



บทที่ 4

การวิเคราะห์และผลการทดสอบ

4.1 ข้อมูลที่ใช้ในการทดสอบ

ในการทดสอบการรู้จำเสียงพูดนี้จะแบ่งกลุ่มเสียงพูดออกเป็น 2 กลุ่มดังนี้คือ

ก. เสียงพูดกลุ่มที่ 1 จะประกอบด้วยคำพูดตัวเลขภาษาไทยจำนวน 10 คำ คือ ตัวเลขศูนย์ถึงเก้า

ข. เสียงพูดกลุ่มที่ 2 จะประกอบด้วยเสียงพูดกลุ่มที่ 1 ทั้งหมด รวมทั้งเสียงพูดคำว่า "สิบ" , "เอ็ด" , "ยี่" , "ร้อย" , "พัน" , "หมื่น" , "แสน" และ "ล้าน" ซึ่งเป็นคำเรียกของตัวเลขในภาษาไทย ในกลุ่มเสียงกลุ่มที่ 2 นี้จะมีคำทั้งสิ้น 18 คำ

ในการเก็บเสียงพูดเพื่อนำมาใช้ในการทดสอบนี้ จะทำการบันทึกจากนักศึกษาจำนวน 40 คน ซึ่งในจำนวนนี้จะแบ่งออกได้เป็น 2 กลุ่มดังนี้คือ กลุ่ม A จะทำการบันทึกเสียงของทั้งหมด 18 คำที่กล่าวมาข้างต้น โดยจะทำการบันทึกคำละ 3 ครั้ง จำนวนคนที่นำมาบันทึกจำนวน 20 คน โดยในกลุ่ม A นี้จะแบ่งเป็นกลุ่มย่อย ๆ 3 กลุ่ม กำหนดให้เป็น A1,A2,A3 ซึ่งแทนการบันทึกเสียงครั้งที่ 1 ครั้งที่ 2 และครั้งที่ 3 ตามลำดับ จำนวน 1080 คำ โดยที่ในกลุ่ม A1 นี้จะนำไปใช้ในการสร้างแบบอ้างอิง และกลุ่ม B จำนวน 20 คน โดยในกลุ่มนี้จะทำการบันทึกเสียงเฉพาะตัวเลข 0 ถึง 9 โดยจะทำการบันทึกแต่ละคำจำนวน 2 ครั้ง จำนวน 400 คำ

ในการบันทึกเสียงเหล่านี้จะทำการบันทึกผ่านการอัดเสียง ในสภาวะปกติไม่ได้ควบคุมสภาวะภายนอกใด ๆ ทั้งยังไม่ได้กำหนดความดังของเสียงคือให้ผู้พูดพูดปกติ โดยจะใช้ความถี่ของการสุ่ม 8 KHz ในการวิเคราะห์เสียงใช้ดีสครีตฮาร์ตเลย์ทรานส์ฟอร์มในการวิเคราะห์หา parameter จากนั้นกำหนดระดับของแอมพลิจูดของสัญญาณเพื่อใช้เป็น parameter ของระบบแล้วทำการเปรียบเทียบเสียงโดยใช้ DTW และกฎการตัดสินใจเป็น Nearest Neighbor (NN) เนื่องจากในวิทยานิพนธ์นี้เลือกใช้แบบอ้างอิงเพียงชุดเดียว

4.2 ผลการทดสอบ

จากผลการทดสอบในภาคผนวก ก ค่าของพารามิเตอร์ต่าง ๆ ที่นำมาใช้ในการทดสอบนี้คือ ในส่วนของการตัดค่า พารามิเตอร์ a,b, และ c จะมีค่าเป็น 0.3 ,0.1 และ 0.5 ตามลำดับ และช่วงความถี่ประมาณ 300 ถึง 3000 Hz ซึ่งตรงกับพารามิเตอร์ k ที่ 10 ถึง 191 โดยที่ไม่มีการปรับค่าขนาดของข้อมูล

