

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาพดิจิทัลเบื้องต้น

2.1.1 ภาพดิจิทัล (Digital Image) ข้อมูลภาพในคอมพิวเตอร์มีการแสดงผล 2 แบบ คือ บิตแมพหรือแรสเตอร์หรือแผนที่จุดภาพ (Bitmap หรือ Raster หรือ Pixel maps) และเวกเตอร์ (Vector) (Kay และ Levine, 1992: 1)

2.1.1.1 บิตแมพ เป็นวิธีที่นิยมใช้กันมากกว่า เพราะง่ายแก่การนำไปใช้จริงและสามารถทำงานกับภาพชนิดใดก็ได้ เมื่อเริ่มที่จะประมวลผลข้อมูลภาพด้วยคอมพิวเตอร์นั้น จะต้องมีการเปลี่ยนข้อมูลภาพให้อยู่ในรูปข้อมูลดิจิทัลซึ่งคอมพิวเตอร์สามารถเข้าใจได้ กระบวนการดังกล่าวนี้เรียกว่า ดิจิไทเซชัน (Digitization) ภาพดิจิทัลชนิดบิตแมพนี้จะแสดงผลโดยใช้แนวความคิดในการจัดการชุดของตัวเลขให้อยู่ในรูปของแถวลำดับหรือเมทริกซ์ (Matrix) ที่มีระยะห่างเท่าๆ กัน โดยมีแถว (Row) และสดมภ์ (Column) เป็นตัวบ่งชี้ตำแหน่งของจุดในภาพ ซึ่งเรียกว่า สมาชิกของภาพ (Image element หรือ Picture element) หรือใช้ตัวย่อว่า จุดภาพ (Pixel) หรือ เพล (Pel) จุดภาพนี้จะ เป็นหน่วยที่เล็กที่สุดของภาพ โดยที่ค่าของจุดภาพจะเป็นตัวเลขจำนวนนับซึ่งจะเป็นค่าที่บอกระดับความเทา (Gray level) หรือระดับความสว่าง (Brightness) หรือแอมพลิจูด (Amplitude) ของจุดที่อยู่ตรงตำแหน่งนั้น จำนวนบิตในจุดภาพนี้ เรียกว่า ความลึกแอมพลิจูดหรือความลึกจุดภาพ (Amplitude depth หรือ Pixel depth) อาจจะประกอบไปด้วย 1, 2, 4, 8, 12, 16 หรือ 24 บิตก็ได้ (ศรีสุดา จารีก, 2536: 5)

ในการแสดงผลภาพ จะมองภาพเป็นฟังก์ชันสองมิติของความเข้มแสง $f(x, y)$ เมื่อ x และ y แทนพิกัดต่างๆ บนภาพ เรียกว่าเป็นสเปซโคออดิเนต (Spatial coordinate) โดยแต่ละตำแหน่งคือจุดภาพซึ่งจะอาศัยความสัมพันธ์ของแถวและสดมภ์ ดังรูปที่ 2.1 ที่แสดงให้เห็นถึงการตั้งแกนพิกัดอ้างอิงที่ใช้ในการประมวลผลภาพ

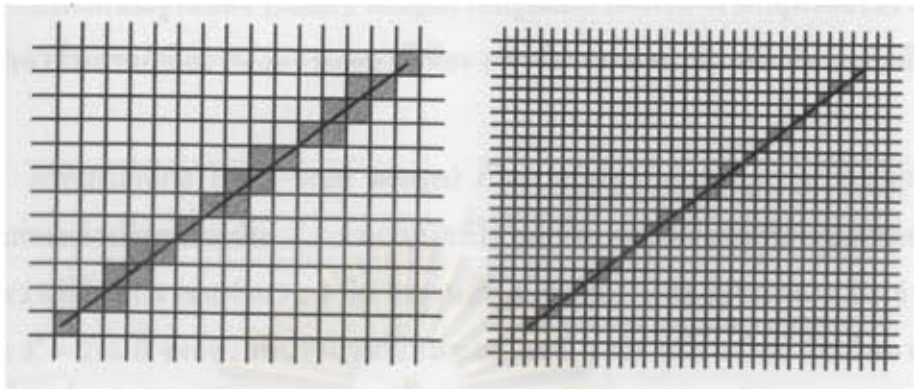


รูปที่ 2.1 แสดงพิกัดอ้างอิงที่ใช้ในการประมวลผลภาพ

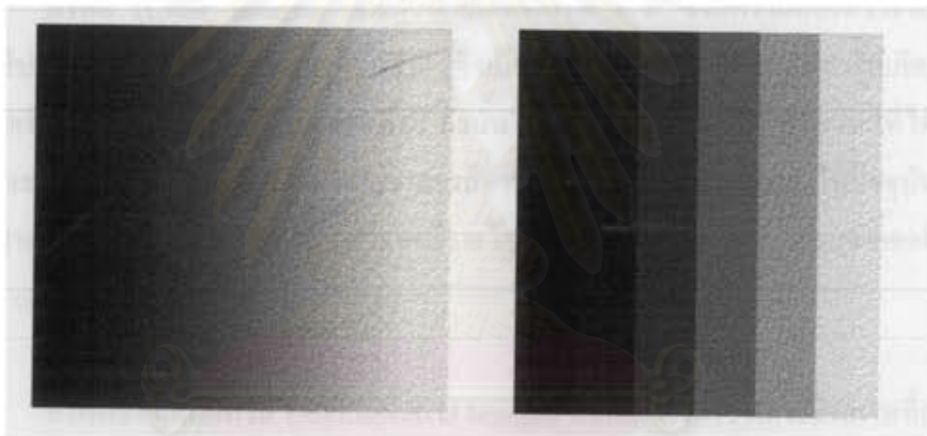
การดิจิไทซ์หรือการแปลงภาพจากภาพแอนะล็อกให้เป็นภาพดิจิทัล ประกอบไปด้วย การชักตัวอย่าง (Sampling) และควอนไทเซชัน (Quantization) (ดวงอาทิตย์ ศรีบุญ และ ชงสุทร เอก อริยทรัพย์, 2539: 3-5)

การชักตัวอย่าง ภาพที่เรามองเห็นจะมีความต่อเนื่องในตำแหน่งต่างๆ แต่ข้อมูลที่มี ลักษณะต่อเนื่องเช่นนี้ไม่สามารถเก็บมาประมวลผลด้วยคอมพิวเตอร์ได้ ซึ่งจำเป็นต้องแปลงภาพ ทางด้านตำแหน่งก่อน โดยการแบ่งภาพออกเป็นตารางกริด (Grid) โดยช่องต่างๆ ในตารางก็คือ ตำแหน่ง (x, y) ของฟังก์ชัน $f(x, y)$ นั่นเอง จากรูปที่ 2.2 จะแสดงให้เห็นการชักตัวอย่างภาพเส้น เียด จะเห็นได้ว่ายิ่งจำนวนกริดมากขึ้น ก็จะทำให้รายละเอียดของตำแหน่งต่างๆ มากขึ้นด้วย แต่จะทำ ให้จำนวนข้อมูลที่ใช้เก็บแทนภาพจะมากขึ้นด้วย

ควอนไทเซชัน นอกจากภาพในธรรมชาติจะมีความต่อเนื่องทางด้านตำแหน่งแล้ว สีหรือระดับความเข้มแสงก็มีความต่อเนื่อง และมีค่าไม่จำกัดอีกด้วย ซึ่งก็คือข้อมูลแอนะล็อก แต่ข้อมูลดิจิทัลจะมีความไม่ต่อเนื่องและมีค่าจำกัด โดยขึ้นอยู่กับจำนวนบิตที่ใช้แทนข้อมูลแอนะล็อกนั่นเอง จากรูปที่ 2.3 แสดงลำดับขั้นของความเข้มแสงที่ใช้แทนแต่ละจุดภาพที่ได้จากการชักตัวอย่าง



รูปที่ 2.2 แสดงการชักตัวอย่างภาพด้วยจำนวนกริดต่างกัน
(ดวงอาทิตย์ ศรีมูล และ ธงยุทธ เอกอริชทรัพย์, 2539: 4)



รูปที่ 2.3 เปรียบเทียบภาพที่เห็นในธรรมชาติกับภาพที่ผ่านควอนไทเซชัน
(ดวงอาทิตย์ ศรีมูล และ ธงยุทธ เอกอริชทรัพย์, 2539: 5)

2.1.1.2 **เวกเตอร์** จะแสดงผลภาพบนคอมพิวเตอร์ด้วยชุดของเส้นตรงหรือรูปทรงทางเรขาคณิตต่างๆ บางครั้งบางบริเวณจะถูกทำให้ที่บ้คัวระดับต่างๆ ของสี ข้อมูลภาพชนิดเวกเตอร์จะดูเหมือนโปรแกรม หรือเหมือนคำสั่ง และข้อมูลในแอสกี ซึ่งจะใช้ได้กับภาพชนิดที่มีรูปทรงอย่างง่าย เช่น โปรแกรมวาดภาพแคด(CAD drawing), แผนภูมิสถิติ, แผนที่, กราฟต่างๆ, ภาพที่มีรูปร่างและสีอย่างง่าย เป็นต้น

2.1.2 ชนิดของภาพ (Image types) (Umbaugh, 1998: 24-31)

ภาพถักรหัสฐานสอง (Binary Images) เป็นรูปแบบของภาพที่ง่ายที่สุด โดยใช้จำนวน 1 บิตต่อจุดภาพ เป็นภาพทีเดียวน (Monochrome) มีเพียง 2 ค่า คือ “1” และ “0” แทนสีขาวและสีดำ

ภาพเกรย์สเกล (Gray-Scale Images) เป็นภาพทีเดียวนที่เก็บข้อมูลระดับความเทาหรือระดับความส่องสว่างที่แตกต่างกัน จำนวนบิตที่ใช้ในแต่ละจุดภาพขึ้นอยู่กับความแตกต่างของความส่องสว่างที่ต้องการ เช่น จำนวน 4 บิต จำนวนบิตที่ได้เท่ากับ $2^4 = 16$ บิต จำนวน 8 บิต จำนวนบิตที่ได้เท่ากับ $2^8 = 256$ บิต ซึ่งถือว่าเป็นจำนวนที่เหมาะสมเพราะ 8 บิต หรือ 1 ไบต์ ถือเป็นมาตรฐานของหน่วยที่เล็กที่สุดในโลกดิจิทัลคอมพิวเตอร์ แต่ในบางกรณี เช่น งานทางด้านการแพทย์หรืองานด้านดาราศาสตร์ อาจใช้ 12 หรือ 16 บิตต่อจุดภาพก็ได้

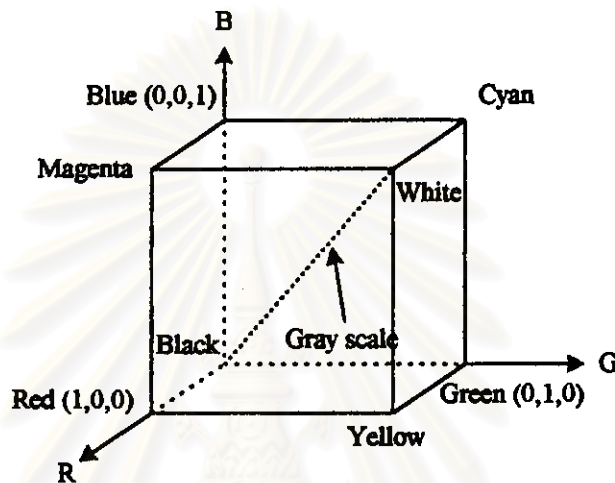
ภาพสี (Color Images) เป็นส่วนผสมของภาพ 3 สี ซึ่งมีความส่องสว่างไม่เท่ากัน ประกอบไปด้วยสีแดง, สีเขียว และสีน้ำเงิน ซึ่งเป็นสีหลักที่ใช้ในการผสมสีที่ทำให้เกิดภาพในคอมพิวเตอร์ โดยใช้จำนวนบิต 8 บิตต่อหนึ่งสี ดังนั้น หนึ่งจุดภาพใช้ 24 บิต จำนวนบิตที่ได้เท่ากับ 2^{24} สี (ประมาณ 16 ล้านสี) ซึ่งก็คือ แบบจำลองสีอาร์จีบี (RGB color model) ในปัจจุบันมีแบบจำลองสีมาตรฐานอีกมากมาย เช่น แบบจำลองสีวายไอคิว (YIQ color model), แบบจำลองสีเอชไอเอส (HIS color model) เป็นต้น

ภาพหลายสเปกตรัม (Multispectral Images) จากที่กล่าวมาข้างต้นเป็นภาพที่สามารถมองเห็นได้โดยตาของมนุษย์ซึ่งอยู่ในรูปส่วนประกอบอาร์จีบีซึ่งมี 3 คลื่นสเปกตรัม แต่ภาพชนิดนี้เป็นภาพที่มนุษย์ไม่สามารถมองเห็นได้เพราะมีคลื่นสเปกตรัมมากกว่า 3 คลื่น เช่น อินฟราเรด, อัลตราไวโอเล็ต, เอ็กซ์เรย์, เสียง หรือเรดาร์ เป็นต้น ซึ่งต้องมีการประมวลผลให้อยู่ในรูปแบบที่มองเห็นได้ แหล่งข้อมูลของภาพชนิดนี้ ได้แก่ ระบบดาวเทียม, ระบบโซนาร์ใต้น้ำ, ระบบภาพสีอินฟราเรด, และระบบภาพทางการแพทย์ที่ใช้ในการวินิจฉัยโรค เป็นต้น

2.1.3 แบบจำลองของสี (Color model)

2.1.3.1 แบบจำลองสีอาร์จีบี (The RGB color model) (Gonzalez และ Woods, 1992: 226-227) แบบจำลองสีนี้ถูกเสนอโดยเอ็นทีเอสซี (NTSC หรือ The National Television Systems Committee) เพื่อแสดงผลภาพสีที่ใช้กับจอภาพสีอาร์จีบี ในแบบจำลองสีนี้มีแม่สีหลัก (Primary color) ซึ่งเรียกว่าแม่สีบวก (Additive Primaries) ได้แก่ สีแดง สีเขียวและสีน้ำเงิน เมื่อนำแม่สีเหล่านี้

มาผสมกันจะทำให้ได้ผลลัพธ์เป็นสีต่างๆ ขึ้นดังแสดงในรูป 2.4 ได้แก่ สีม่วงแดง สีเขียวอมน้ำเงิน สีเหลือง ซึ่งตามสีนี้เรียกว่าสีรอง (Secondary color) นอกจากนี้มีสีค่าที่ตำแหน่งจุดก้นบนแกนด้วยคู่ค่าดับ (0,0,0) และสีขาวแทนด้วยคู่ค่าดับ (1,1,1) เส้นทแยงมุมของแบบจำลองที่ลากระหว่างสีขาวกับสีดำจะหมายถึงสีเทาในระดับต่างๆ ซึ่งมีสัดส่วนของแม่สีสีแดง สีเขียว และสีน้ำเงินที่มีค่าเท่าๆ กัน



รูปที่ 2.4 แสดงแบบจำลองสีอาร์จีบี (Gonzalez และ Woods, 1992: 226)

2.1.3.2 แบบจำลองสีซีเอ็มวาย (The CMY color model) (Gonzalez และ Woods, 1992: 227-228) แบบจำลองสีนี้ถูกออกแบบมาให้ใช้กับการพิมพ์ภาพสีบนเครื่องพิมพ์และพล็อตเตอร์ แม่สีในแบบจำลองสีซีเอ็มวายจะเป็นสีตรงข้ามของแม่สีในแบบจำลองอาร์จีบี ซึ่งเป็นสีที่ได้จากการหักลบสีออกจากแสงสีขาว จึงเรียกแม่สีในแบบจำลองนี้ว่าแม่สีลบ (Subtractive Primaries) ได้แก่ สีม่วงแดง (Magenta มาจากสีแดงผสมสีน้ำเงิน) สีเขียวอมน้ำเงิน (Cyan จากสีเขียวผสมสีน้ำเงิน) สีเหลือง (จากสีแดงผสมสีเขียว) เราสามารถหาค่าแบบจำลองสีซีเอ็มวายได้จากแบบจำลองสีอาร์จีบี ดังนี้

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

ในทำนองเดียวกัน หากเราต้องการหาค่าแบบจำลองสีอาร์จีบีจากแบบจำลองสีซีเอ็มวาย ได้ดังนี้

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$

เนื่องจากการพิมพ์ภาพของอุปกรณ์แสดงผลด้วยแบบจำลองสีซีเอ็มวายนี้ จะทำให้การพิมพ์สีค่าต้องเปลืองหมึกมากเพราะสีด้านเกิดจากการผสมสีเขียวอมน้ำเงิน สีม่วงแดง และเหลือง และคุณภาพหมึกที่ได้ก็ไม่สมบูรณ์ จึงทำให้แบบจำลองนี้ไม่สามารถสร้างสีค่าได้อย่างถูกต้อง ดังนั้นจึงได้เกิดแบบจำลองสีใหม่ที่มีการรวมสีค่าลงไปด้วยโดยไม่ผ่านการผสมสีทั้งสาม คือ แบบจำลองสีซีเอ็มวายเค (The CMYK color model) โดยที่ K แทนสีค่าซึ่งเป็นสีที่สี่ มีความสัมพันธ์กับแบบจำลองสีเดิม ดังนี้

$$K = \min(C, M, Y)$$

$$C = C - K$$

$$M = M - K$$

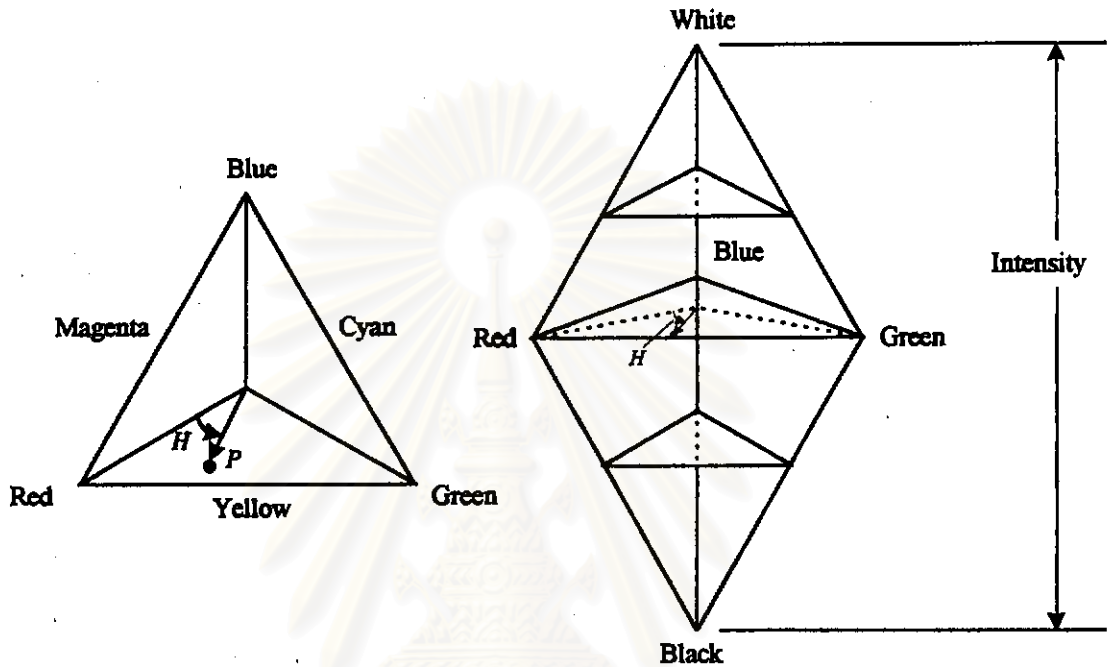
$$Y = Y - K$$

2.1.3.3 แบบจำลองสีวายไอคิว (The YIQ color model) (Gonzalez และ Woods, 1992: 228-229) แบบจำลองนี้ถูกออกแบบมาให้ใช้การถ่ายทอด้วยสัญญาณโทรทัศน์ระบบเอ็นทีเอสซี ซึ่งเป็นการเปลี่ยนแปลงสัญญาณสีในแบบจำลองสีอาร์จีบีให้เป็นแบบจำลองสีวายไอคิว ซึ่งสามารถแสดงผลทั้งในจอภาพสีกับจอภาพขาวดำ โดยจอภาพขาวดำจะใช้เพียงองค์ประกอบ Y จากแบบจำลองสีวายไอคิวเท่านั้น เราสามารถเปลี่ยนค่าสีจากแบบจำลองสีอาร์จีบีให้เป็นสัญญาณสีในแบบจำลองวายไอคิวได้ดังนี้

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

องค์ประกอบ Y หมายถึง ลูมิแนนซ์ (Luminance) หรือค่าความส่องสว่าง เกิดจากการแปลงค่าสีแดง สีเขียว และสีน้ำเงินให้เป็นค่าความส่องสว่าง หรือค่าระดับความเทา นอกจากนี้สมการนี้ยังใช้ในการเปลี่ยนภาพสีให้เป็นภาพเกรย์สเกล ซึ่งระดับพลังงานของสัญญาณ Y จะเป็น 93% องค์ประกอบ I และ Q หมายถึง โครมิแนนซ์ (Chrominance) หรือค่าสัญญาณสี จะมีระดับพลังงานของสัญญาณเท่ากับ 5% และ 2% ตามลำดับ

2.1.3.4 แบบจำลองสีเอชไอเอส (The HIS color model) (สมจิต กติบแสง, 2539: 12-13) เป็นแบบจำลองสีที่นิยมใช้กันมาก เพราะสัมพันธ์กับพฤติกรรมการรับรู้สีของมนุษย์ จากรูปที่ 2.5 แสดงสามเหลี่ยมของแบบจำลองสีเอชไอเอส



รูปที่ 2.5 แสดงสามเหลี่ยมของแบบจำลองสีเอชไอเอส (Gonzalez และ Woods, 1992: 230)

เราสามารถเปลี่ยนค่าสีจากแบบจำลองสีอาร์จีบีให้เป็นสัญญาณสีในแบบจำลองเอชไอเอสโดยพิจารณาจากรูปที่ 2.5 กำหนดให้ R, G และ B แทนค่าสีแดง สีเขียว และสีน้ำเงินตามลำดับ องค์ประกอบ I (Intensity) คือ ค่าความเข้ม, H (Hue) คือค่าสีต้น และ S (Saturation) คือ ความอิ่มตัว สี ค่าความเข้มและค่าสีต้น สามารถหาได้จากสมการดังนี้

$$I = \frac{R+G+B}{3}$$

$$H = \arccos \frac{0.5 * (R-G) + (R-B)}{(R-G)^2 + (R-B) * (G-B)}^{\frac{1}{2}}$$

ถ้า $\frac{B}{I}$ มีค่ามากกว่า $\frac{G}{I}$ ค่ามุม H จะมากกว่า 180 องศา เนื่องจาก arccos มีค่าในช่วง 0 ถึง 180 องศา ดังนั้น H จะมีค่าเป็น $360 - H$ องศา ส่วนค่าอิ่มตัวสีเป็นระยะห่างของสีจากจุดศูนย์

กลางของสามเหลี่ยมเอชไอเอส ที่ที่อยู่ขอบนอกสุดจะมีความอึมดำมากที่สุด ค่าความอึมดำหาค่าได้จากสมการดังนี้

$$S = 1 - \frac{3 * \min(R, G, B)}{R + G + B}$$

2.2 การบีบข้อมูลภาพ (Image compression)

ในโลกเทคโนโลยีที่มีการรับส่งข้อมูลข่าวสารอยู่ตลอดเวลา โดยเฉพาะข้อมูลภาพซึ่งมีขนาดใหญ่ และใช้เวลาในการประมวลผลและเวลาในการรับส่งข้อมูลมาก จึงจำเป็นต้องมีกระบวนการบีบข้อมูลภาพเพื่อประหยัดเนื้อที่และเวลาที่ใช้

ความสนใจในการบีบข้อมูลภาพเริ่มขึ้นเมื่อประมาณกว่า 25 ปีที่แล้วในโลกของแอนะล็อกเพื่อลดแบนวิธ (Bandwidth) ในการส่งสัญญาณวิดีโอ โดยกระบวนการนี้ถูกเรียกว่า การบีบแบนวิธ (Bandwidth compression) จากนั้นเมื่อเทคโนโลยีด้านดิจิทัลพัฒนาขึ้นทำให้เกิดการบีบข้อมูลทางดิจิทัลขึ้น และได้รับการพัฒนาเพิ่มขึ้นอย่างรวดเร็ว (Gonzalez และ Woods, 1992: 307)

หลักในการบีบข้อมูลภาพคือ พยายามลดหรือกำจัดส่วนของข้อมูลที่เกินความจำเป็นหรือซ้ำซ้อนกัน (Data redundancy) โดยยังคงข่าวสารไว้ได้เหมือนเดิม ซึ่งจะทำให้ข้อมูลภาพมีขนาดลดลงจากเดิม และสามารถนำภาพกลับมาแสดงภาพหลังโดยผ่านกระบวนการคลาย (Decompression)

ประเภทของการบีบข้อมูลภาพแบ่งได้เป็น 2 แบบ คือ การบีบข้อมูลแบบไม่มีการสูญเสีย (Lossless compression หรือ Bit-preserving หรือ Versible compression) และการบีบข้อมูลแบบมีการสูญเสียบางส่วน (Lossy compression หรือ Irreversible compression) (Nelson และ Gailly, 1996)

2.2.1 แบบจำลองระบบการบีบ (Compression System Model) (Umbaugh, 1998: 242-247) ประกอบด้วย 2 ส่วน คือ การบีบ (Compression) และ การคลาย (Decompression)

2.2.1.1 การบีบ จากรูปที่ 2.6 แสดงแบบจำลองระบบการบีบในส่วนของ การบีบ ซึ่งประกอบด้วย

1. ขั้นตอนก่อนการประมวลผล (Preprocessing stage) เป็นขั้นตอนก่อนนำภาพไปเข้ารหัสโดยการแปลงรูปแบบของภาพ (Transformation) หรือแยกส่วนประกอบภาพ (Image decomposition) ให้เหมาะสมแก่การเข้ารหัส ประกอบด้วยขั้นตอนย่อยๆ ดังนี้

1.1 การลดทอนข้อมูล (Data reduction) เป็นขั้นตอนของการลดค่าระดับความเทาหรือลดข้อมูลที่มีความซ้ำซ้อนกัน

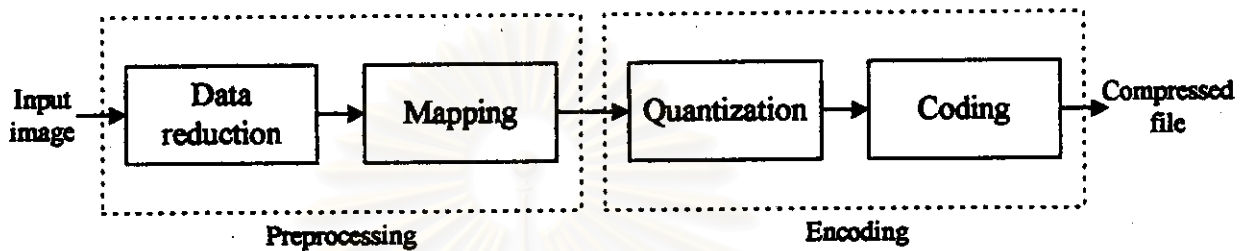
1.2 แมพทิ่ง (Mapping) คือ การแมพหรือแปลงข้อมูลภาพที่ลดความซ้ำซ้อนของข้อมูลแล้วลงในรูปแบบทางคณิตศาสตร์แบบต่างๆ ที่ง่ายแก่การบีบข้อมูล โดยแปลงข้อมูลภาพจากพื้นที่สัญญาณเป็นพื้นที่การแปลงซึ่งจะอยู่ในรูปสัมประสิทธิ์การแปลง เช่น การแปลงฟูร์เรียร์หรือเอฟเอฟที (Fast Fourier Transform หรือ FFT), คิฟเฟอเรนเชียลพัลส์โคดมอดูเลชันหรือดีพีซีเอ็ม (Differential Pulse Code Modulation หรือ DPCM), การแปลงคอสคริทโคไซน์หรือดีซีที (Discrete Cosine Transform หรือ DCT), ซับแบนด์ (Subband) หรือ เวฟเลต เป็นต้น

2. ขั้นตอนการเข้ารหัส (Encoding stage) ประกอบด้วย

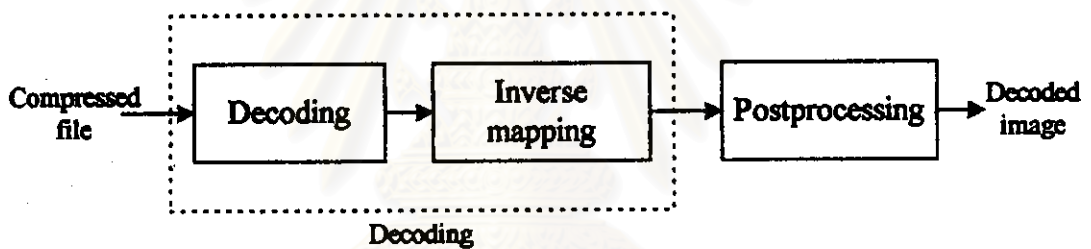
2.1 ควอนไทเซชัน เป็นขั้นตอนการแปลงข้อมูลจากรูปแบบข้อมูลต่อเนื่อง (Continuous data) มากี่บดลงในรูปแบบข้อมูลไม่ต่อเนื่อง (Discrete data) โดยลดค่าสัมประสิทธิ์ลงเพื่อลดจำนวนบิตของข้อมูลลง เป็นขั้นตอนที่อยู่ในการบีบแบบสูญเสียข้อมูลบางส่วนเท่านั้น (ดูรูปที่ 2.8 ซึ่งแสดงการเปรียบเทียบขั้นตอนต่างๆ ของทั้ง 2 แบบ) และเป็นจุดที่เกิดข้อผิดพลาดหรือสูญเสียข้อมูลของภาพในระหว่างการบีบ ชนิดและดีกรี (Degree) ของควอนไทเซชันมีผลกระทบต่ออย่างสูงกับอัตราบิต (Bit rate) และคุณภาพของภาพที่สร้างขึ้นใหม่ ตัวอย่างของควอนไทเซชัน เช่น ยูนิฟอร์มหรือนอนยูนิฟอร์มสเกลาร์ควอนไทเซชัน (Uniform หรือ Nonuniform scalar quantization) หรือ เวกเตอร์ควอนไทเซชัน (Vector quantization)

2.2 การเข้ารหัสสัญลักษณ์ (Symbol encoding) คือ การเข้ารหัสสัมประสิทธิ์ที่ได้ผ่านขั้นตอนควอนไทเซชัน โดยเก็บอยู่ในรูปแบบชุดสายอักขระ (string) ของ 0 และ 1 ซึ่งสามารถส่งออกไปในเครือข่ายหรือเก็บข้อมูล สามารถเข้ารหัสแบบรหัสคง (ขนาดรหัสเท่ากันหมด) (fixed-length code หรือ equal length code) หรือแบบรหัสความยาวแปรได้ (ขนาดรหัสไม่เท่ากัน) (variable-length code หรือ unequal length code) โดยปกติ การเข้ารหัสแบบรหัสความยาวแปรได้ จะมีประสิทธิภาพในการบีบสูงกว่า แต่จะมีค่าโอเวอร์เฮด (overhead) มากในขั้นตอนการ

เข้ารหัสและถอดรหัส ขั้นตอนวิธีในการบีบภาพไม่จำเป็นจะต้องประกอบไปด้วยทุกๆ ขั้นตอน อาจจะมีเพียงหนึ่งหรือสองขั้นตอนหรืออาจจะรวมขั้นตอนต่างๆ เข้าไว้ด้วยกันก็ได้ เช่น ระบบการบีบค่าแตกต่างของวิธีโคสิซิม์ จะรวมขั้นตอนการแปลงและควอนไทซ์เป็นขั้นตอนเดียวกันหรือประมวลผลพร้อมกัน เป็นต้น



รูปที่ 2.6 แสดงแบบจำลองระบบการบีบในส่วนของการบีบ (Umbaugh, 1998: 243)



รูปที่ 2.7 แสดงแบบจำลองระบบการบีบในส่วนของการคลาย (Umbaugh, 1998: 244)

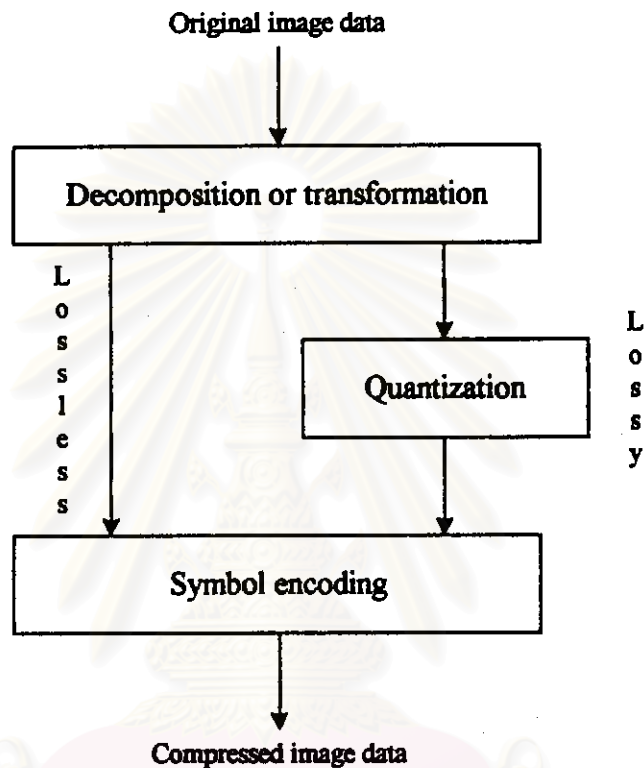
2.2.1.2 การคลาย จากรูปที่ 2.7 แสดงแบบจำลองระบบการบีบในส่วนของการคลาย ซึ่งประกอบด้วย

1. ขั้นตอนการถอดรหัส (Decoding stage) ประกอบด้วย

1.1 การถอดรหัสสัญลักษณ์ (Symbol decoding) จะนำภาพที่บีบแล้วถอดรหัสเพื่อให้ได้ค่าสัมประสิทธิ์กลับมา

1.2 อินเวอร์ตแมพพิ่ง (Inverse mapping) เป็นการทำย้อนกลับกับแมพพิ่งในขั้นตอนการเข้ารหัส โดยให้ค่าสัมประสิทธิ์กลับมีค่าเท่าเดิม ข้อมูลภาพจะถูกแปลงกลับจากพื้นที่การแปลงเป็นพื้นที่สัญญาณ

2. ขั้นตอนหลังการประมวลผล (Postprocessing stage) ในบางกรณีอาจมีการประมวลผลบางอย่าง ที่ต้องประมวลผลย้อนกลับกับขั้นตอนก่อนการประมวลผล เช่น การขยายส่วนภาพที่เกิดจากการลดความซ้ำซ้อนของข้อมูล เป็นต้น เพื่อให้ข้อมูลภาพสมบูรณ์ยิ่งขึ้น



รูปที่ 2.8 แสดงการบีบข้อมูลภาพระหว่างแบบ ไม่มีการสูญเสียข้อมูลและแบบสูญเสียข้อมูลบางส่วน (Dougherty, 1994: 264)

2.2.2 การบีบข้อมูลแบบไม่มีการสูญเสีย (Nelson และ Gailly, 1996) (ดวงอาทิตย์ ศรีมุต และ ชงยุทธ เอกอริชทรัพย์, 2539: 9-14)

การบีบข้อมูลแบบนี้ เมื่อบีบข้อมูลและคลายกลับจะได้ข้อมูลกลับมาที่เหมือนเดิมกับต้นฉบับ วิธีนี้ให้อัตราส่วนการบีบต่ำแต่ข้อมูลไม่มีการสูญเสีย เหมาะสำหรับข้อมูลที่ผิดพลาดไม่ได้เลย เช่น โปรแกรมคอมพิวเตอร์ แฟ้มข้อมูลของฐานข้อมูล แฟ้มข้อมูลแผนตารางทำการ (Spreadsheets) แฟ้มข้อมูลการประมวลคำ (Word processing) ข้อมูลที่ส่งตามเครื่องแฟกซ์ เป็นต้น ซึ่งข้อมูลที่เก็บในแฟ้มเหล่านี้ล้วนมีความสำคัญทุกๆ ไบต์ การบีบข้อมูลแบบนี้ มีอยู่ 2 ประเภท คือ เทคนิคคิกซ์นาร์เบต (Dictionary-based techniques) และการเข้ารหัสเชิงสถิติ (Statistical encoding methods)

