

BAB II

PENGENALAN KOMPRESI MPEG

Kompresi, reduksi bitrate dan reduksi data merupakan istilah yang pada dasar memiliki maksud yang sama. Pada dasarnya adalah informasi dibawa atau dikirim dengan jumlah atau kecepatan data yang lebih kecil. Pada MPEG ini berarti penggunaan bitrate yang lebih kecil.

Di bawah ini merupakan alasan atau sebab digunakannya kompresi :

1. Kompresi memperpanjang waktu dengan besar penyimpanan yang sama
2. Memungkinkan untuk miniaturisasi. Dengan data yang lebih sedikit, tempat penyimpanan dapat dibuat lebih kecil
3. Dalam hal transmisi, bandwidth yang dipakai akan lebih kecil

2.1 MPEG

MPEG adalah singkatan dari *Motion Picture Expert Group* yang merupakan suatu komite yang dibentuk oleh ISO (*International Organization for Standardization*) untuk membuat suatu standar kompresi untuk audio dan video digital dengan nama asli ISO/IEC JTC1 SC29 WG11.

ISO: *International Organization for Standardization*

IEC: *International Electrotechnical Commission*

JTC1: *Joint Technical Committee 1*

SC29: *Sub-committee 29*

WG11: *Working Group 11 (moving pictures with audio)*

Sejauh ini telah dibuat 3 Standar yaitu MPEG-1, MPEG-2 dan MPEG-4. Dan lebih lanjut sedang dikembangkan MPEG-7 dan MPEG-21.

Standarisasi yang ditentukan oleh ISO untuk kompresi ini bukan untuk menentukan *decoder* atau *encoder*-nya melainkan *bitstream*-nya. Dengan begitu perubahan *encoder/decoder* di masa depan dengan beberapa perbaikan tidak akan menjadi masalah dengan *bitstream* yang dibuat sekarang, begitu pula sebaliknya.

MPEG-1 merupakan standar kompresi audio video yang pertama. Kompresi ini ditujukan untuk membuat data video yang cukup untuk ditempatkan ke dalam suatu CD (*Compact Disc*) dengan kecepatan 1 kali, yaitu $650\text{MB}/74\text{menit} = 8,78\text{ MB per menit}$ atau kira-kira 1,17 mega bit per detik. Walaupun pada kenyataan *bitstream* file MPEG dengan kecepatan 10 MB / menit yang akan dimasukkan ke dalam CD tersebut. Hal ini dimungkinkan karena parameter yang dipakai sudah standar sehingga informasi yang ada di depan setiap segmen dihilangkan.

MPEG-2 adalah kompresi generasi selanjutnya dengan beberapa tambahan seperti *interlacing*, *Variable BitRate*, dan resolusi lebih tinggi. Dalam audio MPEG-2 dapat memiliki suara surround, AC3 atau MPEG Audio multichannel. Beberapa aplikasi MPEG-2 adalah DVD (*Digital Video Disc/Digital Versatile Disc*), sama seperti CD tetapi tingkat kerapatan datanya lebih besar, dan digital broadcasting.

MPEG-4 adalah kompresi audio video dengan menganggap audio dan video sebagai satu objek yang utuh, tidak seperti MPEG-1/2 yang dapat dipisahkan antara objek video dan audionya (*multiplex*).

Kebanyakan orang menyangka MP3 adalah MPEG-3, tetapi yang benar adalah MP3 merupakan kependekan dari MPEG audio layer III yang kompresi audio dari MPEG yang paling rumit dengan menggunakan variabel bitrate. Sedangkan MPEG-3 adalah kelanjutan dari MPEG-1/2 yang ditujukan untuk HDTV dengan resolusi 1920 x 1080 dengan frekuensi frame 30 Hz dan bitrate antara 20 dan 40 Mbit/sec, tetapi hal ini dapat dicapai dengan menggunakan MPEG-2 maka dari itu standar MPEG-3 ditiadakan.

