



Article

A Semi-Autonomous Hierarchical Control Framework for Prosthetic Hands Inspired by Dual Streams of Human

Xuanyi Zhou ^{1,2,3} , Jianhua Zhang ⁴, Bangchu Yang ^{1,2}, Xiaolong Ma ^{1,2} , Hao Fu ^{1,2}, Shibo Cai ^{1,2} and Guanjun Bao ^{1,2,*}

- ¹ College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China; zhouxuanyi@zjut.edu.cn (X.Z.); mxslf@126.com (X.M.); fhneversettle@163.com (H.F.); ccc@zjut.edu.cn (S.C.)
- ² Key Laboratory of Special Purpose Equipment and Advanced Processing Technology, Ministry of Education and Zhejiang Province, Zhejiang University of Technology, Hangzhou 310023, China
- ³ State Key Laboratory of Chemical Engineering, College of Chemical and Biological Engineering, Zhejiang University, 38 Zheda Road, Hangzhou 310027, China
- ⁴ School of Mechanical Engineering, Beijing University of Science and Technology, Beijing 100083, China; jhzhang@ustb.edu.cn
- * Correspondence: gjbao@zjut.edu.cn
- † Member of IEEE.

Abstract: The routine use of prosthetic hands significantly enhances amputees' daily lives, yet it often introduces cognitive load and reduces reaction speed. To address this issue, we introduce a wearable semi-autonomous hierarchical control framework tailored for amputees. Drawing inspiration from the visual processing stream in humans, a fully autonomous bionic controller is integrated into the prosthetic hand control system to offload cognitive burden, complemented by a Human-in-the-Loop (HIL) control method. In the ventral-stream phase, the controller integrates multi-modal information from the user's hand-eye coordination and biological instincts to analyze the user's movement intention and manipulate primitive switches in the variable domain of view. Transitioning to the dorsal-stream phase, precise force control is attained through the HIL control strategy, combining feedback from the prosthetic hand's sensors and the user's electromyographic (EMG) signals. The effectiveness of the proposed interface is demonstrated by the experimental results. Our approach presents a more effective method of interaction between a robotic control system and the human.

Keywords: prosthetic hand; control strategy; grasp; manipulation; human inspired



Citation: Zhou, X.; Zhang, J.; Yang, B.; Ma, X.; Fu, H.; Cai, S.; Bao, G. A Semi-Autonomous Hierarchical Control Framework for Prosthetic Hands Inspired by Dual Streams of Human. *Biomimetics* **2024**, *9*, 62. <https://doi.org/10.3390/biomimetics9010062>

Academic Editors: Raphael Milanezi de Andrade and Xiaojie Wang

Received: 11 December 2023
Revised: 3 January 2024
Accepted: 6 January 2024
Published: 22 January 2024



Copyright: © 2024 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

It is anticipated that future humanoid robots will perform various complex tasks through communication with human users [1]. Despite significant advancements in prosthetics technology in recent years, only 50 to 60 percent of amputees are willing to wear prosthetics [2], with rejection rates as high as 40 percent [3]. Brain dynamics experiments may provide insight into why this “prosthetic rejection” occurs [4]. Since prosthetic hands are less comfortable than human hands, using them imposes great cognitive burden on amputees, leading to brain fatigue and psychological frustration that ultimately results in prosthetic hand rejection [5].

Recent studies shed new light on the concept of “cognitive load” when using prosthetic hands. Prolonged use of prosthetic hands while handling tools can increase the strength of (electroencephalographic) EEG alpha waves in the brain [6]. This process can lead to an increase in cognitive load, resulting in fatigue and reduced responsiveness to other objects [7], which is the key reason for the high rejection rate of prostheses. The significant increase in EEG alpha-wave power indicates that users consciously exert greater control over their prosthetic hands [8]. This phenomenon is not observed in ordinary individuals when using their hands naturally and skillfully. User experience research reports confirm

the existence of “cognitive load” [9]. Equipping prosthetic hands with sensor feedback to reduce the visual dependence of amputees is a promising solution to alleviate cognitive load [10].

An inspiration for reducing the “cognitive load” is the ventral stream and dorsal stream, which are two visual systems hypotheses of the human brain. Recently, evidence was provided for a functional segregation of dorsal and ventral streams, supporting the hypothesis [11]. Anatomical studies have proved the interaction between ventral and dorsal streams, especially for skilled grasping [12]. As the demanding precision of the grasp increases, these physiological interconnections gradually become more active [13,14].

The ventral stream, known as the “what pathway,” is primarily responsible for object recognition and perception. The dorsal stream, referred to as the “where pathway,” is primarily involved in the processing of spatial awareness and movement guidance [15]. The primary function of the dual stream is shown in Table 1. In the context of controlling prosthetic hands, the ventral stream can be utilized to extract information about the user’s hand–eye coordination and their viewing field. By analyzing the information, the controller can determine the intention behind the user’s movements and enable the manipulation of primitive switches in the variable domain of view. The dorsal stream plays a crucial role in integrating visual information with motor control. In the context of controlling prosthetic hands, the dorsal stream enables precise force control. Moreover, the dorsal stream needs to obtain detailed information about the identity of the object stored in the ventral-stream region when object attributes require complex fine-tuning of the grasp. Correspondingly, the ventral stream may obtain the latest grab-relevant information from the dorsal-stream region to improve the internal representation of the object. Based on the hypothesis, incorporating both the ventral- and dorsal-stream principles into the control framework of prosthetic hands makes it possible to create a more sophisticated and intuitive control system. The ventral stream contributes to the perception and recognition of objects, while the dorsal stream facilitates movement guidance and precise force control. This framework should improve the performance of prosthetic hands and user experience.

Table 1. The primary function of the dual stream.

	The Ventral Stream	The Dorsal Stream
Function	Identification	Visually guided movement
Sensitive features	High sensitivity to spatial	High sensitivity to time
Memory features	Long-term memory	Short-term memory
Reaction speed	Slow	Quick
Comprehension	Very fast	Very slow
Reference frame	Object-Centric	Human-Centric
Visual input	Fovea or parafovea	Entire retina

2. Related Work

Prosthetic hands primarily rely on EMG signals for control [16,17]. Although this method effectively utilizes residual muscles in the amputated limb, the absence of corresponding tactile feedback necessitates users to rely on visual compensation, resulting in an increased cognitive load. This phenomenon has emerged as a significant factor in prosthetic rejection among amputees. To address this issue, researchers have developed a gaze-training method known as Gazing Training. This method assists amputees in adapting to prosthetic hand usage by reducing the need for conscious control and alleviating cognitive load [18]. However, it is important to note that while Gazing Training demonstrates partial success in mitigating cognitive load during the rehabilitation process, it does not completely eliminate the underlying challenge [19].

