

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอการออกแบบระบบรู้จำเสียงพูดคำโดดบนอุปกรณ์ FPGA ซึ่งมีกระบวนการทำงานสามขั้นตอน คือ การประมวลสัญญาณเบื้องต้น การดึงค่าลักษณะสำคัญด้วยวิธีการหาค่าสัมประสิทธิ์เคปสตรัมบนความถี่เมล และการรู้จำรูปแบบโดยใช้วิธีของโครงข่ายประสาทเทียม สำหรับขั้นตอนการออกแบบระบบบนเครื่องคอมพิวเตอร์ใช้โปรแกรม MATLAB ในการสร้างและฝึกโครงข่ายประสาทเทียมเพื่อเก็บค่าถ่วงน้ำหนัก ส่วนขั้นตอนการออกแบบระบบบนอุปกรณ์ FPGA ใช้โปรแกรม Xilinx Platform Studio สร้างหน่วยประมวลผล MicroBlaze ขึ้นภายในอุปกรณ์ FPGA โดยระบบของ MicroBlaze จะเชื่อมต่อกับชุดวงจรแปลงสัญญาณอนาล็อกเป็นดิจิทัลและชุดหน่วยความจำซึ่งใช้เก็บค่าถ่วงน้ำหนัก ในการเก็บตัวอย่างเสียงพูดคำโดดจากผู้พูด 15 คน จำนวน 1750 คำ แบ่งเป็นสามกลุ่ม คือ กลุ่มตัวอย่างสำหรับการฝึกจำนวน 700 คำ กลุ่มทดสอบการรู้จำแบบขึ้นกับผู้พูดจำนวน 700 คำ และกลุ่มทดสอบการรู้จำแบบไม่ขึ้นกับผู้พูดจำนวน 350 คำ ในขั้นตอนการทดสอบประสิทธิภาพการรู้จำ พบว่าการรู้จำแบบขึ้นกับผู้พูดได้ผลการรู้จำเฉลี่ยร้อยละ 99.57 ส่วนการรู้จำแบบไม่ขึ้นกับผู้พูดได้ผลการรู้จำเฉลี่ยร้อยละ 91.99

233805

Abstract

This thesis proposes a design of an FPGA-based isolated word recognition system, which consists of three processes: (i) preprocessing (ii) feature extraction using Mel-Frequency Cepstral Coefficients (MFCC) and (iii) pattern recognition using Artificial Neural Networks (ANN). MATLAB computer program is used to create and train the neural networks to obtain the weight and bias. The process of FPGA design uses Xilinx Platform Studio (XPS) to create a MicroBlaze system, which will interface with an Analog to Digital Converter (ADC) module and a flash memory used for storing the weight and bias. Speech samples (1750 samples) recorded from 15 speakers are used. These samples are divided into three sets, the training set (700 samples), the speaker-dependent testing set (700 samples) and the speaker-independent testing set (350 samples). The experiment result shows that the recognition performance of the speaker-dependent testing is about 99.57% while the speaker-independent testing is about 91.99%.