浙江大学计算机学院
数字媒体与网络技术

Digital Asset Management
数字媒体资源管理

2. Introduction to Digital Media Format

任课老师：张宏鑫

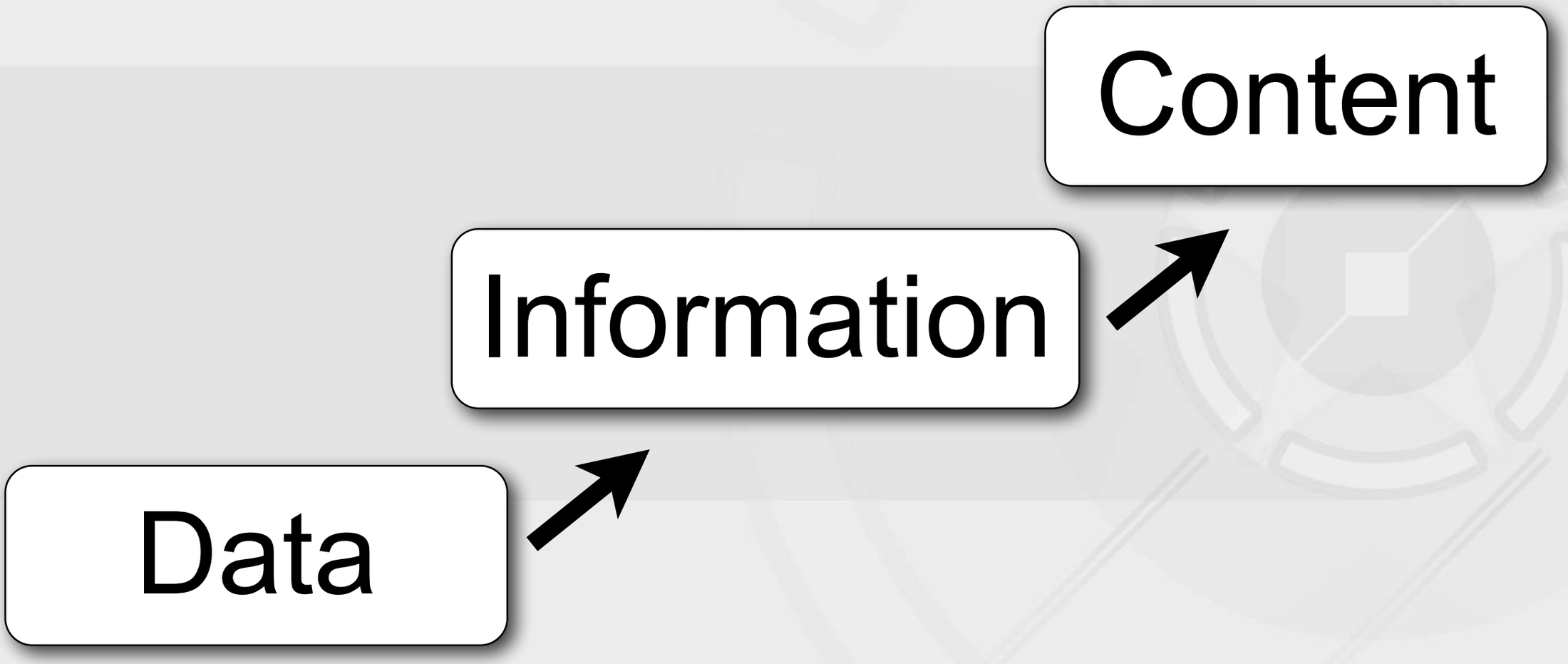
2020-09-17

基于Python的教学

- 推荐使用的工具链：
  - ipython
  - ipython notebook
  - ImageMagick
  - Flask …

Review

Content



What is Digital Asset Management?

- Tools for organizing, storing and retrieving content in digital format
  - downloading, renaming, backing up, rating, grouping, archiving, optimizing, maintaining, thinning, and exporting ...

- Includes:
  - text, video, images, movies, sound, and 3D content

Content

content = essence + metadata

内容＝素材＋元数据

Digital media data types

Table. File format used in Macromedia Director

File import					**File export**		**Native**
Image	Palette	Sound	Video	Animation	Image	Video	
BMP GIF JPG PICT PNG PNT PSD TGA TIFF WMF	PAL ACT	AIFF AU MP3 WAV	AVI MOV	DIR FLA FLC FLI GIF PPT	BMP	AVI MOV	DIR DXR EXE

Outline

- Image format and coding methods
- Audio format and coding methods
- Video format and coding methods
- Introduction to HTML and XML
- Graphics format and coding methods

Key points

- To grasp features of different types of digital media
- To understand principles of coding different types of digital media

浙江大学计算机学院
数字媒体与网络技术

2.1 Image format and coding methods

Common image formats

- General types:
  - –GIF
  - –JPEG
  - –PNG
  - –TIFF
  - –TGA
- Raw data:
  - –RAW
  - –DNG

- Platform spec.:

➡BMP (Win)

➡PAINT&PICT (Mac)

➡PPM (X-Win)

- Vector data:

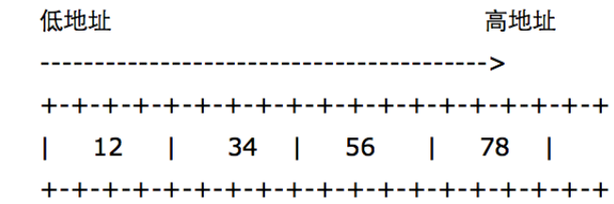
➡WMF (Win)

➡PS and PDF

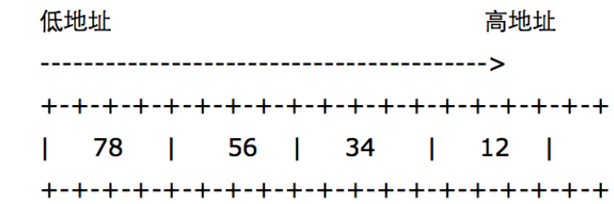
Common image formats

- Key points of storage
  - –Color space
  - –Coding methods
  - –Byte order： hardware dependent
    - MSB/LSB (most/least significant byte)
    - MSB - Power PC CPU
    - LSB - Intel X86 CPU

Big Endian 0x12345678



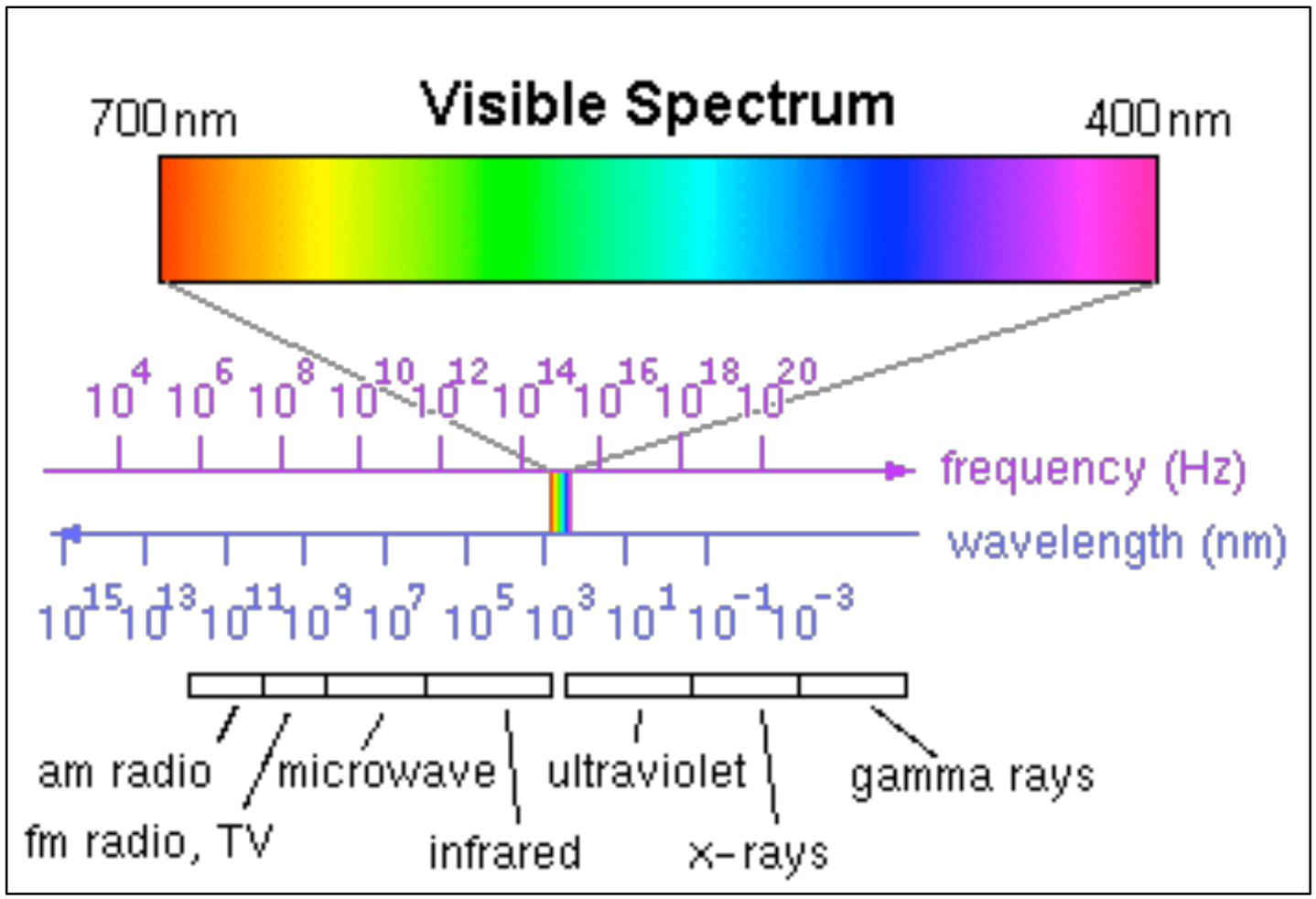
Little Endian



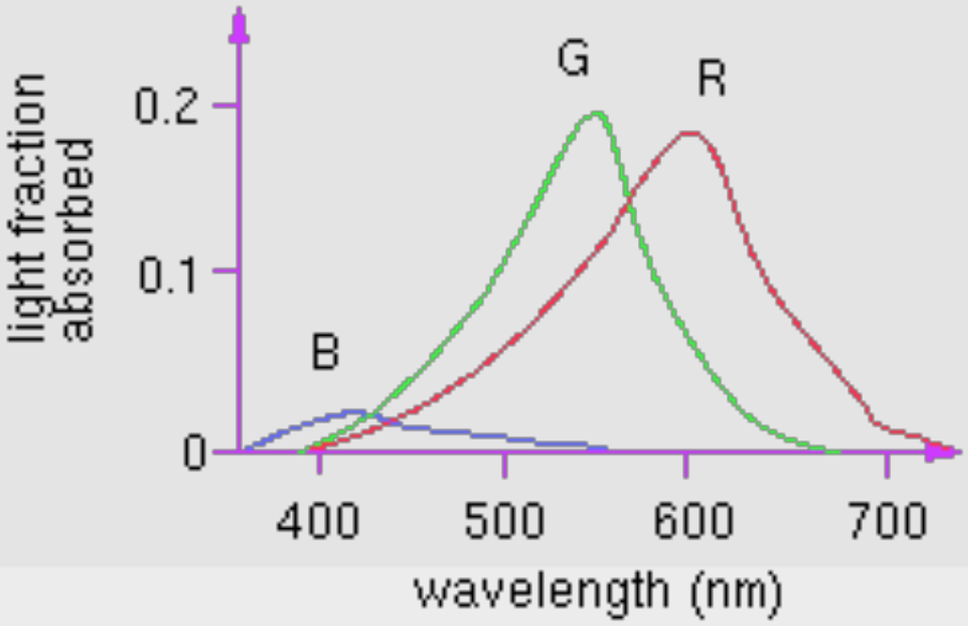
2.1.1 Color spaces

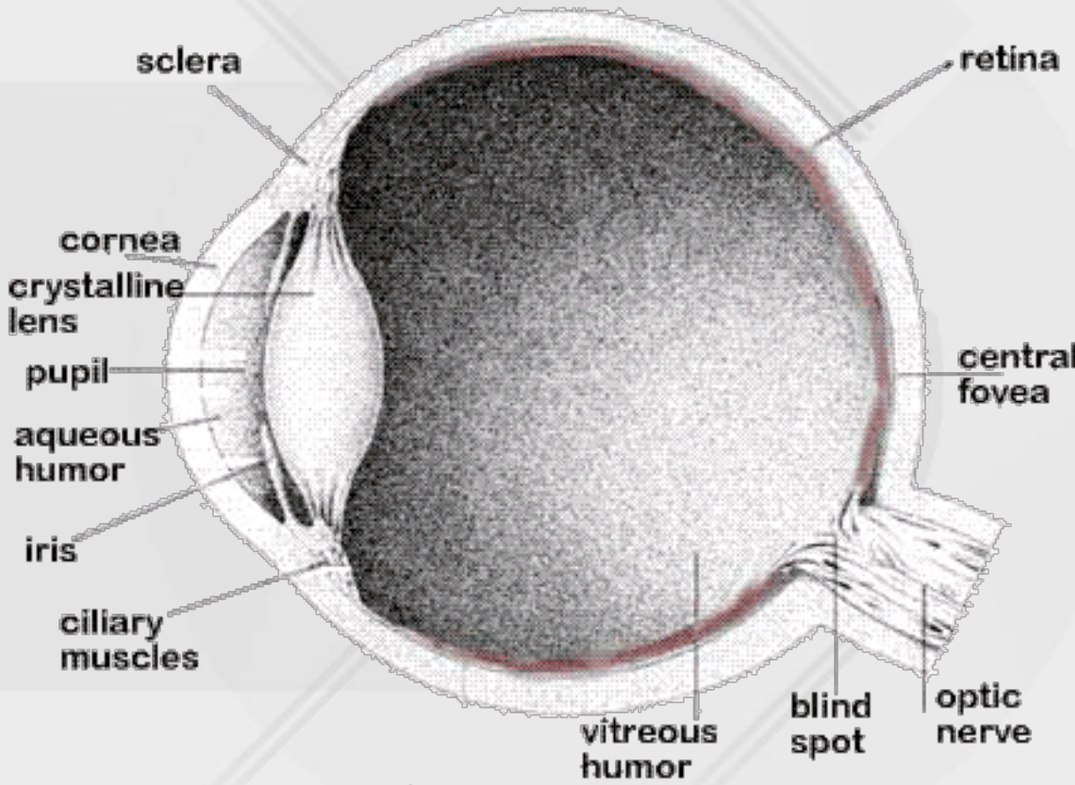
Color systems and color models

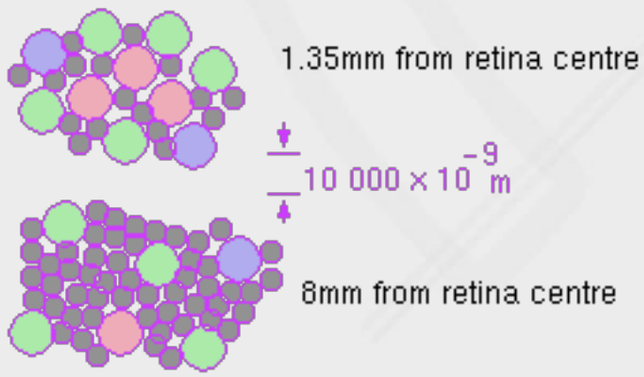
- Common systems
  - –RGB
  - –XYZ
  - –Lab
  - –YUV
  - –HSV



Color Model


light fraction absorbed
0.2
0.1
0
G
R
B
400
500
600
700
wavelength (nm)


sclera
retina
cornea
crystalline lens
pupil
aqueous humor
iris
ciliary muscles
central fovea
vitreous humor
blind spot
optic nerve


1.35mm from retina centre
$10\,000 \times 10^{-9}$ m
8mm from retina centre

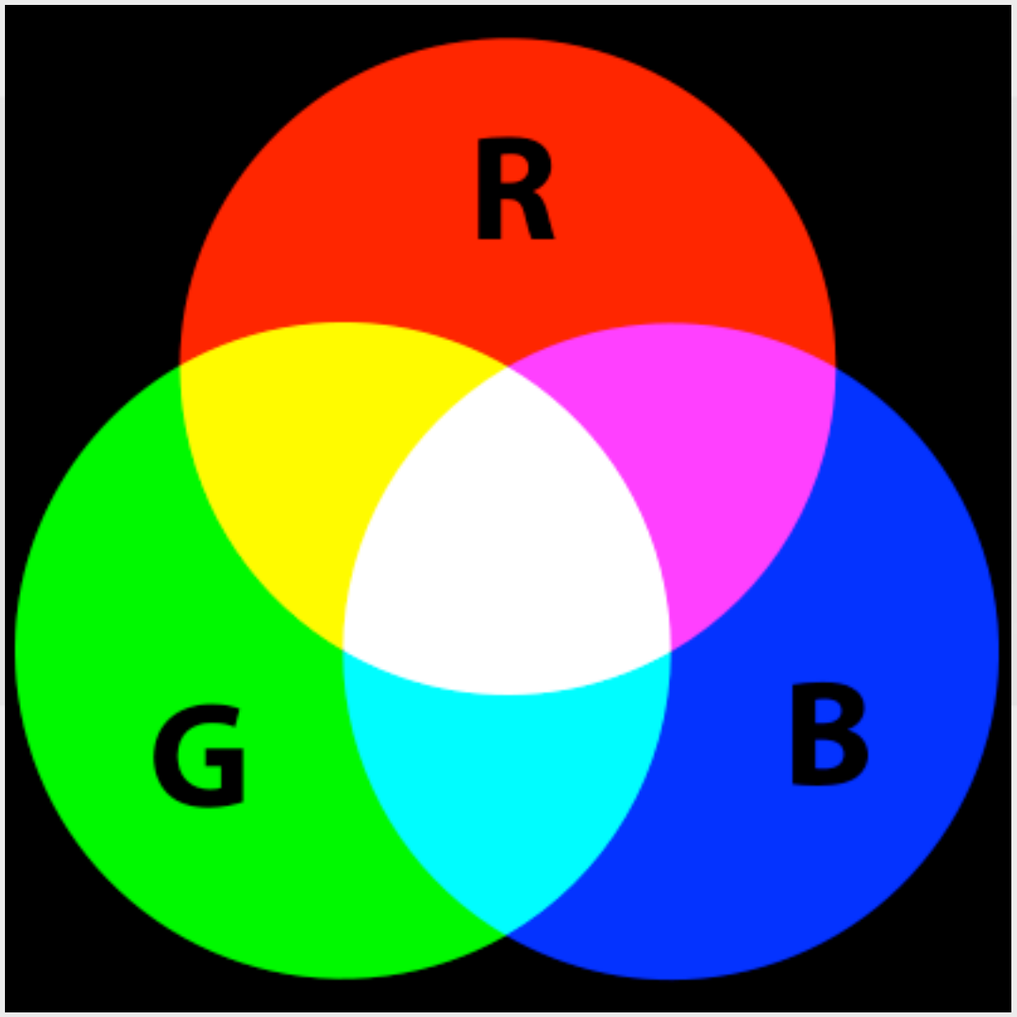
Color perception

- Three types of cones:

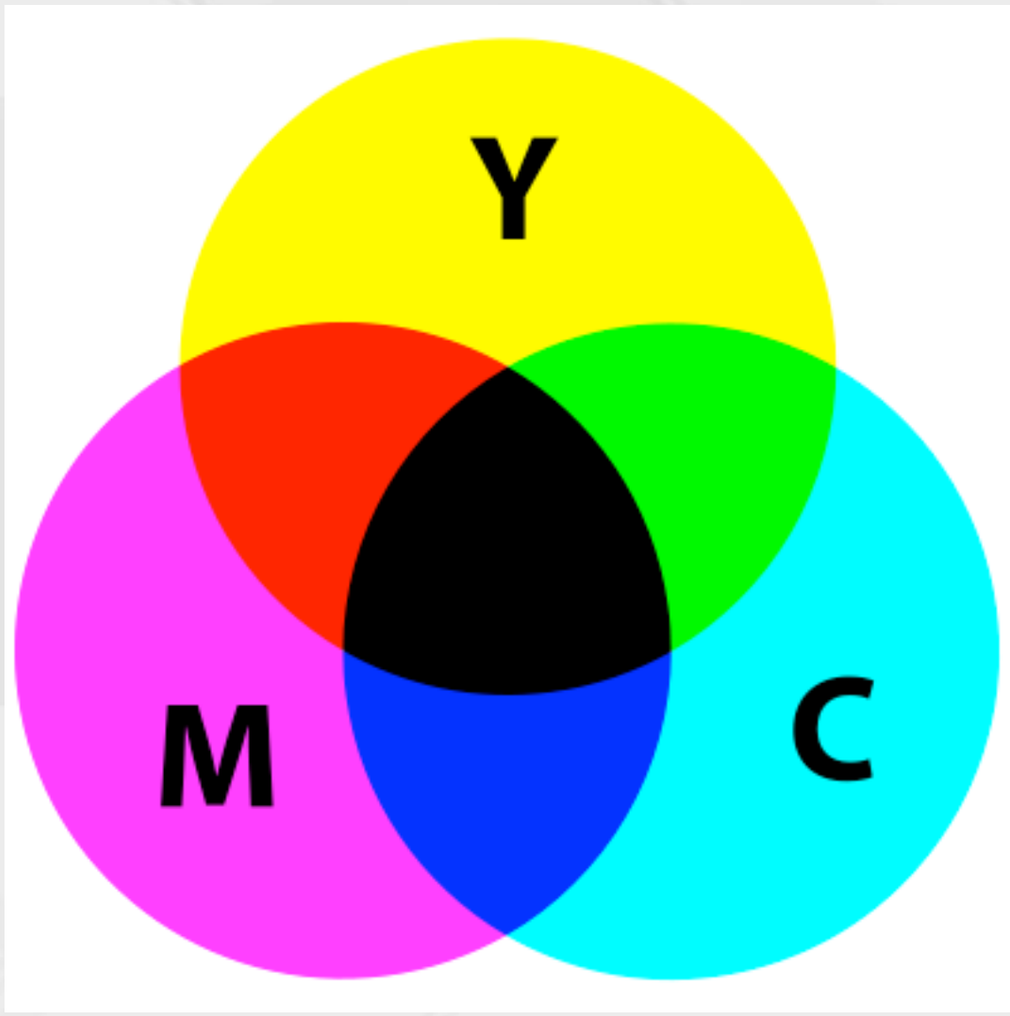
S	M	L	
Blue	Green	Red	roughly approximate
430nm	560nm	610nm	peak sensitivities

- Colorblindness results from a deficiency of one cone type.

RGB & CMYK



Additive color mixing



Subtractive color mixing

CMYK => RGB

$$t_{CMYK} = C, \mathrm{M}, \Upsilon, \mathrm{K}$$

RGB => CMYK?

