Localization under Topological Uncertainty for Lane Identification of Autonomous Vehicles

Samer B. Nashed[1,2], David M. Ilstrup[2], and Joydeep Biswas[1]

***Abstract*— Autonomous vehicles (AVs) require accurate metric and topological location estimates for safe, effective navigation and decision-making. Although many high-definition (HD) roadmaps exist, they are not always accurate since public roads are dynamic, shaped unpredictably by both human activity and nature. Thus, AVs must be able to handle situations in which the topology specified by the map does not agree with reality. We present the Variable Structure Multiple Hidden Markov Model (VSM-HMM) as a framework for localizing in the presence of topological uncertainty, and demonstrate its effectiveness on an AV where lane membership is modeled as a topological localization process. VSM-HMMs use a dynamic set of HMMs to simultaneously reason about location within a set of most likely current topologies and therefore may also be applied to topological structure estimation as well as AV lane estimation. In addition, we present an extension to the Earth Mover's Distance which allows uncertainty to be taken into account when computing the distance between belief distributions on simplices of arbitrary relative sizes.**

I. Introduction

Localization is an essential capability for autonomous mobile robots, including autonomous vehicles (AVs). Most localization algorithms use metric maps as aids in the localization process, which represent features in continuous coordinates. Topological maps represent space as discrete components (vertices) and their logical-spatial relationship (edges) where vertices and edges are taken in the graph theoretic sense. The motivating example in this paper is an AV which, in addition to requiring a metric location estimate, also requires a topological location estimate at the lane level. For example, the AV may need to know not just that it is on Pleasant Street, but whether or not it is in the left turn only lane. Moreover, events such as construction, traffic accidents, natural disasters, and native map errors may result in discrepancies between the topology suggested by the map and reality. The localization algorithm on the AV must be able to reason about this possibility.

Reasoning *globally* about all possible topologies is computationally intractable, since the number of unique topologies scales exponentially with the number of locations. Furthermore, there may be uncertainty in the number of locations. Moreover, global topological information is rarely present. Instead, we propose a method for reasoning about location and structure within the *local*, observable topology. Restricting the scope allows inference algorithms to reason about multiple topologies with varying numbers of nodes.

[1]Samer B. Nashed and Joydeep Biswas are with the College of Information and Computer Sciences, University of Massachusetts, Amherst, MA 01003, USA. Email: {snashed, joydeepb}@cs.umass.edu [2] David M. Ilstrup is at Nissan Research Center, 1215 Bordeaux Drive, Sunnyvale, CA 94089 Email: {david.ilstrup}@nissan-usa.com

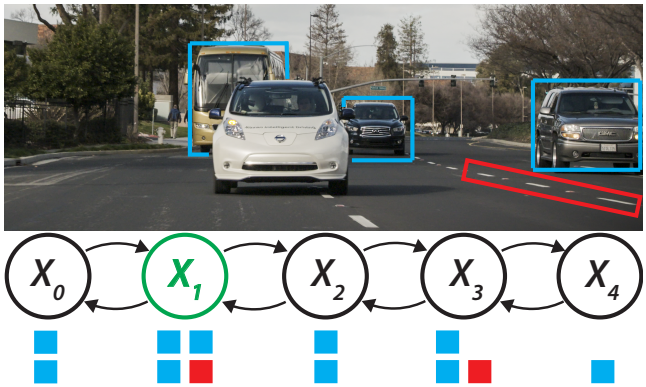

Fig. 1: Agreement of observations to HMM lane-states. State x_1, representing being between the right and center lanes, is the only state for which all three vehicle detections (blue) and the lane line detection (red) are likely for the AV (green).

The discrete nature of topological location, combined with the requirement to reason about multiple possible realities simultaneously, motivates our approach. The first key idea, shown in Fig. 1, is to use Hidden Markov Models (HMMs) for localization by modeling expected observations and transitions at and between topological nodes. The second key idea is to use a variable set of HMMs to model a variety of possible realities. Here, HMMs model the transition dynamics and observation models of a topological map, analogous to how Kalman filters model these aspects of a tracked object. Thus, we call this approach Variable Structure Multiple Hidden Markov Models (VSM-HMM); its function being similar to Variable Structure Multiple Models [11]. An important distinction is that VSMMs allow tracking of objects which follow a variety of process models, whereas the VSM-HMM approach reasons about multiple world models.

This paper presents three contributions. First, we demonstrate a method for applying HMMs to lane-level localization on an AV (§III). Second, we describe our VSM-HMM approach to managing a dynamic set of HMMs, each of which estimates a location within a unique local topology, allowing us to reduce dependence on high-definition (HD) maps (§IV). Third, we extend the Earth Mover's Distance (EMD) in order to handle distributions which have domains of different sizes during model belief initialization (§V).

Our approach is evaluated in simulation as well as on 6 real-world data sets gathered on public, multi-lane roads in Silicon Valley. Results presented in §VI show that the VSM-HMM model can provide accurate topological location estimates, as well as detect disagreements between the topology specified by the map and that supported by observation.

II. Related Work

There are several different approaches to topological localization. The problem of global topological localization may be modeled as a Partially Observable Markov Decision Process (POMDP). However, approaches using POMDPs, either in conjunction with geometric landmarks [25] or fingerprints from visual input [22], do not scale well with the number of topological nodes.

To reduce computational complexity, some approaches eliminate most of the reasoning about uncertainty by directly matching the robot's current view against representative feature vectors from topological locations in the map. A variety of map representations, feature extraction methods, and distance measures have been examined, including Generalized Voronoi Graphs [3], SIFT and SURF features [1], and Jeffrey divergence [24], respectively. These approaches work well for environments in which nodes can be visited often, and their representative feature vectors kept up-to-date [4]. However, this is rarely possible for an AV.

Localization algorithms specifically for AVs have typically focused on metric location, relying on high-definition (HD) maps for reliable, global data association. There are many variants, including approaches which use vanilla particle filters [18], Rao-Blackwellized particle filters [10], and Kalman Filters [19], [21], [5]. However, these approaches require HD maps and compute both metric and topological location in a single pass, which our approach does not.

In contrast, the proposed VSM-HMM model allows decoupling of metric and topological estimates, creating a hybrid metric-topological problem similar to [15], [16], although this paper focuses only on the topological component. Similar to FastSLAM [13], VSM-HMM maintains multi-modal belief over not only topological location, but also the local topological structure. Until now, particle filters were the only viable multi-modal topological localization framework. Moreover, particle filters resampling ignores local topological structure, whereas the VSM-HMM exploits this.

Dynamic Bayes Nets [14] are common tools for localization [8], [2], and HMMs specifically have been used for topological localization before. In [9] HMMs are used as an optional layer to process video feed in the case of a low-confidence nearest neighbor match, and in [7] HMMs are trained to detect and classify specific railroad turnouts using sequences of signals from eddy current sensors. Multiple HMMs have been suggested for other tasks where the number of classes is high, such as genome sequencing [6], and hierarchical HMMs [23] have been used for language, handwriting, and speech recognition.

III. Lane Identification using HMMs

Hidden Markov Models are promising candidates for topological localization for two primary reasons. First, HMMs are well understood theoretically and support many efficient modes of inference. In our application, we use the Forward algorithm because it affords low computational cost. Second, HMMs support learning and are easily designed for specific sensor features and or topological structure. Additionally, methods such as [20], can be used to refine observations prior to evaluation by the HMM.

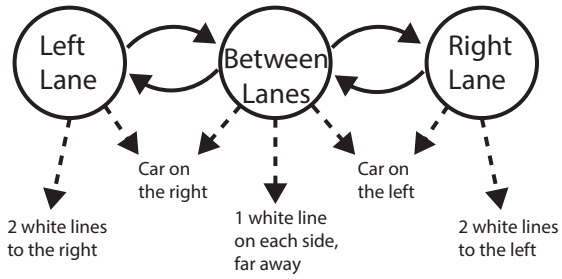

Fig. 2: Simplified example lane-state HMM. Solid lines represent non-zero transition probabilities between states. Self-loops are not shown. Dashed lines represent observation or emission probabilities, some of which are exclusive for a single state. Note that the transition structure of the HMM models the topological structure of the environment. Some types of obervations have been omitted for simplicity.

Topological location is modeled as occupancy of a state $x_i \in X$ within an HMM. The application of this work is toward lane-level localization, and thus the states X in the HMM correspond to either being in the center of a lane or being between two lanes. For example, an HMM representing a two lane road, detailed in Fig. 2, has three states x_0, x_1, and x_2. x_0 and x_2 correspond to occupying the right lane and the left lane, respectively. x_1 represents having some portion of the car over the lane divider. We call states like x_1 *switching states*. In general, an HMM representing a road with L lanes will have $2L - 1$ states.

A particular strength of HMMs in topological localization is their ability to efficiently model real-world dynamics via their transition function. Since the AV can only move from one lane to an adjacent lane by moving through an immediately adjacent switching state, the transition matrix representing the transition function is sparse. We define the state transition matrix for an n-state HMM, $\tau^{\mathbf{n}}$, as

$$\tau^n_{ij} = \begin{cases} t_r \text{ if } i = j \\ t_s \text{ if } |i - j| = 1 \\ 0 \text{ otherwise,} \end{cases} \tag{1}$$

where t_r and t_s are parameters for the probability of remaining in the same state and switching to an adjacent state, respectively. This transition matrix reduces reports of physically impossible events, such as instantaneous changes across multiple lanes, which is a key advantage over other multimodal approaches such as particle filters.

In addition to GPS and inertial sensors, we use lane line detections and observed relative locations of other vehicles to inform our topological estimate, shown in Fig. 3. For lane lines we use a combination of learned parameters and information from a map to parametrize a Gaussian Mixture Model which describes the likelihood of observing a lane line at positions and orientations relative to the AV, given a particular lane-state.

Since observations of lane lines are unreliable due to occlusions, weather and lighting conditions, and absence of the

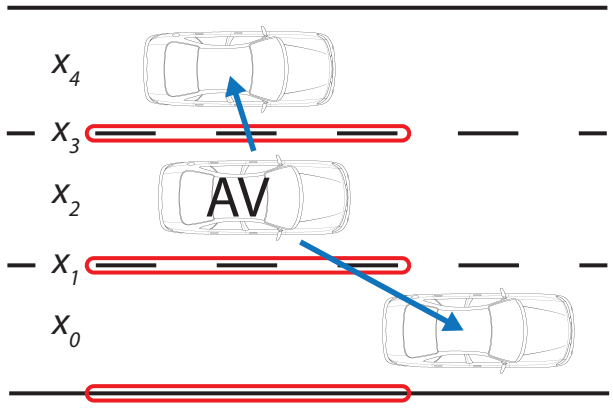

Fig. 3: Diagram of lane-states and observable features. Lane-states $x_0, \ldots x_4$ correspond to distinct topological regions on the road for which there are expected observations. x_2 is the only state which both lane line measurements (red) and vehicle detections (blue) support. Lane line measurements alone would result in equal belief in states x_2 and x_4.

markings themselves on many roads, we also use the relative positions of nearby tracked vehicles to support occupancy of certain lane-states. The key idea is that, although other vehicles move both globally and relatively with respect to the AV, they do so according to a specific local pattern defined by lane membership. For example, if the AV is traveling on a two-lane road and senses a vehicle immediately to its right, then there is a high probability that the AV is in the left lane, since the observed vehicle is far more likely to be traveling in the right lane than to be traveling beyond the edge of the road. We denote the observation function for state x_i and sensor q as $\phi(x_i, q)$.

Although single HMMs are reliable for topological localization when the map is correct and the number of states in the HMM matches the number true lane-states, they can fail when this is not true. To deal with this, we introduce the Variable Structure Multiple Hidden Markov Model.

IV. Variable Structure Multiple HMMs

The Variable Structure Multiple Hidden Markov Model (VSM-HMM) is similar to the variable-structure multiple-model set of Kalman Filters used in many tracking domains. However, whereas the multiple-model approaches in the tracking and control literature predict the current dynamical mode of the tracked object, the VSM-HMM approach estimates the current structure of the local topology. That is, the VSM-HMM hypothesizes about the outside world state rather than an internal process model. This is necessary when the local topological map has non-zero uncertainty.

We define a VSM-HMM as a set of all possible models U, an active model set U_A and an inactive model set U_N. Every $u \in U$ is an HMM with a unique transition matrix defined by the number of lanes. $U_A \cup U_N = U$, and $U_A \cap U_N = \emptyset$.

At every time step, U_A is determined using a variation of the Likely Model Set (LMS) algorithm [12], outlined in Algorithm 1. For every $u \in U$, we compute a model likelihood (line 6). $\Pr(u|M)$ is the probability of model u given the map M. If there was no uncertainty in M, then $\Pr(u|M)$ would be 1 for the model suggested by the map, u_M, and 0 otherwise. In our implementation, $\Pr(u|M) = \alpha 2^{-|(|u|-|u_M|)/2|}$, where $|u_M|$ is the number of states in u_M, and α is a normalizing constant. $\Pr(z|u)$ is the maximum probability of observing z given some state in u, $\textsc{Max}_{x_i} \Pr(z|x_i \in X_u)$. In Fig. 3, $\Pr(z|u)$ would be low for models with fewer than three lanes since z contains features which are unexpected in models with fewer than three lanes.

After model likelihood is computed, up to κ models are chosen so long as the ratio between their likelihood and the maximum likelihood of all models is above a threshold T_{active} (lines 8-14). κ is chosen based on computational constraints. Last, belief is copied from active models and initialized for inactive models (lines 15-19). Initializing belief is done using the Extended Earth Mover's Distance (§V).

Algorithm 1 Likely Model Set

1: **Input:** All models U, active models U_A, observations z, max number of models κ, threshold T_{active}, map M
2: **Output:** Set of updated most likely models U'_A
3: $U'_A \leftarrow \emptyset$
4: $S \leftarrow [\,]$
5: **for all** $u \in U$ **do**
6: $\quad \mathcal{L} \leftarrow \Pr(u|M) \times \Pr(z|u)$
7: $\quad S \leftarrow S.\textsc{append}(\mathcal{L}, u)$
8: $\mathcal{L}_{max} \leftarrow Max_{\mathcal{L}}(S)$
9: **for all** $\mathcal{L}, u \in S$ **do**
10: $\quad$ **if** $U'_A = \emptyset$ **then**
11: $\quad\quad U'_A \leftarrow U'_A \cup u$
12: $\quad$ **else**
13: $\quad\quad$ **if** $|U'_A| < \kappa$ and $\frac{\mathcal{L}}{\mathcal{L}_{max}} > T_{active}$ **then**
14: $\quad\quad\quad U'_A \leftarrow U'_A \cup u$
15: **for all** $u' \in U'_A$ **do**
16: $\quad$ **if** $u' \in U_A$ **then**
17: $\quad\quad u' \leftarrow \textsc{copyExistingBelief}(u')$
18: $\quad$ **else**
19: $\quad\quad u' \leftarrow \textsc{initNewBelief}(u')$
return U'_A

Discrepancies between the map and reality are detected by calculating the entropy, H, of the posterior probability (belief) over the states in each model:

$$H(bel(X)) = -\frac{1}{\log(|X|)} \sum_{i=1}^{|X|} bel(x_i) \log(bel(x_i)).$$

If the model suggested by the map has a high entropy compared to another model, the map is likely incorrect since high entropy indicates no state in the suggested topology can explain the observations. Note that the normalized equation for entropy calculates equal values for all models when no information is present. Thus, having a lack of information altogether will not cause the algorithm to flag the map as having an error. Only observations which both contradict the map's topology *and* may be explained by a different topology will result in a high entropy ratio. If $|U_A| > 1$, serving a localization requests amounts to picking the most likely state from the model with the lowest entropy.

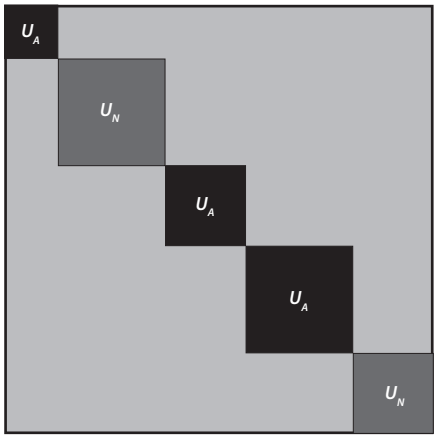

Fig. 4: Transition matrix for single HMM analog to VSM-HMM. Light gray areas hold probabilities of switching models, t_m. Block diagonals represent all models in U. Note that the VSM-HMM reasons about only the active models, represented by the black sub-blocks.

In theory, one could devise a single HMM with a dense transition matrix τ^*, which models the same problem. We can define τ^* in terms of the sub-blocks representing each distinct topological hypothesis, τ^1, τ^2, ..., τ^n (models in the VSM-HMM), and the transitions between them. Let each sub-block τ^k lie on the diagonal of τ^*. Further, let the off-diagonal blocks hold the probabilities of switching between sub-blocks, t_m. Note that t_m is a notational placeholder for a range of values corresponding to the probabilities of transitioning between two specific sub-blocks (models).

$$\tau^*_{ij} = \begin{cases} \tau^k_{lm} \text{ if } i,j \text{ index } l,m \text{ in sub-block } k \\ t_m \text{ otherwise.} \end{cases} \tag{2}$$

In general, this results in a $p \times p$ transition matrix, where $p = \sum_{u \in U} |X_u|$. Similarly, X^* and ϕ^* are defined by the union of all state spaces and observation functions, respectively. However, even when $U_A = U$, the VSM-HMM is more computationally efficient than its single HMM analog since calculating $\Pr(x_t|x_{t-1})$ only considers the block diagonals instead of the entire dense matrix, and calculating when to switch models depends only on the block diagonal size. In practice, $U_A \subset U$ is much more common. In this case, the VSM-HMM approximates the equivalent HMM by reasoning over a subset of the most likely belief points, shown in Fig. 4.

V. Extended Earth Mover's Distance

Whenever a model is initialized, it needs a starting belief. Suppose an AV has a belief, β, about its current topological position in a local topology. β is discrete and lies on the n-simplex, Δ^n, where $n+1$ is the number of local topological states. Here, $n+1 = 2L-1$, where L is the number of lanes on the road the AV is traveling. Further, suppose the AV is nearing an intersection or merge in which the number of lanes on the road the AV will end up on is L'. Once on the new road the AV will initialize a new model and need a new belief, β', about its topological location. However, if $L' \neq L$, β and β' will be over different numbers of states. The question is, if $L' \neq L$ how do we initialize β', given β.

One option is to erase all previous belief and start over from uniform, $\beta' = \mathcal{U}(0, m)$. Another option is to heuristically initialize β', such as right- or left-alignment of lane-states. Both options are computationally efficient, but do not perform optimally in many cases. A third option is to initialize β' as the 'closest' distribution in Δ^m to β, where 'closeness' is defined by some statistical metric. This is preferable, but there are no metrics which satisfy the constraints of the problem since there is no isomorphism between Δ^n and Δ^m, and the mapping from some belief point in Δ^n to the corresponding point Δ^m is uncertain. One example of uncertainty is a two-lane road which becomes a three-lane road across an intersection. It is unclear whether the two lanes in the first road correspond to the two rightmost, two leftmost, or some other combination of lanes in the three-lane road.

Thus, we introduce a statistical metric based on the Earth Mover's Distance (EMD) [17], called the Extended Earth Mover's Distance (EEMD), which measures the *expected* distance between distributions on simplices of arbitrary relative size, given the probability of all mappings between simplices. Proof of metric properties is in the Appendix. We initialize β' such that EEMD(β, β') is minimized.

Before defining EEMD, we introduce some notation. Let $\mathbb{P}^n$ and $\mathbb{P}^m$ be normalized distributions on Δ^n and Δ^m, respectively, and define $N = n+1$, and $M = m+1$. Without loss of generality, suppose $n > m$. Let the function $\mathbf{f}^{m,n} : \Delta^m \to \Delta^n$ be defined as

$$f^{m,n}_j(\mathbb{P}^m) = \begin{cases} \mathbb{P}^m_j \text{ if } j \leq M \\ 0 \text{ if } j > M. \end{cases} \tag{3}$$

Thus, $\mathbf{f}^{m,n}$ pads $\mathbb{P}^m$ with dimensions with zero belief, making it the same size as $\mathbb{P}^n$. We denote this new distribution, now on Δ^n, $\mathbb{P}^{m'}$. We can now use the original formulation of EMD to compute distance between $\mathbb{P}^{m'}$ and $\mathbb{P}^n$. However, in general, there may be uncertainty in the mapping between the two distributions. It may be that $\mathbb{P}^{m'}_j$ and $\mathbb{P}^n_j$ do not correspond to the same real world state. There are N^N possible mappings from $\mathbb{P}^{m'}$ to $\mathbb{P}^n$. We calculate the expected distance by summing over all possibilities for $\mathbb{P}^{m'}$, and compute the EMD weighted by the probability of each mapping, $\Pr(\mathbb{P}^{m'_i})$. Thus, we define the Extended Earth Mover's Distance between $\mathbb{P}^n$ and $\mathbb{P}^m$ as

$$\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = \sum_{i=1}^{N^N} \Pr(\mathbb{P}^{m'_i}) \text{EMD}(\mathbb{P}^n, \mathbb{P}^{m'_i}). \tag{4}$$

It is assumed that $\Pr(\mathbb{P}^{m'_i})$ is known and normalized. In practice we calculate it using information from the map, and our problem has structure allowing us to ignore most of the summands since the corresponding $\Pr(\mathbb{P}^{m'_i})$ term is 0. We use the EEMD as a principled guide to constructing distributions for model initialization. β' is calculated such that EEMD(β, β') is minimized.

VI. Results

To test the VSM-HMM framework, we perform two experiments. The first measures localization accuracy, and

the second tests the framework's ability to reason about local topological structure and detect discrepancies between the map and reality.

Localization accuracy was tested on 6 hand-annotated datasets gathered by an AV on public, multi-lane roads near Nissan Research Center in Silicon Valley. Each dataset was recorded over about a mile of stop-and-go traffic and ranged in time from 2 to 6 minutes. All road segments had between 3 and 6 lanes, corresponding to between 5 and 11 states. In these experiments no metric location information, such as GPS, was used, and topological ground truth was provided.

Because of the intermittent nature of the lane line and vehicle detections, not all timesteps possess enough observations to disambiguate lane-states. Thus, in the results presented in Table I we do not consider instances in which either no observations were recorded or the observations voted for at least half of all states, such as seeing only a single lane line immediately to the left of the vehicle. These instances are labeled "Missing Observations". Given sequences of timesteps with little or no observations, it is possible to have multiple states tie for the same belief. Localization is considered correct if the true state is among those with maximum belief, and incorrect otherwise. Length and observation quality are shown so as to give an idea about the difficulty of the dataset. Predictions are made at 100Hz.

Dataset	Length (mins)	Missing Obs.	Accuracy
1	2.0	11%	83%
2	3.7	18%	74%
3	4.8	6%	83%
4	4.4	9%	95%
5	3.5	30%	81%
6	5.7	20%	73%

TABLE I: Location estimation results

Testing topological structure estimation was done using simulated data, since there were too few cases in the real world data to draw concrete conclusions. To simulate false positive data, lane line and vehicle detections are generated according to the topology given by the map with probability P_M, and according to some other, randomly selected topology with probability $(1 - P_M)$. Further, to simulate the intermittent nature of real-world data, with probability $(1 - P_E)$ no observations are emitted. To see how our approach handles increased sensor noise, we tested different levels of variance. Given a variance σ based on real-world data, simulated observations are generated with variance $K_\sigma \sigma$, where K_σ is an experimental parameter.

Locations of lane lines and vehicle detections are sampled from multivariate normal distributions with means as a function of the given topology, and variances $K_\sigma \sigma$. The observation generation process runs independently for each lane line and vehicle detection. Table II displays the results of local topological structure estimation. Combined, these results demonstrate VSM-HMM as an effective framework for dealing with topological uncertainty.

P_E		0.9			0.7			0.5		
K_σ		1	2	3	1	2	3	1	2	3
P_M	0.9	96	95	92	94	88	83	83	78	72
	0.8	87	85	80	84	79	72	63	60	54
	0.7	75	74	72	66	65	61	55	50	50
	0.6	66	63	63	62	63	59	51	48	49

TABLE II: Local topological structure estimation accuracy. Results are reported as the percent of timesteps during which the correct topological structure was estimated with highest probability (lowest entropy). P_M is probability of sampling from the correct topology. P_E is the probability of emitting observations. K_σ is the amount by which the variance is scaled. Each entry in the table was computed from performance over 1000 timesteps.

VII. Conclusion

This paper presents a framework, Variable Structure Multiple Hidden Markov Models (VSM-HMM), for topological localization in the presence of topological uncertainty. We present empirical results from both simulated and real-world data on an autonomous vehicle which support VSM-HMM's effectiveness. Future work includes automating the generation and maintenance of observation and transition models through learning, as well as integrating this approach with other map representations.

References

[1] H. Andreasson and T. Duckett. Topological localization for mobile robots using omni-directional vision and local features. In *IAV*, 2004.

[2] J. Biswas and M. Veloso. Episodic non-markov localization: Reasoning about short-term and long-term features. In *International Conference on Robotics and Automation (ICRA)*. IEEE, 2014.

[3] H. Choset and K. Nagatani. Topological simultaneous localization and mapping (slam): toward exact localization without explicit localization. *IEEE Transactions on robotics and automation*, 17(2):125–137, 2001.

[4] F. Dayoub and T. Duckett. An adaptive appearance-based map for long-term topological localization of mobile robots. In *Intelligent Robots and Systems (IROS)*. IEEE, 2008.

[5] U. Franke, D. Pfeiffer, C. Rabe, C. Knoeppel, M. Enzweiler, F. Stein, and R. Herrtwich. Making bertha see. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2013.

[6] J. Gough, K. Karplus, R. Hughey, and C. Chothia. Assignment of homology to genome sequences using a library of hidden markov models that represent all proteins of known structure. *Journal of molecular biology*, 313(4):903–919, 2001.

[7] S. Hensel, C. Hasberg, and C. Stiller. Probabilistic rail vehicle localization with eddy current sensors in topological maps. *IEEE Transactions on Intelligent Transportation Systems*, 2011.

