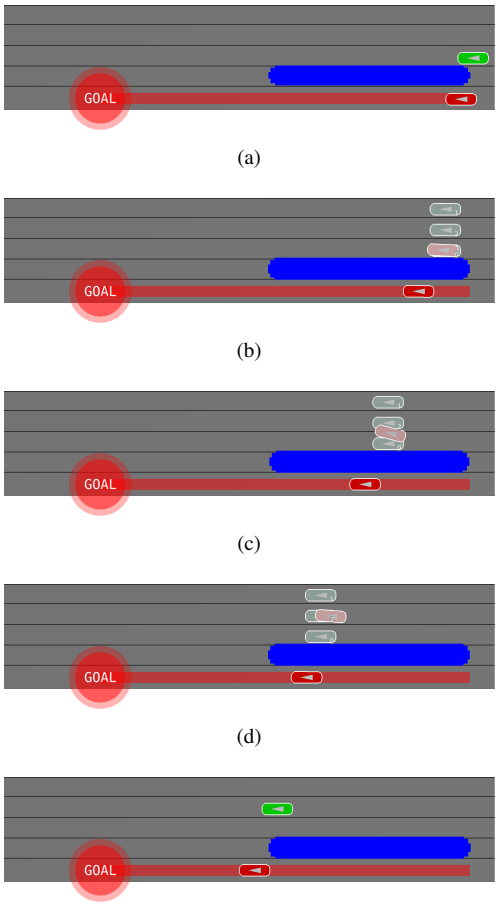


Fig. 6: Multiple-hypothesis augmented tracking under occlusions on a simulated highway segment. (a) The ego-vehicle (dark red) proceeds along its current lane while it tracks another vehicle (green). (b) As the tracked vehicle is blocked from our sensor field-of-view by an occluding object (blue, e.g., a large truck driving along the adjacent lane), the discrete propagation model $f^D(\cdot)$ instantiates a hGMM with three components (light gray), one for each possible hypothesis over each possible lane. (c-d) During the occlusion, the actual occluded vehicle (light red) changes lanes. (e) On re-observing the tracked vehicle, our method is able to correctly associate the observation with the appropriate hypothesis despite the lane change. Note: occluded regions not explicitly shown for clarity.

D. Data Association under Dynamics Modeling Error

We now evaluate the effect of dynamics modeling error on data association in the simulated scenario shown in Fig. 9, involving nine tracked vehicles passing through an occlusion. Initial covariances for all vehicles are set to $\Sigma = \mathrm{diag}([0.5, 1.0, 0.01, 0.05])$, and a single hypothesis is considered per vehicle. Before performing data association, we add zero-mean Gaussian noise to the mean longitudinal position of the occluded state estimates according to the standard deviation of the dynamics modeling error measured on real-world data (Fig. 8). That is, with standard deviation ranging from $\sigma = 0.63\mathrm{m}$ to $\sigma = 6.01\mathrm{m}$. For each σ value, we run 5000 Monte Carlo simulations and count the percentage of correct data associations.

