International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 2, January-February 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Hand Gesture Recognition using OpenCV and Python

Surya Narayan Sharma, Dr. A Rengarajan

Department of Master of Computer Applications, Jain Deemed to be University, Bengaluru, Karnataka, India

ABSTRACT

Hand gesture recognition system has developed excessively in the recent years, reason being its ability to cooperate with machine successfully. Gestures are considered as the most natural way for communication among human and PCs in virtual framework. We often use hand gestures to convey something as it is non-verbal communication which is free of expression. In our system, we used background subtraction to extract hand region. In this application, our PC's camera records a live video, from which a preview is taken with the assistance of its functionalities or activities.

KEYWORDS: *Gesture recognition, OpenCV, human-computer interaction, python, machine learning*

How to cite this paper: Surya Narayan Sharma | Dr. A Rengarajan "Hand Gesture Recognition using OpenCV and Python" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-2, February 2021, pp.346-352, URL: www.ijtsrd.com/papers/ijtsrd38413.pdf

IJTSRD38413

INTRODUCTION

Hand gestures are unprompted and also robust transmission mode for Human Computer Interaction (HCI). Keyboard, mouse, joystick or touch screen are some input device for connection with the computer but they don't provide appropriate interface whereas, the current system will contain either desktop or laptop interface in which hand gesture can be done by wearing data gloves or web camera used for snapping hand image. The first step towards this gesture recognition is hand capturing and analyzing. Sensors are used in Data-Glove methods for initializing fingers movement and other sensor will program hand movements. In comparison the vision based method only needs a camera and hence identifying the actual interaction between human and computer without using any other devices. The challenges of this system are constant background, sometimes person and lighting also. Different procedure and algorithms which are used in this system are elaborated here along with the recognition techniques. The method of searching a connecting region in the picture with particular specification, being it color or intensity, where a pattern and algorithm is adjustable is known as segmentation.

PROBLEM DEFINITION

Gesture recognition has been reshaped for different research applications being it face movements gestures or whole body gestures (Dong, Yan, & Xie, 1998). Few applications has developed and created a hard requirement for this kind of recognition system (Dong et al., 1998). Coming to static recognition system, it is a design recognition problem, for instance, an important part of design recognition pre-processing level, called, feature extraction, must be controlled or managed before any standard pattern or design recognition process can be applied on it. Features correlate to the most preferential information regarding the image under specific lighting criteria. A good amount of research has been done on various aspects of feature extraction (Bretzner, Laptev, & Lindeberg, 2002; Gupta, Jaafar, & Ahmad, 2012; Parvini & Shahabi, 2007; Vieriu, Goraș, & Goraș, 2011). A method for identifying static and dynamic hand gestures by recognizing the movements analyzed by sensors attached with human hands has been proposed by Parvini and Shahabi, and the method achieved more than 75% of recognition rate on the ASL signs (Parvini & Shahabi, 2007). Furthermore, a user have to follow and use glove-based interface to extract the features of hand movements which controls their user-friendliness in the real world applications because a user needs gloves to interact with the system.

Developing recognition system which is efficient of working under different conditions is tough, but it is more possible because these hurdles exist in real-world environment. These criteria includes different compound and illumination background as well as few effects of translations, rotations and scaling by particular angles. Another condition that should be thought about is the expense of computing. Few feature extraction techniques have the disadvantage of being unpredictable and because of which it devours additional time, such as Gabor filters with combination of PCA (Gupta et al., 2012) which may limit their utilization in real-world applications. But the fact being is, the tradeoff between the accurate and computing cost in the hand gesture method should be taken into consideration (Chen, Fu, & Huang, 2003). Whereas, most of the hand gesture recognition

system focuses only on the accurateness for the assessment. It is needed, in the last phase of result evaluation, to consider two things, firstly, accurateness and the other being, computing cost to recognize their robustness and shortcomings and to advance their prospective applications (Chen et al., 2003).

