

INTRODUCTION TO INNOVATIVE PROJECTS - PHY1999
SLOT: TE2

HMI Enhanced Game Using Hand Gestures

GROUP 2

Sagar Sethumadhavan (19BCE2460)
Manishraj Khan (19BCE0737)
Ankesh Singh (19BCE2502)
Sudhanshu Agarwal (19BCE2203)
Shaikh Mohmmadali Akbar (19BCE0738)
Zihan Azad (19BCE2442)

01



Background

During these unprecedented pandemic situations, many new ill habits have become the new normal. With lockdown imposition across the country we are hardly able to step out of our homes and thus end up spending almost the entire day either on the couch or the chair with our hands either on the keyboard or mouse or the TV remote, leading to risks of developing disorders like Carpal Tunnel syndrome. So it is important that we find a way to do a major part of the tasks, if not all, without having to use the conventional methods.



Abstract

Hand gesture recognition systems provide users an enhanced interaction experience as it integrates the virtual and the real world object. Gesture recognition based interactions, provide a more realistic and immersive interaction compared to traditional peripherals. The gesture based interaction interface showcased here can be applied towards many applications like virtual reality, communication techniques and video games.



The focus of our project is on games as the application domain for this interaction method. Gestures, particularly hand gestures are also faster and possibly could be more accurate than using the keyboard – mouse combination of peripherals. The non-touch 4 system is a modern method of computer-interface technology capable of revolutionizing human-computer interaction.

Objectives

04

- To create a human machine interaction system as an alternative for the classical control devices.
- To demonstrate the application of hand gestures using a video game.



