Digitálne spracovanie zvuku, obrazu a biosignálov - DSZOB

JPEG compression

Prof. Vanda Benešová, PhD.

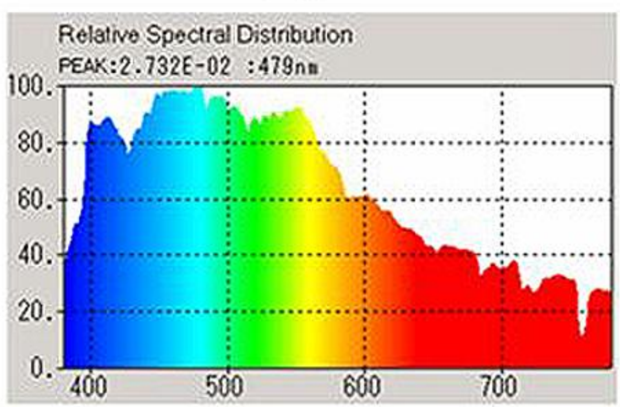
Color

Color

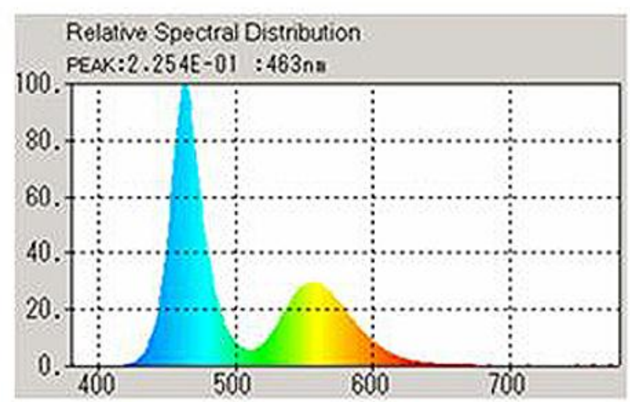
Color is the perceptual result of light in the visible region of the spectrum, having wavelengths in the region of 400 nm to 700 nm.

Color stimuli are assumed to be uniquely defined by their radiant power spectral distributions of $\{P_\lambda d_\lambda\}$ – (often in x components each representing a 10 nm(5nm) band).

Daylight



White LED

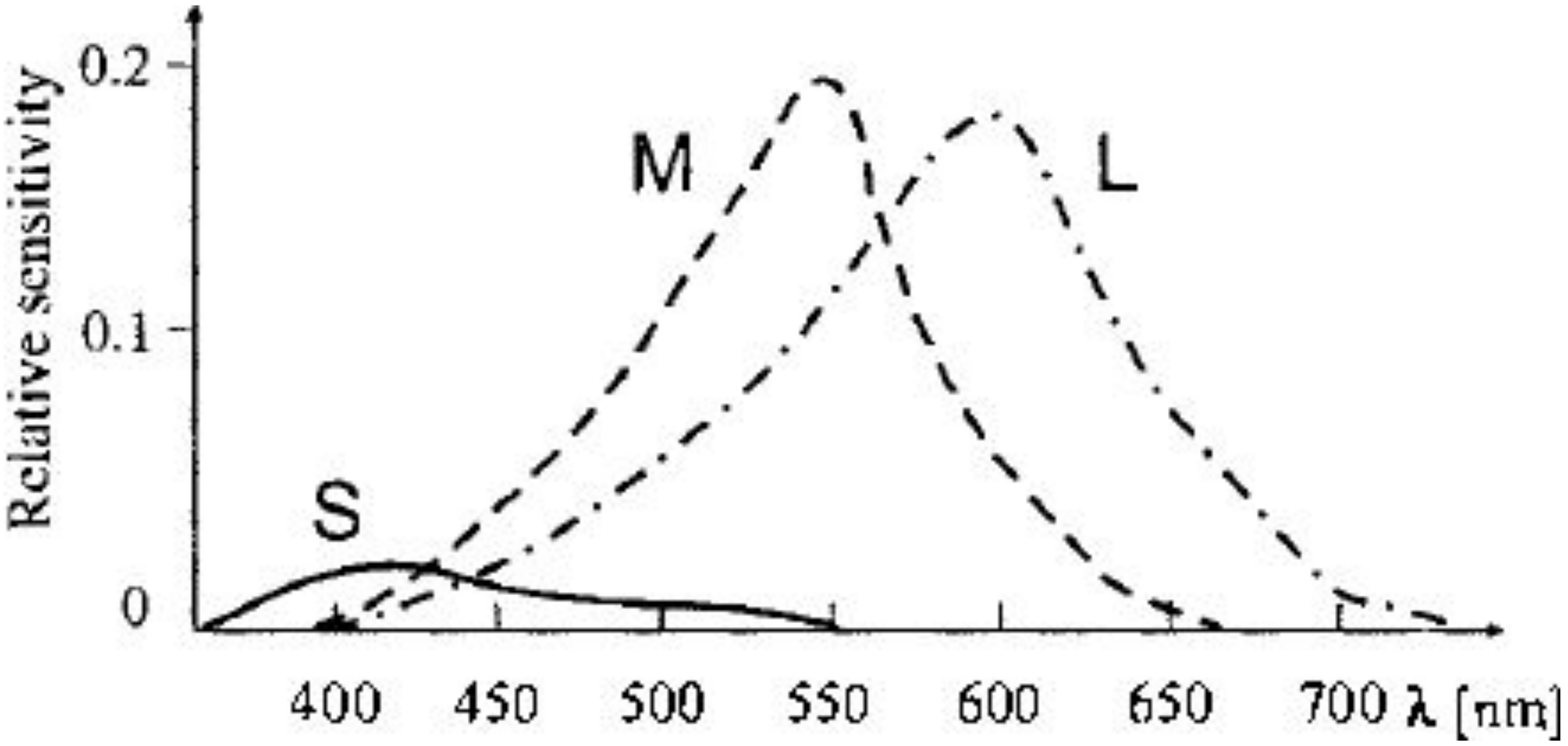


Colour perceived by humans

Human colour perception adds a subjective layer on top of underlying objective physical properties - the wavelength of electromagnetic radiation.

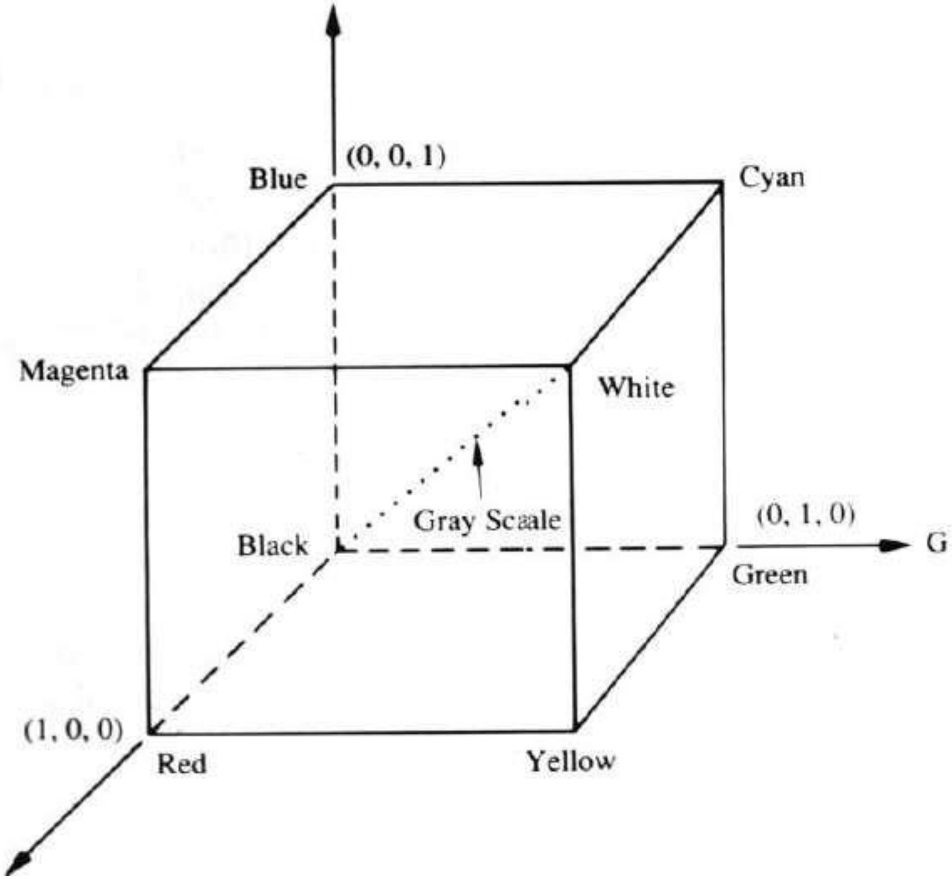
Three types of sensors receptive to the wavelength of incoming irradiation have been established in humans, thus the term trichromacy.

