Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

StaR Maps: Unveiling Uncertainty in Geospatial Relations

Benedict Flade, Simon Kohaut, Julian Eggert, Devendra Dhami, Kristian Kersting

2024

Preprint:

This is an accepted article published in 27th IEEE International Conference on Intelligent Transportation Systems. The final authenticated version is available online at: https://doi.org/10.1109/ITSC58415.2024.10920021 Copyright 2024 IEEE

StaR Maps: Unveiling Uncertainty in Geospatial Relations

Benedict Flade[1,*], Simon Kohaut[2,*], Julian Eggert[1], Devendra Singh Dhami[3], Kristian Kersting[2,4,5,6]

***Abstract*— The growing complexity of intelligent transportation systems and their applications in public spaces has increased the demand for expressive and versatile knowledge representation. While various mapping efforts have achieved widespread coverage, including detailed annotation of features with semantic labels, it is essential to understand their inherent uncertainties, which are commonly underrepresented by the respective geographic information systems. Hence, it is critical to develop a representation that combines a statistical, probabilistic perspective with the relational nature of geospatial data. Further, such a representation should facilitate an honest view of the data's accuracy and provide an environment for high-level reasoning to obtain novel insights from task-dependent queries. Our work addresses this gap in two ways. First, we present Statistical Relational Maps (StaR Maps) as a representation of uncertain, semantic map data. Second, we demonstrate efficient computation of StaR Maps to scale the approach to wide urban spaces. Through experiments on real-world, crowd-sourced data, we underpin the application and utility of StaR Maps in terms of representing uncertain knowledge and reasoning for complex geospatial information.**

***Index Terms*— Statistical Relational AI, Probabilistic Logic, Navigation**

I. Introduction

As intelligent transportation systems become increasingly sophisticated, we must carefully represent the agent's environment and accurately localize relevant features. Such systems fundamentally rely on environment representations that offer geometrical and relational information, ranging from static entities like roads to dynamic elements such as other traffic participants and pedestrians. Modern maps have been shown to include detailed 3D point clouds and semantic maps marked by polylines, such as road lane markings. Such maps serve diverse purposes and have been refined to meet nearly every accuracy requirement, provided that the hardware and conditions are adequate. For example, advancements in localization technologies, such as Real-Time Kinematics with Global Navigation Satellite Systems, have significantly improved accuracy in map creation. However, obtaining high-quality map data over large areas, such as entire cities or regions, can quickly become an intractable task due to the cost of the equipment and frequent changes in the respective areas. Instead, today's and future systems must work with data sources of mixed levels of accuracy, i.e., from crowd-sourced to high-definition, while incorporating their noisy perception of the environment.

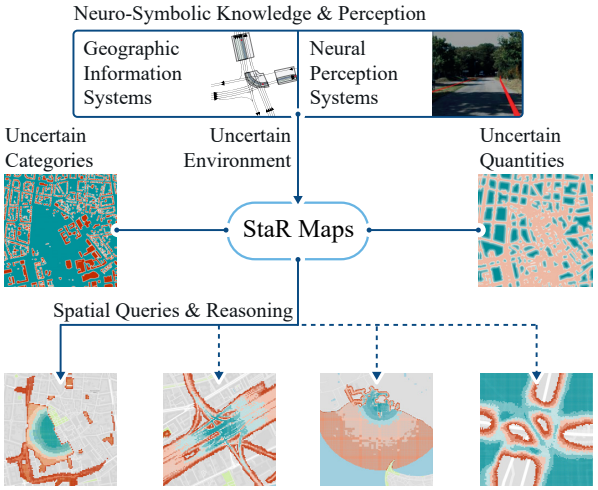


Fig. 1: **Statistical Relational Maps (StaR Maps) capture uncertain environments:** StaR Maps provide a unified interface to heterogeneous background knowledge. Given symbolic, geographic data, and neural perception of the navigation space, they offer semantic, probabilistic answers to first-order logic based spatial queries.

Traditional paper maps often include valuable accuracy statements such as diagrams and estimates of horizontal and vertical positional errors. Regrettably, such uncertainty information is frequently absent in the digital age [1]. When drawing conclusions from map data of multiple sources, we believe that an understanding of the accuracy of each individual map element can significantly enhance spatial reasoning by providing uncertainty information (see Figure 1).

However, with digital maps offering entirely new possibilities for data management, more than merely including an accuracy diagram is required. Besides categorical and quantitative uncertainties, it is important to facilitate reasoning over the data that attends to the attached uncertainties. For example, consider a scenario where the permitted driving speed is related to the distance to the next junction. With uncertainty of the agent's location and the exact location of the junction, it is important to provide a probabilistic perspective into the agent's decision-making process.

Nevertheless, the integration of detailed statistical data into maps has been limited. This oversight may be due to

* Authors contributed equally
[1] Honda Research Institute Europe GmbH, Carl-Legien-Str. 30, 63073 Offenbach, Germany
fist.last@honda-ri.de
[2] Artificial Intelligence and Machine Learning Lab, Department of Computer Science, TU Darmstadt, 64283 Darmstadt, Germany
[3] Uncertainty in Artificial Intelligence Group, Department of Mathematics and Computer Science, TU Eindhoven, 5600 MB Eindhoven, Netherlands
[4] Hessian AI
[5] Centre for Cognitive Science
[6] German Center for Artificial Intelligence (DFKI)

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021

the previously adequate levels of accuracy for the needs of most systems or the focus on meeting specified accuracy requirements rather than the need to analyze and evaluate its uncertainty. However, with the combination of statistical and probabilistic approaches being on the rise [2], it is appropriate to consider whether and how environmental representation should be adapted to benefit from this.

To this end, we introduce Statistical Relational Maps (StaR Maps). StaR Maps provide hybrid probabilistic and spatial relations for navigation and demonstrate their broad range of applications in designing a novel probabilistic framework for planning and complex geospatial queries.

