REAL-TIME HAND GESTURE CONTROL OF A ROBOTIC ARM WITH PROGRAMMABLE MOTION MEMORY

Daniel Giraldi Michels[1], Davi Giraldi Michels[1], Lucas Alexandre Zick[2,3], Dieisson Martinelli[1], André Schneider de Oliveira[1], and Vivian Cremer Kalempa[1]

[1] Federal University of Technology — Paraná, Curitiba, Paraná, Brasil

[2] Department of Information Systems, Universidade do Estado de Santa Catarina (UDESC), São Bento do Sul, Brazil

daniel.michels@edu.udesc.br, davi.michels@edu.udesc.br, zick@alunos.utfpr.edu.br, dmartinelli@alunos.utfpr.edu.br, andreoliveira@utfpr.edu.br, vivian.kalempa@udesc.br

ICINCO 2025

INSTICC

UDESC PLANALTO NORTE

INTRODUCTION

Current teleoperation methods often rely on physical devices like joysticks or data gloves (Zick *et al.*, 2024; Martinelli *et al.*, 2020), but are frequently limited by high cost, complexity, and a lack of intuitive control, hindering their widespread adoption (Moniruzzaman *et al.*, 2022).

To overcome these shortcomings, this paper proposes an innovative control method for articulated robotic arms. The proposed method utilizes real-time recognition of the user's hand position and movements using a single camera, as can be seen in Figure 1.

Figure 1 – System usage.

A distinctive feature of our system is its ability to store and reproduce movement sequences, which greatly simplifies the automation of repetitive tasks by enabling "Programming by Demonstration".

EXPERIMENTAL VALIDATION

To evaluate the system's effectiveness, precision, and repeatability, we designed a "pick and place" test, simulating a common industrial task. The experiment required the robotic arm to pick up a 4 cm cube from a platform elevated by 13 cm and place it on a ground-level surface approximately 30 cm away. An operator first demonstrated the task once using hand gestures, which the system recorded. Subsequently, the system was commanded to autonomously execute the stored movement for 50 consecutive cycles, during which the number of successful attempts and the time per cycle were logged for analysis. A demonstration video of the test is available at:

Figure 2 – A demonstration video of pick and place test.

CONCLUSION

This work validated a viable and effective solution for intuitive robotic programming, successfully bridging the gap between human intention and robotic execution. The high success rate (98%) and consistency in cycle time (0.0126s standard deviation) empirically confirm the system's robustness, precision, and repeatability. We demonstrate that a low-cost, camera-based approach can significantly reduce the complexity of traditional robotic programming, making automation more accessible. Future work will focus on implementing conditional logic based on gestures and abstracting the software to ensure portability across different robot models.

PROPOSED STRATEGY

Our system's architecture comprises three modules: Gesture Capture, Robot Command Mapping, and Trajectory Recording. The high-level control logic was implemented in Python, using an Interbotix PincherX-100 robotic arm controlled via the Robot Operating System (ROS) (Open Source Robotics Foundation, 2025). Real-time gesture capture is achieved with a standard camera and the MediaPipe library (Google, 2025), which tracks 21 3D hand landmarks. The hand's position is mapped to the robot's base and elbow joints, while a calculated depth variable controls the shoulder's reach. The gripper is actuated by the distance between the thumb and index fingertips. This intuitive control, facilitated by OpenCV (Bradski, 2025) for image processing and the `interbotix_xs_modules` library for arm communication, is enhanced by a programming module that allows the operator to Record (`'R'`), Play (`'P'`), and Loop (`'L'`) a demonstrated trajectory.

RESULTS

The system demonstrated high robustness and effectiveness, achieving a 98% success rate, with 49 out of 50 trials completed successfully presented in Table1. The quantitative analysis also confirmed the excellent precision and repeatability of the method. The mean cycle time was 13.35 seconds, with a very low standard deviation of only 0.0126 seconds. The execution time remained remarkably stable across all 50 repetitions, indicating no performance degradation over time and confirming the system's reliability for repetitive tasks.

Table 1 – "Pick and Place" Test Results (50 Repetitions)

Metric	Value
Total Trials	50
Successful Trials	49
Failures	1
Success Rate	98%
Mean Cycle Time (s)	13.35
Time Std. Deviation (s)	0.0126

REFERENCES

BRADSKI, G. The OpenCV Library. **Dr. Dobb's Journal of Software Tools**, 2025.

GOOGLE. **MediaPipe: A Framework for Building Perception Pipelines.**: 2025. `https://mediapipe.dev/`. Visited on: 9 Dec. 2025.

MARTINELLI, D. *et al.* Human-robot interface for remote control via iot communication using deep learning techniques for motion recognition. *In:* **2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)**. Natal: IEEE, 2020. p. 1–6.

MONIRUZZAMAN, M. *et al.* Teleoperation methods and enhancement techniques for mobile robots: A comprehensive survey. **Robotics and Autonomous Systems**, Elsevier, v. 150, p. 103973, 2022.

OPEN SOURCE ROBOTICS FOUNDATION. **Robotic Operating System (ROS).**: 2025. Available from: `https://www.ros.org/`. Visited on: 9 Dec. 2025.

ZICK, L. A. *et al.* Teleoperation system for multiple robots with intuitive hand recognition interface. **Scientific Reports**, Nature Publishing Group, v. 14, n. 1, p. 1–11, 2024.

ACKNOWLEDGEMENTS

The funding organizations, for the support received for the development of this work and participation in this event:

Declaration of Responsibility: the author(s) are solely responsible for the information contained in this document.