

การค้นหเส้นอ้างอิงในฟอร์มเอกสารโดยใช้การแปลงเวฟเล็ต

REFERENCE LINE EXTRACTION IN DOCUMENT FORMS  
USING WAVELET TRANSFORM



วิทยานิพนธ์นี้เขียนขึ้นเพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2544

ISBN 974-648-362-5

การค้นหาเส้นอ้างอิงในฟอร์มเอกสารโดยใช้การแปลงเวฟเล็ต

REFERENCE LINE EXTRACTION IN DOCUMENT FORMS  
USING WAVELET TRANSFORM



เสกสรรค์ พลศรี  
SEKSAN PHONSRI

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2544

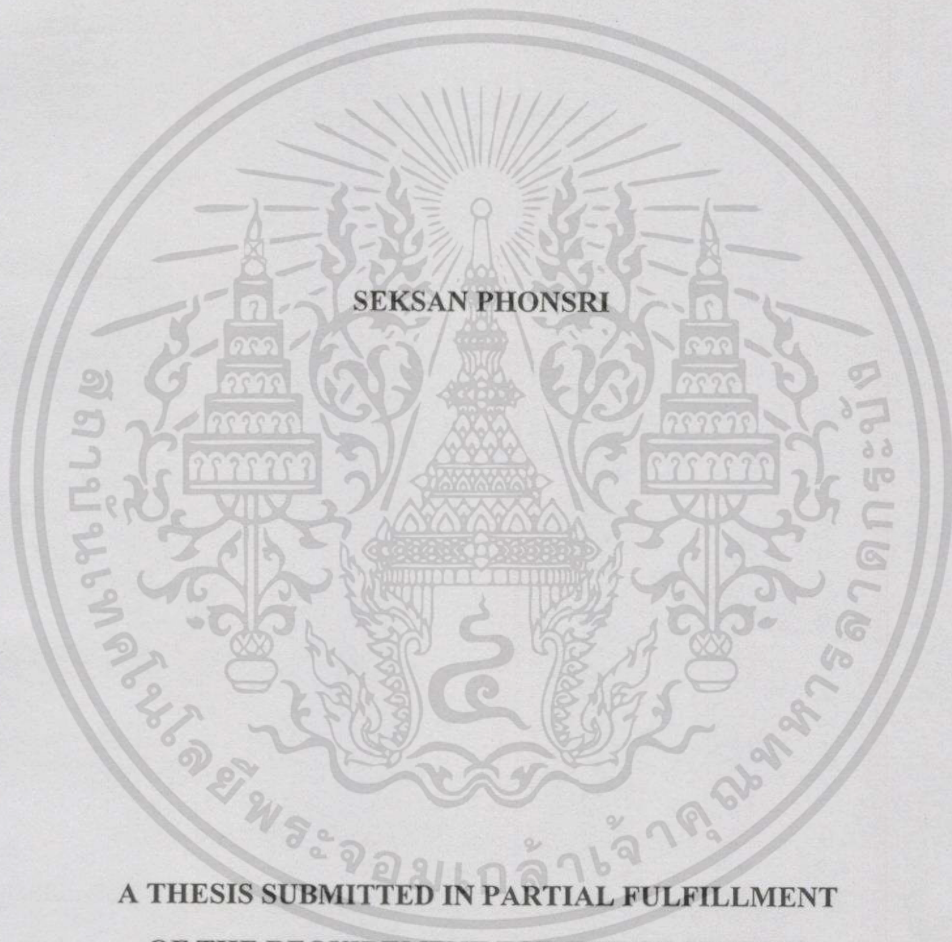
ISBN 974-648-362-5

เลขหมู่.....  
เลขทะเบียน **40622**  
วัน, เดือน, ปี **1.8 ค.ค. 2544**

.b.....
.i.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**REFERENCE LINE EXTRACTION IN DOCUMENT FORMS  
USING WAVELET TRANSFORM**



**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING  
SCHOOL OF GRADUATE STUDIES  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

**2001**

**ISBN 974-648-362-5**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2001**

**SCHOOL OF GRADUATE STUDIES**

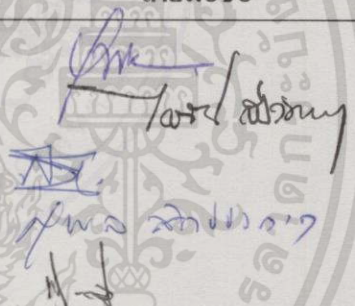
**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บัณฑิตวิทยาลัย  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การค้นหาเส้นอ้างอิงในฟอร์มเอกสาร โดยใช้การแปลงเวฟเล็ต  
REFERENCE LINE EXTRACTION IN DOCUMENT FORMS USING  
WAVELET TRANSFORM

ชื่อนักศึกษา นายเสกสรรค์ พลศรี  
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชา วิศวกรรมไฟฟ้า  
อาจารย์ผู้ควบคุมวิทยานิพนธ์ ผศ.ดร.ปัญญา จิตมีขนิมา

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.ดร.ยุทธพงษ์	รังสรรค์เสรี	
ผศ.ดร.ไกรสิน	สังวัฒนา	
ผศ.ดร.สุรพันธุ์	เอื้อไพบูรณ์	
รศ.ดร.สุวิพล	สิทธิชีวกา	
ผศ.ดร.ปัญญา	จิตมีขนิมา	

วัน/เดือนปี ที่สอบ 24 กรกฎาคม 2544 เวลา 10.00-12.00 น.

สถานที่สอบ ณ อาคาร 12 ชั้น ชั้น 4 (ห้อง E12-404)

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.บุญวัฒน์ อัคร)

คณบดีบัณฑิตวิทยาลัย

วันที่.....๕.....เดือน.....พ.ศ. ๒๕๔๔

หัวข้อวิทยานิพนธ์

การค้นหาเส้นอ้างอิงในฟอร์มเอกสารโดยใช้การแปลงเวฟ  
เล็ต

นักศึกษา

นายเสกสรรค์ พลศรี

รหัสประจำตัว

40061085

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา

วิศวกรรมไฟฟ้า

พ.ศ.

2544

อาจารย์ผู้ควบคุมวิทยานิพนธ์

ผศ.ดร.ปัญญา จูติมัชฌิมา

### บทคัดย่อ

เอกสารฟอร์มส่วนมากนิยมใช้เส้นตรงเป็นแนวอ้างอิงตำแหน่งสำหรับการกรอกข้อมูล ระบบป้อนข้อมูลอัตโนมัติสำหรับประมวลผลเอกสารเช่นนี้ต้องมีความสามารถในการค้นหาเส้นอ้างอิงเหล่านี้เพื่อทราบตำแหน่งของข้อมูลภายในฟอร์มเอกสาร วิทยานิพนธ์นี้เสนออัลกอริทึมสำหรับการค้นหาเส้นอ้างอิงโดยใช้การแปลงเวฟเล็ต (Wavelet Transform) ซึ่งมีคุณสมบัติในการแปลงภาพเอกสารออกเป็นรูปภาพที่มีช่วงความถี่ของสัญญาณแตกต่างกัน โดยหนึ่งในรูปภาพนี้จะปรากฏรายละเอียดของภาพเฉพาะในแนวนอน ซึ่งก็คือเส้นอ้างอิงที่ต้องการนั่นเอง ผลการทดลองกับภาพสังเคราะห์ที่สร้างขึ้น และภาพฟอร์มเอกสารที่ใช้งานจริง แสดงให้เห็นกับประสิทธิภาพของอัลกอริทึมดังกล่าว

<b>Thesis Title</b>	Reference Line Extraction in Document Forms Using Wavelet Transform
<b>Student</b>	Mr. Seksan Phonsri
<b>Student ID</b>	40061085
<b>Degree</b>	Master of Engineering
<b>Programme</b>	Electrical Engineering
<b>Year</b>	2544
<b>Thesis Advisor</b>	Asst. Prof. Dr. Punya Thisimajshima

### ABSTRACT

It is common that most documentary forms opt for the use of straight line as a reference position for filled information. Automated data-entry systems of such documents require an ability to search these reference lines so that the location of information in the forms can be known. This thesis proposes an algorithm for extracting the reference line by using wavelet transform. The wavelet decomposition provides 4 different frequency subbands of the input document image. One of these 4 images exhibits the horizontal detail in the image which is essentially match to the reference line that we are looking for. The experimental results on both synthetic and real document images demonstrate the effectiveness of such an algorithm

## กิตติกรรมประกาศ

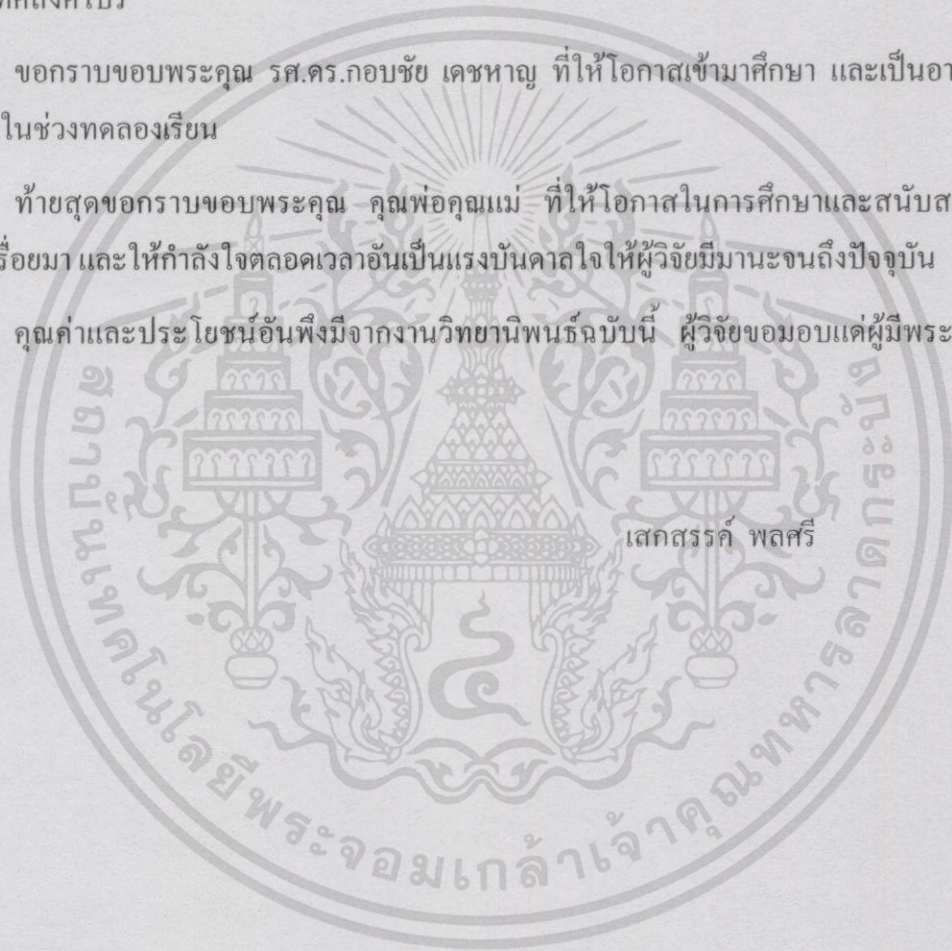
วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี โดยได้รับความกรุณาและคำแนะนำจากหลายฝ่าย ผู้วิจัย  
ขอกราบขอบพระคุณอย่างสูงต่อท่าน ดังต่อไปนี้

ขอกราบขอบพระคุณ ผศ.ดร.ปัญญา จิตมิมชิตา และ รศ.ดร.บุทรพงษ์ รังสรรค์เสรี ที่กรุณา  
ให้คำปรึกษาและแนะนำในด้านการศึกษาและการวิจัย รวมทั้งการเรียบเรียงวิทยานิพนธ์ จนทำให้  
วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยดี และยังสนับสนุนให้ทุนไปนำเสนอผลงานการประชุมวิชาการ  
ที่ประเทศสิงคโปร์

ขอกราบขอบพระคุณ รศ.ดร.กอบชัย เดชหาญ ที่ให้โอกาสเข้ามาศึกษา และเป็นอาจารย์ที่  
ปรึกษาในช่วงทดลองเรียน

ท้ายสุดขอกราบขอบพระคุณ คุณพ่อคุณแม่ ที่ให้โอกาสในการศึกษาและสนับสนุนทุน  
ทรัพย์เรื่อยมา และให้กำลังใจตลอดเวลาอันเป็นแรงบันดาลใจให้ผู้วิจัยมีมานะจนถึงปัจจุบัน

คุณค่าและประโยชน์อันพึงมีจากงานวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุก  
ท่าน



# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญภาพ.....	VII
บทที่ 1 บทนำ.....	1
คำนำ.....	1
1.1 วัตถุประสงค์การวิจัย.....	2
1.2 ขอบเขตของการวิจัย.....	3
บทที่ 2 การแปลงเวฟเล็ท.....	4
2.1 ทฤษฎีพื้นฐานของเวฟเล็ท.....	4
2.1.1 ทฤษฎีพื้นฐานเวฟเล็ทในการสเกลและการเลื่อนตำแหน่ง.....	7
2.1.2 ความเป็นมาของการแปลงเวฟเล็ท.....	8
2.2 การแปลงเวฟเล็ทแบบต่อเนื่อง.....	13
2.3 การแปลงเวฟเล็ทแบบดิสครีท.....	15
2.3.1 การแปลงเวฟเล็ทดิสครีท 1 มิติ.....	16
2.3.2 การแปลงเวฟเล็ทแบบดิสครีทใน 2 มิติ.....	20
2.3.3 เวฟเล็ทแบบออโธ โจนัลและแบบไบออโธ โจนัล.....	24
2.4 การแปลงเวฟเล็ทแบบสเตชันนารี.....	25
บทที่ 3 การค้นหาเส้นอ้างอิงในฟอร์มเอกสาร.....	29
3.1 หลักการค้นหาเส้นอ้างอิง.....	29
3.2 การประยุกต์ใช้งานการแปลงเวฟเล็ทในการค้นหาเส้นอ้างอิง.....	30
3.2.1 การแปลงเวฟเล็ทดิสครีท.....	31
3.2.2 การแปลงเวฟเล็ทสเตชันนารี.....	33

## สารบัญ (ต่อ)

	หน้า
3.3 เทคนิคการทำเรซโซลต์.....	35
3.3.1 การหาค่าเรซโซลต์โดยการกำหนดค่าล่วงหน้า.....	36
3.3.2 การหาค่าเรซโซลต์โดยพิจารณาจากฮิสโตแกรม.....	38
3.3.3 การหาค่าเรซโซลต์จากค่า Root Mean Square (RMS).....	40
3.4 การเข้ารหัสรันเล็งซ์.....	42
บทที่ 4 ผลการทดลอง.....	46
4.1 ผลการทดลอง.....	46
4.2 สรุปผลการทดลอง.....	53
บทที่ 5 บทสรุปและแนวทางในการพัฒนา.....	54
5.1 สรุปผลการวิจัย.....	54
5.2 แนวทางในการพัฒนาต่อไป.....	55
เอกสารอ้างอิง.....	56
ภาคผนวก ผลงานวิจัยที่ได้รับการตีพิมพ์.....	58
ประวัติผู้เขียน.....	59

# สารบัญตาราง

ตารางที่	หน้า
2.1 ตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ทแบบอโอ โทโกนัลชนิด Daubechies ที่มีความยาวของฟิลเตอร์เท่ากับ 4, 6, 8 และ 10 ตามลำดับ.....	25
2.2 เงื่อนไขการแปลงเวฟเล็ทแบบคิสคริท.....	25



# สารบัญรูป

รูปที่	หน้า
1.1 การประมวลผลจากระบบการวิเคราะห์ฟอรัมเอกสาร.....	1
2.1 ตัวอย่างของ Morlet Wavelet: $g(t) = e^{-\frac{t^2}{64}} \cos(t)$ .....	4
2.2 สัญญาณซาย์นุซอซัดัล : $\cos(t)$ .....	5
2.3 ฟังก์ชันหน้าต่าง : $e^{-\frac{t^2}{64}}$ .....	5
2.4 เวฟเล็ตแม่ที่ถูกสเกลเลื่อนตำแหน่งไปที่ค่า $a, b$ ต่างๆ กัน.....	6
2.5 การแยกองค์ประกอบสัญญาณและการรวมองค์ประกอบของสัญญาณเวฟเล็ต.....	7
2.6 สัญญาณที่มีหลายความถี่ที่ตำแหน่งเวลาต่างกัน.....	7
2.7 ระนาบเวลาและความถี่แบบคงที่ (Constant Time-Frequency plan).....	10
2.8 ระนาบเวลาและความถี่สำหรับการแปลงเวฟเล็ต.....	11
2.9 การแยกองค์ประกอบสัญญาณไปยังระดับความละเอียดต่างๆ.....	11
2.10 ฟังก์ชันเวฟเล็ตแม่ชนิดต่างๆ.....	12
2.11 การแปลงเวฟเล็ตในลักษณะของการ inner product.....	14
2.12 การแจกแจงสัมประสิทธิ์ด้วยการแปลงเวฟเล็ต.....	15
2.13 ผลตอบสนองความถี่เชิงอุดมคติของ	
(ก) การกรองความถี่ต่ำ (ข) การกรองความถี่สูง.....	16
2.14 (ก) ตัวดำเนินการลดข้อมูล (ข) การวิเคราะห์สัญญาณด้วยฟิลเตอร์เบงค์.....	17
2.15 (ก) ตัวดำเนินการเพิ่มข้อมูล (ข) การสังเคราะห์สัญญาณด้วยฟิลเตอร์เบงค์.....	17
2.16 โครงสร้างการวิเคราะห์สัญญาณด้วยฟิลเตอร์เบงค์แบบหลายแบนด์.....	19
2.17 โครงสร้างการสังเคราะห์สัญญาณด้วยฟิลเตอร์เบงค์แบบหลายแบนด์.....	20
2.18 การแยกองค์ประกอบของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์เบงค์.....	21
2.19 การรวมองค์ประกอบของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์เบงค์.....	21
2.20 (ก)-(ง) ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต.....	22
2.21 (ก)-(ค) ภาพการแปลงเวฟเล็ตดิสครีท.....	23
2.22 การแยกองค์ประกอบเวฟเล็ตของการแปลงเวฟเล็ตสเตชันนารี.....	26
2.23 (ก)-(ข) ภาพการแปลงเวฟเล็ตสเตชันนารี.....	27

## สารบัญรูป (ต่อ)

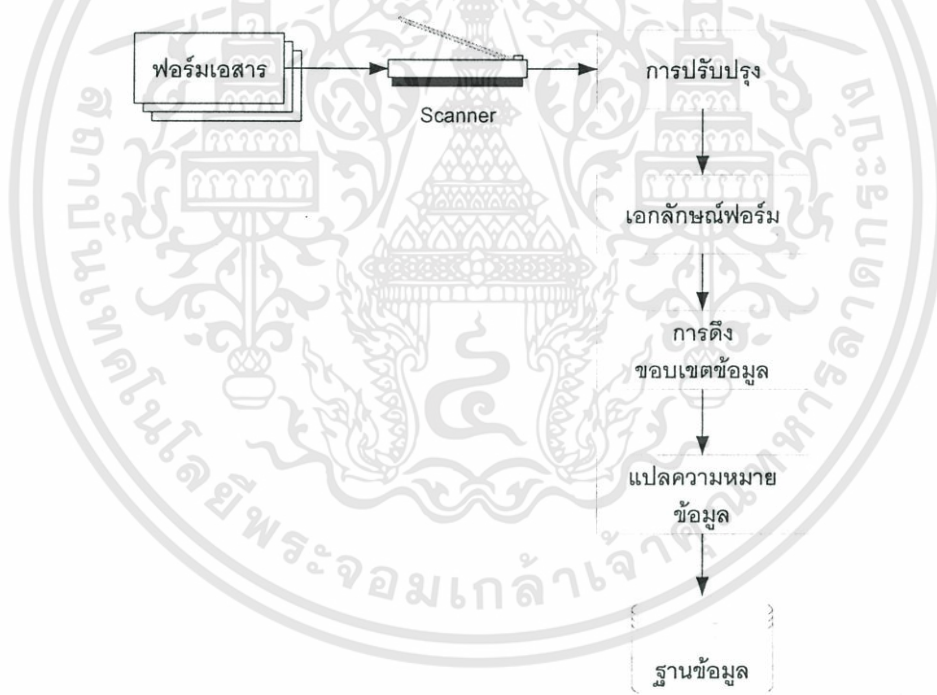
รูปที่	หน้า
3.1	แผนภาพของการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร.....30
3.2	แผนภาพการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ของการแปลงเวฟเล็ทดีสครีท.....32
3.3	ตัวอย่างภาพจากการแปลงเวฟเล็ทดีสครีท (ก) ภาพต้นแบบ (ข) ภาพที่ได้จากการแปลงเวฟเล็ทดีสครีท.....32
3.4	แผนภาพการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ของการแปลงเวฟเล็ทสเตชันนารี.....33
3.5	ตัวอย่างภาพจากการแปลงเวฟเล็ทสเตชันนารี (ก) ภาพต้นแบบ (ข) ภาพที่ได้จากการแปลงเวฟเล็ทสเตชันนารี.....34
3.6	แผนภาพหลักการสร้างภาพไบนารี.....35
3.7	(ก)-(จ) ผลลัพธ์ของการสร้างภาพไบนารีโดยวิธีการกำหนดค่าธรชโซลด์.....37
3.8	แนวความคิดการหาค่าธรชโซลด์ โดยวิธีพิจารณาจากฮิสโตแกรม.....38
3.9	(ก)-(ค) ขั้นตอนการสร้างภาพไบนารีโดยวิธีพิจารณาจากฮิสโตแกรม.....39
3.10	(ก)-(ข) การทำธรชโซลด์ โดยการหาค่าธรชโซลด์จากค่าเฉลี่ย Root Mean Square.....41
3.11	รหัสรันเล็งจ์ของภาพไบนารี.....43
3.12	ตัวอย่างการประยุกต์ใช้งานการเข้ารหัสรันเล็งจ์.....43
3.13	(ก)-(ข) การค้นหาเส้นอ้างอิงโดยการประยุกต์ใช้งานร่วมกันของการแปลง เวฟเล็ทสเตชันนารี เทคนิคการทำธรชโซลด์ และอัลกอริธึมการเข้ารหัสรันเล็งจ์.....45
4.1	(ก)-(ข) ผลการทดลองการค้นหาเส้นอ้างอิง จากข้อมูลภาพใบรับฝากถอนเงินของธนาคารไทยพาณิชย์.....49
4.2	(ก)-(ง) ผลการทดลองการค้นหาเส้นอ้างอิง จากข้อมูลภาพฟอร์มเอกสารเช็คของธนาคารกสิกรไทย.....51
4.3	(ก)-(ง) ผลการทดลองการค้นหาเส้นอ้างอิง จากข้อมูลภาพเช็คของบัญชีของธนาคารไทยพาณิชย์.....53

# บทที่ 1

## บทนำ

### 1.1 คำนำ

เนื่องจากในปัจจุบัน การประมวลผลฟอร์มเอกสาร หรือการจัดการกับข้อมูลในฟอร์มเอกสารเป็นการทำงานที่มีความจำเป็นสำหรับทุกหน่วยงาน การประมวลผลโดยการรับรู้ระบบข้อมูลจากฟอร์มเอกสารจำนวนมากเช่นเช็คธนาคาร สลิปจ่ายเงิน และบิลใบเสร็จต่างๆ ต้องทำการประมวลผลทุกวัน ถ้าฟอร์มเอกสารเหล่านี้สามารถป้อนเข้าไปในคอมพิวเตอร์ และทำการประมวลผลแบบอัตโนมัติ จะทำให้ย่นระยะเวลาในการทำงานลงได้ ดังนั้นการประมวลผลโดยการรับรู้ข้อมูลแบบอัตโนมัติจึงได้เป็นเรื่องสำคัญ ซึ่งหมายถึงการหาวิธีที่จะทำการคัดลอกข้อมูลที่กรอกในฟอร์มเอกสารเพื่อมาทำการจดจำตัวอักษร (OCR) แล้วจัดเก็บลงในระบบฐานข้อมูล



รูปที่ 1.1 การประมวลผลจากระบบการวิเคราะห์ฟอร์มเอกสาร

การประมวลผลจากระบบการวิเคราะห์ฟอร์มเอกสารโดยทั่วไปมีบล็อกไดอะแกรมดังรูปที่ 1.1 ซึ่งประกอบด้วยการสแกนโดยที่การสแกนจะแปลงข้อมูลในกระดาษไปเป็นข้อมูลภาพ อิเล็กทรอนิกส์หรือภาพดิจิทัล การปรับปรุงภาพจะเป็นการประมวลผลต่อมา ซึ่งเป็นการพัฒนาภาพให้ดียิ่งขึ้น เช่นการกำจัดสิ่งไม่พึงประสงค์ในภาพ การทำเรซซโซลด์ การแก้ไขความเอียงของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพเอกสาร ความเป็นเอกลักษณ์ของฟอร์มเอกสารคือการหาโมเดลหรือรูปแบบชนิดของฟอร์มเอกสารที่ต่างจากแบบอื่น ซึ่งจะเป็นการพิจารณาถึงกรอบข้อมูลทุกประเภทที่จำเป็นต่อการประมวลผล โดยที่กรอบข้อมูลนั้นจะประกอบไปด้วยเส้นแนวนอนและแนวตั้งล้อมรอบพื้นที่ขอบเขตของข้อมูลที่ต้องการ การดึงขอบเขตข้อมูลจะเป็นการหาค่าแห่งและบริเวณขอบเขตที่สำคัญของข้อมูลบนฟอร์มเอกสารซึ่งก็คือพื้นที่ภายในกรอบขอบเขตของข้อมูล การแปลความหมายข้อมูลจะเป็นการแปลงบริเวณที่ถูกดึงจากการแสดงภาพไปสู่การแสดงข้อความ เช่น ASCII (American Standard Code for Information Interchange) แล้วทำการจัดเก็บลงฐานข้อมูล ดังนั้นกรอบข้อมูลหรือเส้นอ้างอิงจึงเป็นสิ่งสำคัญในการวิเคราะห์ฟอร์มเอกสารในคอมพิวเตอร์ ทั้งนี้เพราะเส้นตรงหรือเส้นอ้างอิงในภาพฟอร์มเอกสาร คือจุดภาพค่าหลายๆ จุดที่เรียงต่อเนื่องกันไปทั้งในแนวนอนและแนวตั้ง ดังนั้นแม้ว่าตำแหน่งเส้นตรงจะมีการเปลี่ยนแปลงไปอย่างไร เราก็ยังสามารถหาแนวเส้นตรงได้จากเส้นตรงที่หาได้ในภาพก็จะใช้เป็นตัวแทนประเภทของแบบฟอร์มเอกสาร เส้นตรงที่จะใช้ในการจำแนกประเภทแบบฟอร์มมี 2 ประเภทคือเส้นตรงในแนวนอน และเส้นตรงในแนวตั้ง ซึ่งจะพิจารณาถึงความยาวของเส้นตรงแต่ละเส้น ระยะห่างเส้นตรงสองเส้นที่อยู่ใกล้กันและวางอยู่ในแนวเดียวกัน ดังนั้นแนวเส้นตรงจึงเป็นปัจจัยสำคัญในการประมวลผลจากการวิเคราะห์ฟอร์มเอกสาร [1]