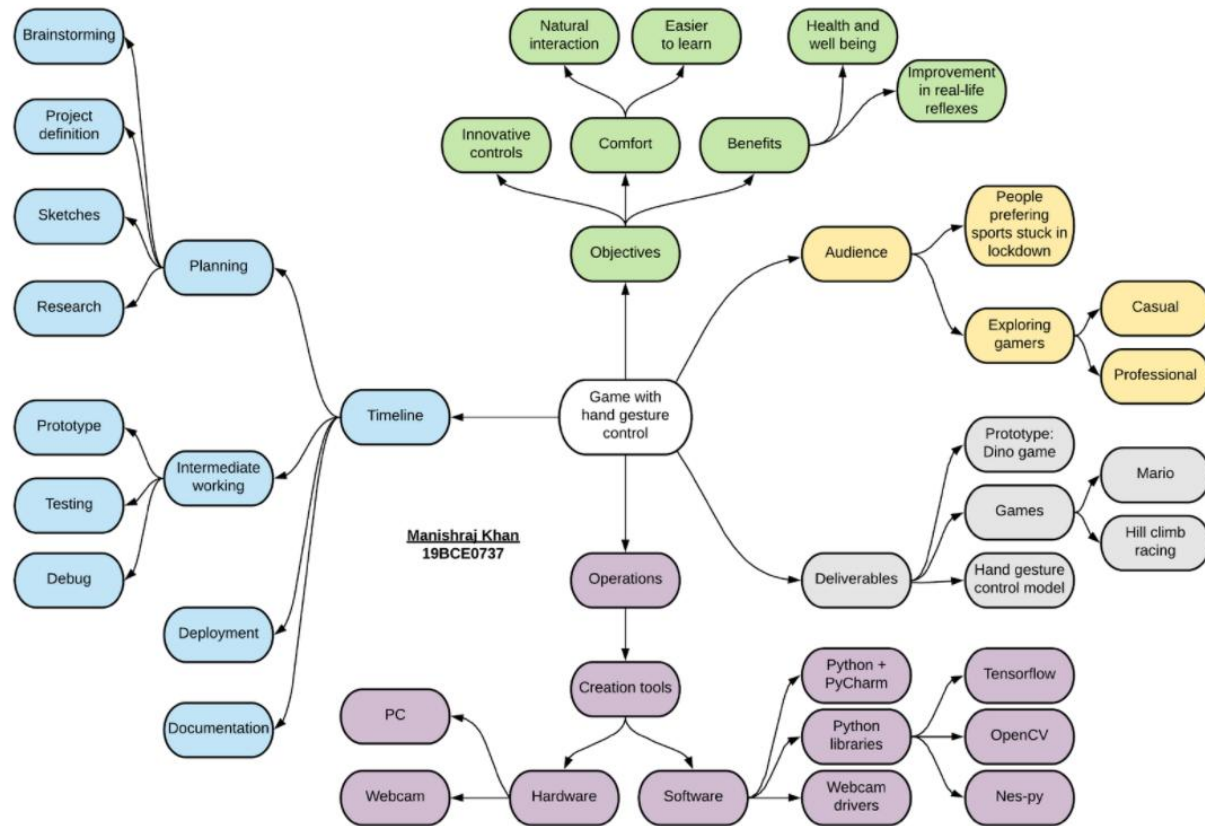


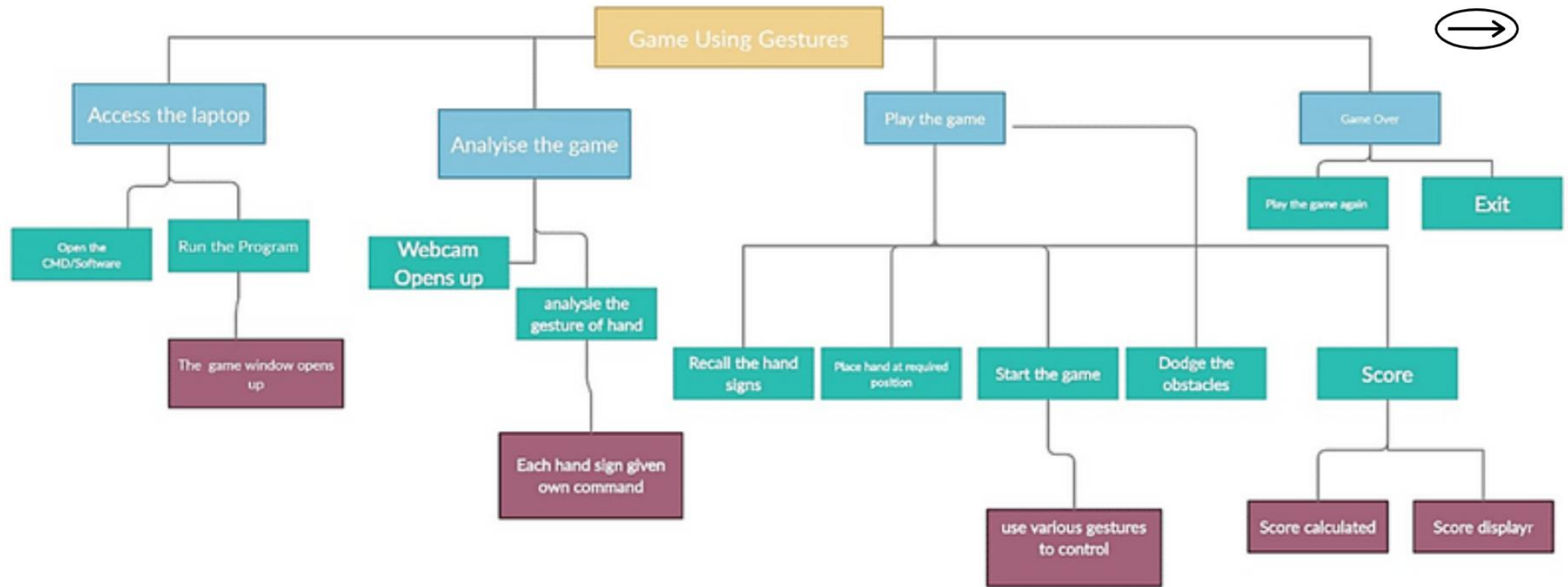
05

Proposed work



The main domain of the project lies in developing a game with the application of gesture recognition system. The usage of hand gestures to promote virtual activity as one does in the real world, results in the main advantage that the game can be played in a virtual space with enhanced interaction much better than conventional peripherals.





Hierarchical task analysis

Innovation and novelty

08

- Gesture recognition based interactions provide a more realistic and immersive interaction compared to traditional peripherals, and thus, provide a fun and interesting way of playing a video game.
- The game controls are much more natural and adaptable than the usual keyboard controls, and therefore are easier to learn for people of any age or of any level of technical proficiency.
- Compared to traditional controllers, gesture controls have decreased risk of developing medical disorders due to prolonged use. In fact, as it promotes physical activity, it helps to improve health and general wellbeing.



