

SPEECH ANALYSIS USING THE WAVELET TRANSFORM

SOMPRASONG NAKNAIMEUNG

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

2000

ISBN 974-665-087-4

COPYRIGHT OF MAHIDOL UNIVERSITY

TH
S697b
2000
C.2

.1129A273x C.2

3937486 SCPY/M: MAJOR : PHYSICS ; M.Sc. (PHYSICS)

KEY WORDS : WAVELET / VOICE RECOGNITION / PITCH PERIOD

SOMPRASONG NAKNAIMEUNG : SPEECH ANALYSIS USING THE WAVELET TRANSFORM. THESIS ADVISORS : PIYAPONG SITHIKONG, Ph.D., PONGSAKORN SAIPECH, Ph.D. 54 P. ISBN 974-665-087-4

In this study we use wavelet analysis to attempt to identify the owner of a voice. This is done by trying to find quantities derived from the wavelet analysis that can be used to reliably distinguish the speech waveforms from different speakers. Two approaches are used. The first is based on the method of Kadambe and Boudreaux-Bartels and is referred to as the pitch period method. It relies on obtaining the pitch period as a function of time for the waveform. The second method relies on obtaining quantities that can be used to distinguish speech waveforms from different people directly from a list of the values of the wavelet coefficients. We call it the direct method. We find that our pitch period method does not give good enough results to be used for voice identification. However, we show that our direct method can be used as a basis for identifying voices.

3937486 SCPY/M: สาขาวิชา : ฟิสิกส์ ; วท.ม. (ฟิสิกส์)

สมประสงค์ นาคในเมือง : การวิเคราะห์เสียงพูดคนโดยใช้การแปลงแบบเวฟเล็ท (SPEECH ANALYSIS USING THE WAVELET TRANSFORM).

คณะกรรมการควบคุมวิทยานิพนธ์ : ปิยพงษ์ สิทธิคง, Ph.D., พงศกร สายเพชร, Ph.D.
54 หน้า. ISBN 974-665-087-4

การวิจัยนี้เป็นการศึกษาความเป็นไปได้ในการจำแนกเสียงคน โดยใช้การวิเคราะห์เชิงเวฟเล็ท (wavelet analysis) จากการวิเคราะห์สัญญาณเสียงพูดได้ค่าความถี่และตำแหน่งของเวลาที่ความถี่นั้นเกิดขึ้น ค่าเหล่านี้สามารถนำไปวิเคราะห์เพื่อหาลักษณะเฉพาะของสัญญาณเสียงที่อาจใช้จำแนกเสียงพูดของแต่ละคนได้ จากการศึกษาการวิเคราะห์เชิงเวฟเล็ทพบว่า ฟังก์ชันฐาน (basis function) ที่ใช้ได้ดีและง่ายที่สุดก็คือ ฮาร์ เวฟเล็ท (Haar wavelet) การวิจัยนี้ใช้ 2 วิธี ในการหาลักษณะเฉพาะของเสียง คือ การหาค่าคาบของระดับเสียง (pitch period) และ การหาค่าสัมประสิทธิ์ของเวฟเล็ท (wavelet coefficients) จากผลการวิจัยพบว่าเสียงพูดหลายๆ ครั้งที่เป็นค่าเดียวกันของคนต่างกันมีค่า pitch period ต่างกันและเสียงพูดค่าเดียวกันของคนคนเดียวก็มีค่า pitch period ต่างกัน แต่เสียงพูดหลายๆ ครั้งที่เป็นค่าพูดเดียวกันของคนคนเดียวก็มีค่าสัมประสิทธิ์ของเวฟเล็ทใกล้เคียงกันและเสียงพูดของคนที่แตกต่างกันมีค่าสัมประสิทธิ์ของเวฟเล็ทต่างกัน จึงเป็นไปได้ที่จะจำแนกเสียงคนโดยใช้ค่าสัมประสิทธิ์ของเวฟเล็ท