M U L T I M E D I A

KOMPRESI AUDIO/VIDEO

Kompresi audio/video adalah salah satu bentuk kompresi data yang bertujuan untuk mengecilkan ukuran file audio/video dengan metode

- Lossy → format : Vorbis, MP3;
- Loseless → format : FLAC; pengguna : audio engineer, audiophiles

Kompresi dilakukan pada saat ***pembuatan*** file audio/video dan pada saat ***distribusi*** file audio/video tersebut!

Kendala kompresi audio:

- Perkembangan sound recording yang cepat dan beranekaragam
- Nilai dari audio sample berubah dengan cepat

Losless audio codec tidak mempunyai masalah dalam kualitas suara, penggunaannya dapat difokuskan pada:

- Kecepatan kompresi dan dekompresi
- Derajat kompresi
- Dukungan hardware dan software

Lossy audio codec penggunaannya difokuskan pada:

- Kualitas audio

- Faktor kompresi
- Kecepatan kompresi dan dekompresi
- Inherent latency of algorithm (penting bagi real-time streaming)
- Dukungan hardware dan software

Metode Kompresi Audio

- Metode Transformasi
    - Menggunakan algoritma seperti MDCT (Modified Discreate Cosine Transform) untuk mengkonversikan gelombang bunyi ke dalam sinyal digital agar tetap dapat didengar oleh manusia (20 Hz s/d 20kHz) , yaitu menjadi frekuensi 2 s/d 4kHz dan 96 dB.
- Metode Waktu
    - Menggunakan LPC (Linier Predictive Coding) yaitu digunakan untuk speech (pidato), dimana LPC akan menyesuaikan sinyal data pada suara manusia, kemudian mengirimkannya ke pendengar. Jadi seperti layaknya komputer yang berbicara dengan bahasa manusia dengan kecepatan 2,4 kbps

Teknik kompresi audio dengan format MPEG (Moving Picture Expert Group)

- MPEG-1 menggunakan bandwidth 1,5 Mbits/sec untuk audio dan video, dimana 1,2 Mbits/sec digunakan untuk video sedangkan 0,3 Mbits/sec digunakan untuk audio.

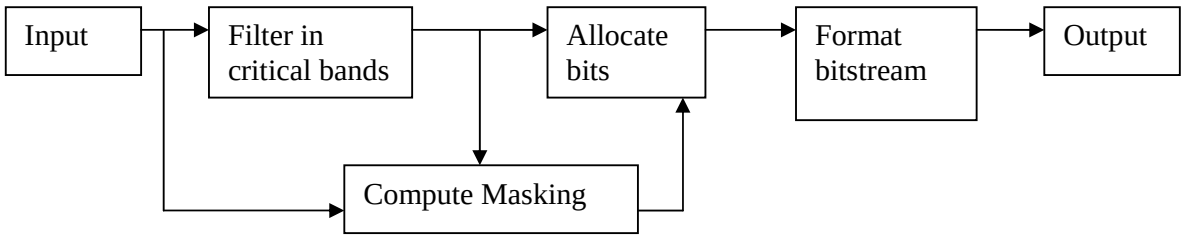
  Nilai 0,3 Mbits/sec ini lebih kecil dibandingkan dengan bandwidth yang dibutuhkan oleh CD Audio yang tidak terkompres sebesar 44100 samples/sec x 16 bits/sample * 2 channel > 1,4 Mbits/sec yang hanya terdiri dari suara saja.
- Untuk ratio kompresi 6:1 untuk 16 bit stereo dengan frekuensi 48kHz dan bitrate 256 kbps CBR akan menghasilkan ukuran file terkompresi

kira-kira 12.763 KB, sedangkan ukuran file tidak terkompresinya adalah 75.576 KB

- MPEG-1 audio mendukung frekuensi dari 8kHz, 11kHz, 12kHz, 16kHz, 22kHz, 24 kHz, 32 kHz, 44kHz, dan 48 kHz. Juga mampu bekerja pada mode mono (single audio channel), dual audio channel, stereo, dan joint-stereo

Algoritma MPEG Audio

- Menggunakan filter untuk membagi sinyal audio: misalnya pada 48 kHz, suara dibagi menjadi 32 subband frekuensi.
- Memberikan pembatas pada masing-masing frekuensi yang telah dibagi-bagi, jika tidak akan terjadi intermodulasi (tabrakan frekuensi)
- Jika sinyal suara terlalu rendah, maka tidak dilakukan encode pada sinyal suara tersebut
- Diberikan bit parity yang digunakan untuk mengecek apakah data tersebut rusak atau tidak (yang mungkin disebabkan oleh gangguan / noise), apabila rusak, maka bit tersebut akan digantikan bit yang jenisnya sama dengan bit terdekatnya.