PROJECT SCOPE AND OBJECTIVES

The scope of the project is to construct a synchronous gesture classifying system that can recognize gestures in lighting circumstances spontaneously. To achieve this goal, a synchronous gesture which based on real time is generated to recognize gestures. An intention of this project is to generate a complete system which can identify, spot and explain the hand motioning through computer sight. This structure will work as one of the envisioning of computer sight and AI with user interaction. It create function to identify hand motion based on various arguments. The topmost preference of the structure is to make it easy to use, simple to handle and user amiable without producing any specific hardware. All functions will appear on same Computer or workstation. Only some specific hardware will be used to digitalize the picture.

Literature survey

Literature Survey on Glove Based Approach

In this approach we attach sensor to mechanical or optical gloves that convert inflection of fingers into electrical signals for hand posture determination and additional sensor for position of the hand. This approach is in utilization for hand gesture recognition method using magnetic field which is attached to the glove.The use of gestures among humans, both in the form of pantomime (ridiculous sitiation) or by using sign language, is closely linked to speech and represents an effective way of communication, used even prior to talking. The formality of the set of rules chosen in each case is related to the validity of the performed gestures, which means that a ridiculous situation gesture could be commending speech in an unplanned manner.

Table 1: Literature review on Glove Based Analysis

Authors	Year	Description
D. J. Sturman and D. Zeltzer (Sturman & Zeltzer, 1994)	1994	The authors proposed technologies such as position tracking, optical tracking, marker systems, silhouette analysis, magnetic tracking or acoustic tracking.
L. Dipietro and A. M. Sabatini and P. Dario (Dipietro, Sabatini, & Dario, 2008)	2008	The authors analyzes the characteristics of the devices, provides a road map of the evolution of the technology, and discusses limitations of current technology
Abhishek, K. S, Qubeley, L. C. Fai and Ho, Derek (Abhishek et al., 2016)	2016	The authors proposed a prototype that recognizes gestures for the numbers 0 to 9 and the 26 English alphabets, A to Z using capacitive touch sensor.

Literature Survey on Vision Based Approach

Vision based approach has the prospective to come up with natural and non-contact solutions, and is built on the way humans explicate and interpret information about their surroundings. It is in all probability the most tough approach to execute(H. Hasan & Abdul-Kareem, 2014). A bare hand is used to extract data needed for recognition, and there is direct interaction between the user and with the system. For acquiring data needed for gesture analysis it uses some image characteristics like color and texture.

Table 2: Literature review on Vision Based Analysis

Authors	Year	Description
P. Garg, N. Aggarwal, and S. Sofa(Garg, Aggarwal, & Sofat, 2009)	2009	This paper is a review about Vision based Hand Gesture Recognition techniques for human computer interaction, combining the various available approaches, listing out their general advantages and disadvantages.
G. Murthy and R. Jadon,(Murthy & Jadon, 2009)	2009	The authors introduced the field of gesture recognition as a mechanism for interaction with computers.
M. K. Ahuja and A. Singh(Ahuja & Singh, 2015)	2015	The authors proposed a scheme using a database-driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching using PCA.

Literature Survey on Colored Marker approach

This approaches uses marked gloves wore in the hand and be colored to be helpful during the hand snapping to capture the fingers and palm. This glove frames the shape of hand by using the geometric-features. In (Lamberti & Camastra, 2011) utilized a wool glove with three different colors to represent the palms and fingers. This methodology considers basic and not costly whenever contrasted using Sensor or Data Glove (Lamberti & Camastra, 2011), however the basic interaction among human and computer still is insufficient.

Table 3: Literature review on colored marker approach

Authors	Year	Description
Wang, Robert Y Popovi, Jovan (Wang & Popović, 2009)	2009	The authors proposed an easy-to-use and inexpensive system that facilitates 3-D articulated user-input using the hands. Their approach uses a single camera to track a hand wearing an ordinary cloth glove that is imprinted with a custom pattern.
Lamberti, L Camastra, Francesco (Lamberti & Camastra, 2011)	2011	Their recognizer is formed by three modules. The first module, fed by the frame acquired by a webcam, identifies the hand image in the scene. The second module, a feature extractor, represents the image by a nine-dimensional feature vector. The third module, the classifier, is performed by means of *Learning Vector Quantization*.
Hasan, Mokhtar M Mishra, Pramod K (M. M. Hasan & Mishra, 2012)	2012	The authors had focused on the researches gathered to achieve the important link between human and his made machines, also they had provided their algorithms for overcoming some shortcomings existed in some mentioned algorithms.

