Digit. Signal Process. 171 (2026) 105836

Contents lists available at ScienceDirect

Digital Signal Processing

journal homepage: www.elsevier.com/locate/dsp

Context and intent enhanced target tracking

Lubos Vaci [a], Marco Mari [b,*], Lauro Snidaro [b], Gian Luca Foresti [b]

[a] Emootion Ltd., Data Fusion Consulting, Sheffield, S6 4FZ, U.K.
[b] Department of Mathematics, Computer Science and Physics, University of Udine, via delle Scienze, 206, Udine, 33100, Italy

ARTICLE INFO

Keywords:
Tracking
Context
Machine learning
Decision support

ABSTRACT

Exploitation of contextual knowledge has recently emerged as a promising approach to increase the performance of Information Fusion systems. Despite pioneering efforts in context assisted target tracking, the realm is still in its infancy, as the frameworks combining context reasoning with target tracking are not abundant. We here postulate that, in addition to physical constraints, such as the road network, knowledge of common patterns that targets pursue can significantly improve tracking accuracy and continuity. In the presented approach, we address the problem of tracking ground targets in complex urban environments, which generally poses a challenge to modern airborne surveillance systems. A target's actions are modeled as a Markov chain with relevant context defining transition and emission probabilities. Target's kinematics are estimated by the Interacting Multiple Models (IMM) filter that estimates the mode transition probability matrix (TPM) at each recursion step. The TPM posterior is computed by a Quasi-Bayesian estimator conditioned on the prior and the likelihood originating from target's measurements and the context. Through extensive simulations, we demonstrate that incorporating contextual information into TPM estimation significantly improves the filtering performance compared to both the IMM filter with a fixed TPM and adaptive TPM estimation without considering contextual information.

1. Introduction

Enhancing tracking algorithms by exploiting Contextual Information (CI) of a target and its surrounding environment can significantly improve track accuracy and continuity [1]. However, not all context that surrounds the target of interest is beneficial and, if utilized incorrectly, it could even hinder system performance. Therefore, context assisted solutions have to be selective and use only relevant context that correctly helps in understanding the estimated situation [1]–[5].

The scenario visualized in Fig. 1(a) represents a vehicle tracking problem with an airborne sensing equipment.

The real world alternatives of such a scenario are likely to be affected by low visibility, complex dynamics of the targets, high clutter and target density, that further increase the tracking complexity [2], [6]. In Fig. 1(a), estimated $\hat{\boldsymbol{x}}_k$ and true location (of measurement origin) $\boldsymbol{y}_k$ of the target at time k is shown within the ellipse expressing an uncertainty of the target's location. In order to maintain track accuracy in these challenging conditions, that is being able to continuously associate the state $\hat{\boldsymbol{x}}_k$ from time $(k-1)$ up to $(k+4)$ with its measurement source sequence $\tilde{\boldsymbol{y}}_k$ a variety of tracking algorithms that exploit *a priori* knowledge of the environment, such as topographic background information and maps Fig. 1(b) that helps in reducing the uncertainty of target estimate $\hat{\mathbf{x}}_k$, has been proposed by the information fusion (IF) community [6–11] and our own work [12] and [13], respectively. At this level, CI is generally considered a constraining factor that affects the evolution of kinematic parameters of the target [14]. Knowledge of soft or hard constraints, that are classified from features, could be applied in Bayesian recursion either in the prior density function [15] or in the likelihood [9]. We will be referring to all these solutions as object/target tracking with CI exploitation at the object assessment level [1,2,16].

However, we are not only interested in detection and exploitation of spatial context but all available context such as spatio-temporal relationships, i.e., behavioral patterns of the target (Fig. 1(c)), and event context, i.e. features related to behavior patterns (Fig. 1(d)), that can exist between target and environment [3]–[5]. There are two main challenges that need to be addressed. First, tracking models have to be capable of representing, learning and inferring the complex spatio-temporal relationships occurring between objects and events. Second, the tracking algorithm has to be able to process the newly obtained CI during the estimation. In this work, we develop a unified framework for context exploitation that integrates Markov models with Interacting Multiple Model (IMM) filtering. By adaptively incorporating CI, the proposed framework enhances robustness and accuracy under uncertain and dynamic conditions, enabling rapid adaptation to target behavior and

* Corresponding author.
E-mail address: mari.marco@spes.uniud.it (M. Mari).

https://doi.org/10.1016/j.dsp.2025.105836

Available online 17 December 2025

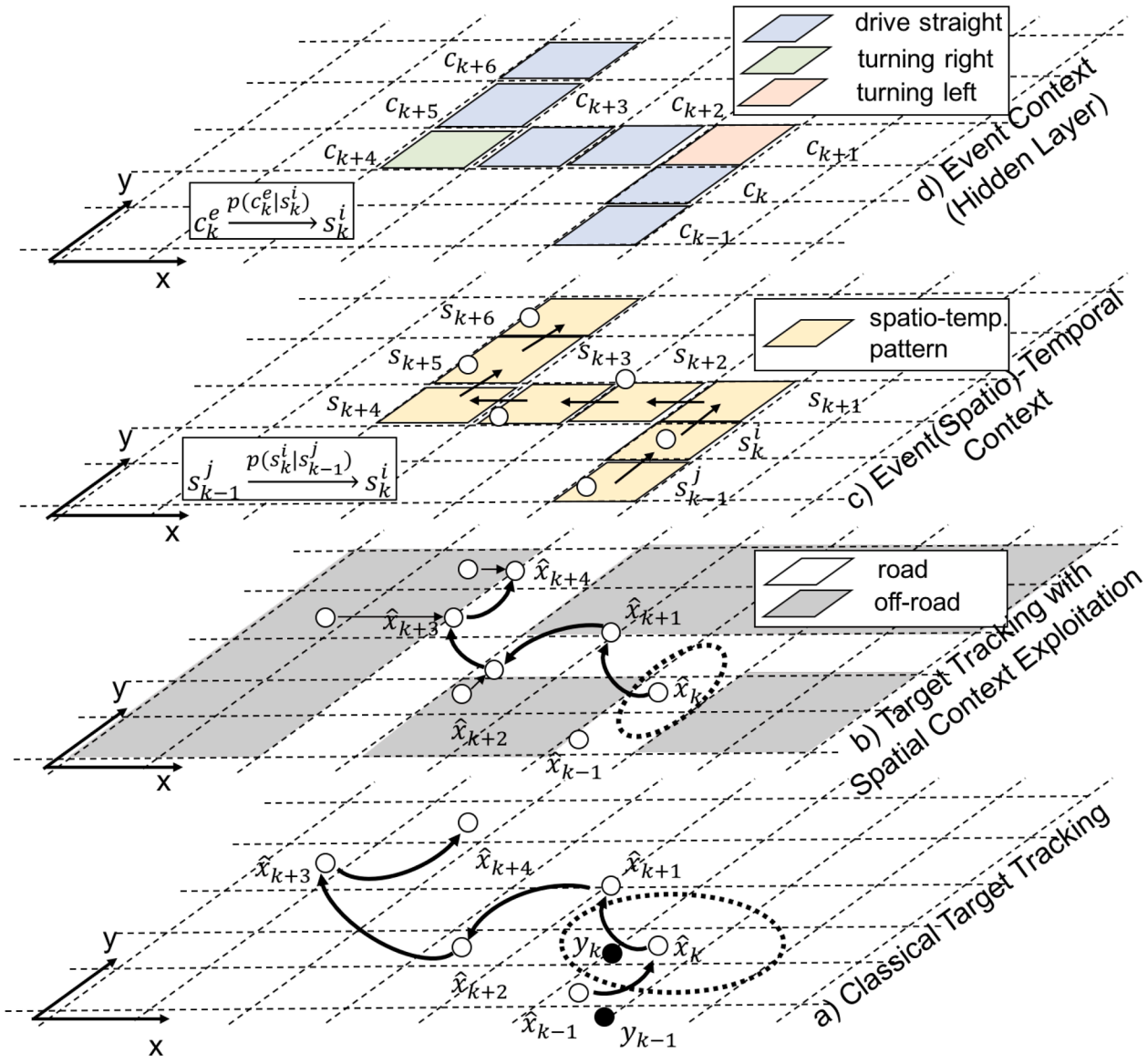

Fig. 1. Visualization of a tracking scenario, spatial, spatio-temporal and event context.

significantly improving tracking performance. Extensive simulation results confirm the effectiveness of the approach, demonstrating its capability to coherently integrate high-level reasoning with low-level fusion within a unified multi-layered architecture.

1.1. Framework for event context exploitation

The estimation of a target kinematics in the scenario depicted in Fig. 1(a) is referred to as a maneuvering target tracking problem under motion-mode and measurement-origin uncertainties (Section 2.1), best described by a semi-Markov process [17], [18]. In short, the semi-Markov process is uniquely defined by transition probabilities $\pi^{ji}_{k|k-1}$ that govern the evolution, i.e. sojourn time, of a mode state from m_{k-1} to m_k. In this work, mode state m_k at given time instance k could correspond to one of the nearly constant velocity models *"CV mode 1, 2"* or the coordinated turn models *"CT mode 1, …, 4"*. For these cases, sojourn-time dependent Markov (STDM) cooperating multiple model (CMM) algorithms are arguably, according to [17], the most suitable solutions. Among the CMM realizations, the IMM [19] are the most popular and reliable estimators. In IMM setting, the time behavior of the mode state m_k is modeled as a Markov chain with a fixed time-invariant TPM, which contradicts the definition of the semi-Markov process. Therefore, an extension of the IMM configuration to the case of STDM process was proposed in [20]. However, the STDM IMM assumes the knowledge of Transition Probability Matrix (TPM) to be *a priori* known before the real-time track can commence. As a solution to this shortcoming, it was proposed to augment the IMM recursion by a TPM estimation step. This modification allows for recursive update of mode dependent conditional means $\mathrm{E}[\mathbf{x}_k|m_{k-1}, \mathbf{x}_{k-1}, \mathbf{y}_{k-1}]$, covariances $\mathrm{Cov}[\mathbf{x}_k|m_{k-1}, \mathbf{x}_{k-1}, \mathbf{y}_{k-1}]$ and mode transition probabilities $p(m_k|m_{k-1})$ from prior values. The idea of estimating the TPM recursively has been proposed by Jilkov et al. [21] and Orguner et al. [22] in terms of the multiple models' probabilities and likelihoods.

In this work, we show that the IMM recursion can be further refined by inclusion of context into mode dependent state estimate $p(\mathbf{x}_k|m_{k-1}, \mathbf{x}_{k-1}, \mathbf{y}_{k-1}, \mathbf{c}_{k-1})$ and into TPM estimation process. In the former case, based on map constraints (Fig. 1(b)) and target patterns (Fig. 1(c)), CI makes certain modes less likely while others more dominant. In the latter case, the knowledge of CI, spatio-temporal (Fig. 1(c)) and event context (Fig. 1(d)), reduces the ambiguities occurring between mode transitions from m_{k-1} to m_k and helps compensating for the lack of sojourn time knowledge in the TPM.

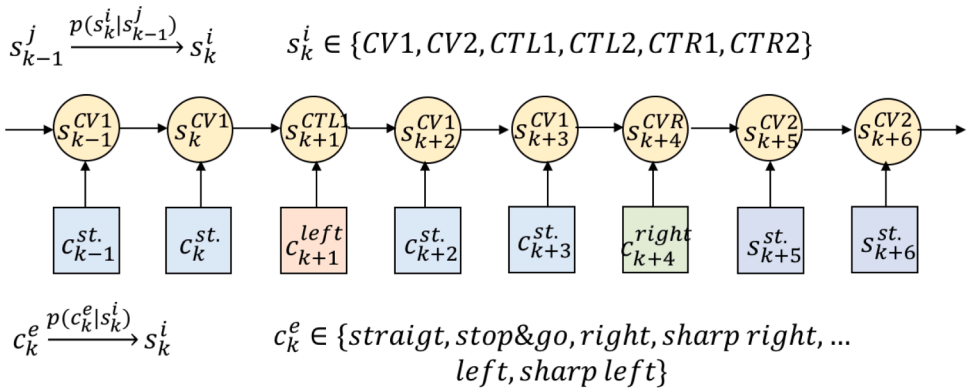

Fig. 2. Structure of the event sequence in Fig. 1(c and d) represented as HMM.

By considering the tracking scenario depicted in Fig. 1, we can describe causalities between target modes and events (Fig. 1(d)), that are considered to be optimal for a particular event, as a Hidden Markov Model (HMM) (Fig. 2). Individual nodes, modal s^i_k or eventual c^e_k, are combined together in space and time in order to form sequences $\mathbf{s}_{1:K} = [s^i_1, s^i_2, \ldots, s^i_K]^\top$ and $\mathbf{c}_{1:K} = [c^e_1, c^e_2, \ldots, c^e_K]^\top$. When the vehicle reaches an intersection it can perform, in accordance with the HMM model, one of the three types of maneuvers c^e_k, e.g. turning *"right"*, turning *"left"* and continuing *"straight"*, which could be either of *"high"* or *"low"* kinematics. From the tracking algorithm perspective, these maneuvers are perceived as a context that is associated to a certain modal state s^i_k, i.e. the nearly constant velocity models *"CV mode 1, 2"* and the coordinated turn models *"CT mode 1, …, 4"*, expressed by a likelihood $p(c^e_k|s^i_k)$.

By denoting the mode state of an IMM estimator by m^i_k and mode sequence of a Markov chain by s^i_k, we distinguish between Gaussian Model (GM) and HMM. Despite these two models are both belonging to the same family of Bayesian estimators, they are fundamentally different and their estimated mode probabilities $p(m^i_k|m^j_{k-1})$ and $p(s^i_k|s^j_{k-1})$ are not necessarily equal. The former refers to the kinematic evolution of the target state, while the latter reflects solely on the mode evolution of the target that constrains the former.

Assuming that the target follows a certain pattern, the future actions of a target could be anticipated by predicting that a certain goal is being pursued by the target. All possible sequences in the tracking scenario are referred to as $\mathcal{S}_k$ and $\mathcal{C}_k$ where $\mathcal{S}_k = \{\mathbf{s}^1_{1:k}, \ldots, \mathbf{s}^{N_s}_{1:k}\}$ and $\mathcal{C}_k = \{\mathbf{c}^1_{1:k}, \ldots, \mathbf{c}^{N_e}_{1:k}\}$ representing the set of all the N_s and N_e sequences up to time k of the states and events, respectively. Each sequence $\mathbf{s}^{n_s}_{1:k}$ can be associated with a certain objective that individual target or group seeks to reach in a road network represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ as discussed in our previous work [13]. The most probable hidden state sequence $\hat{\mathbf{s}}_{1:k}$ of all possible sequences $\mathcal{S}_k$, subject to $\mathcal{G}(\mathcal{V}, \mathcal{E})$, is the one which maximizes the contextual observation likelihood $p(\mathcal{C}_k)$, that is the maximum of the loss function $\hat{\mathbf{s}}_{1:k} = \arg\max_{\mathcal{S}_k} p(\mathcal{C}_k|\boldsymbol{\theta}_{k|k-1})$. The belief that a vehicle follows the sequence $\hat{\mathbf{s}}_{1:k}$ towards its objective is used in aiding the mode selection process of the IMM filter and for adaptive estimation of the TPM.

1.2. Related work

Recent studies have demonstrated the value of incorporating CI to enhance object tracking performance across various domains and environments [2]. In ground vehicle tracking, contextual constraints such as road geometry have been leveraged to improve estimation accuracy and robustness [15,23,24], particularly in complex urban environments or during abrupt maneuvers. For example, spatial context can mitigate occlusion effects in Doppler blind zones [6], while road-aligned motion models improve trajectory estimation by aligning predictions with known infrastructure [7]. Similarly, digital terrain elevation models can serve as CI to provide constraints on ground target location and velocity [25].

CI obtained from naval maps, e.g., water depth and ship paths, was used for abnormal behaviour detection and tracking in [8]. In [11], the authors investigated the possibility of combining context-driven decision making for maneuvering object detection and tracking. By knowing the spatial distribution of a target, the authors augmented sensor detection probability from meteorological and oceanographic products (context) in order to improve feature detection and classification.

Mode estimations obtained by classifiers from sensor measurements are, in a Bayesian sense, related to the states and target modes via a joint posterior distribution [26]. A model for image-based observation for discrete target mode estimation was presented in [27]. According to [28], the optimal filter capable of coping with such a hybrid state estimation is computationally prohibitive. Therefore, the authors proposed the image-enhanced IMM filter as a practical solution for effective reduction of the number of components in the Gaussian mixture. However, the exploitation of CI in the field of target tracking is still rather limited to low-level information fusion, i.e., object assessment [1–5].

The role of CI in probabilistic inference is twofold. First, it creates additional links between events and entities. Secondly, it adjusts the confidence that an entity is following a certain plan [29].

The capabilities of HMMs were demonstrated in [30], in which the authors proposed an IMM framework for intelligent vehicle trajectory prediction that combines a vehicle physics model for accurate short-term forecasting with an HMM for maneuver recognition, enabling probabilistic reasoning about future driving behaviors and improving long-term prediction under uncertainty. By integrating these models, the IMM approach enhances robustness to noise and model errors across varying driving maneuvers.

The causality of relationships, modeled by factor graphs [31], could also be expressed by multiple hierarchical levels of Bayesian Networks (BN) and at different time slices, i.e., dynamic Bayesian networks (DBN) and their equivalents. Recently, a hidden semi-Markov model was proposed by [32] to track the change of targets' interests. This model allows for capturing the different durations of target stays in a latent state, which can better model the heterogeneity of target drifting intentions. Frameworks for event context exploitation are the subject of intensive research, as shown in [33]. For instance, the synergy between tracking methods and semantic modeling that can enhance object labeling, scene understanding, and awareness was proposed in [34]. The framework exploits semantic Web technologies to extend the assertional knowledge extracted by video analysis. In the works of Bashar et al. ([3]–[5]), an intent prediction framework based on learning patterns-of-life has been proposed for Bayesian tracking. The so called bridging framework capitalizes on the premise that the path of the object must end at the intended destination. Since the endpoint is unknown *a priori*, a bridge for each possible destination must be constructed based on premeditated actions guided by intent. The authors are utilizing the likelihood of

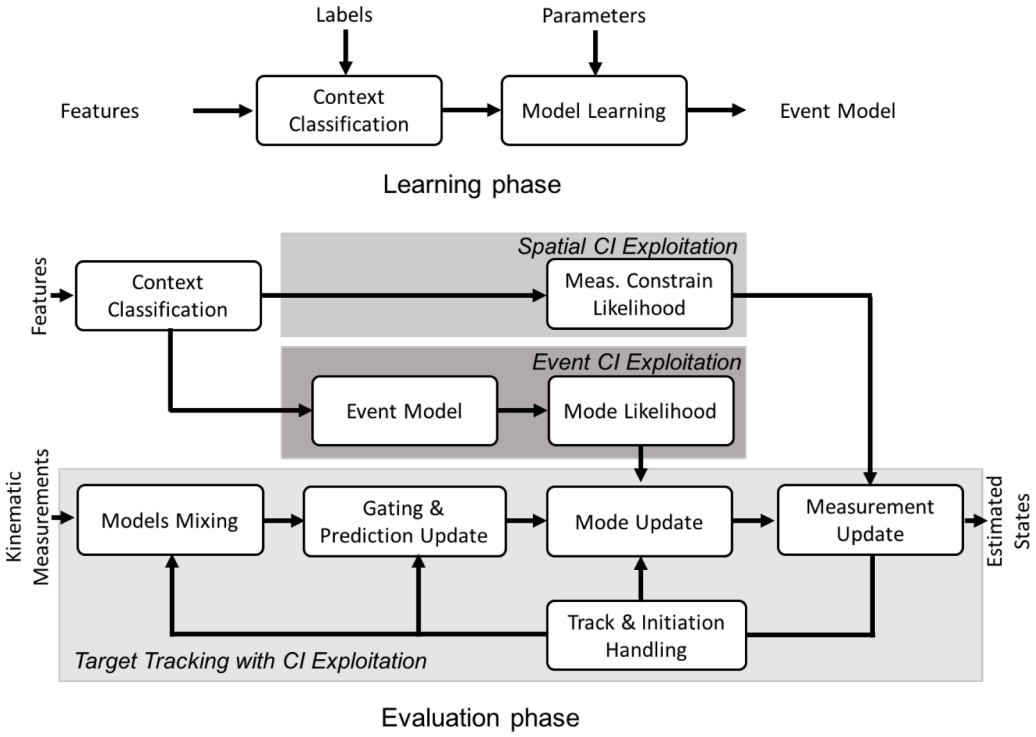

Fig. 3. Target tracking with CI exploitation.

observing partial tracks from a particular bridge to aid the target tracking process. While our work also builds on the premise of predicting target intents by Markov networks, the utilization of this newly obtained knowledge for the benefit of object tracking is fundamentally different. We are altering the prediction step of the filter as opposed to applying hard constraints on the measurement likelihood given the hypothesis. We argue that applying hard constraints makes tracking filters more susceptible to intent prediction errors. For this reason, we amend the likelihood of the transition probabilities between tracking modes as opposed to making hard decisions based on intent.

To the best of our knowledge, BNs were for the first time proposed to aid object tracking by Hautaniemi and Saarinen [35]. They enhanced the classic Multiple Target Tracking (MTT) algorithm, based on IMM filter and Probabilistic Data Association (PDA), with quantities other than kinematic measurements, i.e., context.

In [36], a DBN with hybrid discrete-continuous variables was used to model ocean animal behavior and adapt the bandwidth of a multiple model bootstrap filter based on the most probable motion modes. Similarly, in [37,38], a Bayesian filtering framework combining particle filters and BN was proposed to integrate uncertain context information, such as mobility constraints, sensor conditions, and target behavior, thereby improving tracking precision in complex scenarios like wildlife monitoring.

Schubert and Wanielik [39] argued that context in [35] does not influence the tracking directly, or in the case of [36], that such a tracking procedure is computationally very expensive. Therefore, the authors proposed their approach for incorporating additional information into the IMM called the Meta Model filter. With a structure similar to a BN, the Meta Model represents causality of events in the form of an adaptive TPM. Likewise, in [40,41], a context-based Bayesian network is suggested as a way to substitute sojourn time-dependent transition probabilities for the fixed ones in the standard model.

The TPM is a design parameter whose choice can significantly influence the IMM estimation process. Therefore, algorithms which can identify the TPM during tracking have already been studied in literature [22,42–46], and more recently [47–49]. Furthermore, variable structure algorithms, referred to as Expected-Mode Augmentation (EMA), for multiple-model estimation have been developed [50]. In the EMA approach, the original model set is augmented by a variable set of models intended to match the expected value of the unknown true mode. An online maximum likelihood estimator for the transition probabilities associated with a Jump Markov Linear Process (JMLP) was proposed by Orguner et al. [22], leveraging the reference probability method to derive recursive expressions for otherwise intractable expectations. This estimator employs the Expectation Maximization (EM) algorithm to maximize the likelihood of the TPM. More recently, these models have been constructed adaptively in real time as probabilistically weighted sums over the mode set, as demonstrated in [44] and [46].

Alternatively, variational inference has been successfully used to derive adaptive TPM for Jump Markov Process (JMP) [51,52]. Along with the consolidated improvements made in data-driven solutions, Neural Networks (NNs) have also been used extensively in the field of Object Tracking [53], where deep learning-based approaches are currently the state-of-the-art in extracting meaningful information from high-dimensional data [54]. Although significant progress has been made in Visual Object Tracking, the problem becomes considerably more challenging when dealing with non-visual data. Unlike visual inputs, non-visual modalities often lack rich, high-dimensional information, making it substantially more difficult to extract meaningful features for accurate target characterization and tracking. However, as NNs have demonstrated impressive results in the analysis and predictions of time series data [55], they have been successfully applied for state estimation problems [56,57]. As a consequence, some recent approaches have integrated data-driven methodologies to adaptively estimate TPM in IMM estimator [58], while some others have completely replaced the need for its recursive identification by estimating the target motion mode [59–61].

While NN-based approaches are widely adopted for extracting and exploiting CI from high-dimensional sensory inputs, such as images and videos, for accurate trajectory prediction [62], their application remains limited when CI must be inferred from low-level sensor data, where embedding situational awareness at the object assessment level is

significantly more challenging. Moreover, even in high-dimensional domains, these approaches predominantly focus on spatial context, often overlooking higher-level, event-based reasoning that is critical for robust situational awareness and long-term predictive accuracy.

Building on the insights discussed above, we propose a novel context-aware architecture (Fig. 3) that combines probabilistic reasoning techniques [39], [63] with conventional target tracking. While all works mentioned in the above short survey have reported the usefulness of event models in support of tracking, most approaches exploit only spatial context, which represents only a small portion of the available background information. Our framework simultaneously leverages both spatial and situational (event-based) context: spatial features are directly embedded in the tracking process, whereas the semantic meaning inferred from event reasoning is incorporated during the mode probability update of the IMM filter. By modeling the relationships among contextual observations, sensor measurements, and state variables as conditional dependencies, our method eliminates the need to construct joint probability distributions in Bayes recursion [26,28]. Unlike approaches relying on DBNs with pre-defined TPMs [35,36,39], our approach treats transition probabilities as parameters that are recursively estimated within the IMM framework. Building upon [21,22,44,46], we further extend TPM estimation to explicitly account for CI, allowing the filter to dynamically adapt to evolving situational cues. Through the adaptive incorporation of CI, the proposed framework improves tracking performance under uncertainty and varying situational conditions. These advantages are supported by extensive simulation results that validate the effectiveness of the method.

1.3. Main goals and contributions

We have introduced and evaluated the framework for event context inference and exploitation (Fig. 3) for the purpose of ground objects tracking in [13]. This manuscript extends and improves on our initial work for the following cases:

- In our initial work [13], the event context was treated as a label that was inferred on the basis of event patterns at each road intersection. The significance of this variable was modeled as a relevance function that exponentially depended on the target's spatial distance to the next on road intersection. This modeling detection relied on *a priori* knowledge of the context relevance functions given the street map layout and structure. In this work, the event context influences the spatio-temporal context that relates to the mode of the tracking filter as a dissimilarity measure between mode state and context sequences, thus without requiring a relevance function that would be very difficult to model in practice.
- Spatio-temporal and event context are used not only for recursive estimation of the mode transition probabilities, as shown in [13], but also to constrain the mode probability update step of the IMM filter. This helps the filter react to sudden trajectory changes, which increases the overall track accuracy and continuity.
- In order to improve clarity, derivation of the context-aware adaptive TPM (Section 2.2) utilized in the IMM filtering algorithm (Section 2.1) was extended and reformulated.
- The experimental part was expanded with two new scenarios (Section 3.2). We compared the proposed methodology when the system kinematics is modeled as scalar jump Markov linear system, jump Markov linear system, and semi-Markov process. We have conducted over 600 Monte Carlo experiments in order to justify the benefits of the proposed context-aware architecture for target tracking.
- Furthermore, we have also conducted the convergence analysis (Fig. 7) where the accuracy of individual elements of the TPM are compared against the reference TPM. By doing so, we showed that the proposed solution is feasible for online tracking.

The remainder of the paper is organized as follows: The design of the proposed Bayesian framework for mode-dependent context observations is detailed in Section 2, where situational context is inferred using an HMM [13] and formulated into contextual constraint factors (Section 2.4). Sections 2.2 and 2.3 describe the adaptive estimation of the TPM within the IMM filter. Section 3.1 presents the evaluation scenarios, and Section 3 discusses the experimental results. Finally, Section 4 summarizes the main findings and concludes the paper.

2. Model formulation

In this section, we first introduce a Bayesian recursion for the maneuvering target tracking problem that allows exploitation of contextual information as a mode dependent variable. Differently from the state of the art (Section 1.2), a mode state of the tracking filter and its evolution is conditioned not only on received (physical) observations but also based on the context. In order to accommodate CI in the tracking process, the methodology for adaptive mode transition computation is introduced in Sections 2.2 and 2.3 while the formalism of contextual observations is presented in Section 2.4.

