IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Nex: Next-Gen Exo-Humanoid Ai-Based Robot

[1]Kathiresan A, [2]Deepak B, [3]Madesh Aadithya S B, [4]Sarvesh K
[1]Head of the Department, [2]Student, [3]Student, [4]Student
[1]Information Technology,
[1]PSG Polytechnic College, Coimbatore, India

Abstract: This paper presents NEX, an AI-powered Exo-humanoid robot designed for navigation and personalized assistance in institutions such as universities, hospitals, and museums. The robot integrates real-time navigation using the A* algorithm, obstacle detection with ultrasonic sensors, facial recognition, and natural language interaction to enhance user experience. Its unique multi-modal interaction capabilities and advanced AI-powered systems make it a next-gen solution for simplifying complex layouts, providing guidance, and improving accessibility. This study outlines the design, implementation, and testing phases, demonstrating its potential as an effective, adaptive guide.

Index Terms: AI-based robot, Exo-humanoid, navigation system, A* algorithm, facial recognition, ultrasonic sensors, personalized assistance, natural language processing, multi-modal interaction.

I.INTRODUCTION:

1.1 Background:

Advancements in artificial intelligence (AI) and robotics have enabled the development of systems that can mimic human capabilities in navigation, communication, and assistance. In institutions like universities, hospitals, and museums, effective navigation and personalized assistance are critical. Traditional systems often lack the adaptability and user-friendliness required in such environments, creating a demand for smarter, more interactive solutions.

1.2 Overview of the NEX Project:

NEX is an AI-powered Exo-humanoid robot that bridges the gap between conventional navigation aids and advanced AI-driven systems. Equipped with real-time navigation using the A* algorithm, obstacle detection via ultrasonic sensors, facial recognition, and natural language interaction, NEX aims to redefine user interaction by simplifying complex layouts and offering personalized assistance.

1.3 Problem Addressed:

Institutions often face challenges in guiding visitors through intricate layouts and providing personalized support. Traditional signage or static systems can be insufficient, particularly for individuals requiring accessibility features. NEX addresses these challenges by offering dynamic, real-time assistance, ensuring users can navigate efficiently while enhancing their overall experience.

II. EXISTING AND PROPOSED SYSTEMS:

2.1 Existing Systems

Existing systems commonly rely on traditional methods like static signage, information desks, and basic robotic systems that offer limited interaction capabilities. These solutions typically lack the ability to adapt to changing environments or user needs. Static signage and information desks are fixed in location, requiring users to actively seek out information, which can be inefficient, especially in large or complex spaces. Basic robotic systems often have rudimentary functions with limited interactivity, such as simple task execution or movement within predefined paths. These systems are typically unable to provide real-time navigation or respond to obstacles dynamically. Additionally, they lack multi-modal interaction options, such as voice, gesture, or facial recognition, which are crucial for providing a personalized and seamless experience. As a result, these systems fall short in terms of accessibility, efficiency, and user satisfaction, especially in environments that require high levels of flexibility and personalization.

2.2 Proposed System

The proposed NEX system addresses the limitations of existing systems by integrating advanced AI-powered modules that enable real-time navigation, obstacle detection, facial recognition, and natural language interaction. These capabilities allow NEX to provide dynamic guidance to users, ensuring that they can navigate complex environments efficiently. The real-time obstacle detection ensures that the system can adjust its path instantly, avoiding collisions and ensuring smooth movement. Facial recognition adds a personalized layer to the system, allowing it to identify users and offer tailored information based on their preferences or past interactions. Additionally, NEX's natural language interaction capabilities enable users to communicate with the system using voice commands, making it more intuitive and user-friendly. The modular architecture of NEX further enhances its potential by ensuring scalability and adaptability across a wide range of institutional settings, from hospitals and universities to corporate offices and public spaces. This flexibility allows NEX to be customized for specific needs, making it a powerful solution for environments that require high levels of user engagement and accessibility.