2.2 Kompresi Video MPEG

Dalam kompresi data dikenal dua jenis kompresi yaitu *lossless compression* dan *lossy compression*. *Lossless compression* adalah kompresi dimana tidak ada perbedaan antara data sebelum dan sesudah kompresi. *Lossless compression* sering digunakan untuk menyimpan data dari sebuah tabel. Sedangkan *Lossy Compression* adalah kompresi dengan perbedaan antara data sesudah dan sebelum kompresi seperti MPEG ini. Hal ini dilakukan karena karakteristik dari indra manusia yaitu mata dan telinga.

Kompresi Video MPEG dibagi menjadi 2 bagian yaitu kompresi intraframe (*Spatial coding* atau *intra-coding*) dan kompresi interframe (*Temporal/inter-coding*). Dan dalam standar MPEG ini komite tersebut tidak menentukan bagaimana membuat *encoder* dan *decoder*, tetapi *bitstream* seperti apa yang akan dihasilkan oleh *encoder* dan bagaimana *bitstream* tersebut diartikan oleh *decoder*.

2.2.1 Kompresi Intra Frame

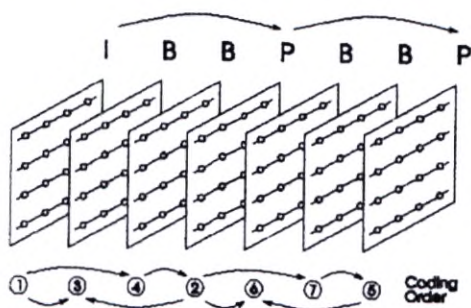
Teknik kompresi IntraFrame (*intra = within*, dalam) dalam MPEG diadaptasi dari kompresi citra tak bergerak (*still image*) JPEG, yang juga merupakan sebuah komite yang berada di bawah ISO.

Tahapan dalam kompresi IntraFrame adalah sebagai berikut:

1. Pembentukan *MacroBlock* (makroblok) 16 x 16 pixel
2. YCbCr 4:2:2 atau 4:2:0 *downsampling*, jika sumber menggunakan RGB
3. Setiap blok 8 x 8 dari Y, Cb dan Cr ditransformasikan dengan DCT menjadi DCT koefisien
4. Kuantisasi
5. RLE (*Run Length Encoding/Zig Zag Scan*)
6. VLE (*Variable Length Encoding/Entropy Coding*)

2.2.2 Kompresi Inter Frame

Dalam kompresi Inter Frame, frame (*picture*) yang dikirimkan merupakan perbedaan dari frame sebelumnya. Hal ini dididapai dengan membentuk kelompok frame, *Group of Picture* (GOP), yang berisi frame satu I dan beberapa frame P dan B.



Gambar 2.1 Contoh frame dalam satu GOP

Frame I (intra) merupakan frame yang dibuat sebagaimana aslinya yang merupakan referensi untuk frame berikutnya dalam satu GOP. Frame P merupakan frame yang berisikan perbedaan dari frame I atau frame P sebelumnya. Sedangkan frame B merupakan frame yang berisikan perbedaan dari frame I atau Frame P sebelumnya dan frame P yang ada sesudahnya. Frame B tidak pernah digunakan sebagai referensi.

Dalam pembentukan frame P dan B ada suatu teknik yang dinamakan *motion estimation* atau *motion compensation*. Pada frame P dan B beberapa makroblok dihilangkan dan diganti dengan vektor gerakan (*motion vector*) yang menunjukkan koordinat makroblok dari frame referensi.

MPEG tidak mengharuskan penggunaan frame IPB pada bitstreamnya, bitstream dapat berisi frame IP atau I saja tergantung dari orang yang membuat bitstreamnya. Begitu pula banyaknya frame P atau B yang dipakai dalam sebuah GOP dengan syarat:

1. GOP memiliki 1 frame I
2. Frame B harus berada antara I dan P atau P dan P

2.3 YCbCr *Downsampling*

Downsampling ini ditujukan untuk mengurangi banyak DCT koefisien yang dibutuhkan. Dalam RGB 16 x 16 pixel makroblok akan menghasilkan 12 8 x 8 DCT koefisien, 4 DCT koefisien pada masing – masing warna. Sedangkan pada YCbCr

