

Article

Hybrid Target Selections by “Hand Gestures + Facial Expression” for a Rehabilitation Robot

Yi Han ^{1,2,†}, Xiangliang Zhang ^{1,†}, Ning Zhang ³, Shuguang Meng ⁴, Tao Liu ¹ , Shuoyu Wang ², Min Pan ⁵, Xiufeng Zhang ^{3,*} and Jingang Yi ⁶

- ¹ The State Key Laboratory of Fluid Power and Mechatronic Systems, School of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China
- ² Department of Intelligent Mechanical Systems Engineering, Kochi University of Technology, 185 Miyanakuchi, Tosayamada-Cho, Kochi 782-8502, Japan
- ³ Key Laboratory of Rehabilitation Technical Aids Technology and System of the Ministry of Civil Affairs, National Research Center for Rehabilitation Technical Aids, Beijing 100176, China
- ⁴ Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650000, China
- ⁵ Department of Mechanical Engineering, University of Bath, Bath BA2 7AY, UK
- ⁶ Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, NJ 08854, USA
- * Correspondence: zhangxiufeng@nrcrta.cn
- † These authors contributed equally to this work.

Abstract: In this study we propose a “hand gesture + face expression” human machine interaction technique, and apply this technique to bedridden rehabilitation robot. “Hand gesture + Facial expression” interactive technology combines the input mode of gesture and facial expression perception. It involves seven basic facial expressions that can be used to determine a target selecting task, while hand gestures are used to control a cursor’s location. A controlled experiment was designed and conducted to evaluate the effectiveness of the proposed hybrid technology. A series of target selecting tasks with different target widths and layouts were designed to examine the recognition accuracy of hybrid control gestures. An interactive experiment applied to a rehabilitation robot is designed to verify the feasibility of this interactive technology applied to rehabilitation robots. The experimental results show that the “hand + facial expression” interactive gesture has strong robustness, which can provide a novel guideline for designing applications in VR interfaces, and it can be applied to the rehabilitation robots.

Keywords: facial expression; hybrid control gestures; interactive tasks; rehabilitation robot; target selection



Citation: Han, Y.; Zhang, X.; Zhang, N.; Meng, S.; Liu, T.; Wang, S.; Pan, M.; Zhang, X.; Yi, J. Hybrid Target Selections by “Hand Gestures + Facial Expression” for a Rehabilitation Robot. *Sensors* **2023**, *23*, 237. <https://doi.org/10.3390/s23010237>

Academic Editor: Eduardo Palermo

Received: 18 November 2022

Revised: 21 December 2022

Accepted: 22 December 2022

Published: 26 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gestures and facial expressions have been the focus of a lot of recent research. They offers a compelling platform for touchless interactions. The advantages and applications of face detection are well-explored in the literature [1–8]. The advantages of facial expressions including the Face Switch have been well documented and investigated in research [9]. It’s a device that can interact with a computer through a combination of facial expressions and gestures.

In this paper, a new computer interactive hybrid gesture control technique is proposed. User facial expression data and gesture data are collected based on the camera and Leap Motion [10–12], providing a new input dimension for interaction. This design realizes the combination of facial expressions and gestures, the selection of targets and the control of rehabilitation robots and other operations (Figure 1).

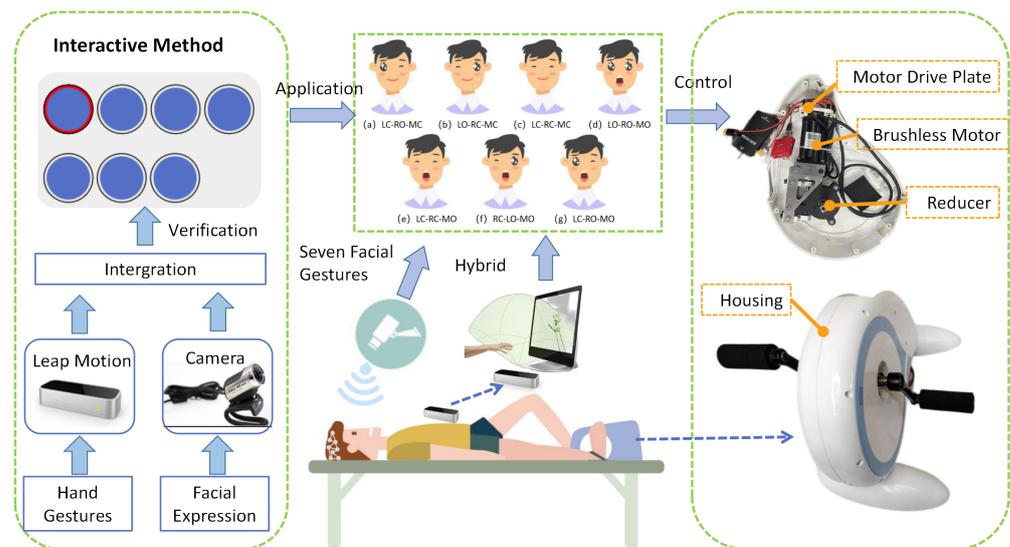


Figure 1. Hybrid Target Selections by “Hand Gestures + Facial Expression” for a Rehabilitation Robot.

The development of face detection technology based on artificial intelligence provides a technical basis for the exploitation of facial expressions. Pilarczyk et al. [3] proposed a computer algorithm based on a CNN/MMD face detector that recognizes human emotions through network cameras. Chu and Peng [7] explored the user identification problems associated with facial recognition technology to unlock mobile devices. Kim and Cha [13] proposed a hands-free natural user interface, which is a facial expression study based on a head-mounted display (HMD) sensor, and users need to wear an HMD device. By contrast, our research does not require users to wear any equipment, thus increasing user comfort versatility. David et al. [9] proposed a FaceSwitch accessibility software system designed to provide facilitated computer interaction for users with compromised upper limb mobility. This software uses a deformable face tracker to track the landmark features of the user’s face. This system enables the user to map a specific facial expression to custom computer control commands that interact with the computer based on a camera and an eye tracker.

Florian et al. [14] proposed a way to trigger interaction with the contours of a medium object using a gaze gesture. Zhang, Harish, and Kulkarni [15] proposed a gaze peak application that could explain eye gestures in real-time. These gestures were decoded into predictive movements and provided the users with a non-denying interface to facilitate user communication. Sun et al. [16] proposed a spatiotemporal fusion model and multi-mode hierarchical fusion strategy. Li et al. [17] proposed an HRI model of a robot arm based on gesture and body motion recognition algorithms. Al-Hammadi et al. [18] proposed an effective deep convolutional neural network method for gesture recognition. Anja and Gebhard [19] conducted a usability study employing head movements and head poses to aid disabled people who cannot use their upper limbs. Garcia et al. [20] used body gestures to navigate interactive maps.

