HAND GESTURE RECOGNITION FOR THE AUDITORY IMPAIRED

AsmaParveen
Mentor
VESIT
Mumbai
asma.parveen@ves.ac.in

VedantChoudhary
Student
VESIT
Mumbai
vedant.choudhary@ves.ac.in

KritiHeda
Student
VESIT
Mumbai
kriti.heda@ves.ac.in

RaunakKhithani
Student
VESIT
Mumbai
raunak.khitani@ves.ac.in

Atharva Kulkarni
Student
VESIT
Mumbai
atharva.kulkarni@ves.ac.in

Abstract -
More than 500 million people of the world suffer from some physical, sensory or mental disability. Often their lives are handicapped by physical and social barriers which hamper their full participation from society and the enjoyment of equal rights and opportunities. Many mute people use sign language. Sign language uses gestures instead of sound to convey meaning, simultaneously combining hand shapes, orientations and movement of the hands, arms or body and facial expressions to express fluidly a speaker's thoughts. Signs are used to communicate words and sentences to audience. This project focuses on developing an application for recognition of hand gestures with reasonable accuracy, where the input to the pattern recognition system will be given from the hand. It recognizes the pattern and the display the pattern in the form of the text. The information given by the camera are collected and stored in the database. The exact meaning of the information is matched with the samples stored previously in the database. The image is processed considering the parameters like the number of fingers used, the angles between them.

I. INTRODUCTION

A gesture is defined as an energetic movement of hands and creating signs with them such as alphabets, numbers, words and sentences. Gestures are classified into two type static gestures and dynamic gestures. Static gestures refer to certain pattern of hand and finger orientation whereas dynamic gestures involve different movement and orientation of hands and face expressions largely used to recognize continuous stream of sentences.

The aim of the gesture recognition is to enable humans to interact with the human made machines in natural way without any mechanical devices and the mathematical equations will be the translator that translates the poses between the gestures and the telerobotic . The gesture recognition is very difficult and complex task since the full recognition system should be able to identify the hand in different scales, positions, orientations, contrasts, luminosity, and others. Automatic sign languages recognition can be of great significance for communication with deaf people and has also its application in virtual reality, machine control in the industrial field

A. Gesture.

It is a motion of the body that is intended to communicate with other agents. For a successful communication, a sender and a receiver must have the same set of information for a particular gesture. As per the context of the project, gesture is defined as an expressive movement of body parts which has a particular message, to be communicated precisely between a sender and a receiver.

A gesture is scientifically categorized into two distinctive categories:

- dynamic gesture: A dynamic gesture is intended to change over a period of time
- Static gesture: A static gesture is observed at the spurt of time.

Example: A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time.

This complex process is called **gesture recognition**.

B. Sign Language

Sign languages are being used extensively in international sign use of deaf and dumb, in the world of sports, for religious practices and also at work places . Gestures are one of the first forms of communication when a child learns to express its need for food, warmth and comfort. It enhances the emphasis of spoken language and helps in expressing thoughts and feelings effectively.

A simple gesture with one hand has the same meaning all over the world and means either 'hi' or 'goodbye'. Many people travel to foreign countries without knowing the official language of the visited country and still manage to perform communication using gestures and sign language. These examples show that gestures can be considered international and used almost all over the world. In a number of jobs around the world gestures are means of communication

In our project we are using hand gesture recognition to identify and understand the sign language used by the deaf and dumb to communicate.

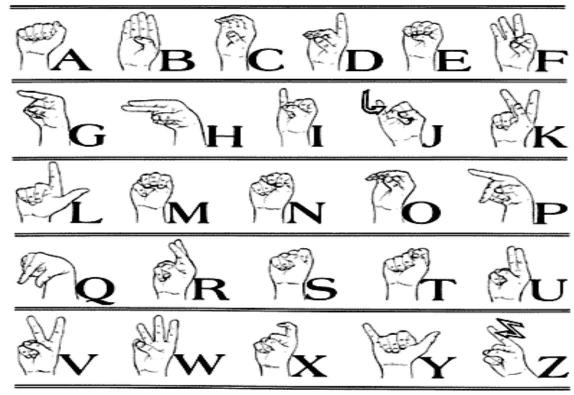


Fig 1 : American Sign Language

II. REQUIREMENTS

A. Functional Requirements

- Functional requirements specifies a function that a system or system component must be able to perform. It can be documented in various ways. The most common ones are written descriptions in documents, and use cases.
- Use cases can be textual enumeration lists as well as diagrams, describing user actions. Each use case illustrates behavioural scenarios through one or more functional requirements. Often, though, an analyst will begin by eliciting a set of use cases, from which the analyst can derive the functional requirements that must be implemented to allow a user to perform each use case.
- Functional requirements is what a system is **supposed to accomplish**. It may be
    - Calculations
    - Technical details
    - Data manipulation
    - Data processing
    - Other specific functionality

  A typical functional requirement will contain a unique name and number, a brief summary, and a rationale. This information is used to help the reader understand why the requirement is needed, and to track the requirement through the development of the system.

