VOLUME 20, N° 2 2026
Journal of Automation, Mobile Robotics and Intelligent Systems

COMMUNICATION WITH THE MANIPULATOR USING GESTURES TO ENHANCE THE MANIPULABILITY OF PERSONS WITH REDUCED MOBILITY

Submitted: 20th January 2025; accepted: 13th May 2025

Sebastian Koryl, Katarzyna Zadarnowska, Krzysztof Arent

DOI: 10.14313/jamris-2026-015

Abstract:
This paper presents a robotic system that assists people with reduced mobility in the activities of picking up and putting down objects out of their reach. Human-robot communication is non-verbal, using gestures that have been specifically selected for the robot's use. Gestures are read out using an RGB-D camera while the commands they express are executed online by a small UR3 cobot. The evaluation of the system has shown that it is useful and safe in the sense of the SUS and GQS, respectively.

Keywords: *assistive robots, manipulators, 3D vision systems, gesture recognition, human-robot interactions*

1. Introduction

Robots as aids for people with disabilities are of wide interest to users, researchers, and engineers [1]. This is a consequence of the strong diversity of needs of people with disabilities, the high societal relevance of this branch of robotics, and the rapid advances in science and technology. A systematic review of assistive robots is included in [1, 2]. The classification proposed in [1] distinguishes manipulation aids. They are categorised into fixed base and wheelchair manipulator arm systems. MySpoon [3], Mico, Jaco [4], iARM [5] and BATEO [6] can serve as examples. In contrast, the classification in [2] is based on applications: assistance in activities of daily living, fetch and bring activities, food and beverage service, medication delivery services, user-related delivery services, telepresence and communication, monitoring safety, and navigation. This article concerns fixed base manipulator arm systems designed for fetch and bring activities.

The potential of assistive robotic manipulation to improve the quality of life for people with motor impairments is a motivation for work [7]. A Body-Machine Interface (BMI) has been proposed. It is based on on a vest that is equipped with four MTx motion trackers in order to capture shoulder movements. BMI uses a person's residual motor capabilities to generate control signals for a robotic arm. In this way, a high-DOF MICO manipulator can assist people with motor impairments to perform everyday tasks. It is noted in [8] that communication with a robot can be done through gestures, and there are circumstances when this is justified. The authors propose a wireless, human gesture based controlled robotic arm system for tool handling (pick and place) and other applications where human reach is elusive. The gestures are captured using IMU and flex sensors placed on the human hand, and they are in sync with the manipulators movement. The solution is intended primarily for industrial applications, not for assistive robotics. The article [9] presents a system for controlling a robot arm by human fingers and hand movements. The sensory hardware is composed of gyroscpes and flex connected to Arduino microcontroller. Sensory data processing is based on fuzzy logic. This is more of a technology demonstrator than a final solution but it is undoubtedly a step in the search for a good system for communicating with a robot using gestures. In [10] the authors propose a vision-based HCI architecture for the robotic arm by identifying somatosensory motion. The input of the model is collected by Kinect sensors from human body movement. The evaluation of the proposed system was preliminary and limited to the virtual environment.

Recently, a number of works have appeared, [11–13], based on the Google MediaPipe [14] that is an open-source framework that offers developers a platform for building real-time multimedia applications. A gesture control interface for laparoscopic surgery is discussed in [11]. Four RGB cameras are used to capture hand movements. Although the study offers a solid mathematical background and includes user evaluation, it lacks real-world testing. A gesture controlled robotic arm for surgical tool assistance is proposed in [12]. The system is based on YOLOv5 for accurate surgical tool detection, MediaPipe Framework for real-time hand gesture tracking, and the UFactory Lite 6 robotic arm. RGB-D and RGB cameras are used for hand movement capture and tool detection, respectively. The recognized gestures have the meaning of instructions to the manipulator, no telemanipulation takes place. The evaluation was carried out in a virtual RobotDK environment and was based exclusively on selected system performance indicators. The concept of a gesture control interface for teleoperation of quadruped robots with a robotic arm is presented in [13]. An RGB-D camera (Intel RealSense) captures human gestures. They have the meaning of motion commands for the Unitree Go 1 mobile platform. The position of the effector of myCobot manipulator can be teleoperated by hand movement using a custom teleoperatin algorithm. However, evaluation of the system is limited to a basic feasibility study.

There are reports in the literature about research on other forms of human-manipulator communication. In [15], the authors present and discuss a tongue drive system in conjunction with augmented reality for an assistive manipulator. A P300-based brain-computer interface is proposed and analysed in [16]. A minimalist interface for a proof-of-concept control system with higher autonomy for head care tasks using the PR2 robot is proposed in [17]. In the article [18], hand configuration in combination with voice commands are used to define the actions of the manipulator. The RGB-D camera is used to segment the scene, detect objects, including the hand, its position, and the object it is pointing at in the scene.

