

Presentation Slide Control Based on Hand Gestures

Nelson Sinaga
Faculty of Engineering
Universitas Negeri Medan
Medan, Indonesia

Baharuddin
Faculty of Engineering
Universitas Negeri Medan
Medan, Indonesia

Hesti Fibriasari
Faculty of Language and Art
Universitas Negeri Medan
Medan, Indonesia

Bakti Dwi Waluyo
Faculty of Engineering
Universitas Negeri Medan
Medan, Indonesia

Joni Syafrin Rambey
Faculty of Engineering
Universitas Negeri Medan
Medan, Indonesia

Abstract: In today's digital world, presentation using a slideshow is an effective and attractive way that helps speakers to convey information and convince the audience. There are ways to control slides with devices like a mouse, keyboard, or laser pointer. The drawback is that one should have previous knowledge about the devices to manage them. Gesture recognition acquired importance a couple of years prior and is utilized to control applications like media players, robot control, and gaming. The hand gesture recognition system is built with gloves and markers. However, using such gloves or tags expands the expense of the system. An Artificial intelligence-based hand gesture detection methodology is proposed in this proposed system. By hand gestures, users can change the presentation slides in forward and backward directions. The use of hand gestures causes the connection simple and helpful and does not need any additional gadgets. The suggested method is to help speakers with a product presentation with naturally improved communication with the computer. Specifically, the proposed system is more viable than a laser pointer since the hand is more apparent and thus can better grab the audience's attention.

Keywords: opencv; hand gesture recognition; presentation slides; machine learning

1. INTRODUCTION

In recent decades Hand gesture recognition has been considered a new technique of Human-Computer Interaction because of its automatic, natural and easiness without requiring input from devices like a keyboard and mouse. For example, detecting spoken language can be done by analyzing lip movements; gaming also uses hand gestures [1]. Now a day's, there are several techniques for hand gesture recognition are in existence on wearable devices like a ring, armbands, gloves, leap motion, and controllers-based motion recognition such as Wii-mote, and ordinary web-camera, stereo cameras, and even using radar, but they still need improvement [2]. So, gesture recognition has gained tons of importance and will not apply. Presentation software control various is one among the various which hand gestures will control.

The machine captures gestures and recognizes them to perform the task. The machine will capture the hand gesture through the camera and recognize it. First background from the captured foreground. This will image recognized motion and remove the filters-out then used for verifying the sign of the gestures. This proposed work aims to implement A.I. in a hand gesture recognition system and use it to control digital presentations using hand gestures.

2. LITERATURE REVIEW

As per the analysis of many other techniques referred to by the researchers, the main aim is to help speakers for a compelling presentation with naturally improved interaction with the computer.

Damiete O. Lawrence and Dr. Melanie J. Ashleigh, author, presented "Impact of Human-Computer Interaction (HCI) on

Users in Higher Educational System: Southampton University as A Case Study". In this paper, Human-Computer Interaction (HCI) perception and impact in the University of Southampton, U.K., an advanced literacy terrain, were measured. The impact of HCI at Southampton University has been positive, showing that getting familiar with HCI generalities ameliorates a stoner's commerce and effectiveness. In conclusion, it can be said that HCI has impacted the literacy terrain as it has impacted other corresponding surroundings [4]. Sebastian Raschka, Joshua Patterson and Corey Nolet presented "Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence". They covered extensively used libraries and generalities, collected for holistic comparison, with the thing of educating the anthology and driving the field of Python machine learning forward [9]. Xuesong Zhai, Xiaoyan Chu, Ching Sing Chai, Morris Siu Yung Jong, Andreja Istenic, Michael Spector, Jia-Bao Liu, Jing Yuan, and Yan Li author presented a Review of Artificial Intelligence (A.I.) in Education from 2010. This study handed a content analysis of studies aiming to expose how artificial intelligence (A.I.) has been applied to the education sector and explore the implicit exploration trends and challenges of A.I. in education [10].

