

การบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทต์
และการชดเชยข้อผิดพลาดของระบบ

MEDICAL IMAGE COMPRESSION USING VECTOR QUANTIZATION
AND SYSTEM ERROR COMPENSATION



วิทยานิพนธ์นี้สำหรับการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2558

KMITL-2015-EN-D-018-016

การบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทต์
และการชดเชยข้อผิดพลาดของระบบ

MEDICAL IMAGE COMPRESSION USING VECTOR QUANTIZATION
AND SYSTEM ERROR COMPENSATION



T139557



ธนะศักดิ์ พันธุ์ประสิทธิ์

TANASAK PHANPRASIT

ศพ.

ศ 1647

๒๐๐๘

๒. 12724476

เลขหมู่.....**139557**
เลขทะเบียน.....
วันเดือนปี.....**110 2119 2558**

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

วิทยานิพนธ์นี้สำหรับการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.2558

MEDICAL IMAGE COMPRESSION USING VECTOR QUANTIZATION
AND SYSTEM ERROR COMPRESSION



A THESIS SUBMITTED IN FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2015

KMITL-2015-EN-D-018-016

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2015


FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทต์และการชดเชยข้อผิดพลาดของระบบ
Thesis Title Medical Image Compression Using Vector Quantization and System Error Compensation
นักศึกษา นายธนะศักดิ์ พันธุ์ประสิทธิ์
รหัสประจำตัว 52610101
ปริญญา วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษาวิทยานิพนธ์ รศ.ดร.ชูชาติ ปิณฑวิรุจน์
หมายเลขวิทยานิพนธ์ KMITL-2015-EN-D-018-016

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ผศ.ดร.กิติพล	ชิตสกุล	
ผศ.ดร.ยุทธนา	คิดใจเดียว	
ศ.ดร.โกสินทร์	จำนงไทย	
รศ.ดร.สุรพันธุ์	เอื้อไพบุลย์	
รศ.ดร.ชูชาติ	ปิณฑวิรุจน์	

วัน / เดือน / ปี ที่สอบ วันศุกร์ที่ 20 กุมภาพันธ์ พ.ศ. 2558 เวลา 10.00-12.00 น.
สถานที่สอบ ณ อาคาร A ชั้น 5 ห้องประชุม 4

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิศวกรรมศาสตร์ รับรองแล้ว



(รองศาสตราจารย์ ดร. คมสัน มาลีสี)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ฉบับนี้ คณะวิศวกรรมศาสตร์
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
วันที่ 20 กุมภาพันธ์ พ.ศ. 2558

หัวข้อวิทยานิพนธ์

การบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์และการชดเชยข้อผิดพลาดของระบบ

นักศึกษา

นายธนะศักดิ์ พันธุ์ประสิทธิ์

รหัสนักศึกษา

5210101

ปริญญา

วิศวกรรมศาสตรดุษฎีบัณฑิต

สาขาวิชา

วิศวกรรมไฟฟ้า

พ.ศ.

2558

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รศ.ดร.ชูชาติ ปิณฑวิรุจน์

อาจารย์ที่ปรึกษาวิทยานิพนธ์ (ร่วม)

รศ.ดร. มนัส สังวรศิลป์

บทคัดย่อ

ในวิทยานิพนธ์ฉบับนี้ ผู้วิจัยได้นำเสนอวิธีการบีบอัดข้อมูลภาพทางการแพทย์แบบใหม่โดยใช้เว็ทเตอร์ควอนไทต์ ข้อดีของเทคนิคที่นำเสนอไม่เฉพาะทำให้อัตราการลดขนาดข้อมูลที่สูง แต่ยังรักษาคุณภาพของสัญญาณข้อมูลต่อสัญญาณรบกวน (Peak Signal to Noise Ratio, PSNR) วิธีใหม่นี้นำไปสู่สามขั้นตอน 1) นำเสนอการออกแบบ Codebook โดยใช้วิธี Discrete Wavelet Transform (DWT) Fuzzy C-Mean (FCM) และ Support Vector Machine (SVM) 2) ปรับปรุงอัตราบิต (Bit rate) โดยใช้วิธีการเข้ารหัส Huffman coding เพื่อลดการซ้ำกันของจำนวน Index 3) นำเสนอการปรับปรุงค่า PSNR ของระบบโดยใช้การชดเชยข้อผิดพลาดของระบบ (System Error Compensation, SEC) จากวิธี ที่นำเสนอสามารถปรับปรุงอัตราบิตได้ 24.00 % และ PSNR ได้ 10.96 % เมื่อเทียบกับวิธีทั่วไป

Thesis Title	Medical Image Compression Using Vector Quantization and System Error Compensation
Student	Mr. Tanasak Phanprasit
Student ID.	52610101
Degree	Doctor of Engineering
Program	Electrical Engineering
Year	2015
Thesis Advisor	Assoc. Prof. Dr. Chuchart Pintavirooj
Thesis Co-Advisor	Assoc. Prof. Dr. Manas Sangworasil

ABSTRACT

A novel medical image compression scheme based on vector quantization (VQ) is proposed in this thesis. The advantages of the technique are not only that it yields high compression ratio but also that it maintains a peak signal to noise ratio (PSNR). This new method involves three steps. First, we present a codebook design using Discrete Wavelet Transform (DWT), Fuzzy C-Means (FCM), and Support Vector Machine (SVM) algorithms. Second, we improve the bit rate using the Huffman coding theme as a method of eliminating the redundant index. Finally, we supplement the system with error compensation to improve the PSNR. With the proposed method, we are able to achieve a bit rate improvement of 24.00 % and a PSNR of 10.96 % over the conventional method

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงลงได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษา รศ.ดร.ชูชาติ ปิณฑวิรุจน์ และอาจารย์ที่ปรึกษาร่วม รศ.ดร.มนัส สังวรศิลป์ ที่ได้ให้โอกาส ความช่วยเหลือ คำชี้แนะ และช่วยแก้ปัญหาต่างๆ ตลอดจนกระทั่งให้ความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า

ขอขอบพระคุณ ศาสตราจารย์คาซุฮิโกะ ฮามามาโตะ (Kazuhiko Hamamoto) แห่ง มหาวิทยาลัยโตไก ผู้ที่ให้คำปรึกษาและคำแนะนำต่างๆ ในเรื่องของงานวิจัย

ขอขอบพระคุณ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, มหาวิทยาลัย กรุงเทพมหานคร ที่ให้การสนับสนุนในการศึกษาครั้งนี้

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร. อาทร สรรพพานิช สำหรับความช่วยเหลือและคำแนะนำต่างๆ เกี่ยวกับงานวิจัย

ขอขอบคุณ เพื่อน ผอง น้อง พี่ แห่ง BIOSIS Lab ที่เอื้อเฟื้อและให้ความช่วยเหลือตลอดระยะเวลาที่ศึกษาอยู่ ณ สำนักเรียนลาดกระบังแห่งนี้

ขอกราบขอบพระคุณบิดาและมารดาของข้าพเจ้า ผู้เป็นเสมือนดังพระอรหันต์ในใจของลูก ที่ได้ให้การเลี้ยงดู การศึกษาอบรม ให้เป็นผู้ที่ไฝ่ในทางที่ดีที่ชอบ ให้มีสติปัญญาในการฝ่าฟันอุปสรรคทุกอย่างไปได้

ท้ายสุด ผู้วิจัยขอขอบคุณ คุณแม่ ปราณีย์ พันธุ์ประสิทธิ์ และภรรยา นางดวงรัตน์ พันธุ์ประสิทธิ์ อันเป็นครอบครัวที่รักยิ่ง ที่คอยเป็นกำลังใจและเฝ้าคอยความสำเร็จของข้าพเจ้าด้วยความอดทนและเข้าใจตลอดมา

ธนะศักดิ์ พันธุ์ประสิทธิ์

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VIII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	3
1.3 สมมติฐานของการศึกษาวิจัย.....	3
1.4 แนวคิดที่ใช้ในงานวิจัย.....	4
1.5 ขอบเขตของงานวิจัย.....	4
1.6 ขั้นตอนการศึกษาวิจัย.....	4
1.7 ส่วนประกอบของวิทยานิพนธ์.....	5
บทที่ 2 ทฤษฎีและความรู้พื้นฐานที่เกี่ยวข้อง.....	6
2.1 บทนำ.....	6
2.2 การบีบอัดข้อมูล (Data Compression).....	6
2.2.1 การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression).....	6
2.2.1.1 การเข้ารหัสฮัฟฟ์แมน (Huffman Coding).....	9
2.2.1.2 การเข้ารหัสรันเลนจ์ (Run-Length Coding).....	10
2.2.1.3 การเข้ารหัสเลขคณิต (Arithmetic Coding).....	10
2.2.1.4 การเข้ารหัสยูนิเวอร์แซล (Universal Coding and Lempel-Ziv Codes).....	12
2.2.1.5 การเข้ารหัสตามระดับของบิต (Bit Plane Coding).....	12
2.2.1.6 การเข้ารหัสทำนาย (Lossless Predictive Coding).....	12
2.2.2 การบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression).....	13
2.2.2.1 การควอนไทซ์แบบสเกลาร์ (Scalar Quantization, SQ).....	13
2.2.2.2 การเข้ารหัสทำนายแบบที่มีการสูญเสีย (Lossy Predictive Coding).....	15
2.2.2.3 การเข้ารหัสโดยการแปลง (Transform Coding).....	15
2.2.2.4 การเข้ารหัสแบนย่อย (Sub-band Coding).....	16
2.2.2.5 การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ (Vector Quantization, VQ).....	16
2.2.2.5.1 การเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์.....	17
2.2.2.5.2 การถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์.....	20
2.2.2.5.3 การออกแบบตัวเก็บรหัสด้วยโมเดลทางคณิตศาสตร์.....	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.2.2.5.4 การสร้างตัวเก็บรหัสที่ดีที่สุดจากการเรียนรู้ชุดข้อมูลที่นำมา ทำการฝึก (Training Set).....	20
2.2.2.5.5 การกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส.....	23
2.2.2.5.5.1 วิธีการแบบสุ่ม (Random)	23
2.2.2.5.5.2 วิธีการแบบยูนิฟอร์มควอนไทเซอร์ (Uniform Quantize).....	23
2.2.2.5.5.3 วิธีการแตกตัว (Splitting algorithm).....	24
2.3 การแบ่งประเภทของข้อมูล (Data Classification).....	27
2.3.1 บทนำ (Introduction).....	27
2.3.2 การแบ่งกลุ่มข้อมูลแบบเป็นลำดับชั้น (Hierarchical clustering).....	28
2.3.2.1 การแบ่งลำดับชั้นแบบรวมกลุ่ม (Agglomerative hierarchical clustering).....	28
2.3.2.2 การแบ่งลำดับชั้นแบบแยกกลุ่ม (Divisive hierarchical clustering).....	30
2.3.3 การแบ่งกลุ่มข้อมูลแบบตัดเป็นส่วน (Partition clustering).....	30
2.3.3.1 ฮาร์ดพาร์ทิชัน (Hard partition).....	32
2.3.4 การประเมินความถูกต้องของการแบ่งกลุ่ม (Cluster Validity).....	37
2.4 เกณฑ์การวัดประสิทธิภาพของการบีบอัดข้อมูล.....	39
2.4.1 ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error, MSE).....	39
2.4.2 สัญญาณต่อสัญญาณรบกวนและค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด...40	
2.4.3 การวัดประสิทธิภาพของการบีบอัดข้อมูลภาพด้วยความเข้าใจ (Perceptual Measures)	41
2.4.4 อัตราบิต (bit rate)	41
2.4.5 อัตราการบีบอัดข้อมูล (Compression ratio, Cr).....	42
2.5 การแปลงเวฟเลต (Wavelet Transform).....	42
2.5.1 ความรู้เบื้องต้นเกี่ยวกับการสเกลและการเลื่อนตำแหน่งของเวฟเลต	43
2.5.2 การแปลงเวฟเลตแบบเต็มหน่วย (Discrete Wavelet Transform)	44
2.5.3 การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต (Mollet's Multiresolution analysis or Mollat' MRA)	44
2.5.4 การแปลงเวฟเลตโดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงก์ (Filter bank)	49
2.5.5 การฟิลเตอร์และการลดอัตราการซีกค่าตัวอย่างลง (Down sampling)	50
2.5.6 การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง.....	53
2.5.7 การฟิลเตอร์และการเพิ่มอัตราการซีกค่าตัวอย่าง (Up sampling).....	54

สารบัญ (ต่อ)

หน้า

2.5.8 การแปลงเวฟเลตกับข้อมูลภาพ	55
2.6 ฟัซซีซีมีน (Fuzzy C-Means).....	57
2.6.1 อัลกอริธึมของฟัซซีซีมีน	57
2.6.2 การแบ่งกลุ่มข้อมูลโดยการวัดระยะทาง	58
2.7 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM).....	61
2.7.1 การคำนวณหาไฮเปอร์เพลนที่ดีที่สุด (Optimal Hyper plane).....	64
บทที่ 3 วรรณกรรมงานวิจัยที่เกี่ยวข้อง	68
3.1 บทนำ	68
3.2 ทบทวนวรรณกรรมการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสียและการบีบอัดข้อมูลแบบที่มีการสูญเสีย	68
3.3 ทบทวนวรรณกรรมการปรับปรุงคุณภาพของภาพและการปรับปรุงอัตราบิต	69
3.4 งานวิจัยนี้เรานำเสนอการบีบอัดข้อมูลภาพแบบใหม่บนหลักการชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC)	71
บทที่ 4 รูปแบบระบบควบคุมแบบระบบเปิดและระบบปิด (Open / Close Loop Control System).....	73
4.1 บทนำ	73
4.2 ระบบควบคุมแบบระบบเปิด.....	73
4.3 ระบบควบคุมแบบระบบปิด.....	74
บทที่ 5 การจำลองระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทต์และการชดเชยค่าผิดพลาดของระบบ (Medical Image Compression Using Vector Quantization and System Error Compensation).....	77
5.1 บทนำ	77
5.2 ปัจจัยที่สำคัญของอัลกอริทึมเวกเตอร์ควอนไทต์ (The important factors of VQ Algorithm)	77
5.3 ระบบการเข้ารหัส (Coding System)	80
5.4 การออกแบบตัวเก็บรหัส (Codebook design)	81
5.5 การปรับปรุงอัตราบิต (Improved bit rate).....	89
5.6 การชดเชยค่าผิดพลาดของระบบ (System Error Compensation).....	89
บทที่ 6 ผลการจำลอง	96
6.1 ผลการจำลองค่าเฉลี่ยเวลาในการคำนวณ	97
6.2 ผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ	

เอกสารนี้เป็นเอกสาร CLC+SEC พร้อมเอกสารทั้งหมดเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ใดๆ 98

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

6.3 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC.....	98
6.4 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC.....	99
6.5 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC.....	100
6.6 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN	101
6.7 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM	101
6.8 ผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC.....	105
6.9 ผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC+SEC.....	105
6.10 ผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับวิธี CLC	106
6.11 ผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับวิธี CLC+SEC	107
6.12 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC+SEC เปรียบเทียบกับการเข้ารหัสแบบ PCA อย่างเดียว	108
บทที่ 7 อภิปรายผล.....	110
7.1 เมื่อปรับปรุงเวลา	110
7.2 เมื่อปรับปรุงเฉพาะการปรับปรุงคุณภาพ	110
7.3 เมื่อปรับปรุงเฉพาะการปรับปรุงอัตราบิต	111
7.4 เมื่อพิจารณาทั้งการปรับปรุงคุณภาพและการปรับปรุงอัตราบิต	112
บทที่ 8 สรุปผลงานวิจัยและข้อเสนอแนะ	113
8.1 สรุปผลการทดลอง.....	113
8.2 ข้อเสนอแนะ.....	114
เอกสารอ้างอิง	116
ภาคผนวก.....	120
ประวัติผู้เขียน.....	121

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1	การแบ่งเขตตามความน่าจะเป็น 10
2.2	การแบ่งกลุ่มข้อมูล..... 33
2.3	เครื่องหมายที่ใช้แทนใน Validity index 38
5.1	เวกเตอร์สามมิติ 8 แบบ..... 85
5.2	ตัวอย่างการออกแบบตัวเก็บรหัส 86
6.1	ค่าเฉลี่ยเวลาในการคำนวณ..... 97
6.2	ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC..... 98
6.3	ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC..... 99
6.4	ตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC..... 100
6.5	ตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC 100
6.6	ตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN 101
6.7	ตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM 103
6.8	บิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC 105
6.9	บิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC+SEC.... 106
6.10	บิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับ วิธี CLC 107
6.11	บิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับ วิธี CLC+SEC 108
6.12	ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC+SEC เปรียบเทียบกับ การเข้ารหัสแบบ PCA อย่างเดียว..... 109

สารบัญรูป

รูปที่	หน้า
2.1 ตัวอย่างของซอฟต์แวร์ที่ใช้สำหรับการเข้ารหัสและถอดรหัส	8
2.2 ตัวอย่างการเข้ารหัสอัฟแมน	9
2.3 การแบ่งขอบเขตตามข้อมูลที่เพิ่มขึ้น	11
2.4 การเข้ารหัสทำนาย	13
2.5 การถอดรหัสการทำนาย	13
2.6 ระดับการควอนไทซ์แบบยูนิฟอร์ม	14
2.7 บล็อกไดอะแกรมการควอนไทซ์แบบเวกเตอร์	17
2.8 การแบ่งภาพเป็นบล็อกย่อยขนาด 4x4	18
2.9 ตัวอย่างการจัดเรียงจุดภาพของบล็อกย่อยขนาดต่าง ๆ ให้อยู่ในรูปของเวกเตอร์	18
2.10 การเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์	19
2.11 โครงสร้างของตัวเก็บรหัส	19
2.12 การถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์	20
2.13 โพลีชาร์ตของอัลกอริทึม LBG	21
2.14 โพลีชาร์ตของการสร้างตัวเก็บรหัส โดยการกำหนดค่าเริ่มต้นแบบ Random	24
2.15 โพลีชาร์ตของการสร้างตัวเก็บรหัส โดยการกำหนดค่าเริ่มต้นแบบวิธีการแตกตัว (Splitting)	25
2.16 การแยกตัวเก็บรหัสให้มีขนาดเป็น 2 เทา	26
2.17 ข้อมูลดิบที่ยังไม่ได้ทำการแบ่งกลุ่ม	28
2.18 ตัวอย่างโดยทั่วไปของการแบ่งลำดับชั้นแบบ Dendrogram	29
2.19 การสุ่มข้อมูลเพื่อทำการแบ่งกลุ่ม	34
2.20 ระยะทางของข้อมูลกับข้อมูลเริ่มต้นของกลุ่ม	35
2.21 ข้อมูลที่ได้รับการแบ่งโดยอาศัย k	35
2.22 จุดศูนย์กลางของข้อมูลในแต่ละกลุ่มข้อมูล	36
2.23 การวัดระยะทางของข้อมูลกับจุดศูนย์กลางข้อมูล (Centroid)	36
2.24 กลุ่มข้อมูลใหม่	37
2.25 คลื่นเวฟเลต	42
2.26 ระยะเวลาและระนาบความถี่การเปลี่ยนแปลงช่วงเวลาและความถี่ที่สัมพันธ์กันของการแปลงเวฟเลต	44
2.27 กลุ่มของปริภูมิเวกเตอร์	46
2.28 ปริภูมิเริ่มต้นที่ถูกแผ่คลุม	47
2.29 สัญลักษ์ณ์ Down Sampling	51
2.30 ตัวกรองความถี่ตามหลักฟิลเตอร์แบงค์	52
2.31 ฟิลเตอร์แบงค์ ที่ระดับ C_{j-1}	52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
2.32 การแยกสัญญาณที่ V_3	53
2.33 การแยกสัญญาณทางด้านความถี่.....	53
2.34 สัญลักษณ์การเพิ่มอัตราการใช้ค่าเป็น 2 เท่า.....	54
2.35 ไดอะแกรมของการซีกค่าโดยใช้หลักการของฟิลเตอร์แบงค์	55
2.36 การสังเคราะห์สัญญาณจากระดับ $j-1$ ไปสู่ j	55
2.37 แผนผังของการแปลงเวฟเลต	55
2.38 ลักษณะของการแบ่งแบนด์ย่อยของภาพ.....	56
2.39 การแปลงเวฟเลตหลาย ๆ ระดับ.....	56
2.40 ลักษณะการแบ่งแบนด์ย่อยหลาย ๆ ระดับ	56
2.41 ความเป็นไปได้ของโครงสร้างข้อมูล.....	57
2.42 การนำพีชชีมาทำการแบ่งกลุ่มข้อมูล.....	58
2.43 การกำหนดข้อมูลให้อยู่ในกลุ่มข้อมูลแต่ละกลุ่ม	59
2.44 คุณสมบัติของ Fuzzy C- Partition.....	61
2.45 การแม็พเวกเตอร์เข้าสู่ Feature Space เพื่อทำการตัดแยก.....	62
2.46 ตัวอย่าง SVM ใน 2 มิติ.....	62
2.47 ตัวอย่าง SVM ใน 3 มิติ.....	63
2.48 ตัวอย่างการแม็พข้อมูลอินพุตไปยังสเปซใหม่ที่ซึ่งสามารถใช้ไฮเปอร์เพลนในการตัดแยกได้ ..64	64
2.49 ผลของการตัดแยกโดยใช้ฟังก์ชันเคอร์เนลต่างๆ	64
2.50 ไฮเปอร์เพล.....	64
4.1 การควบคุมแบบวงเปิด	73
4.2 แบบสมการวงเปิด	74
4.3 การควบคุมแบบวงปิด.....	74
4.4 การควบคุมแบบวงปิดสำหรับข้อมูลภาพ.....	75
5.1 การวิเคราะห์สัญญาณในโดเมนความถี่.....	78
5.2 การนำอัลกอริทึม SEC มาชดเชยอัลกอริทึม CLC	83
5.3 การนำเทคนิค DWT มาสร้าง 3D vector.....	83
5.4 มุมและขนาดของการเวกเตอร์สามมิติ.....	84
5.5 ตัวอย่างการออกแบบอินเตอร์พิกเซลตัวเก็บรหัส	88
5.6 บล็อกการเข้ารหัส.....	91
5.7 บล็อกการถอดรหัส	91
5.8 การเข้ารหัส PCA.....	91
5.9 การออกแบบ “basis vector”	92
5.10 การเข้ารหัสการแปลงสองระดับ DWT	93

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
5.11 การถอดรหัสการแปลงกลับสองระดับ IDWT	94
5.12 ข้อมูลภาพถ่าย PCA	94
5.13 การถอดรหัส PCA	95
6.1 ทดสอบค่า PSNR สำหรับวิธี CLC, CLC+SEC กับสองภาพทางการแพทย์ ขนาดภาพ 256*256 จุดภาพ อัตราบิต 0.50 bpp.....	103
6.2 ทดสอบค่า PSNR สำหรับวิธี CLC, CLC+SEC กับสามภาพทางการแพทย์ ขนาดภาพ 256*256 จุดภาพ อัตราบิต 0.50 bpp.....	104



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการสื่อสารข้อมูลเป็นยุคของข้อมูลข่าวสารที่มีการแข่งขันกันสูง และมีบทบาทเป็นอย่างมากต่อชีวิต ดังจะเห็นได้จากความต้องการใช้ทรัพยากรของช่องทางการสื่อสารเพื่อการจัดเก็บข้อมูลหรือการรับ-ส่งข้อมูลให้เกิดประสิทธิภาพสูงสุด ข้อมูลภาพเป็นสื่อประเภทหนึ่งที่มนุษย์นิยมนำมาใช้ในการสื่อสาร เช่น ข้อมูลภาพทางการศึกษาสำหรับการวิจัยและพัฒนาประเทศ ข้อมูลภาพทางการแพทย์ ข้อมูลภาพทางการแพทย์ เป็นต้น ด้วยเหตุนี้จึงได้มีการศึกษาวิจัยและพัฒนากระบวนการบีบอัดข้อมูลต่างๆที่จะนำมาใช้กับข้อมูลภาพเพื่อให้การสื่อสารข้อมูลเป็นไปอย่างรวดเร็วและมีประสิทธิภาพสูงสุด

การบีบอัดข้อมูล (Data Compression) เป็นสาขาหนึ่งของศาสตร์การสื่อสารข้อมูลเพื่อที่จะหาวิธีการบีบอัด (การเข้ารหัส) ของการจัดเก็บและลดเวลาในการเข้าถึงข้อมูล การบีบอัดข้อมูลจึงได้ถูกพัฒนาอย่างต่อเนื่อง และได้มีการประยุกต์ใช้งานกันมาก เช่น ภาพ เสียง วิดีโอและอื่นๆอีกมาก การเข้ารหัสภาพโดยใช้เวกเตอร์ควอนไทซ์ (Vector Quantization) [1-2] เป็นแนวทางหนึ่งในวิธีการบีบอัดของข้อมูลภาพที่ได้มีการศึกษากันมารวมถึงเป็นวิธีที่ได้ถูกนำมาศึกษาและปรับปรุงต่อจากเดิม

ในปัจจุบันเทคนิคของการบีบอัดข้อมูล แบ่งออกเป็นสองวิธีคือ 1) การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression) 2) การบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression) ในงานวิจัยนี้จะใช้เทคนิคของการบีบอัดข้อมูลแบบที่มีการสูญเสียโดยการใช้เวกเตอร์ควอนไทซ์บีบอัดข้อมูลภาพทางการแพทย์ เพราะว่าวิธีนี้มีข้อดีดังนี้ 1) เป็นเทคนิคของการบีบอัดข้อมูลที่ยาก 2) ให้อัตราส่วนของการบีบอัดข้อมูลสูง (High Compression Ratio) 3) กำหนดรูปร่างได้สะดวก (Shape Flexibility in $L-D$ Space) 4) มีจำนวนบิตต่อการสุ่มที่ต่ำ (Fractional bit per sample)

ในอดีตที่ผ่านมาเทคนิคการบีบอัดข้อมูลภาพโดยใช้เวกเตอร์ควอนไทซ์ นักวิจัยจะมุ่งเน้นที่เรื่องสำคัญสามส่วนคือ 1) การปรับปรุงคุณภาพของสัญญาณข้อมูลต่อสัญญาณรบกวน (Peak Signal to Noise Ratio, PSNR) ดังเช่น นักวิจัยกลุ่มหนึ่งในอดีตที่ผ่านมาได้ทำการปรับปรุงคุณภาพของสัญญาณข้อมูลโดยใช้ตัวเก็บรหัส (codebook) หลายระดับ (multi-resolution codebook) และใช้การแปลงเวฟเล็ตแบบเต็มหน่วย (Discrete Wavelet Transform, DWT) เพื่อลดขนาดของหน่วยความจำ [3] นักวิจัยกลุ่มถัดมาทำการปรับปรุงคุณภาพของสัญญาณข้อมูลโดยใช้ Genetic Algorithm (GA) ร่วมกับ Competitive Learning Network (GCLN) ในโดเมนของการแปลงเวฟเล็ตแบบเต็มหน่วยและเรียกวิธีการนี้ว่า WT+GGCLN [4] 2) การปรับปรุงอัตราบิต (bit rate) หรือเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(bit per pixel, bpp) นักวิจัยกลุ่มนี้นำเสนอการปรับปรุงอัตราบิตโดยใช้การสร้างตัวเก็บรหัสขนาดเล็กจากตัวเก็บรหัสที่เป็นตัวหลักที่มีความสัมพันธ์กับเวกเตอร์ข้างเคียง [5] และนักวิจัยกลุ่มถัดมาใช้ Orthonormal basis ที่ขึ้นอยู่กับประมาณภาพส่วนใหญ่กับเวกเตอร์ที่ไม่เป็นศูนย์ [6] 3) การปรับปรุงเวลา (Time in process) นักวิจัยกลุ่มนี้นำเสนอการปรับปรุงเวลาโดยใช้การค้นหาข้อมูลแบบเร็วด้วยวิธี Triangle Inequality Elimination (TIE) ร่วมกับการควบคุมหลายจุด และการแปลงเวฟเลตแบบเต็มหน่วย [7] และนักวิจัยกลุ่มถัดมานำเสนอวิธีการค้นหาข้อมูลแบบใหม่เรียกว่า Partial-Search บนหลักการสร้างด้วยกราฟ [8] จากอดีตจนถึงปัจจุบันยังไม่มีนักวิจัยวิจัยและพัฒนาเทคนิคเวกเตอร์ควอนไทซ์ได้ครบถ้วนสามเรื่องสำคัญ ดังนั้นเพื่อให้การบีบอัดของข้อมูลด้วยเทคนิคเวกเตอร์ควอนไทซ์มีประสิทธิภาพที่สูงขึ้นกว่าในอดีตที่ผ่านมา ในงานวิจัยนี้จึงนำเสนอวิธีการบีบอัดข้อมูลภาพทางการแพทย์แบบใหม่โดยใช้เวกเตอร์ควอนไทซ์ ข้อดีของเทคนิคที่นำเสนอไม่เฉพาะทำให้ อัตราการบีบอัดข้อมูลที่สูง แต่ยังรักษาคุณภาพของสัญญาณข้อมูลต่อสัญญาณรบกวน (Peak Signal to Noise Ratio, PSNR) วิธีใหม่นี้ไปสู่สามขั้นตอน 1) นำเสนอการออกแบบ Codebook โดยใช้ทฤษฎีการแปลงเวฟเลตแบบเต็มหน่วย (Discrete Wavelet Transform, DWT) ร่วมกับทฤษฎีฟัซซีซีมีน (Fuzzy C-Means) และ ใช้อัลกอริทึม (algorithm) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) 2) ปรับปรุงอัตราบิต (Bit rate) โดยใช้วิธีการเข้ารหัส Huffman coding เพื่อลดการซ้ำกันของจำนวน Index 3) นำเสนอการปรับปรุงค่า PSNR ของระบบโดยใช้การชดเชยข้อผิดพลาดของระบบ (System Error Compensation, SEC) ซึ่งเป็นวิธีการที่ใหม่ผลทำให้สามารถรักษาคุณภาพของสัญญาณข้อมูลต่อสัญญาณรบกวนสูงขึ้นและใช้อัตราบิตที่ต่ำสำหรับนำมาใช้สร้างคืนภาพทางการแพทย์อีกด้วย

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อศึกษาการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทซ์

1.2.2 เพื่อศึกษาการออกแบบการชดเชยข้อผิดพลาดของระบบ (System Error) โดยใช้การแปลงการแปลงเวฟเลตแบบเต็มหน่วย (Discrete Wavelet Transform DWT) ร่วมกับการวิเคราะห์องค์ประกอบของข้อมูล (Principal Component Analysis, PCA)

1.2.3 เพื่อศึกษาและนำเสนอการปรับปรุงคุณภาพของภาพที่สร้างกลับคืนโดยใช้การบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทซ์และการชดเชยข้อผิดพลาดของระบบ

1.2.4 เพื่อศึกษาการปรับปรุงอัตราบิตโดยใช้การบีบอัดข้อมูลภาพแบบไม่มีค่าสูญเสีย (Huffman Coding) เพื่อกำจัดการซ้ำกันของดัชนี (Index)

1.3 สมมติฐานของการศึกษาวิจัย

1.3.1 ในเรื่องการปรับปรุงคุณภาพของภาพที่สร้างกลับคืน เนื่องจากการบีบอัดข้อมูลภาพแบบเวกเตอร์ควอนไทซ์เป็นการบีบอัดข้อมูลภาพแบบที่มีค่าการสูญเสียของระบบ ดังนั้นถ้าเราชดเชยเวกเตอร์ควอนไทซ์ด้วยค่าสูญเสียของระบบก็จะทำให้ได้คุณภาพของภาพที่สร้างกลับคืนสูงขึ้น

1.3.2 1 ในเรื่องการปรับปรุงอัตราบิต เนื่องจากการซ้ำกันของดัชนี (Index) ที่เป็นตัวแทนของแต่ละบล็อกย่อย ดังนั้นถ้าเราลดการซ้ำกันของดัชนีได้โดยที่ไม่มีค่าของการสูญเสียเกิดขึ้น (Lossless Compression) ผลลัพธ์ที่ได้ก็น่าจะสามารถปรับปรุงค่าอัตราบิตได้

1.4 แนวคิดที่ใช้ในงานวิจัย

แนวคิดหลักหรือทฤษฎีที่ใช้ในการทำวิจัยประกอบด้วยการใช้หลักการดังนี้ 1) ในเรื่องการปรับปรุงคุณภาพของภาพที่สร้างกลับคืนใช้หลักการดังนี้ 1.1) ชดเชยเวกเตอร์ควอนไทซ์ด้วยค่าสูญเสียของระบบ โดยที่ค่าสูญเสียของระบบถูกสร้างขึ้นมาจากทฤษฎีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) และทฤษฎีการแปลงเวฟเล็ดแบบเต็มหน่วย (DWT) โดยจะสามารถสร้างค่าสูญเสียของระบบได้ และในงานวิจัยนี้ใช้จำนวนบิตที่ต่ำ (สี่บิต) เพราะเนื่องจากค่าสูญเสียของระบบมีค่าความละเอียดที่ต่ำ (low resolution) 1.2) สร้างตัวเก็บรหัสโดยใช้ทฤษฎีฟuzzy ซีมีน (Fuzzy C-Means) และกำหนดค่าเริ่มต้น (Initial condition) โดยใช้อัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) 2) ในเรื่องการปรับปรุงอัตราบิตใช้หลักบีบอัดข้อมูลแบบไม่มีค่าการสูญเสีย (Lossless compression) ด้วยวิธีการเข้ารหัสฮัฟแมน (Huffman Coding) จากวิธีทั้งสองดังที่กล่าวมาโดยจะสามารถปรับปรุงคุณภาพของภาพที่สร้างกลับคืนและปรับปรุงอัตราบิตให้ดีขึ้นได้

1.5 ขอบเขตของงานวิจัย

วิทยานิพนธ์ฉบับนี้ได้ทำการศึกษาวิจัยการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เวกเตอร์ควอนไทซ์และการชดเชยข้อผิดพลาดของระบบ โดยผู้วิจัยจะขอนำเสนอผลการศึกษาวิจัยเฉพาะในรูปแบบของการจำลอง (Simulations) ระบบการสร้างคืนภาพเท่านั้น โปรแกรมเขียนด้วย MATLAB โดยจะเป็นการศึกษา 1) การออกแบบตัวเก็บรหัส 2) การปรับปรุงอัตราบิต 3) การชดเชยข้อผิดพลาดของระบบ ตามลำดับ

1.6 ขั้นตอนการศึกษาวิจัย

ในการศึกษาวิจัยมีขั้นตอนการศึกษาวิจัยดังนี้

1.6.1 ศึกษาหลักการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์

1.6.2 ออกแบบระบบการบีบอัดข้อมูลภาพทางการแพทย์แบบปิดโดยใช้เว็ทเตอร์ควอนไทต์ ร่วมกับการชดเชยข้อผิดพลาดของระบบ ออกแบบตัวเก็บรหัสและสร้างคีนภาพเอาร์ทพุต

1.6.3 จำลองการสร้างคีนภาพเอาร์ทพุตด้วยระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์ร่วมกับการชดเชยข้อผิดพลาดของระบบ

1.6.4 เปรียบเทียบผลการจำลอง สรุปผลการทดลอง

1.7 ส่วนประกอบของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้แบ่งเนื้อหาออกเป็น 8 บท โดยแต่ละบทมีเนื้อหาโดยสังเขปดังต่อไปนี้

บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา, วัตถุประสงค์ของงานวิจัย, สมมติฐานของการศึกษาวิจัย, แนวคิดและทฤษฎีที่ใช้ในงานวิจัย, ขอบเขตของงานวิจัย, ขั้นตอนการศึกษาวิจัย และ ส่วนประกอบของตัวเล่มวิทยานิพนธ์

บทที่ 2 กล่าวถึงทฤษฎีและความรู้พื้นฐานที่เกี่ยวข้องกับการทำวิทยานิพนธ์ ซึ่งแบ่งออกเป็น 7 ส่วนหลักประกอบไปด้วย ส่วนแรกได้แก่ความรู้เบื้องต้นของการบีบอัดข้อมูล การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย การบีบอัดข้อมูลแบบที่มีการสูญเสีย ส่วนที่สองได้แก่ความรู้พื้นฐานของการแบ่งประเภทของข้อมูล การแบ่งกลุ่มข้อมูลแบบเป็นลำดับขั้น การแบ่งกลุ่มข้อมูลแบบตัดเป็นส่วน การประเมินความถูกต้องของการแบ่งกลุ่ม ส่วนที่สามได้แก่ความรู้เบื้องต้นเกี่ยวกับการเกณฑ์การวัดประสิทธิภาพของการบีบอัดข้อมูล ค่าความผิดพลาดกำลังสองเฉลี่ย สัญญาณต่อสัญญาณรบกวน การวัดประสิทธิภาพของการบีบอัดข้อมูลภาพด้วยความเข้าใจ อัตราบิต อัตราการบีบอัดข้อมูล ส่วนที่สี่ได้แก่ความรู้เบื้องต้นของการสเกลและการเลื่อนตำแหน่งของเวฟเลต ส่วนที่ห้าได้แก่ความรู้เบื้องต้นของการแปลงเวฟเลตแบบเต็มหน่วย การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต การแปลงเวฟเลตโดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงค์ การฟิลเตอร์และการลดอัตราการชักราค่าตัวอย่างลง การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง การฟิลเตอร์และการเพิ่มอัตราการชักราค่าตัวอย่าง การแปลงเวฟเลตกับข้อมูลภาพ ส่วนที่หกได้แก่ความรู้เบื้องต้นของอัลกอริทึมฟัชซีซีมีน การแบ่งกลุ่มข้อมูลโดยการวัดระยะทาง และส่วนที่เจ็ดได้แก่ซัพพอร์ตเวกเตอร์แมชชีน ความรู้เบื้องต้นของการคำนวณหาไฮเปอร์เพลนที่ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3 กล่าวถึงการทบทวนวรรณกรรมที่เกี่ยวข้องกับงานวิจัย ทั้งในส่วนที่เกี่ยวข้องกับการทบทวนวรรณกรรมการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสียและการบีบอัดข้อมูลแบบที่มีการสูญเสีย การปรับปรุงคุณภาพของภาพและการปรับปรุงอัตราบิต

บทที่ 4 กล่าวถึงรูปแบบระบบควบคุมเปิดและระบบปิด (Open / Close Loop Control System)

บทที่ 5 กล่าวถึงการจำลองระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์และการชดเชยข้อผิดพลาดของระบบ ปัจจัยที่สำคัญของอัลกอริทึมเว็ทเตอร์ควอนไทต์ การออกแบบตัวเก็บรหัส การปรับปรุงอัตราบิต การชดเชยข้อผิดพลาดของระบบ

บทที่ 6 กล่าวถึงผลการจำลองด้วยระบบการปรับปรุงเวลา การปรับปรุงอัตราบิต การปรับปรุงคุณภาพของภาพซึ่งจะประกอบไปด้วย กรณีที่ 1) ระบบเว็ทเตอร์ควอนไทต์ทดสอบกับ ภาพมาตรฐานและตัวเก็บรหัสสร้างจากภาพแต่ละภาพ ภาพมาตรฐานและตัวเก็บรหัสสร้างจากภาพมาตรฐานสี่ภาพ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพแต่ละภาพ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพทางการแพทย์สามภาพ กรณีที่ 2) ระบบเว็ทเตอร์ควอนไทต์และการชดเชยข้อผิดพลาดของระบบทดสอบกับ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพแต่ละภาพ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพทางการแพทย์สามภาพ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพทางการแพทย์สองภาพ ภาพทางการแพทย์และตัวเก็บรหัสสร้างจากภาพทางการแพทย์แต่ละภาพ เปรียบเทียบกับการบีบอัดข้อมูลโดยการใช้ Principal Component Analysis (PCA) อย่างเดียว

บทที่ 7 เป็นการอภิปรายผล ในส่วนของ เมื่อปรับปรุงเวลา เมื่อปรับปรุงเฉพาะการปรับปรุงคุณภาพ เมื่อปรับปรุงเฉพาะการปรับปรุงอัตราบิต เมื่อพิจารณาทั้งการปรับปรุงคุณภาพและการปรับปรุงอัตราบิต

บทที่ 8 เป็นการสรุปผลงานวิจัย และข้อเสนอแนะ

บทที่ 2

ทฤษฎีและความรู้พื้นฐานที่เกี่ยวข้อง

2.1 บทนำ

ในบทนี้จะเป็นการกล่าวถึงทฤษฎีและความรู้พื้นฐานในเบื้องต้นของเรื่องที่เกี่ยวข้องกับวิทยานิพนธ์ โดยจัดแบ่งเนื้อหาออกเป็น 6 ส่วนหลักๆ ดังนี้

- ส่วนแรก ได้แก่ความรู้เบื้องต้นของการบีบอัดข้อมูล (Data compression), การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression), การบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression)

- ส่วนที่สอง ได้แก่ความรู้พื้นฐานของการแบ่งประเภทของข้อมูล (Data Classification), การแบ่งกลุ่มข้อมูลแบบเป็นลำดับขั้น (Hierarchical clustering), การแบ่งกลุ่มข้อมูลแบบตัดเป็นส่วน (Partition clustering), การประเมินความถูกต้องของการแบ่งกลุ่ม (Cluster Validity)

- ส่วนที่สาม ได้แก่ความรู้เบื้องต้นเกี่ยวกับเกณฑ์การวัดประสิทธิภาพของการบีบอัดข้อมูล, ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error, MSE), สัญญาณต่อสัญญาณรบกวนและค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด อัตราบิต (bit rate) อัตราการบีบอัดข้อมูล (Compression ratio, Cr)

- ส่วนที่สี่ ได้แก่ความรู้พื้นฐานของการแปลงเวฟเลต (Wavelet Transform), ความรู้เบื้องต้นเกี่ยวกับการสเกลและการเลื่อนตำแหน่งของเวฟเลต, การแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform), การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต (Mollet's Multi resolution analysis or Mollat' MRA), การแปลงเวฟเลตโดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงก์ (Filter bank), การฟิลเตอร์และการลดอัตราการซีกค่าตัวอย่างลง (Down sampling), การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง, การฟิลเตอร์และการเพิ่มอัตราการซีกค่าตัวอย่าง (Up sampling) และแปลงเวฟเลตกับข้อมูลภาพ

- ส่วนที่ห้า ได้แก่ความรู้พื้นฐานของทฤษฎีฟuzzy ซีมีน (Fuzzy C-Means), ทฤษฎีฟuzzy ซีมีน, การแบ่งกลุ่มข้อมูลโดยการวัดระยะทาง และอัลกอริทึมของ Fuzzy C-Means (FCM)

- ส่วนที่หก ได้แก่ความรู้พื้นฐานของอัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) และการคำนวณหาไฮเปอร์เพลนที่ดีที่สุด (Optimal Hyper plane)

2.2 การบีบอัดข้อมูล (Data compression)

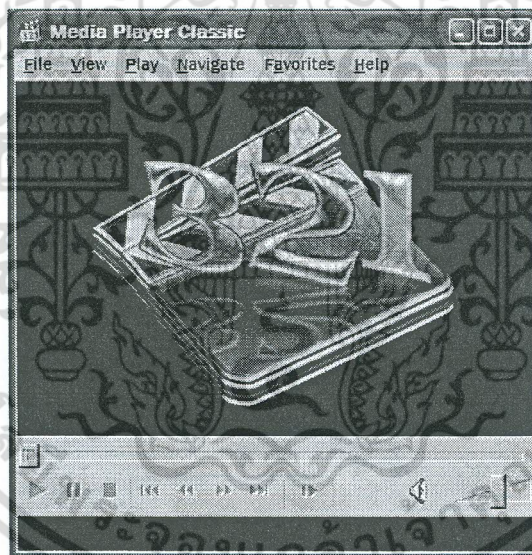
ข้อมูลภาพเชิงตัวเลข (Digital Image) ประกอบด้วยชุดของบิตข้อมูล ตัวอย่างเช่น ภาพระดับความ เป็นเทาที่ประกอบด้วยอาเรย์ของข้อมูลที่ได้มาจากการสุ่มขนาด 256×256 ซึ่งในแต่ละกลุ่มจะ ประกอบด้วยค่าความเข้มจำนวน 8 บิต ดังนั้นจึงต้องการเนื้อที่สำหรับทำการเก็บข้อมูลจำนวน 524,288 บิต หรือ 65,536 ไบท์ ข้อมูลการสุ่มที่เราเรียกกง่าย ๆ เป็นจุดภาพ (Pixel) ซึ่งกรณีที่เป็น ภาพสีขนาดเดียวกันนี้ ก็จะใช้หน่วยความจำในการจัดเก็บข้อมูลเพิ่มขึ้นเป็น 3 เท่า เนื่องจากความ ต้องการในการส่งข้อมูลภาพจะเป็นสัดส่วนกับจำนวนของบิตข้อมูล โดยในการส่งภาพเชิงตัวเลข ตามปกตินั้นจะใช้ความกว้างแถบ (Bandwidth) ที่กว้าง และใช้ระยะเวลาในการส่งภาพ จำนวน 524,288 บิต จะใช้เวลา 54.6 วินาที สำหรับการส่งข้อมูลผ่านโมเด็มที่ความเร็วในการส่ง 9,600 บิต วินาที ซึ่งภาพเชิงตัวเลขอื่นๆ อาจจะมีขนาดใหญ่กว่านี้ เช่น ภาพที่ได้จากรังสีเอกซ์ขนาด $2,048 \times 2,048$ จุดภาพ ที่ค่าความเข้ม 12 บิต ก็จะใช้เวลา 87 นาทีสำหรับการส่งที่ความเร็ว 9,600 บิต วินาที

ค่าใช้จ่ายที่เป็นต้นทุนจะมีความสัมพันธ์กับการจัดเก็บลงหน่วยความจำ และการส่งข้อมูล ภาพ จึงทำให้การบีบอัดข้อมูลมีความจำเป็นที่สำคัญ การบีบอัดข้อมูลภาพเป็นกระบวนการลดจำนวน บิตที่ใช้ในการแสดงภาพ การบีบอัดข้อมูลภาพจะคำนึงถึงข้อมูลทางสถิติของภาพร่วมกับความเหมือน จริงที่ยอมรับได้ด้วยการสังเกตด้วยสายตาของมนุษย์ โดยการลดจำนวนของบิตที่ต้องการใช้ในการ แสดงภาพ ทำให้การบีบอัดข้อมูลภาพสามารถนำไปประยุกต์ใช้ในงานอื่นๆ ได้อย่างมีประสิทธิภาพ การจัดเก็บข้อมูล สามารถพัฒนาให้ดีขึ้นได้ด้วยการบีบอัดข้อมูล เพื่อทำการลดความต้องการ หน่วยความจำ การส่งภาพสามารถพัฒนาให้ดีขึ้นได้ด้วยการลดความต้องการความกว้างแถบ หรือการ เพิ่มความเร็วในการส่งให้กับความกว้างแถบที่ใช้อยู่ การบีบอัดข้อมูลยังสามารถออกแบบให้มีการส่ง แบบโปรแกรมมิ่ง ซึ่งสามารถนำไปใช้ประโยชน์ได้ เมื่อผู้สังเกตต้องการดูภาพที่มีรายละเอียดต่ำเร็ว ที่สุดเท่าที่จะเป็นไปได้ เพื่อที่จะสามารถทำนายภาพที่ถูกส่งต่อมาได้ เช่น การแสดงภาพในการสื่อสาร อินเทอร์เน็ต

เทคนิคในการบีบอัดข้อมูลภาพสามารถแบ่งออกได้เป็น 2 กลุ่ม คือการบีบอัดข้อมูลแบบที่ไม่ มีการสูญเสีย (Lossless Compression) และการบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression) การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสียจะสามารถนำข้อมูลกลับมาได้ใหม่อย่าง สมบูรณ์ ซึ่งมีชื่อเรียกต่างๆ กัน คือ การเข้ารหัสแบบไม่มีการสูญเสีย (Noiseless) การเข้ารหัสที่ สามารถสร้างกลับมาได้ (Invertible) การเข้ารหัสเอนโทรปี (Entropy Coding) และการบีบอัดข้อมูล (Data Compression) การบีบอัดข้อมูลโดยปราศจากการสูญเสียมีความจำเป็นสำหรับข้อมูลบาง ชนิด เช่น ข้อมูลทางการธนาคาร และข้อมูลโปรแกรมคอมพิวเตอร์ เพราะถ้าแม้เพียงบิตเดียวของ ข้อมูลมีการเปลี่ยนแปลงไปในระหว่างกระบวนการบีบอัดข้อมูล ก็สามารถทำให้ข้อมูลเหล่านี้เกิดความ เสียหายขึ้นได้อย่างมาก ตัวอย่างของการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย เช่น โปรแกรม Compress และ Compact บนระบบยูนิกซ์ (UNIX) ซึ่งโดยทั่วไปจะมีอัตราการบีบอัดข้อมูลภาพแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ทางการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม่มีการสูญเสียในช่วงตั้งแต่ 2:1 ถึง 4:1 การบีบอัดข้อมูลแบบไม่มีการสูญเสียที่ใช้เป็นมาตรฐานสากล ได้แก่การบีบอัดข้อมูลแบบ JPEG-LS (lossless/near-lossless compression standard) [9] จะมี อัตราการบีบอัดข้อมูลสำหรับภาพสี ที่มีความซับซ้อนพอสมควรประมาณ 2:1 ดังนั้นเพื่อให้อัตราการ บีบอัดข้อมูลที่สูงขึ้น จึงจำเป็นที่จะต้องใช้การบีบอัดข้อมูลแบบที่มีการสูญเสีย การบีบอัดข้อมูลแบบ ที่มีการสูญเสียนี้จะทำให้เกิดความเพี้ยน (Distortion) กับข้อมูลที่ได้ถูกสร้างกลับขึ้นมาใหม่ หัวใจใน การออกแบบระบบการบีบอัดข้อมูลแบบที่มีการสูญเสีย คือการทำให้มีค่าความเพี้ยนต่ำที่สุดสำหรับ อัตราบิต (Bit Rate) หรือ อัตราการบีบอัดข้อมูล (Compression Ratio) หนึ่งๆ ซึ่งวิธีการบีบอัด ข้อมูลในแบบต่าง ๆ สามารถที่จะนำมารวมกันได้ เรียกว่าเป็นการบีบอัดข้อมูลแบบผสมผสาน (Hybrid Compression) ตัวอย่างซอฟต์แวร์ที่ใช้สำหรับการบีบอัดสื่อหรือข้อมูลดิจิทัลดังเช่นซอฟต์แวร์ ชื่อ “CODEC” ย่อมาจาก Compressor - DE Compressor หมายถึงกลไกทางฮาร์ดแวร์หรือ ซอฟต์แวร์ที่ใช้สำหรับการเข้ารหัสและถอดรหัสข้อมูลหรือสัญญาณดิจิทัล โดยข้อมูลที่ถูกบีบอัดด้วย ซอฟต์แวร์ CODEC อาจอยู่ในรูปแบบต่างกัน เช่น AVI , Ogg , ASF หรือ MP4 ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 ตัวอย่างของซอฟต์แวร์ที่ใช้สำหรับการถอดรหัส

