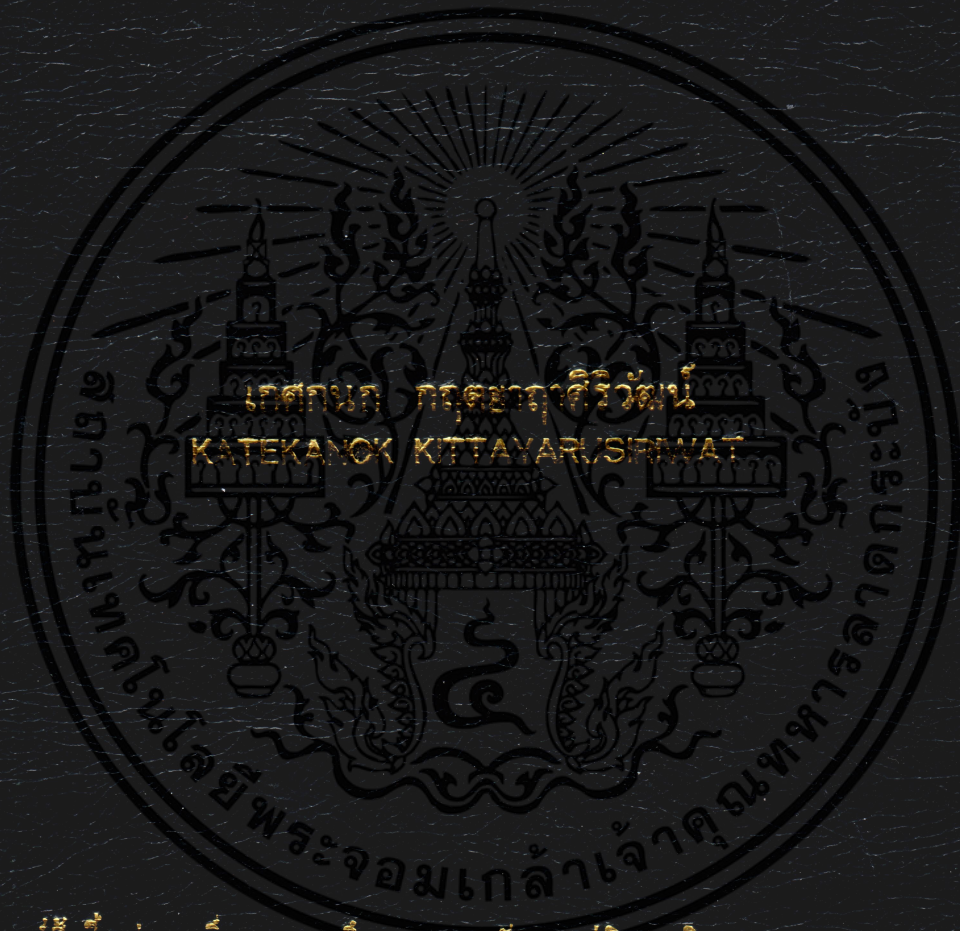


การบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์  
แบบอะแดปทีฟด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต

ADAPTIVE VECTOR QUANTIZATION CODING ON WAVELET  
INFORMATIONS FOR DATA COMPRESSION



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2548

ISBN 974-622-999-1

การบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไตซ์  
แบบอะแด็ปทีฟด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต

ADAPTIVE VECTOR QUANTIZATION CODING ON WAVELET  
INFORMATIONS FOR DATA COMPRESSION



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2543

ISBN 974-622-699-1

ดขทมี.....

ดขทเบีน 35428

ัน, เดือน, ปี 25 ๒๕๔ 2543

เพื่อการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ADAPTIVE VECTOR QUANTIZATION CODING ON WAVELET  
INFORMATIONS FOR DATA COMPRESSION



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  
2000

ISBN 974-622-699-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2000

SCHOOL OF GRADUATE STUDIES

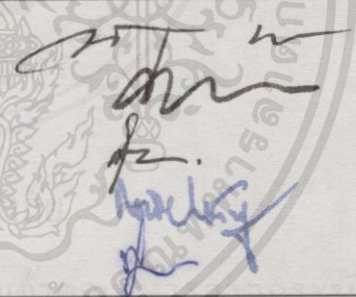
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บัณฑิตวิทยาลัย  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบอะแด็ปทีฟ  
ด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต  
ADAPTIVE VECTOR QUANTIZATION CODING ON WAVELET  
INFORMATIONS FOR DATA COMPRESSION

ชื่อนักศึกษา นางสาวเกศกนก กฤตยาฤทธิวัฒน์  
รหัสประจำตัว 40061059  
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชา วิศวกรรมไฟฟ้า  
อาจารย์ผู้ควบคุมวิทยานิพนธ์ รศ.ดร.ฟูศักดิ์ ชิวสุวิทย์

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.พิพัฒน์	เลาหสงคราม	
รศ.ศุติ	บรรจงจิตร	
ผศ.สมศักดิ์	มิตะธา	
รศ.ดร.กอบชัย	เดชหาญ	
รศ.ดร.ฟูศักดิ์	ชิวสุวิทย์	

วัน/เดือน/ปี ที่สอบ 14 มีนาคม 2543 เวลา 11.00-12.00 น.

สถานที่สอบ ณ ห้องสอบวิทยานิพนธ์ คณะวิศวกรรมศาสตร์ ดึก 12 ชั้น ชั้น 4 ห้อง (E12-404)

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.มนต์ สัจวารสิลป์)

คณบดีบัณฑิตวิทยาลัย

วันที่ 10 เดือน พฤษภาคม พ.ศ. 2543

หัวข้อวิทยานิพนธ์

การบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบ  
อะแดปทีฟด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต

นักศึกษา

นางสาวเกศกนก กฤตยาภาศิริวัฒน์

รหัสประจำตัว

40061059

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา

วิศวกรรมไฟฟ้า

พ.ศ.

2543

อาจารย์ผู้ควบคุมวิทยานิพนธ์

รศ.ดร. พุศัคดี ชิวสุวิทย์

### บทคัดย่อ

จากข้อดีของการส่งข้อมูลผ่านระบบเครือข่ายการสื่อสาร เทคนิคการแปลงเวฟเล็ตก็เป็นวิธีที่รู้จักกันดีวิธีหนึ่งที่น่ามาประยุกต์ใช้ในการบีบอัดข้อมูลเพื่อการจัดส่ง ในการแปลงเวฟเล็ต 1 ครั้ง จะได้ภาพแบนด์ย่อยที่มีข้อมูลในย่านความถี่ต่ำ 1 ส่วน และข้อมูลในย่านความถี่สูง 3 ส่วน โดยภาพแบนด์ย่อยดังกล่าวจะถูกกล่าวถึงในชื่อ LL, LH, HL และ HH สำหรับการบีบอัดข้อมูลด้วยวิธีการเดิม [1,2] นั้น ภาพแบนด์ย่อยในส่วนความถี่สูง (HH) จะถูกตัดทิ้งไป ทำให้การสร้างภาพกลับคืนจากข้อมูลความถี่ส่วน LL, LH และ HL มีการสูญเสียรายละเอียดของขอบภาพ อีกทั้งยังส่งผลให้ค่าความผิดพลาดกำลังสองเฉลี่ยมีค่าสูง เพื่อเป็นการแก้ปัญหาดังกล่าว วิทยานิพนธ์นี้จึงได้นำเสนอเทคนิคการบีบอัดข้อมูลภาพด้วยการนำค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต มาทำเวกเตอร์ควอนไทซ์แบบแบบอะแดปทีฟ ซึ่งจะช่วยลดค่าความผิดพลาดกำลังสองเฉลี่ย และยังคงรักษาค่าเฉลี่ยของจำนวนบิตต่อจุดภาพ (bit rate) ของภาพที่ถูกบีบอัดเอาไว้ได้ในขณะเดียวกัน ภาพแบนด์ย่อยที่มีความถี่ต่ำสุดจะถูกกำหนดจำนวนบิตในการเข้ารหัสต่อจุดเป็น 8 บิต ส่วนภาพแบนด์ย่อยที่เหลือ 15 ส่วน จะนำมาคำนวณหาค่าพลังงาน a.c. ด้วยการแปลงดิสครีตโคไซน์ (discrete cosine transform) แล้วนำผลที่ได้มาแบ่งเป็น 4 กลุ่ม โดยเรียงลำดับจากน้อยไปหามาก สำหรับ 2 กลุ่มแรกที่มีค่าพลังงานต่ำสุดจะถูกเข้ารหัสด้วยค่าของจำนวนบิตเป็นศูนย์ ส่วนกลุ่มที่ 3 ที่มีค่าพลังงานอยู่ในช่วง 50-75 เปอร์เซ็นต์ของพลังงานรวมทั้งหมด จะถูกเข้ารหัสด้วยเวกเตอร์รหัสขนาด 256 เวกเตอร์ และขนาดของบล็อกเวกเตอร์เป็น 4x4 พิกเซล มีอัตราบิตเป็น 0.5 บิตต่อพิกเซล และในส่วนสุดท้ายที่มีค่าพลังงาน a.c. สูงสุด จะถูกเข้ารหัสด้วยเวกเตอร์รหัสขนาด 256 เวกเตอร์ และขนาดของบล็อกเวกเตอร์เป็น 2x2 พิกเซล มีอัตราบิตเป็น 2 บิตต่อพิกเซล ในการคำนวณหาอัตราบิตรวมทั้งภาพนั้นหากได้ผลการคำนวณน้อยกว่าค่าที่กำหนด จะทำการเลื่อนภาพแบนด์ย่อยที่มีค่าพลังงาน a.c. สูงสุดในกลุ่มที่ 3 ขึ้นไปอยู่ในกลุ่มที่ 4 ดังนั้นจะถูกเข้ารหัส 2 บิตต่อพิกเซล ในทางกลับกันหากค่าอัตราบิตรวมทั้งภาพที่ได้มีค่ามากกว่าค่าที่กำหนด ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบนด์ย่อยที่มีค่าพลังงาน a.c. ต่ำสุดในกลุ่มที่ 3 จะถูกเลื่อนลงไปอยู่ในกลุ่มที่ 2 ซึ่งจะถูกเข้ารหัสด้วยอัตราบิตเป็นศูนย์ การคำนวณหาอัตราบิตรวมนี้จะทำซ้ำไปเรื่อยๆ จนกว่าจะได้อัตราบิตที่ไม่สูงกว่าค่าที่กำหนด และภาพที่ได้จากวิธีการที่นำเสนอนี้ จะสามารถลดค่าความผิดพลาดกำลังสองเฉลี่ยได้เป็นอย่างดี



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Adaptive Vector Quantization Coding on Wavelet Informations for data Compression
Student	Miss. Katekanok Kittayarusirawat
Sudent ID.	40061059
Degree	Master of Engineering
Programme	Electrical Engineering
Year	2000
Thesis Advisor	Assoc.Prof.Dr.Fusak Cheevasuvit

### ABSTRACT

Due to its advantages for image transmittly in communication network, wavelet transform is one of the well-known method for image compression. For one state of wavelet transform, The primary subband images consist of one low-pass and three high-pass. These subband images are referred as LL, LH, HL and HH. For image compression by traditional method, the high frequency subband image (HH) will be discarded. However, the reconstructed image from LL, LH, HL will be lost the detail of edges and given high mean square error. To solve this problem, this thesis proposes a method of adaptive vector quantization coding on wavelet information obtained from the two state transformation. All obtained 16 subband images were also employed in order to decrease the mean square error and preserve the compression bit rate in the same time. The lowest frequency subband image will be assigned 8 bits for encoding each pixel. Then, a.c. energy of each subband image, from the remaining 15 subband images, will be calculated by using discrete cosine transform. After that, each of them will be sorted from minimum to maximum. The accumulation of a.c. energy from 15 subband images will be divided into 4 classes. The first two lower energy classes will be encoded with zero bit. The third containing energy from 50% to 75% will be encoded with 256 code vectors for 4x4 pixels block size. This vector quantization code book provides the bit rate of 0.5 bits per pixel. Finally, with the high a.c. energy will be encoded with 256 code vectors for 2x2 pixels block size. This vector quantization code book provides the bit rate of 2 bits per pixel. Then, the total subband image will be

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

calculated the total bit rate. If the total bit rate is less than the given bit rate, the highest energy subband image of the third class will be pushed up into the fourth class. So, this subband image will be encoded with 2 bits per pixel. On the other hand, if the total bit rate is greater than the given bit rate, the lowest energy subband image of the third class will be pushed down into the second class for zero bit encoding. The bit rate adjustment process will be iteratively adapted in order to prevent the obtained bit rate greater than the given bit rate. The result image from the proposed method will be clearly improved by the mean square error (MSE).



# กิตติกรรมประกาศ

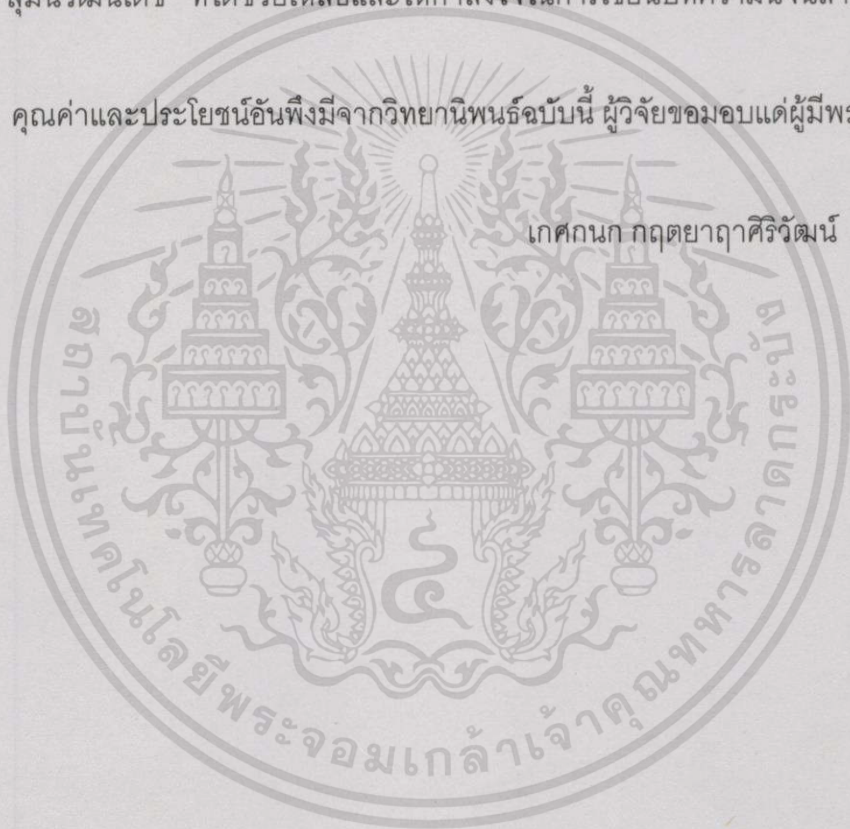
ขอกราบขอบพระคุณ รองศาสตราจารย์.ดร.ฟูศักดิ์ ชิวสุวิทย์ ที่ได้ให้คำแนะนำปรึกษา และดูแลเอาใจใส่ ในการทำวิทยานิพนธ์ฉบับนี้เป็นอย่างดี

ขอขอบคุณ ศูนย์บริการภาพถ่ายจากดาวเทียม สภาวิจัยแห่งชาติที่ได้อนุเคราะห์ข้อมูล เพื่อใช้ในการวิจัย

ขอขอบคุณอาจารย์สักรียา ชิตวงศ์, คุณศิริพงษ์ วงษ์คาร, คุณกาญจนา วงศ์ศรีทอง และ พี่ๆ น้องๆ ที่ห้องวิจัยและบริษัทกรุงเทพคอมพิวเตอร์ เซอร์วิส เซล จำกัด โดยเฉพาะอย่างยิ่งคุณ บัณฑิต สุมณวัฒน์เดช ที่ได้ช่วยเหลือและให้กำลังใจในการเขียนบทความนี้จนสำเร็จลุล่วงไปได้ ด้วยดี

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

เกศกนก กฤตยาภาศิริวัฒน์



# สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	III
กิตติกรรมประกาศ.....	V
สารบัญ.....	VI
สารบัญตาราง.....	VIII
สารบัญภาพ.....	IX
บทที่ 1 บทนำ.....	1
1.1 คำนำ.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	1
1.3 ขอบเขตของการวิจัย.....	2
บทที่ 2 หลักการเบื้องต้นในการลดข้อมูลภาพ.....	4
2.1 การลดข้อมูลแบบที่ไม่มีการสูญเสีย.....	4
2.2 การลดข้อมูลแบบที่มีการสูญเสีย.....	7
2.3 เกณฑ์การวัดประสิทธิภาพของการลดข้อมูลภาพ.....	9
บทที่ 3 การแปลงเวฟเล็ต.....	12
3.1 ทฤษฎีพื้นฐานของการแปลงเวฟเล็ต.....	12
3.2 ความเป็นมาของการแปลงเวฟเล็ต.....	17
3.3 การแปลงเวฟเล็ตแบบต่อเนื่อง.....	22
3.4 การแปลงเวฟเล็ตแบบแบ่งช่วง.....	24
บทที่ 4 การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์.....	36
4.1 การควอนไทซ์แบบเวกเตอร์สำหรับภาพเชิงตัวเลข.....	37
4.2 การหาอัตราบิตของการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์.....	40
4.3 การหาอัตราการบีบอัดข้อมูลของการลดข้อมูลภาพ ด้วยการควอนไทซ์แบบเวกเตอร์.....	40
4.4 การสร้างตัวเก็บรหัส.....	42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

หน้า

บทที่ 5 การแปลงดิสครีตโคไซน์ (DCT).....	48
5.1 คุณสมบัติและนิยามของการแปลงดิสครีตโคไซน์ .....	48
5.2 การแปลงดิสครีตโคไซน์แบบ 2 มิติ .....	54
บทที่ 6 การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ .....	57
6.1 การกำหนดจำนวนบิตแบบอะแดปทีฟ .....	60
6.2 ผลการทดลอง.....	66
6.3 สรุปผลการทดลอง .....	75
บทที่ 7 บทสรุปและแนวทางการพัฒนา.....	76
7.1 สรุปผลการวิจัย.....	76
7.2 ปัญหาและข้อเสนอแนะ.....	77
เอกสารอ้างอิง.....	78
ภาคผนวก ผลงานวิจัยที่ได้รับการตีพิมพ์.....	80
ประวัติผู้เขียน.....	81

# สารบัญภาพ

ภาพที่	หน้า
2.1 แสดงตัวอย่างการเข้ารหัสฮัฟแมน.....	5
2.2 แสดงการเข้ารหัสทำนาย.....	7
2.3 แสดงการถอดรหัสทำนาย.....	7
3.1 แสดงตัวอย่างลักษณะของ Morlet Wavelet.....	12
3.2 แสดงลักษณะของสัญญาณซายน์ซอซัดล : $\cos(t)$ .....	13
3.3 แสดงลักษณะของฟังก์ชันหน้าต่าง $e^{-t^2/64}$ .....	13
3.4 แสดงลักษณะของเวฟเล็ตแม่ที่ถูกสเกลและเลื่อนตำแหน่งไปที่ค่า a,b ต่างๆ กัน.....	14
3.5 แสดงลักษณะของการแตกกระจายสัญญาณและรวมกลับสัญญาณของเวฟเล็ต.....	15
3.6 แสดงลักษณะของสัญญาณที่มีหลายความถี่ที่ตำแหน่งเวลาต่างกัน.....	16
3.7 แสดงลักษณะของผลการ Correlation ระหว่าง $f(t), g((t-b)/a)$ .....	17
3.8 แสดงลักษณะของระนาบเวลาความถี่แบบคงที่.....	19
3.9 แสดงลักษณะของระนาบเวลาและความถี่สำหรับการแปลงเวฟเล็ต.....	20
3.10 แสดงลักษณะการแตกกระจายสัญญาณไปยังระดับความละเอียดต่างๆ.....	20
3.11 แสดงลักษณะของฟังก์ชันเวฟเล็ตแม่ชนิดต่างๆ.....	21
3.12 แสดงการแปลงเวฟเล็ตในลักษณะของการ Inner product.....	24
3.13 แสดงลักษณะของการแจกแจงสัมประสิทธิ์ด้วยการแปลงเวฟเล็ต.....	25
3.14 แสดงสัญลักษณ์ของฟิลเตอร์เบงค์แบบ 2 ช่องสัญญาณ หรือ Bands separator.....	26
3.15 แสดงลักษณะโครงสร้างต้นไม้ในการวิเคราะห์หรือออกเทพของฟิลเตอร์เบงค์.....	26
3.16 แสดงลักษณะของการแยกแบนด์ความถี่ของออกเทพฟิลเตอร์เบงค์.....	27
3.17 แสดงฟิลเตอร์เบงค์แบบ 2 ช่องสัญญาณในการคำนวณหาสัมประสิทธิ์เวฟเล็ต.....	29
3.18 แสดงลักษณะของ DWT โดยใช้ Dyadic tree structure (Octave filter bank).....	30
3.19 แสดงลักษณะของ IWT โดยการใช้ Dyadic tree structure.....	30
3.20 แสดงแผนภาพ Image Wavelet Decomposition.....	32
3.21 แสดงแผนภาพ Image Wavelet Reconstruction.....	32
3.22 แสดงภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ตหลายครั้ง.....	33
3.23 แสดงภาพต้นแบบภาพถ่ายจากดาวเทียม LANDSAT ขนาด 512x512x8 บิต.....	34
3.24 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตระดับที่ 1.....	34

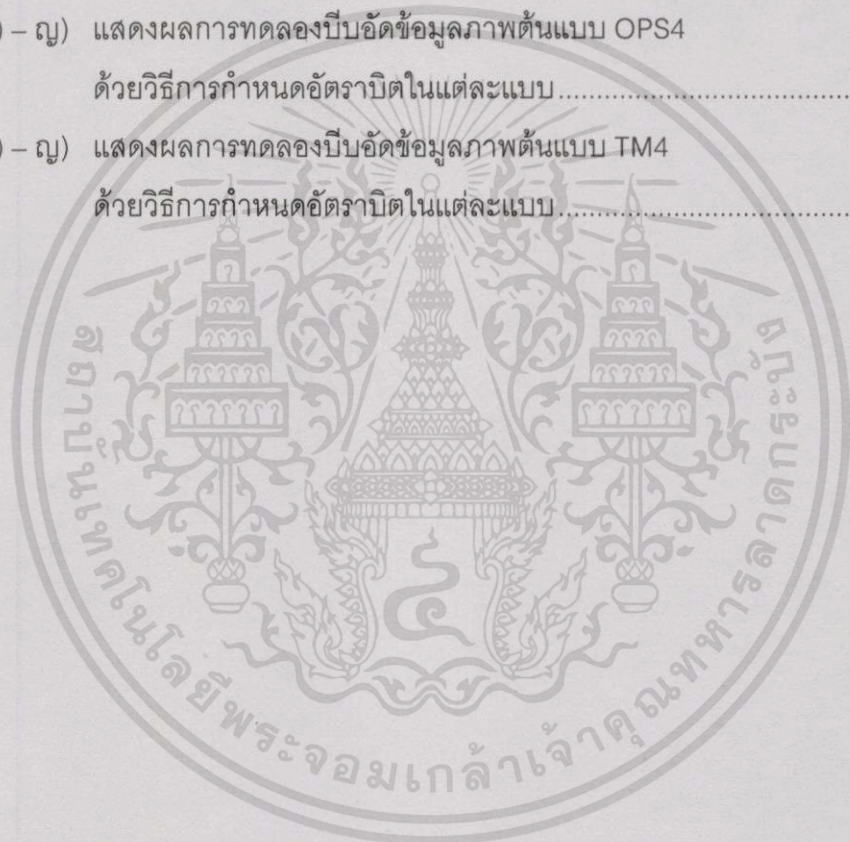
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญภาพ (ต่อ)

ภาพที่	หน้า
3.25 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ดระดับที่ 2	34
3.26 แสดงภาพถ่ายต้นแบบ Barbara ขนาด 512x512x8 บิต	35
3.27 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ดระดับที่ 1	35
3.28 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ดระดับที่ 2	35
4.1 แสดงบล็อกไดอะแกรมของการควอนไทซ์แบบเวกเตอร์ทั้งส่วนเข้ารหัสและส่วนถอดรหัส	36
4.2 แสดงการแบ่งภาพเป็นบล็อกย่อยขนาด 4x4	38
4.3 แสดงตัวอย่างการจัดเรียงจุดภาพของบล็อกย่อยขนาดต่างๆ ให้อยู่ในรูปของเวกเตอร์	38
4.4 แสดงส่วนเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์	38
4.5 แสดงโครงสร้างของตัวเก็บรหัส	39
4.6 แสดงส่วนถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์	40
4.7 แสดงไฟล์ชาร์ตของอัลกอริทึม LBG	43
4.8 แสดงการแยกตัวเก็บรหัสเป็น 2 เทา	47
5.1 แสดงส่วนประกอบของผลลัพธ์ที่ได้จากการแปลงดิสครีตโคไซน์ 2 มิติ	55
5.2 ก. แสดงข้อมูลขนาด 8*8	55
5.3 ข. แสดงผลลัพธ์ที่ได้จากการแปลงดิสครีตโคไซน์	55
6.1 แสดงแผนภาพการบีบอัดข้อมูลภาพร่วมกันระหว่างการควอนไทซ์แบบเวกเตอร์และการแปลงเวฟเล็ด	57
6.2 แสดงการนำข้อมูลกลับจากการบีบอัดข้อมูลภาพร่วมกันระหว่างการควอนไทซ์แบบเวกเตอร์และการแปลงเวฟเล็ด	58
6.3 แสดงผลการแปลงเวฟเล็ดระดับที่ 1 และ 2	58
6.4 วิธีการกำหนดอัตราบิตแบบทั่วไปที่อัตราบิตรวม = 1.03125 bpp	58
6.5 วิธีการกำหนดอัตราบิตแบบละเอียดที่อัตราบิตรวม = 1.03125 bpp	59
6.6 แสดงส่วนเข้ารหัสด้วยการควอนไทซ์แบบเวกเตอร์กับการแปลงเวฟเล็ด	60
6.7 แผนภาพแสดงการทำงานของวิธีการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ	61
6.8 แสดงภาพการแปลงเวฟเล็ด 2 ระดับ	62
6.9 แสดงภาพการตรวจสอบหา AC Energy ด้วยการแปลง DCT	63

## สารบัญญภาพ (ต่อ)

ภาพที่	หน้า
6.10 แสดงการจัดแบ่งกลุ่มของแต่ละบล็อกย่อยและเปอร์เซ็นต์ของ AC Energy ในแต่ละกลุ่ม .....	64
6.11 แสดงผังงานในการปรับอัตราบิดรวมให้ได้อัตราบิดตามที่ต้องการ .....	65
6.12 ก) – ญ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ LENA ด้วยวิธีการกำหนดอัตราบิดในแต่ละแบบ .....	69
6.13 ก) – ญ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ OPS4 ด้วยวิธีการกำหนดอัตราบิดในแต่ละแบบ .....	72
6.14 ก) – ญ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ TM4 ด้วยวิธีการกำหนดอัตราบิดในแต่ละแบบ .....	75



# สารบัญตาราง

ตารางที่	หน้า
4.1 แสดงอัตราบิดต่อจุดภาพของการบีบอัดข้อมูลด้วยการควอนไทซ์แบบเวกเตอร์ที่มีตัวเก็บรหัสและบล็อกย่อยขนาดต่างๆ.....	41
4.2 แสดงอัตราการบีบอัดข้อมูลของการลดข้อมูลด้วยการควอนไทซ์แบบเวกเตอร์ที่มีตัวเก็บรหัสและบล็อกย่อยขนาดต่างๆ.....	42
6.1 แสดงการกำหนดอัตราบิด ขนาดเวกเตอร์ และขนาดบล็อกของแต่ละเวกเตอร์.....	64
6.2 เปรียบเทียบค่า MSE และค่า PSNR จากผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิดของแต่ละวิธี .....	75



# บทที่ 1

## บทนำ

### 1.1 คำนำ

เนื่องจากในปัจจุบันการส่งข้อมูลผ่านระบบเครือข่ายการสื่อสารได้รับความนิยมค่อนข้างสูง ทั้งนี้อาจเนื่องมาจากเทคโนโลยีด้านไมโครอิเล็กทรอนิกส์มีการพัฒนาไปมาก ทำให้เครื่องคอมพิวเตอร์มีประสิทธิภาพการทำงานที่ดีขึ้น เช่น การแสดงภาพเชิงตัวเลขที่ใช้หน่วยความจำขนาดใหญ่ เป็นต้นว่าภาพถ่ายดาวเทียมหรือภาพถ่ายเอกซเรย์ขนาดใหญ่ ก็สามารถที่จะกระทำได้อย่างรวดเร็วและมีประสิทธิภาพมากขึ้น แต่การส่งผ่านข้อมูลผ่านระบบเครือข่ายการสื่อสารด้วยข้อมูลจำนวนมากดังกล่าว จำเป็นต้องมีเทคนิคมาช่วยในเรื่องของการบีบอัดข้อมูลเพื่อให้การจัดส่งเป็นไปอย่างถูกต้อง ลดเวลาและค่าใช้จ่าย โดยเฉพาะอย่างยิ่งถ้าเป็นการส่งข้อมูลผ่านเครือข่ายสาธารณะ เช่น ผ่านทางสายโทรศัพท์

ดังนั้นในวิทยานิพนธ์ฉบับนี้ จึงได้เสนอเทคนิคการบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบอะแดปทีฟเข้ากับค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต ซึ่งจะเป็นวิธีการบีบอัดข้อมูลที่มีประสิทธิภาพสูง สามารถนำไปประยุกต์ใช้ในการบีบอัดข้อมูลที่มีการกระจายของข้อมูลได้อย่างหลากหลาย โดยพิจารณาตามค่าพลังงานความหนาแน่นของปริมาณข่าวสาร นอกจากนี้ยังสามารถแปลงกลับจากข้อมูลที่ถูกบีบอัดแล้วให้กลับเป็นข้อมูลขนาดเท่าเดิมได้อย่างรวดเร็วอีกด้วย

### 1.2 วัตถุประสงค์ของการวิจัย

เป็นการนำเสนอเทคนิคการบีบอัดข้อมูลภาพด้วยเทคนิคการแปลงเวกเตอร์ควอนไทซ์แบบอะแดปทีฟให้กับค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต โดยจะเป็นการศึกษาถึงการเข้ารหัสภาพแบนด์ย่อยต่างๆ ด้วยจำนวนบิตที่แตกต่างกันตามค่าพลังงาน a.c. ซึ่งเป็นตัวบอกปริมาณข่าวสารของภาพในแต่ละแบนด์ย่อย โดยมีการวัดประสิทธิภาพของการบีบอัดข้อมูลภาพด้วยค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error, MSE) และค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak-Signal-to-Noise Ratio, PSNR)

### 1.3 ขอบเขตของการวิจัย

ในวิทยานิพนธ์ฉบับนี้ ได้นำเสนอเทคนิคการบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบอะแดปทีฟให้กับค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ต โดยเนื้อหาและรายละเอียดของการวิจัย แยกออกได้เป็น 7 บท ดังต่อไปนี้

#### บทที่ 1 บทนำ

กล่าวถึงความเป็นมาและความสำคัญของงานวิจัย วัตถุประสงค์ของการทำวิจัย และเนื้อหาของวิทยานิพนธ์

#### บทที่ 2 หลักการเบื้องต้นในการลดข้อมูลภาพ

กล่าวถึงทฤษฎีที่เกี่ยวข้องและหลักการพื้นฐานของการลดข้อมูลภาพ เพื่อให้เกิดความเข้าใจขั้นตอนต่างๆ ของการลดข้อมูลภาพ โดยจะอธิบายวิธีการต่างๆ ในการลดข้อมูลภาพ ทั้งการลดข้อมูลภาพแบบที่มีการสูญเสีย และการลดข้อมูลภาพแบบที่ไม่มีการสูญเสีย วิธีการแปลงเวฟเล็ตของภาพเชิงตัวเลข และกล่าวถึงหลักเกณฑ์ต่างๆ ที่ใช้ในการวัดประสิทธิภาพของการลดข้อมูลภาพ