Jadhav & Lobo proposed that static and dynamic gestures are used together to control PowerPoint presentations. To capture and recognize images, Segmentation methodology is used. It also introduces a motion detection feature to change slides [1]. Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang author presented "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor". They presented a vital part-a grounded hand gesture recognition system using the Kinect detector. A new distance metric, FingerEarth Mover's Distance (F.E.M.D.), is used for diversity measure, which represents the hand shape as a hand with each outlet part as a cluster and

penalizes the empty cutlet- holes. More specifically, our FEMD-grounded hand gesture recognition system achieves 93.2 mean delicacy and runs in 0.0750s per frame when using the thresholding corruption cutlet discovery system [6]. Harika et al. authors proposed and approached a method by using vision-based gesture recognition for a computer-assisted slide presentation. Techniques like the Kalman filter, HSL color model, and Skin color sampling are used. Let us consider the accuracy of this proposed model. Skin color detection has an overall success of about 72.4%, Single fingertip detection has an accuracy of 74.0%, the success rate in moving slides is 77%, and the Success rate in controlling finger-pointing is 80% [2]. Wahid et al. proposed and approached a method to Recognize Hand Gestures Using Machine-Learning Algorithms. Considering the accuracy of this proposed model, The SVM algorithm yielded the highest classification accuracies using both original E.M.G. features (97.56%). It normalized E.M.G. features (98.73%) among N.B., R.F., K.N.N. and DA [3].

Ajay Talele, Aseem Patil, and Bhushan Barse presented "Detection of Real author Time Objects Using TensorFlow and OpenCV". This paper introduced a cutting-edge computer innovative and prescient-based impediment detection technique for cellular generation and its packages. Each character image pixel is classified as belonging to an impediment based totally on its look. This paper presented a brand-new approach for impediment detection with a single webcam digital camera [7]. Ahmed Kadem Hamed AlSaedi, Abbas H. Hassin Al Asadi author, presented "A new hand gestures recognition system". They introduced a low-cost system to fete the hand gesture in real-time. Generally, the system is divided into five ways, one for image accession, alternate to pre-processing the Image, third for discovery and segmentation of hand region, four to features birth and five to count the figures of fritters and gesture recognition. The paper answered the challenge of gyration, exposure, and scaling and got the same results when using the right or left hand. The handed system uses only bare hands and the laptop's webcam, which is veritably flexible for the stoner. The system results show that the recognition rate was 96.6%, which is considered veritably good compared with other exploration papers [8]. Dhall et al. Discovered Hand gesture technology and a Convolutional neural network to build a hand gesture recognition application in the "Automated Hand Gesture Recognition using a Deep Convolutional Neural Network model" paper. In this paper, the author used a CNN, which has specific layers such as the input layer and output layer, and in between them, there are some hidden layers. The first hidden layer is the convolutional layer, which detects and extracts features from images, then the Max pooling layer is used for dimensionality reduction [5].

3. METHODOLOGY

3.1 Image Pre-Processing

The point of pre-processing is to improve the standard of the Image so that we will examine it in an exceptionally better manner. By pre-processing, we will smother undesired distortions and upgrade a few essential elements for the application we are working for. Those features might vary for different applications.

Steps for Image pre-processing:

- Select a boundary of the input image within which we will scan for the presence of a person's hand.

- Produce a mask by opting only pixels that match a specified color range.
- Blur the mask image so that missing data points can be filled.
- Draw a hand contour and use Open CV to identify the fingers.

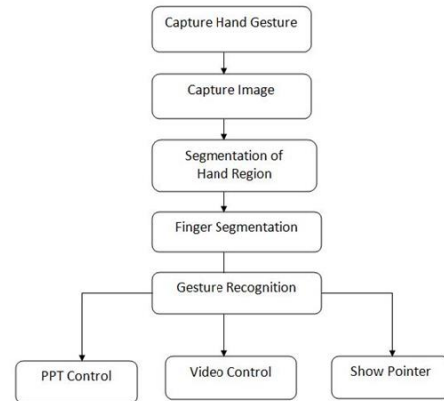


Figure 1. The flow of the system

3.2 Anaconda Framework

Anaconda is used for scientific computing like data science, A.I. applications, massive scope information handling, and predictive analysis, which expects to improve package executives and management. PyCharm is a devoted Python Integrated Development Environment (IDE) executing an enormous scope of required devices for Python engineers, firmly incorporated to establish a helping climate for helpful Python, web, and data science improvement.