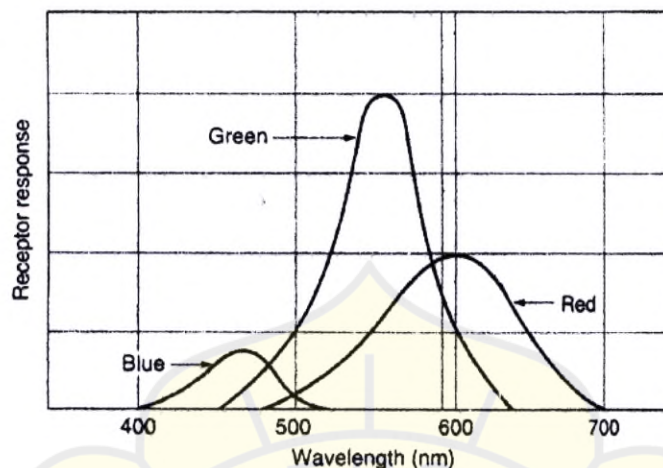
4:2:2 akan dihasilkan 8 DCT koefisien (4 Y dan 2 pada masing-masing Cb dan Cr) dan 6 DCT koefisien pada YCbCr 4:2:0 (4 Y, 1 Cb, 1 Cr).

2.3.1 YcbCr (Composite dan S-video)

Monitor komputer dan televisi menggunakan dua sinyal warna yang berbeda. Pada monitor komputer digunakan sinyal warna RGB sedangkan televisi pada umumnya menggunakan sinyal warna YUV atau YCbCr.

Sinyal warna YUV dibagi menjadi 2 jenis sinyal yaitu luminasi (luminance) dan krominasi (chrominance). Luminasi (Y) merupakan tingkat keputihan yang dapat ditentukan dengan menjumlahkan sinyal RGB dengan perbedaan pada masing-masing warna, ini dikarenakan sensitifitas mata terhadap masing-masing warna. Krominasi (UV) adalah sinyal yang menentukan warna. Pada YCbCr, Cb menentukan perbedaan warna ke arah biru dan Cr menentukan perbedaan warna ke arah merah.

Sinyal RGB dijumlah dengan faktor yang berbeda disebabkan perbedaan sensitifitas mata terhadap warna. Pada gambar 2.2 diperlihatkan bahwa mata lebih sensitif terhadap warna hijau.



Gambar 2.2

Sensitivitas pada warna pada panjang gelombang tertentu

Karena perbedaan tersebut maka persamaan yang menghasilkan sinyal Y yang diambil dari sinyal RGB ditentukan sebagai berikut:

