2023 IEEE International Conference on Robotics and Biomimetics (ROBIO) | 979-8-3503-2570-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ROBIO58561.2023.10354698

Learning Hand Gestures using Synergies in a Humanoid Robot

Parthan Olikkal, *Student Member, IEEE*, Dingyi Pei, *Student Member, IEEE*, Bharat Kashyap Karri, Ashwin Satyanarayana, Nayan M. Kakoty, *Senior Member, IEEE,* and Ramana Vinjamuri, *Senior Member, IEEE*

***Abstract*— Hand gestures are a natural way of communication and integrating them into robots could allow for more efficient human-robot collaboration. In recent years, researchers and roboticists have attempted to replicate human hand motor control using the concept of synergies. In this paper, we present a new approach to obtaining kinematic synergies from hand gestures using a single RGB camera. We capture real-time hand gestures using the MediaPipe framework and convert them to joint angular velocities. We then use dimensionality reduction to obtain kinematic synergies from the joint angular velocities. These synergies can be used to reconstruct new hand gestures. We translate these reconstructed hand movement patterns into a humanoid robot, Mitra. Our results show that it is possible to control most of the joints of the robot for performing hand gestures using only a few synergies. This is more efficient than other contemporary methods. Furthermore, robots and prosthetics that use synergy models could enable near-natural human-robot collaboration.**

Keywords—MediaPipe, hand kinematics, kinematic synergies, biomimetic robots, human robot interaction, bioinspired robots

I. Introduction

Bipedal locomotion in humans along with opposable thumb have promoted the extensive usage of hands for grasping, reaching, and dexterous manipulation. The human hand comprising of joints, tendons, muscles, and nerves coordinate together to achieve gestures, and activities that we perform in our daily life with ease. A simple kinematic model of human hand has more than 20 degrees of freedom (DoF) making it an extremely challenging problem to be replicated in robots. The study of human hand movements thus, has been a significant area of research for three decades and both researchers and roboticists have been actively trying to address this challenge of replicating human hand dexterity.

With the advancement in technology, conventional devices for interaction with computers are replaced with more natural communication approaches such as oral communication and body language. Of these two methods, the most efficient means of natural communication is body language interaction. Of the different parts of the body, hands are the most effective non-verbal means for interaction. When we communicate with others, our hand movements tend to give more impact to messages. For example, for some gestures the coordinated movement of all four fingers and thumb are needed whereas for other gestures, individual finger movements are required. Thus, our range of hand gestures for communication extends from simple hand movements to complex hand movements.

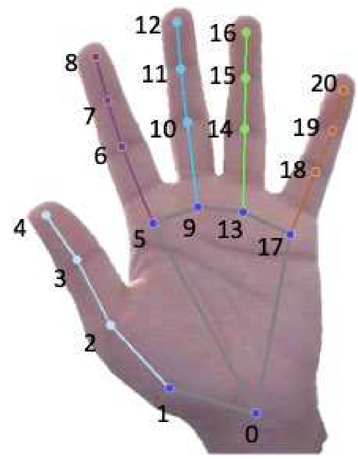

Fig. 1. Twenty-one hand-knuckle landmarks obtained from MediaPipe and their corresponding anatomical areas are illustrated here. Each dot represents the joints and line represents the Euclidean distance from each joint. Corresponding labels in the illustration are 0. Wrist, 1. CMC-Thumb, 2. MCP-Thumb, 3. IP-Thumb, 4. TIP-Thumb, 5. MCP-Index Finger, 6. PIP-Index Finger, 7. DIP-Index Finger, 8. TIP-Index Finger, 9. MCP-Middle Finger, 10. PIP-Middle Finger, 11. DIP-Middle Finger, 12. TIP-Middle Finger, 13. MCP-Ring Finger, 14. PIP-Ring Finger, 15. DIP-Ring Finger, 16. TIP-Ring Finger, 17. MCP-Pinky, 18. PIP-Pinky, 19. DIP-Pinky, 20. TIP-Pinky

By combining hand gestures as an interaction tool and the ability to classify hand gestures into meaningful symbols/values, more intuitive human-robot interaction (HRI) and human-computer interaction (HCI) interfaces can be developed that can potentially assist individuals with motor impairments. Hand gesture-based interaction systems have thus become a magnetic area of research since its introduction in 1970s. A myriad applications of human computer interactive systems have been developed using hand gesture control such as sign language recognition [1], improving motor skills [2] and user guide interactive applications [3].