2.1. Context enhanced maneuvering target tracking

Suppose a maneuvering target is moving according to a kinematic model defined by a discrete JMP

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_k(m_k^i)\mathbf{x}_{k-1} + \mathbf{g}_k(m_k^i)\mathbf{v}_k \quad \text{or} \quad p(\mathbf{x}_k, m_k^i|\mathbf{x}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}_k(m_k^i)\mathbf{x}_k + \mathbf{w}_k \quad \text{or} \quad p(\mathbf{y}_k|m_k^i, \mathbf{x}_k) \\ c_k^j &= \boldsymbol{\phi}_k(m_k^i) \quad \text{or} \quad p(c_k^j|m_k^i) \\ m_k^i &= \boldsymbol{\pi}_{k|k-1}^{ji} m_{k-1}^j \quad \text{or} \quad p(m_k^i|m_{k-1}^j). \end{aligned} \tag{1}$$

In the above equations, $\mathbf{x}_k$ is the state vector while $\mathbf{y}_k$ and $\mathbf{c}_k$ denotes sensor and context observations, respectively. Parameter m_k represents a mode state of the target at time k that can take any of the modes $1, 2, \dots, N_m$. Superscripts, i and j indicate the mode state of the filter in time m_k^i and m_{k-1}^j, and they can also indicate discrepancy between observed m_k^i and estimated mode m_k^j at the time k, respectively. Functions $\mathbf{f}(m_k^i)$, $\mathbf{g}(m_k^i)$ and $\mathbf{h}(m_k^i)$ are mode dependent state transitions. Vectors $\mathbf{v}_k \sim \mathcal{N}(\mathbf{v}_k; 0, \mathbf{Q})$ and $\mathbf{w}_k \sim \mathcal{N}(\mathbf{w}_k; 0, \mathbf{R})$ are mutually uncorrelated Gaussian distributed white noises with covariances $\mathbf{Q}$ and $\mathbf{R}$, respectively.

In this work, contextual information $\mathbf{c}_k$ is a constrain factor on mode state m_k^i and it is represented as a discrete mode dependent point process $\boldsymbol{\phi}(m_k^i)$ for $k = 1, 2, \dots, N_k$ [27] where each c_k^j indicate the inferred mode m_k^i of the target $\mathbf{x}_k$. We allow $c_k^j = 0$ to indicate situations where contextual observation does not provide any useful information about the target. We assume that CI, expressed as an observation likelihood $p(c_k^j|m_k^i)$ that consists of a non-zero mode rate estimate with probability λ_i and mode discernibility with probability d^{ij} for $i, j \in \{1, 2, \dots, N_m\}$, or

$$p(c_k^j|m_k^i) = \begin{cases} 1 - \lambda_k^i & \text{if} \quad c_k^i = 0, \\ d_k^{ij}\lambda_k^i & \text{otherwise.} \end{cases} \tag{2}$$

Since the first appearance of the target, a total number of N_m different mode histories $m_{1:k}$ have been observed until time k. The history of the ith mode state $m_{1:k}^i$, $i \in \{1, 2, \dots, N_m\}$ can be modeled as a homogeneous Markov chain with a TPM $\boldsymbol{\pi}_{k|k-1}^{ji} = p(m_k^i|m_{k-1}^j)$. By assuming a JMP setting defined by (1), the goal of a Bayesian recursive estimation (3–9) is to infer on the target's posterior $p(\mathbf{x}_k|m_k^i, \mathbf{y}_{1:k}, \mathbf{c}_{1:k})$ from observations $\mathbf{y}_{1:k}$, $\mathbf{c}_{1:k}$ and all possible mode sequences $m_k \in \{1, 2, ., N_m\}$ that end in the mode m_k^i.

$$p(m_{k-1}^j|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) \xrightarrow{\text{Mixing}} p(m_k^i|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) \tag{3}$$

$$p(\mathbf{x}_{k-1}|m_{k-1}^j, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) \xrightarrow{\text{Mixing}} p(\mathbf{x}_{k-1}|m_k^i, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) \tag{4}$$

$$p(\mathbf{x}_{k-1}|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \xrightarrow[\text{Update}]{\text{Prediction}} p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \tag{5}$$

$$p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \xrightarrow[\text{Mode Update}]{\text{Measurment}} p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k-1}) \tag{6}$$

$$p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k-1}) \xrightarrow[\text{Mode Update}]{\text{Context}} p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k}) \tag{7}$$

$$p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \xrightarrow[\text{Update}]{\text{Measurement}} p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k-1}) \tag{8}$$

$$p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k-1}) \xrightarrow[\text{Update}]{\text{Context}} p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k}) \tag{9}$$

Given the mode sequence $m_{1:k}^i$, the resulting distribution is the product of posteriors over mode $p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k})$ (6) and system state $p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k})$ (9) is by the law of total probability equal to

$$p(\mathbf{x}_k|\mathbf{y}_{1:k},\mathbf{c}_{1:k}) = \sum_{i=1}^{N_m} p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k})p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k}), \tag{10}$$

i.e. the desired output.

In accordance to the JMP definition (1), contextual observations $\mathbf{c}_{1:k}$ in (9) contain no information about the state vector $\mathbf{x}_k$, thus the context update step can be, in this case, safely omitted. Measurement mode (6), context mode (7) and measurement updates (8) are computed from their prior distributions $p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$, $p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ and measurements $\mathbf{y}_{1:k}$, $\mathbf{c}_{1:k}$ by Bayes' theorem under the assumption of 1st order Markov property which result in (11) and (12), respectively.

$$\begin{aligned} &p(m_k^i|\mathbf{y}_{1:k},\mathbf{c}_{1:k}) = \\ &= \frac{p(\mathbf{y}_k|m_k^i,\mathbf{c}_k)p(\mathbf{c}_k|m_k^i)p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1},\mathbf{c}_k)p(\mathbf{c}_k|\mathbf{c}_{1:k-1})} \end{aligned} \tag{11}$$

$$\begin{aligned} &p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k}) = \\ &= \frac{p(\mathbf{y}_k|m_k^i,\mathbf{x}_k,\mathbf{c}_k)p(\mathbf{c}_k|m_k^i,\mathbf{x}_k)p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(\mathbf{y}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_k)p(\mathbf{c}_k|m_k^i,\mathbf{c}_{1:k-1})} \end{aligned} \tag{12}$$

Note that densities $p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k})$ and $p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k},\mathbf{c}_{1:k-1})$ in (12) are equal since $\mathbf{c}_k = \boldsymbol{\phi}_k(m_k^i)$ does not dependent on $\mathbf{x}_k$. For the same reason the following equality $p(\mathbf{c}_k|m_k^i,\mathbf{x}_k) = p(\mathbf{c}_k|m_k^i)$ holds.

The evolution of the state density (5) from $p(\mathbf{x}_{k-1}|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ to $p(\mathbf{x}_k|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ is calculated by the Chapman-Kolmogorov equation.

The evolution of a mode state (4), in view of the total probability theorem, can be expressed as

$$\begin{aligned} &p(\mathbf{x}_{k-1}|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) = \\ &\sum_{j=1}^{N_m} p(m_k^i|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \times p(\mathbf{x}_{k-1}|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}), \end{aligned} \tag{13}$$

where

$$\begin{aligned} &p(m_k^i|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) = \\ &\frac{p(m_k^i|m_{k-1}^j)p(m_{k-1}^j|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}))}. \end{aligned} \tag{14}$$

Evolution of density from $p(m_{k-1}^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ to $p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ in (3) is performed by evaluating the Chapman-Kolmogorov integral for Markov chain or so called mixing (15):

$$p(m_k^i|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) = \sum_{j=1}^{N_m} p(m_k^i|m_{k-1}^j)p(m_{k-1}^j|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \tag{15}$$

Probability $p(\mathbf{x}_{k-1}|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ is known *a priori* and is normally distributed i.e.

$$p(\mathbf{x}_{k-1}|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) = \mathcal{N}(\mathbf{x}_{k-1};\hat{\mathbf{x}}_{k-1|k-1}^i,\boldsymbol{\Sigma}_{k-1|k-1}^i). \tag{16}$$

Substituting (16) into (13) and evaluating the whole recursion (3–8) yields a Gaussian mixture $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ in (10) with N_m^2 components. In order to avoid the exponential growth of Gaussian mixture components, numerous techniques based on either pruning or merging were proposed [17]. Among them, the IMM filter introduced by Blom and Bar-Shalom [19] is arguably the most popular solution. IMM approximates posterior mode probabilities $p(\mathbf{x}_{k-1}|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})$ in (13) with

$$\begin{aligned} &p(\mathbf{x}_{k-1}|m_k^i,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) = \\ &= \sum_{j=1}^{N_m} p(m_k^i|m_{k-1}^j,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1}) \times \mathcal{N}(\mathbf{x}_{k-1};\hat{\mathbf{x}}_{k-1|k-1}^i,\boldsymbol{\Sigma}_{k-1|k-1}^i), \\ &\approx \mathcal{N}(\mathbf{x}_{k-1};\hat{\mathbf{x}}_{k-1|k-1}^{0i},\boldsymbol{\Sigma}_{k-1|k-1}^{0i}), \end{aligned} \tag{17}$$

where the merged mean $\hat{\mathbf{x}}_{k-1|k-1}^{0i}$ and covariance $\boldsymbol{\Sigma}_{k-1|k-1}^{0i}$ are obtained by moment matching as

$$\begin{aligned} \hat{\mathbf{x}}_{k-1|k-1}^{0i} &= \sum_{j=1}^{N_m} p(m_{k-1}^i|m_{k-1}^j)\hat{\mathbf{x}}_{k|k}^j, \\ \boldsymbol{\Sigma}_{k|k}^{0i} &= \sum_{j=1}^{N_m} p(m_{k-1}^i|m_{k-1}^j)[\boldsymbol{\Sigma}_{k-1|k-1}^j + \\ &\quad + (\hat{\mathbf{x}}_{k-1|k-1}^j - \hat{\mathbf{x}}_{k-1|k-1}^{i0})(\hat{\mathbf{x}}_{k-1|k-1}^j - \hat{\mathbf{x}}_{k-1|k-1}^{i0})^T]. \end{aligned} \tag{18}$$

Furthermore, the mode transition probability $\boldsymbol{\pi}_{k|k-1}^{ji} = p(m_k^i|m_{k-1}^j)$ in (15) is assumed to be known *a priori* and to be a constant time invariant matrix.

With such an approximation the overall posterior mean $\hat{x}_{k|k}$ and covariance $\Sigma_{k|k}$ become:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \sum_{i=1}^{N_m} \boldsymbol{\mu}_k^i \hat{\mathbf{x}}_{k|k}^i, \\ \boldsymbol{\Sigma}_{k|k} &= \sum_{i=1}^{N_m} \boldsymbol{\mu}_k^i [\boldsymbol{\Sigma}_{k|k}^i + (\hat{\boldsymbol{x}}_{k|k}^i - \hat{\mathbf{x}}_{k|k})(\hat{\boldsymbol{x}}_{k|k}^i - \hat{\mathbf{x}}_{k|k})^T], \end{aligned} \tag{19}$$

respectively. The mode conditional means $\{\hat{x}_{k|k}^i\}_{i=1}^{N_m}$, covariances $\{\Sigma_{k|k}^i\}_{i=1}^{N_m}$ and mode probabilities $\{\mu_k^i\}_{i=1}^{N_m} = \{p(m_k^i|\mathbf{y}_{1:k})\}_{i=1}^{N_m}$ must be calculated recursively from their previous values $\{\hat{x}_{k-1|k-1}^i, \Sigma_{k-1|k-1}^i, \mu_{k-1}^i\}_{i=1}^{N_m}$.

2.2. Context adaptive transition probability matrix

In this work, the mode TPM $\boldsymbol{\Pi}_k = [\pi_{k|k-1}^{ji}]_{N_m \times N_m}$ or $\pi_{k|k-1}^{ji} = p(m_k^i|m_{k-1}^j)$ in (15) is not fixed but is adapting to observations, formulated as measurement and context likelihood functions, recursively. Differently from [11,21,22], we exploit CI during the mode mixing step (3) of the IMM filter (Section 2.1) as follows:

$$\begin{aligned} &p(\boldsymbol{\Pi}_k|\mathbf{y}_k,\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1}) \\ &= \frac{p(\boldsymbol{\Pi}_k,\mathbf{y}_k,\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1})}{p(\mathbf{y}_k,\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1})} \\ &= \frac{p(\mathbf{y}_k|\boldsymbol{\Pi}_k,\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1})p(\mathbf{c}_k|\boldsymbol{\Pi}_k,\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1})p(\mathbf{c}_k|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})} \\ &\quad \frac{p(\boldsymbol{\Pi}_k|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})p(\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}, \end{aligned} \tag{20}$$

We seek to further simplify the TPM posterior $p(\boldsymbol{\Pi}_k|\mathbf{y}_k,\mathbf{y}_{k-1},\mathbf{c}_k,\mathbf{c}_{k-1})$ in (20) by assuming a conditional independence between contextual c_k and measurement y_k observations, which leads to (21)

$$\begin{aligned} &p(\boldsymbol{\Pi}_k|\mathbf{y}_k,\mathbf{y}_{1:k-1},\mathbf{c}_k,\mathbf{c}_{1:k-1}) \\ &= \frac{p(\mathbf{y}_k|\boldsymbol{\Pi}_k,\mathbf{y}_{1:k-1})\ p(\mathbf{c}_k|\boldsymbol{\Pi}_k,\mathbf{c}_{1:k-1})\ p(\boldsymbol{\Pi}_k|\mathbf{y}_{1:k-1},\mathbf{c}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})\ p(\mathbf{c}_k|\mathbf{c}_{1:k-1})}. \end{aligned} \tag{21}$$

Here, probability $p(\mathbf{y}_k|\boldsymbol{\Pi}_k,\mathbf{y}_{1:k-1})$ denotes a measurement likelihood, $p(\mathbf{c}_k|\boldsymbol{\Pi}_k,\mathbf{c}_{1:k-1})$ is a context likelihood, $p(\boldsymbol{\Pi}_k|\mathbf{y}_{k-1},\mathbf{c}_{k-1})$ is the TPM prior,

terms $p(\mathbf{y}_k|\mathbf{y}_{1:k-1})$ and $p(\mathbf{c}_k|\mathbf{c}_{1:k-1})$ are measurement and context normalization densities, respectively.