$$Y = 0,3R + 0,59G + 0,11B$$

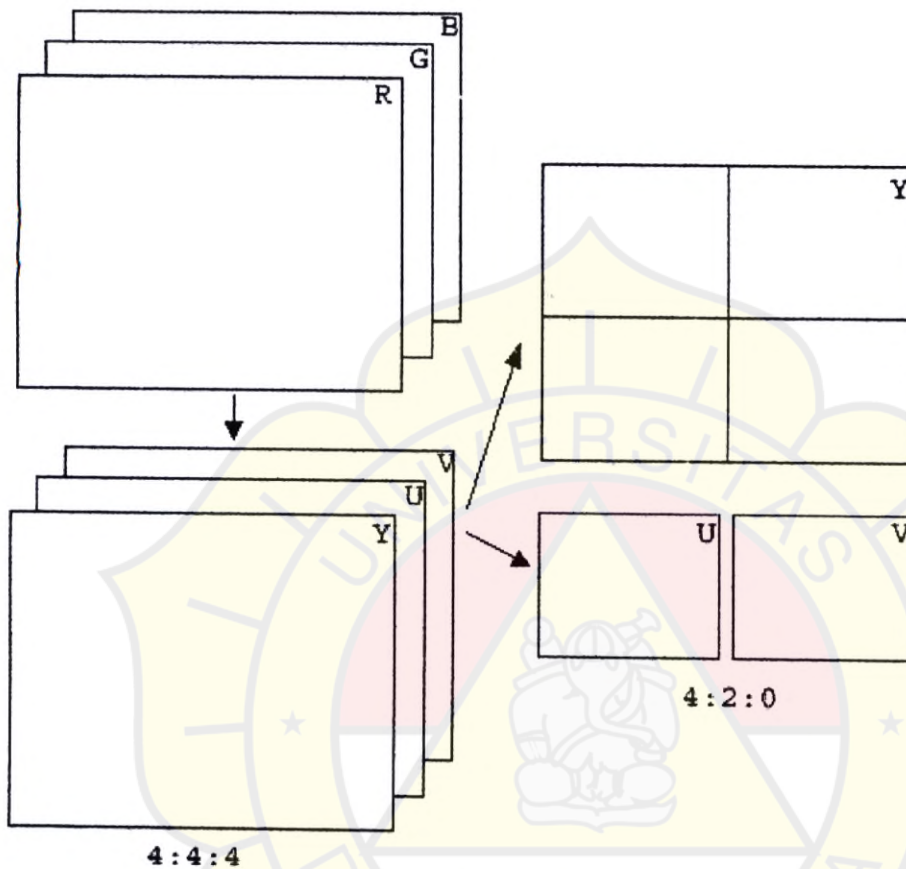
Untuk Cb dan Cr ditentukan sebagai berikut:

$$Cr = -0.421 G - 0.079 B + 0.500 R$$

$$Cb = -0.331 G + 0.500 B - 0.169 R$$

2.3.2 Downsampling CbCr (UV)

Seperti kita ketahui sebelumnya bahwa MPEG maupun JPEG menggunakan makroblok 16 x 16 pixel dan kita menggunakan YCbCr. Untuk MPEG dapat digunakan downsampling 4:2:2 atau 4:2:0 tetapi untuk kompresi maksimum kita gunakan downsampling 4:2:0. Proses downsampling dapat kita lihat pada contoh di bawah.



Gambar 2.3
Proses downsampling YCbCr dari RGB

Jadi jika proses dibalik maka 4 pixel Y akan menggunakan 1 Cb dan 1 Cr (UV) untuk membentuk 4 pixel RGB.

2.4 Transformasi Cosinus Diskrit (DCT)

Transformasi cosinus diskrit atau Discrete Cosine Transform(DCT) merupakan yang mirip dengan Transformasi Forier Diskrit tetapi hanya menggunakan komponen cosinus saja. $N \times N$ 2 Dimensional DCT didefinisikan sebagai berikut

$$F(u, v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \dots(2.1)$$

N x N 2 Dimensional DCT

dengan $u, v, x, y = 0, 1, 2, \dots, N-1$

dimana x, y adalah koordinat spasial (dua dimensi) dari domain sample

u, v adalah koordinat dalam domain transform

(frekuensi)

$N \times N$ adalah nilai matriks

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{untuk } u \text{ atau } v = 0 \\ 1 & \text{lainnya} \end{cases}$$

Gambar di bawah memberikan contoh pengaruh dari masing-masing koefisien DCT terhadap sinyal yang dihasilkan.

Untuk blok 8×8 yang akan digunakan dalam kompresi MPEG didefinisikan sebagai berikut

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \dots(2.2)$$

8 x 8 2 Dimensional DCT

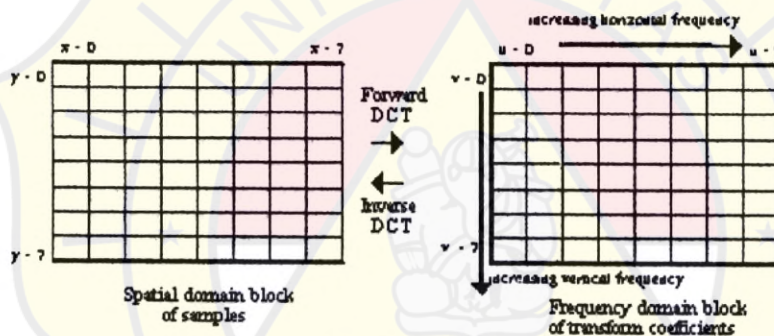
$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u, v) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \dots(2.3)$$

