# VISVESVARAYA TECHNOLOGICAL UNIVERSITY

**"JNANA SANGAMA", BELAGAVI -**590 018.

**2020 – 2021**

Project Report
(Phase - I)
on

# "Conversion of sign language into audio and text"

**Submitted in partial fulfilment for the award of the degree of**

**BACHELOR OF ENGINEERING**

in

**ELECTRONICS & COMMUNICATION ENGINEERING**

**Submitted By**


B Shamanth Kowshik [1JB17EC081]
Uday K V [1JB17EC096]


**Under the guidance of**


**Ms. Supriya M**
Assistant Professor, Dept. Of ECE



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING**

**SJB INSTITUTE OF TECHNOLOGY**

#67, B G S HEALTH AND EDUCATION CITY, Kengeri, Bengaluru-560 060.

||Jai Sri Gurudev||

Sri Adichunchanagiri Shikshana Trust ®

# SJB INSTITUTE OF TECHNOLOGY

**BGS Health & Education City, Kengeri, Bengaluru-560 060.**

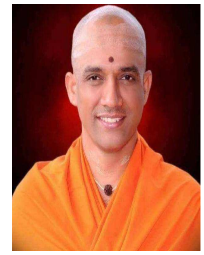
**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING**

## CERTIFICATE

Certified that the project work phase – I entitled ***"Conversion of sign language into audio and text"*** carried out by B Shamanth Kowshik **[1JB17EC081],** Uday K V **[1JB17EC096]** are bonafide students of **SJB Institute of Technology** in partial fulfilment for the award of **"BACHELOR OF ENGINEERING"** in **ELECTRONICS AND COMMUNICATION ENGINEERING** as prescribed by **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the academic year **2020 – 21**. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of project work phase – I prescribed for the said degree.

| | | |
|---|---|---|
| **Mrs. Supriya M** | **Dr. K V Mahendra Prashanth** | **Dr. Ajai Chandran C K** |
| **Guide & Designation** | **Professor & Head** | **Principal** |
| **Dept. of ECE, SJBIT** | **Dept. of ECE, SJBIT** | **SJBIT** |






# ACKNOWLEDGEMENT

I would like to express our profound grateful to **His Divine Soul Jagadguru Padmabhushan Sri Sri Sri Dr. Balagangadharanatha Mahaswamiji** and His Holiness **Jagadguru Sri Sri Sri Dr. Nirmalanandanatha Swamiji** for providing me an opportunity to be a part of this esteemed institution.

I would also like to express our profound thanks to **Revered Sri Sri Dr. Prakashnath Swamiji,** Managing Director, SJB Institute of Technology, for his continuous support in providing amenities to carry out this project (Phase – I) in this admired institution.

I express our gratitude to **Dr. Ajai Chandran C K**, Principal**,** SJB Institute of Technology, for providing me excellent facilities and academic ambience; which helped me in satisfactory completion of project work (Phase – I).

I extend our sincere thanks to **Dr. K V Mahendra Prashanth**, Professor & Head, Department of ECE; for providing us invaluable support throughout the period of my project work (Phase – I).

I wish to express our heartfelt gratitude to my guide, **Mrs. Supriya M**, for his valuable guidance, suggestions and cheerful encouragement during the entire period of my project work (Phase – I).

I express my truthful thanks to **Dr. Mahantesh K & Mr. Bhaskar B,** project coordinators, Dept. of Electronics and Communication for their valuable support.

Finally, I take this opportunity to extend my earnest gratitude and respect to our parents, teaching & technical staff of the department, the library staff and all my friends, who have directly or indirectly supported me during the period of my project work (Phase – I).

**Regards,**
**B Shamanth Kowshik [1JB17EC081]**
**Uday K V [1JB17EC096]**

# DECLARATION

I hereby declare that the entire work embodied in this project report has been carried out under the supervision of **Mrs. Supriya M** in partial fulfilment for the award of "BACHELOR OF ENGINEERING" in ELECTRONICS AND COMMUNICATION ENGINEERING as prescribed by VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI during the academic year 2020 – 21.

B Shamanth Kowshik [1JB17EC081]
Uday K V [1JB17EC096]

# ABSTRACT

Sign language is a means of communication for the acoustically disabled individual, which involves simultaneous use of facial expressions, orientation and movement of hand, finger spellings, body language, head movement and eye gazes in order to effectively convey the message and thought of a person.

Even with usage of sign language, hearing and speech impaired people would face a lot of difficulties in communicating their thoughts and ideas, especially to those who have meager or no knowledge of sign language.