2.2.1 การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression)

การบีบอัดข้อมูลแบบที่ไม่มีการสูญเสียนั้น มักจะนำเอาประโยชน์ของส่วนเกินของข้อมูล (Redundancy) มาใช้ซึ่งวิธีการหนึ่งที่จะนำมาใช้วัดส่วนเกินของข้อมูล คือการหาความน่าจะเป็นของการกระจายของค่าความเข้มของจุดภาพ โดยที่เอนโทรปีอันดับที่หนึ่ง (First Order Entropy) สามารถหาได้จาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

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

เมื่อ $p(i)$ คือความน่าจะเป็นของค่าความเข้ม i ของจุดภาพที่เกิดขึ้น

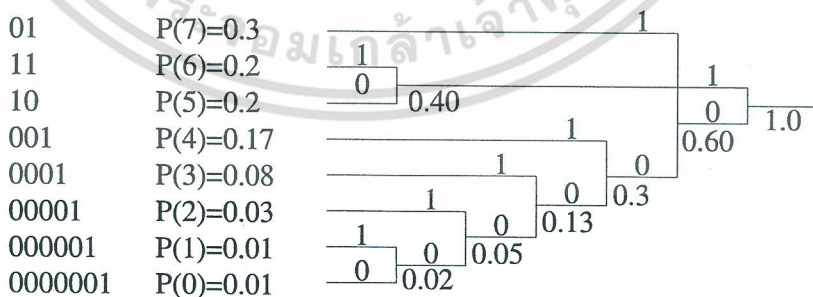
เอนโทรปีจะใช้อธิบายความไม่แน่นอนหรือ Randomness ของข้อมูล และจะแสดงค่าเฉลี่ยของจำนวนบิตที่จะใช้แทนข้อมูล ภาพจะถูกลดข้อมูลโดยการเข้ารหัสค่าความเข้มของจุดภาพที่เกิดขึ้นบ่อยด้วยจำนวนบิตที่น้อย และเข้ารหัสค่าความเข้มของจุดภาพที่เกิดขึ้นน้อยกว่าด้วยจำนวนบิตที่มาก ซึ่งโดยเฉลี่ยแล้วจะเป็นการลดจำนวนบิตที่ใช้แทนข้อมูล

2.2.1.1 การเข้ารหัสฮัฟแมน (Huffman Coding)

การเข้ารหัสฮัฟแมน (Huffman: 1952) เป็นการบีบอัดข้อมูลที่ไม่มีการสูญเสียวิธีหนึ่ง หลักการที่สำคัญของการเข้ารหัสฮัฟแมนคือ ข้อมูลที่มีค่าความน่าจะเป็นสูงกว่าจะถูกเข้ารหัสด้วยจำนวนบิตที่น้อยกว่า

การเข้ารหัสฮัฟแมน [10] สามารถอธิบายได้ดังนี้

- 1) ทำการนับความถี่เพื่อหาความน่าจะเป็นของข้อมูลทุกตัว และเรียงลำดับข้อมูลตามความน่าจะเป็นจากน้อยไปหามาก
- 2) นำความน่าจะเป็นที่มีค่าต่ำสุด 2 ค่ามารวมกันเพื่อสร้างเป็นโนด (Node) ใหม่ แล้วนำความน่าจะเป็นใหม่ที่ได้อาเรียงลำดับใหม่ร่วมกับความน่าจะเป็นเดิมที่เหลือ
- 3) กำหนดค่าไบนารีให้กับโนดคู่ที่ถูกกำหนดให้เป็นเส้นทางของโนดใหม่นั้น โดยโนดหนึ่งมีค่าเป็น 0 ส่วนอีกโนดมีค่าเป็น 1
- 4) ทำซ้ำขั้นตอนที่ 2 และ 3 จนกระทั่งได้ความน่าจะเป็นสุดท้ายเป็น 1.0 ดังแสดงในรูปที่ 2.2



รูปที่ 2.2 ตัวอย่างการเข้ารหัสฮัฟแมน

หลังจากที่ได้รับรหัส Huffman แล้ว การเข้ารหัสและการถอดรหัสสามารถทำได้โดยใช้การเปิดตาราง (Look-Up Table) และรหัสที่ได้ก็สามารถที่จะทำการถอดรหัสได้ทันทีโดยไม่ต้องอ้างอิงถึงข้อมูลตัวอื่น ๆ

2.2.1.2 การเข้ารหัสรันเลนจ์ (Run-Length Coding)

เป็นวิธีการที่ง่ายที่สุดในการเข้ารหัสข้อมูลไบนารีที่มีมักจะเกิดกลุ่มของ 0 และ 1 เช่น ในข้อมูลภาพหรือข้อความสำหรับการส่งแฟกซ์ (Facsimile) หรืออาจใช้กับภาพที่จุดภาพที่อยู่ใกล้กันมีระดับการเปลี่ยนแปลงระดับความเข้มไม่มาก และมีการซ้ำเป็นจำนวนมาก โดยจะเก็บจำนวนของจุดภาพที่ซ้ำกันแทนที่จะเก็บค่าระดับความเข้มของจุดภาพเหล่านั้น [11]

2.2.1.3 การเข้ารหัสเลขคณิต (Arithmetic Coding)

ในการเข้ารหัสเลขคณิตนั้นข่าวสารต้นแบบจะถูกแทนด้วยช่วงระยะห่างของจำนวนจริงระหว่าง 0 และ 1 โดยข้อมูลแต่ละตัวของข่าวสารจะทำให้ช่วงระยะห่างนั้นแคบลง ซึ่งระยะห่างที่เล็กลงนี้จะทำให้ใช้จำนวนของบิตมากขึ้น ตัวข้อมูลที่ต่อเนื่องกันของข่าวสารจะลดขนาดของระยะห่างตามความน่าจะเป็นของข้อมูล ข่าวสารที่มีความน่าจะเป็นสูงจะทำให้ระยะห่างถูกลดลงได้น้อยกว่าข่าวสารที่มีความน่าจะเป็นต่ำ ดังนั้นจึงทำให้ข่าวสารที่มีความน่าจะเป็นสูงใช้จำนวนบิตในการเข้ารหัสที่น้อยกว่า [12], [13]

กระบวนการจะเริ่มจากข้อมูลต้นแบบที่ไม่ได้มีการเรียงลำดับความน่าจะเป็น จากนั้นจะทำการแบ่งช่วงระยะห่างระหว่างระหว่าง 0 กับ 1 เป็นช่วงย่อย ๆ ตามความน่าจะเป็นสะสม (Cumulative Probability) เช่น กำหนดให้ข้อมูลต้นแบบคือ {A, B, C, D, !} มีความน่าจะเป็น {0.3, 0.1, 0.2, 0.3, 0.1} ตามลำดับ การแบ่งช่วงระยะห่างระหว่าง 0 กับ 1 แสดงไว้ในตารางที่ 2.1

ตารางที่ 2.1 การแบ่งเขตตามความน่าจะเป็น

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

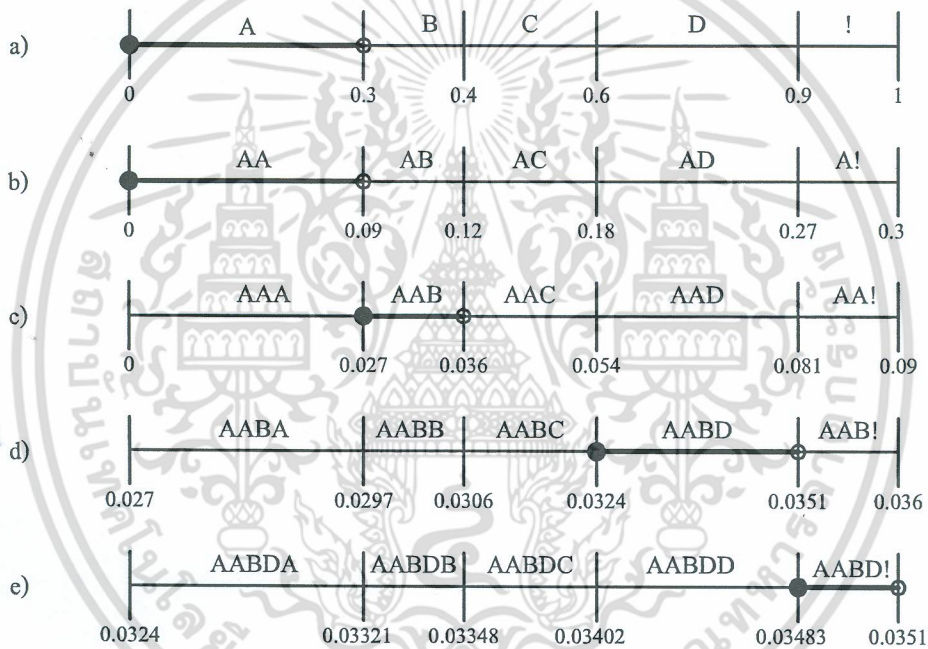
ตัวอย่างการเข้ารหัส AABD! แสดงไว้ในรูปที่ 2.3 โดยมีขั้นตอนดังนี้

- 1) A ตัวแรกจะลดขอบเขตของระยะห่างระหว่าง 0 กับ 1 ลงเป็น [0, 0.3) ตามรูปที่ 2.3 a)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2) A ตัวที่สองจะลดขอบเขตของระยะห่างระหว่าง 0 กับ 0.3 ลงเป็น $[0, 0.09)$ ตามรูปที่ 2.3 b)
- 3) B จะลดขอบเขตระยะห่างระหว่าง 0 กับ 0.09 ลงเป็น $[0.027, 0.36)$ ตามรูปที่ 2.3 c)
- 4) D จะลดขอบเขตระยะห่างระหว่าง 0.027 กับ 0.036 ลงเป็น $[0.0324, 0.0351)$ ตามรูปที่ 2.3 d)
- 5) ! จะลดขอบเขตระยะห่างระหว่าง 0.0324 กับ 0.0351 ลงเป็น $[0.03483, 0.0351)$ ตามรูปที่ 2.3 e)

โดยที่ $[0, 0.3)$ หมายถึงระยะห่างที่อยู่ในช่วงมากกว่าหรือเท่ากับ 0 และน้อยกว่า 0.3 เป็นต้น ดังนั้นข้อมูล AABD! สามารถเข้ารหัสได้โดยการแทนด้วยช่วงระยะห่างหรือตัวเลขจำนวนใด ๆ ที่อยู่ในช่วง $[0.03483, 0.0351)$



รูปที่ 2.3 การแบ่งขอบเขตตามข้อมูลที่เพิ่มขึ้น

สมมติว่าตัวถอดรหัสได้รับรหัสข้อมูลเป็น ! = 0.0350 ดังนั้นขบวนการถอดรหัสจะมีขั้นตอนดังต่อไปนี้

- 1) เริ่มจากการแบ่งช่วง $[0, 1)$ ตามความน่าจะเป็น ! = 0.0350 อยู่ในช่วง $[0, 0.3)$ ซึ่งเป็นช่วงตัวอักษร A ดังนั้นจะได้ตัวอักษรแรกคือ A จากนั้นจะลดขอบเขตของ $[0, 1)$ ลงเป็น $[0, 0.3)$
- 2) แบ่งช่วง $[0, 0.3)$ ตามความน่าจะเป็น แล้วพบว่า ! = 0.0350 อยู่ในช่วง $[0, 0.09)$ ซึ่งเป็นช่วงตัวอักษร A ดังนั้นจะได้ตัวอักษรตัวที่สองคือ A จากนั้นจะลดขอบเขตของ $[0, 0.3)$ ลงเป็น $[0, 0.09)$

3) แบ่งช่วง $[0, 0.09)$ ตามความน่าจะเป็น แล้วพบว่า $! = 0.0350$ อยู่ในช่วง $[0.027, 0.036)$ ซึ่งเป็นช่วงตัวอักษร B ดังนั้นจะได้ตัวอักษรตัวที่สามคือ B จากนั้นจะลดขอบเขตของ $[0, 0.09)$ ลงเป็น $[0.027, 0.036)$

4) แบ่งช่วง $[0.027, 0.036)$ ตามความน่าจะเป็น แล้วพบว่า $! = 0.0350$ อยู่ในช่วง $[0.0324, 0.0351)$ ซึ่งเป็นช่วงตัวอักษร D ดังนั้นจะได้ตัวอักษรตัวที่สามคือ D จากนั้นจะลดขอบเขตของ $[0.027, 0.036)$ ลงเป็น $[0.0324, 0.0351)$

5) แบ่งช่วง $[0.0324, 0.0351)$ ตามความน่าจะเป็น แล้วพบว่า $! = 0.0350$ อยู่ในช่วง $[0.03483, 0.0351)$ ซึ่งเป็นช่วงตัวอักษร! ดังนั้นจะได้ตัวอักษรตัวที่สามคือ ! จากนั้นจะลดขอบเขตของ $[0.0324, 0.0351)$ ลงเป็น $[0.03483, 0.0351)$

2.2.1.4 การเข้ารหัสยูนิเวอร์แซล (Universal Coding and Lempel-Ziv Codes)

เป้าหมายของการเข้ารหัสยูนิเวอร์แซล (Universal Coding) คือการบีบอัดขนาดของข้อมูลที่ไม่ทราบข้อมูลทางสถิติ ซึ่งอัลกอริทึมจะต้องปรับตัวไปตามข้อมูล ตัวอย่างที่สำคัญของการเข้ารหัสยูนิเวอร์แซลคืออัลกอริทึม Lempel-Ziv [14] ลำดับของรหัสข้อมูลจะใช้ลักษณะของตารางของลำดับ ซึ่งถูกปรับปรุงในระหว่างกระบวนการเข้ารหัสโดยการปรับตัวตามข้อมูลอินพุต ดังนั้นจึงสามารถเรียกได้ว่าเป็นการเข้ารหัสของตารางที่ปรับตัว (Adaptive Dictionary Encoding) ได้ โดยปกติแล้ว การเข้ารหัสเลขคณิตมักจะมีข้อเสียเรื่องความยุ่งยากในการคำนวณ ในขณะที่การเข้ารหัส Lempel-Ziv จะมีความง่ายกว่า

2.2.1.5 การเข้ารหัสตามระดับของบิต (Bit Plane Coding)

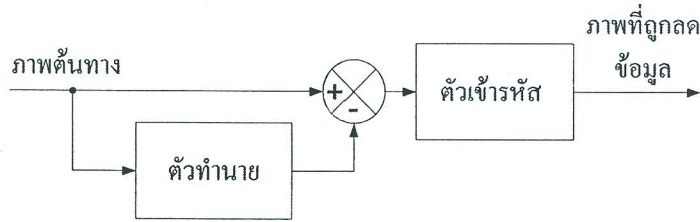
สมมติว่าแต่ละค่าของจุดภาพในภาพระดับเทาภาพหนึ่งถูกแทนด้วยข้อมูล k บิต ดังนั้นเราสามารถสร้างภาพไบนารีจำนวน k ภาพ ซึ่งบิตแต่ละตำแหน่งจะสร้างได้ 1 ภาพ โดยภาพที่ได้มาจากบิตตำแหน่งที่มีนัยสำคัญสูงกว่าจะประกอบด้วยข่าวสารที่สำคัญกว่า และจะมี Spatial Redundancy ที่สูงกว่า ดังนั้นจึงทำให้ง่ายต่อการบีบอัดข้อมูล วิธีการเข้ารหัสตามระดับของบิตจะอาศัยการระดับเทาของภาพในการเพิ่มความสัมพันธ์ (Correlation) ของจุดภาพ จากนั้นจึงนำการเข้ารหัสแบบรันเลนจ์หรือการเข้ารหัสเลขคณิตมาใช้ในการบีบอัดข้อมูล ซึ่งรายละเอียดสามารถศึกษาได้จาก [12], [13]

2.2.1.6 การเข้ารหัสทำนาย (Lossless Predictive Coding)

ปกติสำหรับภาพโดยทั่วไปแล้ว จุดภาพที่อยู่ใกล้กันจะมีการสัมพันธ์ต่อกันสูง (High Correlation) นั้นหมายความว่าข้อมูลของจุดภาพสามารถหาได้โดยการตรวจสอบจุดภาพที่อยู่โดยรอบ ๆ ของจุดภาพที่เราสนใจ ด้วยคุณสมบัตินี้เองที่ได้นำมาใช้ในการเข้ารหัสแบบทำนายโดยพยายามที่จะทำการคำนวณหาค่าจุดภาพที่อยู่รอบ ๆ [12], [13]

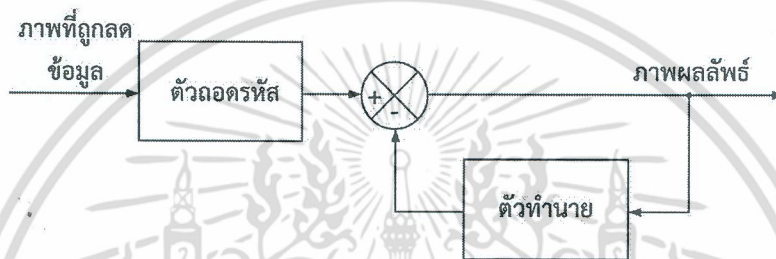
การเข้ารหัสแบบทำนายจะทำการบีบอัดข้อมูลส่วนเกินระหว่างจุดภาพ (Interpixel Redundancy) แล้วเข้ารหัสเฉพาะข่าวสารใหม่ของแต่ละจุดภาพ โดยข่าวสารใหม่ของจุดภาพนี้สามารถหาได้จากความแตกต่างระหว่างค่าจริงของจุดภาพกับค่าที่ได้มาจากการทำนาย ดังสามารถแสดงได้ในรูปที่ 2.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 การเข้ารหัสทำนาย

ส่วนการถอดรหัสนั้นจะนำภาพที่ถูกบีบอัดข้อมูลแล้วมาบวกกับค่าของตัวทำนายที่มีฟังก์ชันของตัวทำนายเช่นเดียวกันกับตัวเข้ารหัส ซึ่งผลที่ได้นั้นจะเรียกว่าภาพผลลัพธ์ดังแสดงในรูปที่ 2.5



รูปที่ 2.5 การถอดรหัสการทำนาย

2.2.2 การบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression)

การบีบอัดข้อมูลแบบที่มีการสูญเสียจะทำให้เกิดความเพี้ยนของข้อมูลขึ้น แต่มีอัตราการลดข้อมูลที่สูงกว่าการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย ซึ่งมีอยู่หลายวิธีดังนี้

2.2.2.1 การควอนไทซ์แบบสเกลาร์ (Scalar Quantization, SQ)

การควอนไทซ์แบบสเกลาร์เป็นวิธีการที่ง่ายที่สุดของการบีบอัดข้อมูลแบบที่มีการสูญเสีย โดยการควอนไทซ์จะทำการตรวจสอบข้อมูลอินพุตที่ได้รับ และเลือกค่าประมาณที่ได้จากฐานข้อมูลที่กำหนดไว้ล่วงหน้าให้ใกล้เคียงที่สุด โดยปกติแล้วข้อมูลอินพุตจะอยู่ในรูปอนาล็อกหรือสัญญาณต่อเนื่อง ส่วนเอาต์พุตที่ได้จะอยู่ในรูปดิจิทัล โดยสามารถนิยามการควอนไทซ์แบบสเกลาร์ได้ว่าเป็นการ Mapping $Q: R \rightarrow C$ เมื่อ R คือค่าจำนวนจริงและ $C = \{y_i; i = 1, 2, \dots, N\}$ เป็นชุดของข้อมูลเอาต์พุต y_i หรือตัวเก็บรหัส (Codebook) ที่มีขนาดเป็น N

ในการสร้างตัวควอนไทซ์ขนาด N จุด จะต้องทำการแบ่งเส้นจริง R ออกเป็น N ส่วน เรียกว่า R_i เมื่อ $i = 1, 2, \dots, N$ และค่าของส่วนที่ i สามารถหาได้จาก $i \neq j$

$$R_i = \{x \in R; Q(x) = y_i\} \equiv Q^{-1}(y_i) \tag{2.2}$$

โดยที่แต่ละส่วนจะมีคุณสมบัติดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\bigcup_{i=1}^N R_i = R \text{ และ } R_i \cap R_j = \emptyset \text{ เมื่อ } i \neq j \quad (2.3)$$

การควอนไทซ์จะแบ่งระดับออกเป็นช่วง ๆ โดยที่แต่ละช่วงจะแทนระดับของข้อมูลที่เป็นค่าเดียว ๆ ซึ่งสามารถแบ่งได้เป็น 2 ประเภท คือ

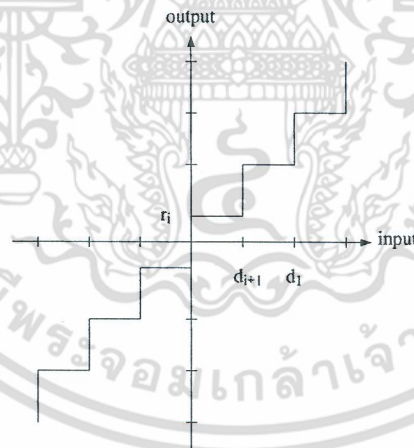
1) การควอนไทซ์แบบยูนิฟอร์ม (Uniform Quantization) เป็นการควอนไทซ์ที่มีรูปแบบที่ง่ายที่สุด โดยการแบ่งย่านของข้อมูลออกเป็นช่วงเท่า ๆ กัน ดังแสดงในรูปที่ 2.6 ในการออกแบบตัวควอนไทซ์จะมี d_i เป็นระดับการตัดสินใจ (decision level) และ r_i เป็นระดับของข้อมูลผลลัพธ์ (Reconstruction level) เมื่อ $i=0,1,2,\dots,L$ โดยที่

$$r_i = \frac{d_i + d_{i+1}}{2} \quad (2.4)$$

และ จะได้ช่วงห่างของการควอนไทซ์ (Δ) เป็น

$$\Delta = d_{i+1} - d_i \quad (2.5)$$

ดังนั้น จะเห็นได้ว่าการควอนไทซ์แบบยูนิฟอร์มมีค่าของ Δ เท่ากันตลอดย่านของข้อมูลอินพุต



รูปที่ 2.6 ระดับการควอนไทซ์แบบยูนิฟอร์ม

2) การควอนไทซ์แบบนอนยูนิฟอร์ม (Non-Uniform Quantization) เป็นการออกแบบการควอนไทซ์ให้มีความเพี้ยนเกิดขึ้นน้อยที่สุด โดยอัลกอริทึม Least Squares Quantization ได้ถูกพัฒนาขึ้นโดย Lloyd [15] และ Max [16] ได้วิเคราะห์สัญญาณรบกวนจากการควอนไทซ์โดยใช้ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error, MSE) และได้พยายามลดสัญญาณรบกวนนี้ลงโดยกำหนดความน่าจะเป็นของความหนาแน่นสัญญาณในช่วงหนึ่งมีค่าไม่คงที่ โดยที่ความเพี้ยนทั้งหมดที่เกิดขึ้นจากการควอนไทซ์หาได้จาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$E = \sum_{k=0}^N \int_{d_k}^{d_{k+1}} (y - r_k)^2 p(y) dy \quad (2.6)$$

เมื่อ $p(y)$ คือฟังก์ชันความน่าจะเป็นของข้อมูล y

และกำหนดให้อนุพันธ์ของค่าความเพี้ยนเมื่อเทียบกับระดับ r_i และช่วง d_i เท่ากับศูนย์ทำได้

$$d_i = \frac{r_i + r_{i-1}}{2} \quad (2.7)$$

และ

$$r_i = \frac{\int_{d_i}^{d_{i+1}} yp(y) dy}{\int_{d_i}^{d_{i+1}} p(y) dy} \quad (2.8)$$

2.2.2.2 การเข้ารหัสทำนายแบบที่มีการสูญเสีย (Lossy Predictive Coding)

ในการเข้ารหัสแบบทำนายโดยทั่วไป ความสัมพันธ์ระหว่างจุดภาพที่อยู่ใกล้กันจะถูกนำมาใช้ในการทำนายค่าของแต่ละจุดภาพ วิธีหนึ่งในการลดข้อมูลที่นำเอาการเข้ารหัสแบบทำนายมาใช้คือการใช้การมอดูเลตแบบรหัสพัลส์ผลต่าง (Difference Pulse Code Modulation, DPCM) โดยจะนำเอาค่าความผิดพลาดจากการทำนายที่ได้มาจากการลบค่าการทำนายออกจากค่าจริงของจุดภาพมาทำการควอนไทซ์แบบสเกลาร์ซึ่งจะทำให้เกิดความเพี้ยนของข้อมูลขึ้น แต่ก็จะสามารถลดข้อมูลได้ในอัตราที่มากขึ้น

2.2.2.3 การเข้ารหัสโดยการแปลง (Transform Coding)

การบีบอัดข้อมูลด้วยการเข้ารหัสการแปลงจะทำการแปลงข้อมูลอินพุตที่อยู่ในรูปของ Spatial Domain ให้อยู่ในรูปของสัมประสิทธิ์ของพลังงานความถี่ ที่ความถี่ต่ำจะมีพลังงานสูง และที่ความถี่สูงพลังงานจะลดลงไป สำหรับข้อมูลภาพโดยทั่ว ๆ ไปนั้นจะประกอบด้วยส่วนที่เป็นความถี่สูงคือส่วนที่เป็นรายละเอียดหรือขอบวัตถุภายในภาพ และส่วนที่เป็นความถี่ต่ำของภาพคือข้อมูลส่วนที่เป็นพื้นของภาพ ซึ่งจำนวนบิตที่ใช้ในการเข้ารหัสของแต่ละช่วงความถี่จะไม่เท่ากัน ดังนั้นเมื่อต้องการลดข้อมูลให้ได้มาก ค่าของพลังงานความถี่สูงจะถูกตัดทิ้งไปเป็นส่วนใหญ่ ทำให้รายละเอียดส่วนที่เป็นขอบในภาพขาดหายไป ภาพที่ได้จะเบลอ (Blur) ขาดความคมชัด

การแปลงที่ใช้ในการบีบอัดข้อมูลภาพมีอยู่หลายวิธี การแปลงฟูริเยอร์แบบเร็ว (Fast Fourier Transform) การแปลงดีสครีตไซน์ (Discrete Sine Transform) การแปลงฮาร์ (Haar Transform) และการแปลงคาร์ฮูเนนเลอเฟ (Karhunen Loeve Transform) เป็นต้น ซึ่งแต่ละวิธีจะมีข้อดี และข้อเสียแตกต่างกันไป แต่วิธีที่นิยมใช้ในการบีบอัดข้อมูลกันมากที่สุดคือ การแปลงดีสครีตโคไซน์ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Discrete Cosine Transform, DCT) [12], [13], [17] เนื่องจากเป็นวิธีที่สามารถคำนวณได้ง่ายและเร็ว ตัวอย่างของการบีบอัดข้อมูลที่ใช้การแปลงโคไซน์คือ การบีบอัดข้อมูลแบบ JPEG ซึ่งปัจจุบันวิธีการนี้เป็นวิธีการมาตรฐานที่ใช้ในการลดข้อมูล

2.2.2.4 การเข้ารหัสแบนย่อย (Sub-band Coding)

ในการเข้ารหัสแบนย่อย [18] ภาพต้นแบบจะถูกกรองตามช่วงของความถี่เพื่อสร้างเป็นภาพย่อยซึ่งภาพเหล่านี้ถูกเรียกว่าภาพแบนย่อย (Sub bands) เนื่องจากแต่ละภาพแบนย่อยนี้ได้ถูกลดความกว้างแถบ (Bandwidth) เมื่อเปรียบเทียบกับภาพต้นแบบ ทำให้สามารถที่จะทำการลดจำนวนการสุ่ม (Down Sampling) ลงได้ ซึ่งสามารถทำได้ด้วยกระบวนการในการกรอง (filtering) ส่วนการลดจำนวนการสุ่มนี้เรียกว่าเป็นขั้นตอนการวิเคราะห์ (Analysis Stage) หลังจากนั้นภาพแบนย่อยที่ได้จะถูกเข้ารหัสด้วยตัวเข้ารหัสแบบต่าง ๆ โดยที่อัตราบิต (Bit Rate) ต่าง ๆ หรือแม้แต่วิธีการเข้ารหัสต่าง ๆ กันก็สามารถนำมาใช้กับภาพแต่ละแบนย่อยได้

ในการสร้างภาพกลับคืนมาอีกครั้งสามารถทำได้โดยการเพิ่มจำนวนการสุ่ม (Up Sampling) ของภาพแบนย่อยที่ได้ถอดรหัสแล้ว จากนั้นนำมาผ่านตัวกรอง แล้วรวมภาพที่สร้างกลับคืนมาแต่ละแบนด์เข้าด้วยกัน กระบวนการนี้เรียกว่าขั้นตอนการสังเคราะห์ (Synthesis Stage)

จะสังเกตได้ว่าการสร้างภาพแบนย่อยจะไม่ทำให้เกิดการลดลงของข้อมูล เนื่องจากจำนวนของการสุ่ม (Sample) ของทุกแบนด์ย่อยรวมกันแล้วยังคงเท่ากับภาพต้นแบบ แต่จะมีข้อดีในแง่ที่แต่ละภาพแบนด์ย่อยนั้นสามารถที่จะทำการเข้ารหัสได้ด้วยประสิทธิภาพที่ดีกว่าภาพต้นแบบเดิม

2.2.2.5 การบีบอัดข้อมูลภาพด้วยการควอนไทซ์แบบเวกเตอร์ (Vector Quantization, VQ)

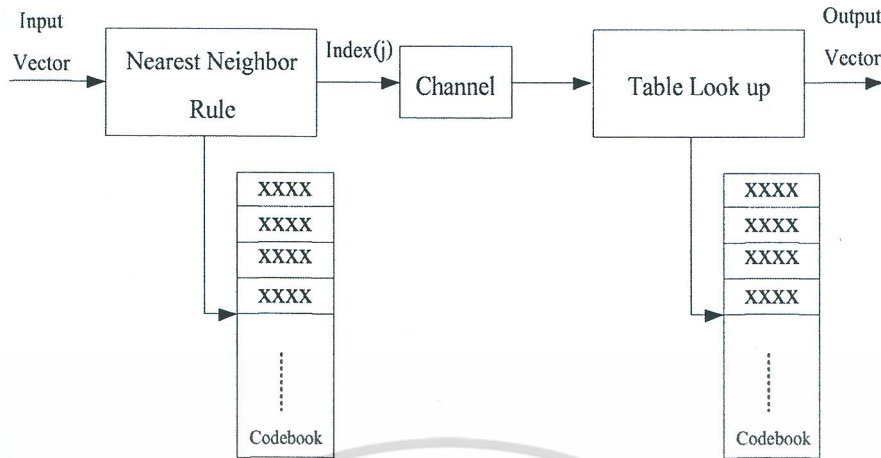
การบีบอัดข้อมูลรูปภาพด้วยวิธีการควอนไทซ์แบบเวกเตอร์ (Vector Quantization, VQ) เป็นวิธีการบีบอัดข้อมูลแบบที่มีการสูญเสียวิธีหนึ่ง ซึ่งนิยามของการควอนไทซ์แบบเวกเตอร์ [19], [20] คือการสร้างความสัมพันธ์ (Q) ระหว่างข้อมูลของ ปริภูมิยูคลิเดียน (Euclidian space) K มิติ (R^K) ลงในเซตย่อย Y ของ R^K

$$Q: R^K \rightarrow Y \quad (2.9)$$

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

และ N คือจำนวนของเวกเตอร์ใน Y

การควอนไทซ์แบบเวกเตอร์ประกอบด้วย 2 ส่วน ดังแสดงในรูปที่ 2.7



รูปที่ 2.7 บล็อกไดอะแกรมการควอนไทซ์แบบเวกเตอร์

ส่วนเข้ารหัสจะทำหน้าที่เปรียบเทียบกับอินพุตเวกเตอร์ที่ต้องการจะเข้ารหัสกับเวกเตอร์รหัส (Code Vector หรือ Code word) ทั้งหมดที่เก็บอยู่ในตัวเก็บรหัส (Codebook) โดยกลุ่มของเวกเตอร์รหัสเหล่านี้จะแทนเวกเตอร์ที่เป็นไปได้ทั้งหมดของอินพุตเวกเตอร์ โดยการเปรียบเทียบนี้จะได้ผลลัพธ์คือดัชนี (Index) ที่แทนเวกเตอร์รหัสที่มีความเพี้ยนน้อยที่สุด (Minimum Distortion) กับอินพุตเวกเตอร์

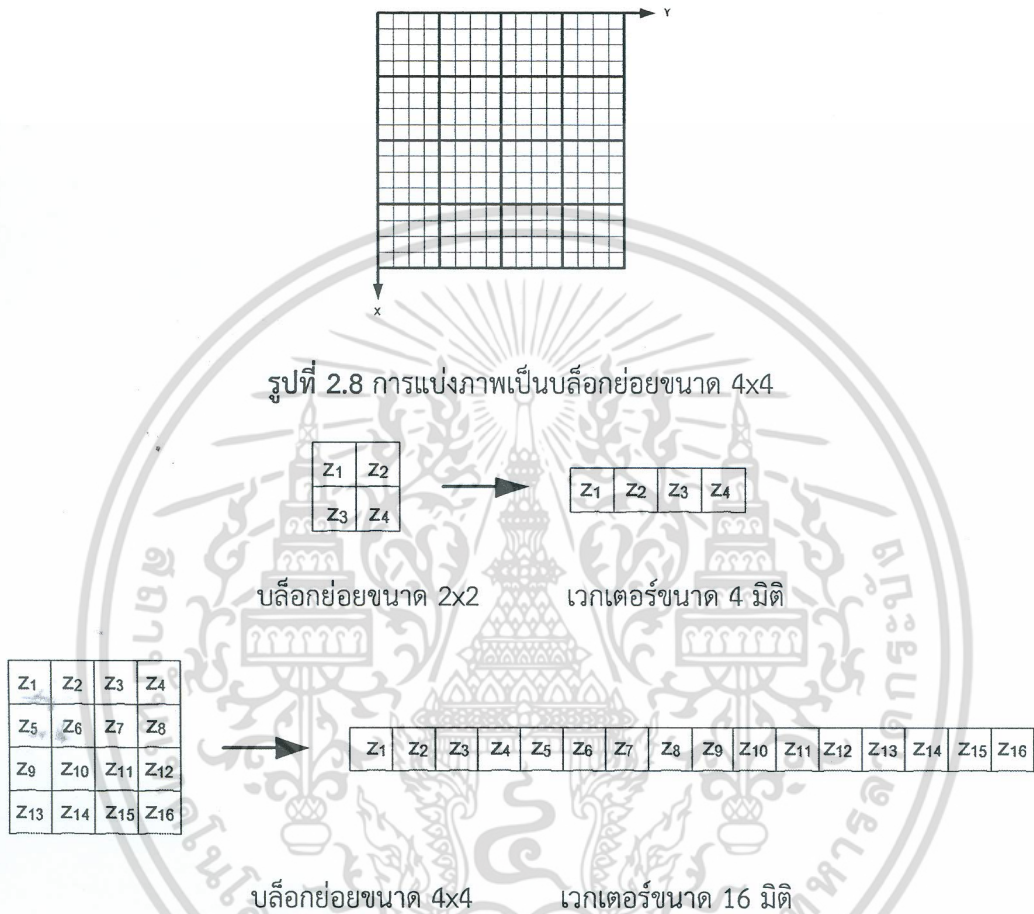
ส่วนถอดรหัสจะทำหน้าที่นำเอาค่าดัชนีมาชี้เวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัส ซึ่งจะต้องเป็นตัวเก็บรหัสเดียวกันกับตอนเข้ารหัส โดยเวกเตอร์รหัสที่ได้จะเป็นเอาท์พุตเวกเตอร์ และในการควอนไทซ์แบบเวกเตอร์สำหรับภาพเชิงตัวเลขเป็นวิธีการบีบอัดข้อมูลภาพวิธีการหนึ่งโดยอาศัยหลักการแทนที่บล็อกย่อยของข้อมูลภาพต้นแบบทั้งชุดข้อมูลที่ใช้แทนบล็อกย่อยนั้น ซึ่งมักจะใช้ดัชนีเป็นเลขจำนวนนับ และมีค่าจำกัดอยู่ในระดับหนึ่ง

2.2.2.5.1 การเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

ในการเข้ารหัสภาพเชิงตัวเลขด้วยการควอนไทซ์แบบเวกเตอร์ ภาพต้นแบบจะถูกแยกองค์ประกอบให้เป็นเวกเตอร์ของภาพขนาด K มิติ ซึ่งเวกเตอร์เหล่านี้สามารถสร้างด้วยวิธีการต่าง ๆ หลายวิธี เช่น แบ่งภาพให้เป็นบล็อกของจุดภาพขนาด $l * m = K$ ซึ่งจะทำการจัดเรียงให้อยู่ในรูปเวกเตอร์ขนาด K มิติได้ หรือองค์ประกอบของสี RGB ของภาพก็สามารถจัดให้อยู่ในรูปเวกเตอร์ขนาด 3 มิติได้ เป็นต้น

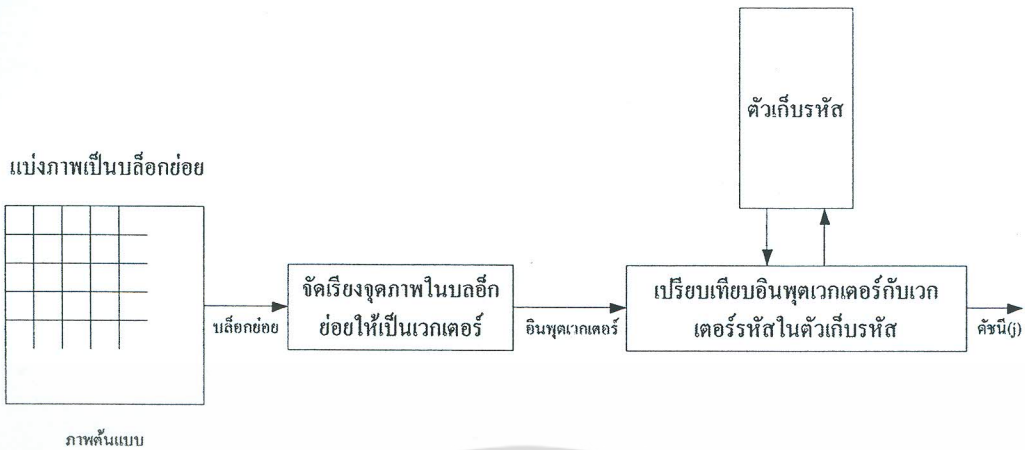
โดยทั่วไปอินพุตเวกเตอร์ที่จะนำมาทำการควอนไทซ์ แบบเวกเตอร์มักจะได้มาจากการแบ่งภาพออกเป็นบล็อกย่อย ๆ ของจุดภาพที่ไม่ซ้อนทับกัน ซึ่งมักจะแบ่งเป็นบล็อกสี่เหลี่ยมจัตุรัสขนาดเท่า ๆ กัน เนื่องจากมีความง่ายในการจัดแบ่งให้ครอบคลุมได้ทั้งภาพ ดังตัวอย่างในรูปที่ 2.8 จากนั้นทำการจัดเรียงข้อมูลจุดภาพในแต่ละบล็อกย่อย ให้อยู่ในรูปของเวกเตอร์ดังตัวอย่างในรูปที่ 2.9

ดังนั้นจะพบว่าถ้าภาพต้นแบบมีขนาดเป็น 256×256 จุดภาพ จะสามารถทำการแบ่งเป็นบล็อกย่อยขนาด 2×2 ที่ไม่ทับซ้อนกันได้จำนวนทั้งหมด 16,384 บล็อก และถ้าแบ่งเป็นบล็อกย่อยขนาด 4×4 ที่ไม่ทับซ้อนกันได้จำนวนทั้งหมด 4,096 บล็อก เป็นต้น ซึ่งจำนวนบล็อกย่อยที่แบ่งได้พอดีนี้จะเท่ากับจำนวนอินพุตเวกเตอร์ที่จะนำมาทำการควอนไทซ์แบบเวกเตอร์ต่อไป



รูปที่ 2.9 ตัวอย่างการจัดเรียงจุดภาพของบล็อกย่อยขนาดต่าง ๆ ให้อยู่ในรูปของเวกเตอร์

จากนั้นส่วนเข้ารหัสจะทำการเปรียบเทียบ อินพุตเวกเตอร์ที่ได้มาจากภาพต้นแบบกับเวกเตอร์รหัสทุกตัวที่เก็บอยู่ในตัวเก็บรหัส โดยตัวเก็บรหัสจะประกอบด้วยกลุ่มของเวกเตอร์รหัสที่แทนเวกเตอร์ที่เป็นไปได้ทั้งหมดของอินพุตเวกเตอร์ ในการเปรียบเทียบหาเวกเตอร์รหัสที่มีความเพี้ยนจากอินพุตเวกเตอร์น้อยที่สุด ผลลัพธ์ที่ได้จากการเปรียบเทียบนี้คือเลขดัชนีที่ใช้แทนเวกเตอร์รหัสที่มีความเพี้ยนน้อยที่สุดดังรูปที่ 2.10



รูปที่ 2.10 การเข้ารหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

เวกเตอร์แต่ละเวกเตอร์ของภาพ X จะถูกนำไปเปรียบเทียบกับชุดข้อมูลที่จะใช้แทนหรือเรียกว่าเวกเตอร์รหัส (Code vector) $(\hat{X}_i, i = 1, 2, 3, \dots, N_c)$ เมื่อ N_c คือ จำนวนของเวกเตอร์รหัส ซึ่งจะได้มาจากตัวเก็บรหัสที่ได้สร้างเอาไว้ก่อนแล้ว ซึ่งมีโครงสร้างดังแสดงในรูปที่ 2.11 โดยเวกเตอร์รหัสจะมีขนาดเป็น K มิติด้วยเหมือนกัน ซึ่งการเปรียบเทียบจะทำการเลือกเวกเตอร์รหัสที่เหมาะสมที่สุดโดยใช้กฎความเพี้ยนที่น้อยที่สุด คือจะทำการเลือกเวกเตอร์รหัส \hat{X}_k ที่ทำให้ $d(X, \hat{X}_k) \leq d(X, \hat{X}_j)$ สำหรับ $j = 1, 2, 3, \dots, N_c$ เมื่อ $d(X, \hat{X}_k)$ คือความเพี้ยนที่เกิดขึ้นจากการแทนเวกเตอร์ต้นแบบ X ด้วยเวกเตอร์รหัส \hat{X}

การหาความเพี้ยนนี้ทำได้จากการหาค่าความผิดพลาดกำลังสองเฉลี่ย (MES) ของอินพุตเวกเตอร์กับเวกเตอร์รหัส ดังสมการต่อไปนี้

$$d(x, \hat{x}_i) = \frac{1}{K} \sum_{m=1}^K [x(m) - \hat{x}_i(m)]^2 \tag{2.10}$$

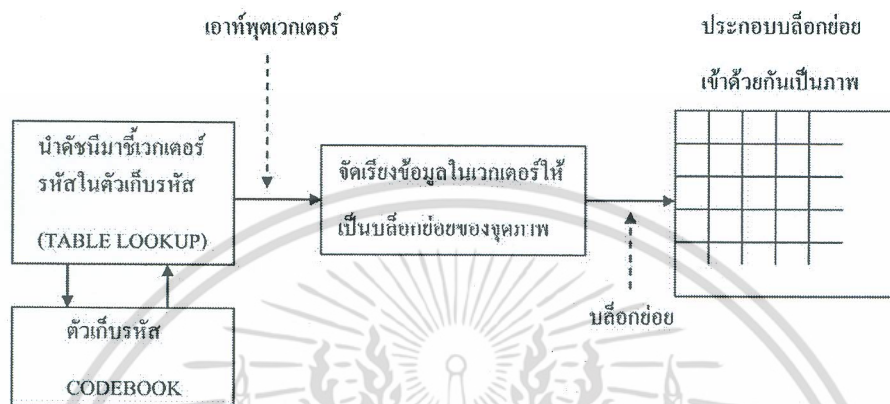
Index(i)	Codevector(\hat{x}_i)				
$i = 1$	$\hat{x}_1(1)$	$\hat{x}_1(2)$	$\hat{x}_1(3)$	\dots	$\hat{x}_1(k)$
$i = 2$	$\hat{x}_2(1)$	$\hat{x}_2(2)$	$\hat{x}_2(3)$	\dots	$\hat{x}_2(k)$
$i = 3$	$\hat{x}_3(1)$	$\hat{x}_3(2)$	$\hat{x}_3(3)$	\dots	$\hat{x}_3(k)$
				\vdots	
i	$\hat{x}_i(1)$	$\hat{x}_i(2)$	$\hat{x}_i(3)$	\dots	$\hat{x}_i(k)$
				\vdots	
$i = N_c$	$\hat{x}_{N_c}(1)$	$\hat{x}_{N_c}(2)$	$\hat{x}_{N_c}(3)$	\dots	$\hat{x}_{N_c}(k)$

รูปที่ 2.11 โครงสร้างของตัวเก็บรหัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2.5.2 การถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

การถอดรหัสด้วยการควอนไทซ์แบบเวกเตอร์ เป็นการนำเอาค่าดัชนีมีซีตำแหน่งข้อมูลเวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัสเดียวกันกับตอนเข้ารหัส โดยผลลัพธ์ที่ได้จะเป็นเอาท์พุทเวกเตอร์ จากนั้นนำเอาเอาท์พุทเวกเตอร์ที่ได้มาจัดเรียงใหม่ให้เป็นบล็อกย่อยของจุดภาพ ดังรูปที่ 2.12



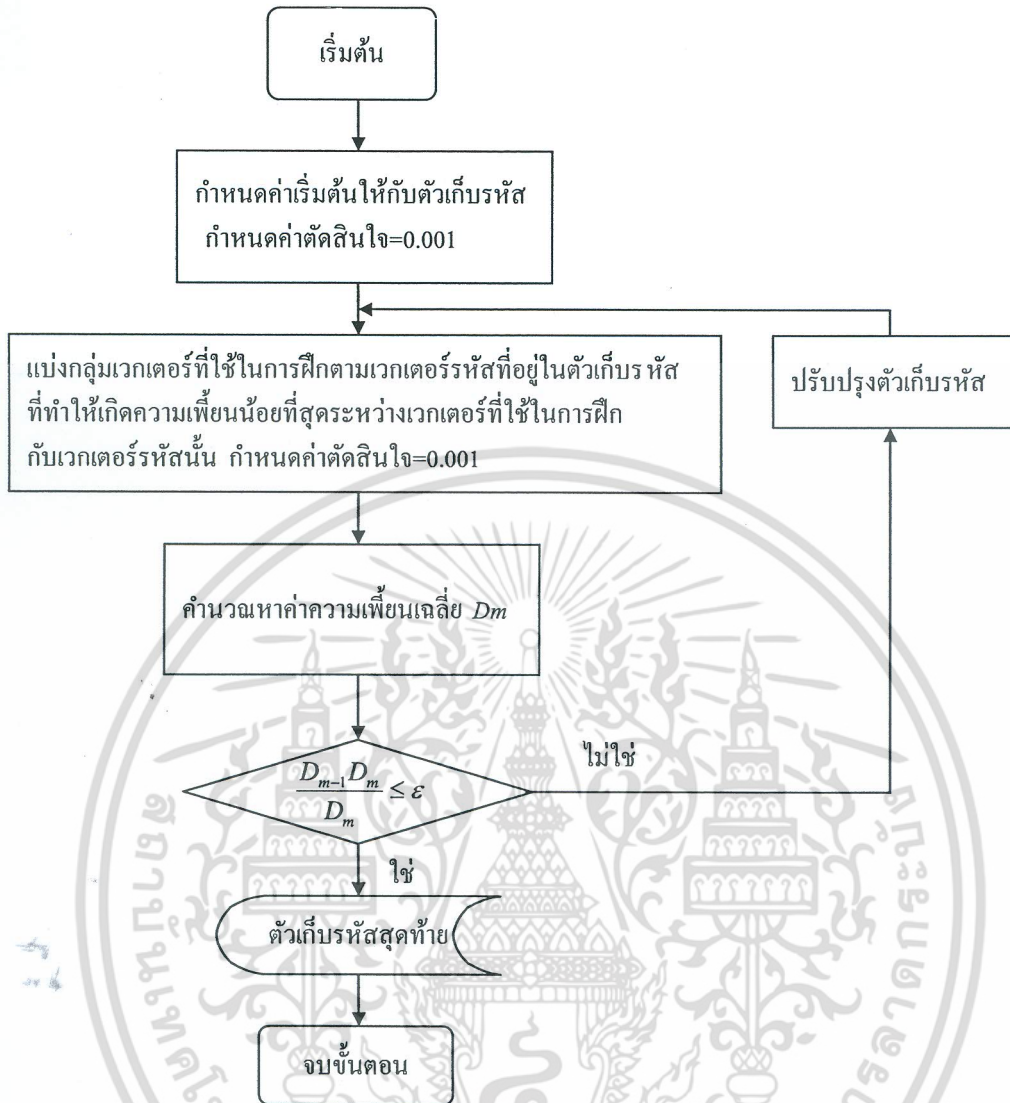
รูปที่ 2.12 การถอดรหัสภาพด้วยการควอนไทซ์แบบเวกเตอร์

2.2.2.5.3 การออกแบบตัวเก็บรหัสด้วยโมเดลทางคณิตศาสตร์

การออกแบบตัวเก็บรหัสด้วยโมเดลทางคณิตศาสตร์ จะได้การควอนไทซ์ที่เป็นแบบยูนิฟอร์ม (Uniform) โดยทำการกำหนดโครงสร้างที่เป็นลักษณะปกติทั่วไปของภาพให้กับตัวเก็บรหัสเพื่อทำให้การเข้ารหัส และการถอดรหัสสามารถกระทำได้รวดเร็ว โดยปกติจำเป็นต้องรู้โมเดลความน่าจะเป็นของเวกเตอร์ของภาพ และต้องการตัวเก็บรหัสที่มีขนาดใหญ่เพื่อให้ได้ประสิทธิภาพที่ดี

