รายงานการวิจัยฉบับสมบูรณ์ โครงการวิจัยเรื่อง

การศึกษาคุณสมบัติอคูสติกส์ของเสียงที่แปรผันกับอารมณ์ในฐานะตัวบ่งบอกระดับแนวโน้มการฆ่าตัวตายในบุคคลที่ซึมเศร้า

A Study of Affect-Related Vocal Acoustics in Speech as Indicators of Suicidal Risk in Depressed Persons

ดำเนินการวิจัยโดย

ดร. ทวีศักดิ์ ยิ่งถาวรสุข

รายงานนี้ได้รับการสนับสนุนจากเงินงบประมาณ ประจำปี 2552

บทคัดย่อ

คุณสมบัติเชิงความถี่ที่เรียกว่า Spectral Characteristics ของสัญญาณเสียงพูดได้ถูกศึกษาอย่างกว้างขวางทั้งในอดีตและปัจจุบัน และมีการรายงานถึงข้อมูลผลงานวิจัยที่เกี่ยวข้องที่เป็นประโยชน์จากการศึกษาเกี่ยวกับสถานะทางอารมณ์ในบุคคลที่ป่วยด้วยโรคซึมเศร้า พลังงานเชิงความถี่สเปกตรัมหรือที่เรียกว่า Spectral Energy ซึ่งกระจายอยู่ในแถบความถี่ต่างๆ ในช่วงความถี่เสียงพูดที่มีค่าไม่เกิน 5KHz ถูกพิจารณาสำหรับการวิเคราะห์คุณสมบัติทางอคูสติกส์ในข้อมูล ซึ่งเปิดเผยให้เห็นถึงคุณสมบัติที่สามารถใช้พยากรณ์และตีความสอดคล้องกับระดับอาการของโรคที่ผู้ป่วยเผชิญอยู่ พิจารณาได้จากการเคลื่อนของพลังงานจากแบนด์วิดท์หนึ่งไปสู่อีกแบนด์วิดท์หนึ่ง ขนาดของพลังงานที่เปลี่ยนแปลงในแต่แบนด์วิดท์ งานวิจัยศึกษาการวิเคราะห์คุณสมบัติอคูสติกส์ของเสียงพูดมุ่งเน้นศึกษา Porwer Spectral Density (PSD) เป็นค่าประมาณพลังงานในสัญญาณเสียงพูด ซึ่งเป็นคุณสมบัติอคูสติกส์ถอดรหัสจากสัญญาณจากฐานข้อมูลสัญญาณเสียงพูด บันทึกจากกลุ่มผู้ป่วยฟื้นฟูจากการเป็นโรคซึมเศร้า (Remitted) ผู้ป่วยซึมเศร้า (Depressed) และผู้ป่วยซึมเศร้าเสี่ยงต่อการฆ่าตัวตาย (High-Risk Suicidal) ผลจากการคัดแยกกรองยืนยันให้เห็นว่า PSD สามารถจำแนกผู้ป่วยระหว่างกลุ่มต่างๆที่ศึกษาอย่างมีประสิทธิภาพ ผลการวิเคราะห์ยังแสดงให้เห็นถึงความแน่นอนของผลวิเคราะห์ที่ได้ ถึงความเปลี่ยนแปลง ความคลาดเคลื่อนของคุณสมบัติอคูสติกส์ของเสียงพูดที่สัมพันธ์กับอาการป่วยจากโรคซึมเศร้า โดยมีผลกระทบต่อการทำงานของระบบการกำเนิดเสียงพูดอย่างชัดเจน ในงานวิจัยด้านสุขภาพ ด้านการแพทย์เป็นที่ทราบกันดี สมองส่วนกลางถูกกระทบได้โดยตรงจากอาการป่วยโรคซึมเศร้า และรวมถึงโรคอื่นๆ ที่มีผลต่อพฤติกรรมทางอารมณ์ และส่งผลกระทบต่อเนื่องกับสุขภาพจิตของผู้ป่วย

KEY WORDS: Speech, Power Spectral Density, Depression, Classification

Abstract

Spectral characteristics of speech signal have formerly been studied and reported for its quantitative information in relating the emotional state in persons suffering depression to the level of symptom affect. The spectral energy distributing in the particular bandwidths of audio frequency less than 5 KHz extracted from the subjects' speech sample reveals as the distinct parameter with a property that can represent the severity of symptom in term of energy shifting and amount change found in this study empirically. The acoustical analysis focused on Power Spectral Density (PSD) was performed on speech samples extracted from three different databases collected from the remitted, depressed, and suicidal patient groups. The result of classifications suggests the significant potential of PSD's in separating all categorized patient groups which consistently verifies that the psychomotor disturbance affects speech production system in speaker, clinically known as the central element in depression and other high-risk emotional disorders.

KEY WORDS: Speech, Power Spectral Density, Depression, Classification

Table of Contents

Page

ABSTRACT 2
TABLE OF CONTENTS 3
LIST OF TABLES 4
LIST OF FIGURES 2
บทนำ 6
วัตถุประสงค์ของโครงการวิจัย 10
ขอบเขตของโครงการวิจัย 11
ประโยชน์ที่คาดว่าจะได้รับ 13
ทฤษฎีที่เกี่ยวข้อง 14
วิธีการดำเนินการวิเคราะห์สัญญาณ (Methodology) 17
ผลการทดลอง (Experimental Results) 19
ข้อวิจารณ์ (Discussion) 22
สรุปและข้อเสนอแนะ (Conclusion and Suggestion) 25
ผลงานวิจัยตีพิมพ์ (Publications) 26
Reference 31

List of Tables

Page

TAB.7.1 PSD DISTRIBUTION FOR ALL SUBJECTS – INTERVIEW SESSION..... 20
TAB.7.2 PSD DISTRIBUTION FOR ALL SUBJECTS – READING SESSION 20
TAB.7.3 SUMMARIZED STATISTICS – INTERVIEW SESSION 20
TAB.7.4 SUMMARIZED STATISTICS – READING SESSION 20
TAB.7.5 SUMMARIZED COMPARATIVE RECOGNITION PERFORMANCES FROM ANALYSES OF SPONTANEOUS SPEECH..... 21
TAB.7.6 SUMMARIZED COMPARATIVE RECOGNITION PERFORMANCES FROM ANALYSES OF AUTOMATIC SPEECH 21

List of Figures

Page

FIG. 5.1 SPEECH PRODUCTION SYSTEM 14
FIG. 6.1 EXTRACTION PROCESS 18
FIG.10.1 COPY OF PROCEEDING COVER 26
FIG.10.2 PHOTO COPY OF PROCEEDING PRESS 26
FIG.10.3 PHOTO COPY OF ORGANIZING COMMITTEES PAGE 27
FIG.10.4 PHOTO COPY OF PREFACE PAGE 27
FIG.10.5 PHOTO COPY OF TABLE OF CONTENTS 28
FIG.10.6 PHOTO COPY OF TABLE OF CONTENTS (CONT…) 28
FIG.10.7 PHOTO COPY OF THE PUBLISHED RESEARCH WORK PAGE 1 29
FIG.10.8 PHOTO COPY OF THE PUBLISHED RESEARCH WORK PAGES 2 & 3 29
FIG.10.9 PHOTO COPY OF THE PUBLISHED RESEARCH WORK PAGES 4 & 5 30

บทที่ 1 บทนำ

1. ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

จากผลงานวิจัยที่ถูกนำเสนอในรายงานการประชุมทางวิชาการระดับนานาชาติและวารสารตีพิมพ์ที่เกี่ยวกับปัญหาสุขภาพจิต และศาสตร์สาขาวิชาการด้านการประมวลผลสัญญาณเสียงพูด รวมถึงผลงานวิจัยที่ตีพิมพ์เผยแพร่ในเรื่องของ "อารมณ์ความรู้สึกของผู้พูดขณะเปล่งเสียง" (Emotional Speech) ได้แสดงให้เห็นการเชื่อมโยงที่สัมพันธ์กันระหว่างการเปลี่ยนแปลงของคุณสมบัติบางประการของตัวแปรทางอคูสติกส์ที่ปรากฏอยู่ในเสียงพูดกับสภาวะการแปรปรวนที่เกิดขึ้นทางอารมณ์ สภาวะจิต และความคิดของตัวผู้พูดซึ่งมีสาเหตุเนื่องมาจากความเครียด ความกังวล หรืออาการป่วยทางจิตปรากฏให้เห็น และได้รับการพิสูจน์จนเป็นที่ยอมในระดับหนึ่งถึงการเชื่อมโยงที่เกิดขึ้นดังกล่าว อิทธิพลจากการแปรปรวนที่เกิดขึ้นสามารถส่งผลกระทบต่อการทำงานในระบบกำเนิดเสียง (Speech Production System) ภายในตัวผู้พูดส่งผลให้เสียงที่เปล่งออกมาถูกแฝงด้วยอิทธิพลจากอารมณ์ สภาวะจิต รวมถึงอาการป่วยจากโรคทางจิต บุคคลผู้ฟังเสียงอาจสามารถสังเกตและรับรู้ในเบื้องต้นได้ถึงการเปลี่ยนแปลงในระดับความดังเสียง (Loudness) ความร่าเริง (Liveliness) ระดับเสียงที่ราบเรียบ (Monotone) ความชัดเจนสดใส (Clarification) ระดับพลังงานในเสียง (Energy in Sound) อัตราความเร็วในการพูด (Speaking Rate) อัตราการหยุดนิ่งระหว่างการพูดสนทนา (Pause rate) และช่วงเวลาโต้ตอบในการสนทนา (Responding Time) เป็นต้นกรณีการป่วยเนื่องจากโรคที่ส่งผลกระทบอย่างเห็นได้ชัดโดยตรงต่อระบบการกำเนิดเสียง เช่น อาการป่วยจากไข้หวัด เป็นตัวอย่างที่เห็นได้ชัดถึงผลกระทบของโรคซึ่งส่งผลกระทบต่อระบบการกำเนิดเสียงของตัวผู้ป่วยและทำให้เสียงผิดเพี้ยนไปจากเดิมอย่างเห็นได้ชัดเมื่อผู้สนทนาร่วมได้ฟังเสียงของผู้ป่วยไข้หวัดการใช้คอมพิวเตอร์ในการวิเคราะห์และแปลข้อมูลที่ได้จากแหล่งต่างๆ ที่เกี่ยวข้องกับผู้ป่วยที่เข้าการบำบัดสุขภาพจิต รวมถึงจิตแพทย์ที่มีประสบการณ์เป็นวิธีการเดียวที่ใช้ดำเนินการในปัจจุบัน การวิเคราะห์โดยอาศัยเครื่องมือคอมพิวเตอร์ สำหรับโครงงานวิจัยที่เสนอนี้ดำเนินการวิจัยมุ่งเน้นศึกษา และทำความเข้าใจเกี่ยวกับการเปลี่ยนแปลงของคุณสมบัติในตัวแปรอคูสติกส์ชนิดไดนามิกส์อันเนื่องมาจากอิทธิพลของอาการป่วยทางจิตโดยจะศึกษาเปรียบเทียบกับคุณสมบัติของอคูสติกส์ที่วิเคราะห์จากเสียงพูดของบุคคลที่ผ่านการวินิจฉัยโดยนักจิตเวชแล้วว่ามีสภาวะจิตที่เป็นปกติเป็นข้อมูลเพื่อการอ้างอิงระดับการเปลี่ยน

แปลงที่แตกต่างไปในสภาวะสังคมมีการแข่งขันสูงไม่ว่าในด้านการทำงาน การทำมาหาเลี้ยงชีพ การศึกษา การดำเนินชีวิตอยู่อาศัยส่งผลให้เกิดความเครียด ความกังวล ความสิ้นหวัง เศร้าโศก ซึมเศร้า เสียใจ เพิ่มสะสมมากขึ้น และส่งผลให้สุขภาพจิตเปลี่ยนแปลงไปและอาจพัฒนาไปสู่อาการป่วยทางจิตขั้นรุนแรงที่ส่งผลให้ผู้ป่วยมีพฤติกรรมเก็บตัวเงียบอยู่ตามลำพัง รู้สึกสิ้นหวังในชีวิต เบื่ออาหาร (Less Appetizers) มีระดับเสียงพูดที่ราบเรียบ (Monotonous Melody) ไม่เน้นเสียงคำขณะพูด (Lack of Emphatic Accent) อาการป่วยที่กล่าวมาเป็นอาการป่วยของโรคที่รู้จักกันอย่างกว้างขวางที่เรียกว่า "โรคซึมเศร้า" (Clinical Depression) ซึ่งเป็นภัยเงียบที่สามารถคุกคามสุขภาพจิตและความนึกคิดของผู้ป่วย จากผลงานวิจัยที่ผ่านมาได้มีการเสนอแนะไว้ว่าผู้ป่วยโรคซึมเศร้าส่วนใหญ่หากไม่ได้รับการรักษาจากจิตแพทย์ในช่วงต้นของการป่วย อาการป่วยโรคซึมเศร้าจะมีแนวโน้มที่จะพัฒนาไปสู่อาการป่วยระยะขั้นสุดท้าย คือการตัดสินใจฆ่าตัวตาย (Committing Suicide) ดังนั้น การวินิฉัยโรคที่มีความถูกต้องจะทำให้เราสามารถทราบและการป้องกันการฆ่าตัวตาย จึงเป็นเรื่องที่มีความสำคัญต่อชีวิต ผลที่คาดว่าจะได้รับจากการทำโครงการวิจัยนี้จะก่อให้เกิดประโยชน์อย่างสูงต่อการวินิฉัยโรคผู้ป่วยโรคซึมเศร้าโดยนักจิตเวช โครงการวิจัยนี้มุ่งนำเสนอการออกแบบและพัฒนาชุดโปรแกรมคอมพิวเตอร์ที่สามารถวิเคราะห์ และประมวลผลข้อมูลเพื่อช่วยการวินิฉัยโรคให้สามารถหาผลสรุปจากการวินิฉัยได้รวดเร็วและมีประสิทธิภาพสูงตามหลักการทางวิศวกรรมคอมพิวเตอร์และวิทยาศาสตร์ผนวกกับการวิเคราะห์คำนวณที่มีความแม่นยำสูงทั้งยังสามารถช่วยลดความยุ่ง

