ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

REAL TIME GAME PLAY USING HEAD POSE ESTIMATION

Mrs.L.N.B.Jyotsna[1], N.Jagadeesh[2], G.Vamsi Kumar Reddy[3], S.J Pradeesh[4], N.Haran Raj[5],
[1]Assistant Professor, [2,3,4,5]Students,
Dept. Of Computer Science & Engineering, Dhanekula Institute of Engineering and Technology, A.P., India.

Abstract — Voice, and Gestures are the widely used ways of communication. We have voice-controlled gadgets like Alexa and Siri-enabled devices to control our homes using voice commands. We do have gesture-controlled TVs. Even the dashboards on top-of-the-segment luxury cars are gesture-controlled. The video games are gesture controlled. Most gesture-controlled games use hand gestures. Another kind of gesture is head gestures. In this system, we are using head pose estimation, which mainly concentrates on the movement of the head monitored using a webcam or externally connected camera to the computer. This system supports real-time gameplay using head movement. We are using the python OpenCV library and MediaPipe Face Mesh solution framework to recognize the head movement. This system mainly provides an enjoyable gaming experience for physically disabled people (1% of the world's population). We are considering a popular game "Subway Surfers" as a demo for the system. This is because it is very popular and has more than one billion downloads from the google play store. The head movements are used as controls for the game avoiding physical touch with the gaming devices.

Keywords — Head Pose Estimation, Gestures, Gesture controlled games, MediaPipe, OpenCV, Head movement detection, Facial landmarks detection, Computer Vision.

INTRODUCTION

The gesture is defined as "a movement of part of the body, especially a hand or the head, to express an idea or meaning" [1]. Gestures serve many roles in communication, learning, and understanding both for those who view them and those who create them. Gestures can be understood by machines to recognize the user's commands. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Gesture recognition enables humans to communicate with the machine and interact naturally without any mechanical devices. Playing games via gesture makes the experience much deeper and rich, making it preferable to using the old keyboard and mouse.

Games are primarily played on computer systems using various input devices of which keyboard and mouse are the most often used. There also exist other input devices which

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

translate a user's body movement into actions in the game (Example Kinect). This makes the playing experience better and makes the user connected to the game. Also, these devices may be expensive to purchase for the average user hence not everyone can use them. An alternative to playing games through traditional input devices is using gestures that the system can understand. Hence this topic, "Real time game play using Head Pose Estimation" comprises using head gesture recognition for playing games.

Fig. 1: Real time game play using Hand gesture

Being technologically advanced is all that a developing world needs to be. This made us control machines with gestures from controlling them with traditional input devices. Gesture controlled games are one of our achievements in growing technology. Most gesture-controlled games use hand gestures as shown in Fig.1. Based on a 2011 WHO survey, the world consists of nearly 75 million people who are physically challenged, which approximately equals 1% of the world's population. Every person in the world seeks relaxation by playing any kind of game. So, we are proposing a system that allows users to play games with their head movements(gestures) and does not require any use of physical devices or hands on the computer. The main motto of this system is to provide an enjoyable gaming experience for physically challenged people.

RELATED WORK

Head pose is an important cue in computer vision when using facial information. Over the last three decades, methods for head pose estimation have received increasing attention due to their application in several image analysis tasks. A research article published in Signal Processing: Image Communication Volume 99 by Khalil Khan [2] surveyed advancements in head pose estimation in the past 10 years. Many researchers are working on finding the head movements of a human. Anastasiya Zharovskikh [3] who published an article on Head Pose

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Estimation with computer vision in InData Labs in January 2021 had written clearly about head pose estimation and its approach using computer vision. In January 2022 Andrey Sheka and Victor Samun [4] proposed a response-based method of knowledge distillation (KD) for the head pose estimation problem. Their method consists of two stages neural network (NN) and knowledge distillation (KD). One of the Head pose estimation applications is driver movement monitoring. A paper written by Zhongxu Hu, Chen Lv, Yanxin Zhou, Yiran Zhang, and Wenhui Huang [5] in March 2021 proposed a low-cost and mark-less driver head tracking framework based on the head pose estimation model, which makes the view of the simulator can automatically align with the driver's head pose.

