

วิทยานิพนธ์นี้เสนอวิธีการ Pitch Extraction ในการรู้จำหน่วยเสียงวรรณยุกต์สำหรับภาษาไทยในสภาวะแวดล้อมที่มีเสียงรบกวน โดยใช้วิธี Autocorrelation Method using Center Clipping (AUTOC) ในการคำนวณหาค่าคาบเวลาพิทช์ ซึ่งจะทำให้การคลิปปอดของสัญญาณในระดับที่ต่างกัน ค่าคาบเวลาพิทช์ที่ได้จากการคลิปปอดของสัญญาณในระดับที่ต่างกันนี้จะถูกแปลงเป็นค่าความถี่มูลฐานซึ่งเป็นตัวบ่งชี้ระดับสูง-ต่ำของเสียง ซึ่งลำดับของความถี่มูลฐานที่ได้จะถูกปรับปรุงให้มีความต่อเนื่องของข้อมูลโดยใช้มีเดียเตอร์ จากนั้นทำการหาค่าการเปลี่ยนแปลงของความถี่มูลฐานของพิทช์นั้นๆเทียบกับเวลา โดยทำการควอนไทซ์การเบี่ยงเบนออกเป็น 3 ระดับตามทิศทางการเพิ่มขึ้นหรือลดลงของความถี่มูลฐาน ซึ่งค่าที่ได้จากการควอนไทซ์นี้จะถูกนำไปใช้เป็นข้อมูลฝึกสอนให้กับการสร้างแบบจำลองการรู้จำของหน่วยเสียงวรรณยุกต์ทั้ง 5 ระดับด้วยวิธี Hidden Markov Model (HMM)

งานวิจัยนี้ได้ทำการทดสอบการรู้จำหน่วยเสียงวรรณยุกต์ที่เป็นสัญญาณรบกวนเกาส์เซียนขาวเข้าไปในเสียงพูดวรรณยุกต์ในภาษาไทยพยางค์เดียวทั้ง 5 ระดับ เพื่อทดสอบความทนทานในการรู้จำเสียงวรรณยุกต์ในภาษาไทยต่อเสียงรบกวนและหาค่าการคลิปปอดของสัญญาณที่เหมาะสมสำหรับการหาค่าคาบเวลาพิทช์ในสภาวะที่มีเสียงรบกวน โดยใช้ข้อมูลเสียงที่ได้จากเพศชาย 5 คนและเพศหญิง 5 คน ผลที่ได้คือการคลิปปอดของสัญญาณที่ระดับ 30 เปอร์เซ็นต์ให้ผลการรู้จำเสียงวรรณยุกต์ภาษาไทยแม่นยำถึง 90.00 เปอร์เซ็นต์ ในระดับสัญญาณต่อสัญญาณรบกวนที่ 10 dB

This thesis presents pitch extraction method for recognition of Thai tone in a noisy environment. The Autocorrelation using Center Clipping (AUTOC) is used to find the pitch period with different levels of clipping. The sequence of fundamental frequency from each word is smooth out by using median filter. The observed sequence of pitch levels are preprocessed to find the pitch differences and the sequence of pitch differences are then grouped into three quantized levels. The resultant sequence is used as bases for training a Hidden Markov Model and recognition of 5 tones.

The white gaussian noise is added into the in isolated monosyllable Thai word to find optimal estimation clipping level of signal for pitch period calculation. The tests are carried out on 5-male and 5-female Thai subject. It shows that clipping level at 30 percent is accuracy of recognition to 90 percent at signal to noise ratio of 10 dB.