Therefore, some researchers aspire to combine intelligent autonomous control methods with human users’ electromyographic (EMG) signals by integrating control techniques from the field of robotics [20,21]. This integration aims to establish a semi-autonomous controller that combines autonomous control capabilities with human EMG signals. The

objective of this semi-autonomous control approach is to partially or even entirely transfer the cognitive load to the controller, enabling autonomous assistance for amputee patients in accomplishing daily tasks. With advancements in both the understanding of the human brain and computer technology in recent years, this is possibility gradually transforming into reality [22].

In 2015, Markovic proposed a framework for a semi-autonomous controller, which consists of an autonomous control unit and an EMG control unit [23]. The autonomous control unit employs computer vision sensors to capture depth and red green blue (RGB) information, as well as proprioceptive feedback from the prosthetic hand, for data processing and information fusion. On the other hand, the EMG control unit utilizes electromyographic signals to reflect user intentions and facilitate manual control of the prosthetic hand. The combination of the EMG control unit and the autonomous control unit forms a semi-autonomous controller that integrates autonomous control capabilities with human EMG signals, allowing for switching between control modes [24].

Bu proposed a visually guided approach to assist patients in achieving semi-autonomous manipulation of prosthetic hands, with the primary goal of alleviating users' cognitive load [25]. Chunyuan introduced global visual information based on EMG signals [26]. Machine vision was employed to extract object shape, size, type, and appearance information, which was then integrated with pre-hand shape for joint motion planning using visual and dual antagonistic-channel EMG signals. Wang proposed an RNN network by incorporating features of the user's gaze point [27]. The semi-autonomous control of the prosthetic hand is achieved by the automatic recognition of the used tools and motion primitives. The integration of computer vision into the semi-autonomous control of prosthetic hands imposed significant computation demands on the controller [28]. To address the issue of computational load in prosthetic hand operations, Fukuda proposed a distributed control system to enhance the real-time performance of the semi-autonomous controller [29]. Moreover, the inertial measurement unit (IMU) was employed to obtain the prosthetic hand's state, while the visual system was utilized to perceive object states [30]. Vorobev proposed a semi-autonomous control method for prosthetic hands. Motion commands were transmitted to the prosthetic hand's main controller using sensors triggered by foot movement inside the shoe [31].

In summary, addressing the issue of cognitive load, the Gazing Training approach is proposed as a rehabilitation training method for amputees to adapt to prosthetic hands [32]. The Gazing Training partially mitigates the cognitive load but does not completely eliminate the problem. In recent years, some researchers have proposed a semi-autonomous control strategy from the perspective of robotics engineering. In these approaches, the controller acquires information about object shape, size, etc., and translates it into corresponding motion primitives, while the user's EMG signals generate grasping commands for the prosthetic hand. This semi-autonomous hierarchical control strategy aims to transfer a portion of the user's cognitive load to the controller, thereby reducing the user's alpha-wave power. Building upon the foundation of traditional semi-autonomous controllers, this paper primarily presents two optimizations to the semi-autonomous hierarchical control strategy for prosthetic hands. The innovations of the paper are summarized as follows:

- (1) A controller is constructed based on the pathway of the human ventral–dorsal nerves. Object semantic segmentation and convolutional neural network (CNN) recognition are categorized as the ventral stream, while the motion tracking of the limb is introduced as the control of the dorsal stream. Moreover, the dorsal stream and ventral stream are integrated to ensure accurate motion primitives.
- (2) In order to reduce the cognitive burden, a semi-autonomous controller is proposed. Feedback of the prosthetic hand is integrated to enhance the perceived experience. EMG signals of the user are obtained to realize the human in the loop control.

3. Methodology

In the guidance stage of the human visual neural stream, the perception-motion guidance is carried out by the ventral stream to recognize and locate objects. The ventral-stream information is matched to the related object memory and long-term action primitives. Inspired by the ventral stream, the prosthetic hand control system is applied to locate the object. Moreover, the initial user intentions are obtained through the posture of the grasping task. The visual information of objects can be further obtained through the CMOS image sensor of the head-mounted device. Since the CMOS image sensor in the headset follows the vision of the human user, the headset is able to locate the objects and prosthetic hands.

In the human dorsal-stream guidance stage, the human user guides the prosthetic hand to grasp and manipulate in a vision-aided manner. At the same time, the CMOS image sensor of the headset will obtain the real-time distance between the prosthetic hand and the object. When the prosthetic hand approaches the object, the controller will drive the prosthetic hand to perform grasp motions with the EMG signal of the human body simultaneously. At this stage, humans mainly use the information of the dorsal neural stream for guidance. Based on the characteristics of dorsal-stream information, the position and force cloud of the grasping process are fed back to the human user. Finally, the EMG signals of the human are combined to realize the human in the loop control. The framework of two nerve streams of visual information is shown in Figure 1.

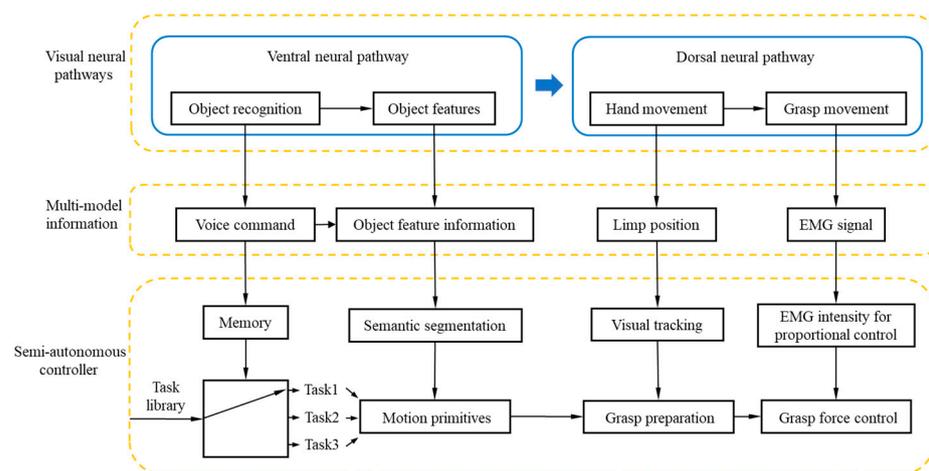


Figure 1. The framework of two nerve streams of visual information.

