Indian Sign Language Converter System Using An Android App.

Pranali Loke
Department of Computer
Vidyalankar Institute of Technology
Wadala(E), Mumbai, India.
Email: pranali.loke18@gmail.com

Juilee Paranjpe
Department of Computer
Vidyalankar Institute of Technology
Wadala(E), Mumbai, India.
Email: juileeparanjpe11@gmail.com

Sayli Bhabal
Department of Computer
Vidyalankar Institute of Technology
Wadala(E), Mumbai, India.
Email: saylibhabal26@gmail.com

Ketan Kanere
Department of Computer
Vidyalankar Institute of Technology
Wadala(E), Mumbai, India.
Email: ketan.kanere9@gmail.com

***Abstract*—In this paper, we introduce a sign language converter system using hand gesture recognition feature to recognize the gestures in Indian sign language and convert them to a natural language. Sign language is a large set of hand gestures that are used to communicate with the hearing impaired and mute. The proposed system uses Hue, Saturation, Intensity (HSV) color model for hand tracking and segmentation. We have used supervised learning for training the neural network which is responsible for data classification. An android application is used to capture images of hand gestures which are sent to a web hosting server, from where they are given as input to the neural network in MATLAB for pattern recognition. The hand gesture is mapped to its equivalent in natural language in MATLAB and the converted text is sent back to the users device. It is an easy-to-use and inexpensive approach to recognize single handed as well as double handed gestures accurately. This system will facilitate communication for millions of deaf and mute people and aid in communicating with people who dont understand sign language.**

***Index Terms*—Hand gesture recognition, HSV color model, neural network, supervised learning, pattern recognition, hand tracking and segmentation.**

I. Introduction

Gesture recognition is an interpretation of gestures using mathematical algorithms. It comprises of any gestures made by face or hands. This technique has a wide range of applications and uses. Gesture recognition can be used to control devices merely through gestures without any physical link with the actual machine. Using gesture recognition, a person can point at the computer screen or at a mobile screen and use sign language to select and use different applications in the device. Sign language used by deaf and mute people can be interpreted using this technique and converted to text thus enabling and aiding better communication between the deaf and mute and people interacting with them.

Image processing is usually used for gesture recognition as it provides features like pattern recognition, texture understanding, content-based retrieval, compression and many more. The HSV (Hue, Saturation and Value) color model in image processing is used skin detection. The openCV library in Matlab is used to convert the RGB color space to HSV one.

An android application with a user-friendly graphical interface that captures gestures and interprets them to text in a natural language will enable the disabled people to connect and interact with the masses without the need of a controller. All input images are captured by camera present in an android device. The captured images are then sent to the Matlab via an online server and are converted in binary form of size defined so that it takes less time and memory space during pattern recognition. The converted text is then sent back to the users device.

Rest of the paper is organized as follows. Section II covers the existing technologies and need for a new technology for hand gesture recognition. Section III talks about the proposed system and discusses the architecture used for implementing the system. Section IV discusses implementation of android application to send the images to the server. Section V talks about image processing and data classification using neural network in MATLAB. Section VI cover sending data to the neural network from a smart phone via web hosting server and sending the converted text back to users device. The paper ends with conclusion and future scope.

II. LITERATURE SURVEY

Applications of image processing in diverse areas are growing rapidly. One of the applications of image processing is hand gesture recognition using pattern recognition algorithms. For tracking and recognizing hand gestures different approaches are currently used.

Previously, systems have been developed to convert sign languages like the American sign language (ASL) to natural languages. However, no such system exists for converting

Indian sign language (ISL) to natural language. Two such existing technologies for recognizing hand gestures have been discussed below. The two main approaches used are Data Glove based technology and Vision-based approach.

