

GESTURE RECOGNITION SYSTEM

Mrs Ayesha Butalia
MIT-College of Engineering

Ms Divya Shah
MIT-College of Engineering

Dr. R.V Dharaskar
GHRCE, Nagpur

Abstract Gestures are a major form of human communication. Hence gestures are found to be an appealing way to interact with computers, as they are already a natural part of how we communicate. A primary goal of gesture recognition is to create a system which can identify specific human gestures and use them to convey information for device control and by implementing real time gesture recognition a user can control a computer by doing a specific gesture in front of a video camera linked to the computer.

A primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control.

This project covers various issues like what are gesture, their classification, their role in implementing a gesture recognition system, system architecture concepts for implementing a gesture recognition system, major issues involved in implementing a simplified gesture recognition system, exploitation of gestures in experimental systems, importance of gesture recognition system, real time applications and future scope of gesture recognition system. The algorithm used in this project are Finger counting algorithm, X-Y axis (To recognize the thumb).

1. INTRODUCTION

A primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control. To help understand what gestures are, an examination of how other researchers view gestures is useful. How do biologists and sociologists define "gesture"? How is information encoded in gestures? We also explore how humans use gestures to communicate with and command other people. Furthermore, engineering researchers have designed a variety of "gesture" recognition systems - how do they define and use gestures?

1.1 Biological and Sociological Definition and Classification of Gestures

From a biological and sociological perspective, gestures are loosely defined, thus, researchers are free to visualize and classify gestures as they see fit. Speech and handwriting recognition research provide methods for designing recognition systems and useful measures for classifying such systems. Gesture recognition system which are used to control memory

and display devices in a local environment and remote environment are examined for the same reason.

People frequently use gestures to communicate. They are used for pointing to a person, to get his attention & convey information about space and temporal characteristics. Evidence indicates that gesturing does not simply embellish spoken language, but is part of the language generation process.

Biologists define "gesture" broadly, stating, "the notion of gesture is to embrace all kinds of instances where an individual engages in movements whose communicative intent is paramount, manifest, and openly acknowledged". Gestures associated with speech are referred to as gesticulation. Gestures that function independently of speech are referred to as autonomous. Autonomous gestures can be organized into their own communicative language, such as American Sign Language (ASL). Autonomous gestures can also represent motion commands.

2. RELEVANCE

2.1 Need

GRS is not only beneficial for performing my computer task but its scope is very vast in day-to-day technical solutions.

2.1.1 Cybernet's Powerpoint

Cyber net in Ann Arbor has created a gesture-recognition device that translates hand gestures into PowerPoint commands. Cyber net's latest research has focused on getting rid of the suit to make virtual worlds more realistic. That means creating software that can "read" a person's body movements and sync it with his virtual surroundings. A by-product of this research is the development of an interface that enables the user to control a PowerPoint presentation wirelessly with simple hand gestures. In fact, it makes use of a gesture recognition software and a camera and you can run an entire presentation without touching anything -- no remotes, no buttons, nothing.

2.1.2 Pointing Device

Many gesture recognition systems are used as pointing devices. The use of the hand as a pointing device similar to a laser pointer. The system presented by Cantzler and Hoile is an example of this. Their system is designed to replace conventional 2D pointing devices like those such as touch pads, trackballs, and mice.

2.2 Goal

The goal of this project is to develop a program implementing real time gesture recognition. At any time, a user can exhibit his hand doing a specific gesture in front of a video camera linked to a computer. However, the user is not supposed to be exactly at the same place when showing his hand. The program has to collect pictures of this gesture, analyze it and identify the sign. It has to do it as fast as possible, given that real time processing is required. In order to lighten the project, it has been decided that the identification would consist in counting the number of fingers that are shown by the user in the input picture.

3. PRESENT THEORY AND PRACTICES

3.1. Previous System

The need for a gesture recognition system (GRS) has been very prevalent over the past decade. Many different problems have found their solutions in gesture recognition. More so, many unique applications have arisen from this technology. Previously, it was inconvenient to interact with the system during presentation and seminars.

The need for a perceptual human computer interface has been around for sometime now. With the onset of computer encounters and interactions becoming more prevalent in our daily life, the need for an intuitive interface to interact with a computer has increased.

3.1.1 Identification of problem in existing system:

Previously, it was inconvenient to interact with the system during presentation ,seminars etc.

Some problem in existing system is as follows:

- It is difficult for non-technical people to interact with computer system for which our module proves to be useful.
- In Existing system there was no efficient & convenient way

with which we could perform the simple My Computer task during the presentation/seminar etc.