- The player would play the game with the help of his/her fist or palm.
- When the palm is gestured up, the character moves up and vice versa.
- When an appropriate hand gesture is initiated, the required operation is performed by the machine.
- The computer's front camera is the sensor which detects motion.



User journey

10

When the user runs the program the game in the command prompt window opens up and accesses the webcam of the user. Once the game starts, the player needs only to move his hand to make his game character move and dodge the obstacles and score points. If the user fails to dodge any obstacle the timer stops and displays the time and total score with a message “GAME OVER”. The user is at his will to begin the game again or to quit the game. The developed game will create randomised obstacles that the player needs to dodge in order to score points and beat his last highest score. The game will also have a timer indicating the time which starts from the moment the player starts the game and would stop if he fails to dodge any obstacle.



Execution

TOOLS

- PyCharm python IDE
- Webcam drivers
- Python libraries
 - Tensorflow
 - OpenCv
 - Nespy

METHODOLOGY

Some of the processes used for gesture recognition -

- Skin Colour Detection
- Hand Region Segmentation
- Gesture Recognition and comparison.





Methodology

Palm recognition

In order to detect the skin from the video, we need to find out the character of the skin. Detecting skin-coloured pixels, although seems a straightforward easy task, has proven to be quite a challenging task in images that are captured under complex unconstrained imaging conditions. So we developed a method based on the colour feature for most humans.

The recognition criteria:

- $R > 85$
- $R - B > 10$
- $R - G > 10$

Based on such criteria, we can efficiently segment the skin from the background, which can be considered to be the human part. Then we can convert the image into a binary image. The results are shown as figures.



Methodology

Pre-processing

Due to the nature of the image, the hand region may have holes and cracks, which will definitely affect the accuracy of hand gestures. Usually the binary image will be noisy, so image pre-processing is necessary, which fills the holes. In our algorithm, the diffusion coefficients are defined according to the distance and direction between the damaged pixel and its neighbourhood pixel.

Contour Extraction and Hand Region Segmentation

After we remove the noise in the image, we need to extract contours. We consider each point cluster as a contour. Among these contours, there is only one contour which represents the hand region. Furthermore, the hand region and face region are the largest two contours. Based on such a fact, the problem of finding the hand from the contours becomes the problem of separating hands from faces. So, we made a specific Box inside which the hand is placed by the user and is detected.



Methodology

Gesture recognition

- The four different scale spatial pyramids are pooled to get the size of 1/4, 1/8, 1/16, 1/32 feature map respectively. So the different scale features can be captured. Then, global average pooling is used to obtain the weights of global abstract features as channel dimensions at lower levels. Finally, the final probability score of each class is obtained by using a fully connected layer and softmax.
- If $\text{angle} \leq 90$ and $\text{count defects} \geq 4$, then the Recognized image is Open Hand. And this condition is not satisfied then the Recognized Image is Closed Hand.