2.2.2.5.4 การสร้างตัวเก็บรหัสที่ดีที่สุดจากการเรียนรู้ชุดข้อมูลที่นำมาทำการฝึก (Training Set)

วิธีนี้มีความเหมาะสมในการใช้งานตามความเป็นจริงมากกว่า เหมาะสำหรับตัวเก็บรหัส ที่มีขนาดเล็ก โดยไม่จำเป็นต้องรู้โมเดลความน่าจะเป็นของเวกเตอร์ของภาพ ซึ่งวิธีการที่นิยมใช้กัน ได้แก่ การออกแบบตัวเก็บรหัสโดยใช้อัลกอริทึมของ Linde-Buzo-Gray หรือ LBG อัลกอริทึม LBG [19], [20] เป็นอัลกอริทึมที่นิยมใช้ในการสร้างตัวเก็บรหัส โดยอัลกอริทึม จะเริ่มต้นจากการกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส แล้วทำการปรับปรุงตัวเก็บรหัสโดยการแทนเวกเตอร์รหัสแต่ละตัวที่อยู่ในตัวเก็บรหัสด้วยเวกเตอร์ตัวใหม่ที่จะทำให้เกิดความเพี้ยนลดลงไปเรื่อย ๆ โดยจะทำการปรับปรุงวนซ้ำไปจนกว่าอัตราการเปลี่ยนแปลงของความเพี้ยนที่ลดลงมีค่าต่ำกว่าขอบเขตที่กำหนดไว้ อัลกอริทึมของ LBG สามารถเขียนเป็นโปรแกรมได้ดังรูปที่ 2.13



รูปที่ 2.13 โพลีชาร์ตของอัลกอริทึม LBG

ลำดับขั้นตอนการทำงานของอัลกอริทึม LBG สามารถอธิบายโดยละเอียดได้ดังต่อไปนี้

ขั้นตอนที่ 1 นำภาพที่จะมาทำการฝึก (Training) แบ่งเป็นบล็อกย่อยขนาดเท่า ๆ กันจำนวน แล้วจัดให้อยู่ในรูปเวกเตอร์สำหรับการฝึก (Training Vectors) x_j

โดยที่ $x_j = [x_j(1) x_j(2) x_j(3) \dots x_j(K)]$ เมื่อ $j = 1, 2, 3, \dots, n$

n คือ จำนวนบล็อกย่อยหรือจำนวนเวกเตอร์ที่ได้มาจากภาพที่จะนำมาทำการฝึก

K คือ ขนาดมิติของเวกเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 ทำการกำหนดค่าเริ่มต้นให้กับอัลกอริทึม ดังนี้

2.1 กำหนด m เป็นจำนวนรอบของการคำนวณ = 0

2.2 กำหนดค่าเริ่มต้นให้กับ ตัวเก็บรหัส $Y_m = \{\hat{x}_i; i = 1, 2, 3, \dots, N_c\}$

2.3 กำหนด D_m เป็นตัวแปรวัดความเพี้ยนเฉลี่ย (Average Distortion) ของ

เวกเตอร์ทั้งหมดให้มีค่ามาก ๆ

2.4 กำหนด ε เป็นค่าตัดสินใจของอัตราการเปลี่ยนแปลงค่าความเพี้ยนเฉลี่ย

=0.001

ขั้นตอนที่ 3 ทำการจัดกลุ่มของเวกเตอร์ที่นำมาฝึก x_j ตามจำนวนของเวกเตอร์รหัส

$$P(Y_m) = (S_i; i = 1, 2, \dots, N_c) \quad (2.11)$$

เมื่อ Y_m คือ ตัวเก็บรหัสของรอบการคำนวณที่ m

N_c คือ จำนวนของเวกเตอร์รหัสที่เก็บอยู่ในตัวเก็บรหัส

S_i คือ เซ็ตของเวกเตอร์ที่มีความเพี้ยนน้อยที่สุดกับ เวกเตอร์รหัส x

โดยการจัดกลุ่มจะใช้หลักความเพี้ยนน้อยที่สุดดังนี้ คือ

$$x_j \in S_k \text{ ถ้า } d(x_j, x_i) \leq d(x_j, x_i) \text{ สำหรับทุกค่าของ } i$$

เมื่อ

$$d(x_j, \hat{x}_i) = \frac{1}{K} \sum_{m=1}^K [x_j(m) - \hat{x}_i(m)]^2 \quad (2.12)$$

ขั้นตอนที่ 4 ทำการคำนวณหาค่าความเพี้ยนเฉลี่ย (Average Distortion) D_{m+1} โดยที่

$$D_{m+1} = D[(Y_m, P(Y_m))] = \frac{1}{n} \sum_{j=1}^n \min d(x_j, \hat{x}) \quad (2.13)$$

ขั้นตอนที่ 5 ทำการคำนวณหาอัตราการเปลี่ยนแปลงของค่าความเพี้ยนเฉลี่ยของรอบการคำนวณที่แล้ว กับรอบการคำนวณปัจจุบันซึ่งสามารถคำนวณได้จาก $\frac{D_m - D_{m+1}}{D_{m+1}}$ แล้วนำมาเปรียบ เทียบกับค่าการตัดสินใจ (ε)

5.1 ถ้า $\frac{D_m - D_{m+1}}{D_{m+1}}$ มีค่ามากกว่าค่าการตัดสินใจ ϵ ให้ทำการปรับปรุงค่าของ

เวกเตอร์รหัสในตัวเก็บรหัส โดยการนำแต่ละกลุ่มของเวกเตอร์ที่จัดไว้ในขั้นตอนที่(3) มาทำการหาค่าเฉลี่ยในแต่ละมิติของเวกเตอร์ ดังนั้นจะได้ตัวเก็บรหัสที่ปรับปรุงแล้วเป็น

$$\hat{X}(P(Y_m)) = (\hat{X}(S_i); i = 1, 2, \dots, N_c) \quad (2.14)$$

เมื่อ

$$\hat{X}(S_i) = \frac{1}{\|S_i\|} \sum_{j=x_j \in S_i} x_j \quad (2.15)$$

ดังนั้นจะได้ตัวเก็บรหัสใหม่เป็น $Y_{m+1} = \hat{X}(S_i)$ (2.16)

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

5.2 ถ้า $\frac{D_m - D_{m+1}}{D_{m+1}}$ น้อยกว่าหรือเท่ากับ ϵ จะเป็นการจบอัลกอริทึม และได้ตัวเก็บ

รหัสคือ Y_m

2.2.2.5.5 การกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัส

การกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัสจะมีผลต่อการทำงานของอัลกอริทึม LBG โดยในการกำหนดค่าเริ่มต้นให้กับตัวเก็บรหัสสามารถกำหนดให้ตัวเก็บรหัส มีขนาดได้ตามที่ต้องการ ซึ่งมักจะกำหนดตามจำนวนบิตของเลขดัชนีที่จะใช้ เช่น ถ้าต้องการให้เลขดัชนีที่มีขนาด 1 บิต จะได้ตัวเก็บรหัสที่มีขนาดเป็น 2^1 เวกเตอร์ หรืออีกวิธีคือเราสามารถกำหนดให้ตัวเก็บรหัสขนาดเล็ก ๆ ก่อนที่จะทำการคำนวณแบบเรียกตัวเอง (Recursive) เพื่อสร้างตัวเก็บรหัสที่ใหญ่ขึ้น จนได้ตามขนาดที่ต้องการได้

2.2.2.5.5.1 วิธีการแบบสุ่ม (Random)

วิธีการนี้เป็นวิธีการที่ง่าย โดยจะเริ่มต้นจากการนำภาพที่ต้องการจะใช้ในการฝึกมาแบ่งเป็นบล็อกย่อย แล้วจัดเรียงบล็อกย่อยให้เป็นเวกเตอร์ จากนั้นนำ เวกเตอร์แรกของเวกเตอร์ที่ได้มา กำหนดเป็นค่าเริ่มต้นให้กับตัวเก็บรหัส แล้วนำไปผ่านอัลกอริทึม LBG ต่อไป ขั้นตอนทั้งหมดสามารถอธิบายได้ตามโพลีชาร์ตที่แสดงไว้ในรูปที่ 2.14

ข้อเสียของวิธีการนี้คืออาจจะไม่ได้เวกเตอร์ที่มีการกระจายที่ดีพอ

2.2.2.5.5.2 วิธีการแบบยูนิฟอร์มควอนไทเซอร์ (Uniform

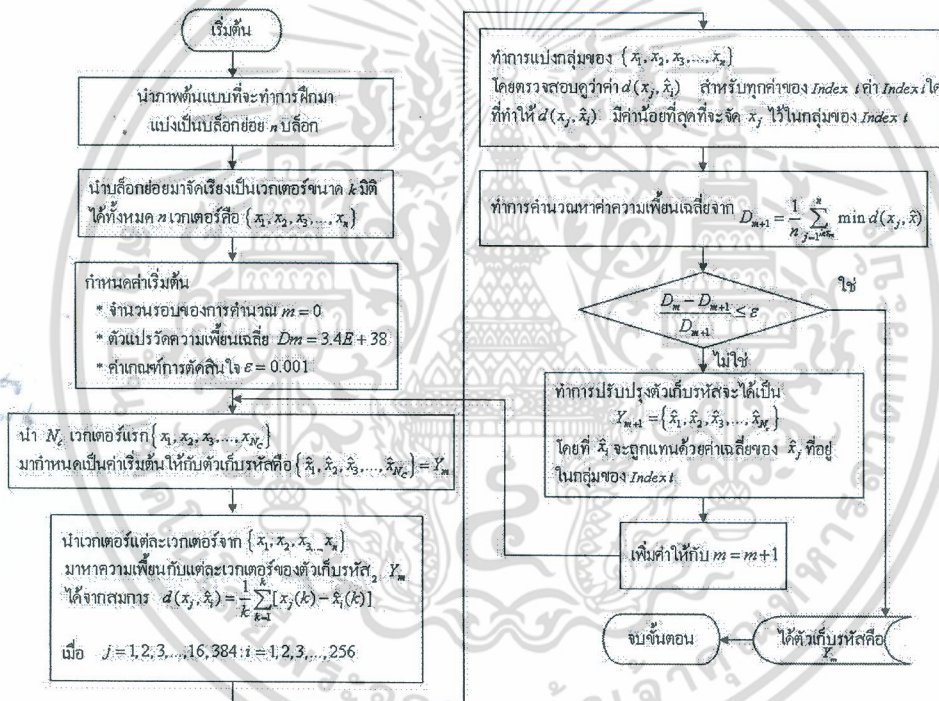
Quantize)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

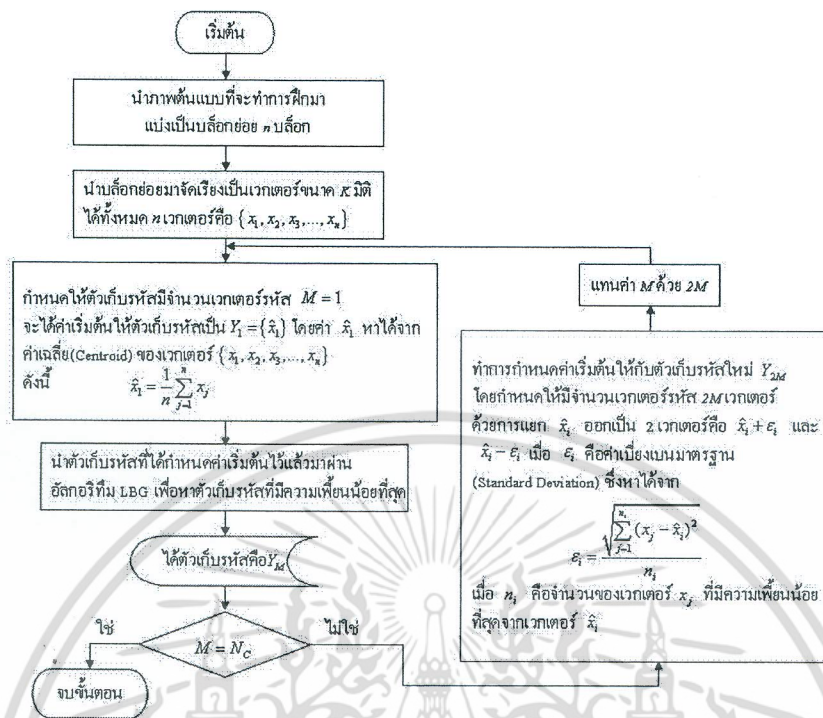
การกำหนดค่าเริ่มต้นด้วยยูนิฟอร์มควอนไทเซอร์ (Uniform Quantizer) เป็นการกำหนดค่าเริ่มต้นให้ครอบคลุมข้อมูลทั้งหมดที่จะนำมาทำการฝึก ซึ่งวิธีนี้จะใช้กับข้อมูลที่มีการกระจายแบบเกาส์ (Gaussian)

2.2.2.5.3 วิธีการแตกตัว (Splitting algorithm)

วิธีการนี้จะมีประโยชน์ในการออกแบบตัวควอนไทซ์ที่มีอัตรา (Rate) สูงขึ้นอย่างต่อเนื่องจนกระทั่งได้ระดับความผิดพลาดที่ต้องการ เราสามารถกำหนดตัวควอนไทซ์ M ระดับ ด้วย $M = 2^R$ เมื่อ $R = 0, 1, \dots$, จนกระทั่งได้ค่าเริ่มต้นสำหรับตัวควอนไทซ์ N_c ระดับ ในรูปที่ 2.15 จะแสดงโฟลว์ชาร์ตของการกำหนดค่าเริ่มต้นโดยการแตกตัว



รูปที่ 2.14 โฟลว์ชาร์ตของการสร้างตัวเก็บรหัส โดยการกำหนดค่าเริ่มต้นแบบ Random



รูปที่ 2.15 โฟลต์ชาร์ตของการสร้างตัวเก็บรหัส โดยการกำหนดค่าเริ่มต้นแบบวิธีการแตกตัว (Splitting)

การสร้างตัวเก็บรหัสขนาด N_c เวกเตอร์โดยการกำหนดค่าเริ่มต้นด้วยวิธีการแตกตัวมีขั้นตอน โดยละเอียดดังต่อไปนี้

ขั้นตอนที่ 1 นำภาพที่จะมาทำการฝึก (Training) แบ่งเป็นบล็อกย่อยขนาดเท่า ๆ กัน แล้วจัดให้อยู่ในรูปเวกเตอร์สำหรับการฝึก (Training vectors) x_j

โดยที่ $x_j = [x_j(1), x_j(2), x_j(3), \dots, x_j(K)]$ เมื่อ $j = 1, 2, 3, \dots, n$

n คือ จำนวนบล็อกย่อยหรือจำนวนเวกเตอร์ที่ได้มาจากภาพที่จะนำมาทำการฝึก

K คือ ขนาดมิติของเวกเตอร์

ขั้นตอนที่ 2 กำหนดค่าเริ่มต้นให้กับตัวเก็บรหัสโดยการกำหนดจำนวนของเวกเตอร์

รหัส $M = 1$ และกำหนดตัวเก็บรหัส $Y_1 = \{\hat{x}_1\}$ เมื่อ \hat{x}_1 คือ ค่าจุดศูนย์กลางหรือ Centroid ของเวกเตอร์สำหรับการฝึก ซึ่งสามารถหาได้ ดังนี้

$$\hat{x}_1 = \frac{1}{n} \sum_{j=1}^n x_j \quad \text{เมื่อ } j=1,2,3,\dots,n \quad (2.17)$$

ขั้นตอนที่ 3 นำตัวเก็บรหัสที่ได้ทำการกำหนดค่าเริ่มต้นแล้วไปผ่านอัลกอริทึม LBG เพื่อให้ได้ตัวเก็บรหัสที่มีความเพี้ยนน้อยที่สุด

ขั้นตอนที่ 4 ทำการตรวจสอบว่า M เท่ากับ N_C หรือไม่

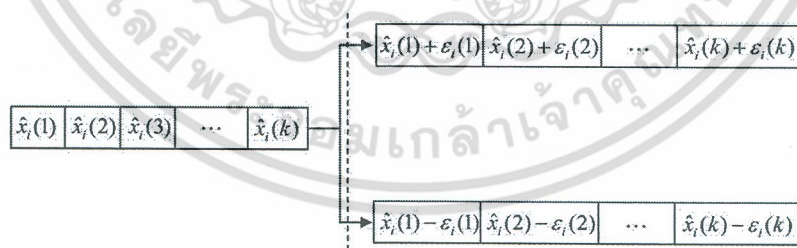
ถ้า $M = N_C$ จะเป็นการจบอัลกอริทึม และได้ตัวเก็บรหัสคือ Y_M

ถ้า $M < N_C$ ให้ไปที่ขั้นตอนที่ 5

ขั้นตอนที่ 5 นำตัวเก็บรหัส Y_M ซึ่งประกอบด้วยเวกเตอร์รหัสจำนวน M เวกเตอร์ คือ $Y_M = \{\hat{x}_i; i=1,2,3,\dots,M\}$ ทำการแยกเวกเตอร์ \hat{x}_i ทุกตัวให้เป็น 2 เวกเตอร์ที่ใกล้เคียงกันคือ $\hat{x}_i + \varepsilon_i$ และ $\hat{x}_i - \varepsilon_i$ เมื่อ ε_i คือเวกเตอร์ที่ใช้ในการแยกซึ่งสามารถหาค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ดังนี้

$$\varepsilon_i = \frac{\sqrt{\sum_{j=1}^{n_i} (x_j - \hat{x}_i)^2}}{n_i} \quad (2.18)$$

เมื่อ n_i คือจำนวนของเวกเตอร์ x_j ที่มีความเพี้ยนน้อยที่สุดกับเวกเตอร์รหัส \hat{x}_i ทำให้ได้ตัวเก็บรหัส $Y_{2M} = \{\hat{x}_i + \varepsilon_i, \hat{x}_i - \varepsilon_i; i=1,2,3,\dots,M\}$ ประกอบด้วยจำนวนเวกเตอร์รหัส $2M$ เวกเตอร์ ดังแสดงในรูปที่ 2.16 จากนั้นแทนค่า M ด้วย $2M$ จากนั้นกลับไปทำที่ขั้นตอนที่ 3



รูปที่ 2.16 การแยกตัวเก็บรหัสให้มีขนาดเป็น 2 เท่า

การสร้างตัวเก็บรหัสโดยการกำหนดค่าเริ่มต้นด้วยการแตกตัว ภายหลังจากการแบ่งกลุ่มเวกเตอร์สำหรับการฝึกที่มีความเพี้ยนน้อยที่สุดกับเวกเตอร์รหัสทุกเวกเตอร์ อาจจะมีเวกเตอร์รหัสที่ไม่มีเวกเตอร์สำหรับการฝึกเป็นสมาชิกอยู่เลย ซึ่งทำให้เวกเตอร์รหัสนั้นไม่เกิดประโยชน์ในเวลาเข้ารหัส ดังนั้นจึงต้องกำจัดเวกเตอร์รหัสนั้นออก แล้วสร้างเวกเตอร์ใหม่โดยการนำเวกเตอร์รหัสที่มี

สมาชิกมากที่สุด (\hat{x}_{\max}) มาทำการแตกตัวออกเป็นเวกเตอร์รหัส 2 เวกเตอร์ที่ใกล้เคียงกันคือ $\hat{x}_{\max} + \varepsilon_{\max}$ และ $\hat{x}_{\max} - \varepsilon_{\max}$ แทนที่เวกเตอร์ \hat{x}_{\max} เดิมและเวกเตอร์รหัสที่ถูกกำจัดออกไป

2.3 การแบ่งประเภทของข้อมูล (Data Classification)

2.3.1 บทนำ (Introduction)

การแบ่งกลุ่ม (Clustering) คือ การแบ่งประเภท (Classification) ของข้อมูลออกเป็นกลุ่มที่แตกต่างกัน โดยจะแบ่งชุดข้อมูล (มักจะเป็นเวกเตอร์) ออกเป็นกลุ่ม (Cluster) นำข้อมูลที่มีคุณลักษณะเหมือนกัน หรือคล้ายกันจัดไว้ในกลุ่มเดียวกัน ขั้นตอนวิธีที่ใช้ในการแบ่งกลุ่มจะอาศัยความเหมือน (Similarity) หรือ ความใกล้ชิด (Proximity) โดยคำนวณจากการวัดระยะระหว่างเวกเตอร์ของข้อมูล การแบ่งกลุ่มข้อมูลเป็นเทคนิคทั่วไปที่ใช้สำหรับการวิเคราะห์ข้อมูลทางสถิติซึ่งถูกนำไปใช้ในหลายๆสาขางานรวมทั้ง Machine learning, Data mining, Pattern recognition, Image analysis และ Bioinformatics

การแบ่งกลุ่มข้อมูลจะแตกต่างจากการแบ่งประเภทข้อมูล (Classification) โดยจะแบ่งกลุ่มข้อมูลจากความคล้าย โดยไม่มีการกำหนดประเภทของข้อมูลไว้ก่อน

นอกจาก Data clustering แล้วยังมีอีกหลายชื่อที่มีความหมายเหมือนกัน คือ Cluster analysis, Automatic classification, Numerical taxonomy, Botryology and Typological analysis

ขั้นตอนที่สำคัญอย่างหนึ่งในการแบ่งกลุ่มข้อมูลก็คือ การเลือกวิธีการวัดระยะห่าง ซึ่งเป็นตัวกำหนดว่า 2 องค์ประกอบที่ค่านอนั้นเหมือนกันอย่างไร การแบ่งกลุ่มนั้นใช้การวัดระยะห่างได้ 2 แบบ คือ สมมาตร (Symmetric) และไม่สมมาตร (Asymmetric) Euclidean space มีคุณสมบัติเป็นระยะห่างแบบสมมาตร (ระยะห่างจาก A ไป B เท่ากับจาก B ไป A) ส่วนในแบบอื่น ๆ นั้นไม่อยู่ในกรณีนี้ วิธีการวัดระยะห่างนี้จะมีผลต่อรูปแบบลักษณะของกลุ่มข้อมูล เพราะบางองค์ประกอบอาจจะอยู่ใกล้กับองค์ประกอบอื่น ทั้งนี้ขึ้นอยู่กับระยะห่างระหว่างองค์ประกอบนั้นๆ กับองค์ประกอบอื่น ตัวอย่างเช่น ในระบบพิกัด 2 มิติ ระยะห่างระหว่างจุด $(x=1, y=0)$ และจุดกำเนิด $(x=0, y=0)$ เท่ากับ 1 เสมอตาม Norms ธรรมดา (Usual norm) แต่ระยะห่างระหว่างจุด $(x=1, y=1)$ และจุดกำเนิดเท่ากับ 2, $\sqrt{2}$ หรือ 1 ตามลำดับ

การวัดระยะทาง

- Euclidean distance (หรือเรียกว่า ระยะทางแบบ Crow flies หรือ ระยะทางแบบ 2-norm) จากการพิจารณาการวิเคราะห์กลุ่มข้อมูลในงานวิจัยทางด้านสุขภาพจิตที่ได้มีการตีพิมพ์พบว่าการวัดระยะทางซึ่งรู้จักกันดีที่สุดในการศึกษาทางด้านนี้ ก็คือ Euclidean distance หรือ squared Euclidean distance

- Manhattan distance (หรือเรียกว่า Taxicab norm หรือ 1-norm)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Infinity-norm
- Maha lanobis distance การแบ่งกลุ่มข้อมูลให้ถูกต้องตาม scales ที่ต่างกัน และความสัมพันธ์ร่วมกันของตัวแปร

อัลกอริทึมในการแบ่งกลุ่มข้อมูล โดยทั่วไปแบ่งได้เป็น 2 ประเภทใหญ่ๆ คือ การแบ่งแบบเป็นลำดับชั้น (Hierarchical) และ การแบ่งแบบตัดเป็นส่วน (Partition) การแบ่งแบบเป็นลำดับชั้นนั้น จะทำการแบ่งกลุ่มจากกลุ่มย่อยที่ถูกแบ่งไว้ก่อนหน้านั้นซ้ำหลายครั้ง ส่วนการแบ่งแบบตัดเป็นส่วนนั้น การแบ่งข้อมูลทั้งหมดจะทำเพียงครั้งเดียว

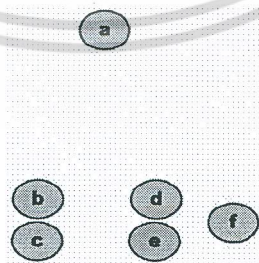
2.3.2 การแบ่งกลุ่มข้อมูลแบบเป็นลำดับชั้น (Hierarchical clustering)

การแบ่งแบบเป็นลำดับชั้น จะมี 2 ลักษณะคือแบบรวมกลุ่มหรือล่างขึ้นบน (Agglomerative หรือ Bottom-up) คือ เป็นการแบ่งแบบรวมกลุ่มจากกลุ่มย่อยให้ใหญ่ขึ้นไปเรื่อยๆ โดยเริ่มจากกลุ่มเล็กสุดคือในแต่ละกลุ่มมีข้อมูลเพียงตัวเดียว และ แบบแยกกลุ่มหรือบนลงล่าง (Divisive หรือ Top-down) คือ เป็นการแบ่งแยกกลุ่มจากกลุ่มใหญ่ให้ย่อยไปเรื่อยๆ โดยเริ่มจากกลุ่มใหญ่ที่สุด คือกลุ่มเดียวมีข้อมูลทุกตัวอยู่ในกลุ่ม

การแบ่งแบบเป็นลำดับชั้นนั้นสร้าง (รวมกลุ่ม) หรือแบ่ง (แบ่งแยกกลุ่ม) ลำดับชั้นของกลุ่มข้อมูล ตัวอย่างโดยทั่วไปของลำดับชั้นนี้ก็คือ ลำดับชั้นแบบ Tree diagram (Dendrogram) ซึ่งแต่ละองค์ประกอบของลำดับชั้นสุดท้าย และกลุ่มเดี่ยวนี้ประกอบไปด้วยทุกๆองค์ประกอบของลำดับชั้นอื่นๆ ขั้นตอนการรวมกลุ่มเริ่มที่ส่วนบนสุดของ Tree diagram ในทางตรงกันข้ามขั้นตอนการแบ่งแยกกลุ่มเริ่มที่ส่วนล่างสุดของ Tree diagram

2.3.2.1 การแบ่งลำดับชั้นแบบรวมกลุ่ม (Agglomerative hierarchical clustering)

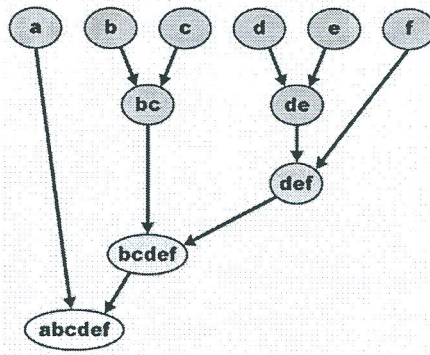
สมมติว่าข้อมูลนี้เป็นข้อมูลที่น่ามาแบ่งกลุ่มและ Euclidean distance อยู่ในรูป Distance metric ดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ข้อมูลดิบที่ยังไม่ได้ทำการแบ่งกลุ่ม

จะได้การแบ่งลำดับชั้นแบบ Dendrogram ของกลุ่มข้อมูลนี้ดังรูปที่ 2.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.18 ตัวอย่างโดยทั่วไปของการแบ่งลำดับชั้นแบบ Dendrogram

ขั้นตอนนี้สร้างลำดับชั้นขึ้นมาจากองค์ประกอบแต่ละองค์ประกอบ ด้วยการรวมกลุ่มไปเรื่อยๆ ดังในตัวอย่างของเราจะเห็นว่าเรามี 6 องค์ประกอบคือ {a}, {b}, {c}, {d}, {e} และ {f} ขั้นตอนแรกทำการพิจารณาว่าองค์ประกอบไหนที่จะนำมารวมกลุ่มกัน โดยปกติแล้วเราจะพิจารณาองค์ประกอบ 2 องค์ประกอบที่อยู่ใกล้กันมากที่สุดตามระยะห่างระหว่าง 2 องค์ประกอบนั้น

อีกทางเลือกหนึ่งในขั้นตอนนี้คือสร้าง Distance metric ซึ่งค่าใน Metric แถวที่ i -th หลักที่ j -th คือ ระยะห่างระหว่างองค์ประกอบที่ i -th และ j -th ดังนั้น ขณะที่ทำการรวมกลุ่มไปเรื่อยๆ แถวและหลักถูกรวมเป็นกลุ่ม และระยะห่างก็จะถูกปรับปรุง ซึ่งนี่ก็เป็นวิธีที่รู้จักกันดีในการแบ่งกลุ่มข้อมูล และมีข้อดีของการเก็บระยะห่างระหว่างกลุ่ม

สมมติ เราทำการรวมองค์ประกอบ 2 องค์ประกอบ b และ c ที่อยู่ใกล้กันมากที่สุด ตอนนี้เราก็จะได้กลุ่มข้อมูล {a}, {b, c}, {d}, {e}, {f} และเราต้องการที่จะรวมกลุ่มข้อมูลเหล่านี้ต่อ ซึ่งทำได้โดยเราจำเป็นต้องหาระยะห่างระหว่าง {a} และ {b, c} แล้วก็อธิบายถึงระยะห่างระหว่างกลุ่มข้อมูล 2 กลุ่มนั้น โดยปกติแล้วระยะห่างระหว่าง กลุ่มข้อมูล 2 กลุ่ม A และ B คือหนึ่งในคำอธิบายข้างล่างนี้

- ระยะห่างที่มากที่สุดระหว่างองค์ประกอบของแต่ละกลุ่ม (หรือเรียกว่า Complete linkage clustering) :

$$\max\{d(x, y) : x \in A, y \in B\} \quad (2.19)$$

- ระยะห่างที่น้อยที่สุดระหว่างองค์ประกอบของแต่ละกลุ่ม (หรือเรียกว่า complete linkage clustering):

$$\min\{d(x, y) : x \in A, y \in B\} \quad (2.20)$$

- ระยะห่างเฉลี่ยระหว่างองค์ประกอบของแต่ละกลุ่ม (หรือเรียกว่า Average linkage clustering):

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\frac{1}{\text{card}(A) \text{ card}(B)} \sum_{x \in A} \sum_{y \in B} d(x, y) \quad (2.21)$$

- ผลรวมของทุก Variance ภายในกลุ่ม
- การเพิ่มขึ้นของความแปรปรวนสำหรับกลุ่มข้อมูลที่จะถูกรวมกลุ่ม (Ward's criterion)
- ความน่าจะเป็นที่กลุ่มข้อมูลที่เลือกมาจากฟังก์ชันการกระจายเดียวกัน (V-linkage)

ในการรวมกลุ่มแต่ละครั้งทำให้เกิดระยะห่างระหว่างกลุ่มข้อมูลเพิ่มมากขึ้น มากกว่าตอนที่ทำการรวมกลุ่มไว้ก่อนหน้านี้ และจะหยุดทำการรวมกลุ่ม เมื่อกลุ่มนั้นห่างกันเกินไปที่จะรวมกันได้ (หลักการทางระยะห่าง) หรือ เมื่อมีจำนวนกลุ่มที่น้อยพอ (หลักการทางจำนวน)

2.3.2.2 การแบ่งลำดับขั้นแบบแยกกลุ่ม (Divisive hierarchical clustering)

การแบ่งกลุ่มข้อมูลแบบแยกกลุ่มจะตรงข้ามกับการแบ่งกลุ่มข้อมูลแบบรวมกลุ่ม โดยจะแบ่งแยกกลุ่มจากกลุ่มใหญ่ให้ย่อยไปเรื่อยๆ โดยเริ่มจากกลุ่มใหญ่ที่สุด คือกลุ่มเดียวมีข้อมูลทุกตัวอยู่ในกลุ่ม โดยจะพบว่าขั้นตอนจะเริ่มที่ส่วนล่างสุดของ tree diagram คือทิศทางของลูกศรจะชี้ในทางตรงกันข้ามกับแบบรวมกลุ่มนั่นเอง

ขั้นตอนนี้สร้างลำดับขั้นขึ้นมาจากองค์ประกอบแต่ละองค์ประกอบ ด้วยการแยกกลุ่มไปเรื่อยๆ โดยถ้าพิจารณาจากรูปที่ 2.18 กลุ่มข้อมูลที่ใหญ่ที่สุดประกอบด้วย {a, b, c, d, e, f} จากนั้นก็ทำการแยกกลุ่มย่อยไปอีก เป็น {a}, {b, c}, {d, e, f} ไปเรื่อยๆ ในการแยกกลุ่มแต่ละครั้งทำให้เกิดระยะห่างระหว่างกลุ่มข้อมูลน้อยลง น้อยกว่าตอนที่ทำการแยกกลุ่มไว้ก่อนหน้านี้ และจะหยุดทำการแยกกลุ่ม เมื่อกลุ่มนั้นใกล้กันเกินไปที่จะแยกกันได้ (หลักการทางระยะห่าง) หรือ เมื่อมีจำนวนกลุ่มที่มากพอ (หลักการทางจำนวน)

2.3.3 การแบ่งกลุ่มข้อมูลแบบตัดเป็นส่วน (Partition clustering)

การแบ่งกลุ่มข้อมูลแบบตัดเป็นส่วนสามารถแบ่งตามลักษณะความเป็นสมาชิกของข้อมูลได้เป็น 2 แบบด้วยกันคือ ฮาร์ดพาร์ติชัน (Hard Partition) และ ฟัซซีพาร์ติชัน (Fuzzy partition) ซึ่งสามารถอธิบายให้เข้าใจถึงนิยามความแตกต่างได้ดังนี้

ให้ $X = \{x_1, x_2, \dots, x_n\}$ เป็นข้อมูลดิบซึ่ง n เป็นจำนวนข้อมูลที่ต้องการจะแบ่งกลุ่มและ x_k เป็น x ใดๆ โดย $x_k \in \mathcal{R}^p$ (จำนวนจริงมีสถิติเท่ากับ p) ให้ $P(x)$ คือเพาเวอร์เซต (Power set) ของ x หรือก็คือเซตของเซตย่อย (Subset) ทุกอันที่เป็นไปได้ของ x จะได้ฮาร์ดพาร์ติชันของ X คือ $\{A_i \in P(x); 1 \leq i \leq c\}$ ซึ่งหมายความว่า เป็นเซตของข้อมูลที่อยู่ในกลุ่มเดียวกันและ c คือจำนวนกลุ่มข้อมูลที่ต้องการจะแบ่ง ในกรณีของฮาร์ดพาร์ติชันสามารถเขียนฟังก์ชันสมาชิกภาพ (Membership function) ของ x_k ใน A_i ดังใน (2.22) โดย u_{ik} คือค่าสมาชิกภาพ (Membership value) ซึ่งใช้บอกระดับความเป็นสมาชิกของ x_k ในเซต A_i

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$u_{ik} = \begin{cases} 1, & x_k \in A_i \\ 0, & x_k \notin A_i \end{cases} \quad (2.22)$$

ในกรณีของฮาดซีฟาที่ชั้น x_k จะต้องเป็นสมาชิกของกลุ่มข้อมูล เพียงกลุ่มเดียวเท่านั้นให้ $x_k \in X, A_i \in P(x), i=1,2,\dots,c$ และ $k = 1,2,\dots, n$ จะสามารถเขียนเป็นสมการ (2.22) และ (2.23)

$$u_{ik} \in \{0,1\}, 1 \leq i \leq c, 1 \leq k \leq n \quad (2.22)$$

$$\sum_{i=1}^c u_{ik} = 1, \forall k \in \{1,2,\dots,n\} \quad (2.23)$$

เนื่องจากกลุ่มข้อมูลแต่ละกลุ่มจะต้องมีสมาชิกอย่างน้อย 1 ตัวและไม่เกินจำนวนข้อมูลดิบทั้งหมด สามารถเขียนเป็น (2.24)

$$0 < \sum_{k=1}^n u_{ik} < n, \forall i \in \{1,2,\dots,c\} \quad (2.24)$$

นิยามที่ 2.1 ฮาดซีฟาที่ชั้น กำหนดให้ $X = \{x_1, x_2, \dots, x_n\}$ เป็นเซตของข้อมูลดิบและ V_{cm} คือเซตของเมตริกจำนวนจริง $U = [u_{ik}]$ ขนาด $c \times n$ แสดงดังสมการ (2.25) และ c เป็นเลขจำนวนเต็มซึ่งมีค่าอยู่ในช่วง $2 \leq c \leq n$ หรือก็คือจำนวนกลุ่มข้อมูลที่ต้องการจะแบ่งต้องมากกว่า 1 กลุ่มและต้องไม่เกินจำนวนข้อมูลดิบทั้งหมด ดังนั้นเซตฮาดซีฟาที่ชั้นสำหรับ X แสดงดังสมการ (2.26)

$$V_{cm} = \{U_1, U_2, \dots, U_m / u_{ik} \in \{0,1\} \text{ และ (2.26) เป็นจริง}\} \quad (2.25)$$

$$M_c = \{U \in V_m\} \quad (2.26)$$

ตัวอย่างที่ 2.1 ให้ X เป็นเซตของรถ 3 คัน $X = \{x_1 = Ford, x_2 = Toyota, x_3 = Chrysler\}$ ถ้าต้องการแบ่งเป็น 2 กลุ่มข้อมูล ($c = 2$) จะได้ฮาดซีฟาที่ชั้นของ X_3 ดังนี้

$$U_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad U_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

ถ้าโคทีเรีย (Criteria) ของการแบ่งกลุ่มข้อมูลเป็นการแบ่งกลุ่มข้อมูลระหว่างรถอเมริกาและญี่ปุ่นแล้ว U_3 น่าจะเป็นพาที่ชั้นที่เหมาะสมที่สุด ในความเป็นจริงเซตของรถอเมริกาและรถญี่ปุ่นไม่แน่นอน เพราะอุปกรณ์บางส่วนของรถอเมริกาเป็นของญี่ปุ่น และอุปกรณ์บางส่วนของรถญี่ปุ่นเป็นของอเมริกา ดังนั้นการแก้ปัญหาคือให้ u_{ik} มีค่าอยู่ในช่วง $[0, 1]$ หรือเป็นแบบฟัซซีนั่นเองซึ่งสามารถเขียนนิยามของฟัซซีพาที่ชั้นได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นิยามที่ 2.2 ฟัชซีเซตพหุชั้น ให้ X และ c เป็นไปตามนิยามที่ 2.1

$$V_{cm} = \{U_1, U_2, \dots, U_m / u_{ik} \in [0,1] \text{ และ } (2.28) \text{ เป็นจริง}\} \quad (2.27)$$

$$M_{fc} = \{U \in V_{cm}\} \quad (2.28)$$

ดังนั้นจากตัวอย่างที่ 5.1 จะสามารถเขียนในรูปแบบฟัชซีเซตพหุชั้นได้เป็น

$$U_1 = \begin{bmatrix} 0.9 & 0.2 & 0.9 \\ 0.1 & 0.8 & 0.1 \end{bmatrix}$$

2.3.3.1 ฮาร์ดพาร์ติชัน (Hard partition)

การแบ่งกลุ่มข้อมูลแบบฮาร์ดพาร์ติชันมีอยู่หลายวิธีด้วยกัน ในที่นี้จะกล่าวเฉพาะวิธีที่สนใจคือ

2.3.3.1.1 การแบ่งกลุ่มข้อมูลแบบเคมีน K-means

อัลกอริทึมแบบเคมีน (K-means clustering) เป็นอัลกอริทึมที่ใช้ในการจัดกลุ่มข้อมูลประเภทหนึ่งที่มีการประยุกต์ใช้งานกันอย่างแพร่หลาย โดยชุดของข้อมูลจะถูกจัดออกเป็นกลุ่ม k กลุ่ม รูปแบบขั้นตอนในการจัดกลุ่มของเคมีนจะอาศัยหลักการคำนวณระยะห่างเฉลี่ยระหว่างพิกัดของข้อมูลกับพิกัดของจุดศูนย์กลางของกลุ่ม (centroids) ข้อมูลจะถูกจัดแยกให้อยู่ในกลุ่ม k ใดๆ ขึ้นอยู่กับระยะห่างที่สั้นที่สุดระหว่างพิกัดของข้อมูลแต่ละตัวกับพิกัดของจุดศูนย์กลางกลุ่มนั้นๆ เมื่อเทียบระยะห่างระหว่างจุดศูนย์กลางกลุ่มจุดอื่นๆ ทุกจุด การจัดกลุ่มจะทำการวนซ้ำกันไปเรื่อยๆ โดยจุดศูนย์กลางของแต่ละกลุ่มก็จะมีเปลี่ยนแปลงไปมาขึ้นอยู่กับ ระยะห่างที่คำนวณได้ในแต่ละรอบเมื่อคิดเทียบกับพิกัดของข้อมูลทุกๆ จุด การจัดกลุ่มจะสิ้นสุดลงก็ต่อเมื่อ จุดศูนย์กลางของกลุ่มข้อมูลทั้ง k กลุ่มไม่มีการเปลี่ยนแปลงอีก เมื่อเราพิจารณาค่า object function ของอัลกอริทึมแบบเคมีน จะได้ตามสมการที่ (2.29)

$$J = \sum_{j=1}^k \sum_{i=1}^n x_i^{(j)} - c_j^2 \quad (2.29)$$

โดยที่ J เป็นระยะห่างระหว่างจุดพิกัดของข้อมูล x ใดๆ กับจุดศูนย์กลางของกลุ่ม C เมื่อ k เป็นจำนวนกลุ่มที่จะจัด และ n เป็นจำนวนของข้อมูลทั้งหมด

ขั้นตอนของ อัลกอริทึม K-means มีดังต่อไปนี้

- ขั้นตอนที่ 1 กำหนดจุดศูนย์กลางกลุ่ม (centroid) จำนวน k จุดโดยเลือกให้ห่างซึ่งกันและกันมากที่สุด
- ขั้นตอนที่ 2 จัดวัตถุให้อยู่ในกลุ่มของจุดศูนย์กลางกลุ่มถ้าวัตถุนั้นอยู่ใกล้จุดศูนย์กลางกลุ่มนั้นๆ มากกว่า จุดศูนย์กลางกลุ่มอื่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 3 ทำการคำนวณจุดศูนย์กลางกลุ่มใหม่จากค่าเฉลี่ยพิสัยของวัตถุที่อยู่ในกลุ่ม

ขั้นตอนที่ 4 ทำขั้นตอน (3) และ (4) ซ้ำ จนจุดศูนย์กลางกลุ่มใหม่ไม่เปลี่ยนแปลง

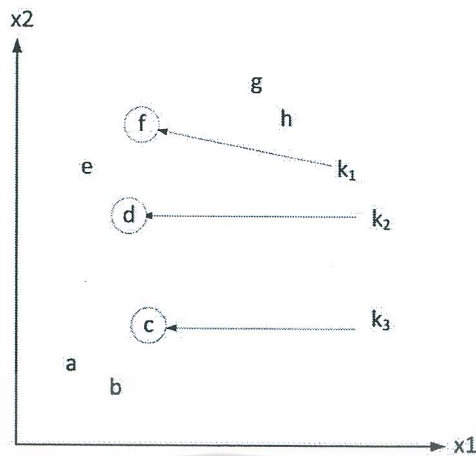
ตัวอย่างที่ 2.2 ตัวอย่างนี้จะแสดงขั้นตอนการแบ่งกลุ่มข้อมูลโดยเคมีน โดยใช้ข้อมูลดังตารางที่ 2.2 ต่อไปนี้

ตารางที่ 2.2 การแบ่งกลุ่มข้อมูล

รายการ (Item)	สังเกต (Observation)	
	X1	X2
A	1	2
B	2	1
C	3	3
D	2.5	5
E	1.5	6
F	3	7
G	5	8
H	6	7

ขั้นตอนวิธีการของเคมีน

1. ทำการสุ่มส่วนของข้อมูลเป็นกลุ่มเซตย่อย k เซต โดยที่ผู้ทำการแบ่งกลุ่มข้อมูล (cluster) จะต้องทราบจำนวนกลุ่ม (k) ที่ต้องการแบ่ง ซึ่งสามารถแสดงได้ดังรูปที่ 2.19



รูปที่ 2.19 การสุ่มข้อมูลเพื่อทำการแบ่งกลุ่ม

จากรูปที่ 2.19 มีการแบ่งข้อมูลออกเป็น 3 กลุ่ม (k) โดยทำการสุ่มจุดเริ่มต้นของแต่ละกลุ่ม (k) นำข้อมูลทุกตัวมาคำนวณระยะทางกับ k การคำนวณระยะทางทำได้หลายวิธีเช่น ระยะทางยูคลิเดียน (Euclidean distance) โดยถ้า d คือระยะทางยูคลิเดียนแล้วจะได้ระยะทางคือ

$$d(i, j) = \sqrt{(|x_{i1} - x_{j1}|^2 + \dots + |x_{ip} - x_{jp}|^2)} \quad (2.30)$$

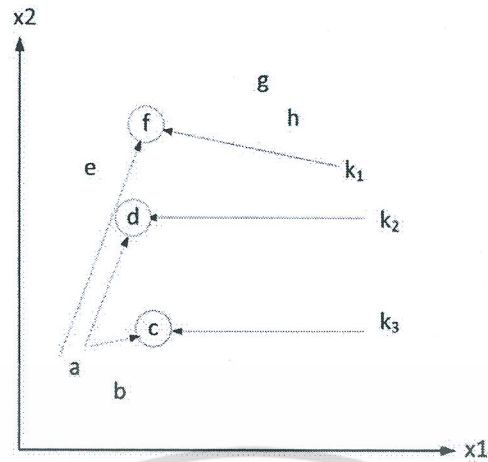
โดย $d(i, j) \geq 0$, $d(i, i) = 0$, $d(i, j) = d(j, i)$, $d(i, j) \leq d(i, k) + d(k, j)$ ซึ่งสามารถคำนวณค่าระยะทางยูคลิเดียน โดยใช้ข้อมูลจากตารางที่ 3.1 และ รูปที่ 2.18 ซึ่งได้มีการกำหนดจุดเริ่มต้นของแต่ละกลุ่มข้อมูลคือ จุด c, จุด d และจุด f ดังนั้นเมื่อคำนวณระยะทางจากจุด a ไปยังจุดเริ่มต้นของแต่ละกลุ่มข้อมูลจะได้ระยะทาง ดังนี้

$$\text{ระยะทางจาก a ไป c} = \sqrt{(|1-3|^2 + |2-3|^2)} = 2.24$$

$$\text{ระยะทางจาก a ไป d} = \sqrt{(|1-2.25|^2 + |2-5|^2)} = 3.12$$

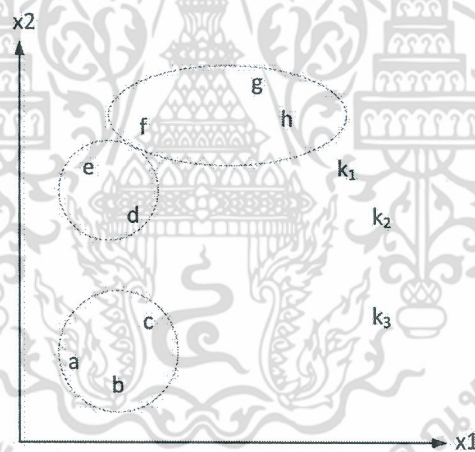
$$\text{ระยะทางจาก a ไป f} = \sqrt{(|1-3|^2 + |2-7|^2)} = 5.39$$

จากการคำนวณระยะทางระหว่างข้อมูลกับกลุ่มข้อมูล(k) นั้นสามารถแสดงได้ดังรูปที่ 2.20



รูปที่ 2.20 ระยะทางของข้อมูลกับข้อมูลเริ่มต้นของกลุ่ม

หลังจากนั้นทำการหาระยะทางของข้อมูลทุกจุดกับกลุ่มข้อมูล (k) ทุกกลุ่มและตรวจดูว่าข้อมูลห่างจากกลุ่มข้อมูล (k) ชัดไหนน้อยที่สุดก็นำข้อมูลไปสังกัดอยู่ในกลุ่มข้อมูล (k) นั้น และได้กลุ่มข้อมูลขึ้นมา โดยแสดงดังรูปที่ 2.21



รูปที่ 2.21 ข้อมูลที่ได้รับการแบ่งโดยอาศัย k

2. เมื่อได้กลุ่มข้อมูลแต่ละกลุ่มแล้ว ทำการคำนวณหาจุดศูนย์กลางของกลุ่มที่เรียกว่า เซ็นทรอยด์ (Centroid) โดยคำนวณจากสมการที่ (2.34)

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ik} \quad (2.34)$$

โดยที่ $x = \{x \in k | k \text{ คือ กลุ่มข้อมูล}\}$

$$x = \{x_1, x_2, x_3, \dots, x_i\}$$

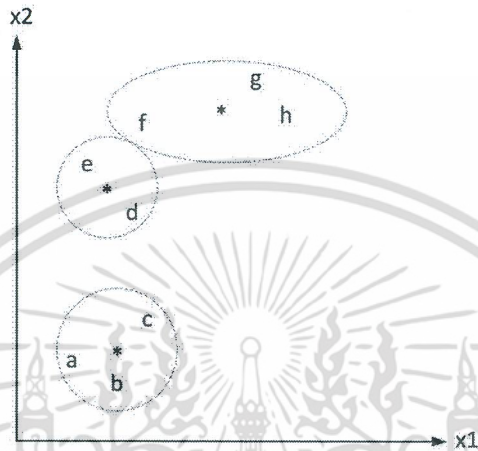
M_k คือ ค่ากลางของข้อมูลในกลุ่มที่ k

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

n_k คือ จำนวนข้อมูลของกลุ่มที่ k

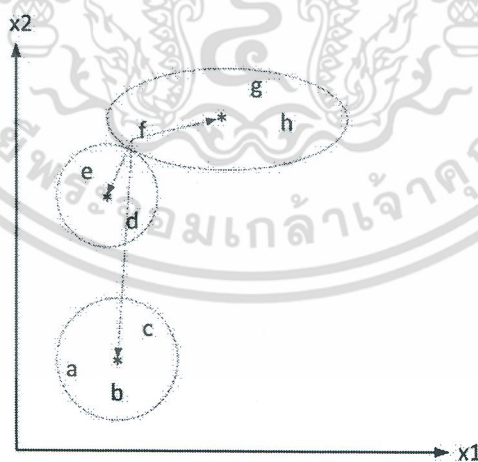
x_{ik} คือ ข้อมูล x ตัวที่ i ของกลุ่มข้อมูลที่ k

เมื่อได้จุดศูนย์กลางของกลุ่มข้อมูลแล้ว จึงนำข้อมูลทุกตัวในแต่ละกลุ่มข้อมูล (Cluster) ไปคำนวณกับจุดศูนย์กลาง (Centroid) ของทุกกลุ่มข้อมูล (Cluster) ดังแสดงในรูปที่ 2.22

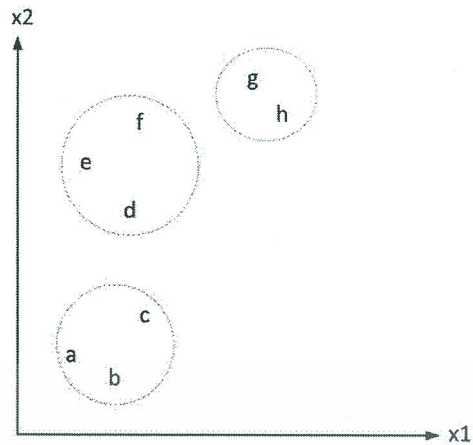


รูปที่ 2.22 จุดศูนย์กลางของข้อมูลในแต่ละกลุ่มข้อมูล

3. ตรวจสอบระยะทางระหว่างข้อมูลกับจุดศูนย์กลางของข้อมูล (Centroid) โดยข้อมูลที่ห่างจากจุดศูนย์กลางข้อมูลไหนสั้นที่สุดก็นำข้อมูลไปสังกัดกับกลุ่มข้อมูลดังกล่าว จากจุดนี้จะทำให้เกิดการย้ายกลุ่มข้อมูลซึ่งสามารถแสดงได้ดังรูปที่ 2.23



รูปที่ 2.23 การวัดระยะทางของข้อมูลกับจุดศูนย์กลางข้อมูล (Centroid)



รูปที่ 2.24 กลุ่มข้อมูลใหม่

4. ทำการคำนวณจุดศูนย์กลางของข้อมูลและกำหนดข้อมูลให้กลุ่มข้อมูลใหม่ ดังรูปที่ 2.24 ไปเรื่อยๆจนกว่าจะถึงเงื่อนไขให้หยุด (Stopping criterion) โดยเงื่อนไขให้หยุดคือ 1) ข้อมูลในแต่ละกลุ่มถึงสถานะเสถียรไม่สามารถเปลี่ยนกลุ่มที่ดีกว่านี้ได้ 2) ค่าผิดพลาดที่ต่ำสุด (Minimum square error) สำหรับทุกกลุ่มข้อมูลโดยหาค่าผิดพลาดที่ต่ำสุดได้จาก

$$E_k^2 = \sum_{k=1}^K e_k^2 \quad \text{โดยที่} \quad e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2 \quad (2.35)$$

โดยที่ $x = \{x \in k | k \text{ คือ กลุ่มข้อมูล}\}$

$$x = \{x_1, x_2, x_3, \dots, x_i\}$$

x_{ik} คือ ข้อมูลที่ i ของกลุ่มที่ k

n_k คือ จำนวนข้อมูลของกลุ่มที่ k

e_k^2 คือ ค่าผิดพลาดในกลุ่มที่ k

E_k^2 คือ ค่าผิดพลาดรวมของทุกกลุ่มข้อมูล

2.3.4 การประเมินความถูกต้องการแบ่งกลุ่ม (Cluster Validity)

กระบวนการของการประเมินผลของอัลกอริทึมแบ่งกลุ่มจะมีหลักเกณฑ์อยู่ 2 แบบ ที่มีเพื่อนำเสนอสำหรับการประเมินผลและการเลือกวิธีแบ่งกลุ่มที่ดีที่สุด นั่นคือ

1. Compactness : ซึ่งจำนวนสมาชิกของภายในกลุ่มจะต้องมีความใกล้ชิดกันมากที่สุดเท่าที่จะเป็นไปได้ วิธีการวัดที่นิยมใช้กันทั่วไปคือค่าความแปรปรวน

2. Separation : กลุ่มแต่ละกลุ่มควรที่จะถูกแยกให้ห่างกันมากที่สุด ซึ่งมีวิธีอยู่ 3 วิธีในการวัดระยะห่างระหว่างสองกลุ่มที่แตกต่างกัน หรือ ระยะห่างระหว่างสมาชิกที่ใกล้ที่สุดของกลุ่ม หรือ ระยะห่างระหว่างสมาชิกที่ไกลที่สุด และ ระยะห่างระหว่างจุดศูนย์กลางของจุดศูนย์กลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นับญาติให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากเหตุผลดังกล่าวจะมีวิธีการอยู่ 3 ประเภทสำหรับใช้ประเมินผลของอัลกอริทึมการแบ่งกลุ่ม

1. External criteria
2. Internal criteria
3. Relative criteria

โดยทั้งเกณฑ์แบบ Internal และ External จะอยู่บนพื้นฐานของวิธีการทางสถิติและต้องการ การคำนวณที่ซับซ้อน วิธีการวัดความถูกต้องแบบ External จะประเมินการแบ่งกลุ่มบนพื้นฐานตามสัญชาตญาณของผู้ใช้ สำหรับแบบ Internal จะอยู่บนพื้นฐานของเมตริกซ์ซึ่งเกี่ยวข้องกับกลุ่มข้อมูลและรูปแบบการแบ่งกลุ่ม ข้อเสียหลักของสองวิธีนี้คือ ความซับซ้อนของการคำนวณ

สำหรับวิธีการแบบ Relative จะเป็นการเปรียบเทียบรูปแบบของการแบ่งกลุ่มที่ต่างกัน อัลกอริทึมการแบ่งกลุ่มหนึ่งชนิดหรือมากกว่าจะเป็นการดำเนินการหลายครั้งด้วยอินพุตพารามิเตอร์ที่ต่างกันบนกลุ่มข้อมูลเดียวกัน เป้าหมายของวิธีแบบ Relative จะเป็นการเลือกรูปแบบการแบ่งกลุ่มที่ดีที่สุดจากผลลัพธ์ที่ต่างกัน โดยการใช้การเปรียบเทียบของดัชนีความถูกต้อง (Validity index) โดยวิธีต่างๆ ตารางที่ 2.3 แสดงความหมายของสัญลักษณ์ที่ใช้ในการประเมินนี้

ตารางที่ 2.3 เครื่องหมายที่ใช้แทนใน Validity index

Notation	Meaning
n_c	Number of cluster
d	Number of dimension
$d(x, y)$	Distance between two data element
\bar{X}_j	Expected value in the j^{th} dimension
$\ X\ $	$\sqrt{X^T X}$, where X^T is a column vector
n_{ij}	Number of element in i^{th} cluster j^{th} dimension
n_j	Number of element in j^{th} dimension in the whole data set
v_i	Centre point of the i^{th} cluster
c_i	i^{th} cluster
$\ c_i\ $	Number of element in i^{th} cluster

1. Davies Bouldin Index เป็นการวัดโดยอยู่บนพื้นฐานบนการวัดความเหมือนกันของกลุ่ม R_{ij} ซึ่งอยู่บนหลักการของการวัดการกระจายของกลุ่ม s_i และการวัดความไม่เหมือนกันของกลุ่ม d_{ij} สำหรับการวัดความเหมือนกันของกลุ่ม R_{ij} สามารถถูกนิยามได้อย่างอิสระแต่มันจะ Satisfy สถานะต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- $R_{ij} \geq 0$
- $R_{ij} = R_{ji}$
- if $s_i = 0$ และ $s_j = 0$ ดังนั้น $R_{ij} = 0$
- if $s_j > s_k$ และ $d_{ij} = d_{ik}$ ดังนั้น $R_{ij} > R_{ik}$
- if $s_j = s_k$ และ $d_{ij} < d_{ik}$ ดังนั้น $R_{ij} > R_{ik}$

ปกติแล้ว R_{ij} เป็นการนิยามได้ ดังนี้

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \quad (2.36)$$

$$d_{ij} = d(v_i, v_j), \quad s_i = \frac{1}{\|C_i\|} \sum_{x \in C_i} d(x, v_i) \quad (2.37)$$

ดังนั้น DB index จะหาได้ ดังนี้

$$DB = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i, \quad (2.38)$$

$$R_i = \max_{j=1..n_c, i \neq j} (R_{ij}), \quad i=1..n_c \quad (2.39)$$

ค่าดัชนี DB จะวัดค่าเฉลี่ยของความเหมือนระหว่างแต่ละกลุ่ม และกลุ่มที่เหมือนกันมากที่สุด ดังนั้นค่าดัชนี DB จะมีค่าต่ำหากความสามารถในการแยกกลุ่มของแต่ละกลุ่มข้อมูลมีประสิทธิภาพ

2.4 เกณฑ์การวัดประสิทธิภาพของการบีบอัดข้อมูล

ในการวัดประสิทธิภาพของการลดข้อมูล ในเชิงคณิตศาสตร์ สามารถกระทำได้หลายวิธีดังนี้

2.4.1 ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error, MSE)

เราสามารถใช้ค่าผิดพลาดกำลังสองเฉลี่ยในการวัดประสิทธิภาพของการลดข้อมูล คือถ้าความผิดพลาดกำลังสองเฉลี่ยมีค่าน้อย ภาพที่ได้จากการแปลงกลับที่ถูกลดข้อมูลแล้วมีความผิดเพี้ยนไปจากภาพต้นแบบน้อย ซึ่งเราสามารถหาค่าความผิดพลาดกำลังสองเฉลี่ยของการลดข้อมูลภาพขนาด $M * N$

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2 \quad (2.40)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ M คือจำนวนจุดภาพตามความกว้างของภาพ

N คือจำนวนจุดภาพตามความสูงของภาพ

$f(i, j)$ คือระดับความเทา (Gray Level) ที่ตำแหน่ง $((i, j))$ ของภาพต้นแบบ

$\hat{f}(i, j)$ คือระดับความเทาที่ตำแหน่ง $((i, j))$ ของภาพที่ถูกสร้างกลับขึ้นมาใหม่

ในบางครั้งเราอาจพบว่ามีการใช้ค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ย (Root Mean Square, $RMSE$) ในการวัดประสิทธิภาพการลดข้อมูลแทนค่า MSE ซึ่งค่า $RMSE$ สามารถหาได้จาก

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2} \quad (2.41)$$

2.4.2 สัญญาณต่อสัญญาณรบกวนและค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด

อีกวิธีหนึ่งของการอธิบายประสิทธิภาพของอัลกอริทึมคือ ค่าสัดส่วนสัญญาณต่อสัญญาณรบกวน (Signal-to-Noise-ratio, SNR) ในบางครั้งเรียกว่าค่าสัญญาณต่อสัญญาณรบกวนการควอนไทซ์ (Signal-to-Quantization-Noise) ซึ่งในการวัดค่าจะมีความสัมพันธ์ของกำลังของสัญญาณกับความเพี้ยนของสัญญาณ สามารถหาได้จากการนำค่าเฉลี่ยของภาพต้นแบบยกกำลังสองแล้วหารด้วย MSE โดยมีหน่วยเป็น (dB) ดังนี้

$$SNR = 10 \log_{10} \frac{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j)]^2}{MSE} \quad (\text{dB}) \quad (2.42)$$

$$SNR = 10 \log_{10} \frac{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j)]^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2} \quad (\text{dB}) \quad (2.43)$$

การวัดที่คล้าย ๆ กันคือ ค่าสัดส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak-Signal-to-Noise Ratio, $PSNR$) โดยสามารถหาได้ดังนี้

$$PSNR = 10 \log_{10} \frac{Peak^2}{MSE} \quad (\text{dB}) \quad (2.44)$$

เมื่อ $Peak$ แทนค่าสูงสุดของขอบเขตของข้อมูล ทั้ง SNR และ $PSNR$ เป็นวิธีที่นิยมใช้อธิบายคุณภาพของการลดข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับภาพที่มีจุดภาพแทนด้วยข้อมูลขนาด n บิต หรืออาจกล่าวได้ว่ามีจำนวนระดับความเทา (Gray Scale) เป็น 2^n บิตจะได้

$$PSNR = 10 \log_{10} \frac{(2^n - 1)}{MSE} \quad (\text{dB}) \quad (2.45)$$

ดังนั้นภาพที่มีจุดภาพแทนด้วยข้อมูลขนาด 8 บิต หรือมีจำนวนระดับความเทาเป็น 256 จะได้

$$PSNR = 10 \log_{10} \frac{(2^8 - 1)^2}{MSE} \quad (\text{dB}) \quad (2.46)$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (\text{dB}) \quad (2.47)$$

$$PSNR = 10 \log_{10} \frac{255}{RMSE} \quad (\text{dB}) \quad (2.48)$$

2.4.3 การวัดประสิทธิภาพของการบีบอัดข้อมูลภาพด้วยการสังเกตด้วยสายตา (Perceptual Measures)

เป็นการยากที่จะใช้คณิตศาสตร์ในการบอกปริมาณความผิดพลาด ตามการสังเกตด้วยสายตาของมนุษย์ และการวัดเชิงคณิตศาสตร์ก็ไม่มีความสัมพันธ์โดยตรงกับคุณภาพของภาพ หมายความว่า อัลกอริทึมอาจจะทำให้เกิดความผิดพลาดเชิงคณิตศาสตร์มากกว่า แต่เมื่อสังเกตด้วยสายตาแล้ว อาจจะมีคุณภาพดีกว่า ซึ่งกรณีนี้อาจเกิดขึ้นได้ถ้าอัลกอริทึมให้เกิดความผิดพลาดในพื้นที่หรือความถี่มากขึ้นซึ่งตาไม่สามารถวัดได้ โดยทั่ว ๆ ไปรหัสที่สร้างขึ้นด้วยการควอนไทซ์แบบหยาบ ๆ เช่น ผลจากบล็อก (Block Effect) ผลจากทางเดิน (Contour Effect) ผลจากวงแหวน (Ringing Effect) เป็นต้น จะไม่แสดงความเพี้ยนทางคณิตศาสตร์ให้เห็นอย่างชัดเจน แต่สามารถที่จะสังเกตเห็นได้ง่ายด้วยสายตาของมนุษย์

2.4.4 อัตราบิต (bit rate)

อัตราบิต คือค่าเฉลี่ยของจำนวนบิตต่อจุดภาพ (Bits Per Pixel, BPP) ของภาพที่ถูกลดข้อมูล โดยสามารถคำนวณได้จากอัตราส่วนระหว่างจำนวนบิตข้อมูลทั้งหมดของภาพที่ถูกลดข้อมูลต่อจำนวนจุดภาพทั้งหมดของภาพต้นแบบ ดังสมการต่อไปนี้

$$BPP = \frac{\text{Bits_Total}}{\text{Pixels_Total}} \quad (2.49)$$

โดยที่ Bits_Total = จำนวนบิตทั้งหมดของจุดภาพที่ถูกลดข้อมูล

Pixels_Total = จำนวนจุดภาพทั้งหมดของภาพต้นแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.5 อัตราการบีบอัดข้อมูล (Compression ratio, Cr)

อัตราการบีบอัดข้อมูลคืออัตราส่วนระหว่างจำนวนบิตข้อมูลที่ใช้แทนภาพต้นแบบ และจำนวนบิตข้อมูลที่ใช้แทนภาพที่ถูกทำการบีบอัดข้อมูล โดยสามารถหาได้ดังสมการต่อไปนี้

$$Cr = \frac{Uncompress_Bits}{Compress_Bits} \quad (2.50)$$

โดยที่ $Uncompress_Bits$ = จำนวนบิตข้อมูลของภาพต้นแบบ

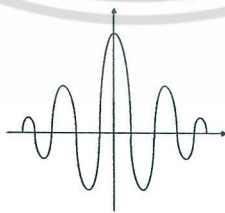
$Compress_Bits$ = จำนวนบิตข้อมูลของภาพที่ถูกลดข้อมูล

2.5 การแปลงเวฟเลต (Wavelet Transform)

การใช้ Wavelet เป็นการวิเคราะห์สัญญาณที่ระดับความละเอียดหลายระดับ (Multi Resolution Analysis) โดยการแปลงเวฟเลตจากภาพจะได้ผลลัพธ์เป็น สัมประสิทธิ์ของเวฟเลต (Wavelet Coefficient) การทำ Discrete Wavelet Transform (DWT) (Jensen and Cour-Harbo : 2001) จะได้จำนวนสัมประสิทธิ์เท่ากับจำนวนพิกเซลในภาพต้นฉบับ เพียงแต่แยกข้อมูลออกเป็นสองส่วน คือ ส่วนที่เรียกว่าส่วน Approximation กับส่วน Prediction Residual ซึ่งก็หมายถึงส่วนความถี่ต่ำกับ ความถี่สูง การแปลงเวฟเลต สามารถทำได้หลายลำดับชั้น ทำให้เราวิเคราะห์ภาพได้ในหลาย ๆ ช่อง ความถี่ (Mallet : 1989) ซึ่งอาจจะมีเพียงสัมประสิทธิ์เพียงช่วงเท่านั้นที่มีค่ามากและสำคัญ บางช่วง สามารถตัดทิ้งได้ เนื่องจากมีค่าต่ำ การตัดทิ้งไม่มีผลกระทบต่ออารมณ์มองเห็นมากนักในเวลาทำการแปลงสัญญาณกลับ (Reconstruction)

ทฤษฎีเวฟเลต (Daubechies : 1992) เป็นคณิตศาสตร์ที่ใช้อธิบายการสร้างโมเดลของสัญญาณระบบหรือกระบวนการทางฟิสิกส์ใด ๆ ซึ่งประกอบขึ้นด้วยเซตของสัญญาณเฉพาะหรือระบบย่อยมารวมกันเป็นสัญญาณหรือระบบนั้น ๆ สัญญาณเฉพาะนี้จะเป็นคลื่นเล็ก ๆ ซึ่งถูกเรียกว่า “เวฟเลต”

ลักษณะของเวฟเลตจะเป็นคลื่นที่มีการเปลี่ยนแปลงอย่างต่อเนื่อง (Oscillatory) และขนาดของคลื่นจะลดลงสู่ศูนย์อย่างรวดเร็วทั้งสองด้าน ดังรูปที่ 2.25



รูปที่ 2.25 คลื่นเวฟเลต

การนำเวฟเลตหลาย ๆ อันมารวมกันเป็นเซต เพื่ออธิบายสัญญาณใด ๆ โดยที่คลื่นเวฟเลตแต่ละตัวจะมีโครงสร้างมาจากฟังก์ชันเดียวกัน ซึ่งฟังก์ชันนี้จะเป็นเวฟเลตต้นกำเนิดที่เรียกว่าเวฟเลต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แม่ (Mother Wavelet) คลื่นเวฟเลตแต่ละอันจะอยู่ในเซตของเวฟเลตนี้ โดยคลื่นแต่ละคลื่นจะเกิดจากการสเกล (Scaling) a และการเลื่อนตำแหน่ง (Translation or Shifting) b พารามิเตอร์ a จะหมายถึงการสเกลหรือการหดเข้าหรือขยายออก ส่วนพารามิเตอร์ b จะเป็นการเลื่อนตำแหน่งการเกิดคลื่นของเวฟเลตบนแกนเวลา นอกจากนี้เวฟเลตที่ถูกสเกลจะต้องทำการนอร์มอลไลซ์ด้วย $\frac{1}{\sqrt{a}}$ เสมอเพื่อให้เวฟเลตที่ถูกสเกลไปแล้วยังมีพลังงานเช่นเดียวกับเวฟเลตแม่ ถ้าให้ $g(t)$ เป็นฟังก์ชันของเวฟเลตแม่ จะสามารถเขียนเป็นสมการทั่วไปของเวฟเลตที่ตำแหน่ง a, b ใด ๆ ได้ซึ่งมีความสัมพันธ์กับเวฟเลตแม่ดังนี้

$$g_{a,b}(t) = \frac{1}{\sqrt{a}} g\left(\frac{t-b}{a}\right) \quad (2.51)$$

2.5.1 พื้นฐานของการสเกลและการเลื่อนตำแหน่งของเวฟเลต

ทฤษฎีเวฟเลตจะใช้ในการอธิบายสิ่งใดสิ่งหนึ่ง เหมือนการแตกสิ่งเหล่านั้นออกเป็น ส่วนประกอบเล็กๆสัมพันธ์กัน จะอยู่ในรูปของเวฟเลตที่ถูกสเกลและเลื่อนตำแหน่งและมีค่าต่าง น้ำหนักต่างกันขึ้นอยู่กับเวฟเลต หรือที่เรียกอีกอย่างว่า การแตกกระจายเวฟเลต (Wavelet Deconstruction) ส่วนการแปลงกลับเวฟเลต (Inverse Wavelet Transform, IWT) หรือเรียกอีกอย่างหนึ่งว่าการรวมกลับเวฟเลต (Wavelet Reconstruction)

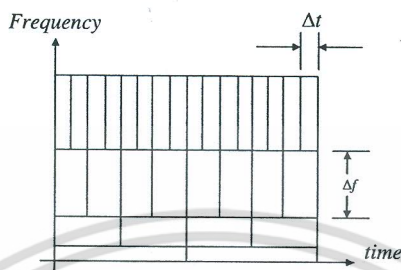
ถ้าให้ $f(t)$ เป็นสัญญาณที่ประกอบด้วยหลายๆความถี่เกิดขึ้นถ้าเราพิจารณาในแกน 2 มิติโดย แกนหนึ่งเป็นแกนเวลาและแกนหนึ่งเป็นแกนความถี่ดังนั้นสามารถเขียนสมการสำหรับการแปลงเวฟเลตแบบต่อ-เนื่อง (CWT) ที่ใช้ในการกระจาย (Decomposition) สัญญาณของฟังก์ชันใดๆ $f(t)$ ด้วยเวฟเลตแม่ $g(t)$ เป็นดังนี้

$$W_g f(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) g^*\left(\frac{t-b}{a}\right) dt \quad (2.52)$$

โดย “*” หมายถึง Complex Conjugate ของ $g(t)$, a เป็นพารามิเตอร์ของการสเกลและ b เป็นพารามิเตอร์ของการเลื่อนตำแหน่ง สัญลักษณ์ $W_g f(ab)$ จะหมายถึง สัญญาณอินพุตที่ขึ้นอยู่กับตัวแปรอิสระ t จะถูกส่ง (Map) ไปเป็นฟังก์ชัน 2 มิติที่ขึ้นอยู่กับตัวแปรอิสระ a, b โดยที่ $W_g f(a,b)$ จะเป็นสัมประสิทธิ์ของเวฟเลต และการสเกลกับการเลื่อนตำแหน่ง เขียนได้เป็น $t = \frac{t-b}{a}$

การแปลงเวฟเลตซึ่งจะเป็นการทรานส์ฟอร์มผ่านฟังก์ชันหน้าต่างจะสามารถเลื่อนตำแหน่งและปรับเปลี่ยนช่วงเวลาให้เหมาะสมกับความถี่ที่จะวิเคราะห์โดยเรียกฟังก์ชันหน้าต่างนี้ว่า “ฟังก์ชันเวฟเลต” (Wavelet Function)

$g(t)$ จะเป็นฟังก์ชันเวฟเลตแม่ ซึ่งถูกเลื่อนตำแหน่งและถูกสเกลโดยพารามิเตอร์ a และ b ซึ่งระยะเวลาและความถี่จะมีลักษณะการเปลี่ยนแปลงช่วงเวลาและความถี่ที่สัมพันธ์กันคือ ในช่วงความถี่สูงจะใช้ช่วงการวิเคราะห์เพียงสั้นและในช่วงความถี่ต่ำๆจะมีช่วงเวลาในการวิเคราะห์ที่ยาวขึ้น (Δf ต่ำ Δt สูง)จะเป็นไปดังรูปที่ 2.26 ซึ่งเหมาะสมกับสัญญาณในทางปฏิบัติ



รูปที่ 2.26 ระยะเวลาและระนาบความถี่การเปลี่ยนแปลงช่วงเวลาและความถี่ที่สัมพันธ์กัน

ของการแปลงเวฟเลต

2.5.2 การแปลงเวฟเลตแบบเต็มหน่วย (Discrete Wavelet Transform)

ปริภูมิเวกเตอร์ในการวิเคราะห์สัญญาณ พิจารณา a, b เป็นจำนวนจริงใด ๆ และ i, j เป็นเวกเตอร์หนึ่งหน่วย (Unit Vector) ในปริภูมิเวกเตอร์ (Vector Space: V) ถ้า $ai + bj \in V$ แล้วจะได้ว่า เวกเตอร์ใด ๆ ที่ประกอบขึ้นจากเวกเตอร์หนึ่งหน่วยตัวหรือหลายตัว แล้วเวกเตอร์นั้นยังคงอยู่ในปริภูมิเวกเตอร์นี้ (V) ดังนั้นเวกเตอร์หนึ่งหน่วยดังกล่าวจะเรียกว่าเป็น เวกเตอร์พื้นฐาน (Basis Vector) แต่เราจะพิจารณาเวกเตอร์เป็นฟังก์ชัน นั่นคือจะเรียกเวกเตอร์พื้นฐานนี้ใหม่ว่า ฟังก์ชันพื้นฐาน (Basis Function : b.f.) และถ้าฟังก์ชันพื้นฐานของเวกเตอร์ทำผลคูณภายใน (Inner product) กัน เช่น $\langle i, j \rangle$ แล้วมีค่าเป็นศูนย์จะได้ว่าฟังก์ชันพื้นฐานดังกล่าวจะแสดงคุณสมบัติออร์ทอโกนัล “Orthogonal” และถ้าฟังก์ชันพื้นฐานของเวกเตอร์ทำผลคูณภายในตัวเองแล้วมีค่าเป็นหนึ่ง ฟังก์ชันพื้นฐานที่มีคุณสมบัติทั้งสองก็จะเรียกว่า ออร์ทอนอโมล (Orthonormal)

ถ้าพิจารณาในลักษณะของสัญญาณอาจจะกล่าวได้ว่า ฟังก์ชันพื้นฐานเป็นสัญญาณพื้นฐานเล็ก ๆ สามารถนำมาประกอบกันเป็นปริภูมิของสัญญาณใด ๆ ดังนั้นถ้าให้ U_i เป็นฟังก์ชันพื้นฐานที่เป็นออร์ทอนอโมลและ C_i เป็นสัมประสิทธิ์ของ U_i จะได้สัญญาณ v ใด ๆ นั้นจะสามารถประกอบขึ้นจากฟังก์ชันพื้นฐานที่ออร์ทอนอโมลที่มีค่าต่าง ๆ รวมกันดังสมการ

$$v = \sum_{i=-\infty}^{\infty} C_i U_i, v \in V \quad (2.53)$$

2.5.3 การวิเคราะห์แบบหลายระดับความละเอียดของมอร์เลต (Mollet's Multi Resolution analysis or Mollet'MRA)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติให้ V_j เป็นปริภูมิเวกเตอร์ที่มี j แสดงถึงระดับความละเอียดและจำนวนของฟังก์ชันพื้นฐาน พิจารณาปริภูมิเวกเตอร์ที่ระดับความละเอียด j จะมีฟังก์ชันพื้นฐานที่จะประกอบกันเป็นสัญญาณ การประมาณของ $f(t)$ เมื่อ $f(t) \in V_j$ ซึ่งเรียกฟังก์ชันพื้นฐานเหล่านี้ว่า ฟังก์ชันสเกลลิง (Scaling Function) ซึ่งจัดเป็นสัญญาณพื้นฐานย่อย ๆ ที่เกิดที่ตำแหน่งที่เวลาต่าง ๆ กันภายใน V_j และมีความถี่ต่างกันภายใน V_j เดียวกัน ดังนั้นถ้าสามารถกำหนดเซตของฟังก์ชันสเกลลิงในเทอมของการเลื่อนตำแหน่งเป็นจำนวนเต็มของฟังก์ชันสเกลลิง พื้นฐานที่ปริภูมิ V_0 ได้คือ $\phi_k(t) = \phi(t-k); k \in \mathbb{Z}$ และ $\phi \in L^2$ โดย \mathbb{Z} คือเซตของจำนวนเต็ม

โดยที่ k จะเป็นตำแหน่งการเกิดของ $\phi(t)$ แต่ละตัวและปริภูมิย่อยของ $L^2(\mathbb{R})$ (โดย \mathbb{R} เป็นเซตของจำนวนจริง) ถูกแผ่คลุม (Span) โดยฟังก์ชันสเกลลิงเหล่านี้ที่กำหนดโดย $V_0 = \text{span} \{ \phi_k(t) \}$ สำหรับทุกจำนวนเต็ม k จาก ∞ ถึง $-\infty$

ดังนั้นจะสามารถประมาณฟังก์ชันสัญญาณ $f(t)$ ภายในปริภูมิ V_0 ได้คือ

$$f(t) = \sum_k a_k \phi_k(t); f(t) \in V_0 \quad (2.54)$$

เนื่องจากความถี่จะเปลี่ยนแปลงขึ้นอยู่กับระดับความละเอียดของปริภูมิ j นั้น ๆ โดยอาศัยลักษณะคุณสมบัติที่ว่า $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}; j \in \mathbb{Z}$ จะได้ว่า

$$f(x) \in V_0 \Leftrightarrow f(2x) \in V_1 \text{ เมื่อ } j=0 \quad (2.55)$$

นั่นคือที่ระดับ V_1 มีความถี่เป็นสองเท่าของความถี่ที่ระดับ V_0 ลักษณะดังกล่าวคือการสเกลทางเวลาของสเกลลิงฟังก์ชัน ดังนั้นเซตของสเกลลิงฟังก์ชันที่เกิดจากการทำสเกลและเลื่อนตำแหน่งของฟังก์ชันสเกลลิงพื้นฐานที่ระดับ j และ $j+1$ ไต ๆ เขียนเป็นสมการได้คือ

$$\phi_{j,k}(t) = \text{const} * \phi(2^j t - k); \text{const} = \text{constant} \quad (2.56)$$

const เป็นการทำการนอร์มอลไลซ์ฟังก์ชันสเกลลิงที่เกิดจากการสเกล (Scaling) ที่เกิดจากคุณสมบัติออร์ทอگونัลซึ่งกันและกันของฟังก์ชันสเกลลิง ที่พิจารณาจาก

$$\langle \phi_{j,k}, \phi_{j,l} \rangle = \int \phi_{j,k}(t) \phi_{j,l}(t) dt = 0; k \neq l \quad (2.57)$$

จากสมการที่ (2.57) เมื่อ $k = l$ เป็นการหาขนาดหรือนอร์มอลไลซ์ของตัวเอง ถ้าขนาดเท่ากับหนึ่งจะได้ว่า

$$\int \{ \phi_{j,k}(t) \}^2 dt = 1 \quad (2.58)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นำสมการ (2.56) มาแทนใน (2.58) เพื่อหา Constant จะได้ว่า

$$\int \text{const}^2 \phi^2(2^j t - k) dt = 1$$

$$2^{-j} \text{const}^2 \int \phi^2(2^j t - k) d(2^j t) = 1$$

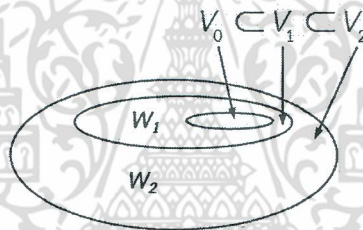
จากสมการที่ (2.58) จะได้ $2^{-j} \text{const}^2 = 1$

$$\text{const} = 2^{j/2}$$

จากสมการที่ (2.56) สามารถเขียนได้ใหม่คือ

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (2.59)$$

จากคุณสมบัติที่ว่า $V_{-\infty} \subset V_{-1} \subset V_0 \subset V_1 \subset V_{\infty}$ สามารถแสดงกลุ่มของปริภูมิเวกเตอร์ได้ ดังรูปที่ 2.27



รูปที่ 2.27 กลุ่มของปริภูมิเวกเตอร์

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

$$\phi(t) = \sum_k h(k) \sqrt{2} \phi(2t - k) \quad (2.60)$$

$h(k)$ เป็นเซตสัมประสิทธิ์สเกลลิงฟังก์ชัน หาได้จาก $\langle \phi(t), \sqrt{2} \phi(2t - k) \rangle$ และจะเรียกสมการที่ (2.60) ว่า สมการรีเคอร์ซีฟ (Recursive Equation) หรือบางครั้งเรียกว่า สมการรีฟายเมนต์ (Refinement Equation) หรือสมการไดเลชัน (Dilation Equation)

จากรูปที่ (2.27) ภายในปริภูมิ V_1 จะประกอบด้วยปริภูมิ V_0 และปริภูมิที่เป็นส่วนประกอบออร์ทอกอนัล (Orthogonal Complement) ของ V_0 คือ W_0 นั่นคือ

$$V_1 = V_0 \oplus W_0 \quad (2.61)$$

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

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$V_2 = V_1 \oplus W_1 \quad (2.62)$$

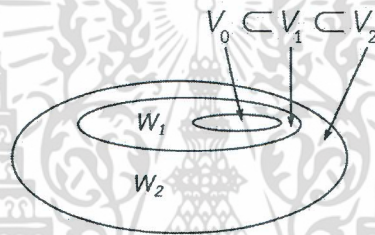
และเมื่อนำ สมการที่ (2.61) มาแทนใน สมการที่ (2.62) จะได้

$$V_2 = V_0 \oplus W_0 \oplus W_1 \quad (2.63)$$

จากคุณสมบัติ $\text{close}_{L^2}(U_{V_j}) = L^2(IR); j \in ZZ$ จะได้ว่า

$$L^2 = V_0 \oplus W_0 \oplus W_1 \oplus \dots \quad (2.64)$$

นั่นคือภายในปริภูมิของ V_{j+1} จะประกอบด้วยปริภูมิ V_j และปริภูมิ W_j โดยที่สมาชิกในปริภูมิ W_j ซึ่งจะออร์โทกอนัลกับสเกลลิงฟังก์ชัน ในปริภูมิ V_j เรียกว่าเวฟเลตฟังก์ชัน (Wavelet Function : $\psi(t)$) ดังรูปที่ 2.28



$$W_2 \perp W_1 \perp W_0 \perp V_0$$

รูปที่ 2.28 ปริภูมิเริ่มต้นที่ถูกแผ่คลุม

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

$$L^2 = V_{10} \oplus W_{10} \oplus W_{11} \oplus W_{12} \oplus \dots \quad (2.65)$$

หรือที่ความละเอียดต่ำ $j = -5$

$$L^2 = V_{-5} \oplus W_{-5} \oplus W_{-4} \oplus W_{-3} \oplus \dots \quad (2.66)$$

และถ้า $j = \infty$ แล้วเทียบกับ สมการที่ (2.43) จะได้

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

เนื่องจาก $f(t) \in L^2(IR)$ จะได้ว่า $f(t)$ จะประกอบด้วย เวฟเลตฟังก์ชันในปริภูมิ W_j

นั่นคือ สามารถประมาณ $f(t)$ ได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$f(t) = \sum_{j,k} a_{j,k} \psi_{j,k}(t) \quad (2.67)$$

โดยที่ $a_{j,k}$ เป็น สัมประสิทธิ์ ที่เรียกว่า ดีสครีตเวฟเลตทรานส์ฟอร์ม (DWT) ของ $f(t)$ ที่สามารถคำนวณหาได้จากการใช้ผลคูณภายใน คือ

$$a_{j,k} = \langle \psi_{j,k}(t), f(t) \rangle \quad (2.68)$$

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

เมื่อพิจารณาจากสมการ ปริภูมิ W_0 จะมีคุณสมบัติต่าง ๆ เหมือนกับปริภูมิ V_0 เพียงแต่จะเป็นส่วนประกอบย่อยที่ประกอบอยู่ในปริภูมิ V_1 ดังนั้น ที่ระดับความละเอียดใด ๆ (j) และพิจารณาคูณสมบัติเหมือนสเกลลิงฟังก์ชัน จะได้สมการที่คล้ายกับสมการที่ (2.59) คือ

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (2.69)$$

เนื่องจากเวฟเลตเหล่านี้จะอยู่ในปริภูมิที่แผ่คลุมโดยสเกลลิงฟังก์ชันถัดไป นั่นคือ $W_0 \subset V_1$ และในทำนองเดียวกันกับสมการที่ (2.60) และสมการที่ (2.69) สามารถเขียนเวฟเลตฟังก์ชันเป็นลักษณะของผลบวกเชิงเส้นของสเกลลิงฟังก์ชัน (Scaling Function) ของปริภูมิ V_1 คือ

$$\psi(t) = \sum_k h_1(k) \sqrt{2} \phi(2t - k), k \in \mathbb{Z} \quad (2.70)$$

จากสมการที่ (2.70) จะเป็นเซตของสัมประสิทธิ์ เวฟเลตฟังก์ชัน โดยมีความสัมพันธ์กับเซตสัมประสิทธิ์ สเกลลิงฟังก์ชัน ($h(k)$) จากคุณสมบัติที่ว่า $V_i \oplus W_j = V_{j+1}; j \in \mathbb{Z}$ จะได้ว่า

$$\langle \phi_{j,k}(t), \psi_{j,l}(t) \rangle = \int \phi_{j,k}^*(t) \psi_{j,l}(t) dt = 0 \quad (2.71)$$

เมื่อ $j, k, l \in \mathbb{Z}$ จากสมการที่ (2.71) จะได้ความสัมพันธ์ของสัมประสิทธิ์เวฟเลตที่ออร์โทกอนัลกับสัมประสิทธิ์สเกลลิง ดังนี้

$$h_1(k) = (-1)^k h(N-1-k) \quad (2.72)$$

เมื่อ N เป็นจำนวนข้อมูลของเซตสัมประสิทธิ์สเกลลิง $h(k)$

จากสมการ (2.64) เราสามารถสร้างฟังก์ชันใด ๆ จากเซตของฟังก์ชัน $\phi_k(t)$ และ $\psi_{j,k}(t)$ แผ่คลุม $L^2(\mathbb{R})$ จะได้

$$g(t) = \sum_{k=-\infty}^{\infty} c(k)\phi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k)\psi_{j,k}(t) \quad (2.73)$$

จากสมการที่ (2.73) ในการกระจายอนุกรม $g(t)$ เทอมแรกของสมการจะเป็นผลรวมแสดงถึงความละเอียดต่ำ หรือการประมาณที่หยาบ (Coarse Approximation) ของ $g(t)$ สำหรับการเพิ่มของแต่ละ j ในเทอมที่สองจะแสดงถึงความละเอียดที่สูงขึ้นและจะทำให้รายละเอียดของข้อมูล (Detail) เพิ่มขึ้น เหมือนกับอนุกรมฟูรีเยร์ โดยที่สมการที่ (2.73) จะเลือกปริภูมิสเกลลิงฟังก์ชันที่สเกล j ใด ๆ ก็ได้คือ

$$g(t) = \sum_k c_{j_0}(k)\phi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{\infty} d_j(k)\psi_{j,k}(t) \quad (2.74)$$

จากสมการที่ (2.74) สัมประสิทธิ์ ทั้งสองเทอมสามารถคำนวณจากผลคูณภายใน คือ

$$C_j(k) = \langle g(t), \phi_{j,k}(t) \rangle \quad (2.75)$$

$$d_j(k) = \langle g(t), \psi_{j,k}(t) \rangle \quad (2.76)$$

สัมประสิทธิ์ในการกระจายของเวฟเลต จะถูกเรียกว่า Discrete Wavelet Transform ของ $g(t)$

2.5.4 การแปลงเวฟเลตโดยใช้หลักการของการวิเคราะห์ฟิลเตอร์แบงก์ (Filter bank) การวิเคราะห์จากระดับที่มีความละเอียดสูงไปยังระดับที่มีความละเอียดต่ำจากสมการรีเคอร์ซีฟ

$$\phi(t) = \sum_n h(n)\sqrt{2} \phi(2t-n) \quad (2.77)$$

ถ้าทำการสเกลและเลื่อนตำแหน่ง ของตัวแปรเวลาจะได้ว่า

$$\begin{aligned} \phi(2^j t - k) &= \sum_n h(n)\sqrt{2} \phi(2^{j+1} t - k - n) \\ \phi(2^j t - k) &= \sum_n h(n)\sqrt{2} \phi(2^{j+1} t - 2k - n) \end{aligned} \quad (2.78)$$

เปลี่ยนตัวแปรโดยให้ $m = 2k + n$ จะได้

$$\phi(2^j t - k) = \sum h(m-2k)\sqrt{2} \phi(2^{j+1} t - m) \quad (2.79)$$

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

$$V_j = \text{span}\{2^{j/2} \phi(2^j t - k)\} \quad (2.80)$$

$$\text{ดังนั้นถ้า } f(t) \in V_{j+1} \text{ แล้ว } f(t) = \sum_k c_{j+1}(k) 2^{(j+1)/2} \phi(2^{j+1} t - k) \quad (2.81)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการ (2.81) $f(t)$ เกิดจากสเกลลิงฟังก์ชันอย่างเดียวไม่มีเวฟเลต ดังนั้นการสเกลที่ความละเอียดต่ำ ๆ เวฟเลตจำเป็นสำหรับ

$$f(t) = \sum_k c_j(k) 2^{j/2} \phi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k) \quad (2.82)$$

ถ้า $\phi_{j,k}(t)$ และ $\psi_{j,k}(t)$ เป็นออร์ทอกอนัลกัน จะได้ สัมประสิทธิ์ สเกลลิงที่ระดับ $j(c_j(k))$ จะหาได้จากผลคูณภายใน คือ

$$c_j(k) = \langle f(t), \phi_{j,k}(t) \rangle \quad (2.83)$$

เมื่อนำ (2.79) แทนใน (2.83) จะได้

$$c_j(k) = \langle f(t), 2^{j/2} \phi(2^j t - k) \rangle$$

$$c_j(k) = \langle f(t), 2^{j/2} \sum_m h(m-2k) \sqrt{2} \phi(2^{j+1} t - m) \rangle \quad (2.84)$$

$$c_j(k) = \sum_m h(m-2k) \langle f(t), 2^{j+1/2} \phi(2^{j+1} t - m) \rangle$$

เมื่อนำ (2.83) เทียบกับ (2.84) จะได้

$$c_j(k) = \sum_m h(m-2k) c_{j+1}(m) \quad (2.85)$$

และในทำนองเดียวกันจะได้

$$d_j(k) = \sum_m h(m-2k) c_{j+1}(m) \quad (2.86)$$

2.5.5 การฟิลเตอร์และการลดอัตราการซิกค่าตัวอย่างลง (Down sampling)

ในรูปแบบการประมวลผลสัญญาณดิจิทัล, การฟิลเตอร์สัญญาณที่เป็นลำดับ จะใช้การทำคอนโวลูชัน (Convolution) ระหว่างลำดับ (Sequence) กับเซตของจำนวน สัมประสิทธิ์ ตัวฟิลเตอร์

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

$x(n)$ = Input sequence

$h(n)$ = Filter coefficient

$y(n)$ = Output sequence

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือเขียนสมการ (2.87) ให้อยู่ในรูปของการคอนโวลูชันจะได้

$$Y(n) = h(n) \oplus x(n) \quad (2.87)$$

จากสมการ (2.85) เมื่อเขียนให้อยู่ในรูปของการคอนโวลูชันจะได้

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

หรือ

$$c_j(k) = h(-n) \oplus c_{j-1}(n) \quad \text{เมื่อ } n = 2k \quad (2.88)$$

และทำนองเดียวกันจะได้

$$d_j(k) = h_1(-n) \oplus c_{j-1}(n) \quad \text{เมื่อ } n = 2k \quad (2.89)$$

จาก (2.88) และ (2.89) สัมประสิทธิ์ $C_j(k)$ และ $d_j(k)$ จะได้ สัมประสิทธิ์ $C_{j+1}(2k)$ เมื่อพิจารณาเทียบกับหลักการฟิลเตอร์แบงก์กับการฟิลเตอร์สัญญาณดิจิตอลแล้ว เซต สัมประสิทธิ์ของ $h(-n)$ และ $h_1(-n)$ เทียบได้กับ สัมประสิทธิ์ของตัวฟิลเตอร์ทางดิจิตอล และตัวชี้ (index) ที่แสดงถึงลำดับของสัญญาณของ $C_j(k)$ และ $d_j(k)$ ได้จากการคอนโวลูชันกับ $h(-n)$ หรือ $h_1(-n)$ แล้วซ้ค่าตัวอย่างลง 2 เท่า เรียกว่าการลดค่าจากค่าตัวอย่างหรือ ดีซิเมเตอร์ (Decimator) ดังรูปที่ 2.29



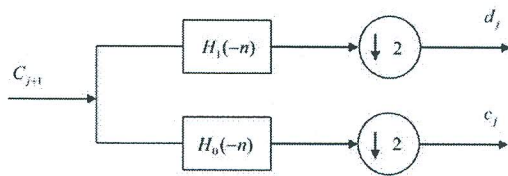
รูปที่ 2.29 สัญลักษณ์ Down Sampling

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

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

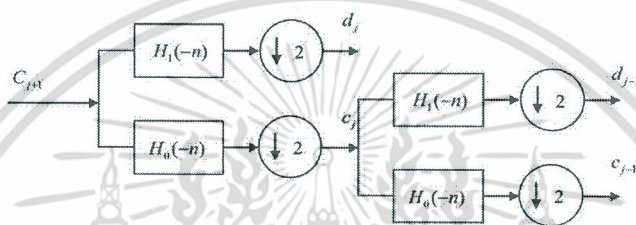
ถ้าจำนวนของตัว สัมประสิทธิ์ ฟิลเตอร์มีจำนวนจำกัดขนาด n แล้ว ตัวฟิลเตอร์จะเรียกว่า การตอบสนองอิมพัลส์ที่จำกัด (Finite Impulse Response : FIR) และจะเรียกว่าการตอบสนองอิมพัลส์ไม่จำกัด (Infinite Impulse Response : IIR) ถ้าจำนวนตัว สัมประสิทธิ์ฟิลเตอร์มีจำนวนไม่จำกัด

ดังนั้นตัวกรองความถี่แบบดิจิตอลและการลดอัตราการซ้ค่าตัวอย่างจะนำมาใช้กับสมการ และตามหลักการของฟิลเตอร์แบงค์ ของสัญญาณดังรูปที่ 2.30



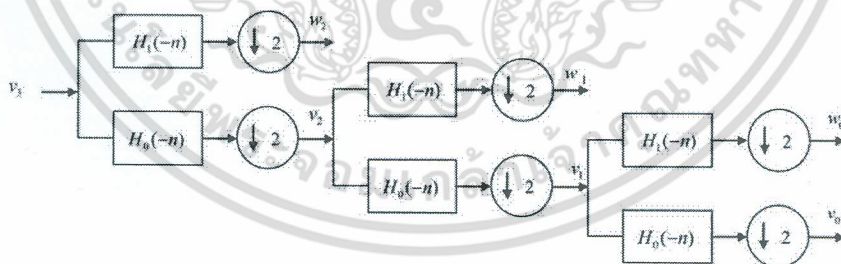
รูปที่ 2.30 ตัวกรองความถี่ตามหลักฟิลเตอร์แบงค์

จากรูปเป็นการกรองสัญญาณหรือการแยกสัญญาณ (Decomposition) C_{j+1} ที่ระดับ $j + 1$ ไปสู่สัมประสิทธิ์ C_j และ d_j ที่ระดับ j ดังนั้นถ้านำฟิลเตอร์แบงค์ดังกล่าวมาแยกสัญญาณ (สัมประสิทธิ์สเกลลิง) อีกเพื่อให้ได้สเกลที่หายากอีก เช่นที่สเกล $j-1$ ดังรูปที่ 2.31



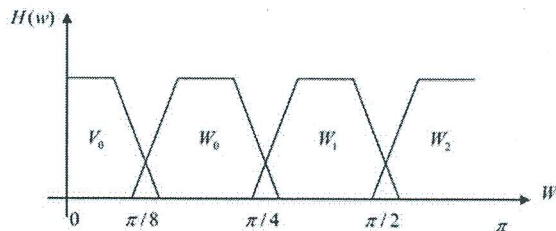
รูปที่ 2.31 ฟิลเตอร์แบงค์ ที่ระดับ C_{j-1}

ดังนั้นถ้าต้องการจะหา j_0 จากระดับสูง โดยในทางปฏิบัติสัญญาณใด ๆ ที่มีช่วงของสัญญาณจำกัด จะมีระดับสูงสุดที่ $j=J$ ด้วยการเริ่มต้นที่การอธิบายสัญญาณที่ระดับความละเอียดสูง (High Resolution) ในเทอมของ สัมประสิทธิ์ c_j และจะวิเคราะห์ลงไปสู่ระดับความละเอียดต่ำ (Low Resolution) $j = j_0$ ดังนั้นจะต้องใช้ขั้นตอนการแยกสัญญาณ $J - j_0$ ขั้นตอนเป็นดังรูปที่ 2.32



รูปที่ 2.32 การแยกสัญญาณที่ V_3

เนื่องจากในปริภูมิ V_3 จะถูกแตกกระจายสัญญาณไปสู่ระดับ $j = 2$ ทั้งในปริภูมิของสเกลลิง V_2 และปริภูมิของเวฟเลต W_2 โดยจากรูป $h_1(-n)$ จะเป็นตัวกรองแยกสัญญาณความถี่สูงหรือเทียบได้กับระดับความละเอียดสูง ส่วนตัวฟิลเตอร์ของ สัมประสิทธิ์ $h_0(-n)$ จะกรองแยกสัญญาณความถี่เมื่อแสดงในลักษณะการวิเคราะห์ทางด้านความถี่ จะแสดงได้ดังรูปที่ 2.33



รูปที่ 2.33 การแยกสัญญาณทางด้านความถี่

ดังนั้นสำหรับ $f(t) \in V_j$ และจากสมการจะได้ว่า

$$f(t) = \sum_k c_j(k) \phi_{j,k}(t) \quad (2.90)$$

$$f(t) = \sum_k c_{j-1}(k) \phi_{j-1,k}(t) + \sum_k d_{j-1}(k) \psi_{j-1,k}(t) \quad (2.91)$$

$$f(t) = \sum_k c_{j-2}(k) \phi_{j-2,k}(t) \quad (2.92)$$

2.5.6 การสังเคราะห์ที่ระดับความละเอียดต่ำไปยังความละเอียดสูง
ในขั้นตอนการสร้างกลับคืนจากการที่เราแตกกระจายออกไป จะสามารถสร้างได้จากการรวม (Combination) ของฟังก์ชันสเกลลิ่งและ สัมประสิทธิ์เวฟเลตที่ความละเอียดต่ำโดยการพิจารณา สัญญาณในปริภูมิของฟังก์ชันระดับ $j+1$ คือ $f(t) \in V_{j+1}$ ซึ่งค่าของสัญญาณสามารถเขียนในเทอมของสเกลลิ่ง f_n ได้ดังนี้

$$f(t) = \sum_k c_{j+1}(k) 2^{j+1/2} \phi(2^{j+1}t - k) \quad (2.93)$$

หรือในเทอมที่ประกอบด้วยฟังก์ชันเวฟเลต ดังนี้

$$f(t) = \sum_k c_j(k) 2^{j/2} \phi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k) \quad (2.94)$$