Relative sensitivity of S ,M, L cones of the human eye



RGB additive colour model

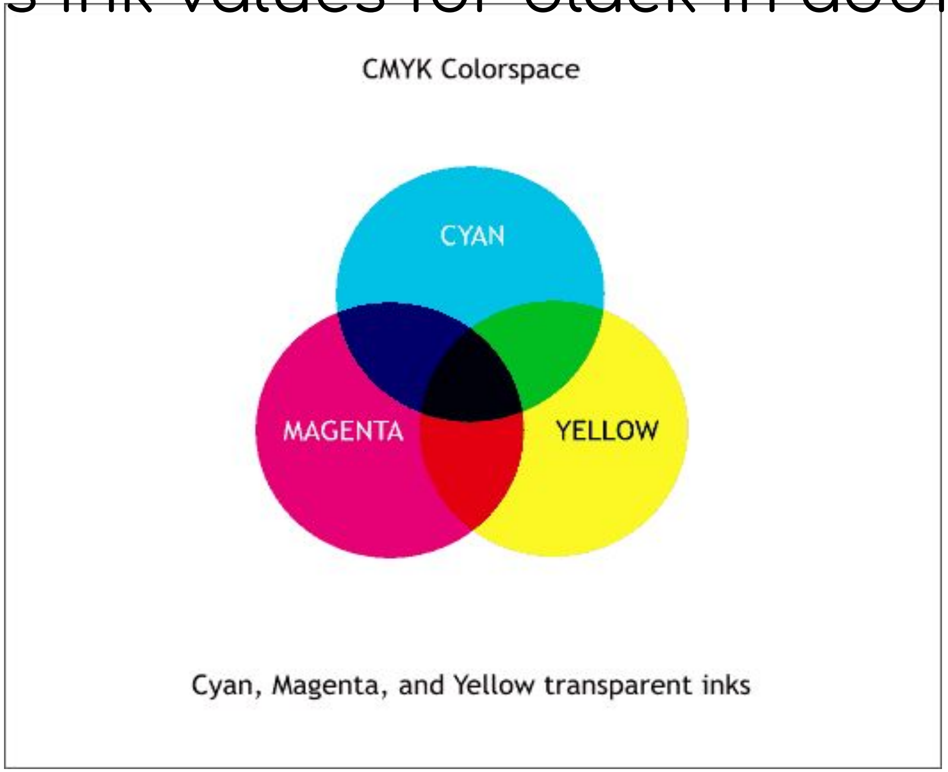
RGB color spaces



CMY colour model Cyan, Magenta, Yellow

subtractive colour mixing which is used in printing processes

CMYK stores ink values for black in addition.



Color Space – YC_bC_r

Used for: digital video encoding

Advantage:

Bandwidth efficiency

Axes:

Y: luma

C_b: blue chroma

C_r: red chroma

Color Space – YC_bC_r

Conversion from RGB:

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

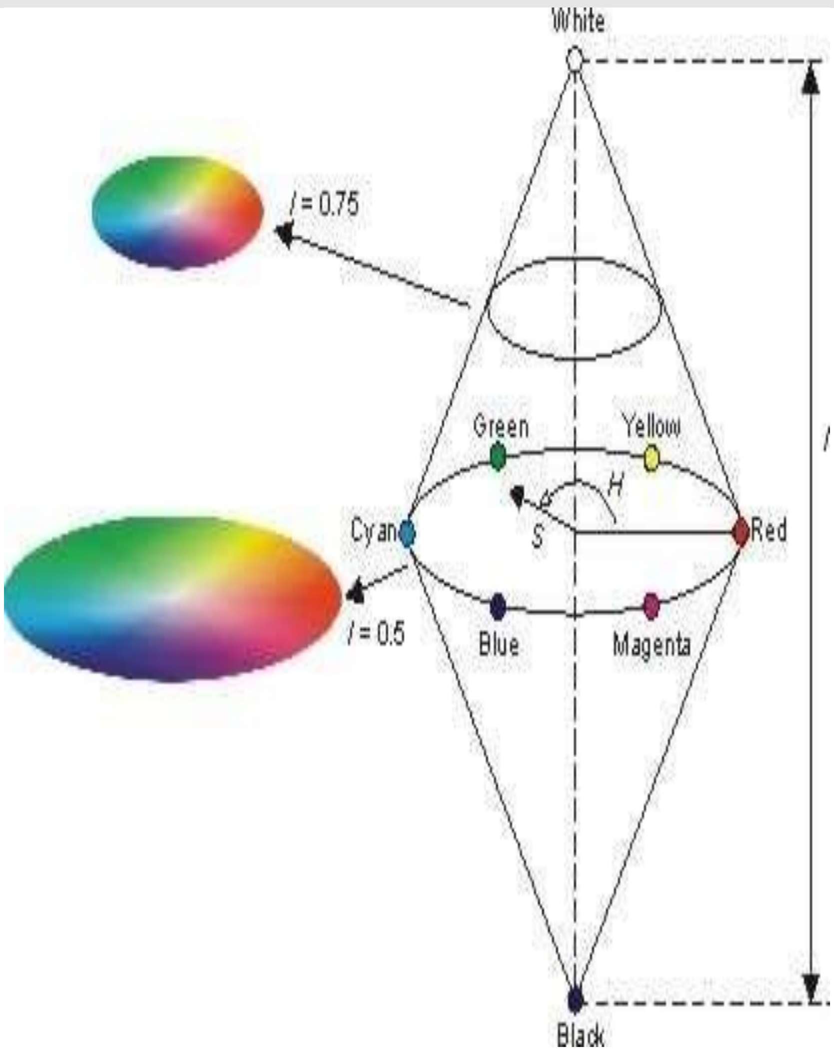
$C_b=0.564(B-Y)$

$C_r=0.713(R-Y)$

Other color spaces

- HSI,
- HSV,
- L*a*b*
- …and more

HSI color space



saturation (S)

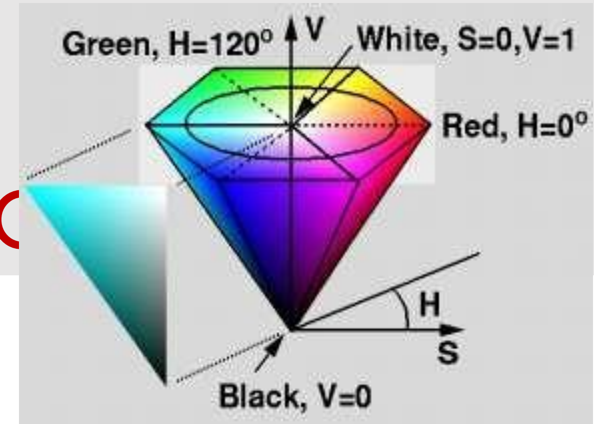
is proportional to radial distance

hue (H)

is a function of the angle in the polar coordinate system.