B. Non-Functional Requirements

Typically non-functional requirements fall into areas such as [1]:

- Performance requirements
- Interface requirements
- Operational requirements
- Resource requirements
- Verification requirements
- Acceptance requirements
- Documentation requirements
- Security requirements
- Portability requirements
- Quality requirements
- Reliability requirements
- Maintainability requirements
- Safety requirements

C. Hardware and Software Requirements

Minimum Requirements:

Hardware:

- RAM 512 MB or higher
- Processor : P4 or higher (2.4 Ghz)
- Hard disk : 15GB

Software:

- Windows 7 or higher
- Microsoft C++ Redistributable 2010 or higher
- MATLAB 7.0 or higher
- .NET Framework

The Proposed system will be developed on:

Hardware:

- RAM 2GB
- Processor : i5(2.4Ghz)
- Hard disk : 15GB
- 1080*1020 full

Software:

- Windows 7
- Microsoft C++ Redistributable 2012
- MATLAB 7.2
- .NET Framework

III. BASIC OPERATIONS

1. Input for the gesture recognition system must be predisposed and decided.
2. **Image preprocessing:** Segmentation and tracking are essential in order to extract useful information from raw gesture images. Thus, it is necessary to be able to recognize the region of foreground and split it from the background in a given gesture image.
3**. Feature vector extraction :** In the third stage, the features (geometric and non-geometric) must be extracted; these features will be used at the time of testing operation.
4**. Efficient classification algorithm**: Classifier is used to identify which trained class the current presented testing gesture belongs to. [2]

IV. TECHNIQUES STUDIED

A. PREPROCESSING

i. RGB to Grayscale
ii. Binarization
Iii Grayscale Filtering
iv Noise removal and smoothing
v Region Filling

B Hand Detection

First of all colored image is read which is captured in image acquisition step. Once we get the image, the dimensions of the image are calculated. Number of color bands should be one. If the image is not a grayscale, convert it to grayscale by only taking green channel. Now find the biggest blobs. This technique results in giving two biggest blobs, ignore the first biggest blob, which is the largest one. The second biggest blob will be the hand. This result in drawing box around the blobs and second biggest blob is separated from the image. The limitation of this technique is that color of clothes and other objects in scene might effect it.

C. Hand Cropping

Once the portion of hand is separated from the Image, the hand is cropped out, for this certain threshold is used. Actually in binarizing of the image a threshold value is used, which only gives out the portion of image with hand and then we can crop out the hand. This image of hand is then stored and passed to the next phase.

D. Feature Extraction

There are various algorithms used for features extractions like Zernike moments and Fourier descriptors. In general, descriptors are some set of numbers that are produced to describe a given shape.

A few simple descriptors are:
1. Area: The number of pixels in the shape.
2. Perimeter : The number of pixels in the boundary of the shape
3. Elongation: Rotate a rectangle so that it is the smallest rectangle in which the shape fits.Then compare its height to its width.
4. Rectangularity: How rectangular a shape is (how much it fills its minimal bounding box) area of the object.
5. Orientation: The overall direction of the shape.
Moments are common in statistics and physics.
What Statistical Moments Are?
1)Mean
2)Variance
3)Skew
4)Kurtosis
Moment of image is weighted average of the images (Intensities of Pixels) they are usually have some attractive property. It is useful to describe shapes in an image (Binary) after segmentation. Using image moments one can find simple properties of an image such as area (intensity), centroid and orientation of object inside an image.
[3]

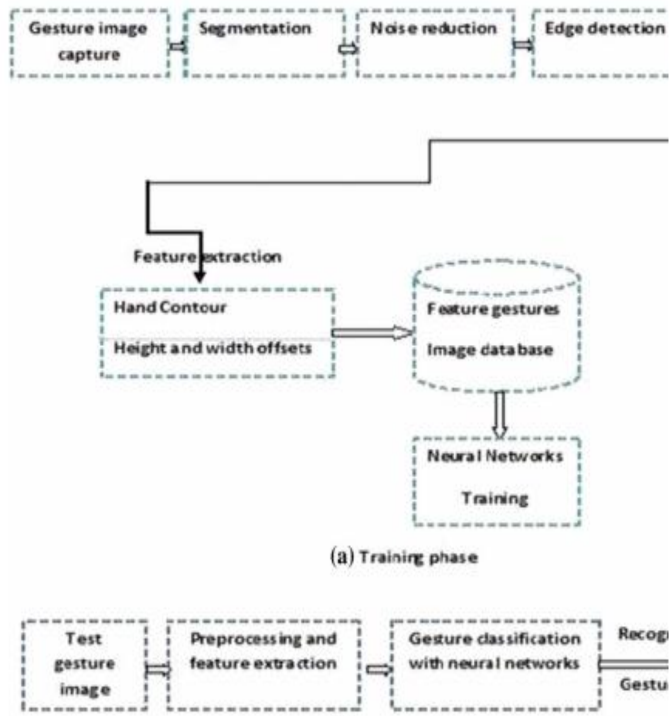
Raw Moments:
Image with pixel intensities I(x, y)

$$M_{ij} = \sum_{X} \sum_{Y} x^i y^i I(x, y)$$

Raw moments of a simple image include:
1) Sum of grey levels or Area (In case of Binary image) : M_{00}
2) Centroid : M_{10}/M_{00} , M_{01}/M_{00}
[3]

E. Recognition
Recognition involve following steps:
1. First the features of the test image are calculated using Hu moments.
2. These features are compared with the training feature set.
3. The algorithm used for classification is KNN (K-nearest neighbor).
4. This algorithm uses neighbors to calculate distance and on the basis of distances it classifies the current record in one of the predefined classes.
5. Euclidean is used for finding the distance by comparison.
Euclidean Distance ((X,Y),(A,B)) = $[(X,A)^2 + (Y,B)^2]^{1/2}$
6.Gesture is classified into the class with which it has minimum distance.
7. K value is selected, which is the number of neighbors taken in account for every calculation.
8. Carefully select the value of K, if the value of K is too small it is sensitive to noise, and if the K value is too large the neighbors might include points that are from other classes. So a normal or medium value of K is selected.
9. One of the limitations of this method is that it will classify the input gesture to at least one of the training class with minimum distance, which results in incorrect classification. So a Threshold is applied.
10. After calculating distance the value is compared with the Threshold.
11. If it passes the threshold it is classified, otherwise it is identified as a new gesture.[4][5]

V. ARCHITECTURE



VII. LITERATURE SURVEY

1. Gesture Recognition: A Survey by SushmitaMitra, Senior Member, IEEE, and Tinku Acharya

In this paper, they provide a survey on gesture recognition with particular emphasis on hand gestures and facial expressions. Applications involving hidden Markov models, particle filtering and condensation,finite-state machines, optical flow, skin color, and connectionist models are discussed in detail. Existing challenges and future research possibilities are also highlighted.

2. Real-Time Hand Tracking and Gesture Recognition System,Nguyen Dang Binh, Enokida Shuichi, Toshiaki Ejima

In this paper, we introduce a hand gesture recognition system to recognize real time gesture in unconstrained environments. The system consists of three modules: real time hand tracking, training gesture and gesture recognition using pseudo two dimensional hidden Markov models (P2-DHMMs).

3. Identifying Fingertips for Human Computer Interaction byR.M. Arunachalam, M. Ashok Gowtham, and R. Aarthi

With the explosive evolution of computer technology, Human-computer interfaces are finding an increased importance in daily life. The use of input devices such as the mouse and keyboard limits the friendliness of the user interaction. Therefore, to increase the interaction with the system, a method is proposed which employs tracking the tip of the fingers using a single camera. It employs a graphical technique involving angle and points. It describes a technique by which the location of fingertip is mapped with that of the monitor. The outcome of this technique in a real time scenario is found to be efficient. Identifying a finger with the help of a fingertip paves way for increased expression of gestures.

4. Finger Detection for Sign Language Recognition by Ravikiran J, Kavi Mahesh, SuhasMahishi, Dheeraj R

This paper introduces an efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet of the American Sign Language. Finger Detection is accomplished based on the concept of Boundary Tracing and Fingertip Detection. The system does not require the hand to be perfectly aligned to the camera or use any special markers or input gloves on the hand.

VIII. REFERENCES

[1] A. Argyros and M. Lourakis, "Vision-based interpretation of hand gestures for remote control of a computer mouse", in Processing of Workshop Computer Human Interaction,Institute of Computer Science, Foundation for Research and Technology,Crete, Greece,vol. 1, pp.4051, July 2006.

[2] Qing Chen, N. D. Georganas, “Hand Gesture Recognition Using Haar-Like Features and a Stochastic Context-Free Grammar”, IEEE transactions on instrumentation and measurement,Ottawa Univ., Ottawa, vol. 57, no. 8,pp.113-117, August 2008.

[3] G. R. S. Murthy and R. S. Jadon, “Hand Gesture Recognition using Neural Networks”, IEEE Transactions,Madhav Institute of Technology and Science, Gwalior, M.P. India, vol. 6, pp.427-431, May 2010.

[4] N.A.Ibrahim and R.Z. Khan, “Survey on Various Gesture Recognition Technologies and Techniques”, International Journal of Computer Applications,A.M.U., Aligarh Uttar Pradesh, India, vol. 50, No. 7, pp1230-1235, May2014.

[5] G. Deepa, et.al “Emotion based Music Retrieval for Gesture Recognition” ,International Journal of Computer Applications,A.M.U., Aligarh Uttar Pradesh, India, vol. 78, no.14, pp.0975 – 8887, September 2013.