Automated Media Player System Using Face and Voice Recognition

Rehan Memon[1], Sahil Khorasi[2], Pratik Rahangdale[3], Dr A.K Sampath[4], Prof. Shiburaj Pappu[5]

Student, Dept. Computer Engineering, Rizvi College of Engineering (RCOE), Bandra(W), Mumbai, India[1,2,3]

Professor, Dept. Computer Engineering, Rizvi College of Engineering (RCOE), Bandra(W), Mumbai, India[4]

Professor, Dept. Computer Engineering, Rizvi College of Engineering (RCOE), Bandra(W), Mumbai, India[5]

ABSTRACT: In this project we are developing an application in order to watch the videos in a hassle-free manner. The automated media player is the key approach with the help of face or voice recognition through which it plays and pauses the video. The web camera is used to check whether the user is looking towards the screen or monitor. If yes, then the video continuous to play. In case if the user is not looking towards the screen then the video gets pause. And the voice commands are implemented with the help of an speech recognition API through which the user can play or stop the videos.

KEYWORDS: Face Detection, Voice recognition, Media player, Voice Command

I. INTRODUCTION

The key factors of today's upcoming world are the advance technologies, thus developing an advanced Media player which plays and pauses by determining whether the user is looking towards the screen. There are at times that a user is being called by someone whenever he/she is watching a video or a movie and get distracted by something and then you have to turn away from the screen and you miss some part of that video, due to which the video gets ahead and so you need to find the part till which it has been seen and again play from there. So, the solution to this problem is an automated media player that stops itself whenever the user isn't looking towards the system. As soon as the user looks at the screen again the media player starts running again. This is done with the help of the camera or web camera. If the camera detects the users face looking towards the screen, the media is played. The player stops as soon as users face is not completely seen. This system also provides the feature of voice recognition. As in case whenever the user wants to do something while looking at the screen and is unable to reach the actual keyboard or mouse in order to pause the video then the user can play or pause the video by its voice command.

II. AIM AND GOAL OF PROJECT

The goal of our project is to create an automated media player based on face and voice recognition.
We have set the following objectives for our automated media player toachieve the goal:

1. The interface of the automated media player should be user friendly.
2. The media player should be accurate in terms of face detection and results.
3. The media player should stop the video as soon as the face is not detected.

The media player should play and pause when voice commands are given.

III. CONCEPT AND EVALUATION

3.1 Existing System

I.

Most of the media player used the the public are the normal/standard player with basic functionalities. There are few existing systems which use eye recognition. Due to which results aren't accurate. While some are based on face recognition and voice detection system which are not implemented properly together and not even individually.

3.2 Proposed System:

An automated media player that stops itself whenever the user isn't looking towards the system. As soon as the user looks at the screen again the media player starts running again. This is done with the help of the camera or web camera. If the camera detects the users face looking towards the screen, the media is played. The player stops as soon as users face is not completely seen. This system also provides the feature of voice recognition. As in case whenever the user

wants to do something while looking at the screen and is unable to reach the actual keyboard or mouse in order to pause the video then the user can play or pause the video by its voice command.

IV. ADVANTAGES

1. User Friendly: The Application will be giving a very user-friendly approach for all users.

2. Efficient and reliable: Since CPU usage is low the load taken by the system will be less to some extent. Therefore, proving its efficiency and applicable over all the platform i.e. Operating System and less configuration requirement making it system reliable.

3. Easy maintenance: **"Automated Media Player"**is design as easy way. So, maintenance is also easy.

V. IMPLEMENTATION METHOD

Face Detection

The face detection is implemented with the help of Emgu CV which is a cross platform for image processing library. It is related to OpenCV in such a manner that Emgu CV act as a .NET wrapper to OpenCV. It uses the Haar Cascade Classifier which uses the machine learning approach where the cascade function is trained from the images which consist of both positive and negative in order to detect the objects in other images. It uses the default xml file to detect the face of the user, resize the images and converting it into grey scale. The images are generally of RGB format but when the OpenCV reads these images, they are stored in BGR (Blue, Green, Red) format. The reason for gray scale is that it is easy to process and is efficient as it contains only 1-channel of black white. It then grabs the area where the face was detected and draws a frame around it in green color with the help of the coordinates passed through the function