Intensity I = (R+G+B)/3

HSV color space
Hue, Saturation, and Va



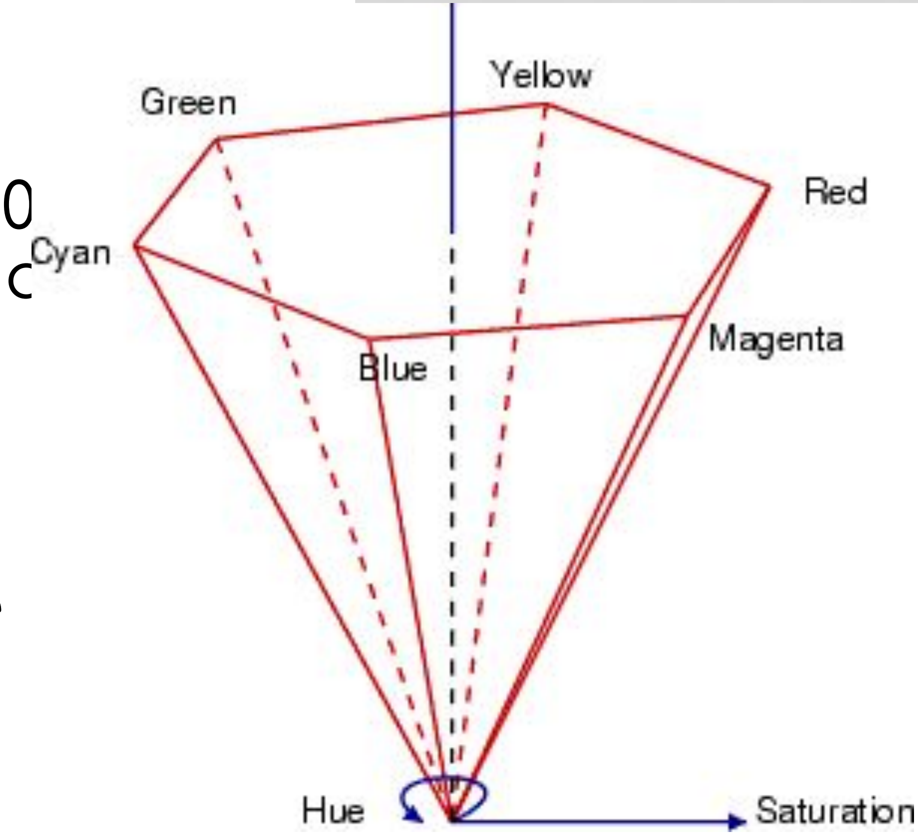
Hue,

Ranges from 0-360

Saturation, Ranges from 0-10
The lower the saturation of c
color, the more "grayness" is
present

Value, the brightness of the
color:

Ranges from 0-100



JPEG

DCT Compression Example – recap.

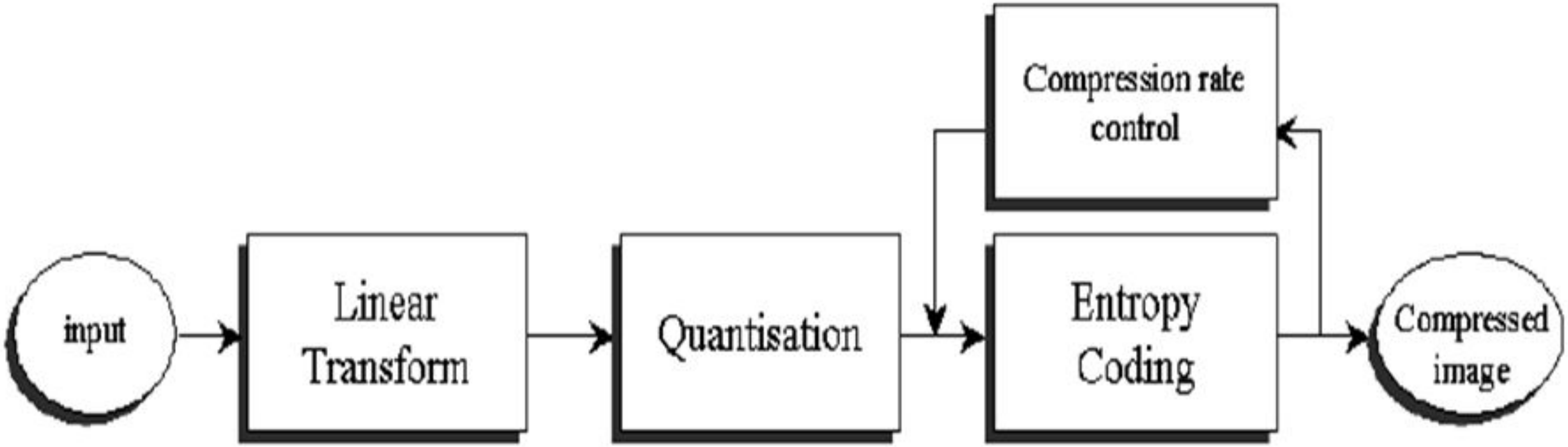
$$
\text{Mask} = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix};
$$

For typical images, many of the DCT coefficients have values close to zero; these coefficients can be discarded without seriously affecting the quality of the reconstructed image.

