Precise positioning and prediction system for autonomous driving based on generative artificial intelligence

Beichang Liu[1],*,[†], Guoqing Cai[1],[†], Zhipeng Ling[2], Jili Qian[3], Quan Zhang[4]

[1]Information Studies, Trine University, Phoenix AZ, USA
[2]Computer Science, University of Sydney, Sydney, Australia
[3]Information Studies, Trine University, Phoenix AZ, USA
[4]Information Studies, Trine University, Phoenix AZ, USA

*Corresponding author: alexfeynman2023@gmail.com
[†]Beichang Liu and Guoqing Cai contributed equally to this paper and should be considered as co-first authors.

Abstract. Self-driving systems collect vast amounts of data through a variety of sensors, including cameras, lidar, millimeter-wave radar, and more. This data needs to be processed in real time to identify obstacles such as roads, vehicles, pedestrians and make decisions accordingly. Therefore, this paper discusses the importance of accurate positioning and prediction system in automatic driving technology, and analyzes the performance of various positioning technologies in automatic driving applications.In addition, the paper explores the application potential of AI technology in autonomous driving and the prospect of combining advanced positioning and prediction systems with generative AI. Overall, this study highlights the importance of algorithm performance improvement and artificial intelligence technology in the development of autonomous driving technology, and provides new ideas and directions for the innovation and development of intelligent transportation systems in the future.

Keywords: Precise Positioning, Prediction System, Autonomous Driving, Generative Artificial Intelligence.

1. Introduction

Autonomous driving technology revolutionizes the automotive industry by harnessing computer and sensor technology to enable vehicles to operate without human intervention. The evolutionary journey of autonomous driving technology can be traced back to its initial phase, marked by the introduction of Adaptive Cruise Control (ACC) in the early 1990s. ACC, utilizing radar and laser sensors, monitored the distance and speed of preceding vehicles, autonomously adjusting the speed and direction of travel based on this data. Building upon the foundation established by ACC, introduced by Toyota in 2003 and subsequently adopted by various automotive manufacturers, these systems integrated [1] GPS, LiDAR, and image recognition technologies to monitor and analyze the vehicle's surroundings. In the realm of autonomous driving, positioning prediction plays a pivotal role, serving as the nexus between sensing and regulation modules. Positioned at the interface between these modules, it facilitates the exchange of crucial information such as target track state and road structure. This predictive capability empowers autonomous vehicles to anticipate and adapt to dynamic driving scenarios, enhancing safety and

efficiency on the road. With a comprehensive understanding of the evolution and significance of autonomous driving technology, the purpose of this research endeavor is to further refine and validate the positioning prediction component. By elucidating the intricate interplay between perception, prediction, and regulation modules, this study aims to enhance the predictive accuracy and responsiveness of autonomous vehicles, thereby advancing the state-of-the-art in autonomous driving technology.

2. Related work

2.1. Autopilot precise positioning

In December 2021, according to the international safety requirements UN-R157 "Automatic Lane Keeping System (ALKS)", the German Federal Motor Transport Authority believes that Mercedes-Benz's L3 level automatic driving system meets the regulations and approves the road, allowing L3 autonomous vehicles to be sold and on the road from the legal level, which is a breakthrough in mass production automatic driving technology. Currently, relying on a single positioning technology cannot meet the high precision required for autonomous vehicles. From the existing Oems to use the program, the basic use of multi-sensor fusion positioning technology, in addition to the use of integrated navigation modules and high-precision maps, will also choose visual SLAM, Lidar and other technologies. Because relative positioning cannot be used with the standard high-precision map, the coordinate system, data format, interface, and timeline of the two are completely different, and the standard high-precision map must be used with absolute positioning. [2] Therefore, the current mainstream positioning technology is GNSS, IMU, and high-precision map, which cooperate and complement each other to form a high-precision positioning system for automatic driving. Some car companies have begun to disassemble the integrated navigation box and integrate the GNSS module and IMU module into their domain controllers. When automatic driving is upgraded to L3+, L4/L5, the high-precision combined positioning module must reach the centimeter level, and needs to meet higher functional safety requirements, and it has the conditions to integrate into the automatic driving domain controller.