Lighting

Increasing the illumination results in greater contrast between skin and background. The intensity must be set to provide ample light for the Charge-Coupled Device in the camera. It was concluded to extract the hand information in standard room lighting.

Camera Orientation and Distance

It's necessary to be attentive about supervision of camera to allow easy alternative of background. Couple of good and more fruitful proposal is to direct camera towards ground or wall. The strength and power of light would be high and the effect of shadow will be low because camera was directed towards down. The interspace of the camera from the hand should be such that it cover ups the whole motion mainly. No effect has been found on the accurateness of the structure if the picture is a focused one or not. Mainly the whole hand area should be covered.

Background selection

Color of background must be different from skin color to maximize differentiation. The ground or floor color used in the work was black. This color was chosen because it showcased minimum amount of self-shadowing issue in comparison with other background colors.

Requirement Analysis

Tools

Anaconda

Anaconda is an open-source distribution for python and R programming language. It is utilized for information science, machine learning, profound learning, and so on. With the availability of more than 300 libraries for information science, it turns out to be genuinely ideal for any developer to work on anaconda for information science.

Hardware Requirements

- Operating System: Windows10
- Processor: Intel(R)Pentium(R) CPU N3710 @1.60GHz
- System Type: 64-bit operating system, x64-basedprocessor
- Installed Ram: 8 GB
- GPU: NVIDIA GeForce GTX 800 or higher
- Web cam (For real-time hand Detection)

Software Requirements

Software used to execute this project is:

Python

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991.It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented, and functional programming.

OpenCV

OpenCV (Open Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage then Itself (which was later acquired by Intel). The library is cross-platform and free for use under the open-source BSD license.

Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

Methods

Proposed Methodology

The overall system comprises of two sections, back end and front-end. The back end framework comprises of three modules: Camera module, Detection module and Interface module as appeared in Fig. 1. They are summed up as follows:

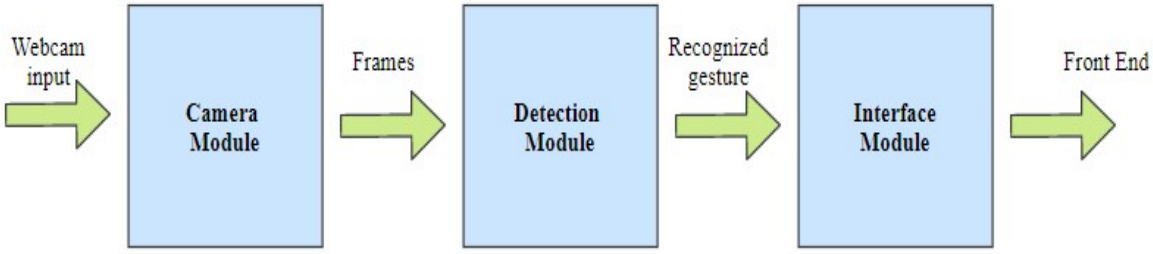

Fig 1 Back and Architecture

Camera module

This module is subject for interfacing and capturing input through the different sorts of picture markers and sends this picture to the detection module for handling as frames. The generally utilized techniques of capturing and recognizing input are hand belts, data gloves and cameras. In our framework, we use the inbuilt webcam which is financially savvy to see both static and dynamic signs.

Detection Module

This module is liable for the image processing. The output from camera module is presented to different image handling methods, for instance, color conversion, noise removal, thresholding following which the image goes through contour extraction. In the event that the image contains defects, at that point convexity defects are found by which the gesture is identified. In the event that there are no defects, at that point the image is classified utilizing Haar cascade to recognize the gesture.