In parallel with the development of human-robot communication technology, the problem of evaluation of such interactive systems has been addressed. An analysis of the previously cited works leads to the conclusion that two types of evaluation parameters can be distinguished: system performance indicators for feasibility study and user-centered UX/HRI measures for users' opinions study. The system performance indicators include task realization success rate [11, 13, 15, 18], task completion time [7, 11, 13, 15, 16, 18–20], response time [11, 12, 16, 18] task execution accuracy [11,18,20]. In turn, user experience has been assessed on the following measures: Quality of Life in Essential Tremor Questionnaire (QUEST; [19]), System Usability Scale (SUS; [11,19,20]), Godspeed Questionnaire Series (GQS) and Attitudes towards Technology Scale (ATTS; [20]), Van Der Laan's technology acceptance scoring, Ergonomics survey [11], the Questionnaire for the Evaluation of Physical Assistive Devices (QUEAD; [21]). This classification provides a structured framework for comparing our system with prior art.

The contribution of this article consists of four elements that are present together.

1) The concept of a manipulator designed to assist individuals with reduced mobility in tasks involving physical objects handling has been proposed. Human-robot interaction is facilitated through gesture-based commands, which are captured by an RGB-D camera and processed using Google MediaPipe Hands software. Crucially, our design eliminates the need for wearable or hand-mounted sensors to indicate to the manipulator effector the position to be tracked and the actions to be performed with the gripper. This sensor-less control mechanism increases user comfort and accessibility.
2) A complete mathematical background is provided to support the gesture-based control logic, enabling robust and responsive interaction between the user and the robotic arm.
3) The gesture control algorithm was implemented and tested on an existing robotics platform comprising hardware and software, confirming its feasibility outside a simulation environment.
4) The evaluation of the gesture control system was carried out in a significant group of 37 naive participants. UX and HRI tests based on live experiments involving a scenario specific to an assistive robotic arm yielded positive results, i.e. the system is sufficiently useful and safe in the sense of SUS and GQS measures.

Initial research for this work was developed in [22] and [20].

The solution proposed in this work has a simpler design and is more user-friendly than the systems discussed in [8–10]. At the same time, it is complementary to the systems presented in [15, 16] and [18], and contributes to the growing field of accessible, sensor-less human–robot interfaces.

2. System Idea

The general concept of the manipulation aid system is shown in Figure 1.

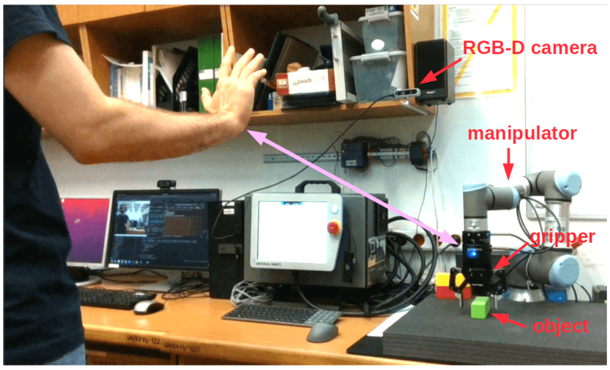


Figure 1. Overall concept of the system

The user can move objects out of their reach into their immediate environment with the support of the manipulator. Commands are given to the manipulator by means of hand gestures, which are recorded using the RGB-D camera.

The set of commands interpreted by the manipulator's sensory system are activation/deactivation, hand tracking, opening and closing of the gripper. Once the gesture command mode is activated, the gripper follows the position of the hand. The gripper can be open or closed, depending on whether the hand is open or closed with a fist. The above makes the manipulator an extension of the user's arm from the user's perspective.

3. System Components

3.1. General System Architecture

The system architecture implementing the concept outlined in Section 3.1 is presented in Figure 2. The user hand is captured by an RGB-D camera. The resulting RGB and depth images are processed by the `/camera_image_processor` module. The extracted hand landmarks positions and recognized gestures are then used to determine the instructions for the robot through the `/robot_controller` module. Consequently, the coordinate system of the gripper follows the coordinate system of the hand and the jaws

Figure 2. General system architecture

of the gripper follow the instructions expressed by the hand gestures.

3.2. Hardware

The hardware platform of the assistive manipulator under consideration consists of the following components:

- UR3 manipulator with CB3 control box and Robotiq 2F-85 two-finger gripper,
- desktop PC,
- Intel RealSense D435 depth camera (RGB-D).

The camera is connected to the PC via a USB cable while the PC communicates with the CB3 control box via TCP/IP.

The UR3 manipulator is a small cobot and can be adapted as an assistive manipulator. The camera built into the UR3 manipulator is of the RGB type and cannot be used for the gripper to track human hand movement in 3D. For this reason, it was necessary to use an external RGB-D camera.

