

# PROCEEDING



The 4<sup>th</sup> International Conference on Information and **Communication Technology** 



May 25 - 27 2016 The Papandayan Hotel, Bandung, Indonesia



### TABLE OF CONTENT













## Design of A Gripping Imitator Robotic Arm for Taking An Object

### Ali Rizal Chaidir, Alfredo Bayu Satriya, Guido Dias Kalandro Electrical Engineering Department Universitas Jember

ali.rizal@unej.ac.id, alfredobayu.teknik@unej.ac.id, guidokalandro89@gmail.com

*Abstract*—This paper presents the design of robotic arm which imitate the human hand movement to grip an object. The robotic arm is remote controlled so that it can be equipped to a robot which can explore inaccessible or hazardous area and do a dangerous task. The operator's hand movement will be captured by camera and will be processed in three level approaches of video processing namely background subtraction, hand and finger detection, and gripping recognition. The video processing result will be converted by inverse kinematics to become joint space of the robotic arm. The joint space data is sent to robotic arm as the signal control using bluetooth module. The test result shows that the system can work properly and in its optimum parameters, it can achieve 85% of success rate.

*Keywords*—*gripping recognition; inverse kinematic; robotic arm; tele-operated; video processing;*

#### I. INTRODUCTION

Robot has become a very beneficial technology in human life. Robot has some characteristics which are very useful in assisting human or even replacing them to do a task as regards both physical activity or decision making [1]. Unlike human and other living things, robot does any task without being tired. Robot is a digital technology so it also has an exact accuracy and precision. Nowadays, robot has been developed for a dangerous task which risks a human life or any task in an inaccessible area for human, such as defusing bomb and underwater or space exploration. The use of robot for space exploration has been discussed since 1980s. National Aeronautics and Space Administration (NASA) even has designed a flying robot in 1983. Robot has been chosen for the outer space mission due to harsh environment in outer space with extreme temperatures, vacuum, radiation, gravity, and great distances, which is hazardous and very difficult for human access [2]. Robot is also proposed for underwater exploration because human beings are not capable of submerging to the ocean depths or staying for longer periods of time underwater [3]. Using robot for defusing or removing bomb is more secure than sending human to do it since it can minimize the number of human victims in case of explosion occurs.

The idea of using robot in defusing bomb, underwater and space exploration is based on the fact that robot is more expendable than human life. Despite having advantages, robot will face some shortcomings for completing these tasks. In space exploration, robot will face uncertainty environment or unexpected situation [4]. Many uncertainty things also might be happened to robot in defusing bomb and underwater exploration. In another side, a robot is designed only for a certain task and certain circumstance, so it will get a difficulty in making decision when it faces many variations of situation. There are two main solution to solve the problem. The first one, the robot can be programmed with a very advanced artificial intelligent to work automatically. Having an artificial intelligent, the robot can learn and adapt in a new environment. Still, this solution is costly in computation and does not guarantee the success of operation. The second solution is equipping the robot with remote control system or teleoperated system [3], [5]. This system allows the interaction between robot and human as its operator for making decision. By using this technique, the robot is fully controlled by human and the operator still does not risk his life.

Our work means to build a robot which is able to be remote controlled by a human operator to explore a hazardous area or to do a dangerous task. In this paper we present an initial part of the project namely gripping imitator robotic arm. The robotic arm is designed to imitate the movement of operator's arm for gripping an object. In exploration, one important task of the robot is collecting a certain object, such as geologic sample [5]. Therefore, a gripping robotic arm is needed in exploration mission. So does the bomb defusing task, the robot needs to have robotic arm to pick, remove or defuse the bomb. The robotic arm imitates the operator's hand gesture by the use of Human Computer Interface (HCI). HCI is an area of research and practice that emerged in the early 1980s, initially as a speciality area in computer science embracing cognitive science and human factors engineering [6]. Actually, there are some devices that can be used as HCI, namely keyboard, mouse, joystick. These devices are not good enough especially to control robotic arm since it has many degrees of freedom (DOF) and variations of movement. On the other hand, HCI device which uses the aspect of human behaviour or state such as speech, sound, facial expression, body position and hand gestures can be powerful for many applications [7].