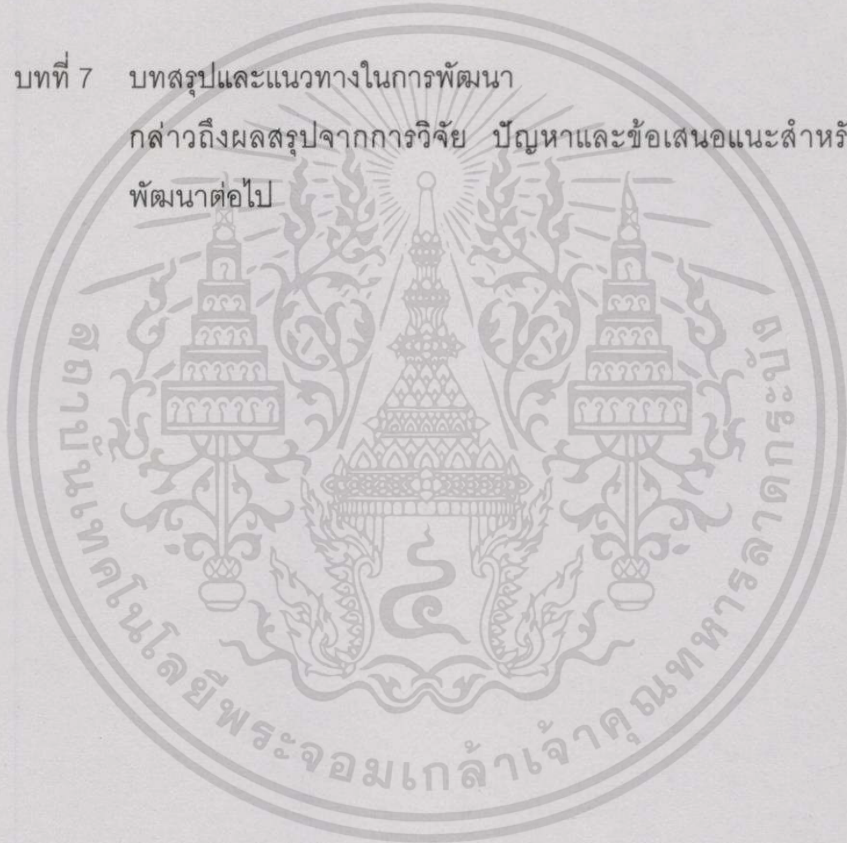
#### บทที่ 3 การแปลงเวฟเล็ต

กล่าวถึงทฤษฎีพื้นฐานของการแปลงเวฟเล็ต ความเป็นมาของการแปลงเวฟเล็ต รายละเอียดของการแปลงเวฟเล็ตแบบต่อเนื่อง และการแปลงเวฟเล็ตแบบแบ่งช่วง

#### บทที่ 4 การลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์

กล่าวถึงการลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ โดยอธิบายถึงความหมายและพื้นฐานของการลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ การสร้างตัวเก็บรหัส การหาอัตราบิตของการเข้ารหัส รวมทั้งการหาอัตราการบีบอัดข้อมูลภาพ และการถอดรหัสภาพเชิงตัวเลขด้วยการควอนไทซ์แบบเวกเตอร์

- บทที่ 5 การแปลงดิสครีตโคไซน์ (DCT)  
กล่าวถึงคุณสมบัติและนิยามของการแปลงดิสครีตโคไซน์ การนำการแปลงดิสครีตโคไซน์มาประยุกต์ใช้ รวมทั้งประสิทธิภาพของวิธีการดังกล่าว
- บทที่ 6 การลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ  
กล่าวถึงการทดลองและผลการทดลอง ที่ได้จากการลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ โดยได้ทำการกำหนดจำนวนบิตแบบอะแดปทีฟ เพื่อนำมาสร้างตัวเก็บรหัสขนาดต่างๆ
- บทที่ 7 บทสรุปและแนวทางในการพัฒนา  
กล่าวถึงผลสรุปจากการวิจัย ปัญหาและข้อเสนอแนะสำหรับแนวทางการพัฒนาต่อไป



## บทที่ 2

# หลักการเบื้องต้นในการลดข้อมูลภาพ

การลดข้อมูลภาพเป็นกระบวนการลดจำนวนของบิตที่จะใช้ในการแสดงภาพ การลดข้อมูลภาพจะคำนึงถึงข้อมูลทางสถิติของภาพ ร่วมกับความสัมพันธ์ที่ยอมรับได้ด้วยการสังเกตด้วยตาของมนุษย์ โดยการลดจำนวนของบิตที่ต้องการใช้ในการแสดงภาพ ทำให้การลดข้อมูลภาพสามารถนำไปประยุกต์ใช้ในงานอื่นๆ ได้อย่างมีประสิทธิภาพ การจัดเก็บข้อมูลสามารถพัฒนาให้ดีขึ้นได้ด้วยการใช้การลดข้อมูลเพื่อทำการลดความต้องการหน่วยความจำ การส่งภาพสามารถพัฒนาให้ดีขึ้นได้ด้วยการลดความต้องการความกว้างแถบ (Bandwidth) หรือการเพิ่มความเร็วในการส่งให้กับความกว้างแถบที่ใช้อยู่

เทคนิคในการลดข้อมูลภาพสามารถแบ่งออกได้เป็น 2 กลุ่มคือ การลดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless compression) และการลดที่มีการสูญเสีย (Lossy compression) การลดข้อมูลแบบไม่มีการสูญเสียจะสามารถนำข้อมูลกลับมาได้ใหม่อย่างสมบูรณ์ ซึ่งมีชื่อเรียกต่างๆ กัน เช่น การเข้ารหัสแบบไม่มีสัญญาณรบกวน (noiseless), การเข้ารหัสที่สามารถสร้างกลับมาได้ (invertible), การเข้ารหัสเอนโทรปี (entropy coding), การอัดข้อมูล (data compaction) เป็นต้น การลดข้อมูลโดยปราศจากการสูญเสียนี้มีความจำเป็นสำหรับข้อมูลบางชนิด เช่น ข้อมูลทางการแพทย์ และข้อมูลโปรแกรมคอมพิวเตอร์ การลดข้อมูลแบบที่ไม่มีการสูญเสียจะให้อัตราการบีบอัดต่ำ ดังนั้นเพื่อให้ได้อัตราการลดข้อมูลที่สูงขึ้น จึงจำเป็นที่จะต้องใช้การลดข้อมูลแบบที่มีการสูญเสีย การลดข้อมูลแบบที่มีการสูญเสียนี้จะทำให้เกิดความเพี้ยน (distortion) กับข้อมูลที่ได้ถูกสร้างกลับขึ้นมาใหม่ หัวใจในการออกแบบระบบการลดข้อมูลแบบที่มีการสูญเสียคือ การทำให้มีค่าความเพี้ยนเฉลี่ยต่ำที่สุดสำหรับอัตราบิต (bit rate) หรืออัตราการลดข้อมูล (compression ratio) ซึ่งวิธีการลดข้อมูลในแบบต่างๆ สามารถที่จะนำมารวมกันได้ เรียกว่าเป็นการลดข้อมูลแบบผสมผสาน (hybrid compression)

### 2.1 การลดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression)

การลดข้อมูลแบบที่ไม่มีการสูญเสียนำเอาประโยชน์ของส่วนเกินของข้อมูล (redundance) มาใช้ ซึ่งวิธีการหนึ่งที่จะมาใช้วัดส่วนเกินของข้อมูลคือ การหาความน่าจะเป็นของการกระจายของค่าความเข้มของจุดภาพ โดยที่เอนโทรปีอันดับที่หนึ่ง (first order entropy) สามารถหาได้จาก

$$H = -\sum_i p(i) \log_2(p(i)) \quad (\text{bits}) \quad (2.1)$$

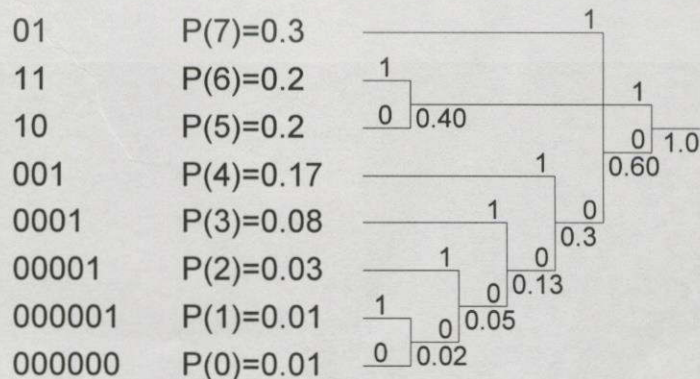
เมื่อ  $p(i)$  คือความน่าจะเป็นของค่าความเข้ม  $i$  ของจุดภาพที่เกิดขึ้น

เอนโทรปีจะใช้อธิบายความไม่แน่นอนหรือ Randomness ของข้อมูล และจะแสดงค่าเฉลี่ยของจำนวนบิตที่จะใช้แทนข้อมูล ภาพจะถูกลดข้อมูลโดยการเข้ารหัสค่าความเข้มของจุดภาพที่เกิดขึ้นบ่อยด้วยจำนวนบิตที่น้อย และเข้ารหัสค่าความเข้มของจุดภาพที่เกิดขึ้นน้อยกว่าด้วยจำนวนบิตที่มาก ซึ่งโดยเฉลี่ยแล้วจะเป็นการลดจำนวนบิตที่ใช้แทนข้อมูล ตัวอย่างการลดข้อมูลแบบที่ไม่มีการสูญเสียวิธีต่างๆ

### 2.1.1 การเข้ารหัสฮัฟแมน (Huffman Coding)

การเข้ารหัสฮัฟแมนเป็นการลดข้อมูลแบบที่ไม่มีการสูญเสียวิธีหนึ่ง หลักการที่สำคัญของการเข้ารหัสฮัฟแมนคือ ข้อมูลที่มีความน่าจะเป็นสูงกว่าจะถูกเข้ารหัสด้วยจำนวนบิตที่น้อยกว่า การเข้ารหัสฮัฟแมน [1] สามารถอธิบายโดยย่อได้ดังนี้

- 1) ทำการนับความถี่เพื่อหาความน่าจะเป็นของข้อมูลทุกตัว และเรียงลำดับข้อมูลตามความน่าจะเป็นจากน้อยไปหามาก
  - 2) นำความน่าจะเป็นที่มีค่าต่ำสุด 2 ค่ามารวมกันเพื่อสร้างเป็นโหนด (node) ใหม่ แล้วนำความน่าจะเป็นใหม่ที่ได้มาเรียงลำดับใหม่ ร่วมกับความน่าจะเป็นเดิมที่เหลือ
  - 3) กำหนดค่าไบนารีให้กับโหนดคู่ที่ถูกกำหนดให้เป็นเส้นทางของโหนดใหม่นั้น โดยโหนดหนึ่งมีค่าเป็น 0 ส่วนอีกโหนดหนึ่งมีค่าเป็น 1
  - 4) ทำซ้ำตามขั้นตอนที่ 2 และ 3 จนกระทั่งได้ความน่าจะเป็นสุดท้ายเป็น 1.0
- ตัวอย่างของการสร้างรหัสฮัฟแมน แสดงได้ดังภาพที่ 2.1



ภาพที่ 2.1 แสดงตัวอย่างการเข้ารหัสฮัฟแมน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากได้รับรหัสฮัฟแมนแล้ว การเข้ารหัสและการถอดรหัสสามารถทำได้โดยใช้การเปิดตาราง (Look-up table) และรหัสที่ได้ก็สามารถที่จะทำการถอดรหัสได้ทันที โดยไม่ต้องอ้างอิงถึงข้อมูลตัวอื่นๆ

### 2.1.2 การเข้ารหัสเลนจ์ (Run-Length Coding)

เป็นวิธีการที่ง่ายที่สุดในการเข้ารหัสข้อมูลไบนารีที่มักจะมีกลุ่มของ 0 และ 1 เช่น ในข้อมูลภาพหรือข้อความสำหรับการส่งแฟกซ์ (facsimile) หรืออาจใช้กับภาพที่ภาพที่อยู่ใกล้กันมีระดับการเปลี่ยนแปลงระดับความเข้มไม่มาก และมีการซ้ำกันเป็นจำนวนมาก โดยจะเก็บจำนวนของจุดภาพที่ซ้ำกัน แทนที่จะเก็บค่าระดับความเข้มของจุดภาพเหล่านั้น [2]

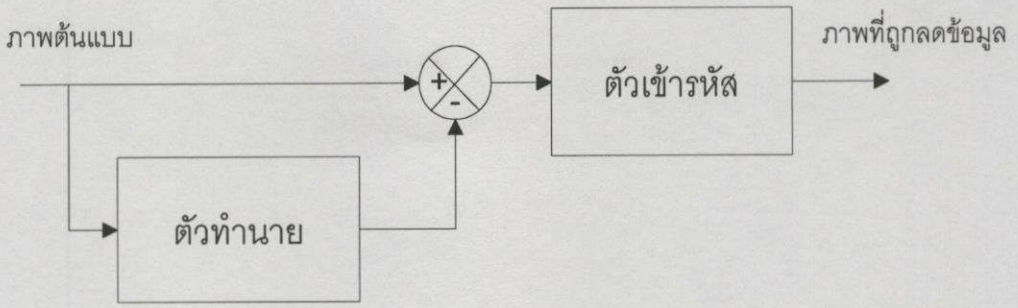
### 2.1.3 การเข้ารหัสตามระดับของบิต (Bit Plane Coding)

สมมติว่าแต่ละค่าของจุดภาพในภาพระดับเทาภาพหนึ่งถูกแทนด้วยข้อมูล  $k$  บิต ดังนั้นเราสามารถสร้างภาพไบนารีจำนวน  $k$  ภาพ ซึ่งบิตแต่ละตำแหน่งจะสร้างได้ 1 ภาพ โดยภาพที่ได้มาจากบิตตำแหน่งที่มีนัยสำคัญสูงกว่า จะประกอบด้วยข่าวสารที่สำคัญกว่า และจะมี spatial redundancy ที่สูงกว่า ดังนั้นจึงทำให้ง่ายต่อการลดข้อมูล วิธีการเข้ารหัสตามระดับของบิตจะอาศัยค่าระดับความเทาของภาพในการเพิ่มความสัมพันธ์ (correlation) ของจุดภาพ จากนั้นจึงนำการเข้ารหัสแบบรันเลนจ์ หรือการเข้ารหัสเลขคณิตมาใช้ในการลดข้อมูล ซึ่งรายละเอียดสามารถศึกษาได้จาก [3,4]

### 2.1.4 การเข้ารหัสทำนาย (Lossless Predictive Coding)

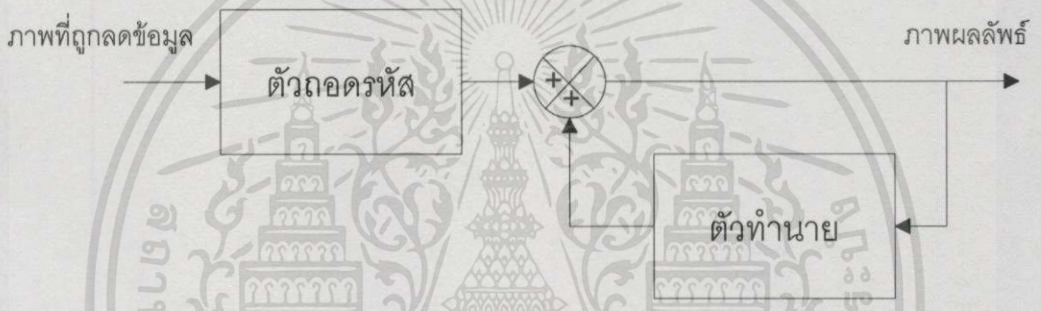
โดยปกติสำหรับภาพทั่วไปแล้ว จุดภาพที่อยู่ใกล้เคียงกันจะมีความสัมพันธ์ต่อกันสูง นั่นหมายความว่าข้อมูลจุดภาพสามารถที่จะหาได้โดยการตรวจสอบจุดภาพที่อยู่บริเวณรอบๆ ของจุดภาพที่เราสนใจ ด้วยคุณสมบัตินี้เองที่ได้นำมาใช้ในการเข้ารหัสแบบทำนายโดยพยายามที่จะทำการคำนวณหาค่าของจุดภาพจากจุดภาพที่อยู่รอบๆ [3,4]

การเข้ารหัสแบบทำนายจะทำการลดข้อมูลส่วนเกินระหว่างจุดภาพ (interpixel redundancy) แล้วเข้ารหัสเฉพาะข่าวสารใหม่ของแต่ละจุดภาพ โดยข่าวสารใหม่ของจุดภาพนี้สามารถหาได้จากความแตกต่างระหว่างค่าจริงของจุดภาพกับค่าที่ได้มาจากการทำนาย แสดงได้ดังภาพที่ 2.2



ภาพที่ 2.2 แสดงการเข้ารหัสทำนาย

ส่วนการถอดรหัสนั้นจะนำภาพที่ถูกลดข้อมูลแล้วมาบวกกับค่าของตัวทำนายที่มีฟังก์ชันของตัวทำนายเช่นเดียวกับกับตัวเข้ารหัส ซึ่งผลที่ได้นั้นจะเรียกว่าภาพผลลัพธ์ ดังภาพที่ 2.3



ภาพที่ 2.3 แสดงการถอดรหัสทำนาย

## 2.2 การลดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression)

การลดข้อมูลแบบที่มีการสูญเสียจะทำให้เกิดความเพี้ยนของข้อมูลขึ้น แต่ก็มีอัตราการลดข้อมูลที่สูงกว่าการลดข้อมูลแบบที่ไม่มีการสูญเสียมาก ซึ่งมีอยู่หลายวิธีการ ดังตัวอย่างต่อไปนี้

### 2.2.1 การเข้ารหัสแบบทำนาย (Lossy Predictive Coding)

ในการเข้ารหัสแบบทำนายโดยทั่วไป ความสัมพันธ์ระหว่างจุดภาพที่อยู่ใกล้เคียงกัน จะถูกนำมาใช้ในการทำนายค่าของแต่ละจุดภาพ วิธีหนึ่งในการลดข้อมูลที่น่าเอาการเข้ารหัสแบบทำนายมาใช้คือ การมอดูเลตแบบรหัสพัลส์เชิงอนุพันธ์ (Differential Pulse Code Modulation, DPCM) โดยจะนำเอาค่าความผิดพลาดจากการทำนายที่ได้มาจากการลบค่าการทำนายออกจากค่าจริงของจุดภาพมาทำการควอนไทซ์แบบสเกลาร์ซึ่งจะทำให้เกิดความผิดเพี้ยนของข้อมูลขึ้น แต่ก็สามารถลดข้อมูลได้ในอัตราที่มากขึ้น

## 2.2.2 การเข้ารหัสโดยการแปลง (Transform Coding)

การลดข้อมูลด้วยการเข้ารหัสการแปลง จะทำการแปลงข้อมูลอินพุตที่อยู่ในรูปของ Spatial domain ให้อยู่ในรูปของ spectral หรือ frequency domain โดยใช้วิธีการแปลงแบบต่างๆ เช่น การแปลงฟูเรียร์ (Fourier Transform) ในการแปลงข้อมูลที่อยู่ในรูปของ spatial domain ให้อยู่ในรูปสัมประสิทธิ์ของพลังงานความถี่ โดยที่ความถี่ต่ำๆ จะมีพลังงานสูง และที่ความถี่สูงๆ พลังงานจะลดลงไป สำหรับข้อมูลภาพโดยทั่วๆ ไปนั้นจะประกอบด้วยส่วนที่เป็นความถี่สูง คือส่วนที่เป็นรายละเอียดหรือขอบภายในภาพ และส่วนที่เป็นความถี่ต่ำของภาพ คือข้อมูลส่วนที่เป็นพื้นของภาพ ซึ่งจำนวนบิตที่ใช้ในการเข้ารหัสของแต่ละช่วงความถี่จะไม่เท่ากัน ดังนั้นเมื่อต้องการลดข้อมูลให้ได้มาก ค่าของพลังงานความถี่สูงจะถูกตัดทิ้งไปเป็นส่วนใหญ่ ทำให้รายละเอียดส่วนที่เป็นขอบในภาพขาดหายไป ภาพที่ได้จะเบลอ ขาดความคมชัด

การแปลงที่ใช้ในการลดข้อมูลภาพมีอยู่หลายวิธี เช่น การแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform), การแปลงวอลซ์-ฮาดามาร์ด (Walsh-Hadamard Transform), การแปลงดิสครีตโคไซน์ (Discrete Cosine Transform), การแปลงฮาร์ (Haar Transform) เป็นต้น ซึ่งแต่ละวิธีจะมีข้อดีและข้อเสียแตกต่างกันไป แต่วิธีที่นิยมใช้ในการลดข้อมูลภาพมากที่สุดคือ การแปลงดิสครีตโคไซน์ (DCT) [3,4,12] เนื่องจากเป็นวิธีที่สามารถคำนวณได้ง่ายและเร็ว ตัวอย่างของการลดข้อมูลที่ใช้การแปลงดิสครีตโคไซน์คือ การลดข้อมูลแบบ JPEG (Joint Photographics Expert Group)

## 2.2.3 การเข้ารหัสแบนด์ย่อย (Subband Coding)

ในการเข้ารหัสแบนด์ย่อย [13,14] ภาพต้นแบบจะถูกกรองตามช่วงของความถี่เพื่อสร้างเป็นภาพย่อย ซึ่งภาพเหล่านี้ถูกเรียกว่าภาพแบนด์ย่อย (subbands) เนื่องจากแต่ละภาพแบนด์ย่อยนี้ได้ถูกลดความกว้างแถบ (bandwidth) เมื่อเปรียบเทียบกับภาพต้นแบบ ทำให้สามารถที่จะทำการลดจำนวนการสุ่ม (downsampling) ลงได้ ซึ่งสามารถทำได้ด้วยกระบวนการในการกรอง (filtering) ส่วนการลดจำนวนการสุ่มนี้เรียกว่าเป็นขั้นตอนการวิเคราะห์ (analysis stage) หลังจากนั้นภาพแบนด์ย่อยที่ได้จะถูกเข้ารหัสด้วยตัวเข้ารหัสต่างๆ

ในการสร้างภาพกลับคืนมาอีกครั้งสามารถทำได้โดยการเพิ่มจำนวนการสุ่ม (upsampling) ของภาพแบนด์ย่อยที่ได้ถอดรหัสแล้ว จากนั้นนำมาผ่านตัวกรองแล้วรวมภาพแต่ละแบนด์ย่อยเข้าด้วยกันเพื่อสร้างเป็นภาพกลับคืนมา กระบวนการดังกล่าวเรียกว่า ขั้นตอนการสังเคราะห์ (synthesis stage)

จะสังเกตได้ว่าการสร้างภาพแบนด์ย่อยจะไม่ทำให้เกิดการลดลงของข้อมูล เนื่องจากจำนวนของการสุ่ม (Sample) ของทุกแบนด์ย่อยรวมกันแล้วยังคงเท่ากับภาพต้นแบบ แต่จะมีข้อดีในแง่ที่แต่ละภาพแบนด์ย่อยนั้นสามารถที่จะทำการเข้ารหัสได้ด้วยประสิทธิภาพที่ดีกว่าภาพต้นแบบเดิม

## 2.3 เกณฑ์การวัดประสิทธิภาพของการลดข้อมูลภาพ

ในการวัดประสิทธิภาพของการลดข้อมูลในเชิงคณิตศาสตร์ สามารถกระทำได้หลายวิธีคือ

### 2.3.1 ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error, MSE)

เราสามารถใช้ค่าผิดพลาดกำลังสองเฉลี่ยในการวัดประสิทธิภาพของการลดข้อมูล คือ ถ้าความผิดพลาดกำลังสองเฉลี่ยมีค่าน้อย แสดงว่าภาพที่ได้จากการแปลงกับภาพที่ถูกลดข้อมูลแล้วมีความผิดเพี้ยนไปจากภาพต้นแบบน้อย ซึ่งเราสามารถหาค่าความผิดพลาดกำลังสองเฉลี่ยของการลดข้อมูลภาพขนาด  $M \times N$  ได้ดังนี้

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2 \quad (2.2)$$

โดยที่  $M$  คือจำนวนจุดภาพตามความกว้างของภาพ  
 $N$  คือจำนวนจุดภาพตามความสูงของภาพ  
 $f(i, j)$  คือระดับความเทา (Gray level) ที่ตำแหน่ง  $(i, j)$  ของภาพต้นแบบ  
 $\hat{f}(i, j)$  คือระดับความเทาที่ตำแหน่ง  $(i, j)$  ของภาพที่ได้ถูกสร้างกลับขึ้นมาใหม่

ในบางครั้งเราอาจจะพบว่ามีการใช้ค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ย (Root mean square error, RMSE) ในการวัดประสิทธิภาพการลดข้อมูลแทนค่า  $MSE$  ซึ่งค่า  $RMSE$  สามารถหาได้จาก

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2} \quad (2.3)$$

### 2.3.2 SNR และ PSNR

อีกวิธีการหนึ่งของการอธิบายประสิทธิภาพของอัลกอริทึมคือ ค่าสัดส่วนต่อสัญญาณรบกวน (Signal-to-Noise Ratio, SNR) ในบางครั้งก็เรียกว่า ค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนการควอนไทซ์ (Signal-to-Quantization-Noise Ratio) ซึ่งในการวัดค่าจะมีความสัมพันธ์ของกำลังของสัญญาณกับความเพี้ยนของสัญญาณ สามารถหาได้จากการนำค่าเฉลี่ยของภาพต้นแบบยกกำลังสองแล้วหารด้วย  $MSE$  โดยมีหน่วยเป็นเดซิเบล (dB) ดังนี้

$$SNR = 10 \log_{10} \frac{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j)]^2}{MSE} \quad (\text{dB}) \quad (2.4)$$

$$SNR = 10 \log_{10} \frac{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j)]^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2} \quad (\text{dB}) \quad (2.5)$$

การวัดที่คล้ายๆ กันคือ ค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak-Signal-to-Noise Ratio, PSNR) โดยสามารถหาได้ดังนี้

$$PSNR = 10 \log_{10} \frac{Peak^2}{MSE} \quad (\text{dB}) \quad (2.6)$$

เมื่อ Peak แทนค่าสูงสุดของขอบเขตของข้อมูล ทั้ง SNR และ PSNR เป็นวิธีที่นิยมใช้อธิบายคุณภาพของการลดข้อมูลภาพ

สำหรับภาพที่มีจุดภาพแทนด้วยข้อมูลขนาด  $n$  บิต หรืออาจกล่าวได้ว่ามีจำนวนระดับความเทา (gray scale) เป็น  $2^n$  จะได้

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \quad (\text{dB}) \quad (2.7)$$

ดังนั้นภาพที่มีจุดภาพแทนด้วยข้อมูลขนาด 8 บิต หรือมีจำนวนระดับความเทาเป็น 256 จะได้

$$PSNR = 10 \log_{10} \frac{(2^8 - 1)^2}{MSE} \quad (2.8)$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (2.9)$$

$$PSNR = 20 \log_{10} \frac{255}{MSE} \quad (2.10)$$

### 2.3.3 อัตราบิต

อัตราบิต (bit rate) คือค่าเฉลี่ยของจำนวนบิตต่อจุดภาพ (bits per pixel, BPP) ของภาพที่ถูกลดข้อมูล โดยสามารถคำนวณได้จากอัตราส่วนระหว่างจำนวนบิตข้อมูลทั้งหมดของภาพที่ถูกลดข้อมูล ต่อจำนวนจุดภาพทั้งหมดของภาพต้นแบบ ดังสมการต่อไปนี้

$$\text{Bit Rate (bits per pixel, BPP)} = \frac{\text{จำนวนบิตทั้งหมดของภาพที่ถูกลดข้อมูล}}{\text{จำนวนจุดภาพทั้งหมดของภาพต้นแบบ}} \quad (2.11)$$

### 2.3.4 อัตราการลดข้อมูล (Compression ratio)

อัตราการลดข้อมูล คืออัตราส่วนระหว่างจำนวนบิตข้อมูลที่ใช้แทนภาพต้นแบบ และจำนวนบิตข้อมูลที่ใช้แทนภาพที่ถูกทำการลดข้อมูล โดยสามารถหาได้ดังสมการต่อไปนี้

$$\text{อัตราการลดข้อมูล} = \frac{\text{จำนวนบิตข้อมูลของภาพต้นแบบ}}{\text{จำนวนบิตข้อมูลของภาพที่ถูกลดข้อมูล}} \quad (2.12)$$

ในวิทยานิพนธ์ฉบับนี้ได้นำเอาการลดข้อมูลแบบที่มีการสูญเสียมาประยุกต์ใช้ โดยใช้วิธีการที่เรียกว่าการแปลงเวฟเลต (Wavelet Transform) ซึ่งเป็นเทคนิคหนึ่งในการเข้ารหัสแบนด์ย่อย การแปลงเวฟเลตที่ใช้เป็นการแปลงเวฟเลตแบบแบ่งช่วง (Discrete Wavelet Transform, DWT) รายละเอียดจะกล่าวถึงในบทที่ 3

# บทที่ 3

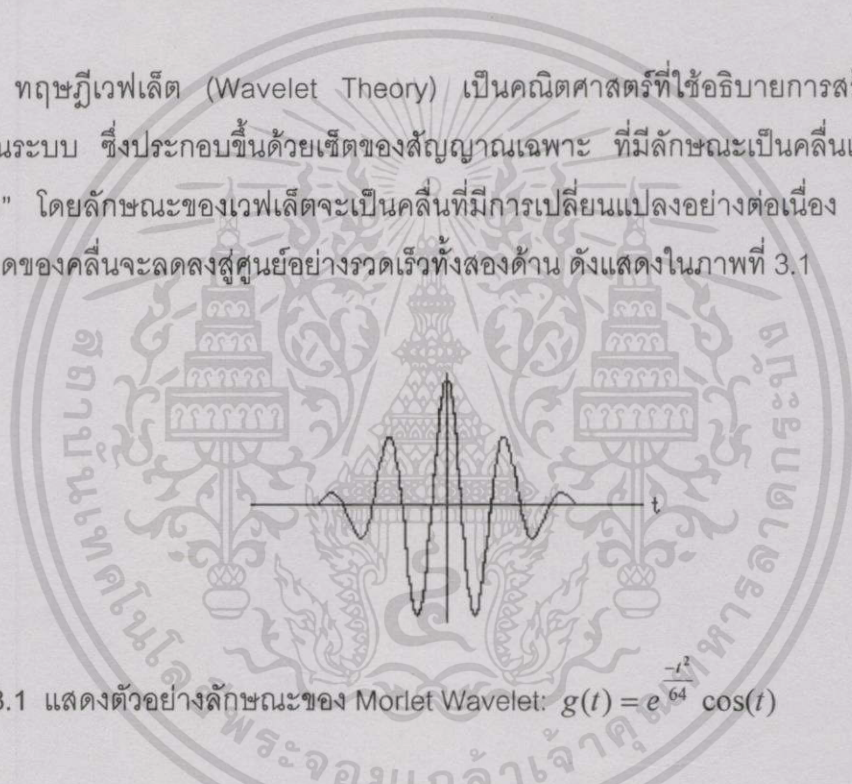
## การแปลงเวฟเล็ต

### 3.1 ทฤษฎีพื้นฐานของการแปลงเวฟเล็ต

ในบทนี้เป็นการอธิบายถึงความหมาย ลักษณะต่างๆ ไป ความแตกต่างระหว่างการแปลงเวฟเล็ตกับการแปลงข้อมูลแบบอื่นๆ และชนิดของการแปลงเวฟเล็ต เป็นต้น

#### 3.1.1 ลักษณะของเวฟเล็ต

ทฤษฎีเวฟเล็ต (Wavelet Theory) เป็นคณิตศาสตร์ที่ใช้อธิบายการสร้างโมเดลของสัญญาณระบบ ซึ่งประกอบขึ้นด้วยเซตของสัญญาณเฉพาะ ที่มีลักษณะเป็นคลื่นเล็กๆ เรียกว่า "เวฟเล็ต" โดยลักษณะของเวฟเล็ตจะเป็นคลื่นที่มีการเปลี่ยนแปลงอย่างต่อเนื่อง (Oscillatory) และขนาดของคลื่นจะลดลงสู่ศูนย์อย่างรวดเร็วทั้งสองด้าน ดังแสดงในภาพที่ 3.1



ภาพที่ 3.1 แสดงตัวอย่างลักษณะของ Morlet Wavelet:  $g(t) = e^{-t^2} \cos(t)$

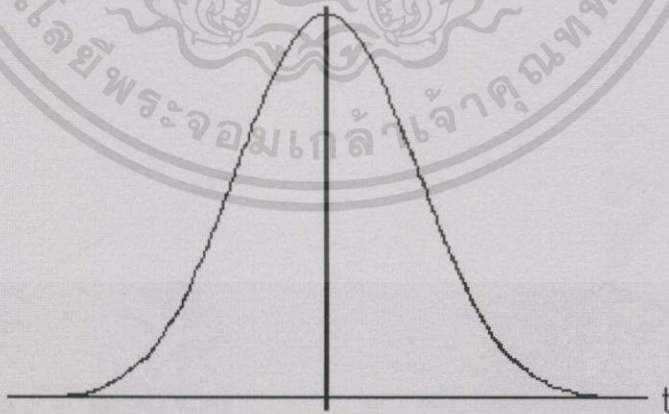
ในภาพที่ 3.1 เป็นตัวอย่างของเวฟเล็ตชนิดหนึ่ง เรียกว่าเวฟเล็ตแบบ Morlet จะสังเกตเห็นว่ามีคุณสมบัติของเวฟเล็ตตามที่กล่าวมาทั้ง 2 เงื่อนไข คือมีการออสซิลเลทในลักษณะของสัญญาณซายน์ซอยดัล (Sinusoidal) ดังแสดงในภาพที่ 3.2 รวมกับเงื่อนไขของการลดลงอย่างรวดเร็วตามลักษณะฟังก์ชันหน้าต่าง (Windows Function) ดังแสดงในภาพที่ 3.3

