Zachary Vella, IT
Institute of Information & Communication Technology
Malta College or Arts, Science & Technology
Corradino Hill
Paola PLA 9032
{zachary.vella}@mcast.edu.mt

***Index Terms*—MCAST, IICT, LaTeX, Project, Paper**

I. Introduction

With the advancements that have been made in the field of vision-based control and real-time motion capture technology, there will be new possibilities for humans to interact with computers. The focus of this research is on the utilisation of a camera-based system that tracks the movements of the person's arm and is then replicated in real-time through a virtual arm. Such a method of communication with computers holds significant potential in several practical applications in fields regarding medical, rescue operations, and the handling of hazardous material, which all require precise and responsive robotic control as a necessity. By integrating computer vision techniques, limb tracking and learning algorithms, this study aims to enhance both accuracy and adaptability for different robotic and limb movement situations. The ability to map human motion accurately onto either a digital or robotic counterpart is a crucial aspect of this research. Another crucial addition for the future aspect of this research is the integration of haptic feedback, specifically through the use of gloves that will allow the user to sense when the virtual or robot arm is making contact with an object. The implementation of a feedback mechanism can have a significant improvement on the user's spatial awareness and overall control in comparison with traditional controllers, which can lead to a more intuitive and natural, almost symbiotic, relationship with robotic systems. To conduct this study, an interpretive philosophical approach will be taken, employing both deductive reasoning and experimental stages. A mixed methods approach will be adopted, combining both qualitative and quantitative data, which will be collected over a longitudinal time frame. A questionnaire will be used to gather feedback on the user experience, performance improvements and possible modifications for enhancing the system. The data collected through these questionnaires will provide valuable insight into the possible uses and effectiveness of vision-based control systems and where their application can have a significant improvement. The primary goal of this research is to evaluate the effectiveness of vision-based control systems to replicate real-time human arm movement with enhanced accuracy and greater responsiveness. With this leverage, computer vision and machine learning will be able to explore improvements in motion tracking, a greater reduction in latency, and overall user interaction. The findings of this research could have significant contributions to multiple fields, including assistive technology, automation, and virtual reality. These again demonstrate the practical benefits of vision-based motion control systems over traditional controller-based systems.

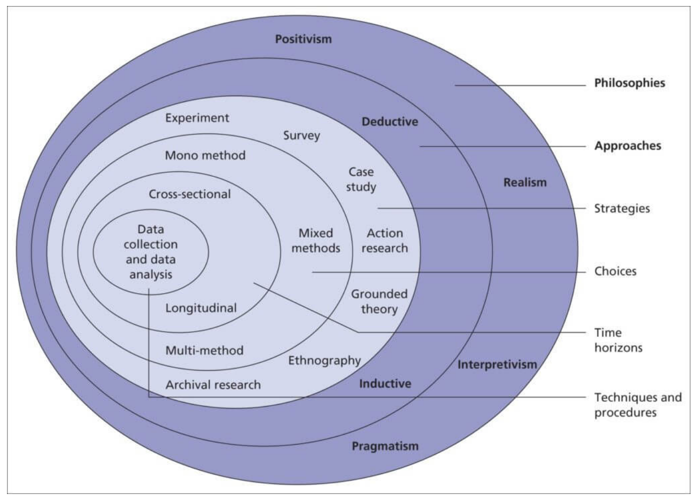

Fig. 1. Research Onion

II. Review of Research Methodology

A. Introduction

In our modern society, people have become accustomed to simple human-computer interaction (HCI), primarily involving input devices such as joysticks, mice and keyboards. These are commonly used in robotics and automation. However, as technology evolves, these user interfaces (UI) need to advance in parallel. As a result, human-robot interactions (HRI), using hands as input devices. This offers more intuitive control, improves safety, accuracy and responsiveness in complex tasks. This review will explore papers containing research and applications on how the implementation of gesture-based systems can be applied across various industries.

