A Video Dataset of Everyday Life Grasps for the Training of Shared Control Operation Models for Myoelectric Prosthetic Hands

Ricardo V. Godoy[1], Bonnie Guan[1], Anany Dwivedi[2], Mahonri Owen[3], and Minas Liarokapis[1]

Abstract— **The control of high degree of freedom prosthetic hands requires high cognitive effort and has proven difficult to navigate particularly as prosthetic hands become more advanced. To reduce control complexity, vision-based shared control methods have been proposed. Such methods rely on image processing and/or machine learning to identify object features and classes and introduce a level of autonomy to the grasping process. However, currently available image datasets lack focus in the area of prosthetic grasping. Thus, in this paper, we present a new dataset capturing user interactions with a wide variety of everyday life objects using a fully actuated, human-like robot hand and an onboard camera. The dataset includes videos of over 100 grasps of 50 objects from 35 classes. A video analysis has been conducted, using established grasp taxonomies to compile a list of grasp types based on interactions of the user with the objects. The resulting dataset can be used to develop more efficient prosthetic hand operation systems based on shared control frameworks.**

I. Introduction

The evolution of prosthetic technology and robotic grasping has led to the development of electromechanical anthropomorphic hands that outperform their predecessors in terms of dexterity, aesthetics, and grasping capabilities. However, complex anthropomorphic prosthetic hands, such as the i-LIMB [1] and the Michelangelo hand [2] often are complex to control due to their many degrees of freedom. Additionally, users need to put more cognitive effort into the operation of such complex devices. To achieve a more simplified control of dexterous prostheses, shared control schemes can be employed [3].

Vision-based approaches, alongside muscle machine interfacing methods such as electromyography (EMG) based control, have thus been proposed to introduce a level of autonomy and improve prosthetic control systems. In such scenarios, a human operator is assisted by intelligent control systems that partially handle task execution and simplify the operation of the device [4]. Such shared control systems can use vision to estimate the size and shape of the object and thus select an object-specific grasp [5], [6]. Others have used machine learning methods to determine the object class and trigger appropriate grasp selection [7]. By reducing unnecessary user input to a prosthesis and lowering user effort, prosthetic acceptance can be improved [8], [9].

[1]R. V. Godoy, B. Guan, and M. Liarokapis are with the New Dexterity Research Group, The University of Auckland, Auckland, New Zealand {rdeg264, bgua324}@aucklanduni.ac.nz, minas.liarokapis@auckland.ac.nz

[2]A. Dwivedi is with Artificial Intelligence Institute, University of Waikato, NZ. E-mail: anany.dwivedi@waikato.ac.nz

[3]Mahonri Owen is with the School of Engineering, The University of Waikato, Hamilton, NZ mahonri.owen@waikato.ac.nz

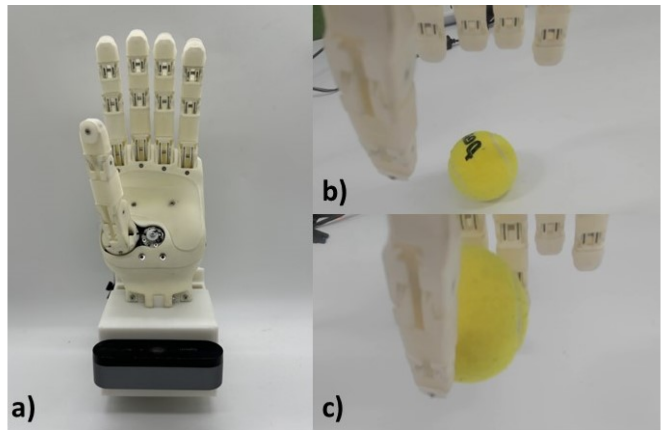

Fig. 1. The prosthetic device used and the camera views. Subfig. a) shows the developed humanlike prosthetic device, b) shows the view from the perspective of the hand while the user is approaching the object, and finally c) the view from the onboard camera when the object is securely grasped.

In our previous work, an affordances-based approach combined with myoelectric decoding for the control of a robotic hand was proposed [10]. This framework employed an object detection model pre-trained on the COCO dataset [11]. The COCO dataset contains a wide variety of 91 object classes. However, for vision-based systems implemented for prosthetic hand control, many of these classes are not necessary (e.g., airplane, horse, bus). Additionally, as the hand gets closer to the object, the image seen from the hand is often much different from the images of objects and that the COCO dataset provides, which can lead to misclassification.