ในวิทยานิพนธ์ฉบับนี้ได้ศึกษาในเรื่องการค้นหาเส้นอ้างอิง (Reference Line) ในฟอร์มเอกสาร โดยหลักในการพิจารณาลักษณะสมบัติที่สำคัญในฟอร์มเอกสาร โดยปกติฟอร์มเอกสารจะประกอบด้วยเส้นตรง ซึ่งส่วนใหญ่แล้วจะอยู่ในทิศทางด้านแนวนอน และแนวตั้งเส้นดังกล่าวนี้เรียกว่า “เส้นอ้างอิง” เพื่อที่จะแสดงตำแหน่งในการกรอกข้อมูล ในทางปฏิบัตินั้นการกรอกหรือการเติมข้อมูลจะอยู่ในตำแหน่งด้านบน ด้านล่าง และด้านข้างของเส้นอ้างอิง โดยที่คุณสมบัติของเส้นอ้างอิงนั้นจะพิมพ์เพื่อแนะนำผู้ใช้กรอกข้อมูลในตำแหน่งที่ถูกต้อง และข้อมูลบางส่วนในฟอร์มเอกสารที่นำมาใช้ประโยชน์ จะถูกป้อนเข้าคอมพิวเตอร์ และทำการประมวลผล ดังนั้นการค้นหาเส้นอ้างอิงจึงเป็นเรื่องสำคัญในการประมวลผลฟอร์มเอกสาร [2]

ดังนั้นในวิทยานิพนธ์ฉบับนี้ได้เสนอวิธีตามหลักการของเวฟเล็ตซึ่งใช้การแปลงเวฟเล็ตแบบสแตชันนารี (Stationary Wavelet Transform) โดยภาพฟอร์มเอกสารที่ผ่านการแปลงเวฟเล็ตจะสามารถแบ่งได้เป็นสี่ภาพแบนด์ และจากสองในสี่ภาพนี้สามารถค้นหาเส้นอ้างอิงจากฟอร์มเอกสารที่มีแบคราวด์หลายระดับเทาได้

## 1.2 วัตถุประสงค์ของการวิจัย

เป็นการนำเสนอเทคนิคการค้นหาเส้นอ้างอิงในภาพฟอร์มเอกสารระดับเทาโดยใช้การแปลงเวฟเล็ต (Wavelet Transform) ซึ่งได้ใช้การแปลงเวฟเล็ตแบบสแตชันนารี (Stationary Wavelet Transform) โดยมีคุณสมบัติในการแปลงภาพเอกสารออกเป็นสี่ภาพที่มีช่วงความถี่ของสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แตกต่างกัน ซึ่งภาพที่ได้จะมีขนาดเท่ากับภาพอินพุต และจากสองในสี่ภาพนี้สามารถค้นหาเส้นอ้างอิงจากฟอร์มเอกสารที่มีแบคกราวด์ระดับเทาได้ พร้อมทั้งยังสามารถกำจัดแบคกราวด์ระดับเทาจากภาพฟอร์มเอกสารได้ แต่ภาพผลลัพธ์ยังคงมีที่สิ่งไม่พึงประสงค์ปะปนมาด้วย (noise) จึงได้นำการสร้างภาพไบนารีและการเข้ารหัสรันเล็งซ์เข้ามาช่วยเพื่อทำการจำแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบคกราวด์ และปรับปรุงภาพโดยกำจัดสิ่งไม่พึงประสงค์ที่ปะปนมาบางส่วนทิ้ง ซึ่งจะทำให้ข้อมูลเส้นอ้างอิงให้ดูเด่นชัดมากยิ่งขึ้น

### 1.3 ขอบเขตของการวิจัย

ในวิทยานิพนธ์ฉบับนี้ ได้นำเสนอเทคนิคการค้นหาเส้นอ้างอิงในฟอร์มเอกสารโดยใช้การแปลงเวฟเล็ต โดยเนื้อหารายละเอียดการวิจัยแยกออกได้เป็น 5 บทดังต่อไปนี้

บทที่ 1 บทนำกล่าวถึงความเป็นมาและความสำคัญของงานวิจัย วัตถุประสงค์ของการทำวิจัย และเนื้อหาของวิทยานิพนธ์ บทที่ 2 การแปลงเวฟเล็ตกล่าวถึงทฤษฎีพื้นฐานของการแปลงเวฟเล็ต การแปลงเวฟเล็ตที่เวลาต่อเนื่อง การแปลงเวฟเล็ตแบบดิสครีต และการแปลงเวฟเล็ตสเตชันนารี บทที่ 3 การค้นหาเส้นอ้างอิงกล่าวถึงความสำคัญของเส้นอ้างอิง และหลักวิธีการในการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร โดยการนำเทคนิคการทำเรซโซลต์ และอัลกอริธึมของการเข้ารหัสรันเล็งซ์มาประยุกต์ใช้งานร่วมกันกับการแปลงเวฟเล็ตสเตชันนารีเพื่อค้นหาเส้นอ้างอิง บทที่ 4 ผลการทดลองกล่าวถึงการทดลองและผลการทดลอง ที่ได้จากการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร โดยได้นำเซตธนาคารที่มีรายละเอียดของเส้นอ้างอิงมาทำการสแกนเป็นภาพระดับเทา แล้วนำภาพที่ได้มาทำการทดลองโดยผ่านการแปลงเวฟเล็ต และปรับปรุงรายละเอียดภาพเส้นอ้างอิงด้วยเทคนิคการทำเรซโซลต์ และอัลกอริธึมของการเข้ารหัสรันเล็งซ์ บทที่ 5 บทสรุปและแนวทางในการพัฒนา กล่าวถึงผลสรุปจากการวิจัย และแนวทางในการพัฒนาต่อไป

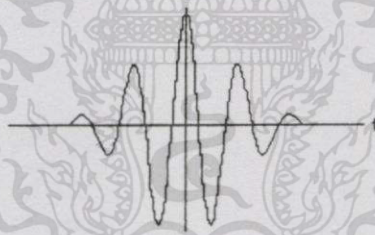
## บทที่ 2

### การแปลงเวฟเล็ท

บทนี้จะกล่าวถึงพื้นฐานการแปลงเวฟเล็ท ความเป็นมาของการแปลงเวฟเล็ท คุณสมบัติของการแปลงเวฟเล็ทแบบดิสครีท (Discrete Wavelet Transform : DWT) 1 มิติและ 2 มิติ ในลักษณะของฟิลเตอร์แบงก์ (Filter bank) การกำหนดผลตอบสนองอิมพัลส์ของฟิลเตอร์ที่นำมาใช้ในขบวนการ DWT ผลตอบสนองอิมพัลส์ที่ได้ทำให้ทราบคุณสมบัติของเวฟเล็ทที่ใช้ คุณสมบัติและเงื่อนไขของเวฟเล็ทแบบออร์โธโกนัล (Orthogonal) และแบบไบออร์โธโกนัล (Biorthogonal) ที่นิยมใช้ทั่วไปในงานด้านการประมวลผลภาพ และการแปลงเวฟเล็ทสเตชันนารี (Stationary Wavelet Transform: SWT)

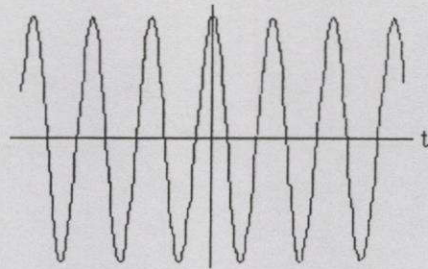
#### 2.1 ทฤษฎีพื้นฐานของเวฟเล็ท

ทฤษฎีเวฟเล็ท (Wavelet Theory) เป็นคณิตศาสตร์ที่ใช้อธิบายการสร้างโมเดลของสัญญาณระบบ ซึ่งประกอบขึ้นด้วยเซตของสัญญาณเฉพาะที่มีลักษณะเป็นคลื่นเล็กๆ เรียกว่า “เวฟเล็ท” โดยลักษณะของเวฟเล็ทจะเป็นคลื่นที่มีการเปลี่ยนแปลงอย่างต่อเนื่อง (Oscillatory) และขนาดของคลื่นจะลดลงสู่ศูนย์อย่างรวดเร็วทั้งสองด้าน ดังแสดงในรูปที่ 2.1

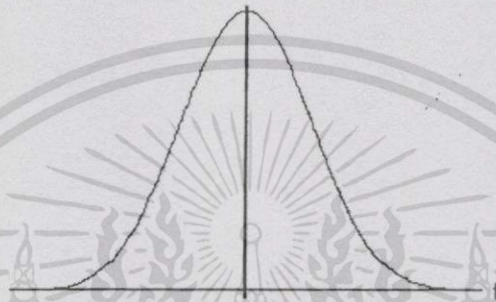


รูปที่ 2.1 ตัวอย่างของ Morlet Wavelet:  $g(t) = e^{-\frac{t^2}{64}} \cos(t)$

ในรูปที่ 2.1 เป็นตัวอย่างของเวฟเล็ทชนิดหนึ่ง เรียกว่าเวฟเล็ทแบบ Morlet จะสังเกตเห็นว่ามีคุณสมบัติของเวฟเล็ทตามที่กล่าวมาทั้ง 2 เงื่อนไข คือมีการออสซิลเลทในลักษณะของสัญญาณไซน์ซอยดัล (Sinusoidal) ดังแสดงในรูปที่ 2.2 รวมกับเงื่อนไขการลดลงอย่างรวดเร็วตามลักษณะฟังก์ชันหน้าต่าง (Windows Function) ดังแสดงในรูปที่ 2.3



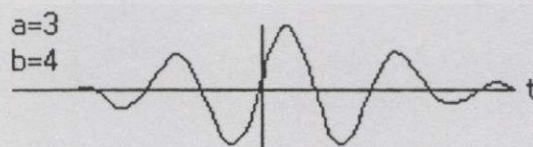
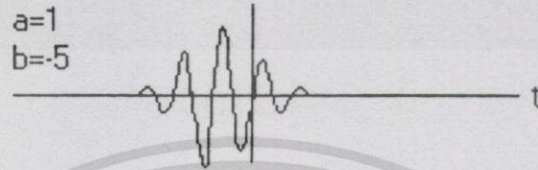
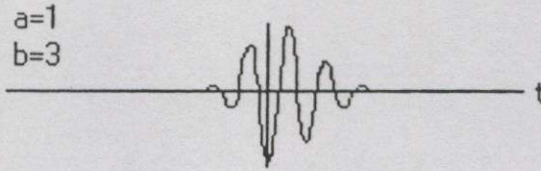
รูปที่ 2.2 สัญญาณไซน์ซอซัด :  $\cos(t)$



รูปที่ 2.3 ฟังก์ชันหน้าต่าง :  $e^{-\frac{t^2}{64}}$

การนำเวฟเล็ตหลายๆ อันมารวมกันเป็นเซต เพื่อจะใช้ในการอธิบายสัญญาณใดๆ โดยที่คลื่นเวฟเล็ตแต่ละตัวมีโครงสร้างมาจากฟังก์ชันเดียวกัน ซึ่งฟังก์ชันนี้จะเป็นเวฟเล็ตต้นกำเนิดที่เรียกว่า เวฟเล็ตแม่ (Mother Wavelets) [3] [4] โดยคลื่นแต่ละคลื่นภายในเซตของเวฟเล็ตจะเกิดจากการสเกลและการเลื่อนตำแหน่งของเวฟเล็ตแม่ การสเกลนี้หมายถึงการหดเข้า (Compressing) หรือการขยายออก (Dilation) โดยใช้ " $a$ " แทนพารามิเตอร์ของการหดหรือการขยายอันเป็นการเปลี่ยนความถี่ดังกล่าว ส่วนการเลื่อนตำแหน่งกำหนดโดยพารามิเตอร์ " $b$ " ซึ่งหมายถึง การเลื่อนตำแหน่งของการเกิดคลื่นบนแกนเวลา ในรูปที่ 2.4 แสดงให้เห็นถึงเวฟเล็ตที่ถูกสเกลและเลื่อนตำแหน่งไปที่ค่า " $a$ ", " $b$ " ต่างๆ กัน โดยที่ยังคงอยู่ภายในเซตของเวฟเล็ตที่สัมพันธ์กับเวฟเล็ตแม่ในรูปที่ 2.1 นอกจากนี้เวฟเล็ตที่ถูกสเกลจะต้องคูณออร์มอลไลซ์ด้วย  $\frac{1}{\sqrt{|a|}}$  เสมอ เพื่อให้เวฟเล็ตที่ถูกสเกลไปแล้วยังคงมีพลังงานเช่นเดียวกับเวฟเล็ตแม่ ดังนั้น ถ้าให้  $\psi(t)$  เป็นฟังก์ชันเวฟเล็ตแม่จะสามารถเขียนสมการแสดงความสัมพันธ์ที่ตำแหน่ง " $a$ ", " $b$ " ใดๆ ได้ดังนี้ [5]

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2.1)$$

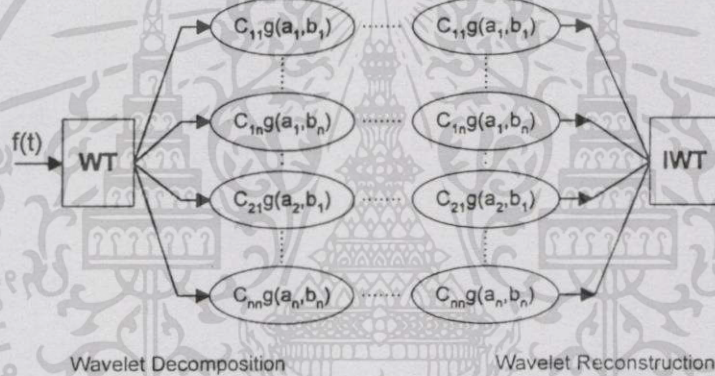


รูปที่ 2.4 เวกเก็ตแม่ที่ถูกลดตำแหน่งไปที่ค่า  $a, b$  ต่างๆ กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

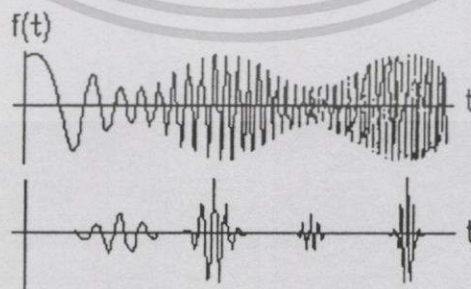
### 2.1.1 ทฤษฎีพื้นฐานเวฟเล็ตในการสเกลและการเลื่อนตำแหน่ง

การอธิบายสิ่งหนึ่งสิ่งใดในทฤษฎีเวฟเล็ตจะเปรียบเสมือนกับการแตกสิ่งเหล่านั้นออกเป็น ส่วนประกอบเล็กๆ ที่สัมพันธ์กันคล้ายกับภาพต่อจิ๊กซอว์ โดยที่ชิ้นส่วนเหล่านั้นจะอยู่ในรูปของ เวฟเล็ตที่ถูกสเกลและเลื่อนตำแหน่ง และมีค่าถ่วงน้ำหนักต่างๆ กันคุณอยู่กับเวฟเล็ตแต่ละตัว ดังนั้นจึงเปรียบเสมือนว่าคลื่นเวฟเล็ตเป็นฟังก์ชันพื้นฐาน (Basis function: b.f.) ในการสร้างสัญญาณ ใดๆ โดยขบวนการแตกกระจายชิ้นส่วนนี้เรียกว่า “การแยกองค์ประกอบเวฟเล็ต” (Wavelet Decomposition) ซึ่งก็คือขั้นตอนในการแปลงเวฟเล็ตนั่นเอง ในทำนองเดียวกันการรวมองค์ ประกอบเวฟเล็ต (Wavelet Reconstruction) ก็คือการแปลงกลับเวฟเล็ต (Inverse Wavelet Transform: IWT) โดยจะเป็นการนำเอาส่วนประกอบย่อยๆ เหล่านี้มาเรียงต่อกันเพื่อประกอบเป็น สัญญาณเดิม ดังแสดงในรูปที่ 2.5 แสดงการแยกองค์ประกอบสัญญาณ  $f(t)$  ออกเป็นคลื่นเวฟเล็ต ที่ความถี่และตำแหน่งต่างกันจำนวน  $n \times n$  ตัว



รูปที่ 2.5 การแยกองค์ประกอบสัญญาณและการรวมองค์ประกอบของสัญญาณเวฟเล็ต

ในรูปที่ 2.6 รูปบนแสดงสัญญาณ  $f(t)$  ที่ประกอบด้วยคลื่นหลายความถี่ ที่ตำแหน่งต่างๆ บนแกนเวลา ส่วนในรูปด้านล่างแสดงถึงการดึงสัญญาณ  $f(t)$  ที่ตำแหน่งเวลาใดๆ ออกมา



รูปที่ 2.6 สัญญาณที่มีหลายความถี่ที่ตำแหน่งเวลาต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการที่สามารถทำการสร้างเซตของเวฟเล็ตที่ได้จากการสเกลและการเลื่อนตำแหน่งของเวฟเล็ตเมื่อนั้น การสเกลและการเลื่อนตำแหน่งดังกล่าวจะถูกเรียกรวมกันว่า Affine Operation [6] ซึ่งจะสามารถเขียนเป็นสมการคณิตศาสตร์ในลักษณะการส่ง (Mapping) ตัวแปรอิสระ  $t$  ไปยังตัวแปรอิสระใหม่  $t'$  ดังแสดงในสมการที่ 2.2

$$t' = \frac{t-b}{a} \quad (2.2)$$

และจะสามารถเขียนสมการสำหรับการแปลงเวฟเล็ตแบบต่อเนื่อง (Continuous Wavelet Transform: CWT) ที่ใช้ในการกระจายสัญญาณของฟังก์ชัน  $f(t)$  ด้วยเวฟเล็ตแม่  $\psi(t)$  ได้ดังนี้ [7]

$$W(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad (2.3)$$

โดยที่ "\*" หมายถึงคอมเพล็กซ์คอนจูเกต " $a$ " เป็นพารามิเตอร์ของการสเกล และ " $b$ " เป็นพารามิเตอร์ของการเลื่อนตำแหน่ง สัญลักษณ์  $W(a, b)$  หมายถึงสัญญาณอินพุตที่ขึ้นกับการส่งตัวแปร  $t$  ไปเป็นฟังก์ชัน 2 มิติที่ขึ้นกับตัวแปรอิสระ  $a$  และ  $b$  โดย  $W(a, b)$  จะเป็นค่าสัมประสิทธิ์ของเวฟเล็ต (Wavelet Coefficient)

### 2.1.2 ความเป็นมาของการแปลงเวฟเล็ต

การแปลงฟูรีเยร์นิยมใช้งานอย่างแพร่หลายซึ่งเป็นการแปลงสัญญาณ  $f(t)$  จากโดเมนเวลาไปเป็นสัญญาณ  $F(\omega)$  ซึ่งอยู่ในโดเมนความถี่หรือเรียกกันทั่วไปว่าเป็นการหาสเปกตรัมของสัญญาณ  $f(t)$  นั้นเองโดยมีสมการเป็น

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (2.4)$$

จะสังเกตเห็นว่าข้อมูลที่ได้จากการทรานส์ฟอร์ม จะมีตลอดย่านความถี่  $-\infty$  ถึง  $+\infty$  ดังนั้นเมื่อเกิดการเปลี่ยนแปลงของสัญญาณอินพุตที่ระยะเวลาใดๆ ก็จะส่งผลกระทบต่อตลอดย่านของสเปกตรัม  $F(\omega)$  ด้วย นอกจากนี้การแปลงฟูรีเยร์ยังไม่สามารถเลือกทำการวิเคราะห์สัญญาณเฉพาะช่วงเวลาหรือความถี่ใดๆ ได้ ดังนั้นวิธีการแปลงฟูรีเยร์จึงให้ข้อมูลที่มากเกินไปนอกเหนือจากนี้ในทางปฏิบัติหากสัญญาณอินพุตจะมีการเปลี่ยนแปลงที่ไม่คงที่ (Non-Stationary Signal) การใช้การแปลงฟูรีเยร์จะทำให้สิ้นเปลืองเวลาในการคำนวณเนื่องจากมีข้อมูลมากเกินไปในการ

คำนวณใหม่ในแต่ละครั้ง ดังนั้นการแปลงฟูรีเยร์จึงเหมาะสำหรับสัญญาณที่มีการเปลี่ยนแปลงคงที่ (Stationary Signal)

การแปลงการบอร์ (Gabor Transform) เป็นการแปลงที่ได้พัฒนาขึ้นมาเป็นแบบแรกเพื่อแก้ปัญหาที่กล่าวมาข้างต้น โดยจะเป็นการทำการแปลงฟูรีเยร์เฉพาะช่วงเวลาและช่วงความถี่ที่กำหนดผ่านฟังก์ชันหน้าต่างต่าง (Window function: wf) ซึ่งหน้าต่างนี้จะใช้ฟังก์ชันแบบเกาส์เซียน (Gaussian function) เท่านั้น

$$g_a(t) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{t^2}{2a}} \quad (2.5)$$

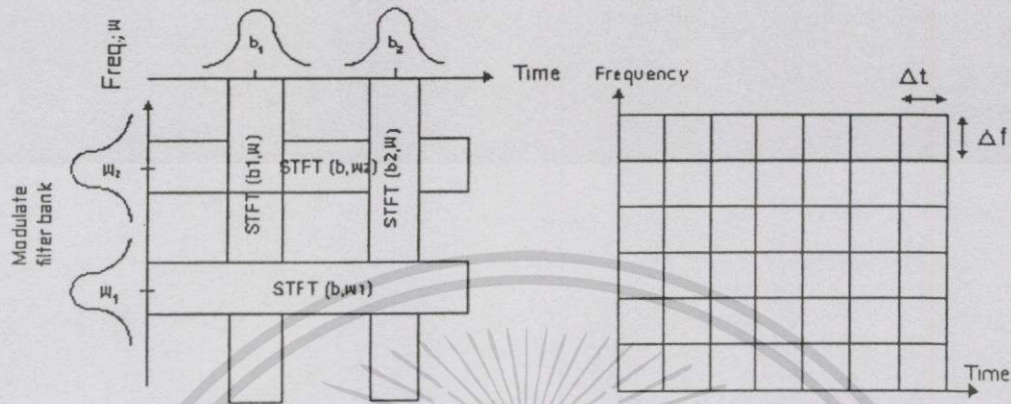
$$G_{a,b}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} f(t) g_a(t-b) dt \quad (2.6)$$

จากสมการที่ (2.5) เป็นฟังก์ชันหน้าต่างแบบเกาส์เซียน และสมการที่ (2.6) เป็นสมการของการแปลงการบอร์ จะเห็นว่าการทรานส์ฟอร์มแบบนี้จะสามารถเลือกตำแหน่งการวิเคราะห์สัญญาณ โดยกำหนดที่พารามิเตอร์  $b$  และเลือกช่วงความถี่ด้วยพารามิเตอร์  $a$  แต่เนื่องจากการแปลงการบอร์ให้ฟังก์ชันหน้าต่างแบบเดียวจึงอาจไม่เหมาะสมกับการประยุกต์ใช้งานในด้านต่างๆ ดังนั้นจึงได้มีการพัฒนาต่อมาเป็นการแปลงฟูรีเยร์ช่วงสั้น (Short time Fourier Transform: STFT) ซึ่งจะสามารถทำการเลือกฟังก์ชันหน้าต่างที่จะใช้ในการวิเคราะห์สัญญาณได้ ทำให้เกิดความยืดหยุ่นในการเลือกใช้ให้เหมาะสมกับสัญญาณอินพุตที่จะวิเคราะห์ได้ดียิ่งขึ้น

$$G_a(\omega) = \int_{-\infty}^{\infty} (e^{-j\omega t} f(t)) \overline{w(t-b)} dt \quad (2.7)$$

