Proximity Perception-Based Grasping Intelligence: Toward the Seamless Control of a Dexterous Prosthetic Hand

Si-Hwan Heo and Hyung-Soon Park, *Member, IEEE*

Abstract—**Achieving the dexterity of a human hand is a major goal in the field of prosthetic hands. To achieve this level of dexterity, a compact high degrees-of-freedom robotic prosthesis and seamless control of the hardware is needed. In this article, to attain seamless control of a highly dexterous prosthetic hand, we propose a novel perception system that provides grasping intelligence to the prosthetic hand. The proximity perception-based grasping intelligence (P2GI) system comprises a proximity sensor system and a prompt decision-making process. The proximity sensors embedded in the prosthetic hand map the point cloud of the object in real time while the prosthetic hand reaches toward the object. Simultaneously, a real-time decision-making algorithm infers the user's intended grasp posture by obtaining the hand-object relation from the point cloud data. The finger motion that stably grasps the target object with the inferred grasp posture is planned accordingly. Consequently, the user can intuitively utilize various grasp postures of the prosthetic hand using a single-channel surface electromyography signal. The P2GI system was evaluated with ten subjects. The results show a grasp posture classification accuracy of 97.8% and a task success rate of 95.7% during the real-time grasping for unknown objects in daily life.**

Index Terms—**Autonomous system, intention recognition, prosthetic hand, proximity perception, shared autonomy.**

I. INTRODUCTION

THE hand is a remarkably versatile and complex body part, capable of performing a wide range of dexterous tasks. Therefore, amputation of the hand results in catastrophic damage to the quality of life for affected individuals. In order to restore their functionality and improve their quality of life, there is a great demand for developing a highly functional prosthetic hand with similar functionality to that of a human. Over the years, significant research efforts have been devoted to enhancing the functionality of prosthetic hands through advancements in various fields, such as mechanisms [1], [2], sensors [3], [4], and control methods [5], [6].

Effectively handling the high degrees-of-freedom (DOF) of a versatile prosthetic hand has remained an important problem regarding the control method [7]. If a user's intention cannot be properly inferred, it becomes a bottleneck for the overall performance of the prosthetic hand. While commercial products such as the Bebionic hand (OttoBock, Duderstadt, Germany) or the i-Limb quantum (Touch Bionics, Livingston, United Kingdom) can perform a wide range of grip postures, they still rely on an EMG (Electromyography)-triggered finite state machine or smartphone interfaces to select and perform various grip postures. The EMG-triggered finite state machine is a widely used method; however, the grasp postures must be manually selected before grasping via EMG signal patterns that the user must memorize [8]. Consequently, users tend to select only a few among the various grasp postures to avoid the cumbersome posture change process; and return to the simple but durable prosthetic hands [9]. As a result, there has been a longstanding demand for the intuitive and seamless control of a high-DOF prosthetic hand [10]. To achieve this, high posture classification accuracy, continuous decision-making, and low latency of less than 100–300 ms have been presented as the key requirements [10], [11].

To address the challenges of intuitive control of highly functional prosthetic hands, various intention recognition interfaces have been developed. Among them, EMG pattern recognition using a multichannel surface EMG sensor is the most widely adopted methodology. To increase the versatility, accuracy, and robustness of motor intention recognition, numerous methods have been developed across various fields, such as signal processing [12], sensor hardware development [13], [14], and surgical methods [15], [16]. However, several issues such as electrode shifting or skin sweating cause difficulties in real-life scenarios [17]. Moreover, the algorithms used to mitigate these issues often result in a long inference time and a large signal processor unit [18].
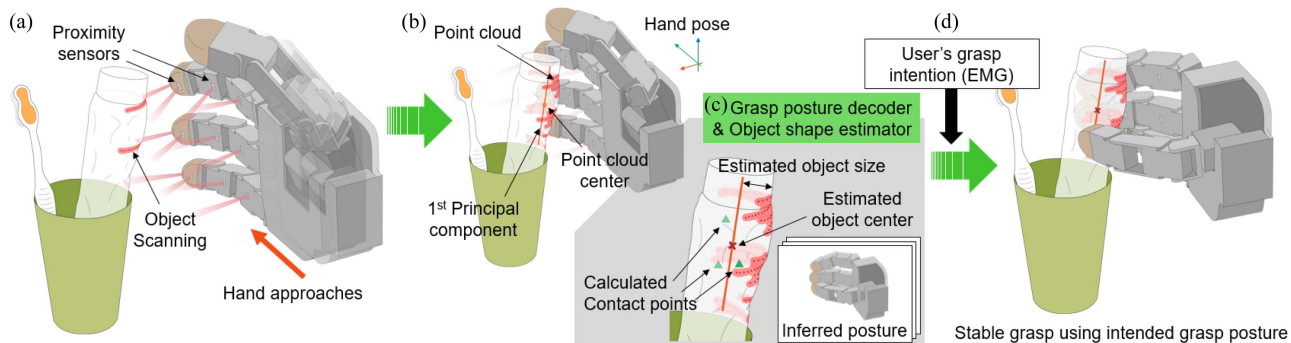
Fig. 1.   Methodology of the P2GI system. (a) Proximity sensors on the palmar side of the prosthetic hand scan the surface of an object as the hand approaches the object. (b) Scanned points form a point cloud of the target object. (c) Grasp posture decoder infers the intended grasp posture of the user; and the object shape estimator estimates the size and exact center position of the object. Based on these decisions, the contact points for stable grasp are calculated. (d) When the user gives the simple grasp command, the prosthetic hand stably grasps the object using the intended grasp posture of the user.

Recent advances in artificial intelligence have opened up new possibilities for control interfaces for prosthetic users. The concept of shared autonomy has been proposed, which autonomously handles the complexity of a high DOF prosthetic hand [19], [20], [21], [22], [23]. In particular, several studies have used a vision-based algorithm to control a dexterous myoelectric prosthetic hand [20], [21], [22], [23]. A mounted camera captures the workspace of the prosthetic hand user and infers the user's intended grasp posture. However, the mobility and compactness of these systems are limited, as the camera must be mounted somewhere, such as the surrounding environment, the prosthetic hand's wrist, or glasses. Moreover, the variety of daily living environments poses a significant challenge for these systems.

Here, we explore another sensory function for the prosthetic hand system-proximity perception, which could be considered as a sixth sense of humans. Proximity perception is a sensory system that can perceive nearby objects using proximity sensors embedded in a robotic system [24]. Although the proximity sensor senses one-dimensional distance data, which is less information than that provided by an RGB-D image, there have been studies that have demonstrated the possibility of inferring object information using proximity sensors [25], [26]. Recent studies have demonstrated systems that can obtain the point cloud of an object by sequentially scanning it with proximity sensors on a low-DOF robotic gripper [27], [28], [29].

In this study, we propose a novel perception system that can perceive the given grasp task and accordingly infer the intended grasp posture of the prosthetic hand user: Proximity perception-based grasping intelligence (P2GI). The P2GI system was designed to collect the point cloud of a target object through an embeddable sensor system and make prompt decisions regarding the grasp strategy. The proposed system offers three main benefits to the users: the ability to utilize various grasp postures with a simple grasp intention, prompt decision-making with a short delay that allows for seamless control, and robust performance in varying environmental conditions. In this study, we demonstrated the P2GI system using a 16 DOF robotic hand

grasping unknown objects with multiple grasp postures under various scenarios.

II. SYSTEM DEVELOPMENT

The methodology of the P2GI system is demonstrated in Fig. 1. The proximity sensors are positioned on the palmar side of the prosthetic hand and measure the distance between the object and each sensor as the prosthetic hand approaches the object. By localizing the proximity sensors in the global space, the distance data from each sensor represents a point in the global space. These points are accumulated to generate a point cloud of the object surface, which is then tracked in real time. Note that the scanning is done simultaneously during the reach-to-grasp motion of the user.

