10. CONCLUSION AND DIRECTION

FOR FURTHER RESEARCH

This thesis contributes largely to the Indian Sign Language recognition in developing algorithms and methods that have immensely met the objectives of creating a comprehensive system which dedicates itself to the Indian deaf and hearing impaired.

Primarily, a dynamic image processing methodology has been formalised consisting of the following processes.
1. Image normalization by bicubic method,
2. Colour Normalization using HUE and gray scale,
3. Image localization such as noise removal and edge detection
4. Background modelling using an adaptive approach,
5. Segmentation to identify regions accurately and ultimately classify IASL gesture,

Secondly, an optimized recognition approach has been proposed with algorithms of high quality which includes

1. Construction of a feature vector containing hand gestures and temporal information for classification using the features found in the individual images.
2. Combinational feature vectors for classification and recognition module,
3. Fusion of classifiers for efficient learning-classification framework for Indian Sign Language Recognition system based on the proposed gesture representation,

10.1 Normalising method

The performance of the system gets enhanced if the signer is centred within each frame using bicubic method after recording the image through digital camera. This in turn helps in achieving comparable recognition rates.
10.2 Colour Normalization

The image intensity normalisation method is used, by assuming that, when the intensity of the lighting source increased by a factor. Each RGB component of each pixel in the image is scaled by the same factor. The distribution of color values in images depends on the illumination which varies widely under different conditions. Interactive adjustment of contrast and brightness by manipulating the range is carried out in experiments.

10.3 Image Localization

While removing the noise disturbance, it is observed that non-linear techniques provide better results than the linear ones for which weighted median filter is proposed. Edge detection is carried out with contemporaneous filters and the best among them is chosen which is then modified to give improved image with a canny operator.

10.4 Background modeling

For separation of background, foreground and background modelling an adaptive approach is used and found to be a more suitable method. It is also proved that background information which are not needed can be deleted in sequence.

10.5 Segmentation

Accurate segmentation is a paramount to the success of the recognition approach. There is no hardware that can contribute much and hence gestures are classified using geometric, textural and statistical information. To maximise the performance of the segmentation methods, a suitable colour space is formed and then segmentation is applied to the desired hand region of the image resized by 256 x 256 pixels. Detection of edges and contours of the gestures are proved by four different methods each having its own significance like pixel, topological, boundary and contour. Optimization of threshold values is done using particle swarm optimization technique for detection of hand signs.
10.6 Feature vector

Feature extraction is a special form of dimensionality reduction. If the features extracted are carefully chosen, it is expected that the feature set will extract the relevant information from the input data in order to perform the desired task. The features are extracted and a database is formed that has the mean intensity, area, perimeter, diameter, centroid, entropy, homogeneity, energy and dissimilarity. Feature combination is also attempted by combining structural, statistical and textural features which is the need of the hour in sign language recognition system.

10.7 Combinational feature vectors

Although it is possible to extract a large set of features, only a small subset of desired parameters are used in the classification due to the curse of dimensionality. It is observed that when the dimensionality increases, the amount of required training data increases exponentially. Moreover, it is noted that there is a strong correlation between different features of the image and therefore, it is necessary to combine them to produce a refined feature vector.

10.8 Classification methods

Three methods are followed to classify the signs. (1) Statistical classifiers (2) Neural network (3) the fusion of the above two methods with support vector machine. An analysis of different classifiers is done out of which the Naive bayes approach is found to be better for Indian Sign Language. However, the results produced are biased by the size of the database, the lack of training data and the large amount of singletons leads to a very difficult task.

To overcome this, initially back propagation procedure is adopted that applies Windrow-Hoff learning rule to multiple layer and non linear differential transfer function which is a relatively efficient method to train the neural nets. Classifier combination methods are also tried and the best combination is accomplished by SVM +KNN.
10.9 Recognition

Unlike the conventional implementations, a new learning algorithm called Extreme Learning Machine (ELM) is proposed for Single-hidden Layer Feed forward Neural networks (SLFNs) which randomly chooses hidden nodes and analytically determines the output weights of SLFNs. It is proved that the above is much faster than the earlier conventional ones.

From the study, it is inferred that application of human knowledge in sign language recognition is a complex task. Hence, an automatic recognizer has been developed to overcome the complexity.

The most important problems mainly faced during the process are:

1) **Data acquisition:** computer vision is not as accomplished as human observers in capturing the complex, static hand gestures that form sign language. This means that the information that is present in a sign movement of a human being may not be observed exactly by an automatic vision analysis system. Thus, the data that humans work with is not necessarily identical to the data of the recognizer work. This needs better techniques that are successful for human signers and an improvisation has been done to achieve the goal.

2) **Differences in basic system architecture.** There is no clear model of human sign recognition especially Indian Sign Language. This makes it more difficult to translate observations from human sign recognition to the automatic recognizer as the signers use techniques that are not compatible with the current architecture of the recognizer.

A more thorough understanding of human sign recognition, more sophisticated computer vision techniques, and a close co-operation between the fields of automatic sign language recognition and human sign perception, seems to be the best way to overcome these obstacles.
IASLR is capable of classifying Indian signs with an accuracy of around 97.35% for the signers used in training, and about 96.92% for other signers. The modular design of the research allows scope for future enhancement of the work both in terms of expanding its vocabulary, and betterment of the recognition accuracy. The major area in which accuracy can be improved is in the database generation and classification of sign location where the performance can be improved by the additional concentration. While developing the steps few attempts have been made to optimise these processes as the primary focus is to attempt to assess the feasibility. Future research can therefore focus on parallelising these processes, which can greatly improve the overall time complexity of the system.

This can be further developed into a system which is integrated as the upcoming telecommunication devices with cameras to bridge the communication gap between the hearing and deaf/hard of hearing communities. The proposed research work can be enhanced in terms of increase in the data processing speed and data storage by using the compression techniques and feature selection/reduction techniques. This work has compositely contributed to the Indian Sign Language, which has very limited research in its automated recognition. Though the techniques are developed specifically for ISL these can be easily adapted to other sign languages and gesture recognition systems. It is also recommended for further research which will be a boom to the deaf and hearing impaired communities.