

Development of Hand Gesture Based Electronic Key Using Microsoft Kinect

Arief Ramadhani^{1*}, Achmad Rizal¹, and Erwin Susanto¹

¹School of Electrical Engineering, Telkom University, 40257 Bandung, Indonesia

Abstract. Computer vision is one of the fields of research that can be applied in a various subject. One application of computer vision is the hand gesture recognition system. The hand gesture is one of the ways to interact with computers or machines. In this study, hand gesture recognition was used as a password for electronic key systems. The hand gesture recognition in this study utilized the depth sensor in Microsoft Kinect Xbox 360. Depth sensor captured the hand image and segmented using a threshold. By scanning each pixel, we detected the thumb and the number of other fingers that open. The hand gesture recognition result was used as a password to unlock the electronic key. This system could recognize nine types of hand gesture represent number 1, 2, 3, 4, 5, 6, 7, 8, and 9. The average accuracy of the hand gesture recognition system was 97.78% for one single hand sign and 86.5% as password of three hand signs.

1 Introduction

Security access using conventional keys is at risk of being lost or duplicated. Applying technology to the door lock system can make the door open only through the authentication process. This authentication process can be in the form of passwords, electronic key, and even biometrics. One of modality that can be used as authentication is hand gesture.

The hand gesture is commonly used in everyday life to communicate or help someone who has a lack of verbal language to communicate [1]. Hand gesture can be understood by looking at the signals that are delivered and then sent to the brain and processed by the brain to mean the meaning of the signal. With the same logic, a machine can also define a signal when it has 'eyes' and 'brains' as in humans. In this context, the use of cameras and processors or controllers can replace the role of the eyes and brain in humans.

Several studies have been conducted related to electronic key based on this hand gesture. An automatic door lock was developed using hand gesture using a webcam [1]. The image segmentation process was carried out with skin detection. The hand image and background were separated based on the color similarity with the skin color. This method could distinguish five different sign namely signals A, B, C, D, and E. The authors in [2] performed hand gesture recognition based on an accelerometer embedded in a user's smart watch. Data from the accelerometer was then processed using the Multi-Layer Perceptron artificial neural network so that the sign could be recognized.

* Corresponding author: ariefmadhani48@gmail.com

2 Related work

Several studies have been conducted regarding hand gesture recognition through different methods. In [3], an RGB camera was used to capture the hand image. The image segmentation process was carried out by determining the limits of red, green, and blue colors. In [4], the authors used Kinect sensor depth in the gesture recognition system. Their system could recognize ten different gestures. By using the support vector machines (SVM), the accuracy could reach more than 95%. The author in [5] also uses Kinect sensor depth in the system. Nine different gestures can be classified using K-Mean Clustering with the accuracy was from 84% to 90%. By using a different approach, the author in [6] transformed geometrical properties of the hand including area, centroid, Euclidian distance into a feature vector. Then, neural network is used for recognition and classification. The average accuracy of this gesture recognition was 85%.

In this study, we developed a hand gesture based electronic key based on the depth sensor of Microsoft Kinect. By utilizing the Kinect sensor depth, lighting conditions did not affect the performance of this system [7]. The difference with previous research in [3-6], we used Raspberry Pi as the processor which makes this system was portable and more efficient in space considering this hand gesture recognition will be applied to the electronic key system. In this study we propose a simpler method of hand gesture recognition by scanning each pixel to detect the presence of the thumb and the number of open fingers. Using this method our system has an accuracy of 97.78 percent for single hand gesture.

3 Proposed method

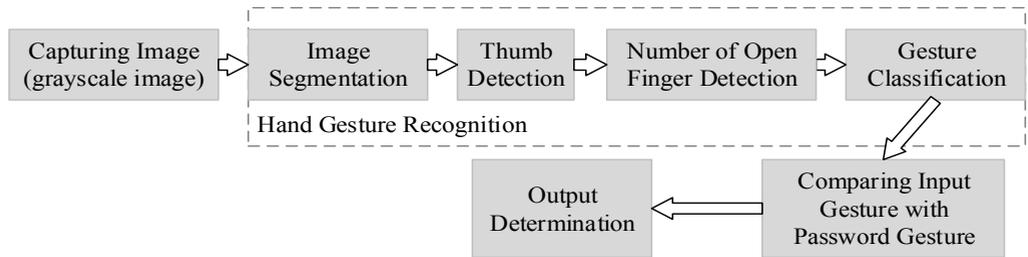


Fig. 1. Proposed method

Fig. 1 displays the proposed method. The hand gesture recognition method proposed in the paper consists of four main components: image segmentation, identification of thumb presence, detection of the number of open fingers and classification of hand signals. The rest subsection explains each process in detail.

