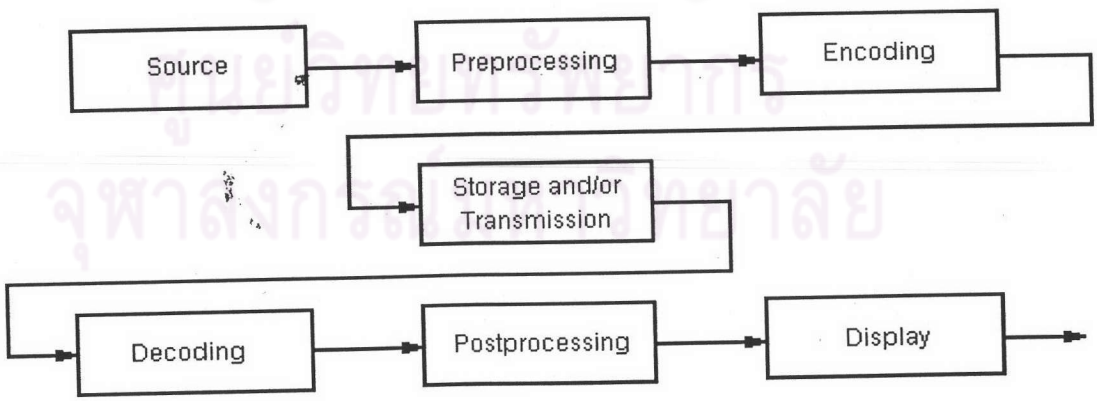
บทที่ 2

หลักการทำงานของ MPEG

ในบทนี้จะกล่าวถึงมาตรฐาน ISO/IEC 11172 หรือที่รู้จักกันดีในชื่อ MPEG (Moving Pictures Expert Group) ซึ่งมาตรฐานนี้ได้ถูกกำหนดขึ้นเมื่อปี ค.ศ 1988 โดยความร่วมมือของ ISO (International Standards Organization) และ IEC (International Electrotechnical Commission) ได้ร่วมกันพัฒนามาตรฐานของการนำเสนอภาพเคลื่อนไหว เสียง และข้อมูลอื่นที่เก็บและค้นคืนในสื่อข้อมูลดิจิตอล โดยเน้นถึงกระบวนการเข้ารหัสและถอดรหัสสัญญาณภาพและโครงสร้างของข้อมูลของ MPEG

2.1 หลักการของสัญญาณภาพ

มาตรฐาน MPEG-1 จะทำงานได้ดีในการสื่อสารข้อมูลที่มีความเร็วของการรับ-ส่งข้อมูล 1.5 เมกะบิตต่อวินาที นั้นหมายถึงว่าภาพเคลื่อนไหวสามารถที่จะจัดเก็บในรูปแบบของข้อมูลคอมพิวเตอร์และรับ-ส่งในระบบเครือข่ายที่มีอยู่ในปัจจุบันและอนาคต รวมถึงการส่งข้อมูลสำหรับโทรทัศน์ระบบดิจิตอลบนสื่อ วิทยาการเส้นใยนำแสง ดาวเทียม เคเบิล ISDN (Integrated Services Digital Network) และ ATM (Asynchronous Transfer Mode) การแสดงภาพสามารถใช้ได้กับโทรทัศน์ทั้งระบบ 625 และ 525 เส้น รวมถึงการแสดงบนจอภาพคอมพิวเตอร์ที่ความละเอียด 350 x 250 จุดใน 1 ภาพ และอัตราการเปลี่ยนภาพ 24-30 ภาพต่อวินาที MPEG ไม่สนใจกระบวนการกวาดตรวจภาพ ดังนั้นถ้าเดิมการแสดงข้อมูลเป็นแบบไขว้กัน (interlaced) จะต้องแปลงเป็นแบบไม่ไขว้กัน (non-interlaced) ก่อนที่จะทำการเข้ารหัส และอาจแปลงกลับเป็นแบบไขว้สำหรับการแสดงผลหลังการถอดรหัส



รูปที่ 2.1 กระบวนการเข้ารหัส และการถอดรหัส

ในรูปที่ 2.1 แสดงถึงกระบวนการเข้าและถอดรหัสของ MPEG ตั้งแต่ข้อมูลภาพเข้ามาจนกระทั่งภาพถูกนำออกแสดง แหล่งของข้อมูลที่ได้มาอาจจะมีได้หลายรูปแบบเช่น แฟ้มข้อมูลคอมพิวเตอร์ แต่โดยทั่วไปแล้วก่อนที่จะเข้ากระบวนการต้องมีการเตรียมข้อมูล (preprocessing) เพื่อปรับความละเอียดของภาพให้เหมาะสม

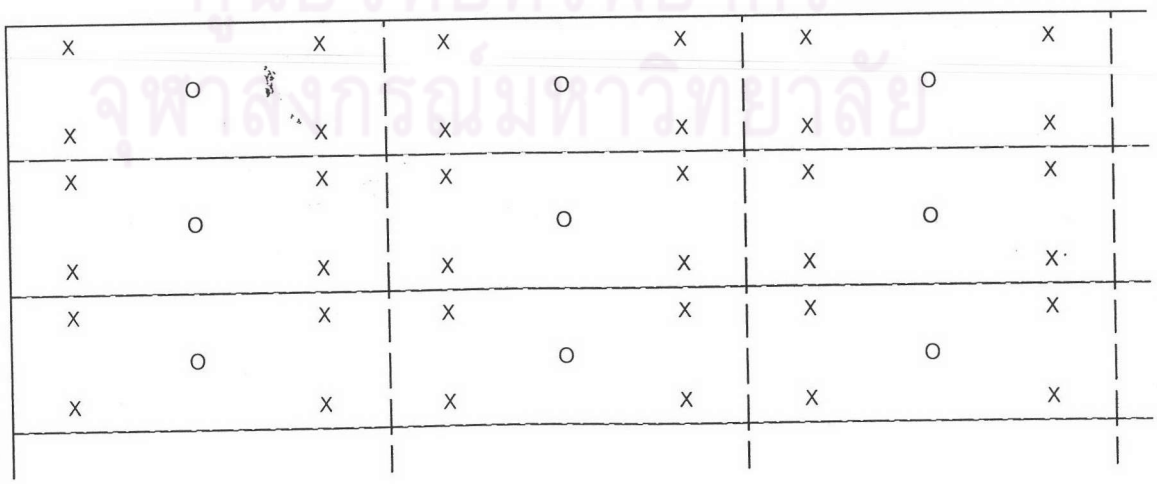
สำหรับภาพที่มีความละเอียดสูงเกินไป ก็จะมีจำนวนข้อมูลที่จะต้องผ่านกระบวนการสูง ทำให้ใช้เวลามากเกินกว่าที่ตัวถอดรหัสจะทำได้ทัน แต่ถ้าความละเอียดต่ำเกินไปรายละเอียดจะหายไป ที่ระดับความเร็วของการส่งข้อมูล 1.5 เมกะบิตต่อวินาที และความเร็วของการแสดงภาพ 24-30 ภาพต่อวินาที ความละเอียดแนวนอนควรอยู่ระหว่าง 250-400 จุดภาพ และ ความละเอียดแนวตั้งควรอยู่ระหว่าง 200-300 จุดภาพ