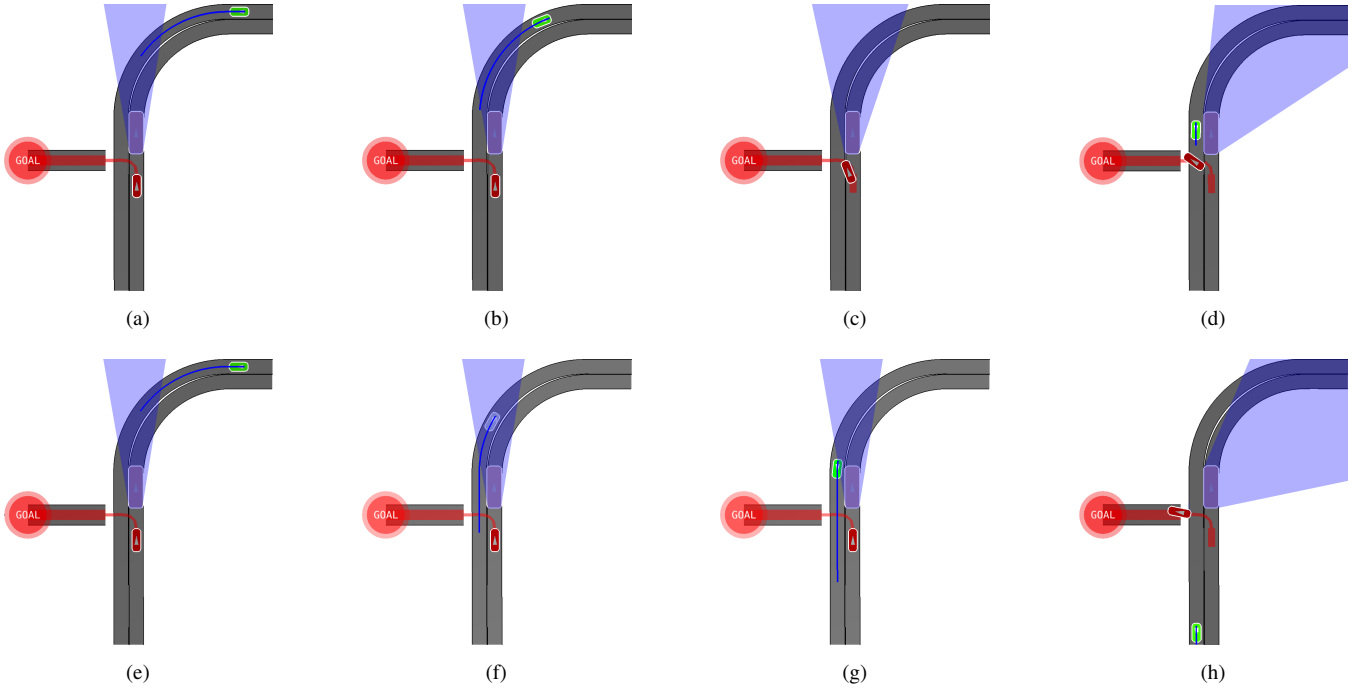


Fig. 5: Simulated intersection handling *without* (top row) and *with* (bottom row) augmented tracking under occlusions. The vehicle under our commmand (dark red) wishes to turn left at the T-intersection, as indicated by its mission plan (transparent red). A large vehicle upfront (light red) occludes our vehicle's sensor field-of-view as it tracks an oncoming vehicle in the opposite lane (green and light gray). Our vehicle uses a simple lane-following prediction scheme to estimate the future trajectory of the oncoming vehicle (blue line). *Without* our method (top row), the tracked vehicle disapears behind the occluding vehicle (c) and hence our controlled vehicle starts turning left after determining clearance. As a result, it leads itself and the oncoming vehicle to an extremely risky situation (d) where the oncoming vehicle is forced to brake abruptly. In contrast, *with* our method (bottom row) the occluded vehicle state is estimated throughout the occlusion (f, light gray), and thus our vehicle can account for the presence of the oncoming vehicle and yield to it before starting the turn.

Results are shown in Fig. 10, where we can observe a data association performance decay above approximately $\sigma = 3\mathrm{m}$, leading to 80% performance at $\sigma = 6\mathrm{m}$. These results show that our proposed data association strategy depends strongly on accurate dynamics modeling, particularly in the longitudinal direction.

VI. Conclusion

We have presented a method to augment standard dynamic object trackers to estimate the occluded state of other traffic agents and associate the occluded estimates with new observations after the tracked object re-enters the visible region of the sensor horizon. Our method performs occluded state estimation using a dynamics model that accounts for the driving behavior of traffic agents and a hGMM to capture multiple hypotheses over distinct discrete behaviors, such as driving along different lanes. Upon new observations, we associate them to existing estimates in terms of the KLD. We evaluated the proposed method in simulation and using a real-world traffic-tracking dataset from an autonomous vehicle platform, showing that our method handles significantly longer occlusions when compared to a standard tracking system.

In future work we plan to evaluate the proposed method in more traffic scenarios, such as intersections and roundabouts, using ground truth positioning from two or more time-synchronized autonomous vehicle platforms. Exploring more robust data association strategies is also a subject for further research.

Acknowledgment

The authors are sincerely grateful to Ryan Wolcott and Stephen Chaves for their insightful comments.

References

[1] M. Montemerlo *et al.*, "Junior: The Stanford entry in the Urban Challenge," *J. Field Robot.*, vol. 25, no. 9, pp. 569–597, 2008.

[2] J. Leonard *et al.*, "A perception-driven autonomous urban vehicle," *J. Field Robot.*, vol. 25, no. 10, pp. 727–774, 2008.

[3] C. Urmson *et al.*, "Autonomous driving in urban environments: Boss and the Urban Challenge," *J. Field Robot.*, vol. 25, no. 8, pp. 425–466, 2008.

