

โปรแกรมรู้จำรู้จกเสียงอักขระในภาษาอังกฤษ

ALPHABET'S VOICE RECOGNITION PROGRAM



นายปูลณรัตน์	ทองธรรม	รหัส 44362663
นายวินัย	เรื่อนเพ็ง	รหัส 44362754
นายศรัณย์	ตันสมบูรณ์	รหัส 44362770

ห้องสมุดคณะวิศวกรรมศาสตร์
วันที่รับ..... 2,9 ส.ย. 2548
เลขทะเบียน..... 4800043
เลขเรียกหนังสือ.....
มหาวิทยาลัยนเรศวร

1506692x C.L
15
16771
2547

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร
ปีการศึกษา 2547

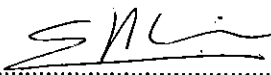


ใบรับรองโครงการวิศวกรรม

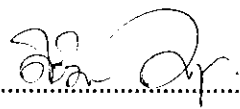
หัวข้อโครงการ	โปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ		
ผู้ดำเนินโครงการ	นาย ปุณณรัตน์ ทองธรรม	รหัส	44362663
	นาย วินัย เรือนเพ็ง	รหัส	44362754
	นาย ศรัณย์ ตันสมบุรณ์	รหัส	44362770
อาจารย์ที่ปรึกษา	อ.พงศ์พันธ์ กิจสนาโยธิน		
สาขาวิชา	วิศวกรรมคอมพิวเตอร์		
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์		
ปีการศึกษา	2547		

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนครสวรรค์ อนุมัติให้โครงการฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะกรรมการสอบโครงการวิศวกรรม


.....ประธานกรรมการ
(อาจารย์พงศ์พันธ์ กิจสนาโยธิน)


.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.สุชาติ-แย้มเม่น)


.....กรรมการ
(อาจารย์ศิริพร เดชะศีลารักษ์)

หัวข้อโครงการ	โปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ		
ผู้ดำเนินโครงการ	นายปุลณรัตน์ ทองธรรม	รหัส	44362663
	นายวินัย เรือนเพ็ง	รหัส	44362754
	นายศรัณย์ ตันเสมอบูรณ์	รหัส	44362770
อาจารย์ที่ปรึกษา	อ.พงศ์พันธ์ กิจสนาโยธิน		
สาขาวิชา	วิศวกรรมคอมพิวเตอร์		
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์		
ปีการศึกษา	2547		

บทคัดย่อ

เป็นความจริงที่ว่าคอมพิวเตอร์ไม่สามารถจดจำเสียงได้ทุกเสียง เสียงแต่ละเสียงมีเอกลักษณ์ประจำเสียง การที่จะทำให้คอมพิวเตอร์จดจำเสียงได้จะต้องนำฟังก์ชันถ่ายโอนและแบบจำลองฮิดเดน มาร์คอฟ มาใช้ในการวิเคราะห์ โดยการนำเสียงที่ต้องการหาผลลัพธ์มาทำการวิเคราะห์และเปรียบเทียบกับฐานข้อมูลเสียง เพื่อหาว่ารูปแบบของเสียงใดในฐานข้อมูลเสียงที่มีความใกล้เคียงกับเสียงที่ต้องการหาผลลัพธ์มากที่สุด แล้วนำเสียงที่ใกล้เคียงที่สุดมาเป็นผลลัพธ์ การวิเคราะห์ด้วยวิธีดังกล่าวทำให้ได้ความถูกต้องของผลลัพธ์มากกว่าร้อยละ 70 เสียงที่มักเกิดความผิดพลาดในการแสดงผลคือเสียงที่มีการออกเสียงคล้ายกัน เช่น b-d-c-e-v, a-j-k, f-h-s-x, m-n, p-t

Project Title Alphabet's Voice Recognition Program
Name Mr. Punnarat Tongtham ID. 44362663
Mr. Winai Ruanpheng ID. 44362754
~~Mr. Saran Tansomboon ID. 44362770~~

Project Advisor Mr. Phongphun Kijsanayothin
Major Computer Engineering

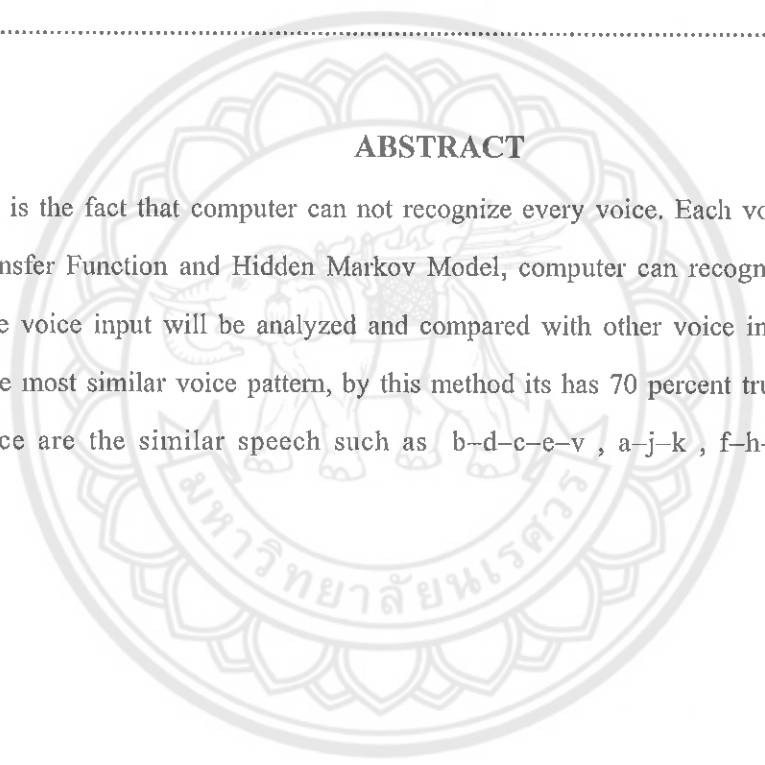
Department Electrical and Computer Engineering

Academic Year 2004

.....

ABSTRACT

It is the fact that computer can not recognize every voice. Each voice has its identify, using Transfer Function and Hidden Markov Model, computer can recognize more pattern of voice. The voice input will be analyzed and compared with other voice in voice database for finding the most similar voice pattern, by this method its has 70 percent true output. The most error voice are the similar speech such as b-d-c-e-v , a-j-k , f-h-s-x , m-n , p-t .



กิตติกรรมประกาศ

โครงการวิศวกรรม โปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ จะไม่สามารถสำเร็จได้ หากไม่ได้รับความช่วยเหลือจากบุคคลจากหลาย ๆ ฝ่าย ซึ่งประกอบไปด้วยอาจารย์ พงศ์พันธ์ กิจสนาโยธิน อาจารย์ที่ปรึกษา ซึ่งคอยให้คำปรึกษาที่ดีเพื่อการพัฒนาโปรแกรมให้มีประสิทธิภาพที่ดีขึ้น นอกจากนี้ยังมีกำลังใจและคำปรึกษาที่ได้รับจากครอบครัวและเพื่อน ๆ วิศวกรรมคอมพิวเตอร์ ซึ่งสิ่งต่าง ๆ นี้คอยเป็นแรงผลักดันให้คณะผู้จัดทำมีกำลังใจในการทำให้งานผ่านไปได้ด้วยดี

สุดท้ายนี้ต้องขอขอบคุณต่อผู้ที่มีส่วนช่วยเหลือคณะผู้จัดทำในทุกๆด้านซึ่งรวมไปถึงบุคคลที่ไม่ได้กล่าวถึงมา ณ ที่นี้ด้วย



นาย ปุณณรัตน์ ทองธรรม
นาย วินัย เรือนเพ็ง
นาย ศรัณย์ ตันสมบุญ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่ 1 บทนำ	
1.1 ที่มาและความสำคัญ.....	1
1.2 วัตถุประสงค์ของ โครงการ.....	1
1.3 ขอบเขตของ โครงการ.....	2
1.4 ขั้นตอนการดำเนินงาน.....	2
1.5 ผลที่คาดว่าจะได้รับ.....	2
1.6 งบประมาณที่ใช้.....	3
บทที่ 2 หลักการและทฤษฎี	
2.1 ความสามารถในการได้ยินของมนุษย์.....	4
2.2 คลื่นเสียง.....	6
2.2.1 คุณสมบัติทางกายภาพของคลื่นเสียง.....	7
2.2.2 ความถี่.....	7
2.2.3 ความยาวคลื่น.....	8
2.3 ทฤษฎีเวฟเล็ท.....	8
2.3.1 ฟิวรีร์ทรานสฟอร์ม.....	9
2.3.2 ซอร์ทไทม์ฟิวรีร์ทรานสฟอร์ม.....	11

สารบัญ(ต่อ)

	หน้า
2.3.3 เวฟเล็ตทรานสฟอร์ม.....	16
2.3.3.1 ความหมายของเวฟเล็ตควาวิซัล(เวฟเล็ต *) และเวฟเล็ตแม่.....	17
2.3.3.2 ความหมายของเวฟเล็ตทรานสฟอร์มและการวิเคราะห์แบบ หลายความละเอียด.....	18
2.3.3.3 คอนทินิวัสเวฟเล็ตทรานสฟอร์ม.....	19
2.3.3.3.1 การคำนวณคอนทินิวัสเวฟเล็ตทรานสฟอร์ม.....	22
2.3.3.3.2 การแปลงคอนทินิวัสเวฟเล็ตทรานสฟอร์มกลับ เป็นสัญญาณเดิม.....	27
2.3.3.3.3 คอนทินิวัสเวฟเล็ตทรานสฟอร์มแบบดิสครีต.....	28
2.3.3.4 ดิสครีตเวฟเล็ตทรานสฟอร์ม.....	31
2.3.3.4.1 การคำนวณดิสครีตเวฟเล็ตทรานสฟอร์ม.....	31
2.3.3.4.2 การแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็น สัญญาณเดิม.....	35
2.3.3.4.3 การวิเคราะห์เวฟเล็ตแพกเกจ.....	39
2.3.3.4.4 การนำดิสครีตเวฟเล็ตทรานสฟอร์มไปใช้งาน.....	39
2.3.3.5 ความแตกต่างและการเลือกใช้คอนทินิวัสเวฟเล็ตทรานสฟอร์ม และดิสครีตเวฟเล็ต ทรานสฟอร์ม.....	39
2.4 ทฤษฎีพื้นฐานของแบบจำลองฮิดเดน มาร์คอฟ.....	41
2.4.1 กระบวนการมาร์คอฟ.....	42
2.4.2 แบบจำลองฮิดเดน มาร์คอฟ.....	44
2.4.3 ส่วนประกอบของ แบบจำลองฮิดเดน มาร์คอฟ.....	45
2.4.4 ปัญหาขั้นพื้นฐานในแบบจำลองฮิดเดน มาร์คอฟ.....	46
2.4.5 การแก้ปัญหาพื้นฐานในแบบจำลองฮิดเดน มาร์คอฟ.....	47
2.4.6 สรุปแบบจำลองฮิดเดน มาร์คอฟ.....	52

สารบัญ(ต่อ)

	หน้า
บทที่ 3 วิธีดำเนินการ	
3.1 การคัดแยกสัญญาณ.....	54
3.2 การวิเคราะห์สัญญาณด้วยเวฟเล็ททรานสฟอร์ม.....	55
3.3 การตัดสัญญาณรบกวน.....	56
3.4 การแบ่งสัญญาณออกเป็นส่วนเล็ก ๆ.....	57
3.5 การสร้างแบบจำลองฮิดเดน มาร์คอฟ.....	59
3.5.1 การหาความน่าจะเป็นของการเป็นสถานะที่จะเป็นสถานะเริ่มต้น.....	59
3.5.2 การหาค่าความน่าจะเป็นของการเปลี่ยนแปลงสถานะ.....	61
3.5.3 การหาค่าความน่าจะเป็นของการเปลี่ยนจากสถานะก่อนหน้ามายังสถานะปัจจุบัน.....	66
3.6 การตัดสินใจ.....	74
บทที่ 4 การทดลองและวิเคราะห์ผลการทดลอง	
4.1 ข้อมูลที่ใช้ในการทดสอบ.....	75
4.2 ผลการทดลอง.....	75
บทที่ 5 บทสรุป	
5.1 สรุปผลการทดลอง.....	80
5.2 ข้อเสนอแนะ.....	81
เอกสารอ้างอิง.....	82
ภาคผนวก	
ภาคผนวก ก ความต้องการของระบบ.....	84
ภาคผนวก ข การติดตั้งโปรแกรม MatLAB 6.1 Release 12.....	85
ภาคผนวก ค การติดตั้งโปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ.....	91
ภาคผนวก ง คู่มือการใช้โปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ.....	94
ประวัติผู้จัดทำโครงการ.....	97

สารบัญตาราง

ตารางที่		หน้า
2.1	เปรียบเทียบฟูเรียร์ทรานสฟอร์ม ซอร์ทไทม์ฟูเรียร์ทรานสฟอร์ม และเวฟเล็ตทรานสฟอร์ม.....	18
3.1	ตารางความน่าจะเป็นของ b_j pine.....	60
3.2	ตารางนับจำนวนของ b_j เฟรมที่ 2.....	62
3.3	ตารางความน่าจะเป็นของ b_j เฟรมที่ 2.....	63
3.4	ตารางนับจำนวนของ b_j เฟรมที่ 3.....	64
3.5	ตารางความน่าจะเป็นของ b_j เฟรมที่ 3.....	65
3.6	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 1.....	67
3.7	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 2.....	67
3.8	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 3.....	68
3.9	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 4.....	68
3.10	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 5.....	69
3.11	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 6.....	69
3.12	ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 7.....	70
3.13	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 1.....	71
3.14	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 2.....	71
3.15	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 3.....	72
3.16	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 4.....	72
3.17	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 5.....	73
3.18	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 6.....	73
3.19	ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 7.....	74
4.1	ผลการทดลองของผู้ใช้คนที่หนึ่ง(เสียงอักษร).....	76
4.2	ผลการทดลองของผู้ใช้คนที่สอง(เสียงอักษร).....	77
4.3	ผลการทดลองของผู้ใช้คนที่สาม(เสียงอักษร).....	78

สารบัญรูป

รูปที่		หน้า
2.1	ความสามารถในการได้ยินเสียงของหูมนุษย์ระดับความดังเสียง (0-130 dB) และช่วงความถี่ (20-20,000 Hz).....	5
2.2	เส้นเท่าระดับเสียง และเส้นถ่วงน้ำหนักแบบ A.....	6
2.3	คลื่นเสียง.....	6
2.4	คลื่นรูปไซน์.....	7
2.5	ความยาวคลื่นกับความถี่.....	8
2.6	สัญญาณสเตรชันนารี ความถี่ 100 , 50 , 25 และ 10 Hz.....	9
2.7	ฟูรีเยร์ทรานสฟอร์ม์ของสัญญาณในรูปที่ 2.6.....	10
2.8	สัญญาณนอนสเตรชันนารี ความถี่ 100 , 50 , 25 , 10 Hz.....	10
2.9	ฟูรีเยร์ทรานสฟอร์ม์ของสัญญาณในรูปที่ 2.8.....	11
2.10	สัญญาณที่มีความถี่ 300 , 200 , 100 และ 50 Hz.....	13
2.11	ความยาวของหน้าต่างขนาด $a = 0.01, 0.001, 0.0001$ และ 0.00001	13
2.12	ความยาวของหน้าต่างขนาด $a = 0.01, 0.001, 0.0001$ และ 0.00001 แบบสามมิติ....	14
2.13	ซอร์ทใหม่ฟูรีเยร์ทรานสฟอร์ม์ของสัญญาณในรูปที่ 2.10 เมื่อความยาวของหน้าต่างขนาด $a = 0.001$	14
2.14	ซอร์ทใหม่ฟูรีเยร์ทรานสฟอร์ม์ของสัญญาณในรูปที่ 2.10 เมื่อความยาวของหน้าต่างขนาด $a = 0.0001$	15
2.15	ซอร์ทใหม่ฟูรีเยร์ทรานสฟอร์ม์ของสัญญาณในรูปที่ 2.10 เมื่อความยาวของหน้าต่างขนาด $a = 0.00001$	15
2.16	สัญญาณที่มีความถี่สูงเป็นช่วงสั้น ๆ.....	17
2.17	ตัวอย่างของเวฟเล็ท โคอิเฟิลต์ส์.....	17
2.18	ตัวอย่างของเวฟเล็ทคาวบิชี่ส์.....	18
2.19	แสดงการเพิ่มและลดสเกลของสัญญาณ โคไซน์เดียวกัน.....	21
2.20	แสดงสัญญาณและเวฟเล็ท $s = 1$ ที่ τ ที่ = 2, 40, 90, 140.....	23
2.21	แสดงสัญญาณและเวฟเล็ท $s = 5$ ที่ $\tau = 20, 60, 110, 140$	23
2.22	แสดงสัญญาณและเวฟเล็ท $s = 20$ ที่ $\tau = 50, 90, 110, 140$	24
2.23	สัญญาณนอนสเตรชันนารีความถี่ 30 , 20 , 10 , 5 Hz.....	25

สารบัญรูป(ต่อ)

รูปที่		หน้า
2.24	แสดงคอนทินิวอัสเวฟเล็ททรานสฟอร์มในมุมมองต่าง ๆ ในระนาบทรานสชัน-สเกล ของสัญญาณในรูป 2.22.....	25
2.25	แสดงคอนทินิวอัสเวฟเล็ททรานสฟอร์ม ในมุมมองต่าง ๆ ในระนาบทรานสเลชัน-สเกล ของสัญญาณในรูป 2.22.....	26
2.26	ความละเอียดของเวฟเล็ททรานสฟอร์ม.....	27
2.27	ตารางไดนามิก.....	29
2.28	แสดงการแปลงสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์มใน 1 ชั้น.....	32
2.29	แสดงการแปลงสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์มซ้ำหลายชั้น.....	33
2.30	แสดงสัญญาณและดิสครีตเวฟเล็ททรานสฟอร์ม โคเอฟฟิเชียนของสัญญาณ.....	34
2.31	แสดงฟิลเตอร์ทั้ง 4 ตัว ในการทำดิสครีตเวฟเล็ททรานสฟอร์ม และ การแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม.....	36
2.32	แสดงดิสครีตเวฟเล็ททรานสฟอร์ม และการเพิ่มจำนวนจุดตัวอย่าง โดยการแทรกจุด 0 ลงในดิสครีตเวฟเล็ททรานสฟอร์มเดิม.....	37
2.33	การแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมด้วยวิธีหลายชั้น.....	37
2.34	การแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมด้วยวิธีที่ 2 หลายชั้น..	38
2.35	การเลือกสถานะแบบมาร์คอฟโดยมีสถานะ 5 สถานะ.....	42
2.36	ความน่าจะเป็นในการทอยเหรียญ.....	44
2.37	แบบจำลองฮิดเดน มาร์คอฟชนิด ไม่ต่อเนื่องของการหยิบลูกบอลจาก ถ้วย โดยมีสถานะ N สถานะ.....	45
2.38	การนำการคำนวณของ $a_i(i)$ ในเทอมของ t และ i	49
2.39	อธิบายลำดับของตัวแปรฟอร์เวิร์ด $a_{i+1}(j)$	49
2.40	วิธีของไวเตอร์บี.....	51
3.1	ขั้นตอนการดำเนินการ.....	53
3.2	แสดงการคัดแยกเฉพาะสัญญาณที่เป็นเสียงพูดจริง.....	55
3.3	การวิเคราะห์ด้วยเวฟเล็ททรานสฟอร์ม.....	55
3.4	แสดงสัญญาณที่วิเคราะห์ด้วยเวฟเล็ททรานสฟอร์ม.....	56
3.5	การนำแต่ย่านความถี่มาคสัญญาณรบกวน.....	56
3.6	ตัวอย่างสัญญาณที่ตัดสัญญาณรบกวนออก.....	57

สารบัญรูป(ต่อ)

รูปที่		หน้า
3.7	การนำสัญญาณมาแบ่งเป็นเฟรม.....	57
3.8	แสดงการแบ่งเฟรมข้อมูล.....	58
3.9	แสดงค่าจำนวนยอดคลื่น.....	58



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญ

ในปัจจุบันมนุษย์มีการติดต่อสื่อสารกับคอมพิวเตอร์ในรูปแบบต่างๆมากขึ้น อาทิเช่น การใช้แป้นพิมพ์ (Keyboard) , เมาส์ (Mouse), จอภาพระบบสัมผัส (Touch Screen) ซึ่งเทคโนโลยีทางด้านคอมพิวเตอร์นี้ช่วยอำนวยความสะดวกในการดำเนินชีวิตของมนุษย์ จึงมีการพัฒนาให้คอมพิวเตอร์มีความสามารถมากขึ้น การพัฒนาที่ออกมาหลายรูปแบบ หนึ่งในรูปแบบนั้นคือการสื่อสารกับคอมพิวเตอร์โดยใช้เสียง ปกติแล้วการสื่อสารของมนุษย์มีกันหลายวิธี เช่น การเขียน การแสดงออกทางท่าทาง และการสื่อสารส่วนใหญ่ที่มนุษย์ใช้กันคือการสื่อสารด้วยเสียง ซึ่งเป็นวิธีการที่สะดวกและง่ายที่สุด ดังนั้นหากคอมพิวเตอร์สามารถรับข้อมูลและคำสั่งจากมนุษย์ผ่านทางเสียงได้ จะทำให้เกิดความสะดวกสบายแก่มนุษย์ในการใช้งานคอมพิวเตอร์

คอมพิวเตอร์จะรับฟังเสียงจากมนุษย์ได้ย่อมเป็นการยาก เนื่องจากอุปสรรคต่าง ๆ เช่น เสียงแต่ละครั้งที่ส่งให้กับคอมพิวเตอร์นั้นจะมีความแตกต่างกัน ซึ่งเป็นรูปแบบที่ไม่แน่นอนของข้อมูลที่รับเข้ามา จึงต้องจัดการกับข้อมูลที่รับเข้ามาให้เป็นรูปแบบที่คอมพิวเตอร์เข้าใจ โดยอาศัยหลักการของฟังก์ชันถ่ายโอน (Transfer Function) และแบบจำลองฮิดเดน มาร์คอฟ (HMM : Hidden Markov Model) จากนั้นนำข้อมูลที่จัดรูปแบบแล้วมาทำการวิเคราะห์เปรียบเทียบกับข้อมูลที่อยู่ในฐานข้อมูลว่าใกล้เคียงกับข้อมูลใด แล้วนำข้อมูลที่ตรงหรือใกล้เคียงออกมาแสดง

1.2 วัตถุประสงค์ของโครงการ

1.2.1 สามารถนำหลักการรู้จำรู้จกรูปแบบ (Pattern Recognition) มาประยุกต์ใช้ในการออกแบบการปรับปรุงข้อมูลที่มีรูปแบบไม่ชัดเจนให้ชัดเจนได้

1.2.2 สามารถสร้างโปรแกรมที่รับฟังเสียงแล้วแสดงผลหรือทำงานตามคำสั่งได้ถูกต้องอย่าง

น้อยร้อยละ 70

1.3 ขอบเขตของโครงการ

1.3.1 สามารถรับฟังเสียงแบบขึ้นกับผู้พูด โดยการบันทึกเสียงของผู้ใช้ 10 ครั้งเพื่อสร้างและฝึกสอนแบบจำลองอ้างอิงของโปรแกรม

1.3.2 สามารถรับฟังเสียงได้ทีละอักขระ โดยครอออกเสียงอักขระตัวถัดไปจำเป็นต้องมีการเว้นระยะเวลา

1.3.3 สามารถเรียนรู้เสียงประจำอักขระเพิ่มเติมได้

1.3.4 โปรแกรมจำเป็นต้องรับข้อมูลเสียงมาก่อนแล้วจึงประมวลผล ไม่ได้ทำงานแบบเวลาจริง

(Real Time)

1.4 ขั้นตอนการดำเนินงาน

กิจกรรม / เดือน	12	1	2	3	4	5	6	7	8	9	10	11
1. ศึกษาและรวบรวมข้อมูล	←————→											
2. ออกแบบโปรแกรม					←————→							
3. เขียนโปรแกรมและทดสอบการทำงาน							←————→					
4. แก้ไขข้อบกพร่อง	←————→											
5. จัดทำคู่มือโครงการ			←————→									
6. ตรวจสอบและปรับปรุงแก้ไขคู่มือ				←————→								
7. ส่งโครงการ					←————→							

1.5 ผลที่คาดว่าจะได้รับ

1.5.1 เข้าใจหลักการรู้จำรู้จัก และการปรับปรุงรูปแบบข้อมูลเสียง ให้เป็นข้อมูลที่มีรูปแบบที่แน่นอน

1.5.2 สามารถนำโปรแกรมที่จัดทำขึ้นไปใช้ได้จริง

1.5.3 เป็นต้นแบบในการพัฒนาประสิทธิภาพ และเสถียรภาพของ โปรแกรมรู้จำรู้จักเสียงในแบบอื่น ๆ ต่อไป

1.6 งบประมาณที่ใช้

- ค่าหนังสือ	900 บาท
- ค่าถ่ายเอกสารและจัดทำรูปเล่ม	600 บาท
ค่าวัสดุคอมพิวเตอร์	1,000 บาท
- ค่าวัสดุสำนักงาน	500 บาท
รวม	3,000 บาท

หมายเหตุ ค่าใช้จ่ายทั้งหมดตัวเดียวกันได้



บทที่ 2

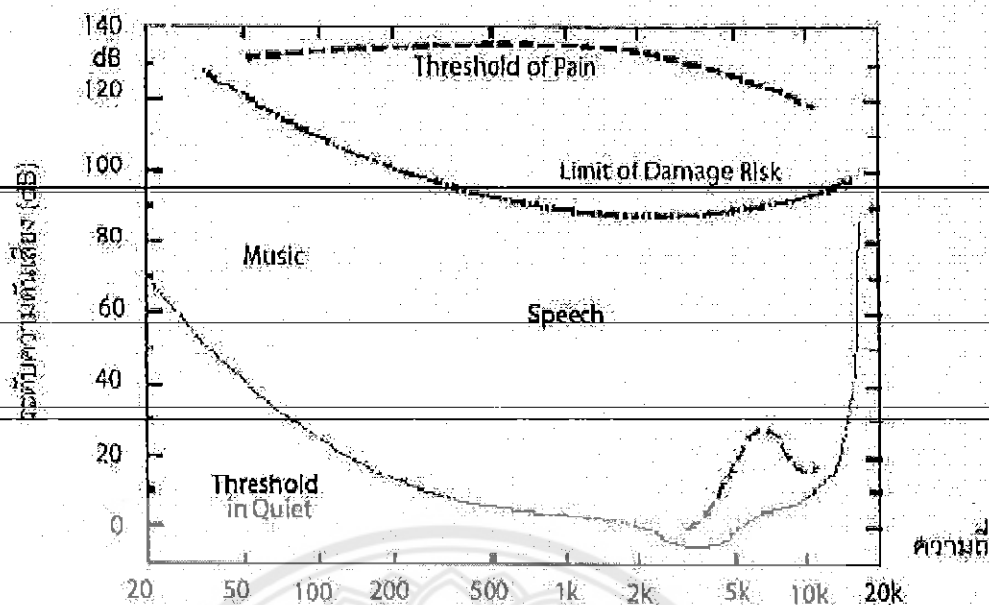
หลักการและทฤษฎี

การที่จะทำให้คอมพิวเตอร์รับข้อมูลเสียงได้ จำเป็นต้องแปลงคลื่นเสียงที่มีความถี่ไม่สม่ำเสมอในแต่ละช่วงเวลาให้มีรูปแบบแน่นอน จึงต้องอาศัยซอฟต์แวร์ทรานส์ฟอร์มมาช่วยแปลงสัญญาณให้ง่ายต่อการวิเคราะห์ เนื่องจากเสียงของแต่ละอักขระนั้นมีรูปแบบเสียงที่แตกต่างกัน จึงต้องแยกเสียงของแต่ละอักขระออกจากกัน โดยการสร้างแบบจำลองเสียงของแต่ละอักขระด้วยหลักการแบบจำลองฮิดเดน มาร์คอฟ เพื่อให้ได้แบบจำลองเสียงของแต่ละอักขระที่ง่ายต่อการวิเคราะห์ โดยจะใช้โปรแกรมแมทแปลเป็นเครื่องมือในการพัฒนา

2.1 ความสามารถในการได้ยินของมนุษย์ (Auditory field and characteristic of human hearing)

มนุษย์ได้ยินเสียง โดยรับคลื่นเสียงทางอากาศมาสู่กระดูก แล้วเปลี่ยนเป็นพลังงานไฟฟ้าและส่งต่อไปยังปลายประสาทสมองรับรู้การได้ยิน จนเกิดการแปลความหมายทำให้มนุษย์เข้าใจความหมายของเสียงที่ได้ยิน การสั่นของเซลล์ขนเร็วหรือช้าสัมพันธ์กับระดับความดันเสียงที่เปลี่ยนแปลง การที่เซลล์ขนแต่ละเซลล์สั่นก็สัมพันธ์กับความถี่ของเสียง จะเห็นได้ว่าความถี่ขององค์ประกอบเพียงอย่างเดียวในการอธิบายให้เข้าใจเรื่องของเสียงและผลกระทบของเสียง ซึ่งเรียกกระบวนการนี้ว่าการได้ยินเสียง โดยการนำเสียงทางอากาศ (Airborne conduction) การได้ยินเสียงอีกอย่างหนึ่งของมนุษย์คือการนำเสียงทางกระดูก (Bone conduction) ซึ่งจะเกิดขึ้นในกรณีที่เสียงมีความถี่ต่ำกว่า 1,000 Hz ลงไป และกรณีที่มีการสั่นสะเทือนของกระดูกในกะโหลกศีรษะ ที่ในสภาวะปกติเราไม่อาจได้ยินหรือได้ยินเบามาก แต่ถ้าเราอุดหูหรือใส่อุปกรณ์ป้องกันเสียงจะได้ยินเสียงนี้อย่างชัดเจน

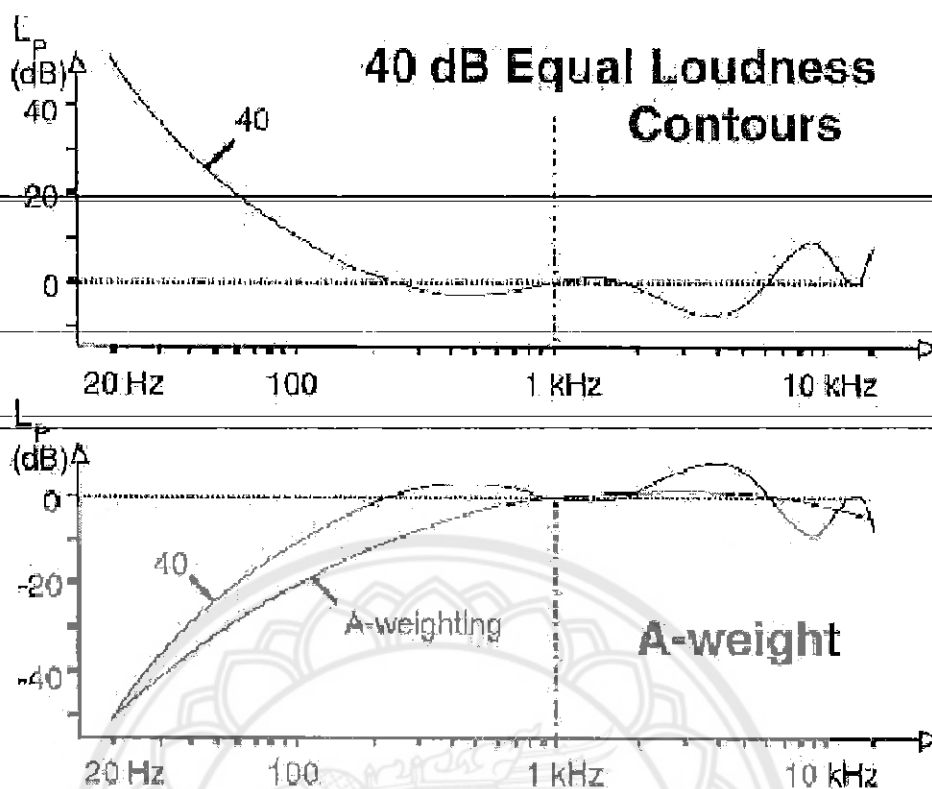
ความสามารถในการได้ยินเสียงของมนุษย์ขึ้นอยู่กับระดับความดังเสียง (0-130 dB) และช่วงความถี่ (20-20,000 Hz) - ดักรูปที่ 2-1 - ซึ่งสัมพันธ์กับकर्สันของเซลล์ขนภายในหูชั้นใน - และมีความสัมพันธ์แบบไม่เป็นเส้นตรง (Non-linear relationship) กล่าวคือ ที่ระดับเสียงเท่ากัน แต่หูมนุษย์จะได้ยินเป็นความดังที่ไม่เท่ากันทุกความถี่ ซึ่งมีนักวิจัยหลายท่านได้พยายามอธิบาย ลักษณะของการได้ยินของมนุษย์นี้ออกมาในรูปของกราฟความสัมพันธ์เพื่อให้เข้าใจถึงลักษณะการได้ยินของหูที่ระดับเสียงและความถี่ต่าง ๆ กัน จะเห็นได้ว่าที่ความถี่ 20 Hz ระดับเสียงต้องดัง 80 dB มนุษย์จึงจะเริ่มได้ยิน แต่ที่ความถี่ 1,000 Hz ดังเพียง 5 dB ก็สามารถได้ยินเสียงนั้นและมนุษย์จะมีความอ่อนไหวหรือไวต่อการได้ยินมากที่สุดที่ความถี่ 4,000 Hz



รูปที่ 2.1 ความสามารถในการได้ยินเสียงของหูมนุษย์ระดับความดังเสียง (0-130 dB) และ ช่วงความถี่ (20-20,000 Hz)

ขอบเขตการได้ยินของหูมนุษย์ (Auditory field) อยู่ในช่วงความถี่ 20-20,000 Hz และ 0-130 dB แต่ละคนได้ยินเสียงไม่เท่ากัน บางคนอาจได้ยินเสียงที่มีความถี่ต่ำมาก ๆ หรือระดับเสียงต่ำมาก ๆ เมื่อเทียบกับคนอื่น ในขณะที่บางคนอาจไม่ได้ยินเสียงที่ระดับความดันต่ำกว่า 25 dB ที่ความถี่ 1,000 Hz เลย ตามปกติขีดเริ่มการได้ยิน (Threshold of hearing) ของหูมนุษย์ในระดับความดังที่ต่ำที่สุดคือ 20 dB เมื่อเปรียบเทียบกับขอบเขตของเสียงดนตรี (Music field) ซึ่งมีระดับเสียงประมาณ 20-90 dB ช่วงความถี่ประมาณ 50-10,000 Hz ขอบเขตเสียงสนทนาของมนุษย์ (Human speech) จะมีระดับเสียงประมาณ 25-80 dB ในช่วงความถี่ประมาณ 100-4,000 Hz และขีดเริ่มของความเจ็บปวด (Threshold of pain) ของหูมนุษย์อยู่ที่ประมาณ 130 dB

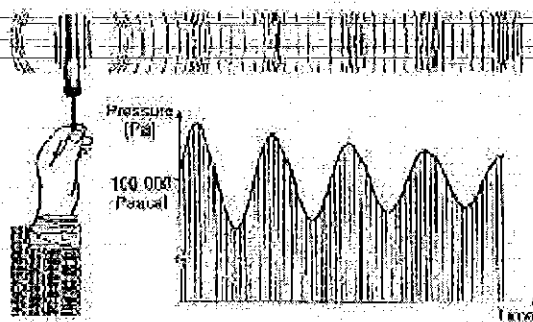
ลักษณะการได้ยินของหูที่มีความสัมพันธ์แบบไม่เป็นเส้นตรงนี้ ทำให้เครื่องมือที่จะใช้ในการตรวจวัดระดับเสียง เพื่อศึกษาผลกระทบที่เกิดกับมนุษย์จำเป็นต้องมีวงจร ถ่วงน้ำหนักความถี่ ให้เครื่องมือสามารถวัดเสียงได้เหมือนลักษณะการได้ยินของมนุษย์มากที่สุด ที่นิยมใช้ในปัจจุบันได้แก่ วงจรถ่วงน้ำหนักแบบ A ซึ่งจะมีค่าใกล้เคียงกับเส้นผกผันของเส้นเท่าระดับเสียงต่ำสุด (40 dB Equal loudness contours) หรือขีดเริ่มการได้ยิน (รูปที่ 2.2)



รูปที่ 2.2 เส้นเท่าระดับเสียง และเส้นถ่วงน้ำหนักแบบ A

2.2 คลื่นเสียง (Sound wave)

คลื่นเสียงเป็นคลื่นตามยาว (Longitudinal wave) ดังรูปที่ 2.3 เมื่อเคาะส้อมเสียงจะมีการสั่นของโมเลกุลในส้อมเสียงทำให้เกิดการเปลี่ยนแปลงความดันบรรยากาศรอบ ๆ ส้อมเสียง เกิดคลื่นออกมาโดยรอบส้อมเสียง คล้ายกับคลื่นน้ำเมื่อเราโยนก้อนหินลงไปในน้ำ เรียกว่า การแผ่กระจายของคลื่นเสียง (Sound propagation)



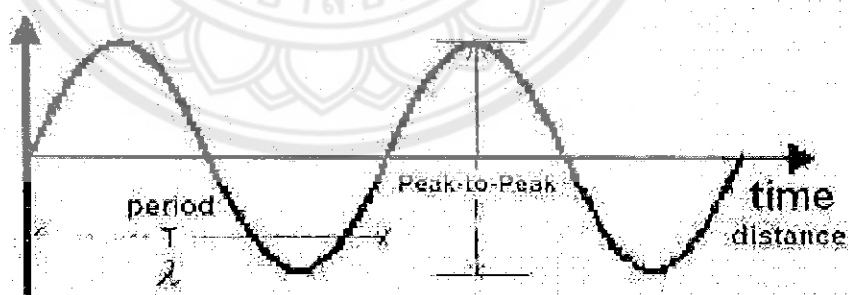
รูปที่ 2.3 คลื่นเสียง (Sound wave)

2.2.1 คุณสมบัติทางกายภาพของคลื่นเสียง (Physics properties of sound wave)

คลื่นเสียงที่เราได้ยินนั้นไม่ว่าจะเป็นเสียงจากแหล่งกำเนิดเสียงใด ส่วนใหญ่จะเป็นเสียงที่ซับซ้อนซึ่งเกิดจากการรวมกันของคลื่นเสียงที่มีความถี่และระดับความดันเสียงต่าง ๆ เข้าด้วยกัน ในการศึกษาเรื่องฟิสิกส์ของคลื่นเสียง สามารถอธิบายเป็นสมการคณิตศาสตร์ให้เข้าใจได้ง่ายขึ้นในรูปแบบของคลื่นรูปไซน์ (Sine Wave) เช่น การเคาะส้อมเสียงเกิดการเคลื่อนที่หรือความสั่นสะเทือนของโมเลกุลอากาศที่เป็นผลมาจากการเปลี่ยนแปลงความดันบรรยากาศโดยรอบ เป็นส่วนอัดส่วนขยาย เรียกว่า ความดันเสียง ซึ่งส่งผ่านทางโมเลกุลของตัวกลางไปยังจุดรับเสียง พารามิเตอร์ทางฟิสิกส์ ที่สำคัญที่จะอธิบายธรรมชาติของการสั่นของโมเลกุลอากาศ จากการเปลี่ยนแปลงความดันบรรยากาศ หรือการเคลื่อนที่ของคลื่นเสียงในตัวกลาง ในที่นี้คืออากาศ ผ่านกระบวนการได้ยินของมนุษย์จนรับรู้และเข้าใจเสียงนั้น มีอยู่ 2 พารามิเตอร์ที่สำคัญคือ ความถี่ของเสียง และความยาวคลื่นเสียง ซึ่งมีความเกี่ยวข้องกับอัตราเร็วของเสียงในอากาศ ทิศทางการแผ่กระจายของเสียง และเวลาที่เสียงเดินทางจากแหล่งกำเนิด ไปยังผู้รับเสียง

2.2.2 ความถี่ (f : Frequency)

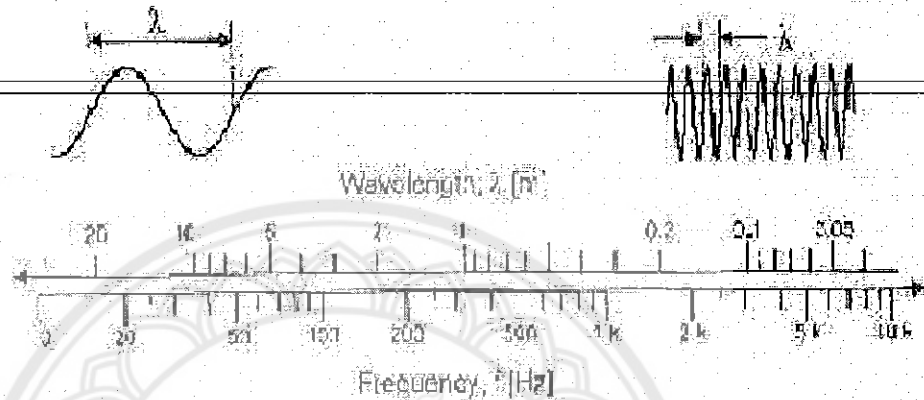
การเคลื่อนที่ของความดันบรรยากาศในตัวอย่างการเคาะส้อมเสียง หรือการเคลื่อนที่ของลูกสูบในท่อ เกิดปรากฏการณ์ส่วนอัดส่วนขยายเดินทางไปตามท่อ ซึ่งนำมาเขียนกราฟระหว่างแอมพลิจูดกับเวลาหรือแอมพลิจูดกับระยะทางจะได้กราฟรูปคลื่นไซน์ (รูปที่ 2.4) จำนวนรอบการเกิดส่วนอัดส่วนขยายเราเรียกว่า ความถี่ของคลื่น มีหน่วยเป็นรอบต่อวินาที (Hz)



รูปที่ 2.4 คลื่นรูปไซน์

2.2.3 ความยาวคลื่น (λ : Wavelength)

จากกราฟการเคลื่อนที่ของคลื่นเสียงระหว่างแอมพลิจูด (Amplitude) กับระยะทาง ความยาวคลื่นเสียง คือระยะทางจากยอดคลื่นหนึ่ง หรือระยะทางของการเกิดคลื่น 1 คลื่น (รูปที่ 2.5) มีหน่วยเป็นเมตร และความยาวคลื่นนั้นมีความสำคัญอย่างมากในการศึกษาการควบคุมเสียง



รูปที่ 2.5 ความยาวคลื่นกับความถี่

2.3 ทฤษฎีเวฟเล็ต (Wavelet)

เนื่องจากเวฟเล็ตและเวฟเล็ตทรานสฟอร์มเป็นแนวคิดใหม่ ที่เพิ่งเกิดขึ้นเพื่อใช้กับสัญญาณหรือภาพที่ต้องการทั้งข้อมูลทางด้านเวลาและความถี่โดยเวฟเล็ตทรานสฟอร์มได้ถูกพัฒนาขึ้นเพื่อใช้ในการแก้ไขปัญหาที่ไม่สามารถกระทำได้ในฟูเรียร์ทรานสฟอร์ม และชอร์ตไทม์ฟูเรียร์ทรานสฟอร์ม แต่เวฟเล็ตทรานสฟอร์มก็มีความคล้ายคลึงกับทั้งฟูเรียร์ทรานสฟอร์มและชอร์ตไทม์ฟูเรียร์ทรานสฟอร์มจึงขอกกล่าวถึงฟูเรียร์ทรานสฟอร์มและชอร์ตไทม์ฟูเรียร์ทรานสฟอร์มไว้ด้วยพอสังเขป

2.3.1 ฟูรีเยร์ทรานสฟอร์ม (FT : Fourier Transform)

เป็นรูปแบบหนึ่งของการแปลงสัญญาณเพื่อให้่ายต่อการวิเคราะห์ โดยจะแปลงสัญญาณจากโดเมน (Domain) ของเวลาไปอยู่ในโดเมนของความถี่

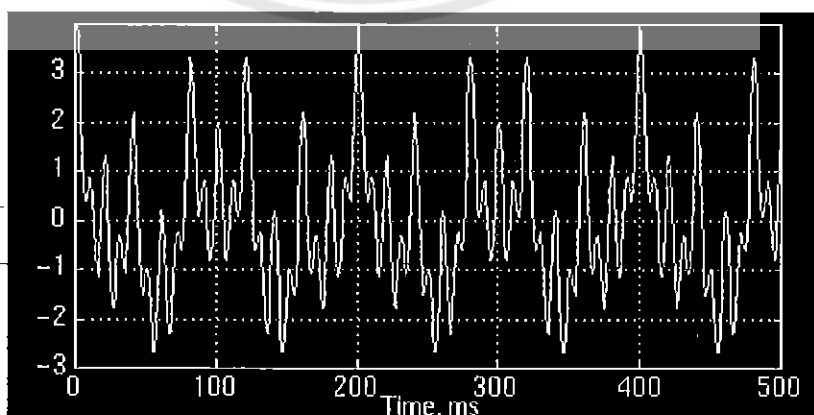
ฟูรีเยร์ทรานสฟอร์มใช้หลักการว่าคลื่นใด ๆ เกิดจากการบวกคลื่นรูปไซน์ที่มีความถี่ไม่คงที่หลาย ๆ ความถี่หลาย ๆ ตัวเข้าด้วยกัน

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2j\pi ft} dt \quad (2.1)$$

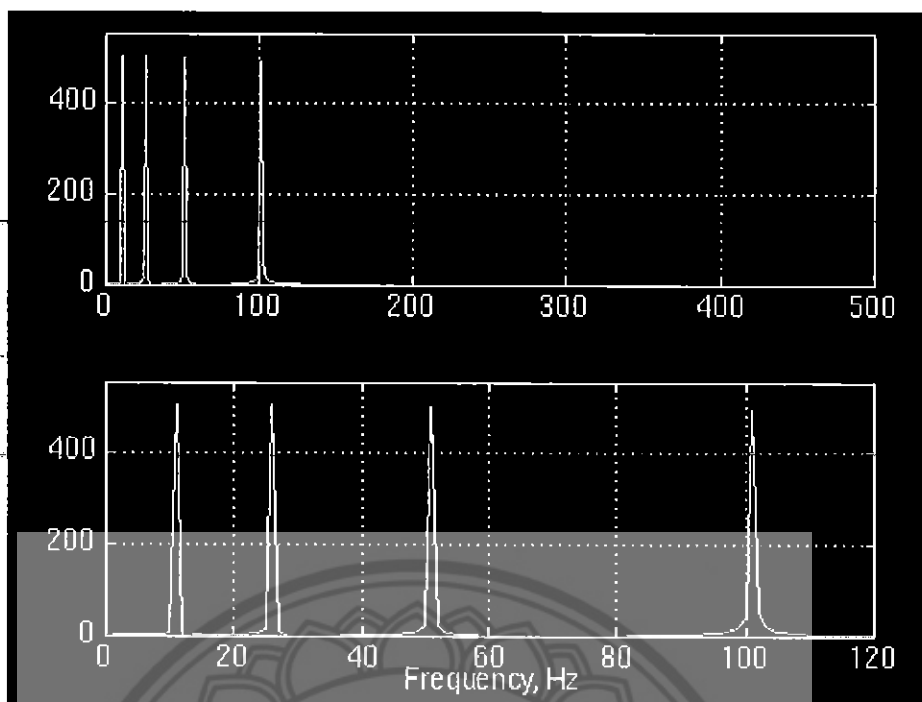
$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{2j\pi ft} df \quad (2.2)$$

สมการ (2.1) เป็นสมการแปลงสัญญาณให้เป็นฟูรีเยร์ทรานสฟอร์มของสัญญาณนั้น ซึ่งจากสมการจะเห็นว่า ถ้าสัญญาณมีความถี่ใดที่ตรงกับคลื่นรูปไซน์ความถี่ f จะให้ผลคูณที่มีค่ามากที่สุดแสดงเป็นค่าความแตกต่างระหว่างห้องคลื่นถึงสันคลื่น (Amplitude peak) ในโดเมนความถี่ของฟูรีเยร์ทรานสฟอร์ม และสมการ (2.2) เป็นสมการแปลงฟูรีเยร์ทรานสฟอร์ม ของสัญญาณนั้นกลับเป็นสัญญาณเดิม

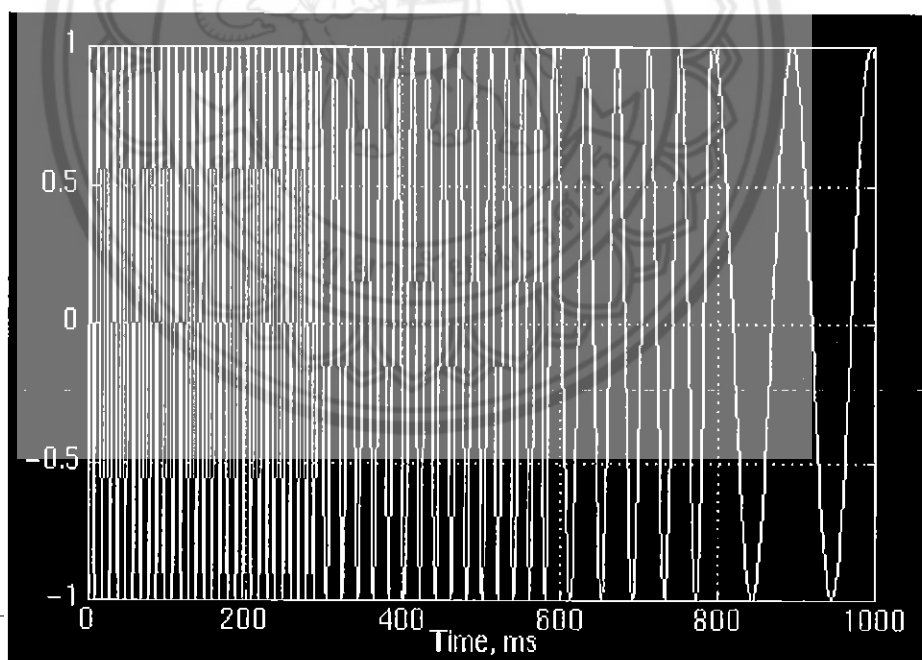
ฟูรีเยร์ทรานสฟอร์มของสัญญาณทำให้ทราบความถี่ของสัญญาณนั้น ๆ แต่ไม่ทราบเวลาที่เกิดความถี่นั้น ๆ ดังแสดงในรูปที่ 2.6 สัญญาณสเตชันนารี (Stationary signal) ความถี่ 100 , 50 , 25 และ 10 Hz , รูปที่ 2.7 ฟูรีเยร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.6 , รูปที่ 2.8 สัญญาณนอนสเตชันนารี (Non-stationary signal) ความถี่ 100 , 50 , 25 , 10 Hz , รูปที่ 2.9 ฟูรีเยร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.8



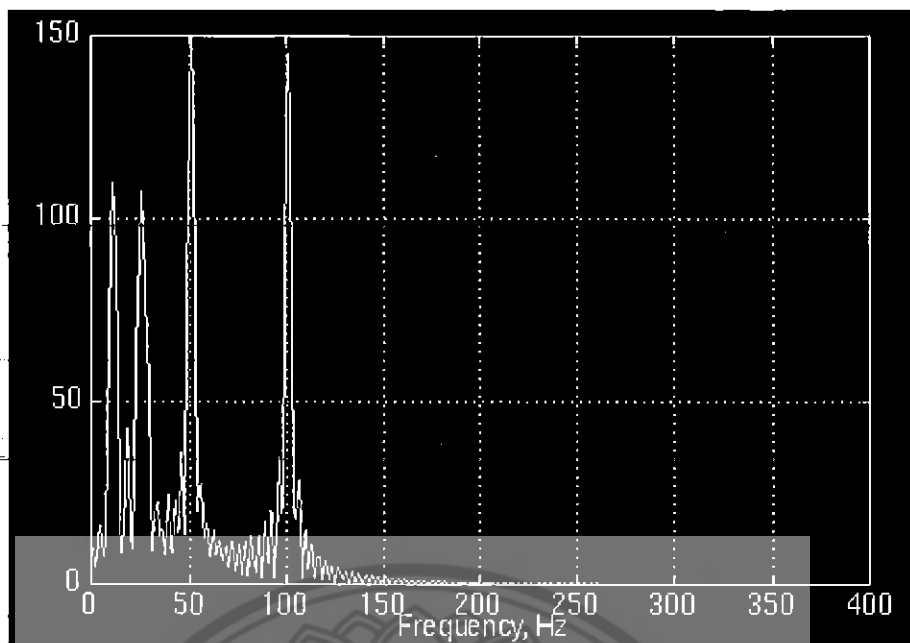
รูปที่ 2.6 สัญญาณสเตชันนารี ความถี่ 100 , 50 , 25 และ 10 Hz



รูปที่ 2.7 พูรีเยร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.6



รูปที่ 2.8 สัญญาณนอนสเตรชันนารี ความถี่ 100 , 50 , 25 , 10 Hz



รูปที่ 2.9 พูรีเยร์ทรานส์ฟอร์มของสัญญาณในรูปที่ 2.8

รูปที่ 2.6 แสดงสัญญาณที่ความถี่ไม่เปลี่ยนแปลงตามเวลา คือความถี่ทุก ๆ ความถี่ที่มีในสัญญาณจะอยู่ทุก ๆ เวลาของสัญญาณ เรียกสัญญาณเช่นนี้ว่า สัญญาณสเตชันนารี จะเห็นว่าสัญญาณในรูปที่ 2.6 และ 2.8 มีความถี่เดียวกันแต่มีรูปสัญญาณไม่เหมือนกัน ยังคงมีฟูรีเยร์ทรานส์ฟอร์มของสัญญาณซึ่งแสดงในรูปที่ 2.7 และ 2.9 เหมือนกัน (โดยเฉพาะค่าความแตกต่างระหว่างท้องถิ่นถึงท้องถิ่น ของฟูรีเยร์ทรานส์ฟอร์ม โดยไม่ดูค่าย่อยอื่น ๆ) เนื่องจากว่าฟูรีเยร์ทรานส์ฟอร์มบอกได้เฉพาะข้อมูลของความถี่แต่ไม่บอกข้อมูลของเวลา จึงทำให้ไม่ทราบว่าสัญญาณที่มีความถี่ใดเกิดที่เวลาใด ดังนั้นฟูรีเยร์ทรานส์ฟอร์มจึงไม่เหมาะกับการวิเคราะห์ที่ต้องการทราบข้อมูลของเวลาด้วย เช่นการวิเคราะห์สัญญาณนอนสเตชันนารี ดังรูปที่ 2.8 ถ้าผู้วิเคราะห์ต้องการทราบช่วงเวลาที่เกิดแต่ละความถี่ ก็ไม่เหมาะที่จะใช้ฟูรีเยร์ทรานส์ฟอร์ม จึงเกิดการปรับปรุงการวิเคราะห์ข้อมูลเป็นซอร์ทไทม์ฟูรีเยร์ทรานส์ฟอร์ม ซึ่งจะทำได้ข้อมูลทั้งด้านความถี่และทางด้านเวลา

2.3.2 ซอร์ทไทม์ฟูรีเยร์ทรานส์ฟอร์ม (STFT : Short Time Fourier Transforms)

เป็นรูปแบบหนึ่งของการแปลงสัญญาณจากโดเมนของเวลา ไปอยู่ในโดเมนของความถี่และเวลาที่เกิดความถี่นั้น ใช้หลักการคล้ายกับฟูรีเยร์ทรานส์ฟอร์ม คือคลื่นใด ๆ เกิดจากการบวกคลื่นรูปไซน์หลาย ๆ ความถี่หลาย ๆ ตัวเข้าด้วยกัน แต่เนื่องจากฟูรีเยร์ทรานส์ฟอร์มเหมาะกับสัญญาณที่เป็นสเตชันนารี ดังนั้นเมื่อทำทรานส์ฟอร์มกับสัญญาณนอนสเตชันนารี (สัญญาณที่ความถี่ขึ้นอยู่กัช่วงเวลา) จึงทำการแบ่งสัญญาณออกเป็นช่วง ๆ แต่ละช่วงสัญญาณย่อย ๆ ที่แบ่งออกมาจะเป็นสัญญาณสเตชันนารี และวิเคราะห์ด้วยฟูรีเยร์ทรานส์ฟอร์มได้

สัญญาณที่แบ่งออกเป็นช่วง ๆ จะใช้หน้าต่าง ๆ (Window หรือ Compactly supported) ในการแบ่ง โดยจะเริ่มวางหน้าต่างที่ต้นสัญญาณ จากนั้นจะทำการคำนวณแบบฟูเรียร์ทรานสฟอร์ม แล้วเลื่อนหน้าต่างถัดไปเพื่อคำนวณต่อไปจนหมดสัญญาณ การเลื่อนหน้าต่างนี้จะทำให้ได้ข้อมูลของเวลา และการคำนวณแบบฟูเรียร์ทรานสฟอร์มจะทำให้ได้ข้อมูลของความถี่

การแปลงสัญญาณเป็นซอร์ทใหม่ฟูเรียร์ทรานสฟอร์ม แสดงได้ดังสมการ (2.3)

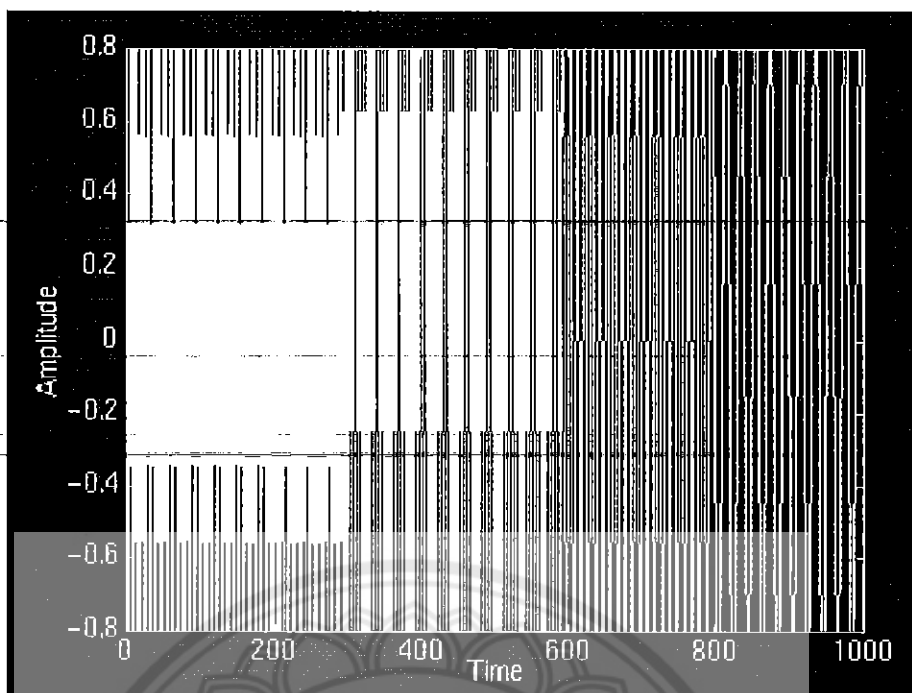
$$STFT_x^{(\omega)}(t', f) = \int [x(t) \cdot \omega * (t - t')] \cdot e^{-j2\pi f t} dt \quad (2.3)$$

ในรูปที่ 2.10 จะแสดงสัญญาณและซอร์ทใหม่ฟูเรียร์ทรานสฟอร์มของสัญญาณที่ใช้หน้าต่างขนาดแตกต่างกัน โดยหน้าต่างจะใช้ฟังก์ชันเกาส์เซียน ($\omega(t)$: Gaussian function) ตามสมการ

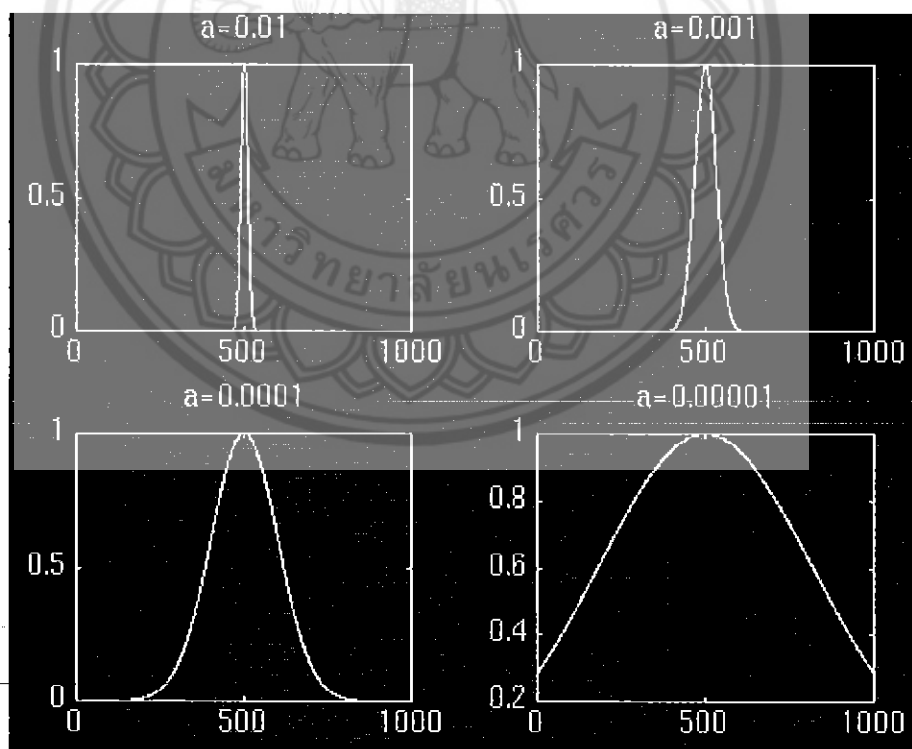
$$\omega(t) = e^{-at^2/2} \quad (2.4)$$

โดย a บอกถึงความยาวของหน้าต่าง และ t คือเวลา

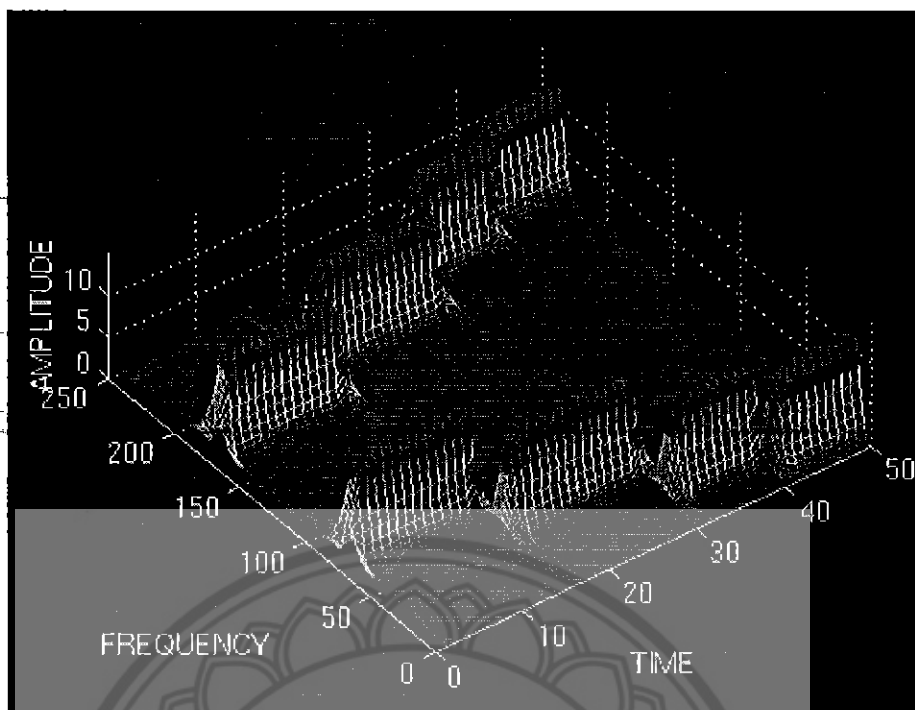
ในซอร์ทใหม่ฟูเรียร์ทรานสฟอร์มจะใช้หน้าต่างที่มีความกว้างเดียวตลอดในการทำทรานสฟอร์มครั้งหนึ่ง ๆ ทำให้ในสัญญาณหนึ่ง ๆ มีความละเอียด (Resolution) ทางด้านเวลาเท่ากันตลอดทั้งสัญญาณ (ความละเอียดแม่นยำของเวลาไม่ว่าจะละเอียดหรือไม่ก็ตาม แต่ความละเอียดจะคงค่านั้นตลอดทั้งสัญญาณ) และความละเอียดด้านความถี่จะเท่ากันตลอดทั้งสัญญาณ (ความละเอียดแม่นยำของความถี่ไม่ว่าจะละเอียดหรือไม่ก็ตาม แต่ความละเอียดจะคงค่านั้นตลอดทั้งสัญญาณ) ไม่ว่าสัญญาณนั้นจะมีความถี่เป็นเท่าใดและมีความถี่นั้นอยู่ที่ใด



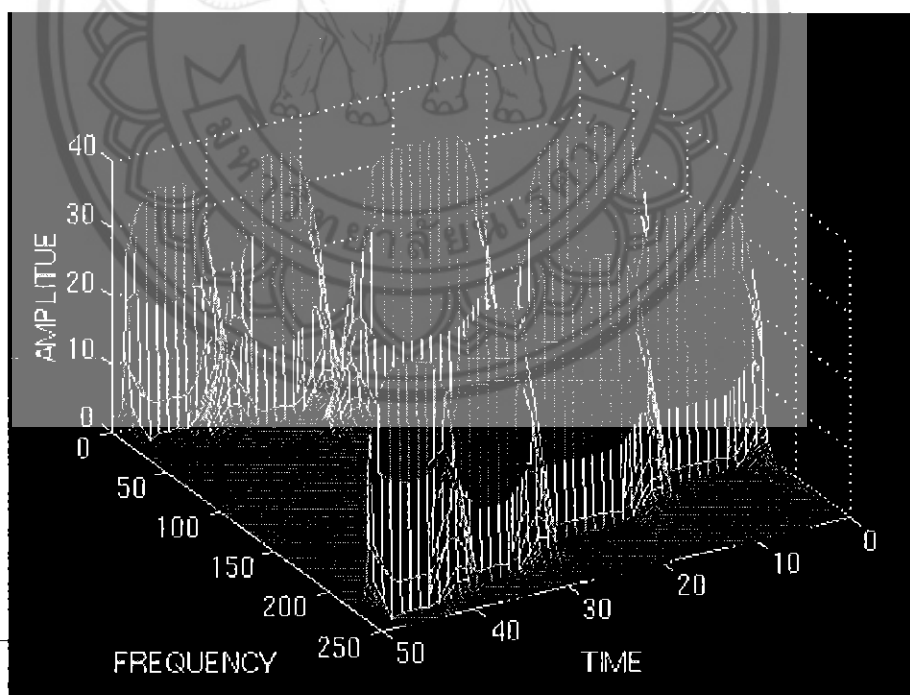
รูปที่ 2.10 สัญญาณที่มีความถี่ 300, 200, 100 และ 50 Hz ตามลำดับ



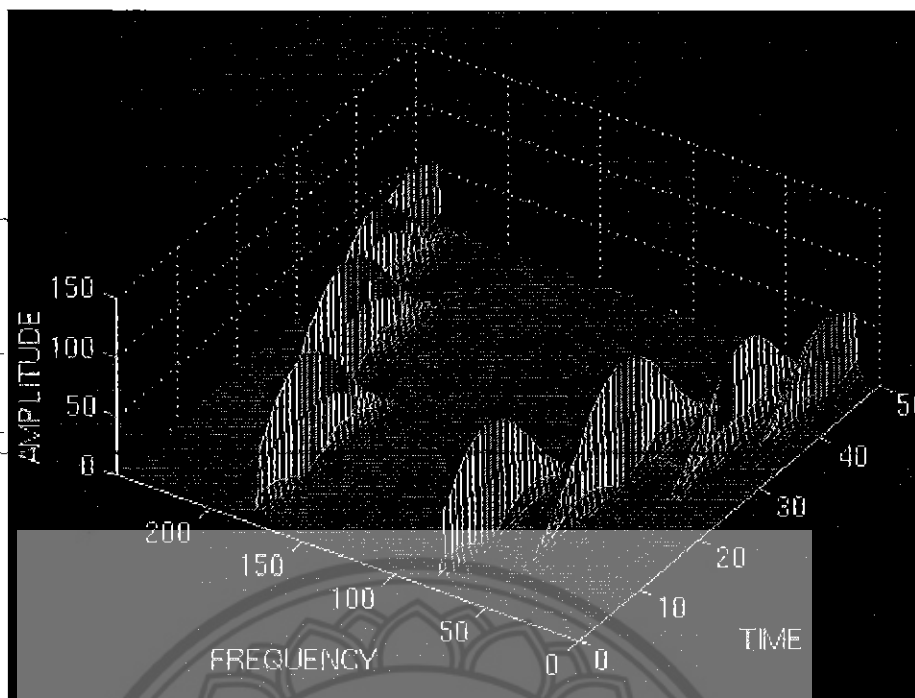
รูปที่ 2.11 ความยาวของหน้าต่างขนาด $a = 0.01, 0.001, 0.0001$ และ 0.00001



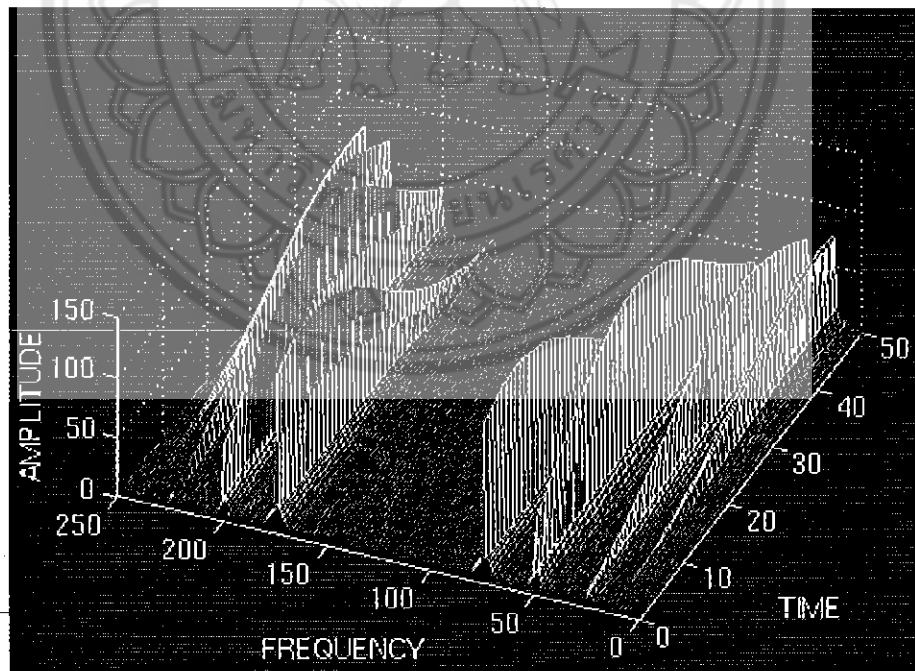
รูปที่ 2.12 ความยาวของหน้าตาขนาด $a = 0.01, 0.001, 0.0001$ และ 0.00001 แบบสามมิติ



รูปที่ 2.13 ซอร์ทไทม์ฟูเรียร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.10
เมื่อความยาวของหน้าตาขนาด $a = 0.001$



รูปที่ 2.14 ซอร์ทโทมฟ์ฟูเรียร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.10
เมื่อความยาวของหน้าต่างขนาด $a = 0.0001$



รูปที่ 2.15 ซอร์ทโทมฟ์ฟูเรียร์ทรานสฟอร์มของสัญญาณในรูปที่ 2.10
เมื่อความยาวของหน้าต่างขนาด $a = 0.00001$