3.3. Software

Ubuntu 20.04 and ROS Noetic form the software base on the PC for this project. The software of the considered assistive manipulator consists of a number of ROS nodes that are associated with Universal Robots ROS Driver, MoveIt Motion Planning Framework, and gripper. In addition, there are two custom nodes: `/camera_image_processor` and `/robot_controller` mentioned in Section 3.1. A part of the ROS graph with these nodes is included in Figure 3.

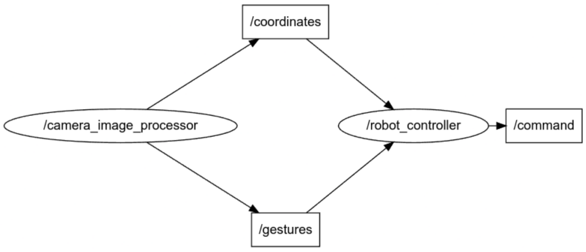
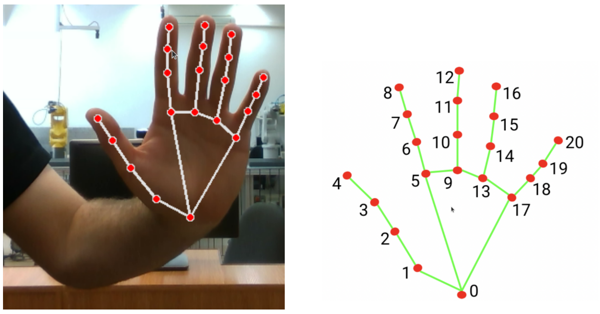


Figure 3. ROS graph

/camera_image_processor is a software which performs RGB image and depth map acquisition and then carries out the data processing.



(a) a human hand and its model in MediaPipe

(b) hand landmarks in the MediaPipe hand model

Figure 4. Google MediaPipe Software Framework: a hand model

The Google MediaPipe Framework [14, 23] was used to implement this node. The work [24] proves that it is a reliable and precise framework for assessing 3D hand movements in clinical applications. Specifically, MediaPipe Hands was used to track hand position and orientation, similar to [11].

The MediaPipe hand landmark recognition model recognizes the hand in the captured data and creates a skeletal model of it, illustrated in Figure 4(a). There are 20 landmarks associated with the hand model shown in Figure 4(b). They can be used to define a coordinate frame $\mathcal{X}_h\mathcal{Y}_h\mathcal{Z}_h$ associated with the hand (a hand coordinate frame).

Let ${}^cp_{h,i}$ denotes a position of the i-th hand landmark in the camera coordinate frame $\mathcal{X}_c\mathcal{Y}_c\mathcal{Z}_c$. The versors $r_{h,x}, r_{h,y}\, r_{h,z}$ of $\mathcal{X}_h\mathcal{Y}_h\mathcal{Z}_h$ are defined as follows:

$$r_{h,z} = \frac{{}^cp_{h,9} - {}^cp_{h,0}}{\|{}^cp_{h,9} - {}^cp_{h,0}\|_2}, \tag{1}$$

$$r_{h,y} = \frac{({}^cp_{h,17} - {}^cp_{h,0}) \times ({}^cp_{h,9} - {}^cp_{h,0})}{\|({}^cp_{h,17} - {}^cp_{h,0}) \times ({}^cp_{h,9} - {}^cp_{h,0})\|_2}, \tag{2}$$

$$r_{h,x} = r_{h,y} \times r_{h,z}. \tag{3}$$

The origin of the hand coordinate frame is located at ${}^cp_{h,0}$. Consequently, a homogeneous transformation matrix from $\mathcal{X}_h\mathcal{Y}_h\mathcal{Z}_h$ to $\mathcal{X}_c\mathcal{Y}_c\mathcal{Z}_c$ takes the following form:

$$T_h^c = \begin{bmatrix} {}^cR_h & {}^cp_{h,0} \\ 0 & 1 \end{bmatrix} \text{ where } {}^cR_h = \begin{bmatrix} r_{h,x} & r_{h,y} & r_{h,z} \end{bmatrix}. \tag{4}$$

The above construct makes some similarities between the coordinate frames of the human hand and the gripper (the latter is denoted by $\mathcal{X}_e\mathcal{Y}_e\mathcal{Z}_e$). In both cases, the z versors have their origin at the wrist and are directed along the fingers to the outside of the arm. In turn, the y versors are directed perpendicular to the plane of the palm on the outside of the inner part of the palm. In the case of a gripper, the inner part of the palm is assumed to be on the camera side of the gripper.

