

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การบีบอัดข้อมูลภาพด้วยเวฟเล็ตทรานส์ฟอร์ม
และการเข้ารหัสภาพรูปแบบเสมือน

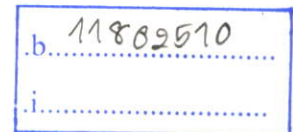
DATA COMPRESSION VIA WAVELET TRANSFORM
AND VIRTUAL PATTERN IMAGE CODING



ศธาวิฑูฒ ทองททา

KHATHAVUT TONGTAR

เลขหมู่.....
เลขทะเบียน..... 74528
วัน,เดือน,ปี..... 3 ต.ค. 2550



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมการวัดคุม
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.2550

**DATA COMPRESSION VIA WAVELET TRANSFORM
AND VIRTUAL PATTERN IMAGE CODING**

KHATHAVUT TONGTAR

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

2007

COPYRIGHT 2007

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

หัวข้อวิทยานิพนธ์	การบีบอัดข้อมูลภาพด้วยเวฟเล็ตทรานส์ฟอร์มและการเข้ารหัสภาพรูปแบบเสมือน
นักศึกษา	นายชราวุฒิ ทองทา
รหัสประจำตัว	45061229
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมการวัดคุม
พ.ศ.	2549
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.ดร.ฟูศักดิ์ ชิวสุวิทย์

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอวิธีการที่มีประสิทธิภาพเพื่อปรับปรุงคุณภาพของการบีบอัดภาพดิจิทัลโดยใช้อัลกอริทึม Visual Pattern Image Coding (VPIC) ในเวฟเล็ตโดเมน เวฟเล็ตเป็นพื้นฐานการบีบอัดภาพที่ให้คุณภาพที่สูงกว่าในการสร้างภาพขึ้นมาใหม่ด้วยค่า PSNR ที่อยู่ในระดับสูงและให้ค่า MSE ต่ำเมื่อเปรียบเทียบกับ VPIC แต่อย่างไรก็ตามวิธีการดังกล่าวให้ค่าอัตราการบีบอัดที่ค่อนข้างต่ำ ดังนั้นวิธีการที่นำเสนอจะเป็นการผสมผสานข้อดีของวิธีการเวฟเล็ตทรานส์ฟอร์มและ VPIC เพื่อให้ได้อัตราการบีบอัดที่สูงขึ้น ในขณะที่ยังคงรักษาคุณภาพของการสร้างภาพขึ้นมาใหม่ วิธีการที่ได้นำเสนอนี้มี 3 ขั้นตอน เริ่มจากใช้เวฟเล็ตทรานส์ฟอร์มเพื่อแยกค่าสัมประสิทธิ์ของความถี่ต่ำและความถี่สูงออกจากกัน ซึ่งค่าสัมประสิทธิ์ของแบนด์ LL จะถูกรักษาไว้ ในขณะที่ความถี่สูง(HL, LH, และ HH) จะถูกเลือกมาใช้งาน โดยอาศัยค่าเบี่ยงเบนมาตรฐาน ซึ่งปกติแบนด์ย่อยที่มีการกระจายตัวของข้อมูลสูงจะประกอบด้วยข้อมูลของขอบภาพที่สามารถใช้ในการปรับปรุงคุณภาพอัลกอริทึมของ VPIC ได้ ในที่สุดแบนด์ LL และแบนด์ที่ถูกเลือกก็จะถูกแยกเพื่อที่จะหาภาพที่ได้จากการบีบอัด จากผลลัพธ์ที่ได้จากการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอมีอัตราการบีบอัดสูงขึ้นในขณะที่ยังคงรักษาคุณภาพของภาพไว้และให้ค่า MSE ต่ำ

Thesis	Data Compression Via Wavelet Transform and Virtual Pattern Image Coding
Student	Mr. Khathavut Tongtar
Student ID	45061229
Degree	Master of Engineering
Program	Instrumentation Engineering
Year	2006
Thesis Advisor	Assoc.Prof.Dr. Fusak Cheevasuvit

ABSTRACT

This research proposes an efficient method to improve the quality of digital image compression using Visual Pattern Image Coding (VPIC) in wavelet domain. Wavelet based compression gives higher quality of the reconstructed image with high PSNR and less amount of minimum mean square error compared to VPIC. However, its compression ratio is reasonably low. Therefore, the proposed method combines the advantages obtained from both wavelet transform and VPIC to achieve high compression ratio while preserve the quality of the reconstructed image. The proposed algorithm can be separated in to three main processes. First, wavelet transform is used to decompose the low and high frequency coefficients. Then, the LL coefficients are reserved, while one of the high frequency sub-bands (HL, LH, and HH) is selected based on its Standard Deviation (SD). The sub-band having the highest information distribution usually contains proper edge information, which could be used to improve VPIC algorithm. Finally, the LL and the selected sub-bands are decomposed to obtain the result image. The experimental results show that the proposed compression algorithm can achieve high compression ratio while still maintain good image fidelity and minimum mean square error.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำและความกรุณาจากอาจารย์ที่ปรึกษา รศ.ดร. พุศัคดี ชิวสุวิทย์ ความช่วยเหลือและคำชี้แนะในการแก้ปัญหาตลอดจนให้ความรู้และประสบการณ์ที่ดีเป็นอย่างยิ่ง ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณครูอาจารย์ที่เคารพทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้และถ่ายทอดประสบการณ์อันเป็นประโยชน์แก่ผู้วิจัย

ขอขอบคุณ เพื่อน ๆ พี่น้องนักศึกษาทุกคนที่ให้ความช่วยเหลือ ให้คำแนะนำและกำลังใจที่ดีอย่างใกล้ชิดตลอดมา

คุณงามความดีและประโยชน์ที่ได้จากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแต่บิดามารดาอันเป็นที่รักและเคารพยิ่ง ตลอดจนผู้มีพระคุณทุกท่าน

ธราวุฒิ ทองทา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญรูป	VIII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	1
1.3 สมมติฐานของการศึกษา	1
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย	2
1.5 ขอบเขตการวิจัย	2
1.6 ขั้นตอนของการศึกษา.....	2
1.7 โครงสร้างของวิทยานิพนธ์.....	3
บทที่ 2 หลักการเบื้องต้นในการบีบอัดข้อมูล.....	4
2.1 การเข้ารหัสโดยการแปลง (Transform Coding).....	4
2.2 การเข้ารหัสแบนด์ย่อย (Sub-band Coding).....	5
2.3 ทฤษฎีข้อมูลและการเข้ารหัสเอนโทรปี (Information Theory and Entropy Coding)...	6
2.3.1 การวัดปริมาณข้อมูล (Measuring Information).....	6
2.3.2 การเข้ารหัสฮัฟแมน (Huffman Coding).....	7
2.3.3 การเข้ารหัสเลขคณิต (Arithmetic Coding).....	8
2.4 ทฤษฎีเวฟเล็ตทรานส์ฟอร์ม.....	10
2.4.1 พื้นฐานของการสเกลและการเลื่อนตำแหน่งของเวฟเล็ต.....	11
2.4.2 การแปลงเวฟเล็ตแบบแบ่งช่วง (Discrete Wavelet Transform).....	12
2.4.3 การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต (Mollat's Multiresolution Analysis or Mollat'MRA).....	13
2.4.4 การแปลงเวฟเล็ต โดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงก์ (Filter Bank).....	18

สารบัญ (ต่อ)

	หน้า
2.4.5 การฟิลเตอร์และการลดอัตราการซีกค่าตัวอย่างลง(Down sampling).....	20
2.4.6 การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง.....	23
2.4.7 การฟิลเตอร์และการเพิ่มอัตราการซีกค่าตัวอย่าง (Up sampling).....	24
2.4.8 การแปลงเวฟเล็ตกับข้อมูลภาพ.....	26
บทที่ 3 Visual Pattern Image Coding (VPIC).....	28
3.1 การเข้ารหัสข้อมูลภาพแบบ VPIC กับภาพเชิงตัวเลขแบบ 2 มิติ.....	28
3.2 ประเภทของบล็อกย่อย.....	29
3.3 การเข้ารหัส.....	30
3.4 การถอดรหัส.....	31
3.5 การวัดประสิทธิภาพของการบีบอัดและการวัดประสิทธิภาพของภาพที่ได้หลังจาก การเรียกคืน.....	32
บทที่ 4 การเข้ารหัสแบบ VPIC บน Wavelet Domain.....	34
4.1 การเลือกแบนด์เพื่อเข้ารหัสแบบ VPIC และการเข้ารหัสแบบ VPIC.....	34
4.1.1 การเลือกแบนด์เพื่อเข้ารหัสแบบ VPIC.....	34
4.1.2 การเข้ารหัสแบบ VPIC.....	35
4.2 การถอดรหัสแบบ VPIC	37
4.3 การวัดประสิทธิภาพการบีบอัดและประสิทธิภาพของภาพที่ได้หลังจากการบีบอัด..	37
บทที่ 5 การทดลองและผลการทดลอง.....	38
5.1 ข้อมูลภาพต้นฉบับที่ใช้ในการทดลอง.....	38
5.2 การหาค่าเบี่ยงเบนมาตรฐาน(Standard Deviation).....	39
5.3 ผลการเข้ารหัสแบบ VPIC.....	43
5.4 การเรียกคืนภาพ.....	45
บทที่ 6 สรุปผลการทดลองและข้อเสนอแนะ.....	49
6.1 สรุปและวิเคราะห์ผลการดำเนินงานวิจัย.....	49
6.2 ปัญหา.....	49

สารบัญ (ต่อ)

	หน้า
6.3 แนวทางในการพัฒนาต่อ.....	50
เอกสารอ้างอิง	51
ภาคผนวก ก ผลงานวิจัยที่ได้รับการตีพิมพ์.....	52
ประวัติผู้เขียน	58

สารบัญตาราง

ตารางที่	หน้า
2.1 การแบ่งขอบเขตความน่าจะเป็นสะสม.....	8
3.1 รายละเอียดและจำนวนบิตของข้อมูล.....	30
3.2 คุณภาพของภาพ.....	33
5.1 แสดงค่าเบี่ยงเบนมาตรฐานของภาพในแบนด์ต่างๆ.....	43
5.2 ขนาดข้อมูลของภาพในแบนด์เมื่อผ่านการเข้ารหัส VPIC.....	44
5.3 ค่า MSE.....	46
5.4 แสดงค่าผลลัพธ์ต่างๆ ของภาพต้นฉบับเมื่อทำการบีบอัดข้อมูลแบบการเข้ารหัสภาพแบบ VPIC โดยมีค่า gradient เท่ากับ 10 และ 30.....	46
5.5 แสดงค่า MSE ของภาพที่เข้ารหัสแบบ VPIC ที่มีค่า gradient 10 และ 30 กับภาพที่ไม่ได้ผ่านการเข้ารหัส.....	47
5.6 แสดงขนาดของข้อมูลที่ได้จากวิธีการต่างๆ.....	47
5.7 ค่า MSE ของภาพที่ได้จากการเข้ารหัสแบบ VPIC และวิธีในงานวิจัย.....	48

สารบัญรูป

รูปที่	หน้า
2.1 การเข้ารหัสโดยการแปลง.....	5
2.2 Huffman Coding.....	7
2.3 แสดงการแบ่งขอบเขตตามข้อมูลที่เพิ่มขึ้น.....	9
2.4 แสดงคลื่นเวฟเล็ด.....	10
2.5 ระบุเวลาและเวลาที่เปลี่ยนแปลงช่วงเวลาและเวลาที่สัมพันธ์กันของการแปลง เวฟเล็ด.....	12
2.6 แสดงกลุ่มของปริภูมิเวกเตอร์.....	15
2.7 แสดงปริภูมิเริ่มต้นที่ถูกแผ่คลุม.....	16
2.8 แสดงสัญลักษณ์ Down sampling.....	21
2.9 แสดงตัวกรองความถี่ตามหลักฟิลเตอร์เบงก์.....	21
2.10 แสดงฟิลเตอร์เบงก์ที่ระดับ C_{j-1}	21
2.11 แสดงการแยกสัญญาณที่ V_3	22
2.12 แสดงการแยกสัญญาณทางด้านความถี่.....	22
2.13 แสดงสัญลักษณ์การเพิ่มอัตราการซีกค่าเป็น 2 เท่า.....	24
2.14 แสดงไดอะแกรมของการซีกค่าโดยใช้หลักการของฟิลเตอร์เบงก์.....	25
2.15 แสดงการสังเคราะห์สัญญาณจากระดับ $j-1$ ไปสู่ j	25
2.16 แสดงแผนผังของการแปลงเวฟเล็ด.....	26
2.17 แสดงลักษณะของการแบ่งแบนด์ย่อยของภาพ.....	26
2.18 แสดงการแปลงเวฟเล็ดหลายๆ ระดับ.....	27
2.19 แสดงลักษณะการแบ่งแบนด์ย่อยหลายๆ ระดับ.....	27
3.1 ฮีสโตแกรมของภาพต้นฉบับ.....	29
3.2 รูปแบบของบล็อกแพทเทินในหมายเลขต่างๆ.....	30
4.1 ไดอะแกรมการเลือกแบนด์และการเข้ารหัสแบบ VPIC.....	35
4.2 รูปแบบของข้อมูลที่ผ่านการเข้ารหัสแบบ VPIC.....	36
4.3 รูปแบบข้อมูลเมื่อผ่านการบีบอัด.....	36
4.4 ไดอะแกรมการถอดรหัสแบบ VPIC และการเรียกคืนภาพ.....	37
5.1 ภาพต้นฉบับที่นำมาใช้ในการทดลอง.....	38
5.2 ฮีสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Lena.....	39
5.3 ฮีสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Girl	40

สารบัญรูป(ต่อ)

รูปที่	หน้า
5.4 ฮีสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Women.....	41
5.5 ฮีสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Cameraman.....	42
5.6 ภาพที่ได้หลังจากนำแบนด์ที่มีค่า STD สูงสุดมาบีบอัดด้วย VPIC.....	43
5.7 ผลลัพธ์ที่ได้จากการเรียกคืน.....	45

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันจะเห็นได้ว่าระบบเครือข่ายการสื่อสารข้อมูลนั้นมีการพัฒนาอย่างต่อเนื่องเพื่อให้สามารถรองรับความต้องการในการรับหรือส่งข้อมูลต่างๆของผู้ใช้งานในระบบนั้นๆ จึงเกิดการค้นคว้าวิจัยในด้านต่างๆ เพื่อที่จะทำให้การสื่อสารข้อมูลมีความสะดวกรวดเร็วเพียงพอต่อความต้องการของผู้ใช้ ซึ่งข้อมูลภาพดิจิทัลเป็นข้อมูลหนึ่งที่ได้รับคามนิยมในการใช้งาน แต่ภาพดิจิทัลที่คืนั้นจำเป็นที่จะต้องมีความละเอียดสูง ทำให้ข้อมูลของภาพจะมีจำนวนมากตามไปด้วย ซึ่งจะส่งผลกระทบต่อ เช่น เมื่อต้องการเก็บภาพดิจิทัลไว้ในสื่อบันทึกต่างๆ หรือต้องการส่งภาพดิจิทัลจากสถานที่หนึ่งไปยังอีกที่หนึ่งผ่านทางระบบเครือข่ายคอมพิวเตอร์ อินเทอร์เน็ต หรือระบบโทรศัพท์มือถือ จะส่งผลให้ใช้เวลาในการประมวลผลและรับส่งข้อมูลมากตามไปด้วย

การบีบอัดข้อมูล(Data Compression) จึงเป็นอีกแนวทางหนึ่งที่ถูกนำมาใช้กับการสื่อสารข้อมูลเพื่อตอบสนองในด้านของความเร็วในการสื่อสารและขนาดของพื้นที่ในการจัดเก็บ การบีบอัดข้อมูลเป็นวิธีการลดขนาดของข้อมูลที่ต้องการจัดเก็บหรือทำการส่งไปยังผู้รับ เมื่อข้อมูลที่ใช้ในการสื่อสารมีขนาดเล็กก็จะทำให้ใช้เวลาในการสื่อสารน้อยเช่นกัน ซึ่งในปัจจุบันการบีบอัดข้อมูลมีประสิทธิภาพที่ดีขึ้นและมีการประยุกต์ใช้กับข้อมูลภาพ เสียง วิดีโอ

การบีบอัดข้อมูลภาพด้วยวิธี Visual Pattern Image Coding (VPIC) บนเวฟเล็ตโดเมน เป็นการประยุกต์วิธีการบีบอัดข้อมูลแบบการเข้ารหัสภาพแบบ VPIC มาใช้กับข้อมูลที่ถูกแบ่งความถี่โดยวิธีเวฟเล็ต

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

- 1.2.1 เพื่อปรับปรุงและประยุกต์วิธีการที่มีอยู่โดยอาศัยสมมติฐานตามหลักการคิดเชิงวิจัยและพัฒนา
- 1.2.2 ประยุกต์ใช้เทคนิคการบีบอัดข้อมูลภาพด้วยเทคนิคเวฟเล็ตทรานส์ฟอร์มมาใช้ร่วมกับ VPIC
- 1.2.3 ประเมินประสิทธิภาพและวัดคุณภาพในเชิงปริมาณ รวมทั้งสรุปวิเคราะห์ผลลัพธ์จากการสืบค้นคืนข้อมูลภาพของงานวิจัย

1.3 สมมติฐานของการศึกษา

ในงานวิจัยนี้ศึกษาวิธีการบีบอัดข้อมูลแบบเวฟเล็ต เพื่อหาความสำคัญของแบนด์ผลลัพธ์ ทั้ง 4 แบนด์โดยหาวิธีการที่สามารถนำมาใช้ในการเลือกแบนด์ที่เหมาะสมและนำแบนด์ที่ได้มาบีบอัดข้อมูลแบบ VPIC และทำการเปรียบเทียบผลลัพธ์ของการบีบอัดข้อมูลระหว่างขนาดของข้อมูล ที่ผ่านการบีบอัดกับค่าความผิดเพี้ยนของภาพเมื่อทำการเรียกคืน

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

เพื่อที่จะให้บรรลุผลตามวัตถุประสงค์ในการบีบอัดข้อมูลภาพ จะต้องอาศัยหลักการและทฤษฎี ดังต่อไปนี้

1. การประมวลผลภาพเบื้องต้น
2. การแปลงเวฟเล็ตทรานส์ฟอร์ม
3. การเข้ารหัสข้อมูลภาพด้วยวิธีการ VPIC

1.5 ขอบเขตการวิจัย

งานวิจัยนี้ทำการศึกษาหลักการการเข้ารหัส(บีบอัด) และการถอดรหัส(คลาย) ที่อยู่บนพื้นฐานของการเข้ารหัสแบบ VPIC บนเวฟเล็ตโดเมน โดยจะทำการเก็บแบนด์ LL ไว้และทำการนำแบนด์ LH ,HL ,HH ที่ได้จากการแปลงเวฟเล็ตมาเพียงหนึ่งแบนด์ โดยจะนำแบนด์ที่มีการกระจายตัวของข้อมูลมากที่สุดมาทำการเข้ารหัสแบบ VPIC เพื่อลดขนาดข้อมูลของแบนด์นั้น ส่วนแบนด์ที่เหลือนั้นสามารถตัดทิ้งไปได้เนื่องจากมีการกระจายตัวของข้อมูลต่ำเสมือนว่ามีความสำคัญน้อยกว่า เมื่อทำการเรียกคืนภาพจะทำการถอดรหัสข้อมูลด้วย VPIC และนำข้อมูลที่ได้จากการถอดรหัสมาทำการแปลงกลับ Wavelet ร่วมกับแบนด์ LL เพียงสองแบนด์เท่านั้นโดยแบนด์ที่เหลือให้ค่าเป็นศูนย์ จากนั้นนำภาพที่ได้จากการแปลงกลับ Wavelet มาทำการหาค่าความผิดเพี้ยนจากภาพต้นฉบับเปรียบเทียบกับขนาดของข้อมูลที่ได้จากการบีบอัดแบบการเข้ารหัสภาพแบบ VPIC บนเวฟเล็ตโดเมน ว่ามีค่าความผิดเพี้ยนมากน้อยเพียงใดต่ออัตราการลดขนาดข้อมูล

1.6 ขั้นตอนของการศึกษา

- 1.6.1 กำหนดวัตถุประสงค์และขอบเขตของงานวิจัย ว่าต้องการนำเสนอหรือศึกษางานวิจัยในหัวข้อนี้มากน้อยแค่ไหนและอย่างไร
- 1.6.2 ศึกษาค้นคว้าทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการบีบอัดข้อมูลภาพ แล้ววิเคราะห์ข้อดี ข้อด้อยและประเด็นที่น่าสนใจของงานวิจัยอื่นๆที่สามารถนำมาปรับปรุงและประยุกต์ใช้ เพื่อก่อให้เกิดประโยชน์แก่งานวิจัยนี้ได้

- 1.6.3 ตั้งสมมติฐานของการศึกษาและกำหนดวางแผนความคิดของงานวิจัย โดยมีการอ้างอิงทฤษฎีหรือหลักการที่เกี่ยวข้องเพื่อที่จะบรรลุตามวัตถุประสงค์ที่ได้กำหนดไว้
- 1.6.4 เตรียมข้อมูลภาพ เพื่อนำมาใช้เป็นข้อมูลพื้นฐานสำหรับการวิจัย โดยใช้ภาพที่นิยมใช้ในการบีบอัดข้อมูลเพื่อนำมาใช้ในการตรวจสอบประสิทธิภาพของงานวิจัยได้อย่างเหมาะสม
- 1.6.5 พัฒนาโปรแกรมเพื่อทดลองบีบอัดข้อมูลภาพด้วยอัลกอริทึมตามแนวคิดที่งานวิจัยนี้ได้นำเสนอ
- 1.6.6 ทำการทดลองบีบอัดข้อมูลภาพด้วยโปรแกรมที่พัฒนาขึ้นมา กับข้อมูลภาพที่ได้เตรียมไว้ โดยในขณะที่ทำการทดลองจะเก็บข้อมูลและผลลัพธ์ของทุกๆ ขั้นตอนเอาไว้ เพื่อนำมาวิเคราะห์ปรับปรุงงานวิจัยต่อไป
- 1.6.7 นำผลลัพธ์จากการทดลองมาวิเคราะห์และประเมินงานวิจัย ทั้งในแง่ของคุณภาพและประสิทธิภาพ โดยเปรียบเทียบกับงานวิจัยอื่นๆ แล้วสรุปผลเพื่อนำเสนอผลงานวิจัย

1.7 โครงสร้างของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้แบ่งออกเป็น 6 บท แต่ละบทประกอบด้วยเนื้อหาดังต่อไปนี้

บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา ความมุ่งหมายและวัตถุประสงค์ของการศึกษา สมมติฐานของการศึกษา รวมทั้งทฤษฎีหรือแนวคิดที่ใช้ในการศึกษา ขอบเขตของการศึกษา ขั้นตอนของการศึกษา

บทที่ 2 กล่าวถึงทฤษฎีเบื้องต้นสำหรับการบีบอัดข้อมูลภาพ

บทที่ 3 กล่าวถึงทฤษฎีของ VPIC

บทที่ 4 กล่าวถึงวิธีการประยุกต์ใช้ VPIC ร่วมกับเวฟเล็ตทรานส์ฟอร์มในการบีบอัดข้อมูลภาพ

บทที่ 5 ผลการทดลอง

บทที่ 6 สรุปผลการทดลอง ข้อเสนอแนะ และแนวทางที่สามารถปรับปรุงต่อไปในอนาคต

บทที่ 2

หลักการเบื้องต้นในการบีบอัดข้อมูล

ในปัจจุบันการสื่อสารมีการพัฒนาอย่างรวดเร็วทำให้การรับส่งข้อมูลมีประสิทธิภาพมากขึ้นด้วยเหตุนี้การบีบอัดข้อมูลจึงเข้ามามีบทบาทสำคัญในการลดขนาดของข้อมูลที่ต้องการส่งไปยังผู้รับหรือข้อมูลที่ต้องการจัดเก็บเพื่อลดเวลาในการสื่อสารหรือประหยัดพื้นที่ในการจัดเก็บข้อมูล การบีบอัดข้อมูลมีอยู่หลายวิธีด้วยกัน วิธีที่กล่าวถึงจะเป็นการบีบอัดข้อมูลแบบการเข้ารหัสในรูปแบบต่างๆ

2.1 การเข้ารหัสโดยการแปลง (Transform Coding)

สมมติว่าบล็อกของสัญญาณที่สุ่มตัวอย่าง (Sampling) มาได้ และต้องการเข้ารหัสภายในจำนวนบิตที่กำหนด ให้ เป็นเวกเตอร์ตัวอย่างนี้

$$X = (X_1, X_2, X_3, \dots, X_k)'$$
 (2.1)

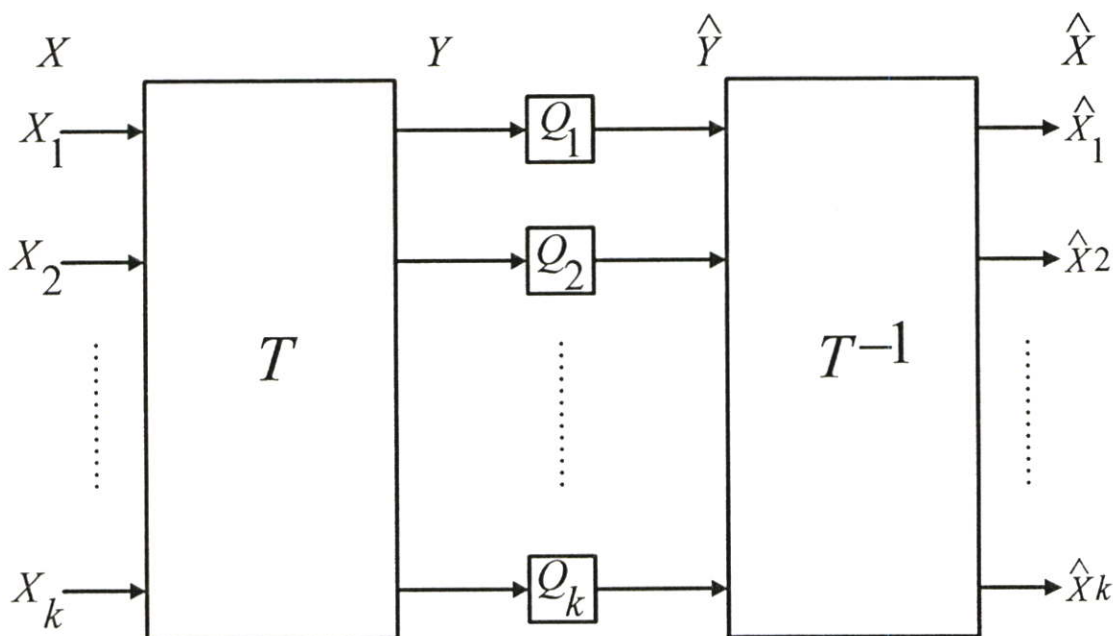
โดยทั่วไปจะพบว่า แต่ละรายการในเวกเตอร์นี้จะมีความสัมพันธ์ (Correlation) กัน การควอนไตซ์แยกกันทีละรายการนั้น จะไม่มีประสิทธิภาพเพียงพอในกรณีที่กำหนดจำนวนโควตาของเนื้อที่จัดเก็บเป็นจำนวนบิตมาให้

แนวคิดของการเข้ารหัส คือ การแปลงเชิงเส้น (Linear Transformation) กับชุดข้อมูลที่มีอยู่ได้แก่เวกเตอร์ ซึ่งจะทำได้เวกเตอร์ใหม่คือ ที่มีจำนวนเท่ากันเรียกว่า Transform Coefficients โดยที่คุณสมบัติคือ Coefficients เหล่านี้จะมีความสัมพันธ์กันน้อยกว่าเวกเตอร์เดิม (ซึ่งก็คือ X) นอกจากนี้ข้อมูลอาจจะมีลักษณะที่ "Compact" มากขึ้น ซึ่งก็คืออาจจะได้ส่วนที่สำคัญโดยใช้ Coefficients เพียงไม่กี่ตัว ซึ่งคาดว่า การลดความซ้ำซ้อนลงนั้น จะช่วยให้การควอนไตซ์มีประสิทธิภาพมากขึ้น

จากการศึกษาพบว่า ยิ่งตัวแปรต่างๆ ไม่มีความสัมพันธ์หรือเป็นอิสระต่อกันมากขึ้นเท่าไร สเกลาร์ควอนไตซ์เซชัน (Scalar Quantization / SQ) ก็จะมีประสิทธิภาพที่ดีขึ้น เมื่อเทียบกับเวกเตอร์ควอนไตซ์เซชัน (Vector Quantization / VQ) หรือก็คือ สัดส่วนประสิทธิภาพของ VQ ต่อ SQ จะลดลง

การ Reconstruction ของสัญญาณซึ่งอยู่ในรูปของเวกเตอร์นั้น จะทำการแปลงกลับ (Inverse Transform) สัญญาณที่ผ่านการแปลงไป (Forward Transform) และควอนไตซ์แล้ว รูปที่ 2.1 แสดงโครงสร้างของการเข้ารหัสโดยการแปลง โดยที่ T คือเมตริกซ์ขนาด $K \times K$ ซึ่งทำการ

แปลงเชิงเส้น สัมประสิทธิ์แต่ละตัวที่ได้จากการแปลงอาจจะต้องการ Quantizer ที่แตกต่างกัน เนื่องจากมี pdf(Point Density Function) ที่ต่างกันนั้นหมายความว่า แต่ละตัวต้องการจองเนื้อที่จัดเก็บไม่เท่ากัน จึงต้องมีเรื่องของ Bit Allocation เข้ามาเกี่ยวข้องซึ่งว่าด้วยการจองเนื้อที่จำนวน N bits สำหรับเวกเตอร์โดยที่จัดสรรให้เกิดประโยชน์สูงสุด



รูปที่ 2.1 การเข้ารหัสโดยการแปลง

เหตุผลที่มาสันนิษฐานการเข้ารหัสโดยการแปลง ก็คือระบบการรับรู้ทางประสาทหูและตาของมนุษย์จะรับรู้ได้ในโดเมนแปลงแบบหนึ่งและจะรู้สึกว่ามีคุณภาพดีถ้ามีการเข้ารหัสที่ดีในโดเมนนั้น การแปลงที่นิยมมากคือ การแปลงฟูริเยร์ หรือการแปลงดิสครีตโคไซน์ ซึ่งนิยมใช้ในการแปลงภาพและเสียง โดยที่อัตราบิต(Bit rate) เท่ากัน สัญญาณที่เข้ารหัสโดยการแปลงจะให้คุณภาพที่ดีกว่า ซึ่งการวัดค่าผิดเพี้ยน (Distortion Measuring) นิยมใช้ Overall Square Error Distortion

2.2 การเข้ารหัสแบนด์ย่อย (Sub-band Coding)

การเข้ารหัสแบนด์ย่อย [1] มีหลักการคือ ภาพต้นฉบับจะถูกกรองตามช่วงความถี่เพื่อสร้างเป็นภาพย่อยซึ่งภาพเหล่านี้เรียกว่า ภาพแบนด์ย่อย เนื่องจากว่าแต่ละภาพแบนด์ย่อยได้ถูกลดความกว้างของแถบ (Bandwidth) เมื่อเปรียบเทียบกับภาพต้นฉบับทำให้สามารถลดจำนวนการสุ่ม (sampling) ได้ซึ่งสามารถทำได้โดยการกรอง (Filtering) ส่วนการสุ่มนี้เรียกว่า ขั้นตอนการวิเคราะห์ (Analysis Stage) หลังจากนั้นภาพแบนด์ย่อยที่ได้จะถูกเข้ารหัสด้วยตัวเข้ารหัสต่างๆ ที่อัตราบิตต่างๆ หรือแม้แต่วิธีการเข้ารหัสแบบต่างๆ กันก็สามารถใช้กับแต่ละแบนด์ย่อยได้

ในการสร้างภาพกลับคืนนั้นสามารถทำได้โดยการเพิ่มจำนวนการสุ่มของภาพแบนด์ย่อยที่ได้ถอดรหัสแล้ว จากนั้นนำมาผ่านตัวกรองแล้วรวมภาพที่สร้างขึ้นแต่ละแบนด์ย่อยเข้าด้วยกัน กระบวนการนี้เรียกว่า ขั้นตอนการสังเคราะห์ (Synthesis stage)

จะสังเกตได้ว่าการสร้างภาพแบนด์ย่อยจะไม่ทำให้เกิดการลดลงของข้อมูลเนื่องจากจำนวนของการสุ่มของทุกแบนด์ย่อยรวมกันแล้วยังคงเท่ากับภาพต้นฉบับ แต่จะมีข้อดีในแง่ที่แต่ละภาพแบนด์ย่อยนั้นสามารถที่จะทำการเข้ารหัสได้ด้วยประสิทธิภาพที่ดีกว่าภาพต้นฉบับ

2.3 ทฤษฎีข้อมูลและการเข้ารหัสเอนโทรปี (Information Theory and Entropy Coding)

2.3.1 การวัดปริมาณข้อมูล (Measuring Information)

Shannon [2] ได้นำเสนอวิธีการวัดปริมาณของข้อมูลเอาไว้ว่า เหตุการณ์ที่เกิดขึ้นโดยมีความน่าจะเป็นเท่ากับ $P(E)$ มีปริมาณข้อมูล (Unit of Information) ดังนี้

$$I(E) = \log \frac{1}{P(E)} = -\log P(E) \quad (2.2)$$

$I(E)$ เรียกว่า self-information ของเหตุการณ์ E เรียกค่านี้ว่า Entropy ของชุดข้อมูล มีหน่วยเป็นจำนวนบิต ใช้อธิบายถึงความไม่แน่นอนหรือ Randomness ของชุดข้อมูล หลักการที่นำมาใช้กับการลดข้อมูลคือ การเข้ารหัสเหตุการณ์ที่เกิดขึ้นบ่อยๆ ด้วยจำนวนบิตน้อยๆ และเข้ารหัสเหตุการณ์ที่เกิดขึ้นไม่บ่อยด้วยจำนวนบิตที่มาก หรือเรียกวิธีการนี้ว่า variable length coding สำหรับการนำมาใช้กับข้อมูลภาพก็คือ เราจะให้ค่าความสว่าง (Intensity) ของจุดภาพที่เกิดขึ้นเป็นเหมือนเหตุการณ์ E ซึ่งค่าความสว่างแต่ละระดับก็จะมีค่าความน่าจะเป็นในการเกิดเป็น $P(E)$ (พิจารณาจาก Histogram) เมื่อนำมาใช้กับหลักการของ variable length coding ก็จะทำให้เราเข้ารหัสค่าจุดภาพแต่ละจุดด้วยจำนวนบิตต่างๆ กันได้โดยไม่จำเป็นต้องใช้จำนวน n บิตเท่ากันหมด (เช่น $n = 8$) ดังนั้นปริมาณข้อมูลเฉลี่ยจะมีค่าดังนี้

$$H(z) = -\sum_{j=1}^J P(a_j) \log P(a_j) \quad (2.3)$$

เมื่อ $\{a_1, a_2, \dots, a_j\}$ เป็นเซตของ Symbol ทั้งหมดของข้อมูลต้นทาง

ในที่นี้แต่ละ Symbol คือแต่ละระดับความสว่าง จุดประสงค์ของ variable length coding คือทำให้ปริมาณข้อมูลเฉลี่ยต่อจุดน้อยที่สุด เทียบกับการเข้ารหัสแบบความยาวคงที่ (Fixed-length coding) ที่มีความยาวเฉลี่ยเป็น อย่างไรก็ตามค่าเฉลี่ยนี้เป็นค่า lower bound เท่านั้นเนื่องจากเวลาที่

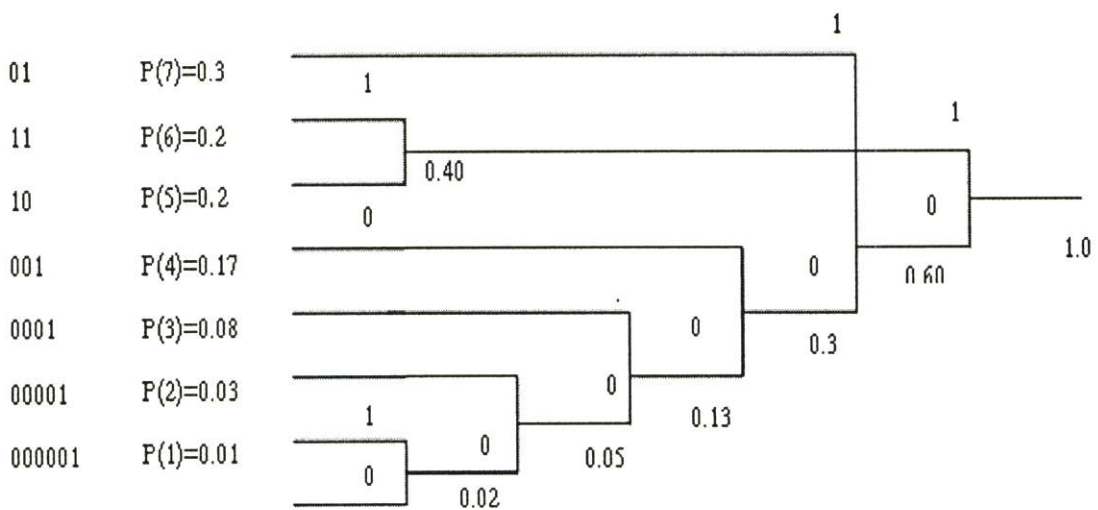
เราเข้ารหัส จำนวนบิตสำหรับแต่ละ Symbol จะต้องเป็นจำนวนเต็มเท่านั้น วิธีการเข้ารหัสโดยใช้ข้อมูลทางสถิติมีหลายวิธีเช่น Huffman coding Arithmetic coding

2.3.2 การเข้ารหัสฮัฟแมน (Huffman Coding)

การเข้ารหัสฮัฟแมน [3] เป็นการลดข้อมูลแบบที่ไม่มีการสูญเสียวิธีหนึ่ง หลักการที่สำคัญของการเข้ารหัสฮัฟแมนคือข้อมูลที่มีความน่าจะเป็นสูงกว่าจะถูกเข้ารหัสด้วยจำนวนบิตที่น้อยกว่า ดังแสดงในรูปที่ 2.2

การเข้ารหัสฮัฟแมน สามารถอธิบายโดยย่อได้ดังนี้

1. ทำการนับความถี่เพื่อหาความน่าจะเป็นของข้อมูลทุกตัวและเรียงลำดับข้อมูลตามความน่าจะเป็นจากน้อยไปหามาก
2. นำ Symbol ที่มีความน่าจะเป็นต่ำสุด 2 ตัว มารวมค่าความน่าจะเป็นเดียวกัน เพื่อสร้างเป็นโหนดใหม่ แล้วนำความน่าจะเป็นใหม่ที่ได้อาเรียงลำดับใหม่รวมกับค่าความน่าจะเป็นที่เหลือ
3. กำหนดค่าไบนารีให้กับโหนดคู่ที่ถูกกำหนดให้เป็นเส้นทางของโหนดใหม่นั้น โดยโหนดหนึ่งกำหนดให้มีค่าเป็น 0 อีกโหนดหนึ่งมีค่าเป็น 1
4. ทำซ้ำตามขั้นตอนที่ 2 และ 3 จนกระทั่ง ได้ความน่าจะเป็นสุดท้ายค่าเดียวคือ 1.0



2.3.3 การเข้ารหัสเลขคณิต (Arithmetic Coding)

ในการเข้ารหัสเลขคณิตนั้น [4] ข่าวสารต้นฉบับจะถูกแทนด้วยช่วงระยะห่างของจำนวนจริงระหว่าง 0 และ 1 โดยข้อมูลแต่ละตัวของข่าวสารจะทำให้ช่วงระยะห่างนั้นแคบลง ซึ่งระยะห่างที่แคบลงนี้จะทำให้ใช้จำนวนของบิตมากขึ้น ตัวข้อมูลที่ต่อเนื่องกันของข่าวสารจะลดขนาดของระยะห่างตามความน่าจะเป็นของข้อมูล ข่าวสารที่มีความน่าจะเป็นสูงจะทำให้ระยะห่างถูกลดลงได้น้อยกว่าข่าวสารที่มีความน่าจะเป็นต่ำ ดังนั้นจึงทำให้ข่าวสารที่มีความน่าจะเป็นสูงใช้จำนวนบิตในการเข้ารหัสน้อยกว่า

ขบวนการจะเริ่มจากข้อมูลต้นแบบที่ไม่ได้เรียงลำดับความน่าจะเป็น จากนั้นจะทำการแบ่งช่วงระยะห่างระหว่าง 0 กับ 1 เป็นช่วงย่อยๆ ตามความน่าจะเป็นสะสม (Cumulative Probability) เช่น กำหนดให้ข้อมูลต้นฉบับคือ A, B, C, D และ! มีความน่าจะเป็นเท่ากับ 0.3, 0.1, 0.2, 0.3 และ 0.1 ตามลำดับ การแบ่งช่วงระยะห่างระหว่าง 0 กับ 1 ดังตารางที่ 2.1

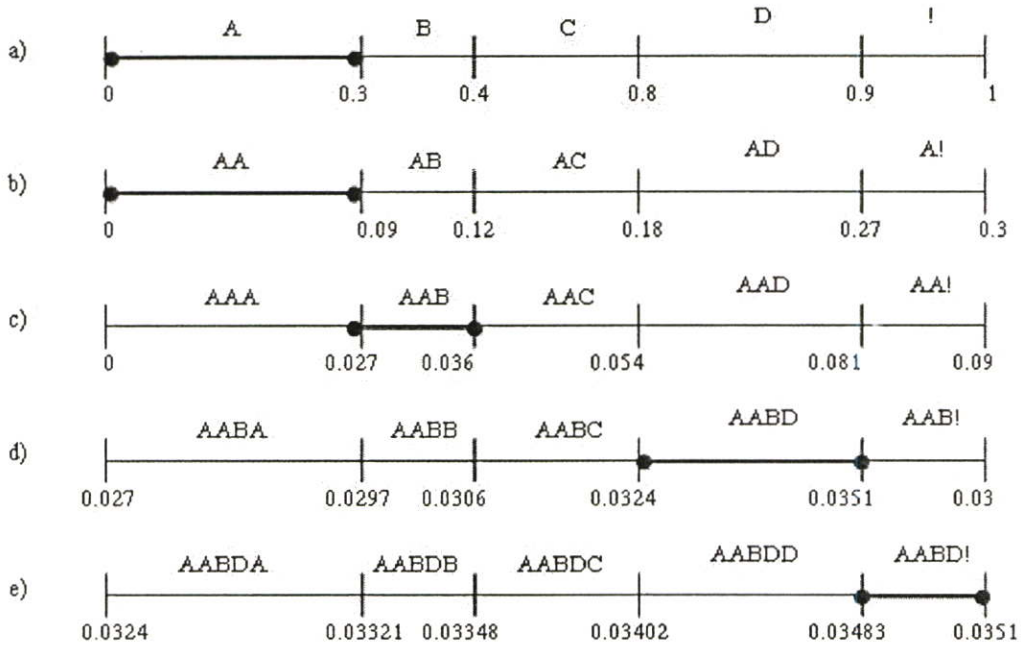
ตารางที่ 2.1 การแบ่งขอบเขตความน่าจะเป็นสะสม

ข้อมูลต้นแบบ	ความน่าจะเป็น (Probability)	ความน่าจะเป็นสะสม (Cumulative Probability)	ขอบเขต(range)
A	0.3	0.3	[0, 0.33)
B	0.1	0.4	[0.3, 0.4)
C	0.2	0.6	[0.4, 0.6)
D	0.3	0.9	[0.6, 0.9)
!	0.1	1.0	[0.9, 1.0)

ตัวอย่างการเข้ารหัส AABD! แสดงไว้ในรูปที่ 2.3 โดยมีขั้นตอนดังต่อไปนี้

1. A ตัวแรกจะลดขอบเขตของระยะห่างระหว่าง 0 กับ 1 ลงเป็น [0,0.3] ตามรูปที่ 2.3a)
2. A ตัวที่สองจะลดขอบเขตของระยะห่างระหว่าง 0 กับ 0.3 ลงเป็น [0,0.09] ตามรูปที่ 2.3b)
3. B จะลดขอบเขตระยะห่างระหว่าง 0 กับ 0.09 ลงเป็น [0.027,0.036] ตามรูปที่ 2.3c)
4. D จะลดขอบเขตระยะห่างระหว่าง 0.027 กับ 0.036 ลงเป็น [0.0324,0.0351] ตามรูปที่ 2.3d)
5. ! จะลดขอบเขตของระยะห่างระหว่าง 0.027 กับ 0.036 ลงเป็น [0.0324,0.0351] ตามรูปที่ 2.3e)

โดยที่ $[a, b]$ หมายถึงระยะห่างที่อยู่ในช่วงมากกว่าหรือเท่ากับ a แต่น้อยกว่า b เป็นต้น ดังนั้นข้อมูล AABD! สามารถเข้ารหัสได้โดยการแทนด้วยช่วงระยะห่างหรือตัวเลขจำนวนใดๆ ที่อยู่ในช่วง $[0.03483, 0.0351]$



รูปที่ 2.3 แสดงการแบ่งขอบเขตตามข้อมูลที่เพิ่มขึ้น

สมมติว่าตัวถอดรหัสดำเนินการได้รับข้อมูลรหัสเป็น $i = 0.035$ ดังนั้นกระบวนการถอดรหัสจะมีขั้นตอนดังต่อไปนี้

1. เริ่มจากแบ่งช่วง $[0, 1]$ ตามความน่าจะเป็น แล้วพบว่า $i = 0.035$ อยู่ในช่วง $[0, 0.3]$ ซึ่งเป็นช่วงของตัวอักษร A ดังนั้นจะได้อักษรตัวแรกคือ A จากนั้นลดขอบเขตของ $[0, 1]$ ลงเป็น $[0, 0.3]$
2. แบ่งช่วง $[0, 0.3]$ ตามความน่าจะเป็น แล้วพบว่า $i = 0.035$ อยู่ในช่วง $[0, 0.09]$ ซึ่งเป็นช่วงตัวอักษร A ดังนั้นจะได้อักษรตัวที่สองคือ A จากนั้นลดขอบเขตของ $[0, 0.3]$ ลงเป็น $[0, 0.09]$
3. แบ่งช่วง $[0, 0.09]$ ตามความน่าจะเป็น แล้วพบว่า $i = 0.035$ อยู่ในช่วง $[0.027, 0.036]$ ซึ่งเป็นช่วงของตัวอักษร B ดังนั้นจะได้อักษรตัวที่สามคือ B จากนั้นลดขอบเขตของ $[0, 0.09]$ ลงเป็น $[0.027, 0.036]$
4. แบ่งช่วง $[0.027, 0.036]$ ตามความน่าจะเป็น แล้วพบว่า $i = 0.035$ อยู่ในช่วง $[0.0324, 0.0351]$ ซึ่งเป็นช่วงของตัวอักษร D ดังนั้นจะได้อักษรตัวที่สี่คือ D จากนั้นลดขอบเขตของ $[0.027, 0.036]$ ลงเป็น $[0.0324, 0.0351]$

