การพัฒนาวงจรเข้ารหัสดิจิตอลจากสัญญาณเสียง

นายสุพงศ์ เกตะนันทน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
พ.ศ. 2526

ISBN 974-562-480-2

009820

i 15432920

DEVELOPMENT OF A DIGITAL ENCODER FROM VOICE SIGNAL

Mr. Supong Pekanan

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science

Department of Computer Engineering

Graduate School

Chulalongkorn University

1983

หัวข้อวิทยานิพนธ์ การพัฒนาวงจรเข้ารหัสดิจิตอลจากสัญญาณเสียง

โดย นายสุพงศ์ เภกะนันทน์

ภาควิชา วิศวกรรมคอมพิวเตอร์

อาจารย์ที่ปรึกษา รองศาสตราจารย์ สมชาย ทยานยง

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร. สุประดิษฐ์ บุนนาค)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(รองศาสตราจารย์ เดือน สินธุ์พันธ์ประทุม)

..... กรรมการ

(รองศาสตราจารย์ สมชาย ทยานยง)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ สุยุชน์ สัตยประกอบ)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ วิชาญ เลิศวิภาตระกูล)

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

หัวข้อวิทยานิพนธ์	การพัฒนาวงจรเข้ารหัสดิจิตอลจากสัญญาณเสียง
ชื่อนิสิต	นายสุพงศ์ เภกะนันทน์
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ สมชาย ทยานยง
ภาควิชา	วิศวกรรมคอมพิวเตอร์
ปีการศึกษา	2525

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อ

การสร้างวงจรเข้ารหัสดิจิตอลจากสัญญาณเสียง เป็นลักษณะของการสร้างระบบประมวลผลด้วยเครื่องคอมพิวเตอร์ ซึ่งประกอบด้วยหน่วยรับข้อมูล หน่วยประมวลผล และหน่วยแสดงผลข้อมูล วงจรที่สร้างขึ้นแบ่งออกเป็นสองภาค คือ วงจรภาคสัญญาณเสียง เป็นหน่วยรับข้อมูลเสียง ทำหน้าที่แยกวิเคราะห์สัญญาณตามความถี่ แล้วเปลี่ยนให้เป็นข้อมูลทางดิจิตอล วงจรคอมพิวเตอร์เป็นหน่วยควบคุม ประมวลผล และแสดงผล ทำหน้าที่ควบคุมการสุ่มข้อมูลเสียงจากวงจรภาคสัญญาณเสียง รับข้อมูล ประมวลผล เพื่อสร้างต้นแบบอ้างอิงของเสียงพูดแต่ละคำ เปรียบเทียบเสียงที่ไม่ทราบความหมายกับต้นแบบอ้างอิง แล้วตัดสินใจเลือกความหมายรวมทั้งแสดงผล

การวิจัยนี้เป็นส่วนหนึ่งที่เกี่ยวข้องกับระบบจดจำคำพูด โดยได้หาทางวิเคราะห์ชี้แนวทางปัญหาและข้อเสนอแนะ ซึ่งเป็นแนวทางไปสู่การพัฒนา เพิ่มขีดความสามารถให้กับคอมพิวเตอร์ในการสื่อสารกับมนุษย์ให้ทำได้ง่ายขึ้น ซึ่งทำให้การประยุกต์นำเอาคอมพิวเตอร์ไปใช้งานได้อย่างกว้างขวาง จากผลการวิจัยพบว่า สำหรับการนับศูนย์ถึงเก้าของภาษาไทยทุกคำสามารถจำได้โดยไม่มีปัญหา ยกเว้น หนึ่ง สอง และสี่ ซึ่งยังมีปัญหาในทางปฏิบัติทำให้ผู้พูดจะต้องระมัดระวังเป็นพิเศษ ในระบบที่ออกแบบมาเพื่อการวิจัยนี้พบว่าถ้าใช้การพูดศูนย์ ถึงเก้า ตามธรรมดาแล้ว จะได้ผลลัพท์ถูกต้อง ร้อยละ 60 และถ้าผู้ใช้ได้รับการฝึกฝนเป็นพิเศษจะสามารถจำได้ถูกต้องมากกว่าร้อยละ 90

