Conclusion and future research

In this chapter, we first summarize the approaches applied for the word recognition system in this thesis and lastly the future research directions are discussed.

5.1 Conclusion

Free style handwriting recognition is a difficult problem, not only because of the great amount of variations in handwriting, but also because of the overlapping and the interconnection of the neighboring characters. Further more, when observed in isolation, characters are often ambiguous and require context to minimize the classification error. Thus, current research aims at developing systems for limited domain applications. Recognition strategies heavily depend upon the nature of the data to be recognized.

Challenges for handwritten word recognition are higher in Indian languages due to the large number of vowels, consonants and conjunctions, its inflectional nature and complex character grapheme. And lack of standard database. Keeping all this in mind and the complex nature of the problem we have tried to address our work for Devanagari script, handwritten word recognition.

The research presented in this thesis has described techniques for segmenting and recognizing handwritten words. The main focus of this thesis was the development and investigation of segmentation technique, structural based classification of characters. Character recognition classifiers were designed,
implemented, tested, and compared and finally all technique were integrated to perform the task of word recognition.

In chapter 3 we have tried to address few problems in Devanagari script based handwritten word recognition. As no standard database is available we started by collecting the data. To get the features such as hanger, stretcher, aligner we used SVD but it failed to generate the features as per the requirement. The invariant moments were taken as feature extraction due the insensitive nature to translation, scale change, mirroring, and rotation and classified using minimum distance classifier. As the results were unsatisfactory due to the varying style of writing a structure based system was developed. It is a novel technique where we have considered the shape of the character as feature, but not directly the shape was taken but box were fixed according to scan of the character done from top to bottom and then left to write. This approach generated template for each character categories by taking the letter shape into consideration by fitting the box to the characters which also acted as the classifiers. The patterns generated by this approach were recognized using rule based approach. The growth in recognition rate compared to invariant moment was 45 to 50% and the overall result achieved is 80%. The implementation involves a large number of assumptions and restrictive conditions. Therefore, the recognition results are very much dependent on the implementation issues. A good improvement was seen in the recognition rate as compared to the invariant moments based approach.

The major constraint of the proposed system were the limited vocabulary used, manual segmentation was performed and even the box fitting procedure in structure based recognition was not done through automated process. Modifiers and header line removal for maintaining the simplicity and uniformity of the database, if not done will add to the complexity of the problem and ambiguity can be generated for the character identification.
5.2 Future Research

A FPGA based hardware circuit is designed to display on LCD the English synonym for the Marathi and Hindi words, which can be improved more in sense that NN learning done for the patterns in Matlab i.e. software based converted to the hardware implementation for better adaptability and design to display Marathi or Hindi.

Fully automated segmentation and box fitting algorithms can be developed for fast computation for solving such complex problems. Increasing the size of the vocabulary while forming the database. Even consider the modifier and partial consonants while feature extraction and recognition purpose. A Large Vocabulary can be taken for further development. Consideration of Natural language processing which can form the integral part of such system when the large vocabulary is used for designing the database.