Artificial Intelligence and Human-Computer Interaction
Y. Ye and P. Siarry (Eds.)

doi:10.3233/FAIA240166

The Prosthetic Hand Can Grasp Occluded Objects with the Help of 3rd-Eye Based on Artificial Vision and AR Technology

Chunyuan SHI [a] and Jingdong ZHAO [a,b 1]
[a] *State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150081, China*
[b] *Artificial Intelligence Laboratory, Harbin Institute of Technology, Harbin 150081, China*
ORCiD ID: Chunyuan Shi https://orcid.org/0000-0002-1479-0926
ORCiD ID: Jingdong Zhao https://orcid.org/0000-0002-2074-6534

Abstract. Grasping occluded objects is a large challenge for hand amputees because of the hand's perception loss. This paper proposed an alternative approach based on artificial vision and augmented reality technology, termed 3rd-Eye, to solve the problem. In detail, a live stream of an RGB camera attached to the prosthetic hand was presented to the user who wore an augmented reality helmet, HoloLens 2. The 3rd-Eye was integrated into a hybrid control prosthetic system (i-MYO proposed in our previous work). It was tested in a special reach-and-grasp experiment where the subject needed to grasp an occluded object in cluttered environments (one to six objects in total) using the prosthetic hand. The new approach has achieved satisfactory results (success rate: 100% in all).

Keywords. Prosthetic hand, computer vision, augmented reality

1. Introduction

Grasping an occluded object is a major challenge for hand amputees wearing a commercial prosthetic hand. A healthy person can process the occluded object thanks to the wealth of sensors in the skin of the palm and finger and the excellent neurofeedback system [1]. Unfortunately, the amputee loses the perception ability of the hand due to the amputation. Wearing a commercial prosthetic hand can restore the grasp function but can't regain perception.

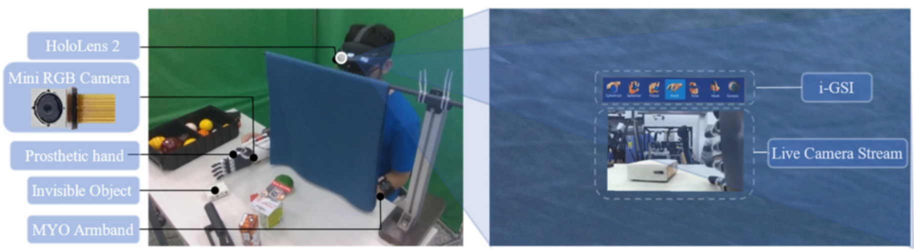
Compared with the natural hand, the prosthetic hand in research is rudimentary in sensors and feedback interface. The information of the prosthetic hand is often encoded into limited modes to give feedback to the user [2]. The information includes the one from interacting with objects (e.g., force [2–4] and proprioceptive information (e.g., hand aperture [2]). The feedback method includes vibrotactile [4], transcutaneous electrical stimulation [3], force applicator [5], etc.

The narrow information bandwidth is the biggest problem in the above perception alternative methods [2]. These methods make it difficult to explore the environment

[1]Corresponding Author: Jingdong Zhao, zhaojingdong@hit.edu.cn

because the feedback information is limited and not intuitive enough. On the one hand, the prosthetic hand sensors cannot get enough information due to technical limitations. It is generally only equipped with tactile sensors at the fingertips [6]. On the other hand, only part of the information can be feedbacked to the user. Due to the surface space limitation of the stump, limited feedback devices can be placed on it. The user can only distinguish between a few patterns that the electrical stimulator or shock sensor stimulates on the forearm stump.

The vision information plays an important role in the reach-and-grasp task, including the hand information and the surrounding information of objects [7]. After proprioception is removed due to amputation, amputees rely heavily on inherent visual supervision to manipulate a prosthetic hand [3]. So, it is challenging for amputees to grasp an occluded object because the inherent vision is also removed. Some studies have integrated a camera into the prosthetic hand [8–10], but the aim is to obtain low-level control information from the image, e.g., the grasp types and aperture of the hand.



a) System Setting b) First-person Perspective

Figure 1. System composition diagram. The i-GSI is a grasp-type switching interface based on eye-tracking and augmented reality technology proposed in our previous work2.

We proposed a simple, practical, alternative approach based on artificial vision and AR technology, termed 3rd-Eye, to restore the hand's ability to explore the environment [11]. In detail, as shown in Figure 1, we attach an RGB camera to the prosthetic hand and present the live camera stream to the user who wears an AR helmet (HoloLens 2). The 3rd-Eye was integrated with a hybrid prosthetic control system proposed in our recent work [12]. With the new system, the user can use the i-GSI [13] to toggle (show or hide) the live camera stream and switch the grasp types of the prosthetic hand, and use an EMG to control the hand opening or closing. The new approach achieved fairly good results (success rate:100% in all tests) in a special reach-and-grasp experiment requiring the user to grasp an occluded object using the prosthetic hand.

