**Chapter 1**

**Introduction**

**1.1 Sign Language**

The World Health Organization (WHO) estimated that, 250 million people in the world are deaf as well as dumb.[1] These group people of use symbolic language to communicate with other people. This symbolic language is called sign language. Sign Language is a build for communication used worldwide among hard of hearing, and deaf peoples. Sign language is not a unique language signed consistently in different countries. Sign language is not recent improvement. There is proof that speaking through gestures has been around since the start of human development [20]. Different counties have their own sign language such as American Sign Language, French Sign Language, Indian Sign Language and Puerto Rican Sign Language to name a few. Table 1 give information about different sign languages used in western continent. Gesture based communication is dependent on region and has significant differences from other languages. It is very important to understand sign language when we communicate with deaf or young children and their families. Lack of understanding result in significant challenges in understanding this community and may result miscommunication.

Table 1: Sign Language in the Americas [20]

|  |  |  |
| --- | --- | --- |
| **North America** | **Central America** | **South America** |
| •American Sign Language  •Inuit Sign Language  •Quebec Sign Language  •Puerto Rican Sign Language | •Costa Rican Sign Language  •Guatemalan Sign Language  •Honduras Sign Language  •Mayan Sign Language  •Mexican Sign Language  •Nicaraguan Sign Language  •Panamanian Sign Language  •Salvadorian Sign Language  •Tijuana Sign Language | •Argentine Sign Language  •Bolivian Sign Language  •Brazilian Sign Language  •Chilean Sign Language  •Colombian Sign Language  •Ecaudorian Sign Language  •Paraguayan Sign Language  •Peruvian Sign Language  •Uruaguayan Sign Language  •Venezuelan Sign Language |

Sign Language is a language which is used to convey messages by hand movements, facial expression and body language for communication. It is mainly used by deaf and people who can hear but cannot speak. Sometime family members and relatives must learn sign language to interpret which enables deaf and wider communities to communicate with each other.

**1.2 Mythologies and Misunderstandings about Sign Language**

The Sign language is enclosed by many mythologies and misunderstanding. Most people who are not disable think that Sign language is just simple a manual representation of the spoken language which is not true. In fact, our language and sign language of the deaf have little in common. Sign language has the difficulty of the verbal language, but it is self-determining from the alphabets. The best example is British Sign Language and American Sign Language which are meaningless although the facts show that disable people from United States and Britain perfectly understand each other.

Another common misunderstanding about sign language is that it is globally understandable which is of course not true. As explained above, the Sign language that is used by the deaf in Unites States and Britain are not the same. The different sign languages might be similar in some alphabets but a deaf person from one country to country communicate as fluently as hearing people from the two countries.

Since sign language is a language of its distinct language, finger spelling or the use of guidebook alphabet cannot be used as an alternative to sign. It is utilized in marking the words with a non-existent sign or when the sign isn't known. In addition, Deaf person would take hours to convey a few minutes of messages through finger spelling.

**1.3 Objective:**

My main objective of my thesis is to help deaf community to increases their self-esteem and IQ level and improve their communication skills. Students who are deaf or have a deaf parent or have a close relative with deaf individual will learn by themselves about sign language alphabets and numbers. Deaf community will learn their first step towards the American sign language. Although correct usage of sign gesture plays very important part in effective communication, deaf students are also encouraged to establish connection to deaf community and to carry their new knowledge and skill beyond the class room and into the community at large. Some of important objective are discussed in Chapter 3.

**1.4 Methodology**

In this thesis, Image classification and machine learning have been used for Interpreting American sign language. For image classification, computer vision algorithms were used to capture images and to process data set for filtering as well as reducing noise from images. Finally, data set is trained Sign machine learning algorithm, conventional neural network for measuring accuracy of training data set, result of thesis algorithm is explained in chapter 7. The abstract view of the derived approach combining the image classification and machine learning for American sign language is shown in Figure 1.

Feature Vector

Prepare Image

Feature Extraction

Neural Network

Feature Extraction

Convert RGB to GRAY

Classification

Apply Threshold/Edge Detection

Classified Image

**Figure 1: Project overview for American sign language**

**1.5 Outline:**

**Chapter 2**

**2.1 Related Work**

**2.1.1 American Sign Language using Machine Learning**

American sign language recognition is not a new machine learning problem. During recent decades, different researches already worked on different classifiers such as linear classifiers, neural networks and Bayesian networks [2-11].

As per research point of view a linear classifier is easy to work with because linear classifier are relatively simple models, it requires sophisticated feature extraction and preprocessing methods to get good results [2, 3, 4]. Singha and Das [2] achieved an accuracy of 96% on Ten classes for images of gestures of one hand using Karhunen-Loeve Transforms. These translate and rotate the axes to build up a new framework based on the variance of the data. This technique is useful after using a skin color detection, hand cropping and edge recognition on the images. They use a linear classifier to recognize number sign including thumbs up, first and index finger pointing left and right, and numbers only. Sharma [4] has done research using Support Vector Machines (SVM) and k-Nearest Neighbors (KNN) to illustrate each color channel after background noise deletion and noise subtraction [4]. Their research suggests using contours, which is very useful to represent hand contours. They got an accuracy of 62.3% using a Support Vector Machines on the segmented color channel model.

Machine learning is most commonly used for image recognition. Hidden Markov Model (HMM) and Dynamic Time Warping (DTW), two kinds of machine learning methods, are widely applied to achieve high accuracies [5, 6, 7]. These are mostly good at capturing time-based patterns, but they require clearly characterized models that are defined before learning. Starner and Pentland [5] used a Hidden Markov Model and a 3-Dimesional glove that detect hand movement. Since the glove can attain 3-Dimesional detail from the hand regardless of spatial orientation, they achieved the best accuracy of 99.2% on the test set. Using Hidden Markov Model uses timeseries data to track hand actions and classify based on the position of the hand in recent frames.

Suk [6] suggested a system for detecting hand gestures in a continuous video stream using a dynamic Bayesian network or DBN model. They try to classify moving hand gestures, such as creating a circle around the body or waving. They attain an accuracy of nearly 99%, but it is worth noting that all hand gestures are different from each other and are not American Sign Language. However, the motion-tracking feature would be applicable for classifying the dynamic letters of ASL: j and z.

Artificial Neural networks (ANN) have been used to capture American Sign language transformation [8, 9, 10, 11]. Possibly, the most important advantage of artificial neural networks is that they represent the most important classification structures.However, ANN require significantly more time and data to train. Up to the present time**,** most have been comparatively low.Mekala [8] classified video of ASL alphabet into text using unconventional feature abstraction and a three-layer Neural Network.They extracted features using hand situation and movement.In the past, American sign language classification, could recognise the presence and position of 6 “points of interest” in the hand, each finger and the center of the palm. Mekala also used Fourier Transforms of the images to classify what section of the frame the hand is positioned in.Whereas they claim correctly categorize 100% of images with this framework, there is no indication of whether this result was reached in the training, validation or test set.

Admasu and Raimond [9] classified Ethiopian Sign Language and achieved 88.5% accuracy result using a feed Forward Neural Network.They use a substantial amount of image preprocessing, including image size standardization, image background deduction, contrast adjustment, and image segmentation**.** Gabor Filter and Principal Component Analysis method was used to extract features.The most related work up to date is by L. Pigou’s [11] research of ANN’s to categorise 20 Italian gestures from the “ChaLearn 2014 Looking at People” gesture recognising competition.They used a Microsoft Kinect on whole body images of person performing the gestures and reach a cross-validation accuracy of 91.7%.With the 3-Dimensional glove, the Kinect allows detection of depth features, which helps significantly in classifying American sign language.

Non-Vision based technology such as Glove-based handshape recognition normally contains the person wearing glove and a certain quantity of wires to connect this glove to a computer. These methods are very hard and non-natural way to communicate with the computer [15]. This device required electricity or electromagnetic interference to get data about the hand, which is sufﬁcient to provide a description of a handshape gesture [16]. Scientists refer to data gloves in different ways, e.g. CyberGlove and Accele Glove.

Figure 2.1 shows the position of the sensors in a data glove proposed by Bedregal[17]. Basically, a timeline of frames can characterise any movement. Thus, a timeline of hand arrangement represents a hand movement using a data glove. An arbitrary generated hand conﬁguration was used to replicate the data transfer [17]. Each expression of the handshape is represented by a tuple of interval angles from each sensor. The detection was applied to Brazilian Sign Language (LIBRAS), using Fuzzy logic.

****

**Figure: 2.1. A Data Glove design with Sensor.**

In this paper they developed a similar hardware device called the Accele Glove. In their research they used a microelectronic mechanical system (MEMS) to extract hand conﬁguration. They have been functional on Vietnamese Sign Language for twenty-three gestures with Fuzzy logic. They achieved the results by handshape, with an overall 98% precision. The relative angles between palm and finger is the data found from the sensing device. The glove covers six accelerometers and a BASIC Stamp microcontroller as in Figure 2.2[18-19].