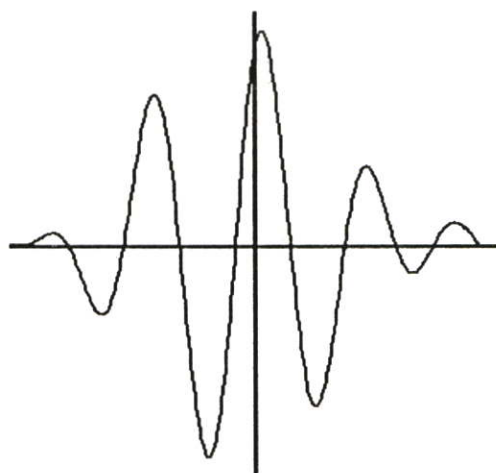
5. แบ่งช่วง $[0.0324, 0.0351]$ ตามความน่าจะเป็น แล้วพบว่า $i = 0.035$ อยู่ในช่วง $[0.03483, 0.0351]$ ซึ่งเป็นช่วงของตัวอักษร! ดังนั้นจะได้อักษรตัวที่ห้าคือ ! จากนั้นลดขอบเขตของ $[0.0324, 0.0351]$ ลงเป็น $[0.03483, 0.0351]$

2.4 ทฤษฎีเวฟเล็ต

การใช้เวฟเล็ตเป็นการวิเคราะห์สัญญาณที่ระดับความละเอียดหลายระดับ (Multi Resolution Analysis) โดยการแปลงเวฟเล็ตจากภาพจะได้ผลลัพธ์เป็นสัมประสิทธิ์ของเวฟเล็ต (Wavelet Coefficient) การทำ Discrete Wavelet Transform(DWT) [5]จะได้จำนวนสัมประสิทธิ์เท่ากับจำนวนพิกเซลในภาพต้นฉบับ เพียงแค่แยกข้อมูลออกเป็นสองส่วน คือส่วนที่เรียกว่าส่วน Approximation กับส่วน Prediction Residual ซึ่งก็หมายถึงส่วนความถี่ต่ำกับความถี่สูง การแปลงเวฟเล็ตสามารถทำได้หลายลำดับชั้น ทำให้สามารถวิเคราะห์ภาพได้ในหลายๆช่วงความถี่ [6] ซึ่งอาจจะมีเพียงสัมประสิทธิ์เพียงช่วงเท่านั้นที่มีค่ามากและสำคัญ บางช่วงสามารถตัดทิ้งได้เนื่องจากมีค่าต่ำ การตัดทิ้งไม่มีผลกระทบต่อการณ์มองเห็นมากนักในเวลาทำการแปลงสัญญาณกลับ (Reconstruction)

ทฤษฎีเวฟเล็ต [7] เป็นคณิตศาสตร์ที่ใช้อธิบายการสร้างโมเดลของสัญญาณระบบหรือกระบวนการทางฟิสิกส์ใดๆ ซึ่งประกอบขึ้นด้วยเซตของสัญญาณเฉพาะหรือระบบย่อยมารวมกันเป็นสัญญาณหรือระบบนั้นๆ สัญญาณเฉพาะนี้จะเป็นคลื่นเล็กๆ ซึ่งถูกเรียกว่า “เวฟเล็ต”

ลักษณะของเวฟเล็ตจะเป็นคลื่นที่มีการเปลี่ยนแปลงอย่างต่อเนื่อง(Oscillatory) และขนาดของคลื่นจะลดลงสู่ศูนย์อย่างรวดเร็วทั้งสองด้านดังรูปที่ 2.4



รูปที่ 2.4 แสดงคลื่นเวฟเล็ต

การนำเวฟเลตหลายๆ อันมารวมกันเป็นเซต เพื่ออธิบายสัญญาณใดๆ โดยที่คลื่นเวฟเลตแต่ละตัวจะมีโครงสร้างมาจากฟังก์ชันเดียวกัน ซึ่งฟังก์ชันนี้จะเป็นเวฟเลตต้นกำเนิดที่เรียกว่า เวฟเลตแม่ (Mother Wavelet) คลื่นเวฟเลตแต่ละอันจะอยู่ภายในเซตของเวฟเลตนี้ โดยคลื่นแต่ละคลื่นจะเกิดจากการสเกล (Scaling) a และการเลื่อนตำแหน่ง (Translation or Shifting) a พารามิเตอร์ b จะหมายถึง การสเกลหรือการหดเข้าหรือขยายออก ส่วนพารามิเตอร์ b จะเป็นการเลื่อนตำแหน่งการเกิดคลื่นของเวฟเลตบนแกนเวลา นอกจากนี้เวฟเลตที่ถูกสเกลจะต้องทำการนอร์มอลไลซ์ด้วย $\frac{1}{\sqrt{a}}$ เสมอ เพื่อให้เวฟเลตที่ถูกสเกลไปแล้วยังมีพลังงานเช่นเดียวกับเวฟเลตแม่ ถ้าให้ $g(t)$ เป็นฟังก์ชันของเวฟเลตแม่จะสามารถเขียนเป็นสมการทั่วไปของเวฟเลตที่ตำแหน่ง a, b ใดๆ ได้ซึ่งมีความสัมพันธ์กับเวฟเลตแม่ดังนี้

$$g_{a,b}(t) = \frac{1}{\sqrt{a}} g\left(\frac{t-b}{a}\right) \quad (2.4)$$

2.4.1 พื้นฐานของการสเกลและการเลื่อนตำแหน่งของเวฟเลต

ทฤษฎีเวฟเลตจะใช้ในการอธิบายสิ่งใดสิ่งหนึ่ง เหมือนการแตกสิ่งเหล่านั้นออกเป็นส่วนประกอบเล็กๆ สัมพันธ์กัน จะอยู่ในรูปของเวฟเลตที่ถูกสเกลและเลื่อนตำแหน่งและมีค่าถ่วงน้ำหนักต่างๆ กันขึ้นอยู่กับเวฟเลตแต่ละตัว การทำการแปลงเวฟเลตหรือเรียกอีกอย่างว่า การแตกกระจายเวฟเลต (Wavelet Deconstruction) ส่วนการแปลงกลับเวฟเลต (Inverse Wavelet Transform: IWT) หรือเรียกอีกอย่างหนึ่งว่าการรวมกลับเวฟเลต (Wavelet Reconstruction)

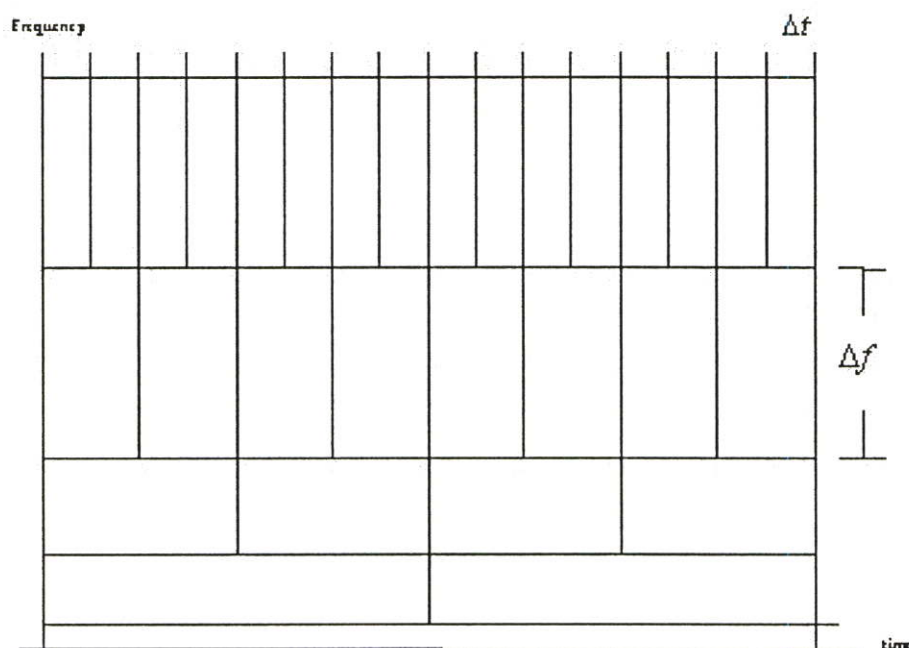
ถ้าให้ $f(t)$ เป็นสัญญาณที่ประกอบด้วยหลายๆ ความถี่เกิดขึ้น ถ้าเราพิจารณาในแกน 2 มิติ โดยแกนหนึ่งเป็นแกนเวลาและแกนหนึ่งเป็นแกนของความถี่ ดังนั้นสามารถเขียนสมการสำหรับการแปลงเวฟเลตแบบต่อเนื่อง (CWT) ที่ใช้ในการกระจาย (Decomposition) สัญญาณของฟังก์ชันใดๆ $f(t)$ ด้วยเวฟเลตแม่ $g(t)$ เป็นดังนี้

$$W_g f(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) g^* \left(\frac{t-b}{a} \right) dt \quad (2.5)$$

โดย “*” หมายถึง Complex Conjugate ของ $g(t)$, a เป็นพารามิเตอร์ของการสเกลและ b เป็นพารามิเตอร์ของการเลื่อนตำแหน่ง สัญลักษณ์ $W_g f(a,b)$ จะหมายถึงสัญญาณอินพุตที่ขึ้นอยู่กับตัวแปรอิสระ t จะถูกส่ง (map) ไปเป็นฟังก์ชัน 2 มิติที่ขึ้นอยู่กับตัวแปรอิสระ a, b โดยที่ $W_g f(a,b)$ จะเป็นสัมประสิทธิ์ของเวฟเลต และการสเกลกับการเลื่อนตำแหน่งเขียนได้เป็น $t = \frac{t-b}{a}$

การแปลงเวฟเล็ต ซึ่งจะเป็นการทรานส์ฟอร์มผ่านฟังก์ชันหน้าต่าง แต่ฟังก์ชันหน้าต่างจะสามารถเลื่อนตำแหน่งและปรับเปลี่ยนช่วงเวลาให้เหมาะสมกับความถี่ที่จะวิเคราะห์โดยเรียกฟังก์ชันหน้าต่างนี้ว่า “ฟังก์ชันเวฟเล็ต” (Wavelet Function)

$g(t)$ จะเป็นฟังก์ชันเวฟเล็ตแม่ ซึ่งจะถูกเลื่อนตำแหน่งและถูกสเกลโดยพารามิเตอร์ a และ b ซึ่งระนาบเวลาและความถี่จะมีลักษณะการเปลี่ยนแปลงช่วงเวลาและความถี่ที่สัมพันธ์กันคือ ในช่วงความถี่สูงจะใช้ช่วงการวิเคราะห์เพียงสั้นและในช่วงความถี่ต่ำๆจะมีช่วงเวลาในการวิเคราะห์ที่ยาวนานขึ้น (Δf ต่ำ Δt สูง) จะเป็นไปดังรูปที่ 2.5 ซึ่งเหมาะสมกับลักษณะสัญญาณในทางปฏิบัติ



รูปที่ 2.5 ระนาบเวลาและความถี่การเปลี่ยนแปลงช่วงเวลาและความถี่ที่สัมพันธ์กันของการแปลงเวฟเล็ต

2.4.2 การแปลงเวฟเล็ตแบบแบ่งช่วง

ปริภูมิเวกเตอร์ในการวิเคราะห์สัญญาณ

พิจารณา a , b เป็นจำนวนจริงใดๆ และ i , j เป็นเวกเตอร์หนึ่งหน่วย (Unit Vector) ในปริภูมิเวกเตอร์ (Vector Space : v) ถ้า $ai + bj \in v$ แล้วจะได้ว่า เวกเตอร์ใดๆ ที่ประกอบขึ้นจากเวกเตอร์หนึ่งหน่วยหนึ่งตัวหรือหลายตัว แล้วเวกเตอร์นั้นยังคงอยู่ในปริภูมิเวกเตอร์นี้ (v) ดังนั้นเวกเตอร์หนึ่งหน่วยดังกล่าวจะเรียกว่าเป็น เวกเตอร์พื้นฐาน (Basis Vector) แต่เราจะพิจารณาเวกเตอร์เป็นฟังก์ชัน นั่นคือจะเรียกเวกเตอร์พื้นฐานนี้ใหม่ว่า ฟังก์ชันพื้นฐาน (Basis Function : b.f.) และถ้าฟังก์ชันพื้นฐานของเวกเตอร์ทำผลคูณภายใน (Inner Product) กัน เช่น $\langle \hat{i}, \hat{j} \rangle$ แล้วมีค่าเป็น

ศูนย์จะได้ว่าฟังก์ชันพื้นฐานของดังกล่าวจะแสดงคุณสมบัติออร์ทอโกนัล “orthogonal” และถ้าฟังก์ชันพื้นฐานของเวกเตอร์ทำผลคูณภายในตัวเองแล้วมีค่าเป็นหนึ่ง ฟังก์ชันพื้นฐานที่มีคุณสมบัติทั้งสองก็จะเรียกว่า ออร์ทอนอโมล “orthogonal”

ถ้าพิจารณาในลักษณะของสัญญาณอาจจะกล่าวได้ว่า ฟังก์ชันพื้นฐานเป็นสัญญาณพื้นฐานเล็กๆ สามารถนำมาประกอบกันเป็นปริภูมิของสัญญาณใดๆ ดังนั้นถ้าให้ U_i เป็นฟังก์ชันพื้นฐานที่เป็นออร์ทอนอโมลและ C_i เป็นสัมประสิทธิ์ของ U_i จะได้สัญญาณ v ใดๆนั้นจะสามารถประกอบขึ้นจากฟังก์ชันพื้นฐานที่ออร์ทอนอโมลที่มีค่าต่างๆรวมกันดังสมการ

$$v = \sum_{i=-\infty}^{\infty} C_i U_i, v \in V \quad (2.6)$$

2.4.3 การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต (Mollat's Multiresolution Analysis or Mollat'MRA)

สมมติให้ V_j เป็นปริภูมิเวกเตอร์ที่มี j แสดงถึงระดับความละเอียดและจำนวนของฟังก์ชันพื้นฐาน พิจารณาปริภูมิเวกเตอร์ที่ระดับความละเอียด j จะมีฟังก์ชันพื้นฐานที่จะประกอบกันเป็นสัญญาณการประมาณของ $f(t)$ เมื่อ $f(t) \in V_j$ ซึ่งเรียกฟังก์ชันพื้นฐานเหล่านี้ว่า ฟังก์ชันสเกลลิ่ง (Scaling Function) ซึ่งจัดเป็นสัญญาณพื้นฐานย่อยๆที่เกิดที่ตำแหน่งที่เวลาต่างๆกันภายใน V_j และมีความถี่ต่างกันภายใน V_j เดียวกัน ดังนั้นถ้าสามารถกำหนดเซตของฟังก์ชันสเกลลิ่ง ในเทอมของการเลื่อนตำแหน่งเป็นจำนวนเต็มของฟังก์ชันสเกลลิ่งพื้นฐานที่ปริภูมิ V_0 ได้คือ $\phi_k(t) = \phi(t - k); k \in ZZ$ และ $\phi \in L^2$ โดย ZZ คือ เซตของจำนวนเต็ม

โดยที่ k จะเป็นตำแหน่งการเกิดของ $\phi(t)$ แต่ละตัวและปริภูมิย่อยของ $L^2(IR)$ (โดยเป็นเซตของจำนวนจริง) ถูกแผ่คลุม (Span) โดยฟังก์ชันสเกลลิ่งเหล่านี้ที่กำหนดโดย $v_0 = span\{\phi_k(t)\}$ สำหรับทุกจำนวนเต็ม k จาก ∞ ถึง $-\infty$

ดังนั้นจะสามารถประมาณฟังก์ชันสัญญาณ $f(t)$ ภายในปริภูมิ V_0 ได้คือ

$$f(t) = \sum_k a_k \phi_k(t); f(t) \in V_0 \quad (2.7)$$

เนื่องจากความถี่จะเปลี่ยนแปลงขึ้นอยู่กับระดับความละเอียดของปริภูมิ j นั้นๆโดยอาศัยลักษณะคุณสมบัติที่ว่า $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}; j \in ZZ$ จะได้ว่า

$$f(x) \in V_0 \Leftrightarrow f(2x) \in V_1 \text{ เมื่อ } j = 0 \quad (2.8)$$

นั่นคือที่ระดับ V_1 มีความถี่เป็นสองเท่าของความถี่ที่ระดับ V_0 ลักษณะดังกล่าวคือการสเกลทางเวลาของสเกลลิงฟังก์ชัน ดังนั้นเซตของสเกลลิงฟังก์ชันที่เกิดจากการทำสเกลและเลื่อนตำแหน่งของฟังก์ชันสเกลลิงพื้นฐานที่ระดับ j และ $j+1$ ใดๆเขียนเป็นสมการได้คือ

$$\phi_{j,k}(t) = \text{const} \times \phi(2^j t - k); \text{const} \quad (2.9)$$

const เป็นการนำนอร์มอลไลซ์ฟังก์ชันสเกลลิงที่เกิดจากการสเกล(Scaling) ที่เกิดจากคุณสมบัติออร์โทกอนัลซึ่งกันและกันของฟังก์ชันสเกลลิง ที่พิจารณาจาก

$$\langle \phi_{j,k}, \phi_{j,l} \rangle = \int \phi_{j,k}^*(t) \phi_{j,l}(t) dt = 0; k \neq l \quad (2.10)$$

จากสมการ(2.10) เมื่อ $k=1$ เป็นการหาขนาดหรือนอร์มอลไลซ์ของตัวเอง ถ้าขนาดเท่ากับหนึ่งจะได้ว่า

$$\int \{ \phi_{j,k}(t) \}^2 dt = 1 \quad (2.11)$$

นำสมการ(2.9) มาแทนในสมการ (2.11) เพื่อหา const จะได้ว่า

$$\begin{aligned} \int \text{const}^2 \phi^2(2^j t - k) dt &= 1 \\ 2^{-j} \text{const}^2 \int \phi^2(2^j t - k) d(2^j t) &= 1 \end{aligned}$$

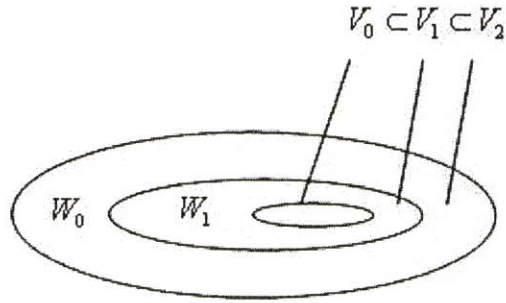
จากสมการ(2.11) จะได้

$$\begin{aligned} 2^{-j} \text{const}^2 &= 1 \\ \text{const} &= 2^{j/2} \end{aligned}$$

จากสมการที่(2.9) สามารถเขียนใหม่ได้คือ

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (2.12)$$

จากคุณสมบัติที่ว่า $V_{-\infty} \dots \subset V_{-1} \subset V_0 \subset V_1 \dots \subset V_{\infty}$ สามารถแสดงกลุ่มของปริภูมิเวกเตอร์ได้ดังนี้



รูปที่ 2.6 แสดงกลุ่มของปริภูมิเวกเตอร์

สเกลลิงฟังก์ชันในปริภูมิ V_0 สามารถเขียนได้เหมือนกับการรวมเชิงเส้นของฟังก์ชันพื้นฐาน $\sqrt{2\phi(2t-k)}$ ของ V_1 นั่นคือ

$$\phi(t) = \sum_k h(k) \sqrt{2\phi(2t-k)} \quad (2.13)$$

$h(k)$ เป็นเซตสัมประสิทธิ์สเกลลิงฟังก์ชัน หาได้จาก $\langle \phi(t), \sqrt{2\phi(2t-k)} \rangle$ และจะเรียกสมการที่ (2.13) ว่า สมการรีเคอร์ซีฟ (Recursive equation) หรือบางครั้งเรียกว่า สมการรีฟายเมนต์ (refinement equation) หรือสมการไดเลชัน (dilation equation)

จากรูปที่ 2.6 ภายในปริภูมิ V_1 จะประกอบด้วยปริภูมิ V_0 และปริภูมิที่เป็นส่วนประกอบออร์ทอกอนัล (Orthogonal complement) ของ V_0 ก็คือ W_0 นั่นคือ

$$V_1 = V_0 \oplus W_0 \quad (2.14)$$

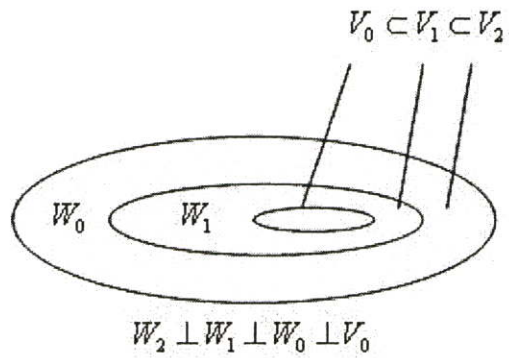
และเมื่อพิจารณาจากคุณสมบัติที่ว่า $V_i \oplus W_i = V_{i+1}$ จะได้ว่า

$$V_2 = V_1 \oplus W_1 \quad (2.15)$$

จากคุณสมบัติ $close_{L^2}(U_{V_j}) = L^2(\mathbb{R}); j \in \mathbb{Z}$ จะได้ว่า

$$L^2 = V_0 \oplus W_0 \oplus W_1 \oplus \dots \quad (2.16)$$

นั่นคือภายในปริภูมิของ V_{j+1} จะประกอบด้วยปริภูมิ V_j และปริภูมิ W_j โดยที่สมาชิกในปริภูมิ W_j ซึ่งจะออร์ทอกอนัลกับสเกลลิงฟังก์ชัน ในปริภูมิ V_j เรียกว่าเวฟเลตฟังก์ชัน (Wavelet function: $\psi(t)$) ดังรูปที่ 2.7



รูปที่ 2.7 แสดงปริภูมิเริ่มต้นที่ถูกแยกคลุม

จะเห็นว่า V_0 เป็นปริภูมิเริ่มต้นที่ถูกแยกคลุมด้วยสเกลลิงฟังก์ชัน $\phi(t-k)$ ดังนั้นเมื่อพิจารณาพร้อมกับคุณสมบัติ $V_{-\infty} \dots \subset V_{-1} \subset V_0 \subset V_1 \dots \subset V_{\infty}$ จะได้ว่าสเกลของปริภูมิเริ่มต้นอาจจะเริ่มที่ระดับความละเอียดสูงเช่น $j=10$

$$L^2 = V_{10} \oplus W_{10} \oplus W_{11} \oplus W_{12} \oplus \dots \quad (2.17)$$

หรือที่ความละเอียดต่ำ $j=-5$

$$L^2 = V_{-5} \oplus W_{-5} \oplus W_{-4} \oplus W_{-3} \oplus \dots \quad (2.18)$$

และถ้า $j = \infty$ แล้วเทียบกับสมการที่(2.17) จะได้

$$W_{-\infty} \oplus \dots \oplus W_{-1} = V_0$$

เนื่องจาก $f(t) \in L^2(\mathbb{R})$ จะได้ว่า $f(t)$ จะประกอบด้วยเวฟเลตฟังก์ชันในปริภูมิ W_j นั่นคือสามารถประมาณ $f(t)$ ได้ดังนี้

$$f(t) = \sum_{j,k} a_{j,k} \psi_{j,k}(t) \quad (2.19)$$

โดยที่ $a_{j,k}$ เป็นสัมประสิทธิ์ที่เรียกว่า ดิสครีตเวฟเลตทรานส์ฟอร์ม(DWT) ของ $f(t)$ ที่สามารถคำนวณหาได้จากการใช้ผลคูณภายในคือ

$$a_{j,k} = \langle \psi_{j,k}(t), f(t) \rangle \quad (2.20)$$

โดยที่ $\psi_{j,k}(t)$ คือเวฟเลตฟังก์ชันที่ได้จากการสเกล j และการเลื่อนตำแหน่ง k ของเวฟเลตแม่ (Mother Wavelet: $\psi(t)$)