Because of the complex anatomy of the human hand, and the underlying joints and muscles, there is a huge possibility to perform one movement task such as picking up a bottle of water, using several different coordinated combinations of muscles and joints and there are multiple ways to perform the same movement. But how does the human brain choose which combination of muscles and joints to be recruited to perform such tasks? Modularity hypothesis introduced by Bernstein [4] was able to address most of the challenges of the large DoFs and thereby the large number of redundant choices for performing a simple task. The neuroscientific reasoning for this strategy is the finding that, despite the complexity of the

*Research supported by National Science Foundation CAREER Award HCC-2053498 and IHFC-NSF grant GP/2021/RR/017.

P. Olikkal, D. Pei, B. K. Kashyap and R. Vinjamuri are with the Vinjamuri lab in the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, 21220, USA (e-mail: rvinjam1@umbc.edu). A. Satyanarayana is with City Tech at CUNY. N. M. Kakoty is with Tezpur University, Assam, India.

human hand, fewer variables can adequately account for most of the variation in patterns of human hand configurations and movements. Bernstein in his modularity hypothesis called these variables as synergies. Kinematic and muscle synergies obtained from joint kinematics and muscle movements have gained much popularity among the other investigated synergies. Here, in this study, we would be focusing on kinematic synergies obtained from joint kinematics while performing hand movements.

The attempt to execute a targeted motor action such as picking up a bottle of water can be divided into reaching, grasping/manipulation and release phases as noted in [5]. The angular joint velocity profiles of these three phases collected from joint recording devices such as CyberGlove are well represented and investigated in [5]–[9]. These studies reinforce that when the subject attempts to reach a target from an initial reference position, a rise in angular velocity profile can be observed in the reaching phase. During the grasping phase, a steady angular velocity profile is noted and finally during the release phase, the angular velocity falls to the initial reference state as elaborated in [5]. Upon close observation, as indicated in [5], it can be noticed that for each gesture, angular joint velocity profile was collected from sensors that correspond to each of the finger joints.

Emphasizing on only one joint, it can be observed that the angular velocity profile for any movement can be split similarly into a start, target, and return phases where the target phase represents the flexion or extension of the joint to achieve a desired state along its DoF and the start and return phases represents the initial state of that joint. When performing any hand gestures, the angular joint velocity profiles of these joints tend to form a gaussian curve as the target state is achieved. Thus, we take inspiration from the above studies that substantiates this observation of a gaussian curve made by each joint during any sort of hand movements.

Technological breakthroughs have encouraged numerous robotic devices to mimic human arm and finger movements by observing the kinematic patterns. These coordinated kinematic patterns are usually extracted and embedded in these devices to assist in performing activities of daily living tasks. Several promising rehabilitative exoskeletons using kinematic synergies are detailed in [10]. However, not many studies have been conducted to understand the efficiency of kinematic synergies in humanoids. Hauser et al., in their study [11] was able to use few non-linear kinematic synergies from lower body to reduce the balance control problem to linear in a humanoid robot during slow movements. [12] validates the human inspired kinematic synergy as a potential candidate for balance control among the group of control concepts. To the best of our knowledge, humanoid robots that perform upper limb movements using kinematic synergies have not yet been explored. Unlike the humanoid, Pepper [13], through this study we attempt to provide a preliminary analysis on using biologically inspired human kinematic synergies on a humanoid robot for hand movements.

Hand gesture recognition based on the extracted features and different recognition approaches are described in [3]. Conventional motion capturing sensors and devices are now replaced with intuitive frameworks which simplify gesture recognition applications. MediaPipe presented by Google is one of the open-source frameworks replacing conventional methods that offers several machine learning solutions. From the several solutions provided for vision tasks such as object detection, face detection, gesture recognition, hand landmark detection, for this paper, we chose hand landmark detector. MediaPipe hand landmark detector enables to identify the hand landmarks in an image. Thus, this model allows one to apply graphic effects over the hand image and pinpoint key hand regions. Using such a framework for gesture recognition helps not only identify hand landmarks in complex environments with different illumination, backgrounds but also enable adequate focus and attention in deriving joint movement kinematics and postures.

Although hand gesture recognition has been a compelling field of research, to our understanding only a few studies have used kinematic synergies for hand gesture classification and recognition. In this paper, we aim to use kinematic synergies obtained from end postures of hand gestures to predict end postures of new hand gestures. These kinematic synergies thus found, and the new hand gestures predicted will all be transferred to a humanoid robot to observe and compared the reconstruction patterns. Results from this paper may provide new insights to understanding hand kinematic synergies which may potentially help improve the design of humanoid robots and rehabilitative robotic devices, prosthetics, and exoskeletons.