Other researchers have created image/video datasets that are more suited to grasping. In [12], RBG-D data was collected from a camera attached to the user's arms as they performed tasks such as food preparation and housekeeping. This is not practical for prosthetic use as such hands often have limitations and cannot perform activities with the same dexterity as the human hand. In [13], the authors created a dataset with prosthetic applications in mind. Similarly, RGB-D data was collected from a camera attached to the user's forearm while limiting the variation in grasp types for better grasp transferring to prosthetic hands. In [14], authors developed a classification model for a vision-based shared control scheme, in which they created their own private dataset with thirteen objects from the YCB object set [15]. However, there is no publicly available dataset containing videos focusing on grasping of different everyday life objects with a prosthetic device (from the field of view of the prosthetic device).

Fig. 2. Set of objects grasped by the prosthetic device. The dataset consists of videos recorded using the onboard camera while the operator uses them.

In this paper, we present a dataset created to support grasping and manipulation applications using prosthetic devices. The dataset comprises videos of users grasping different objects using a prosthetic interface built to operate a humanlike prosthetic hand [16] that is shown in Fig. 1. The videos have been captured from the perspective of the hand using an onboard camera attached to the prosthetic device. The onboard camera provides information on the grasped object during all stages of the grasping process. The video recordings have been analyzed, and a set of grasp types has been extracted for each object based on a comprehensive grasp taxonomy, representing a list of all the possible grasps for each object examined. Furthermore, the analysis includes the identification of the most frequently employed grasps for each class of objects. This dataset can be used for training robust object detection models employed to be used from the perspective of the prosthetic hand. The dataset can help develop shared control frameworks for the operation of prosthetic hands, providing extra information for the user or control module, leading to an easier and more efficient deployment of such interfaces.

II. Methods

The data collection involves the acquisition of grasping data that is composed of video data from an onboard camera and the motor states from a prosthetic device. After collection of the data, the videos are annotated. These video annotations, along with the motor states, are used to establish a list of grasp types for each object class.

A. Prosthetic Device Design and Control

In this work, we modified one of our prosthetic hands, the OpenBionics robot hand [16]. The OpenBionics hand is an adaptive, humanlike robot hand. The previous hand version used three motors to control the fingers: one for finger flexion/extension, one for thumb flexion/extension, and one for abduction/adduction. To achieve individual control of all fingers, three extra Dynamixel XM430-W350-R motors were added to perform finger flexion/extension for the index, middle, and ring fingers. A hand base was specially designed to fit the extra motors. The base was made to be attached to the interfaces presented in [17]. In addition to accommodating the extra motors, this interface includes joysticks and an adapter to connect the base to a wrist brace worn by the user. The user can control the connected prosthetic hand by operating the control interfaces. Furthermore, a Logitech BRIO camera was attached to the interface so as to record video during the execution of the grasping tasks, resulting in a video dataset of objects and grasps from the perspective of the hand. The prosthetic device can be seen in Fig. 1a).

B. Data Acquisition

A user was asked to grasp different objects found in everyday life using the prosthetic device in all the logical and feasible ways they could, while having the Cutkosky's grasp taxonomy [18] as a reference (with the exception of grasps employing an adducted thumb, due to grasping limitations of the prosthetic hand). The video data from the onboard camera was recorded throughout the experiments at a rate of 27 fps, and the data was synchronized with the recordings of all motor states, sampled at 10 fps, and saved to a .csv file with timestamps. Synchronization and recording of all data is implemented in the Robot Operating System (ROS) [19].

C. Establishing Grasp Types

Based on the data acquired, a list of grasp types is obtained. For this paper, the goal is to determine a list of possible grasps that the user can employ to securely handle a determined object. To achieve this goal, a human annotator was employed to analyze the videos. For each time step, the annotator selects the object type that is being grasped and the grasp type that is employed, as defined by the previously chosen grasp taxonomy [18]. Through the camera videos and the identification of the motor states, a list of all grasp types employed for each object class was retrieved.

TABLE I

OBJECTS USED AND EMPLOYED GRASP TYPES. OBJECTS ARE CLASSIFIED AS **BLUE - SMALL**, **GREEN - MEDIUM SIZED**, AND **YELLOW - LARGE**. OBJECTS ARE SORTED FROM SMALLEST TO LARGEST.

