Advances in Machinery, Materials Science and Engineering Application IX
M. Chen et al. (Eds.)

doi:10.3233/ATDE230534

PALMitate: Hand Gesture Robotic Hand Emulator

Jobenilita R. CUÑADO [a,1], Ariel Van C. GERTOS [a], Karl Gian JIMENEZ [a] and Jomarie C. GRANADA [a]
[a] *Department of Engineering Education, University of Mindanao – Tagum, Visayan Village, Tagum City, Davao del Norte 8100, Philippines*

Abstract. Safety-critical situations inevitably require human intervention, thereby, exposing personnel to formidable risks. This encompasses various domains, such as bomb disposal, emergency response, and hazardous material handling, all underlining the paramount importance of personnel safety. This research is dedicated to the development of a wireless robotic hand that replicates predefined gestures, facilitated by image processing algorithms. The hand gesture classifier attains an accuracy rate of 94.67% in identifying and distinguishing gestures. The implementation of hand gesture robotic hand emulators in safety-critical scenarios emerges as a pivotal strategy to potentially replace human participation. This substitution substantially mitigates risks for personnel. Moreover, it stands to enhance safety standards by virtue of its automated and controlled operation, thus minimizing the vulnerability associated with manual intervention. Fundamentally, this study acts as empirical validation of the substantial transformative capabilities inherent in robotics. This has the potential to completely reshape how dangerous tasks are handled in various industries which has a direct impact on keeping people safe and protecting valuable lives. Even though this research investigates a constrained selection of hand gestures and operates within a time limit for emulation, it establishes the fundamental basis for future progress in human-robot interaction and the realm of gesture recognition.

Keywords. Robotic hand emulator, video and image processing, robotics, mimicking robot

1. Introduction

According to the report by the Integrated Network for Societal Conflict Research (INSCR) on High Casualty Terrorist Bombings (HCTB), spanning from September 11, 1989, to March 10, 2019, there were a staggering 1,367 instances of bomb attacks worldwide, targeting both civilian and non-civilian populations. Tragically, these attacks led to an alarming death toll of 54,156 individuals. The statistics highlight the devastating impact of such acts of violence on societies across the globe [1]. In safety-critical situations, dedicated personnel are deployed to respond promptly, often at great personal risk. To address these dangers, researchers have developed robotic applications that mimic human movements, effectively substituting personnel and minimizing the risk of fatalities.

---

[1] Jobenilita R. Cuñado, Corresponding author, Department of Engineering Education, University of Mindanao – Tagum, Visayan Village, Tagum City, Davao del Norte 8100, Philippines; E-mail: jobenilita_cunado@umindanao.edu.ph.

Robot mimicking has emerged as a prominent research topic in recent years. In the early 2000s, researchers primarily relied on basic controls such as joysticks and computer terminals to operate mimicking robots [2][3]. With rapid technological advancements, researchers now utilize advanced sensor technologies to capture human motion accurately, attaching sensors to various parts of the human body. Motion sensors, including accelerometers, play a crucial role in enabling precise human motion capture for mimicking robots, which can be easily interfaced with microcontrollers. They excel at measuring tilt in the x and y planes, converting the data into an analog signal [4][5]. For capturing hand motion, researchers use data gloves equipped with flex sensors which accurately recognize specific finger movements [6][7]. In recent times, researchers have embraced video and image processing techniques to further enhance the capabilities of mimicking robots. In video processing, a video clip is meticulously analyzed by breaking it down into individual frames, which are then compared to a specific dataset for accurate motion recognition [8][9]. On the other hand, image processing involves using various techniques, such as object identification and object tracking, to analyze static images, enabling more comprehensive data extraction [10][11]. For hand gesture classification, the Leap Motion Controller has emerged as a valuable tool. Specifically designed for tracking hand and finger movements in three-dimensional space, it finds prominent use in human-computer interaction applications [12][13]. Bluetooth, and RF (Radio Frequency) are commonly used by most researchers for communication, though these approaches have their limitations. RF signals can be susceptible to interference during operation, potentially affecting performance and reliability. Meanwhile, Bluetooth signals only allow short-range communication, typically limited to two devices [14][15]. Despite these challenges, continuous advancements in communication technologies are expected to address such issues and enable more seamless and robust interactions between humans and mimicking robots.

The use of motion sensors attached to the human hand to capture human motion faced challenges, as external factors like sweat, temperature, and aging could introduce inaccuracies into the data. To address these limitations, this research explores alternative approaches for capturing human motion without the need to attach any sensors to the human body by adopting non-intrusive methods.