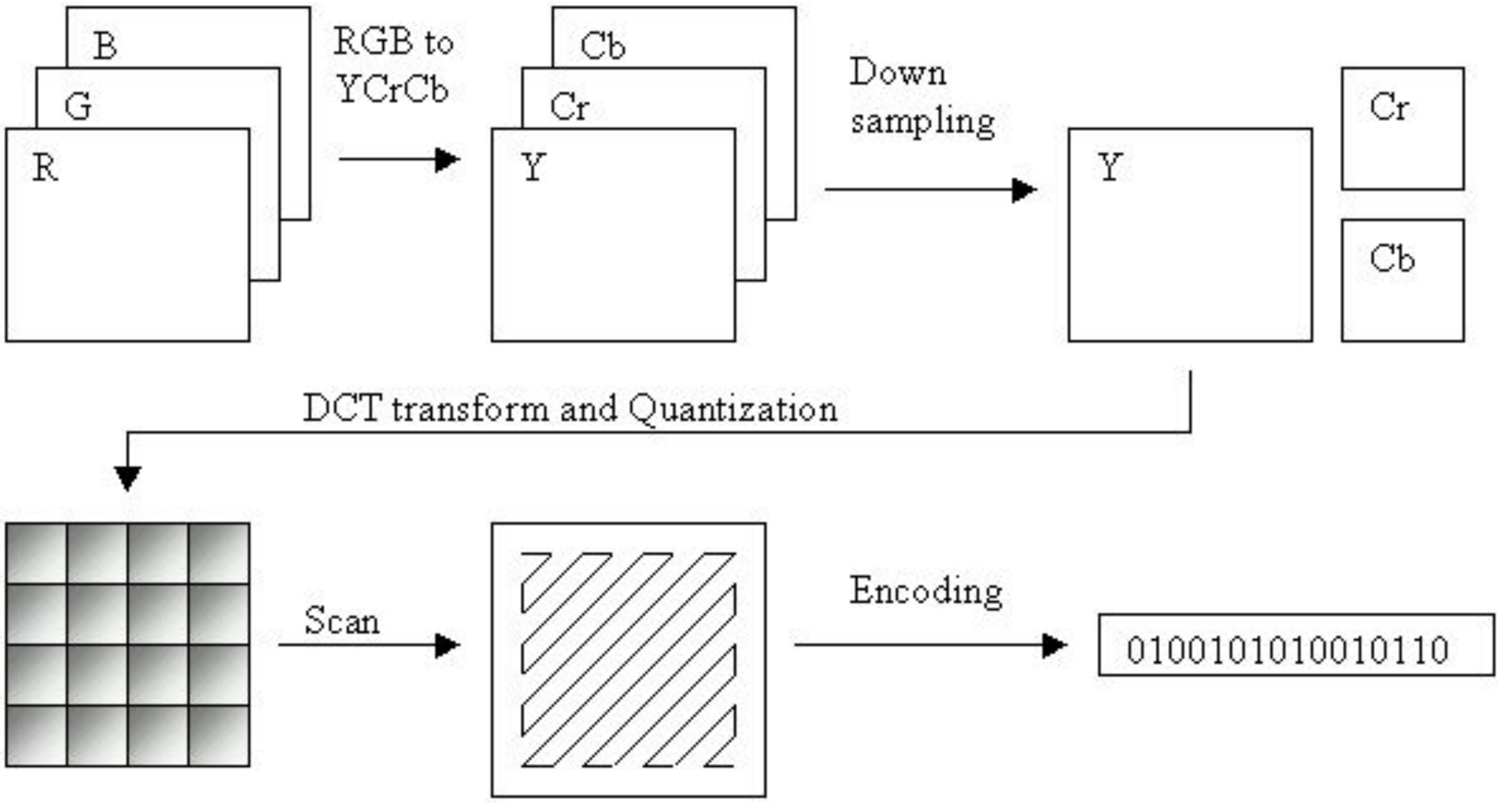
=> 85% of the DCT coefficients were discarded

Image Courtesy of MIT

Common image compression pipeline



JPEG algorithm



JPEG algorithm

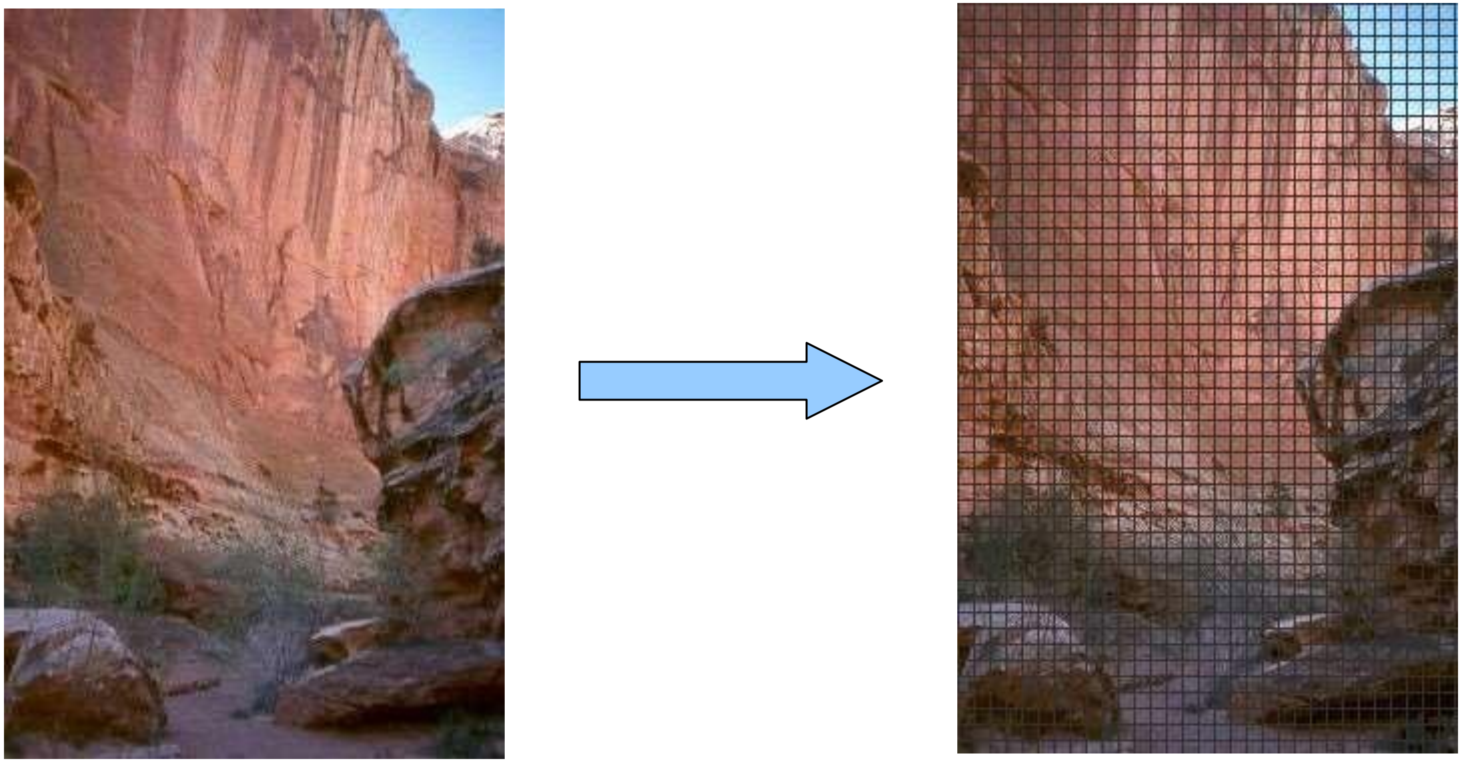
In the JPEG image compression algorithm, the input image is divided into 8-by-8 or 16-by-16 blocks, and the two- dimensional DCT is computed for each block.

The DCT coefficients are then quantized, coded, and transmitted.

The JPEG receiver (or JPEG file reader) decodes the quantized DCT coefficients, computes the inverse two-dimensional DCT of each block, and then puts the blocks back together into a single image.