Let T_c^b denotes a homogeneous transformation matrix from $\mathcal{X}_c\mathcal{Y}_c\mathcal{Z}_c$ to the manipulator base coordinate frame $\mathcal{X}_b\mathcal{Y}_b\mathcal{Z}_b$, [25], and $T_h^b := T_c^b T_h^c$. The structure of the resulting matrix is as follows:

$$T_h^b = \begin{bmatrix} {}^bR_h & {}^bp_{h,0} \\ 0 & 1 \end{bmatrix}. \tag{5}$$

The matrix ${}^bR_h$ represents the orientation of $\mathcal{X}_h\mathcal{Y}_h\mathcal{Z}_h$ in $\mathcal{X}_b\mathcal{Y}_b\mathcal{Z}_b$ while ${}^bp_{h,0}$ represents the position of the human wrist in $\mathcal{X}_b\mathcal{Y}_b\mathcal{Z}_b$.

T_h^b is used to determine a reference path for $\mathcal{X}_e\mathcal{Y}_e\mathcal{Z}_e$ in $\mathcal{X}_b\mathcal{Y}_b\mathcal{Z}_b$. Let ${}^b\bar{p}_{h,0}$, ${}^b\bar{R}_h$, ${}^b\bar{p}_e$, ${}^b\bar{R}_e$ are the positions and the orientations of a human hand and a gripper respectively when the control system is activated. Define

$$\bar{p} := {}^b\bar{p}_e - {}^b\bar{p}_{h,0}, \tag{6}$$

$$\bar{R} := ({}^b\bar{R}_h)^{-1} \cdot {}^b\bar{R}_e. \tag{7}$$

Then the reference path for the gripper can be expressed as follows:

$${}^bp_{e,\mathrm{ref}} = {}^bp_{h,0} + \bar{p}, \tag{8}$$

$${}^bR_{e,\mathrm{ref}} = {}^bR_h \cdot \bar{R}. \tag{9}$$

Sometimes it is good to control only the position of the effector, especially when a twitching human palm makes the gripper oscillates and once fixed the orientation of the gripper is sufficient. In this case, the expression (9) takes the form

$${}^bR_{e,\mathrm{ref}} = {}^b\bar{R}_e. \tag{10}$$

Note, that similar mappings are used in direct and bilateral teleoperation [26]. Consequently, two working modes can be distinguished:

full teleoperation, defined by (1)÷(9);

reduced teleoperation, defined by (1)÷(8), (10).

${}^bp_{e,\mathrm{ref}}$, ${}^bR_{e,\mathrm{ref}}$ are the quantities that are passed to the `/coordinates` topic.

The MediaPipe hand gesture recognition model can recognize several static gestures. The gestures supported by the system are: *closed fist*, *open palm*, *thumb up*, *victory*, *none of them*. They are illustrated in Figure 5. The gesture type is published to the `/gestures` topic once it has been recognized.

Both the position of the wrist and the recognized gestures have an assigned meaning in terms of commands for the manipulator. The position of the wrist determines the position of the gripper. The *open palm* and *closed fist* gestures denote open and closed gripper jaws, respectively. The *thumb up* or *victory* gestures mean activation/deactivation of the gesture teleoperation command mode.

(a) thumb up (b) victory

(c) open palm (d) closed fist

Figure 5. Google MediaPipe Software Framework: recognizable static gestures

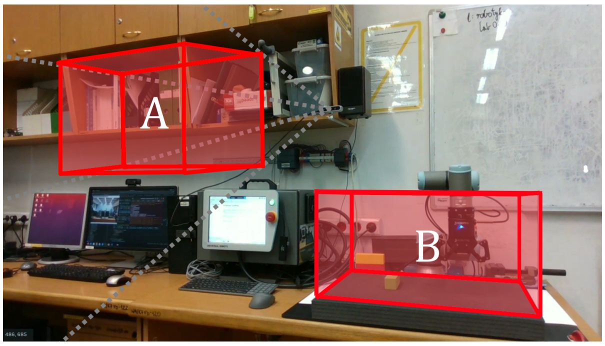


Figure 6. Camera workspace (A) and gripper workspace (B)

The system operates within predefined workspaces in 3D space, as illustrated in Figure 6. Workspace A corresponds to the region within the camera field of view, where user gestures are captured and processed as commands for the robot. Workspace B defines the area within which the gripper operates. Its position remains confined to this designated region. Both workspaces are rectangular in shape, and their exact dimensions can be adjusted to suit specific requirements.

/robot_controller is a high-level software module that processes the data subscribed from the `/coordinates` and `/gestures` topic into commands for the manipulator.

The hand coordinates are converted into coordinates of the robot's base coordinate frame. These coordinates are then set as the target position for the gripper. The manipulator's point-to-point movement is handled by the MoveIt motion planning framework. It provides access to algorithms from the Open Motion Planning Library (OMPL). Robot movement trajectories are calculated with the Rapidly-exploring Random Trees (RRT) Connect algorithm in this system.

The hand gesture dictates whether the gripper opens or closes, and if the control system is activated or deactivated. The resulting data representing the desired motion is stored in the fields of a variable, which is then published to the `/command` topic.