2. Materials and methods

2.1. System components

The system consists of a robot hand, a mini camera, an augmented reality helmet, EMG electrodes, and a standard computer. The mini camera and augmented reality helmet were the main part of 3rd-Eye.

A robot hand with five independent movement fingers (Inspired hand, China) was used to simulate a prosthetic hand. A look-table-up table method was used to control the hand to close, forming six grasp types [14] (Cylindrical, Spherical, Pinch, Tripod, Lateral, Hook). The robot hand was position-controlled, and the position parameters of six grasp types were saved in an offline table. These parameters were obtained by trying and testing.

A normal USB RGB camera (30 Hz, field of view 68°, mm focal length, 2592 × 1944 resolution) was attached to a robot hand to provide live streaming containing information about the hand's surroundings. The vision of the hand and the hand's surroundings profoundly affect the hand's operation [7]. Figure 1(a) shows that the camera was placed at the palm and wrist junction to observe the prosthetic hand and hand's surroundings.

HoloLens 2 (Microsoft, the USA), an augmented reality helmet based on holographic projection (see-through display), was used to present the live camera streaming to the user. The live streaming was shown on a holographic canvas (17 cm × 10 cm) in front of the user at about 0.5 m, as shown in Figure 1(b). The canvas is relatively fixed to the head. The i-GSI, a grasp-type switching interface based on eye-tracking and augmented reality, was used to switch the grasp type of the hand, which was proposed in our previous work and ran in the augmented reality helmet [13]. In this paper, we expand the i-GSI by adding another GazeButton, the 3rd-Eye GazeButton, to the grasp-type panel with six GazeButtons. A GazeButton refers to a button interactive with gaze [15]. The new GazeButton is used to toggle (show or hide) the canvas.

An MYO armband with eight EMG electrodes was used to collect the myoelectric signal of the user's forearm to control the hand opening/closing. The commercial application of the armband can use a built-in supervised learning algorithm to recognize five EMG patterns, including wave-in and wave-out EMG patterns. The wave-in pattern was used to close the robot hand, and the wave-out pattern was used to open it in this paper. The armband was worn on the left forearm of the user. The armband is convenient thanks to its Bluetooth communication with the PC and can also be replaced by a two-site EMG and threshold algorithm like our previous work [16].

2.2. Control flow

Figure 2 shows the control follow of the new system. A user first glances at the camera GazeButton to show the live camera streaming when they want to grasp occluded objects with the prosthetic hand. Then, the user observes objects through the live streaming while they manipulate the prosthetic hand to patrol the objects' area. After deciding which object to grasp or finding the target object, the user glances at the grasp-type GazeButton to switch the grasp type. Finally, the user manipulates the prosthetic hand to a suitable position relative to the target object and activates a myoelectric signal to control the prosthetic hand to grasp the target object.

2.3. Experimental protocol

A special reach-and-grasp experiment was used to test the new system to see if the system can help the user grasp an occluded object, as shown in Figure 3. The user, with the HoloLens 2 on the head, sits comfortably next to the desk. There were some daily-life objects on the desk, and a curtain blocked the user's field of vision, preventing him from seeing objects in front of him. The user was required to manipulate the robot hand to pick up a target object and put it in a black box on the right. The experiment was set up in four different situations to test the performance of the new system as follows:

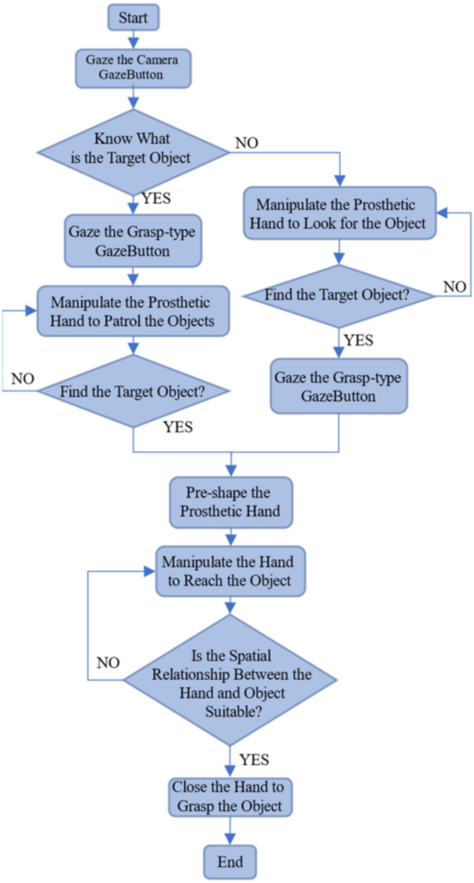
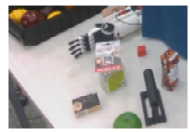


