Kompresi Audio dan Video

Irawan Afrianto

Pendahuluan (1)

- **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, audiophile

- 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

Pendahuluan (2)

- 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 kedalam 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

Kompresi Audio MPEG (Moving Picture Expert Group)

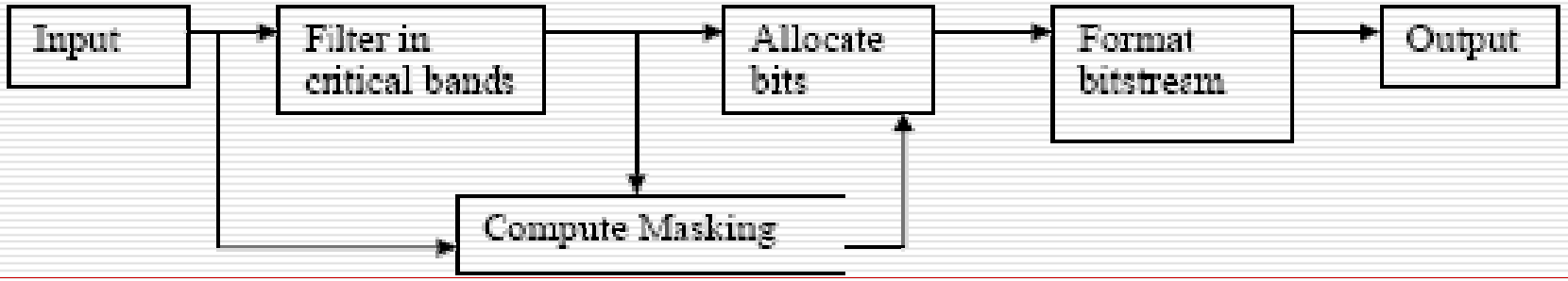
- MPEG-1 menggunakan bandwidth 1,5 Mbits/sec untuk audio dan video, dimana 1,2 Mbits/sec digunakan untuk video sedangkan 0,3 Mbits/sec digunakan untuk audio.

  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 x 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

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 (1)

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.

Teknik kompresi MP3 (2)