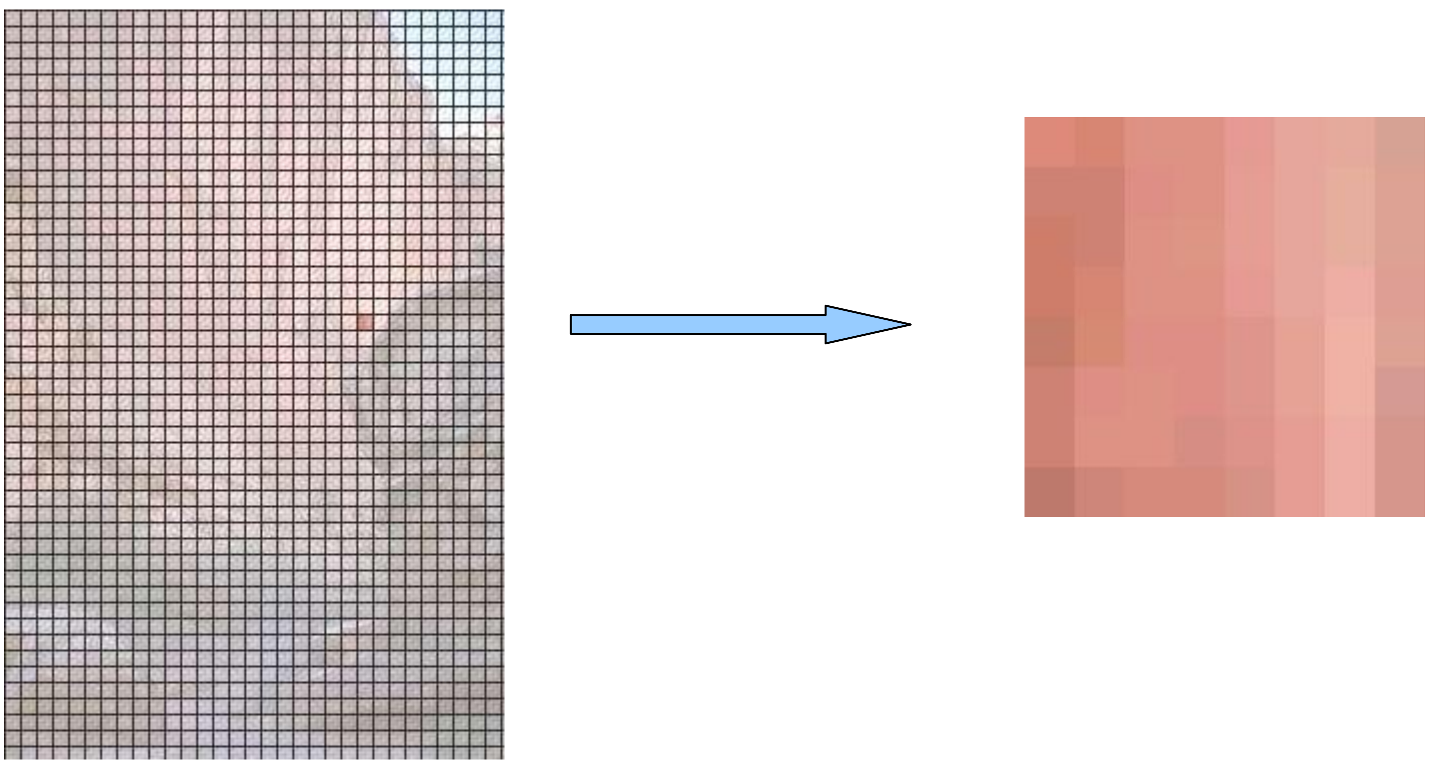
The JPEG compression algorithm

The image is subdivided into pixel blocks of size 8x8 or 16x16



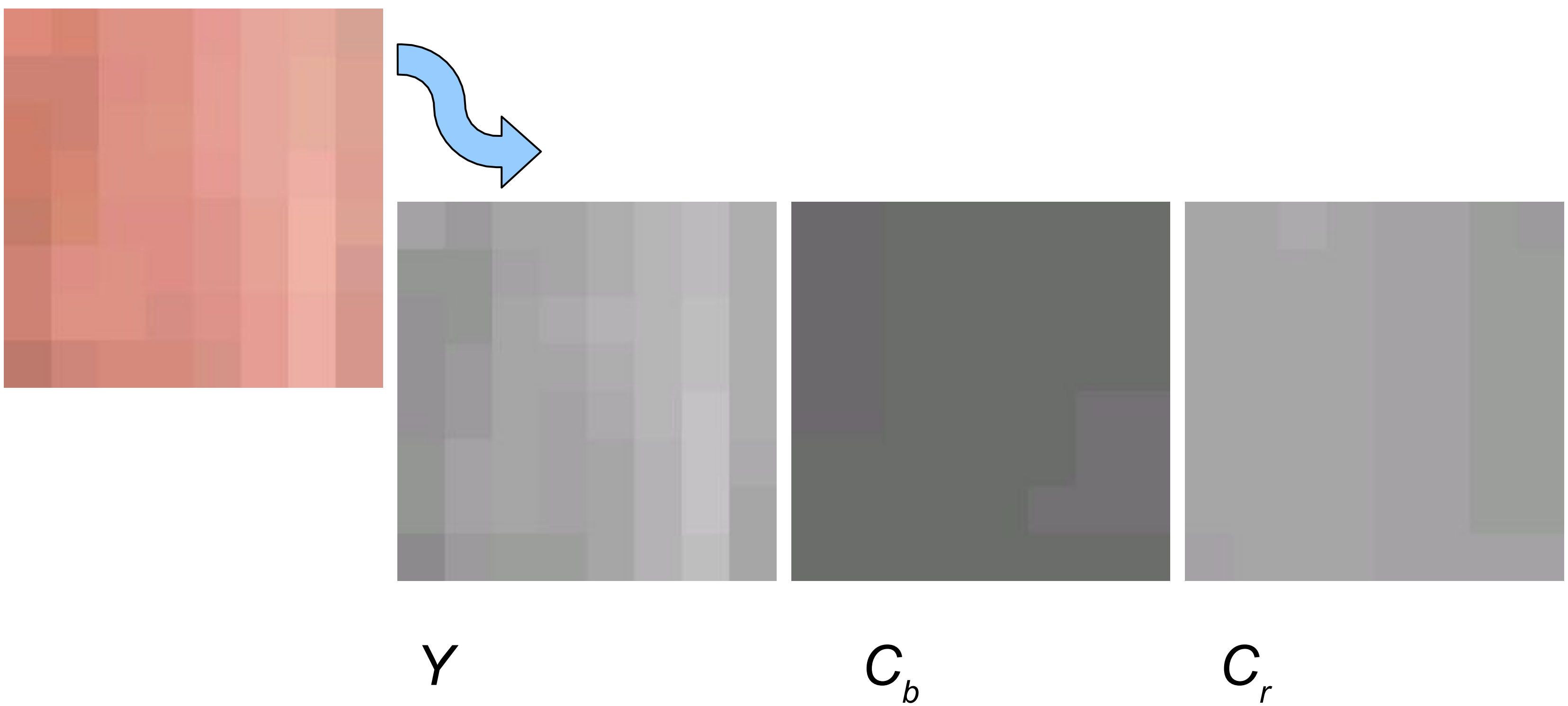
The JPEG compression algorithm

Since each block is processed without reference to the others, we'll concentrate on a single block:



R,G,B -> Y, Cb, Cr color transformation

We apply the transformation from R,G,B -> Y, Cb, Cr to each pixel in the block



Downsampling / chroma subsampling

The chrominance information can (optionally) be downsampled.

Due to the densities of color- and brightness-sensitive receptors in the human eye, humans can see considerably more fine detail in the brightness of an image (the Y' component) than in the hue and color saturation of an image (the Cb and Cr components).