เมื่อพิจารณาจากสมการ ปริภูมิ W_0 จะมีคุณสมบัติต่างๆ เหมือนกับปริภูมิ V_0 เพียงแต่จะเป็นส่วนประกอบย่อยที่ประกอบอยู่ในปริภูมิ V_1 ดังนั้นที่ระดับความละเอียดใดๆ (j) และพิจารณาคูณสมบัติเหมือนสเกลลิ่งฟังก์ชัน จะได้สมการที่คล้ายกับสมการที่ (2.12) คือ

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (2.21)$$

เนื่องจากเวฟเลตเหล่านี้จะอยู่ในปริภูมิที่แผ่คลุมโดยสเกลลิ่งฟังก์ชันถัดไป นั่นคือ $W_0 \subset V_1$ และในทำนองเดียวกันกับสมการที่ (2.21) สามารถเขียนเวฟเลตฟังก์ชันเป็นลักษณะของผลบวกเชิงเส้นของสเกลลิ่งฟังก์ชันพื้นฐาน (basic scaling function) ของปริภูมิ V_1 คือ

$$\psi(t) = \sum_k h_1(k) \sqrt{2} \phi(2t - k); k \in \mathbb{Z} \quad (2.22)$$

จากสมการที่ (2.22) จะเป็นเซตของสัมประสิทธิ์ เวฟเลตฟังก์ชัน โดยมีความสัมพันธ์กับเซตสัมประสิทธิ์ สเกลลิ่งฟังก์ชัน ($h(k)$) จากคุณสมบัติที่ว่า $V_i \oplus W_j = V_{j+1}; j \in \mathbb{Z}$ จะได้ว่า

$$\langle \phi_{j,k}(t), \psi_{j,l}(t) \rangle = \int \phi_{j,k}^*(t) \psi_{j,l}(t) dt = 0 \quad (2.23)$$

เมื่อ $j, k, l \in \mathbb{Z}$ จากสมการที่ (2.23) จะได้ความสัมพันธ์ของสัมประสิทธิ์เวฟเลตที่ออร์ทอกอนัลกับสัมประสิทธิ์สเกลลิ่งดังนี้

$$h_1(k) = (-1)^k h(N-1-k) \quad (2.24)$$

เมื่อ N เป็นจำนวนข้อมูลของเซตสัมประสิทธิ์สเกลลิ่ง $h(k)$

จากสมการ (2.17) เราสามารถสร้างฟังก์ชันใดๆ จากเซตของฟังก์ชัน $\phi_k(t)$ และ $\psi_{j,k}(t)$ แผ่คลุม $L^2(\mathbb{R})$ จะได้

$$g(t) = \sum_{k=-\infty}^{\infty} c(k) \phi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k) \psi_{j,k}(t) \quad (2.25)$$

จากสมการที่ (2.25) ในการกระจายอนุกรม $g(t)$ เทอมแรกของสมการจะเป็นผลรวมแสดงถึงความละเอียดต่ำ หรือการประมาณที่หยาบ (coarse approximation) ของ $g(t)$ สำหรับการเพิ่มของแต่ละ j ในเทอมที่สองจะแสดงถึงความละเอียดที่สูงขึ้นและจะทำให้รายละเอียดของข้อมูล (detail) เพิ่มขึ้น เหมือนกับอนุกรมฟูริเยร์ โดยที่สมการ (2.25) จะเลือกปริภูมิสเกลลิ่งฟังก์ชันที่สเกล j ใดๆ ก็ได้คือ

$$g(t) = \sum_k C_{j_0}(k) \phi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{\infty} d_j(k) \phi_{j,k}(t) \quad (2.26)$$

จากสมการที่ (2.26) สัมประสิทธิ์ ทั้งสองเทอมสามารถคำนวณจากผลคูณภายในคือ

$$C_j(k) = \langle g(t), \phi_{j,k}(t) \rangle \quad (2.27)$$

$$d_j(k) = \langle g(t), \phi_{j,k}(t) \rangle \quad (2.28)$$

สัมประสิทธิ์ในการกระจายของเวฟเล็ต จะถูกเรียกว่า Discrete wavelet transforms ของ $g(t)$

2.4.4 การแปลงเวฟเล็ต โดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงก์(Filter Bank)
การวิเคราะห์จากระดับที่มีความละเอียดสูงไปยังระดับที่มีความละเอียดต่ำจากสมการรีเคอร์ซีฟ

$$\phi(t) = \sum_n h(n) \sqrt{2} \phi(2t - n) \quad (2.29)$$

ถ้าทำการสเกลและเลื่อนตำแหน่งของตัวแปรเวลาจะได้ว่า

$$\begin{aligned} \phi(2^j t - k) &= \sum_n h(n) \sqrt{2} \phi(2(2^j t - n)) \\ \phi(2^j t - k) &= \sum_n h(n) \sqrt{2} \phi(2^{j+1} t - 2k - n) \end{aligned} \quad (2.30)$$

เปลี่ยนตัวแปรโดยให้ $m = 2k + n$ จะได้

$$\phi(2^j t - k) = \sum h(m - 2k) \sqrt{2} \phi(2^{j+1} t - m) \quad (2.31)$$

เมื่อพิจารณา V_j ในทำนองเดียวกันกับสมการ $V_0 = \text{span}\{\phi_k(t)\}$ จะได้ว่า

$$V_j = \text{span}\left\{2^{j/2} \phi(2^j t - k)\right\} \quad (2.32)$$

ดังนั้นถ้า $f(t) \in V_{j+1}$ แล้ว

$$f(t) = \sum_k c_{j+1}(k) 2^{(j+1)/2} \phi(2^{j+1} t - k) \quad (2.33)$$

จากสมการ (2.33) $f(t)$ เกิดจากสเกลลิงฟังก์ชันอย่างเดียวไม่มีเวฟเลต fn ดังนั้นการสเกลที่ความละเอียดต่างๆ เวฟเลตจำเป็นสำหรับ

$$f(t) = \sum_k c_j(k) 2^{j/2} \phi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \varphi(2^j t - k) \quad (2.34)$$

ถ้า $\phi_{j,k}(t)$ และ $\varphi_{j,k}(t)$ เป็นออร์ทอกอนัลกัน จะได้สัมประสิทธิ์สเกลลิงที่ระดับ j ($c_j(k)$) จะหาได้จากผลคูณภายในคือ

$$c_j(k) \langle f(t), \phi_{j,k}(t) \rangle \quad (2.35)$$

เมื่อนำสมการ (2.31) แทนในสมการ (2.35) จะได้

$$\begin{aligned} c_j(k) &= \langle f(t), 2^{j/2} \phi(2^j t - k) \rangle \\ c_j(k) &= \langle f(t), 2^{j/2} \sum_m h(m-2k) \sqrt{2} \phi(2^{j+1} t - m) \rangle \\ c_j(k) &= \sum_m h(m-2k) \langle f(t), 2^{j+1/2} \phi(2^{j+1} t - m) \rangle \end{aligned} \quad (2.36)$$

เมื่อนำสมการ (2.35) เทียบกับ (2.36) จะได้

$$c_j(k) = \sum_m h(m-2k) c_{j+1}(m) \quad (2.37)$$

และในทำนองเดียวกันจะได้

$$d_j(k) = \sum_m h_1(m-2k) c_{j+1}(m) \quad (2.38)$$

2.4.5 การฟิลเตอร์และการลดอัตราการซักราค่าตัวอย่างลง(Down sampling)

ในรูปแบบการประมวลผลสัญญาณดิจิทัล การฟิลเตอร์สัญญาณที่เป็นลำดับ จะใช้การทำการคอนโวลูชัน(Convolution) ระหว่างลำดับ (Sequence) กับเซตของจำนวนสัมประสิทธิ์ตัวฟิลเตอร์

$$\text{เมื่อ } y(n) = \sum_{k=0}^{n-1} h(k)x(n-k)$$

$$x(n) = \text{Input sequence}$$

$$h(n) = \text{Filter coefficient}$$

$$y(n) = \text{Output sequence}$$

หรือเขียนสมการ (2.39) ให้อยู่ในรูปของการคอนโวลูชันจะได้

$$Y(n) = h(n) \oplus x(n) \quad (2.39)$$

จากสมการ (2.37) เมื่อเขียนให้อยู่ในรูปของการคอนโวลูชันจะได้

$$c_j(k) = h(-2k) \oplus c_{j+1}(2k)$$

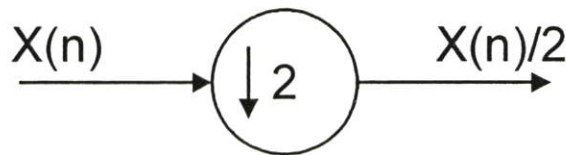
หรือ

$$c_j(k) = h(-n) \oplus c_{j+1}(n) \quad \text{เมื่อ } n = 2k \quad (2.40)$$

และทำนองเดียวกันจะได้

$$d_j(k) = h_1(-n) \oplus c_{j+1}(n) \quad \text{เมื่อ } n = 2k \quad (2.41)$$

จากสมการ (2.41) และ(2.42) สัมประสิทธิ์ $C_j(k)$ และ $d_j(k)$ จะได้สัมประสิทธิ์ $C_{j+1}(2k)$ เมื่อพิจารณาเทียบกับหลักการฟิลเตอร์แบ่งครึ่งกับการฟิลเตอร์สัญญาณดิจิทัลแล้ว เซตสัมประสิทธิ์ของ $h(-n)$ และ $h_1(-n)$ เทียบได้กับสัมประสิทธิ์ของตัวฟิลเตอร์ทางดิจิทัล และตัวชี้(index) ที่แสดงถึงลำดับของสัญญาณของ $C_j(k)$ และ $d_j(k)$ ได้จากการคอนโวลูชันกับ $h(-n)$ หรือ $h_1(-n)$ แล้วซักราค่าตัวอย่างลง 2 เท่าเรียกว่าการลดค่าจากค่าตัวอย่างหรือดีซิเมเตอร์ (Decimator) ดังรูปที่ 2.8



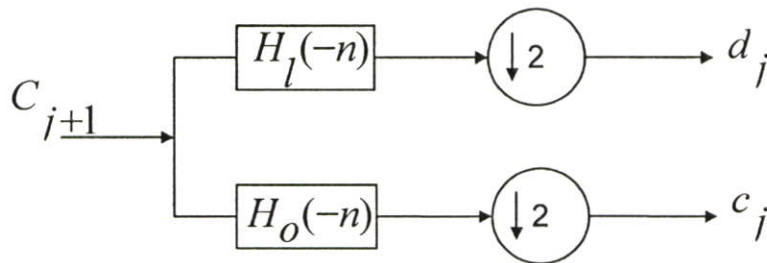
รูปที่ 2.8 แสดงสัญลักษณ์ Down sampling

ตัวอย่าง $x(n) = a, b, c, d, e, f, \dots$

$x(n/2) = a, c, e, \dots$

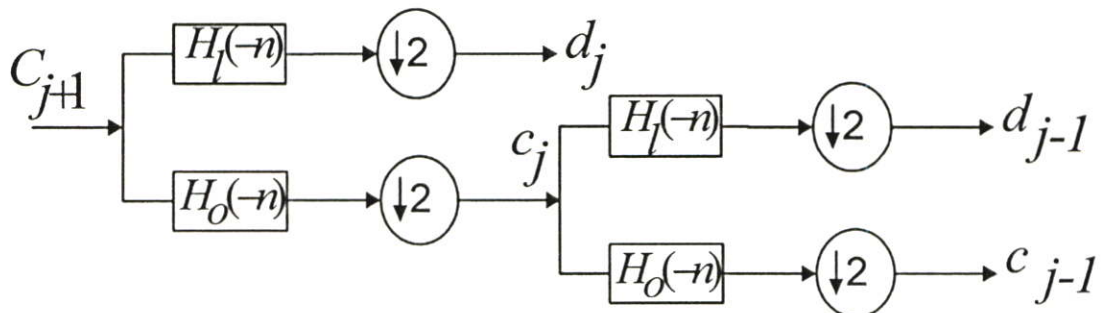
ถ้าจำนวนของตัวสัมประสิทธิ์ฟิลเตอร์มีจำนวนจำกัดขนาด n แล้ว ตัวฟิลเตอร์จะเรียกว่า การตอบสนองอิมพัลส์ที่จำกัด (Finite impulse response: FIR) และจะเรียกว่าการตอบสนองอิมพัลส์ไม่จำกัด (Infinite impulse response: IIR) ถ้าจำนวนตัวสัมประสิทธิ์ฟิลเตอร์มีจำนวนไม่จำกัด

ดังนั้นตัวกรองความถี่แบบดิจิทัลและการลดอัตราการซัดตัวอย่างจะนำมาใช้กับสมการ และตามหลักการของฟิลเตอร์แบ่งคี่ของสัญญาณดังรูปที่ 2.9

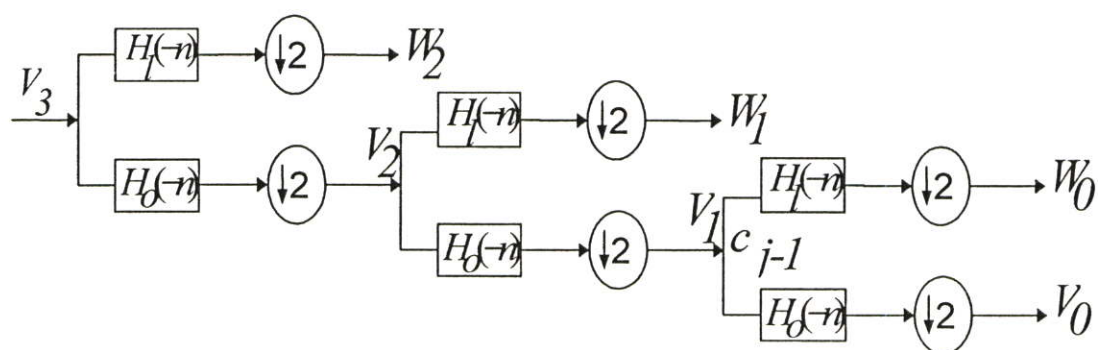


รูปที่ 2.9 แสดงตัวกรองความถี่ตามหลักฟิลเตอร์แบ่งคี่

จากรูปเป็นการกรองสัญญาณหรือการแยกสัญญาณ (Decomposition) C_{j+1} ที่ระดับ $j+1$ ไปสู่ตัวสัมประสิทธิ์ C_j และ d_j ที่ระดับ j ดังนั้นถ้านำฟิลเตอร์แบ่งคี่ดังกล่าวมาแยกสัญญาณ (ตัวสัมประสิทธิ์สเกลลิง) อีกเพื่อให้ได้สเกลที่ขยายอีก เช่น ที่สเกล $j-1$ ดังรูปต่อไปนี้

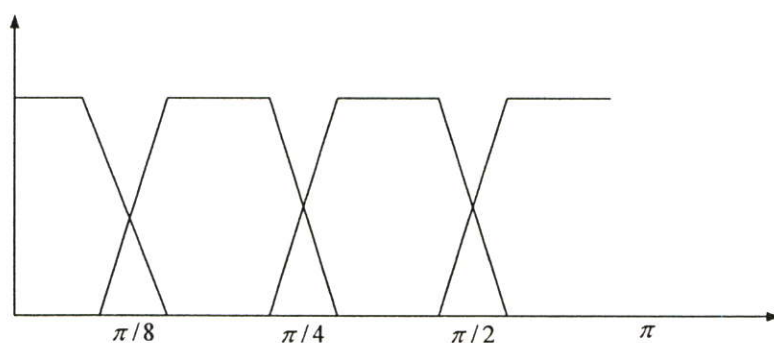
รูปที่ 2.10 แสดงฟิลเตอร์แบ่งคี่ที่ระดับ C_{j-1}

ดังนั้นถ้าต้องการจะหา j_0 จากระดับสูง โดยในทางปฏิบัติสัญญาณใดๆ ที่มีช่วงของสัญญาณจำกัด จะมีระดับสูงสุดที่ $j = J$ ด้วยการเริ่มต้นที่การอธิบายสัญญาณที่ระดับความละเอียดสูง (high resolution) ในเทอมของสัมประสิทธิ์ c_j และจะวิเคราะห์ลงไปสู่ระดับความละเอียดต่ำ (low resolution) $j = j_0$ ดังนั้นจะต้องใช้ขั้นตอนการแยกสัญญาณ $J - j_0$ ขั้นตอนเป็นดังรูปที่ 2.11



รูปที่ 2.11 แสดงการแยกสัญญาณที่ V_3

เนื่องจากในปริภูมิ V_3 จะถูกแตกกระจายสัญญาณไปสู่ระดับ $j = 2$ ทั้งในปริภูมิของสเกลถึง V_2 และปริภูมิของเวฟเลต W_2 โดยจากรูป $h_1(-n)$ จะเป็นตัวกรองแยกสัญญาณความถี่สูงหรือเทียบได้กับระดับความละเอียดสูง ส่วนตัวฟิลเตอร์ของสัมประสิทธิ์ $h_0(-n)$ จะกรองแยกสัญญาณความถี่ต่ำเมื่อแสดงในลักษณะการวิเคราะห์ทางด้านทางด้านความถี่ จะแสดงได้ดังรูปที่ 2.12



รูปที่ 2.12 แสดงการแยกสัญญาณทางด้านความถี่

ดังนั้นสำหรับ $f(t) \in v_j$ และจากสมการจะได้ว่า

$$f(t) = \sum_k c_j(k) \phi_{j,k}(t) \quad (2.42)$$

$$f(t) = \sum_k c_{j-1}(k) \phi_{j-1,k}(t) + \sum_k d_{j-1}(k) \phi_{j-1,k}(t) \quad (2.43)$$

$$f(t) = \sum_k c_{j-2}(k) \phi_{j-2,k}(t)$$

2.4.6 การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง

ในขั้นตอนการสร้างกลับคืนจากการที่เราแตกกระจายออกไป จะสามารถสร้างได้จากการรวม(Combination) ของฟังก์ชันสเกลลิงและสัมประสิทธิ์เวฟเล็ตที่ความละเอียดต่ำโดยการพิจารณาสัญญาณในปริภูมิของฟังก์ชันระดับคือ $j+1 (f(t) \in V_{j+1})$ ซึ่งค่าของสัญญาณสามารถเขียนในเทอมของสเกลลิง f_n ได้ดังนี้

$$f(t) = \sum_k c_{j+1}(k) 2^{j+1/2} \phi(2^{j+1}t - k) \quad (2.44)$$

หรือในเทอมที่ประกอบด้วยฟังก์ชันเวฟเล็ต ดังนี้

$$f(t) = \sum_k c_j(k) 2^{j/2} \phi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \varphi(2^j t - k) \quad (2.45)$$

โดยแทนสมการ (2.21) และ (2.29) ลงใน (2.45)

$$\begin{aligned} f(t) &= \sum_k c_j(k) \sum_n h_0(n) 2^{j+1/2} \phi(2^{j+1}t - 2k - n) \\ &\quad + \sum_k d_j(k) \sum_n h_1(n) 2^{j+1/2} \varphi(2^{j+1}t - 2k - n) \end{aligned} \quad (2.46)$$

จากสมการที่(2.44) สามารถหาค่า $c_{j+1}(k)$ จากการคูณด้วย $2^{j+1/2} \phi(2^{j+1}t - k)$ จะได้

$$c_{j+1}(k) = \langle f(t), \phi_{j+1,k}(t) \rangle \quad (2.47)$$

จาก (2.46) แทนตัวแปร k ด้วย m จะได้

$$\begin{aligned} f(t) &= \sum_m c_j(m) \sum_n h_0(n) 2^{j+1/2} \phi(2^{j+1}t - 2m - n) \\ &\quad + \sum_m d_j(m) \sum_n h_1(n) 2^{j+1/2} \varphi(2^{j+1}t - 2m - n) \end{aligned} \quad (2.48)$$

จาก (2.48) สามารถสร้างกลับคืน สัมประสิทธิ์ $c_{j+1}(k)$ ได้เมื่อกำหนดให้ $2m + n = k$

นั่นคือ

$$f(t) = \sum_m c_j(m) \sum_k h_0(k-2m) 2^{j+1/2} \phi(2^{j+1}t-k) + \sum_m d_j(m) \sum_k h_1(k-2m) 2^{j+1/2} \phi(2^{j+1}t-k) \quad (2.49)$$

เมื่อพิจารณาหาผลคูณภายในของสมการ (2.49) กับ $2^{j+1/2} \phi(2^{j+1}t-k)$ จะได้

$$[f(t), \phi_{j+1}(k)] = \sum_m c_j(m) h_0(k-2m) + \sum_m d_j(m) h_1(k-2m) \quad (2.50)$$

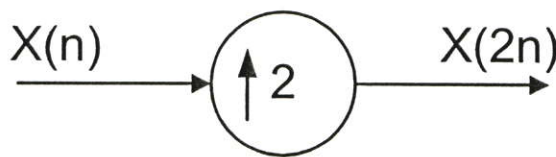
เมื่อเทียบสมการ (2.47) กับสมการ (2.50) จะได้ว่า

$$c_{j+1}(k) = \sum_m c_j(m) h_0(k-2m) + \sum_m d_j(m) h_1(k-2m) \quad (2.51)$$

2.4.7 การฟิลเตอร์และการเพิ่มอัตราการซีกค่าตัวอย่าง (Up sampling)

ในการสังเคราะห์สัญญาณกลับ จะตรงกันข้ามกับการวิเคราะห์สัญญาณด้วยการแยกช่วงออกจากกันทีละสองช่วง โดยทำการเพิ่มอัตราการซีกค่าตัวอย่างให้มีค่าเป็น 2 เท่าจากเดิมก่อน ซึ่ง จะทำการแทรกค่าศูนย์ระหว่างเทอมของสัญญาณดังสมการที่ (2.52)

$$\begin{aligned} Y(2n) &= X(n) \\ Y(2n+1) &= 0 \end{aligned} \quad (2.52)$$



รูปที่ 2.13 แสดงสัญลักษณ์การเพิ่มอัตราการซีกค่าเป็น 2 เท่า

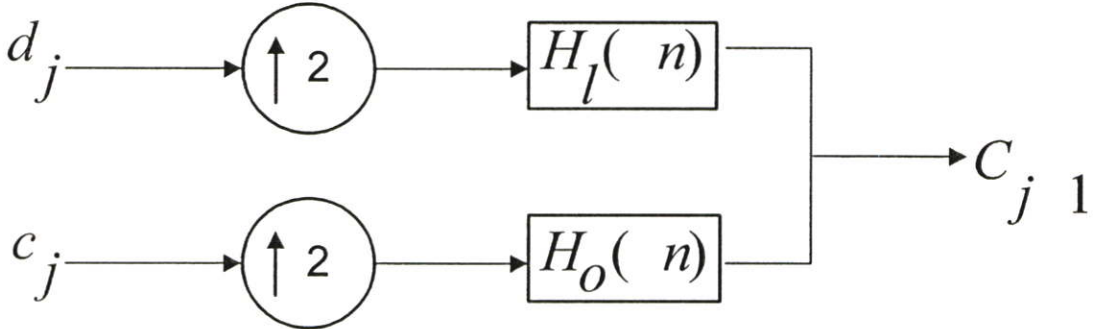
จากสมการ (2.52) พิจารณาผลรวมของแต่ละเทอมในรูปแบบการคอนโวลูชันร่วมกับสมการที่ (2.28) และสมการที่ (2.29) จะเขียนได้ใหม่เป็น

$$c_{j+1}(k) = c_j(k) \otimes h_0(k) \Big|_{k=2m} + d_j(k) \otimes h_1(k) \Big|_{k=2m} \quad (2.53)$$

