Efficient LiDAR-trajectory affinity model for autonomous vehicle orchestration.

MEKALA, M.S., DHIMAN, G., VIRIYASITAVAT, W., PARK, J.H. and JUNG, H.-Y.

2023

Efficient LiDAR-Trajectory Affinity Model for Autonomous Vehicle Orchestration

M. S. Mekala, *Member, IEEE*, Gaurav Dhiman, *Senior Member, IEEE*, Wattana Viriyasitavat, *Senior Member, IEEE*, Ju H. Park, *Senior Member, IEEE*, and Ho-Youl Jung

***Abstract*— Computation and memory resource management strategies are the backbone of continuous object tracking in intelligent vehicle orchestration. Multi-object tracking generates enormous measurements of targets and extended object positions using light detection and ranging (Lidar) sensors. Designing an adequate object-tracking system is a global challenge because of dynamic object detection and data association uncertainties during scene understanding. In this regard, we develop an intelligent multi-objective tracking (IMOT) system with a novel measurement model, called the box data association inflate (BDAI) model, to assess each target's object state and trajectory without noise by using the Bayesian approach. The box object filter method filters ambiguous detection responses during data association. The theoretical proof of the box object filter is derived based on binomial expansion. Prognosticating a lower-dimension object than the original point object reduces the computational complexity of vehicle orchestration. Two datasets (NuScenes dataset and our lab dataset) are considered during the simulations, and our approach measures the kinematic states adequately with reduced computation complexity compared to state-of-the-art methods. The simulation outcomes show that our proposed method is effective and works well to detect and track objects. The NuScenes dataset contains 28130 samples for training, 6019 examples for validation and 6008 samples for testing. IMOT achieves 58.09% tracking accuracy and 71% mAP with 5 ms pre-processing time. The Jetson Xavier NX consumes 49.63% GPU and 9.37% average power and exhibits 25.32 ms latency compared to other approaches. Our system trains a single pair frame in 169.71 ms with affinity estimation time of 12.19 ms, track association time of 0.19 ms and mATE of 0.245 compared to state-of-the-art approaches**

***Index Terms*— Mobile computing, BDAI model, box object filter method, deep learning, multi-object tracking system.**

I. Introduction

A MULTI-OBJECT tracking (MOT) mechanism is an essential paradigm for automated vehicles to strengthen intelligent transport systems. Researchers and scientists have obtained many findings on measuring the state of objects (location of an object and its size in 2D space or volume in 3D space) and extended objects classified based on their characteristics, such as size, direction, and shape. Therefore, an efficient data association approach needs to be designed to streamline the targeted and extended object measurements generated from light detection and ranging (LiDAR) sensors. MOT methods are categorised based on filters, such as the probabilistic hypothesis density filter [1], [2] and multi-Bernoulli filter [3]. In [4], a symmetric positive definite random matrix was used to denote the target extent through random variables. However, it is valid only if the kinematic state is iteratively observed. In [5], an implicit function was considered to represent the object shape as a replacement for the parametric form. Likewise, the Gaussian process was modelled to define the object state through the spatio-temporal Gaussian process covariance function [6].

Traditional tracking systems consider motion prediction strategies, such as the Kalman filter and bipartite graph models [7]. However, these methods are inappropriate during cross-turning and sudden braking scenarios because of unsatisfactory results. For instance, if the detection process misses an object, the corresponding box is assigned to a different object, which causes tracking errors.

Fig. 1 illustrates a scenario to measure traffic objects using the deployed lidar sensors to track suspicious or targeted objects through a component-based measurement system using the onboard unit (OBU) and roadside unit (RSU) frameworks. Hypothesis data and current state data of an object are considered to detect and localize the objects by assigning bbox; consequently, the tracking process is carried out based on the tracking component, which helps to decide on the vehicle's next move. Past examinations concentrated on designing a robust object appearance model to enhance the object tracking accuracy rate. For instance, the scale-invariant feature transform feature, histogram of oriented gradient feature and

Manuscript received 19 November 2021; revised 20 March 2022, 13 June 2022, 21 September 2022, and 14 December 2022; accepted 23 January 2023. This work was supported in part by the Basic Science Research Programs of the Ministry of Education under Grant NRF-2018R1A2B6005105 and in part by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) under Grant 2019R1A5A8080290. The Associate Editor for this article was S. Mumtaz. *(Corresponding authors: M. S. Mekala; Ho-Youl Jung.)*

M. S. Mekala is with the Department of Information and Communication Engineering and the RLRC for Autonomous Vehicle Parts and Materials Innovation, Yeungnam University, Gyeongsan 38544, South Korea, and also with the School of Computing, Robert Gordon University, AB10 7GE Aberdeen, U.K. (e-mail: msmekala@yu.ac.kr).

Gaurav Dhiman is with the Department of Electrical and Computer Engineering, Lebanese American University, Byblos 1526, Lebanon, also with the Department of Computer Science, Government Bikram College of Commerce, Patiala 147001, India, also with the University Centre for Research and Development, Department of Computer Science and Engineering, Chandigarh University, Gharuan, Mohali 140413, India, and also with the Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India (e-mail: gdhiman0001@gmail.com).

Wattana Viriyasitavat is with the Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok 10330, Thailand (e-mail: hardgolf@gmail.com).

Ju H. Park is with the Department of Electrical Engineering, Yeungnam University, Gyeongsan 38544, South Korea (e-mail: jessie@ynu.ac.kr).

Ho-Youl Jung is with the Department of Information and Communication Engineering and the RLRC for Autonomous Vehicle Parts and Materials Innovation, Yeungnam University, Gyeongsan 38544, South Korea (e-mail: hoyoul@yu.ac.kr).

Digital Object Identifier 10.1109/TITS.2023.3242900

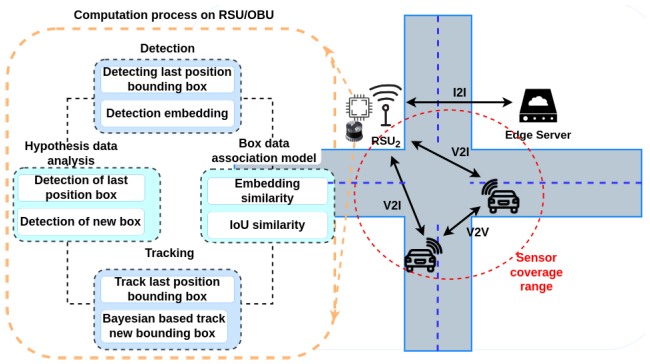


Fig. 1. Continuous object detection and tracking framework.

histogram are considered to design object identification models with various affinity estimations, such as coefficient correlation and χ^2 distance factors. The cloud-based intelligent sensor deployment network was designed based on the Lyapunov approach, which helps enhance service reliability [8]. In [9], a hierarchical association model was designed for efficient MOT based on two-level (local and global) target representation. Local patches are essential to represent the target at the local level, and the global level represents the target with double-bounding boxes. Most affinity models' motion information is linear, but in real time, the target motion is non-linear, specifically in the case of occlusion. In this regard, a non-linear affinity model was designed for MOT based on an affinity score [10], where the network node represented the tracklist and its edge represented the likelihood of neighbouring tracklets. Affinity metrics and object feature quality significantly impact data association model accuracy by comprehending complex object motions and measuring differences in object appearance. Consequently, an adequate data association model diminishes the occlusion ratio and the decision mistakes made in terms of object detection. A few studies have been conducted recently in line with these objectives; for example, a deep affinity network (DAN) was designed for object tracking based on the affinities of joint modelling object appearances per frame. The affinities were measured based on the pairing permutations of selected features [11]. In continuation, a multiple ship tracking system was designed for an adequate assessment of complex marine scenes (long-term occlusions), where a DAN was used to improve scale, region (joint global region modelling module for assessing region dimensions) and motion (motion-matching optimisation module for assessing motion dimensions) aspects of the object for effective tracking [12]. An ant colony heuristic method was used to construct the data association model to comprehend the uncertainty measurements for effective object tracking [13].

Motivation: Object behaviour estimation is a continuous, time- varying and dynamic processes to localise an object through effective data association models using two frameworks (tracking-by-detection (TBD) and detection-by-racking (DBT)). Usually, the TBD framework comprises batch tracking or an online tracking strategy for effective MOT. However, the formulation of data association issues remains challenging because object detection is based on object hypothesis data. One of the main concerns is constructing an affinity model by estimating the object structure of each motion in each frame, which requires enormous measurements. Size, shape and motion pattern measurements are essential components in the affinity model for localising the extended object. Subsequently, eliminating ambiguous object structure detections is necessary to minimise the computation workload, which helps to increase system speed. In this regard, the Bayesian network is considered to accomplish the targets because it is a directed acyclic graph model that helps search for the objects, and maximises the detection probability based on Bayesian statistics with limited available information. However, intelligent computing devices are resource-limited, and designing a box filter method to avoid ambiguous measurements to meet their computation capacity is challenging. A box data association inflate (BDAI) model and a box object filter is derived based on binomial expansion to address this issue. Therefore, a novel measurement model is designed to streamline object detection and data association issues. Our main contributions are as follows:

1) Develop a novel BDAI model to assess each target's object state and trajectory without noise, based on the Bayesian approach.
2) Develop a box object filter to avoid ambiguous detection responses based on the correlation catalecticant square matrix during data association.
3) Design an intelligent vehicle orchestration to optimise computation complexity for continuously monitoring and tracking objects.

