

ระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร

นายกมลชัย โชคชัยชุตติกุล



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2539

ISBN 974-633-827-7

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

I 17528859

THAI DOCUMENT-TO-SPEECH SYSTEM

Mr. Kamolchai Chokchaichutikul

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science
Department of Computer Engineering

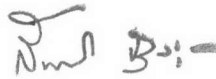
Graduate School

1996

ISBN 974-633-827-7

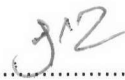
หัวข้อวิทยานิพนธ์ ระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร
โดย นายกมลชัย โชคชัยชุตติกุล
ภาควิชา วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา รองศาสตราจารย์สมชาย ทยานยง

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้
เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต



.....คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.สันติ ฤงสูววรรณ)

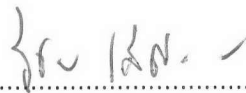
คณะกรรมการสอบวิทยานิพนธ์



.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์สุเมธ วัชรชัยสุรพล)



.....อาจารย์ที่ปรึกษา
(รองศาสตราจารย์สมชาย ทยานยง)



.....กรรมการ
(ผู้ช่วยศาสตราจารย์วิชาญ เลิศวิภาตระกูล)



.....กรรมการ
(อาจารย์จารุมาตร ปิ่นทอง)

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

กมลชัย โชคชัยชุตติกุล : ระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร
(THAI DOCUMENT-TO-SPEECH SYSTEM) อ.ที่ปรึกษา : รศ.สมชาย
ทยานอง, 75 หน้า. ISBN 974-633-827-7

ในยุคของข้อมูลข่าวสารเช่นปัจจุบัน สื่อต่างๆโดยเฉพาะสื่อที่เป็นสิ่งพิมพ์ นับว่าเป็นสื่อที่มีความสำคัญมากของบุคคล แต่สำหรับคนตาบอด ซึ่งไม่สามารถอ่านหนังสือได้ ทำให้ขาดโอกาสในการรับรู้ข้อมูลข่าวสารเหล่านี้

งานวิจัยนี้จึงมีวัตถุประสงค์ เพื่อพัฒนาระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร เพื่อให้คนตาบอดสามารถอ่านสิ่งพิมพ์ต่างๆได้ด้วยตัวเอง โดยการใช้เครื่องกวาดตรวจด้วยแสง (SCANNER) อ่านภาพเอกสารเข้าไปเก็บไว้ในเครื่องคอมพิวเตอร์ ระบบจะทำการวิเคราะห์ภาพเอกสารนั้นว่าประกอบด้วยตัวอักษรอะไรบ้าง แล้วทำการอ่านออกเสียงออกมาเป็นคำตามข้อความในเอกสารนั้น


จากผลการทดสอบพบว่า ระบบสามารถอ่านออกเสียงจากแฟ้มข้อความได้ถูกต้องเกือบ 100 เปอร์เซ็นต์ คุณภาพของเสียงที่ได้ก็มีความคมชัด เนื่องจากเป็นเสียงที่บันทึกจากเสียงคนจริงๆ ส่วนการอ่านออกเสียงจากภาพเอกสาร ความถูกต้องในการอ่านออกเสียงจะขึ้นอยู่กับคุณภาพของภาพเอกสารที่นำมาอ่าน ว่ามีสัญญาณรบกวนมากน้อยเพียงใด ซึ่งจากการทดสอบอ่านจากภาพเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์ พบว่ามีความถูกต้องประมาณ 85 เปอร์เซ็นต์ โดยในการวิเคราะห์ภาพเอกสารขนาดหนึ่งหน้ากระดาษพิมพ์ A4 ใช้เวลาประมาณ 1 นาที 24 วินาที และใช้เวลาอ่านประมาณ 7 นาที 12 วินาที

ภาควิชา วิศวกรรมคอมพิวเตอร์

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ปีการศึกษา 2538

ลายมือชื่อนิติ 

ลายมือชื่ออาจารย์ที่ปรึกษา 

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

C518264 : MAJOR COMPUTER SCIENCES

KEY WORD: DOCUMENT/SPEECH/RECOGNITION

KAMOLCHAI CHOKCHAICHUTIKUL : THAI DOCUMENT-TO-SPEECH SYSTEM. THESIS ADVISOR : ASSO. PROF. SOMCHAI TAYANYONG. 75 pp., ISBN 974-633-827-7.

In this present time of information era, there are various types of media, print media especially, is essential for our life. But for people who are blind and unable to read, have missed the chance to receive this kind of information.

This research's objective is to develop a computer system that can read thai language document and generate voice output. By using scanner to read the printed documents and store it in a computer, and after that the system will analyse and produce voice output from those printed documents.