3.3 Libraries

3.3.1 OpenCV

OpenCV is used in computer vision, image processing, and machine learning applications. OpenCV guides many programming languages like Python, C++, Java, etc. The identification of objects, faces, or even the handwriting of a human can be made by the action of images and videos. It is an open-source library used to accomplish tasks like face detection, object tracking, landmark detection and many others. It is used to capture the video and to perform the hand detection process:

```
cap = cv2.VideoCapture(0)
```

3.3.2 PyAutoGui

The mouse and keyboard can be controlled to do different things. It is a cross-stage G.U.I. automation Python module for human reality. This third-party library can be installed by using the command pip install pyautogui. It is used to press HotKeys:

```
pg.hotkey('escape')
```

3.3.3 MediaPipe

MediaPipe is a framework that is used for working in a machine-learning pipeline. It is a high-fidelity hand and finger-hunting solution. It is just a part of a single frame that employs machine learning (ML) to derive a 3D marker of a hand. It is used in the hand detection process:

```
self.mpHands = mp.solutions.hands
self.hands = self.mpHands.Hands(self.mode,
self.maxHand,self.detectionCon,self.trackCon)
self.mpDraw = mp.solutions.drawing_utils
```

3.3.4 Keyboard

It records the keyboard action and assists in entering keys; thus can also block the keys until a stated key is entered and affect the keys. It captures onscreen keyboard events as it takes all the keys. This module provides group hotkeys. It is used to press keys:

```
keyboard.press_and_release('ctrl + 1')
```

3.3.5 Numpy

NumPy may be used to carry out various mathematical operations on arrays. It provides effective information systems to Python to ensure treasured computations with arrays and matrices. It offers a massive library of high-stage mathematical features to perform on data structures like arrays and matrices. It is used for Creating a 3*3 kernel:

```
kernel = np.ones((3,3),np.uint8)
```

To define the range of skin color in HSV:

```
lower_skin = np.array([0,20,70], dtype = np.uint8)
upper_skin = np.array([20,255,255], dtype = np.uint8)
```

Convert to the coordinates:

```
x3 = np.interp(x1, (frameR, wCam - frameR), (0, wScr))
y3 = np.interp(y1, (frameR, hCam - frameR), (0, hScr))
```

3.3.6 Time

The Python time module helps represent time in code; representation can be in objects, strings, and numbers. Using this module, other functionality can be implemented, like representing time and measuring the efficiency of your code. It is used to set the frame rate:

```
cTime = time.time()
fps = 1/(cTime - pTime)
pTime = cTime
```

3.3.7 Math

The Python Math Library implements some math features and constants in Python. These features can be used in code to perform complicated mathematical computations. The library does not require installation as it is an essential Python module. It is used to find the length of all sides of the triangle:

```
a = math.sqrt((end[0] - start[0])**2 + (end[1] - start[1])**2)
b = math.sqrt((far[0] - start[0])**2 + (far[1] - start[1])**2)
c = math.sqrt((end[0] - far[0])**2 + (end[1] - far[1])**2)
s = (a+b+c)/2
ar = math.sqrt(s * (s - a) * (s - b) * (s - c))
```

To apply the cosine rule:

```
angle = math.acos((b**2+c**2-a**2)/(2*b*c)) * 57
```

3.4 Finger Count Using Distance Transform

3.4.1 Calculation of finger count

After the skin segmentation, the binary Image of the segmented hand region I.B. is obtained and processed using distance transform. The distance transform method gives the distance (Euclidean distance) of each pixel from the nearest boundary pixel. The distance from the boundary to a pixel in the hand region increases as the pixel is away from the boundary. Using this distance value, the centroid of the palm region can be calculated.

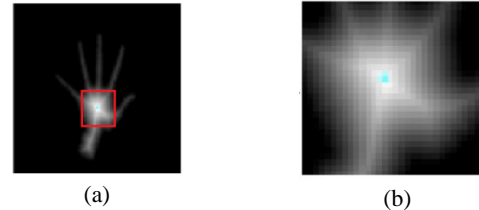


Figure 2. (a) image showing hand region after applying distance transform; (b) the enlarged Image of the region with the red rectangle

Fig. 2(a) shows the image I.D. of the hand after applying the distance transform, and Image Fig. 2(b) shows the enlarged view of the region within the red rectangle. The white color in the center is intense, and the color fades as the distance from the center increases. From this, it is evident that the pixels near the boundary have lower values for distance and those away from the boundary have higher values for distance. This middle region with the highest value for distance is considered the centroid.