3.1. Task-Centric Planning

The task planning is divided into two parts: motion primitive and sequence planning. The controller is proposed in the task-centric task, which is inspired by the two visual streams, realizing the sets and independent planning control of prosthetic hand motion primitives. Firstly, multiple motion primitives are stored in the semi-autonomous controller. Different motion primitives are applied in the semi-autonomous controller for the grasping features of different objects. The object's category and the prosthetic hand's configuration are input into the controller as decision factors. The controller drives the prosthetic hand to complete the grasping task. In the decision-making stage, it is necessary to analyze the object character and obtain the location information in the scene. The decision is calculated by the convolutional neural network and semantic analysis. In order to realize real-time control, the SSD-Mobilenet-V2 convolutional neural network is used to realize semantic segmentation and object recognition for visual images. SSD-Mobilenet-V2 is used for semantic segmentation while Mobilenet is used for object recognition. The Single-Shot Multi-box Detector (SSD) neural network generates constraint squares of fixed size through the forward-propagation CNN network, compared to the intersection over union (IOU) of different Anchor boxes generated on the feature graph, obtaining boundary boxes close to 0.5.

Images captured from CMOS image sensors are segmented and associated with semantics. Specifically, after obtaining the boundary boxes through the SSD method mentioned above, a mobilenet-V2 convolution network is introduced to identify objects in the boundary boxes and generate index information. To improve the operational efficiency of the neural network for embedded devices, the mobilenet-V2 convolutional network is adopted to simplify the calculation process through a deeply separable convolutional operation method. The main advantage of this method is that Linear Bottleneck replaces ReLU to activate a function, achieving channel reductions depth wise and dimension reductions point wise, respectively. At the same time, to reduce the effect of feature reduction caused by linear bottleneck dimension reduction, the addition and nonlinearity of features are realized by inverted residuals. Similarly, an inverted residual neural network has the property of “short-circuit”. The procedure has been simplified as the lightweight software requirement of mobile devices.

In different scenes, objects will correspond to a different set of motion primitives. Taking three grasping and manipulating tasks as examples, the varieties of manipulated primitives involved in this paper are shown in Table 2.

Table 2. The sets of motion primitives.

	Toggle Switch	Screw Cap	Grasp Cap
The rest gesture	gesture 1	gesture 1	gesture 1
Pre-shape gesture	gesture 12	Prepare and pre-envelope	Prepare and pre-envelope
Manipulate gesture	gesture 12	gesture 2	gesture 2

For different grasp and manipulate tasks, motion primitives are required for the different manipulate time sequence planning. Gesture 1 is no contact and no motion of the hand. Gesture 2 is motion of the hand without and within the hand. Gesture 12 is the within-hand movement [33]. To ensure the suitable sequence of motion primitive switching, sequential planning uses the end state of the previous primitive as the beginning state for the next primitive. Considering a common motion in daily life, “object grasping”, the sequence of manipulation can be planned as follows:

- (1) The beginning of the task. The initial motion primitive is in the free state, which is gesture 1.
- (2) Object recognition stage. When the target recognition is completed, the prosthetic hand forms the pre-grasp posture according to the feature information of the object. If the target image is lost, it is estimated that the user gives up grasping, and the prosthetic hand restores to the free primitive state.
- (3) Pre-grasp stage. The head-mounted CMOS image sensor obtains the spatial position relationship between the prosthetic hand and the object in real time. When the space distance between the hand and object is less than the threshold value, it is judged that the hand and object are in contact and enter the grasping primitive stage. When the space distance between the hand and object is greater than the threshold value, the hand and object are considered to be separated and return to the pre-hand type stage. The user can change his view field to restore the original free primitive at this stage.
- (4) Grasp stage. Control is performed on the prosthetic hand after fusing EMG signal and object vision information feedback.
- (5) After grasping, the prosthetic hand ends the grasping task and returns to the initial stage. After completing the grasping task, the prosthetic hand is controlled to separate from the object and restored to the initial free primitive state, changing the view field of the CMOS sensor. In case of unsuccessful separation of the hand and object or an emergency, voice command can perform an emergency reset. We restrict the prosthetic hand from switching directly to the free primitive state when performing grasp primitives for the user’s safety. The aforementioned manipulate sequence planning for the controller is shown in Figure 2.

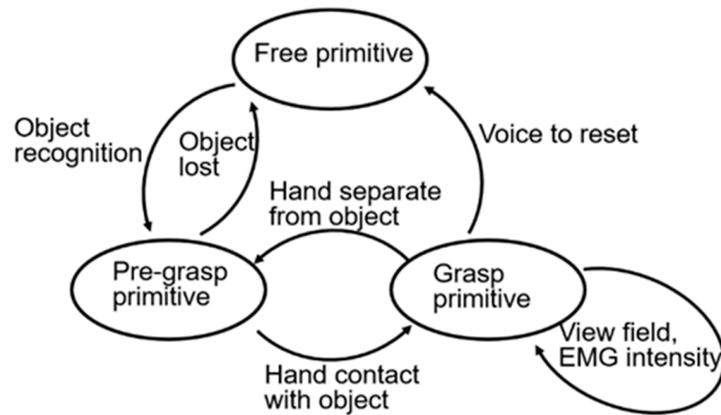


Figure 2. Manipulate sequence planning for the controller.

3.2. Precise Force Control Strategy

The precise force control of the prosthetic hand during contact is proposed for object grasping. In the grasping process, human users are able to participate in the control loop, forming the human-in-loop control mode. The pressure cloud image is shown in the headset. With the pressure feedback, the human users can adjust the grasping force to realize closed-loop control between the human and the prosthetic hand.

Human signals are generated by the flexor digitorum profundus. Signal preprocessing is carried out, including signal amplification, peak-to-peak detection, envelope processing, mean filtering, Fourier transform, and bandpass filtering. The value of the EMG signal is obtained in real time.