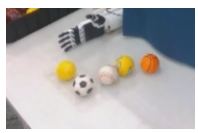
Figure 2. Control flow diagram.

a) Experimental scene

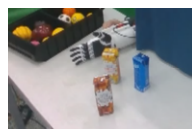
b) SCENE1



c) SCENE2



d) SCENE3



e) SCENE4

Figure 3. Experimental setup.

a) Objects used in SCENE1

b) Objects used in SCENE3

c) Objects used in SCENE4

Figure 4. Objects for experiment.

1) SCENE1: Only an object was on the desk in one test, as shown in Figure 3(b). There are six distinct candidate objects, as shown in Figure 4(a). The shape and size of these objects are obviously different, and the robot needs to grasp them using different grasp types.

2) SCENE2: In every test, there were six objects on the desk at the same time, and the user needed to grasp and transfer the target object, as shown in Figure 3(c). The objects are the same as the ones in SCENE1. Especially, all objects are rearranged in every test.

3) SCENE3: the same as SCENE2, except that five similar toy balls, as shown in Figure 3(d), replace the six objects in SCENE2. The balls' shape, size, and material are the same, but the surface texture and color are slightly different, as shown in Figure 4(b).

4) SCENE4: the same as SCENE3, except three very similar bottles, as shown in Figure 3(e), replace the balls in SCENE3. The three bottles differ only in color, as shown in Figure 4(c).This paper compares the new method with the human hand. The user also uses his hand to complete the SCENE 2–4 experiment. The superiority of the new method can be more fully demonstrated by comparing it with the results of the human hand.

Before the experiment, an assistant calibrates sensors for the user and trains simply the user to become familiar with the new system. An operating system built-in application was used to calibrate the eye-tracking sensor in HoloLens 2 for the user. Some labelled data of EMG was used to calibrate the MYO armband. The assistant introduced the system briefly to the user and helped the user try less than five tests in fifteen minutes. The user needs to complete six sets of tests for each scene. There are a total of 120 trials (SCENE1: 36 times, SCENE2: 36 times, SCENE3: 30 times, SCENE4: 18 times).

In SCENE1, the assistant randomly placed an object on the table and instructed the user to begin the test. In SCENE2-4, after rearranging all objects, the assistant randomly chose a target object and instructed the user to begin. Two metrics were defined to evaluate performance as follows:

1) The time to complete the task in one test: the time was from the experiment beginning to object releasing.

2) The task success rate: it was considered successful that the user picked up the target object using optimal grasp type classification in Figure 4 and put it in the black box.

3. Results & discussions

One user completed 120 tests of SCENE1-4 using the new system and 76 tests of SCENE2-3 using his hand. After trying some tests, he clearly claimed that he could not use his hand to complete the test in SCENE4. Because only with the tactile and force sensor in hand, it is impossible to distinguish different colors. Figure 5 shows a complete trial in SCENE2, where the user used the prosthetic hand to pinch a small cube (1.8 cm) in cluttered environments and put it in the box.

The new system with the 3rd-Eye performs satisfactorily in completing the experiment task, even better than the human hand. Figure 6 shows that the user successfully found the target object through the 3rd-Eye in the prosthetic hand and transferred the target object using the prosthetic hand to the black box in all tests. But, as the difficulty of the test increases, it becomes more difficult for the human hand to distinguish the target object. Although the user can find all the objects by groping on the table in SCENE3-4, he only successfully grasped the target object in 70% of tests in

SCENE3 and was completely unsure which one was the target object in SCENE4. Although the user could use his hand to transfer the target objects in SCENE2, other unrelated objects were often affected when the hand groped the target object. Figure 7 shows that the user accidentally made the thin object drop while groping for the small red cube.