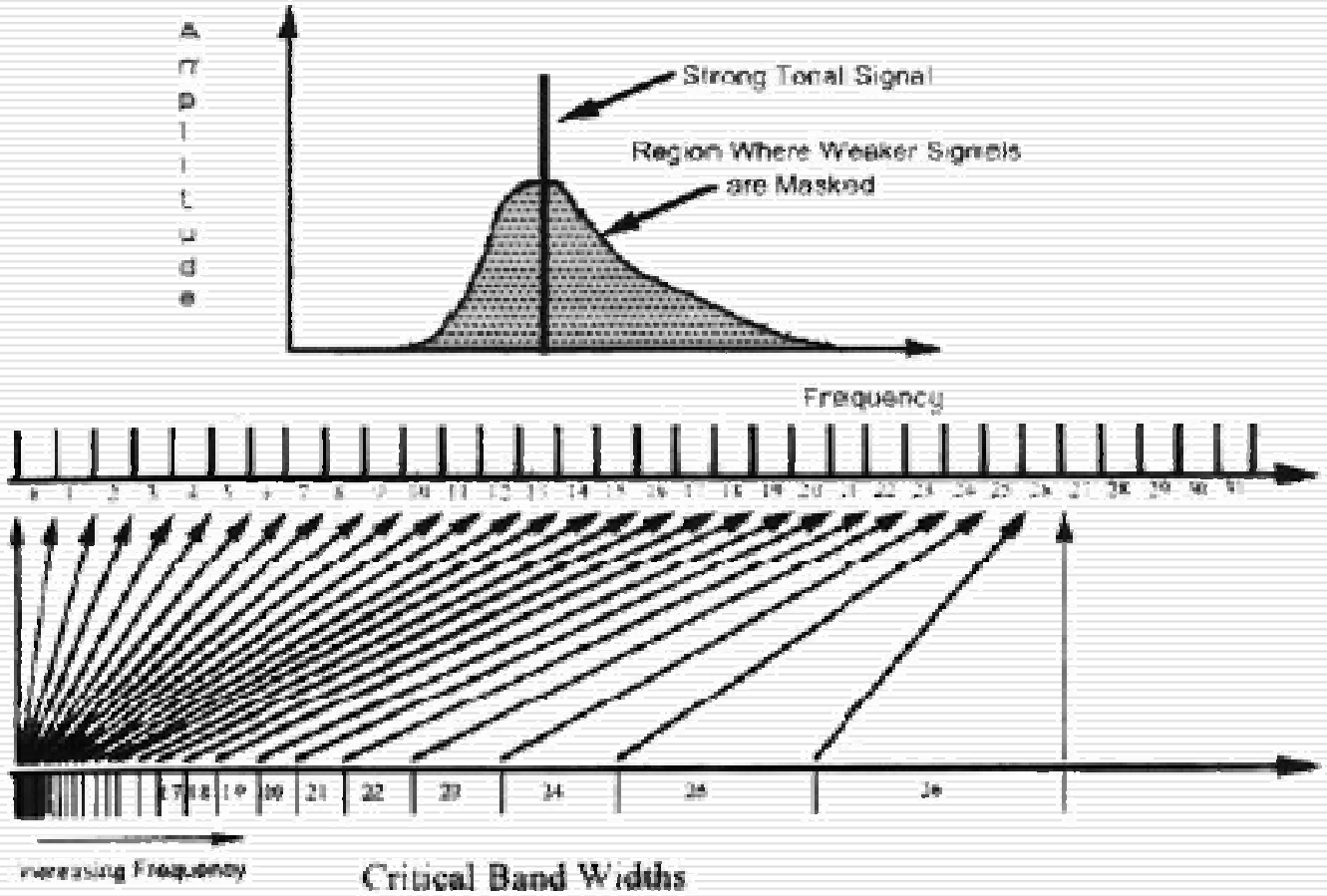
3. Critical band

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

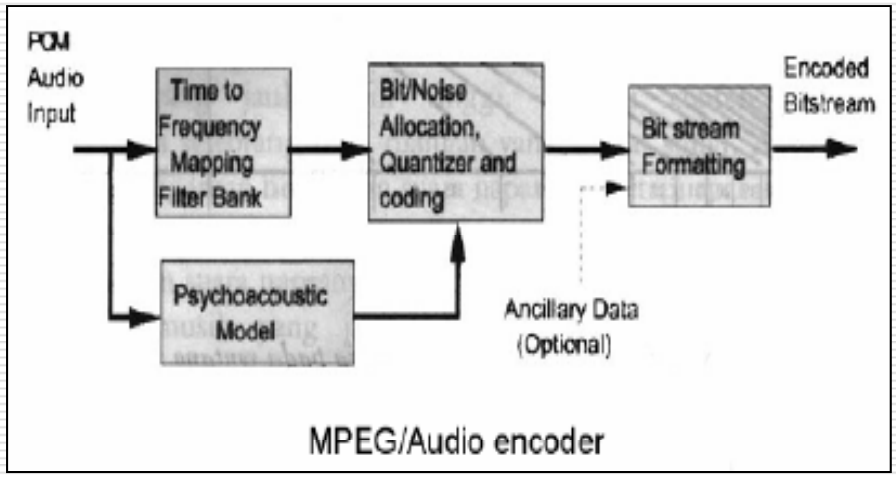
Terkadang dual channel stereo mengirimkan informasi yang sama. Dengan menggunakan joint stereo, informasi yang sama ini cukup ditempatkan dalam salah satu channel saja dan ditambah dengan informasi tertentu. Dengan teknik ini bitrate dapat diperkecil.

Teknik kompresi MP3 (3)

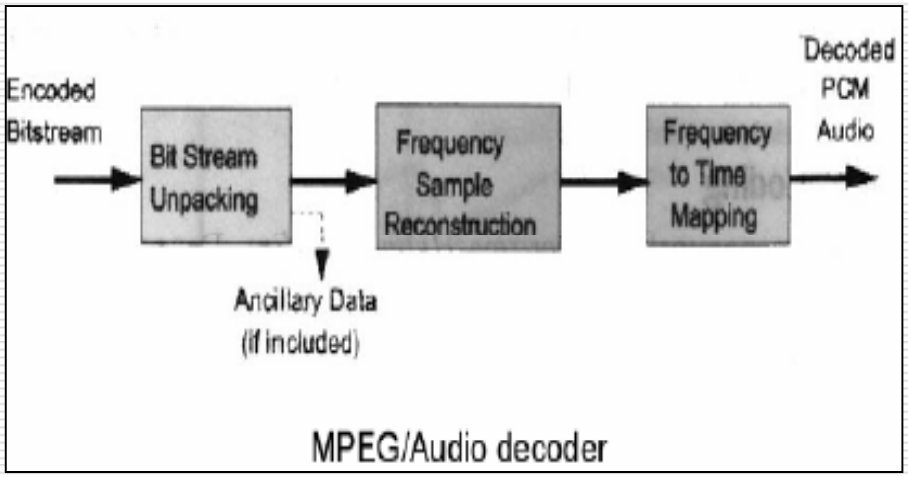


Encoder/Decoder MP3 (1)

