

## บทที่ 4 ผลการวิจัย

ในบทนี้จะกล่าวถึงผลการรู้จำเสียงพูดตัวเลข และคำพยางค์เดียว โดยจำแนกเป็นชุดเสียงพูดเพื่อฝึกฝน (Training Set) และชุดเสียงพูดเพื่อทดสอบ (Testing Set) และทำการเปรียบเทียบเมื่อเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดน มาร์คอฟ และจำนวนชุดรหัสอ้างอิง

### 4.1 ผลการรู้จำคำพูดภาษาไทย

การรู้จำคำพูดภาษาไทย โดยอาศัยแบบจำลองฮิดเดน มาร์คอฟนี้ แบ่งชุดพารามิเตอร์แบบจำลองฮิดเดน มาร์คอฟ เป็น 2 ชุด ได้แก่ ชุดพารามิเตอร์ของคำศัพท์ตัวเลขศูนย์ถึงเก้า กับชุดพารามิเตอร์คำศัพท์พยางค์เดียวจำนวน 20 คำ ในการรู้จำจะเปรียบเทียบเสียงพูดทดสอบกับชุดรูปร้องต้นแบบอ้างอิง โดยเสียงพูดตัวเลขศูนย์ถึงเก้าจะทดสอบกับชุดรูปร้องต้นแบบอ้างอิงศูนย์ถึงเก้า ในทำนองเดียวกัน เสียงพูดคำศัพท์พยางค์เดียวจะทดสอบกับชุดรูปร้องต้นแบบอ้างอิงคำศัพท์พยางค์เดียว เมื่อทำการเปรียบเทียบทั้งชุดจะเรียงลำดับความน่าจะเป็นจากมากไปหาน้อย คำศัพท์ที่รู้จำได้คือคำศัพท์ที่มีค่าความน่าจะเป็นสูงสุด

อัตราความแม่นยำในการรู้จำ (Recognition Accuracy) หรือ อัตราการรู้จำ (Recognition Rate) คือ อัตราความถูกต้องในการรู้จำคำศัพท์แต่ละคำ โดยคิดจากร้อยละของอัตราส่วนจำนวนคำศัพท์ที่รู้จำถูกต้องกับจำนวนคำศัพท์ทั้งหมด

ผลการรู้จำคำพูดภาษาไทยของชุดคำศัพท์เลขศูนย์ถึงเก้า และชุดคำศัพท์พยางค์เดียว โดยจำแนกตามชุดเสียงพูดที่นำมาทดสอบได้แก่ ชุดเสียงพูดเพื่อฝึกฝน (Training Set) และชุดเสียงพูดเพื่อทดสอบ (Testing Set) ดังตารางที่ 4.1

ตารางที่ 4.1 ผลอัตราการรู้จำ

ชุดคำศัพท์	อัตราการรู้จำคำพูด (ร้อยละ)	
	ชุดฝึกฝน	ชุดทดสอบ
ชุดคำศัพท์ตัวเลขศูนย์ถึงเก้า	100.000	93.00
ชุดคำศัพท์พยางค์เดียว	100.000	88.50
อัตราการรู้จำเฉลี่ย	100.000	90.75

ผลการรู้จำคำพูดภาษาไทยในงานวิจัยนี้แบ่งออกเป็น 4 วิธี คือ วิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ วิธีที่ 2 ใช้ LSF โดยประมาณเป็นลักษณะสำคัญ วิธีที่ 3 ใช้ LSF โดยตรงและ

ผ่านตัวกรองมัธยฐาน ในแต่ละวิธีแบ่งออกเป็น 5 แบบพารามิเตอร์ และวิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง แบ่งออกเป็น 4 แบบพารามิเตอร์ ในแต่ละแบบแบ่งเป็น 2 ชุด คือ ชุดเสียงพูดเพื่อฝึกฝน และ ชุดเสียงพูดเพื่อทดสอบ และใช้จำนวนเสียงพูดทดสอบ 20 คน และในแต่ละแบบพารามิเตอร์จะใช้เวลาคำนวณเพื่อการรู้จำโดยเฉลี่ยต่อหนึ่งเสียง เมื่อใช้เครื่องคอมพิวเตอร์ AMD K7-700 หน่วยความจำ 128 MB Harddisk ขนาด 20 GB

ตารางที่ 4.2 รายละเอียดค่าพารามิเตอร์สำหรับการรู้จำวิธีที่ 1 ถึง วิธีที่ 3 และเวลาคำนวณเพื่อรู้จำต่อหนึ่งเสียง