โดยใช้ตัวแปร n ตัวแปรใหม่แทนตัวแปร k ในสมการ(2.53) จะได้

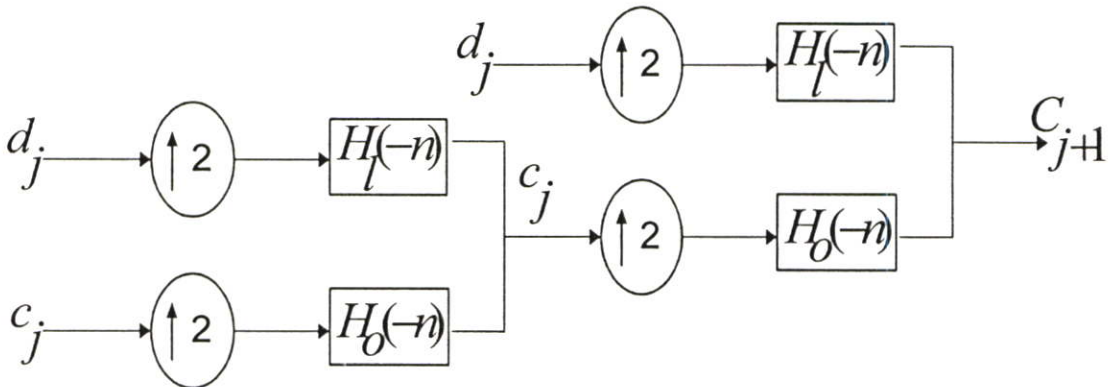
$$c_{j+1}(n) = c_j(n) \otimes h_0(n) \Big|_{n=2m} + d_j(n) \otimes h_1(n) \Big|_{n=2m} \tag{2.54}$$

จากสมการที่ (2.54) สามารถนำมาสร้างในลักษณะฟิลเตอร์แบงก์ จะได้



รูปที่ 2.14 แสดงไดอะแกรมของการซัดค่าโดยใช้หลักการของฟิลเตอร์แบงก์

ในทำนองเดียวกันสามารถทำการสังเคราะห์สัญญาณจากระดับ $j-1$ ไปสู่ระดับ j ได้ดังรูปที่ 2.15

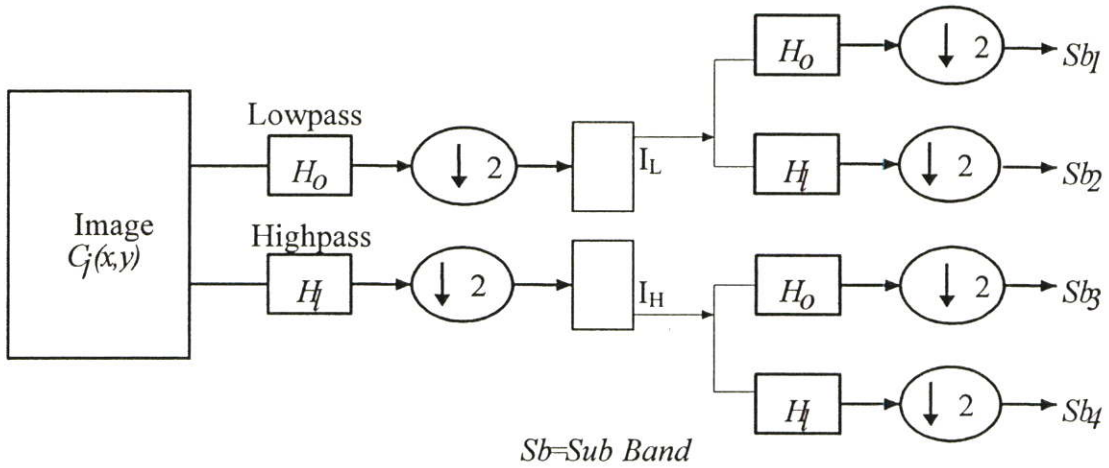


รูปที่ 2.15 แสดงการสังเคราะห์สัญญาณจากระดับ $j-1$ ไปสู่ j

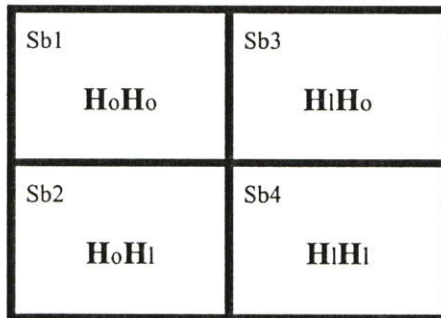
รูปที่กล่าวมาทั้งหมดตั้งแต่ Decimator จนถึงรูปข้างบนนี้จะเรียกว่า การแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete wavelet transform : DWT) ในรูปที่ 2.12 และ 2.13 จะเรียกว่า การแปลงกลับเวฟเลตแบบไม่ต่อเนื่อง (Inverse Discrete wavelet transform : IDWT) จะเห็นได้ว่า $\phi(t)$ และ $\varphi(t)$ จะเป็นฟังก์ชันที่มีช่วงเวลาจำกัด โดยจะมีค่าเท่ากับ 0 ในกรณีที่ไม่อยู่ในช่วงของ $0 \leq t \leq N-1$ ดังนั้น $h_o(-n)$ และ $h_l(-n)$ ในสมการที่ (2.11) และสมการที่ (2.21) จะมีลักษณะเป็นฟิลเตอร์ที่มีการตอบสนองอิมพัลส์ที่จำกัด (FIR Filter) ซึ่งจะเป็นผลให้เฟสของสัญญาณที่ผ่านการกรองความถี่มีลักษณะเชิงเส้นด้วย

2.4.8 การแปลงเวฟเล็ตกับข้อมูลภาพ

ข้อมูลภาพที่มีลักษณะเป็นข้อมูล 2 มิติ ดังนั้นการทำการแปลงเวฟเล็ตกับภาพในทางแนวแกน x และแกน y สลับกันในแต่ละครั้งสามารถเขียนเป็นแผนผังได้ดังรูปที่ 2.16 และมีลักษณะของการแบ่งแบนด์ย่อยของภาพดังรูปที่ 2.17



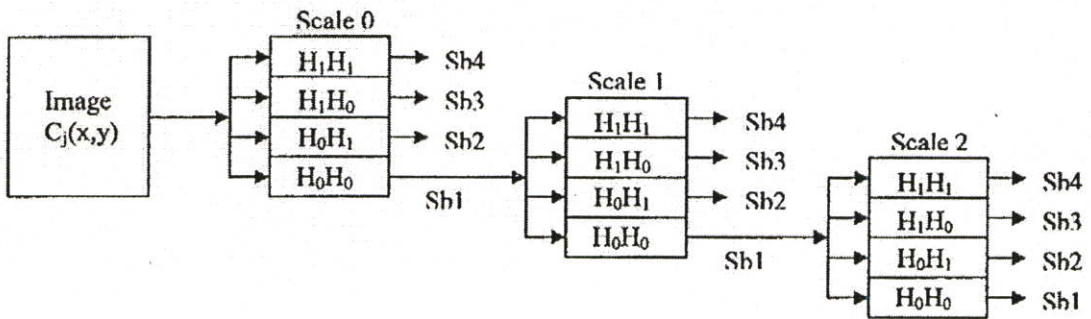
รูปที่ 2.16 แสดงแผนผังของการแปลงเวฟเล็ต



รูปที่ 2.17 แสดงลักษณะของการแบ่งแบนด์ย่อยของภาพ

รูปที่ 2.17 แสดงการแปลงเวฟเล็ตกับข้อมูลภาพนี้จะเป็นการแปลงเพียงครั้งเดียว จากรูปที่ 2.18 เป็นการแปลงเวฟเล็ต 2 มิติ จะเป็นการแปลงในแนวแกน x ตามจำนวนแถวของจุดภาพ ซึ่งทำให้ภาพถูกแบ่งออกเป็น 2 แบนด์ย่อย จากนั้นทำการแปลงเวฟเล็ตในแนวแกน y ตามจำนวนคอลัมน์ของจุดภาพ ซึ่งทำให้ภาพมีลักษณะเป็น 2 มิติ ที่ผ่านการแปลงเวฟเล็ตหนึ่งครั้งต้องถูกแบ่งออกเป็น 4 ส่วน (4 แบนด์ย่อย) แบนด์ย่อยแรกจะเป็นส่วนภาพที่ผ่านการกรองความถี่ต่ำผ่านทั้ง 2 แนวแกน แบนด์ย่อยที่ 2 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านในแนวนอนและความถี่ต่ำผ่านในแนวตั้ง แบนด์ย่อยที่ 3 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านในแนวตั้งและความถี่

ต่ำผ่านในแวนอน แบนด์ย่อยที่ 4 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านทั้ง 2 แวนอน ดังนั้นส่วนที่จะแสดงรายละเอียดของภาพจะเป็นส่วนข้อมูลที่อยู่ในแบนด์ย่อยที่ 2, 3, 4 นอกจากนี้เมื่อนำเอาแบนด์ย่อยที่ 1 ไปทำการแปลงเวฟเล็ตต่อไปเรื่อยๆ ก็จะสามารถทำการแยกแบนด์ย่อยของภาพออกไปได้มากขึ้น



รูปที่ 2.18 แสดงการแปลงเวฟเล็ตหลายๆ ระดับ

Sb ₁	Sb ₃	Sb ₃	Sb ₃
Sb ₂	Sb ₄	H ₁ H ₀	
Sb ₂	Sb ₄	H ₁ H ₀	
H ₀ H ₁	H ₁ H ₁		
Sb ₂	Sb ₄		
H ₀ H ₁	H ₁ H ₁		

รูปที่ 2.19 แสดงลักษณะการแบ่งแบนด์ย่อยหลายๆ ระดับ

จากรูปที่ 2.19 เป็นการนำแบนด์ย่อยมาแยกออกเป็นแบนด์ย่อยๆ โดยจะแบ่งเฉพาะแบนด์ย่อยที่เป็น HOHO เท่านั้น

บทที่ 3

Visual Pattern Image Coding (VPIC)

การเข้ารหัสข้อมูลภาพด้วย VPIC ถูกนำเสนอครั้งแรกโดย D.Chen และ A.C.Bovik [8] มีลักษณะการเข้ารหัสที่พยายามรักษาค่าเฉลี่ยและเกรเดียนท์ของภาพเอาไว้ ทำให้บริเวณขอบหลังการถอดรหัสยังคงไม่ได้ถูกทำลายไป โดยนำข้อมูลของภาพต้นฉบับมาทำการแบ่งออกเป็นบล็อกย่อย ซึ่งขนาดของบล็อกย่อยจะขึ้นอยู่กับขนาดของบล็อกรูปแบบที่นำมาใช้ จากนั้นจะนำบล็อกย่อยที่ได้มาเปรียบเทียบกับบล็อกรูปแบบ เพื่อหาบล็อกแพทเทิร์นที่มีความเหมือนกับบล็อกย่อยมากที่สุด และหมายเลขของบล็อกรูปแบบนั้นจะเป็นรหัสไบนารีของบล็อกย่อย

ส่วนค่าความเข้มแสง (Intensity) ในบล็อกย่อยแต่ละบล็อกที่ได้จะนำมาใช้ในการคำนวณเป็นค่าความเข้มเฉลี่ย ค่า Block gradient magnitude ($|\nabla b|_{i,j}$) เพื่อใช้เป็นข้อมูลและค่า block orientation ($\angle \Delta b_{i,j}$) เพื่อใช้ในการเปรียบเทียบบล็อกรูปแบบกับบล็อกย่อย

3.1 การเข้ารหัสข้อมูลภาพแบบ VPIC กับภาพเชิงตัวเลขแบบ 2 มิติ

ในการเข้ารหัสภาพแบบ VPIC นั้น เมื่อทำการแบ่งภาพต้นฉบับออกเป็นบล็อกย่อยขนาด 4×4 พิกเซลแล้วจะทำการคำนวณค่าความเข้มเฉลี่ย จากสมการที่ 3.1

$$Mean = \left(\sum_{i=1}^4 \sum_{j=1}^4 I(i,j) \right) / 16 \quad (3.1)$$

เมื่อ $I(i,j)$ เป็นบล็อกย่อยของภาพต้นฉบับ

และ

$$\Delta x b_{i,j} = AVE\{I_{n,m} : 4i+2 \leq n \leq 4i+3, 4j \leq m \leq 4j+3\} - AVE\{I_{n,m} : 4i \leq n \leq 4i+1, 4j \leq m \leq 4j+3\} \quad (3.2)$$

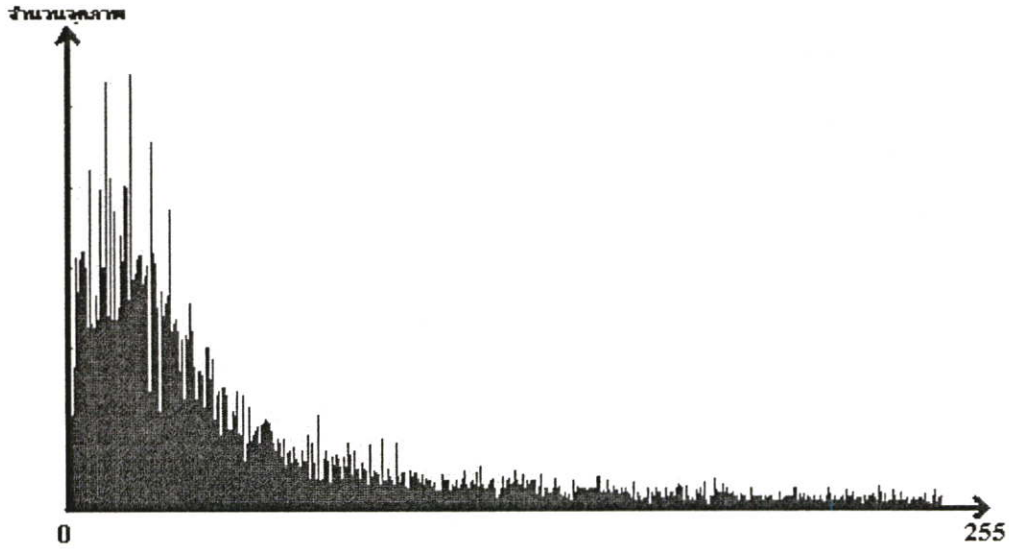
$$\Delta y b_{i,j} = AVE\{I_{n,m} : 4i \leq n \leq 4i+3, 4j+2 \leq m \leq 4j+3\} - AVE\{I_{n,m} : 4i \leq n \leq 4i+3, 4j \leq m \leq 4j+1\} \quad (3.3)$$

เมื่อ $I_{n,m}$ เป็นค่าความเข้มแสงของพิกเซล ในบล็อกย่อย

ค่าความเข้มเฉลี่ยจะถูกใช้เป็นหนึ่งในการคำนวณผลลัพธ์ที่ได้จากการเข้ารหัสของบล็อกย่อย ส่วน $\Delta x b_{i,j}$ และ $\Delta y b_{i,j}$ จะนำไปคำนวณ $|\nabla b|_{i,j}$ และ $\angle \Delta b_{i,j}$ เพื่อใช้แบ่งประเภทของบล็อกย่อยในส่วนถัดไป

3.2 ประเภทของบล็อกย่อย

การเข้ารหัสแบบ VPIC จะไม่ใช้บล็อกย่อยทุกบล็อกมาทำการเข้ารหัส เนื่องจากเมื่อแบ่งภาพต้นฉบับออกเป็นบล็อกย่อยในบางบล็อกย่อยอาจมีค่าความเข้มแสงที่ไม่แตกต่างกันเมื่อมองด้วยสายตา



รูปที่ 3.1 ฮิสโตแกรมของภาพต้นฉบับ

จากฮิสโตแกรมของภาพต้นฉบับในรูปที่ 3.1 จะแบ่งช่วงเทรสโสดออกเป็นเทรสโสดสูงและเทรสโสดต่ำ โดยบล็อกย่อยที่มีค่า $|\nabla b|_{i,j}$ อยู่ระหว่างเทรสโสดต่ำและเทรสโสดสูงจะถือว่าเป็นบล็อกย่อยที่เป็นขอบของภาพ และจะนำบล็อกย่อยเหล่านี้มาทำการเข้ารหัสต่อไป

สมการที่ (3.4) เป็นสมการที่ใช้ในการหาค่า $|\nabla b|_{i,j}$ ของบล็อกย่อย

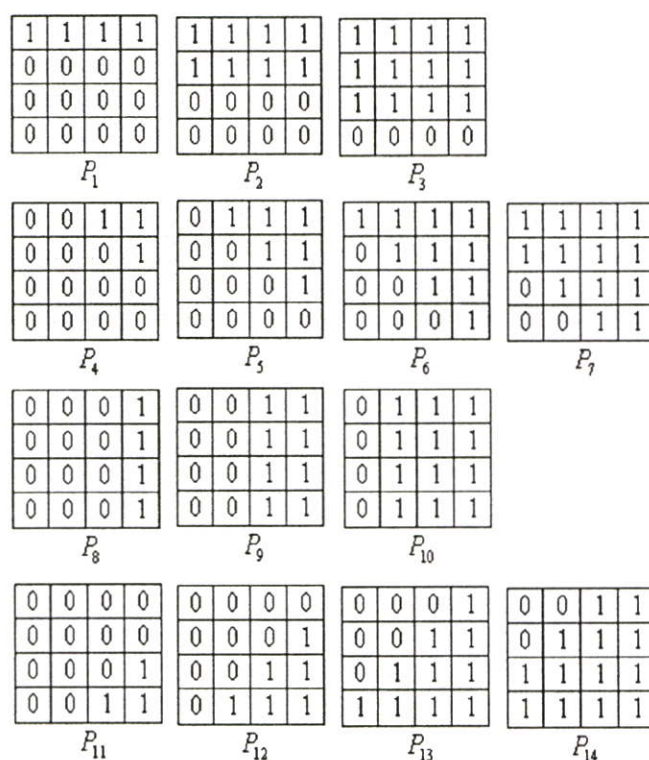
$$|\nabla b|_{i,j} = \sqrt{(\Delta x b_{i,j})^2 + (\Delta y b_{i,j})^2} \quad (3.4)$$

ใช้ค่า $|\nabla b|_{i,j}$ ของบล็อกย่อยเป็นตัวแบ่ง ถ้าค่า $|\nabla b|_{i,j} > |\nabla I|$ จะถือว่าเป็นบล็อกย่อยนั้นเป็นบล็อกย่อยที่เป็นขอบของภาพ และถ้าค่า $|\nabla b|_{i,j} \leq |\nabla I|$ ก็จะถือว่าเป็นบล็อกย่อยที่ไม่เป็นขอบ จะนำเฉพาะบล็อกย่อยที่เป็นขอบของภาพมาทำการเปรียบเทียบกับ บล็อกรูปแบบในรูปที่ 3.2 เพื่อหาบล็อกรูปแบบที่มีความเหมือนกับบล็อกย่อยที่เป็นขอบของภาพมากที่สุด และนำหมายเลขของบล็อกรูปแบบนั้นมาทำการเข้ารหัสข้อมูล

3.3 การเข้ารหัส

จากบล็อกย่อยที่เป็นขอบของภาพนำมาหาค่า $\angle \Delta b_{i,j}$ จากสมการที่ (3.5) เพื่อใช้ในการนำบล็อกย่อยที่เป็นขอบของภาพมาเปรียบเทียบกับบล็อกรูปแบบที่มีทิศทางเดียวกัน โดยหาบล็อกรูปแบบที่มีความเหมือนกับบล็อกย่อยที่เป็นขอบของภาพมากที่สุด ว่าเป็นบล็อกรูปแบบหมายเลขใด

รูปที่ 3.2 แสดงรูปแบบของบล็อกรูปแบบพื้นฐานทั้ง 14 แบบ โดยมีมุมเป็น 90° , 45° , 0° และ -45° ตามลำดับ หมายเลขของบล็อกรูปแบบที่ได้จากการเปรียบเทียบกับบล็อกย่อยที่เป็นขอบของภาพมีลักษณะเป็นเลขไบนารีขนาด 5 บิต



รูปที่ 3.2 รูปแบบของบล็อกรูปแบบในหมายเลขต่างๆ

ตารางที่ 3.1 รายละเอียดและจำนวนบิตของข้อมูล

	Uniform Block(n_u)	Edge Block(n_e)
Block Type	1	1
Mean Intensity	6	6
Gradient Magnitude	0	6
Pattern Number	0	5
Total	7	18

$$\angle \Delta b_{i,j} = \tan^{-1} \left(\frac{\Delta y b_{i,j}}{\Delta x b_{i,j}} \right) \quad (3.5)$$

ข้อมูลของบล็อกย่อยที่เป็นขอบของภาพ จะใช้ค่าความเข้มเฉลี่ย ค่า Gradient Magnitude หมายเลขของบล็อกรูปแบบและบิตแสดงสถานะของข้อมูล

บล็อกย่อยที่ไม่เป็นขอบของภาพจะไม่ถูกนำมาเปรียบเทียบกับบล็อกรูปแบบ ข้อมูลของบล็อกย่อยที่ไม่เป็นขอบของภาพ จะมีเพียงค่าความเข้มเฉลี่ย และบิตแสดงสถานะของข้อมูลขนาด 1 บิต

3.4 การถอดรหัส

ในขั้นตอนการถอดรหัส เมื่อตรวจสอบบิตแสดงสถานะของข้อมูล ถ้าเป็นบล็อกที่ไม่เป็นขอบของภาพจะทำการสร้างบล็อกย่อยขนาด 4x4 พิกเซลแล้วทำการแทนค่าความเข้มเฉลี่ย ลงไปในทุกๆ พิกเซลของบล็อกย่อย

ถ้าเป็นบล็อกที่เป็นขอบของภาพจะทำการเรียกบล็อกรูปแบบที่มีหมายเลขตรงกับ Pattern Number ของข้อมูลในแต่ละชุด แล้วทำการแทนค่า High Intensity ลงในพิกเซลที่เป็น “1” แทนค่า Low Intensity ลงในพิกเซลที่เป็น “0” แล้วนำบล็อกย่อยแต่ละบล็อกมาต่อกันก็จะได้ภาพที่มีความคิดเห็นจากภาพต้นฉบับไม่มาก แต่มีขนาดของข้อมูลที่เล็กลงจากข้อมูลของภาพต้นฉบับมาก

การคำนวณหาค่า High Intensity และ Low Intensity สามารถคำนวณได้จากสมการที่ (3.6) และสมการที่ (3.4)

$$xH + yL = 16 * Mean \quad (3.6)$$

เมื่อ x เป็นจำนวน “1” ในบล็อกรูปแบบหมายเลขนั้นๆ

เมื่อ y เป็นจำนวน “0” ในบล็อกรูปแบบหมายเลขนั้นๆ

ค่า High Intensity และ Low Intensity นั้นจะเป็นค่าเฉพาะของแต่ละบล็อกแพทเทิน เช่น เมื่อต้องการหาค่า High Intensity และ Low Intensity ของบล็อกแพทเทิน “P1” ในรูปที่ 3.2 เมื่อแทนค่าในสมการ (3.6) จะได้

$$4H + 12L = 16 * Mean \quad (3.7)$$

และสมการที่ (3.4)

$$|\nabla b|_{i,j} = \sqrt{(\Delta x b_{i,j})^2 + (\Delta y b_{i,j})^2} \quad (a)$$

$$\text{เมื่อ } (\Delta x b_{i,j}) = \frac{4H - 4L}{8} \text{ และ } (\Delta y b_{i,j}) = 0$$

$$\begin{aligned} |\nabla b|_{i,j} &= \sqrt{\left(\frac{4H - 4L}{8}\right)^2} \\ 8|\nabla b|_{i,j} &= 4H - 4L \end{aligned} \quad (b)$$