- 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

Encoder/Decoder MP3 (2)

- **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

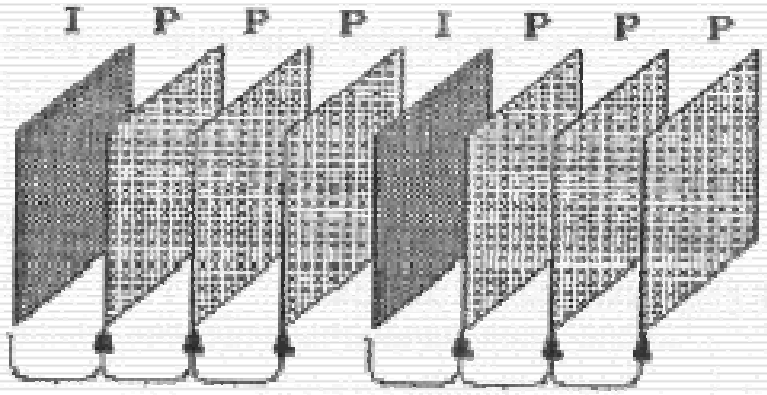
- Video memiliki 3 dimensi:
  - 2 dimensi spatial (horisontal dan vertikal),
  - 1 dimensi waktu.
- Di dalam video terdapat 2 hal yang dapat dikompresi yaitu frame (still image) dan audionya.
- Data video memiliki:
  - redundancy spatial (warna dalam still image)
  - redundancy temporal (perubahan antar 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.

Teknik Video Coding (H261 & H263) - 1

- 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

Teknik Video Coding (H261 & H263) - 2

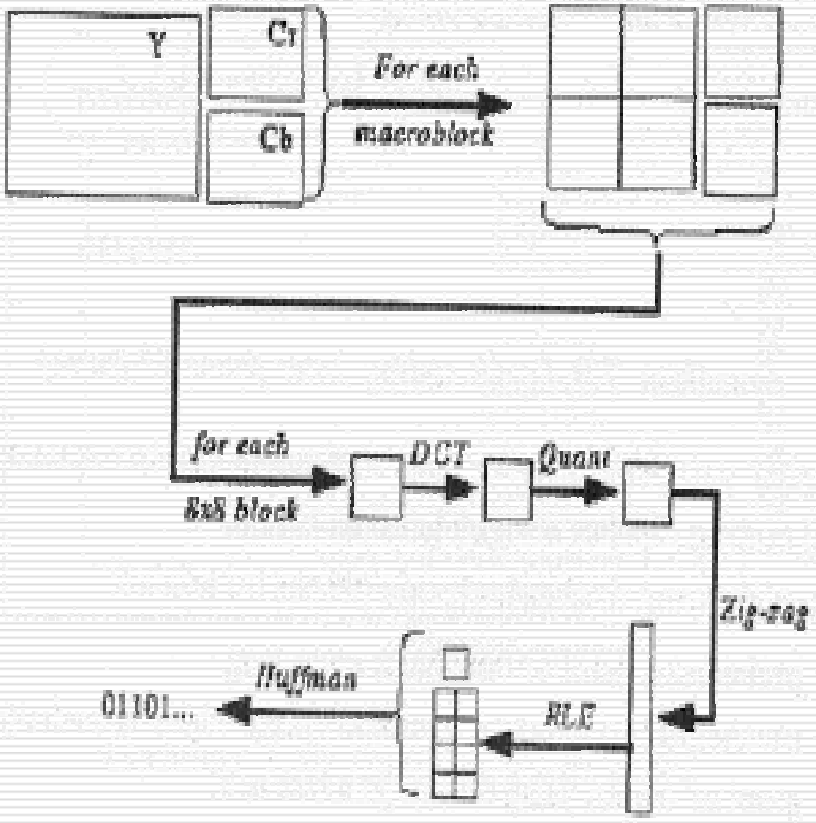
- Mempunyai 2 tipe frame yaitu: Intra-frame (I-frame) dan Interfame (P-frame)
  - I-frame digunakan untuk mengakses banyak pixel
  - P-frame digunakan sebagai “pseudo-differences“ dari frame yang sebelumnya ke frame sesudahnya, dimana antar frame terhubung satu sama lain.



