Turkish Online Journal of Qualitative Inquiry (TOJQI)

Volume 12, Issue 3, July 2021:2523- 2531

Research Article

Smart Media Player on Gesture Recognition using Real Time Image Processing

Avaneesh Kumar Pathak[1], Tanishq Babbar[2], R. Thenmozhi[3*]

Abstract

This paper proposes an advanced media player which is capable of playing and pausing video by identifying if the user is viewing at the display or not. When the user is viewing towards the computer display, then the media player will play the video. On the other hand, if the user is not looking at the screen or looks away from it or the player is unable to identify the user's face, then the video stream will be paused with immediate effect. It can also be used to control different highlights of the media player viz., volume up and down, forward and rewind and all using different hand gestures by real time image processing.

Keywords: *Gesture recognition, media player, face detection, real time ımage processing.*

[1] UG Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India.

[2] UG Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India.

[3*] Assistant Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India, thenmozr@srmist.edu.in

Introduction

Many times, what happens is that when a user is watching a video and receives a call or gets distracted, in that case the user needs to look away from the screen because of which some portion of the video might get missed. After that, user desires to drag back the video stream. As a result, a gesture oriented media player system is proposed which can play and pause video based on the user looking at it or not. The media player uses the web camera to detect the user's face and hand for gesture-based controls. When the user is viewing at the display, then only it will allow the video to be played, in case the user is not viewing at the display screen, then immediately the video stream playback will be paused with immediate effect. Also, the user has the ability to play/pause, volume up/down, forward/rewind video by using hand gestures.

Objectives

The objective of this task is to make a high-level media player system dependent on look and hand gestures. The accompanying targets for the system are to accomplish:

- The media player ought to be precise as far as result.
- The UI of media player should be effective and easy to use.
- The hand signals ought to be caught precisely and activities related to them ought to be performed consummately.
- The media player should stop the video when the client face isn't recognized.

Literature Survey

S. No.	Title of the Paper	Author Names	Results	Advantages	Disadvantages
1	Controlli ng Multimed ia Player with Eye Gaze Using Webcam [1].	Faysal Ahme d, Faysal Ahme d, Shahriar Rahm an, Neephat Benaz ir, AZM Ehtesham Chowdhury,	Can easily control Media player using eye gaze.	It uses image processing for iris detection which is a faster approach for the same	1. Does not work properly in low light. 2. Makes use of VLC media player which is heavy on resources. 3. Uses only eye for

		Md Al Imran			interacting with media player which may not be feasible with low resolution cameras.
2	A Vision Based Hand Gesture Interface for Controlli ng VLC Media Player [2].	Sidharth Rauta ray, Anup am Agrawal.	VLC Media player system can be easily controlled with Hand gestures.	1. Hand segments will be utilized 2. Moving points of hand will be identified.	The complex algorithms viz., k- nearesrt, neighborhood pyramid, lucas kanade optical flow are used
3	Emotion Detection Utilizing Facial Expression [4].	Jyoti Rani, Kanwal Garg.	1. Facial Express ion Recognition System is automated 2. Facial Recognition 3. Recognition of Emotions	1. Several feature extraction methods have been implemented 2. Numerous fields such as science, medicine, & psychology within area of interest	1. Delay in the display of results. 2. Distraction in factors like styles, facial hairs and glass wears
4	Video- Based Face Recognition using Adaptive Hidden Markov Model [3].	Xiao ming Liu, Tsuha n Chen	1. Identification of Face recognition from video sequence. 2. Use of adaptiv e HMM for video based recognit ion	1. Enhanced facial modeling. 2. Improved image based identification	Irrelevant result due to dynamic image recognition
5	Hand Gesture Recognition System to Control Slide Show Navigati on [5].	D. Jadha v, Prof. L.M.R Lobo.	1. The slide show navigation is controlled with static and dynamic gestures. 2.Real time management of of Hand gestures	1. Circular profiling is controlled. 2. Training is not requasired	1. High level Complexity 2. More Time consuming process.
6	Controlling Windows Media player using Hand Recognition System [6]	N. Krishna, Chaithanya, R. Janardhan Roa	Accuracy Rate: Volume up & down: 80% Forward & Backward: 90%	Media player system is controlled using hand gestures	Time consumption is more and Highly Complex

System Model and Evaluatıon

Existing System

In general, the offered frameworks apply eye acknowledgment that results in inaccurate output. Face acknowledgment and hand motions are not carried out as expected together and not even separately.

Proposed System

This task utilizes face recognition and hand movements for regulating the media player system. Face acknowledgment is utilized to stop and play the system. Various hand movements are utilized for regulating the different elements of media player system.

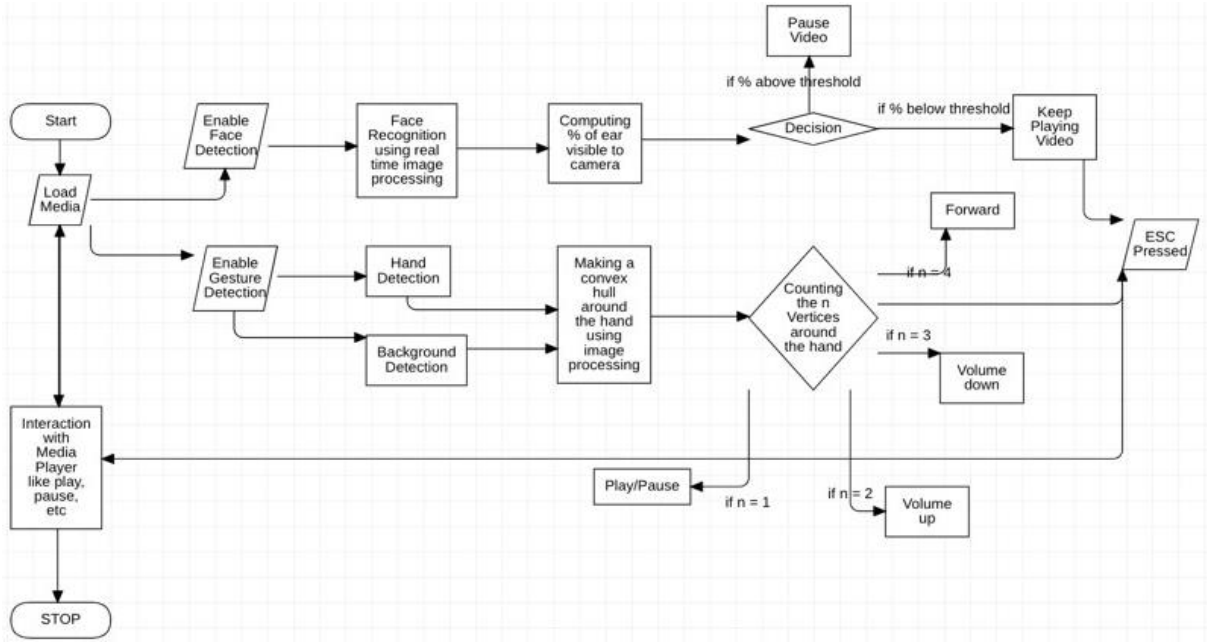


Figure 1. Flowchart of the system

Implementatıon Method

HAAR Cascade Classifiers

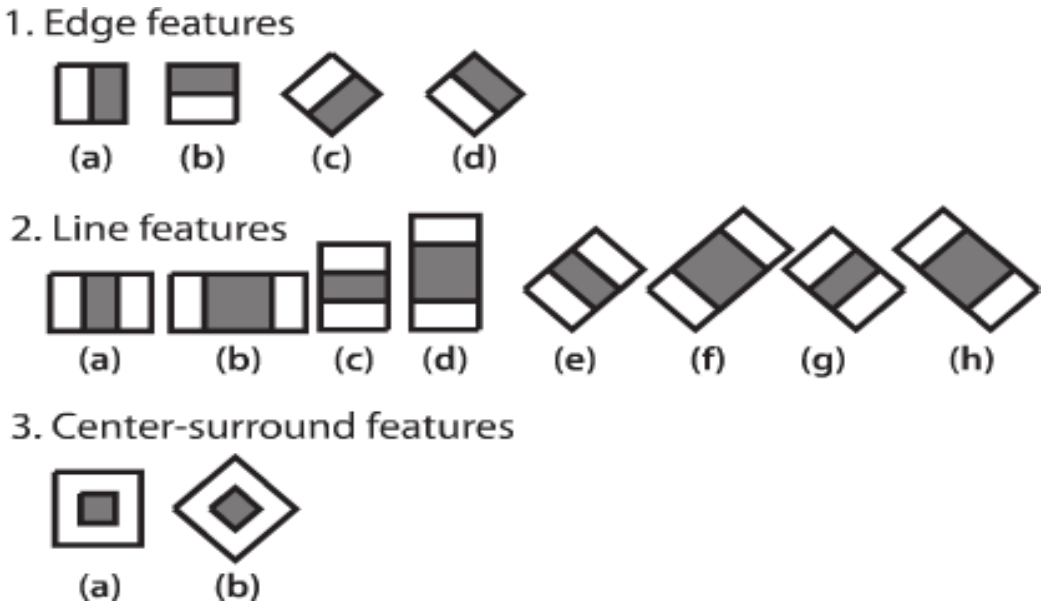
The Haar-like highlights [7] are accordingly coordinated as a classifier cascade to figure out a solid learner or classifier. The vital benefit of a Haar-like element over other methods is its

estimation rate. They are complex picture highlights used in object acknowledgment. These wavelets are natural and applied in initial real-time face detector.

In the suggested discovery structure, a windowpane of the intent size is moved over the key in picture, and for each part of the image; the Haar like element is estimated. This feature is compared with an erudite limit that isolates non-objects from objects. As the Haar-like element is a feeble learner or classifier, numerous Haar-like factors are significant to represent an entity with adequate accuracy.

The Haar- like elements given as input to the basic classifiers:

- Element = x1 × RecSum(y1) + w2 × RecSum(y2)
- The values of the Weights are either positive or negative
- Area and Weights are proportional to one another
- Calculations done at each image point and scale.



Convex Hull Algorithm

A convex polygon is said to be a convex hull when the polygon is able to cover all the points on a given plane.

Consider the input array starts with 0 to n-1points.

1. Each point at ‘y’ coordinate is compared and the point with least value is selected. If the y value of two points is clashing, then select the ‘x’ coordinate with least value. P0 is the point with least value which acquires the first location in the output hull.
2. The remaining n-1 points will be sorted by the order of angle in anti-clockwise direction. When the two points intersect, then the nearest point will be selected first.
3. Then, the two or more points with same angle will be checked to eliminate all same angle points excluding the point with greatest distance from P0. The resultant is the array with size m.
4. When m < 3, then convex hull is impossible.
5. An empty stack ‘S’ is created with first, second and third points S.
6. Take care of the remaining m-3 points one after another. Repeat the following steps for every point[i]

   Remove points from stack

   While the 3 points orientation is not in anti clockwise,

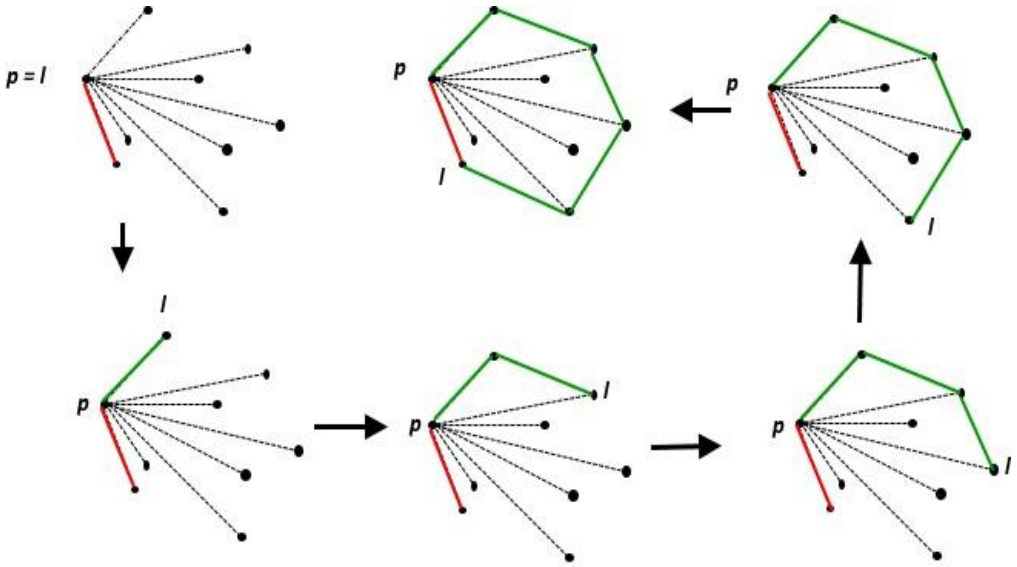
   a. next to top in stack
   b. at the top of stack
   c. points[i]

   Push points[i] -> S
7. Display the data in S

The given algorithm is segregated into two parts.