Building on Jilkov's work [21], we utilize the total probability theorem and the Chapman-Kolmogorov equation for Markov chain (15) in order to derive the expressions for context and measurement likelihoods in (21) as

$$\begin{aligned} &p(\mathbf{c}_k|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \\ &= \sum_{i=1}^{N_m} p(\mathbf{c}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1})\; p(m_k^i|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1}), \\ &= \sum_{i=1}^{N_m} p(\mathbf{c}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \\ &\quad \times \sum_{j=1}^{N_m} p(m_k^i|m_{k-1}^j, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \times p(m_{k-1}^j|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1}). \end{aligned} \tag{22}$$

In the expression for context likelihood $p(\mathbf{c}_k|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1})$ (22), probabilities $p(\mathbf{c}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1})$ and $p(m_{k-1}^j|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1})$ expect the knowledge of TPM $\mathbf{\Pi}_k$ at time k, which is not feasible to compute within the same recursive cycle $(k|k-1)$. Therefore, $\mathbf{\Pi}_k$ is commonly approximated with the TPM estimate from the previous cycle, e.g. $\hat{\mathbf{\Pi}}_{k-1}$, which alters (22) as

$$\begin{aligned} &p(\mathbf{c}_k|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \\ &\approx \sum_{i=1}^{N_m} p(\mathbf{c}_k|m_k^i, \hat{\mathbf{\Pi}}_{k-1}, \mathbf{c}_{1:k-1}) \\ &\quad \times \sum_{j=1}^{N_m} p(m_k^i|m_{k-1}^j, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \\ &\quad \times p(m_{k-1}^j|\hat{\mathbf{\Pi}}_{k-1}, \mathbf{c}_{1:k-1}), \\ &\approx \sum_{i=1}^{N_m} \tilde{\Gamma}_k^i \sum_{j=1}^{N_m} \pi_{k|k-1}^{ji} \tilde{\mu}_{k-1}^j = \tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Gamma}}_k. \end{aligned} \tag{23}$$

Similarly, the measurement likelihood $p(\mathbf{y}_k|\mathbf{\Pi}_k, \mathbf{y}_{1:k-1})$ becomes

$$p(\mathbf{y}_k|\mathbf{\Pi}_k, \mathbf{y}_{1:k-1}) \approx \tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Lambda}}_k. \tag{24}$$

In summary, the following approximations being applied in derivations of context (23) and measurement likelihoods (24)

$$\begin{aligned} p(\mathbf{c}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) &\approx p(\mathbf{c}_k|m_k^i, \hat{\mathbf{\Pi}}_{k-1}, \mathbf{c}_{1:k-1}) = \tilde{\Gamma}_k^i, \\ \tilde{\mathbf{\Gamma}}_k &= \left[\tilde{\Gamma}_k^1, \tilde{\Gamma}_k^2, \ldots, \tilde{\Gamma}_k^{N_m}\right]^T, \end{aligned} \tag{25}$$

$$\begin{aligned} p(\mathbf{y}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{y}_{1:k-1}) &\approx p(\mathbf{y}_k|m_k^i, \hat{\mathbf{\Pi}}_{k-1}, \mathbf{y}_{1:k-1}) = \tilde{\Lambda}_k^i, \\ \tilde{\mathbf{\Lambda}}_k &= \left[\tilde{\Lambda}_k^1, \tilde{\Lambda}_k^2, \ldots, \tilde{\Lambda}_k^{N_m}\right]^T, \end{aligned} \tag{26}$$

$$\begin{aligned} p(m_{k-1}^j|\mathbf{\Pi}_k, \mathbf{y}_{1:k-1}) &\approx p(m_{k-1}^j|\hat{\mathbf{\Pi}}_{k-1}, \mathbf{y}_{1:k-1}) = \tilde{\mu}_{k-1}^j, \\ \tilde{\boldsymbol{\mu}}_{k-1} &= \left[\tilde{\mu}_{k-1}^1, \tilde{\mu}_{k-1}^2, \ldots, \tilde{\mu}_{k-1}^{N_m}\right]^T. \end{aligned} \tag{27}$$

The context $p(\mathbf{c}_k|\mathbf{c}_{1:k-1})$ and measurement $p(\mathbf{y}_k|\mathbf{y}_{1:k-1})$ normalization factors are under local linear approximation (23) and (24) defined by (28) and (29), respectively.

$$\begin{aligned} p(\mathbf{c}_k|\mathbf{c}_{1:k-1}) &= \int p(\mathbf{c}_k|\mathbf{\Pi}_k, \mathbf{c}_{1:k-1})\; p(\mathbf{\Pi}_k|\mathbf{c}_{1:k-1}) d\Pi \\ &\approx \int \tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Gamma}}_k\; p(\mathbf{\Pi}_k|\mathbf{c}_{1:k-1}) d\Pi \\ &\approx \tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Gamma}}_k \end{aligned} \tag{28}$$

$$p(\mathbf{y}_k|\mathbf{y}_{1:k-1}) = \tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Lambda}}_k \tag{29}$$

By substituting (23), (24), (28), (29) into (21), we obtain the simplified Bayes update formula of the form

$$\begin{aligned} &p(\mathbf{\Pi}_k|\mathbf{y}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_k, \mathbf{c}_{1:k-1}) \\ &= \frac{(\tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Lambda}}_k)(\tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Gamma}}_k)}{(\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Lambda}}_k)(\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Gamma}}_k)} p(\mathbf{\Pi}_k|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}). \end{aligned} \tag{30}$$

Till this point, likelihoods (23, 24) and normalization factors (28, 29) in (30) become linear with respect to the TPM prior $\hat{\mathbf{\Pi}}_{k-1}$, but they are still highly non-linear across the whole observation horizon $1:k$. In the next step, we are going to utilize the fact that likelihoods $p(\mathbf{y}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{y}_{k-1}) \approx \tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Lambda}}_k$ and $p(\mathbf{y}_k|m_k^i, \mathbf{\Pi}_k, \mathbf{c}_{1:k-1}) \approx \tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Gamma}}_k$ can be approximated as a kth degree polynomial of the TPM $p(\mathbf{\Pi}_k|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1})$ elements over time $1:k$, i.e.

$$\begin{aligned} &p(\mathbf{\Pi}_k|\mathbf{y}_k, \mathbf{c}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) = \\ &\propto \prod_{k=1}^{K} (\tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Lambda}}_k)(\tilde{\boldsymbol{\mu}}_{k-1}^T \mathbf{\Pi}_k \tilde{\mathbf{\Gamma}}_k) p(\mathbf{\Pi}_{k-1}). \end{aligned} \tag{31}$$

Towards this goal, recursion (30) is further simplified with respect to the marginal PDFs $p(\mathbf{\Pi}_k|\mathbf{y}_k, \mathbf{y}_{k-1}, \mathbf{c}_k, \mathbf{c}_{k-1})$ for each row $i = m_k$ of the TPM $\boldsymbol{\pi}_k^i$, i.e.

$$\begin{aligned} &p(\boldsymbol{\pi}_k^i|\mathbf{y}_k, \mathbf{c}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) = \\ &\quad = \int_{k=1}^{K} p(\mathbf{\Pi}_k|\mathbf{y}_k, \mathbf{c}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}) d\mathbf{\Pi}/d\boldsymbol{\pi}^i, \\ &\quad = \frac{A^i}{(\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Lambda}}_k)(\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Gamma}}_k)}, \end{aligned} \tag{32}$$

where $d\mathbf{\Pi}/d\boldsymbol{\pi}^i = d\boldsymbol{\pi}^1, \ldots, d\boldsymbol{\pi}^{i-1}, d\boldsymbol{\pi}^{i+1}, \ldots d\boldsymbol{\pi}^{N_m}$. Term A^i refers to a marginal integral of the ith row of $p(\boldsymbol{\pi}_k^i|\mathbf{y}_k, \mathbf{c}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1})$ and can be solved as follows

$$\begin{aligned} A^i &= \int_{k=1}^{K} (\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k-1} \tilde{\mathbf{\Lambda}}_k)(\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k-1} \tilde{\mathbf{\Gamma}}_k) \\ &\qquad p(\mathbf{\Pi}_k|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1})\; d\mathbf{\Pi}/d\boldsymbol{\pi}^i \\ &= \sum_{l=1}^{N_m} \tilde{\mu}_{k-1}^{l\,T} \int_{k=1}^{K} \boldsymbol{\pi}_k^{l\,T} p(\boldsymbol{\pi}_k^i|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1})\; d\mathbf{\Pi}/d\boldsymbol{\pi}^i \tilde{\mathbf{\Lambda}}_k \tilde{\mathbf{\Gamma}}_k^T \end{aligned} \tag{33}$$

By assuming the marginalization of the form

$$\begin{aligned} &\int_{l=1}^{N_m} \boldsymbol{\pi}_k^{l\,T}\; p[\boldsymbol{\pi}^1, \ldots, \boldsymbol{\pi}^{N_m}|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}] d\boldsymbol{\pi}^1, \ldots, d\boldsymbol{\pi}^{N_m} \\ &\qquad = \hat{\boldsymbol{\pi}}_{k|k-1}^{l\,T}\; p[\boldsymbol{\pi}_k^i|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}] \end{aligned} \tag{34}$$

expression (33) further simplifies to

$$\begin{aligned} A^i &= \Big[\sum_{i \neq l}^{N_m} \tilde{\mu}_{k-1}^l \hat{\boldsymbol{\pi}}_{k|k-1}^{l\,T} + \tilde{\mu}_{k-1}^i \boldsymbol{\pi}_k^{i\,T}\Big] \\ &\qquad \times \tilde{\mathbf{\Lambda}}_k \tilde{\mathbf{\Gamma}}_k^T\; p[\boldsymbol{\pi}_k^i|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}] \\ &= \{\tilde{\boldsymbol{\mu}}_{k-1}^T \hat{\mathbf{\Pi}}_{k|k-1} \tilde{\mathbf{\Lambda}}_k \tilde{\mathbf{\Gamma}}_k^T \\ &\qquad + \tilde{\mu}_{k-1}^i [\boldsymbol{\pi}_k^{i\,T} - \hat{\boldsymbol{\pi}}_{k|k-1}^{i\,T}] \tilde{\mathbf{\Lambda}}_k \tilde{\mathbf{\Gamma}}_k^T\}\; p[\boldsymbol{\pi}_k^i|\mathbf{y}_{1:k-1}, \mathbf{c}_{1:k-1}]. \end{aligned} \tag{35}$$

The results of (32) and (35) represent the polynomial approximation of posterior density $p(\mathbf{\Pi}_k|\mathbf{y}_k, \mathbf{y}_{1:k-1}, \mathbf{c}_k, \mathbf{c}_{1:k-1})$ estimation which is the subject of the discussion in the following section.

2.3. A Quasi-Bayesian estimator for the TPM

In this section, we describe an approximate solution that finds the maximum a posteriori probability $\hat{\mathbf{\Pi}}_k$ given the TPM likelihood $p(\mathbf{y}_{1:k}, \mathbf{c}_{1:k}|\mathbf{\Pi}_k)$ within a single recursion step of finite mixture components [21], i.e. $\hat{\mathbf{\Pi}}_k = \arg\max_{\Pi_k} p(\mathbf{y}_{1:k}, \mathbf{c}_{1:k}|\mathbf{\Pi}_k)$. We start the derivations by formulating the marginal TPM likelihood $p(\mathbf{y}_k, \mathbf{c}_k|\boldsymbol{\pi}_k^i, \mathbf{y}_{k-1}, \mathbf{c}_{k-1})$ (Section 2.2) as the multivariate Gaussian mixture

$$p(\mathbf{y}_k, \mathbf{c}_k|\boldsymbol{\pi}_k^i, \mathbf{y}_{k-1}, \mathbf{c}_{k-1}) = \sum_{j=1}^{N_m} \pi_k^{ji} g_k^{ji} \tag{36}$$

where

$$g_k^{ji} \approx \tilde{g}_k^{ji}\; p(\mathbf{y}_k|\mathbf{y}_{1:k-1})\; p(\mathbf{c}_k|\mathbf{c}_{1:k-1}) \tag{37}$$

with marginal likelihood functions

$$\tilde{g}_k^{ji} = 1 + \eta_k^i [\lambda_k^j \gamma_k^j - \boldsymbol{\pi}_{k-1}^i \mathbf{\Lambda}_k \mathbf{\Gamma}_k^T]. \tag{38}$$

It is worth noting that TPM likelihood can be represented as a mixture (36). This allows us to reformulate the MMSE estimation of marginalized transitional probabilities $\hat{\boldsymbol{\pi}}^i$ into a classification problem referred to as prior probability estimation (PPE) of finite mixtures, i.e.

$$f(\mathbf{y}_k|\boldsymbol{\Pi}) = \sum_{j=1}^{N_m} \pi_k^j \; g_k^j(\mathbf{y}_k) \tag{39}$$

By assuming that mixture components $g_k^j(\mathbf{y}_k)$ are *a priori* known, the goal of PPE estimation is to calculate weights $\hat{\pi}_k^j$ for a given sequence of observations $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k$ with PDF $f(\mathbf{y}_k|\boldsymbol{\Pi})$, where $\sum_{j=1}^{N_m} \hat{\pi}_k^j = 1$. Based on the conclusions in [21], the MMSE TPM estimation criteria could be met by a Quasi-Bayesian TPM estimator [64]. Similarly to the Gaussian mixture approximations, the underlying idea of the Quasi-Bayesian (QB) approach is to approximate the posterior weighted sum of N_m Dirichlet distributions by a single Dirichlet distribution at each time step. Within an iteration, the QB estimator computes the recursive estimate of TPM $\hat{\boldsymbol{\pi}}_k^j$ from a Dirichlet prior $p(\boldsymbol{\pi}_{k-1}^j)$ and likelihood (39) via (40).

$$\hat{\boldsymbol{\pi}}_k^j = \frac{\boldsymbol{\alpha}_k^j}{k + \alpha_0} \qquad \boldsymbol{\alpha}_k^j = \boldsymbol{\alpha}_{k-1}^j + \frac{\boldsymbol{\alpha}_{k-1}^j \boldsymbol{g}_k^j}{\sum_l^m \boldsymbol{\alpha}_{k-1}^l \boldsymbol{g}_k^l} \tag{40}$$

2.4. Context as a constrain factor

Similarly to [27], we describe the capability of contextual information to amend target's mode estimate m_k^i on the basis of c_k^j, i.e. context likelihood $p(c_k^j|m_k^i)$, in terms of mode-sensing rate λ_k (41) and mode-processing accuracy $\mathbf{D}_k^{ij}$. The first parameter, the context observation rate, characterizes the dependence of CI and mode statement $\boldsymbol{s}_k^i$ generation process upon the complexity of information, sensitivity of the detector, sophistication of processing, availability of CPU, etc.

$$\begin{aligned} \lambda_k &= [\lambda_k^1, \lambda_k^2, \dots, \lambda_k^{N_m}], \\ \lambda_k^i &= p(\iota_k = 1|\phi_k^i = 1) = E(\iota_k = 1|\phi_k^i = 1). \end{aligned} \tag{41}$$

In the above equation, indicator function ι_k expresses the appearance of a relevant context in the tracking area

$$\iota_k = \begin{cases} 1, & \text{relevant context is available,} \\ 0, & \text{relevant context is not available.} \end{cases} \tag{42}$$

In this work, the relevant context is a percentage (a random value between $50 - 99\,\%$ for each simulation run) of all contextual samples. Similarly, ϕ_k^i is an indicator vector that characterize the completeness of mode inference process, i.e. generation of s_k^i from c_k^j, as

$$\phi_k^i = \begin{cases} 1, & s_k^i = i, \\ 0, & \text{otherwise.} \end{cases} \tag{43}$$

The result of mode inference process, on the basis of c_k^j and event model parameters $\boldsymbol{\theta}_{k|k-1}$, that is $\boldsymbol{s}_k = \arg\max_{\boldsymbol{\theta}_{k|k-1}} p(\mathbf{c}_k|\boldsymbol{\theta}_{k|k-1})$ with $\mathbf{S}_k \subset \boldsymbol{\theta}_{k|k-1}$, is one of the N_m possible statements denoted by

$$\boldsymbol{s}_k = \{1, 2, \dots, N_m\}. \tag{44}$$

Note that we denoted the mode state of a context inference process by s_k^i instead of m_k^i in order to distinguish HMM from Gaussian Model (GM). Similarly to (43) an indicator vector ρ_k^i is defined for $\boldsymbol{m}_k^i$.