Thesis Title	Development of A Digital Encoder from Voice Signal
Name	Mr. Supong Pekanan
Thesis Advisor	Associate Professor Somchai Thayanyong
Department	Computer Engineering
Academic Year	1982

ABSTRACT

The construction of digital encoder from voice signal is a construction of computerized processing system which is composed of input unit, processing unit and output unit. The encoder circuit is divided into two parts, audio circuit, which receives voice signal and analyzes it by frequencies then converts results to digital data, and computer circuit, as a control, a processor and a display, the functions of which are controlling the sampling of voice data from audio circuit, accepting the data and process to build reference pattern for each speech, comparing unknown voice pattern with those reference, then selecting suitable meaning and displaying results.

This research, which partially concerns with automatic speech recognition system, gives some basic idea about acoustic analysis together with problems and recommendations for future works which can be used to increase the ability of computer system for more convenience in communication with human being and which will result in more widely use of computer. The result of this

work is that, in recognition of Thai vocabularies 0-9, most of the vocabularies can be recognized easily by the system except 1, 2 and 4. The problem is that, in practical use, the speaker is needed to be more careful in pronunciation. For normal pronunciation of 0-9, the results are more than 60 percents correct and more than 90 percents correct with trained speaker.

กิตติกรรมประกาศ

การทำวิทยานิพนธ์นี้ ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ สมชาย หยานยง เป็นอย่างยิ่งที่ได้กรุณาแนะนำหัวข้อวิทยานิพนธ์ ตลอดจนแนวทางในการวิจัย ให้คำปรึกษา และจัดหาเงินทุนในการทำวิจัย ทำให้การวิจัยนี้สำเร็จลงได้ รวมทั้งคณาจารย์ในภาควิชาวิศวกรรมคอมพิวเตอร์ทุกท่าน ที่ได้กรุณาให้ความรู้ และอำนวยความสะดวกต่าง ๆ ในการทำวิจัยเป็นอย่างดียิ่ง

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ช

รายการตารางประกอบ ..... ฌ

รายการรูปประกอบ ..... ฌ

บทที่

1 บทนำ 1

1.1 ความเป็นมาของปัญหา 1

1.2 วัตถุประสงค์ 3

1.3 ประโยชน์ที่จะได้จากการวิจัย 3

1.4 ขอบเขตของการวิจัย 3

1.5 ขั้นตอนและวิธีการดำเนินการวิจัย 4

2 การศึกษาระบบประมวลผลเสียงพูด 5

2.1 การติดต่อสื่อสารโดยใช้เสียง 5

2.2 การประมวลผลสัญญาณ 5

2.3 ประวัติการค้นคว้าเกี่ยวกับระบบจดจำเสียง 8

2.4 ระบบจดจำคำพูด 9

3 การออกแบบวงจร 12

3.1 แนวทางในการออกแบบ 12

3.2 การออกแบบวงจรภาคสัญญาณเสียงและแยกสัญญาณเชิงความถี่ 13

3.3 วงจรไมโครโพรเซสเซอร์ 26

3.4 ผังการควบคุมของระบบ 28

4 การศึกษาลักษณะของคำและการพัฒนาขั้นตอนในการจดจำ 30

4.1 การศึกษาลักษณะของคำ 30

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ หน้า

4.2 การพัฒนาขั้นตอนในการจดจำ 38
4.3 ผังงานของระบบ 43
4.4 ผังงานของระบบจดจำ 46
5 สรุปผลการวิจัยและข้อเสนอแนะ 54
5.1 สรุปผลการวิจัย 54
5.2 ปัญหาในการวิจัย 55
5.3 ข้อเสนอแนะ 56
เอกสารอ้างอิง 59
ภาคผนวก
ก. แผนผังวงจรโดยละเอียด 60
ข. รายละเอียดโปรแกรมควบคุม 70
ค. ตัวอย่างข้อมูลเสียงและข้อมูลวิเคราะห์ 133
ง. ข้อมูลจำเพาะของอุปกรณ์ 143

