

วิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อศึกษาวิธีการปรับปรุงอัตราการเรียนรู้จำเสียงพูดของผู้ไร้กล่องเสียงแบบพูดโดยใช้หลักทฤษฎีในแบบจำลอง Hidden Markov Model ที่ใช้เสียงผู้พูดปกติเป็นฐานข้อมูล โดยการปรับปรุงลักษณะเด่นของเสียง ซึ่งเทคนิคที่นำเสนอเป็นกระบวนการภาคหน้าเพื่อการปรับปรุงลักษณะเด่นของเสียงของผู้ไร้กล่องเสียงในงานวิจัยนี้คือ การแปลงลักษณะเด่นของเสียงและเทคนิคการนอร์มัลไลซ์สัมประสิทธิ์เซปตรัม ซึ่งในงานวิจัยนี้นำเสนอสามเทคนิคได้แก่ การนอร์มัลไลซ์เซปตรัมด้วยค่าเฉลี่ย การนอร์มัลไลซ์เซปตรัมด้วยค่าเฉลี่ยและค่าสูงสุด และการนอร์มัลไลซ์เซปตรัมด้วยค่าคะแนนมาตรฐาน ซึ่งกระบวนการแปลงลักษณะเด่นของเสียงเป็นการปรับปรุงลักษณะเด่นของเสียงของผู้ไร้กล่องเสียงโดยการแปลงลักษณะเด่นของเสียงผู้ไร้กล่องเสียงไปเป็นลักษณะเด่นของเสียงผู้พูดปกติโดยใช้กฎความสัมพันธ์ของคนทั้งสองเป็นเงื่อนไขในการแปลง ส่วนเทคนิคการนอร์มัลไลซ์สัมประสิทธิ์เซปตรัมเป็นเทคนิคที่ช่วยปรับปรุงลักษณะเด่นของเสียงโดยเพิ่มความคงทนของเวกเตอร์ลักษณะเด่นของเสียงพูดต่อการรบกวนอย่างง่าย จากการทดลองเมื่อใช้การแปลงลักษณะเด่นของเสียงปรับปรุงลักษณะเด่นของเสียงของผู้ไร้กล่องเสียงสามารถปรับปรุงอัตราการเรียนรู้จำเสียงพูดของผู้ไร้กล่องเสียงจากร้อยละ 32 เป็นร้อยละ 67 และเมื่อใช้เทคนิคการแปลงลักษณะเด่นของเสียงร่วมกับเทคนิคการนอร์มัลไลซ์สัมประสิทธิ์เซปตรัมทำให้อัตราการเรียนรู้จำเสียงพูดของผู้ไร้กล่องเสียงเพิ่มเป็นร้อยละ 75

The purpose of this thesis is to improve speech recognition rate for Thai esophageal speech in speech recognition system built by normal speech corpus. A feature conversion technique and three feature normalization techniques, called Cepstral Mean Normalization (CMN), Cepstral Mean Max Normalization (CMMN) and Cepstral Standard-Score Normalization (CZN) are applied as front-end process for improve esophageal speech features in this application. A feature conversion technique for improvement of esophageal speech features is done by changing esophageal speech features into normal speech features in order to improve a recognition rate of esophageal speech in a speech recognition system. It is trained by normal speech corpus based on Hidden Markov Models (HMMs). Three feature normalization techniques are applied in order to prevent variation bias in each signal. A system that combines feature conversion technique and cepstral normalization technique is proposed. From experimental results, the feature conversion technique can significantly improve the recognition rate of esophageal speech from 32% to 67%. Moreover, the recognition rate is raised up to 75% when using the concatenation of feature conversion technique and cepstral normalization technique