CONVERSION OF GESTURES TO VOICE AND TEXT MESSAGE IN REGIONAL LANGUAGE

V. Sathya Preiya, Associate Professor,
S.Menaka, P.Preethi,Sheshadri Roshni Sruthi,
Department of Computer Science Engineering,
Panimalar Engineering College, Chennai, Tamilnadu, India.

ABSTRACT:

Speech and text is the main medium for human communication. A person needs vision to access the information in a text. However those who have poor vision can gather information from voice. This paper proposes a camera based assistive text reading to help visually impaired person in reading the text present on the captured image. The faces can also be detected when a person enter into the frame by the mode control. The proposed idea involves text extraction from scanned image using Tesseract Optical Character Recognition (OCR) and converting the text to speech by e-Speak tool, a process which makes visually impaired persons to read the text. This is a prototype for blind people to recognize the products in real world by extracting the text on image and converting it into speech. Computer vision is one of the emerging technologies that can be used to aid visually impaired people for navigation (both indoor and outdoor), accessing printed material, etc. This paper describes an approach to extract and recognize text from scene images effectively using computer vision technology and to convert recognized text into speech so that it can be incorporated with hardware to develop Electronic travel aid for visually impaired people in future.

KEYWORDS: hand gesture recognition, voice conversion, gesture to speech, speech to gesture conversion.

I .INTRODUCTION:

Recent developments in computer software andrelated hardware technology have provided a value added service to the users. In everyday life, physical gestures are a powerful means of communication. They can economically convey a rich set of facts and feelings. For example, waving one's hand from side to side can mean anything from a "happy goodbye" to "caution". Use of the full potential of physical gesture is also something that most human computer dialogues lack.The task of hand gesture recognition is one the important and elemental problem in computer vision. With recent advances in information technology and media, automated human interactions systems are build which involve hand processing task like hand detection, hand recognition and hand tracking.This prompted my interest so I planned to make a software system that could recognize human gestures through computer vision, which is a sub field of artificial intelligence. The purpose of my software throughcomputer vision was to program a computer to "understand" a scene or features in an image.A first step in any hand processing system is to detect and localize hand in an image. The hand detection task was however challenging because of variability in the pose, orientation, location and scale. Also different lighting conditions add furthervariability.

LITERATURE SURVEY:

Hand Gesture Recognition and Voice Conversion for Deaf and Dumb:

Sign language plays a major role for dump people to communicate with normal people. It is very difficult for mute people to convey their message to normal people. Since normal people are not trained on hand sign language. In emergency time conveying their message is very difficult. So the solution for this problem is to convert the sign language into human hearing voice. There are two major techniques available to detect hand motion or gesture such as vision and non-vision technique and convert the detected information into voice through raspberry pi. In vision based technique camera will be used for gesture detection and non-vision based technique sensors are used. In this project non-vision based technique will be used. Most of the dumb people are deaf also. So the normal people's voice can be converted into their sign language. In an emergency situation the message will automatically send to their relation or friends.

Design and implementation of a sign-to-speech/text system for deaf and dumb people:

This paper presents an approach for designing and implementing a smart glove for deaf and dumb people. There have been several researches done in order to find an easier way for non-vocal people to communicate with vocal people and express themselves to the hearing world. Developments have been made in sign language but mainly in American Sign Language. This research aims to develop a sign to Arabic language translator based on smart glove interfaced wirelessly with microcontroller and text/voice presenting devices. An approach has been developed and programmed to display Arabic text. The whole system has been implemented, programmed, cased and tested with very good results.

Gesture Aided Speech for Deaf and Mute:

Speech impaired people make use of sign language to communicate along with normal people. Common people also face difficulties to understand the gesture language. In order to minimize these real time issues, an attempt has been made to develop a system which consists of Flex sensors attached with the Data gloves along all fingers. These will be wearable gloves converts hand gestures into an audio output to interpret the expression. This reduces the communication gap that exists between mute and ordinary people. The proposed system plays corresponding recorded voice on an android phone connected to the hardware system via Bluetooth Module as an interpretation of the gesture. It is

achieved by integrating flex sensor and 8051 microcontroller. The phone will use a text to speech converter to produce the output of corresponding gesture. This system offers high reliability and fast response as microcontroller works at a speed of 12 MHz. Processing time of microcontroller is much less than the time taken for a human to change from one gesture to another, the system works efficiently in real time scenario hence making the system faster.