3.2. Proposed System

To solve the problem in existing system, a system is provided called gesture recognition system. A primary goal of gesture recognition system is to create a system which can identify specific human gestures and use them to convey information or for device control.

The algorithm “Weighted Averaging Analysis“ used in our project gives more efficient result. About 96% of the signs are correctly classified. Such results are so efficient that it does not need any training like in neural networks. The processing is quite fast given that no sophisticated calculus is required while running the program. The memory requirement is also very less as we are doing statistical analysis and not the database matching like in traditional image processing.

4. WORKING

During implementation one thing was clear that a system is going to be developed which can capture a hand gesture performed by the user in front of web Cam, this capture image is then proceed to identify the valid gesture through specific algorithm & execute the corresponding operation. The overall implementation process is described as follows:

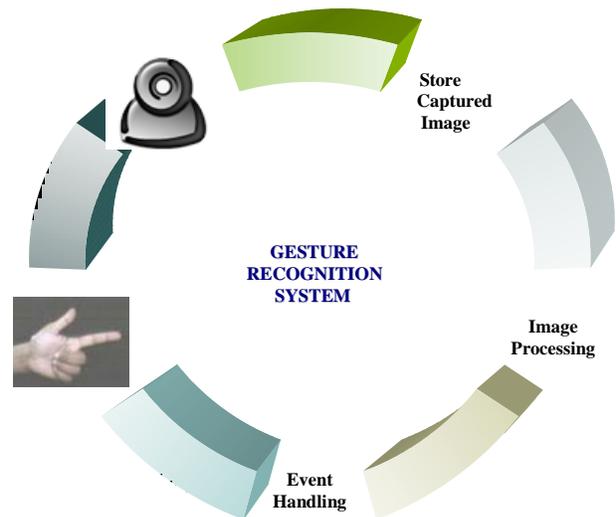


Fig 1 Interaction among the components

4.1. Human Generated Gesture:

As a first step of implementation user will show one gesture. The gesture should be constant for some period of time, which is necessary for dynamic processing. These gestures should be already defined as valid gesture for processing.

4.2. Web Camera

The purpose of web camera is to capture the human generated hand gesture and store it in memory. The package called Java Media Framework is used for storing image in memory and again calling the same program after particular interval.

4.3. Image Processing Algorithm:

This carries the major portion of implementation. First the captured image is preprocessed by techniques like making binary, zooming, cropping and standard resizing. Such preprocessed image is given to the image-processing algorithm. The algorithm will count the number of fingers shown by user, which will work as input for next processing.

4.4. Event Handling:

Once the gesture is identified the appropriate command for it will be executed.

This includes opening, traversing my computer contents as per user requirement.

Shortcut for applications like notepad, WordPad are also provided. Other control

commands include shutdown and restart facilities using gestures.

4.5. Back To Capturing Gestures:

Gesture recognition is a dynamic process so once particular gesture is identified and appropriate control command is executed it will again go to capture next image and process it accordingly.

5. IMPLEMENTATION

5.1 System Description

5.1.1 Collecting the pictures

First of all, it will be necessary to collect pictures of all the gestures that will be supported by the Gesture recognition system. Every gesture is associated a particular task through which the user can control the computer.

5.1.2 Finding the hand

We need then to analyze the picture, and to find the relevant part of the picture. Indeed the user will never put his hand in the same area of the picture.

In any cases, the user expects the program to count same no of fingers. So it is necessary to find the hand in the picture, that is to say to zoom on the hand, and to pick it up.

5.1.3 Zooming on the Hand

- Finding the hand in the picture

According to the requirements, the web camera is not supposed to move. This piece of information gives a huge advantage that allows simplifying the zooming process. Indeed, it implies that the background is more or less always the same. In all what follows, it will be supposed that in the picture, we can just find the hand and the background: no other object should be present.

Please note that all correct sign are suppose to include thumb finger, which is used as a reference for zooming on the hand. After processing noise removal, the resulting picture will be black almost everywhere except where the hand is. So, zooming can then be easily realized by cropping areas whose pixel values are close to 0.