Many researchers have applied gesture to rehabilitation robots [21–24]. Segal et al. [22] presented the design and testing of a gesture-controlled rehabilitation robot (GC-Rebot) to estimate its potential for monitoring user performance and providing entertainment while conducting physical therapy. Michael et al. [25] presented a new gesture-based human interface for natural robot control. RongWen et al. [26] proposed a cooperative surgical robot system, guided by hand gestures and supported by an augmented reality (AR)-based surgical field, for robot-assisted percutaneous treatment. Geng et al. [27] proposed a novel gesture recognition system for intelligent interaction with a nursing-care assistant robot. Marek et al. [28] applied body posture tracking device to rehabilitation robot. Jessica et al. [29] presented the development of an interactive virtual reality (VR)-based framework that allows one to simulate the execution of rehabilitation tasks and

robotic assistance through a robotic standing wheelchair. Fusco et al. [30] study indicates how the combination of robotic treatment with VR is effective in enhancing the recovery of cognitive function in patients with ABI, also improving disability and muscular function. Feng et al. [31] presented a dual-modal hybrid self-switching control strategy automatically determine the exercise mode of patients, i.e., passive and assistive exercise mode. Dong et al. [32] developed three different rehabilitation training methods adapt to the patients at different stages of rehabilitation, namely, passive exercise, active exercise and resistance exercise, respectively. Prieto et al. [33] designed a wearable immersive virtual reality device for promoting physical activity in parkinson's disease patients. Sun et al. [34] proposed a facial emotion recognition system for exoskeleton rehabilitation robot. Bien et al. [35] developed a rehabilitation robotic system with various human-robot interaction interfaces for the disabled.

There are already a number of rehabilitation robots for lower limbs on the market, which can be divided into standing, multi-posture and sitting and horizontal according to different training posture. The standing lower limb rehabilitation robot is mainly divided into two categories, one is the hanging rehabilitation robot represented by Lokomat [36] developed by Hocoma company in Switzerland, the other is the wearing lower limb rehabilitation robot represented by Elegs [37] developed by Berkeley Company in the United States. The multi-type lower limb rehabilitation robot mainly adopts the rehabilitation strategy combining the standing lower limb rehabilitation robot and standing bed. In the aspect of sitting-horizontal rehabilitation robot, the Motion Maker lower limb rehabilitation robot developed by SWORTEC of Switzerland has been put into clinical application [38].

In summary, these technologies have enriched the pool of multi-channel human-computer interactions. The interactive technology in virtual space (meta universe) needs to be explored. However, few study has yet combined facial expressions and hand gestures; thus, inspiring our foray into hybrid gestures technology. The hybrid control gestures technology presented in this study developed a new type of interaction that combines facial expressions and hand gestures. It can receive the simultaneous interaction of hand gestures and facial expressions to perform interactive tasks. The technology provides design guidelines for VR exoskeleton-based interactions.

The accessibility of the hybrid control gestures technique with the following features is explored:

- A hybrid control gestures technique with two input modalities from hand gestures and facial expression is described;
- It is a touchless interactive technique;
- It can be mapped to the interactive mode of the virtual reality environment;
- Facial expressions correspond to basic control functions;
- Hand gestures are detected by Leap Motion, providing a virtual hand technique for cursor movement control.

This paper is organized as follows. Section 2 describes the optimization method, software platform, and hardware equipment of hybrid control gesture technology. In Section 3, the empirical analysis of hybrid control surgery technology is carried out and the application of interactive technology in upper limb rehabilitation robots. In Section 4, the experimental results and design guidelines are discussed. In Section 5, the conclusion is drawn.

2. Materials and Methods

2.1. Interactive Method

2.1.1. Interactive Hardware

A Lenovo computer with the Windows 10 operating system was used in the experiments. A second-generation Leap Motion sensor was used in the experiments to detect hand movements. The Leap Motion sensor was placed 15 cm from the center edge of the monitor base. The external camera, an ANC Core HD 1080p-Y3 webcam, was used to provide a continuous image stream of the subject's face to a face tracking application

described later. It was located at the top of the monitor, 8 cm from the left edge of the display. This distance was only set for consistency of the experimental conditions to reduce deviation resulting from different distances.

2.1.2. Seven Facial Gestures

Seven facial expressions were defined according to the detected facial features, as shown in Figure 2.

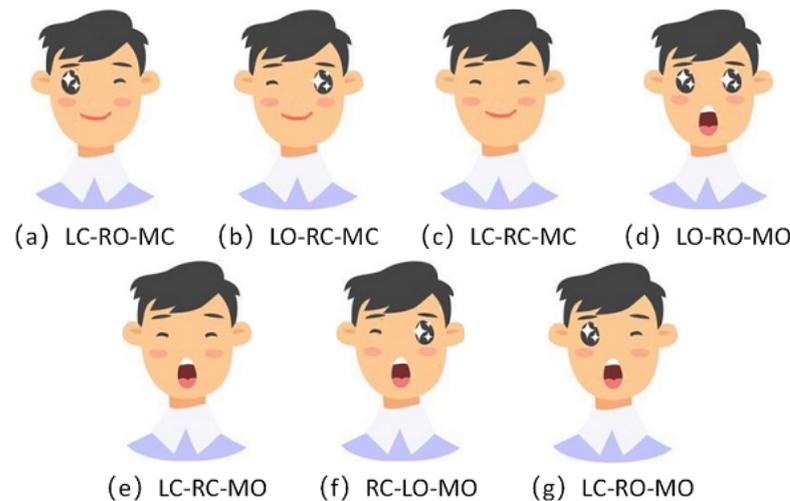


Figure 2. Facial gestures and commands: (a) Closing of the left eye; Opening of the right eye; Closing of the mouth. (b) Opening of the left eye; Closing of the right eye; Closing of the mouth. (c) Closing of the left eye; Closing of the right eye; Closing of the mouth. (d) Opening of the left eye; Opening of the right eye; Opening of the mouth. (e) Closing of the left eye; Closing of the right eye; Opening of the mouth. (f) Closing of the right eye; Opening the left eye; Opening of the mouth. (g) Closing of the left eye; Opening of the right eye; Opening of the mouth.

2.1.3. Software and Method

Hand capture software was used for hand interaction purposes. This software provided gestural information to track the user's hand movements while interacting with the computer. The user activated a task by continuously moving his/her hands. Subsequently, the user could move a cursor using his/her hands at any location on the screen.

A camera was used to track the user's face while interacting with the computer. The camera tracked several specific facial features, such as the opening and closing of the mouth and the opening and closing of the eyes. The facial expression recognition system was used to monitor facial expressions in real-time and to correspond them to specific control commands. Our software provided a friendly and natural way to customize and chart specific gestures to different mouse events.

The interactive zone of Leap Motion was a box. Provided the user's hand or finger was within this box, it could be tracked by Leap Motion. Because there is a certain amount of jitter in the direct control of the mouse movement through Leap Motion, the Kalman filter [39] method is used to control the jitter, and the data collected by Leap Motion and the cursor state data at the previous moment are used to generate the current cursor state through the Kalman filter method [36]. To avoid the influence of noise caused by Leap Motion's sensor noise and hand jitter.

The Kalman filter uses the cursor position state at the previous moment and the instantaneous velocity at the current moment to predict the a priori estimate of the cursor state at the current moment X_t' :

$$X_t' = X_{t-1} + aU_t + W_t \quad (1)$$

X_{t-1} is the cursor position state at the previous moment; U_t is the instantaneous speed of the system; W_t is hand jitter noise; a is the time interval Delta t .

And the process error covariance P_t' formula of the synchronization update is:

$$P_t' = bP_{t-1} + Q_t \quad (2)$$

Q is hand jitter noise variance. Then calculate the Kalman gain K_t :

$$K_t = P_t' / P_t' + R \quad (3)$$

R is the sensor noise variance.