Object	Grasp type											
	Large Diameter	Small Diameter	Medium Wrap	Thumb-4 Finger	Thumb-3 Finger	Thumb-2 Finger	Thumb-Index Finger	Disk (Power)	Sphere (Power)	Disk (Precision)	Sphere (Precision)	Tripod
Glasses2				✓	✓	✓	✓					
Pen						✓	✓					
Glasses1				✓	✓	✓	✓					
Watch												✓
Bucket		✓										
Toothbrush		✓										
Battery		✓				✓						
Marker2		✓		✓	✓	✓						
Jug		✓										
Fork		✓		✓	✓							
Marker1		✓		✓	✓	✓						
Spatula		✓										
Dustpan			✓									
Hairbrush			✓									
Marble												✓
Wallet			✓					✓		✓		
Screwdriver1		✓				✓						
Screwdriver2		✓				✓						
Headphone			✓	✓								
Cleaning brush			✓									
Ball3							✓					✓
Bottle3			✓									
Shaker		✓			✓							
Drill1			✓									
Umbrella			✓	✓								
Plier		✓	✓									
Bath toy							✓					✓
Drill2			✓									
Spray bottle2			✓									
Spray bottle1			✓									
Mouse1									✓		✓	✓
Mouse2									✓		✓	✓
Food box	✓				✓							
Bottle2			✓							✓		
Ball1									✓		✓	✓
Can			✓							✓		
Pencil case2			✓	✓	✓							
Clamp					✓	✓						
Jar			✓							✓		
Pencil case1			✓	✓	✓							
Ball2									✓		✓	
Cup1			✓							✓		
Bottle1			✓					✓		✓		
Sponge	✓			✓	✓	✓	✓					
Socks			✓									
Cup3			✓							✓		
Container	✓			✓	✓							
Cup2			✓							✓		
Tape1					✓	✓				✓		
Tape2					✓	✓				✓		

III. EXPERIMENTS AND RESULTS

The experiments consisted of the user grasping a total of 50 objects from 35 different object classes, shown in Fig. 2 and listed in Table I. The final video dataset consists of more than 48 minutes of video recordings, where more than 100 grasps were captured from the perspective of the hand using the onboard camera. The dataset zip file can be found at:

http://www.newdexterity.org/everydaygraspdataset

The dataset .zip file consists of two separate folders, one containing the videos for each object and one containing the .csv files with the motor states and the grasp type employed by the user for each timestamp. More information can be found in the readme.txt file in the dataset folder.

From this dataset, a human annotator classified each grasp used by the user for each timestamp of the video. As explained in Section II, the list of possible grasps for each object is derived. The list of grasp types is shown in Table I. In this table, the objects are also sorted based on their smallest dimension, which is usually where the user grasps the object. Dimensions from 0 to 25 mm were grouped as small, 25 mm to 60 mm were grouped as medium, and 60 mm or higher were grouped as large objects. Some grasp types, such as disk (precision) and large wrap become more prevalent as the object size increases.

For most objects, more than one grasp type is achievable using the prosthetic device; however, a prosthetic device will likely not be able to achieve as many grasps as the human hand. Moreover, different prosthetic hands will have different capabilities. Still, a variety of precision and power grasps were performed, resulting in 12 grasp types employed by the user using the prosthetic device. From Table I it can be seen that for some objects within the same class (e.g., bottles), a different set of grasps is employed. This suggests that grouping such objects in the same class might make sense in the context of using the video dataset for training object detection models, but in terms of grasping strategy needed they need to be treated differently.

Fig. 3 shows the occurrences of the grasp types employed during data collection. It can be noticed that the most employed grasp for this dataset was the medium wrap, as many objects presented radial symmetry (mostly due to the

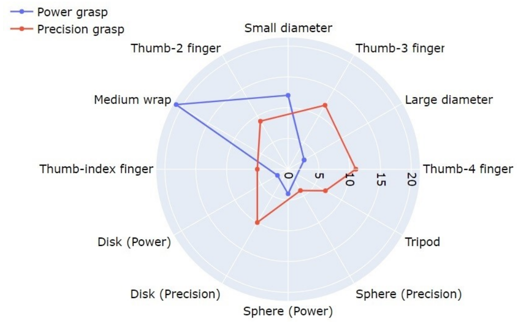

Fig. 3. Occurrence of each grasp type. The radar plot shows the number of times each grasp type is employed to grasp the objects during the experiments. Power grasps are mostly employed when the user requires a grasp that provides more security and stability, whereas precision grasps are employed when an emphasis on dexterity and sensibility is needed.

presence of handles that facilitate the grasping). However, when all the grasp types are grouped into power and precision grasps, Fig. 3 makes it clear that more precision grasps were used. This can be explained by the fact that several objects presented a small geometry, especially relative to the size of a prosthetic hand, which is bulkier compared to the human hand. Of course, the variety of precision and power grasps employed by the user demonstrates the dexterity and grasping capabilities of the examined prosthetic device, and the radar plot will be different for different prosthetic hands.

