Information and Coding

Armando J. Pinho

Departamento de Electrónica, Telecomunicações e Informática

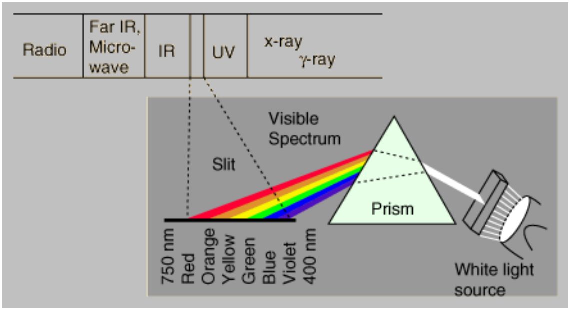
Universidade de Aveiro

ap@ua.pt

Contents

The visible spectrum

- The typical human eye senses electromagnetic wavelengths between 400 and 700 nm, and has maximum sensitivity around the 555 nm (green zone).

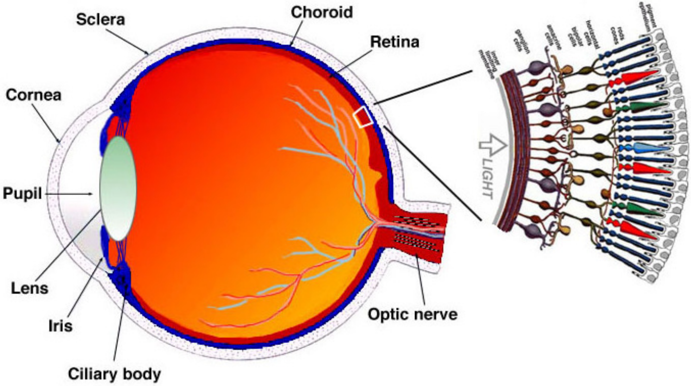


The human perception of color

- Normally, the characteristics that allow colors to be distinguished are:
  - The **brightness** (how bright is the color).
  - The **hue** (the dominant color).
  - The **saturation** (how pure is the color).
- Together, the hue and the saturation define the **chromaticity**.
- Therefore, a color can be characterized by the brightness and the chromaticity.

The human perception of color

- The human eye has **photoreceptors** that are sensitive to short wavelengths (S), medium wavelengths (M) and long wavelengths (L), also known as the blue, green and red photoreceptors.



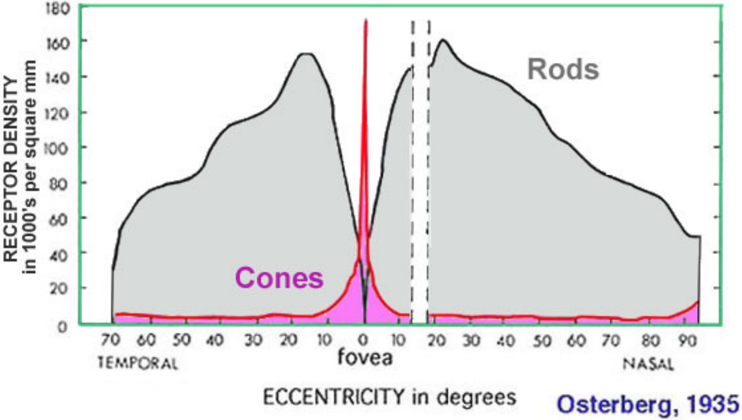
The photoreceptors: cones and rods

- The **cones**:
  - They provide the photopic vision.
  - They are between 6 and 7 million.
  - They are responsible for the perception of color.
  - There are three types:
    - Sensitive to the blue ($\approx$ 2%)
    - Sensitive to the green ($\approx$ 33%)
    - Sensitive to the red ($\approx$ 65%)
  - They are positioned mainly in the central part of the retina (fovea $\approx$ 0.3 mm diameter).

The photoreceptors: cones and rods

- The **rods**:
  - They provide the scotopic vision (under low light conditions).
  - They are between 75 and 150 million.
  - They are much more sensitive than the cones, but they are unable to distinguish colors.
  - They allow vision at low levels of light.
  - Because several rods are connected to the same nerve, they provide less spatial resolution.