สมการที่ (2.7) จะเป็นการแปลงฟูรีเยร์ช่วงสั้น โดยที่ตำแหน่งเวลาในการวิเคราะห์จะกำหนดโดยค่า " $b$ " และช่วงความถี่ในการวิเคราะห์จะกำหนดด้วยความถี่ของฟังก์ชันหน้าต่างนั้นๆ แต่จะสังเกตเห็นว่าทั้งการแปลงการบอร์และการแปลงฟูรีเยร์ช่วงเวลาสั้น จะมีลักษณะของช่วงการวิเคราะห์ที่คงที่ ดังแสดงในรูปที่ 2.7 แสดงช่วงเวลาและช่วงความถี่ที่คงที่ (Fixed resolution transform) ทั้งนี้ฟังก์ชันหน้าต่างไม่มีการเปลี่ยนแปลงคุณสมบัติขณะทำการแปลงซึ่งไม่เหมาะสมในทางปฏิบัติ เนื่องจากสัญญาณที่มีความถี่สูงจะมีเวลาที่แคบ ในขณะที่สัญญาณความถี่ต่ำจะมีช่วงคาบเวลาที่กว้างกว่า ดังนั้นการแปลงการบอร์และการแปลงฟูรีเยร์ช่วงเวลาสั้น อาจจะทำให้ข้อมูลที่ได้ออกมาจากการแปลงออกมามากหรือน้อยเกินไป ถ้าเป็นในกรณีหลังจะทำให้ต้องมีการแปลงจากหลายช่วงเวลาแล้วนำมาประกอบกัน จึงจะได้ข้อมูลครบถ้วนในช่วงเวลาที่ต้องการ ดังนั้นช่วงเวลาควรจะ

ต้องมีการเปลี่ยนแปลงสัมพันธ์กับช่วงความถี่ที่จะทำการวิเคราะห์ นั่นคือในช่วงความถี่ที่สูงขึ้นจะมีช่วงเวลาในการวิเคราะห์ที่น้อยลง และในช่วงความถี่ที่ต่ำลงจะใช้ช่วงเวลาในการวิเคราะห์ที่มากขึ้น [8] [9]

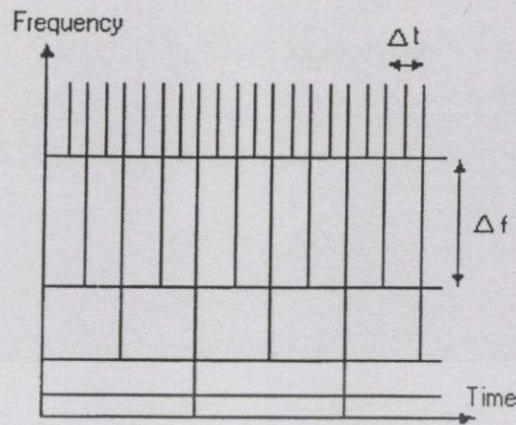


รูปที่ 2.7 ระยะเวลาและความถี่แบบคงที่ (Constant Time-Frequency plan)

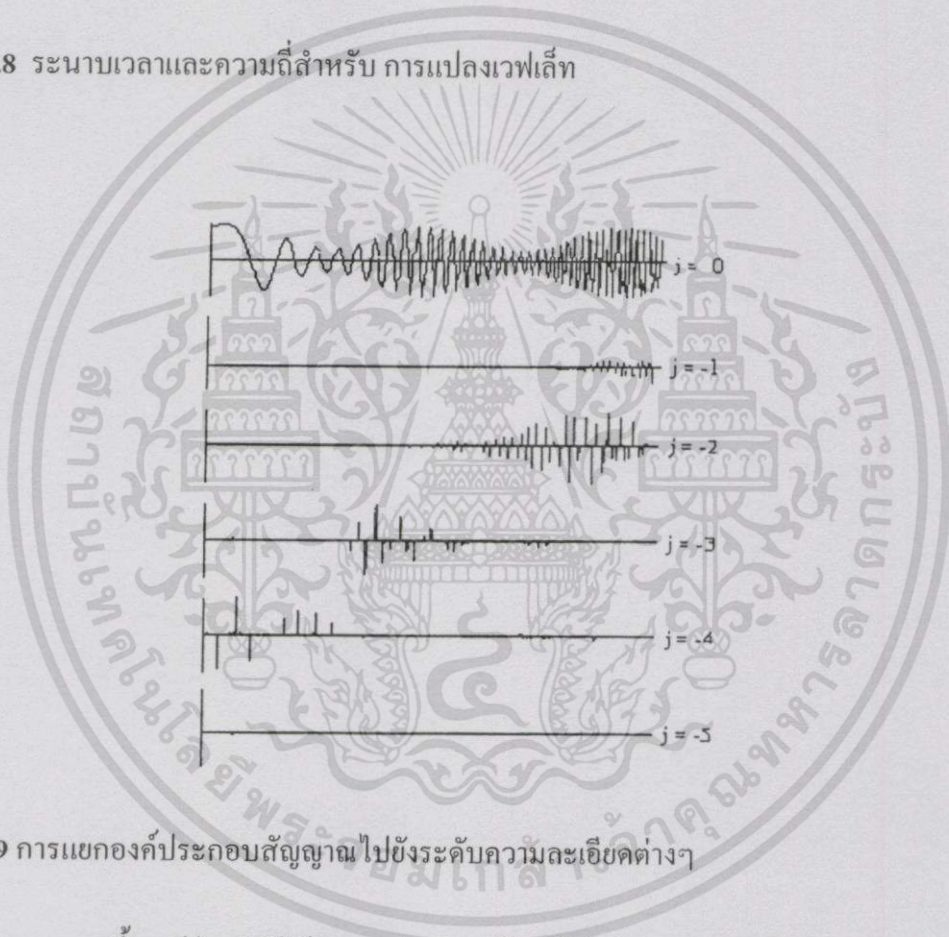
จากเหตุผลดังกล่าวข้างต้นทำให้มีการพัฒนาเป็นการแปลงในรูปแบบใหม่ที่เรียกว่า “การแปลงเวฟเล็ต” ซึ่งจะเป็นการแปลงผ่านฟังก์ชันหน้าต่าง โดยฟังก์ชันหน้าต่างสามารถเลื่อนตำแหน่งและปรับเปลี่ยนช่วงเวลาให้เหมาะสมกับช่วงความถี่ที่จะวิเคราะห์ได้ โดยเรียกฟังก์ชันหน้าต่างแบบนี้ว่า “ฟังก์ชันเวฟเล็ต” (Wavelet function)

$$W(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad (2.8)$$

$\psi(t)$  เป็นฟังก์ชันเวฟเล็ตแม่ ซึ่งจะถูกลดและเลื่อนตำแหน่งโดยพารามิเตอร์ "a", "b" ตามลำดับ โดยช่วงเวลาและความถี่ในการทรานส์ฟอร์มจะสัมพันธ์กันดังที่ได้กล่าวมาแล้วข้างต้น พิจารณารูปที่ 2.8 จะพบว่าระยะเวลาและความถี่จะมีลักษณะการเปลี่ยนแปลงของช่วงเวลา และความถี่ที่สัมพันธ์กัน ในช่วงความถี่สูงจะใช้ช่วงเวลาในการวิเคราะห์เพียงสั้นๆ ( $\Delta f$  สูง  $\Delta t$  ต่ำ) แต่ในช่วงความถี่ต่ำจะมีช่วงเวลาในการวิเคราะห์ที่นานกว่า ( $\Delta f$  ต่ำ  $\Delta t$  สูง) [10] ซึ่งจะเหมาะสมกับลักษณะสัญญาณในทางปฏิบัติ ในรูปที่ 2.9 จะเป็นการแสดงลักษณะการแยกองค์ประกอบสัญญาณออกไปในระดับความถี่ที่ต่างกัน โดยในระดับที่  $j = -1$  จะมีเฉพาะส่วนความถี่สูง และในระดับที่  $j$  ต่ำลง สัญญาณจะมีค่าความถี่ที่ต่ำลงด้วย ซึ่งจะเห็นว่าสอดคล้องกับรูปของระยะเวลาและความถี่



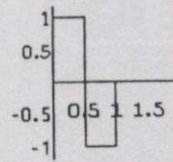
รูปที่ 2.8 ระยะเวลาและความถี่สำหรับการแปลงเวฟเล็ท



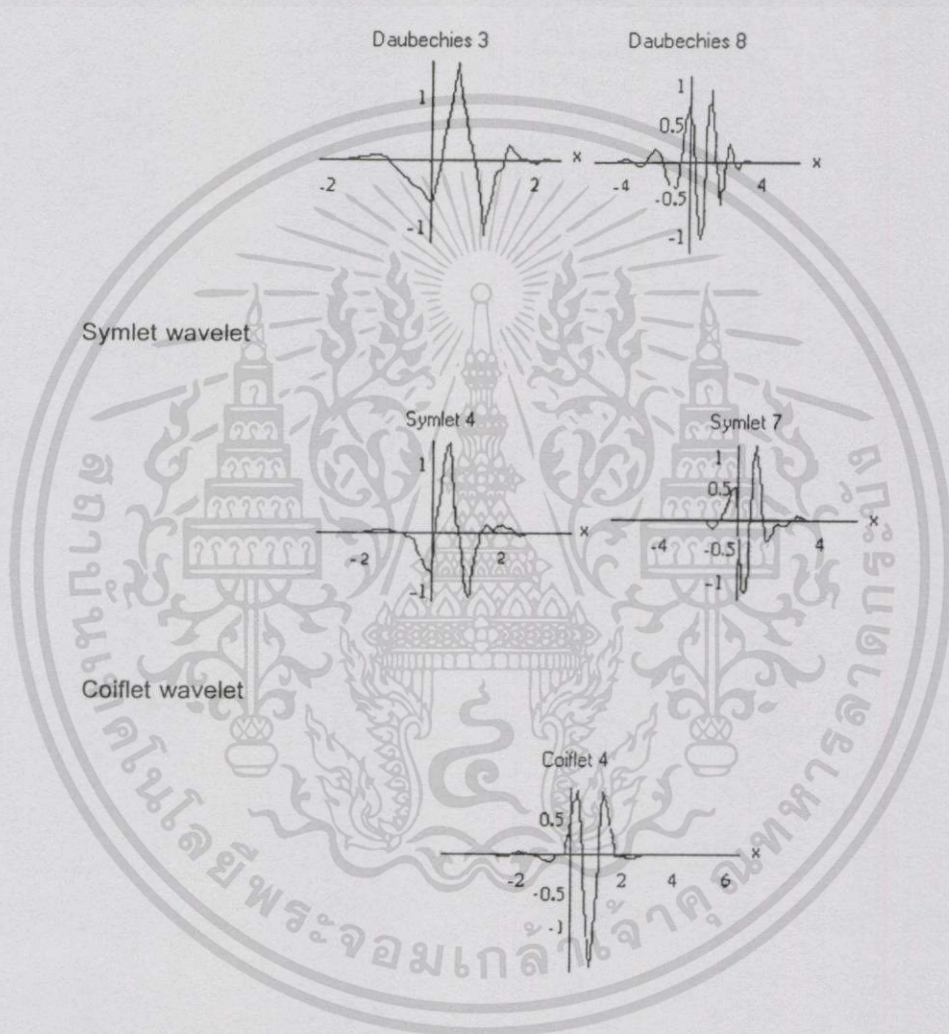
รูปที่ 2.9 การแยกองค์ประกอบสัญญาณไปยังระดับความละเอียดต่างๆ

นอกจากนี้  $\psi(t)$  ยังมีได้หลายรูปแบบแต่ต้องเป็นไปตามคุณสมบัติของการแปลงเวฟเล็ท ดังนั้นจึงทำให้การแปลงเวฟเล็ทมีความยืดหยุ่นและสามารถนำไปประยุกต์ใช้กับงานที่มีความเหมาะสมได้เป็นอย่างดีดังรูปที่ 2.10 แสดงลักษณะของฟังก์ชันเวฟเล็ทแม่ชนิดต่างๆ ที่นิยมใช้ในปัจจุบัน

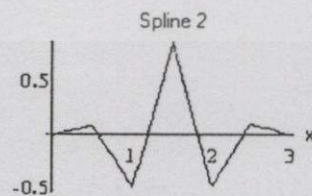
Harr wavelet



Daubechies wavelet



Spline wavelet



รูปที่ 2.10 ฟังก์ชันเวฟเล็ทแม่ชนิดต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2 การแปลงเวฟเล็ทแบบต่อเนื่อง

จากการวิเคราะห์สัญญาณด้วยการแปลงฟูรีเยร์ช่วงสั้น จะสังเกตเห็นว่าหน้าต่างเวลา-ความถี่ (Time-frequency window) จะมีพื้นที่ในการวิเคราะห์ที่คงที่เสมอไม่ว่าจะเป็นการพิจารณาที่ความถี่ในช่วงใดๆ ขนาดช่วงความกว้างของความถี่และช่วงเวลาจะขึ้นอยู่กับลักษณะของฟังก์ชันหน้าต่างที่ใช้ ซึ่งจะไม่มี การเปลี่ยนแปลงคุณสมบัติในขณะการแปลง แต่ในทางปฏิบัติความถี่ของสัญญาณจะเป็นสัดส่วนโดยตรงกับจำนวนไซเคิลต่อหนึ่งหน่วยเวลา ที่ความถี่สูงจะมีการเปลี่ยนแปลงสัญญาณในช่วงเวลาที่แคบ ส่วนการเปลี่ยนแปลงสัญญาณในช่วงความถี่ต่ำจะใช้ช่วงเวลานานกว่า ดังนั้นถ้าใช้การแปลงฟูรีเยร์ช่วงสั้น และเลือกฟังก์ชันหน้าต่างที่มีช่วงความถี่และช่วงเวลาที่ไม่เหมาะสมจะทำให้ต้องมีการทำการแปลงหลายครั้งที่หลายตำแหน่งเพื่อให้ได้ข้อมูลในช่วงเวลาทั้งหมดที่ต้องการ หรือถ้าความกว้างของช่วงเวลาของฟังก์ชันหน้าต่างมากเกินไปจะทำให้ได้รับข้อมูลมากเกินไปจนเกินไป ดังนั้นฟังก์ชันหน้าต่างที่ดี ความกว้างของช่วงเวลาและช่วงความถี่จะต้องมีการเปลี่ยนแปลงเป็นสัดส่วนผกผันที่สัมพันธ์กัน นั่นคือเมื่อทำการวิเคราะห์สัญญาณที่ความถี่สูง ความกว้างของช่วงเวลาควรจะแคบและมีช่วงความถี่ที่กว้าง ในทางกลับกันถ้าวิเคราะห์สัญญาณที่ความถี่ต่ำ ความกว้างของช่วงเวลาควรจะกว้างและมีช่วงความถี่ที่แคบ เพื่อที่จะทำให้ได้ระดับข้อผิดพลาดที่ครบถ้วนและใช้ช่วงเวลาในการแปลงน้อยลง

จากที่กล่าวมาทั้งหมดจะเห็นได้ว่าการแปลงฟูรีเยร์ช่วงสั้น ไม่เหมาะสมสำหรับการวิเคราะห์สัญญาณที่มีการเปลี่ยนความถี่สูงต่ำในสัญญาณตัวเดียวกัน แต่ในการแปลงเวฟเล็ท จะอาศัยฟังก์ชันหน้าต่าง ซึ่งสามารถเปลี่ยนแปลงคุณสมบัติของช่วงเวลาและช่วงความถี่ให้เหมาะสมกับการวิเคราะห์สัญญาณในช่วงความถี่และช่วงเวลาที่น่าสนใจใดๆ ซึ่งฟังก์ชันหน้าต่างที่มีคุณสมบัติดังกล่าวเรียกว่า "ฟังก์ชันเวฟเล็ท"

กำหนดให้  $\psi(t) \in L^2(\mathbb{R})$  และเป็นไปตาม Admissibility condition [11] คือ

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \quad (2.9)$$

เรียก  $\psi(t)$  ว่าเป็นรากฐานเวฟเล็ท (Basis wavelet) หรือเวฟเล็ทแม่ ซึ่งมีคุณสมบัติในการเลื่อนตำแหน่งและการเปลี่ยนความถี่ได้ ดังแสดงในสมการที่ (2.10)

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2.10)$$

และเมื่อนำมาเขียนเป็นสมการการแปลงฟูรีเยร์ที่พิจารณาผ่านฟังก์ชันหน้าต่าง ที่ตำแหน่งเวลาและความถี่ในขณะนั้น ได้ดังสมการที่ (2.11) ซึ่งก็คือการแปลงเวฟเล็ทแบบต่อเนื่องนั่นเอง

$$W_\psi(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|a|}} \psi^* \left( \frac{t-b}{a} \right) dt \tag{2.11}$$

โดยที่  $a, b \in \mathbb{R}, a \neq 0$  และ  $f(t) \in L^2(\mathbb{R})$

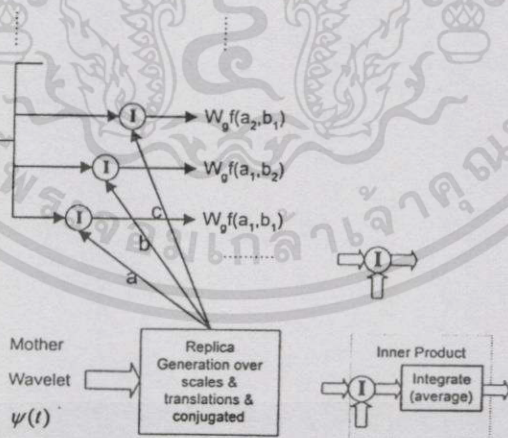
เทอม  $\frac{1}{\sqrt{|a|}}$  จะเป็นการนอร์มอลไลซ์ (Normalization) ของ  $\psi(t)$  เพื่อให้พลังงานของ  $\psi(t)$  ที่  $a, b$  ใดๆ ยังคงมีค่าเท่ากับหนึ่งเสมอ ทำให้การแปลงเวฟเล็ตที่ตำแหน่ง  $a, b$  ใดๆ พลังงานของ  $\psi(t)$  จะไม่มีผลกระทบต่อสัญญาณ  $f(t)$  ในช่วงนั้นๆ และจะสามารถเขียนสมการแปลงเวฟเล็ตแบบต่อเนื่องในลักษณะของ inner product ได้ดังนี้

$$W_\psi(a, b) = \langle f, \psi_{a,b} \rangle \tag{2.12}$$

สังเกตว่าจากการกำหนดให้  $\psi(t) \in L^2 \mathbb{R}$  นั่นคือ  $\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$  แสดงว่า  $\psi(t)$  จะต้องมีขนาดและพลังงานได้เสมอ เพราะฉะนั้น  $\psi(t)$  จะมีการลดลงสู่ศูนย์ทั้งสองด้าน ถ้าให้  $\hat{\psi}(0) = 0$  ซึ่งจะต้องมีเงื่อนไขอีกอันหนึ่งเกิดขึ้นคือ

$$\begin{aligned} \hat{\psi}(\omega) &= \int_{-\infty}^{\infty} e^{-i\omega t} \psi(t) dt \\ \hat{\psi}(0) &= \int_{-\infty}^{\infty} \psi(t) dt = 0 \end{aligned} \tag{2.13}$$

แสดงว่าพื้นที่ใต้กราฟของ  $\psi(t) = 0$  นั่นก็คือฟังก์ชันดังกล่าวต้องมีการออกสขเลขเสมอ



หมายเหตุ  
 $a = \psi^*[(t-b_1)/a_1]$   
 $b = \psi^*[(t-b_2)/a_1]$   
 $c = \psi^*[(t-b_2)/a_2]$

รูปที่ 2.11 การแปลงเวฟเล็ตในลักษณะของการ inner product

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าพิจารณาหน้าต่างเวลา-ความถี่ ของการแปลงเวฟเล็ท โดยที่  $\psi(t), \hat{\psi}(\omega)$  เป็นไปตามเงื่อนไข  $tW(t) \in L^2(\mathbb{R})$  และกำหนดให้จุดศูนย์กลางและรัศมีความกว้างของ  $\psi(t)$  เป็น  $t^*$  และ  $\Delta\psi$  ตามลำดับ ดังนั้น  $\psi_{a,b}$  จะมีจุดศูนย์กลางอยู่ที่  $b+at^*$  และรัศมีเป็น  $a\Delta\psi$  โดยข้อมูลของสัญญาณ  $f(t)$  จะถูกพิจารณาผ่าน Time-window ที่มีความกว้างของช่วงเวลาดังนี้ [9]

$$[b+at^* - a\Delta\psi, b+at^* + a\Delta\psi] \quad (2.14)$$

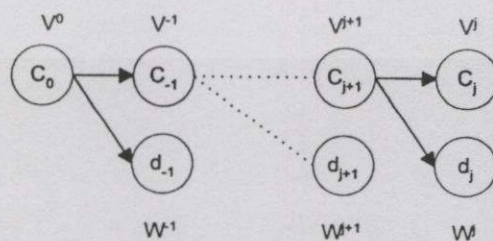
### 2.3 การแปลงเวฟเล็ทแบบดิสครีท

ในหัวข้อนี้จะเป็นการอธิบายถึง การแปลงเวฟเล็ทแบบดิสครีท (Discrete Wavelet Transform: DWT) โดยจะทำการพิจารณาการแปลงเวฟเล็ทแบบดิสครีทในลักษณะของ Filter banks ในทางปฏิบัติสัญญาณที่วิเคราะห์ด้วยคอมพิวเตอร์จะเป็นการสุ่ม (Sampling) สัญญาณเข้ามาในจำนวนที่จำกัด ดังนั้นถ้ากำหนดให้  $c_0(m)$  เป็นเซตของสัมประสิทธิ์ของสัญญาณอินพุตจำนวน  $m$  ตัว และกำหนดให้ระดับความละเอียดของสัญญาณเริ่มต้นที่ศูนย์ จะสามารถเขียนเป็นความสัมพันธ์ได้ว่า

$$f(t) \in V_0; f(t) \in L^2(\mathbb{R}) \quad (2.15)$$

$$V_0 = W_{-1} \oplus \dots \oplus W_{j+2} \oplus W_{j+1} \oplus W_j \oplus V_j \quad (2.16)$$

สมการ (2.16) จะเป็นสมการแยกองค์ประกอบสัญญาณ  $f(t)$  ในสเปซ  $V^0$  ไปจนถึงระดับความละเอียดที่  $j$  ดังนั้นเทอมสัมประสิทธิ์  $c_0(m)$  จะถูกแตกเป็นสัมประสิทธิ์ในเซตของ  $c_j(m)$  ในสเปซของ  $V^j$  และกลุ่มของเซต  $d_{-1}(m), d_{-2}(m), \dots, d_j(m)$  ซึ่งอยู่ในเวฟเล็ทสเปซที่ระดับความละเอียดต่างๆ กัน ขบวนการดังกล่าวนี้จะเป็นลักษณะวิธีการ “การแปลงเวฟเล็ทแบบดิสครีท” (Discrete Wavelet Transform: DWT) นั่นเอง ดังแสดงในรูปที่ 2.12

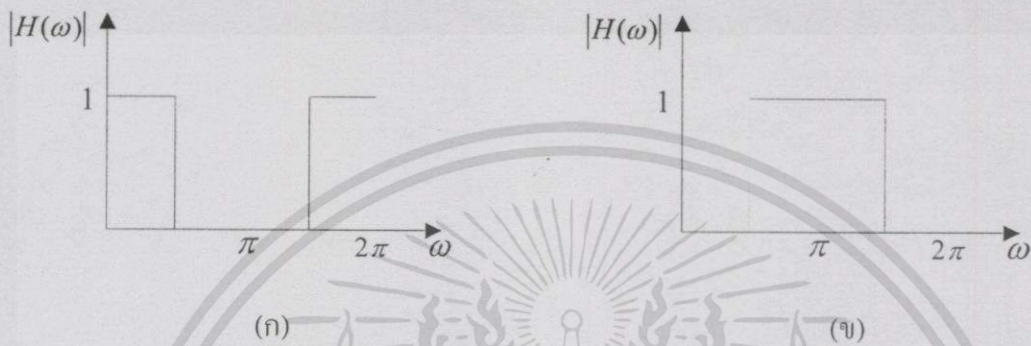


รูปที่ 2.12 การแจกแจงสัมประสิทธิ์ด้วยการแปลงเวฟเล็ท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1 การแปลงเวฟเล็ทดิสครีท 1 มิติ

การแยกข้อมูลไปเป็นลำดับชั้นของความละเอียดต่ำ (Coarse) และความละเอียดสูง (Detail) ต้องการตัวฟิลเตอร์ที่มีคุณสมบัติในการกรองความถี่ต่ำและความถี่สูงเชิงอุดมคติ ดังรูปที่ 2.13 กำหนดให้  $x(n)$  คือลำดับของข้อมูลอินพุท เงื่อนไขการออกแบบตัวฟิลเตอร์ของการกรองความถี่สูง  $H_0$  และการกรองความถี่ต่ำ  $H_1$  เป็นดังนี้



รูปที่ 2.13 ผลตอบสนองความถี่เชิงอุดมคติของ (ก) การกรองความถี่ต่ำ (ข) การกรองความถี่สูง

2.3.1.1 ฟิลเตอร์แบบค้แบบ 2 แบนด์

พิจารณาจากรูปที่ 2.14(ข) และกำหนดให้  $h_0(k)$  และ  $h_1(k)$  คือผลตอบสนองอิมพัลส์ที่จำกัด (Finite impulse response : FIR) เขียนความสัมพันธ์ได้ดังนี้ [11] [12]

$$y_0 = \sum_l x(l)h_0(n-l)|_{n=2k} \text{ และ } y_1 = \sum_l x(l)h_1(n-l)|_{n=2k} \tag{2.17}$$

