

Posture-Based Human-Mimicking Humanoid Robot Interaction System: A Review

Dr. A. V. Dehankar, Dipanjali Warade, Pooja Raut , Pooja Chandewar , Pratiksha Motghare , Yashika Wanjari

Department of Computer Technology Priyadarshini College of Engineering – Nagpur

Abstract: The convergence of advanced computer vision techniques and modern robotics has ushered in a new era of human-mimicking robots, capable of replicating human actions with remarkable accuracy. This project summary outlines a detailed framework for the development of a sophisticated humanoid robot, integrating cutting-edge computer vision capabilities with a robust microcontroller architecture. The literature review highlights the evolution of human-robot interaction, with a specific focus on human motion recognition and the increasing relevance of Kinect-based robotics research. By emphasizing the importance of these advancements, the project delineates a strategic approach centered on creating an organized physical system. Key components, including the camera controller, motor driver controller, and microcontroller, are essential to the system's design, ensuring seamless integration and efficient operation. The primary objective of this project is to facilitate intuitive and precise interactions between humans and robots, enhancing safety and efficiency across a diverse array of industries. By leveraging the latest technological advancements and insights from the literature review, this initiative aims to expand the possibilities of human-robot collaboration, ultimately paving the way for transformative applications in various domains such as healthcare, manufacturing, and entertainment. Additionally, the project explores the critical role of human pose estimation in analysing human activities, allowing for the execution of complex tasks based on recognized gestures.

revolutionized various industries, including healthcare, customer service, manufacturing, and entertainment. With their ability to mimic human actions, these robots have the potential to perform complex tasks that were previously exclusive to humans. They can assist in delicate surgical procedures, provide personalized care to the elderly and individuals with disabilities, engage customers in retail environments, and even act as companions for people in need of social interaction. These robots are equipped with sensors, actuators, and intricate programming that enable them to perceive their environment and respond accordingly. They often feature a sophisticated range of facial expressions, voice modulation, and body movements, allowing them to establish a sense of connection and empathy with humans. Despite their advanced capabilities, human mimicry robots raise ethical and societal questions. Issues such as the potential displacement of human workers, the boundaries of human-robot relationships, and concerns about privacy and security are actively debated within the field of robotics and AI ethics. As the technology continues to evolve, human mimicry robots are expected to play an increasingly integral role in various aspects of human life, fostering a new era of collaboration between humans and machines. With ongoing advancements in AI and robotics, the future holds the promise of further seamless integration between humans and humanoid robots, transforming the way we live, work, and interact with technology.

I. INTRODUCTION

Human mimicry robots, also known as humanoid robots, are cutting-edge machines designed to replicate human actions and behaviors. These robots are engineered with the goal of imitating human movements, gestures, and even expressions to a high degree of accuracy. By employing advanced artificial intelligence (AI), machine learning, and sophisticated mechanical engineering, these robots can interact with humans in a way that simulates human-like communication and behavior. The development of human mimicry robots has

II. RELATED WORK

A. Human-Robot Interaction for robot control

The main purpose of Human-Robot Interaction (HRI) is the understanding, designing and evaluating robotic systems in order to establish a communication between humans and robots. This is important for the interest on keeping human operators in the control loop, being still important assets for an efficient industrial productivity even though the investment of

robots for automating processes is increasing dramatically. Robot manufacturers are developing different types of robots that can be integrated in different domains and are not restricted only for industrial use. Nowadays, there are available in the market mobile robots, which can work along with the human or can be manipulated from great distances to protect human from coming in contact with hazardous situations. Another case to consider the social robots that can interact with people for different purposes, such as providing elderly care[2]. There are a few surveys in the literature recognizing human activities. The development of human-mimicking robots has its roots in several interdisciplinary fields, including robotics, artificial intelligence, biomechanics, and psychology.

HRI is a field that attracts the interest of many researches across the world, and as a result, a lot of research work have been in conducted in this field. The European Union is not an exception as it funds research works concerning HRI. In the scope of the EU project ROBO-PARTNER, Michalos et al. , present three cases of HRI, for executing assembly tasks in an automotive production line. In the first one, robot and human are executing different assembly tasks, while they share the workspace[2]. The second case involves a mobile robot that does not have a direct intervention to the process, but provides to the human operator necessary tools and parts to complete a task.

B. Gesture recognition utilized in robotic applications

Gesture recognition in the field of robotics is a topic of interest for many researchers worldwide. Different methods are used for gesture recognition to achieve HRI. Movements and motions of human's body such as hand, head or the whole body, are captured using mainly cameras, depth sensors or data gloves .

