

Introduction to Digital Television

1. บทนำ
2. เทคโนโลยีของการแพร่ภาพโทรทัศน์แบบอนาล็อก
 - 2.1 การ Scanning
 - 2.2 จอภาพแบบ CRT , LCD และ PDP
 - 2.3 สีและการมองเห็น
 - 2.4 Image Sensor : CCD vs CMOS
 - 2.5 Aspect Ratio (อัตราส่วนของภาพ)
 - 2.6 Bandwidth and Definition
 - 2.7 Video Signal
 - 2.8 การแพร่ภาพโทรทัศน์แบบอนาล็อก
 - 2.9 การตรวจวัดสัญญาณ (Measurement) โทรทัศน์อนาล็อก
3. ITU-R BT. 601 Digital Video Signal
4. Transform to and from Frequency Domain
5. Video Coding
 - 5.1 Video Compression
 - 5.2 Evolution of the Compression Standard
 - 5.3 MPEG -2 Video Coding
 - 5.4 MPEG -4 Video Coding
 - 5.5 H.264 / AVC Video Coding
6. Audio Coding
 - 6.1 MPEG -1 Audio Coding
 - 6.2 MPEG -2 Audio Coding
 - 6.3 MPEG -4 Audio Coding
7. Comparison of Digital Video Standard
8. MPEG Data Stream
9. Digital Modulation Basic
10. Digital Television Transmission Scenario
11. Digital Terrestrial Television Standard

1. บทนำ

เป็นเวลาหลายสิบปีแล้ว ที่การส่งโทรทัศน์และการส่งข้อมูล (data) พัฒนาคู่ขนานกันมาพร้อม ๆ กัน แต่แยกเส้นทางในการพัฒนากันอย่างสมบูรณ์ แม้ว่าในยุคปี ค.ศ.1980 จะมีการใช้จอเครื่องรับโทรทัศน์เป็นมอนิเตอร์ของเครื่องคอมพิวเตอร์ในบ้าน นี่คือการครั้งแรกที่เป็นการทำงานร่วมกันของสองกิจการ ปัจจุบันเป็นการยากอย่างยิ่งที่จะแยกสื่อทั้งสองชนิดคือโทรทัศน์ และคอมพิวเตอร์ ซึ่งหลอมรวมกันมากขึ้น ในยุคของสื่อประสม (Multimedia)

จะเห็นได้ว่าปัจจุบันมีการใช้ TV สำหรับเครื่อง PC เพื่อเปลี่ยนแปลง PC ให้สามารถรับสัญญาณโทรทัศน์เพื่อแสดงผลบนจอ PC ได้เป็นอย่างดี เครื่อง PC จึงกลายเป็นโทรทัศน์ อีกเครื่องหนึ่ง ในทางกลับกันมีการนำเสนอบริการ Teletext ตั้งแต่ยุคปี ค.ศ.1980 เพื่อเป็นสื่อที่นำเสนอข้อมูลดิจิทัลผ่านจอเครื่องรับโทรทัศน์แบบอนาล็อก

บัดนี้เราได้อาศัยอยู่ในยุคของโทรทัศน์ระบบดิจิทัล ที่เริ่มมาตั้งแต่ ปี ค.ศ.1995 ซึ่งโดยความจริงแล้วเสมือนว่าความแตกต่างระหว่างข้อมูล (data) และโทรทัศน์ จะหายไป เมื่อสามารถพบทั้งสองสื่อคือสื่อสารข้อมูล (data) และโทรทัศน์รวมกันในสัญญาณข้อมูล (Data Signal) เดียวกัน หรือในการบริการข้อมูลเดียวกัน เช่น การเข้าถึง Internet ความเร็วสูง ผ่านช่องสัญญาณที่เตรียมไว้สำหรับกิจการโทรทัศน์ดิจิทัล ปัจจุบันร่วมกันนำไปสู่การหลอมละลาย นี้ คือความต้องการอัตราข้อมูลความเร็วสูง ดังเช่นที่ใช้ในการบริการโทรทัศน์ดิจิทัล มาใช้ในกิจการโทรคมนาคม เพราะว่าเมื่อเปรียบเทียบกิจการ โทรศัพท์เคลื่อนที่ระบบ GSM ที่มีอัตราข้อมูล 9600 bit/sec และ UMTS มีอัตราข้อมูล 2 M bit/sec หรือช่องสัญญาณโทรศัพท์พื้นฐาน ISDN เท่ากับ 2 เท่าของ 64 K bit/sec กับอัตราข้อมูลของโทรทัศน์ความชัดเจนปกติ (Standard Definition TV : SDTV) ที่ไม่บีบอัดสัญญาณคือ 270 M bit/sec และโทรทัศน์ความชัดเจนสูง (High Definition TV : HDTV) ที่ไม่บีบอัดสัญญาณประมาณ 1 G bit/sec แล้วจะเห็นว่าต่างกันมาก สามารถเรียกกิจการโทรทัศน์ ว่าเป็นเทคโนโลยี แบบบรอดแบนด์ เพราะไม่ว่ากิจการโทรทัศน์ดิจิทัลหรือโทรทัศน์อนาล็อก ที่มีความกว้างช่องสัญญาณ 6,7 และ 8 MHz หรือโทรทัศน์ผ่านดาวเทียม ที่มีความกว้างช่องสัญญาณ 36 MHz

รากฐานของโทรทัศน์อนาล็อกเริ่มตั้งแต่ปี 1883 แต่มีการให้บริการจริง ๆ แบบ ขาว-ดำ ในยุค ปี 1950 สำหรับประเทศไทยเริ่มให้บริการโทรทัศน์อนาล็อกแบบ ขาว-ดำ เมื่อวันที่ 24 มิถุนายน 2498 และพัฒนาเป็นโทรทัศน์สีอนาล็อก ระบบ CCIR-PAL 625 /50 ในปี 2510

ได้มีการพัฒนาเทคโนโลยีการให้บริการโทรทัศน์อนาล็อก เพื่อให้มีคุณภาพดียิ่งขึ้น จอกว้างมากขึ้น ทั้งในทวีปยุโรป ที่ได้พัฒนาเป็นระบบ D2 MAC (Multiplex Analog Component) หรือในประเทศญี่ปุ่นและสหรัฐอเมริกา ที่พยายามพัฒนาไปสู่ โทรทัศน์ความชัดเจนสูง HDTV ระบบอนาล็อก เช่น MUSE HDTV ในประเทศญี่ปุ่น แต่ทั้งหมดก็ไม่ประสบความสำเร็จในแง่เทคโนโลยีและความนิยมในการให้บริการ

การผลิตรายการ โทรทัศน์ในห้องส่งเริ่มนำเอาสัญญาณโทรทัศน์แบบดิจิทัล มาใช้ตั้งแต่ปี 1990 สัญญาณโทรทัศน์แบบดิจิทัล ดังกล่าวเป็นไปตามมาตรฐาน “ CCIR 601 ” หรือในปัจจุบัน คือ ITU-BT R601 มีอัตราข้อมูล 270 Mbit/sec สัญญาณโทรทัศน์ดิจิทัล ดังกล่าวเหมาะสำหรับการผลิตรายการ เท่านั้น แต่ไม่เหมาะในการการบันทึกจัดเก็บและ ส่งผ่านสื่อต่างๆ ทั้งผ่านดาวเทียม เคเบิลทีวี หรือโทรทัศน์ภาคพื้นดิน ไปยังผู้รับสัญญาณปลายทาง เนื่องจากมีข้อมูลมากเกินไป ไม่สามารถส่งผ่านสื่อได้ จำเป็นต้องบีบอัดสัญญาณ (Compression) ให้มีข้อมูลเล็กลง เหมาะสมกับขนาดช่องสัญญาณ

เหตุการณ์ที่สำคัญในแวดวงโทรทัศน์ดิจิทัล ที่สมควรบันทึกไว้ คือการกำหนดมาตรฐานการบีบอัดสัญญาณ JPEG ที่ย่อมาจาก Joint Photographic Expert Group มีความเชี่ยวชาญการบีบอัดสัญญาณภาพนิ่ง โดยใช้เทคนิค DCT : Discrete Cosine Transform ซึ่งยังคงได้รับความนิยมจนกระทั่งปัจจุบันนี้

DCT ยังคงเป็นขั้นตอนพื้นฐานสำหรับกลุ่ม MPEG : Moving Picture Expert Group ในการพัฒนาการบีบอัดสัญญาณแบบ MPEG-1 ในปี 1993 และมาตรฐาน MPEG-2 ในปี 1995 มาตรฐาน MPEG-1 มีเป้าหมายขนาดข้อมูลประมาณ 1.5 Mbit/sec สำหรับใช้บันทึกข้อมูลบน CD ส่วน MPEG-2 มีเป้าหมายขนาดข้อมูลสูงกว่าเหมาะสำหรับกิจการโทรทัศน์ดิจิทัล SDTV และ HDTV ได้รับมาตรฐาน ISO สำหรับอธิบายโครงสร้างข้อมูลคือ ISO / IEC 13818-1 และ ISO / IEC 13818-2 เป็นวิธีการบีบอัดสัญญาณภาพเคลื่อนไหว ส่วน และมาตรฐาน ISO / IEC 13818-3 เป็นวิธีการบีบอัดสัญญาณเสียง MPEG-2 ปัจจุบันมีการใช้ มาตรฐาน MPEG-2 แพร่หลายทั่วโลก เพื่อบีบอัดขนาดสัญญาณภาพโทรทัศน์ดิจิทัล จาก 270 Mbit/sec ลงมาเหลือเพียง 2-6 Mbit/sec และบีบอัดขนาดสัญญาณเสียง จาก 1.5 Mbit/sec เหลือเพียง 100-400 Kbit/sec เพื่อให้เหมาะสมกับขนาดช่องสัญญาณ TV ขนาด 8 MHz

ในเวลาเดียวกันก็ได้มีการพัฒนาวิธีการบีบอัดสัญญาณ ให้มีประสิทธิภาพดีขึ้น เช่นคุณภาพเท่ากัน แต่ขนาดข้อมูลน้อยกว่า ได้แก่มาตรฐาน MPEG-4 , VC1, Divac (VC2) เป็นต้น

ช่วงต้นของยุคปี 1990 เริ่มมีการบริการโทรทัศน์ดิจิทัลผ่านดาวเทียม และ บริการโทรทัศน์ดิจิทัลผ่านเคเบิล หรือที่เรียกว่าดิจิทัลเคเบิลทีวี ก่อนที่จะพัฒนาสู่การบริการโทรทัศน์ดิจิทัลภาคพื้นดิน ในปี 1998 ทั้งในทวีปยุโรปและประเทศสหรัฐอเมริกา โดยอาศัยเทคโนโลยีการบีบอัดสัญญาณ MPEG-2

ในปี 1999 ได้เริ่มมีการบันทึกสัญญาณภาพและเสียงแบบดิจิทัล ลงในแผ่น DVD หรือ Digital Versatile Disc โดยการบีบอัดสัญญาณแบบ MPEG-2

ปัจจุบันเริ่มมีการให้บริการโทรทัศน์ดิจิทัล โดยอาศัยการบีบอัดสัญญาณ MPEG-4 / AVC ทั้งการให้บริการ ผ่านดาวเทียม บริการดิจิทัลเคเบิลทีวี และโทรทัศน์ดิจิทัลภาคพื้นดิน ทั้งแบบ SDTV และ HDTV

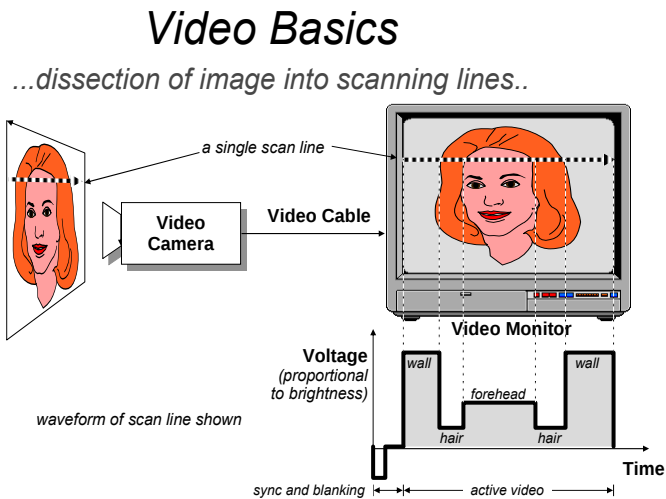
2. เทคโนโลยีของการแพร่ภาพโทรทัศน์แบบอนาล็อก

กิจการ broadcasting โดยเฉพาะโทรทัศน์ ในโลกปัจจุบันมีหลายระบบ เริ่มต้นจากการให้บริการโทรทัศน์ระบบอนาล็อกในแบบ ขาว-ดำ ซึ่งมีหลายระบบ ให้บริการแตกต่างกันในแต่ละประเทศ ตามความแตกต่างทั้งในด้านการเมือง เศรษฐกิจ และทางเทคนิค เมื่อพัฒนาเป็นโทรทัศน์สีจึงต้องต่อยอดเทคโนโลยีจาก โทรทัศน์ขาว-ดำ โดยการแพร่สัญญาณภาพส่วนที่เป็นภาพสีฝากไปกับแถบความถี่ภายในช่องความถี่ที่กำหนด เพื่อให้เครื่องรับโทรทัศน์ขาว-ดำ ยังคงรับสัญญาณจากการส่งโทรทัศน์สีได้ ปัจจุบันมีระบบโทรทัศน์สี 3 ระบบ คือ NTSC, PAL และ SECAM ทั้ง 3 ระบบมีส่วนที่เหมือนกัน คือ ภาพ (ขาว-ดำ) เกิดจากการสแกน (Scan) ทางแนวนอน จำนวนหลายเส้นสแกน (Scan) จนครบเฟรมภาพ(Frame) ในทางทฤษฎีเส้นสแกน (Scan) ทางแนวนอนจะต่อเนื่องกัน นั่นคือ ความชัดเจนของภาพทางแนวนอนไม่มีขีดจำกัด แต่ในทางปฏิบัติความชัดเจนของภาพทางแนวนอน จะถูกจำกัดด้วยความกว้างของแถบความถี่ของช่องโทรทัศน์

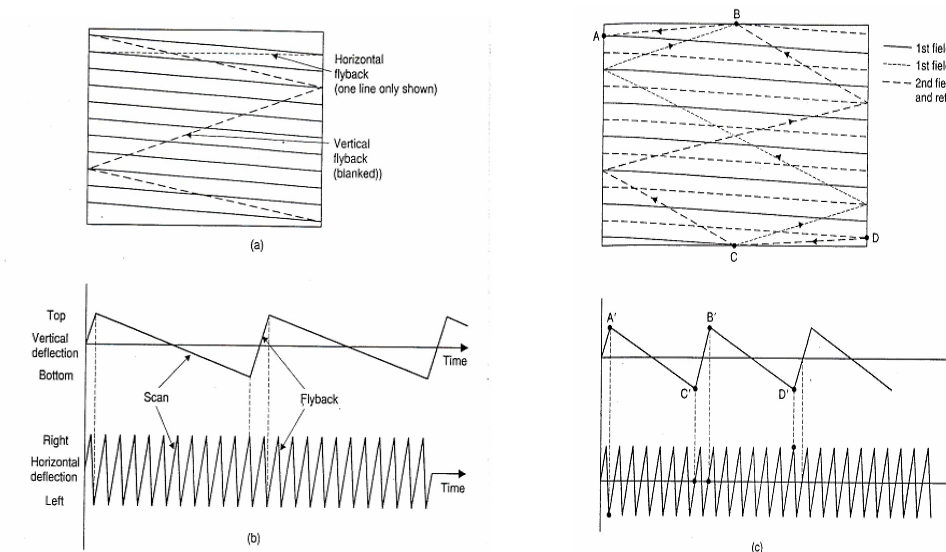
2.1. การ Scanning

โทรทัศน์สี อนาล็อกใช้วิธีสแกน (Scan) ภาพแบบ Interlaced ซึ่งจะต่างกับจอคอมพิวเตอร์ ที่เป็นการสแกนแบบ Progressive การ สแกนแบบ Interlaced ภาพแต่ละเฟรมจะถูกแบ่งออกเป็น Field ที่ และ Field คู่

ทั้งสอง Field จะถูกนำมาแสดงสลับกันด้วยอัตรา จำนวน ภาพ(Frame) ต่อวินาที ที่อ้างอิงกับมาตรฐานภาพยนตร์ 24 ภาพ(Frame) ต่อวินาที คือ 25 ภาพ(Frame) ต่อวินาที ในประเทศที่ใช้กระแสไฟฟ้า 50 Hz และ 30 ภาพ (Frame) ต่อวินาที ในประเทศที่ใช้กระแสไฟฟ้า 60 Hz ทั้งนี้เพื่อหลีกเลี่ยงการวูบวาบ (Flicker) ของภาพ ที่เกิดขึ้นจากการรบกวนเมื่อ จังหวะในการสแกนภาพโทรทัศน์ไม่สัมพันธ์กับแสงสว่างจากหลอดไฟในห้องที่รับชม โทรทัศน์อยู่ ดังแสดงในรูป 1 และ 2



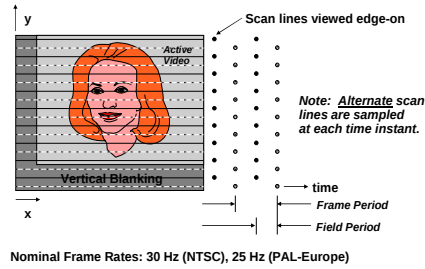
รูปที่ 1 แสดงการสแกนภาพโทรทัศน์



รูปที่ 2 แสดงการสแกนภาพโทรทัศน์แบบ Interlaced ที่ถูกควบคุมโดย Sawtooth waveform ทั้งทางแนวนอนและทางแนวตั้ง

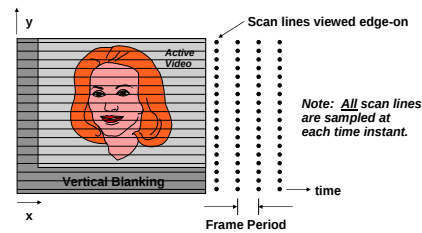
การสแกนภาพแบบ Interlaced สามารถส่งสัญญาณภาพเคลื่อนไหวได้ดี และประหยัดแบนด์วิดท์ของช่องสัญญาณโดยยังคงความชัดเจนทางแนวตั้งในส่วนที่ไม่เคลื่อนไหวไว้ได้ การส่งโทรทัศน์ในปัจจุบันยังคงใช้การสแกนแบบนี้ แต่มีข้อดีคือการปรับเปลี่ยนขนาดของภาพ หรือเปลี่ยนอัตราเฟรมภาพค่อนข้างยุ่งยาก และความชัดเจนทางแนวตั้งลดลงเมื่อภาพเคลื่อนไหวไว้วาง ทั้งส่วนของขอบภาพเช่นขอบตัวหนังสือจะกระพริบเล็กน้อย และนี่คือเหตุที่คอมพิวเตอร์ไม่ใช้การสแกนภาพแบบ Interlaced แต่เลือกใช้แบบ Progressive แทน

The Interlaced Raster



รูปที่ 3 การแสดงผล Interlaced Scan

The Progressive Raster



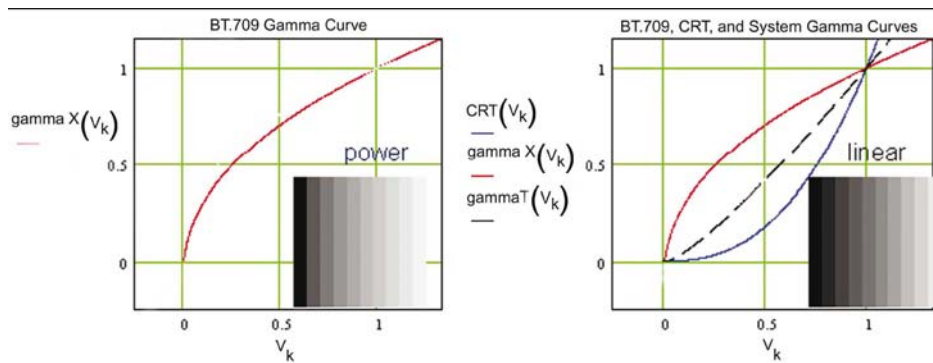
รูปที่ 4 การแสดงผล Progressive Scan

2.2. จอภาพแบบ CRT , LCD และ PDP

ภาพบนจอภาพแบบ CRT (Cathode Ray Tube) เกิดจากการแสกนลำอิเล็กตรอน ภายใต้การควบคุมการแสกนผ่าน สนามแม่เหล็กไฟฟ้า จาก Deflection Coil ความสว่างบนจอขึ้นอยู่กับความเข้มของลำอิเล็กตรอน ซึ่งถูกกำหนดโดยความต่างศักย์ไฟฟ้าของคาโอดต่อศักย์ไฟฟ้ากริดของหลอดภาพ CRT สำหรับศักย์ไฟฟ้าของ Anode จะสูงประมาณ 15-25 KV

เนื่องจากความสัมพันธ์ระหว่างสัญญาณภาพที่ป้อนให้หลอดภาพ CRT กับความสว่างที่เกิดขึ้นบนจอภาพจะไม่แปรผันแบบเส้นตรง (linear) แต่จะเป็นแบบ Exponential มีค่ายกกำลังเรียกว่าค่าแกมมา ประมาณ 2.2-2.8

เพื่อให้กระบวนการถ่ายทำและการแสดงผลบนจอภาพ CRT แปรผันแบบเส้นตรง (linear) จำเป็นต้องทำให้สัญญาณภาพในกระบวนการถ่ายทำเสียรูปไปจากของเดิม โดยการนำกระบวนการ Inverse Gamma มาใช้แก้ไขสัญญาณภาพ ในกล้องโทรทัศน์ ดังรูปที่ 5



รูปที่ 5 Gamma correction complement CRT display response

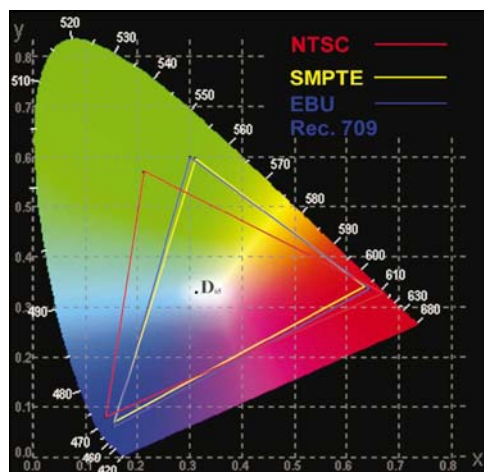
สำหรับจอภาพแบบ LCD : Liquid Crystal Display และ PDP : Plasma Display Panel ทั้งสองชนิดเป็นจอภาพแบบจอแบน โดยเฉพาะ LCD มีหลายขนาด ตั้งแต่ขนาดเล็กสำหรับนาฬิกา โทรศัพท์มือถือและขนาดใหญ่สำหรับ Notebook PC และที่มีขนาดใหญ่กว่า CRT เช่นเครื่องรับโทรทัศน์ HDTV ส่วน PDP จะมีขนาดใหญ่กว่า CRT ทั้ง LCD และ PDP จะมีค่าแกมมาใกล้เคียงกับ CRT แต่มีส่วนที่ดีกว่าคือ ภาพลึชัดเจนไม่ลิไม่เหลื่อมกัน (Convergence error) Linearity ของภาพดีกว่า CRT ไม่มีภาพกระพริบแบบ Flicker สามารถรับสัญญาณได้ทั้งแบบอนาล็อก Component (R,G,B) และสัญญาณดิจิทัลทั้ง DVI และที่ผ่านหัวต่อ HDMI เหมาะสำหรับโครงข่ายสื่อสารภายในบ้าน (IHDN : In-Home Digital Network) ยุคใหม่

ปัจจุบัน (2009) LCD และ PDP ยังมี Resolution ต่ำกว่า CRT จึงมีการพัฒนาการแสดงผลให้มีความชัดเจนเท่าๆ กับ CRT โดยการเพิ่มการสแกนทางแนวตั้งเป็น 2-3 เท่า คือ เป็นแบบ 100 Hz และ 150 Hz แทน

2.3. สีและการมองเห็น

เนื่องจากตาคนจะมีความรู้สึกต่อการเปลี่ยนแปลงสีของแสงน้อยกว่าการเปลี่ยนแปลงความสว่างของแสง และจะมีขีดจำกัดในการมองเห็นความคมชัดหรือความละเอียดใน ปริภูมิ (เชิงพื้นที่ บริเวณ) และไม่สามารถมองเห็นรายละเอียดที่เล็กมากๆ ได้ ที่สำคัญตาคนไม่สามารถจับตามองตามการเคลื่อนที่เร็ว ๆ ของวัตถุได้ดี หากมองภาพที่มีการกระพริบมากกว่า 24 ครั้ง / วินาที (หรือ 24 Hz) ก็จะเห็นเป็นภาพเคลื่อนไหว คุณสมบัติเหล่านี้จึงถูกนำมาพิจารณาออกแบบระบบการส่งโทรทัศน์สี

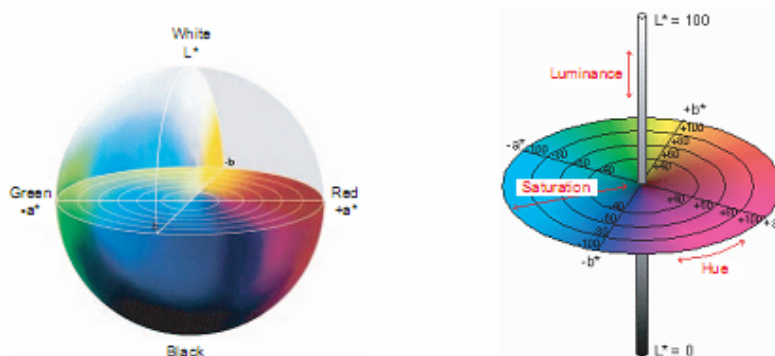
แสงตามธรรมชาติที่ตามองเห็น มีความยาวคลื่นตั้งแต่ 400 นาโนเมตร ถึง 700 นาโนเมตร คือแสงสีม่วง สีสคราม สีนํ้าเงิน จนถึงสีแดง สี (Hue) ของแสงจะแตกต่างกันตามความคลื่นของแสงนั้น ๆ ดังรูปที่ 6



รูปที่ 6 CIE Chromaticity (1931)

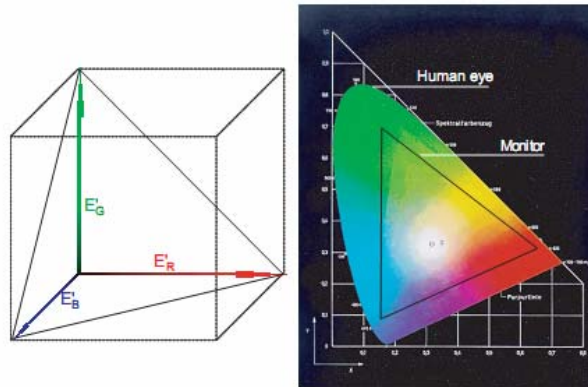
ความอิ่มตัวของสี (Saturation) ที่แสดงผลบนจอมอนิเตอร์หรือจอเครื่องรับโทรทัศน์ NTSC หรือระบบ PAL จะแตกต่างกับ แสงตามมาตรฐาน SMPTE (Society of Motion Picture for Television Engineering) และ EBU (European Broadcasting Union) และแสงจะอิ่มตัวน้อยกว่าแสงตามธรรมชาติที่ตามองเห็น