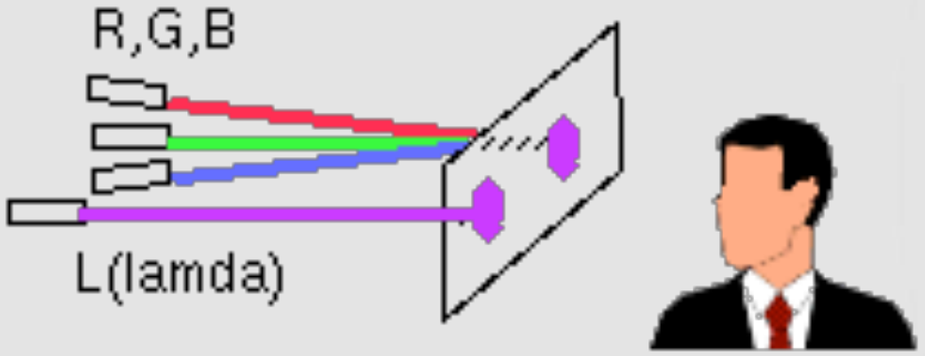
$$t_{CMY} = C', \mathrm{M}', \Upsilon'$$
$$= C(1-\mathrm{K})+\mathrm{K}, \mathrm{M}(1-\mathrm{K})+\mathrm{K}, \Upsilon(1-\mathrm{K})+\mathrm{K}$$

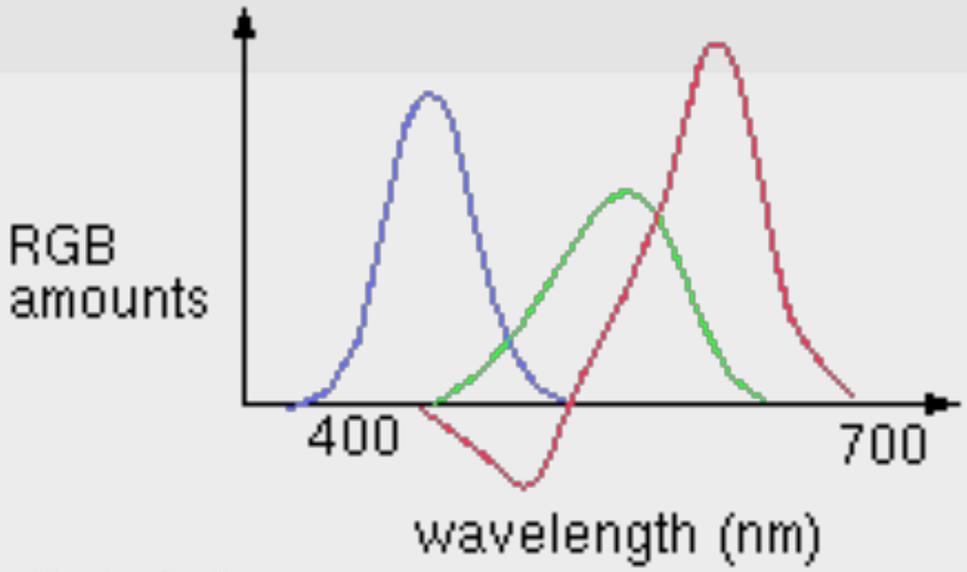
$$t_{RGB} = R, G, \mathrm{B}$$
$$= 1-C', 1-\mathrm{M}', 1-\Upsilon'$$

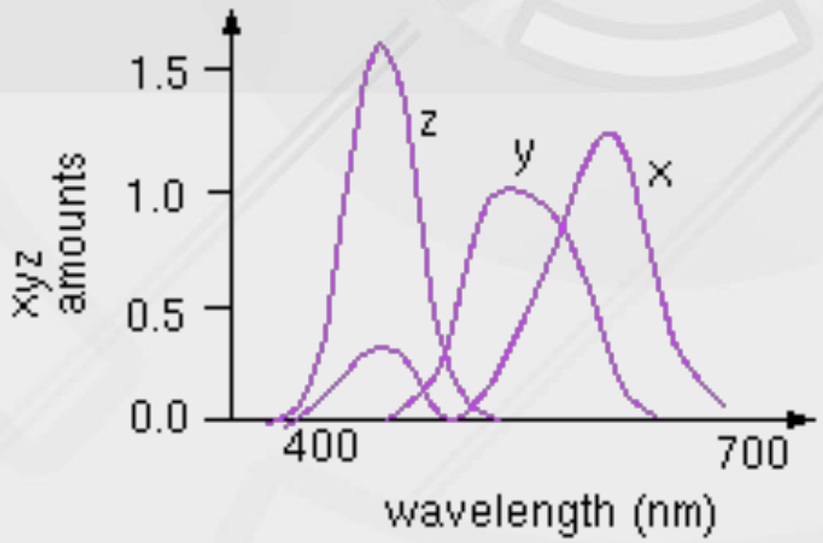
$$t_{RGB} = \{1-(C(1-\mathrm{K})+\mathrm{K}), 1-(\mathrm{M}(1-\mathrm{K})+\mathrm{K}), 1-(\Upsilon(1-\mathrm{K})+\mathrm{K})\}$$
$$= \{1-C(1-\mathrm{K})-\mathrm{K}, 1-\mathrm{M}(1-\mathrm{K})-\mathrm{K}, 1-\Upsilon(1-\mathrm{K})-\mathrm{K}\}$$

CIE XYZ space

- CIE: Commission Internationale d'Eclairage"