Teknik Video Coding (H261 & H263) - 3

- Intraframe coding

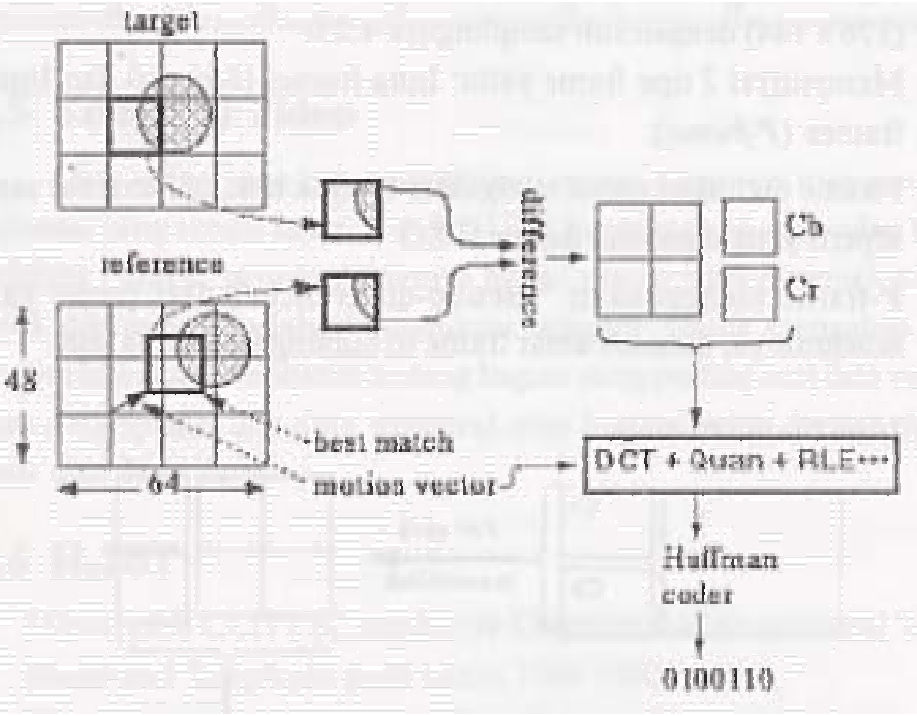
  Makroblok yang digunakan pada gambar asli adalah 16 x 16 pixel perblok, dimana Y menggunakan 4 blok, U (Cr) menggunakan 1 blok, dan V (Cb) menggunakan 1 blok.



Teknik Video Coding (H261 & H263) - 3

- Interframe coding

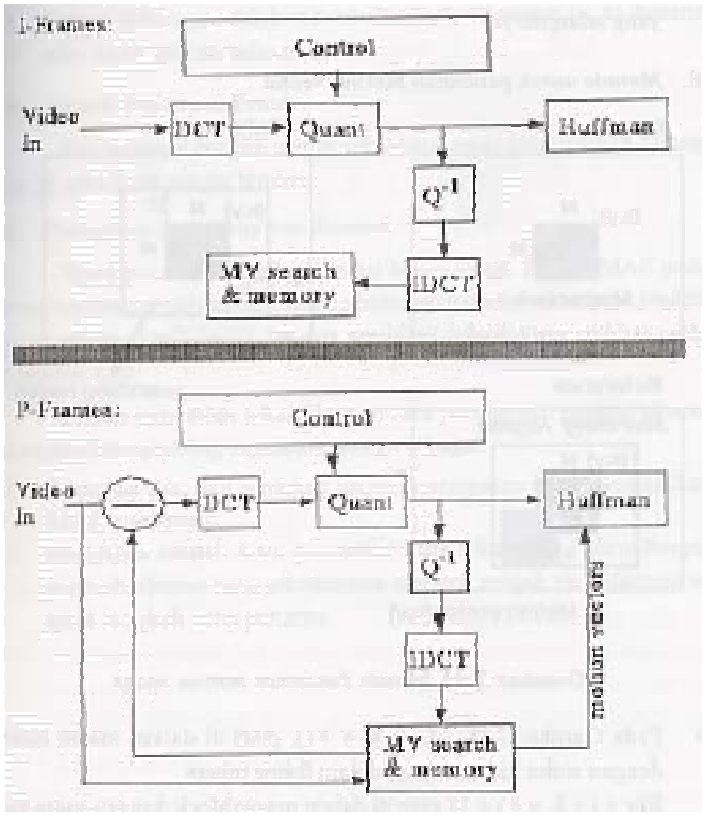
Gambar sebelumnya dijadikan gambar acuan yang akan dibuat gambar hasilnya, dengan menggunakan RMSE untuk mencari tingkat error yang paling kecil.