ยากในขั้นตอนการรวบรวมข้อมูลและประมวลผลข้อมูลที่เกี่ยวข้องกับคนป่วยเช่นประวัติการเข้ารับการบำบัดทางจิตข้อมูลบันทึกการสัมภาษณ์ในระหว่างการพบจิตแพทย์ข้อมูลเบื้องต้นของผู้ป่วยในด้านครอบครัว หน้าที่การงาน การเงิน ประสบการณ์เลวร้ายที่พบเจอในอดีต บันทึกรายงานจากทางเจ้าหน้าตำรวจที่ผู้ป่วยอาจมีส่วนร่วม หรือเกี่ยวข้องในเหตุการณ์ต่างๆ เช่น อุบัติเหตุ อาชญากรรม ฆาตรกรรม ในฐานะผู้ประสบเหตุ หรือพยานเห็นเหตุการณ์จะเห็นได้ว่าการวินิฉัยโรคนั้นมีองค์ประกอบปัจจัยจากหลายๆด้านซึ่งทำให้ต้องอาศัยเวลาอย่างมากในการรวบรวมข้อมูลประกอบการสรุปผลการวินิฉัย โดยนักจิตเวช ทำให้ไม่สามารถได้ข้อสรุปผลการวินิฉัยอย่างทันท่วงทีซึ่งอาจไม่ทันการในกรณีผู้ป่วยตัดสินใจฆ่าตัวตายสำหรับผู้ป่วยที่มีอาการป่วยอยู่ในสภาวะขั้นสุดท้ายของโรคการใช้งานคอมพิวเตอร์ในการวิเคราะห์เบื้องต้นและแปลความหมายข้อมูลที่ได้จากแหล่งต่างๆที่เกี่ยวข้องกับผู้ป่วยที่เข้ารับการบำบัดทางจิต รวมถึงการดำเนินการวินิฉัยโรคโดยจิตแพทย์ที่มีประสบการเป็นวิธีการเดียวที่ใช้ในการวินิฉัยโรค ณ ปัจจุบันการวิเคราะห์ผลการวินิฉัยโรคโดยใช้โปรแกรมคอมพิวเตอร์ที่มีระดับความสามารถในการระบุสภาวะทางจิตของผู้ป่วยผ่านเสียงพูด

ยังไม่ปรากฏชัดเจน และใช้แพร่หลายตามสถานที่โรงพยาบาลศูนย์บำบัดจิตเวชศูนย์บริการให้คำปรึกษาปัญหาด้านสุขภาพจิต ศูนย์สาธารณสุขการเฝ้าระวังและป้องกันการฆ่าตัวตายเป็นงานที่มีความสำคัญต่อสังคมส่วนรวมหากสามารถดำเนินการได้ทันท่วงทีก็จะช่วยลดการสูญเสียที่เกิดขึ้นต่อทรัพยากรมนุษย์ซึ่งเป็นทรัพยากรที่มีค่าและสำคัญยิ่ง เมื่อเปรียบเทียบกับทรัพยากรในด้านอื่นๆ

โครงการวิจัยที่จัดเสนอนี้จะเป็นการวิจัยสืบค้นและทดลองเพื่อยืนยันความเชื่อมโยงที่มีอยู่ระหว่างอคูสติกส์ของเสียงพูดกับอาการป่วยทางจิตเวช โรคซึมเศร้า ซึ่งเป็นสาเหตุต้นๆ ของการตัดสินใจฆ่าตัวตายในบุคคลซึมเศร้า คุณลักษณะเฉพาะทางอื่นๆของอคูสติกส์ของเสียงพูดที่ได้รับการทดลอง และถูกพิสูจน์แล้วว่ามีคุณสมบัติที่สามารถใช้เป็นตัวบ่งชี้สภาวะทางจิตของผู้พูด เช่น โครงสร้างองค์ประกอบฟอร์แมนท์ (Formant Structure) ความถี่หลัก (Fundamental Frequency) พิทช์คอนทัวร์ (Pitch Contour) สัมประสิทธิ์เมลเคปสตรัล (Mel–Cepstral Coefficients) อัตราพลังงานต่อย่านความถี่ (Frequency-Band Energy) กอลตัลไซเคิล (Glottal Cycle) ได้รับความนิยมในระดับหนึ่งแต่ยังขาดความเชื่อมั่นในเรื่องความถูกต้องแม่นยำในการแบ่งแยกประเภทระดับอาการป่วย และยังไม่เป็นที่ยอมรับอย่างกว้างขวางเนื่องจากความแตกต่างในเรื่องวิธีการวิเคราะห์และวิธีการถอดรหัสคุณลักษณะเฉพาะของอคูสติกส์ต่างๆที่ได้กล่าวแล้ว ดังนั้นคุณสมบัติบางประการของตัวแปรอคูสติกส์เองอาจไม่ถูกต้องเที่ยงตรงตามคุณสมบัติตามธรรมชาติของตัวมัน ซึ่งสามารถก่อให้เกิดความคลาดเคลื่อนในการตีความหมายของผลลัพธ์ที่ได้ และอาจส่งผลกระทบต่อการแยกประเภทกลุ่มระดับอาการป่วยให้มีระดับความเชื่อถือทางสถิติต่ำลงได้ สำหรับโครงการวิจัยนี้จะเป็นการวิจัยศึกษาคุณลักษณะเชิงอคูสติกส์ที่มีการตอบสนอง หรือมีคุณสมบัติแบบไดนามิกส์ (Dynamic Acoustic Response) ด้วยเหตุผลที่ว่าขั้นตอนกรรมวิธีของการถอดเอาตัวแปรอคูสติกส์ประเภทนี้ จะเป็นการวิเคราะห์และประมาณโดยตรงจากส่วนภายในสุดของระบบการกำเนิดสัญญาณเสียงพูด หรือที่รู้จักกันในสาขาวิชาการประมวลผลเสียงพูดในฐานะ "แหล่งจ่าย" (Source) แหล่งจ่ายนี้จะถูกแปรเปลี่ยนตามการกระตุ้น(Excitation) ที่เกิดจากการสั่งการจากระบบสมอง แต่ในการสั่งนี้จะถูกควบคุมจากสภาวะจิตในตัวผู้พูด ดังนั้นเสียงพูดที่เปล่งออกมาจะถูกแฝงด้วยคุณลักษณะทางอคูสติกส์ที่มีการเปลี่ยนแปลงคุณสมบัติสัมพันธ์โดยตรงกับแหล่งจ่ายของระบบกำเนิดเสียงพูด ดังนั้นตัวแปรอคูสติกส์ที่ได้มาย่อมให้ความเชื่อมั่นเชิงสถิติสูงในการแปรผันที่สัมพัธ์โดยตรงกับสภาวะจิตขณะพูด จากเหตุผลที่กล่าวมาแล้วนั้นผู้เสนอโครงการวิจัยคาดหวังอย่างมากถึงสมมุติฐานที่สามารถเป็นจริงได้ตามหลักการทางวิทยาศาสตร์และหลักวิชาการที่เกี่ยวข้องต่างๆผลลัพธ์ที่คาดว่าจะได้จากงานวิจัยนี้จะเป็นแนวทางให้ กับวิทยาศาสตร์การแพทย์ของไทยได้ริเริ่มวางแผน ออกแบบ ดำเนินการสร้าง และพัฒนาต่อยอดเทคโนโลยีทางการแพทย์ที่สามารถช่วยนักจิตเวชในการวินิจฉัยอาการโรคของผู้ป่วยโรคซึมเศร้าโดยเฉพาะกรณีอาการป่วยจากโรคซึมเศร้าขั้นรุนแรงที่มีอัตราการเสี่ยงต่อการฆ่าตัวตายที่

สูง การป้องกันการสูญเสียจากการฆ่าตัวตายในรายผู้ป่วยซึมเศร้าที่ดีที่สุดก็คือ การตรวจสอบและระบุระดับอาการโรคในผู้ป่วย ดังนั้นการวินิจฉัยโรคที่มีความรวดเร็ว ถูกต้องแม่นยำ และสามารถสรุปผลได้ทันทีหลังผู้ป่วยเข้ารับการตรวจสอบและผ่านขั้นตอนการตรวจวินิจฉัยอาการโรคโดยนักจิตเวช ย่อมเป็นสิ่งที่ต้องการสำหรับการตรวจวินิจฉัยโรคดังนั้นชุดโปรแกรมคอมพิวเตอร์ประมวลผลที่สามารถจำแนกระดับอาการโรค ซึ่งนำเสนอสำหรับโครงการวิจัยนี้ ประกอบกับประสบการณ์ความเชี่ยวชาญของนักจิตเวช จะก่อให้เกิดความร่วมมือที่มีประสิทธิภาพเมื่อเปรียบเทียบกับการวินิจฉัยโรคในวิธีดำเนินการโดยนักจิตเวชเพียงลำพัง ผลลัพธ์ที่ได้จากโครงการวิจัยที่นำเสนอนี้ย่อมก่อให้เกิดประโยชน์อย่างยิ่งต่อการช่วยเหลือ และการป้องกันการสูญเสียชีวิตของบุคคลที่ป่วยโรคซึมเศร้า

บทที่ 2 วัตถุประสงค์ของโครงการวิจัย

2.1 เพื่อรวบรวมข้อมูลเสียงพูดแบบอิสระ (Free Speech) ที่ถูกบันทึกระหว่างการสัมภาษณ์ของจิตแพทย์กับผู้ป่วยและใช้เป็นกลุ่มตัวอย่างที่แทนจำนวนประชากรของกลุ่มผู้ป่วยซึมเศร้าและกลุ่มผู้ป่วยที่มีแนวโน้ม จะฆ่าตัวตาย

2.2 เพื่อศึกษาหาคุณลักษณะทางอคูสติกส์ของเสียงที่สามารถบ่งบอกความแตกต่างระหว่างกลุ่มผู้ป่วยซึมเศร้ากับกลุ่มผู้ป่วยที่มีแนวโน้มจะฆ่าตัวตาย

2.3 เพื่อพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้สำหรับการประมาณและวิเคราะห์หาคุณลักษณะทางอคูสติกส์ที่เจาะจง

2.3.1 สัดส่วนพลังงานของสเป็คตรัมในช่วงความถี่ 0-2,000 Hz

2.3.2 โครงสร้างสเป็คตรัมของเสียงที่กำเนิดในหลอดเสียง

2.3.3 ปริมาณข้อมูลข่าวสารเฉลี่ยต่อแหล่งกำเนิดของเสียง

2.4 เพื่อพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้สำหรับวัดหาอำนาจการจำแนกของคุณลักษณะทางอคูสติกส์ที่ได้มาและสามารถแสดงผลของการแบ่งแยกกลุ่มผู้ป่วยตามสภาวะทางจิตที่แท้ได้อย่างถูกต้อง

บทที่ 3 ขอบเขตของโครงการวิจัย

3.1 ติดต่อประสานงานกับทางโรงพยาบาลเพื่อเก็บข้อมูลเสียง จำนวนผู้ป่วยที่เข้าร่วมโครงการทั้งสิ้น 20 คน แบ่งเป็นผู้ป่วยหญิงซึมเศร้า 10 คน ผู้ป่วยหญิงซึมเศร้าที่มีประสบการณ์จากการฆ่าตัวตายแต่ไม่สำเร็จ หรือผู้ป่วยที่มีแนวโน้มจะฆ่าตัวตาย 10 คน

3.2 ออกแบบและพัฒนาโปรแกรมการจัดการขั้นต้นกับข้อมูลเสียง (Preprocessing) โดยข้อมูลเสียงของผู้ป่วยแต่ละราย จะผ่านขั้นตอนกรรมวิธีต่างๆต่อไปนี้

3.2.1 การกำจัดสัญญาณต่างๆที่นอกเหนือจากเสียงพูดของผู้ป่วยและจิตแพทย์ เช่น สัญญาณรบกวน (Background Noises, Sound Artifact or Harmonics of Power Line) และ ช่วงเริ่มต้นของการบันทึกเสียง (Heading of Audio Recording)

3.2.2 คัดแยกเสียงของผู้ป่วยออกจากเสียงของผู้สัมภาษณ์ (Sound Editing)

3.2.3 ตรวจสอบหาเฉพาะส่วนของเสียงที่มีพลังงาน และส่วนที่ไม่มีพลังงาน (Voice Detection: Separating voiced and unvoiced segments of speech data)

3.2.4 จัดการแบ่งกลุ่มและรวบรวมข้อมูลเสียง ที่ผ่านขั้นตอนต่างๆมา และจัดเก็บเป็นตัวอย่าง (Samples) ที่จะใช้แทนประชากรของผู้ป่วยในกลุ่ม High-Risk Suicidal Patients และกลุ่ม Depressed Patients

3.3 ออกแบบและพัฒนาโปรแกรมการถอดคุณลักษณะทางอคูสติกส์จากเสียงพูด (Feature Extraction Algorithm) เป็นต้นว่า

