

ชื่อเรื่องวิทยานิพนธ์

การเพิ่มอารมณ์ในเสียงพูดสังเคราะห์โดยใช้แบบจำลอง
พีชชี

ผู้เขียน

นายปริญญา สุวรรณศรีคำ

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รศ. ขจรศักดิ์ คันธพนิต

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอวิธีการเพิ่มอารมณ์ในเสียงพูดสังเคราะห์โดยใช้แบบจำลองพีชชี เพื่อช่วยให้การพูดของหุ่นยนต์มีความใกล้เคียงกับมนุษย์มากขึ้น ในงานวิจัยนี้ได้ศึกษาว่าอารมณ์ในเสียงพูดนั้นสื่อออกมาทางองค์ประกอบของเสียงพูด 3 องค์ประกอบด้วยกันคือ ความดังของเสียงพูด พิตช์ และ ความเร็วของเสียงพูด [9] [12] [16] ดังนั้นการเพิ่มอารมณ์ในงานวิจัยนี้จึงมุ่งไปที่การปรับองค์ประกอบทั้งสาม โดยใช้ระบบพีชชีมาช่วยในการจำลองความสัมพันธ์ของอารมณ์กับองค์ประกอบของเสียงพูด และทำการสร้างเสียงที่พูดที่มีอารมณ์จากเสียงพูดที่เป็นกลาง (Neutral Speech) จากการที่ได้นำเสียงพูดที่ได้รับเพิ่มอารมณ์จากกระบวนการนี้ไปให้ผู้ฟังจำนวน 20 คน ฟังพบว่าผู้ฟังสามารถรับรู้อารมณ์โกรธได้ 60% สำหรับเสียงผู้ชายและ 65% สำหรับเสียงผู้หญิง รับรู้อารมณ์เศร้าได้ 65% สำหรับเสียงผู้ชายและ 65% สำหรับเสียงผู้หญิง รับรู้อารมณ์สนุกได้ 45% สำหรับเสียงผู้ชายและ 40% สำหรับเสียงผู้หญิง รับรู้อารมณ์กลัวได้ 40% สำหรับเสียงผู้ชายและ 40% สำหรับเสียงผู้หญิง

Thesis Title Adding Emotion to Synthesized Speech Using Fuzzy Model

Author Mr. Parinya Suwansrikham

Degree Master of Engineering (Electrical Engineering)

Thesis Advisor Assoc. Prof. Kajornsak Kantapanit

ABSTRACT

This thesis presents a method to add emotion to synthesized speech using fuzzy model to improve the speaking styles of a robot. In this research we have modeled emotions parameters in speech production. The parameters of speech are intensity, pitch and time duration of speech. To add emotion into speech, we modify the three parameters of sample speeches. Fuzzy system is introduced to model emotional parameters in speech production. We have created emotional speech from neutral speech. From the experiment with 20 listeners, it is found that the recognition rates of synthesized anger speech are 60% for male voice and 65% for female voice. For synthesized sad speech, recognition rates are 65% for male voice and 65% for female voice. For synthesized happy speech, recognition rates are 45% for male voice and 40% for voice speech. For synthesized fear speech, recognition rates are 40% for male voice and 40% for female voice.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved