COMPUTER WORD RECOGNITION- IMAGE PROCESSING USING SIGN GESTURES- CREATION OF A NOVEL SIGN LANGUAGE

Dr. Nellulta Sasikala[(1)], Mr. T.akhil[(2)], Mr.M.Vijay kumar[(3)], Mr.G.Sai Suraj[(4)], Mr.A.Satyavardhan[(5)]
[(1)] Associate Professor, [(2)], [(3)], [(4)], [(5)] Project Research Students
Kamala Institute of Technology& Science, Singapur, Huzurabad

ABSTRACT: Sign language is used by deaf and hard hearing people to exchange information between their own community and other people. Computer recognition of Sign language deals from Sign gesture acquisition and continues till text/speech generation. Sign gestures can be classified as static and dynamic. However static gesture recognition is simpler than dynamic gesture recognition but both recognition systems are important to the human community. However, most people do not know the sign language. In this project, we aim to automatically recognize sign language Alphabet / Numbers / Words in image-based methodology using MATLAB.

The problem addressed is based on Digital Image Processing using Skin Detection, Image Segmentation, Image Filtering, and cross-correlation method. This system expects to achieve recognizing gestures of ISL (Indian Sign Language) by special people and converting into speech/text. The statistic of the result of the implementation, it is therefore concluded that the method is used for cross-correlation and color segmentation work with hand gesture recognition. The results obtained are applicable and can be implemented in a mobile device smart phone having a frontal camera. The vision of an efficient system to translate sign language to text is quite achievable, but the challenges lie in optimization.

Key Words: Computer recognition, Skin Detection, ISL (Indian Sign Language)

I. INTRODUCTION

Sign language, any means of communication through bodily movements, especially of the hands and arms, used when spoken communication is impossible or not desirable. The practice is probably older than speech. Sign language may be as coarsely expressed as mere grimaces, shrugs, or pointings; or it may employ a delicately nuanced combination of coded manual signals reinforced by facial expression and perhaps augmented by words spelled out in a manual alphabet. Wherever vocal communication is impossible, as between speakers of mutually unintelligible languages or when one or more would-be communicators is deaf, sign language can be used to bridge the gap.

There are over 135 different sign languages all around the world including American Sign Language (ASL), Australian Sign Language (Auslan) and British Sign Language (BSL). There are also signed representations of oral languages such as Signed Exact English (SEE) and mixes such as Pidgin Signed English (PSE). Sign language is commonly used as the main form of communication for people who are Deaf or hard of hearing, but sign languages also have a lot to offer for everyone.

Sign language is the bridge that connects us to the world of those who have an impaired hearing; or verbal ability. An array of gestures made using hands, fingers, arms, head and also facial expressions; which also helps the deaf and dump to communicate with the people around them and vice versa. It allows them to understand the world around them through visual descriptions and, as a result, contribute to society.

THE BENEFITS OF SIGN LANGUAGES

1. Assists the deaf and also the dump in communicating with others and amongst themselves.
2. Aids in the social inclusion of those who are deaf or hard of hearing.
3. It also gives deaf children the opportunity to educate themselves.
4. Increases the self-esteem of the disabled

Instill a sense of social responsibility as well as sensitivity in non-deaf volunteers; who volunteer to learn sign language in order to communicate with the disabled. It also makes deaf people's lives easier. Because the muscles in their hands develop faster than the muscles in their mouth, deaf children can be taught sign language more easily than verbal languages. As they grow, they can pick up on signs faster and learn to express themselves more effectively. Studies show that children who learn to use signs at a young age are smarter than those who do not. Sign language has proven to be a huge help for autistic children who struggle to express themselves verbally.

II. METHODOLOGY

The project is mainly designed with different commands in MATLAB. Here we develop the prototype of regional sign languages. .Here we mainly use computer vision toolbox functions to determine the results. Besides, USB Webcams support package and image processing toolbox also have a key role in the implementation. The idea is to obtain image from webcam, process the image using required functions and display the output along with the suitable title.

Digital image processing allows one to enhance image features of interest while attenuating detail irrelevant to a given application, and then extract useful information about the scene from the enhanced image.

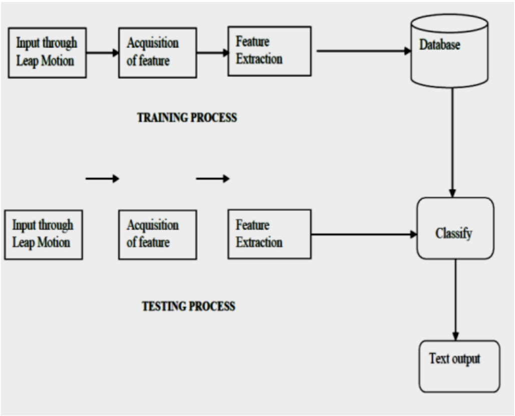
An Introduction to Digital Image Processing Bill Silver Chief Technology Officer Cognex Corporation, Modular Vision Systems Division Digital image processing allows one to enhance image features of interest while attenuating detail irrelevant to a given application, and then extract useful information about the scene from the enhanced image. This introduction is a practical guide to the challenges, and the hardware and algorithms used to meet them. Images are produced by a variety of physical devices, including still and video cameras, x-ray devices, electron microscopes, radar, and ultrasound, and used for a variety of purposes, including entertainment, medical, business (e.g., documents), industrial, military, civil (e.g., traffic), security, and scientific. The goal in each case is for an observer, human or machine, to extract useful information about the scene being imaged.

Often the raw image is not directly suitable for this purpose, and must be processed in some way. Such processing is called image enhancement; processing by an observer to extract information is called image analysis. Enhancement and analysis are distinguished by their output, images vs. scene information, and by the challenges faced and methods employed. Image enhancement has been done by chemical, optical, and electronic means, while analysis has been done mostly by humans and electronically. Digital image processing is a subset of the electronic domain wherein the image is converted to an array of small integers, called pixels, representing a physical quantity such as scene radiance, stored in a digital memory, and processed by computer or other digital hardware. Digital image processing, either as enhancement for human observers or performing autonomous analysis, offers advantages in cost, speed, and flexibility, and with the rapidly falling price and rising performance of personal computers it has become the dominant method in use.

Signal processing is a discipline in electrical engineering and in mathematics that deals with analysis and processing of analog and digital signals, and deals with storing, filtering, and other operations on signals. These signals include transmission signals, sound or voice signals, image signals, and other signals etc.

Out of all these signals, the field that deals with the type of signals for which the input is an image and the output is also an image is done in image processing. As its name suggests, it deals with the processing on images. It can be further divided into analog image processing and digital image processing.

III. IMPLEMENTATION



INPUT THROUG LEAP MOTION :

There are many motion sensing devices available in the market place. The Leap Motion controller was chosen for this project because of its accuracy and low price. Unlike the Kinect, this is a full body sensing device, the Leap Motion controller specifically captures the movements of a human hand, albeit using similar IR camera technology.

ACQUISITION OF FEATURE:

The general aim of Image Acquisition is to transform an optical image (Real World Data) into an array of numerical data which could be later manipulated on a computer, before any video or image processing can commence an image must be captured by camera and converted into manageable entity.

The Image Acquisition process consists of three steps:

1. Optical system which focuses the energy

2. Energy reflected from the object of interest

3. A sensor which measure the amount of energy.

Image Acquisition is achieved by suitable camera. We use different cameras for different application. If we need an x-ray image, we use a camera (film) that is sensitive to x-ray. If we wan infrared image, we use camera which are sensitive to infrared radiation. For normal images (family pictures etc.) we use cameras which are sensitive to visual spectrum. Image Acquisition is the first step in any image processing system.

FEATURE EXTRACTION:

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. So when you want to process it will be easier. The most important characteristic of these large data sets is that they have a large number of variables. These variables require a lot of computing resources to process them. So Feature extraction helps to get the best feature from those big data sets by select and combine variables into features, thus, effectively reducing the amount of data. These features are easy to process, but still able to describe the actual data set with the accuracy and originality.

