Hand Gesture Reorganization using Virtual Canvas

Vanita M. Mestry[1] Rachna P. Hindlekar[2] Dheeraj S. Bari[3] Prof. Aruna Pavate[4]
[4]Project Guide
[1,2,3,4]Department of Computer Science & Engineering
[1,2,3,4]Atharva College of Engineering, University of Mumbai, Mumbai, India

Abstract— Computer vision based hand tracing can be used to interact with machines in a new inventive way. The input components of a normal computer system uses keyboard, mouse and joysticks. Which can be avoided using gesture recognition. Gestures are revelatory body movements involving physical moves of the fingers, hands, arms, head, face, or body with the purpose of: 1) assigning meaningful information or 2) interacting with the environment. They make up one interesting small subspace of possible human motion which tells someone to move or act in a certain way. A gesture may also be seen by the environment as a compact technique for the information to be transfuse elsewhere and subsequently adapted by the receiver. In this paper, we designed a robust real-time hand gesture recognition method. In our method, first step is capturing an image using webcam and then apply pre-processing steps for extracting features. The second stage involves the classification of the features set to the appropriate gesture using the trained classifier that was computed in the setup stage.

Key words: Virtual Canvas, Hand Gesture Reorganization

I. INTRODUCTION

A. What are Gestures?

Gestures can be distinct as the act of the body that is proposed 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 our project, gesture is defined as an expressive movement of body parts which has a some message, to be communicated accurately between a sender and a receiver. A gesture can be categorized into two different categories: dynamic and static. A dynamic gesture is changes does not changes over a period time. 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 clarify all the static and dynamic gestures over a period of time. This multifaceted process is called gesture recognition. Gesture recognition is the progression of recognizing and interpreting a rivulet continuous sequential gesture from the known set of input data.

B. Need

A functioning sign language recognition system could provide an opportunity for the deaf to interact with non-signing people without the need for an interpreter. It could be used to generate speech or text making for the deaf people. Unfortunately, there has not been any system with these capabilities so far. In this project our aim is to develop a system which can classify sign language accurately. To interact with system without getting much closer to it.

C. Applications

Gesture recognition has wide-ranging applications. such as the following:

- Developing aids for the deaf people;
- Enabling very young kids to interact with computers
- Designing approach for forensic identification
- Man machine interface can be made soft and smooth
- Recognizing sign language
- Medically monitoring patients
- Communicating in video conferencing
- Distance learning/tele-teaching assistance Can be used for playing computer games. [1]

II. METHODS

A. Sampling and Quantization

In order to process the image, it must be saved on computer. The image output of most sensors is continuous voltage waveform.

But computer deals with digital pictures or images not with continuous images, thus: continuous images should be converted to digital form.

Continuous Image (in real life) → Digital Image (computer)

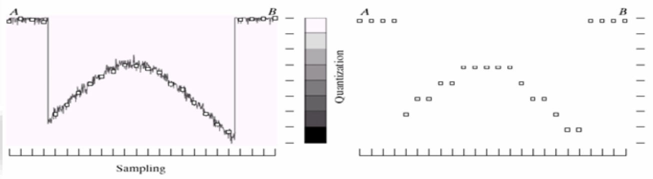

Fig. 1: Sampling Process

An image may be defined as a continuous wave with respect to the x and y coordinates, and also in amplitude. To convert it to the digital form, we have to sample the function in both coordinate and in amplitude. Digitizing the organize values is called sampling. Digitizing the amplitude values is called quantization.

B. Binarization

Binarization is method of binarizingan image by extracting brightness, density as a feature amount from the image. When a pixel is selected in an image, sensitivity is added to and/or subtracted from the value concerning the Y value of the selected value range. Image binarization converts an image of up to 256 gray levels to a black and white image.
A binary image is a digital image that has only two possible values for each pixel. Typically the two colors used for a binary image are black and white though any two colors can be used. The color used for the object(s) in the image is the foreground color while the rest of the image is the background color.

Once the hand is detected, the color of hand is collected from the neighborhood of features mean position. Considering tradeoff between computational cost and accuracy of description, we use a single Gaussian model to describe hand color in HSV color space.[2]The simplest form of segmentation is probably "Otsu thresholding" which assigns pixels to foreground or background based on greyscale intensity. A binary image can be stored in memory as a bitmap, a packed array of bits.

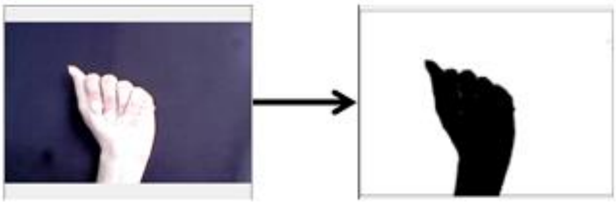

Fig. 2: Given an image shows output of binarization

The simplest form of segmentation is almost certainly "Otsu thresholding" which assigns pixels to foreground or background based on greyscale intensity. A binary image can be stored in memory as a bitmap, a packed array of bits.

C. Matching

One way for finding matching image within a collection of images is to excerpt features from the contest image and all the images in the memory, and then find matching features one by one. size.For a better performing approach we can use Fast Approximate Nearest Neighbor (FLANN) is a technique which is mostly used for the purpose of doing fast nearest neighbor search in large data-sets.

D. SURF (Speed Up Robust Features)

Once image is computated, we match up that image with the dataset images. So here we use Speeded Up Robust Features (SURF) which is one of the best feature-based image matching algorithms. In this we first detect the feature of an image and then describe it. To detect the feature we extract Hessian matrix-based interest points and generates a distribution-based descriptor, and is a scale- and rotation-invariant algorithm. SURF uses Hessian Matrix to select the candidate points in different sizes then it uses Haar wavelet filters and the integral of the image to speed up the filtering operation. It is good at handling images with blurring and rotation. The method is very fast because of the use of an integral image where the value of a pixel (x, y) is the sum of all values in the rectangle defined by the origin and (x, y).

E. FLANN (Fast Approximate Nearest Neighbor)

A SURF descriptor is a 128-dimensional vector normalized to length one. In a typical application, a large number of SURF descriptors extracted from one or many images are stored in a database. So query involves finding the best matched descriptor vector(s) in the database to a SURF descriptor extracted from a query image. A useful data-structure for finding nearest-neighbor queries for image descriptors is the KD-tree, which is a form of balanced binary search tree.

E. Edge Detection

Edges in images are areas with strong intensity contrasts – a jump in intensity from one pixel to the next. Edge detecting an image significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image.

Edge Detection can be use for extracting information about image E.g. Location of objects present in the image, their shape, size, image sharpening and enhancement.

F. Sobel Edge Detection

The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input gray scale image.

The Sobel edge detector uses a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows). A convolution mask is usually much smaller than the actual image. As a result, the mask is slid over the image, manipulating a square of pixels at a time.

$$\underbrace{\begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}}_{Gx} \qquad \underbrace{\begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}}_{Gy}$$

Fig. 3:

Typically, an approximate magnitude is computed using:

$$|G| = |Gx| + |Gy|$$

This is much faster to compute.The angle of orientation of the edge (relative to the pixel grid) giving rise to the spatial gradient is given by:

$$\Theta = \arctan(Gy/Gx)$$

The edge detected image can be obtained from the sobel gradient byusing a threshold value.

the sobel gradient values are lesser than the threshold value then replace it with the threshold value.

III. PROPOSED SYSTEM

We develop our system based on the understanding of the gestures for natural interaction fig.4 shows an overview of our system design.

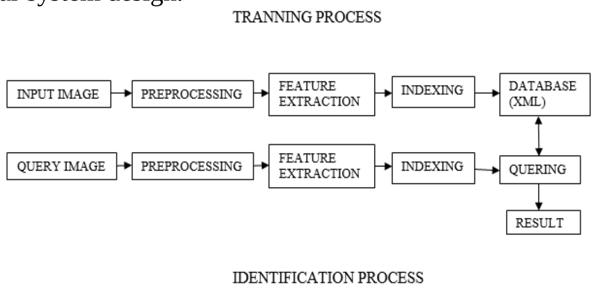

Fig. 4: Proposed System

We have stored some predefined gesture images and their binzarize images too. To achieve accuracy in recognition. Predefine set has 84 set of images into memory.

IV. REQUIREMENTS

Project requires a computer with some particular hardware and software required for both the developer and for the user who uses the system. The hardware needs an essential computer or laptop with a webcam attached whether integral or external. Software included for development of project needs visual studio 2010 or higher and Dot Net Framework which is included in visual studio 2010.

The device types supported would be:

- Pentium 2.8.GHZ or higher microprocessor

- Memory required 512MB RAM
- Standard Keyboard
- PS/2 Mouse
- Basic Webcam (Inbuilt Or External)

V. RESULT AND ANALYSIS

The system is trained on 20 samples for each sign sample of different size and other feature like uniform lighting and background, varying distance is taken into consideration while discarding the other and we tested the system with different lighting, size and distance but the fixed background with samples of 10 peoples. The analysis of this table shows that the accuracy of the system is 81%.

VI. RECOGNITION RATE

The system recital can be evaluated based on its ability to correctly classify samples to their corresponding classes the recognition rate can be defined as the ratio of the number of correctly classified samples to the total number of samples and can be given as:

Recognition rate = (Number of Correctly classified sign/ Total number of signs) * 100%

SIGN	RECOGNIZED SAMPLES	MISCLASSIFIED SAMPLES	RECOGNITION RATE (%)
A	9	1	90%
B	10	0	100%
C	9	1	90%
D	8	2	80%
E	4	6	40%
F	8	2	80%
G	6	4	60%
H	7	3	70%
I	9	1	90%
J	9	1	90%
K	8	2	80%
L	9	1	90%
O	9	1	90%
R	7	3	70%
U	6	4	60%
V	8	2	80%
W	9	1	90%
X	8	2	80%
Y	9	1	90%
Z	8	2	80%

Table 1:

A. Graphical User Interface (Gui)

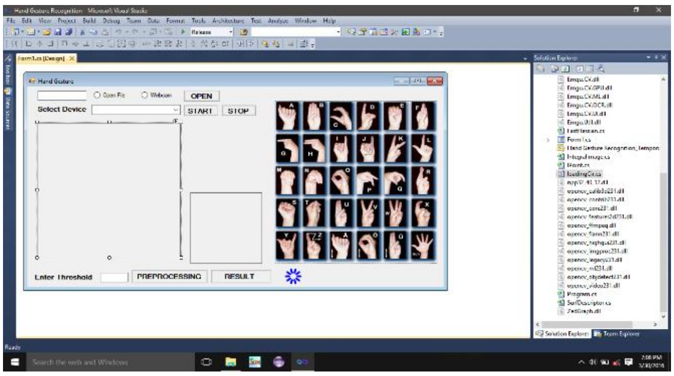

Fig. 5: Home Screen

B. GUI Showing When a Gesture Is Made by User

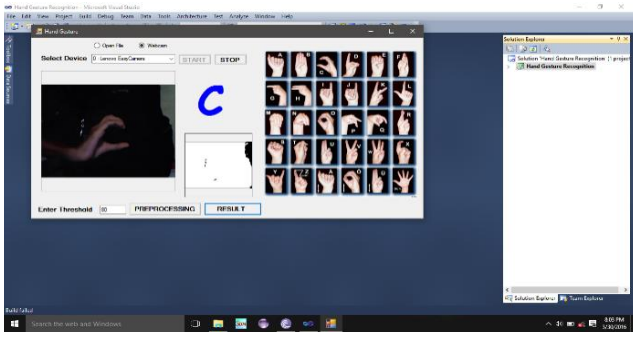

Fig. 6: GUI for Gesture Made

VII. CONCLUSION

The work presented in this project recognizes ASL static signs only. The work can be comprehensive to be able to recognize dynamic signs of ASL. The system deals with images with uniform background, but it can be made background independent. The system can be skilled to other types of images for making application such as games, social site with gesture, media player handling through gestures .it is important to consider increasing the data size, so that it can have more accurate and high performance system.

VIII. FUTURE SCOPE

Adding more gestures, we can handle all My Computer operations like Cut, Copy, Paste and Undo etc.

Integrating our system with voice recognition system we can embed it in ROBOTS.

We can also enhance our system to control PowerPoint application.

We are also able to handle dynamic image processing and event handling accordingly.

REFERENCES

[1] SushmitaMitra, “Gesture Recognition A Survey” Senior Member, IEEE, 3 MAY 2007
[2] Yikai Fang and Kongqiao Wang, “A REAL-TIME HAND GESTURE RECOGNITION METHOD”National Lab of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China
[3] Pranav Mistery, “Sixth Sense”
[4] Digital Image processing by Rafael C. Gonzalez, Richard E. Woods.
[5] Ray Lockton, “Hand Gesture Recognition Using Computer Vision”, Balliol College Oxford University
[6] Sanjay Meena, “A Study on Hand Gesture Recognition Technique”.