III. System Architecture

3.1 Hardware Components

The hardware components of the system include an Arduino controller for managing and controlling the overall operations. A custom-designed PCB board facilitates connections between various parts. The system utilizes four Johnson motors (12V, 60 RPM, grade A) with motor clamps to drive movement, while a NEMA 17 stepper motor is employed for precise control. The L298N Dual H-bridge motor driver powers the stepper motor, ensuring smooth operation. Communication between components is handled by an RC transmitter and receiver. An ultrasonic sensor aids in obstacle detection, while a web camera supports computer vision tasks. A two-pin screw terminal is used for secure electrical connections. The system is powered by a 12V, 7Ah lead acid battery, charged through a 12V, 8Ah charger, with a buck converter regulating power to 12V at 10Ah. Additional components include a speaker for audio feedback, a motor driver (L298N) for motor control, a voltage indicator for monitoring power levels, stepper motor couplings for secure connections, and wheels (100mm diameter, 40mm width) for mobility.

3.2 Software Components

The system's software stack includes Python for running AI algorithms, enabling intelligent decision-making and interaction. OpenCV is used for computer vision and facial recognition, enhancing user interaction and navigation capabilities. ROS (Robot Operating System) integrates all system components, allowing them to work together seamlessly. The Arduino IDE is used for programming and controlling the hardware, while the Natural Language Toolkit (NLTK) enables voice interaction with the system. A real-time operating system coordinates various tasks and schedules operations to ensure the system functions efficiently and reliably.

IV. MODULES

4.1 Camera Module

The Camera Module serves as the system's primary visual input. It leverages a web camera in conjunction with OpenCV to enable advanced facial recognition and environment mapping. This module identifies and authenticates users, tracks their movements, and interprets visual cues from the environment. Additionally, it collects data to assist other modules, such as gesture recognition and navigation.

4.2 Speech Module

The Speech Module facilitates dynamic natural language interaction. Utilizing speech-to-text and text-to-speech capabilities, this module employs NLTK for linguistic processing and integrates with various voice APIs for real-time communication. It supports multiple languages and adapts to user-specific accents and speech patterns for a personalized conversational experience.

4.3 Gesture Module

The Gesture Module provides a hands-free interaction option by processing user gestures as commands. It utilizes a combination of computer vision and machine learning techniques to detect and classify gestures. This feature enhances accessibility for users with limited mobility or speech impairments and enables intuitive control over the system's functionalities.

4.4 Navigation Module

The Navigation Module employs the A* algorithm to calculate optimal paths in real-time. It integrates ultrasonic sensors to detect and circumvent obstacles, ensuring safe and efficient movement in dynamic environments. This module is critical for enabling autonomous navigation in complex institutional layouts, such as hospitals, airports, and universities.

4.5 Information Guiding Module

The Information Guiding Module delivers personalized assistance by leveraging location data, user preferences, and contextual inputs. It combines data from the Camera and Speech Modules to provide relevant guidance and information, enhancing user satisfaction and engagement. This module also adapts its responses based on historical user interactions.

4.6 Integration

The seamless integration of hardware and software components is fundamental to the system's functionality. A robust communication framework ensures efficient data exchange between modules. Real-time processing capabilities enable the system to analyze inputs, execute commands, and provide feedback instantaneously. This holistic integration supports the system’s multitasking capabilities, ensuring uninterrupted user interaction.

V. SYSTEM FLOW

5.1 User Approach Detection

The system’s user approach detection begins with the activation of camera sensors, which continuously scan the environment for signs of user presence. Once a user is detected within the system’s designated proximity, the sensors trigger the initialization of the interaction process. This involves activating several key modules—such as the Camera, Speech, and Gesture Modules—to prepare for multi-modal inputs. The detection algorithm has been meticulously fine-tuned to minimize false positives, ensuring that the system only activates when an actual user approaches. Furthermore, response times are optimized, so the system can quickly transition into active interaction mode as soon as the user is detected, providing a seamless and efficient experience. The system is also designed to handle varying lighting conditions and user movement speeds, ensuring reliable detection in diverse environments.