The primary focus of this study is to create a robotic hand capable of accurately replicating essential hand gestures. Specifically, the objectives include (1) designing a versatile robotic hand capable of mimicking five (5) distinct hand gestures, encompassing gripping, counting, and hand signal applications; and (2) employing video and image processing algorithms to effectively classify and recognize these hand gestures.

The integration of robotic mimicry, remote access, and image processing represents a groundbreaking advancement in technology with profound implications for human tasks. This combination empowers humans to carry out hazardous tasks from a safe distance, shielding them from potential harm. By employing image processing techniques, the system ensures consistent and reliable results without the need for attaching sensors, thereby minimizing potential inaccuracies caused by external factors.

This study focuses on the classification and emulation of five (5) specific hand gestures, namely gripping, counting, and hand signal applications. Following the classification process, the robotic hand replicates the identified hand gesture from a pre-defined initial state, demonstrating its capability within a maximum time delay of fifteen (15) seconds. Although this research explores a limited set of hand gestures and imposes a time constraint for emulation, it lays the groundwork for future developments in human-robot interaction and gesture recognition.

2. Methodology

2.1. System Architecture Overview

Figure 1 presents an architectural overview of the Hand Gesture Robotic Hand Emulator, providing for a comprehensive system that captures, processes, and emulates hand gestures, showcasing the potential of human-robot interaction and gesture recognition in practical applications. It is comprised of six interconnected components that facilitate seamless communication: camera, monitor, transceiver, microcomputer, microcontroller, and robotic arm.

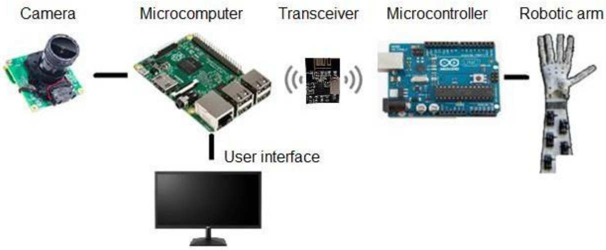


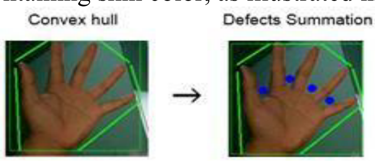
Figure 1. Hand gesture robotic hand emulator system.

2.2. Hand Gesture Classifier

The hand gesture classification process involves two distinct layers of segmentation: Region of Interest (ROI) Detection and skin color extraction:

(1) Region of Interest (ROI) Detection. In figure 2, the region of interest is delineated. To ensure comprehensive coverage despite potential gaps, a convex hull is drawn over the entire region. For determining the actual finger count, a circle is strategically placed between any gaps with an angle less than 90° and greater than 30°. Each time a gap is detected, the finger count is incremented by one. In cases where no defects are detected, a contour ratio is utilized to discern whether the hand gesture corresponds to one or zero fingers.

(2) Skin Color Extraction: In this method, skin color extraction is achieved by defining a range of skin colors in terms of HSV (hue, saturation, value) values, spanning from the lower to the upper range. Subsequently, all pixels falling within this defined range are set to white, while pixels outside this range are set to black. The resulting binary image highlights the regions of the image containing skin color, as illustrated in figure 3.



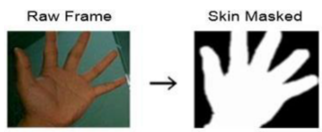


Figure 2. ROI detection.

Figure 3. Skin color extraction.

2.3. Robotic Hand

The robotic arm is driven by five servo motors, with each servo motor dedicated to controlling a specific finger. These servo motors operate by attaching strings from them to specific points on each finger. The control sequence for the servo motors is coded in the Microcontroller, which varies for each specific hand gesture, following instructions from the Microcomputer. The overall structure of the robotic hand is represented in figure 4.

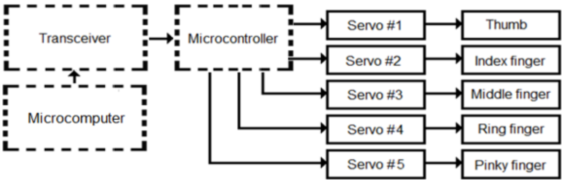


Figure 4. Robotic hand structure block diagram.

