

Sabri A. Mahmoud and Ashraf S. Mahmoud,
“Arabic Character Recognition using Modified Fourier Spectrum (MFS).”
Proceeding of the International Conference Geometric Modelling & Imaging-New Trends,
5-6 July, 2006, London, UK, Page(s) 155 - 159.

Key words: Arabic Character Recognition, OCR, Modified Fourier Spectrum, Fourier Transform

Arabic Character Recognition using Modified Fourier Spectrum (MFS)

Sabri A. Mahmoud¹, Ashraf S. Mahmoud²

¹Information and Computer Science, ²Computer Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, {masaad, ashraf}@ccse.kfupm.edu.sa}

Abstract

Arabic character recognition algorithm using Modified Fourier Spectrum (MFS) is presented. The MFS descriptors are estimated by applying the Fast Fourier Transform (FFT) to the Arabic character primary part contour. Ten descriptors are estimated from the Fourier spectrum of the character primary part contour by subtracting the imaginary part from the real part (and not from the amplitude of the Fourier spectrum as is usually the case). These descriptors are then used in the training and testing of Arabic characters. The computation of the MFS descriptors requires less computation time than the computation of the Fourier descriptors. Experimental results have shown that the MFS features are suitable for Arabic character recognition. Average recognition rate of 95.9% was achieved for the model classes. The analysis of the errors indicates that this recognition rate can be improved by using the “hole” feature of a character and use cleaning corrupted data.

Additionally, three characters can have a zigzag like stroke (Hamza). The dots and Hamza are called secondaries and they may be located above or below the character primary. Written Arabic text is cursive in both machine-printed and hand-written text. The shape of an Arabic character depends on its position in the word; a character might have up to four different shapes depending on its position. The character size varies according to its position in the word. The following figures show some of the characteristics of Arabic text.

- (a) بَعْضُ صِفَاتِ اللُّغَةِ الْعَرَبِيَّةِ
(b) ح د خ

Figure 1. (a) An example of an Arabic sentence indicating some characteristics of Arabic text; (b) Three Arabic characters with similar (same)