A. Data Glove Technology

Data glove is an electronic device that senses the movements of the hand and ,also, the fingers individually and sends these movements to the computer in the form of analog or digital signals. These digital signals are then mapped to the task to be performed in the virtual environment[5]. On this glove, various sensors are placed to detect the global position and relative configurations of the hand[1]. Data glove is used in fields like 3D animation, 3D sketching and also in the medical world. It is basically used to interact with the 3D models[5]. It is an expensive device because of the multiple sensors used. Because of the shortcoming of data glove based systems, vision-based systems are used, which are wireless and the only thing needed is one or multiple cameras. Using motion sensing input devices such as Microsoft Kinect is also another method but this device is also expensive.

B. Vision-based Techniques

There are several vision-based gesture recognition techniques commonly used for static and dynamic gesture recognition. We have discussed a few of those below:

1) Support Vector Machine: Support Vector Machine(SVM) is a supervised machine learning algorithm used in pattern recognition to analyse data and recognize patterns. It is used for classification and regression analysis. It builds a training model and predicts the category of an unknown sample. The objective of SVM modelling is to construct a hyperplane as the decision plane, which separates the positive and the negative binary classes with the largest margin. SVM is not just limited to 2 class classification problem; it can also be used for multi-class classification problem using pair-wise top-down, bottom-up classification methods[6].

2) K-Nearest Neighbours(K-NN): K-nearest neighbour classifier is a machine learning algorithm which relies on the distance between feature vectors. It classifies an unknown data object by finding the most common class among the k-nearest examples in the feature space. The algorithm stores feature vectors and labels of the training images. A major disadvantage of K-NN algorithm is that it uses all the features equally while computing for similarities; which can lead to classification errors ,especially, when a small subset of features is used for classification[7].

C. Convolutional Neural Network

Another important technique used for hand gesture recognition and classification is use of convolutional neural networks. Convolutional neural networks are similar to normal neural networks with few major changes. Convolutional network has raw image pixels to one end and class scores to other end. Convolutional network architectures make the explicit assumption that the inputs are images, which allows to encode certain properties into the architecture. Hence, the implementation of the forward function becomes efficient and this reduces the amount of parameters in the network greatly. Also the neurons in convolutional neural network are arranged in 3 dimensions: width, height and depth. But there is one disadvantage associated with this technique that it involves high computational cost hence to train this neural network computer with good GPU is required[4].

III. PROPOSED SYSTEM

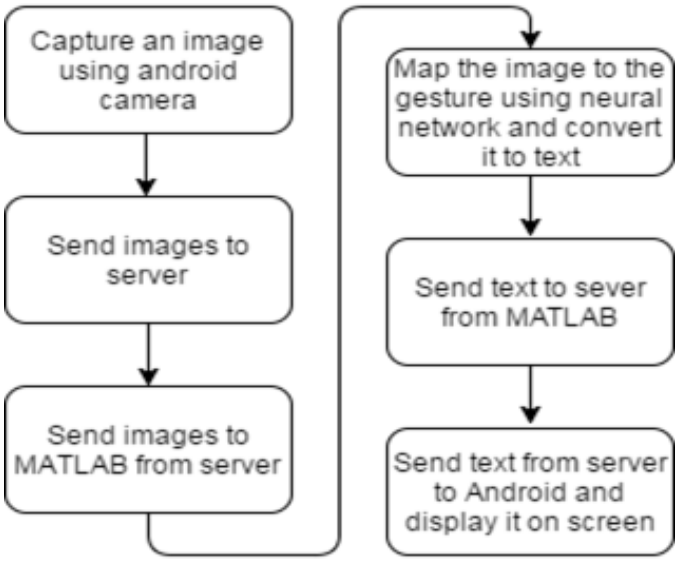


Fig. 1: Flowchart of Proposed System

We propose to implement the sign language converter system as an android application. This application will be able to detect gestures in the Indian Sign Language (ISL). The user will have to capture images by positioning camera it in front of the person who is making the gestures using this app. These images will then be sent to the server. The server will send these images to MATLAB where gesture recognition using neural network will take place. The neural network is trained to recognize ISL gestures. The image will be mapped to the corresponding ISL gesture and converted to text. The converted text will be sent back to android application via server and the output will be displayed in the users smartphone.

