Autonomous Truck Navigation With Trailer Integration Via Natural Language Processing

Junseo(Jason) Kim, Rishabh Raja*, Ateeb Jawaid*, and Raymond Ha**
Supervisor: Dr. Sajad Saeedi

***Abstract*— This paper presents an innovative approach to autonomous truck and trailer navigation, integrating Natural Language Processing (NLP) with advanced sensor fusion technologies to facilitate human-vehicle interaction. The system achieves precise controls and adaptability across various environmental contexts by leveraging the potential of Lidar, Inertial Measurement Units (IMU), and an Ackerman steering mechanism. Finite Element Analysis (FEA) presents structural integrity and operational efficiency, ensuring the system's robustness. Emphasizing a comprehensive development strategy, this study bridges mechanical engineering and computational intelligence, highlighting NLP's pivotal role in enhancing navigational commands and decision-making processes. We discuss multidisciplinary methodology, design rationale, and the integration of cutting-edge technologies that promise a user-friendly autonomous driving experience.**

I. INTRODUCTION

Simultaneous Localization and Mapping (SLAM) is a foundational topic of autonomous robotics, enabling navigation and environmental mapping without pre-existing maps. This paper introduces a pioneering approach to autonomous navigation in truck and trailer systems by leveraging Natural Language Processing (NLP) and advanced sensor fusion.

The project's core innovation lies in the application of NLP for interpreting user commands for navigation, simplifying human-machine interactions significantly. NLP allows for an intuitive interface where operators can direct the autonomous system using natural language, which significantly enhances the traditional manual or predefined command inputs. The system demonstrates precise maneuverability and adaptability to diverse environmental conditions, coupled with an Ackerman steering-based robotic platform and integrated sensor technologies, including Lidar and Inertial Measurement Units (IMU).

The main contribution of this work is the development of a robust mechanical and computational framework that synergizes hardware robustness with software intelligence. Through comprehensive Finite Element Analysis (FEA), the structural integrity and reliability of the truck and trailer system are validated, ensuring its durability and efficiency. The integration of mechanical precision with advanced sensor fusion and NLP technologies provides insight into the future of autonomous transportation.

Fig. 1. Demonstration of Map Generation using Lidar-IMU sensor fusion. The left below shows the Map data from the Sensors using the Navigation stack.

The introduction of NLP as a cognitive engine for autonomous systems represents a significant step forward, enabling direct translation of user commands into actionable navigational decisions. This advancement ensures adaptable and intelligent operation across various scenarios, overcoming the limitations of predefined algorithms and sensor data unable to capture the dynamic complexities of real-world environments. See the videos on the website of the project[1].

The rest of this work is organized as follows:

Sec. II presents the literature review. Sec. III describes the hardware design and analysis of the system. Sec. IV proposes the Lidar-IMU sensor fusion for map generation. Sec. V extends NLP for the truck navigation and control. Sec. VI presents the experimental results. Finally, Sec. VII discusses future works and concludes the paper.

II. LITERATURE REVIEW

The following section summarizes the background and recent advances in autonomous transportation in terms of kinematics, sensor integration, and the use of NLP.

---

*Toronto Metropolitan University, Toronto, Canada

[1] https://sites.google.com/view/autonomous-navigation-trailer/

A. *Kinematics of Truck-Trailer System*

In autonomous truck and trailer systems, understanding the kinematics between the truck and trailer is crucial for stability and maneuverability. Studies have focused on four-bar linkage configurations with articulated designs for stabilization. Experiments with different linkage arrangements showed that forward converging configurations offer superior stability, highlighting the significance of the trailer's positioning relative to its axle[1]. Additionally, research into optimal turning mechanisms based on Ackerman steering principles provided insights into calculating minimum turning radii and ensuring smooth turns [2]. Explorations into Dublin's path for optimal route selection further demonstrated the efficiency of using logical classifications to reduce path analysis time [3].

B. *Sensor Fusion in Autonomous Systems*

Sensor fusion is becoming increasingly crucial to autonomous systems, significantly enhancing safety, reliability, and efficiency in technologies like self-driving vehicles [4]. One approach to vehicle localization combines GNSS, IMU, DMI, and LiDAR with the Unscented Kalman Filter, notably improving localization accuracy and the ability to detect curbs in real-time [5]. Similarly, integrating GNSS, IMU, Odometer, and LiDAR-SLAM has been shown to enhance navigation precision in varied environments, reducing position drift error significantly [6]. This sensor fusion strategy also extends to UAVs, where combining LiDAR and IMU data via the Kalman filter enables effective indoor navigation and localization[7].