3.3.1 สัดส่วนของพลังงานสเป็คตรัมในย่านความถี่ย่อยๆ 500 Hz ต่อพลังงานทั้งหมดในช่วงความถี่ 0-2,000 Hz

3.3.2 ความถี่ฟอร์แมน (Formant Frequency) แถบความถี่ฟอร์แมน (Formant Bandwidth) และขนาดฟอร์แมน (Formant Amplitude) จากโครงสร้างความถี่ต่อเวลาของเสียงพูด (Formant Pattern) โดยประมาณหาด้วยวิธีการทางสถิติ (Gaussian Mixture Model) จากสัญญาณการตอบสนองทางความถี่ของหลอดทางเดินเสียง (Vocal Tract) ที่ได้จากการประมาณเคปสตรัม (Cepstrum Estimation) ของเสียงพูด

3.4 ออกแบบและพัฒนาโปรแกรมการวิเคราะห์ทางสถิติสำหรับตรวจวัดอำนาจการแยกแยะของคุณลักษณะทางอคูสติกส์ที่ศึกษา

3.4.1 ตรวจหาคุณลักษณะทางอคูสติกส์ที่มีความโดดเด่น (Primary Feature) โดยใช้หลักการทางสถิติ

F-ratio ในการวัดหาค่าอำนาจจำแนกระหว่างกลุ่ม (Separation Power) เพื่อลดจำนวนพารามิเตอร์ที่ไม่สำคัญลงให้เหมาะสมกับจำนวนตัวอย่างที่ (Dimensionality Reduction)

3.4.2 ทดสอบการแบ่งกลุ่มผู้ป่วย ตามสภาวะทางจิตที่แท้จริง ซึ่งแฝงมาในพารามิเตอร์ทางเสียง (vocal parameters) โดยใช้แบบสองกลุ่มเปรียบเทียบ (Pairwise Classification) วิธีต่างๆที่จะนำมาใช้เช่น Quadratic Discrimination, Cross-Validation, Maximum likelihood Classification และ Gaussian Mixture Speaker Model โดยทำการฝึกหัด (Training) และทดสอบ (Testing) ตัวแบ่งกลุ่ม (Classifier) ที่มีคุณลักษณะอคูสติกส์ที่ศึกษาเป็นโมเดลสำหรับการประเมินความถูกต้องของการระบุสภาวะทางจิตของกลุ่มผู้ป่วยต่างๆ (Speaker Identification) ที่พารามิเตอร์ทางเสียงแสดงออกมา

บทที่ 4 ประโยชน์ที่คาดว่าจะได้รับ

4.1 เผยแพร่ผลงานวิจัยสู่ระดับสากล ตีพิมพ์ในวารสารระดับนานาชาติ

4.2 โปรแกรมคอมพิวเตอร์ที่มีความสามารถวิเคราะห์และแสดงผลระดับแนวโน้มการฆ่าตัวตายในกลุ่มผู้ป่วยซึมเศร้า และสามารถนำไปใช้ร่วมได้จริงกับการวินิจฉัยโรคของจิตแพทย์ตามโรงพยาบาล ศูนย์บำบัดสุขภาพจิต หรือ คลีนิกที่เกี่ยวข้องกับการบำบัดจิต

บทที่ 5 ทฤษฎีที่เกี่ยวข้อง

5.1 Model of Speech Production System

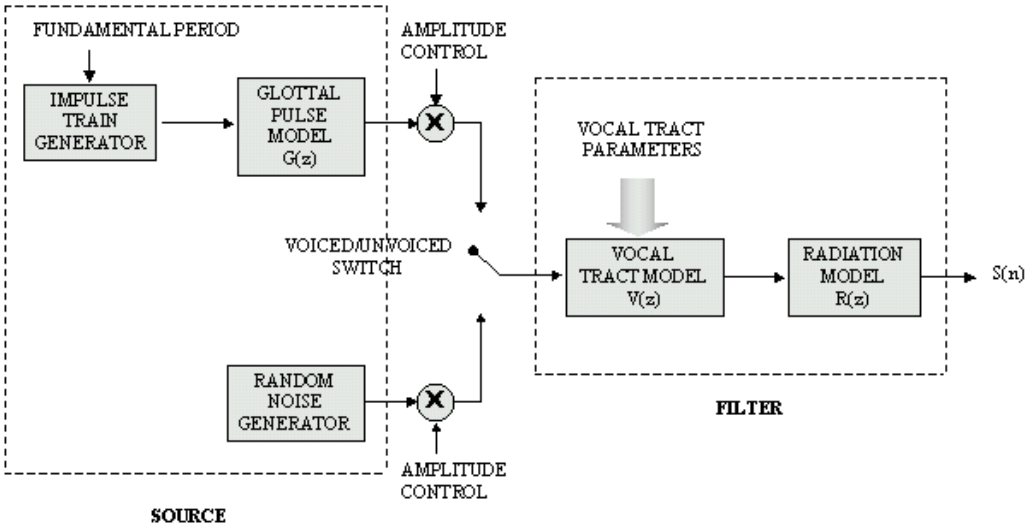


Fig. 5.1 Speech Production System

ส่วนที่จะศึกษาในโครงงานวิจัยนี้คือส่วนที่เรียกว่า ทางเดินของเสียง (Filter Model) และจะเน้นศึกษาเฉพาะสำหรับส่วนของหลอดเสียง (Vocal Tract) สำหรับโครงการวิจัยนี้ โมเดลระบบกำเนิดเสียงของมนุษย์สามารถเขียนเป็นสมการทางคณิตศาสตร์ โดยอาศัยหลักการการแปลงโดเมน (Z-Transformation) ได้ว่า $S(z) = G(z)V(z)R(z)$

โดยที่ $S(z)$ - Speech waveform

$G(z)$ - Glottal pulse train

$V(z)$ - Upper vocal tract (Formant frequencies)

$R(z)$ - Lip radiation

และส่วนที่เน้นศึกษา $V(z)$ มีสมการเป็น

$$V(z) = \frac{G}{1 - \sum_{k=1}^{N} \alpha_k z^{-k}}$$

โดยที่ G = Gain

$\{\alpha_k\}$ = Filter coefficients

N = Number of poles in the unit circle

5.2 Power Spectral Density

พลังงานในสเปคตรัมทางความถี่ของสัญญาณเสียงพูด สามารถแสดงในรูปสมการได้เป็น

$$\hat{R}_x(\omega_k) = \frac{1}{M}\sum_{m=0}^{M-1} |DFT_k(x_m)|^2 \triangleq \{|X_m(\omega_k)^2|\}_m$$

โดยที่

$\{\cdot\}_m$ = Time averaging across frames of speech data (blocks) indexed by *m*

M = Number of frames of speech data

x_m = Frame of speech data,

$X_m(\omega_k)$ = Discrete Fourier Transform (DFT) of each frame of speech data,

5.3 Gaussian Mixture Model (GMM)

การประมาณฟอร์แมนท์พารามิเตอร์ของเสียงพูด โดยอาศัยหลักการที่เกี่ยวข้องกับพารามิเตอร์ทางสถิติ (Mean, Variance, Prior and Posterior Probability) สมการทางคณิตศาสตร์ของ Gaussian mixture distribution สามารถเขียนได้เป็น

$$f_{\mathbf{X}}(x;\theta) = \sum_{m=1}^{M} c_m \cdot N(x;\theta_m)$$

โดยที่

c_m = Mixture weight of each density component of the M mixture, which must satisfy the following constraints: that

$$\sum_{m=1}^{M} c_m = 1;\ c_m \geq 0,\ \theta_m = [\sigma_m^2,\ \mu_m]$$

และสมการ Gaussian distribution สามารถเขียนได้เป็น

$$N(x;\sigma^2,\mu) = \frac{1}{\sqrt{2\pi}\sigma_m} \cdot \exp^{-\frac{1}{2}\left(\frac{x-\mu_m}{\sigma_m}\right)^2}$$

5.4 Modified EM-Algorithm

ซึ่งใช้ในการประมาณหา พารามิเตอร์ทางสถิติ โดยวิธีการ Expectation และ Maximization ค่าพารามิเตอร์ต่างๆ ที่ต้องการหาจากโครงสร้างของการตอบสนองทางความถี่ของหลอดทางเดินเสียง (Frequency Response of The Vocal Tract)

Expectation-step (E-step) :

$$\hat{p}^{(t)}(m \mid \underline{\omega}_k, \hat{\lambda}^{(t)}) = \frac{\hat{c}_m^{(t)} \cdot N(\underline{\omega}_k; \hat{\sigma}_m^{2(t)}, \hat{\mu}_m^{(t)})}{\sum_{m=1}^{M} \hat{c}_m^{(t)} \cdot N(\underline{\omega}_k; \hat{\sigma}_m^{2(t)}, \hat{\mu}_m^{(t)})}$$

Maximization-step (M-step) :

$$\hat{c}_m^{(t+1)} = \sum_{k=1}^{\Omega} \phi(\underline{\omega}_k) \cdot \hat{p}^{(t)}(m \mid \underline{\omega}_k, \hat{\lambda}^{(t)})$$

$$\hat{\mu}_m^{(t+1)} = \frac{1}{\hat{c}_m^{(t+1)}} \sum_{k=1}^{\Omega} \underline{\omega}_k \cdot \phi(\underline{\omega}_k) \cdot \hat{p}^{(t)}(m \mid \underline{\omega}_k, \hat{\lambda}^{(t)})$$

$$\hat{\sigma}_m^{2(t+1)} = \frac{1}{\hat{c}_m^{(t+1)}} \sum_{k=1}^{\Omega} (\underline{\omega}_k - \hat{\mu}_m^{(t)})^2 \cdot \phi(\underline{\omega}_k) \cdot \hat{p}^{(t)}(m \mid \underline{\omega}_k, \hat{\lambda}^{(t)})$$

5.5 F-Ratio

หลักการวัดหาอำนาจจำแนก (F-ratio) F-ratio of the ith feature: $F_i = \frac{B_i}{W_i}$

โดยที่

B_i = Between–class variance

W_i = Pooled within–class variance of the ith feature

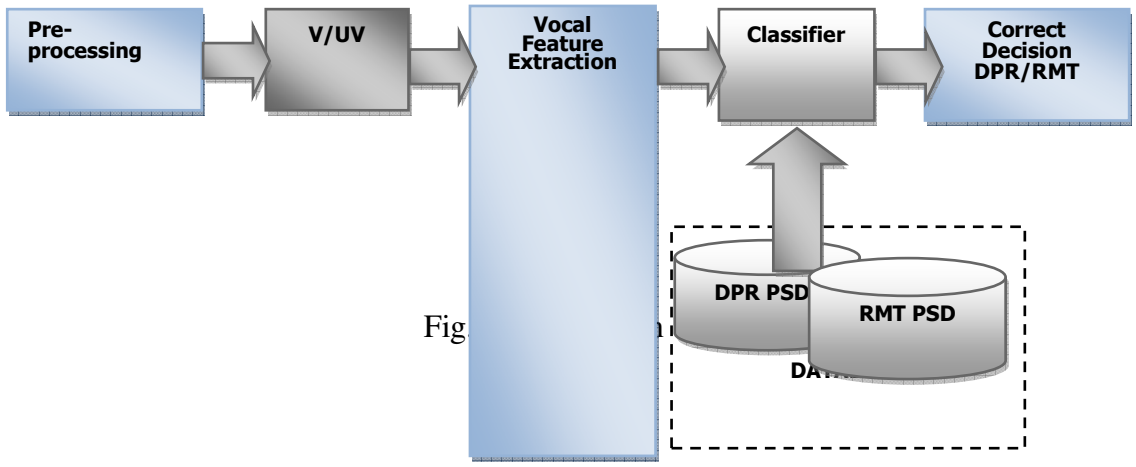
บทที่ 6 วิธีการดำเนินการวิเคราะห์สัญญาณ (Methodology)

ผู้ป่วยแบ่งเป็น 3 กลุ่ม 1.ผู้ป่วยซึมเศร้า 2.ผู้ป่วยเสี่ยงต่อการฆ่าตัวตาย 3.กลุ่มผู้ป่วย ผู้ป่วยที่ฟื้นฟูจากการเป็นโรคซึมเศร้า โดยจะเป็นเพศชายทั้งหมด แต่ละคนจะถูกสัมภาษณ์และมีการบันทึกเทปโดยแบ่งเป็น 2 ส่วน คือส่วนแรกจะเรียกว่า Interview Session โดยจิตแพทย์จะเป็นผู้สอบถามและทำการบันทึกการสอบถามและคำตอบจากผู้ป่วยตามข้อคำถามที่กำหนดไว้ ส่วนที่ 2 เป็นการบันทึก Post Session ในตอนท้ายซึ่งผู้ป่วยจะต้องอ่านข้อความจากหนังสือที่กำหนดซึ่งในที่นี้เรียกว่า Rainbow Passage โดยหนังสือตัวนี้จะมีหน้าที่ในการให้ผู้ป่วยเปล่งเสียง แสดงถึงข้อคำพูดที่ใช้วัดในส่วนของการออกเสียงแบบ phonetic ในเสียงสำหรับวัดความสมดุลมากน้อยที่เกิดขึ้น ข้อมูลที่บันทึกทั้งหมดจะนำไปทำการประมวลผลโดยมีขั้นตอนตามรูปที่แสดงด้านล่าง โดยข้อมูลเสียงที่บันทึกนี้จะเป็นลักษณะการวิเคราะห์แบบ Offline คือเป็นไฟล์ที่บันทึกแล้วนำมาวิเคราะห์แต่ละไฟล์โดยผ่านโปรแกรม Matlab ช่วย ขั้นตอนเบื้องต้นการวิเคราะห์คือ สัญญาณเสียงแต่ละเสียงจะถูกทำการแปลงเป็นดิจิตอล โดยใช้ A/D Converter ที่ความถี่อัตราการสุ่มเท่ากับ 10 KHz และใช้ฟิลเตอร์ชนิด Low-Pass Filter ที่ Cutoff Frequency 5 KHz สัญญาณรบกวนและอื่นๆที่นอกเหนือจากเสียงผู้ป่วยถูกขจัดทิ้งหรือเอาออกจากไฟล์ข้อมูล โดยผ่านโปรแกรม GOLDWAVE Version 5.10 Audio Editor สัญญาณที่ได้หลังจากผ่านขั้นตอนดังกล่าวจะนำไปแบ่งเป็น Segments ย่อยๆละ 20 วินาที แต่ละSegment จะนำไปทดสอบโดยผู้ใช้งานพลังงานของแต่ละ Segment ว่าจะจัดอยู่ในประเภทของเสียงพูดที่เรียกว่า Voiced Speech, Unvoiced Speech หรือ Silence

6.1 การถอดรหัสลักษณะ (Feature Extraction)

Power Spectral Densities (PSD) จะถูกประมาณจากสัญญาณเสียงที่ถูกตรวจจับเป็น Voiced Segment ด้วยวิธี Welch Method และพิจารณาที่ทุกๆ 100-Point โดยใช้ Hamming Window เป็นตัวกำหนดไม่มีการพิจารณา Overlapping การประมาณดำเนินการโดยคอมพิวเตอร์โปรแกรม Matlab โดยใช้ฟังก์ชั่น FFT ที่ 1024-point และ 40ms Window โดยประมาณบน Segment ต่างๆที่เราบันทึกได้ คุณลักษณะ 6 ตัวพารามิเตอร์ ที่จะพิจารณาจะประกอบไปด้วย Peak Power ขนาดเพาเวอร์สูงสุดที่เกิดขึ้นในสเป็คตรัม Frequency location ตำแหน่งความถี่ที่Maximum Peak เกิด Sub-band PSD ของ Sub-band ทุกๆ 500Hz ในช่วง 0-2,000Hz โดยจะกำหนดเป็น PSD1, PSD2, PSD3 และ PSD4 ซึ่งทั้งหมดนี้จะนำไปใช้เป็น input feature สำหรับการคัดกรองระหว่างกลุ่มผู้ป่วยต่างๆ โดยตัวคัดแยก

ประกอบไปด้วย Linear Quadratics Classifier และรวมถึง F-Ratio การคัดกรองพิจารณา Pair-wise โดยทุกๆ 20 วินาทีของสัญญาณเสียงพูดจะนำมาพิจารณาในการคัดกรอง ในส่วนของสถิติจะใช้โปรแกรมที่เรียกว่า Pattern Analysis System ในการวัดคุณสมบัติทางสถิติต่างๆ โดยจะแบ่งวัดเป็น Suicidal กับ Depressed, Depressed กับ Remitted) และ Remitted กับ Suicidal ทั้งหมด 3 คู่ ค่าสถิติที่ตรวจวัดจากพารามิเตอร์ที่ศึกษา จะพิจารณาประกอบไปด้วย Sensitivity, Specificity, Positive, Predictive



Fig

บทที่ 7 ผลการทดลอง (Experimental Results)

ค่าเฉลี่ย Means และ ค่าเบี่ยงเบน Standard Deviation S.D. แต่ละ คุณลักษณะ Acoustical Feature เช่น Peak Power, Peak Location (Hz) แสดงดังตารางที่ 1 เสียงผู้ป่วยเสี่ยงต่อการฆ่าตัวตายสามารถระบุ หรือสังเกตได้เด่นชัดโดย Peak Location เกิดขึ้นที่ความถี่ต่ำกว่าเทียบกับ Peak Location ของ กรณี Remitted Case เห็นได้ชัดเจนว่า ในกรณีศึกษาเสียงจาก Reading Case จะปรากฏในทิศทางเดียวกับกรณี Interview เสียงผู้ป่วยซึมเศร้าจะแสดงให้เห็นถึงการเพิ่มขึ้นของ PSD1, PSD3 และการลดลงของ PSD2 ในกรณี Interviewing Session เมื่อเปรียบเทียบกับกรณีเสียงของผู้ป่วยที่ฟื้นฟูกลับมาจากกรณีโรคซึมเศร้า ทิศทางนี้ยังคงสามารถเห็นได้ในกรณี Reading Case สำหรับผลของการศึกษาคัดกรองระหว่าง Suicidal และ Remitted ค่า PSD1 ในช่วง Sub-band แรก ย่านความถี่ 0-500Hz มีค่าลดลงสำหรับกรณี Remitted ในขณะที่ PSD2 และ PSD3 ใน Sub-band ที่สูงขึ้น มีแนวโน้มเพิ่มขึ้น ผลวิเคราะห์เหล่านี้สามารถสังเกตได้เด่นชัดในกรณี Reading Session ยกเว้น PSD3 ซึ่งไม่ปรากฏความแตกต่างใน Energy Percentage ผลของการวิเคราะห์คัดกรองสรุปแสดงในตารางที่ 3 และ 4 สำหรับกรณี Interview และ Reading Cases ตามลำดับ ผู้ป่วยซึมเศร้าจะสังเกตได้ว่ามีความแตกต่างจากกรณีผู้ป่วยที่ถูกวินิจฉัยอยู่ในกลุ่ม Suicidal โดยพิจารณาได้จากความถูกต้องการคัดกรองที่ระดับ 77% สำหรับกรณี Interview Case อย่างไรก็ตาม ผู้ป่วยซึมเศร้า และผู้ป่วย ที่เสี่ยงต่อการฆ่าตัวตายสามารถคัดกรองแยกกลุ่ม ได้อย่างมีประสิทธิภาพที่ 94% และ 85% เมื่อเปรียบเทียบกับกรณี Remitted Case สมรรถนะเชิงสถิติต่างๆที่ใช้วัดการคัดกรองเช่น Sensitivity, Specificity, PPV และ NPV ได้ถูกคำนวณและพิจารณาเช่นกัน ผลของการคัดกรองจากกรณี Reading Case เปิดเผยให้เห็นว่ามีแนวโน้มตรงข้ามกันจากตารางที่ 4 ความแตกต่างกรณีเปรียบเทียบระหว่าง Suicidal และ Depressed สามารถที่คัดแยกกลุ่มได้ที่ระดับความถูกต้อง 82% ส่วนกรณีเปรียบเทียบอื่นๆมีค่าระดับความถูกต้องการคัดแยกกลุ่มที่น้อยลงโดยรวม

Tab.7.1 PSD Distribution for All Subjects – Interview session

	Suicidal	Depressed	Remitted
Peak Power	**(20.88, 2.70)**	**(20.63, 3.34)**	**(20.92, 1.70)**
Peak Location (Hz)	**(284.47, 84.89)**	**(292.02, 58.55)**	**(331.17, 67.98)**
% Power in PSD_1	**(0.79, 0.08)**	**(0.79, 0.08)**	**(0.74, 0.05)**
% Power in PSD_2	**(0.19, 0.07)**	**(0.18, 0.08)**	**(0.23, 0.04)**
% Power in PSD_3	**(0.01, 0.01)**	**(0.03, 0.02)**	**(0.02, 0.01)**

Tab.7.2 PSD Distribution for All Subjects – Reading session

	Suicidal	Depressed	Remitted
Peak Power	**(21.52, 2.22)**	**(20.80, 1.59)**	**(21.53, 2.08)**
Peak Location (Hz)	**(298.83, 112.84)**	**(296.10, 66.11)**	**(351.65, 74.24)**
% Power in PSD_1	**(0.78, 0.08)**	**(0.82, 0.06)**	**(0.75, 0.09)**
% Power in PSD_2	**(0.19, 0.08)**	**(0.16, 0.05)**	**(0.23, 0.09)**
% Power in PSD_3	**(0.01, 0.01)**	**(0.02, 0.01)**	**(0.01, 0.01)**

Tab.7.3 Summarized Statistics – Interview session

Pairwise	% Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	**77**	**0.89**	**0.63**	**0.76**	**0.8**
Dep/Rem	**94**	**0.94**	**0.94**	**0.94**	**0.94**
Sui/Rem	**85**	**0.91**	**0.76**	**0.84**	**0.86**

Tab.7.4 Summarized Statistics – Reading session

Pairwise	% Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	**82**	**0.73**	**0.88**	**0.82**	**0.82**
Dep/Rem	**73**	**0.85**	**0.58**	**0.71**	**0.76**
Sui/Rem	**75**	**0.73**	**0.76**	**0.73**	**0.76**

TAB.7.5 SUMMARIZED COMPARATIVE RECOGNITION PERFORMANCES FROM ANALYSES OF SPONTANEOUS SPEECH

Pairwise	% Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	**85.58**	**0.81**	**0.88**	**0.89**	**0.85**
Dep/Rem	**72.25**	**0.70**	**0.75**	**0.63**	**0.85**
Sui/Rem	**81.08**	**0.90**	**0.60**	**0.86**	**0.83**

TAB.7.6 SUMMARIZED COMPARATIVE RECOGNITION PERFORMANCES FROM ANALYSES OF AUTOMATIC SPEECH

Pairwise	% Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	**85.50**	**0.91**	**0.88**	**0.99**	**0.93**
Dep/Rem	**92.00**	**0.86**	**0.97**	**0.98**	**0.89**
Sui/Rem	**90.25**	**0.89**	**0.93**	**0.93**	**0.92**

บทที่ 8 ข้อวิจารณ์ (Discussion)

โดยทั่วไปผลทางด้านการวัดสมรรถนะ (Performance) ของขบวนการคัดกรองระบุกลุ่มผู้ป่วย สรุปผลในตารางที่3 แสดงให้เห็นว่าการใช้คุณลักษณะ PSD หลากหลายพารามิเตอร์กรณี Interview Case จะให้การคัดกรองที่มีประสิทธิภาพดีระหว่างผู้ป่วยซึมเศร้า กับผู้ป่วยฟื้นฟูปกติจากอาการป่วยโรคซึมเศร้า และระหว่างผู้ป่วยซึมเศร้าเสี่ยงต่อการฆ่าตัวตาย High-Risk Suicidal และผู้ป่วยซึมเศร้า Depressed เสียงที่ถูกบันทึกในกรณี Reading Case สามารถแสดงให้เห็นถึงประสิทธิภาพในการคัดกรอง พิจารณาได้จากตารางที่ 4 ผลการคัดกรองที่ดีที่สุดด้วยความถูกต้องสูงสุด สามารถพิจารณาได้จากกรณีคัดกรองระหว่างผู้ป่วยเสี่ยงต่อการฆ่าตัวตายกับผู้ป่วยซึมเศร้า ในกรณี Reading Case มีค่าความถูกต้องของการคัดกรองน้อยลงในกรณีผู้ป่วยซึมเศร้า Depressed กับผู้ป่วยฟื้นฟู Remitted ในกรณี Interviewing Case และภาพรวมโดยเฉลี่ย Overall จะมีค่าความถูกต้องการคัดกรองลดลงตามลำดับ อาจจะพิจารณาได้จากสาเหตุเนื่องจากผู้ป่วยอาจจะมีการวางท่าทางในการอ่านข้อความที่แตกต่างกัน กรณีการนั่งหรือการวางท่าทางของการบันทึกเสียงที่ผู้ป่วยแต่ละคนจะมีลักษณะแตกต่างกันตามความเคยชินของตัวบุคคล เช่น ระยะการอ่าน ระยะห่างจากไมโครโฟน และการเอนเอียงของลำตัวที่อาจจะมีทิศทาง ที่ไม่ทำให้การแปล่งเสียงเกิดขึ้นอย่างสมบูรณ์ อาจเกิดจากการไปบล็อกหรือกั้นทิศทางเดินของเสียง มีส่วนทำให้เกิดความแตกต่างในการคัดแยกกรองกรณี Interview Case ตารางที่ 3 แสดงค่าเฉลี่ยของการคัดแยกกรองที่ระดับความถูกต้อง 94% ได้จากกรณีศึกษาผู้ป่วยซึมเศร้า Depressed และผู้ป่วยฟื้นฟู Remitted เป็น Highest Correct Classification Percentage ระดับความถูกต้องสูงสุดที่พบจากกรณีศึกษา Interview Case สำหรับกรณีการอ่าน Reading Case ค่า Highest Correct Classification Percentage ระดับความถูกต้องสูงสุดที่ได้คือ 82% จากกรณีศึกษาการคัดแยกกรองระหว่างผู้ป่วยซึมเศร้าเสี่ยงต่อการฆ่าตัวตาย Suicidal และผู้ป่วยซึมเศร้า Depressed ที่ได้ทางสถิติ โดยดูจากค่าคำนวณ Specificity อยู่ที่ 0.63 สำหรับกรณี Interview Case มีค่าน้อยกว่ากรณี Reading Case ผลลัพธ์ที่ได้นี้สามารถตีความหมายได้ว่า ผลจากการวัดผู้ป่วยซึมเศร้านั้นมีค่าคัดกรองผิดพลาด Misclassified Accuracy เป็นเสียงผู้ป่วยเสี่ยงต่อการฆ่าตัวตายมีมากกว่ากรณีของ Reading Case ผลทดลองแสดงให้เห็นว่า Sub-band PSD based Acoustical Features หลายๆพารามิเตอร์สามารถช่วยให้การคัดกรองมีประสิทธิภาพมากขึ้น ซึ่งสังเกตได้จากกรณีของผู้ป่วยเสี่ยงฆ่าตัวตาย Suicidal กับผู้ป่วยซึมเศร้า Depressed การศึกษาต่อยอดในอนาคตอันใกล้

