Interaction to computer system based on eye and head gesture recognition using Deep Learning

[1]Praphull Maurya, [2]Narendra Rajpoot, [3]Rishabh Shukla, [4]Atul Kr. Uttam, [5]Dhananjay Bisen, [6]Siddhartha Kr. Arjaria

[1]praphullmaurya123@gmail.com, [2]nrajpoot1146@gmail.com, [3]rishabhshukla475@gmail.com, [4]atuluttam150998@gmail.com, [5]bisen.it2007@gmail.com, [6]arjarias@gmail.com

Abstract

This paper is centred around building up a model to control computer system by utilizing eyes and head gestures. The face recognition system is used in combination with eyes and head gesture recognition system to control the movement of mouse and therefore, the computer system operates in real time. The facial recognition model is followed by a Convolution Neural Network (CNN) based element extractor and Support Vector Machine (SVM) to perceive the face. To capture the blink of eyes, we are using the python library named D-Lib. The reason for this task is to develop an eyes and head motion-controlled system that can successfully utilize face parts like eyes and nose and empower the user to perform activities mapped to explicit eye blink/head motions by utilizing PC webcam. It must be consistently accurate so that user can comfortably use it like other frequent gadgets.

Keywords

Human-Computer Interaction, Gesture Recognition, Support Vector Machine, Facial Recognition, Convolutional Neural Network, D-Lib

1. Introduction

The performance of Human-Computer Interaction (HCI) has impacted the interaction with PC over the previous years and with the huge rise of computer system among general public, it has become essential to identify a medium through which human and computer system can interact with one another, as PCs have become a significant part in our day to day lives. Among different techniques, such as hand gestures are very popularly used, but the people who do not have a hand cannot interact with the computer system. The head gesture has the potential to communicate with an intelligent system. This paper shows a vision-based multimodal HCI utilizing eye blink followed by head movement that proposes an interface which can viably substitute the job of a current gadget, for example, a mouse.

1.1 HCI - Human-Computer Interaction

The research in the field of Human-computer interaction (HCI) has improved the design and computer technologies that are used to develop the interfaces between human being and computers. Scientists in the field of HCI observe the human behaviour as how they interact with computer system and develop technologies that can remove the complexities and enables the human to interact with computer system in new ways. There are many ways in which humans interact with computers and interface is the crucial part of that interaction. Graphical User Interface (GUI) are used by Desktop applications, handheld computers, internet browsers, ERP and computer kiosks.

Human-computer interaction considers the manners by which people utilize or don't utilize computational resources, frameworks, and foundations. Improvement in the usability of computer interfaces is an area of research.

1.2 Gesture Recognition

Gesture recognition is a computer science topic dealing with the interpretation of human gestures via mathematical algorithms. Any motion of the body can be considered as gestures, prominent sources of gestures being the face or hand.

Gesture recognition features: More accurate**,** Higher stability**,** Time-saving to unlock a device

Gesture recognition can be carried out with advance technologies like computer vision and image processing. In relation to gesture control, the touchless UI is a emerging kind of innovation. Touchless UI (TUI) is the way toward instructing the PC through body movement and motions without contacting a mouse, a keyboard, or screen.

1.3 CNN - Convolutional Neural Network

A Convolutional Neural Network (CNN/ConvNet) is a deep learning algorithm in the class of deep neural networks, most frequently used to analyse visual imagery. CNN is a classification algorithm that can take images as input and classify images, group images by similarities and accomplish object recognition within scenes. For example, convolutional neural networks (ConvNets or CNNs) are used to identify faces, street signs, individuals, tumours, platypuses and many other characteristics of visual data. CNNs also have applications in recommender systems, medical image analysis, and natural language processing.

1.4 SVM - Support Vector Machine

A Support Vector Machine (SVM, a supervised learning model) is a discriminative classifier formally described through a separating hyperplane. In other words, given labelled training data (supervised learning), the algorithm yields an ideal hyperplane which categorizes new examples. In two-dimensional space this hyperplane is a line isolating a plane into two sections wherein each class lay on either side.

2. Literature Survey

The survey is concentrated to address the points, focus on the research questions, understanding of the research area, clarification of the meaning of the terms and proper identification of the framework and planning of the data collection approach. The important thing is to understand the domain of research in which head movement and eye detection is involved. During the literature survey, the attention is on the most proficient method to build up a model which can satisfy the necessities of physically disabled people and the model should be easy to understand.

Author Name	Proposed Technique	Key findings	Final Results		
Anush Ananthak umar (2018) [1]	Efficient Gesture and Face Recognition for Time Bound Application	**1**-Face extraction is done using the Viola-Jones facial recognition system that uses Haar like attribute along with Adaboost training. **2**-Color segmentation is used for gesture recognition, that extract the tone of skin of detected face and also detect the position of hand.	To check the accuracy of the system with different conditions and image quality, the proposed face recognition model has been tested on many databases and the results yielded on few of which is as shown below:		
			Dataset name	Training time	Accuracy

		3-Kanade-Lucas-Tomasi (KLT) algorithm is used to obtain the gesture by traking the hand. 4-This research uses the background subtraction model reduce the misclassification and extract the foreground. This type of technique highly improves the facial and gesture recognition performance in cluttered and surroundings.	ORL database of faces	150.2 sec	98.61%
			Caltech Faces 1999 database	220.5 sec	97.15%
			Faces96 dataset	376.34 sec	97.01%
			So, the result shows that with less training time the proposed method can achieve high accuracy.		
S. M. Ariful Hoque MD. Sadun Haq Md. Hasanuzza man[2]	CV Based Gesture Recognition for Object Manipulation of System. **Technique-** This research proposes a method to recognize static hand signals. Perceived motions at that point can be utilized to control work area objects.	1-Picture procurement: Image is obtained by Kinect 2.0. Kinect SDK render the obtained picture up to 30 fps. 2-In this training, 9 unique gestures are trained. For preparing, 5 diverse individuals are utilized. This framework is assessed on an exceptional dataset from contribution of these 5 people. This framework incorporates 9 predefined motions. In the wake of running the framework, first the COP is recognized utilizing fingertip. At that point the framework holds for 2 seconds for the motion to begin. after motion begins, the framework catches the data of 30 back to back casings.	The fundamental objective of paper is to build up a robust and intuitive framework utilizing gestures. These motions are then used to control and control distinctive desktop objects. This framework utilizes contour detection algorithm and graham scan algorithm to detect fingertip.9 unique gestures are be utilized to control motions. A lot more can be executed as feature vector that can recognize each of the 5 fingertips with unique excellence, yet for effortlessness just 9 has been picked.		
Mohd. Baqir Khan, Kavya Mishra, Mohamme d Abdul Qadeer (2017) [3]	Gesture Recognition using Open-CV **Technique:** To perform a specific task first an image is captured by the camera then applied to preprocessing to minimize noise and convert image RGB to grayscale. After that feature	According to this paper, we can map some specific operations or tasks with a particular gesture. For example, if the Hand is detected then main window will open.	By using GRS this idea developed an efficient human interaction with machine. this will have many applications like Remote control, Aid to physically challenged, Mouse control. GRS is implemented by using glove approach and colored marker approach.		

|			
	extraction performs for finding the shape of the palm. Hidden Marcov Model, and k-curvature are used for this.	4	
Saransh Sharma, Samyak Jain, Khushboo (2019) [4]	A Static Hand Gesture and Face Recognition System for Blind People **Technique-** This paper introduces a technique, which can be useful for a visually impaired individual. Face and Hand gesture recognition system has been executed by in this experiment using various algorithms and different task are done using this.	**1-**Contour Detection which is one of the best edge detections has been used to extract the features. **2-**For the real time Hand gesture detection Convex defects and Convex hull algorithm has been used. . **3-**Haar cascade classifiers is used for face detection and LBPH recognizers has been used for identification in the real time. **4-**In this paper YCbCr color space. is used for Hand gesture with Skin detection.	The model presented right now be useful for a visually impaired individual and can go about as a virtual helper for it. Face recognition and identification system are done with an accuracy of 92% & Hand gesture is perceived with an exactness of 95.2%. .