2.2 ระบบสีของ MPEG

MPEG ใช้ระบบสี luminance/chrominance ซึ่งคิดค้นโดย NTSC (National Television Standards Committee) ซึ่งเป็นผู้คิดค้นระบบสัญญาณโทรทัศน์ ในการส่งสัญญาณภาพทั่วไปเพื่อให้ภาพที่ปรากฏบนจอเครื่องรับโทรทัศน์เหมือนกับภาพที่สถานีส่ง ทางสถานีจำเป็นจะต้องส่งทั้งรายละเอียดและความเข้มของสัญญาณสี (luminance) ออกอากาศ สัญญาณสี (chrominance) คือ แม่สีทั้งสาม ได้แก่ สีแดง สีเขียว และสีน้ำเงิน แต่เนื่องจากถ้าส่งสัญญาณสีออกอากาศทั้งสามสี จะต้องใช้คลื่นพาหะถึง 3 ความถี่ด้วยกัน อันจะทำให้ย่านการส่งสัญญาณโทรทัศน์สีกว้างเกินกว่าที่กำหนด ระบบ NTSC จะเปลี่ยนสัญญาณแม่สีทั้งสามให้เหลือเป็นสัญญาณความต่างสีเพียงสองสัญญาณ คือ สัญญาณความต่างสีแดง (Cr) และสัญญาณความต่างสีน้ำเงิน (Cb) ส่วนสัญญาณสีเขียวจะไม่ส่งออกอากาศโดยตรงแต่ใช้วิธีฝากไปกับสัญญาณความต่างทั้งสอง ระบบสี luminance/chrominance นี้บางคนอาจรู้จักกันในชื่อของระบบ YUV หรือ HSV (Hue-Saturation-Value)

MPEG เลือกใช้ระบบ NTSC เพราะจากการศึกษาพบว่าสายตาของมนุษย์จะไวต่อความสว่างมากกว่าสี ดังนั้นจึงเก็บค่าความสว่างครบแต่เก็บค่าสีเพียงหนึ่งในสี่ องค์ประกอบของภาพประกอบด้วยเมตริกซ์ขนาด 8x8 ของค่าความสว่าง (Y) และเมตริกซ์ แสดงสี 2 ตัว (Cb กับ Cr) จำนวนเมตริกซ์วายจะเป็นจำนวนคู่ไม่ว่าจะเป็นด้านแนวตั้ง หรือแนวนอน ส่วน Cb และ Cr จะเป็นจำนวนครึ่งหนึ่งของ เมตริกซ์วายทั้งแนวตั้งและแนวนอน

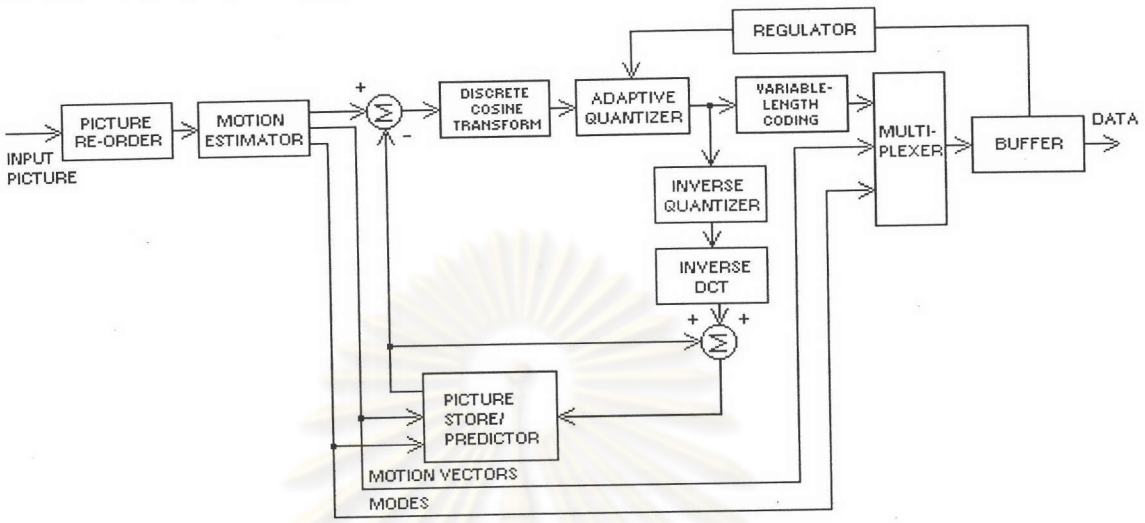
ตัวอย่างตำแหน่งของค่าความสว่าง และค่าของสี ดังแสดงในรูปที่ 2.2 โดยที่ "X" แทนตำแหน่งของ ค่าความสว่าง (Y) และ "O" แทนตำแหน่งของค่าสี Cb และ Cr



รูปที่ 2.2 ตำแหน่งของค่าความสว่าง และค่าของสีบนภาพ

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

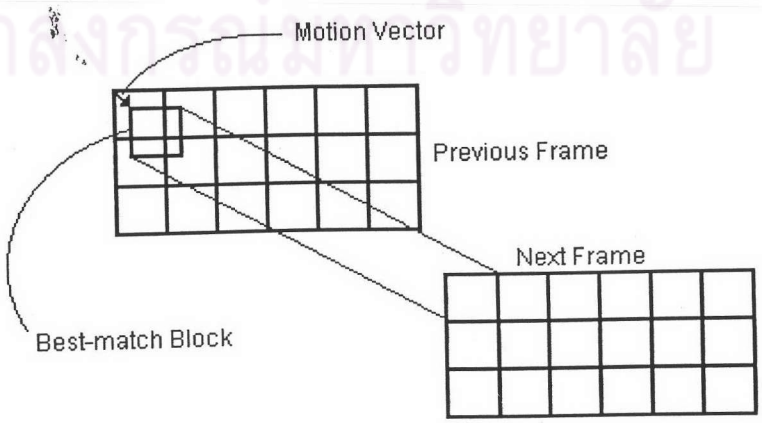
2.3 กระบวนการเข้ารหัสสัญญาณภาพของ MPEG



รูปที่ 2.3 แผนภาพบล็อกการเข้ารหัสสัญญาณภาพ

กระบวนการในการเข้ารหัสข้อมูลภาพของ MPEG นั้นมีหลายขั้นตอน ดังแสดงในรูปที่ 2.3 เริ่มตั้งแต่การกำหนดชนิดของภาพให้แต่ละภาพ ซึ่งชนิดของภาพจะแบ่งออกได้เป็น 3 ประเภทด้วยกันคือ (1) I-Picture หรือภาพที่เข้ารหัสภายใน (Intra Coded Picture) ซึ่งจะเข้าและถอดรหัสโดยไม่เกี่ยวข้องกับภาพอื่นๆ (2) P-Picture หรือภาพที่เข้ารหัสโดยการทำนายภาพจากอดีต (Predictive Coded Picture) ซึ่งเข้าและถอดรหัสโดยใช้ข้อมูลจาก I-Picture หรือ P-Picture ซึ่งเป็นภาพในอดีต และ (3) B-Picture หรือ ภาพที่เข้ารหัสโดยการทำนายแบบสองทาง (Bidirectionally Predictive Coded Picture) ซึ่งเข้าและถอดรหัสโดยใช้ข้อมูลจาก I-Picture และ P-Picture ที่เป็นภาพจากทั้งในอดีตและอนาคต

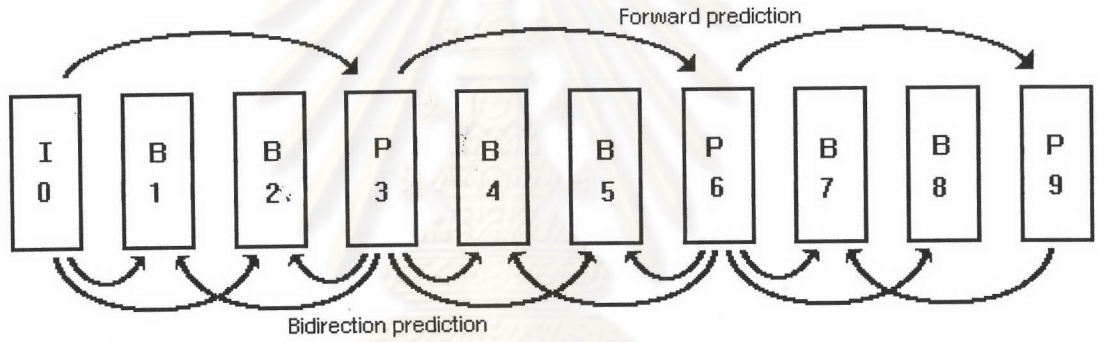
วัตถุประสงค์ของการแบ่งชนิดของภาพนั้น เพื่อที่จะเก็บข้อมูลให้น้อยที่สุดโดยการเก็บภาพชนิด I-Picture จะเก็บทั้งภาพ แต่ถ้าเป็น P-Picture หรือ B-Picture จะเก็บเฉพาะส่วนต่างจากภาพอ้างอิง กระบวนการนี้เรียกว่าการสร้างค่าชดเชยการเคลื่อนที่ (motion compensation) ดังแสดงในรูปที่ 2.4



