Abstract

This paper deals with the recognition of printed basic Telugu characters using the discrete curves and approximation string matching. The features are extracted from smoothed images, obtained after the thinning operation. As by only thinning, spines may arise which will affect the recognition of the character. The features are the discrete curves, specified using the $3 \times 3$ regions of connected component representation. We represent the discrete curves in the form of a string, so the set of discrete curves result a set of strings. Then using the string matching operation we compare the string obtained from the stored character with the string obtained from the extracted character. As we are dealing with the characters so there may be the
presence of noise which will affect the performance of the method so we are considering the approximation string matching instead of the exact string matching. The extracted features of the character are represented as a string and the string is stored in a trie data structure so that a uniform time will take to compare the strings. For the efficient approximate string matching we are using the Look ahead branch and bound scheme with the trie. We apply our method on 42 printed basic Telugu characters for demonstration and it gives promising results. However more extensive study on realistic data is required for betterment of the approach.

References

Character Recognition using Discrete Curve with the use of Approximate String Matching

Index Terms

Computer Science

Pattern Recognition

Keywords

Discrete Curve  Approximate String Matching  Trie  Look Ahead Branch And Bound