IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

HAND GESTURE RECOGNITION TECHNIQUES FOR HUMAN COMPUTER INTERACTION USING OPENCV

[1]Prapti Mandal, [2] Akansha S. Choubey, [3]Vibha Pandey

[1]MTech Scholar, [2,3]Assistant Professor
[1]Department of Computer Science
[1] Shri Shankaracharya College of Engineering and Technology, Bhilai, India

Abstract. Hand gesture recognition is one technology that can figure out what a person is doing with their hands in a live video. The way the hand moves fits into a given field of study. In this study, creating a system that can recognize hand gestures is one of the hard tasks that combines two big problems. First, there is hand detection. Another challenge is making a sign that can be utilized with just one hand at a time. This study focuses on how a system might detect, recognize, and understand hand gestures by computer vision, even when the position, orientation, location, and size of the hands can change. In this system, several kinds of gestures, such numbers and sign languages, need to be made for this project to work successfully. Before the image processing is done, the picture from the real-time video is analyzed with a Haar-cascaded Classifier to find the hand gesture or, in other words, to find the appearance of a hand in a frame. In this project, the hand will be found by utilizing Python programming and the ideas of Region of Interest (ROI). The part of the findings that will be explained in detail will be the simulation, as the only difference between the simulation and the hardware implementation is the source code to read the real-time input video. Using the ideas of hand segmentation and the hand detection system that uses the Haar-cascade classifier, hand gesture recognition may be built using Python and OpenCV.

Keywords-hand gesture, human computer interaction (HCI), contour, convex hull, convexity defects, gesture recognition, python, openCV.

I. INTRODUCTION

Hand gestures are a strong and unprompted way for humans and computers to communicate. (HCI). Some ways to connect to a computer are through a keyboard, mouse, joystick, or touch screen, but these don't provide the right interface. In contrast, the current system will have either a desktop or a laptop interface, where hand gestures can be made by wearing data gloves or using a web camera to take a picture of the hand. Taking a picture of and analyzing the hand is the initial step towards this gesture identification. In Data-Glove approaches, sensors are utilized to start moving fingers, and additional sensors will programmer hand movements. The vision-based solution, on the other hand, merely needs a camera and can figure out how a person and a computer actually communicate without any additional tools. The problems with this technique are that the background is always there, and occasionally people and illumination present as well. This section goes into detail on the many methods and algorithms utilized in this system, as well as the several ways to recognize things. Segmentation is the process of looking for a connected area in a photograph that meets certain criteria, such as color or brightness, using a pattern and algorithm that may be changed..

II. PROBLEM DEFINITION

Whether it's for recognising facial expressions or a whole range of body motion, gesture recognition has evolved to meet the needs of many fields of study. (Dong, Yan, & Xie, 1998). There are just a handful of use cases that have established firm demands for this sort of identification system. (Dong et al., 1998). A typical pattern or design recognition procedure cannot be applied to a static recognition system without first controlling or managing an essential aspect of the design recognition pre-processing phase, called, feature extraction. The most salient characteristics of a picture under a certain set of lighting conditions are correlated with the features. Extensive studies have been conducted on several facets of feature extraction (Bretzner, Laptev, & Lindeberg, 2002; Gupta, Jaafar, & Ahmad, 2012; Parvini & Shahabi, 2007; Vieriu, Goraş, & Goraş, 2011). Parvini and Shahabi devised a method for distinguishing static and dynamic hand gestures by recognising the motions analysed by sensors attached with human hands; their method obtained a recognition rate of greater

than 75% on the ASL signals. (Parvini & Shahabi, 2007). Additionally, a user must follow and utilise a glove-based interface to extract the characteristics of hand movements that regulates their usability in practical applications.

Real-world challenges make it all the more viable to create a recognition system that can function well in a wide variety of settings. Among them include the presence or absence of certain effects caused by translations, rotations, and scaling at specific angles, as well as the presence or absence of a distinct chemical and lighting backdrop. The cost of computation is another factor to take into account. Some feature extraction methods, such as Gabor filters combined with principal component analysis (PCA) (Gupta et al., 2012), can be unreliable and time-consuming, limiting their practical use. However, the reality remains that the hand gesture approach involves a compromise between accuracy and computing expense. (Chen, Fu, & Huang, 2003). However, evaluation accuracy is typically overlooked by most hand gesture detection systems. The final step of evaluating results requires thinking about two things: first, how accurate they are, and second, how much computer power is required to fully understand their strengths and weaknesses and further develop their potential uses. (Chen et al., 2003).