โดยแทนสมการ (2.70) และ (2.78) ลงใน (2.94)

$$f(t) = \sum_k c_j(k) \sum_n h_0(n) 2^{j+1/2} \phi(2^{j+1}t - 2k - n) + \sum_k d_j(k) \sum_n h_1(n) 2^{j+1/2} \psi(2^{j+1}t - 2k - n) \quad (2.95)$$

จาก (2.93) สามารถหาค่า $c_{j+1}(k)$ จากการคูณด้วย $2^{j+1/2} \phi(2^{j+1}t - k)$ จะได้

$$c_{j+1}(k) = \langle f(t), \phi_{j+1,k}(t) \rangle \quad (2.96)$$

จาก (2.95) แทนตัวแปร k ด้วย m จะได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$f(t) = \sum_m c_j(m) \sum_n h_0(n) 2^{j+1/2} \phi(2^{j+1}t - 2m - n) + \sum_m d_j(m) \sum_n h_1(n) 2^{j+1/2} \phi(2^{j+1}t - 2m - n) \quad (2.97)$$

จาก(2.97) สามารถสร้างกลับคืน สัมประสิทธิ์ $c_{j+1}(k)$ ได้เมื่อกำหนดให้ $2m + n = k$ นั่นคือ

$$f(t) = \sum_m c_j(m) \sum_{k=0} h(k-2m) 2^{j+1/2} \phi(2^{j+1}t - k) + \sum_m d_j(m) \sum_k h_1(k-2m) 2^{j+1/2} \phi(2^{j+1}t - k) \quad (2.98)$$

เมื่อพิจารณาหาผลคูณภายในของสมการ (2.98) กับ $2^{j+1/2} \phi(2^{j+1}t - k)$ จะได้

$$[f(t), \phi_{j+1}, k(t)] = \sum_m c_j(m) h_0(k-2m) + \sum_m d_j(m) h_1(k-2m) \quad (2.99)$$

เมื่อเทียบ (2.96) กับ (2.99) จะได้ว่า

$$c_{j+1}(k) = \sum_m c_j(m) h_0(k-2m) + \sum_m d_j(m) h_1(k-2m) \quad (2.100)$$

2.5.7 การฟิลเตอร์และการเพิ่มอัตราการซิกค่าตัวอย่าง (Up sampling)

ในการสังเคราะห์สัญญาณกลับ จะตรงกันข้ามกับการวิเคราะห์สัญญาณด้วยการแยกช่วงความถี่ออก จากกันทีละสองช่วง โดยทำการเพิ่มอัตราการซิกค่าตัวอย่างให้มีค่าเป็น 2 เท่าจากเดิมก่อน ซึ่งจะทำให้ การแทรกค่าศูนย์ระหว่างเทอมของสัญญาณ ดังสมการ (2.101) และรูปที่ 2.34

$$Y(2n) = X(n); Y(2n+1) = 0 \quad (2.101)$$



รูปที่ 2.34 สัญลักษณ์การเพิ่มอัตราการซิกค่าเป็น 2 เท่า

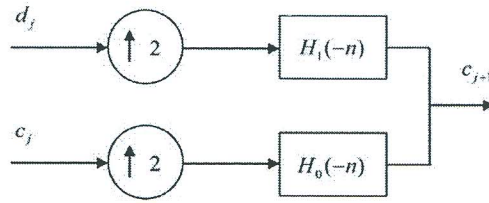
จาก (2.101) พิจารณาผลรวมของแต่ละเทอมในรูปแบบการคอนโวลูชันร่วมกับ (2.77) และ (2.78) จะเขียนได้ใหม่เป็น

$$c_{j+1}(k) = c_j(k) \otimes h_0(k) \Big|_{k=2m} + d_j(k) \otimes h_1(k) \Big|_{k=2m} \quad (2.102)$$

โดยใช้ตัวแปร n ตัวแปรใหม่แทนตัวแปร k ในสมการ (2.102) จะได้

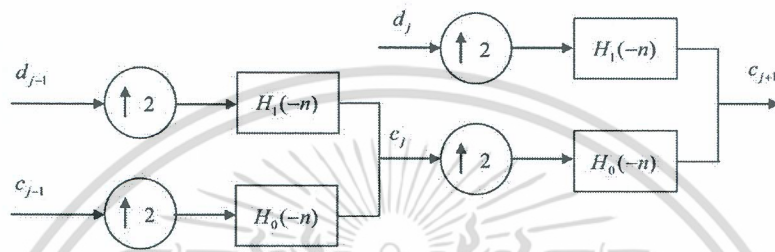
$$c_{j+1}(n) = c_j(n) \otimes h_0(n) \Big|_{n=2m} + d_j(n) \otimes h_1(n) \Big|_{n=2m} \quad (2.103)$$

จากสมการที่ (2.103) สามารถนำมาสร้างในลักษณะฟิลเตอร์แบงค์ จะได้ดังรูปที่ 2.35



รูปที่ 2.35 ไตอะแกรมของการซีกค่าโดยใช้หลักการของฟิลเตอร์แบงค์

ในทำนองเดียวกัน สามารถทำการสังเคราะห์ สัญญาณจากระดับ $j-1$ ไปสู่ระดับ j ได้ดังรูปที่ 2.36

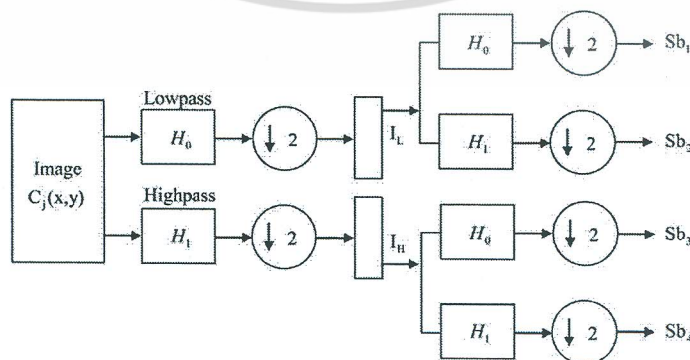


รูปที่ 2.36 การสังเคราะห์สัญญาณจากระดับ $j-1$ ไปสู่ j

รูปที่กล่าวมาทั้งหมดตั้งแต่ Decimator จนถึงรูปข้างบนนี้จะเรียกว่าการแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform: DWT) และในการแปลงกลับเวฟเลตแบบไม่ต่อเนื่องจะมีการทำงานตรงกันข้ามกับรูปที่ (2.35) และ (2.36) จะเรียกว่า การแปลงกลับเวฟเลตแบบไม่ต่อเนื่อง (Inverse Discrete Wavelet Transform : IDWT) จะเห็นได้ว่า $\phi(t)$ และ $\psi(t)$ จะเป็นฟังก์ชันที่มีช่วงเวลาจำกัด โดยจะมีค่าเท่ากับ 0 ในกรณีที่ไม่อยู่ในช่วงของ $0 \leq t \leq N-1$ ดังนั้น $h_0(-n)$ และ $h_1(-n)$ สมการจะมีลักษณะเป็นฟิลเตอร์ที่มีการตอบสนองอิมพัลส์ที่จำกัด (FIR Filter) ซึ่งจะเป็นผลให้เฟสของสัญญาณที่ผ่านการกรองความถี่มีลักษณะเชิงเส้นด้วย

2.5.8 แปลงเวฟเลตกับข้อมูลภาพ

ข้อมูลภาพที่มีลักษณะเป็นข้อมูล 2 มิติ ดังนั้นเมื่อทำการแปลงเวฟเลตกับภาพจะเป็นการทำการแปลงเวฟเลตในทางแนวแกน x และแกน y สลับกันในแต่ละครั้ง สามารถเขียนเป็นแผนผังได้ดังรูปที่ 2.37



รูปที่ 2.37 แผนผังของการแปลงเวฟเลต

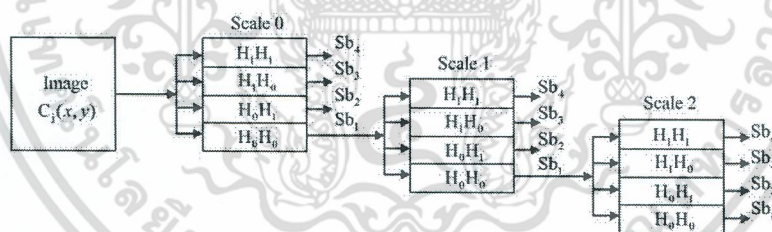
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะของการแบ่งแบนด์ย่อยของภาพจะได้ดังรูปที่ 2.38

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

รูปที่ 2.38 ลักษณะของการแบ่งแบนด์ย่อยของภาพ

จากรูปที่ 2.38 เป็นการแปลงเวฟเลต 2 มิติ จะเป็นการแปลงในแนวแกน x ตามจำนวนแถวของจุดภาพ ซึ่งทำให้ภาพถูกแบ่งออกเป็น 2 แบนด์ย่อย จากนั้นทำการแปลงเวฟเลตในแนวแกน y ตามจำนวนคอลัมน์ของจุดภาพ ซึ่งทำให้ภาพมีลักษณะเป็น 2 มิติ ที่ผ่านการแปลงเวฟเลตหนึ่งครั้ง ต้องถูกแบ่งออกเป็น 4 ส่วน (4 แบนด์ย่อย) แบนด์ย่อยแรกจะเป็นส่วนภาพที่ผ่านการกรองความถี่ต่ำผ่านทั้ง 2 แนวแกน แบนด์ย่อยที่ 2 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านในแนวนอนและความถี่ต่ำผ่านในแนวตั้ง แบนด์ย่อยที่ 3 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านในแนวตั้งและความถี่ต่ำผ่านในแนวนอน แบนด์ย่อยที่ 4 จะเป็นส่วนภาพที่ผ่านการกรองความถี่สูงผ่านทั้ง 2 แนวแกน ส่วนที่แสดงรายละเอียดของภาพจะเป็นส่วนข้อมูลที่อยู่ในแบนด์ย่อยที่ 2, 3, 4 นอกจากนี้เมื่อนำเอาแบนด์ย่อยที่ 1 ไปทำการแปลงเวฟเลตต่อไปเรื่อย ๆ ก็จะสามารถทำการแยกแบนด์ย่อยของภาพออกไปได้มากขึ้นดังรูปที่ 2.39 และรูปที่ 2.40



รูปที่ 2.39 การแปลงเวฟเลตหลาย ๆ ระดับ

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

รูปที่ 2.40 ลักษณะการแบ่งแบนด์ย่อยหลาย ๆ ระดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

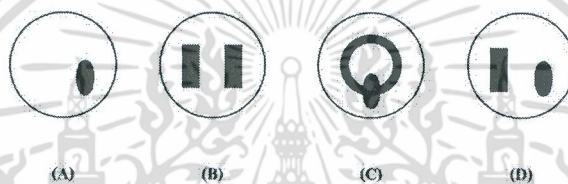
จากรูปเป็นการนำแบนด์ย่อย มาแยกออกเป็นแบนด์ย่อย ย่อยในที่นี่แบ่งเฉพาะ แบนด์ย่อยที่เป็น H_0H_0 เท่านั้น

2.6 ทฤษฎีฟัซซีซีมีน (Fuzzy C-Means)

ในการแบ่งกลุ่มข้อมูลของวิชาการจดจำรูปแบบ (Pattern Recognition) [21] จะหมายถึงการหา ลักษณะการรวมตัวของข้อมูลว่ามีโครงสร้างอย่างไร สำหรับตัวอย่างการแบ่งข้อมูลสามารถแสดงได้ดัง รูปที่ 2.41

2.6.1 อัลกอริทึมของฟัซซีซีมีน

อัลกอริทึมของฟัซซีซีมีนแบบกำหนดค่าเริ่มต้นด้วยค่าสมาชิกภาพ (Membership values: u) มีความ เป็นไปได้ของโครงสร้างข้อมูลดังแสดงในรูปที่ 2.41



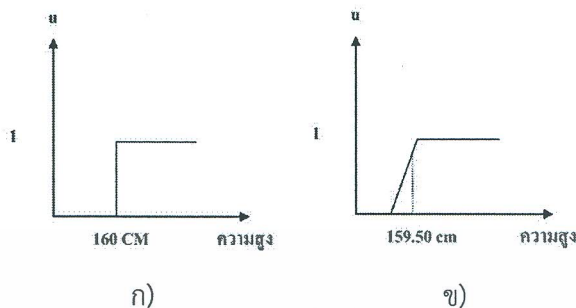
รูปที่ 2.41 ความเป็นไปได้ของโครงสร้างข้อมูล

จากรูปได้แสดงถึงโครงสร้างข้อมูล 4 แบบ ซึ่งในแต่ละแบบเรายังสามารถแบ่งข้อมูลออกเป็น อีก 2 กลุ่ม โดยที่นิยามคำว่ากลุ่มในแต่ละแบบยังมีลักษณะที่แตกต่างกันอีก เช่นในรูป 2.41 (A) จะเป็นกลุ่มที่ต้องอาศัยระยะทางระหว่างจุดข้อมูล (Data Point) คือจุดข้อมูลจะมีระยะทางที่สั้นเมื่ออยู่ในกลุ่มเดียวกัน ส่วนกลุ่มในรูป 2.40 (B) กลุ่มของข้อมูลน่าจะแบ่งตามความต่อเนื่องกันของจุดข้อมูล นั่นก็คือจุดข้อมูลที่อยู่ติดกันจะถือว่าอยู่ในกลุ่มเดียวกัน และกลุ่มในรูป 2.41 (C)-(D) จะถูกนิยามของการแบ่งกลุ่มโดยการรวมคุณสมบัติของระยะทางและการเชื่อมต่อกันของจุดข้อมูล ซึ่งในวิชาการรู้จำรูปแบบ คำว่ากลุ่มของข้อมูลจะถูกเรียกว่า คลัสเตอร์ (Cluster)

ในความเป็นจริงจุดข้อมูลจะไม่มีกระจายในลักษณะที่มีโครงสร้างที่เป็นระเบียบแบบนี้ซึ่ง ไม่มีกฎเกณฑ์ตายตัวในการแบ่งกลุ่มข้อมูล จากความคลุมเคลือของข้อมูลอาจทำให้จุดข้อมูลอยู่ในกลุ่มได้มากกว่า 1 กลุ่ม โดยมีค่าสมาชิกภาพ (Membership Values: u) ของข้อมูลเป็นตัวชี้ว่าจุดข้อมูลนั้นจะมีโอกาสอยู่ในกลุ่มใดบ้างซึ่งจะกล่าวในรายละเอียดต่อไป

การใช้ระบบฟัซซี (Fuzzy) มาทำการแบ่งกลุ่มข้อมูลก็เนื่องจากระบบฟัซซีมีความสามารถในการจัดการกับข้อมูลที่มีความคลุมเครือของการแบ่งกลุ่มได้มากกว่าระบบไม่ใช่ฟัซซี (Non Fuzzy) เพราะในระบบดังกล่าวจะมีค่าสมาชิกภาพ (u) ที่เป็นตัวชี้ว่าข้อมูลจะอยู่ในกลุ่มใดจะมีค่า 0 หรือ 1 (ใช่หรือไม่ใช่) เท่านั้น ยกตัวอย่าง เช่น ถ้ามีการตั้งกฎเกณฑ์ว่า คนที่มีความสูง 160 cm ขึ้นไปจะถือว่า

อยู่ในกลุ่มคนสูง ($u=1$) และคนที่มีความสูงต่ำกว่า 160 cm ลงมาจะถือว่าเป็นกลุ่มคนเตี้ย ($u=0$) ดังที่แสดงได้ดังในรูปที่ 2.41(A)



(ก) การแบ่งกลุ่มข้อมูลโดยระบบ non Fuzzy (ข) การแบ่งกลุ่มข้อมูลโดยระบบ Fuzzy

รูปที่ 2.42 การนำฟัซซีมาทำการแบ่งกลุ่มข้อมูล

จากรูปที่ 2.42 (ก) จะเห็นว่ามีความสูงค่า u อยู่สองค่า คือ 0 กับ 1 เท่านั้นซึ่งที่เป็นมุมมองในระบบที่ไม่ใช่ฟัซซี แต่ในระบบของฟัซซีจะมีความสูงค่า u อยู่ระหว่าง 0 ถึง 1 $[0,1]$ ซึ่งจะทำให้การแบ่งกลุ่มข้อมูลมีความยืดหยุ่นกว่า จากตัวอย่างเดิมถ้าคนที่สูง 159.5 cm จะเห็นว่ามีความสูงพอๆกับคนที่สูง 160 cm ซึ่งในระบบที่ไม่ใช่ฟัซซี จะถือว่าเป็นคนเตี้ย แต่ในระบบฟัซซี จะถือว่าคนนี้มีค่าระดับความเป็นสมาชิกภาพ (u) เกือบเท่ากับ 1 (เป็นคนเกือบสูง) ซึ่งจะเห็นว่ามันสามารถจัดการแบ่งกลุ่มข้อมูลได้ละเอียดกว่า ดังในรูปที่ 2.42 (ข)

จากทั้งระบบที่เป็นฟัซซีและไม่เป็นฟัซซีที่นำมาแบ่งกลุ่มข้อมูลนั้นจะมีวิธีที่นำระบบทั้งสองมาใช้กันก็คืออัลกอริทึมของ Fuzzy C-Means และ Hard C-Means ตามลำดับ ซึ่งอัลกอริทึมทั้งสองนี้จะเป็นวิธีการแบ่งกลุ่มข้อมูล โดยอาศัยหลักการวัดระยะทางสั้นที่สุดระหว่างข้อมูล (Data) และจุดศูนย์กลางของข้อมูล (Cluster Center) เป็นตัวกำหนดว่าข้อมูลใดควรที่จะอยู่ในกลุ่มไหน ซึ่งจะได้กล่าวในหัวข้อต่อไป

2.6.2 การแบ่งกลุ่มข้อมูลโดยการวัดระยะทาง

ในการแบ่งกลุ่มข้อมูลโดยวิธีของ Hard และ Fuzzy C-Means นั้นสิ่งที่เป็นหัวใจในการกำหนดว่าข้อมูลจะอยู่ในกลุ่มใดนั้นก็คือ การวัดระยะทางที่สั้นที่สุดระหว่าง ข้อมูลถึงจุดศูนย์กลางของกลุ่มข้อมูล โดยที่ระยะทางดังกล่าวจะเรียกว่า “ระยะทางยูคลีเดียน” (Euclidean Distance) ส่วนสมการที่ใช้หาระยะทางยูคลีเดียนสามารถกำหนดได้ดังสมการที่ 2.104

$$d_{ik}^2 = \|x_k - v_i\|^2 \quad (2.104)$$

โดยที่ d_{ik} เป็นระยะทางยูคลีเดียนระหว่างข้อมูลและจุดศูนย์กลางของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

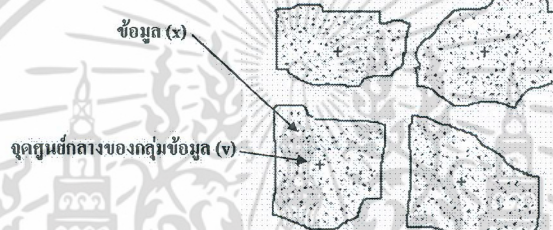
x_k เป็นข้อมูลตัวที่ k

v_i เป็นจุดศูนย์กลางของข้อมูลตัวที่ i

$1 \leq k \leq n$ เมื่อ n คือจำนวนของข้อมูลทั้งหมด

$1 \leq i \leq c$ เมื่อ c คือจำนวนกลุ่มที่ต้องการแบ่ง

จากสมการที่ 2.104 การกำหนดว่าข้อมูลตัวใดควรอยู่ในกลุ่มไหนนั้น จะต้องมีการคำนวณระยะทาง จากข้อมูล (x_k) ทุกตัวไปยังจุดศูนย์กลางของกลุ่มข้อมูล (v_i) ทุกๆกลุ่มดังแสดงในรูปที่ 2.43 ถ้าระยะทางจาก x_k ไปถึง v_i กลุ่มใดสั้นที่สุดเมื่อเทียบกับ v_i กลุ่มอื่นก็แสดงว่าข้อมูลตัวดังกล่าวจะเป็นสมาชิกของกลุ่มนั้น



รูปที่ 2.43 การกำหนดข้อมูลให้อยู่ในกลุ่มข้อมูลแต่ละกลุ่ม

ขั้นตอนอัลกอริทึมของ Fuzzy C-Means (FCM)

ขั้นที่ 1 กำหนดค่าเริ่มต้นของ $u^{(0)}$ และจำนวนกลุ่ม (c) ที่ต้องการจะแบ่ง โดยที่ $2 \leq c \leq n$

ขั้นที่ 2 คำนวณจุดศูนย์กลางข้อมูล (v_i) โดยใช้สมการ 2.105

$$v_i = \frac{\sum_{k=1}^n x_k (u_{ik})^m}{\sum_{k=1}^n (u_{ik})^m}; 1 \leq i \leq c \quad (2.105)$$

เมื่อ m เป็นค่า Weight Constant ที่กำหนดขึ้นและอยู่ในช่วงเปิด $(1, \alpha)$

ขั้นที่ 3 คำนวณหาค่าสมาชิกในลำดับต่อไป ($u^{(\ell+1)}$) ดังสมการที่ 2.106

$$u_{ik}^{(\ell+1)} = \frac{\left[\frac{1}{\|x_k - v_i\|_G^2} \right]^{\frac{1}{(m-1)}}}{\sum_{j=1}^c \left[\frac{1}{\|x_k - v_j\|_G^2} \right]^{\frac{1}{(m-1)}}}; 1 \leq i \leq c, 1 \leq k \leq n \quad (2.106)$$

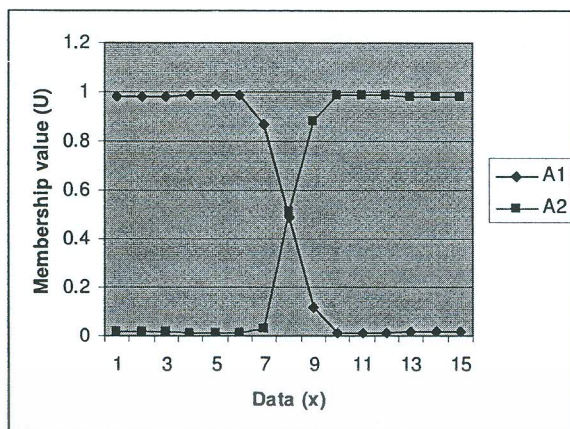
ขั้นที่ 4 คำนวณ $\Delta = \|u^{(\ell+1)} - u^{(\ell)}\|$ ถ้า $\Delta > \varepsilon$ (ε = ค่าผิดพลาดที่ยอมรับได้) ให้ $\ell = \ell + 1$ และกลับไปทำขั้นตอนที่ 2 แต่ถ้า $\Delta \leq \varepsilon$ ให้สิ้นสุดกระบวนการ

จากอัลกอริทึม FCM จะพบว่าในสมการที่ใช้หาค่าจุดศูนย์กลางของกลุ่มข้อมูล (ขั้นตอนที่ 2) จะมีพารามิเตอร์เพิ่มอีก 1 ตัวเมื่อเทียบกับอัลกอริทึมของ HCM คือ m (Weight Constant) ซึ่งมีค่าอยู่ในช่วง $1 < m \leq \infty$ โดยที่ค่า m จะเป็นตัวควบคุมความคลุมเครือ (Fuzziness) ของการประมวลผลทั้งหมด หรือก็คือทำให้ค่าสมาชิกภาพ (u) ของข้อมูล (ในขั้นตอนที่ 3) มีค่าอยู่ในช่วง 0 ถึง 1 นั้นเอง แต่ถ้า $m \rightarrow 1$ จะทำให้อัลกอริทึมของ FCM มีความใกล้เคียงกับอัลกอริทึมของ HCM มาก (คือเกือบทำให้ค่าสมาชิกมีค่า 1 หรือ 0 เท่านั้น) เมื่อ m มีค่าสูงขึ้นจะทำให้การกำหนดค่า u ในแต่ละกลุ่มมีความละเอียดมากขึ้นแต่ก็ต้องใช้เวลานานในการประมวลผล ซึ่งในปัจจุบันยังไม่มีทฤษฎีใดมาเป็นตัวกำหนดค่า m ที่เหมาะสม แต่อย่างไรก็ตามระบบจะลู่เข้าสู่คำตอบอย่างแน่นอน เมื่อ $1 < m \leq \infty$

ถ้าเราใช้อัลกอริทึมของ FCM โดยทดลองแทนค่า $c=2$, $m=2.02$, $\varepsilon=0.01$ ลงในสมการที่ 2.106 จะได้ค่า u ในลำดับสุดท้าย เมื่อ $\ell=5$ (คือจำนวนรอบ) นั่นคือจะได้ค่าสมาชิก $u^{(5)}$ ดังแสดงในสมการที่ 2.107

$$u^{(5)} = \left[\frac{.98.98.98.99.99.99.87.49.12.01.01.01.02.02.02}{.02.02.02.01.01.01.03.51.88.99.99.99.98.98.98} \right] \quad (2.107)$$

จากผลคำตอบของค่า สามารถนำมาเขียนกราฟได้ดังในรูปที่ 2.44 เพื่อแสดงคุณสมบัติของ fuzzy c-partition



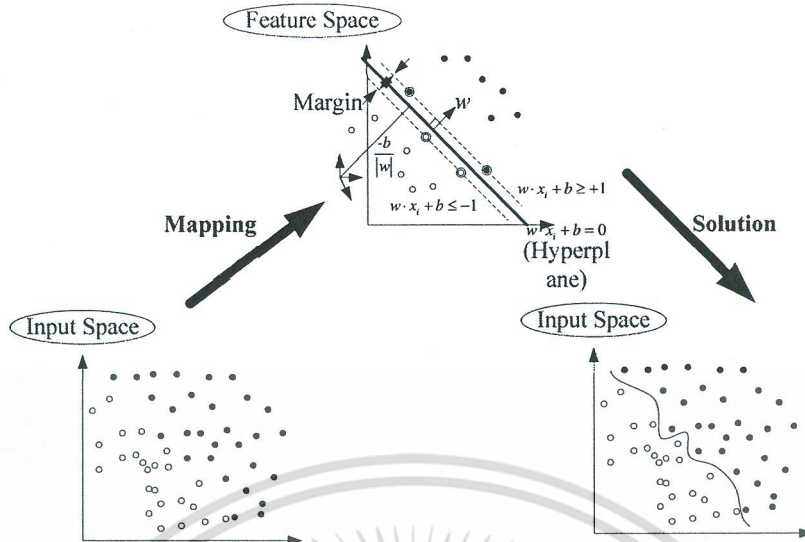
รูปที่ 2.44 คุณสมบัติของ Fuzzy C- Partition

เมื่อกำหนดเครื่องหมาย * แทนค่า u ในกลุ่ม A_2 จะเห็นว่าค่าของ u นั้นจะมีค่าอยู่ระหว่าง 0 ถึง 1 ซึ่งเครื่องหมาย * แทนค่า u ในกลุ่ม A_1 ซึ่งค่าของ u ก็มีลักษณะเดียวกันกับค่า u ในกลุ่ม A_2 คือมีค่าอยู่ระหว่าง 0 ถึง 1, เช่นกัน และเครื่องหมาย □ แทนค่าผลบวกของ u ในกลุ่ม A_1 และ A_2 ซึ่ง คือมีค่าเท่ากับ 1

2.7 อัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM)

SVM เป็นอัลกอริทึมที่ใช้ในการตัดแยก ที่เริ่มมีการนำมาใช้กันอย่างกว้างขวาง ในด้านการประมวลผลภาพดิจิทัล [22], [23] หลักการของ SVM คือการให้อินพุตที่ใช้ฝึกเป็นเวกเตอร์ในปริภูมิ N มิติ เช่น ถ้าในกรณีของ 2 มิติและ 3 มิติ จะเป็นจุดที่อยู่ในระนาบ xy และปริภูมิ xyz ตามลำดับ จากนั้นทำการสร้างไฮเปอร์เพลน (Hyper plane) ที่จะแยกกลุ่มของเวกเตอร์อินพุต ออกเป็นประเภทต่าง ๆ ตามที่กำหนด ในกรณีของ 2 มิติและ 3 มิติ ไฮเปอร์เพลนคือเส้นตรง และระนาบตามลำดับ ข้อเด่นของ SVM คือกรณีไม่สามารถใช้ไฮเปอร์เพลนในการตัดแยก ประเภทได้ตามปกติ SVM จะทำการแมป (Map) เวกเตอร์ในปริภูมิอินพุตให้เข้าสู่ Feature Space โดยใช้ฟังก์ชัน หรืออาจเรียกว่าฟังก์ชันเคอร์เนล (Kernel function) ชนิดต่าง ๆ เช่น ฟังก์ชันโพลีโนเมียล (Polynomial function) ฟังก์ชันเกาส์เซียนเรเดียล (Gaussian Radial function) ฟังก์ชันโพลีโนเมียลเรเดียล (Polynomial Radial function) เป็นต้น ใน Feature Space ดังกล่าวเวกเตอร์อินพุต สามารถ แยกประเภทได้โดยไฮเปอร์เพลน ดังแสดงในรูป 2.45 และ SVM มีความสัมพันธ์อย่างใกล้ชิดกับระบบเครือข่ายนิวรอน (Neural Network) กล่าวคือ SVM ที่ใช้ฟังก์ชันซิกมอยด์ (Sigmoid function) ในการแมป เทียบเท่ากับระบบเครือข่ายนิวรอน แบบป้อนไปหน้า (Feed forward network) ที่มี 2 ชั้น (Two layer feed-forward neural network) ข้อแตกต่างระหว่างซัพพอร์ตเวกเตอร์แมชชีนและระบบเครือข่ายนิวรอนก็คือในซัพพอร์ตเวกเตอร์แมชชีน การแก้สมการหาค่าน้ำหนักใช้การแก้สมการกำลังสอง (Quadratic Equation) ที่มีข้อจำกัดเชิงเส้น (Linear Constrained) แทนที่จะเป็นการหาค่าต่ำสุด (minimization) เหมือนกับของระบบเครือข่ายนิวรอน

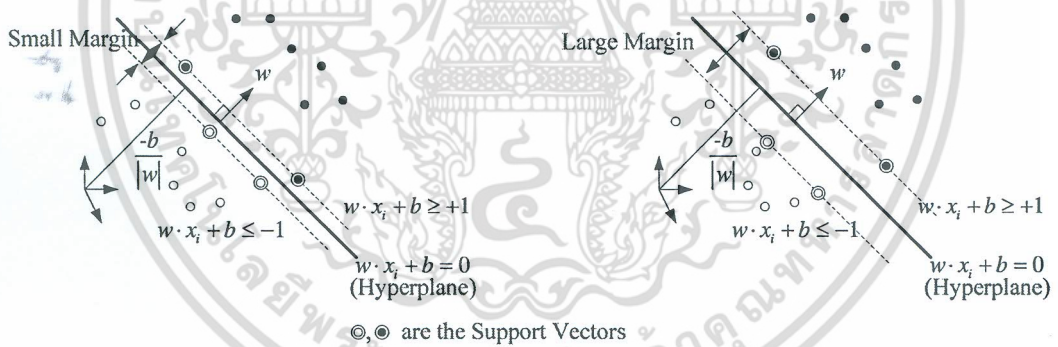
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.45 การแม็พเวกเตอร์เข้าสู่ Feature Space เพื่อทำการตัดแยก

ตัวอย่างการตัดแยกประเภทข้อมูลในสองมิติโดยใช้ SVM

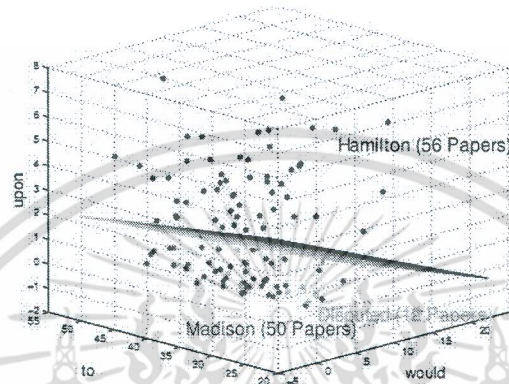
เพื่อให้เข้าใจในหลักการของ SVM พิจารณากรณีของการตัดแยกประเภทในสองมิติดังแสดงในรูปที่ 2.46



รูปที่ 2.46 ตัวอย่าง SVM ใน 2 มิติ

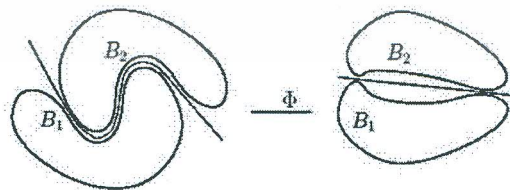
สมมติว่าเราต้องการตัดแยกประเภทอินพุตออกเป็น 2 กลุ่มโดยใช้ไฮเปอร์เพลนที่เป็นเส้นตรง จะเห็นว่ามีเส้นตรงจำนวนมากที่สามารถตัดแยกได้ แต่เส้นตรงเส้นไหนที่ดีที่สุด (Optimal Line) รูปที่ 2.46 แสดงตัวอย่างของ 2 เส้นตรง เราจะนิยามมาร์จิน (Margin) เป็นผลรวมของระยะห่างจากเส้นตรงที่เป็นไฮเปอร์เพลน (เส้นทึบในรูปที่ 2.46) ถึงเส้นตรงที่ผ่านอินพุตที่ไกลที่สุดและขนานกับไฮเปอร์เพลน ของทั้งสองกลุ่ม (เส้นประในรูปที่ 2.46) ระยะดังกล่าวอาจมองเป็นเวกเตอร์และมีชื่อว่า Support Vector อัลกอริทึม SVM จะเลือกไฮเปอร์เพลนที่ให้ค่า Margin มีค่าสูงสุด ดังแสดงในด้านขวามือของรูปที่ 2.46

กรณีของ 3 มิติจะเป็นทำนองเดียวกัน อัลกอริทึม SVM จะทำการสร้างไฮเปอร์เพลนที่เป็นระนาบเพื่อทำการแบ่งกลุ่มข้อมูลที่ให้มาจริง ซึ่งเป็นผลรวมของระยะห่างจากระนาบที่เป็นไฮเปอร์เพลนถึงระนาบที่ผ่านอินพุตที่ใกล้ที่สุด และขนานกับไฮเปอร์เพลนของทั้งสองกลุ่ม รูปที่ 2.47 แสดงตัวอย่างของการคัดแยกด้วย SVM ใน 3 มิติ



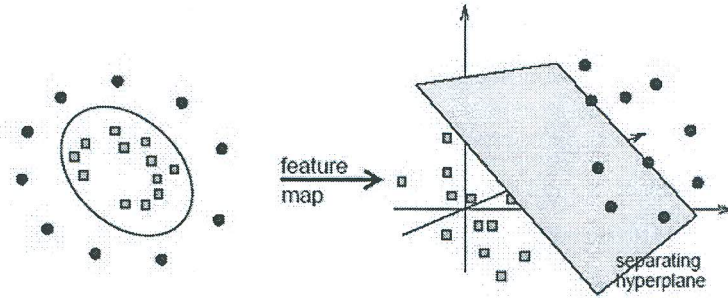
รูปที่ 2.47 ตัวอย่าง SVM ใน 3 มิติ

สำหรับกรณีที่กลุ่มของข้อมูลที่ไม่อยู่ในแนวเส้นตรง และไม่สามารถใช้ไฮเปอร์เพลนในการแยกกลุ่มข้อมูลได้ SVM จะไม่ใช้หลักการการประมาณเส้นโค้งที่ไม่เป็นเชิงเส้น (Non-linear Curve) เพื่อแบ่งกลุ่มข้อมูล แต่จะใช้ฟังก์ชันเคอร์เนล (Kernel Function) เพื่อแมปข้อมูลอินพุตไปยังปริภูมิใหม่ที่สามารถใช้ไฮเปอร์เพลนในการคัดแยกได้ รูปที่ 2.48 (ก) แสดงตัวอย่างของการแมปอินพุตที่มี 2 กลุ่ม (B_1 และ B_2) ด้วยฟังก์ชันเคอร์เนล ϕ ไปยังปริภูมิใหม่ที่สามารถแบ่งกลุ่มได้โดยใช้ไฮเปอร์เพลน รูปที่ 2.48 (ข) แสดงอีกตัวอย่างหนึ่ง



(ก) การแมปอินพุตที่มี 2 กลุ่ม (B_1 และ B_2) ด้วยฟังก์ชันเคอร์เนล ϕ ไปยังปริภูมิใหม่

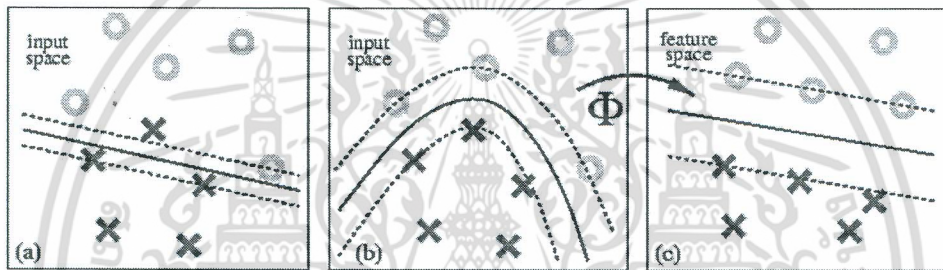
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ข) กรณีที่มีมิติที่ต่ำยากที่จะจำแนกได้ไปสู่มิติที่สูงและง่ายที่จะจำแนกได้

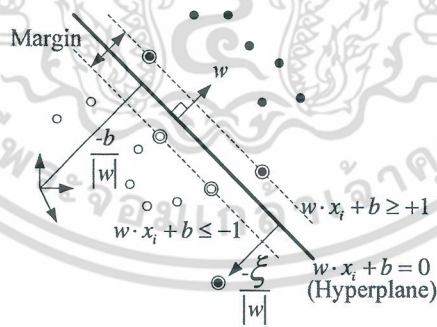
รูปที่ 2.48 ตัวอย่างการแมปข้อมูลอินพุตไปยังปริภูมิใหม่ที่สามารถใช้ไฮเปอร์เพลนในการตัดแยกได้

ในรูปที่ 2.49 แสดงตัวอย่างของการใช้ฟังก์ชันเคอร์เนลต่างๆในการแมปและผลการตัดแยกที่ได้



(ก) เชิงเส้น (ข) โพลีโนเมียล (ค) Feature Space

รูปที่ 2.49 ผลของการตัดแยกโดยใช้ฟังก์ชันเคอร์เนลต่างๆ



รูปที่ 2.50 ไฮเปอร์เพลน

2.7.1 การคำนวณหาไฮเปอร์เพลนที่ดีที่สุด (Optimal Hyper plane)

จากรูปที่ 2.50 พิจารณาปัญหาของการแยกเซตของเวกเตอร์ให้เป็น 2 กลุ่ม

$$D = \{(x_1, y_1), \dots, (x_n, y_n)\}, \quad x \in \mathcal{R}^n, y \in \{1, -1\} \tag{2.108}$$

ด้วยไฮเปอร์เพลน $w \cdot x_i + b = 0$ (2.109)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่แสดงในรูป 2.50

เรากล่าวว่าเซตของเวกเตอร์ถูกแบ่งด้วยไฮเปอร์เพลนอย่างเหมาะสมถ้าระยะระหว่างเวกเตอร์ที่ใกล้ที่สุดถึงไฮเปอร์เพลนมีค่าสูงสุด เพื่อความสะดวกเราจะใช้ Canonical Hyperplane ที่มีข้อจำกัดดังนี้

$$\min_i \langle w, x_i \rangle + b = 1 \quad (2.110)$$

กล่าวอีกนัยหนึ่งคือขนาด (นอร์ม) ของเวกเตอร์ค่าน้ำหนัก w ควรจะเท่ากับสัดส่วนผกผันกับระยะของจุดที่ใกล้ไฮเปอร์เพลนที่สุด นั่นคือไฮเปอร์เพลนในรูปแบบ Canonical ต้องเป็นไปตามข้อบังคับต่อไปนี้

$$y_i [\langle w, x_i \rangle + b] = 1; \quad i=1, \dots, l \quad (2.111)$$

ระยะ $d(w, b; x)$ ของจุด x ถึงไฮเปอร์เพลน (w, b) คือ

$$d(w, b, x) = \frac{|\langle w, x \rangle + b|}{\|w\|} \quad (2.112)$$

ไฮเปอร์เพลนที่ดีที่สุดคือไฮเปอร์เพลนซึ่งทำให้ค่ามาร์จิน (Margin), ρ , มีค่ามากที่สุดภายใต้ข้อบังคับสมการ (2.111) โดยที่มาร์จินถูกนิยามโดย

$$\begin{aligned} \rho &= \min_{x_i, y_i = -1} d(w, b, x) + \min_{x_i, y_i = 1} d(w, b, x) \\ &= \min_{x_i, y_i = -1} \frac{|\langle w, x_i \rangle + b|}{\|w\|} + \min_{x_i, y_i = 1} \frac{|\langle w, x_i \rangle + b|}{\|w\|} \\ &= \frac{1}{\|w\|} \min_{x_i, y_i = -1} |\langle w, x_i \rangle + b| + \min_{x_i, y_i = 1} |\langle w, x_i \rangle + b| \end{aligned}$$

โดยใช้สมการ (2.111)

$$\rho = \frac{2}{\|w\|} \quad (2.113)$$

ดังนั้นไฮเปอร์เพลนที่ดีที่สุดคือไฮเปอร์เพลนที่ทำให้เทอมต่อไปนี้มีค่าต่ำที่สุด

$$\Phi(w) = \frac{1}{2} \|w\|^2 \quad (2.114)$$

คำตอบ (หาค่าต่ำสุด) สำหรับสมการ (2.114) สามารถใช้จุด Saddle ของฟังก์ชัน Lagrange

$$\Phi(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i [\langle w, x_i \rangle + b] - 1) \quad (2.115)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ α คือ Lagrange Multipliers

สมการ Lagrange ต้องให้ค่าต่ำสุดเมื่อเทียบกับ w, b และค่าสูงเมื่อเทียบกับ α ซึ่งเรียกว่า Dual Problem (เนื่องจากมีทั้ง min และ max) นั่นคือ

$$\max_{\alpha} w(\alpha) = \max_{\alpha} (\min_{w,b} \Phi(w,b,\alpha)) \quad (2.116)$$

ค่าต่ำสุดของสมการ (2.115) เมื่อเทียบกับ w, b หาได้จาก

$$\begin{aligned} \frac{\partial \Phi}{\partial b} = 0 &\Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \\ \frac{\partial \Phi}{\partial w} = 0 &\Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i = 0 \end{aligned} \quad (2.117)$$

จากสมการ (2.115), (2.116) และสมการ (2.117) เราสรุป Dual Problem คือ

$$\max_{\alpha} w(\alpha) = \max_{\alpha} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle - \sum_{k=1}^l \alpha_k \quad (2.118)$$

คำตอบของสมการ(2.118) สามารถหาได้จาก

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle + \sum_{k=1}^l \alpha_k \quad (2.119)$$

ภายใต้ข้อกำหนด

$$\begin{aligned} \alpha_i &\geq 0; \quad i=1, \dots, l \\ \sum_{j=1}^l \alpha_j y_j &= 0 \end{aligned}$$

ดังนั้นไฮเปอร์เพลนที่เหมาะสมคือ

$$\begin{aligned} w^* &= \sum_{i=1}^l \alpha_i y_i x_i \\ b &= \frac{1}{2} \langle w^*, x_r + x_s \rangle \end{aligned}$$

โดยที่ X_r และ X_s คือ Support Vector ของแต่ละกลุ่ม

ถ้าในกรณีที่มีการแบ่งประเภทไม่สามารถทำได้โดยไฮเปอร์เพลนเชิงเส้น (Linear Separator) เราอาจใช้ฟังก์ชันเคอร์เนล $k(X_r, X_s)$ ในการแมป X_j ไปยัง Feature Space ตามที่กล่าวถึงก่อนหน้านี้ ดังนั้นสมการ (2.119) สามารถเขียนใหม่ได้เป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{k=1}^l \alpha_k \quad (2.120)$$

ตั้งนั้นไฮเปอร์เพลนที่เหมาะสมคือ

$$w^* = \sum_{i=1}^l \alpha_i y_i k(x_i, x) = 0$$

$$b = \frac{1}{2} \langle w^*, k(x_r + x) + k(x_s, x) \rangle$$

ฟังก์ชันเคอร์เนลที่นิยมใช้คือ

(i) ฟังก์ชันโพลีโนเมียล $k(x, x_i) = (\langle x, x_i \rangle + 1)^d$

(ii) ฟังก์ชันเกาส์เซียนเรเดียล (Gaussian Radial)

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)$$

(iii) ฟังก์ชันโพลีโนเมียลเรเดียล (Polynomial Radial)

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|}{2\sigma^2}\right)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วรรณกรรมงานวิจัยที่เกี่ยวข้อง

3.1 บทนำ

ในบทนี้จะเป็นการทบทวนวรรณกรรมงานวิจัยที่เกี่ยวข้องกับการทำวิทยานิพนธ์ โดยจะแบ่งเนื้อหาออกเป็นสองส่วนคือ ส่วนแรกเป็นการทบทวนวรรณกรรมงานวิจัยที่เกี่ยวข้องกับการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย (Lossless Compression) และ การบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression) ในส่วนที่สองจะเป็นงานวิจัยที่เกี่ยวข้องกับการปรับปรุงคุณภาพของภาพ (Peak Signal to Noise Ratio, PSNR) และการปรับปรุงอัตราบิต (Bit rat) ในระบบ

3.2 ทบทวนวรรณกรรมการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสียและการบีบอัดข้อมูลแบบที่มีการสูญเสีย

การลดขนาดข้อมูลภาพเป็นขั้นตอนที่สำคัญในกระบวนการของภาพทางการแพทย์ โดยเฉพาะเมื่อข้อมูลภาพเป็นภาพ CT และ ภาพ MRI ค่าตอบที่สำคัญของการลดขนาดข้อมูลภาพและสื่อประสม (multimedia) คือการลดขนาดไฟล์ข้อมูลภาพและในขณะเดียวกันยังคงรักษาคุณภาพ (quality) ของภาพให้เป็นที่น่าพอใจที่สุด ดังนั้นขนาดไฟล์ (file size) ของข้อมูลภาพจะมีขนาดเล็กในการจัดเก็บและใช้เวลาและพลังงานในการส่งน้อย

การลดขนาดข้อมูลภาพแบ่งออกเป็นสองชนิดคือ การลดขนาดข้อมูลภาพแบบที่ไม่มีการสูญเสีย (Lossless Compression) และ การลดขนาดข้อมูลภาพแบบที่มีการสูญเสีย (Lossy Compression) การลดขนาดข้อมูลภาพแบบที่ไม่มีการสูญเสียคือการลดขนาดข้อมูลภาพที่ไม่มีค่าการสูญเสียโดยมีการใช้ขนาดไฟล์ (file size) ของข้อมูลภาพจะมีขนาดเล็กเมื่อเทียบกับตอนที่ยังมีการลดขนาดข้อมูลและด้วยเหตุผลนี้การสร้างข้อมูลภาพกลับคืนจะเหมือนกับข้อมูลภาพอินพุต (original input image) ในทางตรงกันข้าม การลดขนาดข้อมูลภาพแบบที่มีการสูญเสียคือการลดขนาดข้อมูลภาพที่มีค่าการสูญเสีย ดังเช่น การสูญเสียส่วนที่เป็นรายละเอียดที่สำคัญของข้อมูลภาพ ดังนั้นจะให้ค่าอัตราการลดข้อมูลภาพ (compression rate) ที่สูง วิธีการควอนไทซ์แบบเวกเตอร์ (Vector Quantization, VQ) เป็นการลดขนาดข้อมูลภาพแบบที่มีการสูญเสียอีกหนึ่งวิธีที่ถูกนำมาใช้กันอย่างกว้างขวาง และเป็นวิธีการลดขนาดข้อมูลที่มีประสิทธิภาพสำหรับภาพดิจิทัลในการใช้กันอย่างแพร่หลายสำหรับการรับ (transmission) และการจัดเก็บ (storage) [24] ข้อดีของ VQ คือจะให้ อัตราการลดขนาด (compression rate) ข้อมูลที่สูง และในการถอดรหัสจะใช้วิธีการทางฮาร์ดแวร์ที่ง่าย การลดขนาดข้อมูลภาพใน VQ จะเป็นการสร้างความสัมพันธ์ระหว่างข้อมูล (mapping) ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูล (data set) ไปสู่บิต (bit stream) ดังนั้นการลดจำนวนของบิตที่ต้องการจะเป็นการแทนชุดข้อมูลเริ่มต้น (Original data) ผลลัพธ์ที่ได้จะใช้พื้นที่ในการจัดเก็บ (storage space) และเวลาในการสื่อสาร (communication time) ที่น้อย ในอัลกอริทึม VQ ภาพอินพุต (input image) จะถูกแบ่งออกเป็นบล็อกเล็ก ๆ แต่ละบล็อกจะเป็นการสร้างความสัมพันธ์ระหว่างข้อมูลไปสู่เวกเตอร์รหัส (code word) ของตัวเก็บรหัส (codebook) เพื่อกำหนดเป็นเวกเตอร์ ขั้นตอนนี้จะเรียกว่าขั้นตอนการเข้ารหัส (encoding process) ขณะที่ขั้นตอนที่ตรงกันข้ามกับขั้นตอนการเข้ารหัสจะเป็นขั้นตอนการสร้างภาพกลับคืน (reconstructed image) จะถูกเรียกว่าขั้นตอนการถอดรหัส (decoding process)

3.3 ทบทวนวรรณกรรมการปรับปรุงคุณภาพของภาพและการปรับปรุงอัตราบิต

ในช่วงสิบปีที่ผ่านมาจำนวนของงานวิจัยจะมีการมุ่งเน้น (focused) ไปที่การพัฒนาอัลกอริทึมไปสู่การปรับปรุงคุณสมบัติ (performance) และมีสองงานวิจัยหลักที่พยายามที่จะวิจัยคือ 1) ทำอย่างไรจะกำเนิดตัวเก็บรหัส (code book) ที่เหมาะสม (optimal codebook) ที่มีค่าความผิดพลาด (error) ต่ำระหว่างภาพอินพุตและภาพที่สร้างกลับ 2) ทำอย่างไรจะลดเวลาการสื่อสารในขั้นตอนการกำเนิดเวกเตอร์รหัส (code vector or code word) ที่อยู่ในตัวเก็บรหัส ดังมีการพัฒนาในอัลกอริทึม [3], [20]