8 x 8 2 Dimensional iDCT

Dalam beberapa buku $f(x,y)$ dapat ditulis sebagai $f(i,j)$

Persamaan di atas dipergunakan untuk mengubah nilai masing-masing krominasi dan luminasi pada blok pixel 8×8 menjadi blok 8×8 koefisien DCT. Dan seperti telah dibahas sebelumnya bahwa YcbCr 4:2:0 akan menghasilkan 6 blok pixel 8×8 , yang terbagi atas 4 blok Y, 1 blok Cb dan 1 blok Cr. Dan keenam blok akan ditransformasikan menjadi koefisien DCT.

Pada koefisien DCT, koefisien yang paling ujung atas kiri (0,0) diberi nama

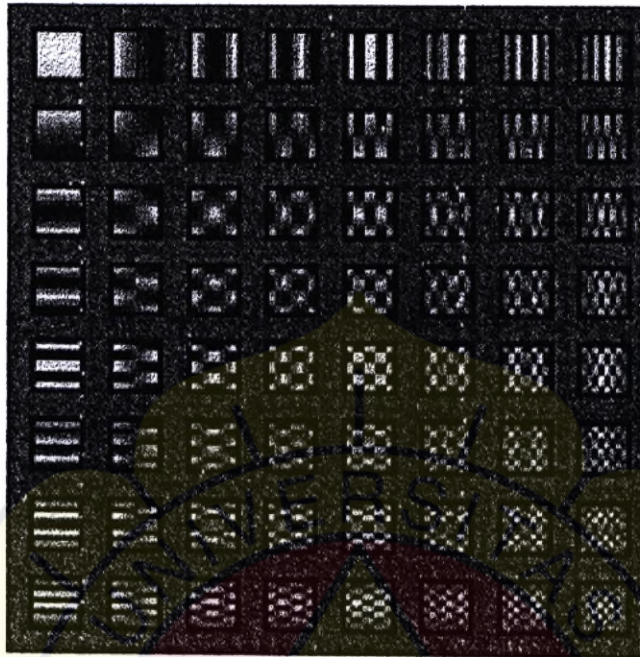


Gambar 2.4

Penggunaan DCT untuk mengubah spatial domain ke frequency domain

komponen DC yang merupakan penentu nilai rata-rata dari blok 8×8 , sedangkan yang lain diberi nama komponen AC. Komponen AC menentukan besarnya frekuensi yang semakin besar dari kiri ke kanan, mewakili frekuensi horizontal, dan semakin besar dari atas ke bawah, yang mewakili frekuensi vertikal

Dan Gambar 2.5 diperlihatkan frekuensi dari masing-masing komponen DCT



Gambar 2.5
Pengaruh koefisien DCT terhadap frekuensi sample

Pemilihan blok 8×8 dikarenakan jika lebih besar dari itu akan sangat sensitif terhadap perubahan nilai pada komponen AC.

MPEG sebenarnya hanya menentukan penggunaan iDCT untuk mengubah blok koefisien DCT kembali menjadi blok pixel 8×8 . Jika digunakan sebuah persamaan atau matriks yang lebih baik untuk menentukan blok koefisien DCT, ini dapat dipakai pada encoder MPEG.

2.5 Kuantisasi, RLE dan VLE

Kuantisasi, RLE dan VLE merupakan tahap akhir dari penghemat ruang penyimpanan dan bandwidth pada MPEG.

Kuantisasi adalah membagi masing-masing koefisien DCT dengan bilangan yang sama atau berbeda. Kuantisasi dengan nilai yang sama disebut kuantisasi seragam (Uniform Quantization), sedangkan kuantisasi dengan nilai berbeda pada masing-masing koefisien dengan suatu tabel tertentu disebut dengan kuantisasi tak seragam (Non-uniform Quantization).