ไคอะแกรม CIE (Commission Internationale d'Eclairage) Chromaticity นี้ พิมพ์เผยแพร่ ปี 1931



รูปที่ 7 CIE Lab colour space

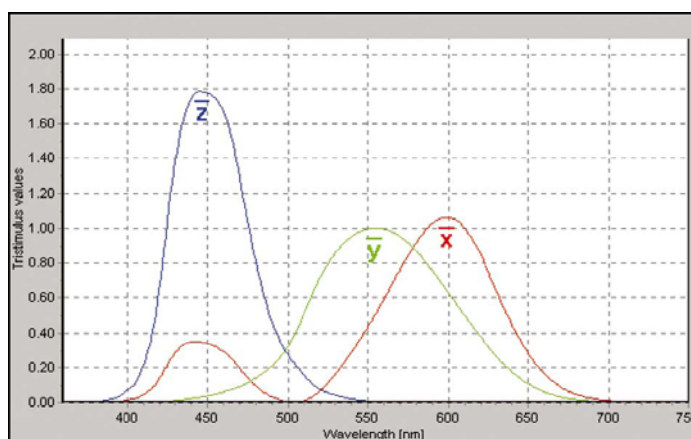
รูปที่ 7 แสดงผลการรับรู้ต่อทุกสีเท่า ๆ กันในแบบ 3 มิติ คือแกนตั้งแสดงความสว่าง (Brighness) แกนนอนเป็นความเข้มของสี และเฟสแสดงสี (Hue) เริ่มเผยแพร่ในปี 1976



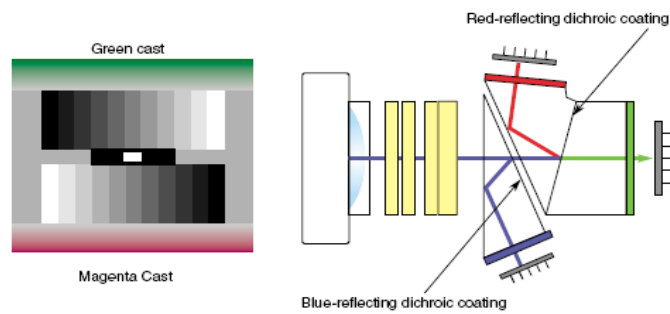
รูปที่ 8 Correlation between RGB cube and CIE xy diagram

รูปที่ 8 แสดงความสัมพันธ์ระหว่างลูกเต๋า RGB และ CIE XY ไคอะแกรม ค่า colour space ที่นำมาใช้ในการโทรทัศน์จะต่างกับเล็กน้อยกับที่ใช้ในวงการคอมพิวเตอร์ (ที่พัฒนาโดย Microsoft และHP) จึงจำเป็นต้องทำความเข้าใจก่อนที่จะการนำภาพคอมพิวเตอร์กราฟิก มาใช้ในการโทรทัศน์ หรือการใช้คอมพิวเตอร์ตัดต่อ (Edit) รายการโทรทัศน์ เพราะสีจะแตกต่างกันเล็กน้อยโดยเฉพาะค่าอิ่มตัวและความสว่างของสี

ค่าสเปกตรัมที่แตกต่างกันของสี RGB ตามรูปที่ 9 นำมาใช้ในการออกแบบปริซึมเพื่อแยกแสง RGB สำหรับกล้องโทรทัศน์สี ตามรูปที่ 10 รวมทั้งนำมาใช้ประโยชน์ในหลอดไฟฟ้าแสงสว่างแบบประหยัดพลังงาน โดยใช้สารเคลือบที่เปล่งเฉพาะแสง ตาม RGB สเปกตรัม เท่านั้น ก็จะไม่เปล่งแสงตลอดค่าสเปกตรัมที่ตามองเห็น จึงประหยัดพลังงานกว่า



รูปที่ 9 ค่าสเปกตรัมที่แตกต่างกันของสี RGB



รูปที่ 10 ปริซึมเพื่อแยกแสง RGB สำหรับกล้องโทรทัศน์สี

2.4. Image Sensor : CCD vs CMOS

Image Sensor เป็นเครื่องมือในการเปลี่ยนแสงให้เป็นสัญญาณไฟฟ้า ปัจจุบันมีให้เลือก 2 แบบ คือ

CCD : Charge Couple Device

CMOS : Complementary Metal Oxide Semiconductor

CCD ออกแบบมาสำหรับใช้ในกล้องโทรทัศน์โดยเฉพาะ มีใช้มานานกว่า 21 ปี CCD จะเปลี่ยน Photon เป็น Electron ก่อนจะส่งผ่าน Electron เป็น Voltage Output แบบอนาล็อกของ CCD สัญญาณนี้จะต้องวงจรอื่นๆ ในตัวกล้องโทรทัศน์ เพื่อเปลี่ยนเป็นสัญญาณดิจิทัล

ส่วน CMOS พัฒนามาบนเทคโนโลยีมาตรฐานการผลิต Memory Chip จึงมีราคาถูกกว่า และกำลังพัฒนาให้มีคุณภาพให้ดีขึ้น กระบวนการเปลี่ยน Photon เป็น Electron ก่อนจะส่งผ่าน Electron เป็น Voltage Output เป็นแบบดิจิทัล ภายในตัว CMOS เอง จึงไม่ต้องใช้วงจรภายนอกช่วย

ดังตารางเปรียบเทียบประสิทธิภาพและคุณสมบัติ

Feature	CCD	CMOS
Signal Out of pixel	Electron packet	Voltage
Signal out of chip	Voltage (Analog)	Bits (Digital)
Signal out of Camera	Bits (Digital)	Bits (Digital)
System Noise	Low	Moderate
System Complexity	Low	Low
Sensor Complexity	Low	Low
Camera Component	Sensor + multiple Support chip+Lens	Sensor + Lens + But additional Support chip common
Relative R&D cost	Lower	Higher

Performance	CCD	CMOS
Responsivity	Moderate	Slightly better
Dynamic Range	High	Moderate
Uniformity	High	Low to Moderate
Uniform Shuttering	Fast, Common	Poor
Speed	Moderate to High	Higher

2.5. Aspect Ratio (อัตราส่วนของภาพ)

ในยุคแรกของการโทรทัศน์ การผลิตหลอดภาพ CRT ก่อนข้างยุ่งยากและมีข้อจำกัดเกี่ยวกับขนาดของจอภาพ จึงเลือกรูปแบบสี่เหลี่ยมผืนผ้า ที่มีอัตราส่วนภาพ 4 : 3 ใกล้เคียงกับขนาดสัดส่วนของภาพยนตร์ในขณะนั้น ปัจจุบันการออกแบบและการผลิตหลอดภาพ CRT ได้ก้าวหน้าไปมาก สามารถลดความเพี้ยนของภาพบริเวณมุมจอได้ดี เพิ่มอัตราส่วนภาพ 4 : 3 ได้โดยไม่สูญเสียคุณภาพ ขนาดที่พอมะคือ 16 : 9 ใกล้เคียงกับขนาดสัดส่วนของภาพยนตร์จอกว้าง



รูปที่ 11 HDTV 16:9 Viewing angle 30 degree



รูปที่ 12 SDTV 4:3 Viewing angle 10 degree

ปัจจุบันอัตราส่วนภาพ 4 : 3 ใช้เป็นมาตรฐาน SDTV และอัตราส่วนภาพ 16 : 9 ใช้เป็นมาตรฐานทั้ง SDTV และ HDTV

2.6. Bandwidth and Definition

จำนวนเส้นสแกนจนครบภาพทางแนวตั้ง เป็นตัวกำหนดความชัดเจนของภาพ ในระบบ 625 เส้น 25 Hz จำนวนเส้นสแกนที่มองเห็นคือประมาณ 600 เส้น ส่วนอีก 25 เส้นที่มองไม่เห็นเป็นส่วนสลับกลับ (Vertical Blanking) หาก 300 เส้นเป็นสีขาวหรือสว่างและอีก 300 เส้นเป็นสีดำหรือมืด นั่นคือจะมี 300 Cycle / Picture Height (300 c/ph) ในทางปฏิบัติความชัดเจนที่ขึ้นย่นได้เพียง 70% ดังนั้นค่าตามอุดมคติ (หรือค่า Kell factor) คือ 210 c/ph

เนื่องจากอัตราส่วนภาพคือ 4 : 3 ดังนั้นการแสดงผลความชัดเจนบนจอตามอุดมคติควรจะเป็น 400 c/ph ตามความกว้างของภาพ ซึ่งลดลงตามค่า Kell factor จนเหลือ 300 c/ph เมื่อรวมส่วนสับคิกกลับเส้นแสดกนทางแนวนอน ค่าที่ต้องการแสดงผล จะเพิ่มจาก 300 c/ph เป็น 360 c/ph

ในระบบ 625 เส้น 25 Hz ความถี่เส้นแสดกนแนวนอน $f_h = 625 \times 25 \text{ Hz} = 15,625 \text{ Hz}$

หาก Video Waveform มีค่าที่ต้องการแสดงผล 360 c/ph

$$\begin{aligned}\text{Bandwidth ของ สัญญาณภาพ Video} &= 15,625 \text{ Hz} \times 360 \text{ c/ph} \\ &= 5.625 \text{ MHz}\end{aligned}$$

นั่นคือ สัญญาณภาพ Video มี Bandwidth กว้าง 5.625 MHz จะมีความชัดเจน (Definition) อย่างน้อย 5.5 MH

2.7. Video Signal

สัญญาณภาพโทรทัศน์สีแบบอนาล็อก ที่เรียกว่า Composite Signal ในระบบโทรทัศน์สี คือ CVBS ได้มาจากการรวมสัญญาณแม่สี R,G,B จากกล้องโทรทัศน์ ให้เป็นสัญญาณ Component Y, R-Y,B-Y ก่อนที่สัญญาณ Colour Difference R-Y, และ B-Y จะมอดูเลตกับแบบ QAM กับ คลื่นพาห้รอง (Colour Sub-Carrier) ความถี่ 4.43 MHz ในระบบ PAL หรือ 3.58MHz ในระบบ NTSC กลายเป็นสัญญาณสี (Chrominance) U&V ในระบบ PAL และ I & Q ในระบบ NTSC ก่อนรวมกับสัญญาณความสว่าง (Luminance) Y และสัญญาณ Sync รวมเป็นสัญญาณภาพ CVBS

2.7.1. สัญญาณภาพ Analog Component Y, R-Y,B-Y

กำหนดให้สัญญาณความสว่าง (Luminance) $Y = 0.299R + 0.587G + 0.114B$

$$\text{Component} \quad R-Y = 0.701R - 0.587G - 0.114B \quad \text{และ} \quad Cr = 0.713(R-Y)$$

$$\text{Component} \quad B-Y = -0.299R + 0.587G + 0.886B \quad \text{และ} \quad Cb = 0.564(B-Y)$$

ในระบบ PAL สัญญาณสี (Chrominance) $U = 0.493(B-Y)$ และ $V = 0.877 (R-Y)$ สัญญาณ U และ V

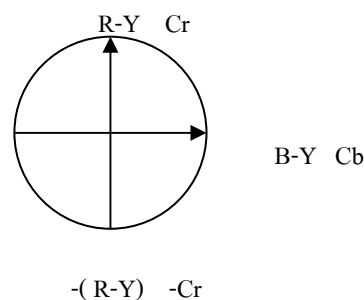
ใช้ในการแพร่สัญญาณ (Transmission) ส่วน สัญญาณ Cr, Cb ใช้ในการเปลี่ยนเป็นสัญญาณดิจิทัล

$$\begin{array}{|c|} \hline Y \\ \hline Cr \\ \hline Cb \\ \hline \end{array} = \begin{array}{|ccc|} \hline 0.3 & 0.59 & 0.11 \\ \hline 0.5 & -0.42 & -0.08 \\ \hline -0.17 & -0.33 & 0.5 \\ \hline \end{array} \begin{array}{|c|} \hline R' \\ \hline G' \\ \hline B' \\ \hline \end{array}$$

$$\begin{array}{|c|} \hline Y = 0.30 R + 0.59 G + 0.11 B \\ \hline Cr \\ \hline Cb \\ \hline \end{array} = \begin{array}{|cc|} \hline 0.00 & 0.71 \\ \hline 0.56 & 0.00 \\ \hline \end{array} \begin{array}{|c|} \hline B'-Y \\ \hline R'-Y \\ \hline \end{array}$$

$$Cr = 0.71 (B-Y)$$

$$Cb = 0.56 (R-Y)$$



รูปที่ 13 สัญญาณ Y ,B-Y ,+/- (R-Y) และ Phase Diagram

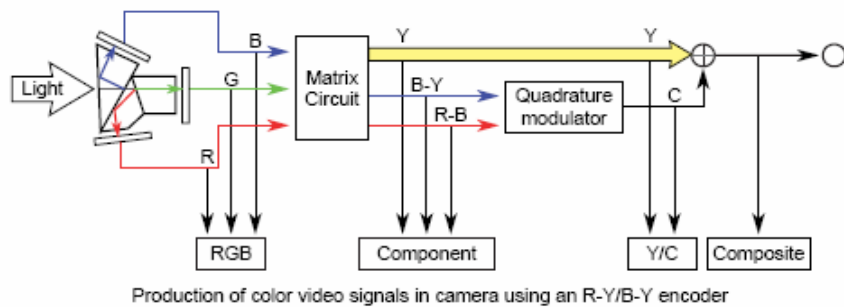
2.7.2. คลื่นพาห้รอง (Colour Sub-Carrier)

เพื่อให้เครื่องรับโทรทัศน์ ขาว-ดำ สามารถรับสัญญาณโทรทัศน์สีได้ จึงไม่สามารถเพิ่มคลื่นพาห้ได้อีก จำเป็นต้องใช้คลื่นพาห้รองแบบ Suppress Carrier เพื่อมอดูเลตแบบ QAM กับสัญญาณ Colure Difference คือ สัญญาณ U และ V ในระบบ PAL

กำหนดให้คลื่นพาห้รอง ห่างจากคลื่นพาห้ภาพและคลื่นพาห้เสียงพอสมควร มีค่าเท่ากับ $1135/4 \times$ ความถี่การสแกนทางแนวนอน และ Off-set อีก 25 Hz เพื่อป้องกันมิให้ สเปกตรัมรบกวนกัน แต่สอดแทรกกระหว่างสเปกตรัม คลื่นพาห้ภาพและคลื่นพาห้เสียง คือ

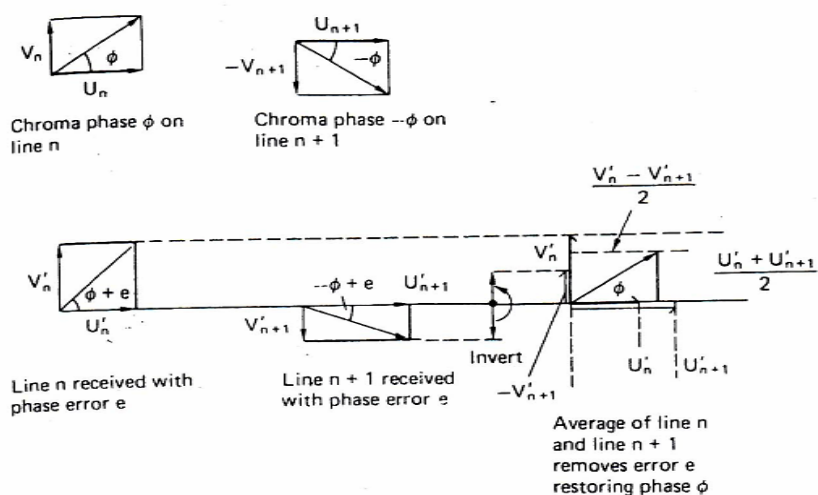
$$1135/4 \times 15625\text{Hz} + 25\text{Hz} = 4.43316875 \text{ MHz}$$

2.7.3. การเข้ารหัสสัญญาณภาพโทรทัศน์สี PAL (PAL Encoding)



รูปที่ 14 สัญญาณภาพโทรทัศน์ CVBS, Y/C, Component Y,B-Y,R-Y และ RGB

การเข้ารหัสสัญญาณภาพโทรทัศน์ แบบอนาล็อก PAL (PAL Encoding) คือการนำสัญญาณ Component B-Y, +/- R-Y หรือสัญญาณ U และ +/- V มอดูเลตกับคลื่นพาห้รองแบบ QAM จะได้สัญญาณ Chroma หรือเรียกว่าสัญญาณ C ที่มีความถี่ 4.43 Hz โดยสัญญาณ +/- R-Y หรือ +/- V จะสลับเฟสทุกเส้นสแกน เมื่อนำไปรวมกับสัญญาณ Y เรียกว่าสัญญาณ Composite หรือ CVBS



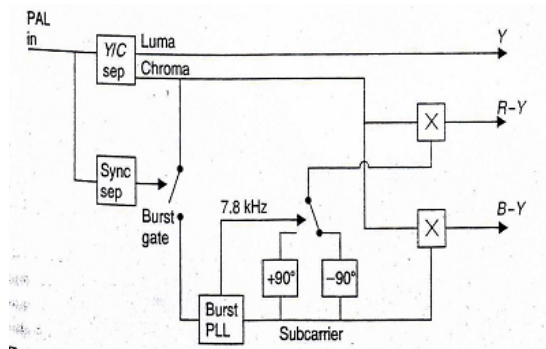
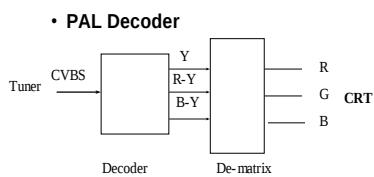
รูปที่ 15 ภาพบน แสดงสัญญาณ Chroma ที่ส่งออก มี Phase = ϕ ภาพล่าง สัญญาณที่รับ Phase เปลี่ยนไป = e เมื่อสลับเฟสไปรวมกัน จะได้ค่าเฉลี่ยที่ เฟสจะไม่เพี้ยน



สัญญาณ Composite CVBS เหมาะที่จะใช้ส่งโทรทัศน์สีแบบอนาล็อกภาคพื้นดิน แต่ไม่เหมาะต่อการผลิตรายการในห้องส่ง รวมทั้งการติดต่อ เนื่องจากการรวมกันของสัญญาณ Y ที่มีแถบความถี่กว้างกว่า 5 MHz กับสัญญาณสี C ที่มีแถบความถี่แคบกว่า คือ 1.5 MHz การส่งผ่านสายเคเบิลจะใช้เวลาไม่เท่ากัน สัญญาณ Y จะช้ากว่า ต้องมีการชดเชย มิฉะนั้นสีจะเลื่อนกัน

การถอดรหัสสัญญาณภาพโทรทัศน์สี CVBS คือการแยกสัญญาณ Y และสัญญาณ C ออกจากกัน ซึ่งมีหลายวิธีแล้วแต่จะต้องการคุณภาพสัญญาณระดับไหน เช่น

- Simple Y/C Separation โดยใช้ Notch filter กรองเอาส่วนที่ประกอบด้วยความถี่ Sub-Carrier fsc 4.43MHz ออกมา มีคุณภาพพอใช้ได้หากสัญญาณสีไม่อึดตัวมากหรือ มีรายละเอียดภาพสีมาก ๆ สีจะเหลื่อมกันบ้าง
- Comp filter Y/C Separation เนื่องจากสเปกตรัมของสัญญาณภาพโทรทัศน์สีจะเกิดขึ้นซ้ำๆ พลังงานของสัญญาณ Luminance Y เส้นแสดก โดยมีเส้นแสดกของคลื่นพาห้สัญญาณสี สอดแทรกในระหว่างเส้นแสดกหลัก Comp filter จึงเหมาะที่จะใช้กรองสัญญาณที่เกิดขึ้นซ้ำ ๆ กัน เนื่องจาก Comp filter มีการตอบสนองความถี่ที่เกิดเป็นช่วงเวลาที่แน่นอน ซึ่งได้ผลดีในระบบ NTSC แต่ในระบบ PAL จะ สูญเสียความชัดเจนของภาพในแนวตั้งบ้าง เนื่องจากการสลับเฟส +/- V และ Comp filter เป็นการกรองเส้นแสดกหลักแนวนอน (Line based)
- Multi-dimensional filtering การใช้ Comp filter แนวนอน (Line based) หรือแบบแนวตั้ง (Field based) อย่างใดอย่างหนึ่งเพียงมิติเดียว กับสัญญาณภาพโทรทัศน์สี PAL จะไม่ได้ผลดีนัก จึงพิจารณาใช้ Comp filter แบบหลายมิติ เช่น 3 มิติ

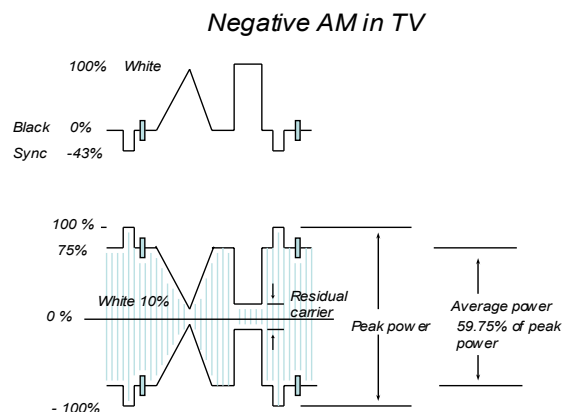


รูปที่ 17 การถอดรหัสสัญญาณภาพโทรทัศน์สี PAL

การเข้ารหัสสัญญาณโทรทัศน์สี ระบบ PAL จึงได้แก้ไขข้อบกพร่องนี้โดยการสลับเฟสสัญญาณ R-Y หรือ V ทุกเส้นสแกน ในการถอดรหัสจึงต้องกลับเฟสแล้วนำสัญญาณมารวมกัน ดังนั้นเมื่อเกิดการ Fading ของสัญญาณ ค่าสัญญาณที่เฉลี่ยรวมกันก็จะต่างกันเฉพาะค่า Amplitude หรือค่า Saturation เท่านั้นแต่สี (Hue) หรือเฟสสีไม่เปลี่ยน

2.8. การแพร่ภาพโทรทัศน์แบบอนาล็อก

สัญญาณภาพโทรทัศน์สีแบบอนาล็อก CVBS จะมอดูเลตกับคลื่นความถี่วิทยุแพร่กระจายจากสถานีส่งไปสู่สายอากาศทางเครื่องรับ เพื่อลดสัญญาณรบกวน การมอดูเลตสัญญาณภาพโทรทัศน์สีแบบอนาล็อกกับคลื่นความถี่วิทยุส่วนมากจะเป็นแบบ Negative Amplitude Modulation ด้วยวิธี Vestigial Sideband (VSB) เพื่อลดขนาดความกว้างของช่องสัญญาณความถี่วิทยุแต่ยังคงคุณสมบัติของ DSB ในการแยกสัญญาณภาพจากคลื่นพาห่ที่มีประสิทธิภาพอย่างง่าย



รูปที่ 18 Negative Amplitude Modulation

ส่วนวิธีการส่งสัญญาณเสียงในระบบโทรทัศน์สีแบบอนาล็อก ในการมอดูเลตกับคลื่นความถี่วิทยุจะแยกกันกับสัญญาณภาพ แต่ทั้งสองสัญญาณจะรวมกันในเครื่องขยายสัญญาณ(ใช้เครื่องขยายสัญญาณร่วมกัน)

หรือก่อนจะป้อนสู่ระบบสายอากาศ (แยกกันใช้เครื่องขยายสัญญาณ) โดยทั่วไปสัญญาณเสียง Mono จะเป็นการมอดูเลตสัญญาณเสียงแบบ FM กับคลื่นความถี่วิทยุ ส่วนสัญญาณเสียง ภาษาที่ 2 หรือเสียงระบบ Stereo จะเป็น Digital NICAM อย่างที่ใช้ประเทศอังกฤษ, ระบบ FM-FM ในประเทศญี่ปุ่น, ระบบ A2 Stereo หรือมาตรฐาน IRT ในประเทศเยอรมัน, และ ระบบ BTSC หรือ MTS ในประเทศสหรัฐอเมริกา

โทรทัศน์สีแบบอนาล็อกทั่วโลก แพร่ภาพในย่านความถี่ VHF และ UHF เหมือนๆกัน แต่การจัดสรรขนาดความกว้างของช่องสัญญาณความถี่วิทยุ จำนวนเส้นสแกน ความถี่ด้านภาพและเสียง จะแตกต่างกันตามมาตรฐานการส่งโทรทัศน์ที่เลือกใช้

ITU ได้จัดสรรความถี่ย่าน VHF และ UHF สำหรับกิจการ Broadcast (แต่หลายประเทศอาจใช้ความถี่เหล่านี้กับกิจการอื่นๆ เช่น Fix หรือ Mobile ร่วมกับกิจการ Broadcast ด้วย) ไว้ดังนี้ คือ

VHF Band I	ความถี่ 47-72 MHz
VHF Band II	ความถี่ 87-108 MHz
VHF Band III	ความถี่ 174-230 MHz
UHF Band IV,V	ความถี่ 470-890 MHz

จะด้วยเหตุผลทางประวัติศาสตร์หรือเหตุผลอื่นๆ ก็ตาม หลายประเทศใช้มาตรฐานการส่งโทรทัศน์ต่างๆกันบนย่านความถี่ UHF มากกว่าย่านความถี่ VHF

ภายหลังการประชุม ITU ณ กรุง Stockholm ในปี 1961 ITU ได้กำหนดมาตรฐานการส่งโทรทัศน์ ตามตัวอักษร A-N ผสมกับระบบโทรทัศน์สี (NTSC,PAL และ SECAM) เช่น PAL-B, NTSC-M เป็นต้น ดังปรากฏในตารางที่ 1 ดังนี้

System	Lines	Frame Rate	BW (MHz)	Visual (MHz)	Sound (MHz)	VSB(MHz)	Vision Mod.	Sound Mod.
A	405	25	5	3	-3.5	0.75	Pos.	AM
B	625	25	7	5	+5.5	0.75	Neg.	FM
C	625	25	7	5	+5.5	0.75	Pos.	AM
D	625	25	8	6	+6.5	0.75	Neg.	FM
E	819	25	14	10	+11.5	2.00	Pos.	AM
F	819	25	7	5	+5.5	0.75	Pos.	AM
G	625	25	8	5	+5.5	0.75	Neg.	FM
H	625	25	8	5	+5.5	1.25	Neg.	FM
I	625	29.97	8	5.5	+5.9996	1.25	Neg.	FM
J	525	25	6	4.2	+4.5	0.75	Neg.	FM
K	625	25	8	6	+6.5	0.75	Neg.	FM

L	625	25	8	6	+6.5	1.25	Pos.	AM
M	525	29.97	6	4.2	+4.5	0.75	Neg.	FM
N	625	25	6	4.2	+4.5	0.75	Neg.	FM

หมายเหตุ ระบบในแถบสีเทา เป็นระบบที่เลิกใช้แล้ว คือ

A ระบบ 405 เส้น ความถี่ VHF ในประเทศอังกฤษ

C ระบบเดิม ความถี่ VHF ในประเทศเบลเยียม

E ระบบ 819 เส้น ความถี่ VHF ในประเทศฝรั่งเศส ความชัดเจนใกล้เคียง HDTV ปัจจุบัน

F ระบบ 819 เส้น แต่ BW 7 MHz ความถี่ VHF ในประเทศเบลเยียมและลักเซมเบิร์ก

ระบบ ที่ยังคงใช้อยู่ในปัจจุบัน ได้แก่

B ทั่วไป ความถี่ VHF เช่น PAL-B, ยกเว้นทวีป ออสเตรเลีย มีทั้ง VHF และ UHF

D ทั่วไปย่านความถี่ VHF เช่น SECAM-D, ยกเว้นจีน PAL-D ทั้ง VHF และ UHF

G ทั่วไปความถี่ UHF เช่น PAL-G ใช้คู่กับ B ในความถี่ VHF ยกเว้นทวีป ออสเตรเลีย

H ความถี่ UHF เช่น PAL-H ใช้ในประเทศเบลเยียม ลักเซมเบิร์ก และกลุ่มประเทศยูโกสลาเวียเดิม

I ความถี่ UHF เช่น PAL-I ใช้ในประเทศอังกฤษ ไอร์แลนด์ อาฟริกาใต้ มาเก๊า และฮ่องกง

J ความถี่ VHF เช่น NTSC-J และ UHF ใช้ในประเทศญี่ปุ่น

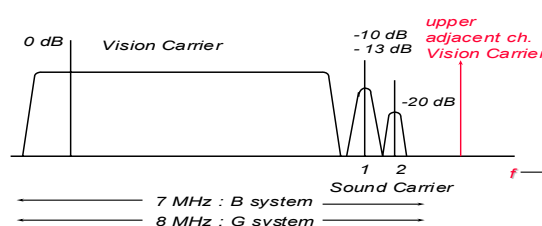
K ทั่วไป ความถี่ UHF เช่น SECAM-K ใช้คู่กับ D ในย่านความถี่ VHF

L ความถี่ VHF Band I เช่น SECAM-L ในประเทศฝรั่งเศส

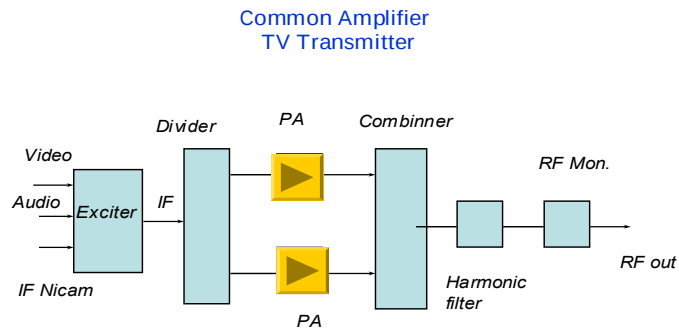
M ความถี่ VHF และ UHF เช่น NTSC-M ใช้ในทวีปอเมริกาเหนือ ฟิลิปปีนส์ เกาหลีใต้ ไต้หวัน และ เมียนมาร์ ส่วนในประเทศบราซิล เป็นระบบ PAL-M

N ความถี่ VHF และ UHF เช่น PAL-N ใช้ในทวีปอเมริกาใต้ อาเจนตินา ปารากวัย อุรุกวัย

สำหรับประเทศไทยและประเทศในกลุ่มอาเซียน ใช้ระบบ PAL-B และ PAL-G ยกเว้น เมียนมาร์ และฟิลิปปินส์ ใช้ระบบ NTSC-M



รูปที่ 19 การส่งโทรทัศน์สี ระบบ PAL B & G



รูปที่ 20 เครื่องส่งโทรทัศน์สี ระบบ PAL B & G

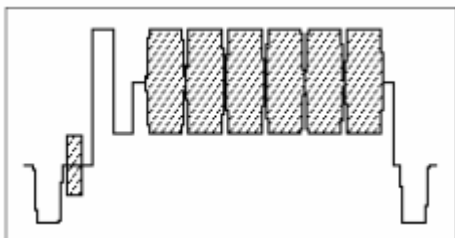
2.9. การตรวจวัดสัญญาณ (Measurement) โทรทัศน์อนาล็อก

คุณภาพในการส่งสัญญาณแพร่ภาพโทรทัศน์ ขึ้นอยู่กับความเพี้ยนของสัญญาณในการส่งและการรับสัญญาณ ซึ่งมีอยู่ 2 ชนิด คือ

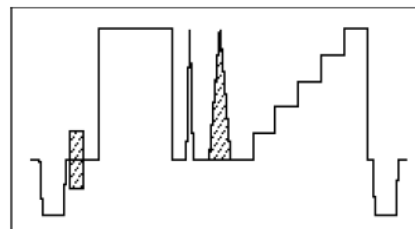
1. Linear Distortion
 - 1.1. Frequency Response
 - 1.2. Group Delay
2. Non- Linear Distortion
 - 2.1. Differential Gain / Differential Phase
 - 2.2. Luminance Non-Linear

การตรวจวัดคุณภาพใช้สัญญาณทดสอบมาตรฐานที่มีคุณสมบัติเหมาะสมในการทดสอบในแต่ละหัวข้อ เช่น

- Multi-burst test signal สำหรับทดสอบ Frequency Response
- 2 T Pulse & Bar Test Signal สำหรับทดสอบ DG / DP, Luminance Non-linear



รูปที่ 21 Multi-burst test signal



รูปที่ 22 CCIR Line 17- Bar ,2 T, Mod 20T, Stare Step

$$\begin{aligned}\text{Cr (อัตราส่วน 2)} &= 360 \text{ pixel} \times 576 \text{ เส้น} \times 25 \text{ ภาพ/วินาที} \times 10 \text{ Bits} = 52.84 \text{ Mbps} \\ \text{จำนวนอัตราบิตทั้งหมด} &= (720 \times 576 + 2 \times 360 \times 576) \times 25 \times 10 = 207.36 \text{ Mbps} \\ &= 207.36 + \text{Audio} + \text{Data} = 270 \text{ Mbps}\end{aligned}$$

สัญญาณทั้งหมดจะมัลติเพล็กซ์ตามลำดับคือ Cb,Y,Cr,Y,Cb,Y.....สตรีมข้อมูลจะเริ่มด้วยรหัส SAV :

Start of Active Video และจบลงด้วยรหัส EAV : End of Active Video ระหว่าง EAV และ SAV จะเป็นสัญญาณ Blanking จะไม่มีข้อมูลเกี่ยวกับ Video เช่นสัญญาณคิงจิตอล ไม่มี Sync Pulse จึงสามารถส่งข้อมูลอื่น ๆ เช่น สัญญาณเสียงคิงจิตอล ได้

การใช้ข้อมูลแบบ 8 Bits หรือ 10 Bits เป็นสัญญาณแบบ 256 และ 1024 ระดับ แต่สัญญาณ Y,Cb,Cr ไม่สามารถใช้ข้อมูลเต็มระดับได้ จำเป็นต้องเผื่อไว้เป็น Head Room และแสดงตัวของ EAV และ SAV

Y ใช้ค่าตั้งแต่ 16 - 235 แบบ 8 Bits และ 64-940 ในแบบ 10 Bits แทนระดับอนาล็อก 0-700 mV

Cb,Cr ใช้ค่าตั้งแต่ 16 - 240 แบบ 8 Bits และ 64-960 ในแบบ 10 Bits แทนระดับอนาล็อก -350-350 mV

สัญญาณ BT.601 มีทั้งแบบขนานที่มีหัวต่อแบบ D-Sub 25 ขา และแบบอนุกรมหรือ SDI : Serial Digital Interface ความต้านทาน 75 Ohms หัวต่อแบบ BNC เหมาะสำหรับงานผลิตรายการโทรทัศน์ในห้องส่ง แต่ไม่เหมาะในการส่งผ่านสื่อ เนื่องจากมีขนาดใหญ่

สัญญาณ BT.709 คือมาตรฐาน HD มี 4 แบบ คือ

System 1 หรือ S 1 ขนาดภาพ 1280 pixel x 720 line Progressive Scan 50 Hz

System 2 หรือ S 2 ขนาดภาพ 1920 pixel x 1080 line Interlace Scan 25 Hz

System 3 หรือ S 3 ขนาดภาพ 1920 pixel x 1080 line Progressive Scan 25 Hz

System 4 หรือ S 4 ขนาดภาพ 1920 pixel x 1080 line Progressive Scan 50 Hz

กำหนดให้ใช้ Sampling Frequency 74.25 MHz สำหรับสัญญาณ Luminance Y

และ 37.125 MHz สำหรับสัญญาณ Colour Difference B-Y, R-Y หรือ Cb, Cr โดยกำหนดความชัดเจนขนาด 8 Bits หรือ 10 Bits และกำหนดรูปแบบสัญญาณสี 4 : 2 : 2 จะได้จำนวนอัตราบิต สำหรับ S2 ดังนี้

สัญญาณ Y (อัตราส่วน 4) = 1920 pixel X 1080 เส้น X 25 ภาพ/วินาที X 10 Bits = 518.40 Mbps

Cb (อัตราส่วน 2) = 860 pixel X 1080 เส้น X 25 ภาพ/วินาที X 10 Bits = 259.20 Mbps

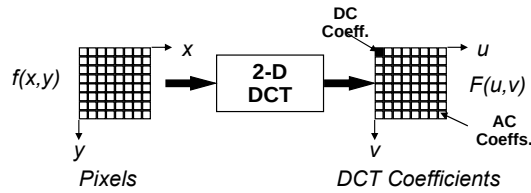
Cr (อัตราส่วน 2) = 860 pixel X 1080 เส้น X 25 ภาพ/วินาที X 10 Bits = 259.20 Mbps

จำนวนอัตราบิตทั้งหมด = (1920 X 1080 + 2 X 860 X 1080) X 25 X 10 = 1036.8 Mbps

4. Transform to and from Frequency Domain

หลักการของแปลงสัญญาณกลับไปกลับมาจากโดเมนความถี่ ถูกนำมาใช้โดยทั่วไปในกิจการสื่อสารโทรคมนาคม ความรู้เรื่องนี้สำคัญมากในการทำควมเข้าใจหลักการ การเข้ารหัสสัญญาณภาพวิดีโอ เสียง และกระบวนการ OFDM เช่น การแปลงฟูเรียร์ (FT) การแปลง Discrete Fourier (DFT) การแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform : FFT) และการแปลง Discrete Cosine (DCT) โดยเฉพาะ DCT ที่นำมาใช้แปลงสัญญาณในโดเมนเวลาในช่วงเวลาหนึ่ง จะได้สัญญาณแอมพลิจูดของ Cosine ในโดเมนความถี่ ซึ่งได้ผลบวกแทนสัญญาณในโดเมนเวลา ค่าสัมประสิทธิ์ที่เท่ากับ “0” สอดคล้องกับส่วนประกอบของ DC ในส่วนของสัญญาณนั้นๆ

DCT and IDCT Formulas



Forward DCT:

$$F(u,v) = (2/N) C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos[(2x+1)u\pi/2N] \cos[(2y+1)v\pi/2N]$$

Inverse DCT:

$$f(x,y) = (2/N) \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u,v) \cos[(2x+1)u\pi/2N] \cos[(2y+1)v\pi/2N]$$

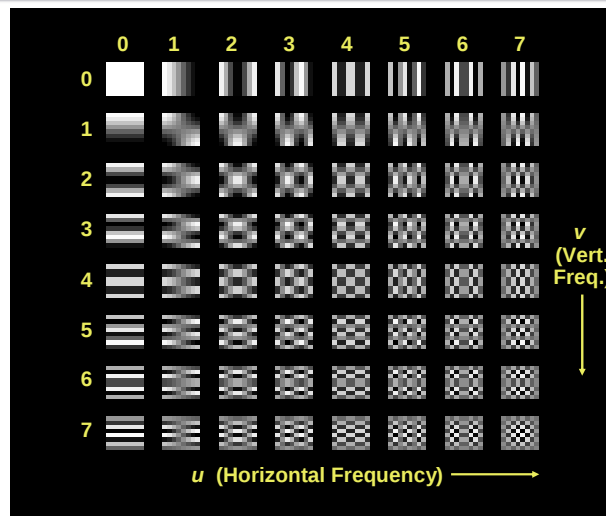
$$\text{where: } \begin{cases} C(u), C(v) = \{1/\sqrt{2} \text{ for } u,v = 0; 1 \text{ otherwise}\} \\ N=8 \end{cases}$$

54

© 1995-99
Sarnoff Corporation

รูปที่ 24 สมการ DCT และ IDCT

2-D DCT Basis Images



55

© 1995-99
Sarnoff Corporation

รูปที่ 25 สัมประสิทธิ์ของ DCT แบบ 2 มิติ 00 คือ DC

DCT ถูกนำมาใช้ในการบีบอัดสัญญาณมาตรฐาน JPEG , MPEG 1 และ MPEG-2

5. Video Coding

สัญญาณดิจิทัลที่ส่ง ความชัดเจนมาตรฐาน มีอัตราข้อมูล 270 Mbit/s มีขนาดใหญ่เกินไปที่ส่งสัญญาณแพร่ภาพ (Broadcast) จึงจำเป็นต้องบีบอัดสัญญาณให้มีขนาดเล็กลง ประมาณ 2-6 Mbit/s ก็พอจะส่งสัญญาณแพร่ภาพ การบีบอัดสัญญาณมีโอกาสเป็นไปได้สูง เนื่องจากข้อมูลภาพมีความซ้ำซ้อนกัน (Redundant) บางส่วน และยังสามารถใช้วิธีการลดข้อมูลบางส่วนที่ไม่สัมพันธ์กัน (Irrelevance) ลงได้อีก

5.1. Video Compression

ข้อมูลภาพมีความซ้ำซ้อนกัน(Redundant) หมายถึงมีมากเกินไปไม่มีประโยชน์ เอาออกได้ สำหรับข้อมูลบางส่วนที่ไม่สัมพันธ์กัน (Irrelevance)หมายถึงข้อมูลที่ไม่จำเป็น ที่ตาคนทั่วไปรับรู้ได้ เช่นภาพที่มีความคมชัดมากๆ สามารถลดลงได้ด้วยการลดแบนด์วิดท์ สัญญาณสี การบีบอัดสัญญาณแบบ JPEG , MPEG ก็ใช้วิธีนี้ อย่างไรก็ตามการลดข้อมูลที่ไม่สัมพันธ์กัน (Irrelevance) จะสูญเสียข้อมูลและคุณภาพบางส่วนไป สรุปคือ

การลดข้อมูล (Data reduction) มี 2 แบบ

- ลดข้อมูลภาพที่มีความซ้ำซ้อนกัน(Redundant) ไม่สูญเสียข้อมูลและคุณภาพ
- ลดข้อมูลที่ไม่จำเป็น ไม่สัมพันธ์กัน (Irrelevance)จะสูญเสียข้อมูลและคุณภาพบางส่วน

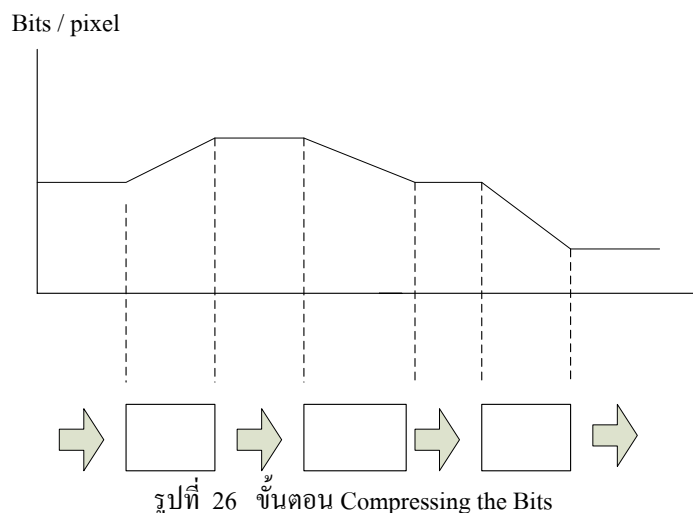
การบีบอัดสัญญาณแบบ MPEG มีขั้นตอนการลดขนาดข้อมูล คือ

1. ลด Resolution จาก 10 Bits เหลือ 8 Bits (จะสูญเสียข้อมูลและคุณภาพบางส่วน)
2. ตัดช่วง Horizontal และ Vertical Blanking ออก (ไม่สูญเสียข้อมูลและคุณภาพ)
3. ลด Resolution สัญญาณสี ทางแนวตั้งลงเป็น 4 : 2 : 0 (จะสูญเสียข้อมูลและคุณภาพบางส่วน)
4. มอดูเลตภาพเคลื่อนไหวด้วยวิธี Differential Pulse Code Modulation :DPCM (ไม่สูญเสียข้อมูลและคุณภาพ)
5. ใช้วิธีแปลงข้อมูลแบบ DCT ตามด้วยการแบ่งนัย (Quantization) ซึ่งจะสูญเสียข้อมูลและคุณภาพบางส่วน
6. ใช้วิธีสแกนแบบ Zig-Zag พร้อมกับเข้ารหัสแบบปรับเปลี่ยนความยาวรหัส Run –Length Coding (ไม่สูญเสียข้อมูลและคุณภาพ)
7. เข้ารหัสแบบ Variable –Length Coding (Huffman Coding ,ไม่สูญเสียข้อมูลและคุณภาพ)

ขั้นตอนการลดขนาดข้อมูล 1-3 สามารถลดขนาดข้อมูลสัญญาณ BT.601 ได้ดังนี้

- | | |
|--|----------------|
| ☐ ITU-R BT.601 | = 270 MBit/s |
| ☐ ลดจาก 10 Bits เป็น 8 Bits (-20%) | = 216 MBit/s |
| ☐ Hor & Ver Blanking (ประมาณ- 25%) | = 166 MBit/s |
| ☐ 4 : 2 : 0 แทน 4:2:2 (-25%) | = 124.5 MBit/s |

ดังนั้นจะต้องใช้ขั้นตอนที่ 4-7 เพื่อลดขนาดข้อมูลจาก 124.5 MBit/s ให้เหลือประมาณ 2-6 MBit/s หรือสูงสุดไม่เกิน 15 MBit/s ตามที่ต้องการใช้ในการ Transmission



5.2. Evolution of the Compression Standard

5.2.1. JPEG

การบีบอัดสัญญาณวิดีโอได้มีวิวัฒนาการตั้งแต่การบีบอัดภาพนิ่ง คือมาตรฐาน JPEG พัฒนาโดยกลุ่มผู้เชี่ยวชาญภาพนิ่ง (Joint Photographic Expert Group) เป็นการเริ่มต้น ใช้ DCT และการแบ่งนัย (Quantization) เป็นเครื่องมือ ต่อมาพัฒนาเป็น M-JPEG หรือ Motion- JPEG เพื่อการบีบอัดภาพเคลื่อนไหว บีบอัดเฉพาะ I เฟรม จึงบีบอัดได้น้อยแต่เสียเวลาหน่วงเหนี่ยวน้อย เหมาะสำหรับงานตัดต่อ ผลิตรายการในห้องส่ง

5.2.2. H.261

H.261 ออกแบบเพื่อใช้ในการสื่อสาร Teleconference มีเป้าหมายขนาดข้อมูล 64 kbits หรือ ผลคูณของ 64 kbits (n x 64 kbits) โดยการนำวิธีการซัดเซกการเคลื่อนที่ แปลงข้อมูลแบบ DCT เฟรม I และเฟรม P บีบอัดได้พอสมควร และเสียเวลาหน่วงเหนี่ยวน้อย

5.2.3. MPEG-1

ในปี 1992 MPEG : Motion Photographic Expert Group ได้ออกแบบมาตรฐาน MPEG-1 เพื่อใช้ในการบันทึกข้อมูลบนแผ่น CD หรือ HDD สำหรับการแสดงผลแบบ Progressive Scan มาตรฐานภาพ SIF : Standard Interface Format บนจอคอมพิวเตอร์ ขนาดภาพ 352 x 288 pixel 25 Hz (จึงไม่จำเป็นต้องใช้เครื่องมือเพื่อ Interlaced Scan) ขนาดข้อมูลไม่เกิน 1.5 Mbits แปลงข้อมูลแบบ DCT เฟรม I, P และเริ่มใช้ B เฟรม ทั้งการพยากรณ์แบบไปข้างหน้าและแบบถอยหลัง จึงบีบอัดสัญญาณได้มาก แต่เสียเวลาหน่วงเหนี่ยวพอสมควร ใน MPEG-1 มีเฉพาะ Program Stream :PS ซึ่งประกอบด้วย Packet Elementary Stream: PES ที่รวมสัญญาณภาพและเสียงไว้ด้วยกัน VLC มีขนาดสูงสุด 64 Kbyte ในรูปแบบสตรีมข้อมูล จึงเหมาะสำหรับจัดเก็บในแผ่น CD, และ HDD ไม่ใช้ในการ Transmission MPEG-1 ได้รับการรับรองเป็นมาตรฐาน ISO/IEC ดังนี้

Part 1 System มาตรฐาน ISO/IEC 11172-1 “ PES Layer”

Part 2 video มาตรฐาน ISO/IEC 11172-2

Part 3 audio มาตรฐาน ISO/IEC 11172-3

5.2.4. MPEG-2

MPEG-2 พัฒนาต่อจาก MPEG-1 และ H.261 ในปี 1995 โดยมีเป้าหมายความชัดเจนระดับมาตรฐานภาพ BT.601 ขนาดภาพ 720 X 576 Pixel 25 Hz สำหรับการแสดงผลแบบ Interlaced Scan เพื่อใช้สำหรับแพร่ภาพโทรทัศน์ SDTV และ DVD ขนาดข้อมูล 2-10 Mbit/s และ 16-50 Mbit/s สำหรับแพร่ภาพ HDTV จากการพัฒนาอย่างต่อเนื่อง จึงนิยมใช้ MPEG-2 เป็นมาตรฐานในบันทึกลงใน HDD เพื่อการผลิตรายการ การกระจายสัญญาณ ทั้งแบบ Contribution และ Distribution ทั้ง SDTV และ HDTV ทั้ง MPEG-1 และ MPEG-2 เป็นมาตรฐาน ISO / IEC ดังนี้

Part 1 System มาตรฐาน ISO/IEC 13818-1 “Transportation”

Part 2 video มาตรฐาน ISO/IEC 13818-2

Part 3 audio มาตรฐาน ISO/IEC 13818-3

5.2.5. MPEG-4

องค์กรระหว่างประเทศสององค์กร คือองค์กรด้าน Broadcasting ISO / IEC และองค์กรโทรคมนาคม ITU ร่วมมือกันพัฒนา MPEG-4 ต่อเนื่องจาก MPEG-2 และ H.263 เพื่อใช้ในการสื่อสารสองทาง (Interactive) สำหรับ Multimedia ขนาดข้อมูลเล็ก ๆ เหมาะกับการสื่อสารผ่านโครงข่ายอยู่กับที่ และโครงข่ายไร้สาย

โดยเฉพาะการ Streaming ผ่าน Internet แต่การพัฒนาได้ผลดีเกินคาด คือสามารถใช้ขนาดข้อมูลตั้งแต่ 28 -500 Kbit/s สำหรับการ Streaming ผ่าน Internet , Mobile Phone และข้อมูลขนาด 0.8-4 Mbit/s สำหรับแพร่ภาพ SDTV MPEG-2 ได้รับการรับรองเป็นมาตรฐาน ISO/IEC ดังนี้

Part 1 System มาตรฐาน ISO/IEC 14496-1

Part 2 Visual มาตรฐาน ISO/IEC 14496-2

Part 3 audio มาตรฐาน ISO/IEC 14496-3

5.2.6. H.264 / AVC

ITU และ ISO / IEC ได้ตั้งคณะทำงานร่วมกันเพื่อพัฒนา มาตรฐานการบีบอัดสัญญาณ ที่มีประสิทธิภาพสูงขึ้น จึงได้พัฒนา H.264 / AVC ขึ้นมาจาก MPEG-4 คือ MPEG-4 part 10 ที่มาตรฐาน ITU เรียกว่า H.264 และด้าน Broadcast มาตรฐาน Part 10 ISO / IEC 14496-10 H.264/AVC : Advance Video Coding

H.264 / AVC มีเป้าหมายความชัดเจนระดับมาตรฐานภาพ BT.601 ขนาดภาพ 720 X 576 Pixel 25 Hz สำหรับการแสดงผลทั้งแบบ Interlaced Scan และ Progressive Scan เพื่อใช้สำหรับการแพร่ภาพโทรทัศน์ SDTV มีขนาดข้อมูล 0.8-4 Mbit/s และข้อมูลขนาด 6-12 Mbit/s สำหรับแพร่ภาพ HDTV และ HD DVD

5.2.7. MPEG-7

MPEG-7 พัฒนาในปี 2001 เพื่อเป็นคำอธิบายเนื้อหาของสื่อประสมเพื่อประโยชน์ในการเชื่อมโยงข้อมูล Metadata XML Based ISO/IEC 15938 “Multimedia Content Description Interface”

5.2.8. MPEG-21

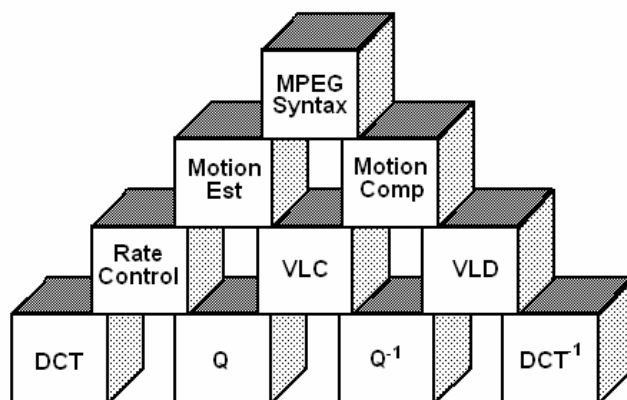
MPEG-21 พัฒนาในปี 2002 เพื่อกำหนดโครงสร้างของ สื่อประสมได้รับการรับรองเป็นมาตรฐาน ISO/IEC ดังนี้

Additional “Tools” ISO/IEC 21000

5.3. MPEG -2 Video Coding

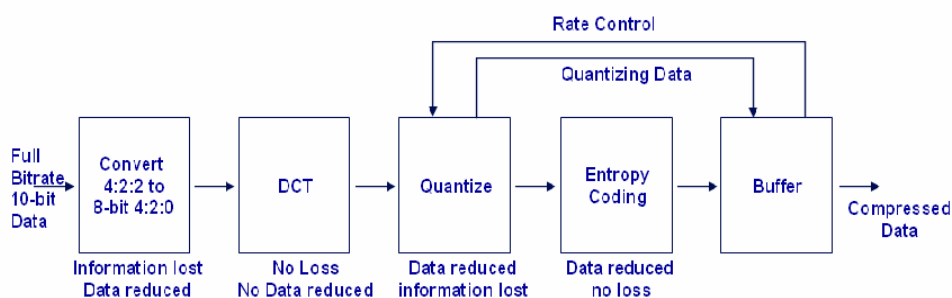
มีกระบวนการหลักในการบีบอัดสัญญาณ MPEG -2 อยู่ 4 กระบวนการคือ

1. กระบวนการเตรียมสัญญาณเพื่อการบีบอัดที่มีประสิทธิภาพ
2. การลดความซ้ำซ้อนเชิงพื้นที่ (Spatial Redundancy) จากเฟรมภาพ หรือฟิลด์ภาพเรียกว่า การบีบอัดแบบ Intra-frame
3. การลดความซ้ำซ้อนเชิงเวลา (Temporal Redundancy) โดยใช้ข้อมูลจากหลาย ๆ เฟรมภาพ เรียกว่าการบีบอัดแบบ Inter-frame
4. กระบวนการทั้งสองกระบวนการจะปรับอัตราบิตโดยระบบควบคุมอัตราบิตเพื่อให้เป็นแบบอัตราบิตคงที่หรือเป็นแบบอัตราบิตปรับค่าได้ ที่สามารถรักษาคุณภาพให้เกือบคงที่ได้



รูปที่ 27 แสดง MPEG Building Blocks

5.3.1. Intra-frame Coding

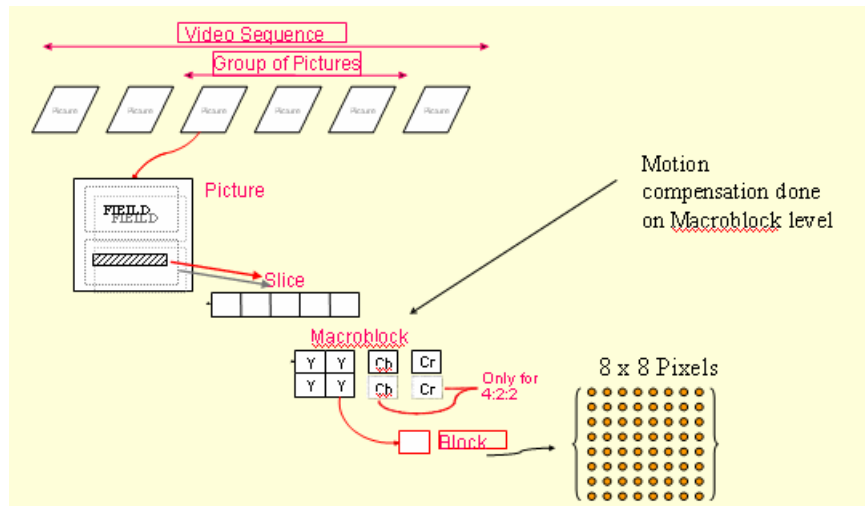


รูปที่ 27 แสดงกระบวนการบีบอัดแบบ Intra-frame

กระบวนการบีบอัดแบบ Intra-frame เพื่อจะลดข้อมูลในภาพเฟรมหรือฟิลด์เดียว มีทั้งแบบสูญเสียข้อมูลและแบบไม่สูญเสียข้อมูล นอกจากจะสูญเสียข้อมูลในการลดอัตราบิต จาก 10 Bits จากเหลือ 8 Bits และสูญเสียความชัดเจนสัญญาณสีจาก 4:2:2 เหลือแบบ 4:2:0 แล้ว จะได้ประสิทธิภาพการบีบอัดสูงจากการควอนไทซ์ (Quantization) สัมประสิทธิ์ที่ได้จากการแปลง DCT เพื่อบีบอัดข้อมูลที่ถูกลด ซึ่งอาจต้องใช้ถึง 11 Bits ที่จะนำเอาสัมประสิทธิ์ DC ไปมอดูเลตอย่างมีประสิทธิภาพเช่น DPCM : Differential Pulse Code Modulation จำนวนบิตจะลดลงเมื่อค่าสัมประสิทธิ์สูงขึ้น

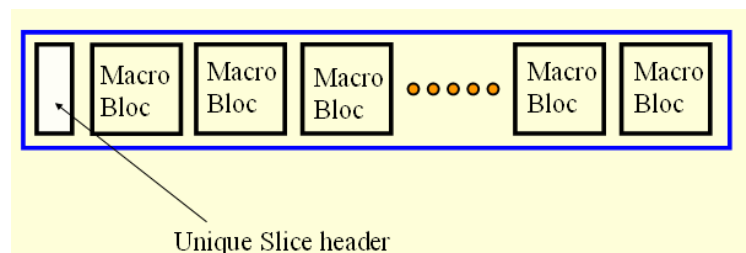
ข้อมูลจะถูกป้อนสู่กระบวนการบีบอัดแบบไม่สูญเสียข้อมูล โดยวิธี RLC : Run Length Coding หรือ VLC : Variable- Length Coding ลำดับในการส่งค่าสัมประสิทธิ์ทำให้ได้ประสิทธิภาพของ RLC ที่เหมาะสม VLC จะลดรหัสให้สั้นลง ลำดับในการส่งค่าสัมประสิทธิ์จะแตกต่างกันระหว่างภาพแบบฟิลด์กับภาพแบบเฟรม ค่าสัมประสิทธิ์ที่ไม่เท่ากับ “0” และที่มีค่าเป็น “0” เกือบทั้งหมดสามารถเข้ารหัสอย่างมีประสิทธิภาพด้วย VLC และการใช้ VLC ทำให้มี Output แบบอัตราบิตปรับค่าได้ (VBR)

5.3.2. MPEG Structure



รูปที่ 28 แสดงโครงสร้าง MPEG 8 x 8 Block

- O ภาพจะถูกแบ่งเป็น 8 x 8 Pixel เรียกว่า Block , ซึ่งจะถูกละเปลี่ยนเป็นโดเมนความถี่แบบ DCT
- O Block จำนวน 4 Block (Y) รวมกับ 4 Block (2Cb+2Cr ในกรณี 4:2:2) เป็น Macroblock ภาพขนาด 720 pixel x 576 Lines จะมี Macroblock = 45 x 36 Macroblocks ในกรณี 4:2:2
- O ลำดับของ Macroblock รวมเป็น Slice
- O โครงสร้างของ Slice อนุญาตให้เครื่องรับทำการ Resynchronization ตอนเริ่มต้นของแต่ละ Slice



รูปที่ 29 แสดง Slice และ Slice header

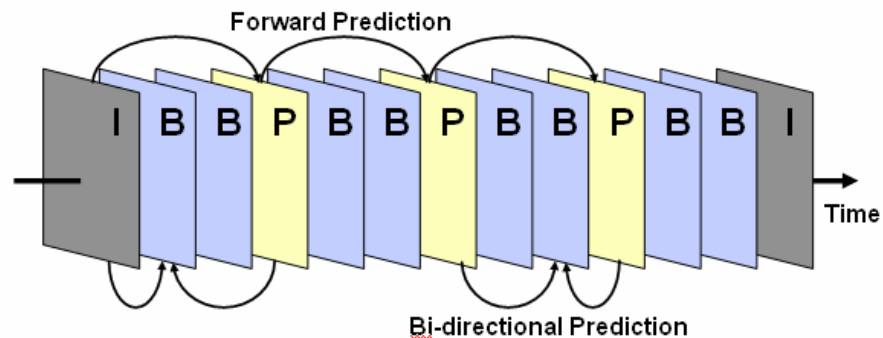
5.3.3. Group of Picture

ชนิดของรูปภาพแบ่งตามลักษณะของการเข้ารหัสสามารถแบ่งได้ 3 ชนิด ได้แก่ I, P และ B โดยมีรายละเอียดดังนี้

1. I (Intra) เป็นภาพที่เข้ารหัสโดยไม่ต้องอาศัยข้อมูลจากเฟรมอื่น ดังนั้นลำดับภาพของวิดีโอที่สลับเฟรมแรกจะเป็นภาพแบบ I เสมอ โดยเป็นเฟรมที่ถูกบีบอัดภายในเฟรม ภาพแบบ I จะถูกส่งทุกๆ ครั้งวินาที มีอัตราการบีบอัดข้อมูลต่ำที่สุดเมื่อเทียบกับภาพทั้งสามแบบ
2. P (Predicted) เป็นภาพที่เข้ารหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า และสามารถถอดรหัสโดยอาศัยข้อมูลของเฟรมก่อนหน้า ซึ่งเฟรมก่อนหน้าเรียกว่า เฟรมอ้างอิงซึ่งอาจจะเป็น I หรือ P เฟรมก็ได้ สำหรับการอ้างอิงถึงข้อมูลของเฟรมอ้างอิงต้องระบุเวกเตอร์การเคลื่อนที่มาด้วย ผลของการบีบอัดภาพแบบ P มีขนาดประมาณร้อยละ 30 ถึง 50 ของเฟรมแบบ I
3. B (Bidirectional Predicted) เป็นภาพที่เข้ารหัสและถอดรหัสโดยอาศัยข้อมูลทั้งจากอดีตและอนาคต รวมกับค่าการเคลื่อนที่ของเฟรมดังกล่าวทั้งสองเฟรม โดยละเลยค่าความเพี้ยนของภาพที่

ได้จากการประมาณ ดังนั้นเฟรมชนิดนี้จะไม่ใช้เป็นเฟรมอ้างอิง ภาพชนิดนี้ให้การบีบอัดข้อมูลสูงสุด กล่าวคือ มีขนาดเหลือเพียงร้อยละ 15 ถึง 25 ของเฟรมแบบ I เท่านั้น

- Contains three types of pictures:
 - Intra (I) pictures
 - Predicted (P) pictures
 - Bi-directional (B) pictures
- intraframe-only spatial DCT*
DCT with forward prediction
DCT with bi-directional prediction



รูปที่ 30 แสดง I,B,P เฟรม และทิศทางการทำนาย

5.3.4. Inter-frame Coding

กระบวนการบีบอัดแบบ Inter-frame คือการลดความซ้ำซ้อนเชิงเวลา (Temporal Redundancy) โดยใช้ข้อมูลจากหลาย ๆ เฟรมหรือหลายฟิลด์ภาพ การเข้ารหัสระหว่างเฟรม (Inter Coding) ซึ่งเป็นการลดความซ้ำซ้อนเชิงเวลา ขึ้นอยู่กับการกำหนดชนิดของรูปภาพในแต่ละเฟรมที่จะทำการเข้ารหัส ถ้าเป็นเฟรมชนิด I (Intra Frame) จะทำการเข้ารหัสภายในเฟรมเท่านั้น แต่ถ้าเป็นเฟรมชนิด P (Predicted Frame) หรือเฟรมชนิด B (Bidirectional Predicted Frame) จะทำการเข้ารหัสระหว่างเฟรม

ตัวอย่างของกระบวนการบีบอัดแบบ Inter-frame ภาพจะซ้ำๆ กันหลาย ๆ เฟรม จะมีภาพบางส่วนเท่านั้นที่จะเปลี่ยนแปลง แม้จะเปลี่ยนแปลงตำแหน่งแต่ข้อมูลรูปร่างอาจไม่เปลี่ยน และภาพจากหลังยังคงเดิมไม่เปลี่ยนเช่นกัน การลดความซ้ำซ้อนเชิงเวลา (Temporal Redundancy) คือการส่งเฉพาะข้อมูลที่เปลี่ยนแปลง หรือหาก Block ภาพเดิมยังคงอยู่แต่เปลี่ยนตำแหน่ง ให้ส่งเฉพาะความสัมพันธ์ตำแหน่งใหม่กับตำแหน่งเดิม (Motion Vectors) หากเป็น Block ของภาพใหม่ก็ให้ส่งทั้ง Block

5.3.5. Motion Compensation (Motion Vector Encoding)

การเคลื่อนที่ของภาพส่วนมากมักจะทำนายได้ การลดขนาดการเคลื่อนที่ก็สามารถเอาชนะความจริงนี้ได้ Motion Compensation จะใช้ Motion Vectors ทั้งใน Encoder และ Decoder เพื่อเสนอภาพที่ได้ทำนายไว้จากการสร้างภาพใหม่ ตาม P เฟรม

Motion Estimation /Compensation

คือกระบวนการคำนวณหา Motion Vectors แต่เป็นคำนวณค่อนข้างละเอียด

สถาปัตยกรรม Motion Estimation Motion Vectors อย่างหยาบๆ คำนวณจากภาพ input ส่วน Motion Vectors อย่างละเอียด เช่นละเอียดระดับ $\frac{1}{2}$ pixel คำนวณจากภาพที่สร้างใหม่

5.3.6. Profile and Level

Level คือตัวบ่งชี้คุณภาพและความชัดเจนของภาพ ส่วน Profile หมายถึงขีดจำกัดของขั้นตอนความยุ่งยากในการเข้ารหัสและการถอดรหัส

MPEG-2 มีหลาย Profile และหลาย Level คือ

High	N/A	4:2:0 1440x1152 60Mbps	4:2:2 1440x1152 300Mbps
Main	4:2:0 720x576 15Mbps	4:2:0 720x576 15Mbps	4:2:2 720x608 50Mbps
Low	4:2:0 352x288 4Mbps	4:2:0 352x288 4Mbps	N/A
Level Profile	Simple	Main	Professional 4:2:2

รูปที่ 31 แสดง MPEG-2 Level & Profile

Simple Profile Format 4:2:0 เฟรมภาพ I,B,P คุณภาพระดับ VHS

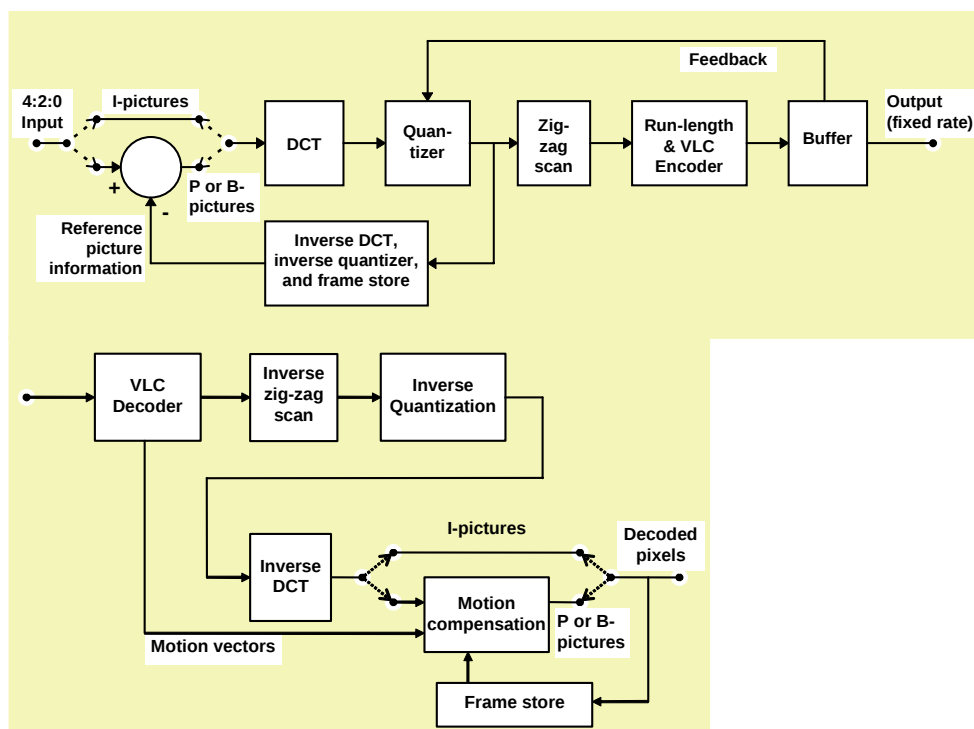
Main Profile Format 4:2:0 เฟรมภาพ I,B,P สำหรับการแจกจ่ายสัญญาณ และแพร่ภาพ
SDTV และ HDTV เช่นโทรทัศน์ผ่านดาวเทียม

422 Professional Format 4:2:2 เฟรมภาพ I,B,P สำหรับงานผลิตรายการในห้องส่ง

ปัจจุบันใช้ MP@ML : Main Profile @ Main Level Format 4:2:0 สำหรับการแจกจ่ายสัญญาณ และ
แพร่ภาพโทรทัศน์ดิจิทัลความชัดเจนมาตรฐาน (SDTV) ขนาดข้อมูล 3-6 Mbit/s

สำหรับ HDTV จะเป็น MP@HL : Main Profile @ High Level Format 4:2:0 ขนาดข้อมูล 16-22 Mbit/s

5.3.7. MPEG-2 Encoder & Decoder (Simplified)



รูปที่ 32 แสดง MPEG-2 Decoder (simplified)

5.4 MPEG-4 Video Coding

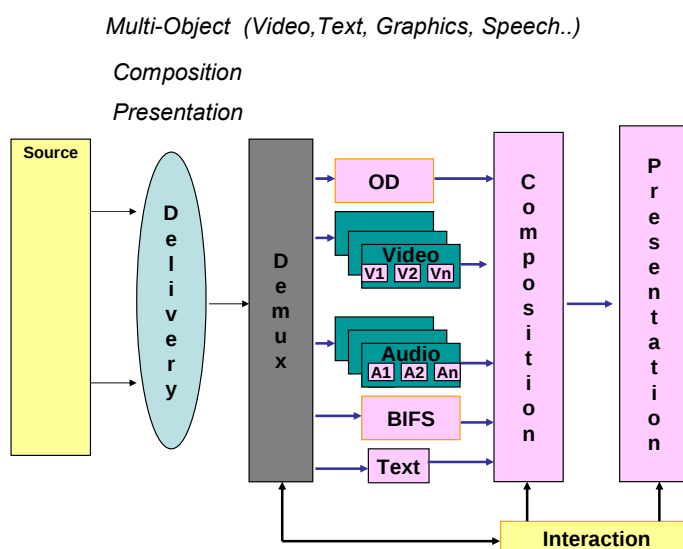
องค์กรระหว่างประเทศสององค์กร คือองค์กรด้าน Broadcasting ISO / IEC และองค์กรโทรคมนาคม ITU ร่วมมือกันพัฒนา MPEG-4 ต่อเนื่องจาก MPEG-2 และ H.263 เพื่อใช้ในการสื่อสารสองทาง (Interactive) สำหรับ Multimedia ขนาดข้อมูลเล็ก ๆ เหมาะกับการสื่อสารผ่านโครงข่ายอยู่กับที่ และโครงข่ายไร้สาย โดยเฉพาะการ Streaming ผ่าน Internet เพื่อความเข้าใจเทคนิคในการบีบอัดสัญญาณ MPEG-4 ควรเปรียบเทียบกับเทคนิคที่ใช้ใน MPEG-2 ซึ่งใช้เป็นมาตรฐานในการแพร่ภาพโทรทัศน์ระบบดิจิทัลในปัจจุบัน

ในขณะที่ MPEG-2 การบีบอัดสัญญาณอยู่บนพื้นฐานเฟรมภาพ โดยการลดสิ่งที่ซ้ำซ้อนและตัดสิ่งที่ไม่ มีประโยชน์ออกไป อาศัยการทำนายการเคลื่อนไหวของภาพในเฟรมภาพข้างหน้าและเฟรมภาพที่ผ่านมาแล้ว เพื่อ ลดข้อมูลลง และนำข้อมูลมาการชดเชยให้ถูกต้อง

5.4.1 MPEG-4 Object

ส่วนการบีบอัดสัญญาณ MPEG-4 จะมีขั้นตอนยุ่งยากมากกว่า เนื่องจากกำหนดให้ภาพที่มองเห็นมิได้มี เพียงวิดีโอ แต่จะมีกราฟิก และตัวหนังสือ (Text) ด้วย เช่นเดียวกับสัญญาณเสียงที่ได้ยินมิได้มีเพียงเสียงพูด แต่จะมีเสียงดนตรี ตามธรรมชาติและเสียงดนตรีที่สร้างขึ้นมาจากอุปกรณ์อิเล็กทรอนิกส์ MPEG-4 จึงมองภาพ และเสียงเหล่านั้นเป็นเป้าหมาย (Object) ที่จะต้องบีบอัดหรือลดปริมาณข้อมูลลงมา

สำหรับการจัดการกับเป้าหมาย (Object) จะต้องอธิบายองค์ประกอบของเป้าหมายอย่างมีระบบ กล่าวคือ องค์ประกอบและข้อมูลการนำเสนอจะรวมอยู่ในเนื้อหาที่นำเสนอ



รูปที่ 34 แสดง Multi-Object (Video,Text,Graphic,Speech..) องค์ประกอบ และการนำเสนอ

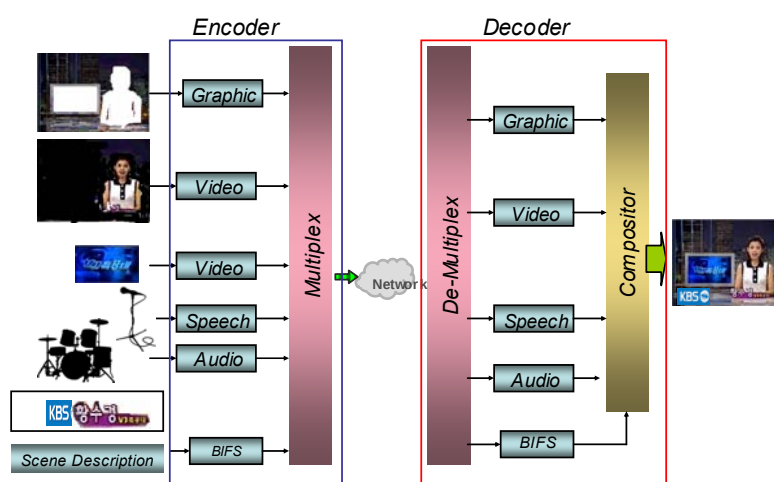
5.4.2 Parts of MPEG-4

เช่นเดียวกับ MPEG-2 ส่วนของ MPEG-4 จะมี ส่วนที่ 1 เป็นระบบ (parts 1 : System) ส่วนที่ 2 และ 3 คือ ภาพ และเสียงตามลำดับ ส่วนของ MPEG-4 ปัจจุบันมี 22 ส่วน ดังนี้คือ

- ส่วนที่ 1 ระบบ
- ส่วนที่ 2 ภาพที่เห็น (Visual)
- ส่วนที่ 3 เสียง (Audio)
- ส่วนที่ 4 การสอดคล้องกัน(Conformance) ของส่วนต่างๆ รวมทั้งข้อมูลการทดสอบ

- ส่วนที่ 5 Software อ้างอิง
- ส่วนที่ 6 DMIF : Delivery Multimedia Integration Framework ส่วนเชื่อมต่อกับระเบียบการส่งผ่านข้อมูล (Transmission Protocol)
- ส่วนที่ 7 Software คำอธิบายอุปกรณ์เข้ารหัสที่เหมาะสม
- ส่วนที่ 8 กำหนดวิธีใช้ MPEG-4 บน IP
- ส่วนที่ 9 คำอธิบายอุปกรณ์ Hardware
- ส่วนที่ 10 Advance Video Coding
- ส่วนที่ 11 คำอธิบายฉาก (BIFS : Binary Format for Scenes) และเครื่องมือในการประยุกต์ใช้
- ส่วนที่ 12 รูปแบบเพิ่มข้อมูล (File format) ตามมาตรฐาน ISO
- ส่วนที่ 13 ส่วนต่อขยาย IPMP
- ส่วนที่ 14 รูปแบบเพิ่มข้อมูล MP4 (ตามมาตรฐาน ส่วนที่ 12)
- ส่วนที่ 15 รูปแบบเพิ่มข้อมูล AVC (ตามมาตรฐาน ส่วนที่ 12)
- ส่วนที่ 16 Animation Framework eXtension : AFX
- ส่วนที่ 17 Streaming Text Format
- ส่วนที่ 18 Font Compression and Streaming
- ส่วนที่ 19 Synthesized Texture Streaming
- ส่วนที่ 20 Lightweight Application Scene Representation (Laser)
- ส่วนที่ 21 MPEG-J Graphical Framework eXtension
- ส่วนที่ 22 Open Font Format

ส่วนที่เป็นหลักสำคัญที่กล่าวถึง คือ ส่วน 1-10



รูปที่ 35 แสดง Over-all Architecture of MPEG-4

5.4.3 MPEG-4 Part 2 Visual

กรอบความคิดเรื่องเป้าหมายคล้ายกับภาษาการเขียนโปรแกรมคอมพิวเตอร์ ใน MPEG-4 เป้าหมายจะแยกกันเด็ดขาดสามารถเลือกเฉพาะเป้าหมายที่สามารถปรับเปลี่ยนได้ จากรูปที่ 35 จะมีฉากหลัง กราฟฟิก และกราฟฟิกที่เป็นโลโก้ ชื่อสถานี หากภาพที่แสดงในรูปเป็นสัญญาณที่มาจาก MPEG-2 จะไม่สามารถลบภาพส่วนที่เป็นกราฟฟิกหรือที่เป็นโลโก้ได้ เพราะถูกบีบอัดทั้งเฟรม แต่หากเป็น MPEG-4 จะสามารถแยกภาพแต่ละส่วน แล้วนำมาประกอบกันเป็นภาพใหม่และสามารถใช้ซ้ำๆ ได้ แม้ว่าในโลกของความจริงการแยกภาพออกเป็นส่วนๆ จากภาพเดิมจะยากมาก แต่รายการโทรทัศน์ต้องการส่วนประกอบต่างๆ มากขึ้น โดยเฉพาะในกระบวนการตัดต่อ เช่น เพิ่มตัวอักษรข้อความ ภาพกราฟฟิกและเสียงดนตรีประกอบ ความคิดแบบนี้จึงมีความหมายสำหรับสร้างสรรค์เนื้อหาและจัดการกับรายการโทรทัศน์ได้ดีขึ้น ในเวลาเดียวกันเมื่อเปรียบเทียบ MPEG-2 ที่มีคุณภาพเท่าๆ กันแล้ว MPEG-4 part 2 Visual จะมีขนาดข้อมูลน้อยกว่าประมาณ 15-20%

5.4.4 Major Functionality in MPEG-4

หน้าที่หลักของแต่ละส่วน ใน MPEG-4 ได้แก่

(1) Transport

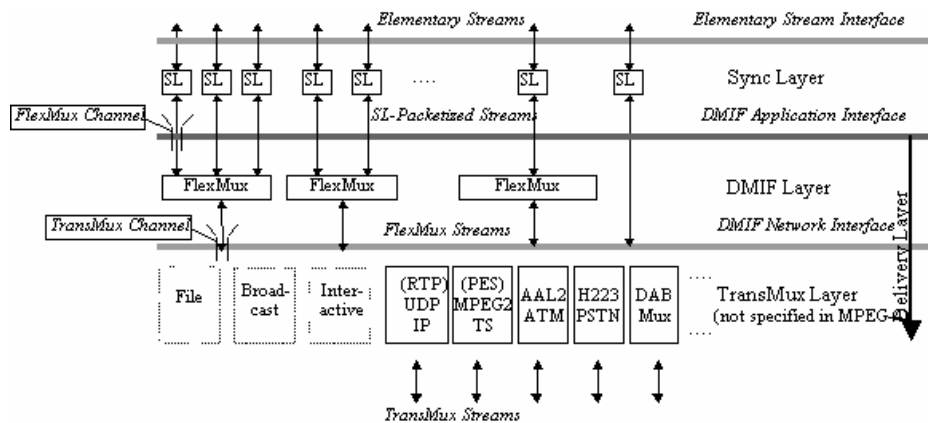
โดยหลักการ MPEG-4 ไม่กำหนดขั้นตอนการขนส่งข้อมูล แต่จะปรับตัวให้เข้ากับระบบการขนส่งข้อมูลเป็นกรณีๆ ไป เช่น

- การขนส่งข้อมูลบน MPEG-2 ทรานสปอร์ตสตรีม (ปรับเข้าหา MPEG-2 System)
- การขนส่งข้อมูลบน IP (ร่วมกับ IETF : Internet Engineering Task Force)

(2) DMIF

Delivery Multimedia Integration Framework คือการเชื่อมต่อระหว่างการประยุกต์ใช้และการขนส่ง ซึ่งยอมให้ผู้พัฒนาการประยุกต์ใช้ MPEG-4 ไม่ต้องกังวลกับการขนส่งข้อมูล การประยุกต์ใช้แบบใดๆ สามารถใช้งานบนชั้นการขนส่งที่ต่างกันได้ เมื่อได้รับการสนับสนุนจาก DMIF ที่ถูกต้อง ตัวอย่างเช่น MPEG-4 DMIF สนับสนุนหน้าที่ต่างๆ

- การประยุกต์ใช้ MPEG-4 DMIF อย่างปลอดภัยโดยไม่ต้องคำนึงถึงว่า จะเชื่อมต่อแบบสื่อสารทางไกลระหว่างผู้ที่เท่าเทียมกัน การแพร่ภาพกระจายเสียงหรือกับหน่วยบันทึกสื่อ
- ควบคุมการจัดตั้ง ช่องสัญญาณ FlexMux
- การใช้โครงข่ายที่มีลักษณะเหมือนกัน ระหว่างการสื่อสารระหว่างผู้ที่เท่าเทียมกันบน IP, ATM, โทรศัพท์เคลื่อนที่, PSTN, Narrowband ISDN
- สนับสนุนโครงข่ายโทรศัพท์เคลื่อนที่ โดยร่วมพัฒนากับ ITU-T
- ผู้ใช้สั่งการพร้อมการรับรู้ข้อความ
- การบริหารจัดการ MPEG-4 –ข่าวสาร Sync Layer



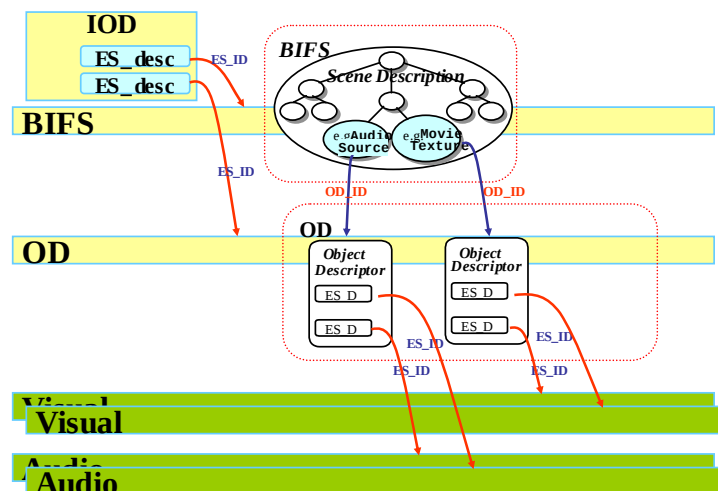
รูปที่ 36 แสดง MPEG-4 System Layers

(3) System

ดังที่ได้อธิบายตั้งแต่ตอนต้น MPEG-4 ได้กำหนดเครื่องมือเพื่อการบีบอัดสัญญาณที่ก้าวหน้าที่มีลำดับขั้นตอนสำหรับข้อมูลสัญญาณภาพและเสียง กระแสข้อมูลเบื้องต้น (Elementary Stream:ES) ซึ่งเป็นผลมาจากกระบวนการเข้ารหัส สามารถแยกส่งหรือแยกบันทึกกันได้ โดยจำเป็นต้องนำมาประกอบเข้าด้วยกัน เพื่อสร้างสื่อประสมที่นำมาแสดงบนเครื่องรับได้

MPEG-4 System จะจัดการ การอธิบายความสัมพันธ์ของชิ้นส่วนภาพและเสียงที่จะนำมาประกอบกันเป็นภาพและเสียงฉากใหม่ ความสัมพันธ์นี้จะอธิบาย ใน 2 ระดับ

- ในระดับต่ำลงมา OD : Object Descriptor จะกำหนดความสัมพันธ์ระหว่าง Elementary Stream:ES ให้เข้าเรื่องกับเป้าหมาย OD ได้เตรียมข้อมูลเพิ่มเติม เช่น URL ที่จำเป็นในการเข้าถึง ES คุณลักษณะของเครื่องถอดรหัส จำเป็นต้องวิเคราะห์ ทฤษฎีสันทางปัญญาและอื่นๆ



รูปที่ 37 แสดง MPEG-4 Decoding

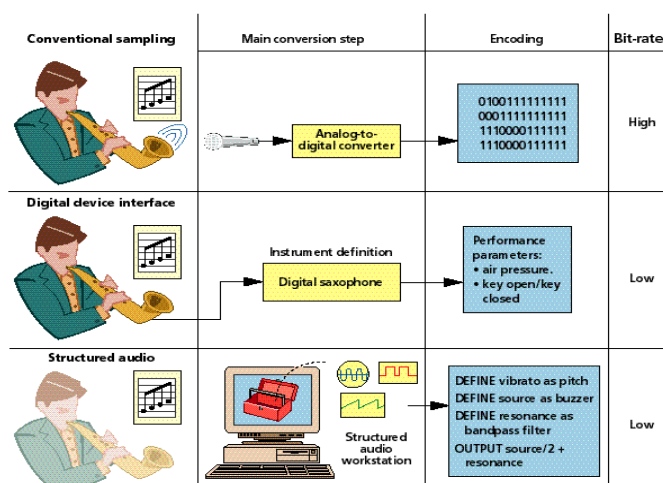
- BIFS จะอธิบายการจัดเรียงชิ้นส่วนภาพที่เคลื่อนไหวตามเป้าหมายในฉาก ผู้ชมอาจมีส่วนร่วมในการโต้ตอบกับเป้าหมาย เช่น การจัดเรียงภาพแต่ละฉากเสียใหม่ หรือเปลี่ยนมุมมอง แบบ 3 มิติตามสิ่งแวดล้อมเสมือนจริง การบรรยายฉากได้จัดเตรียมฉาก และกราฟฟิกขั้นพื้นฐานไว้มากมาย สำหรับผู้ประกอบฉากแบบ 2 มิติ หรือ 3 มิติ

ส่วนประเด็นอื่นๆ ที่จัดการโดย MPEG-4 System ได้แก่

- File มาตรฐานที่สนับสนุนการแลกเปลี่ยน การอนุญาตให้ใช้ เนื้อหา MPEG-4
- การสื่อสารระหว่างกัน รวมทั้งการสื่อสารระหว่างเครื่องลูกข่ายกับ Serverแม่ข่าย
- ใช้ Java (MPEG-J) เพื่อให้สามารถสอบถามสถานีปลายทางภายใต้สิ่งแวดล้อมที่สนับสนุนอยู่ และยังมีการประยุกต์ใช้เครื่องมือ Java เพื่อเข้ารหัส 'MPEGlets'
- เป็นเครื่องมือเพื่อจัดสตรีมหลายๆ สตรีมให้เหลือสตรีมเดียว ได้รวบรวมข้อมูลเวลาเอาไว้ด้วย (FlexMux tool).
- เป็นเครื่องมือเพื่อเก็บบันทึกข้อมูล MPEG-4 ในแฟ้มข้อมูล MPEG-4 คือ 'MP4'
- เชื่อมต่อกับสถานีปลายทางและโครงข่ายในหลายแง่มุม ในรูปแบบการประยุกต์ใช้ Java (MPEG-J)
- ขั้นตอนการขนส่งมีอิสระในการจัดเส้นทางของการขนส่งที่สัมพันธ์กันแบบซ้อนๆกัน เหมือนกับ (RTP)/UDP/IP หรือ MPEG-2 TS ที่จะถูกกำหนดร่วมกับองค์กรที่รับผิดชอบเรื่องมาตรฐาน
- ตัวอักษร .(Text) สนับสนุนภาษาสากลทุกภาษา รวมทั้งสามารถเลือกรูปแบบตัวอักษรด้วย เวลาและการเกิดขึ้นในเวลาเดียวกัน
- เริ่มต้นดำเนินการบริหารจัดการ ส่วน buffersของเครื่องรับปลายทาง และบริหารจัดการอย่างต่อเนื่อง
- แสดงเวลาและควบคุมการเกิดขึ้นพร้อมกัน รวมทั้งมีกลไกในการกู้คืน
- ชุดข้อมูลแสดงสิทธิในทรัพย์สินทางปัญญา สัมพันธ์กับสื่อเป้าหมาย

(4) Audio

ใน MPEG-4 ได้จัดให้มีการประยุกต์ใช้ Audio หลายๆแบบได้สะดวกขึ้น ตั้งแต่เสียงพูดจนถึงเสียงแบบหลายช่องเสียงที่มีคุณภาพ และจากเสียงตามธรรมชาติถึงเสียงที่สร้างขึ้นมา โดยเฉพาะสนับสนุนการทดแทนเป้าหมายเสียงที่มีประสิทธิภาพสูง ประกอบด้วย



Source: Eric Scheiner, Massachusetts Institute of Technology

รูปที่ 38 แสดง Structure Audio

- General Audio Signals

สนับสนุนการเข้ารหัสสัญญาณเสียงตั้งแต่อัตราบิตต่ำจนถึงสัญญาณเสียงที่มีคุณภาพสูงโดยใช้เทคนิคการแปลงรหัส มีอัตราบิตตั้งแต่ 6 KBits/s ขนาดแบนด์วิดท์น้อยกว่า 4KHz จนถึงคุณภาพระดับกระจายเสียง โดยเสียเวลาในการเข้ารหัสน้อย

- Speech signals

ในการเข้ารหัสเสียงพูดสามารถใช้อัตราบิตตั้งแต่ 2 KBits/s จนถึง 24 KBits/s โดยใช้เครื่องมือการเข้ารหัสเสียงพูด อัตราบิตต่ำๆ เฉลี่ย 1.2 kbit/s ก็สามารถใช้ได้เมื่ออนุญาตให้ใช้การเข้ารหัสแบบเปลี่ยนแปลงอัตราบิตได้สามารถประยุกต์ใช้ในกิจการสื่อสารได้ดีเพราะเสียเวลาในการเข้ารหัสน้อย เมื่อใช้เครื่องมือแบบ HVXC (Harmonic Vector eXcitation Coding) ความเร็วและระดับเสียง สามารถปรับเปลี่ยนในขณะที่เล่นกลับได้

- Synthetic Audio

โครงสร้างของ MPEG-4 Audio คือภาษาที่ใช้อธิบาย “เครื่องดนตรี” (โปรแกรมขนาดเล็กเพื่อให้กำเนิดเสียง) และ “โน้ตเพลง” (ข้อมูลที่ป้อนเข้าไปขับเคลื่อนเป้าหมายนั้นๆ) เป้าหมายนั้นอาจไม่จำเป็นต้องเป็นเครื่องดนตรี แต่เนื้อแท้คือสมการคณิตศาสตร์ เพื่อให้กำเนิดเสียงเปียโน เช่นเสียงน้ำตก หรือเสียงแบบที่ไม่เคยได้ยินในธรรมชาติ

- Synthesized Speech

อุปกรณ์เข้ารหัสที่วัดอัตราบิตได้ TTS ใช้อัตราบิตตั้งแต่ 200 Bits/s จนถึง 1.2 KBits/s โดยการป้อนข้อมูลแบบตัวอักษร (Text) เพื่อสังเคราะห์เสียงพูดที่เข้าใจได้ดี

(5) Visual

มาตรฐาน MPEG-4 Visual อนุญาตให้เข้ารหัสแบบผสมกันระหว่างการเข้ารหัสภาพและวิดีโอตามธรรมชาติ(บนพื้นฐาน Pixel) กับฉากที่สังเคราะห์ขึ้นมา (ด้วยคอมพิวเตอร์) นั่นคือจะมีเครื่องมือและชุดของคำสั่งที่สนับสนุนการเข้ารหัสภาพนิ่งและวิดีโอตามธรรมชาติ(บนพื้นฐาน Pixel) และจะมีเครื่องมือที่สนับสนุนการบีบอัดสัญญาณกราฟฟิก 2 มิติ และ 3 มิติ จากตัวแปรเรขาคณิตที่สังเคราะห์ขึ้นมา หัวข้อย่อย ต่อไปนี้ จะอธิบายมาตรฐาน MPEG-4 Visual คือ

- Formats Supported

- MPEG-4 Visual สนับสนุนและอัตราบิตตั้งแต่ 5 kbit/s และเกิน 1 Gbit/s (จึงStreamingได้)
- ในรูปแบบ การแสดงภาพวิดีโอแบบ progressive และวิดีโอแบบ interlaced
- ความชัดเจนภาพจากระดับ QCIF ถึงความชัดเจนภาพระดับห้องส่ง (4k x 4k pixels)

- Compression Efficiency

- สำหรับทุกๆอัตราบิตที่กล่าวถึง ชุดคำสั่งมีประสิทธิภาพสูงสุด การเข้ารหัสโครงสร้างพื้นฐานแบบกระชับ ที่ปรับคุณภาพได้ระหว่างยอมรับได้ "acceptable" สำหรับแบบอัตราบีบอัดสูง ถึงระดับ เกือบไม่สูญเสีย "near lossless".
- ประสิทธิภาพบีบอัดสัญญาณของโครงสร้างพื้นฐานสำหรับแบบร่างโครงสร้างตาข่าย (Mesh) แบบ 2 มิติ และ 3 มิติ
- การเข้าถึงวิดีโอแบบสุ่มเพื่อใช้คำสั่ง pause, fast forward, fast reverse วิดิทัศน์ที่เก็บไว้

- Content-Based Functionalities

- การเข้ารหัสภาพและวิดีโอตามเนื้อหา จึงสามารถแยกการถอดรหัสและประกอบภาพขึ้นมาใหม่จากรูปร่างที่ไม่มีลักษณะของวิดีโอเป้าหมาย
- การเข้าถึงวิดีโอแบบสุ่มเนื้อหาเพื่อใช้คำสั่ง pause, fast forward and fast reverse วิดีโอที่เก็บไว้
- ส่วนขยายการปรับเนื้อหาวิดีโอให้เหมาะสม เพื่อทำการบิด งอ ภาพสังเคราะห์หรือภาพตัวอักษรตามธรรมชาติ โครงร่าง ภาพและวิดีโอที่ซ้อนทับบนเนื้อหาที่สร้างขึ้นใหม่ ตัวอย่างเช่น ตัวหนังสือซ้อนบนวิดีโอเป้าหมายซึ่งตัวหนังสือจะเคลื่อนที่ไปพร้อมกับเป้าหมาย
- **Scalability of Textures, Images and Video**
 - ความสามารถในการปรับเปลี่ยนความยุ่งยากในอุปกรณ์เข้ารหัส ทำให้อุปกรณ์เข้ารหัสสามารถสร้างบิตสตรีมที่มีเหตุผล มีความหมาย ตามรูปร่างของภาพและวิดีโอ
 - ความสามารถในการปรับเปลี่ยนความยุ่งยาก ทำให้อุปกรณ์ถอดรหัส สามารถถอดรหัสตามรูปร่างของภาพและวิดีโอบิตสตรีมโดยอุปกรณ์ถอดรหัสที่มีความยุ่งยากแตกต่างกัน คุณภาพของภาพที่สร้างขึ้นใหม่ โดยทั่วไปสัมพันธ์กับความยุ่งยากของอุปกรณ์ถอดรหัส ซึ่งเป็นผลให้อุปกรณ์ถอดรหัสที่มีความสามารถน้อยจะถอดรหัสภาพออกมาได้เฉพาะบางส่วนของบิตสตรีม การปรับเปลี่ยนขนาดพื้นที่ในอุปกรณ์ถอดรหัสทำให้สามารถถอดรหัสบิตสตรีมย่อยที่สร้างโดยอุปกรณ์เข้ารหัส เพื่อจะได้สร้างและแสดงภาพและวิดีโอเป้าหมายที่สร้างขึ้นใหม่ จากบิตสตรีมที่ลดความชัดเจนลง อุปกรณ์เข้ารหัสสนับสนุนการปรับเปลี่ยนขนาดพื้นที่ได้ถึง 11 ระดับ ที่เรียกว่า 'fine-granularity scalability'
 - ความสามารถในการปรับเปลี่ยนตำแหน่งชั่วคราวทำให้อุปกรณ์ถอดรหัส สามารถถอดรหัสบิตสตรีมย่อยที่สร้างโดยอุปกรณ์เข้ารหัส เพื่อจะได้สร้างและแสดงภาพและวิดีโอเป้าหมายที่สร้างขึ้นใหม่ จากบิตสตรีมที่ลดความชัดเจนลง อุปกรณ์เข้ารหัสสนับสนุนการปรับเปลี่ยน ถึง 3 ระดับ
 - ความสามารถในการปรับเปลี่ยนคุณภาพ สามารถแยกบิตสตรีมออกเป็นชั้นๆ ที่มีอัตราบิตต่างกัน โดยที่การรวมกันของส่วนย่อยของแต่ละชั้นยังคงถอดรหัสออกมาเป็นสัญญาณที่มีความหมาย การแยกบิตสตรีมออกอาจเกิดขึ้นในการส่งสัญญาณหรือในอุปกรณ์ถอดรหัส คุณภาพในการสร้างขึ้นใหม่สัมพันธ์กับจำนวนชั้นที่ใช้ในการถอดรหัสและการสร้างขึ้นใหม่
 - ความสามารถในการปรับเปลี่ยนอย่างละเอียด(Fine Grain) การรวมกันของส่วนย่อยอย่างละเอียดมี 11 ระดับ
- **Shape and Alpha Channel Coding**
 - การเข้ารหัสรูปร่างช่วยอธิบายและจัดองค์ประกอบภาพธรรมดาต่างๆ ไปและวิดีโอ เช่นเดียวกับรูปร่างวิดีโอเป้าหมายอย่างไม่มีลักษณะ การประยุกต์ใช้ที่ได้ประโยชน์

จากแผนที่รูปร่างทวิภาคพร้อมกับภาพเป็นภาพที่มีพื้นฐานเนื้อหา ทดแทนภาพพื้นฐานข้อมูล Interactive Games, surveillance, และ animation

- ‘Gray Scale’ หรือ ‘alpha’ การเข้ารหัสรูปร่าง(Shape Coding) ระบาย alpha จะกำหนดความโปร่งใส ‘transparency’ ของเป้าหมาย ซึ่งไม่จำเป็นต้องเป็นแบบเดียวกัน

- **Robustness in Error Prone Environments**

ความผิดพลาดที่ยืดหยุ่นได้ ยอมให้การเข้าถึงภาพและวิดีโอที่ซ่อนอยู่เหนือขอบเขตที่กว้างขวางของการเก็บและการส่งสื่อ ทั้งนี้รวมถึงการใช้ประโยชน์จากการดำเนินการของลำดับขั้นตอนการบีบอัดภาพและวิดีโอที่อัตราบิดต่ำๆ ในสิ่งแวดล้อมที่มีแนวโน้มจะเกิดความผิดพลาด

- **Face and Body Animation**

เครื่องมือ‘Face and Body Animation’ ตามมาตรฐานยอมให้ส่งพารามิเตอร์ที่จะกำหนดการแสดงเครื่องหมายการวัดและการสังเคราะห์ภาพหน้าและร่างกายที่มีชีวิตชีวาขึ้น รูปแบบจำลองนี้ไม่เป็นมาตรฐานโดย MPEG-4, แต่พารามิเตอร์เป็นมาตรฐาน แม้ว่าจะมีวิธีส่ง เช่น a well-defined face ไปสู่อุปกรณ์ถอดรหัส

เครื่องมือ ต่างๆ ประกอบด้วย

- Definition and coding of face and body animation parameters (model independent):
- Definition and coding of face and body definition parameters (for model calibration):
- Facial texture coding

- **Coding of 2-D Meshes with Implicit Structure**

การเข้ารหัสแบบโครงข่าย 2 มิติ ประกอบด้วย

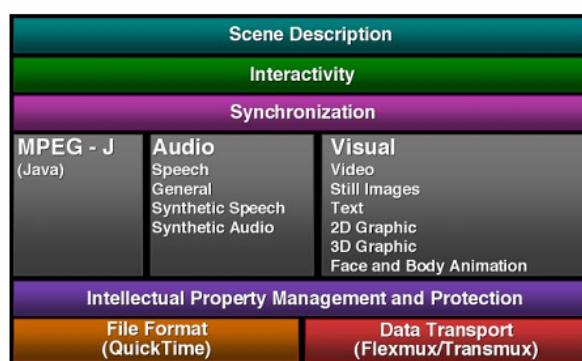
- การทำนาย Mesh-based และการเปลี่ยนรูปร่างที่มีชีวิตชีวา
- การบีบอัดรูปแบบเรขาคณิตสำหรับ เวกเตอร์การเคลื่อนที่

- **Coding of 3-D Polygonal Meshes**

MPEG-4 จัดเตรียมชุดเครื่องมือสำหรับการเข้ารหัส โครงข่ายหลายเหลี่ยมแบบ 3 มิติ มีการใช้โครงข่ายหลายเหลี่ยมทดแทนเป้าหมาย แบบ 3 มิติอย่างกว้างขวาง



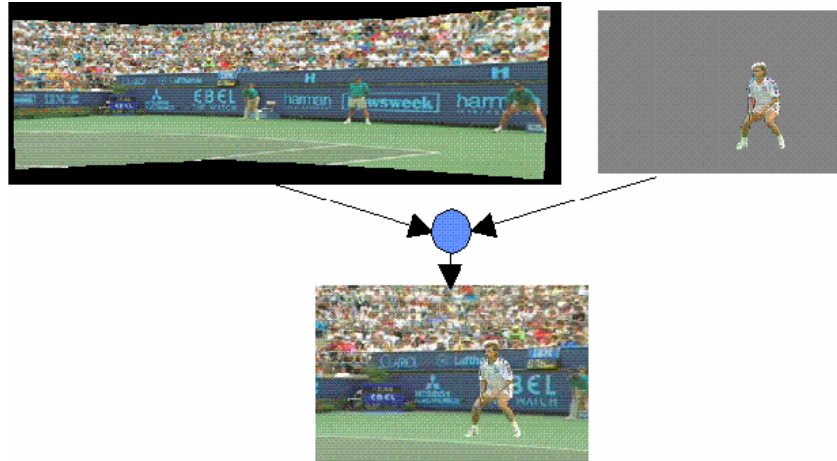
รูปที่ 39 แสดง 2 D Mesh



รูปที่ 40 แสดง MPEG-4 's Component

โดยสรุปแล้ว Video Object จะแบ่งโครงสร้างภาพ เป็นส่วนฉากหลัง (Background) หรือ Sprite มีภาพเป้าหมาย (Object) เป็น Foreground ดังรูปที่ 41 นักเทคนิคคือ Foreground เป้าหมาย สามารถแยกออกมาจากภาพฉากหลัง (Background) หรือ Sprite ตามเหตุการณ์ที่เกิดขึ้นก่อนจะเข้ารหัส ในภาพนี้ที่สปีดภาพฉากหลังเสมือนภาพนิ่ง จะถูกส่งไปยังเครื่องรับเป็นภาพแรกของเหตุการณ์ที่เกิดขึ้นเพียงครั้งเดียวเพื่ออธิบายฉากหลัง ฉากหลัง ยังอยู่ต่อไปเนื่องจากถูกเก็บไว้ในบัฟเฟอร์ Sprite ในแต่ละเฟรมภาพที่ต่อเนื่อง

มาตามลำดับ เฉพาะภาพที่สัมพันธ์กับฉากหลังเท่านั้นที่จะถูกส่งออกไปยังเครื่องรับ โดยยอมให้เครื่องรับสร้างภาพฉากหลังแต่ละเฟรมของเหตุการณ์ที่เกิดต่อเนื่องกัน ตาม Sprite ส่วนภาพForeground เป้าหมายที่เคลื่อนไหวจะแยกส่งตามรูปร่างภาพเป้าหมายที่ไม่มีกฎเกณฑ์ เครื่องรับจะประกอบภาพฉากหลังและภาพForeground เป้าหมายทั้งสองภาพเข้าด้วยกัน เป็นภาพการแข่งขันเทนนิส ที่สมบูรณ์



รูปที่ 41 แสดงตัวอย่าง Sprite Coding of Video Sequence)

การเรียงลำดับภาพในฉากเป้าหมาย ได้รับการระบุจาก Scene Descriptor และ BIFS (Binary for Scene) ทั้งภาพหรือวิดิทัศน์ตามธรรมชาติหรือที่สังเคราะห์ขึ้น ทั้งแบบ 2 มิติ และ 3 มิติ สามารถเข้ารหัสได้เช่นเดียวกัน จากรูปที่ 41 และ รูปที่ 42 สามารถสรุปการเข้ารหัส MPEG-4 ได้ดังนี้ คือ

- **Shape Coding** คือการเข้ารหัสรูปร่าง ที่สามารถเข้ารหัสวิดิทัศน์เป้าหมาย รูปร่างแบบไม่มีกฎเกณฑ์ จากรูปที่ 41 นักเทนนิส จะถูกแยกเข้ารหัสแบบแยกวิดิทัศน์เป้าหมาย รูปร่างของวิดิทัศน์เป้าหมาย ถูกกำหนดโดย Alfa Channel ในตอนแสดงภาพรูปร่างบริเวณขอบภาพที่อยู่จะถูกเข้ารหัสแบบ Gray Scale ขนาดความชัดเจน 8 Bits รูปร่างง่าย ๆ อาจแทนด้วยเป้าหมายแบบทวิภาค (คือมีภาพหรือไม่มี)
- **Texture Coding** คือการเข้ารหัสองค์ประกอบ โดยทั่วไปองค์ประกอบของวิดิทัศน์เป้าหมาย ถูกเข้ารหัสโดย DCT คล้ายกับ H.263, MPEG-1 และ MPEG-2
- **Motion Coding** คือการเข้ารหัสภาพเคลื่อนไหว เพิ่มขึ้นในส่วนการทำนายภาพเคลื่อนไหว ถ้าภาพเคลื่อนไหวทั้งภาพ แบบหันกล้อง(pan) ถ่ายภาพ ไม่ใช่การเข้ารหัสเวกเตอร์เคลื่อนที่ สำหรับแต่ละมาโครบล็อก พารามิเตอร์การเคลื่อนที่ของเป้าหมายทั้งหมดจะถูกส่งออกไป ซึ่งก็เป็นข้อมูลที่กระชับมากๆ

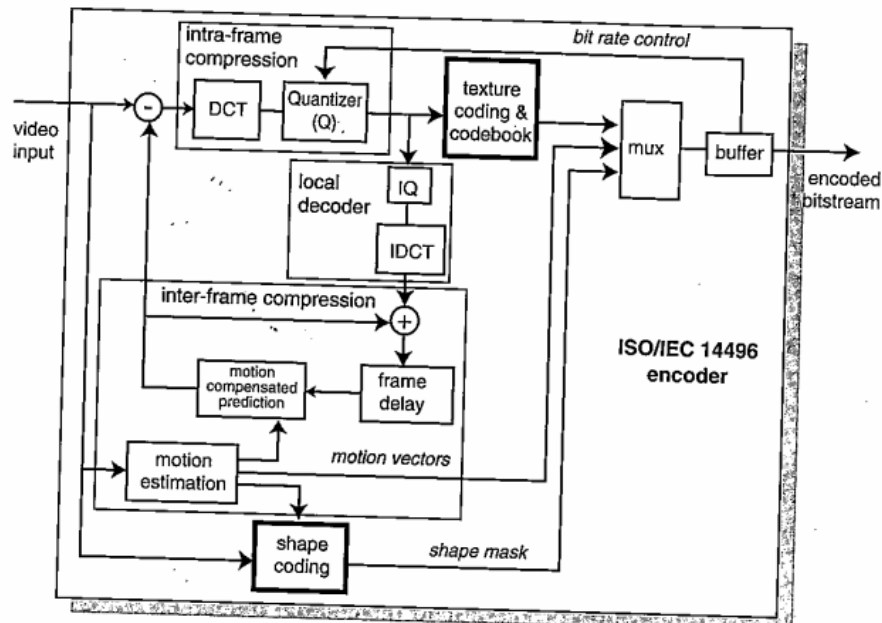
5.4.5 Video Encoder

การกำหนดมาตรฐานอุปกรณ์เข้ารหัส MPEG-4 ก็เพื่อกำหนดมาตรฐานเครื่องเล่นกลับ (Player) หรือสถานีปลายทาง(Terminal) เช่นเดียวกับมาตรฐาน MPEG-1 และ MPEG-2 รายละเอียดสำหรับอุปกรณ์เข้ารหัสนั้น จะยอมให้บริษัทผู้ผลิตพัฒนาเอง ตราบเท่าที่การเข้ารหัสบีตสตรีมสามารถเล่นกลับบนสถานีปลายทาง MPEG-4 ได้

รูปแบบอุปกรณ์เข้ารหัสต่างๆ ไป ดังแสดงตามรูปที่ 42 อุปกรณ์การเข้ารหัสองค์ประกอบ (Texture) และเข้ารหัสรูปร่าง (Shape Coding) เป็นส่วนที่เพิ่มขึ้นมาใหม่ สถาปัตยกรรมของอุปกรณ์การเข้ารหัสขั้นพื้นฐานก็เหมือนกับ MPEG-1 และ MPEG-2 แต่มีประสิทธิภาพดีกว่า 10-15 %

MPEG-4 มีโครงสร้าง (Profile) สำหรับเนื้อหาของภาพตามธรรมชาติหลายแบบ ที่เป็นที่ยอมรับแพร่หลายคือ

- (1) **Simple Profile** เหมาะสำหรับการไร้สายที่ใช้แบนด์วิดท์แคบๆ ใช้เวลาน้อย ประยุกต์ใช้ในกิจการโทรศัพท์ภาพ และ Video-Conference
- (2) **Advanced Simple Profile** สำหรับกิจการสื่อสารระหว่างกัน ที่ใช้แบนด์วิดท์ตั้งแต่ 56 KBit/s จนถึงกิจการโทรทัศน์ความชัดเจนสูงขนาด 8 MBit/s



รูปที่ 42 ระบบของอุปกรณ์เข้ารหัส MPEG-4 Part 2

5.5 H.264 / AVC

H.264 / AVC เป็นโครงการพัฒนาร่วมกันระหว่าง ITU : International Telecommunications Union โดยกลุ่ม Video Coding Experts Group (VCEG) และ ISO/IEC Moving Picture Experts Group (MPEG). รวมกันเป็นกลุ่มทำงานร่วม โดยมีเป้าหมายทำให้เป็นมาตรฐาน ด้วยความพยายามพัฒนาให้ออกแบบการเข้ารหัสง่าย ๆ ไม่ซับซ้อน มีประสิทธิภาพการบีบอัดสูง เชื่อมต่อกับโครงข่ายได้ง่ายแบบ “network-friendly” เช่นการใช้งานวิดีโอทัศน์แบบพูดคุยกัน “conversational” (โทรศัพท์ภาพ) และแบบ “non-conversational” (การเก็บ การบรรดาศ หรือ สตรีมมิ่ง)

มาตรฐานที่ได้จากการทำงานร่วมกัน มีชื่อว่า ITU-T Recommendation H.264|ISO/IEC 14496-10 AVC, หรือ MPEG-4 Part 10 หรือ AVC : Advance Video Coding สำหรับงานบรรดาศ การเก็บ(Storage) หรือการสตรีมมิ่ง หรือเรียกสั้นๆในกิจการโทรคมนาคม คือ H.264.

5.5.2 H.264 / AVC : A Revolution in Quality and Coding Efficiency

มาตรฐาน ITU-T H.264 ได้ถูกนำเสนอมาเพื่อให้ได้ซึ่งประสิทธิภาพการบีบอัดที่สูงกว่าเมื่อเปรียบเทียบกับมาตรฐาน MPEG-1, MPEG-2 และ MPEG-4 Part 2 ที่มีมาก่อนหน้านี้ H.264 มีประสิทธิภาพการเข้ารหัสสูง เช่นมีคุณภาพเท่ากับ MPEG-2 โดยใช้จำนวนอัตราข้อมูลเพียงครึ่งหนึ่ง เมื่อเปรียบเทียบกับ MPEG-4 Part 2 แล้ว H.264 จะได้ภาพที่มีขนาดโตเป็น 4 เท่า เมื่อใช้จำนวนอัตราข้อมูลเท่ากัน

H.264 มีประสิทธิภาพการเข้ารหัสสูงกว่ามาตรฐานที่มีมาก่อนเมื่อได้คุณภาพถึงจุดพิกัดสูงสุด เมื่ออัตรา การบีบอัดสูงขึ้น แทนที่ภาพจะแยกออกเป็นบล็อกๆภาพเป็นหยักๆ เช่นที่เกิดขึ้นกับ มาตรฐาน MPEG-2, H.264 จะค่อยๆลดความชัดเจนลงมา

มาตรฐาน H.264 ถูกพัฒนาขึ้นมาจากมาตรฐานก่อนหน้า ดังนั้นกรรมวิธีในการเข้ารหัสวิดีโอจะอยู่บน พื้นฐานเดียวกัน อันได้แก่ การแปลงเพื่อลดสหสัมพันธ์เชิงพื้นที่ การควอนไทซ์เพื่อควบคุมอัตราบิต การทำนาย การขจัดเซกเมนต์เคลื่อนที่เพื่อลดสหสัมพันธ์เชิงเวลา และการเข้ารหัสเอนโทรปีเพื่อลดสหสัมพันธ์เชิงสถิติ อย่างไรก็ตาม ประสิทธิภาพของการเข้ารหัสจาก H.264 ได้มาจากการเปลี่ยนแปลงในรายละเอียดของกรรมวิธีการเข้ารหัส พื้นฐาน ยกตัวอย่างเช่น การทำนายภาพแบบ intra การแปลงแบบ integer 4x4 การใช้เฟรมอ้างอิงหลายเฟรม และการปรับปรุงการเข้ารหัสแบบเอนโทรปี แต่ประสิทธิภาพที่เพิ่มขึ้นมานั้น ก็ต้องแลกกับความซับซ้อนของตัว เข้ารหัสและตัวถอดรหัสที่เพิ่มขึ้นด้วยเช่นกัน นอกจากนี้ H.264 ยังได้มีการกำหนด profiles และ levels สำหรับการ กำหนดค่าที่เหมาะสมของพารามิเตอร์ต่างๆ ที่ใช้กับงานประยุกต์แบบต่างๆไป

H.264 มีประสิทธิภาพการเข้ารหัสสูงกว่าทุกมาตรฐานที่เคยมี สามารถประยุกต์ใช้งานได้ตั้งแต่งาน บรอดคาสและ DVD ซึ่งปัจจุบันใช้ MPEG-2 จนถึงการประชุมสัมมนาทางวิดีโอ ซึ่งใช้ H.263 ถึง VOD วิดีทัศน์ ตามสั่ง สตรีมมิ่งและข่าวสารสื่อสารซึ่งปัจจุบันใช้ MPEG-4 Part 2 ทั้งหมดสามารถใช้ H.264 แทนได้แบบมี คุณภาพและประสิทธิภาพที่ดีกว่า พูดได้ว่าใช้ได้ตั้งแต่ 3G ถึง HD

5.5.2 H.264 / AVC Application

H.264 / AVC มีการกำหนดทาง syntax ที่เรียบง่าย มีการเชื่อมต่อกับโครงสร้างโปรโตคอลและการ มัลติเพล็กซ์ผ่าน โครงข่ายที่มีการใช้งานในปัจจุบัน เพราะฉะนั้น H.264 / AVC สามารถสนับสนุนการประยุกต์ใช้ ได้มากมาย เช่น

- บรอดคาสผ่านเคเบิล, ดาวเทียม, เคเบิล โมเด็ม, DSL, บรอดคาสภาคพื้นดิน
- Interactive หรือการจัดเก็บลงบนอุปกรณ์ที่บันทึกด้วยแสงหรือด้วยสนามแม่เหล็ก, DVD
- การบริการแบบพุดคุยกัน บน ISDN, Ethernet, LAN, DSL, ไร้สายและโครงข่ายเคลื่อนที่ หรือหลายแบบผสมกัน
- บริการวิดีโอแบบสั่งได้ VOD, หรือบริการสตรีมมิ่งสื่อสาร บน ISDN, เคเบิล โมเด็ม, LAN, DSL, โครงข่ายไร้สาย
- MMS บน ISDN, DSL, Ethernet, LAN, ไร้สายและโครงข่ายเคลื่อนที่

เนื่องจาก H.264 / AVC ถูกออกแบบให้ครอบคลุม Video Coder Layer : VCL เพื่อแทนที่เนื้อหาวิดีโอ อย่างมีประสิทธิภาพ และครอบคลุม Network Abstraction Layer : NAL เพื่อจัดรูปแบบ VCL แทนที่เนื้อหา วิดีทัศน์และจัดให้มีข่าวสารข้อมูลนำหน้าด้วยวิธีการที่เหมาะสมเพื่อการขนส่งโดยชั้นการขนส่งหรือสื่อสำหรับเก็บ ข้อมูล ดังรูปที่ 43

5.5.4 Network Abstraction Layer : NAL

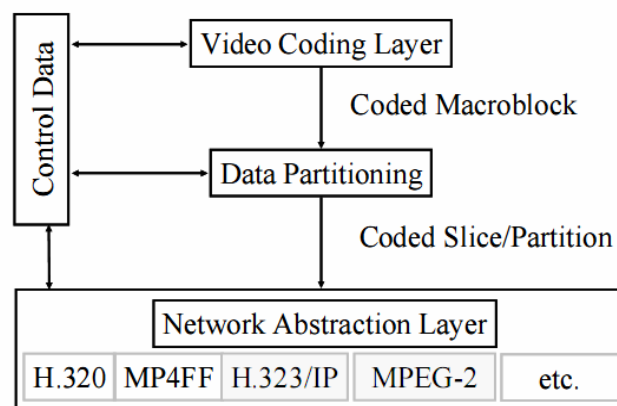
VCL ซึ่งจะได้อธิบายต่อไป ถูกเจาะจงให้แทนเนื้อหาวิดีโออย่างมีประสิทธิภาพ ส่วน NAL เพื่อ จัดรูปแบบ VCL แทนที่เนื้อหาวิดีโอและจัดให้มีข่าวสารข้อมูลนำหน้าด้วยวิธีการที่เหมาะสมเพื่อการขนส่งโดย ชั้นการขนส่ง(Transport) หรือสื่อสำหรับเก็บข้อมูล(Storage Media) ข้อมูลทั้งหมดอยู่ใน NAL แต่ละ ส่วนประกอบด้วย bytes ที่มีจำนวนเต็ม

โดยทั่วไปหน่วยของ NAL จะจําแนกรูปแบบสําหรับใช้ทั้งแบบบรรจุหีบห่อ (packet) และระบบบีบอัดสทริม ทั้งสองแบบเหมือนกันยกเว้น แต่ละหน่วยของ NAL สามารถนํานํานาโดยรหัสเริ่มต้นในชั้นการขนส่งระบบบีบอัดสทริม

NAL ถูกออกแบบเพื่อเตรียม “Network Friendliness” เพื่อให้สามารถทำตามคำสั่ง VCL : Video coding layer ได้ง่ายและมีประสิทธิภาพ สําหรับระบบที่มีหลายรูปแบบอย่างกว้างขวาง

NAL ทำให้ความสามารถในการจัดวางข้อมูล H.264 / AVC VCL บนชั้นการขนส่ง ต่างๆ(ตามรูป 43) เช่น

- RTP/IP สําหรับทุกชนิดของการบริการ Internet แบบเวลาจริงทั้งระบบข่ายสายและไร้สาย (Conversation and Streaming)
- รูปแบบ File เช่น ISO MP4 เพื่อเก็บในหน่วยความจํา และ MMS
- H.32X สําหรับการบริการแบบโต้ตอบ (Conversational) ในระบบข่ายสายและไร้สาย
- MPEG-2 สําหรับการกํากับการกระจายเสียงและแพรวภาพ



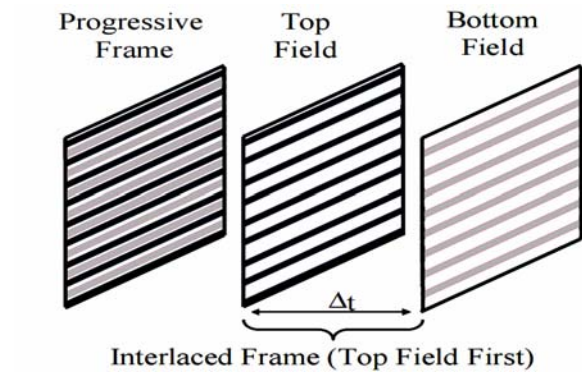
รูปที่ 43 Structure of H.264 / AVC

5.5.4.1 Video coding layer

ชั้นการเข้ารหัสวิดีโอของ H.264/AVC มีเจตนาเหมือนกับ MPEG-2 ซึ่งประกอบด้วยการเข้ารหัสแบบผสม(Hybrid) บนพื้นฐานบล็อกภาพ แบบ Luma และ Chroma ที่เรียกว่ามาโรบล็อก โดยการทำนายเชิงเวลา (Temporal) และเชิงพื้นที่ (spatial) ที่เกิดขึ้นพร้อมกันกับการแปลงรหัส (ดังแสดงในรูป 48)

(A) Picture, frame and field

H.264/AVC ออกแบบเพื่อสนับสนุนการเข้ารหัสวิดีโอ แบบ 4:2:0 ทั้งภาพ Progressive หรือ interlaced เช่นเดียวกับ MPEG-2 ซึ่งอาจจะผสมกันในช่วงเหตุการณ์เดียวกัน ประสิทธิภาพภาพจะมีสองฟิลด์ สอดแทรกกันอยู่ แต่แยกกันตามช่วงเวลาแต่ละฟิลด์ (เวลาครึ่งหนึ่งของเฟรม) อาจจะแยกกันเข้ารหัสเป็นสองฟิลด์ หรือรวมกันเข้ารหัสเป็นเฟรมเดียว ส่วนภาพ Progressive จะเข้ารหัสเป็นเฟรม แต่ยังคงพิจารณาเป็นสองฟิลด์ ณ ช่วงเวลาหนึ่ง



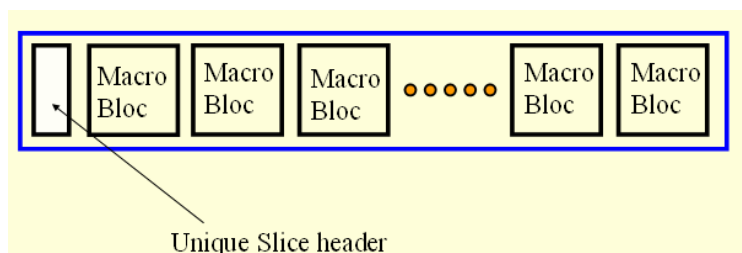
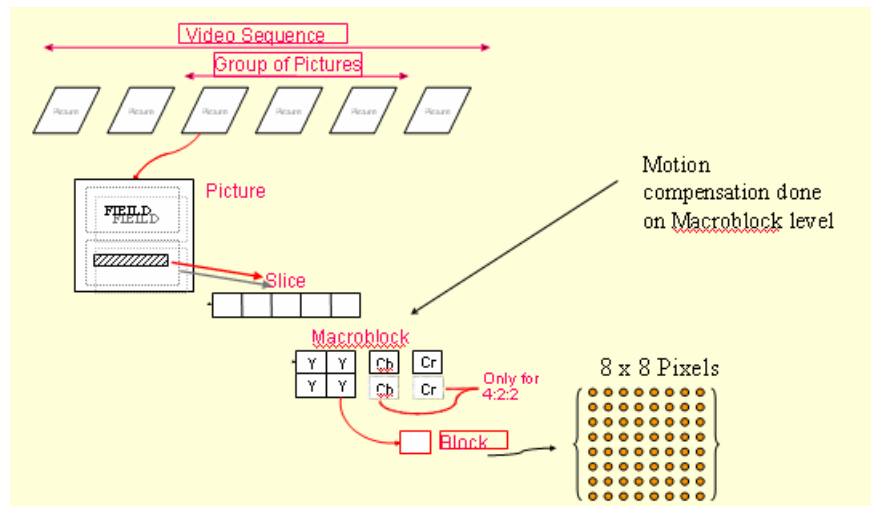
รูปที่ 44 Progressive and interlaced frames and fields

(B) YCbCr color space and 4:2:0 sampling

เช่นเดียวกับ MPEG-2 มาตรฐาน H.264 /AVC ใช้ระยะห่างสัญญาณสี YCbCr ที่ลดความชัดเจนสัญญาณสี คือ 4:2:0 ตามความสามารถรับรู้ส่วนภาพสีของตามนุษย์ โดยเรียกสัญญาณ Y ว่า Luma แทนความสว่างของภาพ และเรียก Chroma component Cb Cr ว่า Chroma แทนสีต่างๆในภาพ

(C) Division of the picture into macroblock

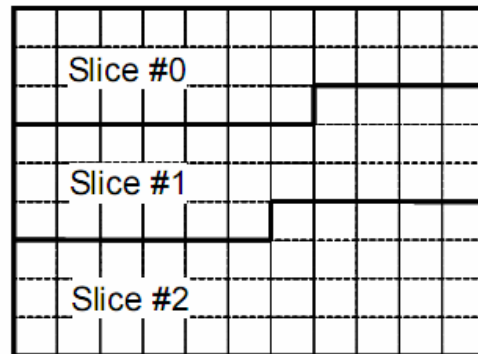
โดยทั่วไปสัญญาณวิดีโอที่เข้ามาในแต่ละเฟรมหรือฟิลด์จะถูกแบ่งให้อยู่ในรูปแบบของบล็อก ที่เหลี่ยมผืนผ้าขนาด 16X16 จุดภาพ สำหรับสัญญาณความสว่าง(Y:Luma) และบล็อกขนาด 8X8 จุดภาพ สำหรับสัญญาณสี(Chroma) ทั้งสองสัญญาณ ที่เรียกว่ามาโครบล็อก (Macroblock) ในแต่ละมาโครบล็อกจะทำการเข้ารหัสภายในเฟรม (Intra frame Coding) ซึ่งเป็นการลดความซ้ำซ้อนเชิงพื้นที่ หรือ การเข้ารหัสระหว่างเฟรม (Inter Coding) ซึ่งเป็นการลดความซ้ำซ้อนเชิงเวลา เป็นวิธีเดียวกันที่ใช้ใน H.261, MPEG-1&2



รูปที่ 45 Macroblock in Slice

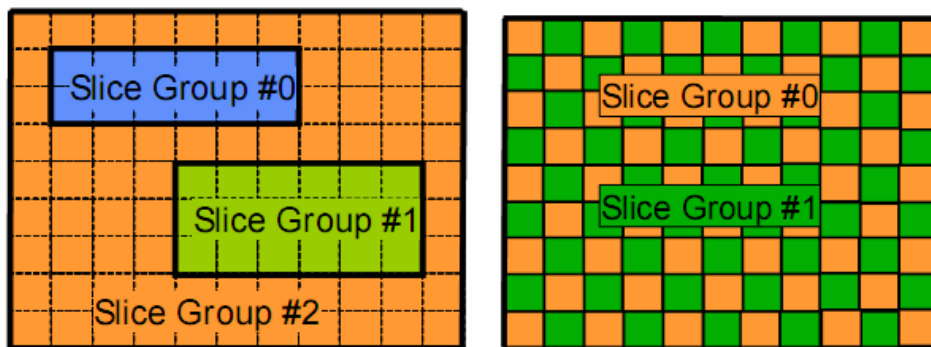
(D) *Slice and slice group*

Slice คือผลการเรียงลำดับมาโครบล็อกตามการสแกนภาพเมื่อไม่ใช้ FMO: Flexible macroblock ordering



รูปที่ 46 Subdivision of a picture into slices when not using FMO.

มาโครบล็อกจะถูกรวบรวมเป็น slices เพื่อแทนชั้นเซตของภาพซึ่งจะถอดรหัสด้วยตัวเอง ลำดับการส่งของมาโครบล็อกในบิตสตรีมขึ้นอยู่กับ Macroblock Allocation Map และไม่จำเป็นต้องเรียงลำดับแบบเส้นสแกน H.264/AVC สนับสนุน slice-coding ที่แตกต่างกัน 5 วิธี ไม่ว่าจะเป็น FMO หรือไม่ คือ



รูปที่ 47 Subdivision of a QCIF frame into slices when utilizing FMO

- เฟรมชนิด I slice (I คือ intra). จะทำการเข้ารหัสภายในเฟรมเท่านั้น ไม่อ้างอิงกับภาพอื่น
- เฟรมชนิด P slice (Predicted Frame) ที่เพิ่มขึ้นมาจากการเข้ารหัสแบบ I slice มาโครบล็อกบางส่วนของ P- slice สามารถเข้ารหัสแบบระหว่างเฟรมทำนาย (Inter-prediction) อย่างน้อยหนึ่งสัญญาณทำนายการขจัดเซกการเคลื่อนที่ (Motion Estimation) ต่อหนึ่งบล็อกการทำนาย
- เฟรมชนิด B slice (Bidirectional Frame) ที่เพิ่มขึ้นมาจากการเข้ารหัสแบบ P slice มาโครบล็อกบางส่วนของ B slice สามารถเข้ารหัสแบบระหว่างเฟรมทำนาย (Inter-prediction) อย่างน้อยสอง สัญญาณทำนายการขจัดเซกการเคลื่อนที่ ต่อหนึ่งบล็อกการทำนาย

เฟรมทั้ง 3 แบบ ที่กล่าวมาเหมือนกับมาตรฐานการบีบอัด MPEG-1&2 ส่วนอีก 2 วิธีคือ

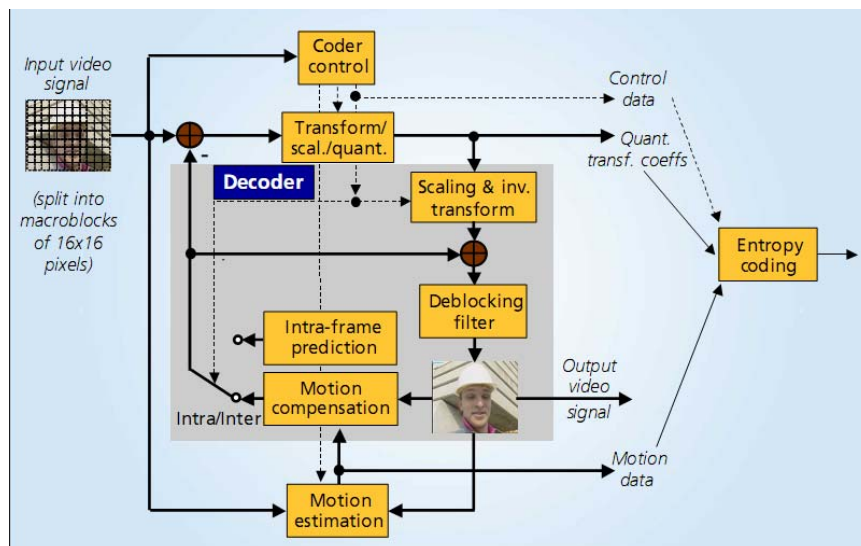
SP slice (switching P) ซึ่งเจาะจงเพื่อประสิทธิภาพการสวิตช์ระหว่างการเข้ารหัสบิตสตรีมที่มีหลายอัตรา บิต สัญญาณการทำนายบิตสตรีมใน Inter Frame สำหรับ SP เฟรมที่เลือกไว้จะถูกควอนไทซ์ในโดเมนการแปลง บีบให้อยู่ในขอบเขตแอมพลิจูดแบบหยวบๆ ซึ่งขอบเขต

- แอมพลิจูดแบบหลายๆจะยอมให้มีการเข้ารหัสแบบอัตราบิตต่ำๆของสัญญาณแตกต่างระหว่างบิตสตรีม
- SI slice เรียกว่า switching I, ซึ่งยอมให้มีการจับคู่ที่สมบูรณ์แบบ(exact match)ของมาโครบล็อกในเฟรม SP slice โดยมีวัตถุประสงค์เพื่อการเข้าถึงแบบสุ่มและเพื่อแก้ไขความผิดพลาด เมื่อการทำนาย Inter เฟรมไม่สามารถใช้ได้เนื่องจากความผิดพลาดในการส่งสัญญาณ

จึงวิธีที่มีประสิทธิภาพสำหรับปกปิดช่องสัญญาณที่มีแนวโน้มจะมีความผิดพลาด ที่ประยุกต์ใช้แบบเสียเวลาน้อย (low delay) คือวิธีที่เรียกว่า FMO : Flexible Macroblock Ordering

(E) Encoding and decoding process for macroblock

โดยสรุป ภาพจะแบ่งเป็นบล็อก ภาพแรกของภาพที่ต่อเนื่องกันหรือจุดแรกที่เข้าถึงแบบสุ่มตามแบบฉบับการเข้ารหัส “Intra” ตัวอย่าง เช่น ไม่ใช้ข้อมูลอื่นนอกจากข้อมูลในภาพนั้นๆ แต่ละตัวอย่างของบล็อกในเฟรม Intra จะถูกทำนายเชิงพื้นที่โดยใช้ตัวอย่างข้างเคียงของบล็อกที่เข้ารหัสก่อนหน้านี้ กระบวนการเข้ารหัสจะเลือกตัวอย่างข้างเคียงอันไหนและอย่างไรสำหรับทำนายเชิงพื้นที่ ซึ่งจะถูกนำไปพร้อมๆกันสู่อุปกรณ์เข้ารหัสและอุปกรณ์ถอดรหัสจะใช้ข้อมูลการทำนาย Intra ที่ส่งมา



รูปที่ 48 shows a block diagram of the video coding layer for a macroblock

สำหรับภาพที่เหลือของภาพที่ต่อเนื่องกัน หรือภาพระหว่างจุดสุ่มภาพตามแบบฉบับการเข้ารหัส “Intra” ได้ใช้ การเข้ารหัสเฟรม Inter ใช้การทำนาย (motion compensation) จากภาพที่ถอดรหัสก่อนหน้านี้กระบวนการเข้ารหัสการทำนายเฟรม Inter (motion estimation) ประกอบด้วยทางเลือกข้อมูลการเคลื่อนที่ รวมถึงภาพอ้างอิงและเชิงพื้นที่ที่เคลื่อนออกไป ซึ่งจะใช้กับตัวอย่างทั้งหมดของบล็อก ข้อมูลการเคลื่อนที่ซึ่งจะถูกส่งแบบข้อมูลข้างเคียงจะถูกนำไปใช้โดยอุปกรณ์เข้ารหัสและอุปกรณ์ถอดรหัส ที่จัดหาสัญญาณการทำนายไว้ให้พร้อมกัน

ส่วนที่เหลือจากการทำนาย (ทั้ง Intra หรือ Inter) ซึ่งแตกต่างกันระหว่างต้นฉบับและบล็อกทำนาย จะถูกแปลง สัมประสิทธิ์ในการแปลงจะถูกประเมินและแบ่งนัย สัมประสิทธิ์ในการแปลงที่ถูกแบ่งนัย(quantized)จะเข้ารหัสเอนโทรปี และถูกส่งพร้อมกับข้อมูลข้างเคียง ทั้งการทำนายแบบ Intra Frame และการทำนายใน Inter Frame

ในอุปกรณ์เข้ารหัสจะมีอุปกรณ์ถอดรหัส เพื่อนำไปสู่การทำนายสำหรับบล็อกต่อไปหรือภาพต่อไป ดังนั้น สัมประสิทธิ์ในการแปลงที่ถูกแบ่งนับ จะเป็นส่วนกลับของการประเมินและส่วนแปลงกลับ(Inverse Transform) เช่นเดียวกับด้านอุปกรณ์ถอดรหัส มีผลให้การทำนายผลการถอดรหัสยังคงเหลืออยู่ ซึ่งจะถูกนำไปรวมกับการทำนายอีก ผลการรวมจะส่งกลับไป Deblocking filter เพื่อเตรียมสัญญาณทางออกเป็น วิดิทัศน์

(F) Adaptive frame/field coding operation

ใน Interlace เฟรม ที่ภาพบางส่วนมีการเคลื่อนที่ หรือกล้องเคลื่อนไหว เส้นแสมกนสองแถวที่อยู่ติดกันจะ แสดงการลดความแตกต่างการพึ่งพาทางสถิติลงมา เมื่อเปรียบเทียบกับเฟรมภาพ Progressive ในกรณีนี้การแยก ฟิลด์กันเพื่อบีบอัดสัญญาณ น่าจะมีประสิทธิภาพสูงกว่า H.264 / AVC จึงยอมให้ Encoder ตัดสินใจเข้ารหัสเฟรม 3 แบบ คือ

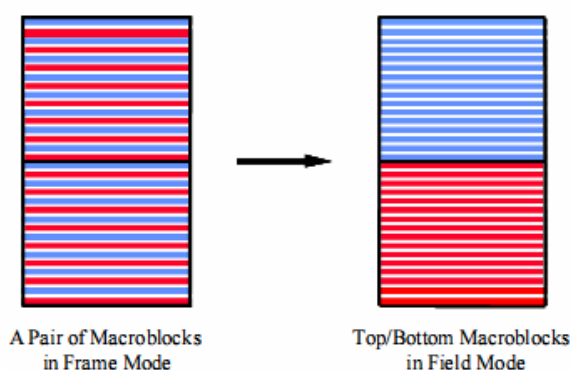
1. รวม 2 ฟิลด์เข้าด้วยกัน แล้วเข้ารหัสแบบเฟรม (frame mode)
2. เข้ารหัสแบบแยกฟิลด์ (field mode)
3. รวม 2 ฟิลด์เข้าด้วยกัน แล้วเข้ารหัสแบบเฟรม แต่ตอนเข้ารหัส ให้แยกคู่ของมาโครบล็อกที่อยู่ ติดกัน เป็นคู่ของ 2 ฟิลด์ หรือมาโครบล็อกเฟรมก่อนเข้ารหัส

การเลือก ข้อ 1 และ 2 คือการเข้ารหัสแบบ Picture Adaptive Frame / Field (PAFF) เมื่อเฟรมถูกแยกเข้ารหัส เป็น 2 ฟิลด์ แต่ละฟิลด์แบ่งเป็นมาโครบล็อกและเข้ารหัสเช่นเดียวกับเฟรม ยกเว้นในกรณีเหล่านี้

- i. การขดเชยการเคลื่อนที่ ใช้การอ้างอิงกับฟิลด์ มากกว่าการอ้างอิงกับเฟรม
- ii. การซิกแซกแสมกน สัมประสิทธิ์ของการแปลง (Transform coefficients) จะแตกต่างออกไป
- iii. การดำเนินการเกิดบล็อกอย่างรุนแรง จะไม่นำมาใช้ในการกรองสัญญาณขอบภาพทางแนวนอนของมาโครบล็อกในฟิลด์ภาพ เนื่องจากแถวของภาพในฟิลด์ มีเนื้อที่เป็นสองเท่าตราบเท่าที่แถวของภาพในเฟรมและความยาวของการกรองสัญญาณที่ครอบคลุมเนื้อที่ส่วนใหญ่

ระหว่างการพัฒนา H.264/AVC มีรายงานว่า PAFF สามารถลดขนาดข้อมูลได้ 16-20% เมื่อเปรียบเทียบกับ ใช้การเข้ารหัสแบบเฟรมอย่างเดียว

ถ้าเฟรมประกอบด้วยภาพที่เคลื่อนไหวและภาพอยู่กับที่ผสมกัน โดยทั่วไปการเข้ารหัสจะมีประสิทธิภาพสูงกว่าหากเข้ารหัสภาพอยู่กับที่มีแบบเฟรมโหมด และภาพเคลื่อนที่แบบฟิลด์โหมด ดังนั้นการตัดสินใจเข้ารหัสแบบฟิลด์ / เฟรม จึงสามารถแยกอย่างอิสระสำหรับคู่ของมาโครบล็อกทางแนวดิ่ง (16x32 ในเขต Luma) ในเฟรมภาพ เข้ารหัสแบบนี้คือ Macroblock-Adaptive Frame / Field (MBAFF)



รูปที่ 49 Conversion of a frame macroblock pair into a field macroblock pair

ไม่เหมือนกับ MPEG-2 ที่การตัดสินใจเข้ารหัสแบบฟิลด์ / เฟรม เกิดขึ้นที่กับคู่ของของมาโครบล็อก มากกว่าในระดับมาโครบล็อก

เนื่องจากเฟรมที่ประกอบด้วยภาพที่เคลื่อนไหวและภาพอยู่กับที่ผสมกัน เกิดขึ้นในเฟรม MBAFF ดังนั้นวิธีการซึ่งจะใช้กับหัวข้อต่อไปนี้จะต้องปรับเปลี่ยนให้เหมาะสม วิธีการเหล่านั้น คือ

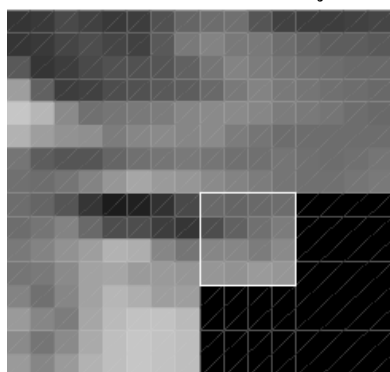
1. Zig-zag scanning
2. prediction of motion vectors
3. prediction of intra prediction Mode
4. intra frame sample prediction
5. deblocking filtering
6. context modeling entropy coding

ระหว่างการพัฒนา H.264/AVC มีรายงานว่า MBAFF สามารถลดขนาดข้อมูลได้ 14-16%

(G) Intra-frame Prediction

แต่ละมาโครบล็อกสามารถส่งโดยการเข้ารหัสแบบเดียวหรือเข้ารหัสหลายแบบแล้วแต่นิยามการเข้ารหัสของ slice สนับสนุนการเข้ารหัสภายในเฟรม (Intra) แบบ 4x4 หรือแบบ 16x16 พร้อมการทำนาย Chroma และการทำนายโหมด I_PCM

สำหรับบล็อกส่วนประกอบความสว่างขนาด 4X4 จุดภาพที่จะทำการเข้ารหัสแสดงดังรูปที่ 50 โดยแต่ละจุดภาพ (a ถึง p) สามารถแสดงดังรูปที่ 51



Q	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				

รูปที่ 50 ส่วนประกอบความสว่าง (Luma) ขนาด 4X4 รูปที่ 51 ตำแหน่งส่วนประกอบความสว่างขนาด 4X4

เมื่อใช้โหมด Intra แบบ 4x4 แต่ละ 4x4 บล็อกจะถูกทำนายจากตัวอย่างเชิงพื้นที่ข้างเคียง 16 ตัวอย่างคือ จาก a – p ของ 4X4 บล็อกจะถูกทำนายจากตัวอย่างข้างเคียง คือ A-Q ที่ถอดรหัสแล้ว

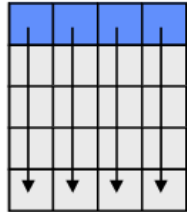
สำหรับมาตรฐานการเข้ารหัสสัญญาณวิดีโอ H.264 การเข้ารหัสภายในเฟรมจะแบ่งออกเป็นบล็อกส่วนประกอบความสว่าง(Luma) ขนาด 4X4 จุดภาพทั้งหมด 9 โหมดและบล็อกส่วนประกอบความสว่าง(Luma) ขนาด 16x16 จุดภาพ ทั้งหมด 4 โหมดและสำหรับบล็อกส่วนประกอบสี(Chroma)ขนาด 4x4 รวม 4 โหมด โดยการเลือกโหมดของการเข้ารหัสเลือกจากการหาค่าความแตกต่างสัมบูรณ์รวม (Sum Absolute Difference Error) ที่น้อยที่สุด

รูปที่ 52 แสดง 5 แบบ ใน 9 แบบ โหมดการทำนาย Intra 4x4 โดยแต่ละโหมดมีรายละเอียดดังนี้

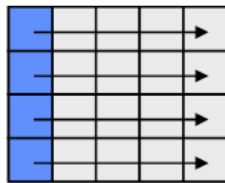
- โหมด 0 สามารถคำนวณจากการประมาณในแนวดิ่งจากจุดภาพ A B C และ D
- โหมด 1 สามารถคำนวณจากการประมาณในแนวนอนจากจุดภาพ I J K และ L
- โหมด 2 การทำนาย DC ตัวอย่างข้างเคียงถูกเฉลี่ย ตามที่แสดงในภาพ

- โหมด 3 สามารถคำนวณจากแนวทแยงซ้ายที่ 45 องศา จากซ้ายลงมาทางขวา
- โหมด 4 สามารถคำนวณจากแนวทแยงขวาที่ 45 องศา จากขวาลงมาทางซ้าย
- โหมดที่เหลือ 5-8 สามารถคำนวณแนวทแยง แนวตั้งทางขวาจากซ้ายลงมาทางขวา, ต่ำกว่าแนวนอนจากซ้ายลงมาทางขวา, แนวตั้งทางขวาจากขวาลงมาทางซ้าย, สูงกว่าแนวนอนจากขวาลงมาทางซ้าย

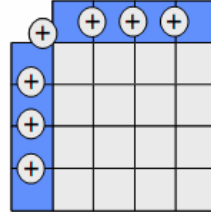
Mode 0 - Vertical



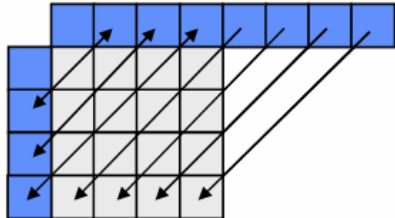
Mode 1 - Horizontal



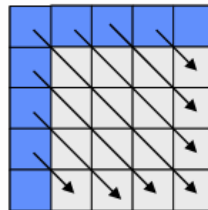
Mode 2 - DC



Mode 3 - Diagonal Down/Left



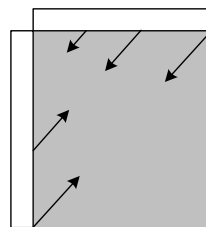
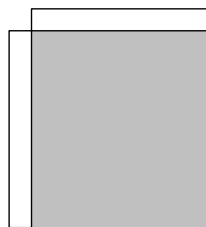
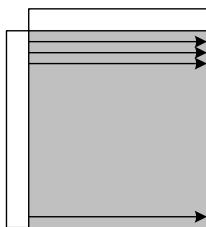
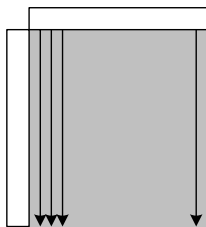
Mode 4 - Diagonal Down/Right



รูปที่ 52 Five of the nine Intra_4x4 prediction modes.

สำหรับบล็อกส่วนประกอบความสว่างขนาด Intra 16X16 สามารถหาได้จาก 4 โหมด รายละเอียดดังนี้

- โหมด 0 สามารถคำนวณจากการประมาณในแนวตั้ง H
- โหมด 1 สามารถคำนวณจากการประมาณในแนวนอน V
- โหมด 2 สามารถคำนวณจากค่าเฉลี่ยจากกลุ่มจุดภาพแนวตั้ง H และแนวนอน V เช่นเดียวกับแบบ Intra 4x4
- โหมด 3 สามารถคำนวณจากกลุ่มจุดภาพแนวนอน H ในแนวทแยงจากขวามาซ้ายและกลุ่มจากจุดภาพแนวตั้ง V ในแนวทแยงจากซ้ายไปขวา



รูปที่ 53 เข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16X16 โหมด 0 ถึง 3

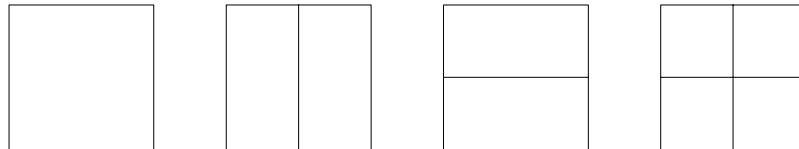
สำหรับการเข้ารหัสภายในเฟรมของส่วนประกอบความเข้มสีขนาด 8X8 มีทั้งหมด 4 โหมดลักษณะคล้ายกับการเข้ารหัสภายในเฟรมของส่วนประกอบความส่องสว่างขนาด 16x16

(H) *Inter-frame Prediction*

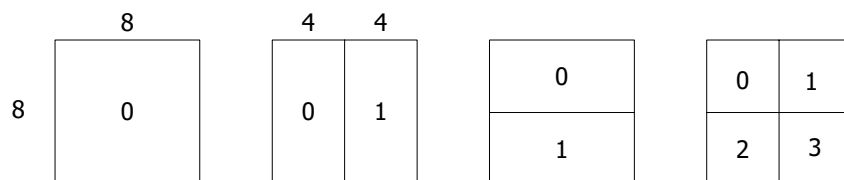
(L) Inter-frame Prediction in P Slice

ส่วนที่เพิ่มขึ้นมาจากการเข้ารหัสแบบภายในเฟรม Intra มาโครบล็อก การทำนายหรือการเข้ารหัสแบบการ
 ชดเชยการเคลื่อนที่ ถูกกำหนดเป็น มาโครบล็อกแบบ P

โดยทั่วไปขนาดของบล็อกที่ใช้ในการชดเชยการเคลื่อนที่แบบ P มีขนาด 16x16 หรือ 8x8 จุดภาพ ซึ่งใน
 มาตรฐาน H.264/AVC สำหรับบล็อกที่มีขนาด 16x16 สามารถแบ่งขนาดบล็อกให้มีขนาดเล็กได้ทั้งหมด 4 แบบ
 คือ 16x16 16x8 8x16 8x8 แสดงดังรูปที่ 54 และ สำหรับบล็อกที่มีขนาด 8x8 สามารถแบ่งให้มีขนาดเล็กได้อีก 4
 แบบ คือ 8x8 8x4 4x8 และ 4x4 แสดงดังรูปที่ 55

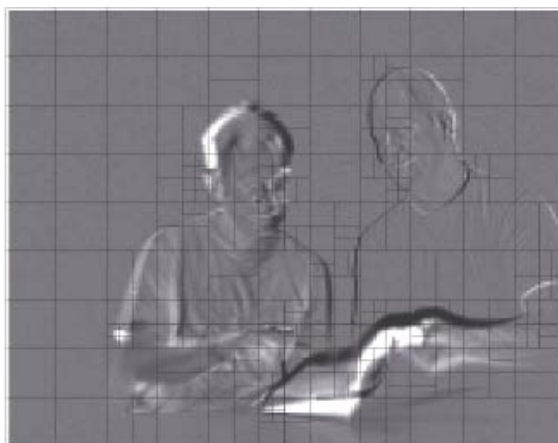


รูปที่ 54 การแบ่งบล็อกขนาด 16X16



รูปที่ 55 การแบ่งบล็อกขนาด 8X8

สำหรับการเลือกขนาดของบล็อกโดยทั่วไปไปบล็อกขนาดใหญ่เหมาะสำหรับพื้นที่ที่มีลักษณะเอกพันธ์
 (Homogeneous) และบล็อกขนาดเล็กเหมาะสำหรับพื้นที่ที่มีรายละเอียดมาก **16** บล็อกแสดงดังรูปที่ 56 **8** **8**



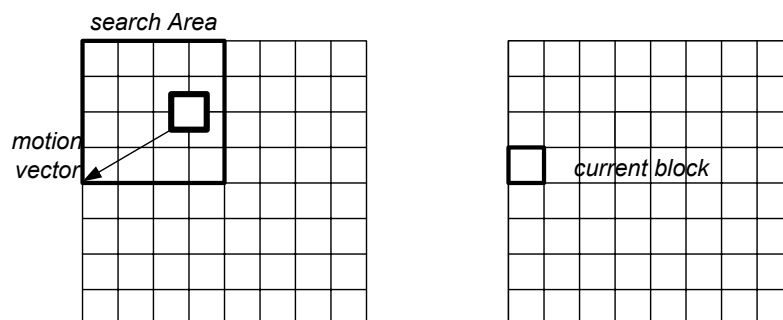
0 **0** **1**

รูปที่ 56 แสดงการแบ่งบล็อกภายในภาพ

เมื่อทำการหาบล็อกที่มีจุดภาพใกล้เคียงมากที่สุดระหว่างเฟรมปัจจุบันและเฟรมก่อนหน้า ในการค้นหา
 บล็อกที่ใกล้เคียงสามารถหาได้ในพื้นที่การค้นหา (Search area) โดยวิธีที่ใช้วัดความใกล้เคียงระหว่างบล็อกทั้ง
 สองนั้นมีหลายวิธี เช่น MAD (Mean Absolute Difference) SSD (Sum Square Difference) SAD (Sum Absolute
 Difference) แสดงดังรูปที่ 57 โดยตำแหน่งที่เปลี่ยนแปลงเรียกว่า เวกเตอร์การเคลื่อนที่

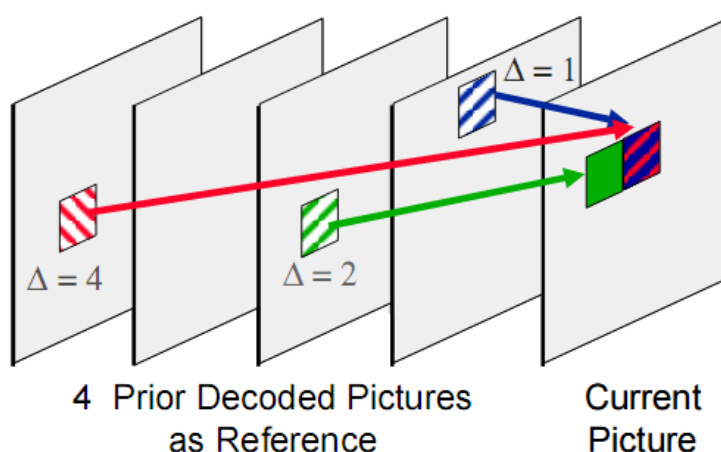
สัญญาณการทำนายสำหรับการเข้ารหัสการทำนาย M x N บล็อก Luma แต่ละบล็อกจะได้มาโดยการ
 แทนพื้นที่ภาพอ้างอิงที่กำหนดโดยการแปลงเวกเตอร์การเคลื่อนที่และค่าอ้างอิงภาพ นั่นคือหากเข้ารหัสมาโคร

บล็อกโดยแบ่งบล็อก 8x8 ออกเป็นสี่ส่วน แต่ละส่วนของ 8x8 ถูกแบ่งออกเป็น 4x4 จำนวนสูงสุด คือ 16 เวกเตอร์ การเคลื่อนที่ จะถูกส่งในหนึ่ง P มาโครบล็อก



รูปที่ 57 การประมาณการเคลื่อนที่

สำหรับมาตรฐาน H.264 /AVC ได้เพิ่มระดับการประมาณการเคลื่อนไหวจากเดิมแบบครึ่งจุดภาพ (Half-pixel) เป็น 1/4 และ 1/8 จุดภาพซึ่งสามารถทำให้ประมาณการเคลื่อนที่ได้แม่นยำขึ้นสำหรับการประมาณการเคลื่อนไหวจากเดิมแบบครึ่งจุดภาพ



รูปที่ 58 Multi-frame motion compensation. In addition to the motion vector, also picture

reference parameters Δ are transmitted. The concept is also extended to B slices as described below.

การทำนายการชดเชยการเคลื่อนที่ แบบ Multi-frame ต้องให้ทั้ง Encoder และ Decoder ต้องเก็บเฟรมภาพอ้างอิง เพื่อใช้สำหรับการทำนายระหว่างเฟรม ไว้ใน Buffer ที่สามารถเก็บได้หลายภาพ Decoder จะทำสำเนาภาพที่อยู่ใน Buffer ที่สามารถเก็บได้หลายภาพ ของ Encoder ตามการจัดการส่วนความจำที่ควบคุมการทำงาน ที่ถูกกำหนดโดยบิตสตรีม พารามิเตอร์ดัชนีอ้างอิงถูกส่ง ทุกๆ การชดเชยการเคลื่อนที่ แบบ 16x16, 16x8, 8x16, หรือ 8x8 Luma บล็อก

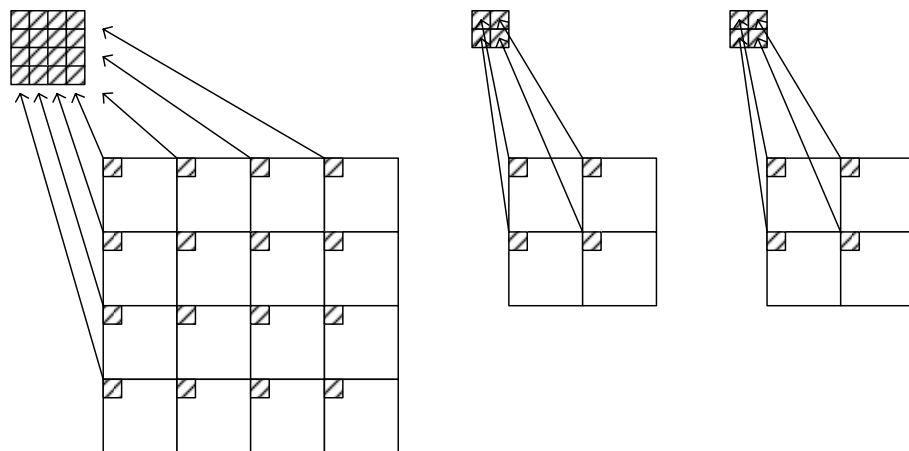
(L) Inter-frame Prediction in B Slice

เมื่อเปรียบเทียบกับมาตรฐานบีบอัดสัญญาณก่อนหน้านี้ ความคิดเกี่ยวกับ B Slice จะถูกกล่าวทัวๆ ไปใน H.264 /AVC กล่าวคือภาพอื่นๆ สามารถอ้างอิง B Slice เพื่อการทำนายการชดเชยการเคลื่อนที่ซึ่งขึ้นอยู่กับ การจัดการส่วนความจำที่ควบคุมการทำงาน ของ Buffer ที่สามารถเก็บได้หลายภาพ ความแตกต่างอย่างมากย ระหว่าง B และ P Slice คือ B Slice ถูกเข้ารหัสโดยที่บางมาโครบล็อกหรือบล็อก จะใช้ค่าเฉลี่ยของการทำนายการชดเชยการเคลื่อนที่ ที่แตกต่างกันอย่างชัดเจน 2 ค่า มาถ่วงน้ำหนัก เพื่อสร้างสัญญาณชดเชยการเคลื่อนที่

B Slice ใช้การเข้ารหัสแบบแยกมาโครบล็อก แบบเดียวกับ P Slice นอกจาก แบบ 16x16, 16x8, 8x16, หรือ 8x8 และการเข้ารหัสแบบภายในเฟรมแล้ว ยังมีการเข้ารหัสแบบการทำนายแบบสองทิศทางและการทำนายแบบทางตรงอื่นๆ ด้วย สำหรับ แบบ 16x16, 16x8, 8x16, หรือ 8x8 แต่ละแบบ สามารถแยกเลือกวิธีเข้ารหัสได้

(I) Transform, Scaling and Quantization

สำหรับมาตรฐาน H.264 / AVC ใช้การแปลงแบบจำนวนเต็ม (Integer Transform) ขนาดบล็อก 4x4 การแปลงแบบจำนวนเต็มใช้พื้นฐานจากการแปลงแบบ DCT ซึ่งกรณีที่ทำอินเวอร์สการแปลงจะมีประสิทธิภาพที่ดีกว่ารูปแบบการแปลงแบบ DCT โดยถ้าบล็อกนั้นเข้ารหัสแบบภายในเฟรม ค่าสัมประสิทธิ์ DC ของบล็อกส่วนประกอบความสว่าง(Luma) จะถูกนำมาจัดเรียงและทำการส่งไปก่อน แล้วจึงทำการส่งข้อมูลของบล็อกที่เหลือแสดงดังรูปที่ 59



รูปที่ 59 การจัดเรียงค่าสัมประสิทธิ์

การควอนไทซ์เป็นขั้นตอนของการสูญเสียข้อมูลบางส่วน การสูญเสียข้อมูลขึ้นอยู่กับค่าควอนไทซ์ กล่าวคือ ถ้าค่าควอนไทซ์มีค่ามากสูญเสียข้อมูลไปมาก ค่าควอนไทซ์มีค่าน้อยสูญเสียข้อมูลไปน้อย ทั้งนี้การเลือกค่าควอนไทซ์เป็นขั้นตอนที่มีความสำคัญต่อจำนวนบิตที่ทำการเข้ารหัสเฟรมนั้นๆอีกด้วย สำหรับมาตรฐาน H.264 / AVC ได้กำหนดขั้นตอนของการควอนไทซ์ (quantization step) ไว้ที่ 51 ระดับ (52 ค่า) ค่าสัมประสิทธิ์ของการแปลงแบบจำนวนเต็ม ค่าสัมประสิทธิ์ที่อยู่ทางด้านบนซ้ายมือจะแทนค่าสัมประสิทธิ์ DC มีค่ามากที่สุดและที่เหลือเรียกว่า สัมประสิทธิ์ AC และเมื่อผ่านการควอนไทซ์เรียงค่าสัมประสิทธิ์ให้เป็นมิติเดียวโดยเริ่มจากค่าสัมประสิทธิ์ DC ซึ่งในมาตรฐาน H.264 / AVC นี้จะมีวิธีเรียงค่าสัมประสิทธิ์อยู่ 2 วิธีคือ การกวาดซิกแซก (Zigzag scan) เป็นวิธีที่ใช้ อยู่โดยทั่วไปทั้งในมาตรฐานก่อนหน้านี้ วิธีที่ 2 คือ การกวาดแบบคู่ (Double scan) ลักษณะการเรียงค่า สัมประสิทธิ์ทั้ง 2 วิธี แสดงดังรูปที่ 5.33

DC

(L) Entropy Coding

ในมาตรฐาน H.264/AVC สนับสนุนการเข้ารหัสเอ็นโทรปี 2 แบบ สำหรับการส่งการควอนไทซ์ค่าสัมประสิทธิ์การแปลงที่มีประสิทธิภาพสูงกว่า แบบแรกคือวิธี Context-Adaptive Variable Length Coding (CAVALLC)

0	1	2	3
4	5	6	7
8	9	10	11

ตามแผนการนี้ ตาราง VLC สำหรับสับเปลี่ยน(Switch) องค์ประกอบของ Syntax ที่มีหลายองค์ประกอบ โดย องค์ประกอบของ Syntax ที่ถูกส่งออกไป

ประสิทธิภาพการเข้ารหัสเอ็นโทรปีสามารถปรับปรุงให้มีประสิทธิภาพยิ่งขึ้นหากใช้ อีกวิธีคือ Context-Adaptive Binary Arithmetic Coding (CABAC) ซึ่งเมื่อเปรียบเทียบกับ CAVALC แล้ว CABAC จะมี ประสิทธิภาพสูงกว่า 5-15% ประสิทธิภาพสูงสุดเมื่อเข้ารหัสสัญญาณ Interlaced TV

(L) In-Loop Deblocking Filter

ภาพเส้นขอบของบล็อกเป็นแบบฉบับของภาพที่ถูกสร้างขึ้นเมื่อการถอดรหัสมีความแม่นยำน้อยกว่า pixel ที่ ถูกสัญญาณรบกวน โดยทั่วไปภาพบล็อกก็จะเป็นภาพที่เกิดจากการบีบอัดสัญญาณ ด้วยเหตุนี้ H.264/AVC จึง กำหนดให้สัญญาณวนรอบ(in-loop) เครื่องกรองสัญญาณด้านการเกิดบล็อกก็แบบปรับตัวได้ โดยความเข้มแข็ง ของ เครื่องกรองสัญญาณ ถูกควบคุมโดยค่าของหลายองค์ประกอบของ Syntax ภาพที่เป็นบล็อกจะลดลงใน ขณะที่ความชัดเจนของภาพคงเดิมไม่เปลี่ยน Deblocking Filter สามารถลดอัตราบิดลง 5-10%



Without Filter

รูปที่ 60 แสดงผลของ Deblocking



With H264/AVC Deblocking

(L) Hypothetical Reference Decoder

ข้อดีของมาตรฐาน H.264/AVC คือการประกันว่า อุปกรณ์ถอดรหัส (Decoder) ที่ปฏิบัติตามมาตรฐาน จะ สามารถถอดรหัสสัญญาณที่เข้ารหัสตามมาตรฐาน

5.5.4.2 Profiles and levels

แต่ละ Profiles จะมีการกำหนดขีดจำกัดของบิตสตรีม Syntax ที่สนับสนุนโดยตัวถอดรหัส H.264/AVC เวอร์ชันแรก ประกอบด้วย 7 profiles ได้แก่

- Baseline profile (BP) จะเหมาะสมกับงานประยุกต์การสนทนาเวลาจริง เช่นการประชุมสัมมนาทางวิดีโอ ททัศน์ และ โทรศัพท์ภาพ
- Main profile (MP) ถูกออกแบบมาให้ใช้กับการเก็บข้อมูลลงสื่อดิจิทัล และ การbroadcastสัญญาณโทรทัศน์ แต่ลดความสำคัญลงเมื่อมีการพัฒนา High profile ขึ้นมา
- Extended profile (XP) เหมาะกับการสตรีมมิ่งวิดีโอ เช่นบริการมัลติมีเดียผ่านโครงข่ายอินเทอร์เน็ต บีบอัดสัญญาณมาก ทนทานต่อการสูญเสียข้อมูล

- High profile (HiP) ออกแบบมาให้ใช้กับการbroadcastสัญญาณโทรทัศน์ความชัดเจนสูง (HDTV)การเก็บข้อมูลลงสื่อดิจิทัล เช่น HD-DVD และ Blu-ray Disc
 - High 10 profile (Hi10P) ออกแบบมาให้ใช้กับ เก็บข้อมูลลงสื่อดิจิทัล โทรทัศน์ความชัดเจนสูง (HDTV) ใช้ข้อมูล 10 Bits
 - High 4:2:2 profile (Hi422P) พัฒนาต่อจาก High 10 profile ถูกออกแบบมาให้ใช้กับงานประยุกต์ แบบ Interlaced Scan ในการส่งและแจกจ่ายคอนเทนต์ ใช้ข้อมูล 10 Bits
 - High 4:4:4 Predictive profile (Hi444PP) พัฒนาต่อจาก High 4:2:2 profile เพื่อใช้กับงานประยุกต์ การตัดต่อในสตูดิโอ และการประมวลผลหลังการผลิต
นอกจากนี้ ยังมี การกำหนด high profiles ทั้งหมด 4 profiles ได้แก่
 - High 10 Intra Profile สำหรับงานที่ใช้ Intra-frameล้วน เช่นการเก็บข้อมูลลงสื่อดิจิทัล โทรทัศน์ความชัดเจนสูง (HDTV) เพื่อง่ายต่อการตัดต่อ (Edit)
 - High 4:2:2 Intra profile สำหรับงานที่ใช้ Intra-frame ล้วน มี Chroma Resolution เพิ่มขึ้น เช่นการเก็บข้อมูลลงสื่อดิจิทัล โทรทัศน์ความชัดเจนสูง (HDTV) เพื่อง่ายต่อการตัดต่อ (Edit)
 - High 4:4:4 Intra profile สำหรับงานที่ใช้ Intra-frame ล้วน
 - CAVLC 4:4:4 Intra profile สำหรับงานที่ใช้ Intra-frame ล้วน และการเข้ารหัส Entropy แบบ CAVLC
- ตารางที่ 1 แสดง profiles ของ H.264/AVC และ MPEG-4 ตารางที่ 2 แสดงรายละเอียดของ profiles ในแต่ละ profile ใดๆ จะมีการกำหนด levels ซึ่งสอดคล้องกับกำลังในการประมวลผลและหน่วยความจำที่รองรับในตัวเข้าและถอดรหัส ทั้งนี้แต่ละ level อาจจะสนับสนุนขนาดภาพหลายขนาด เช่น QCIF, CIF, ITU-R 601 (SDTV), HDTV, S-HDTV และ D-Cinema รวมทั้งแต่ละ level จะมีการกำหนดอัตราบิต ขนาดของภาพ และขนาดของบัฟเฟอร์ เป็นต้น ซึ่งถูกแสดงในตารางที่ 3

Application	Requirements	H.264 Profiles	MPEG-4 Profiles
Broadcast television	Coding efficiency, reliability (over a controlled distribution channel), interlace, low-complexity decoder	Main	ASP (Advanced Simple)
Streaming video	Coding efficiency, reliability (over a uncontrolled packet-based network channel), scalability	Extended	ARTS (Advanced Real Time Simple) or FGS (Fine Granular Scalability)
Video storage and playback	Coding efficiency, interlace, low-complexity encoder and decoder	Main	ASP
Videoconferencing	Coding efficiency, reliability, low latency, low-complexity encoder and decoder	Baseline	SP (Simple)
Mobile video	Coding efficiency, reliability, low latency, low-complexity encoder and decoder, low power consumption	Baseline	SP
Studio distribution	Lossless or near-lossless, interlace, efficient transcoding	Main High	Studio Profile

ตารางที่ 1 การเปรียบเทียบ profiles ระหว่างมาตรฐาน MPEG-4 และ H.264

Profile	Applications
Baseline	Video Conferencing Videophone
Main	Digital Storage Media Television Broadcasting
Extended	Streaming Video
High	Content contribution Content distribution Studio editing Post processing

Level number	Picture type & frame rate
1	QCIF @ 15fps
1.1	QCIF @ 30fps
1.2	CIF @ 15fps
1.3	CIF @ 30fps
2	CIF @ 30fps
2.1	HHR @15 or 30fps
2.2	SDTV @ 15fps
3	SDTV: 720x480x30i, 720x576x25i 10Mbps(max)
3.1	1280x720x30p
3.2	1280x720x60p
4	HDTV: 1920x1080x30i, 1280x720x60p, 2Kx1Kx30p 20Mbps(max)
4.1	HDTV: 1920x1080x30i, 1280x720x60p, 2Kx1Kx30p 50Mbps(max)
4.2	HDTV: 1920x1080x60i, 2Kx1Kx60p
5	SHDTV/D-Cinema: 2.5Kx2Kx30p
5.1	SHDTV/D-Cinema: 4Kx2Kx30p

ตารางที่ 3 การกำหนดพารามิเตอร์ของ Level

ตารางที่ 4 แสดงความสัมพันธ์ระหว่าง Profile กับเครื่องมือ (Tools)

	Baseline	Extended	Main	High	High 10	High 4:2:2	High 4:4:4 Predictive
I and P Slices	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B Slices	No	Yes	Yes	Yes	Yes	Yes	Yes
SI and SP Slices	No	Yes	No	No	No	No	No
Multiple Reference Frames	Yes	Yes	Yes	Yes	Yes	Yes	Yes
In-Loop Deblocking Filter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAVLC Entropy Coding	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CABAC Entropy Coding	No	No	Yes	Yes	Yes	Yes	Yes
Flexible Macroblock Ordering (FMO)	Yes	Yes	No	No	No	No	No
Arbitrary Slice Ordering (ASO)	Yes	Yes	No	No	No	No	No
Redundant Slices (RS)	Yes	Yes	No	No	No	No	No
Data Partitioning	No	Yes	No	No	No	No	No
Interlaced Coding (PicAFF, MBAFF)	No	Yes	Yes	Yes	Yes	Yes	Yes
4:2:0 Chroma Format	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monochrome Video Format (4:0:0)	No	No	No	Yes	Yes	Yes	Yes
4:2:2 Chroma Format	No	No	No	No	No	Yes	Yes
4:4:4 Chroma Format	No	No	No	No	No	No	Yes
8 Bit Sample Depth	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9 and 10 Bit Sample Depth	No	No	No	No	Yes	Yes	Yes
11 to 14 Bit Sample Depth	No	No	No	No	No	No	Yes
8x8 vs. 4x4 Transform Adaptivity	No	No	No	Yes	Yes	Yes	Yes
Quantization Scaling Matrices	No	No	No	Yes	Yes	Yes	Yes
Separate Cb and Cr QP control	No	No	No	Yes	Yes	Yes	Yes
Separate Color Plane Coding	No	No	No	No	No	No	Yes
Predictive Lossless Coding	No	No	No	No	No	No	Yes
	Baseline	Extended	Main	High	High 10	High 4:2:2	High 4:4:4 Predictive

6. Audio Coding

6.1 คุณสมบัติของเสียง (Loudness)

หูมนุษย์สามารถรับรู้ Loudness (Sound Amplitude / Volume) ความดังของเสียง จากเกือบเงียบจนถึงดังสุด Loudness มีหน่วยวัดเป็น dB (decibel) Scale ของ Decibel เป็น Logarithm เพิ่มขึ้นเป็นแบบ exponentially ระดับเสียงในชีวิตประจำวันคือ เสียงพูด สนทนา 40-60 dB เสียงกริ่งโทรศัพท์ 80 dB , เสียงรถกำลังจอดบนถนน 90 dB Rock Concert 120dB ความดังของเสียงทั้ง 3 ระดับหลัง นี้สามารถทำลายประสาทการได้ยินได้

การที่มนุษย์มี 2 หู ทำให้มนุษย์สามารถแยกแยะทิศทางและระยะทางของเสียงได้ เพราะเสียงจากทิศทางหนึ่ง จะเดินทางมาถึงหูมนุษย์ได้ไกลสุด ใช้เวลาเดินทางไม่เกิน 1/500 วินาที

ระดับเสียงที่ต่างกันของหูทั้งสองข้าง เพราะความถี่ของเสียงต่างกัน เสียงความถี่สูงจะระบุแหล่งกำเนิดเสียง ได้ดีกว่าความถี่ต่ำ

Pitch หรือ Tone ของเสียงกำหนดโดยความถี่ของเสียง เช่น 20 Hz เป็น Low Pitch ส่วน 20,000 Hz เป็น High Pitch หูมนุษย์จะรับรู้ได้ดีในย่าน 20 Hz - 20,000 Hz เช่นเสียงพูดในย่าน 200 Hz - 8,000 Hz

6.2 Digital Audio

สัญญาณเสียงดิจิทัล คือ ไบท์สตรีม ประกอบด้วย Amplitude ของข้อมูล Pitch หรือ Tone ของเสียง กำหนดโดยความถี่ของเสียง Pitch คือ ผลการเกิดซ้ำ ๆ (periodic repetition) เป็นช่วง ๆ ของ Amplitude มีการเกิดความเข้าใจผิดว่า Bits Depth หรือจำนวนบิตมาก หรือ อัตราSampling สูงๆ คุณภาพเสียงจะดีกว่า ซึ่งไม่จริงเสมอไป แต่ Bits Depth หรือจำนวนบิตมาก สำคัญกว่า อัตราSampling

เช่น CD ใช้ Bits Depth หรือจำนวนบิต 16 บิต และมี อัตราSampling 44.1 KHz

$$= 2 \text{ ch} \times 44,1 \text{ KHz} \times 16 \text{ Bits/s} = 1.4 \text{ Mbps}$$

ในขณะที่แผ่น CD มีความจุ ประมาณ 0.8 GBytes

$$\text{เสียงภาพยนตร์ } 6 \text{ ch} \times 48 \text{ KHz} \times 16 \text{ Bits/s} = 4.6 \text{ Mbps หรือประมาณ } 1 \text{ GBytes / Hour}$$

ด้วยเทคนิคการบีบอัดสัญญาณในปัจจุบัน เสียงภาพยนตร์ 6 ch จะเหลือเพียงไม่กี่ร้อย Kbps

แต่การบีบอัดสัญญาณแบบง่ายๆ ทั้งแบบไม่สูญเสีย (Lossless) และแบบสูญเสีย (Lossy) ยังไม่สามารถลดขนาดสัญญาณได้มากนัก จึงต้องพิจารณาวิธีอื่นๆ

6.3 Psychoacoustics Model

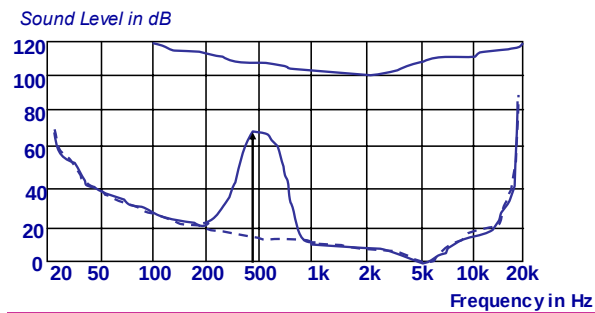
การศึกษา Psychoacoustics Model เพื่อจะทราบว่ามนุษย์รับรู้การได้ยินอย่างไร มีขีดจำกัดแค่ไหน ปรากฏว่า ย่านความถี่เสียงที่ได้ยินอยู่ในย่าน 40 Hz-14 KHz แต่ที่ต่ำที่สุดคือ ในย่าน 2 - 4 KHz โดยมี Dynamic Range คือความสามารถในการได้ยินเสียงเบาที่สุด จนถึงดังที่สุด กว้าง ประมาณ 96 dB เสียงพูดปกติอยู่ในย่าน 80 Hz-8 KHz แต่เสียงพูดต้องการเสียงเฉพาะในย่าน 300-3,200 Hz เท่านั้นที่ทำให้เข้าใจคำพูดได้ชัดเจน

ความจริงหูมนุษย์จะรับรู้เสียงบางชนิดได้ดีกว่า เมื่อมีเสียงหลายๆ ความถี่สัญญาณเสียงที่แรงกว่าอาจจะรับรู้ลง หากความถี่เสียงใกล้เคียงกัน นี่คือการปรากฏการณ์ที่เรียกว่า Frequency Mask

เมื่อหูมนุษย์ได้ยินเสียงที่ดังชั่วขณะหนึ่ง ก็จะรับรู้เสียงที่ดังเบากว่าได้ ปรากฏการณ์นี้เรียกว่า Temporal Mask

มนุษย์มีขีดจำกัดในช่วงความถี่เสียงในการได้ยินเสียง มาตรการ การได้ยินความถี่เสียงที่เหมือนกัน แสดงในรูปแบบความกว้าง 'Critical Band' แคบกว่า 100 Hz ที่ความถี่ต่ำสุดที่ได้ยิน และมากกว่า 4 KHz เมื่อรวมกันแล้ว สามารถแบ่งย่านความถี่ออกเป็น 25 Critical Band

Psycho-Acoustical Data Masking



รูปที่ 63 แสดง Temporal Masking

6.4 Perceptual Coding

ใช้รูปแบบ Psycho-Acoustical Model เพื่อระบุสัญญาณที่ถูกบดบัง โดยผลของ Frequency Masking และ Temporal Masking

สัญญาณที่ไม่ถูกส่งหรือส่งไม่ก็บดบังเพื่อไม่ให้เสียงเพี้ยน สัญญาณที่อยู่ในย่านที่รับฟังได้ชัดเจน จะจัดให้มีบิตมากขึ้น สัญญาณจะได้รับการวิเคราะห์ เพื่อกำหนดระดับเสียงที่เป็นจุดเริ่มต้นการได้ยิน โดยใช้ Perceptual Model

6.5 Audio Main Standard

MPEG ใช้การเข้ารหัสด้วยวิธี Perceptual คือ

- MPEG-1

Layer 1	อัตราบีบอัด 1:4	384 Kbps	สำหรับสัญญาณสเตอริโอ
Layer 2	อัตราบีบอัด 1:6 – 1:8	256-192 Kbps	สำหรับสัญญาณสเตอริโอ
Layer 3 (MP3)	อัตราบีบอัด 1:10-1:12	128 -112 Kbps	สำหรับสัญญาณสเตอริโอ

- MPEG-2

-Back Compatible คืออุปกรณ์ถอดรหัสสามารถถอดรหัส มาตรฐานการบีบอัดแบบก่อน (MPEG-1) ได้

-AAC (Non Compatible) : Advance Audio Coding

6.5.1 MPEG-1

ขนาดข้อมูล 1.5 Mbps ทั้งภาพและเสียง ประกอบด้วย สัญญาณวิดีโอ 1.2 Mbps สัญญาณ เสียง 0.3 Mbps ,Compression Factor Range 2.7-6.4 MPEG Audio สัมพันธ์กับ อัตราSampling แบบ 32, 44.1 และ 48KHz

MPEG-1 สเตอริโอ มี โหมด 4 คือ

1. Mono
2. Dual Mono เช่นเสียง 2 ภาษา
3. Stereo
4. Joint Stereo นำเอาความสัมพันธ์ระหว่าง 2 ช่องมาเป็นประโยชน์

ทั้ง 3 Layer ใช้ Model ในการบีบอัดสัญญาณเหมือนกัน แต่ Codec จะมีความซับซ้อนยิ่งขึ้น ในแต่ละ Layer ทั้ง 3 Layer จะแบ่งข้อมูลเป็นเฟรม ๆ ละ 384 ตัวอย่าง 12 ตัวอย่าง สำหรับ 32 Filtered Sub-Band

- **Layer 1**

Layer 1 ใช้ DCT แบบ Filter ชนิดเฟรมเดียว ทั้งแถบแบนด์มีความถี่เท่ากัน รูปแบบ Psychoacoustic ชนิด Frequency Mask ที่ 384 Kbps / Ch ,เข้ารหัส 32 Sub-Band โดยแบ่ง input เป็น 12 ตัวอย่าง ทำค่าสัมประสิทธิ์ให้เป็นมาตรฐาน แต่ละบล็อกเลือกจาก 15 Quantization ไม่มีการเข้ารหัสแบบ entropy หลังจากการเข้ารหัสการแปลง อุปกรณ์การถอดรหัสจะง่ายกว่าอุปกรณ์เข้ารหัสมาก

- **Layer 2**

- ใช้ 3 เฟรมต่อฟิลเตอร์ (ก่อนหน้า ปัจจุบัน หลัง) รวม 1152 ตัวอย่าง สำหรับ Temporal Masking
- Transparent 256 Kbps
- ปรับปรุง perception Model ให้ดีขึ้น
- การควอนไทซ์แบบละเอียด (Finer resolution Quantization)

- **Layer 3**

ใช้ Critical Band Filter ที่ดีกว่า (Frequency ไม่เท่ากัน) Psychoacoustic Model ที่รวมเอา Temporal Masking และใช้การเข้ารหัส Huffman Coding และใช้วิธีลดความซ้ำซ้อนในสัญญาณสเตอริโอด้วย

- Transparent 96 Kbps/Ch
- ใช้วิธีปรับขนาด (Variable Size) มาเปลี่ยนแปลง DCT บนตัวอย่างแต่ละ Sub-Band Channel
- ระบบยุ่งยากกว่า Layer 1&2
- การเข้ารหัสเพื่อลดความซ้ำซ้อนในสัญญาณสเตอริโอ (Stereo Redundancy Coding)
 - ทั้ง 3 Layer สามารถใช้ การเข้ารหัสเพื่อลดความซ้ำซ้อนในสัญญาณสเตอริโอ (Stereo Redundancy Coding)
 - Middle / Side Stereo Coding เข้ารหัส Middle $M=(L+R)$ และส่วนต่าง Side $S=(L-R)$ ที่มีบิตน้อยมาก Decoder จะสร้างสัญญาณ MS Stereo ใหม่ คือ

$$L = (M+S) / \sqrt{2} \quad R = (M-S) / \sqrt{2}$$

6.5.1 Advance Audio Coding

- เพื่อแทน MP3
- กำหนดในมาตรฐาน ดังนี้
 - MPEG-2 Parts 7 1997
 - MPEG-4 Parts 3 1999
 - HE-AAC (AAC + SBR) 2001
 - HE-AAC v2 (AAC with Parametric Stereo) 2005

7. Comparison of Digital Video Standard

ตารางที่ 5 เปรียบเทียบมาตรฐานบีบอัดสัญญาณ

ตารางที่ 5 เปรียบเทียบมาตรฐานบีบอัดสัญญาณ

Feature/Standard	MPEG-1	MPEG-2	MPEG-4 part 2 (visual)	H.264/MPEG-4 part 10
Macroblock size	16x16	16x16 (frame mode) 16x8 (field mode)	16x16	16x16
Block Size	8x8	8x8	16x16, 16x8, 8x8	16x16, 8x16, 16x8, 8x8, 4x8, 8x4, 4x4
Transform	8x8 DCT	8x8 DCT	8x8 DCT/Wavelet	4x4, 8x8 Int DCT 4x4, 2x2 Hadamard
Quantization	Scalar quantization with step size of constant increment	Scalar quantization with step size of constant increment	Vector quantization	Scalar quantization with step size increase at the rate of 12.5%
Entropy coding	VLC	VLC	VLC	VLC, CAVLC, CABAC
Motion Estimation & Compensation	Yes	Yes	Yes	Yes, more flexible Up to 16 MVs per MB
Playback & Random Access	Yes	Yes	Yes	Yes

Feature/Standard	MPEG-1	MPEG-2	MPEG-4 part 2 (visual)	H.264/MPEG-4 part 10
Pel accuracy	Integer, ½-pel	Integer, ½-pel	Integer, ½-pel, ¼-pel	Integer, ½-pel, ¼-pel
Profiles	No	5	8	4
Reference picture	one	one	one	multiple
Bidirectional prediction mode	forward/backward	forward/backward	forward/backward	forward/forward forward/backward backward/backward
Picture Types	I, P, B, D	I, P, B	I, P, B	I, P, B, SP, SI
Error robustness	Synchronization & concealment	Data partitioning, FEC for important packet transmission	Synchronization, Data partitioning, Header extension, Reversible VLCs	Data partitioning, Parameter setting, Flexible macroblock ordering, Redundant slice, Switched slice
Transmission rate	Up to 1.5Mbps	2-15Mbps	64kbps - 2Mbps	64kbps -240Mbps
Compatibility with previous standards	n/a	Yes	Yes	No
Encoder complexity	Low	Medium	Medium	High

Features	H.261	MPEG-1	MPEG-2	H.263	MPEG-4	H.264
Picture coding type	I, P	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B
Entropy Coding	VLC	VLC	VLC	VLC, SAC	VLC	UVLC, CAVLC, CABAC
MV resolution	Int. Pel	½ pel	½ pel	½ pel	¼ pel	¼ pel
Transform	8x8 DCT	8x8 DCT	8x8 DCT	8x8 DCT	8x8 DCT	4x4 & 8x8 integer
Vector Block size	16x16	16x16	16x16, 16x8	16x16, 8x8	16x16, 8x8	16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4
Spatial Intra Prediction	No	No	No	No	No	Yes
Formats supported	Prog	Prog	Prog/Intr	Prog	Prog/Intr	Prog/Intr
Prediction Modes	Frame	Frame	Field & Frame	Frame	Field & Frame	Field & Frame
De-blocking filter	In-loop	None	Post	Annex J in-loop	Post	In-loop