C. *Natural Language Processing for Navigation*

Natural Language Processing (NLP) is evolving beyond traditional text-based applications to play a pivotal role in autonomous driving systems. By adapting NLP techniques for environmental interpretation, autonomous vehicles can now understand complex commands and navigate through diverse scenarios based on sensor inputs from Lidar, radar, and cameras[8]. This integration with sensor fusion technologies enhances vehicles' environmental perception, allowing for direct responses to obstacles and abstract commands[9]. Additionally, NLP facilitates predictive modeling, enabling vehicles to assess and select the most efficient and safe navigation paths[10].

III. DESIGN AND ANALYSIS

The design of the autonomous truck and trailer system combines Ackerman steering, which doubles as differential steering, with a three-floor structure for component housing, as shown in Fig. 2. This layout optimizes space for wiring and aligns the Lidar with the IMU to facilitate sensor fusion, primarily requiring adjustment for vertical alignment. Structural supports are strategically placed to minimize deflection under load. Failure Mode and Effects Analysis (FMEA) identified two primary failure points: the hitch's rotational component and the 3D-printed axles. Both issues stem from the orientation during 3D printing, which aligns print lines with the direction of applied force, increasing susceptibility to breakage. Despite operational parameters designed to prevent actual failure, manufacturing improvements, such as adjusting print orientation or utilizing injection molding, are advised for enhanced durability. Finite Element Analysis (FEA) for stress and load shown in Fig. 3 validated the design, showing the system's structural robustness with potential vulnerabilities. This approach to design and analysis ensures the system's reliability and robustness, providing innovative navigation capabilities with a durable physical architecture.

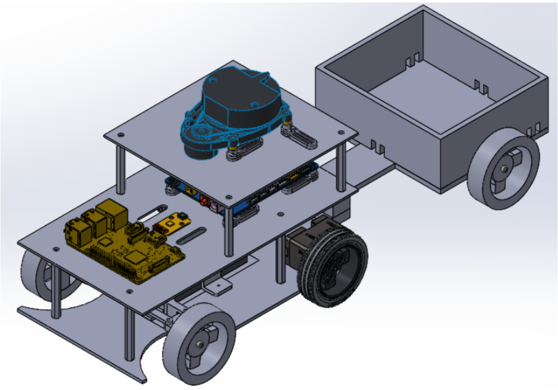


Fig. 2. Final 3D CAD Model of the Truck-Trailer System incorporating both differential drive and Ackerman steering.

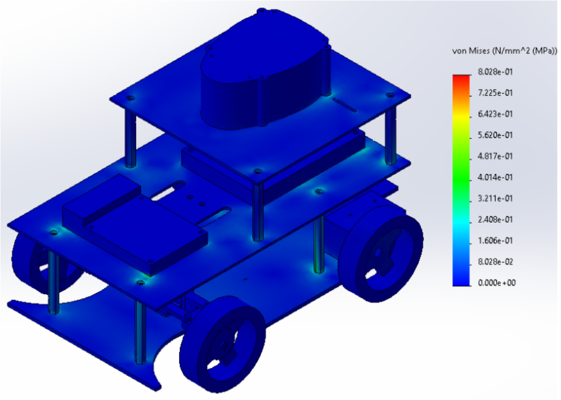


Fig. 3. Diagram demonstrating the FEA of the system using von Mises yield criterion..

Algorithm 1 LiDAR-IMU Sensor Fusion and Mapping

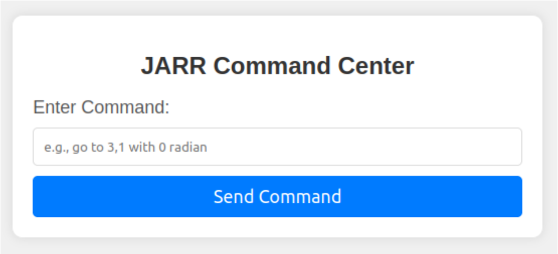
```
Input: LiDAR data $L$, IMU data $I$,
Output: Map $M$, Robot pose $P$
Initialize map $M$ and robot pose $P$
while robot is navigating do
    $L_{points} \leftarrow$ ExtractPointCloud($L$)
    $I_{measure} \leftarrow$ MeasureMotion($I$)
    $P_{predicted} \leftarrow$ PredictPose($P$, $I_{measure}$)
    if MatchFeatures($L_{points}$, $M$) then
        $P \leftarrow$ FuseEstimation($P_{predicted}$, $L_{points}$, $M$)
    else
        $P \leftarrow P_{predicted}$
    end if
    $M \leftarrow$ UpdateMap($M$, $L_{points}$, $P$)
end while
return $M$, $P$
```