รูปที่ 2.4 การสร้างค่าชดเชยการเคลื่อนที่

กระบวนการนี้เป็นการทำนายค่าของบล็อกในภาพ โดยการวางบล็อกลงไปบนภาพอ้างอิงที่ทราบอยู่แล้ว การวางจะอาศัยเส้นแสดงตำแหน่งและทิศทางการเคลื่อนที่ (motion vector) ซึ่งเป็นตัวระบุสถานที่ของบล็อก โดยได้จากการใช้บล็อกของภาพอ้างอิงในการหาค่าของบล็อกในปัจจุบัน ตัวอย่างง่ายๆคือ ภาพที่ถ่ายโดยกล้องอยู่กับที่ และวัตถุในภาพไม่เคลื่อนที่ดังนั้นตำแหน่งของภาพจะคงที่ในทุกๆภาพ ค่าของ motion vector ของแต่ละบล็อกเป็น 0 การใช้เทคนิคนี้ใช้ได้กับภาพลำดับสั้นๆที่ในภาพวัตถุส่วนใหญ่อยู่กับที่ แต่มีบางส่วนเคลื่อนไหว

ถ้าภาพที่เข้ารหัสเป็น B-Picture การเข้าและถอดรหัสจะอาศัยข้อมูลจากภาพ I-Picture และ P-Picture นั่นหมายถึงใช้ข้อมูลจากภาพทั้งในอดีตและอนาคต (ดูรูปที่ 2.5 ประกอบ) ดังนั้นในลำดับในการส่งภาพจะต้องส่ง P-Picture ซึ่งเป็นภาพในอนาคตไปก่อนที่จะส่ง B-Picture เพื่อทำให้ผู้รับสามารถถอดรหัสภาพ B-Picture ได้



รูปที่ 2.5 ความสัมพันธ์ระหว่าง I-Picture B-Picture และ P-Picture

รูปที่ 2.5 แสดงถึงลำดับการแสดงของภาพ ลูกศรแสดงความสัมพันธ์ของ P-Picture และ B-Picture เนื่องจากแต่ละภาพจะมีความเกี่ยวเนื่องกัน ลำดับภาพที่ส่ง เก็บ หรือรับมาอาจจะไม่ใช่ลำดับเดียวกับที่แสดงผล ที่เป็นเช่นนี้เพราะว่า B-Picture ขึ้นกับ I-Picture และ P-Picture โดยที่ I-Picture และ P-Picture จะต้องถูกส่งและถอดรหัสก่อนหน้า B-Picture ตัวอย่างลำดับของรูปที่แสดงจะเป็น

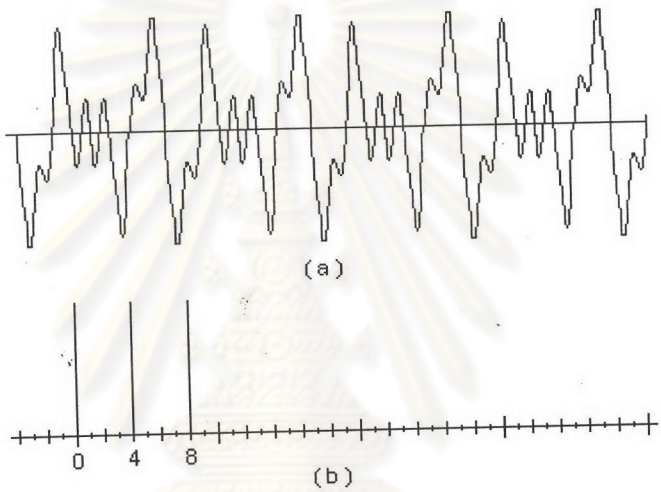
I	B	B	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B	P	B	B
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

ในขณะที่ลำดับของภาพที่ส่งไปถอดรหัสจะเป็น

I	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B	P	B	B	P	B
0	3	1	2	6	4	5	9	7	8	12	10	11	15	13	14	18	16	17	20	19

ดังนั้นตัวถอดรหัสจะต้องเปลี่ยนแปลงลำดับของภาพที่ได้รับให้ถูกต้องก่อนนำไปแสดง

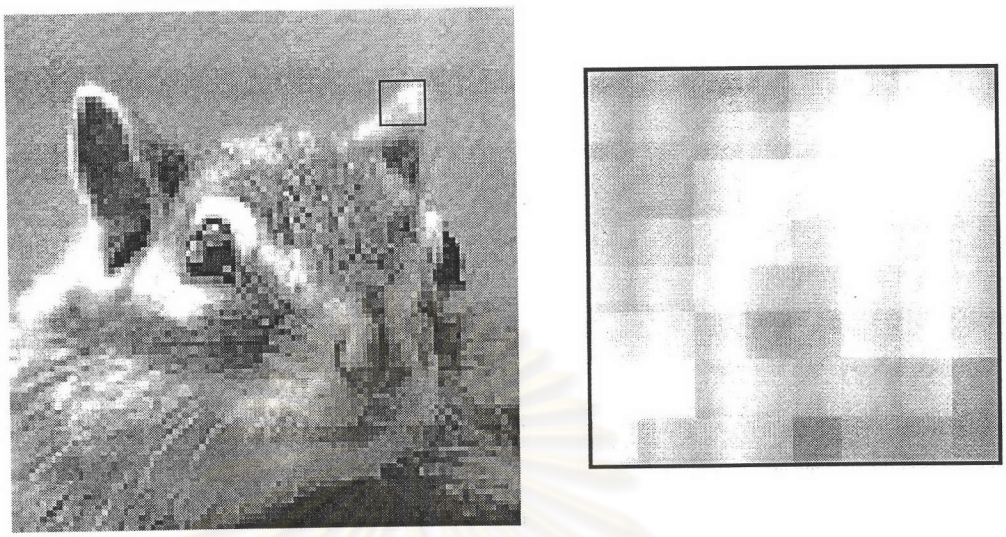
กระบวนการต่อไปที่จะกล่าวถึงเป็นส่วนสำคัญที่สุด เพราะกระบวนการนี้จะใช้การคำนวณมากที่สุดคือเทคนิคการทำ DCT (Discrete Cosine Transform) ซึ่งเป็นกระบวนการทางคณิตศาสตร์ที่คล้ายกับ FFT (Fast Fourier Transform) เนื่องจากเป็นการเอาสัญญาณมาแปลงจากรูปแบบหนึ่งเป็นอีกรูปแบบหนึ่งใช้มากกับการประมวลผลข้อมูลดิจิตอล การแปลงค่าด้วยเทคนิค DCT จะเริ่มจากการรวบรวมค่าซักตัวอย่างที่ได้จากสัญญาณเสียงหรือภาพที่เป็นสัญญาณแอนะล็อก นำมาแสดงเทียบกับแกนเวลา แต่ละจุดบนแกนเวลาจะมีค่าที่แสดงถึงระดับโวลเตจ จากนั้นจะทำการแปลงกลุ่มของจุดเหล่านี้เป็นรูปค่าของความถี่ ที่สามารถอธิบายถึงสัญญาณเดิมที่เข้ามา