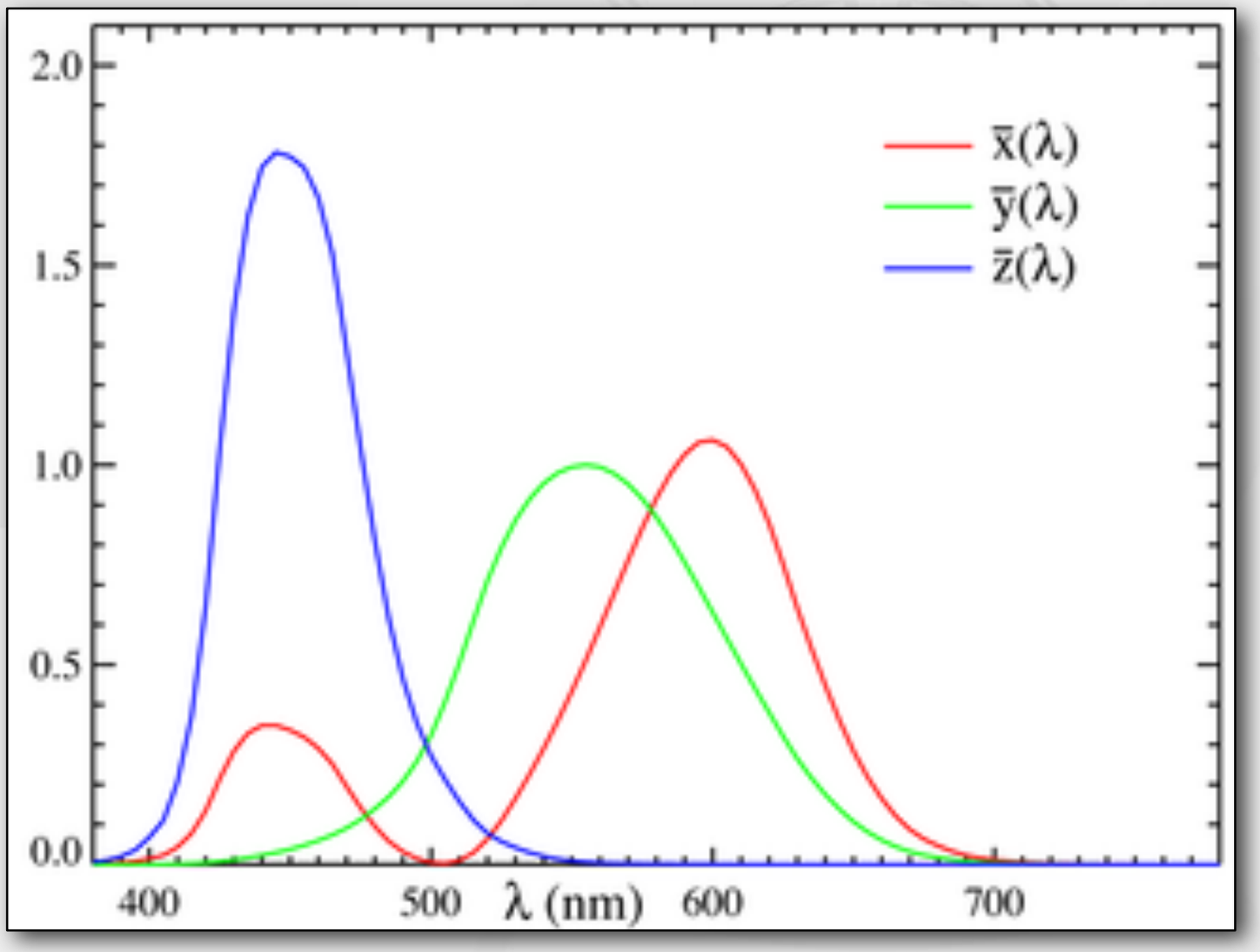
CIE XYZ space

- Color matching function

$$X = \int_0^\infty I(\lambda)\,\overline{x}(\lambda)\,d\lambda$$

$$Y = \int_0^\infty I(\lambda)\,\overline{y}(\lambda)\,d\lambda$$

$$Z = \int_0^\infty I(\lambda)\,\overline{z}(\lambda)\,d\lambda$$

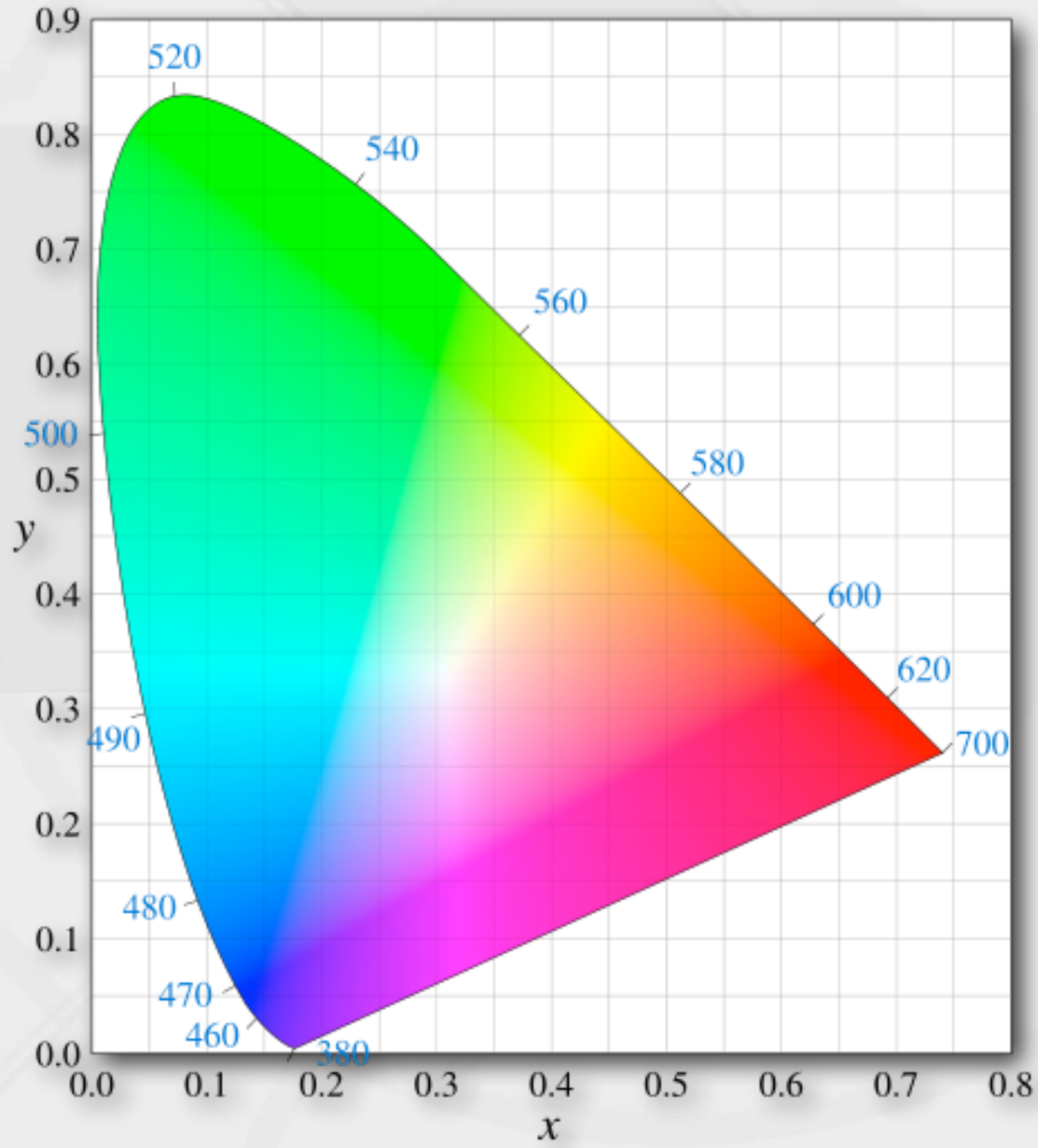


CIE XYZ space

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$

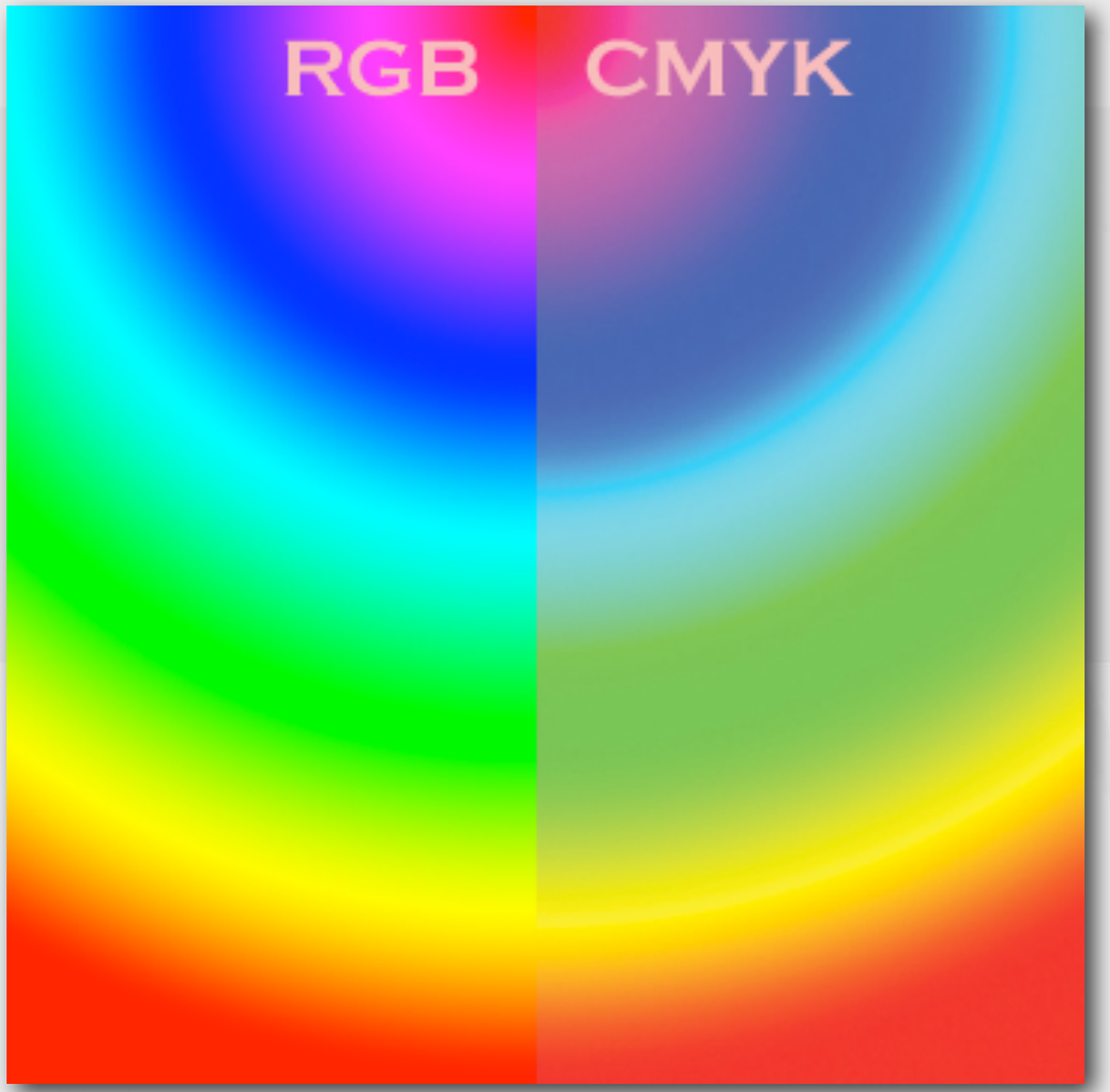


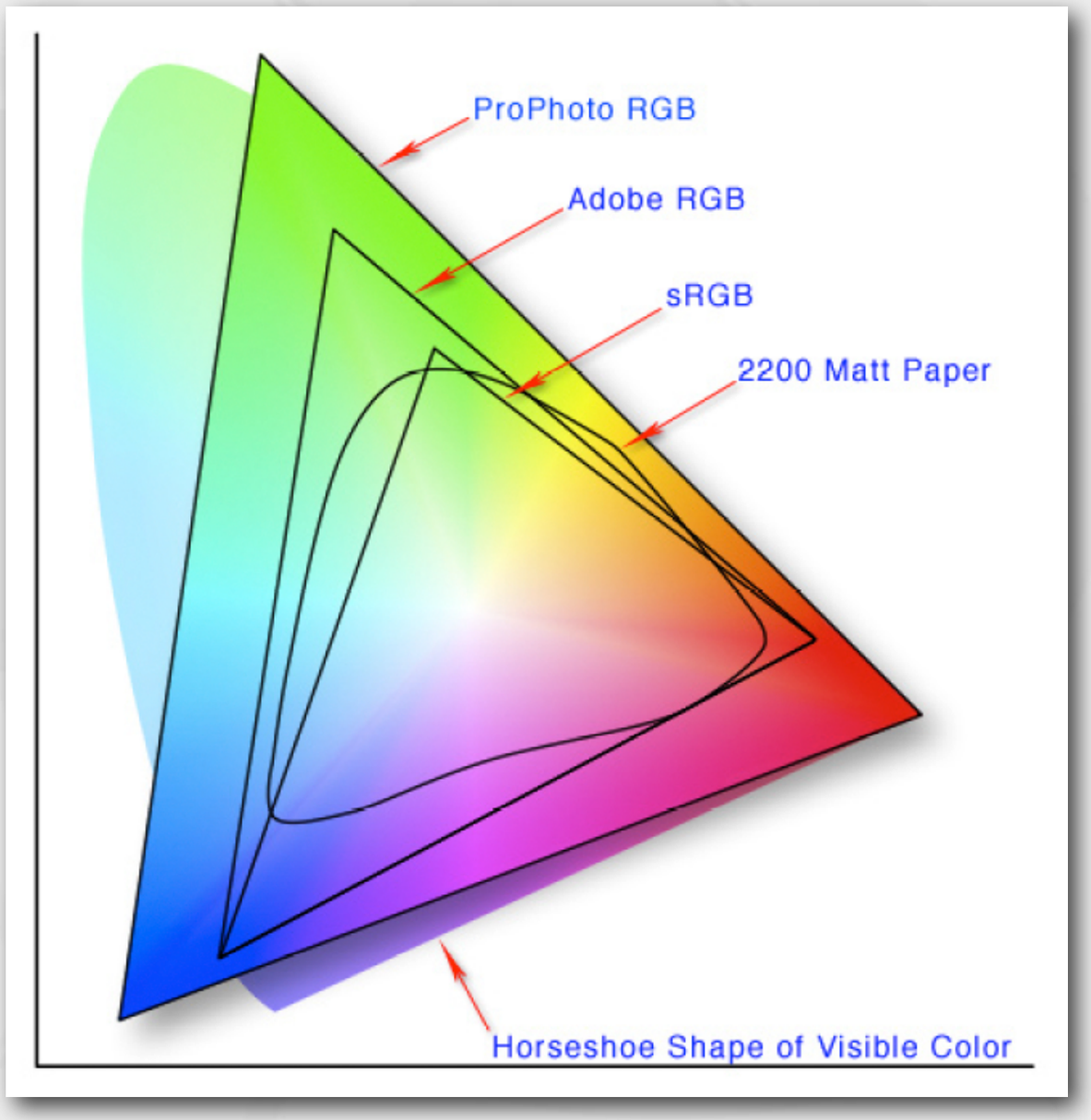
RGB vs. XYZ

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{b_{21}} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Since 1931