At the same time, the data obtained by Leap Motion is used to modify the value of the prior estimation to achieve the optimal estimation of the cursor state at the current moment X_t :

$$X_t = X_{t'm} + K_t(Z_t - X_{t'}) \quad (4)$$

Z_t is the status data collected by Leap Motion.

The updated observation noise covariance P_t formula is:

$$P_t = (1 - K_t')P_t' \quad (5)$$

Figure 3 describes the operation of the program. The system began to divide into two bifurcations, the hand and the face. In the hand system, first, the Leap Motion sensor is activated by the hand. Obtain hand information, and through Kalman filtering, the mouse cursor can be stably controlled by the movement of the hand. In the face system, the ANC camera obtains facial expressions and performs facial expression recognition through Face++. Participants move the mouse cursor with their hands to select the experimental target. After the target is selected, the system analyzes the facial expression data recognized by Face++, and click the event to complete.

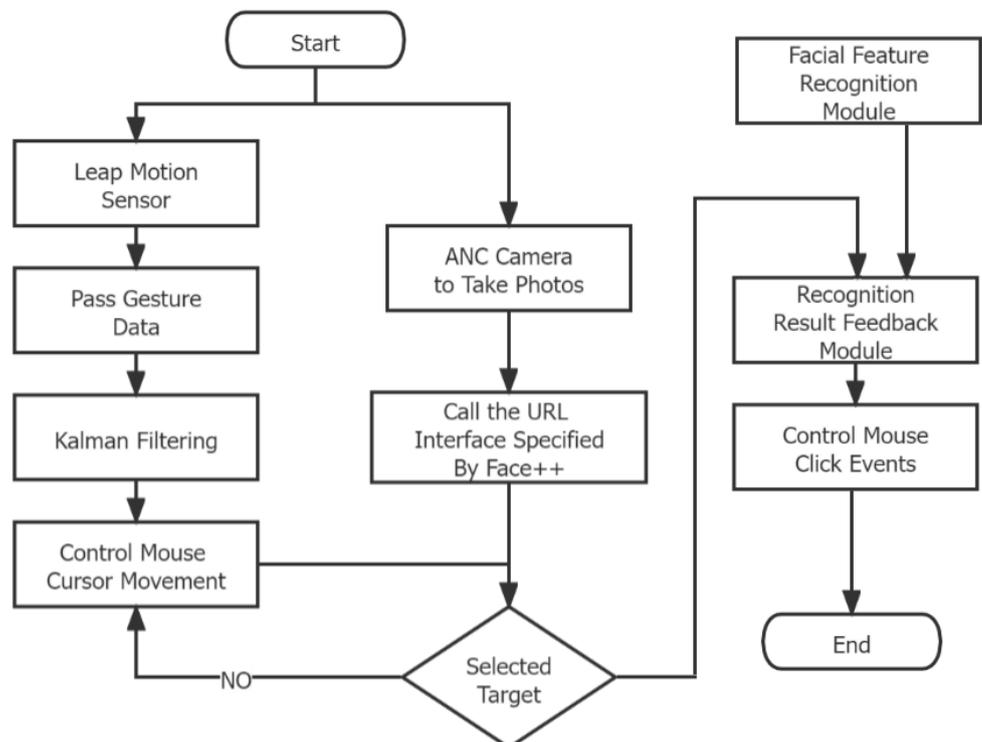


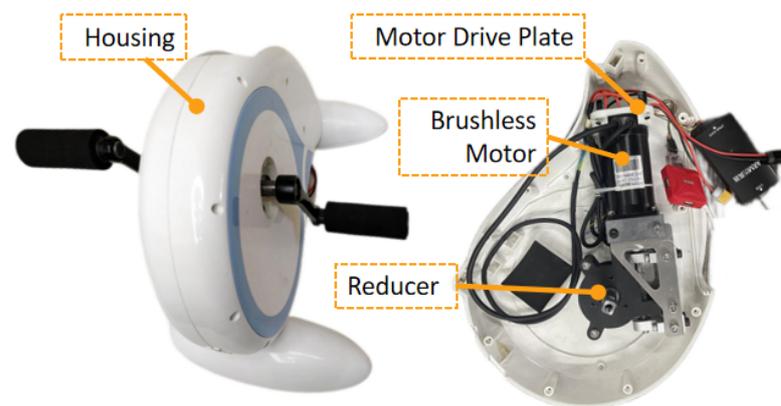
Figure 3. Flow chart of hybrid gesture control.

2.2. Rehabilitation Robot

Using the interaction method proposed above, the present work designed and developed a lower limb rehabilitation robot based on Hand Gestures + Facial Expression, which includes two training modes: active training mode and passive training mode. It is mainly used for bed rehabilitation training of patients with insufficient muscle strength in the early stage of rehabilitation. The present work switched between training modes based on gestures. In the active training mode, the resistance was adjusted according to the facial expression. In passive training mode, adjust the speed based on facial expressions.

2.2.1. Rehabilitation Robot Hardware

The lower limb rehabilitation robot developed at present uses the sensor FOC algorithm to drive the permanent magnet synchronous motor. The main control MCU selects STM32G431RBT6 chip from ST company, and the motor driver selects DRV8323RS three-phase grid driver chip from TI company. The motor Angle is obtained by TLE5012BE1000 magnetic encoder. In addition, it also includes a dissipation circuit module, 24V power supply module, ADC sampling circuit module, and interface module. After the motor drive plate and worm gear reducer was connected, place them in the existing housing of rehabilitation devices, and install the handle and base as shown in Figure 4.



Physical Image of Rehabilitation Robot

Figure 4. The prototype's image of the proposed rehabilitation robot.

2.2.2. Rehabilitation Robot Software

The software of the lower limb rehabilitation robot is divided into a lower computer driver and upper computer control program. The lower computer driver mainly drives the permanent magnet synchronous motor. The development language is C language. The upper computer control program mainly includes the functions of switching training mode, visualizing the training state, training data, etc. The development language is Python, and the third-party libraries used are PyQt5 and pygame. The upper computer control program switches between different training modes and training stops based on gesture, adjusts the resistance based on facial expression in active training mode, and adjusts the rotation speed based on facial expression in passive training mode. There are three gears in the active training mode. The torque of gear A is 15N and activated by LO-RO-MO expression; the torque of gear B is 30N and activated by LC-RC-MO expression; and the torque of gear C is 45N and activated by LC-RC-MC expression. The passive training mode is also divided into three gears. Gear A is activated by LO-RO-MO expression when the speed is 15 rpm, gear B is activated by LC-RC-MO expression when the speed is 30 rpm, and gear C is activated by LC-RC-MC expression when the speed is 50 rpm.

3. Experiments and Results

3.1. Participants

The present work designed and conducted two experiments. Experiment 1 verifies the interaction method proposed above, and experiment 2 verifies the feasibility of applying the proposed new interaction method to the lower limb rehabilitation robot. Ten participants (6 males and 4 females) were recruited from our laboratory for two experimental tests. All participants were in good health and had experience interacting with facial gesture software.

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Zhejiang University (2021-39).

3.2. Task and Procedure

3.2.1. Experiment 1

In this experiment, six sets of circular buttons with differently-sized radii were set. These radii were 10 pixels, 15 pixels, 20 pixels, 25 pixels, 30 pixels, and 35 pixels in size. Each set of target buttons was then placed at three different heights, 6.5 cm, 16 cm, and 26.5 cm, above the target button interface. Height refers to the distance from the top of the target interface to the bottom of the monitor screen. Eighteen sets of tests were completed by placing buttons with different radii at three different heights (i.e., three heights with six radii).

