Low-Cost 3D Holographic Display With Gesture Control

Saima Bibi
Department of Electrical Engineering
Bahria University Karachi Campus
Karachi, Pakistan
hassansaima590@gmail.com

Ayesha Imran
Department of Electrical Engineering
Bahria University Karachi Campus
Karachi, Pakistan
ayeshaimran874@gmail.com

Safwan Saud
Department of Electrical Engineering
Bahria University Karachi Campus
Karachi, Pakistan
safwansaud210@gmail.com

Mukesh Kumar Maheshwari
Department of Electrical Engineering
Bahria University Karachi Campus
Karachi, Pakistan
muekshkumar.bukc@bahria.edu.pk

Bhawani Shankar Chowdhry
National Center of Robotics & Automation
Condition Monitoring Systems Lab
Mehran University of Engineering &
Technology, Jamshoro, Pakistan
bsc_itman@yahoo.com

Waleed Ejaz
Department of Electrical Engineering
Lakehead University
Barrie, ON, Canada
wejaz@lakeheadu.ca

***Abstract*—A way behind the past where teaching and learning were only by video, images, and 2D animations. Now we have live sessions where a classroom is bounded on just a laptop screen. What happened if that live session and 3D imaging are presented in the form of real sculpture by using Augmented Reality? In this paper, we are introducing a low-cost 3D display that presents both recorded and live sessions. Also, this display will be controlled with our fingertips. This application covers sectors of education, medical imaging, and advertisements with good effects and cheap cost.**

***Index Terms*—Hologram, Augmented Reality, Gesture Control**

I. Introduction

The exponential growth of Information and Communication Technology (ICT) has changed the way of life. The development of wearable devices, visual effects, two-dimensional (2D) displays and three-dimensional (3D) has gained attraction in medical, education, business and advertisement. For 3D display hologram technology is used. It is a floating image technology that can display 3D object within its space. The 3D hologram technology is increasing its popularity as it is suitable to be used in developing visualisation skills among learners [1].

Hologram technology is a three-dimensional display that can be viewed by everyone without any need for glasses. A holographic display is a reflection of light, as created by projection to make a holographic 3D image in space. The Hologram projector is used to capture the 3D display on the basis of Pepper's Ghost illusion technique [2]. The Pepper's Ghost technique works using the reflection of mirror [1]. A glass beaker is placed on the stage at an angle of 45 degrees to get a reflection from under the stage. The viewer can see the figure appearing on stage directly through the air and interact with other actors [1], [2].

Visualization skills have a great impact on human understanding and learning. Authors in [3] present the design of the floating hologram method with a reverse pyramid Type for CT and MR diagnosis in a clinical room. The authors considered a laser source and beam splitter for 3D Hologram. Roslan et al., [1] evaluate the student's performance in visualization skills test using 3D Spatial Visual Skills Training (3D SVST) applications. The authors considered types of tests: 3D visualization skills test using paper and 3D visualization skills test using hologram pyramid. The study showed that the hologram pyramid has a positive impact on visualization skills [1]. Authors in [4] present the Holographic Near-Eye 3D Display Method Based on Large-Size Hologram. A comprehensive review of potentials and trends of augmented reality & holography in education is present in [5], [6].

A review piece in [6] explains a vast era of holographic display in medical field imaging. Peak practice in training of doctors and students to convert the human body in 3D model image is appropriately done and future allows to bring these models into real sculpture using augmented reality. Projected AR device in [5] is an example to guide surgeons in performing osteotomie, which makes a clear 3D vision to an operator.

In the above studies lens, laser, and 3D glasses are used for 3D holographic display. A hologram using laser and 3D glass is an expensive solution. Moreover, in the above studies control of holograph display using gestures is missing. Motivated by above studies in this work we present 3D hologram without use of 3D glasses and control it through gesture. The contributions of this work are described as follows:

- Utilize AI technique for gesture recognition by detection of hand key points. Figure 5, 6, 7, and 8 are the following functions of key points:
- Combining recorded and live steaming:
  To make video into four parts as discussed in section B of the implementation part in [12] , the authors use four cameras at four different angles. Here we will just

use only one camera to capture video or image and run the following algorithm to make that screening into four parts.