Kompresi Audio MP3

- Asal-usul MP3 dimulai dari penelitian IIS-FHG (Institut Integriette Schaltungen-Fraunhofer Gesellschaft), sebuah lembaga penelitian terapan di Munich, Jerman dalam penelitian coding audio perceptual.
- Penelitian tersebut menghasilkan suatu algoritma yang menjadi standard sebagai ISO-MPEG Audio Layer-3 (MP3)

sound quality	bandwith	mode	bitrate	reduction ratio
telephone sound	2.5 kHz	mono	8 kbps	96 : 1
better than shortwave	4.5 kHz	mono	16 kbps	48 : 1
better than AM radio	7.5 kHz	mono	32 kbps	24 : 1
similar to FM radio	11 kHz	stereo	56..54 kbps	26..24 : 1
near CD	15 kHz	stereo	96 kbps	16 : 1
CD	>15kHz	stereo	112..128 kbps	14..12 : 1

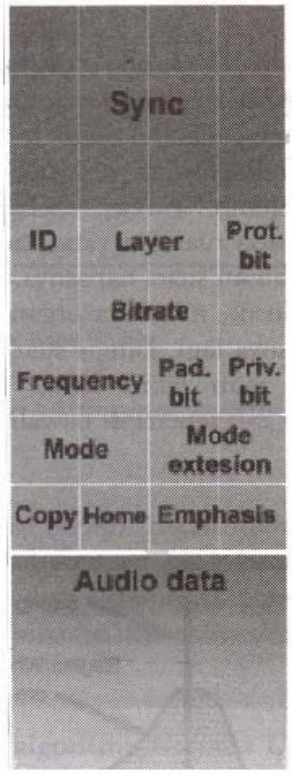
Tabel kemampuan kompresi MPEG Layer 3
dengan kualitas suara yang dihasilkan

Format Header MP3

File MP3 terdiri atas 2 bagian data:

- **Header** : berfungsi sebagai tanda pengenal bagi file MP3 agar dapat dibaca oleh MP3 player yang berukuran 4 byte

  Beberapa karakteristik yang dibaca komputer adalah bit ID, bit layer, bit sampling frequency dan bit mode.
- **Data audio :** berisi data file mp3.



Frame MP3

Bit value	Layer
00	not defined
01	layer 3
10	layer 2
11	layer 1

Isi bit layer

mode value	mode
00	stereo
01	joint stereo
10	dual channel
11	mono

Isi bit mode

frequency value	MPEG-1	MPEG-2
00	44100 Hz	22050 Hz
01	48000 Hz	24000 Hz
10	32000 Hz	16000 Hz
11		

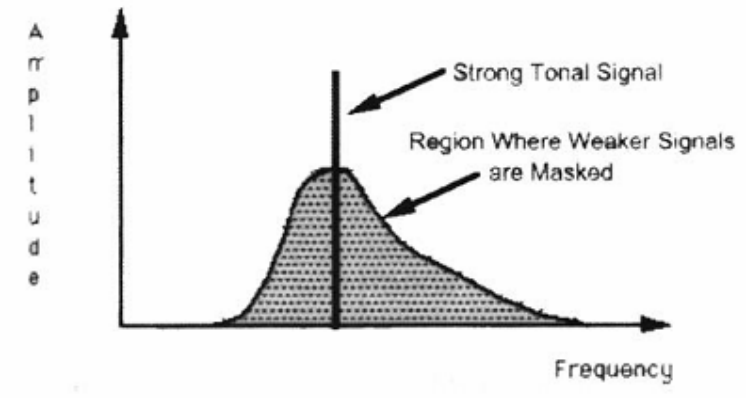
Isi bit frequency sampling

Teknik kompresi MP3

Beberapa karakteristik dari MP3 memanfaatkan kelemahan pendengaran manusia.

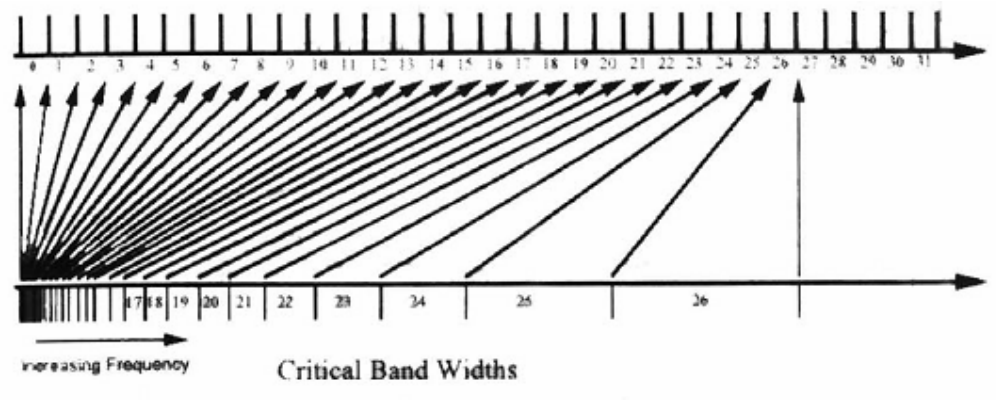
1. **Model psikoakustik**
   - Model psikoakustik adalah model yang menggambarkan karakteristik pendengaran manusia.
   - Salah satu karakteristik pendengaran manusia adalah memiliki batas frekuensi 20 Hz s/d 20 kHz, dimana suara yang memiliki frekuensi yang berada di bawah ambang batas ini tidak dapat didengar oleh manusia, sehingga suara seperti itu tidak perlu dikodekan.
2. **Auditory masking**

   Manusia tidak mampu mendengarkan suara pada frekuensi tertentu dengan amplitudo tertentu jika pada frekuensi di dekatnya terdapat suara dengan amplitudo yang jauh lebih tinggi.



3. **Critical band**

   Critical band merupakan daerah frekuensi tertentu dimana pendengaran manusia lebih peka pada frekuensi-frekuensi rendah, sehingga alokasi bit dan alokasi sub-band pada filter critical band lebih banyak dibandingkan frekuensi lebih tinggi.

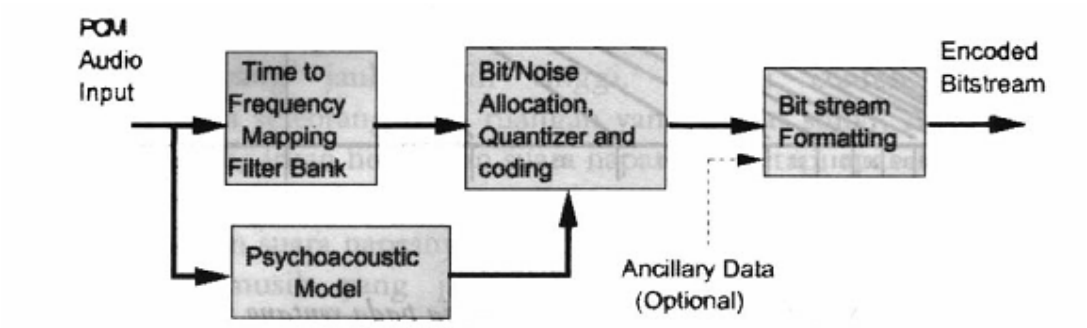


4. Joint stereo

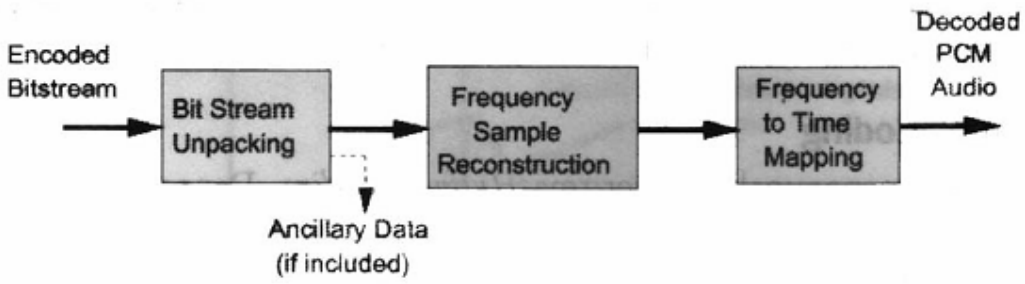
Terkadang dual channel stereo mengirimkan informasi yang sama. Dengan menggunakan joint stereo, informasi yang sama ini cukup ditempatkan dalam salah satu channel saja dan ditambah dengan informasi tertentu. Dengan teknik ini bitrate dapat diperkecil.

Beberapa persyaratan dari suatu encoder/decoder MP3:

- Ukuran file terkompresi harus sekecil mungkin
- Kualitas suara file yang telah terkompresi haruslah sedekat mungkin dengan file asli yang belum dikompresi
- Tingkat kesulitan rendah, sehingga dapat direalisasikan dengan aplikasi yang mudah dibuat dan perangkat keras yang ‘sederhana’ dengan konsumsi daya yang rendah



MPEG/Audio encoder



MPEG/Audio decoder

Filter Bank, adalah kumpulan filter yang berfungsi memfilter masukan pada frekuensi tertentu, sesuai dengan critical band yang telah didefinisikan. Filter yang dipakai adalah gabungan dari filter bank polyphase dan Modified Discrete Cosine Transform (MDCT)

Perceptual Model, dapat menggunakan filter bank terpisah atau penggabungan antara perhitungan nilai energi dan filter bank utama. Keluaran model ini adalah nilai masking treshold. Apabila noise berada dibawah masking treshold, maka hasil kompresi tidak akan dapat dibedakan dari sinyal aslinya.

Quantization/Coding, merupakan proses kuantisasi setelah sinyal disampling. Proses ini dilakukan oleh power-law quantizer, yang memiliki sifat mengkodekan amplitudo besar dengan ketepatan rendah, dan dimasukkannya proses noise shaping. Setelah itu nilai yang telah dikuantisasi dikodekan menggunakan Huffman Coding.

Encoding Bitstream, merupakan tahap terakhir dimana bit-bit hasil pengkodean sampling sinyal disusun menjadi sebuah bitstream.

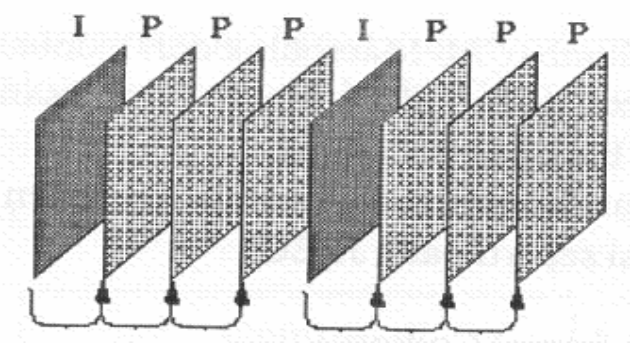
KOMPRESI VIDEO

- Video memiliki 3 dimensi:
  - 2 dimensi spatial (horisontal dan vertikal),
  - 1 dimensi waktu.
- Di dalam video terdapat 2 hal yang dapat dikompresi yaitu frame (still image) dan audionya.
- Data video memiliki:
  - redundancy spatial (warna dalam still image)
  - redundancy temporal (perubahan antar frame)
- Penghilangan redundancy spatial (**spatial / intraframe compression**) dilakukan dengan mengambil keuntungan dari fakta bahwa mata manusia tidak terlalu dapat membedakan warna dibandingkan dengan brightness, sehingga image dalam video bisa dikompresi (teknik ini sama dengan teknik kompresi lossy **color reduction** pada image)
- Penghilangan redundancy temporal (**temporal / interframe compression**) dilakukan dengan mengirimkan dan mengenkode frame yang berubah saja sedangkan data yang sama masih disimpan.

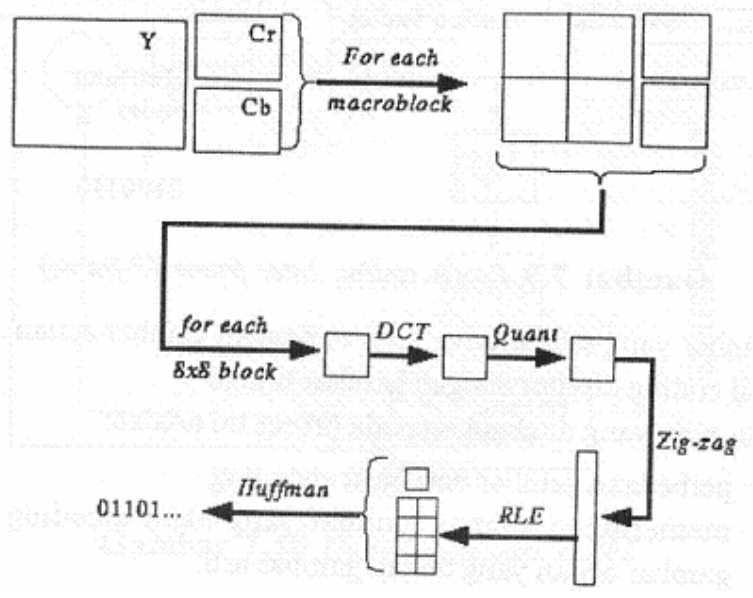
Beberapa Teknik Video Coding

- H.261 dan H.263
  - Merupakan standar video coding yang dibuat oleh CCITT (Consultative Commitee for International Telephone and Telegraph) pada tahun 1988-1990
  - Dirancang untuk video conferencing, aplikasi video telepon menggunakan jaringan telepon ISDN

- Kecepatan bitrate antara p x 64 Kbps. Dimana p adalah frame rate (antara 1 sampai 30)
- Susunan frame H.261 berurutan dimana tiap-tiap 3 buah frame (I) dibatasi dengan 1 buah inter-frame (P)
- Tipe frame gambar yang didukung adalah CCIR 601 CIF (352 x 288) dan QCIF (176 x 144) dengan chroma sub sampling 4:2:0
- Mempunyai 2 tipe frame yaitu: Intra-frame (I-frame) dan Inter-fame (P-frame)
    - I-frame digunakan untuk mengakses banyak pixel
    - P-frame digunakan sebagai “pseudo-differences“ dari frame yang sebelumnya ke frame sesudahnya, dimana antar frame terhubung satu sama lain.

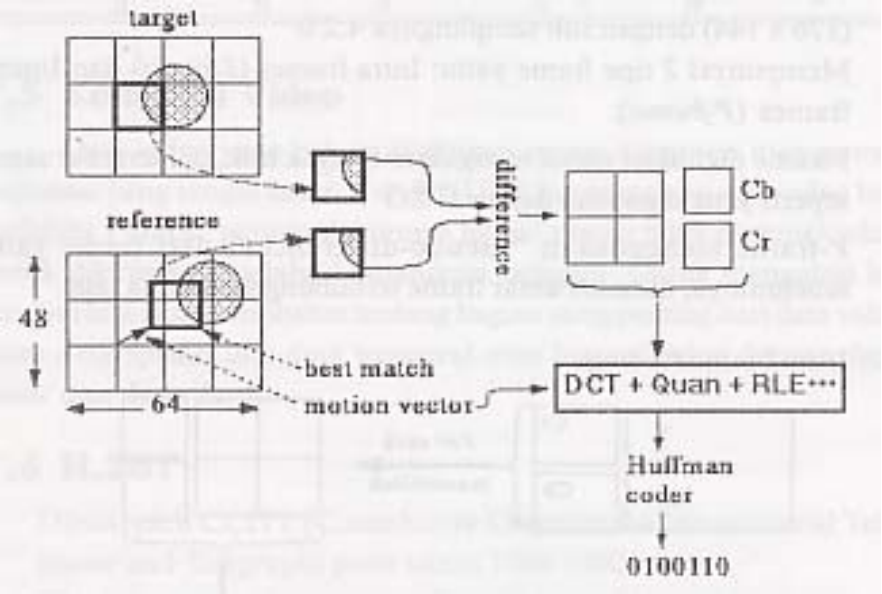


- Intraframe coding



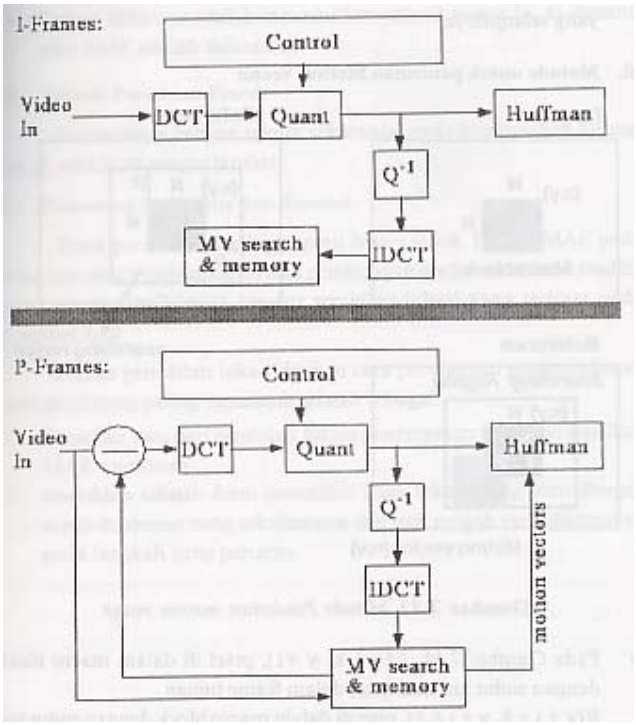
Makroblok yang digunakan pada gambar asli adalah 16 x 16 pixel perblok, dimana Y menggunakan 4 blok, U (Cr) menggunakan 1 blok, dan V (Cb) menggunakan 1 blok.

- Interframe coding



Gambar sebelumnya dijadikan gambar acuan yang akan dibuat gambar hasilnya, dengan menggunakan RMSE untuk mencari tingkat error yang paling kecil.

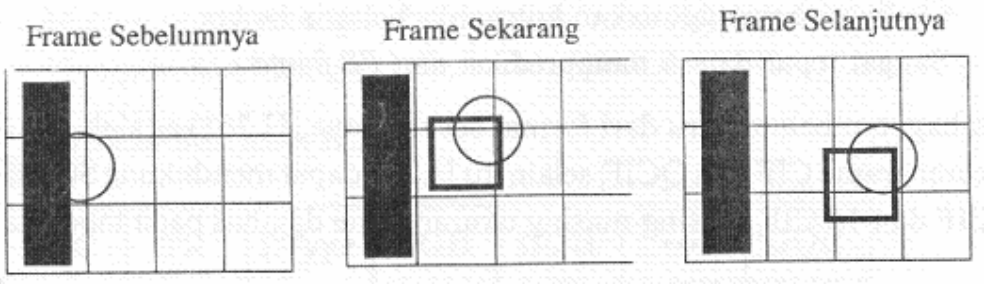
- Encoder H.261



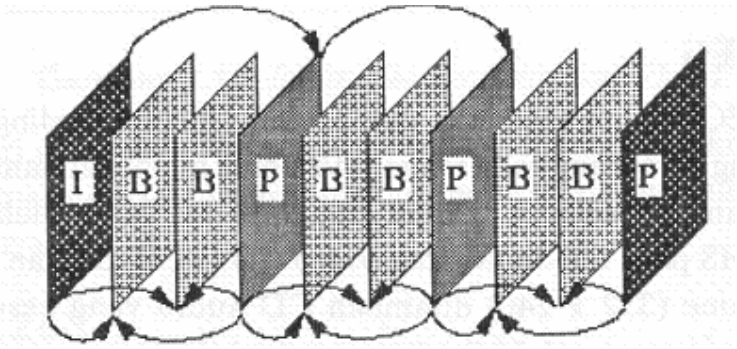
    - Control berfungsi untuk mengatur kecepatan bit rate, jika buffer pengirim penuh, maka bit rate akan dikurangi
    - Memory digunakan sebagai tempat penyimpanan blok gambar yang telah direkonstruksi untuk penciptaan gambar pada P-frame selanjutnya.
  - Kemudian dikemangkan H.263 untuk encoding video pada bitrate rendah

Video Format	Ukuran Resolusi gambar	Mendukung H.261	Mendukung H.263	Bit-rate (Mbit/s) (jika tidak- dikompresi, 30 fps)		Max bit per picture (BPP max, Kb)
				B/W	Bewarna	
SQCIF	128 X 96	n/a	disrankan	3,0	4,4	64
QCIF	176 X 144	disarankan	disarankan	6,1	9,1	64
CIF	352 x 288	Optional	Optional	24,3	36,5	256
4CIF	704 x 576	n/a	Optional	97,3	146,0	512
16CIF	1048 x 1152	n/a	Optional	389,3	583,9	1024

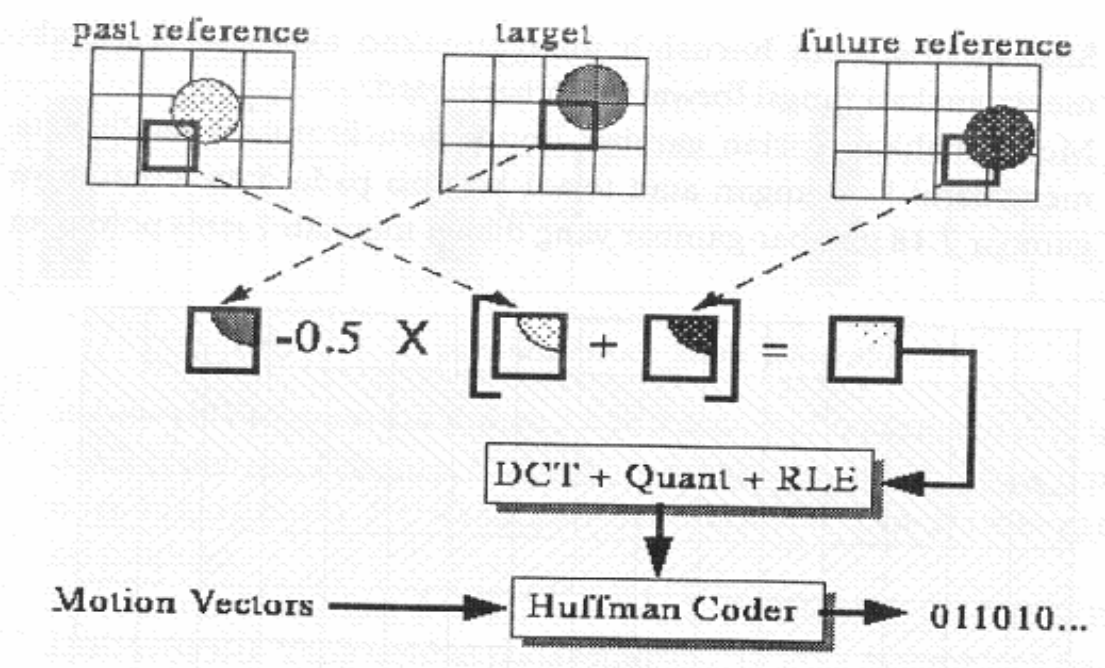
- MPEG audio-video
  - Moving Picture Expert Group dirancang pada tahun 1998 untuk standar audio video transmission
  - MPEG-1 bertujuan membuat kualitas VHS pada VCD dengan ukuran 352 x 240 ditambah kualitas audio seperti CD Audio dengan kebutuhan bandwidth hanya 1,5 Mbits/sec
  - Komponen penting adalah:
    - Audio
    - Video
    - Sistem pengontrol stream video
  - Permasalahan pada frame makroblok



- MPEG menambahkan frame dalam makroblok seperti pada H.261/H.263 yang bernama B-frame (bidirectional frame) sehingga strukturnya adalah:



- Perbedaan dengan H.261
  - Mempunyai jarak yang lebih lebar dibandingkan antara frame I dan frame P sehingga diperlukan perluasan pada vector motion yang digunakan
  - Vektor motion harus berukuran ½ x pixel yang ada

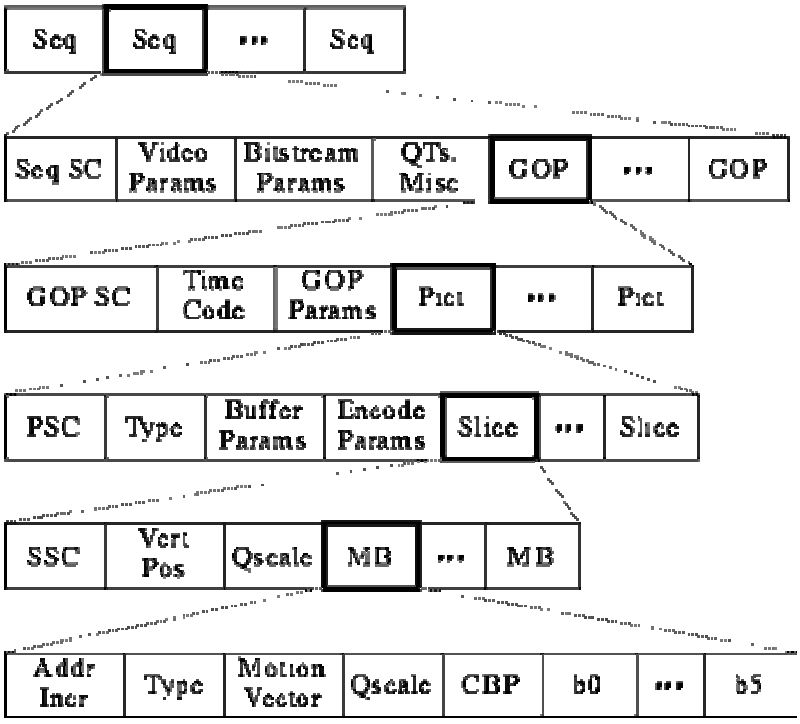


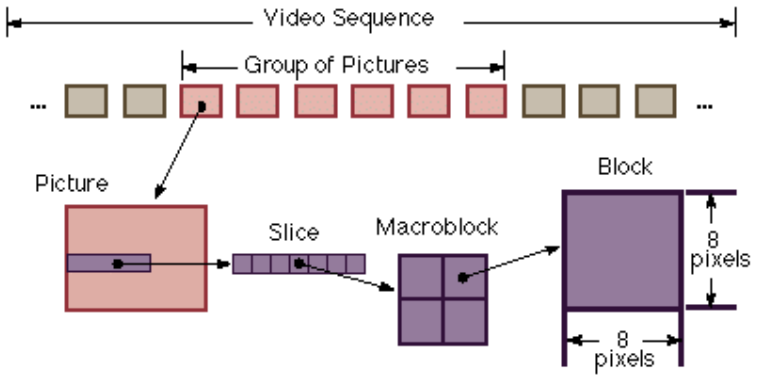
- MPEG-2
    - Merupakan standar pada TV Digital yang dikhususkan untuk HDTV dan DVD

Tabel 7.4 *Tingkatan pada MPEG-2*

Level Aplikasi	Resolusi Maksimum	Maks frame rate (fps)	Maksimum pixel/sec	Maksmum code Data rate (Mb/s)	Pengguna
Tape kecepatan rendah	352 x 288	30	3 M	4	konsumen
Utama	720 x 576	30	10 M	15	Tv studio
Tinggi 1440 HDTV	1440 x 1152	60	47 M	60	Konsumen
Produksi yang tinggi	1920 x 1152	60	63 M	80	Film

    - Perbedaan dengan MPEG-1:
        - Dapat melakukan prediksi isi data dan prediksi frame
        - Ukuran frame bisa lebih dari 16383 x 16383
    - Bagian:
        - **Part 1** - Systems specifies the system coding layer of the MPEG-2
        - **Part 2** - Video specifies the coded representation of video data and the decoding precess required to reconstruct pictures
        - **Part 3** - Audio specifies the coded representation of audio data
        - **Part 4** - Conformance test
    - Video Stream Data Hierarchy:

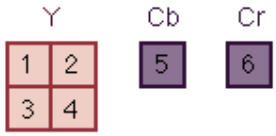




Bagian :

- Video Sequence →diawali dengan sequence header, berisi satu group gambar atau lebih, diakhiri dengan kode end-of-sequence
- GOP (Group of Pictures) → sebuah header dan rangkaian satu gambar atau lebih

- Picture → primary coding unit dari video sequence. Merepresentasikan nilai luminance (Y) dan 2 chrominance (Cb dan Cr)
- Slice → satu atau lebih macroblock. Urutannya dari kiri-kanan, atas-bawah. Penting untuk error handling. Bila terjadi error maka akan di-skip ke slice berikutnya.
- Macroblock → basic coding unit pada algoritma MPEG. 16x16 pixel segment dalam sebuah frame. Macroblock terdiri dari 4 luminance, 1 Cr, dan 1 Cb.

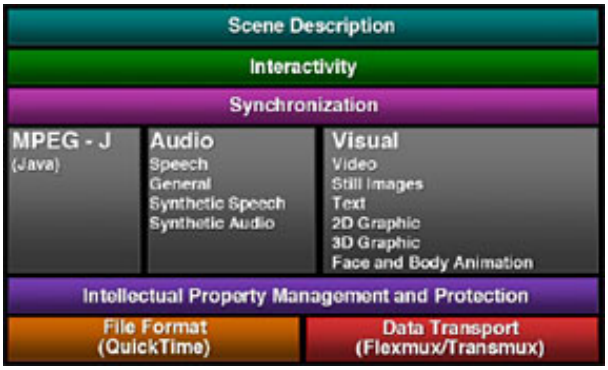


- Block →coding unit terkecil pada algoritma MPEG. 8x8 pixel, dapat berupa salah satu dari luminance rec chrominance, atau blue chrominance.

- MPEG-4
  - Versi 1 dipublikasikan Oktober 1998 sedangkan versi 2 dipublikasikan Desember 1999
  - Untuk komunikasi bitrate yang sangat rendah (4,8 sampai 64 Kb/sec): video dengan bit rate 5 Kb/s s/d 10 Mb/s dan audio dengan bit rate 2 Kb/s s/d 64 Kb/s
  - Sangat baik untuk audio/video dalam jaringan (streaming)
  - Mendukung digital rights management
  - Audio dan video adalah basis dasar dari MPEG-4, di samping itu MPEG-4 dapat mendukung objek 3D, sprites, text dan tipe media lainnya
  - Player : QuickTime (free QuickTime → play back, QuickTime Pro → author MPEG-4 content, QuickTime Streaming Server →

strean .mp4 files, Darwin Streaming Server → stream mp4 files, QuickTime Broadcaster → produce live events, making QuickTime workflow)

- Internet Streaming Media Alliance (ISMA) : Apple, Cisco, IBM, Kasenna, Philips, Sun Microsystems, AOL Time Warner, Dolby Laboratories, Hitachi, HP, Fujitsu, dan 20 perusahaan lainnya → dukungan untuk MPEG-4
- Kategori :
  - MPEG-4 Part 2 (simple profile)
  - MPEG-4 Part 10 / H.264 (high quality, low data rates, small file size, video conference with 3G, kualitas setara MPEG-2, data rate $^{1}/_{3}$ sampai ½ MPEG-2, resolusi sampai 4 kali MPEG-4 part 2)



MPEG Comparison

MPEG-1

- Approved November 1991
- VHS-quality
- Enabled Video CD
- Enabled CD- ROM
- Medium Bandwidth (up to 1.5Mbits/sec)
  - 1.25Mbits/sec video 352 x 240 x 30Hz
  - 250Kbits/sec audio (two channels)
- Non-interlaced video

MPEG-2

- Approved November 1994
- DVD-quality
- Enabled Digital TV set-top boxes
- Enabled Digital Versatile Disk (DVD)
- Higher Bandwidth (up to 40Mbits/sec)
- Up to 5 audio channels (i.e. surround sound)
- Wider range of frame sizes (including HDTV)
- Can deal with interlaced video

MPEG-3

- MPEG-3 was for HDTV application with dimensions up to 1920 x 1080 x 30Hz, however, it was discovered that the MPEG-2 and MPEG-2 syntx worked very well for HDTV rate video. Now HDTV is a part of MPEG-2 High-1440 Level and High Level toolkit.

MPEG-4

- Approved October 1998
- Scalable quality
- Based on QuickTime File Format
- Scalable delivery - from cell phones to satellite television.
- Very Low Bandwidth (64Kbits/sec)
- 176 x 144 x 10Hz
- Optimized for videophones

- AAC (Advanced Audio Coding)
    - Dasar dasri MPEG-4, 3GPP, dan 3GPP2
    - Pilihan untuk audio codec internet, wireless, dan digital broadcast
    - Mendukung audio encoding dengan kompresi lebih efisien dibandingkan MP3, dan mempunyai kualitas hampir setara CD Audio
    - Dikembangkan oleh Dolby, Fraunhofer, AT&T, Sony dan Nokia
    - Audio codec : QuickTime, iTunes, iPod
    - Kelebihan:
        - Peningkatan kompresi dengan kualitas lebih baik dan ukuran file lebih kecil

- Mendukung multichannel audio, mendukung sampai 48 full frequency channel
- High resolution audio, sampling rate sampai 96 kHz
- Peningkatan efisiensi proses decoding, pengurangan processing power untuk decoding