Similar prior works have shown a good result of hand gesture imitation to robotic arm [8], [9], [10]. In [8], the system uses corner feature extraction and Lucas-Kanade Algorithm to track the hand gesture. The result shows a real time hand gesture imitation, but the system only tracks any movement in the captured frames without include any hand detection. It can make an ambiguity in the system when there is another nonhand object moving in the frames. In [9], the system has an accurate active finger estimation based on hand contour. But despite showing a specific robotic arm movement, the result data just show the robotic arm motion in the state of serial signal. Moreover, none of these works has applied the robot to pick an object.



Fig. 1. Hand Gripping Imitator Robotic Arm Complete System

The complete system of this work is shown in Fig.1. The gripping gesture of the operator's hand is captured using web camera. The video data is then being processed in the computer using Visual Basic programming so it can detect the hand gesture. The detected gesture data is transformed using inverse kinematics as the robotic arm signal control. The signal control is sent to robotic arm via USB bluetooth transmitter and bluetooth receiver module.

The rest of the paper is organized as follows. In Section II, we present the system model of the gripping imitator robotic arm. In Section III, we show the result and discussion of our work. Finally, the conclusions are drawn in Section IV.

#### II. METHODS AND MODELS

The procedure of gripping imitator data processing requires video processing for gripping recognition, and inverse kinematic.

#### *A. Video Processing Algorithm*

There are three levels of approach in the video processing algorithm, namely background subtraction, hand and finger detection and gripping recognition. The complete video processing algorithm is shown in Fig. 2.

*1) Background Subtraction:* The captured operator's hand and gripping video is sampled into many frames of image. The speed of video sampling is varied from 0.2 fps to 200 fps so its effect can be known. The frame image is first processed so that hand and finger can be detected. The process of hand and finger detection (foreground) from the background in a frame image is called background subtraction. There are some methods for background subtraction, namely Codebook Method, Gaussian Method, Skin Detection Method and Color Co-Occurrence Method [10]. In this paper, we use skin detection method due to its low complexity and simplicity. The captured video from camera is in RGB color model and it is transformed into YCbCr model. The YCbCr model is better in recognition than the RGB model [11] and has less intersection between skin colour and non-skin colour in any variation of illumination [12], [13]. The pixel color model transformation from RBG into YCbCr is shown in equation (1).



Fig. 2. Flowchart of Video Processing in Gripping Imitator Robotic Arm

$$
\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2568 & 0.5041 & 0.0979 \\ -0.1482 & -0.2910 & 0.4392 \\ 0.4392 & -0.3678 & -0.0714 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}
$$
 (1)

After the transformation into YCbCr model, the image is enhanced by median filter. Median filter is used to reduce the noise in the image by replacing the value of a pixel with the median value in the neighbourhood of that pixel. This filter is very effective in the presence of impulse noise or salt-andpepper noise [14]. The median filter equation is defined as:

$$
Y[a,b] = median (I_{orig}[i,j], i,j \in nbor[a,b])
$$
 (2)

where  $Y[a, b]$  is the result of filter, a and b is the column and row pixel index of the resulted image,  $I_{\text{oria}}[i, j]$  is the original YCbCr image,  $i$  and  $j$  is the column and row pixel index of the original image and *nbor*  $[a, b]$  is the sub image of the original YCbCr image.

The last step of background subtraction is transforming the YCbCr image into binary image (image consists of black and white pixel). In this process, the foreground will be represented by the white pixel (255) while the background will be represented by black pixel (0). The black and white image is gotten by applied the threshold value of each  $Y$ ,  $Cb$ and Cr as defined in equation (3).



Fig. 3. Region of Interest Parameter in The Binary Image of The Operator's Hand

$$
G(a,b) = \begin{cases} 1 & if & (T_Y max \ge Y(a,b) \ge T_Y min) \cap \\ & (T_{Cb} max \ge Y(a,b) \ge T_{Cb} min) \cap \\ & (T_{Cr} max \ge Y(a,b) \ge T_{Cr} min) \\ 0 & if & (T_Y max < Y(a,b) < T_Y min) \cap \\ & (T_{Cb} max < Y(a,b) < T_{Cr} min) \cap \\ & (T_{Cr} max < Y(a,b) < T_{Cr} min) \end{cases} \tag{3}
$$

where  $G(a, b)$  is a binary image,  $Y(a, b)$  is the original image,  $T_Y$  max and  $T_Y$  min are maximal and minimal threshold of Y component,  $T_{Cb}max$  and  $T_{Cb}min$  are maximal and minimal threshold of Cb component, and  $T_{Cr}$  and  $T_{Cr}$  min are maximal and minimal threshold of Cr component.