II. Methods and Analysis

A. MediaPipe Framework

As described by Zhang [14], using a single RGB camera, a real-time hand gesture recognition system has been designed that can predict the skeleton of a human hand. MediaPipe hand landmark detector makes use of two modes – a palm detector model and a hand landmark model. The palm detector model works on detecting the palm by analyzing the entire image and produces the image with an oriented binding frame of the hand. The hand landmark takes in the cropped binding box image as input and through regression returns 3D hand key points on the image. The model returns 21 key points on the 3D hand-knuckle skeletal image on the hand. Each of the identified landmark is composed of distinct relative x, y, and z coordinates where x and y are normalized by image width and height whereas z represents the depth of the landmark. These were illustrated in Fig. 1 with dots as joints and lines represent

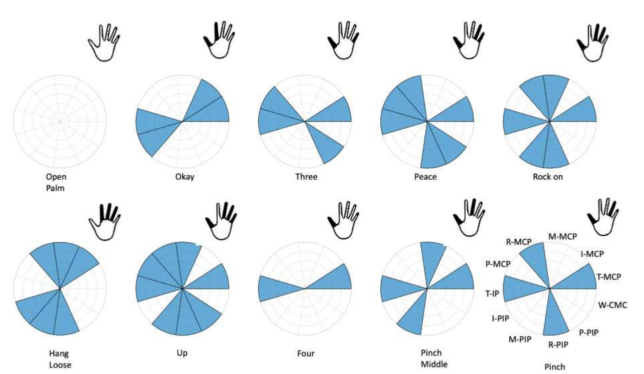

Fig. 2. Ten hand gestures are illustrated here. Each pie represents one hand gesture with 11 sensors and each sector represents one sensor. Those sensors that are activated are shaded in blue and their corresponding finger representation is shown in the top right posture.

Euclidean distance. Euclidean distance measure is calculated between each landmark that act as a condition on each of the dots for identifying any hand gesture.

B. The Humanoid Robot – Mitra

For this study, a humanoid robot, Mitra (Invento Research Inc, Plano TX) was included. Mitra is a custom build robot with 21 DoFs – Fingers (5), Wrist (1), Elbow (1), Shoulders (2) per arm with Head (1) and Base (2) as shown in Fig. 5. Fitted on top of the head is an RGB camera with a resolution of 1280 × 720 pixels for capturing real-time images and videos. The right-hand digits are provided with additional support to allow for grasping heavy objects. It also includes a LiDAR system to map the surroundings and provides several ways to connect such as voice commands, web interfaces, touch screen, joystick, and scripts. For hand gesture modeling, scripting method was used to communicate with Mitra.

C. Experiment

In this study, hand gestures were identified using MediaPipe hand landmark detection model from an RGB camera mounted on Mitra when shown. Once the landmarks were identified, based on the Euclidean distance from the wrist, the open state or closed state of thumb, index, middle, ring, and pinky fingers were found. Based on the open and closed state of the digits, hand gestures are identified. 10 active hand gestures – okay, open palm, three, peace, up, rock on, hang loose, four, pinch middle, pinch ring (where the thumb meets the middle/ring finger) were included in this study. These hand gestures were shown to Mitra from an initial reference posture of a relaxed idle hand posture.

D. Derivation of Synergies

Synthetic Angular Joint Velocities

From the hand gestures shown to Mitra, the end postures of each of the gesture was converted to angular joint velocities using a gaussian function as shown in the equation below from the reference posture.

$$\phi = \frac{1}{\sqrt{2\,\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Here, ϕ is the generated gaussian curve, σ is the standard deviation and μ is the mean of x formed from a sample of 10000 randomly generated data points. A total of 11 such velocity profiles were created corresponding to ten joints of the hand and one carpometacarpal joint of wrist – metacarpophalangeal (MCP) and interphalangeal (IP) joints of the thumb and MCP and proximal interphalangeal (PIP) joints of the other four digits along with an additional sensor to indicate the carpometacarpal (CMC) joint of the wrist as shown in Fig. 2.

Synthetic Kinematic Synergies

Joint angular velocities from the end postures of 10 hand gestures are synthetically generated using the gaussian equation. For those gestures that involve certain joint flexion, corresponding sensors out of 11 were applied with gaussian function. Once the angular joint velocities were generated for the selected 10 hand gestures, the dataset was split into training set consisting of 7 gesture tasks and testing set with 3 gesture tasks. Following [7], an angular velocity matrix was created

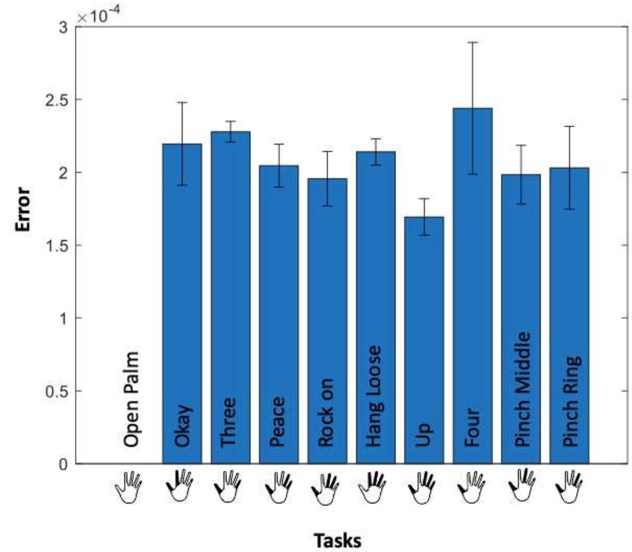

Fig. 3. Mean reconstruction error obtained while reconstructing the 10 hand gestures using synthetic kinematic synergies is illustrated here.