B. Literature Review of Methodologies

These papers have been studied to investigate the diverse methodologies used to investigate the gesture-based control systems:

- Sobhani et al. (2022) implemented a mixed-reality robot teleoperation system using VR headsets to control a robotic arm that has three joints (shoulder, elbow, and wrist). The study was measured using the System Usability Scale (SUS) scoring, head tracking and task performance metric using human participants in assessing performance in remote environments that present high-risk conditions as described in the decommissioning of a nuclear facility [1].

- Bai et al. (2023) created an automatic control system for a 7-DOF robotic arm, using an RGB camera and object pose estimation powered by neural networks. Their testing involved real-time grasping tasks and robot manipulation in a controlled environment [2].
- Nguyen et al. (2022) proposed a real-time skeleton-based gesture recognition model utilising TD-Net architecture alongside MediaPipe for hand pose estimation. They evaluated their system using the IPN dataset, measuring recognition accuracy and inference speed for continuous hand gesture control [3].
- Gourob et al. (2021) developed a low-cost robotic hand system controlled by vision-based gesture recognition. Built using an Arduino UNO and five servo motors, the system recognised gestures through a live feed camera and was tested for assistive technology applications [4].
- Yu et al. (2017) implemented a gesture-controlled UAV system using ROS. Hand gestures were recognised through RGB input and translated into UAV flight commands. Drone-operated simulations were conducted to evaluate results [5].

C. Distinguishing Academic and Non-Academic Sources

All cited papers referenced are academic, peer-reviewed, and this is because they are published by IEEE. Peer-reviewed sources ensure rigour, replicability and a well-structured methodology. In contrast, non-academic sources such as blogs and or video demonstrations, would lack peer review and scientific validation, and thus are removed from the analysis to maintain research quality.

D. Recommended Peer-Reviewed Articles

- M. Sobhani et al., "Usability Study of a Novel Triple-arm Mixed-Reality Robot Teleoperation System," 2022 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR).
- T.-T. Nguyen et al., "A Continuous Real-time Hand Gesture Recognition Method based on Skeleton," 2022 IEEE ICCAIS.
- S. Bai et al., "Automatic Control System for Reach-to-Grasp Movement of a 7-DOF Robotic Arm Using Object Pose Estimation with an RGB Camera," 2023 IEEE ROBIO.
- J. Hossain Gourob et al., "A Robotic Hand: Controlled With Vision-Based Hand Gesture Recognition System," 2021 IEEE ACMI.
- Y. Yu et al., "ROS-based UAV control using hand gesture recognition," 2017 Chinese Control and Decision Conference (CCDC).

E. Contextualization of Literature

- Sobhani et al. investigate HRI in high-risk industrial applications, using VR for safe teleoperation.
- Nguyen et al. advance continuous gesture recognition, critical for dynamic, real-time systems.
- Bai et al. address robotic manipulation, bridging prosthetics and automation with precision pose estimation.
- Gourob et al. emphasise cost-effective assistive solutions for accessibility and healthcare.
- Yu et al. extend gesture control to UAVS, enhancing flexibility in outdoor and emergency scenarios.

F. Critical Comparison and Knowledge Gaps Strengths

- Rich experimental frameworks: (Sobhani et al., Bai et al.) integrating hardware, VR, and teleoperation metrics. This allows users a better visual of their environment and easier manipulation of equipment.
- Real-time recognition: (Nguyen et al.) with high accuracy and robust model performance. Reduces the possibilities of mistakes or injuries due to an error in operation.
- Practical applications: (Yu et al., Gourob et al.) in UAV control and low-cost prosthetics. Can provide an alternative if manual labour is not available or if there is a shortage of it.

G. Weaknesses

- Environmental constraints: Many studies (e.g., Sobhani, Bai) were conducted under controlled lab environments, reducing generalizability. Thus, the data collected will be inaccurate when performing in different real-world conditions.
- Hardware-Software Disconnect: Nguyen's work focused heavily on software modelling without integrating real robotic systems. This can cause problems if the model does not match the robotic system it is applied to, which can lead to orders although correct in the model appearing differently for the robotic hardware.
- Limited user studies: Some projects (e.g., Gourob) lacked extensive user testing or long-term performance metrics. Problems can arise when it comes to the integration of these systems if they are not adaptable to a diverse group of users.