III. PROJECT SCOPE AND OBJECTIVES

The goal of this research is to develop a system capable of instantaneously recognizing motions regardless of the surrounding illumination conditions. To do this, we create a gesture recognition system that works in real time and is based on synchronous movements. The goal of this work is to develop a comprehensive system that can, using computer vision, recognize and make sense of hand gestures. This design is one way that computer vision and AI might be envisioned in tandem with human input. Using those parameters, it generates a function that can recognize hand motion. The design's primary focus is on making the framework accessible and user-friendly without requiring the creation of any specialized hardware. The same computer or workstation will serve as the hub for all operations. The digitization of the film will be limited to a select few pieces of gear.

IV. Literature survey

Literature Survey on Glove Based Approach

In this method, sensors are affixed to mechanical or optical gloves, which translate finger flexion into electrical impulses for use in determining hand posture, and an extra sensor is used to track the hand's location. Magnetic fields coupled to the glove are used in this method for recognizing hand gestures. Humans have been employing gestures as a method of communication since long before the invention of spoken language, whether through pantomime (acting out an absurd situation) or sign language. A absurd scenario gesture might be praising speech in an unintentional way, depending on the formality of the rule set chosen in that specific instance.

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.
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.
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 forovercoming some shortcomings existed in some mentioned algorithms.

V. 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:

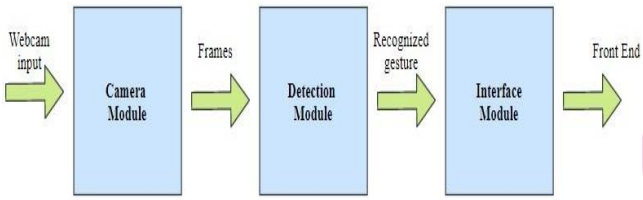

Fig 1 Back and Architecture

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 threshold 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.

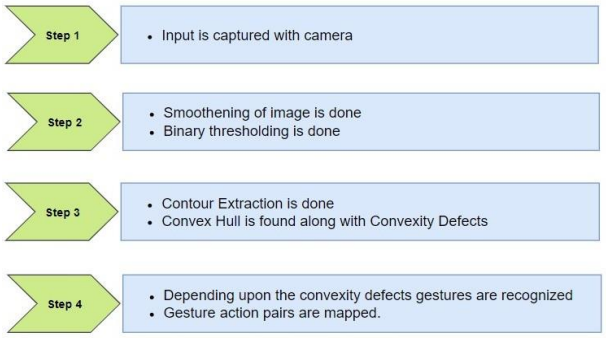

Fig. 2 Proposed method for our gesture recognition system

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.

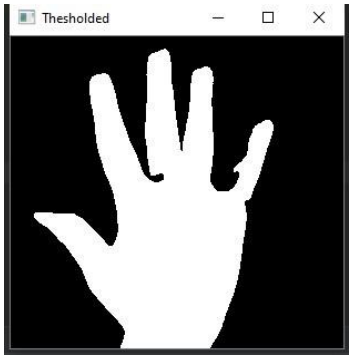

Fig. 3 Hand region segmentation process

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.

VI. 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.

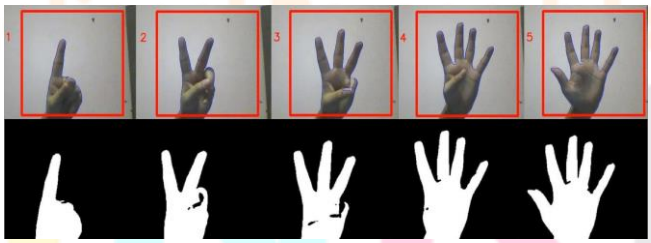

Fig 4 Five hand gestures

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

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

Conclusion and Future Scope

For a prolonged hour, an issue of differentiating movementwas 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 actualfact, a person while gazing to a certain picture can effortlessly identify what is in it whereas, the same things ismuch tough for the computer if it looks at the same picturedue to its functionality of dealing with a picture as a three dimensional matrix.

In future we would like to improve the accuracy further andadd 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 ElectronDevices 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 andTechnology 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] [4] Chen, F.-S., Fu, C.-M., & Huang, C.-L. (2003). Hand gesture recognition using a real-time tracking methodand hidden Markov models. *Image and vision computing, 21*(8), 745-758.

[5] Dipietro, L., Sabatini, A. M., & Dario, P. (2008). ASurvey of Glove-Based Systems and TheirApplications. *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 handgesture recognition for human-vehicle interaction.* Paper presented at the Proc. of the Internationalconference 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: areview. *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 ImageAnalysis and Processing.

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

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

[14] Parvini, F., & Shahabi, C. (2007). An algorithmic approach for static and dynamic gesture recognitionutilising mechanical and biomechanical characteristics. *International journal of bioinformaticsresearch 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., Goraş, B., & Goraş, 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.