The manuscript continues as Section II briefly explains the research gaps and problem statements of the extant approaches. Section III describes the proposed system and its mathematical models with novel algorithms in detail. Section IV, evaluates the investigation outcomes and Section V concludes the manuscript.

II. Related Work

In this section, the previously proposed MOT systems are examined, and the possible shortfalls in their effective formulation are briefly described. A random matrix and joint probabilistic data-association filter were considered together to measure the targeted object motions and status, but this model is affected by inadequate computation complexity during the object-tracking analysis mechanism [14]. Therefore, ellipsoid gates were considered to avoid data association issues, but this scheme is not feasible in a large heterogeneous environment. A robust online motion affinity model was designed based on the tracklet confidence function to optimise the data association issue for MOT [15]. Tracklet confidence was measured based on tracklet continuity and detectability, and a deep appearance learning model was designed to increase association reliability between tracklets. In [16], a relational appearance features and motion patterns learning-based data association model was designed to generate tracks with the reference of one object and its feature differences compared to other objects. In [17], an MOT model was designed to track and detect moving objects by eliminating features lying on tracked objects. A

simple, instinctive cost function was considered to streamline the real-time performance of the visual odometry system. In [18], a constant acceleration motion model was implemented to track the future positions of tracking objects. In [19], the author focused on learning algorithms for MOT in connection with linear processes to streamline multi-dimension assignments. A probabilistic assessment method and end-to-end optimisation factors were derived to assess the learning differences during data association. However, the object localisation rate is not feasible for lightweight environments. An online optical pose association technique was designed to track objects based on a camera system. The occlusion issue was attempted to be resolved based on the local pose-matching strategy. PWC-Net was considered to measure the extended pose based on the differences between the current and previous frames [20]. In [21], MOT frameworks were comprehensively discussed based on their assessments, formulations, principles, drawbacks and scopes to be focused on in future examinations with quantitative comparisons. The authors in [22] addressed optimal object trajectory issues based on two maximum-a-posteriori methods for MOT through object detection and data association measurements; however, these methods are not suitable for complex association probability models because of the association ambiguity caused by noise and frequent inter-object communications. A multi-person tracking algorithm was designed based on object detection and data association strategies. Here, the YOLOv3 algorithm was used for identifying the pedestrian target and the Kalman filter was used for tracking and predicting the target. The Hungarian algorithm used pedestrian features (depth and motion) to detect and predict the results for tracking multi-pedestrian targets. A hybrid track association algorithm was designed to track the distances through an incremental Gaussian mixture model to calculate the association cost [23]. The authors in [24] discussed the process of data association with clustering-based algorithms and the feature group method to reduce the occlusion drop rate without changing the original framework. In [25], a non-local attention association methodology was designed for online MOT, which depended on single object tracking and data association. It also included spatial and temporal features to resolve drawbacks such as noise, occlusion and repeated interactions between targets. In [26], a multi-joint integrated track-splitting tracker was implemented based on a multi-target track-splitting structure. Here, a tree was assembled, with each track component identified with data associations in multiple scans. The hypothesis-testing based tracking (HTBT) method was designed based on the spatio-temporal interaction graph model to construct an effective data association system [27]. A probabilistic 3D multi-object tracking (P3DMOT) system was designed based on feature extraction, Mahalanobis and feature differences to track unmatched objects [28]. However, state-of-the-art approaches have not designed an effective affinity model or path-reliability estimation method to consolidate the data association. To address these issues, we designed an intelligent multi-objective tracking (IMOT) system with a novel measurement model, called the BDAI model, to assess each target's object state and trajectory without noise by using the Bayesian approach.

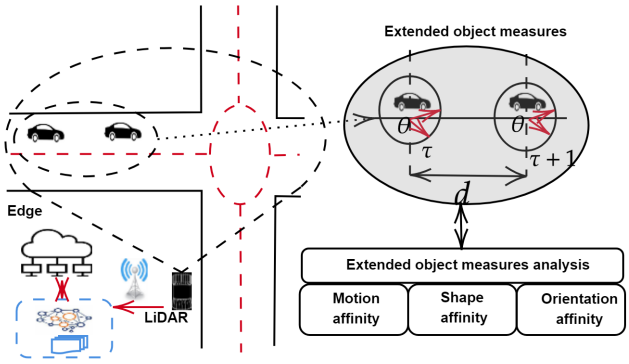


Fig. 2. Extended object tracking with characteristics.

III. Proposed System Model

The proposed IMOT system is designed based on the Bayesian theory for continuous tracking of objects by the intelligent vehicle. Fig. 2 illustrates the measurement of an extended object with size, shape, direction and position. We observe that a car is localised at time τ, which is considered an initial object measurement, and the extended object localises with direction, size and shape. In this regard, the designed novel data association model formulates object-tracking uncertainties, enabling the box object filter to assess the object location for effective tracking. The designed BDAI model and box object filter are both derived in the following sections.

A. Box-Object Filter Model

The multi-object tracking issue is formulated by constructing object and data association model based on the object region or surface. Let us consider, the object states are maintain with a single vector $v_\tau = \left(v_{1,\tau}^T, v_{2,\tau}^T, \ldots, v_{N,\tau_f}^T\right)^T$, where $\tau \in \{1, 2, \ldots, \tau_f\}$ refers the time slot for each measure and N is number of objects. The sub-state vector for extended objects is $v_i^\tau = \left(V_{1,\tau}^T, U_{1,\tau}^T\right)$, where $V_{1,\tau}^T$ refers kinematics of motion (position, velocity), $U_{1,\tau}^T$ refers the considered parameters to model extent object with shapes, and both are independent variables. An object measurement set is denoted with $M = \left\{m_1^\tau, m_1^\tau, \ldots, m_{Q_\tau}^{\tau_f}\right\}$ and the measurements are disordered, and $Q_\tau = \sum_{i=1}^{N} Q_i^s + Q_i^c$ refers total measurements at time τ, where Q_i^s is an object surface measurements with mean a_i^s, Q_i^c refers cluster measurements with mean a_i^c, respectively. The extent object, p_τ, existence is referred with binary value and it is denoted as $p_\tau = \left(p_1^\tau, p_2^\tau, \ldots, p_i^\tau\right)^T$ for individual object tracking, and $p_i^\tau \in \{0, 1\}$ where 0 indicates the object does not exist, and 1 indicates object presence. The probability of extent object localization is formulated with probability density factor (PDF) and is measured as

$$\phi p_i^\tau | p_i^{\tau-1} = \begin{cases} \phi_p, & p_i^\tau = \Gamma_i \\ 1 - \phi_p, & otherwise \end{cases} \tag{1}$$

$$\phi_p = \exp\left(-\frac{1}{p}\sum_{l=1}^{L}\sum_{i=1}^{N}\sqrt{\mathbb{Z}_{i,l}^{j}(\delta_i)\cdot\mathbb{Z}_{i+1,l}^{j+1}(\delta_{i+1})}\right) \tag{2}$$

where $\mathbb{Z}^{j}_{i,l}(\delta_i), \mathbb{Z}^{j+1}_{i+1,l}(\delta_{i+1})$ are color histogram features to measure the target object localization with path or trajectory j, ϕ_p is a probability of existence, δ_i is reliability of path which estimates with Eq. 7, $p^{\tau}_i = \Gamma_i$ evolves with Markov chain model, l is a bin $l \in L$ and $L = 64$ for each histogram color space.

B. Problem Formulation

The extant and extended objects measurements together denoted with $vp_\tau = \left(v^T_\tau, p^T_\tau\right)^T$, and initial probability of object localization is $\phi(vp_0)$. The PDF of precedence state $\phi(vp_\tau|M_{1:\tau})$ is an important factor to localize the extended object at each frame. It is formulated as follows

$$\begin{aligned}\phi(vp_\tau|M_{1:\tau}) &= \frac{\phi(M_\tau|vp_\tau)\,\phi\left(vp^{\tau|\tau-1}_i|M^{\tau}_i\right)}{\phi(M_\tau|M_{1:\tau-1})}\\ \phi(vp_\tau|M_{1:\tau-1}) &= \int_1^{\tau}\phi(M_\tau|vp_\tau)\times\phi\left(vp^{\tau|\tau-1}_i|M^{\tau}_i\right)dx\end{aligned} \tag{3}$$

Where $\phi(vp_\tau|M_{1:\tau-1})$ is the PDF of precedence state and $\phi(M_\tau|M_{1:\tau-1})$ is a normalized policy, M^{τ}_i is a object position parameter set, $\phi(M_\tau|vp_\tau)$ is a probability of expected motion state which is estimated with Eq. 4, $\phi\left(vp^{\tau|\tau-1}_i|M^{\tau}_i\right)$ is a current motion state probability which estimates with Eq. 5. The probability rate of precedence state estimates with Eq. 3, which has a significant impact on MoT and the probability rate of extended object state is mathematically derived in the following sections.