IV. SENSOR INTEGRATION

In SLAM (Simultaneous Localization and Mapping), sensor data integration plays a pivotal role in accurately determining a robot's position within an unknown environment and constructing a detailed map of its surroundings. The system uses Lidar SLAM to capture environmental data as 2D point clouds, enabling precise movement estimation through sequential point cloud registration. Concurrently, the Inertial Measurement Unit (IMU) contributes to localization by tracking the object's movement in space, factoring in acceleration and angular velocity. The fusion of Lidar and IMU data is critical, with IMU measurements refining Lidar outputs by mitigating errors such as motion blur. This integrated approach allows for the extraction of relevant features from the sensor data, like geometric patterns from Lidar scans or key points in visual imagery, which are essential for both mapping the environment and localizing the robot within it. The identified features match existing map elements or are recognized as new environmental features, which is crucial for map updates and precise pose estimation. The robot's pose is continually updated by assimilating sensor data over time, leading to the progressive enhancement of the map with newly discovered information and the refinement of existing landmarks. The detailed algorithm is explained in Algorithm 1.

V. NATURAL LANGUAGE PROCESSING

In the autonomous truck and trailer system, Natural Language Processing (NLP) serves as a cognitive engine, adeptly converting user commands delivered in natural language into actionable navigation instructions. This integration is grounded on parsing techniques that analyze user inputs, extracting essential navigational parameters such as destinations, orientations, and waypoints. Once extracted, these parameters undergo a validation process to ensure their accuracy and relevance, bridging the gap between human communication and machine interpretation.



JARR Command Center

Enter Command:

e.g., go to 3,1 with 0 radian

Send Command

Fig. 4. User Interface for the NLP. The natural language command is then extracted and parsed to retrieve meaningful information, which is then passed to the navigation stack.

This NLP framework is a Flask-based server architecture designed for interaction between the user's natural language commands and the vehicle's autonomous navigation system shown as Fig. 4. This server acts as a conduit, parsing and structuring commands into a format interpretable by the Robot Operating System (ROS). It passes these structured commands to the ROS navigation stack, responsible for the vehicle's path planning and real-time localization tasks.

The integration process is comprehensively designed, illustrating how commands, after being processed by the NLP system, transform into inputs for the navigation module. This guarantees that the autonomous truck and trailer system accurately interprets the specific demands of each command and directly employs this knowledge for dynamic and intelligent navigation

VI. EXPERIMENTAL RESULTS

Simulations with the TurtleBot3 model within the Gazebo environment played a pivotal role in developing and testing our autonomous truck and trailer system. Utilizing the comprehensive functionalities of the Robot Operating System (ROS) navigation stack, these simulations provided a controlled yet versatile platform for algorithm testing and refinement shown in Fig. 5. The global planner within the stack, powered by Dijkstra's algorithm, computes the most efficient path from the vehicle's current location to its designated destination. This process ensures that the long-term navigational strategy is optimal and efficient, considering various factors

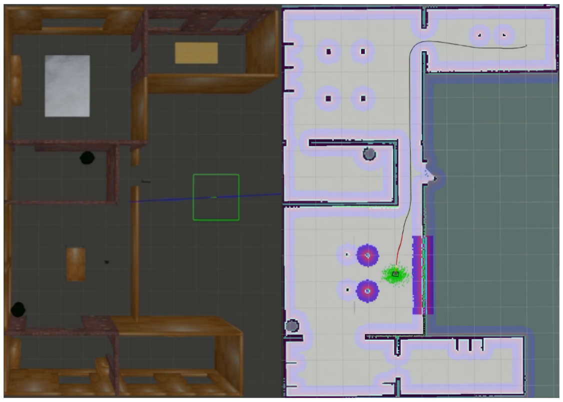


Fig. 5. Simulated experimental results: Gazebo simulator environment (left) and the navigation stack using Dijkstra's algorithm and DWA planner (right).

such as distance, obstacles, and environmental layout.

