**Abstract**

Sign gesture recognition is an important problem in human computer interaction with signiﬁcant societal influence. However, it is a very complex task, since sign gestures are naturally deformable objects. Gesture recognition contains unsolved problems since last two decades, such as low accuracy or low speed, and despite many proposed methods, no perfect result has been found to explain these unsolved problems. In this paper, we suggest a machine learning approach to translating sign gesture language into text.

In this study, we have introduced self generated image data set for American sign language (ASL). This dataset was a collection of 36 characters which contain A to Z alphabets and 0 to 9 number digits. The proposed system can recognize static gestures. This system can learn and classify specific sign gesture of any person. We used a convolutional neural network algorithm for classified image to text.

**Introduction:**

The World Health Organization (WHO) estimated that, 250 million people in the world are deaf as well as dumb [1]. These group of people of group use symbolic language to communicate with other people. This symbolic language is called sign language. Sign Language is a built for communication used worldwide among hard of hearing and deaf people. 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 a name a few. Table 1 gives 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 results in significant challenges in understanding this community and may result in miscommunication. 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.

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 using machine learning algorithm, conventional neural network for measuring accuracy of training data set. The abstract view of the derived approach combining the image classification and machine learning for American sign language is shown in Figure 1.

**Related Work:**

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.

**Date set and Variables:**

I have created my own data set. This dataset was a collection of 36 characters which contain A to Z alphabets and 0 to 9 number digits. I used right hand to capture 1000 images for specific alphabets and numbers. The height and width ratios vary significantly but average approximately 50X50 pixels. The dataset contains over 36,000 images in grey scale color. Additionally, people can add their images to this dataset. Below figure shows an image of A to F alphabet.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A** |  | **D** |
|  | **B** |  | **E** |
|  | **C** |  | **F** |

Figure 2: Data set images

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

Table 1: Dataset Description and Image property

**My Approach for Hand Detection**

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.