Spatial distribution of the photoreceptors



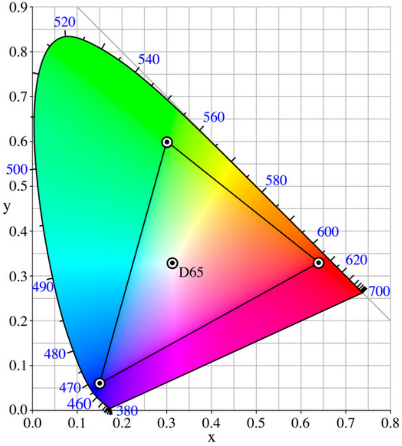
Additive primaries

- The red, green and blue are the three additive primary colors.

- Adding these three colors produces white.

The standard *RGB* (*sRGB*) color space

- Chromaticity diagram and corresponding color gamut of *sRGB* (proposed by HP and Microsoft):



The *sRGB* color space

R component

G component

B component

The *CMY* color space

- *CMY* is based on the subtractive properties of inks.
- The cyan, magenta and yellow are the **subtractive primaries**. They are the complements, respectively, of the red, green and blue. For example, the cyan subtracts the red from the white.

- Conversion from *RGB* to *CMY*: $C = 1 - R$, $M = 1 - G$, $Y = 1 - B$.

The *CMY* color space

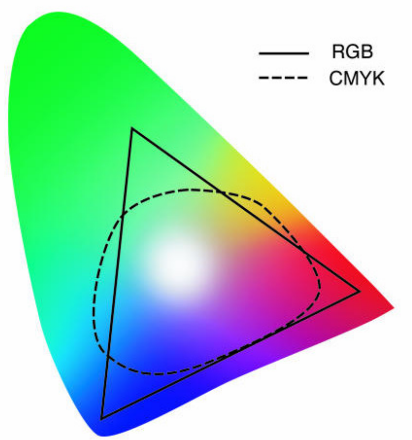
C component

M component

Y component

The *CMYK* color space

- Due to technological difficulties regarding the reproduction of black, it is generally used the *CMYK* color space for printing.



The *CMYK* color space

C component

M component

Y component

K component

The YUV and YC_bC_r color spaces

- The YUV color space is used by the PAL television standard.
- Y is the luminance component:

$$Y = 0.299R + 0.587G + 0.114B$$

- Components U and V represent the chrominance:

$$\begin{aligned} U &= -0.147R - 0.289G + 0.436B = 0.492(B - Y) \\ V &= 0.615R - 0.515G - 0.100B = 0.877(R - Y) \end{aligned}$$

- YC_bC_r is a family of color spaces, related to the YUV color space, used mainly in **digital image/video** systems.

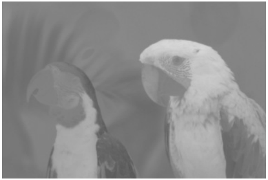
Some advantages of the *YUV* color space

- The *YUV* color space allowed to maintain the compatibility with the old “black and white” television receivers.
- The human eye is more sensitive in the green zone, which is represented mainly by the Y component (the U and V components are related to the blue and red.
- Because the human eye is less sensitive to the blue and red, it is possible to reduce the bandwidth used to represent the U and V components, without introducing significant perceptual degradation.

The YC_bC_r color space

Y component

C_b component



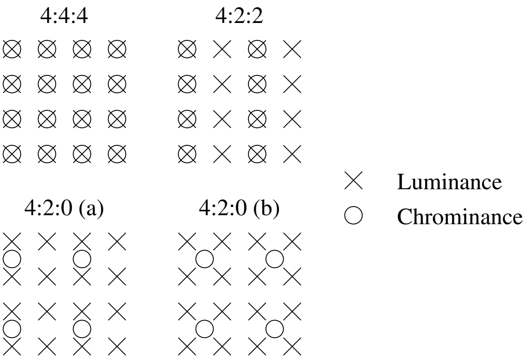
C_r component

Chrominance sub-sampling

- The YUV or YC_bC_r color spaces separate the chrominance component (UV / C_bC_r) from the luminance component (Y).
- The human eye is more sensitive to the greens, which are represented mainly by the Y component.
- For this reason, it is common to sub-sample the chrominance components UV / C_bC_r, producing a reduction in the data rate.
- This reduction is used by both the video coding standards (H.261, MPEG-1, MPEG-2, . . .) and the image coding standards (JPEG).