การนำเวฟเล็ตหลายๆ อันมารวมกันเป็นเซต เพื่อจะใช้ในการอธิบายสัญญาณใดๆ โดยที่คลื่นเวฟเล็ตแต่ละตัวมีโครงสร้างมาจากฟังก์ชันเดียวกัน ซึ่งฟังก์ชันนี้จะเป็นเวฟเล็ตต้นกำเนิดที่เรียกว่า เวฟเล็ตแม่ (Mother Wavelets) โดยคลื่นแต่ละคลื่นภายในเซตของเวฟเล็ตจะเกิดจากการสเกลและการเลื่อนตำแหน่งของเวฟเล็ตแม่ การสเกลในที่นี้หมายถึง การหดเข้า (Compressing) หรือการขยายออก (Dilation) โดยใช้ "a" แทนพารามิเตอร์ของการหดหรือขยายอันเป็นการเปลี่ยน

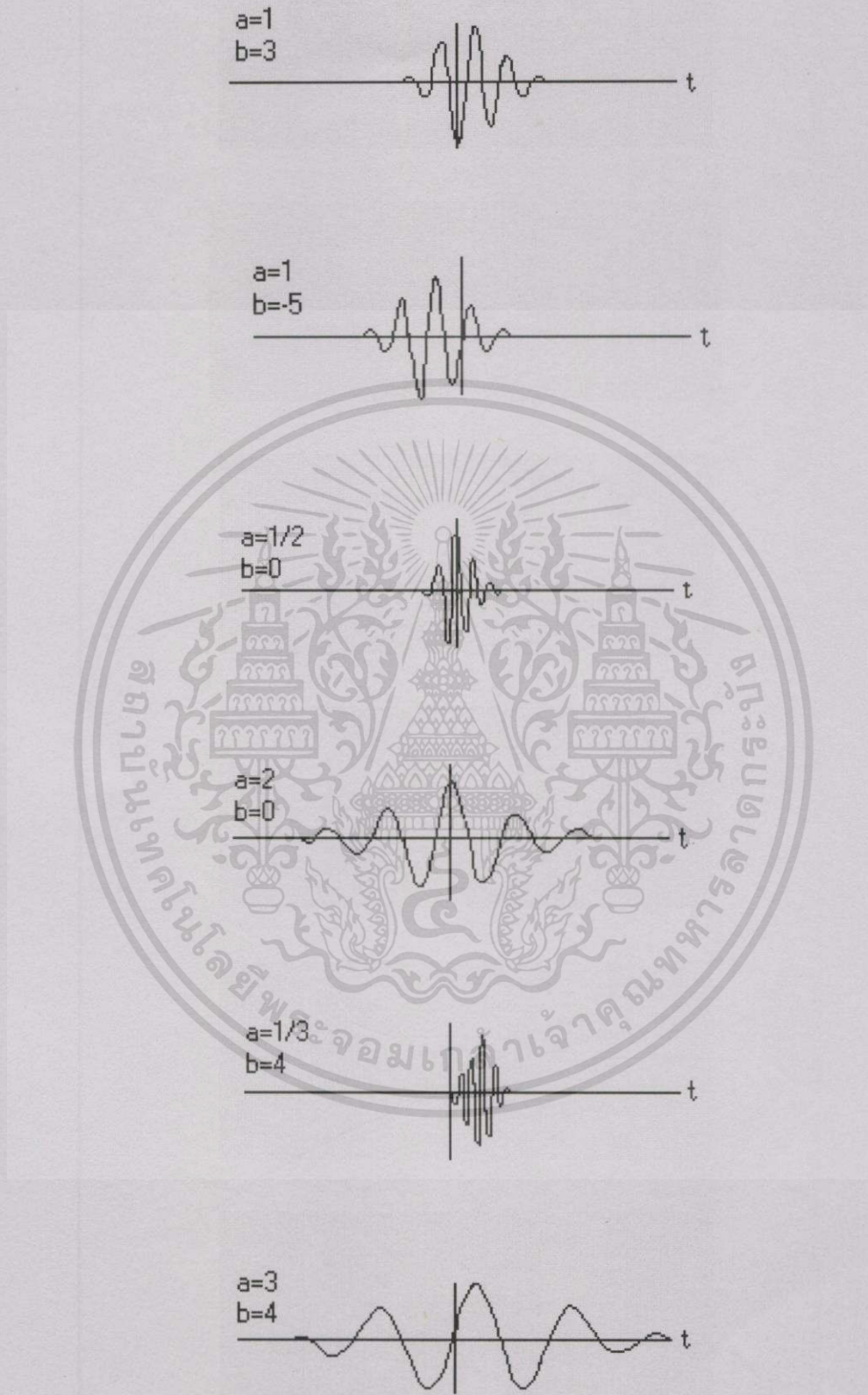
ความถี่ดังกล่าว ส่วนการเลื่อนตำแหน่งกำหนดโดยพารามิเตอร์ "b" ซึ่งหมายถึง การเลื่อนตำแหน่งของการเกิดคลื่นบนแกนเวลา ในภาพที่ 3.4 แสดงให้เห็นถึงเวฟเลิตที่ถูกสเกลและเลื่อนตำแหน่งไปที่ค่า "a", "b" ต่างๆ กัน โดยที่ยังคงอยู่ภายในเซตของเวฟเลิตที่สัมพันธ์กับเวฟเลิตแม่ในภาพที่ 3.1 นอกจากนี้เวฟเลิตที่ถูกสเกลจะต้องถูกนอร์มอลไลซ์ด้วย  $\frac{1}{\sqrt{a}}$  เสมอ เพื่อให้เวฟเลิตที่ถูกสเกลไปแล้วยังคงมีพลังงานเช่นเดียวกับเวฟเลิตแม่ ดังนั้นถ้าให้  $g(t)$  เป็นฟังก์ชันเวฟเลิตแม่จะสามารถเขียนสมการแสดงความสัมพันธ์ที่ตำแหน่ง "a", "b" ใดๆ ได้ดังนี้

$$g_{b,a}(t) = \frac{1}{\sqrt{a}} g\left(\frac{t-b}{a}\right) \quad (3.1)$$

ภาพที่ 3.2 แสดงลักษณะของสัญญาณชานูชอยด์ล:  $\cos(t)$



ภาพที่ 3.3 แสดงลักษณะของฟังก์ชันหน้าต่าง  $e^{-t^2/64}$

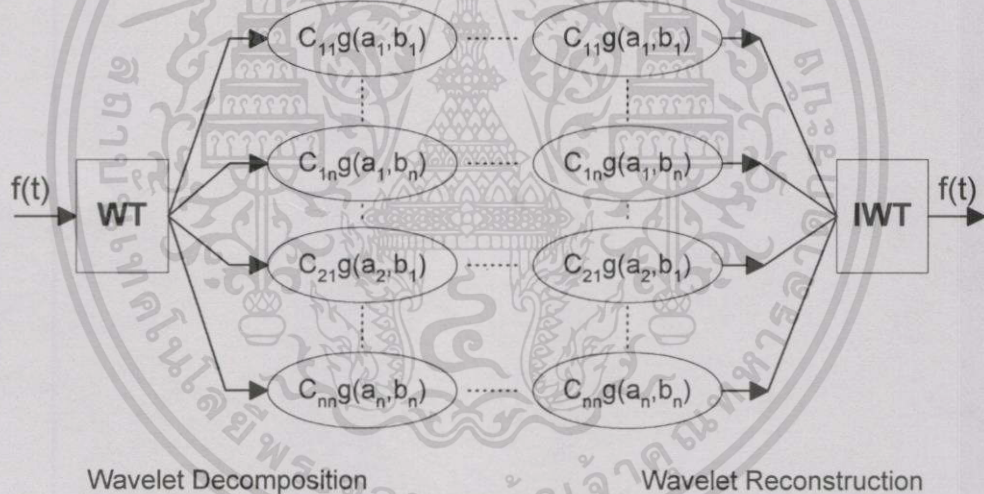


ภาพที่ 3.4 แสดงลักษณะของเวฟเล็ตแม่ที่ถูกสเกลและเลื่อนตำแหน่งไปที่ค่า  $a, b$  ต่างๆ กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

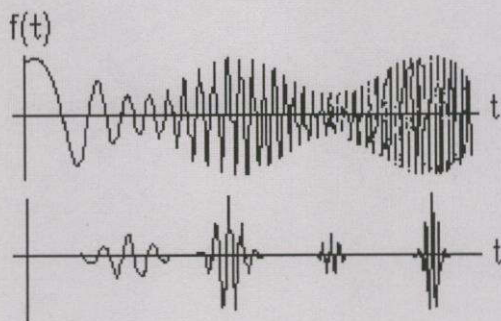
### 3.1.2 ทฤษฎีพื้นฐานเวฟเล็ตในการสเกลและการเลื่อนตำแหน่ง

การอธิบายสิ่งหนึ่งสิ่งใดในทฤษฎีเวฟเล็ตจะเปรียบเสมือนกับการแตกสิ่งเหล่านั้นออกเป็นส่วนประกอบเล็กๆ ที่สัมพันธ์กันคล้ายกับภาพต่อจิ๊กซอว์ โดยที่ชิ้นส่วนเหล่านั้นจะอยู่ในรูปของเวฟเล็ตที่ถูกสเกลและเลื่อนตำแหน่ง และมีค่าถ่วงน้ำหนักต่างๆ กันคุณอยู่กับเวฟเล็ตแต่ละตัว ดังนั้นจึงเปรียบเสมือนว่าคลื่นเวฟเล็ตเป็นฟังก์ชันพื้นฐาน (Basis function: b.f) ในการสร้างสัญญาณใดๆ โดยขบวนการแตกกระจายชิ้นส่วนนี้เรียกว่า "การแตกกระจายเวฟเล็ต" (Wavelet Decomposition) ซึ่งก็คือขั้นตอนการแปลงเวฟเล็ตนั่นเอง ในทำนองเดียวกัน การรวมกลับเวฟเล็ต (Wavelet Reconstruction) ก็คือการแปลงกลับเวฟเล็ต (Inverse Wavelets Transform: IWT) โดยจะเป็นการนำเอาส่วนประกอบย่อยๆ เหล่านี้มาเรียงต่อกันเพื่อประกอบเป็นสัญญาณเดิม ดังแสดงในภาพที่ 3.5 แสดงการแตกกระจายสัญญาณ  $f(t)$  ออกเป็นคลื่นเวฟเล็ตที่ความถี่และตำแหน่งต่างกันจำนวน  $n \times n$  ตัว



ภาพที่ 3.5 แสดงลักษณะของการแตกกระจายสัญญาณและรวมกลับสัญญาณของเวฟเล็ต

ในภาพที่ 3.6 ภาพบนแสดงสัญญาณ  $f(t)$  ที่ประกอบด้วยคลื่นหลายความถี่ ณ ตำแหน่งต่างๆ บนแกนเวลา ส่วนในภาพด้านล่างแสดงถึงการดึงสัญญาณ  $f(t)$  ที่ตำแหน่งเวลาใดๆ ออกมา



ภาพที่ 3.6 แสดงลักษณะของสัญญาณที่มีหลายความถี่ที่ตำแหน่งเวลาต่างกัน

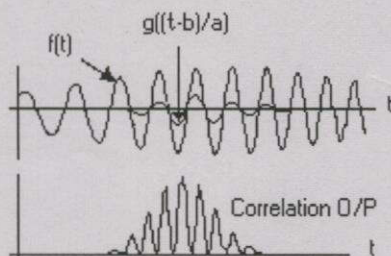
จากการที่สามารถทำการสร้างเซตของเวฟเล็ตได้จากการสเกลและการเลื่อนตำแหน่งของเวฟเล็ตแม่นั้น การสเกลและการเลื่อนตำแหน่งดังกล่าวจะถูกเรียกรวมกันว่า Affine Operation [2] ซึ่งจะสามารถเขียนเป็นสมการคณิตศาสตร์ในลักษณะการส่ง (Mapping) ตัวแปรอิสระ  $t$  ไปยังตัวแปรอิสระตัวใหม่  $t'$  ดังแสดงในสมการที่ (3.2)

$$t' = \frac{t-b}{a} \quad (3.2)$$

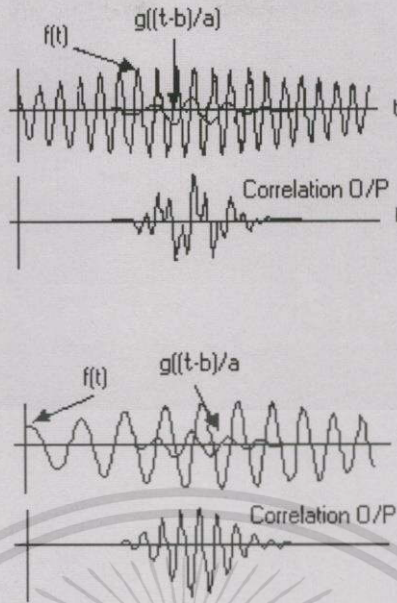
และจะสามารถเขียนสมการสำหรับการแปลงเวฟเล็ตแบบต่อเนื่อง (Continuous wavelets Transform: CWT) ที่ใช้ในการกระจายสัญญาณของฟังก์ชัน  $f(t)$  ด้วยเวฟเล็ตแม่  $g(t)$  ได้ดังนี้

$$(W_g f)(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) g^* \left( \frac{t-b}{a} \right) dt \quad (3.3)$$

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



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.7 แสดงลักษณะของผลการ Correlation ระหว่าง  $f(t), g((t-b)/a)$

### 3.2 ความเป็นมาของการแปลงเวฟเล็ด

การแปลงฟูเรียร์นิยมใช้งานอย่างแพร่หลายซึ่งเป็นการแปลงสัญญาณ  $f(t)$  จากโดเมนส์เวลาไปเป็นสัญญาณ  $F(\omega)$  ซึ่งอยู่ในโดเมนส์ความถี่หรือที่เรียกกันทั่วไปว่าเป็นการหาสเปกตรัมของสัญญาณ  $f(t)$  นั้นเอง โดยมีสมการเป็น

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (3.4)$$

จะสังเกตพบว่าข้อมูลที่ได้จากการทรานส์ฟอร์ม จะมีตลอดย่านความถี่  $-\infty$  ถึง  $+\infty$  ดังนั้นเมื่อเกิดการเปลี่ยนแปลงของสัญญาณอินพุต ณ ขณะเวลาใดๆ ก็จะมีส่งผลกระทบต่อตลอดย่านของสเปกตรัม  $F(\omega)$  ด้วย นอกจากนี้การแปลงฟูเรียร์ยังไม่สามารถเลือกทำการวิเคราะห์สัญญาณเฉพาะช่วงเวลาหรือความถี่ใดๆ ได้ ดังนั้นวิธีการแปลงฟูเรียร์จึงให้ข้อมูลที่มากเกินไปจนเกินไป นอกจากนี้ในทางปฏิบัติหากสัญญาณอินพุตจะมีการเปลี่ยนแปลงที่ไม่คงที่ (Non-Stationary Signal) การใช้การแปลงฟูเรียร์จะทำให้สิ้นเปลืองเวลาในการคำนวณเนื่องจากมีข้อมูลมากเกินไปในการคำนวณใหม่ในแต่ละครั้ง ดังนั้นการแปลงฟูเรียร์จึงเหมาะสำหรับสัญญาณที่มีการเปลี่ยนแปลงคงที่ (Stationary Signal)

การแปลงกabor (Gabor Transform) เป็นการแปลงที่ได้พัฒนาขึ้นเป็นแบบแรกเพื่อแก้ปัญหาที่กล่าวมาข้างต้น โดยจะเป็นการทำการแปลงฟูเรียร์เฉพาะช่วงเวลาและช่วงความถี่ที่

กำหนดผ่านฟังก์ชันหน้าต่าง (Window function: wf) ซึ่งหน้าต่างนี้จะใช้ฟังก์ชันแบบเกาส์เซียน (Gaussian function) เท่านั้น

$$g_a(t) = \frac{1}{2\sqrt{\pi a}} e^{-\frac{t^2}{4a}} \quad (3.5)$$

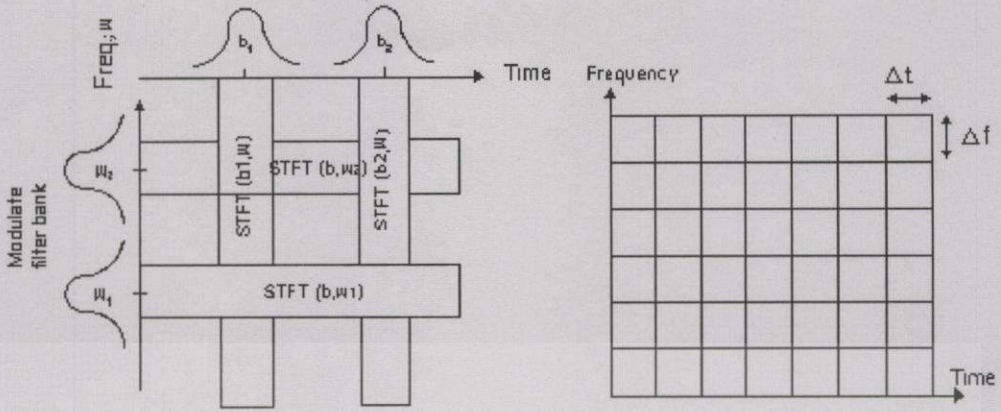
$$(G_b^a f)(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) g_a(t-b) dt \quad (3.6)$$

จากสมการที่ (3.5) เป็นฟังก์ชันหน้าต่างแบบเกาส์เซียน และสมการที่ (3.6) เป็นสมการของการแปลงกาบอร์ จะเห็นว่าการทรานส์ฟอร์มแบบนี้จะสามารถเลือกตำแหน่งการวิเคราะห์สัญญาณ โดยกำหนดที่พารามิเตอร์  $b$  และเลือกช่วงความถี่ด้วยพารามิเตอร์  $a$  แต่เนื่องจากการแปลงกาบอร์ให้ฟังก์ชันหน้าต่างแบบเดียวจึงอาจจะไม่เหมาะสมกับการประยุกต์ใช้งานในด้านต่างๆ ดังนั้น จึงได้มีการพัฒนาต่อมาเป็นการแปลงฟูเรียร์ช่วงเวลาสั้น (Short Time Fourier Transform: STFT) ซึ่งจะสามารถทำการเลือกฟังก์ชันหน้าต่างที่จะใช้ในการวิเคราะห์สัญญาณได้ ทำให้เกิดความยืดหยุ่นในการเลือกใช้ให้เหมาะสมกับสัญญาณอินพุตที่จะวิเคราะห์ได้ดีขึ้น

$$G_a f(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) w(t-b) dt \quad (3.7)$$

สมการที่ (3.7) จะเป็นสมการการแปลงฟูเรียร์ช่วงเวลาสั้น โดยที่ตำแหน่งเวลาในการวิเคราะห์จะกำหนดโดยค่า " $b$ " และช่วงความถี่ในการวิเคราะห์กำหนดด้วยความถี่ของฟังก์ชันหน้าต่างนั้นๆ แต่จะสังเกตเห็นว่าทั้งการแปลงกาบอร์และการแปลงฟูเรียร์ช่วงเวลาสั้น จะมีลักษณะของช่วงการวิเคราะห์ที่คงที่ ดังแสดงในภาพที่ 3.8 แสดงช่วงเวลาและช่วงความถี่คงที่ (Fixed resolution transform) ทั้งนี้ฟังก์ชันหน้าต่างไม่มีการเปลี่ยนแปลงคุณสมบัติขณะทำการแปลงซึ่งไม่เหมาะสมในทางปฏิบัติ เนื่องจากสัญญาณที่มีความถี่สูงจะมีช่วงคาบเวลาที่แคบ ในขณะที่สัญญาณความถี่ต่ำจะมีช่วงคาบเวลาที่กว้างกว่า ดังนั้นการแปลงกาบอร์และการแปลงฟูเรียร์ช่วงเวลาสั้น อาจจะทำให้ข้อมูลที่ได้จากการแปลงออกมามากหรือน้อยเกินไป ถ้าเป็นในกรณีหลังจะทำให้ต้องมีการแปลงจากหลายช่วงเวลาแล้วนำมาประกอบกัน จึงจะได้ข้อมูลครบถ้วนในช่วงเวลาที่ต้องการ ดังนั้นช่วงเวลาควรจะต้องมีการเปลี่ยนแปลงสัมพันธ์กับช่วงความถี่ที่จะทำการวิเคราะห์ นั่นคือในช่วงความถี่ที่สูงขึ้นจะมีช่วงเวลาในการวิเคราะห์ที่น้อยลง และในช่วงความถี่ที่ต่ำลงจะใช้ช่วงเวลาในการวิเคราะห์ที่มากขึ้น [2]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

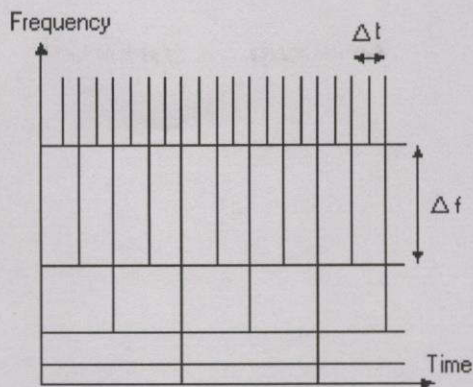


ภาพที่ 3.8 แสดงลักษณะของระนาบเวลาความถี่แบบคงที่ (Constant Time-Frequency plane)

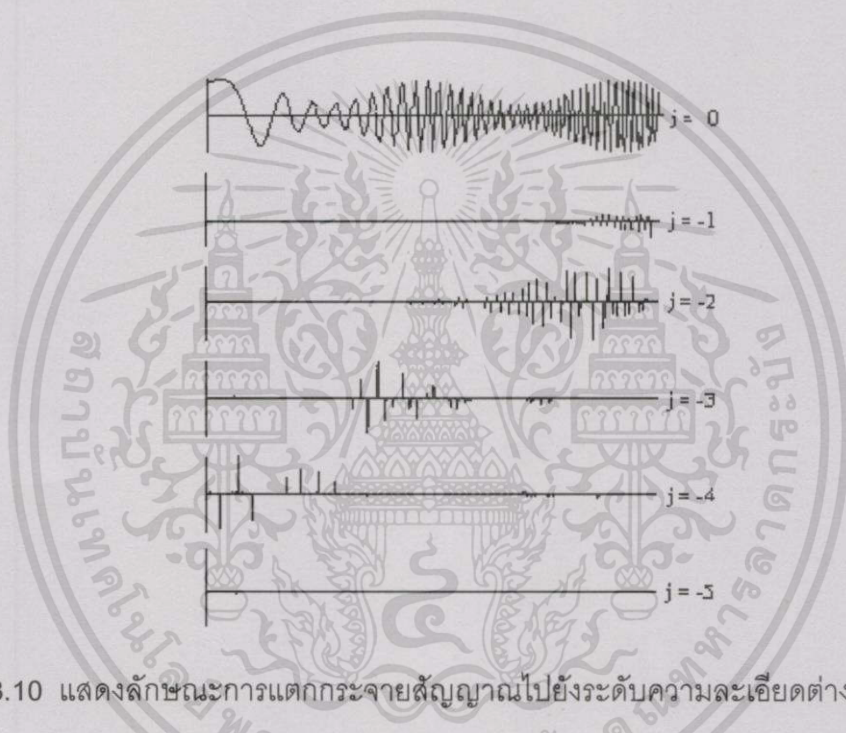
จากเหตุผลดังกล่าวข้างต้นทำให้มีการพัฒนาเป็นการแปลงในรูปแบบใหม่ที่เรียกว่า “การแปลงเวฟเล็ต” ซึ่งจะเป็นการแปลงผ่านฟังก์ชันหน้าต่าง โดยฟังก์ชันหน้าต่างสามารถเลื่อนตำแหน่งและปรับเปลี่ยนช่วงเวลาให้เหมาะกับช่วงความถี่ที่จะวิเคราะห์ได้ โดยเรียกฟังก์ชันหน้าต่างแบบนี้ว่า “ฟังก์ชันเวฟเล็ต” (Wavelet function)

$$(W_g f)(a, b) = |a|^{-1} \int_{-\infty}^{\infty} f(t) g^* \left( \frac{t-b}{a} \right) dt \tag{3.8}$$

$g(t)$  เป็นฟังก์ชันเวฟเล็ตแม่ ซึ่งจะถูกสเกลและเลื่อนตำแหน่งโดยพารามิเตอร์ "a", "b" ตามลำดับ โดยช่วงเวลาและความถี่ในการทรานส์ฟอร์มจะสัมพันธ์กันดังที่ได้กล่าวมาแล้วข้างต้น พิจารณาภาพที่ 3.9 จะพบว่าระนาบเวลาและความถี่จะมีลักษณะการเปลี่ยนแปลงของช่วงเวลา และความถี่ที่สัมพันธ์กันคือ ในช่วงความถี่สูงจะใช้ช่วงเวลาในการวิเคราะห์เพียงสั้นๆ ( $\Delta f$  สูง  $\Delta t$  ต่ำ) แต่ในช่วงความถี่ต่ำจะมีช่วงเวลาในการวิเคราะห์ที่นานกว่า ( $\Delta f$  ต่ำ  $\Delta t$  สูง) ซึ่งจะเหมาะสมกับลักษณะสัญญาณในทางปฏิบัติ ในภาพที่ 3.10 จะเป็นการแสดงลักษณะการแตกกระจายสัญญาณออกไปในระดับความถี่ที่ต่างกัน โดยในระดับที่  $j = -1$  จะมีเฉพาะส่วนความถี่สูง และในระดับที่  $j$  ต่ำลง สัญญาณจะมีค่าความถี่ที่ต่ำลงด้วย ซึ่งจะเห็นว่าสอดคล้องกับภาพของระนาบเวลาและความถี่



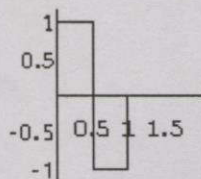
ภาพที่ 3.9 แสดงลักษณะของระนาบเวลาและความถี่สำหรับการแปลงเวฟเล็ต



ภาพที่ 3.10 แสดงลักษณะการแตกกระจายสัญญาณไปยังระดับความละเอียดต่างๆ

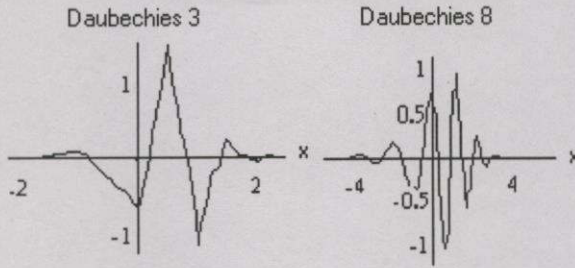
นอกจากนี้  $g(t)$  ยังมีได้หลายรูปแบบแต่ต้องเป็นไปตามคุณสมบัติของการแปลงเวฟเล็ต ดังนั้นจึงทำให้การแปลงเวฟเล็ตมีความยืดหยุ่นและสามารถนำไปประยุกต์ใช้กับงานที่มีความเหมาะสมได้เป็นอย่างดีดังภาพที่ 3.11 แสดงลักษณะของฟังก์ชันเวฟเล็ตแม่ชนิดต่างๆ ที่นิยมใช้ในปัจจุบัน

#### Harr wavelet

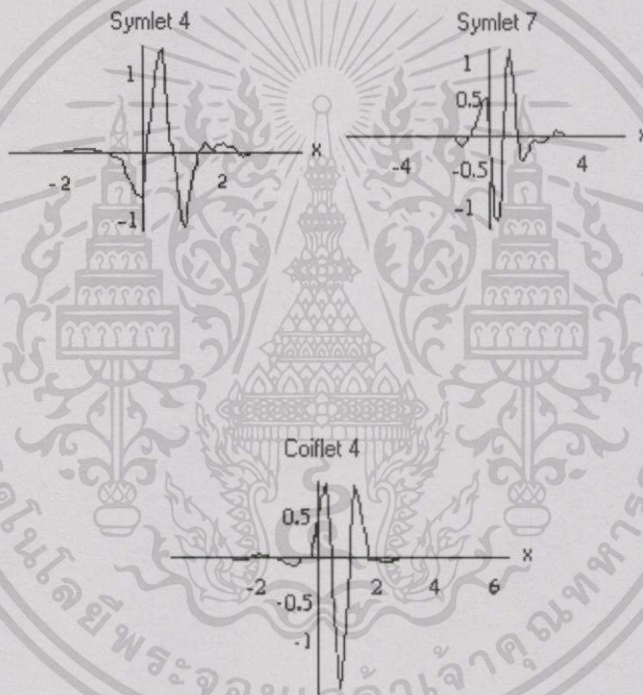


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Daubechies wavelet

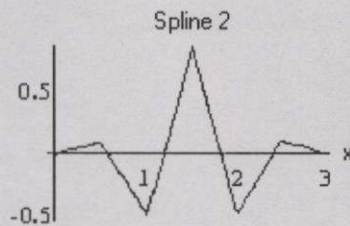


## Symlet wavelet



## Coiflet wavelet

## Spline wavelet



ภาพที่ 3.11 แสดงลักษณะของฟังก์ชันเวฟเลตแม่ชนิดต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3 การแปลงเวฟเล็ตแบบต่อเนื่อง

จากการวิเคราะห์สัญญาณด้วยการแปลงฟูเรียร์ช่วงเวลาสั้น จะสังเกตเห็นว่าหน้าต่างเวลา-ความถี่ (Time-frequency window) จะมีพื้นที่ในการวิเคราะห์ที่คงที่เสมอไม่ว่าจะเป็นการพิจารณาที่ความถี่และช่วงเวลาใดๆ ขนาดช่วงความกว้างของความถี่และช่วงเวลาจะขึ้นอยู่กับลักษณะของฟังก์ชันหน้าต่างที่ใช้ ซึ่งจะไม่มีการเปลี่ยนแปลงคุณสมบัติในขณะการแปลง แต่ในทางปฏิบัติความถี่ของสัญญาณจะเป็นสัดส่วนโดยตรงกับจำนวนไซเคิลต่อหนึ่งหน่วยเวลา ที่ความถี่สูงจะมีการเปลี่ยนแปลงสัญญาณในช่วงเวลาที่แคบ ส่วนการเปลี่ยนแปลงสัญญาณในช่วงความถี่ต่ำจะใช้ช่วงเวลาที่นานกว่า ดังนั้นถ้าใช้การแปลงฟูเรียร์ช่วงเวลาสั้น และเลือกฟังก์ชันหน้าต่างที่มีช่วงความถี่และช่วงเวลาไม่เหมาะสมจะทำให้ต้องมีการทำการแปลงหลายครั้งทีละตำแหน่ง เพื่อให้ได้ข้อมูลในช่วงเวลาทั้งหมดที่ต้องการ หรือถ้าความกว้างของช่วงเวลาของฟังก์ชันหน้าต่างมากเกินไปจะทำให้ได้รับข้อมูลมากเกินไปจนความจำเป็น ดังนั้นฟังก์ชันหน้าต่างที่ดี ความกว้างของช่วงเวลาและช่วงความถี่จะต้องมีการเปลี่ยนแปลงเป็นสัดส่วนผกผันที่สัมพันธ์กัน นั่นคือเมื่อทำการวิเคราะห์สัญญาณที่ความถี่สูง ความกว้างของช่วงเวลาควรแคบและมีช่วงความถี่ที่กว้าง ในทางกลับกันถ้าวิเคราะห์สัญญาณที่ความถี่ต่ำ ความกว้างของช่วงเวลาควรกว้างและมีช่วงความถี่ที่แคบ เพื่อที่จะทำให้ได้รับข่าวสารที่ครบถ้วนและใช้เวลาในการแปลงน้อยลง

จากที่กล่าวมาทั้งหมดจะเห็นได้ว่าการแปลงฟูเรียร์ช่วงเวลาสั้น ไม่เหมาะสมสำหรับการวิเคราะห์สัญญาณที่มีการเปลี่ยนความถี่สูงต่ำในสัญญาณตัวเดียวกัน แต่ในการแปลงเวฟเล็ต จะอาศัยฟังก์ชันหน้าต่าง ซึ่งสามารถเปลี่ยนแปลงคุณสมบัติของช่วงเวลาและช่วงความถี่ให้เหมาะสมกับการวิเคราะห์สัญญาณในช่วงความถี่และช่วงเวลาที่น่าสนใจใดๆ ได้ ซึ่งฟังก์ชันหน้าต่างที่มีคุณสมบัติดังกล่าวเรียกว่า “ฟังก์ชันเวฟเล็ต”

กำหนดให้  $\psi(t) \in L^2(\mathbb{R})$  และเป็นไปตาม Admissibility condition [1] คือ

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \quad (3.9)$$

เรียก  $\psi(t)$  ว่าเป็นรากฐานเวฟเล็ต (Basis wavelets) หรือเวฟเล็ตแม่ ซึ่งมีคุณสมบัติในการเลื่อนตำแหน่งและเปลี่ยนความถี่ได้ ดังแสดงในสมการที่ (3.10)

$$\psi_{b,a}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-a}{b}\right) \quad (3.10)$$