จะเห็นว่าการใช้หน้าต่างแคบ เช่นในรูปที่ 2.14 จะให้ความละเอียดด้านเวลาดี คือสามารถแบ่งเวลาได้ชัดเจนเป็นช่วง ๆ ขณะที่ความละเอียดด้านความถี่ไม่ดี คือในแต่ละช่วงเวลามีความถี่ที่กว้างและไม่ชี้ชัดว่าความถี่ในช่วงเวลานี้เป็นเท่าใด เมื่อหน้าต่างกว้างขึ้น ความละเอียดทางด้านเวลาจะลดลง คือข้อมูลทางด้านเวลาจะคลุมเครือขึ้น ไม่สามารถชี้เวลาได้ชัดเจน ขณะที่ความละเอียดทางด้านความถี่จะเพิ่มขึ้น คือ สามารถระบุความถี่ได้ชัดเจนยิ่งขึ้น

สามารถสรุปปัญหาของซอร์ทไทม์ฟูเรียร์ทรานสฟอร์มได้ดังนี้

1. เมื่อใช้หน้าต่างแคบ ความละเอียดด้านเวลาดี ความละเอียดด้านความถี่ไม่ดี
2. เมื่อใช้หน้าต่างกว้าง ความละเอียดด้านเวลาไม่ดี ความละเอียดด้านความถี่ดี และทำให้สัญญาณไม่เป็นสเตชันนารี

จะเห็นว่าซอร์ทไทม์ฟูเรียร์ทรานสฟอร์มมีปัญหาเรื่องความละเอียด ดังนั้นจึงต้องหาวิธีการวิเคราะห์ในรูปแบบใหม่ซึ่งก็คือเวฟเล็ตทรานสฟอร์ม

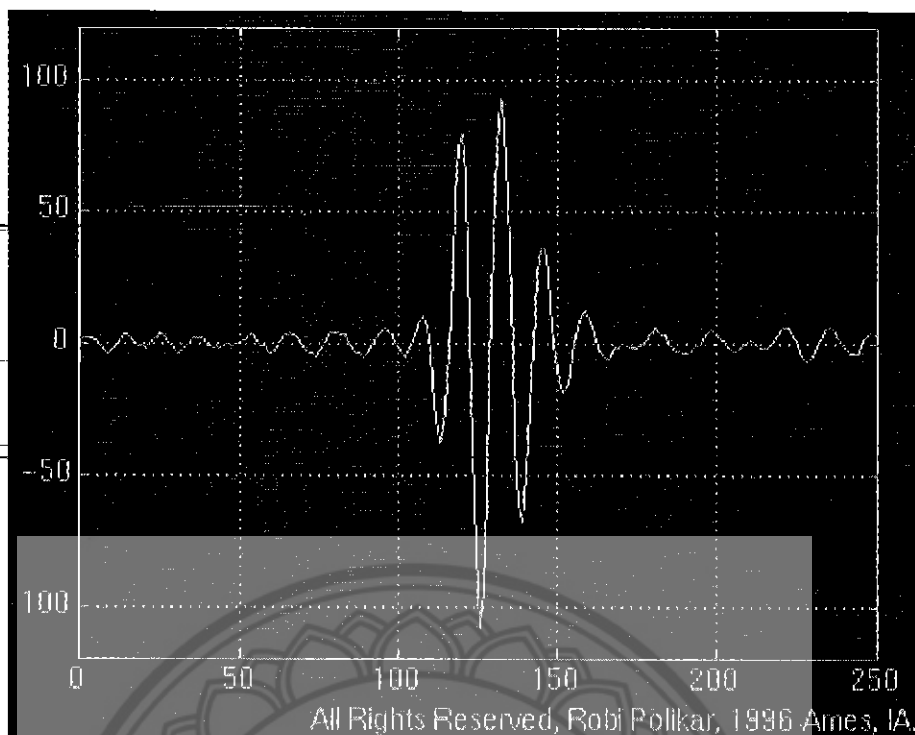
2.3.3 เวฟเล็ตทรานสฟอร์ม (WT : Wavelet Transform)

จากการวิเคราะห์สัญญาณที่ผ่านมา ฟูเรียร์ทรานสฟอร์มให้ข้อมูลเฉพาะความถี่ไม่ได้ให้ข้อมูลของเวลาที่เกิดความถี่นั้น และซอร์ทไทม์ฟูเรียร์ทรานสฟอร์มให้ข้อมูลทั้งความถี่และเวลา แต่มีปัญหาเรื่องความละเอียดแม่นยำคือเมื่อให้ความละเอียดของเวลาสูง จะทำให้ความละเอียดของความถี่ต่ำ และเมื่อให้ความละเอียดของความถี่สูง ก็จะทำให้ความละเอียดของเวลาต่ำ จึงได้มีการวิเคราะห์สัญญาณแบบใหม่ที่เรียกว่าการวิเคราะห์แบบหลายความละเอียด (MRA : Multiresolution Analysis) เป็นการวิเคราะห์สัญญาณที่แต่ละช่วงมีความถี่ต่างกัน ด้วยความละเอียดที่แตกต่างกัน การวิเคราะห์เช่นนี้ต่างจากซอร์ทไทม์ฟูเรียร์ทรานสฟอร์มที่วิเคราะห์สัญญาณหนึ่ง ๆ ด้วยความละเอียดด้านเวลาและความถี่ดีหรือไม่ดีเท่า ๆ กันกับทุก ๆ ความถี่ในสัญญาณนั้น

การวิเคราะห์แบบหลายความละเอียดจะให้ความละเอียดดังนี้

1. ในช่วงที่เป็นความถี่สูงของสัญญาณ : การวิเคราะห์แบบหลายความละเอียดจะให้ความละเอียดของเวลาดี แต่ให้ความละเอียดของความถี่ไม่ดี
2. ในช่วงที่เป็นความถี่ต่ำของสัญญาณ : การวิเคราะห์แบบหลายความละเอียด จะให้ความละเอียดของเวลาไม่ดี แต่ให้ความละเอียดของความถี่ดี

การที่ให้ความละเอียดแตกต่างกันเช่นนี้ เป็นผลดีกับสัญญาณที่มีความถี่สูงเป็นช่วงสั้นๆ นอกนั้นเป็นความถี่ต่ำ ซึ่งสัญญาณที่พบและนำมาใช้ในความเป็นจริงมักจะอยู่ในลักษณะดังรูปที่ 2.16

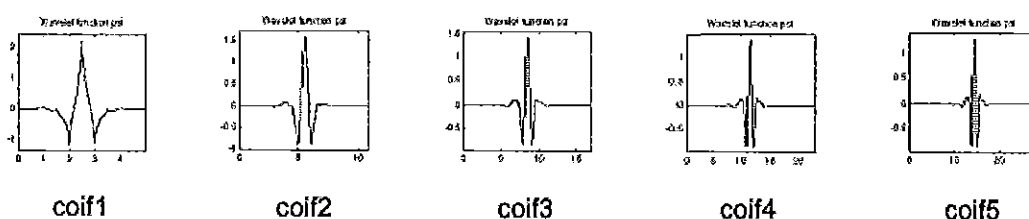


รูปที่ 2.16 สัญญาณที่มีความถี่สูงเป็นช่วงสั้น ๆ

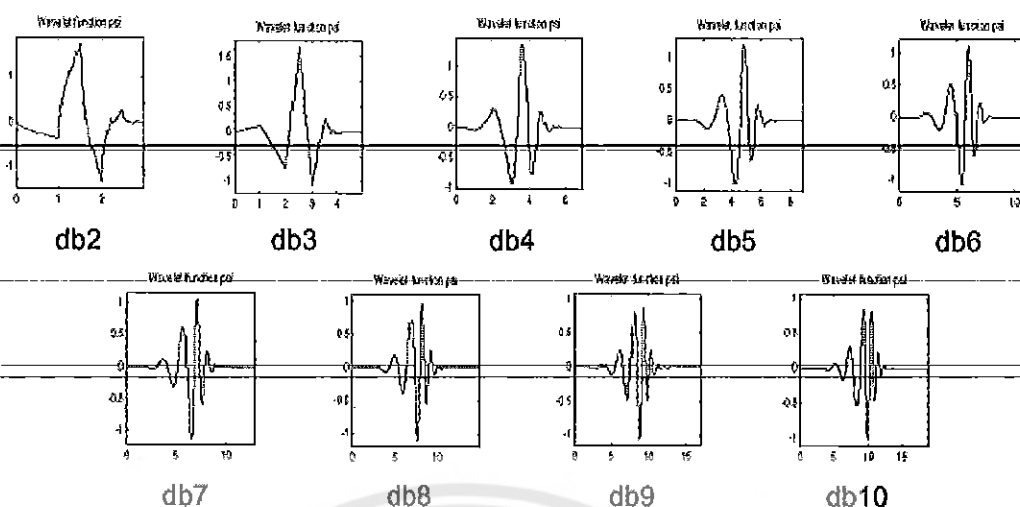
2.3.3.1 ความหมายของเวฟเล็ทดาวบีชีล์(เวฟเล็ท *) และเวฟเล็ทแม่ (Mother wavelet)

เวฟเล็ท หมายถึง คลื่นเล็ก ๆ ซึ่งมีความยาวและพลังงานจำกัด ใช้เป็นเครื่องมือในการวิเคราะห์สัญญาณที่มีการเปลี่ยนแปลงความถี่ไปตามเวลา (Time-varying) สัญญาณที่มีการเปลี่ยนแปลง โดยฉับพลัน (Transient) สัญญาณที่ไม่ได้มีทุกความถี่ในสัญญาณผสมกันอยู่ตลอดช่วงสัญญาณ หรือเป็นสัญญาณที่ไม่ได้มีลักษณะเป็นคาบ เวฟเล็ททำหน้าที่คล้ายกับฟังก์ชันหน้าต่างต่าง และฟังก์ชันเอ็กซ์โปเนนเชียล (e^x : Exponential function) ในซอร์ทโทมัวร์ฟูเรียร์ทรานสฟอร์ม เปรียบเทียบคลื่นรูปไซน์และเวฟเล็ทดาวบีชีล์

เวฟเล็ทแม่ ($\Psi(t)$: Mother wavelet) คือ ตัวเวฟเล็ทต้นแบบที่ใช้สร้างเวฟเล็ทตัวอื่น ๆ โดยการยืดหรือหดเวฟเล็ทแม่ ตัวอย่างเช่น ดาวบีชีล์ (Daubechies), โคอิเฟ็ล็ตส์ (Coiflets)



รูปที่ 2.17 ตัวอย่างของเวฟเล็ท โคอิเฟ็ล็ตส์ (Coiflets)



รูปที่ 2.18 ตัวอย่างของเวฟเล็ทดาบิวชีส์ (Daubechies)

2.3.3.2 ความหมายของเวฟเล็ททรานสฟอร์มและการวิเคราะห์แบบหลายความละเอียด

เวฟเล็ททรานสฟอร์ม คือสัญญาณที่ถูกแปลงให้อยู่ในโดเมนของความถี่และเวลา โดยการนำเอาสัญญาณเดิมซึ่งอยู่ในโดเมนของเวลามาผ่าน สมการทรานสฟอร์มและใช้เวฟเล็ทที่เหมาะสม จะได้ผลลัพธ์เป็นเวฟเล็ททรานสฟอร์มของสัญญาณเดิม ซึ่งเวฟเล็ททรานสฟอร์มนี้จะบอกได้ถึง ความถี่ที่มีในสัญญาณนั้น ๆ และเวลาที่เกิดความถี่นั้นขึ้น

เวฟเล็ททรานสฟอร์มมีความคล้ายคลึงกับฟูริเยร์ทรานสฟอร์มและซอร์ทใหม่ฟูริเยร์ทรานสฟอร์มสามารถเปรียบเทียบได้ดังตารางที่ 2.1

ตารางที่ 2.1 เปรียบเทียบฟูริเยร์ทรานสฟอร์ม ซอร์ทใหม่ฟูริเยร์ทรานสฟอร์ม และเวฟเล็ททรานสฟอร์ม

FT	STFT	WT
1. ใช้สมการ FT	1. ใช้สมการ STFT	1. ใช้สมการ WT
2. ใช้ฟังก์ชันเอ็กซ์โปเนนเชียล $e^{-2j\pi ft}$ หรือก็คือฟังก์ชัน $\cos(2\pi ft) + j\sin(2\pi ft)$ เป็นแบบเบสิสฟังก์ชัน (Basis function)	2. ใช้ฟังก์ชันเอ็กซ์โปเนนเชียล เป็นเบสิสฟังก์ชัน เหมือนกับ FT และมีการใช้หน้าต่าง $w * (t - t')$	2. ใช้เวฟเล็ท $\psi_{r,s}^*(t)$ เป็นเบสิสฟังก์ชัน

ตารางที่ 2.1(ต่อ) เปรียบเทียบฟูเรียร์ทรานสฟอร์ม ซอรัทใหม่ฟูเรียร์ทรานสฟอร์ม และเวฟเล็ตทรานสฟอร์ม

FT	STFT	WT
3. ใช้หลักการว่าคลื่นใด ๆ เกิดจากการบวกคลื่นไซน์ ซอยด์หลาย ๆ ความถี่ หลาย ๆ ตัวเข้าด้วยกัน	3. ใช้หลักการแทนคลื่นใด ๆ ด้วยคลื่นไซน์ซอยด์บวกกัน เหมือน FT แต่แตกต่างกัน ตรงที่ STFT ทำการแบ่งคลื่นใด ๆ ออกเป็นช่วง ๆ แล้วจึงค่อยคิดทำให้ได้ทั้งเวลาและความถี่ของคลื่นใด ๆ นั้น	3. ใช้หลักการแทนคลื่นใด ๆ ด้วยคลื่นเล็ก ๆ ที่มีควมยาว และพลังงานจำกัด (เวฟเล็ต) บวกกัน ทำให้ได้ทั้งเวลาและความถี่ของคลื่นใด ๆ นั้น
4. ได้ทรานสฟอร์มเป็นฟังก์ชันของความถี่ จะทราบว่าเป็นคลื่นใด ๆ นั้นมีความถี่ใด อยู่บ้าง	4. ได้ทรานสฟอร์มเป็นฟังก์ชันของความถี่และเวลาจะทราบว่าในคลื่นใด ๆ มีความถี่ใดอยู่ที่เวลาใด โดยความละเอียดของเวลาและความถี่จะเท่ากันทุก ๆ ความถี่ในคลื่นหนึ่ง ๆ	4. ได้ทรานสฟอร์มเป็นฟังก์ชันของความถี่และเวลาจะทราบว่าในคลื่นใด ๆ มีความถี่ใดอยู่ที่เวลาใด โดยความละเอียดของเวลาและความถี่จะเปลี่ยนแปลงไปตามความถี่ในคลื่นดังกล่าว ความละเอียดของ MRA ที่กล่าวไว้ข้างต้น
5. ไม่เป็น MRA	5. ไม่เป็น MRA	5. เป็น MRA

ในที่นี้ ศึกษาเวฟเล็ตทรานสฟอร์ม 2 แบบ คือ

1. คอนทินิวอัลเวฟเล็ตทรานสฟอร์ม (CWT : Continuous Wavelet Transform)
2. ดิสครีตเวฟเล็ตทรานสฟอร์ม (DWT : Discrete Wavelet Transform)

2.3.3.3 คอนทินิวอัลเวฟเล็ตทรานสฟอร์ม (CWT : Continuous Wavelet Transform)

คอนทินิวอัลเวฟเล็ตทรานสฟอร์ม ได้รับการพัฒนาขึ้นมาเพื่อแก้ไขปัญหาด้านความละเอียดของซอรัทใหม่ฟูเรียร์ทรานสฟอร์ม การวิเคราะห์ด้วยคอนทินิวอัลเวฟเล็ตทรานสฟอร์มมีความคล้ายคลึงกับการวิเคราะห์ด้วยซอรัทใหม่ฟูเรียร์ทรานสฟอร์ม โดยที่คอนทินิวอัลเวฟเล็ตทรานสฟอร์มจะคุณสมบัติความถี่ด้วยเวฟเล็ต ซึ่งคล้ายกับซอรัทใหม่ฟูเรียร์ทรานสฟอร์มตรงที่สัญญาณเดิมถูกคูณด้วยฟังก์ชันหน้าต่าง (Window function) และในการคำนวณสัญญาณเดิมจะถูกแบ่งวิเคราะห์

ออกเป็นส่วน ๆ แยกขาดจากกัน แต่ซอร์ทโทมัสฟูเรียร์ทรานสฟอร์มกับคอนทินิวอัลสเวฟเล็ททรานสฟอร์มก็มีข้อแตกต่างกัน 2 ข้อหลัก คือ

1. ความกว้างของหน้าต่างในเวฟเล็ททรานสฟอร์ม (ซึ่งก็คือตัวเวฟเล็ท) จะเปลี่ยนไปเรื่อย ๆ ทุก ๆ ความถี่และทำการคำนวณทรานสฟอร์ม ซึ่งข้อนี้เป็นคุณสมบัติที่สำคัญที่สุดข้อหนึ่งของเวฟเล็ททรานสฟอร์ม

2. สัญญาณที่ถูกนำมาผ่านฟังก์ชันหน้าต่างแล้ว (Windowed signals) ไม่ถูกทำฟูเรียร์ทรานสฟอร์ม ดังนั้น คลื่นรูปไซน์ครึ่งลูกจึงดูเหมือนคลื่นรูปไซน์เต็มลูก นั่นคือความถี่ที่เป็นลบไม่ได้ถูกคำนวณ

สมการของคอนทินิวอัลสเวฟเล็ททรานสฟอร์ม

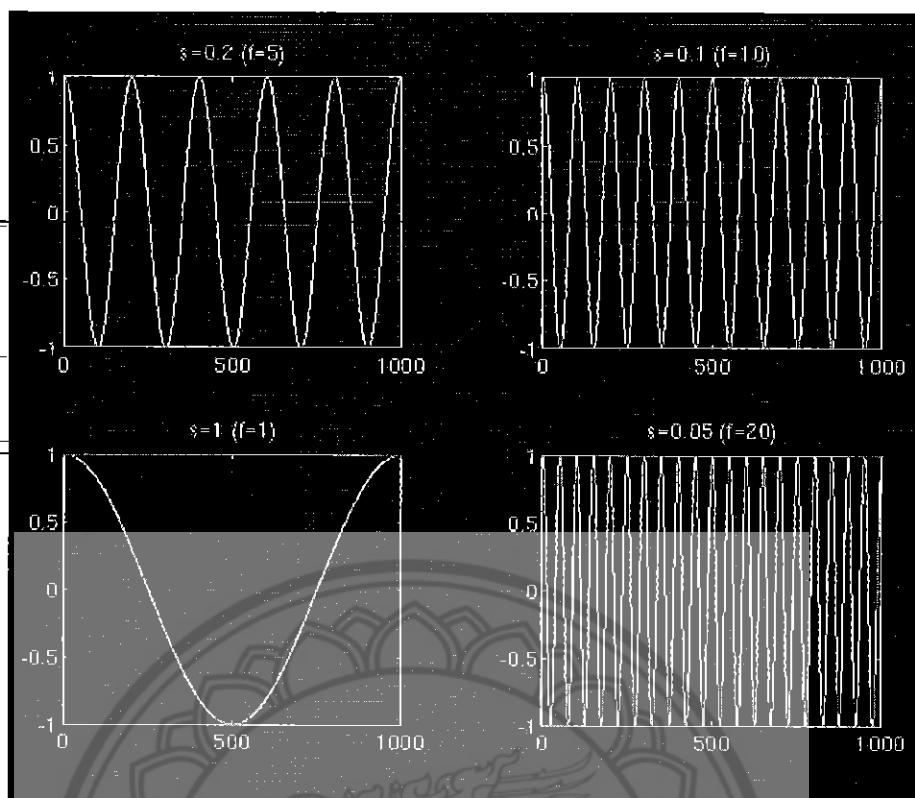
$$CWT_x^\psi(\tau, s) = \Psi_x^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi * \left(\frac{t - \tau}{s}\right) dt \quad (2.5)$$

จากสมการ (2.5) ข้างต้น สังเกตได้ว่า

- CWT เป็นฟังก์ชัน 2 ตัวแปร คือ τ (Translation) และ s (Scale)

- ทรานสเลชัน (Translation) คือตำแหน่งที่เวฟเล็ทเคลื่อนที่ไป ซึ่งจะเป็นข้อมูล 'เวลา' ในทรานสฟอร์มโดเมน
- สเกล (Scale) คือคาบ (1/ความถี่) การเพิ่มสเกล ($s > 1$) จะเป็นเหมือนการยืดสัญญาณออก และการลดสเกล ($s < 1$) จะเป็นเหมือนการบีบสัญญาณเข้ารูปที่ 2.19 จะแสดงถึงการเพิ่มและลดสเกลของสัญญาณ โคไซน์เดียวกัน

- $\psi(t)$ คือ เวฟเล็ทแม่



a	b
c	d

รูปที่ 2.19 แสดงการเพิ่มและลดสเกลของสัญญาณ โคไซน์เดียวกัน โดย (c) , (a) , (b) , (d) จะเรียงจากสเกลสูง ไปยังสเกลต่ำตามลำดับ

จากรูป จะเห็นว่าถ้าสเกลสูง เช่นรูปที่ 2.19(c) มีสเกล = 1 จะทำให้สัญญาณยืดออก (มีความถี่ต่ำ) ขณะที่ถ้าสเกลต่ำ เช่นรูปที่ 2.19(d) มีสเกล = 0.05 จะทำให้สัญญาณหดเข้า (มีความถี่สูง) สัญญาณในความเป็นจริงส่วนใหญ่จะเป็นสัญญาณที่มีสเกลสูง (ความถี่ต่ำ) และมีส่วนที่เป็นสเกลต่ำ (ความถี่สูง) อยู่แทรกเป็นช่วงสั้น ๆ

2.3.3.3.1 การคำนวณคอนทินิวอัสเวฟเล็ตทรานสฟอร์ม

การทำคอนทินิวอัสเวฟเล็ตทรานสฟอร์มคือการทำคอร์รีเลท (Correlation) ระหว่างเวฟเล็ตหลาย ๆ สเกลกับสัญญาณ เพื่อวัดความเหมือนของเวฟเล็ตและสัญญาณ

จากสมการ (2.5) ให้ $x(t)$ เป็นสัญญาณที่จะทำทรานสฟอร์ม จากนั้นเลือกเวฟเล็ตแม่เพื่อให้เป็นต้นแบบของหน้าต่างในการคำนวณ โดยหน้าต่างที่ใช้ทั้งหมดในการคำนวณจะเกิดจากการยืดหรือหดเวฟเล็ตแม่

ในขั้นแรกจะทำการคำนวณเวฟเล็ตแม่ที่ $s = 1$ ก่อน จากนั้นจึงค่อยคำนวณที่ s น้อยกว่าหรือมากกว่า 1 ต่อไปตามความเหมาะสมของสัญญาณเพียงแค่เลือกช่วงค่า s ที่เหมาะสมกับสัญญาณนั้นก็พอ(ซึ่งโดยทั่วไปต้องคำนวณทุกค่า s)

ในที่นี้จะยกตัวอย่างการคำนวณโดยเริ่มจาก $s = 1$ (เวฟเล็ตจะแคบ) แล้วค่อย ๆ เพิ่มค่า s (เวฟเล็ตจะค่อย ๆ ขยายขึ้น) คือจะทำการคำนวณจากความถี่สูง ไปยังความถี่ต่ำ

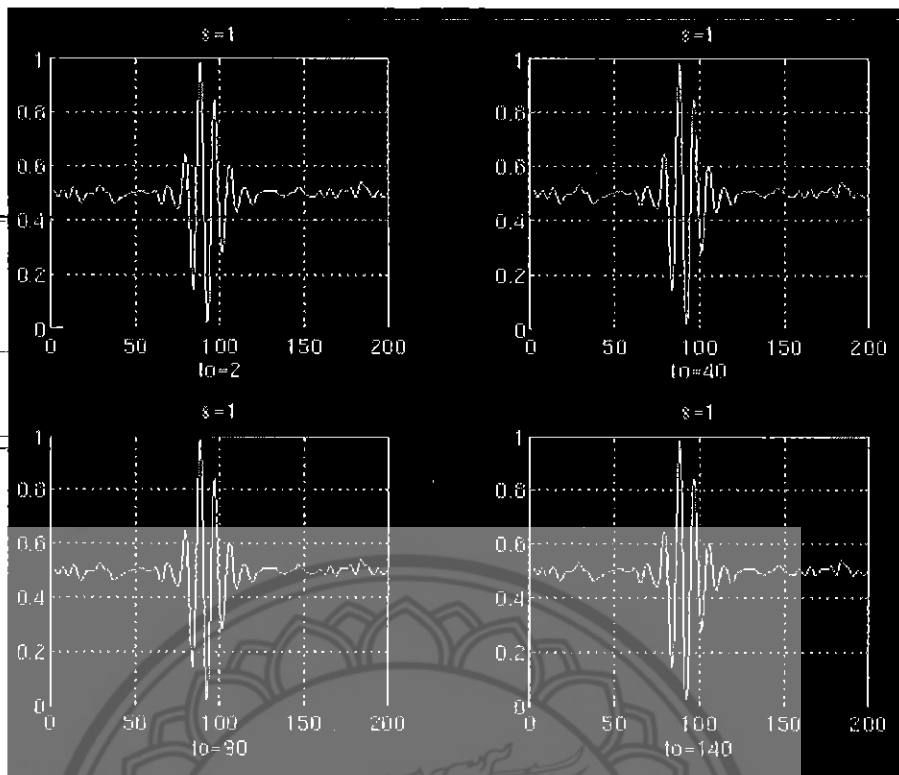
การคำนวณจะมีขั้นตอนดังนี้

1. วางเวฟเล็ตที่มี $s = 1$ ไว้ที่จุดเริ่มต้นของสัญญาณ (เวลา = 0)
2. คูณสัญญาณด้วยเวฟเล็ต
3. รวมผลคูณ โดยบวกทุกจุดตลอดทั้งสัญญาณ

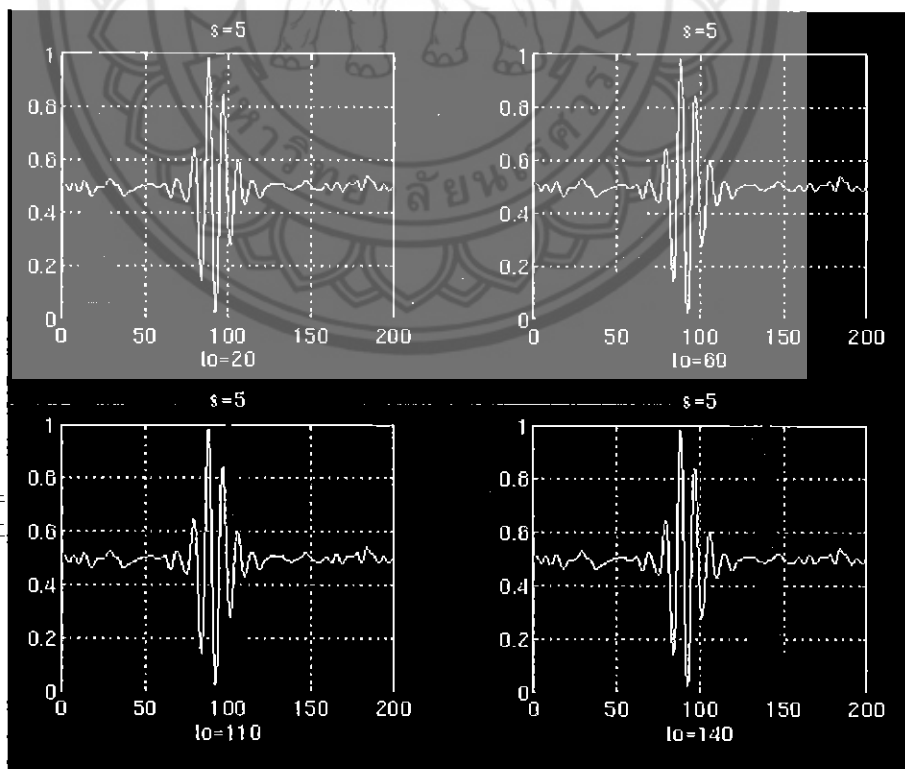
4. นำผลลัพธ์ของการบวกมาคูณกับ $1/\sqrt{s}$ เพื่อทำการนอร์มอลไลซ์ (Normalize) พลังงานให้สัญญาณที่ทรานสฟอร์มแล้ว มีพลังงานเท่า ๆ กันทุกสเกล จะได้ผลลัพธ์เป็นตัวทรานสฟอร์มของสัญญาณเดิม ซึ่งก็คือค่าของ CWT ที่เวลา $t=0, s=1$ ในระนาบเวลา-สเกล (Time-scale plane)

5. จากนั้นทำการเลื่อนเวฟเล็ต $s = 1$ ไปทางขวาของสัญญาณเป็นระยะ τ หน่วย เวฟเล็ตจะอยู่ที่ $t = \tau$ แล้วทำการคำนวณตามขั้นตอน 1-4 อีกครั้ง จะได้ค่า CWT ที่ $t = \tau, s = 1$ ในระนาบเวลา-สเกล ทำขั้นตอนที่ 5 ซ้ำไปจนหมดสัญญาณ จะได้ค่า CWT ของสัญญาณในระนาบเวลา-สเกล หนึ่งแถว คือแถว $s = 1$ ดังรูปที่ 2.19

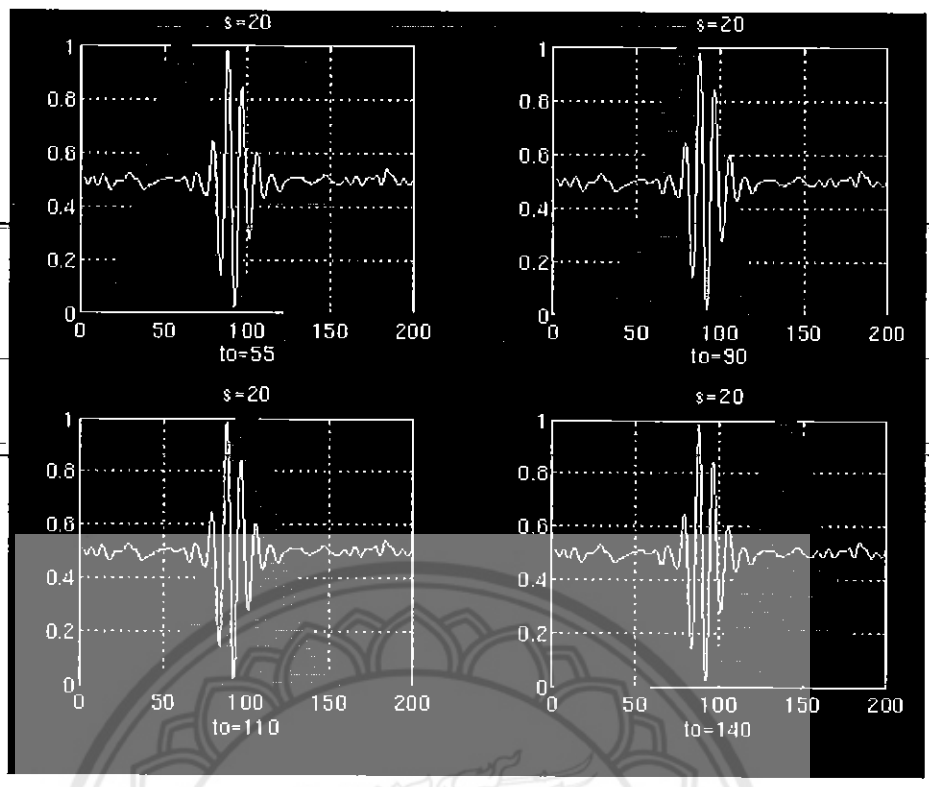
6. เพิ่มค่าสเกลของเวฟเล็ต (เพิ่มค่า s) แล้วคำนวณตามขั้นตอน 1-5 ซ้ำ จนได้ค่า s ที่ต้องการครบ จะทำให้ได้ค่า CWT ทั้งหมดของสัญญาณ



รูปที่ 2.20 แสดงสัญญาณ และเวฟเลต $s = 1$ ที่ τ ที่ = 2, 40, 90, 140



รูปที่ 2.21 แสดงสัญญาณและเวฟเลต $s = 5$ ที่ $\tau = 20, 60, 110, 140$



รูปที่ 2.22 แสดงสัญญาณ และเวฟเล็ท $s = 20$ ที่ $T = 50, 90, 110, 140$

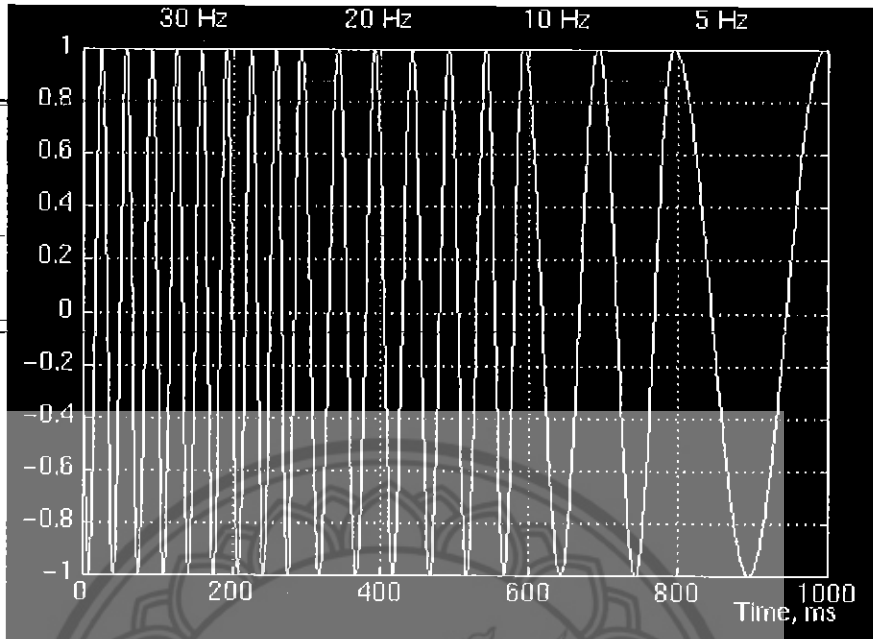
จากรูปสังเกตว่าผลคูณจะไม่เป็น 0 ในจุดที่สัญญาณถูกเวฟเล็ททับเท่านั้น เมื่อเวฟเล็ทเคลื่อนที่จะได้ข้อมูลเกี่ยวกับเวลา และเมื่อเปลี่ยนแปลงขนาดสเกลของเวฟเล็ท จะได้ข้อมูลของความถี่ ถ้าในสัญญาณมีความถี่ที่ตรงกับความถี่ของเวฟเล็ทในขณะนี้ (ขณะทำการคำนวณ) จะทำให้จุดนั้นได้ผลคูณที่มีค่ามาก ถ้าในสัญญาณไม่มีความถี่ที่ตรงกับความถี่ของเวฟเล็ทในขณะที่ยคำนวณผลคูณจะมีค่าน้อย

ดังนั้น สัญญาณในรูปที่ 2.20, 2.21 และ 2.22 ที่เวฟเล็ทสเกลต่ำ (ความถี่สูง) จะให้ผลคูณกับสัญญาณมีค่ามากที่รอบ ๆ $t = 100$ ms และที่เวฟเล็ทสเกลสูง (ความถี่ต่ำ) จะให้ผลคูณกับสัญญาณมีค่ามากที่จุดอื่น ๆ ทั่ว ๆ ไปในสัญญาณ เนื่องจากสัญญาณนี้มีความถี่ต่ำเป็นองค์ประกอบเป็นส่วนใหญ่

4800043

ตัวอย่าง คอนทิวนิวส์เวฟเล็ททรานสฟอร์มของสัญญาณอนสเตรชันนารี

156692x e. 2

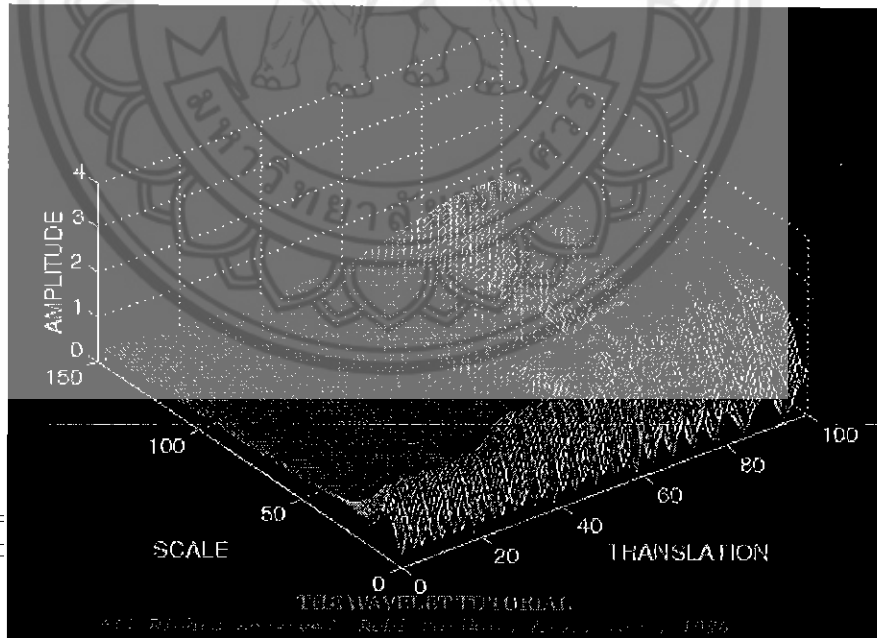


ร.ร.