3. Results and Discussion

The development of a hand gesture robotic hand emulator involved the creation of a robotic hand, which is controlled by a microcontroller. The microcontroller receives precise instructions from the microcomputer after processing the images obtained from the camera. A Hand Haar Cascade algorithm is implemented in the classifier to accurately identify and interpret hand gestures in the robotic emulation. The system is equipped to recognize and emulate five pre-defined hand gestures, as depicted in figure 5. Additionally, it also has a well-defined pre-defined initial state. This initial state serves as the starting point for each gesture emulation.

Figure 5. Pre-defined hand gestures and initial state.

During the fabrication of the robotic hand, a cost-effective approach was employed by utilizing an open-source 3D model printed to create the intricate components of the hand. Each finger was assembled using three 3D printed pieces: the bottom, middle, and top parts. To enable the finger mechanism, two (2) strings are attached from the bottom to the top parts of each finger. These strings are then connected to the servo's shaft at the other end. The servo's capability for both forward and reverse rotation allows it to pull the strings on either side, effectively emulating finger flexion or extension, as demonstrated in figure 6.

In the hand gesture classifier, a method is employed to evaluate the area of contour and the convex hull, as illustrated in figure 7. These measurements are then utilized to calculate the area ratio of each pre-defined hand gesture, as outlined in equation 1.

$$area\ ratio = \left(\frac{area\ of\ convex\ hull - area\ of\ contour}{area\ of\ contour}\right) \quad (1)$$

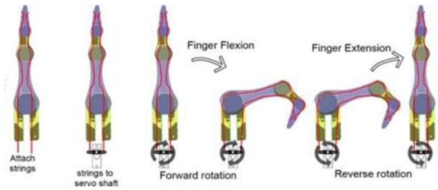


Figure 6. Finger mechanism.

Figure 7. Areas of contour and convex hull.

Area ratio is also used for differentiation because the area ratio of every hand gesture is unique. The area ratio results are shown in table 1.

Table 1. Area ratio results per gesture.

Hand Gestures	Area Ratio
Gripping	11.686
One	23.765
Two	29.896
Three	32.99
Four	31.153
Initial State	51.853

While table 1 yielded some closely matched results, it is essential to note that hand gestures can also be effectively differentiated by identifying defects. Defects refer to regions that are not covered by the hand within the convex hull. To facilitate this process, each defect can be associated with a triangle. By utilizing the cosine rule, the angle between gaps can be accurately obtained, as depicted in equation 2.

$$\cos A = \frac{b^2+c^2-a^2}{2bc} \tag{2}$$

To address the issue of defects caused by noise, a filtering approach is implemented. The system calculates the distance from the convex hull to each angle associated with a defect. If this distance is found to be less than the average distance, which is predefined as 30 units, the defect is excluded from consideration. The area of the triangle formed by the defect is determined using Heron's formula, as outlined in equations 3 and 4.

$$s = \frac{(a+b+c)}{2} \tag{3}$$

$$Area = \sqrt{s(s-a)(s-b)(s-c)} \tag{4}$$

After getting the area, the distance from the angle to the opposite side can be solved using equation 5.

$$distance = \frac{2 * Area}{opposite\ side} \tag{5}$$

A defect is identified when the angle formed by a triangle is equal to or less than 90 degrees and if the distance of the adjacent side is higher than 7.94 mm. The results of the angle and adjacent distance measurements are summarized in table 2.

Table 2. The angle and adjacent side per defect.

Defects	Angle	Adjacent Side
#1	45.76691864	12.91118188
#2	37.59883833	16.02063598
#3	28.82830032	17.95359804
#4	47.89847172	11.69189285

With the aid of the identified defects, the hand gesture classifier can recognize the number of fingers that are spread out using equation 6, where N represents the number of fingers spread out.

$$N = number\ of\ defects + 1 \tag{6}$$

In situations where the hand gestures depicted do not have any defects, the hand gesture classifier employs the area ratio results obtained from table 1 to effectively differentiate between them.

Following the hand gesture classification process, communication between components is achieved using an HC-06 Bluetooth module. With the HC-06 Bluetooth module, the Microcontroller can receive instructions from the Microprocessor.

During the prototype testing of the hand gesture classifier, it was noted that maintaining a specific distance helps emphasize the finger complexion and effectively isolates the hand from external noise before feeding the image to the Microprocessor. The results of the hand gesture classifier's accuracy relative to the distance from the camera to the user are depicted in figure 8. During the actual testing phase, the experiments are conducted in a controlled environment—a closed room with a specific light source. By maintaining a consistent lighting condition, the system achieves more reliable and consistent results during hand gesture recognition. table 3 presents the hand gesture classification accuracy of 94.67%.