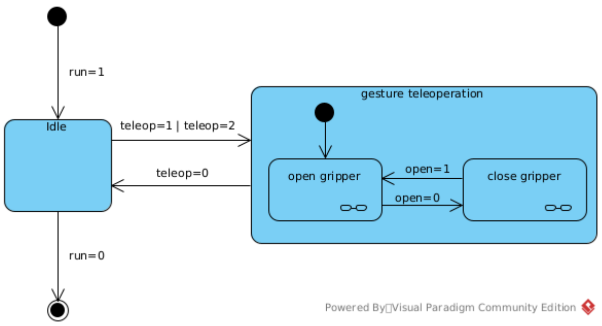


Figure 7. UR3 manipulator gesture control representation using finite state machine

The behaviour of the system is represented by the FSM diagram shown in Figure 7. At startup, the system enters the idle state. The *thumb up* gesture (see Figure 5(a)) toggles the value of the `teleop` variable from 0 to 1, while the *victory* gesture (see Figure 5(b)) toggles it from 0 to 2. A value of 1 corresponds to reduced teleoperation, while a value of 2 corresponds to full teleoperation. When either value is set, the gesture teleoperation state becomes active.

In this state, the manipulator's gripper is teleoperated based on the user's hand gestures according to algorithms (1)–(8), (9), or (10). The system recognises both an *open palm* gesture (see Figure 5(b)) and a *closed fist* gesture (see Figure 5(c)).

Initially, the open gripper state is active. A *closed fist* gesture toggles the value of the open variable between 1 and 0, activating the open gripper and close gripper states, respectively. These states correspond to the open or closed gripper of the manipulator.

The gesture teleoperation state is deactivated when the value of the `teleop` variable is reset to 0 using either the *thumb up* or *victory* gesture. At this point, the system returns to the idle state.

4. System Evaluation

4.1. Research Question

The study aims to answer the following research questions in relation to the system presented in Section 2:

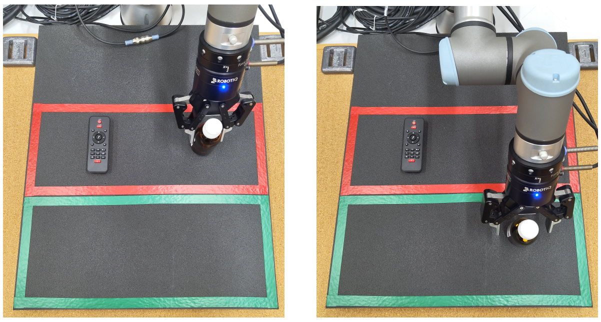
RQ1 Is the system usable in the user's view?

RQ2 Is the system safe from the user's perspective?

4.2. Method

Participants The study involved 37 participants aged 22-25 years old (Med $= 23, Q_1 = 22, Q_3 = 23$). There were 4 females and 33 males in the group. All of them were students of control engineering and robotics at Wrocław University of Science and Technology.

Informed consent was obtained from all subjects involved in the study after they had been acquainted with the research objectives and procedures used in the study.



(a) the beginning (b) the end

Figure 8. Snapshots of the experimental task being carried out

Scenario The task scenario is illustrated in Figure 8. The participant is asked to move everyday objects, a TV remote control and a syrup bottle, from a further away area to a closer location. The further and closer areas are marked with a red and green border, respectively. A similar task in the evaluation scenario can be found in [7] and [16]. All commands addressed to the robot must be expressed in gestures. The set of commands consists of: activation/deactivation of the system, opening and closing of the gripper, and movement of the gripper (see Fig. 7 in Section 3.3).

The reduced teleoperation was selected to be the working mode during the study.

Procedure The team first trained the participant how to operate the robot. As part of the training, the participant had to practise the task of carrying and bringing back an item, in the form of a plastic block. Then the participant performed the task from the scenario. After the experiment, the participant had to fill out the questionnaires. The experiment was conducted in a fully simulated environment in the Robotic Laboratory at Wrocław University of Science and Technology.

The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Research Ethics Committee of Wrocław University of Science and Technology, Opinion No. O-24-63.

Measures The users' impressions have been collected by means of questionnaires. To answer our research questions concerning usability, we used the System Usability Scale (SUS), a well established questionnaire for assessing the usability of the system containing 10 items on a five point Likert scale ranging from totally disagree to totally agree [27]. To assess perceived safety, we used the Godspeed (GQS) questionnaire for perceived safety (3 items using semantic differential scales) [28]. We also used the Negative Attitude towards Robots scale (NARS-PL) [29], Polish version, in order to gather information about participants' attitude towards technology (containing also 13 items on a five point likert), which could impact the results.

4.3. Results

Internal Reliability of the Questionnaires

- For the 10 SUS items the Cronbach's Alpha was 0.9575 (excellent reliability).
- For the 13 NARS items the Cronbach's Alpha was 0.9561 (excellent reliability).
- For the 3 GS items the Cronbach's Alpha was 0.8917 (very good reliability).