DATABASE:

A database is an organized collection of structured information, or data, typically stored electronically in a computer system. A database is usually controlled by a database management system (DBMS). Together, the data and the DBMS, along with the applications that are associated with them, are referred to as a database system, often shortened to just database. After classification through database the text output will be displayed on the screen for the respective sign which has been trained and tested.

IMAGE FROM WEBCAM:

A webcam is a video camera that feeds an image or streams and video in real time. Webcams do not have their own power supply and internal memory.

- It gets power from computer to which it is attached.
- There are different types of webcams such as integrated, standalone, network cameras, and webcams with microphones.
- Here we use the integrated webcam of laptop.
- USB webcams support package helps the Webcam to obtain the image of person in mat lab.
- It is performed by taking a screenshot of image during the live video stream of webcam.
- It is made possible by functions webcam () and snapshot ().
- Syntax for webcam() and snapshot()
- **cam = webcam()**
- This creates the webcam object cam and connects to the single Webcam on your system.
- When the webcam object is created, it connects to the camera, establishes exclusive access, and starts streaming data.
- You can then preview the data and acquire images using the snapshot function.
- **e=c. snapshot**

- This acquires a single image from the Webcam cam and assigns it to the variable

- The snapshot function returns the current frame. Calling snapshot in a loop returns a new frame each time. The returned image is always an RGB image.

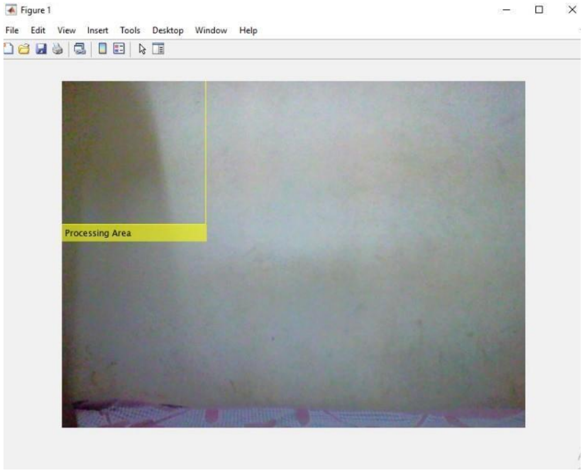
Snapshot uses the camera's default resolution or another resolution that you specify using the Resolution property.

PROCESSING AREA:

Syntax:
bboxes=[x y height width]

Where, x,y are the coordinates of the processing area, height, width are the respective dimensions of the processing area.
bboxes is the variable This syntax is used for opening the processing area on the screen
that is assigned for that syntax in the code.
In our paper we are assigning x=0 , y=0 (i.e; origin) here the processing area starts from origin further it

can be modified according to the user, height=200 , width=200 here the processing area will be created of dimensions 200X200 further it can be modified according to the user.



Processing Area

In the figure window the processing area will display extreme top left of the screen because we have given starting point as origin(x and y) of dimensions 200X200(height and width).

TRAINING:

Syntax:

IFaces=insertObjectAnnotation(e,’rectangle’,bboxes,’Processing Area’)

This syntax is used for training the signs in the processing area.

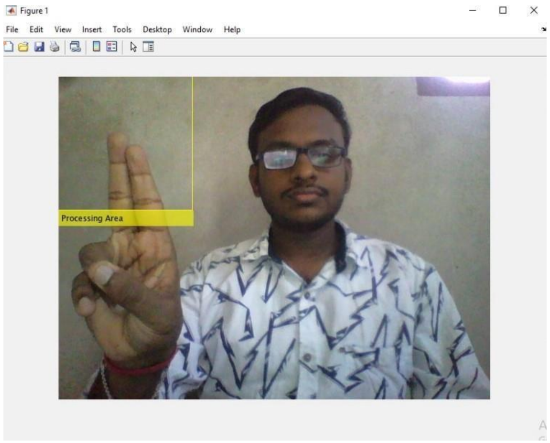
• In bboxes syntax we have created processing area of dimensions 200X200.

• When we run this code a processing area appears on figure window then we have to show the respective sign that user want to train.

• The gestures should be constant until the webcam captures the total image in processing area.

• Here we are giving **e=snapshot(50).** This means the webcam takes 50 images for single sign. It is the user’s choice to set the number of images for training.
The same process continues for all the signs that user want to train. In our project we are training five signs.
• In our code we are giving the data rate as **“0.001”.**With this data speed the images are trained.

• The time taken for training depends on data rate and number of images.In our project we are training five signs for each sign fifty images then total two fifty images with the data rate is 0.001.• The time taken for training our project is approximately sixty minutes (i.e; 1 hour).



Training symbol for word Hello

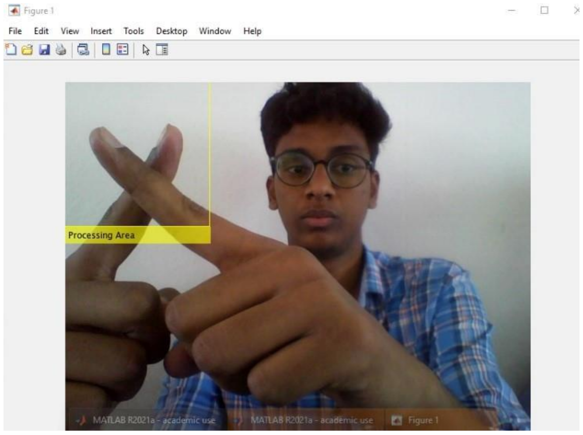
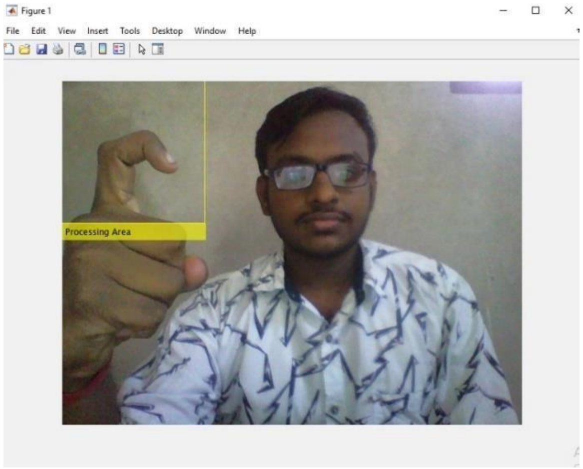
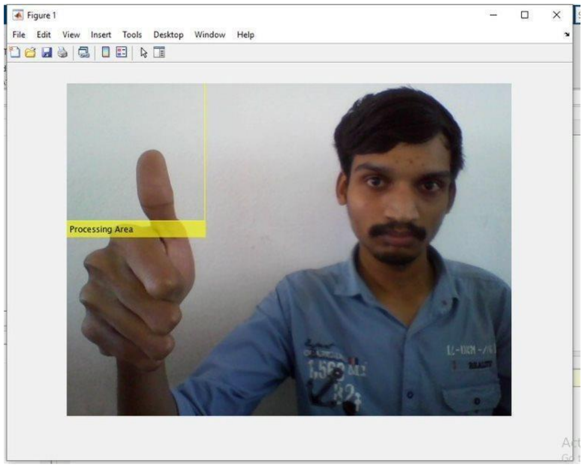


Figure 1
File Edit View Insert Tools Desktop Window Help
Processing Area
MATLAB R2021a - academic use
MATLAB R2021a - academic use
Figure 1

