

APPLICATION OF OPTICAL CHARACTER RECOGNITION FOR ARABIC PRINTED

Muhammad Faisal, S.Kom
mameck@n2mail.com

Sekolah Tinggi Teknik Surabaya
Surabaya, Indonesia
September, 2000

ABSTRACT

Every language has different characteristics and one of them is the form of its writing. One of the language that has a different font from Latin (roman) is Arabic language. Arabic writing has a unique characteristic and needs a different handling from the Latin writing.

This Final Project explains about the design and implementation of Arabic printed recognition system. The observation and handling of specific things that only happens in Arabic characters, such as hole, 'dot', and the separation of connected font, also explained. The character segmentation process consists of two parts; single character segmentation and connected character segmentation. Single Arabic character can have a different shape depends on its position in the writing.

The recognition method used in this Final Project are pixel population method and the approach of character profile analyzing through left, right, top and bottom projection. This Final Project also used the method of analyzing specific features in every Arabic characters. The result of this recognition is a text file with Arabic fonts.