Chrominance sub-sampling

- The most common types of chrominance sub-sampling:



- The 4:2:0 mode has two variants: (a) used by MPEG-2; (b) used by JPEG, MPEG-1, H.261,...

Example YUV 4:2:0

RGB

YC_bC_r 4:2:0

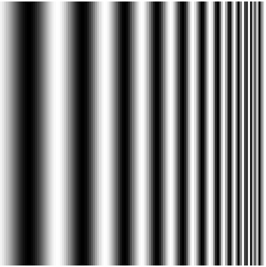
Y component

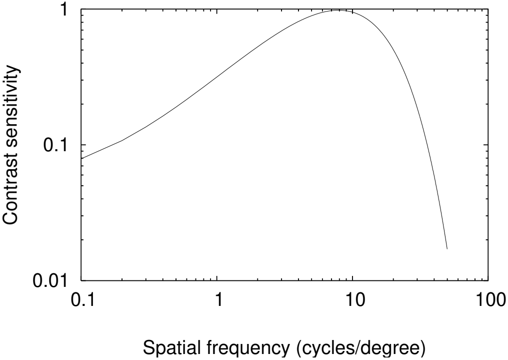
C_b component

C_r component

Spatial frequency

- The human visual system is characterized by a **bandpass** behavior in the spatial frequency domain:



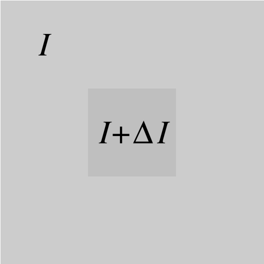


Weber's law

- Non-linear response to the light intensity (Weber's law):

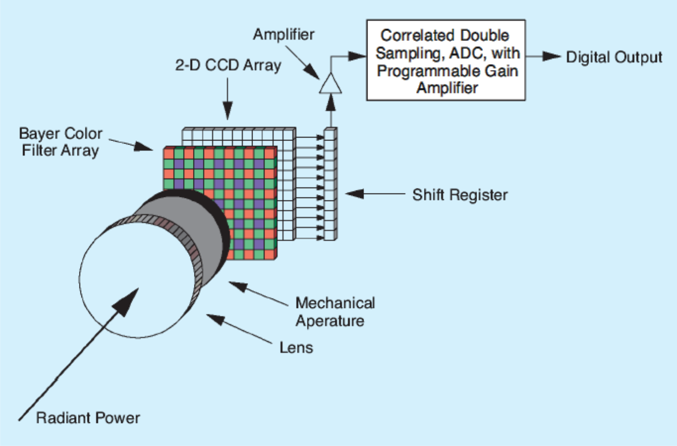
$$\frac{\Delta I}{I} \approx d(\log I) \approx \text{const.}$$

where ΔI represents the minimum intensity variation that can be perceived on a background of intensity I.

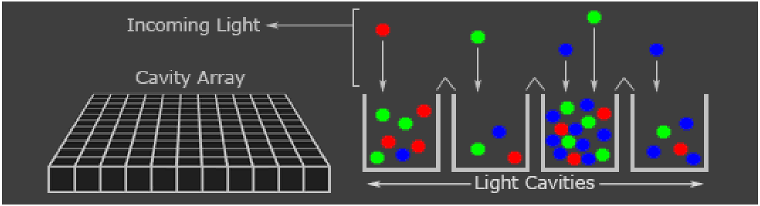


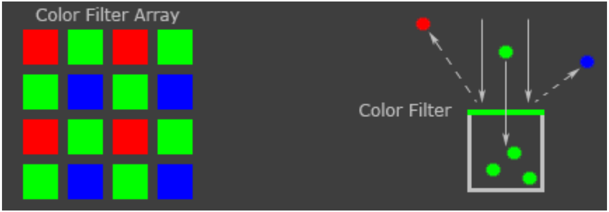
Digital camera

- Image acquisition using a digital camera:
  (IEEE SP Magazine, Jan 2005)