Back in 2005, a gesture-based interface for a robot competition was introduced. In this work, two cameras in fixed positions capture the gestures to control the robot. The interface extracted the skin color and processed the position of the head and the hands, and translated the pose of user into the corresponding gesture. Robot received a command to follow the direction or action depending on the recognized gesture. Nickel's and Stiefelhagen's , used a stereo camera to detect head and hands of the human operator as long as his/hers head orientation to control ARMAR humanoid robot. The conducted experiments of this research work, proves that head

orientation features reduce the error for gesture recognition.

III. EXISTING METHODS

3.1. Kaspar

Another popular solution is Kaspar (Kinesics And Synchronization in Personal Assistant Robotics), which is almost a child-sized humanoid robot with unique features designed to help children with autism develop their social and interaction skills. This robot is still going through a lot of research based enhancement and expected to be available soon. It still carries a high price tag. There are many similar attempts being made in various research communities but primarily for enhancing the social and communication skills.

The proposed solution is aimed at improving on the social skill set of a variety of children with ASD, but at the same time provides them with a toy-like playmate to involve them in activities that need them to use more hand-eye coordination as well as enhances focusing capabilities in an activity. This is done by combining a low-cost, off-the-shelf available toy robot with the power of wireless control provided with easily developed activities custom-built for that specific child. The most intriguing part of this solution is that the activities can be custom-build by the parents, teachers, therapists, and other care-givers for such children with very easy procedures and almost no programming background being required. [2]And yet, the cost has been kept very low in order to make the solution very affordable.

3.2. Robotic arm

The robotic arm was based on the InMoov robot designed by Gaël Langevin [15]. The build of the arm contains several servo motors. Two TowerPro MG946R motors were used for the Shoulder and Bicep joints. These servo motors were converted to continuous servo motors by replacing the internal potentiometer that evaluates current positions with a voltage divider. Six TowerPro MG996R motors were used to actuate the fingers and wrist[3].

3.3. Gesture-Based Robot

A gesture-based interface for a robot competition was introduced . In this work, two cameras in fixed positions capture the gestures to control the robot. The interface extracted the skin color and processed the position of the head and the hands, and translated the pose of user into the corresponding gesture. Robot

received a command to follow the direction or action depending on the recognized gesture.

Nickel's and Stiefelhagen's, used a stereo camera to detect head and hands of the human operator as long as his/hers head orientation to control ARMAR humanoid robot. The conducted experiments of this research work, proves that head-orientation features reduce the error for gesture recognition.

Gesture based applications relied on monocular and stereo cameras until the introduction of devices that integrate depth sensors and infrared (IR) cameras allowing the development of systems based on body gestures for HRI. A humanoid robot can imitate the motion and movements of the human upper-body using Kinect sensor [30]. The same sensor was used by Qian et al, to control a dual-arm robot by gestural commands[4]. The dual-arm robot mimics human actions such as move left or right. Out of this work, Qian et al., achieve 85% gestural recognition accuracy.

3.4. Machine Learning and AI-Based Methods

Machine learning (ML) and artificial intelligence (AI)-based methods play a pivotal role in advancing posture mimicking systems by enabling robots to replicate human movements with high accuracy and adaptability. These methods are data-driven and allow robots to learn from large datasets of human movements, making them more dynamic and capable of mimicking even complex and subtle postures in real-time. Unlike traditional, rule-based systems, ML and AI leverage algorithms to identify patterns, predict future actions, and continuously improve posture replication through experience. The process begins with the collection of extensive datasets, often comprising images, video footage, or sensor data that capture human movements. These datasets are annotated to include essential information about joint positions, body orientations, and movement patterns. Once this data is gathered, machine learning models are trained to process it. Pose estimation, a crucial step in posture mimicking, is achieved through deep learning techniques, especially convolutional neural networks (CNNs). These networks analyze visual data to detect key points on the human body, such as the head, shoulders, elbows, wrists, hips, knees, and ankles. For more accurate posture replication, both 2D and 3D pose estimation methods are employed, with 3D models providing depth and spatial information for better motion replication. Once the human posture is captured, AI algorithms translate this information into commands for the robot's

actuators, ensuring the robot's joints and limbs mirror the detected movements.

Another powerful aspect of AI in posture mimicking is its ability to predict future movements. By analyzing sequential data, recurrent neural networks (RNNs) or transformer-based models can anticipate the next actions or postures a person will adopt, ensuring that the robot smoothly transitions between movements. This predictive ability is especially useful in applications such as physical therapy, where robots need to anticipate and react to human motions in a continuous, fluid manner.