Teknik Video Coding (H261 & H263) - 4

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 dikembangkan H.263 untuk encoding video pada bitrate rendah

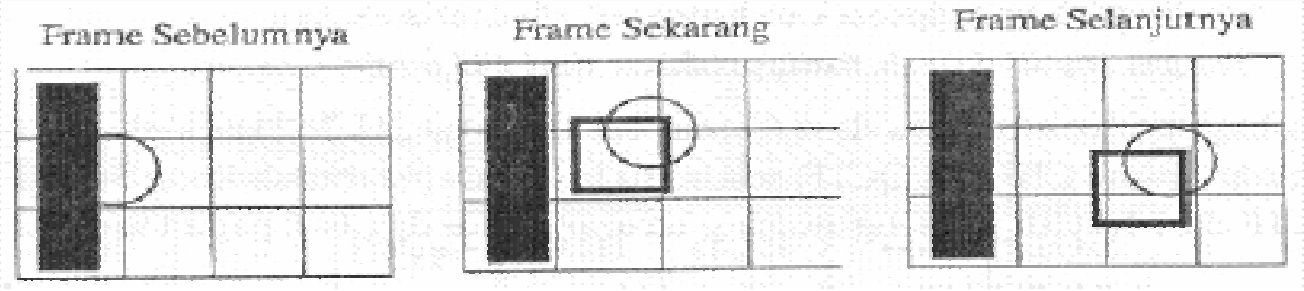


Teknik Video Coding (H261 & H263) - 5

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	disarankan	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

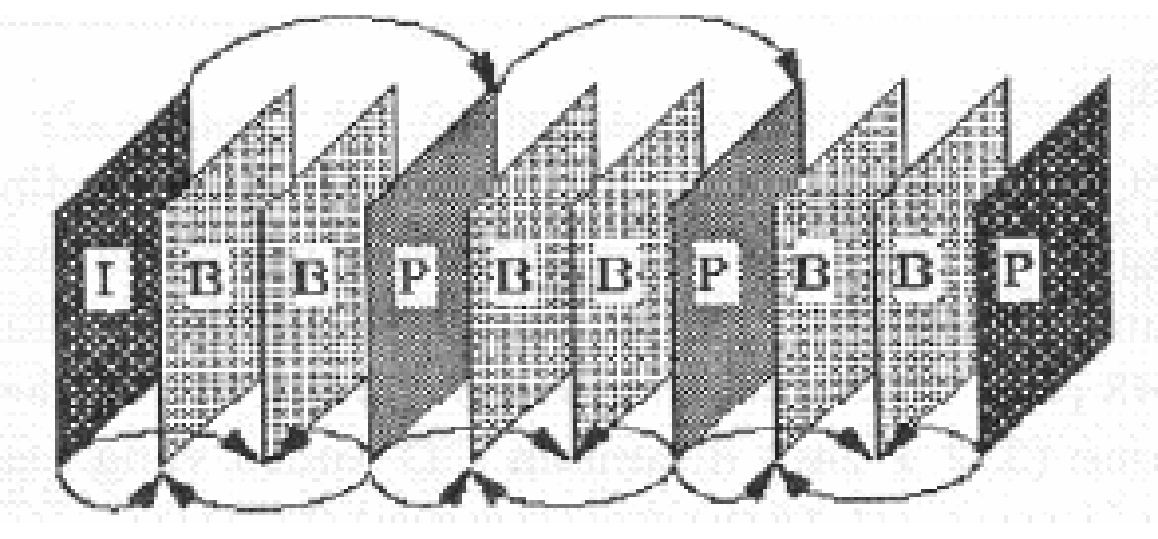
MPEG audio-video

- 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 audio-video

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



MPEG COMPARATION (1)

- **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 COMPARATION (2)

- **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 COMPARATION (3)

- **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

MPEG COMPARATION (4)

- **AAC (Advanced Audio Coding)**
  - Dasar dari 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

Terima Kasih