

Motion and Gesture Controlling

Jayvir Kher, Sunny Rajpurohit, Himanshu Patidar, Gourav Patidar, Kaushal Pathak

Computer Science and Engineering, Parul Institute of Technology, Vadodara, Gujarat, India

ABSTRACT

Hand gesture is an innovative technique of interaction between human and computer. Hand gesture recognition is a technology that is becoming increasingly relevant, given the recent growth and popularity of Virtual and Augmented Reality technologies. In this paper, we represent a novel approach for better human computer interaction (HCI) where we are using hand gesture recognition using real-time camera in order to overcome the drawbacks of existing methods and more user friendly. Our method is to use a camera and computer vision technology, such as image segmentation and gesture recognition, to control keyboard tasks and Hand motions are acquired using web camera based on color detection.

Keywords : Hand Gestures, Colour Detection, Human Computer Interaction (HCI), Web camera.

I. INTRODUCTION

This paper aims to explore the existing options for hand gesture recognition in a common context. Most people nowadays own a laptop with a front-facing camera. If we could tap into this, we could possibly bring a more natural method of interaction to the masses. Moreover, as virtual reality devices become more common, the laptop camera may also become a viable complementary interaction device, capturing a field of view separate from the virtual reality device.

In this gesture recognition system, it is very possible to capture & track the fingertip of hand with a webcam or built-in cam which is bearing a color cap or color sticky note paper and the system track the color and movement of the hand & move cursor with it and for keyboard shortcut control, static hand gestures will be used [1].

This system is implemented in Python programming language using the Computer Vision based library OpenCV. This system has the potential to replace the typical mouse and also the remote controller of

machines. The only barrier is the lighting condition. That's why the system still can't be enough to replace the traditional mouse as most of the computers are used in poor lighting conditions.

1.1 Camera

The system works on the frames captured by the webcam on the computer machine or built-in camera on a laptop. By creating the video capture object, the system will capture video using webcam in real-time. In order to use a single camera with this system the device index would be "0". To add additional camera device index would be 1, 2 and so on. This camera will capture frame by frame and will pass it to the system.

1.2 Capturing

By using an infinite loop, the webcam captures each and every frame till the program termination. The frames from the real-time video are processed from BGR to HSV color space. In this gesture recognition system, it is very possible to capture & track the fingertip of hand with a webcam or built-in cam which is bearing a color cap or color sticky note paper and the system track the color and movement of the

hand & move cursor with it and for keyboard shortcut control , static hand gestures will be used

1.3 Data Notation

Hand Gesture	Label used
Fist of hand	0
1 finger	1
2 fingers	2
3 fingers	3
4 fingers	4
5 fingers	5

II. METHODS AND MATERIAL

Gesture Recognition

Humans are able to recognize body and sign language easily. This is possible due to the combination of vision and synaptic interactions that were formed along brain development [2]. In order to replicate this skill in computers, some problems need to be solved: how to separate objects of interest in images and which image capture technology and classification technique are more appropriate, among others.

Gesture recognition methodologies are usually divided into two categories: static or dynamic [3]. Static gestures are those that only require the processing of a single image at the input of the classifier, the advantage of this approach is the lower computational

cost. This paper aims to explore static gesture recognition.

2.1 Convolutional neural networks

In this work, we used two image bases of 5 gestures, some segmentation techniques and the use of convolutional neural networks (CNNs) for classification. Thus, with the proposed methodology, we demonstrated that with simple architectures of convolutional neural networks, it is possible to achieve excellent results for static gesture classification.

Convolutional neural networks, or CNNs, are widely used for image classification, object recognition, and detection [4]. Three types of layers can summarize its structure: convolution, pooling, and classification, as shown in Figure 1. The CNN architecture must be defined according to the application and is usually defined by the number of alternate convolution and pooling players, number of neurons in each layer, and choice of activation function.

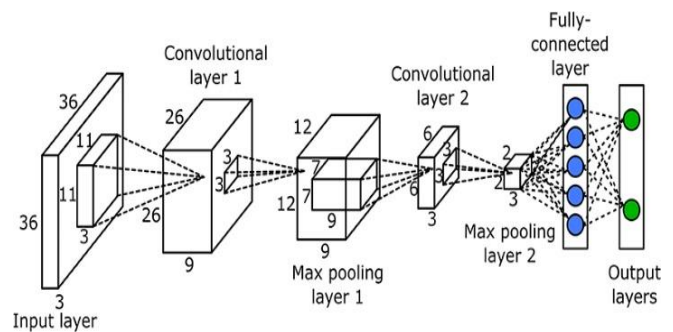


Figure 1. CNN

In the context of image classification, the input for a CNN is an image represented by an arbitrary colour model. At the convolution layer, every neuron is associated with a kernel window that is convolved with the input image during CNN training and classification. This convolution kernel is composed of the weights of each associated neuron. The output of this convolution step is a set of N images, one for each of the N neurons. Because of convolution, these new images can contain negative values. In order to avoid this issue, a rectified linear unit (ReLU) is used to

replace negative values by zero. The outputs of this layer are called feature maps.