The System Usability Scale It follows from Chi-square goodness-of-fit test that the cummulative SUS score comes from a normal distribution (a 100-point scale, $M = 69.5, SD = 7.0, \alpha = 0.05, p = 0.37$). Furthemore, the one-sample t-test allows to conclude that the mean of SUS score is 70, which can be considered as good according to [30] ($\alpha = 0.05, p = 0.19$).

Godspeed Questionnaire Series: Perceived Safety (PL) It follows from Chi-square goodness-of-fit test that the cummulative GQS score comes from a normal distribution (a 18-point scale, $M = 13.5, SD = 2.3, \alpha = 0.05$, $p = 0.05$). Furthemore, the one-sample t-test allows to conclude that the mean of GQS safety score is 13 ($\alpha = 0.05, p = 0.19$).

Negative Attitude Towards Robots Scale (PL) It follows from Chi-square goodness-of-fit test that the cummulative NARS-PL score comes from a normal distribution (a 48-point scale $M = 16.7, SD = 6.4, \alpha = 0.05$, $p =< 0.001$). Furthemore, the one-sample t-test allows to conclude that the mean of NARS-PL score is 17 ($\alpha = 0.05, p = 0.76$).

4.4. Discussion

With regard to the research questions RQ1 and RQ2, it can be concluded that participants with low negative attitudes toward robots assess the tested system as sufficiently useful and sufficiently safe.

When evaluating a system consisting of a human and a robotic assistive arm, user feedback appears to be most relevant. If the results of the users' study are satisfactory, then it can be tentatively assumed that system performance indicators are also satisfactory. The implication in the oposit direction is not obvious.

The evaluation in most previous work on gesture communication with the assistive robotic arm and with other robots is based on system performance indicators [7, 11, 13, 15, 16, 18–20]. The use of UX and HRI studies can be noted in less numerous and more recent works [11, 19–21]. This makes it challenging to do a reliable comparative analysis with results presented in other works.

The satisfactory result of the SUS test is confirmation of the soundness of the proposed gesture-based control of the manipulator by the users. A satisfactory score of perceived safety means that there was no behavior of the robotic system during the execution of the task by the participant that would interfere with his/her sense of security.

The results obtained, although positive, leave some room for improvement. It would seem that changing the appearance of the manipulator to be less industrial and the gripper to be more hand-like would increase the level of perceived safety. BATEO [6], iARM [5], MICO [7] can serve as examples.

Gesture-based manipulator control systems discussed in [8] and [9] require appropriate sensors to be mounted on the hand. The solution proposed in this paper, which does not require such sensors, is definitely more user-friendly. A similar solution with an RGB-D (Kinect) sensor for gesture identification and arm movement is discussed in [10], but the sensory system proposed there was not integrated into the physical robot and no evaluation was performed with the users (which is the case here).

The works [4, 15, 16] focus on different forms of communication (using tongue, brainwaves, joystick) therefore the gesture-based communication proposed here can be seen as a complement to these. All the previously mentioned forms of communication can form the basis of a simple manual control mode for the robot, but they can also be part of a more elaborate semi-autonomous control system [3].

In the work [18] there is no hand telemanipulation as here, but voice commands are responsible for the movement of the manipulator. The hand is used to indicate target locations for the manipulator's gripper. This solution may be difficult to use for users with speech disabilities. On the other hand, both solutions can coexist and complement each other.

The gesture control system presented in this article uses MediaPipe Hands similary to [11, 12] and [13] and Intel RealSense RGB-D camera similary to [12] and [13]. However, in this paper, the target user is a person with reduced mobility, while in the others, these are doctors or indeterminate users.

In this work the manipulator is meant to be an extension of the user's hand, and for this reason, telemanipulation using gestures is of primary importance from the viewpoint of potential applications. This issue is explored here in more depth compared to [11, 12] and [13]. What's more, the evaluation here is based on live experiments, based on a scenario close to the activities of daily living with a larger group of study participants. Therefore the result obtained here is more reliable compared to the conclusions that can be drawn from [11–13], where the number of study participants was significantly smaller and the studies themselves were conducted in a virtual environment [11, 12].

In our context, MediaPipe Hands, combined with depth data from the RGB-D camera, enables robust detection of complex hand gestures without requiring users to wear any sensors or markers. This significantly improves user comfort and reduces setup time, while maintaining high gesture recognition accuracy.

5. Conclusion

It has been proposed and studied a robotic system that can assist people with reduced mobility in the activities of picking up and putting down objects out of reach. The manipulator in this system acts as an elongation of a person's arm. The movement of the manipulator's gripper is done by telemanipulation by the user's hand, whose movement is captured by the camera. A few commands given to the robot are expressed using gestures. The sensory part of the system is based on the RGB-D camera (Intel RealSense D435) and Google MediaPipe Software Framework. During the evaluation, the system worked reliably and participants assessed it as useful and safe.

