Humanoid Robotics: Recent Developments, Challenges, and Ethical Issues in Future Integration

Abhendra Pratap Singh[1*], Uma Gautam[2], Md.Ehsan Asgar [3], Rajeev Ranjan[4], Nandini Sharma[5], Pawan Shukla[6], and Akshit Verma[7]

[1,2,3,4]Department of Mechanical Engineering, HMRITM, Hamidpur, New Delhi,110036, India
[5,6,7] Department of Computer Science and Engineering, HMRITM, Hamidpur, New Delhi,110036, India
*abhendrapratap94@gmail.com

* Corresponding author
doi: https://doi.org/10.21467/proceedings.7.6.40

ABSTRACT

Advancing humanoid robotics is one of the grandest ventures of contemporary technological research, which aspires to create machines that mimic closely the human form, behaviour, and cognition. The article examines the current progress and challenges in the area with an emphasis on fusion of such fields as artificial intelligence, sensor systems, and mechanical designs to allow humanoid robots to perform complex tasks and interact with human environments. Whereas many real milestones have been achieved in mobility, manipulation, and social interaction, big hurdles remain to be overcome to realize the truly autonomous, adaptive robots that can respond with emotional intelligence and make real-time decisions. The paper further explores where humanoid robots present growing opportunities in other sectors such as healthcare, eldercare, customer support, and disaster response. Overall, the results indicate that while humanoid robotics is indeed on the verge of mass applications, the successful integration of the same into society needs not only perseverance in technological advancements but also mindfulness of the ethical, societal, and human-centred issues. In summary, humanoid robots promise to enhance human capabilities and enrich the quality of life, thus becoming real collaborators in the furtherance of a more integrated and intelligent future.

Keywords: autonomous robots, human-robot interaction, robotics ethics

1 Introduction

Humanoid robots have gained more attention in research circles. This growing interest stems from their wide range of potential uses. These robots could help older folks or people with physical challenges. They might also play a role in healthcare, public entertainment, personal care, and education. Search-rescue operations and manufacturing are other areas where they could prove useful. People have noticed that many of these robots copy human actions. They can walk, talk, and grab things just like a normal person [1].
The paper outlines a humanoid robotics system which enables understanding, performing and self-enhancing of tasks given in natural language. Incremental learning through human feedback enables the robots to adapt and improve their actions over time. The robot's language processing capabilities are intertwined with active learning techniques for an enhanced, self-sustained evolution of skill mastery. Natural language dialog plays a pivotal role in enabling intuitive human robot interaction, facilitating not only the expression of intent but also real time correction and guidance. To build more adaptive and resilient robotic systems, it is crucial that robots incrementally learn from such interactions improving their behaviour over time and reducing repeated errors. The team set up the system on the ARMAR-6 humanoid robot and tested it in simulations and actual experiments. The results showed that bit by bit learning boosts how well it does tasks and how well it can adapt [2].

Proceedings DOI: 10.21467/proceedings.7.6; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-989164-9-5

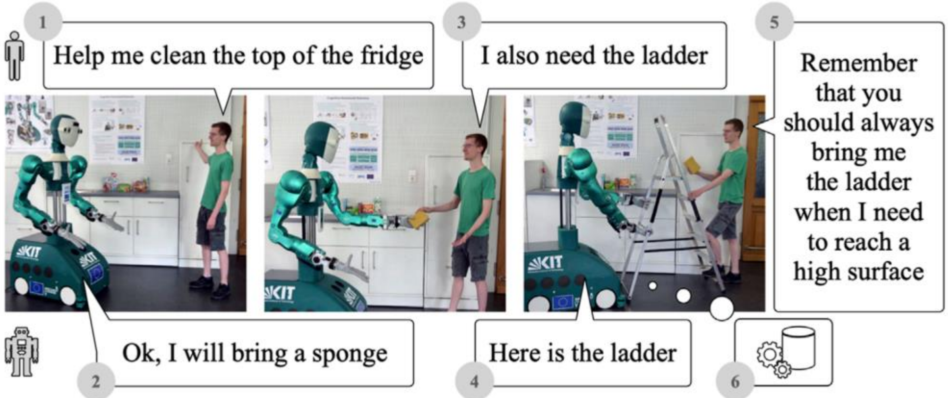

Fig.1. ARMAR-6 incrementally learns behaviour from natural interactions [2]

Advanced robotics will enable automation to transform the relationship between humans and robots in design, empathy, collaboration, and congruity for future advanced manufacturing, incorporating integration for human-robot knowledge transfer. To make it easier for humans and robots to share knowledge, there is a requirement of a single system to represent what is known. This helps them understand tasks in the same way. Assembling top-notch products in industrial settings needs a versatile and thorough picture of the assembly steps. This approach helps humans and robots work together making sure the job gets done right and meets quality standards [3].