The STM32 microcontroller is applied as the central semi-autonomous controller. An Arduino microcontroller was used to obtain EMG values of flexor digitorum Profundo in real time. These two microcontrollers communicate through a serial port and transmit data in hexadecimal format. The central semi-autonomous controller will request for the intensity of the EMG signals obtained by the Arduino controller in real time. Considering the diversity of tasks, environments, and personal behaviors in grasping, the semi-autonomous controller will continuously detect the action potential of the user’s digital-flexor deep muscle. A typical original EMG signal and bandpass filtering EMG intensity are shown in Figure 3. The proportional control output u will be obtained when the EMG intensity P exceeds the threshold, where k_1 is the proportional coefficient.

$$u = k_1 P \tag{1}$$

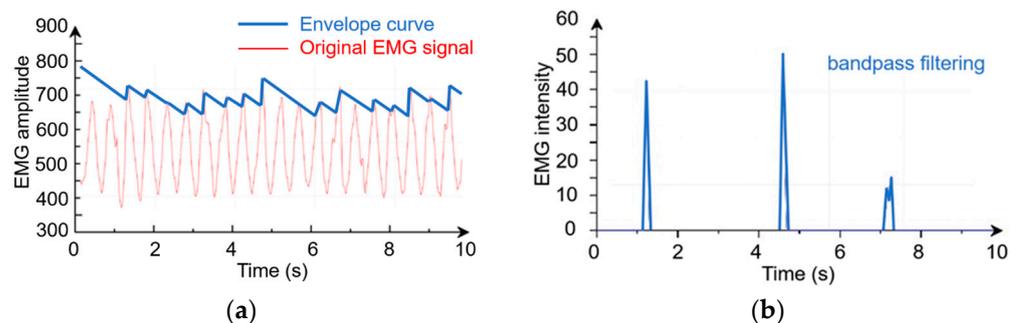


Figure 3. The myoelectric signal. (a) Original EMG signal. (b) Bandpass filtering.

The control output value u of the semi-autonomous controller is transmitted to the prosthetic hand system through Bluetooth. The prosthetic hand system converts the proportional control output u into pulse-width modulation (PWM) signal. It controls five servo motors to realize the movement of the finger joint of the prosthetic hand. Pressure sensors are embedded in each finger to read the contact state between the prosthetic

hand and the object. The haptic pressure distribution cloud is transmitted to a semi-autonomous controller via Bluetooth built into the prosthetic hand system. To tackle the challenge of providing grip force feedback in prosthetic hands, this study leverages computer graphics techniques utilizing the open-source graphics library OpenGL. Within the semi-autonomous controller, a pressure cloud map of the prosthetic hand is generated. This map is subsequently projected onto the user's retina through a head-mounted device, facilitating closed-loop control. The fundamental approach involves the transmission of pressure sensor information from the prosthetic hand system to the semi-autonomous controller via the embedded Bluetooth of the microcontroller unit. The semi-autonomous controller, in turn, employs OpenGL to render the pressure cloud map. The combined controller information and prosthetic hand data are then projected onto the user's retina using the micro display of the head-mounted device.

Users will dynamically adjust the incremental control output of the prosthetic hand system based on real-time feedback from the controller and visual perception. According to Equation (1), once the proportional control output u is generated, the semi-autonomous controller will continuously monitor the user's electromyographic signals. If the user perceives that a stable grasp has not yet been achieved, the proportional control output u will persistently accumulate control output increments.

$$u_i = u_{i-1} + k_2 P \quad (2)$$

where k_2 is the incremental coefficient, which plays a role in amplifying and shrinking the signal during the incremental proportional control. The procedures of the proportional control mode are shown in Figure 4.

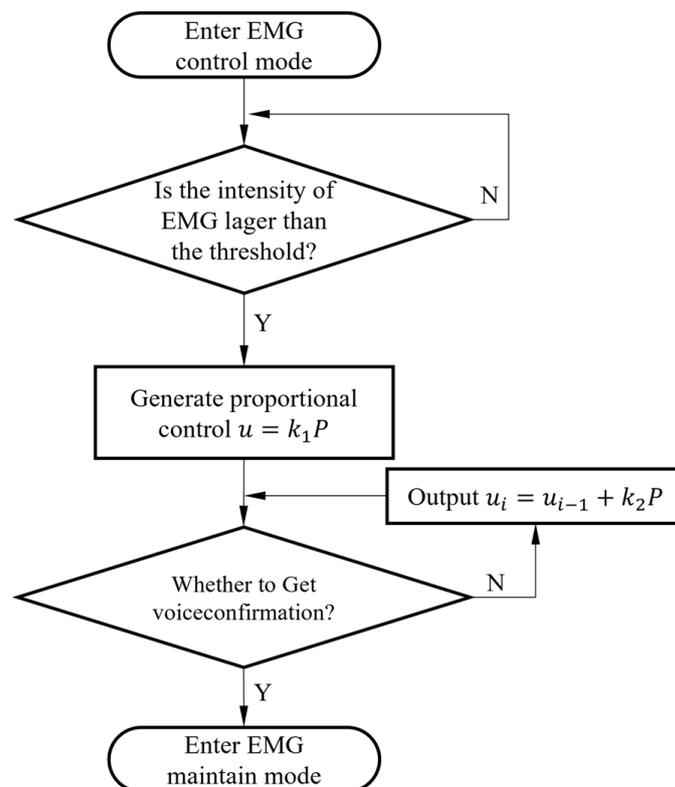


Figure 4. The procedures of the proportional control mode.

To solve the grasping force acquisition of prosthetic hands, we draw a pressure cloud map of prosthetic hands in the semi-autonomous controller, using computer graphics technology (OpenGL). This process costs less in terms of CPU resources. The pressure cloud figure is displayed in a headset and projected onto the user's retina. In this way, the

human brain is connected to the control loop of the prosthetic hand. Rendering the pressure cloud image in the semi-autonomous controller mainly uses the GLWight library under the Qt framework; the generated rendering image is transmitted to the head-mounted device through an HDMI cable. The micro-display device in the headset uses a Vufine+ wearable display for visual projection.

OpenGL’s geometric shader function interface is a four-channel array representing R (red region), G (green region), B (blue region), and α (transparency component). The following formula will convert the contact pressure information transmitted into the OpenGL into a cloud map. The pressure cloud figure will be displayed on the head mount display. When the contact pressure value c_i is less than 0.5-times the contact pressure threshold L , the data relationship of contact pressure information to the cloud map is shown in Formula (3):

$$\begin{cases} R_i = 0 \\ G_i = \frac{255c_i}{0.5L} \\ B_i = 255 - \frac{255c_i}{0.5L} \end{cases} \quad (3)$$

When the contact pressure value c_i is larger than 0.5-times of the contact pressure threshold L , the data relationship of contact pressure information to the cloud map is shown in Formula (4):

$$\begin{cases} R_i = \frac{255c_i}{0.5L} \\ G_i = 255 - \frac{255c_i}{0.5L} \\ B_i = 0 \end{cases} \quad (4)$$

where c_i is the contact force value of the i th finger; L is the contact force threshold; R_i, G_i, B_i represent the color components of the i th finger in red, green, and blue cloud images, respectively.

