Abstract

The goal of this paper is to present an overview about the thinning problem in Arabic text recognition. Thinning "Skeletonization" is a very crucial stage in the ACR, it simplifies the text shape and reduces the amount of data that needs to be handled and it is usually used as a pre-processing stage for recognition and storage systems. The skeleton of Arabic text can be used for each of the baseline detection, character segmentation, and features extraction and also ultimately supporting the classification. Choosing or designing the effective thinning algorithm for Arabic text is crucial in ACR. In this paper, the importances of the thinning for the ACR and the usage of the text skeleton in ACR system are discussed and presented. As well as the challenges that have an impact on the thinning of Arabic text are discussed. The methods of Arabic text thinning are discussed and reviewed based on the technique used, and the methods advantages and drawbacks are discussed in details.

References

The Thinning Problem in Arabic Text Recognition – A Comprehensive Review

recognition using HMMs with explicit state duration. EURASIP Journal on Advances in Signal Processing (1-13).

Analysis & Applications. 5(1): 31-45.
Index Terms

Computer Science
Pattern Recognition
Keywords
  Thinning  Skeleton  Iterative  Non-iterative  Parallel  Sequential  Pre-processing
Arabic character recognition.