Each participant performed 18 different experiments. In the experiments, target selection and facial recognition were performed according to the order of the buttons from left to right in a sequence. The desired target required a land-on approach, which meant selecting a target directly below the cursor.

The sequence of the facial expressions corresponding to the seven buttons was as follows: LO-RC-MC, LC-RO-MC, RC-LO-MO, LO-RO-MO, LC-RC-MC, LC-RC-MO, LC-RO-MO. (The sequences of the buttons were: from left to right, from top to bottom.)

Before the experiments are performed, the flow of the experiment was explained in detail to the participants. The participants were allowed 10 practice trials.

As shown in Figure 5, when participants moved the cursor by their hand gestures to select the target, a red circle around the target appeared. A peripheral red circle defined a wrong selection. The blue circle itself defined the correct selection area. During the experiment, participants attempted to place the cursor in the blue area of the target button.

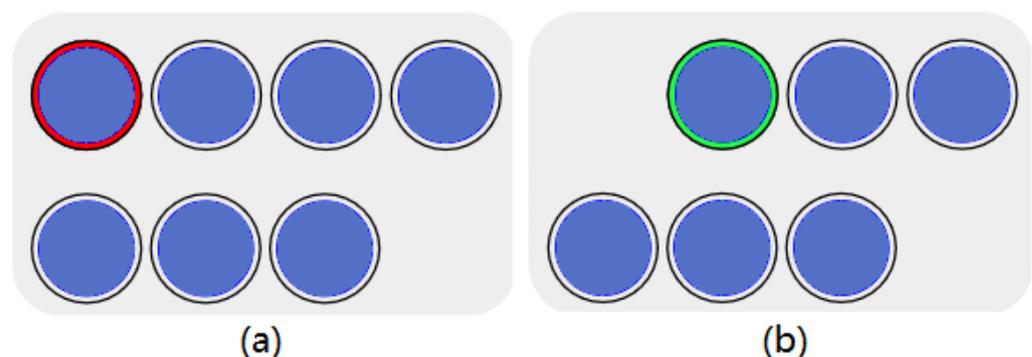


Figure 5. Experimental interface: (a) When the wrong target button is selected, the outer circle of the button is red; (b) The outer circle of the next target button to be recognized is green.

The corresponding facial expression would be recognized only after the correct button was identified by the cursor. After successful recognition, if the cursor selected the red peripheral circle, the backstage data recorded an error in the selection information. When the cursor selected the blue area, the backstage data recorded the correct selection information. The target button disappeared after the cursor selection was completed and the facial

recognition was successful. A green peripheral circle would then surround the subsequent target. This green circle directed the participants to the next target button.

Each participant needed to conduct 18 sets of experiments in the “hands + facial expressions” interactive mode. There were seven target buttons in each set of experiments and each target button corresponded to a facial expression. The sequence of facial gestures corresponding to each target button in each group of experiments was the same.

The overall design of the experiment included the following:

- Ten participants;
- Six target radii;
- Three target heights;
- Seven target buttons (In each group);
- 1260 overall target selections.

Every new interactive technology needs to be measured by human-computer interaction. The proposed work detected the accuracy of target selection through 3 s time pressure (Restrict the target selection task to be completed within 3 s, uncompleted will be regarded as a failure).

3.2.2. Experiment 2

In this experiment, gestures were used to set the switch of two training modes and the emergency stop switch. The left side is the active training mode, and the right side is the passive training mode. As long as gestures are recognized during the training process, emergency stop is indicated. There are three gears in the active training mode. This experiment used three facial expressions to switch gears. The torque of gear A is 15N, which is activated by LO-RO-MO expression, the torque of gear B is 30N, which is activated by LC-RC-MO expression, and the torque of gear C is 45N, which is activated by LC-RC-MC expression. The passive training mode is also divided into three gears. Similarly, used three facial expressions to switch gears. The speed of gear A is 15 rpm, which is activated by LO-RO-MO expression, the speed of gear B is 30 rpm, which is activated by LC-RC-MO expression, and the speed of gear C is 50 rpm, which is activated by LC-RC-MC expression.

The experimental process was divided into two steps. The first step was to use the gesture selection training mode, and the experimenter needed to complete the gesture selection within 3 s. The second step is to use facial expressions to switch training gears or use gestures to stop in an emergency. The specific task of switching gears is shown in Table 1. The proposed work have preset the training mode and gear switching process in advance, and recorded whether the experimenter made the correct choice in real time during the experiment. Experimenter and experiment 1 are the same, each experimenter needs to carry out 10 sets of experiments, each set of experiments includes two mode selection, 12 gear changes in active mode and 12 gear changes in passive mode. Before the experiment began, the participants were given a detailed explanation of the experiment process and were allowed to conduct three practical experiments. The experimental scene is shown in Figure 6.

Table 1. Gear shift task.

Serial No	Gear Shift	Serial No	Gear Shift
1	Shift gear A to stop	7	Shift gear B to stop
2	Shift gear C to stop	8	Shift gear A to B
3	Shift gear B to A	9	Shift gear C to A
4	Shift gear A to C	10	Shift gear B to C
5	Shift gear C to B	11	Stop switching to gear A
6	Stop switching to gear B	12	Stop switching to gear C

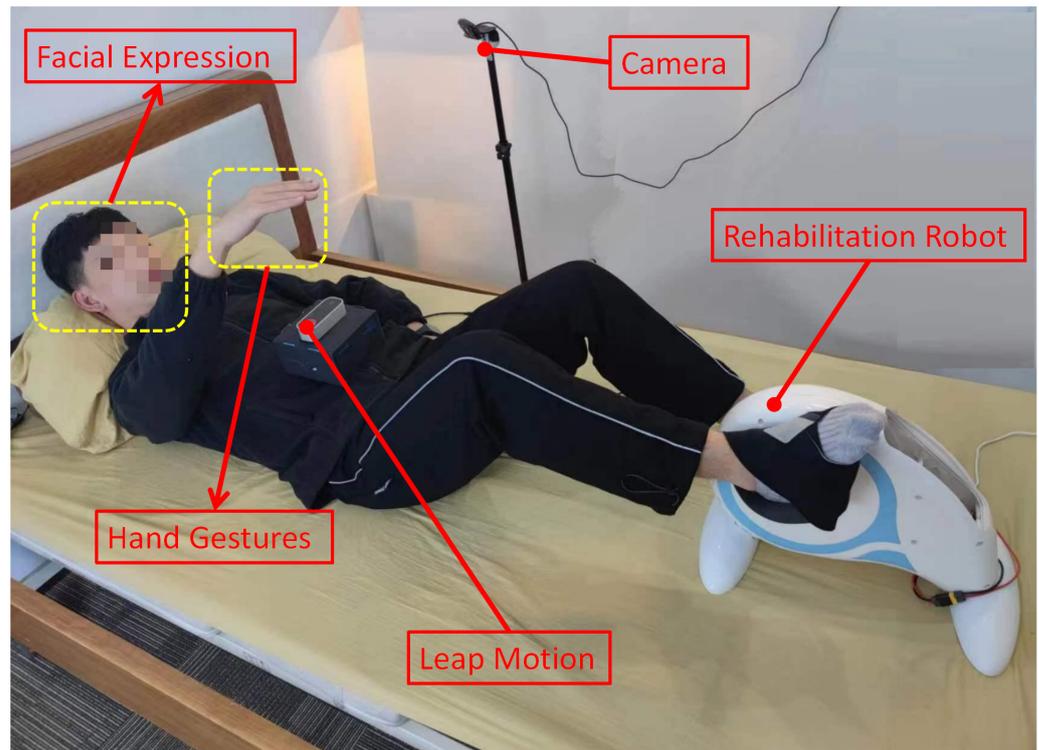


Figure 6. The experimental scene.