ชุดที่	จำนวนเสียงพูดเพื่อฝึกฝน	ขนาดชุดรหัส	จำนวนสถานะของแบบจำลอง	เวลาคำนวณเพื่อรู้จำต่อหนึ่งเสียง (วินาที)
1	20	128	3	0.0343
2	20	256	3	0.0396
3	20	256	5	0.0409
4	20	256	15	0.0459
5	40	256	15	0.0459

ในตารางที่ 4.2 เป็นรายละเอียดค่าพารามิเตอร์สำหรับการรู้จำแต่ละชุด เพื่อใช้ในการเปรียบเทียบอัตราการรู้จำ เมื่อเปลี่ยนพารามิเตอร์ให้มีค่าต่างกัน

ตารางที่ 4.3 รายละเอียดค่าพารามิเตอร์สำหรับการรู้จำวิธีที่ 4 และเวลาคำนวณเพื่อรู้จำต่อหนึ่งเสียง

ชุดที่	จำนวนเสียงพูดเพื่อฝึกฝน	ขนาดชุดรหัส	จำนวนสถานะของแบบจำลอง	เวลาคำนวณเพื่อรู้จำต่อหนึ่งเสียง (วินาที)
1	20	128	3	0.0175
2	20	128	5	0.0187
3	20	128	15	0.0237
4	40	128	15	0.0237

ตารางที่ 4.4 ผลการรู้จำจากวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ

ชุดที่	รายละเอียดพารามิเตอร์	ชุดเสียงพูด	ชุดคำศัพท์ตัวเลข	ชุดคำศัพท์พยางค์เดียว	ค่าเฉลี่ย
1	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	99.00	98.00	98.50
		ชุดเสียงพูดทดสอบ	88.00	72.25	80.13
2	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	99.50	99.75
		ชุดเสียงพูดทดสอบ	85.00	76.75	80.88
3	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 5 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	83.00	77.25	80.13
4	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	85.50	81.50	83.50
5	จำนวนเสียงพูดทดสอบ 40 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	93.00	88.50	90.75

ตารางที่ 4.5 ผลการรู้จำจากวิธีที่ 2 ใช้ LSF โดยประมาณเป็นลักษณะสำคัญ

ชุดที่	รายละเอียดพารามิเตอร์	ชุดเสียงพูด	ชุดคำศัพท์ตัวเลข	ชุดคำศัพท์พยางค์เดียว	ค่าเฉลี่ย
1	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	97.50	98.25	97.88
		ชุดเสียงพูดทดสอบ	77.50	68.50	73.00
2	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	99.00	98.75	98.88
		ชุดเสียงพูดทดสอบ	82.00	63.50	72.75
3	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 5 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	99.50	100.00	99.75
		ชุดเสียงพูดทดสอบ	81.00	72.812	72.38
4	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	79.50	70.00	74.75
5	จำนวนเสียงพูดทดสอบ 40 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	99.75	99.88
		ชุดเสียงพูดทดสอบ	89.50	83.25	86.38

ตารางที่ 4.6 ผลการรู้จำจากวิธีที่ 3 ใช้ LSF โดยตรงผ่านตัวกรองมัธยฐานเป็นลักษณะสำคัญ

ชุดที่	รายละเอียดพารามิเตอร์	ชุดเสียงพูด	ชุดคำศัพท์ตัวเลข	ชุดคำศัพท์พยางค์เดียว	ค่าเฉลี่ย
1	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	98.50	97.25	97.88
		ชุดเสียงพูดทดสอบ	85.50	71.75	78.63
2	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	99.50	99.75
		ชุดเสียงพูดทดสอบ	85.00	71.50	78.25
3	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 5 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	99.75	99.88
		ชุดเสียงพูดทดสอบ	86.00	72.25	79.13
4	จำนวนเสียงพูดทดสอบ 20 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	86.00	72.25	79.13
5	จำนวนเสียงพูดทดสอบ 40 คน 256 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	99.75	99.88
		ชุดเสียงพูดทดสอบ	93.00	83.00	88.00

ตารางที่ 4.7 ผลการรู้จำจากวิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง

ชุดที่	รายละเอียดพารามิเตอร์	ชุดเสียงพูด	ชุดคำศัพท์ตัวเลข	ชุดคำศัพท์พยางค์เดียว	ค่าเฉลี่ย
1	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 3 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	96.50	97.50	97.00
		ชุดเสียงพูดทดสอบ	74.00	58.00	66.00
2	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 5 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	99.50	99.25	99.38
		ชุดเสียงพูดทดสอบ	73.50	61.75	67.63
3	จำนวนเสียงพูดทดสอบ 20 คน 128 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	80.50	64.50	72.50
4	จำนวนเสียงพูดทดสอบ 40 คน 128 ชุดรหัส 15 สถานะ	ชุดเสียงพูดเพื่อฝึกฝน	100.00	100.00	100.00
		ชุดเสียงพูดทดสอบ	89.00	75.75	82.38

## 4.2 วิเคราะห์การรู้จำคำพูดภาษาไทย

จากรายละเอียดผลการรู้จำในตารางที่ 4.4 ถึงตารางที่ 4.7 แบ่งออกเป็น 4 วิธี คือ วิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ วิธีที่ 2 ใช้ LSF โดยประมาณเป็นลักษณะสำคัญ วิธีที่ 3 ใช้ LSF โดยตรงและผ่านตัวกรองมัธยฐาน และวิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง สามารถวิเคราะห์ผลได้ดังนี้

1. ผลการรู้จำจากวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ วิธีที่ 2 LSF โดยประมาณเป็นลักษณะสำคัญ วิธีที่ 3 ใช้ LSF โดยตรงผ่านตัวกรองมัธยฐานเป็นลักษณะสำคัญ วิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง แสดงผลดังตารางที่ 4.4 ถึง 4.7 ตามลำดับ โดยแจกแจงพารามิเตอร์ทั้งหมด 5 ชุด และ 4 ชุด ดังตารางที่ 4.2 และ 4.3 ตามลำดับ

2. ค่าเฉลี่ยสูงสุดของอัตราการรู้จำ คือ ค่าเฉลี่ยของอัตราการรู้จำในวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ ชุดเสียงพูดฝึกฝนมีค่าร้อยละ 100.00 และชุดเสียงพูดทดสอบมีค่าร้อยละ 90.75 ซึ่งสูงกว่าวัตถุประสงค์ของงานวิจัยที่ประเมินไว้ที่ร้อยละ 85.00 และค่าเฉลี่ยของอัตราการรู้จำในวิธีที่ 2 ถึงวิธีที่ 4 มีค่ามากกว่าวัตถุประสงค์ของงานวิจัยด้วย เป็นที่น่าสังเกตว่าวิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง เป็นวิธีที่มีการคำนวณน้อยมาก กล่าวคือ ใช้เพียงบิตข้อมูลของมาตรฐาน G.729 เป็นข้อมูลขาเข้าในแบบจำลองฮิดเดน มาร์คอฟ มีอัตราการรู้จำเฉลี่ยร้อยละ 82.38

3. ในขั้นตอนของการรู้จำ เลือกใช้พารามิเตอร์ที่เกี่ยวข้องกับคุณลักษณะสำคัญเชิงความถี่เท่านั้น ในส่วนของพลังงานและ คาบการสั่นของเสียงจะใช้ในส่วนของ การหาจุดสิ้นสุดของเสียง ซึ่งในงานวิจัยนี้เป็นเสียงตัวเลข และเสียงคำพยางค์เดียว ทำให้การหาจุดสิ้นสุดของเสียงทำได้โดยง่าย และมีความแม่นยำสูง

4. จากการเปรียบเทียบผล เมื่อเปลี่ยนแปลงขนาดของชุดรหัสจาก 128 ชุดรหัส กับ 256 ชุดรหัส ได้อัตรารู้จำเฉลี่ยเพิ่มขึ้น เป็นผลมาจากความละเอียดของชุดรหัสที่จะใช้จำแนกความแตกต่างของเสียงพูดที่ใช้ในการรู้จำ

5. จากการเปรียบเทียบผล เมื่อเพื่อจำนวนสถานะของแบบจำลองฮิดเดน มาร์คอฟ จาก 3 สถานะ เป็น 5 สถานะ และ 15 สถานะ ตามลำดับให้ค่าอัตราการรู้จำใกล้เคียงกัน เพราะความยาวของชุดคำศัพท์มีค่าไม่มากนัก ทำให้ไม่จำเป็นต้องใช้จำนวนสถานะของแบบจำลองฮิดเดน มาร์คอฟสูง