The second modelling parameter characterizes the discrepancy between ϕ_k^i (HMM) and ρ_k^j (GM). The discernibility matrix $\boldsymbol{D}$, denoted by element probabilities

$$\begin{aligned} \delta_k^{ij} &= p(m_k^j|s_k^i), \\ \delta_k^{ij} &= \{\rho_k^j = 1|\iota_k = 1, \phi_k^i = 1\}, \end{aligned} \tag{45}$$

gives a measure of the ability of the tracking filter to correctly recognize mode of the observed target δ^{ij} and the probability of mode misclassification $(\delta_k^{ij}, j \neq i)$.

By considering both mode-sensing rate λ_k and mode-processing accuracy $\mathbf{D}_k^{ij}$ the context constrain is given by

$$\lambda_k^{ij} = p(\rho_k^j = 1|\phi_k^i = 1) = \delta^{ij} \lambda^i \tag{46}$$

and all λ_k^{ij} constitute the rate matrix

$$\boldsymbol{\Lambda} = \boldsymbol{D}^T \mathrm{diag}(\lambda) \tag{47}$$

Given (47) the expression for (9) becomes as follows

$$\begin{aligned} p(m_k^i|\mathbf{y}_{1:k}, \mathbf{c}_{1:k}) = \Big[&(1 - \lambda_k^i) I(c_k^i = 0) + \dots \\ &\lambda_k^i \sum_{l=1}^{d} d_k^{il} I(c_k^l)\Big] p(m_k^i|\mathbf{y}_{1:k}, \mathbf{c}_{1:k-1}) \end{aligned} \tag{48}$$

3. Experimental results

Within this section, we are going to provide details about the experiment design and analyze the performance of the tracking algorithm in cases when the CI was present and absent. Our vision on context-aware target tracking (Fig. 3), motivated and briefly outlined in Section (1.1), combines the situation assessment module with the target tracking algorithm into a multi-layer data fusion architecture.

3.1. Simulation design

We have proposed the context-aware target tracking architecture (Fig. 3) based on the assumption that a target's trajectory within the road network (Fig. 1) can be uniquely defined by a modal context $\mathbf{s}_k$ within a period $k = 1 : K$.

However, modal context $\mathbf{s}_k$, i.e. behaviour patterns of the target (Fig. 1(c)), are not directly observed and therefore they have to be inferred from a sequence of events, i.e. event context, $\mathbf{c}_{1:K}$ (Fig. 1(d)). The HMM (Fig. 2) is tasked to infer the probability of mode context $p(s_k^i)$ given the likelihood $p(\mathbf{c}_k|s_k^i)$ that relates event observations with modes, and the context sequence $\mathbf{c}_{1:K}$. This is achieved by computing the forward probabilities of the mode context $\hat{\boldsymbol{s}}_k = \arg\max_{\boldsymbol{\theta}_{k|k-1}} p(\mathbf{C}_k|\boldsymbol{\theta}_{k|k-1})$ where $\mathbf{S}_k \subset \boldsymbol{\theta}_{k|k-1}$. Here, $\boldsymbol{\theta}_{k|k-1}$ is an emission probability of the observation c_k^e associated with the state s_k^i.

As stated by the Bayesian recursion, defined by (3–9), the context information constrains the mode evolution step (3) as (15), via adaptive TPM estimation process (32, 35), and context mode update step (7) expressed as (11) of the context enhanced maneuvering tracking algorithm (Fig. 3). In order to utilize context in above equations, CI has to be first formalized as the likelihood $p(c_k^e|m_k^i)$ which is, in this work, defined as a function of mode-sensing rate λ_k (41) and mode-processing accuracy, i.e. discernibility matrix, $\mathbf{D}_k^{ij}$ (45). The first parameter, the context observation rate $\lambda_k = [\lambda_k^1, \lambda_k^2, \dots, \lambda_k^{N_m}]^T$, with $N_m = 6$, characterizes the dependence of CI on mode statement $\boldsymbol{s}_k^i$ generation process specified by the probabilities $\lambda_k^i = p(\iota_k|\phi_k^i)$, where ι_k and ϕ_k^i are indicator functions taking values either 0 or 1 defined by (42) and (43), respectively. Mode inference process is being considered as complete, i.e. $\phi_k^i = 1$, when one of the N_m possible statements $\hat{\boldsymbol{s}}_k = \{1, 2, \dots, N_m\}$ is generated by $\hat{\boldsymbol{s}}_k = \arg\max_s p(\mathbf{c}_k|\boldsymbol{\theta}_{k|k-1})$ with $\mathbf{S}_k \subset \boldsymbol{\theta}_{k|k-1}$ on the basis of c_k^e and event model parameters $\boldsymbol{\theta}_{k|k-1}$. Simulation of the mode statement generation process $\lambda_k^i = p(\iota_k|\phi_k^i)$ was performed by setting ι_k to induce a random percentage $(50\,\% - 99\,\%)$ of missing contextual measurements prior to any simulation run.

The second modelling parameter, discernibility $\mathbf{D}_k^{ij}$ characterizes the discrepancy between mode estimated by the IMM filter m_k^j from the one inferred by the HMM s_k^i. Elements of the discernibility matrix $\mathbf{D}$ are probabilities $\delta_k^{ij} = p(m_k^j|s_k^i) = \{\rho_k^j = 1|\iota_k = 1, \phi_k^i = 1\}$ (45), where ϕ_k^i and ρ_k^j are indicator functions for HMM s_k^i and IMM m_k^j, respectively. For example, if the mode of the IMM filter m_k^j corresponds to $j \in CV^1$ and mode inferred by HMM s_k^i is different, i.e. $i \in CTL^1$, then row of matrix $\mathbf{D}_k^i$ is being updated by counting the correctly recognised target

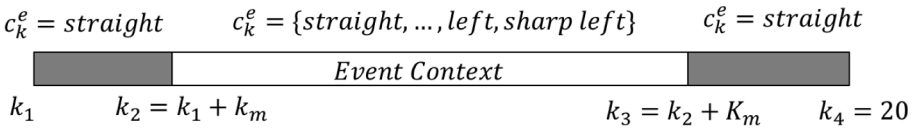

Fig. 4. Visualization of a generic trajectory segment.

modes δ^{ij} and the probability of mode misclassification ($\delta_k^{ij}, j \neq i$). Discernibility matrix $\mathbf{D}_k$ is being updated every time a context inference is performed while only limited history of past mode segments, in our work 5 past time instances, $\mathbf{s}^i_{(k-5):k}$ is being kept in order to achieve a good trade-off between mode consistency and adaptability. Note that if only present context is used in the discernibility update process, matrix $\mathbf{D}_k^i$ will become skewed as there will be insufficient amounts of element counts. On the other hand, keeping the whole history of observation counts will make $\mathbf{D}_k^i$ less sensitive to new context observations, as few incorrectly made estimates will be statistically insignificant to the whole history of observed modes. By considering both mode-sensing rate λ_k and mode-processing accuracy $\mathbf{D}_k^{ij}$ the context constrain is given by $\mathbf{\Lambda} = \mathbf{D}^T \mathrm{diag}(\lambda)$ (47).

Let us justify the importance of the mode context for the target tracking by an example. Suppose, the mode transition probability $p(m_k^{CT^1}|m_{k-1}^{CV^1})$ of the jump from $m_{k-1}^{CV^1}$ to $m_k^{CT^1}$ be very low, i.e. $\boldsymbol{\pi}_{k|k-1}^{ji} = p(m_k^{CT^1}|m_{k-1}^{CV^1}) \approx 0$. However, the target performs a maneuvere corresponding to such an unlikely mode transition. In this case, the chance of a track loss is significantly increased as the IMM tracker considers such a mode change to be very unlikely $p(m_k^{CT^1}|m_{k-1}^{CV^1}) \approx 0$. In real world instances of such a scenario, certain transitions can happen quite frequently while others might be sparse. In our work, context is used to control the mode ambiguity prior to an event thus 'preparing' the filter for such an unlikely mode transition, that is help to decrease and increase the probability of $m_{k-1}^{CV^1}$ and $m_k^{CT^1}$ (Fig. 1), respectively.

Two types of synthetic datasets are considered for the evaluation of the proposed methodology: the Jump Markov Linear Process (JMLP) and a more realistic semi-Markov process. In order to assess the performance of the TPM estimators [21], [22], we decided to model target kinematics as JMLP (1). The modes m_k in JMLP are changing according to a Markov chain $\mathbf{m}_k^i \in \{m_1^i, \ldots, m_K^i\}$ of length K for $\forall i \in 1, \ldots, N_m$ with N_m modes. The sequential Markov process is uniquely defined by a symmetric time-invariant TPM $\mathbf{\Pi}_k = \boldsymbol{\pi}_{k|k-1}^{ji}$ that is also positive-definite.

In order to model the target behaviour in a real tracking environment, i.e. sequence $\mathbf{m}_k \in \{m_1^i, \ldots, m_K^i\}$ of length K $\forall i \in 1, \ldots, N_m$, in accordance to the semi-Markov process we utilize the HMM $\mathcal{G}(\boldsymbol{\pi}_{k|k-1}^{ji}, \boldsymbol{\theta}_{k|k-1}^{ji})$ as an underlying source of context observations. Since the modal context of target s_k^i is in HMM inferred on the basis of CI c_k^e the length of semi-Markov process sequence corresponds to the amount of intersections, e.g. nodes of the graph, the target passes through. The HMM models the transition from s_{k-1}^i to s_k^i through an edge that we model as trajectory segments. A generic trajectory segment is visualised in Fig. 4.

Without loss of generality, we assume that the target can traverse any trajectory segments within $20s$ and that only a single context event c_k^e can occur in this time. Note that while the time needed to traverse any edge in the road topology is assumed to be constant, the distances between the nodes are not the same, as a consequence of different velocity profiles of individual targets. As shown in Fig. 4, at time k_1 the target is moving according to CV model, i.e. $s_k^{CV^1}$, which is also indicated by a context observation $c_k^{straight}$. At random time $k_2 \in \{2, 15\}$ a context event occurs c_k^e which lasts $k_3 + K_m$ long, where K_m is random such that $k_3 \leq 18s$ at most, before the final event returns the filter into straight tracking mode.

In our early work [13], a hypothetical scalar jump Markov linear system with three models was used in the initial set of experiments in order to compare the convergence of the Quasi-Bayesian estimator with existing works [21] and [22]. In this work, we emulate the real target behaviour by combining a nearly constant velocity (49) and coordinated turn (50) models with a pre-defined set of parameters.

$$\begin{bmatrix} p_{k+1}^x \\ p_{k+1}^y \\ v_{k+1}^x \\ v_{k+1}^y \end{bmatrix} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_k^x \\ p_k^y \\ v_k^x \\ v_k^y \end{bmatrix} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} a_k^x \\ a_k^y \end{bmatrix} \tag{49}$$

$$\begin{bmatrix} p_{k+1}^x \\ p_{k+1}^y \\ v_{k+1}^x \\ v_{k+1}^y \end{bmatrix} = \begin{bmatrix} 1 & 0 & \sin\left(\frac{\omega_k T}{\omega_k}\right) & -\frac{1-\cos(\omega_k T)}{\omega_k} \\ 0 & 1 & \frac{1-\cos(\omega_k T)}{\omega_k} & \sin\left(\frac{\omega_k T}{\omega_k}\right) \\ 0 & 0 & \cos(\omega_k T) & -\sin(\omega_k T) \\ 0 & 0 & \sin(\omega_k T) & \cos(\omega_k T) \end{bmatrix} \begin{bmatrix} p_k^x \\ p_k^y \\ v_k^x \\ v_k^y \end{bmatrix} \tag{50}$$

We model the uncertainties in the state space model (49) and (50) as a zero mean normally distributed acceleration noise $\mathbf{a}_k = \mathcal{N}(0, \sigma_a^2)$. Variance σ_a is for low and high noise nearly constant velocity models CV^1 and CV^2 set to $\sigma_a = 0.4ms^{-2}$ and $\sigma_a = 4ms^{-2}$, respectively. Four coordinated turn models for low CTX^1 or high kinematic motion CTX^2 and right CTR^x or left hand side turns CTL^x are considered in the simulation (Fig. 2). While coordinated turn models share the same noise acceleration variance $\sigma_a = 0.04rads^{-1}$ the parameter ω_k is used for differentiation between individual modes, i.e. $\omega_k(CTR^1) = -0.08rads^{-1}$, $\omega_k(CTR^2) = -0.20rads^{-1}$, $\omega_k(CTL^1) = 0.08rads^{-1}$ and $\omega_k(CTL^2) = 0.20rads^{-1}$. We assume that ω_k is an *a priori* known parameter.