Implementation of gesture based voice and language translator for dumb people:

Dumb persons communicate through gestures which are not understood by the majority of people. Gesture is a movement of part of the body, especially a hand or the head, to express an idea or meaning. This paper proposes a system that converts gestures given by the user in the form of English alphabets into corresponding voice and translates this English voice output into any other Microsoft supported languages. The system consists of MPU6050 for sensing gesture movement, Raspberry pi for processing, three button Keypad and speaker. It is implemented by using trajectory recognition algorithm for recognizing alphabets. Raspberry pi generates voice output for the text in multiple languages using voice RSS and Microsoft translator. When tested, the system recognized A-Z alphabets and generated voice output based on the gestures in multiple languages.

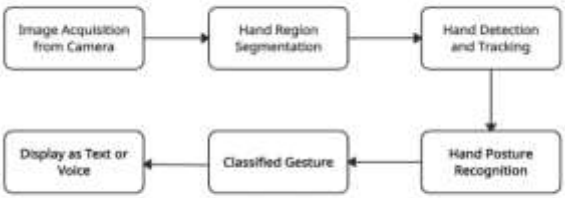
EXISTING SYSTEM:

In recent decades, due to computer software and hardware technologies of continuous innovation and breakthrough, the social life and information technology have a very close relationship in the twenty-first century. In the future, especially the interfaces of consumer electronics products (e.g. smart phones, games and infota1inment systems) will have more and more functions and be complex. How to develop a convenienthuman-machineInterface (HumanMachine Interaction/Interface, HMI) for each consumer electronics product has become an important issue. The traditional electronic input devices, such as mouse, keyboard, and joystick are still the most common interaction way. However, it does not mean that these devices are the most convenient and natural input devices for most users. Since ancient times, gestures are a major way for communication and interaction between people. People can easily express the idea by gestures before the invention of language. Nowadays, gestures still are naturally used by many people and especially are the most major and nature interaction way for deaf people [1]. In recent years, the gesture control technique has become a new developmental trend for many humanbased electronics products, such as computers, televisions, and games. This technique let people can control these classifiers selection are a major issue in most researches. The third stage is to analyze sequential gestures to identify users' instructs or behaviors.

PROPOSED SYSTEM:

Most gesture recognition methods usually contain three major stages. The first stage is the object detection. The target of this stage is to detect hand objects in the digital images or videos. Many environment and image problems are needed to solve at this stage to ensure that the hand contours or regions can be extracted precisely to enhance the recognition accuracy. Common image problems contain unstable brightness, noise, poor resolution and contrast. The better environment and camera devices can effectively improve these problems. However, it is hard to control when the gesture recognition system is working in the real environment or is become a product. Hence, the image processing method is a better solution to solve these image problems to construct an adaptive and robust gesture recognition system. The second stage is object recognition. The detected hand objects are recognized to identify the gestures. At this stage, differentiated features and effective.

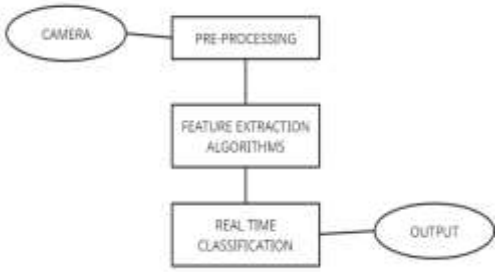
BLOCK DIAGRAM:



SYSTEM IMPLEMENTATION:

Hand gesture recognition system can be dividedinto following modules :

- Pre-processing
- Featureextraction of the processedimage
- Real timeclassification



Value measure intensity or brightness. This is well enough to choose single colour but it ignores complexity of colour appearance. It trade off computation speed mean computationally expensive and perceptual relevance.

a) Pre Processing

Like many other pattern recognition tasks, pre-processing is necessary for enhancing robustness and recognition accuracy.

The pre-processing prepares the image sequence for the recognition, so before calculating the diagonal Sum and other algorithms, a pre-processing step is performed to get the appropriate image, which is required for real time classification. So it consists of some steps. The net effect of this processing is to extract the hand only from the given input because once the hand is detected from the given input it can be recognized easily. So pre- processing step mainly consists of following tasks:

- SkinModeling.
- Removal of Background.
- Conversion from RGB tobinary.
- HandDetectionmodules.

SkinModelling:

There are numerous method used for skin detection such as RGB (Red, Green, Blue), YCbCr (Luminance Chrominance) and HSV (Hue, Saturation, Value).

- **RGB:**

RGB is a 3D color space pixel where each pixel has combination of three colors Red, Green and Blue at specific location. This technique widely used in image processing for identifying skinregion.

YCbCr (LuminanceChrominance):

This color space is used in digital video color information represent two color Cb and Cr. Cb is difference between Blue and Cr is difference between Red component references of value. This is basically RGB transformation to YCbCr for separation of luminance and chrominance for color modelling.

HSV (Hue, Saturation andValue):

In HSV, Hue detect dominant color and Saturation define colourfulness whilst

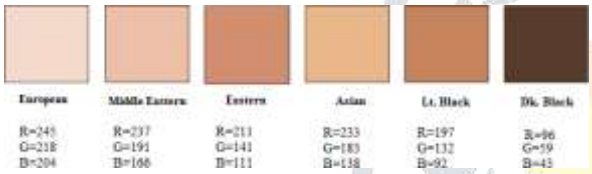
My approach for this thesis is to work with RGB to binarization techniques to Explicitly Defined skin Region.

Skin Detection:

The skin color detection is one of important goal in hand gesture recognition. Skin color detection decision rules which we have to build that will discriminate between skin portion and non-skin portion pixels. This is accomplished usually by metric introduction, which measure distance of the pixel color. This metric type is knows as skin modelling.

Explicitly Defined SkinRegion:

Following are some common ethnic skin groups and there RGB color space:



Different Ethnic Group Skin Patches

To build a skin classifier is to define explicitly through a number of rules the boundaries of skin color cluster in some color space. The advantage of this method is the simplicity of skin detection rules that leads to the construction of very rapid classifier.

In this classifier threshold defined to maximize the chance for recognizing the skin region for each color. If we see in that Red color in every skin sample is greater than 95, Green is greater than 40 and Blue is greater than 20 in**.** So threshold can make this classifier easily detect almost all kind of skin.

This is one of the easiest methods as it explicitly defines skin-color boundaries in different color spaces. Different ranges of thresholds are defined according to each color space components in as the image pixels that fall between the predefined ranges are considered as skin pixels. The advantage of this method is obviously the simplicity which normally avoids of attempting too complex rules to prevent over fitting data. However, it is important to select good color space and suitable decision rules to achieve high recognition rate with this method.

Removal of Background:

I have found that background greatly affects the results of hand detection that's why I have decided to remove it. For this I have written our own code in spite of using any built-in ones.

BeforeAfter

Removal of Background

a) **Conversion from RGB toBinary:**

All algorithms accept an input in RGB form and then convert it into binary format in order to provide ease in recognizing any gesture and also retaining the luminance factor in an image.

b) **Handdetection:**

Image could have more than one skin area but we required only hand for further process. For this I choose criteria image labeling which is following:

c) **Labeling:**

To define how many skin regions that we have in image is by labelling all skin regions. Label is basically an integer value have 8 connecting objects in order to label all skin area pixel. If object had label then mark current pixel with label if not then use new label with new integer value. After counting all labelled region (segmented image) I sort all them into ascending order with maximum value and choose the area have maximum value which I interested because I assume that hand region in bigger part of image. To separate that region which looked for, create new image that have one in positions where the label occurs and others set tozero.

BeforeAfter

Labeling Skin Region

MODULES:

- Dataset collection.
- Dataset pre-processing.
- Machine Learning Training.
- Capture Real time video.
- Eliminate Background.
- Identify Gesture.

Dataset Collection:

Dataset collection involves collecting or creating test dataset for gesture and respective terminology. This is very exhaustive task as we need to collect more samples for gesture. We collect dataset from kaggle website for the project in scope.

Dataset Pre-Processing:

Dataset preprocessing involves removing noise in images and converting to numpy arrays.

Machine Learning:

Machine learning involves training and classification using

CNN algorithms.