สมการ (a) – (b)

$$\begin{aligned} 16L &= 16 * \text{Mean} - 8|\nabla b|_{i,j} \\ L &= \frac{16 * \text{Mean} - 8|\nabla b|_{i,j}}{16} \end{aligned}$$

จากสมการ (a) แทน L ลงไปจะได้

$$H = \frac{16 * \text{Mean} - 12L}{4}$$

3.5 การวัดประสิทธิภาพของการบีบอัดและการวัดประสิทธิภาพของภาพที่ได้หลังจากการเรียกคืน

การวัดประสิทธิภาพของการบีบอัดระหว่าง Edge Block (n_e) และ Uniform Block (n_u) จะวัดจากค่าอัตราการบีบอัด ซึ่งคำนวณได้ตามสมการที่ (3.7)

$$\frac{128}{n_e} : 1 \leq C_{VPIC} \leq \frac{128}{n_u} : 1 \quad (3.7)$$

ส่วนของการวัดประสิทธิภาพของการบีบอัด จะวัดจากค่าอัตราการบีบอัด C_r [9] ซึ่งคำนวณได้ตามสมการที่ (3.8)

$$C_r = \frac{\text{จำนวนบิตที่ใช้เข้ารหัสภาพต้นฉบับ}}{\text{จำนวนบิตที่เข้ารหัสข้อมูลภาพที่ถูกบีบอัด}} \quad (3.8)$$

โดยทั่วไปจะมีการชดเชยข้อดีข้อเสียระหว่างอัตราการบีบอัดกับคุณภาพของภาพ (Picture quality) ในกรณีที่อัตราการบีบอัดสูงจะให้ภาพผลลัพธ์หลังการบีบอัดมีคุณภาพต่ำอย่างไรก็ตามพบว่าคุณภาพและอัตราการบีบอัดอาจเปลี่ยนแปลงไปตามคุณลักษณะหรือ รายละเอียดของภาพ

สำหรับการวัดคุณภาพของภาพด้วยค่า N_b ซึ่งนำเสนอโดย G Wallace [10] จะได้จากการนับจำนวนบิตที่ใช้เข้ารหัสในหนึ่งจุดภาพจากภาพหลังการบีบอัด บางครั้ง N_b นี้ อาจจะใช้คำว่า อัตราบิต (Bit per pixel) แทนก็ได้ ซึ่งอัตราบิตเขียนย่อเป็น bpp สำหรับสูตรในการคำนวณ N_b ได้จาก

$$N_b = \frac{\text{จำนวนบิตทั้งหมดที่ใช้เข้ารหัสภาพที่ถูกบีบอัด}}{\text{จำนวนจุดภาพทั้งหมดของภาพที่ถูกบีบอัด}} \quad (3.9)$$

จากการวัดคุณภาพของภาพด้วยสมการที่ (3.9) นี้ Wallace ได้แบ่งคุณภาพออกเป็น 4 กลุ่มใหญ่ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 คุณภาพของภาพ

N_b (จำนวนบิตต่อจุดภาพ)	คุณภาพของภาพ
0.25 - 0.5	คุณภาพดีพอสมควร
0.5 - 0.75	ดีดีมาก
0.75 - 1.0	ดีเยี่ยม
1.0 - 2.0	เกือบไม่แตกต่างจากภาพต้นฉบับ

การวัดประสิทธิภาพของภาพที่ได้หลังจากการเรียกคืน สามารถคำนวณได้จากกำลังสองเฉลี่ย (Mean Square Error) ของค่าผิดพลาดหรือบางครั้งเรียกว่าค่าความผิดพลาดของความเหมือน (Fidelity error) โดยใช้สมการที่ (3.10)

$$MSE = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N [f(x, y) - \hat{f}(x, y)]^2 \quad (3.10)$$

เมื่อ $M \times N$ เป็นจำนวนจุดภาพทั้งหมด

$f(x, y)$ เป็นค่าระดับสีเทาของจุดภาพต้นฉบับที่ตำแหน่ง

$\hat{f}(x, y)$ เป็นค่าระดับสีเทาของจุดภาพหลังการแก้การบีบอัด (Decompression)

ค่า MSE นี้เป็นการวัดความแตกต่างของภาพต้นฉบับกับภาพหลังการกู้ (Recovery) ภาพคืนจากการบีบอัดข้อมูล โดยปกติจะพบว่าภาพหลังการกู้ภาพคืนจากการบีบอัดที่ให้ค่า MSE ต่ำ ๆ จะมีคุณภาพที่ดีกว่าภาพหลังการแก้การบีบอัดที่ให้ค่า MSE สูง ๆ

บทที่ 4

การเข้ารหัสแบบ VPIC บน Wavelet Domain

การบีบอัดข้อมูลภาพโดยอยู่บนพื้นฐานของการแปลงเวฟเล็ต(Wavelet-based Image Coding) ได้รับความสนใจเป็นอย่างมาก (Antonini et.al.: 1992), (Lewis and Knowles: 1992 ; Shapiro : 1993 ; Said and Pearlman : 1996) เนื่องจากมีจุดเด่นคือ ลดการเกิดรอยต่อระหว่างบล็อกของภาพเหมือนกับเทคนิคอื่นๆ อย่างเช่น JPEG หรือเวกเตอร์ควอนไทซ์เซชันบนสเปเชียลโดเมน (Spatial Domain)

ในงานวิจัยนี้จึงได้นำเสนอแนวคิดในการบีบอัดข้อมูลภาพดังนี้คือ การเข้ารหัส(บีบอัด) และการถอดรหัส(คลาย) ที่อยู่บนพื้นฐานของการเข้ารหัสแบบ VPIC บนเวฟเล็ต โดเมน(Wavelet Domain) โดยจะทำการเก็บแบนด์ LL ไว้และทำการนำแบนด์ LH ,HL ,HH ที่ได้จากการแปลงเวฟเล็ตมาเพียงหนึ่งแบนด์ โดยจะนำแบนด์ที่มีการกระจายตัวของข้อมูลมากที่สุดมาทำการเข้ารหัสแบบ VPIC เพื่อลดขนาดข้อมูลของแบนด์นั้น ส่วนแบนด์ที่เหลือนั้นสามารถตัดทิ้งไปได้เนื่องจากมีการกระจายตัวของข้อมูลต่ำเสมือนว่ามีความสำคัญน้อยกว่า เมื่อทำการเรียกคืนภาพจะทำการถอดรหัสข้อมูลด้วย VPIC และนำข้อมูลที่ได้จากการถอดรหัสมาทำการแปลงกลับเวฟเล็ต ร่วมกับแบนด์ LL เพียงสองแบนด์เท่านั้น โดยแบนด์ที่เหลือให้ค่าเป็นศูนย์ จากนั้นนำภาพที่ได้จากการแปลงกลับเวฟเล็ตมาทำการหาค่าความผิดเพี้ยนจากภาพต้นฉบับเปรียบเทียบกับขนาดของข้อมูลที่ได้จากการบีบอัดแบบการเข้ารหัสภาพแบบ VPIC บนเวฟเล็ตโดเมน ว่ามีค่าความผิดเพี้ยนมากน้อยเพียงใดต่ออัตราการบีบอัดข้อมูล

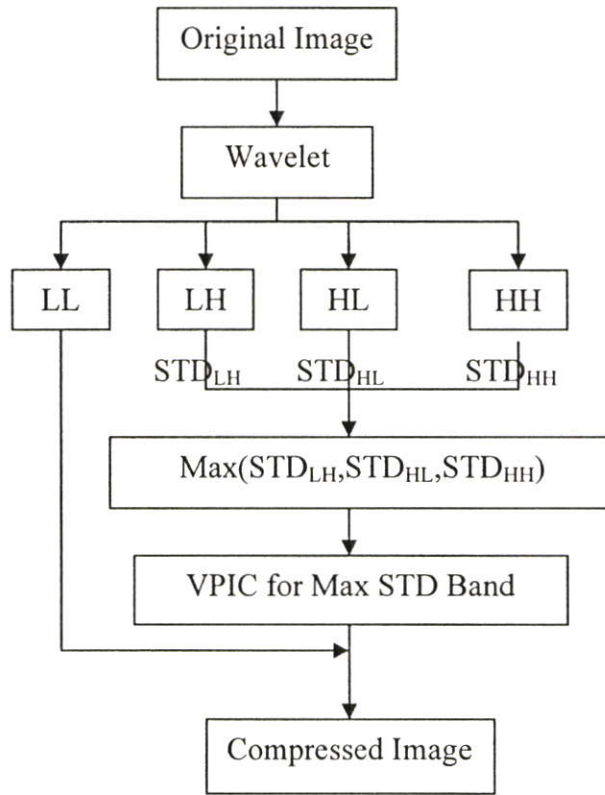
4.1 การเลือกแบนด์เพื่อเข้ารหัสแบบ VPIC และการเข้ารหัสแบบ VPIC

รูปแบบการเข้ารหัสและถอดรหัสสำหรับวิธีการเข้ารหัสภาพแบบ VPIC บนภาพแบนด์ย่อยนั้น จากทฤษฎีของเวฟเล็ตรายละเอียดหรือความสำคัญของภาพส่วนใหญ่จะอยู่ที่แบนด์ LL ส่วนรายละเอียดรองลงมาจะอยู่ที่ลักษณะของภาพซึ่งอยู่ในแบนด์ที่เหลือ ดังนั้นก่อนนำข้อมูลมาเข้ารหัสแบบ VPIC จะต้องทำการเลือกแบนด์ที่มีความสำคัญรองลงมา มีโครงสร้างดังรูปที่ 4.1

4.1.1 การเลือกแบนด์เพื่อเข้ารหัสแบบ VPIC

การแปลงเวฟเล็ตกับภาพจะเป็นการแปลงข้อมูลทางแกน x และแกน y สลับกันในแต่ละครั้ง ผลลัพธ์ที่ได้จากการแปลงจะมี 4 แบนด์ คือ แบนด์ LL แบนด์ LH แบนด์ HL และแบนด์ HH สำหรับแบนด์ LL นั้นจะเป็นแบนด์ที่มีส่วนประกอบสำคัญของภาพอยู่มากที่สุดจึงคงแบนด์ LL ไว้ ส่วนแบนด์ LH แบนด์ HL และ HH นั้นจะทำการเลือกมาเพียงหนึ่งแบนด์ ในการเลือกแบนด์นั้นจะ

เลือกแบนด์ที่มีส่วนประกอบของภาพรองลงมาจากแบนด์ LL โดยจะใช้ค่าเบี่ยงเบนมาตรฐานเป็นเกณฑ์ในการเลือก เนื่องจากค่าเบี่ยงเบนมาตรฐานนั้นเป็นสมการการกระจายตัวของข้อมูลดังสมการ (4.1) เมื่อแบนด์ใดมีค่าเบี่ยงเบนมาตรฐานสูงที่สุดจะเลือกแบนด์นั้นมาทำการเข้ารหัสแบบ VPIC ต่อไป



รูปที่ 4.1 โค้ดแแกรมการเลือกแบนด์และการเข้ารหัสแบบ VPIC

$$s = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{\frac{1}{2}} \quad (4.1)$$

เมื่อ $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ และ n เป็นจำนวนของสมาชิกในภาพ

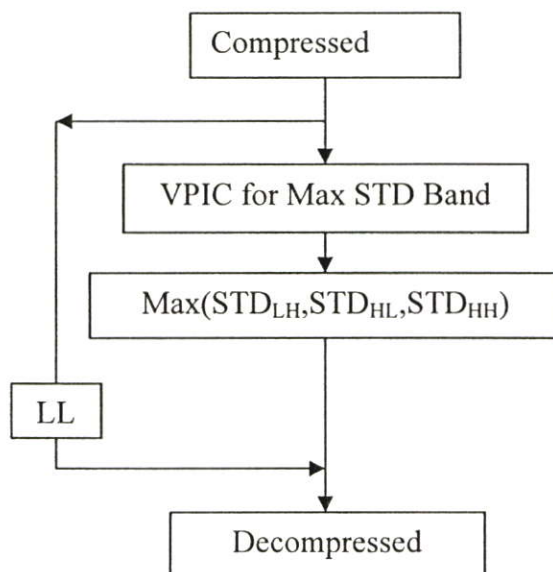
4.1.2 การเข้ารหัสแบบ VPIC

เมื่อได้แบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงที่สุดแล้ว จะนำแบนด์นั้นมาทำการเข้ารหัสแบบ VPIC ตามขั้นตอนในบทที่ 3 โดยในงานวิจัยนี้จะกำหนดค่า gradient เท่ากับ 10 และ 30 [1] ข้อมูลที่ได้จะมีลักษณะดังรูปที่ 4.2

4.2 การถอดรหัสแบบ VPIC

เมื่อภาพได้ผ่านการบีบอัดเป็นข้อมูลแล้ว ในส่วนการเรียกคืนภาพจะทำตามขั้นตอนในรูปที่

4.4



รูปที่ 4.4 โค้ดอะแกรมการถอดรหัสแบบ VPIC และการเรียกคืนภาพ

เมื่อภาพได้ผ่านการบีบอัดเป็นข้อมูลแล้ว ในส่วนการเรียกคืนภาพจะทำงานดังนี้

1. แยกข้อมูลที่ประกอบด้วยข้อมูลของแบนด์ LL กับข้อมูลที่ผ่านการเข้ารหัสแบบ VPIC ออกจากกัน ทำการถอดรหัสข้อมูลที่ผ่านการเข้ารหัสจะได้ข้อมูลเสมือนกับข้อมูลในแบนด์ที่มีค่าเป็นเบนมาตรฐานสูงสุด
2. นำแบนด์ LL และแบนด์ที่มีค่าเป็นเบนมาตรฐานสูงสุดมาทำการแปลงกลับเวฟเล็ดเพียงสองแบนด์เท่านั้น ส่วนแบนด์ที่เหลือให้เป็นศูนย์ จะได้ภาพที่ผ่านการบีบอัด

4.3 การวัดประสิทธิภาพการบีบอัดและประสิทธิภาพของภาพที่ได้หลังจากการบีบอัด

ในงานวิจัยนี้จะมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ทั้งสองแบบด้วยกัน

1. ค่าอัตราการบีบอัด C_r ซึ่งคำนวณจากสมการที่ (3.8) มาทำการวัดประสิทธิภาพการบีบอัดข้อมูล
2. การวัดประสิทธิภาพของภาพที่ได้หลังจากการบีบอัด สามารถคำนวณได้จากสมการที่ (3.10) เป็นการวัดความแตกต่างของภาพต้นฉบับกับภาพหลังการเรียกคืน

ค่าทั้งสองนี้จะมีความสัมพันธ์กันคือ เมื่ออัตราการบีบอัดสูงโดยปกติภาพผลลัพธ์หลังการบีบอัดจะมีคุณภาพต่ำ แต่อย่างไรก็ตามคุณภาพของภาพผลลัพธ์และอัตราการบีบอัดอาจเปลี่ยนแปลงไปตามคุณลักษณะหรือรายละเอียดของภาพได้

บทที่ 5

การทดลองและผลการทดลอง

5.1 ข้อมูลภาพต้นฉบับที่ใช้ในการทดลอง

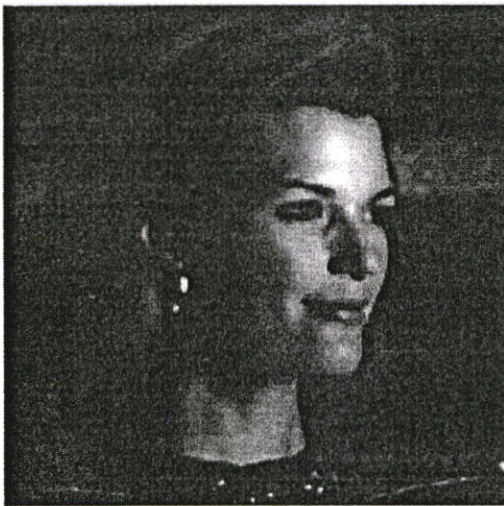
ในการทดลองภาพที่นำมาทดลองมี 4 ภาพ ดังรูปที่ 5.1 ซึ่งเป็นภาพขนาด 512 X 512 พิกเซล ระดับสีเทา 256 ระดับ (8 บิต) แต่จะภาพจะมีขนาดข้อมูลเท่ากัน คือ $512 \times 512 = 262,144$ ไบท์



(a) Lena



(b) Girl



(c) Women



(d) Cameraman

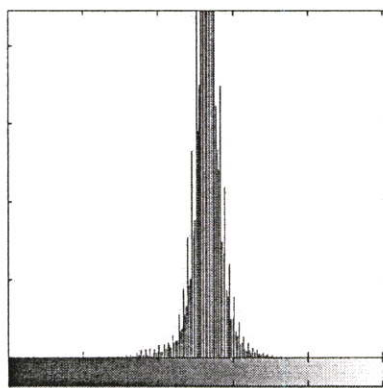
รูปที่ 5.1 ภาพต้นฉบับที่นำมาใช้ในการทดลอง

5.2 การหาค่าเบี่ยงเบนมาตรฐาน(Standard Deviation)

การหาค่าเบี่ยงเบนมาตรฐานจะนำภาพในแบนด์ต่างๆ ที่ได้จากการแปลงเวฟเล็ตแบบ db9 มาหาค่าเบี่ยงเบนมาตรฐาน เพื่อหาแบนด์ที่มีการกระจายตัวของข้อมูลที่คี่ที่สุดและนำแบนด์นั้นมาทำการเข้ารหัสแบบ VPIC ซึ่งการกระจายตัวของข้อมูลในแต่ละแบนด์จะขึ้นอยู่กับลักษณะของภาพที่นำมาใช้



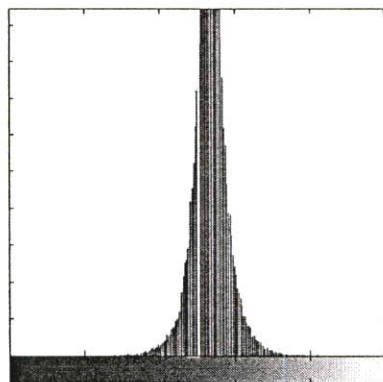
LH Band



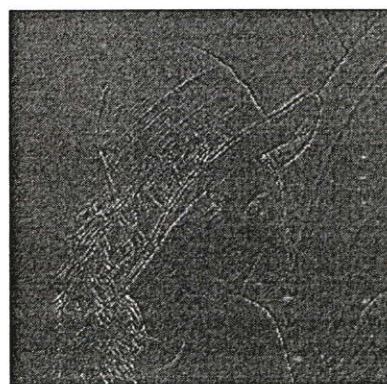
STD = 1.7861



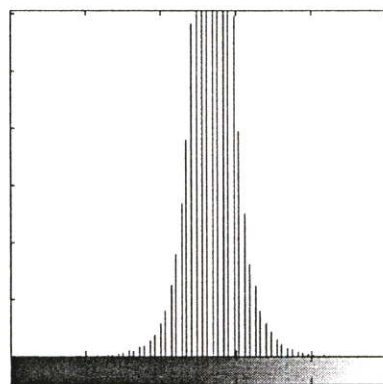
HL Band



STD = 4.7285

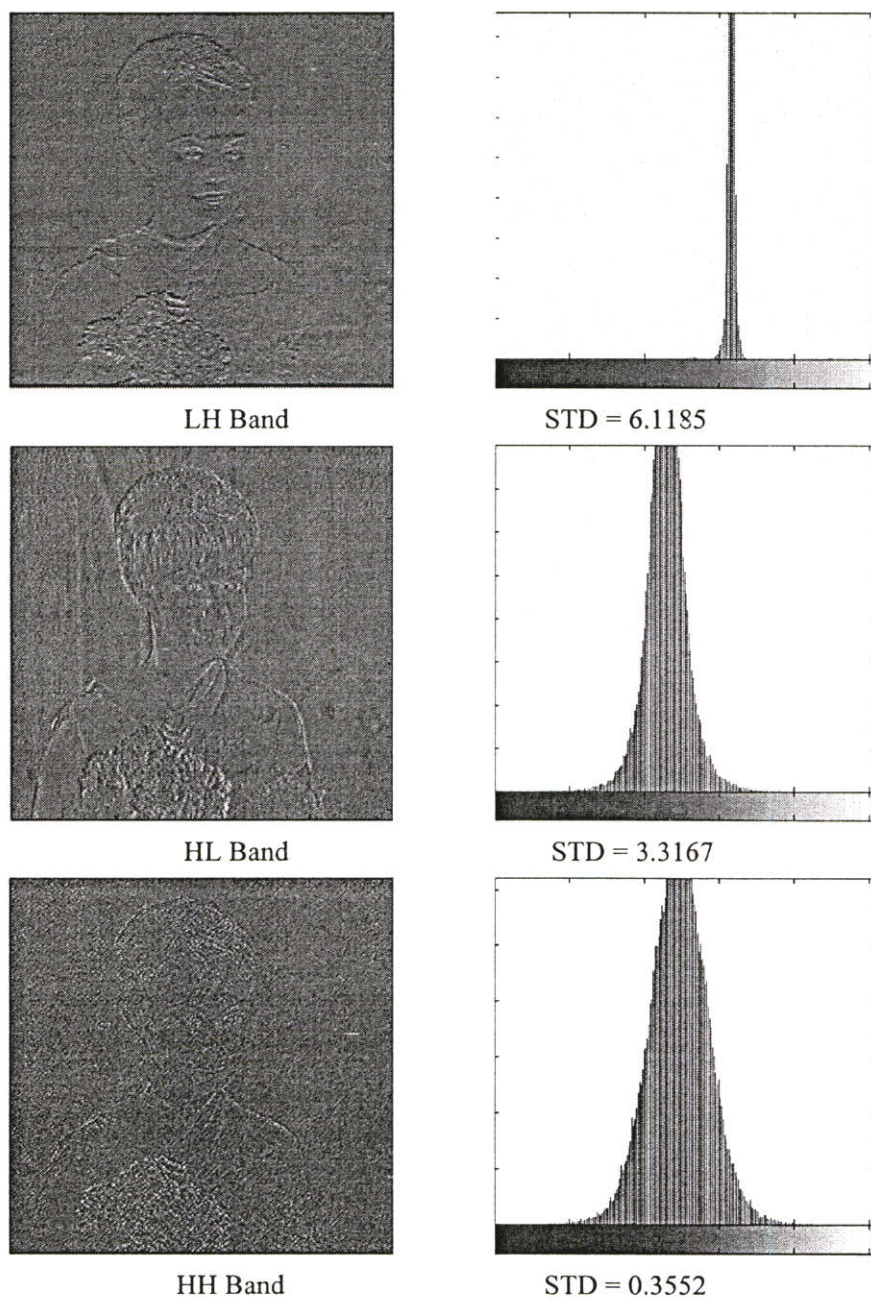


HH Band



STD = 0.2979

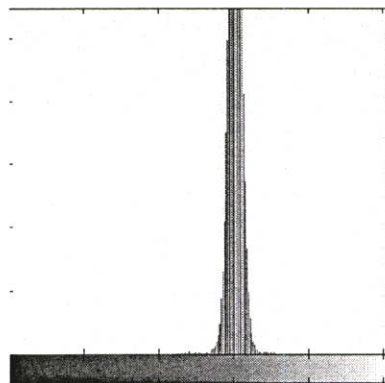
รูปที่ 5.2 ฮิสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Lena