C. Extended Object Motion Analysis

The current and expected motion state analysis methods are essential in data association to construct the effective affinity model. In this regard, estimating the current and extended motion states of object is significant and they are derived as follows

1) Expected Motion State Analysis Method: The probability of expected motion state $\phi(M_\tau|vp_\tau)$ estimation is a significant factor to predict object presence in the successive frames and it is formulated as follows.

$$\phi(M_\tau|vp_\tau) = \frac{e^{-a^s_i}}{Fac(Q_\tau)}\begin{pmatrix}\Delta^{Q_\tau} + \sum_{q=1}^{Q}\sum_{i=1}^{N_a}\Delta^{Q_\tau-q}\times\\ \prod_{q=1}^{Q} a^s_i\phi\left(M^{q_{com}}_\tau|v^{q_e}_\tau\right)\end{pmatrix} \tag{4}$$

Where, $M^{q_{com}}_\tau$ is the complete measurements of object at time τ, $v^{q_e}_\tau$ is existent object sequence association measurements, and the cluster density is $\Delta^{Q_\tau} = a^s_i / cluster\,area$

2) Object Motion Analysis Method: The probability of object's current state is formulated as follows

$$\phi\left(vp^{\tau|\tau-1}_i|M^{\tau}_i\right) = \prod_{i=1}^{N}\begin{bmatrix}\phi\left(\left(v^{\tau}_i|v^{\tau-1}_i\right),\left(p^{\tau}_i,p^{\tau-1}_i\right)\right)\\ \times\phi\left(p^{\tau}_i|p^{\tau-1}_i\right)\end{bmatrix} \tag{5}$$

Subsequently, motion forecasting of i^{th} object is defined as follows

$$\begin{aligned}&\phi\left(\left(v^{\tau}_i|v^{\tau-1}_i\right),\left(p^{\tau}_i,p^{\tau-1}_i\right)\right)\\ &= \begin{cases}\phi_{co}\left(v^{\tau}_i\right) & \left\{p^{\tau}_i,p^{\tau-1}_i\right\} = \{1,0\}\\ \phi_e\left(v^{\tau}_i\right) & \left\{p^{\tau}_i\right\} = \{0\}\\ \phi\left(v^{\tau}_i|v^{\tau-1}_i\right) & \left\{p^{\tau}_i,p^{\tau-1}_i\right\} = \{1,1\}\end{cases}\end{aligned} \tag{6}$$

Where $\phi_{co}\left(v^{\tau}_i\right)$ is object confirmation, $\phi_e\left(v^{\tau}_i\right)$ object end-status, and $\phi\left(v^{\tau}_i|v^{\tau-1}_i\right)$ is prognosticate motion of object. The above analysis methods are essential to construct the expected path and object presence based on the association matrix for increasing tracking efficiency.

D. BDAI Model Formulation

The path or trajectory measurement is an important factor for continuous object tracking. Therefore, validating the path expected measurements are required to diminish the tracking error rate and to enhance reliability. Hence, the reliability of the path or trajectory is measured as follows

$$\delta_i = \frac{1}{\hbar_i}\left(\sum_{fs=1}^{fr\in[fs,fe]}\phi(\delta_i,v_i)\right)\cdot a^s_i\cdot\varepsilon\cdot\wp \tag{7}$$

where $\varepsilon = ls - fs - \hbar_i$ is a total number of object-missed frames, $\hbar_i$ is the length of object path, $\phi(\delta_i, v_i)$ is a probability rate between detected object and path, $\wp$ is a balancing weight factor, and ls, fr, fs, fe refers last frame, total frames, starting frame, ending frame, respectively. The path reliability rate is $\geqslant 0.5$, then it considers as a high path reliability rate, which is in the bounded range of $[0, 1]$. As usual, the high-reliability rate objects and their paths are maintained in a matrix which is denoted with $\left[\delta^j_i\right]_{J\times M}$, where $\delta^j_i = \frac{1}{\hbar_i}\log\left(\phi\left(\delta_{i,h}, v_i\right)\right)$.

The extended object measurements enable three factors (object localization (ϕ^i_p), motion (ϕ^i_m), size (ϕ^i_s)) to assess the probability of each object path quality and it is defined as follows

$$\phi(\delta_i) = \prod_{i=1}^{N}\phi^i_p\cdot\phi^i_s\cdot\phi^i_m \tag{8}$$

where ϕ^i_p is measure with Eq. 2, $\phi^i_m = \phi\left(\left(v^{\tau}_i|v^{\tau-1}_i\right),\left(p^{\tau}_i,p^{\tau-1}_i\right)\right)$ is measure with Eq. 6, and the size (ϕ^i_s) is defined as follows

$$\phi^i_s = \exp\left(-0.5\times\left(\frac{x_i - x_{i+1}}{x_i + x_{i+1}} + \frac{y_i - y_{i+1}}{y_i + y_{i+1}}\right)\right) \tag{9}$$

where x, y is object's height and weight. The dynamic motion of objects at each frame is formulated as follows, where $\delta^{j,j+1}_{i,i+1}$ is a joint path between two objects $i, i+1$.

$$\phi\left(v^{\tau}_i|v^{\tau-1}_i\right) = \left(\frac{\delta^j_i + \delta^{j+1}_{i+1}}{\delta^{j,j+1}_{i,i+1}}\right) - 1 \tag{10}$$

A Bayesian network is an enforced directed acyclic graph with a set of random variables that enhances the efficiency of the

data association model by estimating the detection probability within a timestamp. Each object's hypothesis data are used to assess the object tracking based on contextual affinity methods. The data association model emphasises differential probability errors between the predicted and current object states for effective tracking. Therefore, PDF is gleaned based on the Bayesian mechanism for the effective construction of the data association model as follows:

$$\phi\left(\left(v_i^\tau, V_i^\tau\right), p_i^\tau, \gamma_i | M_i^\tau\right) = \phi\left(v_i^{\tau|\tau-1} | M_i^\tau\right) \times \phi\left(V_i^\tau | M_i^\tau\right) \times \phi\left(\gamma_i^\tau | M_i^\tau\right) \tag{11}$$

where, $\phi\left(v_i^{\tau|\tau-1} | M_i^\tau\right)$ is a probability of predicted state (PPS) of object which is estimated with Eq. 5, $\phi\left(V_i^\tau | M_i^\tau\right)$ is a probability of predicted extension (PPE) of object, $\phi\left(\gamma_i^\tau | M_i^\tau\right)$ is a probability of predicted attributes (PPA) of object, V_i^τ is an object extension set, M_i^τ is an object position parameter set, γ_i is a random scalar of PDF.

1) Measurement Complexity: The computation complexity measurement is required to create a trade-off between cost and system efficiency. Consequently, measuring each object trajectory under a heterogeneous environment is a Hercules process, since the number of cycles is n^2. The complexity of this process is $O\left(n^2\right)$. Subsequently, the probability estimation of each extended object and the construction of a correlation matrix based on hypothesis data play important roles in path measurement and tracking. The process complexity is expressed as $O\left(n\log_2 n\right)$, while the optimal complexity is expressed as $O\left(n^2\right)$ because of $O\left(n^2\right) \gg O\left(n\log_2 n\right)$. Therefore, the path-tracking complexity is expressed as $O\left(n^2\right)$.

Updating the extended object measurements concerning the time window is essential to construct the object's reliability matrix and hypothesis data. Therefore, the time and extended object vectors are revamped as follows:

2) Time Update: The data association stage design is a global challenge under multi-target tracking problems, which can be resolved based on the measurement of the targeted object probability. The measurement of probability predicted extension of the object plays a vital role in evolving the object path track through its kinematics for accelerating detection accuracy, and is formulated as follows:

$$\phi\left(V_i^\tau | M_i^\tau\right) = \exp\left(V_i^\tau; \Lambda_i^\tau, \widehat{V}_i^{\tau|\tau-1}\right) \tag{12}$$

where $\widehat{V}_i^\tau$ is an extended object measurement, $v_i^s > d-1$ is a degree of freedom, σ_i is a scalar that defines the effect of object extension and d is a physical space dimension. where Λ_i^τ is used to measure the object origin based on the degree of association with other objects and space dimensionality, as follows:

$$\Lambda_i^\tau = 2\left((d+2) + \frac{1}{\sigma_i^2}\left(\frac{v_i^s\left(\sigma_i+1\right)\left(\sigma_i-2\right)}{\left(\sigma_i+v_i^s\right)}\right)\right) \tag{13}$$

The extended object shape data measurement plays a vital role in extended object detection and classification at each frame, and the measurement is amended at each iteration based on the object state, $\mathbb{R}_i$ invertible matrix and object characteristics, as follows:

$$\widehat{V}_i^\tau = \frac{v_i^s}{\sigma_i - 1}\left(v_i^\tau - d\right)\mathbb{R}_i \times \widehat{V}_i^\tau \mathbb{R}_i^T \tag{14}$$

The pre-estimation of extended object attributes is essential to predict the presence of an object at each frame by considering the eccentricity of the ellipse-bounded range, angle of direction and extent object semi-coordinates. The probability of the predicted attributes $\phi\left(\gamma_i^\tau | M_i^\tau\right)$ of the object is derived as follows:

$$\phi\left(\gamma_i^\tau | M_i^\tau\right) = \exp\left(\gamma_i, \widehat{\phi}_i^p, \widehat{\phi}_i^s, (0, +\infty)\right) \tag{15}$$

where,

$$\begin{aligned} \widehat{\phi}_p &= \widehat{\psi}_i^{\tau,\tau-1} \times \widehat{\phi}_{i,p}^{\tau|\tau-1} \\ \widehat{\phi}_s &= \left(\widehat{\psi}_i^{\tau,\tau-1}\right)^2 \times \widehat{\phi}_{i,s}^{\tau|\tau-1} + \Delta^{Q_p} \end{aligned} \tag{16}$$

where $\widehat{\psi}_i^{\tau,\tau-1}$ is extended object state parameter derived as follows.

$$\widehat{\psi}_i^{\tau,\tau-1} = \frac{\widehat{b}_{\tau,\tau-1}}{\sqrt{1-\left(\widehat{e_i}\cdot\cos\left(\Theta_i-\theta_i\right)\right)^2}} \tag{17}$$

$$\widehat{e_i} = \sqrt{1 - \frac{\left(\widehat{b}_i^{\tau,\tau-1}\right)^2}{\left(\widehat{a}_i^{\tau,\tau-1}\right)^2}},\ \therefore\ 0 < \widehat{e_i} < 1 \tag{18}$$

where e_i, Θ_i, $\widehat{b}_i^{\tau,\tau-1}$, $\widehat{b}_i^{\tau,\tau-1}$ refers a eccentricity of ellipse, inflate direction angle, semi-axes defined by $\widehat{V}_i^\tau$.

3) Extended Object Measurement Update: The Bayesian method formulates the PDF of an extended object with three factors (object localisation density, extent prediction density and random scalar density) for effective identification and tracking. The random scholar denotes the size of the extended object at each frame. The three factors' product formulates and updates the extent of object measurements for each frame within the time window, as follows:

$$\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^\tau, M_i^{\tau|\tau-1}\right) \triangleq \eta\left(v_i^\tau\right)\eta\left(V_i^\tau\right)\eta\left(\gamma_i^\tau\right) \tag{19}$$

The forecasting density of each factor is defined as follows and the object localization density is

$$\eta\left(v_i^\tau\right) \propto \int \ln\left(\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^{\tau|\tau+1}, M_i^{\tau|\tau-1}\right)\right) \times \eta\left(V_i^\tau\right)\eta\left(\gamma_i^\tau\right) dV_i^\tau d\gamma_i^\tau \tag{20}$$

The predicted extension of object density is

$$\eta\left(V_i^\tau\right) \propto \int \ln\left(\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^{\tau|\tau+1}, M_i^{\tau|\tau-1}\right)\right) \times \eta\left(\gamma_i^\tau\right)\eta\left(v_i^\tau\right) d\gamma_i^\tau dv_i^\tau \tag{21}$$

The elaborated measurements random scalar density is derived as follows

$$\eta\left(\gamma_i^\tau\right) \propto \int \ln\left(\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^{\tau|\tau+1}, M_i^{\tau|\tau-1}\right)\right) \times \eta\left(V_i^\tau\right)\eta\left(v_i^\tau\right) dV_i^\tau dv_i^\tau \tag{22}$$

The path joint prediction density is factorized as follows

$$\phi\left(v_i^\tau, V_i^\tau, p_i^\tau, \gamma_i, M_i^{\tau|\tau-1}\right) = \phi\left(m_i^\tau|v_i^\tau, V_i^\tau, \gamma_i^\tau\right) \times\phi\left(\left(v_i^\tau, V_i^\tau\right), p_i^\tau, \gamma_i|M_i^\tau\right) \quad (23)$$

where, the $\phi\left(\left(v_i^\tau, V_i^\tau\right), p_i^\tau, \gamma_i|M_i^\tau\right)$ is derived from Eq. 11, and $\phi\left(p_i^\tau|v_i^\tau, V_i^\tau, \gamma_i^\tau\right)$ is derive as follows

$$\phi\left(m_i^\tau|v_i^\tau, V_i^\tau, \gamma_i^\tau\right) \propto \exp\left(\widehat{m_i}; d_i v_i + \Theta_i\gamma_i, \frac{\mathbb{R}_i V_i^\tau \mathbb{R}_i^T}{Q_n^\tau}\right) \times w\left(M_i; q_i - 1, \mathbb{R}_i V_i^\tau \mathbb{R}_i^T\right) \quad (24)$$

where, Θ_i is object measurement inflate direction, Q_i^τ is an independent of v_i,

$$\widehat{m_i} = \frac{1}{Q_q}\sum_{q=1}^{Q_q} m_{i,q}^\tau,$$

$$\widehat{m_i} = \sum_{q=1}^{Q_q}\left(m_{i,q}^\tau - \widehat{m}_{i,q}\right)\left(m_{i,q}^\tau - \widehat{m}_{i,q}\right)^T \quad (25)$$

and Eq. 23 is derived as follows

$$\begin{aligned}
&\log\left(v_i^\tau, V_i^\tau, p_i^\tau, \gamma_i, M_i^{\tau|\tau-1}\right)\\
&\propto \log\left|V_i^\tau\right| + \widehat{m_i^\tau} - d_i^\tau v_i^\tau - \Theta_i^\tau\gamma_i^\tau\\
&\times\left(\frac{\mathbb{R}_i V_i^\tau \mathbb{R}_i^T}{\vartheta_i}\right) + \left(\vartheta_i^\tau - d_i^\tau - 1\right)\times\log\left|M_i^\tau\right|\\
&-tr\left(\mathbb{R}_i V_i^\tau \mathbb{R}_i^{-T} M_i^\tau\right)\\
&-\log\left(m_{i|i-1}\right) + \left(v_i^\tau - \widehat{v}_{i|i-1}\right)^T\\
&-0.5\left(\vartheta_i^\tau + u_i^\tau + 1\right)\log\left|V_i^\tau\right| - 0.5tr\left(\widehat{V}_i^\tau V_i^{\tau-1}\right)\\
&-0.5\log\left(\widehat{\phi}_s^\tau + \widehat{\phi}_s^\tau\left(\gamma_i^\tau - \widehat{\phi}_p^\tau\right)^2\right)(0, +\infty)
\end{aligned} \quad (26)$$

The v_i is formulated by substituting Eq. 26 in eq. 20 to estimate the expected object detection density as follows

$$\eta\left(v_i^\tau\right) = \prod_{i=1}^{N}\left[\begin{array}{c}\phi\left(\left(v_i^\tau|v_i^{\tau-1}\right), \left(p_i^\tau, p_i^{\tau-1}\right)\right)\\ \times\phi\left(p_i^\tau|p_i^{\tau-1}\right)\end{array}\right] \quad (27)$$

The V_i is formulated by substituting Eq. 26 in eq. 21 to estimated the density of extended object for effective tracking as follows and the mathematically derivation is amended in Appendix.

$$\eta\left(V_i^\tau\right) = \exp\left(V_i^\tau; \Lambda_i^\tau, \widehat{V}_i^{\tau|\tau-1}\right) \quad (28)$$

The γ_i is formulated by substituting Eq. 26 in eq. 22 to measure the random scalar density as follows

$$\eta\left(\gamma_i^\tau\right) = \exp\left(\gamma_i, \widehat{\phi}_i^p, \widehat{\phi}_i^s, (0, +\infty)\right) \quad (29)$$

The affinity model estimates the object structure (size, shape and motion pattern) to localise the extended object by algorithm 1. The hypothesis data analysis component eliminates the ambiguous structure detections and redundant bboxs by analysing the successive frames to diminish the workload. Consequently, tracking objects frame-by-frame based on object-centric measurements is carried by algorithm 2, where

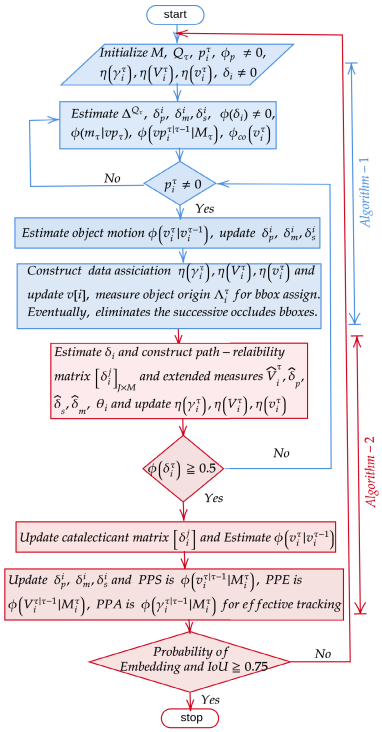


Fig. 3. IMOT flowchart.