รายการตารางประกอบ

ตารางที่ หน้า

3.1 แสดงค่าพารามิเตอร์ของวงจรกรองแถบความถี่จากการคำนวณ 18

4.1 แสดงช่วงเวลาในการออกเสียงเมื่อนับ 0-9 30

4.2 แสดงกลุ่มของคำพูดแยกตามระยะเวลาการออกเสียง 32

4.3 แสดงค่าเฉลี่ยของสัญญาณเสียงต่อหน่วยเวลาของคำพูด 0-9 34

5.1 แสดงผลการทดลองเมื่อให้ระบบจำคำพูดจำนวน 7 คำ 54

รายการรูปประกอบ

รูปที่		หน้า
2.1	แสดงขั้นตอนโดยทั่วไปของระบบข่าวสาร	6
2.2	แสดงขอบเขตของอัตราข้อมูลทางดิจิตอลที่ต้องใช้แทนสัญญาณคำพูด	7
2.3	แสดงการประยุกต์การประมวลเสียงพูดไปใช้งาน	7
2.4	แสดงขั้นตอนของระบบจดจำคำพูด	11
3.1	แสดงรูปแบบของระบบจดจำ	12
3.2	วงจรกรองผ่านความถี่ต่ำ	13
3.3	แสดงวงจรตรวจวัดระดับสัญญาณ	15
3.4	แสดงวงจรกรองผ่านแถบความถี่แบบแอคทีฟ	16
3.5	แสดงวงจรเรียงลำดับเต็มคลื่นพร้อมวงจรกรองผ่านความถี่ต่ำ	18
3.6	แสดงวงจรสร้างสัญญาณนาฬิกา	19
3.7	วงจรนับเวลาและหารความถี่	21
3.8	แสดงวงจรเปลี่ยนสัญญาณอนาล็อกเป็นดิจิตอล	23
3.9	แสดงวงจรตรวจจับการผ่านศูนย์	24
3.10	แสดงวงจรสร้างสัญญาณควบคุมการนับ	24
3.11	แสดงวงจรนับ	25
3.12	แสดงแผนผังของระบบไมโครโพรเซสเซอร์	26
3.13	แสดงการใช้หน่วยความจำ	27
3.14	แสดงผังการควบคุมระบบจดจำ	29
4.1	แสดงรูปสัญญาณของคำพูด 0-9	31
4.2	(ก) แสดงภาพสัญญาณที่มีลักษณะคล้ายคลึงกัน	32
	(ข) แสดงภาพสัญญาณที่มีลักษณะแตกต่างกัน	33
4.3	แสดงภาพขยายของสัญญาณเสียง "9" ในช่วงต้นของคำ	33
4.4	แสดงการกระจายของพลังงานเสียงไปตามแถบความถี่ต่าง ๆ	34
4.5	แสดงรูปสัญญาณที่ปรากฏในแต่ละแถบความถี่ของคำพูด 0-9	36
4.6	แสดงความแตกต่างของรูปสัญญาณในการพูดคำเดียวกัน 2 ครั้ง	37

รูปที่ หน้า

4.7 แสดงระเบียนข้อมูลเสียง 39
4.8 แสดงการจัดระเบียนข้อมูลวิเคราะห์ 40
4.9 แสดงการจัดแฟ้มข้อมูลตัวอย่างเสียงในหน่วยความจำ 41
4.10 แสดงตารางสถิติค่าเฉลี่ยของสัญญาณเสียง..... 42
5.1 แสดงการวัดตำแหน่งและการปรากฏของพลังงานเสียง
ในแถบความถี่ต่าง ๆ 57