เมื่อ  $y_1(k)$  คือเอาท์พุทขององค์ประกอบความถี่ต่ำดังรูปที่ 2.14(ข) ความสัมพันธ์ของ  $y_0(k)$  และ  $y_1(k)$  โดยที่  $k = 0,1,2,\dots$  ในรูปของเวกเตอร์คือ

$$\begin{bmatrix} \vdots \\ y_0(0) \\ y_1(0) \\ y_0(1) \\ y_1(1) \\ y_0(2) \\ y_1(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & h_0(2) & h_0(1) & h_0(0) & h_0(-1) & h_0(-2) & \dots \\ \dots & h_1(2) & h_1(1) & h_1(0) & h_1(-1) & h_1(-2) & \dots \\ \dots & \dots & \dots & h_0(2) & h_0(1) & h_0(0) & \dots \\ \dots & \dots & \dots & h_1(2) & h_1(1) & h_1(0) & \dots \\ \dots & \dots & \dots & \dots & \dots & h_0(2) & \dots \\ \dots & \dots & \dots & \dots & \dots & h_1(2) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ x(0) \\ x(1) \\ x(2) \\ \vdots \end{bmatrix} \tag{2.18}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือ 
$$y = T_i x \tag{2.19}$$

โครงสร้างของ  $T_i$  ประกอบด้วยผลตอบสนองอิมพัลส์ของ  $H_0$  และ  $H_1$  จากนั้นทำการลดข้อมูลลง 2 เท่า (Downsampling) เพื่อสร้างแถวข้อมูลใหม่ ทำนองเดียวกันพิจารณาจากรูปที่ 2.15(ข) แสดงโครงสร้างการรวมกันของ  $y_0(k)$  และ  $y_1(k)$  ที่ผ่านฟิลเตอร์การกรองความถี่สูง  $G_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $G_1$  เพื่อสร้างลำดับข้อมูลเอาท์พุท  $\hat{x}$  สัญลักษณ์ของการเพิ่มข้อมูล 2 เท่า (Upsampling) ด้วยตัวดำเนินการ  $M$  แสดงดังรูปที่ 2.15(ก) กรณีนี้ใช้  $M = 2$  และทำการเพิ่มศูนย์จำนวน  $M - 1$  ระหว่างค่าสัมประสิทธิ์เวฟเล็ท 2 ค่าจะได้ลำดับอินพุทของฟิลเตอร์การกรองความถี่สูง  $G_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $G_1$  มีลักษณะดังนี้

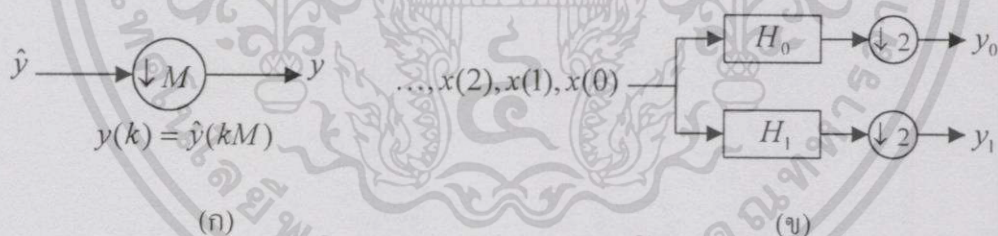
จะได้ว่า

$$\begin{matrix} \dots & 0 & y_0(0) & 0 & y_0(1) & 0 & y_0(2) & 0 & \dots \\ \dots & 0 & y_1(0) & 0 & y_1(1) & 0 & y_1(2) & 0 & \dots \end{matrix}$$

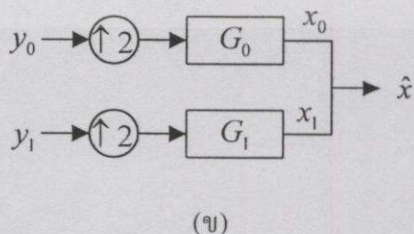
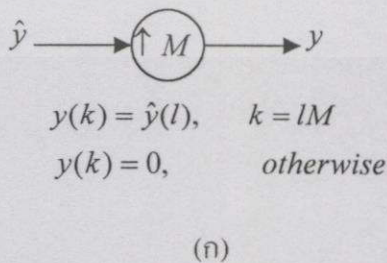
$$x_0(n) = \sum_k y_0(k) g_0(n - 2k) \tag{2.20}$$

$$x_1(n) = \sum_k y_1(k) g_1(n - 2k) \tag{2.21}$$

$$\hat{x}(n) = x_0(n) + x_1(n) \tag{2.22}$$



รูปที่ 2.14 (ก) ตัวดำเนินการลดข้อมูล (ข) การวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงค์



รูปที่ 2.15 (ก) ตัวดำเนินการเพิ่มข้อมูล (ข) การสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงค์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟิลเตอร์  $G$ , คือฟิลเตอร์ของการสังเคราะห์สัญญาณ (Synthesis filters) และสัมพันธ์กับฟิลเตอร์  $H_i$  หรือที่เรียกว่าฟิลเตอร์ของการวิเคราะห์สัญญาณ (Analysis filters) การรวมกันของ  $\hat{x}(n)$  ทุกๆ ค่าเป็นดังนี้

$$\begin{bmatrix} \vdots \\ \hat{x}(0) \\ \hat{x}(1) \\ \hat{x}(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & g_0(0) & g_1(0) & g_0(-2) & g_1(-2) & \cdots \\ \cdots & g_0(1) & g_1(1) & g_0(-1) & g_1(-1) & \cdots \\ \cdots & g_0(2) & g_1(2) & g_0(0) & g_1(0) & \cdots \\ \cdots & g_0(3) & g_1(3) & g_0(1) & g_1(1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ y_0(0) \\ y_1(0) \\ y_0(1) \\ y_1(1) \\ y_0(2) \\ y_1(2) \\ \vdots \end{bmatrix} \quad (2.23)$$

หรือ  $\hat{\mathbf{x}} = T_o \mathbf{y}$  (2.24)

ถ้า  $\hat{\mathbf{x}} = \mathbf{x}$  ดังนั้น

$$T_o T_i = I = T_i T_o \quad (2.25)$$

การคูณแถวของ  $T_i$  ด้วยคอลัมน์ของ  $T_o$  เท่ากับ

$$\sum_n h_i(2k-n)g_j(n-2l) = \delta_{kl} \delta_{ij}, \quad i, j = 0, 1 \quad (2.26)$$

สามารถเขียนเป็นผลคูณภายใน (Inner product) ได้ดังนี้

$$\langle h_i(2k-n), g_j(n-2l) \rangle = \delta_{kl} \delta_{ij} \quad (2.27)$$

หมายถึงผลคูณภายในของเวฟเล็ทจะเป็นหนึ่งได้เมื่อเป็นเวฟเล็ทที่ระดับของการแปลงและตำแหน่งเดียวกันเท่านั้น ถ้าฟิลเตอร์แบบ 2 แบนด์เป็นไปตามเงื่อนไขของสมการที่ (2.26) สามารถกล่าวได้ว่าฟิลเตอร์แบบ 2 แบนด์มีลักษณะของการสร้างข้อมูลกลับคืนมาอย่างสมบูรณ์ (Perfect reconstruction) และ  $\hat{x}(n) = x(n)$  ดังนั้นจะได้ว่า

$$x(n) = \sum_k y_0(k)g_0(n-2k) + \sum_k y_1(k)g_1(n-2k) \quad (2.28)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมการที่ (2.28) คือการกระจายของ  $x(n)$  ออกเป็นเซตลำดับมูลฐาน (Basis sequence)

$$\{g_0(n-2k), g_1(n-2k)\}, \quad k \in Z \quad (2.29)$$

เมื่อ  $Z$  คือเซตของจำนวนเต็ม ในส่วนของ  $y_0(k)$  และ  $y_1(k)$  คือค่าสัมประสิทธิ์เวฟเล็ตของการกระจาย ดังนั้นลักษณะของฟิลเตอร์แบงค์แบบ 2 แบนด์ที่สร้างกลับคืนมาอย่างสมบูรณ์ต้องสอดคล้องกับเงื่อนไขในสมการที่ (2.26) การแปลงสามารถกำหนดได้ดังนี้

$$y_i(n) = \sum_k x(n)h_i(2k-n) \quad (2.30)$$

$$x(n) = \sum_{i=0}^1 \sum_k y_i(k)g_i(n-2k) \quad (2.31)$$

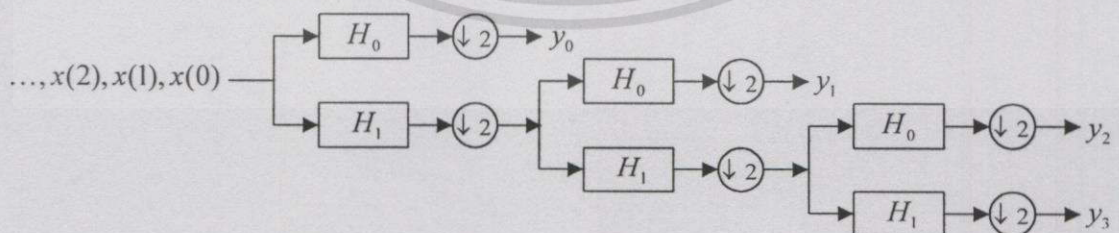
### 2.3.1.2 ฟิลเตอร์แบงค์แบบหลายแบนด์

จากรูปที่ 3.4 แสดงการวิเคราะห์และสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงค์ โดยใช้แนวคิดจากฟิลเตอร์แบงค์แบบ 2 แบนด์ กำหนดให้  $f_i(n)$  เป็นผลตอบสนองอิมพัลส์ของฟิลเตอร์  $F_i$  การแจกแจงลำดับของ  $y_i(k)$  ไปเป็นเอาต์พุต  $\hat{x}(n)$  คือ

$$x_i(n) = \sum_k y_i(k)f_i(n-2^{i+1}k) \quad i = 0, 1, 2, \dots, J-2 \quad (2.32)$$

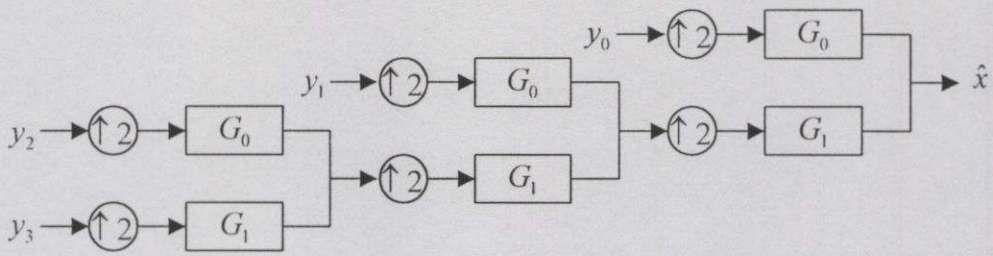
$$x_{J-1}(n) = \sum_k y_{J-1}(k)f_{J-1}(n-2^{J-1}k) \quad (2.33)$$

$$\hat{x}(n) = \sum_{i=0}^{J-1} x_i(n) \quad (2.34)$$



รูปที่ 2.16 โครงสร้างการวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงค์แบบหลายแบนด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 โครงสร้างการสังเคราะห์สัญญาณด้วยฟิลเตอร์เบงก์แบบหลายแบนด์

เมื่อ  $J$  คือจำนวนของแบนด์ จากรูปที่ 2.16 และ 2.17 จะได้  $J = 4$  ถ้าฟิลเตอร์  $G_0$  และ  $G_1$  ใช้ในการสังเคราะห์สัญญาณและฟิลเตอร์  $H_0$  และ  $H_1$  ใช้ในการวิเคราะห์สัญญาณที่สอดคล้องกับเงื่อนไขของสมการที่ (2.26) ดังนั้นการวิเคราะห์และสังเคราะห์สัญญาณของฟิลเตอร์เบงก์ที่ระดับ  $J$  จะมีลักษณะของการสร้างข้อมูลกลับคืนมาได้สมบูรณ์นั่นคือ

$$\hat{x}(n) = x(n) = \sum_{i=0}^{J-2} \sum_k y_i(k) f_i(n - 2^{i+1}k) + \sum_k y_{J-1}(k) f_{J-1}(n - 2^{J-1}k) \quad (2.35)$$

โดยที่

$$y_i(k) = \sum_n x(n) \hat{f}_i(2^{i+1}k - n), \quad i = 0, 1, \dots, J-2 \quad (2.36)$$

$$y_{J-1}(k) = \sum_n x(n) \hat{f}_{J-1}(2^{J-1}k - n) \quad (2.37)$$

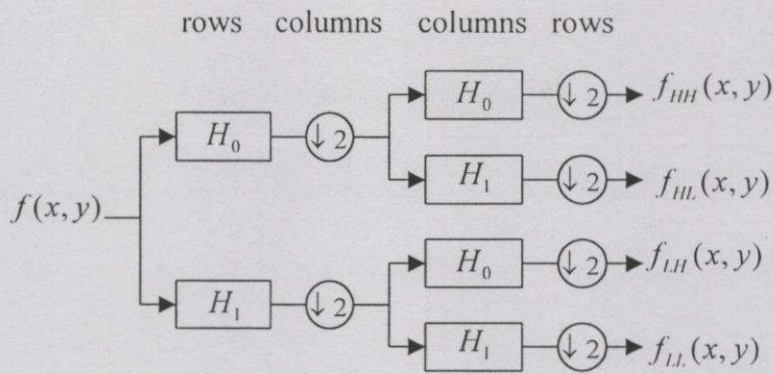
### 2.3.2 การแปลงเวฟเล็ตแบบดิสครีตใน 2 มิติ

ในที่นี้กล่าวถึงการแปลงข้อมูลภาพ 2 มิติ แบ่งออกเป็น 2 ชั้นตอนคือ

#### 2.3.2.1 การแยกองค์ประกอบเวฟเล็ต (Wavelet decomposition)

เป็นการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ดังรูปที่ 2.18 โดยเริ่มจากภาพต้นแบบ  $f(x, y)$  จะถูกกรองไปตามแนวแถว (Row) ด้วยฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  จากนั้นนำผลที่ได้มาลดการจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูล 1 คอลัมน์ (Column) จากข้อมูล 2 คอลัมน์ จะได้เอาที่ทุกความถี่สูงและความถี่ต่ำ จากนั้นเอาที่ทุกทั้งสองจะถูกกรองตามแนวคอลัมน์และลดจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูลมา 1 แถวจากข้อมูล 2 แถว จะได้ภาพแบนด์ย่อยความถี่ต่ำ  $f_{LL}(x, y)$  จำนวน 1 ภาพและภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพคือ  $f_{LH}(x, y)$   $f_{HL}(x, y)$  และ  $f_{HH}(x, y)$  โดยแต่ละภาพจะมีขนาดลดลงเป็น  $1/4$  เท่าของภาพต้นแบบ [13]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

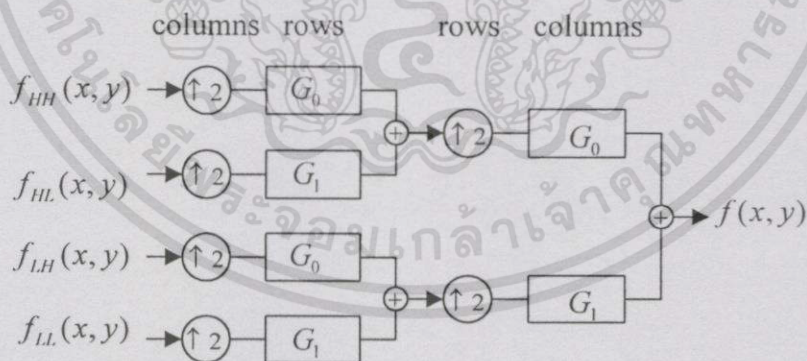


รูปที่ 2.18 การแยกองค์ประกอบของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์แบงก์

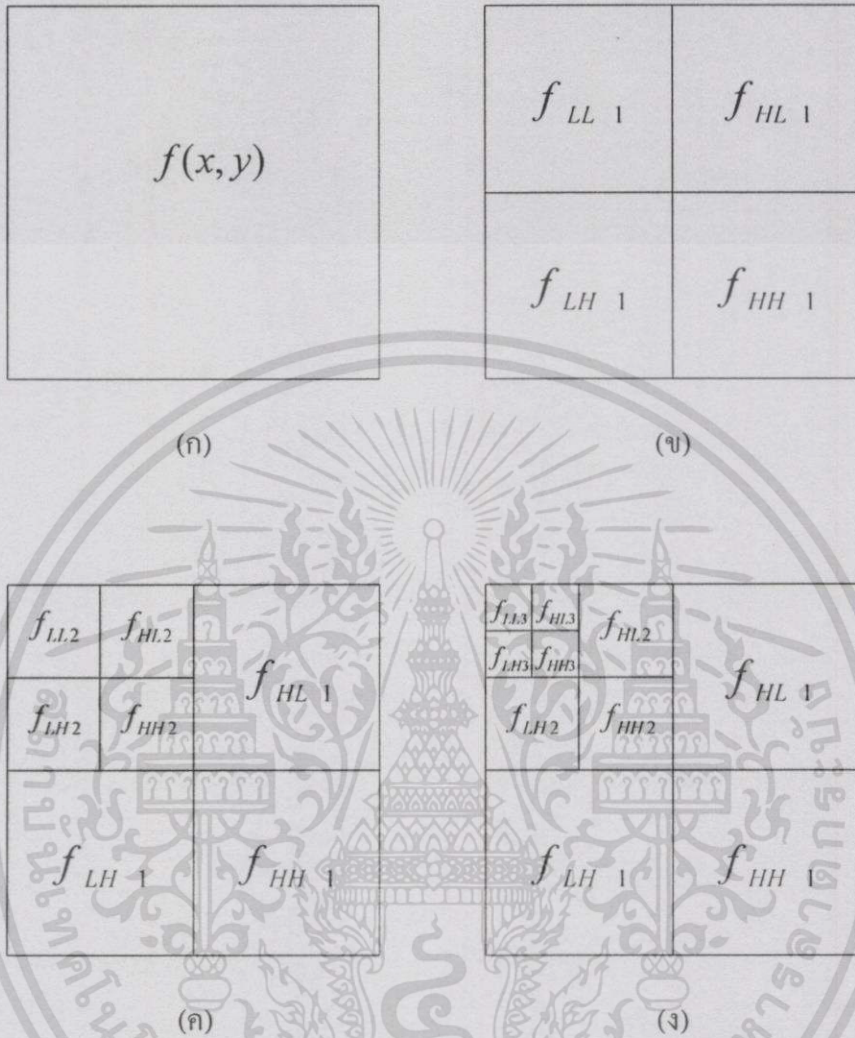
### 2.3.2.2 การรวมองค์ประกอบของเวฟเล็ท (Wavelet reconstruction)

เป็นการสร้างภาพขึ้นมาใหม่จากองค์ประกอบแบนด์ย่อย ดังรูปที่ 2.19 โดยเริ่มจากภาพแบนด์ย่อยทั้ง 4 ภาพจะถูกเพิ่มจำนวนข้อมูลเป็น 2 เท่า โดยการเติมศูนย์เข้าไปในระหว่างแต่ละคู่ของคอลัมน์ หลังจากเพิ่มจำนวนข้อมูลแล้วผลลัพธ์ที่ได้จะถูกกรองไปตามแถวและบวกเข้าด้วยกัน จากนั้นทำซ้ำไปตามแถวและคอลัมน์จะได้ผลลัพธ์ที่สร้างกลับขึ้นมาใหม่

การแปลงเวฟเล็ทของภาพ 1 ครั้ง (Stage) จะได้ภาพที่มีขนาดเท่ากันจำนวน 4 ภาพ และเป็นภาพส่วนที่เป็นข้อมูลรายละเอียดหรือข้อมูลความถี่สูงจำนวน 3 ภาพและเป็นภาพระดับความสว่างหรือข้อมูลความถี่ต่ำจำนวน 1 ภาพ ขั้นตอนการแปลงเวฟเล็ทสามารถกระทำกับข้อมูลภาพที่องค์ประกอบความถี่ต่ำไปเรื่อยๆ ดังรูปที่ 2.20 ส่วนรูปที่ 2.21 แสดงตัวอย่างการแยกองค์ประกอบของภาพ Lena โดยใช้การแปลงเวฟเล็ทดิสครีทของภาพ 2 ครั้งซึ่งใช้ฟิลเตอร์ Daubechies D4



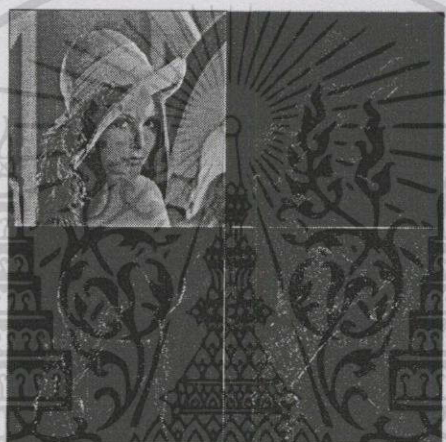
รูปที่ 2.19 การรวมองค์ประกอบของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์แบงก์



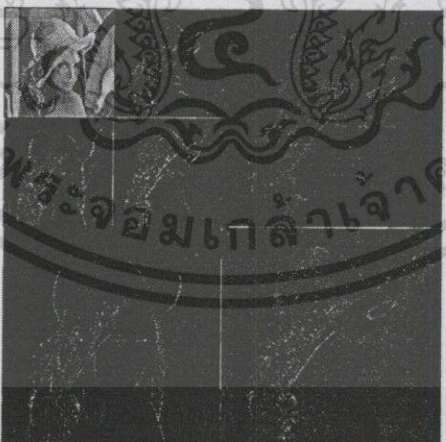
รูปที่ 2.20 ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต (ก) ข้อมูลภาพต้นแบบ (ข) การแปลงเวฟเล็ต 1 ครั้ง (ค) การแปลงเวฟเล็ต 2 ครั้ง (ง) การแปลงเวฟเล็ต 3 ครั้ง



(ก)



(ข)



(ค)

รูปที่ 2.21 ภาพการแปลงเวฟเล็ตดิสครีท (ก) ภาพต้นแบบ (ข) ภาพผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตดิสครีท 1 ครั้ง และ (ค) ภาพผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตดิสครีท 2 ครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.3.3 เวฟเล็ตแบบออร์โธโกนัลและแบบไบออร์โธโกนัล

เวฟเล็ตแบบออร์โธโกนัลและแบบไบออร์โธโกนัลพิจารณาได้จากเซตของฟังก์ชันมูลฐานมีความสัมพันธ์ดังนี้

$$h_i(2k-n) \equiv \phi_{ik}(n), \quad g_j(n-2l) \equiv \psi_{jl}(n), \quad i, j = 0, 1 \quad k, l \in Z$$

จากสมการที่ (2.26) มีคุณสมบัติเชิงตั้งฉาก (Orthogonal) ระหว่าง  $\phi_{ik}(n)$  และ  $\psi_{jl}(n)$  นั่นคือ

$$\langle \phi_{ik}(n), \psi_{jl}(n) \rangle = \delta_{ij} \delta_{kl}$$

เรียกว่าเวฟเล็ตแบบไบออร์โธโกนัล การแปลงเวฟเล็ตแบบดิสครีทในสมการที่ (2.31) คือการกระจายของไบออร์โธโกนัล (Biorthogonal expansion)

เมื่อลำดับของ  $\phi_{ik}(n) = h_i(2l-n)$  มีลักษณะเชิงตั้งฉากซึ่งกันและกัน นั่นคือ

$$\sum_n h_i(2k-n)h_j(2l-n) = \delta_{ij}\delta_{kl}, \quad i, j = 0, 1 \quad k, l \in Z$$

ดังนั้น  $g_i(n) = h_i(-n)$  หรือฟิลเตอร์ของการสังเคราะห์ก็คือฟิลเตอร์ของการวิเคราะห์สัญญาณที่เวลาย้อนกลับและมีเซตของลำดับ  $h_i$  อย่างเดียวกันนั้นในสมการที่ (2.31) เรียกว่าเวฟเล็ตแบบออร์โธโกนัล ตารางที่ 2.1 แสดงตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ตแบบออร์โธโกนัลชนิด Daubechies ( $h_1(n)$ ) และฟิลเตอร์การกรองความถี่สูงหาได้จาก  $h_0(n) = (-1)^n h_1(-n+2L-1)$  เมื่อ  $L$  คือความยาวของฟิลเตอร์หรือจำนวนของค่าสัมประสิทธิ์ฟิลเตอร์ สามารถสรุปความสัมพันธ์ของการแปลงเวฟเล็ตแบบดิสครีทได้ดังตารางที่ 2.2

ตารางที่ 2.1 ตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ตแบบออร์โธโกนัลชนิด Daubechies ที่มี ความยาวของฟิลเตอร์เท่ากับ 4, 6, 8 และ 10 ตามลำดับ

$n$	$h_1^4(n)$	$h_1^6(n)$	$h_1^8(n)$	$h_1^{10}(n)$
0	0.4829629	0.33267	0.2303778	0.1601024
1	0.8365163	0.806891	0.7148466	0.6038293
2	0.2241439	0.459877	0.6308808	0.7243085
3	-0.1294095	-0.135011	-0.0279838	0.1384281
4		-0.08544	-0.1870348	-0.2422949
5		0.03522	-0.0308414	-0.0322449
6			0.0328830	0.00775715
7			-0.0105974	-0.0062415
8				-0.0125807
9				0.0033357

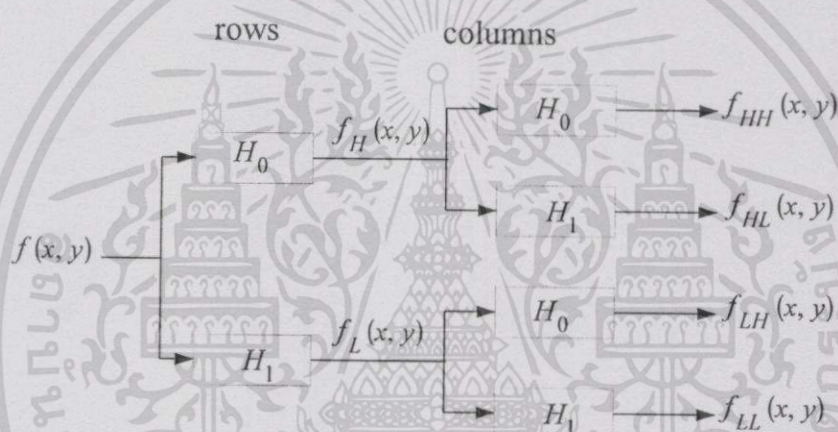
ตารางที่ 2.2 เงื่อนไขการแปลงเวฟเล็ตแบบดิสครีท