The sensor model utilized during the simulation is defined as follows:

$$\begin{bmatrix} y_k^x \\ y_k^y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_k^x \\ p_k^y \\ v_k^x \\ v_k^y \end{bmatrix} + \begin{bmatrix} w_k^x \\ w_k^y \end{bmatrix} \tag{51}$$

We model the measurement uncertainties as a normally distributed white noise $\mathbf{w}_k \approx \mathcal{N}(0, \sigma_w = 10m)$. The detection probability of measurements originated from the target is set to $P_D = 0.9$. The clutter is Poisson distributed with the rate $\beta_{FA} = 1.10^{-7}$. The target can appear in the region $[x, y] \in (0, 10km)$ at any uniformly distributed time between $t_s \in (0s, 50s)$.

3.2. Results and discussion

This section demonstrates the performance of the quasi-Bayesian estimator on two simulated scenarios in the cases of absence or presence of CI.

3.2.1. Scenario A

The first set of experiments assumes a target that is following the hypothetical JMLP process in a 2D scenario. Target kinematics are modeled in accordance with the nearly constant velocity model (49), i.e. CV^2 with $\sigma_a = 4ms^{-2}$, and coordinated turn models (50) for clockwise and $\omega_k(CTR^2) = -0.20rads^{-1}$ counter-clockwise rotations $\omega_k(CTL^2) = 0.20rads^{-1}$. A simplified sensor model is considered, which consists directly from position observations (51) corrupted only by a white noise $\mathbf{w}_k \approx \mathcal{N}(0, \sigma_w = 10m)$.

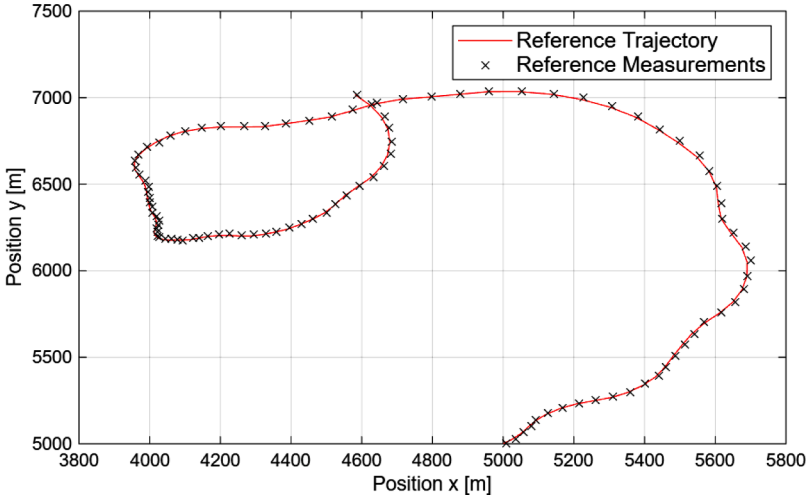

Fig. 5. Example of a trajectory generated by the JMLP described in Section 3.2.1.

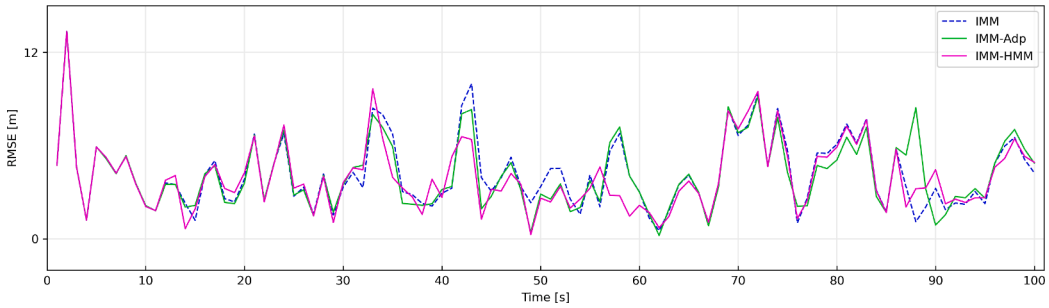

Fig. 6. Comparison of the RMSEs of IMM, IMM-QB adaptive, and IMM-QB-HMM adaptive estimates of the JMLP process.

Modes of the target $\hat{s}^i_{1:k}$ are modeled as a Markov sequence $\hat{\mathbf{s}}_{1:k}$ generated on the basis of the reference TPM:

$$\mathbf{\Pi}^{\star} = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.25 & 0.5 & 0.25 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} \tag{52}$$

An example of the JMLP generation process is illustrated in Fig. 5.

In this scenario, CI, in a form of a likelihood $p(\mathbf{c}_k|\mathbf{\Pi}_k, \mathbf{c}_{k-1})$ expressed as (48), is considered in aiding the TPM estimation process. Here, the context relevance, i.e. weighting factor $\mathbf{w}_k$, is a Gaussian distributed state dependent function $\mathcal{N}(\mathbf{x}_k; \mu, \Sigma)$ without any transitions. This is due to the fact that mode state $\mathbf{s}_{1:k}$ of the JMLP process $\mathbf{s}_{1:k} \in \{s^i_1, \ldots, s^i_k\}$, where $i \in \{CV^2, CTR^2, CTL^2\}$, jumps between modes every time step k, and thus transition models are not required. This time, however, a total number of 50 JMLP sequences $\{\mathbf{s}_{1:k}\}_1^{50} = \mathbf{S}_{1:k}$ were generated where the true sequence $\hat{\mathbf{s}}_{1:k}$ is a subset of $\mathbf{S}_{1:k}$, i.e. $\hat{\mathbf{s}}_{1:k} \subset \mathbf{S}_{1:k}$. The set $\mathbf{S}_{1:k}$ was utilized in the event model learning. The HMM is learned form the parameters specified as: the number of possible states $\sum_{i=1}^{3} s^i_k$, the number of possible context events $\sum_{e=1}^{3} c^e_k$ where $e \in \{$ *straight*, *right*, *left* $\}$, the context event sequences $\{\mathbf{c}_{1:k}\}_1^{50} = \mathbf{C}_{1:k}$, the initial state transition matrix $\mathbf{\Pi}_0$ and the initial emission probability matrix $\mathbf{\Theta}_0$. Note that the set of context sequences $\mathbf{C}_{1:k}$ needs to be generated for the mode sequences $\mathbf{S}_{1:k}$ given the emission probability matrix $\mathbf{\Theta}_k$, in this case chosen at random. Such a mapping is computed as the probability of seeing observations $\mathbf{c}_k = \{c^e_1, c^e_2, \ldots, c^e_K\}$ while being in state i at time k, i.e $\alpha^i_k = p(\mathbf{c}_k, s^i_k|\boldsymbol{\theta}_{k|k-1})$, evaluated by the E-step of the Baum-Welch algorithm.

In Fig. 6 we present a plot of the RMSE in time for each of the evaluated filters w.r.t. the 2D JMLP process [13]. In this particular run, the IMM-QB adaptive tracker outperformed the conventional IMM by a rather small margin of 1.59 % as the mean RMSE of the adaptive IMM filter reached $\text{RMSE}^{IMM-QB} = 4.1102\ m$ while the IMM mean RMSE was $\text{RMSE}^{IMM} = 4.1766\ m$. The inclusion of the CI significantly improved the performance of the IMM-QB-HMM adaptive estimator, i.e. the mean RMSE of $\text{RMSE}^{IMM-HMM} = 3.9936\ m$ is reported. The accuracy was increased by 4.38 % and 2.86 % against the IMM and the IMM-QB adaptive solution, respectively.

This performance gain can be credited to context which boosted the convergence and reduced the steady state error of the estimate of the TPM elements (Fig. 7). Observe that the informativeness of the likelihood (39), a conditional distribution of two independent sources $\mathbf{\Lambda}_k$ and $\mathbf{\Gamma}_k$, is a key factor in the PPE of any TPM matrix represented as a finite mixture. The QB estimator computes a TPM estimate $\hat{\boldsymbol{\pi}}^j_k$ recursively from a Dirichlet prior $p(\boldsymbol{\pi}^j_{k-1})$ and a likelihood (39), (40). In this recursion, the combined likelihood $\mathbf{\Lambda}_k\mathbf{\Gamma}_k$ is considered as a weighting factor, and thus any increase of its informativeness has an immediate impact on the estimator performance.

Similarly to the first scenario, we present the results of 600 Monte Carlo runs of tracking the target, whose kinematics follow a JMLP process, in Fig. 8 and with corresponding statistics presented in Table 1. It can be concluded that utilization of an adaptive TPM estimate can improve the performance of the IMM filter by roughly 7.88 % and 14.96 % while context is absent and present, respectively.

To evaluate the feasibility of running the proposed IMM-QB-HMM adaptive TPM estimation in real time, we analyzed the average execution time per update step for each filter. Since the estimation of the TPM elements occurs only when sensor measurements are available, the computational load is confined to these instances. The filters were executed on an Intel Core i9-10980XE processor (16 cores, 3.0 GHz). The IMM filter exhibited an average execution time per update step of 2.6ms, while the adaptive TPM estimation using the QB estimator required 2.7ms. Incorporating the CI for TPM estimation in the IMM-QB-HMM filter

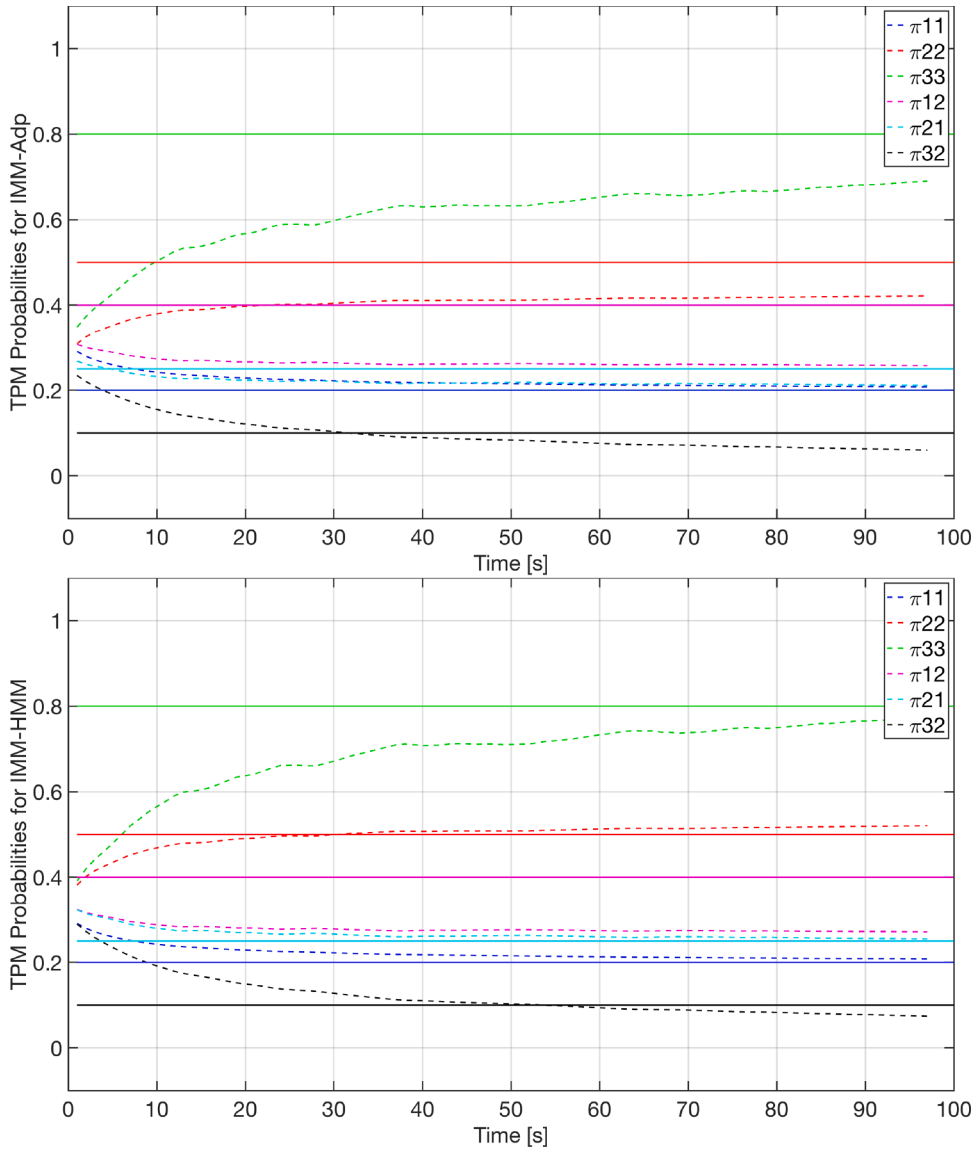

Fig. 7. IMM-QB adaptive (top) and IMM-QB-HMM (bottom) estimates of TPM elements $\hat{\pi}^{ij}_{k|k-1}$ for the JMLP process. In both the plots, the solid lines represent the values of the reference TPM $\mathbf{\Pi}^{\star}$ defined in (52) used to generate the JMLP state sequences. .

Table 1
Comparison of the tracking of a target governed by JMLP process.

Target Tracking	RMSE Mean [m]	RMSE Std. [m]
Kalman filter	4.7665	1.7545
IMM [19]	2.4303	0.6095
IMM-QB Adap [21]	2.2389	0.4905
IMM-QB-HMM Adap. [Ours]	2.0667	0.3942

increased the average execution time per update step to 15.7ms. Despite the additional computational cost introduced by the CI, these results indicate that the IMM-QB-HMM filter remains suitable for real-time implementation.

3.2.2. Scenario B

In the second scenario, tracking of a target behaving according to a semi-Markov process is examined and discussed. In this section, we demonstrate the potential and usefulness of the event context processed by the tracking framework (Fig. 3) on an approximation of the real-world tracking scenario. The process of semi-Markov sequence generation $\mathbf{m}_k \in \{m^i_1, \dots, m^i_K\}$ of length K for $\forall i \in 1, \dots, N_m$ is governed by TPM matrix that corresponds to the state transition matrix of the HMM (Fig. 2). Target specific semi-Markov sequences are generated as samples of the HMM with pre-trained transition $\boldsymbol{\pi}^{ij}_{k|k-1}$ and emission $\boldsymbol{\theta}^{ij}_{k|k-1}$ probability matrices. The resulting dataset corresponds to a set of JMLP sequences that comprise the trajectory segments (Fig. 4) for certain context events. Similarly to the second scenario, a simplified sensor model (51) is utilized for data generation. An example of the trajectory derived from the semi-Markov process described in Section 3.1 is shown in Fig. 9.