Tailored to specific grasping or operational tasks, the semi-autonomous controller dynamically monitors the real-time electromyographic signal strength of the user’s deep flexor muscles. It triggers the electromyographic control phase when this intensity surpasses a predefined threshold. Considering that deep flexor muscle signals originate from the deeper muscle groups of the human body and surface electrode signal acquisition is susceptible to interference, various techniques, including differential amplification, peak detection, envelope processing, mean filtering, Fourier transformation, and bandpass filtering, are employed to extract real-time electromyographic signal intensity values. These intensity values are communicated in real time via serial communication to the STM32 embedded in the semi-autonomous controller for proportional control.

When the prosthetic hand contacts the target object, the contact force value and force distribution position of the tactile sensor embedded in the prosthetic hand are changed. This tactile information will be processed by the processor embedded in the prosthetic hand and then transmitted to the semi-autonomous controller, which uses OpenGL to draw the pressure cloud. A micro-display on the headset projects the pressure cloud image onto the user’s retina, enabling the user to know the motion state and contact force distribution of the prosthetic hand system in real time for further incremental control until the user confirms the stable grasp, as shown in Figure 5.

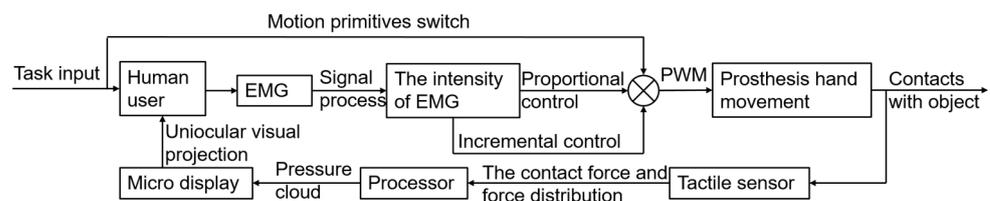


Figure 5. Block diagram for grasping force control of prosthetic hand with the human in the control loop.

4. Experiment

This section sets up a grasping platform for prosthetic hands to verify the feasibility of using the semi-autonomous controller and its robust performance. Users can realize grasping and manipulating tasks of a prosthetic hand using the semi-autonomous controller in the way of “hierarchical visual stream driven”.

As shown in Figure 6 (from the recorder’s perspective), the subject of this experiment is a 25-year-old Chinese adult male with normal vision. Elements of the semi-autonomous control framework are listed below:

- Jetson Nano (Nvidia, Santa Clara, CA, USA), which has 128-core NVIDIA Maxwell™ architecture GPU and Quad-core ARM® Cortex®-A57 MPCore processor and the semi-autonomous controller are integrated in the Jetson Nano.
- WX151HD CMOS image sensor (S-YUE, Shenzhen, China), which has 150-degree wide angle.
- ZJUT prosthetic hand (developed by Zhejiang University of Technology, Hangzhou, China) is equipped with 5 actuators.
- DYHW110 micro-scale pressure sensor (Dayshensor, Bengbu, China) is integrated in the prosthetic hand to obtain the touch force. It has a range of 5 kg, and the combined error is 0.3% of the full scale (F.S.).
- Vufine+ wearable display (Vufine, Sunnyvale, CA, USA) is a high-definition, wearable display that seamlessly integrates with the proposed control framework.

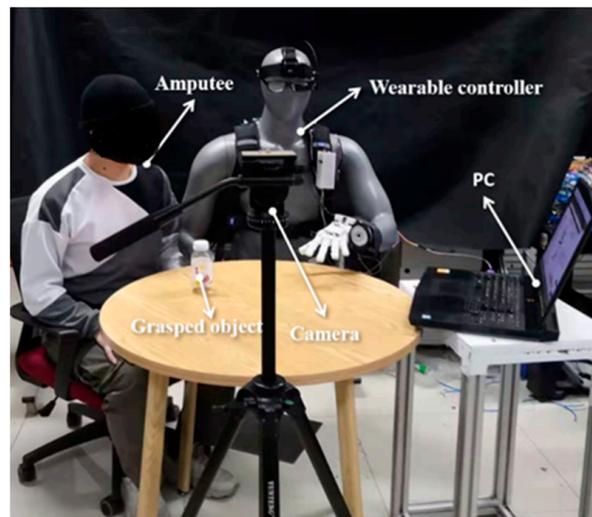


Figure 6. Experimental setup.

On the right side of the experimental setup is the semi-autonomous controller independently developed by the laboratory for this study. This controller follows the user’s multi-modal information and facilitates the grasping function of the prosthetic hand. The detailed structure of the semi-autonomous controller was thoroughly discussed in Section 3. Additionally, the prosthetic hand system, also independently developed by the laboratory, is affixed to the damaged hand of the mannequin wearing the semi-autonomous controller. Other components within the experimental environment include a Graphic User Interface (GUI) on a personal computer. This GUI serves the convenience of subjects, experiment operators, safety officers, and recorders, enabling them to monitor the state of the semi-autonomous controller and make timely adjustments. The object being manipulated in the experiment is a plastic beverage bottle, mimicking the user’s routine grasping actions for daily beverage needs. The experiment recorder controls and adjusts the process based on the recorded experimental footage, as illustrated in Figure 7.

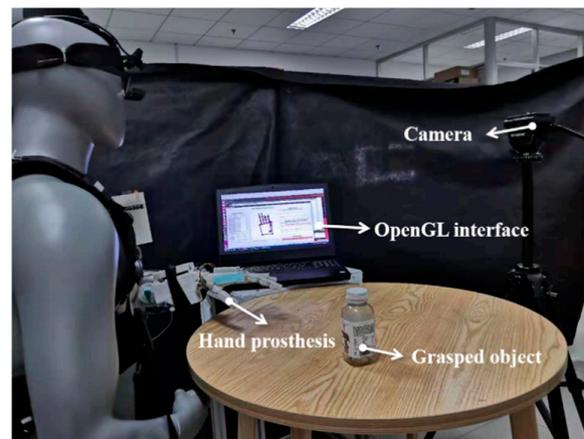


Figure 7. The view field from the subject's perspective.