# Table of Contents

# List of figures



# List of Tables

# CHAPTER 1

## INTRODUCTION

In the 21st century field of science and technology has reached such a level that people are expecting more comfortable and useful things, which can make their lives easier. Nowadays, homes with voice recognition built in with the sense of gestures have already been conceived. Even our mobiles are loaded with all similar technologies. Nevertheless, there are people who are less fortunate than us and are physically challenged, may it be deafness or being aphonic. Such people lag behind their non-handicapped peers in using these technologies. These people have some expectations from the researchers and mostly from a computer scientist that we, computer scientists can provide some machine/model which help them to communicate and express their feelings with others. Very few researchers' have them in mind and provide their continuous works for such people. One might expect digital technologies will play a huge role in human's daily routines and whole world will be interacting via machines either with the means of gestures or speech recognition within a few decades. If we are in a position to predict such a future, we ought to think about the physically challenged and do something for them. Sign language is the natural language of the deaf and aphonic people. It is the basic method for the communication of deaf person. American Sign Language (ASL) is the language chosen by almost all the deaf communities of United States of America. Different Sign languages are evolved depending on the regions such as GSL (German Sign Language), CSL (Chinese Sign Language), Auslan (Australian Sign Language), ArSL (Arabic Sign Language), and many more [1].

### 1.1 CHARACTERISTICS OF SIGN LANGUAGE

Characterization of sign language is between two parameter one being manual and other non-manual. The manual parameter consists of motion, location, hand shape, and hand 2 orientation. The non-manual parameter includes facial expression, mouth movements, and motion of the head [2]. Sign language does not include the environment which kinesics does. Few terms are use in the sign language like signing space, which refers to signing taking place in 3D space and close to truck and head. Signs are either one-handed or two-handed. When only the dominant hand is in use to perform the signs, they are denoted as one-hand signs else when the non-dominant hand also comes in the phase it is termed as two-handed signs [3]. Sign language when evolved is different from spoken language so the grammar of the sign language is primarily different from spoken language. In spoken language, the structure of the sentence is one-dimensional; one word followed by another, while in sign language, a simultaneous structure exists with a parallel temporal and spatial configuration.

As based on these characteristics, the syntax of sign language sentence is not as strict as in spoken language. Formation of a sign language sentence includes or refers to: time, location and person, base. In spoken languages, a letter represents a sound. For deaf nothing comparable exists. Hence the people, who are hearing or speech impaired by birth or became hearing or speech impaired early in their lives, have very limited vocabulary of spoken language and faces great difficulties in reading and writing.

ASL is the fourth most commonly used language in the USA. American Sign Language is the language, which is extensively used by deaf people, and this language is officially acquired by the deaf society of United States. ASL is a complete, complex language that employs signs made by moving the hands combined with facial expressions and postures of the body. ASL is not defined as the world language but it has its roots in English speaking parts of Canada, few regions of Mexico, and all over United States of America. A signer of ASL would have trouble in understanding any other Sign language of any region or country, as they are very different according to grammar and signs. Even ASL has its own grammar. ASL did not take the English grammar. Grammar of ASL provides more elasticity in arranging words. ASL consists of almost 6000 gestures of common words with finger spelling used to communicate proper nouns. Finger enchanting uses single hand and 26 gestures to communicate the 26 letters of the alphabets whereas the user of sign language prefers complete word signs almost everywhere this provides them accession and overrides the pace of conversational English. Sign language is not universal--each community had developed their own sign language. For example In Britain and in America people speak English but both ASL and British Sign language are very different and both language signers have the difficulty in understanding the signs. But, it is interesting to note that ASL shares lots of vocabulary terms with Old French Sign Language (LSF) it is because the first teacher of the deaf in United States was Laurent Clerc [4], in nineteenth century.