[8] M. Kaess, V. Ila, R. Roberts, and F. Dellaert. The bayes tree: An algorithmic foundation for probabilistic robot mapping. In *WAFR*, pages 157–173. Springer, 2010.

[9] J. Kosecká and F. Li. Vision based topological markov localization. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, volume 2, pages 1481–1486. IEEE, 2004.

[10] K. W. Lee, S. Wijesoma, and J. I. Guzmán. A constrained slam approach to robust and accurate localisation of autonomous ground vehicles. *Robotics and Autonomous Systems*, 55(7):527–540, 2007.

[11] X.-R. Li and Y. Bar-Shalom. Multiple-model estimation with variable structure. *IEEE Transactions on Automatic control*, 1996.

[12] X. R. Li and Y. Zhang. Multiple-model estimation with variable structure. v. likely-model set algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):448–466, 2000.

[13] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. In *Aaai/iaai*, pages 593–598, 2002.

[14] K. P. Murphy. Dynamic bayesian networks. *Probabilistic Graphical Models, M. Jordan*, 7, 2002.

[15] A. Ranganathan, D. Ilstrup, and T. Wu. Light-weight localization for vehicles using road markings. In *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*. IEEE, 2013.

[16] A. Ranganathan, S. Matsumoto, and D. Ilstrup. Towards illumination invariance for visual localization. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013.

[17] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, 40(2):99–121, 2000.

[18] A. Schindler. Vehicle self-localization with high-precision digital maps. In *Intelligent Vehicles Symposium*. IEEE, 2013.

[19] M. Schreiber, C. Knöppel, and U. Franke. Laneloc: Lane marking based localization using highly accurate maps. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 449–454. IEEE, 2013.

[20] S. Sivaraman and M. M. Trivedi. Integrated lane and vehicle detection, localization, and tracking: A synergistic approach. *IEEE Transactions on Intelligent Transportation Systems*, 14(2):906–917, 2013.

[21] Z. Tao, P. Bonnifait, V. Fremont, and J. Ibanez-Guzman. Mapping and localization using gps, lane markings and proprioceptive sensors. In *Intelligent Robots and Systems (IROS)*. IEEE, 2013.

[22] A. Tapus, N. Tomatis, and R. Siegwart. Topological global localization and mapping with fingerprints and uncertainty. *Experimental Robotics IX*, pages 99–111, 2006.

[23] S. F. Y. S. N. Tishby. The hierarchical hidden markov model: Analysis and applications.

[24] I. Ulrich and I. Nourbakhsh. Appearance-based place recognition for topological localization. In *International Conference on Robotics and Automation, 2000.*

[25] F. Zanichelli. Topological maps and robust localization for autonomous navigation. In *IJCAI Workshop on Adaptive spatial representations of dynamic environments*, 1999.

VIII. Appendix

Theorem 1. *EEMD is a metric, having the properties of non-negativity, identity, symmetry, and the triangle inequality.*

Non-negativity: Since EMD is a metric, it is always positive. $\Pr(\mathbb{P}^{m_i'})$ is always non-negative. Thus, their product is always non-negative, and the sum of non-negative elements is also non-negative.

Identity:
$(\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = 0 \implies \mathbb{P}^n = \mathbb{P}^m)$.
If $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = 0$, then for all summands, either $\Pr(\mathbb{P}^{m_i'}) = 0$ and or $\text{EMD}(\mathbb{P}^n, \mathbb{P}^{m_i'}) = 0$. Since $\Pr(\mathbb{P}^{m_i'})$ is a distribution, there must be at least one i such that $\Pr(\mathbb{P}^{m_i'}) > 0$. If there are more than one such i, then $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m)$ cannot be zero, since each $\mathbb{P}^{m_i'}$ is unique and EMD is a metric, violating the assumption $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = 0$. If there is a single i such that $\Pr(\mathbb{P}^{m_i'}) > 0$, then since EMD is a metric and must be 0, $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = 0 \implies \text{EMD}(\mathbb{P}^n, \mathbb{P}^{m_i'}) = 0 \implies \mathbb{P}^n = \mathbb{P}^m$.

$(\mathbb{P}^n = \mathbb{P}^m \implies \text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = 0)$.
This is clear from the definition of EEMD and $\Pr(\mathbb{P}^{m_i'})$. All $\Pr(\mathbb{P}^{m_i'})$ will be 0 except when $\mathbb{P}^{m_i'} = \mathbb{P}^m$. For this term, the corresponding EMD will be 0 since EMD is a metric.

Symmetry: The smaller dimension is always augmented, regardless of order. Thus, since $\Pr(\mathbb{P}^{m_i'})$ is constant, the exact same calculation is performed for both $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m)$ and $\text{EEMD}(\mathbb{P}^m, \mathbb{P}^n)$. So, $\text{EEMD}(\mathbb{P}^n, \mathbb{P}^m) = \text{EEMD}(\mathbb{P}^m, \mathbb{P}^n)$.

Triangle Inequality: Let m, n, and k be non-negative integers, and consider the three simplices, Δ^m, Δ^n, and Δ^k. Without loss of generality, let $m < n < k$ and define $N = n + 1$, $M = m + 1$, and $K = k + 1$.

Lemma 1.
$\text{EEMD}(\mathbb{P}^n, f^{m,n}(\mathbb{P}^m)) = \text{EEMD}(f^{n,k}(\mathbb{P}^n), f^{m,k}(\mathbb{P}^m))$.

From the definition of EEMD,

$$\text{EEMD}(\mathbb{P}^n, f^{m,n}(\mathbb{P}^m)) = \sum_{i=1}^{N^N} \Pr(\mathbb{P}^{m_i'})\text{EMD}(\mathbb{P}^n, \mathbb{P}^{m_i'}) \quad (5)$$

and

$$\text{EEMD}(f^{n,k}(\mathbb{P}^n), f^{m,k}(\mathbb{P}^m)) = \sum_{i=1}^{K^K} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}). \quad (6)$$

We call dimensions $1{:}N$ *essential* dimensions, and dimensions $(N+1){:}K$ *extra* dimensions. The sum on the RHS of equation (6) can be decomposed into two parts: one sum, with N^N terms, which corresponds to all mappings in which the extra $K - N$ dimensions map only amongst themselves, and another sum, with $K^K - N^N$ terms, which corresponds to all mappings where at least one of the extra dimensions maps to one of the essential dimensions. Thus, we can rewrite the RHS of equation (6) as

$$\sum_{i=1}^{N^N} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}) + \sum_{i=N^N+1}^{K^K} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}). \quad (7)$$

Since none of the extra dimensions has any meaning in the original problem and therefore cannot possibly map to any essential dimension, $\Pr(\mathbb{P}^{m_i'}) = 0$ for all $i > N^N$. Thus, the second sum is 0 and equation (6) becomes

$$\sum_{i=1}^{K^K} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}) = \sum_{i=1}^{N^N} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}). \quad (8)$$

Furthermore, since all extra dimensions have weight 0, their contribution to all summands in equation (8) is 0. So, equation (8) becomes

$$\sum_{i=1}^{K^K} \Pr(\mathbb{P}^{m_i''})\text{EMD}(\mathbb{P}^{n'}, \mathbb{P}^{m_i''}) = \sum_{i=1}^{N^N} \Pr(\mathbb{P}^{m_i'})\text{EMD}(\mathbb{P}^n, \mathbb{P}^{m_i'}), \quad (9)$$

establishing the lemma. Now, let $\mathbb{P}^m$, $\mathbb{P}^n$, and $\mathbb{P}^k$ be distributions on Δ^m, Δ^n, and Δ^k, respectively. Consider

$$\text{EEMD}(\mathbb{P}^m, \mathbb{P}^k), \text{ and}$$
$$\text{EEMD}(\mathbb{P}^m, \mathbb{P}^n) + \text{EEMD}(\mathbb{P}^n, \mathbb{P}^k).$$

By Lemma 1, we can rewrite these as

$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), \mathbb{P}^k), \text{ and}$$
$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), f^{n,k}(\mathbb{P}^n)) + \text{EEMD}(f^{n,k}(\mathbb{P}^n), \mathbb{P}^k).$$

All distributions now lie on Δ^k. Since all nodes in the simplex are unit distance from all other nodes, then the distance calculated by EMD for weight moving from one node to any other node will be the magnitude of the weight. Thus, if $\mathbb{P}^n$ maintains belief on any nodes with different magnitude than $\mathbb{P}^m$ or $\mathbb{P}^k$, then

$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), \mathbb{P}^k) <$$
$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), f^{n,k}(\mathbb{P}^n)) + \text{EEMD}(f^{n,k}(\mathbb{P}^n), \mathbb{P}^k).$$

Otherwise,

$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), \mathbb{P}^k) =$$
$$\text{EEMD}(f^{m,k}(\mathbb{P}^m), f^{n,k}(\mathbb{P}^n)) + \text{EEMD}(f^{n,k}(\mathbb{P}^n), \mathbb{P}^k).$$