using the gestures under the training set such that each of the 11 sensors were cascaded one after the other. Thus, each row of the angular velocity matrix represents one gesture.

From the several models available, we used both time-invariant and time-variant synergy models. Similar to our previous works [8], [15], it was noted here also that time-variant synergy model provided the best results, hence for this paper we make use of time-variant synergy model. Principal component analysis (PCA) was applied on the above cascaded matrix to obtain PCs that accounted for maximum variance. To find the optimal number of PCs, based on our prior works [8], [15], we chose those PCs that accommodated to 0.9 (90%) of total variance using the equation expressed as

$$\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_m}{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \geq 0.9$$

Where λ represents the magnitude of the corresponding PCs and m represents the optimal number of synergies out of n available synergies. When this fraction reaches to 0.9, the corresponding m identifies as the optimal number to be chosen. These chosen optimal m PCs were termed as synthetic kinematic synergies.

E. Reconstruction of Hand Gestures

The angular joint velocities of 3 gesture tasks grouped under the testing set were reconstructed using the derived synthetic kinematic synergies. Reconstruction of the angular joint velocities were performed by using l_1-norm minimization detailed in [7]. Reconstruction error between the synthetic angular velocities (M_i) and the reconstructed patterns (X) using time-variant synergies were determined as follows.

$$err = \frac{\sum_i (M_i - X)^2}{\sum_i M_i^2}$$

F. Translating Hand Gestures to Mitra

The reconstructed pattern of gestures, the obtained three synthetic kinematic synergies, and the test hand gestures were further translated into Mitra. A moving average function and a

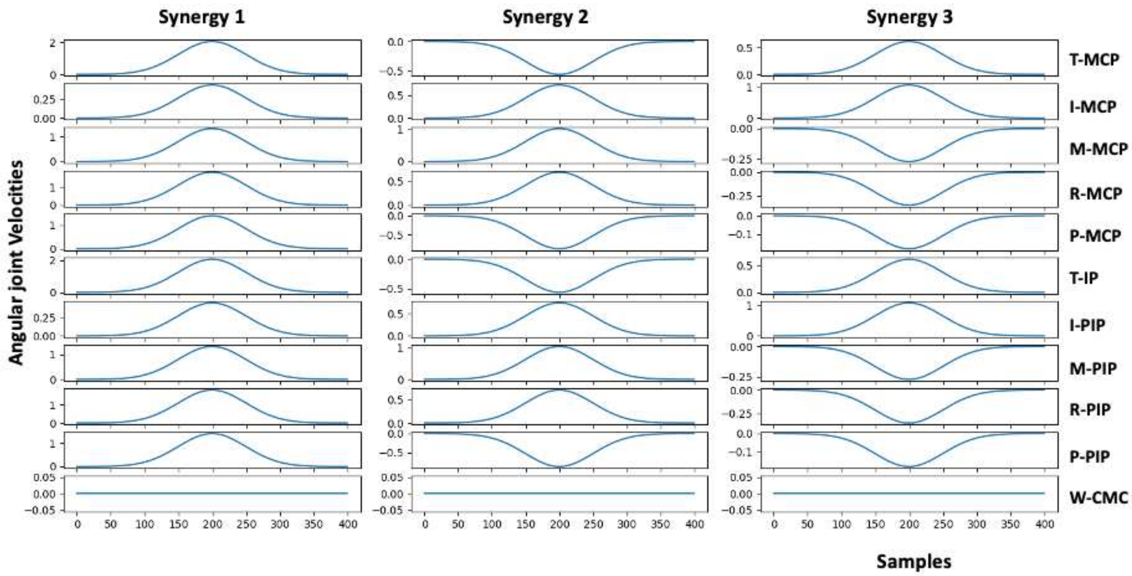

Fig. 4. Angular joint velocities of first three synthetic kinematic synergies of the 11 joints extracted from the training data is illustrated here.
Here, T-Thumb, I-Index, M-Middle, R-Ring, P-Pinky, MCP-Metacarpophalangeal, PIP- Proximal Interphalangeal, IP-Interphalangeal, CMC-Carpometacarpal joints

scaling coefficient was applied to map to the joints of Mitra. Continuous input from the reconstructed patterns and test data were given to Mitra during the execution.