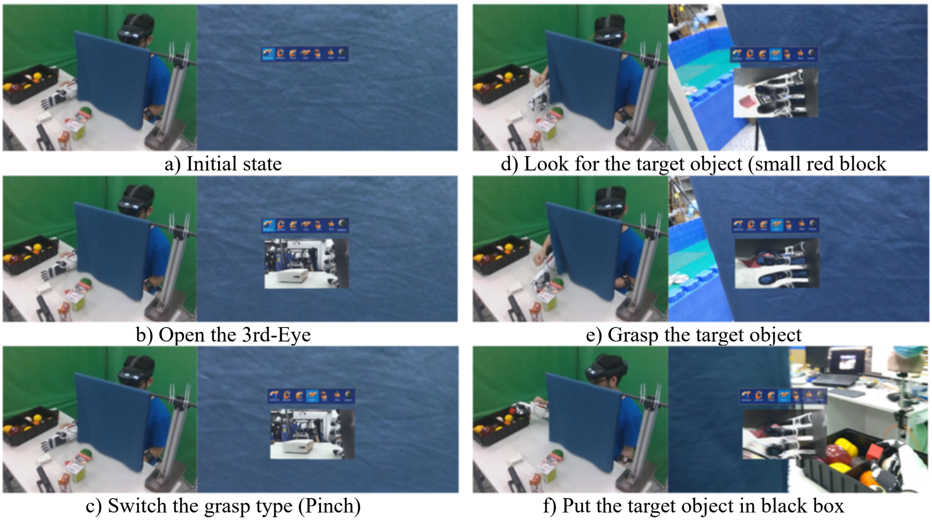


Figure 5. Schematic diagram of the actual scenes.

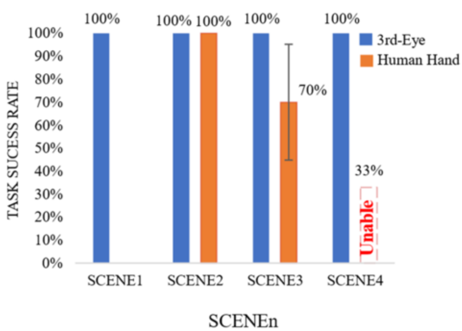


Figure 6. Success rate result in different scenes.

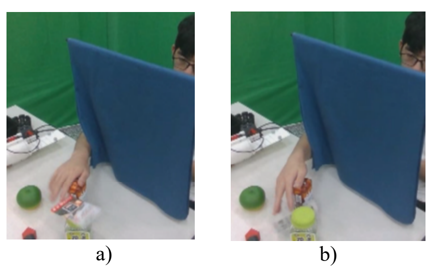


Figure 7. The user accidentally made other object drops while groping the target object with his hand.

The new system spent less time on difficult tasks than the human hand while spending more time on simple tasks. Figure 8 shows that the user spent about 12 s to finish one test using the prosthetic hand in SCENE1-2 and about 8 s using his hand in SCENE2. The user stated that it was easier to find the target object using the 3rd-Eye, but it took more time to manipulate the prosthetic hand to reach the target object and close the hand, especially for objects grasped by Lateral and Pinch. The user spent about 8 s to complete one test using the prosthetic hand in SCENE3-4 and about 11 s using his hand in SCENE3. The user stated that there was little difference in finding the target object using the 3rd-Eye in all scenes, but it required more time to distinguish the texture features of the object surface using the human hand in SCENE3.

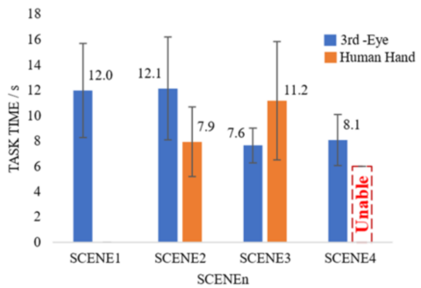


Figure 8. Snapshot on prosthesis transporting four objects in a group.

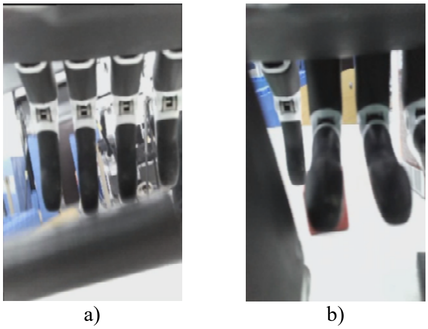


Figure 9. Images of hand camera. (a) The user prepared to grasp an object using the Hook grasp type. (b) The user prepared to grasp the red cube using the Pinch grasp type.

Although the third eye can restore the patient's hand perception well, some deficiencies remain. The live camera streaming on plane canvas loses the three-dimensional space information, so the user still needs a certain amount of trial before closing the hand to ensure that the hand can successfully grasp the object. For example, the user was difficult to judge the front-to-back position of the fingers and the object in Figure 9(a). It may be better to use a point cloud [17] to replace the live camera streaming, which is our next work.

In addition, the intuitiveness of the 3rd-Eye can be further improved. The camera view is not intuitive because it differs from the inherent eye. Moreover, when the prosthetic hand closes to grasp the object, it is easy to block the camera's field, as shown in Figure 9(b). It may be solved using the 3D reconstruction [18] and holographic perspective based on augmented reality technology, which is also our next work.

In this system, we used a 1440P resolution camera with a frame rate of 30FPS as a case for the new method. It can further optimize the system by utilizing a higher-speed capture device and adopting a superior video compression method, such as H.265 (HEVC) instead of H.264 (AVC), to enhance the user experience and improve the scene reconstruction rate.

4. Conclusions

This paper presents a perception feedback method, termed 3rd-Eye, in which artificial vision was used instead of artificial tactile for the prosthetic hand, and augmented reality was used to feedback the vision information to the user. In a special reach-and-grasp experiment requiring the user to grasp an occluded object in cluttered environments, the 3rd-Eye performed better, even better than a human hand in difficult tasks. Our future studies include improving the intuitiveness of the method and validating it on amputees.

References

[1] MacKenzie CL, Iberall T. The Grasping Hand. 1994.

[2] Dosen S, Markovic M, Farina D, Karnal H, Graimann B. GLIMPSE: Google Glass interface for sensory feedback in myoelectric hand prostheses. J Neural Eng. 2007; 14(3): 036007–036007.

[3] Yang B, Jiang L, Ge C, Cheng M, Zhang J. Control of myoelectric prosthetic hand with a novel proximity-tactile sensor. Sci China Technological Sci. 2022; 65(7): 1513–1523.

[4] Witteveen HJ, Rietman HS, Veltink PH. Vibrotactile grasping force and hand aperture feedback for myoelectric forearm prosthesis users. Prosthet Orthot Int. 2014 ;39 (3): 204–212.

[5] Kim K, Colgate JE, Santos-MunnéJ J, Makhlin A, Peshkin MA. On the design of miniature haptic devices for upper extremity prosthetics. IEEE Asme Transactions Mechatronics. 2010; 15(1): 27–39.

[6] Jiang L, Zeng B, Fan S, Sun K, Zhang T, Liu H. A modular multisensory prosthetic hand.2014 IEEE Int Conf Information Automation Icia, no. 2014 July; 648–653.

[7] Bozzacchi C, Brenner E, Smeets JB, Volcic R, Domini F. How removing visual information affects grasping movements.Exp Brain Res. 2018; 236(4): 985–995.

[8] Došen S, Cipriani C, Kostić M, Controzzi M, Carrozza MC, Popović DB. Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation. J Neuroeng Rehabil. 2010;7(1): 42–42.

[9] Ghazaei G, Alameer A, Degenaar P, Morgan G, Nazarpour K. Deep learning-based artificial vision for grasp classification in myoelectric hands. J Neural Eng. 2017; 14(3): aa6802–aa6802.

[10] Castro MN, Dosen S. Continuous semi-autonomous prosthesis control using a depth sensor on the hand. Front Neurorobotics. 2022; 16: 814973.

[11] Callier T, Saal HP, Davis-Berg EC, Bensmaia SJ. Kinematics of unconstrained tactile texture exploration. J Neurophysiol. 2014;113(7): 3013–3020.

[12] Shi C, Yang D, Zhao J, Jiang L. I-MYO: A hybrid prosthetic hand control system based on eye-tracking, augmented reality and myoelectric signal. Arxiv .2022.

[13] Shi C, Yang D, Qiu S, Zhao J. I-GSI: A novel grasp switching interface based on eye-tracking and augmented reality for multi-grasp prosthetic hands. IEEE Robot. Autom. Lett. 2023; 8(3): 1619–1626.

[14] TAYLOR CL, SCHWARZ RJ. The anatomy and mechanics of the human hand. Artif Limbs. 1955; 2(2): 22–35.

[15] Krejtz K, et al. GazeButton: Enhancing buttons with eye gaze interactions. Proc 11th Acm Symposium Eye Track Res Appl. 2019; 1–7.

[16] Shi C, Yang D, Zhao J, Liu H. Computer vision-based grasp pattern recognition with application to myoelectric control of dexterous hand prosthesis. IEEE T Neur Sys Reh. 2020; 28(9): 2090–2099.

[17] Wang H, Tian Y. Sequential point clouds: A Survey. Arxiv. 2022.

[18] Jang M, Lee S, Kang J, Lee S. Technical consideration towards robust 3d reconstruction with multi-view active stereo sensors. Sensors. 2022; 22(11):4142.