[4] M. Likhachev and D. Ferguson, "Planning long dynamically feasible maneuvers for autonomous vehicles," *Int. J. Robot. Res.*, vol. 28, no. 8, pp. 933–945, 2009.

[5] D. Dolgov, S. Thrun, M. Montemerlo, and J. Diebel, "Path planning for autonomous vehicles in unknown semi-structured environments," *Int. J. Robot. Res.*, vol. 29, no. 5, pp. 485–501, 2010.

[6] J. Hardy and M. Campbell, "Contingency planning over probabilistic obstacle predictions for autonomous road vehicles," *IEEE Trans. on Robotics*, vol. 29, no. 4, pp. 913–929, 2013.

[7] Y. Yang and A. Fathy, "See-through-wall imaging using ultra wideband short-pulse radar system," in *Proc. IEEE Antennas and Propagation Society Int. Symp.*, 2005, pp. 334–337.

[8] T. Nothdurft, P. Hecker, T. Frankiewicz, J. Gacnik, and F. Koster, "Reliable information aggregation and exchange for autonomous vehicles," in *Proc. IEEE Vehic. Tech. Conf.*, Sept 2011, pp. 1–5.

[9] E. Galceran, A. G. Cunningham, R. M. Eustice, and E. Olson, "Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction," in *Proc. Robot.: Sci. & Syst. Conf.*, 2015.

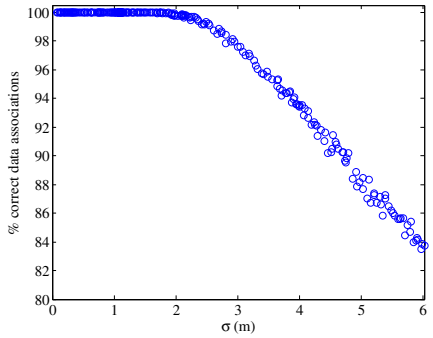
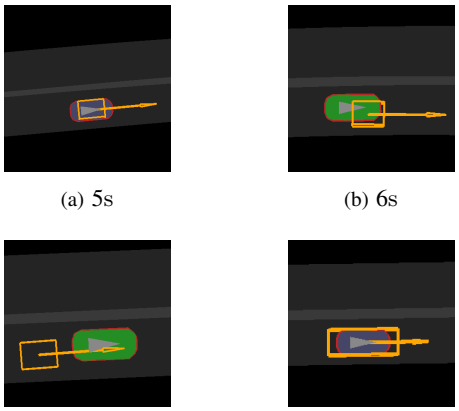


Fig. 10: Data association results under dynamics modeling error. Each data point corresponds to 5000 Monte Carlo simulations, where we count the percentage of correct data associations under longitudinal position error given by zero-mean Gaussian noise with standard deviation following the real-world error from Fig. 8.



(a) 5s (b) 6s

(c) 25s (d) 25.1s

Fig. 7: Sample prolonged augmented tracking on real-world data. Before the start of the injected occlusion (a), the estimated state (blue icon) and the dynamic object track (orange bounding box) coincide. After the occlusion starts (b), our method estimates the object's occluded state (green icon, indicating occlusion). After approximately 19s into the occlusion (c), the diversion between the estimated state and the true state peaks before associating the re-observation with the estimated state after the occlusion ends (d).

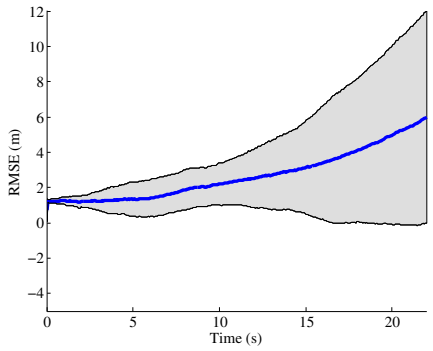
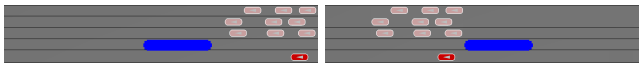


Fig. 8: Mean RMSE (blue) and 1-σ bounds (shaded gray) of the proposed occluded state estimation method evaluated on real-world traffic tracking data. On average, we are able to keep the estimation error below 6m after 20s of occlusion time.