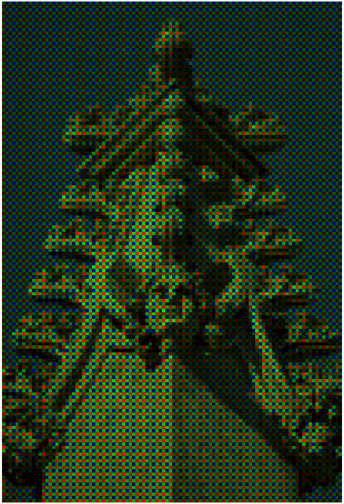


The Bayer matrix





The Bayer matrix



Contents

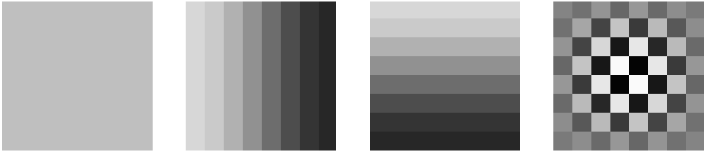
Motivation

- The main objective of using transforms in the context of data compression is to convert the original data into a new data set **more simple to quantize and encode**.
- Transforms are used **to reduce the statistical dependencies** among the original data (Ideally, the resulting coefficients should be statistically independent).
- Transforms are also used **to separate the relevant information from the irrelevant**, in order to permit course quantization or even removal of the irrelevant information.

The DCT 2D

- The DCT 2D is obtained through a separable extension of the 1D version:

$$S_{q_1,q_2} = (\mathbf{s}_{q_1,q_2})_{p_1,p_2} = c_{q_1} c_{q_2} \cos\left(2\pi f_{q_1}(p_1 + 0.5)\right) \cos\left(2\pi f_{q_2}(p_2 + 0.5)\right)$$



$S_{0,0}$ $S_{0,1}$ $S_{1,0}$ $S_{7,7}$

- For calculating a single coefficient of a non-separable transform we need m^2 operations.
- In the separable case, only $2m$ operations are required.

The DCT 2D

- Let us see how the DCT attains energy compaction, considering, for example, the following 8×8 block of pixels:

$$\mathbf{X} = \begin{bmatrix} 183 & 160 & 94 & 153 & 194 & 163 & 132 & 165 \\ 183 & 153 & 116 & 176 & 187 & 166 & 130 & 169 \\ 179 & 168 & 171 & 182 & 179 & 170 & 131 & 167 \\ 177 & 177 & 179 & 177 & 179 & 165 & 131 & 167 \\ 178 & 178 & 179 & 176 & 182 & 164 & 130 & 171 \\ 179 & 180 & 180 & 179 & 183 & 169 & 132 & 169 \\ 179 & 179 & 180 & 182 & 183 & 170 & 129 & 173 \\ 180 & 179 & 181 & 179 & 181 & 170 & 130 & 169 \end{bmatrix}$$

The DCT 2D

- The coefficients (rounded to the integers), resulting from applying the DCT to the block (after subtracting $2^7 = 128$ to each pixel), are:

$$\mathbf{Y} = \begin{bmatrix} 313 & 56 & -27 & 18 & 78 & -60 & 27 & -27 \\ -38 & -27 & 13 & 44 & 32 & -1 & -24 & -10 \\ -20 & -17 & 10 & 33 & 21 & -6 & -16 & -9 \\ -10 & -8 & 9 & 17 & 9 & -10 & -13 & 1 \\ -6 & 1 & 6 & 4 & -3 & -7 & -5 & 5 \\ 2 & 3 & 0 & -3 & -7 & -4 & 0 & 3 \\ 4 & 4 & -1 & -2 & -9 & 0 & 2 & 4 \\ 3 & 1 & 0 & -4 & -2 & -1 & 3 & 1 \end{bmatrix}$$

The JPEG standard

- The JPEG (Joint Photographic Experts Group) standard is a family of coding methods for images of continuous tones of grays or colors.
- The group was established in 1986, the standard was proposed in 1992 and approved in 1994 (ISO 10918-1).
- The JPEG standard comprises **four coding methods**: sequential, progressive, hierarchical and lossless.
- The JPEG standard is based on a number of compression techniques, such as the DCT, statistical coding and predictive coding.

