29 DEC 1999

A PROTOTYPE OF VOICE PERSONALITY TRANSFORMATION

PONGRAT PASATCHASARTNUKOON

With compliments
of
บัณฑิตวิทยาลัย มหิดล

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE (COMPUTER SCIENCE) FACULTY OF GRADUATE STUDIES MAHIDOL UNIVERSITY 1999

ISBN 974-662-954-2

COPYRIGHT OF MAHIDOL UNIVERSITY

Copyright by Mahidol University

TH
P796 ปก
1999

43291 ค.2

3936994 SCCS/M : MAJOR : COMPUTER SCIENCE ; M.Sc. (COMPUTER)
KEY WORD : SOUND TRANSFORMATION / LINEAR PREDICTIVE CODING
PONGRAT PASATCHASARTNUKOON : A PROTOTYPE OF VOICE PERSONALITY TRANSFORMATION. THESIS ADVISORS : ASSOC. PROF. SUPPACHAI TANGWONGSARN, Ph.D., ASST. PROF. DAMRAS WONGSAWANG, Ph.D. 111 P. ISBN 974-662-954-2

This research work presents a prototype of voice personality transformation. This means changing the speech sound from one to another without changing the content. For example, changing the sound from a female voice to a male voice, or a child's voice, but with different tonation from the original person, so that one could not distinguish him/her as the originator.

The prototype was developed and based on two key features of sound transformation, namely pitch period and formant frequency. The model system consists of three main processes : sound analysis, sound transformation and sound synthesis. The sound analysis extracts the features that represent the value of speech signal by using the linear predictive coding algorithm (LPC). Then in the sound transformation, we attempt to change the key features with the preset percentage values depending on the desired voice types. Finally in the sound synthesis, we use new values of the features obtained from the previous step to create speech sounds.

The experiment done by transforming a set of recorded speech sounds with the preset percentage values of pitch period and formant frequency, confirms our approach is justifiable. Moreover, the experimental result shows the gender of speech signal can be detected by observing its pitch period, and also the sound quality can be controlled by comparing the values of formant frequency with those of the threshold.

Copyright by Mahidol University

3936994 SCCS/M : สาขาวิชา : วิทยาการคอมพิวเตอร์ ; วทม. (วิทยาการคอมพิวเตอร์)
คำสำคัญ : การเปลี่ยนแปลงเสียงพูด / การทำนายเชิงเส้น

พงศ์รัฐ เภสัชศาสตร์นุกูล : ต้นแบบการเปลี่ยนลักษณะการพูดของมนุษย์ (A PROTOTYPE OF VOICE PERSONALITY TRANSFORMATION). คณะกรรมการควบคุมวิทยานิพนธ์ รศ. ศุภชัย ตั้งวงศ์ศานต์, Ph.D., ผศ. ดำรัส วงศ์สว่าง, Ph.D. 111 หน้า. ISBN 974-662-954-2

การวิจัยเรื่อง Voice personality transformation คือการเปลี่ยนแปลงเสียงพูดจากเสียงหนึ่งไปเป็นเสียงอื่น โดยไม่มีการเปลี่ยนแปลงเนื้อหาของคำพูดและไม่สามารถจดจำเสียง Original ได้ เช่นการเปลี่ยนแปลงเสียงพูดจากเพศหญิงเป็นเพศชายหรือเสียงเด็ก

ต้นแบบของ Sound transformation ได้พัฒนาและใช้ลักษณะเฉพาะที่สำคัญ คือ Pitch period และ Formant frequency ระบบนี้ประกอบด้วย 3 ส่วนหลัก ได้แก่ การวิเคราะห์เสียง (Sound Analysis), การเปลี่ยนเสียง (Sound Transformation) และการสังเคราะห์เสียง (Sound Synthesis) เริ่มจากการวิเคราะห์เสียง ทำการแยกค่าลักษณะเฉพาะโดยวิธีของการทำนายเชิงเส้น (Linear Predictive Coding) ที่เป็นการแทนค่าของเสียงพูดออกจากสัญญาณเสียงพูด และทำการเปลี่ยนแปลงค่าของลักษณะเฉพาะด้วยค่าที่กำหนดเพื่อให้เป็นเสียงประเภทที่ต้องการ ท้ายที่สุดทำการสังเคราะห์ค่าลักษณะเฉพาะที่ได้ทำการเปลี่ยนแปลงเพื่อสร้างเป็นสัญญาณเสียงพูด

ผลการทดลองโดยการนำสัญญาณเสียงพูด ที่ทำการบันทึกไว้มาผ่านการทำ Sound transformation โดยการเปลี่ยนแปลงค่า Percentage ต่างๆ ตามที่กำหนดไว้ ของค่า Pitch period และ formant frequency ผลที่ได้สามารถเปลี่ยนลักษณะของเสียงเป็นเสียงอื่นที่ไม่สามารถจดจำเสียงเดิมได้โดยไม่มีการเปลี่ยนแปลงเนื้อหาของคำพูด และทำการตรวจพบได้ว่าสัญญาณเสียงพูดนั้นเป็นเสียงของเพศใดจากการสังเกตค่า Pitch period และสามารถที่จะทำการควบคุมคุณภาพของเสียงที่ผ่านการเปลี่ยนเสียง โดยการเปรียบเทียบค่าของ formant frequency กับค่าของ threshold

Copyright by Mahidol University