The algorithm of the P2GI system was designed to infer the hand-object relation using the point cloud and accordingly infer the intended grasp posture of the user. There is a high correlation between the intended grasp posture and the resultant hand-object relation [30], [31]. For instance, when a human intends to perform the power grasp, they position their hand close to the object until the palm and object almost touch each other. Similarly, for a precision pinch, the fingertips tend to be closer to the object than the palm. By using this principle, the P2GI system determines the grasp posture that matches the user's intention. Furthermore, the P2GI system estimates the size and relative position of the target object using the point cloud data. Based on the decisions, the P2GI system calculates the contact points for the stable grasp and corresponding finger motion path. Consequently, when the user gives the grasp command via a simple user interface, such as single-channel EMG, the prosthetic hand can grasp the target object stably with the intended grasp posture.

A. Hardware of the P2GI System

The P2GI system was implemented on an Allegro hand (Wonik Robotics, Seongnam, Korea) (see Fig. 2). The Allegro hand has 16-DOF on the opposable thumb and three fingers, enabling it to perform various grasp postures [32]. Note that the
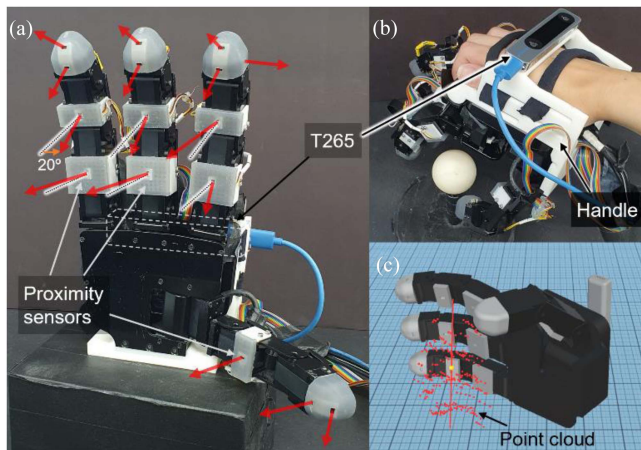
Fig. 2. P2GI system implemented on the allegro hand. (a) Palmar side of the robotic hand. There were 16 proximity sensors and red arrow represents the orientation of each sensor. (b) Dorsal side of the robotic hand. (c) Scanned point cloud when grasping the cylindrical object.

P2GI system is an accessory system that provides grasping intelligence to the prosthetic hand and can be modified and applied depending on the characteristics of the prosthetic hand such as DOF, joint schematics, and size. The system was evaluated with ten non-disabled subjects and the subjects donned the robotic hand using a handle attached to the dorsal side of the robotic hand [see Fig. 2(b)]. The handle was designed to minimize the gap between the wrist side of the robotic hand and the user's wrist.

The proximity sensors were placed on the palmar side of the prosthetic hand [see Fig. 2(a)]. The sensors were integrated into the silicone skin structure. In this study, VL6180X (STMicroelectronics, Geneva, CH), which is a time-of-flight type proximity sensor that can sense an object within ∼10 cm, was used. This sensor provides robust measurement with a 2 mm noise level for various surface and ambient light conditions.

The number and arrangement of the sensors can be customized according to the configuration of the prosthetic hand. In this study, 16 sensors were placed on the palmar side of the hand [see Fig. 2(a)]. Each segment of the fingers has one proximity sensor and each fingertip has two sensors. The proximity sensors on the finger segments are tilted 20 degrees from the perpendicular direction to avoid overlapping sensing points on the same finger. For the fingertip, one sensor heads toward the palmar side and the other heads toward a 45-degree tilted direction to explore a wider area. In addition, an extra sensor was placed on the lateral side of the index finger's fingertip to detect the object in that direction.

The palmar side of the prosthetic hand has limited space as it serves as the direct interface where the object contacts with the hand, and the motors and link mechanisms are compactly placed in this area. Hence, studies that utilize an RGB-D camera for object recognition typically place the camera on the wrist or dorsal side of the prosthetic hand [20], [22] or place a single miniature camera on the thenar eminence part [21]. However,

the proximity sensor unit used in this study has a small form factor with a thickness of only 1.7 mm (see Supplementary Fig. 1). This allowed us to embed multiple sensors into the skin structure that covers the palmar side of the prosthetic hand. By placing multiple proximity sensors in different locations, the P2GI system is able to observe the target object from multiple points of view; even if the object is close to the palmar side of the prosthetic hand and blocks the field of view of some sensors.

To obtain the pose of the prosthetic hand, a T265 (Intel RealSense, CA, USA) was employed as a pose-tracking sensor, enabling a portable and fully embeddable system. The T265 pose tracking sensor was placed at the dorsal side of the prosthetic hand and tracked the real-time pose of the hand [see Fig. 2(b)]. The position tracking error is less than 5% and the drift about the cyclic motion is less than 1% when conducting the reach-to-grasp motion.

B. Point Cloud Mapping

Using the pose tracking data and joint angles of the prosthetic hand, each sensor's global position and orientation were tracked in real time. A virtual model mirroring the real-time posture of the prosthetic hand was used, and the point cloud was mapped in the virtual space [see Fig. 2(c)]. The virtual space was updated at over 100 Hz frame rate, and each sensor value was synchronized via this virtual space.

Upon mapping the point cloud, three rules were applied to refine the acquired point cloud. First rule concerned the minimum distance between each point. If newly sensed points were closer to existing points than the minimum distance, they were not added to the point cloud, ensuring uniform density in a scanned area. The minimum distance was determined based on the computational power of a system and scale of the target object group [see Supplementary Note 1 for more details].

The second rule ensured that only points within the graspable space were considered as part of the target point cloud. To achieve this, two cubic-shaped areas were positioned on the palmar and lateral sides of the prosthetic hand, defining the graspable space [see Supplementary Note 2 and Supplementary Fig. 2]. This rule allowed for the rejection of points that may have been accumulated when the prosthetic hand passed by nontarget objects during the reach-to-grasp motion, thus avoiding their inclusion in the point cloud data.

The third rule was designed to distinguish points near the bottom side as those belonging to the floor. In daily life, objects are often placed on surfaces such as a desk or shelves. In these cases, the proximity sensors may detect not only the object but also the surface on which it rests. Assuming that the floor is a flat and horizontal surface, points belonging to the floor exhibit similar height and are located at the bottom side of the overall point cloud data. Consequently, the point clouds of the floor and target object were differentiated using the height of the points [see Supplementary Note 3 and Supplementary Fig. 3 for the distinguishing algorithm]. This rule enabled the P2GI system to identify the target object even when a small object was placed on the floor.
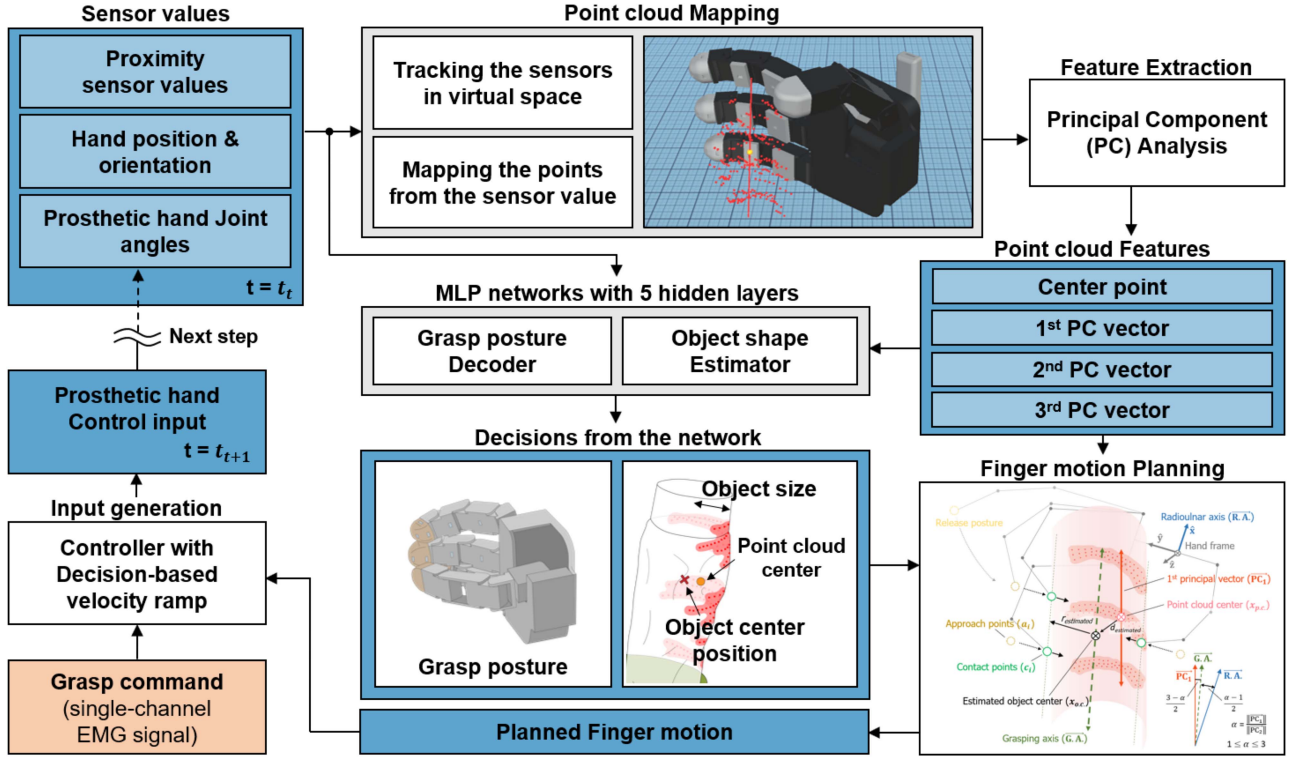
Fig. 3. Algorithm framework of the P2GI system. Using the sensor values gathered from the hardware, the point cloud of the object surface is mapped in a virtual space. The principal components (PCs) of the point cloud are obtained using the PCA and used as the features that represent the point cloud. Using the point cloud features and measured sensor values, the grasp posture decoder inferred the intended grasp posture, and the object shape estimator estimates the size of the object and offset between the center position of the object and the center point of the point cloud. Based on these decisions, the path planning algorithm plans the optimal finger motion path for a stable grasp. Following the planned finger motion path, the controller generates the control input as the grasp command is given from the user via single-channel EMG signal.

C. Real-Time Decision-Making and Control Algorithm

Fig. 3 presents the algorithm framework of the P2GI system. Once the object point cloud is mapped through the P2GI system hardware and point cloud mapping algorithm, shape-related features are extracted from the point cloud using principal component analysis (PCA). These extracted features serve as input for the grasp posture decoder and object shape estimator, which were designed to determine the intended grasp posture and estimate the object size and center position. Through this feature extraction, the point cloud is converted to the features that directly indicate the hand-object relation and overall shape of the object; while the minor characteristics of the point cloud are ignored. Consequently, the P2GI system, which perceives shape-related characteristics of the object, can handle unknown objects using the knowledge acquired from the representative shapes included in the training object set.

The posture decoder and shape estimator make the real-time decisions regarding the grasp strategy and the finger motion planning was executed accordingly. The entire process took less than 25 ms in this study, while using a single PC (with an i5-4690 CPU, one NVIDIA GTX750Ti GPU, and 32-GB RAM) which has a similar computational power to that of a modern smart-phone. While the P2GI system performs real-time finger motion planning, the prosthetic hand user can give a grasp command as desired. In this study, a single-channel EMG was employed to convey the user's grasp intention. The user controlled the closing and opening of the prosthetic hand. The motion of the prosthetic hand followed the planned finger motion path determined by the P2GI system. The decision-based velocity ramp [33] was applied for the effective EMG velocity control [see Supplementary Note 4].

D. Structure of the P2GI System Decoders

The grasp posture decoder employed a multilayer perceptron (MLP) classifier with five hidden layers [see Supplementary Note 5 for detailed structure]. Network inputs included features extracted from the point cloud data and other features, such as proximity sensor values. The primary goal of these input features was to represent the current hand-object relation between the prosthetic hand and the target object.

The point cloud's center and principal components obtained through PCA were used as features representing the acquired point cloud. They were calculated with respect to the prosthetic hand's coordinate system, indicating the relative pose of the target object. In particular, the principal components of the point cloud represent the object's shape and orientation. A larger principal component value signifies that the object is relatively large. The ratio between each principal component reveals the object's symmetry, enabling inferences about whether the object is closer in shape to a sphere or a bar.

The sensed distances from the proximity sensors were used to directly indicate the target object's current position. If the object is located at the distal side of the prosthetic hand, the sensors placed at the fingertip sense the object.

During grasping, the hand posture changes, causing the proximity sensor values to change despite the hand-object relation remaining the same. Thus, a grasp phase variable was used to indicate the degree of openness and closure of the prosthetic hand. This variable was set to a value between 0 and 1, where 0 indicates the fully open state, and 1 indicates the fully closed state irrespective of the grasp posture.

Consequently, the input vector for the grasp posture decoder had a size of 32×1 and contained the following components: relative center position of the point cloud (3×1), the first, second, and third principal vectors and values (three 4×1), distance values from the 16 sensors (16×1), and the grasp phase variable (1×1). The MLP classifier was trained using these features. The output classes were the three grasp postures used in this study.

Note that each input feature contributed to fully describing the current situation. According to the ablation study [see Supplementary Note 6 and Supplementary Figs. 4 to 7], without the point cloud features, the posture decoder became confused when the target object was small so that only a few or no sensors sensed the object. Consequently, the overall accuracy decreased approximately 16.5%. Without the sensed distances from the proximity sensors, the decoders were easily overfitted to the training dataset so that the generalization ability of the decoder decreased. Consequently, the overall accuracy decreased approximately 9.1%.

For the object shape estimator, which estimates the size and center position of the object, an MLP regression network with five hidden layers was used [see Supplementary Note 5 for detailed structure]. The input features used for the grasp posture decoder were also used for the object shape estimator. In addition, the joint angles of each joint were used to indicate the current hand posture. Note that these joint angles should not be used for the grasp posture decoder because they are strongly related to the previous decisions made by the posture decoder. Thus, using joint angles as the input feature can bias the posture decoder to follow the previous decision, even if it was incorrect. As the Allegro hand used in this study has 16 joints, the input vector for the object shape estimator had a size of 48×1, consisting of the 32×1 size input vector for the grasp posture decoder and the 16×1 size joint angle vector.

Based on these input features, the MLP regression network estimated the size of the object and the offset between the center of the point cloud and the exact center position of the object. The point cloud center was calculated as the average position of all the points within the point cloud. When the prosthetic hand approached the object, the object surface that faced the palmar side of the prosthetic hand was dominantly scanned. Consequently, there was an offset between the center point of the scanned point cloud and the exact center position of the object. The object shape estimator was used to estimate this offset. Using the estimated offset and point cloud center, the P2GI system tracks the object's center in real time.

