Multi-Vehicle Collaborative Planning Technology under Automatic Driving

Songsong Rong 1*, Ruifeng Meng 1, Junhong Guo 1, Pengfei Cui 2 and Zhi Qiao 2

1 College of Aviation, Inner Mongolia University of Technology, Hohhot 010051, China; caroline_rong@163.com (S.R.); junhongguo@126.com (J.G.)
2 Inner Mongolia High-Grade Highway Construction and Development Company, Hohhot 010051, China; pfcuiim@126.com (P.C.); 13789416847@163.com (Z.Q.)
* Correspondence: mrfnmgcn@imut.edu.cn

Abstract: Autonomous vehicles hold the potential to significantly improve traffic efficiency and advance the development of intelligent transportation systems. With the progression of autonomous driving technology, collaborative planning among multiple vehicles in autonomous driving scenarios has emerged as a pivotal challenge in realizing intelligent transportation systems. Serving as the cornerstone of unmanned mission decision-making, collaborative motion planning algorithms have garnered increasing attention in both theoretical exploration and practical application. These methods often follow a similar paradigm: the system initially discerns the driving intentions of each vehicle, subsequently assesses the surrounding environment, engages in path-planning, and formulates specific behavioral decisions. The paper discusses trajectory prediction, game theory, following behavior, and lane merging issues within the paradigm mentioned above. After briefly introducing the background of multi-vehicle autonomous driving, it provides a detailed description of the technological prerequisites for implementing these techniques. It reviews the main algorithms in motion planning, their functionalities, and applications in road environments, as well as current and future challenges and unresolved issues.

Keywords: autonomous driving; multi-vehicle coordination; motion planning; development prospect

1. Introduction

As urban ground traffic congestion becomes increasingly severe, it has evolved into a significant challenge hindering urban development. Congestion in large urban areas results in increased time and fuel costs for residents, leading to economic losses amounting to USD 121 billion [1]. The indirect losses caused by traffic congestion are also rising annually. The air pollution generated by urban traffic can lead to health damages ranging from GBP 4.5 billion to GBP 10 billion [2]. Therefore, implementing effective measures to alleviate traffic congestion is urgently needed.

One important way to handle problems like traffic congestion, energy usage, and environmental pollution is through multi-vehicle collaboration technology [3]. It efficiently lowers energy usage, exhaust pollutants, and traffic congestion by optimizing vehicle driving routes and speeds. In contrast to single vehicles, multi-vehicle collaboration demonstrates heightened efficacy in various tasks such as transportation, emergency response, surveying, mapping, and mining, often displaying superior efficiency and operational capabilities [4,5]. Advancements in internet of vehicles technology and vehicle-road collaboration technology facilitate the further development of communication technology, thereby presenting opportunities for coordinated control of multiple vehicles. The efficiency of multi-agent systems can be enhanced without compromising individual vehicle safety or efficiency. Consequently, this technology contributes to the enhancement of environmental quality and facilitates the promotion of sustainable transportation development [6].
In autonomous vehicle systems, the autonomous driving decision-making process comprises four hierarchical levels: path-planning, behavioral planning, motion planning, and local feedback control [7] (Figure 1). Path-planning involves the acquisition of real-time traffic data and road information, where network parameters are defined and an optimal path is computed through the utilization of path search algorithms.

Figure 1. Hierarchical division of planning under automatic driving.

At the behavioral planning stage, the vehicle’s onboard sensors detect other traffic participants and consider traffic regulations and additional information to formulate local path-planning like lane changes and vehicle following. Challenges encountered at this level include planning and tracking trajectories amidst uncertainty and limited information [8].

The subsequent phase in the short-term strategy, behavioral planning, takes into account factors such as vehicle dynamics to devise a rational path for the vehicle. At the lowest tier of local feedback control, adjustments and corrections to the vehicle’s movements are executed in real-time based on the vehicle’s sensor data and status information. This ensures that the vehicle adheres to the anticipated trajectory and behavior during motion.

Currently, technologies based on multi-vehicle collaboration integrate various research achievements in perception, communication, decision-making, and planning, forming relatively comprehensive but still rapidly developing research direction. Despite some reviews focusing on autonomous vehicle planning, they predominantly center around single-vehicle intelligence. There is currently no comprehensive literature review on the perception, communication, and planning aspects of multi-vehicle collaborative technologies. We categorize multi-vehicle collaborative motion planning techniques into rule-based collaborative methods and non-regular motion planning approaches. This paper aims to address this gap by comprehensively analyzing and summarizing the relevant research progress, with the hope of providing insights for the development of this field.