After a convolution layer, it is common to apply a pooling layer. This is important because pooling reduces the dimensionality of feature maps, which subsequently reduces the network training time. Segmentation subdivides an image into regions, so that it is possible to highlight regions that contain characteristics of interest [5]. Therefore, segmentation algorithms can be implemented to separate colors, textures, points, lines, discontinuities, borders, among others. The segmentation process varies according to the problem.

In the case of gesture recognition, the entire background region of the image is not of interest, so only the set of pixels with the presence of the human hand must be maintained Figure 2. One method for this segmentation is the implementation of background removal, where image samples are collected from the environment and then objects are added to the scene. In this way, the pixels of the new images are compared with the images of the scenario.



Figure 2. Threshold Image

In this gesture recognition system, it is very possible to capture & track the fingertip of hand with a webcam or built-in cam which is bearing a color cap or color sticky note paper and the system track the color and movement of the hand & move cursor with

2.2 Methodology

The proposed methodology can be visualized in the flow chart of Figure 3. The images are obtained from the database. Then, the images are used to train a CNN and assess the performance of the technique with cross validation. The images go through an image processing stage, in which the following operations contour generation, and polygonal approximation, to remove image noise.

After segmentation, binary images are obtained, so a logical AND operation is performed between these images and the originals, in order to preserve the information contained in the fingers and the surface of the hand then, the validation results are analyzed. After these steps Finally gesture recognized and operation performs accordingly.

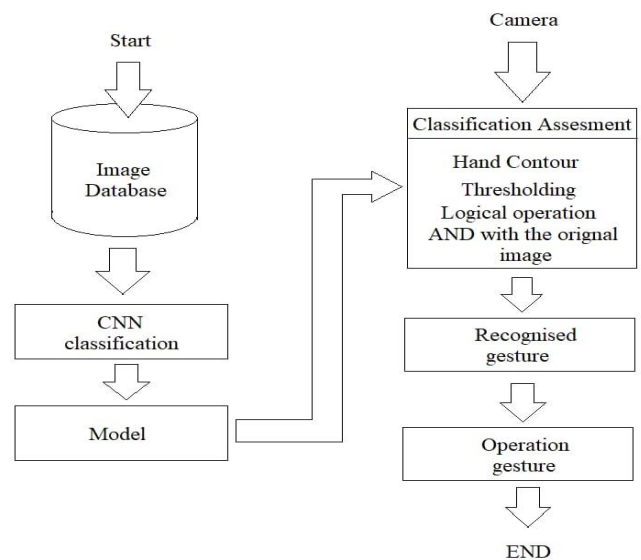


Figure 3. Steps of Gesture Recognition

III. RESULTS AND DISCUSSION

Virtual mouse

A virtual mouse is software that allows users to give mouse inputs to a system without using an actual mouse.

To the extreme it can also be called as hardware because it uses an ordinary web camera. A virtual mouse can usually be operated with multiple input devices, which may include an actual mouse or a computer keyboard. Virtual mouse which uses web camera works with the help of different image processing techniques.

In our work, we have tried to control mouse cursor movement and click events using a camera based on color detection technique. Here real time video has been captured using a Web-Camera. The user wears colored tapes to provide information to the system. Individual frames of the video are separately processed. The processing techniques involve an image subtraction algorithm to detect colors. Once the colors are detected the system performs various operations to track the cursor and performs control actions [6].

3.1 Methodology

Following are the steps in working of our project :

1. Capturing real time video using Web- Camera.
2. Processing the individual image frame.
3. Flipping of each image frame.
4. Conversion of each frame to a grey scale i.
5. Color detection and extraction of the different colors (RGB) from flipped gray scale image.
6. Conversion of the detected image into a binary image.
7. Finding the region of the image and calculating its centroid.
8. Tracking the mouse pointer using the coordinates obtained from the centroid.
9. Simulating the left click and the right click events of the mouse by assigning different color pointers.

A. Capturing the Real Time Video

For the system to work we need a sensor to detect the hand movements of the user. The webcam of the computer is used as a sensor. The webcam captures the real time video at a fixed frame rate and resolution which is determined by the hardware of the camera.

The frame rate and resolution can be changed in the system if required.

Computer Webcam is used to capture the Real Time Video. Video is divided into Image frames based on the FPS (Frames per second) of the camera.

B. Color Detection

This is the most important step in the whole process. The red, green and blue color object is detected by subtracting the flipped color suppressed channel from the flipped Gray-Scale Image. This creates an image which contains the detected object as a patch of grey surrounded by black space.

C. Conversion of gray scale Image into Binary scale Image

The grey region of the image obtained after subtraction needs to be converted to a binary image for finding the region of the detected object. A grayscale image consists of a matrix containing the values of each pixel. The pixel values lay between the ranges 0 to 255 where 0 represents pure black and 255 represents pure white color.