From the testing of the system it was found that nearly 100% accuracy from reading the text files and the quality of voice was very clear since it was recorded by human voice. The accuracy of the reading depended on the quality of documents to be read. As the document printed from laser printer was found about 85% of accuracy. It took about 1 minute 24 second to analyse, and about 7 minutes 12 second to read an A4 printed sheet.

ภาควิชา.....วิศวกรรมคอมพิวเตอร์.....

สาขาวิชา.....วิทยาศาสตร์คอมพิวเตอร์.....

ปีการศึกษา.....2538.....

ลายมือชื่อนิสิต.....

ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....



กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ ได้สำเร็จลุล่วงไปด้วยความช่วยเหลืออย่างยิ่งของ รศ. สมชาย ทยานง อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้กรุณาให้คำแนะนำและข้อคิดเห็นต่างๆในการทำวิจัยมาด้วยดีตลอด

ขอขอบคุณ อาจารย์จรรยาพร ปิ่นทอง, ผศ. วิชาญ เลิศวิภาตระกูล และ ผศ. สุเมธ วัชรชัยสุรพล ที่ได้กรุณาให้คำแนะนำและข้อคิดเห็นในการทำวิจัยครั้งนี้ด้วย ขอขอบคุณคณาจารย์ทุกท่านในภาควิชาวิศวกรรมคอมพิวเตอร์ ที่ได้ให้โอกาสผู้วิจัย เข้าศึกษาและทำวิจัยจนสำเร็จลุล่วงไปด้วยดี นอกจากนี้ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ชวนิสิตปริญญาโททุกท่านที่คอยเป็นกำลังใจตลอดมา

ท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา มารดา รวมทั้งพี่ และ น้อง ที่คอยให้กำลังใจแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญภาพ.....	ญ
<u>บทที่ 1 บทนำ.....</u>	1
- ความเป็นมาของปัญหา.....	1
- วัตถุประสงค์ของการวิจัย.....	2
- ขอบเขตของการวิจัย.....	2
- เครื่องมือที่ใช้ในการวิจัย.....	2
- ขั้นตอนและวิธีดำเนินงาน.....	3
<u>บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....</u>	4
2.1 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1.1 ทฤษฎีที่เกี่ยวกับการรู้จำตัวอักษร.....	4
2.1.1.1 การจับคู่เปรียบเทียบกับอักขระต้นแบบ.....	4
2.1.1.2 การตรวจรู้จากข้อมูลทางสถิติ.....	5
2.1.1.3 การตรวจรู้จากหลักโครงสร้างของอักขระ.....	5

2.1.2 ทฤษฎีที่เกี่ยวข้องกับการอ่านออกเสียงจากเพิ่มข้อความ.....	6
2.1.2.1 ขั้นตอนที่ทำกรรับข้อความเข้ามา.....	7
2.1.2.2 ขั้นตอนที่ทำหน้าที่สังเคราะห์เสียงตามชุดสัญลักษณ์ แทนเสียงและพารามิเตอร์.....	8
2.2 งานวิจัยที่เกี่ยวข้อง.....	9
2.2.1 งานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษร.....	9
2.2.2 งานวิจัยที่เกี่ยวข้องกับการอ่านออกเสียงจากเพิ่มข้อความ.....	10
<u>บทที่ 3 หลักการทำงานของโปรแกรม.....</u>	13
3.1 การวิเคราะห์ข้อมูลนำเข้า.....	13
3.2 ขั้นตอนการแปลงภาพเอกสารให้เป็นข้อความ.....	19
3.2.1 การตัดแบ่งบรรทัดตัวอักษร.....	19
3.2.2 การตัดแบ่งตัวอักษร.....	20
3.2.3 การเปรียบเทียบกับอักษรต้นแบบ.....	24
3.2.4 การจัดเรียงลำดับตัวอักษรใหม่.....	30
3.3 ขั้นตอนการแปลงข้อความให้เป็นเสียง.....	33
3.3.1 การขยายคำย่อในข้อความให้เป็นคำเต็ม.....	33
3.3.2 การค้นหาคำอ่านออกเสียงในต้นไม้คำศัพท์.....	34
3.3.3 การอ่านออกเสียง.....	39
<u>บทที่ 4 โครงสร้างข้อมูล.....</u>	40
4.1 โครงสร้างเพิ่มข้อมูลคอนฟิกูเรชันของระบบ.....	40
4.2 โครงสร้างข้อมูลที่เกี่ยวข้องกับการรู้จำตัวอักษร.....	40
4.3 โครงสร้างข้อมูลที่เกี่ยวข้องกับการอ่านออกเสียง.....	45