เทคนิคคลิกชั้นนารีเบตจะสร้างแฟ้มข้อมูลที่เก็บรหัสแบบตรงซึ่งจะมีจำนวนบิตของแต่ละรหัสคงที่ และเท่ากันหมด (มักจะเป็น 12 หรือ 16 บิต) และแต่ละรหัสจะแทนชุดข้อมูลในแฟ้มข้อมูลต้นฉบับ

ส่วนการเข้ารหัสเชิงสถิติจะใช้วิธีแทนข้อมูลที่เกิดขึ้นบ่อยๆ ในแฟ้มข้อมูลด้วยจำนวนบิตที่ น้อยกว่าจำนวนบิตที่ใช้แทนข้อมูลที่เกิดขึ้น ไม่บ่อยนัก ซึ่งจะได้ขนาดรหัสไม่เท่ากัน โดยพิจารณา จากความน่าจะเป็นในการพบข้อมูลนั้นๆ

2.2.2.1 เทคนิคคลิกชั้นนารีเบต ในเทคนิคนี้รหัสที่ใช้แทนข้อมูลแต่ละชุดจะมีความยาว เท่ากันหมด จึงเป็นไปได้ที่รหัสนั้นจะแทนชุดข้อมูลจริงที่มีความยาวมากกว่าความยาวของรหัส หรือแทนข้อมูลจริงที่มีความยาวน้อยกว่าความยาวรหัสก็ได้ เช่น รหัสขนาด 12 บิต อาจใช้แทนชุด ข้อมูลที่มีขนาดความยาว 64 บิต หรืออาจจะแทนข้อมูลเพียง 8 บิต ก็ได้ ดังนั้นขนาดรวมของข้อมูล อาจจะเพิ่มขึ้นหรือลดลงก็ได้ขึ้นกับวิธีการ และลักษณะของข้อมูล การเข้ารหัสวิธีนี้มีหลายแบบ ได้แก่

2.2.2.1.1 การเข้ารหัสแบบรันเลงท์หรืออาร์แอลซี (Run-Length coding or RLC) (Umbaugh, 1998: 253-257) เป็นวิธีการเข้ารหัสที่ง่ายที่สุดของเทคนิคคลิกชั้นนารีเบต ด้วยหลักการที่ ง่ายส่วนใหญ่ บริเวณจุดภาพใกล้เคียงกันจะมีระดับความเทาหรือความเข้มแสงเหมือนกัน ดังนั้น แทนที่จะเก็บค่าสีตัวกันซ้ำๆ หลายๆ ครั้ง จึงเก็บค่านั้นเพียงครั้งเดียวแล้วตามด้วยค่าที่บอกจำนวนที่ ค่านั้นวิ่งไป (Run-Length) วิธีอาร์แอลซีนีถูกนำมาใช้กับภาพลักษณะฐานสองและภาพเกรย์สเกล ให้ อัตราส่วนการบีบสูงเมื่อใช้กับภาพที่มีวัตถุอยู่บนพื้นที่เป็นสีเดียวกันทั้งหมด

สำหรับภาพลักษณะฐานสองสามารถใช้ฮอริซอนทัลอาร์แอลซี (Horizontal RLC) นับตามแนวนอน(row) หรือเวอติคัลอาร์แอลซี (Vertical RLC) นับตามแนวตั้ง (column) ตัวอย่าง ภาพลักษณะฐานสองขนาด 8x8 ซึ่งต้องการ 3 บิตสำหรับรหัสแต่ละตัว ดังตารางที่ 2.1

เมื่อใช้วิธีฮอริซอนทัลอาร์แอลซีจะได้	บรรทัดแรก	8
	บรรทัดสอง	0, 4, 4
	บรรทัดสาม	1, 2, 5
	บรรทัดสี่	1, 5, 2
	บรรทัดห้า	1, 3, 2, 1, 1

บรรทัดหก	2, 1, 2, 2, 1
บรรทัดเจ็ด	0, 4, 1, 1, 2
บรรทัดแปด	8

ที่บรรทัดสองและบรรทัดเจ็ด ตัวแรกเป็น 0 หมายถึง มีบิต 0 จำนวน 0 ตัว

0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0
0	1	1	0	0	0	0	0
0	1	1	1	1	1	0	0
0	1	1	1	0	0	1	0
0	0	1	0	0	1	1	0
1	1	1	1	0	1	0	0
0	0	0	0	0	0	0	0

ตารางที่ 2.1 แสดงตัวอย่างภาพลักษณะพื้นฐานสองขนาด 8x8

เมื่อนำวิธีนี้ใช้กับภาพเกรย์สเกล จะเรียกรูปนี้ว่าบิตเพลนอาร์แอลซี (Bit-Plane RLC) วิธีนี้ทำงานโดยพื้นฐานจากวิธีอาร์แอลซี ให้แต่ละบิตเพลน (bit plane) เข้ารหัสเป็นอิสระต่อกัน และแทนที่จะใช้ทศนิยมในการวิ่ง ก็จะใช้สองค่าในการวิ่ง คือ (G, L) โดยที่ G คือ gray-scale value หรือค่าระดับความเทา และ L คือ run-length หรือจำนวนบิตที่มีค่าระดับความเทาเท่ากันโดยที่ไม่มีการบีบแบบอื่นเกี่ยวข้องด้วย ภาพ 8 บิตสีเดียวจะให้อัตราส่วนการบีบ 0.5 ถึง 1.2 ต่อ 1 เทคนิคนี้มีประสิทธิภาพดีกับภาพที่มีระดับความเทาจำนวนน้อย ตัวอย่างภาพเกรย์สเกล ขนาด 8x8 ที่มี 4 บิตต่อจุดภาพ ดังตารางที่ 2.2

เมื่อเข้ารหัสจะได้	บรรทัดแรก	10, 8
	บรรทัดสอง	10, 5, 12, 3
	บรรทัดสาม	10, 5, 12, 3
	บรรทัดสี่	0, 3, 10, 3, 0, 2
	บรรทัดห้า	5, 3, 0, 5
	บรรทัดหก	5, 3, 10, 2, 9, 2, 10, 1
	บรรทัดเจ็ด	5, 3, 4, 3, 0, 2
	บรรทัดแปด	0, 8

และจะถูกเก็บในรูปแบบการบีบอาร์แอลซีดังนี้ 10, 8, 10, 5, 12, 3, 10, 5, 12, 3, 0, 3, 10, 3, 0, 2, 5, 3, 0, 5, 5, 3, 10, 2, 9, 2, 10, 1, 5, 3, 4, 3, 0, 2, 0, 8

10	10	10	10	10	10	10	10
10	10	10	10	10	12	12	12
10	10	10	10	10	12	12	12
0	0	0	10	10	10	0	0
5	5	5	0	0	0	0	0
5	5	5	10	10	9	9	10
5	5	5	4	4	4	0	0
0	0	0	0	0	0	0	0

ตารางที่ 2.2 แสดงตัวอย่างภาพเกรย์สเกลขนาด 8x8

วิธีนี้ไม่ใช่เทคนิคการบีบที่คิดสำหรับภาพจับจ้อง แต่มันมีประโยชน์กับภาพง่าๆ เช่น แผนที่ภาพถ่ายฟิสิก ผลลัพธ์จากการบีบโดยใช้วิธีนี้สามารถประยุกต์ให้ดีขึ้นโดยการลดจำนวนระดับความเทาก่อนการประมวลผล แต่การบีบจะไม่เป็นแบบสูญเสียข้อมูลอีก

มาตรฐานของอาร์แอลซีเสนอโดยไอทียูทีอาร์ (ITU-R หรือ The International Telecommunications Union-Radio) แต่เดิมเป็นซีซีทีไอทีที (CCITT หรือ Consultative Committee for International Telegraph and Telephone) มาตรฐานนี้เดิมถูกออกแบบมาเพื่อใช้กับการส่งแฟกซ์ แต่ได้กลายเป็นที่นิยมกับการบีบภาพลักษณะฐานสองโดยใช้วิธีฮัฟแมนหรืออาร์แอลซี เมื่อมีการประมวลผลผลลัพธ์ด้วยวิธีอาร์แอลซีร่วมกับวิธีการเข้ารหัสฮัฟแมน (Huffman Coding) จะทำให้ขนาดแฟ้มข้อมูลลดลงในอัตราส่วนการบีบที่ 15 ถึง 20 ค่อนข้าง

2.2.2.1.2 การเข้ารหัสแบบแลมเพิลซีเวลชหรือแอลแซคดับเบิล (Lempel-Ziv-Welch coding or LZW) (Nelson และ Gailly, 1996: 255-265) (Dougherty, 1998: 272-274) การเข้ารหัสแอลแซค (LZ coding) เป็นเทคนิคการบีบแบบไม่มีการสูญเสียที่ Jacob Ziv และ Abraham Lempel คิดขึ้นเป็นครั้งแรกในปี 1977-1978 คือ LZ77 และ LZ78 ต่อมา Terry Welch ได้นำมาปรับปรุงเพิ่มเติมจนกลายเป็นแอลแซคดับเบิลซึ่งมีวิธีการคล้ายกับอาร์แอลซี เป็นวิธีที่นิยมมากวิธีหนึ่งซึ่งกลายเป็นมาตรฐานให้กับการเก็บภาพแบบต่างๆ เช่น ทิฟ (TIFF), จีฟ (GIF) เป็นต้น รวมทั้งโปรแกรมการบีบข้อมูล เช่น พีเคซีพ (PKZIP) เพราะเป็นวิธีที่บีบข้อมูลแบบไม่มีการสูญเสีย เร็ว มีประสิทธิภาพ และสามารถกระทำกับภาพที่มีความลึกบิต (bit-depth) เท่าใดก็ได้ อีกสาเหตุสำคัญที่

ทำให้เทคนิคแอดแซคดับบลิวเป็นที่นิยม คือ ไม่มีความจำเป็นต้องส่งตารางสายอักขระ (string table) ให้กับแอดแซคข้อมูลผลลัพธ์หรือส่งให้ฝ่ายรับข้อมูล เพราะตารางสายอักขระนี้สามารถสร้างขึ้นมาได้ทั้งในระหว่างการเข้ารหัสและการถอดรหัสโดยใช้ข้อมูลที่ถูกรหัสที่เข้ามาแล้ว ดังนั้นแอดแซคข้อมูลที่เก็บจะมีขนาดเล็ก

เทคนิคนี้จะทำการเข้ารหัสสายของอักขระ โดยสร้างตารางของสายอักขระ (codebook) และรหัสที่ใช้แทนสายอักขระ (codeword) เทคนิคนี้จะสร้างสายอักขระผลลัพธ์และปรับเปลี่ยนค่าในตารางของสายอักขระอยู่ตลอดเวลา ดังนั้นตารางสายอักขระจะเปลี่ยนแปลงอักขระหรือสายอักขระต่างๆ ของภาพที่จะเข้ารหัส ตารางสายอักขระนี้สามารถปรับขึ้นได้ถึง 4096 รหัส และแต่ละรหัสสามารถปรับได้ถึง 12 บิต (กรณีเพิ่มข้อมูลมีขนาด 8 บิต) โดยใช้ 256 รหัสแรกในการแทนรหัสที่เกิดขึ้นได้ของไบต์เดียว รหัสหมายเลขที่ 256 เก็บรหัสพิเศษคือรหัสว่าง (Clear code) รหัสหมายเลขที่ 257 เก็บ End of information (EOI) ซึ่งแสดงจุดสิ้นสุด (end of file) ของภาพ รหัสหมายเลขที่ 258 ถึง หมายเลข 4095 เก็บสายของอักขระหรือรูปแบบของสายอักขระที่พบในข้อมูลขณะทำการบีบ

Characters Input	Code Output	New code value and associated string
"W"	' '	256 = "W"
"E"	'W'	257 = "WE"
"D"	'E'	258 = "ED"
" "	'D'	259 = "D "
"WE"	256	260 = "WE"
" "	'E'	261 = "E"
"WEE"	260	262 = "WEE"
" W"	261	263 = "E W"
"EB"	257	264 = "WEB"
" "	B	265 = "B"
"WET"	260	266 = "WET"
<EOF>	T	

ตารางที่ 2.3 แสดงตัวอย่างการเข้ารหัสแบบแอดแซคดับบลิว (Nelson และ Gailly, 1996: 263)

การเข้ารหัสเริ่มต้นโดยตารางสายอักขระจะเก็บอักขระที่ถูกรหัสแต่ละตัว ค่อยจากนั้นลำดับของข้อมูลเข้าจะถูกทดสอบครั้งละหนึ่งอักขระ โดยตรวจสอบกับสายอักขระที่ยาวที่สุดในตารางว่าพบหรือไม่ โดยสายอักขระส่วนที่พบก็จะถูกเข้ารหัสเป็นค่าชี้ (pointer) และใช้เป็นผลลัพธ์แทนการเข้ารหัส ถ้าไม่พบก็จะเพิ่มอักขระจากข้อมูลเข้ามาเพื่อสร้างสายอักขระใหม่และใส่ใน

ตารางสายอักขระ ข้อมูลเข้าก็จะถูกกระจายค่าออกไปเรื่อยๆ จนกระทั่งหมดเพิ่มข้อมูล จากตัวอย่างข้อมูลที่ให้นำมาเข้ารหัส ดังนี้ “ WED WE WEE WEB WET ” การเข้ารหัสจะแสดงดังตารางที่ 2.3 และเมื่อนำมาเข้ารหัสจะได้ “ WED<256>E<260><261><257>B<260>T”

ส่วนการถอดรหัสจะเริ่มต้นที่ตารางสายอักขระเหมือนกับการเข้ารหัส แต่ละรหัสที่ได้รับจะถูกแปลจากตารางสายอักขระลงในสายอักขระผลลัพธ์ ยกเว้นสัญลักษณ์แรก ทุกๆ ครั้งที่ได้รับรหัส ตารางสายอักขระจะมีการปรับเปลี่ยนตลอดเวลา การถอดรหัสจะสร้างตารางใหม่ที่เหมือนกับการเข้ารหัส จากตัวอย่างการเข้ารหัส เมื่อนำข้อมูลที่เข้ารหัสแล้วมาถอดรหัสจะได้ดังตารางที่ 2.4

Input/New code	Old code	String/Output	Character	New table entry
' '	' '	“ ”		
'W'	' '	“W”	'W'	256 = “W”
'E'	'W'	“E”	'E'	257 = “WE”
'D'	'E'	“D”	'D'	258 = “ED”
256	'D'	“W”	' '	259 = “D ”
'E'	256	“E”	'E'	260 = “WE”
260	'E'	“WE”	' '	261 = “E”
261	260	“E”	'E'	262 = “WEE”
257	261	“WE”	'W'	263 = “E W”
'B'	257	“B”	'B'	264 = “WEB”
260	'B'	“WE”	' '	265 = “B”
'T'	260	“T”	'T'	266 = “WET”

ตารางที่ 2.4 แสดงตัวอย่างการถอดรหัสแบบแอดแมชคียบบิตว (Nelson และ Gailly, 1996: 265)

2.2.2.2 การเข้ารหัสเชิงสถิติ การเข้ารหัสวิธีนี้เป็นอีกเทคนิคที่ใช้กันมาก โดยอาศัยความน่าจะเป็นของการพบตัวอักษรหรือสัญลักษณ์นั้นๆ ถ้าตัวอักษรหรือสัญลักษณ์นั้นพบบ่อยก็จะมีรหัสสั้น แต่ถ้าตัวอักษรหรือสัญลักษณ์นั้นพบไม่บ่อยนักก็จะมีรหัสที่ยาวกว่า ดังนั้นจะเห็นได้ว่ารหัสที่ใช้ในวิธีนี้จะไม่ได้มีขนาดเท่ากัน การเข้ารหัสวิธีนี้ ได้แก่

2.2.2.2.1 การเข้ารหัสฮัฟฟ์แมน (Dougherty, 1994) (Gonzalez และ Woods, 1992: 343-348) เป็นเทคนิคที่นิยมที่สุดสำหรับการเข้ารหัสที่ขจัดความซ้ำซ้อนของข้อมูล พัฒนขึ้นโดย D.Huffman ในปี 1952 สามารถลดขนาดไฟล์ลงได้ 10 ถึง 50% (1.1: 1 ถึง 1.5: 1) บางครั้งอาจให้ถึง

2:1 หรือ 3:1 วิธีนี้จะใช้ต้นไม้แบบทวิภาค (Binary tree) โดยกำหนดรหัสแทนค่าที่เกิดบ่อยๆ ด้วยจำนวนบิตที่น้อยกว่าค่าที่ไม่ค่อยเกิดขึ้น ขั้นตอนการเข้ารหัสสามารถอธิบายได้ดังนี้

ขั้นตอนเริ่มต้นด้วยการสร้างชุดของการลดทอนต้นฉบับ (Source reduction) โดยเรียงลำดับตามความน่าจะเป็น (Probability) ของสัญลักษณ์ต่างๆ จากนั้นรวมค่าความน่าจะเป็นของสัญลักษณ์สองตัวที่น้อยที่สุดแล้วใส่ในการลดทอนต้นฉบับลำดับถัดไป ขั้นตอนนี้อธิบายได้จากตารางที่ 2.5 ที่ด้านซ้ายของตารางมีชุดของสัญลักษณ์ต้นฉบับ (Source symbol) และค่าความน่าจะเป็นที่ลดลง ค่าของความน่าจะเป็นสองตัวต่ำสุด คือ 0.06 และ 0.04 จะถูกรวมกันโดยมีค่าความน่าจะเป็น 0.1 ขั้นตอนนี้จะกระทำไปเรื่อยๆ จนกระทั่งเหลือสองสัญลักษณ์สุดท้ายดังเช่นที่มุมบนขวาของตาราง

ค่าจากต้นฉบับ		การลดทอนต้นฉบับ			
สัญลักษณ์	ความน่าจะเป็น	1	2	3	4
A2	0.4	0.4	0.4	0.4	0.6
A6	0.3	0.3	0.3	0.3	0.4
A1	0.1	0.1	0.2	0.3	
A4	0.1	0.1	0.1		
A3	0.06	0.1			
A5	0.04				

ตารางที่ 2.5 ขั้นตอนการลดทอนต้นฉบับของฮัฟฟ์แมน (Gonzalez และ Woods, 1992: 344)

ขั้นตอนที่สองคือการเข้ารหัสแต่ละสัญลักษณ์ โดยเริ่มต้นที่แหล่งข้อมูลที่เล็กที่สุด (จากตารางที่ 2.5 การลดทอนต้นฉบับลำดับที่ 4) ทำให้อ่อนกลับไปหาแหล่งข้อมูลต้นฉบับด้านขวาสุดได้รหัสที่สั้นที่สุดคือ 0 และ 1 จากตารางที่ 2.6 ให้สังเกตที่ความน่าจะเป็น 0.6 ซึ่งได้จากการรวมค่า 0.3 และ 0.3 เลขฐานสองถูกใช้เข้ารหัส และค่อท้ายด้วย 0 และ 1 ที่การลดทอนต้นฉบับลำดับที่ 3 ซึ่งจะกระทำต่อไปซ้ำๆ ในแต่ละการลดทอนต้นฉบับ จนกระทั่งได้ข้อมูลต้นฉบับ (Original source) รหัสที่ได้สุดท้ายจะอยู่ในตารางที่ 2.6 ทางด้านซ้าย

ค่าจากต้นฉบับ			การลดทอนต้นฉบับ			
สัญลักษณ์	ความน่าจะเป็น	รหัส	1	2	3	4
A2	0.4	1	0.4 1	0.4 1	0.4 1	0.6 0
A6	0.3	00	0.3 00	0.3 00	0.3 00	0.4 1
A1	0.1	011	0.1 011	0.2 010	0.3 01	
A4	0.1	0100	0.1 0100	0.1 011		
A3	0.06	01010	0.1 0101			
A5	0.04	01011				

ตารางที่ 2.6 การกำหนดรหัสในวิธีการเข้ารหัสฮัฟฟ์แมน (Gonzalez และ Woods, 1992: 345)

$$\begin{aligned}
 \text{ความยาวเฉลี่ยของรหัส (L}_{avg}\text{)} &= (0.4*1) + (0.3*2) + (0.1*3) + (0.1*4) + (0.06*5) + (0.04*5) \\
 &= 2.2 \text{ บิตต่อสัญลักษณ์}
 \end{aligned}$$

ด้วยวิธีของฮัฟฟ์แมนนี้ แม้ว่าความยาวของรหัสที่ใช้แทนแต่ละสัญลักษณ์จะไม่เท่ากัน เมื่อนำรหัสมาเรียงต่อกันก็ไม่ก่อให้เกิดความสับสนในการตีความรหัสแต่อย่างใด เช่น 010100111100 สัญลักษณ์แรกคือ 01010 ซึ่งก็คือรหัส A3 ถัดมาคือ 011 ซึ่งก็คือรหัส A1 ดังนั้นจะได้รหัส A3A1A2A2A6

2.2.2.2.2 การเข้ารหัสเลขคณิต (Arithmetic Coding) (Dougherty, 1994) (สุทธิศักดิ์ พงษ์นาทาณิช, 2538: 268-275) จากที่กล่าวมาในวิธีการเข้ารหัสฮัฟฟ์แมนยังไม่ใช่วิธีที่ดีที่สุด เพราะถึงจะเป็นวิธีการเข้ารหัสแบบความยาวแปรได้หรือไม่จำกัดจำนวนบิตก็ตาม จำนวนบิตที่ใช้ในการเข้ารหัสข้อมูลยังเป็นจำนวนเต็ม เช่น 1 หรือ 2 บิต เป็นต้น โดยการปิดหรือตัดเศษให้เป็นจำนวนเต็ม แต่วิธีการเข้ารหัสเลขคณิตจะให้จำนวนบิตในการเข้ารหัสไม่เป็นจำนวนเต็ม แต่จะเป็นจำนวนจริงที่ใช้ในการบีบ เช่น 2.5, 3.2 เป็นต้น ทั้งนี้เพราะทฤษฎีของการเข้ารหัสเลขคณิตเป็นการนำเลขทศนิยมจุดลอยตัว (floating-point) มาใช้ในการเข้ารหัสข้อมูลของสายอักขระ ดังนั้นถ้าหากสายอักขระมีความซับซ้อนและมีความยาวมาก ๆ จำนวนเลขนัยสำคัญของเลขจุดลอยตัว ก็จะมากขึ้นด้วย หรือหมายความว่าความละเอียด (precision) ของเลขจุดลอยตัวที่ได้ที่เข้ารหัสสายอักขระยิ่งสูงขึ้นด้วย ซึ่งเลขจุดลอยตัวนี้จะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าดับขั้นตอนการเข้ารหัสมีขั้นตอนดังนี้

ขั้นแรกข้อมูลเข้าจะถูกคำนวณค่าความน่าจะเป็นของแต่ละตัวอักษรในสายอักขระโดยให้เรียงลำดับตัวอักษรในสายอักขระที่เข้ารหัสจากน้อยไปหามาก(เรียงตามลำดับคัมค่ารหัสแอสกี) ตัวอย่างเช่น ข้อมูลที่ถูกเข้ารหัสมีดังนี้ "BILL GATES" สายอักขระประกอบด้วย 10 ตัวอักษร ดังนั้นค่าความถี่และค่าความน่าจะเป็นของแต่ละตัวอักษรจะมีค่าดังตารางที่ 2.7

ตัวอักษร	ค่าความถี่	ค่าความน่าจะเป็น	ช่วงความน่าจะเป็น
SPACE	1	0.1	$0.0 \leq p < 0.1$
A	1	0.1	$0.1 \leq p < 0.2$
B	1	0.1	$0.2 \leq p < 0.3$
E	1	0.1	$0.3 \leq p < 0.4$
G	1	0.1	$0.4 \leq p < 0.5$
I	1	0.1	$0.5 \leq p < 0.6$
L	2	0.2	$0.6 \leq p < 0.8$
S	1	0.1	$0.8 \leq p < 0.9$
T	1	0.1	$0.9 \leq p < 1.0$

ตารางที่ 2.7 แสดงความถี่และค่าความน่าจะเป็นจากตัวอย่างด้วยวิธีการเข้ารหัสเลขคณิต
(สุทธิศักดิ์ พงษ์รณาทานิช, 2538: 268)

ขั้นที่สอง เนื่องจากค่าความน่าจะเป็นของทุกๆ ตัวอักษรในสายอักขระเคียวกันจะต้องมีค่าในช่วง 0 ถึง 1 ดังนั้นเมื่อเราบวกค่าความน่าจะเป็นของทุกตัวในสายอักขระเคียวกันเข้าด้วยกันก็จะมีค่าเป็น 1 เสมอ ดังนั้นเราจึงสามารถนำค่าความน่าจะเป็นของแต่ละตัวอักษรมาจัดเรียงให้อยู่ในช่วงตั้งแต่ 0 ถึง 1 เพื่อเป็นการจัดลำดับขั้นของแต่ละตัวอักษร โดยที่ความกว้างของแต่ละขั้นอาจจะเท่ากันหรือไม่เท่ากันก็ได้ ทั้งนี้ก็ขึ้นกับค่าความถี่ของแต่ละตัวอักษร เมื่อนำตัวอักษรในสายอักขระมาจัดลำดับก็จะได้ผลลัพธ์ดังตารางที่ 2.7

ขั้นสาม เริ่มต้นจากการเข้ารหัสตัวอักษรตัวแรกของสายอักขระ ซึ่งจากตัวอย่างคือ "B" ดังนั้นรหัสที่ได้จะมีค่าตั้งแต่ 0.2 ถึง 0.3

ขั้นที่ ๒ เข้ารหัสข้อมูลตัวอักษรถัดๆ ไปของสายอักขระ ซึ่งจากตัวอย่างคือ “ILL GATES” โดยยึดหลักการว่า ช่วงตัวเลขจุดยอดตัวของรหัสที่ถูกเข้ารหัสจะได้จากช่วงค่าความถี่ของตัวอักษรนั้นๆ ที่ถูกนำมาคิดลงในช่วงค่าต่ำ-ค่าสูงของตัวอักษรที่ผ่านมา ดังสมการนี้

$$\text{ช่วงความน่าจะเป็น} = \text{ค่าสูง} - \text{ค่าต่ำ}$$

$$\text{ค่าสูงของตัวอักษรที่ถูกเข้ารหัส} = \text{ค่าความถี่ช่วงต่ำของตัวอักษรที่ผ่านมา} + (\text{ช่วงความน่าจะเป็น} * \text{ค่าความถี่ช่วงสูงของตัวอักษรที่ถูกเข้ารหัส})$$

$$\text{ค่าต่ำของตัวอักษรที่ถูกเข้ารหัส} = \text{ค่าความถี่ช่วงต่ำของตัวอักษรที่ผ่านมา} + (\text{ช่วงความน่าจะเป็น} * \text{ค่าความถี่ช่วงต่ำของตัวอักษรที่ถูกเข้ารหัส})$$

ตัวอักษร	ค่าต่ำ	ค่าสูง
	0.0	1.0
B	0.2	0.3
I	0.25	0.26
L	0.256	0.258
L	0.2572	0.2576
SPACE	0.25720	0.25724
G	0.257216	0.257220
A	0.2572164	0.2572168
T	0.25721676	0.25721680
E	0.257216772	0.257216776
S	0.2572167752	0.2572167756

ตารางที่ 2.8 แสดงค่าสูงและค่าต่ำจากตัวอย่างด้วยวิธีการเข้ารหัสเลขคณิต
(สุทธิศักดิ์ พงษ์ธนพานิช, 2538: 268)

เช่น จากตัวอย่างสำหรับ “B” ช่วงความน่าจะเป็น = $1.0 - 0.0 = 1.0$

ค่าสูงของตัวอักษรที่ถูกเข้ารหัส = $0.0 + (1.0 * 0.3) = 0.3$

ค่าต่ำของตัวอักษรที่ถูกเข้ารหัส = $0.0 + (1.0 * 0.2) = 0.2$

ให้นำค่าต่ำของการเข้ารหัสตัวอักษรตัวสุดท้ายเป็นผลลัพธ์ ดังนั้น จากตารางที่ 2.8 ซึ่งแสดงความถี่และค่าความน่าจะเป็นจากตัวอย่าง ผลลัพธ์การเข้ารหัสจะมีค่าเท่ากับ 0.2572167752

เลขจำนวนจริงเข้ารหัส	ตัวอักษร	ค่าต่ำ	ค่าสูง	ช่วงความน่าจะเป็น
0.2572167752	B	0.2	0.3	0.1
0.572167752	I	0.5	0.6	0.1
0.72167752	L	0.6	0.8	0.2
0.6083876	L	0.6	0.8	0.2
0.041938	SPACE	0.0	0.1	0.1
0.41938	G	0.4	0.5	0.1
0.1938	A	0.2	0.3	0.1
0.938	T	0.9	1.0	0.1
0.38	E	0.3	0.4	0.1
0.8	S	0.8	0.9	0.1
0.0				

ตารางที่ 2.9 แสดงการถอดรหัสด้วยวิธีการเข้ารหัสเลขคณิต (สุทธิศักดิ์ พงษ์รนาพาณิช, 2538: 268)

ส่วนการถอดรหัสทำได้โดยวิธีการถอดรหัสทุกๆ ตำแหน่งทศนิยมของเลขจุดลอยตัวด้วยวิธีวนซ้ำคั้งขั้นตอนต่อไปนี้

เริ่มต้นด้วยการค้นหาตัวอักษรตัวแรกของสายข้อมูลที่เข้ารหัส โดยดูจากเลขจุดลอยตัวที่มีค่าอยู่ในช่วงความน่าจะเป็นช่วงใด เช่น จากตัวอย่าง เลขจำนวนจริง 0.2572167752 ซึ่งมีค่าอยู่ระหว่าง 0.2-0.3 ดังนั้นอักษรตัวแรกของสายอักขระคือ "B" เพราะ "B" มีค่าความน่าจะเป็นอยู่ในช่วง 0.2-0.3 ต่อจากนั้นคำนวณดังสมการนี้

เลขจุดลอยตัวเข้ารหัส = (เลขจุดลอยตัวเข้ารหัส - ค่าต่ำของตัวอักษรที่ถูกถอดรหัส) / ช่วงค่าความน่าจะเป็น เช่น จากตัวอย่างคือ $(0.2572167752 - 0.2) / 0.1 = 0.572167752$

ต่อมาแนวคิดจุดยอดตัวเข้ารหัสมาทำการถอดรหัสตัวต่อๆ ไป โดยมีขั้นตอนตาม สมการ โดยดูว่าค่าของจุดยอดตัวมีค่าอยู่ในช่วงความน่าจะเป็นต่ำและสูงของตัวอักษรอะไร ให้กระทำ ไปเรื่อยๆ จนกว่าค่าของเลขจำนวนจริงเข้ารหัสจะมีค่าเป็น 0 ซึ่งหมายความว่า สายอักขระได้ถูก ถอดรหัสจนหมดแล้ว จากนั้นคอนดิงกล่าว สามารถถอดรหัสได้ดังตารางที่ 2.9

2.2.3 การบีบข้อมูลแบบมีการสูญเสียบางส่วน (Nelson และ Gailly, 1996)

การบีบข้อมูลแบบมีการสูญเสียเป็นการบีบโดยยอมให้มีการสูญเสียรายละเอียดของข้อมูล บางส่วน ซึ่งโดยทั่วไปมักจะถูกนำไปใช้กับการบีบข้อมูลแอนะล็อกที่ถูกจัดเก็บในรูปแบบดิจิทัล เช่น ข้อมูลภาพและเสียง ข้อมูลประเภทนี้มักจะมีขนาดใหญ่มาก ซึ่งการบีบโดยวิธีนี้จะให้อัตราส่วนการบีบ ที่สูงมาก แม้จะมีความคิดที่สนใจของข้อมูลไปบางส่วน แต่ข้อมูลที่ได้อีกยังคงสมบูรณ์ในระดับใช้งาน ได้ โดยปกติ การบีบข้อมูลแบบนี้จะมีขั้นตอนที่สำคัญ 2 ส่วน คือ การเปลี่ยนข้อมูลภาพให้อยู่ในรูป สัญญาณระดับความถี่สูงโดยใช้ฟังก์ชันการประมวลผลทางคณิตศาสตร์ซึ่งจะแปลงสัญญาณเหล่านี้ ให้อยู่ในรูปโดเมนของความถี่ (Frequency domain) โดยใช้ขั้นตอนวิธีคล้ายๆ กับการแปลงฟูริเยร์หรือเฟอเฟท และต่อมาข้อมูลความถี่จะถูกบีบโดยใช้เทคนิคการบีบข้อมูลแบบไม่มีการสูญเสีย ในช่วงการเปลี่ยนข้อมูลภาพให้อยู่ในรูปความถี่ ตามมาตรฐานมีการกำหนดค่าแฟกเตอร์ควอลิตี (Quality factor) และยอมให้ควมบีบข้อมูลสามารถกำหนดค่าแฟกเตอร์คุณภาพที่จะใช้ ถ้าต้องการให้ ภาพที่บีบมีคุณภาพสูงเท่าไร ข้อมูลมีการสูญเสียน้อย ข้อมูลที่นำมาเข้ารหัสก็จะมาก ซึ่งจะให้อัตราส่วนการบีบต่ำ

การบีบข้อมูลแบบนี้มีบทบาทมากในระบบสื่อหลายแบบ เช่น การบีบข้อมูลภาพทั้งแบบ ภาพนิ่ง (Still pictures), ภาพเคลื่อนไหว (Moving pictures) และข้อมูลเสียง (Audio) โดยมีองค์การ ระหว่างประเทศว่าด้วยการออกมาตรฐาน เช่น ไอโซ (The International Standards Organization หรือ ISO) และจีไอทีทีที่ร่วมมือกันจัดตั้งคณะกรรมการขึ้นมา 2 ชุด คือ The Joint Photographic Experts Groups (JPEG) ซึ่งเสนอมาตรฐานการบีบข้อมูลภาพเจพีค (JPEG) และ The Moving Pictures Experts Group (MPEG) เสนอมาตรฐานการบีบข้อมูลทั้งภาพและเสียงเอ็มพีคทู (MPEG-II)

การบีบภาพประเภทนี้มีหลายแบบ เช่น ดิฟเฟอเรนเชียลคอดิงหรือดีทีซีเอ็ม การเข้ารหัสการแปลง (Transform image compression) เป็นต้น ซึ่งในวิทยานิพนธ์นี้จะวิจัยในส่วน

การบีบภาพแบบสูญเสียข้อมูลบางส่วน ในแบบการเข้ารหัสการแปลง ได้แก่ เทคนิคการบีบภาพแบบ JPEG และเวฟเลต

2.2.3.1 ดิฟเฟอเรนเชียลพัลส์โคดมอดูเลชันหรือดีพีซีเอ็ม (Dougherty, 1994: 283-285) (Gonzalez และ Woods, 1992: 358-365) (Umbaugh, 1998: 270-275) เป็นวิธีที่มีพื้นฐานจากเทคนิคการเข้ารหัสทำนาย (Predictive coding) เริ่มจากพัลส์โคดมอดูเลชันหรือพีซีเอ็ม (Pulse code modulation หรือ PCM) ซึ่งถูกคิดค้นโดย Reeves ในปี 1938 ใช้ในการส่งข้อมูลสัญญาณแอนะล็อกไปเป็นสัญญาณดิจิทัล เหมาะกับภาพสัญญาณวิดีโอหรือสัญญาณเสียง ค่อมานแทนที่จะควอนไทซ์แต่ละสัญญาณที่อ่านเข้ามา จะใช้วิธีทำนายค่าจุดภาพถัดไปและควอนไทซ์ค่าแตกต่างกันระหว่างค่าที่จริงและค่าทำนาย ซึ่งค่าแตกต่างระหว่างค่าที่จริงและค่าทำนาย เรียกว่า ผลต่างการทำนายหรือค่าผิดพลาด (Prediction difference หรือ error term) วิธีการดังกล่าวเป็นพื้นฐานของวิธีเดลตามอดูเลชัน (Delta modulation) คิดค้นโดย Deloraine ในปี 1946 ต่อมาในปี 1952 วิธีดีพีซีเอ็มก็เกิดขึ้นโดย Cudler โดยแท้จริงแล้ว วิธีเดลตามอดูเลชันเป็นกรณีพิเศษหนึ่งของวิธีดีพีซีเอ็ม คือ เป็นกรณีหนึ่งบิตควอนไทเซอร์ (1-bit quantizer) วิธีเดลตามอดูเลชันเป็นวิธีที่เหมาะสมกับภาพที่ไม่มีการเปลี่ยนแปลงค่าจุดภาพรวดเร็วนักหรือภาพที่มีความถี่ต่ำ ค่าความแตกต่างที่ถูกเข้ารหัสจะใช้เพียง 1 บิตเท่านั้น (Lynch, 1985: 98)

การเข้ารหัสวิธีดีพีซีเอ็ม จากรูปที่ 2.9 ซึ่งแสดงการบีบภาพแบบดีพีซีเอ็ม เริ่มต้นโดยการอ่านค่าที่ละหนึ่งจุดภาพต่อหนึ่งกระบวนการ จากนั้นเก็บค่าจุดภาพแรกโดยไม่ต้องทำนาย หลังจากนั้นทำนายค่าจุดภาพถัดไปจากค่าที่อ่านไปข้าง โดยที่มีพื้นฐานจากจุดภาพค่าก่อนๆ และเข้ารหัสค่าความแตกต่างระหว่างค่าที่จริงและค่าทำนาย ดังนั้นจากรูปที่ 2.9 ตัวทำนาย (Predictor) จะต้องเป็นวงวนป้อนกลับ (Feedback loop) ซึ่งตัวทำนายของการบีบและการคลายจะต้องตรงกัน เทคนิคนี้ใช้ข้อดีที่ว่าค่าจุดภาพที่ใกล้เคียงกันจะมีความสัมพันธ์กันสูงหรือมีค่าใกล้เคียงกัน ซึ่งหมายความว่า ค่าแตกต่างระหว่างจุดภาพจะมีค่าน้อย และเพราะค่าแตกต่างกันเล็กน้อยนี้เองที่ทำให้จำนวนบิตที่ใช้ในการเก็บข้อมูลน้อยลงไปด้วย เช่น จากภาพต้นฉบับ 8 บิตต่อจุดภาพ จะได้ภาพคุณภาพคมชัดซึ่งใช้จำนวนบิต 4 บิตต่อจุดภาพ และภาพที่ขอมรับคุณภาพได้ที่ 3 บิตต่อจุดภาพ เป็นต้น การทำนายค่าจุดภาพจะใช้สมการง่ายๆ ซึ่งทำให้สามารถทำนายค่าจุดภาพถัดไป จากนั้นหาค่าแตกต่างของค่าที่จริงและค่าทำนาย, ควอนไทซ์ค่าแตกต่างนั้น, เข้ารหัสข้อมูลเพื่อการส่งผ่านหรือเก็บตามลำดับ ถ้าการทำนายค่าได้ใกล้เคียงค่าจริงจะทำให้ค่าแตกต่างนั้นเล็ก และการเข้ารหัสข้อมูลก็จะมีขนาดเล็กลง รูปแบบสมการการทำนายเป็นดังนี้

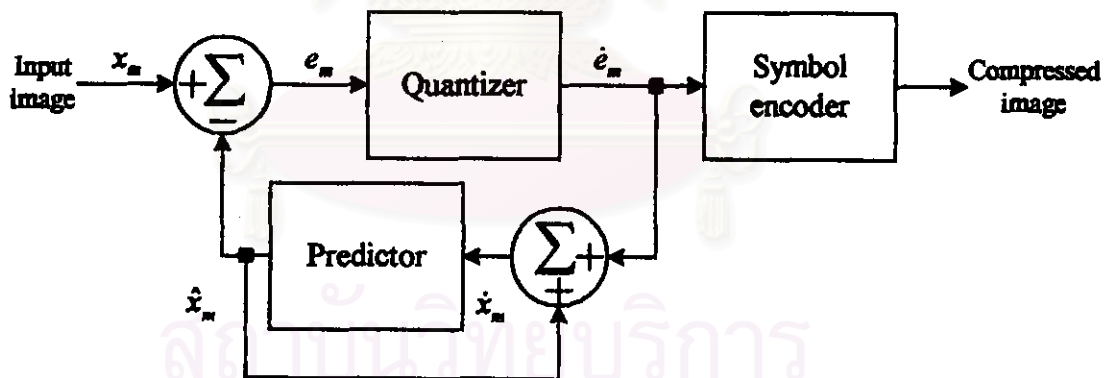
$$\hat{x}_m = \sum_{i=0}^{m-1} \alpha_i x_i$$

โดยที่ \hat{x}_m คือ ค่าจุดภาพที่ทำนาย
 α_i คือ สัมประสิทธิ์การทำนายหรือเวทิงแฟกเตอร์ (Prediction coefficient หรือ Weighting factor) ซึ่งระบุได้โดยผู้ใช้
 x_i คือ ค่าจุดภาพที่แถว m ค่า
 m คือ จำนวนจุดภาพที่ใช้ทำนาย

หลังจากการทำนายแล้ว หาค่าแตกต่างระหว่างค่าที่จริงและค่าทำนายจากสมการดังนี้

$$e_m = x_m - \hat{x}_m$$

โดยที่ e_m คือ ค่าแตกต่างระหว่างค่าที่จริงและค่าทำนาย
 x_m คือ ค่าจุดภาพแท้จริง

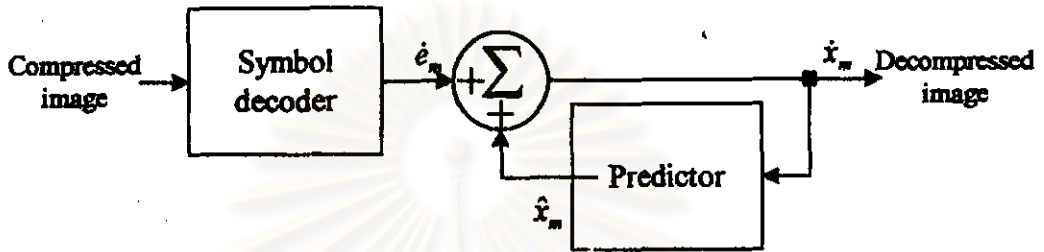


รูปที่ 2.9 แสดงการบีบภาพแบบคิทีซีเอ็ม (Umbaugh, 1998: 271)

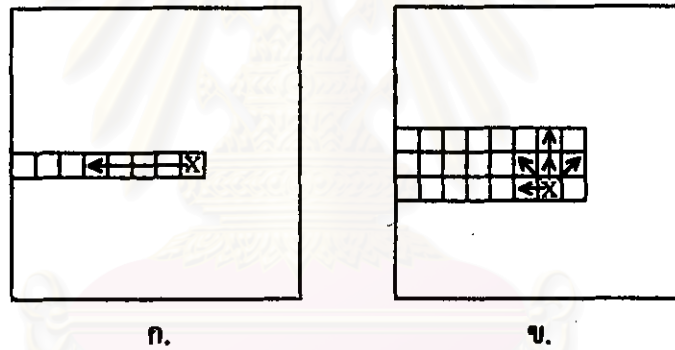
การถอดรหัสจะประกอบไปด้วยตัวทำนายและตัวเพิ่มค่าจุดภาพ จากรูปที่ 2.10 แสดงการถอดภาพแบบคิทีซีเอ็ม เริ่มต้นโดยการอ่านค่าข้อมูลเข้าถอดรหัสจะได้ค่าแตกต่างระหว่างค่าที่จริงและค่าทำนายที่ผ่านการควอนไทซ์แล้ว เพิ่มค่าข้อมูลที่ได้จากการทำนายลงไป เพื่อสร้างภาพให้กลับมาใหม่ ซึ่งคำนวณได้จากสมการดังนี้

$$\hat{x}_m = \hat{e}_m + \hat{x}_m$$

โดยที่ \hat{x}_m คือ ค่าจุดภาพแท้จริงซึ่งอาจจะไม่เท่ากับค่าจากภาพต้นฉบับ ขึ้นอยู่กับค่าความแตกต่างนั้นผ่านการควอนไทซ์หรือไม่
 e_m คือ ค่าแตกต่างระหว่างค่าแท้จริงและค่าทำนายที่ผ่านการควอนไทซ์แล้วและผ่านการถอดรหัส



รูปที่ 2.10 แสดงการคลายแบบดิฟิซิเอ็ม (Umbaugh, 1998: 271)



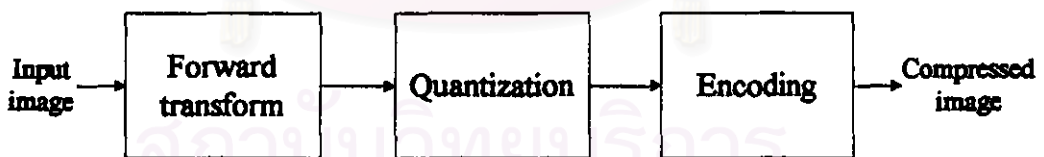
รูปที่ 2.11 แสดงการบีบภาพแบบดิฟิซิเอ็ม รูป ก. หนึ่งมิติ (one-dimensional predictor)

รูป ข. สองมิติ (two-dimensional predictor) (Umbaugh, 1998: 274)

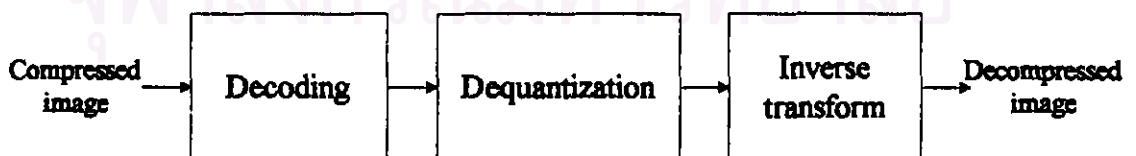
วิธีดิฟิซิเอ็มสามารถบีบข้อมูลภาพเป็นแบบไม่สูญเสียข้อมูลหรือแบบสูญเสียข้อมูลบางส่วนก็ได้ ขึ้นอยู่กับการควอนไทซ์ โดยแบบไม่สูญเสียข้อมูลจะไม่มีควอนไทซ์และจะเข้ารหัสข้อมูลด้วยวิธีที่ไม่มีการสูญเสียข้อมูล เช่น วิธีการเข้ารหัสฮัฟฟ์แมน เป็นต้น สามารถนำไปใช้ได้จริงกับฮาร์ดแวร์และซอฟต์แวร์ สามารถทำนายค่าได้ทั้ง 1 มิติ, 2 มิติ หรือแม้แต่ 3 มิติ ยิ่งมิติมากขึ้นวิธีการจะซับซ้อนมากขึ้น และทำให้ข้อมูลแน่นขึ้นมากขึ้นด้วย แต่จากการศึกษากับภาพโทรทัศน์ไม่ควรเกิน 3 มิติเพราะจะทำให้ซับซ้อนมากเกินไปและไม่ได้เพิ่มประสิทธิภาพแต่อย่างใด จากวิธี 1 มิติ เราจะทำนายค่าเฉพาะบรรทัดปัจจุบันเท่านั้น วิธี 2 มิติจะทำนายค่าจากบรรทัดปัจจุบันและบรรทัดที่แนวดิ่งด้วย ดังเช่นรูปที่ 2.11 ส่วน 3 มิติ จะทำนายค่าที่เฟรมปัจจุบันและเฟรมที่แนวดิ่งด้วย

2.2.3.2 การเข้ารหัสการแปลง (Dougherty, 1994: 285-293) (Gonzalez และ Woods, 1992: 374-389) (Umbaugh, 1998: 276-281) การบีบที่ได้กล่าวในวิธีดัดแปลงเป็นการกระทำกับกับข้อมูลจุดภาพโดยตรงซึ่งเป็นโดเมนของสเปซเชิงพื้นที่ (Spatial domain) แต่ในวิธีการเข้ารหัสการแปลงนี้จะเป็นการแปลง (Transform) ข้อมูลภาพซึ่งอยู่ในรูปโดเมนของสเปซเชิงพื้นที่ให้อยู่ในรูปโดเมนของความถี่ (Frequency domain) เช่น การแปลงฟูริเยร์ (Fourier), ฮาดามาร์ด (Hadamard), ฮาร์ (Haar), เวลช (Welsh), ซายน์ (Sine) และ โคซายน์ (Cosine) เป็นต้น ซึ่งจะทำให้สามารถแยกส่วนที่มีเนื้อหาข่าวสารมากและน้อยของภาพได้ การเข้ารหัสการแปลงมีคุณสมบัติที่สำคัญ 2 ประการ คุณสมบัติที่หนึ่ง เนื่องจากค่าสัมประสิทธิ์ที่ได้หลังการแปลงที่มีพลังงานสูงจะอยู่รวมกันเป็นกลุ่ม ซึ่งเรียกว่า “Energy compaction property” (ภาสกร ประถมบุตร, 2535: 18) จะสามารถเลือกเฉพาะสัมประสิทธิ์กลุ่มที่มีพลังงานสูงซึ่งมีความสำคัญต่อการแปลงกลับในขั้นตอนการคลายภาพ มาเข้ารหัสแทนการเข้ารหัสสัมประสิทธิ์ทุกตัว จึงมีผลทำให้เนื้อที่ในการเก็บและเวลาที่ใช้ในการส่งผ่านข้อมูลน้อยลง คุณสมบัติที่สอง คือ “Correlation reduction property” คือ การลดความสัมพันธ์ระหว่างสัมประสิทธิ์การแปลง

การแปลงส่วนมากสามารถนำมาใช้กับงานด้านการบีบข้อมูลภาพ และมักนำผลมาเปรียบเทียบกันในทางประสิทธิภาพและความสามารถในการนำภาพกลับมาให้เหมือนหรือใกล้เคียงกับต้นฉบับเดิมมากที่สุด ซึ่งการแปลงในแต่ละแบบก็ให้ผลแตกต่างกัน แต่ขั้นตอนโดยทั่วไปของการเข้ารหัสการแปลงจะเหมือนๆ กัน ซึ่งอธิบายได้ดังรูปที่ 2.12 (Dougherty, 1994)



รูปที่ 2.12 แสดงการเข้ารหัสการแปลง (Fluckiger, 1995)



รูปที่ 2.13 แสดงการคลายจากการเข้ารหัสการแปลง (Fluckiger, 1995)

ภาพที่ถูกเข้ารหัสแล้วจากรูปที่ 2.12 สามารถนำข้อมูลนี้ส่งผ่านช่องทางสื่อสารต่างๆ หรือเก็บในหน่วยเก็บข้อมูล เมื่อต้องการดูภาพสามารถทำการคลายจากการบีบภาพเพื่อให้ได้ภาพกลับคืนมา ซึ่งอธิบายได้ดังรูปที่ 2.13

ภาพต้นฉบับจะถูกแปลงให้อยู่ในรูปของความถี่ตามแคว้นการแปลงนั้นๆ ซึ่งตั้งที่ได้หลังจากการแปลง คือ สัมประสิทธิ์การแปลง (Transform coefficients) จากนั้นจะนำค่าสัมประสิทธิ์ที่ได้จากการแปลงเลือกขึ้นมาบางส่วนจากทั้งหมด โดยทั่วไปจะเลือกเอาบริเวณที่มีค่าสัมบูรณ์ของสัมประสิทธิ์สูง ซึ่งเป็นสัมประสิทธิ์ที่มีความสำคัญต่อการคลายกลับของภาพ แล้วนำค่าสัมประสิทธิ์ที่เลือกมาเหล่านั้นผ่านการควอนไทซ์อีกครั้งเพื่อลดจำนวนบิตของข้อมูลลง ซึ่งบางครั้งค่ามันอาจถูกควอนไทซ์ให้เป็นศูนย์ จากนั้นจะทำการเข้ารหัสสัมประสิทธิ์ที่ถูกเลือก

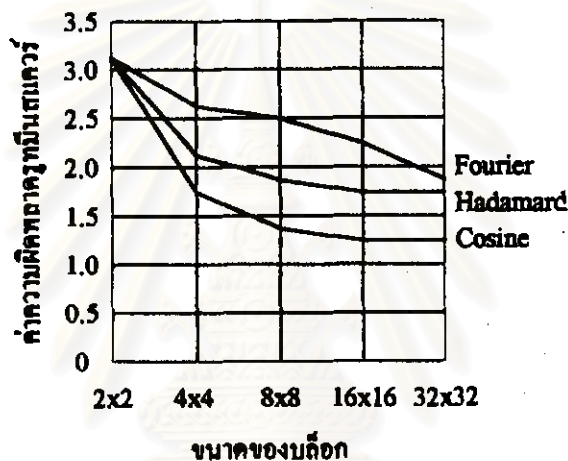
ในการนำภาพกลับมา จะต้องทำกระบวนการที่ตรงข้ามกัน คือ ในขั้นแรกจะนำภาพที่ถูกเข้ารหัสมาถอดรหัสและควอนไทซ์เพื่อให้มีจำนวนบิตเท่ากับต้นฉบับ จากนั้นจะเติมค่าคงที่ใดๆ ลงไปในตำแหน่งที่ถูกตัดค่าออกไปในตอนที่บีบ ซึ่งค่าคงที่ที่ที่ใ้ใช้มักจะเป็นศูนย์ แล้วนำไปแปลงกลับ โดยทำการผกผันการแปลงเพื่อให้ได้ภาพกลับมา

ขั้นตอนการเข้ารหัสการแปลงมีดังต่อไปนี้

1. การแบ่งภาพ (Subimage decomposition) คือ การแบ่งภาพออกเป็นภาพย่อยๆ (Subimage) ให้ได้ขนาด $m \times n$ จุดภาพต่อบล็อก จากนั้นจะทำการเข้ารหัสการแปลงแต่ละบล็อกหรือส่วนย่อยแยกจากกัน ซึ่งจะทำให้เกิดชุดของสัมประสิทธิ์การแปลงขนาด $m \times n$ ต่อบล็อก ด้วยวิธีนี้จะทำให้สามารถลดเวลาในการคำนวณได้มาก นอกจากนี้การเข้ารหัสการแปลงในแต่ละบล็อกจะช่วยลดความสูงฮาคับร็อนเมื่อนำมาทำโคซฮาร์ดแวร์ แต่การแบ่งเป็นบล็อกเมื่อนำผลจากการแปลงไปควอนไทซ์ ผลของภาพที่นำกลับมามีเกิดการเปลี่ยนแปลงระดับความเข้มแสง ที่เห็นชัดคือบริเวณขอบที่เป็นรอยต่อระหว่างบล็อก เรียกว่ากิบต็อกกิงอาร์ทิแฟท (Blocking artifact) ถ้าขนาดของบล็อกเล็กลงเท่าใดก็ยิ่งทำให้เกิดจำนวนบิตต็อกกิงอาร์ทิแฟทมากขึ้น

ด้วยเหตุนี้ ขนาดของบล็อกที่เลือกใช้จึงมีความสำคัญต่อการเข้ารหัสการแปลง โดยทั่วไป ระดับของการบีบและความซับซ้อนในการคำนวณจะเพิ่มขึ้นเมื่อขนาดของบล็อกใหญ่ขึ้น ขนาดของบล็อกหรือภาพย่อยที่นิยม คือ 8×8 หรือ 16×16 จุดภาพต่อบล็อก

2. การแปลง (Transformation) นับเป็นขั้นตอนที่สำคัญที่สุดของการเข้ารหัสการแปลง การแปลงเป็นขั้นตอนทางคณิตศาสตร์ที่ทำหน้าที่แปลงข้อมูลจากข้อมูลภาพซึ่งเป็นโดเมนของเวลาไปเป็นข้อมูลในรูปโดเมนของความถี่ หลังการแปลง ข้อมูลจะอยู่ในรูปของสัมประสิทธิ์ การแปลงซึ่งจะมีค่ามากขึ้นจากซ้ายไปขวา บนลงล่าง ตามแนวแกนนอนและแกนตั้งตามลำดับ หลังการแปลงจะสามารถตัดค่าสัมประสิทธิ์บางตัวทิ้งได้โดยไม่มีผลกระทบต่อรูปภาพมากนัก จุดสำคัญคือเพื่อจะแยกความสัมพันธ์ของจุดภาพในแต่ละบล็อกออกจากกัน ซึ่งจะทำให้ค่าสัมประสิทธิ์ที่ได้จากการแปลงนี้มีสหสัมพันธ์น้อยกว่าข้อมูลจุดภาพต้นฉบับ



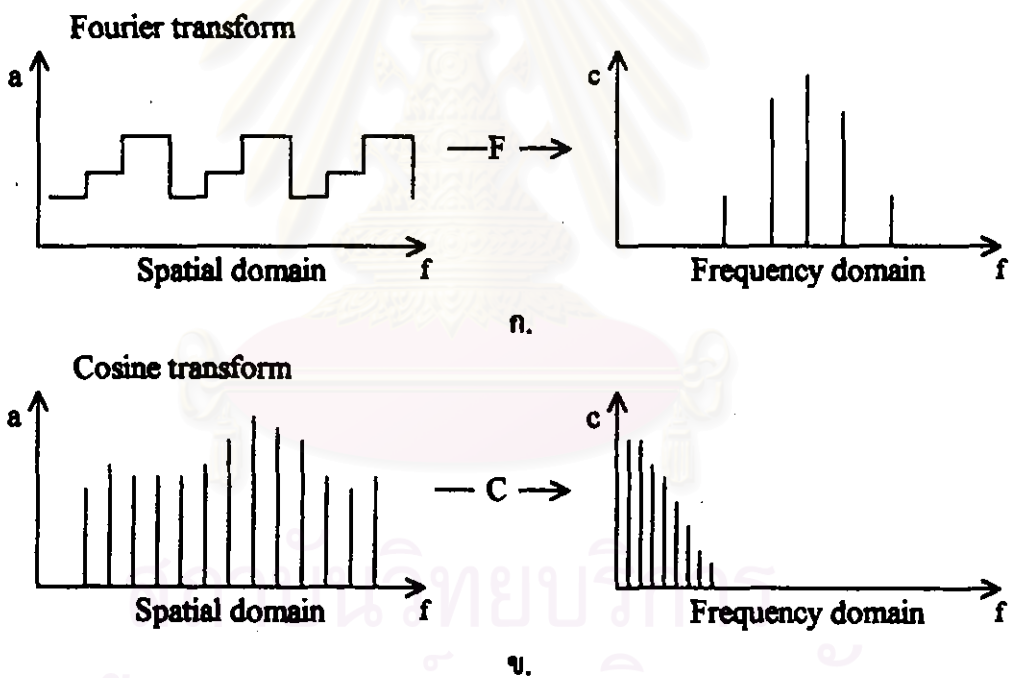
รูปที่ 2.14 แสดงขนาดต่างๆ ของบล็อกและค่าความผิดพลาดจุดภาพมินสมแควร์

(Gonzalez และ Woods, 1992: 381)

การแปลงที่ใช้ในการเข้ารหัสการแปลงมีอยู่หลายวิธี เช่น การแปลงพริ้นซิפלคอมโพเนนต์ (Principal Component Transform หรือ Karhunen-Loeve Transform), การแปลงดิสครีทฟูเรียร์ (Discrete Fourier Transform), การแปลงดิสครีทโคไซน์, การแปลงฮาดามาร์ดหรือฮาร์ (Hadamard หรือ Haar Transform) ฯลฯ วิธี Karhunen-Loeve Transform นับว่าเป็นการแปลงที่ให้ผลดีที่สุดไม่ว่าจะให้ค่าความผิดพลาดจุดภาพมินสมแควร์ (Mean square error) ต่ำที่สุดเมื่อเปรียบเทียบกับ การแปลงชนิดอื่นๆ แต่วิธีนี้ไม่เหมาะสมที่จะนำมาใช้ในทางปฏิบัติ เนื่องจากต้องการข้อมูลทางสถิติซึ่งบางครั้งไม่สามารถทราบได้ หรือไม่สามารถประมาณได้ถูกต้อง นอกจากนี้ยังเป็นการคำนวณที่ซับซ้อนยุ่งยากอีกด้วย

การแปลงที่นิยมใช้ในการเข้ารหัสการแปลง คือ การแปลงดิสครีทโคไซน์หรือ ดิจิทัลคอสายน์ โดย Ahmed และผู้ร่วมวิจัยคนอื่นๆ ในปี 1974 มีเบสิสฟังก์ชันเป็นองค์ประกอบของ

ฟังก์ชันโคซายน์เท่านั้น จัดเป็นฟังก์ชันทางคณิตศาสตร์กลุ่มเดียวกับการแปลงดิคริทฟูเรียร์หรือดีเอฟที ซึ่งถูกพัฒนาโดย Cooley และ Tukey ในปี 1965 ในทางทฤษฎี การแปลงดีซีทีจะให้ประสิทธิภาพในการบีบข้อมูลใกล้เคียงกับ Optimum Transformation ส่วนในทางปฏิบัติ ขั้นตอนวิธีในการคำนวณค่าสัมประสิทธิ์ของดีซีทีสามารถคำนวณได้อย่างรวดเร็วด้วยอัลกอริทึมการแปลงแบบฟาสต์ฟูเรียร์หรือเอฟเอฟที (Fast Fourier Transform หรือ FFT) ซึ่งมีจำนวนครั้งในการคำนวณน้อย จึงไม่มีความซับซ้อนในการคำนวณ ซึ่งจะช่วยให้ประหยัดเวลาได้เป็นอย่างมาก และการแปลงดีซีทีจะทำให้เกิดบล็อกกิ้งอาร์ทีแฟคชั่นเมื่อเปรียบเทียบกับ การแปลงชนิดอื่นๆ ซึ่งทำให้คุณภาพของภาพหลังการคลายกลับใกล้เคียงกับภาพต้นฉบับ นอกจากนี้แล้วค่าสัมประสิทธิ์ของการแปลงด้วยดีซีทียังเป็นเลขจำนวนจริง ทำให้ง่ายต่อการนำไปใช้งาน ด้วยคุณสมบัติที่กล่าวมา การแปลงด้วยดีซีทีจึงกลายเป็นมาตรฐานการบีบของวิธีการเข้ารหัสการแปลง คือ มาตรฐาน JPEG



รูปที่ 2.15 แสดงการแปลงภาพจากโดเมนของสเปซไปเป็นโดเมนของความถี่
รูป ก. การแปลงแบบฟูเรียร์ รูป ข. การแปลงแบบโคซายน์ (Fluckiger, 1995: 531)

เราสามารถพิจารณาการแปลงเป็นการแยกส่วนประกอบของบล็อกละเป็นส่วนประกอบของความถี่ กล่าวคือ ส่วนประกอบความถี่คือชุดของฟังก์ชันทางคณิตศาสตร์หลายๆฟังก์ชันที่ประกอบกันเป็นความถี่นั้นๆ หรือเป็นเสมือนการแยกส่วนประกอบสเปกตรัม โดยแต่ละฟังก์ชันมีสัดส่วนของพลังงานแตกต่างกัน และแต่ละสัมประสิทธิ์ที่ได้จากการแปลงจะได้สัดส่วนกับเศษ

ส่วนของพลังงานในบล็อก เช่น ในกรณีของการแปลงฟูเรียร์จะเป็นองค์ประกอบในทอมของฟังก์ชันชาอน์และโคซาอน์ที่จุดต่างๆ กัน การแยกองค์ประกอบฟังก์ชันคณิตศาสตร์หรือการตีความสเปกตรัมนี้สามารถใช้ในการแปลงแบบอื่นๆ ได้ โดยสามารถแยกเป็นฟังก์ชันอื่นๆ นอกจากชาอน์และโคซาอน์ได้ จากรูปที่ 2.15 แสดงการแปลงจากโดเมนของสเปซไปเป็นโดเมนของความถี่ของการแปลงฟูเรียร์และโคซาอน์โดยที่ a หมายถึง แอมพลิจูด f หมายถึง ความถี่ และ c หมายถึง สัมประสิทธิ์

สิ่งที่สำคัญที่สุด คือ การแปลงนั้นโดยตัวมันเองไม่เพิ่มการบีบแต่อย่างใด เพียงแต่เป็นการเปลี่ยนการแทนที่ของข่าวสารที่บรรจุในภาพให้อยู่ในรูปของความถี่ ซึ่งจะทำให้ข้อมูลเหมาะแก่การบีบ การบีบจะมีประสิทธิภาพเพิ่มขึ้นโดยขั้นตอนควอนไทเซชันและการเข้ารหัสสัมประสิทธิ์การแปลง อัตราส่วนการบีบขึ้นอยู่กับระดับคิกรของการทำควอนไทเซชัน กล่าวคือ ควอนไทเซชันที่หยาบกว่าจะให้ผลดีในการบีบที่สูงกว่า แต่จะทำให้เกิดการค้ำองลง (Degradation) ของคุณภาพของภาพมากขึ้นด้วย

3. การคัดเลือกสัมประสิทธิ์การแปลง (Coefficient transform selection) กระบวนการตั้งแต่การคัดเลือกสัมประสิทธิ์การแปลง ควอนไทเซชัน และการเข้ารหัส เป็นขั้นตอนที่เรียกรวมกันว่า "Bit allocation" การคัดเลือกสัมประสิทธิ์จะเลือกสัมประสิทธิ์ที่มีค่าสัมบูรณ์สูงหรือมีความสำคัญต่อการถอดกลับของภาพ มี 2 วิธี คือ การคัดเลือกสัมประสิทธิ์แบบการเข้ารหัสตามโซน (Zonal coding) และการเข้ารหัสแบบขีดแบ่ง (Threshold coding)

การคัดเลือกสัมประสิทธิ์แบบการเข้ารหัสตามโซน (Dougherty, 1994: 290-291) (Gonzalez และ Woods, 1992: 384-385) เป็นการคัดเลือกเฉพาะค่าสัมประสิทธิ์ที่อยู่ภายในโซนที่กำหนดไว้เพื่อนำมาเข้ารหัส โดยปกติมักเป็นค่าความถี่ต่ำ เพราะค่าสัมประสิทธิ์ที่มีค่าสัมบูรณ์สูงๆ ซึ่งมีความสำคัญต่อการแปลงกลับนี้มักจะตกอยู่ในช่วงใกล้จุดกำเนิดซึ่งเป็นบริเวณค่าความถี่ต่ำ (Low frequency) (ในกรณีที่ภาพนั้นมีค่าของระดับความเข้มแสงหรือระดับความเทาของจุดภาพต่อเนื่องกันหรือใกล้เคียงกัน) ส่วนค่าความถี่สูงจะถูกค้ำค่าให้เป็น 0 ในการกำหนดโซนจะใช้ตัวพรางโซน (Zonal mask) เป็นตัวกำหนด โดยการหาค่าความแตกต่างขององค์ประกอบความถี่ต่างๆ ในแต่ละบล็อก และจากนั้นหาค่าความแตกต่างจากทุกๆ บล็อก ตัวพรางโซนนี้จะป็นบิตแมทของชุดค่า 1 และชุดค่า 0 โดยที่ ค่า 1 คือค่าสัมประสิทธิ์ที่คัดเลือกไว้ และค่า 0 คือค่าสัมประสิทธิ์ที่ขจัดทิ้ง วิธีการนี้มีข้อจำกัดคือต้องมีการกำหนดโซนความถี่ให้เหมาะสมกับความถี่ของรูปภาพไว้ล่วงหน้า

เพราะทุกๆ บล็อกหรือภาพย่อยในภาพจะใช้โหนดความถี่หรือตัวพรางโหนดเดียวกัน ซึ่งจะทำการคัดเลือกตัวประสิทธิ์เกิดข้อผิดพลาดสูงเมื่อค่าของตัวประสิทธิ์มีค่าเบี่ยงเบนสูง ดังนั้นการคัดเลือกตัวประสิทธิ์ของภาพที่มีความถี่ต่างๆ ด้วยโหนดที่เจาะจงเพียงโหนดเดียวจึงมักไม่ให้เกิด จากรูปที่ 2.16 รูป ก. แสดงตำแหน่งของตัวประสิทธิ์ที่คัดเลือกไว้ด้วยตัวพรางโหนดซึ่งจะอยู่รอบๆ บริเวณจุดกำเนิด ส่วนรูป ข. แสดงจำนวนบิตที่ใช้ในการเข้ารหัสตัวประสิทธิ์ที่คัดเลือกไว้

1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ก.

8	7	6	5	4	3	2	1
7	6	5	4	3	2	1	0
6	5	4	3	2	1	0	0
5	4	3	2	1	0	0	0
4	3	2	1	0	0	0	0
3	2	1	0	0	0	0	0
2	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0

ข.

รูปที่ 2.16 แสดงการคัดเลือกตัวประสิทธิ์แบบการเข้ารหัสตามโหนด

รูป ก. แสดงตำแหน่งตัวประสิทธิ์ที่คัดเลือกไว้

รูป ข. แสดงจำนวนบิตที่ใช้ในการเข้ารหัสตัวประสิทธิ์ที่คัดเลือกไว้

(Gonzalez และ Woods, 1992: 385)

การคัดเลือกตัวประสิทธิ์แบบการเข้ารหัสแบบขีดแบ่ง (Dougherty, 1994: 291) (Gonzalez และ Woods, 1992: 385-386) เป็นวิธีที่แก้ปัญหาที่เกิดขึ้นในวิธีการเข้ารหัสตามโหนด คือ เป็นการคัดเลือกเฉพาะค่าตัวประสิทธิ์ที่มีค่าสูงกว่าค่าขีดแบ่งที่ตั้งไว้มาเข้ารหัส วิธีการนี้มีข้อดี คือ สามารถใช้ได้กับภาพที่มีความถี่ต่างๆ กันเนื่องจากตัวประสิทธิ์ที่มีความสำคัญสูงต่อการคลาซกั๊บ จะอยู่ในตำแหน่งต่างๆ ซึ่งไม่ขึ้นอยู่กับความถี่ของภาพ เริ่มต้นจากการกำหนดจุดหรือกลุ่มของตัวพรางขีดแบ่ง (Threshold mask) จากนั้นคัดเลือกตัวประสิทธิ์ให้อยู่ในกลุ่มใดกลุ่มหนึ่งในชุดตัวพรางขีดแบ่งหรือกำหนดให้แต่ละบล็อกมีตัวพรางขีดแบ่งไม่เหมือนกัน ดังนั้นทุกบล็อกจึงต้องส่งตำแหน่งของตัวประสิทธิ์ไปพร้อมกับตัวตัวประสิทธิ์ด้วยเสมอ ซึ่งมีผลทำให้อัตราส่วนการบีบไม่ดีกว่าการคัดเลือกตัวประสิทธิ์แบบเข้ารหัสตามโหนด จากรูปที่ 2.17 รูป ก. แสดงตัวอย่างของตำแหน่งตัวประสิทธิ์ที่คัดเลือกไว้ ด้วยตัวพรางขีดแบ่งซึ่งจะไม่เหมือนกับวิธีการเข้ารหัสตามโหนด เพราะตำแหน่งตัวประสิทธิ์ที่มีค่าสัมบูรณ์สูงมักไม่จำเป็นต้องอยู่บริเวณเดียวกัน ส่วนรูป ข. แสดงลำดับ

ตำแหน่งการเข้ารหัสตัวประติกรที่คัดเลือกไว้ ซึ่งเป็นวิธีการคัดเลือกตัวประติกรตามลำดับแบบรัน-เลนท (run-length coded sequence)

1	1	0	1	0	0	0	0
1	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ก.

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

ข.