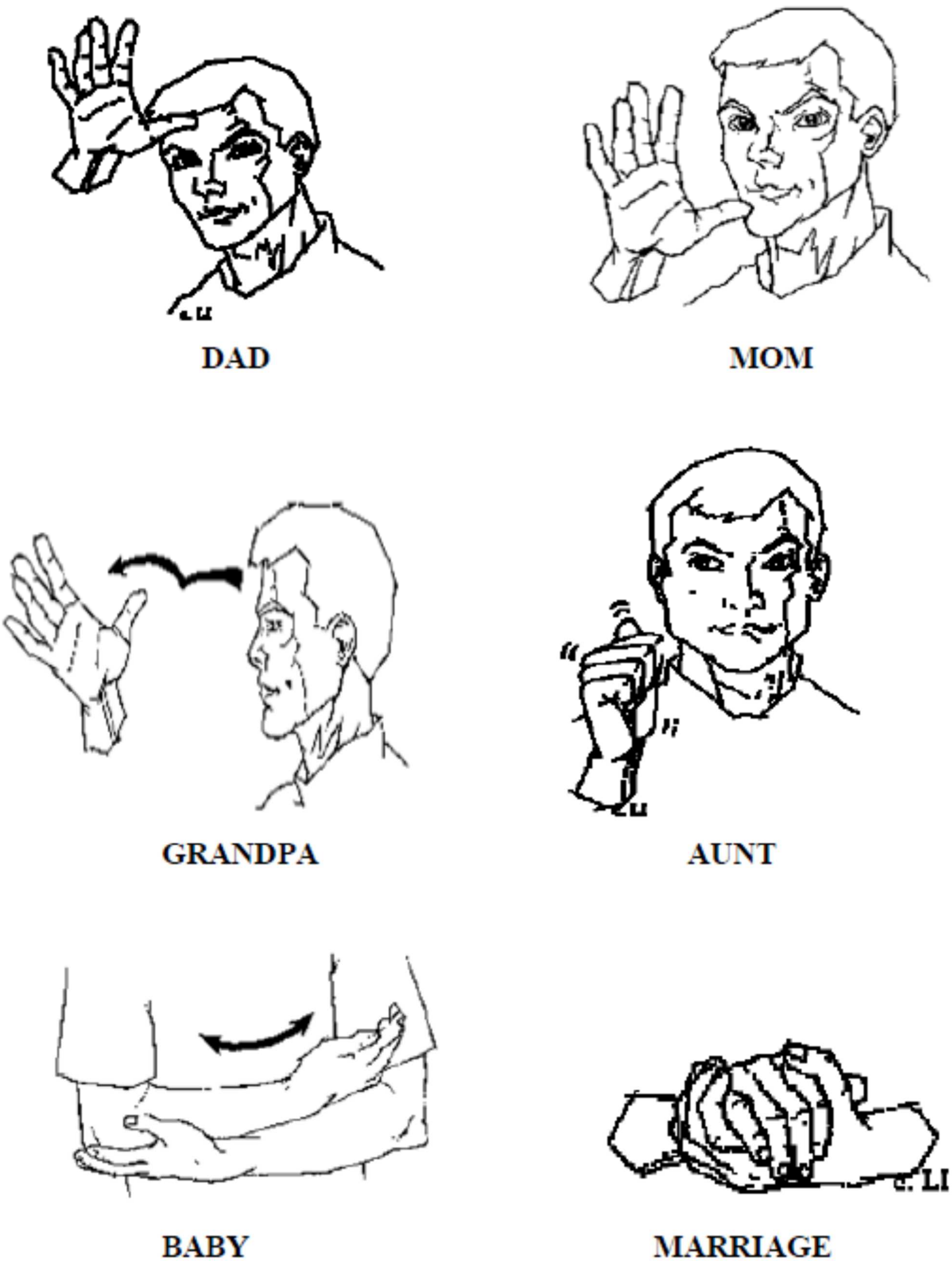



**Figure 1.2. American Sign Language family representation.**

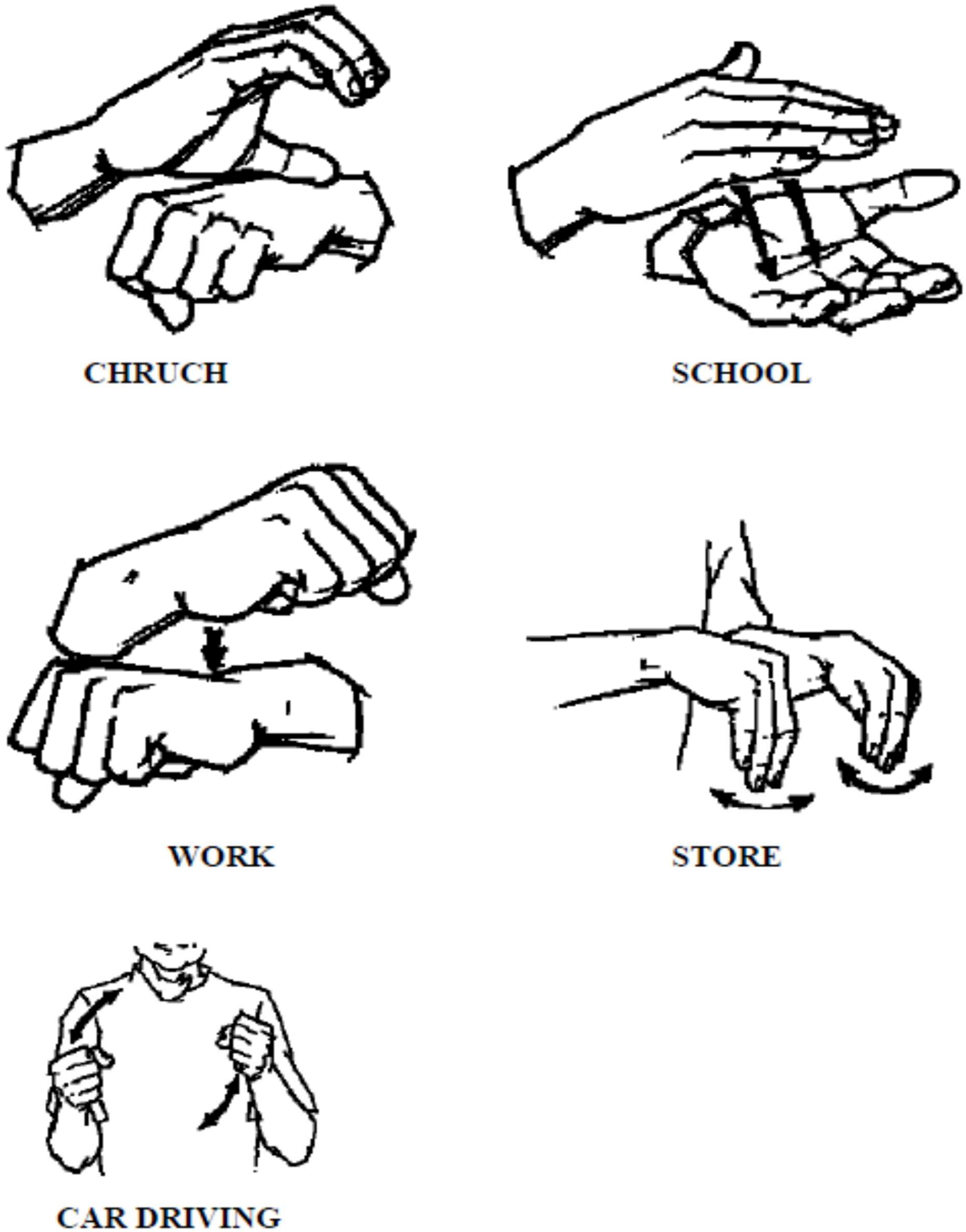



**Figure 1.3. American Sign Language family representation**

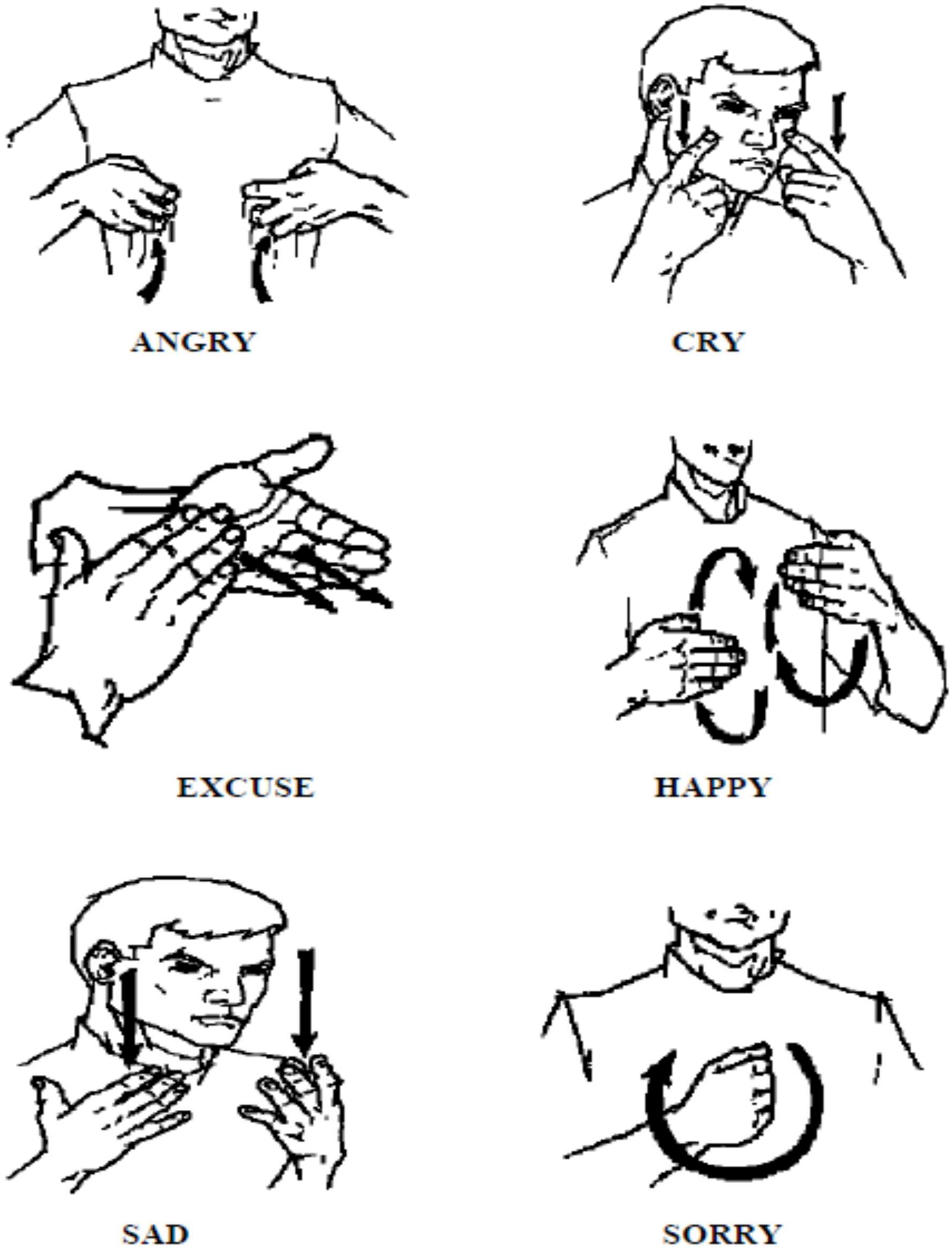



**Figure 1.4. American Sign Language family representation.**

## 1.5 WEBCAM

A webcam is a video camera that feeds or streams its image in real time to or through a computer or computer network. When "captured" by the computer, the video stream may be saved, viewed or sent on to other networks via systems such as the internet, and email as an attachment. When sent to a remote location, the video stream may be saved, viewed or on sent there. Unlike an IP camera (which uses a direct connection using Ethernet or Wi-Fi), a webcam is generally connected by a USB cable, FireWire cable, or similar cable, or built into computer hardware, such as laptops.

