

การรู้จำเสียงพูดภาษาไทยได้มีการวิจัยกันมานานแล้ว แต่ยังไม่มีการวิจัยใดมุ่งประเด็นการเพิ่มจำนวนคำศัพท์ของระบบรู้จำเสียงพูดให้มีจำนวนมากขึ้น งานวิจัยจึงทำการพัฒนาหาแนวทางของระบบรู้จำเสียงพูดภาษาไทยที่สามารถขยายจำนวนคำศัพท์ที่รู้จำมากขึ้น โดยได้ตั้งขอบเขตของจำนวนคำศัพท์ไว้ที่จำนวนคำศัพท์ปานกลางคือ 1,000 คำ เสียงพูดเป็นแบบต่อเนื่องเฉพาะบุคคล ในการพัฒนาระบบรู้จำจะให้ระบบทำการรู้จำหน่วยเสียงของเสียงพูดภาษาไทย ซึ่งมีจำนวนคงที่เสมอไม่ขึ้นกับจำนวนคำศัพท์ กล่าวคือเสียงพูดภาษาไทยประกอบด้วยหน่วยเสียง พยัญชนะต้น สระ พยัญชนะสะกด และวรรณยุกต์ เท่ากับ 33 24 9 และ 5 เสียงตามลำดับ โดยใช้โครงข่ายสมองเทียมเป็นตัวรู้จำ สำหรับลักษณะเด่นของหน่วยเสียงที่ใช้จะเป็น MFCC พลังงาน เวลา และคาบเวลาพิตช์ การสอนโครงข่ายสมองเทียมจะใช้เสียงพูดต้นแบบคำศัพท์ 5 ชุด ๆ ละ 1,000 คำ และทดสอบเสียงพูดคำศัพท์ในกลุ่มทดสอบที่มีจำนวนเท่ากับกลุ่มต้นแบบ ผลการทดลองการรู้จำหน่วยเสียงพยัญชนะต้น สระ พยัญชนะสะกด และวรรณยุกต์ เท่ากับ 56.24% 79.84% 60.32% และ 82.16% ตามลำดับ ค่าการรู้จำหน่วยเสียงที่ได้จากโครงข่ายสมองเทียมจะนำไปใช้ในการรู้จำคำศัพท์ในฐานข้อมูลคำศัพท์ โดยถ้าคำศัพท์ใช้จำนวนแบบหน่วยเสียงที่ประกอบขึ้นเป็นคำศัพท์ 1 แบบ / คำศัพท์ จะได้ผลการรู้จำคำศัพท์ที่กลุ่มทดสอบเท่ากับ 64.22% แต่ถ้าเพิ่มจำนวนแบบเป็น 6 แบบ / คำศัพท์ จะได้ผลการรู้จำคำศัพท์เพิ่มเป็น 84.28%

Abstract

TE 131963

The research on thai speech recognition has been generally conducted without particular aspects on enlarging number of vocabularies on thai speech recognition. From this point leads to the attention of this research to study thai speech recognition with vocabulary enlargement in the moderate boundary of 1,000 words. The speech recognition shall be developed and recognized by neural network to learn thai speech especially continuous and dependent speech in constant practice since sound of thai speech consist of consonants , vowels and tone at level of 33 , 24 , 9 and 5 respectively. For this study MFCC, short-time energy and pitch period shall be outstanding sound unit. Master vocabulary sound speech shall be categorized into 5 groups with individual group of 1,000 words and tested under neural network in same amount of test group. The results reveal 56.24% of consonants recognition, 79.84% of vowel, 60.32% of consonants and 82.16% of tone. The value of sound recognition received from neural network shall be in usage of vocabularies database recognition. In case sound unit comprising to be 1 type/vocabulary Test result of vocabulary recognition shall be 64.22% and increased to be 84.28% when 6 type/vocabulary is in used.