III. Results

From the 10 synthetic joint angular velocities generated through the gaussian function the end postures of the 10 hand gestures were obtained; three synthetic kinematic synergies were extracted from those hand gestures grouped under the training set using PCA. 10-fold cross validation was performed, reducing the variance of the performance of the time-variant synergy model. On an average, for all the 10-fold cross validation in training set, the first synergy accounted for about 57% of the total variance, the second synergy accounted for 20% of the variance while first three synergies together accounted to 90%. It can be noted that, similar to [16] the first synergy was able to account for 50% of the variance and recruiting additional synergies increased the variance. This indicates that from the synthetic joint angular velocities for 7 hand gestures, a relatively small set of synergies could adequately represent the joint movements behind the hand gestures.

Reconstruction of the end posture of the test hand gestures were performed using synthetic kinematic synergies. By implementing 10-fold cross validation, each hand gesture appeared in the testing set thrice. As mentioned previously, reconstruction of the test hand gesture patterns was compared with the synthetically generated joint angular velocities for that hand gesture using least squared error. Fig. 3 represents the reconstruction error of 10 hand gestures reconstructed using synthetic kinematic synergies across all 10-fold cross validation. The reconstructed patterns of the 10 hand gestures were then mapped to Mitra as shown in Fig. 5. As the reconstruction error is significantly smaller and of the magnitude of 10^{-4}, no significant differences were found between the reconstructed hand gestures and generated angular joint velocities for that particular gesture. Synthetic kinematic synergies extracted from the training set, as shown in Fig. 4, were also mapped to Mitra. During the mapping of the synergies, Mitra hands were flexed to 50% which acted an initial reference posture. The movements were mapped such that any value above 50% were considered as flexion and any value below 50% would be considered as extension.

It can be observed from Fig. 4 that the wrist CMC does not have any movements compared to other joints for the three extracted synthetic kinematic synergies through PCA. This is potentially because the hand gestures selected for the experiment does not involve any wrist flexion or extension. Additionally, from Fig. 4 it can be noted that the angular joint velocity profile of the first synergy involves only flexion of varying amplitude, whereas the other two synergy profiles contains both flexion and extension but of smaller amplitudes. This reinforces the results mentioned in [16] and our previous studies that the synergy with the maximum variance may potentially account for the majority of the movement profile followed by the next synergy with the second maximum variance. Combining these synergies as a weighted linear combination, the end postures of the hand gestures were reconstructed with mean reconstruction error as shown in Fig. 3.

IV. Discussion

Numerous investigations [17]–[20] have been done to show that synergies are not a mathematical representation but rather an efficient tool for comprehending how the central nervous system (CNS) organizes motor control and coordination. As a result of such studies, promising results [21]–[25] have led to the use of synergies in several applications including robotics.

This paper presents an approach to generate synthetic joint angular velocities from end postures of 10 hand gestures from a single RGB camera and then derives synthetic kinematic synergies which were later used to reconstruct new hand gestures. To the best of our knowledge, this is one of the first

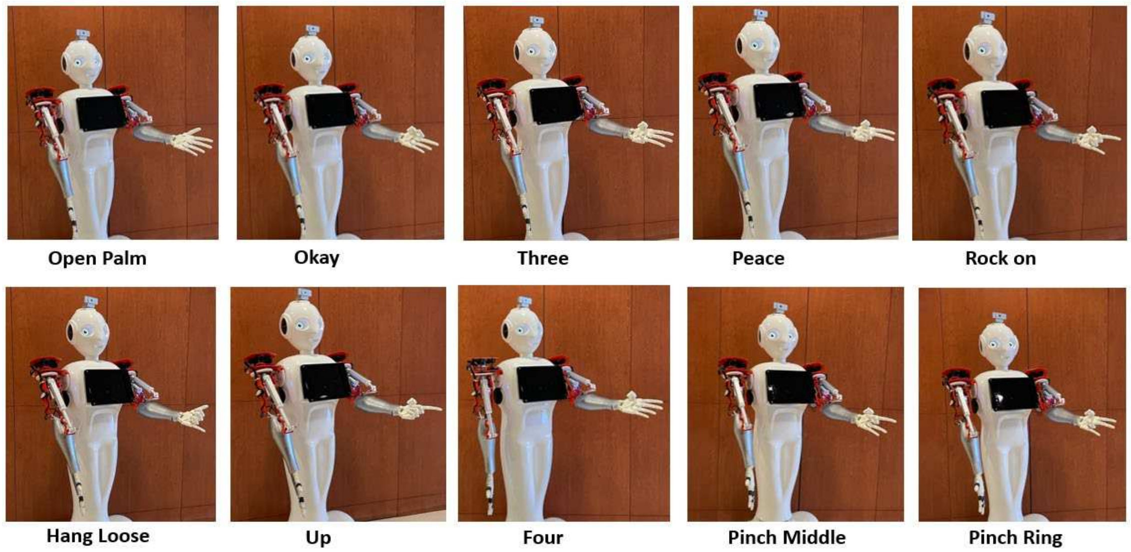

Fig. 5. Reconstructed hand gestures represented in Mitra using first three synthetic kinematic synergies for all the 10-hand gestures.