**Figure 1.5 - Webcam**

Their most popular use is the establishment of video links, permitting computers to act as videophones or videoconference stations. Other popular uses include security surveillance, computer vision, video broadcasting, and for recording social videos.

Webcams are known for their low manufacturing cost and flexibility, making them the lowest cost form of video telephony. They have also become a source of security and privacy issues, as some built-in webcams can be remotely activated via spyware.

# CHAPTER 2

## LITERATURE SURVEY

The history of hand gesture recognition for computer control started with the invention of glove-based control interfaces. Researchers realized that gestures inspired by sign language can be used to offer simple commands for a computer interface. This gradually evolved with the development of much accurate accelerometers, infrared cameras and even fibre optic bend-sensors (optical goniometers). Some of those developments in glove-based systems eventually offered the ability to realize computer vision- based recognition without any sensors attached to the glove. These are the coloured gloves or gloves that offer unique colours for finger tracking ability that would be discussed here on computer vision- based gesture recognition. Over past 25 years, this evolution has resulted in many successful products that offer total wireless connection with least resistance to the wearer.

### 2.1 History of Data Glove

This book is never going to be complete without the historical development of hand gesture recognition based on computer vision without giving the due recognition for the evolution of hand gesture system based on data glove. Data glove in essence is a wired interface with certain tactile or other sensory units that were attached to the fingers or joints of the glove, worn by the user. The tactile switches, optical goniometer or resistance sensors which measure the bending of different joints offered crude measurements as to determine a hand was open or closed and some finger joints were straight or bent. These results were mapped to unique gestures and were interpreted by a computer. The advantage of such a simple device was that there was no requirement for any kind of pre-processing. With very limited processing power on computer back in 1990s, these systems showed great promise despite the limited manoeuvrability due to tethers that connected the glove to the computer.

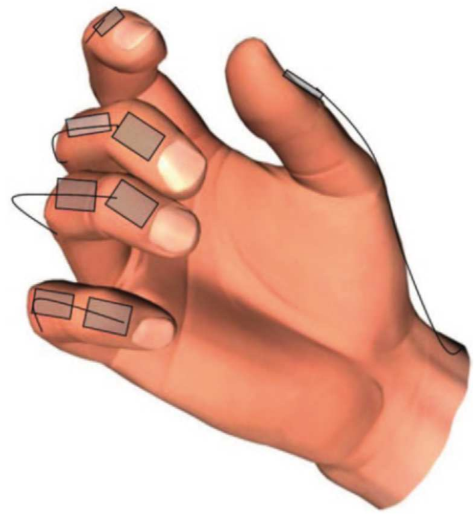


**Figure. 2.1.1** Artistic impression of a sensor glove that places sensors on finger joints

**Figure 2.1.1** shows an artistic impression of a data-glove or a sensor glove that strategically places variety of sensors to monitor the flexing of fingers to form different gestures. Today, there exists gloves that are wireless and easy to wear unlike the ones we had 20 years ago. The following sections of this chapter will discuss the history of some of these devices and their performance scores in interpreting hand gestures.

By looking at the evolution of data gloves, there were two distinct categories emerged over the years.

1. Active data glove—consisted of few or variety of sensors on the glove to measure flexing of joints or acceleration and had a communication path to the host device using wired or wireless technology. These gloves are known to restrain the user of artistic ability.

2. Passive data glove—consisted only of markers or colours for finger detection by an external device such as a camera. The glove did not have any sensors on-board. The first glove prototypes to emerge included the Sayre Glove, the Massachusetts

Institute of Technology (MIT)-LED glove and the Digital Entry Data Glove. The Sayre Glove which was developed in 1977 used flexible tubes with a light source at one end and a photocell at the other, which were mounted along each finger of the glove. Bending fingers resulted in decreasing the amount of light passed between the LED and the photodiode. The system thus detected the amount of finger bending using the voltage measured by a photodiode.

The first glove to use multiple sensors was offered by the ‘Digital Entry Data Glove’ which was developed by Gary Grimes in 1983. It used different sensors mounted on a cloth. It consisted of touch or proximity sensors for determining whether the user's thumb was touching another part of the hand or fingers and four "knuckle-bend sensors" for measuring flexion of the joints in the thumb, index, and little finger. It also had two tilt sensors for measuring the tilt of the hand in the horizontal plane and two inertial sensors for measuring the twisting of the forearm and the flexing of the wrist. This glove was intended for creating "alphanumeric characters" from hand positions. Hand gestures were recognized using hard-wired circuitry, which mapped 80 unique combinations of sensor readings to a subset of the 96 printable ASCII characters. These gloves had limited accuracy and were tethered to computers using cumbersome wiring.

During 1980s, the sensor technology developed rapidly due partly to cold war fears and the natural expansion of industry in many European countries. These sensor technologies paved way for rapid developments in computer technology and peripherals.