2.2. Autonomous driving environment awareness

The autonomous driving system consists of three main modules: perception, decision making and control. Roughly speaking, these three modules correspond to the eyes, brain, and limbs of biological systems. The sensory system (eyes) is responsible for understanding the information of the surrounding obstacles and roads, the decision-making system (brain) determines the next action based on the surrounding environment and set goals, and the control system (limbs) is responsible for executing these actions, such as steering, accelerating, [3] braking, etc. Further, the perception system includes two tasks: environment perception and vehicle positioning. Environmental awareness is responsible for detecting various moving and stationary obstacles (such as vehicles, pedestrians, buildings, etc.), and collecting various information on the road (such as driveable areas, lane lines, traffic signs, traffic lights, etc.). Vehicle positioning is based on environmental perception information to determine the location of the vehicle in the environment, which requires high-precision maps, as well as inertial navigation (IMU) and global positioning system (GPS) assistance. Therefore, in order to obtain more accurate three-dimensional information, Lidar has also been an important part of the autonomous driving perception system, especially for level L3/4 applications. Lidar data is a relatively sparse point cloud, which is very different from the dense grid structure of the image, so the algorithms commonly used in the image field need to be modified to apply the point cloud data. In addition, [4] Convolutional Neural networks (CNNS) in deep learning can also be improved to apply to sparse point cloud structures, such as PointNet or Graph Neural networks. In recent years, researchers have begun to use deep learning to replace classical radar signal processing from lower-level data, and have achieved similar liDAR perception effects through end-to-end learning.

2.3. Intelligent object detection and recognition

1) Traditional methods

In autonomous driving, identifying road elements such as roads, vehicles, and pedestrians, and then making different decisions is the basis for safe vehicle driving. The workflow for object detection and recognition is shown in Figure 1. Tesla uses a combination of wide-angle, medium-focus and telephoto cameras. The wide-angle camera has a viewing Angle of about 150° and is responsible for identifying a large range of objects in the nearby area. The medium focal length camera has a viewing Angle of about 50 and is responsible for identifying lane lines, vehicles, pedestrians, traffic lights and other information. The viewing Angle of the telephoto camera is only about 35°, but the recognition distance can reach 200~250m. It is used to identify distant pedestrians, vehicles, road signs and other information, and to gather road information more comprehensively through a combination of multiple cameras.

2) Methods based on deep learning

Compared with traditional localization object detection and recognition, deep learning requires training based on large data sets, but results in better performance. [5] Deep learning has more powerful feature learning and feature representation capabilities, by learning databases and mapping relationships, processing information captured by cameras into vector Spaces for recognition by neural networks.

3) Depth positioning prediction

In autonomous driving systems, the proper distance is important to ensure the safe driving of the car, so depth estimation from the image is required. The goal of depth estimation is to obtain the distance to the object and ultimately to obtain a depth map that provides depth information for a range of tasks such as [6]3D reconstruction, SLAM, and decision making, and the mainstream distance measurement methods on the market today are monocular, stereoscopic, and RGBD camera-based.

Since it needs to be compared with an established sample database in both the identification and estimation stages, it lacks self-learning capabilities, and the perceived results are limited by the database, and unlabeled targets are usually ignored, which leads to the problem of not being able to identify uncommon targets. At present, monocular camera is gradually becoming the mainstream technology of visual ranging due to its low cost, fast detection speed, ability to identify specific obstacle types, high algorithm maturity and accuracy.

3. Methodology and experimental design

The development of autonomous driving technology cannot be separated from accurate positioning and prediction systems. This section explores how to build accurate positioning and prediction systems based on generative AI techniques and design experiments to verify their effectiveness.

3.1. Methodology

Accurate and real-time positioning is critical for autonomous vehicles (AVs) to drive safely and efficiently. However, there has been a lack of comprehensive reviews comparing the real-time performance of different localization techniques across various hardware platforms and programming languages.Accurate and real-time positioning stands as a cornerstone for the safe and efficient operation of autonomous vehicles [8](AVs). To comprehensively assess the performance of various localization techniques across diverse hardware platforms and programming languages, this study embarks on a meticulous review and analysis. The research aims to bridge the gap in existing literature by investigating state-of-the-art positioning technologies and evaluating their suitability for AV applications. Central to this endeavor is the exploration of the Kalman Filter, a prevalent system state estimation method renowned for its effectiveness in AV positioning and prediction.