ขั้นตอนการกำเนิดเวกเตอร์รหัสมีหลักการพื้นฐานคือการแบ่งกลุ่ม (cluster of training vector) เวกเตอร์สำหรับฝึก (training vector) ที่ได้มาจากข้อมูลภาพอินพุตออกกลุ่มๆ คำตอบที่เหมาะสม (optimal goal) ของการกำเนิดเวกเตอร์รหัสคือการออกแบบตัวเก็บรหัสที่มีค่าอัตราส่วนของสัญญาณสูงสุดต่อสัญญาณรบกวน (peak signal to noise ratio) ของภาพที่สร้างกลับ เช่น อัลกอริทึม Linde- Buzo-Gary (LBG) ใน [20] ถูกนำมาใช้กันเป็นอย่างมากและเป็นอัลกอริทึมที่ง่ายและเร็วสำหรับการกำเนิดการออกแบบตัวเก็บรหัส หลักการของ LBG คล้ายคลึงกับ การเกาะกลุ่มของ K-mean โดยจะเริ่มต้นจาก 1) สุ่มกำเนิด (randomly) k มิติ (k-dimensional) ของเวกเตอร์รหัส อยู่ในตัวเก็บรหัส การเกาะกลุ่มของเวกเตอร์สำหรับฝึกออกเป็น k กลุ่ม บนหลักการวัดระยะทางต่ำที่สุดของระยะทางแบบยูคลีเดียน (Euclidean distance) ระหว่างแต่ละเวกเตอร์สำหรับฝึกและเวกเตอร์รหัส คำนวณจุดศูนย์กลาง โดยการคำนวณศูนย์กลางของเวกเตอร์รหัสใหม่และทำซ้ำขั้นตอนจนกระทั่งลู่เข้าสู่ (convergence) อย่างไรก็ตาม อัลกอริทึม LBG ก็ยังแก้ปัญหาของการให้คำตอบที่ยังไม่ดีพอ ดังที่พบได้ในผลของค่าความผิดพลาดของภาพที่สร้างกลับ เพื่อให้มีปัญหาน้อยที่สุดนักวิจัยจำนวนมากได้มีการพัฒนาและปรับปรุงอัลกอริทึมอย่างต่อเนื่อง ตัวอย่างเช่น การนำเสนอ directed-search binary-splitting (DSBS) ใน [7] การนำเสนอการค้นหาด้วย mean-distance-ordered partial codebook search (MPS) ใน [25] การนำเสนอ double test of principal components (DTPC) ใน [26] การนำเสนอการปรับปรุงด้วย enhanced LBG (ELBG) ใน [27] และการนำเสนอทฤษฎี centroid neural network adaptive resonance theory (CNN-ART) ใน [28]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้การแปลงดีสครีตเวฟเล็ต (Discrete Wavelet Transform, DWT) กับตัวเก็บรหัสที่มีหลายระดับความละเอียด (multi-resolution) สำหรับการออกแบบกับตัวเก็บรหัสถูกนำเสนอใน [3] เพื่อลดขนาดของหน่วยความจำและเพิ่มคุณภาพของภาพ ในการออกแบบตัวเก็บรหัสด้วยการแปลง DWT เริ่มต้นด้วยการกำหนดเซตสำหรับการฝึกที่ประกอบด้วยเวกเตอร์ ที่เกี่ยวกับการตอบสนองของภาพที่แตกต่างกันตามค่าความละเอียดของการกระจาย (decomposition) ของ DTW. LBG ถูกนำมาประยุกต์ใช้สำหรับการกำเนิดตัวเก็บรหัสที่มีค่าความละเอียดหลายระดับโดยประกอบด้วยตัวเก็บรหัสย่อย (sub-codebook) ดังจะเห็นได้ชัดว่าความผิดพลาดของภาพลดลง และแต่ละตัวเก็บรหัสย่อยจะประกอบด้วยเวกเตอร์รหัสจำนวนน้อย และค่าเวลาในการประมวลผล (computational cost) สำหรับขั้นตอนในการเข้ารหัสจำเป็นต้องมีค่าลดลง การประยุกต์นำ DWT สำหรับเทคนิคการเข้ารหัสภาพถูกนำเสนอใน [3] การใช้อัลกอริทึม embedded zero-tree wavelet (EZW) สามารถเพิ่มประสิทธิภาพของการลดขนาดข้อมูลและลดเวลาในการประมวลผลถูกนำเสนอใน [29] การนำเสนอ EZW คือการเข้ารหัสด้วย zero-tree บนหลักการ significance maps ด้วยเลขฐานสอง (binary maps) เพื่อแสดงตำแหน่งของค่าศูนย์และค่าที่ไม่ใช่ค่าศูนย์ของค่าสัมประสิทธิ์ของ DWT หลังจากการควอนไทซ์ (quantized) กับการประมาณแบบต่อเนื่อง (successive approximation) อัลกอริทึม EZW ยังประยุกต์นำการเข้ารหัสเลขคณิต (arithmetic coding) สำหรับอักขระ (string) ของรหัสเวกเตอร์รหัสสุดท้ายซึ่งสามารถลดค่าความผิดพลาดกำลังเฉลี่ย (Mean Square Error, MSE) หรือเพิ่มค่าสัดส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak Signal to Noise Ratio, PSNR) อัลกอริทึม EZW สามารถพบได้ใน [30] และถูกยืนยันว่าทำให้ได้รับผลที่ดีขึ้นในการประยุกต์ร่วมกับอัลกอริทึมเจเนติก (Genetic Algorithm, GA) กับ DWT บนวิธีการกำเนิดเวกเตอร์รหัสสามารถพบได้ใน [4] เป้าหมายของการประยุกต์ใช้ GA เพื่อป้องกันปัญหาของค่าต่ำสุดเฉพาะที่ที่ทำให้เหมาะสมที่สุด (local-minimum optimization) ด้วยการออกแบบตามทฤษฎีเกรย์ (grey-base) ร่วมกับการศึกษาโครงข่าย (Grey-base designed Competitive Learning Network, GCLN) GCLN คือ โครงข่ายประสาทเทียมแบบซูปเปอร์ไวส์ (supervised neural network) ใช้การวิเคราะห์ความสัมพันธ์ตามทฤษฎีเกรย์ เลือกส่วนที่ดีที่สุดบนหลักการความคล้ายคลึงกัน (similarity) ระหว่างเวกเตอร์สำหรับฝึกและคำรหัส ปรับปรุง DWT บนหลักการ GCLN ร่วมกับ GA หรือเรียกว่า WT+GGCLN ซึ่งจะนำไปสู่การหาค่าตอบที่เหมาะสมกับค่าเฉลี่ยของ PSNR ที่ 1.65 dB สูงกว่าวิธีการทั่วไปของ LBG วิธีในลำดับถัดมาของการกำเนิดคำรหัส คือการเข้ารหัสด้วยหลักการ DWT สามารถพบได้ใน [31] ซึ่งใช้หลักการลดการซ้ำกันของข้อมูลภาพทางการแพทย์ (medical image) และ ลดค่าสัมประสิทธิ์ของ DWT โดยการใช้การเข้ารหัสแบบฮัฟแมน (Huffman encoding) เพื่อปรับปรุงประสิทธิภาพของเทคนิค และเรียกว่าเทคนิคนี้ว่า WT+HM ซึ่งเทคนิคถูกนำมาใช้เพื่อปรับปรุง PSNR อัตราการลดข้อมูล (compression ratio) และอัตราบิต (bit rate) หรือจำนวนบิตต่อค่าจุดภาพ (pixel)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวลาในการประมวลผล (computation time) คือคำตอบที่สำคัญอีกอันหนึ่งสำหรับการกำเนิดตัวเก็บรหัส วิธีการเข้ารหัสแบบทั่วไปอย่างเช่น การเข้ารหัสแบบ LBG จะใช้หลักการพื้นฐานของการหาค่าที่เหมาะสมที่สุด (optimization) โดยใช้ขั้นตอนการวนซ้ำ (iterative procedure) เพื่อให้ได้ค่าเวลาที่ใช้ที่เหมาะสมที่สุด (time-consuming) เวลาที่ใช้สำหรับการสร้างตัวเก็บรหัสจะลดลงโดยการใช้อัลกอริทึมการแปลงเวฟเลต (DWT) และสร้างตัวเก็บรหัสที่มีความละเอียดแบบหลายระดับ (multi-resolution codebook) [3] กลุ่มนักวิจัยในลำดับถัดมาแนะนำการเข้ารหัสแบบ Embedded Zero-Tree Wavelet, (EZW) [29], [30], [32] ไม่เฉพาะการปรับปรุงในเรื่องของการบีบอัดข้อมูล แต่ยังคงรวมถึงเวลาที่ใช้ในการสร้างตัวเก็บรหัสด้วยหลักการดังกล่าวเพื่อให้กำเนิดเวกเตอร์รหัสที่สามารถหยุดที่จุดต่าง ๆ ซึ่งเป้าหมาย คือ อัตราบิต (bit rate) หรือ ให้ค่าความผิดพลาดต่ำที่สุด นักวิจัยกลุ่มถัดมา Hsieh และ Liu [7] นำเสนอหลักการลดเวลาของการคำนวณสำหรับการค้นหาเวกเตอร์รหัส โดยการใช่วิธีการค้นหาแบบเร็วบนหลักการ Triangle Inequality Elimination (TIE) กับการควบคุมหลายจุด หลักการ TIE เพิ่มความเร็วขึ้นได้ 39.2 % โดยปราศจากค่าความสูญเสียคุณภาพของภาพ นักวิจัยกลุ่มถัดมา Chang และ Wu [8] นำเสนอวิธีใหม่ชื่อ Partial-Search สำหรับหลักการของ VQ บนหลักการของโครงสร้างกราฟ (graph structure) วิธีที่นำเสนอสามารถปรับปรุงคุณภาพของภาพให้สูงขึ้น 99.98 % ด้วยวิธีการค้นหาทั่วไปแบบเต็ม VQ (Full Search) เรียกว่า FSVQ และใช้เวลาการค้นหาที่ต่ำ

3.4 งานวิจัยนี้เรานำเสนอการบีบอัดข้อมูลภาพแบบใหม่บนหลักการชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC)

ดังที่ทราบกันโดยทั่วไปแล้วว่าการบีบอัดข้อมูลภาพแบบ VQ เป็นการบีบอัดข้อมูลที่มีค่าความสูญเสีย (Lossy Compression) ดังนั้นจึงมีค่าความผิดพลาดของระบบเกิดขึ้น ซึ่งเป็นค่าความผิดพลาดระหว่างภาพต้นฉบับ (Original Image) และภาพที่สร้างกลับคืน (reconstructed Image) สำหรับการบีบอัดข้อมูลแบบ VQ โดยทั่วไปแล้วจะไม่สนใจค่าความผิดพลาดของระบบ อย่างไรก็ตามเพื่อปรับปรุงคุณภาพของภาพที่สร้างกลับคืนกับเทคนิคดังกล่าวร่วมกับค่าผิดพลาดของภาพ (Error Image) ในวิธีการบีบอัดข้อมูล และบวกค่าผิดพลาดของภาพกับภาพที่สร้างกลับคืน ดังนั้นขั้นตอนของวิธี SEC ที่นำเสนอบนหลักการบีบอัดข้อมูลภาพแบบ VQ ประกอบด้วยขั้นตอนดังนี้

ขั้นตอนที่ 1 การชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC) แบบใหม่สำหรับขั้นตอนการกำเนิดเวกเตอร์รหัสของ VQ ซึ่งภาพผิดพลาด (Error Image) จะถูกนำไปบวกรวมกับภาพที่สร้างกลับคืน (Reconstructed Image) ผลลัพธ์ที่ได้ทำให้ได้ค่า PSNR สูงขึ้น

ขั้นตอนที่ 2 เวกเตอร์สำหรับฝึก (Training Vectors) จะถูกแบ่งออกเป็นกลุ่มของจุดภาพที่มีค่าความหนาแน่นเหมือนกัน โดยใช้เทคนิคการแปลง DWT เพื่อกำเนิดเวกเตอร์รหัส ผลลัพธ์ที่ได้ทำให้เวลาที่ใช้ในการกำเนิดเวกเตอร์รหัสลดลง

ขั้นตอนที่ 3 นำ SVM มากำหนดเงื่อนไขเริ่มต้นให้กับ FCM เพื่อใช้ในการแบ่งแยกกลุ่มข้อมูล และ SVM ถูกนำมาใช้สำหรับกำหนดเงื่อนไขเริ่มต้น (Initial Condition) จากที่กล่าวมาแล้วผลทำให้ลดเวลาในขั้นตอนการแบ่งแยกกลุ่มข้อมูล

ขั้นตอนที่ 4 การเข้ารหัสฮัฟแมน (Huffman Coding) ถูกนำมาใช้สำหรับการลดการซ้ำกันของเวกเตอร์รหัสและการปรับปรุงอัตราบิต

องค์ประกอบของงานวิจัยมีดังนี้ ส่วนที่หนึ่ง กล่าวถึง ทบทวนวรรณกรรมงานวิจัยที่เกี่ยวข้อง (Introduction) ส่วนที่สอง กล่าวถึง องค์ประกอบที่สำคัญ (Important Factors) ของ VQ ส่วนที่สามและส่วนที่สี่ กล่าวถึงระบบการเข้ารหัส (Coding System) และการออกแบบตัวเก็บรหัส (Code Book) ส่วนที่ห้าและหก กล่าวถึงการปรับปรุงอัตราบิต (Bit Rate) และการชดเชยค่าผิดพลาดของระบบ (System Error Compensation) ส่วนที่เจ็ด กล่าวถึงการทดลองและผลการทดลอง ส่วนที่แปด กล่าวถึง ข้อเสนอของงานวิจัย



บทที่ 4

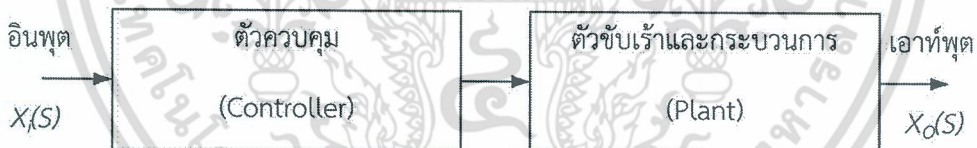
รูปแบบระบบควบคุมระบบเปิดและระบบปิด

4.1 บทนำ

ในบทนี้จะเป็นการกล่าวถึงการทบทวนระบบควบคุมที่เกี่ยวข้องกับการทำวิทยานิพนธ์ โดยจะแบ่งเนื้อหาออกเป็นสองส่วนคือ ส่วนแรกเป็นการทบทวนระบบควบคุมแบบเปิด ในส่วนที่สองจะเป็นการทบทวนระบบควบคุมแบบปิด ซึ่งระบบควบคุมแบบปิดจะถูกนำมาออกแบบเป็นระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์และการชดเชยค่าผิดพลาดของระบบ

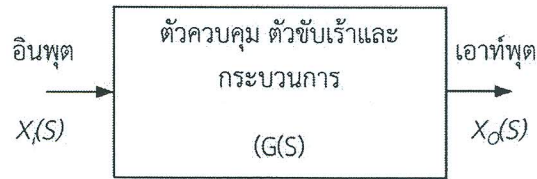
4.2 ระบบควบคุมแบบระบบเปิด (Open Loop Control System)

ระบบควบคุมแบบวงเปิดเป็นระบบที่เอาต์พุตไม่มีผลต่อการควบคุม หรือ กล่าวอีกนัยหนึ่งระบบควบคุมแบบวงเปิดไม่ได้มีการวัดเอาต์พุตเพื่อที่จะป้อนกลับไปเปรียบเทียบกับอินพุต ตัวอย่างการควบคุมในลักษณะนี้ที่เห็นได้ชัดได้แก่ เครื่องซักผ้า คือ การแช่ผ้า การซัก และการล้าง โดยเครื่องจะทำงานตามเวลา เครื่องไม่ได้มีการวัดเอาต์พุตซึ่งคือ ความสะอาดของเสื้อผ้า



รูปที่ 4.1 การควบคุมแบบวงเปิด

ระบบควบคุมแบบวงเปิด เอาต์พุตไม่ได้ถูกเปรียบเทียบกับอินพุตอ้างอิง ดังนั้นจะต้องมีการกำหนดเงื่อนไขการทำงานที่คงที่สำหรับอินพุตอ้างอิงแต่ละตัว ทำให้ความแม่นยำของระบบขึ้นอยู่กับ การปรับแต่ง (Calibration) ถ้ามีการรบกวนในระบบวงเปิด ระบบจะไม่สามารถทำงานตามที่ต้องการได้ การควบคุมแบบวงเปิดในทางปฏิบัติจะเหมาะกับงานที่รู้ความสัมพันธ์ระหว่างอินพุตและเอาต์พุต และต้องไม่มีการรบกวนทั้งภายในและภายนอก เพราะระบบไม่มีการป้อนกลับ ขอให้สังเกตว่าระบบควบคุมใดก็ตามที่มีการทำงานบนพื้นฐานของเวลาหรือทำงานตามเวลาที่กำหนดจะเป็นการควบคุมแบบวงเปิด เช่นการควบคุมสัญญาณไฟจราจรก็เป็นอีกตัวอย่างหนึ่งของการควบคุมแบบวงเปิด



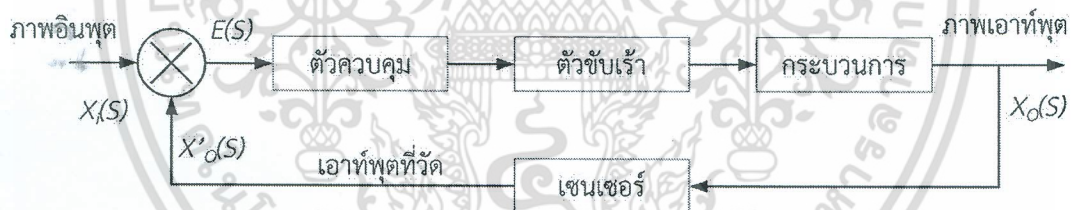
รูปที่ 4.2 แบบสมการวงเปิด

สมการ Open-Loop transfer function, $T(S)$ สามารถเขียนได้ดังต่อไปนี้

$$T(S) = X_o(S) / X_i(S) \quad (4.1)$$

4.3 ระบบควบคุมแบบระบบปิด (Close Loop Control System)

ระบบควบคุมแบบลูปปิด หรือระบบควบคุมแบบป้อนกลับ (Feedback control systems) โดยปกติแล้วระบบควบคุมแบบป้อนกลับก็คือระบบควบคุมแบบวงปิดนั่นเอง ซึ่งคำสองคำนี้สามารถใช้แทนกันได้ ระบบลักษณะนี้จะมีการทำงานที่พยายามลดความแตกต่างระหว่างเอาต์พุตและอินพุต อ่างอิง ความแตกต่างนี้ก็คือความผิดพลาดของระบบ (System error) ซึ่งจะถูส่งให้กับตัวควบคุม เพื่อให้ตัวควบคุมสั่งงานในการลดความผิดพลาดนี้ และนำเอาต์พุตไปที่ค่าที่ต้องการตั้งแสดงในรูปที่ 4.3

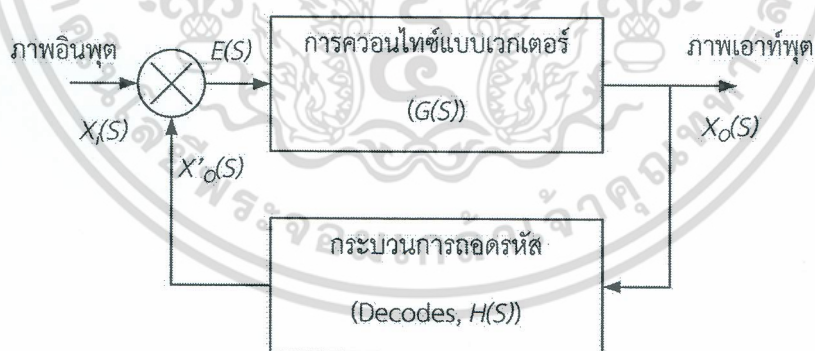


รูปที่ 4.3 การควบคุมแบบวงปิด

ตัวอย่างการควบคุมแบบวงปิดได้แก่ระบบควบคุมอุณหภูมิภายในห้องอุณหภูมิจริงภายในห้องจะถูวัด และนำมาเปรียบเทียบกับอุณหภูมิอ้างอิงหรืออุณหภูมิที่ต้องการแล้วสั่งการให้เทอร์มอสแตต (Thermostat) เปิดหรือปิดอุปกรณ์ทำความร้อนหรือทำความเย็นเพื่อรักษาอุณหภูมิภายในห้องไว้ในระดับต้องการโดยไม่ต้องคำนึงถึงสภาพอุณหภูมิภายนอกห้อง ระบบควบคุมแบบป้อนกลับไม่ได้จำกัดอยู่แค่ในงานด้านวิศวกรรมเท่านั้น แต่สามารถพบได้ในสาขาอื่นๆเช่นกัน เช่นในร่างกายของมนุษย์ เป็นระบบควบคุมแบบป้อนกลับขั้นสูง โดยอุณหภูมิและความดันเลือดของร่างกายจะถูกรักษาให้คงที่ โดยวิธีการป้อนกลับทางสรีรวิทยา ในความเป็นจริงแล้วการควบคุมแบบป้อนกลับเป็นกระบวนการที่สำคัญมากของร่างกายมนุษย์ในการทำให้ร่างกายไม่อ่อนไหวต่อสิ่งแวดล้อมภายนอกและทำให้อวัยวะในร่างกายสามารถทำงานในสภาวะแวดล้อมที่เปลี่ยนแปลงไปได้อย่างถูกต้อง

หลักการพื้นฐานของระบบควบคุมแบบป้อนกลับสามารถกล่าวได้ว่า ระบบควบคุมแบบป้อนกลับประกอบด้วยเส้นทางหรือวงรอบของสัญญาณป้อนกลับซึ่งเป็นสัญญาณเอาต์พุต $X_o(S)$ ตั้งแต่หนึ่งวงรอบขึ้นไป แล้วนำสัญญาณป้อนกลับนี้มาเปรียบเทียบกับสัญญาณสั่งการหรือสัญญาณอ้างอิง $X_i(S)$ จะได้ผลต่างระหว่างสัญญาณทั้งสองเป็น $E(S) = X_i(S) - X_o(S)$ เพื่อนำไปควบคุมสัญญาณเอาต์พุต $X_o(S)$ ให้มีค่าตามที่กำหนดโดยสัญญาณอ้างอิง $X_i(S)$ ระบบควบคุมนี้ประกอบด้วย ส่วน forward (forward path), ส่วนป้อนกลับ (feedback path) และส่วนตรวจจับค่าความคลาดเคลื่อน (error-sensing device) ส่วนตรวจจับค่าความคลาดเคลื่อนนี้จะเปรียบเทียบค่าสัญญาณอินพุตอ้างอิงกับค่าสัญญาณเอาต์พุตจริงๆหรือค่าที่เป็นฟังก์ชันของสัญญาณเอาต์พุต แล้วส่งสัญญาณที่เกิดจากผลต่างของสัญญาณทั้งสองนี้ออกไป

ตัวอย่างการควบคุมแบบวงปิดที่นำมาใช้ในขั้นตอนการบีบอัดข้อมูลภาพกับการบีบอัดข้อมูลภาพแบบการควอนไทซ์แบบเวกเตอร์เรียกว่าการควบคุมแบบวงปิด (Close Loop Control, CLC) สำหรับข้อมูลภาพ ดังแสดงในรูปที่ 4.4 โดยมีตัวแปรข้อมูลดังนี้
เอาต์พุตที่ต้องการ คือภาพอินพุต (Input image, $X_i(S)$)
เอาต์พุตจริง คือภาพเอาต์พุต (Output image, $X_o(S)$)
เซนเซอร์ คือกระบวนการถอดรหัส (Decodes, $H(S)$)
ตัวควบคุม ตัวขับเร็วและกระบวนการ คือการควอนไทซ์แบบเวกเตอร์ $G(S)$
เอาต์พุตที่วัด คือเอาต์พุตจากกระบวนการถอดรหัส $X'_o(S)$



รูปที่ 4.4 การควบคุมแบบวงปิดสำหรับข้อมูลภาพ

สมการ Close-Loop transfer function, $T(S)$ สามารถเขียนได้ดังต่อไปนี้

$$T(S) = X_o(S) / X_i(S) = G(S) / [1 + G(S) H(S)] \quad (4.2)$$

สมการคุณลักษณะ (Characteristic equation) คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$1 + G(S) H(S) = 0 \quad (4.3)$$

จัดรูปสมการใหม่จะได้ว่า

$$G(S) H(S) = -1 \quad (4.4)$$

เนื่องจากทรานสเฟอร์ฟังก์ชัน (transfer function) เป็นตัวแปรเชิงซ้อน (Functions of a Complex Variable) จากสมการ (4.4) จะได้ว่าขนาดของ $G(S)H(S) = 1$ และมุมมีค่าเท่ากับ 180 องศา เมื่อพิจารณาเฉพาะขนาด(gain) จะได้

$$|G(S) H(S)| = 1 \quad (4.5)$$

เมื่อพิจารณาเฉพาะมุม (phase) จะได้ว่า

$$\angle G(S)H(S) = 180 + 360 \quad (4.6)$$

และเมื่อพิจารณา Closed-Loop transfer function ดังต่อไปนี้

$$T(S) = X_o(S) / X_i(S) = [K(S + Z_1)(S + Z_2) \dots (S + Z_m)] / [(S + P_1)(S + P_2) \dots (S + P_n)]$$

จะได้ว่า $-Z_1, -Z_2, \dots, -Z_m$ เป็น zero ของทรานสเฟอร์ฟังก์ชัน และ $-P_1, -P_2, \dots, -P_n$ เป็น pole ของทรานสเฟอร์ฟังก์ชัน อีกทั้งยังเป็นเส้นทางเดินของราก (root locus) ของสมการคุณลักษณะอีกด้วย ถ้า $T(S)$ เป็นผลหารที่ประกอบด้วยเศษและส่วนที่เป็น polynomial จะได้ว่า pole และ zero ของ $T(S)$ ประกอบด้วยจำนวนจริงหรือคู่ conjugate ของจำนวนเชิงซ้อนเท่านั้น

บทที่ 5

การจำลองระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ควเตอร์ควอนไทต์และการชดเชยค่าผิดพลาดของระบบ

5.1 บทนำ

ในบทนี้จะเป็นการกล่าวถึงการจำลองระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ควเตอร์ควอนไทต์และการชดเชยค่าผิดพลาดของระบบ โดยจะแบ่งเนื้อหาออกเป็นสี่ส่วนดังนี้คือ ส่วนแรก กล่าวถึงปัจจัยสำคัญของอัลกอริทึมเว็ควเตอร์ควอนไทต์ ส่วนที่สอง กล่าวถึงการออกแบบตัวเก็บรหัส ส่วนที่สาม กล่าวถึงการปรับปรุงอัตราบิต และส่วนที่สี่ กล่าวถึงการชดเชยค่าผิดพลาดของระบบ

5.2 ปัจจัยที่สำคัญของอัลกอริทึมเว็ควเตอร์ควอนไทต์ (The important factors of VQ Algorithm)

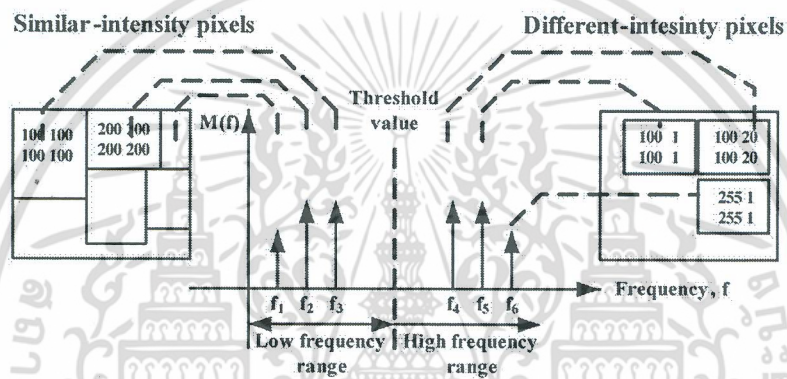
การบีบอัดข้อมูลภาพแบบ VQ ค่าตอบที่สำคัญก็คือการออกแบบตัวเก็บรหัสให้มีค่าอัตราส่วนของสัญญาณข้อมูลต่อสัญญาณรบกวน (Peak Signal to Noise Ratio, PSNR) ที่สูงและในเวลาเดียวกันอัตราบิต (Bit Rate) ที่ต่ำและต้นทุนของเวลา (Computation Time) ที่ใช้ในการคำนวณที่ต่ำ

ตัวเก็บรหัส (Code Book)

การควอนไทซ์แบบเว็ควเตอร์ เป็นการสร้างความสัมพันธ์ (Q) ระหว่างข้อมูลของสเปซยูคลิดิอัน (Euclidian Space) k มิติ (R^k) ลงในเซตย่อยของ R^k ซึ่งประกอบด้วยเซต C ที่มี N เอ็นทิตี และเรียกว่า เว็ควเตอร์รหัส (Code Vector or Code Word) ซึ่ง $Q: R^k \rightarrow C$ โดยที่ $C = (y_1, y_2, \dots, y_N)$ และ $y_i \in R^k$ ของแต่ละ $i \in \mathcal{T} \equiv \{1, 2, \dots, N\}$ โดยที่ C เรียกว่าตัวเก็บรหัส (Code Book) ค่าความละเอียด (Resolution) ของ VQ คือ $r = (\log_2 N)/k$ ซึ่งเป็นการวัดค่าจำนวนบิตต่อเว็ควเตอร์ (bit per vector) ที่เป็นตัวแทนของเว็ควเตอร์อินพุต (input vector) และให้เป็นการแสดงถึงค่าความแม่นยำ (accuracy) ของตัวเก็บรหัส ถ้าตัวเก็บรหัสนั้นถูกออกแบบมาอย่างดี วิธีการทำโดยปกติทั่วไปของการให้กำเนิดเว็ควเตอร์รหัสจะใช้บล็อกขนาด 2×2 ที่อัตราบิตที่ต่ำเราจำเป็นต้องใช้มิติของ k ที่มีขนาดใหญ่ หรือใช้บล็อกที่มีขนาดใหญ่ อย่างไรก็ตาม เมื่อข้อมูลของเว็ควเตอร์รหัสที่กำเนิดจากค่าความหนาแน่นในบล็อกของจุดภาพ (Pixel Intensity) ที่มีขนาดใหญ่ ผลลัพธ์จะทำให้ได้รับค่า PSNR ที่ต่ำ ในทางตรงกันข้ามถ้าเราลดขนาดของบล็อก เราก็จะได้รับค่า PSNR ที่สูงขึ้น ดังนั้นบล็อกที่มีขนาดเล็กก็จะให้ค่าความผิดเพี้ยนต่ำ (Distortions) ระหว่างค่าความหนาแน่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และขนาดของบล็อก ผลลัพธ์ที่ได้คืออัตราบิตที่สูง ในงานวิจัยนี้ เรานำเสนอการเข้ารหัสภาพ VQ ที่มีค่า PSNR สูงที่อัตราบิตที่เท่ากัน เพื่อที่จะให้ได้ค่าความผิดเพี้ยนต่ำ เราจำเป็นต้องใช้บล็อกที่มีขนาดใหญ่และกำหนดให้มีค่าความหนาแน่นของจุดภาพในบล็อกที่มีค่าความหนาแน่นเหมือนกัน เรานำเสนอการใช้ DWT สำหรับการแบ่งกลุ่มของจุดภาพให้กับการกำเนิดตัวเก็บรหัส DWT เป็นเครื่องมือที่มีการนำมาใช้ในการพัฒนาการประมวลผลสัญญาณในช่วง 20 ปีที่ผ่านมา ดังจะพบเห็นได้จากเมื่อมันถูกนำมาใช้กับการบีบอัดภาพ [33], [34] มันเป็นเทคนิคที่มีประสิทธิภาพมากสำหรับการบีบอัดที่ไม่ขึ้นอยู่กับส่วนประกอบของภาพ และมันเป็นที่น่าสนใจในส่วนที่เกี่ยวกับการบีบอัดภาพ เพราะว่ามันสามารถกำหนดตำแหน่งของค่าสัมประสิทธิ์ (coefficient) ในโดเมนสเปเชียล และโดเมนความถี่ (Spatial and Frequency domain) ได้ดี [35]-[37]



รูปที่ 5.1 การวิเคราะห์สัญญาณในโดเมนความถี่

รูปที่ 5.1 แสดงการใช้งานของ DWT เพื่อแยกจุดภาพ ($M(f)$) ที่มีส่วนประกอบของความถี่ต่ำและความถี่สูง บล็อกของจุดภาพขนาด 2×2 ที่มีค่าจุดภาพเหมือนกันอยู่ที่ตำแหน่งในช่วงของความถี่ต่ำ ในทางกลับกันบล็อกของจุดภาพที่มีค่าจุดภาพแตกต่างกันจะอยู่ที่ตำแหน่งในช่วงของความถี่สูง เราสามารถใช้ส่วนประกอบของความถี่เพื่อกำหนดตำแหน่งของบล็อก และกำหนดขนาดในสเปเชียลโดเมน (Spatial Domain)

อัตราบิต (Bit Rate)

อัตราบิตเป็นอีกหนึ่งปัจจัยที่สำคัญในการบีบอัดภาพ สำหรับลำดับของเวกเตอร์อินพุตที่จะถูกควอนไทซ์และการส่งอัตราบิตหรืออัตราการส่ง (Transmission Rate, R) ในบิตต่อเวกเตอร์ (Bits per Vector) กำหนดดังนี้ $R = kr$ โดยที่ r คือ ค่าความละเอียด (Resolution) และ k คือ มิติของเวกเตอร์ (Vector dimension) ถ้า f_v แสดงถึงอัตราของเวกเตอร์ หรือจำนวนของอินพุตเวกเตอร์ที่จะเข้ารหัสต่อวินาที (Second) ดังนั้น อัตราบิต (R_s) ในบิตต่อวินาที กำหนดดังนี้ $R_s = krf_v$ ในอัลกอริทึมเวกเตอร์ควอนไทซ์ ดัชนี (index) คือ การวัดระยะทางยูคลีเดียน (Euclidean Distance) ระหว่างบล็อกย่อยอินพุต (Sub-Input Block) ภาพและแต่ละเวกเตอร์รหัสในตัวเก็บรหัส สำหรับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บล็อกที่มีขนาดใหญ่หรือมิติของเวกเตอร์มีขนาดใหญ่ จำนวนดัดขึ้นหรือจำนวนของอินพุตเวกเตอร์ก็จะ มีน้อย ในทางตรงกันข้ามสำหรับบล็อกที่มีขนาดเล็กจำนวนของดัดขึ้นก็จะมีมาก ตัวอย่าง เช่น ภาพที่มี ขนาด 512×512 จุดภาพและมิติของเวกเตอร์ขนาด 8×8 จุดภาพ จำนวนดัดขึ้นจะมีค่าเท่ากับ $4,096 = (512 \times 512) / (8 \times 8)$ และถ้ามิติของเวกเตอร์ขนาด 2×2 จุดภาพ ภาพมีขนาดเท่ากัน จำนวนของดัดขึ้นจะมีค่าเท่ากับ $65,536 = (512 \times 512) / (2 \times 2)$ ในกรณีที่ต้องการสร้างภาพกลับคืน ที่มีคุณภาพของภาพสูง บล็อกที่มีขนาดเล็กจะให้คุณภาพของภาพกลับคืนดีกว่า เมื่อจำนวนของดัดขึ้นมี จำนวนมากมันก็อาจจะมีดัดขึ้นที่ซ้ำกันซึ่งเป็นสิ่งที่สำคัญมากเช่นกัน เพื่อกำจัดปัญหาการซ้ำกันของ ดัดขึ้น เราจะเลือกการเข้ารหัสแบบฮัฟแมน (Huffman Coding) มาเข้ารหัสดัดขึ้น ซึ่งผลลัพธ์ที่ได้คือการ ปรับปรุงอัตราบิด

คุณภาพของภาพ (PSNR)

คุณภาพของภาพที่ดีเป็นสิ่งที่ต้องการสำหรับการบีบอัดข้อมูล เราต้องการออกแบบตัวเก็บ รหัสที่มีค่าความผิดพลาดต่ำของภาพที่สร้างกลับคืนและรวมถึงค่าคุณภาพของภาพที่สูง การออกแบบ ตัวเก็บรหัสที่เหมาะสมที่สุดมีทางเลือกทำได้ในหลายทาง ทางหนึ่งที่ได้ คือ การเลือกขนาดของบล็อก อัตราการบีบอัด (Compression Ratio) เป็นอีกทางหนึ่งที่มีผลต่อคุณภาพของภาพเพื่อปรับปรุง คุณภาพของภาพ ในงานวิจัยนี้จึงนำเสนอการชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC) ในระบบดังกล่าว ภาพผิดพลาด (Error Image) ระหว่างภาพอินพุต (Input Image) และภาพที่สร้างกลับคืน (Reconstructed Image) จะถูกนำมาเข้ารหัสและถูกส่งไปที่ เอาท์พุต และสุดท้ายจะถูกชดเชยกลับไปภาพที่สร้างกลับคืน ผลการทดลองแสดงให้เห็นว่า อัลกอริทึม SEC ให้คุณภาพของ PSNR ที่ดีเมื่อเปรียบเทียบกับอัลกอริทึมที่รู้จักกันโดยทั่วไป

เวลาที่ใช้ในการคำนวณ (Computational Time)

ในก่อนหน้านั้น อัลกอริทึม VQ จะขึ้นอยู่กับปัญหาของการแบ่งกลุ่ม (Clustering Problem) ซึ่งจะต้องทำให้เหมาะสมที่สุด เวลาที่ใช้ในการคำนวณเป็นปัจจัยสำคัญ ซึ่งอัลกอริทึม VQ จะต้องใช้เวลาในการคำนวณสูงมาก เพื่อที่จะเพิ่มความเร็วในขั้นตอนการสร้างเวกเตอร์รหัส เราจะ นำเสนอการใช้การปรับปรุง FCM ด้วย SVM ในเรื่องของปัญหาของการแบ่งกลุ่ม SVM ถูก นำมาใช้เพื่อให้เวลาที่ใช้ในการคำนวณที่เหมาะสมที่สุด ประโยชน์ของการนำ SVM มาปรับปรุง FCM คือ เวลาที่ใช้ในการคำนวณที่เหมาะสมที่สุด มันเป็นไปได้ที่จะได้รับเวลาที่ใช้ในการคำนวณเฉลี่ย ดีกว่าอัลกอริทึมแบบสุ่ม (Random algorithm) และในลำดับถัดมาเราใช้ DWT ทำการเพื่อแบ่งแยก แต่ละบล็อกย่อยก่อนออกเป็นกลุ่มที่มีค่าความหนาแน่นเหมือนกันโดยใช้ DWT ทำการจัดลำดับ ความสำคัญของค่าสัมประสิทธิ์สำหรับการค้นหาเวกเตอร์รหัส ซึ่งในการทำขบวนการแบ่งแยกแต่ละ บล็อกย่อยก่อน ผลลัพธ์ทำให้ลดเวลาในการคำนวณของขบวนการที่เหมาะสม

5.3 ระบบการเข้ารหัส (Coding System)

ในช่วงหลายปีที่ผ่านมา เทคนิคการบีบอัดข้อมูลแบบที่มีการสูญเสีย (Lossy Compression) ได้มีการพัฒนาอย่างต่อเนื่องและ VQ ก็เป็นเทคนิคหนึ่งในนั้น เทคนิค VQ ภาพจะถูกแปลงออกเป็นรหัสในรูปของเวกเตอร์ และแต่ละอันของเวกเตอร์จะเป็นการประมาณและเป็นตัวแทนของหนึ่งดัชนี ซึ่งมันก็คือสมาชิกที่อยู่ในตัวเก็บรหัสเฉพาะข้อมูลที่ต้องการใช้ในวิธีนี้ก็คือดัชนีของแต่ละบล็อกของข้อมูลภาพเริ่มต้น ซึ่งขั้นตอนในการพัฒนาระบบการเข้ารหัส VQ แบบทั่วไปจะมีสามส่วน ดังนี้

ส่วนที่ 1 (Part 1) การออกแบบตัวเก็บรหัสทำโดยการใช้ DWT โดยการประยุกต์ใช้เทคนิค DWT ในโดเมนสเปเชียล (Spatial domain) สามารถกำหนดกลุ่มของจุดภาพที่มีค่าความหนาแน่นเหมือนกัน กลุ่มของจุดภาพก็คือตำแหน่งในช่วงความถี่ต่ำ ขณะที่ตำแหน่งของจุดภาพก็คือคำตอบสำคัญในการกำหนดขนาดของแต่ละบล็อก ยิ่งกว่านั้นเราใช้เทคนิค DWT ทำการกระจายแต่ละบล็อกภาพที่เป็นบล็อกย่อยของอินพุตให้เป็นส่วนประกอบของแถบย่อย (Sub-Band) ให้เป็นตัวแทนของเวกเตอร์สามมิติ (Three-Dimensional Vectors) ดังแสดงในรูปที่ 5.3 นอกจากนี้ เราใช้อัลกอริทึม FCM ดังใน [38], [39] เพื่อแบ่งแยกขนาดของเวกเตอร์สามมิติ และเงื่อนไขเริ่มต้นของค่าสมาชิกภาพ (Membership) ของ FCM ถูกออกแบบโดยการใช้ SVM ดังใน [40]-[43] เฉพาะ 30 % ของค่าขนาดของเวกเตอร์สามมิติจะถูกกำหนดเป็นเวกเตอร์สำหรับการฝึก (Training vector, u_i) ซึ่งขนาดของเวกเตอร์สามมิติจะถูกกำหนดให้ 100 % ใช้เป็นข้อมูลอินพุต (Input Data, u_j) ของระบบ ขนาดของเวกเตอร์สามมิติจะถูกแบ่งออกเป็น 8 กลุ่ม ตามควอดแดรนต์ (Quadrants) ของสเปเชียลโคออดิเนต (Spatial Coordinate) ซึ่งก็คือเวกเตอร์รหัสในตัวเก็บรหัส ซึ่งขั้นตอนนี้จะเรียกว่า อัลกอริทึมระบบควบคุมแบบปิด (Close Loop Control, CLC)

โดยทั่วไปแล้ว เพื่อให้เวลาของ FCM ลดลง เงื่อนไขเริ่มต้นของ FCM ซึ่งก็คือค่าสมาชิกภาพ (Membership Value) เป็นปัจจัยที่สำคัญอย่างมากคือ เรานำเสนอการใช้ SVM เพื่อมากำหนดค่าสมาชิกภาพเริ่มต้นให้กับ FCM ในการทำดังกล่าวจุดภาพที่เป็นระดับสีเทา (Gray-Level Pixels) ของภาพอินพุตจะถูกแบ่งออกเป็น 8 กลุ่ม (Quadrants, Q_m) ขบวนการนี้จะเหมือนกับขั้นตอนที่หนึ่งถึงขั้นตอนที่สามของการออกแบบตัวเก็บรหัสในหัวข้อที่สี่ ดังแสดงในตารางที่ 5.1. เราพิจารณากลุ่มที่มีจำนวนสูงสุดของเวกเตอร์สามมิติ ในกลุ่มดังกล่าว เราจะใช้ SVM แบ่งแยกกลุ่มของเวกเตอร์สามมิติหรือจุดภาพที่เป็นระดับสีเทาส่งออกเป็นสองกลุ่ม คือ กลุ่มที่ศูนย์ และ กลุ่มที่หนึ่ง จำนวนกลุ่มของ SVM นี้จะถูกกำหนดเป็นค่าสมาชิกภาพของ FCM

ส่วนที่ 2 (Part 2) การปรับปรุงอัตราบิดทำโดยใช้การเข้ารหัสฮัฟแมน ดังแสดงใน [44], [45] เพื่อลดการซ้ำกันของค่าดัชนี

ส่วนที่ 3 (Part 3) ดังที่ทราบกันแล้วนั้นว่า การบีบอัดข้อมูลแบบ VQ เป็นการบีบอัดข้อมูลแบบที่มีค่าความสูญเสีย เราจึงนำเสนออัลกอริทึมใหม่และแตกต่างโดยการใช้การชดเชยค่าผิดพลาดของระบบ (E_2) อัลกอริทึมดังกล่าวจะลดมิติ (Dimension) ค่าผิดพลาดของระบบ (System Error, E)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยการใช้ระบบการเข้ารหัสการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis, PCA) และเทคนิค DWT เพื่อกระทำการกระจาย (Decomposition) สองระดับ (2-Level) อัลกอริทึมนี้เรียกว่า การชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC) การชดเชยค่าผิดพลาดของระบบจะถูกสร้างขึ้นจากขบวนการเข้ารหัสย้อนกลับ (Reverse Encode Process) ซึ่งประกอบด้วย การแปลงกลับเวฟเลต (Inverse Discrete Wavelet Transform, IDWT) และการถอดรหัส (Decode) PCA ภาพที่สร้างกลับและการชดเชยค่าผิดพลาดของระบบจะถูกนำมารวมกันเพื่อสร้างเป็นภาพเอาต์พุต (X_o) ขบวนการรวมกันดังกล่าวทำโดยการใช้การควบคุมแบบปิด (Closed Loop Control) รวมกับการชดเชยค่าผิดพลาดของระบบ (CLC+SEC) ด้วยเหตุนี้เราจึงเรียกวิธีดังกล่าวว่า CLC+SEC อัลกอริทึม VQ แบบใหม่นี้ ดังแสดงในรูปที่ 5.2

โดยที่ X_i คือ ภาพอินพุต (Input Image)

X_r คือภาพที่สร้างกลับคืน (Reconstructed Image)

I คือดัชนี (Index)

X_o คือภาพเอาต์พุต (Output Image)

E คือค่าผิดพลาดระบบ (System Error)

E_1' คือค่า *Final_Data*

E_2' คือค่าเฉลี่ย (Mean Value)

E_3' คือค่าไอเกนเวกเตอร์ (Eigenvector)

E'' คือค่าสัมประสิทธิ์การประมาณ (Approximation Coefficients Value)

E_1 คือค่า *Final_Data**

E_2 คือค่าการชดเชยค่าผิดพลาดของระบบ (System Error Compensation)

5.4 การออกแบบตัวเก็บรหัส (Codebook design)

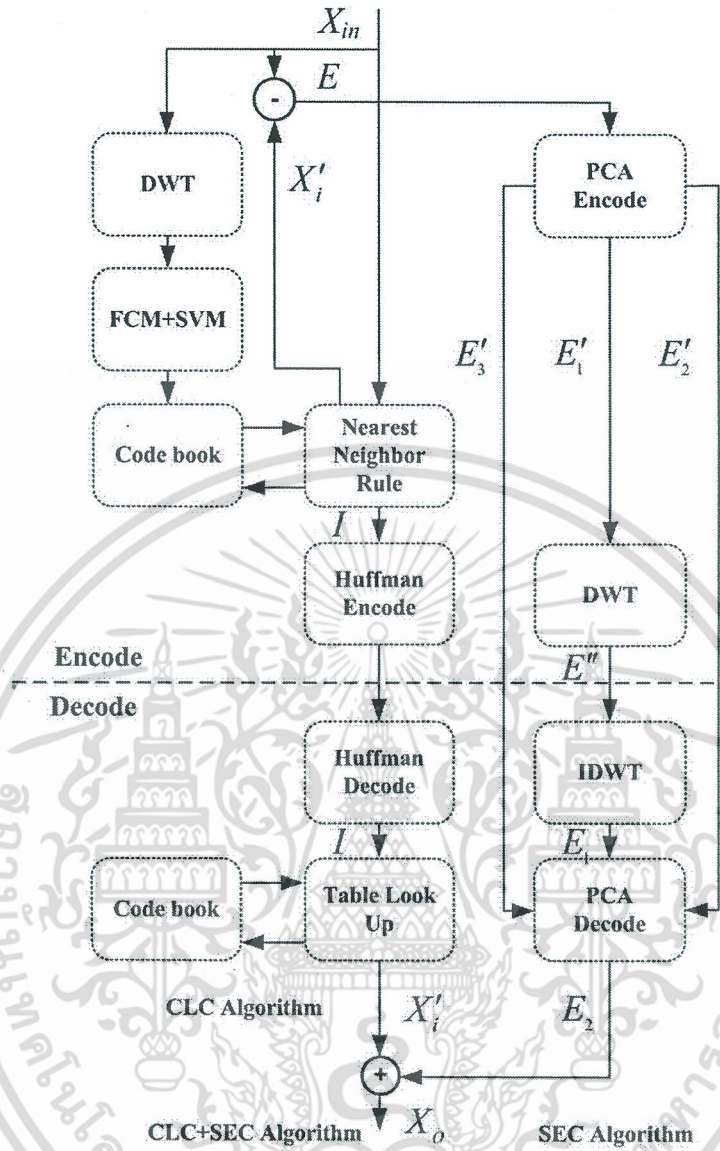
DWT เป็นเทคนิคการแสดงผลที่มีประสิทธิภาพสำหรับการวิเคราะห์เวลาและความถี่และเป็นที่ยอมรับกันอย่างกว้างขวาง คุณลักษณะเฉพาะ (Characteristic) ที่สำคัญอันหนึ่งของการแปลงเวฟเลตคือมีค่าความละเอียด (Resolution) หลายระดับ DWT จะทำการกลับค่าจุดภาพของภาพ ซึ่งจะไม่มีการสูญเสียของข้อมูลในสเปซโดเมน ในบทความวิจัยนี้ DWT ถูกนำมาใช้ในการกำหนดตำแหน่งและขนาดของบล็อกที่ขึ้นกับจุดภาพที่อยู่ใกล้ชิดกัน (Adjacent Pixels) กับค่าความหนาแน่นที่เหมือนกันในการกำหนดภายในจุดภาพของตัวเก็บรหัส (Inter pixel Codebook) ภายในจุดภาพของตัวเก็บรหัสจะประกอบด้วยห้าขั้นตอนหลัก ดังนี้