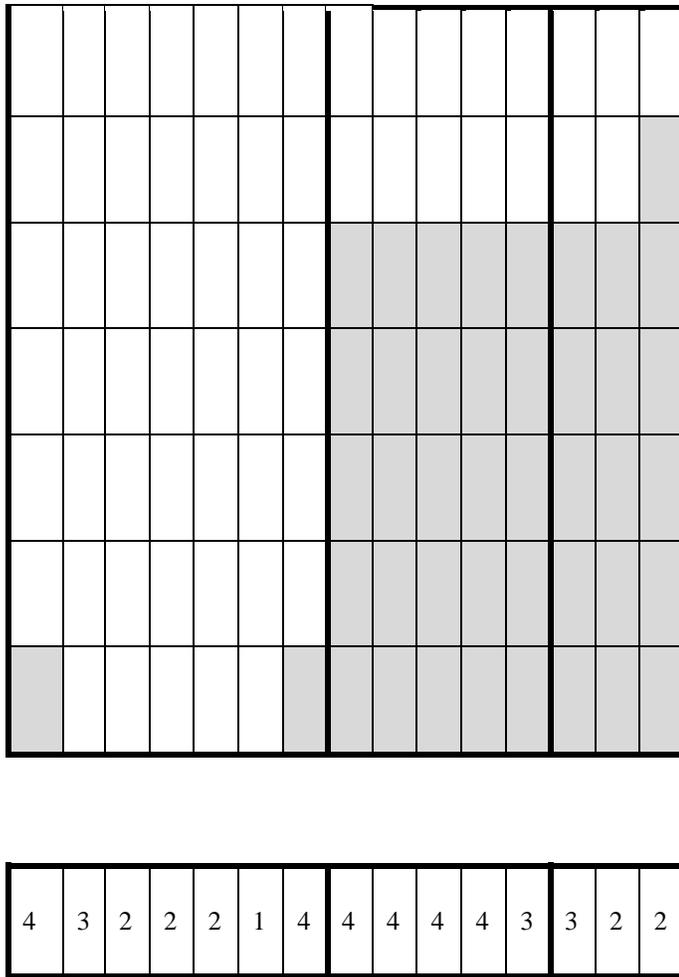
- Creating a Binary Picture

To make all the preprocessing easier, it is better to create a binary picture. To do so, it is necessary to choose a threshold: pixels with value lower than this threshold will be set to 0 (black) and others will be set to 1. The choice of this threshold depends on the web camera properties. Then it is necessary to execute noise-removal functions, else every noisy pixel that its value is too high may be considered as part of the hand and will be included in the zoom-in picture.

- Zooming in the Binary Picture

Let's suppose that the picture is completely black (0), except in the area of the hand in which it is completely white (1), according to the previous few examples.

The problem now is how to crop the area completely black. The fastest method is to compute a line vector in which the i -th element is the sum of the elements of the i -th column of the picture, and a column vector in which the i -th element is the sum of the elements of the i -th line of the picture.



Vector: contains the sum of edges in each column

Fig 4: Analysis of binary picture counting the number of edges in each column

In this schematic example (a 15x15 pixels example, to make it more readable), it is easy to understand that the last columns have edges less relevant than the columns in the middle-right part. The calculus of the number of fingers is quite easy: one finger is the thumb, and count two edges per finger, so finally:

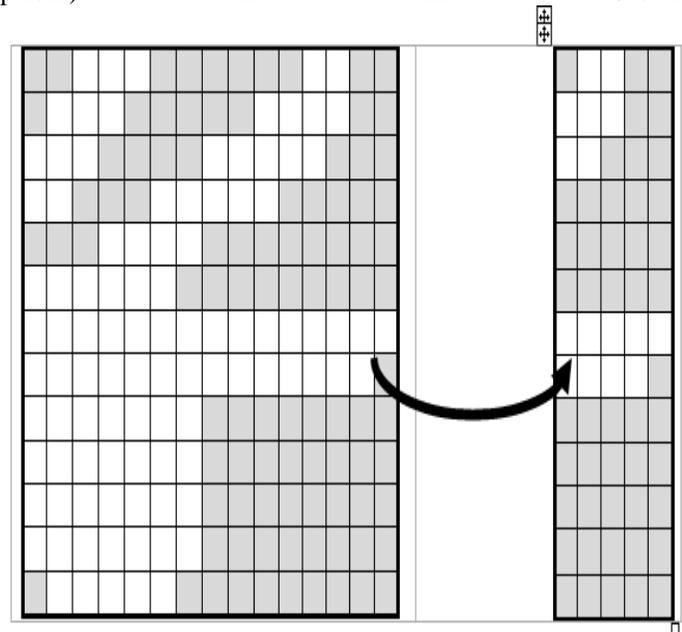
$$\# \text{ fingers} = 1 + \frac{\# \text{ edges}}{2}$$

is the estimated number of fingers. The edge counting method is not subject to constraints like size of the finger because the number of edges should always be twice the number of fingers,

without any dependence on the sensitive changes in the size of the finger: thin and thick fingers have always two edges.

