

Contents

1	Introduction	2
2	Sign language	3
3	Motivation	4
4	Objectives	4
5	Proposed work	5
6	Project Requirements	6
6.1	Python ILDE with video and image processing modules	6
6.1.1	Python image processing tools	6
7	References	8

1 Introduction

Sign language is one of the effective communication tool for the people who are not able to speak or hear anything. It is also useful for the person who are able to speak but not able to hear or vice versa. Sign language is boon for the deaf and dumb people. Sign language is the combination of different gesture, shape and movement of hand, body and facial expression. Deaf people use this different gesture to represent their thoughts. Each gesture or movement of hand, facial expression and body movement has special assigned meaning. In different part of the world, different sign languages are used. Sign language of any particular place depends upon the spoken language and culture of that particular place. For example in USA, American Sign Language (ASL) is used while in England, British sign language is used (BSL). Similarly Japanese sign language (JSL), French sign language (FSL) and Indian sign language (ISL). Work on Indian sign language has just started after ISL got standardized. This paper presents an extensive review on hand gesture recognition used in sign language recognition which is the heart of sign language recognition.

Aim of developing hand gesture recognition system is to establish an interaction between human and computer and hence recognize the hand gesture automatically which can be used later on for controlling robots or for recognizing the sign language. Human –computer interaction (HCI) which is also known as Man-Machine interaction (MMI) is basically representing the relation between human and computer. Hand gesture is used for communicating between human and machine as well as between two different people with the help of sign language. Hand gesture may be of static type or dynamic type. In static type of hand gesture, certain position and posture is used for communication purpose. Static type of gesture has less computational complexity. In dynamic type of hand gesture, sequence of posture is used for communication. Since it is dynamic in nature therefore it requires more computational complexity. Dynamic hand gesture is more suitable for real time application. Some of the information is necessary to recognize the hand gesture. Some recent review work threw some light on the application of hand gesture recognition in our life. Robot control, Gaming surveillance and sign language recognition are some of the common application of hand gesture recognition. The purpose of this paper is to present some review work in the field of hand gesture recognition in the perspective of sign language recognition.

2 Sign language

It is easy to find a wide number of sign languages all over the world and almost every spoken language has its respective sign language, so there are about more than 200 languages available. There are several sign languages available such as American, British, German, French, Italian, and Turkish Sign Language. American Sign Language (ASL) is well-known and the best studied sign language in the world. The grammar of ASL has been applied to other sign languages especially as in British Sign Language (BSL). BSL is not closely related to ASL, so the differences between BSL and ASL are shown in Figure 1. This section is not going to go further with details of a single sign language because each sign language has its own rules. The next section will aim to give a general description of the shared or common characteristics between the different sign languages: origin, phonology, and syntax. Design a sign language translator is not an easy task.



Figure 1: Difference between American and British Sign Languages

3 Motivation

Hand gestures by definition are the motions of the hands to express ideas or to generally communicate. Therefore, hand gesture recognition is the interpretation of ideas expressed visually by the hands. A good example of hand gestures and their recognition is American Sign Language.

Physically disabled people communicate with those who know the sign language, and they can't communicate with those who don't know sign language. So this is the problem faced every day. Hence this project was planned to understand the sign of the physically disabled people.

4 Objectives

- To create a Sign gestures data base, this can be used for further processing.
- To develop an algorithm to compare the real time Sign gesture with the database image to recognize the Sign Language.
- Setting histogram for different environments and training the model.
- Recognizing various ASL Alphabets and some numbers.
- After recognition process, validating the output by displaying recognized text and speech.

5 Proposed work

Capturing the hand gesture of differently abled people (specially those who cannot hear and talk) by camera and processing the video by converting it in to frames and then processing the frames using the image processing methods and comparing the processed images with the saved database images and printing (Displaying) the appropriate message saved with highest compared images and audio output.



Figure 2: Block diagram

Figure2 shows the block diagram of Sign language to text and speech.

- Video processing.
- Image processing.
- Image comparison between processed image and images stored in database.
- First the video of the hand of deaf and dumb people is captured for a short duration of the time. He can press any key to stop the recording as an interrupt.
- Now the video is sent for processing which is converted into frames with the help of OpenCV library.
- The images are processed and the image irregularities such as blur, contrast, resize and normalization is done.
- One by one the image is compared with the database.
- The boundaries of the image are drawn using the Contours Algorithm and Convex Hull Algorithm.
- Database image and processed image is compared using algorithms.
- Earth Movers distance: The earth mover distance is method to evaluate dissimilarity between 2 multidimensional images.

- SSIM (Structural Similarity Index Mean): The Structural Similarity Index (SSIM) is a perceptual metric that quantifies image quality degradation* caused by processing such as data compression or by losses in data transmission.
- SFIT (Scale Invariant Feature Transform): SFIT is methods of extracting feature vector that describe local patches of an image not only are these feature vectors scale invariant but they are also invariant to translation rotation and illumination.
- The image with highest similarity percentage, the corresponding message will be displayed on the output of the screen and corresponding alphabet is converted into speech output.

6 Project Requirements

6.1 Python ILDE with video and image processing modules

Python is an interpreter, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

6.1.1 Python image processing tools

1. Ski-kit image: scikit -image is an open source Python package that works with numpy arrays. It implements algorithms and utilities for use in research, education and industry applications. It is a fairly simple and straightforward library even for those who are new to Python's ecosystem. This code is of high-quality and peer-reviewed, written by an active community of volunteers.
2. Numpy: Numpy is one of the core library in python programming and provides support for arrays. An image is essentially a standard Numpy array containing pixels of data points. Therefore, by using basic NumPy operations, such as slicing, masking and fancy indexing, we can modify the pixel values of an image. The image can be loaded using skimage and displayed using matplotlib.
3. Scipy: scipy is another of Python's core scientific modules like Numpy and can be used for basic image manipulation and processing tasks. In particular, the submodule scipy.ndimage provides functions operating on n-dimensional NumPy arrays. The package currently includes functions for linear and non-linear filtering, binary morphology, B-spline interpolation, and object measurements.

4. OpenCV python: OpenCV (Open Source Computer Vision Library) is one of the most widely used libraries for computer vision applications. OpenCV-Python is the python API for OpenCV. OpenCV-Python is not only fast since the background consists of code written in C/C++ but is also easy to code and deploy(due to the Python wrapper in foreground). This makes it a great choice to perform computationally intensive computer vision programs.
5. Pyttsx3: pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline, and is compatible with both Python 2 and 3.

7 References

1. Rajaganapathy S “Conversation of Sign Language to Speech with Human Gestures” (ISBCC’15) Elsevier published © 2015.
2. Taner Arsan “Sign Language Converter” International Journal of Computer Science Engineering Survey (IJCSES) Vol.6, No.4, August 2015.
3. Sawant Pramada, Deshpande Saylee “Intelligent Sign Language Recognition Using Image Processing” IOSR Journal of Engineering (IOSRJEN) e-ISSN: 2250-3021.
4. Manjushree K, Divyashree B A “Gesture Recognition for Indian Sign Language using HOG and SVM” IRJET e-ISSN: 2395-0056 Volume: 06 Issue: 07 July 2019.