รูปที่ 2.17 แสดงการคัดเลือกตัวประติกรแบบการเข้ารหัสแบบขีดแบ่ง

รูป ก. แสดงตัวอย่างตำแหน่งตัวประติกรที่คัดเลือกไว้

รูป ข. แสดงลำดับการเข้ารหัสตัวประติกรตามตำแหน่งที่คัดเลือกไว้

(Gonzalez และ Woods, 1992: 385)

4. ควอนไทเซชัน (Quantization) เมื่อคัดเลือกตัวประติกรที่ต้องการได้แล้ว จะทำการควอนไทซ์ค่าตัวประติกรเหล่านั้นเพื่อให้สามารถแทนค่าได้ด้วยจำนวนบิตที่กำหนดไว้ก่อนที่ จะทำการเข้ารหัสต่อไป โดยต้องคำนึงถึงคุณภาพของรูปเป็นสำคัญ ความสำคัญของขั้นตอนนี้คือ ต้องการลดค่าตัวประติกรลงเพื่อลดจำนวนบิตของข้อมูลลง โดยทั่วไป จำนวนบิตมากจะใช้กับค่า ตัวประติกรที่มีความดีค่า และจำนวนบิตน้อยจะใช้กับค่าตัวประติกรที่มีความดีสูง ด้วยวิธีการลด จำนวนตัวประติกรลง บางครั้งเมื่อทำการคลายภาพและต้องทำดีควอนไทเซชัน (Dequantization) จะได้ค่าตัวประติกรกลับมาไม่เท่าค่าเดิม ซึ่งหมายถึงรูปที่ได้จากการคลายจะไม่เหมือนต้นฉบับ

5. การเข้ารหัส (Encoding) คือ การเข้ารหัสตัวประติกรที่ได้ผ่านขั้นควอนไทเซชัน โดยมากมักจะใช้วิธีการเข้ารหัสแบบรหัสความยาวแปรได้ (Variable-length code) ได้แก่ การเข้ารหัสฮัฟฟ์แมน การเข้ารหัสเลขคณิต เป็นต้น

2.2.4 การบีบข้อมูลภาพแบบเจพีเอกและเวฟเลต

2.2.4.1 การบีบข้อมูลภาพแบบเจพีเอก (Wallace, 1991: 30-44) (Fluckiger, 1995: 527-536) (Gonzalez และ Woods, 1992: 395-404) (Nelson และ Gailly, 1996: 326-345) (CCITT, 1992)

เทคนิคการบีบภาพแบบเจพีเอกเป็นมาตรฐานการบีบที่ถูกพัฒนาขึ้น โดยกลุ่มไอโซที่ชื่อ "Joint Photographic Expert Group" โดยความร่วมมือกับซีซีไอทีทีซึ่งเป็นหน่วยงานหนึ่งในไอทียู เป็นมาตรฐานที่เกิดขึ้นในปี 1987 และมีการปรับปรุงอยู่ตลอดเวลา จนถึงปี 1990 ก็ได้ออกคำจำกัดความ มาตรฐาน เอกสารเผยแพร่ การทดสอบมาตรฐานและตรวจสอบผลลัพธ์ให้ถูกต้อง จากนั้น จึงได้ออกเป็นมาตรฐานรุ่นล่าสุดในปี 1991 มาตรฐานการบีบภาพเจพีเอกถูกออกแบบมาเพื่อให้บีบภาพ ฟูลคัลเลอร์ (full-color) 24 บิต หรือ ภาพเกรย์สเกลที่สามารถใช้จำนวนบิตได้หลายระดับ เช่น 4-16 บิต สามารถบีบภาพได้ดีมากกว่าภาพเกรย์สเกล เหมาะกับภาพที่เห็นจริงตามธรรมชาติ (natural, real-world scenes) หรือภาพที่มีความต่อเนื่องของระดับความเทา (Continuous-tone images) เช่น ภาพถ่าย ภาพงานพิมพ์ของธรรมชาติ طينดิน แต่ไม่เหมาะกับภาพขาวดำ (Black-and-White images) หรือภาพไบโทนอล (bitonal images) เช่น ภาพการ์ตูนอย่างง่าย ภาพข้อความจดหมาย ภาพลายเส้นอื่นๆ (line drawing) ภาพที่มีการเปลี่ยนแปลงของสีสูง เป็นต้น มาตรฐานการบีบเจพีเอก สามารถบีบภาพได้ทั้งประเภทไม่สูญเสียข้อมูล และสูญเสียข้อมูลบางส่วน ซึ่งสามารถปรับระดับ คุณภาพภาพได้ตั้งแต่หยาบจนละเอียด มีหลักการบีบเช่นเดียวกับการเข้ารหัสการแปลง

2.2.4.1.1 มาตรฐานการบีบภาพแบบเจพีเอก มี 4 ประเภทใหญ่ๆ คือ

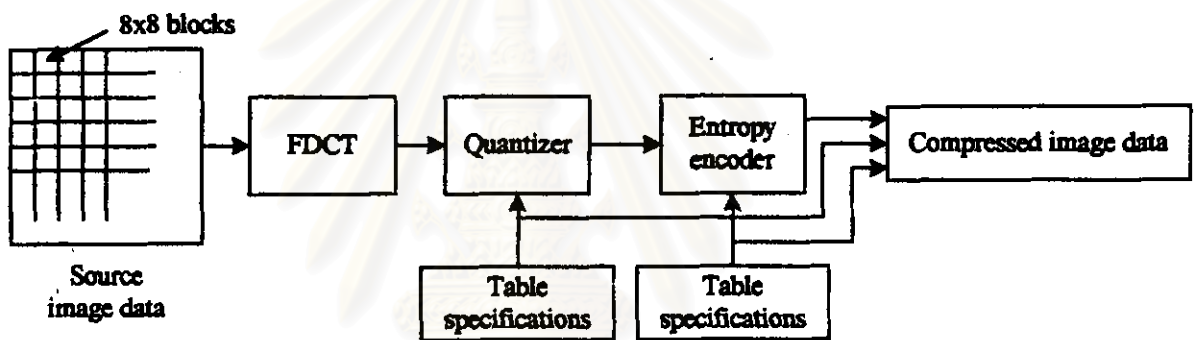
การเข้ารหัสดีซีทีเบสเชิงลำดับหรือเบสไลน์เจพีเอก (Sequential DCT-based encoding หรือ Baseline JPEG) เป็นวิธีที่ง่ายที่สุด คือเป็นการกรวดภาพเพียงครั้งเดียว (Single scan) จากซ้ายไปขวาและบนลงล่าง เป็นการบีบประเภทสูญเสียข้อมูลบางส่วน สามารถบีบภาพได้ ในอัตราส่วนการบีบ 20:1 ถึง 30:1

การเข้ารหัสดีซีทีเบสโพรเกรสซีฟ (Progressive DCT-based encoding) จะใช้ การสร้างภาพใหม่หลายๆ ครั้งจากภาพหยาบสุดไปจนละเอียดสุด ซึ่งจะเป็นการกรวดภาพหลายๆ ครั้ง (multiple scan) เป็นการบีบประเภทสูญเสียข้อมูลบางส่วน ซึ่งออกแบบมาเพื่อใช้ในการส่งผ่าน ข้อมูลระบบทำงานแบบทันที (real-time system) มักใช้ในการส่งภาพทางเครือข่ายต่างๆ

การเข้ารหัสแบบไม่สูญเสียข้อมูลหรือเจทีเคแอลเอส (Lossless encoding หรือ JPEG-LS) เป็นการบีบภาพประเภทไม่สูญเสียข้อมูล บีบได้ในอัตราส่วนการบีบ 2:1 หรือประมาณ 12 บิตต่อจุดภาพ เหมาะกับภาพที่ต้องการให้เหมือนต้นฉบับทุกอย่าง เช่น ภาพถ่ายทางการแพทย์ ภาพถ่ายทางอากาศ เป็นต้น

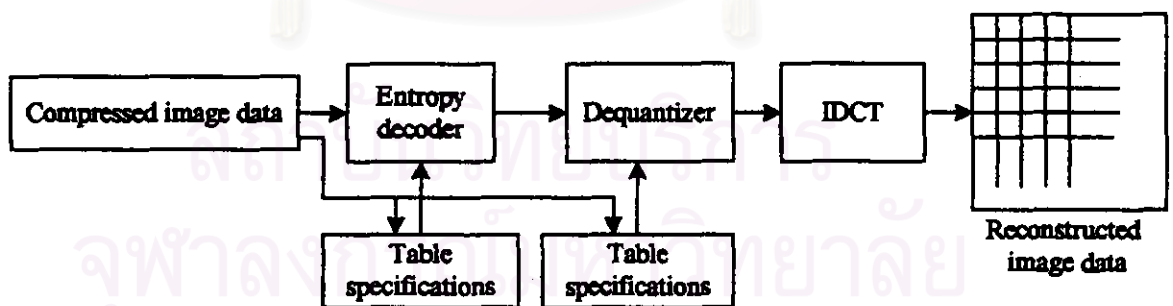
การเข้ารหัสเชิงลำดับชั้น (Hierarchical encoding) จะใช้การสร้างภาพหลายๆ ภาพแต่ละภาพมีความคมชัด (resolution) ต่างกัน ซึ่งขั้นตอนการคลายภาพจะแยกจากกัน

2.2.4.1.2 ขั้นตอนการบีบภาพในประเภทการเข้ารหัสวิธีทีเบตเชิงลำดับ



รูปที่ 2.18 แสดงการบีบภาพด้วยเทคนิคเจทีเคแอลเอสประเภทการเข้ารหัสวิธีทีเบตเชิงลำดับ

(Wallace, 1991)



รูปที่ 2.19 แสดงการคลายจากการบีบภาพด้วยเทคนิคเจทีเคแอลเอสประเภทการเข้ารหัสวิธีทีเบตเชิงลำดับ

(Wallace, 1991)

การเตรียมบล็อกข้อมูล มาตรฐานการบีบภาพเจทีเคแอลเอสจะแบ่งข้อมูลภาพออกเป็นภาพย่อยๆ หรือเป็นบล็อก บล็อกละ 8x8 หรือ 64 ค่า ข้อมูลจะมีลักษณะเป็นแมทริกซ์ค่าแอมพลิจูดของ

สัญญาณ มีค่าสูงสุดเท่ากับ 255 หรือ 8 บิต ในกรณีที่ เป็นภาพสี จะต้องทำขั้นตอน Color space conversion คือการเปลี่ยนภาพสี (Color Space) จากแบบจำลองสีอาร์จีบี (RGB mode) ให้อยู่ในรูปแบบจำลองสีวายซีอาร์จีบี (YCrCb mode) หรือแบบจำลองสีวายยูวี (YUV mode) ซึ่งจะเป็นการแยกข้อมูลลูมิแนนซ์และโครมิแนนซ์ออกจากกัน เพราะมนุษย์จะรับรู้ข้อมูลแสงหรือระดับความส่องสว่าง (ลูมิแนนซ์) มากกว่าข้อมูลสี (โครมิแนนซ์) ทำให้สามารถบีบข้อมูลสีได้มากกว่าข้อมูลระดับความส่องสว่าง ดังนั้นภาพสีบีบได้มากกว่าภาพเกรย์สเกล หลังจากแบ่งภาพเป็นบล็อกแล้ว จะเลื่อนค่าในบล็อกจากค่าในช่วง $[0, 2^P-1]$ หรือ $0-255$ ไปอยู่ในช่วง ช่วง $[-2^{P-1}, 2^{P-1}]$ หรือ -128 ถึง 127 โดยใช้วิธีง่าย ๆ คือลบค่าของจุดภาพด้วย 128 และเข้าสู่ขั้นตอนการแปลงด้วยดีซีที

การแปลงภาพโดยใช้การแปลงโคไซน์ทริโคซายน์ (Discrete Cosine Transform) การแปลงดีซีทีเป็นวิธีการทางคณิตศาสตร์ที่เปลี่ยนข้อมูลจากโดเมนของสเปเชียลไปเป็นโดเมนของความถี่ ด้านแปลงไปข้างหน้า เรียกว่า ฟอว์เวิร์ดดีซีที (Forward discrete cosine transform หรือ FDCT) การแปลงโคไซน์ 1 มิติจะเปลี่ยนข้อมูลที่เป็นอะเรย์ของจำนวนซึ่งจะแทนสัญญาณแอมพลิจูดที่ตำแหน่งต่างๆ ในโดเมนของเวลาไปเป็นอะเรย์ของจำนวนซึ่งจะแทนแอมพลิจูดของความถี่จากสัญญาณต้นฉบับหรือสัมประสิทธิ์ดีซีที (DCT Coefficient) ผลลัพธ์จากการแปลงด้วยดีซีที 1 มิติ จะเป็นอะเรย์ของความถี่ ค่าตัวแรกของอะเรย์จะเป็นสัมประสิทธิ์ดีซีที (DC Coefficient) ซึ่งเป็นค่าเฉลี่ยของทุกๆ ค่าในอะเรย์ และค่าในอะเรย์ที่เหลือจะเป็นสัมประสิทธิ์เอซี (AC Coefficient) จากสมการการแปลงฟอว์เวิร์ดดีซีที 1 มิติ (1-Dimensional FDCT) เป็นดังนี้

$$DCT(i) = C(i) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{(2x+1)i\pi}{2N} \right]$$

และสมการของการแปลงกลับด้วยดีซีที 1 มิติ (Inverse 1-Dimensional DCT) เป็นดังนี้

$$f(x) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i) DCT(i) \cos \left[\frac{(2x+1)i\pi}{2N} \right]$$

โดยที่ $x, i = 0, 1, 2, 3, \dots, N-1$

$$C(i) = \sqrt{\frac{1}{N}} \quad \text{เมื่อ } i = 0$$

$$C(i) = \sqrt{\frac{2}{N}} \quad \text{เมื่อ } i \neq 0$$

ส่วนสมการของการแปลงฟอว์เวิร์ดคิซีที 2 มิติ (2-D DCT) เป็นดังนี้

$$DCT(i, j) = C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

สมการต่อไปนี้เป็น การแปลงกลับด้วยคิซีที 2 มิติ เรียกว่าอินเวอร์สคิซีที (Inverse 2-D DCT)

$$f(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j) DCT(i, j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

โดยที่ $i, j, x, y = 0, 1, 2, 3, \dots, N-1$

$DCT(i, j)$ คือ สัมประสิทธิ์การแปลงที่ตำแหน่งความถี่ i และ j

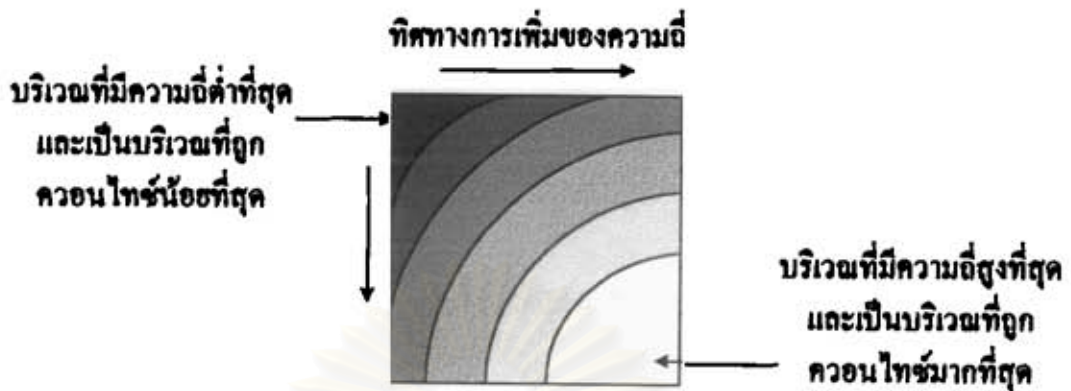
$f(x, y)$ คือ ค่าจุดภาพในบล็อกต้นฉบับ

$$C(i), C(j) = \sqrt{\frac{1}{N}} \text{ เมื่อ } i, j = 0$$

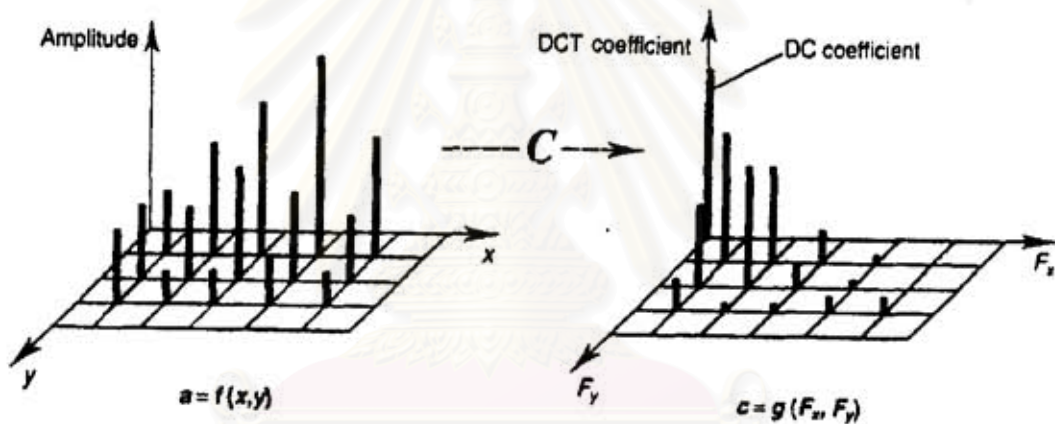
$$C(i), C(j) = \sqrt{\frac{2}{N}} \text{ เมื่อ } i, j \neq 0$$

เมื่อนำการแปลงคิซีที 2 มิติ มาใช้ในขั้นตอนการแปลงของเทคนิคเจเท็ก เราจะพิจารณาแค่บล็อก 8×8 ของภาพเป็นอะเรย์ 2 มิติของค่าความเข้มแสงของจุดภาพ บล็อกข้อมูลภูมิแนวนอนและบล็อกข้อมูลโครมิแนนซ์จะถูกแปลงโดยแยกจากกัน การแปลงคิซีทีโคซายน์จะเปลี่ยนค่าอมพลิจูดหรือค่าความเข้มแสงของจุดภาพไปเป็นสัมประสิทธิ์คิซีที โดยไม่ทำให้เกิดการสูญเสียข้อมูล แต่เป็นการเปลี่ยนโคเมนของภาพให้อยู่ในรูปที่เหมาะสมกับการบีบเท่านั้น หลังการแปลง ตำแหน่งมุมบนซ้ายของแต่ละบล็อก (0,0) คือ สัมประสิทธิ์คิซีที จะเป็นค่าเฉลี่ยของทุกๆ ค่าในแค่บล็อก ส่วนค่าสัมประสิทธิ์ที่เหลือ คือ สัมประสิทธิ์เอซี จะมีค่าตั้งแต่ -1024 ถึง 1024 สำหรับภาพ 8 บิต และความถี่ของสัมประสิทธิ์จะเพิ่มจากซ้ายไปขวาและบนลงล่าง ตามแกนอนและแกนตั้งตามลำดับ ซึ่งแสดงให้เห็นว่าค่าสัมประสิทธิ์ที่มีความสำคัญต่อการแปลงกลับของภาพอยู่ที่สัมประสิทธิ์ความถี่ต่ำ แสดงดังรูปที่ 2.20 (Wallace, 1991) (Fluckiger, 1995)

จากรูปที่ 2.21 แสดงการแปลงคิซีทีโคซายน์ 2 มิติ จากฟังก์ชัน $a = f(x, y)$ x และ y หมายถึง ตำแหน่งของค่าใน 2 มิติ จากฟังก์ชัน $c = g(F_x, F_y)$ c คือสัมประสิทธิ์คิซีที F_x และ F_y คือ ความถี่ที่ตำแหน่ง x และ y



รูปที่ 2.20 แสดงแนวโน้มของค่าสัมประสิทธิ์คิซีทีหลังการแปลง (Mattison, 1994: 166)



รูปที่ 2.21 แสดงการแปลงคิซีทวิท 2 มิติ (Fluckiger, 1995: 531)

ควอนไทเซชัน เป็นขั้นตอนที่มีการสูญเสียข้อมูลบางส่วนครั้งแรก มีจุดประสงค์เพื่อเพิ่มอัตราส่วนการบีบโดยการขจัดข้อมูลที่ไม่มีค่าเป็นต่อข่าวดารออก กระทำได้โดยการหารค่าสัมประสิทธิ์คิซีทีด้วยค่าในตารางควอนไทเซชัน (Quantization table) แต่ละค่าในตารางควอนไทเซชันเรียกว่า "Quantum" มีค่าตั้งแต่ 1 ถึง 255 จากนั้นขจัดตัวเลขทศนิยมให้เป็นตัวเลขจำนวนเต็มที่ใหญ่ที่สุด ซึ่งจะทำให้สัมประสิทธิ์มีค่าน้อยลง ข้อมูลความถี่สูงจะถูกควอนไทซ์ค่าได้มากกว่าข้อมูลความถี่ต่ำ ดังนั้นบริเวณมุมขวาต่าง ค่าสัมประสิทธิ์จะมีค่าเป็นศูนย์หรือใกล้เคียงศูนย์ รูปที่ 2.20 คุณภาพของภาพที่ผู้ใช้ต้องการจะสามารถปรับได้โดยการกำหนดตารางควอนไทเซชันเอง กล่าวคือ ถ้าตารางควอนไทเซชันหยาบหรือแต่ละค่าควอนตัมมีค่าสูง ภาพที่ได้คุณภาพต่ำ ถ้าตารางควอนไทเซชันละเอียดหรือค่าควอนตัมมีค่าต่ำ ภาพที่ได้คุณภาพสูง เช่น ถ้าทุกๆ ค่าในดา

ร่างเป็น 1 หมายความว่า ไม่มีผลกระทบใดๆ เกิดขึ้นกับค่าสัมประสิทธิ์ ถ้า 4 ค่าในตำแหน่งสัมประสิทธิ์มุมซ้ายบนมีค่าเป็น 16 อีก 60 ค่าเป็น 255 ก็จะมีเฉพาะสัมประสิทธิ์ 4 ตัวซ้ายบนเท่านั้นที่เก็บค่าไว้ เป็นต้น ภาพเกรย์สเกลจะใช้ตารางเคียว แต่ภาพสีจะมี 1 ตารางสำหรับองค์ประกอบภูมิแนวน และอีก 1 ตารางสำหรับองค์ประกอบโครมิแนนซ์

สมการต่อไปนี้เป็นการควอนไทซ์ให้ค่าสัมประสิทธิ์ลดลง

$$QuantizedValue(i, j) = IntegerRound \left[\frac{DCT(i, j)}{Quantum(i, j)} \right]$$

สมการควอนไทซ์ดังกล่าว เป็นวิธีที่เรียกว่า ศูนย์ฟอร์มสกาเลาร์ควอนไทซ์ (Uniform scalar quantization) ส่วนฟังก์ชัน IntegerRound หมายถึง ถ้าการหารนั้นให้ค่าเป็นทศนิยมจะต้องปัดค่าที่หารได้ให้เป็นจำนวนเต็มที่ใกล้เคียงที่สุด ส่วนดีควอนไทซ์จะใช้สมการต่อไปนี้

$$DCT(i, j) = Quantum(i, j) * Quantized Value(i, j)$$

โดยที่ $DCT(i, j)$ คือ ค่าสัมประสิทธิ์คิซีที่ได้หลังจากการแปลง

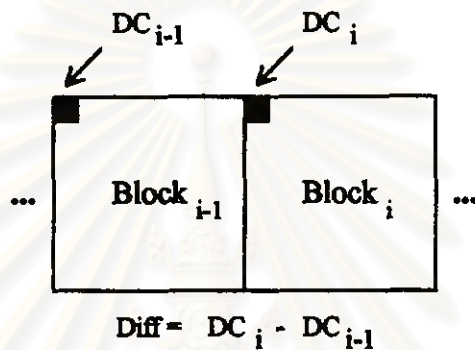
$Quantum(i, j)$ คือ ค่าในตารางควอนไทซ์

$Quantized Value(i, j)$ คือ ค่าที่ได้หลังควอนไทซ์

ซึ่งจะนำไปเข้ารหัส

การเข้ารหัสรันเลนท์และการเข้ารหัสอ็อพทีมัล หลังจากควอนไทซ์แล้ว สัมประสิทธิ์คิซีซึ่งจะมี 1 ค่าคือ 1 บล็อก จะถูกเข้ารหัสในลักษณะที่เรียกว่า "Prediction error" คือ การหาค่าแตกต่างระหว่างค่าสัมประสิทธิ์คิซีตัวปัจจุบันและค่าสัมประสิทธิ์คิซีตัวที่ผ่านมา หลังจากนั้นนำค่าสัมประสิทธิ์ทั้งคิซีและเฮซีเข้ารหัสรันเลนท์ (Run-length encoding) ซึ่งเป็นการแทนที่สัมประสิทธิ์ที่เป็นศูนย์ด้วยจำนวนสัมประสิทธิ์ที่เป็นศูนย์ โดยใช้ลำดับแบบซิกแซก (Zig-zag ordering) เรียงจากสัมประสิทธิ์ความถี่ต่ำไปสัมประสิทธิ์ความถี่สูง จากรูปที่ 2.22 แสดงการเข้ารหัสค่าแตกต่างระหว่างสัมประสิทธิ์คิซี และรูปที่ 2.23 ซึ่งแสดงการเข้ารหัสรันเลนท์ด้วยลำดับแบบซิกแซก

จากนั้นนำค่าสัมประสิทธิ์ทั้งหมดมาเข้ารหัสที่ฟิแมน ซึ่งจะป็นขั้นตอนที่ไม่มี การสูญเสียข้อมูล การเข้ารหัสนี้จะต้องใช้ตารางสำหรับการบีบและการคลายภาพ สำหรับภาพเกรย์ สเกลจะใช้ตารางฮัฟฟิแมนตารางเคียว และภาพสีจะใช้ 1 ตารางสำหรับองค์ประกอบภูมิแนซ และ อีก 1 ตารางสำหรับองค์ประกอบโครมิแนซ เราอาจนำวิธีการเข้ารหัสเลขคณิตมาใช้แทนการเข้ารหัสที่ฟิแมนซึ่งมีประสิทธิภาพดีกว่าประมาณ 5-10% เพราะไม่ต้องใช้ตาราง ใช้สถิติของการเกิด ข้อมูลในภาพในการเข้ารหัสภาพ แต่การคำนวณของการเข้ารหัสเลขคณิตจะซับซ้อนกว่า



รูปที่ 2.22 แสดงการเข้ารหัสค่าแตกต่างระหว่างสัมประสิทธิ์ดีซี (Wallace, 1991)



รูปที่ 2.23 แสดงการเข้ารหัสรันแลนทด้วยลำดับแบบซิกแซก (Fluckiger, 1995)

การเข้ารหัสฮัฟฟ์แมนประกอบด้วย ขั้นแรก เปลี่ยนลำดับซิกแซกของสัมประสิทธิ์ที่ถูกควอนไทซ์แล้วเป็นลำดับสัญลักษณ์อินทอมีเดียท (Intermediate symbol sequence) และขั้นที่สอง เปลี่ยนลำดับสัญลักษณ์อินทอมีเดียทเป็นกระแสข้อมูล (data stream) ดังนั้น สัมประสิทธิ์เอจี้จะถูกแทนที่ด้วยสัญลักษณ์ต่อไปนี้

สัญลักษณ์ที่ 1 : (Runlength, Size)

Runlength คือ จำนวนสัมประสิทธิ์ที่มีค่าศูนย์ที่ต่อเนื่องกันในลำดับซิกแซก ซึ่งนำหน้าสัมประสิทธิ์เอจี้ที่ไม่เป็นศูนย์

Size คือ จำนวนบิตที่ใช้เพื่อเข้ารหัสแอมพลิจูด

สัญลักษณ์ที่ 2 : (Amplitude)

Amplitude คือ ค่าแอมพลิจูดของสัมประสิทธิ์เอจี้ที่ไม่เป็นศูนย์

ในการแทนที่ สัญลักษณ์ที่ 2 ต้องค่อจากสัญลักษณ์ที่ 1 เสมอ ยกเว้นในกรณีที่สัมประสิทธิ์เอจี้ช่วงที่มีค่าศูนย์จนถึงสัมประสิทธิ์เอจี้ตัวสุดท้ายจะใช้สัญลักษณ์ที่ 1 (0,0) ซึ่งหมายถึงจบบล็อก EOB (End of Block) ส่วนสัมประสิทธิ์คิซี้จะถูกแทนที่ด้วยสัญลักษณ์ต่อไปนี้

สัญลักษณ์ที่ 1 : (Size)

Size คือ จำนวนบิตที่ใช้เพื่อเข้ารหัสแอมพลิจูด

สัญลักษณ์ที่ 2 : (Amplitude)

Amplitude คือ ค่าแอมพลิจูดของสัมประสิทธิ์คิซี้

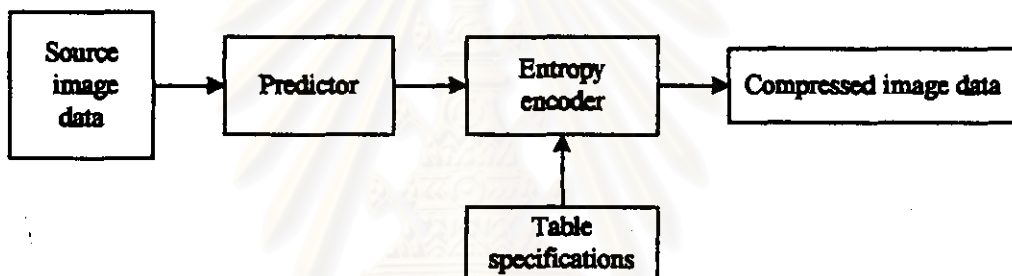
สามารถดูตัวอย่างการแทนที่สัญลักษณ์ในหัวข้อที่ 2.2.4.1.7 ตัวอย่างการบีบข้อมูลภาพด้วยเทคนิคเจทีค

ส่วนของขั้นตอนการคลายภาพจะทำการถอดรหัสด้วยวิธีฮัฟฟ์แมนและรันเลนท โดยย้อนกระบวนการกลับ จากตารางที่ 2.10 แสดงระดับคุณภาพของภาพเมื่อบีบด้วยจำนวนบิตต่อจุดภาพที่แตกต่างกัน ถ้าจำนวนบิตต่อจุดภาพสูง ภาพที่ได้คุณภาพจะดีมาก แต่อัตราส่วนการบีบจะต่ำลง

บิตต่อจุดภาพ	ระดับคุณภาพภาพ
0.25 - 0.5	คุณภาพปานกลางถึงดี
0.5 - 0.75	คุณภาพดีถึงดีมาก
0.75 - 1.5	คุณภาพดีมาก
1.5 - 2.0 หรือมากกว่า	คุณภาพไม่ต่างจากต้นฉบับ

ตารางที่ 2.10 ตารางแสดงคุณภาพของภาพที่จำนวนบิตต่อจุดภาพต่างๆ (Wallace, 1991)

2.2.4.1.3 ขั้นตอนการบีบภาพในประเภทการเข้ารหัสแบบไม่สูญเสียข้อมูล



รูปที่ 2.24 แสดงการบีบภาพด้วยเทคนิคเจทีเคประเภทการเข้ารหัส ไม่สูญเสียข้อมูล (Wallace, 1991)

เมื่อการบีบภาพด้วยเจทีเคต้องการการบีบภาพแบบไม่มีการสูญเสียข้อมูล ซึ่งจะให้ภาพหลังการคลายเหมือนกันในระดับบิตต่อบิต ก็ได้เลือกเทคนิคที่เรียกว่า "Prediction" อย่างง่ายมาใช้ ซึ่งจะให้อัตราส่วนการบีบได้ 2:1 หรือประมาณ 12 บิตต่อจุดภาพสำหรับภาพสี จากรูปที่ 2.24 การบีบภาพเริ่มต้นจากการทำนายค่าจุดภาพที่ต้องการ เช่น จากรูปที่ 2.25 เมื่อเราต้องการทำนายค่า X เราสามารถเลือกใช้สมการการทำนายค่า X ได้จากตารางที่ 2.11 ซึ่งจะทำนายค่า X จากค่า A, B และ C ซึ่งเป็นค่าจุดภาพใกล้เคียง เราสามารถเลือกใช้สมการใดก็ได้จากตารางนี้ หากเลือกสมการแบบ 1 2 และ 3 เป็นการทำนายแบบ 1 มิติ เลือกแบบ 4 5 6 7 เป็นการทำนายแบบ 2 มิติ เลือก 0 จะสามารถใช้ได้กับการเข้ารหัสดิฟเฟอเรนเชียลของประเภทการเข้ารหัสเชิงลำดับขั้นเท่านั้น จากนั้นก็จะลบค่าที่ทำนายจากค่าจุดภาพแท้จริงของ X และนำค่าที่ลบแล้วมาเข้ารหัสแบบไม่มีการสูญเสียข้อมูลด้วยวิธีการเข้ารหัสฮัฟฟ์แมนหรือการเข้ารหัสเลขคณิต ปัจจุบันวิธีนี้ไม่นิยมและไม่มีโปรแกรมตัวจริงประยุกต์ใดๆ สนับสนุนการบีบภาพแบบนี้มากนัก

	C	B		
	A	X		

รูปที่ 2.25 ตัวอย่างการทำนายค่าจากค่าใกล้เคียง (Wallace, 1991)

ค่าที่เลือก	การทำนาย (Prediction)
0	ไม่มีการทำนาย
1	A
2	B
3	C
4	$A+B-C$
5	$A+(B-C)/2$
6	$B+(A-C)/2$
7	$(A+B)/2$

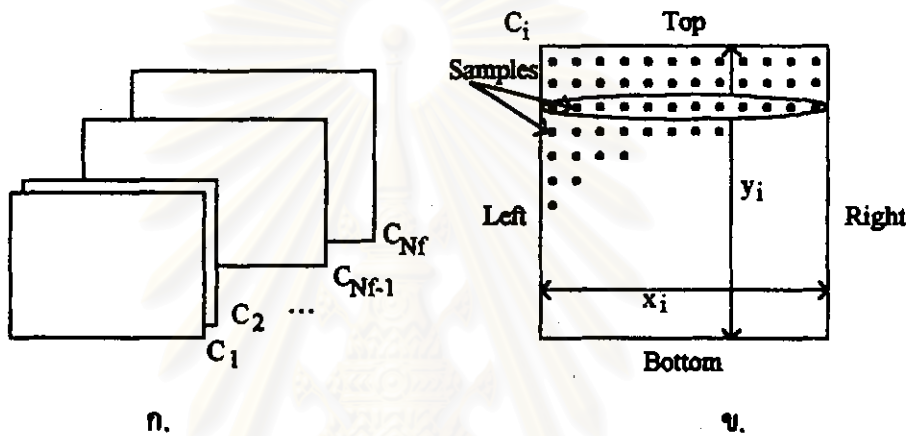
ตารางที่ 2.11 ตารางแสดงการทำนายค่า X จากค่า A, B และ C (Wallace, 1991)

2.2.4.1.4 ภาพหลายองค์ประกอบ (Multiple-component images)

จากที่กล่าวไปแล้วซึ่งเป็นการเข้ารหัสภาพที่มีเพียงองค์ประกอบเดียวหรือภาพ
 เกรย์สเกล แต่ข้อกำหนดของเจ็ทจะสามารถบีบอัดภาพได้มากกว่าหนึ่งองค์ประกอบภาพ
 (Multiple-component) เช่น ภาพสี ซึ่งปกติจะมี 3 องค์ประกอบตามหลักแบบจำลองสีอาร์จีบี และ
 ได้กำหนดให้เปลี่ยนแบบจำลองภาพสีให้เป็นแบบจำลองสีวายซีอาร์จีบีหรือวายซีวี เป็นต้น

จากรูปที่ 2.26 แสดงแบบจำลองภาพคั่นฉบับเจ็ท ภาพคั่นฉบับสามารถมีองค์
 ประกอบของภาพจาก 1 ถึง 255 องค์ประกอบ บางครั้งเรียกว่า สี หรือ คลื่นสเปกตรัม หรือ สัญญาณ
 แต่ละองค์ประกอบประกอบด้วยอะเรย์ของแซมเปิล (Sample) (หรือจุดภาพในตำแหน่งอะเรย์ 2 มิติ

ที่รวมเป็นองค์ประกอบของภาพ) แชนเนลจะถูกกำหนดให้เป็นจำนวนเต็มไม่มีเครื่องหมาย (unsigned integer) ด้วยค่าพริซิชั่น (Precision) P บิต (พริซิชั่น คือ จำนวนบิตที่กำหนดค่าให้แชนเนลเปิดโดยเฉพาะหรือสัมประสิทธิ์คิซิติ) ซึ่งจะมีค่าอยู่ในช่วง $[0, 2^P-1]$ ทุกๆ แชนเนลของทุกๆ องค์ประกอบภายในภาพต้นฉบับเดียวกันจะมีค่าพริซิชั่น P ค่าเดียวกัน กล่าวคือ เท่ากับ 8 หรือ 12 สำหรับการเข้ารหัสแบบดิซิติเบส (DCT-based coding) และเท่ากับ 2 ถึง 16 สำหรับการเข้ารหัสแบบทำนายค่า (Predictive coding)



รูปที่ 2.26 แสดงแบบจำลองภาพต้นฉบับเชิงเทก รูป ก. ภาพต้นฉบับที่มีหลายองค์ประกอบ รูป ข. ลักษณะเฉพาะขององค์ประกอบภาพ (Wallace, 1991)

จากรูปที่ 2.26 ภาพต้นฉบับประกอบด้วยจำนวนองค์ประกอบเท่ากับ Nf แต่ละองค์ประกอบ C_i ประกอบด้วยอะเรย์ของแชนเนลคอดดัมภ์ x_i โดยบรรทัด y_i ขนาดความกว้างยาวขององค์ประกอบได้จากพารามิเตอร์ 2 ค่า คือ X และ Y โดยที่ X คือ ค่ามากที่สุดของ x_i และ Y เป็นค่ามากที่สุดของ y_i สำหรับทุกๆ องค์ประกอบในเฟรม (Frame) สำหรับแต่ละองค์ประกอบคำนวณตามการต่อไปนี้

$$x_i = \left\lceil X \times \frac{H_i}{H_{\max}} \right\rceil \quad y_i = \left\lceil Y \times \frac{V_i}{V_{\max}} \right\rceil$$