Our contributions are two-fold:

- We present Statistical Relational Maps (StaR Maps), hybrid probabilistic, semantic maps that embrace mixed levels of uncertainty in geographic features and provide probabilistic answers to spatial queries.
- We demonstrate how StaR Maps underpin methods that rely on interpretable semantics to constrain and describe navigation tasks, i.e., utilizing a rich vocabulary of spatial relations that StaR Maps provide.

Our results are published as Open Source implementation within our Probabilistic Mission (ProMis) framework at `https://github.com/HRI-EU/ProMis`.

II. Related Work

A. Probabilistic Localization

In this work, we address the topic of uncertainty information in the context of geospatial data. One crucial aspect of this is the topic of localization, i.e., estimating the own or another traffic participant's position and attitude. In absolute localization systems, i.e., with regard to a global reference, approaches have progressed from conventional Global Navigation Satellite Systems (GNSS) to more sophisticated methods such as Real-Time Kinematic (RTK) GNSS [3]. Methods that address challenges in absolute localization, e.g., multipath effects in urban environments, can improve localization accuracy significantly [4]. Similarly, proprioceptive sensors observing internal states, e.g., acceleration [5], or exteroceptive sensors observing the environment, e.g., cameras [6], LiDAR [7], or radar [8], are used to improve localization accuracy. It has been shown that the ego-position, e.g., for street vehicles, can be improved by comparing proprioceptive and exteroceptive observations with features stored in a map [9], [10].

For map-relative localization, environment representations range from 3D point clouds [11] to polyline-level road shapes [12], with road markings presenting one of the most commonly used entities of reference [13], [14]. Furthermore, reaching resource requirements can be achieved by choosing a low-cost approach targeting affordable sensor setups [15]. However, understanding uncertainties in the underlying data is equally crucial for quantifying the localization system's output's trustworthiness. For example, localization approaches based on probabilistic methods, e.g., Bayes filters such as variations of the Kalman filter [16], [17], provide confidence data alongside the expected position information.

B. Statistical Relational Artificial Intelligence

Researchers and engineers have considered various approaches to meet the demands of intelligent transportation systems. To this end, both reasoning, e.g., based on propositional or first-order logic, have been successfully employed for representing and understanding the environment and task at hand. The emerging field of Statistical Relational Artificial Intelligence (StaR AI) has shown meaningful advances in combining tools of those fields into more robust and expressive inference systems [18].

StaR AI approaches have in the past been built based on programmatic reasoning systems in first-order logic such as Prolog [19]. That is, probabilistic extensions that embrace uncertainties in formal logic, e.g., Bayesian Logic Programs [20] and Probabilistic Logic Programs [21], [22], allow for probabilistic inference in first-order logic. While they were not formulated for end-to-end learning with artificial neural networks, Neuro-Symbolic models such as DeepProbLog [23], NeurASP [24] and SLASH [25] have been introduced to close this gap, embracing the advantages of neural information processing and probabilistic reasoning into a single framework.

Applying these methods in building autonomous agents, e.g., robots acting in logically constrained environments, is attractive to ensure safe and reliable decision-making. Notably, systems such as ProMis [26] utilize hybrid probabilistic logic [27], [28], combining categorical and continuous distributions into a spatially independent and identically distributed navigation knowledge and constraints model. As a result of this, such systems provide an interpretable and adaptable framework for planning tasks.

C. Uncertainty-Aware Environment Representation

Two major requirements must be met for intelligent transportation systems to leverage the strengths of current and future StaR AI approaches in spatial tasks, such as laying out optimal paths or searching for spaces that fulfill predetermined requirements. First, the logical component of statistical relational approaches is based on symbols and semantics that allow for modeling entities, their relations, and interactions. Transferred to a representation of the environment, features of the environment need to be assigned precise semantics.

While, e.g., unstructured point cloud maps or unlabeled structures in image data are unsuitable, segmentation approaches such as segment anything [29] and object-centric vision [30], [31] can help to separate objects which can be assigned semantic labels. To this end, work on multi-target tracking [32] and point cloud classifiers [33] have shown successful identification and tracking of objects with their assigned semantics.

A prominent example of creating a semantically annotated, wide-coverage environment representation is the crowd-sourced project OpenStreetMap [34], storing its geospatial

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021

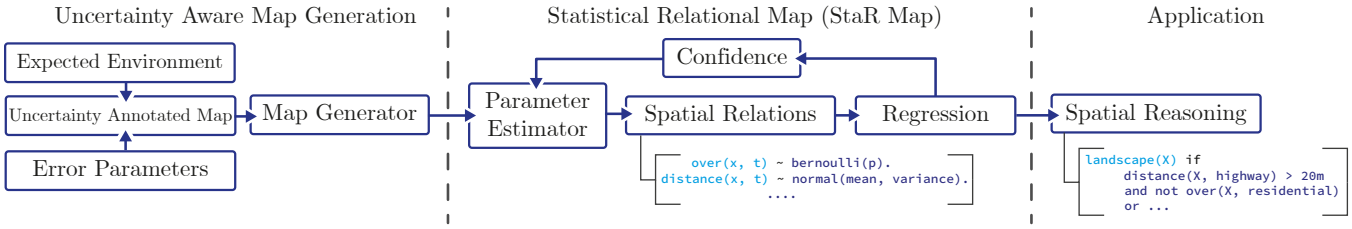


Fig. 2: **The Statistical Relational Maps (StaR Maps) architecture:** Uncertain maps, annotated with translational and transformational statistics, are passed into a density estimator. Given a set of sample locations and the map sampler, generating variations of the annotated data according to its uncertainty, spatial relations are estimated. Using, e.g., moment matching, for each sample request the spatial relations are represented as categorical or continuous distributions. Finally, a scalar field for each relation is interpolated or approximated from the samples, providing a basis for spatial reasoning.

data in nodes and ways alongside a rich tagging system. Further examples of semantic maps include Lanelets [35], atomic interconnected road segments that may carry additional data to describe the static environment. Additionally, approaches such as the Relational Local Dynamic Map [36] represent the environment's static, quasi-static, transient, and dynamic entities as a holistic, interconnected graph.

Many map-creation methods exist. From Lidar- and camera-based mapping [37]–[39] via map data derived from aerial imagery [40], to map generation by analyzing GNSS-based vehicle trajectories [41]. When employing such localization approaches to create maps, the uncertainties of the map-making process also apply to the generated map. In the worst case, the quality of geospatial data in an area, e.g., within a set of local roads, may drastically differ from feature to feature. Even within a single framework such as OpenStreetMap, data from amateur cartographers with affordable equipment intersects with professional data from commercial navigation information providers. More specifically, the authors of [34] criticize that OpenStreetMap does not provide information on the data's positional accuracy. Furthermore, being based on multiple heterogeneous data sources that have extensive coverage introduces not only different variances but distinct error types, and how to account for them simultaneously in a formalized holistic and map-centered way needs to be addressed.

To bridge this gap, we introduce StaR Maps, a novel environment representation and geographic reasoning system that considers heterogeneous sources of map data by incorporating different levels of fine-granular spatial uncertainty into a unified, semantically annotated environment representation. The novelty is a holistic framework that allows the representation of map-related error parameters, i.e., uncertainty information, processes this information, and stores this processed data in view of being subsequently used by applications that rely on a clear understanding of data quality. Hence, StaR Maps allows for a probabilistic worldview and facilitates complex, spatial reasoning and inference about the application and domain-specific queries.

III. Methods

This section introduces the concept of Statistical Relational Maps (StaR Maps). As motivated before, StaR Maps provide an approach to represent and reason on uncertain spatial, relational information. We first define how to combine expectations about the environment with a stochastic error model into Uncertainty Annotated Maps and how they facilitate sampling spatial relations, functions that relate points in navigation space with semantic map features. By further dividing spatial relations into categorical and quantitative relations, e.g., the predicate of a point being over a type of feature or its distance from it. Finally, we show how StaR Maps choose support points to create scalar fields for each of the spatial relations' parameters.

A. Uncertainty Aware Map Generation

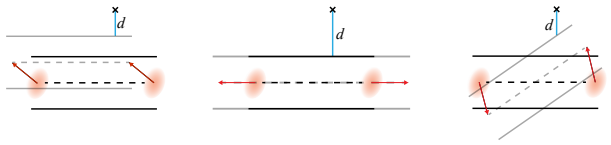
Uncertainty information is crucial when interpreting map data or using it to reason about spatial tasks such as navigation. However, before discussing the implications and utility of considering uncertainty in maps, let us define what we will understand as a map in the following.

Definition 1.1 (Map). A Map $\mathcal{M} = (\mathcal{V}, \mathcal{E}, \tau)$ is a triple of vertices $\mathcal{V}$, edges $\mathcal{E}$ and tagging function τ. Each vertex $\vec{v}$ is in a d-space $\mathbb{R}^d$ of Cartesian coordinates. If a path exists between two vertices in $\mathcal{V}$ across edges in $\mathcal{E}$, they are part of the same *feature*. For each vertex $\vec{v} \in \mathcal{V}$, the function $\tau(\vec{v}) \subseteq \mathbb{T}$ annotates $\vec{v}$ with a set of tags, providing semantics such as road or building types.

Today, many examples of available maps exist, e.g., the crowd-sourced OpenStreetMap [34] or concepts such as the Relational Local Dynamic Map [36]. While these concepts can represent an expectation, they commonly lack the associated uncertainty data. As depicted in Figure 2, we are now introducing uncertainty-aware map generation, combining expectations of the environment represented as a Map with an error model, describing how individual vertices or features likely deviate from the expectation due to the employed sensing methods and equipment. This combination of map data and uncertainty parameters leads us to the following definition.

Definition 1.2 (Uncertainty Annotated Map). An Uncertainty Annotated Map (UAM) $\mathcal{U} = (\mathcal{M}, a, b)$ is a triple of a map $\mathcal{M}$ and two annotator functions a and b.

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021



(a) Translation (b) Scaling (c) Rotation

Fig. 3: **Sampling from UAMs:** Each map element carries an expectation of its true spatial occupancy, annotated with uncertainties that result from, e.g., the employed sensors or measurement methodology. Once a set of maps has been sampled, we probe the environment from a selection of points to fit StaR Maps' spatial relations. Here, sampling the distance of a point to the closest uncertain road with different error models (a - c) is illustrated.

While the map follows Definition 1.1., functions $a(\vec{v}) = \vec{\alpha}$ and $b(\vec{v}) = \vec{\beta}$ assign each vertex $\vec{v}$ transformation parameters $\vec{\alpha}$ and translation parameters $\vec{\beta}$.

Here, $\vec{\alpha}$ and $\vec{\beta}$ are employed as statistical moments. For example, $\vec{\beta}$ might contain mean and variance of a randomly distributed translation. Hence, a UAM represents expectations alongside the parameters of a stochastic error model. While annotating each vertex of a map with individual parameters can be beneficial rather than generating entire map features using the same spatial uncertainty, this choice in granularity will have no systematic impact on our further discussions.

A UAM can be employed as generator of likely maps, e.g., facilitating how we will sample data from spatial relations in Section III-B. To this end, we apply the following affine map as a simple error model, covering perturbations such as translations, rotation, scaling, or shearing effects. Figure 3 illustrates this process and how, e.g., a distance from a fixed point to the closest road changes across sampled maps.

Similar to prior work [42], we consider for each $\vec{v} \in \mathcal{V}$ of map $\mathcal{M}$ the following manipulations:

$$\vec{\Phi}_n \leftarrow \phi(\vec{\alpha}) = \phi(a(\vec{v})) \qquad \text{(Transformation)}$$
$$\vec{t}_n \leftarrow \kappa(\vec{\beta}_i) = \kappa(b(\vec{v})) \qquad \text{(Translation)}$$

Here, ϕ generates matrices $\vec{\Phi}_n$ to apply geometric transformations, e.g., rotations, that keep the center point fixed while κ generates offset vectors $\vec{t}_n$ to apply a translation. Hence, from an original vertex $\vec{v}$ we can generate

$$\vec{v}_n = \vec{\Phi}_n \cdot \vec{v} + \vec{t}_n \tag{1}$$

We are repeating this process across the map N times, resulting in a collection $\mathcal{W} = \{\mathcal{M}_0, ..., \mathcal{M}_N\}$. In the following, we employ $\mathcal{W}$ to estimate the parameters of StaR Maps' spatial relations.

Note that the proposed error models allow the user to make assumptions about the error, e.g., considering how the map is generated. Such uncertainty information could be derived from the sensors used to create the maps, as far as such information is available. However, in cases with limited prior knowledge, we previously showed how to identify such error parameters online for road segments by observing deviations with a hybrid localization system [42].

B. Spatial Relations

Spatial relations are crucial for constraining the operation of intelligent transportation systems, whether on the ground, in the air, or submerged in water. Regardless of application, adhering to precise rules in an uncertain environment is a common requirement. To facilitate a framework for reasoning on such requirements, we show spatial relations as a language to express stochastic relations between points in space, types of map features, and instantiations of UAMs.

We consider the following two models of spatial relations.

$$f: \ \mathbb{R}^d \times \mathbb{T} \times \mathbb{M} \to \mathbb{B} \tag{2}$$
$$g: \ \mathbb{R}^d \times \mathbb{T} \times \mathbb{M} \to \mathbb{R} \tag{3}$$

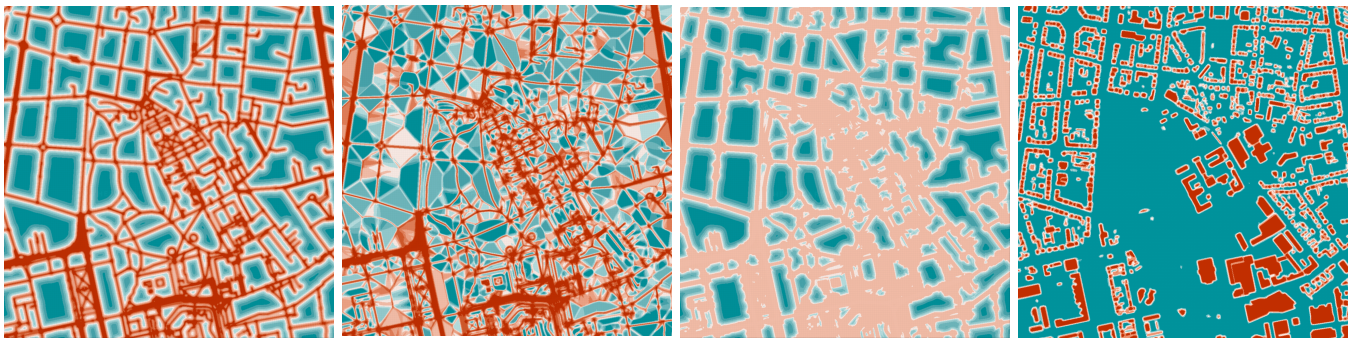
Here, a spatial relation with signature f describes a categorical relationship between a point $x \in \mathbb{R}^d$, a type of map features $\tau \in \mathbb{T}$ and a map $\mathcal{M} \in \mathbb{M}$. Analogously, spatial relations with signature g describe quantitative relationships between the same input data. Many functions will fit either definition; here, we consider the geometric relations *over* and *distance* as running examples of f- and g-type spatial relations, respectively. While the former expresses whether a point in space is above or, in the 2D case, within a map feature of the referenced type, the latter encodes the distance between a point and the closest matching geometry.

Note that these are binary relations where all geospatial features of the same type are considered anonymously, i.e., as a whole. By doing so, rather than computing the distance between a point and all individual features, the following discussions remain computationally tractable. Next, we discuss how the computation of spatial relations f and g over a set of sampled maps $\mathcal{W}$ sets up the data we need to define and apply StaR Maps. While the differentiation between f and g remains important throughout the rest of this work, we will denote function r as a generalization of any spatial relation.

C. Statistical Relational Maps

To summarize, we consider environment expectations alongside a stochastic error model to introduce the concept of UAMs and the process of generating a set of possible instances of maps $\mathcal{W}$ from them. Although we have defined spatial relation r to provide concrete, Boolean, or real-valued quantities, for StaR Maps, we are interested in probabilistic models that express the UAM's encoded uncertainty. To compute these statistics for a specific spatial relation r and a type τ on a point $\vec{x} \in \mathbb{R}^d$, we have to compute $\rho_n = r(\vec{x}, \tau, \mathcal{M}_n)$ for each $\mathcal{M}_n \in \mathcal{W}$. With the set $\mathcal{P} = \{\rho_0, ...\rho_N\}$ at hand, we can then compute the statistical moments, e.g., mean and variance of $\mathcal{P}$. Using moment matching with the desired distribution, for example, a Bernoulli distribution for f- or Gaussian distribution for g-signature relations, we can store the resulting parameters within a StaR Map.

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021



(a) μ_D (b) σ_D^2 (c) $P(D > 30m)$ (d) $P(O)$

Fig. 4: **Scalar and vector fields in StaR Maps express uncertainties in spatial relations:** Queried road network for which random maps have been generated. (a) and (b) show parameters of a normal distribution that model the distance to the closest road, while (c) models the probability of keeping a distance over 30 m. (d) shows the probability of a location of an agent's navigation space being occupied by buildings. Note a color range from red (low) to blue (high) for (a)-(c) and an inverse color range due to visual clarity for (d).

Definition 1.3 (Statistical Relational Map). A Statistical Relational Map (StaR Map) $\mathcal{S} = (\mathcal{F}, \mathcal{G}, u_{i,j}, v_{k,l})$ is made up of sample collections $\mathcal{F} = \{F_{i,j}, ...F_{I,J}\}$ and $\mathcal{G} = \{G_{k,l}, ...G_{K,L}\}$ as well as the functions $u_{i,j} : \mathbb{R}^d \times \mathbb{T} \rightarrow [0,1]^e$ and $v_{k,l} : \mathbb{R}^d \times \mathbb{T} \rightarrow \mathbb{R}^e$. While $F_{i,j}$ and $G_{k,l}$ contain data computed for each pair of spatial relation and type over $\mathcal{W}$, the functions $u_{i,j}$ and $v_{k,l}$ approximate scalar fields for the e parameters of each relation given the data in $F_{i,j}$ and $G_{k,l}$ respectively.

Figure 4 illustrates the StaR Maps generation with a road network example for two spatial queries (*distance* and *over*) that are queried from crowd-sourced data. A set of random maps is generated based on the UAM while considering a translational error model. For each location in the agent's navigation frame, the parameters μ (a) and σ (b) of a normal distribution, modeling the distance to the closest road, are computed. We obtain the probability of a chosen regulatory constraint being met from the cumulative distribution function of $\mathcal{N}(\mu, \sigma^2)$, or more generally via sampling. In this case, we chose to keep a distance of over 30 meters (c). Finally, (d) shows the probability of a location in the agent's navigation space being occupied by buildings.

IV. Experiments

A. Setup

We are employing real-world data that the OpenStreetMap (OSM) community has collectively annotated with semantics such as fine-granular road types, building aspects, and amenities. However, OSM does not contain uncertainty annotations at this moment. Hence, for this evaluation, we annotate all map features with the same synthetic uncertainty such that only a translation error $\vec{t}$ is generated from a Gaussian $\mathcal{N}(0, \text{diag}(10m, 10m))$ model. While this is a stark simplification of the real-world characteristic of such maps, e.g., their heterogeneous accuracy across features, this setup is sufficient to evaluate the creation and application of StaR Maps. The implementation employed for these experiments is available as OpenSource software alongside our Probabilistic Mission Design framework *https://github.com/HRI-EU/ProMis*.

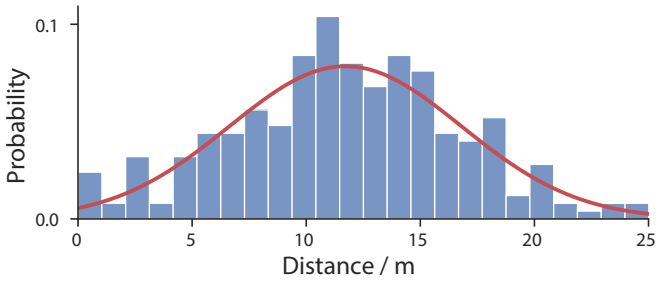


Fig. 5: **Parameter estimation of the distance relation:** Here, a histogram of the sampling process of the distance to the closest road is shown for a single point. From the set of samples we compute mean and standard deviation in order to parameterize the Gaussian that will model the distribution.

B. Creating StaR Maps

We demonstrate the computation and iterative refinement of StaR Maps. As a first step, let us take a look at sampling a spatial relation from a single point. To do so, we generate $N = 50$ maps from the OSM-based UAM as described in Section III and compute the distance to the respective closest road-type element. The result for a single location can be seen in Figure 5, where the set of samples and estimated distribution of the spatial relation are portrayed as histogram and Gaussian, respectively.

Now, let us explore how we can efficiently compute from the data contained in a Star Map, e.g., sets of sampled data in $G_{distance,road}$, containing mean and variance information about the *distance* relation as seen in Figure 4. As a first approach, we compute the parameters from a raster of samples and, as a naive choice for the function $v_{distance,road}$, apply linear interpolation to obtain the scalar field of this parameter. Figure 6 shows how the chosen grid of samples needs to

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021

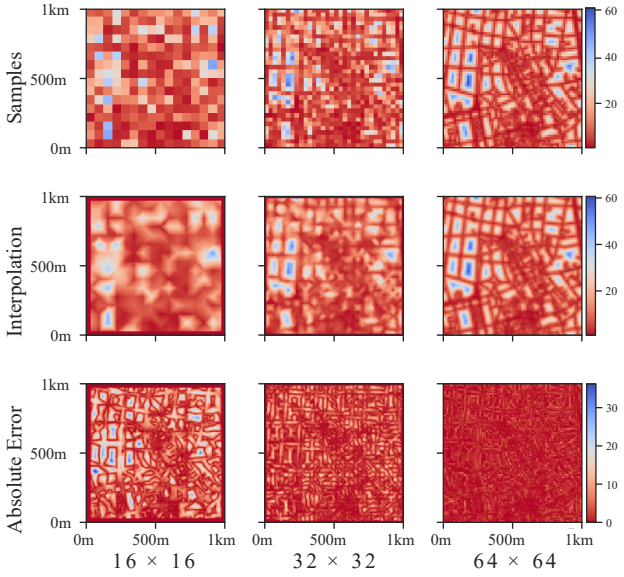


Fig. 6: **Interpolation of *distance* mean on raster of samples:** Distributing sampling points on a regular grid has the advantage of a uniform interpolation quality across the mapped space. A 512×512 raster of samples was taken as the reference image and compared to interpolated low-resolution data.

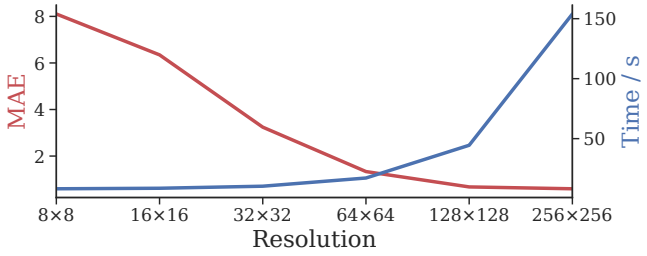


Fig. 7: **Interpolation-based StaR Map creation:** Analogously to Figure 6, we compare the mean absolute error (MAE) of various base resolutions when interpolated and compared to a 512×512 representation. While the error is reduced as expected, computation times grow quadratically due to the sample-grids growing resolution. 43

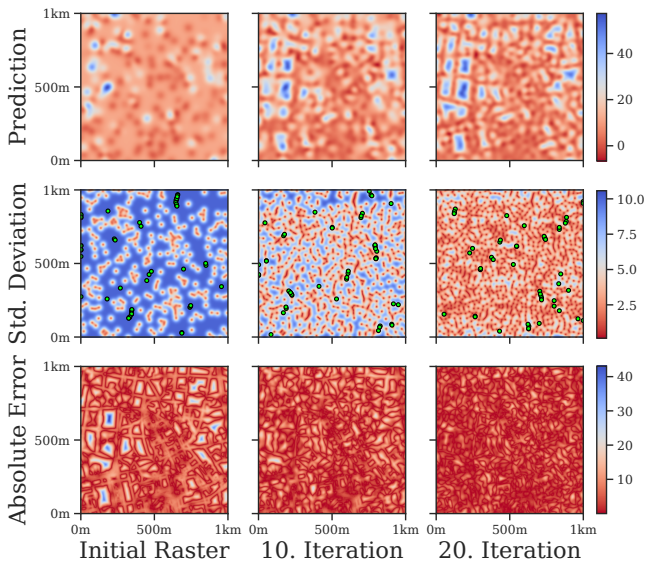


Fig. 8: **Gaussian Process (GP) prediction of *distance* mean using confidence guided refinement:** The GP's prediction capabilities allow us to obtain mean and standard deviation across the mapped area. Sampling at points of low confidence (green dots), we can incrementally refine the GP's training set to capture the spatial relation more accurately.

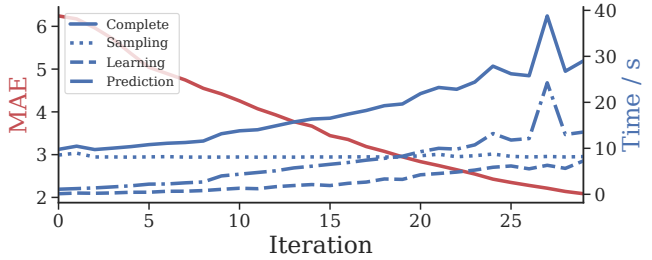


Fig. 9: **Gaussian Process (GP) based StaR Map creation:** In contrast to grid-based sampling as seen in Figure 7, the GP incrementally lowers the error while only taking on few more training points at the actually relevant locations. Although this approach has advantages, one has to note the necessity of tuning hyper-parameters for a reliable application.

be appropriately dense to match a high-resolution reference accurately. Of course, while this strategy can reliably reduce the error, the computational effort grows quadratically, as seen in Figure 7. Note that this approach is further limited to the confinements of the convex hull of the samples.

Instead, consider a more advanced choice for $v_{distance,road}$ by fitting a Gaussian Process (GP) on the samples to (i) predict across a continuous space and (ii) guide an incremental refinement based on the GP's confidence. To do so, we initialize the GP with a random sample of 256 points. Then, by predicting the same high-resolution reference of 512×512 points, we obtain the mean and standard deviation. By choosing points where the standard deviation is high, meaning low confidence of the GP based on his training data, we sample the spatial relation at those locations and add them to the training set. This way, the GP yields a higher accuracy and more sample-efficient result than linear interpolation. Further, considering the GP's confidence across the relevant space, one can decide when to stop the iterative refinements. Figure 8 shows an initial state as well as the prediction, confidence, and remaining error at different stages of refinement. Further, Figure 9 shows the comparatively low demand in time to gather samples and high prediction quality.

C. Reasoning on StaR Maps

Beyond representing the statistics of relations between points in space and the semantics within uncertain maps, StaR Maps facilitate a basis for spatial reasoning. Analogously to prior work [26], one can query spatial requests

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021

```
% Quantitative uncertainties, e.g., distance
distance(x, building) ~ normal(20, 0.5).

% Categorical uncertainties, e.g., over
0.9::over(x, primary).

% The queried space and its constraint
airspace(X) :- over(X, park).
airspace(X) :- distance(X, road) < 15,
    distance(X, pilot) < 250.
```

Listing 1: **StaR Maps facilitate spatial reasoning:** By translating the statistical, relational knowledge of the StaR Map into a hybrid probabilistic logic program, one can reason about spaces where all constraints are likely fulfilled. This example defines a simple airspace for an Unmanned Aerial Vehicle by restricting the valid airspace to be over park areas or close to primary roads and the remote pilot.

to StaR Maps as a statistical relational knowledge base by employing a hybrid probabilistic, first-order logic.

StaR Maps contain two types of knowledge, represented by categorical and continuous distributions, respectively. As basic running examples, making up a considerable proportion of public legislative language, we have shown *over* and *distance* as binary predicates between points in space and different environment semantics. To this end, we show how this vocabulary can be employed for spatial reasoning, i.e., deriving further knowledge beyond what is already encapsulated in the Star Map.

A hybrid probabilistic logic programming language such as presented by Nitti et al. [27] can be employed to represent both categorical and continuous distributions in a declarative modeling language. More critically, they allow for querying the probability of logical conjunctions and disjunctions within the model. To demonstrate the application of StaR Maps, we show such a program in Listing 1. For each chosen point $\vec{x}$ within the navigation space, the StaR Map data has been written out in the form of distributional clauses, e.g., for *distance* and *over* relations. One can further introduce domain knowledge to then query for application dependant probabilities. While in this example, a simple airspace is defined, Figure 10 shows a selection of more complex models being evaluated using StaR Map data of the respective locations.

V. Conclusion

We have presented StaR maps, a novel, hybrid probabilistic and relational environment representation. StaR Maps facilitate a probabilistic, relational worldview and enable complex reasoning about domain-specific queries in first-order logic. Further, StaR Maps improves the representation of spatial relations in contrast to prior work [26] by incorporating a probabilistic regression model on top of the sampling process, leading to confidence-guided refinements and additional information on the data's quality. We further demonstrated how reasoning on StaR Maps through spatial relations provides a foundation for building rich languages for tasks such as navigation or plan validation.

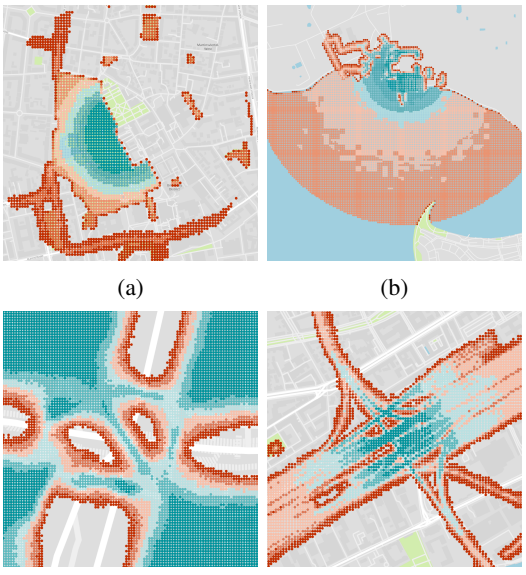


Fig. 10: **Reasoning on StaR Maps:** Through StaR Maps' vocabulary of probabilistic predicates over the navigation space, they facilitate complex semantic queries. Across a variety of environments, namely (a) an urban city park, (b) a bay area, (c) a major junction with a pedestrian crossing, and (d) a central train station for cargo and passenger transport.

Future work includes extensions of our spatial relations into many domains, embracing the diversity of multiple modes of mobility and their demands. While we have shown how to incorporate uncertainties connected to individual map elements and how to obtain the statistics of spatial relations efficiently, guiding the sampling process to intelligently select optimal places for computation and reducing redundancies in the process is essential.

A crucial component of our framework is modeling errors and spatial relations. One approach towards online identification of such parameters in the case of road segments has been demonstrated in [42]. However, further investigation into a more general and robust approach is important to gather uncertainty information on a wider range of geographic features. This task is vital to future map-making and maintenance to lift geographic representations into a probabilistic domain.

Further, it is vital to consider the underlying assumptions of a spatial relation's distribution. While we have utilized, e.g., Gaussian distributions in our presentation, they are not general tools to capture all data accurately. In future work, it will be essential to consider what models to employ and to build up a broader collection of spatial relations alongside them. A proper assessment of these choices' quality will be critical to guide safe and reliable decision-making.

Acknowledgments

The Eindhoven University of Technology authors received support from their Department of Mathematics and Computer Science and the Eindhoven Artificial Intelligence Sys-

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
DOI: 10.1109/ITSC58415.2024.10920021

tems Institute. Map data © OpenStreetMap contributors, licensed under the Open Database License (ODbL) and available from https://www.openstreetmap.org. Map styles © Mapbox, licensed under the Creative Commons Attribution 3.0 License (CC BY 3.0) and available from https://github.com/mapbox/mapbox-gl-styles.

REFERENCES

[1] G. J. Hunter, "Managing uncertainty in GIS," *Geographical information systems*, vol. 2, no. 3, pp. 633–641, 1999.

[2] L. De Raedt, K. Kersting, S. Natarajan, and D. Poole, *Statistical Relational Artificial Intelligence*, ser. Synthesis Lectures on Artificial Intelligence and Machine Learning. Cham: Springer International Publishing, 2016.

[3] T. Ando, W. Kugimiya, T. Hashimoto, F. Momiyama, K. Aoki, and K. Nakano, "Lateral Control in Precision Docking Using RTK-GNSS/INS and LiDAR for Localization," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 1, pp. 78–87, mar 2021.

[4] D. Dong, M. Wang, W. Chen, Z. Zeng, L. Song, Q. Zhang, M. Cai, Y. Cheng, and J. Lv, "Mitigation of multipath effect in GNSS short baseline positioning by the multipath hemispherical map," *Journal of Geodesy*, vol. 90, no. 3, pp. 255–262, mar 2016.

[5] J. Nilsson, J. Fredriksson, and A. C. E. Odblom, "Reliable Vehicle Pose Estimation Using Vision and a Single-Track Model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2630–2643, dec 2014.

[6] Y. Alkendi, L. Seneviratne, and Y. Zweiri, "State of the Art in Vision-Based Localization Techniques for Autonomous Navigation Systems," *IEEE Access*, vol. 9, pp. 76 847–76 874, 2021.

[7] M. Demir and K. Fujimura, "Robust Localization with Low-Mounted Multiple LiDARs in Urban Environments," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, oct 2019, pp. 3288–3293.

[8] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, vol. 2016-Augus, no. Iv. IEEE, jun 2016, pp. 864–870.

[9] O. Pink, "Visual map matching and localization using a global feature map," in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, jun 2008, pp. 1–7.

[10] A. Schindler, "Vehicle self-localization with high-precision digital maps," in *2013 IEEE Intelligent Vehicles Symposium Workshops*. IEEE, jun 2013, pp. 134–139.

[11] T. Caselitz, B. Steder, M. Ruhnke, and W. Burgard, "Monocular camera localization in 3D LiDAR maps," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, oct 2016, pp. 1926–1931.

[12] B. Flade, E. Casapietra, C. Goerick, and J. Eggert, "Behavior-based relative self-localization in intersection scenarios," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, oct 2017, pp. 1–8.

[13] Y. Lu, J. Huang, Y.-T. Chen, and B. Heisele, "Monocular localization in urban environments using road markings," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2017, pp. 468–474.

[14] B. Flade, M. Nieto, G. Velez, and J. Eggert, "Lane Detection Based Camera to Map Alignment Using Open-Source Map Data," in *21st IEEE International Conference on Intelligent Transportation Systems*, 2018, pp. 890–897.

[15] B. Flade, A. Koppert, G. Velez, A. Das, D. Betaille, G. Dubbelman, O. Otaegui, and J. Eggert, "Vision-Enhanced Low-Cost Localization in Crowdsourced Maps," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 3, pp. 70–80, 2020.

[16] S. Rezaei and R. Sengupta, "Kalman Filter-Based Integration of DGPS and Vehicle Sensors for Localization," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 6, pp. 1080–1088, nov 2007.

[17] A. L. Ballardini, D. Cattaneo, R. Izquierdo, I. Parra, M. A. Sotelo, and D. G. Sorrenti, "Ego-lane estimation by modeling lanes and sensor failures," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, oct 2017, pp. 1–7.

[18] G. Marra, S. Dumančić, R. Manhaeve, and L. De Raedt, "From statistical relational to neurosymbolic artificial intelligence: A survey," *Artificial Intelligence*, p. 104062, 2024.

[19] A. Colmerauer, "An introduction to prolog iii," *Communications of the ACM*, vol. 33, no. 7, pp. 69–90, 1990.

[20] K. Kersting and L. De Raedt, "Bayesian logic programming: Theory and tool," 2007.

[21] L. De Raedt, A. Kimmig, and H. Toivonen, "Problog: A probabilistic prolog and its application in link discovery." in *IJCAI*, vol. 7. Hyderabad, 2007, pp. 2462–2467.

[22] D. Fierens, G. Van den Broeck, J. Renkens, D. Shterionov, B. Gutmann, I. Thon, G. Janssens, and L. De Raedt, "Inference and learning in probabilistic logic programs using weighted boolean formulas," *Theory and Practice of Logic Programming*, vol. 15, no. 3, pp. 358–401, 2015.

[23] R. Manhaeve, S. Dumancic, A. Kimmig, T. Demeester, and L. De Raedt, "Deepproblog: Neural probabilistic logic programming," *advances in neural information processing systems*, vol. 31, 2018.

[24] Z. Yang, A. Ishay, and J. Lee, "Neurasp: Embracing neural networks into answer set programming," in *29th International Joint Conference on Artificial Intelligence (IJCAI 2020)*, 2020.

[25] A. Skryagin, D. Ochs, D. S. Dhami, and K. Kersting, "Scalable neural-probabilistic answer set programming," *Journal of Artificial Intelligence Research*, vol. 78, pp. 579–617, 2023.

[26] S. Kohaut, B. Flade, D. S. Dhami, J. Eggert, and K. Kersting, "Mission Design for Unmanned Aerial Vehicles using Hybrid Probabilistic Logic Programs," in *26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 1506–1513.

[27] D. Nitti, T. De Laet, and L. De Raedt, "Probabilistic logic programming for hybrid relational domains," *Machine Learning*, vol. 103, no. 3, pp. 407–449, 2016.

[28] N. Kumar, O. Kuželka, and L. De Raedt, "First-order context-specific likelihood weighting in hybrid probabilistic logic programs," *Journal of Artificial Intelligence Research*, vol. 77, pp. 683–735, 2023.

[29] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollar, and R. Girshick, "Segment Anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, oct 2023, pp. 4015–4026.

[30] F. Locatello, D. Weissenborn, T. Unterthiner, A. Mahendran, G. Heigold, J. Uszkoreit, A. Dosovitskiy, and T. Kipf, "Object-centric learning with slot attention," *Advances in neural information processing systems*, vol. 33, pp. 11 525–11 538, 2020.

[31] T. Kipf, G. F. Elsayed, A. Mahendran, A. Stone, S. Sabour, G. Heigold, R. Jonschkowski, A. Dosovitskiy, and K. Greff, "Conditional object-centric learning from video," in *International Conference on Learning Representations*, 2022.

[32] J. Chen and P. Dames, "Multi-class target tracking using the semantic phd filter," in *The International Symposium of Robotics Research*. Springer, 2019, pp. 526–541.

[33] M. Himmelsbach, F. v. Hundelshausen, and H.-J. Wuensche, "Fast segmentation of 3d point clouds for ground vehicles," in *2010 IEEE Intelligent Vehicles Symposium*. IEEE, 6 2010, pp. 560–565.

[34] M. Haklay and P. Weber, "Openstreetmap: User-generated street maps," *IEEE Pervasive Computing*, vol. 7, pp. 12–18, 10 2008.

[35] P. Bender, J. Ziegler, and C. Stiller, "Lanelets: Efficient map representation for autonomous driving," in *2014 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2014, pp. 420–425.

[36] J. Eggert, D. A. Salazar, T. Puphal, and B. Flade, "Driving situation analysis with relational local dynamic maps (r-ldm)," in *FAST-Zero Symposium*. Society of Automotive Engineers of Japan, 2017.

[37] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4372–4378, 2010.

[38] J. Huang, M. Demir, T. Lian, and K. Fujimura, "An Online Multi-lidar Dynamic Occupancy Mapping Method," in *IEEE Intelligent Vehicles Symposium (IV)*, no. Iv. IEEE, jun 2019, pp. 517–522.

[39] H. Lategahn and C. Stiller, "City gps using stereo vision," *2012 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2012*, pp. 1–6, 2012.

[40] Y. Wei, F. Mahnaz, O. Bulan, Y. Mengistu, S. Mahesh, and M. A. Losh, "Creating semantic hd maps from aerial imagery and aggregated vehicle telemetry for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 15 382–15 395, 9 2022.

[41] D. Betaille and R. Toledo-Moreo, "Creating Enhanced Maps for Lane-Level Vehicle Navigation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 786–798, dec 2010.

[42] B. Flade, S. Kohaut, and J. Eggert, "Error decomposition for hybrid localization systems," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 149–156.

©2024 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

DOI: 10.1109/ITSC58415.2024.10920021