$y_i(k) = \sum_n x(n)\phi_{ik}(n)$	DWT
$x(n) = \sum_i \sum_k y_i(k)\psi_{ik}(n)$	Inverse DWT
$\sum_n \phi_{ik}(n)\psi_{jl}(n) = \delta_{kl}\delta_{ij}$	Biorthogonal expansion
$\psi_{ik}(n) = \phi_{ik}(n)$	Orthogonal expansion
$\sum_n \phi_{ik}(n)\phi_{jl}(n) = \delta_{kl}\delta_{ij}$	

## 2.4 การแปลงเวฟเล็ตแบบสแตชันนารี (Stationary Wavelet Transform)

โดยทั่วไปแล้ว การแปลงเวฟเล็ตแบบดิสครีทจะใช้ในจุดประสงค์ของการบีบอัดสัญญาณ สัมประสิทธิ์เวฟเล็ตได้ถูกแซมปลิงเริ่มต้นที่เกณฑ์ Nyquist การแทนความไม่ซ้ำซ้อน (non-redundant) สัมประสิทธิ์เวฟเล็ต และจำนวนผลรวมของแซมปลิงในการแทนเท่ากับจำนวนผลรวมของพิกเซลภาพ การแปลงเวฟเล็ตแบบดิสครีทนี้ไม่เป็นที่ต้องการสำหรับการวิเคราะห์ปัญหาหรือ การตรวจจับปัญหา ข้อเสียที่สำคัญของการแทนความไม่ซ้ำซ้อนสัมประสิทธิ์เวฟเล็ตนี้มันไม่สามารถที่จะรักษาคุณสมบัติที่สำคัญในการประมวลผลภาพ ซึ่งเป็นการเปลี่ยนแปลงโดยการแปลง คุณสมบัตินี้แน่นอนนั้นระดับที่แสดงในภาพจะถูกแทนด้วยระดับเวฟเล็ตที่เป็นอิสระต่อกันจาก ตำแหน่งในภาพของระดับเวฟเล็ต เพื่อที่จะรักษาการเปลี่ยนแปลงโดยการแปลงจะต้องไม่มีการทำ ความถี่แซมปลิง (Downsampling) และภาพที่ได้จากการแยกส่วนจะรักษาขนาดเท่ากับภาพเดิม และ

เรียกว่าการแปลงเวฟเล็ทแบบสแตชันนารี (stationary wavelet transform) [14] การแยกภาพต้นแบบ ออกเป็นองค์ประกอบย่อยได้แสดงดังแผนภาพรูปที่ 2.22 โดยเริ่มจากภาพต้นแบบ  $f(x, y)$  จะถูก กรองไปตามแนวแถว (Row) ด้วยฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  จะได้เอาต์พุตความถี่สูงและความถี่ต่ำ จากนั้นเอาต์พุตทั้งสองจะถูกกรองตามแนวคอลัมน์ด้วย ฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  อีกครั้งจะได้ภาพแบนด์ย่อย ความถี่ต่ำ  $f_{LL}(x, y)$  จำนวน 1 ภาพและภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพคือ  $f_{LH}(x, y)$   $f_{HL}(x, y)$  และ  $f_{HH}(x, y)$  ตามลำดับ โดยแต่ละภาพจะมีขนาดเท่ากับขนาดของภาพต้นแบบเนื่องจากว่าไม่ต้องมีการลดจำนวนข้อมูลลง 2 เท่า หรือการทำ Downsampling รูปที่ 2.23 แสดงตัวอย่าง การแยกองค์ประกอบของภาพ Tiffany โดยใช้การแปลงเวฟเล็ทสแตชันนารีซึ่งใช้ฟิลเตอร์ Daubechies D4



รูปที่ 2.22 การแยกองค์ประกอบเวฟเล็ทของการแปลงเวฟเล็ทสแตชันนารี



(ก)



(ข)

รูปที่ 2.23 ภาพการแปลงเวฟเล็ตสแตชันนารี (ก) ภาพต้นแบบ (ข) ภาพผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตสแตชันนารี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดลองของการแปลงเวฟเล็ทเบดิสกรีท และการแปลงเวฟเล็ทแบบสเตชันนารี จะเห็นว่าการแปลงทั้งสองแบบจะให้ผลลัพธ์ภาพเอาท์พุทที่เหมือนกันคือจะได้ภาพเอาท์พุทออกมาเป็น 4 แบนด์ภาพคือ  $f_{LL}(x, y)$  จำนวน 1 ภาพ และภาพย่อยความถี่สูงจำนวน 3 ภาพ คือ  $f_{LH}(x, y)$ ,  $f_{HL}(x, y)$  และ  $f_{HH}(x, y)$  จะต่างกันเฉพาะขนาดของภาพเอาท์พุทที่ได้จากการแปลงเท่านั้น ซึ่งภาพจากการแปลงเวฟเล็ทแบบดิสกรีทจะให้ขนาดของภาพเอาท์พุทลดลงเป็น  $1/4$  เท่าของภาพต้นแบบ ส่วนการแปลงเวฟเล็ทแบบสเตชันนารีจะให้ขนาดของภาพเอาท์พุทในแต่ละแบนด์ภาพเท่ากับขนาดของภาพต้นแบบ ซึ่งจากคุณสมบัติของการแปลงเวฟเล็ทแบบสเตชันนารีนี้สามารถนำไปตรวจจับค้นหารายละเอียดในข้อมูลภาพ

การแปลงเวฟเล็ทเป็นการแปลงข้อมูลจากโดเมนเวลา ไปอยู่ในระนาบเวลาและความถี่ ซึ่งเป็นข้อมูลในลักษณะสองมิติ ดังนั้นจึงสามารถทำการเลือกพิจารณาข้อมูลที่ตำแหน่งเวลาและความถี่ใดๆ ก็ได้ ข้อมูลจะถูกพิจารณามานิ่งกั้นหน้าค่างที่สามารถทำการหาคหรือขยายความถี่ได้ เรียกว่า “ฟังก์ชันเวฟเล็ท” โดยที่มีขอบเขตช่วงเวลากว้างหรือแคบเป็นสัดส่วนผกผันกับความถี่ของฟังก์ชันเวฟเล็ท นั่นคือ ฟังก์ชันเวฟเล็ทที่มีความถี่สูงจะมีช่วงขอบเวลาที่แคบ ในขณะที่ความถี่ต่ำจะมีช่วงขอบเขตเวลาที่กว้างขึ้น จากลักษณะสมบัตินี้จะทำให้ระนาบเวลาและความถี่ของการแปลงเวฟเล็ทมีการเปลี่ยนแปลงสเกลในแกนเวลาและความถี่ในลักษณะลอการิทึม ซึ่งสัมพันธ์กับลักษณะการเปลี่ยนแปลงของสัญญาณในทางปฏิบัตินั่นเอง

## บทที่ 3

# การค้นหาเส้นอ้างอิงในฟอร์มเอกสาร

การประมวลผลฟอร์มเอกสาร หรือการจัดการกับข้อมูลในฟอร์มเอกสารเป็นการทำงานที่มีความจำเป็นสำหรับทุกหน่วยงาน การประมวลผลโดยการรับรู้ระบบข้อมูลจากฟอร์มเอกสารจำนวนมากเช่นเช็คธนาคาร สลิปจ่ายเงิน และบิลใบเสร็จต่างๆ ต้องทำการประมวลผลทุกวัน ถ้าฟอร์มเอกสารเหล่านี้สามารถป้อนเข้าไปในคอมพิวเตอร์ และทำการประมวลผลแบบอัตโนมัติ จะทำให้ย่นระยะเวลาในการทำงานลงได้ ดังนั้นการประมวลผลโดยการรับรู้ข้อมูลแบบอัตโนมัติจึงได้เป็นเรื่องสำคัญ [15] ซึ่งหมายถึงการหาวิธีที่จะทำการคัดลอกข้อมูลที่กรอกในฟอร์มเอกสารเพื่อมาทำการจดจำตัวอักษร (OCR) แล้วจัดเก็บลงในระบบฐานข้อมูล ในวิทยานิพนธ์ฉบับนี้ได้ศึกษาในเรื่องการค้นหาเส้นอ้างอิง (Reference Line) ในฟอร์มเอกสาร โดยหลักในการพิจารณาลักษณะสมบัติที่สำคัญในฟอร์มเอกสาร โดยปกติฟอร์มเอกสารจะประกอบด้วยเส้นตรง ซึ่งส่วนใหญ่แล้วจะอยู่ในทิศทางด้านแนวนอน และแนวตั้งเส้นดังกล่าวนี้เรียกว่า “เส้นอ้างอิง” [2] เพื่อที่จะแสดงตำแหน่งในการกรอกข้อมูล ในทางปฏิบัตินั้นการกรอกหรือการเติมข้อมูลจะอยู่ในตำแหน่งด้านบนด้านล่าง และด้านข้างของเส้นอ้างอิง โดยที่คุณสมบัติของเส้นอ้างอิงนั้นจะพิมพ์เพื่อแนะนำผู้กรอกข้อมูลในตำแหน่งที่ถูกต้อง และข้อมูลบางส่วนในฟอร์มเอกสารที่นำมาใช้ประโยชน์ จะถูกป้อนเข้าคอมพิวเตอร์ และทำการประมวลผล ดังนั้นการค้นหาเส้นอ้างอิงจึงเป็นเรื่องสำคัญในการประมวลผลฟอร์มเอกสาร

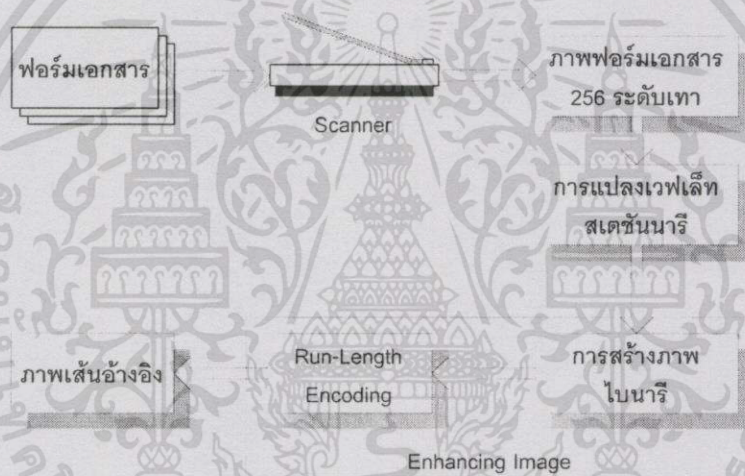
ดังนั้นในวิทยานิพนธ์ฉบับนี้ได้เสนอวิธีตามหลักการของเวฟเล็ตซึ่งใช้การแปลงเวฟเล็ตแบบสเตชันนารี (Stationary Wavelet Transform) โดยภาพฟอร์มเอกสารที่ผ่านการแปลงเวฟเล็ตจะสามารถแบ่งได้เป็นสี่ภาพย่อย และจากสองในสี่ภาพนี้สามารถค้นหาเส้นอ้างอิงจากฟอร์มเอกสารที่มีแบคกราวด์หลายระดับเทาได้

### 3.1 หลักการค้นหาเส้นอ้างอิง

หลักการค้นหาเส้นอ้างอิงในฟอร์มเอกสารสามารถแสดงได้ดังแผนภาพดังรูปที่ 3.1 ซึ่งจะต้องนำฟอร์มเอกสารที่มีรายละเอียดของเส้นอ้างอิง มาทำการสแกนด้วยเครื่องสแกนเนอร์เป็นภาพดิจิทัล 256 ระดับเทา จากนั้นนำภาพที่ได้จากการสแกนมาผ่านขบวนการแปลงเวฟเล็ต โดยได้ใช้การแปลงเวฟเล็ตสเตชันนารี ซึ่งจะให้ผลภาพเอาท์พุทออกมาเป็นสี่ภาพแบนด์ คือ ภาพแบนด์ LL ภาพแบนด์ LH ภาพแบนด์ HL และภาพแบนด์ HH แต่ภาพแบนด์ที่สามารถตรวจจับเส้นอ้างอิงในฟอร์มเอกสารได้นั้นคือภาพแบนด์ LH และภาพแบนด์ HL ซึ่งภาพแบนด์ LH เป็นผลจากการกรองซึ่งปล่อยให้องค์ประกอบความถี่ต่ำผ่านตามทิศทางแนวนอน และความถี่สูงผ่านตามทิศทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แนวตั้ง ทำให้เกิดเป็นภาพที่มีรายละเอียดของเส้นอ้างอิงเฉพาะทางด้านแนวนอน ส่วนภาพในแนวดิ่ง HL จะตรงกันข้ามกับผลของภาพย่อย LH ซึ่งปล่อยให้ความถี่สูงกว่าผ่านตามทิศทางด้านแนวนอน และความถี่ต่ำกว่าผ่านตามทิศทางด้านแนวตั้ง ทำให้เกิดเป็นภาพที่มีรายละเอียดของเส้นอ้างอิงเฉพาะทางด้านแนวตั้ง [2] แต่ข้อมูลภาพเส้นอ้างอิงที่ได้จากการแปลงเวฟเล็ทสเตชันนารีนั้นยังคงมีสิ่งที่ไม่พึงประสงค์ (noise) ปะปนมาด้วย จึงต้องนำหลักในการสร้างภาพไบนารี และอัลกอริทึมของ Run-Length Encoding มาประยุกต์ใช้ในการปรับปรุงภาพ เพื่อทำการจำแนกภาพในส่วนของเส้นอ้างอิงและแบกกราวด์ โดยใช้หลักการของเทคนิคการทำเรซโซลด์เข้ามาช่วยในการปรับปรุงภาพ ซึ่งจะทำได้สามารถจำแนกส่วนระหว่างเส้นอ้างอิงและแบกกราวด์ได้ แต่ภาพที่ได้ยังมีสิ่งที่ไม่พึงประสงค์บางส่วนปะปนมาด้วย เช่นตัวหนังสือบางส่วน เพราะว่าช่วงความถี่ของมันใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง จึงได้ใช้อัลกอริทึมของ Run-length Encoding มาประยุกต์ใช้เพื่อกำจัดสิ่งไม่พึงประสงค์เหล่านี้



รูปที่ 3.1 แผนภาพของการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร

### 3.2 การประยุกต์ใช้งานการแปลงเวฟเล็ทในการค้นหาเส้นอ้างอิง

เวฟเล็ทเป็นฟังก์ชันคณิตศาสตร์ที่แบ่งข้อมูลออกเป็นองค์ประกอบความถี่ที่ต่างกัน และจากนั้นจะศึกษาแต่ละองค์ประกอบซึ่งมีรีโซลูชันแมทซ์กับการสเกลของมัน ดังนั้นตระกูลเวฟเล็ทถูกกำหนดโดยการสเกลและการเลื่อนตำแหน่งของพารามิเตอร์  $a$  และ  $b$  ดังนี้

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (3.1)$$

และการแปลงเวฟเล็ทคือ inner product

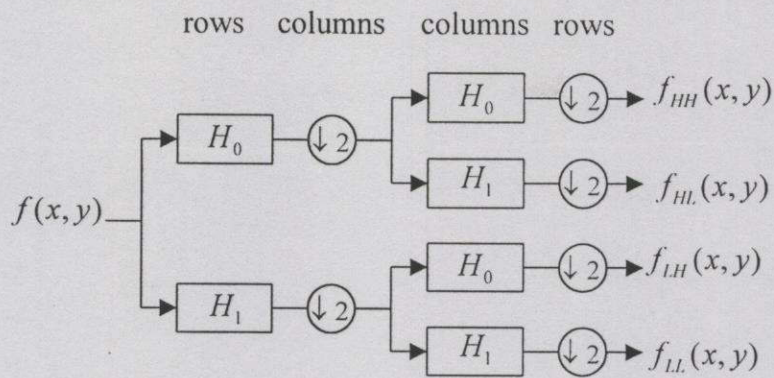
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$W(a,b) = \int_{-\infty}^{\infty} \psi_{ab}(t) f^*(t) dt = \langle \psi_{ab}, f \rangle \quad (3.2)$$

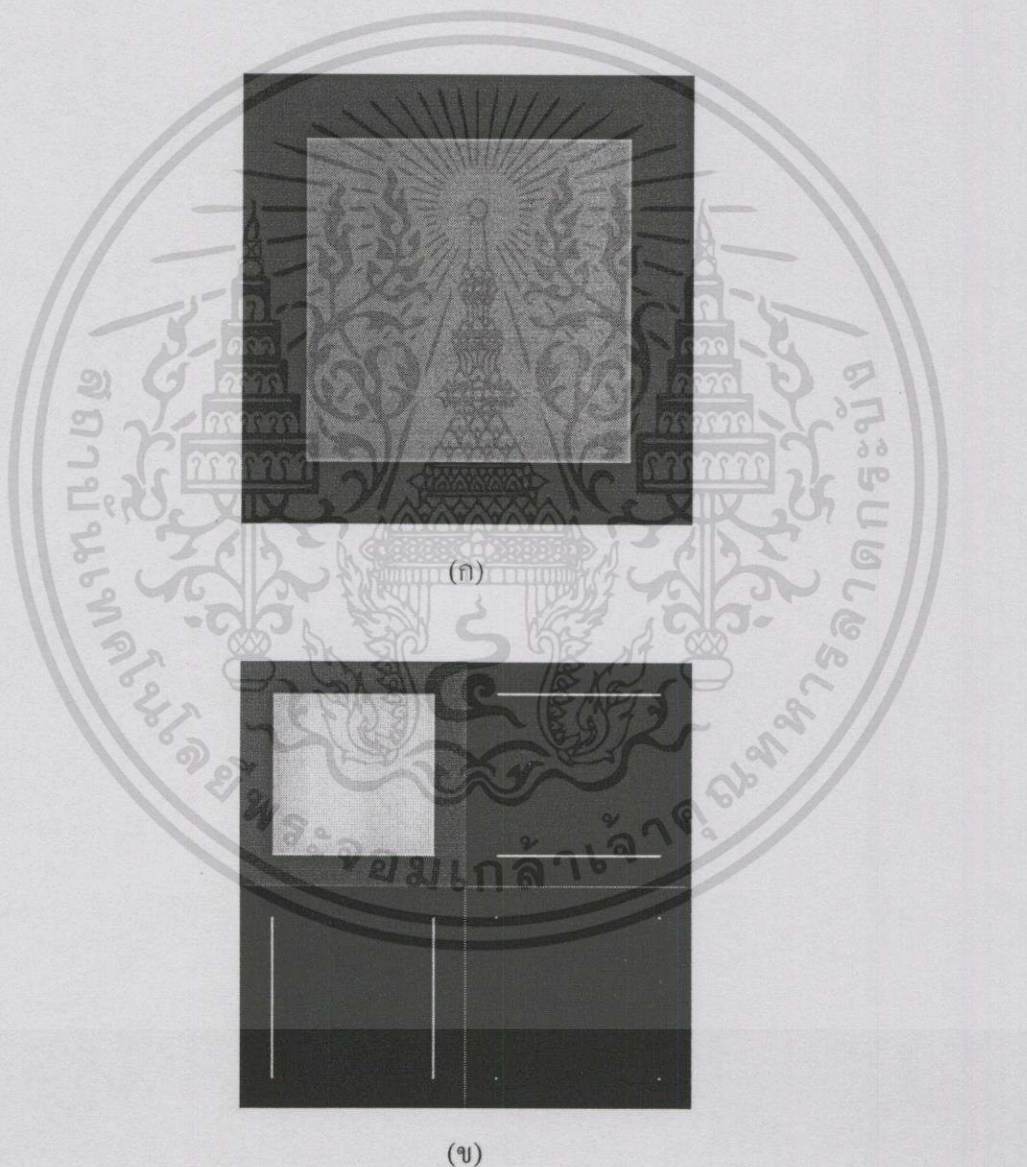
เมื่อ  $a$  มีค่ามาก basis function จะเป็นรูปแบบที่ถูกขยายออกของเวฟเล็ทแม่ นั่นคือฟังก์ชันความถี่ต่ำ ในขณะที่เดียวกันเมื่อ  $a$  มีค่าน้อย basis function จะอยู่ในรูปแบบของการหดเข้าของฟังก์ชันเวฟเล็ท หรือเวฟเล็ทแม่ ฟังก์ชันเวฟเล็ท  $\psi(t)$  ขึ้นอยู่กับพารามิเตอร์  $a$  ซึ่งก็คือการหดเข้าหรือขยายออกในแกนเวลา เป็นสาเหตุให้สอดคล้องกันกับการหดเข้าหรือขยายออกในโดเมนความถี่ [4]

### 3.2.1 การแปลงเวฟเล็ทดิสครีท (Discrete Wavelet Transform)

การแปลงเวฟเล็ทดิสครีทคือการวิเคราะห์สัญญาณในความถี่หลายแบนด์ หรือหลายสเกล ในทางปฏิบัติการวิเคราะห์มีลตรีโซลูชันกระทำโดยการใช้ filter bank ที่ประกอบด้วย ตัวกรองผ่านความถี่ต่ำ (low-pass filter)  $H_1$  และ ตัวกรองผ่านความถี่สูง (high-pass filter)  $H_0$  ในกรณีของภาพการฟิลเตอร์ถูกประยุกต์ใช้ในการแยกฟิลเตอร์ในแนวแถวและหลัก การทำเช่นนี้ให้สัมพันธ์กับการกระจายที่สัมพันธ์กับการแปลงเวฟเล็ทดิสครีทในการประมวลสัญญาณดิจิทัล ในกรณีของภาพดิจิทัลซึ่งเป็นข้อมูลขนาดสองมิติ การกรองความถี่จะทำให้ได้โดยการกรองความถี่ในส่วนที่เป็นแถว และกรองความถี่ในส่วนที่เป็นหลัก โดยที่ทำการการแปลงเวฟเล็ทดิสครีทจะแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ซึ่งแสดงได้ดังแผนภาพรูปที่ 3.2 โดยเริ่มจากภาพต้นแบบ  $f(x,y)$  จะถูกกรองไปตามแนวแถว (Row) ด้วยฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  จากนั้นนำผลที่ได้มาทำการลดจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูล 1 คอลัมน์ (Column) จากข้อมูล 2 คอลัมน์ จะได้เอาที่พหุความถี่สูงและความถี่ต่ำ จากนั้นเอาที่พหุทั้งสองจะถูกกรองตามแนวคอลัมน์และลดจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูลมา 1 แถวจากข้อมูล 2 แถว จะได้ภาพแบนด์ย่อยความถี่ต่ำ  $f_{LL}(x,y)$  จำนวน 1 ภาพและภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพคือ  $f_{LH}(x,y)$   $f_{HL}(x,y)$  และ  $f_{HH}(x,y)$  โดยแต่ละภาพจะมีขนาดลดลงเป็น 1/4 เท่าของภาพต้นแบบตามลำดับ ตัวอย่างของภาพที่ได้จากการการแปลงเวฟเล็ทดิสครีทโดยใช้ฟิลเตอร์ Daubechies D4 แสดงดังรูปที่ 3.3



รูปที่ 3.2 แผนภาพการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อยของการแปลงเวฟเล็ตดิสครีท

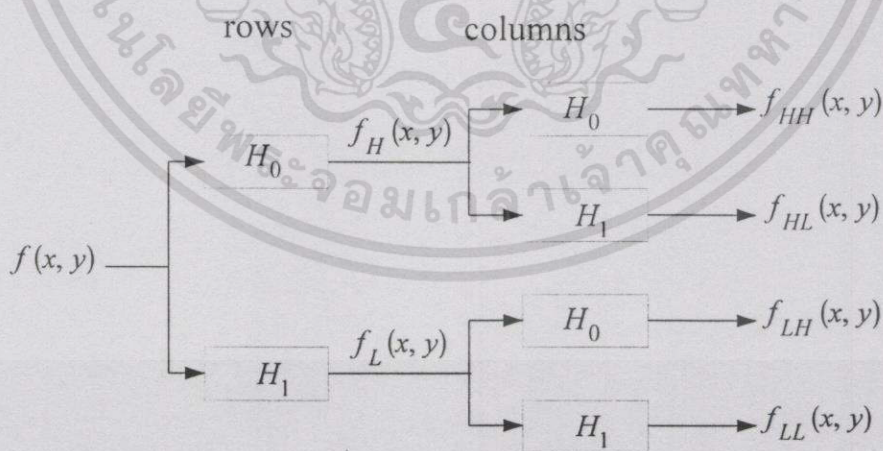


รูปที่ 3.3 ตัวอย่างภาพจากการแปลงเวฟเล็ตดิสครีท (ก) ภาพต้นแบบ (ข) ภาพที่ได้จากการแปลงเวฟเล็ตดิสครีท

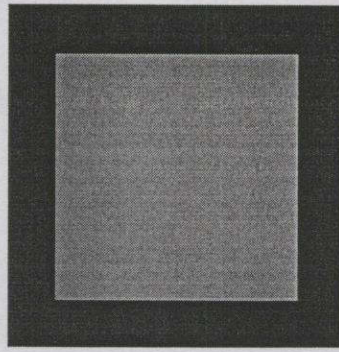
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.2.2 การแปลงเวฟเล็ทสเตชันนารี (Stationary Wavelet Transform)

