

บทคัดย่อ

169736

วิทยานิพนธ์นี้นำเสนอวิธีในการรู้จำหน่วยเสียงสระเสียงเดียวในภาษาไทย และหาวิธีการลดมิติของความเข้มแถบความถี่วิกฤติ(Critical band intensities : CBI) บนสเกลความเข้มแถบความถี่วิกฤติ เพื่อเพิ่มประสิทธิภาพในการรู้จำโดยลดความยุ่งยากซับซ้อน และเวลาในกระบวนการคำนวณลง หน่วยเสียงสระที่จะทำการรู้จำคือเสียง อี, เอะ, แอะ, อี, เออะ, อะ, อุ, โอะ และ เอาะ แต่ละหน่วยเสียงจะถูกนำไปคำนวณสัมประสิทธิ์ของการประมาณเชิงเส้น แล้วทำการแปลงสเปกตรัมของการประมาณเชิงเส้นที่ได้ไปอยู่ในสเกลความเข้มแถบความถี่วิกฤติ ซึ่งจะได้เวกเตอร์คุณสมบัติของความเข้มแถบความถี่วิกฤติ (CBI) 18 มิติ นำมาหาความเข้มแถบความถี่วิกฤติที่มีประสิทธิภาพโดยวิธีการหาค่าสูงสุดของความถูกต้องในการรู้จำเพิ่มทีละมิติ ซึ่งเป็นวิธีการลดขนาดมิติของเวกเตอร์คุณสมบัติที่จะถูกนำไปใช้สร้างและทดสอบแบบอ้างอิงต่อไป การทดสอบแบบอ้างอิงจะอาศัยเทคนิค K-Nearest Neighbor (KNN) ในการตัดสินใจแยกแยะเสียงสระ

ในงานวิจัยนี้เราสามารถลดขนาดมิติที่ใช้ในการรู้จำในแบบอ้างอิงต่างๆ ได้ประมาณ 52 % โดยค่าเปอร์เซ็นต์ความถูกต้องของทุกแบบอ้างอิงจะสูงขึ้นเมื่อเทียบกับการรู้จำเสียงสระเสียงเดียวในภาษาไทยแบบใช้ข้อมูลจากทุกมิติแบบเดิม

ABSTRACT

169736

This thesis presents a recognition model for 9 unmixed vowels in Thai spoken language and a method to reduce the dimension of critical band intensities on Bark Scale. The nine unmixed vowels to be recognized are “i”, “e”, “ε”, “ω”, “γ”, “a”, “u”, “o”, and “ɔ̄”. Each unmixed vowel is represented by 18-dimensional feature vector of Critical Band Intensities (CBI) of the vocal tract transfer function, which is determined from Linear Predictive Coding Coefficients. These 18-dimension vectors are examined to find the effective CBI components that maximize the accuracy. Reference models are tested by applying unknown sounds to the recognition procedure using KNN technique for classification.

In this research, we reduce up to 52 % of the dimensions and get high accuracy in all reference models when compared with the full blown 18-dimension recognition model for 9 unmixed vowels in Thai spoken language.