Interface Module

This module is liable for calibrating the detected hand gestures to their associated actions. These actions are then passed to the suitable application. The front end comprises of three windows. The main window comprises of the video input that is captured from the camera with the corresponding name of the gesture identified. The subsequent window shows the contours found inside the input image. The third window shows the smooth thresholded adaptation of the image. The benefit of including the threshold and contour window as an aspect of the Graphical User Interface is to make the user aware of the background irregularities that would affect the input to the system and consequently they can adjust their laptop or desktop web camera so that it can be avoided. This would bring about better execution.

Proposed Method

The final architecture for any system to recognize the hand gesture could be elaborated as appeared in Fig 2. We proposed a gesture recognition system that follows a very efficient methodology. Our framework contains four steps, which are as followed.

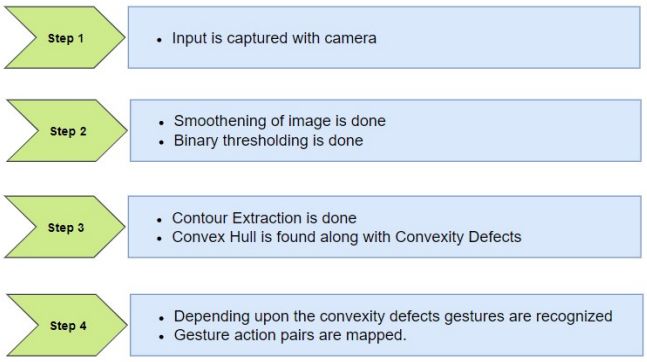

Fig. 2 Proposed method for our gesture recognition system

Image Capturing

In this initial phase we used a webcam to get the RGB image (frame by frame) using bare hand gestures only.

Pre-Processing

Next, in here this step, to minimize the calculation time we have taken just the crucial area instead of the whole frame from the video stream and it is known as Region Of Interest (ROI). Image processing works to manipulate over the color images into a grayscale image to progress the processing and after completing the processing it restores the images to its initial color space, in this way accordingly, we convert region of interest into a grayscale image. Point to be noted that in this step the algorithm will fail in the event that there's any vibration for the camera.

Hand region Segmentation

This phase is important in any process to hand gestures recognition and facilitate in developing the working of the system by eliminating the unnecessary data within the video stream. In basic, there are 2 ways to recognize the hand in image, the initial technique depends on Skin-Color, it is a straightforward but effective by light conditions in the environment and the nature of the background. The second technique, is on the form of hand and get profit from the principle of convexity in detection of the hand. The posture of hand is very important feature in the process of recognition the hand gesture (Li & Zhang, 2012).

There are other many techniques helpful to detect the hand region from the image may be summarized as:

A. Edge-Detection.
B. RGB values as a result of the values of RGB for hand completely different from the background of the image.
C. Subtraction of background

In this background subtraction method is used to separate the hand from the background. The background is identified from made the process target a certain scene for a least of 30 frames and through that generating the running average for the recent frame and all using the provided equation:

dst(x, y) = (1- alpha).dst (x, y) + alpha.src (x, y)

where, src (x,y) is a source photo may be one or three channels and 8-bits or 32-bits floating point, dst (x,y) is destination photo containing similar channels such as the source image and 32-bits or 64-bits floating point. Eventually, alpha is a weight of the source image and might be taken as threshold to generate out the time for calculate the running average over the frames.

After analyzing the background, we put the hand in front of the camera lens, after that calculate absolute difference between the background that calculates by utilizing the running average and the current frame that contains the hand as a foreground object. This method is called background subtraction.

The next step is thresholding the image which is performed after background subtraction in which the result are only gestures of hand in white color. This method is very vital and should be done before the contours get a method to attain high accuracy.