และเมื่อนำมาเขียนเป็นสมการการแปลงฟูเรียร์ที่พิจารณาผ่านฟังก์ชันหน้าต่าง ที่ ตำแหน่งเวลาและความถี่ในขณะนั้นได้ตั้งสมการที่ (3.11) ซึ่งก็คือการแปลงเวฟเล็ตแบบต่อเนื่อง นั้นเอง

$$(W_{\psi} f)(b, a) = |a|^{\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (3.11)$$

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

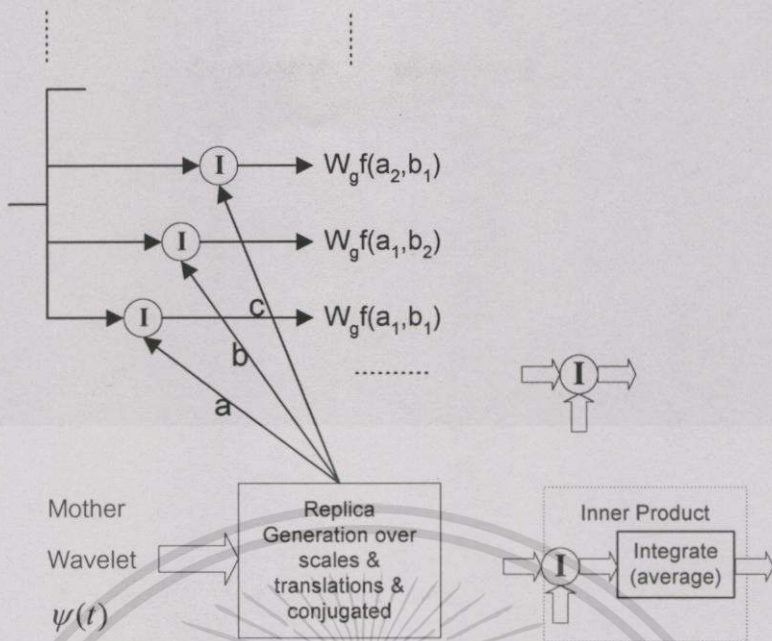
เทอม  $|a|^{\frac{1}{2}}$  จะเป็นการนอร์มอลไลซ์ (Normalization) ของ  $\psi(t)$  เพื่อให้พลังงานของ  $\psi(t)$  ที่  $b, a$  ใดๆ ยังคงมีค่าเท่ากับหนึ่งเสมอ ทำให้การแปลงเวฟเล็ตที่ตำแหน่ง  $b, a$  ใดๆ พลังงานของ  $\psi(t)$  จะไม่มีผลกระทบต่อสัญญาณ  $f(t)$  ในช่วงนั้นๆ และจะสามารถเขียนสมการการแปลงเวฟเล็ตแบบต่อเนื่องในลักษณะของ Inner product ได้ดังนี้

$$(W_{\psi} f)(b, a) = \langle f, \psi_{b,a} \rangle \quad (3.12)$$

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

$$\begin{aligned} \hat{\psi}(\omega) &= \int_{-\infty}^{\infty} e^{-i\omega t} \psi(t) dt \\ \hat{\psi}(0) &= \int_{-\infty}^{\infty} \psi(t) dt = 0 \end{aligned} \quad (3.13)$$

แสดงว่าพื้นที่ใต้กราฟของ  $\psi(t) = 0$  นั่นก็คือฟังก์ชันดังกล่าวต้องมีการออกสลับเสมอ



หมายเหตุ

$$\begin{aligned}
 a &= \psi^*[(t-b_1)/a_1] \\
 b &= \psi^*[(t-b_2)/a_1] \\
 c &= \psi^*[(t-b_2)/a_2]
 \end{aligned}$$

ภาพที่ 3.12 แสดงการแปลงเวฟเล็ตในลักษณะของการ Inner product

ถ้าพิจารณาหน้าต่างเวลา-ความถี่ ของการแปลงเวฟเล็ต โดยที่  $\psi(t), \hat{\psi}(\omega)$  เป็นไปตามเงื่อนไข  $tW(t) \in L^2(\mathbb{R})$  และกำหนดให้จุดศูนย์กลางและรัศมีมีความกว้างของ  $\psi(t)$  เป็น  $t^*$  และ  $\Delta\psi$  ตามลำดับ ดังนั้น  $\psi_{b,a}$  จะมีจุดศูนย์กลางอยู่ที่  $b + at^*$  และมีรัศมีเป็น  $a\Delta\psi$  โดยข้อมูลของสัญญาณ  $f(t)$  จะถูกพิจารณาผ่าน Time-window ที่มีความกว้างของช่วงเวลาดังนี้ [1]

$$[b + at^* - a\Delta\psi, b + at^* + a\Delta\psi] \tag{3.3}$$

### 3.4 การแปลงเวฟเล็ตแบบแบ่งช่วง

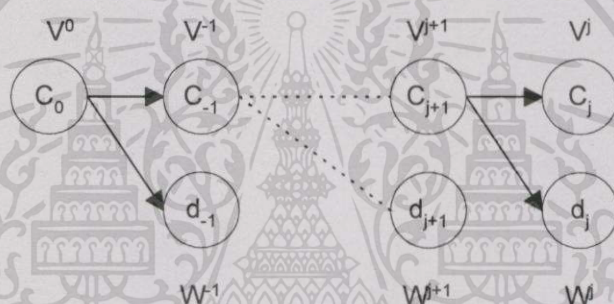
ในหัวข้อนี้จะเป็นการอธิบายถึง การแปลงเวฟเล็ตแบบแบ่งช่วง (Discrete Wavelets Transform: DWT) โดยจะทำการพิจารณาการแปลงเวฟเล็ตแบบแบ่งช่วงในลักษณะของ Filter banks ในทางปฏิบัติสัญญาณที่วิเคราะห์ด้วยคอมพิวเตอร์จะเป็นการสุ่ม (Sampling) สัญญาณเข้ามาในจำนวนที่จำกัด ดังนั้นถ้ากำหนดให้  $c_0(m)$  เป็นเซตของสัมประสิทธิ์ของสัญญาณอินพุตจำนวน  $m$  ตัว และกำหนดให้ระดับความละเอียดของสัญญาณเริ่มต้นที่ศูนย์ จะสามารถเขียนเป็นความสัมพันธ์ได้ว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$f(t) \in V_0; f(t) \in L^2(\mathbb{R}) \tag{3.15}$$

$$V_0 = W_{-1} \oplus \dots \oplus W_{j+2} \oplus W_{j+1} \oplus W_j \oplus V_j \tag{3.16}$$

สมการที่ (3.16) จะเป็นสมการการแตกกระจายสัญญาณ  $f(t)$  ในสเปซ  $V^0$  ไปจนถึงระดับความละเอียดที่  $j$  ดังนั้นเทอมสัมประสิทธิ์  $c_0(m)$  จะถูกแตกเป็นสัมประสิทธิ์ในเซตของ  $c_j(m)$  ในสเปซ  $V^j$  และกลุ่มของเซต  $d_{-1}(m), d_{-2}(m), \dots, d_j(m)$  ซึ่งอยู่ในเวฟเล็ตสเปซที่ระดับความละเอียดต่างๆ กัน ขบวนการดังกล่าวนี้จะเป็นลักษณะของวิธีการ “การแปลงเวฟเล็ตแบบแบ่งช่วง” (Discrete Wavelets Transform: DWT) นั่นเอง ดังแสดงในภาพที่ 3.13



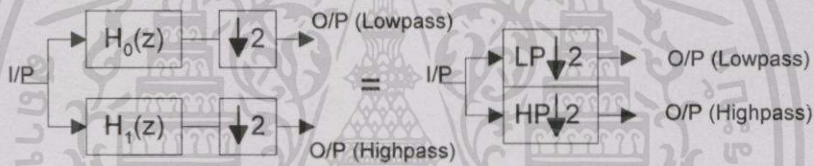
ภาพที่ 3.13 แสดงลักษณะของการแจกแจงสัมประสิทธิ์ด้วยการแปลงเวฟเล็ต

### 3.4.1 การแปลงเวฟเล็ตโดยใช้หลักการวิเคราะห์ของฟิลเตอร์แบงค์

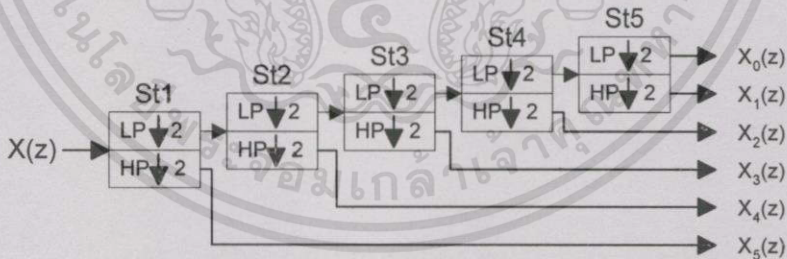
ขบวนการแปลงเวฟเล็ตแบบแบ่งช่วง จะทำการพิจารณาสัญญาณผ่านฟังก์ชันหน้าต่างซึ่งมีการเปลี่ยนแปลงความถี่ในอัตราครึ่งละสองเท่า ก็เปรียบเสมือนกับการนำสัญญาณอินพุตนั้นผ่านวงจรฟิลเตอร์ที่มีแบนด์วิทที่มีอัตราการลดลงสองเท่าเหมือนกับแบนด์วิทของฟังก์ชันหน้าต่างในขณะนั้นนั่นเอง ดังนั้นจึงสามารถนำเอาหลักการของฟิลเตอร์แบงค์ มาใช้ในการสร้างการแปลงเวฟเล็ตแบบแบ่งช่วงในทางปฏิบัติได้ ก่อนที่จะอธิบายถึงรายละเอียดของการแปลงเวฟเล็ตแบบแบ่งช่วง ขออธิบายถึงหลักการพื้นฐานของฟิลเตอร์แบงค์ก่อนดังนี้

### 3.4.1.1 การวิเคราะห์ห่ออกเทพของฟิลเตอร์แบงก์ (Octave Analysis Filter Banks)

ฟิลเตอร์แบงก์แบบ 2 ช่องสัญญาณ (Two channel filter banks) เป็นการแยกสัญญาณอินพุตออกเป็น 2 ส่วน โดยแบ่งเป็นส่วนความถี่ต่ำและส่วนความถี่สูง ดังนั้นฟิลเตอร์แบงก์แบบสองช่องสัญญาณ จึงประกอบด้วยวงจรที่เป็น ตัวกรองความถี่ต่ำ (lowpass filter):  $H_0(z)$  และ ตัวกรองความถี่สูง (Complementary highpass filter):  $H_1(z)$  ซึ่งในส่วนของความถี่สูง (highpass) จะใช้ในการแยกความถี่ในแบนด์ที่ต้องการออกมา (Band-separating filter) ดังแสดงในภาพที่ 3.14, ภาพที่ 3.15 และภาพที่ 3.16 ลักษณะของ การวิเคราะห์ห่ออกเทพของฟิลเตอร์แบงก์ จะเป็นแบบโครงสร้างต้นไม้ (Tree structure) ซึ่งเป็นการนำเอาฟิลเตอร์แบงก์แบบ 2 ช่องสัญญาณมาเรียงต่อกัน โดยใช้สัญญาณเอาท์พุตในส่วนความถี่ต่ำจากวงจรตัวกรองความถี่ต่ำ มาทำการแยกแบนด์ความถี่ออกอีกครั้งหนึ่งดังแสดงในภาพที่ 3.16



ภาพที่ 3.14 แสดงสัญลักษณ์ของฟิลเตอร์แบงก์แบบ 2 ช่องสัญญาณ หรือ Bands separator

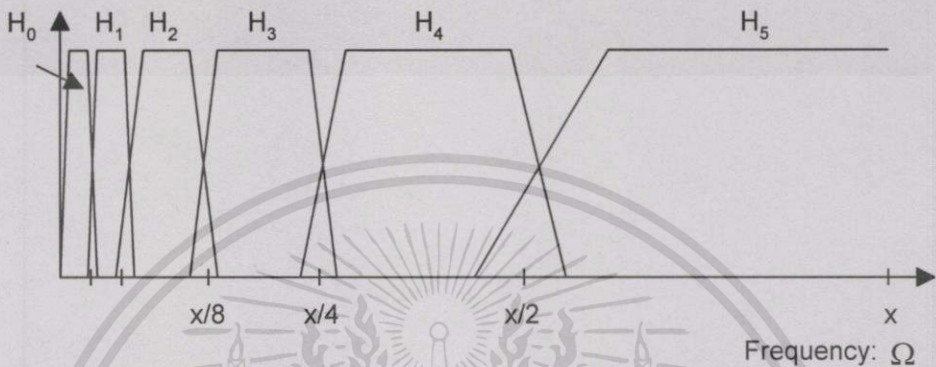


ภาพที่ 3.15 แสดงลักษณะโครงสร้างต้นไม้ในการวิเคราะห์ห่ออกเทพของฟิลเตอร์แบงก์

ในฟิลเตอร์แบงก์แบบ 2 ช่องสัญญาณ ชุดแรก จะทำการแยกสัญญาณความถี่สูงครั้งด้านบน ( $X_5(z)$ ) และสัญญาณความถี่ต่ำครั้งด้านล่างออกจากกัน จากนั้นจะนำสัญญาณความถี่ต่ำครั้งด้านล่างมาทำการแยกแบนด์ความถี่ออกอีกครั้งหนึ่งอีกครั้ง โดยกำหนดให้ครั้งบนทางความถี่สูงเป็นสัญญาณ  $X_4(z)$  และจะนำส่วนความถี่ต่ำไปทำการแยกแบนด์ต่อไปเรื่อยๆ ตามระดับที่ต้องการ ในภาพที่ 3.16 แสดงให้เห็นถึงลักษณะของแบนด์ความถี่ที่ถูกแบ่งออกไปเรื่อยๆ สัมพันธ์กับอัตราการสุ่มของสัญญาณอินพุต  $X(z)$  ซึ่งหมายถึงจำนวนข้อมูลการสุ่มจะลดลง เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สองเท่าในแต่ละสเตจนั่นเอง พังก์ชันทรานส์เฟอร์ของฟิลเตอร์แบงค์ในแต่ละระดับจะหาได้จากสมการ (3.17)

$$H_i(z) = \frac{X_i(z)}{X(z)} \text{ เมื่อ } i = 0, 1, \dots, 5 \quad (3.17)$$



ภาพที่ 3.16 แสดงลักษณะของการแยกแบนด์ความถี่ของออกเทฟฟิลเตอร์แบงค์

Cut off frequency มีค่าลดลงครึ่งละสองเท่า ดังนั้นจึงเรียกว่าออกเทฟฟิลเตอร์แบงค์ (Octave filter banks) และจากภาพที่ 3.16 จะสังเกตว่า Absolute transition bandwidth ซึ่งเป็นเอาทพุทของความถี่ต่ำ จะลดลงในอัตราสองเท่า โดยลดลงไปในทางความถี่ที่ต่ำลง และพังก์ชันทรานส์เฟอร์:  $H_0(z)$  จะเป็นส่วนความถี่ต่ำครั้งสุดท้าย และ  $H_5(z)$  จะเป็นส่วนความถี่สูง ในขณะที่พังก์ชันทรานส์เฟอร์อื่นๆ จะมีลักษณะเป็น Band pass ดังนั้นฟิลเตอร์แบงค์จึงมีความสามารถในการแยกแบนด์ความถี่ต่างๆ ออกมา (Band-separating filter)

### 3.4.1.2 การสร้าง DWT โดยใช้การวิเคราะห์ฟิลเตอร์แบงค์

ในหัวข้อนี้จะแสดงวิธีการพิจารณาการแจกแจงสัมประสิทธิ์ของการโปรเจกต์สัญญาณในสเปซ  $V^{j+1}$  ไปในสเปซย่อย  $V^j$  และ  $W^j$  ซึ่งเป็นการทำ DWT หนึ่งระดับในรูปแบบของการวิเคราะห์ด้วยฟิลเตอร์แบงค์โดยอาศัยหลักการของฟิลเตอร์แบงค์แบบ 2 ช่องสัญญาณ

กำหนดให้สัญญาณอินพุต  $f(t) \in V^{j+1}$  ดังนั้นจะได้ว่า

$$f_{j+1}(t) = \sum_n c_{j+1}(n) \phi_{j+1,n}(t) \in V^{j+1} \quad (3.18)$$

ซึ่งสัมประสิทธิ์  $c_{j+1}(n)$  เป็นค่าสัมประสิทธิ์ข้อมูลสุ่มที่ทราบค่าของ  $f(t)$  จากความสัมพันธ์  $V^{j+1} = V^j \oplus W^j$  จะสามารถเขียนเป็นการโปรเจกต์  $f(t)$  ลงในสเปซ  $V^j$  และ  $W^j$  ได้เป็น

$$f_{j+1}(t) = \sum_m c_j(m) \phi_{j,m}(t) + \sum_m d_j(m) \psi_{j,m}(t) \quad (3.19)$$

ในทางปฏิบัติการโปรเจกต์ชั้นจะเป็นการหาค่าของ  $c_j(m)$  และ  $d_j(m)$  จาก  $c_{j+1}(n)$  และขณะเดียวกันฟังก์ชันพื้นฐาน  $\phi_{j+1,m}(t)$  ของสเปซ  $V^{j+1}$  สามารถถูกแบ่งออกเป็นเซตของฟังก์ชันพื้นฐาน  $\phi_{j,m}(t)$  ในสเปซ  $V^j$  ด้วย โดยใช้สมการที่ (3.20)

$$\phi_{j,m}(t) = 2^{j/2} \sum_v h_0(v) \phi(2^{j+1}t - 2m - v) \quad (3.20)$$

ถ้าสมมติให้  $2m + v \rightarrow n$  และ  $h_0(n) = 2^{-1} h_0 k, k \in \mathbb{Z}$  จะได้ว่า

$$\begin{aligned} \phi_{j,m}(t) &= \sum_n h_0(n-2m) 2^{(j-1)/2} \phi(2^j t - n) \\ &= \sum_n h_0(n-2m) \phi_{j-1,n}(t) \end{aligned} \quad (3.21)$$

ในทำนองเดียวกันฟังก์ชันพื้นฐาน  $\phi_{j+1,m}(t)$  ของสเปซ  $V^{j+1}$  จะถูกแบ่งออกเป็นเซตของฟังก์ชันพื้นฐาน  $\psi_{j,m}(t)$  ในสเปซ  $W^j$  ด้วยเช่นกัน ดังนั้น  $\psi_{j,m}(t)$  สามารถกำหนดได้โดยวิธีเดียวกัน ดังสมการที่ (3.22)

$$\psi_{j,m}(t) = \sum_n h_1(n-2m) \phi_{j+1,n}(t) \quad (3.22)$$

ถ้าต้องการหาสัมประสิทธิ์ของฟังก์ชันสเกลลิง  $c_j(m)$  จะสามารถหาได้จากสมการที่ (3.23) โดยใช้ความสัมพันธ์จากสมการ  $c_j(n) = \langle f(t), \phi_{j,n}(t) \rangle$  และสมการที่ (3.21) ช่วยในการพิสูจน์ดังนี้

$$\begin{aligned} c_j(m) &= \langle f(t), \phi_{j,m}(t) \rangle \\ &= \left\langle f(t), \sum_n h_0(n-2m) \phi_{j+1,n}(t) \right\rangle \end{aligned}$$

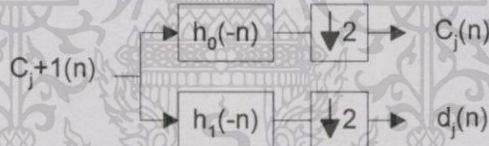
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
&= \sum_n h_0(n-2m) \langle f(t), \phi_{j+1,m}(t) \rangle \\
&= \sum_n h_0(n-2m) c_{j+1}(n) \\
&= h_0(-n) * c_{j+1}(n) |_{n=2m}
\end{aligned} \tag{3.23}$$

เพราะฉะนั้นสัมประสิทธิ์  $c_j(m)$  จะหาได้จากการคอนโวลูชันของสัมประสิทธิ์  $c_{j+1}(n)$  กับลำดับ  $h_0(-n)$  แล้วทำการลดอัตราการสุ่มลง (Down sampling) ด้วยสอง เพื่อให้เอาต์พุตมีจำนวนเท่ากับ  $m$  ในขณะเดียวกันสัมประสิทธิ์ของเวฟเล็ต  $d_j(m)$  ก็จะสามารถหาได้จากวิธีเดียวกันคือ

$$d_j(m) = h_1(-n) * c_{j+1}(n) |_{n=2m} \tag{3.24}$$

จากสมการที่ (3.23) และ (3.24) สามารถเขียนเป็นแผนภูมิในลักษณะของฟิลเตอร์แบบคี่แบบ 2 ของสัญญาณ สำหรับการแจกแจงสัมประสิทธิ์เวฟเล็ตได้ดังภาพที่ 3.17 ซึ่งจะเป็นการโปรเจกชันสัญญาณไปยังสเปซย่อยนั่นเอง

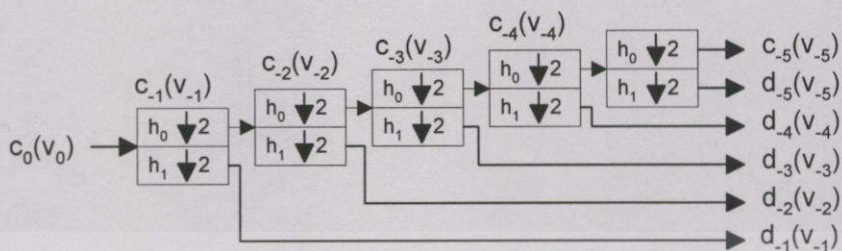


ภาพที่ 3.17 แสดงฟิลเตอร์แบบคี่แบบ 2 ของสัญญาณ ในการคำนวณหาสัมประสิทธิ์เวฟเล็ต

ดังนั้นจึงสรุปได้ว่า การแตกกระจายสัญญาณจะถูกกระทำด้วย Analysis filter banks ที่มี Impulse response เป็น  $h_0(-n)$  และ  $h_1(-n)$  โดยที่สัญญาณ  $f(t) \in V^{j+1}$  จะมีความละเอียดเป็นสองเท่าของสัญญาณในสเปซย่อย  $V^j, W^j$  และสัมประสิทธิ์  $c_j(m), d_j(m)$  จะมีอัตราการสุ่มลดลงครึ่งหนึ่งของ  $c_{j+1}(n)$  นั้นหมายความว่า จำนวนสัมประสิทธิ์ของ  $c_{j+1}(n)$  จะถูกแบ่งให้  $c_j(m), d_j(m)$  อย่างละครึ่ง

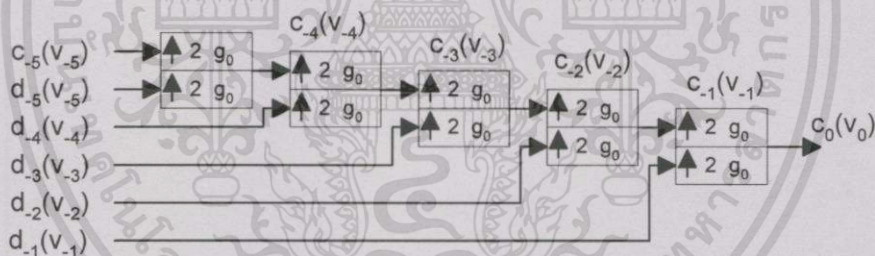
ถ้าพิจารณาคุณสมบัติการสเกลสอง (Two-scale property)  $f(t) \rightarrow f(2t)$  และจากลักษณะสมบัติของการเป็นโครงข่ายของสเปซย่อยๆ  $V^j$  จะพบว่าการโปรเจกชัน  $V^{j+1}$  เป็น  $V^j$  จะเป็นการลดความถี่ของสัญญาณลงครึ่งหนึ่ง จึงมีลักษณะเป็นตัวกรองความถี่ต่ำและการโปรเจกชันไปใน  $W^j$  จะเป็นส่วนของตัวกรองความถี่สูง ดังนั้น  $h_0(-n)$  และ  $h_1(-n)$  จึงเป็น Impulse response ของตัวกรองความถี่ต่ำและตัวกรองความถี่สูง ตามลำดับ

ในกรณีที่มีการทำ DWT ซ้ำในแนวของความถี่ต่ำ จะกลายเป็นลักษณะของ Dyadic Tree structure ซึ่งเหมือนกับขบวนการ Octave filter banks ดังแสดงในภาพที่ 3.18



ภาพที่ 3.18 แสดงลักษณะของ DWT โดยใช้ Dyadic tree structure (Octave filter bank)

สัมประสิทธิ์เวฟเล็ต  $d_{-1}$  ถึง  $d_{-5}$  จะเปรียบเสมือน Discrete-time bandpass signal หมายถึงสัญญาณในส่วนที่ผ่าน Basspass filter ที่แบนด์ความถี่ต่างๆ ในขณะที่  $c_{-5}$  ของฟังก์ชันสเกลลิงจะเป็น Discrete-time lowpass signal ซึ่งดูได้จากภาพที่ 3.16 และจากโครงสร้างในภาพที่ 3.18 เรียกว่า "การแปลงเวฟเล็ตแบบแบ่งช่วง" (Discrete Wavelets Transform: DWT)



ภาพที่ 3.19 แสดงลักษณะของ IWT โดยการใช้ Dyadic tree structure

ในภาพที่ 3.19 แสดงขบวนการ "การแปลงกลับเวฟเล็ต" (Inverse Wavelets Transform: IWT) จะสังเกตเห็นว่า Upsampling digital filter  $g_0(n)$  และ  $g_1(n)$  จะกระทำกับสัมประสิทธิ์  $c_j(m)$  และ  $d_j(m)$  ในกรณีที่เลขคู่เท่านั้น เช่น สัมประสิทธิ์  $c_j(0)$  จะกระทำกับ  $g_0(n)$  และ  $c_j(1)$  จะกระทำกับ  $g_0(n-2)$  เป็นต้น

จากเนื้อหาทั้งหมดที่กล่าวมาทำให้ทราบถึงคุณสมบัติของ การแปลงเวฟเล็ต ความแตกต่างระหว่างการแปลงเวฟเล็ตกับการแปลงข้อมูลแบบอื่นๆ และความเหมาะสมของงานที่จะนำการแปลงเวฟเล็ต ชนิดต่างๆ ไปประยุกต์ใช้

ในวิทยานิพนธ์ฉบับนี้จะใช้การแปลงเวฟเล็ตในการประมวลผลภาพแบบ DWT โดยใช้หลักการแยกองค์ประกอบแบนด์ย่อย (subband decomposition) ของภาพ ซึ่งเป็นเทคนิคของ Mallat [17] โดยภาพต้นแบบจะถูกแยกเป็นองค์ประกอบความถี่สูงและองค์ประกอบความถี่ต่ำ ดังรายละเอียดต่อไปนี้

การแปลงเวฟเล็ตในการประมวลผลภาพสามารถแบ่งออกได้เป็นสองขั้นตอนคือ

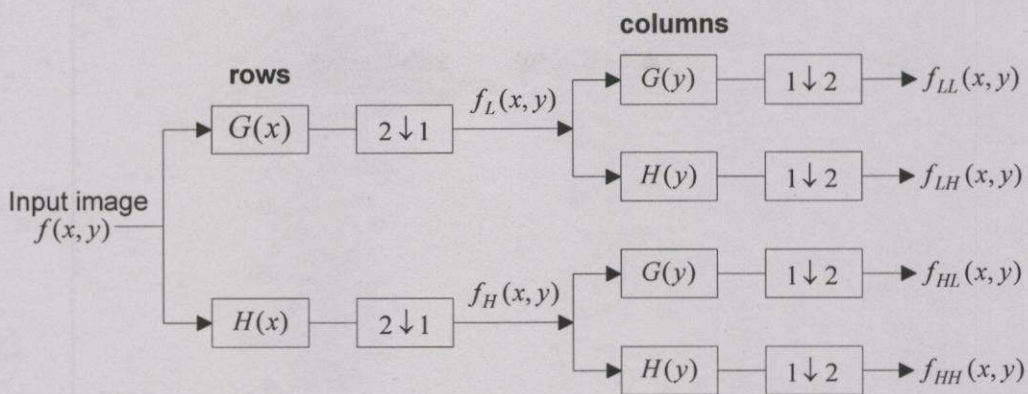
### 1) Wavelet Decomposition

เป็นการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ซึ่งแสดงได้ดังแผนภาพในภาพที่ 3.20 โดยเริ่มต้นจากภาพต้นแบบ  $f(x, y)$  จะถูกกรองไปตามแนวแถว (row) ด้วยตัวกรองผ่านความถี่ต่ำ  $G$  และตัวกรองผ่านความถี่สูง  $H$  จากนั้นจะนำผลที่ได้มาทำการลดจำนวนการสุ่ม (downsampling) ลง 2 เท่า โดยการเก็บข้อมูลมา 1 หลัก (column) จากข้อมูล 2 หลัก ซึ่งจะทำให้ได้เอาท์พุทความถี่สูงและความถี่ต่ำ จากนั้นเอาท์พุททั้งสองจะถูกกรองตามแนวหลัก และลดจำนวนการสุ่มลง 2 เท่า โดยการเก็บข้อมูล 1 แถวจาก 2 แถว ผลลัพธ์ที่ได้จะประกอบด้วยภาพแบนด์ย่อยความถี่ต่ำ  $f_{LL}(x, y)$  จำนวน 1 ภาพ และภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพ คือ  $f_{LH}(x, y), f_{HL}(x, y)$  และ  $f_{HH}(x, y)$  ตามลำดับ โดยแต่ละภาพจะมีขนาดลดลงเป็น  $1/4$  เท่าของภาพต้นแบบ

### 2) Wavelet Reconstruction

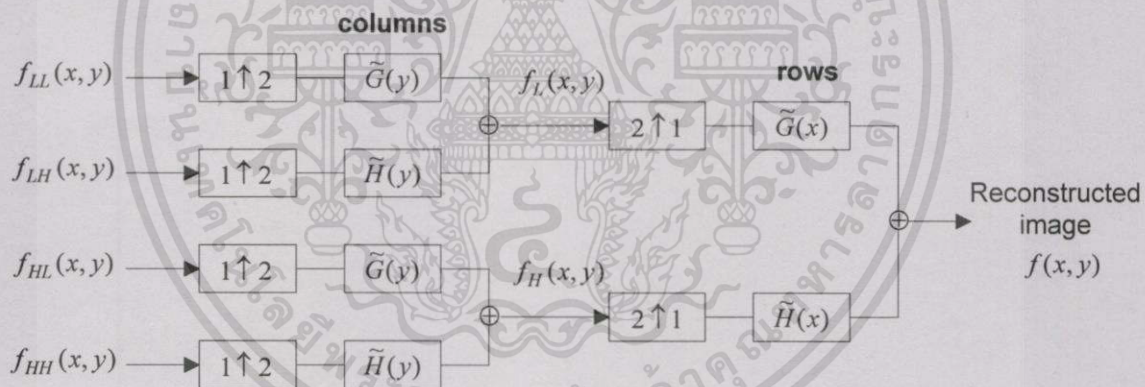
เป็นการสร้างภาพขึ้นมาใหม่จากองค์ประกอบย่อย ซึ่งแสดงได้ดังแผนภาพในภาพที่ 3.21 โดยเริ่มจากภาพแบนด์ย่อยทั้ง 4 จะถูกเพิ่มจำนวนการสุ่ม (upsampling) เป็น 2 เท่า โดยการเติมศูนย์เข้าไปในระหว่างแต่ละคู่ของแถว ภายหลังจากการเพิ่มจำนวนการสุ่มแล้วผลลัพธ์ที่ได้จะถูกกรองไปตามหลักและบวกเข้าด้วยกัน จากนั้นจะทำกระบวนการซ้ำไปตามแถว ซึ่งผลลัพธ์ที่ได้จะเป็นภาพที่ถูกสร้างกลับขึ้นมาใหม่

ในการแปลงเวฟเล็ตสำหรับภาพ 1 ครั้ง (Stage) จะทำให้ได้ภาพที่มีขนาดเท่ากัน 4 ภาพ คือเป็นภาพที่มีข้อมูลส่วนรายละเอียด (ข้อมูลความถี่สูง) จำนวน 3 ภาพ และเป็นภาพที่มีระดับความสว่าง (ข้อมูลความถี่ต่ำ) จำนวน 1 ภาพ ซึ่งขั้นตอนการแปลงเวฟเล็ตนี้สามารถกระทำกับภาพข้อมูลความถี่ต่ำได้ต่อไปเรื่อยๆ ดังแสดงในภาพที่ 3.22



- $G$  คือ การทำคอนโวลูชันตามแถวหรือหลักด้วยตัวกรอง  $G$
- $H$  คือ การทำคอนโวลูชันตามแถวหรือหลักด้วยตัวกรอง  $H$
- $2 \downarrow 1$  คือ การเก็บ 1 หลักออกจาก 2 หลัก
- $1 \downarrow 2$  คือ การเก็บ 1 แถวออกจาก 2 แถว

ภาพที่ 3.20 แสดงแผนภาพ Image Wavelet Decomposition



- $\tilde{G}$  คือ การทำคอนโวลูชันตามแถวหรือหลักด้วยตัวกรอง  $\tilde{G}$
- $\tilde{H}$  คือ การทำคอนโวลูชันตามแถวหรือหลักด้วยตัวกรอง  $\tilde{H}$
- $2 \uparrow 1$  คือ การใส่ศูนย์เข้าไปในระหว่างหลัก
- $1 \uparrow 2$  คือ การใส่ศูนย์เข้าไปในระหว่างแถว

ภาพที่ 3.21 แสดงแผนภาพ Image Wavelet Reconstruction

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$f_{LL1}$	$f_{LH1}$
$f_{HL1}$	$f_{HH1}$

ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต 1 ครั้ง

$f_{LL2}$	$f_{LH2}$	$f_{LH1}$
$f_{HL2}$	$f_{HH2}$	
$f_{HL1}$	$f_{HH1}$	

ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต 2 ครั้ง

$f_{LL3}$	$f_{LH3}$	$f_{LH2}$	$f_{LH1}$
$f_{HL3}$	$f_{HH3}$		
$f_{HL2}$	$f_{HH2}$		
$f_{HL1}$	$f_{HH1}$		

ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต 3 ครั้ง

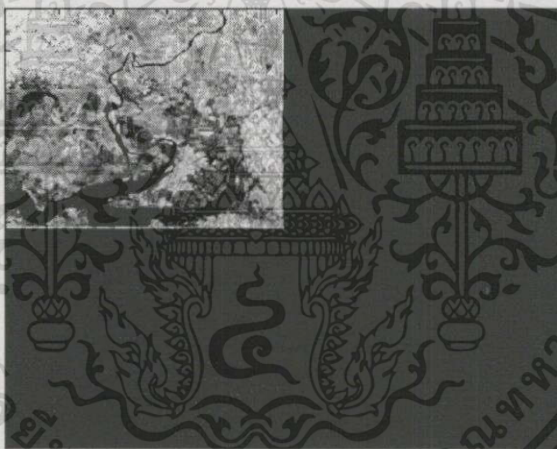
ภาพที่ 3.22 แสดงภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ตหลายครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่าง แสดงภาพการแปลงเวฟเล็ตกับภาพขนาด 512x512 และใช้สัมประสิทธิ์ของ Daubechies order 8



ภาพที่ 3.23 แสดงภาพต้นแบบภาพถ่ายจากดาวเทียม LANDSAT ขนาด 512x512x8 บิต

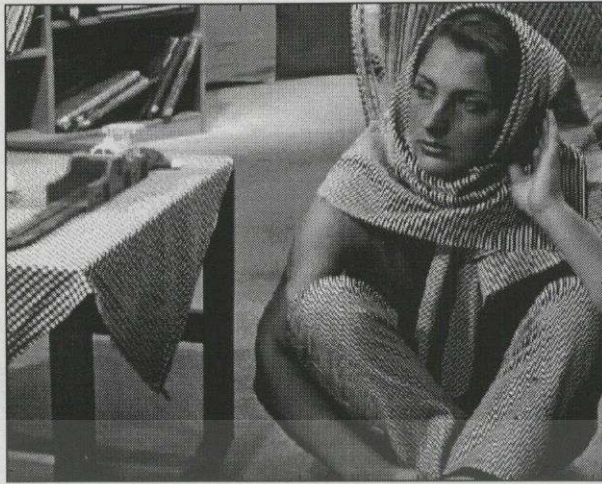


ภาพที่ 3.24 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตระดับที่ 1

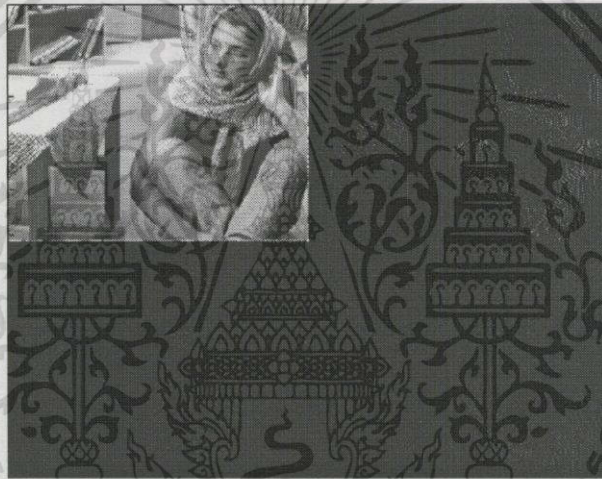


ภาพที่ 3.25 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตระดับที่ 2

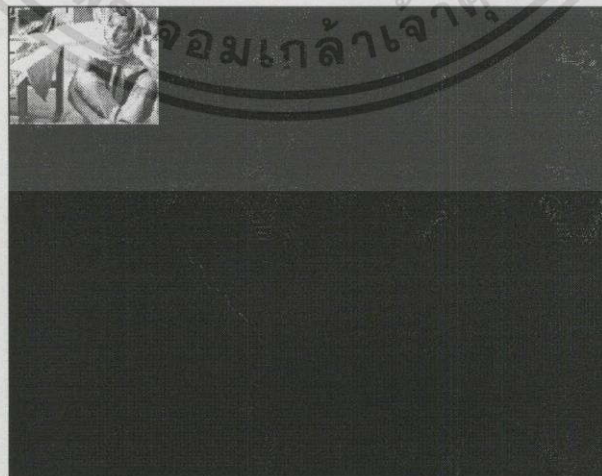
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.26 แสดงภาพถ่ายต้นแบบ Barbara ขนาด 512x512x8 บิต



ภาพที่ 3.27 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตระดับที่ 1



ภาพที่ 3.28 แสดงผลลัพธ์ที่ได้จากการแปลงเวฟเล็ตระดับที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

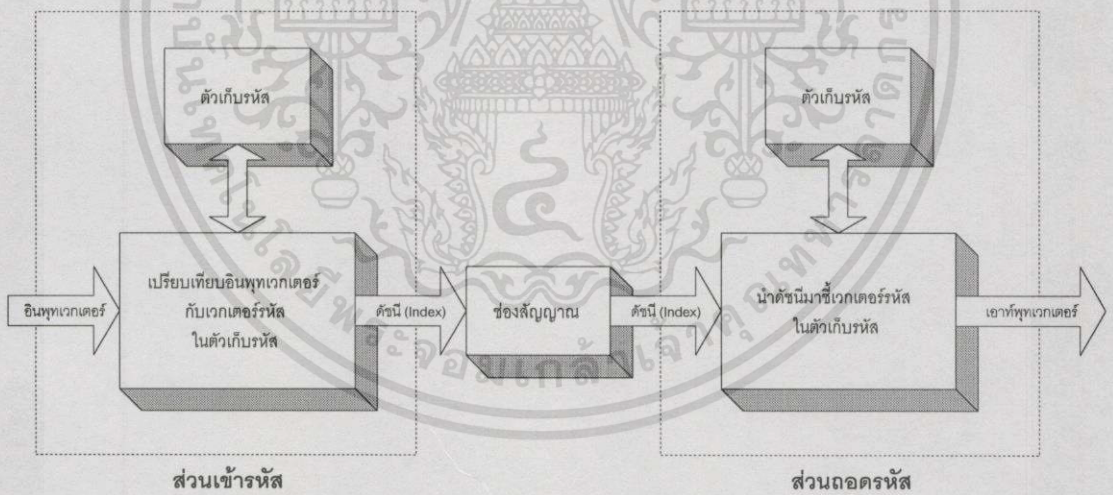
# การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์

การบีบอัดข้อมูลภาพด้วยวิธีการควอนไทซ์แบบเวกเตอร์ (Vector Quantization, VQ) เป็นวิธีการลดข้อมูลแบบที่มีการสูญเสียวิธีหนึ่ง ซึ่งนิยามของการควอนไทซ์แบบเวกเตอร์ คือการสร้างความสัมพันธ์ ( $Q$ ) ระหว่างข้อมูลของสเปซยูคิเดียน  $K$  มิติ ( $R^k$ ) ลงในเซตย่อย  $Y$  ของ  $R^k$  [1]

$$Q: R^k \rightarrow Y \quad (4.1)$$

ซึ่ง  $Y = (\hat{x}_i; i = 1, 2, \dots, N)$  คือ เซตของเวกเตอร์ที่ถูกสร้างกลับขึ้นมาใหม่ และ  $N$  คือจำนวนของเวกเตอร์ใน  $Y$

การควอนไทซ์แบบเวกเตอร์ประกอบด้วย 2 ส่วนคือ ส่วนเข้ารหัส และส่วนถอดรหัส ดังแสดงในภาพที่ 4.1



ภาพที่ 4.1 แสดงบล็อกไดอะแกรมของการควอนไทซ์แบบเวกเตอร์ ทั้งส่วนเข้ารหัสและส่วนถอดรหัส

ส่วนเข้ารหัสจะทำหน้าที่เปรียบเทียบอินพุตเวกเตอร์ที่ต้องการจะเข้ารหัสกับเวกเตอร์รหัส (codevector) ทั้งหมดที่เก็บอยู่ในตัวเก็บรหัส (codebook) โดยกลุ่มของเวกเตอร์รหัสเหล่านี้จะแทนเวกเตอร์ที่เป็นไปได้ทั้งหมดของอินพุตเวกเตอร์ ซึ่งการเปรียบเทียบดังกล่าวจะได้ผลลัพธ์คือดัชนี (index) แทนเวกเตอร์รหัสที่มีความพี้ยนน้อยที่สุดกับอินพุตเวกเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนถอดรหัสจะทำหน้าที่นำเอาค่าดัชนีมาชี้เวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัส ซึ่งจะต้องเป็นตัวเก็บรหัสเดียวกันกับตอนเข้ารหัส โดยเวกเตอร์รหัสที่ได้จะเป็นเอาต์พุตเวกเตอร์

#### 4.1 การควอนไทซ์แบบเวกเตอร์สำหรับภาพเชิงตัวเลข

การควอนไทซ์แบบเวกเตอร์สำหรับภาพเชิงตัวเลขเป็นวิธีการในการบีบอัดข้อมูลภาพวิธีการหนึ่ง โดยอาศัยหลักการแทนที่บล็อกย่อยของข้อมูลภาพต้นแบบด้วยเลขดัชนีของชุดข้อมูลที่ใช้แทนบล็อกย่อยนั้น ซึ่งมักจะใช้ดัชนีเป็นเลขจำนวนเต็ม และมีค่าจำกัดอยู่ในระดับหนึ่ง

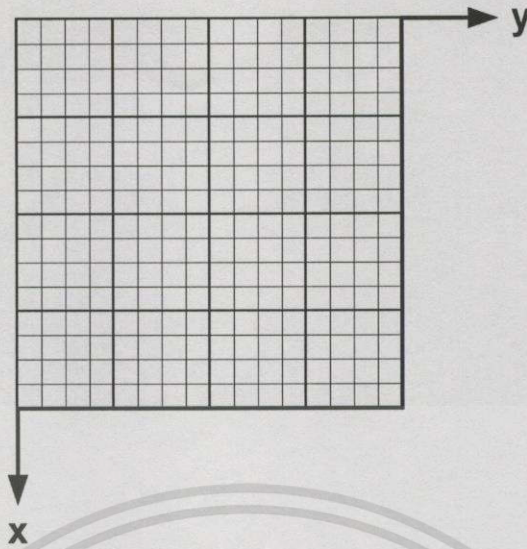
##### 4.1.1 การเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

ในการเข้ารหัสภาพเชิงตัวเลขด้วยการควอนไทซ์แบบเวกเตอร์ ภาพต้นแบบจะถูกแยกองค์ประกอบให้เป็นเวกเตอร์ของภาพขนาด  $K$  มิติ ซึ่งเวกเตอร์เหล่านี้ สามารถสร้างได้ด้วยวิธีการต่างๆ หลายวิธี เช่น แบ่งภาพให้เป็นบล็อกของจุดภาพขนาด  $l \times m = K$  ซึ่งจะทำให้การจัดเรียงให้อยู่ในรูปเวกเตอร์ขนาด  $K$  มิติได้

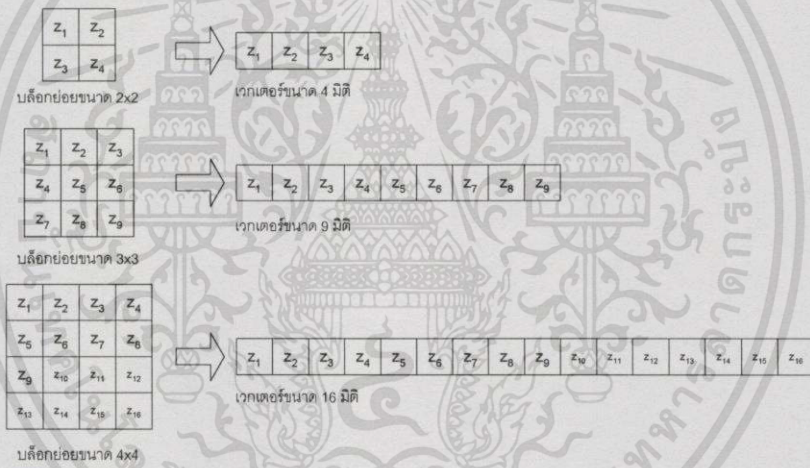
โดยทั่วไปอินพุตเวกเตอร์ที่จะนำมาทำการควอนไทซ์แบบเวกเตอร์จะได้มาจากการแบ่งออกเป็นบล็อกย่อยๆ ของจุดภาพที่ไม่ซ้อนทับกัน ซึ่งมักจะแบ่งเป็นบล็อกสี่เหลี่ยมจัตุรัสขนาดเท่าๆ กัน เนื่องจากมีความง่ายในการจัดแบ่งให้ครอบคลุมได้ทั้งภาพ ดังตัวอย่างในภาพที่ 4.2 จากนั้นจะทำการจัดเรียงข้อมูลจุดภาพในแต่ละบล็อกย่อยให้อยู่ในรูปของเวกเตอร์ ดังตัวอย่างในภาพที่ 4.3

ดังนั้นจะพบว่าถ้าภาพต้นแบบมีขนาดเป็น  $256 \times 256$  จุดภาพ จะสามารถทำการแบ่งเป็นบล็อกย่อยขนาด  $2 \times 2$  ที่ไม่ซ้อนทับกันได้ทั้งหมดจำนวน 16,384 บล็อก และถ้าแบ่งเป็นบล็อกย่อยขนาด  $4 \times 4$  ที่ไม่ซ้อนทับกันก็จะได้ทั้งหมดจำนวน 4,096 บล็อก เป็นต้น ซึ่งจำนวนบล็อกย่อยที่แบ่งได้พอดีนี้จะเท่ากับจำนวนอินพุตเวกเตอร์ที่จะนำมาทำการควอนไทซ์แบบเวกเตอร์ต่อไป

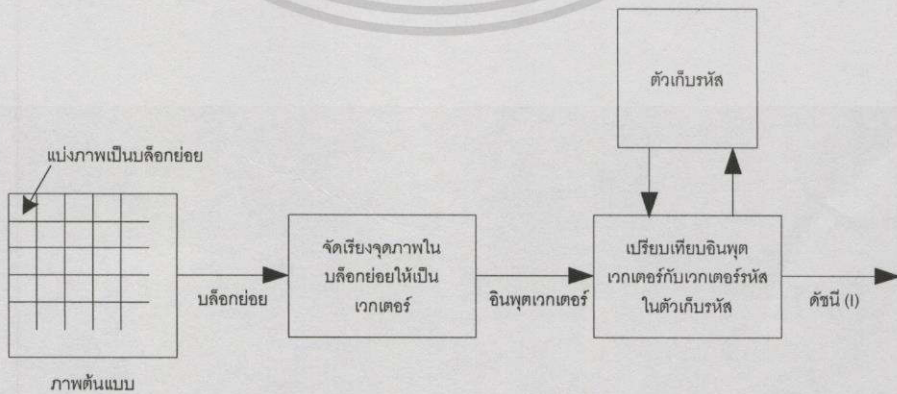
ส่วนเข้ารหัสจะทำการเปรียบเทียบอินพุตเวกเตอร์ที่ได้มาจากภาพต้นแบบกับเวกเตอร์รหัสทุกตัวที่เก็บอยู่ในตัวเก็บรหัส โดยตัวเก็บรหัสจะประกอบด้วยกลุ่มของเวกเตอร์รหัสที่แทนเวกเตอร์ที่เป็นไปได้ทั้งหมดของอินพุตเวกเตอร์ ในการเปรียบเทียบหาเวกเตอร์รหัสที่มีความพี้ยนจากอินพุตเวกเตอร์น้อยที่สุด ผลลัพธ์ที่ได้จากการเปรียบเทียบนี้คือเลขดัชนีที่ใช้แทนเวกเตอร์รหัส ที่มีความพี้ยนน้อยที่สุด ดังภาพที่ 4.4



ภาพที่ 4.2 แสดงการแบ่งภาพเป็นบล็อกย่อยขนาด 4x4



ภาพที่ 4.3 แสดงตัวอย่างการจัดเรียงจุดภาพของบล็อกย่อยขนาดต่างๆ ให้อยู่ในรูปของเวกเตอร์



ภาพที่ 4.4 แสดงส่วนเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

เวกเตอร์แต่ละเวกเตอร์ของภาพ  $X$  จะถูกนำไปเปรียบเทียบกับชุดข้อมูลที่จะใช้แทนหรือเรียกว่า เวกเตอร์รหัส (codevector)  $(\hat{X}_i, i=1,2,3,\dots,N_c)$  เมื่อ  $N_c$  คือจำนวนของเวกเตอร์รหัส ซึ่งจะได้มาจากตัวเก็บรหัสที่ได้สร้างเอาไว้ก่อนแล้ว มีโครงสร้างดังแสดงในภาพที่ 4.5 โดยเวกเตอร์รหัสจะมีขนาดเป็น  $K$  มิติด้วยเหมือนกัน ในการเปรียบเทียบจะทำการเลือกเวกเตอร์รหัสที่เหมาะสมที่สุด โดยใช้กฎความเพี้ยนที่น้อยที่สุด คือจะทำการเลือกเวกเตอร์รหัส  $\hat{X}_k$  ที่ทำให้  $d(X, \hat{X}_k) \leq d(X, \hat{X}_j)$  สำหรับ  $j=1,2,3,\dots,N_c$  เมื่อ  $d(X, \hat{X}_k)$  คือความเพี้ยนที่เกิดขึ้นจากการแทนเวกเตอร์ต้นแบบ  $X$  ด้วยเวกเตอร์รหัส  $\hat{X}_k$

การหาความเพี้ยนนี้หาได้จาก การหาค่าความผิดพลาดกำลังสองเฉลี่ย (MSE) ของ อินพุตเวกเตอร์กับเวกเตอร์รหัส ดังสมการต่อไปนี้

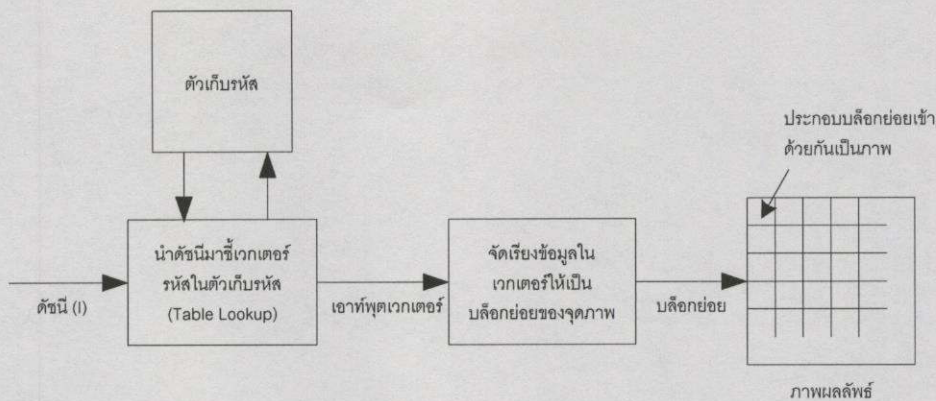
$$d(x, \hat{x}_i) = \frac{1}{K} \sum_{m=1}^K [x(m) - \hat{x}_i(m)]^2 \quad (4.2)$$

Index (i)	codevector $(\hat{x}_i)$				
$i=1$	$\hat{x}_1(1)$	$\hat{x}_1(2)$	$\hat{x}_1(3)$	...	$\hat{x}_1(K)$
$i=2$	$\hat{x}_2(1)$	$\hat{x}_2(2)$	$\hat{x}_2(3)$	...	$\hat{x}_2(K)$
$i=3$	$\hat{x}_3(1)$	$\hat{x}_3(2)$	$\hat{x}_3(3)$	...	$\hat{x}_3(K)$
$\vdots$					
$i$	$\hat{x}_i(1)$	$\hat{x}_i(2)$	$\hat{x}_i(3)$	...	$\hat{x}_i(K)$
$\vdots$					
$i=N_c$	$\hat{x}_{N_c}(1)$	$\hat{x}_{N_c}(2)$	$\hat{x}_{N_c}(3)$	...	$\hat{x}_{N_c}(K)$

ภาพที่ 4.5 แสดงโครงสร้างของตัวเก็บรหัส

#### 4.1.2 การถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

เป็นการนำเอาค่าดัชนีมาซึ่งข้อมูลเวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัสเดียวกันกับตอนเข้ารหัส โดยผลลัพธ์ที่ได้จะเป็นเอาท์พุตเวกเตอร์ จากนั้นนำเอาเอาท์พุตเวกเตอร์ที่ได้มาจัดเรียงใหม่ให้เป็นบล็อกย่อยของจุดภาพ ดังภาพที่ 4.6



ภาพที่ 4.6 แสดงส่วนถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

## 4.2 การหาอัตราบิต (Bit rate) ของการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์

ถ้าตัวเก็บรหัสประกอบด้วยเวกเตอร์รหัสจำนวน  $N_c$  ระดับ จะได้อัตราบิตต่อเวกเตอร์หรือจำนวนบิตของดัชนี

$$I = \log_2 N_c \quad (4.3)$$

และถ้าเวกเตอร์รหัสมีขนาดเป็น  $K$  มิติ หรือกล่าวได้ว่าประกอบด้วยสมาชิก  $K$  จุดภาพ จะได้อัตราบิตต่อจุดภาพ (Bits per pixel, BPP) ดังนี้

$$BPP = \frac{I}{K} = \frac{1}{K} \log_2 N_c \quad (4.4)$$

ดังนั้นจะเห็นได้ว่าอัตราบิตขึ้นอยู่กับตัวเก็บรหัส โดยที่จำนวนของเวกเตอร์รหัสที่อยู่ในตัวเก็บรหัสจะแปรผันตรงกันกับอัตราบิต คือถ้าตัวเก็บรหัสประกอบด้วยเวกเตอร์รหัสจำนวนมากขึ้น จะทำให้มีอัตราบิตที่สูงขึ้น ส่วนขนาดของเวกเตอร์รหัสที่อยู่ในตัวเก็บรหัส จะแปรผกผันกับอัตราบิต คือ ถ้าเวกเตอร์รหัสมีขนาดเพิ่มขึ้นจะทำให้มีอัตราบิตที่ลดลง

## 4.3 การหาอัตราการบีบอัดข้อมูล (Compression rate) ของการลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์

การหาอัตราการบีบอัดข้อมูลสามารถหาได้จากอัตราส่วนระหว่างจำนวนบิตข้อมูลจุดภาพทั้งหมดใน 1 บล็อกย่อยของภาพต้นแบบต่อจำนวนบิตข้อมูลที่ใช้แทน 1 บล็อกย่อย ดังนั้น

สำหรับภาพต้นแบบขนาด 8 บิตต่อจุดภาพ สามารถทำการคำนวณหาอัตราการบีบอัดข้อมูลได้ดังสมการต่อไปนี้

$$\text{อัตราการบีบอัดข้อมูล} = \frac{\text{จำนวนบิตข้อมูลจุดภาพทั้งหมดใน 1 บล็อกย่อย}}{\text{จำนวนบิตข้อมูลที่ใช้แทน 1 บล็อกย่อย}} \quad (4.5)$$

$$= \frac{K \times 8}{I} = \frac{K \times 8}{\log_2 N_c} = \frac{8}{BPP} \quad (4.6)$$

ตารางที่ 4.1 แสดงอัตราบิตต่อจุดภาพของการบีบอัดข้อมูลด้วยการควอนไทซ์แบบเวกเตอร์ที่มีตัวเก็บรหัสและบล็อกย่อยขนาดต่างๆ

ระดับของตัวเก็บรหัส $N_c$	จำนวนบิตของดัชนี $(I)$	อัตราบิตต่อจุดภาพ (Bit Per Pixel, BPP)		
		ขนาดบล็อกย่อย 2x2 $K = 4$	ขนาดบล็อกย่อย 4x4 $K = 16$	ขนาดบล็อกย่อย 8x8 $K = 64$
256	8	2	0.5	0.125
128	7	1.75	0.4375	0.1094
64	6	1.5	0.375	0.09375
32	5	1.25	0.3125	0.078125
16	4	1	0.25	0.0625
8	3	0.75	0.1875	0.046875
4	2	0.5	0.125	0.03125
2	1	0.25	0.0625	0.015625

ตารางที่ 4.2 แสดงอัตราการบีบอัดข้อมูลของการลดข้อมูลด้วยการควอนไทซ์แบบเวกเตอร์ที่มีตัวเก็บรหัสและบล็อกย่อยขนาดต่างๆ

ระดับของตัวเก็บรหัส $N_c$	จำนวนบิตของดัชนี ( $J$ )	อัตราการบีบอัดข้อมูล (Compression Rate)		
		ขนาดบล็อกย่อย 2x2 $K = 4$	ขนาดบล็อกย่อย 4x4 $K = 16$	ขนาดบล็อกย่อย 8x8 $K = 64$
256	8	4	16	64
128	7	4.571	18.286	73.143
64	6	5.333	21.333	85.333
32	5	6.4	25.6	102.4
16	4	8	32	128
8	3	10.667	42.667	170.667
4	2	16	64	256
2	1	32	128	512

#### 4.4 การสร้างตัวเก็บรหัส (Codebook)

การสร้างตัวเก็บรหัสแบ่งได้เป็น 2 กลุ่ม คือ

##### 4.4.1 การสร้างตัวเก็บรหัสด้วยโมเดลทางคณิตศาสตร์

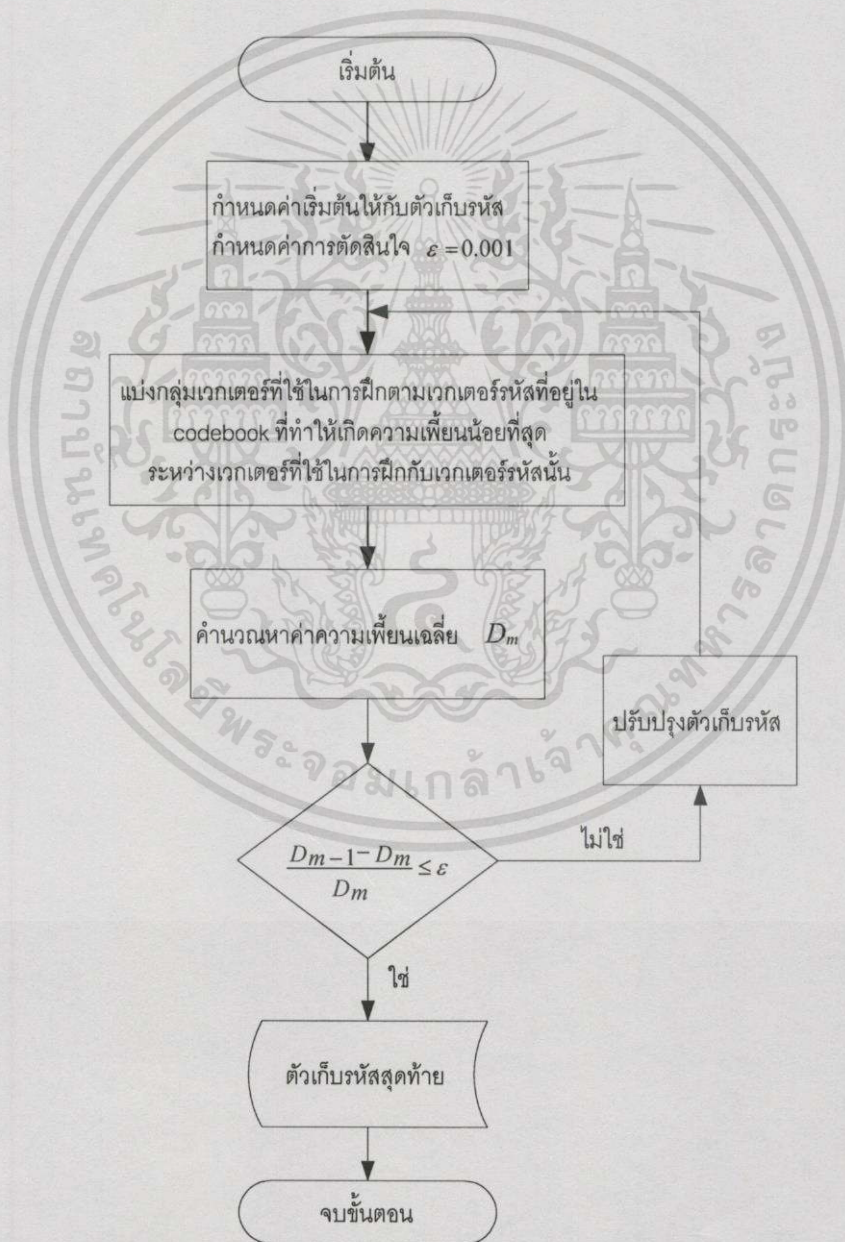
วิธีนี้จะได้ตัวควอนไทซ์ที่เป็นแบบยูนิฟอร์ม (Uniform) โดยการกำหนดโครงสร้างที่เป็นลักษณะปกติทั่วไปของภาพให้กับตัวเก็บรหัสเพื่อทำให้การเข้ารหัสและการถอดรหัสสามารถกระทำได้รวดเร็ว โดยปกติจำเป็นต้องรู้โมเดลความน่าจะเป็นของเวกเตอร์ของภาพ และต้องการตัวเก็บรหัสที่มีขนาดใหญ่ เพื่อให้ได้ประสิทธิภาพที่ได้

##### 4.4.2 การสร้างตัวเก็บรหัสโดยการหาตัวเก็บรหัสที่ดีที่สุดจากการเรียนรู้จากชุดข้อมูลที่นำมาทำการฝึก (Training set)

วิธีนี้มีความเหมาะสมในการใช้งานตามความเป็นจริงมากกว่า โดยเหมาะสำหรับตัวเก็บรหัสที่มีขนาดเล็ก โดยไม่จำเป็นต้องรู้โมเดลความน่าจะเป็นของเวกเตอร์ของภาพ ซึ่งวิธีการที่นิยมใช้กันได้แก่ การสร้างตัวเก็บรหัสโดยใช้อัลกอริทึมของ Linde-Buzo-Gray หรือ LBG

#### 4.4.3 อัลกอริทึม Linde-Buzo-Gray (LBG)

อัลกอริทึม LBG [14,15] เป็นอัลกอริทึมที่นิยมใช้ในการสร้างตัวเก็บรหัส โดยอัลกอริทึมจะเริ่มต้นจากการกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส แล้วทำการปรับปรุงตัวเก็บรหัสโดยการแทนเวกเตอร์รหัสแต่ละตัวที่อยู่ในตัวเก็บรหัสด้วยเวกเตอร์ตัวใหม่ที่จะทำให้เกิดความเพี้ยนลดลงไปเรื่อยๆ โดยจะทำการปรับปรุงวนซ้ำไปจนกว่าอัตราการเปลี่ยนแปลงของความเพี้ยนที่ลดลงมีค่าต่ำกว่าขอบเขตที่กำหนดไว้ อัลกอริทึมของ LBG สามารถเขียนเป็นโฟลว์ชาร์ตได้ดังภาพที่ 4.7



ภาพที่ 4.7 แสดงโฟลว์ชาร์ตของอัลกอริทึม LBG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับขั้นตอนการทำงานของอัลกอริทึมสามารถอธิบายโดยละเอียดได้ดังต่อไปนี้

**ขั้นตอนที่ 1** นำภาพที่จะทำการฝึกแบ่งเป็นบล็อกย่อยๆ แล้วจัดให้อยู่ในรูปเวกเตอร์สำหรับการฝึก (Training vectors)  $x_j$  โดยที่  $j = 0, 2, \dots, n-1$  เมื่อ  $n$  คือจำนวนบล็อกย่อยหรือจำนวนเวกเตอร์ที่ได้มาจากภาพที่จะนำมาทำการฝึก