โดยที่ H_i และ V_i เป็นอัตราการชักตัวอย่าง (Sampling factors), H_{\max} และ V_{\max} เป็นอัตราการชักตัวอย่างสูงสุดสำหรับทุกๆ องค์ประกอบในเฟรม และ $\lceil \cdot \rceil$ เป็นฟังก์ชันฉัดถึง (ceiling function)

ตัวอย่างเช่น พิจารณาภาพที่มี 3 องค์ประกอบ ประกอบด้วยขนาดความกว้างยาวมากที่สุด 512 บรรทัด และ 512 แกรมเปิดต่อบรรทัด ด้วยอัตราการจัดตัวอย่างดังนี้

องค์ประกอบ 0	$H_0 = 4,$	$V_0 = 1$
องค์ประกอบ 1	$H_1 = 2,$	$V_1 = 2$
องค์ประกอบ 2	$H_2 = 1,$	$V_2 = 1$

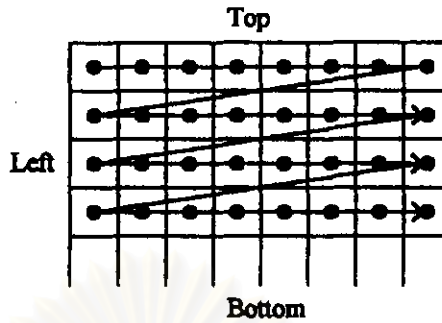
จากนั้น $X = 512, Y = 512, H_{\text{รวม}} = 4, V_{\text{รวม}} = 2$ และ x_i และ y_i สำหรับแต่ละองค์ประกอบเป็น

องค์ประกอบ 0	$x_0 = 512,$	$y_0 = 256$
องค์ประกอบ 1	$x_1 = 256,$	$y_1 = 512$
องค์ประกอบ 2	$x_2 = 128,$	$y_2 = 256$

2.2.4.1.5 ลำดับการเข้ารหัสและเทคนิคการประมวลผลโดยแทรกการทำงานหรืออินเทอร์ลีฟวิ่ง (Interleaving)

มาตรฐานการบีบภาพที่เหมาะสมจะต้องบอกได้ว่าระบบจะประมวลผลข้อมูลระหว่างขั้นตอนการคลาออย่างไร หลายๆ การประยุกต์ (Application) จำเป็นต้องลำดับขั้นตอนของการแสดงผลหรือการพิมพ์ภาพหลายๆ ครั้งในลักษณะพารารลกับขั้นตอนของการคลาอ ดังนั้นองค์ประกอบทั้งหลายจึงต้องมีการอินเทอร์ลีฟ (Interleave) ภายในแฟ้มข้อมูลภาพ

เพื่อสร้างระบบการอินเทอร์ลีฟที่ใช้ทั้งการเข้ารหัสดิคซีทีเบสและการเข้ารหัสการทำนาย เชนทีกได้ให้นิยามของ "data unit" หรือหน่วยข้อมูลคือ แกรมเปิดหรือค่าจุดภาพในวิธีการเข้ารหัสแบบทำนายค่า หรือบล็อกขนาด 8×8 ในวิธีการเข้ารหัสแบบดิคซีทีเบส และคำนิยามของ "Minimum Coded Unit" (MCU) คือ หน่วยเข้ารหัสที่เล็กที่สุดของหน่วยข้อมูลที่เป็นอินเทอร์ลีฟจากรูปแบบแฟ้มข้อมูลในส่วนหัวกราดตรวจ (Scan header) ซึ่งเป็นส่วนที่กำหนดลำดับของหน่วยข้อมูลหน่วยใดจะถูกเข้ารหัสและถูกวางในแฟ้มข้อมูลภาพที่บีบ



รูปที่ 2.27 ลำดับข้อมูลแบบนอนอินเทอร์ลิต (Non-interleaved Data Ordering) (Wallace, 1991)

ถ้าพารามิเตอร์ของส่วนหัวกราดตรวจ ($N_s = 1$) หมายถึง มีเพียงหนึ่งองค์ประกอบ การเข้ารหัสจะเป็นนอนอินเทอร์ลิต (Non-interleaved) องค์ประกอบถูกระบุโดยค่าพารามิเตอร์ C_{s_1} ดังนั้นเอ็มซียู (MCU) จะเป็น 1 หน่วยข้อมูลและลำดับของหน่วยข้อมูลภายในกราดตรวจ จะเรียงจากซ้ายไปขวา และจากบนลงล่าง แสดงดังรูปที่ 2.27 โดยไม่ใช้ค่าพารามิเตอร์ H_x และ V_x

ถ้า $N_s > 1$ หมายถึง มีมากกว่าหนึ่งองค์ประกอบ การเข้ารหัสจะเป็นอินเทอร์ลิต องค์ประกอบจะถูกระบุโดยค่าพารามิเตอร์ C_{s_1} ถึง $C_{s_{N_s}}$ ดังนั้นเอ็มซียูจะเป็นลำดับของหน่วยข้อมูลที่กำหนดโดยอัตราการจัดตัวขององค์ประกอบภายในกราดตรวจ แต่ละองค์ประกอบกราดตรวจ C_{s_j} จะถูกแบ่งเป็นส่วนๆ ลงในอะเรย์สี่เหลี่ยมของหน่วยข้อมูล H_x ตามแนวนอนและหน่วยข้อมูล V_x ตามแนวตั้ง ค่าพารามิเตอร์ H_x และ V_x อยู่ในข้อกำหนดองค์ประกอบส่วนหัวเฟรม (frame header component-specification) เป็นค่าที่บอกจำนวนหน่วยข้อมูลต่อองค์ประกอบ C_{s_j} ในแต่ละเอ็มซียู ภายในแต่ละ H_x โดยอะเรย์ของ V_x หน่วยข้อมูลจะถูกเรียงลำดับจากซ้ายไปขวาและบนลงล่าง และรูปแบบของอะเรย์จะถูกเรียงลำดับจากซ้ายไปขวา และจากบนลงล่างภายในแต่ละองค์ประกอบ ลำดับขององค์ประกอบในกราดตรวจจะตรงกันกับลำดับที่ระบุในส่วนหัวเฟรม (Frame header)

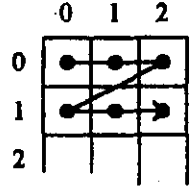
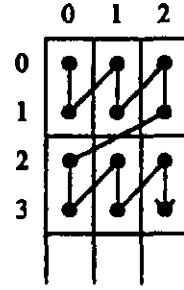
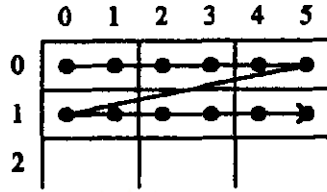
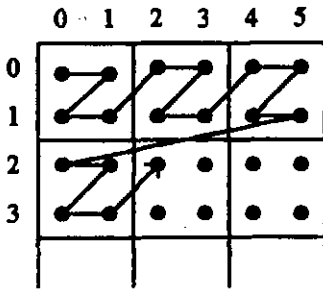
จากรูปที่ 2.28 $N_s = 4$ และ MCU_1 ประกอบด้วยหน่วยข้อมูลที่นำมาจากบริเวณบนซ้ายสุดของ C_{s_1} ตามด้วยหน่วยข้อมูลจากบริเวณเดียวกันของ C_{s_2} , C_{s_3} และ C_{s_4} ตามลำดับ MCU_2 ก็จะมีรูปแบบต่อเนื่องตามลำดับ จากนั้นข้อมูลที่อินเทอร์ลิตจะเรียงลำดับตามเอ็มซียู

$Cs_1: H_1 = 2, V_1 = 2$

$Cs_2: H_2 = 2, V_2 = 1$

$Cs_3: H_3 = 1, V_3 = 2$

$Cs_4: H_4 = 1, V_4 = 1$



$$\begin{aligned}
 MCU_1 &= d_{00}^1 d_{01}^1 d_{10}^1 d_{11}^1 & d_{00}^2 d_{01}^2 & d_{00}^3 d_{10}^3 & d_{00}^4 \\
 MCU_2 &= d_{02}^1 d_{03}^1 d_{12}^1 d_{13}^1 & d_{02}^2 d_{03}^2 & d_{01}^3 d_{11}^3 & d_{01}^4 \\
 MCU_3 &= d_{04}^1 d_{03}^1 d_{14}^1 d_{15}^1 & d_{04}^2 d_{03}^2 & d_{02}^3 d_{12}^3 & d_{02}^4 \\
 MCU_4 &= d_{20}^1 d_{21}^1 d_{30}^1 d_{31}^1 & d_{10}^2 d_{11}^2 & d_{20}^3 d_{30}^3 & d_{10}^4
 \end{aligned}$$

หน่วยข้อมูล Cs_1 Cs_2 Cs_3 Cs_4

รูปที่ 2.28 ตัวอย่างการเรียงลำดับข้อมูลแบบอินเทอร์ลิฟ (Wallace, 1991)

จำนวนของหน่วยข้อมูลที่อยู่ในเอ็มซียู จะถูกกำหนดโดยจำนวนขององค์ประกอบอินเทอร์ลิฟและอัตราการชักตัวอย่างที่สัมพันธ์กับจำนวนมากที่สุดขององค์ประกอบที่สามารถอินเทอร์ลิฟ คือ 4 และจำนวนมากที่สุดของหน่วยข้อมูลในเอ็มซียู คือ 10 แต่จึ่งที่อนุญาตให้บางองค์ประกอบเป็นอินเทอร์ลิฟและบางองค์ประกอบไม่เป็นอินเทอร์ลิฟได้ภายในเพิ่มข้อมูลภาพเดียวกัน สมการต่อไปนี้เป็น การหาจำนวนรวมขององค์ประกอบที่ถูกอินเทอร์ลิฟ

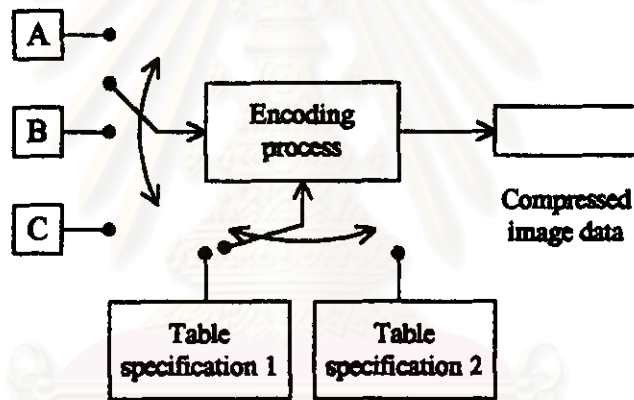
$$\sum_{\text{all } i \text{ in interleave}} H_i \times V_i \leq 10$$

2.2.4.1.6 ตารางหลายชุด (Multiple tables)

การบีบอัดภาพเจทีคจะควบคุมการใช้ตารางให้เหมาะสมกับองค์ประกอบนั้นๆ ตารางควอนไทเซชันและตารางเข้ารหัสฮัฟฟ์ที่แมนเคียวกันจะถูกใช้เพื่อเข้ารหัสในทุกๆ ค่าแซมเปิลในหนึ่งองค์ประกอบ เจทีคจะสามารถเก็บตารางควอนไทเซชันได้ถึง 4 ตารางที่แตกต่างกันและตารางเข้ารหัสฮัฟฟ์ที่แมนถึง 4 ตารางที่แตกต่างกัน แต่การบีบอัดภาพแบบเวสโตนิงลำดับเป็นข้อยกเว้น

ซึ่งเก็บตารางเข้ารหัสฮัฟฟ์แมนได้เพียง 2 ตารางเท่านั้น ด้วยเหตุที่มีหลายตาราง จึงจำเป็นต้องมีการสวิตซ์ซึ่งระหว่างตารางในระหว่างการประมวลผลคลายภาพในลำดับที่จะเลือกตารางที่เหมาะสมให้กับองค์ประกอบที่เหมาะสม เพราะตารางไม่สามารถไหลคในระหว่างประมวลผลการคลายภาพในหนึ่งกราดตรวจ (Scan)

จากรูปที่ 2.29 แสดงการควบคุมการสวิตซ์ซึ่งตารางและตัวอย่างของการประมวลผลการเข้ารหัสว่าจะเลือกวิธีใดให้องค์ประกอบภาพหลายๆ องค์ประกอบเข้ารหัสให้เหมาะสมกับข้อมูลตารางหลายๆ ชุด เมื่อกระทำการประมวลผลเข้ารหัส จากรูปแสดงภาพคั่นฉบับซึ่งประกอบไปด้วย 3 องค์ประกอบ A, B และ C และมีข้อกำหนดตาราง 2 ชุด (ไม่มีข้อแตกต่างระหว่างตารางควอนไทซ์และตารางเข้ารหัสเอ็นโทรปี)



รูปที่ 2.29 องค์ประกอบอินเทอร์ลีฟ (Component-interleave) และการควบคุมสวิตซ์ซึ่งตาราง (Table-switching control) (Wallace, 1991)

2.2.4.1.7 ตัวอย่างการบีบข้อมูลภาพด้วยเทคนิคเจเท็ก

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

ตารางที่ 2.12 ตัวอย่างค่าจุดภาพบางส่วนของภาพคั่นฉบับขนาด 8x8 (Wallace, 1991)

235.6	-1.0	-12.1	-5.2	2.1	-1.7	-2.7	1.3
-22.6	-17.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

ตารางที่ 2.13 ตัวอย่างค่าสัมประสิทธิ์หลังผ่านขั้นตอนฟอร์เวิร์ดคิซิกี้ (Wallace, 1991)

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

ตารางที่ 2.14 ตัวอย่างของค่าในตารางคอนโทเจนซ์ (Wallace, 1991)

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ตารางที่ 2.15 ตัวอย่างค่าสัมประสิทธิ์หลังผ่านคอนโทเจนซ์และนำไปเข้ารหัส (Wallace, 1991)

240	0	-10	0	0	0	0	0
-24	-12	0	0	0	0	0	0
-14	-13	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ตารางที่ 2.16 ตัวอย่างค่าสัมประสิทธิ์ที่ผ่านการคิดคอนโทเจนซ์ (Wallace, 1991)

144	146	149	152	154	156	156	156
148	150	152	154	156	156	156	156
155	156	157	158	158	157	156	155
160	161	161	162	161	159	157	155
163	163	164	163	162	160	158	156
163	164	164	164	162	160	158	157
160	161	162	162	162	161	159	158
158	159	161	161	162	161	159	158

ตารางที่ 2.17 ตัวอย่างค่าคุณภาพที่ได้กลับมาหลังการคลาย (Wallace, 1991)

จากตารางที่ 2.12 ซึ่งแสดงถึงตัวอย่างบางส่วนของภาพบล็อก 8x8 หลังจากลดค่าทุกค่าในบล็อกด้วย 128 และนำเข้าการแปลงด้วยฟอว์เวิร์ดคิซิจิ ได้ผลลัพธ์ตามตารางที่ 2.13 ตารางที่ 2.14 เป็นตัวอย่างของตารางควอนไทซ์สำหรับองค์ประกอบภูมิแนวนอน ตารางที่ 2.15 แสดงถึงสัมประสิทธิ์ที่ถูกควอนไทซ์แล้วและนำไปเข้ารหัสต่อไป ตารางที่ 2.16 แสดงถึงสัมประสิทธิ์ที่ผ่านการดีควอนไทซ์ จากนั้นนำไปผ่านการแปลงกลับด้วยอินเวอร์สคิซิจิ จะได้ผลลัพธ์ตามตารางที่ 2.17 ที่แสดงถึงบล็อกตัวอย่างของภาพหลังจากการคลาย

ตัวอย่างการเข้ารหัสจากตารางที่ 2.15 สัมประสิทธิ์คิซิจิแทนด้วยสัญลักษณ์ (Size) (Amplitude) และสัมประสิทธิ์เอซิทแทนด้วยสัญลักษณ์ (Runlength, Size) (Amplitude) ดังนี้

สมมติให้ค่าสัมประสิทธิ์คิซิจิในบล็อกที่ม้วนมีค่าเท่ากับ 12 ดังนั้นความแตกต่างของค่าสัมประสิทธิ์คิซิจิในบล็อกที่ม้วนและบล็อกตัวอย่างเท่ากับ +3 ดังนั้น จึงแทนด้วย (2)(3) ส่วนสัมประสิทธิ์เอซิทจะแทนที่ด้วยสัญลักษณ์จากลำดับซิกแซกต่อไปนี้ (1,2)(-2), (0,1)(-1), (0,1)(-1), (0,1)(-1), (2,1)(-1), (0,0)

รหัสที่ใช้เข้ารหัสสัมประสิทธิ์คิซิจิในสัญลักษณ์ที่ 1 มาจากค่าในตารางที่ 2.18

(2) 011

รหัสที่ใช้เข้ารหัสสัมประสิทธิ์เอซิทในสัญลักษณ์ที่ 1 มาจากค่าในตารางที่ 2.19

(0,0) 1010

(0,1) 00

(1,2) 11011

(2,1) 11100

จำนวนบิตที่ใช้เข้ารหัสตัวประพจน์ดิซิจนและเอชในสัญลักษณ์ที่ 2 มาจากตารางที่

2.20

(3) 11

(-2) 01

(-1) 0

จะได้รหัสดังนี้ 011111101101000000001110001010 เท่ากับ 31 บิตต่อบล็อก

Size	ความยาวรหัส	รหัสที่ใช้
0	2	00
1	3	010
2	3	011
3	3	100
4	3	101
5	3	110
6	4	1110
7	5	11110
8	6	111110
9	7	1111110
A	8	11111110
B	9	111111110

ตารางที่ 2.18 แสดงรหัสที่ใช้เข้ารหัสภูมิแนนซ์ตัวประพจน์ดิซิจนของตารางอัฟที่แมน
(Mattison, 1994: 341)

Runlength, Size	ความยาวรหัส	รหัสที่ใช้
(0,0)	4	1010
(0,1)	2	00
(0,2)	2	01
(0,3)	3	100
...
(1,2)	5	11011
...
(2,1)	5	11100
...

ตารางที่ 2.19 แสดงบางส่วนของรหัสที่ใช้เข้ารหัสตามนัยสัมประสิทธิ์เอจิบของตารางฮัทฟีแมน
(Mattison, 1994: 332)

Size	Amplitude
0	0
1	-1, 1
2	-3, -2, 2, 3
3	-7...-4, 4...7
4	-15...-8, 8...15
5	-31...-16, 16...31
6	-63...-32, 32...63
7	-127...-64, 64...127
8	-255...-128, 128...255
9	-511...-256, 256...511
A	-1023...-512, 512...1023
B	-2047...-1024, 1024...2047

ตารางที่ 2.20 แสดงจำนวนบิตที่ใช้เข้ารหัสสัมประสิทธิ์คี่และเอจิบที่ค่าแอมพลิจูดต่างๆ
(Mattison, 1994: 330)

2.2.4.1.8 ขั้นตอนการบีบภาพในประเภทการเข้ารหัสดิจิทัลโทรทรรศน์

การบีบในประเภทดิจิทัลโทรทรรศน์ประกอบไปด้วยการแปลงฟอร์เวิร์ดดิจิทัลและควอนไทเซชันที่เหมือนกันกับที่ใช้ในแบบสโตนเจ็ททุกอย่าง แดกต่างกันที่แต่ละส่วนประกอบของภาพจะเข้ารหัสในลักษณะการกราดภาพหลายๆ ครั้ง การกราดภาพครั้งแรกจะได้ภาพหยาบคุณภาพต่ำสุดที่คั่งได้ ที่มองเห็นลักษณะของภาพคร่าวๆ ซึ่งจะทำให้การส่งผ่านเครือข่ายเร็วกว่าการส่งผ่านภาพทั้งภาพซึ่งจะช้ากว่า และการกราดภาพครั้งต่อไปภาพจะละเอียดคมชัดมากขึ้นเรื่อยๆ จนมีคุณภาพใกล้เคียงกับการบีบด้วยแบบสโตนเจ็ท ข้อเสียคือในแต่ละการกราดภาพแต่ละครั้งจะใช้จำนวนครั้งในการคำนวณเช่นเดียวกับการแสดงภาพทั้งภาพแบบสโตนเจ็ท และต้องเพิ่มบิตเฟอริให้หน่วยความจำมากขึ้น สัมประสิทธิ์ที่อยู่ในบิตเฟอริจะเข้ารหัสแยกกันไปในกราดภาพแต่ละครั้ง การเข้ารหัสเพื่อส่งผ่านมี 2 วิธี วิธีแรกเรียกว่า "Spectral selection" จะเลือกสัมประสิทธิ์ในบางช่วงความถี่ เช่น สัมประสิทธิ์ในช่วงความถี่ต่ำมาเข้ารหัสและส่งผ่านก่อน ซึ่งอาจจะส่งผ่านภาพเรื่องความถี่ต่ำชัดซึกซกก็ได้ วิธีที่สองเรียกว่า "Successive approximation" เลือกบิตที่มีความสำคัญคือการลดอันดับของภาพมาเข้ารหัสและส่งผ่านไปก่อน แล้วจึงค่อยส่งบิตที่ไม่มีความสำคัญมากนักต่อไป เราสามารถใช้อำนาจทั้งสองแยกกันหรือรวมกันก็ได้

2.2.4.1.9 ขั้นตอนการบีบภาพในประเภทการเข้ารหัสเชิงลำดับขั้น

การบีบภาพประเภทการเข้ารหัสเชิงลำดับขั้นจะมีการเข้ารหัสภาพในลักษณะ "pyramid" คือ ภาพที่เข้ารหัสมีผลัดพรหลายๆ ภาพโดยที่แต่ละภาพจะมีความคมชัด (resolution) ต่างกัน ขั้นตอนการเข้ารหัสภาพดังกล่าวโดยสรุปดังนี้

1. กรอง (Filter) หรือดาวน์โหลด (Downsample) ภาพต้นฉบับโดยจำนวนที่ต้องการจะต้องเป็นผลคูณของ 2 ทั้งทางแนวนอนและแนวตั้ง
2. เข้ารหัสภาพที่ลดขนาดลงโดยใช้วิธีแบบสโตนเจ็ท, ดิจิทัลโทรทรรศน์หรือการเข้ารหัสแบบไม่สูญเสียข้อมูล
3. ออกรหัสภาพที่ลดขนาดลง และอัพโหลด (Upsample) ด้วย 2 ทั้งทางแนวนอนและแนวตั้ง.
4. ใช้ภาพที่เพิ่มค่าจากขั้นตอน 3 มาทำขนาดค่าจุดภาพต้นฉบับที่ความคมชัดนั้นๆ และเข้ารหัสภาพโดยใช้วิธีแบบสโตนเจ็ท, ดิจิทัลโทรทรรศน์หรือการเข้ารหัสแบบไม่สูญเสียข้อมูล

5. กระทำซ้ำขั้นตอน 3 และ 4 จนกว่าจะได้ภาพที่มีความคมชัดที่สมบูรณ์ที่สุด

การบีบภาพประเภทนี้มีประโยชน์มากในโปรแกรมมอดรูดประโยชน์ที่ภาพมีความคมชัดสูงจะต้องแสดงด้วยอุปกรณ์แสดงผลที่มีความคมชัดต่ำ เช่น ภาพที่กราดและบีบที่ความคมชัดสูงถ้าหรับใช้กับเครื่องพิมพ์ที่มีคุณภาพดีมาก แต่การแสดงผลบนจอภาพมีความคมชัดต่ำ

2.2.4.2 การบีบข้อมูลภาพแบบเวฟเลต

2.2.4.2.1 ความเป็นมาของเวฟเลต (โชติวัฒน์ อุณนาภิรักษ์, 2538) (Grap, 1995) (Umbaugh, 1998: 125-130)

เวฟเลตเกิดขึ้นมาโดยมีพื้นฐานมาจากฟูรีเยร์ (Fourier) หลังจากที่ฟูรีเยร์เกิดขึ้นได้ไม่นาน ฟูรีเยร์เกิดขึ้นในปี 1807 โดย Jean Baptiste Joseph Fourier (1768-1830) นักคณิตศาสตร์ชาวฝรั่งเศสซึ่งได้พัฒนาฟูรีเยร์เพื่อใช้ในการอธิบายการแพร่ของอุณหภูมิและการนำความร้อน และหลังจากนั้นพบว่าฟูรีเยร์มีประโยชน์กับงานหลายๆ ด้าน เช่น การวิเคราะห์ความถี่สัญญาณในงานวิศวกรรมเครื่องกล การวิเคราะห์วงจรในงานวิศวกรรมไฟฟ้า และในงานด้านประมวลผลภาพคอมพิวเตอร์ เป็นต้น แต่ด้วยข้อจำกัดบางประการที่ทำให้งานวิจัยบางงานใช้ได้ไม่เหมาะสม จึงทำให้เกิดเวฟเลตขึ้น ซึ่งมีการกล่าวถึงเวฟเลตครั้งแรกในภาคผนวกในวิทยานิพนธ์ของ A.Hart ในปี 1909 ต่อจากนั้นก็มีการพัฒนาเวฟเลตในหลายๆ สาขาโดยเป็นอิสระต่อกัน เช่น

ในช่วงปี 1930 Paul Levy นักฟิสิกส์ได้ศึกษา Brownian motion, Littlewood, Paley และ Stein ได้พัฒนา Littlewood-Paley theory

ช่วงปี 1960-1980 Guido Weiss และ Ronald R. Coifman ศึกษา atomic decomposition, ในปี 1964 Calderon สาขาคณิตศาสตร์บริสุทธิ์ได้พัฒนา Calderon-Zygmund theory, ในปี 1968 Aslaksen และ Klander สาขาควอนตัมฟิสิกส์, ในปี 1976 Croisier, Esteban และ Galand ได้พัฒนา QMF filter, ในปี 1980 Alex Grossman และ Jean Morlet นักฟิสิกส์และวิศวกร ได้ให้คำจำกัดความเวฟเลตในสาขาควอนตัมฟิสิกส์ซึ่งเป็นจุดเริ่มต้นของทฤษฎีเวฟเลต, ในปี 1980 Stromberg ค้นพบ Orthogonal wavelet ชนิดแรก

ช่วงหลังปี 1980, ในปี 1983 Morlet ในงานการวิเคราะห์ข้อมูลแผ่นดินไหว, ในปี 1985 Stephane Mallat ได้นำแนวคิดเริ่มต้นในงานด้านการประมวลผลสัญญาณดิจิทัล (Digital signal processing) มาพบความสัมพันธ์ระหว่าง Quadrature mirror filters, Pyramid algorithm และ Orthonormal wavelet bases และด้วยผลดุษณีนี้ทำให้ Yves Meyer ได้สร้าง non-trivial wavelet แบบแรก อีก 2 ปีต่อมา Ingrid Daubechies ได้ใช้งานของ Mallat สร้างชุดของ Wavelet basis function ที่ดีที่สุดซึ่งได้กลายเป็นการประยุกต์แนวคิดที่สำคัญที่สุดในปัจจุบัน, ในปี 1986 Vetterli ในงานด้านวิศวกรรมไฟฟ้า เป็นต้น

ในที่นี้ Basis function คือ ฟังก์ชันพื้นฐานหรือฟังก์ชันเล็กที่สุด เช่น เวกเตอร์ 2 มิติ (x, y) เป็นการรวมกันของเวกเตอร์ $(1, 0)$ และ $(0, 1)$ ด้วยการคูณ x กับเวกเตอร์ $(1, 0)$ ได้ผล $(x, 0)$ และการคูณ y กับเวกเตอร์ $(0, 1)$ ได้ผล $(0, y)$ เมื่อรวมกันก็จะได้เวกเตอร์ (x, y) และฟังก์ชันพื้นฐานที่คตินั้นจะต้องมีคุณสมบัติเชิงตั้งฉาก (Orthogonal property) ดังนั้น Wavelet basis function คือ ฟังก์ชันชุดแรกสุดหรือฟังก์ชันที่เล็กที่สุดที่ประกอบเป็นฟังก์ชันเวฟเลต (Grap, 1995)

2.2.4.2.2 การแปลงเวฟเลต (โชคิวัฒน์ อุณาภิรักษ์, 2538) (Rioul, 1991) (Burrus, Gopinath และ Guo, 1998) (Vetterli, 1995) (Grap, 1995)

เวฟเลตเป็นวิธีการทางคณิตศาสตร์ที่ถูกสร้างขึ้นมาเพื่อการแบ่งสัญญาณหรือฟังก์ชันออกเป็นส่วนๆ โดยที่แต่ละส่วนนั้นมีองค์ประกอบในเชิงความถี่แตกต่างกันไป และศึกษาหรือวิเคราะห์แต่ละส่วนด้วยความละเอียด (Resolution) ที่เหมาะสมกับมาตราส่วน (Scale) ของส่วนนั้นๆ

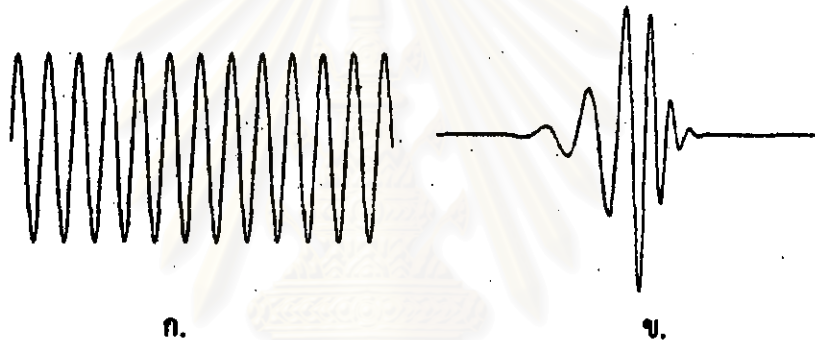
โดยทั่วไปแล้วการวิเคราะห์สัญญาณอาจทำได้ 2 วิธี คือ

1. การวิเคราะห์ในเชิงเวลา (Time domain)
2. การวิเคราะห์ในเชิงความถี่ (Frequency domain)

การวิเคราะห์ในเชิงความถี่นั้นนิยมทำโดยการแปลงสัญญาณเป็นเชิงความถี่ก่อน ซึ่งวิธีที่นิยมใช้กันคือ การแปลงฟูเรียร์ (Fourier transform) การแปลงด้วยฟูเรียร์นี้จะวิเคราะห์สัญญาณเฉพาะในเชิงความถี่เท่านั้น แต่ในขณะที่การแปลงเวฟเลตสามารถวิเคราะห์ในเชิงเวลาและ

ความถี่ของสัญญาณด้วย จึงอาจเรียกได้ว่าเป็นเครื่องมือในการวิเคราะห์สัญญาณในเชิงเวลา-ความถี่ (Time-frequency domain)

การแปลงฟูเรียร์จะมีฟังก์ชันพื้นฐานอยู่ในเทอมของฟังก์ชันชาซน์ และฟังก์ชันโคชาซน์ซึ่งเป็นฟังก์ชันตรีโกณมิติที่มีคุณสมบัติเชิงคั้งฉาก เหมาะสมกับสัญญาณที่มีลักษณะคงตัว (Stationary signal) เท่านั้น แต่ในทางปฏิบัติสัญญาณที่นำมาวิเคราะห์มักจะมีลักษณะเป็นสัญญาณไม่คงตัว (Non-Stationary signal) หรือเป็นสัญญาณคงตัวเป็นช่วง (Quasi-Stationary signal) ดังนั้นการวิเคราะห์โดยใช้การแปลงฟูเรียร์จะไม่เหมาะสมนัก รูปที่ 2.30 แสดงตัวอย่างของฟังก์ชันชาซน์และฟังก์ชันเวฟเลตของ Daubechies



รูปที่ 2.30 แสดงตัวอย่างของเวฟและเวฟเลต รูป ก. ฟังก์ชันชาซน์
รูป ข. ฟังก์ชันเวฟเลตของ Daubechies (Burrus, Gopinath และ Guo, 1998: 1)

Gabor ได้พัฒนา การแปลงฟูเรียร์ช่วงสั้น (Short-Time Fourier Transform หรือ STFT) ซึ่งจะวิเคราะห์สัญญาณในช่วงเวลาสั้นๆ เพื่อหาความถี่ชั่วขณะ (Instantaneous frequency) โดยการมองสัญญาณผ่านหน้าต่างต่าง (window) $g(t)$ โดยมีจุดศูนย์กลางที่ตำแหน่ง τ แล้วจึงทำการแปลงฟูเรียร์กับสัญญาณที่ผ่านหน้าต่างแล้ว $x(t)g^*(t-\tau)$ ด้วยเหตุที่ทำการแปลงในหน้าต่าง การแปลงฟูเรียร์ช่วงสั้นจึงเรียกได้อีกชื่อว่า Windowed Fourier Transform หรือ WFT จะได้สมการการแปลงฟูเรียร์ช่วงสั้นดังนี้

$$STFT(\tau, f) = \int x(t)g^*(t-\tau)e^{-2\pi ft} dt$$

อย่างไรก็ตาม จะเห็นได้ว่าการแปลงฟูเรียร์ช่วงสั้นยังมีข้อจำกัดที่ไม่สามารถเปลี่ยนความละเอียด (Resolution) ในการวิเคราะห์สัญญาณได้ ในบางกรณี สัญญาณที่นำมาวิเคราะห์อาจมีอยู่หลายองค์ประกอบ ถ้าหากว่ามีบางองค์ประกอบเป็นสัญญาณที่คงตัวในช่วง

สั้นๆ และมีบางองค์ประกอบคงตัวอยู่ในช่วงระยะเวลาที่ยาวกว่า จะเห็นได้ว่าเราจะไม่สามารถวิเคราะห์สัญญาณโดยที่องค์ประกอบของสัญญาณทั้งสองส่วนถูกวิเคราะห์ที่ความละเอียดที่เหมาะสมทั้งคู่

เวฟเลตถูกพัฒนาขึ้นเพื่อแก้ไขข้อจำกัดของการแปลงฟูเรียร์ช่วงสั้นในเรื่องของระดับความละเอียดในการวิเคราะห์ที่เปลี่ยนแปลงไม่ได้ โดยเวฟเลตจะถูกใช้อธิบายการสร้างรูปแบบของสัญญาณ ซึ่งประกอบขึ้นด้วยเซตของสัญญาณเฉพาะมารวมกันเป็นสัญญาณนั้นๆ สัญญาณเฉพาะนี้จะเปลี่ยนถี่ๆ เรียกว่า “เวฟเลต” การนำเวฟเลตหลายๆ อันมารวมกันเป็นเซต เพื่อจะใช้ในการอธิบายสัญญาณใดๆ นั้น คลื่นเวฟเลตแต่ละตัวจะต้องมีโครงสร้างมาจากฟังก์ชันเคียวกัน เรียกว่า มานธอร์เวฟเลต (Mother wavelet) $\psi(t)$ หรือฟังก์ชันการวิเคราะห์ (Analyzing function) ฟังก์ชันดังกล่าวจะมีคุณสมบัติดังนี้ คือ

- well localized คือ มีค่าเข้าใกล้ศูนย์เมื่อ เข้าใกล้อนันต์
- การแกว่ง (oscillation) เพื่อแสดงความเป็นคลื่น โดยอินทิกรัลของ $\psi(t)$ และ m โมเมนต์แรกของ $\psi(t)$ มีค่าเป็นศูนย์ ดังนี้

$$0 = \int_{-\infty}^{\infty} \psi(t) dt = \dots = \int_{-\infty}^{\infty} t^{m-1} \psi(t) dt$$

จากมานธอร์เวฟเลต $\psi(t)$ นี้จะสร้างเวฟเลตอื่นๆ $\psi_{a,\tau}(t)$ โดยเวฟเลตแต่ละตัวจะเกิดจากการสเกลหรือมาตราส่วนหรือเปลี่ยนความละเอียด (Scaling or Dilation: “ a ”) ซึ่งปกติมีค่าเป็น 1 และการเลื่อนตำแหน่งในเชิงเวลา (Translation: “ b ”) ซึ่งปกติพิจารณาที่ $b = 0$ ของ มานธอร์เวฟเลต นอกจากนี้ เวฟเลตที่ถูกสเกลจะต้องทำการปรับให้อยู่ในภาวะปกติ (Normalize) ด้วย $\frac{1}{\sqrt{a}}$ เสมอ เพื่อให้เวฟเลตที่ถูกสเกลไปแล้วยังคงมีพลังงานเช่นเดียวกับมานธอร์เวฟเลต ดังนั้น เมซิสเวฟเลต (Basis wavelet) จะเป็นสมการดังนี้