the estimation of path reliability helps to validate the tracking path accuracy. Moreover, algorithm 1 assesses the object presence with accurate measurements where line 1 initialises the entail factors, and line 2 is a loop to check every object per frame. Line 3 measures the object probability density factor. Line 4 is a loop used to assess object motion per frame based on specific parameters that are part of the affinity method; for example, line 5 measures the probability density factor for an extended object. Lines 6–10 assess the comprehensive object measurements (shape, motion, expected and scalar density) described in our approach. Eventually, successive occluded boxes are eliminated to preserve the computation resources. Algorithm 2 optimises the object tracking uncertainties based on the BDAI model through object affinity measurements and a data association model. Line 1 initialises the entail factors, and line 2 is a loop for the iterative analysis of object states. Line 3 measures the reliability of each object path, and the outcomes are updated in the catalecticant matrix. Line 4 helps to analyse hypothetical data. Line 5 assesses the object's dynamic motion to measure the extended state using algorithm 1. The path probability of the extended object is estimated using line 6, and the core programming flow is depicted in Fig. 3. This algorithm resolves the MOT issue by deriving a novel data association method with a variant Bayesian mechanism.

4) Complexity Analysis: Let us assume that the algorithm 1 & 2 diverges into three sub-modules. First, the iterative process of data ambiguous analysis impacts the targeted object accuracy, and the complexity of this process is expressed as $O\left(n^2\right)$. Sorting the objects based on the affinity model is significant, and the complexity of this process is expressed as $O\left(n\log_2 n\right)$. Third, updating the extended object measurements of each frame helps consolidate the accuracy of the data association model. The complexity of this process is expressed as $O\left(n^3\right)$.

Algorithm 1 Reliable Object Filtering Model

input : Object-state set $v[i]$,
$M = \{m_{1,\tau}, m_{1,\tau}, \ldots, m_{Q,\tau_f}\}$, τ, Q_τ
output: Object set filtering without redundancy through accurate measurements

Let $\tau \neq 0$, $Q_\tau = \sum_{i=1}^{N} Q_i^s + Q_i^c$;
for *each* $v_i \in N$ **do**
  Estimate Probability density factor
  $\phi_p = \exp\left(-\frac{1}{p}\sum_{l=1}^{L}\sum_{i=1}^{N}\sqrt{\mathbb{Z}_{i,l}^{j}(\delta_i)\cdot\mathbb{Z}_{i+1,l}^{j+1}(\delta_{i+1})}\right)$;
  while $p_i^\tau \neq 0$ **do**
    Assessment of prognosticated object state
    $\phi\left(v_i^\tau, V_i^\tau\right) = \exp\left(V_i^\tau; \upsilon_i^\tau, \widehat{V}_i^\tau\right)$;
    Estimate extended object shape
    $\widehat{V}_i^\tau = \frac{\upsilon_i^s}{\sigma_i - 1}\left(\upsilon_i^\tau - d\right)\mathbb{R}_i \times \widehat{V}_i^\tau \mathbb{R}_i^T$;
    Dynamic update of each object states:
    Estimate object PDF precedence
    $\underset{\tau}{\phi}\left(\upsilon p_\tau | M_{1:\tau-1}\right) =$
    $\int_1 \phi\left(M_\tau | \upsilon p_\tau\right) \times \phi\left(\upsilon p_i^{\tau|\tau-1} | M_i^\tau\right) dx$;
    Update and estimate the extended object expected density as $\eta\left(V_i^\tau\right) \propto$
    $\int \ln\left(\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^{\tau|\tau+1}, M_i^{\tau|\tau-1}\right)\right) \times$
    $\eta\left(\gamma_i^\tau\right)\eta\left(v_i^\tau\right) d\gamma_i^\tau dv_i^\tau$;
    Estimate elaborated measurements with a random scalar density as $\eta\left(\gamma_i^\tau\right) \propto$
    $\int \ln\left(\phi\left(v_i^\tau, V_i^\tau, \gamma_i^\tau, p_i^{\tau|\tau+1}, M_i^{\tau|\tau-1}\right)\right) \times$
    $\eta\left(V_i^\tau\right)\eta\left(v_i^\tau\right) dV_i^\tau dv_i^\tau$;
  end
  Update $v[i] \leftarrow \left\{\eta\left(v_i^\tau\right) \bigcap \eta\left(V_i^\tau\right) \bigcap \eta\left(\gamma_i^\tau\right)\right\}$;
end

The system complexity is

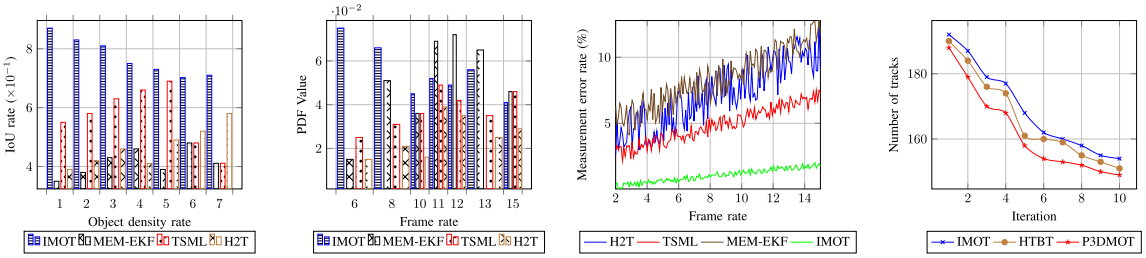
$$O\left(n^2\right) + O\left(n\log_2 n\right) + O\left(n^3\right) \tag{30}$$

IV. Experimental Results

The PC runs 64-bit Ubuntu 18.04.5 LTS on the Intel Core i7-10700 CPU 3.80GHz×16 and NVIDIA GeForce. The MATLAB[1] driving scenario designer provides a simple method of generating road users and their trajectory scenarios, and the MATLAB GPU Coder is used to generate the target code for effective processes over NVIDIA Jetson. The second set of simulations is conducted with TensorFlow, a mmdetection3D platform, for performance cross-validation. The performance of our system is assessed based on the nuScenes dataset [29], and it has complete 360° coverage because of 32-beam LiDAR, six cameras and radars, which provide ground-truth knowledge regarding the targeted object to monitor multiple objects over each frame. The nuScence dataset diverges into 65% of the training set, 20% of the test set and 15% of the validation set. The learning process is controlled by tuning the hyperparameters through the evaluation process based on the validation set, which takes place for each 80-epoch iteration. The proposed system's performance at the device is analysed based on *Jetson Xavier NX* with mode *15W & 6-cores* and mapping module frequency of approximately *1 Hz*; the results are visualised with *Rviz*, which is a robot operating system tool. The dataset has 1000 scenes with 3D bounding boxes, which are annotated at *2Hz* frequency. The 3D bounding boxes are effectively filtered using 20s long scenes based on the classes. The object tracking accuracy and speed are measured based on the track velocity in terms of frames per second (FPS). The multiplicative error model-extended Kalman filter [30], target-specific metric learning (TSML) [31], hypergraphs for multi-object tracking (H2T) [32], HTBT method and P3DMOT model are considered to assess the performance of our proposed system based on the following metrics: ground truth (GT), mostly tracked (MT), mostly lost (ML), false positive (FP), false negative (FN), number of identity switches (IDS), multiple objects tracking precious (MOTP), and multiple object tracking accuracy (MOTA) measurements, as listed in Table I, for both NuScences and real-time datasets.Our data association model is inspired by robust object detection by formulating occluded targets for effective object tracking. Unlike state-of-the-art approaches, the data association model is less complex due to its accurate affinity measurements and efficient matrix construction of extent objects. Note that the occluded target identity assessment and assignment are based on the association link between the detected object and the occluded target. If there is no link, then it is treated as a new target for accomplishing tracking efficiency. The target is eliminated when it vanishes for consecutive frames or continues occlusion. In our simulation, the average threshold value is set as 0.5 to enhance the data association model performance. Table II presents the detailed IMOT performance of each class. For example, bicycle, bus, car, motorcycle, pedestrian, trailer and truck have 33.8%, 66.6%, 73.2%, 57.6%, 68.9%, 61.1% and 45.1% MOTA, respectively.