IV. Conclusion

This paper introduces a new open-access dataset designed to contribute to the development of prosthetic device shared control frameworks. The dataset features videos capturing user interactions with various objects through a wearable interface, employing a humanlike prosthetic hand. In total, 50 objects of 35 different classes were grasped resulting in a comprehensive video dataset capturing more than 100 grasps from the perspective of the hand. The subsequent analysis of the video recordings resulted in the extraction of a comprehensive set of grasps derived from well known grasp taxonomy that delineates all possible grasps. Additionally, objects were grouped into sizes to find correlation between object size and grasps employed. While the video dataset can improve computer vision algorithms applied in prosthetic control, the grasp type list can be used, similarly as in our previous work [10], in the development of shared control frameworks, thereby facilitating a smoother and more efficient deployment of prosthetic operation schemes. In future works, this dataset may be used to train object detection models, such as YOLO [20], which can then be deployed in a shared control scheme with the same or other prosthetic hands. Additionally, the dataset can be expanded to encompass more objects, involving more annotators.

References

[1] "Össur Touch Solutions. Ossur.com — ossur.com," https://www.ossur.com/en-us/prosthetics/touch-solutions, [Accessed 25-01-2024].

[2] "Michelangelo hand — The Michelangelo hand helps you regain extensive freedom — ottobock.com," https://www.ottobock.com/en-us/product/8E500, [Accessed 25-01-2024].

[3] A. Dwivedi, Y. Kwon, A. J. McDaid, and M. Liarokapis, "A learning scheme for emg based decoding of dexterous, in-hand manipulation motions," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 2205–2215, 2019.

[4] A. Dwivedi, D. Shieff, A. Turner, G. Gorjup, Y. Kwon, and M. Liarokapis, "A shared control framework for robotic telemanipulation combining electromyography based motion estimation and compliance control," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 9467–9473.

[5] S. Došen, C. Cipriani, M. Kostić, M. Controzzi, M. C. Carrozza, and D. B. Popović, "Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation," *Journal of neuroengineering and rehabilitation*, vol. 7, no. 1, pp. 1–14, 2010.

[6] M. Markovic, S. Dosen, C. Cipriani, D. Popovic, and D. Farina, "Stereovision and augmented reality for closed-loop control of grasping in hand prostheses," *Journal of neural engineering*, vol. 11, no. 4, p. 046001, 2014.

[7] T. C. Hansen, M. A. Trout, J. L. Segil, D. J. Warren, and J. A. George, "A bionic hand for semi-autonomous fragile object manipulation via proximity and pressure sensors," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 6465–6469.

[8] C. Cipriani, F. Zaccone, S. Micera, and M. C. Carrozza, "On the shared control of an emg-controlled prosthetic hand: Analysis of user–prosthesis interaction," *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 170–184, 2008.

[9] J. Gonzalez-Vargas, S. Dosen, S. Amsuess, W. Yu, and D. Farina, "Human-machine interface for the control of multi-function systems based on electrocutaneous menu: application to multi-grasp prosthetic hands," *PloS one*, vol. 10, no. 6, p. e0127528, 2015.

[10] R. V. Godoy, B. Guan, A. Dwivedi, and M. Liarokapis, "An affordances and electromyography based telemanipulation framework for control of robotic arm-hand systems," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 6998–7004.

[11] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.

[12] A. Saudabayev, Z. Rysbek, R. Khassenova, and H. A. Varol, "Human grasping database for activities of daily living with depth, color and kinematic data streams," *Scientific data*, vol. 5, no. 1, pp. 1–13, 2018.

[13] L. T. Taverne, M. Cognolato, T. Bützer, R. Gassert, and O. Hilliges, "Video-based prediction of hand-grasp preshaping with application to prosthesis control," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 4975–4982.

[14] F. Hundhausen, D. Megerle, and T. Asfour, "Resource-aware object classification and segmentation for semi-autonomous grasping with prosthetic hands," in *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2019, pp. 215–221.

[15] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The ycb object and model set: Towards common benchmarks for manipulation research," in *2015 international conference on advanced robotics (ICAR)*. IEEE, 2015, pp. 510–517.

[16] J. Chapman, A. Dwivedi, and M. Liarokapis, "A dexterous, adaptive, affordable, humanlike robot hand: Towards prostheses with dexterous manipulation capabilities," in *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, 2022, pp. 337–343.

[17] F. Sanches, G. Gao, N. Elangovan, R. V. Godoy, J. Chapman, K. Wang, P. Jarvis, and M. Liarokapis, "Scalable. intuitive human to robot skill transfer with wearable human machine interfaces: On complex, dexterous tasks," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 6318–6325.

[18] M. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 269–279, 1989.

[19] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source Robot Operating System," in *ICRA Workshop on Open Source Software*, 2009.

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.