ในตารางที่ 4.2.1 จะแสดงถึงการทดสอบแบบอ้างอิงของเสียงของตัวเลข 0-9 โดยใช้ข้อมูลกลุ่ม A1 ไปสร้างกลุ่มอ้างอิง ตามจำนวนผู้พูด 5,10,15 และ 20 คน ซึ่งในการทดสอบนี้จะแบ่งการทดสอบออกเป็น 3 กลุ่ม คือ

- ก. ทดสอบกับกลุ่ม A1 ตามจำนวนผู้พูดที่ใช้ในการสร้างแบบอ้างอิง
- ข. ทดสอบกับกลุ่ม A2,A3 ตามจำนวนผู้พูดในกลุ่ม A1 ที่ใช้ในการสร้างแบบอ้างอิง
- ค. ทดสอบกับกลุ่ม B

ตารางที่ 4.2.1 แสดงผลการรู้จำของเสียงพูดกลุ่มที่ 1 (0 - 9) โดยใช้แบบอ้างอิงที่จำนวนผู้พูดต่าง ๆ กัน

กลุ่มทดสอบ	ความถูกต้อง (%)			
	จำนวนผู้พูดที่ใช้ในการสร้างแบบอ้างอิง			
	5 คน	10 คน	15 คน	20 คน
ก. (A1)	98.00	94.00	92.00	90.50
ข. (A2,A3)	81.00	83.50	85.00	86.50
ค. (B)	74.25	76.75	73.75	79.25

จากตารางที่ 4.2.1 นี้แสดงให้เห็นว่า ยิ่งใช้จำนวนผู้พูดมากขึ้น ความถูกต้องในกลุ่มทดสอบ ก ซึ่งเป็นกลุ่มที่ใช้ในการสร้างแบบอ้างอิงจะลดลง เนื่องจากจำนวนผู้พูดมากขึ้น ความแตกต่างของเสียงแต่ละบุคคลมีมากขึ้น แต่กลุ่ม ข และ ค ซึ่งไม่ใช่กลุ่มของเสียงที่ใช้ในการสร้างแบบอ้างอิง จะให้ผลของอัตราการรู้จำดีขึ้น

ในตารางที่ 4.2.2-4.2.4 จะแสดงผลของการรู้จำของกลุ่มทดสอบ A1,A2-A3,และ B ตามลำดับ โดยที่แบบอ้างอิงจะสร้างจากข้อมูลเสียงในกลุ่ม A1 จำนวนผู้พูด 20 คน ซึ่งจะแจกแจงผลของการรู้จำของกลุ่มเสียงกลุ่มที่ 1 ซึ่งจากตารางทั้ง 3 ตารางนี้จะพบว่าเสียงของตัวเลขแต่ละตัวจำนวน 50 ครั้ง ที่ทำให้การรู้จำผิดพลาดมากที่สุดมีอยู่ 3 คู่ คือ "0" กับ "4" , "3" กับ "9" และ "3" กับ "9" โดยมีรายละเอียดดังนี้

- ก. เสียง "0" จะผิดพลาดเป็นเสียง "4" จำนวน 18 ครั้ง
- ข. เสียง "3" จะผิดพลาดเป็นเสียง "5" จำนวน 6 ครั้ง
- ค. เสียง "3" จะผิดพลาดเป็นเสียง "9" จำนวน 16 ครั้ง
- ง. เสียง "4" จะผิดพลาดเป็นเสียง "0" จำนวน 11 ครั้ง
- จ. เสียง "5" จะผิดพลาดเป็นเสียง "3" จำนวน 8 ครั้ง
- ฉ. เสียง "9" จะผิดพลาดเป็นเสียง "3" จำนวน 18 ครั้ง

จากผลการทดสอบที่จะเห็นได้ว่า แบบอ้างอิงของกลุ่มเสียงของตัวเลขของทั้งคู่กลุ่มดังกล่าวที่มีความผิดพลาด จะมีรูปแบบอ้างอิงที่ใกล้เคียงกัน ดังแสดงในภาคผนวก ข

ตารางที่ 4.2.2 แสดงผลการรู้จำของกลุ่ม A1 ต่อแบบอ้างอิงที่สร้างจากผู้พูด จำนวน 20 คน