Here are the following steps for converting one video into four directions for prism:

1. Import Open CV
2. Capture the Video
3. Using command to rotate frame with 0 degree
4. Using command to rotate frame with 90 degree
5. Using command to rotate frame with 180 degree
6. Using command to rotate frame with 270 degree
7. Display step 3 to 6 simultaneously

- Testing visualizing with different cameras:
  Here we test detection rates with respect to the distance and resolution of the camera.II refers to the results that high-resolution capture the best quality streaming and gestures controlling from a near distance give high detection rates.

The rest of the paper is organized as follows. In section

II. Background

In this section, we present the fundamental of holographic display followed by machine learning techniques for gesture control.

A. Open Computer Vision (CV)

For real-sight vision to understand a digital picture or video, Computer Vision (CV) plays an important role. As a part of digital image processing CV working categorized into 3 steps: In the first step, image of analog sign is converted into a digital form of matrix, image acquisition.

In second step, algorithm 1 is applied to detect hands, its structure, depth and finger counting, image processing. In final step, if else condition will be applied, and decide the functionality of the gestures based on program, image analysis.

Open CV is software that implements computer vision algorithms. To replace CV we can also use Point Cloud Library (PCL), NVIDIA CUDA-X, BoofCV, etc. but Open CV is integrated with python and creates the code is simpler and therefore can be easy to implement. Thanks to this, we can read and write images and videos. Face expression, eye movements, and detection of some objects can easily be done easily.

B. MediaPipe

MediaPipe is an open-supply platform for constructing pipelines to method perceptual data of various modalities, together with video and audio. This framework works in the computing device, Android, iOS, and embedded gadgets like Raspberry Pi. MediaPipe is a hand and finger tracking solution that employs gadget learning to deduce 21 landmarks of a hand from a separate frame in 3-d. It consists of three main models working collectively: palm detector version, hand landmark version, and gesture recognizer.

To reduce the data augmentation and allow the network to dedicate most of its capacity towards coordinate prediction accuracy, the palm detector model is used to detect first hand locations and then returns an oriented box of hand. The hand landmark model provides high-fidelity 3D hand key points inside the detected hand region.

To allows reading 21 hand key points in which we set gesture on a particular set of points. Author in [11] describes the classification on which this detection is based on two types of classification, Binary, and 3D classification. Binary will detect whether the hand is left or right on the other side 3D classification use (x,y) coordinates to detect the tips of the fingers.

C. PyAutoGUI

The proposed approach aims to use non-commercial tools that address user concerns about autonomous and free use of computer systems. H. PyAuto GUI. This tool is an open source tool that enables automation and testing of graphical user environments (GUIs) through a computer vision approach. This tool allows you to automate all the graphic information on your computer screen, such as in pur case controlling of keyboard keys with respect to gesture intake by CV. PyAutoGUI allows us to test web applications without user intervention and enables automation of any web application. [13]

D. Recorded and live sessions

Here we brought an option for user to perform display using both option of live streaming and already downloaded video. this user interface helps to switch in both units and display desired type of video.

Section III-A, describes the detail forming of 3D visual. For already recorded video or image samples, we need only one camera to capture user's hand gesture. When it comes to live session, we need two cameras one of which captures the user's hand gestures and the second camera is used to stream a live video to display. Detection rates depend on camera pixels and distance between the user and 3D setup, refer to II.

III. Design and Implementation

In this section we present the construction of 3D holograph display and gesture control.