3.5. Dynamic Balancing and Human-Like Control System

Dynamic balancing refers to the ability of a robot to maintain stability while in motion, such as walking, running, or shifting weight. Unlike static balancing, which focuses on stability when a robot is stationary, dynamic balancing ensures that the robot adjusts its posture and movements in real-time to avoid tipping over. This is achieved through a combination of sensors, such as gyroscopes and accelerometers, and algorithms that predict and adjust the robot's movements to keep its center of mass within a stable support area, like the feet. For instance, techniques like the Zero Moment Point (ZMP) method are used to calculate where forces acting on the robot balance out, ensuring stability during motion.

Human-like control systems aim to replicate the intuitive and adaptable way humans move and interact with their environment. These systems model human biomechanics, allowing robots to move fluidly and adapt to changing circumstances. Feedback loops are essential for adjusting movements based on sensory input, similar to how humans adjust their posture or movements in real-time. Machine learning algorithms help robots learn from inverse kinematics ensure precise limb control. These systems also optimize energy efficiency, allowing robots to perform tasks in a more natural, less energy-intensive manner. By integrating sensory feedback and advanced algorithms, humanlike control systems enable robots to perform dynamic actions with smooth, human-like precision.

3.6. HRNet

The proposed HRNet maintains high-resolution representations throughout the process. We start from a high-resolution convolution stream and add high-to-low resolution convolution streams one by one. The multiresolution streams are connected in parallel.

We obtain a network consisting of several (four in the current design) stages, and the n th stage contains n streams corresponding to n resolutions. We conduct multi-resolution fusions, exchanging the information across the parallel streams over and over[5].

The semantic strength and the spatial precision of the high-resolution representation learned from HRNet lie in the following two aspects: First, our approach connects the convolution streams of high-to-low resolution in a parallel manner rather than in a serial one. Therefore, compared with recovering the high resolution from the low resolution, our approach can maintain the high resolution directly, and thus the learned representation is spatially more precise. While most of the existing schemes fuse the high-resolution low-level and upsampled low-resolution high-level representations, we propose performing repetitive multi-resolution fusions to enhance higher resolution representations using the low-resolution representations and viceversa. Thus, all the representations of high-to-low resolution become semantically stronger[5].

IV. PROPOSED SYSTEM

Reviewing the joint angle regulation by the three latent variables and the behaviour of different grasp types for varying object width some general observations for static grasping postures can be derived.[6]

The proposed system for a posture-based human-mimicking robot is designed to replicate human body movements and postures using a combination of sensing, control, and machine learning techniques. The primary objective is for the robot to mirror human motions in real-time, adapting to dynamic environments and performing tasks with accuracy and fluidity.

The system starts with the robot's ability to perceive human posture using various sensors. Cameras and depth sensors capture images or video data of human movements. This data is processed using pose estimation algorithms, which identify key points on the human body, such as joints and limbs, and track their movement over time. To ensure real-time tracking, the system uses a combination of visual data from cameras and motion data from sensors like accelerometers and gyroscopes, which monitor the robot's own position and orientation. Once the human posture is recognized, machine learning models, especially deep learning algorithms, process the data

and extract relevant features, like joint angles and body positions. This information is then translated into commands for the robot's actuators, allowing it to replicate the human posture. The robot's joint movements are determined through inverse kinematics (IK), which ensures that the robot's limbs align with the human pose, while accounting for the robot's unique structure and joint constraints. Machine learning techniques, such as reinforcement learning, can optimize the process by allowing the robot to improve its posture replication over time, adapting to the nuances of human movement. Dynamic balancing is another critical aspect of the system, enabling the robot to maintain stability while moving, just like a human. As the robot moves, its center of mass shifts, and the system constantly adjusts its posture to ensure that it remains balanced. Algorithms like the Zero Moment Point (ZMP) method predict and adjust the robot's movement to prevent tipping over, allowing it to move smoothly and naturally. The system also incorporates feedback mechanisms experience and refine their movements, while techniques like learning allows the robot to optimize its performance by that allow the robot to adapt to changes in its environment or movements. By using proprioceptive sensors, such as force sensors in the feet or joints, the robot can adjust its posture based on real-time feedback, compensating for any imbalance or external forces that may affect its stability. This feedback loop ensures that the robot remains stable and responsive to its environment. Machine learning plays a significant role in the robot's ability to refine its movements. Through supervised learning, the robot can improve its ability to replicate human postures by analyzing a large dataset of human movements. Reinforcement this process leads to improved accuracy and adaptability. Real-time processing is vital for the success of the system. The robot must process data rapidly, calculate the necessary joint movements, and adjust its posture on the fly to ensure fluid, human-like movement. High-performance computing systems are often required to handle the intensive processing and ensure low latency feedback, allowing the robot to respond to human movements or environmental changes instantly. This system has a wide range of applications. In healthcare, robots could assist patients in rehabilitation by guiding them through specific movements or mimicking therapeutic exercises. In social robotics, the robot could replicate human gestures and body language to create more natural and engaging interactions with people.