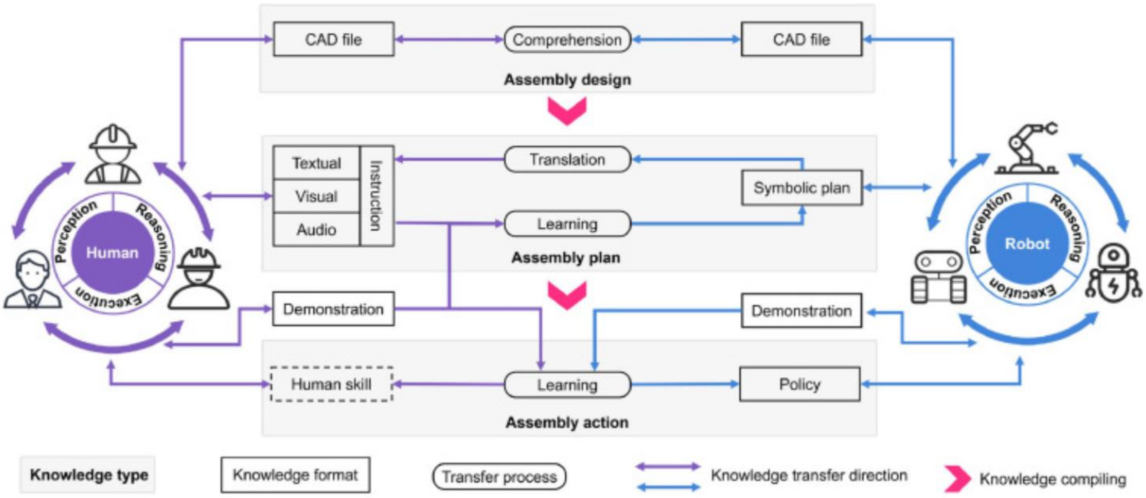

Fig.2. Human-robot assembly knowledge transfer framework [3]

New breakthroughs in big language models (LLMs) have shown exciting uses in robotics to understand context, reason, and plan. LLMs let robots follow spoken commands, but most systems still depend on voice-only interaction with few multi-modal features. Work on human-robot interaction (HRI) remains scattered and lacks a clear approach. This research introduces a new LLM-based robot system that can interpret different human inputs like where people look where they stand, and how they act. It allows robots to respond through coordinated speech and physical actions. Early findings suggest this could change HRI into a more natural and adaptable way for humans and robots to work together [4].

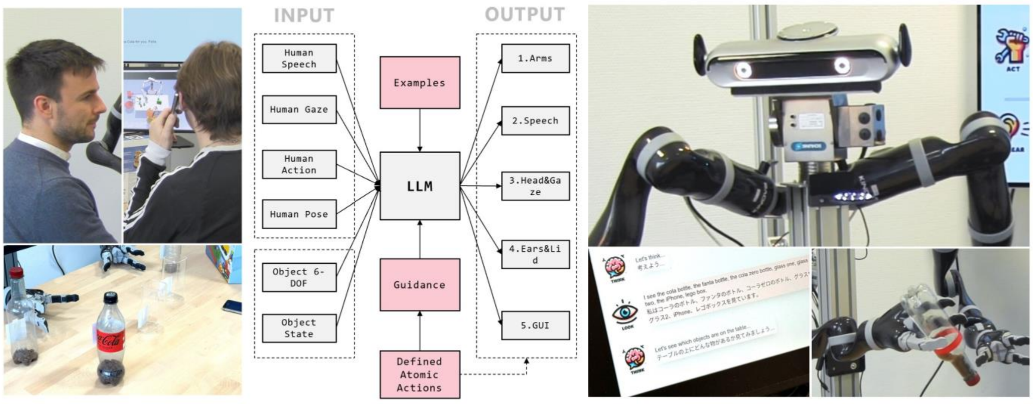

Fig.3. LLM driven human-robot interaction centred around guidance, capabilities, and examples [4]

HOVER, Humanoid Versatile Controller, blends different control styles into one policy. This framework allows for easy changes between styles keeping the intended about each one. It makes controlling humanoid robots better and stronger without having to relearn everything for each style. Traditional ways are out due to its inability to perform. So, HOVER saves time and makes these robots way more flexible for the next requirement. While current ways to control human-like robots are impressive, they stick to formulas made for specific jobs like moving fast for walking or following precise joint moves for more complex motion. This way of doing things equivalents to making new controller types and rewards from scratch each time, which eats up time and gets repetitive. Also, it cramps the robot's style, because they can't hop from one job to another. A robot that's built for walking on two legs, for example, might just fail at a job where it needs to grab and move things. These hurdles show the reason to have one control setup that can handle all sorts of things well and with ease [5].

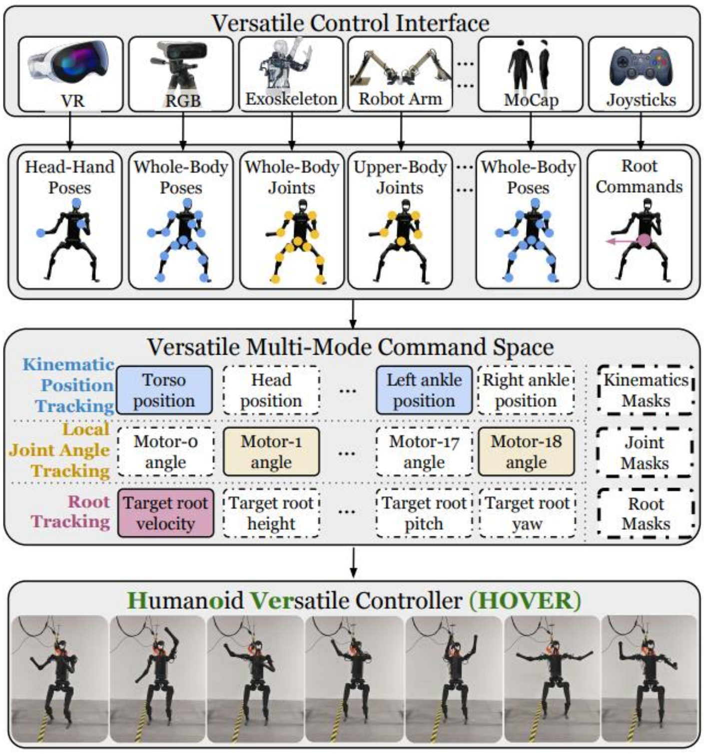

Fig.4. HOVER enables versatile humanoid control with a unified multi-mode command space. The versatile multi-mode command space supports kinematic position tracking (blue), local joint angle tracking (yellow), and root tracking (purple). Highlighted boxes indicate active commands being tracked, while the masks (dashed boxes on the right) allow selective activation of different command spaces to accommodate various tasks [5]

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

Robotics has stepped up its game in the last few decades making robots that can interact with humans in all kinds of places. For instance, NAO is a human-like robot made by the folks at Aldebaran Robotics, but now it's under the wing of Softbank Robotics. Experts have put the NAO bot to use in various areas like schooling, health services, recreation, and studies. Its success depends on folks and robots getting along well. That throws a spotlight on Human-Robot Interaction (HRI) where smart people from different backgrounds work to level up the chat, help, and team up between humans and bots. Major changes in how bots see, think, find their way, and talk have cranked up the chances for some teamwork with humans. The whole point of looking into HRI is to get people and robots to perform tasks together. The main topics they are examining include how individuals interact, learn on the spot, and collaborate as a team. Teams do all sorts of tests which includes real-life user stuff, and pretend runs to figure out HRI. As robot tech keeps getting advancements, making HRI better is important to brew up even smarter bot mates like NAO. The NAO robot stands out as a super expressive human like bot that can recognize voices, show emotions through face and gestures, and can be programmed too, which is the reason it is handy in lots of different jobs. But even with these functions helping it out, it's still tricky when people and robots have must work together. As it looks kind of like a person, folks might expect too much from it, and sometimes it doesn't catch what people are throwing at it because its sensors and how it understands the world aren't top-notch.

Fig.5. Nao Robot [6]

Self-driving robots working together with people make cooperating on changing, custom jobs better. Smarter AI tech like 3D building and big systems for dealing with LLMs make it easier for robots to know parts and get what instructions mean. These LLMs improve robots' grasp of what one tells them, so they can do putting-together work well using their seeing and driving skills [7]. People get hyped about hanging out with social robots because it's a fresh experience, but there is a bit of worry about how long they will stay interested. When they first meet a robot, the newness grabs their attention. Social robots shine when it comes to drawing out significant social behaviours and feelings from people, whether in experiments or out in the real world [8]. The struggle for keeping everyone fed around the globe is complicated. In addition to produce food, farmers also gain from it. The changing climate, lack of enough labour and the sky-high prices of necessaries like fertilizers and bug killers, however could lead to a real tough situation. But things are starting to look up due to some technological upgrades. Farms are getting smarter with devices like sensors that stay put and gear like robots and self- driving machines. These tools are real game changers helping to produce more food, cut down on need of too much labour, and sustain the whole food industry for ages [9].

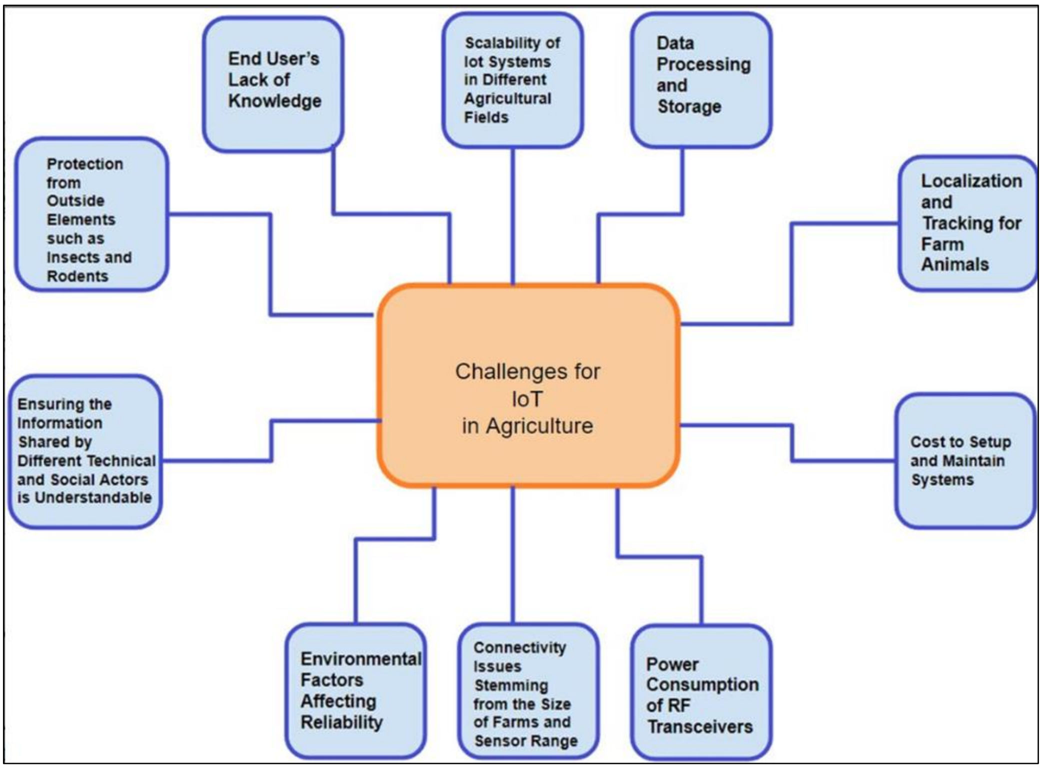

Fig.6. Challenges for IoT in agricultural applications [9]

Teaching humanoid robots to move like humans is still tough. Even though humans can move in substantial ways making robots do the same is tough. Creating control over stuff that can handle smooth changes and mimic real human moves is challenging. To get humanoid robots acting more like real people when they are out in the real world, it is required to close this division. To get smooth movements without making tough reward systems, the "adversarial motion prior" (AMP) helps a lot. In contrast, Deep RL is all the rage these days for getting robots to move in ways, because it plays around with rewards to make the modelling bit automatic. It's done some nifty work with four-legged, two-legged, and human-like robots. But when there are humanoids with loads of moving parts, the stuff RL comes up with can look kind of weird assuming an odd arm that flails and rough landings. So, makers often use things like footstep planners when they want to smooth that out. Robots that look like humans have a tough time copying what people show them because their joints, the way their bodies are set up, and their bone layouts are different from that of humans [10].

2 Literature Review

Jha et al. [11] was keen to crack the code on tracking and identifying the various moving parts of a human body. They set their sights on pinpointing the location of a robotic arm made to mimic a human one, by employing technology that makes educated guesses about body posture. They gathered 3D details of where a person's skeletal joints are using a 'Kinect' depth sensor along with a tool named Media Pipe. Then, they calculated the joint angles in a backward direction, a key for tackling tricky issues. With these angles in hand, they managed to get a robot's neck, shoulder, and elbow to move together. They did this by zipping signals back and forth from a Python program to the machinery via serial communication. When they compared their findings, they noticed that the Media Pipe number work was sort of more accurate, with fewer mistakes than what the 'Kinect' device delivered. An important aspect of joint work between humans and robots is the study of human pose and motion; this knowledge will, in turn, assist in controlling humanoid robots. Tracking human movement (including joint movement) under changing environments is a problem since many factors are influencing the process simultaneously. For the last few decades, many specialists were using depth sensors, and one of the most popular ones for 3D data capture was the Kinect. But nowadays, a new one called Media Pipe is being examined by some very intelligent people since it apparently seems to be more accurate and computationally cheap. Inverse kinematics allows humanoid robots to mimic human motion. It finds the joint angles required to place limbs in certain positions. People have connected this to real-time applications using Python-Arduino communication to allow the robots to respond and interact quickly and smoothly.

Kim et al. [12] emphasized the crucial need of humans to understand the specific capabilities of the robots rather than observing its general activity. Many people narrow their thinking about features of autonomy into how many orders they can execute without them. New research proposes that one must display a better and, in fact, broader outlook on the matter and talks about six different varieties of robot freedom depending on their collaboration with human input before and during performing activities. There is doing-stuff autonomy, planning-stuff autonomy, teamwork autonomy, unpredictable autonomy, thinking-stuff autonomy, and moving-stuff autonomy. Besides this, as these robots become much more intelligent and much more independent, the requirement to understand while addressing the complexities which would come forward with autonomous decisions is significantly expressed. Cheng et al. [13] presented Expressive Whole-Body Control (Ex Body), a method for controlling humanoid robots so that they can reproduce a full repertoire of movements close to those we perform even in the real world. This is accomplished by teaching the entire robot body movement types from dance and other activities observed in real humans and applying reinforcement learning techniques. To ensure that robots could approach capabilities of humans in movement, Ex Body focused more on imitating what the upper body does, while opening some flexibility for the legs to act according to speed commands instead of direct ones. Once learned in the world inside a computer, the skill is expected to materialize in the real world as actions like walking or greeting in new ways or even dancing. It is on the verge of completion, and it passed several tests that it underwent.

Zhang et al. [14] stated that robots that have human shapes are great at performing certain specialized jobs because of RL or Reinforcement Learning. But figuring out the right rewards and teaching them isn't easy. So, the Adam humanoid robot came along, and its patterned build makes it an expert at imitating what it sees. They also implemented an entirely new imitation learning framework based on these adversarial motion priors. This is what enables Adam to walk and move just like a human being. When they did the evaluation, Adam was pulling all kinds of walking moves just like the humans. And that's a big factor for watching real people move robots learning.

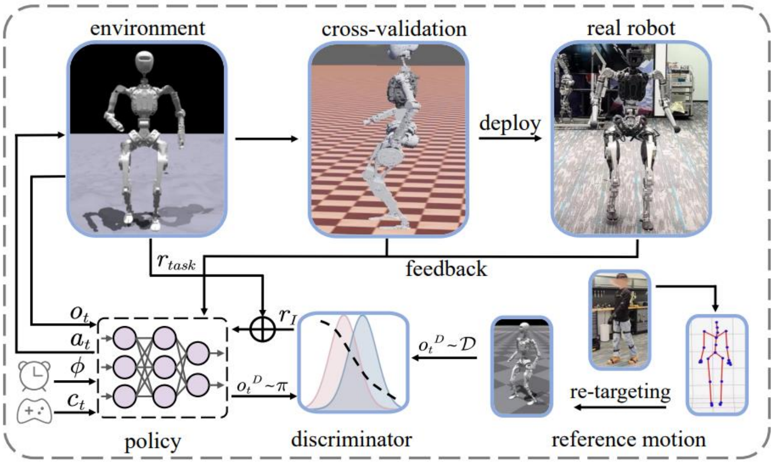

Fig.7. Adversarial Motion Priors Imitation Training Framework of Humanoid Robot [14]

Kim et al. [15] predicts the large language models (LLMs) have so much power to turn conversation with the robots much interesting by having robots converse in smart-sounding ways. But, the requirement is not known to orchestrate the LLMs into robots properly. There was an event wherein 32 people interacted

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

with LLM-embodied robots while also trying out the regular chatty devices they typed to or talked at. They found that the robot companions with LLMs were much more fun to hang out with and felt much more like friends due to their cute behaviours in between talking; however, when it came to reflecting conversation, they underperformed. This whole venture shows the intelligent points that one must consider while incorporating LLMs into robot companions without glitching. Abu-Jassar et al. [16] study implements a computer vision setup for a small mobile humanoid bot. It emphasizes making both hardware and software for controlling motors and vision more independent. The hardware design uses a modified tiny-YOLO NN to detect objects and make decisions in real time. Hardware components communicate according to a well-defined plan concerning mechanism for interconnection along with intelligent algorithms assisting them. Replacing the VGG-16 part of the SSD model with the lightweight tiny-YOLO enables better object detection and faster decision making in the robot, and it still performs well with a relatively low number of objects embedded into it. Testing the system with moving and static objects indicates a high level of even enhanced detection capabilities and gives solid evidence for this approach's success.

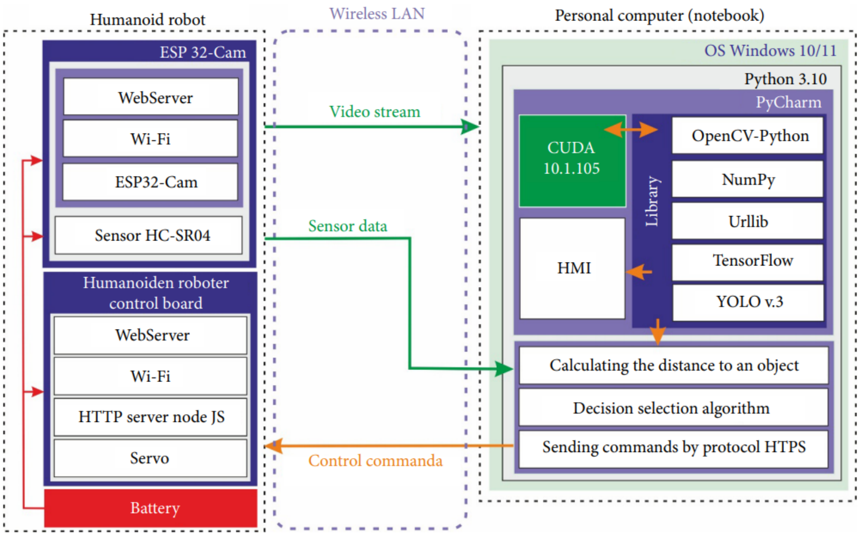

Fig.8. Structure of information interaction camera module and control module with system of management and decision making to control a humanoid robot [16]

Li et al. [17] studied OKAMI in this paper, a novel technique in teaching humanoid bots to handle an object by utilizing only an RGB-D video demonstration of the action. OKAMI incorporates something called object-aware retargeting, allowing for bots to adjust human-type moves for various locations of items when active in tasks. It also calls upon open-world vision systems to spot the crucial objects for the job and tackles how the actual body and hand moves act independently. Notably, OKAMI really works when deployed into a plethora of different environments and beats IML methods. It even boasts this unique capability-teaching robots on reflexive visuomotor policies and achieving a success rate of 79.2%, without any external control.

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

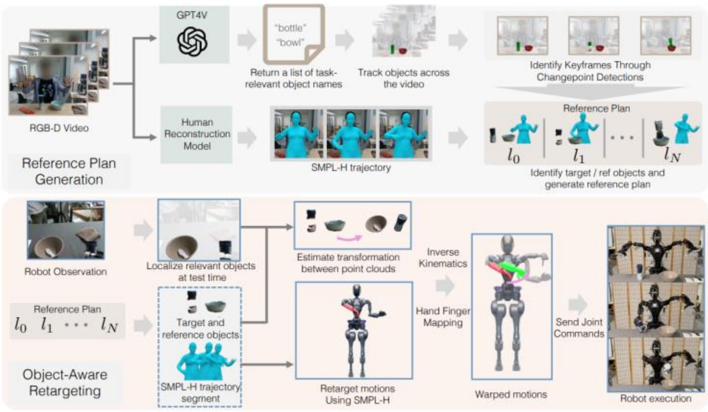

Fig.9. Overview of OKAMI. OKAMI is a two-staged method that enables a humanoid robot to imitate a manipulation task from a single human video. In the first stage, OKAMI generates a reference plan using GPT- 4V and large vision models for subsequent manipulation. In the second stage, OKAMI follows the reference plan, where it retargets human motions onto the humanoid with object awareness. The retargeted motions are converted into a sequence of robot joint commands for the robot to follow [17]

Zhu et al. [18] on R4SJ team, initiated a new framework that utilizes Engineering for Social Justice (E4SJ) ideas to implement equity in Human-Robot Interaction (HRI). They reviewed ten years of HRI studies to measure their impact for the betterment of the society and found some limitations. So, five impactful suggestions were proposed to analyse for bright researchers, builders, and designers. This initiative is inciting the experts and policymakers in the HRI to emphasize equity and develop socially beneficial technologies.

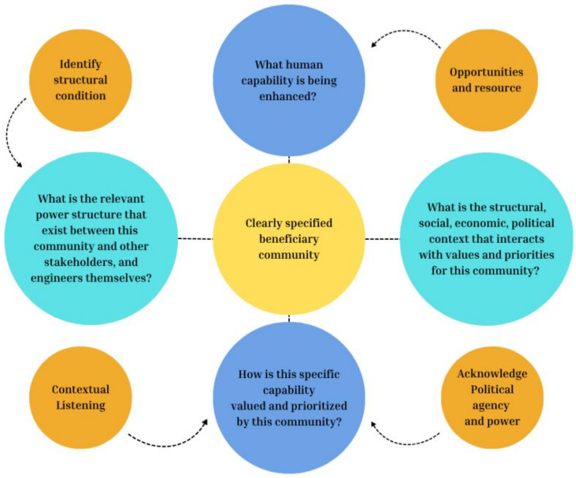

Fig.10. 5 R4SJ recommendations and E4SJ criteria [18]

Hou et al. [19] suggested the issues associated with power dynamics within Human-Robot Interaction (HRI) often receive the short end of the stick, but are nevertheless very important. The scholarly review of HRI literature using big notions of power by social scientists like Fiske and Berdahl, and French and Raven's notions about the bases of power, revealed underlying power structures regarding the authority holders, decision-makers, and those in dominant positions. Hou et al. strongly recommended placing the concept of power at the centre of HRI as it will draw together all these fragmented bits of speculation and lay one strong blueprint for the future researchers through its exposition. Gu et al. [20] suggest that humanoid-Gym, a reinforcement learning framework, is based on Nvidia Isaac Gym for training humanoid robots in locomotion skills allowing zero shot sim to real transfer. The framework provides a sim to sim validation of the policies in different physics simulators (Isaac Gym and Mujoco), providing a robustness check. The experiments conducted on the XBot-S and XBot-L humanoid robots prove that the system can be effectively deployed in the real world with zero additional tuning. Hirose et al. [21] incited from the walking patterns of bipedal creatures, have been primarily engaged in robotic research continuously at Honda since 1986. This work has brought ASIMO, a real advanced robot capable of walking, pivots, and can interact well with humans. ASIMO represents a massive leap in terms of robot design and control. Its five-fingered hands are adept to perform complex tasks, and it walks like a person. Honda didn't make this bot for recreation; they want it to assist people in their everyday lives. With its characteristic tenacity, Honda is continuously pushing the boundaries of innovation to ensure that robots like ASIMO become an integral part of daily life.

Choudhury et al. [22] studied that humanoid robots are gaining importance in health and education. They can act as an aid to improve social skills and calm individuals with autism, cancer, and brain disorders. Aspects like age, gender, brain health, and the nature of how humans relate with each other are crucial in equipping humanoid robots. It seems that particularly socially deprived human beings among children and elderly find more appeal in them. Humanoid robots are consistent, so they work well as an intervention for teaching social skills to children with autism. The trust of the people for their robotic companions is heavily dependent on the purpose for which these robots are used and other parameters as to get along with them, like their background history and social circles. Adams et al. [23] shows Karel Capek's play R.U.R. of 1923, all the way to these current attempts to make robots look and act human. They have come a long way from doing only simple or risky jobs to working intricate tasks in places such as factories, nuclear plants, and even outer space. But making one of these bots like a human, smart and adaptable, is still complex for engineers. Integrating together the mechanical components and wires with the computational equipment, and sensors capable of understanding tasks properly, adapting to new environments, as well as learning new skills, proves to be a huge hurdle in front of engineers. All labs worldwide are actively surveying about the behaviour of these robots, how they understand the world, and the learning process they adopt. This review mentions that developing a humanoid robot does not belong to one individual but collective effort requiring brains and people from engineering, computer design, and instant control systems. Pollard et al. [24] analysed the physical motion of the human upper body in the study to modify it for anthropomorphic humanoid robot mechanisms, which have less flexibility in joint movement and slower speed than humans do. At this point, the focus is on getting the Sarcos humanoid robot to move exactly like humans move and are using both the eyes and numerical analysis to assess deviations between the robot and real human performance. There is need to perfect the ability of robots to carry out activities which humans do, although today's robots are still a long way from achieving this goal. Yokoyama et al. [25] demonstrates the development of a humanoid robot, HRP-2P, designed for cooperation with humans. This robot has a biped walking mechanism, perceives the world using stereo vision, and listens through an audio-based human interface to identify payloads for lifting and carrying along with a human partner. The robot is equipped with arms that operate under a cooperative control framework, whereas the arm decides where to move based on push and twist feedback from sensors in its wrist. These experiments demonstrate that this cooperative approach works quite well in real-world environments.

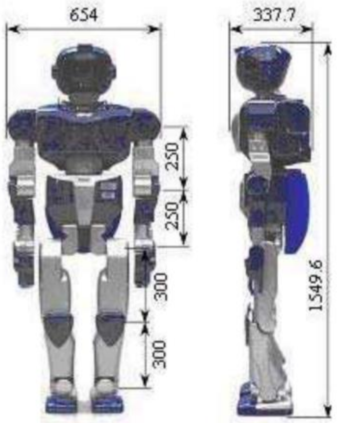

Fig.11. Humanoid Robot HRP-2P [25]

DiSalvo et al. [26] shows the investigation regarding humanoid robots of how their facial characteristics affect human perception of them conveying human behaviour. 48 robots were studied with people asked to evaluate. Studies show that the size of the head and the number of bits of the face that are seen affect whether one sees it as human. The study, therefore, lays down some basic guidelines and rules for making robot faces in the future that would invite people to settle in and regard them as more human-like.

Kaneko et al. [27] showcases the in-depth study investigating the development of the humanoid robot HRP-3 and the next generation of HRP-2. The new-generation model has better mechanical components and body construction designed to keep dust and water splashes out, making it great for outside the lab use. Other new features of HRP-3 include a new wrist and hand for better gripping and lifting. The AI software is improved to support its actual tasks. The study describes HRP-3's inner workings, investigates the enticing technology, and demonstrates some possibly unique tests proving its competence and versatility in various applications.

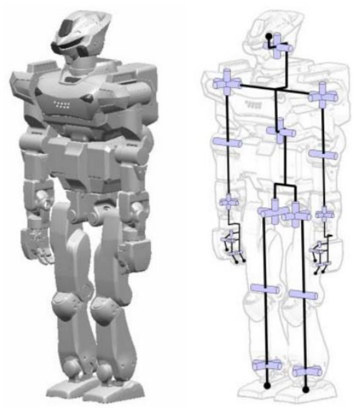

Fig.12. Configuration of HRP-3 [27]

Dimensions	Height	1,606 [mm]
	Width	693 [mm]
	Depth	410 [mm]
Weight inc. batteries		68 [kg]
D.O.F.		Total 42 D.O.F.
	Head	2 D.O.F.
	Arm	2 Arms × 7 D.O.F.
	Hand	2 Hands × 6 D.O.F.
	Waist	2 D.O.F.
	Leg	2 Legs × 6 D.O.F.
Operational Time		120 [min]
Feature		Dust/Splash-proof

Fig.13. Principal Specifications of HRP-3 [27]

Kanda et al. [28] shows that the research indicates the influence of a humanoid robot on how individuals behave when engaged with it socially. Chatting with 48 subjects were two differently looking robots that behaved and spoke in the same way. Apparently, the appearance of the robot did not influence how subjects verbally engaged with it. Yet aspects of their nonverbal response, such as physical proximity to the robot or speed of verbal reply, were heavily influenced. It was assumed that this was due to the atmosphere or the perceived qualities that were just implied by looking at them. This shows, again, how important robot design is for the interaction of human beings and robots.

Breazeal et al. [29] showcases that the research sheds light on the psycho-emotional and expressive actions that elevate social exchange between humans and pseudo-human robots. The team has worked out the science behind the ability of the bot to exhibit feeling and expressiveness. The insights were mixed from the kids learning of a language and emotional senses through patterns of talking. The research studies both parts of the individual expressiveness and the whole nature of the bot while interacting one-to-one with human beings. It was found that wherein the robot showed emotions, it could engage better with humans. This stresses the importance of feelings in human robotics interaction. Brooks et al. [30] suggests crafting the humanoid robot Cog spotlights the importance of having a body and senses that work together when you want to make robots that can socialize with people. Research papers shows that if humanoid robots are going to act like humans, they should have smooth moves, good sensing, and the skill to get and react to the emotions people exhibit around them. Tests on Cog show the progress in tasks like copying perceptions, showing feelings, and responding to individuals appropriately. But it is not done yet. There is need to incorporate these systems together better to develop a robot companion that can work autonomously and learn from the growth and knowledge of its human counterparts. Shiomi et al. [31] discusses effectiveness of interactive humanoid robots in informal learning places like science museum to improve educational experience. The incorporation of RFID technology helped the robots in personalizing their interaction with and guiding the visitors through different exhibits. The trial proved that mixing exhibit guiding with playful and child-like interactions went a long way towards improving the visitor's study-engagement and satisfaction. Such research findings endorse the prospects of humanoid robots nurturing interest in science through dynamic, context-oriented interactions and underline their utility in public educational markets.

3 Conclusions

Advanced research is being carried out that identifies the important forces the humanoids must navigate within structured surroundings for human environments. The possibilities for humanoids are thus seen as transformational not as mechanical contraptions, but rather as soft agents with capabilities for adaptation through interaction by learning and contextual reasoning to assist in agile tasks in complex real-life environments. Although current systems can already perform very impressively on levels such as mobility, perception, or social interaction, the road into attaining completely self-aware humanoids that respond emotionally toward the human is a significant scientific challenge. Bridging this would require a lot of varied means: technological advances, a good understanding of human behaviour, ethical approaches to real-life problems, and meeting human needs. The future of humanoid robotics does not lie only in human mimics but rather in the re-envisioning of enhanced human interactions, assistance, and coexistence with intelligent and affectively receptive machines. While this research will continue to usher fresh avenues of exploration, humanoids will then find themselves moving away from being experimental prototypes, toward being serious contributors to a more humanizing, inclusive, and intelligent future.

References

[1] J. Qi, L. Ma, Z. Cui, and Y. Yu, “Computer vision-based hand gesture recognition for human-robot interaction: A review,” *Complex & Intelligent Systems*, vol. 10, no. 1, pp. 1581–1606, 2024, doi: 10.1007/s40747-023-01173-6.

[2] L. Bärmann, R. Kartmann, F. Peller-Konrad, J. Niehues, A. Waibel, and T. Asfour, “Incremental learning of humanoid robot behavior from natural interaction and large language models,” *Frontiers in Robotics and AI*, vol. 11, p. 1455375, 2024, doi: 10.3389/frobt.2024.1455375.

[3] R. K.-J. Lee, H. Zheng, and Y. Lu, “Human-robot shared assembly taxonomy: A step toward seamless human-robot knowledge transfer,” *Robotics and Computer-Integrated Manufacturing*, vol. 86, p. 102686, 2024, doi: 10.1016/j.rcim.2023.102686.

[4] C. Wang, S. Hasler, D. Tanneberg, F. Ocker, F. Joublin, A. Ceravola, J. Deigmoeller, and M. Gienger, “LaMI: Large language models for multi-modal human-robot interaction,” in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–10, 2024, doi: 10.1145/3613905.3651029.

[5] T. He, W. Xiao, T. Lin, Z. Luo, Z. Xu, Z. Jiang, J. Kautz, et al., “Hover: Versatile neural whole-body controller for humanoid robots,” *arXiv preprint*, arXiv:2410.21229, 2024.

[6] E. Nagy, É. Karl, and G. Molnár, "Exploring the Role of Human-Robot Interactions, within the Context of the Effectiveness of a NAO Robot," *Acta Polytechnica Hungarica*, vol. 21, no. 3, pp. 177–190, 2024, doi: 10.12700/APH.21.3.2024.3.12.

[7] S. Liu, J. Zhang, L. Wang, and R. X. Gao, "Vision AI-based human-robot collaborative assembly driven by autonomous robots," *CIRP Annals*, vol. 73, no. 1, pp. 13–16, 2024, doi: 10.1016/j.cirp.2024.03.004.

[8] G. Laban, A. Kappas, V. Morrison, and E. S. Cross, "Building long-term human–robot relationships: Examining disclosure, perception and well-being across time," *International Journal of Social Robotics*, vol. 16, no. 5, pp. 1–27, 2024, doi: 10.1007/s12369-023-01076-z.

[9] M. O. Yerebakan and B. Hu, "Human–robot collaboration in modern agriculture: A review of the current research landscape," *Advanced Intelligent Systems*, vol. 6, no. 7, p. 2300823, 2024, doi: 10.1002/aisy.202300823.

[10] A. Tang, T. Hiraoka, N. Hiraoka, F. Shi, K. Kawaharazuka, K. Kojima, K. Okada, and M. Inaba, "Humanmimic: Learning natural locomotion and transitions for humanoid robot via Wasserstein adversarial imitation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 13107–13114, 2024, doi: 10.1109/ICRA57147.2024.10610449.

[11] P. Jha, G. P. K. Yadav, D. Bandhu, N. Hemalatha, R. K. Mandava, M. Ş. Adin, K. K. Saxena, and M. Patel, "Human–machine interaction and implementation on the upper extremities of a humanoid robot," *Discover Applied Sciences*, vol. 6, no. 4, art. 152, 2024, doi: 10.1007/s42452-024-05734-3.

[12] S. Kim, J. R. Anthis, and S. Sebo, "A taxonomy of robot autonomy for human-robot interaction," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 381–393, 2024, doi: 10.1145/3610977.3634993.

[13] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, "Expressive whole-body control for humanoid robots," *arXiv preprint*, arXiv:2402.16796, 2024.

[14] Q. Zhang, P. Cui, D. Yan, J. Sun, Y. Duan, G. Han, W. Zhao, et al., "Whole-body humanoid robot locomotion with human reference," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11225–11231, 2024, doi: 10.1109/IROS58592.2024.10801451.

[15] C. Y. Kim, C. P. Lee, and B. Mutlu, "Understanding large-language model (LLM)-powered human-robot interaction," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 371–380, 2024, doi: 10.1145/3610977.3634966.

[16] A. T. Abu-Jassar, H. Attar, A. Amer, V. Lyashenko, V. Yevsieiev, and A. Solyman, "Development and investigation of vision system for a small-sized mobile humanoid robot in a smart environment," *International Journal of Crowd Science*, vol. 9, no. 1, pp. 29–43, Jan. 2025, doi: 10.26599/IJCS.2023.9100018.

[17] J. Li, Y. Zhu, Y. Xie, Z. Jiang, M. Seo, G. Pavlakos, and Y. Zhu, "OKAMI: Teaching humanoid robots manipulation skills through single video imitation," presented at the 8th Annu. Conf. on Robot Learning (CoRL), 2024, arXiv preprint arXiv:2410.11792, doi: 10.48550/arXiv.2410.11792.

[18] Y. Zhu, R. Wen, and T. Williams, "Robots for Social Justice (R4SJ): Toward a more equitable practice of human-robot interaction," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 850–859, 2024, doi: 10.1145/3610977.3634944

[19] Y. T.-Y. Hou, E. Cheon, and M. F. Jung, "Power in human-robot interaction," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 269–282, Mar. 2024, doi: 10.1145/3610977.3634949.

[20] X. Gu, Y.-J. Wang, and J. Chen, "Humanoid-gym: Reinforcement learning for humanoid robot with zero-shot sim-to-real transfer," *arXiv preprint*, arXiv:2404.05695, 2024.

[21] M. Hirose and K. Ogawa, "Honda humanoid robots development," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1850, pp. 11–19, Nov. 2006, doi: 10.1098/rsta.2006.1917.

[22] A. Choudhury, H. Li, C. M. Greene, and S. Perumalla, "Humanoid robot-application and influence," *Archives of Clinical and Biomedical Research*, vol. 2, no. 6, pp. 198–227, Dec. 2018, doi: 10.26502/acbr.50170059

[23] B. Adams, C. Breazeal, R. A. Brooks, and B. Scassellati, "Humanoid robots: A new kind of tool," *IEEE Intelligent Systems and Their Applications*, vol. 15, no. 4, pp. 25–31, Jul./Aug. 2000, doi: 10.1109/5254.867912

[24] N. S. Pollard, J. K. Hodgins, M. J. Riley, and C. G. Atkeson, "Adapting human motion for the control of a humanoid robot," in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, pp. 1390–1397, May 2002, doi: 10.1109/ROBOT.2002.1014737.

[25] K. Yokoyama, H. Handa, T. Isozumi, Y. Fukase, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, and H. Hirukawa, "Cooperative works by a human and a humanoid robot," in *2003 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, pp. 2985–2991, 2003, doi: 10.1109/ROBOT.2003.1242049

[26] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal: The design and perception of humanoid robot heads," in *Proc. 4th Conf. Designing Interactive Systems (DIS)*, pp. 321–326, 2002, doi: 10.1145/778712.778756

[27] K. Kaneko, K. Harada, F. Kanehiro, G. Miyamori, and K. Akachi, "Humanoid robot HRP-3," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2471–2478, 2008

[28] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, and H. Ishiguro, "Analysis of humanoid appearances in human-robot interaction," *IEEE Transactions on Robotics*, vol. 24, no. 3, pp. 725–735, June 2008, doi: 10.1109/TRO.2008.924186.

[29] C. Breazeal, "Emotion and sociable humanoid robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1–2, pp. 119–155, July 2003, doi: 10.1016/S1071-5819(03)00018-1.

[30] R. A. Brooks, C. Breazeal, M. Marjanović, B. Scassellati, and M. M. Williamson, "The Cog project: Building a humanoid robot," in *Int. Workshop on Computation for Metaphors, Analogy, and Agents*, pp. 52–87, Berlin, Heidelberg: Springer, 1998

[31] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Interactive humanoid robots for a science museum," in *Proc. 1st ACM SIGCHI/SIGART Conf. on Human-Robot Interaction*, pp. 305–312, 2006, doi: 10.1145/1121241.1121309.