**ขั้นตอนที่ 2** ทำการกำหนดค่าเริ่มต้นให้กับอัลกอริทึมดังนี้

- 2.1) กำหนด  $m$  เป็นจำนวนรอบของการคำนวณ = 0
- 2.2) กำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส  $Y_0$
- 2.3) กำหนด  $D_m$  เป็นตัวแปรวัดความเพี้ยนเฉลี่ย (average distortion) ของเวกเตอร์ทั้งหมดให้มีความมากๆ
- 2.4) กำหนด  $\varepsilon$  เป็นค่าตัดสินใจของอัตราการเปลี่ยนแปลงค่าความเพี้ยนเฉลี่ย = 0.001

**ขั้นตอนที่ 3** ทำการจัดกลุ่มของเวกเตอร์ที่นำมาฝึกตามจำนวนของเวกเตอร์รหัสที่อยู่ในตัวเก็บรหัส

$$P(Y_m) = (S_i; i = 1, 2, \dots, N_c) \quad (4.7)$$

เมื่อ  $Y_m$  คือตัวเก็บรหัสของรอบการคำนวณที่  $m$  และ  $S_i$  คือเซตของเวกเตอร์ที่มีความเพี้ยนน้อยที่สุดกับเวกเตอร์รหัส  $\hat{x}_i$

โดยการจัดกลุ่มจะใช้หลักความเพี้ยนน้อยที่สุดดังนี้ คือ

$$x_j \in S_k \text{ ถ้า } d(x_j, \hat{x}_k) \leq d(x_j, \hat{x}_i) \text{ สำหรับทุกค่าของ } i$$

เมื่อ

$$d(x, \hat{x}_i) = \frac{1}{K} \sum_{m=1}^K [x(m) - \hat{x}_i(m)]^2 \quad (4.8)$$

**ขั้นตอนที่ 4** ทำการคำนวณหาค่าความเพี้ยนเฉลี่ย (Average distortion)  $D_m$  โดยที่

$$D_m = D[(Y_m, P(Y_m))] = \frac{1}{n} \sum_{j=0}^{n-1} \text{mind}(x_j, \hat{x}) \quad (4.9)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 5 ทำการคำนวณหาอัตราการเปลี่ยนแปลงของค่าความเพี้ยนเฉลี่ยของรอบการคำนวณที่แล้วกับรอบการคำนวณปัจจุบัน ซึ่งสามารถคำนวณได้จาก  $\frac{D_{m-1} - D_m}{D_m}$  แล้วนำมาเปรียบเทียบกับค่าการตัดสินใจ  $\varepsilon$

- 5.1) ถ้า  $\frac{D_{m-1} - D_m}{D_m}$  มีค่ามากกว่าค่าการตัดสินใจ  $\varepsilon$  ให้ทำการปรับปรุค่าของเวกเตอร์รหัสในตัวเก็บรหัส โดยการนำแต่ละกลุ่มของเวกเตอร์ที่จัดไว้ในขั้นตอนที่ 3 มาทำการหาค่าเฉลี่ยในแต่ละมิติของเวกเตอร์ ดังนั้นจะได้ตัวเก็บรหัสที่ปรับปรุแล้วเป็น

$$\hat{X}(P(Y_m)) = (\hat{X}(S_i); i = 1, 2, \dots, N_c) \quad (4.10)$$

เมื่อ

$$\hat{X}(S_i) = \frac{1}{\|S_i\|} \sum_{j: x_j \in S_i} x_j \quad (4.11)$$

ดังนั้นจะได้ตัวเก็บรหัสใหม่เป็น

$$Y_{m+1} = \hat{X}(S_i) \quad (4.12)$$

จากนั้นทำการเพิ่มค่าให้กับตัวจำนวนรอบของการคำนวณ  $m$  ขึ้นอีก 1 ( $m = m + 1$ ) แล้วกลับไปทำซ้ำเริ่มจากขั้นตอนที่ 3

- 5.2) ถ้า  $\frac{D_{m-1} - D_m}{D_m}$  น้อยกว่าหรือเท่ากับ  $\varepsilon$  จะเป็นการจบอัลกอริทึม และได้ตัวเก็บรหัสคือ  $Y_m$

#### 4.4.4 การกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส

การกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัสจะมีผลต่อการทำงานของอัลกอริทึม การกำหนดค่าเริ่มต้นดีจะทำให้ใช้เวลาในการคำนวณน้อย และได้ตัวเก็บรหัสที่ดี ในการกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส สามารถกำหนดให้กับตัวเก็บรหัสตามขนาดที่ต้องการ หรือสามารถกำหนดให้กับตัวเก็บรหัสขนาดเล็กๆ ก่อนที่จะทำการคำนวณแบบเรียกตัวเอง (Recursive) เพื่อสร้างตัวเก็บรหัสที่ใหญ่ขึ้นจนได้ขนาดตามที่ต้องการ

การกำหนดค่าเริ่มต้นของตัวเก็บรหัสให้กับอัลกอริทึมนั้นมีอยู่หลายวิธีคือ

1) การกำหนดค่าเริ่มต้นด้วย  $N$  เวกเตอร์แรกของลำดับการฝึกหัด

วิธีการนี้เป็นวิธีการที่ง่าย แต่มีข้อเสียคือวิธีการนี้อาจจะไม่ได้เวกเตอร์ที่มีการกระจายที่ดีพอ การกำหนดค่าเริ่มต้นด้วยวิธีนี้อาจเรียกได้อีกอย่างหนึ่งว่าเป็นการกำหนดค่าเริ่มต้นแบบ Random

2) การกำหนดค่าเริ่มต้นด้วยยูนิฟอร์มควอนไทซ์เซอร์ (Uniform quantizer) [14]

เป็นการกำหนดค่าเริ่มต้นให้ครอบคลุมข้อมูลทั้งหมดที่จะนำมาทำการฝึก ซึ่งวิธีนี้จะใช้กับข้อมูลที่มีการกระจายแบบเกาส์เซียน (Gaussian)

3) การกำหนดค่าเริ่มต้นโดยการแตกตัว (Splitting)

วิธีการนี้จะมีประโยชน์ในการออกแบบตัวควอนไทซ์ที่มีอัตราสูงขึ้นอย่างต่อเนื่อง จนกระทั่งได้ระดับความแม่นยำที่ต้องการ เราสามารถกำหนดตัวควอนไทซ์  $M$  ระดับด้วย  $M = 2^R$  เมื่อ  $R = 0, 1, \dots$ , จนกระทั่งได้ค่าเริ่มต้นสำหรับตัวควอนไทซ์  $N_c$  ระดับ

การกำหนดค่าเริ่มต้นโดยการแตกตัวมีขั้นตอนดังต่อไปนี้

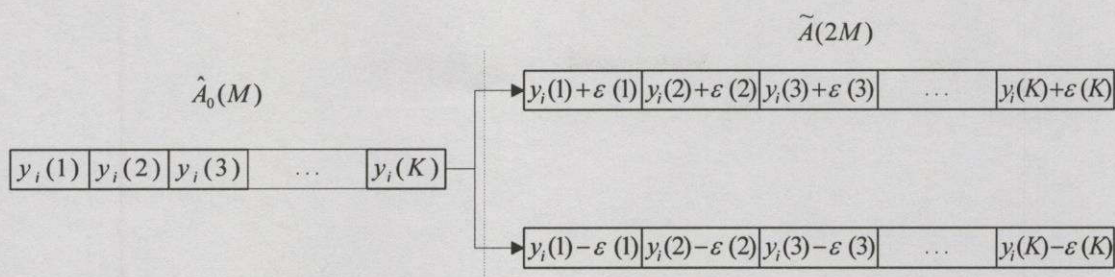
ขั้นตอนที่ 1 กำหนดค่าเริ่มต้นโดยการกำหนด  $M = 1$  และกำหนดตัวเก็บรหัส  $\hat{A}_0(1) = \hat{x}(A)$  เมื่อ  $\hat{x}(A)$  คือค่าเฉลี่ยหรือ centroid ของลำดับการฝึกหัด ซึ่งสามารถหาได้ดังนี้

$$\hat{x}(A) = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{เมื่อ } i = 1, 2, \dots, n \quad (4.13)$$

เมื่อ  $x_i$  คือเวกเตอร์ที่นำมาใช้ในการฝึก (Training vector)

$n$  คือจำนวนเวกเตอร์ที่นำมาใช้ในการฝึก

ขั้นตอนที่ 2 ให้ตัวเก็บรหัส  $\hat{A}_0(M)$  ซึ่งประกอบด้วยเวกเตอร์รหัสจำนวน  $M$  เวกเตอร์ คือ  $\hat{A}_0(M) = \{y_i; i = 1, \dots, M\}$  ทำการแยกทุกเวกเตอร์  $y_i$  ให้เป็น 2 เวกเตอร์ที่ใกล้เคียงกันคือ  $y_i + \varepsilon$  และ  $y_i - \varepsilon$  เมื่อ  $\varepsilon$  คือเวกเตอร์คงที่ ที่ใช้ในการแยก ทำให้ได้ตัวเก็บรหัส  $\tilde{A}(2M) = \{y_i + \varepsilon, y_i - \varepsilon; i = 1, \dots, M\}$  ประกอบด้วย  $2M$  เวกเตอร์ ดังแสดงใน ภาพที่ 4.8 จากนั้นแทนค่า  $M$  ด้วย  $2M$



ภาพที่ 4.8 แสดงการแยกตัวเก็บรหัสเป็น 2 เท่า

ขั้นตอนที่ 3 ทำการตรวจสอบว่า  $M$  เท่ากับ  $N_c$  หรือไม่

- ถ้า  $M = N_c$  ให้กำหนด  $\hat{A}_0 = \tilde{A}(M)$  ซึ่ง  $\hat{A}_0$  จะเป็นค่าเริ่มต้นของตัวเก็บรหัสของอัลกอริทึม LBG สำหรับการควอนไทซ์  $N_c$  ระดับ
- ถ้า  $M > N_c$  ให้ใช้อัลกอริทึม LBG โดยกำหนดค่าเริ่มต้นด้วย  $\tilde{A}(M)$  เพื่อทำการหาตัวเก็บรหัสให้ได้ค่าความเพี้ยนที่น้อยที่สุด จากนั้นกลับไปขั้นตอนที่ 2

การสร้างตัวเก็บรหัสโดยการกำหนดค่าเริ่มต้นด้วยการแตกตัว ภายหลังจากการแบ่งกลุ่มเวกเตอร์สำหรับการฝึกที่มีความเพี้ยนน้อยที่สุดกับเวกเตอร์รหัสทุกเวกเตอร์แล้ว อาจจะมีเวกเตอร์รหัสที่ไม่มีเวกเตอร์สำหรับการฝึกเป็นสมาชิกอยู่เลย ซึ่งทำให้เวกเตอร์รหัสนั้นไม่เกิดประโยชน์ในเวลาเข้ารหัส ดังนั้น จึงต้องกำจัดเวกเตอร์รหัสนั้นออก แล้วสร้างเวกเตอร์ใหม่โดยการนำเวกเตอร์รหัสที่มีสมาชิกมากที่สุด ( $Y_{\max}$ ) มาทำการแตกตัวออกเป็นเวกเตอร์รหัส 2 เวกเตอร์ที่ใกล้เคียงกันคือ  $Y_{\max} + \varepsilon$  และ  $Y_{\max} - \varepsilon$  แทนที่เวกเตอร์  $Y_{\max}$  เดิมและเวกเตอร์รหัสที่ถูกกำจัดออกไป

ในวิทยานิพนธ์ฉบับนี้ได้ทำการทดลองสร้างตัวเก็บรหัส โดยใช้อัลกอริทึม LBG โดยมีกำหนดค่าเริ่มต้นโดยการแตกตัว ดังนี้คือ ทำการสร้างตัวเก็บรหัสโดยเริ่มต้นจากการกำหนดให้ตัวเก็บรหัสมีขนาดเป็น 1 ระดับ (มีจำนวนเวกเตอร์รหัส 1 เวกเตอร์) แล้วนำมาผ่านอัลกอริทึม LBG จากนั้นทำการแยกตัวเก็บรหัสที่ได้ออกเป็น 2 เท่า (มีจำนวนเวกเตอร์รหัส 2 เวกเตอร์) แล้วทำการฝึกตัวเก็บรหัสให้ได้ตัวเก็บรหัสที่ทำให้เกิดความเพี้ยนน้อยที่สุด จากนั้นมาทำการแยกตัวเก็บรหัสเป็น 2 เท่าอีก (มีจำนวนรหัส 4 เวกเตอร์) แล้วนำมาผ่านอัลกอริทึม LBG อีก โดยจะทำเช่นนี้วนไปเรื่อยๆ จนกระทั่งได้ตัวเก็บรหัสมีขนาดเป็น 256 ระดับ (มีจำนวนเวกเตอร์รหัส 256 เวกเตอร์) ดังนั้น จะได้ตัวเก็บรหัสที่มีขนาดเป็น 1, 2, 4, 8, 16, 32, 64, 128 และ 256 ระดับ ตามลำดับ

## บทที่ 5

# การแปลงดิสครีตโคไซน์

### 5.1 คุณสมบัติและนิยามของการแปลงดิสครีตโคไซน์ (Discrete Cosine Transform, DCT)

การลดความซับซ้อนของการคำนวณทางคณิตศาสตร์ทั่วไปจะใช้วิธีการแปลงให้เป็นค่าที่เหมาะสมกับแต่ละการใช้งานเพื่อลดเวลาในการประมวลผล การแปลงฟูเรียร์ถูกนำมาใช้ในการแยกส่วนประกอบของสัญญาณทางเวลาให้อยู่ในรูปของความถี่ในย่านต่าง ๆ ซึ่งได้มีการนำมาประยุกต์ใช้ในงานทางด้านกรประมวลผลภาพในปัจจุบัน ในบทนี้จะกล่าวถึงคุณสมบัติและการพัฒนาสู่ DCT

#### 5.1.1 การแปลงโคไซน์ฟูเรียร์ (Fourier Cosine Transform, FCT)

จากข้อกำหนดของการแปลงฟูเรียร์ [16] โดยกำหนดให้ ฟังก์ชัน  $x(t)$  สำหรับค่า  $-\infty < t < \infty$  จะได้สมการแปลงฟูเรียร์ของฟังก์ชัน  $x(t)$  เป็น

$$X(\omega) \equiv F[x(t)] = \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (5.1)$$

และการแปลงกลับฟูเรียร์ของฟังก์ชัน  $x(t)$  เป็น

$$x(t) \equiv F^{-1}[X(\omega)] = \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} X(\omega)e^{j\omega t} d\omega \quad (5.2)$$

ทำการสร้างฟังก์ชัน  $y(t)$  ในรูปของฟังก์ชัน  $x(t)$  โดยกำหนดค่า  $t \geq 0$  ตามนี้

$$y(t) = \begin{cases} x(t) & t \geq 0, \\ x(-t) & t \leq 0. \end{cases}$$

จะได้การแปลงฟูเรียร์ของฟังก์ชัน  $y(t)$  ดังนี้

$$\begin{aligned}
 F[y(t)] &= \left(\frac{1}{2\pi}\right)^{1/2} \left\{ \int_0^{\infty} x(t)e^{-j\omega t} dt + \int_{-\infty}^0 x(-t)e^{-j\omega t} dt \right\} \\
 &= \left(\frac{1}{2\pi}\right)^{1/2} \int_0^{\infty} x(t)[e^{-j\omega t} + e^{j\omega t}] dt \\
 &= \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(t) \cos(\omega t) dt
 \end{aligned} \tag{5.3}$$

จากสมการที่ (5.3) ซึ่งเป็นสมการของการแปลงโคซายน์ฟูเรียร์ (FCT) ดังนั้นการแปลงโคซายน์ฟูเรียร์ของสมการ  $x(t)$  โดยกำหนดให้

$$X_c(\omega) \equiv F_c[x(t)] = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(t) \cos(\omega t) dt \tag{5.4}$$

และหาฟังก์ชันการแปลงกลับของสมการที่ (5.3) ดังนี้

$$y(t) = x(t) \equiv F_c^{-1}[X_c(\omega)] = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} X_c(\omega) \cos(\omega t) d\omega \quad (t \geq 0) \tag{5.5}$$

จากสมการที่ (5.4) และ (5.5) เป็นฟังก์ชันการแปลงสัมพันธกันและจะได้คุณสมบัติหลายประการตามนี้

- 1) คุณสมบัติการแปลงกลับ (Inversion)

$$F_c \equiv F_c^{-1} \tag{5.6}$$

ซึ่งแสดงดังสมการที่ (2.4) และ (2.5) ได้ว่า

$$F_c^{-1}\{F_c[x(t)]\} = x(t) \quad t \geq 0$$

## 2) ความเป็นเชิงเส้น (Linearity)

$$F_c[\alpha x(t) + \beta y(t)] = \alpha X_c(\omega) + \beta Y_c(\omega) \quad (5.7)$$

เมื่อ  $\alpha$  และ  $\beta$  เป็นค่าคงที่

## 3) การขยายในเวลา (Scaling in time)

$$\begin{aligned} F_c[x(at)] &= \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(at) \cos(\omega t) dt \\ &= a^{-1} \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(\tau) \cos\left(\frac{\omega \tau}{a}\right) d\tau \\ &= \frac{1}{a} X_c\left(\frac{\omega}{a}\right), \quad a > 0 \end{aligned} \quad (5.8)$$

## 4) เลื่อนในเวลา (Shift in time)

$$\begin{aligned} F_c[x(t - \alpha)] &= \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(t - \alpha) \cos(\omega t) dt \\ &= \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(\tau) \cos[\omega(\tau + \alpha)] d\tau \\ &= \cos(\omega\alpha) F_c[x(t)] - \sin(\omega\alpha) F_s[x(t)] \end{aligned} \quad (5.9a)$$

เมื่อ  $F_s$  เป็นการแปลงซายน์ฟูเรียร์ (FST) มีสมการตามนี้

$$F_s[x(t)] \equiv \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} x(t) \sin(\omega t) dt \quad (5.9b)$$

โดยกำหนดว่า  $x(t) = 0$  สำหรับค่า  $t < 0$

## 5) อนุพันธ์ในเวลา (Differentiation in time)

$$\begin{aligned}
 F_c \left[ \frac{d}{dt} x(t) \right] &= \left( \frac{2}{\pi} \right)^{1/2} \int_0^{\infty} \left[ \frac{d}{dt} x(t) \right] \cos(\omega t) dt \\
 &= \left( \frac{2}{\pi} \right)^{1/2} \left\{ [x(t) \cos(\omega t)]_0^{\infty} + \omega \int_0^{\infty} x(t) \sin(\omega t) dt \right\} \\
 &= - \left( \frac{2}{\pi} \right)^{1/2} x(0) + \omega F_s [x(t)]
 \end{aligned} \tag{5.10}$$

โดยสมมุติว่าค่า  $x(t)$  มีค่าเป็นศูนย์เมื่อ  $t$  เข้าสู่ค่าอนันต์

6) คุณสมบัติการประสาน (The convolution property)

จากการประสานกันของฟังก์ชันสองฟังก์ชัน  $x(t)$  และ  $y(t)$  จะได้ว่า

$$x(t) * y(t) = \int_0^{\infty} x(t-u)y(u)du \tag{5.11}$$

การแปลงโคซายน์ฟูเรียร์ของสมการที่ (5.11) คือ

$$\begin{aligned}
 F_c [x(t) * y(t)] &= \left( \frac{2}{\pi} \right)^{1/2} \int_0^{\infty} \int_0^{\infty} x(t-u)y(u) \cos(\omega t) du dt \\
 &= \left( \frac{2}{\pi} \right)^{1/2} [X_c(\omega)Y_c(\omega) - X_s(\omega)Y_s(\omega)]
 \end{aligned} \tag{5.12}$$

5.1.2 นิยาม(Definitions)

จากสมการที่ (5.4) การแปลงโคซายน์ฟูเรียร์มีแกนหลัก (Kernel) โดยกำหนดได้ดังนี้

$$K_c(\omega, t) = \cos(\omega t) \tag{5.13a}$$

กำหนดให้  $\omega_m = 2\pi m \delta f$  และ  $t_n = n \delta t$  เป็นการสุ่มเชิงมุมความถี่และเวลา เมื่อ  $\delta f$  และ  $\delta t$  แทนช่วงเวลาของการสุ่มหนึ่งหน่วยความถี่และเวลาตามลำดับ โดยค่า  $m$  และ  $n$  เป็นค่าจำนวนเต็ม สามารถเขียนสมการที่ (5.13a) ให้อยู่ในรูปใหม่ได้ดังนี้

$$K_c(\omega_m, t_n) = K_c(2\pi m \delta f, n \delta t) = \cos(2\pi mn \delta f \delta t) = K_c(m, n) \quad (5.13b)$$

โดยกำหนดต่อไปว่า  $\delta f \delta t = 1/(2N)$  เมื่อ  $N$  เป็นจำนวนเต็มจะได้ว่า

$$K_c(m, n) = \cos\left(\frac{\pi mn}{N}\right) \quad (5.14)$$

สิ่งที่ได้สมการที่ (5.14) จะเป็นแกนหลักของโคซายน์ฟูเรียร์ช่วง (Discretized Fourier Cosine) ซึ่งมีคุณสมบัติข้อ 1) ถึงข้อ 6) เช่นเดียวกับการแปลงโคซายน์ฟูเรียร์ แต่จะง่ายกว่าในการคำนวณและการนำไปใช้งาน จากสมการที่ (5.14) สามารถกำหนดให้อยู่ในรูปของ Matrix การแปลงขนาด  $N+1, N+1$  โดยแสดงสัญลักษณ์เป็น  $[M]$  โดยมีค่า  $m$  และ  $n$  เป็นส่วนประกอบแต่ละค่าใน Matrix ตามนี้

$$[M]_{mn} = \cos\left(\frac{\pi mn}{N}\right) \quad m, n = 0, 1, \dots, N. \quad (5.15)$$

เมื่อทำการคูณ Matrix  $[M]$  เข้ากับ vector  $x = [x(0), x(1), x(2), \dots, x(N)]^T$  จะได้ vector  $X = [X(0), X(1), \dots, X(N)]^T$  ดังนี้

$$X = [M]x \quad (5.16)$$

เมื่อ

$$X(m) = \sum_{n=0}^N \cos\left(\frac{\pi mn}{N}\right) x(n) \quad m = 0, 1, \dots, N. \quad (5.17)$$

สังเกตเห็นได้ว่า  $x$  คือค่าของการสุ่มสัญญาณเป็นจำนวน  $N$  ครั้ง ส่วนสมการที่ (5.17) เป็นสมการของการแปลงดิสครีตโคซายน์ (Discrete Cosine Transform) ซึ่งนำไปใช้กันมากในงานประมวลผลสัญญาณดิจิทัล (Digital Signal Processing) โดยได้มีการกำหนดนิยามของ DCT ออกเป็น 4 กลุ่ม [17] ด้วยกันดังนี้

(1) DCT-I:

$$[C_{N+1}^I]_{mn} = \left(\frac{2}{N}\right)^{1/2} \left[ k_m k_n \cos\left(\frac{mn\pi}{N}\right) \right] \quad m, n = 0, 1, \dots, N;$$

(2) DCT-II:

$$[C_N^{II}]_{mn} = \left(\frac{2}{N}\right)^{1/2} \left[ k_m \cos\left(\frac{\left(n + \frac{1}{2}\right)m\pi}{N}\right) \right] \quad m, n = 0, 1, \dots, N-1;$$

(3) DCT-III:

$$[C_N^{III}]_{mn} = \left(\frac{2}{N}\right)^{1/2} \left[ k_n \cos\left(\frac{\left(m + \frac{1}{2}\right)n\pi}{N}\right) \right] \quad m, n = 0, 1, \dots, N-1;$$

(4) DCT-IV:

$$[C_N^{IV}]_{mn} = \left(\frac{2}{N}\right)^{1/2} \left[ \cos\left(\frac{\left(m + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\pi}{N}\right) \right] \quad m, n = 0, 1, \dots, N-1;$$

เมื่อ

$$k_j = \begin{cases} 1 & \text{เมื่อ } j \neq 0 \text{ หรือ } N \\ \frac{1}{\sqrt{2}} & \text{เมื่อ } j = 0 \text{ หรือ } N \end{cases}$$

## 5.2 การแปลงดิสครีตโคซายน์ 2 มิติ

การแปลงดิสครีตโคซายน์ 2 มิตินั้น นิยมใช้กันมากในการบีบอัดข้อมูลภาพ เนื่องจากลดขั้นตอนการประมวลผลทางคณิตศาสตร์แล้ว ยังให้ค่า MSE เทียบเคียงได้กับ K-L Transforms [17] โดยสมการการแปลงดิสครีตโคซายน์ 2 มิติ แสดงได้ดังนี้

$$F(u, v) = \frac{2}{N} c(u)c(v) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y) \cos \left[ \frac{(2x+1)u\pi}{2N} \right] \cos \left[ \frac{(2y+1)v\pi}{2N} \right] \quad (5.18)$$

เมื่อ

$u, v$  เป็นตัวแปรทางความถี่ ค่า  $0, 1, \dots, N-1$

$x, y$  เป็นตัวแปรทางเวลาค่า  $0, 1, \dots, N-1$

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & , \quad u = 0 \\ 1 & , \quad u > 0 \end{cases}$$

$$C(v) = \begin{cases} \frac{v}{\sqrt{2}} & , \quad v = 0 \\ 1 & , \quad v > 0 \end{cases}$$

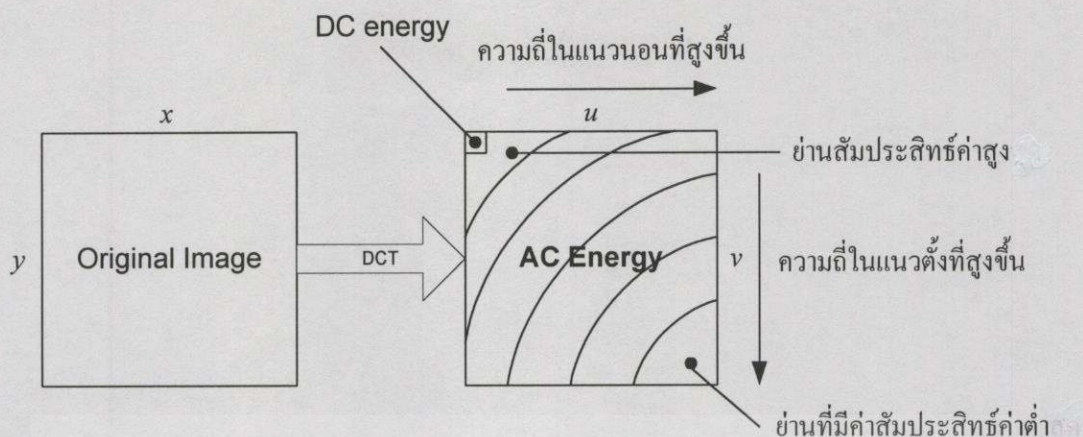
$f(x, y)$  เป็นสมการของภาพขนาด  $N \times N$  ที่ตำแหน่ง  $x, y$  ใด ๆ

$F(u, v)$  เป็นสมการ การแปลงดิสครีตโคซายน์ของฟังก์ชัน  $f(x, y)$

และจะได้สมการแปลงกลับดังนี้

$$\hat{f}(x, y) = \frac{2}{N} c(u)c(v) \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} F(u, v) \cos \left[ \frac{(2x+1)u\pi}{2N} \right] \cos \left[ \frac{(2y+1)v\pi}{2N} \right] \quad (5.19)$$

ผลลัพธ์ที่ได้จากการแปลงดิสครีตโคซายน์ 2 มิติ ของข้อมูลภาพ จะประกอบด้วย 2 ส่วน คือ ส่วนพลังงาน DC (DC component) หรือระดับความสว่าง และ ส่วนพลังงาน AC (AC Component) หรือระดับความคมชัด โดยพลังงาน DC จะเป็นค่าเฉลี่ยของจุดภาพทั้งหมดจะปรากฏที่มุมซ้ายบนของภาพผลลัพธ์หรือ ที่จุด  $0, 0$  ส่วนพลังงานของ AC จะกระจายแบ่งแยกเป็นชั้น ๆ โดยมีค่าสัมประสิทธิ์จากมากไปหาน้อย ส่วนที่อยู่ใกล้จุดพลังงาน DC จะมีค่าสัมประสิทธิ์สูงสุด ส่วนที่อยู่ไกลจากจุดพลังงาน DC จะมีค่าสัมประสิทธิ์ต่ำสุด ดังแสดงได้ดังภาพที่ 5.1 และ ภาพที่ 5.2 แสดงตัวอย่างผลของการแปลง DCT



ภาพที่ 5.1 แสดงส่วนประกอบของผลลัพธ์ที่ได้จากการแปลงดิสครีตโคไซน์ 2 มิติ

140	144	147	140	140	155	179	175
144	152	140	147	140	148	167	179
125	155	136	167	163	162	152	172
168	145	156	160	152	155	136	160
162	148	156	148	140	136	147	162
147	167	140	155	155	140	136	162
136	156	123	167	162	144	140	147
148	155	136	155	152	147	147	136

ก)

151	-18	15	-9	23	-9	-14	-19
21	-34	26	-9	-11	11	14	7
-10	-24	-2	6	-18	3	-20	-1
-8	-5	14	-15	-8	-3	-3	8
-3	10	8	1	-11	-18	18	15
4	-2	-18	8	8	-4	1	-7
9	1	-3	4	-1	-7	-1	-2
0	-8	-2	2	1	4	-6	0

ข)

ภาพที่ 5.2 ก) แสดงข้อมูลขนาด 8\*8 ข) แสดงผลลัพธ์ที่ได้จากการแปลงดิสครีตโคไซน์

จากคุณลักษณะที่ได้จากการแปลงดิสครีตโคไซน์จะเห็นได้ว่า เมื่อความถี่สูงขึ้นค่าสัมประสิทธิ์ในส่วนของพลังงาน AC จะมีค่าน้อยลงหรือเข้าใกล้ศูนย์ ดังนั้นเมื่อต้องการบีบอัดข้อมูลภาพก็ทำการส่งเฉพาะส่วนที่มีค่าสัมประสิทธิ์สูงออกไป และตัดส่วนสัมประสิทธิ์ต่ำที่ไม่มีผลกระทบต่อข้อมูลภาพมากนักทิ้งไปก็จะสามารถลดทอนข้อมูลในการส่งภาพได้ นอกจากนี้พลังงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

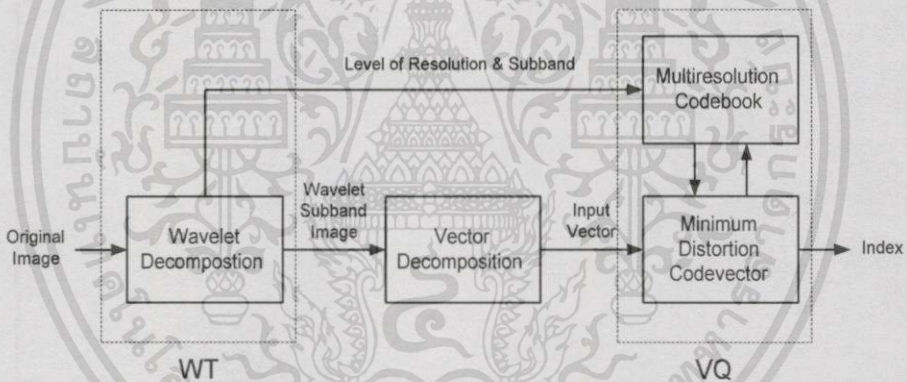
AC ยังบ่งบอกถึงปริมาณข่าวสารของภาพนั้น ๆ ด้วย ภาพที่มีข่าวสารมาก เช่นภาพที่มีความคมชัดสูงหรือภาพที่มีขอบชัดเจน เมื่อทำการแปลงดิสครีตโคไซน์แล้วจะให้ค่าพลังงาน AC ที่สูงกว่าภาพที่มีปริมาณข่าวสารน้อย โดยระดับความสว่างของแต่ละภาพจะไม่มีผลต่อค่าพลังงาน AC เลย จากคุณสมบัติข้อนี้ทำให้สามารถวัดหรือเปรียบเทียบปริมาณข่าวสารของแต่ละภาพได้ แล้วทำการปรับเปลี่ยนการบีบอัดข้อมูลภาพให้เหมาะสม หรือให้ได้อัตราการบีบอัดสูงสุดและให้ค่าความผิดพลาดน้อยที่สุด ซึ่งวิธีการนำไปใช้งานจะกล่าวถึงในบทที่ 6 เรื่องการลดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ



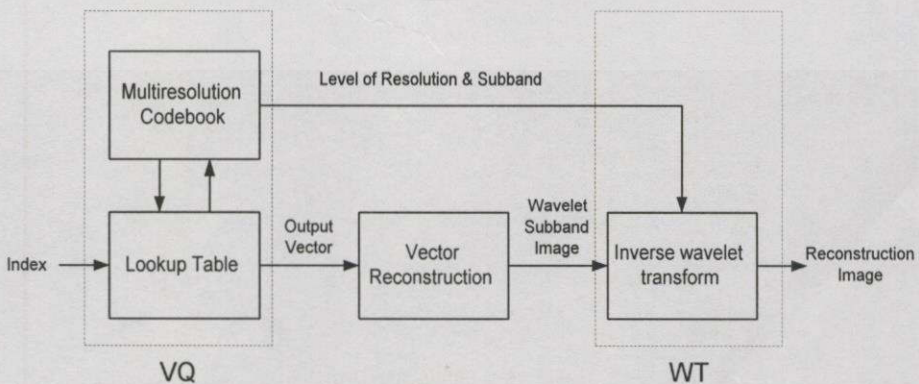
## บทที่ 6

# การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ แบบอะแดปทีฟ

จากการพัฒนาการวิธีการบีบอัดข้อมูลภาพที่ผ่านมานั้น ได้มีการนำวิธีการแปลงหรือเทคนิคการควอนไทซ์ที่นิยมกันในแบบต่าง ๆ มาใช้งาน เช่น การแปลง DCT, การแปลงเวฟเล็ต และการควอนไทซ์แบบเวกเตอร์ จากนั้นจึงได้มีการพัฒนาโดยการนำวิธีการเหล่านี้มาทำงานร่วมกันในการบีบอัดข้อมูลภาพ โดยใช้วิธีการแปลงเวฟเล็ตและการควอนไทซ์แบบเวกเตอร์ [18,19] ทำให้สามารถลดความสูญเสียของข้อมูลภาพที่แปลงกลับมาได้ดีกว่าการใช้การแปลงเวฟเล็ตหรือการควอนไทซ์แบบเวกเตอร์เพียงอย่างเดียว โดยวิธีการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ร่วมกับการแปลงเวฟเล็ตนี้ มีขั้นตอนการทำงานดังภาพที่ 6.1 และการแปลงกลับดังภาพที่ 6.2



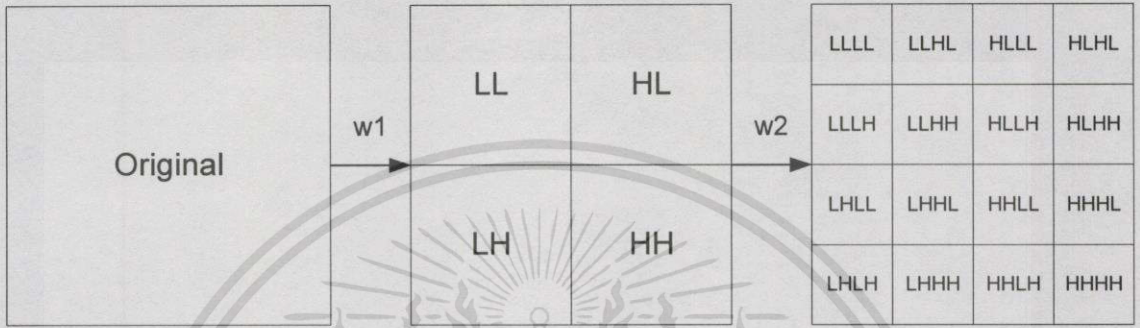
ภาพที่ 6.1 แสดงแผนภาพการบีบอัดข้อมูลภาพรวมทั้งระหว่างการควอนไทซ์แบบเวกเตอร์และการแปลงเวฟเล็ต



ภาพที่ 6.2 แสดงการนำข้อมูลกลับจากการบีบอัดข้อมูลภาพรวมทั้งระหว่างการควอนไทซ์แบบเวกเตอร์และการแปลงเวฟเล็ต

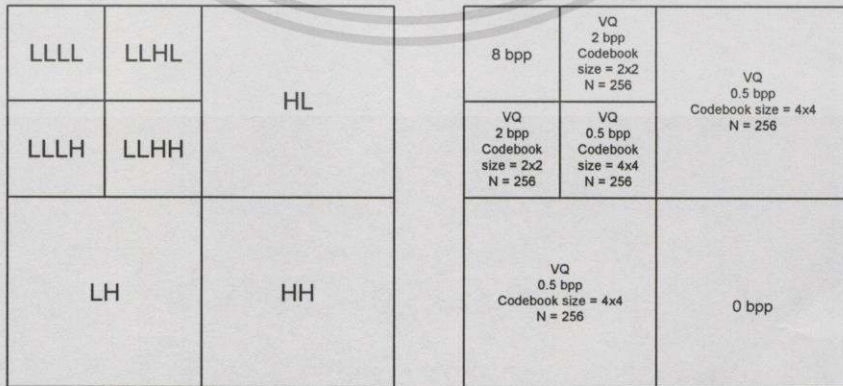
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในกระบวนการแปลงดังภาพที่ 6.1 นั้น กระบวนการจะเริ่มต้นจากการนำข้อมูลภาพต้นแบบมาทำการแปลงเวฟเล็ต 2 ระดับ โดยในการแปลงเวฟเล็ตระดับที่ 1 จะได้ภาพแบนด์ย่อยของเวฟเล็ตขนาดหนึ่งส่วนสองเท่าของภาพต้นแบบจำนวน 4 บล็อก และเมื่อนำมาแปลงเวฟเล็ตระดับที่ 2 จะได้ภาพแบนด์ย่อยขนาดหนึ่งส่วนสี่เท่าของภาพต้นแบบจำนวน 16 บล็อก แสดงได้ดังภาพที่ 6.3



ภาพที่ 6.3 แสดงผลการแปลงเวฟเล็ตระดับที่ 1 และ 2

โดยภาพแบนด์ย่อยของเวฟเล็ตในแต่ละบล็อกจะถูกนำไปสู่กระบวนการควอนไทซ์แบบเวกเตอร์ ในขบวนการนี้จะมีการเลือกบล็อกย่อยและกำหนดอัตราบิตในการบีบอัดให้แต่ละบล็อกย่อยนั้น โดยขึ้นอยู่กับแต่ละวิธีการ ซึ่งในที่นี้จะมีการใช้งานอยู่ 2 แบบ คือการกำหนดอัตราบิตแบบทั่วไป [18] และการกำหนดอัตราบิตแบบละเอียด [19] ในการกำหนดอัตราบิตแบบทั่วไปนั้นจะทำการกำหนดบิตให้กับบล็อกย่อย LLLL, LLHL, LLLH และ LLHH ในแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ตระดับที่ 2 พร้อมทั้งกำหนดแบนด์ย่อยที่ HL และ LH ที่ได้จากการแปลงเวฟเล็ตระดับที่ 1 แสดงได้ดังภาพที่ 6.4



(ก) ภาพแบนด์ย่อย

(ข) การกำหนดอัตราบิต

ภาพที่ 6.4 วิธีการกำหนดอัตราบิตแบบทั่วไป ที่อัตราบิตรวม = 1.03125 bpp

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การกำหนดอัตราบิตแบบทั่วไปนี้ จะไม่พิจารณาถึงบล็อก HH ทำให้เกิดการสูญเสียข้อมูลในส่วนขของรายละเอียดของขอบภาพ จึงได้มีการกำหนดอัตราบิตแบบละเอียดเพื่อลดการสูญเสียของขอบภาพขึ้น โดยวิธีการนี้จะทำการกำหนดอัตราบิตในแต่ละแบนด์ย่อยในระดับที่ 2 จำนวน 13 บล็อกย่อย ยกเว้นเฉพาะบล็อกย่อยที่ HLLH, LHHL และ HHLL ซึ่งในแต่ละบล็อกย่อยนั้นได้มีการกำหนดอัตราบิตที่แตกต่างกันตามความสำคัญของแต่ละแบนด์ย่อย โดยพิจารณาจากสถิติในการกระจายของข่าวสารในแต่ละบล็อกย่อย แสดงได้ภาพที่ 6.5

LLLL	LLHL	HLLL	HLHL	8 bpp	VQ 2 bpp Codebook size = 2x2 N = 256	VQ 0.5 bpp Codebook size = 4x4 N = 256	VQ 0.5 bpp Codebook size = 4x4 N = 256
LLLH	LLHH	HLLH	HLHH	VQ 2 bpp Codebook size = 2x2 N = 256	VQ 0.5 bpp Codebook size = 4x4 N = 256	0 bpp	VQ 0.5 bpp Codebook size = 4x4 N = 256
LHLL	LHHL	HHLL	HHHL	VQ 0.5 bpp Codebook size = 4x4 N = 256	0 bpp	0 bpp	VQ 0.25 bpp Codebook size = 4x4 N = 16
LHLH	LHHH	HHLH	HHHH	VQ 0.5 bpp Codebook size = 4x4 N = 256	VQ 0.5 bpp Codebook size = 4x4 N = 256	VQ 0.25 bpp Codebook size = 4x4 N = 16	VQ 0.5 bpp Codebook size = 4x4 N = 256

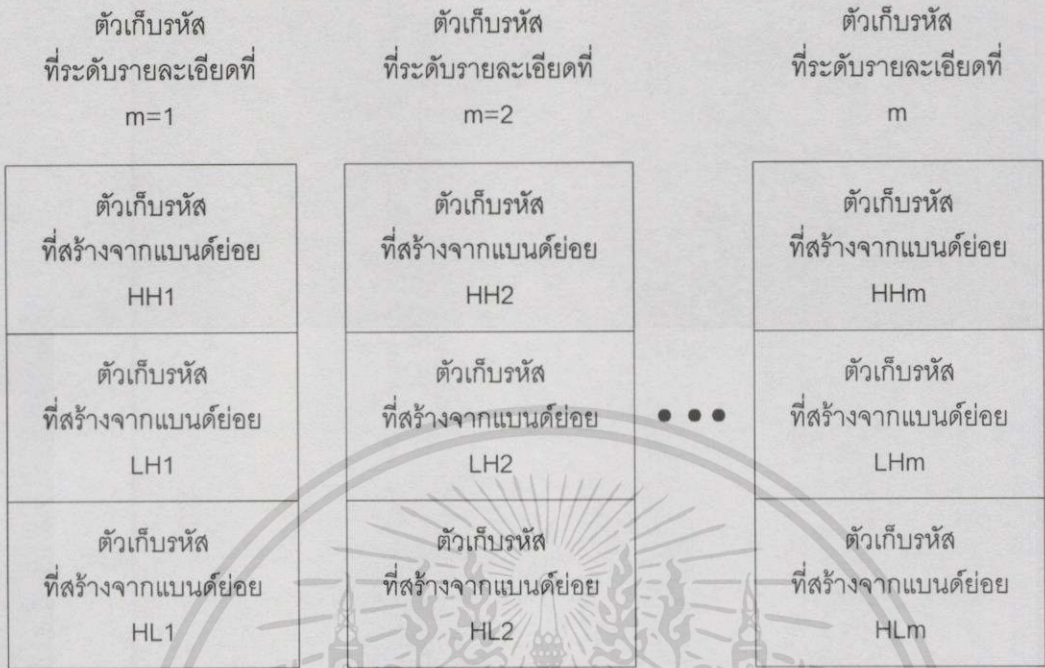
(ก) ภาพแบนด์ย่อย

(ข) การกำหนดอัตราบิต

ภาพที่ 6.5 วิธีการกำหนดอัตราบิตแบบละเอียด ที่อัตราบิตรวม = 1.03125 bpp

เมื่อทำการกำหนดอัตราบิตของแต่ละบล็อกย่อยแล้ว บล็อกย่อยแต่ละบล็อกที่ได้จากการแปลงเวฟเล็ตจะถูกนำไปสร้างตัวเก็บรหัส โดยการสร้างตัวเก็บรหัสนั้นจะสร้างแยกกันตามแต่ละบล็อกย่อยและเนื่องจากการแปลงเวฟเล็ตของภาพเชิงตัวเลขจะได้ภาพแบนด์ย่อยที่มีขนาดเล็กลง โดยที่แต่ละแบนด์ย่อยมีระดับรายละเอียด (Resolution) และความสำคัญของข้อมูลที่แตกต่างกัน และค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ตจะแสดงลักษณะข้อมูลทาง spatial frequency ที่แน่นอนของภาพที่ระดับรายละเอียดและแบนด์ย่อยเดียวกัน [3] จึงทำให้เป็นข้อดีในการสร้างตัวเก็บรหัสตามแต่ละระดับรายละเอียดและแบนด์ย่อย (Multi Resolution Code Book) ดังภาพที่ 6.6

ในขบวนการควอนไตซ์ข้อมูลแต่ละบล็อกย่อย จะตรวจสอบจากตัวเก็บรหัสในแต่ละบล็อกย่อยนั้น เช่น บล็อกย่อยที่ HH ก็จะไปเปิดตัวเก็บรหัสที่ได้จากการสร้างด้วยข้อมูลของบล็อก HH ที่ตัวเก็บรหัส HH1 หรือบล็อกย่อย LLHH ก็จะไปเปิดตัวเก็บรหัสที่ได้จากการสร้างด้วยข้อมูลของบล็อก LLHH ที่ตัวเก็บรหัส HH2 เป็นต้น ผลลัพธ์ที่ได้ก็จะเป็นตัวชี้ในการเปิดตัวเก็บรหัสที่ปลายทาง ซึ่งถือเป็นอันสิ้นสุดขั้นตอนการบีบอัดข้อมูลด้วยวิธีการแปลงเวฟเล็ตร่วมกับการควอนไตซ์แบบเวกเตอร์



ภาพที่ 6.6 แสดงส่วนเข้ารหัสด้วยการควอนไทซ์แบบเวกเตอร์กับการแปลงเวฟเล็ต

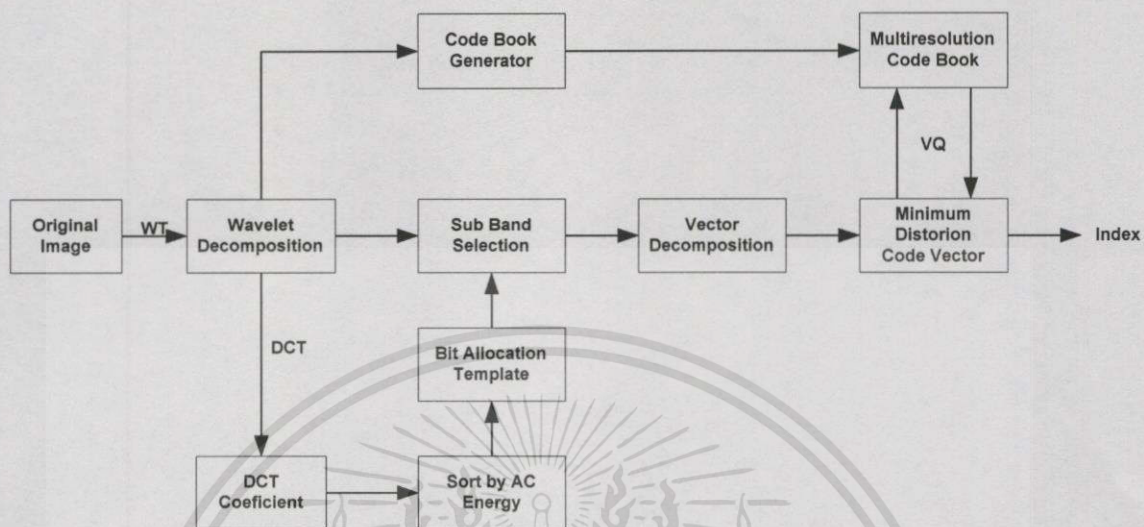
จากวิธีการบีบอัดข้อมูลภาพการแปลงเวฟเล็ตรวมกับการควอนไทซ์แบบเวกเตอร์ที่ได้กล่าวมาแล้วนั้น จุดที่เป็นปัจจัยที่ทำให้เกิดผลการสูญเสียข้อมูลจากขบวนการบีบอัดอยู่ที่วิธีการกำหนดอัตราบิตที่กำหนดให้แต่ละบล็อกย่อยของสัมประสิทธิ์เวฟเล็ต โดยวิธีการกำหนดอัตราบิตทั้งสองวิธีจะเป็นการกำหนดอัตราบิตที่คงตัว สำหรับภาพแต่ละภาพแล้วการกระจายความถี่ของแต่ละแบนด์ย่อยในการแปลงเวฟเล็ตนั้นแตกต่างกันไปและแบนด์ย่อยแต่ละแบนด์มี รูปแบบที่ไม่คงตัว ดังนั้นวิธีการที่มีการกำหนดอัตราบิตให้กับแต่ละแบนด์ย่อยในรูปแบบที่คงตัวจะส่งผลให้การบีบอัดข้อมูลภาพบางภาพเกิดค่า MSE สูงกว่าที่ควรจะเป็น ดังนั้นวิทยานิพนธ์ฉบับนี้จะได้มีแนวคิดที่จะทำการปรับปรุงขบวนการบีบอัดข้อมูลภาพดังกล่าวให้มีความยืดหยุ่นและมีประสิทธิภาพมากขึ้น ดังรายละเอียดในหัวข้อ 6.1

## 6.1 การกำหนดจำนวนบิตแบบอะแดปทีฟ

ขั้นตอนนี้จะขั้นตอนการกำหนดอัตราบิตในแต่ละบล็อกย่อยที่ได้จากการแปลงเวฟเล็ตตามสัดส่วนของข่าวสารในแต่ละบล็อกย่อยนั้น ๆ ซึ่งวิธีการดังกล่าวได้ใช้วิธีการตรวจสอบด้วย DCT เพื่อจัดเรียงอันดับความสำคัญตามปริมาณข้อมูลข่าวสารแต่ละบล็อกย่อย แล้วจึงทำการกำหนดอัตราบิตให้แต่ละบล็อกย่อยให้เหมาะสมเพื่อให้ได้อัตราบิตรวมที่ต้องการ และมีการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สูญเสียน้อยที่สุด ซึ่งขบวนการดังกล่าวคือวิธีการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ สามารถแสดงขั้นตอนการทำงานได้ดังภาพที่ 6.7



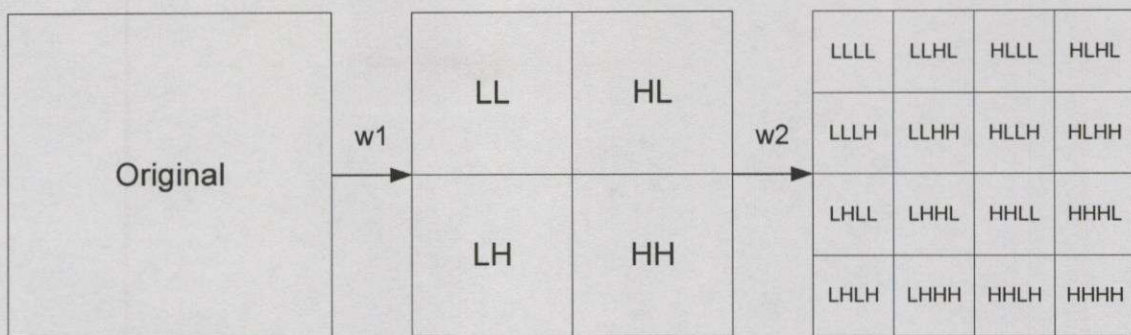
ภาพที่ 6.7 แผนภาพแสดงการทำงานของวิธีการบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์แบบอะแดปทีฟ

รายละเอียดการทำงานของภาพที่ 6.7 แบ่งออกได้เป็น 4 ขั้นตอนด้วยกันคือ

1. ขั้นตอนการแปลงภาพต้นแบบสู่เวฟเล็ต
2. ขั้นตอนตรวจสอบและจัดเรียงบล็อกย่อย
3. ขั้นตอนการกำหนดอัตราบิตให้แต่ละบล็อกย่อย
4. ขั้นตอนการควอนไทซ์แบบเวกเตอร์

#### 1. ขั้นตอนการแปลงภาพต้นแบบสู่เวฟเล็ต

ในขั้นตอนนี้จะนำข้อภาพต้นแบบมาทำการแปลงเวฟเล็ต 2 ระดับ ซึ่งจะได้บล็อกย่อยของสัมประสิทธิ์เวฟเล็ตจำนวน 16 บล็อก ดังภาพที่ 6.8



ภาพที่ 6.8 แสดงภาพการแปลงเวฟเล็ต 2 ระดับ

## 2. ขั้นตอนตรวจสอบและจัดเรียงบล็อกย่อย

ในขั้นตอนนี้จะนำบล็อกย่อยทั้ง 16 บล็อกมาทำการตรวจสอบปริมาณข้อมูลหรือข่าวสารในแต่ละบล็อกย่อย โดยในบล็อกที่ LLLL จะเว้นไว้ทั้งนี้เนื่องจากบล็อกดังกล่าวจะเป็นบล็อกที่เก็บข้อมูลองค์ประกอบหลักของภาพ ซึ่งมีลักษณะเหมือนภาพต้นแบบ แต่มีขนาดเล็กกว่าเท่านั้น ส่วน 15 บล็อกย่อยที่เหลือจะเป็นการกระจายขององค์ประกอบภาพ เช่น จำพวกขอบหรือความถี่ในย่านต่าง ๆ ดังนั้นกระบวนการตรวจสอบจะทำการตรวจสอบเฉพาะ 15 บล็อกย่อยเท่านั้น กระบวนการตรวจสอบบล็อกย่อยนั้น จะทำการตรวจสอบโดยใช้การแปลง DCT ในแต่ละบล็อกย่อย โดยสัมประสิทธิ์ DCT ของแต่ละบล็อกย่อยจะถูกนำมาหาค่าเฉพาะพลังงาน AC ส่วนค่าพลังงาน DC จะไม่นำมาพิจารณา ซึ่งการคำนวณหา AC Energy ของแต่ละบล็อกย่อยแสดงได้ดังนี้

$$AC_i = \sum_{u=0}^j \sum_{v=0}^i |DCT\_Coef_i(u,v)| - DCT\_Coef_i(0,0) \quad (6.1)$$

โดยที่

$$i = 1, 2, \dots, 4^L - 1 \text{ และ } j = \frac{N}{2^L} - 1$$

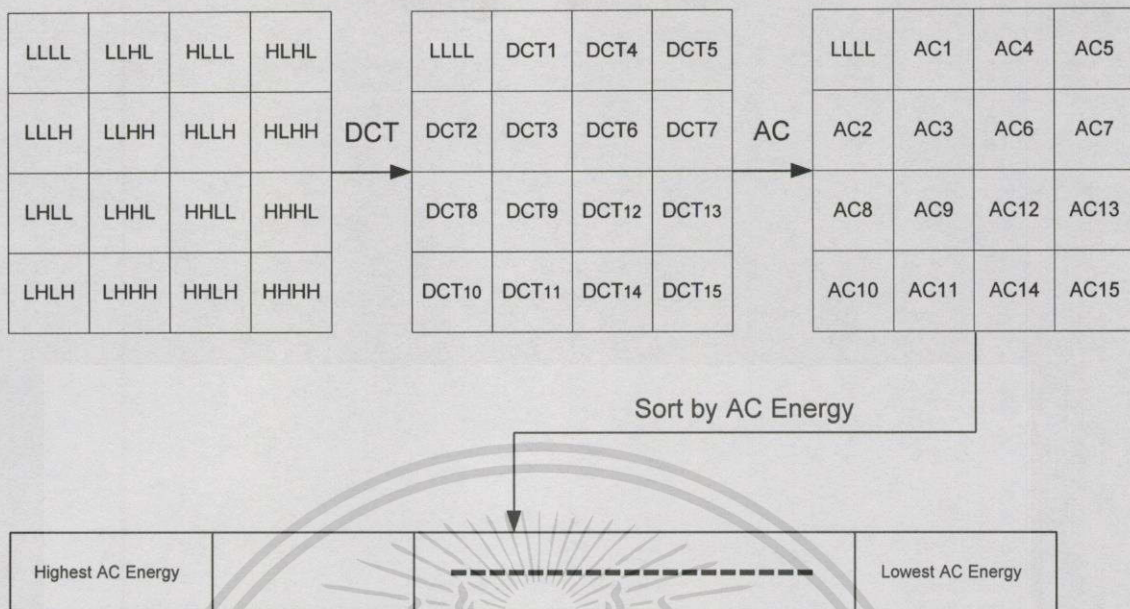
$L$  = ระดับของการแปลง Wavelet ในที่นี้คือระดับที่ 2

และพลังงาน AC Energy รวมทั้งหมดจะเท่ากับ

$$AC_{total} = \sum_{i=1}^{4^L-1} AC_i \quad (6.2)$$

จากนั้นทำการจัดเรียงบล็อกย่อยตามค่าพลังงาน AC จากมากไปหาน้อย ซึ่งกระบวนการทำงานดังกล่าวแสดงไว้ในภาพที่ 6.9

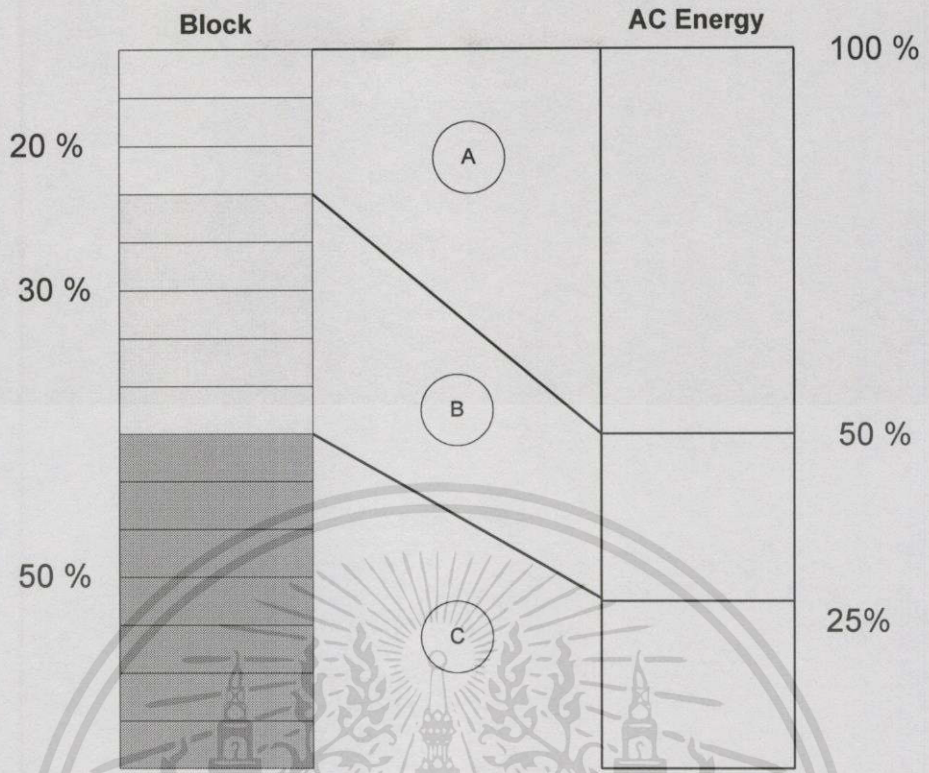
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.9 แสดงภาพการตรวจสอบหาค่าพลังงาน AC ด้วยการแปลง DCT

### 3. ขั้นตอนการกำหนดอัตราบิตให้แต่ละบล็อกย่อย

เมื่อได้ลำดับการจัดเรียงตามค่าพลังงาน AC ของบล็อกย่อยแล้ว ขั้นตอนต่อไปจะเป็นการกำหนดอัตราบิตให้แต่ละบล็อกย่อย โดยกระบวนการจะเริ่มจากการแบ่งกลุ่มของบล็อกย่อย โดยหลักเกณฑ์ในการแบ่งกลุ่มจะนำผลรวมของค่าพลังงาน AC สะสมจากบล็อกที่มีค่าพลังงาน AC สูงสุดรวมกับบล็อกที่มีพลังงาน AC อันดับรองลงไป เมื่อครบตามค่าที่กำหนดไว้แล้ว จะทำการกำหนดกลุ่ม จากนั้นก็ทำกระบวนการเดียวกันกับบล็อกถัดไปจนกระทั่งถึงบล็อกสุดท้าย ผลจากการทดลองพบว่าการกระจายของสัดส่วนค่าพลังงาน AC มีการกระจายไม่สม่ำเสมอในแต่ละบล็อกย่อย โดยพลังงาน AC ส่วนใหญ่จะอยู่ที่บล็อกต้นๆ คือบล็อก LLHL และ LLLH ซึ่งเมื่อนำค่าพลังงาน AC มารวมกันแล้วจะมีค่าใกล้เคียงหรือเกิน 50% ของพลังงาน AC รวม ส่วนพลังงานที่เหลือจะกระจายอยู่ในแต่ละบล็อกแตกต่างกันไป ซึ่งจากการทดลองพบว่าสามารถแบ่งออกเป็น 3 กลุ่มด้วยกัน ดังภาพที่ 6.10 คือกลุ่ม A มีจำนวนบล็อกคิดเป็น 20% ของบล็อกทั้งหมด มีค่าพลังงาน AC เป็น 50% ของพลังงาน AC รวมทั้งหมด กลุ่ม B มีจำนวนบล็อกคิดเป็น 30% ของบล็อกทั้งหมด มีค่าพลังงาน AC เป็น 25% และกลุ่ม C มีจำนวนบล็อกคิดเป็น 50% ของบล็อกทั้งหมด มีค่าพลังงาน AC เป็น 25%



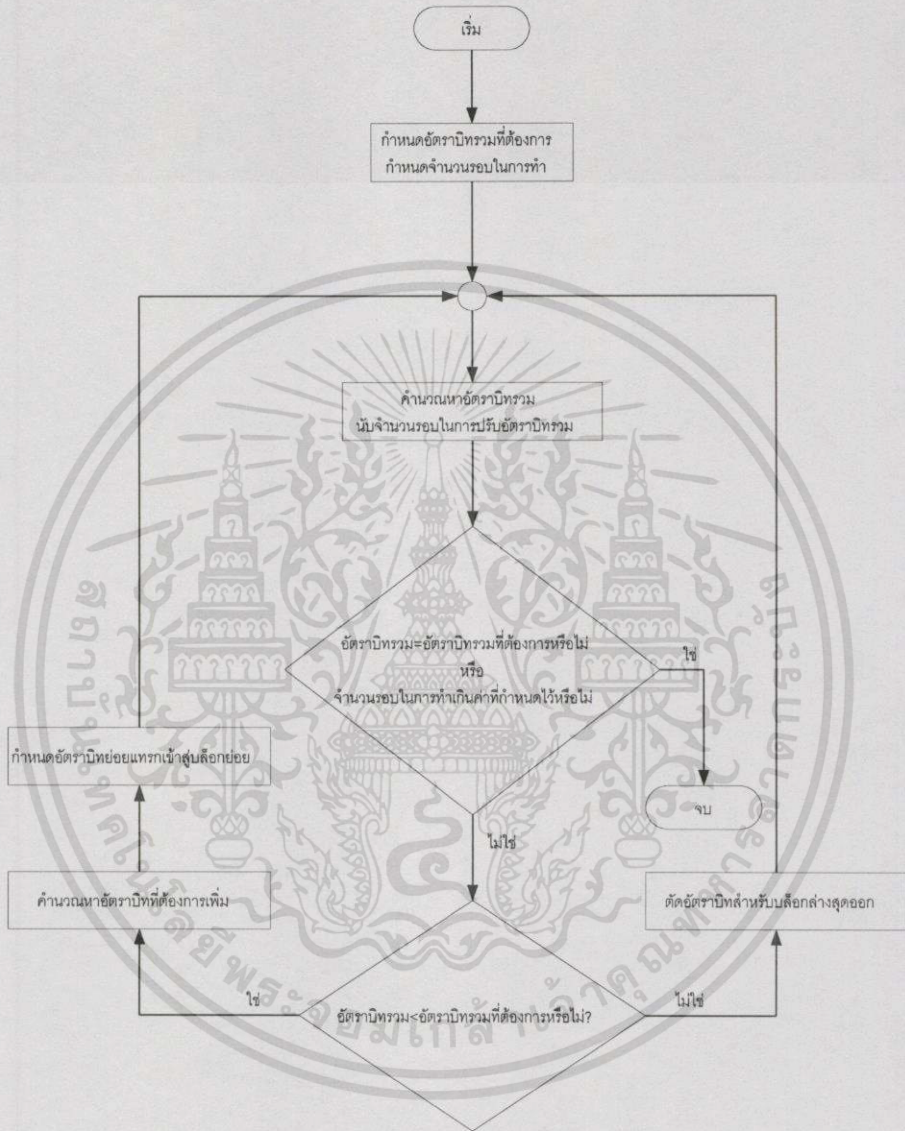
ภาพที่ 6.10 แสดงการจัดแบ่งกลุ่มของแต่ละบล็อกย่อยและเปอร์เซ็นต์ของ AC Energy ในแต่ละกลุ่ม

เมื่อได้กลุ่มของบล็อกย่อยเป็น 3 กลุ่มคือ กลุ่ม A, B และ กลุ่ม C เรียบร้อยแล้วให้ทำการกำหนดอัตราบิตให้เหมาะสมกับสัดส่วนในแต่ละกลุ่ม ซึ่งในกลุ่มจะกำหนดให้มีอัตราบิต ขนาดของเวกเตอร์ และขนาดของบล็อกสำหรับการทำการบีบอัดด้วยการควอนไทซ์แบบเวกเตอร์ โดยกลุ่ม A มีอัตราบิตที่ 2 bpp ขนาดของเวกเตอร์ที่ 8 บิต หรือ 256 เวกเตอร์ และมีขนาดบล็อกของแต่ละเวกเตอร์เท่ากับ 4(2X2) ส่วนกลุ่ม B มีอัตราบิตที่ 0.5 bpp ขนาดของเวกเตอร์ที่ 8 บิต หรือ 256 เวกเตอร์ และมีขนาดบล็อกของแต่ละเวกเตอร์เท่ากับ 16(4X4)และกลุ่ม C ที่อัตราบิต 0 bpp ดังแสดงไว้ในตารางที่ 6.1