รูปที่ 2.6 ตัวอย่าง (a) กราฟ (b) สเปกตัมของกราฟรูป (a)

รูปที่ 2.6 (a) แสดงถึงสัญญาณแอนะล็อกซึ่งประกอบด้วยคลื่นไซน์ที่มีความถี่แตกต่างกัน 3 ความถี่ รวมเป็นสัญญาณเพียงคลื่นเดียวซึ่งมีลักษณะค่อนข้างซับซ้อน แต่ละจุดบนแกนนอนแสดงถึงจุดเวลาของค่าตัวอย่าง ส่วนจุดบนแกนตั้งแสดงถึงโวลเตจหรือแอมพลิจูดของสัญญาณที่เวลาขณะนั้น หากเก็บข้อมูลโดยการบันทึกค่าแอมพลิจูดทุกช่วงเวลาขนาดข้อมูลที่เก็บจะมีปริมาณมาก รูปที่ 2.6 (b) แสดงให้เห็นว่า หลังจากการทำแปลงฟูริเยร์ แต่ละจุดบนแกนนอนจะหมายถึงค่าของความถี่ (frequency) ส่วนจุดบนแกนตั้งแสดงถึงค่าแอมพลิจูดหรือขนาดที่ความถี่นั้น จะเห็นได้ว่าเราสามารถที่จะแสดงรูปของคลื่นที่ได้จากค่าทั้ง 3 ความถี่นี้กลับไปเป็นดังรูป 2.6 (a) ประโยชน์ของการแปลงนี้คือเราสามารถที่จะใช้ข้อมูลเพียง 3 ค่านี้แทนข้อมูลตัวอย่างทั้งหมดได้ทำให้ประหยัดเนื้อที่การเก็บข้อมูลได้มาก

MPEG ใช้เทคนิค DCT แปลงบล็อกเล็กๆ ขนาด 8x8 จุดของข้อมูลภาพให้เป็นเมตริกซ์ 8x8 ของค่าสัมประสิทธิ์ความถี่ทั้งแนวตั้งและแนวนอน โดยเราสามารถใช้กระบวนการแปลงกลับคือการทำ IDCT (Inverse Discrete Cosine Transform) ที่จะแปลงสัมประสิทธิ์ของค่าความถี่กลับมาเป็นบล็อก 8x8 จุดของข้อมูลภาพโดยข้อมูลไม่มีการสูญหาย สัมประสิทธิ์ DCT ที่อยู่ในตำแหน่ง (0,0) (มุมบนซ้าย) ของบล็อกแสดงถึงความถี่ที่เป็นศูนย์ทั้งแนวตั้งและแนวนอน เรียกว่าสัมประสิทธิ์ดีซี โดยสัมประสิทธิ์ดีซีจะเป็นสัดส่วนกับค่าเฉลี่ยของข้อมูล



รูปที่ 2.7 บล็อกเล็ก ๆ ขนาด 8x8 ของข้อมูลภาพ

ภาพในบล็อก เพราะว่าที่จุดนี้ทั้งแกน x และ y มีค่าความถี่เป็นศูนย์ ค่าสัมประสิทธิ์ดีซี นี้เกือบจะเป็นค่าที่มากที่สุด ค่าสัมประสิทธิ์อื่นๆจะถูกเรียกว่าสัมประสิทธิ์เอซี ที่ตำแหน่งยิ่งห่างไกลออกไปจากสัมประสิทธิ์ดีซีมากเท่าไร ค่าก็จะต่ำลง ตัวอย่างต่อไปแสดงผลลัพธ์ที่ได้หลังการทำ DCT โดยข้อมูลเข้าเป็นภาพขาว-ดำ (gray-scale) ค่าของข้อมูลอยู่ระหว่าง 140 -175 เมื่อผ่านการ DCT แล้วสร้างเป็นเมตริกซ์ ได้ดังนี้

เมตริกซ์ของข้อมูลเข้า

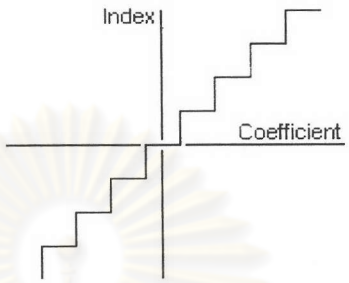
$$\begin{bmatrix} 140 & 144 & 147 & 140 & 140 & 155 & 179 & 175 \\ 144 & 152 & 140 & 147 & 140 & 148 & 167 & 179 \\ 152 & 155 & 136 & 167 & 163 & 162 & 152 & 172 \\ 168 & 145 & 156 & 160 & 152 & 155 & 136 & 160 \\ 162 & 148 & 156 & 148 & 140 & 136 & 147 & 162 \\ 147 & 167 & 140 & 155 & 155 & 140 & 136 & 162 \\ 136 & 156 & 123 & 167 & 162 & 144 & 140 & 147 \\ 148 & 155 & 136 & 155 & 152 & 147 & 147 & 136 \end{bmatrix}$$

เมตริกซ์ของสัมประสิทธิ์ DCT

$$\begin{bmatrix} 186 & -18 & 15 & -9 & 23 & -9 & -14 & 19 \\ 21 & -34 & 26 & -9 & -11 & 11 & 14 & 7 \\ -10 & -24 & -2 & 6 & -18 & 3 & -20 & -1 \\ -8 & -5 & 14 & -15 & -8 & -3 & -3 & 8 \\ -3 & 10 & 8 & 1 & -11 & 18 & 18 & 15 \\ 4 & -2 & -18 & 8 & 8 & -4 & 1 & -7 \\ 9 & 1 & -3 & 4 & -1 & -7 & -1 & -2 \\ 0 & -8 & -2 & 2 & 1 & 4 & -6 & 0 \end{bmatrix}$$

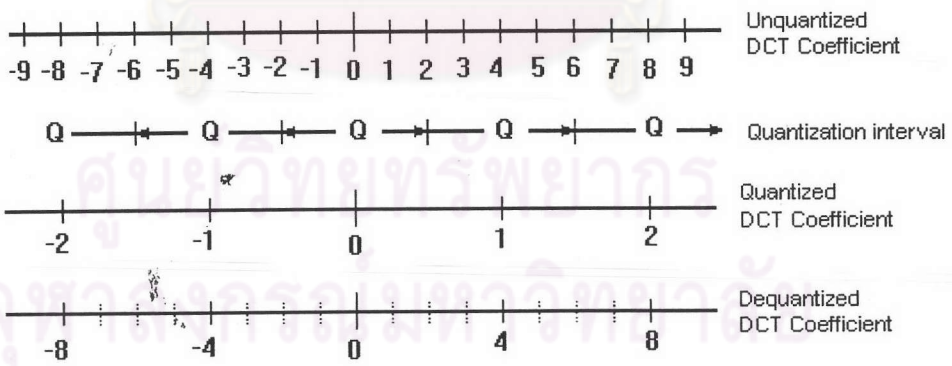
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

กระบวนการถัดจาก DCT คือการควอนไทซ์ (quantization) ซึ่งเป็นการแทนค่าของชุดของข้อมูลที่มีค่าใกล้เคียงกันด้วยค่าค่าเดียว อาจเทียบได้กับการปัดเศษให้เป็นเลขจำนวนเต็ม เพื่อลดจำนวนบิตที่เก็บข้อมูล ดังแสดงในรูปที่ 2.8