$$a = \sqrt{(end[0] - start[0])^2 + (end[1] - start[1])^2}$$

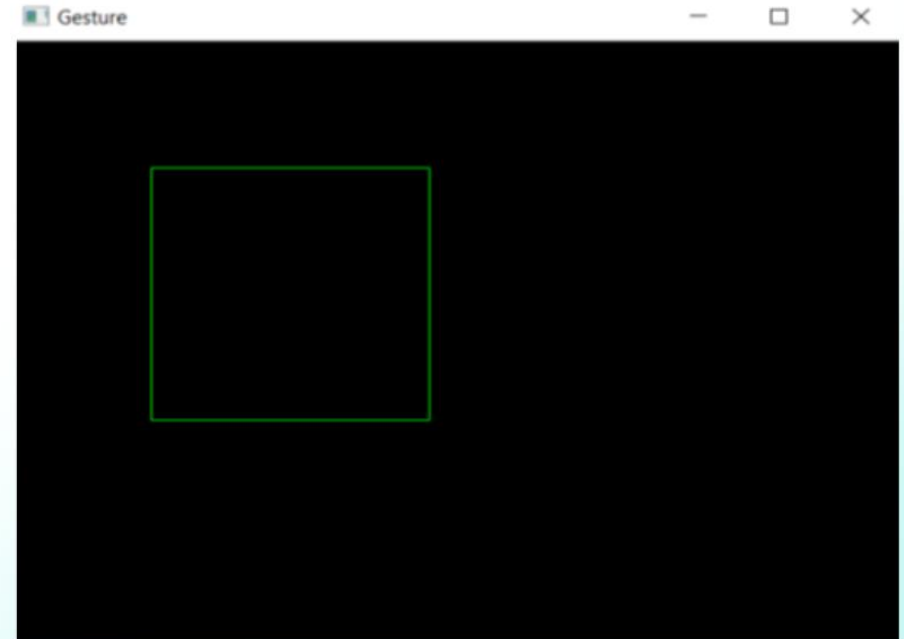
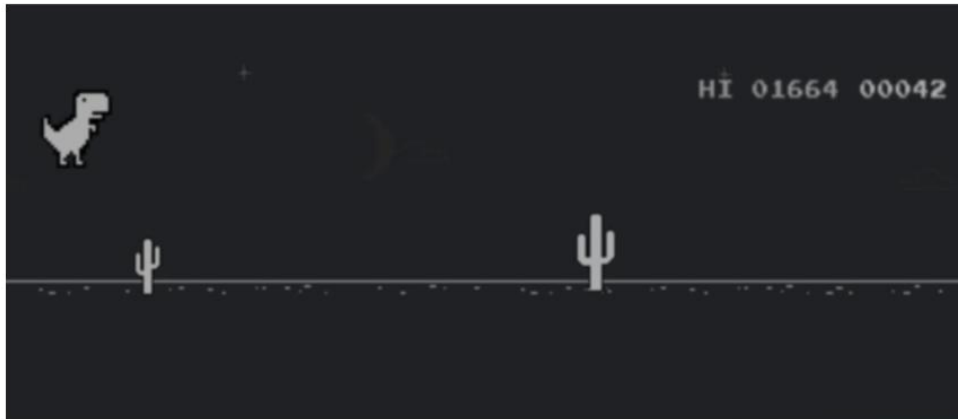
$$b = \sqrt{(far[0] - start[0])^2 + (far[1] - start[1])^2}$$

$$c = \sqrt{(end[0] - far[0])^2 + (end[1] - far[1])^2}$$

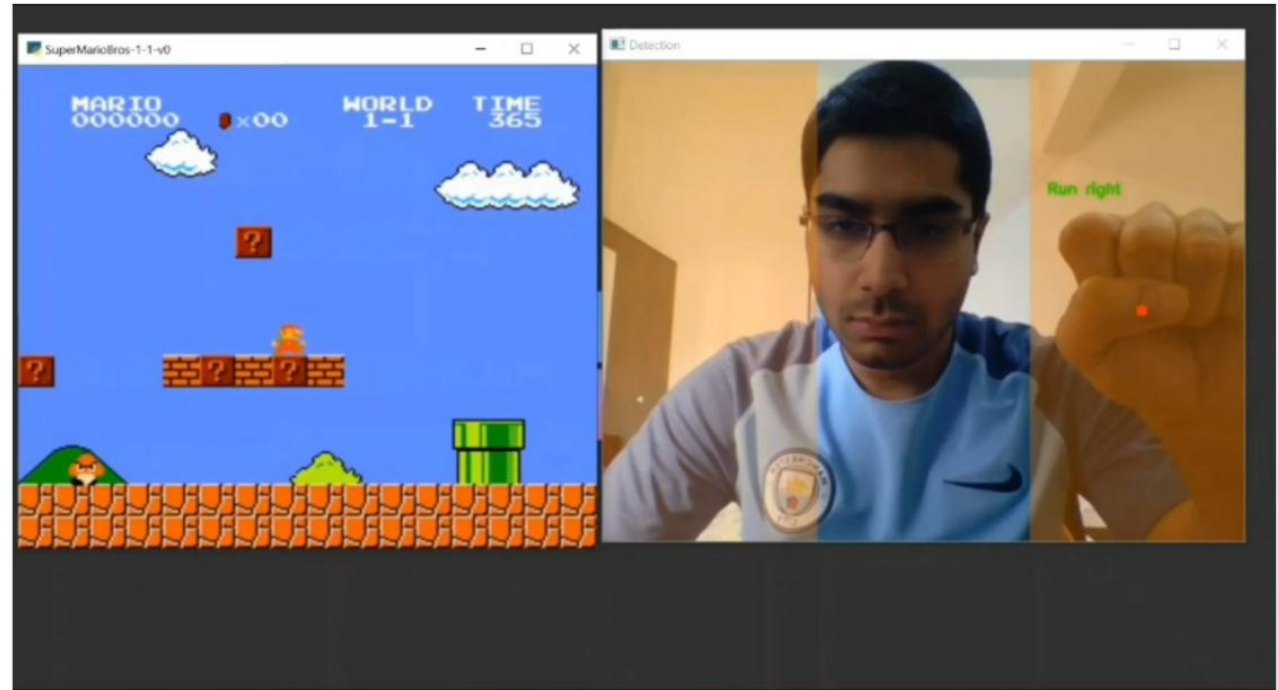
$$\text{angle} = \cos^{-1} \left[\frac{b^2 + c^2 - a^2}{2bc} \right]$$



