

THE IMPLEMENTATION OF SPEECH RECOGNITION SYSTEMS ON FPGA-BASED EMBEDDED SYSTEMS WITH SOC ARCHITECTURE

SHING-TAI PAN¹, CHIH-CHIN LAI² AND BO-YU TSAI²

¹Department of Computer Science and Information Engineering

²Department of Electrical Engineering

National University of Kaohsiung

No. 700, Kaohsiung University Rd., Nanzih District, Kaohsiung 81148, Taiwan

{stpan; cclai}@nuk.edu.tw

Received June 2010; revised November 2010

ABSTRACT. *An implementation of Artificial-Neural-Network (ANN) based speech recognition systems on the embedded platform is explored in this paper. An FPGA chip is adopted as the hardware of the embedded platform with the architecture of SOC. This makes the speech recognition systems applicable on the voice activated systems in toys, games, smart phones, office devices, vehicular communications, etc. Because the fast Fourier transform (FFT) is an important operation in speech recognition, in which a great number of floating-point arithmetic operations are performed, it takes a long time for speech recognition due to the limitation on the computation capability of the embedded platform. In this paper, we use Integer FFT to replace the Floating FFT so that the speed of speech recognition is enhanced without influencing the recognition rate. The experimental results from the FPGA platform reveal that the speech recognition rate of the proposed hardware implementation methods is better than that in existing literatures.*

Keywords: Speech recognition, Embedded system, SOC architecture, FPGA

1. Introduction. The development of speech recognition attracts a lot of researches in recent years [1, 2, 3, 4, 5, 6]. With the development and maturity of speech recognition technique, speech becomes an important tool of man-machine interface. This speech interface is increasingly used in office automation, factory automation and home automation devices [2, 3]. The methods of speech recognition are mainly categorized as three ways: the earliest one is the Dynamic Time Warping (DTW) method [7], which uses time difference in a speech sound frame to achieve the speech recognition. Then, the Artificial Neural Network (ANN) [8] is applied on the speech recognition and then replaces DTW in most of the applications of speech recognition. Finally, Hidden Markov Model (HMM) [9, 10] appears, which uses the statistical method to enhance the speech recognition rate.

Due to the vast improvement on the manufacture of ICs in recent years, the capacity and the calculation ability of a chip are dramatically increased. Consequently, the speech recognition can be implemented in a chip rather than in a computer. This makes the embedded speech recognition systems in many consuming electronics realizable and then applicable on the voice activated systems in toys, games, consumer electronics and office devices [3]. Especially, due to the trend of the recent development in consumer electronics, the speech recognition systems are always embedded in smart phones, vehicular communications and intelligent houses, etc. In [3], a micro-controller 8051 chip is adopted for implementation of a speech recognition system. However, in order to have more extensive applications of speech recognition in consuming electronics, a platform with better computational ability and flexibility, such as embedded systems with SOC structure, is