

Jordan University of Science and Technology

Hardware Implementation of Web Based Arabic Optical Character Recognition Units

Authors: Osama Al-Khaleel, Inad Aljarrah, Abdelrahman Idries, Khaldoon Mhaidat

Abstract: Optical character recognition (OCR) is an important application in the field of pattern recognition. It extracts text from an image document and saves it in an editable form. Examples where OCR is used include library digitization and text searching in scanned documents. Web based applications are main tools for data processing over the net. However, implementing such applications in dedicated hardware systems would increase performance and reliability by many folds over software implementation. In this paper, we present a detailed hardware implementation of the features extraction and character matching units of an Arabic optical character recognition (AOCR) system. The hardware implementation of each of these two units is described in VerilogHDL and functionally tested using ISim from Xilinx. Furthermore, each implementation is synthesized using Xilinx ISE 13.1 targeting Xilinx Spartan6 FPGA family. Experimental results show significant speed up in the hardware implementations over software ones. We further, explore the possibility of accessing these systems over the Web. Thus, they are beneficial to wider range of people.