This article presents a comprehensive examination of multi-vehicle collaborative planning technology within the context of autonomous driving. Section 1 provides an overview of the motivation behind the study, reviews related work in the field, and outlines the contributions of the article. Section 2 discusses the development of rule-based multi-vehicle collaborative planning algorithms, addressing subtasks and game-theoretic issues in autonomous driving. Section 3 focuses on trajectory planning in multi-vehicle coordination, covering both vehicle platoons and non-platoon formations. This section provides a detailed overview of perception-based trajectory prediction, a crucial foundation for collaborative planning, and elaborates on the application characteristics of reinforcement learning in this field. Section 4 examines the challenges and solutions related to vehicle interactions in road traffic environments. Finally, Section 5 presents forward-looking recommendations for future research endeavors.
2. Rule-Based Multi-Vehicle Collaboration Planning

Much akin to human-driven vehicles, autonomous vehicles are bound by a plethora of regulations governing their operation. These regulations impose constraints on motion planning, necessitating that collaborative planning among vehicles consistently adheres to these constraints [9]. Rule-based multi-vehicle collaborative planning algorithms are crafted upon pre-defined rules and logic, enabling autonomous vehicles to execute safe and efficient path-planning. This algorithmic paradigm facilitates cooperative vehicle navigation by adhering to a shared set of rules, obviating the necessity for global optimization or centralized control. Consequently, such approaches enhance traffic efficiency, mitigate accident risks, and ensure the stable operation of vehicles within complex and dynamic traffic environments.

2.1. Behavior Selection in Multi-Vehicle Driving

Rule-based multi-vehicle collaboration algorithms typically derive from heuristic principles, encompassing strategies such as first in first out (FIFO), shortest path first (SPF), earliest arrival first (EAF), priority first, and minimum conflict first (MCF), among others. For instance, Ref. [10] employs FIFO rules to address the challenge of multi-vehicle cooperative merging at highway entrances. Building upon this, Ref. [11] proposes a rule-based adjustment algorithm, striking a favorable balance between computational complexity and traffic efficiency. This approach yields a near-optimal merging sequence, demonstrating applicability beyond highway entrances, including unsignalized intersections as highlighted by [12].

In current driving scenarios, scholars segment the research process into distinct sub-tasks, encompassing activities such as starting, lane keeping, and parking, among others, and formulate corresponding rules. Vehicles can then select behaviors by adhering to predefined switching rules [13]. Addressing the imperative of minimizing collision risks within structured road scenarios such as highways, Ref. [14] proposes a vehicle path-planning algorithm.

Fuzzy logic and expert systems are typical rule-based decision-making methods [15]. Fuzzy logic is a method for handling uncertainty and vagueness, allowing reasoning and decision-making in ambiguous environments. Ref. [16] utilizes fuzzy logic to process fuzzy input data between vehicles, considering variables such as distance from oncoming vehicles, own speed, and oncoming vehicle speed to determine the optimal lane-changing behavior. Ref. [17] employs fuzzy decision-making to facilitate the multi-vehicle merging process. Expert systems utilize domain knowledge and rules to coordinate actions between multiple vehicles. However, in practical operations, a large amount of accurate and detailed empirical knowledge is often required for coordination. The tendency to rely on precise empirical knowledge makes it ineffective in dealing with uncertainty in real-world scenarios.

The maturity of communication technology has facilitated advancements in rule-based multi-vehicle collaborative methods and enriched the completion of subtasks in driving environments. In recent years, road vehicle network technology has experienced rapid development. Dedicated short range communication (DSRC) technology operates in frequency bands ranging from 75 MHz to 5.9 GHz [18]. The technology of cellular vehicle-to-everything (C-V2X) communication, based on LTE-V networks, is characterized by large communication range, low latency, and frequent data exchange rates. Dynamic spectrum access (DSA) meets the growing bandwidth demands. These road communication systems can fulfill various communication tasks [19,20].

Based on V2V communication, a rule-based vehicle lane-changing control strategy proposed by [16] can smoothly navigate around sudden blockages on highways. Ref. [21] introduces a cooperative lane-changing model for scenarios with multiple congested lanes at intersections. Some following vehicles on the target lane decelerate to create enough space for coordinated maneuvers. However, the current state of communication is still immature, with instances of communication delays and interruptions. Ref. [22] presents
an algorithm catering to both autonomous and non-autonomous vehicles. This algorithm, designed for mixed traffic scenarios, obviates the need for inter-vehicle communication and models various maneuvers utilizing empirical formulas.

2.2. Game Theory

In traditional algorithms, path-planning typically aims to devise an optimal or nearly optimal route from a set of origins to destinations. However, when multiple vehicles share the same origin and destination concurrently, employing this method may lead to congestion if all vehicles opt for the same recommended path [23]. Consequently, the calculated optimal solution ceases to be the most advantageous path in practice. This scenario is commonly referred to as a congestion game [24]. In order to solve this problem, Ref. [23] models the vehicle routing problem at the intersection as a population game and proposed a distributed cooperative routing algorithm (DCR) based on evolutionary games to achieve Nash equilibrium.

In traffic control, intersections are divided into signal-controlled intersections and uncontrolled intersections [25]. Coordinating multiple vehicles at uncontrolled intersections presents a more intricate challenge compared to signal-controlled intersections. Opting for overly conservative decisions may lead to vehicles becoming jammed and unable to traverse the intersection, while excessively aggressive decisions can increase the risk of collisions.

Leveraging vehicle-to-vehicle (V2V) communication, vehicle-to-infrastructure (V2I) communication, and a multi-vehicle collaborative planning strategy can significantly enhance the safety and efficiency of intersection traffic (Figure 2).

![Figure 2. Intelligent connected vehicle communication diagram.](image)

Drawing upon the concept of a zero-sum game, Ref. [26] investigates two collision-free control strategies for vehicles. Ref. [27] introduces a framework centered on a novel dynamic game formulation and progressive view optimization to capture vehicle interaction behavior at uncontrolled intersections. This framework explicitly incorporates the dynamic behavior of vehicles and adheres to common traffic rules. In addition to intersection scenarios, several researchers have explored multi-vehicle collaboration using game theory principles in highway settings [28–31]. For instance, Ref. [32] introduces a method based on game theory techniques for highways, proposing a novel collaborative driving prediction and planning framework tailored for highly automated driving in dynamic environments. This approach facilitates model-based interactive sensory motion prediction for all vehicles within the scene, enabling the planning of maneuver sequences over extended time scales. Furthermore, in the context of highway entrance ramp merging scenarios, an Intention-integrated Prediction- and Cost Function-based Algorithm (IPCB) framework is introduced to facilitate autonomous vehicles in executing multi-vehicle collaborative behaviors [33]. Additionally, Ref. [34] delves into the interaction and decision-making processes of autonomous vehicles with diverse driving characteristics, presenting a cooper-
ative decision-making framework grounded in joint game theory to address the multi-lane merging challenges encountered by autonomous vehicles in multi-lane merging areas.

V2V and V2I communications undoubtedly provide great convenience for multi-vehicle collaborative driving. However, due to the high cost of the required infrastructure, network scalability issues and related security issues, there are still certain difficulties in the marketization of multi-vehicle collaborative driving. Moreover, self-driving vehicles must effectively collaborate and interact with vehicles lacking inter-vehicle communication capabilities. Challenges arise due to limitations and uncertainties in information sharing and communication among multiple vehicles, resulting in delays and gaps in vehicle-acquired information. The absence of accurate and real-time data may compromise the effectiveness of game theory methods.

Nevertheless, rule-based methods often exhibit limited applicability. They are tailored to specific scenarios and predefined traffic situations, thereby struggling to fully adapt to the intricacies of complex and dynamic traffic environments [35–37]. Consequently, in certain instances, the integration of other intelligent algorithms becomes imperative, such as deep reinforcement learning. In this regard, Ref. [38] merges rule-based algorithms with reinforcement learning, presenting an integrated decision-making framework (IDF) for autonomous vehicles (Figure 3). In situations where risks are identified within the decision-making network, caution rules are invoked to ensure driving safety. Similarly, manually encoding rules necessitate a substantial corpus of driving data to encompass all conceivable scenarios [39,40]. Addressing this limitation, Ref. [41] proposes a method that circumvents dependency on extensive driving data by employing a hierarchical reinforcement learning approach with labeled driving data. As such, the subsequent section delves into related research endeavors aimed at achieving greater flexibility and intelligence in multi-vehicle collaboration.

![Figure 3. Integrated decision framework of motion decision system.](image)

In decentralized multi-vehicle games, vehicles do not need to rely on infrastructure for communication; instead, they achieve their goals through autonomous decision-making and collaboration. Compared to centralized communication networks, decentralized multi-vehicle games not only reduce communication costs but also enhance system reliability and security. Recent work has applied decentralized methods [42,43]. On one hand, concurrent experience replay trajectory methods are used to share experiences among multiple agents and perform concurrent replays to address the issue of low learning efficiency. On the other hand, adjusting reward mechanisms and replay strategies enhances agents’ understanding of the environment and improves their behavior.

Additionally, in multi-vehicle collaborative planning, considerations such as overall traffic flow efficiency, congestion levels, and resource utilization require attention. Game theory methods often prioritize individual participants’ optimization goals, neglecting the optimization and equilibrium of the entire system. The pareto optimization theory provides a framework for optimizing resource allocation and utility distribution among
multiple vehicles. It ensures the maximization of overall system welfare while considering the diverse goals and constraints of individual agents [44]. Ref. [45] introducing the Pareto optimal method addresses cooperative control problems in autonomous driving vehicles. Future advancements should integrate game theory approaches with system optimization methods to achieve a more comprehensive and cohesive framework for multi-vehicle collaborative planning. By incorporating system-level considerations into the decision-making process, such as optimizing traffic flow efficiency and resource allocation, the resulting methods can better address the complex dynamics of real-world traffic scenarios.


Accurate and efficient trajectory prediction for multi-vehicle autonomous driving is crucial in complex traffic scenarios. To accomplish driving environment subtasks, intelligent vehicles need to infer the precise future trajectories of neighboring vehicles. Therefore, trajectory prediction is a fundamental prerequisite for non-regular multi-vehicle collaboration planning [46]. Non-regular multi-vehicle collaboration is divided into scenarios involving coordinated motion between vehicle formations and those without vehicle formations. Cooperative motion between vehicle formations requires consideration of the mutual influence and interaction between vehicles to achieve coordinated motion [47,48]. Both scenarios necessitate predicting surrounding information in the environment to assess risks associated with vehicle behavior and subsequently take appropriate action [49–51].