โดยวิธีการจัดลำดับคุณสมบัติการจำแนกและวิธีการ Predictive Validity การประเมินความถูกต้อง สามารถทดสอบโดยวิธี Ranking Method ผ่านโปรแกรม PcOLPARS พบว่า Sub-bands ทั้งสาม PSD1, PSD2 และ PSD3 เป็น Spectral Energy Feature ที่มีคุณลักษณะเด่นในเรื่องอัตราการ จำแนก สำหรับการศึกษาต่อเนื่องในอนาคตขนาดของข้อมูล Speech Databases ที่ใช้ ควรมีจำนวนผู้ป่วยมากขึ้นอันจะมีส่งให้การตีความเชิงสถิติของผลวิเคราะห์มีความแม่นยำมากขึ้น ในอนาคตงานวิจัยต่อเนื่องจะต้องทำการเก็บข้อมูลมากขึ้นมากกว่าเดิม เพื่อช่วยส่งเสริมงานวิจัยในด้านแม่นยำ และเชื่อถือได้ในด้านการวิเคราะห์ผลนอกเหนือจากนั้นแล้วการทำให้ผลการคัดแยกกรองมีความถูกต้องและดีขึ้นสามารถที่จะทำได้โดยใช้ตัวคัดแยกกรองที่มีลักษณะแบบ Multi-parameter นั้นคือจะสามารถที่รวบรวมคุณลักษณะต่างๆที่ประมาณได้จากหลายๆ มิติ วิธีการต่างๆ เข้าด้วยกัน เช่น MFCC, Glottal Slope Tilt หรือ F_0 Contour ร่วมเป็นคอมมิเนชั่นเข้าด้วยกันเป็นเซ็ต สามรถเพิ่มประสิทธิภาพการคัดแยกกรองมากขึ้น รวมถึงการใช้เทคนิคใหม่ๆหรือที่ดีจากวิธีการถอดรหัสคุณลักษณะต่างๆจากวิทยาการสาขา Speech processing

นอกจากการศึกษาการเปลี่ยนแปลงของพลังงานในแถบความถี่เสียงในช่วง 0 – 2000Hz ผลของการวิเคราะห์คุณลักษณะอคูสติกส์ของเสียงพูดในกรณีที่เรียกว่าตัวประมาณ Performance อาศัยหลักการ Feature Model ที่กำหนดจาก Mean, Variance, Prior and Posterior Probability ความน่าจะเป็นก่อนและหลังจากการประมาณแบบจำลอง Model เชิงคณิตศาสตร์ ผลการทดลองโดยรวมจากการ Classification คัดแยกกรองมีความแตกต่างอย่างเห็นได้ชัดเจนระหว่างคุณลักษณะเชิงสถิติกรณี Spontaneous หรือ Interviewing Case กับ กรณีศึกษา Reading Case กรณี Interviewing Case เสียงสัมภาษณ์สามารถระบุได้โดยจากค่าคัดแยกกรองสูงสุดในแต่กลุ่มพารามิเตอร์ที่ใช้ในการวัดหา โดยพิจารณาจากค่า F-ratio ของมัน เช่น CF4, CF3 และ PSD2 โดยพบได้ในการคัดกรองระหว่างกลุ่มผู้ป่วยซึมเศร้า และผู้ป่วยซึมเศร้าเสี่ยงต่อการฆ่าตัวตาย WC4, BW4 และ PSD3 พบในกรณีการคัดกรองระหว่างผู้ป่วยฟื้นฟู กับผู้ป่วยซึมเศร้า และ CF4, CF2 และ CF3 พบในกรณีผู้ป่วยเสี่ยงต่อการฆ่าตัวตาย ผลจากกรณีศึกษา Reading Case หรือกรณีการอ่านข้อความเปิดเผยให้เห็นว่า CF4, PSD2 และ CF3 เป็นคุณลักษณะที่เด่นชัดในการคัดกรองระหว่างผู้ป่วยซึมเศร้า PSD3, PSD2 และ BW3 พบได้ในกรณีผู้ป่วยฟื้นฟู กับผู้ป่วยซึมเศร้า CF4, PSD3 และ CF3 พบได้ในกรณีผู้ป่วยฟื้นฟูกับผู้ป่วยซึมเศร้าเสี่ยงฆ่าตัวตาย จากการวิเคราะห์กรณีคัดกรองแบบคู่ กรณีเสียงบันทึกได้ผลคัดแยกกรองสูงสุงที่ 85.58% เมื่อศึกษากรณีเปรียบเทียบระหว่าง Depressed กับ Suicidal ซึ่งเป็นค่าความถูกต้องสูงสุดที่ได้จากกรณีต่างๆ โดยมีค่าSpecificity (SP) อยู่ที่ 0.88 PPV ค่าเท่ากับ 0.89 และ NPV ค่าเท่ากับ 0.85 และค่าความถูกต้องคัดแยกกรองสูงสุด 92% พบได้จากกรณีคัดกรองผู้ป่วยฟื้นฟูจากซึมเศร้าและ

ผู้ป่วยซึมเศร้า กรณี Reading Case โดยมีค่า SP สูงสุดที่ 0.97 PPV ค่าเท่ากับ 0.98 และ Sensitivity (SE) ค่าสูงสุดเท่ากับ 0.86 และ NPV ค่าเท่ากับ 0.89 ตามลำดับ ผลการศึกษาเปรียบเทียบระหว่างกรณีการบันทึกเสียงทั้งสองแบบ Interviewing and Reading Cases พบว่ากรณีการคัดแยกกรองระหว่างผู้ป่วยซึมเศร้าและผู้ป่วยเสี่ยงฆ่าตัวตาย ผลการคัดกรองให้ประสิทธิภาพด้อยกว่า 3% จากความถูกต้องที่ได้ในกรณี Interviewing Case สิ่งที่น่าสนใจคือค่า 20% ของความถูกต้องในการคัดแยกกรองโดยประมาณสามารถเพิ่มขึ้นในกรณีศึกษาระหว่างผู้ป่วยฟื้นฟูและผู้ป่วยซึมเศร้าอยู่ที่ 92% สำหรับกรณี Reading Case เมื่อเปรียบเทียบกับกรณี Interviewing Caseที่ระดับความถูกต้อง 72.25% ผลของการวัดที่ได้ด้านประสิทธิภาพของการคัดแยกกรอง ที่ค่าความถูกต้องสูงสุดแสดงให้เห็นว่าคุณลักษณะที่ศึกษานี้ มีประสิทธิภาพดีพอสำหรับการคัดแยกกรองเสียง ซึ่งสามารถใช้ในการคัดแยกกรองผู้ป่วยกลุ่มซึมเศร้าและกลุ่มเสี่ยงฆ่าตัวตายอย่างถูกต้อง เสียงที่ได้บันทึกจากการสัมภาษณ์โดยจิตแพทย์ที่ใช้ในการวิเคราะห์คุณลักษณะทางอคูสติกส์แสดงให้เห็นถึงความมีประสิทธิภาพอย่างมีนัยยะสำคัญในกรณีการคัดแยกกรองระหว่างกลุ่มผู้ป่วยฟื้นฟูและกลุ่มผู้ป่วยซึมเศร้ากรณี Reading Case

การวิจัยต่อเนื่องในอนาคต จะพิจารณาเรื่องการพัฒนาชุดโปรแกรมคัดแยกกรอง โดยมุ่งเน้นพิจารณาหลักการถอดรหัสคุณลักษณะเฉพาะจากเสียงพูด และการประเมินประสิทธิภาพการวิเคราะห์พารามิเตอร์ในสัญญาณเสียง และการคัดแยกกรอง การพัฒนาฐานข้อมูลเสียงบันทึก (Database) ที่มีขนาดใหญ่กว่าขนาดที่ใช้ในปัจจุบันศึกษาจะช่วยให้ส่งเสริมให้การวินิจฉัยโรคกรณีโรคที่เกี่ยวข้องกับสุขภาพจิตมีความถูกต้องแม่นยำมากขึ้น และเป็นประโยชน์ต่อเครื่องมือช่วยวินิจฉัยโรคทางจิตแก่แพทย์ การวิเคราะห์โรคในรายผู้ป่วยกรณีที่เกี่ยวข้องกับโรคทางโรคประสาทอาจมีความเกี่ยวข้องกับกรณีอาการต่อเนื่องในรายผู้ป่วยซึมเศร้าซึ่งต้องอาศัยวิธีการตรวจวัดที่มีความซับซ้อน และฐานข้อมูลที่มีขนาดใหญ่พอ ที่จะช่วยทำให้การคัดกรองแยกในรายาผู้ป่วยประเภทเสี่ยงรุนแรงและมีแนวโน้มจะฆ่าตัวตาย และมีความจำเป็นที่ต้องดำเนินการวินิจฉัยเร่งด่วนและผลลัพท์ที่แม่นยำถูกต้อง

บทที่ 9 สรุปและข้อเสนอแนะ (Conclusion and Suggestion)

การศึกษา PSD ของ Sub-bands ย่อยๆระหว่าง 0-2kHz ในกรณีเสียงพูดที่บันทึก Interviewing Case ระหว่างผู้ป่วยและนักจิตวิทยาแสดงผลจากากรศึกษามีประสิทธิภาพในการคัดแยกกรองโดยเฉพาะกรณีกลุ่มเปรียบเทียบผู้ป่วยซึมเศร้ากับกลุ่มผู้ป่วยฟื้นฟูรวมถึงกลุ่มผู้ป่วยเสี่ยงฆ่าตัวตาย จากกลุ่มผู้ป่วยซึมเศร้ากรณี Reading Case อ่านบทความ พิจารณาจากผลการศึกษาการคัดแยกกรองที่มีความถูกต้องเป็นเปอร์เซ็นสัดส่วนที่สูงชัดเจน ซึ่งบ่งบอกได้ว่า PSD ที่ศึกษาในงานวิจัยนี้สามารถใช้เป็นตัวบ่งบอก วัดระดับหรือเชื่อมโยงสัมพันธ์ต่อระดับอาการป่วยโรคซึมเศร้าโดยถอดหัสมาจากเสียงพูดของผู้ป่วยที่เป็นโรคผลการศึกษาจากกรณี ผู้พูดแบบReading Case มีประโยชน์และสามารถให้ข้อมูลเสริมพิจารณานอกเหนือจากแบบ Interviewing Case รวมถึงการออกแบบและสร้างตัวคัดกรองแบบ Multivariate-Parameter Classifier ที่มีคุณลักษณะเป็นไฮร์บริด หลากมิติในการขยายความสามารถจากพารามิเตอร์ต่างๆรวมเข้าไว้ด้วยกัน สามารถที่จะทำให้การคัดแยกกรองกลุ่มอาการ โรคที่เกี่ยวข้องกับอาการซึมเศร้ามีความแม่นยำถูกต้องมากขึ้นรวมถึงการออกแบบและพัฒนาขบวนการวิเคราะห์ที่สามารถจะปรับปรุงให้มีผลการทดลองมีประสิทธิภาพสูงขึ้น

บทที่ 10 ผลงานวิจัยตีพิมพ์ (Publications)