attempts to extract synthetic kinematic synergies from human inspired gaussian function generated angular joint velocities. When a simple finger flexion is performed, we can notice that the angular joint velocity profile of that flexion tends to represent a gaussian function, a bell-shaped velocity profile that well accounts for the three phases of movement profile. By employing such gaussian functions to represent flexion of fingers, end postures of 10 hand gestures were generated. By using only a few synergies, hand gestures under the testing task were reconstructed. This reconstruction and the recorded patterns for the hand gestures under test data were then mapped into Mitra.

Of the different mapping approaches such as joint-joint mapping, cartesian space mapping [26], [27], joint-cartesian mapping [28] and object-based mapping [29] as elaborated in [30], in this study we adopted joint-joint mapping approach which has shown promising results [26], [31] for a direct relationship between the corresponding joints of the human hand and Mitra's hand. Such joint-joint mapping approach allows for high mapping capabilities for the hand gesture data used for this study.

Through a simple mapping of the 21 landmarks found from MediaPipe Hand landmark detector [14] or learning from human demonstrations such as behavioral cloning [32], [33] to a robot can essentially perform the 10 hand gestures. This study attempts to show that using synergies, a subset of the 21 landmarks can be used to execute the selected hand gestures. In learning human demonstration, one of the key challenges is to convert the human hand motion into robot hand motion. Assuming a humanoid with 21 joints, the 21 hand landmarks need to be translated the robot to get a single hand gesture. Similarly, through behavioral cloning, motion retargeting of the 21 joints from human demonstration to robot should be performed to achieve the same hand gesture. Moreover, multiple demonstrations of the same gesture need to be recorded and provided to the robot to learn that gesture. To collect these demonstrations of the same gesture requires long hours of intense human effort from setting the angle of the camera to verifying and eliminating outlier data manually. But, through a synergy-based model, by using only a subset of the 21 landmarks, similar hand gestures can be executed by the robot with only one demonstration of the same hand gesture.

In this study, Mitra has only one DoF for each of the digit thus a total of 5 DoF for the 5 digits. MCPs of these 5 digits can be controlled and based on the flexion, the MCPs move in a gradient fashion accounting for the flexion of PIPs and DIPs. Each of the 10 hand gestures were demonstrated only once at the RGB camera of Mitra and the end postures are generated from the MediaPipe framework. Kinematic synergies extracted from the training data (in Fig. 5) of the generated joint angular velocities were also translated to the humanoid using a mapping function. But as kinematic synergies are bipolar in nature, meaning it has both positive and negative activation potential accounting for flexion and extension, we changed the initial reference state to accommodate this property of kinematic synergies. Thus, the MCPs of Mitra were set to 50% flexed as the initial reference state. Upon mapping the selected three synergies to Mitra, those joints that were below this reference were able to indicate the negative activation potential whereas those joints that were above the reference state were able to indicate the positive activation potential. These mapped values were fed continuously to the MCPs of the humanoid. Similarly, reconstructed hand gestures using the synthetic kinematic synergies were translated in a continuous fashion to the MCPs. As the MCPs moved, it brought together PIPs and DIPs to the target position from the initial reference posture. Each of the achieved targeted position of the 10 hand gestures are shown in Fig. 5.

Thus, by integrating Mitra with the synergy-based model, not only can the robot learn the hand gestures from the test data, but also a library of new synergies can be formed based on the new hand gestures shown to Mitra apart from the selected hand gestures.

V. Conclusion

In this paper, we propose a new method to extract synthetic kinematic synergies from end postures of hand gestures from a single RGB camera using MediaPipe, gaussian function and

PCA. To the best of our knowledge, this is one of the first study to translate kinematic synergies into a humanoid robot to perform hand gesture movements. Currently, an offline model is implemented that takes 10 preconfigured hand gestures that includes most of the finger joints. In the near future, we will extend these results to an online model which imitates the hand gestures in real-time and include flexion and extension of the wrist joint as well. Moreover, we widen the scope to include capturing continuous hand gesture movements and real-time mapping into the humanoid. Enabling such synergy-based humanoid and robots can reduce the complexity involved in motion retargeting and potentially enable improvements in industrial robots and assistive robots.

REFERENCES

[1] R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," *Expert Syst. Appl.*, vol. 150, Jul. 2020, doi: 10.1016/J.ESWA.2020.113336.