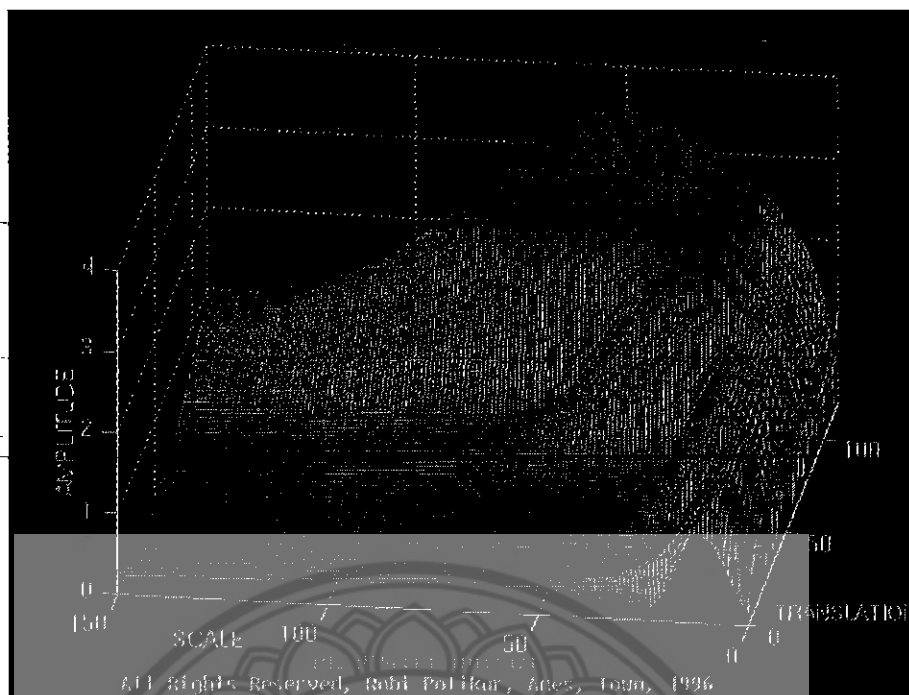
16591

2549

รูปที่ 2.23 สัญญาณอนสเตรชันนารีความถี่ 30 , 20 , 10 , 5 Hz



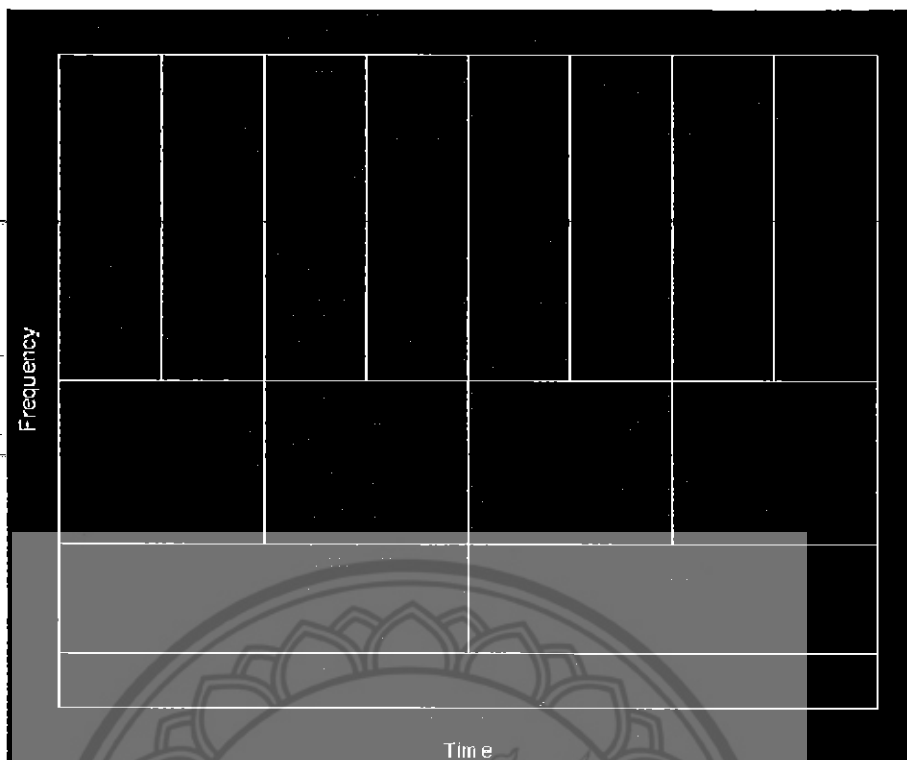
รูปที่ 2.24 แสดงคอนทิวนิวส์เวฟเล็ททรานสฟอร์มในมุมมองต่าง ๆ ในระนาบทรานสเลชัน-สเกล ของสัญญาณในรูป 2.23



รูปที่ 2.25 แสดงคอนทิวอัสเวฟเล็ดทรานสฟอร์ม ในมุมมองต่าง ๆ
ในระนาบทรานสเลชัน-สเกล ของสัญญาณในรูป 2.23

สังเกตคอนทิวอัสเวฟเล็ดทรานสฟอร์มของสัญญาณพบว่าในช่วงทรานสเลชัน 0-30 มีค่าความแตกต่างระหว่างท้องถิ่นถึงท้องถิ่นอยู่ที่สเกลต่ำสุด (ความถี่สูงสุด) และเมื่อทรานสเลชันมากขึ้น ค่าความแตกต่างระหว่างท้องถิ่นถึงท้องถิ่นจะปรากฏที่สเกลสูงขึ้นตามลำดับ (ความถี่ต่ำลง) ซึ่งตรงกับสัญญาณที่มีความถี่สูงสุดตรงช่วงต้นของสัญญาณและเมื่อเวลาเพิ่มขึ้นความถี่ก็ลดลงจาก 30 Hz เป็น 20, 10 และ 5 Hz ที่ปลายสัญญาณ รูปที่ 2.25 แสดงให้เห็นถึงคุณสมบัติเรื่องความละเอียดของเวฟเล็ดทรานสฟอร์มด้วย คือ

- ที่สเกลต่ำ (ความถี่สูง) จะมีช่วงสเกลที่แคบกว่า คือมีความคลุมเครือสเกลน้อยกว่า ซึ่งหมายความว่ามีความละเอียดในด้านสเกลดีกว่า ดังนั้นจะมีความละเอียดในด้านความถี่น้อยกว่า
- ที่สเกลสูง (ความถี่ต่ำ) มีช่วงสเกลที่กว้างกว่า คือมีความละเอียดในด้านสเกลน้อยกว่า ดังนั้น จะมีความละเอียดในด้านความถี่ดีกว่า



รูปที่ 2.26 ความละเอียดของเวฟเล็ททรานสฟอร์ม

รูปที่ 2.26 อธิบายถึงความละเอียดของเวฟเล็ททรานสฟอร์ม แต่ละช่วงแสดงถึงค่าในระนาบเวลา-ความถี่ โดยทุก ๆ ช่องมีพื้นที่เท่า ๆ กัน แตกต่างกันที่ความกว้างและความยาวของแต่ละช่อง

- ที่ความถี่ต่ำ ช่องจะต่ำและกว้าง แสดงให้เห็นถึงการมีความละเอียดในด้านความถี่ดี (เพราะสามารถระบุค่าความถี่ได้ชัดเจนกว่า) แต่ความละเอียดในด้านเวลาไม่ดี
- ที่ความถี่สูง ช่องจะสูงและแคบ แสดงให้เห็นถึงการมีความละเอียดในด้านความถี่ไม่ดี แต่ความละเอียดในด้านเวลาดี

2.3.3.3.2 การแปลงคอนทินิวอัสเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม

(Reconstruction of wavelet transform)

คือการทำเวฟเล็ททรานสฟอร์มให้กลับมาเป็นสัญญาณเดิมโดยใช้สมการ (2.6)

$$x(t) = \frac{1}{c_\psi^2} \iint \Psi_x^\psi(\tau, s) \frac{1}{s^2} \psi\left(\frac{t-\tau}{s}\right) d\tau ds \quad (2.6)$$

โดย c_ψ เป็นค่าคงที่ซึ่งขึ้นกับเวฟเล็ตที่ใช้เรียกว่า ค่าคงที่แอดมิสซิบิลิตี (Admissibility constant) สมการ (2.6) จะขึ้นอยู่กับค่าคงที่นี้ตามเงื่อนไขแอดมิสซิบิลิตี (Admissibility condition) ดังแสดงในสมการ (2.7)

$$c_\psi = \left\{ 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi \right\}^{1/2} < \infty \quad (2.7)$$

$\hat{\psi}(\xi)$ คือฟูริเยร์ทรานสฟอร์มของ $\psi(t)$ สมการ (2.7) บ่งบอกถึง $\hat{\psi}(0) = 0$ ซึ่งก็คือ

$$\int \psi(t) dt = 0 \quad (2.8)$$

สมการ (2.8) จะเป็นจริงได้ก็ต่อเมื่อ เวฟเล็ตมีการแกว่งไปมา (Oscillate) ซึ่งตามปกติ สมการ (2.8) มักจะเป็นจริง

2.3.3.3 คอนทินิวัสเวฟเล็ตทรานสฟอร์มแบบดิสครีต (Discretization of the continuous wavelet transform)

จากการคำนวณคอนทินิวัสเวฟเล็ตทรานสฟอร์มที่กล่าวมาข้างต้น จะเห็นว่าสามารถคำนวณเป็นแบบดิสครีตได้ โดยเป็นการสุ่มตัวอย่างบนระนาบทรานสเลชัน-สเกลซึ่งเป็นการสุ่มตัวอย่างที่เป็นแบบแผน (Uniform sampling rate) แต่ในเวฟเล็ตทรานสฟอร์มจะใช้การเปลี่ยนสเกล (ซึ่งก็คือการเปลี่ยน 1/ความถี่) มาช่วยลดอัตราการสุ่ม (Sampling rate) ตัวอย่างได้ โดยที่สเกลสูง ๆ (ความถี่ต่ำ ๆ) สามารถลดอัตราการสุ่มตัวอย่างได้ตามกฎของไนควิสต์ (Nyquist's rule)

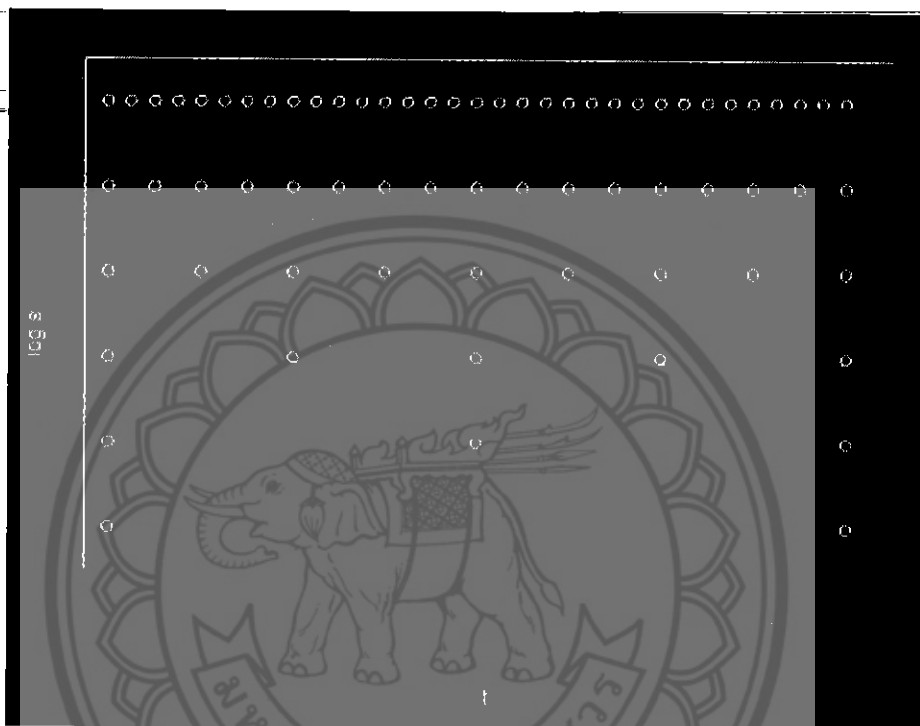
ในระนาบเดียวกัน สามารถใช้อัตราการสุ่มตัวอย่าง N_1 ที่สเกล S_1 และอัตราการสุ่มตัวอย่าง N_2 ที่สเกล S_2 เมื่อ $S_1 < S_2$ ($f_1 > f_2$) สามารถเขียนเป็นความสัมพันธ์ได้ดังนี้

$$N_2 = \frac{S_1}{S_2} N_1 \quad \text{หรือ} \quad N_2 = \frac{f_2}{f_1} N_1 \quad (2.9)$$

การลดอัตราการสุ่มจะช่วยทำให้การคำนวณลดลง โดยการลดอัตราการสุ่มของทรานสฟอร์มของสัญญาณนี้จะทำตามกฎของไนควิสต์ก็ต่อเมื่อเมื่อมีความต้องการที่จะทำทรานสฟอร์มนี้กลับไปเป็นสัญญาณเดิม เนื่องมาจากอัตราการสุ่มตัวอย่างไนควิสต์คืออัตราการสุ่มตัวอย่างที่ต่ำที่สุดที่ทำให้สามารถนำจำนวนตัวอย่างที่สุ่ม (ในที่นี้ตัวอย่างที่สุ่มคือทรานสฟอร์มของสัญญาณ)

กลับไปสร้างเป็นสัญญาณเดิมได้ ถ้าไม่ต้องการนำตัวอย่างที่สุ่มกลับไปสร้างเป็นสัญญาณเดิมก็สามารถที่จะลดอัตราการสุ่มตัวอย่างให้น้อยลงกว่าอัตราการสุ่มในวิธีที่ได้

ในระนาบเวลา-สเกล สามารถลดอัตราการสุ่มได้ทั้งแกนของเวลาและสเกล โดยในแกนสเกลจะถูกสุ่มแบบลอการิทึม (Logarithm) แล้วเวลาในแต่ละสเกลจึงค่อยถูกสุ่มในอัตราที่ลดลงโดยขึ้นอยู่กับสเกล อัตราการสุ่มนี้จะขึ้นไปตามตาราง ไดนามิค (Dynadic grid)



รูปที่ 2.27 ตารางไดนามิค (Dynadic grid)

โดยมากจะใช้ลอการิทึมฐาน 2 ดังนั้น สเกลที่ถูกคำนวณคือ สเกล 2 , 4 , 8 , 16 , 32 , 64 ส่วนในแกนเวลา (หรือแกนทรานสเลชัน) อัตราการสุ่มจะลดลง 2 เท่าขณะที่สเกลเพิ่มขึ้น 2 เท่า สเกลที่ถูกสุ่มและเวลาที่ถูกรวมแสดงในสมการ (2.10) และ (2.11) ตามลำดับ

$$S = S_0^j \quad \text{โดย } S_0 > 1 \quad (2.10)$$

$$\tau = k S_0^j \tau_0 \quad \text{โดย } \tau_0 > 0 \quad (2.11)$$

สมการคอนทินิวอัสเวฟเล็ททรานสฟอร์ม (2.5) และสมการรีคอนสตรัคชัน (2.6) สามารถเขียนให้อยู่ในรูปสมการคอนทินิวอัสแบบดิสครีตโดยใช้ j และ k ได้จาก

$$\psi_{\tau,s} = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (2.12)$$

แทนด้วยสมการ (2.8) และ (2.9) จะได้

$$\psi_{j,k}(t) = S_0^{-j/2} \psi(S_0^{-j}t - k\tau_0) \quad (2.13)$$

ใช้เวฟเล็ทด้วยสมการ (2.13) ได้สมการคอนทินิวอัสเวฟเล็ททรานสฟอร์มแบบดิสครีตเป็นสมการ (2.14)

$$\Psi_x^{\psi_{j,k}} = \int x(t) \psi_{j,k}(t) dt \quad (2.14)$$

และได้สมการรีคอนสตรัคชันเป็นสมการ (2.15)

$$x(t) = c_\psi \sum_j \sum_k \Psi_x^{\psi_{j,k}} \psi_{j,k}(t) \quad (2.15)$$

ถึงแม้จะคิดคอนทินิวอัสเวฟเล็ททรานสฟอร์มเป็นแบบดิสครีต เพื่อลดการคำนวณลงแล้วก็ตาม ก็ยังใช้เวลาในการคำนวณมาก เนื่องมาจากจุดที่นำมาคำนวณยังมีจุดที่เป็นจุดฟุ่มเฟือยอยู่อีกมาก (Redundancy) จึงมีการใช้ดิสครีตเวฟเล็ททรานสฟอร์มในการคำนวณ เพื่อลดจุดที่คำนวณลงให้พอเพียงกับการวิเคราะห์ห้กลับ การใช้เวลาในการคำนวณก็จะลดลง

2.3.3.4 ดิสครีตเวฟเล็ททรานสฟอร์ม

ดิสครีตเวฟเล็ททรานสฟอร์ม เกิดจากการพัฒนาการโค้ดแบบสับแบนด์ (Subband coding) และการโค้ดแบบพีรามิด (Pyramidal coding) หรือรู้จักกันในชื่อว่าการวิเคราะห์แบบหลายความละเอียด

2.3.3.4.1 การคำนวณดิสครีตเวฟเล็ททรานสฟอร์ม

ดิสครีตเวฟเล็ททรานสฟอร์ม มักจะเกิดจากการสุ่มคอนทินิวอัสเวฟเล็ททรานสฟอร์มที่ใช้ตารางโคเนคติก (คือใช้ $S_0 = 2$ และ $r = 1$) - การทำดิสครีตเวฟเล็ททรานสฟอร์มจะเป็นการนำเอาสัญญาณผ่านโลวพาสฟิลเตอร์และไฮพาสฟิลเตอร์ (Low pass and high pass filter) เพื่อแยกสัญญาณออกเป็นส่วนความถี่ต่ำและความถี่สูงตามลำดับ โดยโลวพาสฟิลเตอร์ คือสเกลลิงฟังก์ชัน (Scaling function) และไฮพาสฟิลเตอร์ คือ เวฟเล็ทฟังก์ชัน (Wavelet function) ซึ่งในกรณีที่เวฟเล็ทไม่ได้แบ่งเป็นเวฟเล็ทฟังก์ชัน (ไฮพาสฟิลเตอร์) และสเกลลิงฟังก์ชัน (โลวพาสฟิลเตอร์) มาตั้งแต่แรก จะต้องนำเวฟเล็ทนี้มาผ่านกระบวนการแยกให้เป็นไฮพาสกับโลวพาสฟิลเตอร์ก่อน แล้วจึงนำไปใช้ในการแปลงสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม เวฟเล็ทที่มีเวฟเล็ทฟังก์ชันและสเกลลิงฟังก์ชันก็เช่น เวฟเล็ทเมเยอร์ (Meyer wavelet) เวฟเล็ทที่ไม่มีเวฟเล็ทฟังก์ชันและสเกลลิงฟังก์ชันต้องนำเวฟเล็ทนั้นไปแยกเพื่อให้ได้โลวพาสฟิลเตอร์และไฮพาสฟิลเตอร์ เช่น เวฟเล็ทควาวิชีส์ ในกรณีที่ฟิลเตอร์เป็นฟิลเตอร์ที่แบ่งครึ่งความถี่ได้พอดี (Halfband filter) ผลลัพธ์ที่ได้ออกมาจากฟิลเตอร์จะมีความถี่เป็นครึ่งหนึ่งของสัญญาณเดิม ดังนั้นจากกฎของในควิชที่จึงสามารถลดอัตราการสุ่มลงได้ครึ่งหนึ่งเช่นกัน

ผลลัพธ์ที่ได้ออกมาจากฟิลเตอร์และลดการสุ่มลงไป 2 เท่าแล้วเป็นดังนี้

$$y_{high}[k] = \sum_n x[n] \cdot g[2k - n] \quad (2.16)$$

$$y_{low}[k] = \sum_n x[n] \cdot h[2k - n] \quad (2.17)$$

เรียก $y_{high}[k]$ ว่า ดีเทลโคเอฟฟิเชียน (Cd : Detail coefficients)

เรียก $y_{low}[k]$ ว่า แอปพรอกซิเมชันโคเอฟฟิเชียน (Ca : Approximation coefficients)

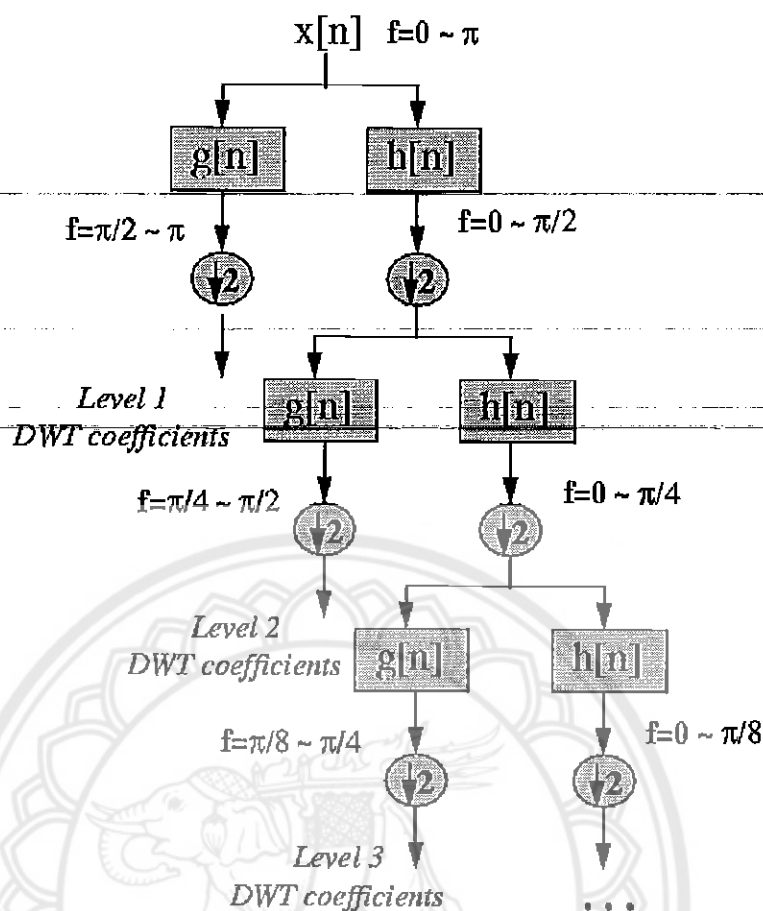
สรุปขั้นตอนการทำดิสครีตเวฟเล็ททรานสฟอร์ม

1. นำสัญญาณมาผ่านไฮพาสฟิลเตอร์ได้แอมพลิจูดเพิ่มขึ้น โคเอฟฟิเชียน แล้วนำมาลดจำนวนจุดตัวอย่าง (Down sampling) ลง 2 เท่า
2. นำสัญญาณเดิมมาผ่านไฮพาสฟิลเตอร์ได้คิเทิลโคเอฟฟิเชียน แล้วนำมาลดจำนวนจุดตัวอย่างลง 2 เท่า
3. นำผลที่ได้จากข้อ 1 มาแยกอีกโดยทำซ้ำแบบข้อ 1 และ 2
4. ทำซ้ำข้อ 3 จะได้ข้อมูล (ความถี่) ที่ต้องการ หรือ ทำซ้ำงานจำนวนจุดตัวอย่างเหลือ 1 จุด

การทำตามขั้นตอนที่ 1 และ 2 จะทำให้ความละเอียดด้านเวลา (Time resolution) ลดลง 2 เท่า เนื่องจากจำนวนตัวอย่างที่สุ่มลดลง 2 เท่า ขณะที่ความละเอียดด้านความถี่ (Frequency resolution) เพิ่มขึ้น 2 เท่า เนื่องจากความถี่มีความคลุมเครืออย่างน้อยลง 2 เท่า เมื่อทำตามขั้นตอนที่ 1 และ 2 ซ้ำต่อไป จะแยกสัญญาณออกได้เป็นชั้น ๆ โดยในแต่ละชั้นจะมีความละเอียดด้านเวลาลดลง 2 เท่า และความละเอียดด้านความถี่เพิ่มขึ้น 2 เท่า



รูปที่ 2.28 แสดงการแปลงสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์มใน 1 ชั้น



รูปที่ 2.29 แสดงการแปลงสัญญาณให้เป็นดิสครีตเวฟเลตทรานสฟอร์มซ้ำหลายชั้น

จากรูปที่ 2.29 $x[n]$ คือสัญญาณเดิม

$h[n]$, $g[n]$ คือ โลว์พาสและไฮพาสฟิลเตอร์

f คือ ช่องทางการส่งข้อมูล (Bandwidth) ในแต่ละชั้น

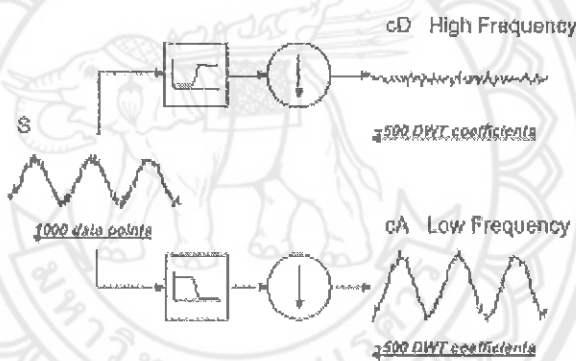
เริ่มแรก $x[n]$ มีจุดตัวอย่างอยู่ 512 จุด ความถี่ $0-\pi$ radians/sec เมื่อผ่านฟิลเตอร์และทำการลดจำนวนจุดตัวอย่างลงครึ่งหนึ่งแล้ว จำนวนจุดตัวอย่างจะเหลือ 256 ในการแยกชั้นที่ 1 นี้ จะมีความละเอียดด้านเวลาลดลง 2 เท่าและมีความละเอียดด้านความถี่เพิ่มขึ้น 2 เท่าจากสัญญาณเดิม จากนั้นนำเอาส่วนความถี่ต่ำ (แอมพลิจูดโคเอฟฟิเชียน) มาแยกต่อไป เมื่อแยกชั้นที่ 2 เสร็จ ความละเอียดด้านเวลาลดลง 4 เท่า และความละเอียดด้านความถี่จะเพิ่มขึ้น 4 เท่าจากสัญญาณเดิม สัญญาณ $x[n]$ จากรูปที่ 2.28 นี้จะแยกได้ทั้งหมด 8 ชั้น ซึ่งจะเหลือตัวอย่างที่สุ่ม 2 จุด

ดังนั้นความยาวของสัญญาณจะมีผลต่อจำนวนชั้นในการแยกด้วย เช่น สัญญาณยาว 1024 จุด จะสามารถแยกสัญญาณนี้ได้ถึง 10 ชั้น ความแตกต่างของฟูเรียร์ทรานสฟอร์มกับคิสิกิตเวฟเล็ดทรานสฟอร์ม คือในคิสิกิตเวฟเล็ดทรานสฟอร์ม ข้อมูลทางด้านเวลาจะไม่สูญเสียไป

- ถ้าข้อมูลหลักของสัญญาณอยู่ที่ความถี่สูง ก็จะมีผลความละเอียดทางด้านเวลามากกว่า เนื่องจากที่ความถี่สูง ๆ จำนวนจุดตัวอย่างยังไม่ถูกลดจำนวนลงไปมากนัก
- ถ้าข้อมูลหลักของสัญญาณอยู่ที่ความถี่ต่ำ ความละเอียดทางด้านเวลาจะไม่ละเอียด เนื่องจากจำนวนจุดตัวอย่างถูกลดให้เหลือน้อยลง

จึงสามารถสรุปได้ดังนี้

- คิสิกิตเวฟเล็ดทรานสฟอร์มที่ความถี่สูงจะให้ความละเอียดด้านเวลาดี
- คิสิกิตเวฟเล็ดทรานสฟอร์มที่ความถี่ต่ำจะให้ความละเอียดด้านความถี่ดี ซึ่งเหมือนกับคอนทินิวัสเวฟเล็ดทรานสฟอร์มและการวิเคราะห์แบบหลายความละเอียด



รูปที่ 2.30 แสดงสัญญาณและคิสิกิตเวฟเล็ดทรานสฟอร์ม โคอэффициเซียนของสัญญาณ

2.3.3.4.2 การแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม

(Reconstruction หรือ Synthesis)

ฟิลเตอร์ในสมการเวฟเล็ททรานสฟอร์ม มีความสัมพันธ์กันดังนี้

$$g[L - 1 - n] = (-1)^n \cdot h[n] \quad (2.18)$$

$g[n]$ เป็น ไฮพาส (High Pass)

$h[n]$ เป็น โลวพาส (Low Pass)

L เป็น ความยาวฟิลเตอร์ (นับเป็นจำนวนจุด)

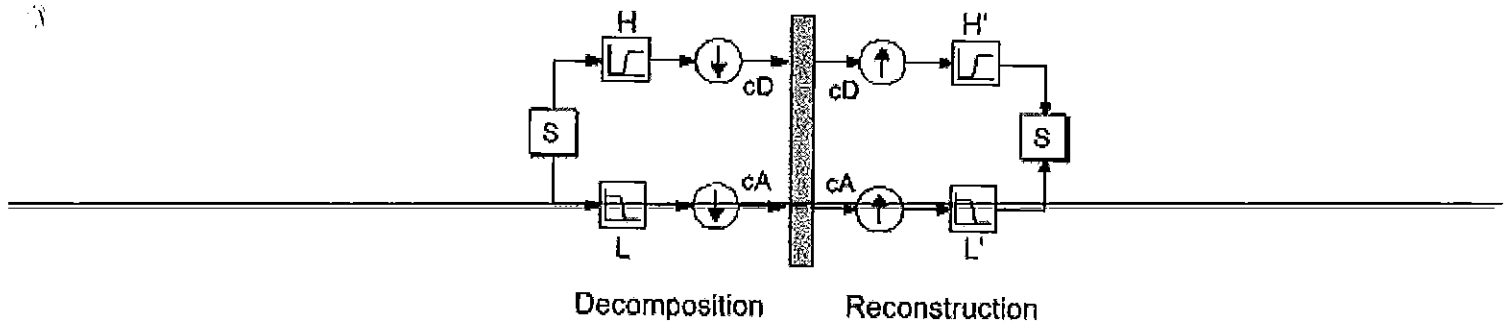
ฟิลเตอร์ที่มีความสัมพันธ์กันตามสมการ (2.18) เรียกว่า ควอดราเจอร์ มิลเรอร์ ฟิลเตอร์ (QMF : Quadrature Mirror Filters)

ในการทำดิสครีตเวฟเล็ททรานสฟอร์ม ให้กลับเป็นสัญญาณเดิมหรือที่เรียกว่ารีคอนสตรัคชันนั้น จะทำย้อนกลับกับตอนที่ทำสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม โดยการรีคอนสตรัคชันจะนำดิสครีตเวฟเล็ททรานสฟอร์มมาเพิ่มจำนวนจุดตัวอย่างขึ้น 2 เท่า (Up sampling) แล้วผ่านไฮพาสฟิลเตอร์หรือโลวพาสฟิลเตอร์ โดยถ้าเป็นแอมพลิจูดซิมเพิลโคเอฟิเชียนจะนำมาผ่านโลวพาสฟิลเตอร์ และถ้าเป็นดิทเทิล โคเอฟิเชียนจะนำมาผ่านไฮพาสฟิลเตอร์ แล้วนำทั้ง 2 ส่วนมาบวกกัน

ไฮพาสฟิลเตอร์และโลวพาสฟิลเตอร์ในการรีคอนสตรัคชันนั้นจะเป็นฟิลเตอร์ 2 ตัวในฟิลเตอร์ 4 ตัว ซึ่งฟิลเตอร์ทั้ง 4 ตัวนี้มีคุณสมบัติเป็นควอดราเจอร์ มิลเรอร์ ฟิลเตอร์

ฟิลเตอร์ทั้ง 4 ตัวที่มีคุณสมบัติเป็นควอดราเจอร์ มิลเรอร์ ฟิลเตอร์ในเรื่องดิสครีตเวฟเล็ททรานสฟอร์มและการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม คือ

1. โลวพาสฟิลเตอร์ในการทำสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม $h[n]$
2. ไฮพาสฟิลเตอร์ในการทำสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม $g[n]$
3. โลวพาสฟิลเตอร์ในการรีคอนสตรัคชันดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม $h'[n]$
4. ไฮพาสฟิลเตอร์ในการรีคอนสตรัคชันดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม $g'[n]$



รูปที่ 2.31 แสดงฟิลเตอร์ทั้ง 4 ตัว ในการทำดิสครีตเวฟเล็ททรานสฟอร์มและการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม

ในกรณีที่ฟิลเตอร์ทั้ง 4 เป็นฮาล์ฟแบนฟิลเตอร์ ฟิลเตอร์ทั้ง 4 คือ

- ฟิลเตอร์ในการทำดิสครีตเวฟเล็ททรานสฟอร์ม $h[2k-n]$, $g[2k-n]$
- ฟิลเตอร์ในการรีคอนสตรัคชัน $h[-n+2k]$, $g[-n+2k]$

จะเห็นว่าในกรณีฮาล์ฟแบนด์ฟิลเตอร์ (Half band filter) ฟิลเตอร์ที่ใช้ในการทำดิสครีตเวฟเล็ททรานสฟอร์ม เหมือนกันกับฟิลเตอร์ในการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมเกือบทุกอย่างแตกต่างกันตรงที่สลับเวลากันเท่านั้น ซึ่งสมการการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมจากฮาล์ฟแบนด์ฟิลเตอร์ข้างต้น สามารถเขียนได้ดังนี้

$$x[n] = \sum_{k=-\infty}^{\infty} (y_{high}[k] \cdot g[-n+2k]) + (y_{low}[k] \cdot h[-n+2k]) \tag{2.19}$$

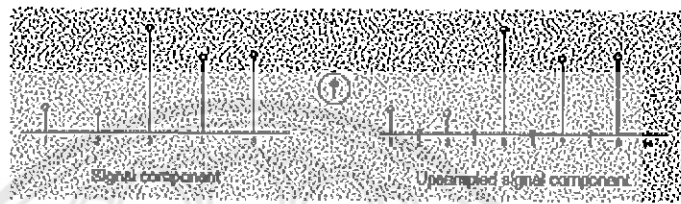
สมการ (2.19) ใช้ในการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม ไปทีละขั้น