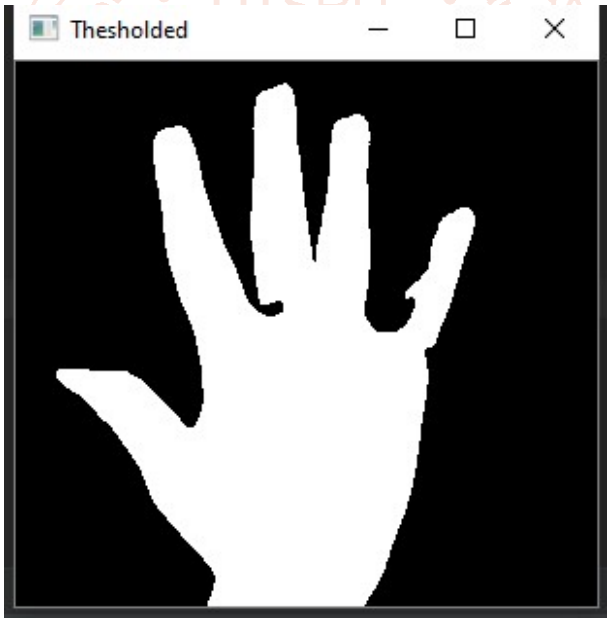

Fig. 3 Hand region segmentation process

Contour Extraction

The contour is outlined as object's (hand) boundary that can be seen in the image. The contour can also be a wave connecting points that has the similar color value and is important in shape analyzing, objects identification method.

Feature Extraction and recognition

The convex hull cluster is of peaks that covers the region of hand. In here, we must clear the principle of the convex Set, which means all lines between any 2 points within hull are entirely within it.

After determining the gesture, the specific functioning is performed. The method of recognizing the movement is a dynamic process. After operating the specific command from the gesture, go back to the initial step to accept other image to be processed and so on.

Result and Discussion

This project recognized the count of fingers as shown in figure. Our initial approach to form a gesture recognition system was through the tactic of background subtraction. Many problems and accuracy issues were faced while implementing recognition

system using background subtraction. Background subtraction cannot take care of sudden, drastic lighting changes resulting in many inconsistencies. The gesture recognition system when used against any plain background was sturdy and performed with good accuracy. This accuracy was maintained no matter the color of the background, provided it's a plain, solid color background empty of any inconsistencies. In cases wherever the background wasn't plain, the objects within the background verified to be inconsistencies to the image capture method, leading to faulty outputs. So it's recommended that this system be used with a clear background to supply the simplest potential results and good accuracy.

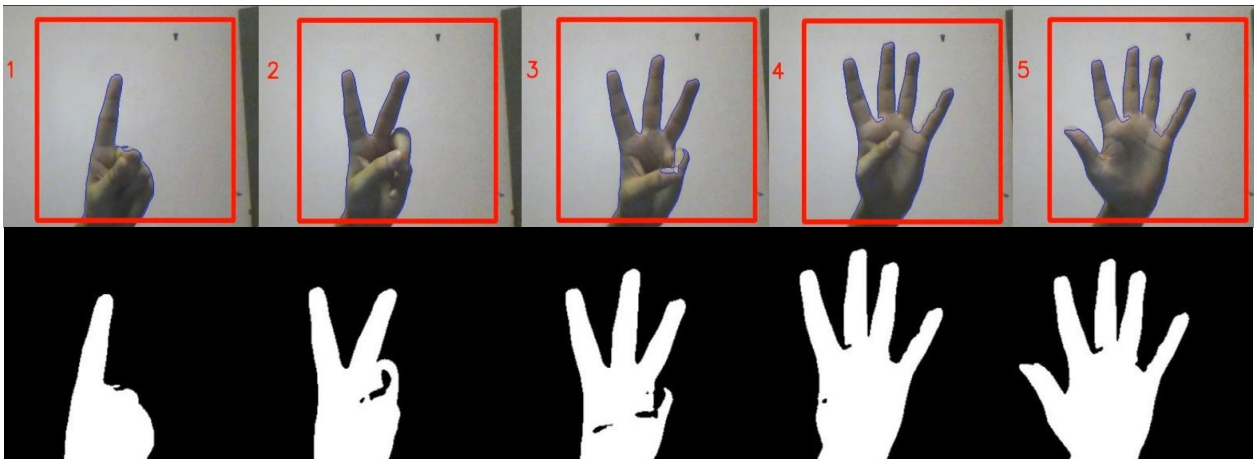

Fig 4 Five hand gestures

Table 4: Accuracy of each gesture with plain background and non-plain background

Gesture	Accuracy with plain background (in %)	Accuracy with non-plain background (in %)
1	95	45
2	94	48
3	96	46
4	91	40
5	92	41