3.3. Results

3.3.1. Experiment 1

To analyze the experimental results, target radius (10 pixels, 15 pixels, 20 pixels, 25 pixels, 30 pixels, 35 pixels) and target height (low: 6.5 cm, between 16 cm, high: 26.5 cm) were used as independent variables. A two-way repeated measures ANOVA (Analysis of Variance) was applied to determine the effect of target button radius and target height variation on the accuracy of the participant to perform the hybrid control gesture recognition. Through studentized residuals and a Shapiro-Wilk test, five groups' data were determined not to be normally distributed; the remaining 13 groups' data were normally distributed ($p > 0.05$).

A one-way repeated measures ANOVA was also applied to analyze whether the data were normally distributed, which proved a more robust indicator of deviation from the normal distribution, especially if the sample size of each group was equal or almost equal. Moreover, a non-normal distribution did not substantially impact the probability of making a type I error; therefore, the data could still be tested directly. Determined by studentized residuals and a standard deviation of 3, there were no outliers in any data set. The interaction term height radius ($X^2 = 62.262$, $p = 0.208$) was obtained from the result of Mauchly's sphericity test when it met Mauchly's spherical assumptions; therefore, the variance-covariance matrix of the dependent variable was equal ($p = 0.208$).

The results were represented as the average standard deviation. After analysis, the interaction between height and radius and the difference in the accuracy of hybrid control gesture recognition was not statistically significant ($F_{5,058} = 349.003$, $p = 0.988$, $p = 0.426$). Therefore, the influence of the two subjects' internal factors (height and radius) was validated.

The main effect of target height on hybrid control gestures recognition error rate was not statistically significant ($F_{2,138} = 0.042$, $p = 0.959$).

The effect of the target radius on the accuracy of recognition of the hybrid control gestures was statistically significant ($F_{5,345} = 7.841$, $p < 0.001$). Since the target radius factor had six variables, pairwise comparisons were required. The difference between the hybrid control gestures recognition error rate between target radii of 10 pixels and 15 pixels

was 0.152 (95% confidence interval: 0.054 0.251, $p < 0.001$), which is statistically significant. The difference between the hybrid control gestures recognition error rate of the target radii of 10 pixels and 20 pixels was 0.181 (95% confidence interval: 0.087 0.275, $p < 0.001$), which was statistically significant. The difference in the hybrid control gestures recognition error rate between the target radii of 10 pixels and 25 pixels was 0.186 (95% confidence interval: 0.094 0.277, $p < 0.001$), which was statistically significant. The difference in the hybrid control gestures recognition error rate of target radii of 10 pixels and 30 pixels was 0.214 (95% confidence interval: 0.122 0.307, $p < 0.001$), which was statistically significant. The difference in the hybrid control gestures recognition error rate between target radii of 10 pixels and 35 pixels was 0.210 (95% confidence interval: 0.119 0.300, $p < 0.001$), which was statistically significant. The difference in the hybrid control gestures recognition error rate between target radii of 15 pixels and 30 pixels was 0.062 (95% confidence interval: 0.010 0.114, $p = 0.008$), which was statistically significant. The difference in the hybrid control gestures recognition error rate between target radii of 15 pixels and 35 pixels was 0.057 (95% confidence interval: 0.003 0.112, $p = 0.032 < 0.05$), which was statistically significant. The comparison result of the recognition error rate among other target radii is not statistically significant, for example, the recognition error rate between the target radius of 15 pixels and 20 pixels is not statistically significant. Figure 7 shows the recognition error rates for different target radii at different target heights.

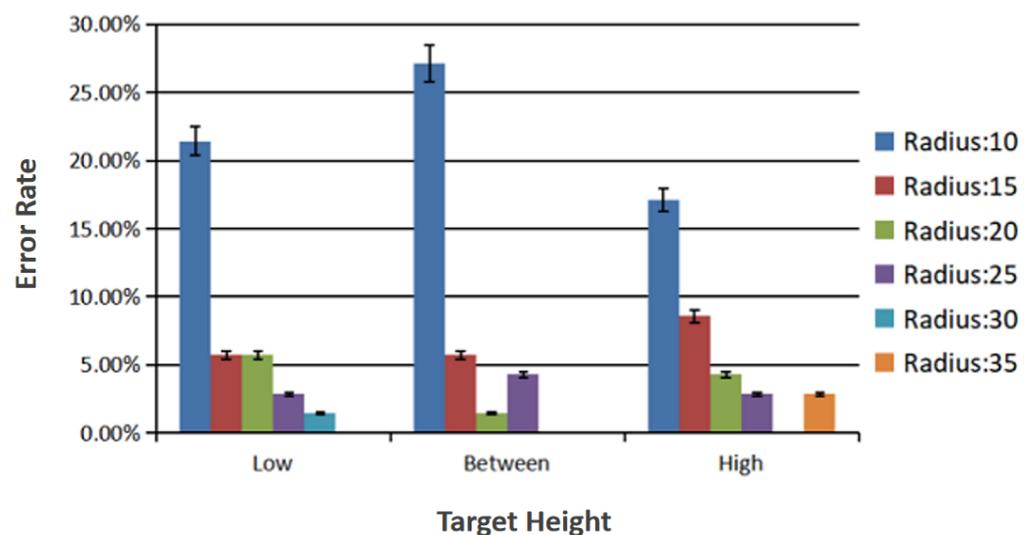


Figure 7. Recognition error rates for different target radii at different target heights.

Figure 8 shows the recognition error rate when the target radius is ignored. The ordinate represents the recognition error rate and the abscissa represents the three different target heights.

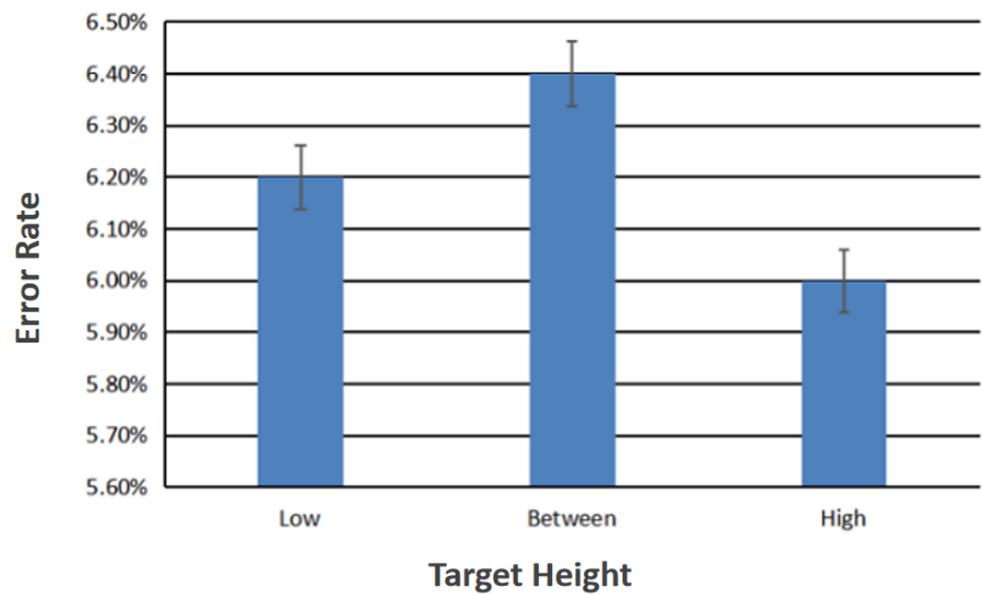


Figure 8. Recognition error rate different target heights.