The Fig. 10 reports the accuracy of IMM filters supplemented with an online TPM estimation process in the presence of context (2.3) w.r.t. the ground truth. The time between mode m^i_k and m^j_{k-1}, i.e. the sojourn time τ^i of mode m^i_k, is chosen from $f(\tau^{ji})$ or $\Pi(dk) = e^{\Lambda dt}$. Similarly to $\mathbf{\Pi}$, the transition density matrix of process $\mathbf{\Lambda}$ has its diagonal elements defines as $\lambda^{ii} = \frac{1}{\tau}$. By knowing that the mean sojourn time of the process, depicted in Fig. 4, is approximately $\tau \approx 20s$, the direct discrete counterparts of main diagonal elements λ^{ii} can be evaluated as $\pi^{ii} = 1 - \frac{1}{\tau}$.

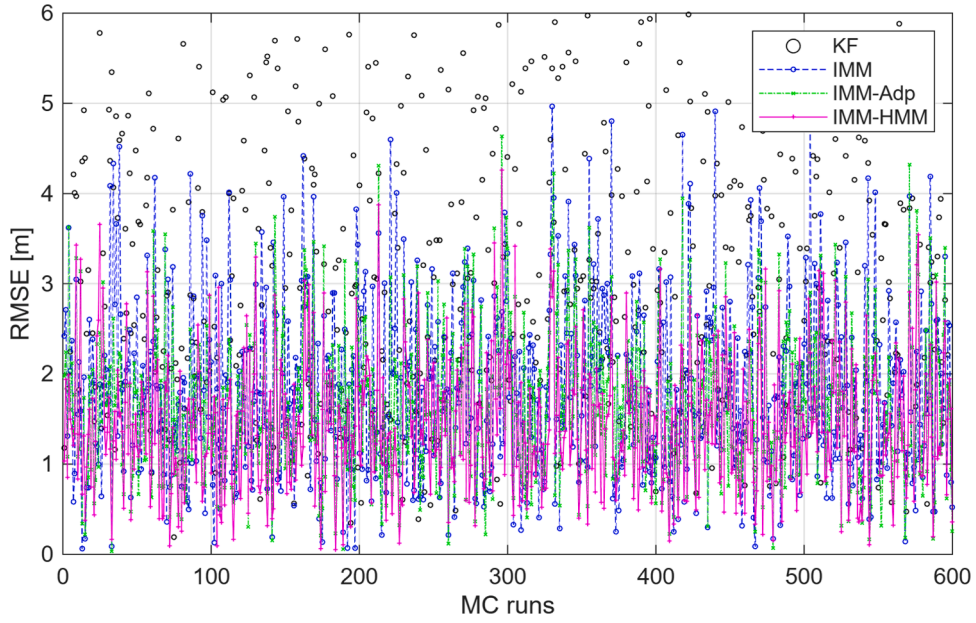

Fig. 8. Monte Carlo simulation of the tracking of a target governed by JMLP process.

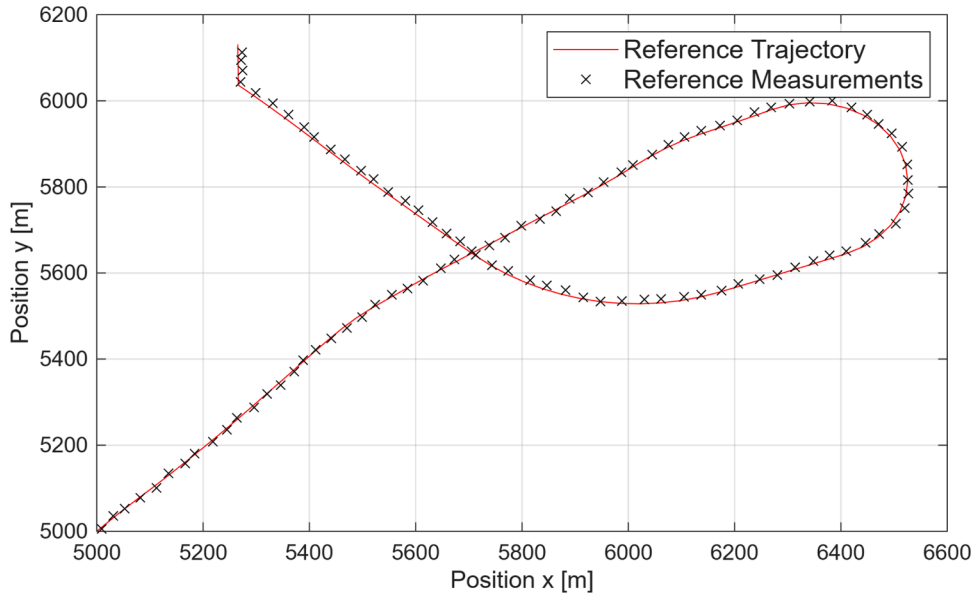

Fig. 9. Example of a trajectory generated by the semi-Markov process described in Section 3.1.

Thus, the best guess of the TPM for SMP equals to matrix $\mathbf{\Pi}_k^{IMM}$ with diagonal elements $\pi_k^{ii} = 0.95$.

In this particular scenario, a gain of 2.41 % was reported, between the IMM-QB adaptive and the IMM solutions with $\text{RMSE}^{IMM-QB} = 2.9389\ m$ and $\text{RMSE}^{IMM} = 2.8680m$, respectively. However, including context the IMM-QB-HMM filter ($\text{RMSE}^{IMM-QB-HMM} = 2.7783m$) returned a gain of 3.12 % and 5.46 % over the IMM and the IMM-QB adaptive filters, respectively. Note that the amount of mode changes that the target performs on its way towards its goal also influences the performance gain of adaptive TPM trackers. This observation reflects the underlying principle of the adaptive TPM estimation process that relies on the count of transitions occurring between the modes. When the number of mode states is increased, significantly longer sequences need to be observed in order to achieve sufficient convergence rates towards the reference TPM.

Inclusion of CI had three distinct advantages. First, context increases the confidence of TPM estimators that a certain mode transition was observed. Second, CI significantly improves the informativeness of the measurement likelihood. Third, CI aids the filter in performing jumps between unlikely modes. The results of 600 Monte Carlo simulation runs of 2D tracking scenario where a target behaves according to the semi-Markov process are shown in Fig. 11 and summarized in Table 2. The results show that adaptive TPM estimation improves tracker accuracy by roughly 11.28 % and 4.88 % when context is present and absent, respectively.

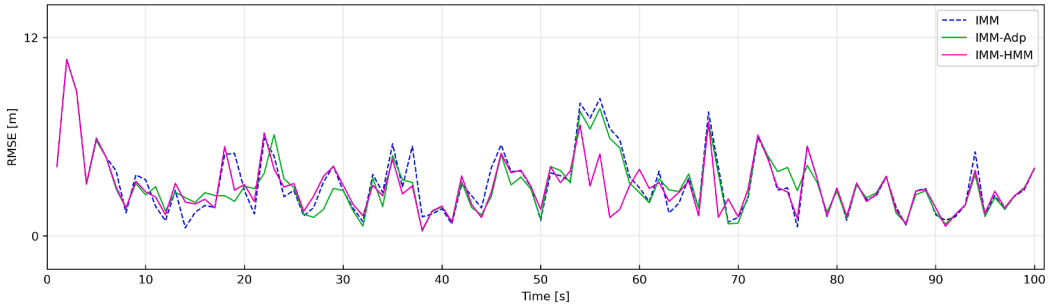

Fig. 10. Comparison of the RMSE of IMM, IMM-QB adaptive, IMM-QB-HMM adaptive estimates of the semi-Markov process.

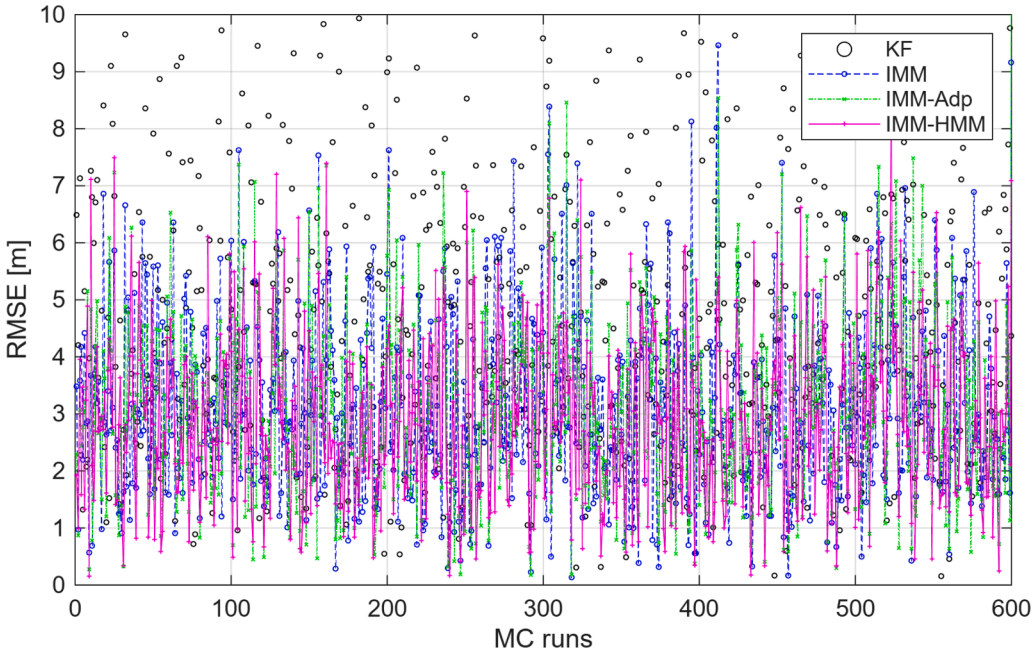

Fig. 11. Monte Carlo simulation of tracking of target governed by scalar semi-Markov process.

Table 2
Comparison of tracking of target governed by a semi-Markov process.

Target Tracking	RMSE Mean [m]	RMSE Std. [m]
Kalman filter	4.8689	3.1148
IMM [19]	2.9667	1.7259
IMM-QB Adap [21]	2.8220	1.7076
IMM-QB-HMM Adap. [Ours]	2.6321	1.6392

4. Conclusions

A target tracking framework capable of exploiting the behavior patterns of targets was proposed and evaluated on synthetic datasets. Behavior patterns were modeled as HMM comprised of the actions that the target performed on its journey. Future actions were predicted from current evidence and knowledge of past history. Predicted actions were used to adaptively estimate the parameters of the mode TPM by a QB estimator within the IMM filter recursion. The sub-optimality of QB estimators gives rise to steady state errors observed in the transition probabilities estimates. Plausible reasons behind these biases could be the approximations involved in the derivation of the measurement likelihood of the TPM, and the mixing process of combining mixture components into a single Gaussian by averaging their means and covariances. It is worth noting that the convergence of the QB estimators is highly dependent on the number of jumps observed between individual modes. For this reason, the informativeness of the likelihood is a key performance factor in the recursive estimation of any TPM matrix represented as a finite mixture.

The underlying role of context in adaptive TPM estimation is twofold: (i) it enhances convergence and mitigates steady-state errors in mode transition estimates, and (ii) it dynamically adjusts mode ambiguity to reduce track losses during unlikely maneuvers. The contextual likelihood is derived from the inferred target mode and a relevance weighting that scales with event proximity. When event outcomes are uncertain, progressive context weighting allows the IMM filter to rebalance observation gains-prioritizing measurements when necessary-to preserve track continuity. Based on the outcome of Monte Carlo simulations, utilization of the adaptive estimate of the TPM matrix in a JMLP tracking scenario improved the average performance of the IMM filter by roughly 14.96 % while context was present and by approximately 7.88 % while context was absent.

In the simulation of the ground object tracking with an airborne sensor, the knowledge of spatial (road topology) and event context (target behaviour patterns) improved average tracking accuracy by 11.28 %. Arguably, the performance gain of the IMM-QB-HMM filter is not final, but it is proportional to the amount of mode transitions for a given event sequence. Longer sequences with a larger amount of mode transitions lead to larger performance gains. Without the presence of context, the accuracy of IMM-QB-HMM and IMM-QB-Adaptive solutions is almost equivalent. Furthermore, the availability of CI $(50\,\% - 99\,\%)$, modeled as a

context statement generation process λ (47), decreased the performance of IMM-QB-HMM proportionally to the amount of missing contextual observation. In a hypothetical case, where all contextual measurements are gone missing, the performance of IMM-QB-HMM and IMM-QB-Adaptive will become equal.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Lubos Vaci: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision; **Marco Mari:** Software, Writing – review & editing, Visualization; **Lauro Snidaro:** Writing – review & editing, Supervision, Conceptualization; **Gian Luca Foresti:** Resources.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by ONRG Grant N62909-17-1-2011.

References

[1] L. Snidaro, J. García, J. Llinas, Context-based information fusion: a survey and discussion, Inf. Fusion 25 (2015) 16–31. https://doi.org/10.1016/j.inffus.2015.01.002
[2] E. Blasch, C. Yang, J. Garcia, L. Snidaro, J. Llinas, Contextual tracking approaches in information fusion, in: Context-Enhanced Information Fusion: Boosting Real-World Performance with Domain Knowledge, Springer International Publishing, Cham, 2016, pp. 73–97. https://doi.org/10.1007/978-3-319-28971-7{_}4
[3] B.I. Ahmad, J.K. Murphy, P.M. Langdon, S.J. Godsill, R. Hardy, L. Skrypchuk, Intent inference for hand pointing gesture-based interactions in vehicles, IEEE Trans. Cybern. 46 (4) (2016) 878–889. https://doi.org/10.1109/TCYB.2015.2417053
[4] B.I. Ahmad, J.K. Murphy, P.M. Langdon, S.J. Godsill, Bayesian intent prediction in object tracking using bridging distributions, IEEE Trans. Cybern. 48 (1) (2018) 215–227. https://doi.org/10.1109/TCYB.2016.2629025
[5] B.I. Ahmad, P.M. Langdon, S.J. Godsill, A Bayesian framework for intent prediction in object tracking, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 8439–8443. https://doi.org/10.1109/ICASSP.2019.8682603
[6] G. Battistello, M. Mertens, M. Ulmke, W. Koch, Context exploitation for target tracking, in: Context-Enhanced Information Fusion, Springer, Cham, 2016, pp. 297–338. https://doi.org/10.1007/978-3-319-28971-7{_}12
[7] F. Gustafsson, U. Orguner, T.B. Schon, P. Skoglar, R. Karlsson, et al., Handbook of Intelligent Vehicles, Springer London, London, London, 2012. https://doi.org/10.1007/978-0-85729-085-4
[8] E.D. Mart, J. Garcia, J.L. Crassidis, Improving multiple-model context-aided tracking through an autocorrelation approach, in: 15th International Conference on Information Fusion, July, Singapore, 2012, pp. 1822–1829.
[9] F. Papi, M. Podt, Y. Boers, On constraints exploitation for particle filtering based target tracking, in: 15th International Conference on Information Fusion, Singapore, 2012, pp. 455–462.
[10] C. Yang, E. Blasch, Fusion of tracks with road constraints, J. Adv. Inf. Fusion 3 (1) (2008) 14–32.
[11] M. Mishra, W. An, X. Han, D. Sidoti, D.F.M. Ayala, K.R. Pattipati, Decision support software for anti-submarine warfare mission planning within a dynamic environmental context, 2014 IEEE Int. Conf. Syst. Man Cybern. 06269 (tier III) (2014) 3390–3393.
[12] L. Vaci, L. Snidaro, G.L. Foresti, Encoding context likelihood functions as classifiers in particle filters for target tracking, in: Proceedings on Multisensor Fusion and Integration, IEEE, Baden-Baden, Germany, 2016, pp. 310–315.
[13] L. Vaci, L. Snidaro, G.L. Foresti, Context-based goal-driven reasoning for improved target tracking, in: 21st International Conference on Information Fusion, IEEE, Cambridge, UK, 2018, pp. 1403–1410.
[14] D. Simon, Kalman filtering with state constraints: a survey of linear and nonlinear algorithms, IET Control Theory Appl. 4 (8) (2010) 1303. https://doi.org/10.1049/iet-cta.2009.0032
[15] M. Ulmke, F. Govaers, Gaussian mixture based target tracking combining bearing-only measurements and contextual information, in: 2018 21st International Conference on Information Fusion (FUSION), 2018, pp. 1–8. https://doi.org/10.23919/ICIF.2018.8455749
[16] E. Blasch, A.N. Steinberg, S. Das, J. Llinas, O. Kessler, E. Waltz, F. White, et al., Revisiting the JDL model for information exploitation, in: 16th International Conference on Information Fusion, Istanbul, 2013, pp. 129–136.
[17] X.R. Li, V.P. Jilkov, Survey of maneuvering target tracking. part V: multiple-model methods, IEEE Trans. Aerosp. Electron. Syst. 41 (4) (2005) 1255–1321. https://doi.org/10.1109/TAES.2005.1561886
[18] T. Connare, E. Blasch, Group IMM tracking utilizing track and identification fusion, in: Proceedings of the Workshop on Estimation, Tracking, and Fusion: A Tribute to Yaakov Bar Shalom, Monerey, California, 2001, pp. 205–220.
[19] H.A.P. Blom, Y. Bar-Shalom, Interacting multiple model algorithm for systems with Markovian switching coefficients, IEEE Trans. Automat. Control 33 (1988) 780–783. https://doi.org/10.1109/9.1299
[20] L. Campo, Y. Bar-Shalom, P. Mookerjee, State estimation for systems with Sojourn-time-dependent Markov model switching, IEEE Trans. Autom. Control 36 (2) (1991) 238–243. https://doi.org/10.1109/9.67304
[21] V.P. Jilkov, X.R. Li, Online Bayesian estimation of transition probabilities for Markovian jump systems, Trans. Signal Process. 52 (6) (2004) 1620–1630. https://doi.org/10.1109/TSP.2004.827145
[22] U. Orguner, M. Demirekler, Maximum likelihood estimation of transition probabilities of jump Markov linear systems, IEEE Trans. Signal Process. 56 (10) (2008) 5093–5108. https://doi.org/10.1109/TSP.2008.928936
[23] C. Liu, B. Li, W.-H. Chen, Particle filtering with soft state constraints for target tracking, IEEE Trans. Aerosp. Electron. Syst. 55 (6) (2019) 3492–3504. https://doi.org/10.1109/TAES.2019.2908292
[24] L. Xu, Y. Liang, Z. Duan, G. Zhou, Route-based dynamics modeling and tracking with application to air traffic surveillance, IEEE Trans. Intell. Transp. Syst. 21 (1) (2020) 209–221. https://doi.org/10.1109/TITS.2018.2890570
[25] D.-U. Kim, W.-C. Lee, H.-L. Choi, J. Park, J. An, W. Lee, Ground moving target tracking filter considering terrain and kinematics, Sensors 21 (20) (2021). https://doi.org/10.3390/s21206902
[26] E.O. Drummond, Feature, attribute, and classification aided target tracking, in: Procedings on Society of Photo-Optical Instrumentation Engineers, SPIE, San Diego, California, 2001, p. 542. https://doi.org/10.1117/12.492746
[27] C. Yang, Y. Bar-Shalom, C.F. Lin, Discrete-time point process filter for mode estimation, Trans. Autom. Control 37 (11) (1992) 1812–1816.
[28] D.D. Sworder, J.E. Boyd, Maneuver sequence identification, Proc. Signal Process. Inf. Fusion 5204 (2004) 155–164. https://doi.org/10.1117/12.502384
[29] P. Nell, A. De Freitas, G. Pavlin, P. De Villiers, Particle-balanced context-based filtering for hypothesis maintenance in sparse sensor coverage situations, in: 2022 25th International Conference on Information Fusion (FUSION), 2022, pp. 1–7. https://doi.org/10.23919/FUSION49751.2022.9841361
[30] H. Gao, Y. Qin, C. Hu, Y. Liu, K. Li, An interacting multiple model for trajectory prediction of intelligent vehicles in typical road traffic scenario, IEEE Trans. Neural Netw. Learn. Syst. 34 (9) (2023) 6468–6479. https://doi.org/10.1109/TNNLS.2021.3136866
[31] Y. Zhu, N.M. Nayak, A.K. Roy-Chowdhury, Context-aware activity modeling using hierarchical conditional random fields, Trans. Pattern Anal. Mach. Intell. 37 (7) (2015) 1360–1372.
[32] H. Zhang, W. Ni, X. Li, Y. Yang, Modeling the heterogeneous duration of user interest in time-dependent recommendation: a hidden semi-Markov approach, IEEE Trans. Syst. Man Cybern. Syst. 48 (2) (2018) 177–194. https://doi.org/10.1109/TSMC.2016.2599705
[33] L. Snidaro, I. Visentini, K. Bryan, Fusing uncertain knowledge and evidence for maritime situational awareness via Markov logic networks, Inf. Fusion 21 (2015) 159–172.
[34] D. Cavaliere, V. Loia, A. Saggese, S. Senatore, M. Vento, Semantically enhanced UAVs to increase the aerial scene understanding, IEEE Trans. Syst. Man Cybern. Syst. (2017) 1–13. https://doi.org/10.1109/TSMC.2017.2757462
[35] S.K. Hautaniemi, J.P.P. Saarinen, Multitarget tracking with the IMM and Bayesian networks : empirical studies, in: Sensor Fusion: Architectures, Algorithms, and Applications V, 2001, pp. 47–57. https://doi.org/10.1117/12.421121
[36] A. Plotnik, S. Rock, 12th International Symposium of Robotics Research ISRR, Robotics Research, 28 of *Springer Tracts in Advanced Robotics*, Springer Berlin Heidelberg, Berlin, Heidelberg, Berlin, Heidelberg, 2007. https://doi.org/10.1007/978-3-540-48113-3
[37] R. Claessens, G. Pavlin, P. de Oude, Context driven tracking using particle filters, in: 2015 18th International Conference on Information Fusion (Fusion), 2015, pp. 1128–1135.
[38] P. de Oude, G. Pavlin, J.P. de Villiers, High-level tracking using bayesian context fusion, in: 2018 21st International Conference on Information Fusion (FUSION), 2018, pp. 1415–1422. https://doi.org/10.23919/ICIF.2018.8455342
[39] R. Schubert, G. Wanielik, A unified bayesian approach for tracking and situation assessment, in: 2010 IEEE Intelligent Vehicles Symposium, 2010, pp. 738–745. https://doi.org/10.1109/IVS.2010.5548072
[40] Z. Tian, Y. Li, M. Cen, H. Zhu, T. Kirubarajan, Context-enhanced vehicle tracking method under the connected environment (poster), in: 2019 22th International Conference on Information Fusion (FUSION), 2019, pp. 1–6. https://doi.org/10.23919/FUSION43075.2019.9011161
[41] Z. Tian, M. Cen, Y. Li, H. Zhu, Ground vehicle tracking using context-based Sojourn time dependent Markov model and pseudo-measurement, IEEE Access 8 (2020) 111536–111552. https://doi.org/10.1109/ACCESS.2020.3001394
[42] B. Trendelkamp-Schroer, F. Noé, Efficient Bayesian estimation of Markov model transition matrices with given stationary distribution, J. Chem. Phys. 138 (16) (2013) 1–16. 1301.2078https://doi.org/10.1063/1.4801325

[43] J. Lan, X.R. Li, Equivalent-model augmentation for multiple-model estimation, Trans. Aerosp. Electron. Syst. 49 (4) (2013) 2615–2630.
[44] E. Ozkan, F. Lindsten, C. Fritsche, F. Gustafsson, Recursive maximum likelihood identification of jump Markov nonlinear systems, Trans. Signal Process. 63 (3) (2015) 754–765. https://doi.org/10.1109/TSP.2014.2385039
[45] R. Guo, M. Shen, D. Huang, X. Yin, L. Xu, Recursive estimation of transition probabilities for jump Markov linear systems under minimum Kullback-Leibler divergence criterion, IET Control Theory Appl. 9 (17) (2015) 2491–2499. https://doi.org/10.1049/iet-cta.2014.0590
[46] A.R. Braga, C. Fritsche, F. Gustafsson, M.G.S. Bruno, Gradient-based recursive maximum likelihood identification of jump Markov non-linear systems, in: Proceedings on 20th International Conference on Information Fusion, IEEE, Xi'an, China, 2017. https://doi.org/10.23919/ICIF.2017.8009651
[47] G. Xie, L. Sun, T. Wen, X. Hei, F. Qian, Adaptive transition probability matrix-based parallel IMM algorithm, IEEE Trans. Syst. Man Cybern. Syst. 51 (5) (2021) 2980–2989. https://doi.org/10.1109/TSMC.2019.2922305
[48] L.B. Cosme, M.F. S.V. D'Angelo, W.M. Caminhas, M.O. Camargos, R.M. Palhares, An adaptive approach for estimation of transition probability matrix in the interacting multiple model filter, J. Intell. Fuzzy Syst. 41 (1) (2021) 155–166. https://doi.org/10.3233/JIFS-201129
[49] I.H. Lee, C.G. Park, An improved interacting multiple model algorithm with adaptive transition probability matrix based on the situation, Int. J. Control Autom. Syst. 21 (10) (2023) 3299–3312. https://doi.org/10.1007/s12555-022-0989-4
[50] X.R. Li, V.P. Jilkov, A survey of maneuvering target tracking - part II: ballistic target models, Proc. SPIE Conf. Signal Data Process. Small Targets (August) (2001) 1–23. https://doi.org/10.1117/12.492753
[51] J. Cao, Y. Liang, L. Liu, Variational Bayesian inference for jump Markov linear systems with unknown transition probabilities, in: 2018 21st International Conference on Information Fusion (FUSION), 2018, pp. 2065–2071. https://doi.org/10.23919/ICIF.2018.8455538
[52] C. Cheng, J.-Y. Tourneret, A variational marginalized particle filter for jump Markov nonlinear systems with unknown transition probabilities, Signal Process. 188 (2021) 108226. https://doi.org/10.1016/j.sigpro.2021.108226
[53] Y. Zhang, T. Wang, K. Liu, B. Zhang, L. Chen, Recent advances of single-object tracking methods: a brief survey, Neurocomputing 455 (2021) 1–11. https://doi.org/10.1016/j.neucom.2021.05.011
[54] M. Fiaz, A. Mahmood, S. Javed, S.K. Jung, Handcrafted and deep trackers: recent visual object tracking approaches and trends, ACM Comput. Surv. 52 (2) (2019). https://doi.org/10.1145/3309665
[55] R. Kumar, M. Bhanu, J.a. Mendes-Moreira, J. Chandra, Spatio-temporal predictive modeling techniques for different domains: a survey, ACM Comput. Surv. 57 (2) (2024). https://doi.org/10.1145/3696661
[56] X.-B. Jin, R.J. Robert Jeremiah, T.-L. Su, Y.-T. Bai, J.-L. Kong, The new trend of state estimation: from model-driven to hybrid-driven methods, Sensors 21 (6) (2021). https://doi.org/10.3390/s21062085
[57] Y. Bai, B. Yan, C. Zhou, T. Su, X. Jin, State of art on state sstimation: Kalman filter driven by machine learning, Annu. Rev. Control 56 (2023) 100909. https://doi.org/10.1016/j.arcontrol.2023.100909
[58] Q. Fu, K. Lu, C. Sun, Deep learning aided state estimation for guarded semi-Markov switching systems with soft constraints, IEEE Trans. Signal Process. 71 (2023) 3100–3116. https://doi.org/10.1109/TSP.2023.3274937
[59] L. Deng, D. Li, R. Li, Improved IMM algorithm based on RNNs, J. Phys. Conf. Ser. 1518 (1) (2020) 012055. https://doi.org/10.1088/1742-6596/1518/1/012055
[60] M. Mari, L. Snidaro, Ensemble of KalmanNets with innovation-based attention for robust target tracking, Inf. Fusion 127 (2026) 103777. https://doi.org/10.1016/j.inffus.2025.103777
[61] Y. Chen, X. Liu, C. Li, J. Zhu, M. Wu, X. Su, UAV Swarm centroid tracking for edge computing applications using GRU-assisted multi-model filtering, Electronics 13 (6) (2024). https://doi.org/10.3390/electronics13061054
[62] F. Marchetti, F. Becattini, L. Seidenari, A.D. Bimbo, Multiple trajectory prediction of moving agents with memory augmented networks, IEEE Trans. Pattern Anal. Mach. Intell. 45 (6) (2023) 6688–6702. https://doi.org/10.1109/TPAMI.2020.3008558
[63] M. Fanaswala, V. Krishnamurthy, Spatio-temporal trajectory models for target tracking, Trans. Aerosp. Electron. Syst. 47 (1) (2014) 2824–2843.
[64] M.F.A. Smith, E.U. Makov, A Quasi-Bayes sequential procedure for mixtures, J. R. Stat. Soc. 40 (1) (1978) 106–112.