The sequential mode of JPEG

- Every codec should include this mode in order to be considered JPEG-compatible (it is also known as the “baseline” mode).
- The sequential mode of JPEG comprises the following steps:
  - Calculation of the DCT.
  - Quantization of the DCT coefficients, in order to eliminate less relevant information, according to the characteristics of the human visual system.
  - Statistical coding (Huffman or arithmetic) of the quantized DCT coefficients.

The sequential mode of JPEG

- Calculation of the DCT:
  - The image is partitioned into 8×8 blocks of pixels. If the number of rows or columns is not multiple of 8, then they are internally adjusted (using padding).
  - Subtract 2^{b-1} to each pixel value, where b is the number of bits used to represent the pixels.
  - Calculate the DCT 2D of each block.

The sequential mode of JPEG

- Quantization of the DCT coefficients:
  - The DCT coefficients are quantized using a quantization matrix, Q, previously scaled by a compression quality factor.
  - Next, the coefficients are organized in a one-dimensional vector according to a zig-zag scan.
- Statistical coding:
  - The non-zero AC coefficients are encoded using Huffman or arithmetic coding, representing the value of the coefficient, as well as the number of zeros preceding it.
  - The DC coefficient of each block is predictively encoded in relation to the DC coefficient of the previous block.

Quantization of the coefficients

- Compression is obtained due to the **low-pass characteristic of the human visual system**.
- Because of this characteristic, generally more bits are assigned to the low frequencies (those appearing in the upper left corner of the transformed block).
- This is done using **threshold coding** (non-linear approximation).

Quantization of the coefficients

- Generally, the elements of Q are 8 bit integers that determine the quantization step according to the position of each coefficient.
- **Example:** quantization matrix of JPEG (luminance):

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 129 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Quantization of the coefficients

- By applying this quantization matrix to the block that we have previously used as example, we obtain the matrix $\widetilde{Y}$:

$$\widetilde{Y} = \begin{bmatrix} 20 & 5 & -3 & 1 & 3 & -2 & 1 & 0 \\ -3 & -2 & 1 & 2 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- In this example, 45 of the 64 coefficients are eliminated.

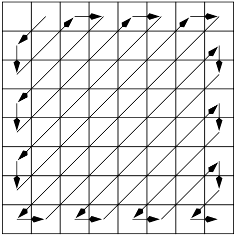
Quantization of the coefficients

- Because the sensitivity of the human eye to the colors is different from that of the luminance, JPEG provides a different quantization matrix for the chrominance components:

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

Coefficient coding

- JPEG uses a zig-zag scanning of $\widetilde{Y}$ in order to encode the quantized coefficients, except for the $(0,0)$ position, i.e., the DC coefficient.



- The objective of this scanning is to group together the zero coefficients, allowing a more efficient representation.
- This efficiency is obtained using a variant of run-length coding.

Coefficient coding

- Using again the same example, a JPEG encoder would generate the following codewords:

$$(0,5),(0,-3),(0,-1),(0,-2),(0,-3),(0,1),(0,1),(0,-1),(0,-1),$$
$$(2,1),(0,2),(0,3),(0,-2),(0,1),(0,1),(6,1),(0,1),(1,1),\mathrm{EOB}$$

$$\begin{bmatrix} 20 & 5 & -3 & 1 & 3 & -2 & 1 & 0 \\ -3 & -2 & 1 & 2 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The visual effect of the coding blocks

- The coding techniques that are based on a partition of the image into blocks are generally affected by a visual phenomenon known as the **blocking artifact**.

This artifact is more visible when the compression ratio is high and happens because the blocks are encoded independently (except for the DC coefficients).

Example: 8×8 DCT, 0.31 bpp.