OCR - Optical character recognition

OCR is widely used in everyday life. The number of characters read ranges from a few (e.g. account numbers on a payment order) to many (e.g. all the books in the world [1]). The steps in an OCR system are shown in the figure from [2] below.

Your task is to implement an OCR system. Focus on the Feature extraction and Classification steps. Read [2] to get information on different features and use a classification method of your choice.

Start with a small amount of printed characters, for example the numbers 0-9. If there is time, continue with hand-written text.
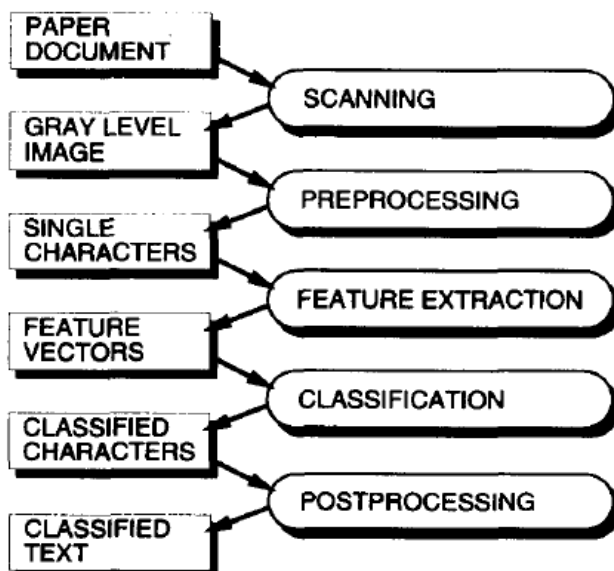

Fig. 2. Steps in a character recognition system.

[1] http://books.google.com/

[2] Due Trier et al. "Feature extraction methods for character recognition-A survey". Pattern recognition, Vol. 29, Issue 4. Pp 641-662, 1996

Contact: Artur Chodorowski

E-mail: artur.chodorowski@chalmers.se