For classification of gestures many applications make use of regression analysis. But regression analysis is based on assumption that the cause and effect relationship between the variables remains unchanged. This might not always be the case and hence estimation of the variable values done on the basis of the regression equation may lead to erroneous results. Hence, in this project, we present a vision-based system which uses the built-in camera provided in android device.

For classification purpose, we are making use of neural network trained using Scaled Conjugate Gradient Back-propagation algorithm. For feature extraction of images and classification using neural network, we will be making use of MATHWORKs Matlab. We collected a vast and varied dataset for training purpose of the neural network. This system will allow user to capture images from remote locations and then these images will be sent to a server which will forward the images to MATLAB for processing where a trained neural network will be present. The response will be sent back to server and server will send back its response in the form of its textual meaning to the android device.

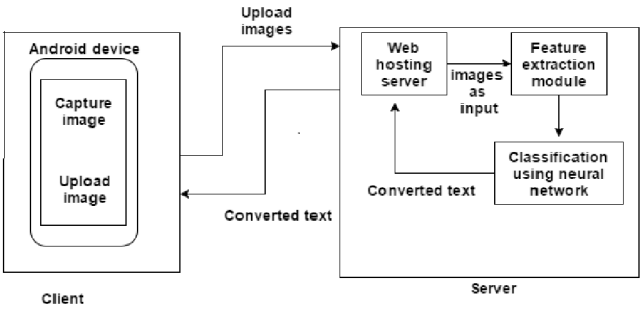


Fig. 2: System Architecture

This system will be implemented using client-server architecture. On the client machine, there will be an android application which will be used to capture images of the gesture. Once the image is captured, it will be uploaded on the server using the android application. Upon reaching the server, feature extraction of the user-uploaded image will take place in the feature extraction module. The features extracted are nothing but the angles between different parts of the edges of the gesture that are calculated. These features in the form of angle-values are sent to neural network, trained for classifying different gestures to their textual meaning. This system will be implemented using client-server architecture[9].

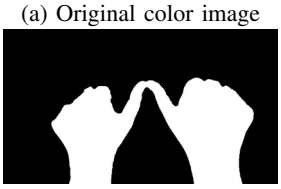
IV. ANDROID IMPLEMENTATION

The android app for the system was developed using Android Studio. In this system, we are using okhttp3, an HTTP library, that makes networking for Android apps easier and faster. It facilitates automatic scheduling of network requests. It integrates easily with any protocol and provides support for raw strings, images, and JSON. For accessing camera in android device, camera2 class provided by Android Studio was used. The android app will invoke the camera in-built in the android device. For this, various permissions like camera access, read and write external storage, read phone state and internet permission were given. The captured images were then sent to the server. These images were then provided as input to Matlab for pattern recognition and classification.

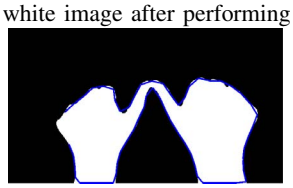
V. DATA CLASSIFICATION USING NEURAL NETWORK IN MATLAB

We perform two major operations on an image to extract information from the image which acts as the input to the neural network. These operations are explained briefly below:

(a) Original color image



(b) Black and white image after performing preprocessing



(c) Tokenized image

Fig. 3: Preprocessing and Tokenizing of alphabet A of ISL.

1)Preprocessing: Here we convert colored image to a b/w image using HSV skin color extraction technique. In this technique, HSV color space is used instead of RGB color space[10]. Hue represents the dominant color, saturation indicates the intensity of the color and value represents the brightness[11]. This technique distinguishes the skin pixel and majority of the gesture is retained and unnecessary background and noise is eliminated. Figure 3b shows a preprocessed image.
2)Tokenizing: We use Sobel edge detection technique to trace the boundary . Sobel edge detection technique perform a 2D spatial gradient quantity on a image and points with high spatial frequency are highlighted[12]. We then sample the points on the edge of the gestures and obtain the angle between the vectors formed by consecutive points. This operation can also be called as feature extraction[2]. Figure 3c shows a tokenized image.

Using complete image as an input to the neural network would be inefficient; hence, it is advisable to extract features from an image and then use them as inputs as they sufficiently represent a particular image and are smaller in dimensions. Detailed explanation of each operation is given below:

Algorithm: Preprocessing.
Input: Image with RGB color map which contains gesture.
Output: Binary image where gesture pixel is 1 and rest is 0.

Steps :
1. Resize input image rgbimg to fixed predefined size: [rows, cols].
2. Convert RGB color channels of rgbimg to HSV channel denoted by hsvimg.
3. Set HSV thresholds for skin color detection[3].
$[0 <= H <= 50]$
$[0.05 <= S <= 0.8]$
$[0.25 <= V <= 1]$
4. Apply these thresholds on hsvimg to filter out only skin colored component (i.e., pixel value 1 if it falls in the threshold).
5. Eliminate blobs that are less than permissible size: [150 px] and store as noisereducedimg.
6. Smooth out the contours of the white blobs in noisereducedimg using structuring element as disk and call it smoothing.
7. Fill all closed blobs which have holes in them.
8. Retain only the blob having the largest filled area filledimg and call the new image gesturebwimg.
9. return gesturebwim.

VI. SENDING TEXT FROM MATLAB TO ANDROID

The mapping of image to a gesture will give us the text corresponding to that gesture. This text will be sent to android application via a server. For this, MATLAB will use RESTful API. REST means representational state transfer. It is a common architectural style for web services. RESTful APIs or interfaces provide standard HTTP methods such as GET, PUT, POST, or DELETE. RESTful API will be used for request-response relationship with the server and JSON to represent MATLAB datatypes.

JSON object is used to send text from server to the android device. JSON stands for JavaScript Object Notation and is used to exchange data to/from a web server. There are four different classes provided by Android to manipulate JSON data. We have used JSONObject class to receive text from the server[8].

VII. CONCLUSION AND FUTURE SCOPE

In this paper, a new approach was proposed for static gesture recognition on the resource constrained devices like smartphones. The focus of this research is applying HSV color model for feature extraction. From the study of skin color segmentation methods, we conclude that there is a problem of hand segmentation using normal RGB camera in skin color background as well as change in the light conditions. Proposed sign language recognition system recognizes the gestures in constrained environment like dark background. The performance of the algorithm used can decrease due to varying lighting conditions, lighter background and background noise. So, further work should focus on hand segmentation method on the resource constrained devices with varying light conditions and skin color background. This system can be further developed to recognize gestures in real time and in video format. Many other gestures of the sign language can also be made a part of the database.

REFERENCES

[1] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat. *Vision Based Hand Gesture Recognition*

[2] Jagdish L. Raheja, A. Singhal*, Sadab*,Ankit Chaudhary *Android based Portable Hand Sign Recognition System.*

[3] V. A. OLIVEIRA, A. CONCI. *Skin Detection using HSV color space*

[4] CS231n Convolutional Neural Networks for Visual Recognition [Online]. Available: http://cs231n.github.io/convolutional-networks/ [Accessed: 23-Mar-2017]

[5] Piyush Kumar, Jyoti Verma and Shitala Prasad *Hand Data Glove: A Wearable Real-Time Device for Human-Computer Interaction.*

[6] Mrs.A.R.Patil, Dr.S.S.Subbaraman *A Review On Vision Based Hand Gesture Recognition Approach Using Support Vector Machines.*

[7] Jinho Kim, Byung-Soo Kim, Silvio Savarese *Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines.*

[8] Android - JSON Parser [Online]. Available: https://www.tutorialspoint.com/android/android_json_parser.htm/ [Accessed: 25-Mar-2017]

[9] Pieter Vermeulen, Todd F. Mozer*Client/server architecture for text-to-speech synthesis .*

[10] Michael W. Schwarz, William B. Cowan, John C. Beatty*An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models.*

[11] Gururaj P. Surampalli, Dayanand J, Dhananjay M *An Analysis of skin pixel detection using different skin color extraction techniques.*

[12] Muthukrishnan R., M.Radha *Edge detection techniques for image segmentation.*