International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

An Autonomous Navigation Approach Based on Birds–Eye View Semantic Maps

Mrs.kavyashree N
Assistant Professor
Computer Science & Engineering
BGS Institute of Technology
Adhichunchanagiri University

Sinchana SY
21CSE403
Computer Science & Engineering
BGS Institute of Technology
Adhichunchanagiri University

ABSTRACT—This paper presents an autonomous navigation approach that draws inspiration from birds' navigation techniques, utilizing bird's eye view semantic maps. These maps encapsulate high-level semantic information about the environment, such as roads, sidewalks, obstacles, and landmarks. By integrating this semantic understanding with a comprehensive aerial perspective, our approach aims to enhance navigation capabilities in dynamic environments. We leverage advanced computer vision techniques to construct detailed semantic maps, enabling autonomous agents to plan and execute maneuvers effectively. The proposed approach offers a promising solution for enhancing navigation robustness and adaptability in various real-world scenarios, including autonomous vehicles, robotics, and unmanned aerial systems.

Keywords: Birds-Eye View, path planning, Motion Control, Real-time Navigation

1. INTRODUCTION

Autonomous navigation in dynamic environments presents a formidable challenge, demanding systems capable of perceiving surroundings and making real-time decisions. Conventional methods often rely on sensor data and predefined maps, which may lack the flexibility needed to navigate complex scenarios. This paper introduces an innovative approach to autonomous navigation utilizing bird's eye view semantic maps. Inspired by the navigation strategies observed in birds, our approach integrates high-level semantic information with a comprehensive view of the environment. By combining semantic understanding with geometric precision, our method aims to enhance navigation robustness and adaptability. This introduction provides an overview of the motivation, methodology, and contributions of our proposed approach, setting the stage for detailed exploration in subsequent sections.

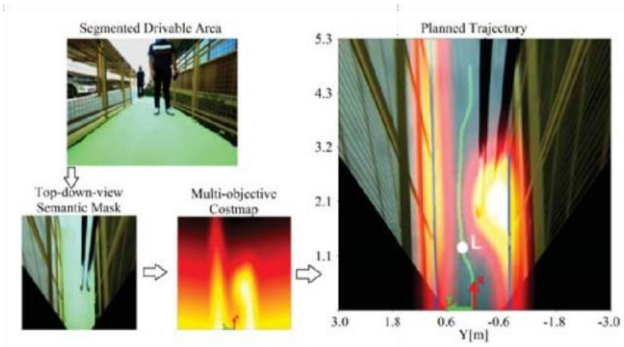

Fig.1. perspective–driven multi-objective path planning system

Autonomous navigation systems are integral to a wide range of applications, from self-driving cars to unmanned aerial vehicles, where the ability to perceive and interpret the surrounding environment accurately is paramount. Traditional navigation methods often rely on predefined maps and sensor data, yet they may struggle to adapt to dynamic and complex scenarios. To address these challenges, this paper proposes an innovative approach inspired by birds' navigational strategies, leveraging bird's

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

eye view semantic maps to enhance autonomous navigation capabilities.

Central to our approach is the construction of bird's eye view semantic maps using advanced computer vision techniques. These maps serve as a rich representation of the environment, capturing both semantic features and geometric layout. By encoding essential environmental information into semantic maps, including road structures, pedestrian pathways, and obstacles, our method enables autonomous agents to perceive and understand their surroundings comprehensively, facilitating informed decision-making during navigation tasks.

The proposed navigation approach holds significant promise for various real-world applications, including autonomous vehicles, robotics, and unmanned aerial systems. By leveraging bird's eye view semantic maps, our method offers the potential to enhance navigation robustness and adaptability, enabling autonomous agents to navigate safely and efficiently in dynamic and challenging environments. In the following sections, we provide a detailed description of our methodology, experimental results, and the implications of our approach for advancing autonomous navigation technologies.

2. RELATED WORK

A. Semantic segmentation

Semantic segmentation has experienced significan advancements in recent years, driven by profoun developments in brain organizations. pivot occurred with the introduction of convolution layers in Completely Convolutional Brai Organizations (FCNN). FCNNs not onl demonstrated remarkable performan improvements but also reduced the parameter cou compared to Completely Associated Brai areas Organizations. Subsequent progressions include the integration of multiscale features and t introduction of context aggregation, whic addressed critical issues faced by FCNNs such as the lack of contextual relationships and global information across diverse domains.

B. Vision-based Navigation using Deep Learning

3. SEMANTIC SEGMENTATION BASED VISUAL NAVIGATION

Vision-based route utilizing deep learning, particularly through end-to-end imitation learning, represents a significant advancement in achieving autonomous navigation solely from visual input. In this approach, a Convolutional Neural Network (CNN) is trained to directly map pixels from a front-facing camera to steering commands. However, the end-to-end nature of this system makes it challenging to interpret the inner workings of the CNN, complicating troubleshooting and correcting undesired behaviors.

Alternatively, some methods compute motion planning and control from a semantic mask generated from the first-person view. These methods leverage recent developments in deep learning-based perception tasks such as semantic segmentation. However, since motion is computed from semantic data at the image level, they lack spatial information about the surroundings necessary for smooth navigation.

1. Objective: The primary goal of the navigation system is to enable AMRs to navigate autonomously using only a front-facing RGB camera, without the need for depth or ranging sensors.
2. Semantic Segmentation: The system begins by processing the input RGB image using deep learning models for semantic segmentation. Specifically, models like PSPNet and FCHarDNet are employed to detect drivable areas in the scene. This segmentation allows the robot to identify obstacles, drivable paths, and other relevant features in its environment.
3. Bird's-Eye View Semantic Map: The drivable detected in the RGB image are transformed into a Bird's-Eye view semantic map. This transformation provides a top-down perspective of the environment, allowing the robot to perceive the

spatial layout and distances to objects more effectively.

4. Multi-objective Cost-map: From the Bird's-Eye view semantic map, a multi-objective cost-map is computed. This cost-map incorporates objectives such as obstacle avoidance, center-line following, and distance to goal reduction. It assigns costs to different areas of the map based on these objectives, guiding the robot in selecting safe and efficient paths.
5. Path Planning: Using the cost-map, the system calculates the lowest-cost path for the robot to follow. Path planning algorithms iterate over possible paths in the semantic map, considering the costs associated with each path, to find the optimal route while avoiding obstacles and adhering to navigation objectives.
6. Motion Control: Once the path is determined, a motion controller is utilized to steer the robot along the path. A pure-pursuit motion controller is employed, ensuring smooth and accurate navigation towards a lookahead point along the planned path.
7. Real-time Implementation: The entire system is implemented to operate in real-time, with computational efficiency considered for each processing step. This allows for responsive and timely navigation decisions by the robot as it encounters changes in its environment

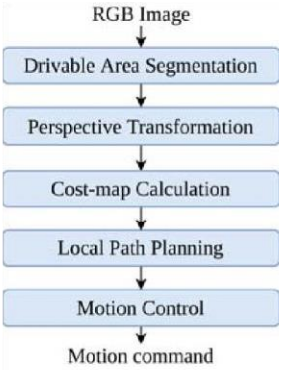

Fig 2: single front- facing image

The described model, FCHarDNet, employs a decoder component utilizing bilinear upsampling and 1x1 convolutional layers to restore diminished picture highlights to their original size. It incorporates residual connections between blocks, batch normalization, and ReLU activation functions. Initially initialized with pre-trained weights, FCHarDNet undergoes further training using Stochastic Gradient Descent and a Bootstrapped Cross-Entropy Loss.

In contrast, PSPNet (Pyramid Scene Parsing Network) is a Convolutional Neural Network (CNN) designed to capture global context information by aggregating region-based context through pyramid pooling modules. By integrating features across different scales, PSPNet effectively covers the entire, half, and smaller portions of the image. PSPNet has demonstrated superior performance compared to Fully Convolutional Neural Networks (FCNNs) on various datasets.

Figure 3 displays the semantic segmentation result, where drivable regions under different conditions are highlighted in green. This segmentation output showcases the effectiveness of the models in accurately identifying and delineating drivable areas within the image.

Fig. 3. Drivable area extraction using Semantic Segmentation. Here, PSPNet was used for Indoor and Outdoor environments, while FCHarDNet was used for our simulated environment

D. *Multi-objective Cost-map*

1.Obstacle Avoidance**:** It seems like you're describing a method for obstacle avoidance for a robot, particularly focusing on staying away from the edges of the drivable region. The formula you provided suggests a convolution operation between a Gaussian kernel and a mask representing the drivable region, which results in a cost function for obstacle avoidance.

The formula for obstacle avoidance

(JOA(x,y)) at a point (x,y) in the Bird's-Eye view semantic map is given by:

$$JOA(x,y)=G(\mu,\sigma 2)*MC(x,y)JOA(x,y)=G(\mu,\sigma 2)*M$$

Where:

- $JOA(x,y)JOA(x,y)$ is the obstacle avoidance cost function at point (x,y).
- $G(\mu,\sigma 2)G(\mu,\sigma 2)$ is a 2D Gaussian kernel with mean (μ) and variance (σ^2).
- $MC(x,y)MC(x,y)$ is the contours mask of the Bird's-Eye view semantic map.

The convolution operation between the Gaussian kernel and the contours mask effectively assigns higher costs to points near the edges of the drivable area, promoting obstacle avoidance behavior during path planning

2. **Center-line Following**: Center-line following is technique used in navigation systems to guide a mobile robot along a predefined path, typically

3. Integration into Cost-map: A term is included in the multi-objective cost-map to promote center-line following behavior. This term reduces the cost along the computed center line, encouraging the robot to stay close to the center of the drivable area during navigation.

3.Distance to goal reduction

Distance to goal reduction" refers to a component of the multi-objective cost-map used in the navigation system described in the paper. This component aims to guide the robot towards its goal position by assigning lower costs to points closer to the goal.

Here's an overview of how "distance to goal reduction" is implemented in the system:

along the center line of a drivable area. In the context of the semantic segmentation-based visual navigation system described in the paper, center-line following is one of the objectives considered i

the computation of the multi-objective cost-map.

Here's Detection of Side Boundaries: Some environments, such as hallways, sidewalks, and roads, have well-defined side an overview of how center-line following is implemented in the system

boundaries. The navigation system detects these boundaries, allowing it to identify the left and right
edges of the drivable area.

1. Computation of Center Line: Once the side boundaries are detected, the system computes the
1. center line of the drivable area. This center line represents the optimal path for the robot to follow within the drivable area. Calculation of Euclidean Distance: The distance
between each point in the Bird's-Eye view semantic map and the goal position is computed using the Euclidean distance formula. This distance represents the proximity of each point to a the goal.
2. Cost Assignment: Points closer to the goal position are assigned lower costs, while points farther away from the goal receive higher costs. The exact cost assignment strategy may vary, but typically it involves inversely proportional relationships, where closer points have lower costs and farther points have higher costs.
3. Integration into Cost-map**:** The costs

4. associated with "distance to goal reduction" are combined with other cost components, such as obstacle avoidance and center-line following, to form the multi- objective cost-map. This ensures that the navigation system considers proximity to the goal as one of the factors when planning the robot's path.

REFERENCES

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.

[2] Clement´ Godard, Oisin Mac Aodha, Michael Firman, and Gabriel J Brostow. Digging into self-supervised monocular depth estimation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3828– 3838, 2019.

[3] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881– 2890, 2017.

[4] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431– 3440, 2015.

[5] Bharath Hariharan, Pablo Arbelaez,´ Ross Girshick, and Jitendra Malik. Hypercolumns for object segmentation and fine-grained localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 447–456, 2015.