SC : Thin Cylinder (D=20mm), MC : Medium Cylinder (D=50mm), LC : Thick Cylinder (D=80mm)
SS : Small Sphere (D=20mm), MS : Medium Sphere (D=50mm), LS : Large Sphere (D=80mm)
PL : Plate (T=10mm), Red border : Power, Green border : Precision, Blue border: Lateral

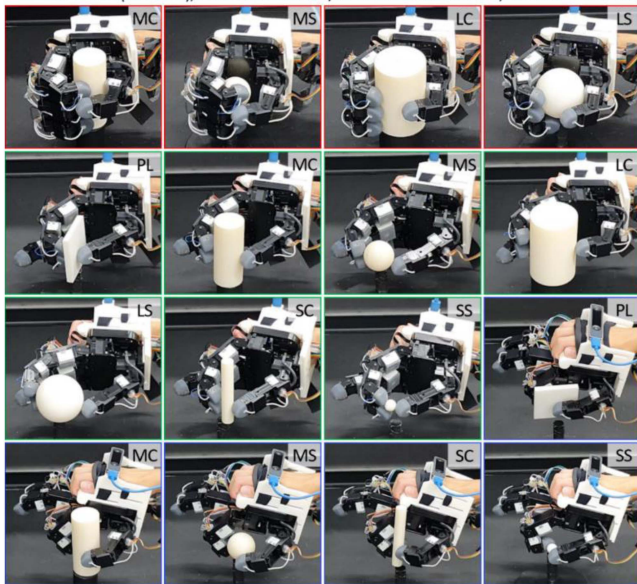Fig. 4. Sixteen Ggasp tasks for the data acquisition. Three grasp postures, power grasp (red border), precision pinch (green border), and lateral pinch (blue border), were used. Seven different objects were used as the target object: three cylinders (SC, MC, and LC) and three spheres (SS, MS, and LS) of different sizes, and a thin plate (PL).

In terms of the network output, the object size was assigned as the diameter or thickness of the object, while the offset value was calculated based on the object's geometry. Under the assumption that the P2GI system scans one side of the object, the offset was determined by calculating the centroid of half of the object's surface [see Supplementary Note 7].

E. Training of the P2GI System Decoders

To train the decoders, 16 grasp tasks representing daily life situations were selected (see Fig. 4). According to the research on human grasp taxonomy [34], [35], there are three primary grasp postures: the power grasp, precision pinch, and lateral pinch. These grasp postures account for more than 70% of human grasp tasks. In addition, the Allegro hand used in this study can perform these three postures. Thus, the P2GI system was designed to handle these three postures in this study. The target postures managed by the P2GI system can vary depending on the possible grasp postures of a prosthetic hand.

In terms of the object set, a diverse collection of objects with varying PCA characteristics was used. According to the research that categorizes objects in daily life [35], [36], object shapes can be classified as equant, prolate, oblate, and bladed objects. It is notable that shape-related characteristics of an object set depend on each individual's living environment. Consequently, the object set for the network training was designed to encompass a broad range of object taxonomy, rather than selecting specific objects from daily life. Seven objects were included: small sphere, medium sphere, large sphere, thin cylinder, medium cylinder, thick cylinder, and plate. Equant objects

were represented by spheres, prolate objects were represented by cylinders, and oblate and bladed objects were represented by the thin plate. Roughness variation was excluded to minimize the number of objects in the object set. The objects' sizes ranged from 20 to 80 mm in diameter, similar to those used in previous studies [5], [37], [38]. Small and thin objects had a 20 mm diameter, medium objects had a 50 mm diameter, and large and thick objects had an 80 mm diameter. All seven objects were grasped using the three grasp postures, except for certain cases such as the lateral pinch on the thick cylinder, considering the object size. Although numerous irregular objects exist in daily life, such as a mayonnaise jar, this object set can represent the shape of each subpart of the object where the user intends to grasp.

To collect the training datasets, the experimenter conducted the 16 grasp tasks shown in Fig. 4 while donning the prosthetic hand. Initially, there was no decoder capable of making grasp decisions. Thus, the experimenter conducted the grasp tasks and collected the dataset while the P2GI system followed preassigned decisions. This dataset was used to train the initial decoder. Subsequently, the experimenter repeated the grasp tasks while the P2GI system made real-time decisions using the initial decoder. In this case, the misclassifications and grasp failures could occur, and they were included in the new dataset for training a baseline decoder that reflects the experimenter's behavioral characteristics. The baseline decoder was then used during the first evaluation, in which ten subjects first used the P2GI system and conducted the 16 grasp tasks. The subject-specific datasets were collected, which reflect each user's unique characteristics. These datasets were used to train personalized decoders for each individual. In the subsequent evaluations, the personalized decoders replaced the baseline decoder [see Supplementary Note 8 and Supplementary Fig. 8 for more details about this step-by-step training process].

F. Finger Path Planning

Based on the decision from the grasp posture decoder, the prosthetic hand's target posture was controlled as the user's intended grasp posture. Moreover, the object shape estimation from the object shape estimator was utilized for the finger path planning. The path planning strategy may vary depending on the DOF and configuration of the prosthetic hand. In this study, the object size and center position were tracked in real time, and the object pose was estimated via principal components of the point cloud. Based on the estimation, while conducting the inferred grasp posture, the contact points were aligned near the estimated object center, and the amount of hand opening was adjusted according to the estimated object size [see Supplementary Note 9 and Supplementary Fig. 9].

III. System Evaluation

The performance of the P2GI system was quantitatively evaluated by conducting various grasp tasks, including 16 grasp tasks shown in Fig. 4 and additional tasks with six unknown objects. Ten right-handed subjects (three females and seven males whose average age was 23.3 ± 2.6 years old) participated in the study.

None of the subjects had prior experience with the myoelectric control systems before the study commenced. The study was approved by the Institutional Review Board of Korea Advanced Institute of Science and Technology (No. IRB-21-087).

A. Evaluation Session Design

At the beginning of the evaluation session, the experimenter provided an explanation of the P2GI system and demonstrated it. A single-channel sEMG sensor was then attached to the flexor muscle of the subject's left forearm. The EMG sensor was calibrated by measuring the maximum voluntary contraction. Subsequently, the subject donned the robotic hand on their right hand. By contracting and relaxing their left forearm, the subjects controlled the closing and opening of the robotic hand at the desired speed. To prevent muscle fatigue from influencing grasp intention, the sEMG sensor was attached to the opposite hand that moves the robotic hand, and some subjects held an object (i.e., a folding hex key set) in their hand instead of merely clenching their fist.

Following the initial setup, a 30-minute practice period ensued. The subjects performed grasp tasks under the experimenter's guidance. The objects were positioned 50 cm away from the subject's torso [see Supplementary Fig. 10], and the experimenter instructed which grasp posture should be used. At the beginning, subjects learned the basics of the P2GI system by performing power grasp, precision pinch, and lateral pinch on the medium cylinder. Then, they performed grasp tasks in the order of grasp tasks shown in Fig. 4. Each grasp task was performed twice or more, depending on the subject. Some subjects spent the entire practice time for this step, while others completed it earlier and spent the remaining time freely conducting tasks they wanted to perform.

After that, the evaluation commenced. For the first part, subjects were asked to conduct the 16 grasp tasks they had practiced. Each task was conducted five times consecutively. Completing all grasp tasks took approximately 20–30 min.

For the second part, the subjects were asked to grasp activities of daily living (ADL) objects that were not included in the training object set. As shown in Fig. 5, a coke can, mayonnaise jar, Rubik's cube, pencil, business card, and file were used. The Rubik's cube was lubricated, making it easily mixed up and, thus, more challenging to grasp. The pencil was placed in a pen holder, allowing for ± 10-degree wobbling due to the slack between the pencil and the holder. The business card was secured in a hand-shape holder, and the file was placed in a file holder. The subjects were instructed to perform a power grasp on the coke can and mayonnaise jar, a precision pinch on the Rubik's cube and pencil, and a lateral pinch on the business card and file. For each ADL object, the grasp trials were conducted five times consecutively.

In this study, each subject participated in three evaluation sessions, spaced 3 to 7 days apart. In the first evaluation session, the baseline decoder, trained using the experimenter's movement data, was employed for all subjects. For the second evaluation, the personalized decoders were trained using movement data collected from the subjects during the first evaluation. Each
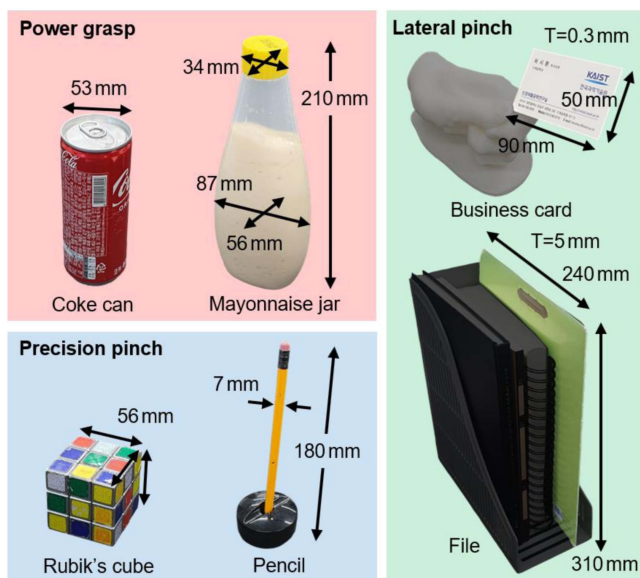
Fig. 5.   ADL objects used for the evaluation.

Fig. 6.   Grasp posture classification accuracy of the P2GI system for 16 evaluation tasks.

subject then performed the evaluation tasks using their personalized decoder in the second session. For the third session, the personalized decoders were trained again using the movement data from both the first and second evaluations and were used during the third evaluation session. Note that during the second and third evaluation sessions, the practice session served only as a reminder of the system and was kept under 10 min.

Supplementary video 1 shows the entire grasp tasks conducted during the evaluation consecutively, with the real-time point cloud mapping and subsequent decisions.

B. Evaluation Measures

The grasp posture classification accuracy, object size estimation error, and task success rate were used as the evaluation measures. The grasp posture classification accuracy is the ratio of correct grasp posture decisions during the real-time grasp trial; and the object size estimation error refers to the estimation error when the object shape estimator estimates the size and the amount of offset between the object center and point cloud center. The task success rate is the ratio of the successful grasp trials among the entire grasp attempts.

The grasp posture decision was made in real time throughout the entire grasp trial. However, since the closing and opening of the prosthetic hand were controlled via the EMG signal, there were moments when the fingers moved in accordance with the decisions and moments when the fingers did not move or moved independently of the decisions. Thus, only decisions that genuinely influence the finger motion were considered when evaluating the classification accuracy.

The object size was estimated at every time step, and the average was used as a representative value to indicate the estimation for each grasp trial. Moreover, there were five grasp trials for each grasp task. Thus, the performance during these
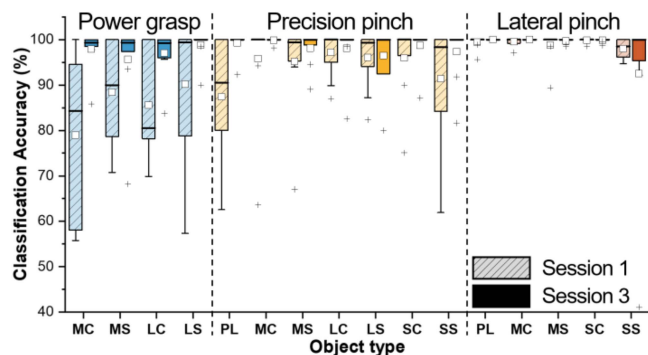
five grasp trials was averaged to represent the performance when each subject used the P2GI system.

Although the classification accuracy and estimation error are good, the resultant success rate could be low if the system is difficult to use. In this study, only trials that stably grasped and lifted an object using the correct grasp posture were counted as successful trials. The trials that missed the object during the grasp or succeeded but employed a different grasp posture were counted as failed trials. For each grasp task, there were five grasp trials, and the task success rate represented the ratio of the successful trials among the entire grasp trials.

The numerical equations for the data processing and plotting methods are described in Supplementary Note 10.

C. Evaluation Results–16 Grasp Tasks Using Known Objects

In most grasp tasks, the grasp posture classification accuracy during the trials was >90% [see Fig. 6 and Supplementary Fig. 11]. The average classification accuracy was 93.6% for the first trial employing the baseline decoder and increased to 98.1% for the third trial employing the personalized decoder. With the personalized decoder, seven subjects showed an average classification accuracy of over 98%, and the subject with the lowest classification accuracy showed 95.9% average classification accuracy. Note that the classification accuracy was significantly improved after the decoder personalization [see Supplementary Fig. 12].

In most cases, the object size estimation error was less than 15 mm, particularly after the personalization [see Fig. 7 and Supplementary Fig. 13]. Before the personalization, there was a tendency for size underestimation, which was mitigated following the personalization. During the first session employing the baseline decoder, the average estimation errors, calculated via the root mean square method, were 11.2 mm for the size, and 4.1 mm for the offset between the object center and point cloud center. During the third session employing the personalized decoder, the average estimation errors were 6.3 mm for the size and 2.4 mm for the offset.

The average task success rate was 85.4% during the first session employing the baseline decoder but increased to 93.1%

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE/ASME TRANSACTIONS ON MECHATRONICS
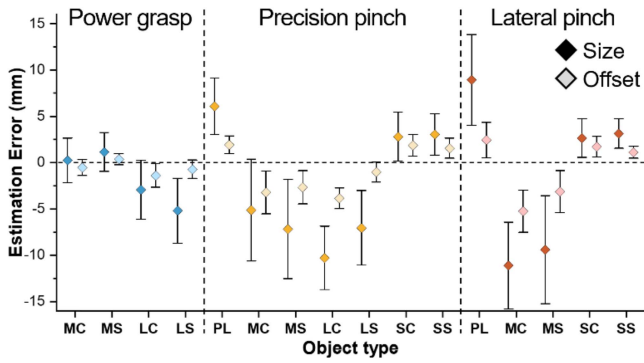

Fig. 7.   Object shape estimation error of the P2GI system with the personalized decoder, during the session 3. The whisker represents 95% confidence interval.
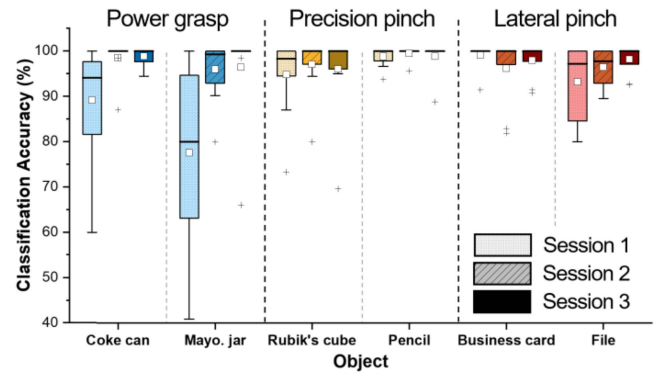

Fig. 9.   Grasp posture classification accuracy of the P2GI system for six grasp tasks with unknown ADL objects.
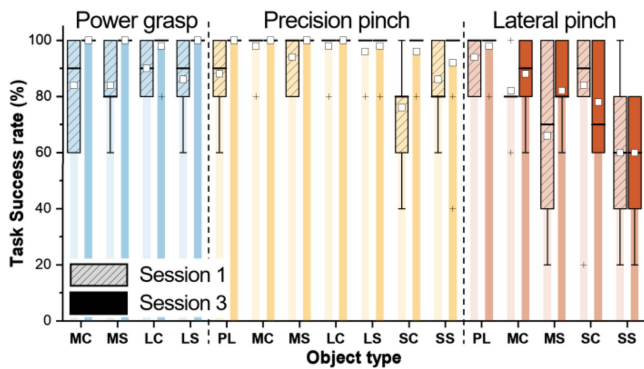

Fig. 8.   Task success rate when using the P2GI system for 16 evaluation tasks.

Fig. 10.   Shape estimation results of the P2GI system for six unknown ADL objects, during the session 3.

during the third session employing the personalized decoder. Nine subjects showed a success rate of >90%, and one showed the success rate of 85% [see Fig. 8 and Supplementary Fig. 14]. Among the grasp tasks, the failures mostly occurred during the lateral pinch, due to the difficulty of the task. During the first evaluation when the baseline decoder was used and the subject first used the prosthetic hand with the P2GI system, the failures also occurred for power grasp and precision pinch. However, these issues were mitigated as the decoder was personalized for each subject.

D. Evaluation Results–Grasp Tasks Using ADL Objects

The evaluation using the ADL objects, which were not included in the training object set, was conducted to verify that the P2GI system can generalize the knowledge acquired from the seven basic objects to various unknown objects.

Similar to the evaluation results with the known objects, the grasp posture classification accuracy was 92.1% for the first session employing the baseline decoder and 97.8% for the third session employing the personalized decoder [see Fig. 9]. With the personalized decoder, eight subjects showed an average classification accuracy of >97% and the subject with the lowest classification accuracy showed 92.6% average classification accuracy.
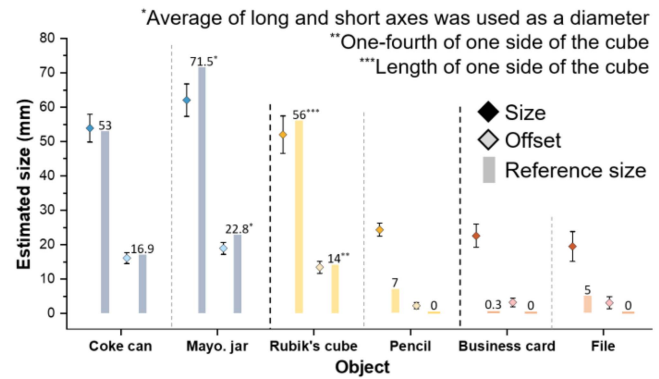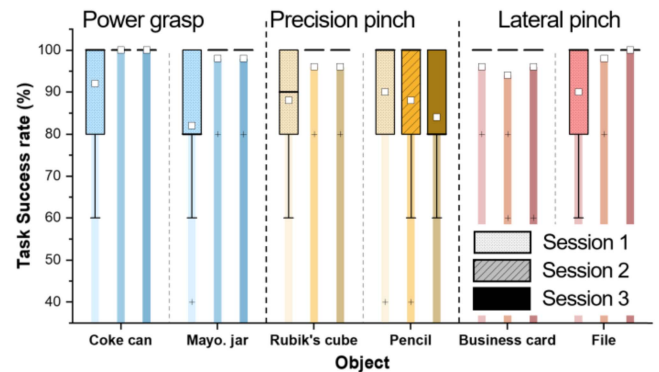

Fig. 11.   Task success rate when using the P2GI system while grasping six unknown ADL objects.

Regarding the size estimation, as the size of the ADL objects could not be represented by a single parameter, the estimated sizes, rather than the errors, are presented in Fig. 10 and Supplementary Fig. 15. Fig. 10 shows the average shape estimation result from ten subjects during the third evaluation. The reference sizes were derived according to the objects' dimensions shown in Fig. 5. In comparison to the reference sizes, the object shape estimations were successfully carried out for ADL objects that were not included in the training object set. Note that since
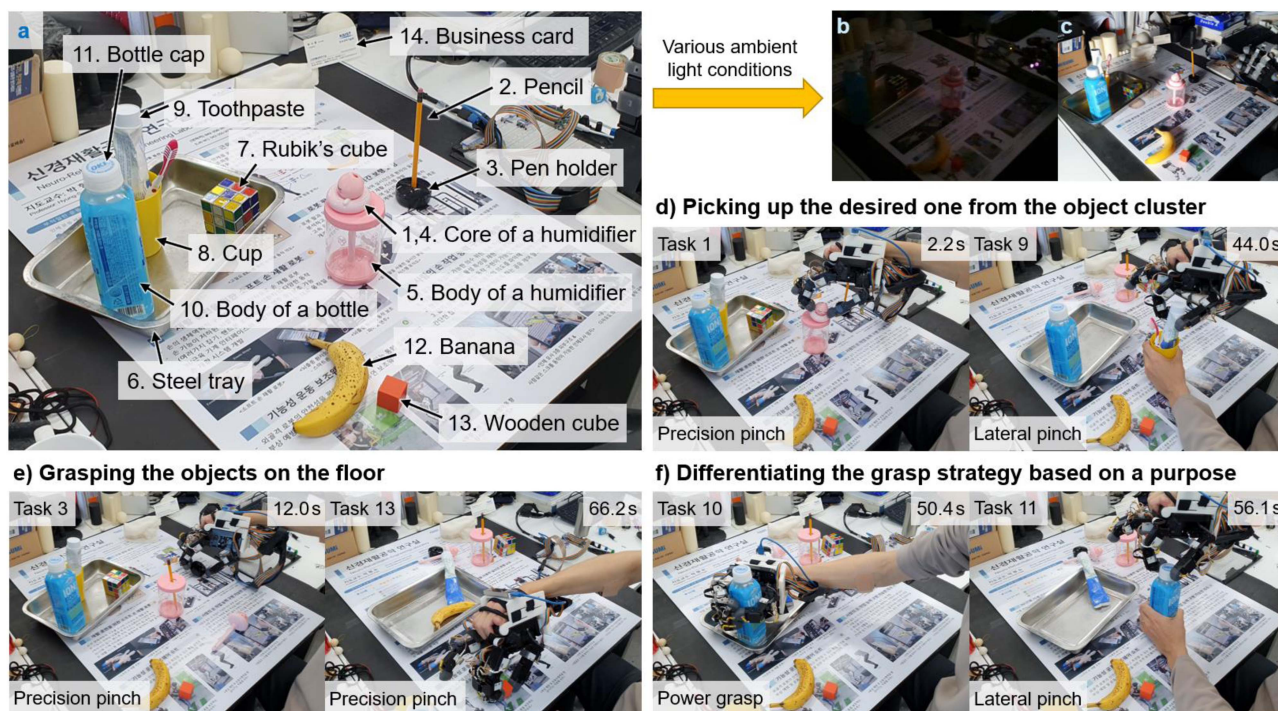
Fig. 12. Series of object manipulation tasks using unknown ADL objects. (a) Initial setup for the demonstration. The number represents the manipulation order. The manipulations were done under the various ambient light conditions: indoor light, (b) low light, and (c) directional light. (d) User can pick up the desired object from the object cluster. (e) User can grasp a small object on the floor. (f) User can vary the grasp strategy depending on a purpose. The time represented on the right-top corner represents the time after the demonstration begins.

the diameter of the smallest objects in the training object set was 20 mm, the size estimation for small objects appeared to be limited to 20 mm.

Consequently, the average task success rate was 89.7% during the first session, and increased to 95.7% during the third session [see Fig. 11]. Nine subjects showed a success rate of >90% and one subject showed a success rate of 86.7%. Since there was no instance of conducting the lateral pinch on the small sphere, which was the most difficult task, the overall success rate was even higher than in the evaluations from the previous section. The object with the lowest success rate was the pencil placed in the pen holder; as it was thin and prone to swaying during the grasp task.

E. Robustness Regarding Various ADL Situations

Given that the prosthetic hand serves as a body part for amputees, it must be able to handle various situations encountered in our daily lives. As the perception system targets daily living environments, the P2GI system should operate robustly irrespective of the object type, surrounding objects, and environmental conditions.

In this section, the robustness of the P2GI system is demonstrated by using various unknown ADL objects and environmental conditions. It is important to note that the baseline decoder, which was trained using the limited dataset collected using the seven basic objects in a controlled environment, was employed for the demonstration. Fig. 12(a) shows the initial setup for the demonstration. Thirteen ADL objects were presented together

in front of the user. Some objects were placed on the floor, while other objects were clustered or assembled together. For the cluttered surrounding environment, the banner containing various figures was used as the floor. In addition, for the variations in the ambient light conditions, the same manipulations were conducted under both low and directional light conditions [see Fig. 12(b) and (c)]. The experimenter performed the demonstration while attaching a single-channel sEMG sensor to the right forearm, where the prosthetic hand was donned.

Under the given ADL situation, a series of manipulation tasks was performed in the following order:

1) picking up the core unit from the body of the humidifier and placing it on the floor;
2) picking up the pencil from the pen holder and pegging it in the body of the humidifier;
3) picking up the pen holder and placing it outside;
4) picking up the core unit of the humidifier and placing it outside;
5) grasping the body of the humidifier and placing it outside;
6) pulling the steel tray containing the objects by pinching the edge of the tray;
7) picking up the Rubik's cube on the tray and placing it outside;
8) grasping the cup containing the toothbrush and toothpaste tube and placing it in front of the user;
9) picking up the toothpaste tube from the cup and placing it on the tray;
10) grasping the bottle and placing it in front of the user;

11) grasping the bottle cap and opening the bottle;
12) picking up the banana and placing it on the tray;
13) picking up the wooden cube and placing it on the tray;
14) picking up the business card from the hand-shape holder.

Each task was conducted consecutively without switching modes or triggering between the tasks.

Supplementary videos 2 and 3 show the demonstrations under the indoor light condition and low and directional light conditions, respectively. Snapshots taken during the demonstration are displayed in Supplementary Fig. 16. Tasks 1, 7, 8, 9, and 10 show that the experimenter picked up or grasped the object which was clustered with other objects. The P2GI system can easily distinguish the target object from the object cluster, as the P2GI system perceives only the object within the graspable space [see Fig. 12(d)]. Tasks 3, 4, 12, and 13 show that the P2GI system can make reliable decisions regarding the small objects on the floor [see Fig. 12(e)]. Tasks 10 and 11 shown in Fig. 12(f) show that the decisions from the P2GI system varied depending on the user's intention. As the user varies the grasp strategy and approach movement based on the purpose, the P2GI system makes the appropriate decisions accordingly. In addition, Task 6 shows that the P2GI system can make decisions by grasping the edge of the steel tray using the precision pinch, by just approaching the prosthetic hand system toward the edge of the steel tray where the target segment. Tasks 9 and 11 were conducted as dual-arm tasks.

Note that the ADL objects and environmental conditions employed in this demonstration were not used for the decoder training. Thanks to the P2GI system hardware, which gathers the point cloud of the object using the proximity sensors, the limited dataset collected with basic objects and controlled environment was sufficient to handle various untrained ADL objects and environmental conditions. Furthermore, the decisions were made continuously and promptly during the series of manipulation tasks. For each manipulation task, the reach-to-grasp and grasp motion took approximately 2.3 s. Throughout the demonstration, the user seamlessly controlled the prosthetic hand using a simple EMG interface; while the decisions regarding the grasp motion were made in real time within a 25 ms delay.

IV. Conclusion

The performance and potential of the prosthetic hand system with the P2GI system were demonstrated through quantitative evaluation and demonstration involving various ADL objects. The P2GI system with the personalized decoder showed a grasp posture classification accuracy of >95% and the subjects using the prosthetic hand with the P2GI system achieved an average task success rate of 93.1% while handling multiple grasp postures for various target objects. The performance was maintained for the six untrained ADL objects. The subjects only needed to contract and relax their forearm muscles to control the closing and opening of the prosthetic hand. The decisions were promptly made in real time as the prosthetic hand approached the target object. The entire grasp procedure was carried out seamlessly. Furthermore, the prosthetic hand users could grasp small and thin objects without much attention, thanks to the real-time estimation of the size and center position of the target object.

The P2GI system utilizes proximity sensors that can be embedded in the prosthetic hand, making it portable and operational without externally located sensors such as a camera. Although the current system in this study employed a PC for the sensor signal acquisition and point cloud mapping in the virtual space, the required computational power is small enough to be managed by a smartphone. The program used for the point cloud mapping was developed using UNITY, which is widely used in mobile application development.

Moreover, the data size managed by the P2GI system is small enough to be processed in real time. The P2GI system implemented in this study collects 38×1 data for every frame, which comprises 16 proximity sensor values, 16 joint angles, and 3-axis global location and 3-axis global orientation of the robotic hand. This is significantly smaller compared to typical RGB-D images, which have a $640 \times 480 \times 4$ data size. In this study, the P2GI system made real-time decisions at an approximately 40 Hz sampling rate while using a single PC (with an i5-4690 CPU, one NVIDIA GTX750Ti GPU, and 32-GB RAM), which has similar computational power comparable to that of a modern smartphone. Consequently, decision-making takes less than 25 ms, considerably faster than the 300 ms reported as the allowable latency for the seamless control of a prosthesis [10].

One of the most significant advantages of the P2GI system is its ability to achieve high robustness with a small dataset collected in a controlled environment. Without requiring a large size dataset featuring various objects and environmental conditions, the P2GI system can generalize the knowledge trained from the limited dataset to a wide range of ADL conditions encountered in daily lives. Moreover, the prosthetic hand is a personalized system that necessitates individualized tuning to match each user's movement characteristics and preferences. In this regard, the requirement of a small data size is also beneficial for the convenience of the personalization. In this study, for the personalization of the P2GI system, the data collection took approximately 30 min using seven basic objects in a controlled environment. With this limited dataset, the performance of the P2GI system was personalized and improved in both classification accuracy and task success rate.

Compared to the decision-making systems employing the RGB-D camera, the proposed system is free from their recurrent challenges; such as image occlusion, the necessity of high computational power and resultant high latency, and the necessity of large amounts of dataset for robust and general performance. In this article, we demonstrated that the proposed system can make reliable decisions in real time about various unknown objects in untrained environmental conditions; while employing a much smaller dataset compared to the dataset collected for the RGB-D camera-based decision-making system [20], [21], [22].

The current stage of the P2GI system uses the basic MLP networks as the decision-making decoders. Thanks to the P2GI system hardware acquiring the input features highly correlated to the output decisions, these basic MLP networks provide accurate and prompt decisions regarding the intended grasp posture and object size. Both point cloud features and sensor value features were successfully collaborated through the MLP network. However, based on the purpose, much research could be

driven related to the decoders. For instance, not only estimating the size of the object but also identifying the edge or valley of the object could be possible by using the acquired point cloud data. For a simplified system, the state-machine using the point cloud features and sensor values as the parameters could be applied instead of the data-driven method. While maintaining the low latency of the system, more complex neural networks could be developed to handle more sophisticated decisions aimed at various ADL manipulations, including in-hand manipulation.

In this study, the Allegro hand was employed as the high-DOF prosthetic hand system that utilizes the P2GI system. It has sufficient dexterity to demonstrate the advantages of the P2GI system, but relatively heavy weight necessitated that the EMG sensor was attached to the opposite arm to avoid fatigue during the evaluation with subjects. Note that the P2GI system does not require the EMG sensor to be attached to the opposite arm. The further evaluations including more ADL and dual-arm tasks would be possible when the P2GI system was applied to other compact prosthetic hand systems with an appropriate weight, as the experimenter did during the ADL demonstration.

REFERENCES

[1] M. Controzzi, F. Clemente, D. Barone, A. Ghionzoli, and C. Cipriani, "The SSSA-MyHand: A dexterous lightweight myoelectric hand prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 5, pp. 459–468, May 2017.

[2] U. Kim et al., "Integrated linkage-driven dexterous anthropomorphic robotic hand," *Nature Commun.*, vol. 12, 2021, Art. no. 7177.

[3] W. Chen, H. Khamis, I. Birznieks, N. F. Lepora, and S. J. Redmond, "Tactile sensors for friction estimation and incipient slip detection-toward dexterous robotic manipulation: A review," *IEEE Sens. J.*, vol. 18, no. 22, pp. 9049–9064, Nov. 2018.

[4] M. Cheng, L. Jiang, F. Fan, B. Yang, J. Dai, and H. Liu, "Development of a multisensory underactuated prosthetic hand with fully integrated electronics," *IEEE/ASME Trans. Mechatron.*, vol. 28, no. 2, pp. 1187–1198, Apr. 2023.

[5] G. C. Matrone, C. Cipriani, E. L. Secco, G. Magenes, and M. C. Carrozza, "Principal components analysis based control of a multi-DOF underactuated prosthetic hand," *J. Neuroeng. Rehabil.*, vol. 7, 2010, Art. no. 16.

[6] O. M. Andrychowicz et al., "Learning dexterous in-hand manipulation," *Int. J. Robot. Res.*, vol. 39, pp. 3–20, 2020.

[7] J. A. George et al., "Biomimetic sensory feedback through peripheral nerve stimulation improves dexterous use of a bionic hand," *Sci. Robot.*, vol. 4, 2019, Art. no. eaax2352.

[8] A. W. Franzke et al., "Users' and therapists' perceptions of myoelectric multi-function upper limb prostheses with conventional and pattern recognition control," *PLoS One*, vol. 14, 2019, Art. no. e0220899.

[9] B. H. Young, "The bionic-hand arms race: High-tech hands are complicated, costly, and often impractical," *IEEE Spectr.*, vol. 59, no. 10, pp. 24–30, Oct. 2022.

[10] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[11] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 15, no. 1, pp. 111–118, Mar. 2007.

[12] B. Rodríguez-Tapia, I. Soto, D. M. Martínez, and N. C. Arballo, "Myoelectric interfaces and related applications: Current stage of EMG signal processing-A systematic review," *IEEE Access*, vol. 8, pp. 7792–7805, 2020.

[13] C. Amma et al., "Advancing muscle-computer interfaces with high-density electromyography," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 929–938.

[14] L. Pan et al., "Improving robustness against electrode shift of high density EMG for myoelectric control through common spatial patterns," *J. NeuroEng, Rehabil.*, vol. 12, 2015, Art. no. 110.

[15] T. A. Kuiken et al., "Targeted reinnervation for enhanced prosthetic arm function in a woman with proximal amputation: A case study," *Lancet*, vol. 369, pp. 371–380, 2007.

[16] A. L. Ciancio et al., "Control of prosthetic hands via the peripheral nervous system," *Front. Neurosci.*, vol. 10, 2016, Art. no. 116.

[17] I. Kyranous, S. Vijayakumar, and M. S. Erden, "Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses," *Front. Neurorobot.*, vol. 17, 2018, Art. no. 58.

[18] W. Li, P. Shi, and H. Yu, "Gesture recognition using surface electromyography and deep learning for prostheses hand: State-of-the-art, challenges, and future," *Front. Neurorobot.*, vol. 15, 2021, Art. no. 621885.

[19] K. Z. Zhuang et al., "Shared human-robot proportional control of a dexterous myoelectric prothesis," *Nature Mach. Intell.*, vol. 1, pp. 400–411, 2019.

[20] S. Dosen et al., "Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation," *J. Neuroeng. Rehabil.*, vol. 7, 2010, Art. no. 42.

[21] J. Strake et al., "Semi-autonomous control of prosthetic hands based on multimodal sensing, human grasp demonstration and user intention," *Robot. Auton. Syst.*, vol. 154, 2022, Art. no. 104123.

[22] Y. He, R. Shima, O. Fukuda, N. Bu, N. Yamaguchi, and H. Okumura, "Development of distributed control system for vision-based myoelectric prosthetic hand," *IEEE Access*, vol. 7, pp. 54542–54549, 2019.

[23] C. Shi, D. Yang, J. Zhao, and H. Liu, "Computer vision-based grasp pattern recognition with application to myoelectric control of dexterous hand prosthesis," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 28, no. 9, pp. 2090–2099, Sep. 2020.

[24] S. E. Navarro et al., "Proximity perception in human-centered robotics: A survey on sensing systems and applications," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1599–1620, Jun. 2022.

[25] S. Ye, K. Suzuki, Y. Suzuki, M. Ishikawa, and M. Shimojo, "Robust robotic grasping using IR net-structure proximity sensor to handle objects with unknown position and attitude," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, 3271–3278.

[26] S. Bazeille, V. Lebastard, S. Lanneau, and F. Boyer, "Model based object localization and shape estimation using electric sense on underwater robots," *IFAC-PapersOnLine*, vol. 50, pp. 5047–5054, 2017.

[27] D. Guo, P. Lancaster, L.-T. Jiang, F. Sun, and J. R. Smith, "Transmissive optical pretouch sensing for robotic grasping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 5891–5897.

[28] P. Lancaster, B. Yang, and J. R. Smith, "Improved object pose estimation via deep pre-touch sensing," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 2448–2455.

[29] Y. Suzuki, "Proximity-based non-contact perception and omnidirectional point-cloud generation based on hierarchical information on fingertip proximity sensors," *Adv. Robot.*, vol. 35, pp. 1181–1197, 2020.

[30] F. Wörgötter, E. E. Aksoy, N. Krüger, J. Piater, A. Ude, and M. Tamosiunaite, "A simple ontology of manipulation actions based on hand-object relations," *IEEE Trans. Auton. Ment. Develop.*, vol. 5, no. 2, pp. 117–134, Jun. 2013.

[31] D. Kim et al., "Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view," *Sci. Robot.*, vol. 4, 2019, Art. no. eaav2949.

[32] D.-H. Lee, J.-H. Park, S.-W. Park, M.-H. Baeg, and J.-H. Bae, "KITECH-hand: A highly dexterous and modularized robotic hand," *IEEE/ASME Trans. Mechatron.*, vol. 22, no. 2, pp. 876–887, Apr. 2017.

[33] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "A decision-based velocity ramp for minimizing the effect of misclassifications during real-time pattern recognition control," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 8, pp. 2360–2368, Aug. 2011.

[34] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The GRASP taxonomy of human grasp types," *IEEE Trans. Hum.-Mach. Syst.*, vol. 46, no. 1, pp. 66–77, Feb. 2016.

[35] T. Feix, I. M. Bullock, and A. M. Dollar, "Analysis of human grasping behavior: Object characteristics and grasp type," *IEEE Trans. Haptics*, vol. 7, no. 3, pp. 311–323, Jul.–Sep. 2014.

[36] T. Zingg, "Beitrag zur Schotteranalyse," Doctoral dissertation, ETH Zürich, Zürich, Switzerland, 1935.

[37] T. Feix, I. M. Bullock, and A. M. Dollar, "Analysis of human grasping behavior: Correlating tasks, objects and grasps," *IEEE Trans. Haptics*, vol. 7, no. 4, pp. 430–441, Oct.–Dec. 2014.

[38] J. T. Belter, M. T. Leddy, K. D. Gemmell, and A. M. Dollar, "Comparative clinical evaluation of the Yale multigrasp hand," in *Proc. IEEE 6th Int. Conf. Biomed. Robot. Biomechatronics*, 2016, pp. 528–535.
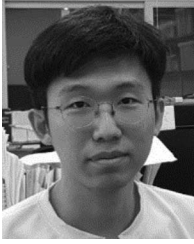
Si-Hwan Heo received the B.S., M.S., and Ph.D. degree from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2016, 2018, and 2023, respectively, all in mechanical engineering.

He is currently a Postdoctoral Researcher with the Korea Advanced Institute of Science and Technology. His research interest includes the prosthetic hand, shared autonomy, and hand rehabilitation.

Hyung-Soon Park (Member, IEEE) received the Ph.D. degree in mechanical engineering from KAIST, Daejeon, Korea, in 2004.

From 2009 to 2013, he was a staff scientist with Rehabilitation Medicine Department, National Instituted of Health, Bethesda, MD, USA. He is now a Professor with the Mechanical Engineering Department, KAIST. His current research interest include application on robotics and control technology for effective neuro-rehabilitation and study of neuromuscular impairments post brain injuries.