1) Kalman filtering

Kalman filter is a very popular system state estimation method, it is quite similar to probabilistic positioning, we learned before Monte Carlo positioning method, the main difference is that Kalman is to estimate a continuous state, while Monte Carlo divided the world into many discrete small pieces, as a result, Kalman gives us a unimodal distribution, Monte Carlo is a multi-modal distribution. Both methods are suitable for locating and tracking other vehicles. In fact, the particle filter is also suitable

for positioning and prediction, and the particle filter is continuously multi-peak distribution. The beauty of the Kalman filter is that it combines less-than-accurate sensor measurements with less-than-accurate motion predictions to get a filtered position estimate that is better than all the estimates from sensor readings or motion predictions alone. We build the state variable based on a Gaussian distribution, so we need two pieces of information at time k: the best estimate (i.e., the mean, often represented by μ elsewhere), and the covariance matrix.

$$\widehat{x_k} = \begin{bmatrix} position \\ velocity \end{bmatrix}, P_k = \begin{bmatrix} \Sigma_{pp} & \Sigma_{pv} \\ \Sigma_{vp} & \Sigma_{vv} \end{bmatrix} \tag{1}$$

We're only using position and speed here, and actually this state can contain multiple variables that represent whatever information you want to represent). Next, we need to predict the next state (k moment) based on the current state (k-1 moment). What about using a matrix to predict the position and velocity at the next moment? The following is a basic kinematic formula to express:

$$p_k = p_{k-1} + \Delta t v_{k-1} \tag{2}$$

In other words:

$$v_k = v_{k-1} \tag{3}$$

A prediction matrix to represent the state at the next time, but we still don't know how to update the covariance matrix. At this point, we need to introduce another formula, if we multiply every point in the distribution by the matrix A, what happens to its covariance matrix as depicted here? Here is the formula:

$$Cov(x) = \Sigma$$

$$Cov(Ax) = A \sum A^T \tag{4}$$

Combining equation (4) and (3), we get:

$$\hat{x}_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \hat{x}_{k-1} = F_k \hat{x}_{k-1} \tag{5}$$

The development of autonomous driving technology hinges upon the implementation of precise positioning and prediction systems. This section delineates the methodology employed to construct accurate positioning and prediction systems leveraging generative AI techniques, along with the experimental design devised to validate their efficacy.

3.2. Experimental design

1) Heterogeneous sensing systems are commonly used in robotics and autonomous vehicles to generate comprehensive environmental information. Commonly used sensors include various cameras [7], 2D/3D Lidar (LIGHT Detection and Ranging), radar (RAdio Detection and Ranging), The combination of these is mainly due to the fact that different sensors have different (physical) characteristics, and each class has its own advantages and disadvantages [6]. First, we describe the various sensors for efficient vehicle perception and positioning, and explain why these sensors were chosen, where to install them, and some of the trade-offs we made in system configuration.

2) In the experiment, we used a range of heterogeneous sensors, including multiple cameras, 2D/3D liDAR, and radar. These sensors are configured on experimental vehicles to achieve comprehensive perception and positioning of diverse driving scenarios in urban and suburban environments. By integrating these sensors into the ROS platform and conducting comprehensive data acquisition in urban and suburban environments, we are able to effectively simulate real-world autonomous driving scenarios, providing a solid foundation for reliability and repeatability of experimental results.
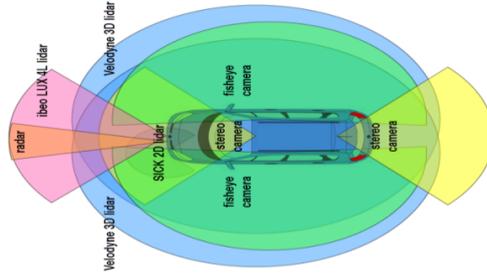
Figure 1. 3D map of perception sensor

Second, we propose a new AD dataset, entirely ROs-based, recorded by our platform in urban and suburban areas, with all sensors calibrated and data approximately synchronized (i.e., at the software level, except for two 3D liDAR's synchronized at the hardware level via communication with positioning satellites), The ground truth track of the vehicle's location recorded by GPS-RTK is also provided. The dataset includes many new features of urban and suburban driving, such as highly dynamic environments (large number of moving objects in vehicle mileage), roundups, slopes, construction detours, aggressive driving, etc., and is particularly suitable for long-term vehicle autonomous driving research because it captures daily and seasonal variations [8].

2) State space modeling:

First, the BEV unifies multimodal data processing dimensions, converting multiple camera or radar data into 3D perspectives for target detection and segmentation, thereby reducing perceptual errors and providing richer outputs for downstream prediction and planning control modules.