Part A (Sorting points): First step is to identify the least value point with which all the other points will get sorted. With the sorted points, a simple closed path is formed.

Part 2 (Acceptance or rejection): Then the concave points in the closed path will be removed by traversing through it. Orientation helps to remove the points and the formation of convex hull is done with the first two points in the sorted array. Then with the track of recent three points, the formation of final angle is done. The three points are denoted as pre(p), cur(p) and nxt(p). The Step by step procedure is shown in the following diagram.



The execution of jarvis's March

Advantages of the Proposed System

The Smart Media Player has the following advantages:

- Not a single portion of the video is missed by the user in case he/she looks away from the screen.
- No need for the user to continue hauling back the video from where they missed because the video stream stops whenever the user changes their observation.
- User can likewise advance and reverse the video stream wherever necessary.
- Saves Power and Time.
- Provide precise results.

Conclusion and Results

The primary goal is making user experience much better while using the media player and easy to be controlled. The robotization of the media player is done to a certain extent so as to accomplish the required objective. Utilization of face acknowledgement and hand gesture recognition for this purpose to control the highlights of the media player, for example, once the user is not viewing towards the display, video stream is stopped for that time frame until the user again looks at the screen (using face acknowledgement), and also controlling features viz., volume up or down, forward and backward (with the help of hand gestures).

The vital benefit of Haar-like component over other highlights is its computation speed. It utilises the necessary pictures so that the Haar-like components can be determined in steady time i.e., around 60 microprocessor commands for a 2-rectangle feature.

Let the number of input points be depicted by n. If a sorting algorithm which takes O(nlonn) time, then the Graham Scan algorithm takes O(nlogn) time. Finding the bottom most point i.e. the first step, takes o(n) time. Sorting points i.e., sencond step takesO(nlogn) time. At most one time, every element is pushed and popped i.e., third step takes O(n) time. Processing points one by one i.e., sixth step takes O(n) time, assume stack processes take constant time. So, the overall complexity is O(n) + O(nlogn) + O(n) + O(n) i.e., O(nlogn).

References

Ahmed, F., Rayhan, M.S.S., Rahman, S., Benazir, N., Chowdhury, A.E., & Al Imran, M. (2019). Controlling Multimedia Player With Eye Gaze Using Webcam. *In 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, 152-156. https://doi.org/10.1109/ICREST.2019.8644103

Rautaray, S.S., & Agrawal, A. (2010). A vision based hand gesture interface for controlling VLC media player. *International Journal of Computer Applications, 10*(7), 11-16.

Liu, X., & Cheng, T. (2003). Video-based face recognition using adaptive hidden markov models. *In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings, 1*(I-I).

Rani, J., & Garg, K. (2014). Emotion detection using facial expressions-A review. *International Journal of Advanced Research in Computer Science and Software Engineering, 4*(4), 465-467.

Jadhav, D., & Lobo, L.M.R.J. (2014). Hand gesture recognition system to control slide show navigation. *International Journal of Application or Innovation in Engineering & Management (IJAIEM), 3*(1), 283-286.

Chaitanya, N.K., & Rao, R.J. (2014). Controlling of windows media player using hand recognition system. *The International Journal of Engineering and Science (IJES), 3*(12), 01-04.

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR, 1*(I-I).

Maheswari, K.U., & Govindarajan, S. (2019). Hybridisation of oppositional centre-based genetic algorithms for resource allocation in cloud. *International Journal of Networking and Virtual Organisations, 21*(3), 307-325.

Umamaheswari, K.M., Chanana, T., & Fernandes, K. (2020). Meme Chat: An Innovatıve Socıal Medıa Platform for Content Monetızatıon & Corporate Outreach Usıng OCR & NLP. *Journal of Critical Reviews, 7*(4), 206-208.

Maheswari, K.U., Roy, S., & Govindarajan, S. (2017). Hybrıd Green Schedulıng Algorıthm Usıng Genetıc Algorıthm and Partıcle Swarm Optımızatıon Algorıthm ın IAAS Cloud. *ARPN Journal of Engineering and Applied Sciences*, 12(12), 3762-3766.