ตารางที่ 6.1 แสดงการกำหนดอัตราบิต ขนาดเวกเตอร์ และขนาดบล็อกของแต่ละเวกเตอร์

กลุ่ม	bpp	l (Vecor size)	K (Block size)	% AC_Energy	% จำนวนบล็อก
A	2	8	4	ประมาณ 50 %	ประมาณ 20 %
B	0.5	8	16	ประมาณ 25 %	ประมาณ 30 %
C	0	0	0	ประมาณ 25 %	ประมาณ 50 %

เนื่องจากอัตราบิตที่กำหนดไว้ในแต่ละบล็อกย่อยยังไม่ได้คำนึงถึงอัตราบิตรวมที่กำหนด ดังนั้นจึงต้องทำการปรับค่าอัตราบิตรวมที่คำนวณได้ให้เท่ากับอัตราบิตรวมตามที่ต้องการ ตามผังงานดังภาพที่ 6.11



ภาพที่ 6.11 แสดงผังงานในการปรับอัตราบิตรวมให้ได้อัตราบิตตามที่ต้องการ

### 3.1 การเพิ่มอัตราบิต กรณีอัตราบิตที่คำนวณได้ต่ำกว่าอัตราบิตรวมที่ต้องการ

#### 1. หาอัตราบิต ที่ต้องเพิ่มจาก

$$\text{อัตราบิตที่ต้องเพิ่ม} = \text{อัตราบิตรวมที่ต้องการ} - \text{อัตราบิตที่คำนวณได้}$$

#### 2. นำอัตราบิตที่ต้องเพิ่มมากำหนดตำแหน่งในบล็อก โดยการเลือกหาอัตราบิตที่ใกล้เคียงจากด้านบนลงล่าง แล้วทำการแทรกค่าอัตราบิตที่คำนวณได้ลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไปยังบล็อก ดังกล่าว พร้อมทั้งทำการเลื่อนค่าเดิมที่มีอยู่ (ในกรณีที่อัตราบิตของบล็อกดังกล่าวไม่เท่ากับ 0 bpp) ให้เลื่อนตำแหน่งลง

3.2 การลดอัตราบิต กรณีที่อัตราบิตคำนวณได้สูงกว่าอัตราบิตรวมที่ต้องการ สามารถทำได้โดยกำจัดบล็อกจากด้านล่างสุด โดยกำหนดให้เป็น 0 bpp แล้วตรวจสอบค่าอัตราบิตรวมอีกครั้งหนึ่ง ว่ามากกว่าหรือน้อยกว่าค่าที่ต้องการ ถ้ามากกว่าก็ใช้วิธีการกำจัดค่าบล็อกด้านล่างสุดอีกครั้ง แต่ถ้าน้อยกว่า ก็ให้กลับไปทำขบวนการเพิ่มอัตราบิตทำอย่างนี้ไปเรื่อยๆ จนกว่าจะได้ bpp ตามที่กำหนด หรือครบจำนวนรอบที่จำกัดไว้

พิจารณาภาพที่ 6.11 ผังงานแสดงการเพิ่มและลดจำนวนอัตราบิตให้ได้ตามอัตราบิตที่ต้องการ

#### 4. ขั้นตอนการควอนไทซ์แบบเวกเตอร์

ขั้นตอนนี้จะเป็นการนำบล็อกย่อยในแต่ละบล็อกไปทำการเข้ารหัสเวกเตอร์แบบควอนไทซ์ ตามการกำหนดอัตราการบีบอัดข้อมูลที่ได้จากขั้นตอนการกำหนดอัตราบิตให้แต่ละบล็อกย่อย ที่แสดงไว้ในตารางที่ 6.1 แล้วจึงนำผลลัพธ์ที่ได้ส่งข้อมูลการบีบอัดออกไปสู่ผู้รับได้

#### 6.2 ผลการทดลอง

ในวิทยานิพนธ์ฉบับนี้ได้ทำการทดลองบีบอัดข้อมูลภาพขนาด 512X512 รายละเอียดต่อจุดขนาด 8 บิตต่อจุด จำนวน 3 ภาพด้วยกันคือ ภาพ LENA ภาพถ่ายดาวเทียมแสดงรายละเอียดพื้นที่ชายฝั่งบริเวณอ่าวคุ้งกระเบน จังหวัดจันทบุรี (ภาพจากดาวเทียม JERS-1) ชื่อ OPS4 และภาพถ่ายดาวเทียมแสดงบริเวณอ่างเก็บน้ำในประเทศแคนาดา (ภาพจากดาวเทียม Landsat) ชื่อ TM4 ด้วยวิธีการกำหนดอัตราบิตในแต่ละวิธี คือวิธีการกำหนดอัตราบิตแบบทั่วไป [18] การกำหนดอัตราบิตแบบละเอียด [19] และการกำหนดอัตราบิตแบบปรับตัวได้ ซึ่งเป็นวิธีใหม่ที่ได้นำเสนอในวิทยานิพนธ์นี้ โดยการทดลองจะกำหนดอัตราบิตรวมคงที่ไว้ที่ 1.03125 bpp ผลลัพธ์แสดงไว้ในภาพที่ 6.11, 6.12 และ 6.13 จากนั้นได้นำผลลัพธ์ทั้งสามวิธีมาทำคำนวณหาค่า MSE และค่า PSNR เปรียบเทียบกันทั้งสามวิธี ซึ่งแสดงผลไว้ดังตารางที่ 6.2



ก) ภาพต้นแบบ LENA ขนาด 512X512, 8 bpp

LLLL	LLHL	HL	8 bpp	VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256
LLLH	LLHH		VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256	
	LH	HH		VQ 0.5 bpp Codebook size=4X4 N=256	0 bpp

ข) ภาพแบนด์ย่อยเวฟเล็ต

ค) การกำหนดอัตราบิตทั่วไป



ง) ภาพผลลัพธ์จากวิธีการกำหนดอัตราบิตทั่วไป

ค่า MSE=30.2816, ค่า PSNR=33.3190, Bitrate=1.03125 bpp

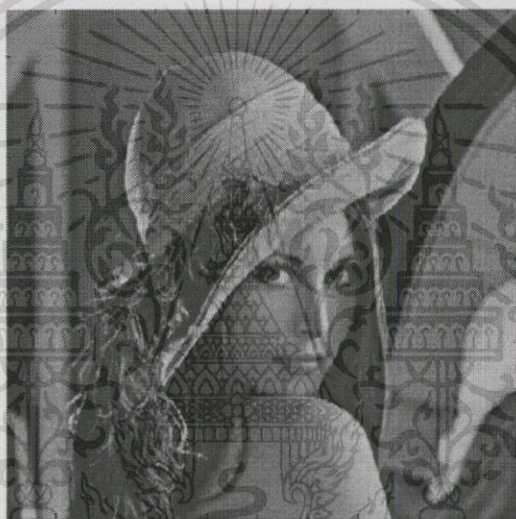
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เพื่อการเรียนการสอนเท่านั้น ไม่อนุญาตให้เผยแพร่ไปโดยไม่ได้รับอนุญาต  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=16	VQ 0.25bpp Codebook size=4x4 N=16

จ) ภาพแบนด์ย่อยเวฟเล็ต

ฉ) การกำหนดอัตราบิตแบบละเอียด



ช) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบละเอียด  
มีค่า MSE=27.1235, ค่า PSNR=33.7973, Bitrate=1.03125 bpp

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 2bpp Codebook size=2x2 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=256
	0 bpp	0 bpp	0 bpp
	0 bpp	0 bpp	0 bpp

ซ) ภาพแบนด์ย่อยเวฟเล็ต

ฅ) การกำหนดอัตราบิตแบบปรับตัวได้

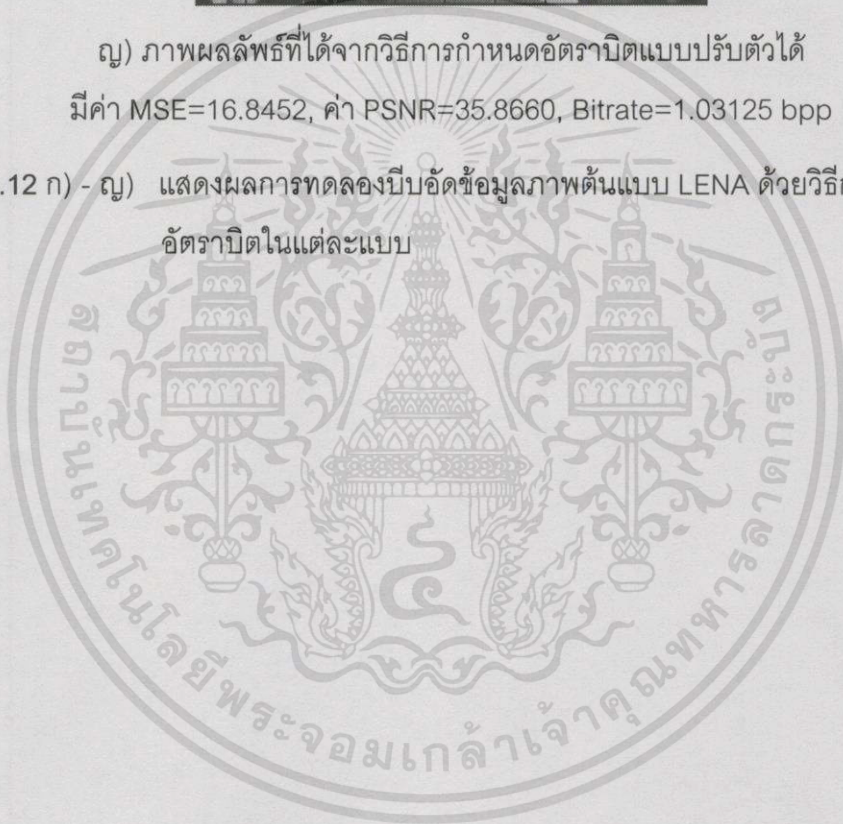
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ณ) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบปรับตัวได้

มีค่า MSE=16.8452, ค่า PSNR=35.8660, Bitrate=1.03125 bpp

ภาพที่ 6.12 ก) - ณ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ LENA ด้วยวิธีการกำหนดอัตราบิตในแต่ละแบบ





ก) ภาพต้นแบบ OPS4 ขนาด 512X512, 8 bpp

LLLL	LLHL	HL	8 bpp	VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256
LLLH	LLHH		VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256	
	LH	HH		VQ 0.5 bpp Codebook size=4X4 N=256	0 bpp

ข) ภาพแบนด์ย่อยเวฟเล็ต

ค) การกำหนดอัตราบิตทั่วไป



ง) ภาพผลลัพธ์จากวิธีกำหนดอัตราบิตทั่วไป

ค่า MSE=60.1278, ค่า PSNR=30.3400, Bitrate=1.03125 bpp

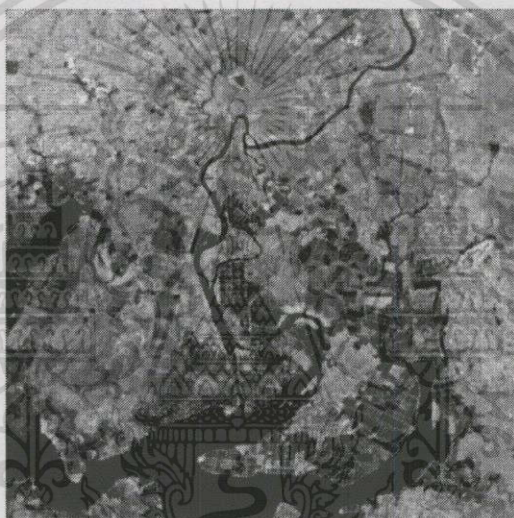
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.25bpp Codebook size=4x4 N=16

จ) ภาพแบนด์ย่อยเวฟเล็ต

ฉ) การกำหนดอัตราบิตแบบละเอียด



ช) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบละเอียด

มีค่า MSE=55.3232, ค่า PSNR=30.7017, Bitrate=1.03125 bpp

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 2bpp Codebook size=2x2 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp

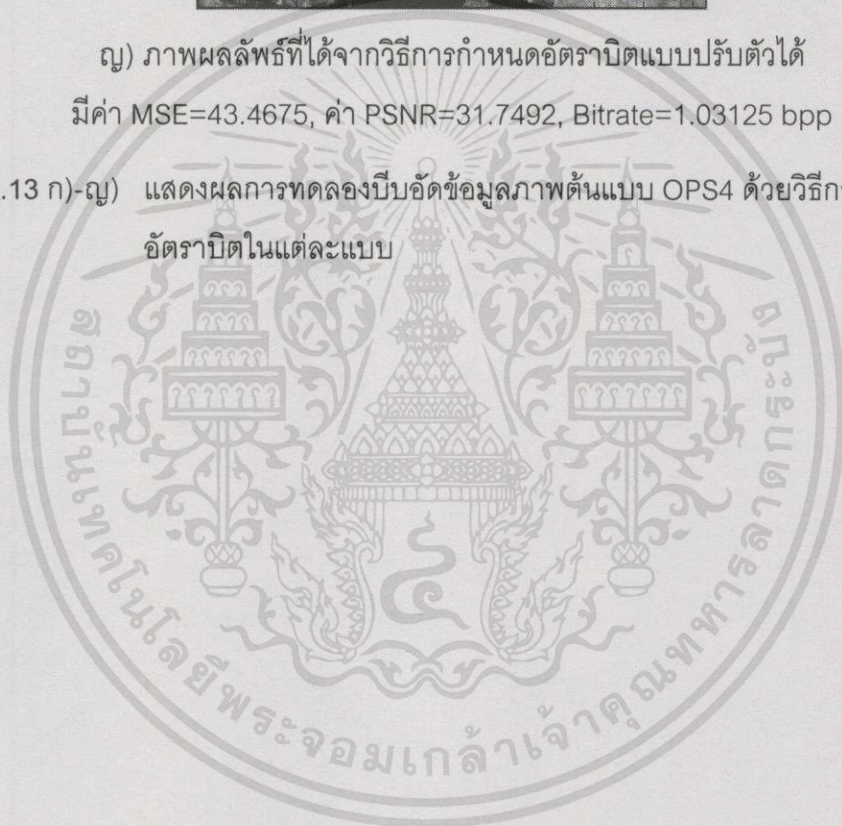
ซ) ภาพแบนด์ย่อยเวฟเล็ต

ฅ) การกำหนดอัตราบิตแบบปรับตัวได้



ณ) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบปรับตัวได้  
มีค่า MSE=43.4675, ค่า PSNR=31.7492, Bitrate=1.03125 bpp

ภาพที่ 6.13 ก)-ณ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ OPS4 ด้วยวิธีการกำหนด  
อัตราบิตในแต่ละแบบ





ก) ภาพต้นแบบ TM4 ขนาด 512X512, 8 bpp

LLLL	LLHL	HL	8 bpp	VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256
LLLH	LLHH		VQ 2 bpp Codebook size=2X2 N=256	VQ 0.5 bpp Codebook size=4X4 N=256	
	LH	HH		VQ 0.5 bpp Codebook size=4X4 N=256	0 bpp

ข) ภาพแบนด์ย่อยเวฟเล็ก

ค) การกำหนดอัตราบิตทั่วไป



ง) ภาพผลลัพธ์จากวิธีกำหนดอัตราบิตทั่วไป

มีค่า MSE=47.6863, ค่า PSNR=31.3469, Bitrate=1.03125 bpp

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

จ) ภาพแบนด์ย่อยเวฟเล็ต

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.5bpp Codebook size=4x4 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	VQ 0.25bpp Codebook size=4x4 N=16	VQ 0.25bpp Codebook size=4x4 N=16

ฉ) การกำหนดอัตราบิตแบบละเอียด



ช) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบละเอียด

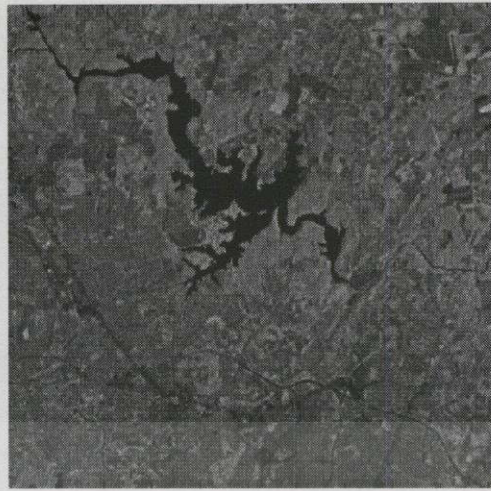
มีค่า MSE=44.2886, ค่า PSNR=31.6679, Bitrate=1.03125 bpp

LLLL	LLHL	HLLL	HLHL
LLLH	LLHH	HLLH	HLHH
LHLL	LHHL	HHLL	HHHL
LHLH	LHHH	HHLH	HHHH

ซ) ภาพแบนด์ย่อยเวฟเล็ต

8 bpp	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	VQ 2bpp Codebook size=2x2 N=256
	VQ 2bpp Codebook size=2x2 N=256	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp
	VQ 0.5bpp Codebook size=4x4 N=256	0 bpp	0 bpp

ฌ) การกำหนดอัตราบิตแบบปรับตัวได้



ญ) ภาพผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบปรับตัวได้

มีค่า MSE=35.9265, ค่า PSNR=32.5767, Bitrate=1.03125 bpp

ภาพที่ 6.14 ก)-ญ) แสดงผลการทดลองบีบอัดข้อมูลภาพต้นแบบ TM4 ด้วยวิธีการกำหนดอัตราบิตในแต่ละแบบ

ตารางที่ 6.2 เปรียบเทียบค่า MSE และค่า PSNR จากผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตของแต่ละวิธี

ภาพต้นแบบ	วิธีการกำหนดอัตราบิต					
	แบบทั่วไป [18]		แบบละเอียด [19]		แบบปรับตัว	
	MSE	PSNR	MSE	PSNR	MSE	PSNR
LENA	30.2816	33.3190	27.1235	33.7973	16.8452	35.8660
OPS4	60.1278	30.3400	55.3232	30.7017	43.4675	31.7492
TM4	47.6863	31.3469	44.2886	31.6679	35.9265	32.5767

### 6.3 สรุปผลการทดลอง

จากผลการทดลองพบว่าผลลัพธ์ที่ได้จากวิธีการกำหนดอัตราบิตแบบปรับตัวที่น่าเสนอนั้น สามารถลดอัตราการสูญเสียของข้อมูลภาพสูงที่สุดเมื่อเทียบกับอีกสองวิธีในอัตราการบีบอัดข้อมูลที่เท่ากัน ทั้งนี้เนื่องจากวิธีการกำหนดอัตราบิตแบบทั่วไปและแบบละเอียดมีการกำหนดอัตราบิตในแต่ละแบนด์ย่อยที่คงตัว ส่งผลให้ในบางกรณีที่แบนด์ย่อยที่มีข้อมูลน้อยกลับได้อัตราบิตที่สูง ส่วนแบนด์ย่อยที่มีข้อมูลมากกลับได้อัตราบิตที่ต่ำหรือเป็น 0 bpp ส่วนวิธีการกำหนดอัตราบิตที่น่าเสนอในวิทยานิพนธ์นี้สามารถจัดการเกี่ยวกับการกำหนดอัตราบีบอัดข้อมูลในแต่ละแบนด์ย่อย ให้ได้อัตราการบีบอัดที่ปรับตัวตามการกระจายของข้อมูลภาพในแต่ละบล็อกย่อย ซึ่งสังเกตเห็นได้ว่ารูปแบบการกำหนดอัตราบิตแต่ละแบนด์ย่อยในแต่ละภาพจะไม่ซ้ำกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทสรุปและแนวทางการพัฒนา

## 7.1 สรุปผลการวิจัย

จากผลการทดลองที่ได้พบว่า การบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบอะแดปทีฟด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเล็ตนั้น สามารถปรับเปลี่ยนค่าของจำนวนบิตในการเข้ารหัสในแต่ละแบนด์ย่อย ตามค่าสัดส่วนของพลังงาน a.c. ที่หาได้จากการแปลงดิสครีตโคไซน์ ซึ่งค่าพลังงาน a.c. นี้จะเป็นตัวบ่งบอกปริมาณข่าวสารที่ปรากฏในภาพ แล้วทำการกำหนดจำนวนบิตในการเข้ารหัสของแบนด์ย่อยต่างๆ ซึ่งจะแปรเปลี่ยนไปตามปริมาณข่าวสารที่บ่งชี้โดยค่าพลังงาน a.c. โดยถ้าภาพแบนด์ย่อยนั้นมีค่าพลังงาน a.c. สูงก็จะเข้ารหัสด้วยจำนวนบิตที่สูง แต่ถ้าภาพแบนด์ย่อยนั้นมีค่าพลังงาน a.c. ต่ำ ก็จะใช้จำนวนบิตที่ต่ำ ลดหลั่นกันลงไปตามสัดส่วนที่คำนวณได้ ซึ่งวิธีการดังกล่าวจะช่วยในส่วนของการรักษาอัตราบิตในการจัดส่งไม่ให้สูงเกินความจำเป็น และมีความผิดพลาดของข้อมูลน้อยที่สุด นอกจากนี้ยังสามารถนำไปประยุกต์ใช้กับภาพที่มีการกระจายของข้อมูลที่มีความหลากหลายได้อย่างมีประสิทธิภาพ โดยพิจารณาจากค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error, MSE) และค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak-signal-to-noise ratio, PSNR)

สรุปปัจจัยที่มีผลต่อจำนวนบิตที่ใช้ในการลดข้อมูล [1] คือ

1. จำนวนบิตของดัชนีที่ใช้ในการเข้ารหัส ซึ่งจะขึ้นอยู่กับจำนวนของเวกเตอร์รหัสหรือระดับการควอนไทซ์ของตัวเก็บรหัส
2. จำนวนดัชนีที่ใช้ในการเข้ารหัส ซึ่งจะขึ้นอยู่กับขนาดมิติของเวกเตอร์รหัสหรือขนาดของบล็อกย่อย

จากผลการทดลองวัดประสิทธิภาพในการลดข้อมูลที่ได้จะพบว่าถ้าอัตราบิตที่ใช้ในการควอนไทซ์แบบเวกเตอร์ลดลงจะทำให้เกิดความเพี้ยนของภาพที่สูงขึ้น คือจะมีค่าความผิดพลาดกำลังสองเฉลี่ยที่สูงขึ้น และมีค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุดที่ลดลง

ข้อดีของการบีบอัดข้อมูลภาพด้วยวิธีนี้ก็คือ

1. เป็นวิธีการบีบอัดข้อมูลที่มีประสิทธิภาพสูง เนื่องจากการเข้ารหัสภาพแบนด์ย่อยต่างๆ ด้วยจำนวนบิตที่แปรเปลี่ยนไปตามค่าพลังงานความหนาแน่นของปริมาณข่าวสารที่พิจารณาจากค่าพลังงาน a.c. ทำให้มีความหลากหลายในการนำไปประยุกต์ใช้กับภาพประเภทต่างๆ
2. สามารถกำหนดอัตราบิตที่จะใช้ในการส่งได้ ทำให้ลดค่าใช้จ่ายและระยะเวลาในการจัดส่ง

3. การเข้ารหัสและถอดรหัสทำได้ง่าย โดยการเข้ารหัสจะใช้การคำนวณหาค่าความผิดพลาดกำลังสองเฉลี่ยแล้วนำมาเปรียบเทียบกัน ส่วนการถอดรหัสก็เป็นเพียงการนำค่าดัชนี มาชี้ข้อมูลในตัวเก็บรหัสโดยไม่มีการคำนวณ

## 7.2 ปัญหาและข้อเสนอนแนะ

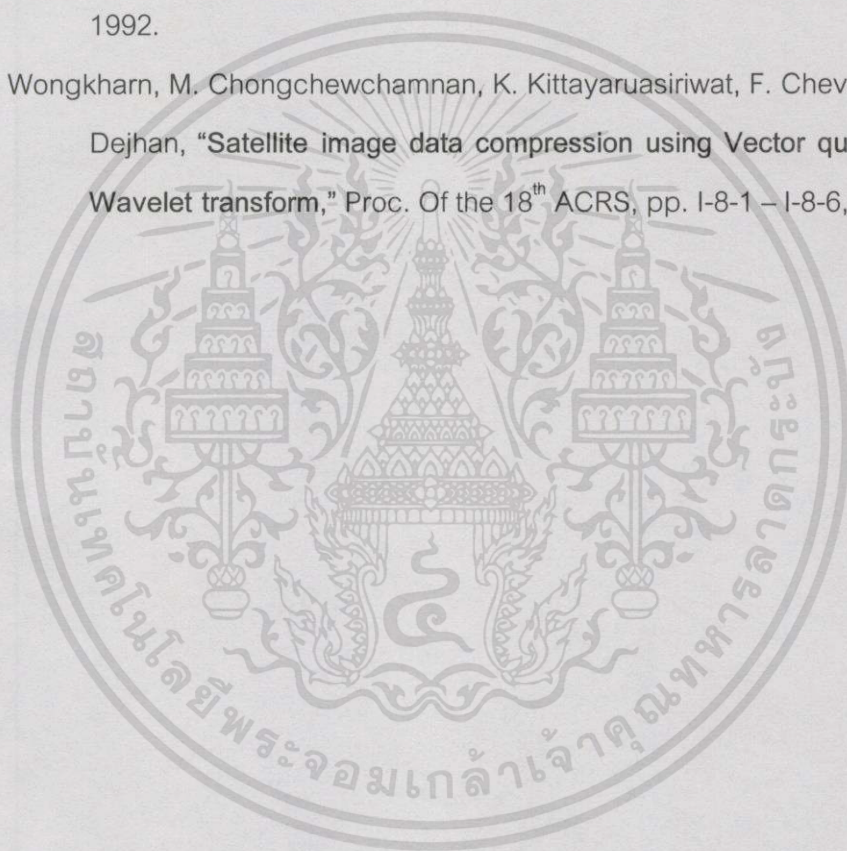
ในการสร้างตัวเก็บรหัส จะใช้อัลกอริทึม LBG ซึ่งเป็นการคำนวณแบบวนรอบ จึงใช้เวลาค่อนข้างนาน โดยเฉพาะสำหรับภาพที่มีขนาดใหญ่และมีความละเอียดสูง จะใช้เวลาในการคำนวณนาน แต่ถ้าเป็นเครื่องคอมพิวเตอร์ที่มีประสิทธิภาพสูงจะช่วยลดเวลาตรงส่วนนี้ได้

ข้อเสนอนแนะ จากการนำเสนอวิธีการบีบอัดข้อมูลภาพด้วยการเข้ารหัสเวกเตอร์ควอนไทซ์แบบอะแดปทีฟด้วยค่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเลตนั้น สามารถลดการสูญเสียของข้อมูลภาพได้เป็นอย่างดี โดยเพิ่มขบวนการตรวจวัดปริมาณข่าวสารของภาพในแต่ละแบนด์ย่อย และอัลกอริทึมในการกำหนดอัตราการบีบอัดข้อมูลให้กับภาพในแต่ละแบนด์ย่อย ซึ่งผู้ทำวิทยานิพนธ์คิดว่าขบวนการดังกล่าวสามารถนำไปประยุกต์ใช้หรือพัฒนาในงานทางด้านการบีบอัดข้อมูลภาพประเภทอื่น และพอจะเป็นแนวทางเริ่มต้นสำหรับนำไปพัฒนาปรับปรุงเพื่อให้ขบวนการบีบอัดข้อมูลภาพมีความเร็วและ ชาญฉลาดในการบีบอัดข้อมูลภาพในแต่ละภาพได้

## เอกสารอ้างอิง

- [1] ศิริพงษ์ วงษ์คาร. 2540. “การลดข้อมูลภาพถ่ายจากดาวเทียมด้วยการควอนไทซ์แบบเวกเตอร์.” วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- [2] ชัยชาญ มัคคุ่น. 2539. “การแก้ไขลักษณะของลายเส้นโดยใช้เวฟเล็ตทรานส์ฟอร์ม.” วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- [3] S. Mallat. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” IEEE Trans. On Patt. Anal. Machine Intel., Vol.11, No.7, 1989, pp674-693
- [4] C.K. Chui. **An Introduction to Wavelets.** Boston: Academic Press. 1992.
- [5] R.K. Young. **WAVELET THEORY AND ITS APPLICATIONS.** Boston: Kluwer Academic Publishers. 1994.
- [6] Oliver Rioul and Martin Vetterli. “Wavelets and Signal Processing.” IEEE SP Magazine, Oct. 1991. pp. 14-88.
- [7] R.G. Gallager. “Variations on a theme by Huffman.” IEEE Trans. Inform. Theory., Nov 1978. IT-25: 668-674.
- [8] H. Gharavi. “Conditional run-length and variable-length coding of digital picture.” IEEE Trans. Comm., June 1987. COM-35 (6): 671-677.
- [9] R.C. Gonzalez and R.E. Woods. **Digital Image Processing.** Addison – Wesley Publish Company, Inc., 1992.
- [10] M. Rabbani and P.W. Jones. **Digital Image Compression Techniques.** Vol. TT7 of SPIE Tutorial Texts in Optical Engineering. SPIE Press. 1991.
- [11] M.A. Sid – Ahmed. **Image Processing, Theory, Algorithm, and Architectures.** McGraw-Hill Inc., 1995.
- [12] J.W. Woods. **Subband Image Coding.** Boston: Kluwer Academic Publishers. 1991.
- [13] J.W. Woods and S.D.O’Neil. “Subband Coding of Images.” IEEE Trans. Acoust. Speech Signal Process., Oct. 1986. ASSP-34: 1278-1288.
- [14] Y.Linde, A. Buzo, and R.M. Gray. “Vector Quantization and Signal Compression.” IEEE Trans. Commun., vol COM-28., Jan 1980. Pp. 84-95.

- [15] A. Gersho and R.M. Gray, **Vector Quantization and Signal compression**. Boston: Kluwer Academic Publishes. 1992.
- [16] K.R. RAO and P. YIP. **DISCRETE COSINE TRANSFORM**. San Diego: Academic Press, Inc. 1990.
- [17] Phillip A. Laplante and Alexander D. Stoyenko. **Real-Time Imaging**. New York: The Institute of Electrical and Electronics Engineers, Inc. 1996.
- [18] M. Antonini, M. Barland, P. Matieu and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. An Image Proc.*, vol.1, pp.205-220, April 1992.
- [19] S. Wongkharn, M. Chongchewchamnan, K. Kittayaruasiriwat, F. Chevasuvit and K. Dejhan, "Satellite image data compression using Vector quantization on Wavelet transform," *Proc. Of the 18<sup>th</sup> ACRS*, pp. I-8-1 – I-8-6, Oct. 1997.



## ภาคผนวก

## ผลงานวิจัยที่ได้รับการตีพิมพ์

- [1] F. Cheevasuvit, A. Somboonkaew, K. Dejhan, S. Chitwong, S. Mitatha and S. Wongkharn. "Adaptive Vector Quantization Coding on Wavelet Informations for data Compression." Proceedings of the 19<sup>th</sup> Asian Conference on Remote Sensing, Manila, Philippines, Nov. 16-20, 1998.
- [2] K. Wongsritong, K. Kittayarusirawat, F. Cheevasuvit, K. dejhan, and A. Somboonkaew. "Contrast Enhancement Using Multipeak Histogram Equalization with Brightness Preserving". IEEE Asia-Pacific Conference on Circuits and System., Chiangmai, Thailand, November, 1998. pp455-458
- [3] F. Cheevasuvit, S. Wongkharn, M. Chongcheawchamnan, K. Dejhan,. "Sattelite Image Data Compression using Vector Quantization on Wavelet Informations." Proceedings of the 18<sup>th</sup> Asian Conference on Remote Sensing, Kuala Lumpur, Malaysia, Oct. 20-24, 1997.

## ประวัติผู้เขียน

นางสาวเกศกนก กฤตยาภาศิริวัฒน์ เกิดเมื่อวันที่ 26 ตุลาคม 2513 ที่จังหวัดขอนแก่น จบชั้นมัธยมศึกษาตอนปลายที่โรงเรียนบ้านไผ่ จังหวัดขอนแก่น เมื่อปี พ.ศ. 2532 และได้รับสิทธิ์พิเศษโครงการช่างเผือก ในการศึกษาต่อระดับปริญญาตรีของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง จนสำเร็จการศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ ปีการศึกษา 2535

หลังจากนั้นได้ทำการศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในปีการศึกษา 2540 ปัจจุบันทำงานในตำแหน่ง System Programmer 2 บริษัทกรุงเทพคอมพิวเตอร์ เซอร์วิสเซส จำกัด