Table 1. Perception Data Log

Timestamp	Latitude	Longitude	Speed (m/s)	Lidar Data (Obstacle Distance)	Camera Data (Object Detection)	Distance Sensors Data
08:00:00.000	51.5074	-0.1278	10.5	5.2	[Object: Car, Distance: 10m]	[Front: 3m, Rear: 4m]
08:00:01.000	51.5075	-0.1279	10.6	5.1	[Object: Pedestrian, Distance: 8m]	[Front: 3.2m, Rear: 3.9m]
08:00:02.000	51.5076	-0.1280	10.7	5.0	[Object: Traffic Light, Distance: 20m]	[Front: 3.5m, Rear: 3.8m]
08:00:03.000	51.5077	-0.1281	10.8	4.9	[Object: Bicycle, Distance: 15m]	[Front: 3.8m, Rear: 3.7m]
08:00:04.000	51.5078	-0.1282	10.9	4.8	[Object: Truck, Distance: 12m]	[Front: 4.0m, Rear: 3.6m]

Secondly, BEV realizes timing information fusion. According to Table 1 the end-to-end optimization is directly completed through the neural network, and the perception and prediction are uniformly calculated in the 3D [8] space, thus effectively reducing the accumulation of error in the serial perception and prediction in the traditional perception task. Transformer's Attention mechanism helps transform 2D image data into 3D BEV space.Therefore, Transformer has a large saturation range. Compared with traditional CNN, Transformer has stronger sequence modeling capabilities and global information perception capabilities, which is very favorable for big data training requirements in AI large models.

3) Prediction step:

Drawing upon the established model of the system and leveraging the available dataset, the predictive capabilities of the Kalman filtering algorithm come to the forefront. By utilizing the dynamic model of the system and the current state estimate, we engage in the process of state prediction through kinematic equations. This foundational step allows us to project the future state of the system, yielding predictive covariance metrics for position and velocity. These metrics serve as vital indicators of the expected uncertainty in the system's trajectory, facilitating informed decision-making in real-time navigation and localization tasks.

Table 2. Predictive Covariance Data Table

Timestamp	Predicted Covariance (Position)	Predicted Covariance (Velocity)
08:00:00.000	0.05	0.02
08:00:01.000	0.06	0.03
08:00:02.000	0.07	0.04
08:00:03.000	0.08	0.05
08:00:04.000	0.09	0.06

Using the dynamic model of the system and the estimated value of the current state, the state prediction is carried out through the kinematic equation. The predicted state and predicted covariance are obtained from this process. The table above presents the predicted covariance data, including the covariance of predicted position and velocity. These values are calculated based on the system's dynamic model and the current estimate of the state, providing an estimation of uncertainty in the predicted position and velocity of the vehicle. Smaller values indicate more accurate predictions of the state.

4) Measurement update:

The comparison between the measured value provided by the sensor and the predicted value serves a pivotal role in the Kalman filtering process. This step is crucial as it allows for the refinement of the system's state estimate and covariance, ultimately leading to an optimal estimation of the system's true state.[9] Through this iterative process of comparing and updating, the Kalman filter effectively fuses information from multiple sources to provide a more accurate and robust estimation of the system's state. This optimal state estimate and covariance are essential for guiding decision-making processes in real-time applications such as autonomous navigation, where precise knowledge of the system's state is paramount for ensuring safe and efficient operation.

Table 3. Sensor Comparison Data Table

Timestamp	Measured Position	Measured Velocity	Predicted Position	Predicted Velocity
08:00:00.000	[51.5074, -0.1278]	10.5	[51.5075, -0.1279]	10.4
08:00:01.000	[51.5075, -0.1279]	10.6	[51.5076, -0.1280]	10.5
08:00:02.000	[51.5076, -0.1280]	10.7	[51.5077, -0.1281]	10.6
08:00:03.000	[51.5077, -0.1281]	10.8	[51.5078, -0.1282]	10.7
08:00:04.000	[51.5078, -0.1282]	10.9	[51.5079, -0.1283]	10.8

In this table, the measured position and velocity values provided by the sensors are compared with the predicted position and velocity values obtained from the Kalman filtering algorithm. By comparing and analyzing these metrics, we are able to verify the advantages of the proposed method and compare it with other methods to determine its performance advantages in the field of autonomous driving.