แบบทดสอบ (A1)	แบบอ้างอิง (A1)									
	0	1	2	3	4	5	6	7	8	9
0	19				1					
1	1	18						1		
2			20							
3				18		1				1
4	2				18					
5				2		18				
6	1	1					18			
7		2					1	17		
8				1					18	1
9				2		1				17

(ก)

ความถูกต้อง	%
ต่ำสุด	70.00
สูงสุด	100.00
เฉลี่ย	90.50

(ข)

ตารางที่ 4.2.3 แสดงผลการรู้จำของกลุ่ม A2,A3 ต่อแบบอ้างอิงที่สร้างจากผู้พูดจำนวน 20 คน

แบบทดสอบ (A2,A3)	แบบอ้างอิง (A1)									
	0	1	2	3	4	5	6	7	8	9
0	33	2			4			1		
1	2	34			1			3		
2			39	1						
3			2	33		3				2
4	8				31		1			
5				3		34			1	2
6		3					37			
7		2					2	36		
8				2		1			36	1
9			1	5					1	33

ความถูกต้อง	%
ต่ำสุด	55.00
สูงสุด	100.00
เฉลี่ย	86.50

(ข)



(ก)

ตารางที่ 4.2.4 แสดงผลการรู้จำของกลุ่ม B ต่อแบบอ้างอิงที่สร้างจากผู้พูดจำนวน 20 คน

แบบทดสอบ (B)	แบบอ้างอิง (A1)									
	0	1	2	3	4	5	6	7	8	9
0	26	1			13					
1		32	1		4		3			
2			36	1			1		2	
3				24		2		1		13
4	1				38	1				
5				3		37				
6	2	3	1				32	2		
7		4					2	34		
8				4		4			31	1
9				11		2				27

ความถูกต้อง	%
ต่ำสุด	50.00
สูงสุด	95.00
เฉลี่ย	79.25

(ข)

(ก)

ในตารางที่ 4.2.5 จะเป็นการเปรียบเทียบผลทดสอบการรู้จำของเสียง โดยใช้แบบอ้างอิงจำนวน 18 แบบของคำในกลุ่มเสียงที่ 2

ตารางที่ 4.2.5 แสดงผลการรู้จำของเสียงกลุ่มที่ 2 ต่อแบบอ้างอิงที่สร้างจากผู้พูดจำนวน 20 คน

กลุ่มทดสอบ	ความถูกต้อง (%)
A1	80.00
A2,A3	71.11
B	65.00

จากผลการทดสอบในตารางที่ 4.2.5 เห็นได้ว่า เมื่อมีจำนวนแบบอ้างอิงมีมากขึ้น เปอร์เซ็นต์ความถูกต้องในการรู้จำจะลดลง สำหรับในตารางที่ 4.2.6-4.2.8 จะเป็นการแจกแจงผลของการรู้จำของเสียงพูดในกลุ่ม A1,A2-A3, และ B ตามลำดับ ที่ใช้แบบอ้างอิงจำนวน 18 แบบ โดยสร้างจากกลุ่มผู้พูดในกลุ่ม A1 จำนวน 20 คน ซึ่งจะพบว่า เสียงพูดแต่ละคำที่เกิดความผิดพลาด สามารถจัดแยกเป็นกลุ่ม ๆ ได้ดังนี้

- ก. กลุ่มที่ 1 เสียง "ศูนย์", "หนึ่ง", "สี่", "ยี่" และ "หมื่น"
- ข. กลุ่มที่ 2 เสียง "สอง" และ "ร้อย"
- ค. กลุ่มที่ 3 เสียง "สาม", "เก้า" และ "ล้าน"
- ง. กลุ่มที่ 4 เสียง "ห้า" และ "ล้าน"
- จ. กลุ่มที่ 5 เสียง "หก", "เจ็ด", "เอ็ด" และ "สิบ"
- ฉ. กลุ่มที่ 6 เสียง "แปด" และ "แสน"