Voice Recognition

We are using a feature in our Media player which enables the user to issue voice commands with the help of the Google Speech API. It uses passing audio reference through url website through which the application would get the input. The url would ask permission for recording the audio. The API helps to convert the speech to text & this texts serves as an input to determine the system to play or pause the video.
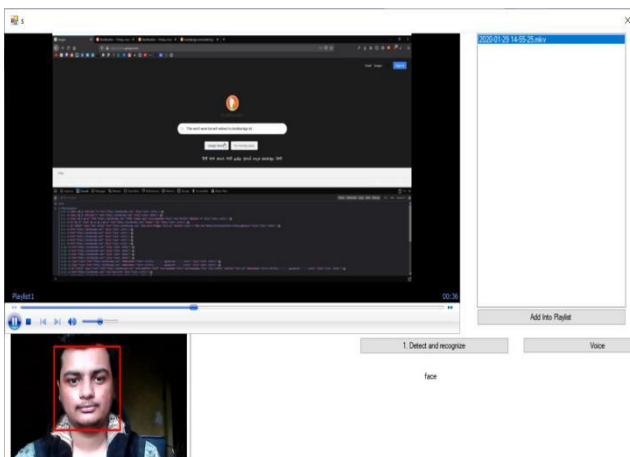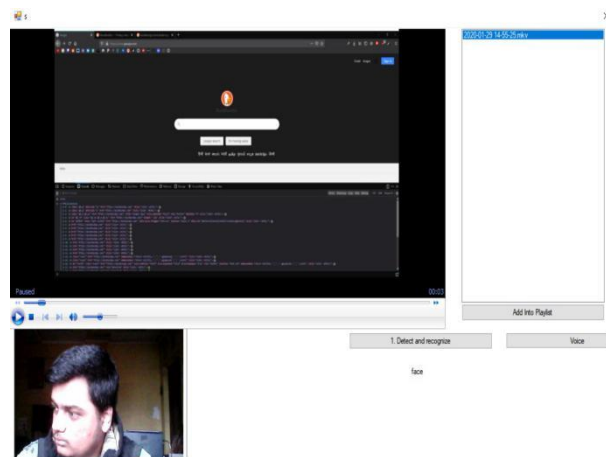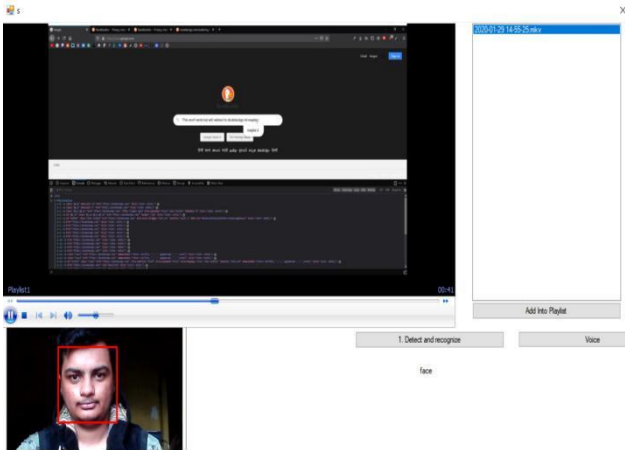
VI. RESULTS
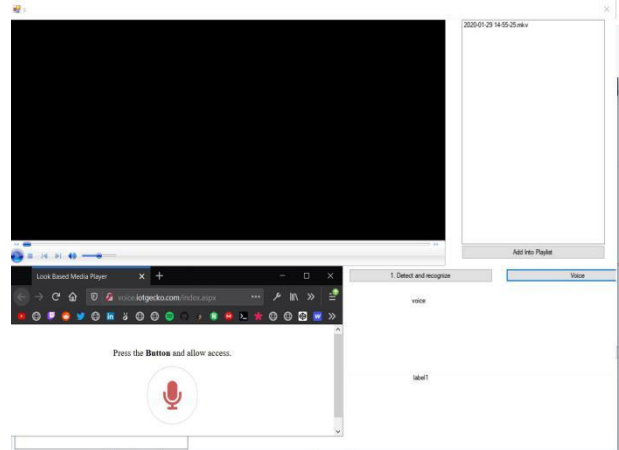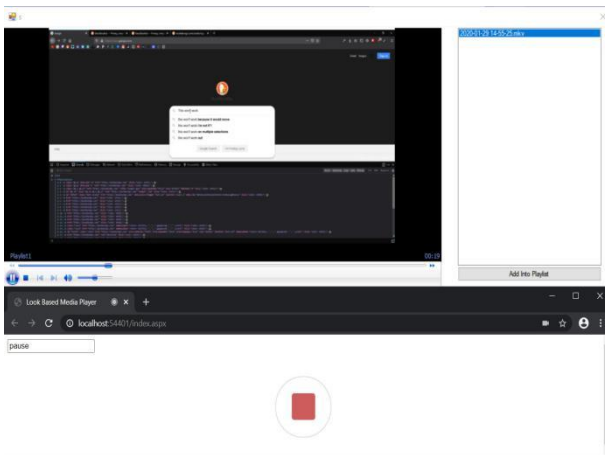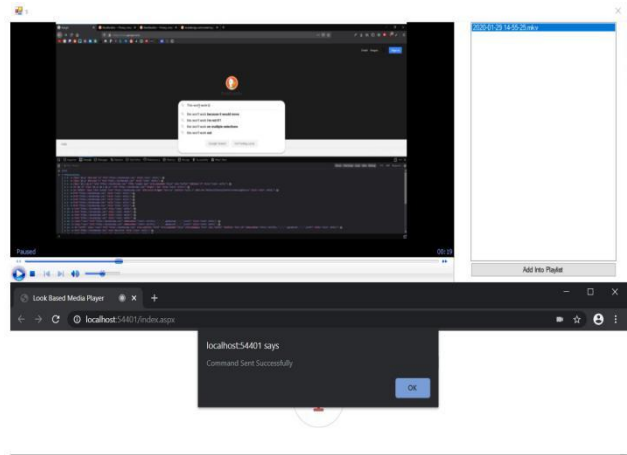