A. 3D Holograph Display

To design a 3D Holograph we considered the concept in [2]. In [2] pyramid is used which is easily manipulating by sheet of plastic (acrylic sheet). Acrylic sheet are used as it is light weight, low price and easily available. The device generates the 3D visual display for the viewer and creates the holographic image or video that visible to everyone without the use of 3D glasses. By interference of light rays creates 3D hologram that can be viewed from any angle and the user can see the object without any use of 3D glasses, lenses or other props from multiple side back, front, left and right. For the Hologram projection, the LCD TV for any smart phone is used to play the video on screen in 4 parts which is $45°$ apart. These video in project in the middle of each pyramid 4 parts,

TABLE I
HAND KEYS NUMBERS AND THEIR FUNCTIONS (LANDMARKS REFERENCE IS FROM [11])

Key points	Finger counts	Function
8	1	Right and left control buttons to move forward and backward video.
12	2	Up and down buttons to control the volume high and low of the video.
16	3	Window + down button to minimize windows screens.
20	4	Space bar to control pause and play activity.

resulting 3D image from refraction and reflection of each side [7] as shown in Fig. 1.

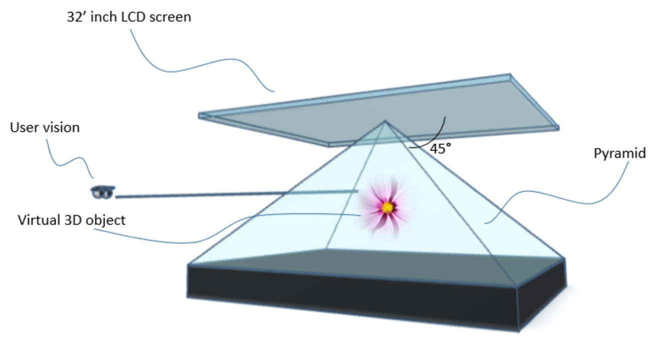

Fig. 1. Sight model of prism and LCD position

In this work for 3D holographic display with gesture control, we have LCD TV of 32-inch on which the video in 4 parts rotate 45-degree apart from each other. These 4 sections of videos are reflected on the 4 sides of the pyramid made of acrylic sheet, although it allow the user to see 3D hologram objects from all sides at the same time near to the real image.

B. Gesture Control

After the hologram 3D display, we use the camera/webcam to get the gesture from user like fingers count in order to control the video with gestures. Here we use six gestures to play, pause, forward, backward, volume up and down for holographic video. All of these gestures are used as input from user and decoded into the controlling display by using python libraries.

To detect an object in real time open CV [8]. For hand gesture tracking MediaPipe [9] with OpenCV is used. MediaPipe is a framework used to build the pipelines for hand tracking on the basis of hand landmark Model. MediaPipe have two states for tracking: (i) palm detection and (ii) hard landmarks. Palm detection wok on complete capture image of webcam and provides the cropped image of hand only, whereas hand landmarks provides the land mark on captured image of hand. There are 21 hand points.

With MediaPipe and Open CV, pyautogui library is used to control the mouse and keyboard of pc to automate interaction with other application [10]. The complete gesture recognition and control procedure is discussed in Algorithm 1. In algorithm 1 first initialize the Open CV to capture the video. After that it detect the hand key point and read the point position.

Algorithm 1 Gesture Recognition Procedure

```
1:  Import Open CV
2:  Capture the Video
3:  Detect hand key points
4:  Read point position
5:  if Position not read then
6:      Check camera connection or Go to 3 again
7:  else
8:      Display the gesture
9:      Control video according to gesture
10:     while Detected hand key points do
11:         if Detected single finger then
12:             Fast forward: Finger move right side
13:             Fast backward: Finger move left side
14:         else if Detected two fingers then
15:             Volume Up: Fingers move upward
16:             Volume Down: Fingers move downward
17:         else if Detected three fingers then
18:             Adjust the brightness to 50%
19:         else if Detected four fingers then
20:             Play and pause the video
21:         end if
22:     end while
23: end if
```

If position is not read then this process is repeated. After that it display the gesture and an action is performed according to gesture. The details of actions for different gestures are presented in algorithm 1. Four fingers counting are responsible of video forward and backward, volume high and low, screen brightness and pause play of pictorials respectively.