B. Finger Counting Algorithm

Another method that has been tried, that implements in a different way of counting the number of fingers. Initially the RGB image is converted into binary as seen. To get the desired gesture the image is cropped and zoomed. After this when the pixel value changes from 1 to 0 and then 0 to 1 we consider it as one finger. We follow the same method to detect more fingers (if present) in the Gesture.



Picture after final cropping with no. of fingers

Fig 5: Counting of fingers after cropping

C. Counting through X-Y axis.

In this algorithm the thumb and fingers are detected separately. X-axis is used for detection of thumb and Y-axis is used for detection of fingers. After having done the research work about which algorithm is to be used, from the above mentioned we have implemented X-Y axis algorithm in our project.

As seen in the diagram below, Thumb being vertical and fingers being horizontal, it is very easy to detect the presence of thumb through X-axis which helps us to see the presence of Thumb with the help of pixel counting acting as Thumbs width. Similarly, for fingers being horizontal, it becomes very easy to detect the presence of finger through Y-axis and if present, it helps us to count the number of fingers present in the gesture.



Fig 6: Detection of Thumb through X-axis



Fig 7: Detection of Finger through Y-axis

The following figure shows **Representation of Gesture in the form of Pixels**

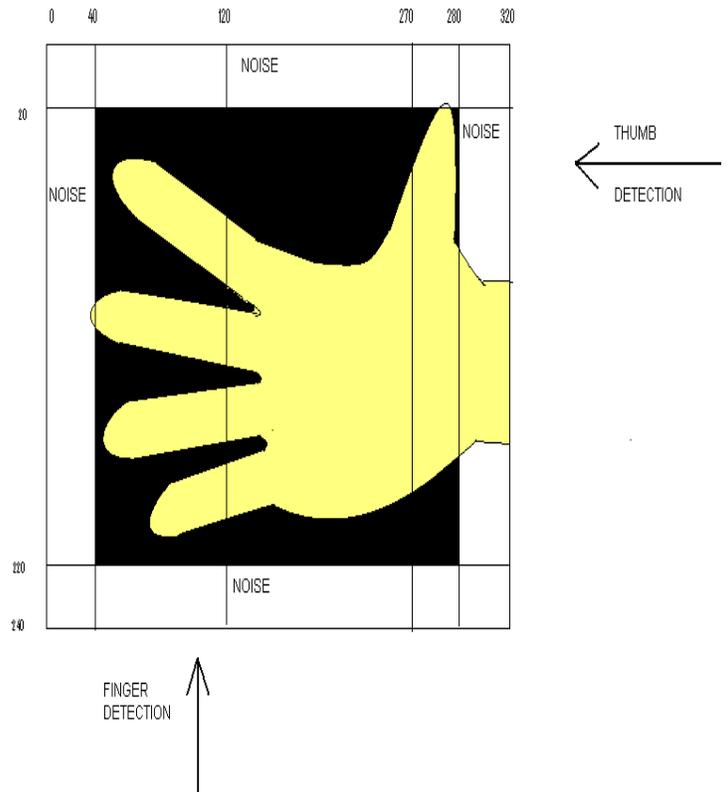


Fig 8: Implementation of counting fingers with the pixels of 320x240.

6. APPLICATIONS AND FUTURE ENHANCEMENT

6.1 Application

GRS is not only beneficial for performing my computer task but its scope is very vast in day-to-day technical solutions.

Some are given below:

- CARPEL TUNNEL SYNDROME
- CYBERNET'S POWERPOINT
- CNOKIA'S POD-PHONE
- POINTING DEVICE

6.2 Future Enhancement

- By adding more gestures, we can handle all My Computer operations like Cut, Copy, Paste and Undo etc.
- By integrating our system with voice recognition system we can embed it in ROBOTS.
- We can enhance our system to control PowerPoint application.
- We are also able to handle dynamic image processing and event handling accordingly.

REFERENCES

- [1] Herbert Schildt, “The Complete Reference (Java 2 Fifth Edition)“, McGraw-Hill Publication.
- [2] Matthew T. Nelson “Java Foundation Classes“, McGraw-Hill Publication.
- [3] Matthew Robinson, Pavel Vorobiev, ”Swing”, O’RELLY.
- [4] Robert Eckstein, Marcloy and Dave Wood, “Java Swing“, O’RELLY.
- [5] David Flanagan, “Java Foundation Classes”.