3.1. Multi-Vehicle Trajectory Prediction Method

Compared to single autonomous vehicles, trajectory prediction in multi-vehicle scenarios is more complex. Particularly in urban environments, road information and interactions with traffic participants can influence the predictive information of the fleet (Shown in Figure 4). Road information encompasses both inherent road infrastructure details in maps (lanes, traffic lights) and corresponding traffic regulations (yielding to pedestrians, traffic light rules). Furthermore, interactions between vehicles and other traffic participants further increase the complexity.

![Figure 4](image)

Figure 4. Multiple vehicles exchange constraint information in a road scenario.

Therefore, the accuracy of perception in the traffic environment forms the foundation of multi-vehicle collaborative planning. Currently, autonomous vehicles are equipped with heterogeneous sensor systems to provide ample environmental information, reducing issues related to unstable or missing information due to external factors. For instance, in environments where GNSS modules fail, optical flow modules can assist vehicles in position estimation. Radar, with its short-wave and high-penetration characteristics, exhibits strong robustness in rainy or snowy weather conditions. Additionally, it accurately perceives the surrounding environment in darkness, compensating for the shortcomings of camera perception. Hence, a multi-sensor fusion approach effectively complements perception information, acquiring as much valid information as possible for autonomous
driving [52,53]. This supports trajectory prediction of surrounding vehicles (illustrated in Figure 5).

Figure 5. The heterogeneous multi-sensor configuration in autonomous vehicles.

In multi-vehicle collaborative planning, the planning algorithm predicts the trajectories of other vehicles independently of its prediction range [54]. In complex traffic scenarios, various traffic participants take different traffic actions, causing the traffic environment to be dynamically changing. If an autonomous vehicle encounters an emergency, it needs to take proactive actions, such as slowing down to allow surrounding vehicles to move in, accelerating to change lanes to overtake, to ensure driving safety and efficiency. Therefore, the vehicle needs to accurately predict the location of adjacent vehicles. Future trajectories to allow risk assessment of vehicle behavior and further appropriate actions to be taken.

In the early days, it was assumed that the vehicle motion state was deterministic. The most primitive trajectory prediction method based on kinematics for multi-vehicle trajectory prediction was the single trajectory simulation method [55,56]. This method is usually based on a fixed prediction scenario or input condition, and cannot handle the uncertainties in the actual traffic environment, such as emergencies, or behavioral changes of other traffic participants. This may cause the planned trajectory to not adapt to the actual situation. Therefore, researchers research on trajectory prediction methods for vehicle motion uncertainty, such as the trajectory prediction method based on the Kalman filter [57], the Monte Carlo method [58], and Gaussian noise simulation. However, this method has poor performance in predicting vehicle motion trajectories in the vehicle operating environment.

In recent years, due to the development of positioning technology and sensor technology, the amount of vehicle natural driving trajectory data collection has been greatly improved compared to the past. Neural networks can automatically learn and capture complex features and patterns during vehicle driving by training on a large number of data samples. Trajectory prediction methods based on neural networks have been used in multi-vehicle collaborative planning [59–63]. As a variant of recurrent neural network (RNN), long-short term memory (LSTM) has a longer memory period and is suitable for long-term training of other vehicles. Tracking and trajectory prediction algorithms, which are favored by scholars, are used to track positions based on ranging sensors [64], identify driver intentions based on trajectory data [65], and predict future vehicle positions [66].

The essence of vehicle trajectory prediction lies in time series data regression [63]. Ref. [67] uses LSTM to predict trajectories for highway scenes. Ref. [66] expands the practicality of the former work and estimated the future positions of surrounding vehicles at a predetermined time. In order to improve the accuracy of prediction, Ref. [68] uses a priori trajectory position sequences to predict vehicle trajectories, which significantly improves the prediction performance and reduces computational losses.
While these methods exhibit noticeable efficacy in accomplishing their respective tasks, they nonetheless encounter challenges such as incomplete trajectory sequences or difficulties in establishing a simple probabilistic framework. Ref. [69] proposes a new vehicle trajectory analysis and prediction technology based on the LSTM encoder-decoder architecture. This method is able to generate future trajectories of surrounding vehicles based on the latest sensor measurement sequences.

However, simple LSTM models have two drawbacks: the inability to simultaneously model the spatial interaction between vehicles and trajectory sequences, and the possibility of encountering the vanishing gradient problem when modeling long-term sequences. Ref. [70] combines the characteristics of high-level spatio-temporal graphs with the sequence learning of RNN, and proposes structural-RNN, which constructs temporal relationships and spatial interactions into different time series to analyze the relationship between vehicles and trajectory sequences. Ref. [71] integrates the spatial expansion characteristics of convolutional neural network (CNNs) with the temporal expansion features of LSTM to forecast the trajectories of surrounding vehicles. The proposed method optimizes the model’s hyperparameters using the grid search algorithm to fulfill the dual precision prediction demands concerning space and time. Concerning the vanishing gradient issue encountered during modeling, Ref. [72] a trajectory prediction model based on spatio-temporal (ST-LSTM). This model incorporates shortcut connections between the input and output of each LSTM layer to directly convey historical trajectory information to subsequent layers. Such a structure aims to mitigate the vanishing gradient problem during backpropagation.

To acquire the aforementioned information and enhance the accuracy of multi-vehicle collaborative trajectory prediction models, existing research investigates various stages including the encoding-decoding phase [73,74], model optimization phase, and trajectory evaluation phase. The encoding phase captures interaction information between agents, laying the foundation for the model to understand the current state and interactions of agents. The model optimization phase improves prediction capability by refining the model structure and utilizing advanced training algorithms [75,76]. The trajectory evaluation phase focuses on assessing the output trajectories. However, there is limited research on trajectory prediction evaluation, with most studies modeling traffic agents as points without imposing constraints. Such unconstrained predictions may introduce significant uncertainty in predicting future states.

3.2. Cooperative Planning Method of Multi-Vehicle Formation

In multi-vehicle collaborative planning, through formation control, vehicles can move in specific geometric shapes or formations to achieve preset collective objectives. Formation control ensures that vehicles maintain reasonable relative positions and spacing, thus enabling efficient and coordinated collective actions. This is a specialized method deployed among vehicles with limited sensing and communication ranges [77]. In recent years, with the continuous iterative optimization of reinforcement learning algorithm models, deployment in the field of autonomous driving has also gradually matured.

Reinforcement learning does not rely on specific rules and models. Through the interaction between the agents and the environment, it can learn and optimize the behavior strategy of the vehicle to achieve the coordinated movement of multiple vehicles under autonomous driving. The Markov decision process describes the relationship between the agent and the environment. A mathematical framework for the environment interaction process. In this process, the agent is in different states and selects an action to interact with the environment based on the current state. The environment transitions to the next state based on the agent’s action and the current state, subsequently providing the agent with a reward or punishment signal. This process is a sequential decision-making process, and the agent’s goal is to maximize the cumulative reward by selecting the optimal action sequence(Figure 6).
In autonomous vehicle path-planning using reinforcement learning, the Q-learning algorithm has been widely employed due to its ability to address stochastic transformations and reward dynamics without requiring adaptation to environmental changes. The Q-learning algorithm also has a certain degree of parallelism and scalability, and can be used in multiple environments. Collaborative learning and path-planning are performed between individual agents to achieve path-planning and decision-making in more complex traffic scenarios. Refs. [78,79] developed a multi-layered architecture for multi-vehicle formation control, utilizing a decentralized control framework for multi-agent systems and employing traditional Q-learning for vehicle path-planning. However, traditional Q-learning exhibits deficiencies such as inadequate exploration-exploitation balance and sluggish convergence rates in multi-vehicle collaboration scenarios [79]. To mitigate these shortcomings, Ref. [80] adjusts dynamic parameters through trial and error, enhancing the Q-learning algorithm by incorporating an action deletion mechanism to better reconcile exploration and exploitation objectives. Such approaches face challenges in complex environments where anticipated outcomes remain elusive. Subsequently, Ref. [81] introduces a novel update method for the Q-function, proposing an enhanced Q-learning algorithm tailored for multi-vehicle cooperative obstacle avoidance in static environments. This refined algorithm facilitates the identification of optimal paths within the same iteration, resulting in shortened final path trajectories.

In recent years, researchers have shown significant interest in the task of path-planning for autonomous multi-vehicle “platoons”. Ref. [82] introduces a novel approach known as distributed receding horizon control (DRHC) architecture, which integrates nonlinear platoon dynamics with a deep reinforcement learning (DRL) scheme. This innovative framework addresses the path-planning challenges encountered by autonomous multi-vehicle platoons navigating urban road networks. By leveraging DRL for route decision-making, the DRHC architecture ensures the safe navigation of the platoon towards its target destination while preventing collisions between adjacent vehicles. Similarly, Ref. [83] recognizes the presence of non-autonomous vehicles within traffic flow and proposes a motion decision-making algorithm based on Nash-Q learning. This algorithm is specifically designed to account for the interactive dynamics between autonomous and non-autonomous vehicles. The decision-making framework is tailored for the study of multi-agent systems, with Q-learning serving as the foundational motion decision-making model, capable of determining optimal state-action sequences for autonomous vehicles.

As the number of vehicles increases and traffic flows become more complex, the importance of information exchange and cooperation among multiple vehicles becomes paramount. Through information exchange, vehicles can share real-time traffic conditions, road conditions, and expected actions, enabling more efficient path-planning and decision-making. This collaborative mechanism can assist vehicles in avoiding traffic jams, reducing traffic accidents, and optimizing overall traffic mobility. Ref. [84] employs a deep Q network (DQN) approach to tackle the hierarchical collaborative multi-vehicle tracking problem. By making decisions grounded in encoded states, DQN dynamically adjusts individual decisions based on real-time traffic information, thus deriving an optimal...
dataset of collective actions. Building upon this methodology, Ref. [85] introduces a collision avoidance algorithm tailored for multi-vehicle systems, effectively addressing scenarios characterized by high vehicle density and unrestricted driving directions. This approach strategically plans routes, thereby mitigating detour inconveniences and equitably distributing traffic flow.

3.3. Multi-Vehicle Collaborative Planning without Formation

As the development of self-driving cars advances, the emergence of hybrid autonomous driving environments, characterized by the coexistence of autonomous vehicles and human-driven counterparts, becomes increasingly likely. Given the unpredictable behavior of human drivers regarding cooperation with autonomous vehicles, the path-planning strategies of autonomous vehicles require meticulous consideration of the dynamics and roles of surrounding vehicles. In multi-vehicle collaboration scenarios, vehicles often contend for scarce resources such as the right of way or road access. Failure to accurately ascertain priority may precipitate contention conflicts, thereby impeding the seamless progression of vehicle operations.

In multi-agent reinforcement learning, the training cycle changes dynamically, and classic Q-learning and other methods cannot solve the problem of unstable agent environment in the system. Thus, the introduction of policy optimization methods becomes imperative to enhance the efficacy of large-scale multi-agent learning, particularly in environments characterized by non-stationarity. Ref. [86] employs deep deterministic policy gradients (DDPG) to facilitate the acquisition of cooperative overtaking strategies among multiple vehicles within simulated highway scenarios. Notably, the reward function is designed to leverage original sensor data from the current time step to inform decision-making processes. By driving agents to explore their environment and circumvent local optima, the RND-DDPG algorithm expedites algorithmic convergence, thereby enhancing overall learning efficiency.

The actor-critic algorithm, predicated on the fusion of value function and strategy optimization, introduces two distinct neural networks: the action network and the critic network. Ref. [87] delves into the mechanisms through which a decentralized reward structure fosters altruistic behaviors among vehicles. To incentivize vehicles to consider the welfare of both autonomous and human-driven counterparts, they introduce a synchronous Advantage Actor-Critic algorithm. Figure 7 illustrates diverse simulation outcomes of autonomous vehicles employing self-serving versus altruistic strategies during lane merging maneuvers.

![Figure 7. Vehicle confluence area.](image)

With the rapid advancements witnessed in artificial intelligence technology, swarm intelligence algorithms have found widespread application in path-planning tasks for unmanned vehicles. Prominent among these intelligent bionic algorithms are particle swarm optimization (PSO) [88–91], ant colony algorithms [92], genetic algorithms [93,94], and artificial neural network algorithms. Leveraging their biomimetic attributes, these methods exhibit enhanced intelligence and efficiency. Swarm intelligence algorithms, inspired by natural processes, replicate cooperative behaviors observed in decentralized self-organizing systems to achieve purposeful activities. These algorithms excel in ad-
dressing complex optimization challenges that are difficult for conventional mathematical methods, particularly those characterized by computational complexities or uncertainties from stochastic processes. However, they are susceptible to encountering issues in path-planning applications, including but not limited to, the pitfalls of local optimal solutions and slow convergence rates.

In a concerted effort to address path-planning and coordination intricacies inherent in multiple autonomous driving scenarios, Ref. [95] introduces a probability-based continuous ant colony optimization method, coupled with a random walk strategy and adaptive waypoint rectification mechanism. This approach meticulously optimizes the trajectory of each unmanned vehicle. Subsequently, the conundrum of obstacle avoidance within a multi-agent coordination framework is deftly resolved through the application of a speed transformation optimization algorithm. Furthermore, recognizing the paramount importance of mitigating energy consumption in multi-vehicle collaborative endeavors, Ref. [96] integrates the ant-based algorithm with a predictive model for power/energy consumption. This innovative fusion introduces the concept of green ant (G-Ant), offering a tailored solution for multi-vehicle collaboration. It facilitates collision-free, energy-efficient shortest path-planning strategies.

Compared to alternative algorithms, the particle swarm optimization (PSO) algorithm demonstrates a significant advantage in swiftly converging towards solutions in task allocation scenarios. In the context of autonomous driving challenges, multi-vehicle task allocation stands out as a pivotal concern. Ref. [97] formulates a comprehensive mathematical model encompassing pertinent constraints and functions relevant to both vehicles and targets. Moreover, this study introduces an innovative particle encoding and decoding methodology aimed at enhancing the performance of the particle swarm algorithm, thus facilitating the derivation of feasible multi-task allocation solutions. Nevertheless, a noteworthy drawback of the PSO algorithm lies in its propensity to prematurely converge towards local optima. Despite this limitation, recent endeavors have seen select researchers endeavoring to mitigate this shortcoming through the integration of PSO into multi-vehicle task distribution frameworks.

### 4. Future Development and Discussion

Existing multi-vehicle autonomous driving cooperative technology has laid the foundation for enhancing the autonomy of future smart city traffic. However, many key technologies still need further development, and there is still a long way to go for autonomous driving vehicles to form a transportation system. The preceding text has conducted a detailed analysis and discussion of the intelligent key technologies of autonomous driving vehicles. This section will expound on future development trends and the obstacles faced, and propose corresponding potential solutions based on experience.

#### 4.1. Environmental Uncertainty

In multi-vehicle collaborative planning, the interaction and collaboration among vehicles are significantly more complex compared to single-vehicle scenarios. The intentions and dynamic behaviors of other vehicles serve as crucial elements in this interaction, contributing to increased complexity and uncertainty in the planning process. Additionally, factors such as traffic flow dynamics, variations in road conditions, and changes in the state of traffic signals, among other environmental variables, undergo temporal and spatial alterations. These dynamic transformations profoundly impact parameters such as vehicular velocities, trajectory selection, and overall traffic conduct, thereby posing formidable challenges to the effectiveness of multi-vehicle collaborative planning efforts.

In uncertain environments, the uncertainty surrounding the future trajectories of other road users can be attributed to several factors. Firstly, the intentions of other vehicles and traffic participants are often unknown, contributing to the unpredictability of their actions. Secondly, the lack of precise longitudinal predictions regarding the future movements of other vehicles further compounds this uncertainty. Additionally, the probabilistic nature of
interactions between other vehicles and the target vehicle introduces another layer of unpredictability. Moreover, limitations inherent in sensor technology utilized for data collection and perception also contribute to the uncertainty surrounding future trajectories [54].

Addressing the challenge of uncertainty pertaining to the future trajectories of other road users, Ref. [98] leverages a partially observable Markov decision process to estimate the behavior of vehicles or pedestrians based on data gathered from noise sensors. Integration of this information into path-planning methodologies enables the realization of safe navigation in uncertain environments.

Multi-vehicle collaboration involves the states and behaviors of multiple vehicles, causing the state space to grow exponentially. Vehicles must navigate high-dimensional state spaces to seek optimal solutions, imposing greater challenges on search algorithms and computational efficiency requirements. The decisions and actions of each vehicle will affect the status and behavior of other vehicles, causing uncertainty to spread and diffuse in the fleet. The interaction between vehicles amplifies the complexity and instability of the entire system. Vehicles must take into account the behaviors and potential decisions of other vehicles to mitigate potential conflicts and collisions.

The vehicle’s own state is also a crucial factor influencing safe driving. The dynamic state of autonomous vehicles varies with changes in the target trajectory, vehicle speed, and front wheel steering angle. Trajectory-tracking control is essential for ensuring the safe and accurate tracking of the target trajectory by intelligent vehicles [99,100]. Tracking control for vehicle platooning is achieved by dispersing the longitudinal dynamic model of individual vehicles [101]. However, there is currently limited research on multi-vehicle trajectory tracking control. Most studies still focus on single-vehicle control, overlooking the interactive effects between vehicles.

In addition, pedestrians play a significant role in road traffic scenarios. Unlike vehicles, pedestrians have unpredictable movement patterns and behaviors, making them challenging for autonomous vehicles to safely navigate around. Typically, autonomous vehicles perceive pedestrians as dynamic obstacles and plan their trajectories to avoid potential collisions within the pedestrians’ movement areas. For instance, understanding pedestrian intent and behavior can help vehicles anticipate their movements and plan safer trajectories accordingly. Moreover, incorporating pedestrian awareness into the decision-making process of autonomous vehicles can lead to more efficient navigation strategies, reducing the risk of accidents and improving overall traffic flow [102].

4.2. Vehicle Collision

In complex urban traffic environments, vehicle conflicts pose a significant challenge. Multi-vehicle conflicts commonly arise in specific road segments or intersections, particularly in congested sections, intersections, highways, parking lots, and other densely trafficked areas (Figure 8). These conflicts, which may not necessarily culminate in collisions, often lead to traffic congestion, reduced efficiency, or undesirable traffic behaviors.

![Figure 8. Dense traffic area scenario. (a) Heavy traffic lanes. (b) Urban intersection. (c) Parking lot.](image)

At intersections, conflicts can arise when multiple vehicles arrive simultaneously and attempt to traverse the intersection without well-defined coordination and priority rules. For example, if vehicles such as car A and car B intend to make conflicting maneuvers, such
as left and right turns, respectively, within the limited green light duration, simultaneous entry into the intersection can lead to conflicts. Similarly, on highway merge lanes, multiple vehicles attempting to merge onto the main thoroughfare without coordination may result in frequent lane changes, decelerations, and accelerations, leading to traffic congestion and accidents. Human factors, excessive traffic volumes, road conditions, and weather represent key contributors to vehicle conflict instances, exerting an impact on the efficacy of multi-vehicle cooperation efforts.

In highly constrained spaces like parking lots, vehicles are particularly susceptible to model errors. Furthermore, these spatial environments are non-convex, and relying on geometric approximations can quickly deplete the available free space. Ref. [103] conceptualizes the self-driving multi-vehicle parking problem as a symmetric mixed integer linear optimization, with a focus on optimizing intentions and driving styles. Furthermore, tackling the inherent nonlinear and non-convex nature of this optimal control problem, conflict resolution is modeled in [104] as a multi-agent partially observable Markov decision process. This framework explores effective conflict resolution strategies within constrained spaces, generating collision-free and kinematically feasible paths. While reinforcement learning has attracted considerable attention in the field of autonomous driving, it still faces numerous challenges in practical applications. The black-box nature of reinforcement learning may lead vehicles to adopt unsafe measures in unforeseen circumstances. Although many studies have employed rule-based constraints to prevent reinforcement learning from executing unsafe strategies, achieving a balance between learned optimal strategies and safety remains a significant challenge.

4.3. Sustainable Development

Intelligent transportation systems have the potential to enhance traffic efficiency, alleviate congestion, and mitigate emissions, consequently reducing adverse environmental effects. Achieving intelligent transportation systems relies heavily on the advancement of autonomous systems, which contribute significantly to the development of autonomous driving and multi-agent control strategies for establishing sustainable and coordinated traffic management. Autonomous vehicles offer notable advantages in various key parameters, including reduced energy intensity, fuel consumption, emissions, travel times, and congestion levels.

Decision-making conflicts and competition between vehicles may lead to an increase in energy consumption. Energy conservation and management require a large amount of vehicle operating data and energy consumption information for modeling and optimization. However, obtaining real-time and accurate vehicle energy consumption data may face difficulties, especially when it comes to data sharing between different models and manufacturers. The lack of sufficient data may limit energy consumption research and practical applications. In multi-vehicle collaboration, conflicts can arise between optimizing energy consumption and minimizing travel time. Finding a balance between energy and time considerations is a crucial challenge to address. One approach to tackle this issue is by leveraging the Markov chain model to implement road energy consumption modeling. This involves considering various factors such as different driving modes, the agent’s surrounding environment, road conditions, and applicable limitations. By incorporating these factors and employing suitable learning strategies, the energy demand of the driving path can be estimated based on input parameters and model-based predictions [105].

Interactions between vehicles and pedestrians are inevitable in future scenarios. Ethical dilemmas faced by autonomous decision-makers will also constrain the development of smart cities. Incorporating ethical considerations into multi-vehicle decision-making frameworks, such as setting objective functions and costs within ethical systems, is a thought-provoking issue that requires careful consideration. In addition, public acceptance of autonomous vehicles is also a significant factor influencing their sustainable development. According to [106], technological, normative, and ethical problems all affect public acceptability. Relying solely on technological progress to address public concerns
is insufficient. Addressing societal perceptions begins with ensuring transparency in the development and deployment of autonomous systems. Additionally, normative issues such as legal frameworks and regulatory standards must be established to govern the use of autonomous vehicles. Under the premise of ensuring safety, autonomous vehicles can be allowed to operate under common driving conditions with behavior patterns similar to human-driven vehicles [107].

5. Conclusions

This collaborative planning paradigm plays a crucial role in aiding vehicles to circumvent collisions in intricate traffic environments, optimize traffic flow, and enhance driving efficiency and safety. With the vigorous development of vehicle technology, the vision of smart cities led by intelligent autonomous vehicles may soon become a reality. Considering the complexity of multi-vehicle cooperative systems, future research must emphasize breakthroughs in intelligent and information technologies, in addition to highlighting planning studies among multiple vehicles. Our literature review has revealed a wealth of technology in multi-vehicle collaborative planning. We aim to showcase the diversity and potential of multi-vehicle cooperative planning. Through these planning efforts, we hope readers can further explore research in related fields.

Author Contributions: R.M. and J.G. were responsible for the finalization and adjustment of the overall framework, and provided guidance on the challenges facing the paper in the future. S.R. was responsible for some literature research and article writing. P.C. and Z.Q. were responsible for some literature research and mapping. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Research on Key Technologies for Big Data Intelligent Analysis and Decision Making in 5G Smart Expressway Operation (2022YFSJ0040) and Research on Big Data Intelligent Analysis Technology for Road Conditions in the Hu-Bao Section of the G6 Expressway (JY20220170).

Institutional Review Board Statement: Not applicable for studies not involving humans or animals.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors Pengfei Cui and Zhi Qiao were employed by the company Inner Mongolia High-Grade Highway Construction and Development Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References


41. Duan, J.; Li, S.E.; Guan, Y.; Sun, Q.; Cheng, B. Hierarchical reinforcement learning for self-driving decision-making without reliance on labelled driving data. IET Intell. Transp. Syst. 2020, 14, 297–305. [CrossRef]

42. Yadav, P.; Mishra, A.; Kim, S. A comprehensive survey on multi-agent reinforcement learning for connected and automated vehicles. Sensors 2023, 23, 4710. [CrossRef]


45. Duan, J.; Li, S.E.; Guan, Y.; Sun, Q.; Cheng, B. Hierarchical reinforcement learning for self-driving decision-making without reliance on labelled driving data. IET Intell. Transp. Syst. 2020, 14, 297–305. [CrossRef]


71. Xie, G.; Shangguan, A.; Fei, R.; Ji, W.; Ma, W.; Hei, X. Motion trajectory prediction based on a CNN-LSTM sequential model. Sci. China Inf. Sci. 2020, 63, 212207. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.