ถ้าฟิลเตอร์ไม่เป็นฮาล์ฟแบนด์ฟิลเตอร์ในอุดมคติ (Ideal Halfband Filter) จะไม่สามารถทำการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมที่สมบูรณ์ได้ แต่ถึงแม้จะไม่สามารถหาฟิลเตอร์ในอุดมคติก็ยังสามารถหาฟิลเตอร์ที่ทำให้การแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิมสมบูรณ์ภายใต้เงื่อนไขได้ เนื่องจากการลดจำนวนจุดตัวอย่างในการทำสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม ก่อให้เกิดความผิดเพี้ยนของสัญญาณที่เรียกว่า แอลเลียซซิ่ง (Aliasing) ปัญหานี้สามารถขจัดได้โดยใช้ฟิลเตอร์ที่เหมาะสมกรองสัญญาณ ทั้งในการแปลงสัญญาณให้เป็นดิสครีตเวฟเล็ททรานสฟอร์ม และการแปลงดิสครีตเวฟเล็ททรานสฟอร์มกลับเป็นสัญญาณเดิม ฟิลเตอร์ทั้งในทางไปและกลับนี้จะมีความสัมพันธ์และคล้ายคลึงกันมาก แต่ไม่

เหมือนกันซึ่งฟิลเตอร์นี้ได้รับการค้นพบและพัฒนาโดย อินกริด ดาเวบชีส์ (Ingrid Daubechies) และ รู้จักกันในชื่อ เวฟเล็ตดาเวบชีส์ (Wavelet Daubechies)

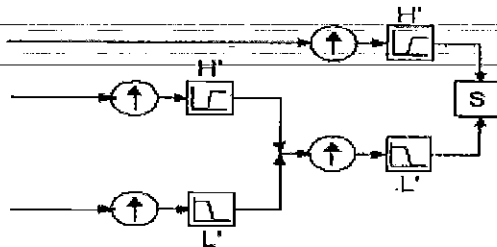
การแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมสามารถแบ่งได้เป็น 2 วิธี คือ
วิธีที่ 1 ทำการแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมทีละขั้น

1. ทำการเพิ่มจำนวนจุดตัวอย่างขึ้น 2 เท่าโดยใช้วิธีแทรกจุด 0 ลงไประหว่างจุดตัวอย่างเดิมของดิสครีตเวฟเล็ตทรานสฟอร์ม เพื่อเพิ่มความยาวของสัญญาณให้ยาวขึ้น 2 เท่า ดังรูป



รูปที่ 2.32 แสดงดิสครีตเวฟเล็ตทรานสฟอร์ม และการเพิ่มจำนวนจุดตัวอย่าง โดยการแทรกจุด 0 ลงในดิสครีตเวฟเล็ตทรานสฟอร์มเดิม

2. นำดีเทิล โคเอฟฟิเชียนที่ผ่านการเพิ่มจุดตัวอย่างขึ้น 2 แล้วมาผ่านไฮพาสฟิลเตอร์
3. นำแอมพลิจูดของโคเอฟฟิเชียนที่ผ่านการเพิ่มจุดตัวอย่างขึ้น 2 เท่า แล้วมาผ่านโลพพาสฟิลเตอร์
4. นำผลจากขั้นตอน 2 และ 3 มาบวกกัน ผลลัพธ์ที่ได้จะเป็นโคเอฟฟิเชียนของขั้นที่อยู่ก่อนหน้าในขั้นตอนนี้ทำดิสครีตเวฟเล็ตทรานสฟอร์มเรานำแอมพลิจูดของโคเอฟฟิเชียนมาแยก ดังนั้นผลลัพธ์ที่ได้จากขั้นตอน 4 นี้คือแอมพลิจูดของโคเอฟฟิเชียนของขั้นที่อยู่ก่อนหน้า
5. นำแอมพลิจูดของโคเอฟฟิเชียนที่หามาได้จากขั้นตอน 4 และดีเทิล โคเอฟฟิเชียน ที่อยู่ในขั้นเดียวกันกับแอมพลิจูดของโคเอฟฟิเชียนนั้น ไปผ่านขั้นตอน 1, 2, 3 และ 4 ซ้ำ
6. ทำขั้นตอน 5 ซ้ำจนได้สัญญาณเดิม



รูปที่ 2.33 การแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิม ด้วยวิธีหลายขั้น

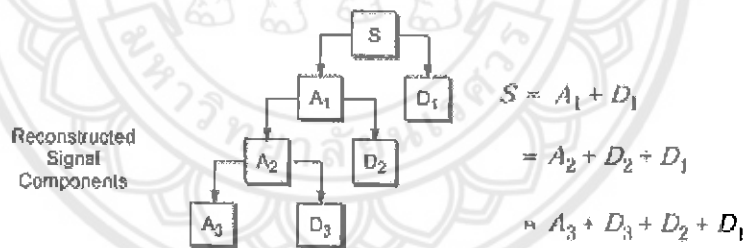
วิธีที่ 2 การแปลงคิสคริตเวฟเล็ดทรานสฟอร์มกลับเป็นสัญญาณเดิม โดยนำแอฟพรอกซิเมชันและดีเทลโคเอฟฟิเชียนที่มีอยู่ก่อนแล้วทั้งหมด จากนั้นนำผลลัพธ์ที่ได้มารวมเป็นสัญญาณ

- การนำแอฟพรอกซิเมชันชั้นโคเอฟฟิเชียนมาเพิ่มจำนวนจุดตัวอย่างขึ้น 2 เท่าแล้วผ่านโลวพาสฟิลเตอร์ และใช้เวกเตอร์ของ 0 มาเพิ่มจำนวนจุดตัวอย่างขึ้น 2 เท่าแล้วผ่านไฮพาสฟิลเตอร์ แทนที่จะใช้ดีเทลโคเอฟฟิเชียน จากนั้นนำผลที่ได้มารวมกันเกิดผลลัพธ์เป็นรีคอนสตรัคชันแอฟพรอกซิเมชัน A1 (Reconstructed approximation A1)

- นำเวกเตอร์ของ 0 มาเพิ่มจำนวนจุดตัวอย่างขึ้น 2 เท่าแล้วผ่านโลวพาสฟิลเตอร์ แทนที่จะใช้แอฟพรอกซิเมชันชั้นโคเอฟฟิเชียน ได้ผลลัพธ์เป็นรีคอนสตรัคชันดีเทล D1 (Reconstructed detail D1) นำ $A1 + D1 = S$

การทำการแปลงคิสคริตเวฟเล็ดทรานสฟอร์มกลับเป็นสัญญาณเดิมด้วยวิธีที่ 2 สามารถใช้กับคิสคริตเวฟเล็ดทรานสฟอร์มหลายชั้นได้เช่นกัน โดยการแปลงคิสคริตเวฟเล็ดทรานสฟอร์มกลับเป็นสัญญาณเดิม โดยรีคอนสตรัคชันโคเอฟฟิเชียนที่แยกออกมาทุกตัวในทุกระดับก่อนแล้วจึงนำมาบวกกันเป็นสัญญาณ จากรูปที่ 2.33 จะได้สัญญาณตามสมการ

$$\begin{aligned} S &= A_1 + D_1 \\ &= A_2 + D_2 + D_1 \\ &= A_3 + D_3 + D_2 + D_1 \end{aligned}$$



รูปที่ 2.34 การแปลงคิสคริตเวฟเล็ดทรานสฟอร์มกลับเป็นสัญญาณเดิมด้วยวิธีที่ 2 หลายชั้น

2.3.3.4.3 การวิเคราะห์เวฟเล็ตแพคเกจจิง (Wavelet package analysis)

เวฟเล็ตแพคเกจจิง เป็นการวิเคราะห์สัญญาณแยกทั้งทางด้านแอมพลิจูดและเฟสโคเอฟฟิเชียน และดีเทิลโคเอฟฟิเชียน แตกต่างจากเดิมที่แยกแต่ทางด้านแอมพลิจูดและเฟสโคเอฟฟิเชียน การวิเคราะห์เวฟเล็ตแพคเกจจิงนี้ ทำให้สามารถแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมได้ด้วยวิธีการที่หลากหลายกว่าเดิม โดยวิธีในการแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมจากเวฟเล็ตแพคเกจจิงจะเป็น 2^n วิธีที่แตกต่างกัน เช่นจากรูปที่ 2.34 สามารถแทน S ได้ด้วย $S = A1 + A1A2D3 + D1A1D3 + D1D2$ ซึ่งจะเห็นว่า S มีได้หลากหลายกว่าเดิม

2.3.3.4.4 การนำดิสครีตเวฟเล็ตทรานสฟอร์มไปใช้งาน

แบ่งเป็น 3 ขั้นตอน คือ

1. นำสัญญาณมาแปลงเป็นดิสครีตเวฟเล็ตทรานสฟอร์ม โคเอฟฟิเชียน
2. นำดิสครีตเวฟเล็ตทรานสฟอร์ม โคเอฟฟิเชียนนี้ไปปรับปรุง เช่น ตัดโคเอฟฟิเชียนบางตัวออก
3. นำดิสครีตเวฟเล็ตทรานสฟอร์ม โคเอฟฟิเชียนที่ปรับปรุงแล้ว มารีคอนสตรัคท์กลับเป็นสัญญาณจะได้สัญญาณที่ปรับปรุงแล้ว

ดิสครีตเวฟเล็ตทรานสฟอร์มสามารถใช้ปรับปรุงสัญญาณ หรือข้อมูลได้หลายอย่าง เช่น ลดสัญญาณรบกวน หรือทำการบีบอัดข้อมูล เมื่อทราบว่าข้อมูลหลักของสัญญาณอยู่ที่ ความถี่ใดก็สามารถตัดความถี่ที่ไม่จำเป็นได้ จะเป็นการลดขนาดสัญญาณหรือภาพโดย ไม่ทำให้ข้อมูลเสียหาย และเมื่อจะทำการแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมก็เดิม 0 แทนในโคเอฟฟิเชียนที่ตัดทิ้งไป เป็นต้น

2.3.3.5 ความแตกต่างและการเลือกใช้คอนทินิวัสเวฟเล็ตทรานสฟอร์มและดิสครีตเวฟเล็ตทรานสฟอร์ม

การใช้ดิสครีตเวฟเล็ตทรานสฟอร์มหรือคอนทินิวัสเวฟเล็ตทรานสฟอร์ม ให้พิจารณาจากคำถาม เช่น ผู้ใช้ต้องการรู้ค่าทุกค่าของทรานสฟอร์มเพื่อจะนำไปแปลงดิสครีตเวฟเล็ตทรานสฟอร์มกลับเป็นสัญญาณเดิมอย่างถูกต้องแม่นยำหรือไม่ ผู้ใช้จะทำการวิเคราะห์แบบไม่มีจุดที่ฟุ่มเฟือย (Non-Redundant analysis) หรือไม่

เมื่อสัญญาณมีพลังงานจำกัด การรีคอนสตรัคชันสามารถทำได้โดยง่าย ไม่มีความจำเป็นต้องใช้ค่าทรานสฟอร์มทุกค่ามาทำการรีคอนสตรัคชัน ในกรณีนี้ใช้ดิสครีตเวฟเล็ตทรานสฟอร์ม ก็พอเพียงต่อความต้องการแล้ว ขณะที่การใช้คอนทินิวัสเวฟเล็ตทรานสฟอร์มจะเป็นการฟุ่มเฟือย

เมื่อสัญญาณถูกบันทึกในเวลาต่อเนื่อง (Continuous time) หรือถูกบันทึกในตารางเวลา (Time grid) ที่ละเอียดมาก ๆ สามารถใช้ได้ทั้งดิสครีตเวฟเล็ตทรานสฟอร์มและคอนทินิวอัลเวฟเล็ตทรานสฟอร์ม โดยทั้งดิสครีตเวฟเล็ตทรานสฟอร์มและคอนทินิวอัลเวฟเล็ตทรานสฟอร์ม ต่างก็มีข้อดีดังนี้

การวิเคราะห์แบบดิสครีต ทำให้ประหยัดเนื้อที่ในการโค้ด (Space-saving coding) และพอเพียงต่อการสังเคราะห์กลับ นอกจากนี้การวิเคราะห์ด้วยดิสครีตเวฟเล็ตทรานสฟอร์มยังแยกส่วนที่เป็นความถี่ต่ำและความถี่สูงออกจากกัน ทำให้ง่ายต่อการวิเคราะห์งานบางอย่างอีกด้วย เช่น การกำจัดสัญญาณรบกวน การบีบอัดข้อมูล เป็นต้น

การวิเคราะห์แบบคอนทินิวอัล มักจะทำให้ดีความง่ายกว่าการวิเคราะห์แบบดิสครีต เนื่องจากความนุ่มนวลในการคำนวณแบบคอนทินิวอัล ช่วยเสริมลักษณะของทรานสฟอร์มให้เห็นข้อมูลทั้งหมดได้ชัดเจนยิ่งขึ้น ซึ่งจะใช้ได้กับข้อมูลที่มีความละเอียดมาก ๆ การวิเคราะห์แบบคอนทินิวอัลนี้ จะมีข้อดีในการดูและตีความง่าย แต่มีข้อเสียที่ไม่ประหยัดเนื้อที่



2.4 ทฤษฎีพื้นฐานของแบบจำลองฮิดเดน มาร์คอฟ (Hidden Markov Model)

ผลที่ได้จากกระบวนการทำงาน (Process) ทั่วไปจะมีลักษณะเฉพาะเหมือนสัญญาณต่าง ๆ ซึ่งสัญญาณเหล่านี้จะเป็นดิสครีตคอนทินิวอัส (Discrete continuous) หรือจะเป็นสัญญาณที่ประกอบด้วยลักษณะเฉพาะต่าง ๆ (Pure-signal), สัญญาณที่ถูกรบกวนโดยแหล่งกำเนิดอื่น ๆ, ผลจากการบิดเบือนของการส่ง (Transmission distortion) หรือเกิดการสะท้อนกลับ เป็นต้น สัญญาณต่าง ๆ เหล่านี้จะมีลักษณะเฉพาะเป็นของตัวเองเสมอ

ปัญหาหนึ่งที่น่าสนใจก็คือลักษณะเฉพาะของสัญญาณในเทอมของแบบจำลองสัญญาณ

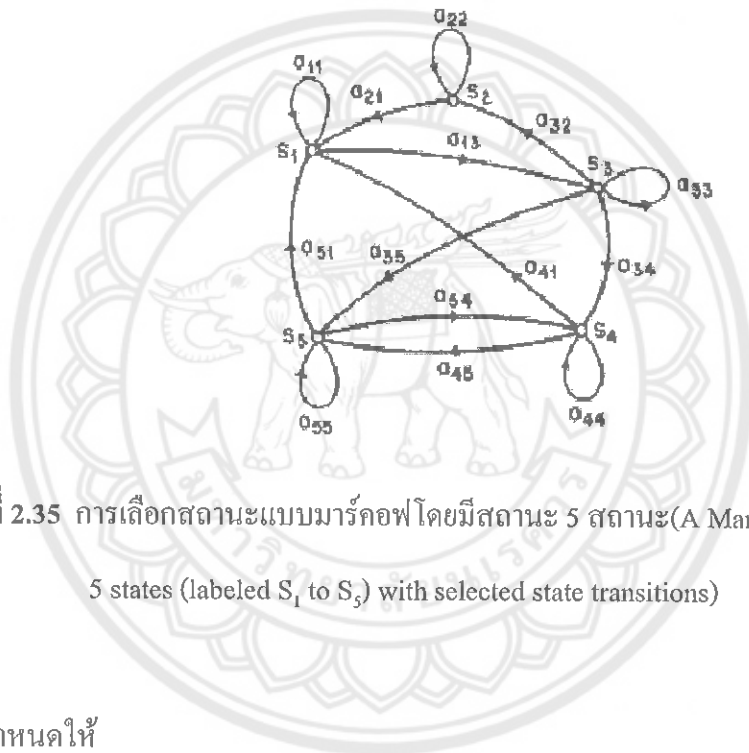
(Signal model) ซึ่งมีอยู่หลายเหตุผลที่อธิบายว่าเหตุใดจึงมีผู้สนใจในการประยุกต์แบบจำลองสัญญาณนี้ เหตุผลหนึ่งก็คือแบบจำลองสัญญาณนั้นสามารถบอกรากฐานสำคัญเพื่อใช้สมมติรูปร่างของระบบ ตัวอย่างเช่น ถ้าหากต้องการปรับปรุงสัญญาณเสียงพูดที่ถูกรบกวนจากคลื่นรบกวน (Noise) และเกิดการบิดเบือนของการส่ง สามารถใช้แบบจำลองสัญญาณในการออกแบบระบบเพื่อกำจัดคลื่นรบกวนและลบเสียงการบิดเบือนของการส่ง เหตุผลที่สองอธิบายว่าเหตุใดแบบจำลองสัญญาณจึงมีความสำคัญ นั่นคือแบบจำลองสัญญาณทำให้ทราบถึงแหล่งกำเนิดสัญญาณต่าง ๆ คุณสมบัตินี้มีความสำคัญอย่างมากเพราะต้นทุนของการสร้างสัญญาณจากแหล่งกำเนิดจริงนั้นมีค่าสูง ในกรณีของโครงสร้างสัญญาณที่ดีสามารถจำลองแหล่งกำเนิดและเรียนรู้ได้มากเท่าที่จะเป็นไปได้จากการจำลองแหล่งกำเนิด เหตุผลที่กล่าวมาทำให้ทราบว่าเหตุใดแบบจำลองสัญญาณจึงมีความสำคัญ เพราะให้ผลการทดลองที่ดีและสามารถทำให้เข้าใจระบบที่ทดลอง เช่น ระบบการคาดเดา (Prediction system) ระบบการรู้จำรู้จำ (Recognition system) ระบบการหาเอกลักษณ์ (Identification system) เป็นต้น

เหตุผลเหล่านี้เป็นทางเลือกที่เป็นไปได้หลายทางที่นำมาใช้สำหรับเลือกชนิดของแบบจำลองสัญญาณเพื่อใช้หาลักษณะเฉพาะในคุณสมบัติของสัญญาณ สามารถแบ่งชนิดของแบบจำลองสัญญาณได้เป็น 2 ประเภทคือ แบบจำลองเชิงกำหนด (Deterministic model) และแบบจำลองเชิงสถิติ (Statistical model) ประโยชน์ทั่วไปที่ได้จากแบบจำลองเชิงกำหนด คือคุณสมบัติเฉพาะบางอย่างของสัญญาณ เช่น สัญญาณนั้นเป็นสัญญาณคลื่นไซน์ (Sine wave) หรือเป็นผลรวมของเอ็กซ์โปเนนเชียล เป็นต้น ในกรณีนี้ต้องการทราบรายละเอียดต่าง ๆ ของแบบจำลองสัญญาณ เช่น ความสูงของคลื่น (Amplitude) ความถี่ (Frequency) เป็นต้น อีกหนึ่งประเภทของแบบจำลองสัญญาณเป็นกลุ่มของแบบจำลองเชิงสถิติ ซึ่งหาลักษณะเฉพาะของสัญญาณจากคุณสมบัติเชิงสถิติ เช่น กระบวนการเกาส์เซียน (Gaussian process), กระบวนการปัวซอง (Poisson process), กระบวนการมาร์คอฟ (Markov process) และกระบวนการฮิดเดนมาร์คอฟ (Hidden Markov process) เป็นต้น

ในโครงการนี้จะกล่าวถึงพื้นฐานสำคัญของทฤษฎีแบบจำลองฮิดเดน มาร์คอฟ (HMM : Hidden Markov Model) ปัญหาพื้นฐานของทฤษฎีนี้ วิธีการคำนวณหาค่าพารามิเตอร์การแก้ปัญหา และการปรับปรุงแบบจำลองให้ดีขึ้น รวมถึงความสามารถในการนำไปประยุกต์ใช้

2.4.1 กระบวนการมาร์คอฟ (Markov Process)

พิจารณาระบบที่อธิบายถึงช่วงเวลาหนึ่งของกลุ่มสถานะที่แน่นอนจำนวน N สถานะ คือ S_1 ถึง S_N ที่แสดงให้เห็นดังรูปที่ 1 โดยกำหนดให้ $N=5$ และค่า a_{ij} เป็นค่าความน่าจะเป็นในการเปลี่ยนสถานะหนึ่งไปยังอีกสถานะหนึ่ง (โดยที่ i เป็นสถานะต้นและ j เป็นสถานะปลาย)



รูปที่ 2.35 การเลือกสถานะแบบมาร์คอฟโดยมีสถานะ 5 สถานะ (A Markov chain with 5 states (labeled S_1 to S_5) with selected state transitions)

กำหนดให้

t คือลำดับเวลาหนึ่งของการเปลี่ยนสถานะ (Discrete time)

q_t เป็นสถานะปัจจุบัน ณ เวลา t

$$P[q_t = S_j | q_{t-1}, q_{t-2} = S_k, \dots] = P[q_t = S_j | q_{t-1} = S_i] \tag{2.20}$$

พิจารณาสมการที่ (2.20) นี้ จะเห็นว่าสมการด้านขวามือจะไม่ขึ้นอยู่กับเวลา ดังนั้นเราจะได้ค่าความน่าจะเป็นของ a_{ij} คือ

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N \tag{2.21}$$

$$a_{ij} \geq 0 \quad (2.22a)$$

$$\sum_{j=1}^N a_{ij} = 1 \quad (2.22b)$$

พิจารณาตัวอย่างแบบจำลองมาร์คอฟของสภาพอากาศที่มี 3 สถานะคือ ฝนตก เมฆมาก และท้องฟ้าโปร่ง โดยจะสมมติสภาพอากาศจะเป็นอย่างไรในหนึ่งวัน กำหนดให้แต่ละสถานะเป็นดังนี้ สถานะที่ 1: ฝนตก

สถานะที่ 2: เมฆมาก

สถานะที่ 3: ท้องฟ้าโปร่ง

จะกำหนดค่าความน่าจะเป็นของการเปลี่ยนสถานะซึ่งแทนด้วยเมทริกซ์ (Matrix) A

$$A = \{a_{ij}\} = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

สมมติให้สภาพอากาศในวันที่ 1 เป็นท้องฟ้าโปร่ง (สถานะที่ 3) แล้วจะถามว่าค่าความน่าจะเป็นที่สภาพอากาศอีก 7 วันจะมีสภาพอากาศเป็น “ท้องฟ้าโปร่ง-ท้องฟ้าโปร่ง-ฝนตก-ฝนตก-ท้องฟ้าโปร่ง-เมฆมาก-ท้องฟ้าโปร่ง” มีค่าเท่าไร จะกำหนดให้ลำดับข้อมูลการเปลี่ยนสถานะเหล่านี้แทนด้วย O โดยที่ $O = \{S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3\}$ ซึ่งตรงกับวันที่ $t = 1, 2, \dots, 8$ จะได้ค่าความน่าที่สภาพอากาศเป็นไป O คือ

$$P(O | Model) = P[S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3 | Model]$$

$$= P[S_3] \cdot P[S_3 | S_3] \cdot P[S_3 | S_3] \cdot P[S_1 | S_3] \cdot P[S_1 | S_3] \cdot P[S_3 | S_1] \cdot P[S_2 | S_3] \cdot P[S_3 | S_2]$$

$$= \pi_3 \cdot a_{33} \cdot a_{33} \cdot a_{31} \cdot a_{11} \cdot a_{13} \cdot a_{32} \cdot a_{23}$$

$$= 1 \cdot (0.8)(0.8)(0.1)(0.4)(0.3)(0.1)(0.2)$$

$$= 1.536 \times 10^{-4}$$

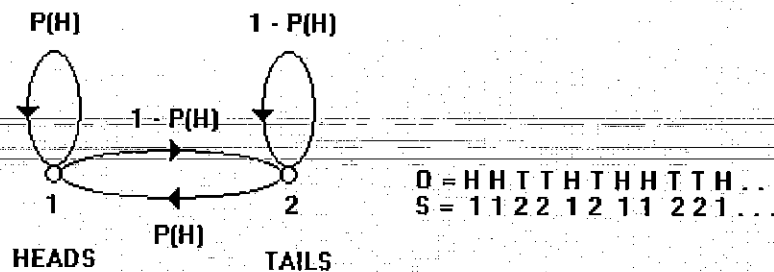
กำหนดค่า π เป็นค่าความน่าจะเป็นตั้งต้นของสถานะโดยที่

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N \tag{2.23}$$

2.4.2 แบบจำลองฮิดเดน มาร์คอฟ (Hidden Markov model)

พิจารณาแบบจำลองมาร์คอฟ (Markov model) ว่าสามารถสังเกตการณ์ที่เกิดขึ้นในการเปลี่ยนแปลงสถานะได้ ซึ่งเหตุผลนี้ที่ทำให้แบบจำลองมาร์คอฟ มีข้อจำกัดในการนำไปใช้มากขึ้นไป ในส่วนนี้ได้เพิ่มเติมแนวความคิดของแบบจำลองมาร์คอฟ รวมเข้ากับกรณีข้อมูลที่ข้อมูลของสถานะนั้นเป็นฟังก์ชันเชิงความน่าจะเป็น (Probabilistic function) ซึ่งสามารถเรียกแบบจำลองนี้ว่าแบบจำลองฮิดเดน มาร์คอฟ คุณสมบัติของแบบจำลองนี้คือ ไม่สามารถสังเกตเหตุการณ์ที่เกิดขึ้นในกระบวนการทำงาน เพื่อให้สามารถเข้าใจได้ง่ายจึงพิจารณาแบบจำลองของการทดลองทอยลูกเต๋า (High Low) ในการทอยลูกเต๋าค่ะแต่ละครั้งลูกเต๋าก็จะถูกเขย่าในถ้วยซึ่งไม่สามารถสังเกตเห็นได้ว่าข้างในนั้นเกิดอะไรขึ้น สิ่งเดียวที่สังเกตได้ก็คือ ผลลัพธ์เมื่อเปิดฝาด้วยออก

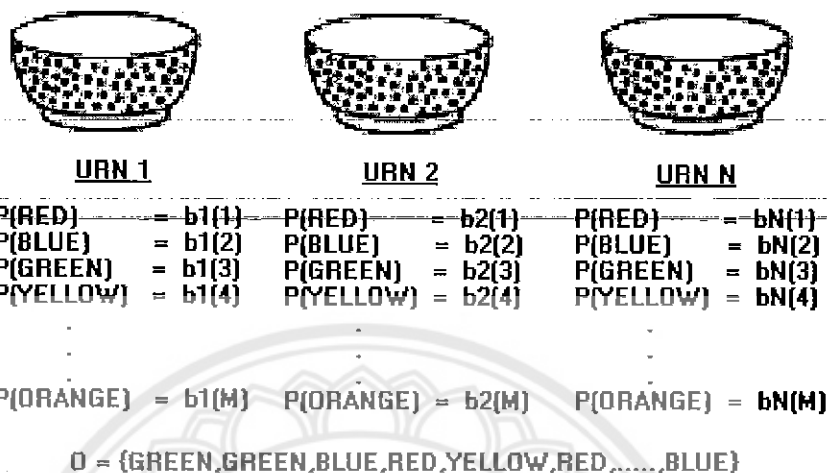
ปัญหาที่น่าสนใจในการสร้างแบบจำลองฮิดเดน มาร์คอฟของการทอยเหรียญ (Coin Toss Model) ขึ้นมาเพื่อบ่งบอกลำดับของผลลัพธ์(หัวและก้อย)นั้น ปัญหาแรกคือการตัดสินใจว่าสถานะในแบบจำลองเกี่ยวข้องกับสิ่งใด และการตัดสินใจต่อไปว่าจะต้องมีจำนวนสถานะเท่าใดในแบบจำลอง กรณีการทอยเหรียญ 1 เหรียญ สามารถกำหนดจำนวนสถานะของแบบจำลองได้ 2 สถานะ คือ หัวและก้อย แบบจำลองนี้แสดงในรูปที่ 2.36 ซึ่งแบบจำลองนี้เป็นเพียงแบบจำลองมาร์คอฟ เพราะสามารถสังเกตเหตุการณ์ในการเปลี่ยนแปลงสถานะจากหัวไปก้อยหรือก้อยไปหัวได้ สิ่งที่น่าสนใจอีกประการหนึ่งคือ เมื่อใช้แบบจำลองฮิดเดน มาร์คอฟ เหตุการณ์นี้จะได้แบบจำลองที่มี 2 สถานะ (สถานะของเหรียญที่ถูกวนกลับ(Bias)) และมีตัวแปรที่ไม่รู้ค่าก็คือการวนกลับของเหรียญ



รูปที่ 2.36 ความน่าจะเป็นในการทอยเหรียญ

เพื่อแสดงแนวคิดเพิ่มเติมของแบบจำลองฮิดเดน มาร์คอฟ จะนำแบบจำลองของลูกบอลในถ้วยมาช่วยอธิบาย สมมติให้มีจำนวนถ้วยอยู่ N ใบ และในถ้วยแต่ละใบจะมีลูกบอล M สี ซึ่ง

จำนวนลูกบอลแต่ละสีในแต่ละถ้วยไม่เท่ากัน ในการทดลองจะหยิบลูกบอลขึ้นมาโดยไม่รู้ว่าหยิบมาจากถ้วยใด สิ่งที่จะรู้เพียงอย่างเดียวคือลูกบอลที่หยิบขึ้นมาสีอะไร ซึ่งแบบจำลองนี้ได้แสดงในรูปที่ 2.37



รูปที่ 2.37 แบบจำลองฮิดเดน มาร์คอฟชนิดไม่ต่อเนื่องของการหยิบลูกบอลจากถ้วยโดยมีสถานะ N สถานะ (An N-state urn and ball model which illustrates the general case of a discrete symbol HMM)

เห็นว่าแต่ละสถานะนั้นไม่ใช่สีของลูกบอลแต่เป็นถ้วยแต่ละใบ เมื่อผลลัพธ์ที่ได้เป็นสีของลูกบอล จึงไม่มีทางรู้เลยว่าการเปลี่ยนแปลงสถานะนั้นเป็นไปในรูปแบบใด เพราะไม่รู้ว่าหยิบลูกบอลมาจากถ้วยใบใด และไม่รู้เลยว่าข้างในแบบจำลองเป็นแบบใดด้วย จากแบบจำลองของลูกบอลในถ้วยให้แนวคิดว่ แบบจำลองฮิดเดน มาร์คอฟคืออะไรและสามารถนำไปประยุกต์ใช้ได้อย่างไร

2.4.3 ส่วนประกอบของ แบบจำลองฮิดเดน มาร์คอฟ

ส่วนประกอบต่าง ๆ ที่ แบบจำลองฮิดเดน มาร์คอฟ ต้องมี

N คือจำนวนของสถานะในแบบจำลอง เช่น จำนวนถ้วยใส่ลูกบอล

M คือจำนวนชนิดของข้อมูล เช่น จำนวนสีของลูกบอล

ค่าการกระจายความน่าจะเป็นของการเปลี่ยนสถานะ ($A = \{a_{ij}\}$) โดยที่

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], \quad 1 \leq i, j \leq N \quad (2.24)$$

ค่าการกระจายความน่าจะเป็นของข้อมูลในสถานะ j ($B = \{b_j(k)\}$) โดยที่

$$b_j(k) = P[V_k \text{ at } t | q_i = S_j], \quad 1 \leq j \leq N$$

$$1 \leq k \leq M \quad (2.25)$$

ค่าเริ่มต้นของสถานะ ($\pi = \{\pi_i\}$) โดยที่

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N \quad (2.26)$$

สามารถบอกได้ว่าค่าต่าง ๆ เหล่านี้เป็นลักษณะเฉพาะของแบบจำลอง ซึ่งค่า N และ M นั้น
รู้ได้จากค่าตัวแปร A และ B ดังนั้นจึงสามารถละไว้ได้ ตัวแปรของแบบจำลองแสดงได้ดังนี้

$$\lambda = (A, B, \pi) \quad (2.27)$$

2.4.4 ปัญหาขั้นพื้นฐานในแบบจำลองฮิดเดน มาร์คอฟ

ปัญหาที่ 1 : การคำนวณหาค่า $P(O | \lambda)$ โดยที่มีลำดับข้อมูลเป็น $O = O_1 O_2 \dots O_T$ และมี
แบบจำลองเป็น $\lambda = (A, B, \pi)$

ปัญหาที่ 2 : การเลือกเส้นทางของลำดับสถานะที่ให้ค่าความเป็นไปได้มากที่สุด
($Q = q_1 q_2 \dots q_T$)

ปัญหาที่ 3 : การปรับค่าของตัวแปรต่าง ๆ ในแบบจำลอง $\lambda = (A, B, \pi)$ เพื่อให้ค่า
 $P(O | \lambda)$ มากที่สุด

ในปัญหาแรกเป็นปัญหาเกี่ยวกับระบบการคำนวณ จะพิจารณาว่าเหตุใดปัญหานี้จึงเป็น
ปัญหาพื้นฐานในแบบจำลองฮิดเดน มาร์คอฟ เนื่องจากแบบจำลองมีลักษณะเป็นหลายมิติ จึงมี
สมการหลายชั้นซึ่งจะมีผลในการนำไปคำนวณโดยใช้คอมพิวเตอร์

ส่วนปัญหาที่ 2 เป็นปัญหาที่ไม่สามารถสังเกตได้ของแบบจำลองฮิดเดน มาร์คอฟ ในการ
หาลำดับของสถานะ เหตุที่ไม่ต้องหาลำดับของสถานะเพราะจะไม่มีทางรู้ได้เลยว่าลำดับของสถานะ
ที่เกิดขึ้นเป็นอย่างไร จึงสมมติขึ้นมาโดยที่ลำดับสถานะนั้นจะให้ค่าความเป็นไปได้มากที่สุด

ส่วนปัญหาที่ 3 เป็นการทำให้แบบจำลองมีความสมบูรณ์มากขึ้น โดยนำชุดข้อมูลแต่ละชุด
ผ่านกระบวนการในแบบจำลองซึ่งจะทำการปรับค่าตัวแปรต่าง ๆ โดยอัตโนมัติ ทำให้มีผลลัพธ์ที่มี
ความแม่นยำเพิ่มขึ้น ซึ่งเรียกกระบวนการนี้ว่าการฝึกฝน ยังมีข้อมูลมากและเป็นข้อมูลที่มีความ
ถูกต้องสูงจะทำให้แบบจำลองมีประสิทธิภาพมากขึ้น

2.4.5 การแก้ปัญหาพื้นฐานในแบบจำลองฮิดเดน มาร์คอฟ

การแก้ปัญหาที่ 1

จะทำการคำนวณหาค่าความน่าจะเป็นของลำดับข้อมูล $O = O_1 O_2 \dots O_T$ จากแบบจำลอง λ จะเขียนแทนค่านี้ด้วย $P(O|\lambda)$ โดยจะกำหนดให้ลำดับของสถานะคือ

$$Q = q_1 q_2 \dots q_T \quad (2.28)$$

ให้ q_1 เป็นสถานะตั้งต้น และค่าความน่าจะเป็นของลำดับข้อมูล O โดยที่มีลำดับสถานะเป็นดังสมการที่ 9 คือ

$$P(O|Q, \lambda) = \prod_{t=1}^T P(O_t | q_t, \lambda) \quad (2.29a)$$

กำหนดให้ลำดับข้อมูลเป็นอิสระต่อกัน จะได้

$$P(O|Q, \lambda) = b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdot \dots \cdot b_{q_T}(O_T) \quad (2.29b)$$

ค่าความน่าจะเป็นของชุดลำดับสถานะ Q สามารถเขียนได้ดังนี้

$$P(Q|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T} \quad (2.30)$$

เพราะฉะนั้นค่า $P(O|\lambda)$ จะสามารถเขียนได้ดังนี้

$$P(O|\lambda) = \sum_{all Q} P(O|Q, \lambda) P(Q|\lambda) \quad (2.31)$$

$$= \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_1}(O_1) a_{q_1 q_2} b_{q_2}(O_2) \dots a_{q_{T-1} q_T} b_{q_T}(O_T) \quad (2.32)$$

เนื่องจากว่าการใช้คอมพิวเตอร์คำนวณหาค่า $P(O|\lambda)$ โดยวิธีนี้ คอมพิวเตอร์จะต้องใช้จำนวนคำสั่งถึง $2T \cdot N^T$ ครั้ง ถ้าสมมติให้แบบจำลองมีสถานะ 5 สถานะ ($N = 5$) และมีชุดลำดับข้อมูล 100 ข้อมูล ($T = 100$) ดังนั้นคอมพิวเตอร์จะต้องให้คำสั่งในการคำนวณเท่ากับ $2 \cdot 100 \cdot 5^{100} \approx 10^{72}$ ครั้ง จึงจำเป็นต้องใช้วิธีอื่น ๆ มาช่วยในการลดจำนวนคำสั่งในการคำนวณหา