รูปที่ 2.8 ลักษณะของการควอนไทซ์

ค่าสัมประสิทธิ์ที่ได้จากการทำ DCT ที่ความถี่สูงทั้งแนวตั้งและแนวนอน ยิ่งความถี่สูงมากค่าจะใกล้เคียงศูนย์มาก และเพื่อที่จะทำให้มีค่าเป็นศูนย์เป็นจำนวนมากจะต้องใช้การแทนค่า วิธีการก็คือนำค่าควอนไทซ์ไปหารทุกๆค่าและทำการปัดเศษทิ้ง รูปที่ 2.9 แสดงถึงตัวอย่างของการควอนไทซ์ ค่าคงที่คือค่า 4 โดยค่ากลางคือ 2 จะถูกบวกเข้าไปก่อนที่จะหารด้วย 4 หลังจากนั้นก็ปัดเศษทิ้งไป กระบวนการที่จะได้ค่าสัมประสิทธิ์คืนมาคือการแปลงกลับควอนไทซ์ (dequantization) โดยการคูณค่าคงที่นั้นกลับเข้าไป แต่ค่าที่ได้จะไม่เหมือนเดิม โดยคุณภาพของภาพจะเป็นสัดส่วนผกผันกับค่าคงที่นี้ ถ้าต้องการให้อัตราการอัดข้อมูลสูงค่าคงที่นี้ควรมีขนาดใหญ่ แต่โดยปกติถ้าค่าคงที่นี้มีขนาดเล็กสายตาของมนุษย์จะไม่สามารถแยกความแตกต่างระหว่างสัมประสิทธิ์ที่ผ่านการควอนไทซ์ (dequantized DCT coefficient) กับค่าที่ไม่ได้ควอนไทซ์ (unquantized DCT coefficient)



รูปที่ 2.9 การควอนไทซ์ และ การแปลงกลับ

โดยปกติผลลัพธ์ของเมตริกซ์สัมประสิทธิ์ DCT จะใช้เนื้อที่มากกว่าข้อมูลภาพเดิมเนื่องจาก ข้อมูลภาพจะเป็นค่าข้อมูล 8 บิต แต่หลังจากแปลงเป็นสัมประสิทธิ์ DCT แล้วค่าจะอยู่ระหว่าง -1,024 ถึง 1,023 ต้องใช้ถึง 11 บิต การควอนไทซ์จะลดจำนวนบิตของการเก็บข้อมูลได้ เราสามารถที่จะลดการเก็บสัมประสิทธิ์ได้มากขึ้นโดยตำแหน่งของสัมประสิทธิ์ที่ห่างไกลจากสัมประสิทธิ์ดีซีมากจะมีความสำคัญน้อย MPEG จะใช้เมตริกซ์การควอนไทซ์ กระทำกับทุกๆตำแหน่งของค่าบนเมตริกซ์สัมประสิทธิ์ DCT ค่าที่อยู่ในเมตริกซ์การควอนไทซ์ คือค่า

ควอนตัม (quantum) ซึ่งแสดงถึงขนาดของขั้น (step size) ที่จะกระทำกับจุดต่างๆบนภาพ ค่าควอนตัมจะอยู่ระหว่าง 1 ถึง 255 สำหรับจุดที่มีความสำคัญมากก็จะถูกเข้ารหัสด้วยค่าที่ต่ำ ค่าควอนตัมจะเพิ่มมากขึ้นเมื่อตำแหน่งนั้นห่างไกลออกมาจากจุดเริ่มต้น (0,0) โดยแสดงสมการได้ดังนี้

$$Quantized\ Value\ (i,j) = \frac{dct(i,j)}{Quantum(i,j)} Rounded\ to\ nearest\ integer \quad \ldots\ldots\ldots\ldots\ (1)$$

จากสมการจะเห็นได้ว่าสัมประสิทธิ์ที่อยู่ที่ความถี่สูงจะถูกปัดเป็นศูนย์ไป ยกเว้นค่าสัมประสิทธิ์ที่มีค่าสูงเกินปกติถึงจะไม่เป็นศูนย์ ระหว่างการถอดรหัส สมการการแปลงกลับควอนไทซ์ จะเป็นดังนี้

$$DCT(i,j) = Quantized\ Value(i,j) * Quantum(i,j) \quad \ldots\ldots\ldots\ldots\ (2)$$

เมื่อเราใช้ค่าควอนตัมที่มีขนาดใหญ่จะทำให้เกิดผิดพลาดสูงตอนที่ทำการถอดรหัส แต่ถ้าข้อผิดพลาดนั้นอยู่ที่ความถี่สูงก็จะไม่มีผลกระทบกับคุณภาพของภาพ การเลือกขนาดการเพิ่มของควอนตัมสูง (high step size) สำหรับสัมประสิทธิ์ DCT ส่วนใหญ่แล้วจะได้อัตราการอัดข้อมูลที่สูง แต่คุณภาพของภาพก็จะไม่ดีนัก ตรงกันข้ามหากเราเลือกขนาดการเพิ่มของควอนตัมต่ำ (low step size) อัตราการอัดข้อมูลที่ได้ก็ไม่ดี แต่คุณภาพของภาพก็จะดี เพื่อเป็นการยืดหยุ่นโดยให้ผู้ใช้สามารถที่จะตัดสินใจเลือกระหว่างอัตราการอัดข้อมูลหรือคุณภาพ โดยสามารถกำหนดตัวประกอบคุณภาพ (quality factor) ซึ่งมีค่าระหว่าง 1 ถึง 25

ตัวอย่างเมตริกซ์การควอนไทซ์ ที่ได้เมื่อตัวประกอบคุณภาพเป็น 2

$$\begin{bmatrix} 3 & 5 & 7 & 9 & 11 & 13 & 15 & 17 \\ 5 & 7 & 9 & 11 & 13 & 15 & 17 & 19 \\ 7 & 9 & 11 & 13 & 15 & 17 & 19 & 21 \\ 9 & 11 & 13 & 15 & 17 & 19 & 21 & 23 \\ 11 & 13 & 15 & 17 & 19 & 21 & 23 & 25 \\ 13 & 15 & 17 & 19 & 21 & 23 & 25 & 27 \\ 15 & 17 & 19 & 21 & 23 & 25 & 27 & 29 \\ 17 & 19 & 21 & 23 & 25 & 27 & 29 & 31 \end{bmatrix}$$

สัมประสิทธิ์ DCT ก่อนทำการควอนไทซ์

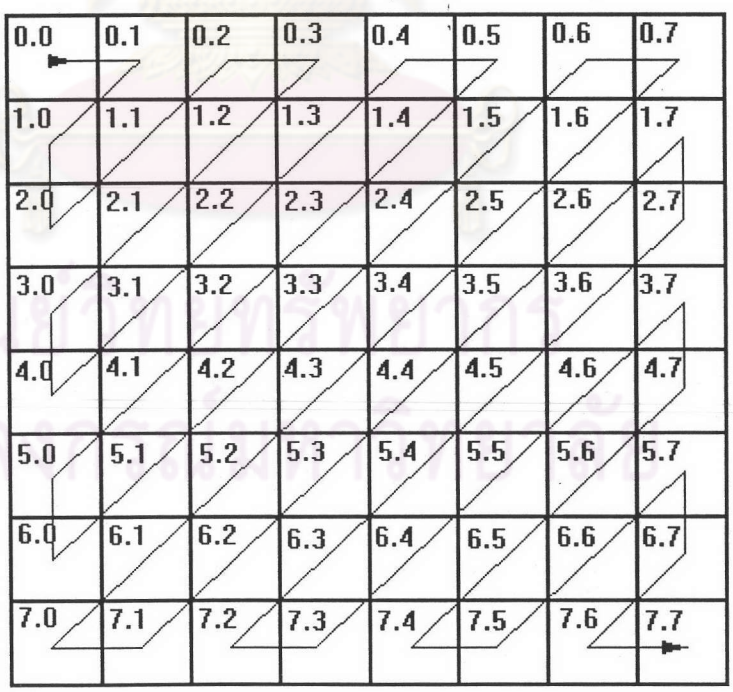
$$\begin{bmatrix} 92 & 3 & -9 & -7 & 3 & -1 & 0 & 2 \\ -39 & -58 & 12 & 17 & -2 & 2 & 4 & 2 \\ -84 & 62 & 1 & -18 & 3 & 4 & -5 & 5 \\ -52 & -36 & -10 & 14 & -10 & 4 & -2 & 0 \\ -86 & -40 & 49 & -7 & 17 & -6 & -2 & 5 \\ -62 & 65 & -12 & -2 & 3 & -8 & -2 & 0 \\ -17 & 14 & -36 & 17 & -11 & 3 & 3 & -1 \\ -54 & 32 & -9 & -9 & 22 & 0 & 1 & 3 \end{bmatrix}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สัมประสิทธิ์ DCT หลังทำการแปลงกลับควอนไทซ์

$$\begin{bmatrix} 90 & 0 & -7 & 0 & 0 & 0 & 0 & 0 \\ -35 & -56 & 9 & 11 & 0 & 0 & 0 & 0 \\ -84 & 54 & 0 & -13 & 0 & 0 & 0 & 0 \\ -45 & -33 & 0 & 0 & 0 & 0 & 0 & 0 \\ -77 & -39 & 45 & 0 & 0 & 0 & 0 & 0 \\ -52 & 60 & 0 & 0 & 0 & 0 & 0 & 0 \\ -15 & 0 & -19 & 0 & 0 & 0 & 0 & 0 \\ -51 & 19 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

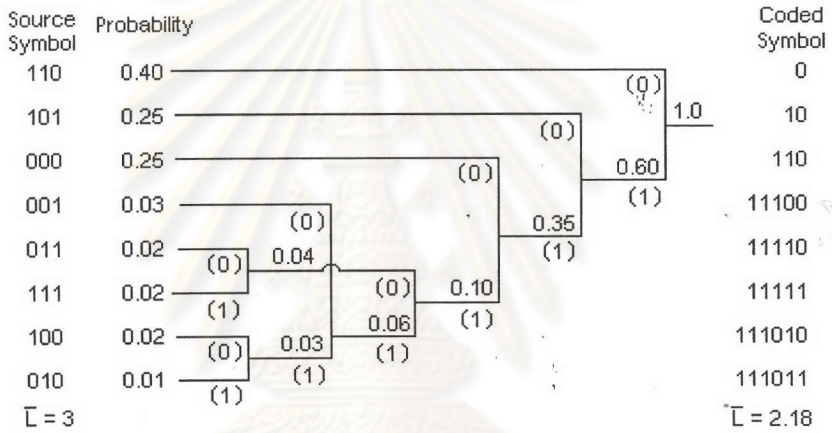
ถัดจากการควอนไทซ์จะสังเกตเห็นว่าข้อมูลจะแยกไปสองเส้นทาง ทางหนึ่งจะผ่านกระบวนการแปลงกลับของ DCT และการแปลงกลับควอนไทซ์ เพื่อนำไปเปรียบเทียบหาค่า motion vector กับภาพอ้างอิงในกรณีที่เป็น P-Picture หรือ B-Picture และส่งเฉพาะ motion vector ไป แต่ถ้าเป็น I-Picture จะใช้เป็นข้อมูลอ้างอิง ในขณะเดียวกับ ข้อมูลที่ผ่านการควอนไทซ์ จะเข้าสู่กระบวนการสุดท้ายของการเข้ารหัสคือ 3 ขั้นตอนของการอัดข้อมูล ขั้นตอนแรกคือการเปลี่ยนค่าสัมประสิทธิ์ดีซีจากค่าสัมบูรณ์ (absolute) เป็นค่าสัมพันธ์ (relative) อ้างอิงกับบล็อกที่มาเปรียบเทียบ ซึ่งส่วนใหญ่จะมีค่าที่ไม่แตกต่างกับบล็อกใกล้เคียงนัก ทำให้ค่าที่เก็บมีขนาดเล็ก ขั้นตอนที่สองคือเก็บค่าสัมประสิทธิ์เอซีอื่นๆในลำดับซิกแซ็ก (zigzag sequence) เส้นทางของซิกแซ็กแสดงไว้ดังรูปที่ 2.10



รูปที่ 2.10 เส้นทางซิกแซ็กของการเก็บข้อมูล

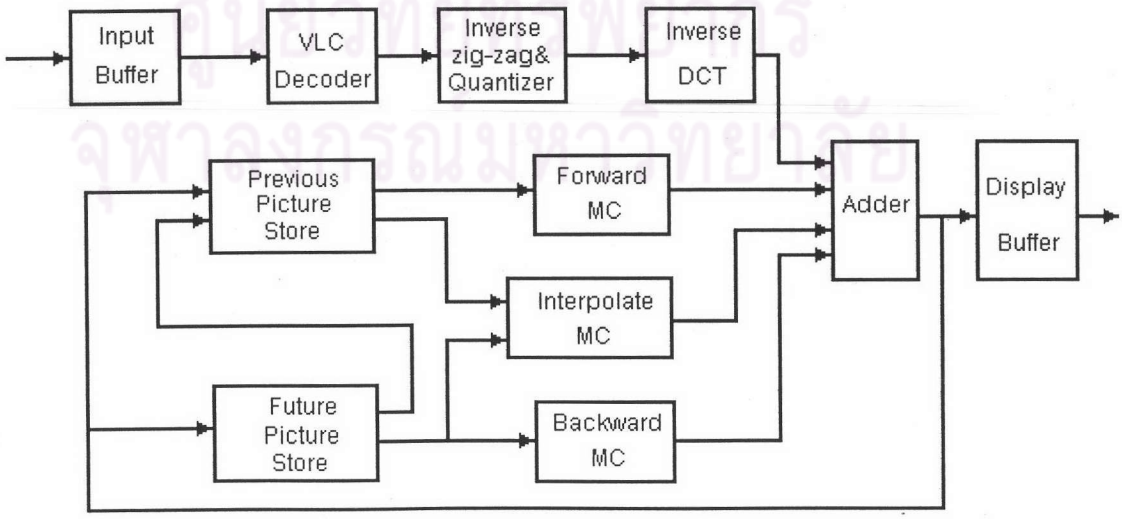
เนื่องจากสัมประสิทธิ์ส่วนใหญ่ภายหลังการควอนไทซ์ถูกปัดเป็นศูนย์ การเก็บข้อมูลจะใช้เทคนิคนับจำนวนค่าที่เป็นศูนย์ (Run-Length Encoding of Zero) การทำให้มีอัตราการอัดข้อมูลสูง โดยวิธีที่จะเพิ่มจำนวนความยาวของเลขศูนย์ ให้จัดเก็บข้อมูลในลำดับซิกแซ็ก แทนที่จะเป็นลำดับแถวแบบปกติ

ขั้นตอนที่สามคือการแทนค่าข้อมูลที่ไม่เป็นศูนย์และจำนวนอื่นๆ โดยใช้การเข้ารหัสที่ความยาวข้อมูลแปรได้ (Variable-Length Coding) เป็นการเข้ารหัสที่ใช้อาศัยการทำสถิติซึ่งจะระบุคำรหัส (codewords) ให้กับค่าที่จะเข้ารหัส ค่าที่เกิดขึ้นบ่อยๆจะใช้คำรหัสที่สั้นและค่าที่เกิดไม่บ่อยจะเป็นคำรหัสที่ยาว โดยปกติคำรหัสที่สั้นจะต้องสั้นกว่าค่าของข้อมูลจริง ตัวอย่างของการเข้ารหัสแบบนี้คือการเข้ารหัสฮัฟฟ์แมน (Huffman Coding) ในรูปที่ 2.11 แสดงการสร้างรหัสสัญลักษณ์(coded symbol) จากค่าของความน่าจะเป็น(probability) ในตัวอย่างค่าเฉลี่ยของจำนวนบิตที่ใช้เหลือเพียง 2.18 ในขณะที่ข้อมูลจริงจะต้องใช้ถึง 3 บิต



รูปที่ 2.11 การเข้ารหัสฮัฟฟ์แมน

2.4 กระบวนการถอดรหัสสัญญาณภาพของ MPEG



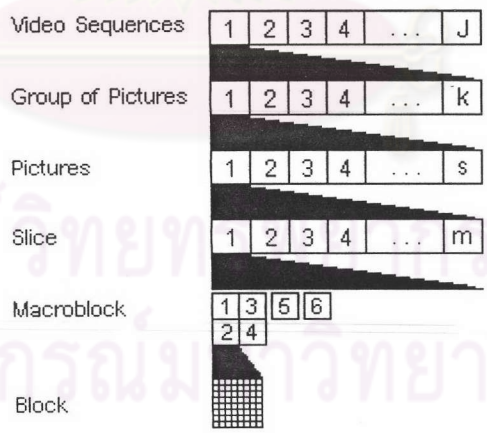
รูปที่ 2.12 แผนภาพบล็อกแสดงการถอดรหัสสัญญาณภาพ

กระบวนการถอดรหัสสัญญาณภาพเป็นกระบวนการย้อนกลับของการเข้ารหัส ดังแสดงในแผนภาพรูปที่ 2.12 ขั้นตอนอย่างคร่าวๆคือ ในขณะที่ตัวถอดรหัสรับข้อมูลในตามลำดับของการส่ง ที่บัฟเฟอร์จะสะสมข้อมูลที่เข้ามาไว้จนกระทั่งมากพอและส่งต่อไปยังตัวถอดรหัส VLC (Variable Length Code) ถ้าชนิดของภาพที่เข้ามาเป็น I-Picture ตัวถอดรหัส VLC จะถอดรหัสต่อไปจนได้สัมประสิทธิ์ควอนไทซ์ (quantized coefficient) ส่วนตัวแปลงกลับควอนไทซ์ (inverse quantizer) จะสร้างสัมประสิทธิ์ DCT โดยการคูณค่าควอนตัมเข้าไป ในที่สุดสัมประสิทธิ์ต่างๆจะถูกแปลงกลับไปเป็นข้อมูลภาพโดย IDCT จากนั้นบล็อกที่ได้จะถูกเก็บอยู่ในที่อยู่ของภาพในอดีต (Previous Picture Store) และที่เก็บของการแสดงภาพ (display buffer) และภาพก็จะถูกนำมาแสดงในเวลาที่เหมาะสมต่อไป

ถ้าภาพที่เข้ามาเป็น P-Picture หรือ B-Picture ในแต่ละบล็อก ตัวถอดรหัส VLC จะแทนที่ motion vector ด้วยค่าบล็อกที่ได้จากภาพที่ถอดรหัสมาก่อนหน้านี้ และถอดรหัสจนได้ สัมประสิทธิ์ควอนไทซ์ ส่วนตัวแปลงกลับควอนไทซ์ จะสร้างสัมประสิทธิ์ DCT ในที่สุดสัมประสิทธิ์ต่างๆจะถูกแปลงกลับไปเป็นข้อมูลภาพ ถ้าภาพนั้นเป็น P-Picture ผลสุดท้ายของบล็อกจะถูกเก็บอยู่ในที่เก็บของภาพในอนาคต (Future Picture Store) และที่เก็บของภาพที่จะแสดงแต่ภาพนี้ยังไม่สามารถที่จะแสดงได้จนกว่า B-Picture ที่อาศัยข้อมูลจาก P-Picture นี้จะได้รับ ถอดรหัส และทำการแสดงไปก่อน

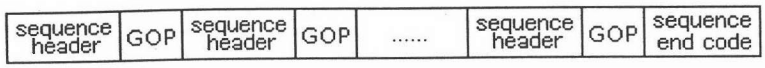
2.5 โครงสร้างกระแสข้อมูลของ MPEG

MPEG ได้แบ่งย่อยข้อมูลออกเป็น 6 ระดับด้วยกัน ดังแสดงในรูปที่ 2.13



รูปที่ 2.13 โครงสร้างกระแสข้อมูลของ MPEG

ลำดับ (Video Sequence) คือ ชั้นที่อยู่บนสุดของการเข้ารหัส ประกอบด้วยส่วนหัว (header) และ GOP (Group Of Pictures) ในส่วนหัวจะมีข้อมูลบอกให้ตัวถอดรหัสเข้าสู่สถานะการเริ่มต้น ทำให้ตัวถอดรหัสสามารถที่จะถอดรหัสลำดับใดๆ โดยไม่มีผลกระทบกับการถอดรหัสของลำดับที่ทำมาแล้วในอดีต ลำดับจะขึ้นต้นด้วยส่วนหัวของลำดับ ตามด้วยกลุ่มของภาพจำนวน 1 กลุ่มขึ้นไป โดยก่อนที่จะขึ้นกลุ่มใหม่ จะต้องมีส่วนหัวของลำดับขึ้นก่อนก่อนทุกครั้งไป และสุดท้ายจบด้วยรหัสสิ้นสุดของลำดับดังแสดงในรูปที่ 2.14



รูปที่ 2.14 ลำดับการส่งข้อมูลของ MPEG

GOP (Group Of Pictures) เป็นกลุ่มของภาพที่เข้ารหัสแล้ว อาจเป็นภาพภาพเดียวหรือหลายภาพก็ได้ โดยที่แต่ละ GOP จะเป็นอิสระต่อกัน GOP ประกอบไปด้วยส่วนหัวซึ่งเก็บข้อมูลเวลาและส่วนที่เป็นภาพ (picture) ภาพแรกในลำดับการส่งจะเป็น I-Picture และภาพสุดท้ายของกลุ่มจะต้องเป็น I-Picture หรือ P-Picture ในแต่ละ GOP จะมีจำนวนกี่ภาพก็ได้ แต่อย่างน้อยต้องมี I-Pioture หนึ่งภาพ

ที่ลำดับการส่งของภาพ และเป็นลำดับภาพในกระแสข้อมูล

1 4 2 3 7 5 6	10 8 9 13 11 12 16 14 15	19 17 18 22 20 21 25 23 24
I P B B P B B	I B B P B B P B B	I B B P B B P B B

ตัวอย่างจากลำดับการแสดงของภาพ

1 2 3 4 5 6 7	8 9 10 11 12 13 14 15 16	17 18 19 20 21 22 23 24 25
I B B P B B P	B B I B B P B B P	B B I B B P B B P

ที่เป็นขีดแสดงถึงขอบเขตของ GOP สังเกตว่าที่กลุ่มแรกจะมีภาพน้อยกว่ากลุ่มอื่น 2 ภาพ โดยไม่มี B-Picture นำ I-Picture อย่างไรก็ดีตามปกติในลำดับของการแสดงอาจจะมี B-Picture นำหน้า I-Picture ได้ถึงแม้ว่าจะเป็นกลุ่มแรกก็ตาม เมื่อถอดรหัสออกมาแล้วจะได้ผลลัพธ์ดังนี้

1 2 3 4 5 6 7	8 9 10 11 12 13 14 15 16	17 18 19 20 21 22 23 24 25

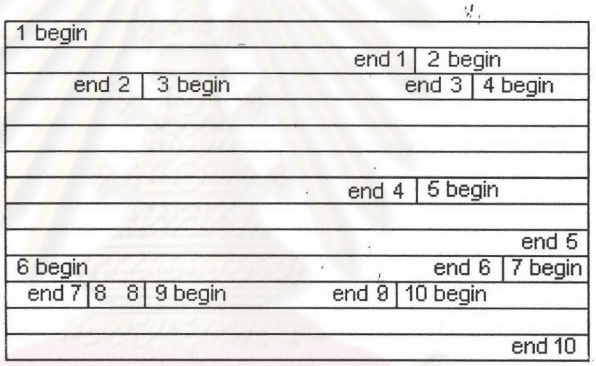
ภาพ (Picture) เป็นภาพเดี่ยวๆ ของภาพเคลื่อนไหวหรือภาพยนตร์ แบ่งออกเป็น 4 ประเภท

