Abstract

This paper presents a simple and efficient approach for the implementation of Optical Character Recognition and translation of scanned images of printed text into machine-encoded text. It makes use of four different image analysis phases followed by image detection via multi layer back propagation neural network. This paper also describes scanning the entire document and recognizing individual characters from image irrespective of their position, size and various
font styles. It deals with recognition of the symbols from English language which can be easily extended for large set of symbols including UNICODE characters; this can be easily achieved by training the network for different symbol set. The phases described in this paper are simple and easy to implement.

Reference

- Neural Networks for Unicode Optical Character Recognition www.projectsatbangalore.com/IEEE.pdf
- “α-Soft: An English Language OCR” (Computer Engineering and Technology (ICCET), 2010 2nd International Conference on 16-18 April 2010)
- “Image Recognition with Neural Networks” Murat Firat http://www.codeproject.com/KB/cs/BackPropagationNeuralNet.aspx
- Backpropagation http://en.wikipedia.org/wiki/Backpropagation
- Creating Optical Character Recognition (OCR) applications using Neural Networks Alex Cherkasov http://www.codeproject.com/KB/dotnet/simple_ocr.aspx
- B.V.Dhandra, Mallikarjun Hangarge, Gururaj Mukarambi, “Spatial Features for Handwritten Kannada and English” IJCA Special Issue on “Recent Trends in Image Processing and Pattern Recognition”

Index Terms

Computer Science

Image Processing
Key words

Optical Character Recognition (OCR)          ANN
Document Scanning

Image Processing