Dasar Citra/Video & Sistem Visual Manusia

Pendahuluan

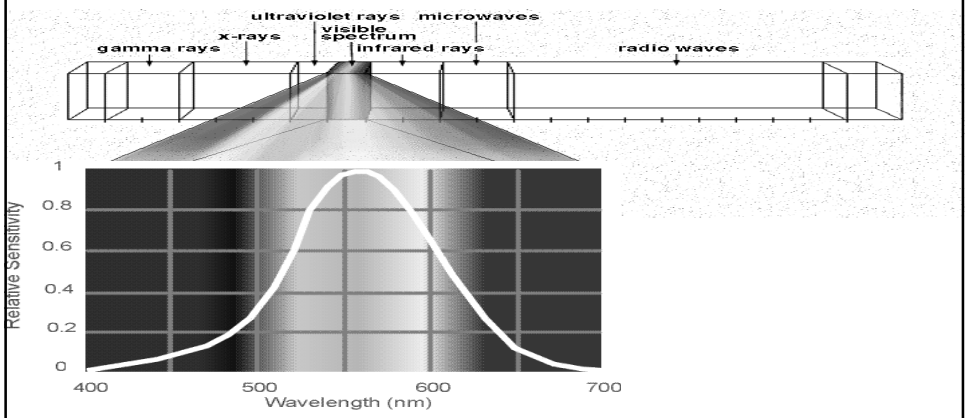
- Citra (image) adalah representasi dua dimensi dari dunia visual
- Harga *(color value*) pada sembarang titik pd citra/frame video mencatat cahaya yg dipancarkan atau direfleksikan dari titik tertentu dari 3-D *scene* yg diobservasi

Cahaya dan Warna

Apakah cahaya?

- Cahaya (*light*) → radiasi elektromagnetik yang dapat menstimulasi human visual response
- Hanya radiasi pada panjang gelombang tertentu visible → 380 nm - 780 nm
- Sumber cahaya memancarkan energi dlm satu range panjang glb dan intensitasnya bervariasi dlm waktu dan ruang → distribusi intensitas cahaya $C(x,y,z,\lambda)$
- Jika sumber cahaya terdiri dari satu panjang gelombang → sumber monochromatic
- Sumber monochrome memp intensitas, yaitu energi yg dipunyai sumber
- Sumber lain (non-monochromatic) merupakan komposisi dari satu range panjang glb, masing-masing memp intensitas sendiri

Spektrum Elektromagnetik



Fungsi sensitivitas luminance manusia

http://www.yorku.ca/eye/photopik.htm

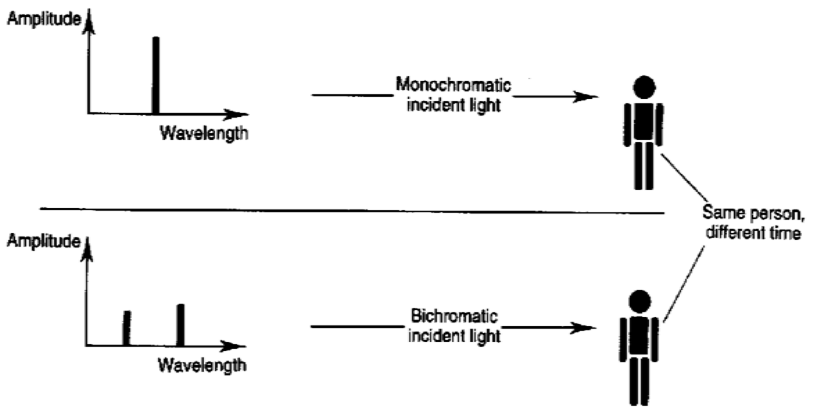
Cahaya dan Warna

Apakah Warna (color)?

- Colors adalah sensasi yang diterima mata
- Persepsi warna yang diterima tergantung pada isi spektral (komposisi λ)
- Mis. Cahaya yg memp energi terkonsentrasi sekitar 700 nm tampak merah
- Cahaya yg memp energi sama pada keseluruhan band visible tampak putih
- Panjang glb yg berbeda menghasilkan impresi/sensasi warna yg berbeda
- Dengan energi yang sama, cahaya monochromatic dari dua panjang gelombang yang berbeda akan menghasilkan visual response yang berbeda
- Sembarang colors dapat diperoleh dari percampuran tiga primary colors berbeda

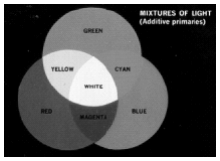
Cahaya dan Warna

- Dua distribusi spectral berbeda dapat menghasilkan sensasi color yang sama

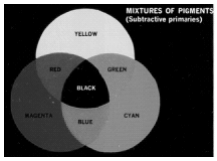


Iluminasi dan Refleksi Cahaya

- Sumber Iluminasi
  - Yang memancarkan cahaya (mis. matahari, bola lampu, monitor)
  - Persepsi warna yg diterima tergantung frekuensi yg dipancarkan
  - Mengikuti additive rule
    R + G + B = White

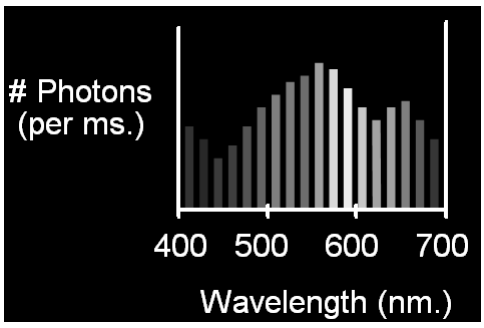


- Sumber Refleksi
  - Merefleksikan cahaya yg datang (mis. pakaian)
  - Persepsi warna yg diterima tergantung frekuensi yg direfleksikan (=emitted freq - absorbed freq)
  - Mengikuti subtractive rule
    C + M + Y = Black



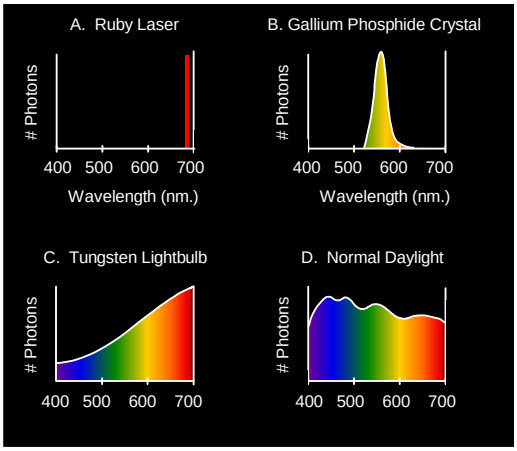
Fisika Cahaya

- Cahaya dp secara lengkap dideskripsikan secara fisik oleh spektrumnya: jumlah photon (per unit waktu) pd tiap panjang gelombang 400 – 700 nm



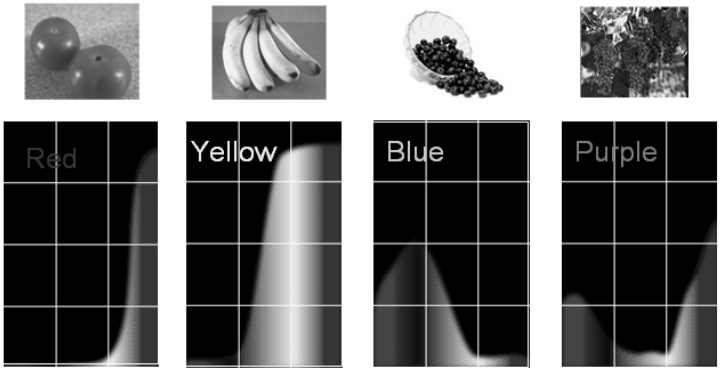
Fisika Cahaya

- Contoh spektra macam-macam sumber cahaya (sumber iluminasi)



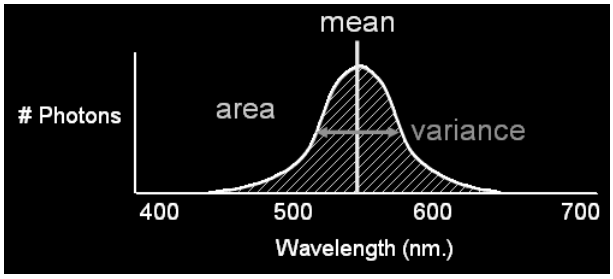
Fisika Cahaya

- Contoh bbrp spektra pantulan permukaan (sumber refleksi)

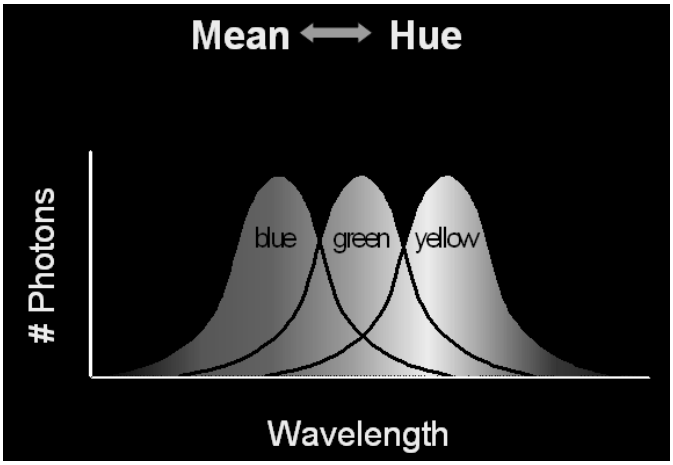


Korespondensi Psychophysical

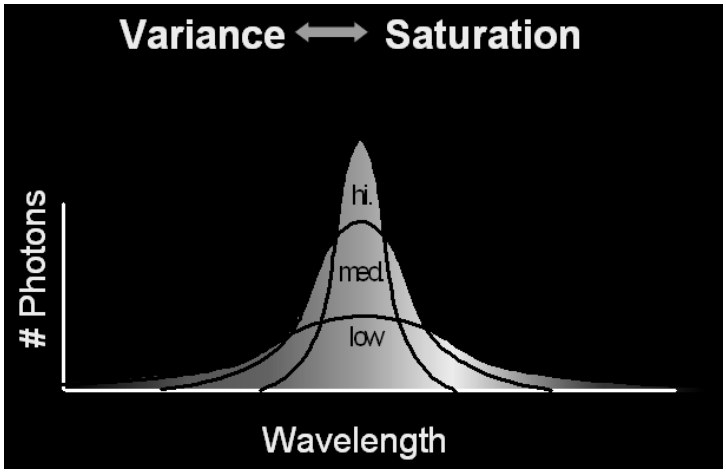
- Tidak ada penjelasan fungsional sederhana utk persepsi warna utk semua cahaya dibawah semua kondisi, tetapi …
- Sebagai bantuan:
  - Perhatikan physical spectra sbg distribusi normal



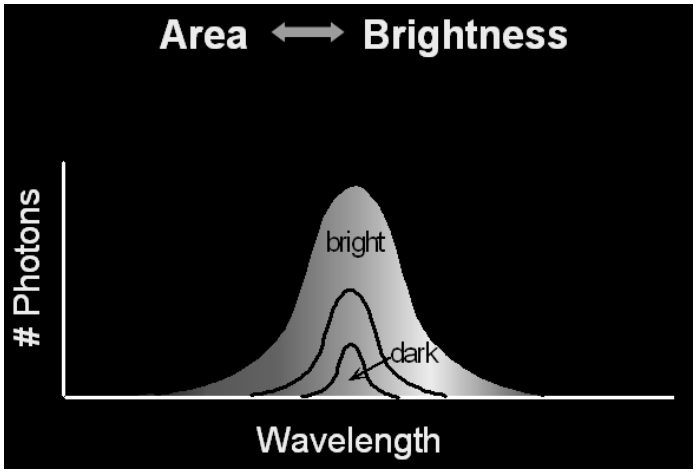
Korespondensi Psychophysical



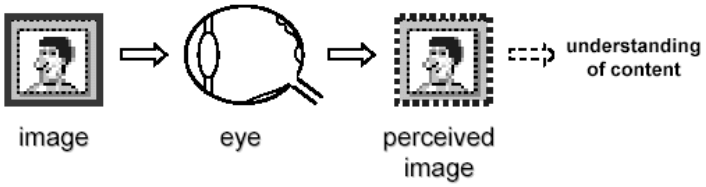
Korespondensi Psychophysical



Korespondensi Psychophysical

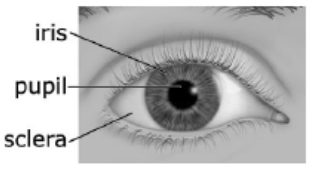


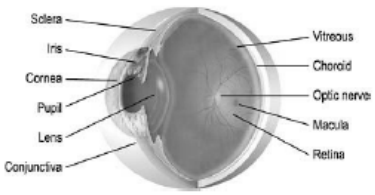
Pemrosesan Informasi oleh Observer Manusia



- Persepsi visual
  - Berhubungan dg bagaimana persepsi thd citra oleh observer manusia
    - Pemrosesan awal oleh mata
    - Pemrosesan lebih jauh oleh otak
  - Penting utk mengembangkan image fidelity measure
    - Diperlukan utk perencanaan & evaluasi algoritma & sistem DIP/DIV

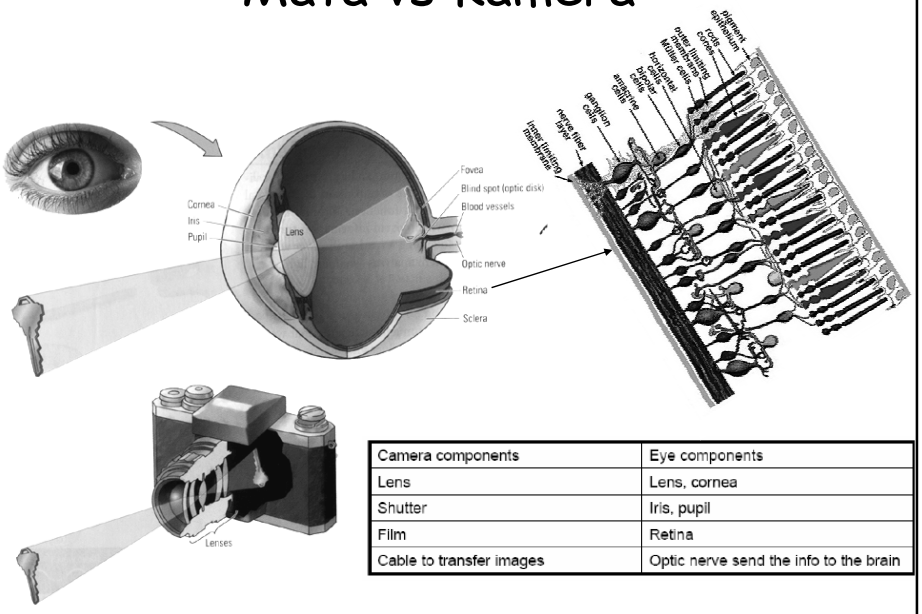
Anatomi Mata





From:
http://www.stlukeseye.com/Anatomy.asp

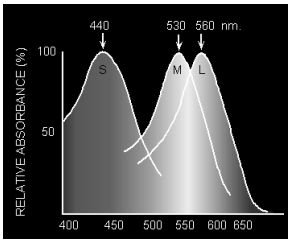
Mata vs Kamera



Camera components	Eye components
Lens	Lens, cornea
Shutter	Iris, pupil
Film	Retina
Cable to transfer images	Optic nerve send the info to the brain

Persepsi Warna Manusia

- Retina berisi photo receptors
  - **Cones**: day vision dp melihat (persepsi) color tone (Hue)
    - Tiga tipe cones (Red, green & blue cones) → overlapping passband dg puncak sekitar merah (560 nm), hijau (530 nm) dan biru (440 nm)
    - Teori Tri-receptors color vision [Young 1802]
  - **Rods**: night vision, persepsi hanya brightness
- Sensasi color dikarakteristikkan oleh
  - Luminance (brightness)
  - Chrominance
    - Hue (color tone)
    - Saturation (color purity)
- Respon receptor thd cahaya dg distribusi C(λ):



$$C_i = \int C(\lambda) a_i(\lambda) d\lambda, \quad i = r, g, b, y$$

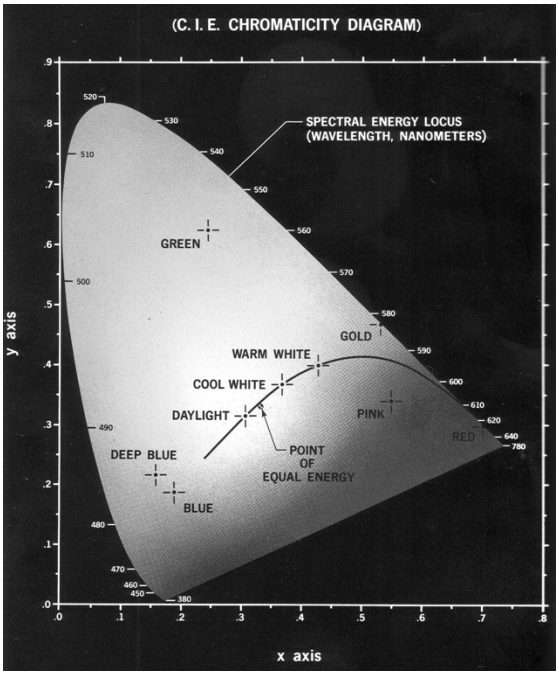
Dimana $a_r(\lambda)$, $a_g(\lambda)$, $a_b(\lambda)$ frek response dari cones

Trichromatic Color Mixing

- Dari struktur mata manusia, semua warna dipandang sebagai kombinasi variabel dari *primary color*: red (R), green (G) dan blue (B)
- Utk standarisasi CIE (Commission Internationale de l'Eclairage – the International Commission on Illumination):
  - Blue = 435,8 nm
  - Green = 546,1
  - Red = 700 nm
- Nilai R, G, dan B yg diperlukan utk membentuk sembarang color disebut nilai tristimulus: X, Y, Z
  - Suatu color dispesifikasikan oleh trichromatic coefficient:

$$x = Z/(X+Y+Z)$$
$$y = Y/(X+Y+Z) \quad \rightarrow \quad x + y + z = 1$$
$$z = Z/(X+Y+Z)$$

C.I.E Chromaticity Diagram



Model Representasi Color

- Model RGB
  - Red, Green dan Blue
- Model HIS
  - Hue, Saturation dan Intensity
  - Jarang digunakan
- Model YUV
  - Y adalah luminance, U dan V adalah komponen chrominance
  - Digunakan pd sistem PALTV
- Model YIQ
  - Y adalah luminance, I dan Q adalah komponen chrominance
  - Digunakan pd sistem TV NTSC
- YDbDr Model
  - Used in SECAM
- YCbCr Model
  - In JPEG and MPEG
  - U and V are shifted to non-negative
- Semua nilai YCC mempunyai korespondensi satu-satu dg RGB, conversion matrix

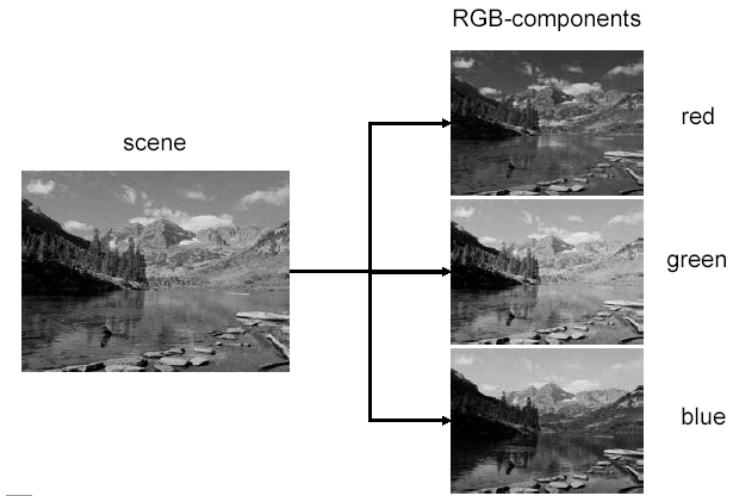
Model Representasi Color

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

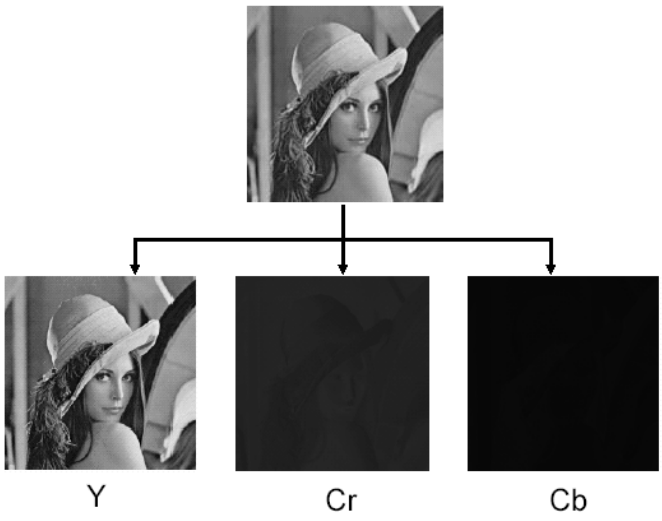
$$\begin{bmatrix} Y \\ Db \\ Dr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.450 & -0.883 & 1.333 \\ -1.333 & 1.116 & -0.217 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

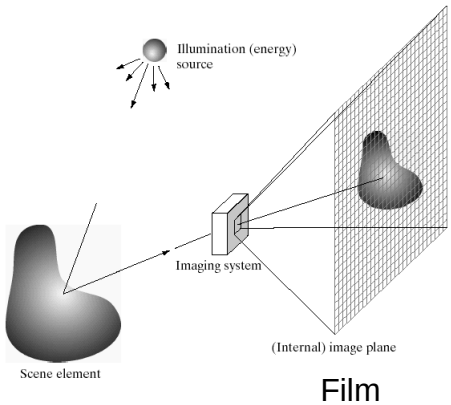
Sinyal Video RGB



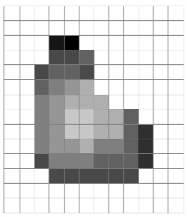
Sinyal Video YC_rC_b



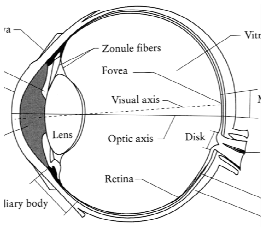
Pembentukan Image



Film



Digital Camera

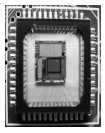


The Eye

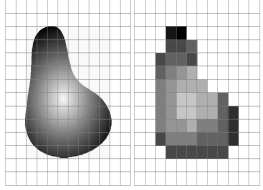
Kamera Digital

- Pada kamera digital film diganti dg sensor array
  - Tiap cell pd array adalah light-sensitive diode yg mengkonversikan photons to electrons
  - Dua tipe yg umum
    - Charge Coupled Device (CCD)
    - CMOS

http://electronics.howstuffworks.com/digital-camera.htm

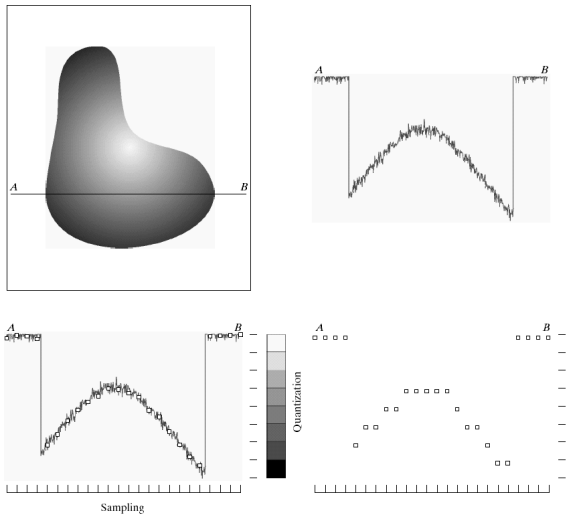


CMOS sensor



a b
FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

Sampling and Quantization

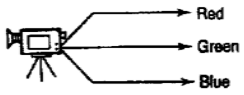


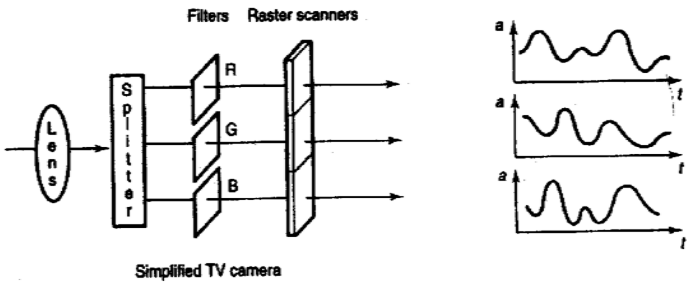
a b
c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Dasar Video : Video Analog

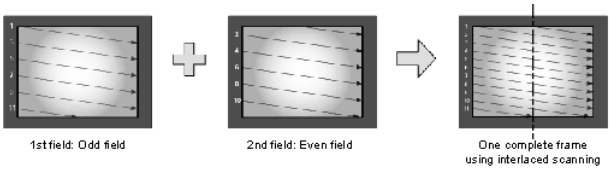
- Kamera analog menghasilkan 3 sinyal kontinyu Red, Green, Blue

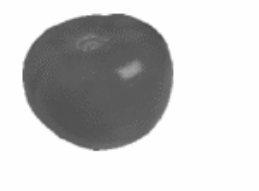


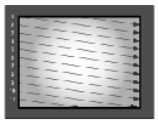


Simplified TV camera

Interlace vs. Progressive Scan







http://www.axis.com/products/video/camera/progressive_scan.htm

Progressive scan

http://www.axis.com/products/video/camera/progressive_scan.htm

Interlace

http://www.axis.com/products/video/camera/progressive_scan.htm

Dasar Video : Video Analog

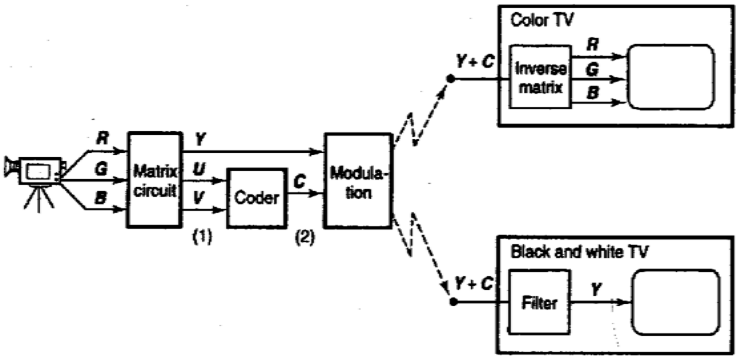
- Sinyal R, G, B dapat ditransformasikan ke tiga sinyal lain:
  - luminance (Y): memuat informasi lightness & brightness
  - dua sinyal color (color difference) : (U,V/I,Q)

    Y = 0,299R + 0,587G + 0,114B

    U = -0,147R + 0,289G + 0,436B = 0,492(B - Y)

    V = 0,615R - 0,515G - 0,100B = 0,877(R - Y)
- Alasan transformasi :
  - Sistem visual manusia kurang sensitif terhadap color (dibandingkan luminance) → sinyal color ditransmisikan/ direpresentasikan dengan BW < sinyal luminance
  - Backward compatibility
- BW luminance 5 Mhz, BW masing-masing color 1,5 Mhz

Dasar Video : Video Analog



(1) *YUV* are the luminance and chrominance signals in PAL. *YIQ* are the corresponding signals in NTSC
(2) The *U* and *V* signals are combined into a chroma signal *C*

Dasar Video : Video Digital

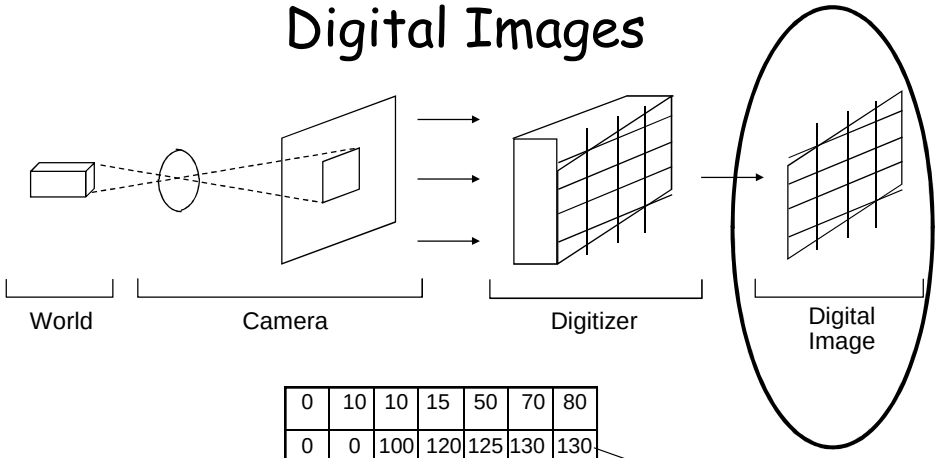
- Video digital didapat dari :
  - Kamera video digital
  - Digitalisasi video analog → CCIR-601 (ITU-R BT.601)
    - Filtering → mencegah aliasing
    - Sampling → luminance 13,5 Mhz, chrominance = 1/2 x luminance
    - Kuantisasi → 8-bit
- Color space CCIR-601:

  Y = 0,257R + 0,504G + 0,098B + 16

  Cb = -0,148R - 0,291G + 0,439B + 128

  Cr = 0,439R - 0,368G - 0,071B + 128

Digital Images

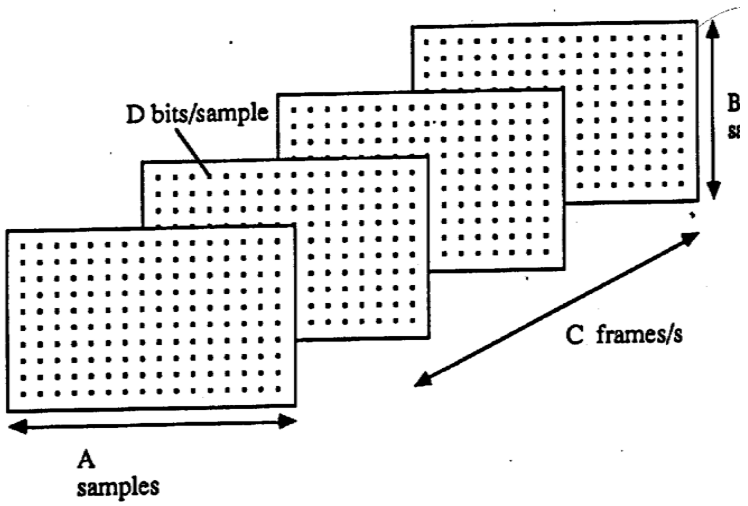


0	10	10	15	50	70	80
0	0	100	120	125	130	130
0	35	100	150	150	80	50
0	15	70	100	10	20	20
0	15	70	0	0	0	15
5	15	50	120	110	130	110
5	10	20	50	50	20	250

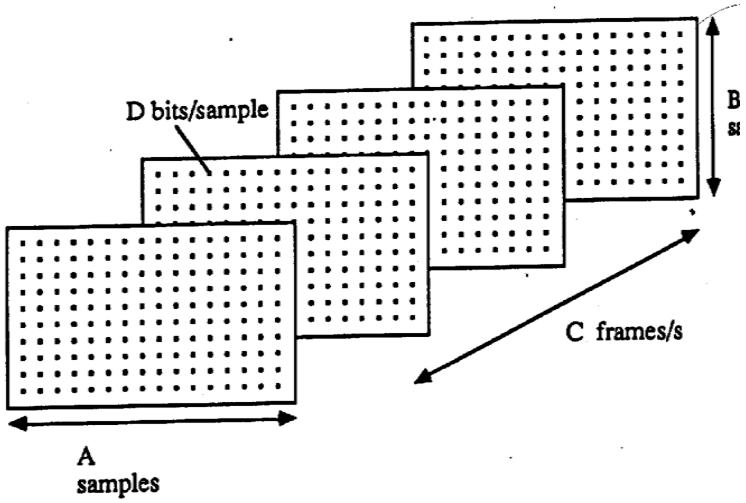
PIXEL
(picture element)

Typically:
0 = black
255 = white

Dasar Video : Video Digital



Dasar Video : Video Digital



Dasar Video : Video Digital

- Beberapa contoh parameter scanning & kuantisasi untuk sinyal luminance

SISTEM	A	B	C	D	ABCD (Mbps)
Low-quality videophone	64	64	8	6	0,2
Videoconference (CIF)	352	288	15	8	12
Digital Broadcast Television	720	576	25	8	83
HDTV	1920	1150	50	8	883
Facsimile/handwriting @ 4 line/mm, per A4 page	1200	800	0,01	1	0,01

Dasar Video: Format Video

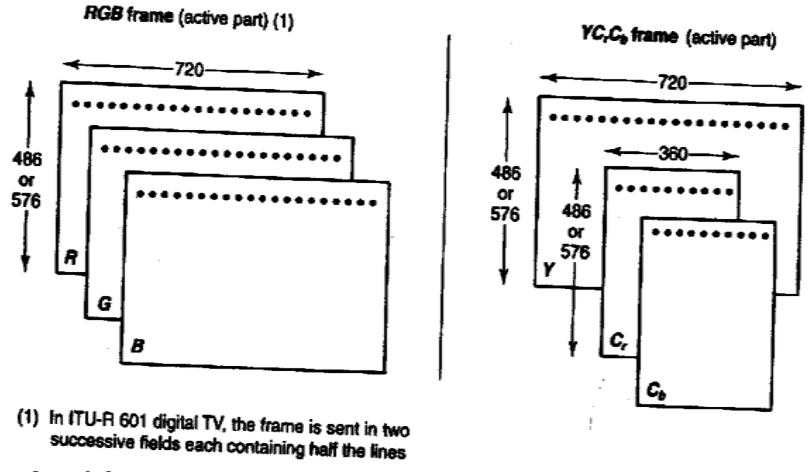
- CCIR-601/625 (Eropa)
  - jumlah garis 625 (aktif 576)
  - jumlah frame per detik 25
  - jumlah pixel per garis 720
- CCIR-601/525 (USA)
  - jumlah garis 525 (aktif 480)
  - jumlah frame per detik 30
  - jumlah pixel per garis 720

→ total pixel per detik = 10.368.000 → bit rate = 165,888,000 bit/s

→ bandwidth total = 216 Mbit/s

Dasar Video : Image Format

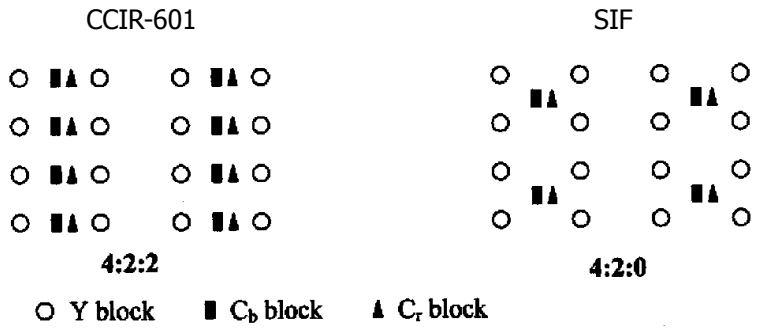
- CCIR-601 → untuk kualitas broadcats



Dasar Video : Image Format

SIF (Source Input Format)

- Eropa: 360 pixel/garis, 288 garis per gambar, 25 gambar perdetik
- USA : 360 pixel/garis, 240 garis per gambar, 30 gambar per detik
- Scanning : progressive
- Pola sampling 4:2:0



Dasar Video : Image Format

CIF (Common Intermediate Format)

- Eropa/USA: 360 pixel/garis, 288 garis per gambar, 30 gambar perdetik
- Scanning : progressive
- Pola sampling 4:2:0

QSIF, QCIF

- untuk aplikasi video over mobile networks, video telephony → kurangi rate → Eropa : 15, 10 dan 7,5 frame/s → USA 12,5 dan 8,3 frame/s
- Resolusi spatial dikurangi setengahnya untuk tiap arah → Quarter-SIF (QSIF) dan Quarter-CIF (QCIF)

Sub-QCIF

- 128 pixel x 96 pixel, 5 frame/sec, 4:2:0 → ukuran image standar paling kecil

Chrominance Subsampling Format

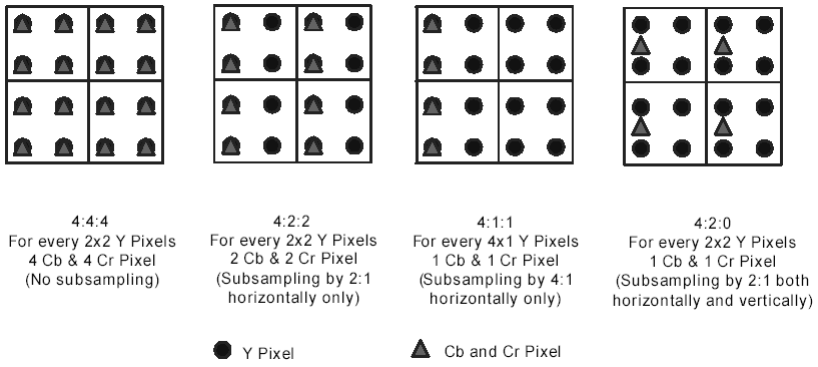
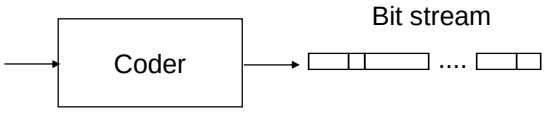
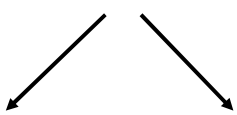


Image Coding



Objektif: Mencari cara utk merepresentasikan citra original tanpa (?) distorsi dg jumlah bit sekecil mungkin

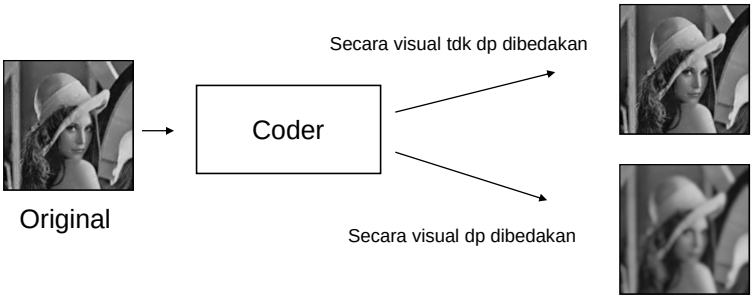


Lossless image coding

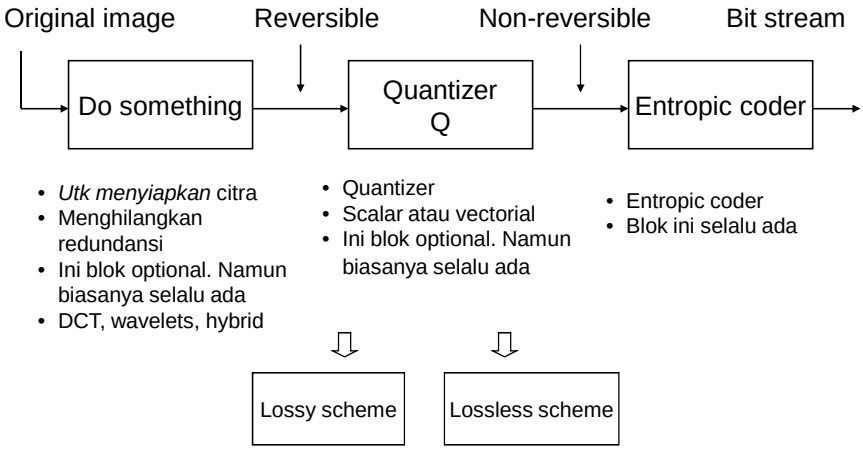
Lossy image coding

Lossless dan Lossy Image Coding

- Lossless image coding :
  image hasil decoding identik pixel per pixel dg yg original
- Lossy image coding :
  image hasil decoding TIDAK idetik pixel per pixel dg yg original



Skim Umum Image Coding (compression)



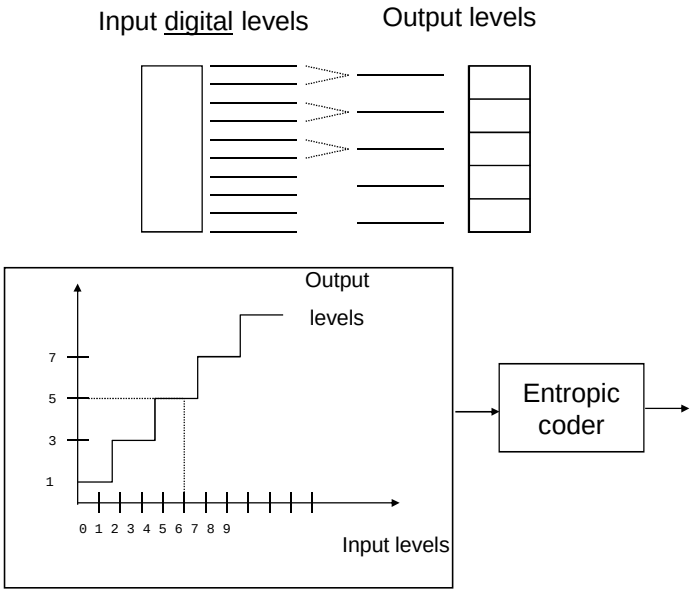
- *Utk menyiapkan* citra
- Menghilangkan redundansi
- Ini blok optional. Namun biasanya selalu ada
- DCT, wavelets, hybrid

- Quantizer
- Scalar atau vectorial
- Ini blok optional. Namun biasanya selalu ada

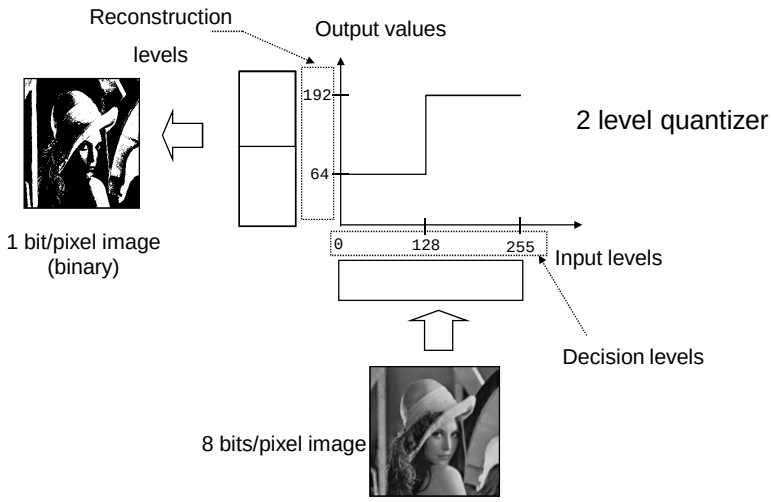
- Entropic coder
- Blok ini selalu ada

Quantization

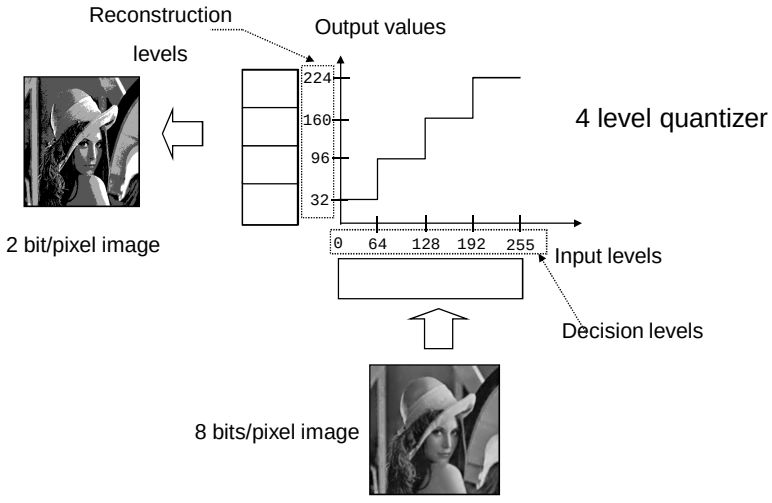
Quantization (1)



Quantization (2)



Quantization (3)



Uniform Quantizer

4 bits

2 bits

3 bits

1 bit

Ukuran Kompresi (1)

$$\text{Compression factor} = \frac{\text{Bits original image}}{\text{Bits compressed image}}$$

$$\text{Bits/pixel} = \frac{\text{Bits compressed image}}{\text{Number of pixels}}$$

Ukuran Kompresi (2)

Original image
256 x 256 x 8 bits

Compressed image
40.000 bits

$$\text{Bits/pixel} = \frac{40.000}{256 \times 256} = 0.61 \text{ bpp}$$

$$\text{C. F.} = \frac{8 \text{ bpp}}{0.61 \text{ bpp}} = 13.1$$

Ukuran Kualitas Kompresi

Coder

X(m,n) Y(m,n)

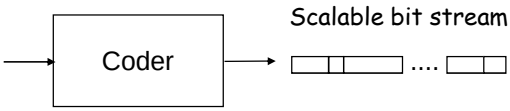
Ukuran subjektif ⇨ Opini dg skala 5

Ukuran objektif ⇨

$$\text{PSNR(dB)} = 10\log_{10}\frac{255^2}{\text{MSE}}$$

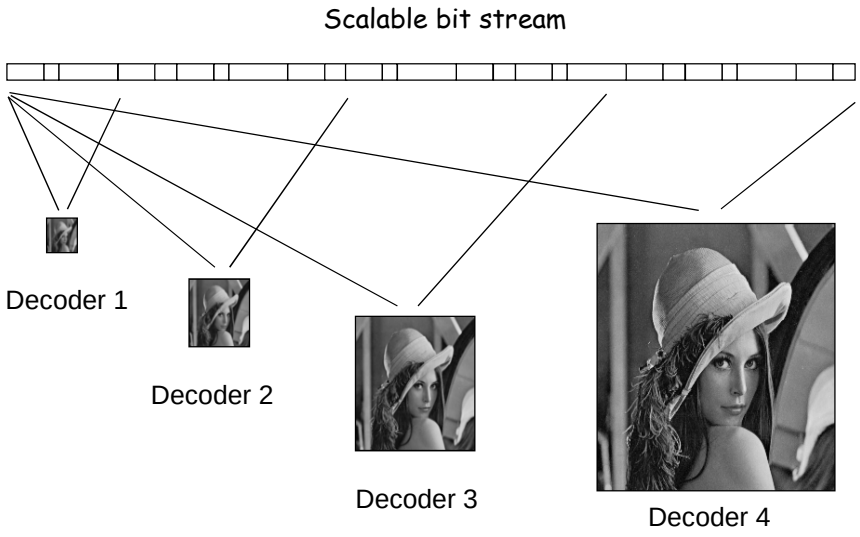
$$\text{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(y_{ij} - x_{ij})^2$$

Scalability

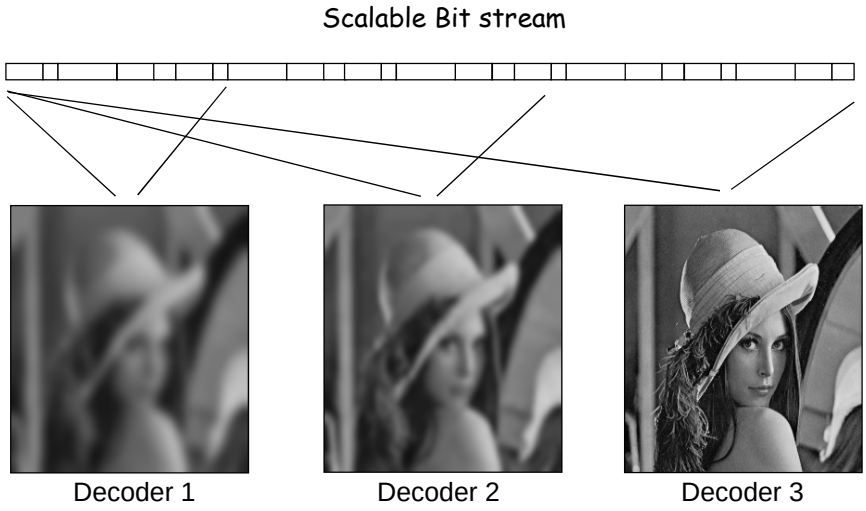


- Spatial scalability
- PSNR scalability (quality)
- Non-scalable

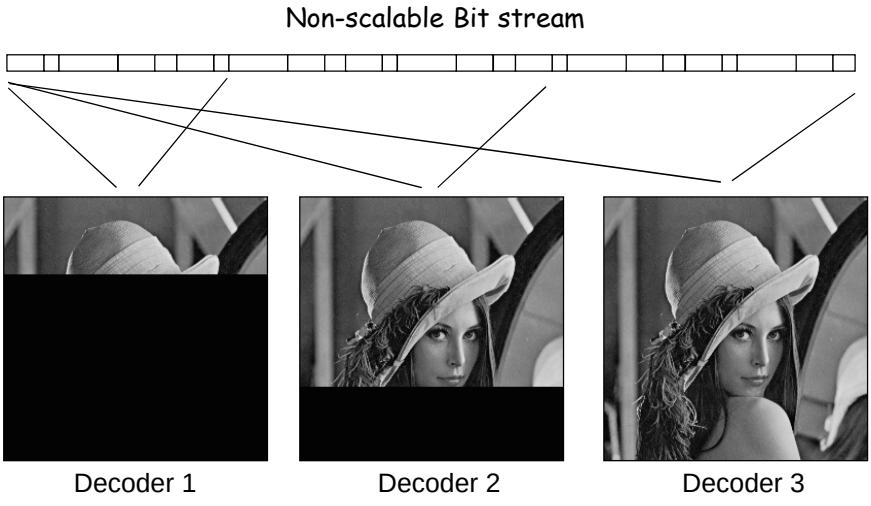
Spatial Scalability



PSNR Scalability (Quality)



Non-Scalable



Asymmetric Coder

- Kompleksitas coder berbeda drpd kompleksitas decoder
- Decoder selayaknya sesederhana mungkin

JPEG: Symmetric

MPEG 1- 2: Very asymmetric