The hand region's width will be approximately twice the distance from the centroid to the nearest boundary pixel (say 2d), as shown in Fig. 3. The width of each finger is approximately one-fourth of the hand's width (i.e., ¼ of 2d). A suitable structuring element S (disc) that can erode the fingers thoroughly is chosen, and erosion is performed on the segmented hand region.

$$Rr_1 = I_B \ominus S \quad (1)$$

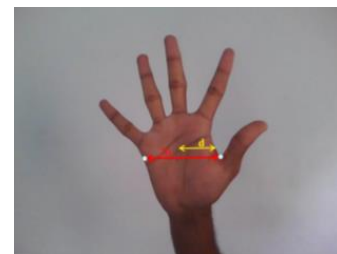


Figure 3. Image showing the width of the hand 2d and d the distance between the center of the hand nearest boundary pixel

The dilated palm region Rp_2 is subtracted from the original binary Image I_B to give the finger regions F_B alone, as shown in Fig. 4(c).

$$F_R = I_R - RP_2 \quad (2)$$

The number of fingers used to represent the gesture is found by drawing a line along the central axis of the segmented finger regions, as shown in Fig. 4(b). The number of lines drawn is equal to the number of active fingers. This count is used to control the slides of PowerPoint.

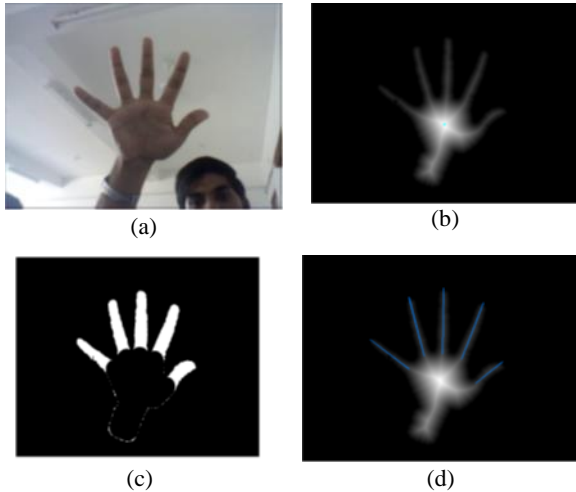


Figure 4. (a) input image I from the user; (b) image I_D of hand region after applying distance transform; (c) the finger regions F_R after erosion, dilation, and subtraction; (d) the fingers count by drawing lines along the central axis of finger regions

Since gesture recognition depends purely on the number of active fingers, any finger can denote a count. Only the count value of the active fingers is taken as input. So, the user can feel free to represent a count irrespective of the finger. Hence value one denoted by the user using the index finger or thumb will be the same as shown in Fig. 5 (a) and (b).

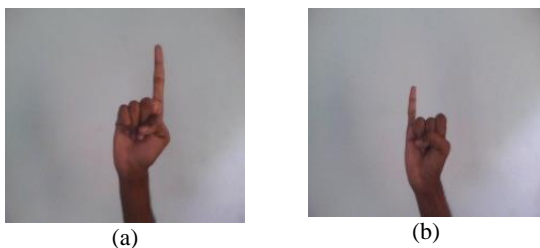


Figure 5. (a) and (b) various gestures to represent the number one

Table 1. Gestures and their functions

No Hand (s)	Finger Count	Functions
One hand	0	The transition from one gesture to another
One hand	1	Next slide
One hand	2	Previous slide
One hand	3	Slide show
One hand	4	Stop slide show
Two hands	The sequence of digit number	Go to the specified slide number

The slideshow is controlled by taking the finger count, calculated using distance transform as input from the user. The gestures used to control the slide show are mentioned in Table 1.

When both hands represent the number of the slide the user needs to navigate, both hands represent the first digit of the number, and both hands are closed. After this, both hands are used again to represent the next digit of the slide number.

4. RESULTS

Place Tables/Figures/Images in the text as close to the reference as possible (see Figure 1). It may extend across both columns to a maximum width of 17.78 cm (7").