The ratios are:

4:4:4 (no downsampling),

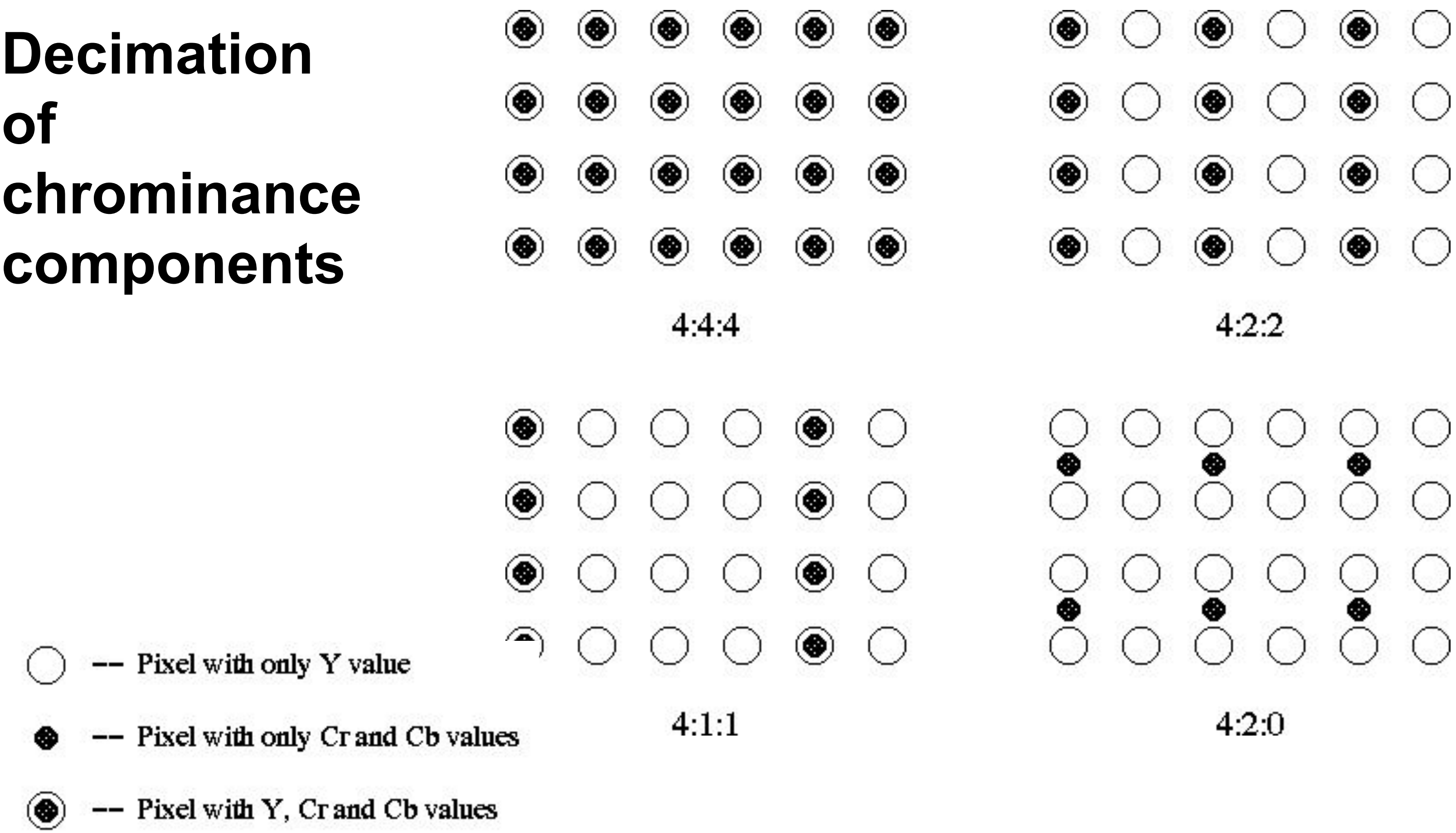
4:2:2 (reduction by a factor of 2 in the horizontal direction), or (most commonly)

4:2:0 (reduction by a factor of 2 in both the horizontal and10

vertical directions).

Downsampling / chroma subsampling

Decimation of chrominance components



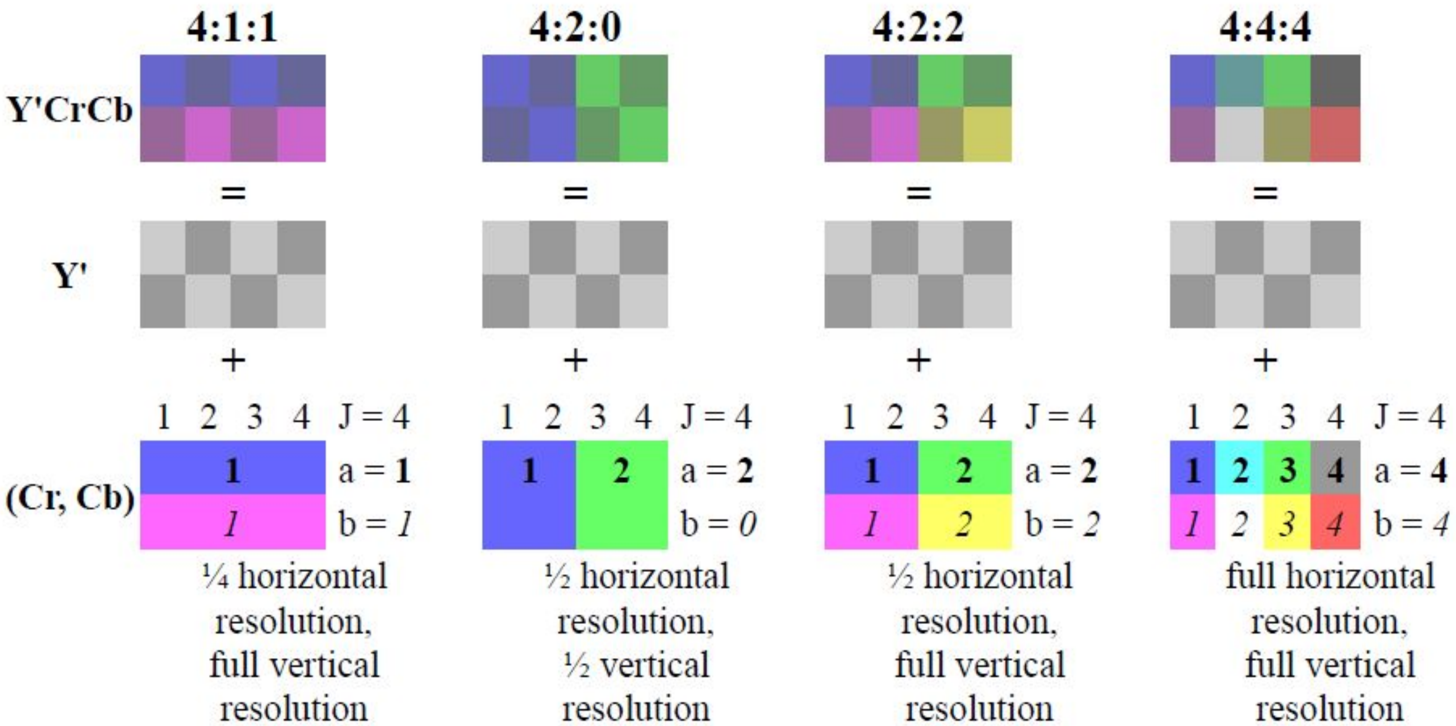
Downsampling / Chroma subsampling

Expressed as a three part ratio **J:a:b (e.g. 4:2:2),**

J: *horizontal sampling reference - usually, 4.*

a: *number of chrominance samples (Cr, Cb) in the 1st row of J pixels.*

b: *number of (additional) chrominance samples (Cr, Cb) in the 2nd row of J pixels.*



Discrete Cosine Transform - recap.