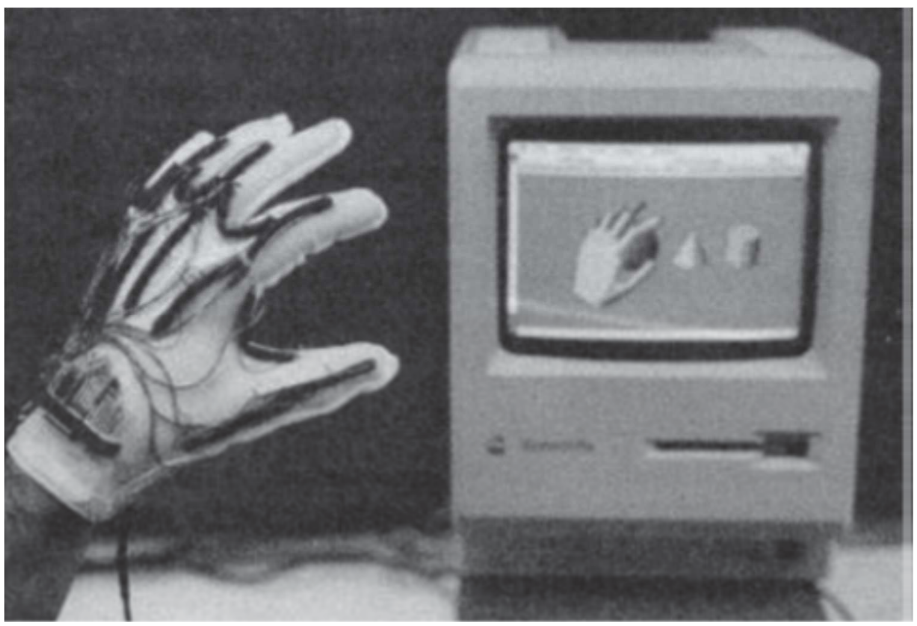


**Fig. 2.1.2** The ZTM Glove developed by Zimmerman

This was an improved version of the first Data-Glove developed by Zimmerman in 1982 which is shown in **Fig. 2.1.2**. The technology was similar to the one used in Sayre Glove in 1977. However, the 1987 version carried fibre optics instead of light tubes and was equipped with 5–15 sensors increasing its ability to distinguish different gestures. The multiple sensors available on the Data-Glove made it popular among researchers of different fields and number of similar devices was developed. Data Glove inspired development of Power Glove, which was commercialized by Mattel Intellivision as a control device for the Nintendo video game console in 1989. The Power Glove used resistive ink to measure the flexion of the finger joints.

There was other development such as Super Glove developed Nissho Electronics in 1995 consisted of sensors and used resistive ink printed on boards sewn on the glove cloth. An updated version of the Power Glove, the P5 Glove, was commercialized by Essential Reality, LLC, in 2002.

### 2.2 What's Out There Today?

The following section details the state of the data glove today. A number of these are now commercially available for different types of human computer interaction (HCI). These data gloves are mainly aimed at researchers to develop sophisticated systems to make the HCI a reality.

### 2.3 Paper 1

**Title**: Sign Language Converter
**Names**: 1. Taner Arsan 2. Oğuz Ülgen
**Year**: 2015
**Methodology**:

1-Database
2-Voice Recognition Procedure
3-Motion Capture Procedure

### 2.4 Paper 2

**Title**: Sign Language Translator for Speech-impaired
**Names**: 1. SATYA PRAKASH 2. KAPIL KUMAR AHUJA
3. RAHUL THAKUR 4. VAMSI KRISHNA
**Year**: 2019
**Algorithm**:

1. Rule Based Classifier
2. Background subtraction method by detecting the colour of skin using HSV (Hue Saturation Value) model.

**Conclusion**: The main objective was to translate sign language to text/speech. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language. This leads to the elimination of the middle person who generally acts as a medium of translation. This would contain a user-friendly environment for the user by providing speech/text output for a sign gesture input.

**Reference**:
https://www.geeksforgeeks.org/project-idea-sign-language-translator-speech-impaired/

**2.5 Paper 3**
**Title**: Conversion of sign language into text
**Name**: Mahesh Kumar N B
**Year**: 2018
**Methodology**:

1. Glove based approaches
2. Vision based approaches.

**Conclusion**: This research work has focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled conditions. By using LDA algorithm for sign recognition operation the dimensionality will be reduced. Due to dimensionality reduction the noise will be reduced and with high accuracy. In future this project will be enhanced by determining the numbers which will be shown in words.
**Reference**: https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf

**2.6 Paper 4**
**Title**: Design and Implementation of Sign Language Translator Using Micro Touch Sensor.
**Names**: 1. Nithyakalyani.K 2. S. Ramkumar
3. K. Manikandan
**Year**: 2020
**Methodology**: The gestures are performed by pressing the fingers either on the palm or other body parts such as forehead or in front of the chest. The words such as, „I", „HELLO", „YES", „NO", „THANK YOU", „SORRY" and „KNOW" are the words for which signs can be performed and recognised using the proposed system.
**Conclusion**: In this project, we have successfully designed and implemented a sign language interpretation system for Indian Sign language with the help of a wearable hand glove. This device allows translation of single-handed signs using micro touch switch and Arduino. The gestures performed by the user are converted into text and speech with the help of input matrix assigned in the micro touch sensor, this can be easily understood by the normal people.
**Reference**:
http://www.ijstr.org/final-print/jan2020/Design-And-Implementation-Of-Sign-Language-Translator-Using-Microtouch-Sensor.pdf