3. Proposed System
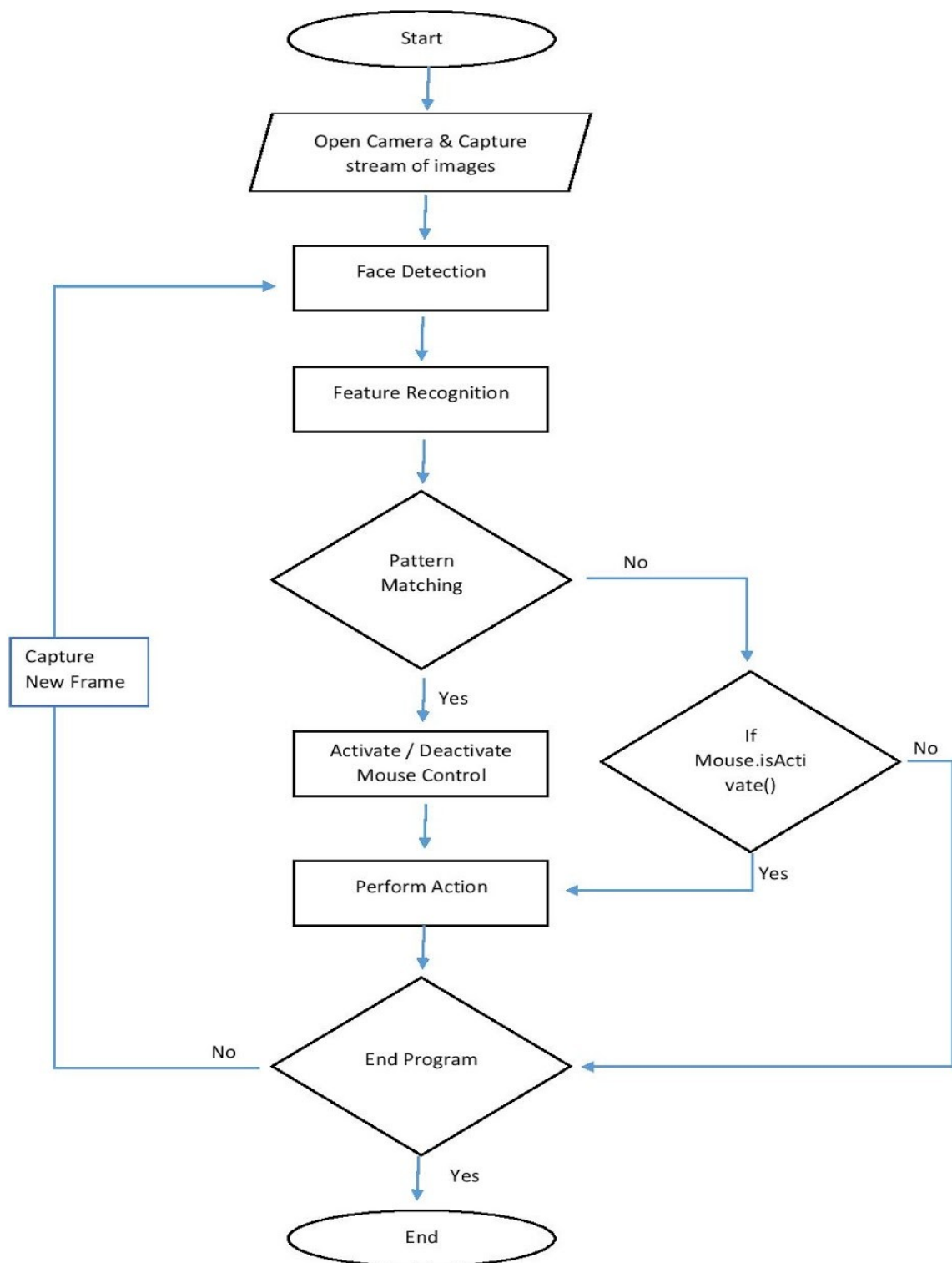
Figure: Flow Chart

(Above Flow Chart shows the sequence of routines that are running inside the system)

Above Flow Chart has following steps-:

1. Start program
2. Turning on the webcam and webcam continuously captures frames of image.
3. The system performs an action to detect face and recognize its features like eyes, nose.
4. The system finds the pattern and match the pattern.
5. If pattern has been matched, then system take an action of activation or deactivation of mouse.
6. Perform Action:

   a) Mouse control by moving the head.
   b) Scroll vertically, scroll horizontally.
   c) Right click, left click and double click on eye blink.

4. Conclusion

In this paper, we proposed a model that interact to computer system based on eye and head gesture recognition using Deep Learning. The experiment is based on a model with recognition accuracy greater than the real-time models. Proposed model is able to recognize blink of eyes and gestures of the head. We are using eye aspect ratio in order to achieve the maximum accuracy and will perform the left, right and double click of mouse operations on respective left, right and both blink of eyes. The principle objective of this paper is to build up a powerful and intuitive model utilizing eye flicker and head gesture. These motions are then used to control the system. This system uses CNN algorithm for feature extraction and SVM algorithm for face recognition.

5. References

[1] Anush Ananthakuma, Georgia Institute of Technology Atlanta, GA, 30332, "Efficient Face And Gesture Recognition For Time-Sensitive Application ", in IEEE 2018.

[2] S. M. Ariful Hoque, MD. Sadun Haq, Md. Hasanuzzaman , " Computer Vision Based Gesture Recognition for Desktop Object Manipulation". International Conference on Innovation in Engineering and Technology (ICIET) 27-29 December, 2018

[3] Mohd. Baqir Khan, Kavya Mishra, Mohammed Abdul Qadeer, " Gesture Recognition using Open-CV" 7th International Conference on Communication Systems and Network Technologies 2017

[4] Saransh Sharma, Samyak Jain, Khushboo." A Static Hand Gesture and Face Recognition System For Blind People". 6th International Conference on Signal Processing and Integrated Networks (SPIN) 2019