The UR3 manipulator used in this study is not the primary focus of the system. The core software component of the proposed system can be easily adapted for use with other manipulators that are compatible with ROS/ROS2 and the MoveIt software environment.

The results obtained justify the continuation of work at a higher level of technological readiness. This work should be based on another collaborative robot (better suited to the requirements of an assistive robotic arm) a much wider and more diverse group of participants in UX and HRI studies, and a wider spectrum of system performance indicators to evaluate the gesture-based communication with an assistive robotic arm.

AUTHORS

Sebastian Koryl – Faculty of Electronics, Photonics and Microsystems, Wrocław University of Science and Technology27 Wybrzeże Stanisława Wyspiańskiego st., 50-370 Wrocław, Poland, e-mail: sebastian.koryl@gmail.com.

Katarzyna Zadarnowska* – Department of Cybernetics and Robotics, Wrocław University of Science and Technology, 27 Wybrzeże Stanisława Wyspiańskiego st., 50-370 Wrocław, Poland, e-mail: katarzyna.zadarnowska@pwr.edu.pl
https://wefim.pwr.edu.pl/pracownicy/wizytowki-pracownikow/profil.html?name=katarzynazadarnowska-1186.

Krzysztof Arent – Department of Cybernetics and Robotics, Wrocław University of Science and Technology27 Wybrzeże Stanisława Wyspiańskiego st., 50-370 Wrocław, Poland, e-mail: krzysztof.arent@pwr.edu.pl
https://wefim.pwr.edu.pl/pracownicy/wizytowki-pracownikow/profil.html?name=krzysztofarent-1219.

*Corresponding author

References

[1] H. F. M. Van der Loos and D. J. Reinkensmeyer. *Springer Handbook of Robotics*, chapter "Rehabilitation and Health Care Robotics", 1223–1251. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[2] A. Nanavati, V. Ranganeni, and M. Cakmak, "Physically Assistive Robots: A Systematic Review of Mobile and Manipulator Robots that Physically Assist People with Disabilities", *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 7, no. Volume 7, 2024, 2024, 123–147, https://doi.org/10.1146/annurev-control-062823-024352

[3] R. Soyama, S. Ishii, and A. Fukase, "8 Selectable Operating Interfaces of the Meal-Assistance Device "My Spoon". In: *Advances in Rehabilitation Robotics: Human-Friendly Technologies on Movement Assistance and Restoration for People with Disabilities*, Springer Berlin Heidelberg, 2004, 155–163, 10.1007/10946978_8

[4] A. Campeau-Lecours, H. Lamontagne, S. Latour, P. Fauteux, V. Maheu, F. Boucher, C. Deguire, and L.-J. C. L'Ecuyer, "Kinova Modular Robot Arms for Service Robotics Applications", *International Journal of Robotics Applications and Technologies*, vol. 5, no. 2, 2017, 10.4018/IJRAT.2017070104

[5] "iARM". Assistive Innovations, 2026. https://www.assistive-innovations.com/robotarm

[6] "BATEO". ACCREA, 2026. https://accrea.com/products/bateo/

[7] S. Jain, A. Farshchiansadegh, A. Broad, F. Abdollahi, F. Mussa-Ivaldi, and B. Argall, "Assistive Robotic Manipulation through Shared Autonomy and a Body-Machine Interface". In: *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*, 2015, 526–531, 10.1109/ICORR.2015.7281253

[8] A. Vijay Joseph, A. Mathur, J. Verma, and A. Singh, "Gesture Based Wireless Control for a Robotic Manipulator", *International Journal of Engineering & Technology*, vol. 7, no. 2.31, 2018, 231–234, 10.14419/ijet.v7i2.31.13449

[9] P. Sihombing, R. B. Muhammad, H. Herriyance, and E. Elviwani, "Robotic Arm Controlling Based on Fingers and Hand Gesture". In: *2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)*, 2020, 40–45, 10.1109/MECnIT48290.2020.9166592

[10] X. Li, "Human–Robot Interaction Based on Gesture and Movement Recognition", *Signal Processing: Image Communication*, vol. 81, 2020, 115686, https://doi.org/10.1016/j.image.2019.115686

[11] W. Rhee, Y. G. Kim, J. H. Lee, J. W. Shim, B. S. Kim, D. Yoon, M. Cho, and S. Kim, "Unconstrained Lightweight Control Interface for Robot-Assisted Minimally Invasive Surgery Using MediaPipe Framework and Head-Mounted

Display", *Virtual Reality*, vol. 28, no. 114, 2024, https://doi.org/10.1007/s10055-024-00986-1

[12] A. R. Khan, A. Khan, and A. Faisal, "Gesture-Controlled Robotic Arm for Real-Time Surgical Tool Assistance". In: *2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T)*, 2025, 886–891, 10.1109/ICPC2T63847.2025.10958621

[13] J. Xie, Z. Xu, J. Zeng, Y. Gao, and K. Hashimoto, "Human–Robot Interaction Using Dynamic Hand Gesture for Teleoperation of Quadruped Robots with a Robotic Arm", *Electronics*, vol. 14, no. 5, 2025, 10.3390/electronics14050860

[14] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann. "MediaPipe: A Framework for Building Perception Pipelines", 2019.