1) *I-Picture* หรือ Intra Coded Picture ซึ่งจะเข้าและถอดรหัสโดยไม่เกี่ยวข้องกับภาพอื่นๆ
2) *P-Picture* หรือ Predictive Coded Picture ซึ่งเข้าและถอดรหัสโดยใช้การชดเชยการเคลื่อนที่จาก I-Picture หรือ P-Picture ภาพในอดีต
3) *B-Picture* หรือ Bidirectionaly Predictive Coded Picture ซึ่งเข้าและถอดรหัสโดยใช้การชดเชยการเคลื่อนที่ จาก I-Picture หรือ P-Picture ซึ่งเป็นภาพในอดีตและภาพในอนาคต
4) *D Picture* เพื่อใช้ในโหมดการค้นหาภาพอย่างเร็วเท่านั้น

โดยทั่วไปการเข้าและถอดรหัสจะใช้การผสมของ I-Picture P-Picture และ B-Picture I-Picture จะเกิดขึ้นทุกๆ ครึ่งวินาที และ 2 B-Picture คั่นระหว่าง I-Picture หรือ P-Picture รูปภาพประกอบด้วยส่วนหัวกับสไลซ์ ส่วนหัวของรูปภาพจะเก็บเวลา ชนิดของภาพ และข่าวสารของการเข้ารหัส

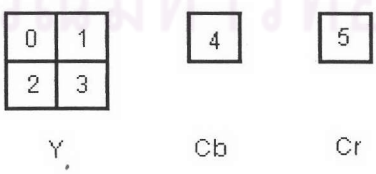
สไลซ์ (Slice) ภาพจะถูกตัดแบ่งออกเป็นสไลซ์ โดยที่สไลซ์คือกลุ่มของแมโครบล็อก เริ่มจากมุมบนซ้ายของภาพและกวาดตรวจจากซ้ายไปขวาและจากบนลงล่าง แมโครบล็อกที่จุดเริ่มและสิ้นสุดของสไลซ์จะต้องเป็นแมโครบล็อกที่มีข้อมูล ทุกๆสไลซ์จะต้องมีอย่างน้อย 1 แมโครบล็อก ใน 1 ภาพสไลซ์ไม่จำเป็นจะต้องมีขนาดเท่ากันทุกสไลซ์ แต่ต้องไม่มีการซ้อนทับกันหรือมีช่องว่างระหว่างสไลซ์ การแบ่งภาพออกเป็นสไลซ์เพื่อประโยชน์ในการส่งข้อมูล หากเกิดผิดพลาดในการส่งข้อมูล ตัวถอดรหัสจะสามารถดึงข้อมูลถัดไปขึ้นมาได้โดยการรอสไลซ์ถัดไปเพื่อที่จะไม่ต้องหยุดแสดงไปทั้งภาพ เพราะสไลซ์สามารถที่จะทำการถอดรหัสโดยอาศัยข้อมูลภายในสไลซ์เองไม่ต้องใช้ข้อมูลจากสไลซ์อื่น ถ้าการส่งข้อมูลเกิดในภาวะที่จะไม่เกิดความผิดพลาดขึ้นเลย ก็สามารถที่จะแบ่งให้ 1 ภาพเป็น 1 สไลซ์ก็ได้เพราะถ้าแบ่งภาพออกเป็นสไลซ์มากๆ ก็จะต้องเสียเนื้อที่เก็บข้อมูลเกี่ยวกับสไลซ์เข้าเพิ่มขึ้น

สไลซ์ประกอบด้วยส่วนหัวกับแมโครบล็อก ที่จุดเริ่มของแต่ละสไลซ์ค่าสัมประสิทธิ์ดีซีและ motion vector จะถูกตั้งใหม่ ส่วนของหัวของสไลซ์จะมีตำแหน่งและข่าวสารอื่นๆ ซึ่งพอเพียงสำหรับการแสดงข้อมูล



รูปที่ 2.15 ตัวอย่างการจัดแบ่งออกเป็นสไลซ์ที่เป็นไปได้

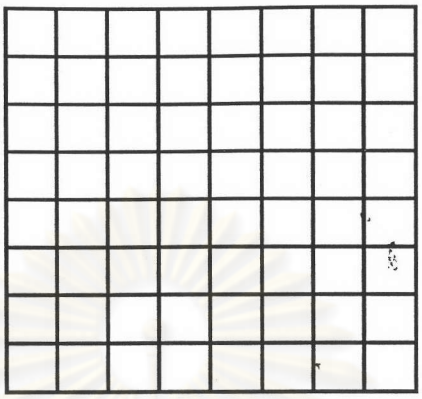
แมโครบล็อก (Macroblock) เป็นพื้นฐานของการทำค่าชดเชยการเคลื่อนที่ และการตั้งค่าขนาดการควอนไทซ์ แต่ละแมโครบล็อกประกอบไปด้วยส่วนหัวและบล็อกของค่าความสว่าง 4 บล็อก บล็อกของสี 2 บล็อก (Cr และ Cb) การเรียงตำแหน่งดังแสดงในรูปที่ 2.16 ส่วนหัวของแมโครบล็อกประกอบด้วยข้อมูลการควอนไทซ์ และค่าชดเชยการเคลื่อนที่



รูปที่ 2.16 โครงสร้างของแมโครบล็อก

บล็อก (Block) เป็นหน่วยย่อยที่เล็กที่สุดและ DCT/IDCT จะทำการอัดข้อมูลที่ระดับบล็อก แต่ละบล็อกจะประกอบไปด้วย 64 จุดข้อมูลมีจำนวนจุดตามแนวตั้งและแนวนอน 8 จุด ดังแสดงในรูปที่ 2.17 ข้อมูล

ของบล็อกอาจจะเป็นค่าข้อมูลที่แสดงความสว่าง (luminance) หรือ อาจจะเป็นค่าของข้อมูลที่แสดงสี (chrominance) ซึ่งแต่ละจุดจะมีข้อมูลแสดงความสว่าง 1 ค่า และค่าของสี 2 ค่า (Cr, Cb)



รูปที่ 2.17 โครงสร้างของบล็อก

2.6 สรุป

กล่าวโดยสรุป MPEG ได้ใช้เทคนิคหลายๆอย่างผสมผสานกัน ในการอัดข้อมูล ไม่ว่าจะเป็นการเก็บข้อมูลโดยการแปลงข้อมูลให้อยู่ในรูปสัมประสิทธิ์ความถี่และด้วยการควอนไทซ์ทำให้ข้อมูลกลายเป็นศูนย์จำนวนมาก โดยปกติแล้วการทำ DCT นั้นข้อมูลจะไม่มีการสูญหายหรือเปลี่ยนแปลง ยกเว้นการควอนไทซ์ที่ทำให้ข้อมูลบางส่วนหายไป แต่ MPEG มีตัวแปรที่เรียกว่า ตัวประกอบคุณภาพ (quality factor) ซึ่งผู้ใช้สามารถกำหนดได้เอง ว่าจะเลือกเอาคุณภาพของภาพหรืออัตราการอัดข้อมูลที่สูง นอกจากนี้การสร้างค่าชดเชยการเคลื่อนที่โดยการเก็บเฉพาะ motion vector นั้นลดจำนวนข้อมูลได้มาก เนื่องจากข้อมูลที่เป็นภาพยนตร์นั้นแต่ละภาพจะมีส่วนแตกต่างกันน้อยมาก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย