

การลดขอบบล็อกลงในภาพ JPEG ด้วยวิธีเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้

REDUCING BLOCKINESS IN JPEG IMAGES
USING ADAPTIVE WAVELET THRESHOLDING

ศิริพร เดชชะสีดารักษ์
SIRIPORN DACHASILARUK

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2543
ISBN 974-622-725-4

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การลดขอบบล็อกรูปภาพ JPEG ด้วยวิธีเวฟเล็ตเชอร์โวลด์ดิ่งที่ปรับตัวเองได้

REDUCING BLOCKINESS IN JPEG IMAGES USING ADAPTIVE WAVELET THRESHOLDING



ศิริพร เดชะศิลาธิรักษ์

SIRIPORN - DACHASILARUK

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2543

ISBN 974-622-725-4

เลขหมู่.....
เลขทะเบียน..... 35718
วัน, เดือน, ปี 19 ต.ย. 2543

**REDUCING BLOCKINESS IN JPEG IMAGES
USING ADAPTIVE WAVELET THRESHOLDING**

SIRIPORN DACHASILARUK

**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 974-622-725-4

COPYRIGHT 2000

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

หัวข้อวิทยานิพนธ์	การลดขอบบล็อคในภาพ JPEG ด้วยวิธีเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้
นักศึกษา	นางสาวศิริพร เดชะศิริรักษ์
รหัสประจำตัว	39061085
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ.ศ.	2543
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ. ดร. บุทธพงษ์ รังสรรค์เสรี

บทคัดย่อ

ความคิดเพี้ยนแบบบล็อคของข้อมูลภาพเกิดขึ้นเมื่อทำการบีบอัดข้อมูลภาพที่ใช้พื้นฐานของการแปลงดิคริตโคไซน์แบบบล็อคที่อัตราการบีบอัดข้อมูลภาพที่สูงหรือที่บิดเบลดำ โดยทั่วไปแล้วการลดขอบบล็อคข้อมูลภาพจะทำในขั้นตอนของการประมวลผลหลังการขยายข้อมูลภาพ เนื่องจากไม่ขึ้นอยู่กับขั้นตอนการเข้ารหัสและถอดรหัสข้อมูลภาพ นอกจากนี้ยังคงรักษาบิดเบลดำไว้เหมือนเดิม ในวิทยานิพนธ์ฉบับนี้ได้นำเสนอการลดขอบบล็อคด้วยวิธีเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้ ซึ่งเป็นวิธีการที่ไม่ซับซ้อนและยังคงรักษาลักษณะที่สำคัญของภาพไว้ นอกจากนี้ได้ทำการศึกษาเปรียบเทียบเทคนิคการลดขอบบล็อควิธีอื่นๆ ได้แก่ การกรองความถี่ต่ำผ่าน การกรอง Wiener ที่ปรับตัวเองได้และการทำเวฟเล็ตเรโซลคิง ผลการทดลองที่ได้จากการลดขอบบล็อคด้วยเทคนิคต่างๆ พบว่าเทคนิคที่ได้นำเสนอจะให้ผลภาพดีที่สุด เมื่อทำการประเมินผลภาพที่ได้จากการวัดค่าอัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) ค่าความแปรปรวนที่บริเวณขอบบล็อค (EV) และแสดงผลภาพที่ได้

Thesis Title	Reducing Blockiness in JPEG Images Using Adaptive Wavelet Thresholding
Student	Ms. Siriporn Dachasilaruk
Student ID.	39061085
Degree	Master of Engineering
Programme	Electrical Engineering
Year	2000
Thesis Advisor	Assoc. Prof. Dr. Yuttapong Rangsanseri

ABSTRACT

Blockiness occurs when using compression techniques based on block discrete cosine transform (BDCT) at high compression ratios or low bit rate. A commonly used approach for image deblocking is proposed as post-processing mechanisms after decompression, because this technique is independent of coding schemes and provides better visual quality with the same bit rate while keeping codec intact. This thesis presents spatial adaptive wavelet thresholding for reducing blockiness. This technique is simple implement and maintains feature of image. Moreover, a comparative study of various post-processing techniques is presented. The techniques investigated include low-pass filtering, adaptive Wiener filtering, wavelet thresholding, and the proposed technique. The results obtained indicate that, among the techniques studied, the proposed technique gives best performance when assessed by using peak signal to noise ratio (PSNR), edge variance (EV), and visual performance.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยความช่วยเหลืออย่างดียิ่งจากหลายๆ ฝ่าย ซึ่งผู้จัดทำใคร่ขอขอบคุณทุกๆ ท่านที่มีส่วนร่วมสนับสนุนช่วยเหลือและแนะนำในทุกๆ ด้าน

ขอขอบพระคุณ รศ.ดร.บุทธิพงษ์ รัสสุรงค์เสรี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษาแนะนำและข้อเสนอแนะที่เป็นประโยชน์ตลอดจนห้องทดลองและการทำงานต่อผู้วิจัยตลอดมา

ขอขอบคุณอาจารย์ปัญญา ฐิติมัทธมาและอาจารย์ทุกท่านที่ได้กรุณาให้คำแนะนำต่างๆ และประสิทธิ์ประสาทความรู้แก่ผู้วิจัย ขอขอบพระคุณท่านกรรมการสอบทุกท่านที่ได้สละเวลาอันมีค่าในการตรวจสอบวิทยานิพนธ์ในครั้งนี้

ขอกราบขอบพระคุณบิดา มารดา และพี่ๆ น้องๆ ทุกคน ที่ให้โอกาสและกำลังใจต่อผู้วิจัย

ขอขอบคุณพี่ๆ และเพื่อนๆ นักศึกษาทุกคนที่ช่วยเหลือให้คำแนะนำต่างๆ จนวิทยานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ด้วยดี

สุดท้ายขอขอบคุณบัณฑิตวิทยาลัย ที่ได้ให้ทุนสนับสนุนการทำวิทยานิพนธ์ครั้งนี้

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

ศิริพร เศษะศิริรักษ์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญภาพ	VIII
บทที่ 1 บทนำ	1
1.1 ที่มาของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	2
1.3 ขอบเขตของการวิจัย	2
บทที่ 2 หลักการพื้นฐานของการบีบอัดข้อมูลภาพ	5
2.1 การจำแนกประเภทของการบีบอัดข้อมูลภาพ	5
2.2 เทคนิคการบีบอัดข้อมูลภาพแบบต่างๆ	9
2.2.1 มาตรฐานการบีบอัดข้อมูลภาพนิ่ง JPEG	9
2.2.2 การบีบอัดข้อมูลภาพด้วยการเข้ารหัสการควอนไทซ์แบบเวกเตอร์ ...	12
2.2.3 การบีบอัดข้อมูลภาพเฟรคทอล	13
2.2.4 การบีบอัดข้อมูลภาพด้วยการแปลงเวฟเล็ต	15
2.3 ความผิดเพี้ยนที่เกิดจากการบีบอัดข้อมูลภาพ	16
2.3.1 ความผิดเพี้ยนแบบบล็อก	16
2.3.2 ความผิดเพี้ยนแบบวงแหวน	17
2.3.3 การลดทอนความคมชัดของภาพหรือภาพเบลอ	17
2.3.4 การเบี่ยงเบนในเนื้อหาของภาพ	18
2.3.5 ความผิดเพี้ยนเชิงเรขาคณิต	18
2.4 การวัดประสิทธิภาพการบีบอัดข้อมูลภาพ	19
2.5 สรุป	20

สารบัญ(ต่อ)

	หน้า
บทที่ 3 การแปลงเวฟเล็ด	22
3.1 การแปลงเวฟเล็ดแบบดิสครีต	22
3.1.1 การแปลงเวฟเล็ดแบบดิสครีตใน 1 มิติ	22
3.1.2 การแปลงเวฟเล็ดแบบดิสครีตใน 2 มิติ	27
3.1.3 เวฟเล็ดแบบอโรโกนัลและแบบไบอโรโกนัล	29
3.2 การหาขอบที่หลายระดับของการแปลงภาพ	31
3.2.1 การหาอนุพันธ์ด้วยการแปลงเวฟเล็ดใน 1 มิติ	31
3.2.2 การหาอนุพันธ์ด้วยการแปลงเวฟเล็ดใน 2 มิติ	35
3.2.3 การหาขอบภาพด้วยการแปลงเวฟเล็ด	39
3.3 สรุป	41
บทที่ 4 การลดขอบบล็อคในข้อมูลภาพ	41
4.1 การจำแนกประเภทวิธีการลดขอบบล็อค	41
4.2 การกรองความถี่	44
4.3 เวฟเล็ดเชรโซลต์ดิง	45
- 4.3.1 รูปแบบการทำเชรโซลต์	45
4.3.2 การเลือกชนิดและระดับของการแปลงเวฟเล็ด	48
4.3.3 การประมาณค่าเชรโซลต์	48
4.4 เวฟเล็ดเชรโซลต์ดิงที่ปรับตัวเองได้	49
4.4.1 การจำแนกลักษณะพื้นที่ของข้อมูลภาพ	49
4.4.2 การทำเวฟเล็ดเชรโซลต์ดิง	52
4.5 การวัดประสิทธิภาพของการลดขอบบล็อค	53
4.6 สรุป	53
บทที่ 5 ผลการทดลอง	54
5.1 ข้อมูลภาพที่ใช้ในการทดลอง	54
5.2 การลดขอบบล็อคด้วยการกรองความถี่	55

สารบัญ(ต่อ)

	หน้า
5.3 การลดขอบบล็อกดด้วยเวฟเล็ดเธอร์โซลดิ้ง	56
5.4 การลดขอบบล็อกดด้วยเวฟเล็ดเธอร์โซลดิ้งที่ปรับตัวเองได้	59
บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ	70
6.1 สรุปผลการวิจัย	70
6.2 ปัญหาที่เกิดขึ้นและข้อเสนอแนะ	71
เอกสารอ้างอิง	72
ภาคผนวก	74
ภาคผนวก ก. การกรอง Wiener ที่ปรับตัวเองได้	75
ภาคผนวก ข. ผลงานวิจัยที่ได้รับการตีพิมพ์	81
ประวัติผู้เขียน	82

สารบัญตาราง

ตารางที่	หน้า
2.1	การจำแนกการเข้ารหัสและวิธีการบีบอัดข้อมูลแบบต่างๆ 6
2.2	ตัวอย่างตารางการควอนไทซ์ 11
3.1	ตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ตแบบออร์โธโกนัลชนิด Daubechies ที่มีความยาวของฟิลเตอร์เท่ากับ 4 6 8 และ 10 ตามลำดับ 30
3.2	เงื่อนไขการแปลงเวฟเล็ตแบบดิสครีต 30
3.3	ผลตอบสนองอิมพัลส์ของ H , G , K และ L 34
3.4	ค่าสัมประสิทธิ์ของการนอร์มอลไลซ์และ $\lambda_j = 1$ เมื่อ $j > 5$ 35
3.5	อัลกอริธึมของการแปลงเวฟเล็ตแบบดิสครีตใน 1 มิติ 35
3.6	อัลกอริธึมของการแปลงเวฟเล็ตแบบดิสครีตใน 2 มิติ 38
5.1	ข้อมูลภาพที่ใช้ในการทดลองที่บิตเรทต่างๆ 54

สารบัญรูป

รูปที่	หน้า
2.1	ประเภทของการบีบอัดข้อมูลภาพ 5
2.2	ขั้นตอนของการเข้ารหัสแบบรันเลนจ์ 7
2.3	การสร้างตารางฮัฟแมน 8
2.4	JPEG ชนิด Sequential Baseline 10
2.5	การเข้ารหัสของค่า DC และลำดับการซิกแซกของการเข้ารหัสเอ็นโทรปี 12
2.6	ขั้นตอนการควอนไทซ์แบบเวกเตอร์ 12
2.7	การแบ่งภาพ LENA เป็นช่วงของบล็อกข้อมูลในรูปสี่เหลี่ยมและสามเหลี่ยม 14
2.8	การบีบอัดและการขยายข้อมูลภาพโดยใช้พื้นฐานการแปลงเวฟเล็ต 15
2.9	ลักษณะของความชันที่ไม่ต่อเนื่องของพิกเซลที่บริเวณขอบบล็อก 17
2.10	ความผิดเพี้ยนของภาพที่ได้จากการบีบอัดข้อมูลภาพ 18
3.1	ผลตอบสนองความถี่เชิงอุดมคติ 22
3.2	ตัวดำเนินการลดข้อมูลและการวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์ 24
3.3	ตัวดำเนินการเพิ่มข้อมูลและการสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์ 24
3.4	โครงสร้างการวิเคราะห์และสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์แบบหลายแบนด์ 26
3.5	การแปลงเวฟเล็ตของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์แบงก์ 28
3.6	ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเล็ต 28
3.7	ตัวอย่างการหาค่าสูงสุดของสัญญาณใน 1 มิติ 33
3.8	การแปลงเวฟเล็ตแบบดิสครีตใน 1 มิติ 34
3.9	การหาขอบของข้อมูลภาพต้นแบบ CIRCLE ขนาด 128x128 พิกเซล ที่ระดับของการแปลงภาพ $1 \leq j \leq 4$ (จากบนลงล่าง) สำหรับพิกเซลสี่ขา สี่เหลี่ยม สี่ด้านแสดงค่าลบ ศูนย์ และค่าบวกตามลำดับของ $W_{2^j}^1 f$ $W_{2^j}^2 f$ $M_{2^j} f$ $A_{2^j} f$ และตำแหน่งของค่ามอดูลัสสูงสุด (พิกเซลสี่ด้าน) 37
3.10	การแปลงเวฟเล็ตแบบดิสครีตใน 2 มิติ 39
3.11	ลักษณะขอบที่ระดับของการแปลงภาพ $1 \leq j \leq 3$ 40
4.1	รูปแบบของการทำเรโซลคิงที่ค่า $\delta = 1$ 47

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.2	อัลกอริทึมของการทำเวฟเล็ตเรซ โชลคั้ง 47
4.3	ภาพ (c) และ (d) แสดงตำแหน่งพิกเซลที่มีค่าสูงสุดและค่าต่ำสุด (พิกเซลสีดำ) ของภาพ (a) และ (b) ตามลำดับ 51
4.4	ความหนาแน่นของค่าสูงสุดและค่าต่ำสุดของภาพตัวอย่าง 4 ภาพ 51
4.5	ตัวอย่างการแยกเนื้อหาของภาพ 51
4.6	อัลกอริทึมของเวฟเล็ตเรซ โชลคั้งที่ปรับตัวเองได้ 52
4.7	ตำแหน่งพิกเซลที่บริเวณขอบบล็อคของการหาค่า EV 53
5.1	ตัวอย่างของภาพ LENA ที่บิตเรทต่างๆ 55
5.2	ตัวอย่างภาพ LENA เมื่อลดขอบบล็อคด้วยการกรองความถี่ต่ำผ่านได้ 55
5.3	ตัวอย่างภาพ LENA เมื่อลดขอบบล็อคด้วยการกรอง Wiener ที่ปรับตัวเองได้ 55
5.4	ค่า PSNR ของภาพ LENA ที่บิตเรทต่างๆ เมื่อใช้เวฟเล็ตต่างชนิดกัน 57
5.5	ค่า EV ของภาพ LENA ที่บิตเรทต่างๆ เมื่อใช้เวฟเล็ตต่างชนิดกัน 58
5.6	ตัวอย่างของภาพ LENA ที่ได้จากการลดขอบบล็อคด้วยเวฟเล็ตเรซ โชลคั้ง 58
5.7	การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 5$ 60
5.8	การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 25$ 61
5.9	การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 10$ 62
5.10	การจำแนกเนื้อหา (พื้นที่สีเทา) และส่วนที่ราบเรียบ (พื้นที่สีขาว) ของภาพ LENA 63
5.11	การจำแนกเนื้อหา (พื้นที่สีเทา) และส่วนที่ราบเรียบ (พื้นที่สีขาว) ของภาพ LENA ที่ 0.31 บิต/พิกเซล เมื่อให้ส่วนที่เป็นเนื้อหาของภาพมีความหนาแน่นของพิกเซลที่มีค่าสูงสุดและค่าต่ำสุดอยู่ในช่วงขอบเขตของค่าบนและค่าล่างเท่ากับ 0.01-0.1 64
5.12	ตัวอย่างผลที่ได้ในจำแนกขอบภาพ (พื้นที่สีดำ) เนื้อหา (พื้นที่สีเทา) และ ส่วนที่ราบเรียบ(พื้นที่สีขาว) ของภาพ LENA ที่ 0.31 บิต/พิกเซล 65
5.13	ตัวอย่างของผลการเปรียบเทียบค่า PSNR เมื่อลดขอบบล็อคด้วยวิธีการต่างๆ 66
5.14	ตัวอย่างผลของการเปรียบเทียบค่า EV เมื่อลดขอบบล็อคด้วยวิธีการต่างๆ 67
5.15	ตัวอย่างของผลการเปรียบเทียบระดับความเข้มของพิกเซล เมื่อลดขอบบล็อคด้วยวิธีการต่างๆ 68

สารบัญรูป (ต่อ)

รูปที่	หน้า
5.16 ตัวอย่างของภาพ LENA ที่ 0.42 บิต/พิกเซล เมื่อลดขอบบล็อคของภาพด้วยวิธีการต่างๆ	69
ก.1 การกรอง Wiener เพื่อหาค่าผิดพลาดกำลังสองเฉลี่ยที่น้อยที่สุดของ $f(n_1, n_2)$ จาก $g(n_1, n_2) = f(n_1, n_2) + v(n_1, n_2)$	76
ก.2 ระบบการสร้างภาพกลับคืนแบบปรับตัวเองได้	76
ก.3 ผลตอบสนองอิมพัลส์ของฟิลเตอร์ที่เปลี่ยนแปลงตามพื้นที่	80

บทที่ 1

บทนำ

1.1 ที่มาของปัญหา

การบีบอัดข้อมูลภาพ (Image compression) เป็นปัญหาที่สำคัญมากสำหรับการใช้งานสื่อสารด้านภาพ ความมุ่งหมายของการบีบอัดข้อมูลภาพเพื่อต้องการลดพื้นที่ในการจัดเก็บข้อมูลภาพ และค่าใช้จ่ายในการสื่อสาร โดยยังคงรักษาคุณภาพของภาพที่ได้ไว้ เทคนิคการบีบอัดข้อมูลภาพที่มีประสิทธิภาพมีหลายเทคนิค แต่ที่ถูกนำมาใช้งานอย่างกว้างขวางคือเทคนิคการเข้ารหัสการแปลงแบบบล็อก เทคนิคอย่างหนึ่งที่นิยมโดยทั่วไปได้แก่ การแปลงดิสครีตโคไซน์แบบบล็อกหรือ BDCT (Block Discrete Cosine Transform) นอกจากนี้ BDCT ได้ถูกพัฒนาขึ้นเป็นเทคนิคที่สำคัญเพื่อการบีบอัดข้อมูลภาพดิจิทัล เช่น มาตรฐานการบีบอัดข้อมูลภาพนิ่ง JPEG มาตรฐานการบีบอัดข้อมูลภาพวิดีโอ MPEG และมาตรฐาน H.261 ของโทรศัพท์ภาพ (Videophone) หรือการประชุมระยะไกล (Teleconference) เป็นต้น อย่างไรก็ตามการบีบอัดข้อมูลภาพแบบ BDCT จะเห็นความผิดเพี้ยนของภาพที่อัตราการบีบอัดข้อมูลภาพที่สูงและที่เห็นได้อย่างชัดเจนคือลักษณะของขอบบล็อกข้อมูลภาพ (Blocking artifact) ซึ่งเป็นผลเนื่องมาจากค่าความผิดพลาดของการควอนไทซ์ในขั้นตอนการเข้ารหัสแบบบล็อก ผลของขอบบล็อกข้อมูลภาพสามารถสังเกตเห็นได้ในลักษณะที่เป็นคาบทั้งแนวนอนและแนวตั้งของภาพและเห็นได้อย่างชัดเจนในบริเวณพื้นที่ของภาพที่มีความราบเรียบ (Smooth region)

ได้มีการพัฒนาเทคนิคต่างๆ มากมายเพื่อทำการลดขอบบล็อกข้อมูลภาพหรือที่เรียกว่า deblocking เทคนิคเหล่านี้ได้มีการนำเสนอในบทความต่างๆ เสมือนเป็นกลไกอย่างหนึ่งในขั้นตอนของการประมวลผลหลังการบีบอัดข้อมูลภาพ (Post-processing) ซึ่งเป็นวิธีที่คืออย่างหนึ่งที่จะสามารถนำไปสู่การพัฒนาในการสื่อสารด้านภาพและวิดีโอเนื่องจากเทคนิคนี้ไม่ขึ้นกับรูปแบบของการเข้ารหัสและถอดรหัสข้อมูลภาพ (Codec coding) รวมทั้งให้คุณภาพของภาพที่ดีกว่า ขณะเดียวกันยังคงให้บิดเบือนและรักษารูปแบบการเข้ารหัสและถอดรหัสข้อมูลไว้เหมือนเดิมทุกประการ เทคนิคของการประมวลผลหลังการบีบอัดข้อมูลภาพขึ้นอยู่กับมุมมองของการแก้ปัญหาการลดขอบบล็อกข้อมูลภาพ ถ้ามองปัญหาการลดขอบบล็อกข้อมูลภาพเช่นเดียวกับปัญหาของการสร้างภาพขึ้นมาใหม่จากข้อมูลภาพเดิมที่มีอยู่ (Restoration problem) ซึ่งวิธีการสร้างภาพขึ้นมาใหม่มีหลายวิธีด้วยกัน เช่น วิธีการ POCS (Projection onto convex set) [1] วิธีการกำหนดค่ากำลังสองที่น้อยที่สุดหรือ CLS (Constrained least square) [2] เป็นต้น วิธีการเหล่านี้เป็นวิธีการที่มีการทำซ้ำและใช้เวลาในการคำนวณที่มาก หรืออาจมองปัญหาการลดขอบบล็อกข้อมูลภาพเช่นเดียวกับ

ปัญหาของการกรองข้อมูลภาพ (Filtering problem) ซึ่งวิธีการอย่างง่ายที่สุดคือการกรองความถี่ต่ำผ่านได้ [3] เป็นต้น

ในวิทยานิพนธ์นี้ได้มองปัญหาการลดขอบบล็อคข้อมูลภาพเช่นเดียวกับปัญหาของการกำจัดสัญญาณรบกวน (Denoising problem) วิธีการอย่างหนึ่งที่ได้รับความสนใจอย่างมากคือเวฟเล็ตเทรโซลด์ดิ้ง (Wavelet thresholding) เป็นวิธีการกำจัดสัญญาณรบกวนโดยการตัดค่าสัมประสิทธิ์เวฟเล็ตทิ้งไปด้วยเทรโซลด์ค่าหนึ่ง อย่างไรก็ตามในการประยุกต์ใช้งานด้านการประมวลผลภาพสัญญาณรบกวนในข้อมูลภาพจะเห็นได้แตกต่างกันตามลักษณะพื้นที่ของข้อมูลภาพ กล่าวคือพื้นที่ของข้อมูลภาพที่มีความราบเรียบ (Smooth region) จะสังเกตเห็นสัญญาณรบกวนได้มากกว่าพื้นที่ส่วนที่เป็นขอบภาพ (Edge region) หรือพื้นที่ส่วนที่เป็นเนื้อหาของภาพ (Texture region) ดังนั้นในวิทยานิพนธ์นี้จะนำเสนอวิธีการทำเวฟเล็ตเทรโซลด์ดิ้งที่ปรับตัวเองได้ (Adaptive wavelet thresholding) เป็นวิธีการทำเทรโซลด์ด้วยค่าเทรโซลด์ที่เหมาะสมตามลักษณะพื้นที่ของข้อมูลภาพในโดเมนเวฟเล็ต ซึ่งเป็นวิธีการที่ง่าย การคำนวณไม่ยุ่งยากซับซ้อน โดยสามารถลดขอบบล็อคข้อมูลภาพและยังคงรักษาลักษณะที่สำคัญของภาพไว้ได้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อนำเสนอวิธีการลดขอบบล็อคข้อมูลภาพที่ได้จากการบีบอัดข้อมูลภาพ JPEG ด้วยเทคนิคของการทำเวฟเล็ตเทรโซลด์ดิ้งที่ปรับตัวเองได้ จากนั้นทำการศึกษาเปรียบเทียบผลที่ได้จากการลดขอบบล็อคด้วยวิธีการอื่นๆ ได้แก่ การกรองความถี่ต่ำผ่านได้ การกรอง Wiener ที่ปรับตัวเองได้และการทำเวฟเล็ตเทรโซลด์ดิ้ง โดยทำการประเมินผลภาพที่ได้จากการวัดค่าอัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) ค่าความแปรปรวนที่บริเวณขอบบล็อค (EV) และแสดงผลภาพที่ได้ นอกจากนี้วิธีการที่นำเสนอสามารถนำไปประยุกต์ใช้กับการลดความผิดเพี้ยนของภาพที่ได้จากการเข้ารหัสข้อมูลด้วยเทคนิคต่างๆ หรือใช้กำจัดสัญญาณรบกวนชนิดอื่นได้

1.3 ขอบเขตของการวิจัย

ในส่วนของการวิจัยได้เสนอวิธีการทำเวฟเล็ตเทรโซลด์ดิ้งที่ปรับตัวเองได้เพื่อลดขอบบล็อคข้อมูลภาพที่ได้จากการบีบอัดข้อมูลภาพ JPEG และทำการศึกษาเปรียบเทียบกับวิธีการอื่นๆ

วิทยานิพนธ์นี้ได้แบ่งออกเป็น 6 บทในแต่ละบทมีหัวข้อและเนื้อหา ดังนี้

บทที่ 1 บทนำ

กล่าวถึงที่มาของปัญหางานวิจัย วัตถุประสงค์ของการทำวิจัยและขอบเขตเนื้อหาของการทำวิทยานิพนธ์

บทที่ 2 หลักการพื้นฐานของการบีบอัดข้อมูลภาพ

กล่าวถึงวัตถุประสงค์ของการบีบอัดข้อมูลภาพ การจำแนกประเภทของการบีบอัดข้อมูลภาพ เทคนิคการบีบอัดข้อมูลภาพที่นิยมใช้ในปัจจุบัน ความผิดพลาดที่เกิดจากการบีบอัดข้อมูลภาพ รวมทั้งเกณฑ์การวัดประสิทธิภาพของภาพที่ได้จากการบีบอัดข้อมูลภาพ

บทที่ 3 การแปลงเวฟเล็ท

กล่าวถึงทฤษฎีพื้นฐานของการแปลงเวฟเล็ทแบบดิสครีต โดยใช้หลักการวิเคราะห์ฟิลเตอร์แบงก์ (Filter bank) คุณสมบัติของเวฟเล็ทแบบอโธโกนัลและแบบไบอโธโกนัลที่นิยมใช้ทั่วไปในงานด้านการประมวลผลภาพ ความสัมพันธ์ของการหาขอบภาพที่หลายระดับของการแปลงเวฟเล็ท โดยกล่าวถึงคุณสมบัติทั่วไปในการหาอนุพันธ์ของการแปลงเวฟเล็ทใน 1 มิติและ 2 มิติ รวมทั้งการนำเวฟเล็ทไปประยุกต์ใช้ในการหาขอบภาพ

บทที่ 4 การลดขอบบล็อคในข้อมูลภาพ

กล่าวถึงการจำแนกประเภทวิธีการลดขอบบล็อค วิธีการลดขอบบล็อคด้วยการกรองความถี่ต่างๆ ได้แก่ การกรองความถี่ต่ำผ่าน การกรอง Wiener ที่ปรับตัวเองได้ เป็นต้น การลดขอบบล็อคด้วยเวฟเล็ทเชรโซลคิงและเวฟเล็ทเชรโซลคิงที่ปรับตัวเองได้ เกณฑ์ที่ใช้ในการวัดประสิทธิภาพของการลดขอบบล็อคคือค่าอัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) และค่าความแปรปรวนที่บริเวณขอบบล็อค (EV)

บทที่ 5 ผลการทดลอง

กล่าวถึงการเตรียมข้อมูลภาพที่ใช้ในการทดลอง การวิเคราะห์ผลที่ได้จากการลดขอบบล็อคด้วยวิธีการต่างๆ ได้แก่ การกรองความถี่ต่ำผ่าน การกรอง Wiener ที่ปรับตัวเองได้ เวฟเล็ทเชรโซลคิงและเวฟเล็ทเชรโซลคิงที่ปรับตัวเองได้ สำหรับวิธีการที่นำเสนอจะแสดงผลที่ได้ในแต่ละขั้นตอนและทำการเปรียบเทียบผลที่ได้กับวิธีการอื่นๆ

บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ

สรุปผลการเปรียบเทียบที่ได้จากการทดลอง พร้อมทั้งวิจารณ์ถึงข้อดีข้อเสียพร้อมทั้งข้อเสนอแนะแนวทางการวิจัยที่สามารถพัฒนาต่อไปได้

และสุดท้ายในภาคผนวกจะเป็นรายละเอียดของการกรอง Wiener ที่ปรับตัวเองได้ เพื่อเป็นรายละเอียดสำหรับผู้ที่สนใจและต้องการเข้าใจในเนื้อหาได้มากขึ้นเพื่อสามารถนำไปใช้ในงานวิจัยต่อไปได้

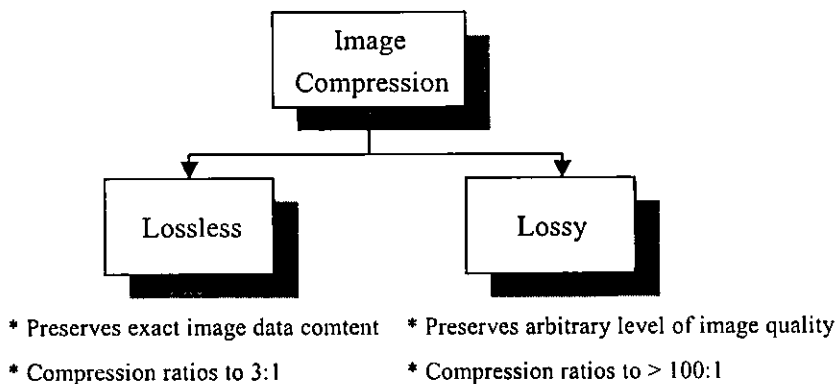
บทที่ 2

หลักการพื้นฐานของการบีบอัดข้อมูลภาพ

ข้อมูลต่างๆ เช่น ภาพนิ่ง เสียง และวิดีโอ จำเป็นต้องใช้พื้นที่ของการเก็บข้อมูลที่มาก ในกรณีของภาพวิดีโอต้องอาศัยเทคโนโลยีของ CD มาใช้ในการเก็บข้อมูลที่มีจำนวนมาก นอกจากนี้ ข้อมูลขนาดใหญ่เป็นอุปสรรคที่สำคัญในการใช้งานด้านการสื่อสารที่มีความจุของช่องสัญญาณที่จำกัดและค่าใช้จ่ายที่สูง ดังนั้นข้อมูลที่มีจำนวนมากจะทำให้อัตราการส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งใช้เวลานานขึ้น ดังนั้นระบบต่างๆ จึงให้ความสำคัญกับการลดขนาดข้อมูลหรือที่เรียกว่า การบีบอัดข้อมูล ซึ่งมีได้หลายวิธีที่แตกต่างกันไป ขึ้นอยู่กับการนำไปประยุกต์ใช้งานด้านต่างๆ ในเนื้อหาบทนี้จะกล่าวถึง การจำแนกประเภทของการบีบอัดข้อมูลภาพ เทคนิคการบีบอัดข้อมูลภาพต่างๆ ที่นิยมใช้ทั่วไป ความผิดพลาดที่เกิดจากการบีบอัดข้อมูลภาพและการวัดประสิทธิภาพการบีบอัดข้อมูลภาพ ตามลำดับ

2.1 การจำแนกประเภทของการบีบอัดข้อมูลภาพ

โดยทั่วไปรูปแบบการบีบอัดข้อมูลภาพ สามารถแบ่งเป็น 2 ประเภท ดังนี้



รูปที่ 2.1 ประเภทของการบีบอัดข้อมูลภาพ

(1) การบีบอัดข้อมูลแบบไม่มีการสูญเสีย (Lossless compression) เป็นวิธีการที่ข้อมูลภาพเมื่อผ่านขั้นตอนการบีบอัดแล้วผลลัพธ์ของข้อมูลภาพที่ได้จะเหมือนภาพต้นแบบ (Original image) ทุกประการ นิยมใช้กับข้อมูลที่มีความสำคัญมากและสูญเสียไม่ได้ เช่น ข้อมูลภาพที่ใช้วินิจฉัยทางการแพทย์ เป็นต้น

(2) การบีบอัดข้อมูลแบบมีการสูญเสีย (Lossy compression) เป็นวิธีการที่ผลลัพธ์ของข้อมูลภาพที่ได้จากการขยายข้อมูลภาพกลับมีความแตกต่างจากภาพต้นแบบ โดยยอมให้มีการสูญเสียข้อมูลได้บ้าง นิยมใช้ในการประมวลผลภาพทั่วไปเป็นอย่างมาก เนื่องจากตาของคนเราไม่สามารถตรวจจับข้อมูลในส่วนที่มีการสูญเสียได้ อย่างไรก็ตามคุณภาพของภาพที่ได้มีความสัมพันธ์กับอัตราการบีบอัดข้อมูลภาพด้วย (Compression Ratio) นั่นคือ ข้อมูลภาพมีอัตราการบีบอัดที่สูง จะได้ขนาดของข้อมูลภาพที่มีขนาดเล็ก แต่ทำให้คุณภาพของภาพที่ได้แย่ลงไปด้วย

วิธีการบีบอัดข้อมูลภาพมีหลายวิธี แต่ละวิธีมีความเหมาะสมในลักษณะของการนำไปใช้งานที่ต่างกัน สามารถจำแนกความแตกต่างของวิธีการบีบอัดข้อมูลภาพได้ดังตารางที่ 2.1

ตารางที่ 2.1 การจำแนกการเข้ารหัสและวิธีการบีบอัดข้อมูลแบบต่างๆ

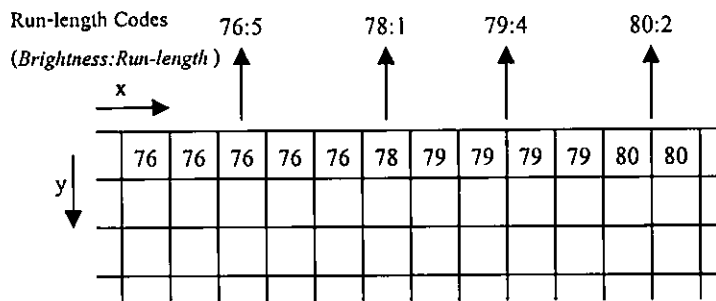
Entropy Encoding	Run-length Coding	
	Huffman Coding	
	Arithmetic Coding	
Source Coding	Prediction	DPCM
		DM
	Transformation	FFT
		DCT
	Layered Coding	Bit Position
		Subsampling
		Sub-band Coding
Vector Quantization		
Hybrid Coding	JPEG	
	MPEG	
	H.261	
	DVI RTV, DVI PLV	

2.1.1 การเข้ารหัสเอนโทรปี (Entropy encoding)

การเข้ารหัสแบบนี้จะไม่คำนึงถึงคุณสมบัติเฉพาะของข้อมูลที่นำมาผ่านกระบวนการ โดยสายข้อมูลจะถูกบีบอัดข้อมูลโดยการใช้ลำดับของข้อมูลทางดิจิทัล การเข้ารหัสแบบนี้เป็นการเข้ารหัสแบบไม่มีการสูญเสีย ได้แก่

(1) การเข้ารหัสแบบรันเลนจ์ (Run-length coding) เป็นวิธีการที่ง่ายที่สุดในการเข้ารหัสข้อมูล โดยข้อมูลภาพที่มีพิกเซลที่อยู่ติดกัน มีการเปลี่ยนแปลงของระดับความเข้ม ไม่มากและมี

จำนวนที่ซ้ำกัน จะทำการเก็บจำนวนของพิกเซลที่ซ้ำกันแทนการเก็บค่าระดับความเข้มของพิกเซลเหล่านั้นดังรูปที่ 2.2

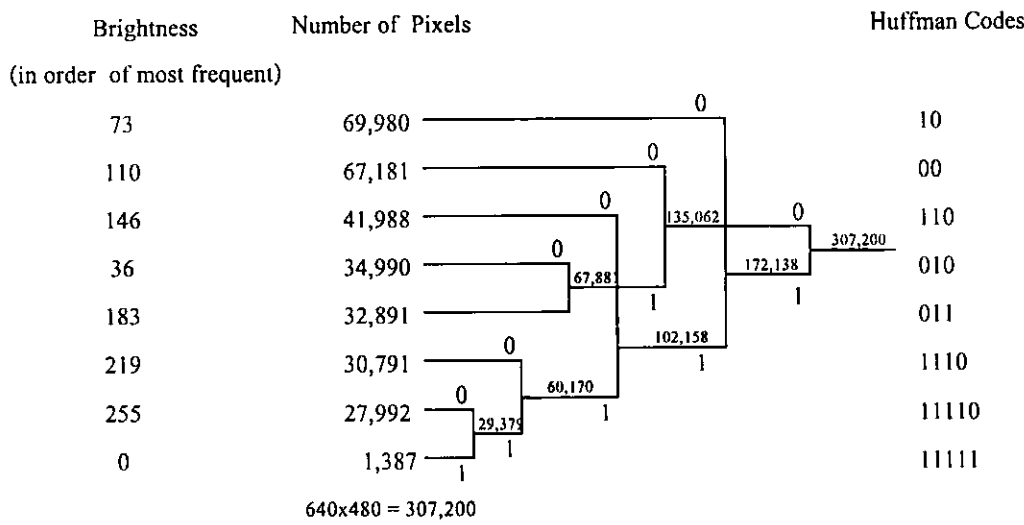


รูปที่ 2.2 ขั้นตอนของการเข้ารหัสแบบรันเลนจ์

(2) การเข้ารหัสฮัฟแมน (Huffman coding) หลักการสำคัญของการเข้ารหัสฮัฟแมนคือข้อมูลที่มีความน่าจะเป็นสูงกว่าหรือข้อมูลที่มีความถี่สูงกว่าจะถูกเข้ารหัสด้วยจำนวนบิตที่น้อยกว่า การเข้ารหัสฮัฟแมนสามารถอธิบายอย่างย่อดังนี้ [5]

1. การนับความถี่ของข้อมูลทุกตัวและเรียงลำดับข้อมูลตามความถี่จากมากไปหาน้อย
2. นำค่าความถี่ที่ต่ำสุด 2 ค่ามารวมกันเพื่อสร้างโหนดใหม่ นำค่าความถี่ใหม่ที่ได้อามาเรียงลำดับร่วมกับความถี่เดิมที่เหลือ
3. กำหนดค่าไบนารีให้กับโหนดทุกคู่ที่ถูกกำหนดให้เป็นเส้นทางของโหนดใหม่ โดยที่โหนดหนึ่งมีค่าเป็น 0 ส่วนอีกโหนดหนึ่งมีค่าเป็น 1
4. ทำซ้ำตามขั้นตอนที่ (2) และ (3) จนกระทั่งได้ความถี่เท่ากับจำนวนพิกเซลทั้งหมดของภาพ ดังรูปที่ 2.3
5. การเข้ารหัสและถอดรหัสสามารถกระทำได้โดยใช้การเปิดตาราง (Look-up table) และรหัสที่ได้สามารถทำการถอดรหัสได้ทันทีโดยไม่ต้องอ้างอิงถึงข้อมูลอื่น

(3) การเข้ารหัสเลขคณิต (Arithmetic coding) การเข้ารหัสเลขคณิตนั้น ข้อมูลจะถูกแทนที่ด้วยช่วงระยะห่างของจำนวนจริงระหว่าง 0 และ 1 โดยข้อมูลแต่ละตัวจะลดขนาดของระยะห่างตามความน่าจะเป็นของข้อมูลนั้น ข้อมูลที่มีความน่าจะเป็นสูงจะทำให้ระยะห่างถูกลดลงได้น้อยกว่าข้อมูลที่มีความน่าจะเป็นต่ำ ดังนั้นข้อมูลที่มีความน่าจะเป็นสูงจะใช้จำนวนบิตในการเข้ารหัสน้อยกว่า



รูปที่ 2.3 การสร้างตารางฮัฟแมน

2.1.2 การเข้ารหัสซอร์ส (Source coding)

พิจารณาความสัมพันธ์ของข้อมูลหรือองค์ประกอบของข้อมูลเริ่มต้นเป็นหลัก โดยส่วนใหญ่เป็นการบีบอัดข้อมูลที่มีการสูญเสีย โดยการเข้ารหัสแบบนี้จะให้อัตราการบีบอัดข้อมูลที่สูงกว่าแบบการเข้ารหัสแบบเอนโทรปี ได้แก่

- (1) การเข้ารหัสทำนาย (Prediction coding) เป็นการใช้นิเทศการทำนาย ทำนายข้อมูลที่เกิดขึ้นในปัจจุบันโดยอาศัยข้อมูลจากอดีตที่ผ่านมา แล้วนำมาเปรียบเทียบกับค่าของข้อมูลเริ่มต้นเพื่อหาค่าผลต่าง เช่น การมอดูเลตรหัสพัลส์เชิงอนุพันธ์ หรือ DPCM (Differential pulse code modulation) และการมอดูเลตเดลด้าหรือ DM (Data modulation)
- (2) การเข้ารหัสการแปลง (Transform coding) เป็นการใช้นิเทศของการแปลงรูปจากโดเมนหนึ่งไปยังอีกโดเมนหนึ่ง เช่น การแปลงดิสครีตโคไซน์ (DCT) การแปลงฟูเรียร์ (FFT) เป็นต้นเป็นการแปลงข้อมูลจากโดเมนสเปเชียลไปยังโดเมนความถี่
- (3) การเข้ารหัสเลเยอร์ (Layer coding)
- (4) การเข้ารหัสด้วยการควอนไทซ์แบบเวกเตอร์ (Vector quantization)

2.1.3 การเข้ารหัสแบบไฮบริด (Hybrid coding)

เป็นการนำหลักการที่ได้จากวิธีที่กล่าวในข้างต้นมาผสมเข้าด้วยกัน เช่น การบีบอัดข้อมูลภาพหนึ่ง JPEG (Joint Photographic Expert Group) ที่อาศัยหลักพื้นฐานของการแปลงแบบดิสครีตโคไซน์ร่วมกับการมอดูเลตรหัสพัลส์เชิงอนุพันธ์ เป็นต้น

2.2 เทคนิคการบีบอัดข้อมูลภาพแบบต่างๆ

2.2.1 มาตรฐานการบีบอัดข้อมูลภาพนิ่ง JPEG

JPEG เป็นมาตรฐานการบีบอัดและขยายข้อมูลภาพนิ่งที่กำหนดขึ้นโดย Joint Photographic Expert Group ซึ่งเป็นความร่วมมือกันระหว่าง ISO (International Organization for Standardization) และ ITU (International Telecommunication Union) โดยมีวัตถุประสงค์ที่จะพัฒนามาตรฐานการบีบอัดข้อมูลภาพดิจิทัลดังนี้ [6]

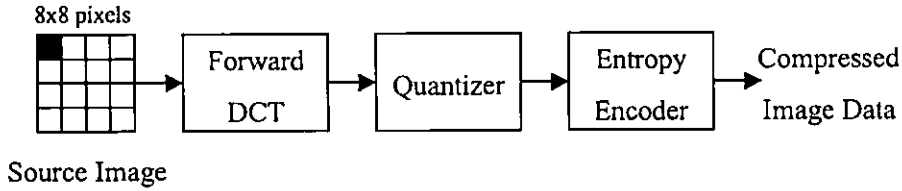
- (1) ผู้ใช้สามารถปรับเปลี่ยนอัตราส่วนของการบีบอัดข้อมูลภาพได้
- (2) สามารถนำไปใช้งานกับภาพดิจิทัลต่างๆ ได้
- (3) มีขั้นตอนการคำนวณที่ยุ่งยากซับซ้อนเพื่อนำไปใช้งานต่างๆ ได้อย่างกว้างขวาง
- (4) JPEG ประกอบด้วยการทำงาน 4 โหมดดังนี้
 - 4.1 การเข้ารหัสแบบ Sequential
 - 4.2 การเข้ารหัสแบบ Progressive
 - 4.3 การเข้ารหัสแบบ Lossless
 - 4.4 การเข้ารหัสแบบ Hierarchical

JPEG มีการบีบอัดข้อมูลภาพแบบมีการสูญเสีย 3 โหมดและแบบไม่มีการสูญเสีย 1 โหมด ในหัวข้อนี้จะเน้นการบีบอัดข้อมูลภาพ JPEG ชนิด Sequential Baseline JPEG จะให้อัตราการบีบอัดข้อมูลภาพที่ดี เมื่อทำการบีบอัดข้อมูลภาพด้วยอัตราการบีบอัดที่ 20: 1 ผลภาพที่ได้ไม่สามารถสังเกตเห็นความแตกต่างได้เมื่อเปรียบเทียบกับภาพต้นแบบ การบีบอัดข้อมูลภาพ JPEG สามารถกำหนดระดับคุณภาพของภาพด้วยค่า Q (Quality factor) โดยค่า Q ที่มีค่าสูงจะให้คุณภาพของภาพที่ได้สูงแต่จะมีขนาดของไฟล์ที่ใหญ่กว่าเมื่อเปรียบเทียบกับค่า Q ที่มีค่าต่ำกว่า ซึ่งจะให้ขนาดของไฟล์เล็กกว่าแต่คุณภาพของภาพที่ได้แย่กว่า และเป็นความสัมพันธ์พื้นฐานของคุณภาพและอัตราการบีบอัดข้อมูลในการบีบอัดข้อมูลแบบมีการสูญเสีย JPEG มีขั้นตอนการบีบอัดข้อมูลแบบสมมาตร (Symmetrical algorithm) เนื่องจากมีขั้นตอนการบีบอัดและขั้นตอนการขยายข้อมูลเท่ากัน ดังรูปที่ 2.4 ดังนั้นเวลาที่ใช้ย่อมเท่ากันด้วย

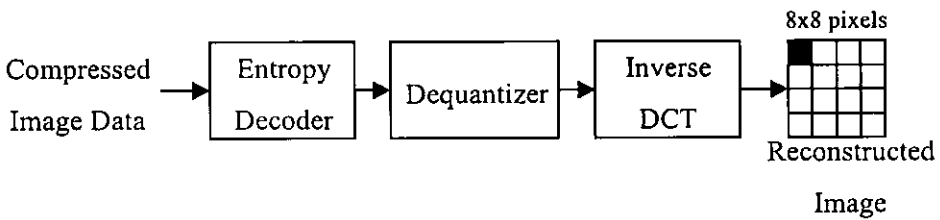
การบีบอัดข้อมูลภาพ JPEG ชนิด Sequential Baseline ประกอบด้วยขั้นตอนการทำงานพื้นฐาน 5 ขั้นตอนดังนี้

- (1) การแปลงระบบสี RGB ไปเป็นระบบสี YCbCr ยกเว้นในกรณีของภาพหลายระดับเทา เพื่อแยกเป็นองค์ประกอบของแสง (Luminance) และองค์ประกอบของสี (Chrominance) เนื่องจากการมองเห็นของคนเรามีข้อจำกัดในการรับรู้การเปลี่ยนแปลงองค์ประกอบของแสงได้ดีกว่าองค์

ประกอบของสี ทำให้ JPEG พยายามลดข้อมูลในส่วนที่เป็นองค์ประกอบของสีมากกว่าองค์ประกอบของแสง



(a)



(b)

รูปที่ 2.4 JPEG ชนิด Sequential Baseline (a) การเข้ารหัส (b) การถอดรหัส

(2) การลดข้อมูลในส่วนที่เป็นองค์ประกอบของสี ขณะที่องค์ประกอบของแสงมีความละเอียดแบบเต็ม (Full resolution) การลดข้อมูล (Subsampling) ในส่วนที่เป็นองค์ประกอบของสี ลง 2 พิกเซลในทิศทางแนวนอนและแนวตั้งเป็นรูปแบบที่นิยมมากอย่างหนึ่ง โดยการหาค่าเฉลี่ยของบล็อกข้อมูลขนาด 2×2 พิกเซล ขั้นตอนนี้จะมีการสูญเสียข้อมูลภาพและลดข้อมูลลงไปครึ่งหนึ่งของข้อมูลเดิม

(3) การแบ่งข้อมูลออกเป็นบล็อกย่อยๆ ขนาด 8×8 พิกเซล และทำการแปลงข้อมูลให้อยู่ในโดเมนของความถี่ด้วยการแปลง Forward DCT ดังสมการที่ (2.1) ข้อมูลตัวแรกของบล็อกเรียกว่าค่า DC เป็นค่าเฉลี่ยของข้อมูลภายในบล็อกนั้นๆ และข้อมูลที่เหลือจำนวน 63 ค่าเรียกว่าค่า AC

สมการ Forward DCT

$$C(u, v) = \frac{1}{\sqrt{2N}} \alpha(u) \alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{(2x+1)v\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (2.1)$$

สมการ Inverse DCT

$$f(x, y) = \frac{1}{\sqrt{2N}} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v)\text{Cos}\left[\frac{(2x+1)v\pi}{2N}\right]\text{Cos}\left[\frac{(2y+1)v\pi}{2N}\right] \quad (2.2)$$

เมื่อ

$C(u, v)$ คือค่าสัมประสิทธิ์ DCT

$f(x, y)$ คือค่าระดับสีในพิกเซล

$\alpha(i) = 1/\sqrt{2}$ เมื่อ $i = 0$ และ $\alpha(i) = 1$ เมื่อ $i > 0$

(4) การควอนไทซ์ค่าสัมประสิทธิ์ DCT โดยทำการควอนไทซ์บล็อกย่อยๆ ด้วยค่าสัมประสิทธิ์ของการควอนไทซ์และทำการปิดเศษค่าให้เป็นเลขจำนวนเต็ม ทำให้ข้อมูลมีค่าเป็นศูนย์มากเพื่อสะดวกในการเข้ารหัสต่อไป ใน JPEG ไม่มีการกำหนดตารางการควอนไทซ์ที่แน่นอน ตัวอย่างของตารางการควอนไทซ์ขององค์ประกอบของแสงและองค์ประกอบของสีแสดงดังตารางที่ 2.2 (a) และ (b) ตามลำดับ

(5) การเข้ารหัสเอนโทรปี ในที่นี้ใช้การเข้ารหัสแบบฮัฟแมน โดยค่า DC ใช้การเข้ารหัสของผลต่างระหว่างค่า DC ของบล็อกก่อนหน้านี้และบล็อกปัจจุบันเนื่องจากค่า DC ของบล็อกที่ใกล้เคียงกันจะมีความสัมพันธ์กันอย่างมาก สำหรับค่า AC จะเรียงลำดับแบบซิกแซก (Zigzag sequence) เพื่อให้ค่าสัมประสิทธิ์ที่มีความถี่ต่ำที่มีค่าส่วนใหญ่ไม่เป็นศูนย์อยู่ติดกันและค่าสัมประสิทธิ์ความถี่สูงที่มีค่าส่วนใหญ่เป็นศูนย์อยู่ติดกันเพื่อสะดวกในการเข้ารหัสแบบรันเลนจ์

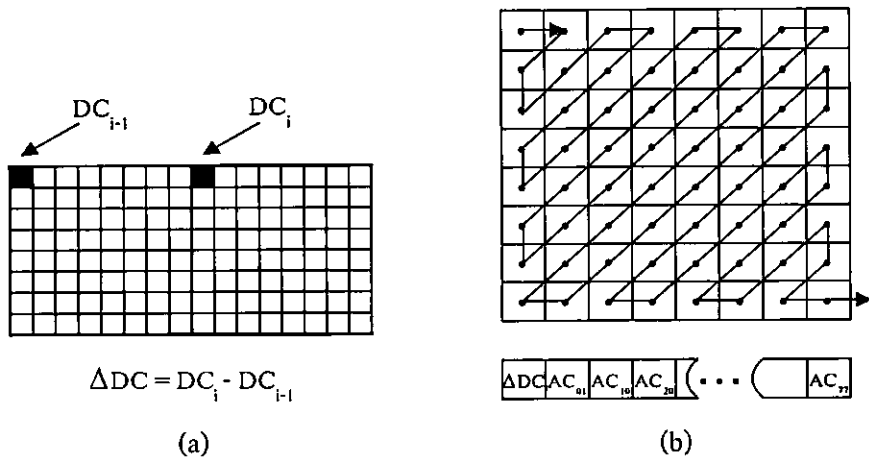
ตารางที่ 2.2 ตัวอย่างตารางการควอนไทซ์ (a) องค์ประกอบของแสง (b) องค์ประกอบของสี

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(a)

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

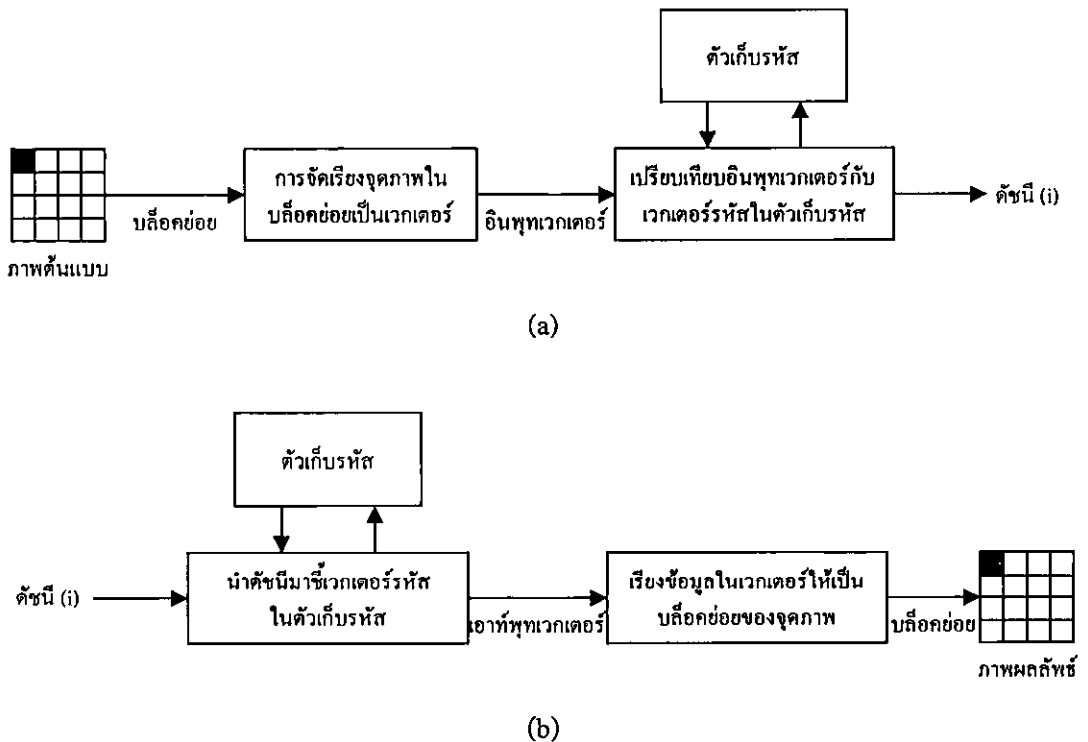
(b)



รูปที่ 2.5 (a) การเข้ารหัสของค่า DC (b) ลำดับการซิกแซกของการเข้ารหัสเอ็นโทรปี

2.2.2 การบีบอัดข้อมูลภาพด้วยการเข้ารหัสการควอนไทซ์แบบเวกเตอร์ [7]

การควอนไทซ์แบบเวกเตอร์มีลักษณะเช่นเดียวกับ JPEG คือการแบ่งข้อมูลออกเป็นบล็อกย่อยๆ ขนาด $n \times n$ พิกเซล และอาศัยหลักการแทนที่บล็อกย่อยของข้อมูลด้วยเลขดัชนีของชุดข้อมูลที่ใช้แทนบล็อกย่อยเหล่านั้น มักจะใช้เลขดัชนีเป็นเลขจำนวนนับและมีค่าจำกัดอยู่ระดับหนึ่ง



รูปที่ 2.6 ขั้นตอนการควอนไทซ์แบบเวกเตอร์ (a) การเข้ารหัส (b) การถอดรหัส

ขั้นตอนในการเข้ารหัสการควอนไทซ์แบบเวกเตอร์นั้น ข้อมูลภาพจะถูกแยกองค์ประกอบให้เป็นเวกเตอร์ขนาด K มิติ เวกเตอร์เหล่านี้สามารถสร้างได้หลายวิธีเช่น การแบ่งข้อมูลภาพออกเป็นบล็อกย่อยๆ ขนาด $n \times n = K$ พิกเซล และทำการเรียงให้อยู่ในรูปเวกเตอร์ขนาด K มิติได้เป็นต้น จำนวนบล็อกย่อยที่แบ่งได้พอดีจะเท่ากับจำนวนอินพุทเวกเตอร์ ในการเข้ารหัสจะเปรียบเทียบอินพุทเวกเตอร์กับเวกเตอร์รหัสทุกตัวที่เก็บอยู่ในตัวเก็บรหัส โดยเปรียบเทียบหาเวกเตอร์รหัสที่แทนด้วยเวกเตอร์ที่มีความเพี้ยนจากอินพุทเวกเตอร์น้อยที่สุด ผลลัพธ์ที่ได้ก็คือเลขดัชนีที่ใช้แทนเวกเตอร์รหัสที่มีความเพี้ยนน้อยที่สุด ดังรูปที่ 2.6(a) ในขั้นตอนการถอดรหัสจะเป็นการนำเอาค่าดัชนีมาชี้ข้อมูลเวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัสเดียวกันกับขั้นตอนการเข้ารหัส ผลลัพธ์ที่ได้จะเป็นเอาท์พุทเวกเตอร์ จากนั้นนำเอาท์พุทเวกเตอร์ที่ได้มาจัดเรียงใหม่ให้เป็นบล็อกย่อยของภาพ ดังรูปที่ 2.6(b)

2.2.3 การบีบอัดข้อมูลภาพแฟรคทอล (Fractal)

การบีบอัดข้อมูลภาพแฟรคทอลจะใช้พื้นฐานของการแบ่งภาพออกเป็นส่วนย่อยๆ (Fractal) และกำหนดให้แต่ละส่วนย่อยที่มีลักษณะเหมือนกันหรือคล้ายกันแสดงในแต่ละพื้นที่ บล็อกที่มีขนาดเล็กกว่าจะทำการหาส่วนที่เหมือนกันภายในภาพได้ง่ายกว่า ภาพจะถูกแบ่งออกเป็นโดเมนและช่วงของบล็อกที่แตกต่างกัน แต่ละบล็อกของโดเมนหลังจากทำการแปลงข้อมูลจะมีช่วงของบล็อกที่ตรงกันหรือใกล้เคียงมากที่สุด

การบีบอัดข้อมูลภาพแฟรคทอลเกี่ยวข้องกับการแปลงข้อมูลภาพซึ่งเป็นฟังก์ชันทางคณิตศาสตร์ประกอบด้วย การหมุน (Rotation) การปรับขนาด (Scaling) การลาดเอียง (Skew) และการเลื่อน (Translation) ตัวอย่างของการแปลงข้อมูลภาพ 2 มิติ สามารถหาได้จากสมการต่อไปนี้

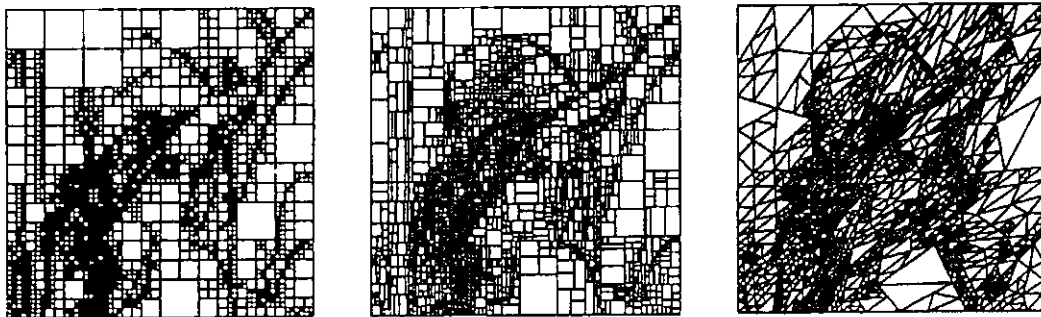
$$W(x, y) = (ax + by + e, cx + dy + f) \quad (2.3)$$

หรืออยู่ในรูปของเมตริกซ์ได้ดังนี้

$$W \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} = \begin{bmatrix} ax + by + e \\ cx + dy + f \end{bmatrix} \quad (2.4)$$

เมื่อเมตริกซ์ $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ แสดงการหมุน การลาดเอียงและการเปลี่ยนขนาด ส่วนเมตริกซ์ $\begin{bmatrix} e \\ f \end{bmatrix}$ แสดงการเลื่อน

การแปลงข้อมูลจะย้ายจากจุด $(0, 0)$ $(1, 0)$ $(0, 1)$ $(1, 1)$ ไปยัง (e, f) $(a+e, c+f)$ $(b+e, d+f)$ $(a+e, c+d+f)$ ตามลำดับ เมื่อค่า a b c d และ f เป็นค่าสัมประสิทธิ์ของการแปลง



รูปที่ 2.7 การแบ่งภาพ LENA เป็นช่วงของบล็อกข้อมูลในรูปสี่เหลี่ยมและสามเหลี่ยม

การบีบอัดข้อมูลภาพแฟรคทอลประกอบด้วย การหาชุดของการแปลงข้อมูลไปทางขวา สำหรับอัลกอริทึมการบีบอัดข้อมูลภาพแฟรคทอลมีขั้นตอนการทำงานดังนี้

(1) ในตอนแรกภาพจะแบ่งออกเป็นส่วนๆ โดยไม่ซ้อนทับกัน กลุ่มโดเมนของพื้นที่ที่ต้องครอบคลุมตลอดทั้งภาพสามารถมีขนาดหรือรูปร่างใดๆ ก็ได้

(2) รวบรวมช่วงของพื้นที่ที่กำหนด ซึ่งมีขนาดใหญ่กว่าโดเมนของพื้นที่และสามารถซ้อนทับกันได้ตลอดทั้งภาพ

(3) เลือกระดับของการแปลงข้อมูลที่ใช้งานในช่วงของพื้นที่ที่ใกล้เคียงกับพื้นที่ของโดเมน ข้อมูลของการแปลง 3 มิติแต่ละค่าจะมีความสัมพันธ์กัน

(4) จัดเก็บข้อมูลในรูปแบบของ FIF (Fractal Image Format) ประกอบด้วยส่วนหัว (Header) และรายละเอียด (Information) ของข้อมูลในการเลือกพื้นที่ตามด้วยกลุ่มรายชื่อของค่าสัมประสิทธิ์ที่เลือกในแต่ละพื้นที่ของโดเมน

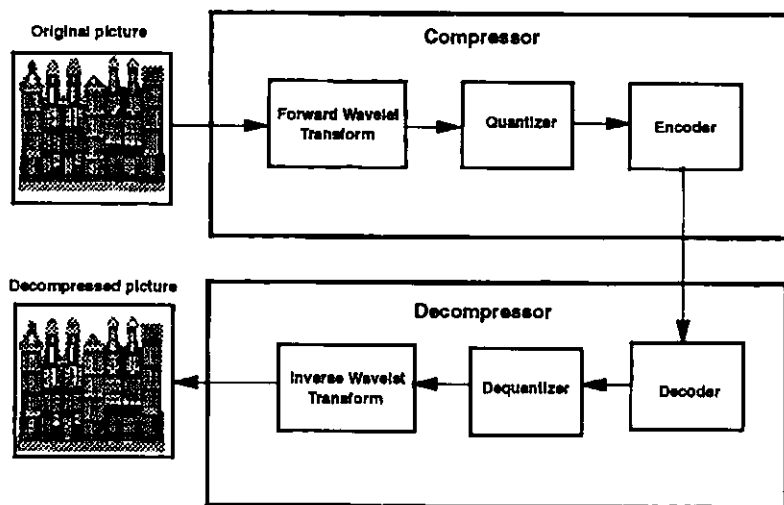
ขั้นตอนการบีบอัดข้อมูลจะให้ไฟล์ข้อมูลที่ไม่ขึ้นกับความละเอียดของภาพต้นแบบ (Original image) การกำหนดค่าพารามิเตอร์จะระบุเป็นสมการสำหรับภาพนั้นๆ เช่น สมการเส้นตรง หาได้ดังนี้ $y = ax + b$ โดยทราบค่าสัมประสิทธิ์ของ a และ b ภายหลังสามารถจะวาดที่ความละเอียดใดๆ ก็ได้เช่นเดียวกับค่าสัมประสิทธิ์ของภาพ อัลกอริทึมการขยายข้อมูลภาพจะสร้างแบบจำลองของแฟรคทอลหรือส่วนย่อยของภาพต้นแบบที่ความละเอียดใดๆ

ขั้นตอนการขยายข้อมูลภาพเป็นขั้นตอนการทำซ้ำ ในแต่ละพื้นที่ของโดเมนจะแทนที่ด้วยข้อมูลที่ผ่านการแปลงจากช่วงของพื้นที่ที่เหมาะสมโดยใช้ค่าสัมประสิทธิ์ที่เก็บไว้สำหรับโดเมนนั้นๆ ขั้นตอนการบีบอัดและขยายข้อมูลภาพแฟรคทอลเป็นลักษณะอสมมาตร (Asymmetric) นั่นคือต้องใช้เวลาการคำนวณในขั้นตอนการบีบอัดมากกว่าการขยายข้อมูล

สรุปได้ว่าการบีบอัดข้อมูลภาพแฟรคทอลมีลักษณะตรงข้ามกับ JPEG โดยที่การบีบอัดข้อมูลภาพแฟรคทอลได้ให้การขยายข้อมูลภาพที่รวดเร็วและไม่ขึ้นกับความละเอียด นอกจากนี้ยังให้อัตราบีบอัดข้อมูลที่สูง มีประโยชน์ในการใช้งานระบบมัลติมีเดียแบบสมมาตรที่ให้ความสำคัญของการเข้าถึงข้อมูลภาพที่รวดเร็วได้