รูปที่ 5.3 ฮีสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Girl



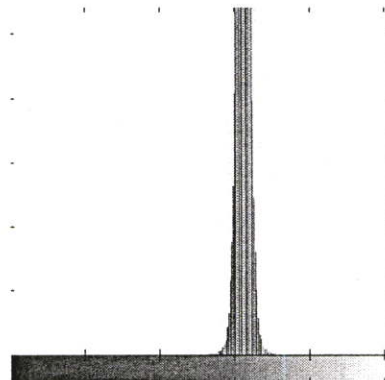
LH Band



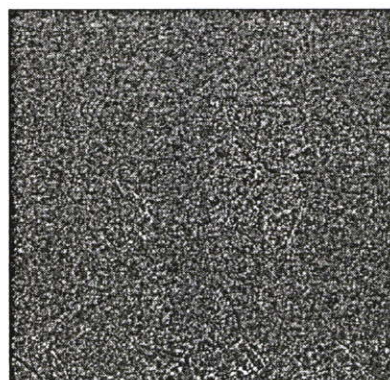
STD = 1.6240



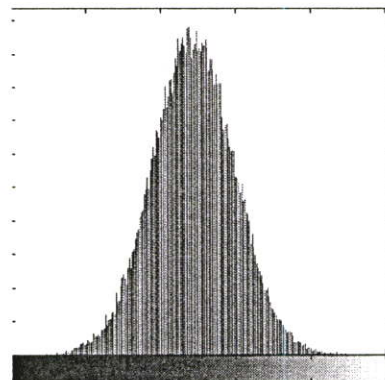
HL Band



STD = 1.6865



HH Band

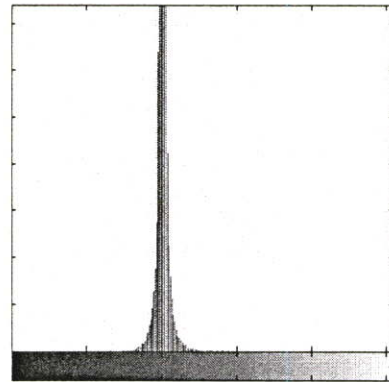


STD = 0.2480

รูปที่ 5.4 ฮิสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของภาพ Women



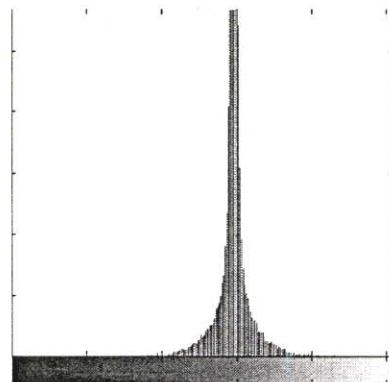
LH Band



STD = 2.8008



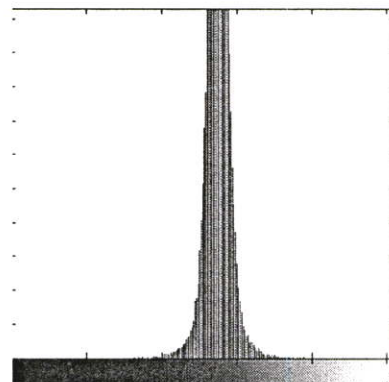
HL Band



STD = 5.8528



HH Band



STD = 0.3910

รูปที่ 5.5 ฮิสโตแกรมและค่าเบี่ยงเบนมาตรฐานที่ได้จากแบนด์ LH HL และ HH ของ ภาพ Cameraman

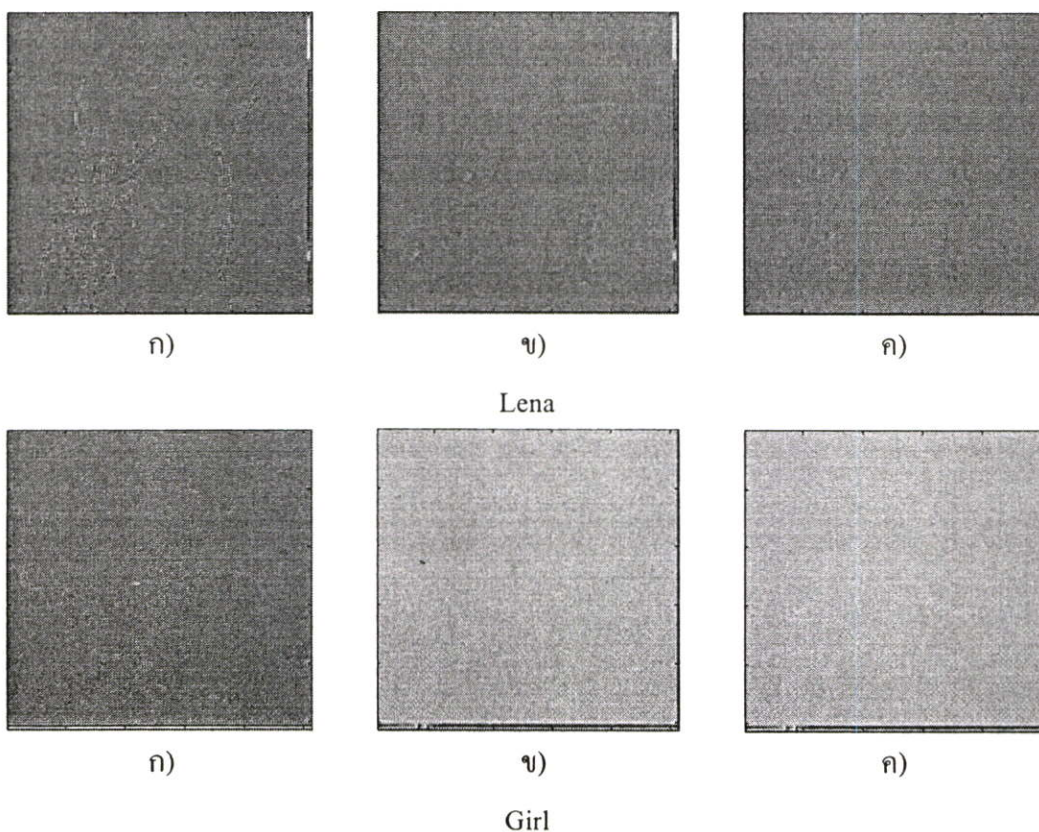
จากรูปที่ 5.2 5.3 5.4 และ 5.5 จะเห็นว่าแบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงสุดนั้นจะมีฮิสโตแกรมที่มีลำดับของกราฟต่อเนื่องกัน และจากตารางที่ 5.1 จะเห็นว่าค่าเบี่ยงเบนมาตรฐานของแบนด์ต่างๆ นั้น จะไม่อยู่ที่แบนด์ใดแบนด์หนึ่งแต่จะขึ้นอยู่กับลักษณะของภาพที่นำมาใช้

ตารางที่ 5.1 แสดงค่าเบี่ยงเบนมาตรฐานของภาพในแบนด์ต่างๆ

ภาพ	ค่าเบี่ยงเบนมาตรฐาน		
	LH	HL	HH
Lena	1.7861	4.7285	0.2979
Girl	6.1185	3.3167	0.3552
Women	1.6240	1.6865	0.2480
Cameraman	2.8008	5.8528	0.3910

5.3 ผลการเข้ารหัสแบบ VPIC

จากค่าเบี่ยงเบนมาตรฐานที่ได้ทำให้สามารถระบุแบนด์ที่จะนำไปเข้ารหัส VPIC โดยการเข้ารหัสแบบ VPIC ในงานวิจัยนี้จะกำหนดค่า gradient = 10 และ 30

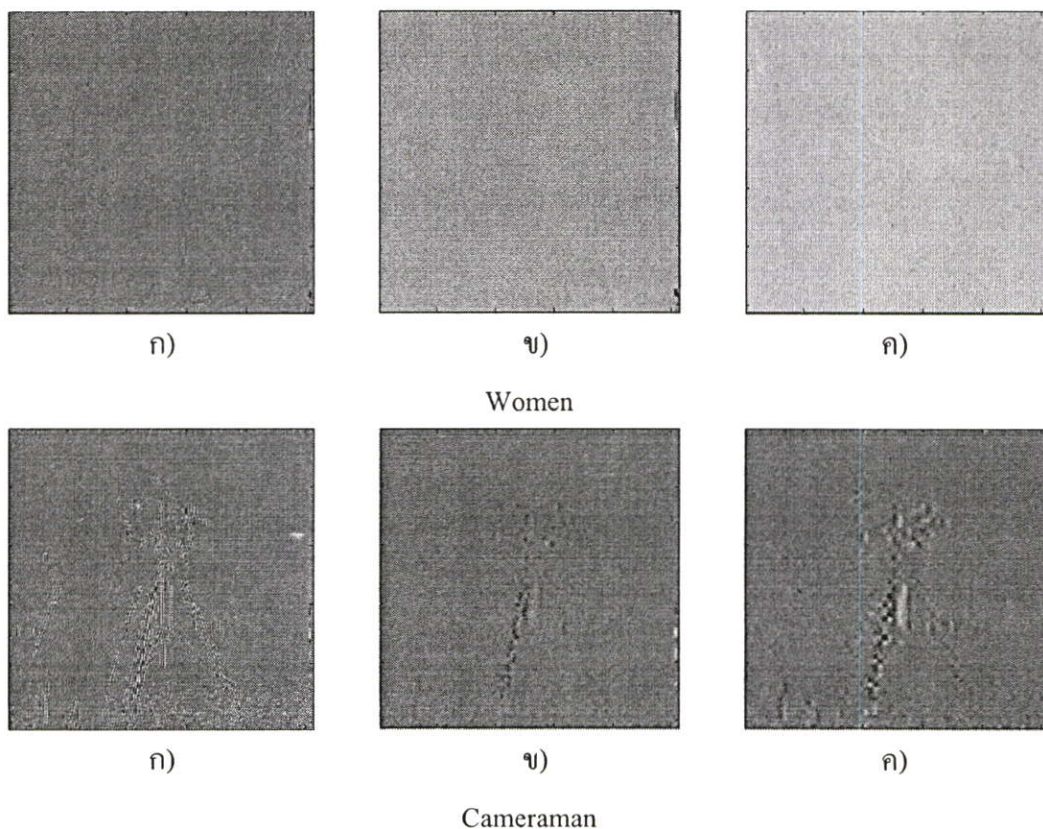


รูปที่ 5.6 ภาพที่ได้หลังจากนำแบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงสุดมาผ่านการบีบอัดด้วย VPIC

(ก) ภาพในแบนด์ต้นฉบับ

(ข) ภาพที่ผ่านการบีบอัดโดยมีค่า gradient = 10

(ค) ภาพที่ผ่านการบีบอัดโดยมีค่า gradient = 30



รูปที่ 5.6 (ต่อ) ภาพที่ได้หลังจากนำแบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงสุดมาบีบอัดด้วย VPIC

(ก) ภาพในแบนด์ต้นฉบับ

(ข) ภาพที่ผ่านการบีบอัดโดยมีค่า gradient = 10

(ค) ภาพที่ผ่านการบีบอัดโดยมีค่า gradient = 30

จากรูปที่ 5.6 จะเห็นว่าการเข้ารหัสแบบ VPIC นั้นเมื่อกำหนดค่า gradient = 10 ภาพที่ได้จะดีกว่าค่า gradient = 30 แต่จะใช้เวลาในการเข้ารหัสมากกว่าและมีอัตราการบีบอัดที่น้อยกว่า ทำให้ขนาดข้อมูลที่ได้ใหญ่กว่าด้วยตารางที่ 5.2 โดยขนาดของข้อมูลของแบนด์ที่นำมาทำการเข้ารหัสเท่ากับ 65,536 ไบต์

ตารางที่ 5.2 ขนาดข้อมูลของภาพในแบนด์เมื่อผ่านการเข้ารหัส VPIC

ภาพ	แบนด์	Compression ratio(Cr)		ขนาดข้อมูล (บิต)	
		10	30	10	30
Lena	HL	17.9066	18.1039	29,279	28,960
Girl	LH	17.8463	17.8463	29,378	29,378
Women	HL	18.1937	18.2704	28,817	28,696
Cameraman	HL	18.0627	18.2844	29,026	28,674

5.4 การเรียกคืนภาพ



ก)



ข)



ค)

Lena



ก)



ข)



ค)

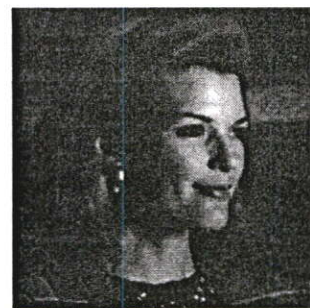
Girl



ก)



ข)



ค)

Women



ก)



ข)



ค)

Cameraman

รูปที่ 5.7 ผลลัพธ์ที่ได้จากการเรียกคืน (ก) ภาพต้นฉบับ (ข) gradient = 10 (ค) gradient = 30

ตารางที่ 5.3 ค่า MSE

ภาพ	MSE	
	10	30
Lena	4.7735	5.0040
Girl	6.2108	6.2108
Women	1.4746	1.5046
Cameraman	10.5215	10.6232

ในการกำหนดค่า gradient นั้น จะมีผลต่อการเข้ารหัสแบบ VPIC ดังแสดงในตารางที่ 5.4 โดยจะเห็นว่า การกำหนดค่า gradient ที่ต่ำนั้นจะทำให้อัตราการเจอบล็อกย่อยที่เป็นขอมมีมากขึ้น เมื่อบล็อกย่อยที่เป็นขอมมีจำนวนเพิ่มขึ้นจะทำให้ต้องใช้เวลาในการเปรียบเทียบบล็อกรูปแบบมากขึ้น และอัตราการบีบอัดข้อมูลก็จะลดลงด้วยเช่นกัน แต่จะมีผลดีกับค่า MSE เพราะจะทำให้ค่า MSE ต่ำลง

ตารางที่ 5.4 แสดงค่าผลลัพธ์ต่างๆ ของภาพต้นฉบับเมื่อทำการบีบอัดข้อมูลแบบการเข้ารหัสภาพแบบ VPIC โดยมีค่า gradient เท่ากับ 10 และ 30

Lena	Gradient	10	30
	Edge block	55	26
	Uniform block	4041	4070
	compression ratios	17.9066	18.1039
	MSE	4.7735	5.0040
Girl	Gradient	10	30
	Edge block	64	64
	Uniform block	4032	4032
	compression ratios	17.8463	17.8463
	MSE	6.2108	6.2108
Woman	Gradient	10	30
	Edge block	13	2
	Uniform block	4083	4094
	compression ratios	18.1937	18.2704
	MSE	1.4746	1.5046

ตารางที่ 5.4(ต่อ)

Cameraman	Gradient	10	30
	Edge block	32	0
	Uniform block	4064	4096
	compression ratios	18.0627	18.2844
	MSE	10.5215	10.6232

จากตารางที่ 5.5 เป็นการเปรียบเทียบค่า MSE ที่ได้จากการนำแบนด์ LL เพียงแบนด์เดียว และแบนด์ LL กับแบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงสุดเปรียบเทียบกับวิธีการบีบอัดข้อมูลในงานวิจัยนี้ จะเห็นว่าค่า MSE ของวิธีการในงานวิจัยนี้มีค่ามากกว่าวิธีการนำแบนด์ LL กับแบนด์ที่มีค่าเบี่ยงเบนมาตรฐานสูงสุดมาแปลงกลับ แต่จะมีขนาดของข้อมูลที่มากกว่าแบนด์ LL เพียงเล็กน้อย ดังแสดงในตารางที่ 5.6

ตารางที่ 5.5 แสดงค่า MSE ของภาพที่เข้ารหัสแบบ VPIC ที่มีค่า gradient 10 และ 30 กับภาพที่ไม่ได้ผ่านการเข้ารหัส

ภาพ	MSE			
	LL	LL + Max STD	LL+ Max STD (VPIC) $ \nabla I = 10$	LL+ Max STD (VPIC) $ \nabla I = 30$
Lena	6.9295	1.0933	4.7735	5.0040
Girl	12.9019	3.0796	6.2108	6.2108
Woman	1.7826	0.8913	1.4746	1.5046
Cameraman	11.1846	2.2766	10.5215	10.6232

ตารางที่ 5.6 แสดงขนาดของข้อมูลที่ได้จากวิธีการต่างๆ

ภาพ	ขนาดของข้อมูล (ไบนารี)			
	LL	LL + Max STD	LL+ Max STD (VPIC) $ \nabla I = 10$	LL+ Max STD (VPIC) $ \nabla I = 30$
Lena	65,536	131,072	69,195.8	69,155.99
Girl			69,208.24	69,208.24
Woman			69,138.13	69,123
Cameraman			69,164.3	69,120.3

ตารางที่ 5.7 ค่า MSE ของภาพที่ได้จากการเข้ารหัสแบบ VPIC และวิธีในงานวิจัย

ภาพ	VPIC/Cr		LL+ Max STD(VPIC) /Cr	
	Threshold		Threshold	
	10	30	10	30
Lena	34.6200/12.2586	45.8000/16.0219	4.7735/3.7884	5.0040/3.7906
Girl	57.3140/12.5104	71.5340/17.2295	6.2108/3.7877	6.2108/3.7877
Woman	9.8510/15.2090	13.2070/17.6457	1.4746/3.7916	1.5046/3.7924
Cameraman	62.4610/12.8424	74.7470/15.9109	10.5215/3.7901	10.6232/3.7925

บทที่ 6

สรุปผลการทดลอง

วิทยานิพนธ์นี้ได้นำเสนอการบีบอัดข้อมูลภาพแนวทางใหม่ โดยการประยุกต์ใช้การแปลงเวฟเล็ตและการเข้ารหัสแบบ VPIC มาทำงานร่วมกัน ซึ่งจะทำให้ขนาดของข้อมูลที่ได้มีขนาดเล็ก ส่วนค่าผิดเพี้ยนยังอยู่ในเกณฑ์ที่ดี

6.1 สรุปและวิเคราะห์ผลการดำเนินงานวิจัย

จากผลการทดลองของงานวิจัยนี้แสดงให้เห็นว่าการแปลงเวฟเล็ตที่ระดับสูงนั้นจะให้ผลดีในการเลือกแบนด์มาเข้ารหัส VPIC ทำให้การเลือกแบนด์มีประสิทธิภาพที่ดี ในส่วนของการเข้ารหัสแบบ VPIC นั้นค่า gradient ที่มากหรือน้อยจะมีผลต่อประสิทธิภาพและความเร็วในการเข้ารหัสเช่นกัน

ในการนำแบนด์ย่อยของเวฟเล็ต มาทำการเข้ารหัสแบบ VPIC นั้น แบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต จะต้องเป็นแบนด์ที่ได้แยกองค์ประกอบของความถี่ต่ำและความถี่สูงออกจากกันอย่างชัดเจน ทำได้โดยการแปลงเวฟเล็ตในระดับสูง เมื่อแต่ละแบนด์ถูกแยกความถี่ออกจากกัน จะทำให้การเลือกแบนด์มาเข้ารหัสแบบ VPIC เป็นไปได้ถูกต้องมากขึ้น

การกำหนดค่า gradient นั้น เมื่อกำหนดค่า gradient เท่ากับ 10 อัตราส่วนที่บล็อกละย่อยจะเป็นบล็อกละย่อยที่เป็นขอบจะมีมาก เมื่อมีจำนวนบล็อกละย่อยที่เป็นขอบมากเวลาที่ใช้ในการเปรียบเทียบกับบล็อกรูปแบบจะใช้นานขึ้นด้วย ค่าความผิดเพี้ยนจะอยู่ในเกณฑ์ที่ดีว่าการกำหนดค่า gradient ที่สูง

เมื่อกำหนดค่า gradient เท่ากับ 30 อัตราส่วนที่บล็อกละย่อยจะเป็นบล็อกละย่อยที่เป็นขอบจะมีน้อยกว่าเมื่อเทียบกับค่า gradient ที่ต่ำกว่า แต่เมื่อจำนวนบล็อกละย่อยที่เป็นขอบน้อยเวลาที่ใช้ในการเปรียบเทียบกับบล็อกรูปแบบก็จะน้อยลงเช่นกัน แต่ค่าความผิดเพี้ยนจะน้อยกว่า

6.2 ปัญหา

จากการทดลอง ขนาดและรายละเอียดของภาพที่จะนำมาเข้ารหัส VPIC นั้นมีความสำคัญมาก เนื่องจากการเข้ารหัสแบบ VPIC เป็นการเข้ารหัสบล็อกละย่อยขนาด 4x4 ถ้าภาพที่จะนำมาเข้ารหัสมีขนาดเล็กหรือมีรายละเอียดซับซ้อน เมื่อทำการแบ่งบล็อกละย่อยออกมาบล็อกละย่อยนั้นอาจจะมีรายละเอียดหรือรูปแบบภายในมาก แต่การเข้ารหัสแบบ VPIC จะทำการเลือกรูปแบบบล็อกรูปแบบที่มีความเหมือนมากที่สุดเพียงแบบเดียว รายละเอียดที่รองลงมาจะหายไป

เมื่อกำหนดค่า gradient ที่แตกต่างกันผลลัพธ์ที่ได้ก็จะแตกต่างกันด้วย ซึ่งค่า gradient ที่มีค่าต่ำอาจจะดีกับภาพบางลักษณะ ส่วนค่า gradient ที่มีค่าสูงก็เช่นกัน การกำหนดค่า gradient นั้น ยังไม่มีความแน่นอน แต่ถ้ากำหนดค่า gradient ที่ต่ำแม้ผลลัพธ์จะออกมาดี แต่เวลาที่ใช้ในการบีบอัดจะมากขึ้น ขนาดของข้อมูลก็จะมากขึ้นด้วย

6.3 แนวทางการพัฒนาต่อ

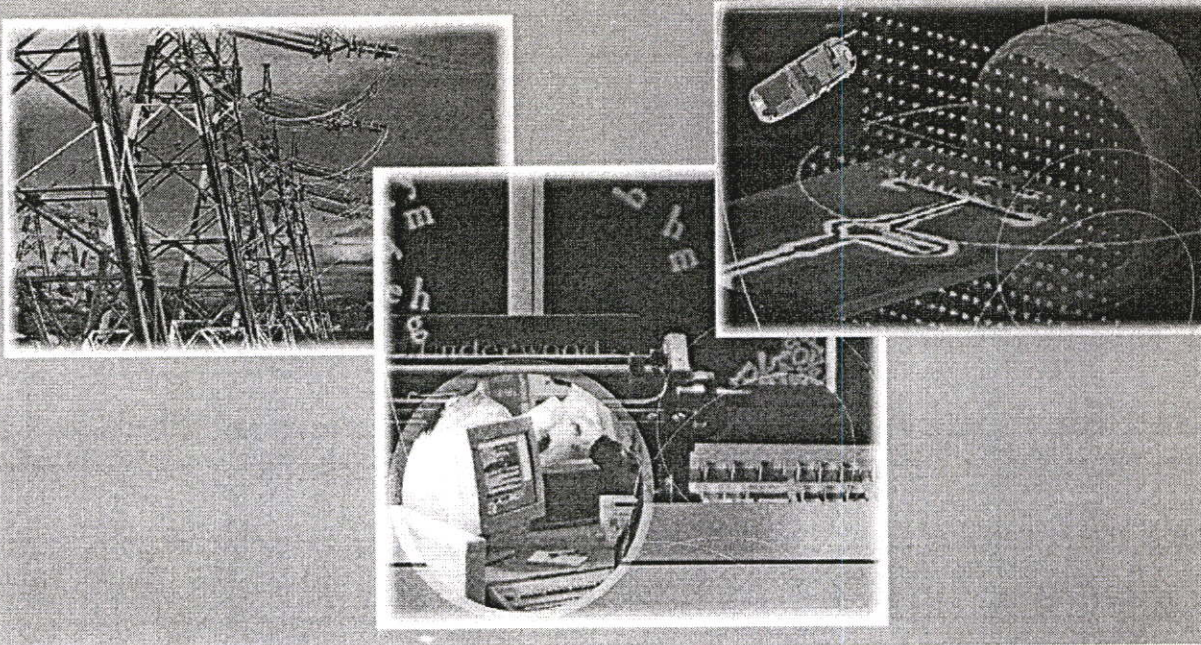
จากผลการทดลอง จะเห็นว่าในขั้นตอนการกำหนดค่า gradient ที่เหมาะสมกับภาพแต่ละภาพนั้นยังไม่มีวิธีการที่แน่นอนเพื่อนำมาใช้ในการกำหนดค่า gradient ให้กับภาพ

เอกสารอ้างอิง

- [1] J.W. Woods and S.D. O'Neil, "Subband Coding of Images," IEEE Trans. Acoust. Speech Signal Process., ASSP-34:1288, Oct. 1986.
- [2] C.E. Shannon, "A Mathematical Theory of Communication," Bell Syst. Tech.J. Vol.27, pp.398-403 , July 1948.
- [3] D.A. Huffman, "A Method for the Construction of Minimum-redundancy Codes." Proc. of The I.R.E., Vol.40, pp.1098-1101, 1952.
- [4] I.H. Witten, R. Neal and J.G. Cleary. "Arithmetic Coding for Data Compression," Communication of the ACM, 30(6):520-541, 1987.
- [5] A. Jensen and A. Cour-Harbo, Ripples in Mathematics : The Discrete Wavelet Transform, Springer Verlag, 2001.
- [6] S.G. Mallet, "A Theory for Multi-Resolution Signal Decomposition: The Wavelet Representation" , IEEE Trans. Pattern Analysis and Machine Intelligence., Vol. 11, No. 7, pp. 674-693, July 1989.
- [7] M. Antonini, M. Barlaud, P. Matchieu and I. Daubechies, "Image Coding using Wavelet Transform", IEEE Trans. On Image Processing, Vol.1, No.2, pp.205-220, April 1992.
- [8] D. Chen and A. C. Bovik, "Visual pattern image coding," IEEE Trans. Communication., Vol. COM-38, No. 12, pp. 2137-2146, Dec. 1990
- [9] P.A. Laplante and A. D. Stoyenko, "Real-Time Imageing : Theory, Techniques, and Applications," IEEE Press, 1996
- [10] G. Wallace, "The JPEG Still picture compression standard," Communication of the ACM, vol. 34, no. 4, April 1991, pp. 30-44.

