An AI Powered Tour Guide Robots: Technologies, Architectures, And Implementation for the Academic Center of Excellence

Vijaya Chavan[1], Komal Jadhav[2], Sandesh Satavnekar[3]

[1] *Vijaya Chavan Mechanical Engineering & VPKBIET Baramati*
[2] *komal Jadhav Mechanical Engineering & VPKBIET Baramati*
[3] *Sandesh Satavnekar Mechanical Engineering & VPKBIET Baramati*

---***---

Abstract -The increasing demand for interactive and autonomous information dissemination in specialized environments such as academic Center of Excellence has led to growing interest in AI-powered tour guide robots. This paper reviews the current state of such systems with a focus on developing a low-cost and efficient solution for the Centre of Excellence in Robotics and Automation under Industry 4.0. The study highlights core technologies including autonomous navigation, obstacle avoidance, and conversational artificial intelligence for real-time human–robot interaction. The review also explores system integration using embedded platforms such as Raspberry Pi 4GB, along with supporting technologies like computer vision, speech recognition, and natural language processing to enhance user engagement. Special attention is given to practical challenges such as localization accuracy, response latency, and operation in dynamic indoor environment Furthermore, key design considerations including cost-effectiveness, scalability, energy efficiency, and ease of deployment are discussed. The paper also outlines performance evaluation metrics such as navigation accuracy, interaction success rate, and user satisfaction. This study aims to provide a foundational reference for undergraduate B.E. projects by identifying research gaps, current trends, and best practices in the development of intelligent robotic tour guide systems.

Key Words: **- Tour Guide Robot, Artificial Intelligence, Autonomous Navigation, Human–Robot Interaction, Computer Vision.**

1. INTRODUCTION

Robots that interact and assist the public conjure up futuristic scenarios from television, cinema, and other media. This research focuses its attention on robots that interact with people indoors, such as waiters, domestic assistants, and museum tour-guides, among others. These types of robots must, however, be sufficiently well programmed for realistic human–machine interaction. In this study, the mobile robot designed to work in indoor environments as a laboratory tour-guide is called Doris. Programming a robot is no easy task in view of the many issues that need to be solved, such as the localization of robot, its movements, its ability to process information in complex environments, the definition of set conditions, the interpretation of an input and the appropriate response, and so forth. Once solved, these tasks will have to work in complete synchrony.

2. Body of Paper

Section 1.

(LITERATURE SURVEY)

The development of autonomous tour guide robots has evolved significantly over the past decades. Thrun et al. [3] demonstrated the feasibility of autonomous robot navigation in public areas. It has various limitations such as computing resources and sensor technology [9]. This developer system mostly relies on laser range finders and pre-mapped environments. It also has limitations such as limited adaptability and increasing deployment costs. This research considers modern approaches in computer vision and machine learning systems [2]. It offers more flexibility and cost-effective solutions. Visual SLAM as another alternative to traditional sensor-based localization methods [10] Redmon et al. [6] incorporated the capabilities of the camera insect localization method. ORB-SLAM is a feature-based method that is used to significantly improve and increase computational efficiency [11, 12]. Recent research focusing on real-time performance and robustness in dynamic environments [11].

ORB-SLAM2 has implemented embedded hardware for practical deployment in the laboratory, which I consider to be significantly cost-effective [3]. Recent advancements in visual SLAM for dynamic environments have shown promising results. Bescos et al. [13] introduced an approach for tracking, mapping, and inpainting in dynamic scenes, which is particularly relevant for tour guide robots operating in environments with moving people and objects. Similarly, Yu et al. [14] presented a semantic visual SLAM system that integrates semantic segmentation to better handle dynamic objects. More recently, Zhu et al. [15] proposed a double-constrained visual SLAM approach for realistic map reconstruction in dynamic scenes, offering improved robustness

For navigation in dynamic environments, deep reinforcement learning approaches have shown significant potential [16], though they often require extensive computational resources not suitable for embedded systems like Raspberry Pi [17]. Deep learning has been used for robotic perception. It gives importance to any objective question and visual perception [18]. The evolution has been from traditional computer vision methods to deep learning methods. Architectures like R-CNN have enabled [10]. Redmon et al. [6] has been used for real-time object detection with high accuracy. YOLOv5 has used a balance of speed and accuracy, YOLO is also suitable for real-time robotic applications in which computing resources are limited. Using advanced technologies, robots are able to understand and perceive the environment more effectively.

Section 2.

(System Architecture)

The system architecture of the proposed AI-based tour guide robot consists of multiple integrated modules that work together to achieve autonomous navigation and intelligent human interaction. The overall system is divided into hardware and software components. The **input layer** includes sensors such as a camera module, ultrasonic sensors, and optional LiDAR for environment perception. The camera captures real-time images, which are processed using computer vision techniques for object detection and path recognition.

