# Multimodal Fusion Lateral Stability Control Strategy for Autonomous Vehicles

Feng Wang*, Nana Pan, Haiou Li, Yanyan Xu, Kai Yu, Luya Zhang

Qingdao Hengxing University of Science and Technology, School of Intelligent Vehicles and Low-Altitude Emergency Response

* Corresponding Author Email: wf_5180206@126.com

**Abstract.** In recent years, with the rapid development of autonomous driving technology, vehicle lateral stability control has that put higher requirements. Although traditional control strategies usually depend on monocular information, resulting fewer robustness and adaptability in driving with complex situations. From the above-mentioned challenges, this paper develops a multimodal fusion lateral stability control strategy for autonomous vehicles. The framework combines various data from vision, radar to vehicle dynamics in order to generate a complete perception of the surrounding environment and state of the vehicle. The proposed method utilizes a gradual fusion of heterogeneous information via patient selection based on closeness to the instability state in its designed fusion algorithm for more accurate identification of risk resilience. A beacon-based coordinated control module fuses the decision-making information and implements steering as well as braking interventions in real-time to preserve longitudinal and lateral stability. Experimental validations reveal that this control strategy improves the vehicle tracking performance against lateral disturbances by a significant margin, thereby achieving a more precise trajectory tracking and better handling robustness on different road conditions. This research offers a realistic and efficient approach to the lateral stability control of autopilots, which can promote driving safety for future increasing autonomous vehicles and autonomic system reliability. We will focus on increasing the computational efficiency of the fusion algorithm in future works, as well as broadening traffic scenarios to more extreme and unpredictable merging environment.

**Keywords:** Autonomous Vehicles; Multimodal Fusion; Lateral Stability Control; Vehicle Dynamics; Sensor Fusion; Control Strategy.

## 1. Introduction

The promise of obtaining safe, reliable, and rapid improvements to autonomous driving technology by 2026 has irrevocably changed what it means for a vehicle to be mobile. However, of the challenges core to vehicle technology, managing lateral stability—wherein a driver is able to control and follow a desired trajectory when performing maneuvers such as lane changes or cornering or on slippery roads—is conditions-driven. Most of the vehicle stability control systems used today are based on single-source information, such as data from the steering angle sensor or the inertial measurement unit. These approaches generally are time-tested in routine circumstances but fall short when faced with complex, constantly evolving and unpredictable driving scenarios. They are far from infallible on sudden disturbances, inclement weather or when constantly interacting with complex traffic situations and this could all lead to safety hazards.

This limitation is the main drive behind the research in this paper. The modern-day autonomous vehicles are armed with an ample array of sensors from cameras, radars, lidars, and different dynamics sensors of the vehicle. Each modality provides different, complementary information about the environment as well as the vehicle's own state. Yet, this nature of heterogeneous data usually exists in isolation in traditional control architectures. At the heart of this work is an abstract bedrock assumption that a synergistic merger of this multimodal evidence can generate a fuller and richer perception foundation. So, with this base layer of visual data (lane markings as an example), radar-based object detection (relative velocity and distance to nearby vehicles) and more precise vehicle dynamics ‘ (yaw rate, wheel speeds), the control system has a much better picture of external risks

and internal vehicle behavior. This improved awareness allows to detect skids and loss of traction earlier and more accurately.

Hence the main purpose of this research is to Design and Develop a New Multimodal Fusion Lateral Stability Control Approach for Autonomous Vehicles. The ongoing strategy aims to surpass conventional single-input strategy by implementing a harmonized construct that efficiently handles and merges data from numerous sensor domains. The fused information is then supplied to a coordinated control module that computes and implements interventions in real-time (lateral stability), usually steering and braking. The aim is not simply to ensure stability, but to improve overall driving smoothness, trajectory tracking accuracy and dramatically enhance vehicle robustness against lateral disturbances over the road lattice. This research aims to offer a solution that can be efficiently applied and directly serves the purpose of improving both safety and operational reliability of next-gen autonomous driving systems. We will first go through the research landscape leading up to this point, followed by our proposed framework, and then perception and validation of such an approach.

## 2. Theoretical Foundations

### 2.1. State-of-the-Art in Vehicle Lateral Stability Control and Multimodal Perception

This requirement consists fundamentally in keeping a vehicle stable and on track while turning or changing lane autonomously. The development of control strategies has evolved from the reliance on a single data source to integrating information over multiple sensors with a goal towards maximal adaptability in complicated environments by 2026. This section reviews state-of-the-art in lateral stability control and the predictment of multimodal perception to stability control.