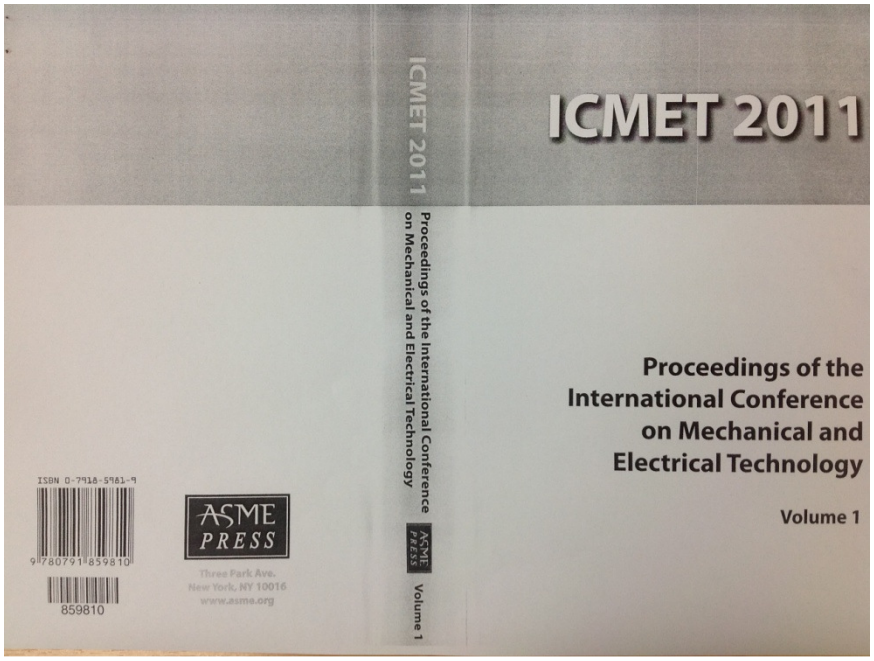


FIG.10.1 COPY OF PROCEEDING COVER

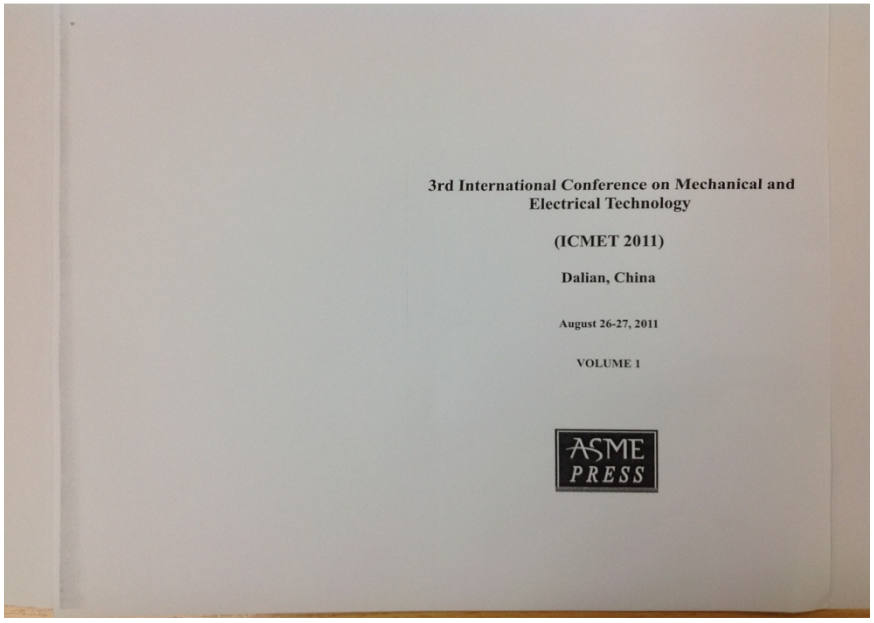
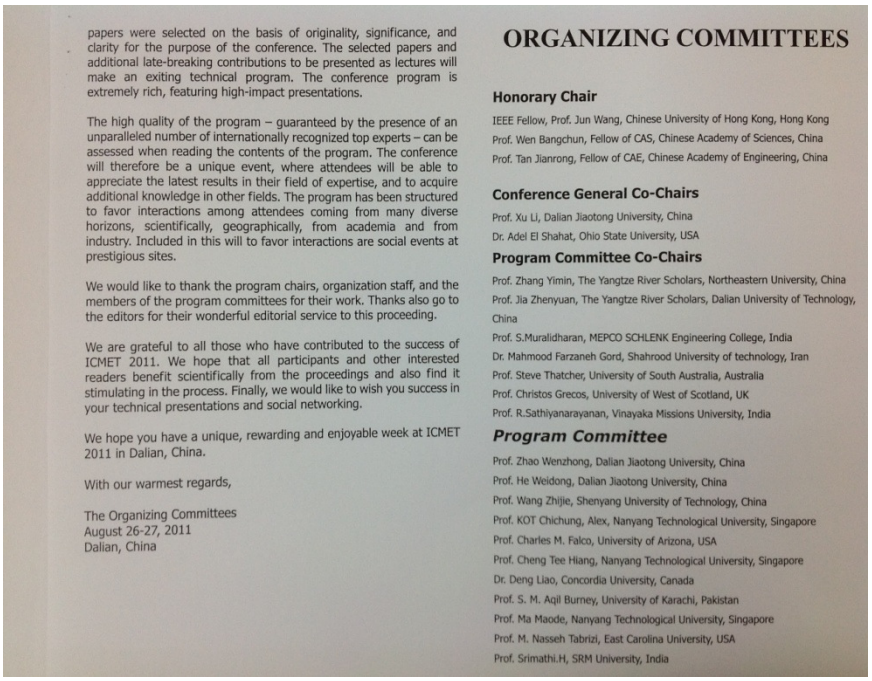


FIG.10.2 PHOTO COPY OF PROCEEDING PRESS



papers were selected on the basis of originality, significance, and clarity for the purpose of the conference. The selected papers and additional late-breaking contributions to be presented as lectures will make an exiting technical program. The conference program is extremely rich, featuring high-impact presentations.

The high quality of the program – guaranteed by the presence of an unparalleled number of internationally recognized top experts – can be assessed when reading the contents of the program. The conference will therefore be a unique event, where attendees will be able to appreciate the latest results in their field of expertise, and to acquire additional knowledge in other fields. The program has been structured to favor interactions among attendees coming from many diverse horizons, scientifically, geographically, from academia and from industry. Included in this will to favor interactions are social events at prestigious sites.

We would like to thank the program chairs, organization staff, and the members of the program committees for their work. Thanks also go to the editors for their wonderful editorial service to this proceeding.

We are grateful to all those who have contributed to the success of ICMET 2011. We hope that all participants and other interested readers benefit scientifically from the proceedings and also find it stimulating in the process. Finally, we would like to wish you success in your technical presentations and social networking.

We hope you have a unique, rewarding and enjoyable week at ICMET 2011 in Dalian, China.

With our warmest regards,

The Organizing Committees
August 26-27, 2011
Dalian, China

ORGANIZING COMMITTEES

Honorary Chair

IEEE Fellow, Prof. Jun Wang, Chinese University of Hong Kong, Hong Kong
Prof. Wen Bangchun, Fellow of CAS, Chinese Academy of Sciences, China
Prof. Tan Jianrong, Fellow of CAE, Chinese Academy of Engineering, China

Conference General Co-Chairs

Prof. Xu Li, Dalian Jiaotong University, China
Dr. Adel El Shahat, Ohio State University, USA

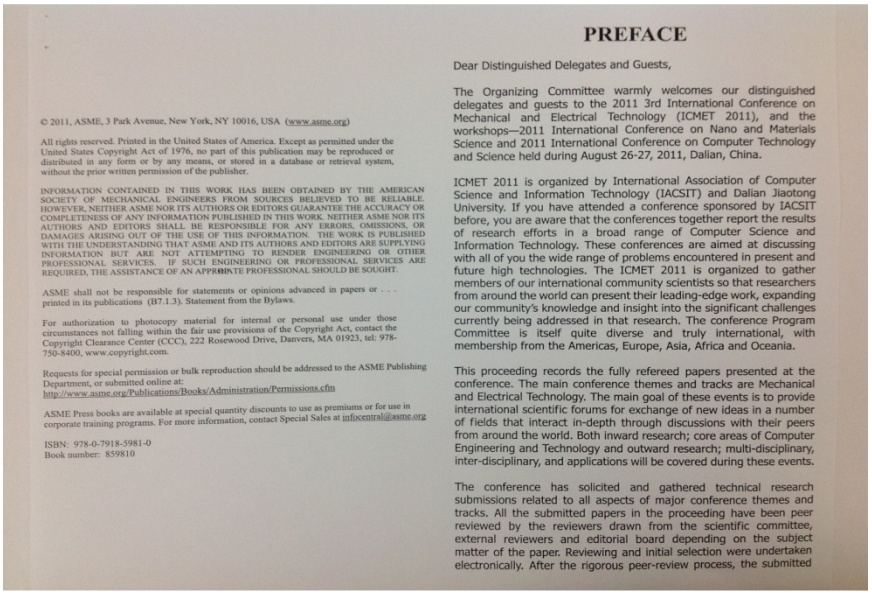
Program Committee Co-Chairs

Prof. Zhang Yimin, The Yangtze River Scholars, Northeastern University, China
Prof. Jia Zhenyuan, The Yangtze River Scholars, Dalian University of Technology, China
Prof. S.Muralidharan, MEPCO SCHLENK Engineering College, India
Dr. Mahmood Farzaneh Gord, Shahrood University of technology, Iran
Prof. Steve Thatcher, University of South Australia, Australia
Prof. Christos Grecos, University of West of Scotland, UK
Prof. R.Sathiyanarayanan, Vinayaka Missions University, India

Program Committee

Prof. Zhao Wenzhong, Dalian Jiaotong University, China
Prof. He Weidong, Dalian Jiaotong University, China
Prof. Wang Zhijie, Shenyang University of Technology, China
Prof. KOT Chichung, Alex, Nanyang Technological University, Singapore
Prof. Charles M. Falco, University of Arizona, USA
Prof. Cheng Tee Hiang, Nanyang Technological University, Singapore
Dr. Deng Liao, Concordia University, Canada
Prof. S. M. Aqil Burney, University of Karachi, Pakistan
Prof. Ma Maode, Nanyang Technological University, Singapore
Prof. M. Nasseh Tabrizi, East Carolina University, USA
Prof. Srimathi.H, SRM University, India

FIG.10.3 PHOTO COPY OF ORGANIZING COMMITTEES PAGE



© 2011, ASME, 3 Park Avenue, New York, NY 10016, USA (www.asme.org)

All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

INFORMATION CONTAINED IN THIS WORK HAS BEEN OBTAINED BY THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS FROM SOURCES BELIEVED TO BE RELIABLE. HOWEVER, NEITHER ASME NOR ITS AUTHORS OR EDITORS GUARANTEE THE ACCURACY OR COMPLETENESS OF ANY INFORMATION PUBLISHED IN THIS WORK. NEITHER ASME NOR ITS AUTHORS AND EDITORS SHALL BE RESPONSIBLE FOR ANY ERRORS, OMISSIONS, OR DAMAGES ARISING OUT OF THE USE OF THIS INFORMATION. THE WORK IS PUBLISHED WITH THE UNDERSTANDING THAT ASME AND ITS AUTHORS AND EDITORS ARE SUPPLYING INFORMATION BUT ARE NOT ATTEMPTING TO RENDER ENGINEERING OR OTHER PROFESSIONAL SERVICES. IF SUCH ENGINEERING OR PROFESSIONAL SERVICES ARE REQUIRED, THE ASSISTANCE OF AN APPROPRIATE PROFESSIONAL SHOULD BE SOUGHT.

ASME shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1.3). Statement from the Bylaws.

For authorization to photocopy material for internal or personal use under those circumstances not falling within the fair use provisions of the Copyright Act, contact the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, tel: 978-750-8400, www.copyright.com.

Requests for special permission or bulk reproduction should be addressed to the ASME Publishing Department, or submitted online at:
http://www.asme.org/Publications/Books/Administration/Permissions.cfm

ASME Press books are available at special quantity discounts to use as premiums or for use in corporate training programs. For more information, contact Special Sales at infocentral@asme.org

ISBN: 978-0-7918-5981-0
Book number: 859810

PREFACE

Dear Distinguished Delegates and Guests,

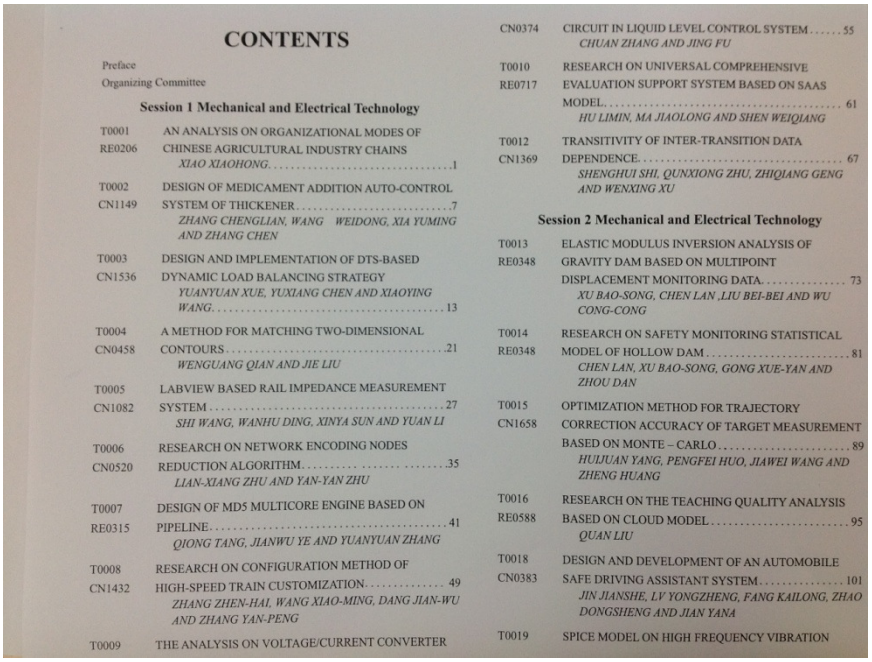
The Organizing Committee warmly welcomes our distinguished delegates and guests to the 2011 3rd International Conference on Mechanical and Electrical Technology (ICMET 2011), and the workshops—2011 International Conference on Nano and Materials Science and 2011 International Conference on Computer Technology and Science held during August 26-27, 2011, Dalian, China.

ICMET 2011 is organized by International Association of Computer Science and Information Technology (IACSIT) and Dalian Jiaotong University. If you have attended a conference sponsored by IACSIT before, you are aware that the conferences together report the results of research efforts in a broad range of Computer Science and Information Technology. These conferences are aimed at discussing with all of you the wide range of problems encountered in present and future high technologies. The ICMET 2011 is organized to gather members of our international community scientists so that researchers from around the world can present their leading-edge work, expanding our community's knowledge and insight into the significant challenges currently being addressed in that research. The conference Program Committee is itself quite diverse and truly international, with membership from the Americas, Europe, Asia, Africa and Oceania.

This proceeding records the fully refereed papers presented at the conference. The main conference themes and tracks are Mechanical and Electrical Technology. The main goal of these events is to provide international scientific forums for exchange of new ideas in a number of fields that interact in-depth through discussions with their peers from around the world. Both inward research; core areas of Computer Engineering and Technology and outward research; multi-disciplinary, inter-disciplinary, and applications will be covered during these events.