ค่า $P(O|\lambda)$ ซึ่งรายงานนี้จะใช้วิธีการลดจำนวนคำสั่งของการคำนวณโดยใช้วิธีฟอร์เวิร์ด โพรซีเยอร์ (Forward Procedure)

วิธีฟอร์เวิร์ด โพรซีเยอร์จะมีตัวแปรเพิ่มขึ้นมาเพื่อจัดระบบสมการใหม่ให้เข้าใจง่ายขึ้น และลดรูปสมการ ซึ่งตัวแปรนี้แทนด้วย

$$a_t(i) = P(O_1 O_2 \dots O_t, q_t = S_i | \lambda) \quad (2.33)$$

วิธีฟอร์เวิร์ด โพรซีเยอร์ประกอบด้วย 3 ส่วน มีดังนี้

1. การกำหนดค่าเริ่มต้น (Initialization)

$$a_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad (2.34)$$

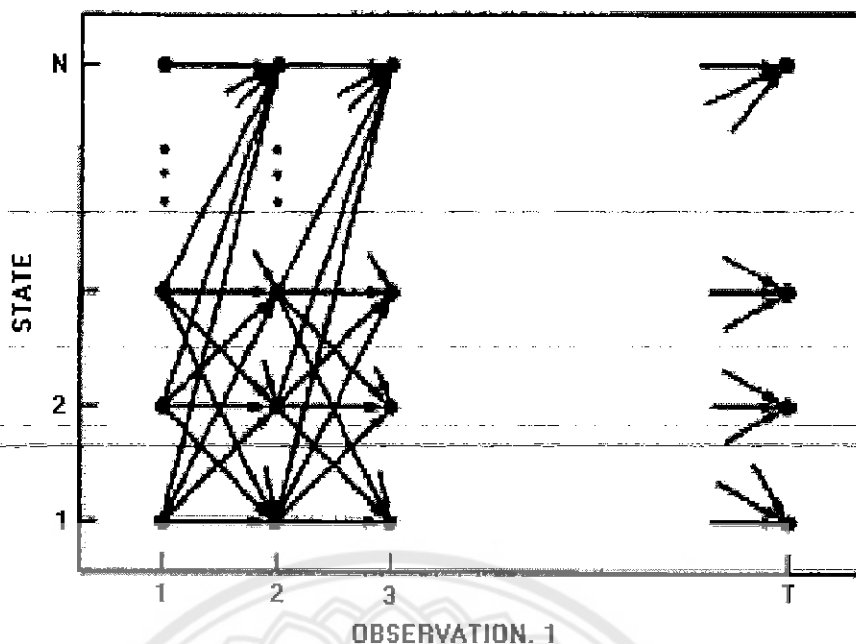
2. การเหนี่ยวนำ (Induction)

$$a_{t+1}(j) = \left[\sum_{i=1}^N a_t(i) a_{ij} \right] b_j(O_{t+1}), \quad \begin{array}{l} 1 \leq t \leq T-1 \\ 1 \leq j \leq N \end{array} \quad (2.35)$$

3. การยกเลิก (Termination)

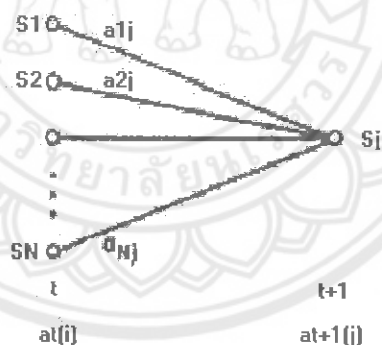
$$P(O|\lambda) = \sum_{i=1}^N a_T(i) \quad (2.36)$$

ข้อ 1 เป็นส่วนของการเริ่มต้นข้อมูลตัวแรก O_1 ของทุก ๆ สถานะในแบบจำลอง และข้อที่ 2 เป็นหัวใจสำคัญของวิธีฟอร์เวิร์ด โพรซีเยอร์ซึ่งเป็นวิธีการคำนวณหาค่า $a_t(i)$ ของข้อมูลถัดจาก O_t ทั้งหมดไปจนถึงตัวสุดท้าย เมื่อการคำนวณดำเนินไปจน $t = T-1$ ก็เป็นอันเสร็จสิ้นของการคำนวณโดยวิธีฟอร์เวิร์ด โพรซีเยอร์โดยข้อที่ 3 จะเป็นสมการที่เสร็จสมบูรณ์แล้วของค่า $P(O|\lambda)$ ซึ่งเท่ากับผลรวมของค่า $a_T(i)$ ของข้อมูลตัวสุดท้ายของทุก ๆ สถานะในแบบจำลองสามารถแสดงลักษณะของการคำนวณโดยวิธีฟอร์เวิร์ด โพรซีเยอร์ด้วยกราฟได้ดังรูปที่ 2.38



รูปที่ 2.38 การนำการคำนวณของ $a_t(i)$ ในเทอมของ t และ i (Implementation of the computation of $a_t(i)$ in terms of a lattice of observation t , and states i)

ในส่วนย่อยของรูปที่ 2.38 มีองค์ประกอบดังรูปที่ 2.39



รูปที่ 2.39 อธิบายลำดับของตัวแปรฟอร์เวิร์ด $a_{t+1}(j)$

หากใช้วิธีฟอร์เวิร์ดโปรซีเยอร์ในการคำนวณหาค่า $P(O|\lambda)$ โดยใช้คอมพิวเตอร์ จะมีจำนวนคำสั่งในการคำนวณทั้งหมดเป็น N^2T เท่านั้น ถ้าสมมติให้แบบจำลองมีสถานะทั้งหมด 5 สถานะ ($N = 5$) และมีชุดลำดับข้อมูล 100 ข้อมูล ($T = 100$) ดังนั้นคอมพิวเตอร์จะต้องใช้คำสั่งในการคำนวณประมาณ 3000 คำสั่งเท่านั้น เมื่อเทียบกับวิธีตรงแล้วจะมีความแตกต่างกันมาก

การแก้ปัญหาที่ 2

สำหรับปัญหาในข้อที่ 2 เป็นการเลือกเส้นทางของการเปลี่ยนแปลงสถานะที่ให้ค่าความเป็นไปได้มากที่สุด รายงานนี้ได้เสนอวิธีของไวเตอร์บี (Viterbi Algorithm) ซึ่งกำหนดให้ชุดลำดับของการเปลี่ยนแปลงสถานะที่ดีที่สุดเป็น $Q = \{q_1, q_2, \dots, q_T\}$ และให้ชุดลำดับของข้อมูลเป็น $O = \{O_1, O_2, \dots, O_T\}$ และต้องหาค่า

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1, q_2, \dots, q_t = i, O_1, O_2, \dots, O_t | \lambda] \quad (2.37)$$

โดยค่า $\delta_t(i)$ เป็นค่าที่ดีที่สุดของหนึ่งเส้นทางในลำดับที่ t และค่า $\psi_t(j)$ คือเส้นทางที่ให้ค่าที่ดีที่สุด วิธีของไวเตอร์บีจะมีส่วนประกอบที่สำคัญ 4 ส่วน คือ

1. การกำหนดค่าเริ่มต้น

$$\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad (2.38a)$$

$$\psi_1(i) = 0 \quad (2.38b)$$

2. การเรียกซ้ำ (Recursion)

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t), \quad \begin{matrix} 2 \leq t \leq T \\ 1 \leq j \leq N \end{matrix} \quad (2.39a)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], \quad \begin{matrix} 2 \leq t \leq T \\ 1 \leq j \leq N \end{matrix} \quad (2.39b)$$

3. การยกเลิก

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)] \quad (2.40a)$$

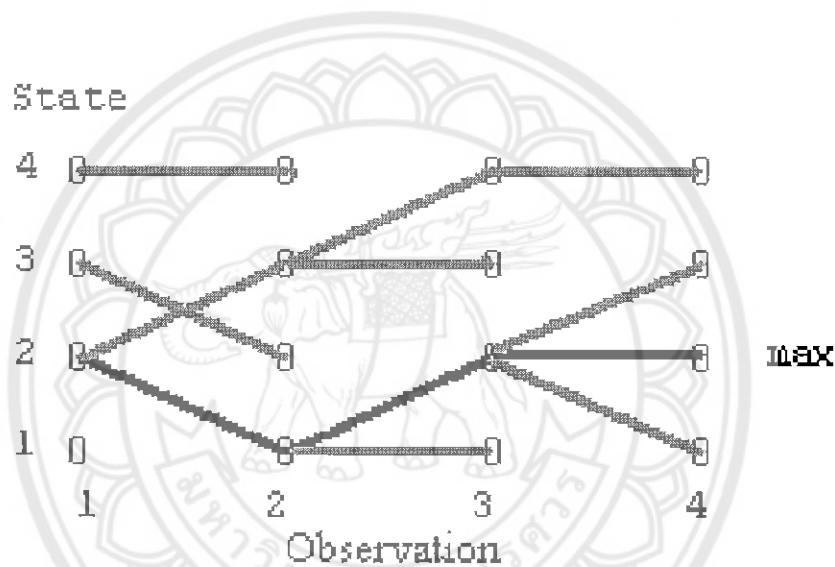
$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad (2.40b)$$

4. เส้นทางเดินย้อนกลับของสถานะ (Path backtracking)

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, \dots, 1 \quad (2.41)$$

ในส่วนที่ 1 จะเป็นค่าเริ่มต้นของข้อมูลตัวแรกของทุก ๆ สถานะในแบบจำลอง ส่วนเส้นทางเดินนั้นยังไม่มีจึงเป็นศูนย์ สำหรับส่วนที่ 2 เป็นการคำนวณค่า $\delta_t(i)$ และ $\psi_t(i)$ ในลำดับต่าง ๆ ถัดจากข้อมูลตัวแรกไปจนถึงข้อมูลตัวสุดท้าย ในส่วนที่ 3 จะเป็นส่วนสุดท้ายของข้อมูล ค่าที่เลือกเส้นทางของการเปลี่ยนแปลงสถานะจะเป็นค่า $\delta_T(i)$ ตัวสุดท้ายที่มีค่ามากที่สุด และจะเลือกเส้นทางเริ่มต้นจากสถานะนั้นด้วย จากนั้นในส่วนที่ 4 ก็จะเป็นการเลือกเส้นทางเดินทั้งหมดของชุดลำดับข้อมูล จะทำการเลือกย้อนกลับไปและมีเพียงเส้นทางเดียวที่ให้เลือกอยู่แล้ว

สมมติให้แบบจำลองมีสถานะ 4 สถานะและจำนวนข้อมูลมี 4 ตัว ดังรูปที่ 2.40 จะเห็นว่าเส้นทางที่มีสีเข้มเป็นเส้นทางที่ให้ค่าความเป็นไปได้มากที่สุด จะเลือกเส้นทางนี้เป็นเส้นทางการเปลี่ยนแปลงสถานะของแบบจำลองนี้



รูปที่ 2.40 วิธีของไวเตอร์บี (Viterbi Algorithm)

การแก้ปัญหที่ 3

แท้จริงแล้วการปรับปรุงค่าตัวแปรต่าง ๆ ในแบบจำลองนั้นไม่สามารถทำได้ หากไม่รู้การเปลี่ยนแปลงของสถานะ ซึ่งเส้นทางของการเปลี่ยนแปลงสถานะที่หาได้จากปัญหาที่ 2 เป็นเพียงเส้นทางที่คิดว่ามีความเป็นไปได้มากที่สุดเท่านั้น ถ้ามีจำนวนชุดข้อมูลที่นำมาฝึกฝนมีมากพอ ก็สามารถปรับค่าตัวแปรต่าง ๆ ในแบบจำลองให้ลู่เข้าสู่ค่าที่ดีที่สุดได้ และจะทำให้ค่า $P(O | \lambda)$ มากที่สุด

วิธีปรับค่าตัวแปรต่างๆในแบบจำลองกำหนดให้มีจำนวนชุดข้อมูล R ชุด (O_1, O_2, \dots, O_R) มาทำการประมวลผลในแบบจำลองและหาเส้นทางของการเปลี่ยนแปลงของสถานะ จากนั้นจะทำการปรับค่าตัวแปรต่าง ๆ คือค่า $\pi, a_j, b_j(k)$

จะได้แบบจำลองที่มีค่าตัวแปรเหล่านี้ขึ้นมาใหม่ ทำการฝึกฝนไปเรื่อย ๆ โดยใช้แบบจำลองที่ทำการปรับค่าตัวแปรเรียบร้อยแล้ว ค่าตัวแปรจะถูกปรับปรุงจนเข้าสู่ค่าใดค่าหนึ่ง ซึ่งเป็นค่าที่ทำให้แบบจำลองมีความสมบูรณ์มากที่สุด และทำให้ค่า $P(O|\lambda)$ มากที่สุด

2.4.6 สรุปแบบจำลองฮิดเดน มาร์คอฟ

จากเนื้อหาที่กล่าวมาทั้งหมดได้แสดงให้เห็นถึงองค์ประกอบต่าง ๆ ในแบบจำลองฮิดเดน มาร์คอฟและการคำนวณหาองค์ประกอบเหล่านั้น แท้จริงแล้วชุดลำดับข้อมูลที่กล่าวในรายงานนี้เป็นชุดข้อมูลที่มีลักษณะไม่ต่อเนื่อง (Discrete) ทั้งหมด ชุดข้อมูลที่มีลักษณะต่อเนื่อง (Continuous) จะมีลักษณะเดียวกัน เพียงแต่ค่าความน่าจะเป็นของข้อมูลในแบบต่าง ๆ ขึ้นอยู่กับว่าข้อมูลนั้นจะเป็นข้อมูลที่มีการกระจายลักษณะใด จากนั้นใช้ความรู้ทางคณิตศาสตร์มาแก้ปัญหาของฟังก์ชันโดยหลักการแบบจำลองฮิดเดน มาร์คอฟยังเหมือนเดิม

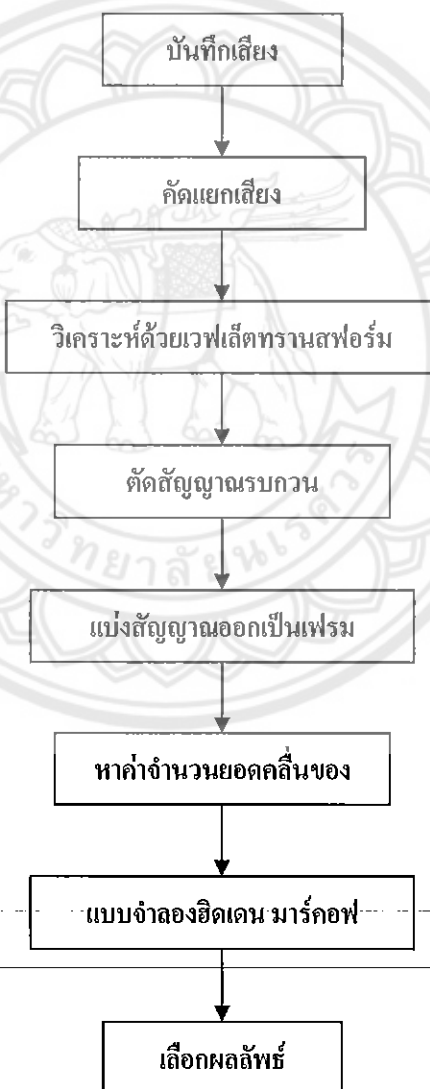
เมื่อทำความเข้าใจทฤษฎีแบบจำลองฮิดเดน มาร์คอฟอย่างแท้จริงจะพบว่าสัญญาณแต่ละสัญญาณจะมีลักษณะเฉพาะ ฉะนั้นหนึ่งลักษณะเฉพาะของสัญญาณสามารถนำมาสร้างแบบจำลองได้หนึ่งแบบจำลอง เช่น สัญญาณเสียงของคำว่า "A" ก็จะเป็นหนึ่งแบบจำลอง "B" ก็จะเป็นอีกหนึ่งแบบจำลอง ถ้าใช้เพียงเสียงผู้พูดคนเดียวมาทำการฝึกฝน แบบจำลองจะรู้จำได้ถึงลักษณะเฉพาะของเสียงผู้พูดคนเดียวเท่านั้น ถ้านำเสียงของผู้พูดหลาย ๆ คนมาทำการฝึกฝนคำว่า "A" แบบจำลองก็จะรู้จำได้ถึงลักษณะเสียงของคำว่า "A" มากขึ้น

การนำทฤษฎีแบบจำลองฮิดเดน มาร์คอฟไปประยุกต์ใช้นั้น จะแบ่งเป็น 3 ลักษณะ คือ ระบบการคาดเดา ระบบการรู้จำรู้จักสัญญาณและระบบการหาเอกลักษณ์ ซึ่งส่วนใหญ่นิยมนำไปใช้ในการรู้จำรู้จักสัญญาณ ตัวอย่างเช่น การรู้จำรู้จักเสียง (Speech Recognition) การประมวลผลภาพ (Image Processing) หรือการระบุตัวบุคคล (Identify) ไม่ว่าจะเป็นเสียง ภาพหรือรูปลักษณะต่าง ๆ

บทที่ 3

วิธีดำเนินการ

การพัฒนาโปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษโดยใช้โปรแกรม MatLAB เป็นเครื่องมือนั้น จำเป็นต้องศึกษาคำสั่งของโปรแกรม MatLAB ทฤษฎีเวฟเล็ตและทฤษฎีแบบจำลองฮิดเดน มาร์คอฟ เมื่อมีความเข้าใจในทฤษฎีที่เกี่ยวข้องแล้วจึงนำทฤษฎีมาประยุกต์ใช้ในการพัฒนาโปรแกรม กระบวนการทำงานของโปรแกรมจะแสดงในรูปที่ 3.1 เพื่อให้เห็นภาพรวมของการทำงานของโปรแกรม



รูปที่ 3.1 ขั้นตอนการดำเนินการ

หลังจากศึกษาเกี่ยวกับกระบวนการและหลักการในการรู้จำรู้จักเสียงแล้ว ขั้นตอนต่อมาคือการพัฒนาโปรแกรม โดยใช้โปรแกรม MatLAB เป็นเครื่องมือในการพัฒนา

วิธีการดำเนินการในส่วนแรกจะเป็นการบันทึกเสียงของผู้ใช้อย่างน้อย 2 ครั้งสำหรับแบบจำลองของอักขระหนึ่งตัว โดยเสียงครั้งแรกจะใช้สำหรับนำไปสร้างแบบจำลองของเสียง ส่วนเสียงครั้งที่สองจะนำมาใช้ปรับปรุงค่าต่าง ๆ ในแบบจำลองที่สร้างจากเสียงครั้งแรกเพื่อให้คู่เข้าสู่เส้นทางที่ถูกต้องมากที่สุด หรือเรียกได้ว่าเป็นการฝึกสอนโปรแกรมให้มีความฉลาดมากยิ่งขึ้น ยังมีการฝึกสอนมากเท่าใดย่อมส่งผลให้โปรแกรมมีความฉลาดมากยิ่งขึ้นตามไปด้วย

เสียงที่ได้จะทำการคำนวณเพื่อปรับบรรทัดฐานด้วยเวฟเล็ต แล้วทำการตัดแบ่งออกเป็นเฟรม (Frame) แล้วจึงทำการสร้างแบบจำลองหรือปรับค่าในแบบจำลองเดิมให้เหมาะสมและตัดสินใจเลือกผลลัพธ์ โดยระบบปัญญาประดิษฐ์ขั้นตอนต่าง ๆ มีรายละเอียดดังนี้

3.1 การคัดแยกสัญญาณ

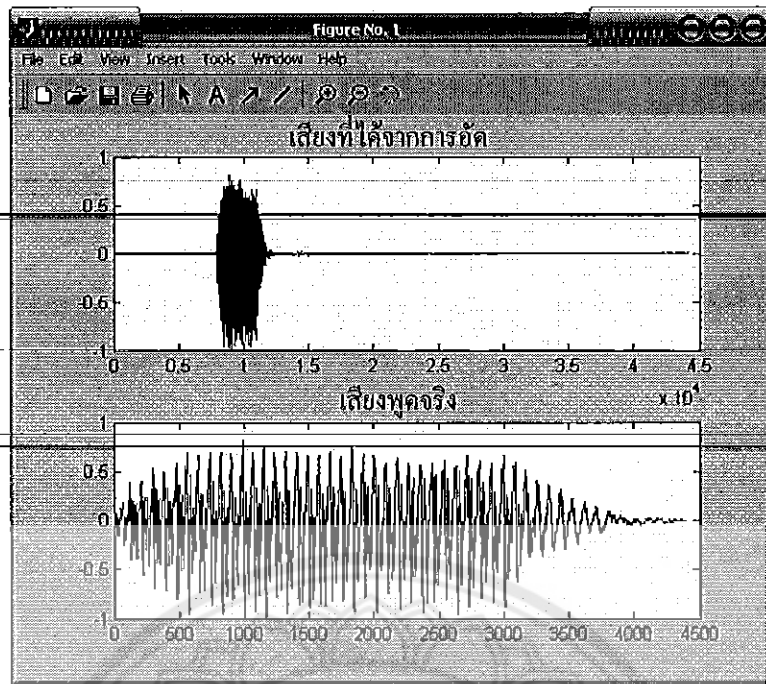
เมื่อบันทึกเสียงเข้ามาแล้ว จะทำการแยกเฉพาะสัญญาณเสียงพูดออกมา โดยใช้ค่าส่วนเบี่ยงเบนมาตรฐานเป็นตัวชี้วัด โดยจะทำการแบ่งสัญญาณเสียง ที่บันทึกเข้ามาออกเป็นเฟรมเฟรมละ 200 จุดข้อมูล จากนั้นทำการคำนวณค่าส่วนเบี่ยงเบนมาตรฐาน ของแต่ละเฟรมข้อมูล หากกลุ่มเฟรมข้อมูลในช่วงใดมีค่าส่วนเบี่ยงเบนมาตรฐานมากจะถือว่าเป็นข้อมูลจริง ที่เหลือจะถูกมองเป็นเสียงเงียบหรือสัญญาณรบกวนและถูกตัดทิ้งไป แล้วนำกลุ่มเฟรมข้อมูลที่เป็นสัญญาณเสียงจริงมารวมกลับเป็นเป็นสัญญาณเดิม รูปที่ 3.2 แสดงการคัดแยกเฉพาะสัญญาณที่เป็นเสียงพูดจริง และสมการที่ใช้หาค่าส่วนเบี่ยงเบนมาตรฐานคือสมการที่ 3.1

$$S.D. = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}} \quad (3.1)$$

โดยที่ X_i คือ ข้อมูลที่ไม่ได้แจกแจงความถี่ตัวที่ i

\bar{X} คือ ค่าเฉลี่ยเลขคณิตของข้อมูล

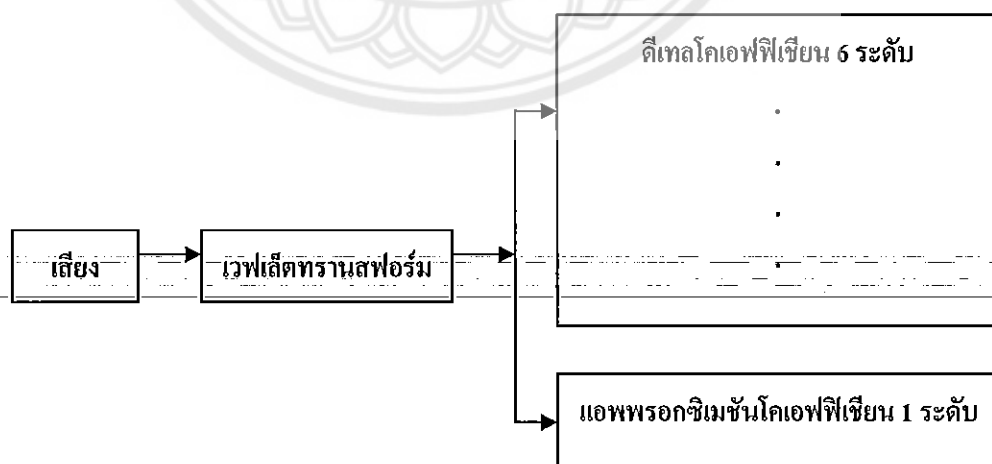
N คือ จำนวนข้อมูลทั้งหมด



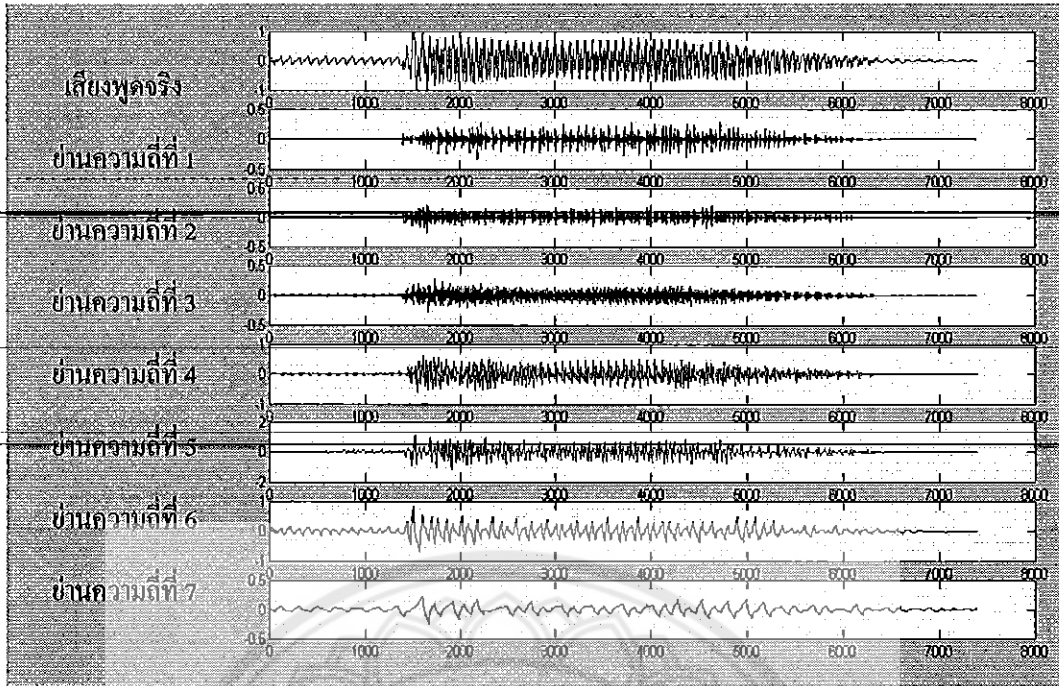
รูปที่ 3.2 แสดงการคัดแยกเฉพาะสัญญาณที่เป็นเสียงพูดจริง

3.2 การวิเคราะห์สัญญาณด้วยเวฟเลตทรานสฟอร์ม

นำสัญญาณเสียงที่บันทึกเข้ามาและถูกคัดแยกเฉพาะสัญญาณเสียงจริงแล้วมาวิเคราะห์ด้วยเวฟเลตทรานสฟอร์ม จะทำให้ได้ช่วงความถี่ของ สัญญาณออกเป็น 7 ย่านความถี่ คือ 6 ดีเทลโคแอฟฟิเซียนและ 1 แอพรอกซิเมชันโคแอฟฟิเซียนดังรูปที่ 3.3 และรูปที่ 3.4



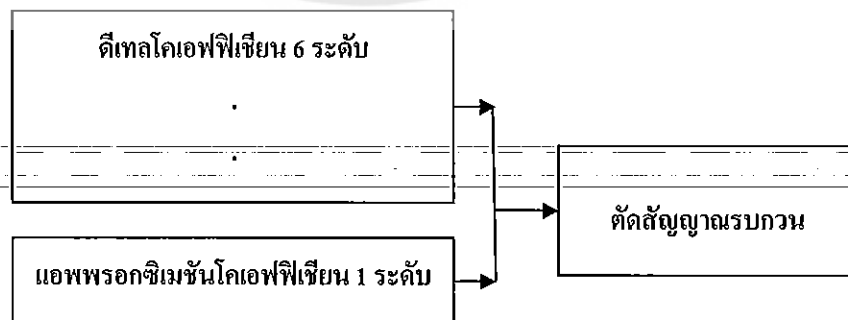
รูปที่ 3.3 การวิเคราะห์ด้วยเวฟเลตทรานสฟอร์ม



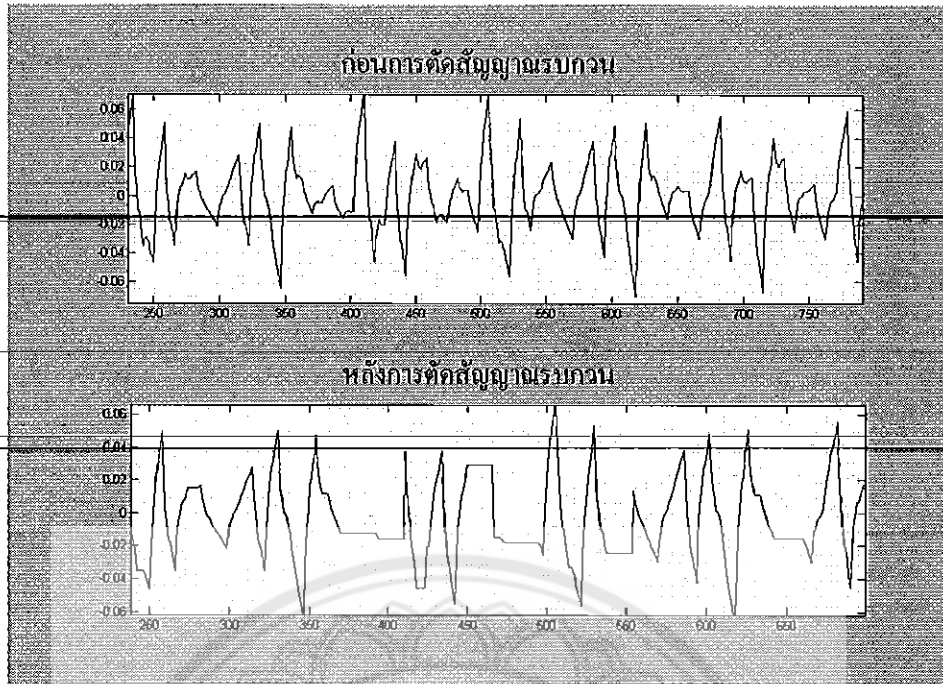
รูปที่ 3.4 แสดงสัญญาณที่วิเคราะห์ด้วยเวฟเลตทรานสฟอร์ม

3.3 การตัดสัญญาณรบกวน

หลังจากวิเคราะห์สัญญาณเสียงด้วยเวฟเลตทรานสฟอร์มและได้ย่านความถี่ทั้ง 7 ย่าน จากนั้นจะนำย่านความถี่ดังกล่าวมาทำการตัดสัญญาณรบกวนออก โดยการหาค่าเฉลี่ยผลต่างระหว่างยอดคลื่นของช่วงเสียงที่เงียบ แล้วจึงนำค่าที่ได้มาเป็นค่ามาตรฐานในการตัดสัญญาณรบกวน ถ้าหากในช่วงที่เป็นเสียงจริงมีผลต่างระหว่างยอดคลื่นน้อยกว่าค่ามาตรฐานที่ได้จะทำการปรับค่ายอดคลื่นทั้งสองให้มีค่าเท่ากัน ดังแสดงในรูปที่ 3.5 และรูปที่ 3.6



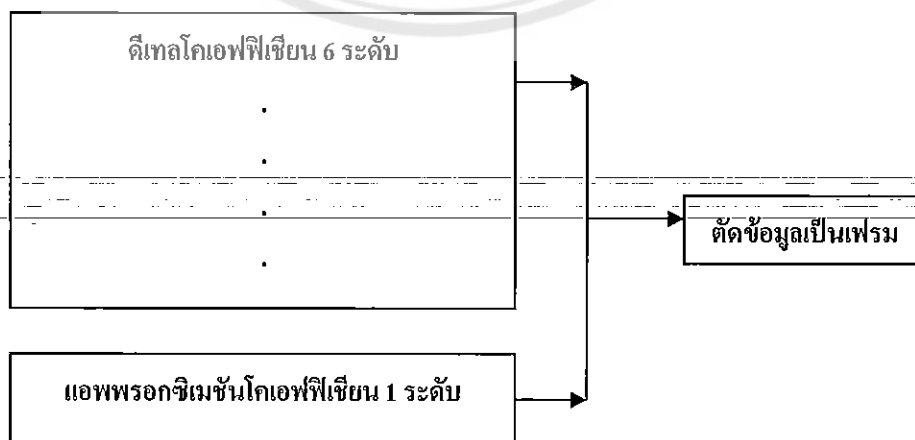
รูปที่ 3.5 การนำแต่ย่านความถี่มาลดสัญญาณรบกวน



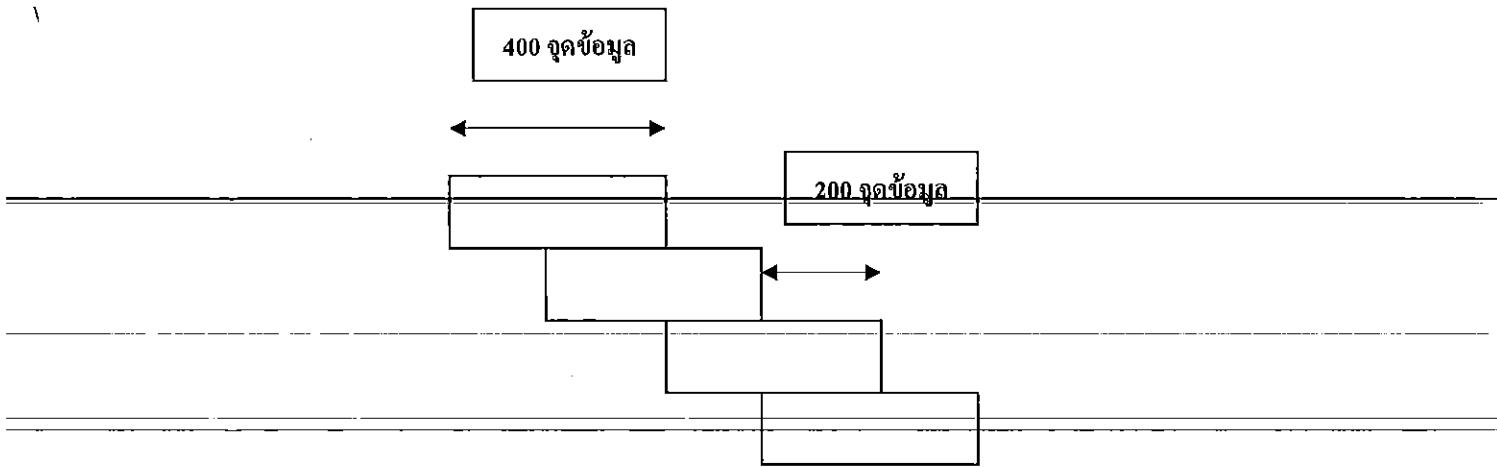
รูปที่ 3.6 ตัวอย่างสัญญาณที่ตัดสัญญาณรบกวนออก

3.4 การแบ่งสัญญาณออกเป็นส่วนเล็ก ๆ (Frame)

เมื่อวิเคราะห์สัญญาณเสียงด้วยเวฟเล็ตทรานสฟอร์ม เพื่อแยกย่านความถี่ออกเป็น 7 ย่านและทำการตัดสัญญาณรบกวน จากนั้นจะทำการตัดสัญญาณทั้ง 7 ย่านความถี่ออกเป็นเฟรมเฟรมละ 400 จุดข้อมูล โดยแต่ละเฟรมข้อมูลจะมีการซ้อนทับกัน 200 จุดข้อมูล ดังแสดงในรูปที่ 3.7 และ รูปที่ 3.8

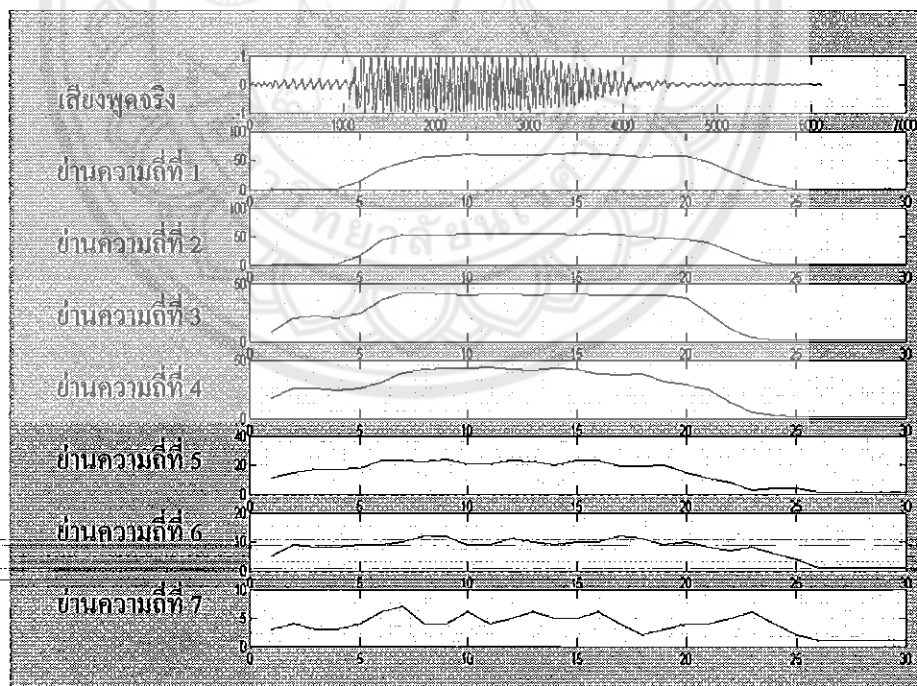


รูปที่ 3.7 การนำสัญญาณมาแบ่งเป็นเฟรม



รูปที่ 3.8 แสดงการแบ่งเฟรมข้อมูล

เมื่อแบ่งสัญญาณออกเป็นเฟรมจากดีเทิล โคเอฟฟีเซียนและแอฟพรอกซิเมชัน โคเอฟฟีเซียนทั้ง 7 ย่านแล้ว จากนั้นจะทำการหาค่าจำนวนยอดคลื่นประจำเฟรมข้อมูล เพื่อนำค่าจำนวนยอดคลื่นเหล่านี้มาใช้กับแบบจำลองฮิดเดน มาร์คอฟ ดังแสดงในรูปที่ 3.9



รูปที่ 3.9 แสดงค่าจำนวนยอดคลื่นในแต่ละย่านความถี่

3.5 การสร้างแบบจำลองฮิดเดน มาร์คอฟ

เมื่อทำการวิเคราะห์สัญญาณด้วยเวฟเลตทรานสฟอร์ม ทำการตัดสัญญาณรบกวน และแบ่งสัญญาณออกเป็นเฟรมที่มีการซ้อนทับกันแล้ว จากนั้นจะเป็นการหาค่าพารามิเตอร์เพื่อใช้กับแบบจำลองฮิดเดน มาร์คอฟ ซึ่งเป็นแบบจำลองเชิงความน่าจะเป็น

3.5.1 การหาความน่าจะเป็นของการเป็นสถานะที่จะเป็นสถานะเริ่มต้น ($b_{j, \text{pine}}$)

ค่า $b_{j, \text{pine}}$ นี้เป็นค่าที่จะบอกให้ทราบว่าที่ค่าจำนวนขดคลื่นนั้น มีค่าความน่าจะเป็นเท่าใดที่จะเป็นสถานะเริ่มต้น หาโดยการนำค่า $b_{j, \text{pine}}$ ของเสียง ที่มีการปรับช่วงค่าจำนวนขดคลื่นให้อยู่ในช่วง 0 – 80 ของทั้ง 7 ย่านความถี่ที่นำมาสร้างแบบจำลอง มาเปรียบเทียบกับว่ามีค่าจำนวนขดคลื่นใกล้เคียงความถี่ใด เช่น เสียง “A” 2 เสียง โดยที่ n คือ ค่าจำนวนขดคลื่น

เสียง “A” เสียงที่ 1	ย่านความถี่ที่ 1 มีค่าจำนวนขดคลื่นใกล้เคียง 80
	ย่านความถี่ที่ 2 มีค่าจำนวนขดคลื่นใกล้เคียง 79
	ย่านความถี่ที่ 3 มีค่าจำนวนขดคลื่นใกล้เคียง 78
	ย่านความถี่ที่ 4 มีค่าจำนวนขดคลื่นใกล้เคียง 3
	ย่านความถี่ที่ 5 มีค่าจำนวนขดคลื่นใกล้เคียง 2
	ย่านความถี่ที่ 6 มีค่าจำนวนขดคลื่นใกล้เคียง 3
	ย่านความถี่ที่ 7 มีค่าจำนวนขดคลื่นใกล้เคียง 1
เสียง “A” เสียงที่ 2	ย่านความถี่ที่ 1 มีค่าจำนวนขดคลื่นใกล้เคียง 80
	ย่านความถี่ที่ 2 มีค่าจำนวนขดคลื่นใกล้เคียง 78
	ย่านความถี่ที่ 3 มีค่าจำนวนขดคลื่นใกล้เคียง 78
	ย่านความถี่ที่ 4 มีค่าจำนวนขดคลื่นใกล้เคียง 2
	ย่านความถี่ที่ 5 มีค่าจำนวนขดคลื่นใกล้เคียง 1
	ย่านความถี่ที่ 6 มีค่าจำนวนขดคลื่นใกล้เคียง 2
	ย่านความถี่ที่ 7 มีค่าจำนวนขดคลื่นใกล้เคียง 1

นำมาเก็บในตารางความน่าจะเป็นของ b_j pine

ย่านความถี่ที่ 1		ย่านความถี่ที่ 2		ย่านความถี่ที่ 3		ย่านความถี่ที่ 4	
n	b_j pine	n	b_j pine	n	b_j pine	n	b_j pine
1	0.01	1	0.01	1	0.01	1	0.01
2	0.01	2	0.01	2	0.01	2	1
3	0.01	3	0.01	3	0.01	3	1
...		
78	0.01	78	1	78	1	78	0.01
79	0.01	79	1	79	0.01	79	0.01
80	1	80	0.01	80	0.01	80	0.01

ย่านความถี่ที่ 5		ย่านความถี่ที่ 6		ย่านความถี่ที่ 7	
n	b_j pine	n	b_j pine	n	b_j pine
1	1	1	0.01	1	1
2	1	2	1	2	0.01
3	0.01	3	1	3	0.01
...		
78	0.01	78	0.01	78	0.01
79	0.01	79	0.01	79	0.01
80	0.01	80	0.01	80	0.01

ตารางที่ 3.1 ตารางความน่าจะเป็นของ b_j pine

3.5.2 การหาค่าความน่าจะเป็นของการเปลี่ยนแปลงสถานะ (b_j)

ค่า b_j นี้เป็นค่าที่จะบอกให้ทราบว่ามีความน่าจะเป็นเท่าใด ที่ค่าจำนวนยอดคลื่นนั้นจะเป็นสถานะ หาโดยการนำค่าจำนวนยอดคลื่นของเฟรมที่สองของเสียง ที่มีการปรับช่วงค่าจำนวนยอดคลื่นให้อยู่ในช่วง 0 - 80 ของทั้ง 7 ย่านความถี่ที่นำมาสร้างแบบจำลอง แล้วเปรียบเทียบกับค่าใกล้เคียงกับค่าจำนวนยอดคลื่นใด และทำเช่นนี้กับเฟรมต่อไปจนหมดทั้งเสียง เช่น เสียง "A" 2 เสียง มีจำนวนเฟรมของแต่ละเสียงเท่ากับ 5 เฟรม ในเฟรมแรกใช้ในการหาค่า b_{j_pine} ดังนั้นจึงเลือกอีก 4 เฟรม คือ เฟรมที่ 2 ถึงเฟรมที่ 5 โดยที่ n คือ ค่าจำนวนยอดคลื่น

เสียง "A" เสียงที่ 1

ย่านความถี่ที่ 1 มีค่าจำนวนยอดคลื่นใกล้เคียง 80, 78, 79, 80 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 2 มีค่าจำนวนยอดคลื่นใกล้เคียง 78, 78, 79, 79 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 3 มีค่าจำนวนยอดคลื่นใกล้เคียง 78, 79, 80, 79 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 4 มีค่าจำนวนยอดคลื่นใกล้เคียง 1, 3, 3, 2 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 5 มีค่าจำนวนยอดคลื่นใกล้เคียง 2, 3, 2, 3 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 6 มีค่าจำนวนยอดคลื่นใกล้เคียง 1, 2, 3, 2 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 7 มีค่าจำนวนยอดคลื่นใกล้เคียง 1, 1, 2, 1 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5

เสียง "A" เสียงที่ 2

ย่านความถี่ที่ 1 มีค่าจำนวนยอดคลื่นใกล้เคียง 78, 80, 79, 78 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 2 มีค่าจำนวนยอดคลื่นใกล้เคียง 78, 79, 78, 80 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 3 มีค่าจำนวนยอดคลื่นใกล้เคียง 79, 79, 80, 80 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 4 มีค่าจำนวนยอดคลื่นใกล้เคียง 2, 1, 3, 1 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 5 มีค่าจำนวนยอดคลื่นใกล้เคียง 3, 2, 1, 3 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 6 มีค่าจำนวนยอดคลื่นใกล้เคียง 2, 1, 2, 2 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5
 ย่านความถี่ที่ 7 มีค่าจำนวนยอดคลื่นใกล้เคียง 1, 2, 2, 1 เรียงจากเฟรมที่ 2 มาเฟรมที่ 5

นำมาเก็บในตารางนับจำนวนของ b_j โดยเก็บแยกทีละเฟรม
เฟรมที่ 2

ย่านความถี่ที่ 1		ย่านความถี่ที่ 2		ย่านความถี่ที่ 3		ย่านความถี่ที่ 4	
n	b_j	n	b_j	n	b_j	n	b_j
1	0.01	1	0.01	1	0.01	1	1.01
2	0.01	2	0.01	2	0.01	2	1.01
3	0.01	3	0.01	3	0.01	3	0.01
...		
78	1.01	78	2.01	78	1.01	78	0.01
79	0.01	79	0.01	79	1.01	79	0.01
80	1.01	80	0.01	80	0.01	80	0.01

ย่านความถี่ที่ 5		ย่านความถี่ที่ 6		ย่านความถี่ที่ 7	
n	b_j	n	b_j	n	b_j
1	0.01	1	1.01	1	2.01
2	1.01	2	1.01	2	0.01
3	1.01	3	0.01	3	0.01
...		
78	0.01	78	0.01	78	0.01
79	0.01	79	0.01	79	0.01
80	0.01	80	0.01	80	0.01

ตารางที่ 3.2 ตารางนับจำนวนของ b_j เฟรมที่ 2

ย่านความถี่ที่ 1 ที่ค่าจำนวนยอดคลื่น 78 , 80 มีความน่าจะเป็นเท่ากับ 1.01/2.8
ย่านความถี่ที่ 2 ที่ค่าจำนวนยอดคลื่น 78 มีความน่าจะเป็นเท่ากับ 2.01/2.8
ย่านความถี่ที่ 3 ที่ค่าจำนวนยอดคลื่น 78 , 79 มีความน่าจะเป็นเท่ากับ 1.01/2.8
ย่านความถี่ที่ 4 ที่ค่าจำนวนยอดคลื่น 1 , 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
ย่านความถี่ที่ 5 ที่ค่าจำนวนยอดคลื่น 2 , 3 มีความน่าจะเป็นเท่ากับ 1.01/2.8
ย่านความถี่ที่ 6 ที่ค่าจำนวนยอดคลื่น 1 , 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
ย่านความถี่ที่ 7 ที่ค่าจำนวนยอดคลื่น 1 มีความน่าจะเป็นเท่ากับ 2.01/2.8

นำมาเก็บในตารางความน่าจะเป็นของ b_j โดยเก็บแยกที่สะพานเฟรมที่ 2

ย่านความถี่ที่ 1		ย่านความถี่ที่ 2		ย่านความถี่ที่ 3		ย่านความถี่ที่ 4	
n	b_j	n	b_j	n	b_j	n	b_j
1	0.01/2.8	1	0.01/2.8	1	0.01/2.8	1	1.01/2.8
2	0.01/2.8	2	0.01/2.8	2	0.01/2.8	2	1.01/2.8
3	0.01/2.8	3	0.01/2.8	3	0.01/2.8	3	0.01/2.8
...		
78	1.01/2.8	78	2.01/2.8	78	1.01/2.8	78	0.01/2.8
79	0.01/2.8	79	0.01/2.8	79	1.01/2.8	79	0.01/2.8
80	1.01/2.8	80	0.01/2.8	80	0.01/2.8	80	0.01/2.8

ย่านความถี่ที่ 5		ย่านความถี่ที่ 6		ย่านความถี่ที่ 7	
n	b_j	n	b_j	n	b_j
1	0.01/2.8	1	1.01/2.8	1	2.01/2.8
2	1.01/2.8	2	1.01/2.8	2	0.01/2.8
3	1.01/2.8	3	0.01/2.8	3	0.01/2.8
...		
78	0.01/2.8	78	0.01/2.8	78	0.01/2.8
79	0.01/2.8	79	0.01/2.8	79	0.01/2.8
80	0.01/2.8	80	0.01/2.8	80	0.01/2.8

ตารางที่ 3.3 ตารางความน่าจะเป็นของ b_j เฟรมที่ 2

เฟรมที่ 3

n	b_j
1	0.01
2	0.01
3	0.01
...	
78	1.01
79	0.01
80	1.01

n	b_j
1	0.01
2	0.01
3	0.01
...	
78	1.01
79	1.01
80	0.01

n	b_j
1	0.01
2	0.01
3	0.01
...	
78	0.01
79	2.01
80	0.01

n	b_j
1	1.01
2	0.01
3	1.01
...	
78	0.01
79	0.01
80	0.01

n	b_j
1	0.01
2	1.01
3	1.01
...	
78	0.01
79	0.01
80	0.01

n	b_j
1	1.01
2	1.01
3	0.01
...	
78	0.01
79	0.01
80	0.01

n	b_j
1	1.01
2	1.01
3	0.01
...	
78	0.01
79	0.01
80	0.01

ตารางที่ 3.4 ตารางนับจำนวนของ b_j เฟรมที่ 3

- ย่านความถี่ที่ 1 ที่ค่าจำนวนยอดคลื่น 78 , 80 มีความน่าจะเป็นเท่ากับ 1.01/2.8
- ย่านความถี่ที่ 2 ที่ค่าจำนวนยอดคลื่น 78 , 79 มีความน่าจะเป็นเท่ากับ 2.01/2.8
- ย่านความถี่ที่ 3 ที่ค่าจำนวนยอดคลื่น 79 มีความน่าจะเป็นเท่ากับ 1.01/2.8
- ย่านความถี่ที่ 4 ที่ค่าจำนวนยอดคลื่น 1 , 3 มีความน่าจะเป็นเท่ากับ 1.01/2.8
- ย่านความถี่ที่ 5 ที่ค่าจำนวนยอดคลื่น 2 , 3 มีความน่าจะเป็นเท่ากับ 1.01/2.8
- ย่านความถี่ที่ 6 ที่ค่าจำนวนยอดคลื่น 1 , 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
- ย่านความถี่ที่ 7 ที่ค่าจำนวนยอดคลื่น 1 , 2 มีความน่าจะเป็นเท่ากับ 2.01/2.8

นำมาเก็บในตารางความน่าจะเป็นของ b_j โดยเก็บแยกทีละเฟรม
เฟรมที่ 3

ย่านความถี่ที่ 1		ย่านความถี่ที่ 2		ย่านความถี่ที่ 3		ย่านความถี่ที่ 4	
n	b_j	n	b_j	n	b_j	n	b_j
1	0.01/2.8	1	0.01/2.8	1	0.01/2.8	1	1.01/2.8
2	0.01/2.8	2	0.01/2.8	2	0.01/2.8	2	0.01/2.8
3	0.01/2.8	3	0.01/2.8	3	0.01/2.8	3	1.01/2.8
...		
78	1.01/2.8	78	1.01/2.8	78	0.01/2.8	78	0.01/2.8
79	0.01/2.8	79	1.01/2.8	79	2.01/2.8	79	0.01/2.8
80	1.01/2.8	80	0.01/2.8	80	0.01/2.8	80	0.01/2.8

ย่านความถี่ที่ 5		ย่านความถี่ที่ 6		ย่านความถี่ที่ 7	
n	b_j	n	b_j	n	b_j
1	0.01/2.8	1	1.01/2.8	1	1.01/2.8
2	1.01/2.8	2	1.01/2.8	2	1.01/2.8
3	1.01/2.8	3	0.01/2.8	3	0.01/2.8
...		
78	0.01/2.8	78	0.01/2.8	78	0.01/2.8
79	0.01/2.8	79	0.01/2.8	79	0.01/2.8
80	0.01/2.8	80	0.01/2.8	80	0.01/2.8

ตารางที่ 3.5 ตารางความน่าจะเป็นของ b_j เฟรมที่ 3

ส่วนเฟรมที่เหลือให้ทำวิธีเดียวกันนี้ จนหมดเสียง

3.5.3 การหาค่าความน่าจะเป็นของการเปลี่ยนจากสถานะก่อนหน้ามาสถานะปัจจุบัน

(b_j state)

ค่า b_j state นี้เป็นค่าที่จะบอกให้ทราบว่ามีความน่าจะเป็นเท่าใดที่ค่าจำนวนยอดคลื่นที่เป็นสถานะก่อนหน้าจะเปลี่ยนความถี่ในสถานะปัจจุบัน หาโดยการนำเสียงที่มีครบในช่วงค่าจำนวนยอดคลื่นให้อยู่ในช่วง 0 – 80 ของทั้ง 7 ย่านความถี่ที่นำมาสร้างแบบจำลอง มาสังเกตว่ามีการเปลี่ยนแปลงสถานะอย่างไร และทำเช่นนี้กับเฟรมต่อไปจนหมดทั้งเสียง เช่น เสียง “ A ” 2 เสียง มีจำนวนเฟรมของแต่ละเสียงเท่ากับ 5 เฟรม โดยที่ n คือ ค่าจำนวนยอดคลื่น ตัวอย่างการวิเคราะห์สัญญาณเสียงของเสียง “A” จะแสดงดังตารางที่ 3.6 ถึงตารางที่ 3.19

เสียง “ A ” เสียงที่ 1

ย่านความถี่ที่ 1	มีค่าจำนวนยอดคลื่นใกล้เคียง 80 , 80 , 78 , 79 , 80	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 2	มีค่าจำนวนยอดคลื่นใกล้เคียง 79 , 78 , 78 , 79 , 79	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 3	มีค่าจำนวนยอดคลื่นใกล้เคียง 78 , 78 , 79 , 80 , 79	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 4	มีค่าจำนวนยอดคลื่นใกล้เคียง 3 , 1 , 3 , 3 , 2	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 5	มีค่าจำนวนยอดคลื่นใกล้เคียง 2 , 2 , 3 , 2 , 3	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 6	มีค่าจำนวนยอดคลื่นใกล้เคียง 3 , 1 , 2 , 3 , 2	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 7	มีค่าจำนวนยอดคลื่นใกล้เคียง 1 , 1 , 1 , 2 , 1	เรียงจากเฟรมที่ 2 ถึง 5

เสียง “ A ” เสียงที่ 2

ย่านความถี่ที่ 1	มีค่าจำนวนยอดคลื่นใกล้เคียง 80 , 78 , 80 , 79 , 78	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 2	มีค่าจำนวนยอดคลื่นใกล้เคียง 78 , 78 , 79 , 78 , 80	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 3	มีค่าจำนวนยอดคลื่นใกล้เคียง 78 , 79 , 79 , 80 , 80	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 4	มีค่าจำนวนยอดคลื่นใกล้เคียง 2 , 2 , 1 , 3 , 1	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 5	มีค่าจำนวนยอดคลื่นใกล้เคียง 1 , 3 , 2 , 1 , 3	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 6	มีค่าจำนวนยอดคลื่นใกล้เคียง 2 , 2 , 1 , 2 , 2	เรียงจากเฟรมที่ 2 ถึง 5
ย่านความถี่ที่ 7	มีค่าจำนวนยอดคลื่นใกล้เคียง 1 , 1 , 2 , 2 , 1	เรียงจากเฟรมที่ 2 ถึง 5

นำมาเก็บในตารางนับจำนวนของ b_j state โดยเก็บแยกเป็นเปลี่ยนจากเฟรมที่ 1 ไปยังเฟรมที่ 2 , จากเฟรมที่ 2 ไปยังเฟรมที่ 3 , จากเฟรมที่ 3 ไปยังเฟรมที่ 4 , ไปเรื่อย ๆ จนหมดเสียงจากเฟรมที่ 1 ไปยังเฟรมที่ 2

ย่านความถี่ที่ 1

$n_{\text{ก่อน}} / n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	0.01		0.01	0.01	0.01
2	0.01	0.01	0.01		0.01	0.01	0.01
3	0.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		0.01	0.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		1.01	0.01	1.01

ตารางที่ 3.6 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 1

ย่านความถี่ที่ 2

$n_{\text{ก่อน}} / n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	0.01		0.01	0.01	0.01
2	0.01	0.01	0.01		0.01	0.01	0.01
3	0.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		1.01	0.01	0.01
79	0.01	0.01	0.01		1.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.7 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 2

ย่านความถี่ที่ 3

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	0.01		0.01	0.01	0.01
2	0.01	0.01	0.01		0.01	0.01	0.01
3	0.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		1.01	1.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.8 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 3

ย่านความถี่ที่ 4

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	0.01		0.01	0.01	0.01
2	0.01	1.01	0.01		0.01	0.01	0.01
3	1.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		0.01	0.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.9 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 4

ย่านความถี่ที่ 5

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	1.01		0.01	0.01	0.01
2	0.01	1.01	0.01		0.01	0.01	0.01
3	0.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		0.01	0.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.10 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 5

ย่านความถี่ที่ 6

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01	0.01	0.01		0.01	0.01	0.01
2	0.01	1.01	0.01		0.01	0.01	0.01
3	1.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		0.01	0.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.11 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 6

ย่านความถี่ที่ 7

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	2.01	0.01	0.01		0.01	0.01	0.01
2	0.01	0.01	0.01		0.01	0.01	0.01
3	0.01	0.01	0.01		0.01	0.01	0.01
...							
78	0.01	0.01	0.01		0.01	0.01	0.01
79	0.01	0.01	0.01		0.01	0.01	0.01
80	0.01	0.01	0.01		0.01	0.01	0.01

ตารางที่ 3.12 ตารางนับจำนวนของ b_j state ย่านความถี่ที่ 7

ย่านความถี่ที่ 1 ที่ค่าจำนวนขอดคลื่น 80 เปลี่ยนเป็น 80 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 1 ที่ค่าจำนวนขอดคลื่น 80 เปลี่ยนเป็น 78 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 2 ที่ค่าจำนวนขอดคลื่น 79 เปลี่ยนเป็น 78 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 2 ที่ค่าจำนวนขอดคลื่น 78 เปลี่ยนเป็น 78 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 3 ที่ค่าจำนวนขอดคลื่น 78 เปลี่ยนเป็น 78 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 3 ที่ค่าจำนวนขอดคลื่น 78 เปลี่ยนเป็น 79 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 4 ที่ค่าจำนวนขอดคลื่น 3 เปลี่ยนเป็น 1 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 4 ที่ค่าจำนวนขอดคลื่น 2 เปลี่ยนเป็น 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 5 ที่ค่าจำนวนขอดคลื่น 2 เปลี่ยนเป็น 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 5 ที่ค่าจำนวนขอดคลื่น 1 เปลี่ยนเป็น 3 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 6 ที่ค่าจำนวนขอดคลื่น 3 เปลี่ยนเป็น 1 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 6 ที่ค่าจำนวนขอดคลื่น 2 เปลี่ยนเป็น 2 มีความน่าจะเป็นเท่ากับ 1.01/2.8
 ย่านความถี่ที่ 7 ที่ค่าจำนวนขอดคลื่น 1 เปลี่ยนเป็น 1 มีความน่าจะเป็นเท่ากับ 2.01/2.8

นำมาเก็บในตารางความน่าจะเป็นของ b_j state โดยเก็บแยกเป็นเปลี่ยนจากเฟรมที่ 1 ไปยังเฟรมที่ 2 , จากเฟรมที่ 2 ไปยังเฟรมที่ 3 , จากเฟรมที่ 3 ไปยังเฟรมที่ 4 , ไปเรื่อย ๆ จนหมดเสียงจากเฟรมที่ 1 ไปยังเฟรมที่ 2

ย่านความถี่ที่ 1

$n_{\text{ก่อน}} / n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		1.01/2.8	0.01/2.8	1.01/2.8

ตารางที่ 3.13 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 1

ย่านความถี่ที่ 2

$n_{\text{ก่อน}} / n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		1.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		1.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.14 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 2

ย่านความถี่ที่ 3

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		1.01/2.8	1.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.15 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 3

ย่านความถี่ที่ 4

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	1.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	1.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.16 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 4

ย่านความถี่ที่ 5

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	1.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	1.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.17 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 5

ย่านความถี่ที่ 6

$n_{\text{ก่อน}}/n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	1.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	1.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.18 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 6

ย่านความถี่ที่ 7

$n_{\text{ก่อน}} / n_{\text{หลัง}}$	1	2	3	...	78	79	80
1	2.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
2	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
3	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
...							
78	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
79	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8
80	0.01/2.8	0.01/2.8	0.01/2.8		0.01/2.8	0.01/2.8	0.01/2.8

ตารางที่ 3.19 ตารางความน่าจะเป็นของ b_j state ย่านความถี่ที่ 7

ส่วนเฟรมที่เหลือให้ทำวิธีเดียวกันนี้ จนหมดเสียง

3.6 การตัดสินใจ

เมื่อนำเสียง 1 เสียงเข้าแบบจำลอง จะได้ค่าความน่าจะเป็น 8 ค่า คือ ค่าความน่าจะเป็นของย่านความถี่ที่ 1 ถึง 7 และค่าความน่าจะเป็นของทุกย่านความถี่คู่กัน ค่าความน่าจะเป็นของแต่ละย่านความถี่ หาโดยการนำค่า b_j pine , b_j , b_j state มาคูณกัน ค่าความน่าจะเป็นของทุกย่านความถี่คู่กัน หาโดยการนำค่าความน่าจะเป็นของแต่ละย่านความถี่มาคูณกัน เมื่อได้แบบจำลองเสียงครบทั้ง 26 เสียง นำมาเรียงลำดับค่าความน่าจะเป็นของแต่ละย่านความถี่จากมากไปหาน้อย นำค่ามากที่สุดมาให้คะแนนความสำคัญ 5 คะแนน รองลงมาให้ 4 คะแนน ลำดับที่ 3 ให้ 3 คะแนน ลำดับที่ 4 ให้ 2 คะแนน ลำดับที่ 5 ให้ 1 คะแนน ส่วนค่าความน่าจะเป็นของทุกย่านความถี่คู่กันจะให้คะแนนเป็น 3 เท่าของคะแนนค่าความน่าจะเป็นของแต่ละย่านความถี่ นำคะแนนที่ได้ของละเสียงมาเปรียบเทียบกันว่าเสียงใดมีคะแนนมากที่สุด ให้เสียงนั้นเป็นเสียงที่ถูกต้อง

บทที่ 4

การทดลองและวิเคราะห์ผลการทดลอง

ในการทดลองผู้พัฒนาได้ทำการเก็บข้อมูลที่ใช้ในการทดสอบและพัฒนาปรับปรุงกฎเกณฑ์การตัดสินใจรวมทั้งปรับปรุงค่าพารามิเตอร์ของแบบจำลองฮิดเดน มาร์คอฟ ให้มีประสิทธิภาพและความแม่นยำเพิ่มมากขึ้น โดยรายละเอียดต่างๆมีดังต่อไปนี้

4.1 ข้อมูลที่ใช้ในการทดสอบ

ในกระบวนการสร้างแบบจำลองอ้างอิง การฝึกสอนแบบจำลองและการทดสอบการรู้จำของแบบจำลอง ผู้พัฒนาได้ทำการทดสอบโดยมีผู้ทดสอบ 3 คน แต่ละคนจะพูดเพื่อสร้างแบบจำลองอ้างอิงของและฝึกสอนแบบจำลอง 10 ครั้ง (สำหรับ 1 เสียงอักขระ) ทำเช่นนี้จนครบทั้งหมด 26 เสียง เมื่อทำการทดสอบจะทดสอบเฉพาะรูปแบบ (Profile) ของแต่ละคน ดังนั้นจะมีแบบจำลองทั้งหมด 3 กลุ่ม กลุ่มละ 26 แบบจำลอง ทุกแบบจำลองได้รับการฝึกสอนจากเจ้าของรูปแบบแล้ว

4.2 ผลการทดลอง

หลังจากทำการสร้างและฝึกสอนแบบจำลองฮิดเดน มาร์คอฟแล้ว จากนั้นทำการทดสอบโปรแกรม โดยให้ผู้ทดสอบทำการพูดแบบเรียงลำดับจากเสียง A-Z แล้วหาความถูกต้องของผลลัพธ์ออกมาเป็นร้อยละ ดังแสดงในตารางที่ 4.1 ถึง 4.3 โดยใช้เวลาในการประมวลผลตั้งแต่เริ่มต้นจนถึงสิ้นสุดเท่ากับ 53 วินาที

ลำดับที่	เสียง	ผลลัพธ์		
		ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
1	a	a	a	a
2	b	b	d	b
3	c	c	e	c
4	d	b	b	d
5	e	e	b	c
6	f	f	f	s
7	g	g	g	e
8	h	h	s	s
9	i	i	i	i
10	j	k	j	j
11	k	k	j	k
12	l	l	l	l
13	m	n	m	m
14	n	n	n	n
15	o	o	o	o
16	p	t	p	p
17	q	q	q	q
18	r	r	r	r
19	s	h	s	x
20	t	p	t	t
21	u	u	u	u
22	v	v	v	b
23	w	w	w	w
24	x	s	x	x
25	y	y	y	y
26	z	z	z	z
ความถูกต้อง (ร้อยละ)		73.07	80.76	76.92

ตาราง 4.1 ผลการทดลองของผู้ใช้คนที่หนึ่ง(เสียงอักษร)

ลำดับที่	เสียง	ผลลัพธ์		
		ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
1	a	a	a	a
2	b	b	d	b
3	c	c	e	c
4	d	d	d	e
5	e	c	e	d
6	f	s	f	f
7	g	g	g	p
8	h	s	h	h
9	i	i	i	i
10	j	j	a	j
11	k	k	k	j
12	l	l	l	l
13	m	m	n	m
14	n	n	n	n
15	o	o	o	o
16	p	p	p	p
17	q	q	q	q
18	r	r	r	r
19	s	x	s	h
20	t	p	t	t
21	u	u	u	u
22	v	v	v	b
23	w	w	w	w
24	x	s	a	x
25	y	y	y	y
26	z	z	z	z
ความถูกต้อง (ร้อยละ)		76.92	80.76	76.92

ตาราง 4.2 ผลการทดลองของผู้ใช้คนที่สอง(เสียงอักษร)

ลำดับที่	เสียง	ผลลัพธ์		
		ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
1	a	a	a	a
2	b	b	d	b
3	c	c	e	c
4	d	d	d	b
5	e	c	e	d
6	f	f	f	x
7	g	g	g	p
8	h	s	h	h
9	i	i	i	i
10	j	j	j	a
11	k	k	k	j
12	l	l	l	l
13	m	m	m	n
14	n	m	n	n
15	o	o	o	o
16	p	p	p	p
17	q	q	q	q
18	r	r	r	r
19	s	x	s	h
20	t	t	p	t
21	u	u	u	u
22	v	v	b	v
23	w	w	w	w
24	x	s	x	h
25	y	y	y	y
26	z	z	z	z
ความถูกต้อง (ร้อยละ)		80.76	84.61	73.07

ตาราง 4.3 ผลการทดลองของผู้ใช้คนที่สาม(เสียงอักษร)

ผลการทดลองจากตาราง 4.1 ถึง 4.3 พบว่าโปรแกรมมีความถูกต้องของการแสดงผลลัพธ์ สำหรับผู้ทดสอบแต่ละคนมากกว่าร้อยละ 70 เสียงที่มักมีการแสดงผลลัพธ์ผิดพลาดจะเป็นเสียงที่มีการออกเสียงคล้าย ๆ กัน เช่น b-d-c-e-v, a-j-k, f-h-s-x, m-n, p-t ทั้งนี้ ความถูกต้องของผลลัพธ์ จะขึ้นอยู่กับความชัดเจนและการเน้นหนักเสียงว่าคล้ายคลึงกับเสียงที่ใช้สร้างแบบจำลองอ้างอิงหรือฝึกสอนแบบจำลองมากน้อยเพียงใด



บทที่ 5

บทสรุป

โครงการนี้เป็นการศึกษาประยุกต์ใช้ดิฟเฟอเรนเชียลอีควชันและทฤษฎีแบบจำลองฮิดเดน มาร์คอฟในการรู้จำเสียงอักขระในภาษาอังกฤษ ซึ่งการค้นคว้าเริ่มจากการศึกษาเวฟเล็ตทรานสฟอร์ม เพื่อใช้ในการสังเคราะห์ลักษณะเด่นของเสียง และทฤษฎีแบบจำลองฮิดเดน มาร์คอฟ เพื่อใช้เก็บลักษณะเด่นของเสียงและอ้างอิงหาผลลัพธ์

คณะผู้จัดทำได้ใช้โปรแกรม MatLAB เป็นเครื่องมือในการพัฒนา ด้วยการบันทึกสัญญาณเสียงที่เข้ามาด้วยความถี่ 11025 จุดข้อมูลต่อวินาทีโดยใช้เวลาในการบันทึก 4 วินาทีจากนั้นทำการคัดแยกสัญญาณเฉพาะสัญญาณที่เป็นเสียงพูดจริง แล้วทำการวิเคราะห์สัญญาณด้วยเวฟเล็ตทรานสฟอร์ม โดยการผ่านตัวกรองความถี่สูงผ่าน (High pass filter) จะได้ดิฟเฟอเรนเชียลอีควชันออกมา 6 ระดับ และแอมพลิจูดซิมเพิลดิฟเฟอเรนเชียลอีควชัน 1 ระดับ จากนั้นทำการตัดสัญญาณรบกวนออกและแบ่งสัญญาณจริงออกเป็นเฟรมที่มีความยาว 400 จุดข้อมูลและแต่ละเฟรมข้อมูลมีการซ้อนทับกัน 200 จุดข้อมูล จากนั้นทำการหาค่าความถี่และค่าความน่าจะเป็นต่าง ๆ ของเฟรมข้อมูลเพื่อนำมาใช้กับแบบจำลองฮิดเดน มาร์คอฟต่อไป จากการทดลองสรุปผลได้ดังนี้

5.1 สรุปผลการทดลอง

1. การใช้โปรแกรม MatLAB ในการพัฒนา พบว่าใช้เวลาในการคำนวณและประมวลผลมากพอสมควร
2. โปรแกรม MatLAB ช่วยลดความยุ่งยากในการประมวลผลสัญญาณ
3. จากการทดลองพบว่าให้ความถูกต้องของผลลัพธ์มากกว่าร้อยละ 70
4. เสียงที่มักมีการแสดงผลผิดพลาด คือ เสียงที่มีการออกเสียงคล้ายคลึงกัน เช่น b-d-c-e-v , a-j-k , f-h-s-x , m-n , p-t ทั้งนี้ความถูกต้องของผลลัพธ์จะขึ้นอยู่กับความชัดเจนและการเน้นหนักเสียงว่าคล้ายคลึงกับเสียงที่ใช้แบบจำลองอ้างอิงหรือฝึกสอนแบบจำลองมากน้อยเพียงใด

5.2 ข้อเสนอแนะ

1. ควรมีการปรับปรุงให้สามารถแสดงผลพร้อมออกมาเป็นไฟล์ข้อความ (.txt, .doc) เพื่อความสะดวกในการใช้งาน