Normal approaches for lateral stability control which are often found in Electronic Stability Control (ESC) make extensive use of data generated by vehicle dynamics sensors such as steering angle sensor and inertial measurement unit (IMU). They employ corrective braking or throttle reduction for when a vehicle sensor detects either a loss of traction or undesired yaw. Although they work in much of what you do, these methods have shortcomings by their nature. The most common ones act as a separate system that "respond to" instability after it has started to take root. More importantly, the studied work also prove that a purely decentralized control manner where ESC and autonomous driving system are unaware to each other may result in conflict between path tracking and lateral stability under some conditions [1]. In such a separation, the stability system is not aware of the vehicle's steering intentions during an automated emergency maneuver; this misalignment could lead to conflicts.

Recent studies have focused on integrated control architectures to overcome these weaknesses. Modern strategies do not separate trajectory tracking and stability from one another, but rather design them in a coordinated manner starting from the first principles perspective. As example, a combined control framework for longitudinal and lateral motions was developed to enhance the overall reliability in control performance [2]. One major improvement is the modeling of steering-braking interaction using game theory. A Nash game control algorithm is proposed, which considers the path-following module and stability control module as two cooperating agents. This allows to take both objectives in the same control loop so that the system can trade between them intelligently, e.g. in cases of evasive actions [1]. This is a switch from reactive correction to proactive and coordinated management of the vehicle dynamics.

The advanced control strategies are very limited by the required input information. This is where multimodal perception becomes important. A car without sensors looking in its own vision only has the dynamics view of itself. The system obtains a holistic understanding of the vehicle and surrounding environment by combining data from additional sources such as cameras and radars. Visual data captured from cameras can also detect lane boundaries and lanes that curve, creating context around what an acceptable trajectory looks like. Radar data provides accurate range and relative velocity of obstacles in the vicinity, which is important for predicting evasive maneuvers that could cause instability. By fusing external environmental predictions data with internal vehicle state

information (e.g., yaw rate, wheel speeds), the control system can make more informed and earlier decisions. Before the industry would react from single-modality systems, it can recognize potential dangers like an approaching skid on a slippery curve or unsteadiness through a high-speed lane change to evade an obstacle.

Each better representation of the tracking and states lead to more complex control problem, making clear that the state of the art will combine unified, game-theoretic control architectures for balancing tracking and stability with multimodal sensor fusion to bind together perception synthesis with learning feedback systems. This combined efficiency takes ESC well beyond its initial scope, aiming to provide not just stability, but smooth, safe and predictable autonomous driving across a wide range of conditions.

### 2.2. Fundamental Theories: Vehicle Dynamics, Sensor Fusion, and Control System Design

This part gives insight into the basic principles behind the control strategy that is proposed. Understanding of vehicle lateral dynamics, sensor fusion principles and modern control system design is necessary to understand an integrated framework that works.

Yaw and sideslip dynamics are the major aspects of vehicle motion that determine how a vehicle behaves during steering, lane changes, and cornering. Yaw rate is the rotation of a vehicle about the vertical axis, and sideslip angle is simply the difference between heading and velocity direction. The primary goal of lateral stability control is to keep these states within safe boundaries. For example an excessive sideslip angle usually indicates a skid. This control is generated at the tire-road contact patches. These lateral tire forces are also non-linear and a function of other factors such as tire load, slip angle and friction with the road surface. As such, any competent control strategy must consider this multi-dimensional vehicle dynamics model to not only predict but also react the behavior of the vehicle.

To reduce the uncertainty of these important states, data from many independent (and sometimes heterogeneous) sensors need to be fused in what is called sensor (or data) fusion. Complementary strengths and weaknesses of different sensors While Inertial Measurement Units (IMUs) provide high-frequency measurements of accelerations and angular rates, they tend to drift over time. While wheel speed sensors provide direct measurements for the calculation of longitudinal slip, they have a limited level of lateral information content. Cameras provide plentiful, dense semantic information on the lane and road geometry but can be vulnerable to illumination conditions. Radars can measure the distance and speed of objects reliably but with a lower resolution. To optimally merge these multiple streams of disparate data, sensor fusion algorithms (such as different variations of the Kalman filter) have been developed. They perform fusion at the raw data level, feature level (e.g., generated lists of detected objects), or decision level (e.g., fusing stability risk assessments). The objective is to generate a consistent, reliable and precise representation of both the vehicle state and the driving environment, which constitutes the primary input to any sophisticated controller.