Estimated Values:

Updated State Estimate: [51.5075, -0.1279] (Position), 10.4 (Velocity)

In this iteration, the updated state estimate indicates that the vehicle's position is slightly adjusted towards the measured position, taking into account the reliability of both the measurement and the prediction. Similarly, the velocity estimate is refined based on the weighted average of the measured and predicted velocities, guided by the Kalman gain. These adjustments enhance the accuracy of the state estimate, ensuring that it accurately reflects the true state of the system amidst dynamic changes and sensor noise.

3.3. Experimental result

The experimental investigation focused on evaluating state-of-the-art positioning technologies for autonomous vehicles (AVs), emphasizing the criticality of accurate and real-time positioning for safe and efficient autonomous driving. Notably, the Kalman Filter, a widely adopted localization algorithm in autonomous driving, was identified as a key component of the analysis.The iterative refinement

process, guided by the Kalman gain, facilitated the adaptation of the covariance matrix to dynamic system changes and sensor measurement errors, thereby enhancing the accuracy of state estimation for precise positioning and prediction in [10]AV applications. It indicates that the proposed method has better performance and reliability in the field of autonomous driving. Compared to previous methods, our approach is able to more accurately determine the location of the vehicle and more accurately predict the future state of the vehicle, thereby improving the overall performance of the autonomous driving system.

4. Conclusion

Autonomous driving technology stands at the forefront of innovation, poised to reshape the landscape of transportation and profoundly impact society. Central to the realization of autonomous driving functions is the perception system, which serves as the cornerstone for accurately sensing and interpreting the dynamic traffic environment. Through advancements in sensor technology and data processing algorithms, autonomous vehicles can navigate complex scenarios with precision and reliability, laying the foundation for a future where road safety is paramount. By harnessing the power of these integrated systems, autonomous vehicles can achieve precise real-time positioning and make informed decisions in dynamic environments. Moreover, the experimental evidence presented in the paper underscores the effectiveness of these systems in mitigating the impact of sensor noise and uncertainties, bolstering the overall performance and safety of autonomous driving systems. Looking ahead, the future of autonomous driving technology holds immense promise, driven by ongoing advancements in precise positioning and prediction systems, coupled with the relentless pursuit of innovation in artificial intelligence. As autonomous driving technology evolves, it is poised to revolutionize the way we commute, ushering in an era where mobility is seamlessly integrated, safer, and more accessible for all.

References

[1]   Abdul Razak Alozi,and Mohamed Hussein. "Enhancing autonomous vehicle hyperawareness in busy traffic environments: A machine learning approach." Accident Analysis and Prevention 198. (2024):

[2]   Chen Ruinan, et al. "A Cognitive Environment Modeling Approach for Autonomous Vehicles: A Chinese Experience." Applied Sciences 13. 6 (2023):

[3]   Liu, Shaoshan, et al. "Edge computing for autonomous driving: Opportunities and challenges." Proceedings of the IEEE 107.8 (2019): 1697-1716.

[4]   Jie Luo. "A method of environment perception for automatic driving tunnel based on multi-source information fusion." International Journal of Computational Intelligence Studies 12. 1-2 (2023):

[5]   Farag Wael. "Self-Driving Vehicle Localization using Probabilistic Maps and Unscented-Kalman Filters." International Journal of Intelligent Transportation Systems Research 20. 3 (2022):

[6]   "Engineering - Naval Architecture and Ocean Engineering; University of Tasmania Researchers Reveal New Findings on Naval Architecture and Ocean Engineering (Experimental and numerical study of autopilot using Extended Kalman Filter trained neural networks for surface vessels)." Network Weekly News (2020):

[7]   Yuanyuan Wang,and Hung Duc Nguyen. "Rudder Roll Damping Autopilot Using Dual Extended Kalman Filter–Trained Neural Networks for Ships in Waves." Journal of Marine Science and Application 18. 4 (2019):

[8]   Daniel Viegas, et al. "Discrete-time distributed Kalman filter design for formations of autonomous vehicles." Control Engineering Practice 75. (2018):

[9]   Kumar, Swarun, Shyamnath Gollakota, and Dina Katabi. "A cloud-assisted design for autonomous driving." Proceedings of the first edition of the MCC workshop on Mobile cloud computing. 2012.

[10]   MOOSMANN F, STILLER C. Velodyne SLAM[C]// 2011 IEEE Intelligent Vehicles Symposium (IV). Baden-Baden, Germany: IEEE, 2011.