

ในการรู้จำเสียงสำหรับบางภาษาตัวอย่างเช่น ภาษาไทย ภาษาเวียดนาม และภาษาจีน เสียงวรรณยุกต์มีบทบาทสำคัญต่อการแยกความหมายของคำศัพท์ ส่งผลให้การตรวจหาเสียงวรรณยุกต์ถูกนำมาใช้เพื่อช่วยรับรองความถูกต้องของคำศัพท์ ดังนั้นวิทยานิพนธ์นี้จึงนำเสนอการรู้จำเสียงแบบคำโดดด้วยการตรวจหาเสียงวรรณยุกต์สำหรับสถาปัตยกรรมวงจรรวมความจุสูงมาก การรู้จำเสียงถูกแบ่งออกเป็น 3 กระบวนการคือ การวิเคราะห์สัญญาณเสียง การรู้จำเสียง และการตัดสินใจสุดท้าย การวิเคราะห์สัญญาณเสียงประกอบไปด้วย ขั้นตอนการตรวจหาระดับเสียงโดยวิธีฟังก์ชันผลต่างของขนาดเฉลี่ย และการวิเคราะห์สัมประสิทธิ์เซปสตรัลบนสเกลเมก การรู้จำเสียงโดยใช้แบบจำลองมาร์คอฟแบบซ่อนถูกนำมาใช้ในกระบวนการรู้จำ ในกระบวนการดังกล่าวการตรวจหาเสียงวรรณยุกต์ถูกรวมเข้าไปในแต่ละขั้นตอนกระบวนการคำนวณของการรู้จำคำศัพท์ นอกจากนี้วิธีการที่นำเสนอยังสามารถพิสูจน์บนสถาปัตยกรรมวงจรรวมความจุสูงมาก ที่ออกแบบจากแบบจำลองมาร์คอฟแบบซ่อนที่มีการคำนวณแบบขนาน ส่วนการตัดสินใจจะถูกคำนวณโดยการรวมผลลัพธ์จากกระบวนการก่อนหน้าเข้าด้วยกัน เพื่อเป็นการทดสอบประสิทธิภาพของการรู้จำที่นำเสนอ การทดลองได้ถูกจัดขึ้นและทดสอบกับคำศัพท์ภาษาไทยที่เลือกมาจากคำสั่งเครื่องควบคุมโทรทัศน์จำนวน 29 คำ และคำศัพท์ที่มีเสียงวรรณยุกต์ออกเสียงใกล้เคียงกันจำนวน 37 คำ ผลการทดสอบแสดงผลการพัฒนาประสิทธิภาพด้านความถูกต้องที่เพิ่มขึ้นร้อยละ 3.68 สำหรับคำสั่งเครื่องควบคุมโทรทัศน์ และร้อยละ 7.57 สำหรับคำศัพท์ที่มีเสียงวรรณยุกต์ออกเสียงใกล้เคียงกัน เมื่อเปรียบเทียบกับการรู้จำแบบเดิม

In speech recognition for some languages such as Thai, Vietnamese and Chinese, tone plays a key role in meaning classification. Tone detection function is needed in order to guarantee the corrected words. This thesis proposes isolated-word speech recognition with tone detection function for a VLSI architecture. The recognition is divided into 3 processes; speech analysis, speech recognition and final decision. The speech analysis consists of AMDF-based pitch detector and MFCC analysis. HMM-based speech recognition is used in recognition process. In this process, tone detection function is added into each computation of word recognition. Moreover, the proposed method is verified by implementing on a VLSI architecture designed with HMM parallel computation. The decision is computed by combining the output from the previous processes together. To evaluate the performance of the proposed method, the experiment is set and performed with 29 Thai words selected from TV remote control commands and 37 Thai indistinct tonal words. The results reveal the improvement of accuracy rate 3.68 percent for remote control commands and 7.57 percent for indistinct tonal words, comparing with the conventional method.