ภาคผนวก ก

ผลงานวิจัยที่ได้รับการตีพิมพ์



ECTI-CON 2007

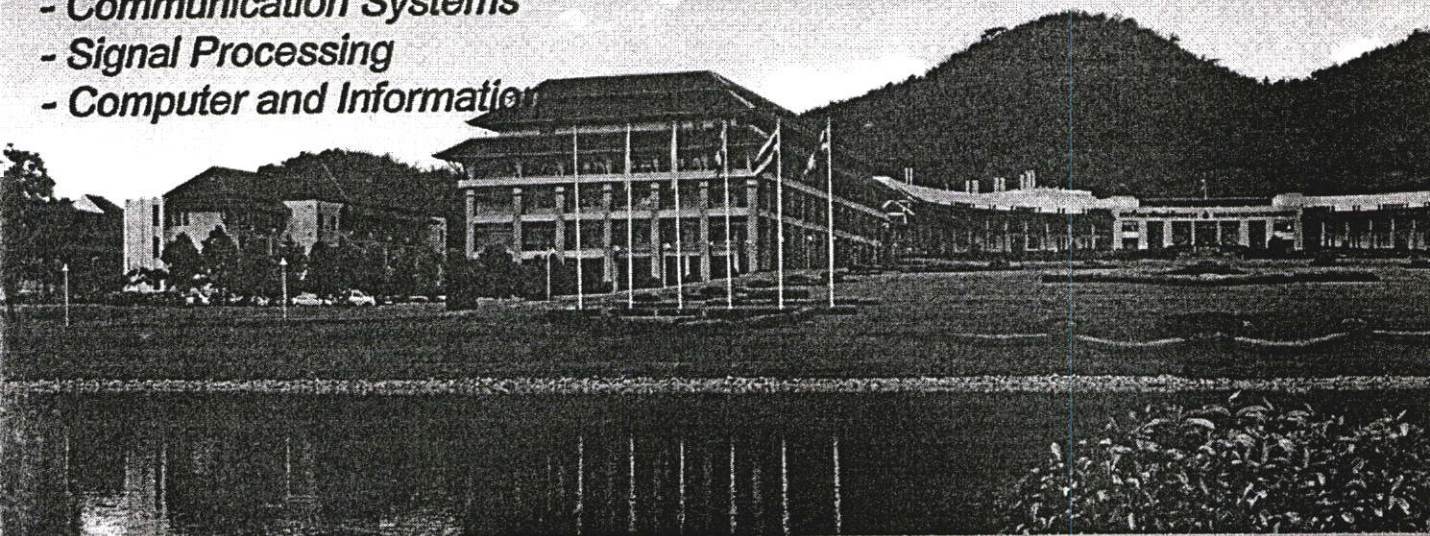
*Mae Fah Luang University, Chiang Rai, Thailand
May 9-12, 2007*

VOLUME 1

- *Circuits and Systems*
- *Control Engineering*
- *Electrical Power Engineering*
- *Other Related Fields*

VOLUME 2

- *Communication Systems*
- *Signal Processing*
- *Computer and Information*



ECTI
Association

IEEE
THAILAND SECTION

NECTEC
a member of NSTDA

WD Western
Digital



Digital Image Compression using VPIC algorithm in Wavelet Domain

K.Tongtar, N. Moongflangklang and F.Cheevasuvit

Department of Instrumentation Engineering, Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang Bangkok, Thailand 10520.
s5061229@kmitl.ac.th, nanfa_blue@yahoo.com, kcfusak@kmitl.ac.th

Abstract-This paper proposes an efficient method to improve the quality of digital image compression using Visual Pattern Image Coding (VPIC) in wavelet domain. Wavelet based compression gives higher quality of the reconstructed image with high PSNR and less amount of minimum mean square error compared to VPIC. However, its compression ratio is reasonably low. Therefore, the proposed method combines the advantages obtained from both wavelet transform and VPIC to achieve high compression ratio while preserve the quality of the reconstructed image. The proposed algorithm can be separated in to three main processes. First, wavelet transform is used to decompose the low and high frequency coefficients. Then, the LL coefficients are reserved, while one of the high frequency sub-bands (HL, LH, and HH) is selected based on its Standard Deviation (SD). The sub-band having the highest information distribution usually contains proper edge information, which could be used to improve VPIC algorithm. Finally, the LL and the selected sub-bands are decomposed to obtain the result image. The experimental results show that the proposed compression algorithm can achieve high compression ratio while still maintain good image fidelity and minimum mean square error.

I. INTRODUCTION

Several algorithms are proposed for digital image compression such as Block Truncation (BTC), Vector Quantization (VQ) and Visual Pattern Image Coding (VPIC). BTC[1], [2] was first proposed in 1979 by Delp and Mitchell, to compress monochrome images. BTC is an efficient method for block correlated signals. It quantizes block data into two values by preserving the first moment and the second moment. The advantages of BTC are its simple computation, easy implementation and fine edge preservation. The main drawback of BTC is its high bit rate. VQ[3] was proposed in 1980 by Linde, Buzo and Gray also called "LBG" algorithm. It separates an image into blocks (4x4 pixels), which we can determine intensity vector of pixels. Drawback of this method is its high complexity coding and low quality of edge image.

BTC and VQ are similar in that the image to be coded is first partitioned into a large number of smaller sub-images or blocks. Each block is then coded either by extracting certain intrinsically high-information quantities which are transmitted in reduced form(as in BTC), or by marking a comparison between the block and a set of preordained vectors or "typical" sub-image, the index of the "closest" vector being transmitted instead of the original block(as in VQ).

VPIC[4] was proposed in 1990 by Chen and Bovik. High compression ratio with low computational complexity is the

main advantage for VPIC. However, the algorithm produces higher minimum mean square error and lower PSNR compared to other techniques.

This paper is organized as follows. Section II describes the overview of proposed method. Section III presents applying of VPIC. Section IV is experimental results and the final is the conclusions.

II. OVERVIEW OF PROPOSED METHOD

In this paper, we presents the compression method applying VPIC encoding with selection of wavelet sub-bands containing 4 separated the frequency bands (LL, LH, HL and HH). Only the LL sub-band is reserved with no modification. One of the other sub-bands is selected based on maximum SD value. This sub-band is used as input to VPIC. Then, the LL sub-band and the result of VPIC are used as input to inverse wavelet transform in order to compressed information as shown in figure 1.

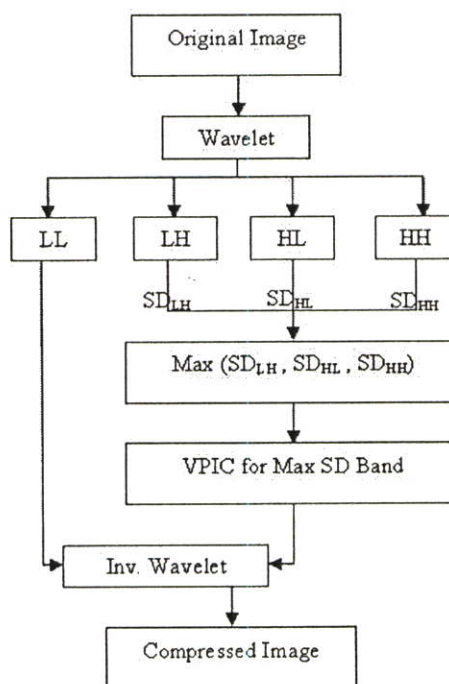


Fig.1 Show the overview of proposed method.

III. IMPLEMENTATION OF THE PROPOSED ALGORITHM

In this section, we apply VPIC to encode image information into binary blocks, and then each binary block is classified as the uniform or edge block based on its gradient magnitude in final stage.

A. Image Coding

In order to reduce the computation time, in the first stage, algorithm is initiated by separating image into sub-blocks (4x4 pixels). Then, the mean intensity of the pixel in each sub-block is calculated as shown in (1). This mean value is used to translate the intensity sub-block into binary sub-block. If the intensity value in the pixel is equal or greater than the mean intensity value, we assign the value of "1" to that pixel, otherwise assign "0".

$$Mean = \left(\sum_{i=1}^4 \sum_{j=1}^4 I(i, j) \right) / 16 \tag{1}$$

where $I(i, j)$ is sub-block of image.

Next stage, determine the value of $\Delta_x b_{i,j}$ and $\Delta_y b_{i,j}$ of each sub-block illustrated in (2) and (3) respectively to calculate the value of gradient magnitude and orientation $\angle \Delta b_{i,j}$, which can be calculated using (4) and (5) respectively. The gradient value is used to separate between uniform sub-block and edge sub-block while the $\angle \Delta b_{i,j}$ is used to determine the direction of edge in each sub-block.

$$\Delta_x b_{i,j} = AVE \{ I_{n,m} : 4i+2 \leq n \leq 4i+3, 4j \leq m \leq 4j+3 \} - AVE \{ I_{n,m} : 4i \leq n \leq 4i+1, 4j \leq m \leq 4j+3 \} \tag{2}$$

$$\Delta_y b_{i,j} = AVE \{ I_{n,m} : 4i \leq n \leq 4i+3, 4j+2 \leq m \leq 4j+3 \} - AVE \{ I_{n,m} : 4i \leq n \leq 4i+3, 4j \leq m \leq 4j+1 \} \tag{3}$$

where $I_{n,m}$ is intensity of pixels in sub-block.

$$|\nabla b|_{i,j} = \sqrt{(\Delta_x b_{i,j})^2 + (\Delta_y b_{i,j})^2} \tag{4}$$

$$\angle \Delta b_{i,j} = \tan^{-1} \left(\frac{\Delta_y b_{i,j}}{\Delta_x b_{i,j}} \right) \tag{5}$$

Ultimately, we compare the gradient magnitude value $|\nabla b|_{i,j}$ with the threshold $|\nabla I|$ [1] to classify the class of sub-block. If the $|\nabla b|_{i,j} \leq |\nabla I|$, that sub-block is defined as uniform block, and can go to image coding stage. Otherwise, it is defined as edge block. The edge block is compared to 28 pattern blocks to find the most similarity for the next encoding process. Figure 2 shows the 14 examples of the pattern blocks.

B. Coding Analysis

In coding stage, the binary coding is constructed based on the class of each sub-block from the previous stage. The coding information composes of 6-bit mean intensity, 6-bit gradient magnitude value, 5-bit number of pattern block and 1-bit sub-block type. Moreover, two additional bits called sub-band ID are used to identify the sub-band, whose the SD value

is maximum. Figure 3 presents the data format obtained after VPIC coding. Table I shows the number of bits for each field of the information coding.

C. Decoding Analysis

At the decoder, the class of sub-block is determined from the received compressed bitstream. If the sub-block is a uniform block, a 4x4 pixel block is created and the mean intensity of each pixel is estimated. In contrast, if the sub-block is an edge block, the pattern block that matches with the information of that sub-block is applied. The mean intensity and gradient magnitude value are used to calculate low and high intensity value. The low intensity value is represented by "0" pixel, high intensity value is represented by "1" as shown in (6).

$$xH+yL=16*Mean \tag{6}$$

where x is "1" in the pattern block number
y is "0" in the pattern block number

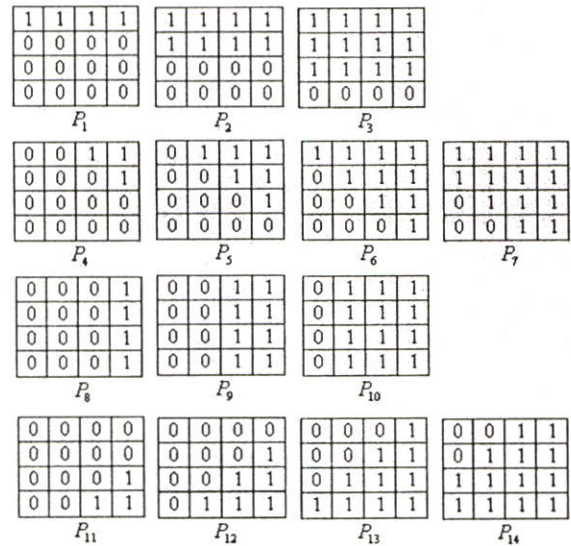


Fig.2 Show 14 examples pattern blocks (a) 90° patterns; (b) 45° patterns; (c) 0° patterns; (d) -45° patterns.

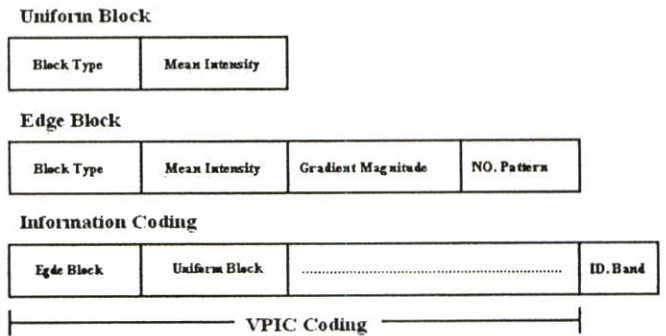


Fig.3 Coding pattern of proposed method.

TABLE I
DETAIL OF INFORMATION CODING.

	Uniform Block	Edge Block
Block type	1	1
Mean intensity	6	6
Gradient	0	6
Pattern number	0	5
Total	7	18

For example, we would like to find the low and high intensity value of pattern block P1 as shown in figure 2.

Given

$$4H+12L=16*Mean \quad (6.1)$$

From (5)

$$|\nabla b|_{i,j} = \sqrt{(\Delta x b_{i,j})^2 + (\Delta y b_{i,j})^2}$$

Where $(\Delta x b_{i,j}) = \frac{4H-4L}{8}$ and $(\Delta y b_{i,j}) = 0$

$$|\nabla b|_{i,j} = \sqrt{\left(\frac{4H-4L}{8}\right)^2}$$

$$8|\nabla b_{i,j}| = 4H-4L \quad (6.2)$$

Given (6.1)-(6.2)

$$16L = 16*Mean - 8|\nabla b|_{i,j}$$

$$L = \frac{16*Mean - 8|\nabla b|_{i,j}}{16} \quad (7)$$

Substitute L in (6.2)

$$H = \frac{16*Mean - 12L}{4} \quad (8)$$

Equation (7) and (8) can be used to calculate the low and high intensity value of P1 pattern block.

IV. EXPERIMENTAL

This section presents the experimental results of the proposed compression techniques. The experiments are conducted using the grayscale images of "Lena", "Girl", "Baboon", "Camera man" and "Woman with 512x512 pixels size. The original images are decomposed using wavelet transform. Daubechies 9 wavelet is used in this paper. The wavelet coefficients of 4 sub-bands (LL, LH, HL, and HH) are coded at 8 bits/component. Figure 3 displays wavelet coefficients for all four sub-bands of Lena image.

As previously mentioned, the LL sub-band is preserved without any modification. The VPIC coding is performed on the chosen high frequency sub-bands, which its SD value is maximum. At the decoder side, both information of the LL and the selected high frequency sub-bands are used to decode and reconstruct the result image.

For example, the SD values of all the high frequency sub-band of Lena image is presented in Table II. The sub-band

having the highest SD value is the HL band. Table III shows the MSE of the reconstructed image using different combinations of LL and one of the high frequency sub-band. The minimum MSE also occurs in the case of the combination of the LL and HL sub-band information.

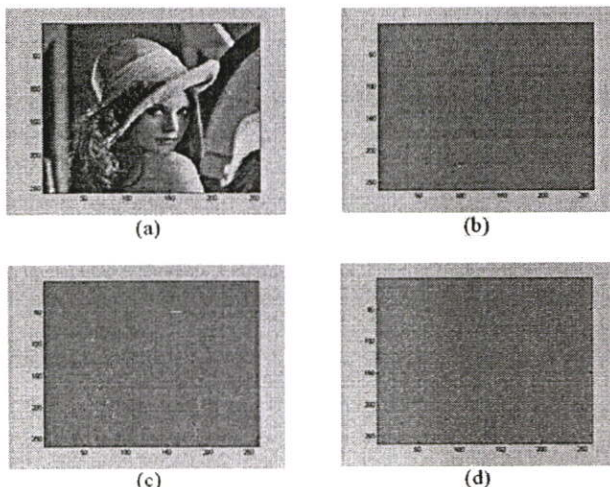


Fig.3 Result of wavelet transform of Lena
(a) LL band; (b) LH band; (c) HL band; (d) HH band.

TABLE II
STANDARD DEVIATION VALUE OF 3 SUB-BANDS OF LENA IMAGE.

Band	Standard Deviation
LH	1.7861
HL	4.7285
HH	0.2979

TABLE III
MSE OF WAVELET TRANSFORM OF LL BAND WITH OTHER BAND.

Wavelet	MSE
LL+LH	5.8584
LL+HL	1.0933
LL+HH	6.9073

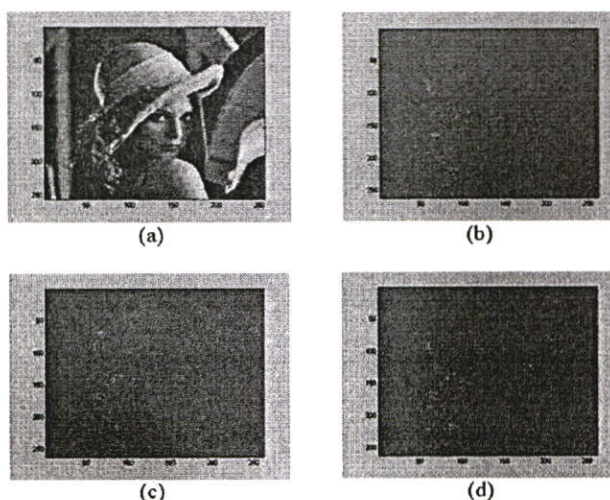


Fig. 4 (a) LL Band; (b) HL Band;
(c) gradient value =10; (d) gradient value = 30.

Figure 4 shows wavelet coefficients of the LL sub-band (a), the HL sub-band (b) of the original Lena image and those of the HL sub-band of the reconstructed image using gradient threshold = 10 (c) and 30 (d). Table IV show the results of the HL sub-band after VPIC coding with different gradient thresholds. The higher the value of the threshold, the lower the chance to find the edge blocks. As the number of the edge blocks decrease, the compression ratio will be higher.

Table V compares the MSE values of different tested images obtained from VPIC and the proposed technique (Wavelet and VPIC) algorithm with various gradient thresholds.

TABLE IV

THE INFORMATION OF LENA IMAGE IN HL BAND WITH PROPOSED METHOD.

	Gradient (threshold) = 10	Gradient (threshold) = 30
Edge block	55	26
Uniform block	4041	4070
compression ratios	17.9066	18.1039

TABLE V

THE MSE VALUE OF VPIC AND THE PROPOSED METHOD

Image Name	VPIC		Wavelet & VPIC	
	Gradient (threshold)		Gradient (threshold)	
	10	30	10	30
Lena	34.62	45.8	4.8	5
Cameraman	62.4613	74.7471	10.5215	10.6232
Woman	9.8507	13.2069	1.4746	1.5046
Girl	57.3144	71.5343	6.2108	6.2108
Baboon	287.5341	327.7233	53.2742	53.4756

TABLE VI

THE RELATION OF INFORMATION SIZES AND QUALITY FACTORS

Feature	Wavelet	VPIC		Wave & VPIC	
		Gradient (threshold)		Gradient (threshold)	
		10	30	10	30
Information size	130KB	21.1KB	16.3KB	68.7KB	68.6KB
MSE	1.0933	34.62	45.8	4.8	5
PSNR	47.7436	32.74	31.5	41.34	41.14

Table VI shows the results of information sizes and quality factors (MSE and PSNR) for three compression techniques (Wavelet, VPIC, and the proposed combination of Wavelet and VPIC). From the table V, the proposed technique offers lower size of the information while producing slightly higher MSE value compared to wavelet based technique. However, the proposed technique produces lower MSE value even though the information size is larger than that of VPIC technique.

V. CONCLUSION

This paper presents an image compression algorithm using a combination of VPIC and wavelet based coding techniques to improve compression quality factors of MSE and PSNR. The value of minimum mean square error and good image fidelity are improved by conserving the coefficients of the low frequency component (LL band) of wavelet transform which represents the smooth of image data. The coefficient of the other three sub-bands provides additional information, which is less important and can be neglected with less amount of perceptually effect. However, if some of the information from high frequency sub-bands is included in the encoded bitstream accompany with the information from LL sub-band, the quality of the reconstructed image could be improved. This paper proposed a method to choose the appropriate high frequency sub-band base on standard deviation. Then, the selected sub-band is encoded using VPIC technique. Consequently, the proposed compression algorithm can achieve high compression ratio while still maintains good image fidelity and minimum mean square error.

REFERENCES

- [1] E. J. Delp and O. R. Mitchell, "Image compression using block truncation coding," IEEE Trans. Commun., vol. COM-27, pp. 1335-1341, Sept. 1979.
- [2] O. R. Mitchell and E. J. Delp, "Multilevel Graphics representation using block truncation coding," Proc. IEEE, vol.68, July 1980
- [3] A. Gersho, R.M. Gray, "Vector Quantization and Signal Compression", Kluwer Academic Publishers 1992.
- [4] D.Chen and A. C. Bovik, "Visual pattern image coding," IEEE Trans. Commun.,vol. COM-38,no. 12, pp. 2137-2146, Dec. 1990

ประวัติผู้เขียน

ชื่อ-นามสกุล	นายคธาวุฒิ ทองทา
วัน เดือน ปีเกิด	10 กรกฎาคม 2516
ที่อยู่	151/3 ซ.วงศ์สว่าง 11 ถ.วงศ์สว่าง บางซื่อ กรุงเทพฯ 10800
ประวัติการศึกษา	2541 ประกาศนียบัตรวิชาชีพชั้นสูง สาขาวิชาอิเล็กทรอนิกส์ มหาวิทยาลัย สยาม 2544 อุดสาหกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์ มหาวิทยาลัยสยาม