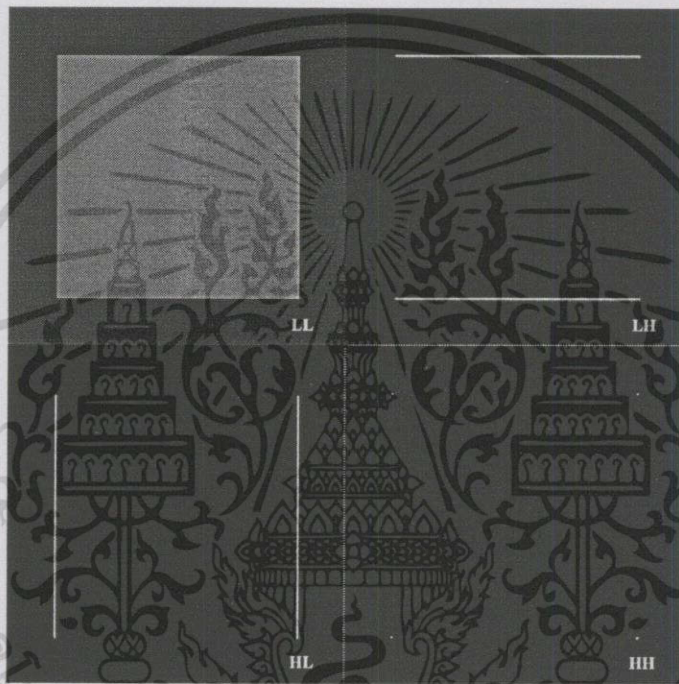
ดังที่ได้เห็นในตัวอย่างของภาพที่ได้จากการแปลงเวฟเล็ทดิสครีทจะมีขนาดลดลงเป็น  $1/4$  เท่าของภาพต้นแบบ ซึ่งโดยส่วนใหญ่แล้วการแปลงเวฟเล็ทดิสครีทจะใช้ในจุดประสงค์ของการบีบอัดสัญญาณข้อมูล สัมประสิทธิ์เวฟเล็ทได้ถูกแซมปลิงที่เกณฑ์ Nyquist การแทนความไม่ซ้ำซ้อน (non-redundant) สัมประสิทธิ์เวฟเล็ท และจำนวนผลรวมของแซมปลิงในการแทนเท่ากับจำนวนผลรวมของพิกเซลภาพ ข้อเสียที่สำคัญของการแทนความไม่ซ้ำซ้อนสัมประสิทธิ์เวฟเล็ทนี้มันไม่สามารถที่จะรักษาคุณสมบัติที่สำคัญในการประมวลผลภาพ ซึ่งคือการเปลี่ยนแปลงโดยการแปลงเวฟเล็ทดิสครีท ดังนั้นการแปลงเวฟเล็ทดิสครีทนี้จึงไม่เป็นที่ต้องการในการตรวจจับข้อมูลในภาพ เพื่อที่รักษาการเปลี่ยนแปลงโดยการแปลงจะต้องไม่มีการทำ downsampling และภาพผลลัพธ์ที่ได้จากการแยกองค์ประกอบย่อยจะรักษาขนาดเท่ากับภาพต้นแบบ ซึ่งเรียกว่าการแปลงเวฟเล็ทสเตชันนารี (Stationary Wavelet Transform) [14] การแยกภาพต้นแบบออกเป็นภาพองค์ประกอบย่อยได้แสดงดังแผนภาพรูปที่ 3.4 โดยเริ่มจากภาพต้นแบบ  $f(x, y)$  จะถูกกรองไปตามแนวแถว (Row) ด้วยฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  จะได้เอาที่ทุกความถี่สูงและความถี่ต่ำ จากนั้นเอาที่ทุกทั้งสองจะถูกกรองตามแนวคอลัมน์ด้วยฟิลเตอร์การกรองความถี่สูง  $H_0$  และฟิลเตอร์การกรองความถี่ต่ำ  $H_1$  อีกครั้งจะได้ภาพแบนด์ย่อยความถี่ต่ำ  $f_{LL}(x, y)$  จำนวน 1 ภาพและภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพคือ  $f_{LH}(x, y)$   $f_{HL}(x, y)$  และ  $f_{HH}(x, y)$  ตามลำดับ โดยแต่ละภาพจะมีขนาดเท่ากับขนาดของภาพต้นแบบเนื่องจากว่าไม่ต้องการมีการลดจำนวนข้อมูลลง 2 เท่า หรือการทำ Downsampling ตัวอย่างของภาพที่ได้จากการแปลงเวฟเล็ทสเตชันนารีโดยใช้ฟิลเตอร์ Daubechies D4 แสดงดังรูปที่ 3.5



รูปที่ 3.4 แผนภาพการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อยของการแปลงเวฟเล็ทสเตชันนารี



(ก)



(ข)

รูปที่ 3.5 ตัวอย่างภาพจากการแปลงเวฟเล็ตสแตชันนารี (ก) ภาพต้นแบบ (ข) ภาพที่ได้จากการแปลงเวฟเล็ตสแตชันนารี

จากรูปที่ 3.5 เป็นการนำการแปลงเวฟเล็ตสแตชันนารีมาประยุกต์ใช้กับภาพที่ได้ทำการสังเคราะห์สร้างขึ้นซึ่งเป็นภาพสี่เหลี่ยมระดับเทาที่ได้แสดงในรูปที่ 3.5 (ก) จะเห็นได้ว่าการแปลงเวฟเล็ตสแตชันนารีซึ่งสามารถแยกภาพต้นแบบออกมาเป็น 4 ภาพแบนด์ดังแสดงในรูปที่ 3.5 (ข) ภาพแบนด์ LL เป็นผลจากการกรองถี่โดยยอมให้ความถี่ต่ำกว่าผ่านได้ในแนวแถว (row) และแนวหลัก (column) ซึ่งจะทำให้ภาพแบนด์ LL มีรายละเอียดใกล้เคียงกับภาพต้นแบบ ภาพแบนด์ HH เป็นผลจากการกรองโดยยอมให้ความถี่สูงกว่าผ่านได้ในแนวแถวและแนวหลัก ซึ่งจะทำให้ภาพแบนด์ HH มีรายละเอียดในแนวทแยงมุม ภาพแบนด์ LH เป็นผลจากการกรองโดยยอมให้ความถี่ต่ำ

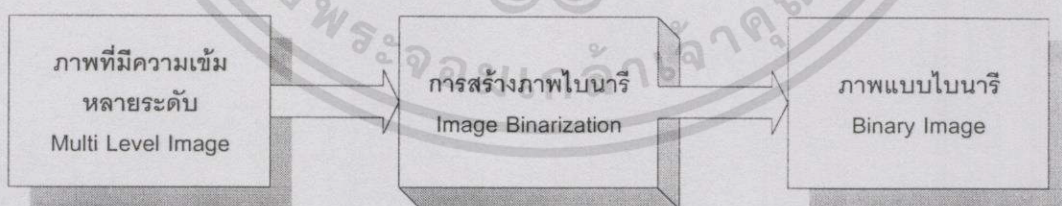
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กว่าผ่านได้ในแนวแถว และความถี่สูงกว่าผ่านได้ในแนวหลัก ซึ่งส่งผลให้ภาพแบนด์ LH มีรายละเอียดเส้นตรงทางด้านแนวนอน เช่นเดียวกันภาพแบนด์ HL จะตรงกันข้ามกับแบนด์ LH นั่นคือภาพที่ได้จากการกรองความถี่จะยอมให้ความถี่สูงกว่าผ่านได้ในแนวแถว และความถี่ต่ำกว่าผ่านได้ในแนวหลัก ซึ่งส่งผลให้ภาพแบนด์ HL มีรายละเอียดเส้นตรงในด้านแนวตั้ง

ในวิทยานิพนธ์นี้ได้สนใจภาพแบนด์ LH และ HL เพราะคุณสมบัติของภาพเหล่านี้สามารถตรวจจับแนวเส้นตรงในภาพฟอร์มเอกสารได้ จากนั้นเราทำการจำแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ และกำจัดข้อมูลสิ่งที่ไม่พึงประสงค์ปะปนมา(noise) โดยใช้เทคนิคการทำเรซโซลต์ และอัลกอริทึมของการเข้ารหัสรันเล็งซ์มาประยุกต์ใช้เพื่อที่จะทำการปรับปรุงภาพเพื่อให้ได้ภาพที่มีรายละเอียดของเส้นอ้างอิงเด่นชัดยิ่งขึ้น

### 3.3 เทคนิคการทำเรซโซลต์ (Thresholding Technique)

เทคนิคการทำเรซโซลต์เป็นเทคนิคที่ใช้ในการสร้างภาพไบนารี ขั้นตอนการสร้างภาพไบนารีจากภาพที่มีระดับความเข้มหลายระดับ สามารถแสดงได้ดังรูปที่ 3.6 จากภาพสามารถอธิบายหลักการทำงานของการสร้างภาพไบนารีได้ดังนี้ ภาพเริ่มต้นเป็นภาพที่มีแต่ละจุดภาพมีระดับความเข้มหลายระดับ ตั้งแต่ 0 ถึง  $L$  ระดับ จากนั้นภาพเริ่มต้นนี้จะถูกนำเข้าไปสู่กระบวนการที่เรียกว่า “อิมเมจไบนารีเซชัน” (Image Binarization) ซึ่งเป็นกระบวนการในการแปลงข้อมูลภาพที่มีระดับความเข้มหลายระดับเป็นไบนารี ภาพผลลัพธ์ที่ได้จากกระบวนการนี้จะได้เป็นภาพไบนารีที่แต่ละจุดภาพจะมีเพียง 2 ระดับคือ 0 (จุดดำ) กับ 1 (จุดขาว) ซึ่งการพิจารณาว่าจุดภาพใดควรจะเป็นจุดขาวหรือจุดดำ จะกระทำโดยการเปรียบเทียบระหว่างจุดภาพเริ่มต้นกับค่าหนึ่งที่เรียกว่า “ค่าเรซโซลต์” (Threshold Value) [16] เทคนิคนี้ใช้กันมากในกรณีที่มีข้อมูลภาพมีลักษณะแตกต่างกันมาก ระหว่างวัตถุ (object) และแบกกราวด์ เช่นภาพตัวอักษรที่ปรากฏบนพื้นกระดาษสีขาว



รูปที่ 3.6 แผนภาพหลักการสร้างภาพไบนารี

สำหรับหลักการทำงานพื้นฐานของเทคนิคนี้ จะทำการแยกข้อมูลภาพออกเป็นสองส่วน โดยอาศัยค่าหนึ่งที่เรียกว่า “ค่าเรซโซลต์” (Threshold Value) โดยค่าของจุดภาพใดๆ ที่มีค่าน้อยกว่าค่าเรซโซลต์จะถูกกำหนดให้เป็น 0 (จุดดำ) และถ้าค่าของจุดภาพใดๆ ที่มีค่ามากกว่า

หรือเท่ากับค่าเรขาคณิตจะถูกเปลี่ยนให้เป็น 1 (จุดขาว) ซึ่งการทำงานสามารถแสดงได้ดังสมการที่ (3.3)

$$b(x, y) = \begin{cases} 1; & g(x, y) > Thr \\ 0; & g(x, y) \leq Thr \end{cases} \quad (3.3)$$

โดยที่

$b(x, y)$	ข้อมูลภาพผลลัพธ์ เป็นภาพไบนารี
$g(x, y)$	ข้อมูลภาพอินพุท ที่มีระดับความเข้ม 0 ถึงระดับ $L$
$Thr$	ค่าเรขาคณิต เป็นค่าคงที่ ที่อยู่ระหว่าง 0 ถึงระดับ $L$
0	จุดดำ
1	จุดขาว

ในการสร้างภาพไบนารีโดยใช้เทคนิคเรขาคณิตเพื่อให้ได้ผลลัพธ์ที่เหมาะสมและคมชัด นั้น สิ่งสำคัญที่สุดคือ ค่าเรขาคณิต เนื่องจากถ้าเลือกค่าเรขาคณิตที่ไม่เหมาะสม (ค่าเรขาคณิตมีค่าน้อยหรือมากเกินไป) ภาพที่ได้อาจจะไม่เหมาะสม ขาดความคมชัด และรายละเอียดบางส่วนขาดหายไป กล่าวคือภาพที่ได้อาจจะมืดจนเกินไป (จุดดำมากเกินไป) หรือสว่างมากเกินไป (จุดขาวมากเกินไป) หรือภาพที่ได้มีสิ่งที่ไม่พึงประสงค์ปนมาด้วย (noise) อันเป็นผลทำให้ภาพผลลัพธ์ที่ได้ไม่สวยงามเท่าที่ควร ดังนั้นปัญหาของการสร้างภาพไบนารีโดยวิธีเรขาคณิตก็คือ ทำอย่างไรจึงจะหาค่าเรขาคณิตที่เหมาะสมสำหรับแต่ละภาพที่จะนำมาทำการสร้างภาพไบนารี ได้มีผู้เสนอวิธีการหาค่าเรขาคณิตหลายวิธี โดยแต่ละวิธีเหมาะสมกับลักษณะการทำงานที่ต่างกันไป เช่นการหาค่าเรขาคณิตโดยการกำหนดค่าล่วงหน้า (Preassigned Threshold Value) การหาค่าเรขาคณิตจากฮิสโตแกรม (Histogram Based Threshold Value) และการหาค่าเรขาคณิตจากค่า RMS (Root Mean Square Threshold Value)

### 3.3.1 การหาค่าเรขาคณิตโดยการกำหนดค่าล่วงหน้า

การหาค่าเรขาคณิตวิธีการกำหนดค่าล่วงหน้า [17] นี้เป็นวิธีที่ง่ายที่สุด เป็นวิธีการคำนวณค่าเรขาคณิตโดยการกำหนดเองจากผู้ใช้ (User) ซึ่งการกำหนดนี้จะขึ้นอยู่กับประสบการณ์ของผู้ใช้นั้นๆ โดยการเลือกค่าคงที่ค่าหนึ่ง ซึ่งเรียกค่านั้นว่าค่าเรขาคณิต โดยค่าที่เลือกมานี้จะเป็นค่าที่อยู่ระหว่างค่าต่ำสุดและค่าสูงสุดของระดับความเข้มของข้อมูลภาพอินพุท เช่นข้อมูลภาพอินพุทที่มีระดับ

เทา 256 ระดับ จะมีค่าระดับเทาได้ตั้งแต่ 0-255 ค่าเรขาคณิตที่สามารถเลือกได้ก็คืออยู่ที่ระหว่าง 0-255 เมื่อเลือกค่าเรขาคณิตได้แล้วสามารถสร้างภาพไบนารีได้โดยใช้สมการที่ (3.3)

ในการทดลองสร้างภาพไบนารีโดยวิธีการหาค่าเรขาคณิตโดยการกำหนดนี้ ได้เลือกใช้ภาพของตัวอักษรบนพื้นกระดาษหนังสือพิมพ์ โดยได้ทำการสแกนเป็นภาพที่มีระดับเทา 256 ระดับ จากนั้นทำการเลือกค่าเรขาคณิตขึ้นมา 3 ค่าคือ 80 150 และ 200 สำหรับผลลัพธ์ของการสร้างภาพไบนารีโดยวิธีดังกล่าวนี้แสดงดังรูปที่ 3.7 รูปที่ 3.7 (ก) เป็นภาพต้นฉบับ 256 ระดับเทา จากนั้นนำไปผ่านเทคนิคการทำเรขาคณิตโดยกำหนดให้ค่าเรขาคณิตเท่ากับ 80 150 และ 200 ซึ่งจะให้ภาพผลลัพธ์แสดงดังรูปที่ 3.7 (ข) (ค) และ (ง) ตามลำดับ จากผลลัพธ์ที่ได้จะเห็นว่าถ้าเรากำหนดค่าเรขาคณิตเอนเอียงไปทางระดับความเข้มที่ต่ำหรือสูงเกินไป ภาพผลลัพธ์ที่ได้ก็จะไม่เหมาะสม เช่นอาจขาวเกินไป (รายละเอียดไม่ครบ) หรือดำเกินไป (เกิดสิ่งไม่พึงประสงค์ปะปนมาด้วย (noise))

ดร.อุษรยาภิปฏิเสธ พ.ด.ท.หัวหน้า  
ศูนย์ฯ ถ้าเป็นคน  
หนึ่งตำรวจด้วย โ  
ความเป็นธรรม

(ก)

ดร.อุษรยาภิปฏิเสธ  
ไม่ได้ซ่อมคดีปล้นรถ

(ข)

ดร.อุษรยาภิปฏิเสธ พ.ด.ท.หัวหน้า  
ศูนย์ฯ ถ้าเป็นคน  
หนึ่งตำรวจด้วย โ  
ความเป็นธรรม

(ค)

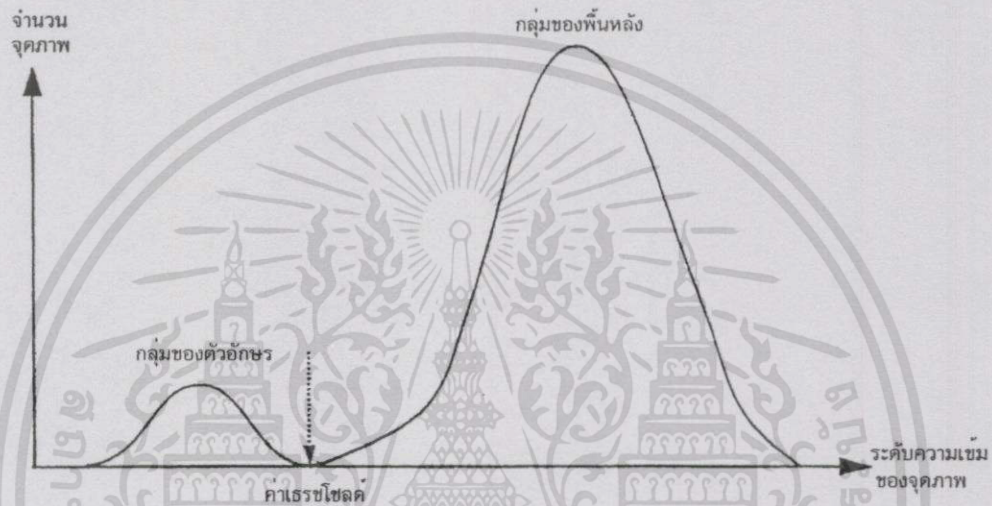
ดร.อุษรยาภิปฏิเสธ พ.ด.ท.หัวหน้า  
ศูนย์ฯ ถ้าเป็นคน  
หนึ่งตำรวจด้วย โ  
ความเป็นธรรม

(ง)

รูปที่ 3.7 ผลลัพธ์ของการสร้างภาพไบนารีโดยวิธีการกำหนดค่าเรขาคณิต (ก) ภาพอินพุต 256 ระดับเทา (ข) ค่าเรขาคณิตเท่ากับ 80 (ค) ค่าเรขาคณิตเท่ากับ 150 (ง) ค่าเรขาคณิตเท่ากับ 200

### 3.3.2 การหาค่าเรขาคณิตโดยพิจารณาจากฮิสโตแกรม

การหาค่าเรขาคณิตโดยพิจารณาจากฮิสโตแกรมนี้ [18] เป็นการพิจารณาจากฮิสโตแกรมระดับเทาของข้อมูลภาพที่เป็นอินพุท โดยที่การหาค่าเรขาคณิตวิธีนี้ ข้อมูลภาพอินพุทที่เหมาะสมต้องมีลักษณะที่สามารถแบ่งแยกข้อมูลออกเป็นสองกลุ่มได้อย่างชัดเจน คือกลุ่มหนึ่งจะเป็นกลุ่มของวัตถุ (Object) และอีกกลุ่มหนึ่งจะเป็นแบกราวด์ ซึ่งแนวคิดในการหาค่าเรขาคณิตของวิธีนี้สามารถแสดงได้ดังรูปที่ 3.8



รูปที่ 3.8 แนวความคิดการหาค่าเรขาคณิต โดยวิธีพิจารณาจากฮิสโตแกรม

สำหรับการหาค่าเรขาคณิตโดยวิธีพิจารณาจากฮิสโตแกรม สามารถสรุปเป็นขั้นตอนได้ดัง

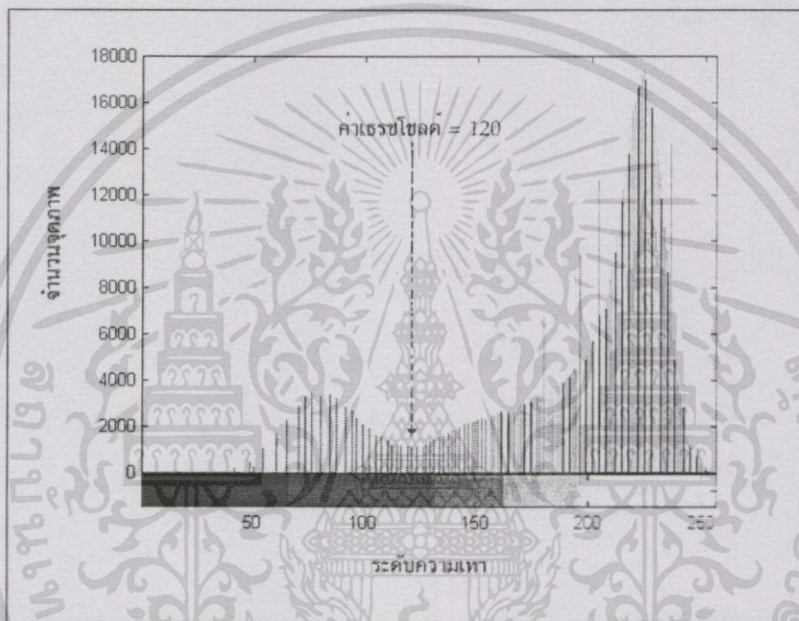
เช่น

1. ข้อมูลภาพอินพุท ที่มีระดับความเข้ม 0 ถึง  $L$  ระดับ
2. หาฮิสโตแกรมระดับความเข้มของภาพหน้าเอกสาร โดยกานับจำนวนจุดภาพที่ระดับความเข้มแต่ละระดับ
3. จากฮิสโตแกรมจะพบว่า จะเกิดกลุ่มของระดับความเข้ม 2 กลุ่ม กลุ่มหนึ่งคือกลุ่มของวัตถุ และอีกกลุ่มหนึ่งคือกลุ่มของแบกราวด์
4. จะเลือกค่าต่ำที่สุดที่อยู่ระหว่างสองกลุ่มนั้น ซึ่งค่านั้นก็คือนั่นเอง
5. สร้างภาพไบนารี โดยนำค่าเรขาคณิตที่ได้มาแทนในสมการที่ (3.3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	ออมทรัพย์	เงินฝากประจำ			
		3 เดือน	6 เดือน	12 เดือน	24 เดือน
กรุงเทพ	2.00	2.50	2.50	3.00	3.50
กรุงไทย	2.00	2.50	2.50	3.00	3.50
กสิกรไทย	2.00	2.50	2.50	3.00	3.50
ไทยพาณิชย์	2.00	2.50	2.50	3.00	3.50
กรุงศรีอยุธยา	2.00	2.50	2.50	3.00	3.50
ทหารไทย	2.00	2.50	2.50	3.00	3.50
ดีเอสไอไทยกยู	1.75	2.25	2.25	3.00	3.50
สแตนดาร์ดชาร์เตอร์ดนคร	1.75	2.50	2.50	2.75	3.00
เอเชีย	1.75	2.25	2.25	2.50	2.75

(ก)



(ข)

	ออมทรัพย์	เงินฝากประจำ			
		3 เดือน	6 เดือน	12 เดือน	24 เดือน
กรุงเทพ	2.00	2.50	2.50	3.00	3.50
กรุงไทย	2.00	2.50	2.50	3.00	3.50
กสิกรไทย	2.00	2.50	2.50	3.00	3.50
ไทยพาณิชย์	2.00	2.50	2.50	3.00	3.50
กรุงศรีอยุธยา	2.00	2.50	2.50	3.00	3.50
ทหารไทย	2.00	2.50	2.50	3.00	3.50
ดีเอสไอไทยกยู	1.75	2.25	2.25	3.00	3.50
สแตนดาร์ดชาร์เตอร์ดนคร	1.75	2.50	2.50	2.75	3.00
เอเชีย	1.75	2.25	2.25	2.50	2.75

(ค)

รูปที่ 3.9 ขั้นตอนการสร้างภาพไบนารีโดยวิธีพิจารณาจากฮิสโตแกรม (ก) ภาพอินพุต 256 ระดับเทา (ข) ภาพฮิสโตแกรมของภาพอินพุต 256 ระดับเทา (ค) ภาพผลลัพธ์ไบนารี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากขั้นตอนการทำงานเบื้องต้น เราสามารถทดลองสร้างภาพไบนารีได้ โดยได้ทดลองกับ ภาพอินพุต 256 ระดับเทาที่เป็นภาพตัวอักษรบนพื้นกระดาษที่สีอ่อนกว่า ดังแสดงในรูปที่ 3.9 (ก) ซึ่งเมื่อทำการสร้างฮิสโตแกรมระดับเทาของภาพแล้วปรากฏว่าเกิดกลุ่มของข้อมูล 2 กลุ่ม คือกลุ่มหนึ่งจะเป็นกลุ่มของข้อความ และอีกกลุ่มหนึ่งจะเป็นกลุ่มของแบกกราวด์ จากนั้นทำการเลือกค่าต่ำสุด (Valley) ระหว่าง 2 กลุ่มนั้นเป็นค่าธรชโซลด์ ซึ่งเท่ากับ 120 ดังแสดงดังรูปที่ 3.9 (ข) จากนั้นนำค่าธรชโซลด์ที่ได้มาสร้างภาพไบนารีได้โดยใช้สมการที่ (3.3) ซึ่งจะได้ภาพผลลัพธ์ดังแสดงในรูปที่ 3.9 (ค) จะเห็นได้ว่าภาพที่ได้จะมีความคมชัดของตัวอักษร และสามารถแยกแยะระหว่างแบกกราวด์และตัวอักษรได้ดี ทำให้ได้ภาพผลลัพธ์ที่มีความคมชัดสวยงามในส่วนที่วัตถุ แต่วิธีนี้ไม่เหมาะสมกับภาพที่ไม่สามารถแยกกลุ่มของสิ่งที่อยู่ในภาพได้อย่างชัดเจน ระหว่างกลุ่มของวัตถุและกลุ่มของแบกกราวด์

### 3.3.3 การหาค่าธรชโซลด์จากค่า Root Mean Square (RMS)

การหาค่าธรชโซลด์จากค่า Root Mean Square (RMS) [19] เป็นการหาค่าธรชโซลด์ที่แตกต่างจากสองวิธีที่กล่าวมา สำหรับวิธีนี้จะเป็นการหาค่าธรชโซลด์โดยอัตโนมัติโดยไม่ต้องให้ผู้ใช้เป็นผู้กำหนดค่าธรชโซลด์ ซึ่งการหาค่าธรชโซลด์วิธีนี้ได้อาศัยการคำนวณพื้นฐานทางสถิติในเรื่องของการหาค่ากลางหรือค่าเฉลี่ย (Mean) มาประยุกต์ใช้ ค่าธรชโซลด์ที่คำนวณได้จะเป็นค่าที่ได้จากค่ากึ่งกลางระหว่างค่าระดับความเข้มสูงสุด และค่าระดับความเข้มต่ำสุดของข้อมูลภาพอินพุต สำหรับการคำนวณหาค่าธรชโซลด์จากค่า Root Mean Square (RMS) นี้สามารถคำนวณได้จากสมการที่ (3.5) โดยนำค่าคงที่ค่าหนึ่งคูณกับค่า RMS

$$RMS = \left[ \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)^2 \right]^{1/2} \quad (3.4)$$

$$Thr = k (RMS) \quad (3.5)$$

โดยที่

$Thr$	ค่าธรชโซลด์
$k$	ค่าคงที่
$RMS$	ค่าเฉลี่ย Root Mean Square
$f(x, y)$	ข้อมูลภาพอินพุต ที่มีระดับความเข้ม 0 ถึงระดับ $L$
$M$	จำนวนจุดภาพตามความกว้างของภาพอินพุต
$N$	จำนวนจุดภาพตามความสูงของภาพอินพุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อกำหนดค่าเรขาคณิตจากค่าเฉลี่ย Root Mean Square (RMS) ได้แล้ว ก็สามารถสร้างภาพไบนารีได้โดยนำค่าเรขาคณิตที่ได้มาแทนค่าในสมการที่ (3.3) ตัวอย่างของการใช้เรขาคณิตจากค่าเฉลี่ย Root Mean Square (RMS) ได้แสดงดังรูปที่ 3.10 โดยที่ได้เลือกใช้ภาพของตัวหนังสือบนพื้นกระดาษหนังสือพิมพ์ ซึ่งได้ทำการสแกนเป็นภาพที่มีระดับเทา 256 ระดับเป็นภาพอินพุตที่แสดงในรูปที่ 3.10(ก) ส่วนภาพผลลัพธ์ที่ได้จากการทำเรขาคณิตโดยการหาค่าเฉลี่ย Root Mean Square ได้แสดงให้เห็นดังรูปที่ 3.10(ข) ซึ่งจะเห็นว่าภาพผลลัพธ์ที่ได้จากวิธีนี้สามารถแยกตัวหนังสือกับแบกกราวด์ได้ แต่ภาพอาจจะมีสิ่งที่ไม่พึงประสงค์ปะปนมาด้วยอันเนื่องมาจากค่าเรขาคณิตที่ได้เป็นการคำนวณค่าเฉลี่ย RMS จากภาพอินพุตที่ต้องการทำเรขาคณิต จึงทำให้ได้ค่าเรขาคณิตที่ไม่เหมาะสมแต่ผลของภาพที่ได้ก็สามารถยอมรับได้ แล้วข้อดีของวิธีนี้คือผู้ใช้ไม่ต้องมากำหนดค่าเรขาคณิตเอง เพราะค่าเรขาคณิตที่ได้เป็นการคำนวณค่าเฉลี่ย RMS ซึ่งเหมาะกับการทำเรขาคณิตแบบอัตโนมัติ

## พ่อเผย.โต้-หมอบใส่ร้าย! ลูกไม่โต้กลับไสยศาสตร์

ดร.ยันคดีพ.ญ.หายตัว ให้ความเป็นธรรมทุกฝ่าย จะไม่ตอบโต้สภาฯ ส่วนที่สามมีไปแถลงดลุ่มเมียทางทีวีนั้นจะไม่นำมาประกอบคดี เพราะถือว่ามีไม่ได้ให้การกับตำรวจ ส่งกำลังเข้าตรวจค้นสำนักเข้าทรง รวมถึงสถานปฏิบัติธรรมที่เมืองจันทน์  อ่านต่อหน้า 8

(ก)

## พ่อเผย.โต้-หมอบใส่ร้าย! ลูกไม่โต้กลับไสยศาสตร์

ดร.ยันคดีพ.ญ.หายตัว ให้ความเป็นธรรมทุกฝ่าย จะไม่ตอบโต้สภาฯ ส่วนที่สามมีไปแถลงดลุ่มเมียทางทีวีนั้นจะไม่นำมาประกอบคดี เพราะถือว่ามีไม่ได้ให้การกับตำรวจ ส่งกำลังเข้าตรวจค้นสำนักเข้าทรง รวมถึงสถานปฏิบัติธรรมที่เมืองจันทน์  อ่านต่อหน้า 8

(ข)

รูปที่ 3.10 การทำเรขาคณิต โดยการหาค่าเรขาคณิตจากค่าเฉลี่ย Root Mean Square (ก) ภาพข้อมูลอินพุต 256 ระดับเทา (ข) แสดงภาพไบนารีที่ได้จากการทำเรขาคณิต

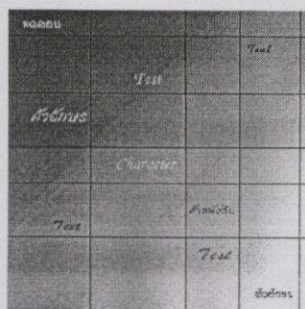
จากที่กล่าวมาสำหรับการสร้างภาพไบนารีโดยเทคนิคเรซโซลด์ จะเหมาะกับภาพที่มีความแตกต่างกันมากระหว่างวัตถุและแบกกราวด์ โดยที่สามารถแยกได้ระหว่างวัตถุและแบกกราวด์ สำหรับวิธีการประมวลผลนั้น สิ่งที่สำคัญที่สุดก็คือค่าเรซโซลด์ โดยที่แต่ละวิธีในการคำนวณหาค่าเรซโซลด์นั้น วิธีการหาค่าเรซโซลด์โดยการกำหนดค่าล่วงหน้า ภาพผลลัพธ์ที่ได้จะขึ้นอยู่กับประสบการณ์ของผู้ใช้นั้นๆ สำหรับวิธีหาค่าเรซโซลด์โดยการพิจารณาจากฮิสโตแกรม จะเห็นได้ว่าวิธีนี้จะเป็นวิธีที่ยืดหยุ่นที่สุดในการคำนวณหาค่าเรซโซลด์ คือค่าเรซโซลด์ที่คำนวณได้นั้น จะเป็นค่าน้อยที่สุดของจำนวนความเข้มที่อยู่กึ่งกลางระหว่างกลุ่มของข้อมูลสองกลุ่ม ดังนั้นค่าที่คำนวณได้มานั้นจะมีความเหมาะสมที่สุดเพื่อที่จะได้ภาพผลลัพธ์ที่มีความเหมาะสม จากการทดลองสร้างภาพไบนารีโดยใช้เทคนิคเรซโซลด์ จะเห็นได้ว่าภาพข้อมูลอินพุทที่จะนำมาสร้างภาพไบนารีโดยเทคนิคนี้ ควรจะเป็นภาพที่มีลักษณะแตกต่างกันมากๆ สามารถแยกได้เป็นสองส่วนคือ ส่วนของวัตถุและส่วนของแบกกราวด์ เช่นภาพหน้าเอกสารที่มีลักษณะที่มีสีเข้มบนพื้นกระดาษที่มีความเข้มอ่อนกว่า ส่วนวิธีการหาค่าเรซโซลด์จากค่าเฉลี่ย Root Mean Square (RMS) ค่าเรซโซลด์ที่คำนวณได้จะเป็นค่าเฉลี่ยที่อยู่ระหว่างความเข้มต่ำสุดและสูงสุดของภาพอินพุท จะพบว่าภาพผลลัพธ์ที่ได้จะเหมาะสมและคมชัดก็ต่อเมื่อค่าเรซโซลด์นี้สามารถแบ่งแยกข้อมูลออกเป็นสองส่วนได้อย่างชัดเจน ซึ่งถ้าค่าเรซโซลด์ที่คำนวณได้ไม่สามารถแบ่งแยกได้ คือค่าเรซโซลด์ที่คำนวณได้เอนเอียงไปทางระดับความเข้มที่ต่ำหรือสูงเกินไป ภาพผลลัพธ์ที่ได้ก็จะไม่เหมาะสม เช่นอาจจะขาวเกินไป (รายละเอียดไม่ครบ) หรือดำเกินไป (เกิดสิ่งไม่พึงประสงค์ปะปนมาด้วย (noise)) แต่ข้อดีของวิธีนี้คือผู้ใช้ไม่ต้องมากำหนดค่าเรซโซลด์ขึ้นเอง เหมาะสำหรับการทำการประมวลผลแบบอัตโนมัติ

### 3.4 การเข้ารหัสรันเล็งซ์ (Run-Length Encoding)

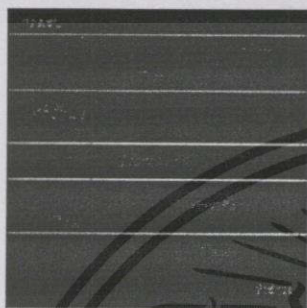
การแทนภาพไบนารีสามารถใช้การเข้ารหัสรันเล็งซ์นำมาแสดงความยาวของจำนวนจุดภาพ ในการแทนเช่นนี้การแสดงจำนวนความยาวการวิ่งของพิกเซล 1 ในภาพถูกใช้แทนภาพ การเข้ารหัสเช่นนี้ได้ถูกใช้ในการส่งข้อมูลภาพ นอกจากนั้นคุณสมบัติบางข้อเช่นพื้นที่ของวัตถุจะถูกคำนวณจากรหัสรันเล็งซ์ของวัตถุโดยตรง [20] วิธีที่ใช้กันทั่วไปในการเข้ารหัสรันเล็งซ์คือกำหนดให้ buffer ( $b$ ) เป็นตัวเก็บความยาวบิต จากนั้นใช้พิกเซลค่า 1 เป็นตำแหน่งเริ่มต้น และให้พิกเซลค่า 1 ที่ตำแหน่งเริ่มต้นนี้เป็นตัวอ้างอิง จากนั้นกำหนดให้ Counter ( $c$ ) เท่ากับ 1 จากนั้นพิจารณาพิกเซลถัดไปที่อยู่ด้านขวา ถ้าพิกเซลนั้นเหมือนกับพิกเซลปัจจุบัน กำหนด  $c = c + 1$  ถ้าพิกเซลนั้นต่างจากพิกเซลปัจจุบันเก็บค่า  $c$  ใน buffer ของความยาว และกำหนดให้  $c = 1$  กระทำต่อไปจนจบแถว ดังนั้นการวิ่งความยาวในแต่ละครั้งจะถูกเก็บโดยการใช้  $b$  บิต รูปที่ 3.11 แสดงรูปแบบรหัสรันเล็งซ์ของภาพไบนารี



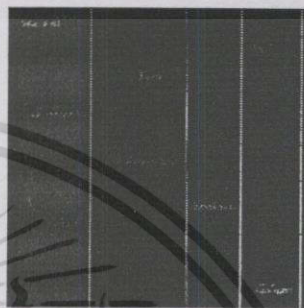
ตัวอย่าง การนำเทคนิคการทำเรซโซลต์และการเข้ารหัสรันเล็งธมาประยุกต์ใช้งานร่วมกันกับการแปลงเวฟเล็ทสเตชันนารีแสดงให้เห็นดังรูปที่ 3.13 ซึ่งได้ทดลองกับภาพที่ได้ทำการสังเคราะห์สร้างขึ้นซึ่งเป็นข้อมูลภาพขนาด  $512 \times 512$  พิกเซลที่มีรายละเอียดของเส้นตรงทั้งทางด้านแนวนอนและแนวตั้งพร้อมทั้งตัวหนังสือ และแบกกราวด์ระดับเทา (Grayscale) ดังรูปที่ 3.13 (ก) จากนั้นเรานำข้อมูลภาพที่เราได้ทำการสังเคราะห์สร้างขึ้นมาผ่านขบวนการแปลงเวฟเล็ทสเตชันนารีซึ่งจะให้ผลภาพเอาท์พุทออกมาเป็น 4 ภาพแบนด์คือภาพแบนด์ LL LH HL และ HH แต่ในที่นี้ได้แสดงเฉพาะภาพแบนด์ LH และ HL เท่านั้นเพราะคุณสมบัติของสองภาพนี้สามารถค้นหาเส้นอ้างอิงได้ดังแสดงในรูปที่ 3.13 (ข) และ 3.13 (ค) ซึ่งเป็นภาพผลลัพธ์ในแบนด์ LH และ HL ตามลำดับ ซึ่งจะเห็นได้ชัดว่ามีเพียงเส้นตรงแนวนอนเท่านั้นที่ยังคงอยู่ในภาพแบนด์ LH ในขณะที่มีเพียงเส้นตรงแนวตั้งเท่านั้นที่อยู่ในภาพแบนด์ HL โดยที่ภาพแบนด์ LH และ HL จะเก็บเฉพาะที่เป็นเส้นตรงเท่านั้น แบกกราวด์ระดับเทาจะไม่ปรากฏให้เห็นในภาพเอาท์พุท เพราะว่าความถี่ของมันสูงกว่าความถี่ของเส้นตรงแต่ข้อมูลภาพที่ได้จากการแปลงเวฟเล็ทสเตชันนารีนั้นยังคงมีสิ่งไม่พึงประสงค์ปะปนมาด้วย จึงต้องนำไปผ่านขบวนการสร้างภาพไบนารีเพื่อทำการจำแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ด้วยเทคนิคการทำเรซโซลต์โดยวิธีการคำนวณหาค่าเรซโซลต์จากค่าเฉลี่ย Root Mean Square (RMS) ในภาพแบนด์ LH และ HL ซึ่งจะได้ผลลัพธ์ดังรูปที่ 3.13(ง) และ 3.13 (จ) ตามลำดับ จากผลภาพที่ผ่านเทคนิคการทำเรซโซลต์จะเห็นว่ามีความถี่ของตัวหนังสือปะปนมาด้วยเนื่องจากว่าช่วงความถี่ของตัวหนังสือใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง ดังนั้นจึงได้ใช้อัลกอริธึมของการเข้ารหัสรันเล็งธมาประยุกต์ใช้เพื่อกำจัดตัวหนังสือเหล่านี้ และเก็บรักษาเฉพาะความยาวของเส้นอ้างอิงเท่านั้น ซึ่งได้แสดงผลลัพธ์ดังรูปที่ 3.13 (ฉ) และ (ช) ตามลำดับ



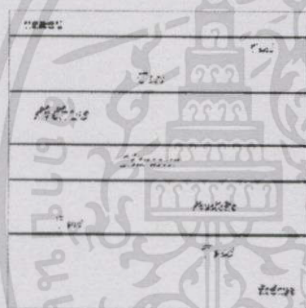
(ก)



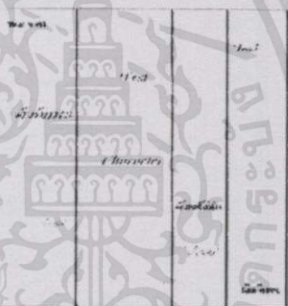
(ข)



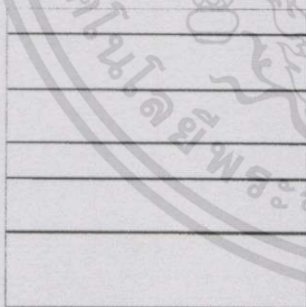
(ค)



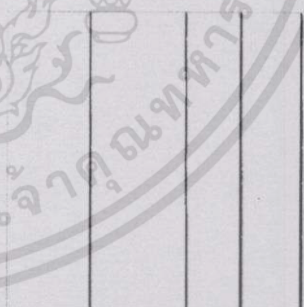
(ง)



(จ)



(ฉ)



(ช)

รูปที่ 3.13 การค้นหาเส้นอ้างอิงโดยการประยุกต์ใช้งานร่วมกันของการแปลงเวฟเล็ตสแตชันนารี เทคนิคการทำเรซโซลต์ และอัลกอริทึมการเข้ารหัสรันเล็งซ์ (ก) ภาพต้นแบบที่ได้ทำการสังเคราะห์สร้างขึ้น (ข) ภาพแบนด์ LH (ค) ภาพแบนด์ HL (ง) ภาพแบนด์ LH ผ่านการทำเรซโซลต์ (จ) ภาพแบนด์ HL ผ่านการทำเรซโซลต์ (ฉ) ภาพแบนด์ LH ที่ผ่านการกำจัดสิ่งไม่พึงประสงค์ และ (ช) ภาพแบนด์ HL ที่ผ่านการกำจัดสิ่งไม่พึงประสงค์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

### ผลการทดลอง

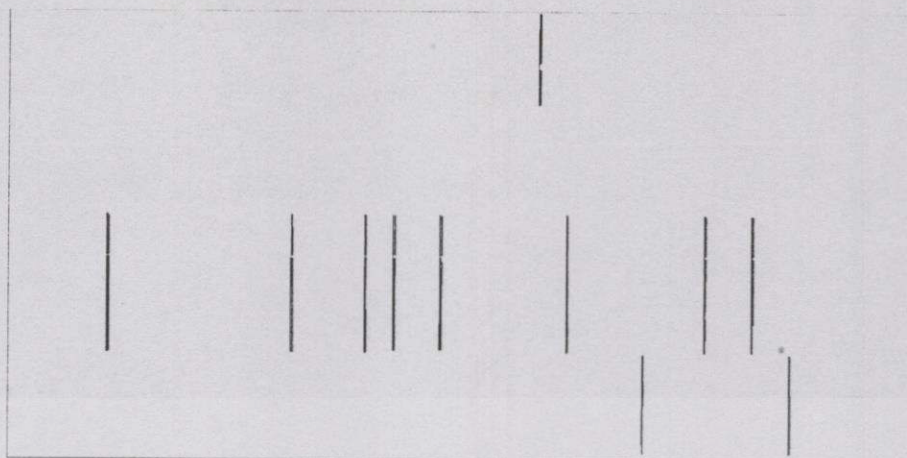
ในบทนี้ได้กล่าวถึงการทดลองโดยใช้หลักการของการแปลงเวฟเล็ทสเตชันนารีและเทคนิคการทำเรซโซลต์ที่ได้กล่าวไว้ในบทที่ 3 มาประยุกต์ใช้กับภาพฟอร์มเอกสารจริง โดยได้นำฟอร์มเอกสารที่มีรายละเอียดของเส้นอ้างอิงเช่นใบรับฝาก-ถอนเงินของธนาคาร หรือ เช็คของบัญชี โดยได้นำฟอร์มเอกสารเหล่านี้มาทำการสแกนด้วยเครื่องสแกนเนอร์ด้วยการดิจิทัลเป็นภาพดิจิทัล 256 ระดับเทา (gray scale) โดยใช้รีโซลูชัน 200 จุดภาพต่อนิ้ว (Dpi.) จากนั้นนำภาพฟอร์มเอกสาร 256 ระดับเทาที่ได้จากการสแกนมาผ่านขบวนการแปลงเวฟเล็ทสเตชันนารีเพื่อทำการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร โดยได้ใช้เทคนิคการทำเรซโซลต์และอัลกอริทึมของการเข้ารหัสรันเล็งส์มาใช้ในการปรับปรุงภาพเส้นอ้างอิงที่ได้จากการแปลงเวฟเล็ท การทดลองทั้งหมดนี้ได้จำลองการทำงานด้วยการเขียนโปรแกรมที่ทำงานด้วย MATLAB เวอร์ชัน 5.1 ประมวลผลบนเครื่อง PC Pentium III 500 MHz

#### 4.1 ผลการทดลอง

ในส่วนนี้ได้ทดลองค้นหาเส้นอ้างอิงในฟอร์มเอกสาร โดยนำใบรับฝากถอนเงินของธนาคารไทยพาณิชย์ที่มีรายละเอียดของเส้นอ้างอิงทั้งในทิศทางด้านแนวนอนและแนวตั้งมาทำการทดลอง โดยนำใบรับฝากถอนเงินของธนาคารไทยพาณิชย์มาผ่านการสแกนโดยดิจิทัลเป็นข้อมูลภาพ 256 ระดับเทาซึ่งแสดงดังรูปที่ 4.1 (ก) จากนั้นนำภาพที่ได้จากการสแกนมาผ่านขบวนการแปลงเวฟเล็ทสเตชันนารีด้วยโปรแกรม MATLAB เวอร์ชัน 5.1 โดยใช้ฟิลเตอร์ของเวฟเล็ทที่ Daubechies-4 (D4) ซึ่งจะให้ผลภาพเอาท์พุทในแบนด์ LH และแบนด์ HL ดังรูปที่ 4.1 (ข) และ (ค) ตามลำดับ จากผลภาพในแบนด์ LH และ HL จะเห็นว่าแต่ละภาพแบนด์สามารถตรวจจับเส้นอ้างอิงได้ โดยที่เส้นอ้างอิงทางด้านแนวนอนปรากฏอยู่ในภาพแบนด์ LH ส่วนเส้นอ้างอิงทางด้านแนวตั้งจะปรากฏอยู่ในภาพแบนด์ HL และแบกกราวด์ระดับเทาถูกกำจัดทิ้งไปในทั้งสองภาพแบนด์นี้ แต่ข้อมูลภาพเส้นอ้างอิงทั้งสองภาพแบนด์ยังคงมีสิ่งที่ไม่พึงประสงค์ปะปนมาด้วย จึงได้ใช้หลักการของการสร้างภาพไบนารีมาทำการจำแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ โดยใช้เทคนิคการทำเรซโซลต์ซึ่งใช้การหาค่าเรซโซลต์จากค่าเฉลี่ย Root Mean Square (RMS) โดยจะได้ภาพผลลัพธ์จากเทคนิคการทำเรซโซลต์ดังรูปที่ 4.1 (ง) และ (จ) ตามลำดับ จากผลลัพธ์ของภาพที่ผ่านการทำเรซโซลต์จะเห็นว่าเทคนิคการทำเรซโซลต์ไม่สามารถกำจัดตัวหนังสือได้หมด เพราะว่าช่วงความถี่ของตัวหนังสือมีความถี่ใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง จึงได้นำหลัก



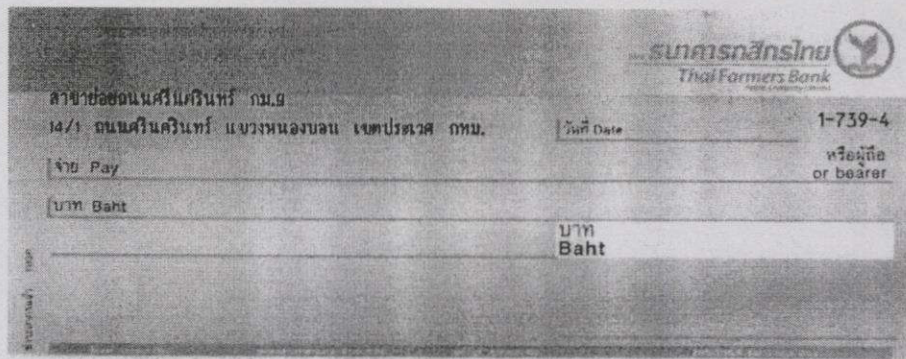




(ข) ภาพแบนด์ HL ที่ผ่านการกำจัดสิ่งไม่พึงประสงค์โดยใช้อัลกอริทึมของรันเล็งซ์

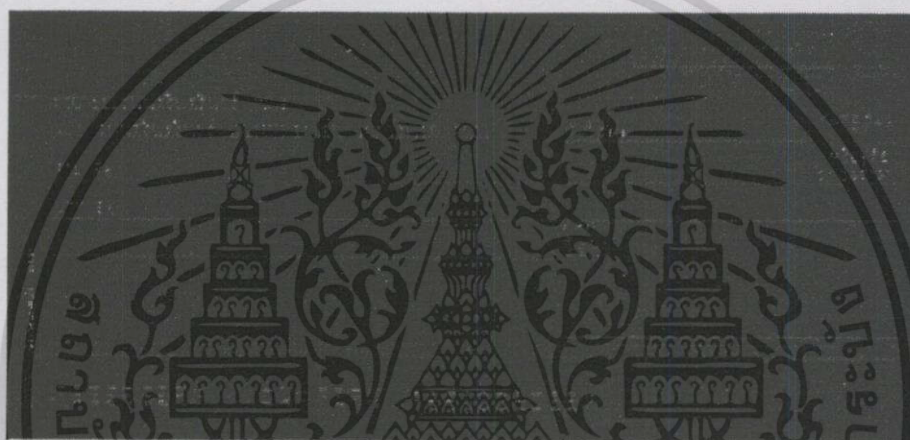
#### รูปที่ 4.1 ก) – ข) ผลการทดลองการค้นหาเส้นอ้างอิงจากข้อมูลภาพใบรับฝากถอนเงินของธนาคารไทยพาณิชย์

ต่อมาได้นำฟอร์มเอกสารเช็คของธนาคารกสิกรไทยมาทำการทดลองซึ่งได้แสดงผลการทดลองดังรูปที่ 4.2 โดยได้ฟอร์มเอกสารเช็คของธนาคารกสิกรไทยที่มีรายละเอียดของเส้นอ้างอิงมาผ่านการสแกนด้วยเครื่องสแกนเนอร์โดยดิจิทัลซ์เป็นข้อมูลภาพ 256 ระดับเทาซึ่งแสดงดังรูปที่ 4.2 (ก) จากนั้นนำภาพที่ได้จากการสแกนมาผ่านขบวนการแปลงเวฟเล็ทสเคชันนารีด้วยโปรแกรม MATLAB เวอร์ชัน 5.1 โดยใช้ความยาวของฟิลเตอร์ของเวฟเล็ทที่ Daubechies-4 (D4) ซึ่งจะให้ผลภาพเอาท์พุทในแบนด์ LH ดังรูปที่ 4.2 (ข) ในที่นี้ได้แสดงเฉพาะภาพผลลัพธ์ในแบนด์ LH เท่านั้น เพราะว่าภาพต้นแบบที่มาทำการทดลองมีรายละเอียดเฉพาะเส้นอ้างอิงแนวนอน จากผลภาพในแบนด์ LH จะเห็นว่าภาพแบนด์ LH สามารถตรวจจับเส้นอ้างอิงในทิศทางด้านแนวนอนได้พร้อมทั้งสามารถกำจัดแบกกราวด์ระดับเทาได้ แต่ข้อมูลภาพเส้นอ้างอิงที่ได้ยังคงมีสิ่งไม่พึงประสงค์ปะปนมาด้วย จึงต้องนำภาพแบนด์ LH ไปสร้างภาพไบนารีเพื่อจำแนกข้อมูลระหว่างเส้นอ้างอิงและแบกกราวด์โดยเทคนิคการทำเรซโซลต์ด้วยการคำนวณหาค่าเรซโซลต์จากการหาค่าเฉลี่ย Root Mean Square (RMS) ของภาพแบนด์ LH ซึ่งจะให้ผลของภาพที่ผ่านขบวนการทำเรซโซลต์ดังรูปที่ 4.2 (ค) จากผลลัพธ์ของภาพที่ผ่านการทำเรซโซลต์จะเห็นว่าเทคนิคการทำเรซโซลต์ไม่สามารถกำจัดตัวหนังสือได้หมด เพราะว่าช่วงความถี่ของตัวหนังสือมีความถี่ใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง จึงได้นำหลักการของการเข้ารหัสรันเล็งซ์มาประยุกต์ใช้ เพื่อกำจัดรายละเอียดของตัวหนังสือและสิ่งที่ไม่พึงประสงค์บางส่วน ซึ่งจะแสดงผลดังรูปที่ 4.2 (ง)

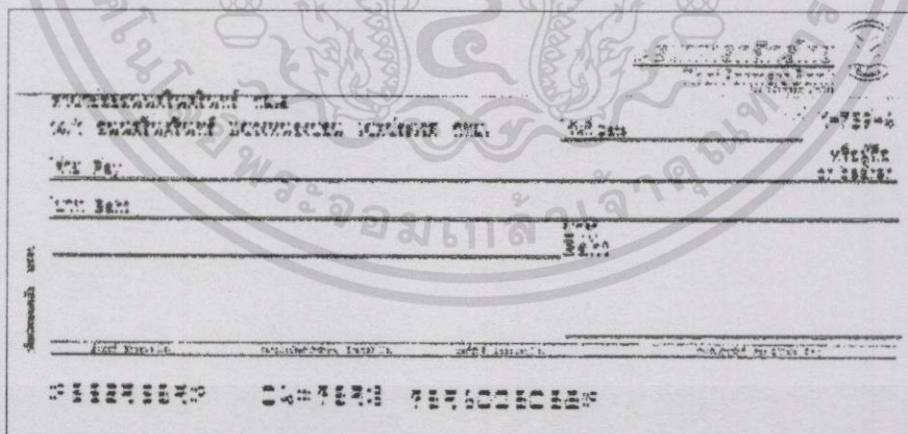


# 2259239# 04 739 739 1003038#

(ก) ภาพต้นแบบที่ได้จากการสแกน

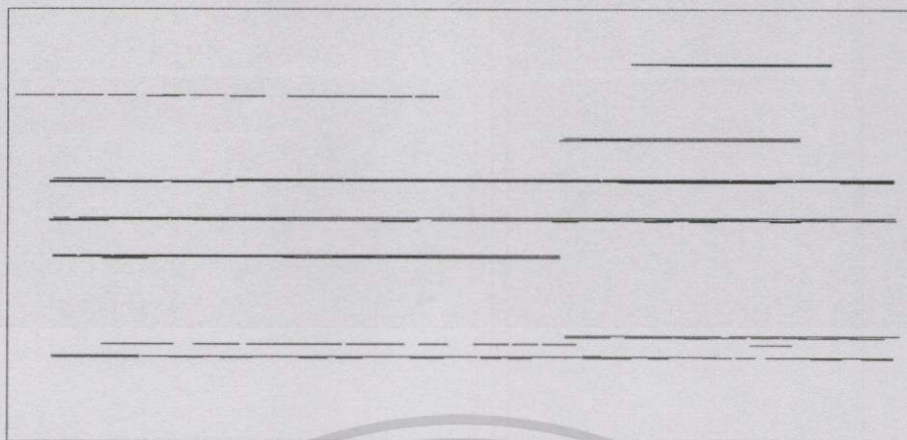


(ข) ภาพแบบค้ LH



(ค) ภาพแบบค้ LH ที่ผ่านการทำเรซโซลต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ง) ภาพแบนด์ LH ที่ผ่านการกำจัดสิ่งไม่พึงประสงค์โดยใช้อัลกอริทึมของรันเล็งส์

#### รูปที่ 4.2 ก) – ง) ผลการทดลองการค้นหาเส้นอ้างอิงจากข้อมูลภาพฟอร์มเอกสารเช็คของธนาคาร กสิกรไทย

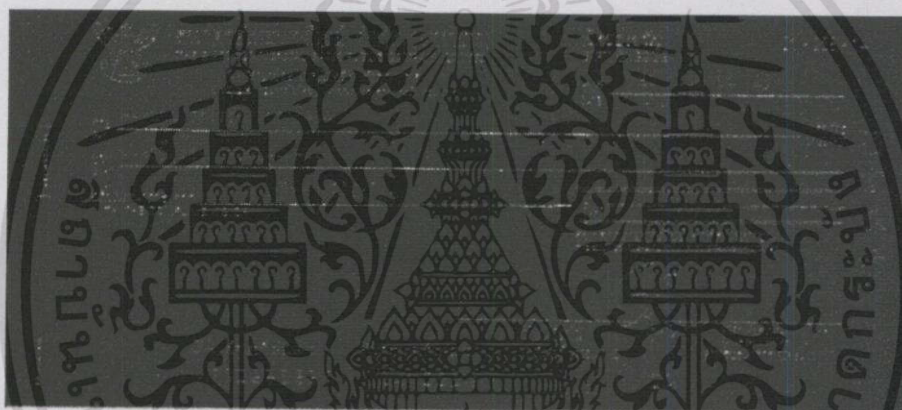
ในส่วนนี้ได้้นำฟอร์มเอกสารเช็คของบัญชีของธนาคารไทยพาณิชย์มาทำการทดลอง ซึ่งแสดงผลการทดลองดังรูปที่ 4.3 โดยฟอร์มเอกสารเช็คของบัญชีของธนาคารไทยพาณิชย์ที่ได้นำมาทำการทดลองนี้มีรายละเอียดของแบกกราวด์ที่เป็นรูปภาพระดับเทา และมีรายละเอียดของเส้นอ้างอิงที่ธนาคารได้พิมพ์ไว้เพื่อให้ผู้ใช้กรอกข้อมูลรายละเอียดตามที่ธนาคารได้กำหนดไว้ โดยนำเช็คของบัญชีของธนาคารไทยพาณิชย์มาทำการสแกนโดยดิจิทัลเป็นข้อมูลภาพ 256 ระดับเทาซึ่งแสดงดังรูปที่ 4.3 (ก) จากนั้นนำข้อมูลภาพที่ได้จากการสแกนมาผ่านขั้นตอนการแปลงเวฟเล็ตสเคชันนารีด้วยโปรแกรม MATLAB เวอร์ชัน 5.1 โดยใช้ความยาวของฟิลเตอร์ของเวฟเล็ตที่ Daubechies-4 (D4) ซึ่งจะให้ผลภาพเอาท์พุทในแบนด์ LH ดังรูปที่ 4.3 (ข) ในที่นี้ได้แสดงเฉพาะรายละเอียดของภาพแบนด์ LH เช่นเดียวกับส่วนที่แล้ว เพราะว่าภาพต้นแบบที่มาทำการทดลองมีรายละเอียดเฉพาะเส้นอ้างอิงแนวนอน จากผลที่ได้ในภาพแบนด์ LH นั้นจะเห็นว่ามีเฉพาะเส้นอ้างอิงในด้านแนวนอนเท่านั้นที่ปรากฏอยู่พร้อมทั้งสามารถกำจัดภาพแบกกราวด์ระดับเทาได้ แต่ข้อมูลภาพเส้นอ้างอิงที่ได้ยังคงมีสิ่งไม่พึงประสงค์ปะปนมาด้วย จึงต้องนำภาพแบนด์ LH ไปสร้างภาพไบนารีเพื่อจำแนกข้อมูลระหว่างเส้นอ้างอิงและแบกกราวด์โดยเทคนิคการทำเรซโซลต์ด้วยการคำนวณหาค่าเรซโซลต์จากการหาค่าเฉลี่ย Root Mean Square (RMS) ของภาพแบนด์ LH ซึ่งจะให้ผลของภาพที่ผ่านขั้นตอนการทำเรซโซลต์ดังรูปที่ 4.3 (ค) จากผลลัพธ์ของภาพที่ผ่านการทำเรซโซลต์จะเห็นว่าเทคนิคการทำเรซโซลต์ไม่สามารถกำจัดตัวหนังสือได้หมด เพราะว่าช่วงความถี่ของตัวหนังสือมีความถี่ใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง จึงได้นำหลักการของการเข้ารหัสรันเล็งส์มาประยุกต์

ใช้ เพื่อกำจัดรายละเอียดของตัวหนังสือและสิ่งที่ไม่พึงประสงค์บางส่วน ซึ่งจะแสดงผลดังรูปที่

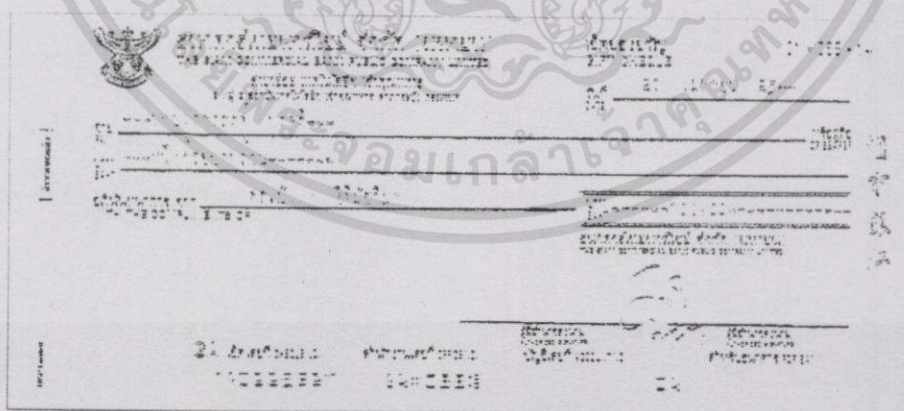
4.3 (ง)



(ก) ภาพต้นแบบที่ได้จากการสแกน

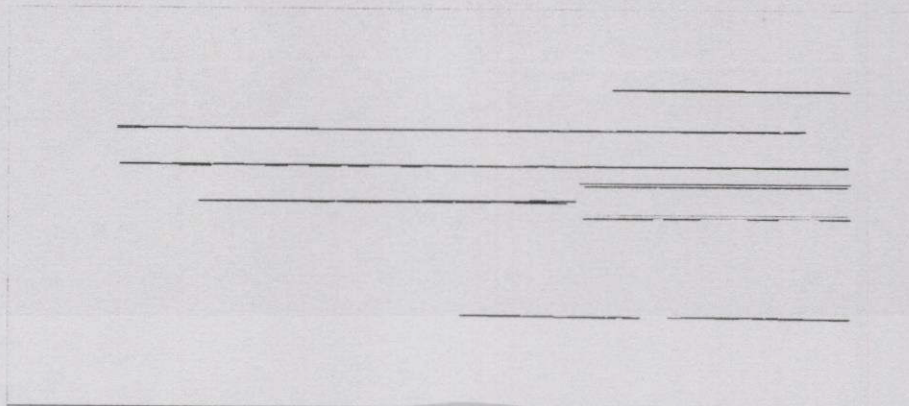


(ข) ภาพแบนด์ LH



(ค) ภาพแบนด์ LH ที่ผ่านการทำเรซโซลด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ง) ภาพแบนด์ LH ที่ผ่านการกำจัดสิ่งไม่พึงประสงค์โดยใช้อัลกอริทึมของรันเล็งซ์

รูปที่ 4.3 ก) – ง) ผลการทดลองการค้นหาเส้นอ้างอิงจากข้อมูลภาพเข็ชของขัญของธนาการไทย พาณิษฐ์

#### 4.2 สรุปผลการทดลอง

จากผลการทดลองพบว่าผลลัพธ์ที่ได้จากขบวนการแปลงเวฟเล็ทสเดชันนารี สามารถค้นหาเส้นอ้างอิงในภาพฟอร์มเอกสารระดับเทาได้ โดยที่ภาพแบนด์ LH สามารถค้นหาเส้นอ้างอิงในทางด้านแนวนอนได้ ส่วนภาพในแบนด์ HL สามารถค้นหาเส้นอ้างอิงในทางด้านแนวตั้งได้ และจากผลลัพธ์ของภาพทั้งในแบนด์ LH และ HL สามารถกำจัดแบกกราวด์ระดับเทาของภาพฟอร์มเอกสารได้ แต่ผลลัพธ์ของข้อมูลภาพเส้นอ้างอิงที่ได้ยังคงมีสิ่งไม่พึงประสงค์ปะปนมาด้วยจึงต้องใช้หลักการของการสร้างภาพไบนารีมาช่วยในการจำแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ โดยนำเทคนิคการทำเรซโซลด์เข้ามาประยุกต์ใช้ ซึ่งผลลัพธ์ที่ได้ของภาพจากการทำเรซโซลด์สามารถจำแนกส่วนของเส้นอ้างอิงให้เด่นชัดยิ่งขึ้น แต่ภาพผลลัพธ์ที่ได้จากเทคนิคการทำเรซโซลด์ไม่สามารถกำจัดตัวหนังสือ หรือสิ่งไม่พึงประสงค์บางส่วนได้หมดเนื่องจากว่า ช่วงความถี่เหล่านั้นใกล้เคียงกับช่วงความถี่ของเส้นอ้างอิง จึงได้นำอัลกอริทึมของการเข้ารหัสรันเล็งซ์มาประยุกต์ใช้ เพื่อกำจัดสิ่งไม่พึงประสงค์เหล่านี้ และเก็บรักษาเฉพาะรายละเอียดของเส้นอ้างอิงเท่านั้น

## บทสรุปและแนวทางในการพัฒนา

## 5.1 สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้ได้ทำการศึกษาและวิจัยเกี่ยวกับการค้นหาเส้นอ้างอิงในภาพฟอร์มเอกสารระดับเทา โดยนำฟอร์มเอกสารมาทำการสแกนด้วยเครื่องสแกนเนอร์เป็นภาพฟอร์มเอกสาร 256 ระดับเทา โดยใช้ความละเอียดของการสแกน 200 จุดภาพต่อนิ้ว แล้วใช้เครื่องคอมพิวเตอร์ทำการประมวลผล โดยนำเสนอวิธีการแปลงเวฟเล็ทซึ่งได้ใช้การแปลงเวฟเล็ทสแตชันนารีมาแปลงภาพฟอร์มเอกสารระดับเทาออกมาเป็นสีภาพแบนด์ และจากสองในสีภาพแบนด์นี้สามารถค้นหาเส้นอ้างอิงในภาพฟอร์มเอกสารที่มีแบกกราวด์หลายระดับเทาได้ โดยที่ผลของภาพเส้นอ้างอิงที่ได้จากแปลงเวฟเล็ทสแตชันนารีนี้ยังไม่ดีเท่าที่ควร เพราะว่ามีสิ่งไม่พึงประสงค์ปะปนมาด้วย เช่น สัญลักษ์ณ์ หรือ ตัวอักษรข้อความต่างๆ ดังนั้นจึงต้องทำการปรับปรุงภาพเส้นอ้างอิงให้มีรายละเอียดของเส้นอ้างอิงให้คมชัดยิ่งขึ้น และกำจัดสิ่งไม่พึงประสงค์ทิ้งไป โดยใช้หลักการสร้างภาพไบนารี และการเข้ารหัสรันทรงศาทำการปรับปรุงภาพเส้นอ้างอิง

การสร้างภาพไบนารีนั้นได้ใช้เทคนิคการทำเรซโซลด์มาทำการจับแแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ โดยการหาค่าเรซโซลด์จากค่าเฉลี่ย Root Mean Square (RMS) จากภาพเส้นอ้างอิงที่ได้จากการแปลงเวฟเล็ท สำหรับการหาค่าเรซโซลด์จากค่าเฉลี่ย RMS ของภาพเส้นอ้างอิงนั้น เราได้เลือกใช้ค่า  $k = 6$  คูณกับค่าเฉลี่ย RMS ซึ่งจะทำให้ภาพที่ผ่านการปรับปรุงโดยเทคนิคการทำเรซโซลด์นี้สามารถจับแแนกข้อมูลภาพระหว่างเส้นอ้างอิงและแบกกราวด์ได้ โดยที่ข้อมูลภาพเส้นอ้างอิงนั้นจะเป็นจุดภาพที่เป็นค่า 0 ที่เรียงต่อเนื่องติดๆ กันไป แต่ข้อมูลภาพที่ผ่านการปรับปรุงภาพด้วยเทคนิคการทำเรซโซลด์นี้ยังมีข้อมูลที่ไม่พึงประสงค์ปะปนมาด้วย เช่น สัญลักษ์ณ์บางอย่าง หรือกลุ่มของตัวอักษร ดังนั้นจึงได้ใช้อัลกอริธึมการเข้ารหัสรันทรงศาประยุกต์ใช้ในการกำจัดในสิ่งเหล่านี้

อัลกอริธึมของการเข้ารหัสรันทรงศาที่นำมาประยุกต์ใช้นี้ จะนำมาตรวจสอบข้อมูลจุดภาพที่เป็นค่า 0 ที่เรียงต่อเนื่องติดๆ กันไปซึ่งก็คือรายละเอียดของเส้นอ้างอิง โดยใช้การเข้ารหัสรันทรงศาตรวจสอบข้อมูลจุดภาพที่เป็นค่า 0 ที่เรียงต่อเนื่องติดๆ กันที่น้อยกว่าหรือเท่ากับ 50 ตัว ให้ทำการเปลี่ยนจุดภาพค่า 0 นั้นเป็นค่า 1 ส่วนข้อมูลจุดภาพที่เป็นเป็นค่า 0 ที่เรียงต่อเนื่องติดๆ กันมากกว่า 50 ตัวให้เป็นค่า 0 เหมือนเดิม ที่เลือกกำหนดจุดภาพอ้างอิงเท่ากับ 50 ก็เพราะว่าข้อมูลภาพที่ทำการสแกนเข้ามาใช้รายละเอียดในการสแกน 200 จุดภาพต่อ 1 นิ้ว ดังนั้นข้อมูลจุดภาพ 50 จุดภาพเท่ากับ 0.25 นิ้ว ซึ่งหมายความว่าถ้าข้อมูลของเส้นอ้างอิงยาวไม่ถึง 0.25 นิ้วให้กำจัดทิ้งไป ถ้าข้อมูลของ

เส้นอ้างอิงที่ยาวเกิน 0.25 นิ้วให้เก็บไว้ ซึ่งจากการทำเช่นนี้สามารถกำจัดสัญลักษณ์ และกลุ่มตัวอักษรที่ไม่ต้องการทิ้งไปได้ และคงไว้แต่เฉพาะข้อมูลของเส้นอ้างอิง

## 5.2 แนวทางในการพัฒนาต่อไป

สำหรับแนวทางในการพัฒนาขั้นตอนต่อไปของการค้นหาเส้นอ้างอิงในฟอร์มเอกสาร ก็คือการหาวิธีที่จะดึงนำข้อมูลที่อยู่บนเส้นอ้างอิงออกมาเพื่อจัดเก็บลงในระบบฐานข้อมูล ซึ่งเมื่อเราสามารถทราบตำแหน่งของเส้นอ้างอิงแล้วจะทำให้เราทราบตำแหน่งของข้อมูลบนเส้นอ้างอิงด้วยเช่นกัน ดังนั้นจึงสามารถนำข้อมูลบนเส้นอ้างอิงมาทำการแปลความหมายแล้วจัดเก็บลงในระบบฐานข้อมูลได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## เอกสารอ้างอิง

- [1] S.L. Taylor, R. Fritzson, and J.A. Pastor. "Extraction of Data from Preprinted Forms". *Machine Vision and Application*. Vol. 5, 1992. pp. 211-222.
- [2] S. Phonsri, P. Thitimajshima and Y. Rangsaneri, "Extraction of Reference Line in Automated Data-Entry System." 2<sup>nd</sup> International Conference on Information Communications and Signal Processing. Singapore, December 1999.
- [3] A. Grossman and J. Morlet. "Decompositions of hardly functions into square integrable wavelets of constant shape." *SIAM Journal of Mathematical Analysis*. Vol. 15, no. 4, July 1984. pp. 723-736.
- [4] I. Daubechies. "Orthonormal bases of compactly supported wavelets." *Communication in Pure and Applied Mathematics*. Vol. 41, November 1988. pp. 909-996.
- [5] L.W. Couch. **Digital and Analog Communication System**. Upper Saddle River, NJ: Prentice-Hall. 1997.
- [6] M.Vetterli and J. Kovacevic. **Wavelet and Subband Coding**. Englewood Cliffs, NJ: Prentice-Hall. 1995.
- [7] I. Daubechies. "The wavelet transform, time frequency localization and signal anlysis." *IEEE Transactions on Information Theory*. Vol. 36, no. 5, September 1990. pp. 961-1005.
- [8] A. Akansu and R. Haddad. **Multiresolution Signal Decomposition: Transform, subbands, wavelet**. San Diego : Academic Press. 1992.
- [9] A. Papoulis. **Signal Analysis**. New York: McGraw-Hill. 1997.
- [10] O. Rioul and P. Duhamel. "Fast algorithms for discrete and continuous wavelet transform." *IEEE Transactions on Information Theory*. Vol.38. no. 2, March 1992. pp. 569-586.
- [11] M. Vetterli and C. Herley. "Wavelet and filter banks: Theory and design." *IEEE Transactions on Signal Processing*. Vol. 10, September 1992. pp. 1414-1429.
- [12] F. Mintzer. "Filter for distortion-free two band multirate filter banks" *IEEE Transactions on Acoustic, Speech and Signal Processing*. Vol. 33, June 1985. pp. 626-630.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [13] S. Mallat. "A theory of multiresolution signal decomposition: The wavelet representation" IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 11, July 1989. pp. 674-693.
- [14] S. Mallat. **A Wavelet Tour of Signal Processing**. San Diego: Academic Press. 1998.
- [15] Y.Y. Tang, C.Y. Suen and C.D. Yan. "Document Processing for Automatic Knowledge Acquisition." IEEE Trans. On Knowledge and Data Engineering., Vol. 6, 1994. pp. 3-21.
- [16] A. Low. **Introductory Computer Vision and Image Processing**. New York: McGraw-Hill Inc. 1991.
- [17] H.R Myler and A.R. Weeks. **Computer Imaging Recipes in C**. Prentice Hall. 1993.
- [18] R. Jain, R. Katsuri., and B.G. Schunck. **Machine Vision**. Singapore: McGraw-Hill Inc. 1995.
- [19] P. Loannis. **Digital Image Processing Algorithms**. Prentice Hall. 1993.
- [20] S.C. Hinds, J.L. Fisher and D.P. D'Amato. "A Document Skew Detection Method Using Run-Length Encoding and the Hough Transform." Proceeding, 10<sup>th</sup> International Conference on Pattern Recognition. Vol. 1, June 1990. Pp. 464-468.

## ภาคผนวก

### ผลงานวิจัยที่ได้รับการตีพิมพ์

- [1] ยุทธพงษ์ รังสรรค์เสรี, ปัญญา จูติม์ขมิมา, กฤษณะ ชินสาร และ เสกสรรค์ พลศรี. “การสร้างโมเดลและการจำแนกแบบฟอร์มสำเร็จรูป” การประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 20.
- [2] เสกสรรค์ พลศรี, เต็มพงษ์ ศรีเทศ และ ปัญญา จูติม์ขมิมา. “การค้นหาเส้นอ้างอิงสำหรับระบบป้อนข้อมูลอัตโนมัติ” การประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 37 สาขาวิศวกรรมศาสตร์.
- [2] S. Phonsri, P. Thitimajshima and Y. Rangsanseri, “Reference Line Extraction for Automated Data-Entry System.” Proceeding of 4<sup>th</sup> IEEE Malaysia International Conference on Communications and 4<sup>th</sup> IEEE Asia Pacific International on Consumer Electronics. Malaysia, November 1999.

## ประวัติผู้เขียน

นายเสกสรรค์ พลศรี เกิดเมื่อวันที่ 22 มีนาคม พ.ศ. 2516 ที่โรงพยาบาลศิริราช สำเร็จการศึกษาระดับปริญญาตรี อุดสาหกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์ จากมหาวิทยาลัยสยาม เมื่อปี พ.ศ. 2538



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้