Capture Real Time Video & Gesture Detection:
We use opencv to capture and process video. Here we capture video and pre-process frames and fit in model to get the gesture

Background Elimination:
We will find the contours to detect edges of the main object and create a mask with numpy zeros for the background and then combine the mask and the image using bitwise and operator.

RESULT:

In result, after execution of the codecamera will be capturing the gesture and it will eliminate backgroundthen output will be shared in voice and text.

CONCLUSION:

The aim of this paper is to build a system to achieve real time gesture recognition with Indian Sign Language alphabets as the use case. In order to do so I created multiple versions of datasets, trained corresponding models with high accuracy and different methods based on camera and background settings. The project has resulted in a pipeline which can be used to train any gesture recognition application as only the dataset has to be changed and other steps remain the same.

REFERENCES:

1. J.Jenkinwinston (96207106036), M.Maria Gnanam (96207106056), R.Ramasamy (96207106306), Anna University ofTechnology, Tirunelveli : Hand Gesture Recognitionsystem Using HaarWavelet.
2. Laura Dipietro, Angelo M. Sabatini, *Senior Member, IEEE,* and Paolo Dario, *Fellow, IEEE,* A Survey of Glove-Based Systems and TheirApplications.
3. Kay M. Stanney HANDBOOK OF VIRTUAL ENVIRONMENTS Design, Implementation, and Applications, Gesture Recognition Chapter #10 byMatthew Turk
4. Daniel Thalman, Gesture Recognition Motion Capture, Motion Retargeting, and ActionRecognition
5. Hatice Gunes, Massimo Piccardi, Tony Ja, 2007, Research School of Information Sciences and Engineering Australian National University Face and Body Gesture Recognition for a Vision-Based Multimodal Analyzer.
6. J. Heinzm ann and A. Zelinsky Robust Real - Time Face Tracking and Gesture Recognition
7. Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva, Moscow State University A Survey on Pixel-Based Skin Color Detection Techniques.
8. TEY YI CHI, Universiti Teknologi Malaysi, FUZZY SKINDETECTION.
9. Robyn Ball and Philippe Tissot, Texas A&M University, Demonstration of Artificial Neural Network in Matlab.
10. Howard Demuth, Mark Beale, Neural NetworkToolbox.
11. Ari Y. Benbasat and Joseph A. Paradiso in MIT Media Laboratory, Cambridge An Inertial Measurement Framework for Gesture Recognition andApplications.
12. Peggy Gerardin, Color Imaging Course Winter Semester 2002-2003, Color Image Segmentation.
13. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum, March 2003, A procedure for developing intuitive and ergonomic gesture interfaces for man- machineinteraction.
14. Oğuz ÖZÜN1, Ö. Faruk ÖZER2, C. Öncel TÜZEL1, Volkan ATALAY, A. Enis ÇETİN2, Dept. of Computer Engineering, Middle East Technical University, Ankara, Turkey, Dept. of Electrical Engineering, Bilkent University, Ankara,Turkey

Faculty of Engineering, Sabanci University, Istanbul, Turkey, Vision based single stroke character recognition for wearable computing.

15. Melinda M. Cerney1, Judy M. Vance, March 28, 2005, Gesture Recognition in Virtual Environments: A Review and Framework for FutureDevelopment.
16. Ashutosh Saxena, Aditya Awasthi, Vaibhav Vaish, and SANKET: Interprets your HandGestures.
17. Yoichi Sato, Institute of Industrial Science, the University of Tokyo, Makiko Saito Hideki Koike, Graduate School of Information Systems, University of Electro- Communications, Tokyo Real-Time Input of 3D Pose and Gestures of a User's Hand and Its Applications forHCI.
18. Mohammad Al-aqrabawi Fangfang Du, 11th March‟ 2000, Human Skin Detection Using ColorSegmentation.
19. Manuel Cabido Lopes Jos´e Santos-Victor Instituto de Sistemas e Rob´otica, Institution Superior T´ecnico, Lisbon, Portugal in IROS Workshop on Robot Programming by demonstration, Las Vegas, SA, Oct 31st, 2003 Motor Representations for Hand Gesture Recognition andImitation.
20. Raymond Lockton and Andrew W. Fitzgibbon, Department of Engineering Science University of Oxford. Real-time gesture recognition using deterministicboosting.