A sensor-based approach is also proposed by Linh Nguyen Viet, Tuan Nguyen Dinh, Hoang Nguyen Viet, Duc Tran Minh, and Long Tran Quoc [6] in 2021. They considered seven main phases, with the beginning being the process of wearing sensors for humans. A new model is called FSANet-Wide and uses this model with a new dataset UET-Headposetrain and two other datasets. The system includes a head-mounted sensor module, an Arduino board, a surveillance camera, a chin rest for fixing the head, and a server to control, store and process data. In this model, the angle sensor fixed to the head of the human transmits the angles of the head position through Arduino and the obtained results are combined and compared with other datasets to obtain the actual results. Developing a gesture recognized gaming is not a new concept many applications provide users to experience gaming using hand gestures. In a paper published by Dr. Parameshachari B D, Rubeena Muheeb, Nagashree R N, Deekshith B N, Keerthikumar M, Rashmi P, and Rachana C R [7] proposed a gesture recognition-based car gaming. By considering these research works we are now proposing a system that facilitates users to play the game using their head movements. There are no or minimal works on real-time gameplay using head gestures. This motivated us to work on head gesture recognition and incorporate the gaming controls with the head movements.

PROPOSED METHOD

Head Pose Estimation: The main objective of head pose estimation is to find the relative orientation (and position) of the human's head concerning the camera. The head pose estimates can provide information on which direction the human head is facing [3]. Despite the head pose estimation task may seem to be easily solved, achieving acceptable quality on it has become possible only with recent advances in Deep Learning. Simply put, head pose estimation means detecting the position of a human head in the image. Particularly, it means detecting the

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

head's Euler's angle – yaw, pitch, and roll. They define the object's rotation in a 3D environment.

MediaPipe: MediaPipe is a Framework for building machine learning pipelines for processing time-series data like video, audio, etc. This cross-platform Framework works on Desktop/Server, Android, iOS, and embedded devices like Raspberry Pi and Jetson Nano [8]. Fig.2 shows the process flow of gesture recognition using the MediaPipe framework.

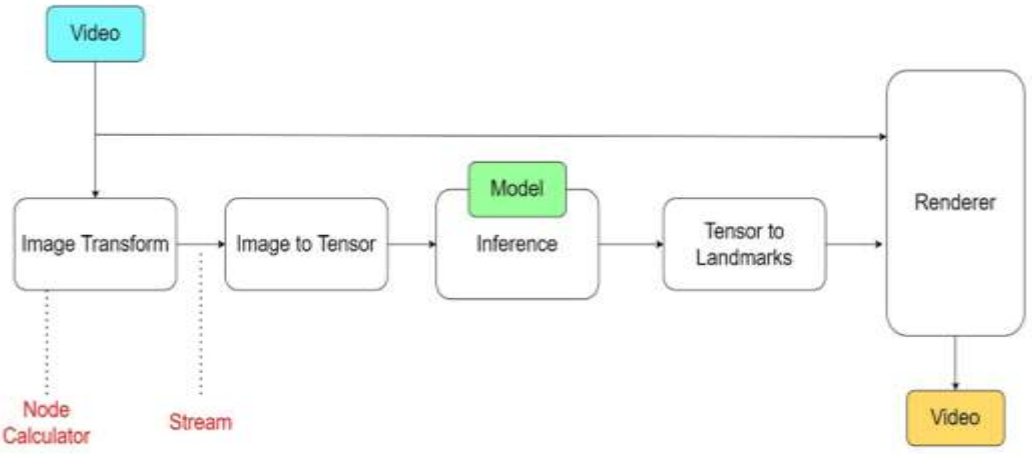

Fig. 2: MediaPipe Framework gesture recognition solution flow chart

Face Mesh Solution Framework: MediaPipe Face Mesh is a solution that estimates 468 3D face landmarks in real-time even on mobile devices. It employs machine learning (ML) to infer the 3D facial surface, requiring only a single camera input without the need for a dedicated depth sensor [9].

In this model, we are using this Face mesh solution framework from the MediaPipe library in python to find out the head movements and the lines of code shown in the below code snippet.

```
mp_face_mesh = mp.solutions.face_mesh
face_mesh = mp_face_mesh.FaceMesh(min_detection_confidence=0.5, min_tracking_confidence=0.5)
cap = cv2.VideoCapture(0)
```

PyAutoGUI: PyAutoGUI is essentially a Python package that works across Windows, macOS X, and Linux which provides the ability to simulate mouse cursor moves and clicks as well as keyboard button presses. Hence, we are using this package to map the head movements to the game controls. The code snippet for the same is shown below.

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

```
74              if y < -10:
75                  text = "Looking Left"
76                  pgi.keyDown("left")
77                  pgi.keyUp("left")
78              elif y > 10:
79                  text = "Looking Right"
80                  pgi.keyDown("right")
81                  pgi.keyUp("right")
82              elif x < -10:
83                  text = "Looking Down"
84                  pgi.keyDown("down")
85                  pgi.keyUp("down")
86              elif x > 10:
87                  text = "Looking Up"
88                  pgi.keyDown("up")
89                  pgi.keyUp("up")
90              else:
91                  text = "Forward"
```

Architectural Design:

Fig. 3 below shows the architectural design of the system. The architecture contains two modules, one for detecting the head movement of the person and another for mapping the head movements to the gaming controls.

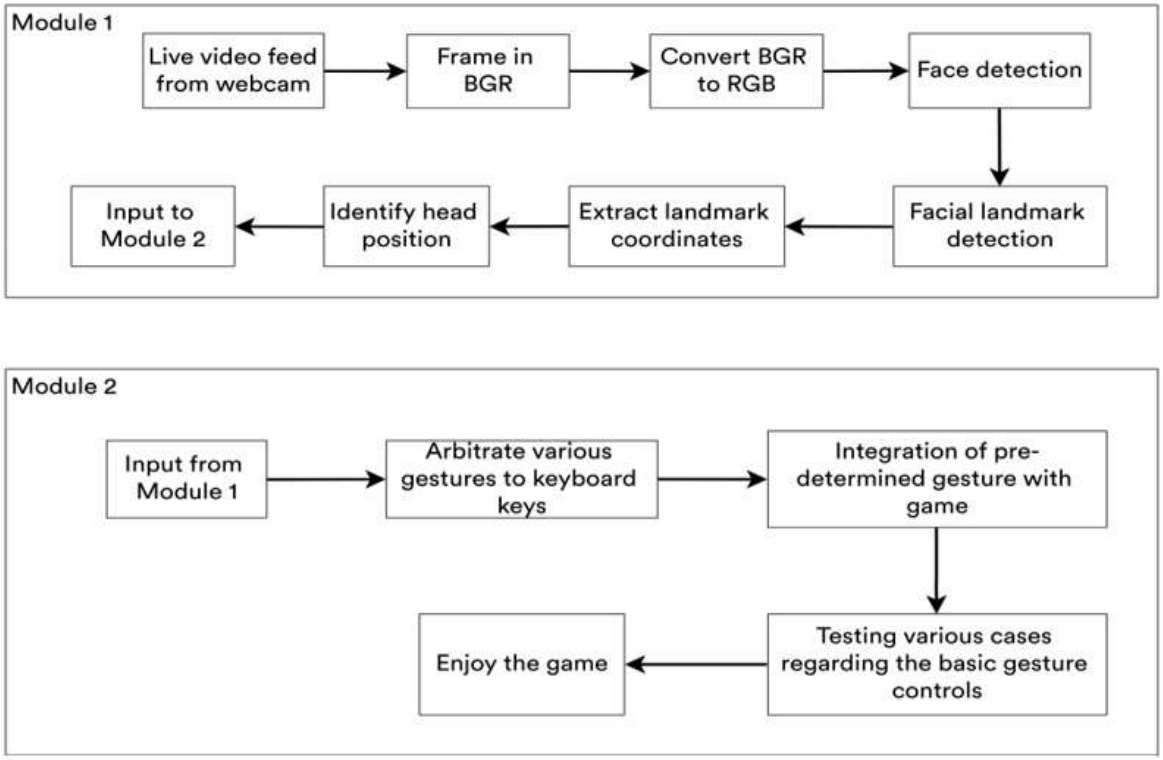

Fig. 3: Architectural Design for the proposed system

Using this system, we can play any game that can be controlled using a computer keyboard. As a demo for the system, we are using one of google play's most installed and played games "Subway Surfers".

