

ระบบอัจฉริยะสำหรับตัดแปลงสำเนียงภาษาอังกฤษของสัญญาณเสียงพูด

โดย นายภูวดล หนูชัย

นายวรมธ จำปาเทศ



บทคัดย่อ

ระบบตัดแปลงสำเนียงภาษาอังกฤษของสัญญาณเสียงพูด ในโครงงานนี้ถูกออกแบบเพื่อปรับสำเนียงเสียงพูดภาษาอังกฤษ ของ User โดยยังคงอัตลักษณ์เสียงพูดของ User ไว้ ระบบนี้ใช้การบันทึกสัญญาณเสียงพูดของ User และอักขร และสคริปต์ของเสียงที่พูดนั้น จากนั้นระบบจะสร้างสัญญาณเสียงอ้างอิง โดยใช้ระบบแปลงข้อความเป็นเสียงพูด (TTS) จากนั้นสัญญาณเสียงพูดของ User จะถูกปรับให้ตรงกับสัญญาณอ้างอิงโดยใช้ Dynamic-Time Warping (DTW) เพื่อให้มีการ Synchronization กันอย่างถูกต้อง เพื่อให้เสียงสระ หรือเสียงพยัญชนะเดียวกัน อยู่ในตำแหน่งที่ตรงกัน ในการปรับสำเนียง ระบบใช้แพ็คเกจ Python ชื่อว่า WORLD ซึ่งวิเคราะห์และแยกสัญญาณทั้งสองออกเป็นสามส่วน คือ Spectral Envelope (sp) , Aperiodicity (ap) และ Fundamental Frequency (f0) ความถี่มูลฐาน จากสัญญาณใหม่ ที่จะตัดแปลง จะถูกสร้างขึ้น โดยมีระยะเวลา และจังหวะที่ตรงกันกับสัญญาณอ้างอิง และค่า f0 จะถูกปรับระดับความสูงเสียง และอนุพันธ์ โดยยังคงรักษาอัตลักษณ์เสียงของ User ระบบนี้ได้รับการทดสอบกับ เสียงพูดภาษาอังกฤษสำเนียงไทย ที่ตัดแปลงเป็นสำเนียงอเมริกัน และ แสดงให้เห็นว่ามีผลลัพธ์ที่สามารถต่อยอดได้ อย่างไรก็ตาม มีข้อจำกัดของระบบ ที่ได้วิเคราะห์ และอภิปรายในรายงานนี้ โดยรวมแล้วระบบตัดแปลงสำเนียงเสียงพูดภาษาอังกฤษนี้ นำเสนอวิธีการที่มีประสิทธิภาพ ในการตัดแปลงสำเนียงภาษาอังกฤษ โดยยังคงรักษาอัตลักษณ์ของเสียงของ User ไว้

Intelligent system for modifying English accent of speech signals

By Mr. Phoowadol Thanoochai

Mr. Woramed Jumpatet



ABSTRACT

The English Accent Modification System presented in this project is designed to modify the accent of a user's speech while still maintaining their unique voice identity. The system utilizes a recording of the user's speech signals and the corresponding text transcript, generating a reference speech signal using a text-to-speech (TTS) system. The user's speech signal is then aligned with the reference signal using dynamic time warping (DTW) to ensure synchronization. To modify the accent, the system employs the use of the Python package WORLD, which analyzes and decomposes both signals into three components - spectral envelope, aperiodicity, and fundamental frequency (f_0). The modified signal is then generated, with its duration and rhythm matching that of the reference signal, and its f_0 modified to match the pitch contour and derivative while still preserving the user's voice characteristics. The system has been tested on Thai accents, modified to the American accent, and has shown promising results. However, limitations of the system are also discussed in this report. Overall, the English Accent Modification System offers a unique and effective approach to modifying accents while preserving the user's voice identity.