3.1 Image Segmentation

Image segmentation is the process of dividing images into meaningful parts and not overlap based on similar criteria [8]. The purpose of image segmentation is to separate the desired object from the background. The segmentation of the grayscale image was done by determining the intensity limit (brightness) of pixels. In the depth image of the Kinect, the closest object to the sensor had the smallest intensity. In this case, the closest object to the

sensor was the hand. So we simply took the smallest intensity in the grayscale image in the segmentation process.



Fig. 2. Image segmentation result

3.2 Thumb detection

Fig. 3 displays the thumb detection process. The thumb detection process was done by scanning the pixels of each column from the leftmost column. The scanning process stopped when it finds the first 255 value and records the pixel row coordinates. The line coordinate data was then compared to the row coordinate limit. If the row coordinate of the leftmost white pixel was smaller than the row coordinate limit (Fig. 3. (a)), it means that there was no thumb on the image. If the row coordinate of the leftmost white pixel was greater than the row coordinate limit (Fig. 3. (b)), it could be assumed that there was a thumb. In some instances, the row coordinates of the leftmost white pixel were greater than the row coordinate limit, but in reality, there is no thumb on the image (Fig. 3. (c)).

To complete this process, we needed to calculate the area of the white area (pixel value of 255) from the column where we found the first 255 value for 20 columns to the right. Figure 4. (a) shows the area (the red line). This area was used to distinguish that the area is part of the thumb or not. If the area obtained did not exceed the limit, then it can be said that there was a thumb, and vice versa. Based on the experiment on ten subjects, we determined the line coordinate limit at 55 and area limit of 400 pixels.

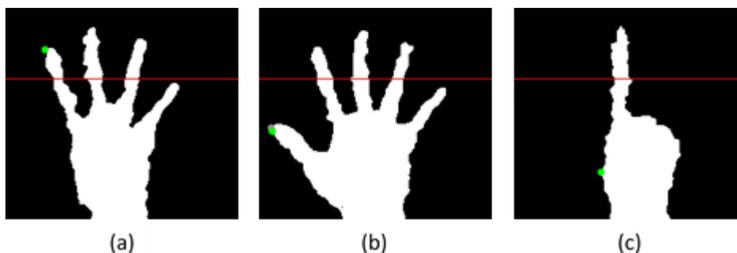


Fig. 3. (a) The coordinate of the leftmost white pixel is smaller than row limit, (b)(c) The coordinate of the left most white pixel is higher than row limit

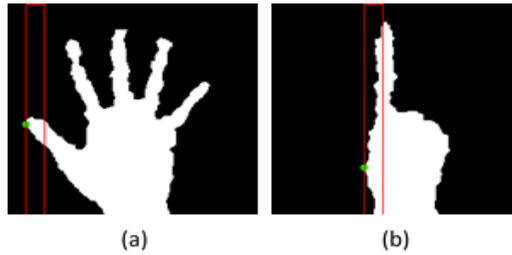


Fig. 4. (a) Area of white pixel on the thumb, (b) Area of white pixel when thumb was not detected.

3.3 Number of open fingers detection

We scanned each pixel from the top of each row. An open finger was obtained if the white pixels with 5 pixels before are black, and the next 4 pixels are white. Each time this condition was fulfilled in the same row, the count would increase by 1. Each time the scan changed row, the count would be reset to 0. The end value of each row would be saved. The end value of each row would be compared, and the maximum value was taken, this value indicated the number of open finger.



Fig. 5. Condition when open finger detected

3.4 Hand gesture classification

We classified the hand gesture base on thumb and number of open fingers. The hand gesture was classified into two conditions, namely when the thumb was detected and when the thumb was not detected. Table 1 presents the signs formed by the hand gesture.

Table 1. Combination between Thumb and number of open finger

Thumb	Number of open finger	Classification target
Not detected	1	Sign 1
Not detected	2	Sign 2
Not detected	3	Sign 3
Not detected	4	Sign 4
detected	4	Sign 5
detected	0	Sign 6
detected	1	Sign 7
detected	2	Sign 8
detected	3	Sign 9

4 Results and discussion

Tests were carried out on 10 subjects with two test scenarios. The first test is performed on the hand gesture classification, the second test is performed on the hand gesture sequence as a password.

4.1 Hand gesture classification testing

Tests were carried out on ten subjects who demonstrated signals 1 to 9 with each signal was taken five times. The distance between hand and sensor was 60-65 cm that produced the best hand image. Table 2 summarizes the classification result of 450 experiment.