**2.7 Paper 5**

**Title**: Automatic language translation software for aiding communication between Indian sign language and spoken English using lab view
**Names**: Yellapu Madhurai
**Year**: 2018

**Algorithm**:

Real time
Vision based
Automatic and continuous operation
Efficient translation

**Conclusion**: In this project the geometric features which are extracted from the dominant hand improve the accuracy of the system. The whole image frames containing all the aspects of the signs are considered. Training the speech recognition for shorter phrases is difficult than longer phrases. Also, the system is trained on a limited database

**Reference**:

https://www.slideshare.net/madhuriyellapu/sign-language-translator-ieee-power-point

| Method | Type Of Input Device | Segmentation Type | Features (Geometric Or Non) | Feature Vector Representation | Classification Algorithm | | Recognition Rate |
|---|---|---|---|---|---|---|---|
| Tin H. | Digital camera | threshold | Non geometric | Orientation histogram | supervised neural network | | 90% |
| Kouichi M. | Data glove | threshold | Non geometric | 13 data item (10 for bending, 3 for coordinate angles) | Two neural networ k system | back propagation network | 71.4% |
| | | | | 16 data item (10 for bending, 3 for coordinate angles, 3 for positional data) | | Elman recurrent network | 96% |
| Stergiopoulou E. | Digital camera | YCbCr color space | geometric | Two angles of the hand shape, compute palm distance | Gaussian distribution | | 90.45% |
| Shweta in | Web-cam | | Non geometric | | supervised neural network | | N/A |
| William F. and Michal R. | Digital camera | | Non geometric | Orientation histogram | Euclidean distance metric | | N/A |
| Hanning Z., et al. | Digital camera | threshold | Non geometric | augmented of the local orientation histogram | Euclidean distance metric | | 92.3%. |
| Wysoski, et al. | Digital camera | skin color detection filter | Geometric | the histogram of radial boundary | MLP+ DP matching | | 98.7% |
| Xingyan L. | Digital camera | threshold | Non geometric | One dimensional array of 13 element | Fuzzy C-Means algorithm | | 85.83% |
| Keskiin C., et al. | Two colored cameras and marker | connected components algorithm with double thresholding | Geometric | sequences of quantized velocity vectors | HMM | | 98.75% |

**Fig 2.8 :-** Literature survey

# CHAPTER 3

## OBJECTIVES

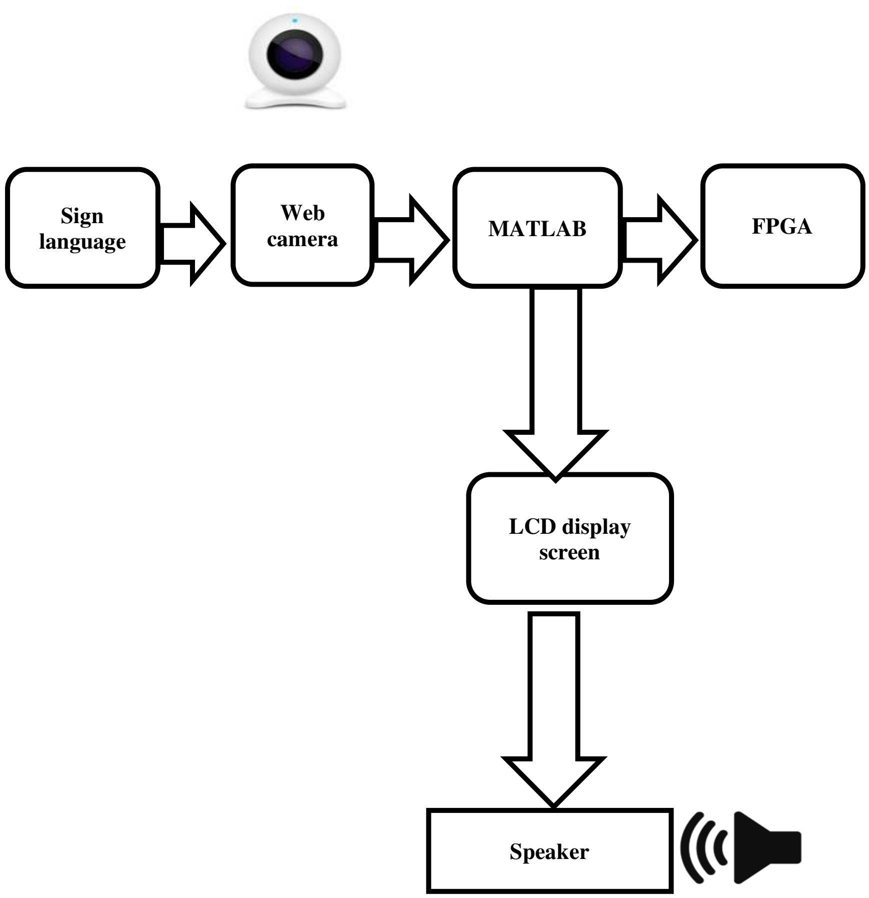
The inability of normal people (who don’t understand sign language) to comprehend sign language gestures results in a communication barrier between normal and hearing and speech impaired people.

