**Chapter 2**

**Related Work**

**2.1 American Sign Language using Machine Learning**

American sign language recognition is not a new machine learning problem. During recent decades, different researchers 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 classifiers 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. 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 detects 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 time series 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 requires 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 the section of the frame the hand is positioned in.Whereas they claim to 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 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 golve 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.[18-19].****

**Figure 2. A Glove device with Sensor**

Researchers [49] have proposed a new system for a gesture-to-speech/text for deaf community, applied to Arabic Sign Language. Abdulla and Manaf [49] 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.