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

RESULTS

To build a game that can be played using head movements first we have to find out the head positions. Here are the outputs of head pose estimation. Fig. 4.1, 4.2, 4.3, 4.4, and 4.5 show the head movement direction of the user. Here we are considering Forward-Looking, Up Looking, Down Looking, Left Looking, and Right Looking directions of the head.

Fig. 4.2: Detecting Looking Up position

Fig. 4.3: Detecting Looking Down position

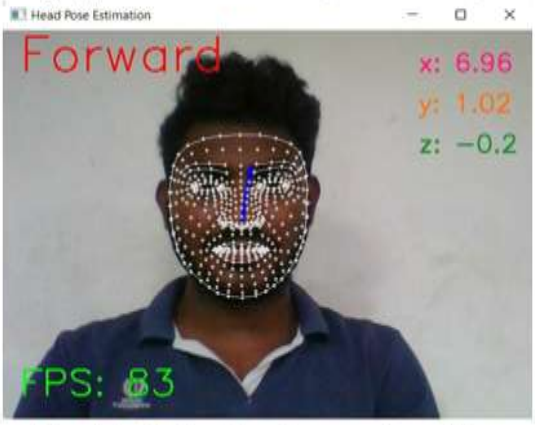

Fig. 4.1: Detecting Forward position

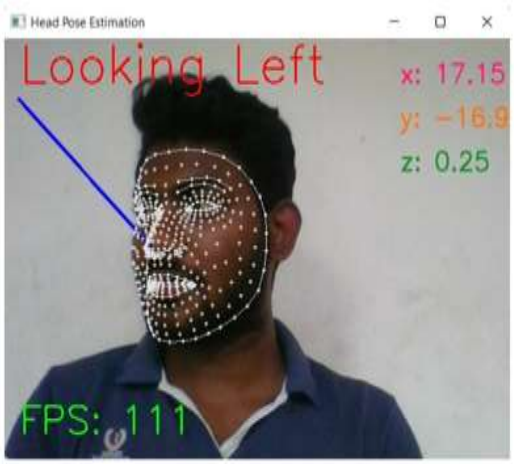

Fig. 4.4: Detecting Looking Left position

Fig. 4.5: Detecting Looking Right position

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Now we have to map the above output directions of the head to the game controls using the PyAutoGUI library in python. Fig. 5.1 and Fig. 5.2 show only two direction controls left and right with the game. The up and down controls are also can be mapped in the same way.