RLE, atau sering juga disebut sebagai ZigZag Scan, adalah suatu cara untuk mengubah blok DCT 8×8 menjadi blok DCT 1×64 dengan menempatkan frekuensi terendah pada bagian awal.

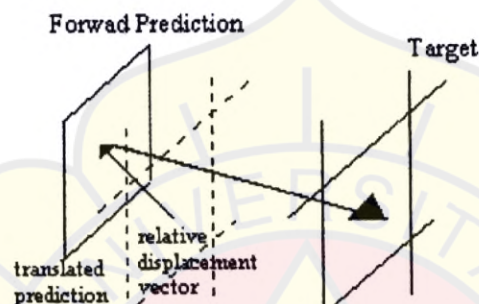
	u							
	0	1	2	3	4	5	6	7
0	0	1	5	6	14	15	27	28
1	2	4	7	13	16	26	29	42
2	3	8	12	17	25	30	41	43
3	9	11	18	24	31	40	44	53
4	10	19	23	32	39	45	52	54
5	20	22	33	38	46	51	55	60
6	21	34	37	47	50	56	59	61
v 7	35	36	48	49	57	58	62	63

Gambar 2.6
Urutan ZigZag Scan mulai dari 0 sampai 63

VLE atau entropy coding atau Variable Length Coding (VLC) adalah cara untuk menghemat ruang penyimpanan karena pada dasarnya pengubahan blok pixel 8×8 ke blok DCT 8×8 tidak mengurangi ruang penyimpanan. Hal ini dicapai dengan menghilangkan 0 yang ada di dalam blok DCT.

2.6 Motion Vector dan Motion Estimation

Pada dasarnya proses kompresi interFrame adalah dengan mengirimkan sebuah frame yang berisi bagian yang berbeda dengan frame sebelumnya. Hal ini dirasakan belum mengurangi ruang penyimpanan secara maksimal.



Gambar 2.7
Motion vector

Cara yang dipakai selanjutnya adalah dengan mengirimkan *motion vector* atau vektor gerakan. Hal ini dilakukan dengan mengirimkan vektor gerakan sebuah makroblok dari frame referensi.

Teknik mencari *motion vector* disebut sebagai *motion estimation*. MPEG tidak menentukan bagaimana cara mencari mencari *motion vector* tersebut, hal ini diserahkan kepada pembuat encoder, tetapi hanya menentukan bagaimana *motion vector* tersebut diterjemahkan oleh decoder

2.7 Tipe Frame

2.7.1 Frame I dan P

Frame P tidak bisa terus menerus digunakan karena ini akan cenderung menyebabkan masalah. Masalah yang timbul adalah ketika suatu transmisi ditangkap

ditengah suatu bitstream, maka decoder tidak membentuk sebuah frame karena tidak adanya frame I maka dari itu frame I dikirimkan secara periodik misalnya 1 dalam 5 frame. Sekumpulan frame dengan satu frame dan beberapa frame lainnya yang berada pada dua frame I disebut Group of Picture atau GOP.

Frame I akan berisi frame sebagaimana aslinya. Tidak ada makroblok yang dilewatkan. Sedangkan frame P berisi perbedaan dan motion vector dari frame sebelumnya.

2.7.2 Frame B

Frame B adalah frame yang berisi perbedaan dan motion vector dari frame sebelum dan/atau sesudahnya tetapi motion vector pada frame B tidak dapat hanya dari frame dari sebelumnya.

Dengan begitu frame B ini dapat memperkecil data sehingga kompresi yang lebih kecil dapat dicapai, tapi frame B ini dapat menambah kompleksitas dalam perhitungan dalam proses decoding.

Seperti yang telah disebutkan sebelumnya dalam GOP bisa terdapat 3 jenis frame yaitu frame I, P dan B. Jika maka urutan dalam pengiriman frame dalam bistream dan pada urutan yang sebenarnya berbeda. Karena decoder harus membentuk frame I dan P atau 2 frame P terlebih dahulu sebelum membentuk frame B yang berada diantaranya maka P yang berada sesudah frame B dikirim terlebih dahulu.