The conference has solicited and gathered technical research submissions related to all aspects of major conference themes and tracks. All the submitted papers in the proceeding have been peer reviewed by the reviewers drawn from the scientific committee, external reviewers and editorial board depending on the subject matter of the paper. Reviewing and initial selection were undertaken electronically. After the rigorous peer-review process, the submitted

Fig.10.4 Photo Copy of Preface page



CONTENTS

Preface
Organizing Committee

Session 1 Mechanical and Electrical Technology

T0001 RE0206 AN ANALYSIS ON ORGANIZATIONAL MODES OF CHINESE AGRICULTURAL INDUSTRY CHAINS *XIAO XIAOHONG* 1

T0002 CN1149 DESIGN OF MEDICAMENT ADDITION AUTO-CONTROL SYSTEM OF THICKENER 7
ZHANG CHENGLIAN, WANG WEIDONG, XIA YUMING AND ZHANG CHEN

T0003 CN1536 DESIGN AND IMPLEMENTATION OF DTS-BASED DYNAMIC LOAD BALANCING STRATEGY
YUANYUAN XUE, YUXIANG CHEN AND XIAOYING WANG 13

T0004 CN0458 A METHOD FOR MATCHING TWO-DIMENSIONAL CONTOURS 21
WENGUANG QIAN AND JIE LIU

T0005 CN1082 LABVIEW BASED RAIL IMPEDANCE MEASUREMENT SYSTEM 27
SHI WANG, WANHU DING, XINYA SUN AND YUAN LI

T0006 CN0520 RESEARCH ON NETWORK ENCODING NODES REDUCTION ALGORITHM 35
LIAN-XIANG ZHU AND YAN-YAN ZHU

T0007 RE0315 DESIGN OF MD5 MULTICORE ENGINE BASED ON PIPELINE 41
QIONG TANG, JIANWU YE AND YUANYUAN ZHANG

T0008 CN1432 RESEARCH ON CONFIGURATION METHOD OF HIGH-SPEED TRAIN CUSTOMIZATION 49
ZHANG ZHEN-HAI, WANG XIAO-MING, DANG JIAN-WU AND ZHANG YAN-PENG

T0009 THE ANALYSIS ON VOLTAGE/CURRENT CONVERTER

CN0374 CIRCUIT IN LIQUID LEVEL CONTROL SYSTEM 55
CHUAN ZHANG AND JING FU

T0010 RE0717 RESEARCH ON UNIVERSAL COMPREHENSIVE EVALUATION SUPPORT SYSTEM BASED ON SAAS MODEL 61
HU LIMIN, MA JIAOLONG AND SHEN WEIQIANG

T0012 CN1369 TRANSITIVITY OF INTER-TRANSITION DATA DEPENDENCE 67
SHENGHUI SHI, QUNXIONG ZHU, ZHIQIANG GENG AND WENXING XU

Session 2 Mechanical and Electrical Technology

T0013 RE0348 ELASTIC MODULUS INVERSION ANALYSIS OF GRAVITY DAM BASED ON MULTIPOINT DISPLACEMENT MONITORING DATA 73
XU BAO-SONG, CHEN LAN ,LIU BEI-BEI AND WU CONG-CONG

T0014 RE0348 RESEARCH ON SAFETY MONITORING STATISTICAL MODEL OF HOLLOW DAM 81
CHEN LAN, XU BAO-SONG, GONG XUE-YAN AND ZHOU DAN

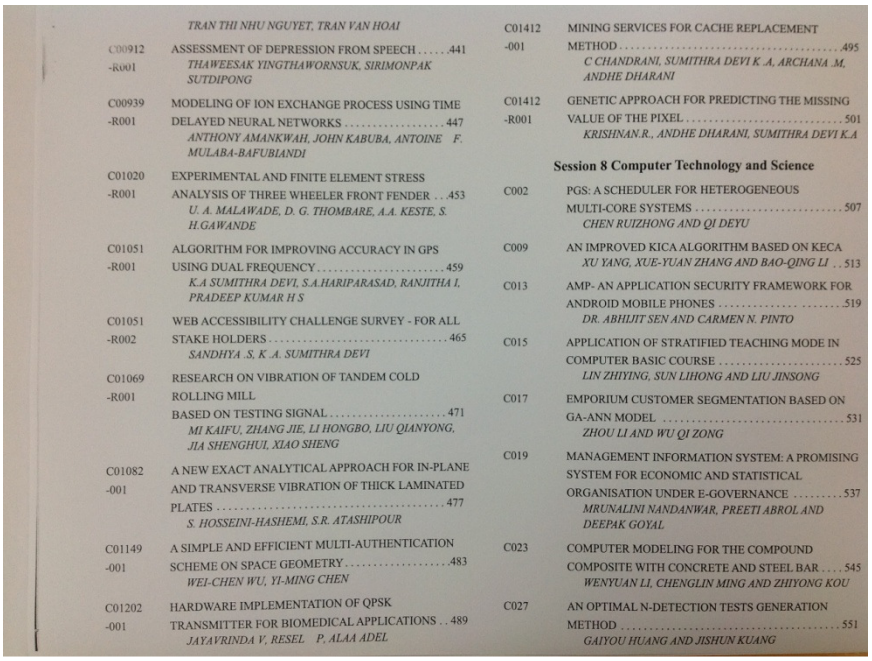
T0015 CN1658 OPTIMIZATION METHOD FOR TRAJECTORY CORRECTION ACCURACY OF TARGET MEASUREMENT BASED ON MONTE – CARLO 89
HUIJUAN YANG, PENGFEI HUO, JIAWEI WANG AND ZHENG HUANG

T0016 RE0588 RESEARCH ON THE TEACHING QUALITY ANALYSIS BASED ON CLOUD MODEL 95
QUAN LIU

T0018 CN0383 DESIGN AND DEVELOPMENT OF AN AUTOMOBILE SAFE DRIVING ASSISTANT SYSTEM 101
JIN JIANSHE, LV YONGZHENG, FANG KAILONG, ZHAO DONGSHENG AND JIAN YANA

T0019 SPICE MODEL ON HIGH FREQUENCY VIBRATION

Fig.10.5 Photo Copy of Table of Contents



TRAN THI NHU NGUYET, TRAN VAN HOAI

C00912 -R001 ASSESSMENT OF DEPRESSION FROM SPEECH 441
THAWEESAK YINGTHAWORNSUK, SIRIMONPAK SUTDIPONG

C00939 -R001 MODELING OF ION EXCHANGE PROCESS USING TIME DELAYED NEURAL NETWORKS 447
ANTHONY AMANKWAH, JOHN KABUBA, ANTOINE F. MULABA-BAFUBIANDI

C01020 -R001 EXPERIMENTAL AND FINITE ELEMENT STRESS ANALYSIS OF THREE WHEELER FRONT FENDER 453
U. A. MALAWADE, D. G. THOMBARE, A.A. KESTE, S. H.GAWANDE

C01051 -R001 ALGORITHM FOR IMPROVING ACCURACY IN GPS USING DUAL FREQUENCY 459
K.A SUMITHRA DEVI, S.A.HARIPARASAD, RANJITHA I, PRADEEP KUMAR H S

C01051 -R002 WEB ACCESSIBILITY CHALLENGE SURVEY - FOR ALL STAKE HOLDERS 465
SANDHYA .S, K .A. SUMITHRA DEVI

C01069 -R001 RESEARCH ON VIBRATION OF TANDEM COLD ROLLING MILL BASED ON TESTING SIGNAL 471
MI KAIFU, ZHANG JIE, LI HONGBO, LIU QIANYONG, JIA SHENGHUI, XIAO SHENG

C01082 -001 A NEW EXACT ANALYTICAL APPROACH FOR IN-PLANE AND TRANSVERSE VIBRATION OF THICK LAMINATED PLATES 477
S. HOSSEINI-HASHEMI, S.R. ATASHIPOUR

C01149 -001 A SIMPLE AND EFFICIENT MULTI-AUTHENTICATION SCHEME ON SPACE GEOMETRY 483
WEI-CHEN WU, YI-MING CHEN

C01202 -001 HARDWARE IMPLEMENTATION OF QPSK TRANSMITTER FOR BIOMEDICAL APPLICATIONS 489
JAYAVRINDA V, RESEL P, ALAA ADEL

C01412 -001 MINING SERVICES FOR CACHE REPLACEMENT METHOD 495
C CHANDRANI, SUMITHRA DEVI K .A, ARCHANA .M, ANDHE DHARANI

C01412 -R001 GENETIC APPROACH FOR PREDICTING THE MISSING VALUE OF THE PIXEL 501
KRISHNAN.R., ANDHE DHARANI, SUMITHRA DEVI K.A

Session 8 Computer Technology and Science

C002 PGS: A SCHEDULER FOR HETEROGENEOUS MULTI-CORE SYSTEMS 507
CHEN RUIZHONG AND QI DEYU

C009 AN IMPROVED KICA ALGORITHM BASED ON KECA
XU YANG, XUE-YUAN ZHANG AND BAO-QING LI 513

C013 AMP- AN APPLICATION SECURITY FRAMEWORK FOR ANDROID MOBILE PHONES 519
DR. ABHIJIT SEN AND CARMEN N. PINTO

C015 APPLICATION OF STRATIFIED TEACHING MODE IN COMPUTER BASIC COURSE 525
LIN ZHIYING, SUN LIHONG AND LIU JINSONG

C017 EMPORIUM CUSTOMER SEGMENTATION BASED ON GA-ANN MODEL 531
ZHOU LI AND WU QI ZONG

C019 MANAGEMENT INFORMATION SYSTEM: A PROMISING SYSTEM FOR ECONOMIC AND STATISTICAL ORGANISATION UNDER E-GOVERNANCE 537
MRUNALINI NANDANWAR, PREETI ABROL AND DEEPAK GOYAL

C023 COMPUTER MODELING FOR THE COMPOUND COMPOSITE WITH CONCRETE AND STEEL BAR 545
WENYUAN LI, CHENGLIN MING AND ZHIYONG KOU

C027 AN OPTIMAL N-DETECTION TESTS GENERATION METHOD 551
GAIYOU HUANG AND JISHUN KUANG

Fig.10.6 Photo Copy of Table of Contents (Cont…)

THAWEESAK YINGTHAWORNSUK
King Mongkut's University of Technology Thonburi, Thailand

SIRIMONPAK SUTDIPONG
King Mongkut's University of Technology Thonburi, Thailand

ABSTRACT

Spectral characteristics of speech signal have formerly been studied and reported for its quantitative information in relating the emotional state in persons suffering depression to the level of symptom affect. The spectral energy distributing in the particular bandwidths of audio frequency less than 5 KHz extracted from the subjects' speech sample reveals as the distinct parameter with a property that can represent the severity of symptom in term of energy shifting and amount change found in this study empirically. The acoustical analysis focused on Power Spectral Density (PSD) was performed on speech samples extracted from three different databases collected from the remitted, depressed, and suicidal patient groups. The result of classifications suggests the significant potential of PSD's in separating all categorized patient groups which consistently verifies that the psychomotor disturbance affects speech production system in speaker, clinically known as the central element in depression and other high-risk emotional disorders.

KEY WORDS

Speech, Power Spectral Density, Depression, Classification

1. INTRODUCTION

Committing suicide is the compulsory action in person resulted from a severe depression elevating critically in his/her state of mind. This phenomenon remains poor understanding and underresearched. The ways to prevent suicide are monitoring the symptom of depression persons are suffering by care taker, and judging on how severe degree of symptom he/she experiences by physician. The particular technique to help pinpoint a person who is at an elevated risk of committing suicide driven by severe depression are sorely needed in clinical practice. The main objective of this study is to present that the pattern of vocal frequency response is affectively changed by depression, and to reveal that there is an existence of relation between acoustical properties in spoken sound of patient and emotional illness that patient experiences. Another task is an attempt to suggest the effective vocal parameter as a single feature or a combination set that can be used to monitor the measurable change of spectral energy due to affect in speech. The findings obtained from the study could lead us to the development of new medical technology capable of monitoring and assessing the risk of suicide in patients with severe degree of depression.

In formerly published pilot works [2], [4], [5], the analytical techniques have been developed to determine if subjects were categorized in one of three mental states: healthy control, non-suicidal depressed, or high-risk suicidal. The initial sets of recordings analyzed in such studies were made in a wide variety of

Fig.10.7 Photo Copy of the published research work page 1

clinical and technical conditions, without the advantages of acoustically controlled environment and high fidelity in recording. In the past studies the clinical tapes recorded in interviews were extracted for audio records by selecting the segments of audio signal based on the practitioners' decision with expertise in speech science. All audio tapes were sampled at random that makes diagnostic groupings entirely separable.

The vocal cues have been properly used as indicators in diagnosing the underlying symptom in patient by experienced clinicians [2, 6], but these skills are not in widespread clinical use. The considerable evidence suggests that the emotional arousal produces changes in the speech production scheme by affecting the respiratory, phonatory, and articulatory processes that in turn are encoded in speech signal itself [7, 8]. Changes in heart rate, blood pressure, respiratory patterns, muscle tension, and motor activity transiently alter respiratory, phonatory, and articulatory functions in speech production in an *acutely* state-related fashion, directly tied to emotions. Consequently, emotional disturbances can be expected to cause measurable changes in speech parameters. Emotional content of the voice can be associated with acoustical variables such as level, range, contour, and perturbation of the fundamental frequency, energy of speech, energy distribution in frequency range, peak location, bandwidth and formant intensity, and a variety of temporal measures [9]. The formerly published researches have shown that depression has a major affect on the acoustic characteristics of voice as compared to that in normal controls. This paper is organized as follows: Second section provides the descriptions on database of speech recordings, patient populations, feature extraction, and statistical analyses. Next section presents results, discussion and conclusion.