(a) (b)

Fig. 9: Simulated scenario for the data association under noisy dynamics model experiment. (a) A self-driving vehicle (dark red) tracks nine other vehicles (light red) as they pass through an occlusion induced by the blue region. (b) Noise is added to the longitudinal position of the occluded state estimates according to the experimental data in §V-C before attempting association with the re-observed vehicles.

[10] C. Richter, J. Ware, and N. Roy, "High-speed autonomous navigation of unknown environments using learned probabilities of collision," in *Proc. IEEE Int. Conf. Robot. and Automation*, Hong Kong, May 2014, pp. 6114–6121.

[11] L. Marcenaro, M. Ferrari, L. Marchesotti, and C. Regazzoni, "Multiple object tracking under heavy occlusions by using kalman filters based on shape matching," in *Proc. Int. Conf. Image Process.*, vol. 3, 2002, pp. 341–344.

[12] H. Gong, J. Sim, M. Likhachev, and J. Shi, "Multi-hypothesis motion planning for visual object tracking," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 619–626.

[13] K. Fragkiadaki, W. Zhang, G. Zhang, and J. Shi, "Two-granularity tracking: Mediating trajectory and detection graphs for tracking under occlusions," in *Computer Vision – ECCV 2012*, ser. Lecture Notes in Computer Science, A. Fitzgibbon *et al.*, Eds. Springer Berlin Heidelberg, 2012, vol. 7576, pp. 552–565.

[14] C.-C. Wang, C. Thorpe, and S. Thrun, "Online simultaneous localization and mapping with detection and tracking of moving objects: Theory and results from a ground vehicle in crowded urban areas," in *Proc. IEEE Int. Conf. Robot. and Automation*, 2003.

[15] D. Ferguson, M. Darms, C. Urmson, and S. Kolski, "Detection, prediction, and avoidance of dynamic obstacles in urban environments," in *Proc. IEEE Intell. Vehicles Symp.*, June 2008, pp. 1149–1154.

[16] J. Leonard *et al.*, "Team MIT Urban Challenge technical report," Massachusetts Institute of Technology Computer Science and Artificial Intelligence Laboratory, Cambridge, Massachusetts, Tech. Rep. MIT-CSAIL-TR-2007-058, 2007.

[17] T.-D. Vu and O. Aycard, "Laser-based detection and tracking moving objects using data-driven markov chain monte carlo," in *Proc. IEEE Int. Conf. Robot. and Automation*, May 2009, pp. 3800–3806.

[18] A. Petrovskaya and S. Thrun, "Model based vehicle detection and tracking for autonomous urban driving," *Autonomous Robots*, vol. 26, no. 2–3, pp. 123–139, 2009.

[19] J. Choi, S. Ulbrich, B. Lichte, and M. Maurer, "Multi-target tracking using a 3d-lidar sensor for autonomous vehicles," in *Proc. IEEE Int. Conf. Intell. Transp. Sys.*, 2013, pp. 881–886.

[20] D. Held, J. Levinson, S. Thrun, and S. Savarese, "Combining 3d shape, color, and motion for robust anytime tracking," in *Proc. Robot.: Sci. & Syst. Conf.*, Berkeley, USA, July 2014.

[21] C. Fulgenzi, C. Tay, A. Spalanzani, and C. Laugier, "Probabilistic navigation in dynamic environment using rapidly-exploring random trees and gaussian processes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, Sep. 2008, pp. 1056–1062.

[22] N. Du Toit and J. Burdick, "Robotic motion planning in dynamic, cluttered, uncertain environments," in *Proc. IEEE Int. Conf. Robot. and Automation*, May 2010, pp. 966–973.

[23] K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 1164–1171.

[24] G. S. Aoude *et al.*, "Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns," *Autonomous Robots*, vol. 35, no. 1, pp. 51–76, 2013.

[25] Q. Tran and J. Firl, "Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression," in *Proc. IEEE Intell. Vehicles Symp.*, 2014, pp. 918–923.

[26] A. Broadhurst, S. Baker, and T. Kanade, "Monte carlo road safety reasoning," in *Proc. IEEE Intell. Vehicles Symp.*, 2005, pp. 319–324.

[27] F. Havlak and M. Campbell, "Discrete and continuous, probabilistic anticipation for autonomous robots in urban environments," *IEEE Trans. on Robotics*, vol. 30, no. 2, pp. 461–474, 2014.

[28] S. Julier and J. Uhlmann, "A new extension of the kalman filter to nonlinear systems," *Int. symp. aerospace/defense sensing, simul.*, 1997.