บทที่ 5 การติดตั้งและวิธีใช้โปรแกรม.....	48
5.1 การติดตั้งโปรแกรม.....	48
5.2 การใช้งานโปรแกรม.....	56
บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ.....	63
6.1 สรุปผลการวิจัย.....	63
6.2 ข้อเสนอแนะ.....	64
รายการอ้างอิง.....	67
ภาคผนวก ก.....	68
ประวัติผู้เขียน.....	75

สารบัญภาพ

		หน้า
รูปที่ 3.1	แสดงการทำงานหลักๆ ของระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร.....	13
รูปที่ 3.2	แสดงการสร้างข้อมูลต้นแบบตัวอักษร “ก”, “ข” และ “ค”.....	17
รูปที่ 3.3	แสดงการทำงานของระบบคอมพิวเตอร์อ่านออกเสียงภาษาไทยจากเอกสาร.....	18
รูปที่ 3.4	การสแกนหาขอบบนของบรรทัดตัวอักษร.....	19
รูปที่ 3.5	การสแกนหาขอบล่างของบรรทัดตัวอักษร.....	19
รูปที่ 3.6	การสแกนหาขอบซ้ายของตัวอักษร.....	20
รูปที่ 3.7	การสแกนหาขอบขวาของตัวอักษร.....	20
รูปที่ 3.8	Flow Chart แสดงการตัดแบ่งตัวอักษร.....	23
รูปที่ 3.9	แสดงตัวอักษรและข้อมูลตัวอักษรต้นแบบ.....	25
รูปที่ 3.10	แสดงการเปรียบเทียบตัวอักษร “ข” กับตัวอักษรต้นแบบ “ก”.....	28
รูปที่ 3.11	แสดงการเปรียบเทียบตัวอักษร “ข” กับตัวอักษรต้นแบบ “ข”.....	29
รูปที่ 3.12	แสดงการเปรียบเทียบตัวอักษร “ข” กับตัวอักษรต้นแบบ “ค”.....	30
รูปที่ 3.13	แสดงการสร้าง Array ของพยัญชนะ, สระบนและสระล่างเพื่อใช้ในการจัดเรียงลำดับตัวอักษร.....	31
รูปที่ 3.14	แสดงการเก็บตัวอักษรที่วิเคราะห์ได้ลงใน Array ทั้ง 3 ตามชนิดตัวอักษร.....	31
รูปที่ 3.15	แสดงการอ่านพยัญชนะตัวที่ 1 จาก Array ของพยัญชนะ.....	32
รูปที่ 3.16	แสดงการอ่านสระล่างจาก Array ของสระล่างทั้งหมดเข้ามาที่ละตัว.	32
รูปที่ 3.17	แสดงการอ่านสระบนจาก Array ของสระบนทั้งหมดเข้ามาที่ละตัว.	33
รูปที่ 3.18	แสดงการอ่านพยัญชนะตัวที่ 2 และตัวต่อไปจาก Array ของพยัญชนะเข้ามาที่ละตัวจนหมด.....	33

รูปที่ 3.19	35
รูปที่ 3.20	36
รูปที่ 3.21	37
รูปที่ 3.22	38
รูปที่ 4.1	โครงสร้างเพิ่มคำศัพท์.....	45
รูปที่ 4.2	โครงสร้างเพิ่มข้อมูลเสียง.....	45
รูปที่ 4.3	โครงสร้างข้อมูลต้นไม้อำคำศัพท์.....	47
รูปที่ 5.1	แสดงหน้าจอโปรแกรม WaveStudio.....	49
รูปที่ 5.2	แสดงหน้าจอการตั้งค่า format ของเสียงที่จะบันทึก.....	50
รูปที่ 5.3	แสดงหน้าจอ Mixer เพื่อปรับความดังของลำโพงและไมโครโฟน....	51
รูปที่ 5.4	แสดงหน้าจอการบันทึกเสียง.....	52
รูปที่ 5.5	แสดงหน้าจอขณะทำการบันทึกเสียง.....	53
รูปที่ 5.6	แสดงหน้าจอคลื่นเสียงที่บันทึกได้.....	54
รูปที่ 5.7	แสดงการคลุมช่วงเงียบของข้อมูลเสียงที่บันทึก.....	55
รูปที่ 5.8	แสดงหน้าจอเมนูหลักของระบบคอมพิวเตอร์อ่านออกเสียง ภาษาไทยจากเอกสาร.....	57
รูปที่ 5.9	หน้าจอแสดงรายชื่อเพิ่มรูปภาพเอกสาร.....	58
รูปที่ 5.10	หน้าจอแสดงรายชื่อเพิ่มข้อความ.....	60
รูปที่ 5.11	แสดงหน้าจอการบันทึกข้อความลงเพิ่ม.....	61
รูปที่ 5.12	หน้าจอแสดงรายชื่อเพิ่มรูปแบบตัวอักษร.....	62