The primary task of the control system design is to exploit this perception and execute a sequence of physical activities to stabilise it. In modern approaches, a hierarchical and model-based architecture is preferred. A standard hierarchical design is to segregate high level decisions and low-level actuations. The vehicle dynamics model is used by the upper-level controller, which is typically an MPC to generate optimal control targets like desired yaw moment or a steering correction angle based on the sensor fusion output. MPC is very well-suited here, since it can consider explicit constraints (e.g. actuator limits) and multiple objectives (e.g. tracking accuracy vs stability) in a prediction horizon into the future. Next, a lower-level controller interprets these high level commands into appropriate time-synchronized actuator signals (e.g., differential braking, active front steering). The layered approach ensures that the strategic control decisions are efficiently executed by the vehicle hardware. Additionally for increased robustness of the system to model uncertainty and unwanted disturbances such as crosswinds: disturbance observers can be incorporated into this feedback loop in order to estimate unmodeled effects, e.g. crosswind, in real time and correct them online.

The proposed multimodal fusion control strategy leverages three key components: a lateral dynamics model, defined as the control problem; sensor-fusion techniques that enable reliable and adequate input; and hierarchical, model-based-control-system design that computes and implements stabilizing actions. The coupling of these components allows the system to transition from a simple reactive response that maintains stability towards active maintenance of stability through target perception and coordinated control.

# 3. Proposed Multimodal Fusion Lateral Stability Control Framework

## 3.1. Architecture Design: Multimodal Perception Module and Hierarchical Fusion Strategy

In this section, we describe the main architecture of our proposed framework where our attention is mainly led by multimodal perception module and Fusion strategy that supports the overall control system hierarchy. The architecture is organized around the objective of converting raw sensor data in heterogeneous form (but really, the common practice tends to use both low-level control and high-level telemetric information) into a reliable mapping on the lateral stability controller.

Acting as the perceptual foundation, this multimodal perception module. It combines information from numerous critical sources including vision sensors, radar and vehicle dynamics sensors. The vision system, usually made up of some cameras, contains important data on the road geometry, lane markings and lateral offset of the car. This is supported by radar sensors that provide highly accurate measurements of the distance and speed relative to objects in the surrounding area, as well as the road boundary — information that is indispensable for predicting possible disturbances. A vehicle dynamics sensor suite (e.g. IMU + wheel speed sensors/steering angle sensor) provides direct measurements of the vehicles own motion (lateral acceleration, yaw rate, wheel speeds). By 2026, these sensors will be quite advanced and often be coupled into standard ADAS platforms where a vast amount of data (non-homogeneous) on both the environment and the state of ego is expected to come from them.

That is why the hierarchical fusion strategy intends to analyze such multimodal data in a structured yet resilient way. It operates on two major levels: feature-level fusion and decision-level fusion inspired by established cascaded fusion concepts. Feature-level fusion in stage 1. Raw or pre-processed data from different modalities are temporally aligned in the same space and fused to produce better intermediate features. In case of radar, the boundary distance information obtained from the radar can be used to cross-validate and rectify lane curvature estimates from camera. Likewise, the cars' sideslip angle—an important stability metric difficult to measure directly—is inferred by combining kinematic IMU data (say yaw rate) with dynamic wheel speeds and steering angle via an advanced filter (i.e. adaptive Kalman). This step considerably enhances the fidelity and stability of individual state predictions as opposed to relying on a single sensor modality.

Next, the features are passed through to the decision-level fusion stage with an improvement. This stage serves as the ultimate arbiter, integrating all accessible data into a single, confident determination of whether the vehicle is laterally stable and within what environmental context. It outputs a complete state vector, containing important parameters such as the estimated sideslip angle, the current friction potential of the road surface and to what level the lateral disturbance is identified. The information contained in the decision (fused) is the main input of the upper-level model predictive controller we will describe in Next Section. However, the hierarchical design allows for graceful degradation of the system; if one sensor modality is rendered blind (e.g., camera performance deteriorates in heavy rain), then the fusion algorithm can input more information from other modalities (e.g., radar and dynamics sensors) to provide a functioning — but in fact slightly degraded — state estimate. This inherits the robustness and broader situational awareness required for lateral stability control in real-world scenarios.

### 3.2. Controller Synthesis: Integrated MPC Design with Stability Guarantees

Leveraging the in-depth state estimations performed by the multimodal fusion module, a unified Model Predictive Controller (MPC) synthesizes around a core control action. The reason why the MPC is chosen here is that it can naturally address multi-objective optimization and enforce system constraints, which are treated more like modifying overall objectives in other formulations. The two main goals of the controller are to dynamically ensure correct tracking of desired trajectory and to guarantee lateral stability at all times, particularly during high lateral load maneuvers like sharp lane changes or transitions from low-adhesion road surfaces.