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad , a > 0, \tau \in R$$

จากสมบัติการเปลี่ยนความละเอียดได้นี้เอง ทำให้เราสามารถใช้อเวฟเลตในการวิเคราะห์สัญญาณแบบหลายละเอียด (Multi-Resolution) โดยใช้อะการแปลงเวฟเลต ซึ่งถ้ามองเวฟเลต

ในลักษณะของตัวกรอง (Filter) ความละเอียดจะเพิ่มขึ้นเป็นสัดส่วนกับความถี่กึ่งกลางของช่วงที่ กำลังพิจารณาอยู่ นั่นคือ

$$\frac{\Delta f}{f} = c$$

ค่า c คือ ค่าคงที่ที่มักใช้ในการประมาณผลสัญญาณซึ่งเรียกว่า Constant-Q

เมื่อเขียนสมการการแปลงเวฟเลตแบบต่อเนื่อง (Continuous Wavelet Transform หรือ CWT) จะได้ดังสมการต่อไปนี้

$$CWT(\tau, a) = \frac{1}{\sqrt{a}} \int x(t) \psi^* \left(\frac{t-\tau}{a} \right) dt \quad , a = \frac{f_0}{f}$$

หรือหากเขียนสมการการแปลงแบบต่อเนื่องให้มีรูปแบบของเบซิซเวฟเลต

$$CWT(\tau, a) = \int x(t) \psi_{a,\tau}^*(t) dt$$

เราสามารถสร้างสัญญาณกลับมาใหม่โดยสมการต่อไปนี้

$$x(t) = c \int CWT(\tau, a) \psi_{a,\tau}(t) \frac{da}{a^2} dt$$

สำหรับการแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform หรือ DWT) จะยึดแนวคิดเบื้องต้นของการแปลงเวฟเลต คือ การพยายามแทนที่ฟังก์ชัน f ด้วยเวฟเลตหลายๆ เวฟเลต โดยแต่ละเวฟเลตจะแยกพิจารณาฟังก์ชันที่ความละเอียดต่างๆ กัน โดยเขียนได้เป็น

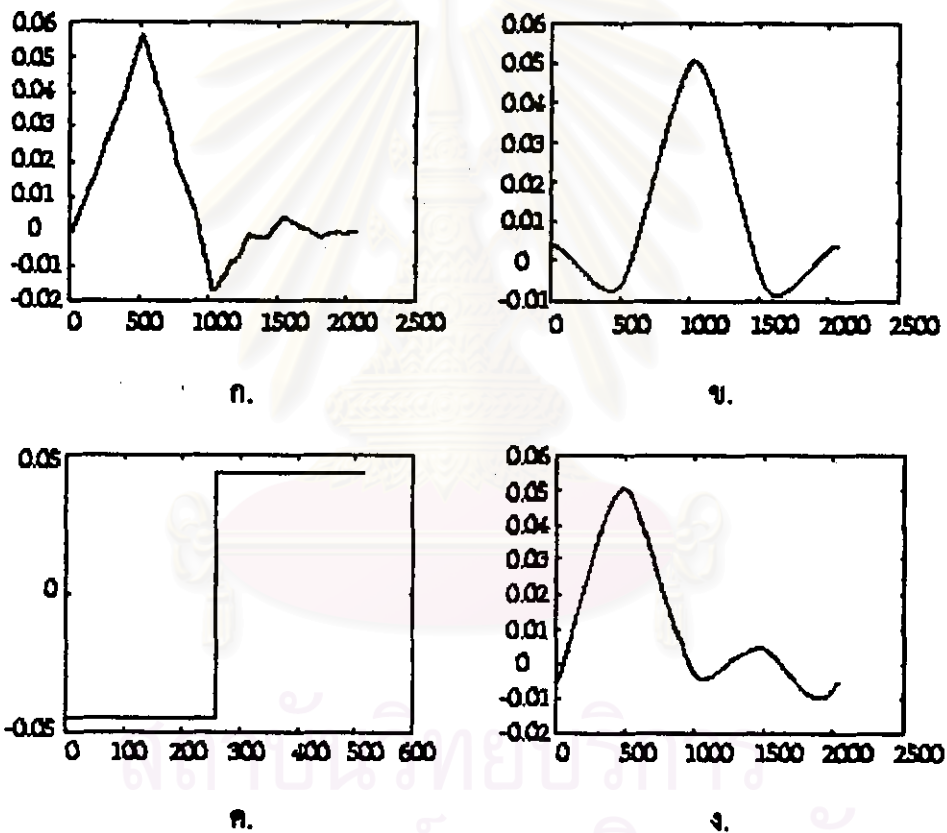
$$f = \sum C_{m,n}(f) \psi_{m,n}$$

โดยที่ $C_{m,n}$ คือ สัมประสิทธิ์

$\psi_{m,n}$ คือ เบซิซฟังก์ชัน

การแปลงเวฟเลตแบบไม่ต่อเนื่องมี 2 ประเภทหลักๆ คือ การเข้ารหัสคลื่นย่อย (Subband coding) มักใช้ในการบีบอัดเสียงหรือภาพ และการเข้ารหัสแบบพีรามิดหรือการวิเคราะห์สัญญาณหลายความละเอียด (Pyramidal coding หรือ Multiresolution signal analysis) มักใช้ในงาน Computer vision คือ การวิเคราะห์รายละเอียดของภาพจากขอบ ไปละเอียด ซึ่งถูกเรียกว่า "Coarse-

to-fine processing strategy" เช่น ภาพทะเล ขึ้นเป็นภาพหยาบก็จะอธิบายของโครงสร้างภาพทั้งหมด ในขณะที่ภาพทะเลอย่างละเอียดจะสามารถระบุทุกๆ สิ่งในภาพได้ เราจะประมวลผลจำนวนของ รายละเอียดน้อยกว่าที่จำเป็นต้องกระทำในงานการรู้จำ (recognition) เราสามารถจำวัตถุได้จากการ อธิบายของภาพหยาบโดยไม่จำเป็นต้องวิเคราะห์ที่ภาพละเอียด ดังนั้น รายละเอียดของภาพที่หยาบ ก็เพียงพอแล้ว วิธีนี้มีประโยชน์ในอัลกอริทึมการรู้จำแบบ (pattern recognition algorithms) วิธีนี้ถูก ใช้ศึกษากับงานประมวลผลภาพที่ระดับความละเอียดต่ำ (low-level image processing) เช่น stereo matching และ template matching เป็นต้น จากรูปที่ 2.31 จะแสดงตัวอย่างของเวฟเลตชนิดต่างๆ



รูปที่ 2.31 แสดงตัวอย่างของเวฟเลตชนิดต่างๆ รูป ก. Daubechies รูป ข. Coiflet

รูป ค. Haar รูป ง. Symmlet (Grap, 1995)

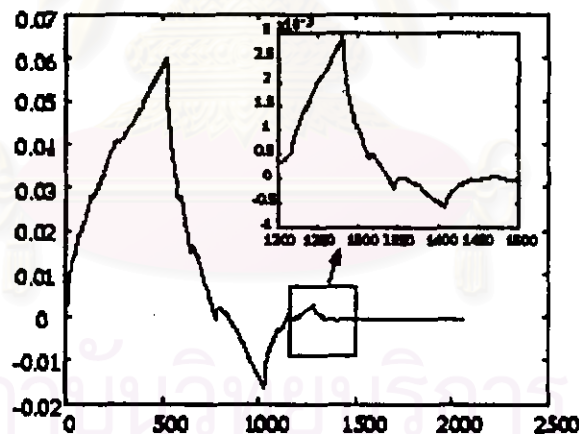
สำหรับการวิเคราะห์สัญญาณหลายความละเอียด จะประกอบด้วยฟังก์ชัน 2 ฟังก์ชัน คือ มานอร์เวฟเลต $\psi(t)$ และฟังก์ชันสเกลลิง (Scaling function) $\phi(t)$ หรือเรียกว่าฟาเธอร์เวฟเลต (Father wavelet) โคช

$$\psi(t) = \sum_{k=0}^{M-1} (-1)^k c_{M-1-k} \phi(2t - k)$$

สมการดังกล่าวเป็นสมการมาซอร์เวฟเลตและในรูปแบบของการกรองคิจิตต์ สมการนี้เท่ากับตัวกรองสัญญาณแบบไฮพาส (High-pass filter) และเพื่อการกระจายข้อมูลที่ตำแหน่งความถี่ที่แตกต่างกันมาซอร์เวฟเลตจะใช้สมการมาซอร์เวฟเลต ซึ่งมีลักษณะเป็นสมการเรียกซ้ำดังนี้

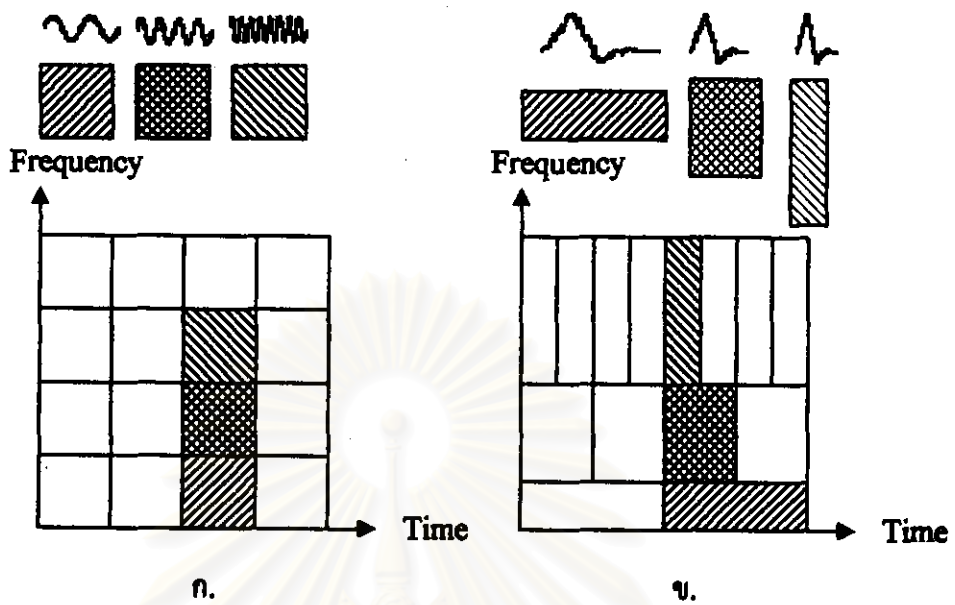
$$\phi(t) = \sum_{k=0}^{M-1} c_k \phi(2t - k)$$

$\phi(t)$ จะเป็นฟังก์ชันสเกลลิ่งสำหรับมาซอร์เวฟเลต $\psi(t)$ และ c_k จะเป็นสัมประสิทธิ์เวฟเลต ในรูปแบบของการกรองคิจิตต์ สมการนี้เท่ากับตัวกรองสัญญาณแบบโลว์พาส (Low-pass filter) รูปที่ 2.32 แสดง Daubechies mother wavelet ซึ่งมีการกระจายข้อมูลที่มีความถี่แตกต่างกัน รูปเล็กที่อยู่ภายในรูปใหญ่เกิดจากการขยายรูปเข้าไปในบริเวณ 1200 ถึง 1500



รูปที่ 2.32 แสดง Daubechies mother wavelet (Grap, 1995)

ด้วยคุณสมบัติเฉพาะในการวิเคราะห์เวฟเลต คือ ความละเอียดในการวิเคราะห์สัญญาณในเชิงของเวลา-ความถี่ เมื่อเปรียบเทียบกับการแปลงฟูเรียร์ จากรูปที่ 2.33 รูป ก. แสดงความละเอียดที่เกิดจากการแปลงฟูเรียร์ รูป ข. แสดงความละเอียดที่เกิดจากการแปลงเวฟเลตที่แตกต่างกันในการวิเคราะห์สัญญาณในเชิงเวลา-ความถี่ ดังนั้นการแปลงเวฟเลตจึงมีประสิทธิภาพสำหรับการวิเคราะห์สัญญาณที่ไม่ต่อเนื่องกันหรือสัญญาณไม่คงตัว



รูปที่ 2.33 แสดงความละเอียดของการแปลงทั้ง 2 แบบ ในการวิเคราะห์สัญญาณในเชิงเวลา-ความถี่

รูป ก. เมซิทฟังก์ชันแบบฟูรีเยร์ (Fourier basis function)

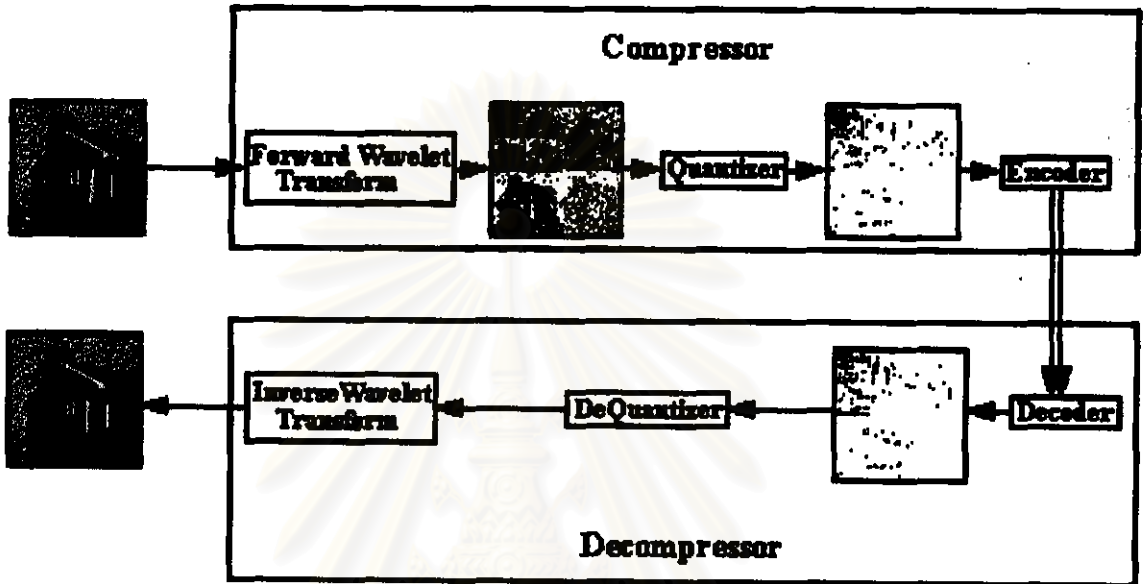
รูป ข. เมซิทฟังก์ชันแบบเวฟเลต (Wavelet basis function) (Grap, 1995)

2.2.4.2.3 การบีบข้อมูลภาพด้วยการแปลงเวฟเลต (Hilton, Jawerth และ Sengupta, 1994) (Sweldens, 1995) (Umbaugh, 1998: 125-130) (Shapiro, 1993)

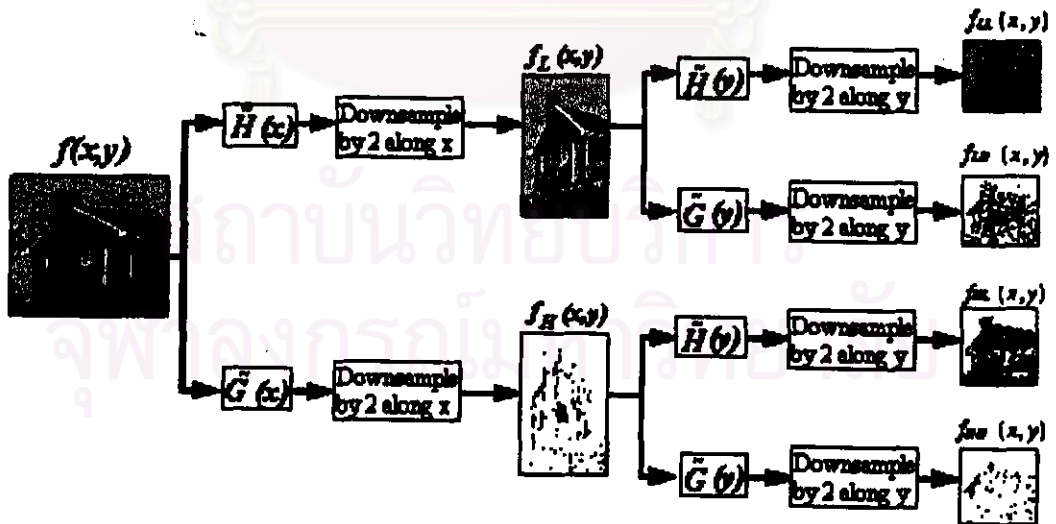
ขั้นตอนการบีบข้อมูลภาพด้วยการแปลงเวฟเลตแสดงในรูปที่ 2.34 ประกอบไปด้วยขั้นตอนการบีบภาพ 3 ขั้นตอน คือ Forward wavelet transform, Quantization และ Encoding และขั้นตอนการคลายภาพจะกระทำย้อนกลับจากการบีบภาพ

การบีบภาพด้วยเวฟเลตจะใช้การแปลงเวฟเลตในขั้นตอนการแปลง โดยทั่วไป จะใช้การแปลงเวฟเลตแบบไม่ต่อเนื่องหรือ DWT การแปลงจะแยกส่วนประกอบของภาพโดยเปลี่ยนข้อมูลภาพให้อยู่ในรูปข้อมูลความถี่ โดยการผ่านตัวกรองสัญญาณแบบไฮพาส (G) และตัวกรองสัญญาณแบบโลว์พาส (H) ส่วนที่ผ่านตัวกรองสัญญาณแบบไฮพาส คือ ข้อมูลบริเวณความถี่สูงหรือรายละเอียดของข้อมูลที่ความละเอียด (Resolution) นั้นๆ และส่วนที่ผ่านตัวกรองสัญญาณแบบโลว์พาส คือ ข้อมูลบริเวณความถี่ต่ำหรือข้อมูลหยาบในระดับความละเอียดนั้น ตัวกรองสัญญาณจะต้องเป็น "Perfect reconstruction filters" ซึ่งหมายความว่า ความผิดเพี้ยนใดๆ ที่เกิดจากการแปลงฟอว์เวิร์ด (forward) จะถูกยกเลิกด้วยการแปลงอินเวอร์ส (inverse) เช่น การแปลงแบบ

Quadrature Mirror Filters (QMF) ให้ (\tilde{H}, \tilde{G}) และ (H, G) เป็นการแปลงฟอว์เวิร์ดและอินเวอร์ตตามลำดับ



รูปที่ 2.34 แสดงการบีบภาพด้วยการแปลงเวฟเลต (Sweidens, 1995)



รูปที่ 2.35 แสดงขั้นตอนการบีบภาพด้วยการแปลงฟอว์เวิร์ดเวฟเลต (Sweidens, 1995)

ในการแปลงเวฟเลต เราจะใช้วิธีการแยกส่วนประกอบภาพในลักษณะเชิงลำดับชั้น (Hierarchical) ที่จะแยกส่วนประกอบภาพแบบเรียกซ้ำ (Recursive) ได้ดังนี้

จากรูปที่ 2.35 แสดงการแปลงฟอว์วอร์ดเวฟเลต (Forward wavelet transform) ของภาพ ภาพ $f(x,y)$ เริ่มต้นจากการกรองสัญญาณทางแกน x ผลลัพธ์ที่ได้เป็นภาพ Lowpass image $fL(x,y)$ และ Highpass image $fH(x,y)$ ซึ่งจะทำให้ $fL(x,y)$ และ $fH(x,y)$ ทางแกน x ถูกแบ่งครึ่งแบบวนวิซของสัญญาณ และทำการวนเปิดโดยไม่มีการสูญเสียข้อมูล จากนั้นทั้ง $fL(x,y)$ และ $fH(x,y)$ จะถูกกรองสัญญาณทางแกน y ผลลัพธ์ได้ 4 ภาพย่อย (Subimages) ประกอบด้วย $fLL(x,y)$, $fLH(x,y)$, $fHL(x,y)$ และ $fHH(x,y)$ และทำการวนเปิดภาพย่อยอีกครั้ง ซึ่งเราจะแยกส่วนประกอบภาพเป็น $fLL(x,y)$ (Average signal หรือ Lowpass-Lowpass) อีก 3 ภาพย่อย คือ Detail signal ได้แก่ $fLH(x,y)$ (Horizontal feature หรือ Lowpass-Highpass), $fHL(x,y)$ (Vertical feature หรือ Highpass-Lowpass) และ $fHH(x,y)$ (Diagonal feature หรือ Highpass-Highpass)

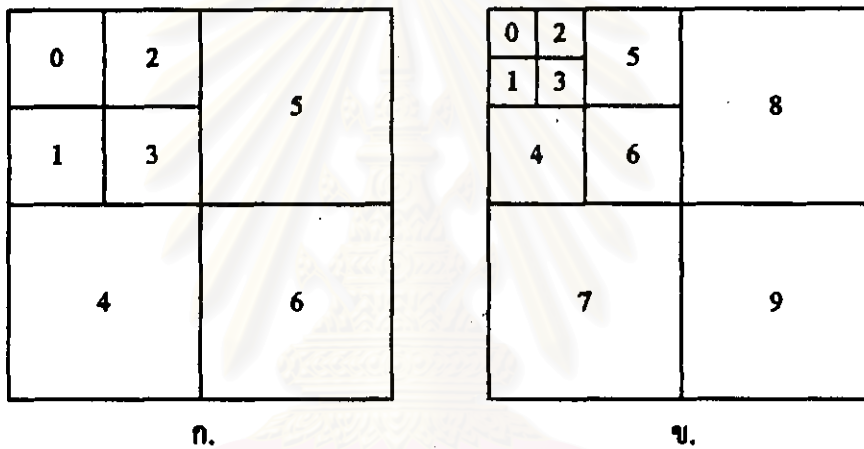
รูปแบบการแสดงผลการแปลงเวฟเลต แสดงดังรูปที่ 2.36 เราสามารถแยกส่วนประกอบภาพหรือทำการแปลงเรียกซ้ำ (Recursive transform) ที่ Average signal หรือ $fLL(x,y)$ (Lowpass-Lowpass) ได้ ซึ่งจะให้ผลลัพธ์จำนวนภาพย่อยเท่ากับ $3N+1$ เช่น 4, 7, 10, 13, 16 เป็นต้น วิธีนี้เรียกว่าการแยกส่วนประกอบความละเอียดหลายระดับ (Multiresolution decomposition) แสดงดังรูปที่ 2.37 จำนวนการแปลงขึ้นอยู่กับปัจจัยต่างๆ เช่น อัตราส่วนการบีบ, ขนาดของภาพต้นฉบับ, ความยาวของตัวกรอง QMF เป็นต้น โดยทั่วไป ถ้าต้องการอัตราส่วนการบีบสูง ก็จะต้องทำการแปลงมากขึ้นด้วย เราสามารถเห็นผลลัพธ์ของภาพหลังการแปลงเป็นแบบสเตเชียลหรือเห็นเป็นภาพได้ เมื่อเปรียบเทียบกับภาพแปลงชนิดอื่นๆ เพราะข้อมูลหลังการแปลงชนิดอื่นๆ จะไม่มีความสัมพันธ์กันกับภาพต้นฉบับเลย

เวฟเลตที่นิยมใช้ในงานบีบอัดภาพ คือ

- Orthogonal Daubechies Wavelet ที่มี 3 vanishing moment มี filter of length 6
- Biorthogonal wavelet ที่มี 2 vanishing moment มี filter with length 3 และ 5

Low/ Low	Low/ High
High/ Low	High/ High

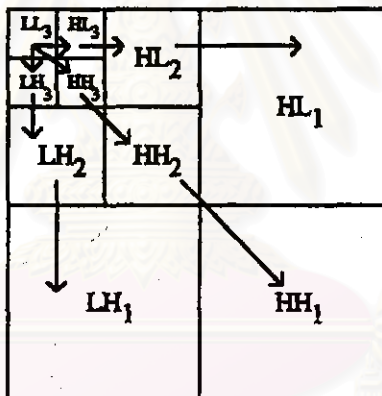
รูปที่ 2.36 แสดงตำแหน่งของคลื่นความถี่ในภาพหลังการแปลงด้วยเวฟเลต (Umbaugh, 1998: 128)



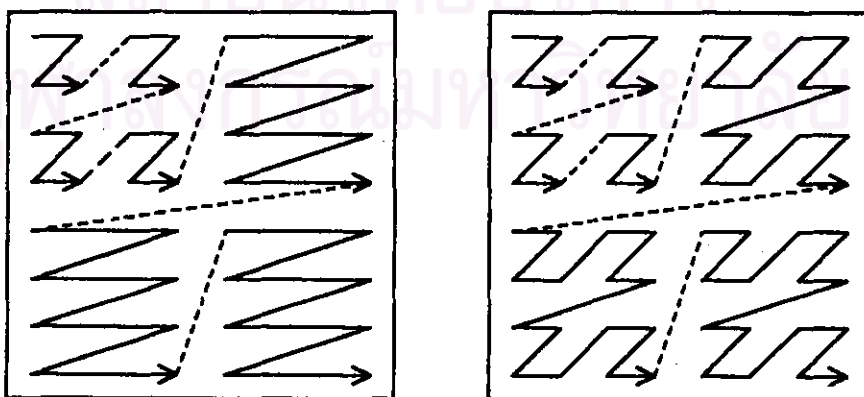
รูปที่ 2.37 แสดงตัวอย่างการแยกส่วนประกอบความละเอียดหลายระดับที่ให้ผลลัพธ์เป็นจำนวนภาพย่อยต่างๆ กัน รูป ก. แสดงจำนวน 7 ภาพย่อย รูป ข. แสดงจำนวน 10 ภาพย่อย (Umbaugh, 1998: 128)

หลังจากการทำการแปลงเวฟเลตแล้ว จะได้เมทริกซ์ของสัมประสิทธิ์ที่ประกอบด้วยสัญญาณค่าเฉลี่ย (Average signal) และสัญญาณรายละเอียด (Detail signal) จากนั้นทำการควอนไทซ์ผลลัพธ์จากการแปลงเพื่อลดจำนวนข้อมูลสัมประสิทธิ์การแปลงลงและเข้ารหัสค่าที่ถูกควอนไทซ์ โดยความเป็นจริงแล้วการแปลงด้วยเวฟเลตจะทำให้ค่าสัมประสิทธิ์มีขนาดใหญ่ขึ้น ซึ่งเป็นการเพิ่มขนาดที่ต้องการใช้ในการเก็บภาพ การบีบไม่เกิดขึ้นในขั้นตอนการแปลง เพียงแต่เป็นการเปลี่ยนรูปแบบข้อมูลภาพที่จะเก็บให้เหมาะสมกับการบีบ ซึ่งการบีบจะเกิดขึ้นจากควอนไทซ์และการเข้ารหัสสัมประสิทธิ์การแปลง

สำหรับการเข้ารหัสสัมประสิทธิ์ จะทำการกราดตรวจผ่าน Parent ก่อน Child ดังรูปที่ 2.38 ซึ่งแสดงความสัมพันธ์ระหว่างกันของคลื่นย่อยแบบ Parent-child ถูกตรวในรูปแบบแสดงความสัมพันธ์จาก Parent ไป Child สัมประสิทธิ์ที่หยากกว่าหรือมีความถี่ต่ำจะเรียกว่า Parent และสัมประสิทธิ์ที่ละเอียดกว่าหรือมีความถี่สูงจะเรียกว่า Child และในทุกๆ Parent จะมี 3 Child การกราดตรวจจะกระทำจาก Parent ไป Child คือ LL LH HL และ HH ดังนั้นการกราดตรวจของสัมประสิทธิ์เวฟเลตที่ N-scale pyramid จะเริ่มต้นที่ความถี่ต่ำสุดคือ LL_N LH_N HL_N และ HH_N ไปจนถึงระดับที่ $N-1$ แต่อย่างไรก็ตาม ในการเรียงลำดับสัมประสิทธิ์เพื่อการเข้ารหัสมีหลายวิธี ดังเช่นรูปที่ 2.39 แสดงตัวอย่างของการเรียงลำดับสัมประสิทธิ์ในการกราดตรวจของการแปลงเวฟเลตที่ 2 ระดับ รูป ก. แบบ Raster scan รูป ข. แบบ Morton order รูป ค. แบบ Peano scan ถูกตรวแบบจุดในรูปแบบแสดงถึงการเชื่อมกันระหว่างคลื่นย่อย สังเกตได้ว่า จะเป็นการกราดตรวจจาก LL , HL , LH และ HH

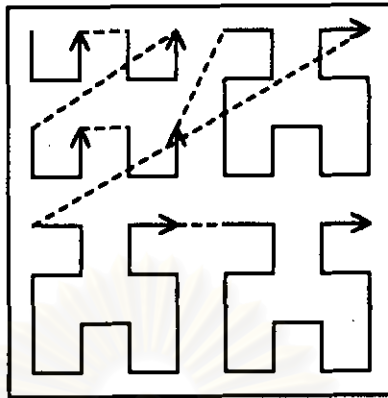


รูปที่ 2.38 แสดงความสัมพันธ์ระหว่างกันของคลื่นย่อยแบบ Parent-child (Shapiro, 1993)



ก.

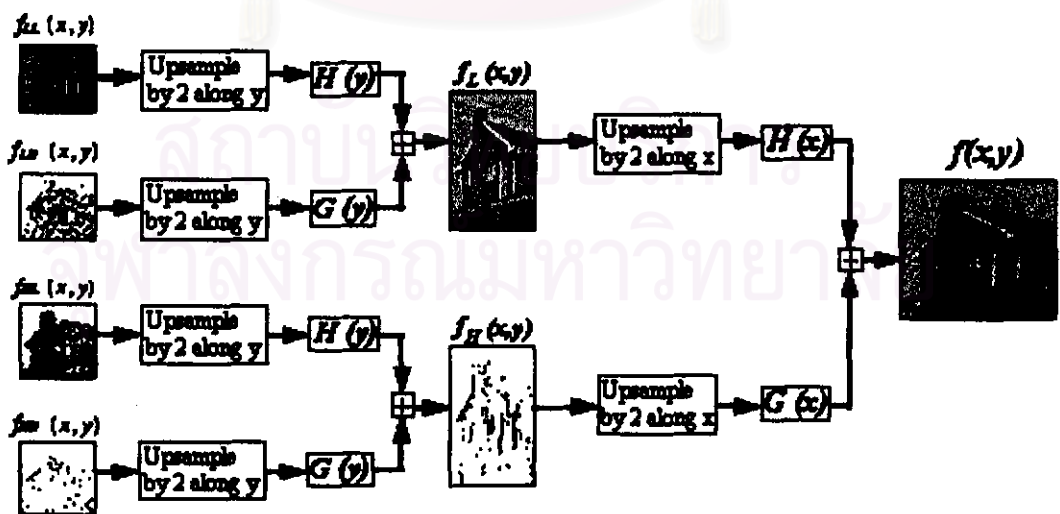
ข.



ก.

รูปที่ 2.39 แสดงตัวอย่างของการเรียงลำดับสัมประสิทธิ์ในการกราดตรวจของการแปลงเวฟเลต 2 ระดับ รูป ก. แบบ Raster scan รูป ข. แบบ Morton order รูป ค. แบบ Peano scan

จากรูปที่ 2.40 แสดงการแปลงอินเวอร์ตเวฟเลต (Inverse wavelet transform) ของภาพ โดยนำสัญญาณค่าเฉลี่ยและสัญญาณละเอียดมาทำอัทแซมเปิดทางแกน y ด้วยตัวกรองสัญญาณแบบอินเวอร์ตโลว์พาส (Lowpass inverse filter) และตัวกรองสัญญาณแบบอินเวอร์ตไฮพาส (Highpass inverse filter) ทั้ง 4 ภาพ การทำอัทแซมเปิดจะกระทำโดยการใส่ค่า 0 ระหว่างคู่ของค่าสัญญาณในแกน y หลังจากนั้น ก็จะทำอัทแซมเปิดทางแกน x และสร้างภาพกลับมาใหม่



รูปที่ 2.40 แสดงขั้นตอนการคลายภาพด้วยการแปลงอินเวอร์ตเวฟเลต (Sweldens, 1995)

การใช้การแปลงเวฟเลตกำลังเป็นที่นิยมอย่างเพิ่มขึ้นสำหรับการบีบภาพในงานวิจัยในปัจจุบัน ความก้าวหน้าของคอมพิวเตอร์ในด้านต่างๆ เช่น ข่าวดารบนซูเปอร์ไฮเวย์ (information superhighway), การประยุกต์สื่อหลายแบบ (multimedia application), และที่วิศวกรรมระดับสูง ทั้งหมดที่กล่าวมาทำให้เกิดความสนใจอย่างมากในการบีบภาพ คุณสมบัติการแยกส่วนประกอบความละเอียดหลายระดับ (Multiresolution decomposition) ของการแปลงเวฟเลต ซึ่งแยกข้อมูลความละเอียดต่ำ (Low-resolution) จากข้อมูลที่ละเอียด ซึ่งเป็นประโยชน์ในการประยุกต์ที่มีความต้องการข่าวดารหลายแบบแต่มีความเร็วสูง เช่น การพิจารณาภาพในฐานะข้อมูล หรือภาพที่มีการส่งผ่านแบบโทรเลขชีพบนเวดจ์ไวด์เว็บ การแปลงเวฟเลตเป็นการแปลงชนิดใหม่ที่มีการใช้กับโปรแกรมสำเร็จรูปประยุกต์การบีบภาพ เช่น การบีบภาพทางการแพทย์ นอกจากนี้ ยังมีการใช้ในการลดสัญญาณรบกวน (noise reduction), การวิเคราะห์การรับรู้ของมนุษย์ (analysis of human perception), เรดาร์ (radar), การทำนายแผ่นดินไหว (earthquake prediction), ธรณีวิทยาที่เกี่ยวข้องกับการเกิดแผ่นดินไหว (seismic geology) เป็นต้น

เมื่อเปรียบเทียบภาพที่บีบด้วยเทคนิคเจทีเคและเวฟเลต จะแสดงดังรูปที่ 2.41 รูป ก. ภาพต้นฉบับ รูป ข. ภาพที่บีบด้วยเทคนิคเจทีเคที่อัตราส่วนการบีบ 36: 1 รูป ค. ภาพที่บีบด้วยเทคนิคเวฟเลตที่อัตราส่วนการบีบ 150: 1 และรูป ง. บีบด้วยเทคนิคเวฟเลตที่อัตราส่วนการบีบ 180: 1

2.3 เกณฑ์ในการเปรียบเทียบคุณภาพของภาพ (Fidelity criteria)

สำหรับการบีบข้อมูลภาพแบบสูญเสียข้อมูลบางส่วน จะทำให้มีข่าวดารของภาพบางส่วนสูญหาย หรือผิดเพี้ยนไปจากต้นฉบับ จึงต้องมีเกณฑ์ในการวัดคุณภาพของภาพหลังการบีบแล้ว เพราะถ้าต้องการภาพที่มีอัตราส่วนการบีบสูง คุณภาพของภาพก็จะเลวลง แต่ถ้านำคุณภาพของภาพที่ดี อัตราส่วนการบีบของภาพก็ต่ำลง

เกณฑ์ที่นิยมใช้แบ่งได้เป็น 2 ประเภท คือ แบบเทียบกับภาพต้นฉบับ (Subjective fidelity criteria) และแบบเทียบกับภาพบีบ (Objective fidelity criteria) (Umbaugh, 1998: 237-242) (ควงอาทิตย์ ศรีมุต และ องอุทร เอกอริชทรัพย์, 2539: 14-16)



ก.



ข.



ก.



ง.