- OK (thumb up): The OK gesture presentation will start in presentation mode (Figure 6).
- Two Fingers (victory): Doing two-finger gesture videos will be played or paused in the presentation slide (Figure 7).
- Good: doing a Good gesture presentation will show the previous slide. Users can change the slide backward (Figure 8).
- Three Fingers: The next slide will be presented by making three fingers gesture (Figure 9).
- Four Fingers: The video volume will decrease by making four fingers gesture (Figure 10).
- Five fingers (Palm): Video volume will be increased by making five fingers gestures (Figure 11).
- One finger: by making the one-finger gesture, the user can point some information using a pointer (Figure 12).
- Fist: doing a fist gesture presentation will exit from presentation mode (Figure 13).



Figure 6. OK gesture

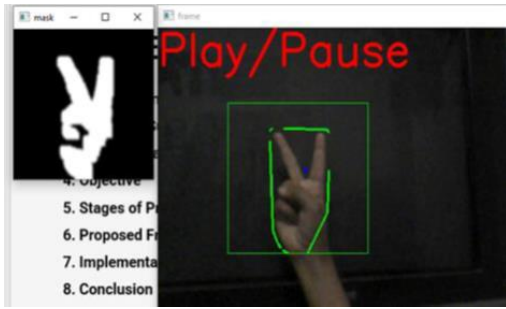


Figure 7. Two fingers gesture

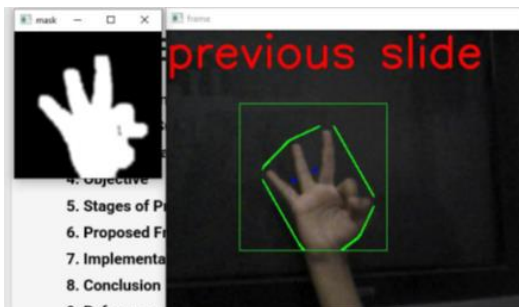


Figure 8. Good gesture

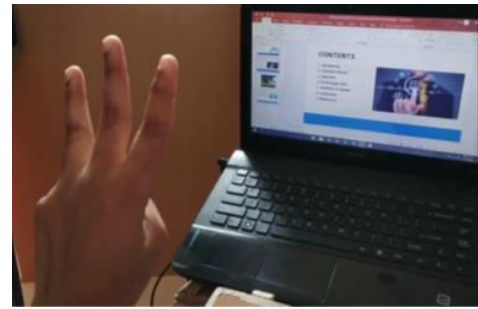
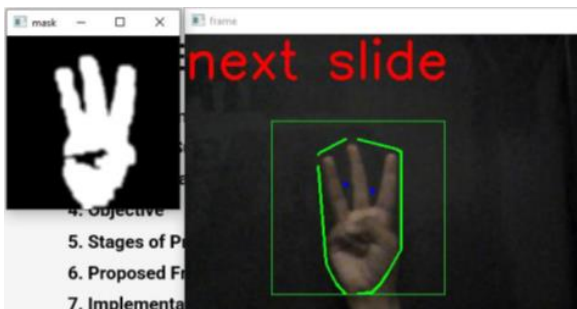