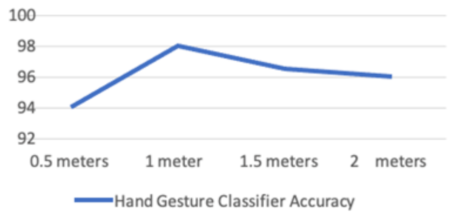


Figure 8. Hand gesture classifier accuracy relative to distance of the camera from the user.

Table 3. Hand gesture classifier's confusion matrix.

		Predicted					
Hand Gestures		1	2	3	4	5	6
#1	1	15					
#2	2		15				
#3	3			14	1		
#4	4			1	13	11	
#5	5					14	
Initial State	6						15

4. Conclusion and Future Works

The research objectives have been successfully accomplished, resulting in the design of a robotic hand capable of mimicking five pre-defined hand gestures and the implementation of video and image processing algorithms for accurate hand gesture classification. The use of high-stall torque servo motors has proven crucial in maintaining precise finger positions for each hand gesture, contributing to the robotic hand's realistic emulation capabilities. Additionally, the elimination of uncontrolled luminance in the hand gesture classifier has significantly improved its accuracy, particularly in detecting defects, and has prevented distortions in the raw image.

The decision to utilize Bluetooth connectivity has proven advantageous, as it allows for a reliable and secure point-to-point connection. Its low power consumption and the ability to disable visibility further enhance the system's performance, minimizing interference and maximizing efficiency.

For future endeavors, incorporating background subtraction techniques, especially when performing robotic emulation in high-luminance environments, will further improve the hand gesture classifier's robustness and accuracy.

By building upon the achievements of this research and considering the suggested improvements, the system can be refined and extended to explore new possibilities in human-robot interaction, virtual reality, and assistive technologies.

References

[1] The INSCR website. Retrieved from: http://www.systemicpeace.org/. 2019.

[2] Harja J, Tikkanen J, Sorvoja H, Myllylä R. Magnetic resonance imaging-compatible, three - degrees - of - freedom joystick for surgical robot. Int. J. Med. Robotics Comput. Assist. Surg. 2007; 3: 365-371. doi:10.1002/rcs.159. 2007.

[3] Saharia T, Bauri J, Bhagabati C. Joystick controlled wheelchair. International Research Journal of Engineering and Technology (IRJET). 2017; 04(07).

[4] Bularka S, Szabo R, Otesteanu M, Babaita M. Robotic arm control with hand movement gestures. Proceedings of the 2018 41st International Conference on Telecommunications and Signal Processing (TSP); Athens, 2018, pp. 1-5, doi: 10.1109/TSP.2018.8441341.

[5] Megalingam RK, Gedela V, Bandyopadhyay S, Rahi M. Robotic arm design, development, and control for agriculture applications. 2017; pp. 1-7. 10.1109/ICACCS.2017.8014623.

[6] Mamun A, Alamgir FM. Flex sensor-based hand glove for deaf and mute people. International Journal of Computer Networks and Communications Security. 2017; 5: 11.

[7] Guo Y, Zhang X, An N. Monitoring neck posture with flex sensors. 2019: 459-463. 10.1109/ICIST.2019.8836806.

[8] Siswanti S, Asmir R. Image processing of hand gesture for augmented reality systems. 10.2991/aisr.k.200424.096.

[9] Long P. Video processing for hand gesture recognition. Thai Nguyen University Journal of Science and Technology. 2016; 14: 81-86.

[10] Sree K, Eyakumar G. An evolutionary computing approach to solve object identification problem in image processing applications. Journal of Computational and Theoretical Nanoscience. 2020; 17: 439-444. 10.1166/jctn.2020.8687.

[11] Tian J, Wang G, Gao W, Fang D. Research and implementation of object tracking. IOP Conference Series: Earth and Environmental Science. 2020; 440: 042098. 10.1088/1755- 1315/440/4/042098.

[12] Khan F, Ong H, Bahar N. A sign language to text converter using leap motion. International Journal on Advanced Science, Engineering, and Information Technology. 2016; 6: 1089. 10.18517/ijaseit.6.6.1252.

[13] Demircioglu B, Bulbul G, Kose H. Turkish sign language recognition with leap motion. 2016: 589-592. 10.1109/SIU.2016.7495809.

[14] Thomas A. RF based wireless bomb defusing manipulator robotic arm. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2016; 5: 3511-3516. 10.15662/IJAREEIE.2016.0505008.

[15] Thomas A. RF based wireless bomb defusing manipulator robotic arm. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 5. 3511-3516. 10.15662/IJAREEIE.2016.0505008. 2016.