IV. RESULTS

To make a Pyramid for hologram projector, we utilized the transparent acrylic sheet for pyramid with 2.5 mm of thickness. This acrylic sheet is light weight, low priced and easily available. [7] The resultant display is an illusion object of projection of lights refracted with 45-degree angle between prism surface and LCD display. A black stand of 4 sided pillars allows one see the object from all directions, shown in Fig. 2.

Figure 3 presents the holograph display in present of light. Figure 4 presents the holograph display without light. The hologram without light is good as compared to hologram display with light.

Different gesture are shown in Fig. 5 to 8. The Fig. 5 present the gesture control using one finger. One finger gesture is

TABLE II
GESTURE DETECTION RATES OF VARIOUS RESOLUTIONS

Camera type	Resolution (pixels)	Detection Rate (%)	Distance (m)
Webcam	640	96.7	1
Cellphone camera	760	98.09	1
Computer camera	320	97.3	0.2
High resolution camera	1080	99.54	1
Logitech webcam	320	95.07	2

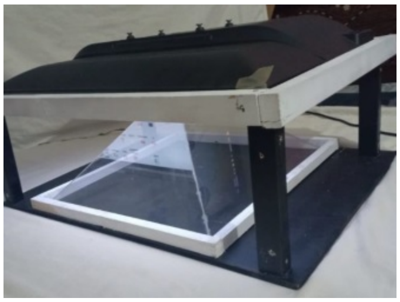

Fig. 2. Holographic model with outside support for LCD and inside prism

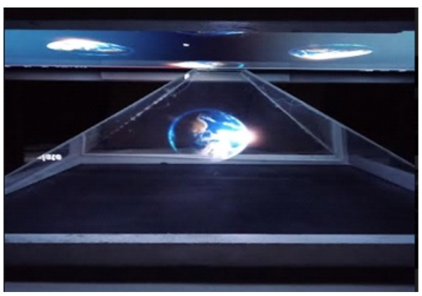

Fig. 3. Model without light interference (formation of the sharp object)

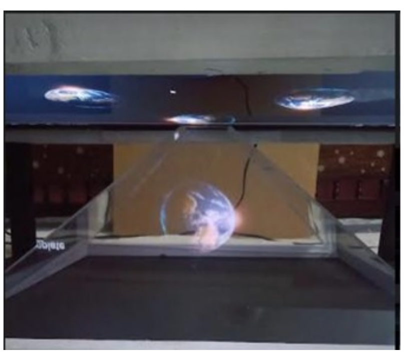

Fig. 4. Model with light interference (formation of the blurred object)

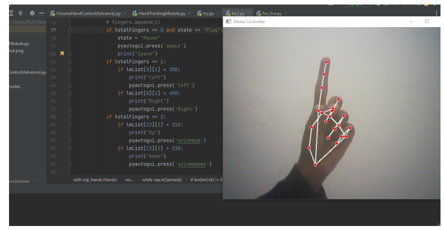

Fig. 5. Gesture control with single finger

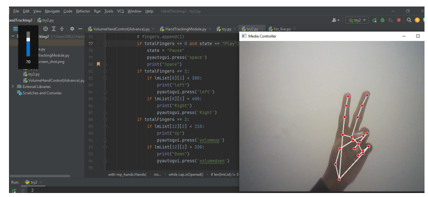

Fig. 6. Gesture control with two fingers

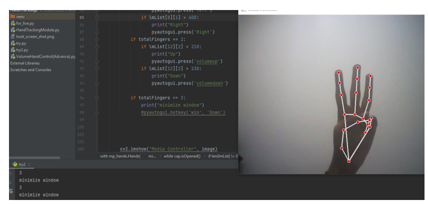

Fig. 7. Gesture control with three fingers

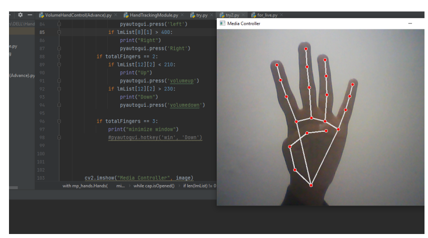

Fig. 8. Gesture control with four fingers

used to rewind video fast forward and fast backward. When we move the finger right side the video rewind fast forward and rewind fast backward on moving finger to the left side. Fig. 6 delineates the two finger gesture control. Two finger gestures is used to adjust the volume of the 3D video or the display screen. When we show the two fingers move upward the volume of display become increases and when we move the fingers downward the volume become decreases, shown in Fig. 6 left top corner. Three finger gesture control and four finger gesture control are shown in Fig. 7 and Fig. 8, respectively. Three finger gestures is used to set the brightness of the 3D hologram video, it set the brightness to 50% of the total. As we use the LCD TV for display, by using this gesture we set the brightness at 50% of LCD TV. Four finger gestures is used to play and pause the 3D video. When we show the four finger gesture the video start play or pause, in actual this gesture work like a space bar. As we use the space bar to start or stop the video. Following are key results points using different resolution camera to detect hand gestures rate with respect to distance, shows higher the resolution and nearer the distance, have high detection rates in Table II.

V. Conclusion

The main objective of this project is to combine human interface with a 3D display at a very low cost, including recorded and live sessions. Same models can build for classrooms, labs, operation theaters, business or conference rooms, and on the top of buildings or shops for advertisement purposes. Requirements are just limited to camera, glass prism, display screen and an embedded tool to run python script e.g. Raspberry pi.

Furthermore, much work remains to be done on the pyramid hologram in the future, in which the best quality high resulted cameras can be used to achieve 100% display results with HD view. In addition to improving the hand key points detection, this work will also be done using GPUs which give full detection and modify the running time of the code. This concept can also move on IOT based techniques where users can operate remotely.

References

[1] R. K. Roslan, et al., “3D spatial visualisation skills training application for school students using hologram pyramid,” *JOIV: International Journal on Informatics Visualization,* 1.4 (2017): 170-174.

[2] ”Peeper Ghost: Hologram illusion”, [Online] https://www.scienceworld.ca/resource/peppers-ghost-hologram-illusion/

[3] M. Kim, et al., “Design of the Floating Hologram Method with a Reverse Pyramid Type for CT and MR Diagnosis in Clinical Room,” *Diagnostics,* 12.5 (2022): 1157.

[4] S. J. Liu, et al., “Holographic near-eye 3D display method based on large-size hologram,” *Frontiers in Materials* (2021): 355.

[5] K. Takrouri, et al., “AR Technologies in Engineering Education: Applications, Potential, and Limitations,” *Digital,* 2.2 (2022): 171-190.

[6] H. W. Yoo, et al., “The potentials and trends of holography in education: A scoping review,” *Computers & Education,* (2022): 104533.

[7] Nazia Fakhruddin.medium.com,11 june 2019. jpeg image, “projection pyramd uses the principle”, https://miro.medium.com/max/966/1*-lle3trdN9US7RE_MlqzlQ.jpeg

[8] Opencv.org, “tutorial”, https://opencv.org/

[9] Github.io, media pipe, “media pipe, hand gesture, key point detection,” https://google.github.io/mediapipe/ 42 42

[10] PyAutoGUI, “welcoom to PyAutoGUI, documentation,” Examples-pyautogui, https://pyautogui.readthedocs.io/en/latest/

[11] Chen, Boxu, et al. ”Visual gesture recognition based on hand key points.” Journal of Physics: Conference Series. Vol. 2024. No. 1. IOP Publishing, 2021

[12] Siang, Chan Vei, et al. ”Interactive holographic application using augmented reality EduCard and 3D holographic pyramid for interactive and immersive learning.” 2017 IEEE Conference on e-Learning, e-Management and e-Services (IC3e). IEEE, 2017

[13] Silva, J. L., Ornelas, J. D., & Silva, J. C. (2016). Supporting GUI exploration through USS tool. Supporting GUI exploration through USS tool, (4), 1-15.