Fig. 4 illustrates object measurement and tracking analysis based on the frame rate. An object dimensionality measurement is essential to track objects in a real-time environment. The IoU parameter helps validate the object-centric measure to ensure targeted object detection accuracy. Fig. 4(a) illustrates the intersection over union (IoU) rate analysis as per the object density of each frame. The IoU rate of IMOT diminishes as the object density rate drastically increases. However, on average, the IoU rate of our approach is determined as $> 0.7\%$ due to the effective measurement of extended objects through the data association method to localise the object for effective tracking. MEM-EKF achieves a moderate IoU rate compared to the other two approaches. When the density rate is < 3, IMOT achieves an accurate IoU measure because of few object detections. However, the 3-extent approaches do not achieve efficient measurements even in the same circumstances.Fig. 4(b) illustrates the PDF impact to track the extended object based on their state measurements per frame. Our system achieves a higher object detection probability rate than the state-of-art approaches. Fig. 4(c) shows the measurement error rate of all approaches, where IMOT achieves a lower error rate

[1]https://de.mathworks.com/help/driving/ref/drivingscenariodesignerapp.html



(a) IoU rate analysis as per object density of each frame (b) PDF rate analysis for each frame (c) Measurement error rate analysis (d) Number of tracks

Fig. 4. Object measurement and tracking comparative analysis with respect to frame rate.

Algorithm 2 BDAI Model for Continues Object Tracking

input : Object-state set $v[i]$, $V_{1,\tau}^{T}$, τ
output: Updated object path
Let $\tau \neq 0$, $\delta_i \neq 0$, $[\delta_i^j]_{J\times M} \neq 0$, $\phi(\delta_i) \neq 0$;
for *each* $v_i \in N$ **do**
Estimate existed objects path reliability $\delta_i = {}^{1}/_{\hbar_i}\left(\sum_{f_s=1}^{fr\in[f_s,f_e]} \phi(\delta_i, \upsilon_i)\right)\cdot a_i^s \cdot \varepsilon$;
Assess each path quality with Inflate based data association $\phi(\delta_i) = \prod_{i=1}^{N} \phi_p^i \cdot \phi_s^i \cdot \phi_m^i$ and Estimate expected object motion with Eq. 4;
Estimate dynamic motion of object as per object states measurement using Algorithm 1;
$\phi\left(v_i^{\tau}|v_i^{\tau-1}\right) = \left(\frac{\delta_i^j+\delta_{i+1}^{j+1}}{\delta_{i,i+1}^{j,j+1}}\right) - 1$;
Construct correlation catalecticant matrix $[\delta_i^j]_{J\times M}$;
Where, $\delta_i^j = \frac{1}{\hbar_i}\log\left(\phi\left(\delta_{i,h}, \upsilon_i\right)\right)$;
Track the object with updated object status measurements;
end

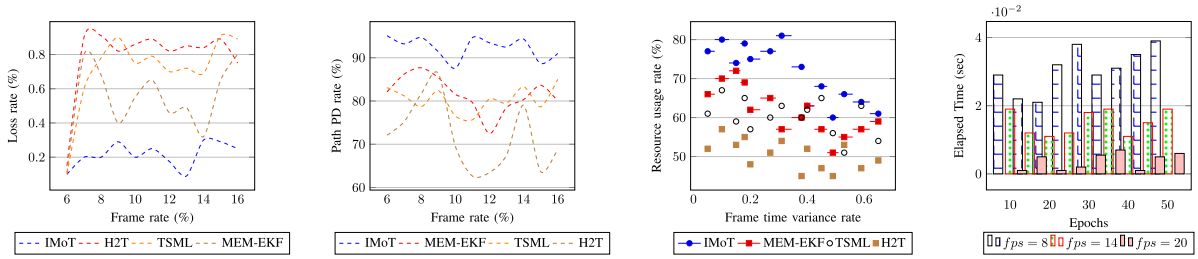
than state-of-the-art approaches. Nevertheless, it increases to 2% as the object density rate increases. However, the error rate is mitigated due to the effectiveness of the BDAI model by avoiding redundant measurements based on an encoded filter. The state-of-the-art models have achieved a $\geq$ 10% error rate, which is inappropriate for deploying the object-tracking method on the vehicle node. Fig. 4(d) illustrates several detected object tracks at each iteration. In our simulation, continuously occluded objects are eliminated, and the object identification ratio decreases as the iteration count increases. IMOT achieves a lower tracking error rate than state-of-the-art approaches because the affinity model is formulated based on the Bayesian approach.

A comparative analysis of the loss rate and probability of object detection (PD) rate is illustrated in Fig 5. The noise and redundancy removal policy causes loss of the entail data because of the inadequate extended object measurement system, which we resolve using the BDAI model. As observed in Fig. 5(a), the IMOT approach achieves a lower loss rate than state-of-the-art approaches because of the extended Bayesian approach formulation; however, MEM-EKF achieves an average better loss rate than TSML and H2T. The Bayesian approach effectively diminishes the loss rate due to an accurate measurement system design to measure extended object characteristics, such as size and shape (part of affinity). Subsequently, Fig. 5(b) shows the object or path detection rate probability concerning the FPS factor. Compared to other approaches, IMOT achieves a PD rate of more than 90% because the catalecticant square matrix is considered to update the state of each object along with path reliability measurements. In most cases, the reliability rate is [0.5, 1]. The H2T model achieves a lower prediction rate due to inadequate computation complexity than the TSML and MEM-EKF models. Figs. 5(c) and 5(d) illustrate the Jetson device resource usage rate in terms of CPU cycles per second concerning the variance of frame time and elapsed time to execute the service. IMOT achieves a lower resource usage rate than other approaches. Fig. 6 shows the experiment setup and extended object measurements, and Fig. 6(a) illustrates the detected object path concerning each approach. The original object is detected at frame 15, where the IMOT measured area is appropriate since the IoU rate is $>$ 0.75 and the HTBT, P3DMOT and H2T models achieve average better measurements of inadequate object detection ratio than the remaining two approaches. Each colour interprets the accuracy of every state-of-the-art approach concerning the proposed approach. Fig. 6(b) represents the OS1-64 channel LiDAR setup with RTX3090 to collect real-time data.

In simulation 2, the performance of the proposed approach is analysed with the HTBT and P3DMOT approaches. The

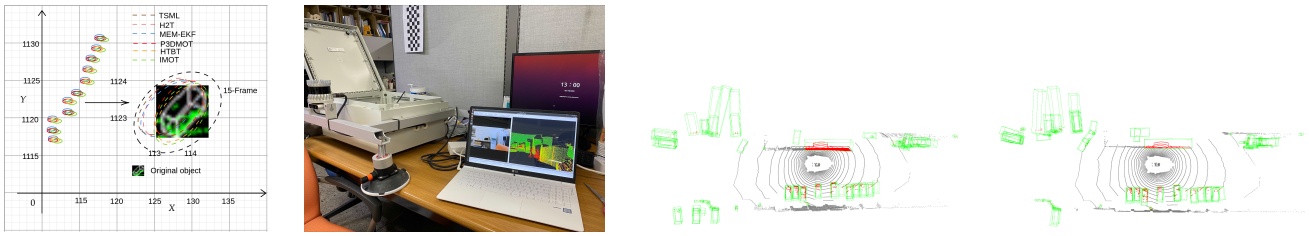
TABLE I
QUANTITATIVE COMPARISON OF 3D MOT EVALUATION RESULTS ON NUSCENES DATASET AND REAL-TIME DATASET

#item	IMOT	HTBT	P3DMOT	MEM-EKF	TSML	H2T
NuScences Dataset evaluation values						
GT	**17081**	17081	17081	17080	17081	17081
MT	**5774**	5278	5063	5213	4516	5319
ML	**1476**	2094	1986	1683	1984	1966
FP	**14817**	17877	17443	14714	18355	16469
FN	**24191**	24013	26430	25894	31057	24092
IDS	**141**	288	1042	1869	1851	557
MOTA	**0.5809**	0.5536	0.5265	0.5354	0.4737	0.5450
MOTP	**0.279**	0.302	0.297	0.298	0.361	0.286
Lab Dataset with 64-beam LiDAR						
GT	**1510**	1470	2950	6800	3500	3900
MT	**752**	635	537	529	334	139
ML	**183**	231	260	296	160	345
FP	**36**	44	85	134	116	117
FN	**39**	61	79	457	349	261
IDS	**3**	1	82	15	12	5
MOTA	**0.948**	0.927	0.916	0.910	0.863	0.901
MOTP	**0.051**	0.072	0.083	0.089	0.136	0.098
Avg-FPS	**21.1**	20	19	8.47	12.95	15.8



(a) Frame rate vs Loss rate (b) Path PD rate vs frame rate (c) Resource usage vs Frame rate (d) Elapsed time vs Epochs

Fig. 5. Comparative analysis of Loss rate and probability of object detection (PD) rate.



(a) Extended object measurement analysis (b) OUSTER-64 beam LiDAR framework (c) Frame-7292 (d) Frame-7298

Fig. 6. OUSTER-64 beam LiDAR framework for extended object measurement analysis.