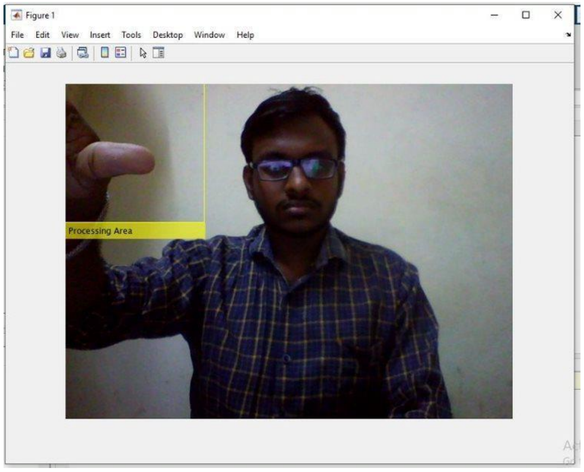
Training symbol for word wrong



Training symbol for word question



Training symbol for word understood



Training symbol for word not understood

IV. Computer Vision Toolbox

Design and test computer vision, 3D vision, and video processing systems

Computer Vision Toolbox™provides algorithms, functions, and apps for designing and testing computer vision, 3D vision, and video processing systems. You can perform object detection and tracking, as well as feature detection, extraction, and matching. You can automate calibration workflows for single, stereo, and fisheye cameras. For 3D vision, the toolbox supports visual and point cloud SLAM, stereo vision, structure from motion, and point cloud processing. Computer vision apps automate ground truth labeling and camera calibration workflows.

You can train custom object detectors using deep learning and machine learning algorithms such as YOLO v2, SSD, and ACF. For semantic and instance segmentation, you can use deep learning algorithms such as U-Net and Mask R-CNN. The toolbox provides object detection and segmentation algorithms for analyzing images that are too large to fit into memory. Pre trained models let you detect faces, pedestrians, and other common objects.

You can accelerate your algorithms by running them on multi core processors and GPUs. Toolbox algorithms support C/C++ code generation for integrating with existing code, desktop prototyping, and embedded vision system deployment. You can accelerate your algorithms by running them on multicore processors and GPUs. Toolbox algorithms support C/C++ code generation for integrating with existing code, desktop prototyping, and embedded vision system deployment.

V. RESULTS

As a prototype we have trained only five signs, further the number of signs can be increased as per users' choice, those five signs output are shown below:

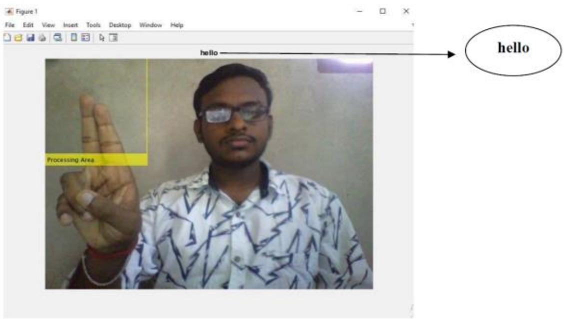
1. hello

2. wrong

3. question

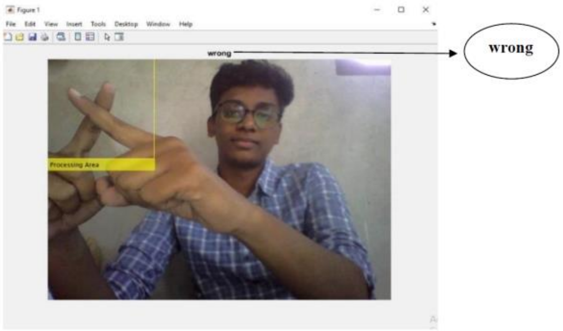
4. understood

5. not understood



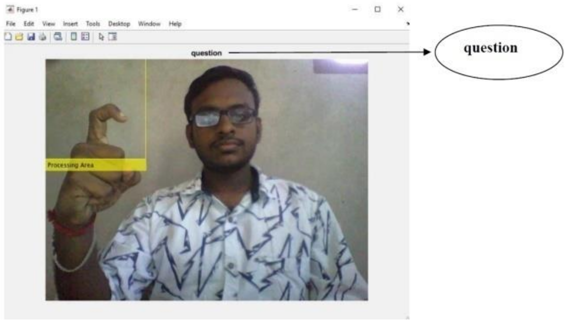
Text output **HELLO** for the given sign

For training of hello sign word, we have shown two straight fingers (i.e.; one index finger and one middle finger) in the processing area , here we are using fifty images for training of this sign word, then system learns (with data rate of 0.001) that sign as hello. while testing we have to show the same sign, so that the system recognizes and gives the output word as hello.



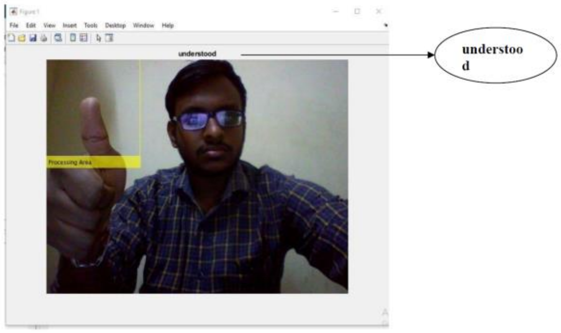
Text output **WRONG** for the given sign

For training of wrong sign word, we have shown two fingers crossed each other (i.e.; index fingers of both hands) in the processing area , here we are using fifty images for training of this sign word, then system learns (with data rate of 0.001) that sign as wrong. while testing we have to show the same sign, so that the system recognizes and gives the output word as wrong.



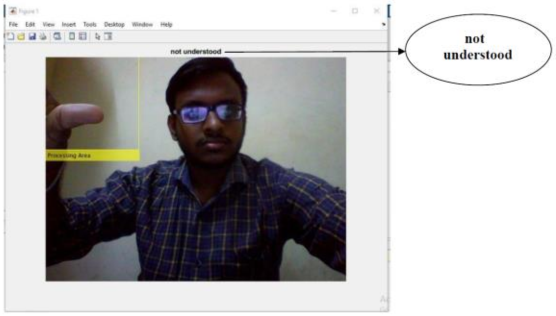
Text output **QUESTION** for the given sign

For training of question sign word, we have shown one finger (i.e.; index finger with slightly folded) in the processing area , here we are using fifty images for training of this sign word, then system learns (with data rate of 0.001) that sign as question. while testing we have to show the same sign, so that the system recognizes and gives the output word as question.



Text output **UNDERSTOOD** for the given sign

For training of understood sign word, we have shown one straight finger (i.e.; thumb finger) in the processing area , here we are using fifty images for training of this sign word, then system learns (with data rate of 0.001) that sign as understood. while testing we have to show the same sign, so that the system recognizes and gives the output word as understood.



Text output **NOT UNDERSTOOD** for the given sign

For training of not understood sign word, we have shown one finger (i.e.; thumb towards left side) in the processing area , here we are using fifty images for training of this sign word, then system learns (with data rate of 0.001) that sign as not understood. while testing we have to show the same sign, so that the system recognizes and gives the output word as not understood.

CONCLUSION

In this paper, Word Recognition using Sign Language –MATLAB. The concept of different methodologies, segmentation, feature extraction, are discussed and compared on recognition and implementation of word recognition using hand gesture. We have trained five signs as a prototype; in future it can extended to many signs as per users’ requirement. we have trained two hundred and fifty images as for five signs(i.e.; fifty images for one sign), for more accuracy number of images can be increased.

There are many approaches to sign gesture recognition and each approach has its strength and weakness, Current systems are mainly focused on manual signs/alphabets/numerical only, but we mainly focused on words because it requires less time to convey the message in an effective and easy manner. This project not only supports standard sign languages but also supports native and regional languages, signs can be set according to the user. In future this application can be developed as app in mobiles and watches.

BIBLIOGRAPHY

1. **Suharjito; Herman Gunawan; Narada Thiracitta; Ariadi Nugroho** "Sign Language Recognition using Modified convolution Neural Network Model" (INAPR) IEEE 2018.

2. **Nitesh S. Soni, Prof. Dr. M. S. Nagmode, Prof.R. D. Komati** "Design of ASL Recognition System Using MATLAB" IJAREEIE Vol. 5, Issue 6, June 2016.

3. **Thair khan** "Hand Gesture Recognition based on Digital Image Processing using MATLAB" IEEE 2015.

4. **Prof. Rajeshri Rahul Itkarkar,** "A Study of Vision Based Hand Gesture Recognition for Human Machine Interaction", International Journal of Innovative Research inAdvanced Engineering, Vol. 1, pp:12, (2014).

5. **Anuja V.Nair, Bindu.V,** "A Review on Indian Sign Language Recognition", International journal of computer applications, Vol. 73, pp: 22, (2013).