The project aims to bridge the communication gap by developing an effective gadget that converts sign-language gestures into audio and textual form providing quick information exchange.

# CHAPTER 4

## METHODOLOGY

A live video feed of the user is obtained using a web camera. Pre-processing involves colour conversion and segmentation. Image segmentation is the process of partitioning a digital image into multiple segments. Hand gesture is detected using suitable algorithm. Gestures are detected and processed by FPGA. After processing of gesture, FPGA kit will be interfaced with speaker.



### 4.1 MIT Data Glove

From its developments in early 1980s, MIT Data Glove has evolved dramatically offering different capabilities with different models. Currently developed under MIT spinoff company AnthroTronix, acceleGlove as shown in Fig. 4.1, is a user programmable glove that records hand and finger movements in 3D. The other models available from them include 5DT's Data Glove for virtual reality that cost between $1000–$5000.

The company initially developed Data Gloves for US Defence for controlling robots. Their acceleGlove is also used in video games, sports training, or physical rehabilitation.

As shown in **Fig. 4.1**, an accelerometer rests just below each fingertip and on the back of the hand. The accelerometers can detect the three-dimensional orientation of the fingers and palm with respect to the gravity when a gesture or any movement is made.

The accuracy of these measurements is within a few degrees which allow programs to distinguish slight changes in hand position. The glove has openings for finger tips which would allow the user to type or write while wearing the glove.


**Fig. 4.1** MIT Acceleglove with its multiple sensors.

## 4.2 GESTURE RECOGNITION TECHNIQUES

The recognition of gesture involves several concepts such as pattern recognition, motion detection and analysis, and machine learning. Different tools and techniques are utilized in gesture recognition systems, such as computer vision, image processing, pattern recognition, statistical modelling.

### 4.2.1 Fuzzy Clustering Algorithm

Clustering algorithms is a general term comprises all methods that partitioning the given set of sample data into subsets or clusters based on some measures between grouped elements. According to this measure the pattern that share the same characteristics are grouped together to form a cluster. Clustering Algorithms have been widely spread because of their ability of grouping complicated data collections into regularly clusters. In fuzzy clustering, the partitioning of sample data into groups in a fuzzy way are the main difference between fuzzy clustering and other clustering algorithm, where the single data pattern might belong to different data groups. Xingyan L. In presented fuzzy c-means clustering algorithm to recognize hand gestures in a mobile remote. A camera was used for acquire input raw images, the input RGB images are converted into HSV colour model, and the hand extracted after some pre-processing operations to remove noise and unwanted objects, and thresholding used to segment the hand shape. 13 elements were used as feature vector, first one for aspect ratio of the hand's bounding box, and the rest 12 parameters represent grid cell of the image, and each cell represents the mean grey level in the 3 by 4 blocks partition of the image, where the mean value of each cell represents the average brightness of those pixels in the image, Then FCM algorithm used for classification gestures.

# CHAPTER 5

# POSSIBLE OUTCOMES

1. This project is a small step towards helping a physically challenged people and lot more can be done to make the product more sophisticated, user friendly & efficient.
2. Using more memory and powerful microprocessor, more languages can be covered.
3. This project can be modified to make it compatible with mobile phones.
4. We can increase the range of product by using more powerful trans-receiver module.

# CHAPTER 6

## REFERENCES

[1] Wikipedia. Sign Language, 2012. http://en.wikipedia.org/wiki/Sign_language, accessed Feb. 2012.

[2] P. Boyes Braem. *Einführung in die Gebärdensprache und ihre Erforschung (Introduction to sign language and its study)*. Signum, Hamburg, Ger., 1995.

[3] S. Prillwitz, R. Leven, H. Zienert, T. Hanke, and J. Henning. Hamburg Notation System for Sign Languages – An Introductory Guide. Institute of German Sign Language and Communication of the Deaf, University of Hamburg, Hamburg, Ger., 1989.

[4]Wikipedia,File:Aslfingerspellalpha,2011. http://en.wikipedia.org/wiki/File:Aslfingerspellalpha.png, accessed Apr. 2012.

[5]Wikipedia.AmericanSignLanguage,2012. http://en.wikipedia.org/wiki/American_Sign_Language, accessed Jan. 2012.

[6]B.Vicars.BasicSignsPictures,n.d.http://www.lifeprint.com/asl101/pages-layout/concepts.htm, accessed Dec. 2011.

[7] Wikipedia. IMAGE PROCESSING, http://en.wikipedia.org/wiki/Image_processing

[8] Wikipedia. RGB COLOR MODDEL, http://en.wikipedia.org/wiki/Color_model

[9] Wikipedia. CAMERA, http://en.wikipedia.org/wiki/Camera

[10] Noor Adnan Ibraheem Department of Computer Science A.M.U., Aligarh Uttar Pradesh, India. *Survey on Various Gesture Recognition Technologies and Techniques.* Volume 50 – No.7, July 2012