*2) Hand and Finger Detection:* Because the most important object is the hand and finger, so it will be defined as the Region of Interest (ROI) of the image. In getting ROI, hand and finger will be cropped from the arm by finding the parameter  $x$ ,  $y$ , width and height as shown in Fig. 3.  $y$  is obtained by counting the difference between the first pixel line and the top of the white pixel.  $x$  is obtained by counting the difference between the first column and the leftmost white pixel. Width is obtained by counting the difference between  $x$ and the rightmost white pixel. Height is obtained by counting the difference between  $y$  and the lowest white pixel. The ROI is shown in Fig. 4. The ROI is then is resized to compensated the hand size in case the distance of hand and camera changes. The middle finger end position is obtained by scanning the image from top till rightmost white pixel is found. The similar way is applied to find the thumb end position.

*3) Gripping Recognition:* There are two movements in the video. The first movement is the movement of the arm. The movement of arm is defined by the movement of the middle finger and it will be converted to robotic arm using inverse kinematics. The second movement is the gripping movement. The process of gripping recognizing is obtained by Template Matching. For achieving a fast computation in template matching process, the Canny Edge Detection is applied to the binary ROI image [15]. The process of Canny Edge Detection is shown in Fig. 5.

The image is filtered using smoothing filter namely Gaussian Convolution. Then, the edge will be detected by the



Fig. 4. Region of Interest of The Hand



Fig. 5. Canny Edge Detection Algorithm

Gradient Calculation using Sobel operator. The Sobel operator  $(G_x$  and  $G_y$ ) is presented in equation (4-6). Detected edge might be too thick, so "Non Maximum Suppression" will be applied to improve the localization. Some 'streaking' might be appear in the image after Non Maximum Suppression. It can be overcome by applying another double threshold technique. The last step is applying hysterisis to remove non-edge component.

$$
G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}; G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}
$$
 (4)

$$
|G| = \sqrt{G_x^2 + G_y^2} \tag{5}
$$

$$
\theta = \arctan(G_x/G_y) \tag{6}
$$

After the edge is perfectly detected, the template matching is applied to recognize the gripping gesture of the operator's hand. Template Matching is an algorithm for object recognition or medical imaging which works by comparing two images [16]. Image after previous processes in a frame will be compared with image in previous thirtieth frame. If the correlation between the two images is bigger than the threshold, it can be concluded that there is no movement. If the correlation is smaller than the threshold, it can be concluded that there is a gripping gesture. The correlation value is obtained using equation (7).

$$
cor = \frac{\sum_{i=0}^{N-1} (I_T - \overline{I_T}) \times (I_S - \overline{I_S})}{\sqrt{\sum_{i=0}^{N-1} (I_T - \overline{I_T})^2 \times \sum_{i=0}^{N-1} (I_S - \overline{I_S})^2}} \tag{7}
$$

where  $I_T$  is the template grey level image,  $\overline{I_T}$  is the average template grey level image,  $I_S$  is the source image section,  $I_S$ is the average grey level in the source image, and  $N$  is the number of pixel in the section image.



Fig. 6. Robotic Arm Presentation in Cartesian Space



Fig. 7. Robotic Arm Presentation in Joint Space using Trigonometry Solution [17]

#### *B. Inverse Kinematics*

The position of middle finger end is presented by the position of the end manipulator of the hand (P). The notation P is in the form of Cartesian Space and presented by  $P_x$  and  $P_y$ . In another side, the robotic arm works in the joint space. In the joint space, the movement of the arm is presented in the form of the joint angle  $(\theta_1$  and  $\theta_2$ ). The kinematics of the robot and its presentation is shown in Fig. 6. Therefore, The data is needed to be converted from the Cartesian Space into Joint Space using inverse kinematics. Consider that the system is in the 2-DOF planar, the inverse kinematics can be solved by the geometry approach using trigonometry as shown in Fig. 7 [17].