Complementing the global planner, the Dynamic Window Approach (DWA) serves as the local planner, focusing on the vehicle's immediate navigational challenges. DWA dynamically adjusts the vehicle's trajectory based on real-time sensor inputs, allowing for swift navigation around unforeseen obstacles and rapidly changing environments. This dual-planner ensures a balanced navigation strategy that combines strategic pathfinding with the agility needed for immediate environmental adaptation.

Transitioning from virtual simulations to real-world applications, experiments were conducted in a physical environment designed to emulate real-world navigation's complexities like Fig. 1. The gmapping algorithm was used during these experiments, creating a highly accurate digital map of the experimental environment. This map was essential for the ROS navigation stack to effectively execute complex navigational tasks, adapting the vehicle's movement strategy based on the layout and obstacles in the real world.

The iterative development process needed the collection and analysis of sensor data throughout both simulated and real-world testing phases. This data provided the system's navigational performance, confirming the effectiveness of algorithmic adjustments. The use of the global and local planners within the ROS navigation stack showcased the system's advanced capability for strategic planning and responsive maneuvering.

VII. FUTURE WORK AND CONCLUSIONS

This paper presents a promising approach to autonomous navigation, integrating Natural Language Processing (NLP) with advanced sensor fusion to enhance user-machine interaction and environmental perception in a truck and trailer system. The project successfully interprets natural language commands for precise navigation and leverages a rotary encoder alongside Lidar and IMU sensors for optimal route planning and obstacle avoidance.

Validated through simulations and real-world tests, the system proves effective in adapting to complex scenarios, highlighting the potential for broader applications in autonomous logistics. Despite current semantic understanding and technical language processing limitations, future work will enhance NLP models and expand datasets to refine the system's interpretative accuracy, particularly for intricate maneuvers like reverse parking. The exploration of Large Language Models (LLMs) and Vision Language Models (VLMs) aims to bridge these gaps, advancing autonomous transportation toward a more intuitive and reliable future.

ACKNOWLEDGEMENTS

We would like to thank Dr. Sajad Saeedi for the guidance and Mr. Matthew Lisondra for the feedback.

REFERENCES

[1] F. Sorge, "Optimization of vehicle-trailer connection systems," in *Journal of Physics: Conference Series*, vol. 744, no. 1. IOP Publishing, 2016, p. 012209.

[2] H. Choset, K. M. Lynch, S. Hutchinson, G. A. Kantor, and W. Burgard, *Principles of robot motion: theory, algorithms, and implementations*. MIT press, 2005.

[3] A. M. Shkel and V. Lumelsky, "Classification of the dubins set," *Robotics and Autonomous Systems*, vol. 34, no. 4, pp. 179–202, 2001.

[4] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, "Sensor and sensor fusion technology in autonomous vehicles: A review," *Sensors*, vol. 21, no. 6, p. 2140, 2021.

[5] X. Meng, H. Wang, and B. Liu, "A robust vehicle localization approach based on gnss/imu/dmi/lidar sensor fusion for autonomous vehicles," *Sensors*, vol. 17, no. 9, p. 2140, 2017.

[6] L. Chang, X. Niu, and T. Liu, "Gnss/imu/odo/lidar-slam integrated navigation system using imu/odo pre-integration," *Sensors*, vol. 20, no. 17, p. 4702, 2020.

[7] G. A. Kumar, A. K. Patil, R. Patil, S. S. Park, and Y. H. Chai, "A lidar and imu integrated indoor navigation system for uavs and its application in real-time pipeline classification," *Sensors*, vol. 17, no. 6, p. 1268, 2017.

[8] T. D. Putri *et al.*, "Intelligent transportation systems (its): A systematic review using a natural language processing (nlp) approach," *Heliyon*, vol. 7, no. 12, 2021.

[9] F. Duffhauß, "Deep sensor data fusion for environmental perception of automated systems," Ph.D. dissertation, Universität Tübingen, 2024.

[10] R. Oterino-Bono, N. Pavón-Pulido, J. D. Blasco-García, J. A. López-Riquelme, M. Jiménez-Muñoz, J. J. Feliu-Batlle, and M. T. Herrero, "Deep learning methods integration for improving natural interaction between humans and an assistant mobile robot in the context of autonomous navigation," in *EPIA Conference on Artificial Intelligence*. Springer, 2022, pp. 537–548.