Figure 9 shows the recognition error rate obtained when the target height is ignored. The ordinate represents the recognition error rate and the abscissa represents the six different target radii.

The data mentioned in this chapter are all accurate values obtained by the two-way repeated measure ANOVA.

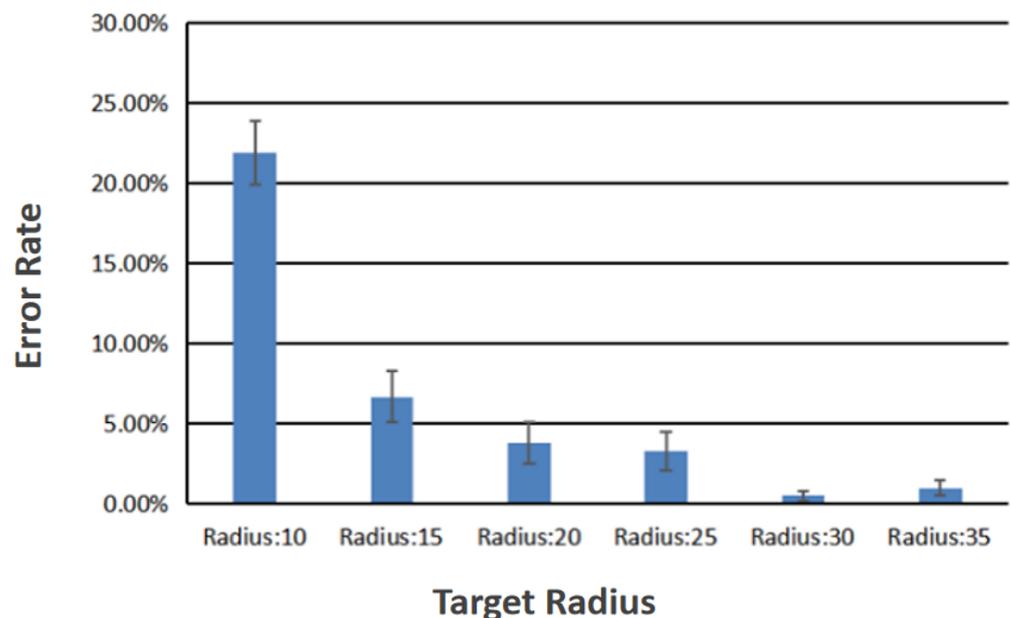


Figure 9. Recognition error rate different target radii.

The above experimental results show that “hand + facial expression” is feasible as an interaction technology. It provides data support for experiment 2 to control rehabilitation robot by gesture and facial expression.

3.3.2. Experiment 2

Experiment 2 is mainly to verify the accuracy of the interactive method control of the rehabilitation robot. This experiment includes 26 subtasks in total, including 2 mode switching subtasks and 24 gear switching subtasks. Each subtask has 100 experiments in

total, and each control subtask will record whether it is correct to calculate the error rate of each subtask.

Figure 10 shows the error rate of each subtask in the active mode. The ordinate identifies the error rate, and the abscissa represents different subtasks.

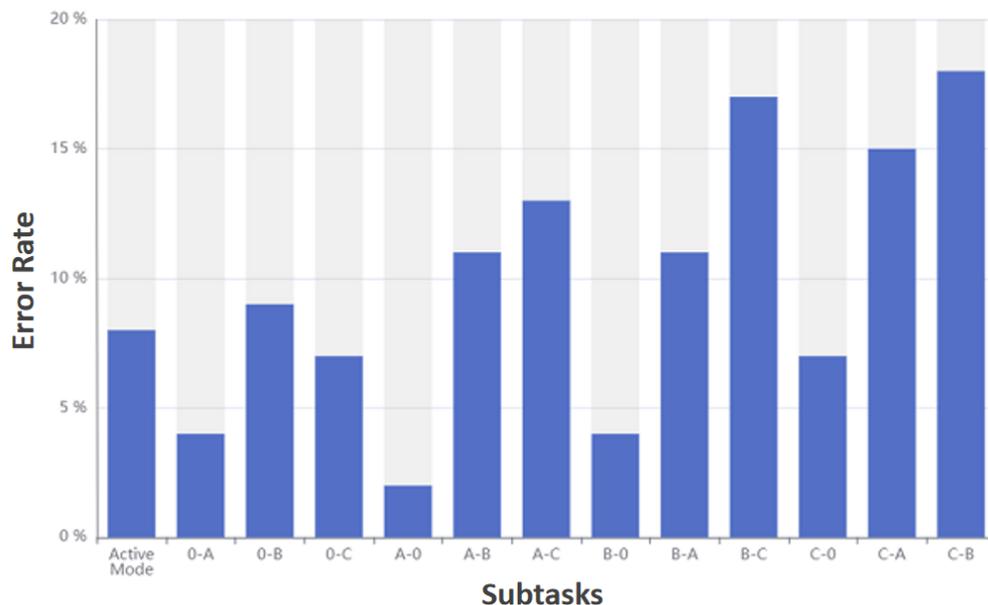


Figure 10. An error rate of each subtask in active mode.

Figure 11 shows error rate of each subtask in the passive mode is displayed. The ordinate identifies the error rate, and the abscissa represents different subtasks.

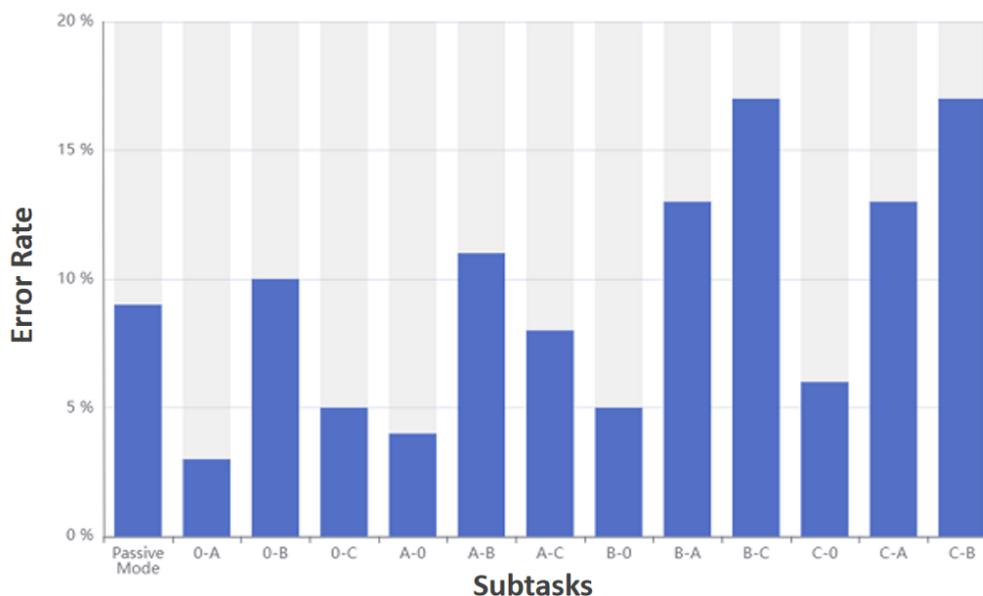


Figure 11. Error rate of each subtask in passive mode.