2.2.4 การบีบอัดข้อมูลภาพด้วยการแปลงเวฟเลต

ช่วงเวลา 7 ปีที่ผ่านมา การบีบอัดข้อมูลภาพและวิดีโอโดยใช้พื้นฐานการแปลงเวฟเลตได้กลายเป็นหัวข้อที่น่าสนใจในงานวิจัยเป็นอย่างมาก ได้มีการนำเสนออัลกอริทึมการบีบอัดข้อมูลภาพด้วยการแปลงเวฟเลตมากมายและมีรูปแบบที่แตกต่างกันตั้งแต่การเข้ารหัสเอนโทรปี การควอนไทซ์แบบเวกเตอร์ การแปลงข้อมูลแบบปรับตัวเองได้ (Adaptive transformation) จนถึงการเข้ารหัสแบบต้นไม้ (Tree coding) และการเข้ารหัสโดยใช้พื้นฐานของขอบภาพ (Edge-based coding) บล็อกไดอะแกรมของอัลกอริทึมการบีบอัดข้อมูลภาพโดยใช้พื้นฐานการแปลงแบบเวฟเลตแสดงดังรูปที่ 2.8 การบีบอัดโดยใช้พื้นฐานการแปลงเวฟเลต ประกอบด้วย 3 ขั้นตอนดังนี้ (1) FWT (Forward wavelet transform) (2) การควอนไทซ์ค่าสัมประสิทธิ์ของการแปลง (3) การเข้ารหัสค่าสัมประสิทธิ์การควอนไทซ์



รูปที่ 2.8 การบีบอัดและการขยายข้อมูลภาพโดยใช้พื้นฐานการแปลงเวฟเลต

2.3 ความผิดเพี้ยนที่เกิดจากการบีบอัดข้อมูลภาพ

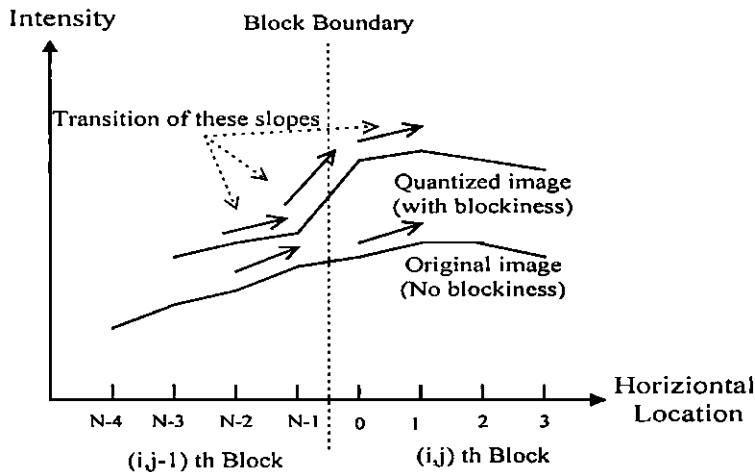
โดยทั่วไปคุณภาพในการมองเห็นของภาพที่ได้จากการบีบอัดข้อมูลแล้ว (Compressed image) ขึ้นอยู่กับองค์ประกอบที่สำคัญ 3 ประการคือ ชนิดของข้อมูล (Data source) บิตเรทของการเข้ารหัสข้อมูล (Coding bit rate) และขั้นตอนการบีบอัดข้อมูล (Compression algorithm) สำหรับรายละเอียดต่างๆ ของการบีบอัดข้อมูลภาพ เช่น สี ขนาดของภาพ เป็นต้น จะถูกเก็บไว้เป็นส่วนหนึ่งของข้อมูลภาพที่ต้องการจำนวนบิตในการจัดเก็บจำนวนหนึ่ง เมื่อผ่านขั้นตอนการบีบอัดข้อมูลที่ให้บิตเรทเดียวกัน พบว่าข้อมูลภาพในพื้นที่ที่มีรายละเอียดมากจะมีความผิดเพี้ยนของภาพมากกว่าพื้นที่ที่มีรายละเอียดน้อย บิตเรทของการเข้ารหัสเป็นองค์ประกอบที่สำคัญอย่างหนึ่งซึ่งแสดงถึงคุณภาพของภาพที่ได้จากการบีบอัดข้อมูลภาพ การบีบอัดข้อมูลภาพแบบมีการสูญเสียขึ้นอยู่กับความสัมพันธ์ระหว่างบิตเรทและความผิดเพี้ยนของภาพที่ได้ โดยที่บิตเรทต่ำกว่าจะให้ความผิดเพี้ยนของภาพมากกว่าเนื่องจากการสูญเสียข้อมูลไป ประเภทความผิดเพี้ยนของภาพขึ้นอยู่กับวิธีที่ใช้ในการบีบอัดข้อมูลภาพและจะปรากฏให้เห็นในลำดับของการบีบอัดข้อมูลภาพวิดีโอเช่นเดียวกัน สามารถแบ่งประเภทความผิดเพี้ยนของภาพได้ดังนี้

2.3.1 ความผิดเพี้ยนแบบบล็อก (Blocking artifact)

สามารถสังเกตได้จากภาพที่ผ่านขั้นตอนการบีบอัดข้อมูลภาพด้วยเทคนิคการเข้ารหัสที่เกี่ยวข้องกับการแบ่งข้อมูลภาพออกเป็นบล็อกย่อยๆ เช่น การบีบอัดข้อมูลภาพที่ใช้การแปลงแบบดิคคริตโคไซน์เป็นพื้นฐาน (JPEG) การควอนไทซ์แบบเวกเตอร์หรือ VQ (Vector Quantization) การเข้ารหัสแบบ BTC (Block Truncation Coding) และการเข้ารหัสข้อมูลภาพแฟรคทอล เป็นต้น ความผิดเพี้ยนแบบบล็อกสามารถสังเกตเห็นได้อย่างชัดเจนในมาตรฐานการบีบอัดข้อมูลภาพ JPEG และ MPEG โดยปรากฏให้เห็นเป็นขอบภาพในแนวนอนและแนวตั้งตลอดทั้งภาพ โดยเฉพาะในพื้นที่ของภาพที่มีความราบเรียบ (Smooth region) เป็นผลเนื่องมาจากการไม่คำนึงถึงความสัมพันธ์ของบล็อกข้อมูลที่อยู่ติดกัน นักวิจัยบางท่าน [8] ได้พิจารณาความผิดเพี้ยนแบบบล็อกที่เกิดขึ้นในลักษณะของความชันแบบไม่ต่อเนื่อง (Slope) ที่บริเวณขอบของบล็อกที่อยู่ติดกัน ถ้าค่าสัมประสิทธิ์ DCT ของบล็อกที่อยู่ติดกันถูกควอนไทซ์ด้วยค่าที่มาก (Coarse) ทำให้เห็นความชันของระดับความเข้ม (Intensity) ของพิกเซลที่บริเวณขอบแตกต่างจากพิกเซลที่ไม่ได้ผ่านการควอนไทซ์ โดยพิกเซลของภาพที่ไม่ได้ผ่านการควอนไทซ์นั้นจะมีการเปลี่ยนแปลงระดับความเข้มสีที่ต่อเนื่องหรือมีลักษณะของความชันไม่เปลี่ยนแปลงมากนักดังรูปที่ 2.9

2.3.2 .ความผิดเพี้ยนแบบวงแหวน (Ringing artifact)

เป็นผลเนื่องมาจากการควอนไทซ์ด้วยค่าที่มาก (Coarse quantization) ที่องค์ประกอบความถี่สูง ทำให้ค่าสัมประสิทธิ์ที่องค์ประกอบความถี่สูงถูกรองทิ้งไป ค่าความผิดพลาดของการควอนไทซ์จะเกิดขึ้นในลักษณะของสัญญาณรบกวนแบบวงแหวนที่บริเวณขอบภาพที่มีการเปลี่ยนแปลงแบบทันทีทันใด (Sharp edge) สัญญาณรบกวนนี้ขึ้นอยู่กับความยาวของฟิลเตอร์ที่ใช้ ถ้าความยาวของฟิลเตอร์มีค่ามาก สัญญาณรบกวนแบบวงแหวนจะแผ่ออกไปกว้างกว่า ตัวอย่างเช่น การบีบอัดข้อมูลภาพ JPEG มีความยาวของผลตอบสนองอิมพัลส์ของการแปลงแบบ DCT เท่ากับ 8 ซึ่งเป็นความยาวของผลตอบสนองอิมพัลส์ในช่วงสั้นจึงสังเกตเห็นความผิดเพี้ยนแบบวงแหวนได้น้อยที่บิตเรตสูงถึงปานกลางแต่จะเห็นได้อย่างชัดเจนที่บิตเรตต่ำ เป็นต้น โดยทั่วไปกล่าวได้ว่าความผิดเพี้ยนแบบวงแหวนเกิดขึ้นกับรูปแบบของการเข้ารหัสที่เกี่ยวข้องกับการควอนไทซ์ในโดเมนความถี่และสังเกตเห็นได้ชัดเจนในการเข้ารหัสด้วยการแปลงเวฟเลตที่บิตเรตต่ำ



รูปที่ 2.9 ลักษณะความชันที่ไม่ต่อเนื่องของพิกเซลที่บริเวณขอบบล็อก

2.3.3 การลดทอนความคมชัดของภาพหรือภาพเบลอ (Blurring)

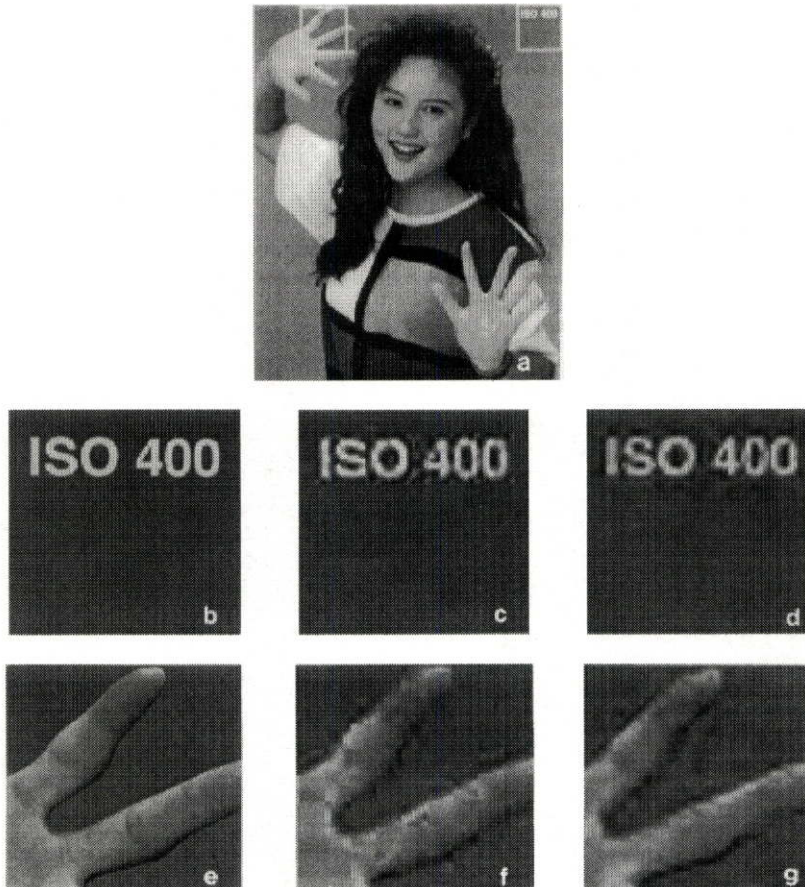
การเบลอของภาพเป็นผลเนื่องมาจากการสูญเสียองค์ประกอบความถี่สูงไปและเกิดขึ้นได้ในเทคนิคของการเข้ารหัสแบบมีการสูญเสียที่บิตเรตต่ำ อย่างไรก็ตามระดับการเบลอของภาพจะไม่รบกวนในการมองเห็นของผู้สังเกตการณ์ ถ้าระยะของการมองเห็นที่มีค่ามากผลของความเบลอในภาพจะลดลง ลักษณะของภาพเบลอปรากฏให้เห็นในรูปแบบที่แตกต่างกัน เช่น บริเวณขอบภาพที่มีความราบเรียบ การเบลอในเนื้อหาของภาพและการเบลอแบบเป็นช่วงๆ ในการเคลื่อนไหวของวัตถุเนื่องจากการประมาณทิศทางการเคลื่อนไหวที่ไม่ดีพอ (Motion compensaion) เป็นต้น

2.3.4 การเบี่ยงเบนในเนื้อหาของภาพ (Texture deviation)

ชนิดของความผิดเพี้ยนของภาพนอกเหนือจากที่กล่าวมานั้นเรียกว่าการเบี่ยงเบนในเนื้อหาของภาพที่เกิดขึ้นเนื่องจากการสูญเสียข้อมูลที่องค์ประกอบความถี่ปานกลางและมีลักษณะของสัญญาณรบกวนเป็นจุดเล็กๆ กระจายอยู่ทั่วไป (Granular noise) ซึ่งตาของคนเรามีความไวต่อสัญญาณรบกวนชนิดนี้น้อยมากในพื้นที่ของภาพที่มีรายละเอียดมาก

2.3.5 ความผิดเพี้ยนเชิงเรขาคณิต (Geometrical deformation)

ในการเข้ารหัสด้วยภาพจำลอง (Model-based coding) ข้อมูลภาพหรือส่วนใดส่วนหนึ่งของข้อมูลภาพ เช่น รูปหน้าคน เป็นต้น จะถูกนำมาสร้างแบบจำลองและพารามิเตอร์ของแบบจำลองนี้ จะใช้สำหรับการสังเคราะห์ข้อมูลภาพ (Image synthesis) เมื่อทำการสังเคราะห์ข้อมูลภาพกลับมา อาจทำให้รูปร่างและตำแหน่งที่เป็นส่วนสำคัญของข้อมูลภาพเปลี่ยนไป นำไปสู่ความผิดพลาดได้



รูปที่ 2.10 ความผิดเพี้ยนของภาพที่เกิดจากการบีบอัดข้อมูลภาพ (a) ภาพต้นแบบ (b) และ (e) ภาพขยายของข้อความและมือ (c) และ (f) ภาพที่ได้จากการเข้ารหัส JPEG ที่ 0.26 บิต/พิกเซล (d) และ (g) ภาพที่ได้จากการเข้ารหัสการแปลงเวฟเล็ดที่ 0.125 บิต/พิกเซล

ในที่นี้แสดงตัวอย่างความผิดเพี้ยนของภาพที่เกิดจากการบีบอัดข้อมูลภาพที่พบเห็นทั่วไป คือ ความผิดเพี้ยนแบบบล็อก ความผิดเพี้ยนแบบวงแหวนและการเบลอของภาพดังรูปที่ 2.10 แสดงการเปรียบเทียบให้เห็นระหว่างภาพต้นแบบของรูปที่ 2.10(a) และแสดงภาพพื้นที่บางส่วน ของภาพต้นแบบ 2 ส่วนคือข้อความและนิ้วมือนี้อย่างรูปที่ 2.10(b) และ 2.10(c) ตามลำดับ ข้อมูลภาพที่ ผ่านการเข้ารหัสแบบ JPEG ที่ 0.26 บิต/พิกเซล จะแสดงให้เห็นความผิดเพี้ยนแบบบล็อกและความ ผิดเพี้ยนแบบวงแหวนดังรูปที่ 2.10(d) และ 2.10(e) เช่นเดียวกันเมื่อข้อมูลภาพผ่านการเข้ารหัสด้วย การแปลงเวฟเล็ตที่ 0.125 บิต/พิกเซล ผลของภาพที่ได้แสดงดังรูปที่ 2.10(f) และ 2.10(g) ซึ่งแสดง ให้เห็นความผิดเพี้ยนแบบวงแหวนได้อย่างชัดเจนที่บริเวณขอบของภาพ และสามารถสังเกตเห็น ลักษณะความเบลอของภาพดังรูปที่ 2.10(c)-(g)

2.4 การวัดประสิทธิภาพการบีบอัดข้อมูลภาพ

ในการบีบอัดข้อมูลภาพแบบมีการสูญเสีย ประสิทธิภาพของเทคนิคต่างๆ ไม่สามารถสรุป ได้จากอัตราการบีบอัดข้อมูลเพียงอย่างเดียว แต่ต้องพิจารณาถึงความสูญเสียข้อมูลที่เกิดขึ้นว่าอยู่ใน ช่วงที่ยอมรับได้หรือไม่ เกณฑ์ที่ใช้ในการวัดประสิทธิภาพของการบีบอัดข้อมูลภาพสามารถแบ่ง ออกเป็น 2 ประเภทใหญ่ ดังนี้

2.4.1 การวัดเชิงคุณภาพ

เป็นวิธีการพื้นฐานแต่ให้ผลที่น่าเชื่อถือมาก โดยใช้สายตาของผู้ที่ไม่เกี่ยวข้องทำการตัดสิน คุณภาพของภาพในฐานะที่เป็นผู้ใช้ธรรมดาและยังใช้สายตาของผู้เชี่ยวชาญที่มีความสามารถในการ สังเกตความบกพร่องเล็กน้อยที่ผู้ใช้ทั่วไปอาจมองข้าม ในการตัดสินใจอาจใช้กฎเกณฑ์ต่างๆ ดังนี้

(1) การตัดสินใจจากภาพทดสอบแล้วให้ระดับคะแนน เมื่อผู้ทดสอบได้มองภาพทดสอบ แล้วระบุระดับคุณภาพตามที่ได้ระบุไว้ โดยอาจมีชุดภาพมาตรฐานที่มีระดับคะแนนระบุไว้อ้างอิง ด้วยก็ได้ ระดับคุณภาพสามารถแสดงได้ทั้งระดับความถูกต้องและระดับความผิดพลาด

(2) การตัดสินใจจากชุดภาพทดสอบแล้วเรียงลำดับคุณภาพของภาพ ซึ่งผู้ใช้จะเรียงลำดับ ภาพทดสอบตามคุณภาพของภาพที่เห็น ผู้ทดสอบจะตัดสินใจในลักษณะเปรียบเทียบว่าดีกว่าหรือ แย่กว่า

2.4.2 การวัดเชิงปริมาณ

การวัดค่าความผิดพลาดที่เกิดจากการบีบอัดข้อมูลสามารถกระทำได้หลายวิธีการ คือ

ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error, MSE)

ในการหาค่าความผิดพลาดเชิงเลข นิยมใช้ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error, MSE) ถ้ามีค่าน้อยแสดงว่าภาพที่ได้กลับคืนมามีความผิดเพี้ยนจากภาพต้นแบบน้อย สามารถหาค่าความผิดพลาดกำลังสองเฉลี่ยของข้อมูลภาพขนาด $M \times N$ ได้ดังนี้

$$MSE = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N [f(x, y) - \hat{f}(x, y)]^2 \quad (2.5)$$

โดยที่ M คือ จำนวนพิกเซลตามความกว้างของภาพ

N คือ จำนวนพิกเซลตามความสูงของภาพ

$f(x, y)$ คือ ค่าของพิกเซลที่ตำแหน่ง (x, y) ของภาพต้นแบบ

$\hat{f}(x, y)$ คือ ค่าของพิกเซลที่ตำแหน่ง (x, y) ของภาพที่สร้างกลับคืนมาใหม่

อัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak signal-to-noise ratio, PSNR)

ในการใช้งานบางประเภท นิยมใช้ค่า PSNR หาได้จาก

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

โดยที่ Peak คือค่าสูงสุดของขอบเขตข้อมูล สำหรับภาพที่มีพิกเซลแทนด้วยขนาดข้อมูล n บิต จะมีจำนวนระดับความเทา (gray scale) เท่ากับ 2^n จะได้ว่า

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

ดังนั้นภาพขนาด 8 บิตหรือมีจำนวนระดับความเทาเท่ากับ 256 ระดับ จะได้ว่า

$$PSNR = 10 \log \frac{255^2}{MSE} \quad (\text{dB}) \quad (2.8)$$

การวัดค่าความสามารถในการบีบอัดข้อมูลสามารถหาได้จาก

บิตเรท (Bit rate)

บิตเรทคือค่าเฉลี่ยของจำนวนบิตต่อพิกเซล (Bits per pixel, bpp) ของภาพที่ถูกลดข้อมูล โดยสามารถคำนวณได้จากอัตราส่วนระหว่างจำนวนบิตของภาพที่ผ่านการบีบอัดข้อมูลทั้งหมดต่อจำนวนพิกเซลของภาพต้นแบบ ดังสมการต่อไปนี้

$$\text{Bit rate}(bpp) = \frac{\text{จำนวนบิตทั้งหมดของภาพที่ผ่านการบีบอัดข้อมูล}}{\text{จำนวนพิกเซลทั้งหมดของภาพต้นแบบ}} \quad (2.9)$$

อัตราการบีบอัดข้อมูลภาพ (Compression ratio, CR)

อัตราการบีบอัดข้อมูลภาพคืออัตราส่วนระหว่างจำนวนบิตข้อมูลภาพต้นแบบและจำนวนบิตข้อมูลภาพที่ผ่านการบีบอัดข้อมูล หาได้จากสมการต่อไปนี้

$$\text{อัตราการบีบอัดข้อมูลภาพ} = \frac{\text{จำนวนบิตทั้งหมดของภาพต้นแบบ}}{\text{จำนวนบิตทั้งหมดของภาพที่ผ่านการบีบอัดข้อมูล}} \quad (2.10)$$

2.5 สรุป

ในปัจจุบันระบบต่างๆ ได้ให้ความสำคัญกับการลดขนาดข้อมูลหรือที่เรียกว่าการบีบอัดข้อมูลภาพ ซึ่งมีหลายวิธีที่แตกต่างกันไป สำหรับงานด้านการประมวลผลภาพนิยมใช้การบีบอัดข้อมูลภาพแบบมีการสูญเสียเนื่องจากให้อัตราการบีบอัดข้อมูลภาพที่สูงขณะเดียวกันทำให้คุณภาพของภาพที่ได้แยกลงหรือเกิดความผิดเพี้ยนของภาพขึ้น ที่เห็นได้อย่างชัดเจนคือความผิดเพี้ยนแบบบล็อกที่เกิดจากการบีบอัดข้อมูลภาพแบบ BDCT ในวิทยานิพนธ์นี้จะเน้นที่ความผิดเพี้ยนแบบบล็อกที่เกิดจากการบีบอัดข้อมูลภาพแบบ JPEG เนื่องจากความผิดเพี้ยนในลักษณะนี้สามารถสังเกตเห็นได้ชัดเจนที่สุดและใช้ในการเตรียมข้อมูลสำหรับการวิจัยเพื่อลดขอบบล็อกรูปข้อมูลภาพต่อไป

บทที่ 3

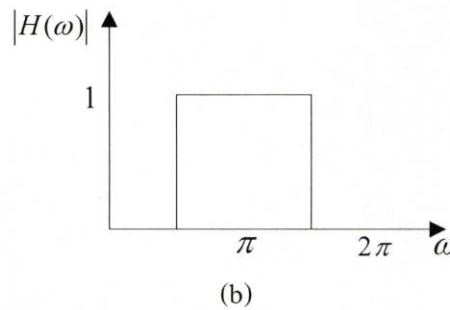
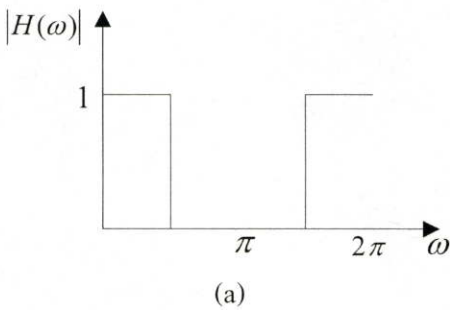
การแปลงเวฟเล็ต

บทนี้จะกล่าวถึงคุณสมบัติของการแปลงเวฟเล็ตแบบดิสครีต (Discrete Wavelet Transform : DWT) 1 มิติและ 2 มิติ ในลักษณะของฟิลเตอร์แบงก์ (Filter bank) การกำหนดผลตอบสนองอิมพัลส์ของฟิลเตอร์ที่นำมาใช้ในขบวนการ DWT ผลตอบสนองอิมพัลส์ที่ได้ทำให้ทราบคุณสมบัติของเวฟเล็ตที่ใช้ คุณสมบัติและเงื่อนไขของเวฟเล็ตแบบออร์โธโกนัล (Orthogonal) และแบบไบออร์โธโกนัล (Biorthogonal) ที่นิยมใช้ทั่วไปในงานด้านการประมวลผลภาพ การนำไปประยุกต์ใช้ในการหาขอบที่หลายระดับของการแปลงภาพโดยการหาอนุพันธ์ของการแปลงเวฟเล็ตใน 1 มิติและ 2 มิติตามลำดับ

3.1 การแปลงเวฟเล็ตแบบดิสครีต [9]

3.1.1 การแปลงเวฟเล็ตแบบดิสครีตใน 1 มิติ

การแยกข้อมูลไปเป็นลำดับชั้นของความละเอียดต่ำ (Coarse) และความละเอียดสูง (Detail) ต้องการตัวฟิลเตอร์ที่มีคุณสมบัติในการกรองความถี่ต่ำและความถี่สูงเชิงอุดมคติ ดังรูปที่ 3.1 กำหนดให้ $x(n)$ คือลำดับของข้อมูลอินพุต เงื่อนไขการออกแบบตัวฟิลเตอร์ของการกรองความถี่ต่ำ H_0 และการกรองความถี่สูง H_1 เป็นดังนี้



รูปที่ 3.1 ผลตอบสนองความถี่เชิงอุดมคติของ (a) การกรองความถี่ต่ำ (b) การกรองความถี่สูง

(1) ฟิลเตอร์แบงก์แบบ 2 แบนด์

พิจารณาจากรูปที่ 3.2(b) และกำหนดให้ $h_0(k)$ และ $h_1(k)$ คือผลตอบสนองอิมพัลส์ที่จำกัด (Finite impulse response : FIR) เขียนความสัมพันธ์ได้ดังนี้

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

เมื่อ $y_1(k)$ คือเอาที่พู่ทขององค์ประกอบความถี่ต่ำดังรูปที่ 3.2(b) ความสัมพันธ์ของ $y_0(k)$ และ $y_1(k)$ โดยที่ $k = 0, 1, 2, \dots$ ในรูปของเวกเตอร์คือ

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

หรือ
$$\mathbf{y} = T_i \mathbf{x} \quad (3.1)$$

โครงสร้างของ T_i ประกอบด้วยผลตอบสนองอิมพัลส์ของ H_0 และ H_1 จากนั้นทำการลดข้อมูลลง 2 เท่า (Subsampling) เพื่อสร้างแถวข้อมูลใหม่ ทำนองเดียวกันพิจารณาจากรูปที่ 3.3(b) แสดงโครงสร้างการรวมกันของ $y_0(k)$ และ $y_1(k)$ ที่ผ่านฟิลเตอร์การกรองความถี่ต่ำ G_0 และฟิลเตอร์การกรองความถี่สูง G_1 เพื่อสร้างลำดับข้อมูลเอาที่พู่ท \hat{x} สัญลักษณ์ของการเพิ่มข้อมูล 2 เท่า (Upsampling) ด้วยตัวดำเนินการ M แสดงดังรูปที่ 3.3(a) กรณีนี้ใช้ $M = 2$ และทำการเพิ่มศูนย์จำนวน $M - 1$ ระหว่างค่าสัมประสิทธิ์เวฟเล็ท 2 ค่า จะได้ลำดับอินพุทของฟิลเตอร์การกรองความถี่ต่ำ G_0 และฟิลเตอร์การกรองความถี่สูง G_1 มีลักษณะดังนี้

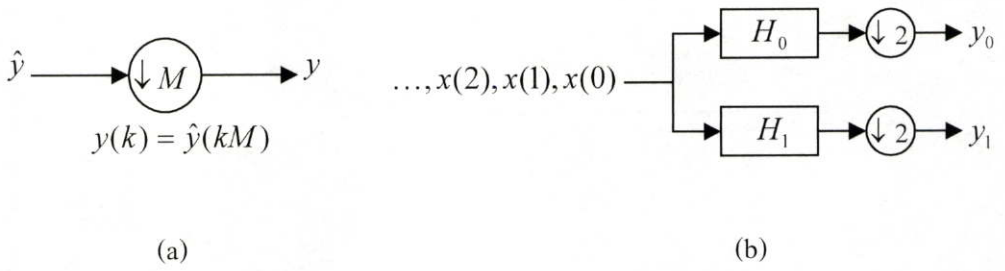
$$\begin{array}{cccccccc} \dots & 0 & y_0(0) & 0 & y_0(1) & 0 & y_0(2) & 0 & \dots \\ \dots & 0 & y_1(0) & 0 & y_1(1) & 0 & y_1(2) & 0 & \dots \end{array}$$

จะได้ว่า

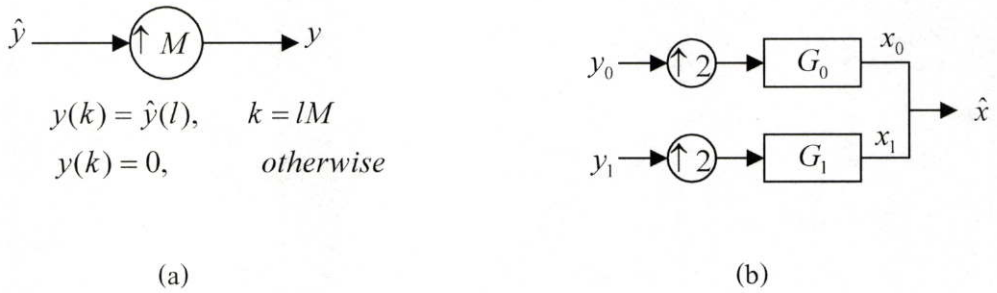
$$x_0(n) = \sum_k y_0(k)g_0(n-2k)$$

$$x_1(n) = \sum_k y_1(k)g_1(n-2k)$$

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



รูปที่ 3.2 (a) ตัวดำเนินการลดข้อมูล (b) การวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์



รูปที่ 3.3 (a) ตัวดำเนินการเพิ่มข้อมูล (b) การสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์

ฟิลเตอร์ G_i คือฟิลเตอร์ของการสังเคราะห์สัญญาณ (Synthesis filters) และสัมพันธ์กับฟิลเตอร์ H_i หรือที่เรียกว่าฟิลเตอร์ของการวิเคราะห์สัญญาณ (Analysis filters) การรวมกันของ $\hat{x}(n)$ ทุกๆ ค่า เป็นดังนี้

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

หรือ $\hat{\mathbf{x}} = T_o \mathbf{y}$ (3.2)

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

$$T_o T_i = I = T_i T_o \quad (3.3)$$

การคูณแถวของ T_i ด้วยคอลัมน์ของ T_o เท่ากับ

$$\sum_n h_i(2k-n)g_j(n-2l) = \delta_{ki}\delta_{jl}, \quad i, j = 0,1 \quad (3.4)$$

สามารถเขียนเป็นผลคูณภายใน (Inner product) ได้ดังนี้

$$\langle h_i(2k-n)g_j(n-2l) \rangle = \delta_{ki}\delta_{jl}$$

หมายถึงผลคูณภายในของเวฟเล็ตจะเป็นหนึ่งได้เมื่อเป็นเวฟเล็ตที่ระดับของการแปลงและตำแหน่งเดียวกันเท่านั้น ถ้าฟิลเตอร์แบงก์แบบ 2 แบนด์เป็นไปตามเงื่อนไขของสมการที่ (3.4) สามารถกล่าวได้ว่าฟิลเตอร์แบงก์แบบ 2 แบนด์มีลักษณะของการสร้างข้อมูลกลับคืนมาอย่างสมบูรณ์ (Perfect reconstruction) และ $\hat{x}(n) = x(n)$ ดังนั้นจะได้ว่า

$$x(n) = \sum_k y_0(k)g_0(n-2k) + \sum_k y_1(k)g_1(n-2k) \quad (3.5)$$

สมการที่ (3.5) คือการกระจายของ $x(n)$ ออกเป็นเซตลำดับมูลฐาน (Basis sequence)

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

เมื่อ Z คือเซตของจำนวนเต็ม ในส่วนของ $y_0(k)$ และ $y_1(k)$ คือค่าสัมประสิทธิ์เวฟเล็ตของการกระจาย ดังนั้นลักษณะของฟิลเตอร์แบงก์แบบ 2 แบนด์ที่สร้างกลับคืนมาอย่างสมบูรณ์ต้องสอดคล้องกับเงื่อนไขในสมการที่ (3.4) การแปลงสามารถกำหนดได้ดังนี้

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

$$x(n) = \sum_{i=0}^1 \sum_k y_i(k)g_i(n-2k) \quad (b) \quad (3.6)$$

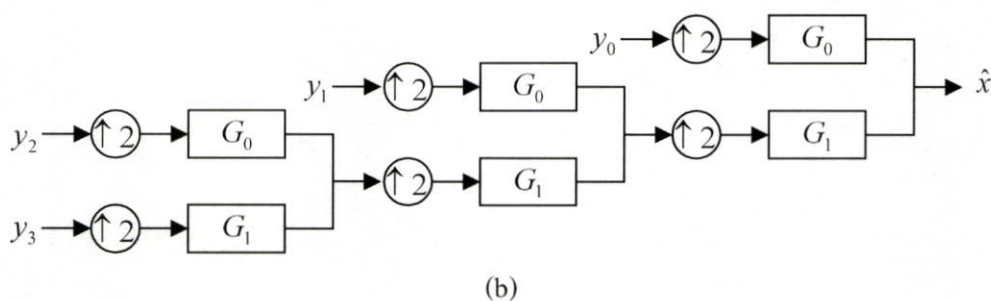
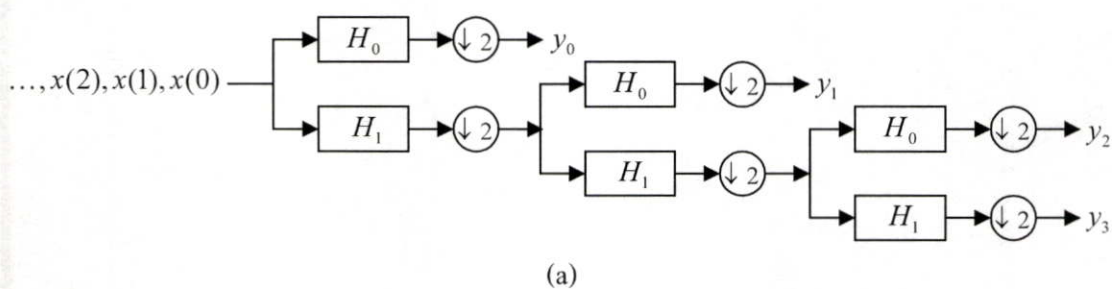
(2) ฟิลเตอร์แบงก์แบบหลายแบนด์

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

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

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

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



รูปที่ 3.4 (a) โครงสร้างการวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์แบบหลายแบนด์ (b) โครงสร้างการสังเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์แบบหลายแบนด์

เมื่อ J คือจำนวนของแบนด์ จากรูปที่ 3.4 จะได้ $J = 4$ ถ้าฟิลเตอร์ G_0 และ G_1 ใช้ในการสังเคราะห์สัญญาณและฟิลเตอร์ H_0 และ H_1 ใช้ในการวิเคราะห์สัญญาณที่สอดคล้องกับเงื่อนไข

ของสมการที่ (3.4) ดังนั้นการวิเคราะห์และสังเคราะห์สัญญาณของฟิลเตอร์แบงก์ที่ระดับ J จะมีลักษณะของการสร้างข้อมูลกลับคืนมาได้อย่างสมบูรณ์นั่นคือ

$$\hat{x}(n) = x(n) = \sum_{i=0}^{J-2} \sum_k y_i(k) f_i(n - 2^{i+1}k) + \sum_k y_{J-1}(k) f_{J-1}(n - 2^{J-1}k) \quad (3.7)$$

โดยที่

$$y_i(k) = \sum_n x(n) \hat{f}_i(2^{i+1}k - n), \quad i = 0, 1, \dots, J-2 \quad (3.8)$$

$$y_{J-1}(k) = \sum_n x(n) \hat{f}_{J-1}(2^{J-1}k - n) \quad (3.9)$$

3.1.2 การแปลงเวฟเลตแบบดิสครีตใน 2 มิติ

ในที่นี้กล่าวถึงการแปลงข้อมูลภาพ 2 มิติ แบ่งออกเป็น 2 ขั้นตอนคือ

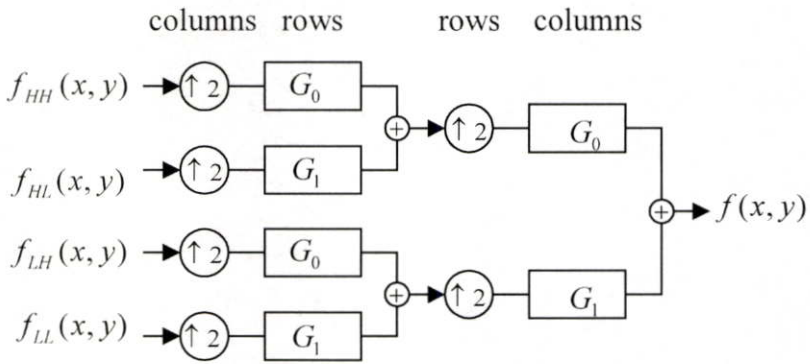
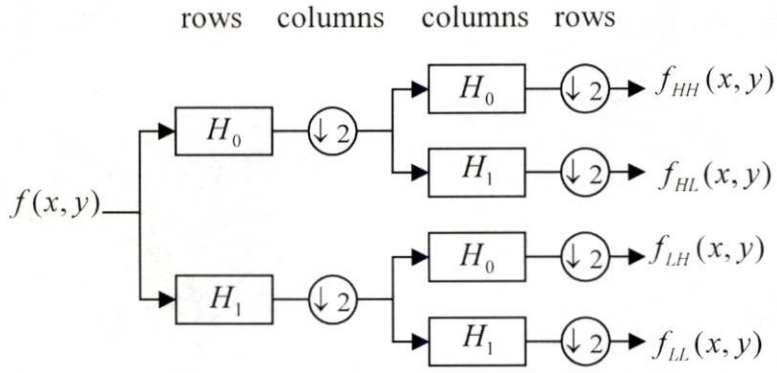
(1) การแยกองค์ประกอบเวฟเลต (Wavelet decomposition)

เป็นการแยกภาพต้นแบบออกเป็นองค์ประกอบย่อย ดังรูปที่ 3.5(a) โดยเริ่มจากภาพต้นแบบ $f(x, y)$ จะถูกกรองไปตามแนวแถว (Row) ด้วยฟิลเตอร์การกรองความถี่สูง H_0 และฟิลเตอร์การกรองความถี่ต่ำ H_1 จากนั้นนำผลที่ได้มาลดการจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูล 1 คอลัมน์ (Column) จากข้อมูล 2 คอลัมน์ จะได้เอาท์พุทความถี่สูงและความถี่ต่ำ จากนั้นเอาท์พุททั้งสองจะถูกกรองตามแนวคอลัมน์และลดจำนวนข้อมูลลง 2 เท่าโดยการเก็บข้อมูลมา 1 แถว จากข้อมูล 2 แถว จะได้ภาพแบนด์ย่อยความถี่ต่ำ $f_{LL}(x, y)$ จำนวน 1 ภาพและภาพแบนด์ย่อยความถี่สูงจำนวน 3 ภาพคือ $f_{LH}(x, y)$ $f_{HL}(x, y)$ และ $f_{HH}(x, y)$ โดยแต่ละภาพจะมีขนาดลดลงเป็น 1/4 เท่าของภาพต้นแบบ

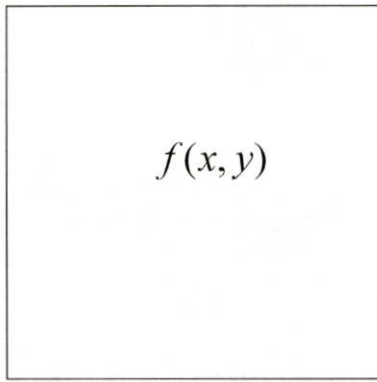
(2) การรวมองค์ประกอบของเวฟเลต (Wavelet reconstruction)

เป็นการสร้างภาพขึ้นมาใหม่จากองค์ประกอบแบนด์ย่อย ดังรูปที่ 3.5(b) โดยเริ่มจากภาพแบนด์ย่อยทั้ง 4 ภาพจะถูกเพิ่มจำนวนข้อมูลเป็น 2 เท่า โดยการเติมศูนย์เข้าไปในระหว่างแต่ละคู่ของคอลัมน์ หลังจากเพิ่มจำนวนข้อมูลแล้วผลลัพธ์ที่ได้จะถูกกรองไปตามแถวและบวกเข้าด้วยกัน จากนั้นทำซ้ำไปตามแถวและคอลัมน์จะได้ผลลัพธ์ที่สร้างกลับขึ้นมาใหม่

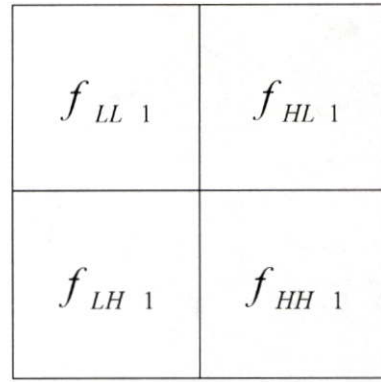
การแปลงเวฟเลตของภาพ 1 ครั้ง (Stage) จะได้ภาพที่มีขนาดเท่ากันจำนวน 4 ภาพ และเป็นภาพส่วนที่เป็นข้อมูลรายละเอียดหรือข้อมูลความถี่สูงจำนวน 3 ภาพและเป็นภาพระดับความสว่างหรือข้อมูลความถี่ต่ำจำนวน 1 ภาพ ขั้นตอนการแปลงเวฟเลตสามารถกระทำกับข้อมูลภาพที่องค์ประกอบความถี่ต่ำไปเรื่อยๆ ดังรูปที่ 3.6



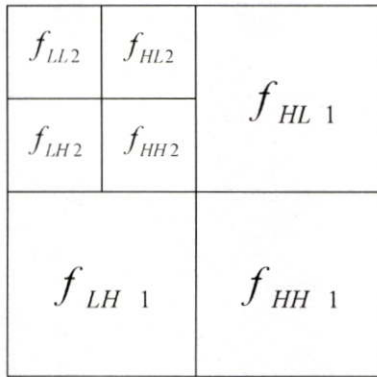
รูปที่ 3.5 การแปลงเวฟเล็ดของข้อมูลภาพโดยใช้หลักการของฟิลเตอร์แบงก์ (a) การแยกองค์ประกอบของเวฟเล็ด (b) การรวมองค์ประกอบของเวฟเล็ด



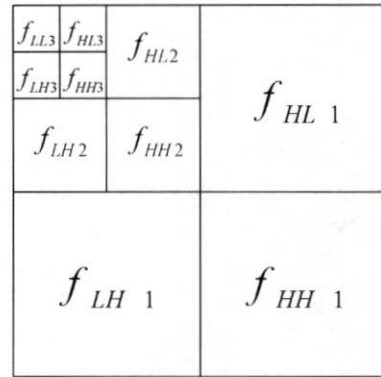
(a)



(b)



(c)



(d)

รูปที่ 3.6 ภาพแบนด์ย่อยที่ได้จากการแปลงเวฟเลต (a) ข้อมูลภาพต้นแบบ (b) การแปลงเวฟเลต 1 ครั้ง (c) การแปลงเวฟเลต 2 ครั้ง (d) การแปลงเวฟเลต 3 ครั้ง

3.1.3 เวฟเลตแบบออร์โธโกนัลและแบบไบออร์โธโกนัล

เวฟเลตแบบออร์โธโกนัลและแบบไบออร์โธโกนัลพิจารณาได้จากเซตของฟังก์ชันมูลฐานมีความสัมพันธ์ดังนี้

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

จากสมการที่ (3.4) มีคุณสมบัติเชิงตั้งฉาก (Orthogonal) ระหว่าง $\phi_{ik}(n)$ และ $\psi_{jl}(n)$ นั่นคือ

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

เรียกว่าเวฟเลตแบบไบออร์โธโกนัล การแปลงเวฟเลตแบบดิสคริตในสมการที่ (3.6) คือการกระจายของไบออร์โธโกนัล (Biorthogonal expansion)

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

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

ดังนั้น $g_i(n) = h_i(-n)$ หรือฟิลเตอร์ของการสังเคราะห์คือฟิลเตอร์ของการวิเคราะห์สัญญาณที่เวลาย้อนกลับและมีเซตของลำดับ h_i อย่างเดียวกันนั้นในสมการที่ (3.6) เรียกว่าเวฟเล็ตแบบออโธโกนัล ตารางที่ 3.1 แสดงตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ตแบบออโธโกนัลชนิด Daubechies ($h_1(n)$) และฟิลเตอร์การกรองความถี่สูงหาได้จาก $h_0(n) = (-1)^n h_1(-n + 2L - 1)$ เมื่อ L คือความยาวของฟิลเตอร์หรือจำนวนของค่าสัมประสิทธิ์ฟิลเตอร์ สามารถสรุปความสัมพันธ์ของการแปลงเวฟเล็ตแบบดิสครีตได้ดังตารางที่ 3.2

ตารางที่ 3.1 ตัวอย่างฟิลเตอร์การกรองความถี่ต่ำของเวฟเล็ตแบบออโธโกนัลชนิด Daubechies ที่มีความยาวของฟิลเตอร์เท่ากับ 4, 6, 8 และ 10 ตามลำดับ

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

ตารางที่ 3.2 เงื่อนไขการแปลงเวฟเล็ตแบบดิสครีต

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

3.2 การหาขอบที่หลายระดับของการแปลงภาพ

จุดที่มีการเปลี่ยนแปลงอย่างทันทีทันใด (Sharp variation point) เป็นลักษณะที่สำคัญในการวิเคราะห์คุณสมบัติของสัญญาณที่ขณะใดขณะหนึ่ง (Transient signal) หรือสัญญาณภาพ โดยทั่วไปการเปลี่ยนแปลงอย่างทันทีทันใดของข้อมูลภาพจะระบุขอบเขตของโครงสร้างที่สำคัญของข้อมูล นักวิจัยด้าน Computer Vision ได้เสนอแนวคิดของการหาขอบที่หลายระดับของการแปลงภาพ (Multiscale edge detection) ระดับของการแปลงภาพจะเป็นตัวกำหนดขนาดของสัญญาณที่บริเวณใกล้เคียงเพื่อคำนวณหาการเปลี่ยนแปลงของสัญญาณนั้นๆ การแปลงเวฟเลตมีความสัมพันธ์อย่างมากกับการหาขอบที่หลายระดับของการแปลงภาพและสามารถเข้าใจขั้นตอนการทำงานได้เป็นอย่างดี

จุดที่มีการเปลี่ยนแปลงอย่างทันทีทันใดในข้อมูลภาพ มีหลายรูปแบบแตกต่างกันไปขึ้นอยู่กับโครงสร้างของภาพ เช่น ขอบภาพที่เกิดจากเงา (Shadow) หรือเนื้อหาลายของภาพ (Texture) เป็นต้น ลักษณะของจุดเหล่านี้จะมีระดับความเข้มของข้อมูลภาพที่แตกต่างกัน เพื่อทำการหาขอบภาพได้อย่างถูกต้องจำเป็นต้องมีการวิเคราะห์คุณสมบัติของจุดเหล่านั้น โดยทั่วไปแล้วในทางคณิตศาสตร์ ลักษณะเฉพาะของจุดเหล่านี้สามารถจำแนกได้ด้วยค่า Lipchitz exponent และทฤษฎีของการแปลงเวฟเลตสามารถพิสูจน์ให้เห็นว่าค่า Lipchitz exponent สามารถคำนวณได้จากค่ามอดูลัสสูงสุดของการแปลงเวฟเลตในทุกๆ ระดับของการแปลงภาพ [11] ในหัวข้อนี้จะอธิบายถึงการหาขอบภาพที่หลายระดับของสัญญาณใน 1 มิติและ 2 มิติตามลำดับ รวมทั้งการจำแนกลักษณะขอบภาพด้วยค่า Lipchitz exponent

3.2.1 การหาอนุพันธ์ด้วยการแปลงเวฟเลตใน 1 มิติ

(1) คุณสมบัติทั่วไป

กำหนดให้ฟังก์ชัน $\psi_{2^j}(x)$ แทนการขยาย (dilation) ของฟังก์ชัน $\psi(x)$ ด้วยแฟกเตอร์ของ 2^j

$$\psi_{2^j}(x) = \frac{1}{2^j} \psi\left(\frac{x}{2^j}\right) \quad (3.10)$$

การแปลงเวฟเลตของฟังก์ชัน $f(x)$ ที่ระดับของการแปลงภาพ 2^j และที่ตำแหน่ง x สามารถเขียนในรูปผลคูณของการคอนโวลูชัน ได้ดังนี้

$$W_{2^j} f(x) = f * \psi_{2^j}(x) \quad (3.11)$$

เรียกว่าลำดับของการแปลงเวฟเลตของฟังก์ชัน

$$(W_{2^j} f(x))_{j \in \mathbb{Z}} \quad (3.12)$$

การแปลงฟูเรียร์ของ $W_{2^j} f(x)$ คือ

$$\hat{W}_{2^j} f(\omega) = \hat{f}(\omega) \hat{\psi}(2^j \omega) \quad (3.13)$$

จะได้ว่า

$$\sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j \omega)|^2 = 1 \quad (3.14)$$

ดังนั้นการขยายของฟังก์ชัน $\hat{\psi}(\omega)$ ด้วยแฟลคเตอร์ของระดับของการแปลงภาพ $(2^j)_{j \in \mathbb{Z}}$ จะครอบคลุมความถี่ทั้งหมดและสามารถสร้างฟังก์ชัน $f(x)$ กลับคืนมาได้จากการแปลงเวฟเล็ตของฟังก์ชันนั้นๆ รูปที่ 3.7(a) และ 3.7(b) แสดงตัวอย่างของสัญญาณและการแปลงเวฟเล็ตของสัญญาณที่เป็นลำดับ ถ้าสัญญาณต้นแบบมีลักษณะไม่ต่อเนื่องจำนวน N ค่า จะใช้การคำนวณในการแปลงเวฟเล็ตของสัญญาณจำนวน $O(N \log(N))$ ครั้ง การสร้างสัญญาณกลับคืนมาจากการแปลงเวฟเล็ตของสัญญาณเดิมต้องใช้การคำนวณเท่ากับ $O(N \log(N))$ ครั้ง เช่นเดียวกัน

ฟังก์ชันเวฟเล็ต $\psi(x)$ บางฟังก์ชันจะมีค่าสูงสุดของการแปลงเวฟเล็ตเช่นเดียวกับจุดที่มีการเปลี่ยนแปลงอย่างทันทีทันใดของสัญญาณ ถ้ากำหนดให้ $\theta(x)$ เป็นฟังก์ชันของความราบเรียบ (Smoothing function) กำหนดให้ $\psi(x)$ เป็นอนุพันธ์อันดับที่ 1 ของฟังก์ชัน $\theta(x)$

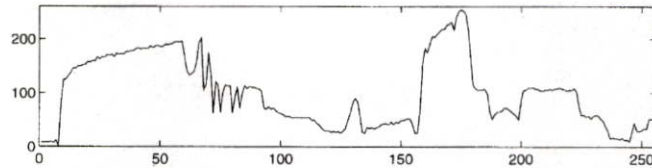
$$\psi(x) = \frac{d\theta(x)}{dx} \quad (3.15)$$

กำหนดให้ $\theta_{2^j}(x) = \frac{1}{2^j} \theta\left(\frac{x}{2^j}\right)$ การแปลงเวฟเล็ตที่ระดับของการแปลงภาพ 2^j หาได้จาก

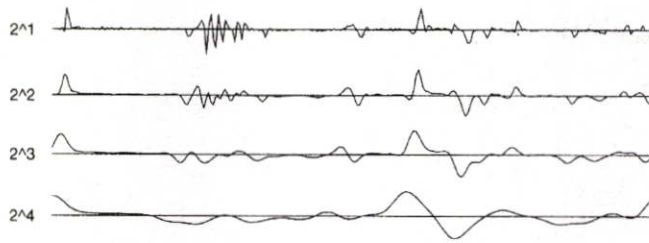
$$\begin{aligned} W_{2^j} f(x) &= f * \psi_{2^j}(x) \\ &= f * \left(2^j \frac{d\theta_{2^j}}{dx}\right)(x) \\ &= 2^j \frac{d}{dx} (f * \theta_{2^j})(x) \end{aligned} \quad (3.16)$$

การแปลงเวฟเล็ต $W_{2^j} f(x)$ เป็นปฏิภาคโดยตรงกับอนุพันธ์อันดับที่ 1 ของฟังก์ชัน $f(x)$ ที่ถูกทำให้เรียบด้วย $\theta_{2^j}(x)$ ดังนั้นแอมพลิจูดสูงสุดของการแปลงเวฟเล็ต $|W_{2^j} f(x)|$ สัมพันธ์กับค่าสูง

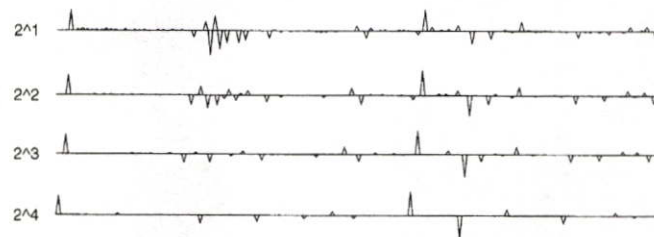
สุดของอนุพันธ์ของฟังก์ชัน $f * \theta_{2^j}(x)$ นั่นคือจุดที่มีการเปลี่ยนแปลงอย่างทันทีทันใดที่ระดับของการแปลงภาพ 2^j ตัวอย่างการหาค่าสูงสุดของการแปลงเวฟเล็ตแสดงดังรูปที่ 3.7(c) ค่าสูงสุดของการแปลงเวฟเล็ตจะระบุตำแหน่งของจุดที่เปลี่ยนแปลงอย่างทันทีทันใดของสัญญาณแต่ละระดับของการแปลงที่แตกต่างกันไป



(a)



(b)



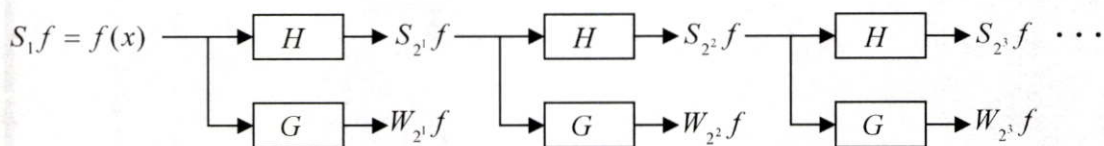
(c)

รูปที่ 3.7 ตัวอย่างการหาค่าสูงสุดของสัญญาณใน 1 มิติ (a) สัญญาณต้นแบบของภาพ LENA จำนวน 256 ค่า (b) องค์ประกอบของสัญญาณที่ระดับของการแปลง $1 \leq j \leq 4$ (c) ค่าแอมพลิจูดสูงสุดของ (b)

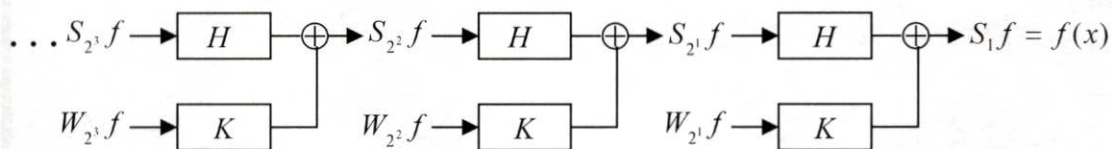
(2) อัลกอริธึมของการแปลงเวฟเล็ตใน 1 มิติ

สมมติเวฟเล็ต $\psi(x)$ มีลักษณะเช่นเดียวกับค่าสัมประสิทธิ์ของฟิลเตอร์ H G และ K ดังตารางที่ 3.3 และกำหนดให้ H_j G_j และ K_j เป็นค่าสัมประสิทธิ์ของฟิลเตอร์ที่เดิมศูนย์จำนวน $2^j - 1$ ค่าระหว่างค่าของฟิลเตอร์ H G และ K ดังนั้นทรานเฟอร์ฟังก์ชันของฟิลเตอร์เหล่านี้คือ $H(2^j \omega)$ $G(2^j \omega)$ และ $K(2^j \omega)$ ตามลำดับ กำหนดให้ $A * B$ เป็นการคอนโวลูชันของ

สัญญาณแบบดิสครีต A และ B อัลกอริทึมของการแปลงเวฟเล็ดแบบดิสครีตใน 1 มิติแสดงดังตารางที่ 3.5 และรูปที่ 3.8 ขั้นตอนการแยกองค์ประกอบเวฟเล็ดจะแยกสัญญาณ $S_{2^j}f$ ไปเป็นองค์ประกอบ $S_{2^{j+1}}f$ และ $W_{2^{j+1}}f$ ที่ระดับของการแปลงภาพ 2^j โดยแต่ละองค์ประกอบจะมีขนาดเดียวกับสัญญาณต้นแบบ ทำนองเดียวกันการรวมองค์ประกอบเวฟเล็ดของสัญญาณ $S_{2^{j-1}}f$ หาได้จากการรวมองค์ประกอบของ $S_{2^j}f$ และ $W_{2^j}f$ ที่ระดับของการแปลงภาพ 2^j



(a)



(b)

รูปที่ 3.8 การแปลงเวฟเล็ดแบบดิสครีตใน 1 มิติ (a) การแยกองค์ประกอบเวฟเล็ด (b) การรวมองค์ประกอบเวฟเล็ด

ตารางที่ 3.3 ผลตอบสนองอิมพัลส์ของ H , G , K และ L

n	H	G	K	L
-3			0.0078125	0.0078125
-2			0.054685	0.046875
-1	0.125		0.171875	0.1171875
0	0.375	-2.0	-0.171875	0.65625
1	0.375	2.0	-0.054685	0.1171875
2	0.125		-0.0078125	0.046875
3				0.0078125

ตารางที่ 3.4 ค่าสัมประสิทธิ์ของการนอร์มอลไลซ์และ $\lambda_j = 1$ เมื่อ $j > 5$

j	λ_j
1	1.50
2	1.12
3	1.03
4	1.01
5	1.00

ตารางที่ 3.5 อัลกอริทึมของการแปลงเวฟเล็ตแบบดิสครีตใน 1 มิติ

การแยกองค์ประกอบเวฟเล็ต	การรวมองค์ประกอบเวฟเล็ต
$j = 0$ <i>while</i> ($j < J$) $W_{2^{j+1}}f = \frac{1}{\lambda_j} \cdot S_{2^j}f * G_j$ $S_{2^{j+1}}f = S_{2^j}f * H_j$ $j = j + 1$ <i>end of while</i>	$j = 1$ <i>while</i> ($j > 0$) $S_{2^{j-1}}f = \lambda_j \cdot W_{2^j}f * K_{j-1} + S_{2^j}f * H_{j-1}$ $j = j - 1$ <i>end of while</i>

3.2.2 การหาอนุพันธ์ด้วยการแปลงเวฟเล็ตใน 2 มิติ

(1) คุณสมบัติทั่วไป

การแปลงเวฟเล็ตสามารถขยายต่อไปใน $L^2(R^2)$ โดยฟังก์ชันเวฟเล็ตมีเงื่อนไขดังนี้ $\Psi^k(x, y) \in L^2(R^2)$ ในที่นี้จะกล่าวถึงการแยกองค์ประกอบของเวฟเล็ต 2 มิติใน 2 ทิศทางเท่านั้น คือแนวนอน (Horizontal) และแนวตั้ง (Vertical) การแปลงเวฟเล็ตนี้จะกำหนดฟังก์ชันเวฟเล็ตจำนวน 2 ฟังก์ชันคือ $\Psi^1(x, y)$ และ $\Psi^2(x, y)$ เป็นดังนี้

$$\Psi^1(x, y) = \frac{\partial \Theta^1(x, y)}{\partial x} \quad \text{และ} \quad \Psi^2(x, y) = \frac{\partial \Theta^2(x, y)}{\partial y} \quad (3.17)$$

เมื่อ $\Theta^1(x, y)$ และ $\Theta^2(x, y)$ คือฟังก์ชันของความราบเรียบ ถ้ากำหนดให้

$$\Psi_{2^j}^1(x, y) = \frac{1}{2^j} \Psi^1\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \quad \text{และ} \quad \Psi_{2^j}^2(x, y) = \frac{1}{2^j} \Psi^2\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \quad (3.18)$$

การแปลงเวฟเล็ทของฟังก์ชัน $f(x, y) \in L^2(R^2)$ ที่ระดับของการแปลงภาพ 2^j ในทิศแนวนอน และแนวตั้งหาได้ดังนี้

$$W_{2^j}^1 f(x, y) = f * \Psi_{2^j}^1(x, y) \quad (3.19)$$

$$W_{2^j}^2 f(x, y) = f * \Psi_{2^j}^2(x, y) \quad (3.20)$$

เรียกการแปลงเวฟเล็ทใน 2 มิติของฟังก์ชัน $f(x, y)$ ว่าเซตของฟังก์ชัน

$$\left(W_{2^j}^1 f(x, y), W_{2^j}^2 f(x, y) \right)_{j \in \mathbb{Z}} \quad (3.21)$$

การเลือกฟังก์ชัน $\Theta^1(x, y)$ และ $\Theta^2(x, y)$ ที่มีลักษณะเฉพาะและฟังก์ชันเวฟเล็ทเป็นไปตามเงื่อนไขดังนี้

$$\sum_{j=-\infty}^{\infty} \left(\left| \hat{\Psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2 + \left| \hat{\Psi}^2(2^j \omega_x, 2^j \omega_y) \right|^2 \right) = 1 \quad (3.22)$$

ข้อกำหนดดังกล่าวจะทำให้การแปลงเวฟเล็ทใน 2 มิติมีคุณสมบัติเช่นเดียวกับการการแปลงเวฟเล็ทใน 1 มิติ ฟังก์ชัน $f(x, y)$ สามารถสร้างกลับคืนมาได้จากการรวมองค์ประกอบเวฟเล็ทของสัญญาณ

กำหนดให้ $\Theta_{2^j}^1(x, y) = \frac{1}{2^j} \Theta^1\left(\frac{x}{2^j}, \frac{y}{2^j}\right)$ และ $\Theta_{2^j}^2(x, y) = \frac{1}{2^j} \Theta^2\left(\frac{x}{2^j}, \frac{y}{2^j}\right)$ ดังนั้นสามารถเขียนสมการที่ (3.18) (3.19) และ (3.20) ได้ใหม่ดังนี้

$$W_{2^j}^1 f(x, y) = 2^j \frac{\partial}{\partial x} (f * \Theta_{2^j}^1)(x, y) \quad (3.23)$$

$$W_{2^j}^2 f(x, y) = 2^j \frac{\partial}{\partial y} (f * \Theta_{2^j}^1)(x, y) \quad (3.24)$$

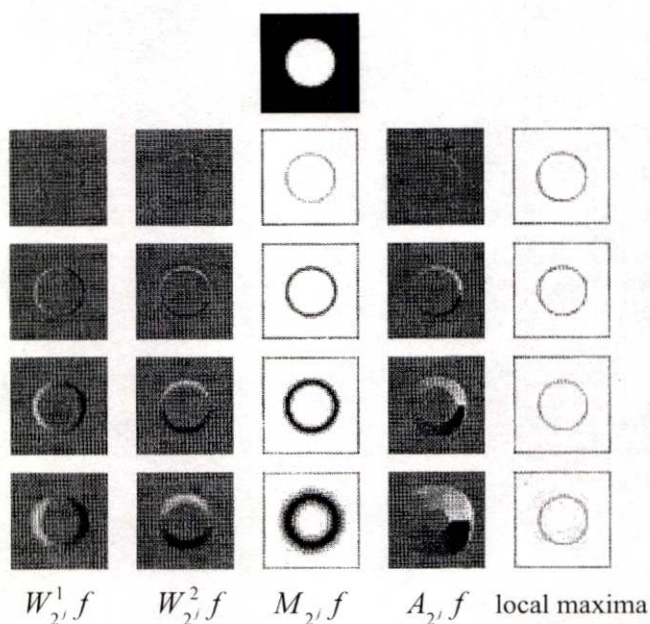
องค์ประกอบเวฟเล็ท $W_{2^j}^1 f(x, y)$ และ $W_{2^j}^2 f(x, y)$ คืออนุพันธ์อันดับย่อยของฟังก์ชัน $f(x, y)$ ในทิศทางแนวนอนและแนวตั้งที่ระดับของการแปลงภาพ 2^j ตามลำดับ หรือมองในลักษณะองค์ประกอบเกรเดียนท์เวกเตอร์ของฟังก์ชัน $f(x, y)$ ที่ทำให้เรียบด้วยฟังก์ชัน $\Theta_{2^j}^1(x, y)$ และ $\Theta_{2^j}^2(x, y)$ ค่ามอดูลัสของเกรเดียนท์เวกเตอร์ในแต่ละระดับของการแปลงภาพ 2^j หาได้จาก

$$M_{2^j} f(x, y) = \sqrt{\left| W_{2^j}^1 f(x, y) \right|^2 + \left| W_{2^j}^2 f(x, y) \right|^2} \quad (3.25)$$

มุมของเกรเดียนต์เวกเตอร์ในทิศทางแนวอนหาได้จาก

$$A_{2^j} f(x, y) = \text{argument} \left(\frac{W_{2^j}^1 f(x, y)}{W_{2^j}^2 f(x, y)} \right) \quad (3.26)$$

จุดที่มีการเปลี่ยนแปลงอย่างทันทีทันใดของฟังก์ชัน $f(x, y)$ ที่ระดับของการแปลงภาพ 2^j จะสัมพันธ์กับค่ามอดูลัสสูงสุดในทิศทางของเกรเดียนต์ การหาค่ามอดูลัสสูงสุดในข้อมูลภาพจะพิจารณาพิกเซลที่มีค่า $M_{2^j} f(x, y)$ มากกว่าค่ามอดูลัสของ 2 พิกเซลที่อยู่ติดกันในทิศทางเดียวกับมุม $A_{2^j} f(x, y)$ ตัวอย่างของการแปลงเวฟเล็ตของข้อมูลภาพต้นแบบ CIRCLE ดังรูปที่ 3.9(a) และแสดงองค์ประกอบเวฟเล็ต $W_{2^j}^{1,d} f$ และ $W_{2^j}^{2,d} f$ ที่ระดับของการแปลงภาพ $1 \leq j \leq 4$ ในคอลัมน์ที่ 1 และ 2 ของรูปที่ 3.9(b) สำหรับค่า $M_{2^j} f(x, y)$ และ $A_{2^j} f(x, y)$ แสดงในคอลัมน์ที่ 3 และ 4 ของรูปที่ 3.9(b) ตามลำดับ และคอลัมน์สุดท้ายของรูปที่ 3.9(b) แสดงตำแหน่งของค่ามอดูลัสสูงสุดของภาพที่ได้จากการแปลงเวฟเล็ต จะเห็นได้ว่าค่ามอดูลัสสูงสุดมีความสัมพันธ์กับส่วนที่เป็นขอบภาพหลักของข้อมูลภาพนั้น



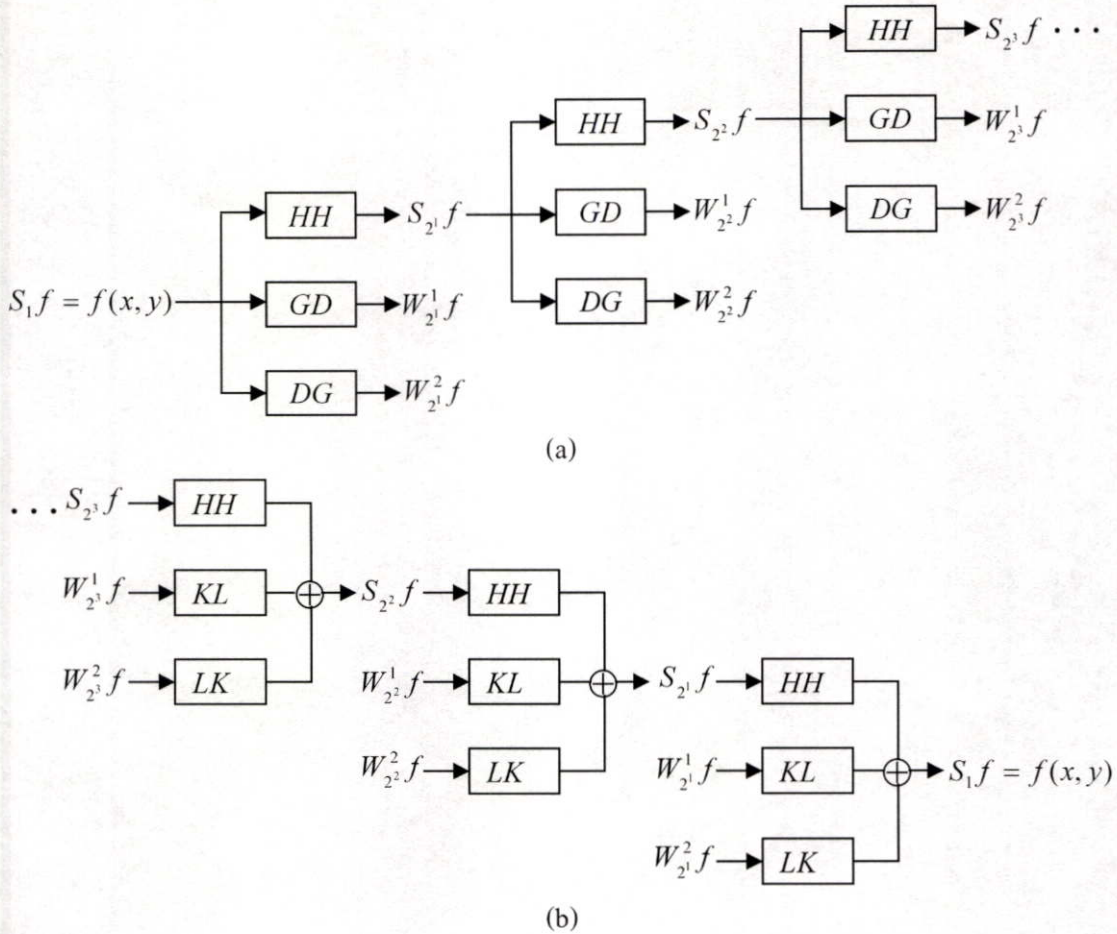
รูปที่ 3.9 การหาขอบของข้อมูลภาพต้นแบบ CIRCLE ขนาด 128x128 พิกเซล ที่ระดับของการแปลงภาพ $1 \leq j \leq 4$ (จากบนลงล่าง) สำหรับพิกเซลสีขาว สีเทา สีดำแสดงค่าลบ ศูนย์ และค่าบวกตามลำดับของ $W_{2^j}^1 f$ $W_{2^j}^2 f$ $M_{2^j} f$ $A_{2^j} f$ และตำแหน่งของค่ามอดูลัสสูงสุด (พิกเซลสีดำ)

(2) อัลกอริธึมของการแปลงเวฟเล็ตใน 2 มิติ

กำหนดฟังก์ชันเวฟเล็ต 2 ฟังก์ชันคือ $\psi^1(x, y)$ และ $\psi^2(x, y)$ มีลักษณะเช่นเดียวกับค่าสัมประสิทธิ์ฟิลเตอร์ของ H G K และ L ดังตารางที่ 3.3 และกำหนดให้ H_j G_j K_j และ L_j เป็นค่าสัมประสิทธิ์ฟิลเตอร์ที่เดิมศูนย์จำนวน $2^j - 1$ ค่าระหว่างค่าฟิลเตอร์ของ H G K และ L ดังนั้นทรานเฟอร์ฟังก์ชันของฟิลเตอร์เหล่านี้คือ $H(2^j \omega)$ $G(2^j \omega)$ $K(2^j \omega)$ และ $L(2^j \omega)$ ตามลำดับ และ D เป็นฟิลเตอร์ดิแรคที่มีผลตอบสนองอิมพัลส์เท่ากับ 1 ที่ $n = 0$ และเท่ากับ 0 ที่ n ใดๆ กำหนดให้ $A^*(H, L)$ เป็นการคอนโวลูชันของแฉกและคอดัมน์ของข้อมูลภาพ A ด้วยค่าฟิลเตอร์ H และ L ที่มีขนาด 1 มิติ อัลกอริธึมของการแปลงเวฟเล็ตแบบดิสครีตใน 2 มิติแสดงดังตารางที่ 3.6 และรูปที่ 3.10 ขั้นตอนการแยกองค์ประกอบเวฟเล็ตจะแยก $S_{2^j} f$ ไปเป็นองค์ประกอบของ $S_{2^{j+1}} f$ $W_{2^{j+1}}^1 f$ และ $W_{2^{j+1}}^2 f$ ที่ระดับของการแปลงภาพ 2^j โดยแต่ละองค์ประกอบนี้จะมีขนาดเท่ากับข้อมูลภาพต้นแบบ ทำนองเดียวกันการรวมองค์ประกอบเวฟเล็ตของ $S_{2^{j-1}} f$ หาได้จากการรวมองค์ประกอบ $S_{2^j} f$ $W_{2^j}^1 f$ และ $W_{2^j}^2 f$ ที่ระดับของการแปลงภาพ 2^j

ตารางที่ 3.6 อัลกอริธึมของการแปลงเวฟเล็ตแบบดิสครีตใน 2 มิติ

การแยกองค์ประกอบเวฟเล็ต	การรวมองค์ประกอบเวฟเล็ต
$j = 0$	$j = J$
$while(j < J)$	$while(j > 0)$
$W_{2^{j+1}}^{1,d} f = \frac{1}{\lambda_j} \cdot S_{2^j}^d f * (G_j, D)$	$S_{2^{j-1}}^d f = \lambda_j \cdot W_{2^j}^{1,d} f * (K_{j-1}, L_{j-1}) +$
$W_{2^{j+1}}^{2,d} f = \frac{1}{\lambda_j} \cdot S_{2^j}^d f * (D, G_j)$	$\lambda_j \cdot W_{2^j}^{2,d} f * (L_{j-1}, K_{j-1}) +$
$S_{2^{j+1}}^d f = S_{2^j}^d f * (H_j, H_j)$	$S_{2^j}^d f * (H_{j-1}, H_{j-1})$
$j = j + 1$	$j = j - 1$
$end\ of\ while$	$end\ of\ while$



รูปที่ 3.10 การแปลงเวฟเลตแบบดิสครีตใน 2 มิติ (a) การแยกองค์ประกอบเวฟเลต (b) การรวมองค์ประกอบเวฟเลต

3.2.3 การหาขอบภาพด้วยการแปลงเวฟเลต

ในที่นี้จะกล่าวถึงทฤษฎี Lipchitz exponent ที่เกี่ยวข้องกับการหาขอบภาพด้วยการแปลงเวฟเลตโดยตรง

ทฤษฎีที่ 1 กำหนดให้ $0 \leq \alpha \leq 1$ และ $K > 0$ ฟังก์ชัน $f(x)$ คือค่า Lipchitz α ที่มีลักษณะสม่ำเสมออยู่ในช่วง $]a, b[$ ดังนั้นที่ทุกๆ $x \in]a, b[$ มีเงื่อนไขดังนี้

$$|W_2^j f(x)| \leq K 2^{j\alpha} \tag{3.27}$$

ทฤษฎีที่ 2 กำหนดให้ $0 \leq \alpha \leq 1$ และ $K > 0$ ฟังก์ชัน $f(x, y)$ คือค่า Lipchitz α ที่มีลักษณะสม่ำเสมอในช่วงเซตเปิดของ R^2 ดังนั้นที่ทุกจุด (x, y) ของเซตเปิด R^2 มีเงื่อนไขดังนี้

$$M_2^j f(x, y) \leq K 2^{j\alpha} \tag{3.28}$$

ทฤษฎีเหล่านี้แสดงให้เห็นว่าค่า Lipchitz α หาได้จากค่าสูงสุดของการแปลงเวฟเล็ตที่ทุกระดับของการแปลงภาพ ตัวอย่างการหาขอบของภาพต้นแบบ LENA ขนาด 256×256 พิกเซล ดังรูปที่ 3.11(a) ที่ระดับของการแปลงภาพ $1 \leq j \leq 3$ ตำแหน่งของค่ามอดูลัสสูงสุดในทิศทางของเกรเดียนต์เวกเตอร์แสดงดังรูปที่ 3.11(b) และตำแหน่งของค่ามอดูลัสตามสมการที่ (3.28) เมื่อกำหนด $\alpha = 1$ และ $K = 5$ แสดงดังรูปที่ 3.11(c)



(a)

 $j = 1$  $j = 2$  $j = 3$

(b)

 $j = 1$  $j = 2$  $j = 3$

(c)

รูปที่ 3.11 ลักษณะขอบที่ระดับของการแปลงภาพ $1 \leq j \leq 3$ (a) ภาพต้นแบบ LENA (b) ตำแหน่งของค่ามอดูลัสสูงสุดในทิศทางของเกรเดียนต์เวกเตอร์ (c) พิกเซลสีขาวแสดงตำแหน่งของค่ามอดูลัสตามสมการที่ (3.28) เมื่อกำหนดค่า $\alpha = 1$ และ $K = 5$

3.3 สรุป

ในบทนี้จะกล่าวถึงการแปลงเวฟเล็ตแบบดิสครีต โดยพิจารณาจากฟิลเตอร์แบงก์และเงื่อนไขของเวฟเล็ตแบบออร์โธนอร์มัลและแบบไบออร์โธนอร์มัลที่นำไปใช้ในการทดลองและวิเคราะห์ผลการลดขอบบดือคของข้อมูลภาพด้วยวิธีการเวฟเล็ตเรโซลคิง นอกจากนี้ทฤษฎีของเวฟเล็ตได้พิสูจน์ให้เห็นว่าการหาตำแหน่งของค่าสูงสุดของการแปลงเวฟเล็ตในแต่ละระดับของการแปลงภาพสามารถหาขอบภาพได้และการพิจารณาค่า Lipchitz exponent สามารถแยกความแตกต่างของขอบภาพได้ ในวิทยานิพนธ์นี้จะใช้การหาขอบภาพด้วยค่า Lipchitz exponent เนื่องจากแสดงลักษณะของขอบภาพได้ดีกว่าการหาตำแหน่งค่าสูงสุดของการแปลงเวฟเล็ตในทิศทางของเกรเดียนท์และมีการคำนวณไม่ยุ่งยากซับซ้อน

บทที่ 4

การลดขอบบล็อคในข้อมูลภาพ

บทนี้จะกล่าวถึงการจำแนกประเภทวิธีการลดขอบบล็อค วิธีการลดขอบบล็อคด้วยการกรองความถี่ต่างๆ ได้แก่ การกรองความถี่ต่ำผ่าน การกรอง Wiener ที่ปรับตัวเองได้ เป็นต้น การทำเวฟเล็ทเชอร์ โซลคิงและเวฟเล็ทเชอร์ โซลคิงที่ปรับตัวเองได้ เภณท์ที่ใช้ในการวัดประสิทธิภาพของการลดขอบบล็อคคือค่าอัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) และค่าความแปรปรวนที่บริเวณขอบ (EV)

4.1 การจำแนกประเภทวิธีการลดขอบบล็อค

ได้มีการพัฒนาเทคนิคต่างๆ มากมายเพื่อลดขอบบล็อคข้อมูลภาพหรือที่เรียกว่า Deblocking สามารถจำแนกออกเป็น 2 ประเภทใหญ่ๆ ดังนี้

4.1.1 การใช้รูปแบบของการเข้ารหัสและถอดรหัสที่ต่างไปจากเดิม

ได้แก่ การสร้างค่าสัมประสิทธิ์ DC กลับคืนมาด้วยเทคนิคการประมาณค่าแบบ MAP (Maximum a posteriori) [12] การทำนายค่าสัมประสิทธิ์ DC โดยการหาค่าผลต่างกำลังสองเฉลี่ยของความชันหรือ MSDS (Mean Square Difference of Slope) ระหว่างพิกเซลที่อยู่ติดกันที่บริเวณขอบบล็อค [8] การเข้ารหัสโดยใช้การแปลงแบบร่วมกันหรือ CTC (Combined Transform Coding) [13] การใช้การแปลงแบบแลพออร์โธโกนัลหรือ LOT (Lapped orthogonal transform) [14] เป็นต้น

4.1.2 การประมวลผลหลังการขยายข้อมูลภาพ (Post-processing)

สามารถมองแนวทางการแก้ปัญหาเพื่อลดขอบบล็อคข้อมูลภาพได้ 2 แนวทางคือ การปรับปรุงภาพ (Image enhancement) และการสร้างภาพขึ้นมาใหม่จากข้อมูลภาพเดิม (Image restoration) อัลกอริทึมที่ใช้ในการปรับปรุงภาพนั้นมีวัตถุประสงค์เพื่อปรับปรุงคุณภาพของภาพที่มองเห็น โดยออกแบบวิธีการปรับปรุงภาพที่ขึ้นอยู่กับโครงสร้างความผิดเพี้ยนของภาพและการตอบสนองต่อการมองเห็นของมนุษย์ (Human visual sensitivity) ได้แก่ การประยุกต์ใช้งานที่เกี่ยวข้องกับการกรองสัญญาณที่บริเวณขอบบล็อคข้อมูลภาพ [3] เป็นต้น สำหรับอัลกอริทึมที่ใช้ในการสร้างภาพขึ้นมาใหม่จากข้อมูลภาพเดิมหรือปัญหาการสร้างข้อมูลภาพกลับคืนมา (Image recovery problem) จะใช้พื้นฐานของความรู้ความเข้าใจในรูปแบบความผิดเพี้ยนของภาพก่อนหน้า (Prior

knowledge) และสิ่งที่สังเกตเห็นได้จากการถอดรหัสข้อมูลภาพ เทคนิคของการสร้างภาพกลับคืนมาด้วยวิธีการแบบดั้งเดิม ได้แก่ วิธีการ POCS (Projection onto convex set) [1] วิธีการกำหนดค่ากำลังสองที่น้อยที่สุดหรือ CLS (Constrained least square) [2] เป็นต้น แต่เทคนิคของการสร้างภาพกลับคืนมามีขั้นตอนการคำนวณซ้ำที่ต้องอาศัยผลที่ได้จากการคำนวณก่อนหน้านี้หรือที่เรียกว่ารีเคอร์ซีฟ (Recursive) และใช้การคำนวณที่มาก

กรณีประเภทแรกนั้นจะต้องมีรูปแบบการเข้ารหัสเฉพาะของตนเองนั่นคือจะไม่ยอมรับการบีบอัดข้อมูลภาพที่เป็นมาตรฐานเช่น JPEG MPEG และ H.261 เป็นต้น ดังนั้นการลดขอบบดื้อคด้วยเทคนิคการประมวลผลหลังการขยายข้อมูลภาพโดยใช้เทคนิคการปรับปรุงภาพจะมีประสิทธิภาพที่ดีกว่าในการที่จะพัฒนาไปสู่การสื่อสารด้านภาพและวิดีโอ เนื่องจากเทคนิคดังกล่าวไม่ขึ้นอยู่กับรูปแบบของการเข้ารหัสและถอดรหัส (Codec coding) และให้คุณภาพของภาพที่ได้ดีกว่าที่บีบอัดเหมือนกัน ในวิทยานิพนธ์นี้ได้ทำการศึกษาปัญหาการปรับปรุงภาพเพื่อทำการลดขอบบดื้อคข้อมูลภาพในมุมมองใหม่ นั่นคือมองปัญหาการลดขอบบดื้อคเช่นเดียวกับปัญหาการกำจัดสัญญาณรบกวนหรือที่เรียกว่า Denoising โดยทั่วไปวิธีการกำจัดสัญญาณรบกวนคือการทำสัญญาณให้เรียบ (Smoothing) และจากแนวคิดที่ว่าข้อมูลภาพส่วนใหญ่มีระดับพลังงานของสัญญาณที่ค่าสัมประสิทธิ์ของการแปลงจำนวนหนึ่ง ส่วนสัญญาณรบกวนกระจายอยู่ทั่วไปที่ความถี่สูงและเป็นค่าสัมประสิทธิ์ที่ไม่สำคัญ วิธีการกำจัดสัญญาณรบกวนที่มาตรฐานอย่างหนึ่งคือการกรอง Wiener ที่ปรับตัวเองได้ [15] ซึ่งสามารถชดเชยสเปกตรัมที่ความถี่สูงแต่ยังคงตัดรายละเอียดของข้อมูลภาพบางส่วนทิ้งไป ดังนั้นเทคนิคของการกำจัดสัญญาณรบกวนในโดเมนเวฟเล็ตจึงเป็นทางเลือกอย่างหนึ่งที่น่าสนใจและมีความเหมาะสมเป็นอย่างมาก

ในปัจจุบันวิธีการกำจัดสัญญาณรบกวนที่ได้รับความสนใจเป็นอย่างมากคือเวฟเล็ตเรโซลคิง ซึ่งเป็นวิธีการกำจัดสัญญาณรบกวนที่ขึ้นอยู่กับค่าเรโซลคิง วิธีการนี้เป็นวิธีการที่ง่ายและมีประสิทธิภาพ Donoho [16] ได้แสดงให้เห็นว่าเวฟเล็ตเรโซลคิงมีคุณสมบัติที่เหมาะสมในการกำจัดสัญญาณรบกวน โดยค่าสัมประสิทธิ์ที่ไม่สำคัญและขึ้นอยู่กับค่าเรโซลคิงจะเปรียบได้กับสัญญาณรบกวน ขณะที่ค่าสัมประสิทธิ์ที่สำคัญจะเปรียบได้กับโครงสร้างหลักของสัญญาณ Gopinath et al. [17] ได้นำวิธีการของ Donoho มาใช้ในขั้นตอนการประมวลผลหลังการขยายข้อมูลภาพแบบ JPEG และให้ผลภาพที่ดีขึ้น แต่เนื่องจากสัญญาณรบกวนสามารถสังเกตเห็นได้ในแต่ละพื้นที่ของข้อมูลภาพมีลักษณะแตกต่างกันนั้นคือบริเวณของภาพที่ราบเรียบจะสังเกตเห็นสัญญาณรบกวนได้มากกว่าบริเวณของภาพที่เป็นขอบภาพหรือเนื้อหาของภาพ ดังนั้นวิธีการทำเวฟเล็ตเรโซลคิงควรจะปรับเปลี่ยนตามลักษณะพื้นที่ของข้อมูลภาพเพื่อสามารถกำจัดสัญญาณรบกวนได้อย่างมีประสิทธิภาพมากขึ้น ในเนื้อหาของบทนี้จะกล่าวถึงวิธีการกรองความถี่ต่ำ การกรอง Wiener ที่ปรับตัวเองได้ เวฟเล็ตเรโซลคิงและเวฟเล็ตเรโซลคิงที่ปรับเปลี่ยนได้ตามลักษณะพื้นที่ของข้อมูลภาพ รวมทั้งการวัดประสิทธิภาพที่ได้จากการลดขอบบดื้อคข้อมูลภาพ

4.2 การกรองความถี่

4.2.1 การกรองความถี่ต่ำ

การกรองความถี่ต่ำผ่านได้กำหนดให้ค่าสัมประสิทธิ์ของฟิลเตอร์มีขนาด 3×3 พิกเซล โดยกระทำที่บริเวณขอบบล็อคข้อมูลภาพ [3] อย่างไรก็ตามวิธีการนี้ยังคงเห็นลักษณะของความไม่ต่อเนื่องของบล็อค ดังนั้นเพื่อให้เห็นความเป็นไปได้ของการลดขอบบล็อคด้วยการกรองสัญญาณแบบเชิงเส้น (Linear filtering) จะใช้ค่าสัมประสิทธิ์ของฟิลเตอร์เท่ากับ $1/16$ $[1 \ 2 \ 1; 2 \ 4 \ 2; 1 \ 2 \ 1]$ โดยกระทำกับทุกพิกเซลตลอดทั้งภาพ

4.2.2 การกรอง Wiener ที่ปรับตัวเองได้

การกรอง Wiener ที่ปรับตัวเองได้มีคุณลักษณะของการกรองความถี่ต่ำผ่าน ถ้ากำหนดให้ค่าสัมประสิทธิ์ของฟิลเตอร์เปลี่ยนแปลงได้ตามลักษณะพื้นที่ของภาพ (Space-variant filter) และ $h(n_1, n_2)$ คือค่าสัมประสิทธิ์ของฟิลเตอร์ที่ใช้สำหรับพื้นที่ใดๆ ในข้อมูลภาพเนื่องจากค่าสัมประสิทธิ์ของฟิลเตอร์ที่มีค่ามากจะไม่มีผลต่อพื้นที่ของภาพที่มีรายละเอียดน้อย (Low detail region) แต่จะมีผลต่อพื้นที่ของภาพที่มีรายละเอียดมาก (High detail region) เช่น ทำให้บริเวณขอบของภาพเบลอ เป็นต้น (รายละเอียดกล่าวไว้ในภาคผนวก ก) ค่าสัมประสิทธิ์ของฟิลเตอร์ $h(n_1, n_2)$ หาได้จาก

$$h(n_1, n_2) = \begin{cases} \sigma_f^2 + \frac{\sigma_v^2}{(2M+1)^2} \\ \frac{\sigma_f^2 + \sigma_v^2}{\sigma_f^2 + \sigma_v^2}, n_1 = n_2 = 0 \\ \frac{\sigma_v^2}{(2M+1)^2} \\ \frac{\sigma_v^2}{\sigma_f^2 + \sigma_v^2}, -M \leq n_1 \leq M, -M \leq n_2 \leq M \quad \text{except } n_1 = n_2 = 0 \\ 0, \quad \text{otherwise} \end{cases} \quad (4.1)$$

เมื่อ σ_f^2 คือค่าความแปรปรวนของข้อมูลภาพ $f(x, y)$ ที่บริเวณใดๆ

σ_v^2 คือค่าความแปรปรวนของความผิดพลาดในกรณีนี้สมมุติว่าทราบค่า

M คือเลขจำนวนเต็มยกเว้นศูนย์

4.3 เวฟเล็ตเชอร์โซลดิ้ง

ถ้าให้ y คือข้อมูลภาพที่สร้างขึ้นใหม่จากข้อมูลภาพต้นแบบ x รูปแบบของปัญหาในการปรับปรุงภาพ สามารถเขียนเป็นสมการได้ดังนี้ [17]

$$y = x + e \quad (4.2)$$

เมื่อ e คือค่าความผิดพลาดของการสร้างข้อมูลภาพขึ้นใหม่หรือความผิดพลาดของความผิดพลาดของข้อมูลภาพ y มีความสัมพันธ์กับขั้นตอนการบีบอัดข้อมูลภาพที่บิตเรตต่างๆ วัตถุประสงค์ของการปรับปรุงภาพเพื่อต้องการประมาณข้อมูลภาพต้นแบบ $\hat{x}(y)$ ที่ได้กำจัดค่าความผิดพลาด e ออกจากข้อมูลภาพ y โดยที่ยังคงรักษาลักษณะที่สำคัญของข้อมูลภาพต้นแบบ x ไว้

จากแนวคิดพื้นฐานของ Donoho ในการทำเวฟเล็ตเชอร์โซลดิ้ง สมมติให้ e คือสัญญาณรบกวนแบบเกาส์เซียน (Gaussian noise) และต้องการหา $\hat{x}(y)$ จากข้อมูลภาพที่สร้างขึ้นใหม่ y ที่ทำให้ค่า $\|\hat{x}(y) - x\|$ มีค่าน้อยที่สุดหรือหาค่า $\hat{x}(y)$ ที่มีค่าใกล้เคียงกับภาพต้นแบบ x มากที่สุด ถ้ากำหนดให้ W และ W^{-1} คือการแปลงเวฟเล็ตแบบออร์โธโกนัล (Orthogonal wavelet transform) และการแปลงเวฟเล็ตกลับตามลำดับ ดังนั้นจากสมการที่ (4.2) สามารถเขียนใหม่ได้ว่า

$$Y = X + E \quad (4.3)$$

อักษรตัวพิมพ์ใหญ่แสดงค่าตัวแปรในโดเมนของการแปลงข้อมูลนั้นคือ $Y = Wy$ ให้ \hat{X} คือค่าการประมาณของ X ที่ได้จากการทำเวฟเล็ตเชอร์โซลดิ้งของ Y ดังนั้นขั้นตอนของการทำเวฟเล็ตเชอร์โซลดิ้งเขียนสมการได้ดังนี้

$$\hat{x}(y) = W^{-1} \hat{X} = W^{-1} T_\delta(Wy) \quad (4.4)$$

เมื่อ T คือรูปแบบของการทำเวฟเล็ตเชอร์โซลดิ้งและ δ คือค่าเวฟเล็ตเชอร์โซลดิ้งที่ได้จากการประมาณค่า Y สิ่งที่สำคัญของการทำเวฟเล็ตเชอร์โซลดิ้งคือการกำจัดสัญญาณรบกวนหรือความผิดพลาดของข้อมูลภาพและยังคงรักษาลักษณะที่สำคัญของข้อมูลภาพต้นแบบไว้

4.3.1 รูปแบบการทำเวฟเล็ตเชอร์โซลดิ้ง [18]

(1) ฮาร์ดเชอร์โซลดิ้ง (Hard thresholding) วิธีการนี้เป็นการเก็บหรือตัดค่าสัมประสิทธิ์เวฟเล็ตทิ้งไป (Keep or kill) ค่าสัมบูรณ์ของค่าสัมประสิทธิ์เวฟเล็ตทุกค่าจะถูกนำมาเปรียบเทียบกับค่า

เชอร์โซลด์ที่กำหนด (δ) ถ้าขนาดของค่าสัมประสิทธิ์เวฟเล็ทมีค่าน้อยกว่าค่าเชอร์โซลด์ δ ดังนั้นค่าสัมประสิทธิ์เวฟเล็ทจะถูกแทนที่ด้วยศูนย์

$$T_H(Y, \delta) = \begin{cases} Y & , |Y| \geq \delta \\ 0 & , |Y| < \delta \end{cases} \quad (4.5)$$

ฟังก์ชันของการทำฮาร์ดเชอร์โซลด์แสดงดังรูปที่ 4.1(a) ฮาร์ดเชอร์โซลด์จะถูกใช้ในการเข้ารหัสด้วยการแปลงเวฟเล็ทเพื่อให้ความยาวของข้อมูลสั้นที่สุดและลำดับของข้อมูลที่เป็นศูนย์ที่ได้จากการเชอร์โซลด์จะทำให้การเข้ารหัสมีประสิทธิภาพมากยิ่งขึ้น

(2) ซอฟท์เชอร์โซลด์ (Soft thresholding) เป็นวิธีการที่ค่าสัมประสิทธิ์เวฟเล็ททุกค่าขึ้นอยู่กับค่าเริ่มต้นของข้อมูล ฟังก์ชันของการทำซอฟท์เชอร์โซลด์แสดงดังรูปที่ 4.1(b) สมการของซอฟท์เชอร์โซลด์หาได้จาก

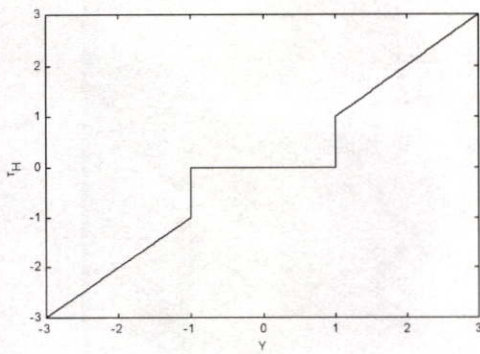
$$T_S(Y, \delta) = \begin{cases} \text{sgn}(Y)(|Y| - \delta) & , |Y| \geq \delta \\ 0 & , |Y| < \delta \end{cases} \quad (4.6)$$

(3) ควอนไทล์เชอร์โซลด์ (Quantile thresholding) หาได้จาก

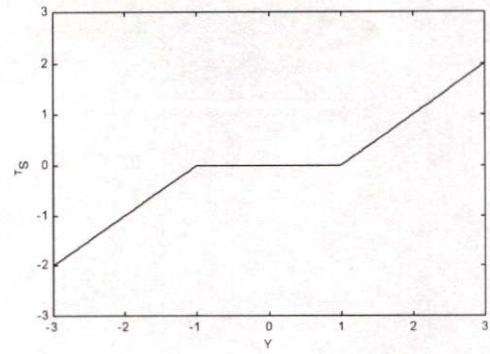
$$T_Q(Y, p) = \begin{cases} Y & : Y \geq p \\ 0 & : Y < p \end{cases} \quad (4.7)$$

โดยที่ p คือ p -quantile ของค่าสัมประสิทธิ์เวฟเล็ททุกค่า เช่น การแทนที่ค่าสัมประสิทธิ์เวฟเล็ทด้วยศูนย์จำนวน 30 % ของค่าสัมประสิทธิ์เวฟเล็ทที่น้อยที่สุด เป็นต้น

(4) ยูนิเวอร์แซลเชอร์โซลด์ (Universal thresholding) Donoho ได้กำหนดค่าเชอร์โซลด์ $\delta = \sigma\sqrt{2\log n}/\sqrt{n}$ ของกลุ่มข้อมูลที่เป็นค่าสัมประสิทธิ์เวฟเล็ท เมื่อ n คือขนาดของข้อมูลและ σ คือค่าเบี่ยงเบนมาตรฐานของสัญญาณรบกวนที่สเกลใดๆ ยูนิเวอร์แซลเชอร์โซลด์สามารถเป็นได้ทั้งฮาร์ดเชอร์โซลด์และซอฟท์เชอร์โซลด์



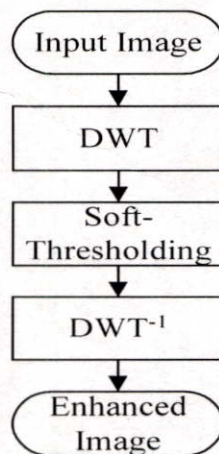
(a)



(b)

รูปที่ 4.1 รูปแบบของการทำแธรโซลด์คิงที่ค่า $\delta = 1$ (a) ฮาร์ดแธรโซลด์คิง (b) ซอฟท์แธรโซลด์คิง

ในระยะเริ่มแรกนั้นวิธีการของเวฟเล็ทแธรโซลด์คิงได้นำมาใช้สำหรับการกำจัดสัญญาณรบกวนแบบเกาส์เซียน Donoho [16] ได้แสดงให้เห็นว่าถ้าค่าความผิดพลาดหรือสัญญาณรบกวนมีขอบเขตที่แน่นอนจะได้ว่าวิธีการทำซอฟท์แธรโซลด์คิงเป็นการกำจัดสัญญาณรบกวนที่เหมาะสมที่สุด วิธีการนี้เริ่มจากการแปลงข้อมูลภาพที่สร้างกลับมาให้อยู่ในโดเมนของเวฟเล็ทดังรูปที่ 3.5 จากนั้นทำการแธรโซลด์ค่าสัมประสิทธิ์เวฟเล็ท โดยค่าสัมประสิทธิ์ที่มีค่าต่ำกว่าค่าแธรโซลด์จะถูกตัดทิ้งไป ขั้นตอนสุดท้ายทำการแปลงเวฟเล็ทกลับจะได้ข้อมูลภาพที่ถูกกำจัดสัญญาณรบกวนแล้วสามารถแสดงขั้นตอนการทำงานได้ดังรูปที่ 4.2 โดยปกติสัญญาณรบกวนมีการกระจายอย่างสม่ำเสมอ (Uniform distribution) ที่ทุกระดับของการแปลงเวฟเล็ท (Level) และเห็นได้อย่างชัดเจนที่แบนด์ย่อยของความถี่สูง (High-frequency subbands) ดังนั้นขั้นตอนการทำซอฟท์แธรโซลด์คิง จะกระทำในทุกแบนด์ย่อยยกเว้นที่แบนด์ย่อยของความถี่ต่ำ (Low-frequency subbands)