D. Finding Centroid of an object and plotting:

For the user to control the mouse pointer it is necessary to determine a point whose coordinates can be sent to the cursor. With these coordinates, the system can control the cursor movement. An inbuilt function in MATLAB is used to find the centroid of the detected region. The output of function is a matrix consisting of the X (horizontal) and Y (vertical) coordinates of the centroid. These coordinates change with time as the object moves across the screen.

IV. CONCLUSION

We simplified the process of human interaction with the computers by the help of hand motion and gesture. From the report we can control, system mouse by colour detection and can control keyboard shortcuts by static hand gestures.

CNN model accuracy

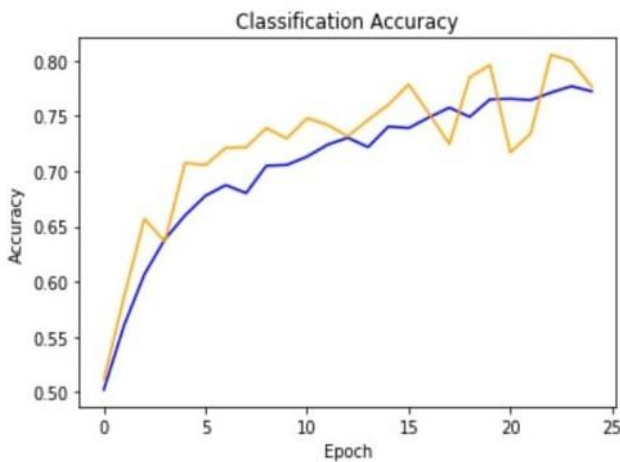


Figure 4. Accuracy of model

CNN model cross entropy loss:

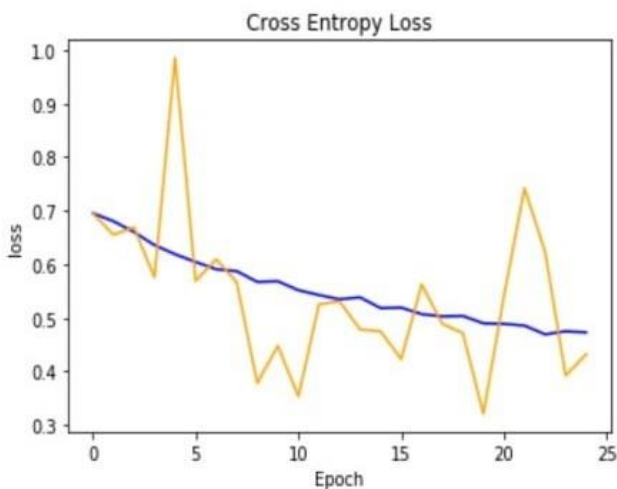


Figure 5. Cross entropy loss

V. REFERENCES

- [1]. Zi Xian Justin “Hand Recognition and Gesture Control using a laptop web-camera ” Yeo Exchange Student with Stanford University
- [2]. M. W. Cohen, N. B. Zikri, and A. Velkovich, “Recognition of continuous sign language alphabet using leap motion controller,” in Proceedings of the 2018 11th International Conference on Human System Interaction (HSI), pp. 193–199, Gdańsk, Poland, July 2018.
- [3]. P. K. Pisharady and M. Saerbeck, “Recent methods and databases in vision-based hand gesture recognition: a review,” Computer Vision and Image Understanding, vol. 141, no. 4356, pp. 152–165, 2015.
- [4]. A. C. G. Vargas, A. M. P. Carvalho, and C. N. Vasconcelos, “Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres,” in Proceedings of the SIBGRAPI—Conference on Graphics, Patterns and Images, Sao Paulo, Brazil, October 2016.
- [5]. R. C. Gonzalez and R. E. Woods, Digital Image Processing, Prentice-Hall, Upper Saddle River, NJ, USA, 3rd edition, 2006.
- [6]. Ranjeet Saroj, Avnish Kumar, Abhilash Moolya, Deepika Pandit, “VIRTUAL MOUSE USING COLOR DETECTION”, VIRTUAL MOUSE USING COLOR DETECTION, Volume: 05, Issue: 04, Apr-2018.
- [7]. Suresh D S, Bhavana I V, “Virtual Mouse Implementation using Color Pointer Detection”, International Journal of Research Studies in Science, Engineering and Technology Volume 1, Issue 5, August 2014, PP 23-32.

Cite this article as :

Jayvir Kher, Sunny Rajpurohit, Himanshu Patidar, Gourav Patidar, Kaushal Pathak, "Motion and Gesture Controlling", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 2, pp. 337-341, March-April 2020. Available at doi : <https://doi.org/10.32628/CSEIT206298>
Journal URL : <http://ijsrcseit.com/CSEIT206298>