RGB :: CMYK :: XYZ color spaces

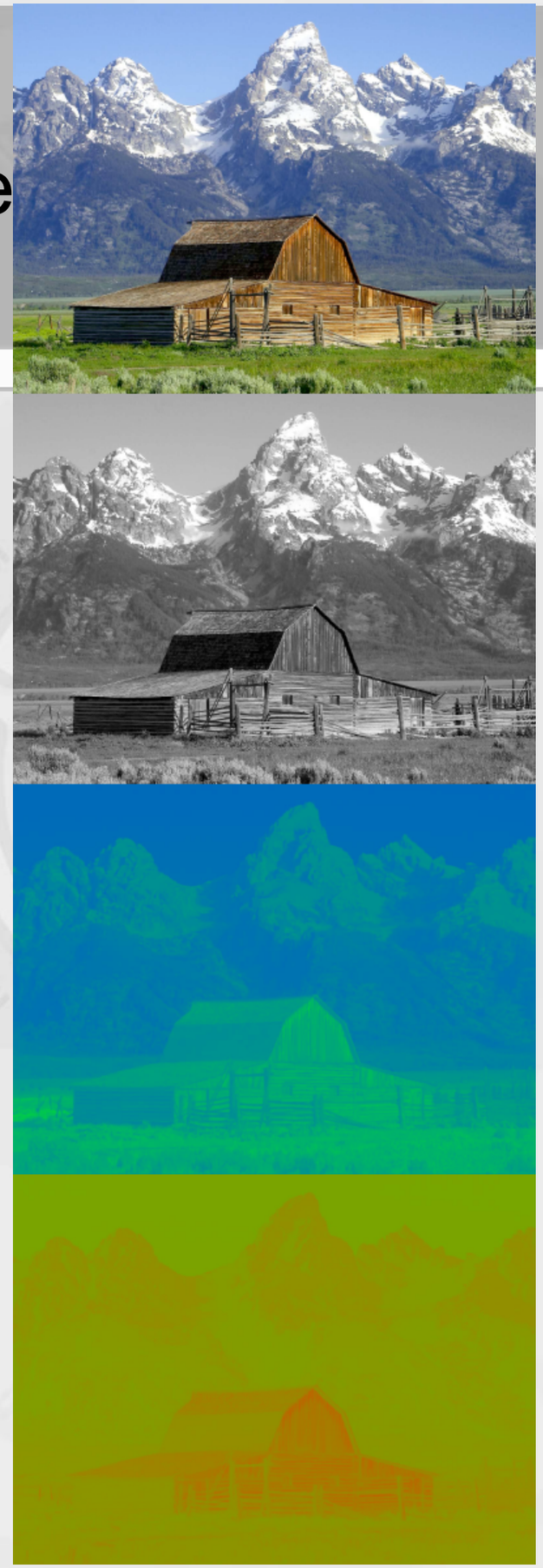




YUV color spaces

- used in most video capture system
- PAL television system

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



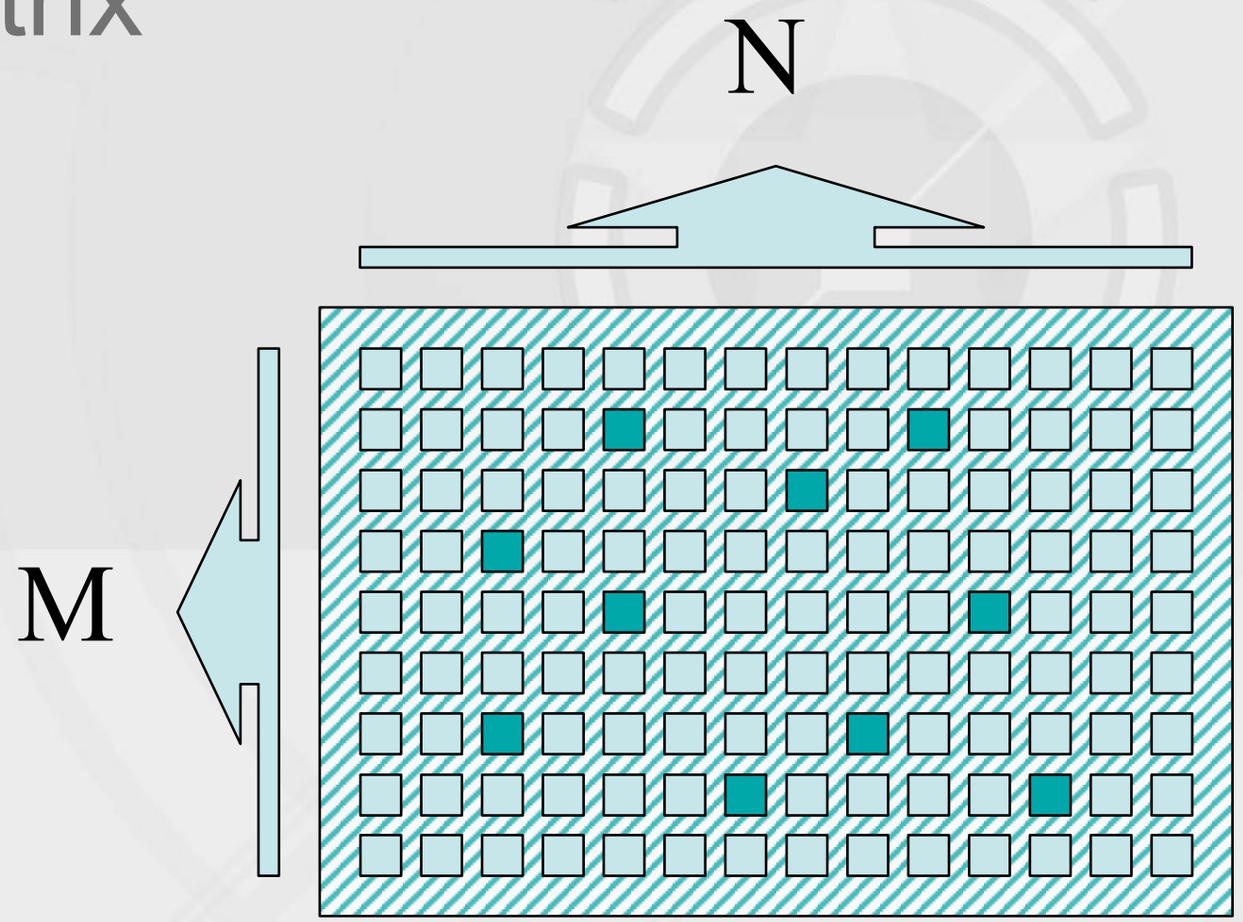
Color spaces: reference

- http://en.wikipedia.org/wiki/Color_space
- http://www.cs.unc.edu/~mcmillan/comp136/Lecture4/Color.html

2.1.2 Image representations

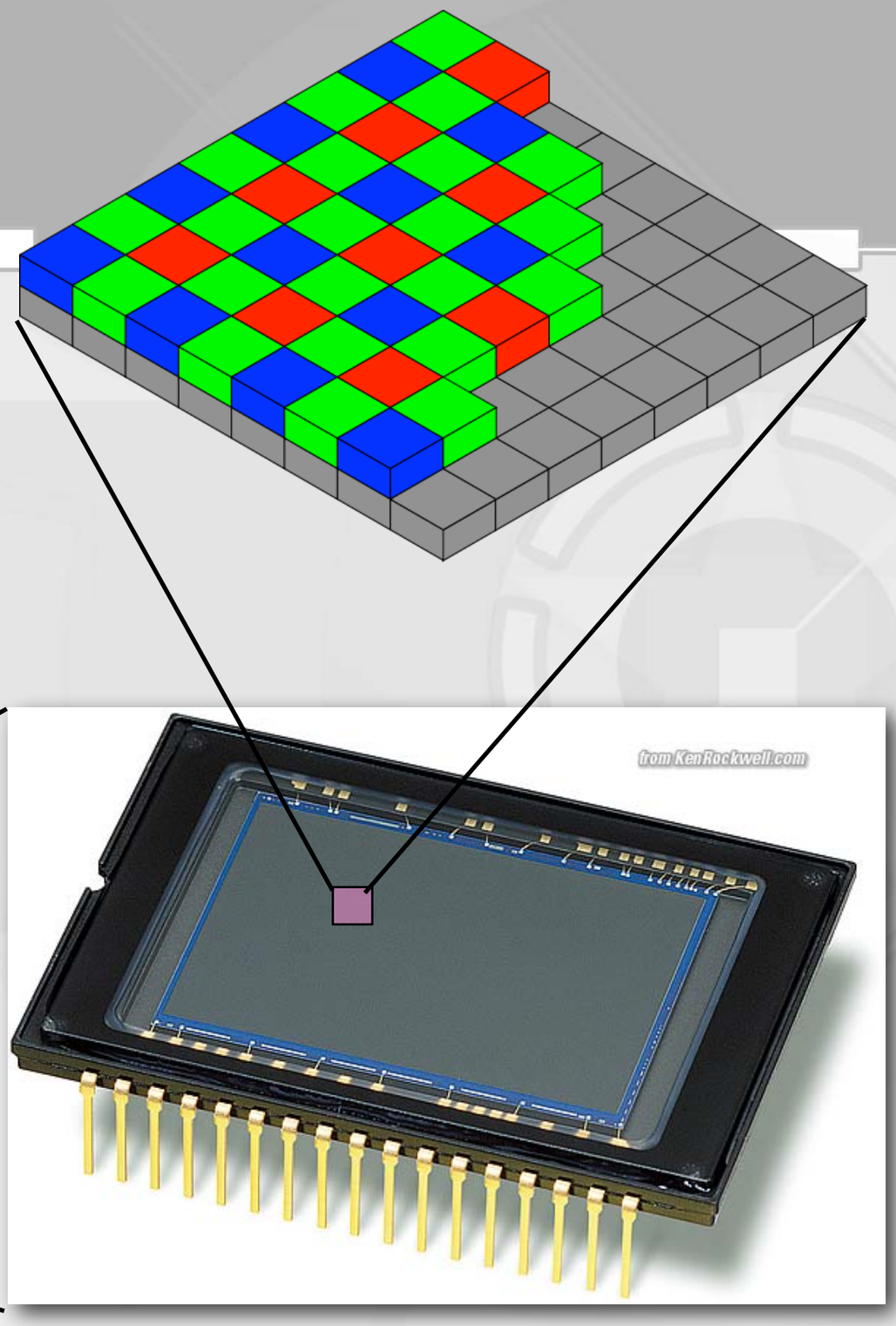
Representation of digital images

- An image can be viewed as a ***N×M*** vector matrix
  - Grayscale image
  - Color image
  - Palette



CCD sensors

Nikon D40



CCD sensor

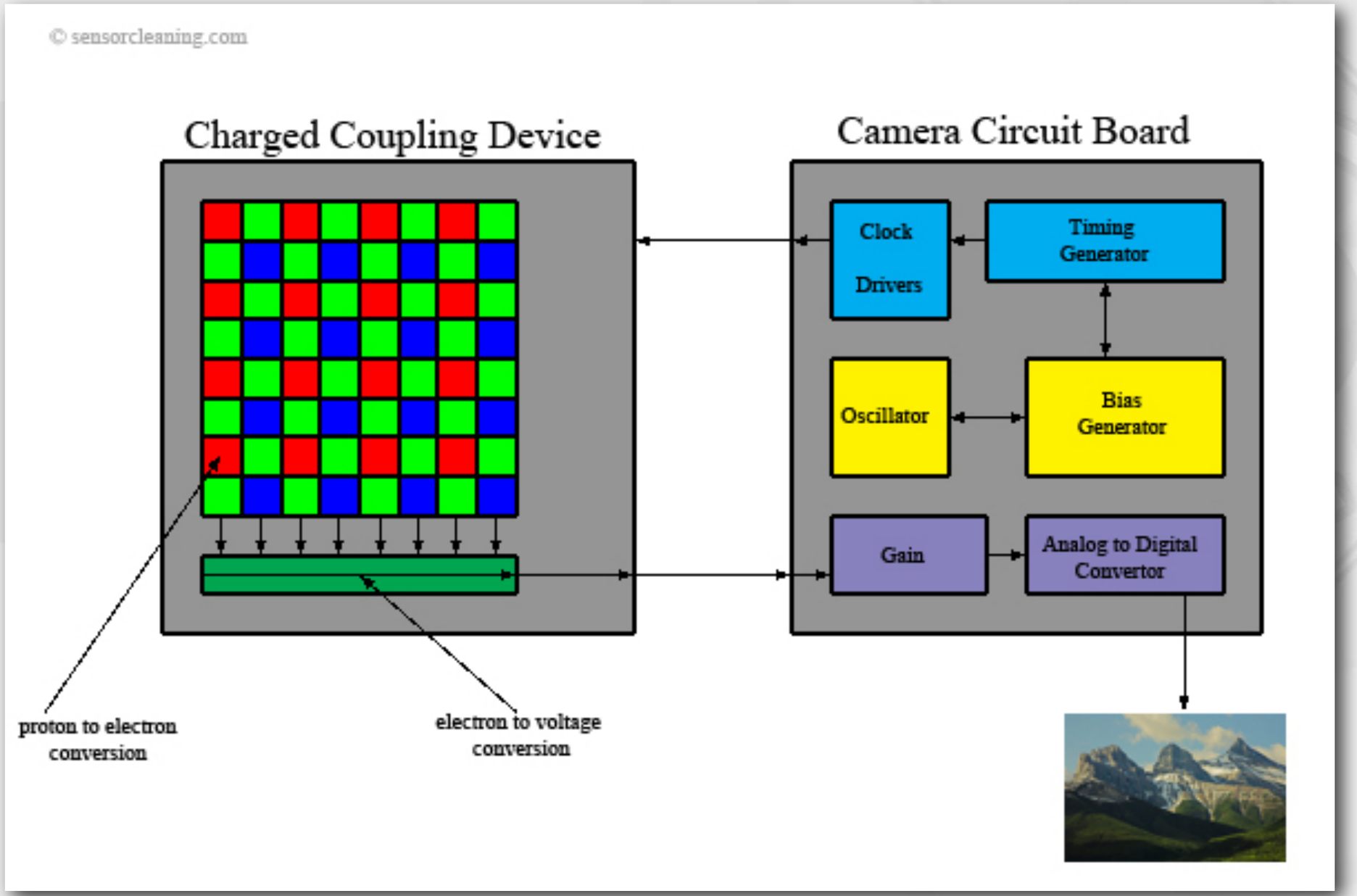
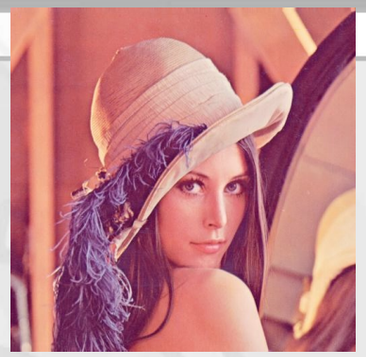

© sensorcleaning.com
Charged Coupling Device
Camera Circuit Board
Clock Drivers
Timing Generator
Oscillator
Bias Generator
Gain
Analog to Digital Convertor
proton to electron conversion
electron to voltage conversion

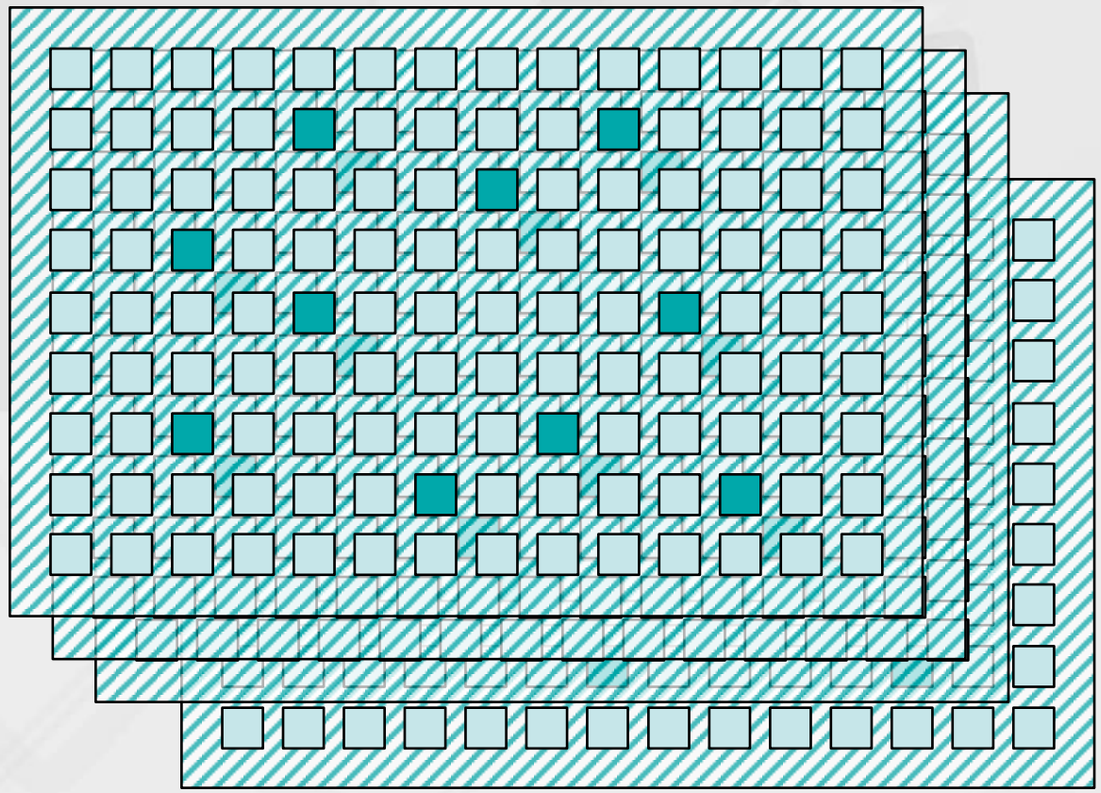
Image resolutions

Dimensions	MEGA pixels	Name	Comments
640x480	0.3	VGA	
720x576	0.4	CCIR 601 DV PAL	PAL DV, and PAL DVDs
768x576	0.4	CCIR 601 PAL full	PAL with square sampling grid ratio
800x600	0.4	SVGA	
1024x768	0.8	XGA	
1280x960	1.2		
1600x1200	2.1	UXGA	
1920x1080	2.1	1080 HDTV	high resolution digital TV format
2048x1536	3.1	2K	Typically used for digital effects in feature films.
4096x3072	12.4	4K (!!!)	

Rep of Images

- Binary image
  - –1 bit = Boolean value
  - –One bit-planes



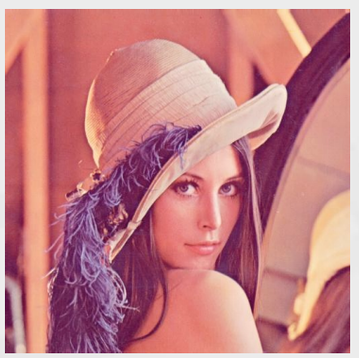


- Common Grayscale image:
  - –8 bits = 256 degrees of grayscale
  - –Eight bit-planes

Rep of Images

- Most used color images
  - –24bit RGB
  - –Red/Green/Blue each channel has 256 degrees of grayscale
  - –Can represent 2^24 = 16,777,216 types of color





Rep of image – Palette

- Some systems and applications can only use 8-bit color images
  - –Solution： Palette (Color look-up table)

High dynamic range image

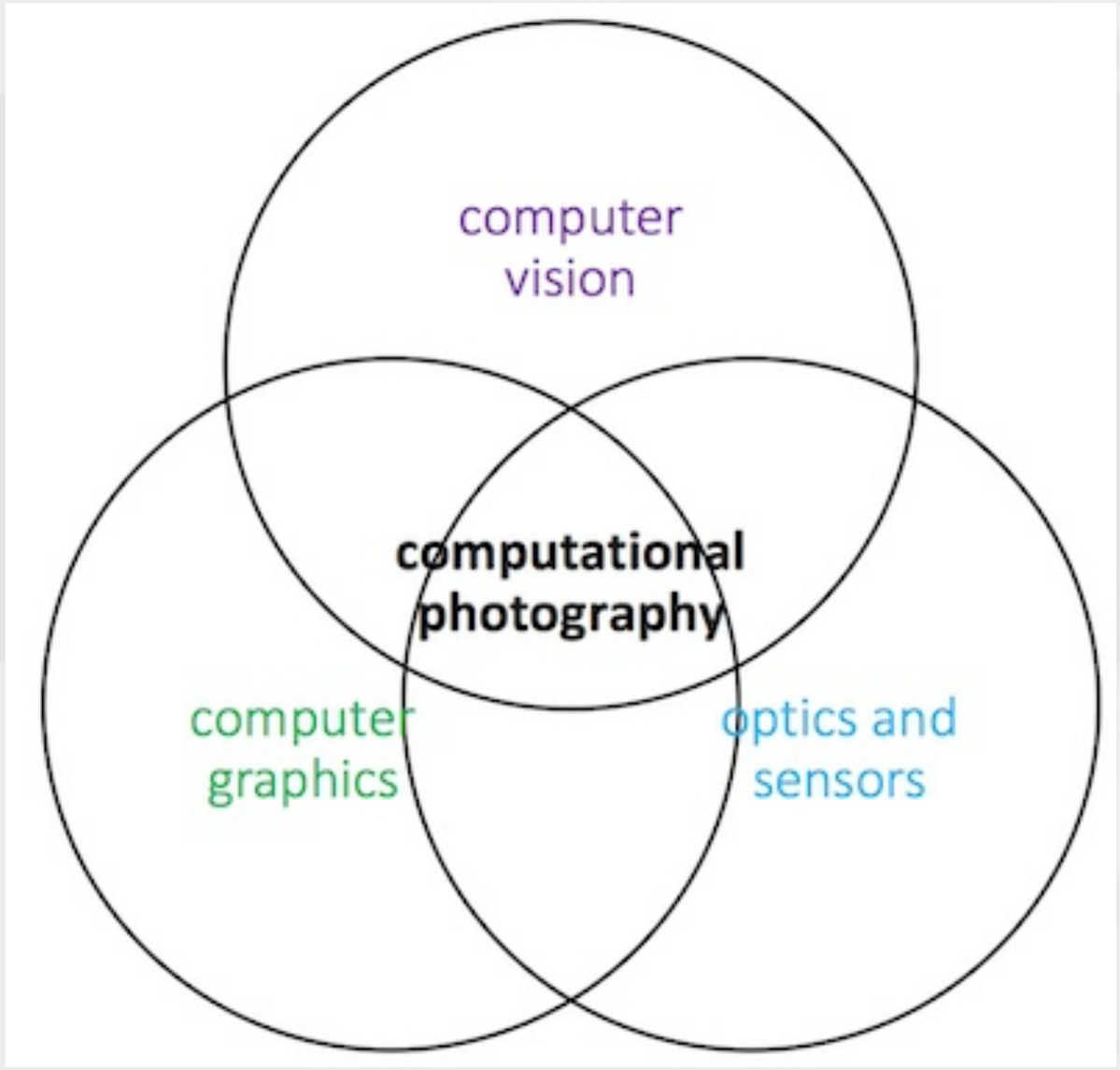
HDRI example: a New York City nighttime cityscape

High dynamic range image

- HDR pixels:
  - 16-bit or 32-bit floating point numbers
  - 10–12 bits luminance (亮度)
  - 8 bits chrominance (色度)
  - 10^−4 to 10^8: the range of visible luminance values
- CMOS image sensors: up to 110dB
- Tone mapping:
  - Typical computer monitors, prints, and other methods of displaying images only have a limited dynamic range