The forward transform of a block xb is given by:

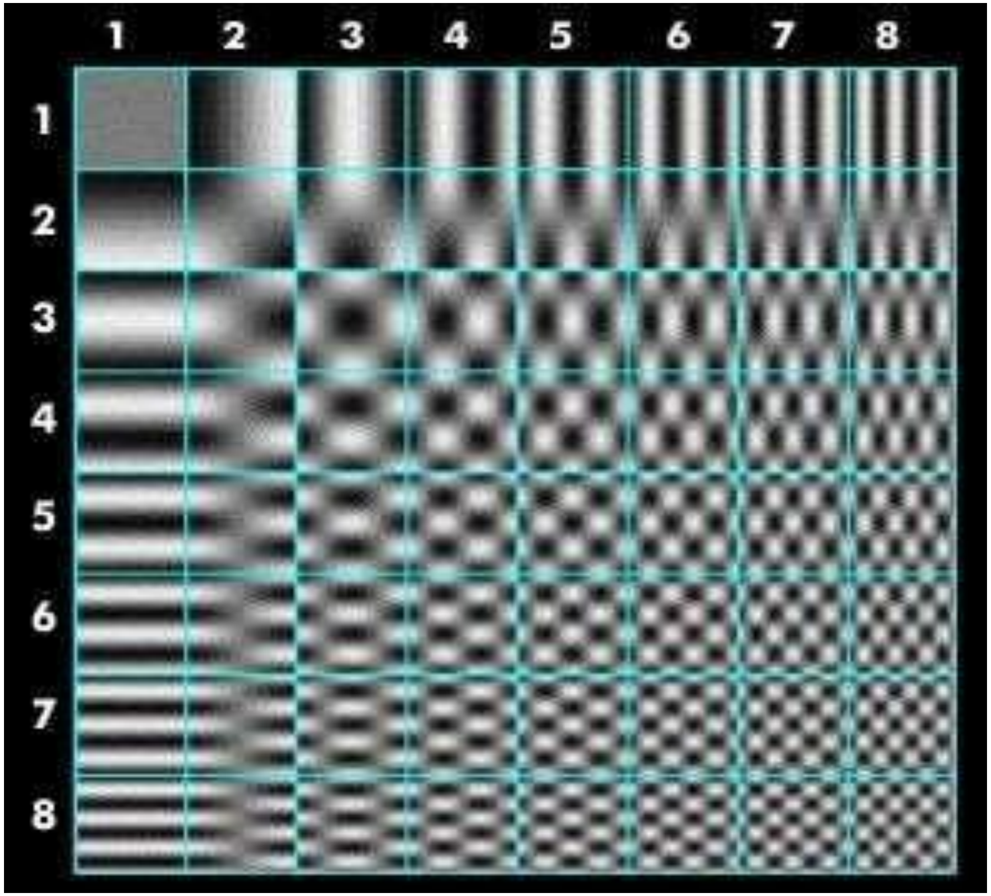
$$(\mathbf{X}_b)_{u,v} = \frac{C(u)}{\sqrt{N/2}} \frac{C(v)}{\sqrt{N/2}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (\mathbf{x}_b)_{i,j} \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N},$$

where $0 \leq u, v < 8$ and

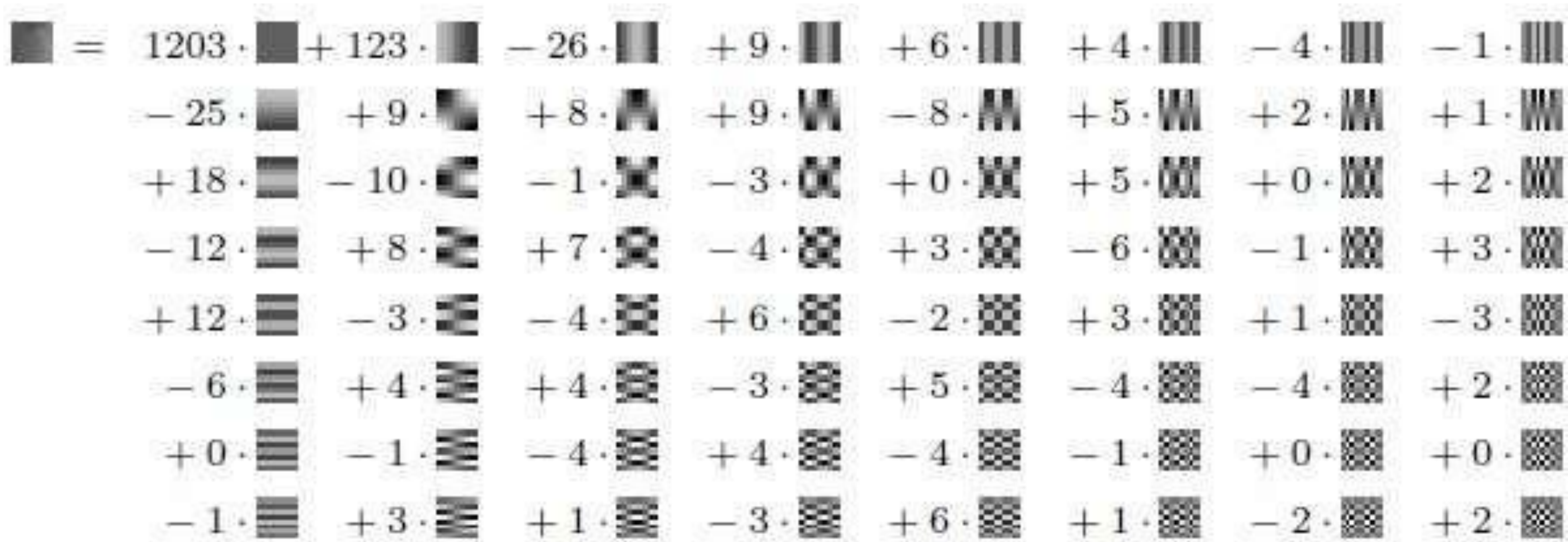
$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & u > 0 \end{cases}.$$

The Discrete Cosine Transform - recap.

Discrete-cosine basis functions for N = 8. The origin of each block is at its top left.



Example of Cosine coeff.



Quantization

The coefficients $F_{w,u}$, are real numbers, which will be stored as

integers.

Rather than simply rounding the coefficients Fw,u, we will first divide by a quantizing factor and then record round($F_{w,u}$/ $Q_{w,u}$)

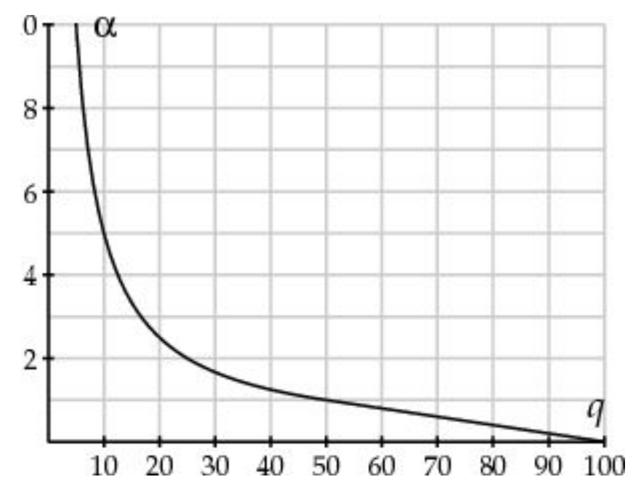
Human eye is not particularly sensitive to rapid variations in the image. This means we may deemphasize the higher frequencies, without significantly affecting the visual quality of the image, by choosing a larger quantizing factor for higher frequencies.

Quantization

When a JPEG file is created, the algorithm asks for a parameter to control the quality of the image and how much the image is compressed.

This parameter, which we'll call q, is an integer from 1 to 100.

$$\alpha = \begin{cases} \frac{50}{q} & \text{if } 1 \leq q \leq 50 \\ 2 - \frac{q}{50} & \text{if } 50 \leq q \leq 100 \end{cases}$$



$$\mathrm{round}(F_{w,u} \;/(\alpha^* Q_{w,u}))$$

Quantization

Here are typical values for $Q_{w,u}$ recommended by the JPEG standard.

for the luminance coefficients

$$Q_l = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

for the chrominance coefficients:

$$Q_c = \begin{bmatrix} 17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\ 18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\ 24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\ 47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \end{bmatrix}$$

Quantization

Quantizing with q = 50 gives the following blocks:

20	-7	-1	1	-2	1	0	0
1	0	0	0	1	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Y

-9	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

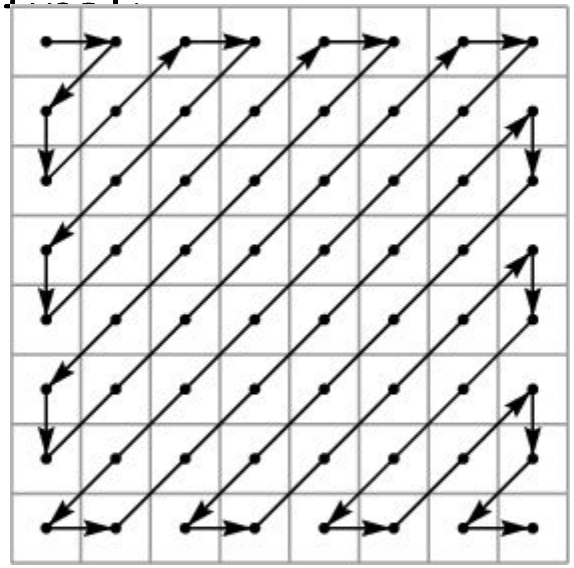
C_b

17	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

C_r

Order the coefficients : zig-zag pattern

We now order the coefficients as shown below so that the lower frequencies appear first:



In particular, for the luminance coefficients we record

20 -7 1 -1 0 -1 1 0 0 0 0 0 0 0 -2 1 1 0 0 0 0 ... 0

Entropy coding

Run-length encoding (RLE)

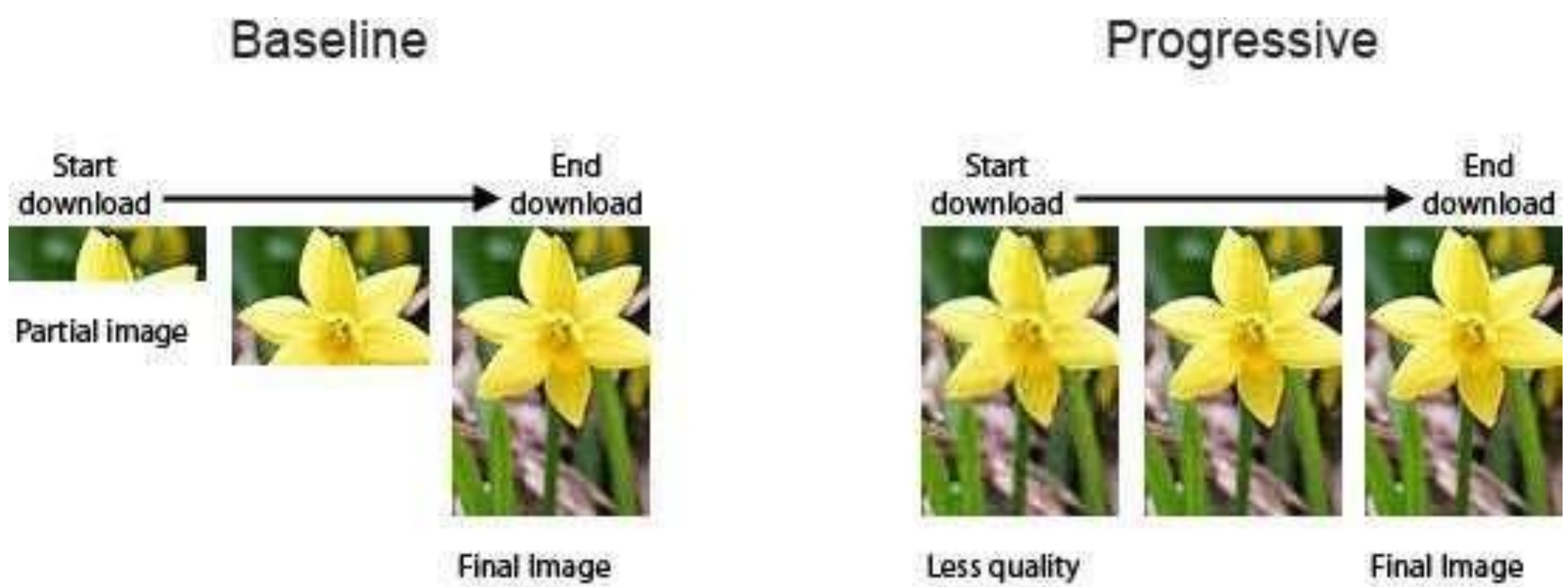
Huffman coding

Arithmetic coding

JPEG baseline and progressive

Baseline JPEG -will start to display the image as the data is made available, line by line.

Progressive JPEG - shows a blurry/low-quality photo in its entirety, and then becomes clearer as the image's data becomes more fully downloaded.



JPEG Markers (Format JPEG)

A JPEG image consists of a sequence of segments, each beginning with a marker, each of which begins with a 0xFF byte followed by a byte indicating what kind of marker it is.

Common JPEG markers[16]

Short name	Bytes	Payload	Name	Comments
SOI	0xFF, 0xD8	*none*	Start Of Image	
SOF0	0xFF, 0xC0	*variable size*	Start Of Frame (Baseline DCT)	Indicates that this is a baseline DCT-based JPEG, and specifies the width, height, number of components, and component subsampling (e.g., 4:2:0).
SOF2	0xFF, 0xC2	*variable size*	Start Of Frame (Progressive DCT)	Indicates that this is a progressive DCT-based JPEG, and specifies the width, height, number of components, and component subsampling (e.g., 4:2:0).
DHT	0xFF, 0xC4	*variable size*	Define Huffman Table(s)	Specifies one or more Huffman tables.
DQT	0xFF, 0xDB	*variable size*	Define Quantization Table(s)	Specifies one or more quantization tables.
DRI	0xFF, 0xDD	4 bytes	Define Restart Interval	Specifies the interval between RST*n* markers, in macroblocks. This marker is followed by two bytes indicating the fixed size so it can be treated like any other variable size segment.
SOS	0xFF, 0xDA	*variable size*	Start Of Scan	Begins a top-to-bottom scan of the image. In baseline DCT JPEG images, there is generally a single scan. Progressive DCT JPEG images usually contain multiple scans. This marker specifies which slice of data it will contain, and is immediately followed by entropy-coded data.
RST*n*	0xFF, 0xD*n* (*n*=0..7)	*none*	Restart	Inserted every *r* macroblocks, where *r* is the restart interval set by a DRI marker. Not used if there was no DRI marker. The low 3 bits of the marker code cycle in value from 0 to 7.
APP*n*	0xFF, 0xE*n*	*variable size*	Application-specific	For example, an Exif JPEG file uses an APP1 marker to store metadata, laid out in a structure based closely on TIFF.
COM	0xFF, 0xFE	*variable size*	Comment	Contains a text comment.
EOI	0xFF, 0xD9	*none*	End Of Image	