6. จากการเปรียบเทียบจำนวนเสียงพูดเพื่อฝึกฝน เมื่อเพิ่มจำนวนเสียงพูดเพื่อฝึกฝน จาก 20 คน เป็น 40 คน จะให้ผลการรู้จำมีประสิทธิภาพเพิ่มขึ้นอย่างมาก ในแต่ละวิธีที่ใช้ในการรู้จำจะให้อัตราการรู้จำเพิ่มขึ้นอีกประมาณร้อยละ 10 ดังนั้นถ้ามีจำนวนเสียงพูดเพื่อฝึกฝนจำนวนมากกว่า 40 คนแล้วอาจให้อัตราการรู้จำสูงกว่านี้อีก

7. จากผลของการเปลี่ยนจำนวนชุดรหัส การเปลี่ยนจำนวนสถานะของแบบจำลอง ฮิดเดน มาร์คอฟ และการเปลี่ยนจำนวนเสียงพูดเพื่อฝึกฝน การเพิ่มจำนวนเสียงพูดเพื่อฝึกฝนมีผลต่ออัตราการรู้จำมากที่สุด

8. ในการรู้จำจากวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ ให้ค่าการรู้จำในชุดเสียงพูดเพื่อฝึกฝนทั้งชุดคำศัพท์ตัวเลข และชุดคำศัพท์พยางค์เดียวสูงมาก (จากตารางที่ 4.4) ในส่วนของชุดเสียงพูดทดสอบให้ผลการรู้จำประมาณร้อยละ 83 ถึง 96 (จากตารางที่ 4.4) ซึ่งเป็นวิธีที่ให้ค่าเฉลี่ยของอัตราการรู้จำสูงสุด

9. ในการรู้จำจากวิธีที่ 2 ใช้ LSF โดยประมาณเป็นลักษณะสำคัญ เป็นการทดสอบว่า ถ้าลดการคำนวณในส่วนของหารลักษณะสำคัญ โดยละการคำนวณบิต L2 และ L3 ของการถอดรหัส G.729 จะมีผลมากน้อยเพียงใด จากตารางที่ 4.5 จะให้อัตราการรู้จำลดลงประมาณร้อยละ 3.5 ถึงร้อยละ 10 ทำให้การประมาณ LSF โดยคิดเฉพาะบิต L0 กับ L1 มีความผิดพลาดเพิ่มขึ้นมาก ทำให้การลดการคำนวณไม่คุ้มกับความผิดพลาดที่เพิ่มขึ้นมาก

10. ในการรู้จำจากวิธีที่ 3 ใช้ LSF โดยตรงและผ่านตัวกรองมัธยฐาน เป็นการเพิ่มการคำนวณในส่วนของตัวกรองมัธยฐานเพื่อดูแนวโน้มของลักษณะการเปลี่ยนแปลงของเส้น LSF จากตารางที่ 4.6 ให้การรู้จำใกล้เคียงกับวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ ดังนั้นการผ่านตัวกรองมัธยฐานไม่ช่วยให้อัตราการรู้จำเพิ่มอย่างมีนัยสำคัญ จึงไม่จำเป็นต้องทำการคำนวณในส่วนของ การผ่านตัวกรองมัธยฐาน

11. ในการรู้จำจากวิธีที่ 4 ใช้ดัชนีของชุดรหัส G.729 โดยตรง เป็นวิธีที่มีการคำนวณน้อยที่สุด เพราะไม่ต้องคำนวณหารลักษณะสำคัญ แล้วทำการควอนไทซ์แบบเวกเตอร์เพื่อเป็นข้อมูลขาเข้าของแบบจำลองฮิดเดน มาร์คอฟ แต่ใช้บิตของมาตรฐาน G.729 ได้โดยตรง ทำให้การคำนวณลดลงประมาณเท่าตัว แม้ว่าให้อัตราการรู้จำน้อยกว่าวิธีอื่น ๆ แต่ยังมีอัตราการรู้จำสูงสามารถนำไปประยุกต์ใช้งานได้

12. คำศัพท์ที่มีการรู้จำผิดพลาดมากที่สุด คือ คำว่า "ปาก" กับ "ตา" ซึ่งทั้งสองคำนี้ จะมีการรู้จำผิดพลาดซึ่งกันและกัน อาจเนื่องจาก ทั้งสองคำ มีการออกเสียงคล้ายกัน และเป็นคำศัพท์ที่สั้น ทำให้ระบบรู้จำไม่สามารถรู้จำได้ว่าเสียงนั้นเป็นเสียงของคำใดแน่นอน