2. DATABASE OF SPEECH RECORDING

All recordings were obtained from three different categorized groups of depressed, high-risk suicidal and remitted subjects. Database consists of thirty males evenly grouped into three categories. Each subject in each group has two different types of speech samples recorded. One is the speech sample collected from main interviewing session with psychiatrist and another is collected from the post session that patient reads a predetermined part of book. The passage used in post session is composed of the standardized texts generally used in speech science since it contains all of normal sounds in spoken English and it is phonetically balanced [3]. In processing state, the raw speech sample was randomly extracted from both databases in representing each subject. All speech samples are off-line analyzed and processed throughout the entire procedure in this study. First each speech signal was digitized via a 16-bit analog-to-digital (A/D) converter at a 10-kHz sampling rate with an anti-aliasing filter (i.e., 5-kHz low-pass). The background noise and other voices not belonged to patient are removed via monitoring through a GOLDWAVE v.5.10 audio editor. Before we extract the vocal feature PSD's, the edited speech signal is segmented into 20-second portions for entire database. After all 20-second segments are tested for voicing, only voiced speech samples are kept for further detrending and normalization steps which necessarily need in processing to compensate for possible differences in recording level among all recorded speech samples.

3. FEATURE EXTRACTION

Power spectral densities (PSD's) are estimated from voiced speech signal based on the classical Welch method associated with non-overlapping 100-point Hamming windows. This estimation was implemented via algorithm written in MATLAB by using the 1024-point Fast Fourier Transforms (FFT) and 40-ms windowing on each segment of speech sample [2]. Six features are collectively determined which are peak power, frequency location, and four power features calculated in four different frequency ranges: 0-500 Hz, 500-1000 Hz, 1000-1500 Hz, and 1500-2000 Hz. The percentages of total power obtained from 500-Hz bandwidths are called sub-bands (PSD_1, PSD_2, PSD_3, and PSD_4) which are a set of input feature vectors for the classifier.

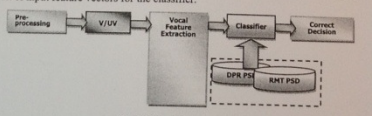


Fig.1 Extraction system model

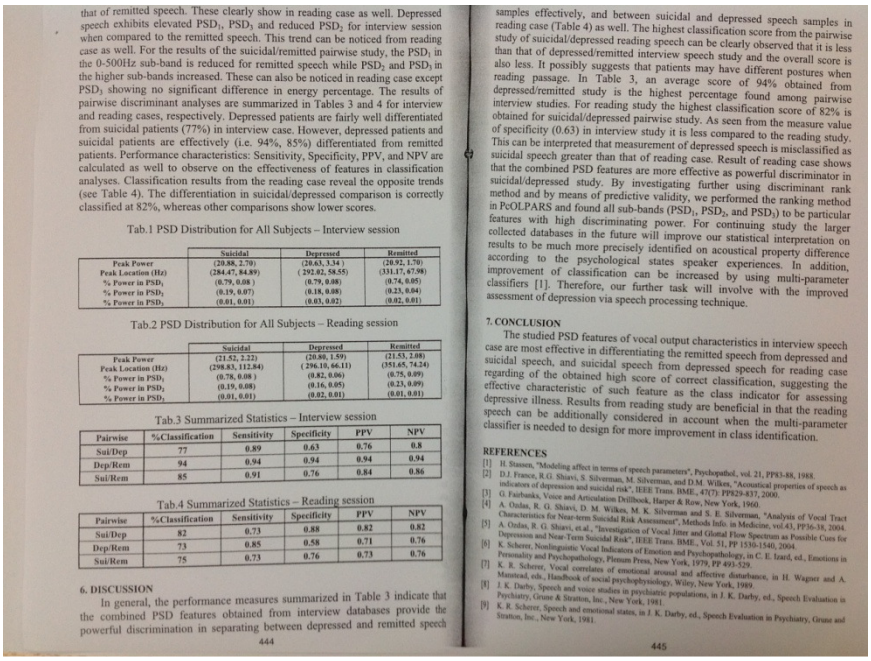
4. PAIRWISE CLASSIFICATION

All estimated acoustical features are arranged and stored in matrix form for statistical analysis. In each subject, each of the five acoustical features (i.e., peak power, peak location, PSD_1, PSD_2, and PSD_3) formed an output vector of features to represent each 20-second speech segment. Each output matrix contained N rows and M columns (N x M matrix), where N is the number of means obtained from each 20-second segment of voiced speech and M is the number of acoustical features. Mathematically, acoustical features estimated from three speech databases are defined into three matrices of feature. The prepared parameter matrix representing each class is then imported into the On-Line Pattern Analysis System (PcOLPARS) and the SYSTAT a statistical package for statistical discriminant analyses. The pairwise (i.e., suicidal/depressed, depressed/remitted, remitted/suicidal) statistical analyses are performed on all vocal parameter matrices. All features are combined to design a classifier. Classification score and performance (i.e., measures of sensitivity, specificity, positive predictive (PPV), negative predictive (NPV) values) are calculated. The pair-wise statistical analyses include the calculation and comparison of class covariance matrices, application of quadratic discriminators accociated with using the hold-one-out method in compensating for small size of datbase, and 95% confidence interval.

5. RESULTS

Means and standard deviations of the set of features, peak power, peak location (Hz), and the percentages of total power from each categorized group of patients for the interview are summarized in Table 1. Suicidal speech is characterized by peak location with significantly lower frequency compared to

Fig.10.8 Photo Copy of the published research work pages 2 &3



that of remitted speech. These clearly show in reading case as well. Depressed speech exhibits elevated PSD_1, PSD_3 and reduced PSD_2 for interview session when compared to the remitted speech. This trend can be noticed from reading case as well. For the results of the suicidal/remitted pairwise study, the PSD_1 in the 0-500Hz sub-band is reduced for remitted speech while PSD_2 and PSD_3 in the higher sub-bands increased. These can also be noticed in reading case except PSD_3 showing no significant difference in energy percentage. The results of pairwise discriminant analyses are summarized in Tables 3 and 4 for interview and reading cases, respectively. Depressed patients are fairly well differentiated from suicidal patients (77%) in interview case. However, depressed patients and suicidal patients are effectively (i.e. 94%, 85%) differentiated from remitted patients. Performance characteristics: Sensitivity, Specificity, PPV, and NPV are calculated as well to observe on the effectiveness of features in classification analyses. Classification results from the reading case reveal the opposite trends (see Table 4). The differentiation in suicidal/depressed comparison is correctly classified at 82%, whereas other comparisons show lower scores.

Tab.1 PSD Distribution for All Subjects – Interview session

	Suicidal	Depressed	Remitted
Peak Power	(20.88, 2.70)	(20.63, 3.34)	(20.92, 1.70)
Peak Location (Hz)	(284.47, 84.89)	(292.02, 58.55)	(331.17, 67.98)
% Power in PSD_1	(0.79, 0.08)	(0.79, 0.08)	(0.74, 0.05)
% Power in PSD_2	(0.19, 0.07)	(0.18, 0.08)	(0.23, 0.04)
% Power in PSD_3	(0.01, 0.01)	(0.03, 0.02)	(0.02, 0.01)

Tab.2 PSD Distribution for All Subjects – Reading session

	Suicidal	Depressed	Remitted
Peak Power	(21.52, 2.22)	(20.80, 1.59)	(21.53, 3.08)
Peak Location (Hz)	(298.83, 112.84)	(296.10, 66.11)	(351.65, 74.24)
% Power in PSD_1	(0.78, 0.08)	(0.82, 0.06)	(0.75, 0.09)
% Power in PSD_2	(0.19, 0.08)	(0.16, 0.05)	(0.23, 0.09)
% Power in PSD_3	(0.01, 0.01)	(0.02, 0.01)	(0.01, 0.01)

Tab.3 Summarized Statistics – Interview session

Pairwise	%Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	77	0.89	0.63	0.76	0.8
Dep/Rem	94	0.94	0.94	0.94	0.94
Sui/Rem	85	0.91	0.76	0.84	0.86

Tab.4 Summarized Statistics – Reading session

Pairwise	%Classification	Sensitivity	Specificity	PPV	NPV
Sui/Dep	82	0.73	0.88	0.82	0.81
Dep/Rem	73	0.85	0.58	0.71	0.76
Sui/Rem	75	0.73	0.76	0.73	0.76

6. DISCUSSION

In general, the performance measures summarized in Table 3 indicate that the combined PSD features obtained from interview databases provide the powerful discrimination in separating between depressed and remitted speech

samples effectively, and between suicidal and depressed speech samples in reading case (Table 4) as well. The highest classification score from the pairwise study of suicidal/depressed reading speech can be clearly observed that it is less than that of depressed/remitted interview speech study and the overall score is also less. It possibly suggests that patients may have different postures when reading passage. In Table 3, an average score of 94% obtained from depressed/remitted study is the highest percentage found among pairwise interview studies. For reading study the highest classification score of 82% is obtained for suicidal/depressed pairwise study. As seen from the measure value of specificity (0.63) in interview study it is less compared to the reading study. This can be interpreted that measurement of depressed speech is misclassified as suicidal speech greater than that of reading case. Result of reading case shows that the combined PSD features are more effective as powerful discriminator in suicidal/depressed study. By investigating further using discriminant rank method and by means of predictive validity, we performed the ranking method in PcOLPARS and found all sub-bands (PSD_1, PSD_2, and PSD_3) to be particular features with high discriminating power. For continuing study the larger collected databases in the future will improve our statistical interpretation on results to be much more precisely identified on acoustical property difference according to the psychological states speaker experiences. In addition, improvement of classification can be increased by using multi-parameter classifiers [1]. Therefore, our further task will involve with the improved assessment of depression via speech processing technique.

7. CONCLUSION

The studied PSD features of vocal output characteristics in interview speech case are most effective in differentiating the remitted speech from depressed and suicidal speech, and suicidal speech from depressed speech for reading case regarding of the obtained high score of correct classification, suggesting the effective characteristic of such feature as the class indicator for assessing depressive illness. Results from reading study are beneficial in that the reading speech can be additionally considered in account when the multi-parameter classifier is needed to design for more improvement in class identification.

REFERENCES

[1] H. Stassen, "Modeling affect in terms of speech parameters", Psychopathol., vol. 21, PP83-88, 1988.

[2] D.J. France, R.G. Shiavi, S. Silverman, M. Silverman, and D.M. Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk", IEEE Trans. BME., 47(7): PP829-837, 2000.

[3] G. Fairbanks, Voice and Articulation Drillbook, Harper & Row, New York, 1960.

[4] A. Ozdas, R. G. Shiavi, D. M. Wilkes, M. K. Silverman and S. E. Silverman, "Analysis of Vocal Tract Characteristics for Near-term Suicidal Risk Assessment", Methods Info. in Medicine, vol.43, PP36-38, 2004.

[5] A. Ozdas, R. G. Shiavi, et.al., "Investigation of Vocal Jitter and Glottal Flow Spectrum as Possible Cues for Depression and Near-Term Suicidal Risk", IEEE Trans. BME., Vol. 51, PP 1530-1540, 2004.

[6] K. Scherer, Nonlinguistic Vocal Indicators of Emotion and Psychopathology, in C. E. Izard, ed., Emotions in Personality and Psychopathology, Plenum Press, New York, 1979, PP 493-529.

[7] K. R. Scherer, Vocal correlates of emotional arousal and affective disturbance, in H. Wagner and A. Manstead, eds., Handbook of social psychophysiology, Wiley, New York, 1989.

[8] J. K. Darby, Speech and voice studies in psychiatric populations, in J. K. Darby, ed., Speech Evaluation in Psychiatry, Grune & Stratton, Inc., New York, 1981.

[9] K. R. Scherer, Speech and emotional states, in J. K. Darby, ed., Speech Evaluation in Psychiatry, Grune and Stratton, Inc., New York, 1981.

Fig.10.9 Photo Copy of the published research work pages 4 &5

References

[1] H. Stassen, "Modeling affect in terms of speech parameters", Psychopathol., vol. 21, PP83-88, 1988.

[2] D.J. France, R.G. Shiavi, S. Silverman, M. Silverman, and D.M. Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk", IEEE Trans. BME., 47(7): PP829-837, 2000.

[3] G. Fairbanks, Voice and Articulation Drillbook, Harper & Row, New York, 1960.

[4] A. Ozdas, R. G. Shiavi, D. M. Wilkes, M. K. Silverman and S. E. Silverman, "Analysis of Vocal Tract Characteristics for Near-term Suicidal Risk Assessment", Methods Info. in Medicine, vol.43, PP36-38, 2004.

[5] A. Ozdas, R. G. Shiavi, et.al., "Investigation of Vocal Jitter and Glottal Flow Spectrum as Possible Cues for Depression and Near-Term Suicidal Risk", IEEE Trans. BME., Vol. 51, PP 1530-1540, 2004.

[6] K. Scherer, Nonlinguistic Vocal Indicators of Emotion and Psychopathology, in C. E. Izard, ed., Emotions in Personality and Psychopathology, Plenum Press, New York, 1979, PP 493-529.

[7] K. R. Scherer, Vocal correlates of emotional arousal and affective disturbance, in H. Wagner and A. Manstead, eds., Handbook of social psychophysiology, Wiley, New York, 1989.

[8] J. K. Darby, Speech and voice studies in psychiatric populations, in J. K. Darby, ed., Speech Evaluation in Psychiatry, Grune & Stratton, Inc., New York, 1981.

[9] K. R. Scherer, Speech and emotional states, in J. K. Darby, ed., Speech Evaluation in Psychiatry, Grune and Stratton, Inc., New York, 1981.