RESEARCH ARTICLE OPEN ACCESS

Handstrokes for Recognition of Alphabets& Digits

Mrudula S Deshmukh,
BE Computer Engineering, Sinhgad College Of Engineering, Pune - 41.

ABSTRACT
Handgestures since ancient times have been used for communication even before speech and language were developed. They have proved to be very effective means of communication. Thus, many new technologies attempt to use hand gestures for developing Human Computer Interfaces. However the practical usage of these interface solutions is limited because of two primary reasons: one, they are either not user friendly, or two, are expensive. Keeping this in mind along with the special requirements of the physically disabled users, a system is proposed which will accept handstrokes as an input from the user and display the corresponding English character [A-Z]/digit [0-9] to the user along with the narration. This system will be helpful for physically disabled users to learn our language, thus enabling them to communicate with us in our language. It can also be used as an innovative way of teaching children.
Keywords - Image Processing, Image Segmentation, Hand Motion Tracking, Feature Extraction, Classification.

I. INTRODUCTION

Human Computer Interaction is recently the area of research where different technologies using variety of techniques are being developed. Some of the interface solutions use speech or keyboard interfaces [1].Recently even multi-touch technology is used. However these technologies have restricted use in everyday environments[2]. Also, the multi-touch technology is useful for people with full finger function. The use of these technologies for people with partial limbs, or for aged people who cannot repeatedly perform the same gestures is restricted[3]. These problems can be overcome by using handstrokes.

Handstrokes are nothing but dynamic gestures. Handstrokes will be given in the form of an input to the system. The handstroke will be in the form a live video stream. The motion traced from the handstroke will be processed by the system. The corresponding alphabet [A-Z]/digit (0-9) will be recognized by the system. This recognized character will be the output which will be displayed and narrated to the user. The narration will help the user to learn the pronunciation of the character.

II. SYSTEM OVERVIEW

The system comprises of three phases –

Fig. 1 System Overview

These phases are explained below.

1. Pre-processing

This is the first phase of the system. The input to this system is the live video stream given by the user in the form of a handstroke. Two operations are performed on the input in this stage. First, segmentation is to be applied to the video input so as to obtain only the hand of the user eliminating the background. Some of the segmentation techniques which can be used are edge detection, histogram, thresholding, skin color segmentation [4]. In edge detection, points or edges of an image are identified depending on the discontinuities in the brightness of the points. This change in brightness is useful in detecting the boundaries of objects in the image. Edge detection helps in reducing the size of data to be processed further as it removes the less relevant data leading to reduction in the processing time. Histogram method is used to simply store data in a given format. It can be used to obtain a view of the underlying data. Thresholding technique can be used to reduce processing data size by eliminating the image pixels which are above or below the specified threshold. Skin color Segmentation is the most commonly used segmentation technique. This technique is very useful in face detection and hand detection as it separates the skin color pixels from the non-skin color pixels which may avoid the exhaustive searches for face & hand in an image thus resulting in greater utility & robustness. The accuracy of skin color segmentation depends on the selection of color space. Normally digital images use RGB color space. But YCbCr color space consists of a wide range of colours that can easily

classify skin & non-skin color pixels irrespective of illumination conditions. Thus, there is a need to convert these images into YCbCr color space.Transformation of RGB to YCrCb is as follows-

$Y = 0.299R + 0.587G + 0.114B$

$Cr = R - Y$

$Cb = B – Y$

Ranges for Y, Cr, Cb values are

$Y>80$,

$85<Cb<135$,

$135>Cr>180$

Where Y, Cb, Cr = [0,255].

Fig. 2 Skin color Segmentation

The second operation we do in this phase is centroid calculation. We calculate the centroid of the hand so that we can trace the hand motion path from the input.

2. Feature Extraction

The input to this stage is the segmented image obtained after pre-processing stage. This image is then mapped to a standard plane of pre-defined size so as to get fixed dimension images for feature extraction. This process is called image normalization. Features are calculated from the normalized image which can be used as the basis of classification. Feature is simply an n-dimensional vector containing the numerical information of an object. Features extracted from an image vary largely depending upon the application for which it is used.The objective of feature extraction is to map input pattern onto points in a feature vector. For the proposed system we can use features like start & end point of the handstroke, mean, standard deviation, eccentricity, no. of horizontal & vertical lines, etc.

3. Classification

Classification maps the points in the feature vector to a predefined class label. In other words, the input is classified depending on the values of its features and the values of features of sample images stored in the database. There are different classifiers viz. Naïve Bayesian Classifier, Decision tree classifier, k nearest neighbour, neural networks, etc. [5] NB models are commonly used in machine learning applications, as they are simple. In this each attribute to contributes towards the final decision equally and is independent of the other attributes. The efficiency of NB is greater due to its simplicity & thus it can be used in various domains. Decision Tree Classifiers are easy to understand & easy to implement. Also, they perform well on continuous as well on discreet classification is done without much computation. Neural networks are as complex as human nervous system. Yet they easily represent complex relationships between input & output. Neural network is not so easy to understand from visual representation & is also difficult to explain.

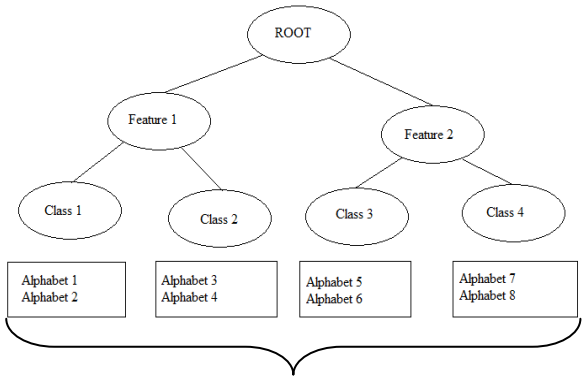

Fig 3. Classification using combination of Decision Trees& Naïve Bayes Classifier

For the proposed system, a combination of two classifiers can be used. Decision tree classifier can be used to classify the input until it gets a class label. Further classification can be done using Naïve Bayes classifier which will check the probability of the input resembling with the characters belonging to that class. The sample data which will have the highest probability of being the same as the specified input will be the resultant character.

III. OUTPUT

The output will be obtained from the classification phase. The corresponding alphabet [A-Z]/digit [0-9] will then be displayed to the user. This will help the user to improve the hand motion while giving the handstroke next time. The output displayed will give the correct way to perform the handstroke.

The displayed output will also be narrated to the user. This will enable the user to pronounce the alphabet/digit. It will specially be helpful for visually impaired people as they will know how to write the characters & how to pronounce it, thus helping in communication.

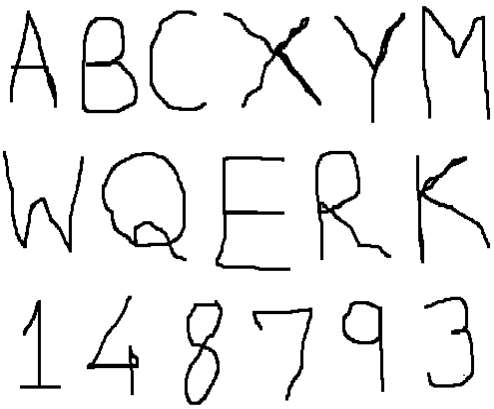

Fig. 3 Alphabets & digits using handstrokes

III. FUTURE WORK

This system can be extended to forming words and sentences so that sign language will not be an obstacle in the communication**.**This system can also be implemented for other languages as well.

IV. CONCLUSION

The proposed system consists of three phases- one, pre-processing which will apply segmentation technique to the image & calculate the centroid. From the centroids of all images obtained from the same video motion path can be obtained. In the second phase normalization is done & features are extracted and stored. In the third & the last stage, classification technique is applied to obtain the output which is then displayed & narrated to the user. The system is cost effective as no other hardware is required. Also it will help reduce the communication gap between us and physically challenged people.

REFERENCES

[1.] V.Shishir Reddy, V.Raghuveer, J.Vamsi Krishna, K.Chandralohit, Finger Gesture based Tablet Interface, IEEE 2012.

[2.] MeenakhiPawar, Hand Gesture based Interface for Aiding Visually impaired, IEEE 2012.

[3.] Markus Modzelewski, Esteban Bayro Kaiser, Hand Gesture Interface For Visually Impaired & Blind People, IEEE 2012.

[4.] Tarek M. Mahmoud, A New Fast Skin Color Detection Technique", World Academy of Science, Engineering and Technology 2008.

[5.] Daniela XHEMALI1, Christopher J. HINDE2 and Roger G. STONE3, Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages, ICJSI,vol 4,no 1,2009.

[6.] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai ,SKIN SEGMENTATION USING COLOR AND EDGE INFORMATION, IEEE 2003.