The **processing unit** consists of a microcontroller or single-board computer such as Raspberry Pi, which acts as the brain of the system. It processes sensor data and implements AI algorithms including SLAM and deep learning models (e.g., YOLO) for navigation and decision-making.

The **interaction module** enables communication between the robot and users. It includes speech recognition and text-to-speech systems, allowing the robot to answer queries and provide information about different locations. The **control system** manages motor drivers and actuators to control the robot's movement. Based on the processed data, it sends signals to motors for forward, backward, and directional motion.

The **output layer** includes display units, speakers, and movement of the robot. The robot delivers audio or visual information to users while guiding them to specific locations. Overall, the system integrates **AI, computer vision, and embedded systems** to create an efficient, low-cost, and intelligent tour guide robot suitable for smart lab environments.

Section 3.

(METHODOLOGY)

The robot platform is assembled with properly selected components. The performance of this project has to balance cost and reliability. Raspberry Pi 4 with 4 GB RAM is a main processing unit in this system. It provides enough computing power for SLAM at the same time. It is helping in object detection and path planning. Arduino Uno is capable of handling motor controls according to high level decisions. Webcam Logitech C920 HD Pro Webcam is used as the primary visual sensor

1) Data Acquisition: The robot collects real-time data using sensors such as a camera and ultrasonic sensors. The camera captures images of the environment, while ultrasonic sensors detect obstacles.

2) Pre-processing: The captured visual data is processed using image processing techniques such as noise reduction and feature extraction, this improves the accuracy of object detection and navigation.

3) Environment Mapping: Simultaneous Localization and Mapping is used to create a map of the environment and determine the robot's current position. This allows the robot to navigate efficiently in unknown areas.

4) Object Detection and Recognition: Deep learning algorithms such as YOLO (You Only Look Once) are used to detect objects, landmarks, and obstacles in real-time. This helps the robot identify important locations within the lab.

5) Path Planning: Based on the generated map and detected obstacles, the system calculates the optimal path using AI-based algorithms. The robot dynamically adjusts its route to avoid collisions

6) Navigation and Control: The control unit sends signals to motor drivers to move the robot in the desired direction. The robot continuously updates its position and path during movement.

7) Human-Robot Interaction: Speech recognition and text-to-speech systems are used to interact with users. The robot can answer queries and provide information about different locations in the lab.

8) Output and Feedback: The robot provides output through speakers or display and guides users to the desired location. Feedback is continuously used to improve navigation and interaction.

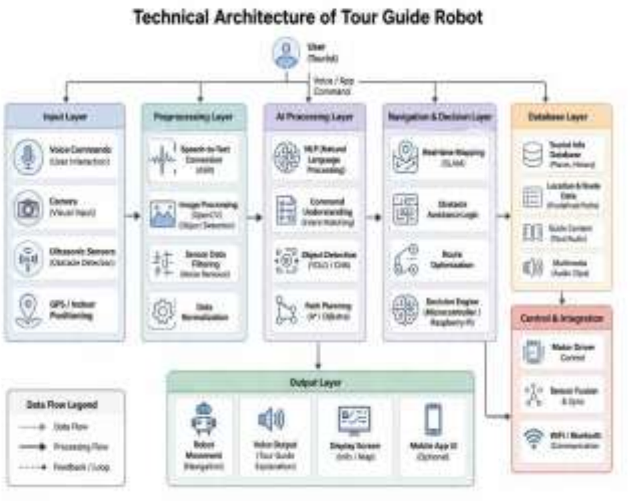
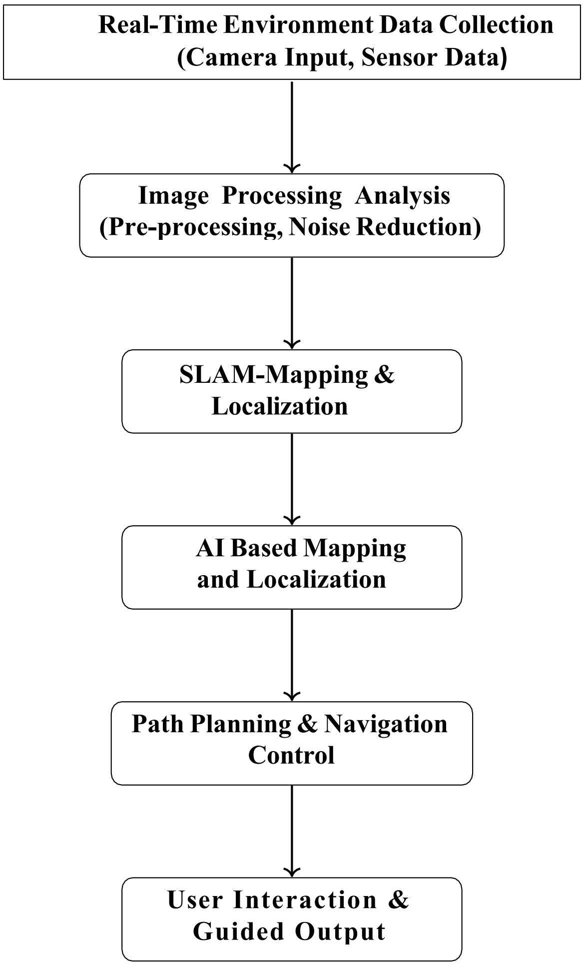


Fig -1: System Architecture

Table -1:



Section 4.
(RESULTS AND DISCUSSION)

Table 1 gives a comparison of AI based SLAM system and traditional line following system. This AI based SLAM system is better in all parameters except computational load. It reduces the time by 40% in two times and it shows that this system is more efficient than traditional system. The developed system reduces the deviation by 65.6% i.e. from 3.2 cm to 1.1 cm. It shows that SLAM provides excellent localization accuracy. This SLAM tracking provides localization success rate of 98.2%. In this system, primary failures occur at 15% turns and 72% in low texture areas. Illumination changes are causing 13% failure results. Regression analysis is used to find the correlation between path deviation and SLAM tracking stability. A strong negative correlation $R^2 = 0.83$ was found between SLAM feature number and path segmentation. If ORB feature tracking is reduced to less than 300 features per frame, the path deviation increases by an average of 2.7 cm. This shows that sufficient feature density is important for navigation accuracy.

3. CONCLUSIONS

The tour guide robot developed in this project provides an efficient and automated solution for guiding visitors **in** various environments such as colleges, museums, and public places. The system successfully integrates sensors, microcontrollers, and basic programming to achieve navigation and obstacle detection.

This robot reduces human effort and improves user experience by offering accurate guidance and information. It is cost-effective, easy to implement, and suitable for real-time applications

In future, the system can be enhanced by incorporating advanced technologies such as artificial intelligence, voice recognition, and mobile application control to make it more intelligent and interactive**.**

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Mechanical Engineering, VPKBIET Baramati, for providing the necessary infrastructure and support that made this research possible. The authors are highly thankful to their project guide, Dr. Vinod Todkari, for her constant guidance, encouragement, and valuable insights throughout the development of the AI-powered tour guide robot. Her support played a vital role in the successful completion of this project. The authors also extend their appreciation to all faculty members and staff of the department for their continuous support, suggestions, and cooperation during the implementation and documentation phases of this work. Finally, the authors would like to thank their colleagues and friends for their motivation and support throughout the project duration

REFERENCES

[1] Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J.J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on Robotics, 32(6): 1309-1332. https://doi.org/10.1109/TRO.2016.2624754

[2] Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A.B., Dellaert, F., Fox, D., Hähnel, D., Rosenberg, C., Roy, N., Schulte, J., Schulz, D. (2000). Probabilistic algorithms and the interactive museum tour-guide robot Minerva. The International Journal of Robotics Research,19(11): 972-999. https://doi.org/10.1177/02783640022067922

[3] Thrun, S., Bennewitz, M., Burgard, W., Cremers, A.B., Dellaert, F., Fox, D., Hähnel, D., Rosenberg, C., Roy, N., Schulte, J., Schulz, D. (1999). Minerva: A second-generation museum tour-guide robot. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Detroit, MI, USA, pp.1999-2005.

[4] Dissanayake, M., Newman, P., Clark, S., Durrant-Whyte, H.F., Csorba, M. (2001). A solution to the simultaneous localization and map building (SLAM) problem. IEEE Transactions on Robotics and Automation, 17(3): 229-241. https://doi.org/10.1109/70.938381

[5] Alvarado Vasquez, B.P.E., Gonzalez, R., Matia, F., De la Puente, P. (2018). Sensor fusion for tour-guide robot localization. IEEE Access, 6: 78947-78964. https://doi.org/10.1109/ACCESS.2018.2885648

[6] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You Only Look Once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779-788.https://doi.org/10.1109/CVPR.2016.91

[7] Hart, P.E., Nilsson, N.J., Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics, 4(2): 100-107.

[8] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature, 521(7553): 436-444. https://doi.org/10.1038/nature14539

[9] Bailey, T., Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): Part II—State of the art. IEEE Robotics & Automation Magazine, 13(3): 108-117. https://doi.org/10.1109/MRA.2006.1678144

[10] Girshick, R., Donahue, J., Darrell, T., Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), USA, pp. 580-587. https://doi.org/10.1109/CVPR.2014.8