[2] S. Cai, G. Zhu, Y. T. Wu, E. Liu, and X. Hu, "A case study of gesture-based games in enhancing the fine motor skills and recognition of children with autism," *Interact. Learn. Environ.*, vol. 26, no. 8, pp. 1039–1052, Nov. 2018, doi: 10.1080/10494820.2018.1437048.

[3] Indriani, M. Harris, and A. S. Agoes, "Applying Hand Gesture Recognition for User Guide Application Using MediaPipe," *Proc. 2nd Int. Semin. Sci. Appl. Technol. (ISSAT 2021)*, vol. 207, pp. 101–108, Nov. 2021, doi: 10.2991/AER.K.211106.017.

[4] N. Bernstein, "The Co-ordination and Regulation of Movement." 1967.

[5] N. J. Jarque-Bou, M. Vergara, J. L. Sancho-Bru, V. Gracia-Ibáñez, and A. Roda-Sales, "A calibrated database of kinematics and EMG of the forearm and hand during activities of daily living," *Sci. Data*, vol. 6, no. 1, pp. 1–11, Dec. 2019, doi: 10.1038/s41597-019-0285-1.

[6] N. J. Jarque-Bou, A. Scano, M. Atzori, and H. Müller, "Kinematic synergies of hand grasps: A comprehensive study on a large publicly available dataset," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, p. 63, May 2019, doi: 10.1186/s12984-019-0536-6.

[7] R. Vinjamuri *et al.*, "Dimensionality Reduction in Control and Coordination of the Human Hand," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 2, pp. 284–295, 2010, doi: 10.1109/TBME.2009.2032532.

[8] P. Olikkal, D. Pei, T. Adali, N. Banerjee, and R. Vinjamuri, "Data Fusion-Based Musculoskeletal Synergies in the Grasping Hand," *Sensors (Basel).*, vol. 22, no. 19, Oct. 2022, doi: 10.3390/S22197417.

[9] P. Olikkal, D. Pei, T. Adali, N. Banerjee, and R. Vinjamuri, "Musculoskeletal Synergies in the Grasping Hand; Musculoskeletal Synergies in the Grasping Hand," *2022 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2022, doi: 10.1109/EMBC48229.2022.9871023.

[10] N. Jarrassé *et al.*, "Robotic Exoskeletons: A Perspective for the Rehabilitation of Arm Coordination in Stroke Patients," *Front. Hum. Neurosci.*, vol. 8, no. DEC, Dec. 2014, doi: 10.3389/FNHUM.2014.00947.

[11] H. Hauser, G. Neumann, · Auke, J. Ijspeert, W. Maass, and A. J. Ijspeert, "Biologically inspired kinematic synergies enable linear balance control of a humanoid robot," *Biol Cybern*, vol. 104, pp. 235–249, 2011, doi: 10.1007/s00422-011-0430-1.

[12] A. V. Alexandrov, V. Lippi, T. Mergner, A. A. Frolov, G. Hettich, and D. Husek, "Human-inspired Eigenmovement concept provides coupling-free sensorimotor control in humanoid robot," *Front. Neurorobot.*, vol. 11, no. APR, p. 257890, Apr. 2017, doi: 10.3389/FNBOT.2017.00022/BIBTEX.

[13] A. K. Pandey and R. Gelin, "A Mass-Produced Sociable Humanoid Robot: Pepper: the First Machine of Its Kind," *IEEE Robot. Autom. Mag.*, vol. 25, no. 3, pp. 40–48, 2018, doi: 10.1109/MRA.2018.2833157.

[14] F. Zhang *et al.*, "MediaPipe Hands: On-device Real-time Hand Tracking," 2020, [Online]. Available: http://arxiv.org/abs/2006.10214

[15] P. Olikkal, D. Pei, T. Adali, N. Banerjee, and R. Vinjamuri, "Musculoskeletal Synergies in the Grasping Hand," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2022-July, pp. 3649–3652, 2022, doi: 10.1109/EMBC48229.2022.9871023.

[16] V. Patel, M. Burns, Z. H. Mao, N. E. Crone, and R. Vinjamuri, "Linear and Nonlinear Kinematic Synergies in the Grasping Hand," *J Bioeng. Biomed. Sci*, vol. 5, no. 2, p. 163, 2015, doi: 10.4172/2155-9538.1000163.

[17] M. C. Tresch, V. C. K. Cheung, and A. D'Avella, "Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets," *J. Neurophysiol.*, vol. 95, no. 4, pp. 2199–2212, Apr. 2006, doi: 10.1152/JN.00222.2005/ASSET/IMAGES/LARGE/Z9K0040673510008.JPEG.

[18] K. M. Steele, M. C. Tresch, and E. J. Perreault, "The number and choice of muscles impact the results of muscle synergy analyses," *Front. Comput. Neurosci.*, vol. 7, no. JUL, Aug. 2013, doi: 10.3389/FNCOM.2013.00105.

[19] A. Santuz, A. Ekizos, L. Janshen, V. Baltzopoulos, and A. Arampatzis, "On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion," *Int. J. Neural Syst.*, vol. 27, no. 5, Aug. 2017, doi: 10.1142/S0129065717500071.

[20] J. Taborri, E. Palermo, D. Masiello, and S. Rossi, "Factorization of EMG via muscle synergies in walking task: Evaluation of intra-subject and inter-subject variability," *I2MTC 2017 - 2017 IEEE Int. Instrum. Meas. Technol. Conf. Proc.*, Jul. 2017, doi: 10.1109/I2MTC.2017.7969775.

[21] P. K. Artemiadis and K. J. Kyriakopoulos, "EMG-based teleoperation of a robot arm in planar catching movements using ARMAX model and trajectory monitoring techniques," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 2006, pp. 3244–3249, 2006, doi: 10.1109/ROBOT.2006.1642196.

[22] P. K. Artemiadis, P. T. Katsiaris, and K. J. Kyriakopoulos, "A biomimetic approach to inverse kinematics for a redundant robot arm," *Auton. Robots*, vol. 29, no. 3–4, pp. 293–308, Nov. 2010, doi: 10.1007/S10514-010-9196-X/METRICS.

[23] T. Cunha, P. M. Vieira, K. Costa, and C. P. Santos, "Looking for motor synergies in Darwin-OP biped robot," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 2016-June, pp. 1776–1781, Jun. 2016, doi: 10.1109/ICRA.2016.7487322.

[24] E. Hocaoglu and V. Patoglu, "Tele-impedance control of a variable stiffness prosthetic hand," *2012 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2012 - Conf. Dig.*, pp. 1576–1582, 2012, doi: 10.1109/ROBIO.2012.6491192.

[25] F. Lunardini, C. Casellato, A. D'Avella, T. D. Sanger, and A. Pedrocchi, "Robustness and Reliability of Synergy-Based Myocontrol of a Multiple Degree of Freedom Robotic Arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 9, pp. 940–950, Sep. 2016, doi: 10.1109/TNSRE.2015.2483375.

[26] J. Rosell and R. Suárez, "Using hand synergies as an optimality criterion for planning human-like motions for mechanical hands," *IEEE-RAS Int. Conf. Humanoid Robot.*, vol. 2015-February, pp. 232–237, Feb. 2015, doi: 10.1109/HUMANOIDS.2014.7041365.

[27] G. Gioioso, G. Salvietti, M. Malvezzi, and D. Prattichizzo, "Mapping synergies from human to robotic hands with dissimilar kinematics: An approach in the object domain," *IEEE Trans. Robot.*, vol. 29, no. 4, pp. 825–837, May 2013, doi: 10.1109/TRO.2013.2252251.

[28] R. Meattini, D. Chiaravalli, L. Biagiotti, G. Palli, and C. Melchiorri, "Combined Joint-Cartesian Mapping for Simultaneous Shape and Precision Teleoperation of Anthropomorphic Robotic Hands," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 10052–10057, Jan. 2020, doi: 10.1016/J.IFACOL.2020.12.2726.

[29] G. Gioioso, G. Salvietti, M. Malvezzi, and D. Prattichizzo, "An Object-Based Approach to Map Human Hand Synergies onto Robotic Hands with Dissimilar Kinematics," *Robotics*, pp. 97–104, 2019, doi: 10.7551/mitpress/9816.003.0018.

[30] G. Salvietti, "Replicating human hand synergies onto robotic hands: A review on software and hardware strategies," *Front. Neurorobot.*, vol. 12, no. JUN, p. 27, Jun. 2018, doi: 10.3389/FNBOT.2018.00027/BIBTEX.

[31] M. T. Ciocarlie and P. K. Allen, "Hand Posture Subspaces for Dexterous Robotic Grasping," *http://dx.doi.org/10.1177/0278364909105606*, vol. 28, no. 7, pp. 851–867, Jun. 2009, doi: 10.1177/0278364909105606.

[32] F. Torabi, G. Warnell, and P. Stone, "Behavioral Cloning from Observation," *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 2018-July, pp. 4950–4957, May 2018, doi: 10.48550/arxiv.1805.01954.

[33] E. De Coninck, T. Verbelen, P. Van Molle, P. Simoens, and B. Dhoedt, "Learning robots to grasp by demonstration," *Rob. Auton. Syst.*, vol. 127, p. 103474, May 2020, doi: 10.1016/J.ROBOT.2020.103474.