4.1. Prosthetic Hand Grasping Experiment

In their daily lives, amputees often grasped objects with their prosthetic hands. The purpose of the experiment is that the semi-autonomous controller can assist the user in grabbing the “bottle” naturally under the multi-mode interaction between the semi-autonomous controller and the human user. The grasping process is as follows, which can be seen in Figure 8:

- (1) The semi-autonomous controller was worn by a dummy model. A prosthetic hand, electromyographic electrode, and head-mounted device on the human subject are used to obtain the human EMG signal and project the signal to the human eye. (Figure 8.1).
- (2) The human subject attached the prosthetic hand to his left hand and exhaled, “grab the bottle”. The subject was looking at the bottle with his arm close to it. The head-mounted device will perform visual semantic segmentation and convolutional neural network recognition for the “bottle” in this process. According to the information returned by the CMOS image sensor and the speech task library, the prosthetic hand can switch the combination of motion primitives. (Figure 8.2).
- (3) When the human subject's arm comes close to the “bottle”, it generates an EMG signal. The semi-autonomous controller implements proportional and incremental control under a specific motion primitive, depending on the EMG strength, until the user is instructed to “determine” the prosthetic hand motion. During this period, users can observe the changes in the pressure cloud of the prosthetic hand in real time. (Figure 8.3–6).
- (4) The experimental subjects grasp the “bottle” and place it in another position on the table. After placement, the prosthetic hand is released and reset by voice. (Figure 8.7–8).

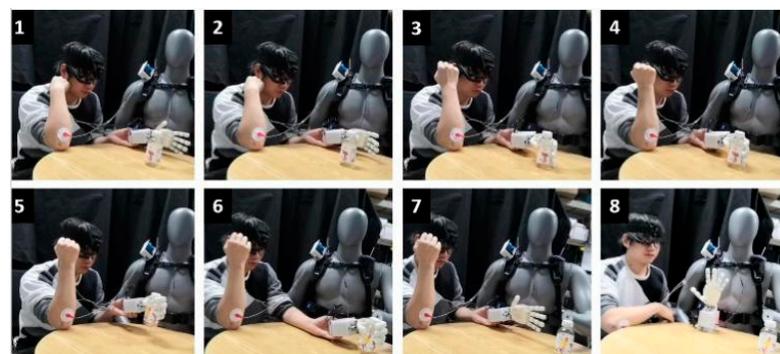


Figure 8. Arm and hand collaborative control experiment.

4.2. Prosthetic Hand and Human Hand Coordinative Manipulation Experiments

This experiment aims to verify that the semi-autonomous controller can assist humans in completing the manipulation of “screw the bottle cap” naturally under the multi-mode interaction between the semi-autonomous controller and the human user. The cooperative manipulation experiment of the prosthetic hand and human hand can be seen in Figure 9:

- (1) The semi-autonomous controller was worn by the dummy. A prosthetic hand, EMG electrode, and head-mounted device were attached to the human subjects. (Figure 9.1).
- (2) The experimental subject fixed the prosthetic hand on his left hand and exclaimed the command “screw the bottle cap” by voice. While the subject is looking at the “bottle”, the user is grabbing the “bottle” with his right hand, and the prosthetic hand with his left arm is approaching the “bottle cap”. The head-mounted device will perform visual semantic segmentation and convolutional neural network recognition for the “bottle” in this process. According to the information returned by the CMOS image sensor and the speech task library, the prosthetic hand can switch the combination of motion primitives. (Figure 9.2–3).
- (3) When the distance between the prosthetic hand and the “bottle cap” is less than the threshold value, the prosthetic hand will grab the “bottle cap”, and the user drives the prosthetic hand to rotate the “bottle cap” to the set angle through his left arm. (Figure 9.4–7).
- (4) When the “bottle cap” is not unscrewed and the distance between the prosthetic hand and the “bottle cap” exceeds a certain distance, the prosthetic hand will return to the pre-hand type. (Figure 9.8–10).
- (5) Repeat Step 3 and Step 4 until the cap is unscrewed.
- (6) Place the unscrewed bottle cap on the desktop and command the voice to reset the prosthetic hand. (Figure 9.8–10).

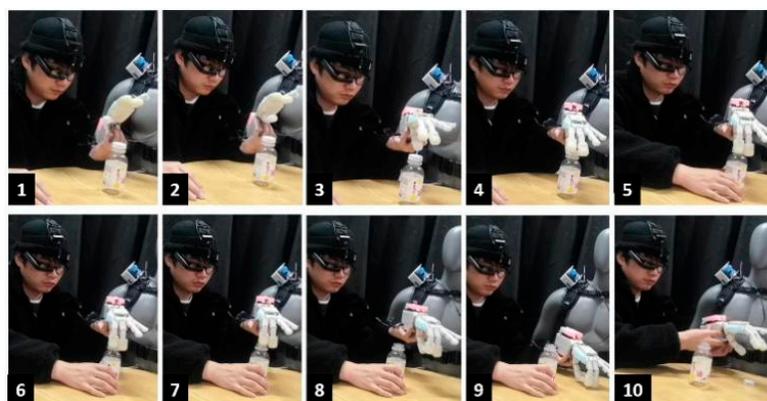


Figure 9. The cooperative manipulation experiment of a prosthetic hand and human hand.

4.3. Results

According to the results of two experiments, the proposed two visual-stream-driven manipulation strategy is in line with the natural manipulation rules of human beings and can effectively assist patients in completing familiar grasping and manipulation tasks in daily life. The experimental results are consistent with the expectations.

The experimental results shown in Table 3 combine the user experience and controller characteristics. The grasping task of a prosthetic hand and the human manipulation task described in this paper are similar in the structure form of the control method. In terms of the task layer, voice interaction is combined with the visual neural network. Due to differences in user intention and environment, the processing content will change the output result. At the planning level, the action primitive of the grasping task mainly switches between grasping hand type and pre-grasping hand type; the state transition condition is primarily used to trigger the user’s man-in-loop control. The motion primitives of manipulation tasks switch between multiple manipulators, and the state transition

condition is mainly used to trigger the arrangement of action primitives. In the motion control layer, the size of manipulated objects is usually tiny, so the switch of motion primitive is mainly realized through arm movement guidance. There are differences in the size, mass, and type of the objects in the grasping task, which requires higher grasping stability; thus, the force control strategy of the human-in-loop is introduced.

Table 3. The experimental results.