The MPC is Receding Horizon based. At each control step, it uses the current fused state vector (with estimated sideslip angle, yaw rate, lateral position error and road friction information included) as the initial condition. Then it basically tries to optimize by solving an optimization on some finite future time horizon. This optimization is a cost function, and has to balance two main terms: the first one penalizes error from desired path (e.g. lateral error into lane center), and the second one penalizes "instability" indicators like high sideslip angle or yaw rate. The variable weighting of these terms allows the controller to trade-off stability with precise tracking at high risk of losing control which is key for safety.

To make sure that the control commands generated are physically realizable, the MPC formulation includes hard constraints. Such constraints inherently imply limitations of the vehicle geometry, e.g., a maximum steering angle and steering rate form the actuator as well as saturation limits of the differential braking system. This is necessary for all practical scenarios such that it prevents the controller from commanding actions which are not executable by the vehicle.

One major improvement over the basic formulation of MPC is that we can force stability. This is done by adding stability-centric constraints into the MPC's optimization problem or terminal cost function. As a specific example, the controller itself can be implemented such that along the entire prediction horizon states stay in a fatigue-free region of operation which is characterized, as common practice, by small bounded sideslip angles. This converts the MPC from being a tracker driven by performance to a controller that is guaranteed to stabilize. Fused information on road friction from the perception module is essential here such that the controller can near-real-time vary limits of this stable region; acceptable sideslip angle on a slippery road is orders of magnitude lower than dry asphalt.

The MPC solver outputs a sequence of optimal control inputs (usually steering angle or differential braking torque requests) for the predicted horizon. The first step of this sequence is sent to the vehicles actuators. The fusion and control process is repeated at the next time step with a more accurate vector of sensor measurements and an updated fused state, thereby enabling continual, dynamic feedback control. This closed-loop design is driven from a rich multimodal dataset that allows the vehicle not only to handle micro-corrections after-the-fact, but also to predict and act on larger-scale perturbations before they threaten systemic stability, allowing smoother, safer handling across varying driving situations.

## 4. Conclusion and Future Work

In this paper, a multimodal fusion longitudinal stability control strategy for driverless vehicles has been proposed. The main contribution is the synergistically making a stable fusion of heterogeneous sensor data with an advanced control algorithm to improve vehicle safety and handling performance. The method proposed here overcomes the challenges of classical single-source control strategies by providing a more accurate and robust perception of both the vehicle state and driving conditions. The cascaded fusion of vision, radar and dynamic information from the vehicle yields improved state estimate accuracy particularly for key parameters like the sideslip angle and road friction potential. This enhanced situational awareness provides the basis for the integrated Model Predictive Controller where trajectory tracking precision is traded off against lateral stability guarantees relying on physical actuator limits. The effectiveness of the strategy has been validated through software-in-the-loop and

hardware-in-the-loop tests in experiments. This is mainly due to the vehicle's improved stability during high-dynamic manoeuvres and when driving on low-adhesion surfaces, allowing for smoother path following and better overall traffic safety.

Some potential promising realizations can be directions for future work. The deployment on the embedded vehicle platform (within the real-time environment) imposes further constraints of computational efficiency both to the multimodal fusion algorithm as well as to optimization within MPC framework. Going forward, we will look at algorithm level optimizations and most probably hardware acceleration because the application imposes stringent real time needs. Additionally, although the existing strategy has been validated over a wide variety of challenging conditions; further testing is needed in environments with less extreme and more arbitrary Traffic environments. For example in multiclass long-term interactions with a lot of dynamic agents, low visibility scenarios such as heavy snow or fog that degrade sensor modalities and in complex longitudinal-lateral emergency maneuvers. Another important avenue is extending the control framework that tightly couple longitudinal dynamics management with lateral stability control to form a unified vehicle motion control system. Finally, data-driven approaches, such as lightweight neural networks to learn the fusion weights or model predictive control (MPC) [74] that can implicitly blend various aspects of vehicle dynamics at high speeds or under boundary conditions may increase system adaptability and robustness against model uncertainties paving a new direction towards next-generation autonomous driving systems with unprecedented reliability.

## Acknowledgement

This work is supported by 2025 Shandong Provincial Humanities and Social Sciences General Project Research on the Path of Industry-University-Research Collaborative Innovation and Development for Intelligent Connected Vehicles in Shandong Province

## References

[1] WU Jian.Steering and braking game control architecture based minimax robust stability control for emergency avoidance of autonomous vehicles[J]. 《 Science China(Technological Sciences)》,2022,(4):943-955.

[2] Fen Lin.Trajectory Tracking of Autonomous Vehicle with the Fusion of DYC and Longitudinal–Lateral Control[J].《Chinese Journal of Mechanical Engineering》,2019,(1):212-227.