Therefore, the relation between  $P = (P_x \text{ and } P_y)$  and joint angle ( $\theta_1$  and  $\theta_2$ ) can be written as:

$$
P_x = I_1 \cos \theta_1 + I_2 \cos(\theta_1 + \theta_2)
$$
 (8)

$$
P_y = I_1 \sin \theta_1 + I_2 \sin(\theta_1 + \theta_2)
$$
\n<sup>(9)</sup>

After some algebraic calculations, the joint angle is defined as:

$$
\theta_1 = \arctan(\frac{P_x}{P_y}) - \arctan(\frac{I_2 \sin \theta_2}{I_1 + I_2 \cos \theta_2})\tag{10}
$$



Fig. 8. Test Procedure of The Robotic Arm Control

$$
\theta_2 = \arccos\left(\frac{P_x^2 + P_y^2 - I_1^2 - I_1^2}{2I_1 I_2}\right) \tag{11}
$$

#### *C. Test Procedure*

The robotic arm and its control algorithm will be tested to take an object and carry it over about 10 cm as shown in Fig 8. The robot will be controlled by the operator's hand with variation of the video sampling or frame rate. Four frame rates are used namely 0.2 frame/second, 2 frame/second, 20 fps, and 200 fps. Moreover, the system will be tested by varying the distance between operator's hand and camera. There are three distances which will be tested, they are 0.5 meters, 0.6 meters and 0.7 meters. Every variable will be tested twenty times and the result will be recorded and presented. While testing the frame rate variable, the distance is fixed in 0.6 m. Vice versa, the frame rate is fixed in 20 fps while the distance is being tested.

#### III. RESULTS

The effect of frame rate to the system performance is shown in Table I. When the frame rate is increased the success rate is become higher. Frame rate 0.2 and 2 has the same success rate namely 15% which is a very low performance. Video processing for detecting the hand movement involves more than one frame so low frame rate makes high time delay in the process. This high time delay in processing makes the system does not operate in real time. Moreover, high time delay produces error which makes the robot does not do the task in high success rate. Both frame rate 20 fps and 200 fps have a high performance namely 85% of success rate. This result shows that 20 fps of frame rate is enough to achieve a high performance. Choosing frame rate above 20 fps will produce more complex computation and risk an error in data processing. It can be conclude that a frame rate about 20 to 200 fps is suitable for the system.

The effect of changing operator's hand to camera is shown in Table II. The result shows that when the distance is increased, the success rate tends to be better. The system becomes better while the distance is increased due to the camera coverage. The further the distance, the larger the coverage of camera. Large coverage of camera facilitate the

TABLE I. TEST RESULT IN VARIATION OF FRAME RATE WITH 0.6 M DISTANCE BETWEEN HAND AND CAMERA

Frame Rate (fps)	Success Rate $(\% )$
0.2	15
	15
20	85
200	Q٢

TABLE II. TEST RESULT IN VARIATION OF OPERATOR'S HAND TO CAMERA DISTANCE WITH 20 FPS OF FRAME RATE



movement of the hand and it makes the robotic arm could move well in order to do the task to take and carry it over. After the addition of the distance from 0.7 m to 0.8 m the success rate does not improve as previous addition. It remains about 85 % of success both in 0.7 m and 0.8 m distance. It means that in the distance of 0.7 m and above the coverage does not give effect to the system. It also means that the distance of 0.7 m between operator's hand and camera has given good performance to the system. Although the result shows that there is a change in performance when the distance changes, but actually the algorithm can compensate the change of distance. The change of performance is only affected by the camera coverage. After the distance is enough to cover the hand movement (0.7 m in our sampling), the system performance tends to be steady (85% in 0.7 m and 0.8 m). It can be concluded that the video processing system can compensate the change of distance but it needs a minimum distance of 0.7 m between operator's hand and camera to obtain a good performance.

The result shows that a perfect performance can not be achieved both by increasing the distance between hand and camera and the frame rate. The maximum success rate in data retrieval is 85%. This is because the failure of the robot in the task is due to non technical reason despite the algorithm. The object usually failed to be carried because both the robotic arm surface and object surface are slippery. The object often falls in the task process.