Output



Output



Link to demo video:

<https://drive.google.com/file/d/1LMtnWvqle02X3QryX5maWkhySHyxMyiH/view?usp=sharing>



Comparison with existing models

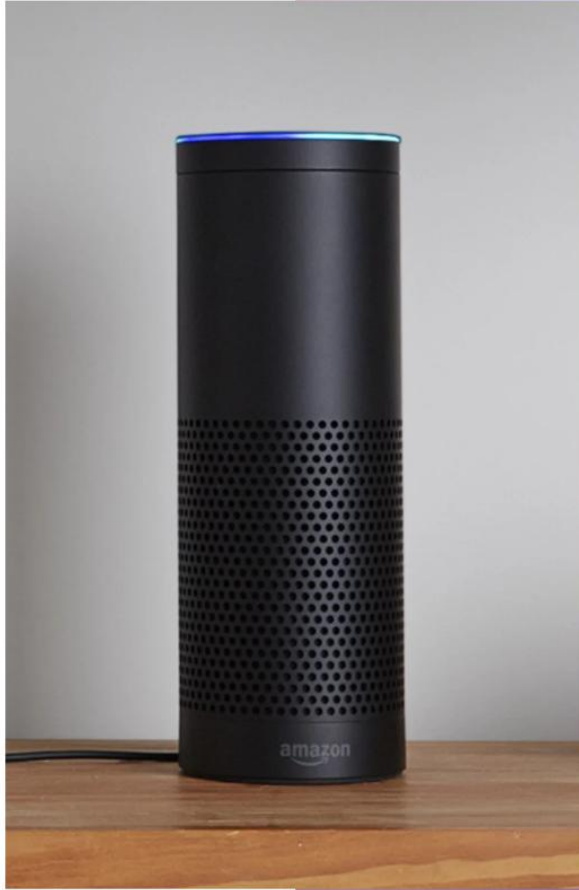


17

1. REMOTE GESTURE DEVICES

- Although these systems are based on gesture recognition, these are dependent on separate physical hardware control devices, and hence not as convenient as our system.
- The risk of developing medical disorders like Carpal Tunnel Syndrome persists.

Comparison with existing models



18

2. VOICE COMMANDS

- Although it may sound different than gesture controls, it can be used to achieve many of the basic controls that we achieve using hand recognition.
- Gesture recognition control is much faster and more efficient compared to voice commands.
- Not appropriate for users with hearing or speaking difficulties or disabilities.
- In case of voice commands, there is a major privacy issue, hence cannot be used in all environments and circumstances.

19

Future prospects



INCLUSION OF VIRTUAL REALITY

In the future, the game can even be developed to include virtual reality for better experience.

CHOOSE A RATE YOU KNOW YOU DESERVE

Presentations are tools that can be used as demonstrations, lectures, speeches, and more. Improved on a large server to incorporate multiplayer levels, hence promoting better real life social experience

BETTER GRAPHICS

The graphics at present are proposed to be simple due to present limitations of the development team, but later better graphics can be used for better experience

3D COVERAGE

Better ML models can be trained with larger datasets that use multiple webcams in order to have a 3D view that can be integrated with either an AR or a 3D projection system to provide complete realistic interaction

20

Thank you