	Task Type	Task Characteristics
Coordinative movement of hand and arm Two hands coordination	Grasp manipulation	Force control Movements switch

5. Conclusions and Future Works

5.1. Conclusions

Different from the traditional semi-autonomous controller design, this work is inspired by the two visual streams of the humanoid control strategy. The summary of this framework can be seen in Figure 10. The main contributions are as follows:

- (1) In terms of the control layer of the motion primitive planning, the traditional object semantic segmentation and CNN recognition are classified as ventral flow, residual arm motion tracking is introduced as dorsal flow, inspired by the human brain. We optimized the information collection method of the motion primitive planning control layer and the state transfer strategy among the movement primitives according to the multimodal information in different stages of ventral flow and dorsal flow.
- (2) In terms of the force control layer, the issue of the user's "cognitive burden" is reduced in the existing semi-autonomous control strategy, so this paper takes the user as a high-dimensional controller and the EMG strength of flexor digitorum profundus as the control quantity, giving feedback to the prosthetic hand body state, realizing precise force control with human-in-loop.

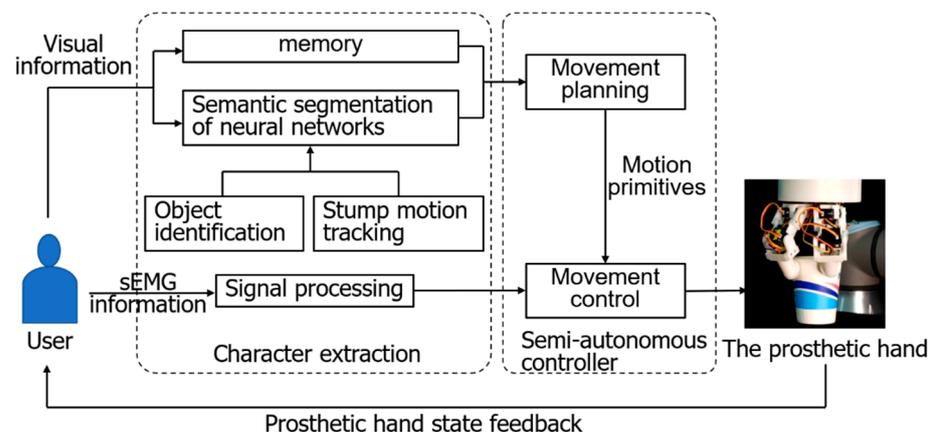


Figure 10. Framework of the semi-automatic controller for the prostheses hand.

5.2. Future Works

The proposed control framework, inspired by the human ventral–dorsal stream visual process, currently emphasizes functional grasping actions within the spectrum of human hand operations. However, this biological process represents just one aspect of the myriad ways humans execute gripping tasks, emphasizing functional manipulations. Ongoing neuroscientific efforts delve into understanding how different brain regions guide grasping operations through visual cues. Building upon the current ventral–dorsal stream, future work will explore bio-inspired control strategies, integrating the latest neuroscientific findings related to the user's dual visual neural pathways. The aim is to deepen the

algorithmic sophistication and broaden the spectrum of multimodal information, fostering a more profound integration between humans and machines.

In future investigations, a pivotal area for exploration revolves around the usability and social acceptance of prosthetic hands in various environments, especially social situations. Current technology, while advancing rapidly, may pose challenges in social integration due to its external and conspicuous nature. To enhance user experience and societal acceptance, future work could focus on developing discreet, aesthetically pleasing designs that seamlessly integrate prosthetic hands into social settings. This involves not only technical advancements but also a nuanced understanding of user preferences, comfort levels, and societal perceptions. Exploring materials, form factors, and user-centric design principles could contribute significantly to reducing the stigma associated with prosthetic devices, fostering a more inclusive and socially integrated environment for users. This research direction aligns with the broader goal of not only improving the functional aspects of prosthetic hands but also enhancing the overall quality of life and social experiences for individuals using this technology.

Author Contributions: Conceptualization, J.Z. and G.B.; methodology, X.Z.; software, B.Y.; validation, B.Y. and X.Z.; formal analysis, X.M.; data curation, H.F.; writing—original draft preparation, X.Z.; writing—review and editing, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Program of Zhejiang (Grant No. 2021C04015), the Key Research Program of Zhejiang (Grand No. LZ23E050005), Natural Science Foundation of Zhejiang Province (Grand No. Q23E050071).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Mouri, T.; Kawasaki, H.; Ito, S. Unknown object grasping strategy imitating human grasping reflex for anthropomorphic robot hand. *J. Adv. Mech. Des. Syst. Manuf.* **2007**, *1*, 1–11. [CrossRef]
2. Stephens-Fripp, B.; Alici, G.; Mutlu, R. A review of non-invasive sensory feedback methods for transradial prosthetic hands. *IEEE Access* **2018**, *6*, 6878–6899. [CrossRef]
3. Biddiss, E.A.; Chau, T.T. Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthet. Orthot. Int.* **2007**, *31*, 236–257. [CrossRef]
4. Ortiz, O.; Kuruganti, U.; Blustein, D. A Platform to Assess Brain Dynamics Reflective of Cognitive Load during Prosthesis Use. MEC20 Symposium. 2020. Available online: https://www.researchgate.net/publication/373294777_A_PLATFORM_TO_ASSESS_BRAIN_DYNAMICS_REFLECTIVE_OF_COGNITIVE_LOAD_DURING_PROSTHESIS_USE (accessed on 10 December 2023).
5. Parr, J.V.V.; Vine, S.J.; Wilson, M.R.; Harrison, N.R.; Wood, G. Visual attention, EEG alpha power and T7-Fz connectivity are implicated in prosthetic hand control and can be optimized through gaze training. *J. Neuroeng. Rehabil.* **2019**, *16*, 52. [CrossRef] [PubMed]
6. Deeny, S.; Chicoine, C.; Hargrove, L.; Parrish, T.; Jayaraman, A. A simple ERP method for quantitative analysis of cognitive workload in myoelectric prosthesis control and human-machine interaction. *PLoS ONE* **2014**, *9*, e112091. [CrossRef] [PubMed]
7. Parr, J.V.; Wright, D.J.; Uiga, L.; Marshall, B.; Mohamed, M.O.; Wood, G. A scoping review of the application of motor learning principles to optimize myoelectric prosthetic hand control. *Prosthet. Orthot. Int.* **2022**, *46*, 274–281. [CrossRef] [PubMed]
8. Ruo, A.; Villani, V.; Sabattini, L. Use of EEG signals for mental workload assessment in human-robot collaboration. In Proceedings of the International Workshop on Human-Friendly Robotics, Delft, The Netherlands, 22–23 September 2022; pp. 233–247.
9. Cordella, F.; Ciancio, A.L.; Sacchetti, R.; Davalli, A.; Cutti, A.G.; Guglielmelli, E.; Zollo, L. Literature review on needs of upper limb prosthesis users. *Front. Neurosci.* **2016**, *10*, 209. [CrossRef]
10. Park, J.; Zahabi, M. Cognitive workload assessment of prosthetic devices: A review of literature and meta-analysis. *IEEE Trans. Hum. Mach. Syst.* **2022**, *52*, 181–195. [CrossRef]
11. Foster, R.M.; Kleinholdermann, U.; Leifheit, S.; Franz, V.H. Does bimanual grasping of the Muller-Lyer illusion provide evidence for a functional segregation of dorsal and ventral streams? *Neuropsychologia* **2012**, *50*, 3392–3402. [CrossRef]