Experimental results show that both modes have the highest error rate when switching between B and C. The overall error rate is lower than the selection error rate in experiment 1.

4. Discussion

4.1. Data Analysis Conclusion

In this experiment, three target heights and six target radii were tested. The differences in accuracy of the hybrid control gesture recognition under these variables were compared. From the analysis of all the data, we drew the following conclusions:

- As shown in Figure 8, if the target radius is ignored, the error rate of recognition when the target height is limited to “high” is the lowest. The error rate of recognition is the highest when the target height is limited to “between”;
- As shown in Figure 9, if the target height is ignored, the error rate of recognition when the target radius is limited to 30 pixels is the lowest. The error rate of recognition is the highest when the target radius is limited to 10 pixels.
- The interaction between height and radius and the difference in the accuracy of hybrid control gesture recognition was not statistically significant. The target height has no significant effect on the selection error rate. The size of the target radius has a significant impact on the selection error rate.

According to the above conclusions, for the target of the hybrid “hand + facial expression” gestures, the position factor of the target can be ignored, and only the target radius factor needs to be concerned. When the target radius is limited to 30 pixels and above, the interactive pattern recognition rate can reach 93.81%, which further are very robust. The present work verified the error rate of this interaction technology applied to the lower limb rehabilitation robot through experiments. The results show that this interaction technology can be applied to the rehabilitation robots, and the interaction accuracy can be improved.

4.2. Design Guideline

Based on the research and experience in hybrid control gestures, the interactive operation of the simulated virtual environment through the camera and Leap Motion has extremely high reference significance for the hybrid gesture control of the virtual environment. Gestures are the most intuitive body language for pointing operations, and facial expressions are used for auxiliary control. For the gesture control, the stability method of this research can be used to improve the ease of gesture control. The basic operations such as confirmation through simple and easy-to-collect facial expressions can reduce the misoperation of gestures and enhance the ease of use of gesture control. Based on the gesture control, a dimensional operation is added to achieve target selection tasks in more complex scenes.

Hybrid control gesture technology removes physical barriers, and can interact with computers in a non-contact way with the help of VR interactive exoskeletons, and even provides an alternative solution for people with disabilities to perform interactive operations. It also provides a new interactive method for solving the interactive task of target selection in the virtual space.

5. Conclusions

The present work design and propose a hybrid gesture target selection technology which is suitable for multiple scene interaction. The cursor movement is controlled by gestures, and facial expressions are used to confirm the operation (ie, “click”). The sensor simulates the scene to collect gesture and facial expression information and evaluates the influence of target size and target height on selection time and accuracy. The results show that the target height (position) has no significant effect on the error rate and task time, while the target radius has a significant effect on it. When the target width is greater than 30 pixels, it can reach a recognition accuracy of 93.81%, which is very robust. The results of the interaction experiments on rehabilitation robots show that the application of this interaction technology on rehabilitation robots is feasible.

Author Contributions: Conceptualization, T.L.; Software, X.Z. (Xiangliang Zhang); Data curation, Y.H., N.Z. and S.M.; Writing—original draft, X.Z. (Xiangliang Zhang); Writing—review and editing, T.L., S.W., X.Z. (Xiufeng Zhang) and J.Y.; Supervision, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Science Foundation of China awards (U1913601, 52175033 and U21A20120), the National Key R&D Program of China (2020YFC2007800), the Zhejiang Provincial Natural Science Foundation under award LZ20E050002, the Key Research and Development Program of Zhejiang under awards 2022C03103 and 2021C03051.

Institutional Review Board Statement: This study was approved by the Medical Ethics Committee of School of Biomedical Engineering and Instrument Science, Zhejiang University (Project identification code:2021-39)

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors would like to thank the editor and anonymous reviewers for their useful comments for improving the quality of this paper.

Conflicts of Interest: An example of Ethical Statements: The animal protocols used in this work were evaluated and approved by the Animal Use and Ethic Committee (CEUA) of the Institute Pasteur Montevideo (Protocol 2009_1_3284). They are in accordance with FELASA guidelines and the National law for Laboratory Animal Experimentation (Law no. 18.611).

References

1. Ding, I., Jr.; Hsieh, M.C. A hand gesture action-based emotion recognition system by 3D image sensor information derived from Leap Motion sensors for the specific group with restlessness emotion problems. *Microsyst. Technol.* **2020**. [[CrossRef](#)]
2. Li, L.; Mu, X.; Li, S.; Peng, H. A review of face recognition technology. *IEEE Access* **2020**, *8*, 139110–139120. [[CrossRef](#)]
3. Pilarczyk, R.; Chang, X.; Skarbek, W. Human Face Expressions from Images. *Fundam. Informaticae* **2019**, *168*, 287–310. [[CrossRef](#)]
4. Lin, J.; Xiao, L.; Wu, T.; Bian, W. Image set-based face recognition using pose estimation with facial landmarks. *Multimed. Tools Appl.* **2020**, *79*, 19493–19507. [[CrossRef](#)]
5. Mosquera, J.H.; Loaiza, H.; Nope, S.E.; Restrepo, A.D. Identifying facial gestures to emulate a mouse: Navigation application on Facebook. *IEEE Lat. Am. Trans.* **2017**, *15*, 121–128. [[CrossRef](#)]
6. Yan, J.; Lu, G.; Bai, X.; Li, H.; Sun, N.; Liang, R. A novel supervised bimodal emotion recognition approach based on facial expression and body gesture. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2018**, *101*, 2003–2006. [[CrossRef](#)]
7. Chu, C.H.; Peng, S.M. Implementation of Face Recognition for Screen Unlocking on Mobile Device. In Proceedings of the 23rd ACM International Conference on Multimedia, Brisbane, Australia, 26–30 October 2015; pp. 1027–1030.
8. Nagi, J.; Giusti, A.; Di Caro, G.A.; Gambardella, L.M. Human control of UAVs using face pose estimates and hand gestures. In Proceedings of the 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Bielefeld, Germany, 3–6 March 2014; pp. 1–2.
9. Rozado, D.; Niu, J.; Lochner, M. Fast human-computer interaction by combining gaze pointing and face gestures. *ACM Trans. Access. Comput. (TACCESS)* **2017**, *10*, 1–18. [[CrossRef](#)]
10. Feng, Y.; Uchidiuno, U.A.; Zahir, H.R.; George, I.; Park, A.E.; Mentis, H. Comparison of kinect and leap motion for intraoperative image interaction. *Surg. Innov.* **2021**, *28*, 33–40. [[CrossRef](#)]
11. Vysocký, A.; Grushko, S.; Oščádal, P.; Kot, T.; Babjak, J.; Jánoš, R.; Sukop, M.; Bobovský, Z. Analysis of precision and stability of hand tracking with leap motion sensor. *Sensors* **2020**, *20*, 4088. [[CrossRef](#)]
12. Li, H.; Wu, L.; Wang, H.; Han, C.; Quan, W.; Zhao, J. Hand gesture recognition enhancement based on spatial fuzzy matching in leap motion. *IEEE Trans. Ind. Inform.* **2019**, *16*, 1885–1894. [[CrossRef](#)]
13. Kim, J.; Cha, J.; Lee, H.; Kim, S. Hand-free natural user interface for VR HMD with IR based facial gesture tracking sensor. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, 8–10 November 2017; pp. 1–2.
14. Jungwirth, F.; Haslgrübler, M.; Ferscha, A. Contour-guided gaze gestures: Using object contours as visual guidance for triggering interactions. In Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, Warsaw, Poland, 14–17 June 2018; pp. 1–10.
15. Zhang, X.; Kulkarni, H.; Morris, M.R. Smartphone-based gaze gesture communication for people with motor disabilities. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 2878–2889.

16. Sun, B.; Cao, S.; He, J.; Yu, L. Affect recognition from facial movements and body gestures by hierarchical deep spatio-temporal features and fusion strategy. *Neural Netw.* **2018**, *105*, 36–51. [[CrossRef](#)] [[PubMed](#)]
17. Li, X. Human–robot interaction based on gesture and movement recognition. *Signal Process. Image Commun.* **2020**, *81*, 115686. [[CrossRef](#)]
18. Al-Hammadi, M.; Muhammad, G.; Abdul, W.; Alsulaiman, M.; Bencherif, M.A.; Mekhtiche, M.A. Hand gesture recognition for sign language using 3DCNN. *IEEE Access* **2020**, *8*, 79491–79509. [[CrossRef](#)]
19. Jackowski, A.; Gebhard, M. Evaluation of hands-free human-robot interaction using a head gesture based interface. In Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, Vienna, Austria, 6–9 March 2017; pp. 141–142.
20. Guerrero-García, J.; González, C.; Pinto, D. Studying user-defined body gestures for navigating interactive maps. In Proceedings of the XVIII International Conference on Human Computer Interaction, Cancun, Mexico, 25–27 September 2017; pp. 1–4.
21. Segal, A.D.; Lesak, M.C.; Suttora, N.E.; Silverman, A.K.; Petruska, A.J. iRebot: An interactive rehabilitation robot with gesture control. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 5158–5161.
22. Segal, A.D.; Lesak, M.C.; Silverman, A.K.; Petruska, A.J. A Gesture-Controlled Rehabilitation Robot to Improve Engagement and Quantify Movement Performance. *Sensors* **2020**, *20*, 4269. [[CrossRef](#)]
23. Gerlich, L.; Parsons, B.N.; White, A.S.; Prior, S.; Warner, P. Gesture recognition for control of rehabilitation robots. *Cogn. Technol. Work* **2007**, *9*, 189–207. [[CrossRef](#)]
24. Kawarazaki, N.; Hoya, I.; Nishihara, K.; Yoshidome, T. 7 cooperative welfare robot system using hand gesture instructions. In *Advances in Rehabilitation Robotics*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 143–153.
25. Wolf, M.T.; Assad, C.; Vernacchia, M.T.; Fromm, J.; Jethani, H.L. Gesture-based robot control with variable autonomy from the JPL BioSleeve. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 1160–1165.
26. Wen, R.; Tay, W.L.; Nguyen, B.P.; Chng, C.B.; Chui, C.K. Hand gesture guided robot-assisted surgery based on a direct augmented reality interface. *Comput. Methods Programs Biomed.* **2014**, *116*, 68–80. [[CrossRef](#)]
27. Yang, G.; Lv, H.; Chen, F.; Pang, Z.; Wang, J.; Yang, H.; Zhang, J. A novel gesture recognition system for intelligent interaction with a nursing-care assistant robot. *Appl. Sci.* **2018**, *8*, 2349. [[CrossRef](#)]
28. Sierotowicz, M.; Connan, M.; Castellini, C. Human-in-the-loop assessment of an ultralight, low-cost body posture tracking device. *Sensors* **2020**, *20*, 890. [[CrossRef](#)]
29. Ortiz, J.S.; Palacios-Navarro, G.; Andaluz, V.H.; Guevara, B.S. Virtual reality-based framework to simulate control algorithms for robotic assistance and rehabilitation tasks through a standing wheelchair. *Sensors* **2021**, *21*, 5083. [[CrossRef](#)]
30. Fusco, A.; Giovannini, S.; Castelli, L.; Coraci, D.; Gatto, D.M.; Reale, G.; Pastorino, R.; Padua, L. Virtual Reality and Lower Limb Rehabilitation: Effects on Motor and Cognitive Outcome—A Crossover Pilot Study. *J. Clin. Med.* **2022**, *11*, 2300. [[CrossRef](#)]
31. Feng, G.; Zhang, J.; Zuo, G.; Li, M.; Jiang, D.; Yang, L. Dual-Modal Hybrid Control for an Upper-Limb Rehabilitation Robot. *Machines* **2022**, *10*, 324. [[CrossRef](#)]
32. Dong, M.; Yuan, J.; Li, J. A Lower Limb Rehabilitation Robot with Rigid-Flexible Characteristics and Multi-Mode Exercises. *Machines* **2022**, *10*, 918. [[CrossRef](#)]
33. Campo-Prieto, P.; Cancela-Carral, J.M.; Rodríguez-Fuentes, G. Wearable Immersive Virtual Reality Device for Promoting Physical Activity in Parkinson’s Disease Patients. *Sensors* **2022**, *22*, 3302. [[CrossRef](#)]
34. Sun, W.; Peng, H.; Liu, Q.; Guo, Z.; Ibrah, O.O.; Wu, F.; Li, L. Research on Facial Emotion Recognition System Based on Exoskeleton Rehabilitation Robot. In Proceedings of the 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 16–18 October 2020; pp. 481–484.
35. Bien, Z.; Kim, D.J.; Chung, M.J.; Kwon, D.S.; Chang, P.H. Development of a wheelchair-based rehabilitation robotic system (KARES II) with various human-robot interaction interfaces for the disabled. In Proceedings of the 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), Monterey, CA, USA, 20–24 July 2003; Volume 2, pp. 902–907.
36. Chaparro-Rico, B.D.; Cafolla, D.; Tortola, P.; Galardi, G. Assessing stiffness, joint torque and ROM for paretic and non-paretic lower limbs during the subacute phase of stroke using lokomat tools. *Appl. Sci.* **2020**, *10*, 6168. [[CrossRef](#)]
37. Díaz, I.; Gil, J.J.; Sánchez, E. Lower-limb robotic rehabilitation: Literature review and challenges. *J. Robot.* **2011**, *2011*, 759764. [[CrossRef](#)]
38. Hu, W.; Li, G.; Sun, Y.; Jiang, G.; Kong, J.; Ju, Z.; Jiang, D. A review of upper and lower limb rehabilitation training robot. In Proceedings of the International Conference on Intelligent Robotics and Applications, Wuhan, China, 16–18 August 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 570–580.
39. Kalman, R.E.; Bucy, R.S. New results in linear filtering and prediction theory. *J. Basic Eng.* **1961**, *83*, 95–108. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.