TABLE II

3D MOT EVALUATION RESULTS ON NUSCENES DATASET

	Bicyle	Bus	Car	Motorcycle	Pedestrian	Trailer	Truck	Total	Avg
GT	2186	1701	68518	1945	34010	2566	8639	119565	**17080.7**
MT	76	61	3565	85	1597	95	295	**5774**	824.8
ML	85	27	856	45	325	36	102	**1476**	210.8
FP	256	56	8062	211	3455	246	2531	**14817**	2116.7
FN	1190	512	10215	610	7012	751	2201	**22491**	3213
IDS	1	0	54	2	81	0	3	**141**	20.1
MOTA	0.338	0.666	0.732	0.576	0.689	0.611	0.451	4.066	**0.5809**
MOTP	0.119	0.321	0.211	0.247	0.283	0.423	0.351	1.955	**0.2792**

TABLE III

QUANTITATIVE RUN-TIME ANALYSIS

Model	#N	Affinity Estimator (ms)	Track Association (ms)	mAP	mATE(m)
HTBT	15	12.21	0.43	0.651	0.275
P3DMOT		13.91	0.77	0.630	0.291
MEMEKF		16.05	0.93	0.543	0.300
IMOT		**11.99**	**0.16**	**0.692**	**0.238**
HTBT	25	12.73	0.51	0.637	0.302
P3DMOT		14.21	0.81	0.601	0.326
MEMEKF		16.93	0.98	0.529	0.337
IMOT		**12.19**	**0.19**	**0.710**	**0.245**

TABLE IV

JESTSON XAVIER NX PERFORMANCE ANALYSIS

Method	precision	Average Latency (ms)	Average Power (W)	Average GPU Usage (%)	mAP(%)
IMOT	Float32	**25.32**	**9.73**	**49.63**	**69.12**
MEM		29.01	13.62	70.56	51.3
P3DMOT		27.91	14.85	55.89	58.9
HTBT		52.3	13.91	69.01	61.5
IMOT	Float16	**26.70**	**10.55**	**53.58**	**67.9**
MEM		36.9	14.63	62.45	52.8
P3DMOT		49.65	12.53	59.91	62.8
HTBT		58.36	11.95	52.63	64.3

average runtime performance analysis report is updated based on several tracking objects (#N) in Table III. The run time is measured with NVIDIA GeForce RTX-3090 GPU. Our approach consumes an average of 169.71 ms to train single pair frames, and the affinity estimation time is 11.99 ms, track association time is 0.16 ms and mATE is 0.238. When $N = 25$, the proposed system consumes the average run time as follows: the affinity estimation time is 12.19 ms, the track association time is 0.19 ms and mATE is 0.245. Table IV shows the embedded device performance analysis results. Our approach achieves low usage of GPU (49.63%) and power (9.73 w) and exhibits latency with 69.12% mean average precision, which is quite optimal due to the system's low complexity for tracking service execution. Subsequently, the HTBT approach achieves adequate resource usage and more effective performance than the other two approaches.

Fig. 6(c) and 6(d) show tracking results of our proposed method. In our simulation, continuous identification of occluded objects is omitted and the tracking of these objects is halted. It can be seen that the background points are filtered effectively. However, the points of the stopped vehicle are also detected at the initial frames, but in further successive frames, these points are also effectively tracked and eliminated to achieve targeted accuracy.

V. CONCLUSION

This document describes an IMOT system based on intelligent vehicle orchestration for continuous object tracking. LiDAR sensors generate considerably extended object measurements during target assessment. The intelligent vehicle is responsible for assessing the decision of the object status by examining the measured data. The proposed BDAI model regulates the computation service based on the Bayesian approach. It plays a vital role in achieving 58.09% accuracy with 20 FPS by avoiding ambiguous detection responses based on the box object filter method. The theoretical proof of the box object filter is derived based on binomial expansion, and comprehensive enactment equations are derived for linear motion analysis to cope with the measurement models. Our model outcomes indicate that our method measures MOTA, MOTP, mAP and mATE as 0.5809, 0.279, 0.710 and 0.245, respectively. The Jetson Xavier NX consumes 49.63% GPU and 9.37% average power and exhibits 25.32 ms latency as compared to other approaches.

The designed model suits lightweight cyber-physical system (CPS) frameworks because of its low complexity and computational resources. Generally, real-time scenarios are unpredictable, as the extended object measurement count

is unknown. For instance, the *mmWave* sensing mechanism breeds massive measurements for indoor object localisation. In this regard, deploying the proposed system in a real-time environment is a challenge that we will consider in future work for continuous object detection and tracking with embedded devices by mapping the requirements of a lightweight mechanism.

APPENDIX

V_i is formulated by substituting 26 in eq. 21; such that,

$$\eta\left(V_i^{\tau}\right) = \chi\left(V_i^{\tau}|\widehat{V}_i^{\tau}; \upsilon_i^{\tau}\right) \tag{31}$$

V_i is derived as follows

$$\begin{aligned}
log\left(\eta\left(v_i^{\tau}\right)\right) &\propto \int \ln\left(\phi\left(v_i^{\tau}, V_i^{\tau}, p_i^{\tau}|p_{i+1}^{\tau}\right)\right) \\
&\quad \times \eta\left(V_i^{\tau}\right)\eta\left(\gamma_i^{\tau}\right) dV_i^{\tau} d\gamma_i^{\tau} \\
&= -0.5 \cdot \vartheta_i \times \left(\widehat{p_i} - d_i v_i + \Theta_i \gamma_i\right) \mathbb{R}_i V_i^{\tau} \mathbb{R}_i^T \\
\Gamma\left[\gamma_i^{\tau}\right] &= \int \gamma_i^{\tau} \eta\left(\gamma_i^{\tau}\right) d\gamma_i^{\tau} = \Gamma_{\eta(\gamma_i^{\tau})} \\
&= -0.5 \cdot \vartheta_i \times \left(\widehat{p_i} - d_i v_i + \Theta_i \gamma_i \right. \\
&\qquad \left. + \Theta_i \Gamma\left[\gamma_i^{\tau}\right] - \Theta_i \Gamma\left[\gamma_i^{\tau}\right]\right)^T \\
&\quad \times \mathbb{R}_i \Gamma\left[V_i^{\tau}\right] \mathbb{R}_i^T \\
&= -0.5 \cdot \vartheta_i \times \left(\widehat{p_i} - d_i v_i + \Theta_i \gamma_i \right. \\
&\qquad \left. + \Theta_i \left[\gamma_i^{\tau} - \Gamma\left[\gamma_i^{\tau}\right]\right]\right)^T \\
&\quad \times \mathbb{R}_i \Gamma\left[V_i^{\tau}\right] \mathbb{R}_i^T \\
&= -0.5 \cdot \vartheta_i \times \left(\widehat{p_i} - d_i v_i + \Theta_i \gamma_i\right)^T \mathbb{R}_i \Gamma\left[V_i^{\tau}\right] \mathbb{R}_i^T \\
\log \eta\left(v_i^{\tau}\right) &= -0.5\left(v_i^{\tau} - \Gamma\left[v_i^{\tau}\right]\right)^T
\end{aligned}$$

Such that,

$$\eta\left(v_i^{\tau}\right) = \Upsilon\left(v_i^{\tau}|\widehat{v}_i^{\tau}, \phi\left(v_i^{\tau}|\widehat{v}_i^{\tau}\right)\right) \tag{32}$$

The rest of the two parameters (V_i, γ_i) are formulated similarly.

REFERENCES

[1] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5657–5671, Nov. 2012.

[2] S. Aradi, "Survey of deep reinforcement learning for motion planning of autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 740–759, Feb. 2022.

[3] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1638–1653, Apr. 2016.

[4] M. Feldmann and W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1042–1059, Jul. 2012.

[5] A. Zea, F. Faion, M. Baum, and U. D. Hanebeck, "Level-set random hypersurface models for tracking nonconvex extended objects," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 2990–3007, Dec. 2016.

[6] W. Aftab, R. Hostettler, A. De Freitas, M. Arvaneh, and L. Mihaylova, "Spatio-temporal Gaussian process models for extended and group object tracking with irregular shapes," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2137–2151, Mar. 2019.

[7] S. Yin, J. H. Na, J. Y. Choi, and S. Oh, "Hierarchical Kalman-particle filter with adaptation to motion changes for object tracking," *Comput. Vis. Image Understand.*, vol. 115, no. 6, pp. 885–900, Jun. 2011.

[8] M. S. Mekala, P. Rizwan, and M. S. Khan, "Computational intelligent sensor-rank consolidation approach for industrial Internet of Things (IIoT)," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2121–2130, Feb. 2023.

[9] H. Bao, M. Lin, and Z. Chen, "Robust visual tracking based on hierarchical appearance model," *Neurocomputing*, vol. 221, pp. 108–122, Jan. 2017.

[10] W. Lu, Z. Zhou, L. Zhang, and G. Zheng, "Multi-target tracking by non-linear motion patterns based on hierarchical network flows," *Multimedia Syst.*, vol. 25, no. 4, pp. 383–394, Aug. 2019.