Table 2. Hand gesture classification result

	Detected sign	Number of errors	Cause of error		Accuracy
			Thumb detection	Number of open finger detection	
Sign 1	1	-	-	-	100%
Sign 2	2	-	-	-	100%
Sign 3	3	-	-	-	100%
Sign 4	4	-	-	-	100%
Sign 5	5 and 9	1	-	1	98%
Sign 6	6	-	-	-	100%
Sign 7	7 and 8	9	-	9	82%
Sign 8	8	-	-	-	100%
Sign 9	9	-	-	-	100%

Based on Table 2, misclassification occurred at most when performing a classification of sign '7' which was classified as sign '8'. There were ten errors from 450 experiment data, produced an accuracy of 97.78%.

4.2 Hand gesture sequence as password testing

At this stage, hand signals were used as passwords with three hand movement sequences. The test was performed on ten subjects who demonstrated three hand signs as a password; signal '3', sign '4' and sign '5'. Each subject demonstrated a password sign 20 times. Table 3 shows the experiment results. There were 27 errors from 200 trials. Therefore, the accuracy of the password recognition was 86.5%.

Table 3. Result of password testing

Hand sign	Target	Detected sign	Number of trial	Number of error detection
First sign	3	3 and 9	200	1
Second sign	4	4 and 3		8
Third sign	5	5, 4 and 9		18

Table 4. Average time required by the system to identify passwords

Subjects	Time to recognize gesture (s)	Total time to unlock the door (s)
Subject 1	1,41	8,85
Subject 2	1,59	9,04
Subject 3	1,49	9,03
Subject 4	1,45	8,91
Subject 5	1,29	8,06

Subject 6	1,66	9,21
Subject 7	1,25	8,73
Subject 8	1,49	8,25
Subject 9	1,54	9,06
Subject 10	1,37	8,80
Average	1.45	8,79

From the test, the average processing time for the recognition of hand signals is 1.45 seconds, while the total time taken by user perform the three hand sign until the key was unlocked is 8.79 seconds. The system built depended on the distance between the hand and the sensor. The distances of hand and sensors that were too close will make the image degrade while if it was too far away, it was difficult to separate hands and background.

The advantage of the system that it did not require machine learning for the classification process but only used pixel analysis. Hand sign recognition was not determined by who the person was but the sequence of hand signs. Meanwhile, the weaknesses of this system are some parameters determined from empirically trained data. In the next study, more data will be tested and longer hand signal sequence as passwords.

5 Conclusion

This paper describes the design of electronic key systems using hand gesture with Kinect sensors as data acquisition devices. Based on the experiment, the system could recognize nine types of hand sign that represent numbers 1 through 9 using the right hand. Optimum distance for this system was 60 cm - 65 cm from the sensor. The hand gesture recognition system would be implemented into an electronic key system that was able to recognize three different hand signs as a password. The average accuracy of the hand gesture recognition system was 97.78% for one single hand sign and the recognition accuracy of password reached 86.5% with the average time processing of 1.45 seconds.

References

1. R. Passarella and M. Fadli, "Hand Gesture Recognition as Password to Open The Door With Camera and Convexity Defect Method," *Proceeding of The 1st International Conference on Computer Science and Engineering*, pp. 69–73 (2014).
2. H. J. Ahn, J. S. Kim, J.-Y. Shim, and J. S. Kim, "Hand Gesture Recognition for Doors with Neural Network," *Proc. Int. Conf. Res. Adapt. Converg. Syst.*, pp. 15–18 (2017).
3. A. Chaudhary, K. Vatwani, T. Agrawal, and J. L. Raheja, "A vision-based method to find fingertips in a closed hand," *J. Inf. Process. Syst.*, vol. 8, no. 3, pp. 399–408 (2012).
4. T. Q. Vinh and N. T. Tri, "Hand gesture recognition based on depth image using kinect sensor," in *2015 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS)*, pp. 34–39 (2015).
5. Yi Li, "Hand gesture recognition using Kinect," *2012 IEEE Int. Conf. Comput. Sci. Autom. Eng.*, pp. 196–199 (2012).
6. N. B. Bhoyar and P. M. Bartere, "Neural Network based Static Hand Gesture Recognition," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 3, no. 2, pp. 112–116 (2014).

7. N. Febriana, A. Rizal, E. Susanto, “Sleep Monitoring System Based on Body Posture Movement using Microsoft Kinect Sensor,” *Int. Symp. Biomed. Eng.*, vol. 3 (2018).
8. J. Chen, B. Guan, H. Wang, X. Zhang, Y. Tang, and W. Hu, “Image Thresholding Segmentation Based on Two Dimensional Histogram Using Gray Level and Local Entropy Information,” *IEEE Access*, vol. 6, no. c, pp. 5269–5275 (2017).