ซึ่งทั้ง 6 กลุ่มนี้จะมีรูปแบบอ้างอิงของเสียงที่คล้ายคลึงกันดังแสดงในภาคผนวก ข โดยที่เสียงตัวเลขในกลุ่มที่ 5 จะมีความแตกต่างจากคำในกลุ่มอื่นตรงที่จะเป็นคำที่สั้นกว่าคำในกลุ่มอื่นทั้งหมด ความผิดพลาดส่วนใหญ่เกิดจากรูปแบบอ้างอิงที่คล้ายคลึงกัน ยกเว้นคำว่า "สิบ" ความผิดพลาดจะไม่ได้เกิดจากการที่มีแบบอ้างอิงที่คล้ายคลึงกัน เพราะ "สิบ" จะผิดพลาดเป็น "ยี่" และ "หนึ่ง" เป็นส่วนมาก ทั้งนี้เป็นผลของทำไดนามิกโถมวาร์ปิงของคำสองคำที่มีความยาวของเสียงแตกต่างกันมาก โดยในการหาไดนามิกโถมวาร์ปิงนี้จะถูกกำหนดข้อมูลที่มีจำนวนน้อยกว่า ในที่นี้ทำให้ข้อมูลเสียงที่เป็นแบบอ้างอิงที่ยาวกว่าถูกตัดออกไป จึงทำให้เกิดความผิดพลาดนี้ขึ้น



ตารางที่ 4.2.6 แสดงผลการรู้จำของกลุ่มผู้พูด A1 ต่อแบบอ้างอิงที่สร้างจากผู้จำนวน 20 คน โดยการทดสอบเสียงพูดในกลุ่มที่ 2

แบบทดสอบ A1	แบบอ้างอิง A1																	
	0	1	2	3	4	5	6	7	8	9	10	เอ็ด	ยี่	ร้อย	พัน	หมื่น	แสน	ล้าน
0	16				1												3	
1		17									2						1	
2			19											1				
3				13						1				2	1			3
4	2				17											1		
5						15										2	1	2
6		1					17					1				1		
7		2						13				4					1	
8				1					17								2	
9				2						17			1					
10		3									15		2					
เอ็ด							1	1			1	17						
ยี่	1				4						1		13		1			
ร้อย			3										1		15	1		
พัน									1						1	17		1
หมื่น	2	2														16		
แสน				1											1		18	
ล้าน				1		1				1						1		16

ตารางที่ 4.2.7 แสดงผลการรู้จำของกลุ่มผู้พูด A2,A3 ต่อแบบอ้างอิงที่สร้างจากผู้จำนวน 20 คน โดยทดสอบเสียงกลุ่มที่ 2

แบบทดสอบ A2,A3	แบบอ้างอิง A1																	
	0	1	2	3	4	5	6	7	8	9	10	เอ็ด	ยี่	ร้อย	พัน	หมื่น	แสน	ล้าน
0	29	2			1						1				1	6		
1		31						1			3	1	1			3		
2			30											10				
3				23		1				1				3	2			10
4	7				28		1									4		
5						28				1				2	2		3	4
6		2					30					6				2		
7		1					2	26			4	6		1				
8				2		1			30								7	
9				4					1	32				2				1
10		6						1			24		7			2		
เอ็ด		1					4	2			4	29						
ยี่	2	1			7								25			5		
ร้อย			4	1										31	3			1
พัน				1		1							1	1	35			1
หมื่น	2	10											2			26		
แสน			1	1			1		1					5			30	1
ล้าน				6		2						1		4	2			25

ตารางที่ 4.2.8 แสดงผลความถูกต้องของกลุ่มผู้พูด B ต่อแบบอ้างอิงที่สร้างจากผู้พูดจำนวน 20 คน โดยทดสอบเสียงกลุ่มที่ 2

แบบทดสอบ B	แบบอ้างอิง A1																	
	0	1	2	3	4	5	6	7	8	9	10	ด	ย	ร	พ	ห	แ	ล
0	26	1			8											5		
1		25					3						3	1		8		
2			30				1	2						7				
3				14						12			1	1	2			10
4					31								6		1	2		
5				2		32						1			1			4
6		3					29					5		1		2		
7		2					2	23		4	7	1				1		
8				1					26	1			2		1		5	4
9				9		1				24			1	2			1	2