[11] S. Sun, N. Akhtar, H. Song, A. S. Mian, and M. Shah, "Deep affinity network for multiple object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 104–119, Jan. 2019.

[12] W. Zhang and X. He, "A robust deep affinity network for multiple ship tracking," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–20, 2021.

[13] B. Xu, M. Lu, J. Shi, J. Cong, and B. Nener, "A joint tracking approach via ant colony evolution for quantitative cell cycle analysis," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 6, pp. 2338–2349, Jun. 2021.

[14] Y. Xie, Y. Huang, and T. L. Song, "Iterative joint integrated probabilistic data association filter for multiple-detection multiple-target tracking," *Digit. Signal Prog.*, vol. 72, pp. 232–243, Jan. 2018.

[15] S.-H. Bae and K.-J. Yoon, "Confidence-based data association and discriminative deep appearance learning for robust online multi-object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 595–610, Mar. 2018.

[16] J. Gwak, "Multi-object tracking through learning relational appearance features and motion patterns," *Comput. Vis. Image Understand.*, vol. 162, pp. 103–115, Sep. 2017.

[17] M. Aladem and S. A. Rawashdeh, "A combined vision-based multiple object tracking and visual odometry system," *IEEE Sensors J.*, vol. 19, no. 23, pp. 11714–11720, Dec. 2019.

[18] H. Wu, W. Han, C. Wen, X. Li, and C. Wang, "3D multi-object tracking in point clouds based on prediction confidence-guided data association," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5668–5677, Jun. 2022.

[19] P. Emami, P. M. Pardalos, L. Elefteriadou, and S. Ranka, "Machine learning methods for data association in multi-object tracking," *ACM Comput. Surv.*, vol. 53, no. 4, pp. 1–34, Jul. 2021.

[20] S. You, H. Yao, and C. Xu, "Multi-target multi-camera tracking with optical-based pose association," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 8, pp. 3105–3117, Aug. 2021.

[21] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, and T.-K. Kim, "Multiple object tracking: A literature review," *Artif. Intell.*, vol. 293, Apr. 2021, Art. no. 103448.

[22] M. Yang, M. Pei, and Y. Jia, "Online maximum a posteriori tracking of multiple objects using sequential trajectory prior," *Image Vis. Comput.*, vol. 94, Feb. 2020, Art. no. 103867.

[23] X. Lin, C.-T. Li, V. Sanchez, and C. Maple, "On the detection-to-track association for online multi-object tracking," *Pattern Recognit. Lett.*, vol. 146, pp. 200–207, Jun. 2021.

[24] J. Xu, C. Bo, and D. Wang, "A novel multi-target multi-camera tracking approach based on feature grouping," *Comput. Electr. Eng.*, vol. 92, Jun. 2021, Art. no. 107153.

[25] H. Wang, S. Wang, J. Lv, C. Hu, and Z. Li, "Non-local attention association scheme for online multi-object tracking," *Image Vis. Comput.*, vol. 102, Oct. 2020, Art. no. 103983.

[26] Y. Huang, T. L. Song, W. J. Lee, and T. Kirubarajan, "Multiple detection joint integrated track splitting for multiple extended target tracking," *Signal Process.*, vol. 162, pp. 126–140, Sep. 2019.

[27] H. Sheng, "Hypothesis testing based tracking with spatio-temporal joint interaction modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 9, pp. 2971–2983, Sep. 2020.

[28] H.-K. Chiu, J. Li, R. Ambrus, and J. Bohg, "Probabilistic 3D multi-modal, multi-object tracking for autonomous driving," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 14227–14233.

[29] H. Caesar, "nuScenes: A multimodal dataset for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 11621–11631.

[30] S. Yang and M. Baum, "Tracking the orientation and axes lengths of an elliptical extended object," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4720–4729, Sep. 2019.

[31] B. Wang, G. Wang, K. L. Chan, and L. Wang, "Tracklet association by online target-specific metric learning and coherent dynamics estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 3, pp. 589–602, Mar. 2017.

[32] L. Wen, Z. Lei, S. Lyu, S. Z. Li, and M.-H. Yang, "Exploiting hierarchical dense structures on hypergraphs for multi-object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 10, pp. 1983–1996, Oct. 2016.

M. S. Mekala (Member, IEEE) received the Ph.D. degree from VIT University, India. He is currently working as an Assistant Professor with the School of Computing, Robert Gordon University, Aberdeen, U.K. He is a Former Post-Doctoral Researcher at RLRC LAB, Yeungnam University, gyeonnsan, Korea, and a Research Coordinator and Member of FEB Laboratory at KL University, india. He has published more than 20 peer-review research papers (indexed in SCI-SCIE). His research interests include service computing, intelligent machine vision, data communication, decision making system design, edge computing, CPS, IoT communication, and Reliability Analysis. He is serving as the Guest Editor of more than two special issues in various peer-reviewed journals. He is a recipient of Best Research award for two consecutive years in 2018 and 2019. He received the research scientist award in 2020. He is currently active reviewer of IEEE the *Internet of Things Journal*, IEEE TRANSACTION ON SERVER COMPUTING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTION ON ROBOTICS, IEEE TRANSACTION ON VEHICULAR TECHNOLOGY, *Measurement-Elsevier*, computer networks, neural computing and applications, Applied soft-computing Methods.

Gaurav Dhiman (Senior Member, IEEE) received the master's degree in computer applications and the Ph.D. degree in computer engineering from the Thapar Institute of Engineering and Technology, Patiala. He is currently working as an Assistant Professor with the Department of Computer Science, Government Bikram College of Commerce, Patiala. He is also an Adjunct Faculty with Chandigarh University and Graphic Era Deemed to be University. He was selected as an outstanding reviewer from Knowledge-Based Systems (Elsevier). He has published more than 200 peerreview research papers (*indexed in SCI-SCIE*) and ten international books. His research interest includes single, multi, many-objective optimization (bio-inspired, evolutionary and quantum), soft computing (type-1 and type-2 fuzzy sets), power systems, and change detection using remotely sensed high-resolution satellite data. His research articles can be found in Knowledge-Based Systems (Elsevier), Advances in Engineering Software (Elsevier), Applied Intelligence (Springer), Journal of Computational Science (Elsevier), Engineering with Computers (Springer), Applied Soft Computing (Elsevier), Engineering Applications of Artificial Intelligence (Elsevier), Journal of Ambient Intelligence and Humanized Computing (Springer), among others. His research can also be seen in http://www.dhimangaurav.com. He is serving as the Guest Editor of more than forty special issues in various peer-reviewed journals. He is also in Editorial Board Members of Current Chinese Computer Science (Bentham Science), Current Dissertations (Bentham Science), *Journal of Fuzzy Logic* and Modeling in Engineering (Bentham Science), *Journal of the Institute of Electronics and Computer* (IEC Science), *Journal of PeerScientist*, Artificial Intelligence Evolution, Acta Scientific Computer Sciences.

Wattana Viriyasitavat (Senior Member, IEEE) received the M.S. degree in computer engineering from Carnegie Mellon University, USA, in 2006, and the D.Phil (Ph.D.) degree in computer science from the University of Oxford in 2013. He is an Associate Professor at the Business Information Technology Division, Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Thailand. His research interests focus on trust, security in service workflows, specification language, blockchain, IoT, and compliance checking.

Ju H. Park (Senior Member, IEEE) received the Ph.D. degree in electronics and electrical engineering from Pohang University of Science and Technology (POSTECH), Pohang, Republic of Korea, in 1997. From May 1997 to February 2000, he was a Research Associate in Engineering Research Center-Automation Research Center, POSTECH. He joined Yeungnam University, Kyongsan, Republic of Korea, in March 2000, where he is currently the Chuma Chair Professor. He has published a number of articles in these areas. His research interests include robust control and filtering, neural/complex networks, fuzzy systems, multiagent systems, and chaotic systems. Since 2015, he has been a recipient of the Highly Cited Researchers Award by Clarivate Analytics (formerly, Thomson Reuters) and listed in three fields, Engineering, Computer Sciences, and Mathematics, in 2019 to2022. He is a Subject Editor, Advisory Editor, Associate Editor and Editorial Board Member of several international journals, including IET Control Theory and Applications, Applied Mathematics and Computation, *Journal of The Franklin Institute, Nonlinear Dynamics, Engineering Reports, Cogent Engineering*, the IEEE TRANSACTION ON FUZZY SYSTEMS, the IEEE TRANSACTION ON NEURAL NETWORKS AND LEARNING SYSTEMS, and the IEEE TRANSACTION ON CYBERNETICS. He is a fellow of the Korean Academy of Science and Technology (KAST).

Ho-Youl Jung received the Ph.D. degree in electronics engineering from the INSA de Lyon (Institute National des Sciences Appliquées de Lyon), France, in 1998. He is currently a Professor with the Department of Information and Communication Engineering, Yeungnam University, Korea. His Teaching and research interests include digital signal processing, computer vision, deep-learning, autonomous vehicle, computer graphics, control signal processing, and IoT.