In collaborative environments, such as industrial settings, the robot could work alongside humans, mimicking their movements for tasks like assembly or caregiving. The system could also be applied in entertainment, where it could be used for motion capture or animating virtual characters, or in sports, where robots could demonstrate movements or assist athletes in improving their performance.

Despite its potential, the proposed system faces challenges. Realtime processing must be optimized to ensure low latency, and the system must be robust enough to recognize postures in various conditions, such as low light or when parts of the body are obscured. Additionally, the robot's movements must be both accurate and safe, particularly in applications where it interacts with humans closely, such as healthcare or collaborative work.

Machine learning algorithms further refine the robot's ability to adapt and improve over time, making it more effective and natural in its movement. While the system great promise, challenges such as realtime processing, robustness to environmental variations, and ensuring safe human make the project highly valuable, paving the way for more advanced and capable robots in the future.

Ultimately, this project demonstrates the potential for creating robots that not only perform tasks but also interact with humans in a way that feels natural and intuitive, offering new possibilities for human-robot cooperation across various industries.

IV. CONCLUSION

In conclusion, the posture-based human-mimicking robot project presents a significant step forward in creating robots that can replicate human movements and adapt to dynamic environments. By integrating advanced sensing technologies, machine learning algorithms, dynamic balancing systems, and real-time control mechanisms, the proposed system can mirror human postures and movements with high accuracy and fluidity. This enables robots to perform a wide range of tasks, from interacting with humans in social or healthcare settings to assisting in industrial environments or even engaging in entertainment and sports applications. The use of vision-based sensors and motion tracking ensures that the robot can accurately perceive and interpret human gestures, while the application of inverse kinematics allows for precise replication of those

movements. Additionally, dynamic balancing and feedback systems ensure that the robot remains stable and responsive, even during complex or unexpected interactions.

REFERENCES

- [1] Hand Gesture-Based On-Line Programming of Industrial Robot Manipulators ,Antonios Syalris, Borja Ramis Ferrer and Jose L.Martinez Lastra FAST-Lab, Faculty of Engineering and Natural Sciences, Tampere University,Finland -2019.
- [2] Wirelessly Controlled Mimicing Humanoid Robot Assisting Children with ASD in learning Focused and Coordinated Behaviors,978-1-4799-7100-8/14/\$31.00 ©2014.
- [3] U. E. Ogenyi, J. Liu, C. Yang, Z. Ju and H. Liu, Physical Human–Robot Collaboration: Robotic Systems, Learning Methods, Collaborative Strategies, Sensors, and Actuators," in IEEE Transactions on Cybernetics - April 2021.
- [4] Gesture-Controlled Robotic Arm Utilizing OpenCV ,Conference Paper in European Journal of Science and Technology – June 2021.
- [5] Pose Estimation Techniques Using Openpose and HRNET, Master Of Technology In Computer Science, Department Of Computer Science And Engineering, JNTUH University College Of Engineering Science And Technology, Hyderabad, Telangana, India. - September 2024.
- [6] Synergy-Based, Data-Driven Generation of Object-Specific Grasps for Anthropomorphic Hands 2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids) Beijing, China, November 6-9, 2018.
- [7] Knee Joint Mechanism That Mimics Elastic Characteristics and Bending in Human Running , 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)Congress Center Hamburg Sept 28 - Oct 2, 2015. Hamburg, Germany.
- [8] B. Ramis Ferrer, W. M. Mohammed, A. Lobov, A. Moreno Galera, and J. L. M. Lastra, "Including Human Tasks as Semantic Resources in Manufacturing Ontology Models," in IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society, 2017.

- [9] D. Lu, Y. Yu, and H. Liu, "Gesture recognition using data glove: An extreme learning machine method," in 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2016.
- [10] M. Diaz, A. Fisher, M.A. Ferrer, R. Plamondon, "Dynamic signature verification system based on real signature", IEEE Transactions on Cybernetics – 2018.
- [11] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li and C.-Y. Su, "Interface Design of a Physical Human–Robot Interaction System for Human Impedance Adaptive Skill Transfer," in IEEE Transactions on Automation Science and Engineering, vol. 15, no. 1, pp. 329-340, Jan. 2018.
- [12] Kim, J.-W.; Choi, J.-Y.; Ha, E.-J.; Choi, J.-H. Human Pose Estimation Using Media Pipe Pose and Optimization Method based on Humanoid Model- 202