~~2. ควรมีการปรับปรุงให้มีเสถียรภาพและความถูกต้องมากขึ้น และพัฒนาให้สามารถแยกแยะจุดเด่นของเสียงที่มีความคล้ายคลึงกันออกจากกันให้ชัดเจนกว่านี้~~

3. ควรมีการพัฒนาให้ใช้กับภาษาไทยที่ไม่ใช่คำโดด

4. ควรมีการติดต่อกับผู้ใช้แบบกราฟฟิก (Graphic user interface)



เอกสารอ้างอิง

- [1] รศ.ดร.มนัส สังวรศิลป์ และ วรรัตน์ ภัทรอมรกุล . คู่มือการใช้งาน Matlab ฉบับสมบูรณ์ .
 นนทบุรี : สำนักพิมพ์อินโฟเพรส , 2543 .
- [2] รศ.ดร.สุทธิชัย เปรมฤดีปรีชาชาญ . การวิเคราะห์ปัญหาคุณภาพไฟฟ้าโดยใช้การแปลงเวฟเลข
 และโครงข่ายประสาทเทียม . มหาวิทยาลัยเชียงใหม่ , 2547 .
- [3] หนึ่งทัย ชัยอากร . กระบวนการลูกโซ่มาร์คอฟและการประยุกต์ . การค้นคว้าอิสระเชิง
 วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต มหาวิทยาลัยเชียงใหม่ , 2541 .
- [4] กาญจนา ทองบุญนาค . การรู้จำเสียงคำโดดด้วยโครงข่ายประสาทเทียม . การค้นคว้าอิสระเชิง
 วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต มหาวิทยาลัยเชียงใหม่ , 2544 .
- [5] Richard J. Mammone . **Artificial Neural Networks for Speech and Vision** .
 Chapman & Hall , 1993 .
- [6] "Wavelet Toolbox." [Online]. Available:
[Http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/acknowle.html](http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/acknowle.html)
- [7] "Matlab routines for Linear Predictive Code (LPC)." [Online]. Available:
[Http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/lpc/html](http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/lpc/html)
- [8] "Contents." [Online]. Available:
[Http:// www.eso.org/projects/esomidass/doc/user/98NOV/volb/node1.html](http://www.eso.org/projects/esomidass/doc/user/98NOV/volb/node1.html)



ภาคผนวก ก
ความต้องการของระบบ

การใช้งานโปรแกรม “รู้จำรู้จักเสียงอักขระในภาษาอังกฤษ” ผู้ใช้จะต้องติดตั้งโปรแกรมและมีระบบฮาร์ดแวร์ ดังนี้

ส่วนของโปรแกรม

1. Operating System : Windows 98 Second Edition
(Windows XP Professional Recommend)
2. MatLAB 6.1 Release 12

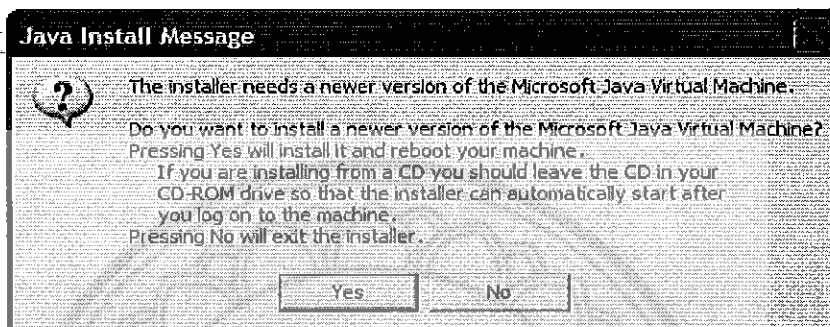
ส่วนของฮาร์ดแวร์

1. Central Processing Unit : Pentium III 1000 MHz (Pentium 4 1800 MHz Recommend)
2. Random Access Memory : 128 MB (256 MB Recommend)
3. Hard Disk Drive : Free space more than 600 MB for MatLAB installation
: Free space more than 700 MB for 1 user profile (Alphabet)
: Free space more than 350 MB for 1 user profile (Numeric)
4. Sound Card
5. Microphone

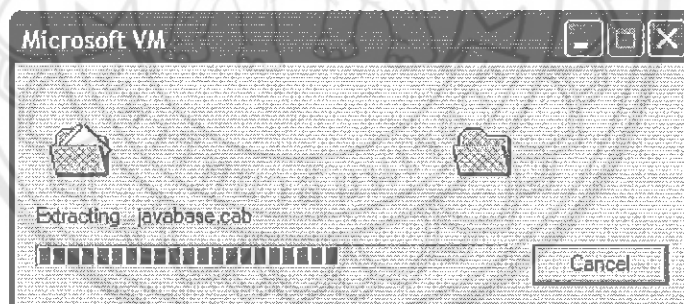
ภาคผนวก ข

การติดตั้งโปรแกรม MatLAB 6.1 Release 12

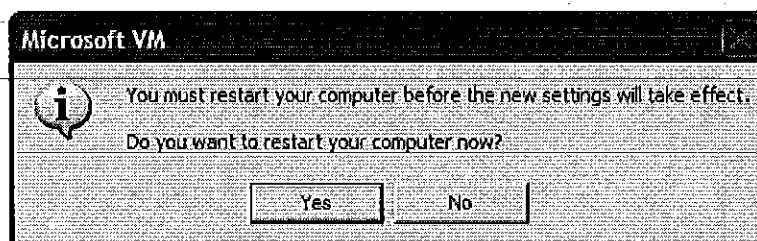
โปรแกรมรู้จักเสี่ยงอักขระในภาษาอังกฤษ ใช้โปรแกรมประมวลผลทางคณิตศาสตร์ที่ชื่อว่า MatLAB มีขั้นตอนการติดตั้งดังต่อไปนี้



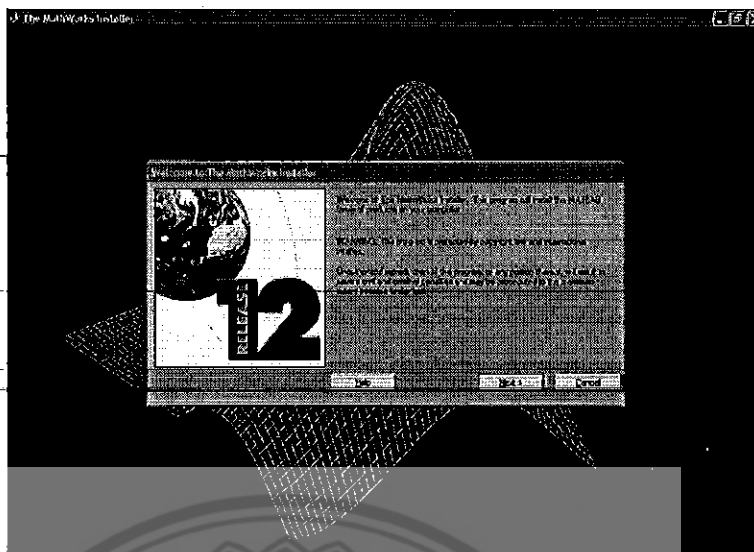
1. โปรแกรมจะถามว่าต้องการติดตั้ง Microsoft Java Virtual Machine ที่ใหม่กว่าหรือไม่ ให้คลิก Yes เพื่อติดตั้ง



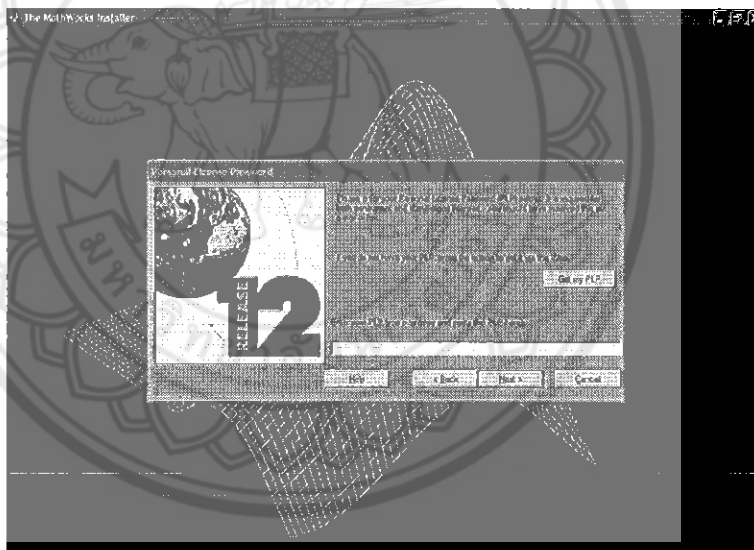
2. โปรแกรมทำการคัดลอกไฟล์ข้อมูล



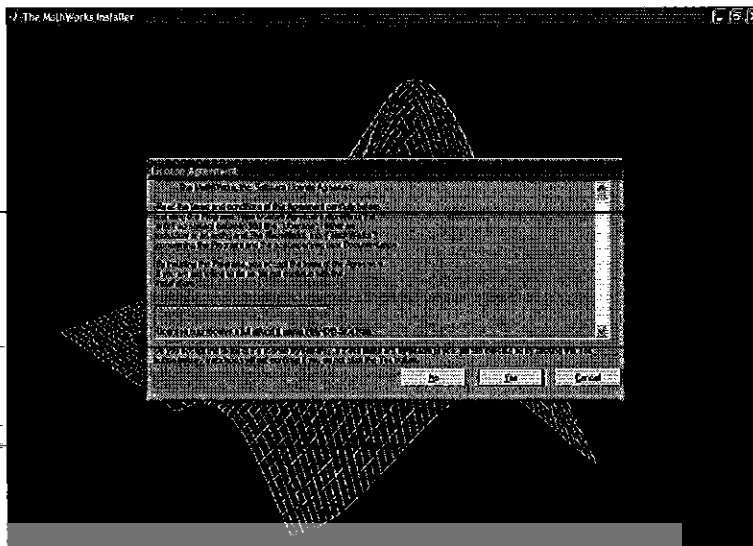
3. คลิก Yes เพื่อรีสตาร์ทเครื่องหลังการติดตั้ง Microsoft Java Virtual Machine เสร็จแล้ว



4. คลิก Next เพื่อเข้าสู่การติดตั้ง MatLAB



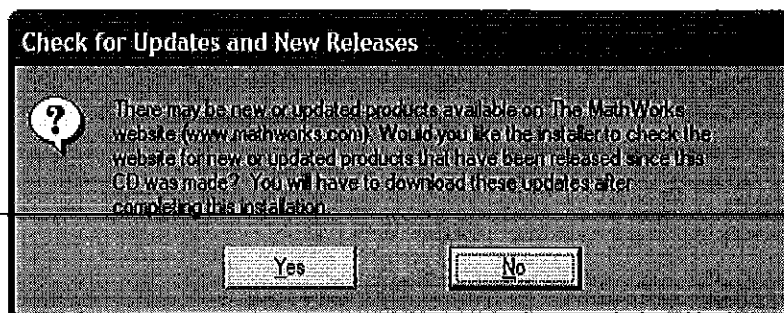
5. ใส่ Serial Number แล้วคลิก Next



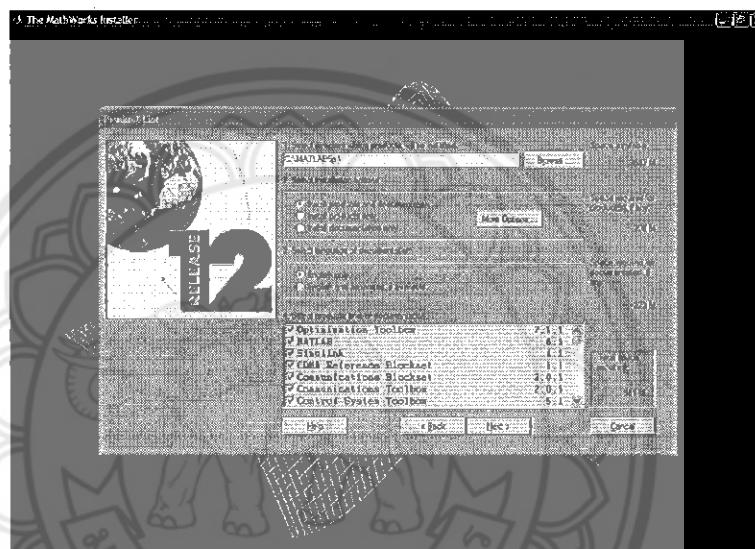
6. คลิก Yes หลังอ่านข้อตกลง



7. ใส่ชื่อผู้ใช้แล้วคลิก Next



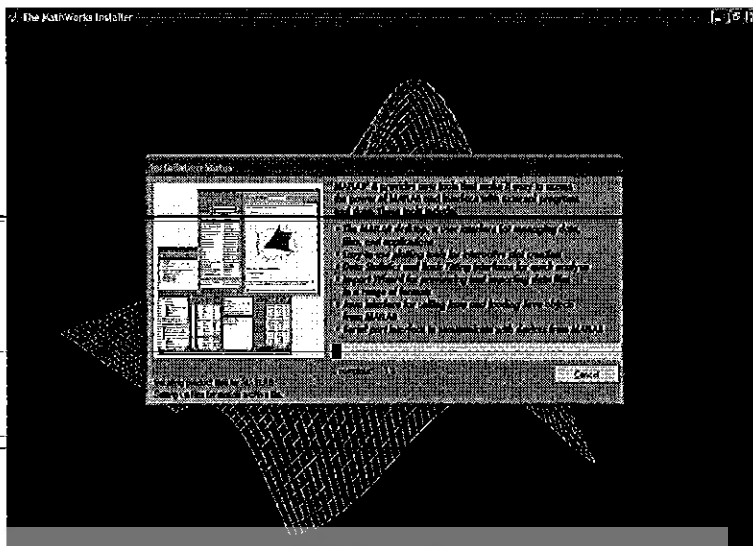
8. คลิก No เพื่อไม่ Update โปรแกรมผ่านเว็บไซต์



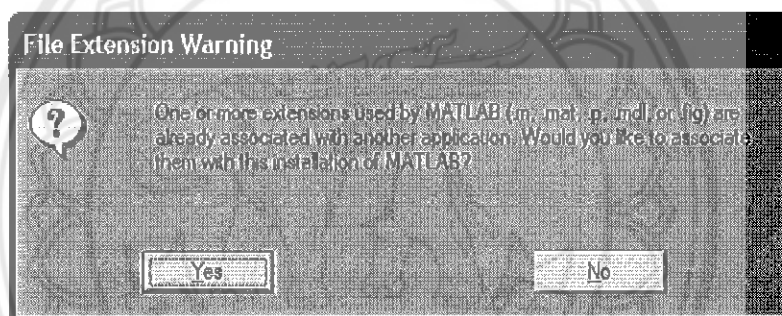
9. คลิก Next หลังจากเลือก Component ของโปรแกรม



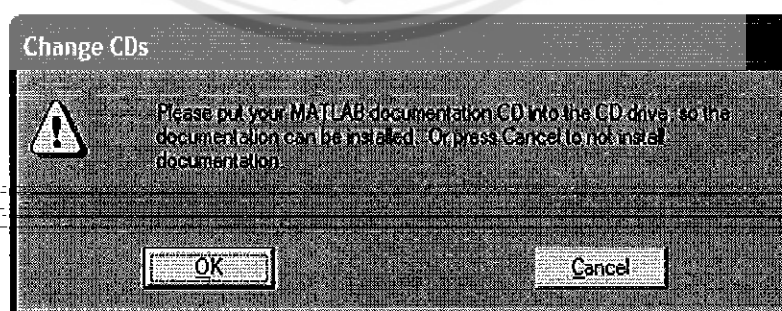
10. คลิก Yes เพื่อสร้างและติดตั้งโปรแกรมไปยังไดเรกทอรี c:\matlab6p1



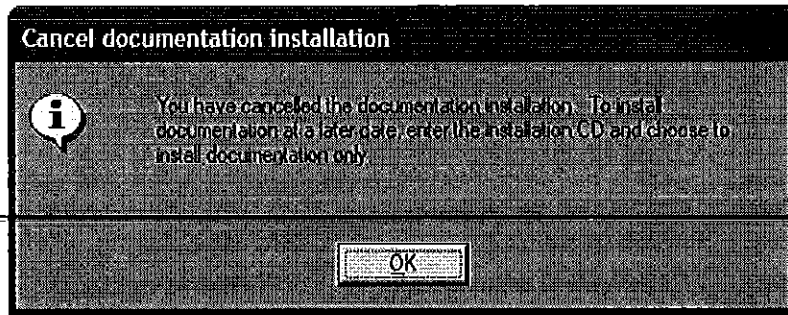
11. โปรแกรมทำการคัดลอกข้อมูล



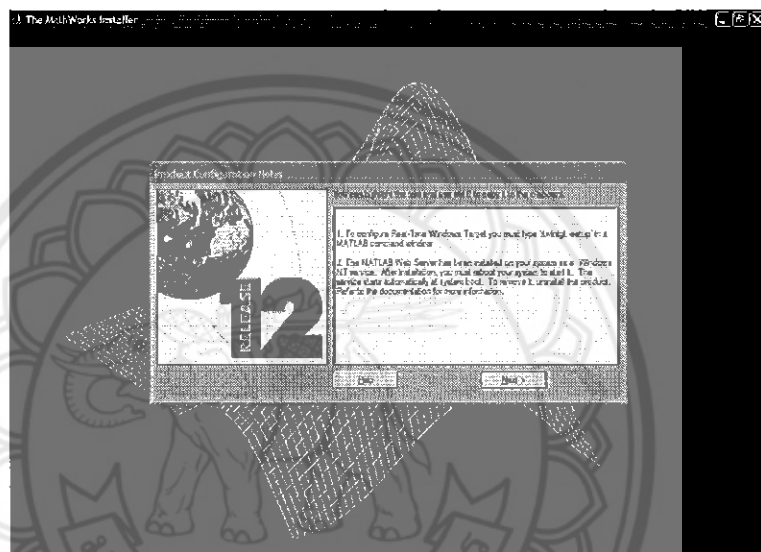
12. คลิก Yes เพื่อให้โปรแกรมติดตั้งองค์ประกอบที่สามารถสร้างไฟล์ข้อมูลที่มีนามสกุล .m, .mat, .p, .mdl, .fig



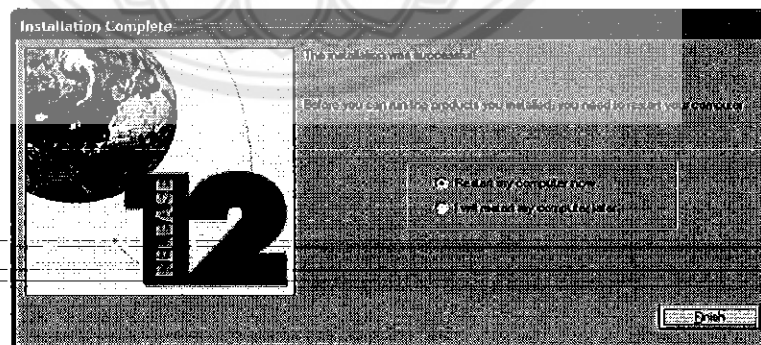
13. คลิก Cancel เพื่อไม่ติดตั้ง MatLAB document



14. คลิก OK เพื่อยืนยันว่าไม่ต้องการติดตั้ง MatLAB document



15. คลิก Next



16. คลิก Finish แล้วรอการรีสตาร์ทเครื่อง

ภาคผนวก ก

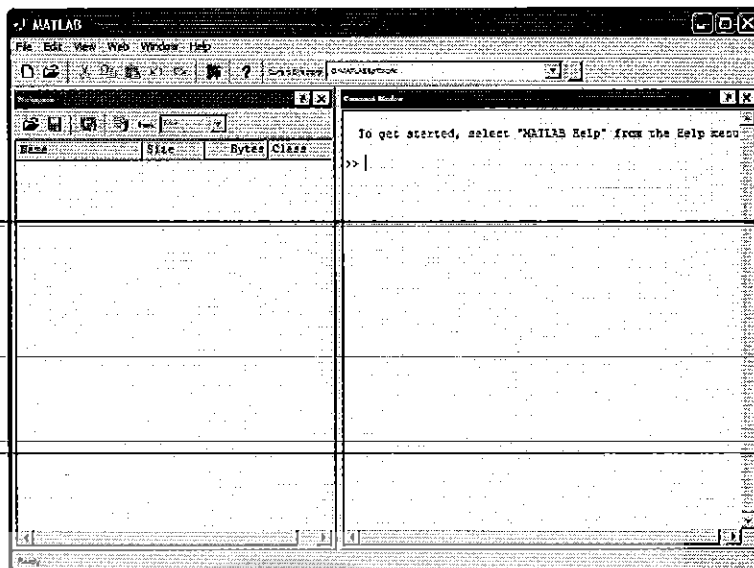
การติดตั้งโปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ

การโปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ จำเป็นต้องทำงานร่วมกับโปรแกรม
MatLAB การติดตั้งโปรแกรมรู้จำรู้จำเสียงอักขระในภาษาอังกฤษ มีขั้นตอนการติดตั้งดังต่อไปนี้

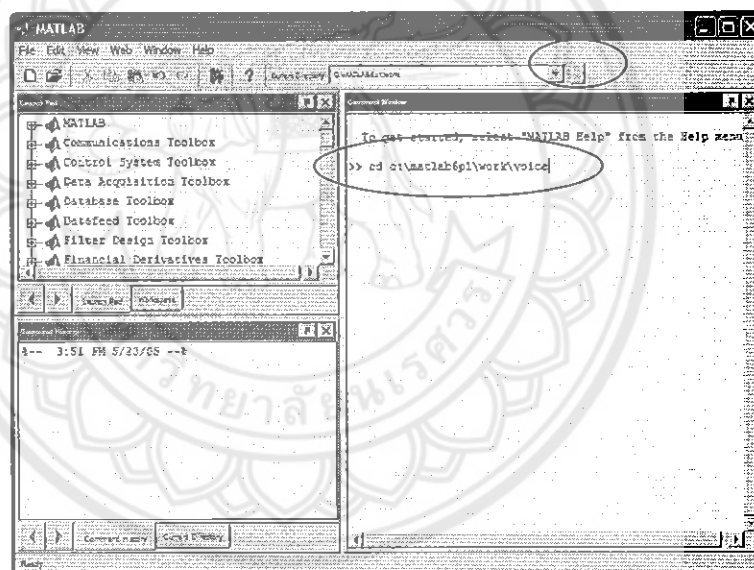
1. ทำการคัดลอกไดเรกทอรี VOICE ที่มีไฟล์ข้อมูลชื่อ

- VOICE
- ANALYZE_VOICE
- CREATE_MODEL
- FIND_FREQ3
- RECOGNITION
- RECORD_SOUND
- TEACH
- UPDATE_Bj
- DENOISE

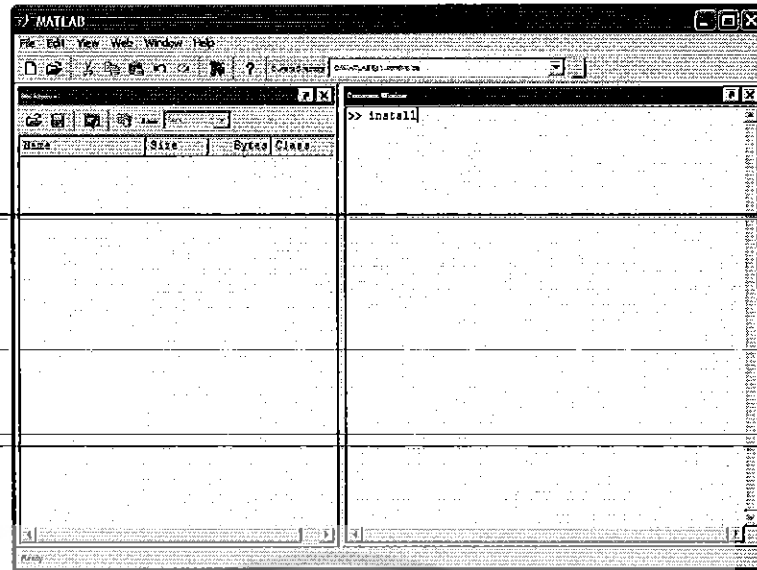
นำไปไว้ที่ใดก็ได้ที่มีเนื้อที่ว่างบนฮาร์ดดิสก์มากกว่า 700 เมกกะไบต์ (สำหรับผู้ใช้ 1 คนจะ
ใช้เนื้อที่ประมาณ 700 เมกกะไบต์)



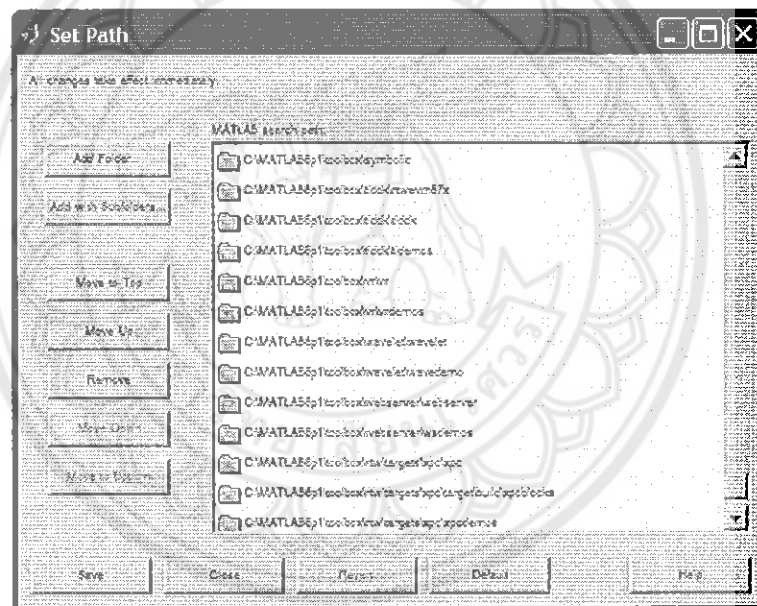
2. เปิดโปรแกรม MatLAB



3. พิมพ์ cd แล้วตามด้วย Path ที่นำไดเรกทอรี VOICE ไปเก็บไว้หรือทำการ Browse ก็ได้



4. พิมพ์ install แล้วกด Enter

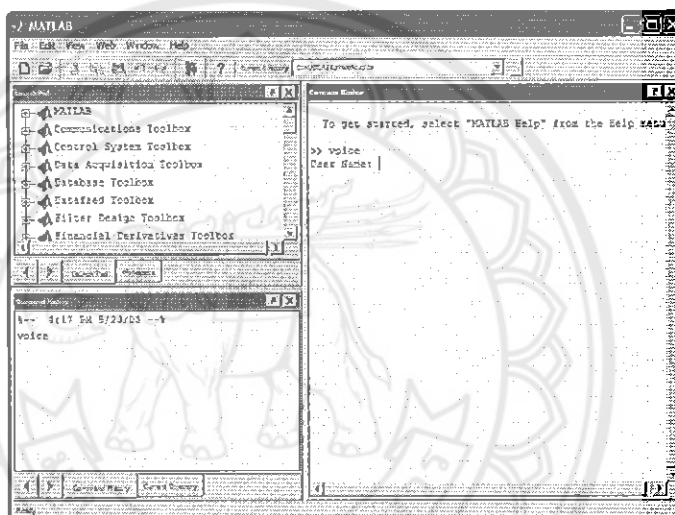


5. เลือก File -> Set path แล้วตรวจสอบ Path ที่ตั้งไว้, คลิกที่ Path นั้นแล้วคลิกปุ่ม Save
เพื่อการใช้งานครั้งต่อไปจะไม่ต้องมีการ Set path ใหม่

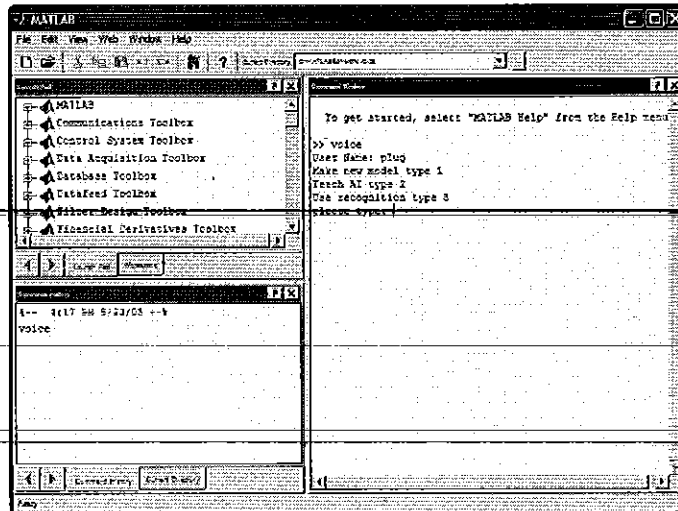
ภาคผนวก ง

คู่มือการใช้โปรแกรมรู้จักเสียงอักขระในภาษาอังกฤษ

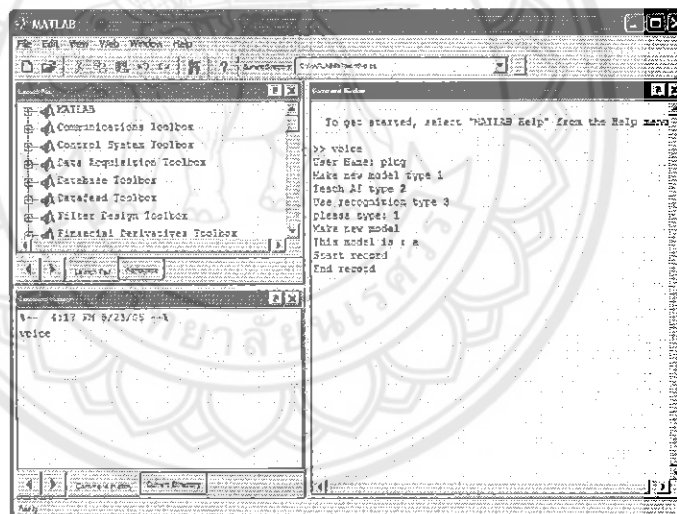
โปรแกรมรู้จักเสียงอักขระในภาษาอังกฤษมีการเรียกใช้แบบ Command line ผ่านโปรแกรม MatLAB มีขั้นตอนการใช้ดังต่อไปนี้



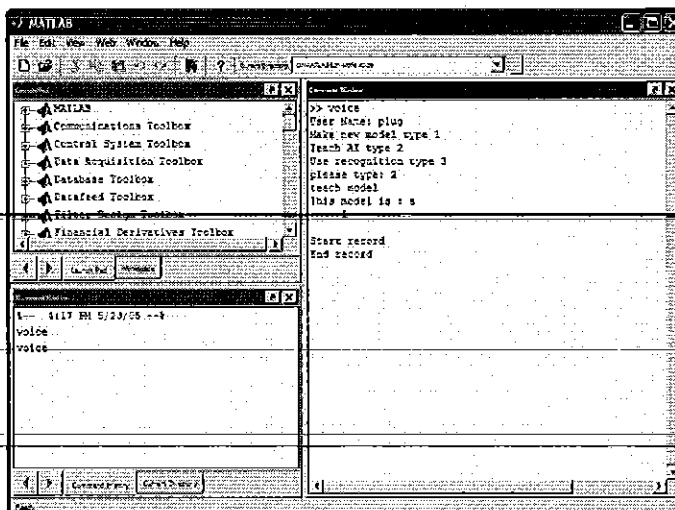
1. พิมพ์ voice เพื่อเรียกใช้โปรแกรม แล้วใส่ User Name ของผู้ใช้



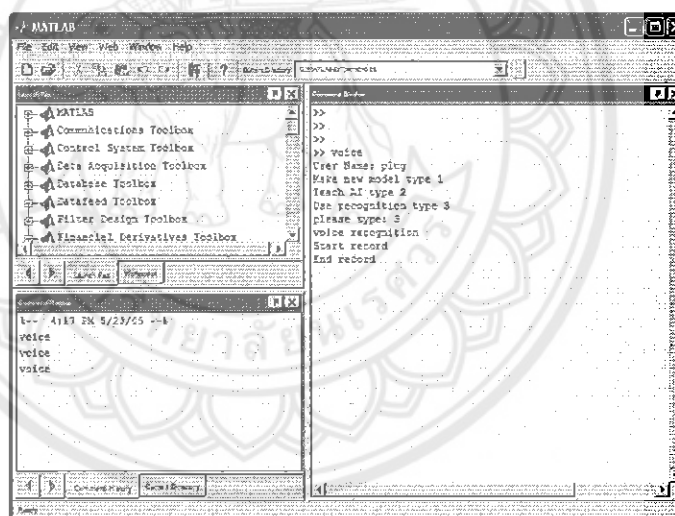
2. เลือก 1 เพื่อสร้างแบบจำลองของเสียง
- 2 เพื่อทำการฝึกสอนแบบจำลอง
- 3 เพื่อใช้การรู้จำรู้จักเสียง



หากเลือก 1 จะต้องใส่ชื่อของแบบจำลองเสียงที่จะสร้างด้วย จากนั้นจะมีการบันทึกเสียงครั้ง เพื่อสร้างและฝึกสอนแบบจำลอง



หากเลือก 2 เพื่อทำการฝึกสอนแบบจำลองที่มีอยู่แล้วให้มีความฉลาดขึ้น จะต้องมีการบอกชื่อแบบจำลองที่ต้องการฝึกสอนด้วย จากนั้นโปรแกรมจะทำการบันทึกเสียง 10 ครั้งเพื่อใช้ในการฝึกสอน

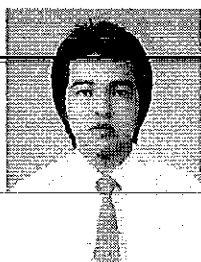


หากเลือก 3 เพื่อใช้การรู้จำรู้จกเสียงจะมีการบันทึกเสียงเข้าไป 1 ครั้งและจะทำการคำนวณหาผลลัพธ์ จากนั้นจะถามว่าผลลัพธ์ที่แสดงออกมาถูกต้องหรือไม่

ตอบ 1 เมื่อผลลัพธ์ถูกต้อง

ตอบ 2 เมื่อผลลัพธ์ผิดพลาด แล้วรับคำตอบที่ถูกต้องเข้าไป เพื่อนำไปฝึกสอนแบบจำลองต่อไป

ประวัติผู้เขียนโครงการ



ชื่อ นาย ปุณรัตน์ ทองธรรม

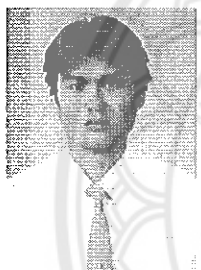
ภูมิลำเนา 71/2 หมู่ 1 ต.ธานี อ.เมือง จ.สุโขทัย 64000

ประวัติการศึกษา

- จบระดับมัธยมศึกษา โรงเรียนจุฬารัตน์ราชวิทยาลัย
พิษณุโลก

- ปัจจุบันกำลังศึกษาอยู่ในระดับปริญญาตรีชั้นปีที่ 4
สาขาวิศวกรรมไฟฟ้าและคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

E-mail : lism_zeroone@yahoo.com



ชื่อ นาย วินัย เรือนเพ็ง

ภูมิลำเนา 67 หมู่ 3 ต.นาบ่อคำ อ.เมือง จ.กำแพงเพชร 62000

ประวัติการศึกษา

- จบระดับมัธยมศึกษา โรงเรียนนาบ่อคำวิทยาคม

- ปัจจุบันกำลังศึกษาอยู่ในระดับปริญญาตรีชั้นปีที่ 4
สาขาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

E-mail : winai_r99@mthai.com



ชื่อ นาย ศรัทธัย ต้นสมบูรณ์

ภูมิลำเนา 145 หมู่ 8 ต.ทุ่งเสลี่ยม อ.ทุ่งเสลี่ยม จ.สุโขทัย 64150

ประวัติการศึกษา

- จบระดับมัธยมศึกษา โรงเรียนทุ่งเสลี่ยมชนูปถัมภ์

- ปัจจุบันกำลังศึกษาอยู่ในระดับปริญญาตรีชั้นปีที่ 4
สาขาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

E-mail : spoon04@hotmail.com