

วิทยานิพนธ์ฉบับนี้ได้นำเสนอระบบรู้จำภาษาไทยโดยการใช้วิธีการแบ่งแบบซอฟต์แวร์ โดยวิธีการแบ่งแบบซอฟต์แวร์นี้ได้อ้างอิงถึงคุณลักษณะของภาษาไทย ซึ่งจะมีสระเป็นแกนกลางของพยางค์ โดยได้ใช้ลักษณะเด่น คือ สัมประสิทธิ์เมลฟรีควเอนซีเคปสตรัมเปรียบเทียบกับค่าการทำนายเชิงเส้นจากเมลฟรีควเอนซี สำหรับหน่วยเสียงพยัญชนะต้น สระ และตัวสะกด และได้ใช้แบบจำลองฮิดเดนมาร์คอฟในการรู้จำหน่วยเสียง โดยที่มีหน่วยเสียงพยัญชนะต้น 21 กลุ่ม หน่วยเสียงสระ 18 กลุ่ม หน่วยเสียงตัวสะกด 9 กลุ่ม สำหรับเสียงวรรณยุกต์ที่มี 5 กลุ่ม ได้ใช้ลักษณะเด่น คือ การปรับเส้นโค้งและใช้โครงข่ายประสาทเทียมในการรู้จำ จากการทดสอบกับผู้พูด 30 คน ทั้งแบบขึ้นกับผู้พูดและไม่ขึ้นกับผู้พูด ค่าความถูกต้องที่ได้เมื่อใช้สัมประสิทธิ์เมลฟรีควเอนซีเคปสตรัมเป็นลักษณะเด่นได้ค่าความถูกต้องรวม 3 อันดับแรก 63.82%, 67.71% และ 72.08% เมื่อใช้ค่าการทำนายเชิงเส้นจากเมลฟรีควเอนซีได้ค่าความถูกต้องรวม 3 อันดับแรก 62.32%, 67.86% และ 79.42% ตามลำดับเรียงจากพยัญชนะต้น สระ และตัวสะกด และค่าถูกต้องรวม 3 อันดับแรกของเสียงวรรณยุกต์คือ 89.69% เมื่อใช้ผู้พูดที่ใช้ในการฝึกฝนระบบรู้จำ ในงานวิจัยนี้ได้ใช้ค่าความถูกต้อง 3 อันดับมาทำการตรวจสอบกับพจนานุกรมที่ได้สร้างขึ้นว่าเป็นพยางค์ที่มีใช้ในภาษาไทยกลางหรือไม่ หากพบว่าเป็นพยางค์ที่ไม่มีใช้ในภาษาไทยกลาง ก็จะทำการเลือกความถูกต้องอันดับต่อไปมาใช้ จากนั้นทำการตรวจเช็คกับพจนานุกรมจนได้พยางค์ที่มีใช้ในภาษาไทยกลาง โดยที่ระบบพจนานุกรมได้ช่วยเพิ่มค่าความถูกต้องของหน่วยเสียงได้ 0.9%

ABSTRACT

203616

In this thesis, a Thai phoneme recognition system with soft phoneme segmentation is proposed. The soft phoneme segmentation technique is based on the characteristics of Thai language in which the vowel is the core of a syllable. Two features, i.e., the Mel frequency cepstrum coefficients and the Mel frequency with perceptual linear prediction are used for comparison. The recognition system utilizes the discrete hidden Markov model to recognize the Thai phonemes, i.e., 21-class initial consonants, 18-class vowels, and 9-class final consonants, while the curve fitting is used as a feature for the 5-class tones and the neural network is applied as a classifier. The recognition data were recorded from 30 speakers. The overall top three classification rates for the initial consonant, the vowel, and the final consonant are 63.82%, 67.71%, and 72.08%, respectively, using the Mel frequency cepstrum coefficients. The rates of 62.32%, 67.86%, and 79.42%, are obtained using the Mel frequency with perceptual linear prediction. The overall top three classification rate of the tone is 89.69% obtained from the speakers in the training set. In this study, those overall top three classification rates are used to examine whether the possible syllable is a general Thai syllable in the dictionary. If those overall top three classification rates produce a syllable which is not in the dictionary, the next order classification rates are used instead. In conclusion, this dictionary system increases the correction of the phoneme classification rates by 0.9%.