รูปที่ 4.2 อัลกอริทึมของการทำเวฟเล็ทแธรโซลด์คิง

อย่างไรก็ตามความผิดพลาดของการเข้ารหัสไม่ได้เป็นสัญญาณรบกวนแบบเกาส์เซียน แต่ข้อมูลภาพที่ผ่านการบีบอัดยังคงรักษาลักษณะที่สำคัญของภาพไว้ เช่น ขอบภาพ เป็นต้น และความผิดพลาดของการเข้ารหัสเป็นข้อมูลที่ไม่สำคัญ ในการแปลงเวฟเล็ตจะเก็บค่าพลังงานในส่วนที่สำคัญของภาพไว้ด้วยค่าสัมประสิทธิ์เวฟเล็ตที่มีขนาดใหญ่ (Magnitude) และค่าพลังงานของความผิดพลาดของการเข้ารหัสมีค่าสัมประสิทธิ์เวฟเล็ตที่มีขนาดเล็กกระจายอยู่ทั่วไป ดังนั้นลักษณะที่สำคัญของภาพและความผิดพลาดของการเข้ารหัสสามารถแยกออกจากกันได้ดีในโดเมนเวฟเล็ต โดยที่ค่าสัมประสิทธิ์เวฟเล็ตที่มีค่ามากจะแสดงถึงข้อมูลภาพต้นแบบและค่าสัมประสิทธิ์เวฟเล็ตที่มีค่าที่น้อยจะแสดงถึงความผิดพลาดของการเข้ารหัสข้อมูลภาพ โดยธรรมชาติแล้วการเก็บค่าสัมประสิทธิ์ที่มีขนาดใหญ่และตัดค่าสัมประสิทธิ์ที่มีขนาดเล็กทิ้งไป (กำหนดให้มีค่าเป็นศูนย์) หรือเป็นการทำเรโซลต์จึงจะสามารถลดค่าพลังงานของความผิดพลาดเนื่องจากการเข้ารหัสข้อมูลได้

4.3.2 การเลือกชนิดและระดับของการแปลงเวฟเล็ต

ในที่นี้ได้เลือกใช้ฟิลเตอร์ของเวฟเล็ต 4 ตระกูล ได้แก่ Daubechies Symlet Coiflet และ Biorthogonal ฟิลเตอร์ของเวฟเล็ตในแต่ละตระกูลจะมีความยาวของฟิลเตอร์ที่แตกต่างกัน (Filter length) และได้เลือกค่าความยาวของฟิลเตอร์ในแต่ละตระกูลที่สามารถลดขอบบลิ๊อคข้อมูลภาพได้ผลดีที่สุดเพื่อนำมาเปรียบเทียบระหว่างตระกูลเวฟเล็ตต่อไป โดยทำการทดลองกับข้อมูลภาพหลายระดับเทาขนาด 256x256 พิกเซลที่ผ่านขั้นตอนการบีบอัดข้อมูลภาพแบบ JPEG ในบิตเรทต่างๆ จำนวนหลายภาพและวัดประสิทธิภาพของผลภาพที่ได้จากค่า PSNR และ EV จะได้เวฟเล็ตแต่ละชนิดเป็น ดังนี้ Daubechies-10 (Db10) Symlet-6 (Sym6) Coiflet-5 (Coif5) Biorthogonal (Bior 4.4) ซึ่งความยาวของฟิลเตอร์ของเวฟเล็ตแต่ละชนิดเท่ากับ 20 12 30 และ 8 ตามลำดับ กำหนดให้ค่าระดับของการแปลงเวฟเล็ตมีค่าเท่ากับ 3 ($L = 3$) เนื่องจากค่าระดับของการแปลงเวฟเล็ตที่มีค่ามากกว่า 3 นั้น จะมีผลต่อการกำจัดสัญญาณรบกวนเพียงเล็กน้อยเท่านั้นและยังใช้การคำนวณที่มากอีกด้วย

4.3.3 การประมาณค่าเรโซลต์

สามารถประมาณค่าเรโซลต์ได้อย่างง่ายด้วยการคำนวณค่าเบี่ยงเบนมาตรฐาน (σ) ในภาพแบนด์ย่อยแต่ละภาพที่ได้จากการแปลงเวฟเล็ต ดังรูปที่ 3.6 ค่าเรโซลต์หาได้จาก

$$\delta_{b,l}(e) = f_a \sigma_{b,l} = f_a \sqrt{\frac{1}{N-1} \sum (f_{b,l} - m)^2} \quad (4.8)$$

เมื่อ N คือจำนวนข้อมูลทั้งหมดในแต่ละแบนด์ย่อยของ $f_{b,l}$

m คือค่าเฉลี่ยของสัมประสิทธิ์เวฟเล็ตในแต่ละแบนด์ย่อยของ $f_{b,l}$

b คือแบนด์ย่อยของความถี่สูง (HH HL และ LH)

l คือระดับของการแปลงเวฟเล็ต ($l=1,2,\dots,L$)

f_a คือแฟกเตอร์ของค่าเรโซลต์ กำหนดให้มีค่าเท่ากับ $1/l$

4.4 เวฟเล็ตเรโซลต์ดิ่งที่ปรับตัวเองได้

4.4.1 การจำแนกลักษณะพื้นที่ของข้อมูลภาพ

(1) การหาพื้นที่ขอบภาพ (Edge region)

การหาขอบที่หลายระดับของการแปลงภาพหรือการหามอดูลัสสูงสุดของการแปลงเวฟเล็ตได้กล่าวรายละเอียดไว้แล้วในบทที่ 3 แนวคิดที่สำคัญของการหาขอบภาพจะกล่าวสรุปในที่นี้ เวฟเล็ตที่ใช้ในการหาขอบภาพดังรูปที่ 3.10 จะไม่มีการลดจำนวนข้อมูลลง (Nonsubsampling) เพื่อแสดงเกรเดียนท์ของภาพในระดับของการแปลงภาพที่แตกต่างกัน สมมุติให้ $f(x, y)$ เป็นข้อมูลภาพที่ได้จากการบีบอัดข้อมูลภาพแบบ JPEG ที่มีขนาด $M \times N$ พิกเซล ดังนั้นในแต่ละระดับของการแปลงภาพ 2^j เมื่อ $j > 0$ และ $S_1 f = f(x, y)$ สามารถแยกเป็นองค์ประกอบย่อยจำนวน 3 แบนด์ดังนี้ องค์ประกอบความถี่ต่ำ $S_{2^j} f$ องค์ประกอบความถี่สูง $W_{2^j}^1 f$ ในทิศทางแนวนอนและองค์ประกอบความถี่สูง $W_{2^j}^2 f$ ในทิศทางแนวตั้ง ดังรูปที่ 3.10(a) ค่าสัมประสิทธิ์ของเวฟเล็ตทั้ง 3 แบนด์ ($S_{2^j} f, W_{2^j}^1 f, W_{2^j}^2 f$) ที่ระดับของการแปลงภาพ 2^j จะมีขนาดเท่ากับ $M \times N$ พิกเซล ซึ่งมีขนาดเดียวกับข้อมูลภาพที่ได้จากการบีบอัดข้อมูลภาพแบบ JPEG และค่าสัมประสิทธิ์ของฟิลเตอร์ที่ใช้ในแต่ละระดับของการแปลงภาพ 2^j เมื่อ $j > 0$ จะเพิ่มค่าศูนย์จำนวน $2^j - 1$ ค่าระหว่างค่าสัมประสิทธิ์ของฟิลเตอร์ที่ระดับของการแปลงภาพ 2^j เมื่อ $j = 0$ สามารถหาค่ามอดูลัสของเกรเดียนท์เวกเตอร์ในแต่ละระดับของการแปลงภาพ 2^j ได้จาก

$$M_{2^j} f(x, y) = \sqrt{|W_{2^j}^1 f(x, y)|^2 + |W_{2^j}^2 f(x, y)|^2} \quad (4.9)$$

จากนั้นหาตำแหน่งของขอบในแต่ละระดับของการแปลงภาพ 2^j ด้วยค่า Lipchitz α นั่นคือถ้าข้อมูลภาพ $f(x, y)$ มีขอบภาพที่ตำแหน่ง (x_0, y_0) ดังนั้นที่จุด (x, y) ใดๆ ที่อยู่ติดกับ (x_0, y_0) จะเป็นไปตามเงื่อนไขของสมการที่ (4.10)

$$M_{2^j} f(x, y) \leq K 2^{j\alpha} \quad (4.10)$$

เมื่อ K คือค่าคงที่ที่มีค่ามากกว่าศูนย์และ $\alpha \in [0,1]$ คุณสมบัตินี้สามารถแยกส่วนที่เป็นขอบภาพและส่วนที่ไม่ใช่ขอบภาพ (Non edge region) ออกจากกันได้ที่ทุกระดับของการแปลงภาพด้วยเวฟเล็ต

(2) พื้นที่ของเนื้อหาและส่วนที่ราบเรียบของภาพ (Texture and smooth region)

ทำการจำแนกพิกเซลของข้อมูลภาพในส่วนที่ไม่ใช่ขอบภาพออกเป็นพื้นที่ของเนื้อหาและส่วนที่ราบเรียบของภาพ การจำแนกข้อมูลภาพเหมือนกับการทำเซกเมนเตชันของข้อมูลภาพนั่นเองและสามารถทำในสเปซเฮลโดเมนได้ดีกว่าโดเมนของการแปลงข้อมูล ในที่นี้สมมุติว่าภาพทุกภาพมีเนื้อหาของภาพด้วยค่าระดับที่เหมาะสมค่าหนึ่ง การหาพื้นที่ในส่วนที่เป็นเนื้อหาของภาพและส่วนที่ราบเรียบของภาพสามารถหาได้จากการกำหนดช่วงของระดับความละเอียด (Coarseness scale) Rosenfeld และ Troy [19] ได้เสนอวิธีการวัดค่าความละเอียดของภาพที่ง่าย รวดเร็วและมีประสิทธิภาพ โดยการหาตำแหน่งของพิกเซลที่มีค่าสูงสุดและต่ำสุด (Local extrema) ตลอดทั้งแถวและคอลัมน์ ความหนาแน่นของค่าสูงสุดและต่ำสุดของภาพ (Density of local extrema) สามารถใช้วัดระดับความละเอียดในเนื้อหาของภาพได้

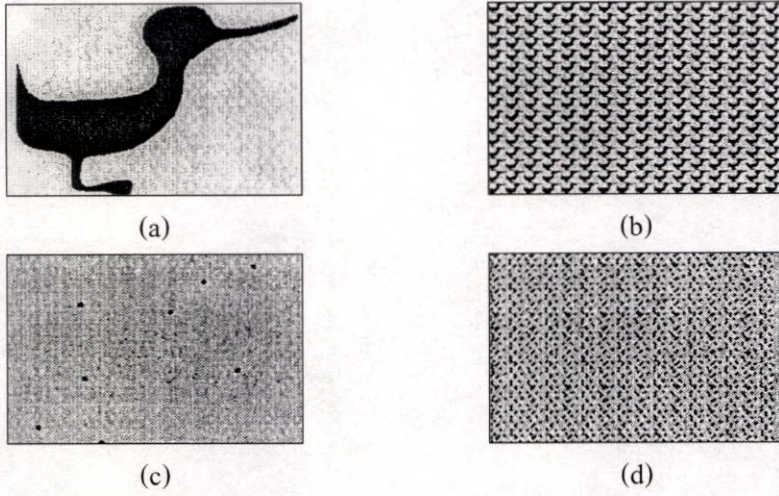
กำหนดให้ $I(r, c)$ คือฟังก์ชันข้อมูลภาพของพิกเซลในแถวที่ r และคอลัมน์ที่ c ให้พิกเซลที่ (r_0, c_0) คือตำแหน่งของพิกเซลที่มีค่าสูงสุดในแถวนั้นๆ ถ้า

$$I(r_0, c_0 - 1) < I(r_0, c_0) > I(r_0, c_0 + 1) \quad (4.11)$$

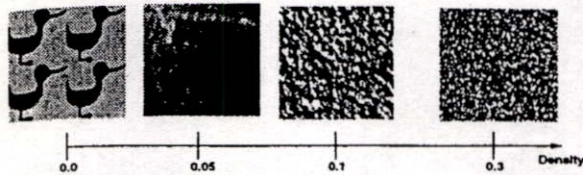
สำหรับตำแหน่งของค่าต่ำสุดในแถว ค่าต่ำสุดและค่าสูงสุดในคอลัมน์สามารถหาได้ด้วยวิธีการเช่นเดียวกัน พิกเซลที่ตำแหน่ง (r_0, c_0) คือพิกเซลที่มีค่าสูงสุดและค่าต่ำสุด (Local extremum) ถ้าตำแหน่งของพิกเซลนั้นมีค่าสูงสุดหรือค่าต่ำสุดในแถว (Local row extremum) และในคอลัมน์ (Local column extremum) แสดงตัวอย่างดังรูปที่ 4.3

ระดับความละเอียดของพิกเซลใดๆ สามารถคำนวณได้จากความหนาแน่นของค่าสูงสุดและค่าต่ำสุดในวินโดว์ขนาด $s \times s$ พิกเซลของข้อมูลภาพ ระดับความละเอียดจะมีค่าตั้งแต่ 0 (ไม่มีพิกเซลใดมีค่าสูงสุดและค่าต่ำสุด) จนถึง 1 (ทุกๆ พิกเซลมีค่าสูงสุดและค่าต่ำสุด) ค่าระดับความละเอียดสามารถกำหนดได้ว่าพื้นที่ใดเป็นเนื้อหาของภาพหรือส่วนที่ราบเรียบของภาพ พิกเซลที่เป็นเนื้อหาของภาพจะมีค่าความหนาแน่นของค่าสูงสุดและค่าต่ำสุดอยู่ระหว่างขอบเขตของค่าบนและค่าล่าง (Upper and lower bound) ของระดับความละเอียดแสดงตัวอย่างดังรูปที่ 4.4 และ 4.5 จะเห็นว่าข้อมูลภาพที่มีรายละเอียดมากจะมีระดับความละเอียดหรือความหนาแน่นของค่าสูงสุดและค่าต่ำสุดสูงกว่าข้อมูลภาพที่มีรายละเอียดน้อย ขนาดของวินโดว์ที่เหมาะสมขึ้นอยู่กับชนิดของภาพ

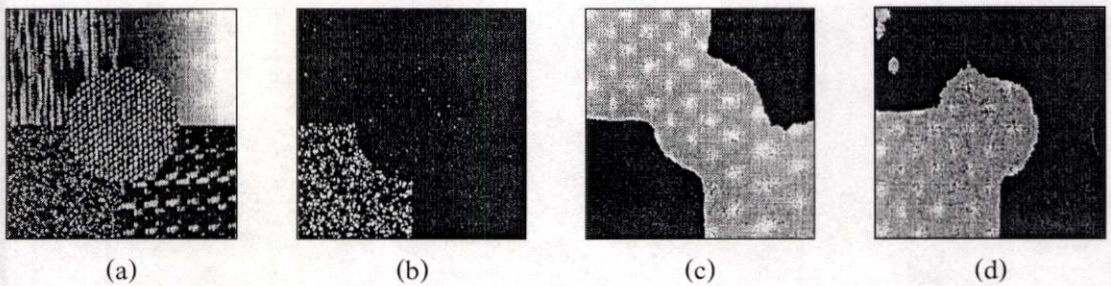
โดยวินโดว์ขนาดเล็กใช้สำหรับเนื้อหาของภาพที่มีรายละเอียดมากและวินโดว์ขนาดใหญ่ใช้สำหรับเนื้อหาของภาพที่มีรายละเอียดน้อย



รูปที่ 4.3 ภาพ (c) และ (d) แสดงตำแหน่งพิกเซลที่มีค่าสูงสุดและค่าต่ำสุด (พิกเซลสีดำ) ของภาพ (a) และ (b) ตามลำดับ



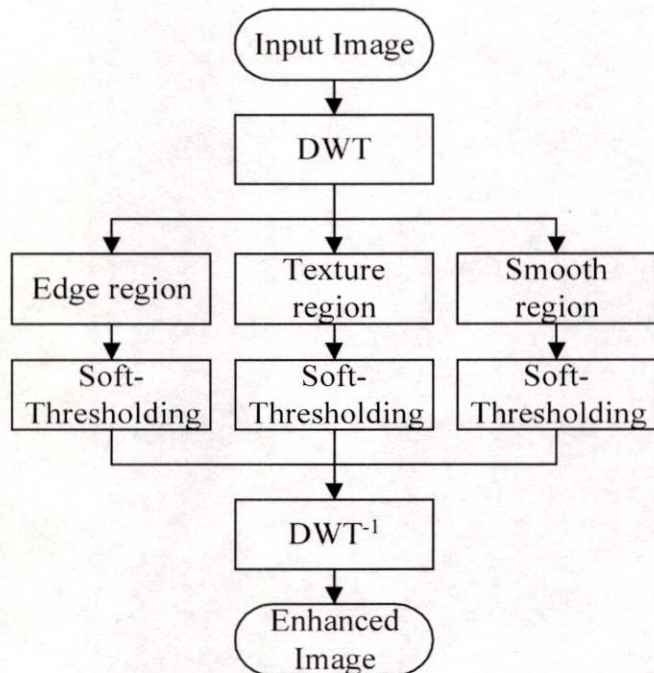
รูปที่ 4.4 ความหนาแน่นของค่าสูงสุดและค่าต่ำสุดของภาพตัวอย่าง 4 ภาพ



รูปที่ 4.5 ตัวอย่างการแยกเนื้อหาของภาพ (a) ข้อมูลภาพต้นแบบ (b) พื้นที่สีขาวแสดงตำแหน่งของพิกเซลที่มีค่าสูงสุดและค่าต่ำสุด (c) พื้นที่สีขาวแสดงตำแหน่งของพิกเซลที่มีความหนาแน่นของค่าสูงสุดและค่าต่ำสุดในช่วง 0.02-0.2 (d) พื้นที่สีขาวแสดงตำแหน่งของพิกเซลที่มีความหนาแน่นของค่าสูงสุดและค่าต่ำสุดในช่วง 0.1-0.5

4.4.2 การทำเวฟเล็ตเชรโซลดิ้ง

ใช้วิธีการทำซอฟต์แวร์โซลดิ้งด้วยค่าเชรโซลด์ที่แตกต่างกันตามลักษณะของพื้นที่ข้อมูลภาพ ดังนี้ค่าเชรโซลด์ของพื้นที่ในส่วนที่เป็นขอบภาพ (δ_E) ส่วนที่เป็นเนื้อหาของภาพ (δ_T) และส่วนที่ราบเรียบของภาพ (δ_S) เนื่องจากสัญญาณรบกวนมีการกระจายอย่างสม่ำเสมอ (Uniform distribution) ที่ทุกระดับของการแปลงเวฟเล็ต (Level) และเห็นได้อย่างชัดเจนที่แบนด์ขององค์ประกอบความถี่สูง ($W_{2^j}^1 f, W_{2^j}^2 f$) ดังนั้นขั้นตอนการทำซอฟต์แวร์โซลดิ้งจะทำในทุกแบนด์ยกเว้นที่แบนด์ขององค์ประกอบความถี่ต่ำ ($S_{2^j} f$) ในที่นี้สมมุติว่าทราบค่าของสัญญาณรบกวนในแต่ละแบนด์ขององค์ประกอบความถี่สูงและค่าเชรโซลด์หาได้จากค่าเบี่ยงเบนมาตรฐาน (σ) ของค่าสัมประสิทธิ์เวฟเล็ตในแต่ละแบนด์ขององค์ประกอบความถี่สูงที่สเกล 2^j นั่นคือ $\delta = \sigma_{b,j}$ เมื่อ b คือแบนด์ขององค์ประกอบความถี่สูงและ j คือระดับการแปลงภาพ ค่าเชรโซลด์ที่หาได้นี้จะทำการปรับค่าให้เหมาะสมตามลักษณะพื้นที่ของข้อมูลที่แตกต่างกันด้วยค่าแฟคเตอร์ที่เหมาะสมค่าหนึ่ง จะได้ว่า $(\delta_E, \delta_T, \delta_S) = \sigma_{b,j} \times (c_E, c_T, c_S)$ และ $(c_E, c_T, c_S) = (0, 0.5, 1)$ ขั้นตอนการทำเวฟเล็ตเชรโซลดิ้งที่ปรับตัวเองได้แสดงดังรูปที่ 4.6

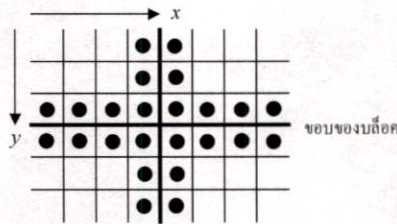


รูปที่ 4.6 อัลกอริทึมของเวฟเล็ตเชรโซลดิ้งที่ปรับตัวเองได้

4.5 การวัดประสิทธิภาพของการลดขอบบล็อค

การวัดประสิทธิภาพของภาพมีบทบาทสำคัญอย่างมากในสาขาการประมวลผลภาพที่เกี่ยวข้องกับการเข้ารหัสข้อมูลภาพ การวัดประสิทธิภาพของภาพของผลภาพที่สร้างกลับมานั้นสามารถหาได้จากค่า PSNR ดังสมการที่ (2.8) อย่างไรก็ตามการวัดค่า PSNR ยังไม่เพียงพอต่อการวัดค่าประสิทธิภาพของผลภาพที่ได้จากการลดขอบบล็อค เนื่องจากการวัดค่าความผิดเพี้ยนของพิกเซลต่อพิกเซลตลอดทั้งภาพ แต่ไม่ได้ทำการวัดค่าการเปลี่ยนแปลงที่บริเวณขอบบล็อค ดังนั้นจึงเลือกใช้การวัดค่าความแปรปรวนที่บริเวณขอบบล็อคหรือ EV [20] ซึ่งเป็นเกณฑ์การวัดประสิทธิภาพของผลภาพอย่างง่ายวิธีการหนึ่ง การหาค่าความแปรปรวนสามารถคำนวณได้จากค่าพิกเซลของข้อมูลภาพ $f(x, y)$ สองพิกเซลที่อยู่ติดกันในแถวหรือคอลัมน์เดียวกันแต่อยู่ในบล็อคที่แตกต่างกัน โดยปกติขอบบล็อคข้อมูลภาพก่อนเข้ารหัสข้อมูลภาพจะมีค่าของพิกเซลทั้งสองใกล้เคียงกันแต่จะมีค่าที่แตกต่างกันอย่างมากเมื่อผ่านขั้นตอนการควอนไทซ์ หาได้จาก

$$EV = \sum_{x=1}^M \sum_{y=1}^{\frac{N}{8}-1} [f(x, 8y) - f(x, 8y+1)]^2 + \sum_{x=1}^{\frac{M}{8}-1} \sum_{y=1}^N [f(8x, y) - f(8x+1, y)]^2 \quad (4.12)$$



รูปที่ 4.7 ตำแหน่งพิกเซลที่บริเวณขอบบล็อคของการหาค่า EV

4.6 สรุป

ในเนื้อหาของบทนี้จะกล่าวถึงการแก้ปัญหาเพื่อลดขอบบล็อคข้อมูลภาพในมุมมองของปัญหาเช่นเดียวกับการกำจัดสัญญาณรบกวนหรือที่เรียกว่า Denoising และวิธีการที่น่าสนใจอย่างหนึ่งคือการทำเวฟเล็ตเรโซลคิงซึ่งเป็นการกำจัดสัญญาณรบกวนในโดเมนเวฟเล็ต แต่เนื่องจากสัญญาณรบกวนในข้อมูลภาพสามารถสังเกตเห็นได้แตกต่างกันตามลักษณะของพื้นที่ภาพ ดังนั้นในวิทยานิพนธ์นี้ได้นำเสนอวิธีการทำเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้ซึ่งเป็นวิธีการที่สามารถกำจัดสัญญาณรบกวนหรือลดขอบบล็อคข้อมูลภาพได้ดีกว่าเวฟเล็ตเรโซลคิงแบบเดิม รวมทั้งรายละเอียดของขั้นตอนการทำเวฟเล็ตเรโซลคิงทั้งสองรูปแบบ เพื่อใช้ในการทดลองและการวิเคราะห์ผลต่อไป

บทที่ 5

ผลการทดลอง

บทนี้จะกล่าวถึงการเตรียมข้อมูลภาพที่ใช้ในการทดลองและแสดงผลการเปรียบเทียบประสิทธิภาพในการลดขอบบล็อคด้วยวิธีการที่ได้กล่าวไว้แล้วในบทที่ 4 โดยแสดงการวัดผลเปรียบเทียบด้วยค่า PSNR (สมการที่ 2.8) และ EV (สมการที่ 4.12) พร้อมทั้งแสดงผลภาพที่ได้จากการลดขอบบล็อคด้วยวิธีการกรองความถี่ต่ำผ่าน การกรอง Wiener ที่ปรับตัวเองได้ เวฟเล็ตเชอร์โซลคิงและเวฟเล็ตเชอร์โซลคิงที่ปรับตัวเองได้ การทดลองทั้งหมดนี้ได้จำลองการทำงานด้วยการเขียนโปรแกรมที่ทำงานด้วย MATLAB เวอร์ชัน 5.1 ประมวลผลบนเครื่อง PC Pentium II 266 MHz

5.1 ข้อมูลภาพที่ใช้ในการทดลอง

ข้อมูลภาพที่ใช้ในการทดลองเป็นข้อมูลภาพมาตรฐานที่พบเห็นได้ทั่วไปในงานวิจัยทางด้านการประมวลผลภาพ ซึ่งเป็นข้อมูลภาพหลายระดับเทา (Grayscale) ขนาด 256×256 พิกเซล จำนวน 6 ภาพได้แก่ ภาพ LENA, BARBARA, BABOON, CMAN, PEPPERS และ AIRPLANE จากนั้นนำข้อมูลภาพเหล่านี้มาผ่านขั้นตอนการบีบอัดและขยายข้อมูลภาพแบบ JPEG ด้วยโปรแกรม MATLAB เวอร์ชัน 5.1 โดยกำหนดคุณภาพของการบีบอัดข้อมูลภาพด้วยค่า Q (Quality factor) เท่ากับ 5, 10, 15, 20, 25 และ 30 ตามลำดับ และให้ภาพที่บิตเรตต่างๆ ดังตารางที่ 5.1 จะเห็นว่าค่า Q ที่มีค่าน้อยจะให้ภาพที่บิตเรตต่ำกว่าและแสดงความผิดเพี้ยนแบบบล็อคมากกว่าค่า Q ที่มีค่าสูงกว่าดังรูปที่ 5.1

ตารางที่ 5.1 ข้อมูลภาพที่ใช้ในการทดลองที่บิตเรตต่างๆ

QUALITY FACTOR	BITRATE					
	LENA	BARBARA	BABOON	CMAN	PEPPERS	AIRPLANE
5	0.31	0.36	0.34	0.30	0.33	0.33
10	0.42	0.48	0.56	0.40	0.45	0.45
15	0.51	0.60	0.74	0.48	0.54	0.55
20	0.59	0.70	0.91	0.56	0.62	0.64
25	0.66	0.82	1.07	0.63	0.70	0.72
30	0.74	0.92	1.21	0.70	0.77	0.80



(a)



(b)



(c)



(d)

รูปที่ 5.1 ตัวอย่างของภาพ LENA ที่บิตเรทต่างๆ (a) ภาพต้นแบบ (b) ภาพ LENA ที่ 0.51บิต/พิกเซล (c) ภาพ LENA ที่ 0.42 บิต/พิกเซล (d) ภาพ LENA ที่ 0.31 บิต/พิกเซล

5.2 การลดขอบบล็อควงการกรองความถี่

จากรูปที่ 5.2 และ 5.3 แสดงตัวอย่างของผลภาพที่ได้จากการลดขอบบล็อควงการกรองความถี่ต่ำและการกรอง Wiener ที่ปรับตัวเองได้ตามสมการที่ (4.1) เมื่อกำหนดค่า M ต่างๆ ตามลำดับ



(a)



(b)

รูปที่ 5.2 ตัวอย่างภาพ LENA เมื่อลดขอบบล็อกรหัสด้วยการกรองความถี่ต่ำผ่านได้ (a) ภาพ LENA ที่ 0.42 บิต/พิกเซล (b) ภาพ LENA ที่ 0.31 บิต/พิกเซล



(a)



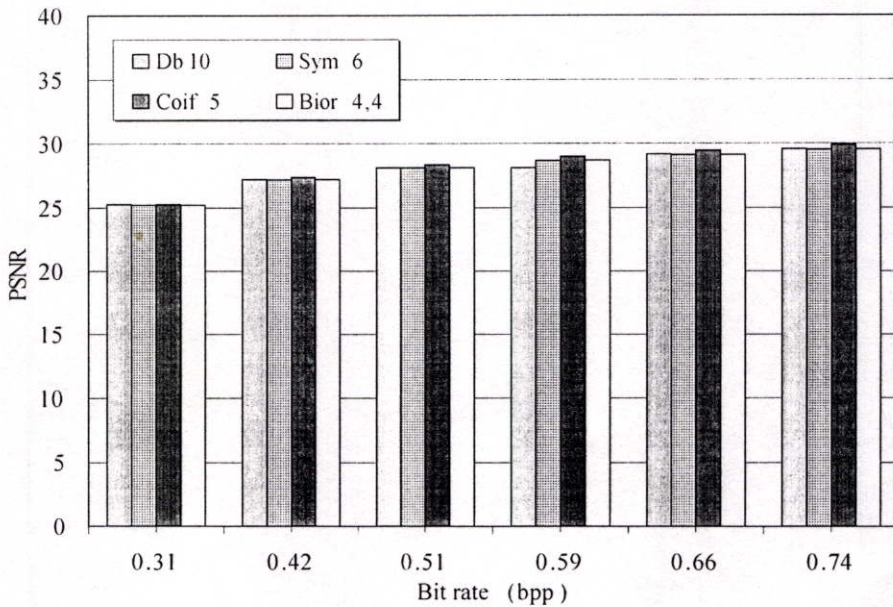
(b)

รูปที่ 5.3 ตัวอย่างภาพ LENA เมื่อลดขอบบล็อกรหัสด้วยการกรอง Wiener ที่ปรับตัวเองได้เมื่อกำหนดค่า M ดังนี้ (a) $M = 1$ (b) $M = 2$

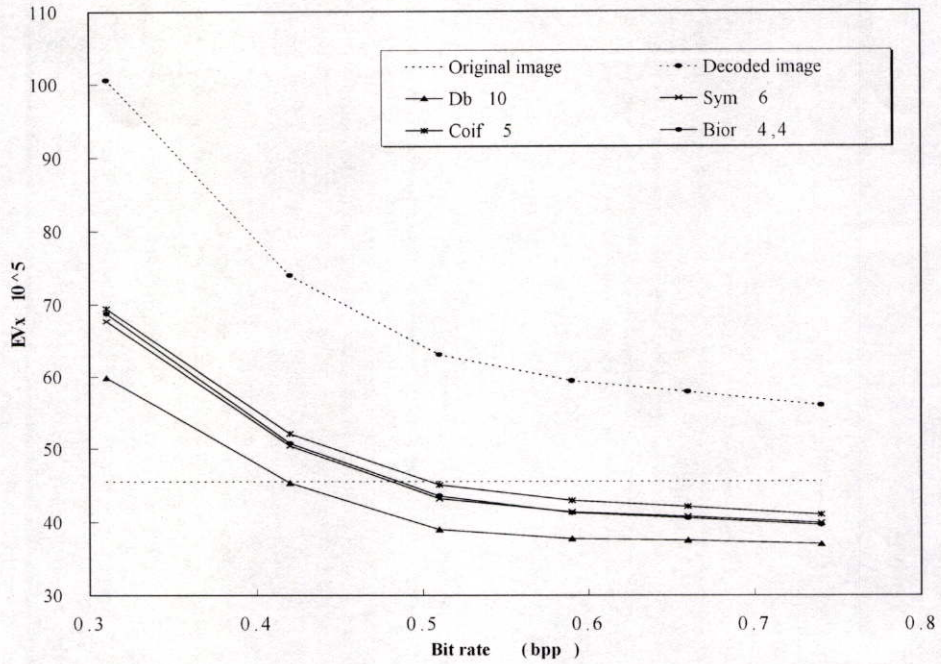
5.3 การลดขอบบล็อกรหัสด้วยเวฟเล็ตเรโซลคิง

ขั้นตอนของการทำเวฟเล็ตเรโซลคิงแสดงดังรูปที่ 4.2 ในที่นี้จะแสดงผลที่ได้จากการทดลองเพื่อลดขอบบล็อกรหัสของภาพ LENA ที่บิตเรทต่างๆ เท่านั้น การลดขอบบล็อกรหัสข้อมูลภาพด้วยเวฟเล็ตเรโซลคิงได้กำหนดค่า $L = 3$ เนื่องจากระดับของการแปลงเวฟเล็ต (L) ที่มีค่ามากกว่า 3

จะมีผลต่อการลดขอบของบล็อกข้อมูลเพียงเล็กน้อยเท่านั้น ทำนองเดียวกันจากการทดลองพบว่าค่าแฟคเตอร์ของค่าเรโซลต์ (f_a) ที่มีค่าเท่ากับ $1/2$ เป็นค่าที่เหมาะสมในฟิลเตอร์ของเวฟเล็ตทุกระยะ การวัดประสิทธิภาพของผลภาพที่ได้จากการลดขอบบล็อกด้วยเวฟเล็ตเรโซลต์ดังโดยใช้ความยาวของฟิลเตอร์ของเวฟเล็ตแต่ละตระกูลที่ให้ผลภาพดีที่สุดมาทำการเปรียบเทียบกันได้แก่ Daubechies-10 (Db10) Symlet-6 (Sym6) Coiflet-5 (Coif5) และ Biorthogonal (Bior 4,4) โดยการวัดค่า PSNR และ EV ของภาพ LENA ที่ผ่านขั้นตอนการบีบอัดและขยายข้อมูลภาพแบบ JPEG ที่มีบิตเรตตั้งแต่ 0.31 ถึง 0.74 บิต/พิกเซล และแสดงผลที่ได้ดังรูปที่ 5.4 และ 5.5 ตามลำดับ รูปที่ 5.4 แสดงให้เห็นว่าฟิลเตอร์ของเวฟเล็ตที่ให้ค่า PSNR ดีที่สุดคือเวฟเล็ตชนิด Coif5 รูปที่ 5.5 แสดงการเปรียบเทียบค่า EV ของข้อมูลภาพต้นแบบและภาพที่ได้จากการลดขอบบล็อกของเวฟเล็ตแต่ละชนิด พบว่าผลจากการลดขอบบล็อกข้อมูลภาพด้วยเวฟเล็ตชนิด Coif5 จะมีค่าใกล้เคียงกับข้อมูลภาพต้นแบบมากที่สุด รูปที่ 5.6 แสดงตัวอย่างของผลภาพที่ได้เมื่อทำเวฟเล็ตเรโซลต์ด้วยเวฟเล็ตชนิด Coif5 ของภาพ LENA ที่บิตเรต 0.42 บิต/พิกเซล [21]



รูปที่ 5.4 ค่า PSNR ของภาพ LENA ที่บิตเรตต่างๆ เมื่อใช้เวฟเล็ตต่างชนิดกัน



รูปที่ 5.5 ค่า EV ของภาพ LENA ที่บิตเรตต่างๆ เมื่อใช้เวฟเล็ตต่างชนิดกัน



(a)



(b)

รูปที่ 5.6 ตัวอย่างของภาพ LENA ที่ได้จากการลดขอบบล็อกรด้วยเวฟเล็ตเรโซลคั้ง (a) ข้อมูลภาพ LENA ที่ 0.42 บิต/พิกเซล (b) ผลการลดขอบบล็อกรด้วยเวฟเล็ตชนิด Coif5

5.4 การลดขอบบล็อควงเวฟเล็ตเชอร์โซลดิ้งที่ปรับตัวเองได้

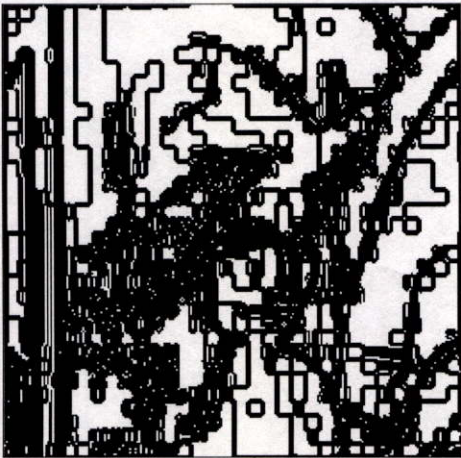
5.4.1 ผลของการจำแนกลักษณะพื้นที่ของข้อมูลภาพ

5.4.1.1 การจำแนกพื้นที่ของขอบภาพ

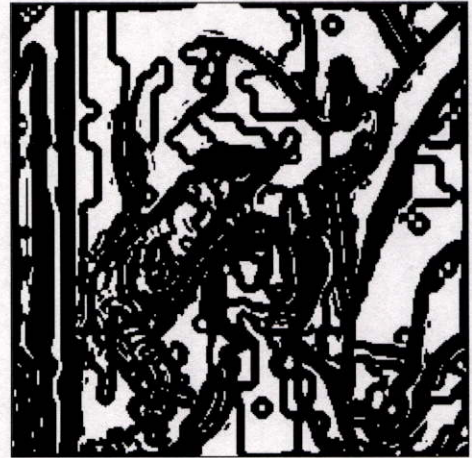
การหาพื้นที่ขอบภาพของการแปลงเวฟเล็ตในแต่ละระดับของการแปลงภาพ 2^j เมื่อ $1 \leq j \leq 3$ จากสมการที่ (4.10) ด้วยการกำหนดค่า Lipchitz $\alpha = 1$ และค่า K ต่างๆ พบว่าถ้าค่า K มีค่าที่น้อยเกินไปจะไม่สามารถแยกขอบภาพออกจากสัญญาณรบกวนหรือความผิดเพี้ยนแบบบล็อควงได้ดังรูปที่ 5.7 ทำนองเดียวกันถ้าค่าของ K มีค่าที่มากเกินไปทำให้ไม่สามารถแยกขอบภาพได้ดีนักดังรูปที่ 5.8 อย่างไรก็ตามสัญญาณรบกวนหรือความผิดเพี้ยนแบบบล็อควงยังคงปรากฏให้เห็นที่ระดับของการแปลงภาพ 2^j เมื่อ $j = 1, 2$ ดังนั้นเพื่อให้ได้ลักษณะของขอบภาพจริงๆ โดยปราศจากความผิดเพี้ยนแบบบล็อควง ในที่นี้จะใช้ผลของการหาขอบที่ระดับของการแปลงภาพ $j = 3$ เมื่อกำหนดค่า $K = 10$ เป็นขอบภาพที่ใช้อ้างอิงแทนขอบภาพที่ระดับของการแปลงภาพ $j = 1, 2$ ซึ่งสามารถแยกขอบภาพได้ดีที่สุดดังรูปที่ 5.9



(a)

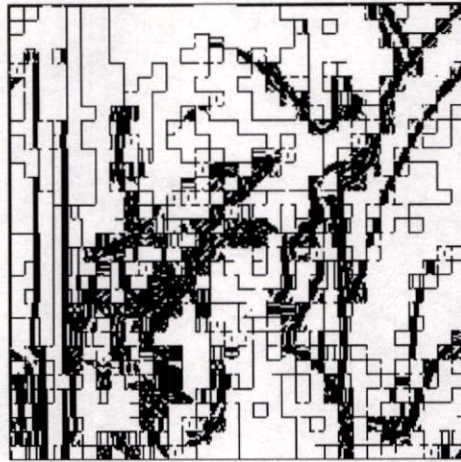


(b)



(c)

รูปที่ 5.7 การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 5$ (a) ระดับของการแปลงภาพที่ $j = 1$ (b) ระดับของการแปลงภาพที่ $j = 2$ (c) ระดับของการแปลงภาพที่ $j = 3$



(a)

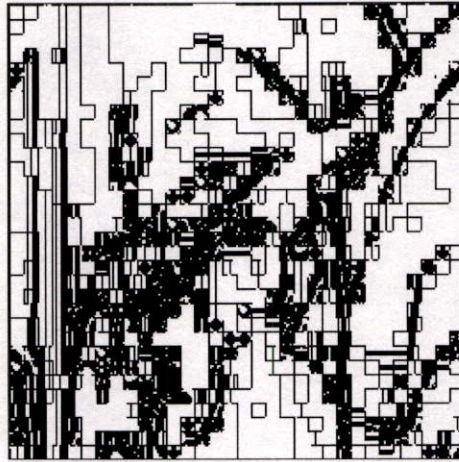


(b)

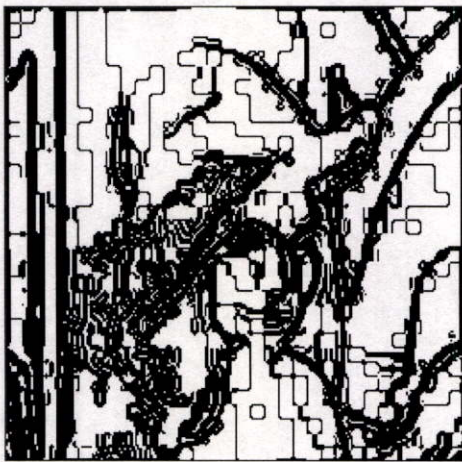


(c)

รูปที่ 5.8 การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 25$ (a) ระดับของการแปลงภาพที่ $j = 1$ (b) ระดับของการแปลงภาพที่ $j = 2$ (c) ระดับของการแปลงภาพที่ $j = 3$



(a)



(d)

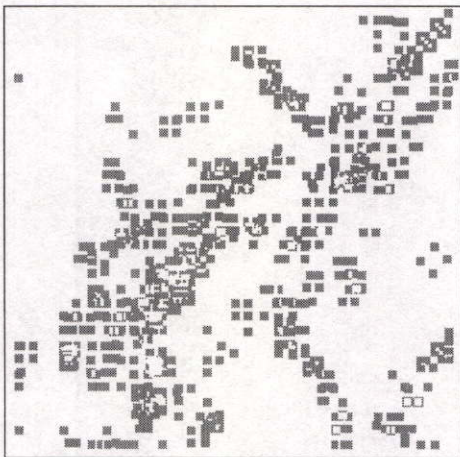


(c)

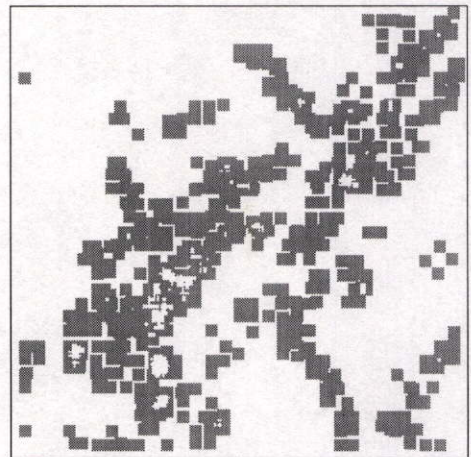
รูปที่ 5.9 การแยกขอบของภาพ LENA ที่ 0.31 บิต/พิกเซลด้วยค่า $K = 10$ (a) ระดับของการแปลงภาพที่ $j = 1$ (b) ระดับของการแปลงภาพที่ $j = 2$ (c) ระดับของการแปลงภาพที่ $j = 3$

5.4.1.2 การจำแนกพื้นที่ส่วนที่เป็นเนื้อหาและส่วนที่ราบเรียบของภาพ

การหาความหนาแน่นของค่าสูงสุดและค่าต่ำสุดของข้อมูลภาพเมื่อใช้ขนาดวินโดว์ดังนี้ 3×3 , 5×5 , 7×7 , 9×9 และ 11×11 จากรูปที่ 5.10 แสดงตัวอย่างของผลภาพที่ได้จากการจำแนกเนื้อหาของภาพ (พื้นที่สีเทา) และส่วนที่ราบเรียบของภาพ (พื้นที่สีขาว) ของวินโดว์ขนาด 5×5 และ 7×7 พิกเซล ของข้อมูลภาพ LENA ที่ 0.31 บิต/พิกเซล ในรูปที่ 5.1(d) เมื่อกำหนดขอบเขตของค่าบนและค่าล่างของส่วนที่เป็นเนื้อหาของภาพเท่ากับ 0.01-0.1 เนื่องจากเป็นค่าที่เหมาะสมที่ได้จากการทดลอง ตัวอย่างของกราฟแสดงความสัมพันธ์ระหว่างความหนาแน่นของพิกเซลที่มีค่าสูงสุดและค่าต่ำสุดเมื่อใช้ขนาดวินโดว์ต่างๆ ของข้อมูลภาพ LENA ที่ 0.31 บิต/พิกเซล และภาพ PEPPERS ที่ 0.33 บิต/พิกเซลพบว่าวินโดว์ขนาด 7×7 พิกเซล จะมีลักษณะที่สม่ำเสมอมากกว่าขนาดวินโดว์อื่นๆ ดังรูปที่ 5.11(a) และ (b) ตามลำดับ

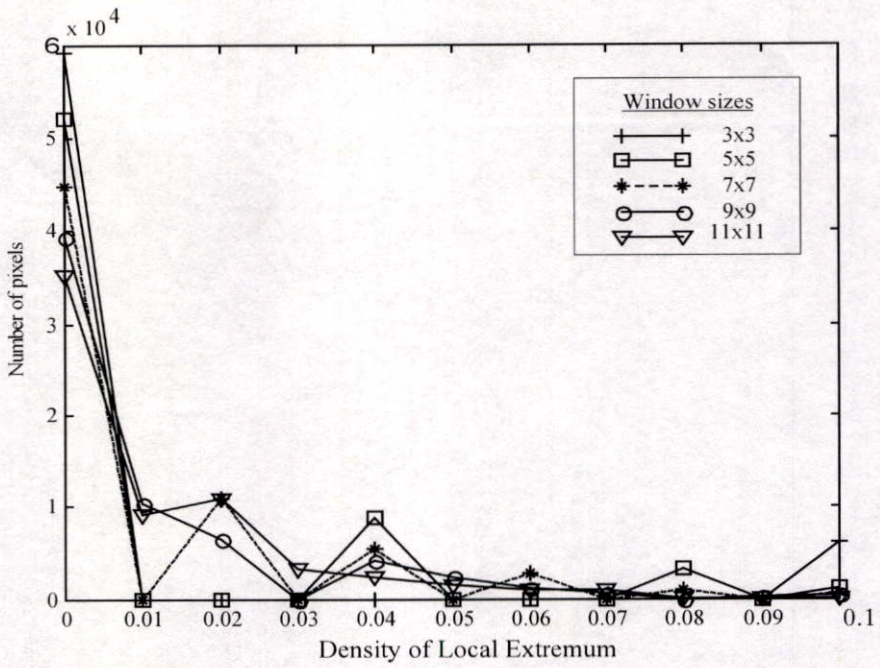


(a)

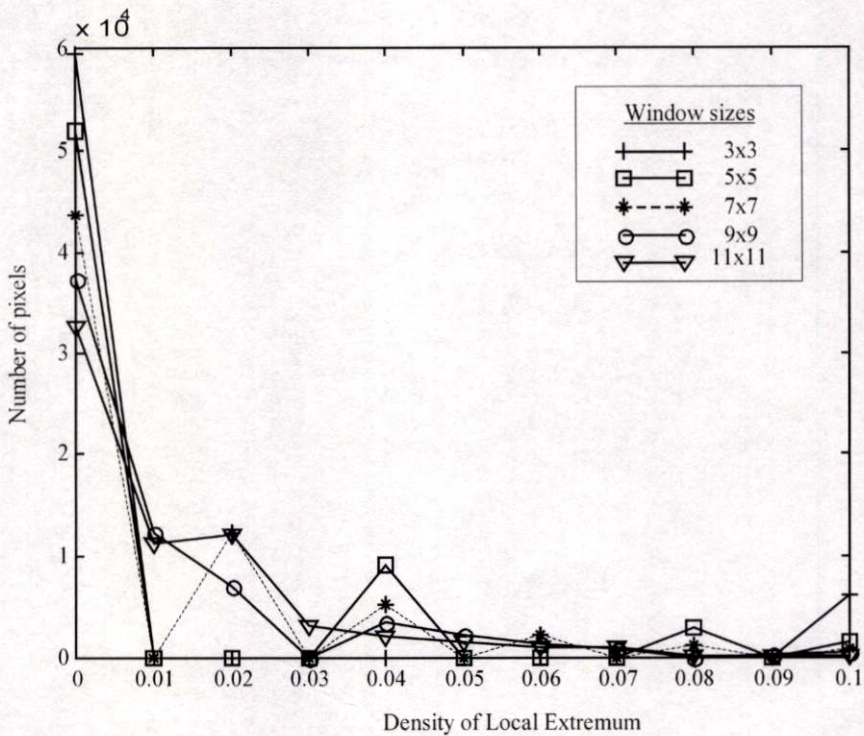


(b)

รูปที่ 5.10 การจำแนกเนื้อหา (พื้นที่สีเทา) และส่วนที่ราบเรียบ (พื้นที่สีขาว) ของภาพ LENA ที่ 0.31 บิต/พิกเซล เมื่อให้ส่วนที่เป็นเนื้อหาของภาพมีความหนาแน่นของพิกเซลที่มีค่าสูงสุดและค่าต่ำสุดอยู่ในช่วงขอบเขตของค่าบนและค่าล่างเท่ากับ 0.01-0.1 (a) ขนาดวินโดว์เท่ากับ 5×5 พิกเซล (b) ขนาดวินโดว์เท่ากับ 7×7 พิกเซล



(a)

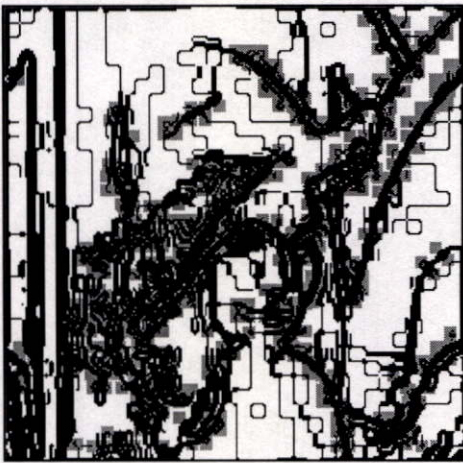


(b)

รูปที่ 5.11 ความสัมพันธ์ระหว่างความหนาแน่นของพิกเซลที่มีค่าสูงสุดและค่าต่ำสุดข้อมูลภาพเมื่อกำหนดขนาดวินโดว์ต่างๆ (a) ภาพ LENA ที่ 0.31 บิต/พิกเซล (b) ภาพ PEPPERS ที่ 0.33 บิต/พิกเซล



(a)



(b)

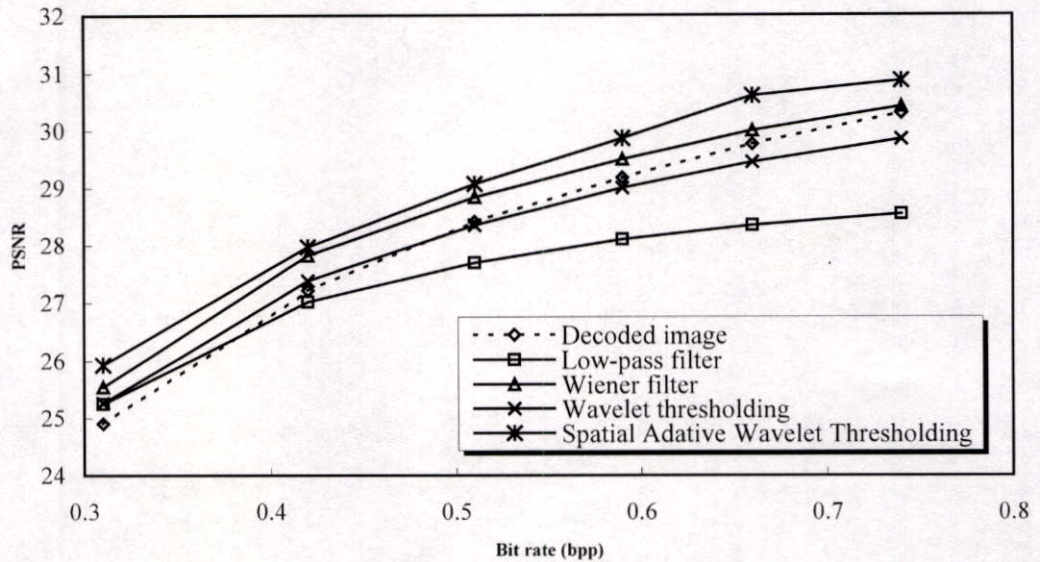


(c)

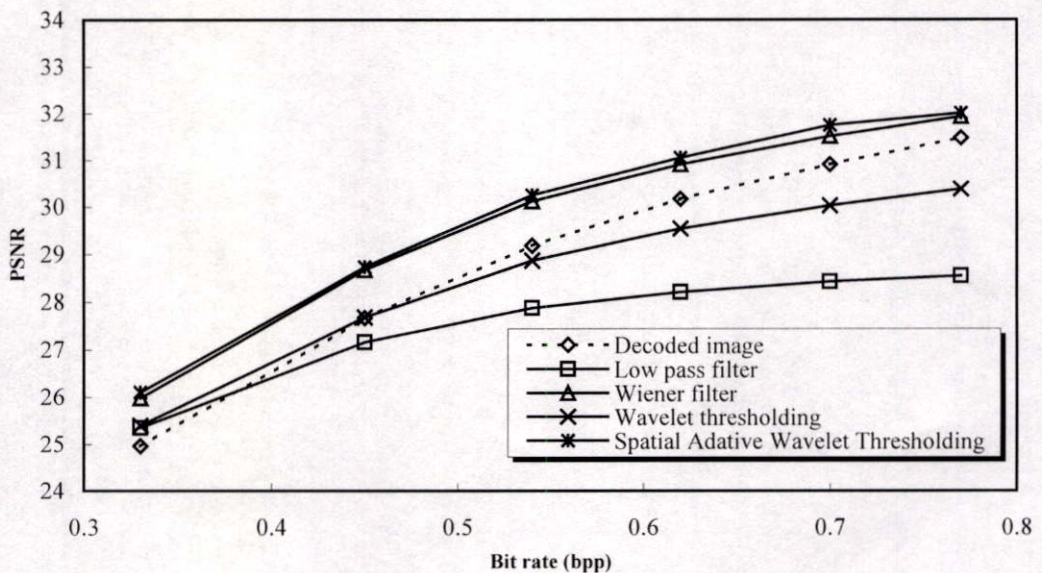
รูปที่ 5.12 ตัวอย่างผลที่ได้ในจำแนกขอบภาพ (พื้นที่สีดำ) เนื้อหา (พื้นที่สีขาว) และส่วนที่ราบเรียบ (พื้นที่สีเทา) ของภาพ LENA ที่ 0.31 บิต/พิกเซล (a) ระดับของการแปลงภาพที่ $j = 1$ (b) ระดับของการแปลงภาพที่ $j = 2$ (c) ระดับของการแปลงภาพที่ $j = 3$

5.4.2 เปรียบเทียบผลของวิธีการที่นำเสนอกับวิธีการอื่น

เปรียบเทียบผลของการทดลองของวิธีการที่นำเสนอกับวิธีการอื่นๆ ได้แก่ การกรองความถี่ต่ำผ่านเมื่อใช้ฟิลเตอร์เท่ากับ $1/16[1\ 2\ 1; 2\ 4\ 2; 1\ 2\ 1]$ การกรอง Wiener ที่ปรับตัวเองได้เมื่อกำหนดให้ $M = 1$ และเวฟเลตเรโซลชันตั้งเมื่อใช้เวฟเลตชนิด Coif5 [22, 23]

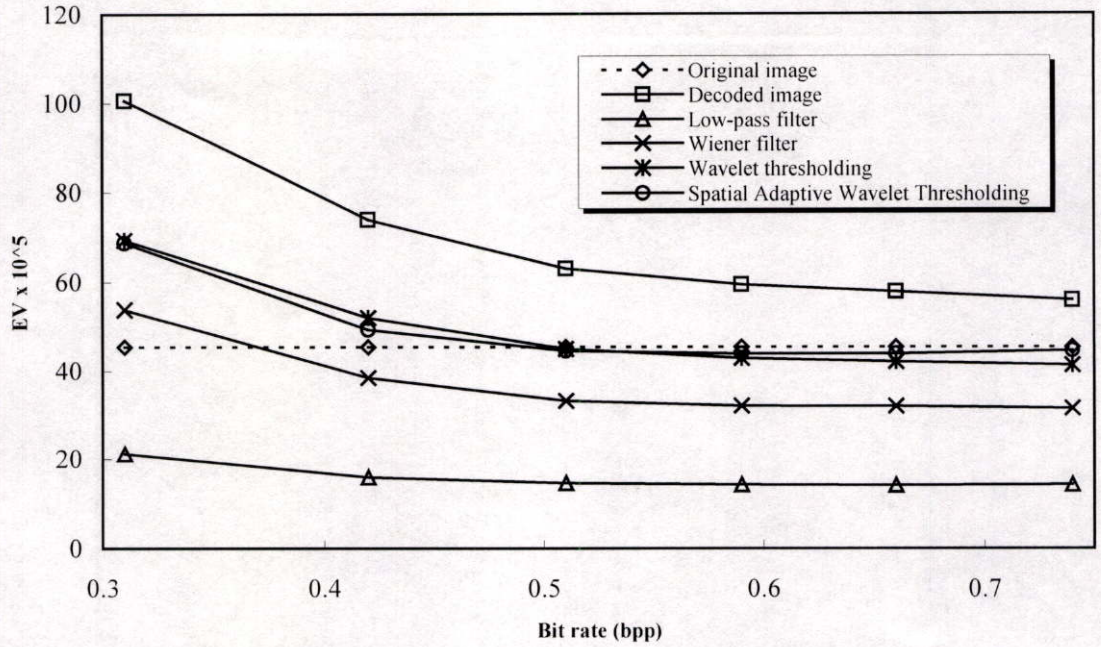


(a)

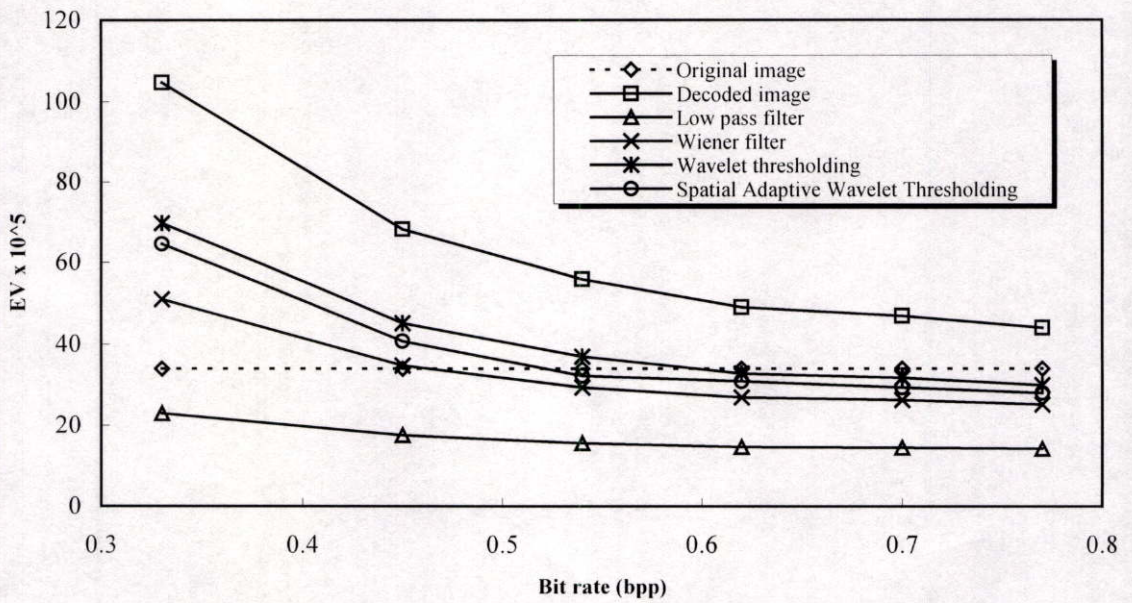


(b)

รูปที่ 5.13 ตัวอย่างของผลการเปรียบเทียบค่า PSNR เมื่อลดขอบบลิ๊อคด้วยวิธีการต่างๆ (a) ภาพ LENA ที่บิตเรทต่างๆ (b) ภาพ PEPPERS ที่บิตเรทต่างๆ

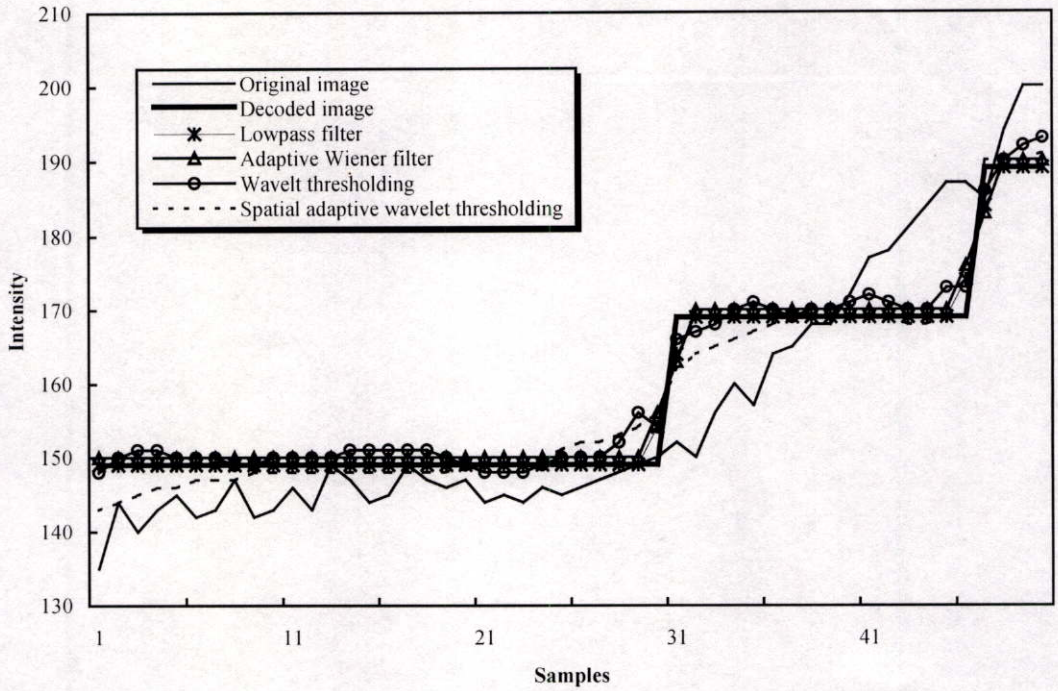


(a)

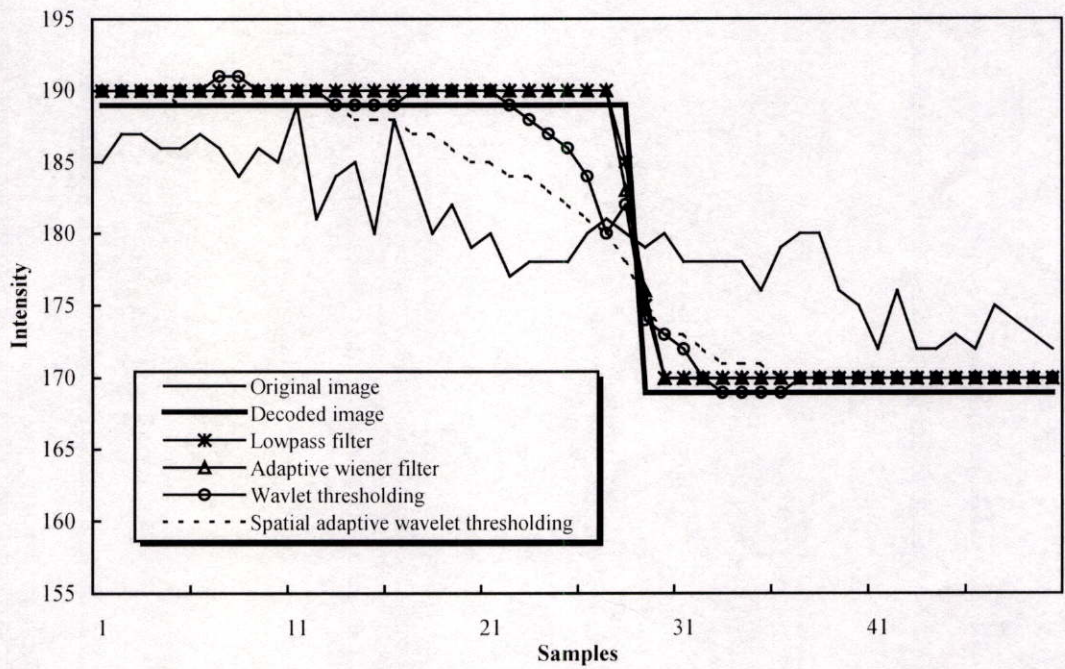


(b)

รูปที่ 5.14 ตัวอย่างผลของการเปรียบเทียบค่า EV เมื่อลดขอบบด็อดด้วยวิธีการต่างๆ (a) ภาพ LENA ที่บิตเรทต่างๆ (b) ภาพ PEPPERS ที่บิตเรทต่างๆ



(a)



(b)

รูปที่ 5.15 ตัวอย่างของผลการเปรียบเทียบระดับความเข้มของพิกเซลเมื่อลดคอบบลิ้อคด้วยวิธีการต่างๆ (a) ภาพ LENA (b) ภาพ CMAN



(a)



(b)



(c)



(d)

รูปที่ 5.16 ตัวอย่างของภาพ LENA ที่ 0.42 บิต/พิกเซล เมื่อลดขอบบดลอกของภาพด้วยวิธีการต่างๆ (a) การกรองความถี่ต่ำผ่าน (b) การกรอง Wiener ที่ปรับตัวเองได้ (c) เวฟเล็ตเชรโซลคิงเมื่อใช้เวฟเล็ตชนิด Coif5 (d) เวฟเล็ตเชรโซลคิงที่ปรับตัวเองได้