12. Cloutman, L.L. Interaction between dorsal and ventral processing streams: Where, when and how? *Brain Lang.* **2013**, *127*, 251–263. [[CrossRef](#)]
13. Van Polanen, V.; Davare, M. Interactions between dorsal and ventral streams for controlling skilled grasp. *Neuropsychologia* **2015**, *79*, 186–191. [[CrossRef](#)] [[PubMed](#)]
14. Brandi, M.-L.; Wohlschläger, A.; Sorg, C.; Hermsdörfer, J. The neural correlates of planning and executing actual tool use. *J. Neurosci.* **2014**, *34*, 13183–13194. [[CrossRef](#)] [[PubMed](#)]
15. Culham, J.C.; Danckert, S.L.; DeSouza, J.F.X.; Gati, J.S.; Menon, R.S.; Goodale, M.A. Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.* **2003**, *153*, 180–189. [[CrossRef](#)] [[PubMed](#)]
16. Meattini, R.; Benatti, S.; Scarcia, U.; De Gregorio, D.; Benini, L.; Melchiorri, C. An sEMG-based human–robot interface for robotic hands using machine learning and synergies. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2018**, *8*, 1149–1158. [[CrossRef](#)]
17. Lange, G.; Low, C.Y.; Johar, K.; Hanapiah, F.A.; Kamaruzaman, F. Classification of electroencephalogram data from hand grasp and release movements for BCI controlled prosthesis. *Procedia Technol.* **2016**, *26*, 374–381. [[CrossRef](#)]
18. Thomas, N.; Ung, G.; Ayaz, H.; Brown, J.D. Neurophysiological evaluation of haptic feedback for myoelectric prostheses. *IEEE Trans. Hum. Mach. Syst.* **2021**, *51*, 253–264. [[CrossRef](#)]
19. Cognolato, M.; Gijssberts, A.; Gregori, V.; Saetta, G.; Giacomino, K.; Hager, A.-G.M.; Gigli, A.; Faccio, D.; Tiengo, C.; Bassetto, F. Gaze, visual, myoelectric, and inertial data of grasps for intelligent prosthetics. *Sci. Data* **2020**, *7*, 43. [[CrossRef](#)]
20. Laffranchi, M.; Boccardo, N.; Traverso, S.; Lombardi, L.; Canepa, M.; Lince, A.; Semprini, M.; Saglia, J.A.; Naceri, A.; Sacchetti, R. The Hannes hand prosthesis replicates the key biological properties of the human hand. *Sci. Robot.* **2020**, *5*, eabb0467. [[CrossRef](#)]
21. Kilby, J.; Prasad, K.; Mawston, G. Multi-channel surface electromyography electrodes: A review. *IEEE Sens. J.* **2016**, *16*, 5510–5519. [[CrossRef](#)]
22. Mendez, V.; Iberite, F.; Shokur, S.; Micera, S. Current solutions and future trends for robotic prosthetic hands. *Annu. Rev. Control Robot. Auton. Syst.* **2021**, *4*, 595–627. [[CrossRef](#)]
23. Markovic, M.; Dosen, S.; Popovic, D.; Graimann, B.; Farina, D. Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis. *J. Neural Eng.* **2015**, *12*, 066022. [[CrossRef](#)] [[PubMed](#)]
24. Starke, J.; Weiner, P.; Crell, M.; Asfour, T. Semi-autonomous control of prosthetic hands based on multimodal sensing, human grasp demonstration and user intention. *Robot. Auton. Syst.* **2022**, *154*, 104123. [[CrossRef](#)]
25. Fukuda, O.; Takahashi, Y.; Bu, N.; Okumura, H.; Arai, K. Development of an IoT-based prosthetic control system. *J. Robot. Mechatron.* **2017**, *29*, 1049–1056. [[CrossRef](#)]
26. Shi, C.; Yang, D.; Zhao, J.; Liu, H. Computer vision-based grasp pattern recognition with application to myoelectric control of dexterous hand prosthesis. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 2090–2099. [[CrossRef](#)]
27. Wang, X.; Haji Fathaliyan, A.; Santos, V.J. Toward shared autonomy control schemes for human-robot systems: Action primitive recognition using eye gaze features. *Front. Neurobotics* **2020**, *14*, 567571. [[CrossRef](#)] [[PubMed](#)]
28. Guo, W.; Xu, W.; Zhao, Y.; Shi, X.; Sheng, X.; Zhu, X. Towards Human-in-the-Loop Shared Control for Upper-Limb Prostheses: A Systematic Analysis of State-of-the-Art Technologies. *IEEE Trans. Med. Robot. Bionics* **2023**, *5*, 563–579. [[CrossRef](#)]
29. He, Y.; Shima, R.; Fukuda, O.; Bu, N.; Yamaguchi, N.; Okumura, H. Development of distributed control system for vision-based myoelectric prosthetic hand. *IEEE Access* **2019**, *7*, 54542–54549. [[CrossRef](#)]
30. He, Y.; Fukuda, O.; Bu, N.; Yamaguchi, N.; Okumura, H. Prosthetic Hand Control System Based on Object Matching and Tracking. The Proceedings of JSME Annual Conference on Robotics and Mechatronics (Robomec), 2019; p. 2P1-M09. Available online: https://www.researchgate.net/publication/338151345_Prosthetic_hand_control_system_based_on_object_matching_and_tracking (accessed on 10 December 2023).
31. Vorobev, E.; Mikheev, A.; Konstantinov, A. A Method of Semiautomatic Control for an Arm Prosthesis. *J. Mach. Manuf. Reliab.* **2018**, *47*, 290–295. [[CrossRef](#)]
32. Ghosh, S.; Dhall, A.; Hayat, M.; Knibbe, J.; Ji, Q. Automatic gaze analysis: A survey of deep learning based approaches. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *46*, 61–84. [[CrossRef](#)]
33. Bullock, I.M.; Ma, R.R.; Dollar, A.M. A hand-centric classification of human and robot dexterous manipulation. *IEEE Trans. Haptics* **2012**, *6*, 129–144. [[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.