Figure 9. Three fingers gesture

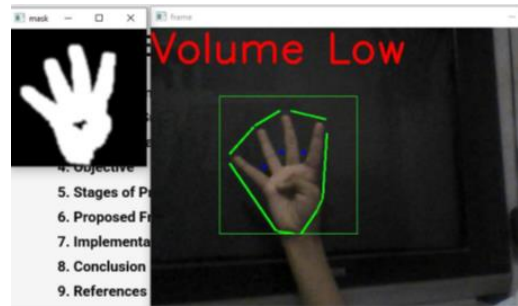


Figure 10. Four fingers gesture

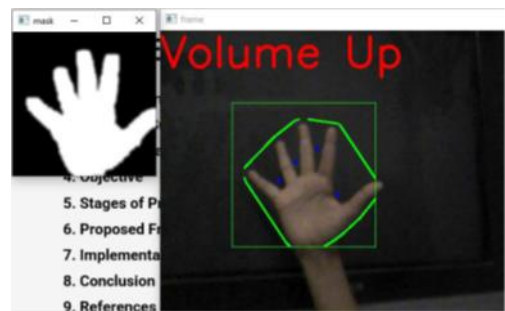


Figure 11. Five fingers gesture



Figure 12. One finger gesture

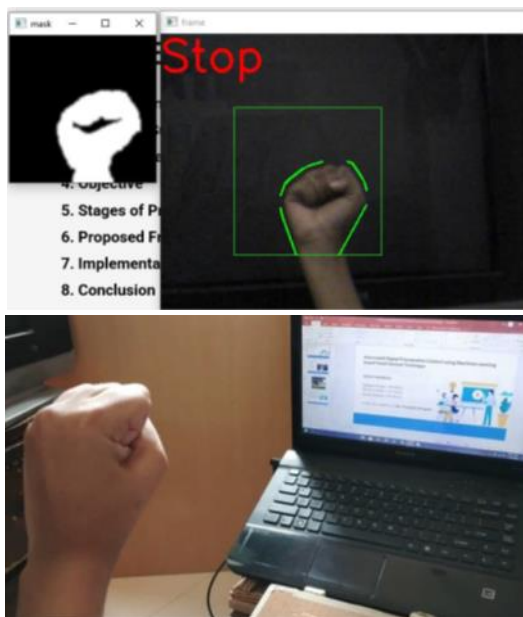


Figure 13. Fist gesture

5. DISCUSSION

Hand gestures are more natural in interaction as they are an essential part of body language, compared with other devices. The use of hand gestures does not require any extra device and makes interaction easy. In this proposed system, an AI-based hand gesture detection methodology is proposed. Using hand gestures would make it easy for the speaker to present more easily. This proposed system aims to develop software that will

help the presenter control the presentation slides by using different hand gestures. With this software, no device like a keyboard, mouse or even remote for changing slides will be needed.

6. CONCLUSION

This proposed system, "Automated Digital Presentation Control Using Hand Gesture Technique," makes presentation easy. The presenter will be able to change slides without using any external device. This will be useful in corporate or institutions where presentation is part of work.

7. REFERENCES

- [1] D. Jadhav, Prof. L.M.R.J. Lobo, Hand Gesture Recognition System To Control Slide Show Navigation I.J.A.I.E.M., Vol. 3, No. 4 (2014)
- [2] M. Harika, A. Setijadi P, H. Hindersah, Finger-Pointing Gesture Analysis for Slide Presentation, Bong-Kee Sin Journal Of Korea Multimedia Society, Vol. 19, No. 8, August (2016)
- [3] Md.F. Wahid, R. Tafreshi, M. Al-Sowaidi, R. Langari, An Efficient Approach to Recognize Hand Gestures Using Machine-Learning Algorithms, IEEE 4th M.E.C.B.M.E., (2018)
- [4] D.O. Lawrence, and Dr. M.J. Ashleigh, Impact Of Human-Computer Interaction (Hci) on Users in Higher Educational System: Southampton University As A Case Study, Vol.6, No 3, pp. 1-12, September (2019)
- [5] I. Dhall, S. Vashisth, G. Aggarwal, Automated Hand Gesture Recognition using a Deep Convolutional Neural Network, 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), (2020)
- [6] Ren, Zhou, et al. Robust part-based hand gesture recognition using kinect sensor, IEEE transactions on multimedia 15.5, pp.1110-1120, (2013)
- [7] Ajay Talele, Aseem Patil, Bhushan Barse on the Detection of Real Time Objects Using TensorFlow and OpenCV, Asian Journal of Convergence in Technology, Vol 5, (2019)
- [8] Ahmed Kadem Hamed AlSaedi, Abbas H. Hassin Al Asadi, A New Hand Gestures Recognition System, Indonesian Journal of Electrical Engineering and Computer Science, Vol 18, (2020)
- [9] Sebastian Raschka, Joshua Patterson and Corey Nolet, Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence, (2020)
- [10] Xuesong Zhai, Xiaoyan Chu, Ching Sing chai, Morris Siu Yung Jong, Andreja Istenic, Michael Spector, Jia-Bao Liu, Jing Yuan, Yan Li, A Review of Artificial Intelligence (A.I.) in Education from 2010 to 2020, (2021)