Conclusion and Future Scope

For a prolonged hour, an issue of differentiating movement was vital in computer perception due to the opposition of removing the targeted object, like the hand from a framework which was making mess in actual time. In actual fact, a person while gazing to a certain picture can effortlessly identify what is in it whereas, the same things is much tough for the computer if it looks at the same picture due to its functionality of dealing with a picture as a three dimensional matrix.

In future we would like to improve the accuracy further and add more gestures to implement more functions.

Bibliography

[1] Abhishek, S, K., Qubeley, Fai, L. C., Ho, & Derek. (2016). Glove-based hand gesture recognition sign language translator using capacitive touch sensor. In *2016 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC)* (pp. 334-337): IEEE.

[2] Ahuja, M. K., & Singh, A. (2015). *Static vision based Hand Gesture recognition using principal component analysis.* Paper presented at the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE).

[3] Bretzner, L., Laptev, I., & Lindeberg, T. (2002). *Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering.* Paper presented at the Proceedings of fifth IEEE international conference on automatic face gesture recognition.

[4] Chen, F.-S., Fu, C.-M., & Huang, C.-L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image and vision computing, 21*(8), 745-758.

[5] Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A Survey of Glove-Based Systems and Their Applications. *Ieee transactions on systems, man, and cybernetics, part c (applications and reviews), 38*(4), 461- 482.

[6] Dong, G., Yan, Y., & Xie, M. (1998). *Vision-based hand gesture recognition for human-vehicle interaction.* Paper presented at the Proc. of the International conference on Control, Automation and Computer Vision.

[7] Garg, P., Aggarwal, N., & Sofat, S. (2009). Vision based hand gesture recognition. *World academy of science, engineering and technology, 49*(1), 972-977.

[8] Gupta, S., Jaafar, J., & Ahmad, W. F. W. (2012). Static hand gesture recognition using local gabor filter. *Procedia Engineering, 41*, 827-832.

[9] Hasan, H., & Abdul-Kareem, S. (2014). Retracted article: Human–computer interaction using vision-based hand gesture recognition systems: A survey. *Neural Computing and Applications, 25*(2), 251-261.

[10] Hasan, M. M., & Mishra, P. K. (2012). Hand gesture modeling and recognition using geometric features: a review. *Canadian journal on image processing and computer vision, 3*(1), 12-26.

[11] Lamberti, L., & Camastra, F. (2011). *Real-time hand gesture recognition using a color glove.* Paper

presented at the International Conference on Image Analysis and Processing.

[12] Li, L., & Zhang, L. (2012). Corner Detection of Hand Gesture. *TELKOMNIKA Indonesia Journal of Electrical Engineering, 10*(8), 2088-2094.

[13] Murthy, G., & Jadon, R. (2009). A review of vision based hand gestures recognition. *International Journal of Information Technology and Knowledge Management, 2*(2), 405-410.

[14] Parvini, F., & Shahabi, C. (2007). An algorithmic approach for static and dynamic gesture recognition utilising mechanical and biomechanical characteristics. *International journal of bioinformatics research and applications, 3*(1), 4-23.

[15] Sturman, D. J., & Zeltzer, D. (1994). A survey of glove-based input. *IEEE Computer Graphics and Applications, 14*(1), 30-39.

[16] Vieriu, R.-L., Goraș, B., & Goraș, L. (2011). *On HMM static hand gesture recognition.* Paper presented at the ISSCS 2011-International Symposium on Signals, Circuits and Systems.

[17] Wang, R. Y., & Popović, J. (2009). Real-time hand-tracking with a color glove. *ACM transactions on graphics (TOG), 28*(3), 1-8.