บทที่ 6

สรุปผลการวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้ได้ทำการศึกษาและวิจัยเกี่ยวกับการลดขอบบล็อคในข้อมูลภาพที่เข้ารหัสด้วยการแปลงดิสครีตโคไซน์แบบบล็อค โดยนำเสนอวิธีการทำเวฟเล็ตเรซโวลคิงที่ปรับตัวเองได้ซึ่งเป็นวิธีการที่ใช้คุณสมบัติของการทำเรซโวลคิงในโดเมนเวฟเล็ตและการมองเห็นสัญญาณรบกวนที่มีความแตกต่างกันในแต่ละพื้นที่ของข้อมูลภาพ การวิจัยได้แบ่งออกเป็น 4 ส่วน คือ

1. การลดขอบบล็อคด้วยการกรองความถี่

ผลการลดขอบบล็อคด้วยการกรองความถี่ต่ำและการกรอง Wiener ที่ปรับตัวเองได้สามารถลดขอบบล็อคได้ในระดับหนึ่งและยังเห็นขอบบล็อคอยู่ นอกจากนี้การกรองความถี่มีผลทำให้ภาพที่ได้เบลอหรือลดทอนความคมชัดของภาพไปและสูญเสียรายละเอียดของภาพบางส่วนไป

2. การลดขอบบล็อคด้วยเวฟเล็ตเรซโวลคิง

การหาความยาวของฟิลเตอร์ของเวฟเล็ตแต่ละชนิดที่ให้ผลภาพดีที่สุดมาทำการเปรียบเทียบกัน ได้แก่ Daubechies-10 (Db10) Symlet-6 (Sym6) Coiflet-5 (Coif5) และ Biorthogonal (Bior 4,4) โดยการวัดค่า PSNR และ EV พบว่าฟิลเตอร์ของเวฟเล็ตชนิด Coif5 ให้ค่า PSNR ดีที่สุดและค่า EV ของภาพที่ได้จากการลดขอบบล็อคด้วยฟิลเตอร์ของเวฟเล็ตชนิด Coif5 มีค่าใกล้เคียงกับภาพต้นแบบมากที่สุด นั่นแสดงให้เห็นว่าลักษณะความคมชัดที่บริเวณขอบภาพมีความใกล้เคียงกับภาพต้นแบบมากที่สุดด้วยเช่นกัน นอกจากนี้ยังให้ผลภาพที่ดีขึ้น

3. การลดขอบบล็อคด้วยเวฟเล็ตเรซโวลคิงที่ปรับตัวเองได้

3.1 การหาพื้นที่ของขอบภาพด้วยค่ามอดูลัสสูงสุดของการแปลงเวฟเล็ตในแต่ละระดับของการแปลงภาพ 2^j เมื่อ $1 \leq j \leq 3$ เมื่อกำหนดค่า Lipchitz $\alpha = 1$ และค่า K ต่างๆ พบว่าเมื่อกำหนดค่า K ด้วยค่ามากขึ้นไปจะไม่สามารถแยกขอบภาพได้ดีหรือค่า K น้อยเกินไปจะพบความผิดเพี้ยนแบบบล็อคในทุกๆระดับของการแปลงภาพ ดังนั้นจากการทดลองพบว่าค่า $K = 10$ และระดับของการแปลงภาพที่ $j = 3$ เป็นค่าที่เหมาะสมกับทุกภาพและไม่มีผิดเพี้ยนแบบบล็อคปรากฏให้เห็น ดังนั้นจึงใช้ขอบของระดับของการแปลงภาพที่ $j = 3$ เป็นขอบภาพอ้างอิงในระดับของการแปลงภาพที่ $j = 1, 2$

3.2 การหาพื้นที่ในส่วนที่เป็นเนื้อหาและส่วนที่ราบเรียบของภาพ ด้วยการวัดระดับความละเอียดของภาพจากความหนาแน่นของค่าสูงสุดและค่าต่ำสุดของภาพที่ขนาดวินโดว์ต่างๆ พบว่า

เมื่อใช้วินโดว์ขนาด 7×7 พิกเซลและการกำหนดช่วงของความหนาแน่นของค่าสูงสุดและค่าต่ำสุดของส่วนที่ราบเรียบและส่วนที่เป็นเนื้อหาของภาพจะอยู่ในช่วง 0-0.01 และ 0.01-0.1 ตามลำดับ สามารถแยกพื้นที่แต่ละส่วนออกจากกันได้อย่างชัดเจน

3.3 การทำเวฟเล็ตเรโซลคิงด้วยค่าเรโซลคิงที่เหมาะสมตามลักษณะพื้นที่ของข้อมูลภาพ โดยเรียงลำดับของค่าเรโซลคิงได้ดังนี้ $\delta_{Edge} < \delta_{Texture} < \delta_{Smooth}$ เพื่อลดขอบบล็อคในพื้นที่ส่วนที่ราบเรียบของภาพให้ได้มากที่สุดขณะเดียวกันยังคงรักษารายละเอียดในส่วนที่เป็นขอบและเนื้อหาของภาพไว้

4. เปรียบเทียบผลของวิธีการที่นำเสนอกับวิธีการอื่นๆ

วิธีการของเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้จะให้ค่า PSNR สูงสุด ค่า EV มีค่าใกล้เคียงกับภาพต้นแบบและวิธีการลดขอบบล็อคด้วยเวฟเล็ตเรโซลคิง เมื่อพิจารณาถึงลักษณะของระดับความเข้มของพิกเซลพบว่าวิธีการที่ได้แนะนำมีค่าระดับความเข้มของพิกเซลใกล้เคียงกับพิกเซลของภาพต้นแบบมากกว่าและผลภาพที่ได้สามารถลดขอบบล็อคได้ดีกว่าวิธีการอื่นๆ

จากการทดลองนี้สามารถเป็นแนวทางสำหรับผู้ที่มีสนใจสามารถนำมาศึกษาค้นคว้าวิจัยเพื่อพัฒนาการกำจัดสัญญาณรบกวนต่างๆ ในข้อมูลภาพต่อไป เช่น การนำไปประยุกต์ใช้กับการรบกวนในข้อมูลภาพ SAR การพัฒนาอัลกอริทึมของการบีบอัดข้อมูลภาพร่วมกับการกำจัดสัญญาณรบกวน เป็นต้น

6.2 ปัญหาที่เกิดขึ้นและข้อเสนอแนะ

ปัญหาที่พบในการลดขอบบล็อคข้อมูลภาพด้วยการทำเวฟเล็ตเรโซลคิงที่ปรับตัวเองได้คือแฟคเตอร์ของค่าเรโซลคิงที่ใช้เป็นค่าคงที่ที่ได้จากการทดลอง ทำให้ได้ค่าเรโซลคิงในแต่ละพื้นที่ของข้อมูลภาพไม่เหมาะสมอย่างแท้จริง ดังนั้นควรใช้วิธีการประมาณค่าเรโซลคิงที่มีประสิทธิภาพด้วยการวัดจากข้อมูลจริง เช่น การประมาณค่าแบบ Bayesian หรือการใช้การกระจายแบบ Laplacian เป็นต้น การวัดค่าประสิทธิภาพของผลภาพที่ได้ควรใช้วิธีการวัดค่าหลายวิธีประกอบกัน เพื่อให้การวิเคราะห์ผลภาพที่ได้มีความชัดเจนมากยิ่งขึ้น เช่น การวัดค่าโดยใช้หลักการมองเห็นของมนุษย์หรือการวัดค่าในโดเมนของความถี่ เป็นต้น การแยกลักษณะของพื้นที่ข้อมูลภาพในส่วนเนื้อหาของและส่วนที่ราบเรียบของข้อมูลภาพจะใช้เวลาในการคำนวณที่นานควรมีการปรับปรุงหรือพัฒนาอัลกอริทึมต่อไป เพื่อสามารถใช้ในการประมวลผลที่เวลาจริงได้ (Real time processing) การเลือกใช้เวฟเล็ตชนิดอื่นๆ ในการทำเวฟเล็ตเรโซลคิง เช่น Wavelet package เป็นต้น

เอกสารอ้างอิง

- [1] H. Paek, R.C. Kim, and S.U. Lee, "On the POCS-Based Postprocessing Technique to Reduce the Blocking Artifacts in Transform Coded Images," IEEE Trans. on Circuits and Systems for Video Technology, vol. 8, pp. 358-367, 1998.
- [2] M. Crouse and K. Ramchandran, "Nonlinear Constrained Least Squares Estimation to Reduce Artifacts in Block Transform-Coded Images," IEEE International Conference on Image Processing (ICIP'95), vol. 1, pp. 462-465, 1995.
- [3] H.C. Reeve and J.S. Lim, "Reduction of the Blocking Effects in Image Coding," Optical Engineering, vol.23, pp. 34-37, 1984.
- [4] R. Steinmetz and K. Nahrstedt. 1995. **Multimedia: Computing, Communications and Applications**. Prentice Hall Series in Innovative Technology.
- [5] Gregory A.Baxes. 1994. **Digital Image Processing**. USA: John Wiley & Sons, Inc.
- [6] Randy Crane. 1997. **A Simplified Approach to Image Processing**. USA: John Wiley & Sons, Inc.
- [7] ศิริพงษ์ วงษ์คาร. "การลดข้อมูลภาพถ่ายจากดาวเทียมด้วยการควอนไทซ์แบบเวกเตอร์," วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2540
- [8] S. Minami and A. Zakhor, "An Optimization Approach for Removing Blocking Effects in Transform Coding," IEEE Trans. on Circuits and Systems for Video Technology, vol. 5, pp. 74-82, Apr. 1995.
- [9] K. Koutroumbas. 1999. **Pattern Recognition**. Academic Press, Inc.
- [10] C. S. Burrus, R. A. Gopinath, and H. Guo. 1998. **Introduction to Wavelets and Wavelet Transform**. Prentice Hall International, Inc.
- [11] S. Mallat and S. Zhang, "Characterization of Signals from Multiscale Edges," IEEE Trans. on Pattern Analysis And Machine Intelligence, vol. 14, spp. 710-732, July 1992.
- [12] F. Tse and W. Cham, "DC Coefficient Restoration Using MAP Estimation Technique," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'98), vol. 5, pp. 2557-2560, 1998.

- [13] Y.Q. Zhang, R.L. Pickholtz, and M.H. Loew, **"A New Approach to Reduce the "Blocking Effect" of Transform Coding,"** IEEE Trans. on Communications, vol. 41, pp. 299-302, Feb. 1993.
- [14] V.E. DeBrunner, L. Chen, and H.J. Li, **"Lapped Multiple Bases Algorithms for Still Image Compression Without Blocking Effect,"** IEEE Trans. on Image Processing, vol. 6, pp. 1316-1321, Sep. 1997.
- [15] S. Grace and M.Vetteli, **"Spatial Adaptive Wavelet Thresholding for Image Denoising,"** IEEE Image Processing, vol. 2, page 374-377, 1997.
- [16] D.L. Donoho, **"De-noising by Soft Thresholding,"** IEEE Trans. on Information Theory, vol. 41, pp. 613-627, 1995.
- [17] R.A. Gopinath, M. Lang, H. Guo, and J.E. Odergard, **"Wavelet-Based Post-Processing of Low Bit Rate Transform Coded Image,"** International Conference on Image Processing (ICIP'94), vol. II, pp.913-917, 1994.
- [18] B. Vidakovic and P. Muller, **"Wavelets for kids A Tutorial Introduction,"** Duke University, 1994.
- [19] K. Karu, A. K. Jain and R. M. Bolle, **"Is there any Texture in the Image?,"** IEEE Proceedings of ICRP'96, pp. 770-774. 1995.
- [20] R. Horng and A.J. Ahumada, **"A Fast DCT Block Smoothing Algorithm,"** SPIE Conference on Visual Communications and Image Processing, vol. 2501, 1995.
- [21] Y. Rangsanseri and S. Dachasilaruk, **"Filter Evaluation in Image Deblocking by Wavelet Thresholding,"** IEEE Region 10 Conference 99 (TENCON'99), Korea, September, 1999.
- [22] Y. Rangsanseri and S. Dachasilaruk, **"Comparative Study of Post-processing Techniques for Reducing Blockiness in Transform Coded Images,"** IEEE International Symposium on Consumer Electronic (ISCE'99), Malaysia, November, 1999.
- [23] S. Dachasilaruk and Y. Rangsanseri, **"Image Deblocking using Spatial Adaptive Wavelet Thresholding,"** Visual Communications and Image Processing 2000 (VC2000), Perth, Australia, June, 2000. (accepted)
- [24] J. S. Lim. 1990. **Two-Dimensional Signal and Image Processing.** Prentice Hall International, Inc.

ภาคผนวก

ภาคผนวก ก.

การกรอง Wiener ที่ปรับตัวเองได้

1. ทฤษฎี

รูปแบบของข้อมูลภาพที่มีความผิดเพี้ยน (Degraded image) เนื่องจากสัญญาณรบกวนแบบสุ่ม (Additive random noise) เป็นดังนี้ [24]

$$g(n_1, n_2) = f(n_1, n_2) + v(n_1, n_2) \quad (1)$$

โดยที่

$g(n_1, n_2)$ คือข้อมูลภาพที่มีความผิดเพี้ยน

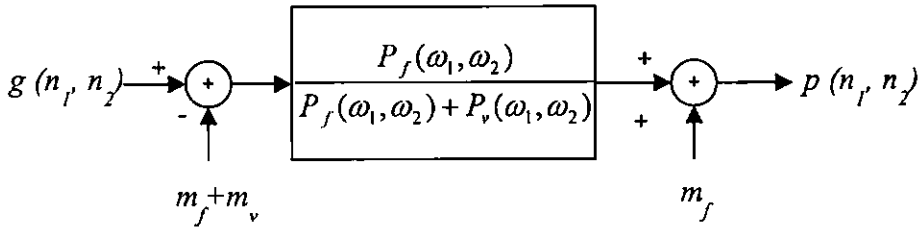
$f(n_1, n_2)$ คือข้อมูลภาพต้นแบบ

$v(n_1, n_2)$ คือสัญญาณรบกวนแบบสุ่ม

การกรอง Wiener เป็นวิธีการที่พัฒนาขึ้นเพื่อลดสัญญาณรบกวนแบบสุ่มในข้อมูลภาพ ถ้าสมมุติให้ $f(n_1, n_2)$ และ $v(n_1, n_2)$ เป็นข้อมูลการสุ่มที่มีค่าเฉลี่ยเท่ากับศูนย์และเป็นเชิงเส้นอิสระ (Linear independent) ถ้าทราบค่ากำลังของสเปกตรัม (Power spectrum) ของ $f(n_1, n_2)$ และ $v(n_1, n_2)$ มีค่าเท่ากับ $P_f(\omega_1, \omega_2)$ และ $P_v(\omega_1, \omega_2)$ ตามลำดับ ดังนั้นการประมาณค่าความผิดพลาดกำลังสองเฉลี่ยที่น้อยที่สุดของ $f(n_1, n_2)$ จะได้จากการกรองสัญญาณ $g(n_1, n_2)$ ด้วยการกรอง Wiener ที่มีผลตอบสนองต่อความถี่เป็น $H(\omega_1, \omega_2)$ และหาได้จาก

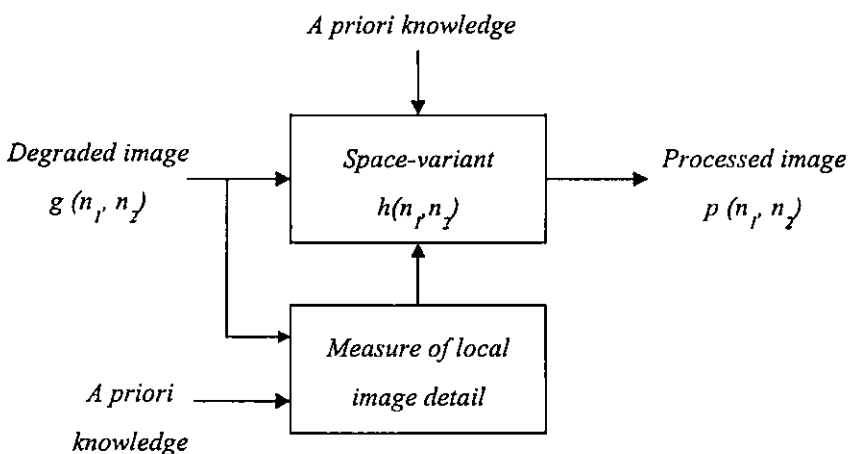
$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_v(\omega_1, \omega_2)} \quad (2)$$

จากลักษณะของการกรอง Wiener ตามสมการที่ (2) จะมีคุณสมบัติเป็นการกรองความถี่ต่ำอย่างหนึ่ง (Lowpass filter) จากรูปที่ 1 ถ้ากำหนดให้ $f(n_1, n_2)$ มีค่าเฉลี่ยเป็น m_f และ $v(n_1, n_2)$ มีค่าเฉลี่ยเป็น m_v สัญญาณที่ได้จาก $g(n_1, n_2) - (m_f + m_v)$ จะถูกกรองด้วยฟิลเตอร์ จากนั้นเพิ่มสัญญาณด้วยค่าเฉลี่ย m_f เพื่อรักษาค่าเฉลี่ยไม่ให้เป็นศูนย์ เป็นการหาค่าผิดพลาดกำลังสองเฉลี่ยที่น้อยที่สุดระหว่าง $f(n_1, n_2)$ และ $p(n_1, n_2)$



รูปที่ ก.1 การกรอง Wiener เพื่อหาค่าผิดพลาดกำลังสองเฉลี่ยที่น้อยที่สุดของ $f(n_1, n_2)$ จาก $g(n_1, n_2) = f(n_1, n_2) + v(n_1, n_2)$

ข้อมูลภาพใดๆ จะมีลักษณะพื้นที่ของข้อมูลภาพในแต่ละส่วนที่แตกต่างกันไป เช่น ท้องฟ้าและผาผนังจะมีค่าความเข้มของระดับสี (Intensity) ที่ใกล้เคียงกันหรือมีรายละเอียดของภาพน้อย (Low-detail image) ขณะที่ต้นไม้และอาคารต่างๆ จะมีการเปลี่ยนแปลงของค่าความเข้มของระดับสี (Intensity) ที่มากกว่าหรือมีรายละเอียดของภาพมากกว่า (High-detail image) ซึ่งสัญญาณรบกวนอย่างเดียวกันจะเห็นได้ชัดเจนกว่าในส่วนที่มีรายละเอียดของภาพน้อย ข้อมูลภาพที่มีความผิดเพี้ยนอาจเกิดจากการเปลี่ยนแปลงของภาพบริเวณหนึ่งไปยังอีกบริเวณหนึ่งด้วย ดังนั้นการประมวลผลภาพโดยใช้คุณสมบัติของภาพในพื้นที่ส่วนย่อย (Local characteristic) จะมีประโยชน์อย่างมากในการสร้างข้อมูลภาพกลับคืนมาจากข้อมูลภาพเดิม (Image restoration) การปรับปรุงภาพ (Image enhancement) และในการประมวลผลภาพอื่นๆ จากลักษณะพื้นที่ของข้อมูลภาพในแต่ละส่วนมีรายละเอียดที่แตกต่างกัน นำไปสู่วิธีการประมวลผลภาพที่มีปรับเปลี่ยนได้ (Adaptive image processing) เพื่อแก้ปัญหาข้อมูลภาพที่มีความผิดเพี้ยนเนื่องจากสัญญาณรบกวนแบบสุ่ม สามารถแสดงความสัมพันธ์ดังรูปที่ 2



รูปที่ ก.2 ระบบการสร้างภาพกลับคืนแบบปรับตัวเองได้ (Adaptive Image Restoration System)

จากรูปที่ 2 การวัดรายละเอียดบางส่วนของข้อมูลภาพจะใช้ค่าความแปรปรวนของข้อมูลภาพในส่วนย่อย (Local variance) ฟิลเตอร์ที่เปลี่ยนแปลงตามลักษณะพื้นที่ของข้อมูลภาพ (Space-variant filter) ในที่นี้กำหนดให้เป็น $h(n_1, n_2)$ และได้ออกแบบขึ้นมาเพื่อลดสัญญาณรบกวนในแถบความถี่กว้าง (Wideband) เนื่องจากพลังงานของภาพส่วนใหญ่จะรวมอยู่ที่บริเวณความถี่ต่ำ (Low-frequency region) และสัญญาณรบกวนแบบสุ่มจะกระจายอยู่ทั่วไปที่แถบความถี่กว้าง ดังนั้น $h(n_1, n_2)$ จึงมีคุณสมบัติเป็นฟิลเตอร์ของการกรองความถี่ต่ำและจากลักษณะของข้อมูลภาพจะได้ว่า บริเวณของภาพที่มีรายละเอียดน้อยจะมีการเปลี่ยนแปลงของสัญญาณเพียงเล็กน้อยจะใช้ฟิลเตอร์ของการกรองความถี่ต่ำที่มีค่ามากโดยไม่มีผลต่อองค์ประกอบของสัญญาณ ส่วนบริเวณของภาพที่มีรายละเอียดมาก เช่น ขอบภาพ เป็นต้น จะมีการเปลี่ยนแปลงของสัญญาณมากจะใช้ฟิลเตอร์ของการกรองความถี่ต่ำที่มีค่าน้อยโดยไม่มีผลต่อองค์ประกอบของสัญญาณหรือไม่ทำให้ภาพเบลอ

ความแตกต่างของอัลกอริทึม ขึ้นอยู่กับการกำหนดการวัดค่าที่ใช้แสดงรายละเอียดของภาพในส่วนย่อย ที่อาจเป็นไปได้หลายวิธี เช่น ค่าเฉลี่ย (mean) ค่าสูงสุด (max) ค่าต่ำสุด (min) ค่ากลาง (median) ค่าเบี่ยงเบนมาตรฐาน (standard deviation) และค่าความแปรปรวน (variance) เป็นต้น

2. การออกแบบฟิลเตอร์ของการกรอง Wiener ที่ปรับตัวเองได้

จากรูปที่ 1 การกรอง Wiener ต้องทราบค่า m_f , m_v , $P_f(\omega_1, \omega_2)$ และ $P_v(\omega_1, \omega_2)$ ในที่นี้จะใช้การกำหนดค่าเหล่านี้โดยการประมาณของภาพในส่วนย่อยต่างๆ แทนที่การกำหนดค่าจากข้อมูลภาพทั้งหมด

สมมุติให้สัญญาณรบกวน $v(n_1, n_2)$ มีค่าเฉลี่ย $m_v = 0$ และมีค่าความแปรปรวนเป็น σ_v^2 ดังนั้นกำลังของสเปกตรัม $P_v(\omega_1, \omega_2)$ จะหาได้จาก

$$P_v(\omega_1, \omega_2) = \sigma_v^2 \quad (3)$$

พิจารณาพื้นที่ส่วนย่อยของข้อมูลภาพ $f(n_1, n_2)$ จะได้ความสัมพันธ์ดังนี้

$$f(n_1, n_2) = m_f + \sigma_f w(n_1, n_2) \quad (4)$$

โดยที่

m_f คือค่าเฉลี่ยของภาพส่วนย่อย

σ_f คือเบี่ยงเบนมาตรฐานของ $f(n_1, n_2)$

$w(n_1, n_2)$ คือค่าสัญญาณรบกวนเฉลี่ยที่เป็นศูนย์ในหน่วยของความแปรปรวน

จากสมการที่ (4) สัญญาณ $f(n_1, n_2)$ ถูกออกแบบขึ้นมาจากผลรวมของค่าเฉลี่ยของภาพ ส่วนย่อย (Space-variant local mean) กับสัญญาณรบกวนที่มีค่าความแปรปรวนเปลี่ยนแปลงตาม พื้นที่ภาพ (Space-variant local variance) ค่าฟิลเตอร์ของการกรอง Wiener ของในภาพส่วนย่อย หาได้จาก

$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_v(\omega_1, \omega_2)} = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} \quad (5)$$

จากสมการที่ (5) ให้ $h(n_1, n_2)$ คือผลตอบสนองอิมพัลส์ที่ถูกปรับขนาด หาได้จาก

$$h(n_1, n_2) = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} \delta(n_1, n_2) \quad (6)$$

จากสมการที่ (6) และรูปที่ 1 ภาพ $p(n_1, n_2)$ ที่ผ่านการประมวลผลจากภาพในส่วนย่อยต่างๆ หาได้จาก

$$\begin{aligned} p(n_1, n_2) &= m_f + (g(n_1, n_2) - m_f) * h(n_1, n_2) \\ &= m_f + (g(n_1, n_2) - m_f) * \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} \delta(n_1, n_2) \\ &= m_f + \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} (g(n_1, n_2) - m_f) \end{aligned} \quad (7)$$

ถ้ากำหนดให้ค่า m_f และ σ_f^2 มีการเปลี่ยนแปลงในทุกๆ พิกเซลจะได้

$$p(n_1, n_2) = m_f(n_1, n_2) + \frac{\sigma_f^2(n_1, n_2)}{\sigma_f^2(n_1, n_2) + \sigma_v^2} (g(n_1, n_2) - m_f(n_1, n_2)) \quad (8)$$

จากสมการที่ (8) จะแสดงการประมวลผลใน 2 กรณีคือ ค่าเฉลี่ยของภาพในส่วนย่อย (Local mean : $m_f(n_1, n_2)$) และความแตกต่างของภาพในส่วนย่อย ($g(n_1, n_2) - m_f(n_1, n_2)$) ทั้งสองกรณีให้ผลร่วมกัน ขณะที่ค่าเฉลี่ยของภาพไม่มีการเปลี่ยนแปลง ความแตกต่างของภาพในส่วนย่อยจะมีการปรับค่าขึ้นอยู่กับขนาดของ σ_f^2 และ σ_v^2

ถ้า $\sigma_f^2 > \sigma_v^2$ ความแตกต่างของภาพในส่วนย่อย $g(n_1, n_2)$ จะสมมุติให้มีค่าเช่นเดียวกับ ข้อมูลภาพต้นแบบ $f(n_1, n_2)$ นั่นคือค่าของภาพส่วนย่อย $g(n_1, n_2)$ ไม่มีการเปลี่ยนแปลง ดังนั้น $p(n_1, n_2) \approx g(n_1, n_2)$

ถ้า $\sigma_f^2 < \sigma_v^2$ ความแตกต่างของภาพในส่วนย่อย $g(n_1, n_2)$ จะสมมุติให้เป็นค่าสัญญาณรบกวน $v(n_1, n_2)$ นั่นคือค่าของภาพในส่วนย่อย $g(n_1, n_2)$ จะมีการเปลี่ยนแปลง ดังนั้น $p(n_1, n_2) \approx m_f$

นอกจากนี้สามารถประมาณค่าเฉลี่ยของภาพในส่วนย่อย $m_f(n_1, n_2)$ จาก $g(n_1, n_2)$ ได้ดังนี้

$$\hat{m}_f(n_1, n_2) = \frac{1}{(2M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} g(k_1, k_2) \quad (9)$$

โดยที่ $(2M+1)^2$ คือ จำนวนพิกเซลของภาพในส่วนย่อย (Local region) ที่ใช้ในการประมาณค่า และ $M=1, 2, 3, \dots$ จากสมการที่ (9) นำค่า $\hat{m}_f(n_1, n_2)$ แทน $m_f(n_1, n_2)$ ในสมการที่ (8) จะได้ว่า

$$p(n_1, n_2) = g(n_1, n_2) * h(n_1, n_2) \quad (10)$$

โดยที่

$$h(n_1, n_2) = \begin{cases} \sigma_f^2 + \frac{\sigma_v^2}{(2M+1)^2}, & n_1 = n_2 = 0 \\ \frac{\sigma_f^2 + \sigma_v^2}{\sigma_f^2 + \sigma_v^2}, & n_1 = n_2 = 0 \\ \frac{\sigma_v^2}{(2M+1)^2}, & -M \leq n_1 \leq M, -M \leq n_2 \leq M \quad \text{except } n_1 = n_2 = 0 \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

จากสมการที่ (11) สามารถแสดงลักษณะของฟิลเตอร์ $h(n_1, n_2)$ ในกรณีที่ $\sigma_f^2 \gg \sigma_v^2$, $\sigma_f^2 = \sigma_v^2$ และ $\sigma_f^2 \ll \sigma_v^2$ เมื่อ $M=1$ ดังรูปที่ 3 การออกแบบฟิลเตอร์ $h(n_1, n_2)$ อย่างง่ายคือการกำหนด

ขนาดของฟิลเตอร์ให้มีขนาดเล็กๆ เช่น 3x3, 5x5, 7x7 เป็นต้น กำหนดให้ $\sigma_g^2 = \sigma_f^2 + \sigma_v^2$ ดังนั้นสามารถประมาณค่า σ_f^2 ได้จาก $g(n_1, n_2)$ ดังนี้

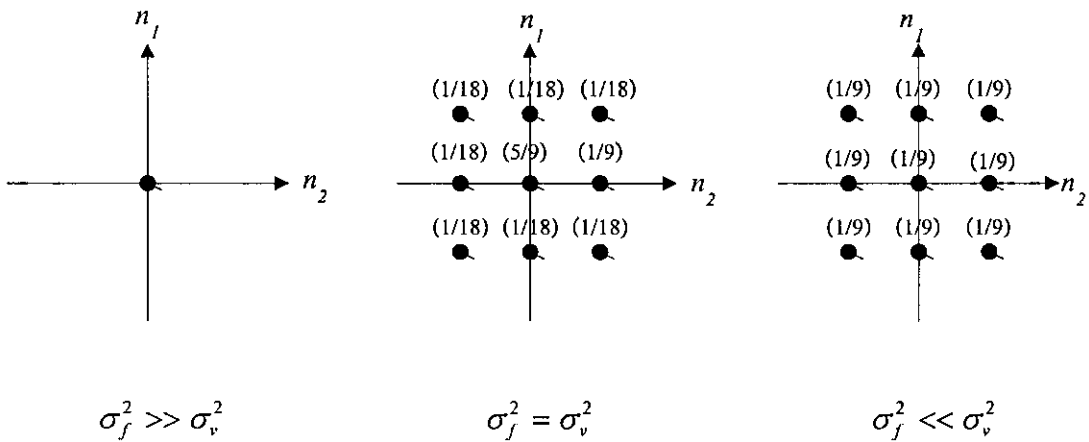
$$\hat{\sigma}_f^2(n_1, n_2) = \begin{cases} \hat{\sigma}_g^2(n_1, n_2) - \sigma_v^2, & \text{if } \hat{\sigma}_g^2(n_1, n_2) > \sigma_v^2 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

โดยที่

$$\hat{\sigma}_g^2(n_1, n_2) = \frac{1}{(2M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} (g(k_1, k_2) - \hat{m}_f(n_1, n_2))^2 \quad (13)$$

$$\sigma_v^2 = \frac{1}{M \times N} \sum_{n_1=0}^{M-1} \sum_{n_2=0}^{N-1} \hat{\sigma}_g^2(n_1, n_2) \quad (14)$$

เมื่อ $M \times N$ คือ จำนวนพิกเซลทั้งหมดของข้อมูลภาพ



รูปที่ ก.3 ผลตอบสนองอิมพัลส์ของฟิลเตอร์ที่เปลี่ยนแปลงตามพื้นที่

ภาคผนวก ข.

ผลงานที่ได้รับการตีพิมพ์

1. Y. Rangsanseri and S. Dachasilaruk, **“Integrated Speckle Reduction and Data Compression of SAR Images in Wavelet Domain,”** 2nd International Symposium on Operationalization of Remote Sensing, Enschede, Netherlands, August, 1999.
2. Y. Rangsanseri and S. Dachasilaruk, **“Filter Evaluation in Image Deblocking by Wavelet Thresholding,”** IEEE Region 10 Conference 99 (TENCON’99), Korea, September, 1999.
3. Y. Rangsanseri and S. Dachasilaruk, **“Comparative Study of Post-processing Techniques for Reducing Blockiness in Transform Coded Images,”** IEEE International Symposium on Consumer Electronic (ISCE’99), Malaysia, November, 1999.
4. S. Dachasilaruk, Y. Rangsanseri, and P. Thitimajshima, **“Application of Multiscale Edge Detection to Speckle Reduction of SAR Images,”** Asian Conference on Remote Sensing (ACRS’99), Hong Kong, November, 1999.
5. Y. Rangsanseri, P. Thitimajshima and S. Dachasilaruk, **“Edge-preserving Despeckle of SAR Imagery,”** The 28th International Symposium on Remote Sensing of Environment (ISRSE), Cape Town, South Africa, March, 2000.
6. S. Dachasilaruk and Y. Rangsanseri, **“Image deblocking using spatial Adaptive wavelet thresholding,”** Visual Communications and Image Processing 2000 (VC2000), Perth, Australia, June, 2000. (accepted)

ประวัติผู้เขียน

นางสาว ศิริพร เดชะศิลาภย์ เกิดเมื่อวันที่ 21 มีนาคม พ.ศ. 2515 ที่จังหวัดพิษณุโลก
สำเร็จการศึกษาระดับปริญญาตรี วิทยาศาสตร์บัณฑิต สาขาวิชาวัสดุศาสตร์ มหาวิทยาลัยเชียงใหม่
เมื่อปี พ.ศ. 2537