****

**Figure 2.2 A Glove device with Sensor**

Researchers [] have proposed a new system for a gesture-to-speech/text for deaf community, applied to Arabic Sign Language. This author includes the design and implementation of a smart glove. Main advantage of this glove is that it does not depend on light conditions, which means it gives good accuracy in dark environments. As per the author, the glove is low price, low power consumption and has full mobility. Another benefit of these gloves is that they attached ﬂex sensors which used a wireless interface to a microcontroller.

**Chapter 3**

**Data set**

**3.1 American Sign Language:**

American sign language is using to communicate between deaf community and normal community. However, there are only 2.5 million ~ 5.0 million speak which significantly limit the number of that they can easily communicate with [12].



**American Sign language Manual Alphabet [13].**

American Sign Language is implemented from French sign language which was introduced by Thomas Hopins Gallaudet in United States. ASL is like French sign language; Individuals who speak American Sign Language can effectively communicate in French Sign Language. A variation of American Sign Language also exits. Similarly, to English which is international language, but it has unique variations between English spoken in England, United States or Australian, there are separate difference that have changed in sign language [12].

****

**Figure 1. American Sign language numbers. [13]**

**3.2 Characteristics of American Sign Language:**

* American Sign language is an entire visual-gestural dialect with its very own language structure, vocabulary, and linguistic structure.
* Like other sign language, it utilizes the hands, the body, and face looks (counting mouth developments) to express significance and the eyes to see meaning.
* Hand - to-hand connection is especially critical in ASL since it has no composed frame. There are, in any case, documentation frameworks that are utilized for recording signs on paper.
* ASL is separate from English and is unique from other signed languages. An example of the distinctiveness of signed languages from each other and from the surrounding spoken language(s) is that, although English is the shared spoken language of the U.S., Canada, and Britain, signers of ASL do not understand signers of British Sign Language (BSL).

**3.3. Statistics** **about sign language use in Canada:**

In Canada, Statistics Canada reports that as indicated by the 2006 Census 8,995 people revealed a gesture-based communication just like their primary language or one of their first languages, as gave beneath.

|  |  |
| --- | --- |
| American Sign Language | 2,485 |
| Quebec Sign Language | 730 |
| Sign languages, not included elsewhere | 5,780 |

**Table 1: Statics about Sign Language as a Mother Tongue [14].**

In addition, Statistics Canada reports that as per the 2006 Census 43,090 people reported knowledge of a gesture-based communication, as provided below.

|  |  |
| --- | --- |
| American Sign Language | 11,110 |
| Quebec Sign Language | 730 |
| Sign languages, not included elsewhere | 5,780 |

**Table 2: Statics about Knowledge of Sign Languages [14].**

**3.4 Dataset and variables:**

I have created my own data set. This dataset was a collection of 36 which contain A to Z alphabet and 0 to 9 numbers digit. In my dataset consist of A to Z alphabet and 0 to 9 numbers where I have used right hand to capture 1200 images for specific alphabet and numbers. After that I implement code which convert flip image to right to left hand image. The height and width ratios vary significantly but average approximately 50X50 pixel. The dataset contains over 100,000 images in gray scale color. Additionally, People who want to add their images to this dataset than they can add. Below figure shows an image of A to Z alphabet.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A** |  | **N** |
|  | **B** |  | **O** |
|  | **C** |  | **P** |
|  | **D** |  | **Q** |
|  | **E** |  | **R** |
|  | **F** |  | **S** |
|  | **G** |  | **T** |
|  | **H** |  | **U** |
|  | **I** |  | **V** |
|  | **J** |  | **W** |
|  | **K** |  | **X** |
|  | **L** |  | **Y** |
|  | **M** |  | **Z** |

**Figure 2: Data set images.**

|  |  |
| --- | --- |
| **Property** | **Description** |
| Alphabets | A to Z |
| Numbers | 0 to 9 |
| Color | Gray Scale |
| Dimensions | 50x50 |
| Height | 50 pixels |
| Width | 50 pixels |
| File type | JPEG |

**Table 3: Dataset Description and Image property**

**3.5 Capturing Images for Dataset**

Used for detecting hand gesture using skin colour, there are different approaches including skin colour-based methods.In my case, after detecting and subtracting the face and other background, skin recognition and a contour comparison algorithm were used to search for the hand and discard other background colour objects for every frame captured from a webcam or video file.Palm to extract their contours and saved the four for evaluation with the contours of the skin detected area of every frame.After detecting the skin area for each frame captured, I compared the contours of the detected areas with the previously saved hand histogram template contours to remove other skin like objects existing in the image.If the contour comparison of the spotted skin area complies with any one of the saved hand histogram contours than it captured only hand gesture only. I have explained more information about my approach to hand detection in chapter 4.