5.2 Voice Command Recognition

Voice command recognition plays a central role in the system's interaction capabilities. The Speech Module captures voice commands, which are then processed through advanced natural language processing (NLP) algorithms. These algorithms are capable of accurately interpreting a wide range of user intents, even in noisy or acoustically challenging environments. The NLP system uses context-aware processing to distinguish between different types of commands and prioritize user requests. Additionally, a feedback mechanism is integrated into the system to handle ambiguous or unclear commands. If the system detects any uncertainty in understanding, it prompts the user for clarification, ensuring robust and reliable communication. This mechanism greatly enhances the system's flexibility, making it capable of handling diverse voices, accents, and speech patterns without sacrificing accuracy.

5.3 Face Recognition

The Camera Module incorporates facial recognition technology to identify and authenticate users. Upon detection, the system uses this data to personalize interactions, providing access to user-specific preferences, histories, and tailored recommendations. This feature ensures that the system can dynamically adjust its responses based on past interactions or specific user needs. The face recognition process is powered by a continuously updated database that ensures high accuracy, even as new users are introduced. Over time, the system learns to recognize subtle variations in facial features, improving its ability to authenticate users reliably. Additionally, continuous updates to the recognition database ensure that the system can adapt to changes in a user's appearance or lighting conditions, maintaining a high level of personalization and security.

5.4 Response Generation

The system generates responses based on user queries and environmental inputs, delivering context-aware answers that adapt to the situation. Depending on user preferences and the context of the interaction, responses can be delivered through speech, text, or visual displays. The response generation process is designed to handle complex queries by integrating data from multiple system modules, including navigation, facial recognition, and environmental sensors. This integration allows the system to provide dynamic, relevant information, such as offering directions in response to location-based queries or providing personalized recommendations based on the user's previous interactions. The system's ability to handle a wide variety of user requests—from simple commands to complex, multi-step tasks—ensures that it can cater to a diverse range of needs in real-time.

5.5 Obstacle Avoidance and Navigation

The Navigation Module leverages ultrasonic sensors to detect obstacles in the system's path, ensuring that the system can navigate its environment safely and efficiently. Once an obstacle is detected, the system recalibrates its path dynamically using the A* algorithm, which continuously computes the most optimal route to the destination while avoiding detected obstacles. In addition to real-time recalibration, the system incorporates predictive algorithms that analyze environmental patterns and anticipate potential hazards, such as a moving object or a sudden change in terrain. By predicting obstacles before they are directly encountered, the system ensures smooth and proactive navigation. This combination of real-time obstacle avoidance and predictive algorithms guarantees both safety and efficiency during the system's operation, whether in open spaces or densely populated areas.

5.6 Gesture Control

Gesture control allows users to interact with the system without the need for verbal communication. The system can process a wide range of gestures, which are interpreted as non-verbal commands. This feature is particularly useful in environments where speech is not feasible, such as in noisy areas, or where users may have difficulty using voice commands due to physical constraints. The system supports a broad array of gestures, including hand movements, body posture changes, or specific finger signs. Furthermore, it allows for customizable gestures tailored to individual user needs, ensuring that the system can be adapted for specific tasks or preferences. Gesture recognition expands the system's versatility, providing an intuitive and hands-free way for users to control their environment, thus enhancing the overall user experience.

5.7 Ongoing Interaction Loop

The ongoing interaction loop is a core aspect of the system's design, ensuring continuous engagement with the user. The system is constantly monitoring user inputs, whether through voice, gestures, or facial recognition, and adjusting its responses based on these inputs. In addition to this, the system also takes environmental changes into account, allowing it to adapt dynamically to factors like shifting obstacles or changing lighting conditions. This continuous feedback loop ensures that the system is always responsive to both user actions and the environment, offering a truly adaptive interaction. Moreover, the system integrates real-time updates and adaptive learning mechanisms that allow it to refine its responses and interactions over time. By learning from user preferences, behaviors, and environmental patterns, the system continually improves its efficiency and accuracy, creating a more personalized and responsive experience for users.