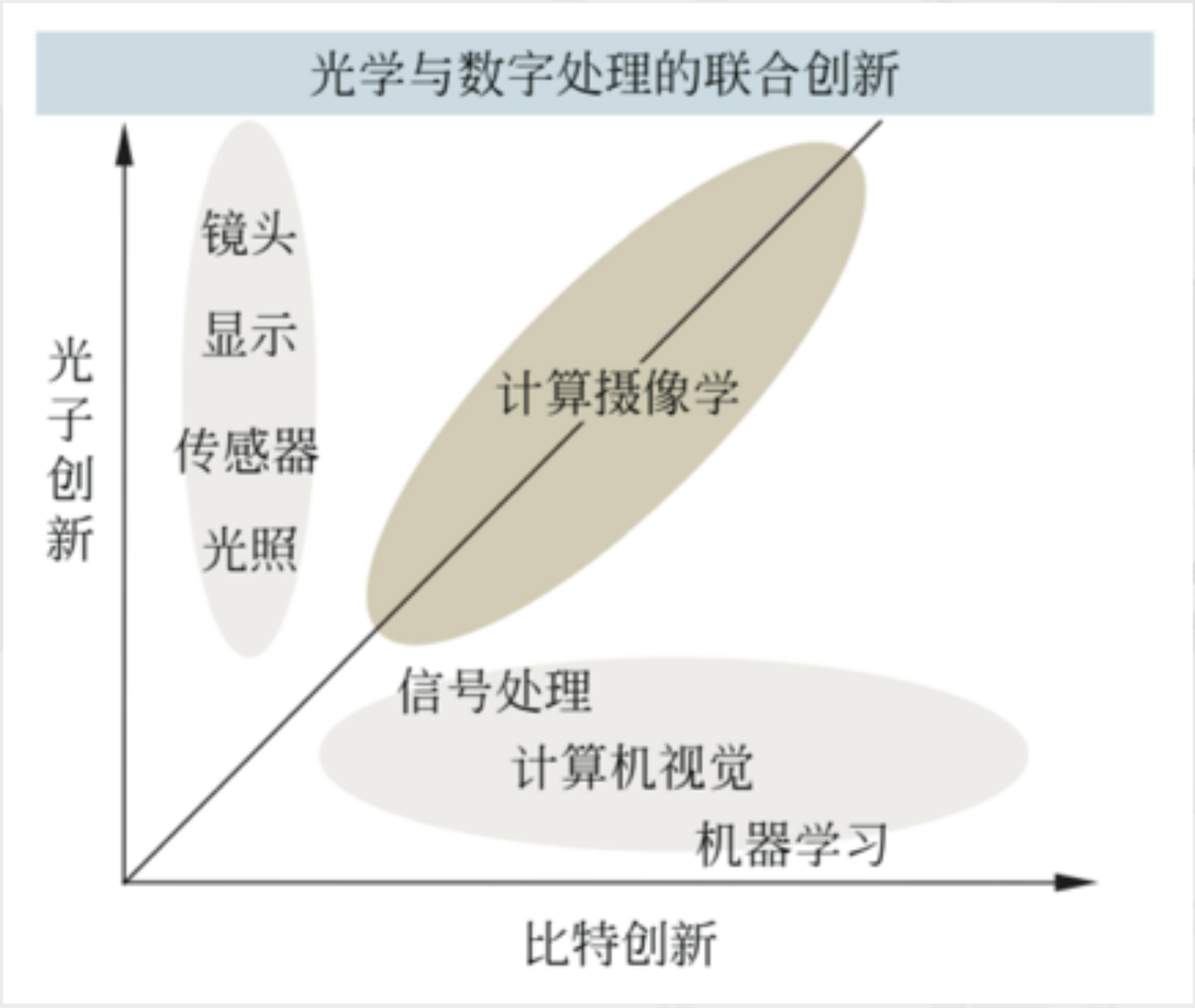
An example of a High Dynamic Range (HDR) photography, made of three different exposures

计算摄影学
COMPUTATIONAL PHOTOGRAPHY



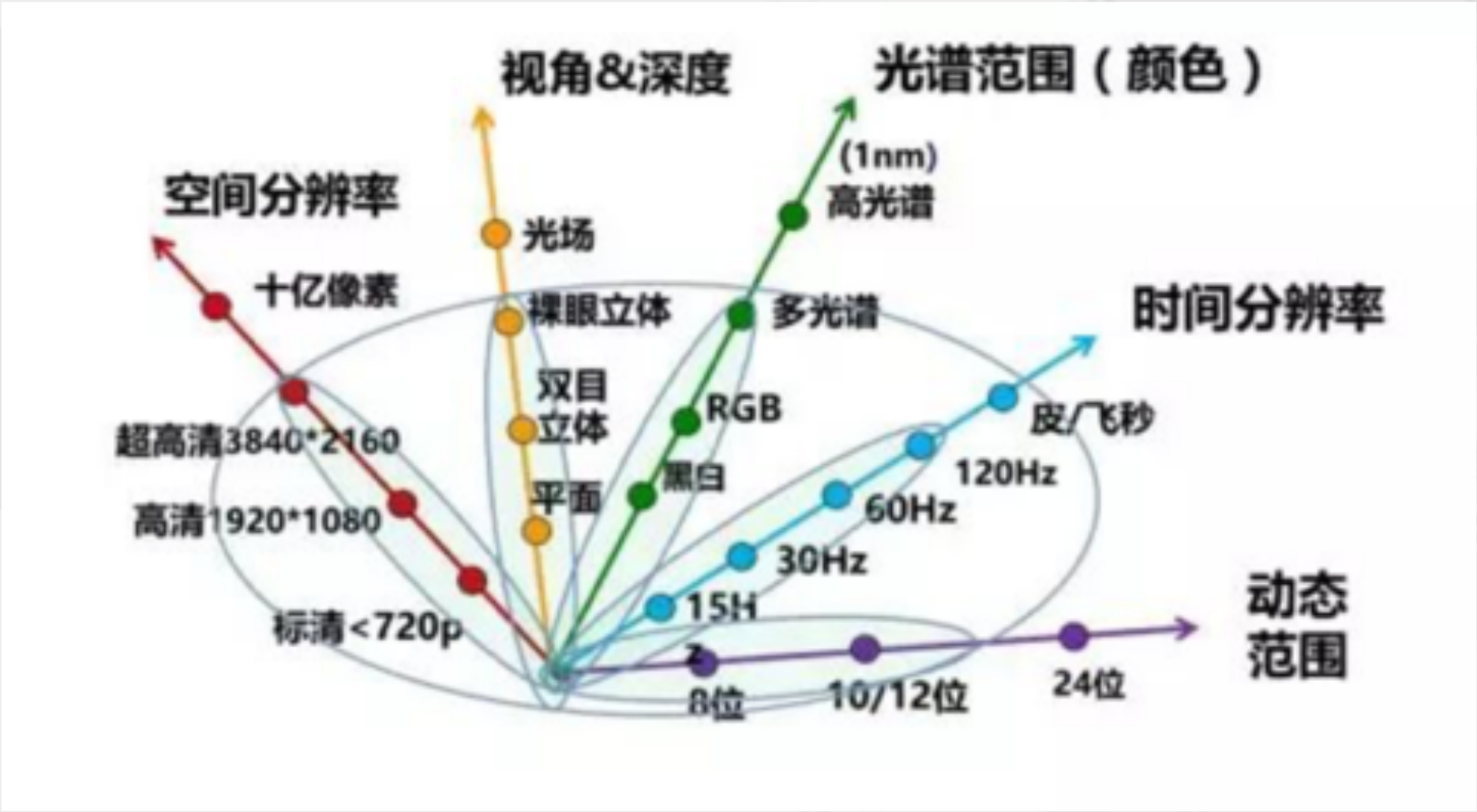
作为计算摄影学的一个重要起源，是1990年代计算机图形学与视觉技术的融合，出现了基于图像的建模与绘制技术（IBMR），形成了可视计算这一交叉性领域的兴起与发展。