[15] F.-J. Chu, R. Xu, Z. Zhang, P. A. Vela, and M. Ghovanloo, "Hands-Free Assistive Manipulator Using Augmented Reality and Tongue Drive System". In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, 5463–5468, 10.1109/IROS.2018.8594508

[16] P. D. Lillo, F. Arrichiello, D. D. Vito, and G. Antonelli, "BCI-Controlled Assistive Manipulator: Developed Architecture and Experimental results", *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 1, 2021, 91–104, 10.1109/TCDS.2020.2979375

[17] K. P. Hawkins, P. M. Grice, T. L. Chen, C.-H. King, and C. C. Kemp, "Assistive Mobile Manipulation for Self-Care Tasks Around the Head". In: *2014 IEEE Symposium on Computational Intelligence in Robotic Rehabilitation and Assistive Technologies (CIR2AT)*, 2014, 16–25, 10.1109/CIRAT.2014.7009736

[18] B. Kulecki, "Multimodal Robot Programming Interface Based On RGB-D Perception and Neural Scene Understanding Modules", *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 17, no. 3, 2024, 29–37, 10.14313/JAMRIS/3-2023/20

[19] C.-S. Chung, K. Williams, S. M. Eckstein, I. Eckstein, R. Cooper, and R. A. Cooper, "Manipulation Performance of an Assistive Robotic Arm Control Interface for Individuals with Cerebral Palsy - A Case Study". In: *Proceedings of the RESNA Annual Conference*, 2022.

[20] K. Arent, J. Jakubiak, M. Drwięga, M. Cholewiński, G. Stollnberger, M. Giuliani, M. Tscheligi, D. Szczęśniak-Stańczyk, M. Janowski, W. Brzozowski, and A. Wysokiński, "Control of Mobile Robot for Remote Medical Examination: Design Concepts and Users' Feedback from Experimental Studies". In: *2016 9th International Conference on Human System Interactions (HSI)*, 2016, 76–82, 10.1109/HSI.2016.7529612

[21] A. Lebrasseur, J. Lettre, F. Routhier, P. S. Archambault, and A. Campeau-Lecours, "Assistive Robotic Arm: Evaluation of the Performance of Intelligent Algorithms", *Assistive Technology*, vol. 33, no. 2, 2021, 95–104, 10.1080/10400435.2019.1601649, PMID: 31070524.

[22] S. Koryl. "Technology for Human-Manipulator Communication Using Gestures". Master's thesis, Wrocław University of Science and Technology, 2024.

[23] Google. *MediaPipe Solutions Guide*, 2024. https://ai.google.dev/edge/mediapipe/solutions/guide

[24] G. Amprimo, G. Masi, G. Pettiti, G. Olmo, L. Priano, and C. Ferraris, "Hand Tracking for Clinical Applications: Validation of the Google MediaPipe Hand (GMH) and the Depth-Enhanced GMH-D Frameworks", *Biomedical Signal Processing and Control*, vol. 96, 2024, 106508, https://doi.org/10.1016/j.bspc.2024.106508

[25] Y. Alemu. "Robotic System Technology to Support the Manipulability of a Person with Motor Disabilities". Master's thesis, Wrocław University of Science and Technology, 2024.

[26] G. Niemeyer, C. Preusche, S. Stramigioli, and D. Lee, "Telerobotics". In: *Springer Handbook of Robotics*, 2016.

[27] R. A. Grier, A. Bangor, P. Kortum, and S. C. Peres, "The System Usability Scale: Beyond Standard Usability Testing", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, 2013, 187–191, 10.1177/1541931213571042

[28] R. Szczepanowski, T. Niemiec, E. Cichoń, K. Arent, M. Florkowski, and J. Sobecki, "Factor Analysis of the Polish Version of Godspeed Questionnaire (Gqs)", *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 16, no. 2, 2023, 30–35, 10.14313/JAMRIS/2-2022/13

[29] G. Pochwatko, J.-C. Giger, M. Różańska-Walczuk, J. Świdrak, K. Kukiełka, J. Możaryn, and N. Piçarra, "Polish Version of the Negative Attitude Toward Robots Scale (NARS-PL)", *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 9, no. 3, 2015, 65–72, 10.14313/JAMRIS_2-2015/25

[30] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale", *Journal of Usability Studies*, vol. 4, no. 3, 2009, 114–123.