VI. IMPLEMENTING AND TESTING:

6.1 Navigation System

The Navigation System was designed to operate efficiently in dynamic and unpredictable environments, such as large institutions with varied layouts. It utilizes the A* algorithm to calculate the shortest path between points while continuously adjusting to obstacles detected in real time. The system incorporates feedback from ultrasonic sensors to make dynamic path adjustments, ensuring smooth navigation even in crowded or changing environments. Extensive field testing in real-world institutional settings, such as hospitals and universities, confirmed its reliability and effectiveness. The system proved capable of efficiently managing complex, dynamic routes while avoiding obstacles, making it suitable for a wide range of applications where precise and adaptive movement is critical.

6.2 Obstacle Detection

The obstacle detection subsystem is powered by high-precision ultrasonic sensors, which provide real-time feedback about the environment. These sensors are strategically placed to detect obstacles at various distances, allowing the system to recognize and respond to objects or obstructions in its path instantly. This capability is crucial for maintaining continuous operation in complex, unpredictable environments. Whether navigating through narrow hallways or crowded spaces, the ultrasonic sensors ensure that the system can avoid collisions and navigate around obstacles effectively, thereby reducing the risk of damage to both the system and its surroundings. This precise obstacle detection feature was rigorously tested to ensure its effectiveness under different conditions, proving its reliability in enhancing the system's safety and efficiency.

6.3 Personalized Interaction

To create a more engaging and personalized experience, the system integrates facial recognition technology. This allows the system to identify users and tailor interactions based on individual preferences, previous interactions, and contextual data. By recognizing returning users, the system can adjust its responses, providing more relevant and personalized information. This feature significantly enhances user satisfaction, particularly in environments where frequent interaction with the system is required. In institutional settings, such as hospitals or schools, this personalized approach can help the system deliver information more effectively, anticipate user needs, and create a more intuitive experience. The integration of facial recognition was thoroughly tested and proven to have high accuracy, ensuring seamless and reliable user identification even in busy or dynamic environments.

6.4 Voice Integration

Natural language processing (NLP) was incorporated into the system to enable seamless voice-based interaction. The voice integration allows users to issue commands and ask questions, with the system able to process and respond to queries in real-time. The Speech Module is capable of understanding complex queries and recognizing contextual nuances, ensuring that the system’s responses are appropriate and accurate. This capability was thoroughly tested across diverse user groups, including people with different speech patterns and accents, to ensure that the system could accommodate a wide range of users. The result is a highly responsive and intuitive voice interface that improves accessibility, particularly for users with mobility challenges or those who prefer hands-free interaction.

6.5 Testing Outcomes

Comprehensive testing of the NEX system confirmed its ability to function reliably in a variety of real-world environments. The navigation system demonstrated consistent performance in complex and dynamic settings, handling both pre-planned and spontaneous path changes effectively. The obstacle detection system was highly successful in preventing collisions, ensuring safety while navigating through narrow or crowded areas. The facial recognition system proved to be accurate and reliable, offering personalized interactions based on user history and preferences. Voice integration also performed well, with the system understanding and responding to complex queries with minimal latency. Additionally, the gesture recognition system was successfully validated, allowing users to execute commands through non-verbal gestures, further enhancing the system's versatility and user-friendliness.

VII. RESULTS:

The NEX system successfully met its objectives by enhancing navigation with integrated Navigation and Obstacle Detection Modules for smooth movement and real-time path recalibration. It improved user interaction through multi-modal input processing, including voice, gesture, and facial recognition, providing a highly accessible and adaptable experience. Personalized assistance was achieved by combining user preferences, facial recognition, and contextual data, significantly boosting user satisfaction. Additionally, NEX streamlined operations in institutional settings by simplifying navigation, enhancing accessibility, and reducing human staff workload, establishing itself as a reliable and efficient solution for next-generation robotic systems.