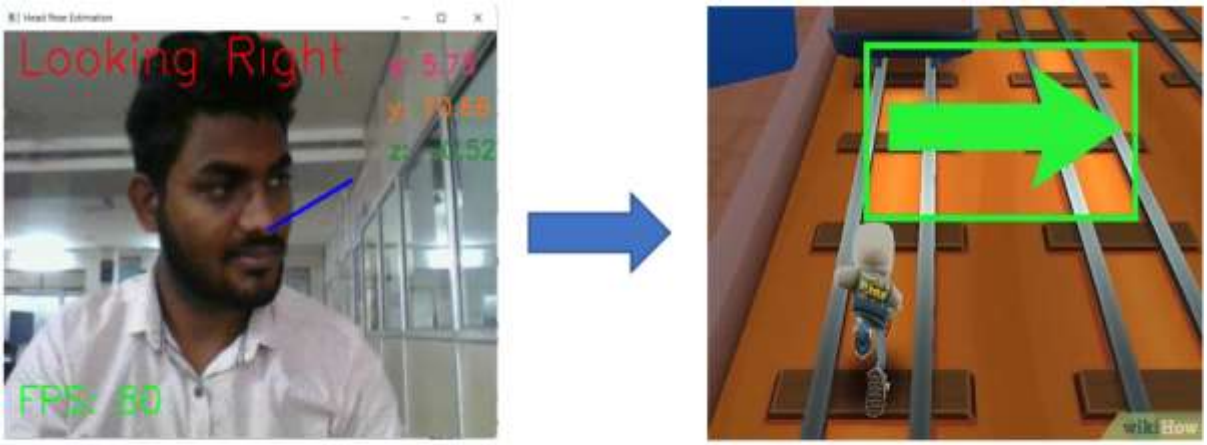

Fig. 5.1: Moving the game character to the right by moving the head to the right

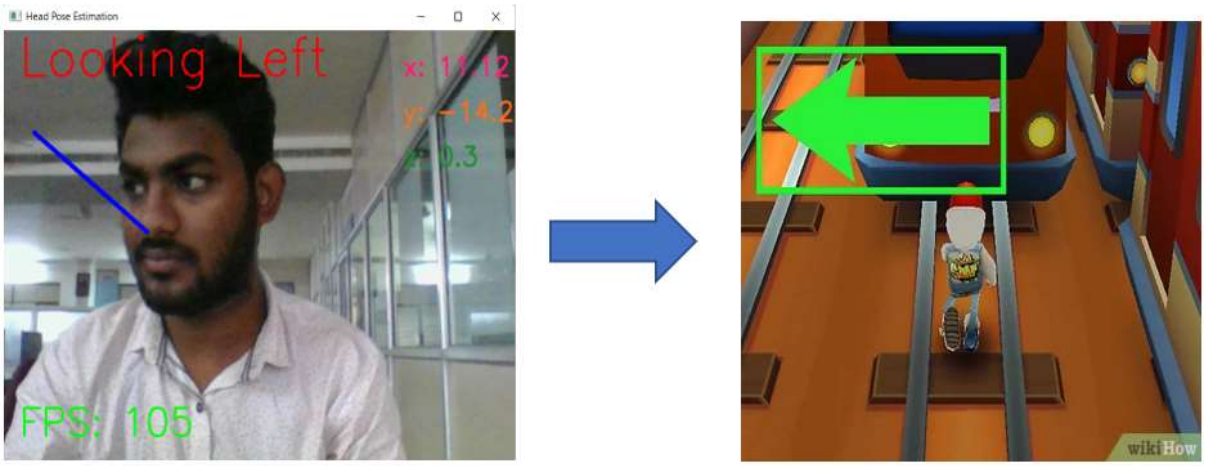

Fig. 5.2: Moving the game character to the left by moving the head to the left

CONCLUSION

Head pose estimation is one of the growing technologies in the field of Computer Science. Developing gesture controlled games are also not a new advancement. This paper proposed a new way of utilizing head gesture recognition technology in the field of entertainment. This is not the end of this head gestures recognition in gaming we can also control mouse actions using our head movements as an extension to this project. This not only reflects in the technology growth but also helps many physically challenged people to enjoy gaming.

REFERENCES

[1] From hands to minds: Gestures promote understanding: Kang and Tversky, Cognitive Research: Principles and Implications (2016) 1:4 DOI 10.1186/s41235-016-0004-9

DOI : 10.36893.JK.2022.V12I01N01.554-561

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

[2] Head pose estimation: A survey of the last ten years Khalil Khan, Rehan Ullah Khan, Riccardo Leonardi, Pierangelo Migliorati, Sergio Benini in Signal Processing: Image Communication Volume 99, November 2021, 116479

[3] Head Pose Estimation with Computer Vision, author: Anastasiya Zharovskikh in InData Labs, 19 January 2021

[4] Boosting of Head Pose Estimation by Knowledge Distillation: Andrey Sheka and Victor Samun, arXiv:2108.09183v2 [cs.CV] 28 Jan 2022

[5] A Marker-free Head Tracker Using Vision-based Head Pose Estimation with Adaptive Kalman Filter: Zhongxu Hu, Member, IEEE, Chen Lv, Senior Member, IEEE, Yanxin Zhou, Yiran Zhang, and Wenhui Huang, arXiv:2103.13006v1 [eess.IV] 24 Mar 2021.

[6] UET-Headpose: A sensor-based top-view head poses: Linh Nguyen Viet, Tuan Nguyen Dinh, Hoang Nguyen Viet, Duc Tran Minh, Long Tran Quoc**,** arXiv:2111.07039v1 [cs.CV] 13 Nov 2021.

[7] Estimation of Human Head Pose using Similarity Measures: Jamie Sherrah, Eng Jon Ong and Shaogang Gong.

[8] Introduction to MediaPipe: Author: Kukil, published an article in LearnOpenCV, 1 March 2022(https://learnopencv.com/introduction-to-mediapipe/).

[9] MediaPipe Fash Mesh solution framework: https://google.github.io/mediapipe/solutions/face_mesh

[10] PyAtuoGUI Module reference documentation: https://pyautogui.readthedocs.io/en/latest/

DOI : 10.36893.JK.2022.V12I01N01.554-561