ขั้นตอนที่ 1 (Step 1) ภาพอินพุต (X_{in}) ขนาด 512×512 จุดภาพถูกแยกออกเป็น $16,384$ ($=128 \times 128$) ไม่มีความสัมพันธ์กัน (Non-Overlapped) ของบล็อกย่อยขนาด 4×4 จุดภาพ แต่ละบล็อกย่อยมีขนาด 4×4 จุดภาพ และกำหนดเป็น Z_1 ถึง $Z_{16,384}$ ดังแสดงในรูปที่ 5.3

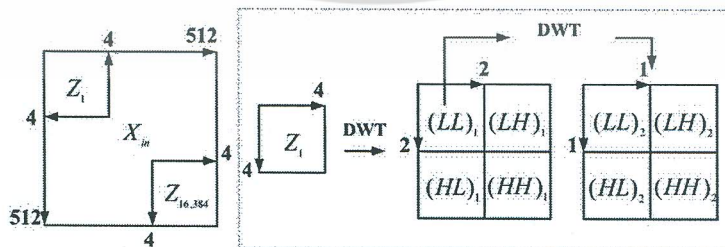
โดยที่ Z แสดงถึงภาพบล็อกย่อยอินพุต (Sub-Input Block) (ภาพอินพุต, X_{in})
 LL_i แสดงถึงค่าสัมประสิทธิ์การประมาณ (Approximation Coefficients)
 LH_i แสดงถึงค่าสัมประสิทธิ์ทางแนวตั้ง (Vertical)
 HL_i แสดงถึงค่าสัมประสิทธิ์ทางแนวนอน (Horizontal)
 HH_i แสดงถึงค่าสัมประสิทธิ์ตามแนวทแยง (Diagonal)
 3D Vector แสดงถึงค่า เวกเตอร์สามมิติ

ขั้นตอนที่ 2 (Step 2) วิเคราะห์คุณลักษณะเฉพาะของแต่ละภาพบล็อกย่อยอินพุตจะถูกวิเคราะห์เพื่อกำหนดเป็นที่อยู่ (Address) ของแต่ละบล็อกย่อยที่สัมพันธ์กับความถี่ของภาพอินพุต เพื่อที่จะกำเนิดการตอบสนองการแปลงค่าสัมประสิทธิ์เราใช้การแปลงสองระดับ (Two-Level) ของ DWT ด้วยการใช้พารามิเตอร์ (parameter) เวฟเล็ต (wavelet) ชื่อ "Haar" แต่ละบล็อกย่อยจะถูกกระจาย (Decomposed) ออกเป็นสองระดับ คือ LL_1, LL_2, LH_2, HL_2 และ HH_2 ดังแสดงในรูปที่ 5.3 เวกเตอร์สามมิติ ($B_1, B_2, B_3, \dots, B_{16,384}$) ถูกสร้างโดยการใช้ค่าสัมประสิทธิ์ของแถบย่อย (Sub-Bands) LH_2, HL_2 และ HH_2 เพื่อกำหนดเป็นเวกเตอร์หนึ่งหน่วย (Unit Vectors) ตามแนวแกน i, j และ k ดังแสดงในรูปที่ 5.4

$$X_{in} = \begin{pmatrix} Z_1 & Z_2 & \dots & Z_{128} \\ Z_{129} & Z_{130} & \dots & Z_{256} \\ \dots & \dots & \dots & \dots \\ Z_{16,257} & Z_{16,258} & \dots & Z_{16,384} \end{pmatrix} \quad (5.1)$$



รูปที่ 5.2 การนำอัลกอริทึม SEC มาชดเชยอัลกอริทึม CLC



- * (discard $(LH)_i, (HL)_i, (HH)_i, (LL)_i$)
- * 3D vector, $\vec{B}_n = (LH)_i \vec{i} + (HL)_i \vec{j} + (HH)_i \vec{k}$
- * $n=1, 2, 3, \dots, 16,384$

รูปที่ 5.3 การนำเทคนิค DWT มาสร้าง 3D vector

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Z_n = \begin{pmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,4} \\ X_{2,1} & X_{2,2} & \dots & X_{2,4} \\ \dots & \dots & \dots & \dots \\ X_{4,1} & X_{4,2} & \dots & X_{4,4} \end{pmatrix} \quad (5.2)$$

โดยที่ X แสดงค่าจุดภาพที่มีระดับเทา (Gray-Level Pixel) และ $n = 1, 2, 3, \dots, 16,384$

- กำหนดเวกเตอร์สามมิติสำหรับภาพบล็อกอินพุต ($Z_1, Z_2, Z_3, \dots, Z_{16,384}$) แสดงดังนี้

$$3Dvector_n = \vec{B}_n = (LH)_2 \vec{i} + (HL)_2 \vec{j} + (HH)_2 \vec{k} \quad (5.3)$$

- กำหนดค่าขนาดของแต่ละเวกเตอร์ ($B_1, B_2, B_3, \dots, B_{16,384}$) ของเวกเตอร์สามมิติ ซึ่งสามารถอธิบายได้ ดังนี้

ค่ามุมหรือค่าขนาดตรีโกณมิติของเวกเตอร์

$$B_n = \|\vec{B}_n\| = \sqrt{((LH)_2)^2 + ((HL)_2)^2 + ((HH)_2)^2} \quad (5.4)$$

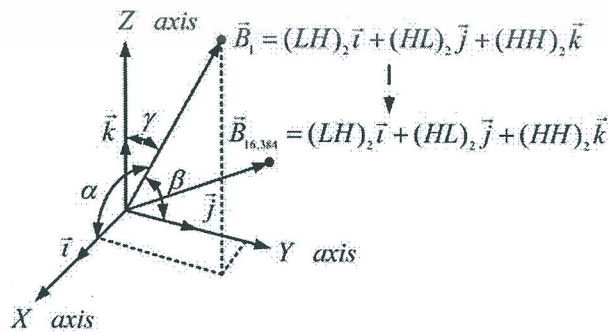
มุมระหว่างตรีโกณมิติของเวกเตอร์และแนวแกน

$$\alpha_n = \arccos\left(\frac{((LH)_2)}{\|\vec{B}_n\|}\right) \quad (5.5)$$

$$\beta_n = \arccos\left(\frac{((HL)_2)}{\|\vec{B}_n\|}\right) \quad (5.6)$$

$$\gamma_n = \arccos\left(\frac{((HH)_2)}{\|\vec{B}_n\|}\right) \quad (5.7)$$

ในการใช้งานเราไม่ใช้ค่าพารามิเตอร์ $\alpha_n, \beta_n, \gamma_n$ และ $\|\cdot\|$ คือ “operator norm”



รูปที่ 5.4 มุมและขนาดของการเวกเตอร์สามมิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 3 (Step 3) บล็อกภาพอินพุตย่อยทั้งหมดจะถูกแบ่งออกเป็นแปดควอดแดรนต์ (Quadrants, Q_m) ตามเครื่องหมายที่แสดงในตารางที่ 5.1 ในแปดครอดแรรันจะรวมเป็นจำนวนทั้งหมดของค่าขนาดของเวกเตอร์สามมิติ เราเรียกมันว่าจำนวนทั้งหมดของค่าขนาดของเวกเตอร์สามมิติ (Total - Group, T) (5.8) ดังแสดงในตารางที่ 5.2

ในแต่ละควอดแดรนต์จะประกอบด้วยจำนวนของค่าขนาดของเวกเตอร์สามมิติ ซึ่งคือค่าขนาดของเวกเตอร์สามมิติก่อนการแบ่งแยกกลุ่มโดยการใช้ FCM + SVM ซึ่งเราจะเรียกว่า “Before FCM - SVM ($Big_Group(m)$)” ดังแสดงใน (5.9) ขนาดของเวกเตอร์สามมิติหลังจากที่ถูกแบ่งแยก เราเรียกว่า “After FCM + SVM ($Small_Group(h)$)” ดังแสดงในตารางที่ 5.2 หลังจากนั้นเวกเตอร์รหัสของตัวเก็บรหัสจะถูกสร้างโดยการใช้ (5.11)

$$Total_group, T = \begin{matrix} \hat{B}_1 \\ \hat{B}_2 \\ \hat{B}_3 \\ \vdots \\ \hat{B}_8 \end{matrix} \quad (5.8)$$

ที่เครื่องหมายเหมือนกันของเวกเตอร์สามมิติ

$$Big_group, \hat{B}_m = \begin{matrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{16,384} \end{matrix} \quad (5.9)$$

ตารางที่ 5.1 เวกเตอร์สามมิติ 8 แบบ

Quadrant (Q_m) $m=1, 2, 3, \dots, 8$	3D vector (\vec{B}_n) $n=1, 2, 3, \dots, 16,384$
1	-i, -j, -k
2	-i, -j, +k
3	-i, +j, -k
4	-i, +j, +k
5	+i, -j, -k
6	+i, -j, +k
7	+i, +j, -k
8	+i, +j, +k

ตารางที่ 5.2 ตัวอย่างการออกแบบตัวเก็บรหัส

Address	3D Vector	Total number of length 3D vector (Total_group)		Code word(C) $C(N) = \frac{1}{k} \sum_{j=1}^k \bar{X}_{i,j}(j); N = 2^b$ $j = 1, 2, \dots, 16.$ $b = \text{bit}, k = h.$
		Before FCM+SVM (Big_group(m))	After FCM+SVM (Small_group(h))	
1	-, -, -	$m = 8;$ $B_1, B_2, B_3, B_4,$ $B_5, B_6, B_7, B_8.$	$h = 4;$ $B_1, B_2, B_3, B_4.$	$B_1(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_2(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_3(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_4(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(1) = \text{Code word}(1)$
			$h = 4;$ $B_5, B_6, B_7, B_8.$	$B_5(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_6(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_7(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_8(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(2) = \text{Code word}(2)$
2	-, -, +	$m = 1; B_9.$	$h = 1; B_9.$	$B_9(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(3) = \text{Code word}(3)$
3	-, +, -	$m = 1; B_{10}.$	$h = 1; B_{10}.$	$B_{10}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(4) = \text{Code word}(4)$
4	-, +, +	$m = 1; B_{11}.$	$h = 1; B_{11}.$	$B_{11}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(5) = \text{Code word}(5)$
5	+, -, -	$m = 4;$ $B_{12}, B_{13}, B_{14}, B_{15}.$	$h = 2;$ $B_{12}, B_{13}.$	$B_{12}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_{13}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(6) = \text{Code word}(6)$
			$h = 2;$ $B_{14}, B_{15}.$	$B_{14}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_{15}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(7) = \text{Code word}(7)$
6	+, -, +	$m = 0.$	-	-
7	+, +, -	$m = 0.$	-	-
8	+, +, +	$m = 2;$ $B_{16}, B_{17}.$	$h = 2;$ $B_{16}, B_{17}.$	$B_{16}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $B_{17}(1) \leftrightarrow \bar{X}_{1,1} \bar{X}_{1,2} \dots \bar{X}_{4,16}$ $C(8) = \text{Code word}(8)$

โดยที่ $S_n = |\hat{X}_{1,1}, \hat{X}_{1,2}, \hat{X}_{1,3}, \dots, \hat{X}_{4,16}|$

$\hat{X}_{i,j}$ แสดงถึงค่าระดับจุดภาพสีเทา

Before FCM+SVM: ที่จำนวนของค่าขนาดของเวกเตอร์สามมิติ (\hat{B}_m) มันจะถูกแบ่งแยกกลุ่มโดยใช้ FCM + SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มจุดศูนย์กลาง (Cluster Center, V_i)

$$V_i = \frac{\sum_{k=1}^n \hat{B}_{m,k} (u_{i,k})^w}{\sum_{k=1}^n (u_{i,k})^w}; 1 \leq i \leq c \quad (5.10)$$

โดยที่ $\hat{B}_{m,k}$ แสดงถึงค่าระดับจุดภาพสีเทาที่มีเครื่องหมายเหมือนกัน และเป็นจำนวนที่ใหญ่ของค่าขนาดของเวกเตอร์สามมิติ

$u_{i,k}$ แสดงถึงเงื่อนไขเริ่มต้น (Initial Condition) ถูกออกแบบโดยใช้ SVM

c แสดงถึงกลุ่ม ($2 < c < n$)

w แสดงถึงค่าน้ำหนักคงที่ (Weigh Constant)

n แสดงถึงจำนวนของเวกเตอร์สามมิติ

After FCM+SVM or *Small_Group(h)*:

เวกเตอร์รหัส (Code Word or Code Vector, $C(N)$)

$$C(N) = \frac{1}{k} \sum_{j=1}^k (\bar{X}_{i,j}(j)); N = 2^b \quad (5.11)$$

โดยที่ $\bar{X}_{i,j}(j)$ แสดงถึงค่าระดับจุดภาพสีเทาหลังการแบ่งแยกกลุ่ม

n แสดงถึง จำนวนของเวกเตอร์รหัส

b แสดงถึง จำนวนบิต

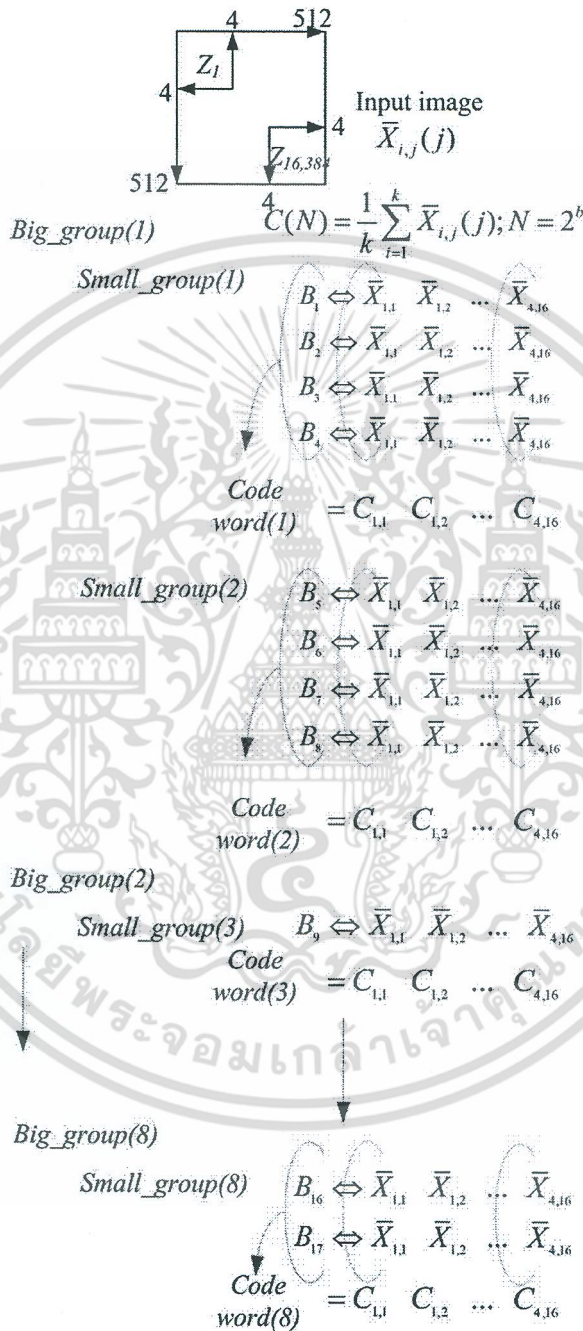
$k = h$ แสดงถึง จำนวนมิติ ($j = 1, 2, 3, \dots, 16.$)

ขั้นตอนที่ 4 (Step 4) เวกเตอร์รหัสจะถูกกำหนดโดยการประยุกต์ใช้อัลกอริทึม FCM เพื่อแบ่งแยกค่าสมาชิกภาพ (Membership) ใน *Big_group(m)* ออกเป็น *Small_group(h)* ในแต่ละ *Small_group(h)* พารามิเตอร์ h จะแสดงถึงตำแหน่งที่อยู่ของแต่ละบล็อกภาพอินพุตย่อย รวมถึงเราประยุกต์ใช้ความน่าจะเป็นของค่าสมาชิกภาพเริ่มต้นโดยการใช้ SVM โดยที่ SVM ถูกนำมาใช้แยกข้อมูล *Big_group(m)* ออกเป็นสองกลุ่มที่มีค่าเป็นบวก (+) และค่าเป็นลบ (-) โดยการสร้างเส้นไฮเปอร์เพลน (Hyper Plane) ที่มีช่องว่าง (Margin) ใหญ่ที่สุด ในงานวิจัยนี้เราเลือกใช้ฟังก์ชันชื่อ "Gaussian Radial Basis Function, RBF" เพื่อทำการแบ่งแยกข้อมูล เพราะว่ามันสามารถตรวจจับกรณีที่อินพุตมีความไม่เป็นเชิงเส้น (Nonlinear) ได้ดี และมีค่าพารามิเตอร์เพียงสองถึงสามตัวเท่านั้น เมื่อเปรียบเทียบกับฟังก์ชันชนิดอื่น ๆ

อินเตอร์พิกเซลตัวเก็บรหัส (Inter Pixel Codebook) สามารถนำเสนอได้ดังแสดงในรูปที่ 5.5 ดังนั้นจาก (5.11) ถ้าจำนวนของบิตที่ถูกใช้เท่ากับแปด จำนวนของเวกเตอร์รหัสจะมีค่าเท่ากับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

256 ในการเข้ารหัสภาพเป้าหมายที่ต้องการคือได้รับค่า PSNR ที่สูงที่อัตราบิตที่ต่ำ ไม่เฉพาะคุณภาพของภาพเอาท์พุทที่ถูกปรับปรุงโดยการใช้อินเตอร์พิกเซลตัวเก็บรหัสและอัลกอริทึม SEC แต่อัตราบิตจะถูกลดลงโดยการใช้การเข้ารหัสแบบฮัฟแมน การเข้ารหัสแบบฮัฟแมนจะถูกกล่าวถึงในลำดับถัดไป



รูปที่ 5.5 ตัวอย่างการออกแบบอินเตอร์พิกเซลตัวเก็บรหัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.5 การปรับปรุงอัตราบิต (Improve Bit Rate)

การปรับปรุงอัตราบิตสามารถทำได้หลายวิธี อัลกอริทึมหนึ่งคือการลดการซ้ำกันของค่าดัชนีโดยใช้การเข้ารหัสแบบไม่มีค่าความสูญเสีย (Lossless Coding) ดังเช่นการเข้ารหัสแบบฮัฟแมน [46] การเข้ารหัสแบบฮัฟแมนถูกพัฒนาโดย David A. Huffman ในปี 1952 เป็นอัลกอริทึมการเข้ารหัสเอนโทรปี (Entropy) ใช้สำหรับการบีบอัดข้อมูลแบบไม่มีค่าความสูญเสีย ซึ่งวิธีนี้ถูกนำมาใช้หลาย ๆ การใช้งานทางการค้า เช่น โทรสาร (Fax Machines) JPEG และ MPEG

5.6 การชดเชยค่าผิดพลาดของระบบ (System Error Compensation, SEC)

เป้าหมายของการใช้อัลกอริทึม SEC ในขบวนการเข้ารหัส คือให้ได้รับภาพที่มีค่า PSNR สูง อัลกอริทึม SEC ประกอบด้วย สองขั้นตอนหลัก ดังนี้

ขั้นตอนที่ 1 (Step 1)

ภาพที่สร้างกลับ (X'_i) จะถูกสร้างโดยการใช้ความสัมพันธ์ระหว่างตัวเก็บรหัสและกฎความสัมพันธ์ใกล้เคียง (Nearest Neighbor Rule) ซึ่งนั่นก็คือการเข้ารหัสภาพอินพุตย่อยระดับสี่เทา (Z_i) ของ HxV จุดภาพ โดยที่ H จะแสดงถึงแกนอน (Horizontal axes) และ V จะแสดงถึงแกนตั้ง (Vertical axes) และผลของบล็อกภาพอินพุตย่อยจะถูกแบ่งออกเป็นบล็อกเล็กที่ไม่มีความสัมพันธ์กันขนาด $v * v$ จุดภาพ โดยที่ $v * v$ จะมีขนาดบล็อกเท่ากับ $4 * 4$

$$Z = \{Z_i | i = 1, 2, 3, \dots, (H * V) / (v * v)\} \quad (5.12)$$

ตัวเก็บรหัสจะถูกนำมาใช้ในการเข้ารหัสแต่ละบล็อก แต่ละบล็อกอินพุตย่อยเราจะคำนวณระยะทางระหว่างบล็อกอินพุตย่อยและแต่ละเวกเตอร์รหัส (C_j) ที่อยู่ในตัวเก็บรหัส โดยการวัดระยะทางยูคลิดเนียน (Euclidean Distance, d) ด้วย (5.13) ดัชนี (i_j) ของเวกเตอร์รหัสจะถูกเลือกเข้ารหัสโดยพิจารณาที่ค่าผิดพลาดของระยะทางต่ำที่สุดเพื่อเข้ารหัสของแต่ละบล็อกอินพุตย่อยดังกล่าว

$$d(Z_i, C_j) = \|Z_{i,l} - C_{j,l}\|^2 = \min_{j=1}^{v*v} (\sum_{l=1}^{v*v} (Z_{i,l} - C_{j,l})^2) \quad (5.13)$$

โดยที่ u แสดงถึงขนาดของตัวเก็บรหัส

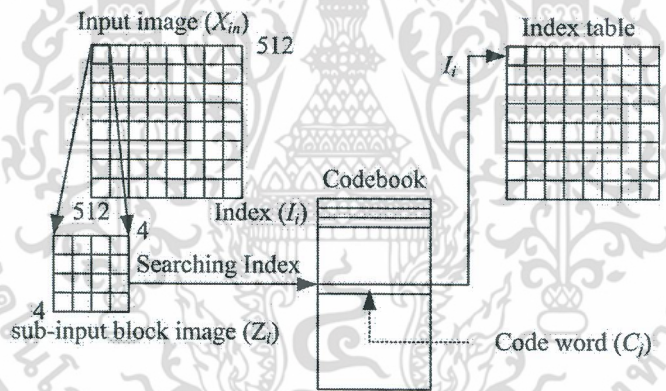
$C_{j,l}$ แสดงถึงค่าจุดภาพของส่วนประกอบลำดับที่ $l - th$ ที่ตำแหน่ง $j - th$ ของเวกเตอร์รหัสในตัวเก็บรหัส

บล็อกอินพุตย่อย คือค่าของจุดภาพลำดับที่ $l - th$ ที่ตำแหน่ง $j - th$ ที่ครอบคลุมบล็อกแผนผังกระบวนการเข้ารหัสดังแสดงในรูปที่ 5.6 เพื่อที่จะสร้างภาพกลับคืน (Reconstructed Image) กระบวนการถอดรหัสจำเป็นต้องมีขั้นตอนที่ตรงกันข้ามกับกระบวนการเข้ารหัส กระบวนการถอดรหัสดังแสดงในรูปที่ 5.7

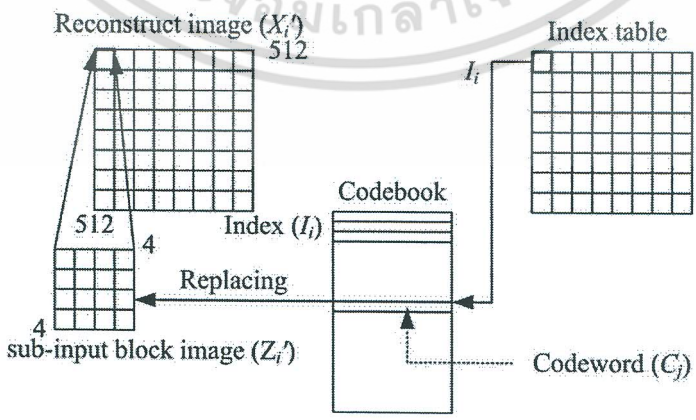
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 (Step 2)

ค่าผิดพลาดของระบบ (System Error) (5.14) คือ ภาพผิดพลาด (Error Image) ระหว่างภาพที่สร้างกลับ (Reconstructed Image) และภาพอินพุต (Input Image) เพื่อนำไปปรับปรุงคุณภาพของภาพเอาต์พุต (Output Image) เพื่อที่จะลดค่าผิดพลาดของระบบหรือเพิ่มค่า PSNR ให้สูงขึ้น ระบบ VQ จึงต้องการอัลกอริทึม SEC เพื่อใช้สำหรับการชดเชยค่าผิดพลาดของระบบ ใน SEC การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis, PCA) ถูกนำมาใช้เป็นอันดับแรกเพื่อทำหน้าที่ลดมิติของค่าผิดพลาดของระบบลง 50 % ดังสมการ (5.15) และทำการกระจายโดยใช้ DWT ออกเป็นสี่แถบย่อย ($LL_1, LH_1, HL_1,$ และ HH_1) เฉพาะแถบย่อย LL_1 จะถูกกระจายออกเป็นสี่แถบย่อย ($LL_2, LH_2, HL_2,$ และ HH_2) (5.16) ซึ่งเฉพาะแถบย่อย LL_2 จะถูกส่งออกไปที่ระบบการถอดรหัส เพื่อทำการชดเชยค่าผิดพลาดของระบบ เพื่อที่จะชดเชยค่าผิดพลาดระบบ การแปลงกลับเวฟเล็ตแบบเต็มหน่วยสองระดับ (Inverse two - Level Discrete Wavelet Transform, IDWT) (5.17) จะถูกถอดรหัสด้วย PCA (5.18) เพื่อสร้างค่าผิดพลาดของระบบกลับคืน บวกภาพที่สร้างขึ้นกับภาพเอาต์พุต (X_o) (5.19) ค่าผิดพลาดของระบบและการชดเชยค่าผิดพลาดของระบบมีการคำนวณได้ดังสมการ (5.14) และ (5.18)



รูปที่ 5.6 บล็อกการเข้ารหัส



รูปที่ 5.7 บล็อกการถอดรหัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$E = X_n - X'_i \tag{5.14}$$

$$(E'_1, E'_2, E'_3, \dots) = PCA_Encode(E) \tag{5.15}$$

และแต่ละภาพจะถูกคำนวณตั้งสมการ

$$E'' = DWT(E'_1) \tag{5.16}$$

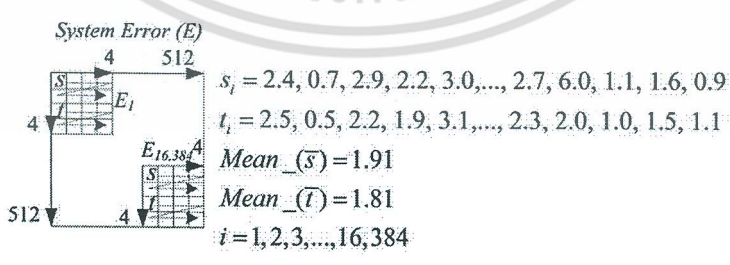
$$E_1 = IDWT(E'') \tag{5.17}$$

$$E_2 = PCA_Decode(E_1, E'_2, E'_3) \tag{5.18}$$

$$X_o = X'_i + E_2; \quad E_2 \neq E_1 \tag{5.19}$$

โดยที่ E'_1 แสดงถึงค่า *Final_Data* ได้มาจากการเข้ารหัส PCA
 E'' แสดงถึงค่าสัมประสิทธิ์จากการแปลงสองระดับของ DWT
 E_1 แสดงถึงค่า *Final_Data** ได้มาจากการแปลงกลับสองระดับ IDWT
 E_2 แสดงถึงค่าการชดเชยค่าผิดพลาดระบบได้มาจากการถอดรหัส PCA
 X_o แสดงถึงค่าภาพเอาต์พุตหลังจากมีการชดเชยค่าผิดพลาดของระบบ

การเข้ารหัสและการถอดรหัส PCA ในส่วนนี้เราจะแสดงการเข้ารหัสและการถอดรหัส
 อย่างไม่ถูกนำมาใช้ในการสร้างค่าการชดเชยค่าผิดพลาดของระบบ สิ่งแรกการเข้ารหัส PCA จะสร้าง
 ค่า *Final_Data* ค่าเฉลี่ย (Mean Value) และค่าไอเกนเวกเตอร์ (Eigenvector Value) เฉพาะค่า
Final_Data จะถูกกระจายออกเป็นค่าสัมประสิทธิ์การประมาณโดยใช้การแปลงสองระดับ DWT ค่า
*Final_Data** จะถูกรวม (Composed) โดยการใช้อัลกอริทึมการแปลงกลับสองระดับ IDWT ค่าสุดท้ายของ
 การชดเชยค่าผิดพลาดของระบบจะถูกสร้างกลับโดยการใช้ค่าเฉลี่ย ค่าไอเกนเวกเตอร์และค่า
*Final_Data** ที่ขั้นตอนของการถอดรหัส อัลกอริทึมการเข้ารหัส PCA ดังรูปที่ 5.8 กระบวนการ
 เข้ารหัสมีดังนี้



รูปที่ 5.8 การเข้ารหัส PCA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(i) หากจุดศูนย์กลาง (Centering) ค่า s และค่า t ทั้งหมดจะถูกลบด้วยค่าเฉลี่ย \bar{s} และค่าเฉลี่ย \bar{t}
โดยที่ \bar{s} และ \bar{t} จะแสดงถึง คือ 8 บิตระดับสี่เทาของค่าผิดพลาดของระบบ

$$X_o = X'_i + E_2; \quad E_2 \neq E_1 \quad (5.20)$$

Data_Adjust แสดงถึงค่าการปรับแต่งของค่าผิดพลาดระบบ

s_i แสดงถึง ค่าสมาชิกภาพของการปรับแต่งค่าในแกน x

t_i แสดงถึง ค่าสมาชิกภาพของการปรับแต่งค่าในแกน y

x, y แสดงถึงค่าสมาชิกภาพของการปรับแต่งข้อมูลในแกน x และแกน y ตามลำดับ

(ii) การทำให้สมบูรณ์ (Whitening) คำนวณหาค่าเวกเตอร์หลัก (Basis Vector) (*Feature Vector*) ค่าไอเกนแวลู (Eigenvalue) ค่าไอเกนเวกเตอร์ ของเมตริกซ์แปรปรวน (Covariance Matrix, *CM*)

$$X_o = X'_i + E_2; \quad E_2 \neq E_1 \quad (5.21)$$

$$FeatureVector = |eigenvalues, eigenvectors| = eig|CM| \quad (5.22)$$

โดยที่ *eig* แสดงถึงการคำนวณค่าไอเกนแวลูและไอเกนเวกเตอร์ที่มีค่าไอเกนแวลูสูง

ตัวอย่าง *Feature Vector* หรือ เวกเตอร์หลักของเมตริกซ์แปรปรวนดังนี้

Eigenvalues		eigenvectors	
0.0491	0.0000	-0.7352	0.6779
0.0000	1.2840	0.6779	0.7352

$$FeatureVector = \begin{vmatrix} 0.0000 & 0.6779 \\ 0.0000 & 0.7352 \end{vmatrix}$$

รูปที่ 5.9 การออกแบบ "basis vector"

โดยที่ *Feature Vector* แสดงถึงเมตริกซ์ของไอเกนเวกเตอร์ที่มีค่าไอเกนแวลูสูง

(iii) การลดมิติ (Dimensional Reduction) ในขั้นตอนนี้มีมิติของตัวแปร *RowData_Adjust* จะถูกลดลงหนึ่งมิติดังแสดงในสมการ (5.25)

$$RowFeatureVector = FeatureVector^T \quad (5.23)$$

$$RowData_Adjust = Data_Adjust^T \quad (5.24)$$

$$Final_Data(E'_1) = RowFeatureVector * RowData_Adjust \quad (5.25)$$

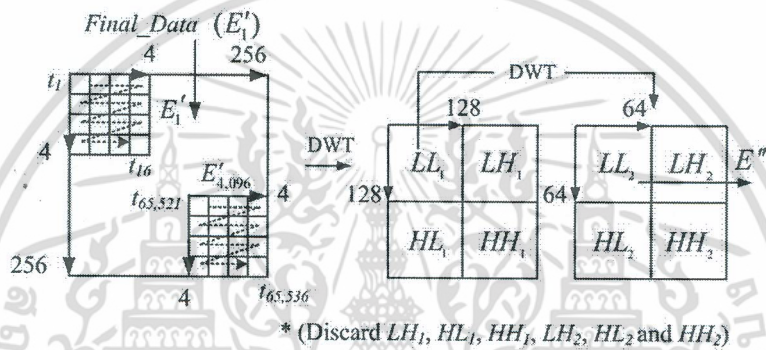
โดยที่ T แสดงถึงการทรานโพสของเมตริกซ์ (Transpose of matrix)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

RowFewtureVector แสดงถึงการทรานโพสของค่าตัวแปร FewtureVector
 RowData_Adjust แสดงถึงการทรานโพสของค่าตัวแปร Data-Adjust
 Final_Data คือผลจากการคูณกันระหว่างสมการ (23) และ (24)

Final_Data ในตัวอย่างนี้เราจะไม่คิดค่าที่มีความสำคัญต่ำเพื่อที่จะลดมิติ ดังแสดงในสมการ (5.26)

$$\begin{matrix} S_i = 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & \dots & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ E'_1 = 0.83 & -1.78 & 0.99 & 0.27 & 1.68 & \dots & 0.91 & -0.10 & -1.15 & -0.44 & -1.22 \end{matrix} \quad (5.26)$$

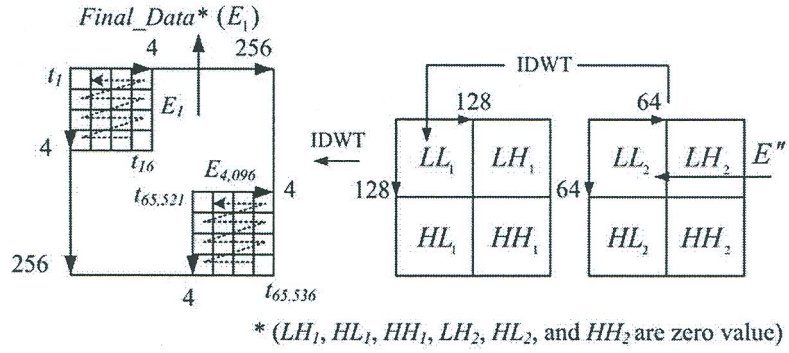


รูปที่ 5.10 การเข้ารหัสการแปลงสองระดับ DWT

ในรูปที่ 5.10 เราแสดงการจัดเรียงค่าตัวแปร Final_Data บล็อกสี่เหลี่ยมขนาด 256x256 จุดภาพ การแปลงเวฟเลต “Haar” ถูกนำมาใช้กับ Final_Data เพื่อที่จะคำนวณค่าสัมประสิทธิ์ของเมตริกซ์ LL₁, LL₂, และรายละเอียดของสัมประสิทธิ์ (Details coefficients matrix) LH₁, LH₁, LH₂, LH₂ และ HH₁, HH₂ (แกนนอน แกนตั้ง แนวแกนทแยง)

อัลกอริทึม ดังแสดงในรูปที่ 5.11 ขั้นตอนของการถอดรหัสสามารถสรุปได้ดังนี้

- (i) แปลงกลับเวฟเลต (Invers Discrete Wavelet Transform, IDWT) ในขั้นตอนแรกการแปลงกลับเวฟเลตสองระดับถูกนำมาใช้ เพื่อสร้างข้อมูลกลับคืนดังแสดงในรูปที่ 5.11

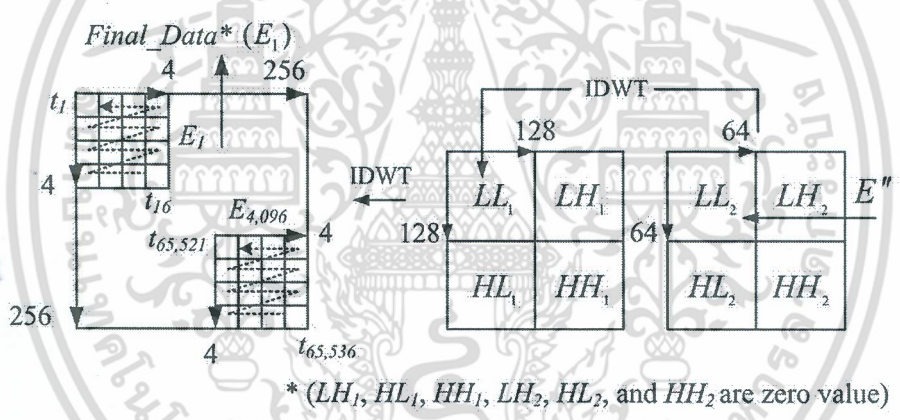


รูปที่ 5.11 การถอดรหัสการแปลงกลับสองระดับ IDWT

(ii) ข้อมูลภาพฉาย PCA (PCA Projection Data) การแปลงข้อมูลกลับ คือภาพฉายบน

RowFeatureVector ดังแสดงในสมการ (5.27) และดังแสดงในรูปที่ 5.12

$$PCA + projection(E_3) = RowFeatureVector^T * Final_Data \tag{5.27}$$



รูปที่ 5.12 ข้อมูลภาพฉาย PCA

โดยที่ E₃ แสดงถึงข้อมูลภาพฉาย PCA

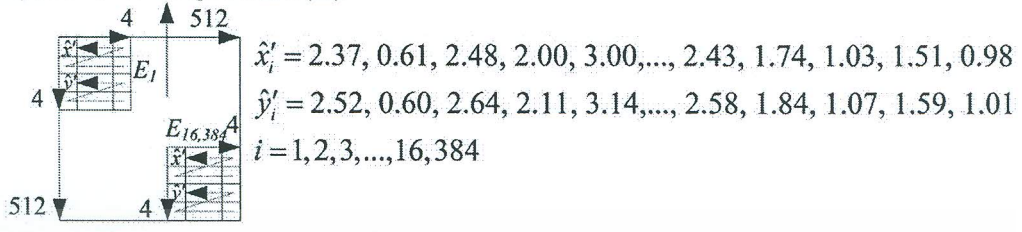
(iii) การถอดรหัส PCA (PCA Decode) การชดเชยค่าผิดพลาดของระบบได้มาจากการใช้สมการดังนี้

$$System\ Error\ Compensation = \begin{vmatrix} s'_i + \bar{s} \\ t'_i + \bar{t} \end{vmatrix} = \begin{vmatrix} \hat{x}' \\ \hat{y}' \end{vmatrix} \tag{5.28}$$

ในรูปที่ 5.13 เราแสดงการคำนวณการชดเชยค่าผิดพลาดของระบบจากข้อมูลภาพฉาย PCA รวมกับค่าเฉลี่ย \bar{s} และ \bar{t} ในการทดลองค่าผิดพลาดของระบบ คือภาพที่แตกต่างระหว่างภาพที่สร้างกลับ และภาพอินพุต (14) ซึ่งค่าดังกล่าวนี้มีค่าต่ำมาก ดังนั้นขบวนการเข้ารหัสและการถอดรหัส

ของค่าผิดพลาดระบบจะใช้จำนวนบิตเท่ากับ 4 บิต ซึ่งมันเพียงพอสำหรับการควอนไทซ์และการ ดี-ควอนไทซ์ (Quantization and de-quantization)

System Error Compensation (E_2)



รูปที่ 5.13 การถอดรหัส PCA



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6

ผลการจำลอง

ในบทนี้จะเป็นการกล่าวถึงผลการจำลองระบบการบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์และการชดเชยค่าผิดพลาดของระบบ โดยจะแบ่งผลการจำลองระบบออกเป็นสิบสองส่วนดังนี้คือ ส่วนที่หนึ่ง กล่าวถึงผลการจำลองค่าเฉลี่ยเวลาในการคำนวณ ส่วนที่สอง กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC ส่วนที่สาม กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC ส่วนที่สี่ กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับสี่ภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC ส่วนที่ห้า กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC ส่วนที่หก กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับสี่ภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN ส่วนที่เจ็ด กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM ส่วนที่แปด กล่าวถึงผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC ส่วนที่เก้า กล่าวถึงผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC+SEC ส่วนที่สิบ กล่าวถึงผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับวิธี CLC ส่วนที่สิบเอ็ดผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับวิธี CLC+SEC กล่าวถึง และส่วนที่สอง กล่าวถึงผลการจำลองตัวเก็บรหัสที่ถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC+SEC เปรียบเทียบกับการเข้ารหัสแบบ PCA อย่างเดียว

เงื่อนไขของการทดลอง คือกำหนดให้ใช้ “Intel (R) Core(TM) i5 CPU M520@2.40GHz 2.40 GHz” สำหรับการเปรียบเทียบ เราทดสอบอัลกอริทึมกับภาพสองชนิด คือ ภาพมาตรฐาน “Airplane, Girl, Lena, and Pepper” และภาพทางการแพทย์ “MR Cardiac Tract, MRI Skull, Ultrasound Liver Cyst, MRI nevit and MRI Brain” ภาพทั้งสองชนิดที่ใช้ในการทดสอบมีขนาด 512x512 จุดภาพ และ 256x256 จุดภาพ และขนาดบล็อก 4x4 จุดภาพ ยิ่งกว่านั้นเราทดสอบอัลกอริทึมกับตัวเก็บรหัสสองชนิด ชนิดที่หนึ่ง คือทดสอบกับภาพแต่ละภาพที่แยกกัน และชนิดที่สอง คือทดสอบกับภาพสองภาพ สามภาพ และ สี่ภาพ

อัตราบิตในส่วนนี้ถูกคำนวณดังนี้
$$\text{bpp} = ((\text{CLC}+\text{Huffman})/(512*512))+(((\text{SEC})+(\text{eigenvector})+(\text{mean value}))/((512*512))) = (((512*512*1/(4*4))/(512*512))+(((64*64*4)+(16)+(16)))/(512*512))= 0.13$$

6.1 ผลการจำลองค่าเฉลี่ยเวลาในการคำนวณ

อัลกอริทึม VQ จะมีปัญหาของการแบ่งกลุ่ม (Clustering Problem) ซึ่งจะต้องทำให้เหมาะสมที่สุด เวลาที่ใช้ในการคำนวณเป็นปัจจัยสำคัญ ซึ่งอัลกอริทึม VQ จะต้องใช้เวลาในการคำนวณสูงมาก เพื่อที่จะเพิ่มความเร็วในขั้นตอนการสร้างเวกเตอร์รหัส เราจึงนำเสนอการใช้การปรับปรุง FCM ด้วย SVM ในเรื่องของปัญหาของการแบ่งกลุ่ม SVM ถูกนำมาใช้เพื่อให้เวลาที่ใช้ในการคำนวณที่เหมาะสมที่สุด ประโยชน์ของการนำ SVM มาปรับปรุง FCM คือ เวลาที่ใช้ในการคำนวณที่เหมาะสมที่สุด มันเป็นไปได้ที่จะได้รับเวลาที่ใช้ในการคำนวณเฉลี่ยดีกว่าอัลกอริทึมแบบสุ่ม (Random algorithm) 14.17 %

ดังที่กล่าวมาก่อนหน้านั้น ตัวเก็บรหัสจะถูกออกแบบบนหลักการปรับอัลกอริทึม FCM โดยใช้ SVM วัตถุประสงค์ของการนำ SVM มาปรับปรุงอัลกอริทึม FCM เพื่อกำหนดค่าสมาชิกภาพเริ่มต้นของอัลกอริทึม FCM ดังนั้นปัญหาของเวลาในการแบ่งกลุ่มก็จะลดลง เพื่อแสดงผลของการปรับปรุงอัลกอริทึมของ FCM โดยใช้ SVM เรามีการเปรียบเทียบเวลาในการให้กำเนิดตัวเก็บรหัสระหว่างอัลกอริทึม FCM แบบทั่วไป และแบบการปรับปรุงอัลกอริทึมของ FCM โดยใช้ SVM จำนวนของเวลาที่ใช้ในการทดสอบภาพดังแสดงในตารางที่ 6.1 ในกรณีของการใช้อัลกอริทึม FCM แบบทั่วไปค่าสมาชิกภาพเริ่มต้นจะถูกกำหนดโดยการใช้วิธีแบบสุ่ม (Randomly Set) สำหรับภาพที่ใช้ในการทดสอบมีขนาด 256x256 จุดภาพและขนาดบล็อก 4x4 จุดภาพ เวลาเฉลี่ยของการกำเนิดตัวเก็บรหัสคือ 92.66 วินาทีสำหรับวิธีการใช้อัลกอริทึม FCM โดยใช้ SVM และ 107.96 วินาที สำหรับอัลกอริทึม FCM แบบทั่วไปดังแสดงในตารางที่ 6.1

ตารางที่ 6.1 ค่าเฉลี่ยเวลาในการคำนวณ

Type	Images	SVM (Sec)	Random (Sec)
Non-Medical Images	Airplane	50.46	60.49
	Girl	45.23	60.18
	Lena	46.27	60.24
	Pepper	46.45	60.43
	Four images	102.67	120.19
Medical images	MR Cardiac Tract	40.76	52.00
	MRI Skull	45.66	60.27
	Ultrasound Liver Cyst	46.25	60.40
	Three images	82.65	95.73

6.2 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

ตารางที่ 6.2 จะแสดงค่า PSNR และ MSE ตัวเก็บรหัสถูกออกแบบมาจากแต่ละภาพมาตรฐาน เนื้อหาของการทดสอบสำหรับอัลกอริทึม CLC และ CLC + SEC มีดังนี้ ภาพมีขนาด 512x512 จุดภาพ ขนาดบล็อก คือ 4x4 จุดภาพ และอัตราบิตที่ใช้ในการทดสอบ คือ 0.13 bpp, 0.19 bpp และ 0.25 bpp (ทั้งหกเงื่อนไขในการทดสอบจะไม่ใช้การเข้ารหัสฮัฟแมน)

ในกรณีของอัลกอริทึม CLC และตัวเก็บรหัสถูกสร้างจากภาพมาตรฐาน ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp มีค่าเท่ากับ 22.64 dB 24.63 dB และ 26.30 dB ตามลำดับ สำหรับอัลกอริทึม CLC+SEC ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp มีค่าเท่ากับ 29.99 dB 30.01 dB และ 30.53 dB ตามลำดับ ผลการจำลองดังแสดงในตารางที่ 6.2

ตารางที่ 6.2 ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

Algorithms	Images 512x512 pixels	Bit rate (bpp)					
		0.13		0.19		0.25	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC	Airplane	21.87	422.75	23.65	280.60	25.17	197.73
CLC+SEC		28.06	101.64	28.15	99.56	29.20	78.17
CLC	Girl	23.89	265.51	26.51	145.24	27.66	111.45
CLC+SEC		31.92	41.79	31.94	41.60	31.95	41.50
CLC	Lena	23.20	311.23	25.47	184.54	26.84	134.61
CLC+SEC		31.15	49.90	31.17	49.67	31.35	47.65
CLC	Pepper	21.61	448.83	22.89	334.26	25.54	181.59
CLC+SEC		28.82	85.33	28.84	84.93	29.61	71.14

6.3 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC

ตารางที่ 6.3 จะแสดงค่า PSNR และ MSE ตัวเก็บรหัสถูกออกแบบมาจากแต่ละภาพทางการแพทย์ ดังนี้ MR Cardiac Tract MRI Skull และภาพ Ultrasound Liver Cyst เนื้อหาของการทดสอบสำหรับอัลกอริทึม CLC และ CLC + SEC มีดังนี้ ภาพมีขนาด 512x512 จุดภาพ ขนาดบล็อก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คือ 4x4 จุดภาพ และอัตราบิตที่ใช้ในการทดสอบ คือ 0.13 bpp, 0.19 bpp และ 0.25 bpp (ทั้งหกเงื่อนไขในการทดสอบจะไม่ใช้การเข้ารหัสฮัฟแมน)

ในกรณีของอัลกอริทึม CLC และตัวเก็บรหัสถูกสร้างจากภาพทางการแพทย์ ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.255 bpp มีค่าเท่ากับ 24.44 dB 26.66 dB และ 27.75 dB ตามลำดับ สำหรับอัลกอริทึม CLC + SEC ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp มีค่าเท่ากับ 33.77 dB 33.81 dB และ 34.14 dB ผลการจำลองดังแสดงในตารางที่ 6.3

ตารางที่ 6.3 ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ

CLC+SEC

Algorithms	Images 512x512 pixels	Bit rate (bpp)					
		0.13		0.19		0.25	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC	MR Cardiac Tract	24.11	252.40	26.78	136.48	28.05	101.88
CLC+SEC		35.00	20.56	35.10	20.10	35.27	19.32
CLC	MRI Skull	26.16	157.43	28.19	98.65	28.82	85.33
CLC+SEC		35.79	17.14	35.80	17.10	36.08	16.04
CLC	Ultrasound Liver Cyst	22.93	331.19	25.01	205.15	26.37	150.00
CLC+SEC		30.53	57.56	30.54	57.42	31.07	50.83

6.4 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

ตารางที่ 6.4 แสดงค่า PSNR และ MSE ตัวเก็บรหัสถูกสร้างจากสีภาพมาตรฐานสำหรับอัลกอริทึม CLC และ CLC+SEC ขนาดภาพมีค่าเท่ากับ 512x512 จุดภาพ ขนาดบล็อกมีค่าเท่ากับ 4x4 จุดภาพ และทำการทดสอบที่อัตราบิตเท่ากับ 0.13 bpp 0.19 bpp และ 0.25 bpp ที่อัตราบิตดังกล่าวจะมีการเข้ารหัสด้วยวิธีฮัฟแมน สำหรับภาพมาตรฐานค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp 0.19 bpp และ 0.25 bpp สำหรับอัลกอริทึม CLC มีค่าเท่ากับ 22.64 dB 24.70 dB และ 26.32 dB ตามลำดับ สำหรับอัลกอริทึม CLC+SEC ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp 0.19 bpp และ 0.25 bpp มีค่าเท่ากับ 29.99 dB 30.05 dB และ 30.72 dB ผลการจำลองดังแสดงในตารางที่ 6.4

ตารางที่ 6.4 ตัวเก็บรหัสถูกออกแบบสำหรับสี่ภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

Algorithms	Images 512x512 pixels	Bit rate (bpp)					
		0.13		0.19		0.25	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC CLC+SEC	Airplane	21.29 28.05	483.15 101.88	22.36 28.14	377.64 99.79	22.71 28.92	348.40 83.38
CLC CLC+SEC	Girl	23.30 31.91	304.15 41.89	24.27 31.92	243.27 41.79	25.64 32.04	177.45 40.65
CLC CLC+SEC	Lena	22.54 31.16	362.31 49.78	23.36 31.14	299.97 50.01	23.69 31.51	278.02 45.93
CLC CLC+SEC	Pepper	19.57 28.83	717.93 85.13	21.37 28.89	474.33 83.96	22.03 29.14	407.46 79.27

6.5 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC

สำหรับอัลกอริทึม CLC เมื่อนำไปใช้กับภาพทางการแพทย์ ค่าเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp 0.19 bpp และ 0.25 bpp มีค่าเท่ากับ 24.02 dB 25.31 dB และ 26.05 dB ตามลำดับ และมีค่าเท่ากับ 33.76 dB 33.80 dB และ 34.11 dB สำหรับอัลกอริทึมของ CLC+SEC ผลการจำลองดังแสดงในตารางที่ 6.5

ตารางที่ 6.5 ตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC

Algorithms	Images 512x512 pixels	Bit rate (bpp)					
		0.13		0.19		0.25	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC CLC+SEC	MR Cardiac Tract	23.97 34.99	260.66 20.61	25.40 35.00	187.53 20.56	25.97 35.29	164.47 19.23
CLC CLC+SEC	MRI Skull	25.42 35.78	186.67 17.18	26.84 35.81	134.61 17.06	27.22 36.14	123.33 15.82
CLC CLC+SEC	Ultrasound Liver Cyst	22.68 30.52	350.82 57.69	23.69 30.58	278.02 56.90	24.95 30.91	208.01 52.73

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.6 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN

ตารางที่ 6.6 แสดงผลของ PSNR และ MSE และตัวเก็บรหัสถูกสร้างจากสีภาพมาตรฐานสำหรับอัลกอริทึมของ CLC, CLC+SEC และ LBG+SEC ซึ่งทั้งหมดจะถูกนำมาเปรียบเทียบกับอัลกอริทึมของ WT+GGCLN อัลกอริทึมของ LBG เป็นที่รู้จักกันอย่างมากใน VQ [1] อัลกอริทึมทั้งหมด (CLC CLC+SEC LBG+SEC และ WT+GGCLN) ใช้ขนาดภาพ 256x256 จุดภาพ ขนาดบล็อกมีค่าเท่ากับ 4x4 จุดภาพ และทำการทดสอบที่อัตราบิต 0.44 bpp หรือ 0.47 bpp ที่อัตราบิตดังกล่าวจะไม่มีการเข้ารหัสด้วยวิธีฮัฟแมนที่อัตราบิต 0.44 bpp ค่าเฉลี่ย PSNR มีค่าเท่ากับ 27.29 dB 30.11 dB 28.9 dB และ 29.45 dB สำหรับ CLC CLC+SEC LBG+SEC และ WT+GGCLN ตามลำดับ ค่าเฉลี่ย PSNR ที่ได้รับจากอัลกอริทึม CLC+SEC (0.44 bpp) คือ 3.99 % สูงกว่าวิธีทั่วไป (LBG+SEC, 0.44 bpp) ยิ่งกว่านั้นอัลกอริทึม CLC+SEC (0.44 bpp) สามารถได้รับค่า PSNR 2.19 % สูงกว่าวิธีทั่วไป (WT+GGCLN, 0.47 bpp) อัตราการบีบอัด (Compression Ratio, CR) คือ $[512*512*8 / (512*512*7 / (4*4))] = 18.29$ สำหรับ CLC และ $[512*512*8 / ((512*512*6 / (4*4))+64*64*4)] = 18.29$ สำหรับ CLC+SEC ผลการจำลองดังแสดงในตารางที่ 6.6

ตารางที่ 6.6 ตัวเก็บรหัสถูกออกแบบสำหรับสีภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN

Algorithms	Compression Ratio (CR)	Images 256x256 pixels							
		Airplane		Girl		Lena		Pepper	
		PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC	18.29	25.97	164.47	28.73	87.11	27.61	112.74	26.84	134.61
CLC+SEC	18.29	28.54	91.01	31.68	44.17	30.20	62.10	30.02	64.73
LBG+SEC	18.29	27.34	119.97	30.58	56.90	29.16	78.90	28.56	90.60
WT+GGCLN	16.89	27.66	111.45	31.90	41.98	29.52	72.62	28.71	85.51

6.7 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM

ตารางที่ 6.7 แสดงถึงค่า PSNR และ MSE สำหรับอัลกอริทึมของ CLC CLC+SEC ตัวเก็บรหัสถูกสร้างจากสามภาพทางการแพทย์ MR Cardiac Tract, MRI Skull และ Ultrasound Liver Cyst ผลลัพธ์ถูกนำมาเปรียบเทียบกับอัลกอริทึม WT+HM ขนาดภาพ 256x256 จุดภาพ ขนาดบล็อก 4x4 จุดภาพ และอัตราบิต 0.25 bpp ที่อัตราบิตดังกล่าวจะไม่มีการเข้ารหัสด้วยวิธีฮัฟแมน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึม CLC, CLC+SEC และ LBG+SEC ที่อัตราบิต 0.25 bpp ค่าเฉลี่ย PSNR มีค่าเท่ากับ 26.13 dB 30.58 dB 30.26 dB และ 27.23 dB อัลกอริทึม CLC, CLC+SEC, LBG+SEC และ WT+HM ตามลำดับ ค่าเฉลี่ย PSNR ถูกสร้างโดย CLC+SEC (0.25 bpp) คือ 10.96 % สูงกว่าวิธีโดยทั่วไป (WT+HM, 0.31 bpp) อัลกอริทึมของ CLC และ CLC+SEC ถูกนำมาใช้กับสองภาพทางการแพทย์ (MRI Nevit และ MRI Brain) เจื่อนไขการทดสอบอัลกอริทึมของ CLC และ CLC+SEC คือขนาดภาพมีค่าเท่ากับ 256x256 จุดภาพ ขนาดบล็อกมีค่าเท่ากับ 4x4 จุดภาพและอัตราบิตมีค่าเท่ากับ 0.5 bpp ค่าเฉลี่ย PSNR สำหรับอัลกอริทึม CLC และ CLC+SEC คือ 28.70 dB และ 32.22 dB ตามลำดับ ดังแสดงในรูปที่ 6.1 ในกรณีที่ตัวเก็บรหัสถูกสร้างจากภาพมาตรฐาน (Airplane, Girl, Lena และ Pepper) ค่าเฉลี่ย PSNR ของอัลกอริทึม CLC CLC+SEC มีค่าเท่ากับ 27.60 dB และ 32.02 dB ตามลำดับ สุดท้ายได้นำอัลกอริทึมมาทำการทดสอบกับสามภาพทางการแพทย์ (MR Cardiac Tract, MRI Skull และ Ultrasound Liver Cyst) ซึ่งอัลกอริทึมดังกล่าวนี้ถูกนำมาเปรียบเทียบกับผลของอัลกอริทึม WT+HM ค่าเฉลี่ย PSNR ของอัลกอริทึม CLC และ CLC+SEC มีค่าเท่ากับ 29.01 dB และ 32.92 dB ตามลำดับ ผลการทดสอบดังแสดงในรูปที่ 6.2

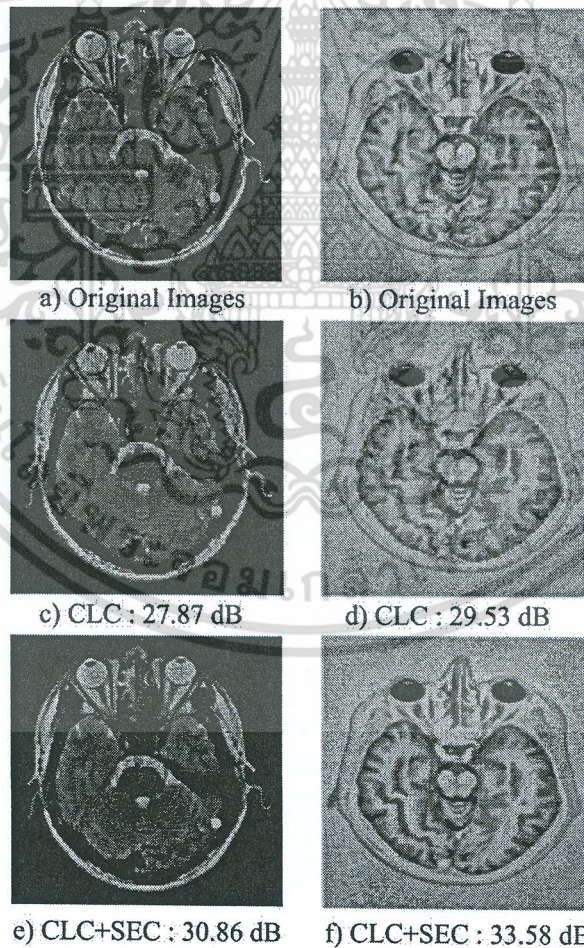


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 6.7 ตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM

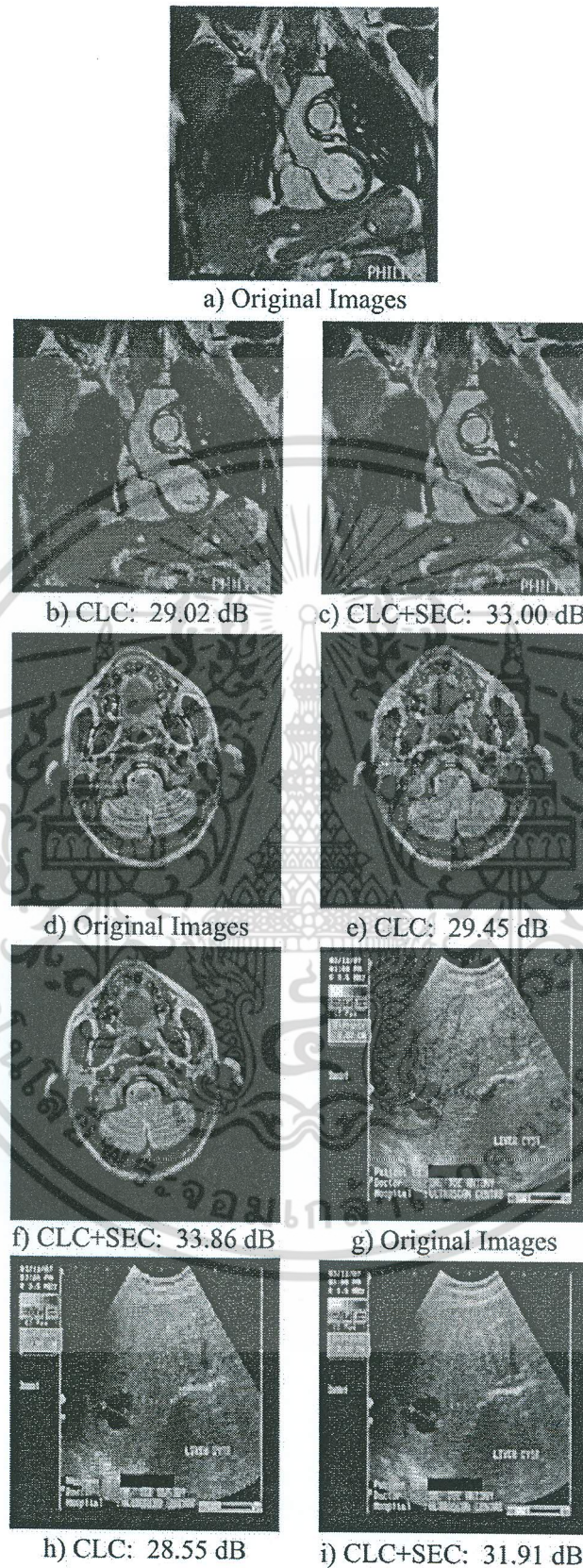
Algorithms	Compression Ratio (CR)	Images					
		MR Cardiac Tract		MRI Skull		Ultrasound Liver Cyst	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
CLC	32.00	25.55	181.17	27.05	128.26	25.37	188.84
CLC+SEC	32.00	30.55	57.29	32.05	40.56	29.19	78.36
LBG+SEC	32.00	30.07	63.99	31.82	42.76	28.88	84.16
WT+HM	CR	27.64	111.96	29.07	80.55	24.97	207.05
		CR = 27.50		CR = 26.89		CR = 23.07	
		S = 256x256		S = 256x256		S = 401x494	

Where: S is the size of image.



รูปที่ 6.1 ทดสอบค่า PSNR สำหรับวิธี CLC, CLC+SEC กับสองภาพทางการแพทย์ ขนาดภาพ 256*256 จุดภาพ อัตราบิต 0.50 bpp

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 ทดสอบค่า PSNR สำหรับวิธี CLC, CLC+SEC กับสามภาพทางการแพทย์ ขนาดภาพ 256*256 จุดภาพ อัตราบิต 0.50 bpp

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.8 ผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมน สำหรับวิธี CLC

ตารางที่ 6.8-6.11 แสดงผลการทดสอบกับภาพมาตรฐาน (Airplane, Girl, Lena และ Pepper) ขนาดภาพ 512x512 จุดภาพและขนาดบล็อก 4x4 จุดภาพ ยิ่งกว่านั้นตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพ อัตราบิตเฉลี่ยคือ $(((512*512*8)/(4*4))/(512*512)) = 0.50$ bpp สำหรับ CLC (ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมน) และ $[107,381.50/(512*512)] = 0.41$ bpp สำหรับ CLC (มีการเข้ารหัสด้วยวิธีฮัฟแมน) ดังนั้น 0.50 bpp ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมน และ 0.41 bpp มีการเข้ารหัสด้วยวิธีฮัฟแมน อัลกอริทึม CLC จะให้ค่า PSNR ที่ 18.07 % และ 18.00 % สูงกว่าวิธีทั่วไป [4]

ตามที่มีการคำนวณของอัตราบิตเฉลี่ย $[(512*512*8)/(512*512)] = 0.5$ bpp ค่าความละเอียด (Resolution) ของ VQ คือ ค่าอัตราบิต $r = \log_2 N/K$ โดยที่ N แสดงถึงจำนวนของเวกเตอร์รหัส K แสดงถึงมิติ Euclidean Space หรือ ขนาดของบล็อก เช่น $16 = 4*4$ ดังนั้นใน 1 ตัวเก็บรหัสเมื่อจำนวนบิตที่ใช้เท่ากับ 8 และขนาดของบล็อกเท่ากับ 4x4 จำนวนเวกเตอร์รหัสจะมีค่าเท่ากับ 256

ตารางที่ 6.8 บิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC

Images 512*512 pixels	Without Huffman Compression (bit)	With Huffman Compression (bit)	Encoding Bits Difference (bit)
Airplane	131,072.00	118,854.00	12,218.00
Girl	131,072.00	109,899.00	21,173.00
Lena	131,072.00	107,571.00	23,501.00
Pepper	131,072.00	93,202.00	37,870.00
Average bit	131,072.00	107,381.50	23,690.50
Average bpp	0.50 bpp	0.41 bpp	0.09 bpp

6.9 ผลการจำลองบิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มี การบีบอัดข้อมูลแบบฮัฟแมน สำหรับวิธี CLC+SEC

ตารางที่ 6.9 แสดงบิตเฉลี่ย $(((8*64*64)+((512*512*8)/(4*4)))/(512*512+(64*64)) = 0.62$ bpp สำหรับอัลกอริทึม CLC+SEC (ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมน) และ $[140,149.5/(512*512)] = 0.54$ bpp สำหรับอัลกอริทึม CLC+SEC (มีการเข้ารหัสด้วยวิธีฮัฟแมน)

ตามลำดับ ในลำดับถัดมาที่อัตราบิต 0.62 bpp (ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมน) และ 0.54 bpp (มีการเข้ารหัสด้วยวิธีฮัฟแมน) เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเข้ารหัสด้วยวิธีฮัฟแมน) อัลกอริทึม CLC+SEC จะให้ค่า PSNR ที่ 14.46 % และ 14.52 % สูงกว่าวิธีทั่วไป [4]

ตารางที่ 6.9 บิตเฉลี่ยและอัตราบิตเฉลี่ยมีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนสำหรับวิธี CLC+SEC

Images 512*512 pixels	Without Huffman Compression (bit)	With Huffman Compression (bit)	Encoding Bits Difference (bit)
Airplane	163,840.00	151,622.00	12,218.00
Girl	163,840.00	142,667.00	21,173.00
Lena	163,840.00	140,339.00	23,501.00
Pepper	163,840.00	125,970.00	37,870.00
Average bit	163,840.00	140,149.50	23,690.50
Average bpp	0.62 bpp	0.54 bpp	0.09 bpp

6.10 ผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมน และปรับค่าอัตราบิตสำหรับวิธี CLC

ตารางที่ 6.10 แสดงบิตเฉลี่ยและ PSNR ของภาพหลังจากมีการเข้ารหัสด้วยวิธีฮัฟแมนกับอัลกอริทึม CLC ซึ่งถูกออกแบบโดยการใช้การปรับขนาดของอัตราบิต (Various Sizes of Bit Rate) ที่อัตราบิต 0.50 bpp บิตเฉลี่ย คือ 107,457.75 บิตค่า PSNR มีค่าเท่ากับ 29.41 dB ที่ 0.44 bpp บิตเฉลี่ยคือ 96,016.75 บิต ค่า PSNR มีค่าเท่ากับ 28.92 dB และที่ 0.38 bpp บิตเฉลี่ยคือ 83,430.50 บิต ค่า PSNR มีค่าเท่ากับ 28.14 dB

ตารางที่ 6.10 บิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิต

สำหรับวิธี CLC

Images : dB 512*512 pixels	With Huffman compression 0.50 bpp	With Huffman compression 0.44 bpp	With Huffman compression 0.38 bpp
Airplane	118,892.00 bits	103,706.00 bits	85,681.00 bits
PSNR	29.19 dB	28.46 dB	27.23 dB
Girl	109,763.00 bits	97,069.00 bits	84,827.00 bits
PSNR	29.76 dB	29.58 dB	28.92 dB
Lena	108,025.00 bits	95,111.00 bits	85,007.00 bits
PSNR	30.07 dB	29.43 dB	28.77 dB
Pepper	93,151.00 bits	88,181.00 bits	78,207.00 bits
PSNR	28.63 dB	28.21 dB	27.63 dB
Average bit	107,457.75 bits	96,016.75 bits	83,430.50 bits
Average bpp	29.41 dB	28.92 dB	28.14 dB

6.11 ผลการจำลองบิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิตสำหรับวิธี CLC+SEC

ตารางที่ 6.11 แสดงถึงบิตเฉลี่ยและ PSNR ของภาพหลังจากมีการเข้ารหัสด้วยวิธีฮัฟแมนกับอัลกอริทึม CLC+SEC ซึ่งถูกออกแบบโดยการใช้การปรับขนาดของอัตราบิตที่อัตราบิต 0.62 bpp, 0.55 bpp และ 0.49 bpp ที่อัตราบิตเฉลี่ยที่ 0.62 bpp บิตเฉลี่ย คือ 141,249.75 บิต ค่า PSNR มีค่าเท่ากับ 33.13 dB ที่อัตราบิตเฉลี่ยที่ 0.55 bpp บิตเฉลี่ยคือ 127,760.75 บิต ค่า PSNR มีค่าเท่ากับ 32.69 dB และที่อัตราบิตเฉลี่ย 0.49 bpp บิตเฉลี่ย คือ 116,198.50 บิต ค่า PSNR มีค่าเท่ากับ 32.26 dB

ตารางที่ 6.11 บิตเฉลี่ยและเฉลี่ย PSNR มีและไม่มีการบีบอัดข้อมูลแบบฮัฟแมนและปรับค่าอัตราบิต

สำหรับวิธี CLC+SEC

Images : dB 512*512 pixels	With Huffman compression 0.62 bpp	With Huffman compression 0.55 bpp	With Huffman compression 0.49 bpp
Airplane	151,660.00 bits	136,474.00 bits	118,449.00 bits
PSNR	32.78dB	32.10 dB	31.52 dB
Girl	142,531.00 bits	129,837.00 bits	117,595.00 bits
PSNR	33.55 dB	33.34 dB	33.02 dB
Lena	144,889.00 bits	123,783.00 bits	117,775.00 bits
PSNR	33.75 dB	33.30 dB	32.78 dB
Pepper	125,919.00 bits	120,949.00 bits	110,975.00 bits
PSNR	32.44 dB	32.01 dB	31.70 dB
Average bit	141,249.75 bits	127,760.75 bits	116,198.50 bits
Average bpp	33.13 dB	32.69 dB	32.26 dB

6.12 ผลการจำลองตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC+SEC เปรียบเทียบกับการเข้ารหัสแบบ PCA อย่างเดียว

เพื่อตรวจสอบว่าอัลกอริทึม CLC+SEC ดีกว่าการเข้ารหัส PCA เพียงอย่างเดียว เราทำการศึกษาระหว่างอัลกอริทึม CLC+SEC และการเข้ารหัส PCA อย่างเดียว ตารางที่ 6.12 แสดงผลการประยุกต์ใช้งานกับภาพทางการแพทย์ระหว่างการเข้ารหัส PCA อย่างเดียว (PCA+DWT = SEC) และอัลกอริทึม CLC+SEC ที่อัตราบิต 0.13 bpp และ 0.25 bpp ผลการศึกษอัลกอริทึม CLC+SEC ให้ค่า PSNR สูงกว่าการเข้ารหัส PCA อย่างเดียว

ตารางที่ 6.12 ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC+SEC

เปรียบเทียบกับวิธีการเข้ารหัสแบบ PCA อย่างเดียว

Algorithms	Images 512x512 pixels	Bit rate (bpp)			
		0.13 PSNR MSE		0.25 PSNR MSE	
PCA+DWT=SEC CLC+SEC	MR Cardiac Tract	34.07 35.00	25.47 20.56	34.33 35.27	23.99 19.32
PCA+DWT=SEC CLC+SEC	MRI Skull	35.75 35.79	17.30 17.14	35.99 36.08	16.37 16.04
PCA+DWT=SEC CLC+SEC	Ultrasound Liver Cyst	26.95 30.53	131.24 57.56	27.49 31.07	115.90 50.83

อัตราบิตที่ใช้ในการจำลอง

อัตราบิตของอัลกอริทึม SEC ที่มีผลกระทบต่อประสิทธิภาพของอัลกอริทึม CLC ยังไม่ใช่ค่าที่เหมาะสมที่สุด อัตราบิตนี้ได้มาจากการทดลองซึ่งจะต้องเป็นค่าที่ยอมรับได้ระหว่างค่า PSNR และจำนวนของบิตที่ใช้ ที่อัตราบิตต่ำจะให้ค่า PSNR ที่ดีซึ่งอันนี้เป็นเพราะว่าข้อผิดพลาดระบบส่วนใหญ่มีค่ามากในขณะที่น่าไปเข้าขบวนการเข้ารหัสและถอดรหัส ในทำนองเดียวกันในงานวิจัยนี้อัลกอริทึม SEC ถูกออกแบบเพื่อให้ทำงานที่อัตราบิตต่ำ เมื่อเรากำหนดค่าอัตราบิตของ CLC+SEC ค่าดังกล่าวจะต้องมีค่าอัตราบิตไม่มากกว่าค่าอัตราบิตของ CLC ยิ่งกว่านั้นการใช้เวฟเลตแม่ (Mother Wavelet) ใน DWT จะต้องถูกเลือกให้เหมาะสมกับค่าของ $Final_Data (E_1')$ ในงานวิจัยนี้เลือกใช้เวฟเลตแม่ชื่อ "Haar" ซึ่งมันสามารถรองค่าพลังงานของสัมประสิทธิ์ที่ความถี่ต่ำออกได้ดี รวมถึงมันให้คุณสมบัติของการบีบอัดพลังงานได้สูง (High Energy Compactness)

บทที่ 7

อภิปรายผล

ในบทนี้จะเป็นการกล่าวถึงการอภิปรายผลการทดลอง โดยเน้นไปที่ผลของการปรับปรุงเวลาของการประมวลผล ผลของการปรับปรุงคุณภาพ ผลของการปรับปรุงอัตราบิตและผลของการปรับปรุงเมื่อพิจารณาทั้งการปรับปรุงคุณภาพและการปรับปรุงอัตราบิต

7.1 เมื่อปรับปรุงเวลา

ผลของการปรับปรุงอัลกอริทึมของ FCM โดยใช้ SVM เรามีการเปรียบเทียบเวลาในการกำหนดตัวเก็บรหัสระหว่างอัลกอริทึม FCM แบบทั่วไป และแบบการปรับปรุงอัลกอริทึมของ FCM โดยใช้ SVM จำนวนของเวลาที่ใช้ในการทดสอบภาพดังแสดงในตารางที่ 6.1 ในกรณีของการใช้อัลกอริทึม FCM แบบทั่วไปค่าสมาชิกภาพเริ่มต้นจะถูกกำหนดโดยการใช้วิธีแบบสุ่ม (Randomly Set) สำหรับภาพที่ใช้ในการทดสอบมีขนาด 256x256 จุดภาพและขนาดบล็อก 4x4 จุดภาพ เวลาเฉลี่ยของการกำหนดตัวเก็บรหัสคือ 92.66 วินาทีสำหรับวิธีการใช้อัลกอริทึม FCM โดยใช้ SVM และ 107.96 วินาที สำหรับอัลกอริทึม FCM แบบทั่วไป และพบว่าเวลาที่ใช้ในการปรับปรุงอัลกอริทึมของ FCM โดยใช้ SVM ใช้เวลาน้อยกว่าอัลกอริทึม FCM แบบทั่วไป 15.3 วินาที

7.2 เมื่อปรับปรุงเฉพาะการปรับปรุงคุณภาพ

กรณีภาพมาตรฐาน ในการจำลองอัลกอริทึม CLC และ CLC+SEC จะใช้การจำลองตัวเก็บรหัสสองวิธี คือ 1) ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพ 2) ตัวเก็บรหัสถูกออกแบบสำหรับสี่ภาพ

- ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

ผลการจำลองดังแสดงในตารางที่ 6.2 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC และ CLC+SEC ค่าผลต่างเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม CLC มีค่าเท่ากับ 7.42 dB, 5.40 dB, 4.23 dB ตามลำดับ

- ตัวเก็บรหัสถูกออกแบบสำหรับสี่ภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC

ผลการจำลองดังแสดงในตารางที่ 6.4 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC และ CLC+SEC ค่าผลต่างเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม CLC มีค่าเท่ากับ 8.31 dB, 7.18 dB, 6.89 dB ตามลำดับ

- ตัวเก็บรหัสถูกออกแบบสำหรับสี่ภาพมาตรฐานสำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+GGCLN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการจำลองดังแสดงในตารางที่ 6.6 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC+SEC และ WT+GGCLN ค่าผลต่างเฉลี่ย PSNR ที่อัตราการบีบอัด (Compression Ratio, CR) เท่ากับ 18.29, 16.79 และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม WT+GGCLN มีค่าเท่ากับ 0.66 dB ตามลำดับ

กรณีภาพทางการแพทย์ ในการจำลองอัลกอริทึม CLC และ CLC+SEC จะใช้การจำลองตัวเก็บรหัสสองวิธีเช่นกัน คือ 1) ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพ 2) ตัวเก็บรหัสถูกออกแบบมาจาก 4 ภาพทางการแพทย์

- ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC ผลการจำลองดังแสดงในตารางที่ 6.3 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC และ CLC+SEC ค่าผลต่างเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม CLC มีค่าเท่ากับ 9.37 dB, 7.15 dB, 6.39 dB ตามลำดับ

- ตัวเก็บรหัสถูกออกแบบมาจาก 3 ภาพทางการแพทย์สำหรับวิธี CLC และ CLC+SEC ผลการจำลองดังแสดงในตารางที่ 6.5 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC และ CLC+SEC ค่าผลต่างเฉลี่ย PSNR ที่อัตราบิต 0.13 bpp, 0.19 bpp และ 0.25 bpp และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม CLC มีค่าเท่ากับ 9.74 dB, 8.52 dB, 8.07 dB ตามลำดับ

- ตัวเก็บรหัสถูกออกแบบมาจาก 3 ภาพทางการแพทย์สำหรับวิธี CLC, CLC+SEC, LBG+SEC เปรียบเทียบกับวิธี WT+HM ผลการจำลองดังแสดงในตารางที่ 6.7 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC+SEC และ WT+HM ค่าผลต่างเฉลี่ย PSNR ที่อัตราการบีบอัด (Compression Ratio, CR) เท่ากับ 32.00, 27.50; 32.00, 26.89; 32.00, 23.07 และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม WT+HM มีค่าเท่ากับ 2.91 dB; 2.98 dB; 4.22 dB ตามลำดับ

7.3 เมื่อปรับปรุงเฉพาะการปรับปรุงอัตราบิต

กรณีภาพมาตรฐาน

ในการจำลองอัลกอริทึม CLC ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพ ผลการจำลองดังแสดงในตารางที่ 6.8 และพบว่าผลของการเปรียบเทียบระหว่างมีและไม่มี การเข้ารหัสด้วยวิธีฮัฟแมน ค่าบิตเฉลี่ย ที่อัตราบิต 0.50 bpp และ 0.41 bpp และอัลกอริทึม CLC มีการเข้ารหัสด้วยวิธีฮัฟแมนต่ำกว่าอัลกอริทึม CLC ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมนมีค่าเท่ากับ 23,690.50 บิต ตามลำดับ

ในการจำลองอัลกอริทึม CLC+SEC ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพ ผลการจำลองดังแสดงในตารางที่ 6.9 และพบว่าผลของการเปรียบเทียบระหว่างมีและไม่มี การเข้ารหัสด้วยวิธีฮัฟแมน ค่าบิตเฉลี่ย ที่อัตราบิต 0.62 bpp และ 0.54 bpp และอัลกอริทึม CLC+SEC มีการเข้ารหัสด้วยวิธีฮัฟแมนต่ำกว่าอัลกอริทึม CLC+SEC ไม่มีการเข้ารหัสด้วยวิธีฮัฟแมนมีค่าเท่ากับ 23,690.50 บิต ตามลำดับ

7.4 เมื่อพิจารณาทั้งการปรับปรุงคุณภาพและการปรับปรุงอัตราบิด

- ตัวเก็บรหัสถูกออกแบบสำหรับแต่ละภาพมาตรฐานสำหรับวิธี CLC และ CLC+SEC มีการเข้ารหัสด้วยวิธีฮัฟแมน

ผลการจำลองดังแสดงในตารางที่ 6.10 และ 6.11 และพบว่าผลของการเปรียบเทียบระหว่างอัลกอริทึม CLC และ CLC+SEC มีการเข้ารหัสด้วยวิธีฮัฟแมน ค่าผลต่างเฉลี่ย PSNR ที่อัตราบิด 0.50 bpp, 0.49 bpp และอัลกอริทึม CLC+SEC มีค่าผลต่างเฉลี่ย PSNR สูงกว่าอัลกอริทึม CLC มีค่าเท่ากับ 2.85 dB ตามลำดับ

- ตัวเก็บรหัสถูกออกแบบสำหรับสามภาพทางการแพทย์ ผลการจำลองดังที่กล่าวมาก่อนหน้านั้นแล้ว ซึ่งเป็นการเปรียบเทียบระหว่างอัลกอริทึม CLC+SEC และ WT+HM



สรุปผลการทดลองและข้อเสนอแนะ

8.1 สรุปผลการทดลอง

ในวิทยานิพนธ์ฉบับนี้ ผู้วิจัยได้ทำการศึกษาวิจัยและนำเสนออัลกอริทึม ใหม่ของการบีบอัด ข้อมูลภาพทางการแพทย์ ซึ่งเป็นการปรับปรุงอัลกอริทึม การบีบอัดข้อมูลเวกเตอร์ควอนไทต์แบบ ทั่วไป (VQ) เพื่อให้ได้เป็นอัลกอริทึม แบบ CLC และ CLC+SEC ชั้นตอนสามชั้นตอนถูกนำมาใช้ในการ ปรับปรุงการบีบอัดข้อมูลเวกเตอร์ควอนไทต์แบบทั่วไปสำหรับสองอัลกอริทึม ดังกล่าวข้างต้นนี้ ชั้นตอนที่หนึ่ง ตัวเก็บรหัสถูกออกแบบโดยใช้การเข้ารหัสสัมประสิทธิ์ของ DWT และ FCM ถูก นำมาใช้แบ่งแยกกลุ่มค่าสมาชิกของสัมประสิทธิ์และอัลกอริทึม SVM ถูกนำมากำหนดค่าเริ่มต้นให้กับ FCM ชั้นตอนที่สองเรานำเสนอการเข้ารหัสฮัฟแมนเพื่อกำจัดการซ้ำกันของค่าดัชนีของตัวเก็บรหัส ชั้นตอนที่สาม การชดเชยค่าผิดพลาดของระบบถูกนำมาใช้เพื่อปรับปรุงค่า PSNR ของระบบ VQ คุณสมบัติของอัลกอริทึม CLC และ CLC+SEC ต่อภาพทั้งสองชนิด คือ ภาพมาตรฐานและภาพทาง การแพทย์ และถูกศึกษาในส่วนของค่า PSNR และอัตราบิตโดยมีรายละเอียด ดังนี้

ผลการศึกษาในส่วนของค่า PSNR เมื่อตัวเก็บรหัสถูกออกแบบโดยใช้ภาพมาตรฐานสีภาพ คุณสมบัติของอัลกอริทึม CLC+SEC ดังแสดงในตารางที่ 6.6 2.19% ดีกว่าเมื่อดูในค่าของ PSNR เมื่อ เปรียบเทียบกับวิธีทั่วไป [4] รวมถึงตัวเก็บรหัสเมื่อถูกออกแบบโดยใช้สามภาพทางการแพทย์ อัลกอริทึม CLC+SEC ดังในตารางที่ 6.7 จะให้คุณสมบัติ 10.96% ดีกว่าเมื่อดูในค่าของ PSNR เมื่อ เปรียบเทียบกับวิธีทั่วไป [31]

ผลการศึกษาในส่วนของค่าอัตราบิต เมื่ออัตราบิตเป็นปัจจัยหลักสำคัญในการพิจารณา ตารางที่ 6.6 แสดงอัลกอริทึม CLC+SEC กับตัวเก็บรหัสถูกสร้างจากสีภาพมาตรฐานจะให้อัตราบิตที่ ดีกว่า 8.47% เมื่อเปรียบเทียบกับวิธีทั่วไป [4] ในกรณีที่ตัวเก็บรหัสถูกสร้างจากสามภาพทาง การแพทย์ดังแสดงในรูปที่ 6.2 อัลกอริทึม CLC+SEC จะให้อัตราบิตที่ดีกว่า 24.00% เมื่อเปรียบ เปรียบเทียบกับวิธีทั่วไป [31]

ผลการศึกษาเมื่อพิจารณาพร้อมกันทั้งค่าของ PSNR และค่าอัตราบิตอัลกอริทึม CLC+SEC จะให้คุณสมบัติที่ดีกว่าวิธีทั่วไป ผลของค่า PSNR แสดงว่าคุณสมบัติของอัลกอริทึม CLC ต่ำกว่า อัลกอริทึม WT+GGCLN แต่ให้อัตราการบีบอัด (Compression Ratio) ที่สูงกว่า เพื่อปรับปรุง คุณสมบัติของอัลกอริทึม CLC ตัวเก็บรหัสจะต้องถูกออกแบบให้เป็นสากล (Universal Codebook) การเปรียบอัลกอริทึม CLC+SEC กับ CLC และวิธีทั่วไปดังแสดงในตารางที่ 6.6 และตารางที่ 6.7 จะ รองรับให้อัตราบิตต่ำและค่า PSNR สูง และดังนั้นมันก็เหมาะสมเพียงที่จะนำไปใช้กับการบีบอัด เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยามให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลภาพทางการแพทย์ได้ สูดท้ายอัลกอริทึม ที่นำเสนอ (CLC, CLC+SEC) เมื่อนำไปประยุกต์ใช้ งานกับภาพทางการแพทย์ ผลลัพธ์ที่ได้จะให้ผลที่ดีกว่าวิธีทั่วไปสำหรับภาพทางการแพทย์ที่ทำการ ทดสอบทั้งหมด

อย่างไรก็ตามเพื่อเป็นการยืนยันถึงความเป็นไปได้จริงของหลักการที่นำเสนอมานี้ จึงควร จะต้องมีการศึกษาวิจัยต่อเนื่องในเรื่องของการออกแบบและสร้างระบบการชดเชยค่าผิดพลาดของ ระบบให้มีประสิทธิภาพที่สูงขึ้น โดยเฉพาะจำนวนบิตที่ใช้ในระบบควรจะต้องมีจำนวนบิตและเวลาที่ ใช้ในการประมวลผลที่ต่ำที่สุด และรวมถึงจะต้องมีโครงสร้างที่ไม่ซับซ้อนก่อนในเบื้องต้น ทั้งนี้เพื่อเป็น การประเมินทิศทางการชดเชยค่าผิดพลาดของระบบจริงต่อไป

8.2 ข้อเสนอแนะ

1. ในงานวิจัยนี้อยู่ในแนวเส้นทางไหนของการบีบอัดข้อมูลใช้เว็ทเตอร์ควอนไทต์

การกำราบบีบอัดข้อมูลภาพทางการแพทย์โดยใช้เว็ทเตอร์ควอนไทต์และการชดเชย ข้อผิดพลาดของระบบนี้เป็นแนวเส้นทางหนึ่งของการบีบอัดข้อมูลภาพแบบเว็ทเตอร์ควอนไทต์หลาย ชั้น (Multi-stage VQ) [37]

2. ทำไมต้องเป็นการบีบอัดการบีบอัดข้อมูลภาพทางการแพทย์

เนื่องจากระบบมีการแยกข้อมูลที่มีค่าจุดภาพซ้ำกันไปอยู่รวมกันที่ความถี่ต่ำและข้อมูลที่มี ค่าจุดภาพต่างกันไปอยู่รวมกันที่ความถี่สูง โดยเฉพาะภาพที่มีบริเวณซ้ำกันส่วนมาก เช่นภาพ CT ภาพ MRI นั้นจึงเป็นคำตอบว่าทำไมต้องเป็นการบีบอัดการบีบอัดข้อมูลภาพทางการแพทย์

3. จุดอ่อนและจุดแข็งของวิธีนี้

จุดอ่อน ของงานวิจัยนี้คือ

1) การออกแบบตัวเก็บรหัสมีความซับซ้อนในการคำนวณสูง ในกรณีที่ต้องการ ใช้เวลาในการคำนวณที่เป็นเวลาจริง (real-time) การออกแบบตัวเก็บรหัสด้วยวิธีที่นำเสนอจะใช้เวลา มาก

2) ในกรณีที่ภาพมีรายละเอียดสูงมากจะใช้เวลานานในการแบ่งกลุ่มของข้อมูล นาน และรวมถึงภาพที่ได้จากการถอดรหัสจะเกิดเป็นตารางสี่เหลี่ยม (blocky effect)

3) ในกรณีที่บล็อกของภาพมีขนาดใหญ่ (มากกว่า 16×16 จุดภาพ) ภาพที่ได้ จากการถอดรหัสจะได้ PSNR ต่ำ เพราะว่าจำนวนบล็อกของภาพที่ใช้ในการเปรียบเทียบมีจำนวน น้อย

จุดแข็ง ของงานวิจัยนี้คือ

1) ในการเข้ารหัสใช้จำนวนบิตต่ำ เพราะว่าในการถอดรหัสเพื่อสร้างภาพ กลับคืนระบบที่นำเสนอในงานวิจัยนี้จะมีการชดเชยค่าผิดพลาดของระบบ และมีการลดการซ้ำกันของ ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) เหมาะสำหรับภาพที่มีบริเวณซ้ำกันส่วนมากเป็น ขาวและดำ เช่น ภาพทางการแพทย์ (ภาพ MRI ภาพ CT) ภาพที่ได้จากการถอดรหัสจะได้ PSRN สูง เพราะวาระบบมีการแยกข้อมูลที่มีค่าจุดภาพซ้ำกันไปอยู่รวมกันที่ความถี่ต่ำและข้อมูลที่มีค่าจุดภาพต่างกันไปอยู่รวมกันที่ความถี่สูง และรวมถึงระบบมีการชดเชยค่าผิดพลาดของระบบอีกด้วย

3) สามารถใช้ได้กับบล็อกของภาพมีขนาดใหญ่ (4x4, 8x8 และ 16x16 จุดภาพ) ภาพที่ได้จากการถอดรหัสจะได้ PSRN จะมีค่าสูงไม่แตกต่างกันมาก เพราะวาระบบมีการชดเชยค่าผิดพลาดของระบบ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] Christopher F. Barnes, Syed A. Rizvi, and Nasser Nasrabadi M. "Advances in Residual Vector Quantization: A Review." *IEEE Trans. Image processing*, vol. 5, no. 2, Feb. 1996. pp. 226-262.
- [2] Vasuki A. and Vanathi P.T. "A review of vector quantization techniques" *IEEE potentials*, Jul/Aug. 2006. pp. 39-47.
- [3] M. Antonini, M. Barlaud, P. Mathieu, and Daubechies I. "Image Coding Using Wavelet Transform." *IEEE Trans. Image Processing*, vol. 1, no. 2, 1992. pp. 205-220.
- [4] L. Chin – Yuan, and Chin – Hsing C. "A genetic grey based neural networks with wavelet transform for search of optimal codebook." *IEICE Trans. Fundamentals*, vol. e86-a, no. 3, 2003. pp. 715-721.
- [5] Shuangteng Z. "Side-Match Vector Quantizers Using Neural Network Based Variance Predictor for Image Coding." *International Journal of Image Processing (IJIP)*, vol. 5, Issue (2), 2011.
- [6] Stephane M. and Frederic F. "Analysis of Low Bit Rate Image Transform Coding." *IEEE Trans. Signal processing*, vol. 46, no. 4, April. 1998. pp. 1027-1042.
- [7] Chaur – Heh H. and Yong – Jzu L. "Fast Search Algorithm for Vector Quantization of Images Using Multiple Triangle Inequalities and Wavelet Transform." *IEEE Trans. Image Processing*, vol. 9., no.3, 2000. pp. 321-328.
- [8] Chin – Chang C. and Wen – Chuan W. "Fast Planar-Oriented Ripple Search Algorithm for Hyperspace VQ Codebook." *IEEE Trans. Image processing*, vol. 16, no. 6, June. 2007. pp. 1538-1547.
- [9] ISO/IEC JTC 1/SC 29/WG1 FCD 14495-public draft. "lossless and near-lossless coding of continuous tone still images." 1997, JPEG-LS.
- [10] Gallager R. G. "Variations on a theme by Huffman." *IEEE Trans. Inform. Theory*, IT 25, Nov 1978. pp. 668-674.
- [11] Gharavi H. "Condition run-length and variable-length coding of digital picture." *IEEE Trans. COMM*, COM-35(6), June 1987. pp. 671-677.
- [12] Gonzale R. C. and Woods R. E. *Digital Image Processing*. Addison-Wesley Company, Inc. 1992.
- [13] Rabbin M. and Jones P.W. "Digital Image Compression Techniques." *SPIE Tutorial Texts in Optical Engineering*, vol. TT7, SPIE Press, 1991.
- [14] Ziv J. and Lempel A. "A universal algorithm for sequential data compression." *IEEE Trans. Inform. Theory*, IT-23, 1977. pp. 337-343.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [15] Lloyd S. P. "Least Squares Quantization in PCM." *IEEE Trans. On Inform. Theory*, vol. IT-28, Mar 1982. pp. 129-137.
- [16] Max J. "Quantizing for Minimum Distortion." *IRE Trans. Inf. Theory*, vol. IT-6, Mar 1960. pp. 7-12.
- [17] Sid-Ahmed M. A. *Image Processing, Theory, Algorithm, and Architectures*, McGrawHill Inc. 1995.
- [18] Woods J. W. *Subband Image Coding*, Kluwer Academic Publishers, Boston, 1991.
- [19] Nasrabadi N. M. and King R. A. "Image coding using vector quantization: A Review." *IEEE Trans. Communication*, vol. COM-36, Aug 1980. pp. 957-971.
- [20] Linde Y. Buzo A. and Gray R. M. "An algorithm for vector quantize design." *IEEE Trans. Communication*, vol. COM-28, Jan 1980. pp. 84-95.
- [21] Li-Xin Wang. "A Course in fuzzy systems and control." USA. Prentice-Hall Inc. Upper Saddle River, NJ 07458, 1997. pp. 342-350.
- [22] Liu Y. H. and Chen Y. T. "Face Recognition Using Tital Margin-Based Adaptive Fuzzy Support Vector Machines." *IEEE Transactions on Neural Networks*, vol. 18, 2007. pp. 178-192.
- [23] Zhang J. P., Li Z. W. and Yang J. "A parallel SVM training algorithm on large-scale classification problem." *Proceedings of 2005 International Conference on Machine Learning and Cybernetics*, vol. 3, 2005. pp. 1637-1641.
- [24] Nasrabadi N. M. and King R. A. "Image coding using vector quantization: A review." *IEEE Trans. on Communications*, vol. 36, No. 8, Aug 1988. pp. 957-971.
- [25] Ra S.-W., Kim J.-K. "A fast mean-distance-ordered partial codebook search algorithm for image vector quantization." *IEEE Trans. on Circuit and systems*, vol. 40, no. 93, September 1993. pp. 576-579.
- [26] Chang C. C., Lin D. C., and Chen T. S. "An improved VQ codebook search algorithm using principal component analysis." *Journal of visual communication and image representation*, vol. 8, no. 1, 1997. pp. 27-37.
- [27] Patane G., Russo M. "The enhanced LBG algorithm." *Pergamon*, vol. 14, 2001. pp. 1219-1237.
- [28] Lin T.-C. and Yu P.-T. "Centroid neural network adaptive resonance theory for vector quantization." *Elsevier, Signal Processing*, vol. 83, 2003. pp. 649-654.

- [29] Shapiro J. M. "Embedded image coding using zero trees of wavelet coefficients." *IEEE Trans. on Signal Processing*, vol. 41, no. 12, 1993. pp. 3445-3462.
- [30] Said A. and Pearlman W.A. "A new fast and efficient image codec based on set partitioning in hierarchical trees." *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6., June 1996. pp. 243-250.
- [31] Ruchika, M. Singh, Singh A. R. "Compression of medical image using wavelet transforms." *IJSCE*, ISSN: 2231-2307, vol.2, Issue-2, May 2012. pp. 339-343.
- [32] Dehkordi V. R., Daou H. and Labeau F. "A channel differential EZW coding scheme for EEG data compression." *IEEE Trans. on Information technology in biomedicine*, vol. 15, no. 6, 2001. pp. 243-250.
- [33] El-Sakka M. *Advanced topics in image compression*, 2002. <http://www.csd.uwo.ca/faculty/elsakka/>
- [34] Demirel H. and Anbarjafari G. "Discrete wavelet transform-based satellite image resolution enhancement." *IEEE Trans. Geoscience and remote sensing*, vol. 49, no. 6, June, 2011. pp.1997- 2004.
- [35] Paul C. *A fifteen minutes introduction of wavelet transform and applications*, 1998. <http://www.gler.noaa.gov/pers/liu/>
- [36] Nagendran R., ArockiaJansi Rani P. "A novel FPCM based vector quantize codebook design for image compression in the wavelet packet domain." *The International journal of Multimedia \& Its Application (IJMA)*, Vol. 2, no. 1, Feb, 2010.
- [37] Jeong G.-H. and Lee I.-S. "Wavelet-based ECG compression using dynamic multi-stage vector quantization." *IEEE Conference on Industrial Electronics and Applications*, 25-27 May, 2009. pp. 2100-2105.
- [38] C.Kuo-B., Huang W.-C., Liu H.-C. and Tseng S.-C. "A novel fuzzy C- means method for hyper spectral image classification." *Geoscience and Remote Sensing Symposium, IGARSS 2008. IEEE International*, vol. 2, 2008. pp. 1002-1005.

- [39] Havens T. C., Bezdek J. C., Leckie C., Hall L. O. and Palaniswami M., "Fuzzy C means algorithms for very large data." *IEEE Trans on fuzzy systems*, vol. 20, no. 6, December, 2012. pp. 1130-1146.
- [40] Hsu C.-H., Chang C.-C., Lin C.-J. **A practical guide to support vector classification**, National Taiwan University, Department of Computer Science, April 3, 2010.
- [41] Gutschoven B. and Verlinde P. "Multi-modal identity verification using support vector machines (SVM)." **The third Int. Conference on Information Fusion**, 2000. pp. 3-8.
- [42] Zhang H., Shi W. and Liu K. "Fuzzy- topology-integrated support vector machine for remotely sensed image classification." *IEEE Transactions on Geoscience and remote sensing*, vol. 50, 2012. pp. 850-862.
- [43] Juang C.-F. and Chen G.-C. "A TS fuzzy system learned through a support vector machine in principal component space for real-time object detection." *IEEE Transactions on Industrial electronics*, vol. 59, no.8, 2012. pp. 3309-3320.
- [44] Li X., Shen Y. "A medical image compression scheme based on low order predictor and most-likely magnitude Huffman code." **Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation**, June, 25-28, Luoyang, China, 2006.
- [45] Yang E.-h., Wang L. "Joint optimization of Run-Length coding, Huffman coding, and quantization table with complete baseline JPEG decoder compatibility." *IEEE Transactions on Image Processing*, vol. 18, 2009. pp. 63-74.
- [46] DAVID A. Huffman, ASSOCIATE, IRE. "A method for the construction of minimum redundancy codes." *Proceedings of the IRE*, volume 40, September, 1952. pp. 1098-1101.

ภาคผนวก

ผลงานวิจัยที่ได้รับการตีพิมพ์

1. Phanprasit T, Leauhatong T, Pintavirooj C and Sangworasil M. "Image Coding Using Vector Quantization Based on Wavelet Transform Fuzzy C-Means and Principle Component Analysis" The 9th International Symposium on Communication and Information Technology 2009(ISCIT 2009), Songdo-lfez ConvensiA, Incheon, Korea, September 28-30, 2009, pp. 448-453.
2. Phanprasit T, Pintavirooj C and Sangworasil M. "Image Compression Using Vector Quantization and System Error Compensation" 2010 International Workshop on Information Communication Technology (ICT 2010), KMITL, Bangkok, Thailand, Aug 24-25, 2010.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล
วัน เดือน ปีเกิด
ที่อยู่

นายชนะศักดิ์ พันธุ์ประสิทธิ์
11 เมษายน 2503
หมู่บ้าน เดอะสวีทการ์เด้นวิลล์ 2 27/17 ถ. หนองระแหง แขวง
สามวาตะวันตก เขตคลองสามวา จ.กรุงเทพมหานคร 10510

ประวัติการศึกษา

2537 วิศวกรรมศาสตรมหาบัณฑิต สาขา วิศวกรรมไฟฟ้า
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
2532 วิศวกรรมศาสตรบัณฑิต สาขา วิศวกรรมอิเล็กทรอนิกส์
มหาวิทยาลัยเอเชียอาคเนย์

ประวัติการทำงาน

2537- ปัจจุบัน อาจารย์ประจำ (รองศาสตราจารย์)
คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมไฟฟ้าและ
อิเล็กทรอนิกส์
มหาวิทยาลัยกรุงเทพ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้