Fig no : 1

Fig no : 2

Fig no : 3

Fig no : 4

Fig no : 5

Fig no : 6

VII. CONCLUSION

The main aim of this project is to get the best user experience through this automated media player. We have tried to achieve this goal by automating the media player in a wide extent. We are doing this with the help of face detection and voice recognition for controlling varied features of the media player such as pausing and starting the video again when the userisn't looking at the screen(for which face detection is used), and controlling functions as play and pause(for which voice recognition is used).

VIII. ACKNOWLEDGMENT

We cannot find the right words today to express our deep gratitude and satisfaction. We are indebted to our inspiring Prof. Dr. A.K Sampath and Prof. Shiburaj Pappu who has extended all valuable Guidance, help and constant encouragement through various difficult stages for the development of project. We express our sincere gratitude to our college for providing us all the facilities.

REFERENCES

[1] Rabiner LR, Sambur MR: An algorithm for determining theendpoints of isolated utterances. Bell Syst. Techn. J 1975,54

[2] Ghosh P, Tsiartas A, Narayanan S: Robust voice activity detection using long-term signal variability. Audio, Speech, and Lang. Proc. IEEE Trans 2011,19 IJCSN - International Journal of Computer Science and Network, Volume 6, Issue 3, June 2017 ISSN (Online) : 2277-5420 www.IJCSN.org Impact Factor: 1.5 394 Copyright (c) 2017 International Journal of Computer Science and Network. All Rights Reserved.

[3] Prasad RV, Sangwan A, Jamadagni HS, Chiranth MC, Sah R, Gaurav V: Comparison of voice activity detection algorithms for VoIP. In Proceedings of the Seventh International Symposium on Computers and Communications (ISCC'02), Washington, DC. Piscataway: IEEE; 2002:530–535.

[4] Davis A, Nordholm S, Togneri R: Statistical voice activity detection using low-variance spectrum estimation and an adaptive threshold. Audio, Speech, Lang. Proc. IEEE Trans 2006,14(2):412– 424.

[5] Madhu N: Note on measures for spectral flatness. Electron. Lett 2009,45(23):1195–1196 Khoa PC: Noise robust voice activity detection. Master's thesis, NangYang Technological University, 2012.

[6] P Voila and M J. Jones, "Rapid object detection using a boosted cascade of simple features [C]", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (ICCVPR 2001), vol. 1, pp. 1-511-1-518, 8-14 Dec. 2001.

[7] Evaluation of Haar Cascade Classifiers Designed for Face Detection.World Academy of Science, Engineering and Technology. International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:6, No:4, 2012.

[8] R. Szeliski. Computer Vision: Algorithms and Applications. Springer 2011.

[9] OpenCV for Computer Vision Applications.Proceedings of National Conference on Big Data and Cloud Computing (NCBDC'15), March 20, 2015.

[10] https://towardsdatascience.com/computer-vision-detecting-objects-using-haar-cascade-classifier-4585472829a9