H. Knowledge Gaps Identified

- Cross-environment testing: Adapting systems for different lighting, weather, or terrain conditions.
- Hardware robustness: Ensuring mechanical durability over extended operation.
- Longitudinal studies: Tracking usability and accuracy over long periods to measure system fatigue and adaptability.

I. Literature Map Gesture-based HRI Methodologies

- Vision-based Recognition
  - Skeleton + TD-Net [2]
  - RGB + Pose Estimation [3,4]

- Interface and Integration
- VR + Robotic Arms + SUS [1]
- ROS + UAV [5]

- Evaluation Methods
  - Quantitative (Accuracy, Completion Time) [1,2,3]
  - Qualitative (User Feedback, SUS) [1]
- Application Context
  - High-risk Industry and Surgery [1,3]
  - Assistive Robotics [4]
  - UAV Control [5]
- Research Gaps
  - Environment Robustness
  - Hardware Integration in ML Models
  - Longitudinal Performance Testing

J. Conclusion

Gesture-based systems show promise for improving human-robot interaction that requires precision and safety. However, for widespread adoption, further development is essential. In the future, the use of Artificial Intelligence (AI) can be used to assist in the identification and tracking of a person's arm. This role can be further expanded to real-time situations as notifying the user using external sensors and translating it to information that the user can understand. These can range from temperature fluctuation, identification of the structural integrity of a building, and radiation. Sites such as roboFlow allow users to annotate certain parts of an image and name it, for identification, such as a hand. This process requires the processing of thousands of images to train the dataset to identify what a hand is and what it is doing. The dataset can then be given to YOLO to run the program and verify its success. At the same time, the field of robotics will also have to advance to the point where its movement can copy ours accurately to perform said tasks as well as a human could without any of the risk. These robots will have to be built to do these specific jobs, Surgery robots will need to be small, nimble and accurate; construction robots will need strong servos, be resistant to the elements and be durable, etc. They can also be modular, allowing for adaptation to specific situations; these can range from being able to replace tools and equipment, to defence from specific environments, temperatures and hostile environments. This can range from protection from wildfires during firefighting, heating for cold climates, armour to defend from shrapnel and high explosives, radiation shielding and medical robotic arms with modular tools for what is needed in surgery.

K. References

M. Sobhani, A. Smith, M. Giuliani and T. Pipe, "Usability Study of a Novel Triple-arm Mixed-Reality Robot Teleoperation System," 2022 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Sevilla, Spain, 2022, pp. 217-223, doi: 10.1109/SSRR56537.2022.10018630. keywords: Headphones;Atmospheric measurements;Robot vision systems;Virtual reality;Streaming media;Cameras;Manipulators;Teleoperation;Mixed-Reality;Human-Robot Interaction;Remote presence;System Usability,

S. Bai, J. Guo, Y. Jiang, H. Yokoi and S. Togo, "Automatic Control System for Reach-to-Grasp Movement of a 7-DOF Robotic Arm Using Object Pose Estimation with an RGB Camera," 2023 IEEE International Conference on Robotics and Biomimetics (ROBIO), Koh Samui, Thailand, 2023, pp. 1-6, doi: 10.1109/ROBIO58561.2023.10354531. keywords: Visualization;Shape;Pose estimation;Robot vision systems;Neural networks;Control systems;End effectors,

T. -T. Nguyen et al., "A Continuous Real-time Hand Gesture Recognition Method based on Skeleton," 2022 11th International Conference on Control, Automation and Information Sciences (ICCAIS), Hanoi, Vietnam, 2022, pp. 273-278, doi: 10.1109/ICCAIS56082.2022.9990122. keywords: Automation;Gesture recognition;Feature extraction;Skeleton;Robustness;Real-time systems;Task analysis;hand gesture recognition;skeleton-based hand gesture recognition;continuous hand gesture recognition,

J. Hossain Gourob, S. Raxit and A. Hasan, "A Robotic Hand: Controlled With Vision Based Hand Gesture Recognition System," 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), Rajshahi, Bangladesh, 2021, pp. 1-4, doi: 10.1109/ACMI53878.2021.9528192. keywords: Service robots;Robot vision systems;Input devices;Human-robot interaction;Gesture recognition;Assistive robots;Older adults;Human Robot Interaction;Hand Gesture Recognition;Robotic Hand,

Y. Yu, X. Wang, Z. Zhong and Y. Zhang, "ROS-based UAV control using hand gesture recognition," 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, China, 2017, pp. 6795-6799, doi: 10.1109/CCDC.2017.7978402. keywords: Gesture recognition;Algorithm design and analysis;Drones;Cameras;Robots;Mechatronics;Automation;ROS;UAV Control;Hand Gesture Recognition,

III. Research Questions

A. Research Questions

- 1. Can vision-based control systems replicate real-time human arm movement compared to traditional input devices?
- 2. Can the combination of VR and hand tracking give the user a better understanding of the robot's environment compared to a simple screen and controls?
- 3. Will there be challenges in applying these systems and machines to different situations, as well as training those who have used an old system for so long

B. Research Objectives

- 1. The development of a vision-based system capable of tracking and replicating human arm movement in real-time.

- 2. Comparing the performance and user experience of gesture-based systems with traditional controllers.
- 3. Conducting usability testing to assess accuracy, latency and responsiveness.
- 4. Analyse limitations and potential improvements

C. Research Philosophies, Approaches and Paradigms

The philosophy that was chosen is Interpretivist, which supports the understanding of the user experiences and the effectiveness of the technology. A deductive approach will be used to test hypotheses taken from existing theories. The reach paradigm will utilise a mixed-methods framework, using both qualitative and quantitative data.

D. Chosen Methodology

The methodology used in the experiment will be mixed-methods. This involves using both qualitative data from the program and output, and quantitative data by using questionnaires.

E. Chosen Methodology and Design

The research involves studying results acquired from a program designed to capture human arm movement through the use of a vision-based camera system and replicating it using a robotic or digital arm. The user will have their hand in front of a camera and have their hand movement tracked and copied onto a digital or robotic hand. To test the difference, a test will be conducted between vision-based control and traditional control methods.

F. Reflections on Validity, Reliability and Generalizability

- 1. Validity: Using real-world scenario testing and standard evaluation methods.
- 2. Reliability: Repeated testing through different environments and scenarios
- 3. Generalizability: These results can then be used a data for domains such as prosthetics, VR or hazardous environment operations.

G. Ethical Considerations

- 1. Informed Consent: The participants will be fully informed of the purpose of this study and procedures.
- 2. Data Privacy: Collected data will remain anonymous and stored securely
- 3. Participant Safety: The participants will be kept at a safe distance, and wearing safety equipment is necessary
- 4. Bias Avoidance: There will be no bias in diversity during the selection phase.

...

References

[1] M. Sobhani, A. Smith, M. Giuliani, and T. Pipe, "Usability study of a novel triple-arm mixed-reality robot teleoperation system," in *2022 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2022, pp. 217–223.

[2] T.-T. Nguyen, N.-C. Nguyen, D.-K. Ngo, V.-L. Phan, M.-H. Pham, D.-A. Nguyen, M.-H. Doan, and T.-L. Le, "A continuous real-time hand gesture recognition method based on skeleton," in *2022 11th International Conference on Control, Automation and Information Sciences (ICCAIS)*, 2022, pp. 273–278.

[3] S. Bai, J. Guo, Y. Jiang, H. Yokoi, and S. Togo, "Automatic control system for reach-to-grasp movement of a 7-dof robotic arm using object pose estimation with an rgb camera," in *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2023, pp. 1–6.

[4] J. Hossain Gourob, S. Raxit, and A. Hasan, "A robotic hand: Controlled with vision based hand gesture recognition system," in *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, 2021, pp. 1–4.

[5] Y. Yu, X. Wang, Z. Zhong, and Y. Zhang, "Ros-based uav control using hand gesture recognition," in *2017 29th Chinese Control And Decision Conference (CCDC)*, 2017, pp. 6795–6799.