#### IV. CONCLUSION

Our system and algorithm are proved to be able to work properly. The system has a good performance with the parameter distance of operator's hand and camera 0.7 m or above and the parameter frame rate between 20 fps to 200 fps with the success rate of 85%. The system is able to compensate the changing of operator's hand and camera distance. There is also failure due to non technical reason such as slippery object surface. For future work, we will also develop a more powerful algorithm and test it in a hazardous environment.

#### **REFERENCES**

- [1] B. Siciliano, L. Sciavicco, G. Oriolo, *Robotics: Modelling, Planning and Control*, Springer, 2009.
- [2] K. Yoshida, "Achievements in Space Robotics", *IEEE Robotics and Automation Magazine*, Vol. 16, Issue 4, pp. 20-28, December 2009.
- [3] R. Saltaren, R. Aracil, C. Alvarez, E. Yime and J. M. Sabater, "Field and service applications - Exploring deep sea by teleoperated robot - An Underwater Parallel Robot with High Navigation Capabilities", *IEEE Robotics and Automation Magazine*, Vol. 14, Issue 3, pp. 65-75, March 2008.
- [4] D. M. Lyons, R. C. Arkin, S. Jiang, T-M. Liu, and P. Nirmal, "Performance Verification for Behavior-Based Robot Missions", *IEEE Transactions On Robotics*, Vol. 31, No. 31, June 2015.
- [5] J. A. Shah, J. H. Saleh, and J. A. Hoffman, "Review and Synthesis of Considerations in Architecting Heterogeneous Teams of Humans and Robots for Optimal Space Exploration", *IEEE Transactions On System, Man, and Cybernetics-Part C: Applications and Review*, Vol. 37, No. 5, September 2007.
- [6] J. M. Carroll "The Encyclopedia of Human Computer Interaction, 2nd edition: Human Computer Interaction - brief info," Interaction Design Foundation.
- [7] Q. Chen, N. D. Georganas, and E. M. Petriu, "Hand Gesture Recognition Using Haar-Like Features and a Stochastic Context-Free Grammar," *IEEE Transactions On Instrumentation and Measurements*, Vol. 57, No. 8, August 2008.
- [8] H.-T. Cheng, Z. Sun, P. Zhang, "Imirok: Real-Time Imitative Robotic Arm Control for Home Robot Applications," *IEEE International Conference Pervasive Computing and Communications Workshops (PERCOM Workshops)*, March 2011.
- [9] G. Choudhary B., C. Ram B. V., "Real Time Robotic Arm Control Using Hand Gestures," *International Conference On High Performance Computing and Applications (ICHPCA)*, 2014.
- [10] S. Karamchandani, S. Sinari, A. Aurora, D. Ruparel"The Gesture Replicating Robotic Arm," *International Symposium on Computational and Business Intelligence*, 2013.
- [11] W. Wang "Hand Segmentation Using Skin Color and Background Information ," *International Conference on Machine Learning and BCybernetics*, July 2012.
- [12] A. Y. Dawod, J. Abdullah, M. J. Alam "Adaptive Skin Color Model for Hand Segmentation ," *International Conference on Computer Applications and Industrial Electronics*, December 2010.
- [13] A. Y. Dawod, J. Abdullah, M. J. Alam "A New Method for Hand Segmentation Using Free-Form Skin Color Model ," *International Conference on Advanced Computer Theory and Engineering*, Augustus 2010.
- [14] J. W. Woods, Multidimensional Signal, Image, and Video Processing and Coding 2nd ed. USA : Academic Press, 2012.
- [15] C. Gentos, C.-L. Sotiropoulou, S. Nikolaidis, N. Vassiliadis "Real- Time Canny Edge Detection Parallel Implementation for FPGAs ," *International Conference on Electronics, Circuits, and Systems*, December 2010.
- [16] I. Pham, R. Jalovecky, M. Polasek"Using Template Matching for Object Recognition in Infrared Video Sequences ," *Digital Avionics Systems Conference*, September 2015.
- [17] S. Cubero et al, Industrial Robotics: Theory, Modelling and Control Germany : pro literatur Verlag, 2007.