รูปที่ 2.41 แสดงตัวอย่างการบีบภาพด้วยเทคนิคเงาเพิกและเวฟเลต รูป ก. ภาพต้นฉบับ Lena

รูป ข. รูปที่บีบด้วยเงาเพิกที่อัตราส่วนการบีบที่ 36: 1

รูป ค. รูปที่บีบด้วยเวฟเลตที่อัตราส่วนการบีบที่ 150: 1

รูป ง. รูปที่บีบด้วยเวฟเลตที่อัตราส่วนการบีบที่ 180: 1

2.3.1 แบบเทียบกับภาพต้นฉบับ (Umbaugh, 1998: 237-242) จะใช้ความรู้สึกจากการรับรู้ภาพเป็นเครื่องมือในการตัดสินคุณภาพและความถูกต้องของภาพ โดยการเปรียบเทียบความรู้สึกของมนุษย์จากการรับรู้ภาพระหว่างภาพต้นฉบับและภาพที่ได้จากการบีบและคลายกลับมา ซึ่งคือวิธีคะแนนค่าเฉลี่ยความคิดเห็น (Mean opinion score หรือ MOS) โดยมีค่าระดับความพอใจภาพหลายระดับ อาจแบ่งเป็นระดับดังนี้คือ {1, 2, 3, 4, 5} และแต่ละค่าแทนความหมายว่า {แย่มาก, แย่, ปานกลาง, ดี, ดีมาก} ตามลำดับ โดยปกติวิธีนี้จะเป็นวิธีที่ดีมาก แต่การทดสอบด้วยวิธีนี้จะทำได้ยากและใช้เวลานานมาก ผลลัพธ์ที่ได้ในแต่ละครั้งก็ไม่แน่นอนขึ้นอยู่กับเงื่อนไขการทดสอบ เช่น เงื่อนไขการรับรู้ภาพ, อุปกรณ์ที่ใช้แสดงภาพ เป็นต้น อีกทั้งยังเป็นวิธีที่ใช้ในขั้นตอนการออกแบบระบบเท่านั้น ไม่สามารถนำไปใช้งานในทันทีที่ต้องการ

Impairment	Quality	Comparison
5-Imperceptible	A-Excellent	+2 much better
4-Perceptible, not annoying	B-Good	+1 better
3-Somewhat annoying	C-Fair	0 the same
2-Severely annoying	D-Poor	-1 worse
1-Unusable	E-Bad	-2 much worse

ตารางที่ 2.21 แสดงการให้คะแนนภาพตามหลักสากลที่นิยมใช้กัน (Umbaugh, 1998: 242)

แบบเทียบกับภาพต้นฉบับนี้แบ่งได้เป็น 3 ประเภทหลักๆ คือ แบบ Impairment tests เปรียบเทียบว่าภาพที่บีบนั้นคุณภาพแย่อย่างไร แบบ Quality tests เปรียบเทียบว่าภาพที่บีบนั้นคุณภาพคืออย่างไร แบบ Comparison tests เปรียบเทียบภาพที่บีบมีความเหมือนกับภาพต้นฉบับอย่างไร จากตารางที่ 2.21 แสดงตัวอย่างการให้คะแนนภาพตามหลักสากลทั้ง 3 แบบที่นิยมใช้กัน

2.3.2 แบบเทียบกับภาพบีบ (Umbaugh, 1998: 237-242) (Gonzalez และ Woods, 1992: 318-320) ด้วยเหตุที่วิธีแบบเทียบกับภาพต้นฉบับมีข้อเสียดังที่กล่าว แบบเทียบกับภาพบีบก็เกิดขึ้นเพื่อแก้ไขปัญหาที่กล่าวมา และเพื่อให้คอมพิวเตอร์สามารถจำลองค่าคุณภาพของภาพได้ วิธีในแบบเทียบกับภาพบีบ ได้แก่ ความผิดพลาดรูทมีนสแควร์ (Root-mean-square error หรือ e_{RMS}), อัตราซิกเนลทูนอยส์รูทมีนสแควร์ (Root-mean-square signal-to-noise ratio หรือ SNR_{RMS}), อัตราพีคซิกเนลทูนอยส์ (Peak-signal-to-noise ratio หรือ SNR_{PEAK}), อัตราส่วนการบีบ (Compression ratio)

และวิธีอื่นๆ เป็นต้น วิธีเหล่านี้จะเป็นตัวชี้ความผิดพลาดของภาพ ด้วยการวัดคุณภาพของภาพโดยการหาค่าความสูญเสียขั้วดาวของภาพระหว่างภาพต้นฉบับและภาพที่ได้จากการบีบและคลายกลับมา และจะวัดได้ดีในกรณีที่ภาพเกิดความผิดพลาดโดยสุ่ม (random error) แต่จะไม่ดีถ้ารับความผิดพลาดโครงสร้าง (structure error) อย่างไรก็ตาม วิธีแบบเดียวกับภาพบีบนี้ก็ยังไม่ใช่วิธีที่ดีที่สุด เพราะค่าที่ได้นี้ไม่ได้เกิดจากการรับรู้ของมนุษย์ แต่ในความเป็นจริงมนุษย์เป็นผู้ใช้ภาพ

2.3.2.1 ความผิดพลาดรูทมีนสแควร์ จำกัดให้ $f(x,y)$ แทนภาพข้อมูลเข้าหรือภาพต้นฉบับ และ $\hat{f}(x,y)$ แทนภาพต้นฉบับที่ผ่านการบีบและคลายกลับมา สำหรับ x และ y ใดๆ แล้ว ค่าความผิดพลาด $e(x,y)$ ระหว่าง $f(x,y)$ และ $\hat{f}(x,y)$ สามารถหาได้ดังนี้

$$e(x,y) = \hat{f}(x,y) - f(x,y)$$

ดังนั้นค่าความผิดพลาดของภาพต้นฉบับขนาด $M \times N$ จะได้

$$\text{ค่าความผิดพลาดทั้งหมด} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]$$

และค่าความผิดพลาดรูทมีนสแควร์ระหว่าง $f(x,y)$ และ $\hat{f}(x,y)$ จะหาได้จากสมการต่อไปนี้

$$e_{RMS} = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2}$$

ถ้าค่า e_{RMS} มีค่าน้อยแสดงให้เห็นว่าภาพที่บีบและคลายกลับมามีคุณภาพดี

2.3.2.2 อัตราจิกเนตของรูทมีนสแควร์ ในบางกรณี เกณฑ์วัดความถูกต้องอาจจะแสดงด้วยอัตราจิกเนตของรูทมีนสแควร์ของภาพที่คลายกลับมา ถ้าพิจารณาภาพที่บีบและคลายกลับมา $\hat{f}(x,y)$ เป็น "signal" และค่าความผิดพลาด $e(x,y)$ เป็น "noise" เราจะหาค่าอัตราจิกเนตของรูทมีนสแควร์นี้ได้จากสมการต่อไปนี้

$$SNR_{RMS} = \sqrt{\frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2}}$$

ถ้าค่า SNR_{RMS} มีค่ามากกว่าแสดงให้เห็นว่าภาพมีคุณภาพดี

2.3.2.3 อัตราผิดซิกเนตยูน้อย สามารถหาได้จากสมการต่อไปนี้

$$SNR_{PRK} = 10 \log_{10} \frac{(L-1)^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2}$$

โดยที่ L คือ จำนวนระดับความเข้มแสง เช่น ภาพ 8 บิต ค่า $L = 256$

2.3.2.4 อัตราส่วนการบีบ (Compression ratio) คือ ความแตกต่างของขนาดแฟ้มข้อมูลต้นฉบับและแฟ้มข้อมูลที่ได้จากการบีบ ซึ่งการคำนวณค่าดังกล่าวมีอยู่ด้วยกันหลายวิธี วิธีหนึ่งที่ถูกใช้อย่างแพร่หลาย คือ การบอกค่าเป็นอัตราส่วนระหว่างข้อมูลเข้าและข้อมูลออก เช่น “4:1” ส่วนอีกวิธีหนึ่งที่ใช้กันมาก คือ การบอกเป็นเปอร์เซ็นต์ ซึ่งมีค่าตั้งแต่ 0 ถึง 100 โดยปกติจะไม่มี 0% ซึ่งจะหมายถึง ไม่มีการบีบข้อมูลภาพหรือเท่ากับเก็บภาพมีขนาดเท่ากับภาพต้นฉบับ ถ้า 100% หมายถึง ขนาดข้อมูลที่ใช้เก็บภาพเท่ากับ 0 ไบต์ ซึ่งเป็นไปไม่ได้ในความเป็นจริง

สมการอัตราส่วนการบีบมีดังต่อไปนี้ สมการแรกจะได้ค่าเป็นอัตราส่วน ส่วนสมการที่สองจะได้ค่าเป็นเปอร์เซ็นต์

$$\text{อัตราส่วนการบีบ} = \frac{\text{ขนาดข้อมูล ต้นฉบับ}}{\text{ขนาดข้อมูล ที่ผ่านการ บีบ}}$$

หรือ

$$\text{อัตราส่วนการบีบ} = \frac{(\text{ขนาดข้อมูลต้นฉบับ} - \text{ขนาดข้อมูลที่ผ่านการบีบ})}{\text{ขนาดข้อมูลต้นฉบับ}} \times 100$$

เนื่องจากทั้งวิธีแบบเทียบกับภาพบิบบและแบบเทียบกับภาพต้นฉบับส่วนแต่ไม่ใช่วิธีที่ดีที่สุด อีกทั้งการนำทั้งสองวิธีมาประยุกต์รวมกันก็เป็นสิ่งที่นักวิจัยทั้งหลายได้พยายามหลีกเลี่ยงเนื่องจากความสลับซับซ้อนในการแก้ปัญหา แต่ในเวลาต่อมาได้มีผู้เสนอวิธีที่สามารถนำทั้งสองวิธีมาประยุกต์รวมกันได้ คือ D.J. Sakrison ซึ่งเป็นผู้เสนอวิธีมาตราส่วนคุณภาพภาพ (Picture Quality Scale หรือ PQS) เป็นครั้งแรก

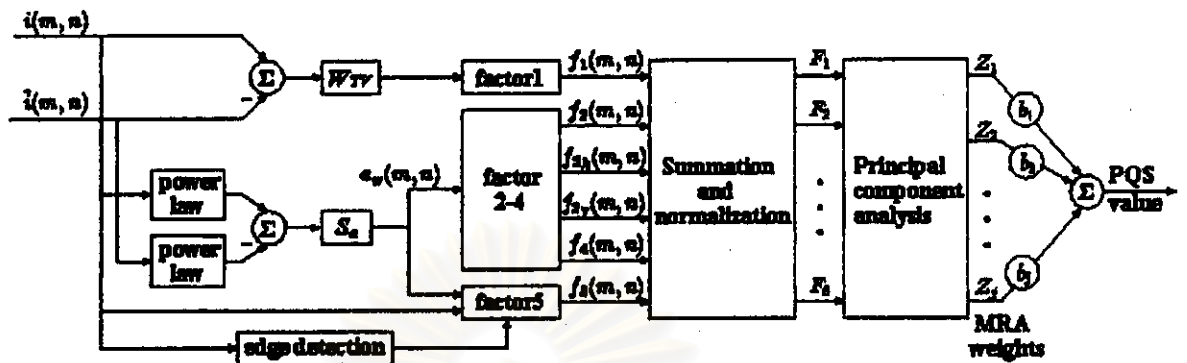
2.3.3 มาตราส่วนคุณภาพภาพ (Miyahara, Kotani และ Algazi, 1996) (Avahanam และ Algazi, 1996) (Gerfeldler และ Muller, 1996) เป็นระเบียบวิธีใหม่ที่น่ามาใช้ในงานวิจัย เกิดจากการรวมกันของแบบเทียบกับภาพต้นฉบับและแบบเทียบกับภาพบิบบเพื่อใช้ในการเปรียบเทียบภาพไม่มีสี (achromatic image) ในขอบเขตคุณภาพของภาพทั้งหมดซึ่งได้จำกัดความโดยวิธีคะแนนค่าเฉลี่ยความคิดเห็น สามารถแสดงความผิดเพี้ยนของภาพได้ทั้งความผิดเพี้ยนโดยสุ่ม และความผิดเพี้ยนโครงสร้าง การประเมินผลคุณภาพภาพด้วยวิธีนี้จะให้ค่าที่ใกล้เคียงกับค่าคะแนนเฉลี่ยความคิดเห็น โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) มากกว่า 0.92 วิธีค่ามาตราส่วนคุณภาพภาพนี้จะมีค่าอยู่ระหว่าง 1 ถึง 5

ขั้นตอนของการทำมาตราส่วนคุณภาพภาพ ประกอบ ไปด้วยขั้นตอนหลักๆ ดังนี้

1. การหาปัจจัยความผิดเพี้ยนของภาพ 5 ปัจจัย (Distortion factor)
2. การวิเคราะห์องค์ประกอบหลัก (Principal component analysis)
3. การวิเคราะห์ถดถอยพหุ (Multiple regression analysis)

จากรูปที่ 2.42 ซึ่งแสดงโครงสร้างของวิธีมาตราส่วนคุณภาพภาพ โดยให้ $I(m,n)$ เป็นภาพต้นฉบับ และ $i(m,n)$ เป็นภาพบิบบที่มีความผิดเพี้ยน เราจะคำนวณรูปแสดงความผิดเพี้ยน (local distortion maps) $\{f_i(m,n)\}$ จากปัจจัยความผิดเพี้ยน $\{F_i\}$ จากนั้นจะแยกความสัมพันธ์ของปัจจัยทั้ง 5 ด้วยวิธีวิเคราะห์องค์ประกอบหลัก หลังจากนั้นใช้วิธีการวิเคราะห์ถดถอยพหุเพื่อรวมผลลัพธ์จากการวิเคราะห์องค์ประกอบหลักลงในตัวเลขตัวหนึ่งซึ่งแทนค่าคุณภาพของภาพ

2.3.3.1 การหาปัจจัยความผิดเพี้ยนของภาพ 5 ปัจจัย (Distortion factor) ปัจจัยความผิดเพี้ยน 5 ปัจจัย เป็นฟังก์ชันของความแตกต่างระหว่างภาพต้นฉบับ และภาพที่บิบบและคลาสนด้วยปัจจัยความผิดเพี้ยนที่ 1 และ 2 แสดงการเกิดความผิดเพี้ยนแบบโดยสุ่ม (random error) ในภาพปัจจัยความผิดเพี้ยนที่ 3, 4, 5 แสดงการเกิดความผิดเพี้ยนแบบโครงสร้าง (structure error) ในภาพ



รูปที่ 2.42 แสดงโครงสร้างของวิธีมาตราส่วนคุณภาพภาพ (Miyahara, Kotani และ Algazi, 1996)

ปัจจัยความผิดเพี้ยนที่ 1 และ 2 จะแสดงถึงการเกิดสัญญาณรบกวน (noise) ในบริเวณต่างๆ ของภาพที่บีบและคลายแล้วเปรียบเทียบกับภาพต้นฉบับ โดยปัจจัยที่ 1 คือ ค่าแสดงความแตกต่างระหว่างภาพต้นฉบับและภาพบีบและคลายแล้ว ซึ่งได้รูปแบบจากมาตรฐานการวัดสัญญาณรบกวนในโทรทัศน์ของ CCIR (International Radio Consultative Committee) และปัจจัยที่ 2 ได้รูปแบบจากแบบจำลองของ HVS (Human Visual System)

ปัจจัยความผิดเพี้ยนที่ 1 ได้จากมาตรฐานการวัดสัญญาณรบกวนในโทรทัศน์ของ CCIR 567-1 นิยามได้ดังนี้

$$e_i(m,n) = i(m,n) - \hat{i}(m,n)$$

และคำนวณภาพแสดงความผิดเพี้ยนด้วยสมการ

$$f_1(m,n) = [e_i(m,n) * W_{TV}(m,n)]^2$$

โดยที่น้ำหนักความถี่ (frequency weighting) ซึ่งนิยามโดย CCIR 567-1 เป็นดังสมการ

$$W_{TV} = \frac{1}{1 + \left(\frac{f}{f_c}\right)^2}, f = \sqrt{u^2 + v^2}, f_c = 5.56 \text{ cycles/degree}$$

ดังนั้น ปัจจัยความผิดเพี้ยนที่ 1 คำนวณได้โดยสมการ

$$F_1 = \frac{\sum_{m,n} f_1(m,n)}{\sum_{m,n} i^2(m,n)} \quad \text{เมื่อภาพมีขนาดจุดภาพเท่ากับ } M \times N$$

จากปัจจัยความผิดเพี้ยนที่ 2 สามารถคำนวณภาพแสดงความผิดเพี้ยนด้วยสมการ

$$f_2(m,n) = I_T(m,n)[e_w(m,n) * S_a(m,n)]^2$$

โดยที่ F_2 จะไม่สนใจ $e_w(m,n)$ ที่มีค่าต่ำกว่าขีดแบ่ง (Threshold หรือ T) ที่กำหนดไว้ ซึ่งใช้ฟังก์ชัน $I_T(m,n)$ และให้ $T = 1$

$$e_w(m,n) = c(x,y) - \hat{c}(x,y)$$

$c(x,y)$ คือ การปรับคอนทราสต์ (Contrast) ของภาพต้นฉบับ

$\hat{c}(x,y)$ คือ การปรับคอนทราสต์ของภาพที่ถูกบีบ

$$S_a(u,v) = S(\omega)O(\omega,\theta)$$

$$S(\omega) = 1.5e^{\frac{-\sigma^2\omega^2}{2}} - e^{-2\sigma^2\omega^2}$$

$$\sigma = 2, \omega = \frac{2\pi f}{60}, f = \sqrt{u^2 + v^2}$$

ซึ่ง u และ v เป็น spatial frequency ตามแนวอนและแนวตั้งตามลำดับ

$$O(\omega,\theta) = \frac{1 + e^{\beta(\omega-\omega_0)} \cos^4 2\theta}{1 + e^{\beta(\omega-\omega_0)}}$$

$$\beta = 8, f_0 = 11.13 \text{ cycles/degree}$$

ดังนั้น ปัจจัยความผิดเพี้ยนที่ 2 คำนวณได้โดยสมการ

$$F_2 = \frac{\sum_{m,n} f_2(m,n)}{\sum_{m,n} \hat{i}^2(m,n)}$$

ปัจจัยความผิดเพี้ยนที่ 3 (End of block disturbances) แสดงถึงความผิดพลาดจากการเข้ารหัสภาพที่เกิดจากการแบ่งภาพเป็นบล็อก ซึ่งจะก่อให้เกิดความไม่ต่อเนื่องในระดับความเข้มแสงหรือมีลักษณะของสีหรือความถี่ไม่ต่อเนื่องกันในบริเวณจุดสิ้นสุดของบล็อก (end of block) ในภาพ ซึ่งเรียกว่า บล็อกกิงอาร์ทีแฟท การบีบภาพแบบการเข้ารหัสการแปลงจะทำให้ภาพเกิดความผิดพลาดชนิดนี้ เช่น วิธีการเข้ารหัสภาพของเจเพ็ก ซึ่งมีการแบ่งภาพเป็นบล็อก บล็อกละ 8×8 หรือ 16×16 เป็นต้น

เราสามารถนิยามปัจจัยความผิดเพี้ยนที่ 3 ว่าเป็นฟังก์ชันของภาพ 2 ภาพที่มีความผิดเพี้ยน ภาพหนึ่งคือ ภาพที่มีความผิดเพี้ยนตามแนวนอน (horizontal block error discontinuities) อีกภาพหนึ่งคือ ภาพที่มีความผิดเพี้ยนตามแนวตั้ง (vertical block error discontinuities) ได้ดังนี้

ภาพที่มีความผิดเพี้ยนตามแนวนอนคำนวณได้ดังสมการ

$$f_{3h}(m, n) = I_h(m, n) \Delta_h^2(m, n)$$

โดยที่ $\Delta_h(m, n) = e_w(m, n) - e_w(m, n+1)$ และ

$I_h(m, n)$ เป็นฟังก์ชันที่จะเลือกเฉพาะค่าความแตกต่างในขอบเขตบล็อกตามแนวนอน

ดังนั้น ปัจจัยความผิดเพี้ยนที่ 3 ตามแนวนอนคำนวณได้ดังสมการ

$$F_{3h} = \frac{1}{N_h} \sum_{m,n} f_{3h}(m, n)$$

โดยที่ $N_h = \sum_{m,n} I_h(m, n)$ เป็นจำนวนจุดภาพที่ถูกเลือกโดยฟังก์ชันดังกล่าว

ภาพที่มีความผิดเพี้ยนตามแนวตั้งคำนวณได้ดังสมการ

$$f_{3v}(m, n) = I_v(m, n) \Delta_v^2(m, n)$$

โดยที่ $\Delta_v(m, n) = e_w(m, n) - e_w(m+1, n)$ และ

$I_v(m, n)$ เป็นฟังก์ชันที่จะเลือกเฉพาะค่าความแตกต่างในขอบเขตบล็อกตามแนวตั้ง



ดังนั้น ปัจจัยความผิดเพี้ยนที่ 3 ตามแนวตั้งคำนวณ ได้ดังสมการ

$$F_{3v} = \frac{1}{N_v} \sum_{m,n} f_{3v}(m,n)$$

โดยที่ $N_v = \sum_{m,n} I_v(m,n)$ เป็นจำนวนจุดภาพที่ถูกเลือกโดยฟังก์ชันคั่งกต่าง

ดังนั้น ปัจจัยความผิดเพี้ยนที่ 3 จะคำนวณ ได้ดังสมการ

$$F_3 = \sqrt{F_{3h}^2 + F_{3v}^2}$$

ปัจจัยความผิดเพี้ยนที่ 4 (Correlated errors) แม้ว่าภาพจะไม่แสดงถึงความผิดพลาดแบบจุดถิ่นสุดของบด็อก ลักษณะและโครงร่าง (textures) ของภาพก็จะแสดงถึงความผิดพลาดที่มีความสัมพันธ์กัน ซึ่งสามารถรับรู้ได้มากกว่าความผิดพลาดแบบโดยสุ่ม เพื่อประเมินความผิดพลาดแบบโครงสร้าง เราจะใช้ local spatial correlation โดยปัจจัยที่ 4 จะนิยามว่าเป็นการรวมกัน (summation) ของ local error correlation ของภาพทั้งหมด หากค่าปัจจัยที่ 4 เป็น 0 หมายถึงไม่มีความสัมพันธ์กัน

ภาพแสดงความผิดเพี้ยนคำนวณ ได้ดังนี้

$$f_4(m,n) = \sum_{(k,l) \in W} |r(m,n,k,l)|^{0.25}$$

โดยที่

$$r(m,n,k,l) = \frac{1}{n-1} \left[\sum e_w(i,j)e_w(i+k,j+l) - \frac{1}{n} \sum e_w(i,j) \sum e_w(i+k,j+l) \right]$$

ดังนั้น ปัจจัยความผิดเพี้ยนที่ 4 คำนวณ ได้ดังสมการ

$$F_4 = \frac{1}{MN} \sum_{m,n} f_4(m,n)$$

ปัจจัยความผิดเพี้ยนที่ 5 (Errors in the vicinity of high contrast image transitions) เป็นความผิดพลาดที่เกิดขึ้นกับการเข้ารหัสภาพแทบทุกชนิด แสดงถึงความผิดพลาดในบริเวณที่มี

การเปลี่ยนแปลงความถี่ในภาพสูง หรือการเปลี่ยนแปลงจากความถี่หนึ่งไปอีกความถี่หนึ่งในภาพหรือขอบต่างๆ การเกิดความผิดพลาดเช่นนี้ ทำให้เกิดผลกระทบ 2 แบบ คือ วิวชดมาตส์กิง (Visual masking) จะทำให้การมองเห็นสัญญาณรบกวน (noise) ลดลง และมองเห็นความไม่เป็นระเบียบหรือเบลอ (Blur) มากขึ้น

แม้ว่าวิวชดมาตส์กิงจะลดการมองเห็นความถี่ของของคุณภาพของภาพในบริเวณที่มีการเปลี่ยนแปลงความถี่สูง เทคนิคการเข้ารหัสจะทำให้เกิดข้อผิดพลาดขนาดใหญ่และถี่รบกวนการรับรู้อื่นๆ ในบริเวณเดียวกัน ปัจจัยความผิดพลาด F_s จะวัดรวมทุกๆ ข้อผิดพลาดและถี่รบกวนในบริเวณที่มีการเปลี่ยนแปลงความถี่สูง ในส่วนที่ไม่เหมือนกับปัจจัยความผิดพลาดอื่นๆ คือ ปัจจัยความผิดพลาดที่ 5 จะใช้ภาพต้นฉบับและ $e_w(m, n)$ เท่านั้น ไม่ใช่ภาพที่ถูกบีบ

ปัจจัยฮอริซอนทัลมาตส์กิง (horizontal masking factor) คำนวณดังสมการ

$$S_h(m, n) = e^{\{-0.04V_h(m, n)\}}$$

ฟังก์ชันฮอริซอนทัลโลคอลลแอคทีวิตี (horizontal local activity function) คำนวณดัง

สมการ

$$V_h(m, n) = \frac{|i(m, n-1) - i(m, n+1)|}{2}$$

ปัจจัยวอติคัลมาตส์กิง (vertical masking factor) คำนวณดังสมการ

$$S_v(m, n) = e^{\{-0.04V_v(m, n)\}}$$

ฟังก์ชันวอติคัลโลคอลลแอคทีวิตี (vertical local activity function) คำนวณดังสมการ

$$V_v(m, n) = \frac{|i(m-1, n) - i(m+1, n)|}{2}$$

เราสามารถคำนวณข้อผิดพลาดแบบมาตส์ (masked error) ที่เกิดในแต่ละจุดภาพได้

โดย

$$f_s(m, n) = I_M(m, n) |e_w(m, n)| (S_h(m, n) + S_v(m, n))$$

โดยที่ $I_M(m, n)$ คือ ฟังก์ชันที่เลือกจุดภาพที่ใกล้เคียงกับบริเวณที่มีการเปลี่ยนแปลง
ความถี่สูง

สุดท้าย ปัจจัยความผิดเพี้ยนที่ s คำนวณได้ดังสมการต่อไปนี้

$$F_s = \frac{1}{N_K} \sum_{m,n} f_s(m, n)$$

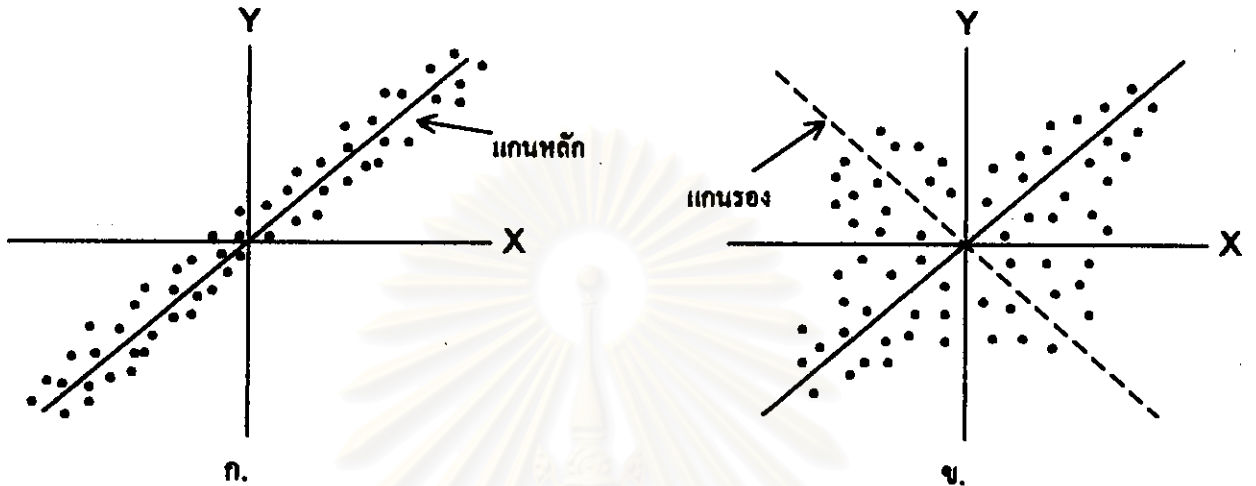
โดยที่ N_K คือ จำนวนจุดภาพที่วิธี 3x3 Kirsch edge ซึ่งเป็นวิธี Edge detection ชนิด
หนึ่ง จะยอมรับค่า $k(m, n) \geq K$ สำหรับค่าขีดแบ่งที่ $K = 400$

เพราะว่าปัจจัยความผิดเพี้ยน (F_s) มีความสัมพันธ์กัน เราจึงต้องแยกความสัมพันธ์
ระหว่างปัจจัยต่างๆ ออกจากกัน โดยใช้วิธีการวิเคราะห์องค์ประกอบหลัก และระบุลักษณะพิเศษ
ขององค์ประกอบหลักนั้นๆ ซึ่งจากการวิเคราะห์องค์ประกอบหลักจะได้เมทริกซ์การผันแปรร่วม
(Covariance matrix) จากนั้นค่ามาตราส่วนคุณภาพภาพจะเป็นผลรวมเชิงเส้น (linear combination)
ของปัจจัยความผิดเพี้ยนที่แยกความสัมพันธ์ออกจากกัน สมบัตินี้จะได้จากวิธีการวิเคราะห์องค์
ประกอบระหว่างคะแนนค่าเฉลี่ยความคิดเห็นและปัจจัยความผิดเพี้ยนที่ผ่านการวิเคราะห์องค์
ประกอบหลัก

2.3.3.2 การวิเคราะห์องค์ประกอบหลัก (Principal component analysis) (ควงพร ชู
รักษ์, 2529: 17-23) (Johnson และ Wichern, 1988) วิเคราะห์ส่วนคุณภาพภาพได้ใช้การวิเคราะห์
องค์ประกอบหลักมาช่วยในการประมาณค่าสัมประสิทธิ์ความอดดอชพหุ เมื่อเกิดพหุสัมพันธ์
ระหว่างตัวแปรอิสระ ด้วยการจัดรูปแบบตัวแปรอิสระเสียใหม่ก่อนที่จะนำไปหาค่าสัมประสิทธิ์
ความอดดอชพหุ โดยให้ตัวแปรใหม่เป็นผลบวกเชิงเส้นของตัวแปรเดิม

การวิเคราะห์องค์ประกอบหลักเป็นวิธีการลดข้อมูลหรือตัวแปรให้น้อยลง โดยอาศัย
หลักความสัมพันธ์เชิงเส้นระหว่างตัวแปรที่ใช้เป็นข้อมูลในการวิเคราะห์ แต่ไม่มีข้อสมมติเกี่ยวกับ
ความสัมพันธ์เชิงสาเหตุและผลระหว่างปัจจัยและตัวแปร เช่น การวิเคราะห์องค์ประกอบหลักที่ใช้
เมื่อมีตัวแปร 2 ตัว คือ X และ Y ก่อนการวิเคราะห์องค์ประกอบหลัก เริ่มจากการศึกษาความ

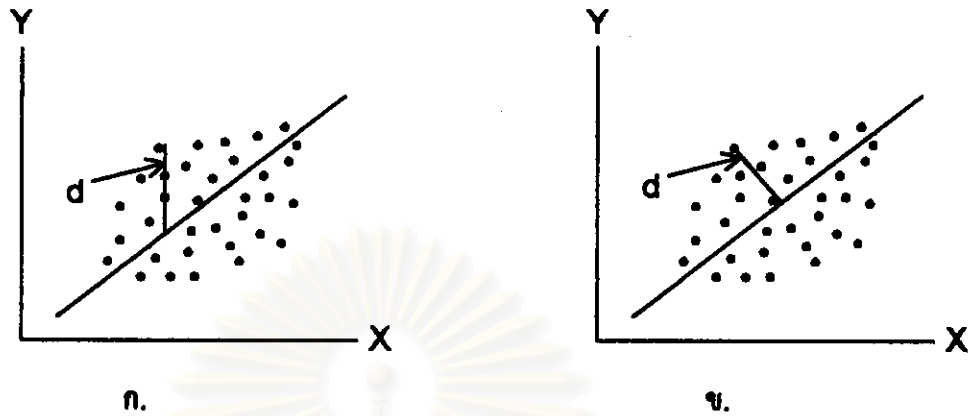
สัมพันธระหว่างตัวแปร สมมติว่า X และ Y มีความสัมพันธ์กัน ซึ่งความสัมพันธ์กันนี้อาจดูได้จาก การลงจุดบนกราฟดังรูปที่ 2.43 รูป ก. และ รูป ข.



รูปที่ 2.43 แสดงความสัมพันธ์ระหว่าง X และ Y รูป ก. X และ Y มีความสัมพันธ์กันสูง
รูป ข. X และ Y มีความสัมพันธ์กันต่ำกว่า (ดวงพร ชูรัตน์, 2529: 17)

จากรูปที่ 2.43 รูป ก. ตัวแปร X และ Y มีความสัมพันธ์กันมาก และเป็นความสัมพันธ์เชิงเส้นตรงในทางบวก ถ้าหากสามารถกำหนดความลาดชันของเส้นตรงนั้น ได้ดังเช่นรูปดังกล่าว ผลลัพธ์ก็จะ ได้เส้นที่สามารถกำหนดค่าของ X เมื่อรู้ค่าของ Y และกำหนดค่าของ Y เมื่อรู้ค่าของ X เส้นที่ได้นี้คือ แกนหลัก (Principal axis) ถ้าจุดต่างๆ นั้นอยู่บนเส้นตรงแกนหลักทั้งหมด แกนหลักก็สามารถที่จะให้ข้อมูลเกี่ยวกับ X และ Y ได้อย่างถูกต้องสมบูรณ์ทุกค่า แต่ถ้าจุดแสดงค่า X และ Y กระจายออกไปมาก ก็ต้องอาศัยแกนเพิ่มอีก 1 แกน ซึ่งแกนที่เพิ่มขึ้นนี้จะต้องมีจุดเริ่มต้นตั้งฉากกับแกนหลักดังที่แสดงไว้ในรูป 2.43 รูป ข. แกนหลักจะลากผ่านจุดต่างๆ ที่ทำให้ระยะระหว่างจุดกับแกนหลัก (โดยการลากเส้นจากจุดมาตั้งฉากกับแกนหลัก) ต่ำที่สุด และทำให้ผลรวมของระยะทางยกกำลังสองมีค่าต่ำสุด การหาค่าต่ำสุดของแกนหลักนี้แตกต่างจากวิธีการหาค่าต่ำสุดของค่าต่ำสุด แสดงดังรูปที่ 2.44 โดยที่การลากเส้นระยะทางตามวิธีกำลังสองน้อยที่สุด เป็นการลากเส้นขนานกับแกน แต่การลากเส้นระยะทางตามวิธีการวิเคราะห์องค์ประกอบหลักจะลากเส้นตั้งฉากกับแกนหลัก

รูปที่ 2.44 รูป ก. และ รูป ข. แสดงให้เห็นความแตกต่างระหว่างการหาค่าต่ำสุดของวิธีกำลังสองน้อยที่สุด และวิธีวิเคราะห์องค์ประกอบหลัก ถึงแม้ว่าทั้งสองวิธีจะพยายามทำให้ระยะทางต่ำสุดเช่นกัน



รูปที่ 2.44 แสดงความแตกต่างระหว่างวิธีการหาค่าต่ำสุด

รูป ก. แสดงการลากเส้นระยะทางต่ำสุดของวิธีการหาค่าต่ำสุดสองค่าต่ำสุด

รูป ข. แสดงการลากเส้นระยะทางต่ำสุดของแกนหลักของวิธีการวิเคราะห์องค์ประกอบหลัก

(ดวงพร ชูรัมย์, 2529: 18)

ถ้ามีจำนวนตัวแปรเพิ่มขึ้นจำนวนมิติของกราฟก็เพิ่มขึ้น เช่น ถ้ามี 3 ตัวแปรก็ต้องเพิ่มเส้นแสดงมิติเพิ่มอีก 1 เส้น และการลงจุดก็ต้องคำนึงถึงค่าของตัวแปร 3 ตัวพร้อมๆ กัน ในการหาแกนหลักก็ต้องการแกนซึ่งสามารถอธิบายการผันแปรของตัวแปรทั้ง 3 ตัวให้ได้มากที่สุด และแกนต่อไป ซึ่งต้องอธิบายการผันแปรที่เหลือให้ได้มากที่สุด

วิธีการแยกแกนตามลำดับจากแกนหลักไปยังแกนรอง หรือการลดตัวแปรตาม วิธีการวิเคราะห์องค์ประกอบหลักนี้ อาศัยสมการที่เรียกว่าสมการไอเกน (Eigen-equation) ซึ่งสามารถเขียนในรูปเมทริกซ์ดังนี้

$$RV = \lambda V$$

โดยที่ R คือ เมทริกซ์ที่ต้องการหา

V คือ เมทริกซ์ของไอเกนเวกเตอร์ (eigenvector)

λ คือ ไอเกนแวลู (eigenvalue)

การแก้สมการ ไอเกนจะได้ไอเกนแวลูและไอเกนเวกเตอร์ที่สัมพันธ์กับเมทริกซ์ข้อมูล ซึ่งวิธีการแก้สมการนี้ อาศัยสมการที่เรียกว่าสมการดีเทอร์มิแนนต์ (Determinant equation)

$$\text{Det}(R - I\lambda) = 0$$

จากตารางที่ 2.22 แสดงเมทริกซ์การผันแปรร่วม (Covariance matrix) ซึ่งเป็นเมทริกซ์ค่าสัมประสิทธิ์ความสัมพันธ์ระหว่างตัวแปร โดยมีค่าสัมประสิทธิ์ของเมทริกซ์ตามแนวเฉียงเป็น 1 เสมอ ตารางที่ 2.23 แสดงค่าไอเกนแวลูและไอเกนเวกเตอร์ของแต่ละองค์ประกอบ

	F ₁	F ₂	F ₃	F ₄	F ₅
F ₁	1.00000	0.99671	0.97530	0.83914	0.58231
F ₂	0.99671	1.00000	0.97432	0.82308	0.56825
F ₃	0.97530	0.97432	1.00000	0.87491	0.56958
F ₄	0.83914	0.82308	0.87491	1.00000	0.67142
F ₅	0.58231	0.56825	0.56958	0.67142	1.00000

ตารางที่ 2.22 แสดงตัวอย่างของเมทริกซ์การผันแปรร่วม

λ_1	λ_2	λ_3	λ_4	λ_5
4.191650	0.591437	0.190214	0.023917	0.002781

L ₁	L ₂	L ₃	L ₄	L ₅
0.475001	-0.244488	0.270834	-0.387094	0.701004
0.471985	-0.268292	0.323841	-0.310442	-0.709931
0.475262	-0.247235	0.019208	0.842713	0.049659
0.450578	0.105979	-0.862187	-0.200654	-0.046043
0.350300	0.892126	0.279357	0.057934	-0.002157

ตารางที่ 2.23 แสดงตัวอย่างของไอเกนแวลูและไอเกนเวกเตอร์

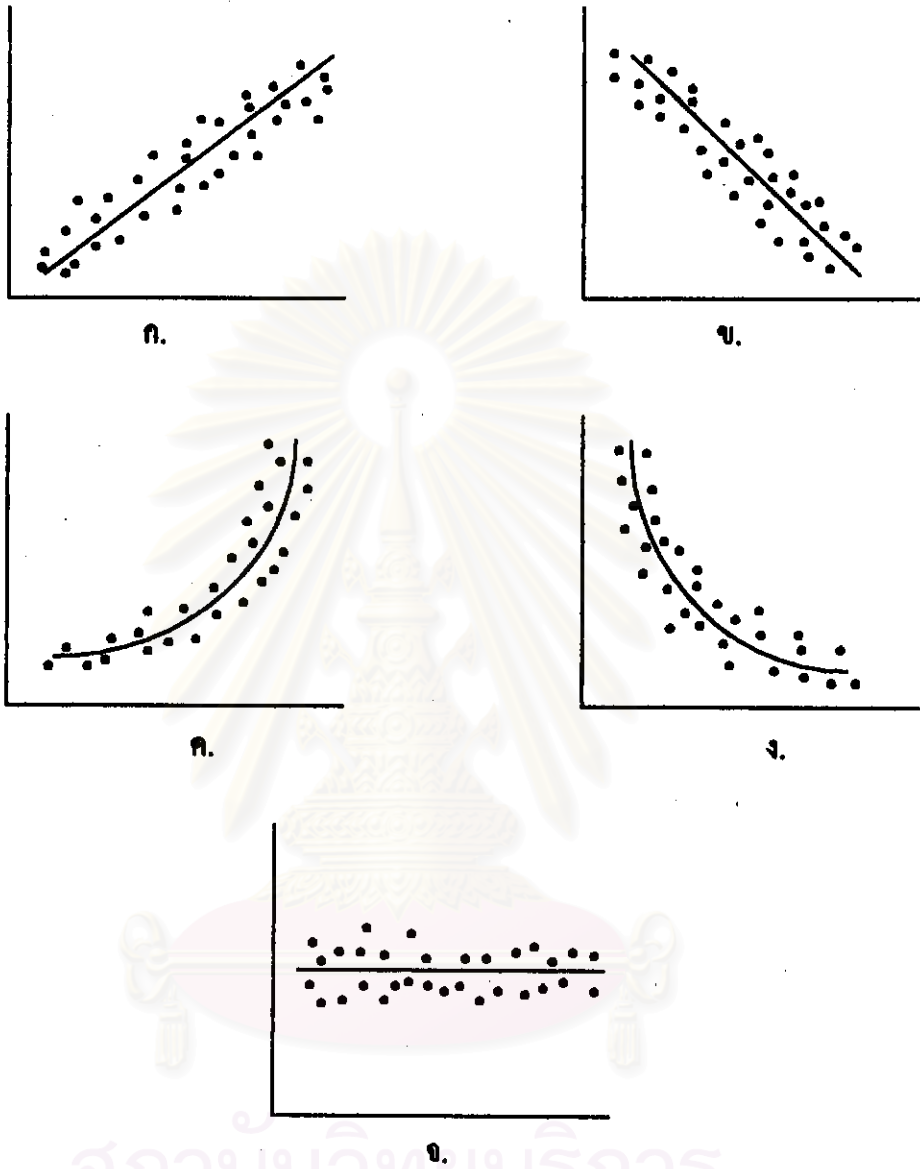
แกนหลักคือ แกนที่ไอเกนแวลูมีค่ามากที่สุด แกนรองจะมีค่าของไอเกนแวลูรองลงมาตามลำดับ ผลรวมของไอเกนแวลูทั้งหมดจะเท่ากับจำนวนตัวแปร ถ้าหารไอเกนแวลูค่าแรกด้วยจำนวนตัวแปรจะได้ค่าของความผันแปรของตัวแปรที่อธิบายได้โดยแกนหลักและค่าผล

หารต่อๆ ไป คือ ค่าของความผันแปรที่อธิบายได้โดยแกนรองตามลำดับ อาจสรุปได้ว่า สัดส่วนของความผันแปรของแกนแต่ละแกนเท่ากับ โยเก็นแวลลิวของแต่ละแกนหารด้วยจำนวนตัวแปร น้ำหนักของตัวแปรที่มีต่อองค์ประกอบ (Principal component loading) คือ อัตราความผันแปรของตัวแปรที่อธิบายได้โดยแกนหลักและแกนรอง ซึ่งเท่ากับผลคูณระหว่าง โยเก็นแวลลิวและรากที่สองของ โยเก็นแวลลิวของแกน

2.3.3.3 การวิเคราะห์ถดถอยพหุ (Multiple regression analysis) (สมบูรณ์ สุขพงษ์ และ เปรมใจ ศรีศรานุกวัฒนา, 2527: 211-241) (อภิรัช บวรกิติวงศ์, 2532: 112-153) การวิเคราะห์ถดถอยพหุเป็นวิธีการทางสถิติที่ใช้ในการวิเคราะห์หารูปแบบของความสัมพันธ์ระหว่างตัวแปรตาม (Dependent variable) หนึ่งตัว ปกติใช้สัญลักษณ์ Y และตัวแปรอิสระ (Independent variable) ตั้งแต่ 2 ตัวขึ้นไป ปกติใช้สัญลักษณ์ X โดยมีข้อสมมติฐานว่า ความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ เป็นความสัมพันธ์เชิงเส้น (Linear relationship) เพื่อที่จะทำให้สามารถนำผลของตัวแปรอิสระแต่ละตัวที่มีต่อตัวแปรตามมารวมกันได้ กล่าวคือ ตัวแปรอิสระแต่ละตัวต้องเป็นเอกเทศจากกันและกัน หรือหากตัวแปรอิสระมีความสัมพันธ์หรือปฏิภริยาคู่กันและกัน เราสามารถแยกปฏิภริยาของความสัมพันธ์ดังกล่าวมาเป็นตัวแปรอิสระตัวหนึ่งหรือหลายตัว และถ้าหากความสัมพันธ์ระหว่างตัวแปรไม่เป็นเชิงเส้น เราจะต้องแปลงความสัมพันธ์ดังกล่าวให้เป็นเชิงเส้นเพื่อให้สอดคล้องกับคุณสมบัติขั้นพื้นฐานของการวิเคราะห์การถดถอย

สิ่งที่ได้จากการวิเคราะห์ถดถอยคือ รูปแบบหรือสมการของความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ที่แสดงให้เห็นว่าตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรตามหรือไม่มากน้อยเพียงใด และเป็นไปในเชิงบวก หรือในเชิงลบ ถ้าตัวแปรอิสระตัวใดมีความสัมพันธ์ในเชิงบวกกับตัวแปรตาม (ซึ่งดูได้จากเครื่องหมายที่อยู่หน้าค่าสัมประสิทธิ์ของความสัมพันธ์หน้าตัวแปรอิสระตัวนั้น) หมายความว่า เมื่อค่าของตัวแปรอิสระตัวนั้นเพิ่มขึ้น ค่าของตัวแปรตามจะเพิ่มขึ้น แต่ถ้าเครื่องหมายเป็นลบ เมื่อค่าของตัวแปรอิสระเพิ่มขึ้น ค่าของตัวแปรตามจะลดลง

จากรูปที่ 2.45 แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม รูป ก. แสดงความสัมพันธ์เชิงเส้นตรงแบบบวก รูป ข. แสดงความสัมพันธ์เชิงเส้นตรงแบบลบ รูป ค. แสดงความสัมพันธ์เชิงเส้นโค้งแบบบวก รูป ง. แสดงความสัมพันธ์เชิงเส้นโค้งแบบลบ และรูป จ. ไม่มีสหสัมพันธ์ต่อกัน



รูปที่ 2.45 แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม
 รูป ก. ความสัมพันธ์เชิงเส้นตรงแบบบวก รูป ข. ความสัมพันธ์เชิงเส้นตรงแบบลบ
 รูป ค. ความสัมพันธ์เชิงเส้นโค้งแบบบวก รูป ง. ความสัมพันธ์เชิงเส้นโค้งแบบลบ
 รูป จ. ไม่มีสหสัมพันธ์ต่อกัน (สมบุญ ฤชพงษ์ และ เปรมใจ ศรีตราวุฒนา, 2527: 215)

เมื่อได้รูปแบบของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามแล้ว จะสามารถนำไปประมาณค่าของตัวแปรตามได้ โดยการกำหนดค่าของตัวแปรอิสระแต่ละตัว จะตัวเดียวหรือหลายตัวก็ได้ กล่าวคือ หากเรารู้ค่า Y ก็สามารถหาค่า X ได้ หรือทราบค่า X ก็สามารถหาค่า Y

ได้ ความถูกต้องของการประมาณค่า ย่อมขึ้นอยู่กับอัตราของความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามว่า มีมากน้อยเพียงใด หากมีความสัมพันธ์กันมาก ก็สามารถประมาณค่าของตัวแปรตามได้ถูกต้องมากกว่ากรณีที่ตัวแปรอิสระมีความสัมพันธ์กับตัวแปรตามน้อย

นอกจากต้องการประมาณความสัมพันธ์ระหว่างตัวแปรและหาค่าประมาณของตัวแปรตามจากตัวแปรอิสระ โดยการแทนค่าตัวแปรอิสระลงไปในสมการ ซึ่งจะทราบค่าประมาณของตัวแปรตามที่ต้องการแล้ววัตถุประสงค์อีกอย่างหนึ่งในการศึกษาเรื่องการถดถอย คือ การวัดค่าความคลาดเคลื่อนจากการใช้เส้นถดถอย (Regression line) เป็นพื้นฐานในการประมาณค่า (เส้นถดถอย คือ เส้นที่ลากผ่านจุดต่างๆ ในกราฟแสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ซึ่งเส้นถดถอยนี้เป็นได้ทั้งเส้นตรงหรือไม่เป็นเส้นตรงก็ได้) ค่าที่วัดออกมานั้น เรียกว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard error of estimate) ซึ่งวัดการกระจายของค่า Y รอบๆ เส้นถดถอย ชั้นเส้นถดถอยนี้ประมาณข้อมูลได้ดี ข้อมูลจะกระจายอยู่รอบเส้นถดถอย ค่าความคลาดเคลื่อนที่วัดได้จะมีค่าต่ำ ในทางตรงกันข้ามถ้าข้อมูลกระจายอยู่รอบเส้นถดถอยมาก ค่าความคลาดเคลื่อนนั้นก็จะมีค่าสูง

สมการการถดถอยพหุมีรูปแบบดังนี้

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon$$

โดยที่ $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ คือ ค่าสัมประสิทธิ์ของความถดถอย (Regression coefficient หรือ Partial regression coefficient)

Y คือ ตัวแปรตาม

$X_1, X_2, X_3, \dots, X_k$ คือ ตัวแปรอิสระจำนวน k ตัว

α คือ พหามิเตอร์ที่ไม่ทราบค่า

ε คือ ค่าผิดพลาด (Error term)

ดังนั้น สมการมาตรฐานคุณภาพภาพสามารถคำนวณได้โดยการวิเคราะห์ถดถอยพหุขององค์ประกอบหลัก (Z) ซึ่งเป็นปัจจัยความผิดเพี้ยนที่ผ่านการวิเคราะห์องค์ประกอบหลัก มีรูปแบบสมการดังนี้

$$PQS = b_0 + \sum_{j=1}^l b_j Z_j$$

โดยที่ b_0 คือ พารามิเตอร์ที่ไม่ทราบค่า

b_j คือ ค่าสัมประสิทธิ์การถดถอยที่คำนวณได้จากการวิเคราะห์ถดถอยพหุ

Z_j คือ องค์ประกอบหลักของปัจจัยทั้ง 5 ที่ผ่านการวิเคราะห์องค์ประกอบหลัก

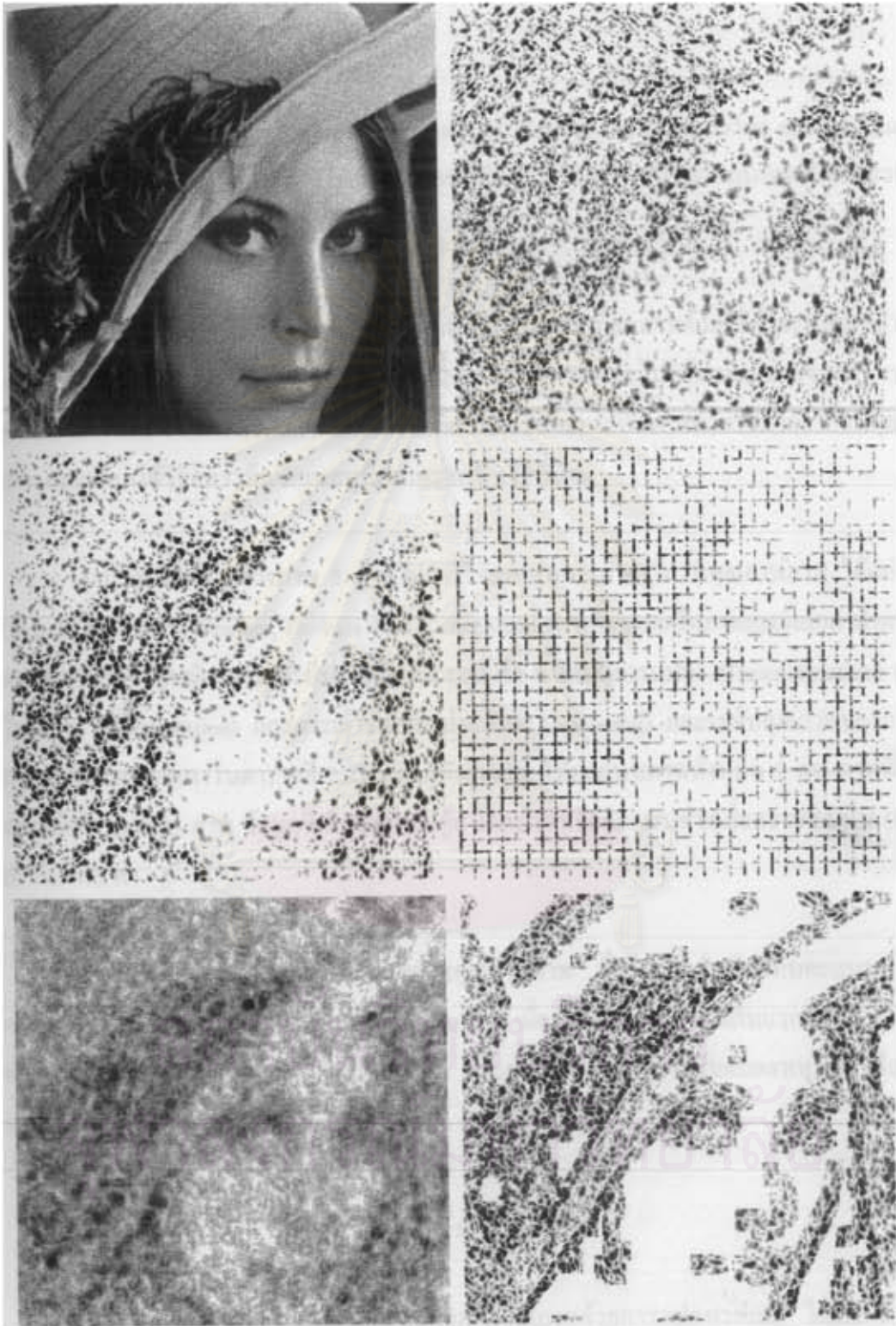
เมื่อปัจจัยทั้ง 5 ผ่านการวิเคราะห์องค์ประกอบหลัก ผลลัพธ์ คือ เลือกเพียง 3 องค์ประกอบหลักก็จะให้ความเชื่อถือที่ 99% โดยมีลักษณะเฉพาะขององค์ประกอบหลักดังนี้ องค์ประกอบที่ 1 จะเป็นข้อผิดพลาดประเภทจำนวนที่ผิดพลาด (The amount of error) องค์ประกอบที่ 2 จะเป็นข้อผิดพลาดประเภทตำแหน่งที่ผิดพลาด (The location of error) และองค์ประกอบที่ 3 จะเป็นข้อผิดพลาดประเภทโครงสร้างที่ผิดพลาด (The structure of error)

ตัวอย่างสมการมาตรฐานส่วนคุณภาพภาพที่ผ่านการวิเคราะห์องค์ประกอบหลักและการวิเคราะห์การถดถอยพหุดังนี้

$$PQS = 5.632 - 0.068Z_1 - 1.536Z_2 - 0.0704Z_3$$

สำหรับภาพที่คุณภาพแย่มาก เป็นไปได้ที่จะได้ค่ามาตรฐานส่วนคุณภาพภาพที่เป็นค่าลบ และถ้าหากต้องการให้วิธีมาตรฐานส่วนคุณภาพภาพมีประสิทธิภาพเพิ่มขึ้น ก็สามารถแบ่งเป็น 2 มาตรฐาน ดังนี้ มาตรฐานสำหรับภาพคุณภาพต่ำถึงปานกลาง และมาตรฐานสำหรับภาพคุณภาพปานกลางถึงสูง

จากรูปที่ 2.46 แสดงภาพตัวอย่าง Lena ที่ผ่านการประมวลผลด้วยวิธีมาตรฐานส่วนคุณภาพภาพ ได้แก่ ภาพแสดงความคิดเห็นที่ 1, 2, 3, 4 และ 5 ตามลำดับ ภาพที่แสดงปัจจัยความคิดเห็นที่ 1 จะแสดงค่าความแตกต่างในแต่ละจุดภาพระหว่างภาพต้นฉบับและภาพที่ผ่านการบีบตามมาตรฐานการวัดสัญญาณรบกวนของโทรทัศน์ CCIR 567-1 ภาพที่แสดงปัจจัยความคิดเห็นที่ 2 จะแสดงค่าความแตกต่างในแต่ละจุดภาพระหว่างภาพต้นฉบับและภาพที่ผ่านการบีบตามมาตรฐานแบบจำลองของ HVS ภาพที่แสดงปัจจัยความคิดเห็นที่ 3 ลักษณะภาพจะเป็นบล็อกเล็กๆ ที่มีขนาด 8x8 จุดภาพ ที่เกิดจากเทคนิคการบีบภาพที่ทำให้เกิดบล็อกกิงอาร์ทิแฟกต์ ภาพที่แสดงปัจจัยความคิดเห็นที่ 4 แสดงข้อผิดพลาดในแต่ละจุดภาพที่มีความสัมพันธ์กันในโครงสร้างของภาพ ภาพที่แสดงปัจจัยความคิดเห็นที่ 5 จะแสดงข้อผิดพลาดของบริเวณที่มีการเปลี่ยนแปลงความเข้มแสงของขอบต่างๆ ในภาพ



รูปที่ 2.46 แสดงภาพที่ผ่านการประมวลผลด้วยวิธีมาตราส่วนคุณภาพภาพ
จากภาพต้นฉบับภาพชายบน ภาพ Lena และปัจจัยความคิดเห็นที่ 1, 2, 3, 4 และ 5 ตามลำดับ
(Miyahara, Kotani และ Algazi, 1996)

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่อง “Comparative Performance of Wavelet and JPEG coders at High Quality” (Algazi and Estes, 1997) เป็นงานวิจัยการเปรียบเทียบสมรรถนะการบีบภาพแบบเวฟเลตและเวฟเลตที่ภาพคุณภาพสูง

ในงานวิจัยนี้ ได้พิจารณาว่าการใช้เกณฑ์การเปรียบเทียบภาพด้วยวิธีค่าความผิดพลาดมีนสมควร์ (MSE) และอัตราทศิกนณดฐนื้ออ (PSNR) ไม่เพียงพอและไม่ให้ความหมายของการเปรียบเทียบภาพ ดังนั้น จึงเลือกใช้เกณฑ์การเปรียบเทียบสมรรถนะการบีบภาพด้วยวิธีมาตราส่วนคุณภาพภาพ และเลือกใช้วิธีอัตราทศิกนณดฐนื้ออ เพราะเป็นวิธีที่นิยม

งานวิจัยนี้ใช้ตัวอย่างภาพ 5 ภาพในการวิจัหาสมการมาตราส่วนคุณภาพภาพ ได้แก่ ภาพ Church, Hairband, Weather, Barbara และ Cameraman ภาพทั้งหมดเป็นภาพเกรย์สเกล นำภาพทั้งหมดมาเปรียบเทียบขั้นตอนการบีบภาพ 3 ขั้นตอน คือ ขั้นตอนการแปลง (Transformation) ควอนไทเซชัน (Quantization) และการเข้ารหัสสัมประสิทธิ์ (Encoding) และระดับคุณภาพของการบีบภาพ 5 ระดับดังแสดงในตารางที่ 2.21 กลุ่มผู้สังเกตและให้คะแนนภาพทั้งหมด 9 คน ภาพที่ใช้ในงานวิจัยทั้งหมด 75 ภาพ ข้อมูลที่เก็บจากกลุ่มผู้สังเกตภาพหรือคะแนนค่าเฉลี่ยความคิดเห็น (MOS) เท่ากับ 675 ค่า

ด้วยการเปรียบเทียบโดยวิธีมาตราส่วนคุณภาพภาพ ซึ่งมีความสัมพันธ์กับคะแนนค่าเฉลี่ยความคิดเห็นที่สัมประสิทธิ์ความสัมพันธ์เท่ากับ 0.93 เมื่อพิจารณาจากค่าไถ่แ่นแวลดิวที่มากที่สุด 3 องค์ประกอบหลักจะให้ความถูกต้องถึง 98% และเมื่อผ่านการวิเคราะห์หัดดอของททุ จะได้สมการมาตราส่วนคุณภาพภาพได้ดังนี้

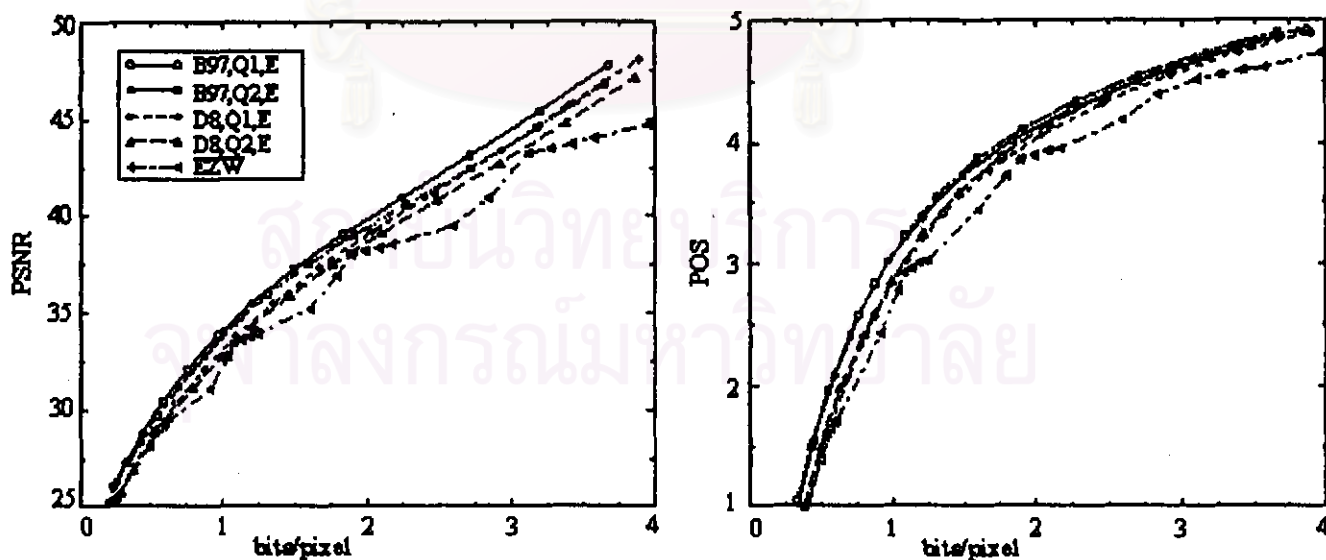
$$PQS = 5.632 - 0.068Z_1 - 1.536Z_2 - 0.0704Z_3$$

ดังนั้น จากงานวิจัยนี้เลือกเปรียบเทียบการบีบภาพด้วยการแปลงเวฟเลต ได้แก่ ตัวกรองสัญญาณแบบ 8 tap orthogonal wavelet ของ Daubechies (โดยใช้ตัวย่อว่า D8) และ “9-7” wavelet ของ Barlaud (ใช้ตัวย่อว่า B97) เลือกใช้มาตราส่วน (Scale) เท่ากับ 4 และแยกได้ 13 คลื่นย่อย

ขั้นตอนควอนไทเซชัน ได้แก่ Single uniform scalar quantization (โดยใช้ตัวย่อว่า Q1) และ The HVS frequency weighted quantizer ของ Lewis และ Knowles (ใช้ตัวย่อว่า Q2) และขั้นตอนการเข้ารหัสสัมประสิทธิ์ที่ถูกควอนไทซ์ ได้แก่ The “color shrinking” (ใช้ตัวย่อว่า E) รวมทั้งวิธี EZW ดังนั้น การบีบภาพทั้งหมดเท่ากับ $\{D8, B97\} \times \{Q1, Q2\} \times \{E\}$ และวิธี Embedded zerotree wavelet algorithm หรือ EZW และหลังจากการเปรียบเทียบภาพด้วยการแปลงเวฟเลตชนิดต่างๆ จะนำวิธีที่ดีที่สุดของการแปลงมาเปรียบเทียบกับวิธีบีบด้วยเทคนิคเชิงเท็ก

ผลลัพธ์ B97 เป็นเทคนิคที่ให้ค่ามาตราส่วนคุณภาพภาพ และอัตราบิตซิกเนลสูงน้อยกว่า D8 แตกต่างกันเท่ากับ 0.43 PQS และ 1.2 dB จากการใช้การแปลงแบบ B97 สามารถประหยัดการใช้จำนวนบิตเท่ากับ 0.2 บิตต่อจุดภาพ สำหรับค่ามาตราส่วนคุณภาพภาพและอัตราบิตซิกเนลสูงที่กล่าวมา เมื่อนำ B97 เปรียบเทียบกับ EZW พบว่า B97 มีสมรรถนะดีกว่า EZW

ควอนไทเซชันแบบ Q1 จะลดค่าวิธีคำนวณผิดพลาดมินตแควร์ ซึ่งผลลัพธ์จาก Q1 จะดีกว่า Q2 เมื่อเปรียบเทียบกับอัตราบิตซิกเนลสูง แต่เมื่อเปรียบเทียบกับค่ามาตราส่วนคุณภาพภาพ Q2 จะดีกว่า Q1 แสดงให้เห็นว่าการเปรียบเทียบกับอัตราบิตซิกเนลสูงไม่เพียงพอ ดังนั้นวิธี (B97, Q2, E) ให้สมรรถนะที่ดีที่สุด การเปรียบเทียบแสดงดังรูปที่ 2.47

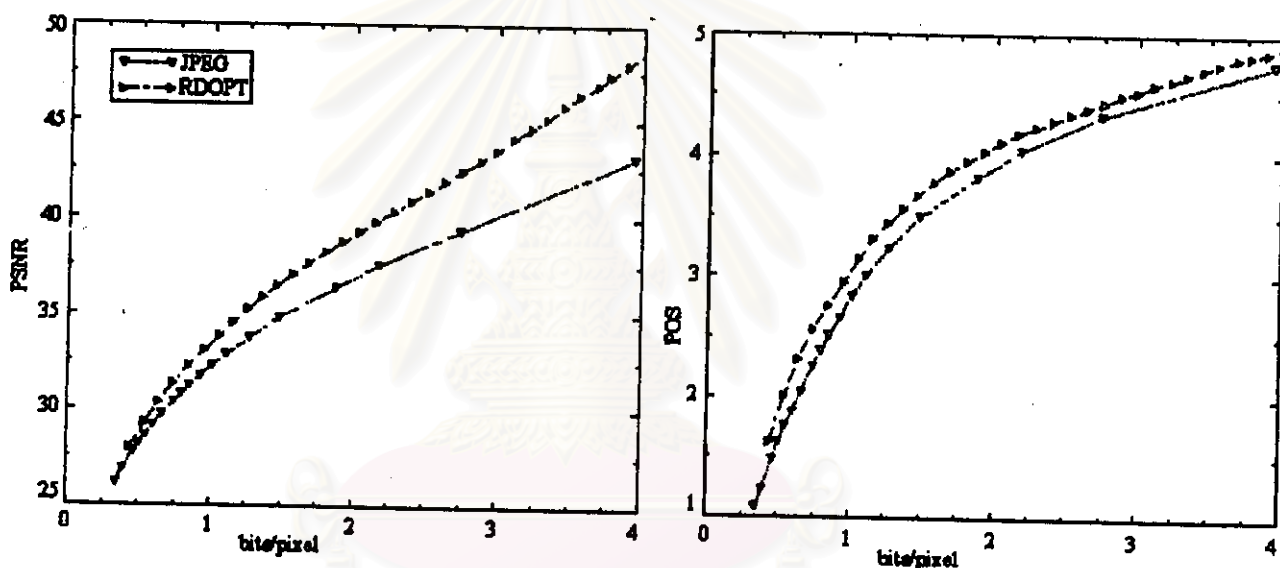


รูปที่ 2.47 แสดงการเปรียบเทียบการเข้ารหัสภาพด้วยเวฟเลตของภาพ “hotel”

(Algazi และ Estes, 1997)

สำหรับการบีบภาพด้วยเจเพ็ก ในงานวิจัยนี้ จะเลือกวิธีการเข้ารหัสแบบเบสไลน์เจเพ็ก เปรียบเทียบกับ Rate-distortion quantizer optimization technique หรือ RDOPT เมื่อใช้การเปรียบเทียบด้วยอัตราที่ลจิกเนตพูนี้อย พบว่า RDOPT สามารถลดอัตราบิตลงได้ถึง 1 บิตต่อจุดภาพและให้ค่าอัตราที่ลจิกเนตพูนี้อยได้ถึง 5 dB จากตัวอย่างภาพ ได้ค่าเท่ากับ 0.5 บิตต่อจุดภาพ และ 3 dB

เมื่อใช้การเปรียบเทียบด้วยค่ามาตราส่วนคุณภาพภาพ การลดลงของอัตราบิตจะน้อยกว่า 0.5 บิตต่อจุดภาพ และให้ค่ามาตราส่วนคุณภาพภาพน้อยกว่า 0.4 PQS จากตัวอย่างภาพ ได้ค่าเท่ากับ 0.25 บิตต่อจุดภาพและ 0.25 PQS การเปรียบเทียบแสดงดังรูปที่ 2.48



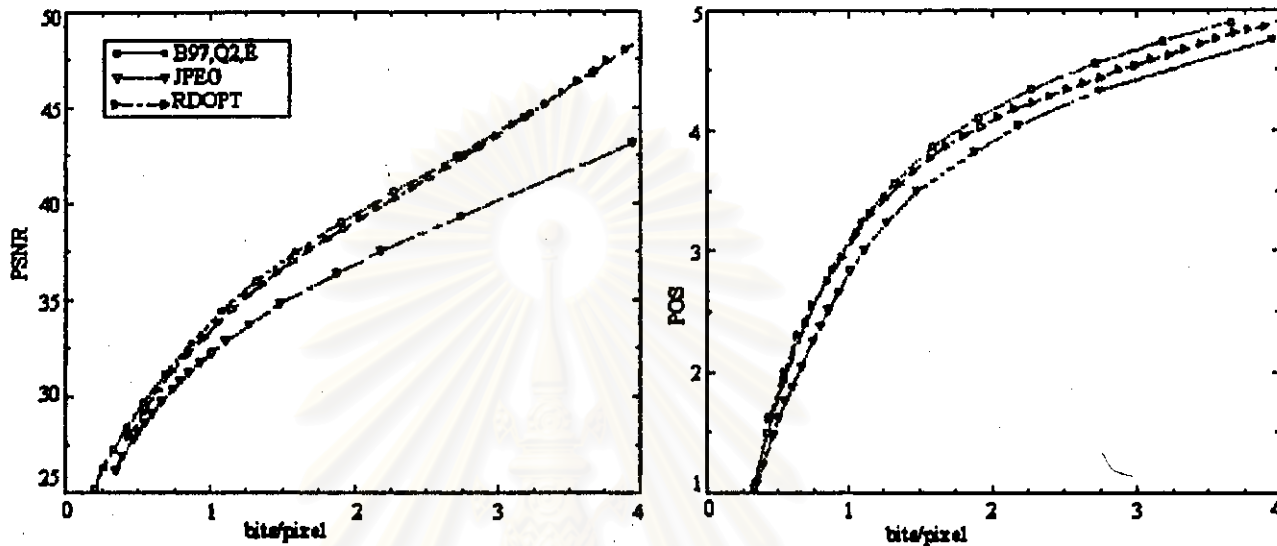
รูปที่ 2.48 แสดงการเปรียบเทียบการเข้ารหัสภาพด้วยเจเพ็กของภาพ "hotel"

(Algazi และ Estes, 1997)

เมื่อนำการแปลงเวฟเลตวิธีที่ดีที่สุด คือ {B97, Q2, E} เปรียบเทียบกับการบีบแบบเบสไลน์เจเพ็กและ RDOPT ด้วยภาพ "hotel" เมื่อพิจารณาที่ค่ามาตราส่วนคุณภาพภาพต่ออัตราบิต จากตัวอย่างภาพจะพบว่า การแปลงเวฟเลตให้ผลลัพธ์ดีกว่าการบีบด้วยเจเพ็กเท่ากับ 0.25 PQS ที่อัตราบิตน้อยกว่าเท่ากับ 0.2 บิตต่อจุดภาพ และดีกว่า RDOPT เท่ากับ 0.1 PQS ที่อัตราบิต 0.4 บิตต่อจุดภาพ และสามารถมีค่าสูงกว่าเท่ากับ 0.3 PQS ที่อัตราบิต 0.7 บิตต่อจุดภาพ

เมื่อเปรียบเทียบภาพทดสอบทุกๆ ภาพ การบีบด้วยการแปลงเวฟเลต {B97, Q2, E} จะดีกว่าจนที่ประมาณ 0.2-0.8 PQS และต้องการอัตราบิตประมาณ 0.15-0.8 บิตต่อจุดภาพ จากตัวอย่าง

ภาพ ค่าที่ได้เท่ากับ 0.4 PQS และอัตราบิด 0.5 บิตต่อจุดภาพ เมื่อเปรียบเทียบกับอัตราพีคซิกเนลที่ น้อยค่าจะแตกต่างกันประมาณ 1.1 บิตต่อจุดภาพ และ 5.5 dB การเปรียบเทียบแสดงดังรูปที่ 2.49



รูปที่ 2.49 แสดงการเปรียบเทียบการเข้ารหัสภาพระหว่างเจทีคและเวฟเลตของภาพ "hotel"

(Algazi และ Estes, 1997)

ดังนั้น จากงานวิจัยนี้ สรุปได้ว่า การบีบภาพด้วยการแปลงเวฟเลตโดยใช้ตัวกรองสัญญาณ "9-7" wavelet ของ Barlaud คอนโทเนียนด้วยวิธี The HVS frequency weighted quantizer ของ Lewis และ Knowles และการเข้ารหัสแบบ The "color shrinking" จะให้ผลลัพธ์จากการเปรียบเทียบ กับเบสไลน์เจทีคดีกว่า และมีสมรรถนะดีกว่า RDOPT เพียงเล็กน้อย

สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย