7.1 Conclusions

There are many cases of system which are useful for video text information extraction system like: vehicle license plate extraction, text based video indexing, video content analysis and video event identification. In this work, we have new approach for character recognition system based on template matching. This system tested the 50 videos with 1250 video frames and 1530 text lines. The system is texture-based approaches to automatic detection, segmentation and recognition of visual text occurrences in images and video frames. The characters are recognized automatically on run-time basis, In a few cases in which 7.85% characters could not get detected but some other character get recognized. The overall empirical performance of this system recognition rate is 92.15%. Empirically show that the recognition rates of VTDAR are compared to branded tools as Tesseract 3.02, Transym OCR 3.3 which is approximately similarly.

We are proposed two different techniques for accomplished result first is template base and second is neural network as SOM based. We achieved accuracy rate of character recognition rate by using SOM is better than Template. We showed that adapting text detection constraints to each image Significantly improves recognition performance. The character recognition is done by the SOM. We achieved accuracy rate of text recognition is 95.7144%.
7.2 Application of this System

- There are many significant of a text extraction such as Keyword based video or image search like.
- Text based image indexing and retrieval.
- Offline vehicle license detection and recognition and Document analysis.
- Page segmentation, technical paper analysis.
- Street signs, name plates, document coding, object identification.
- Text based video indexing, video content analysis etc.

7.3 Visions for Future Work

I explored the idea of analyzing the shapes within a candidate text region to verify that it contains text. For example, the occurrence of text and edges of the shapes in a region could be used to remove very simple shapes unlikely to be text characters. Unfortunately, some characters in some scripts have very simple shapes. I abandoned this idea because I was unwilling to impose constraints on the script of the text to be detected. However I believe the idea of analyzing shapes within a candidate text region deserves further investigation.

The VTDAR algorithm and SOM presented in this dissertation works is well with large font sizes. However, there is a significant amount of text in video that has very small size, sometimes with stroke widths less than one pixel connected component labeling on such small fonts often gives inaccurate results.
In addition to growing, shrinking, and rotating text, other types of “stylized” text can be found in general-purpose video. For example, text can break into pieces, or morph between fonts, or undergo perspective distortion. The tracking algorithm presented in this dissertation works for some of these cases, but further research is required to extend the algorithm to handle more types of stylized text. More research is needed to design OCR modules geared specifically for the unique challenges of text in video. One of the prime challenges for our recognition approach is connected text, either due to the font type or image blur. This is a challenge because our method assumes that each connected component is one discrete character. Future research should explore how to adapt this technique so it does not rely on having separated characters.