计算摄影学
COMPUTATIONAL PHOTOGRAPHY



计算摄影学的学科衍生与交叉

计算摄影学
COMPUTATIONAL PHOTOGRAPHY



2.1.3 Image encoding

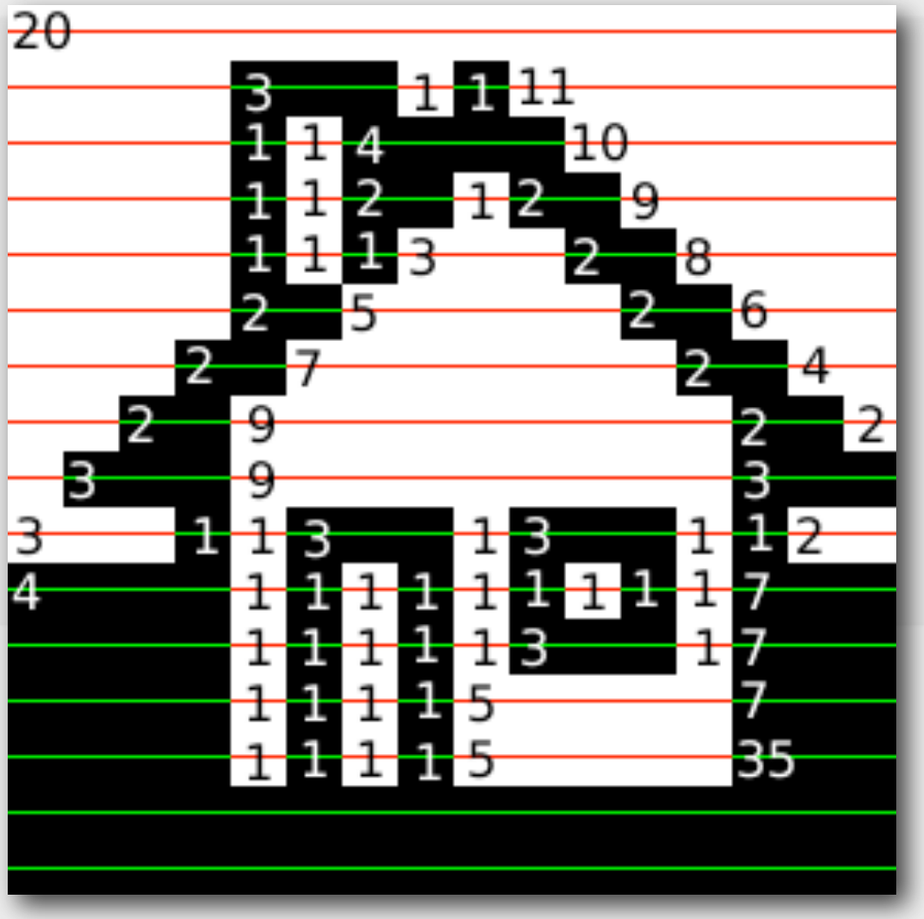
Image compression methods

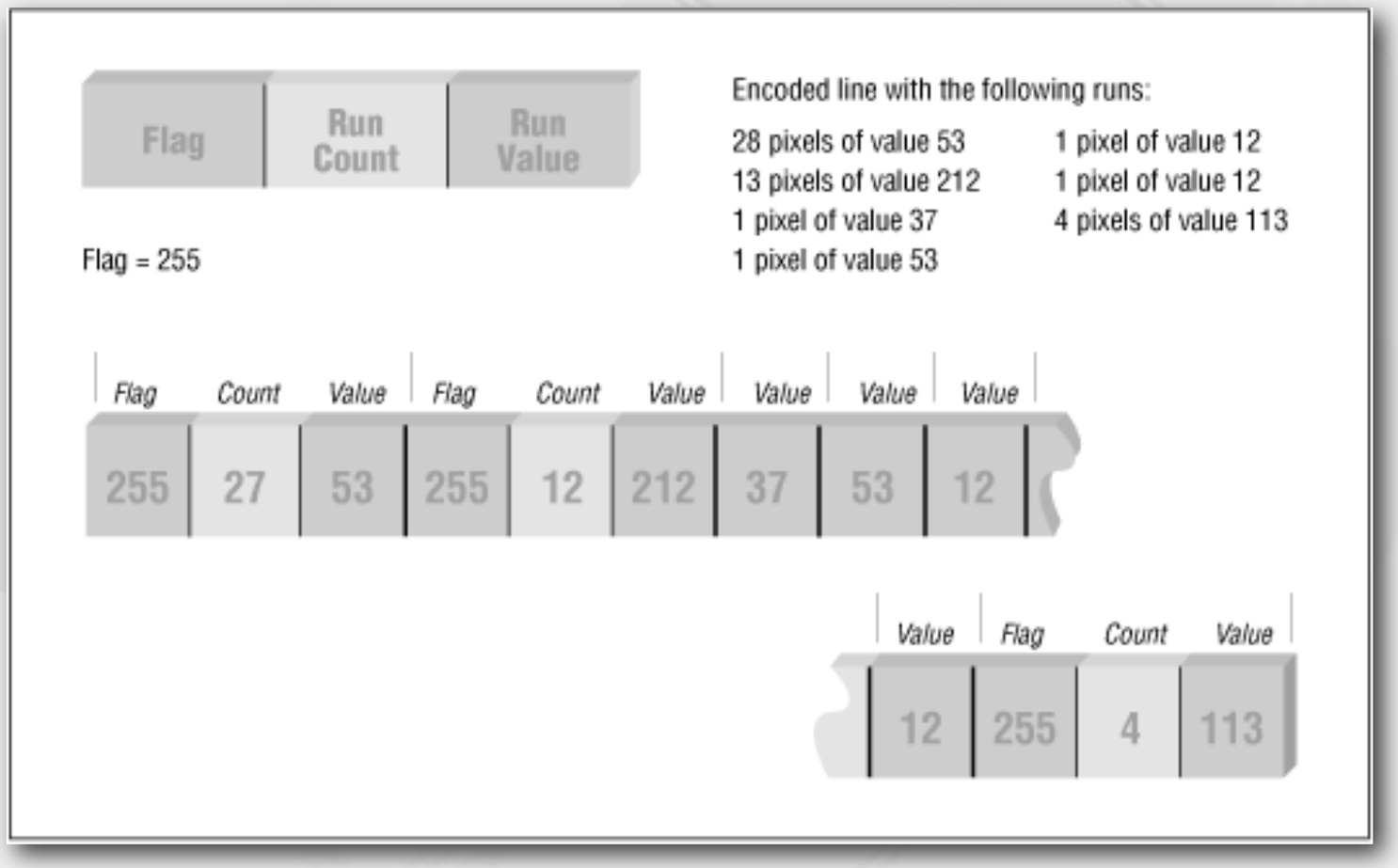
- lossless compression
- lossy compression

Lossless image compression methods

- Based on information theory
- General encoding methods
  - –RLC (Run-Length Coding)
  - –VLC (Variable-Length Coding)
  - –Dictionary Coding
  - –Arithmetic Coding

Run-Length Encoding





LZW : Lempel-Ziv-Welsh

- Universal lossless data compression algorithm
  –by Abraham Lempel, Jacob Ziv, and Terry Welsh

- The compressor algorithm builds a string translation table from the text being compressed

LZW - Compressor

```
w = NIL;
add all possible char codes to the dictionary
for (every character c in the uncompressed data) do
    if ((w + c) exists in the dictionary) then
        w = w + c;
    else
        add (w + c) to the dictionary;
        add the dictionary code for w to output;
        w = c;
    endif
done
add the dictionary code for w to output;
display output;
```

● 原输入数据为：

● A B A B A B A B B B A B A B A A C D A C D A D C A B A A A B A B ...

● 初始标号集为：

0	1	2	3	4	5
A	B	C	D	Clear	End

● 编码过程：

步骤	前缀	后缀	Entry	认识(Y/N)	输出	标号
1		A	(，A)			
2	A	B	(A,B)	N	A	6
3	B	A	(B,A)	N	B	7
4	A	B	(A,B)	Y		
5	6	A	(6,A)	N	6	8
6	A	B	(A,B)	Y		
7	6	A	(6,A)	Y		
8	8	B	(8,B)	N	8	9
9	B	B	(B,B)	N	B	10
10	B	B	(B,B)	Y		
11	10	A	(10,A)	N	10	11
12	A	B	(A,B)	Y		

● 编码结果：

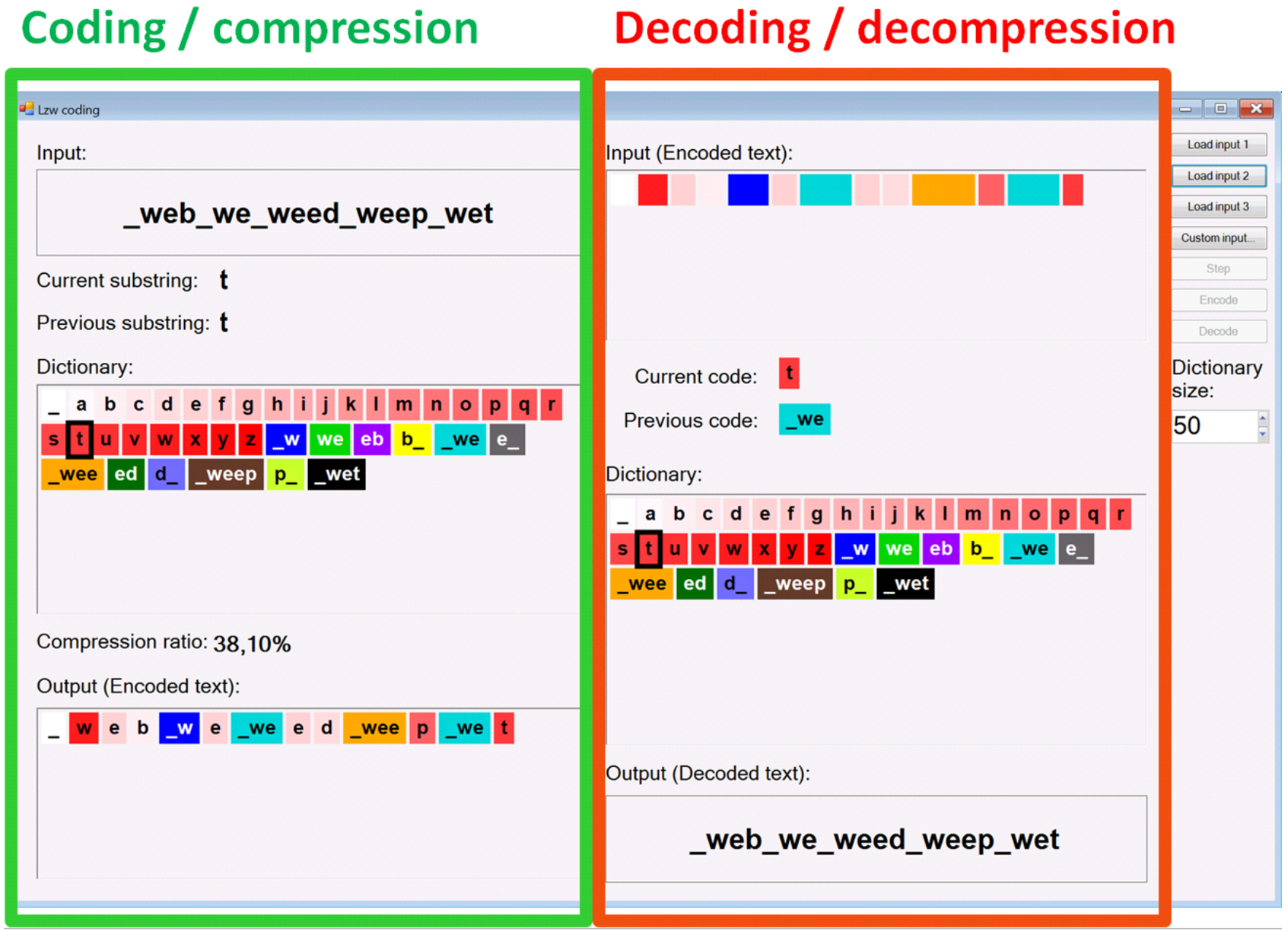
0	1	2	3	4	5	6	7	8	9	10	11
A	B	C	D	Clear	End	AB	BA	6A	8B	BB	10A

LZW - Decompressor

```
read a char k;
output k;
w = k;
while (read a char k) do
    if (index k exists in dictionary) then
        entry = dictionary entry for k;
    else if (k == currSizeDict)
        entry = w + w[0];
    else
        signal invalid code;
    endif
    output entry;
    add w+entry[0] to the dictionary;
    w = entry;
done
```

LZW Demo

http://vgg.fiit.stuba.sk/2015-02/lzw-coding/



Deflate

- a lossless data compression algorithm:
  - –LZ77 algorithm + Huffman coding.
  - –originally defined by Phil Katz for version 2 of his PKZIP archiving tool,
  - –later specified in RFC 1951.
  - –used by gzip, modern versions of zip and as part of the compression process of PNG, PPP, HTTP, SSH

Lossless image compression methods (cont.)

- Other lossless image compression methods
  - –Image different encoding （差分）
  - –Lossless JPEG (JPEG 2000)
    - discrete wavelet transform

Lossy image compression methods

- Quantization
- Transform coding
  - –Discrete Cosine Transform
  - –Discrete Wavelet Transform
  - –Karhune-Loeve Transform (Principle component analysis)

Image compression standards (jpeg.org)

- JPEG
  - –Joint picture encoding group
  - –Discrete Cosine Transform
- JPEG 2000
  - –newer standard
  - –Discrete Wavelet Transform

JPEG compression: main idea

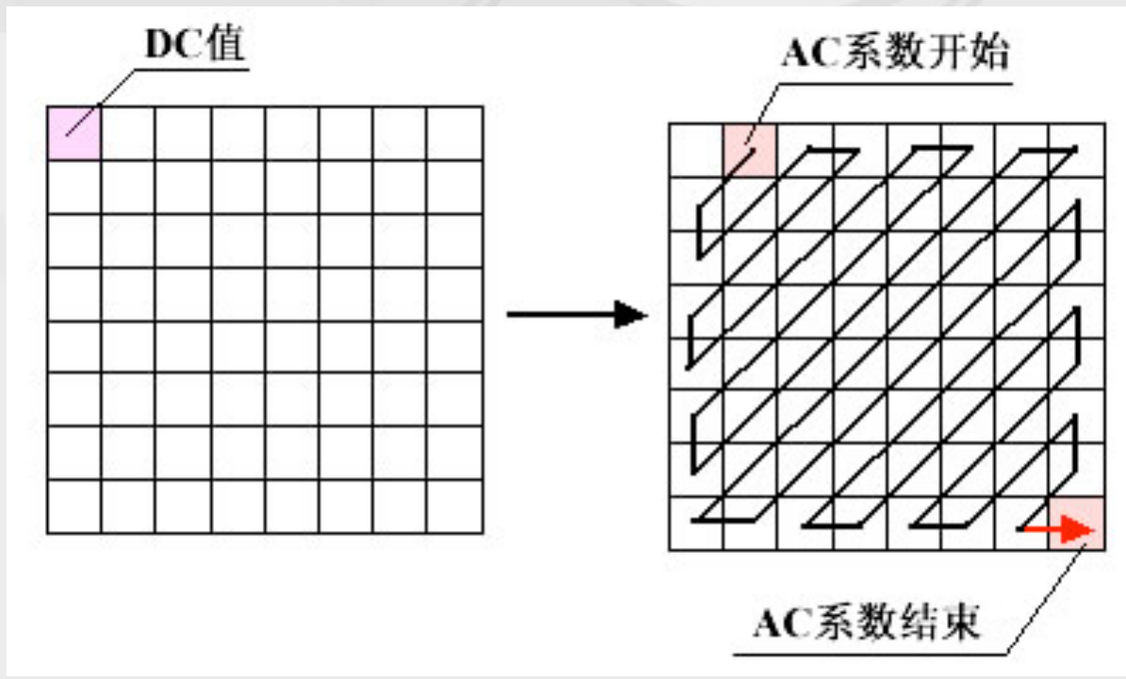
http://zh.wikipedia.org/zh-cn/JPEG

$$\begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 55 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix}$$

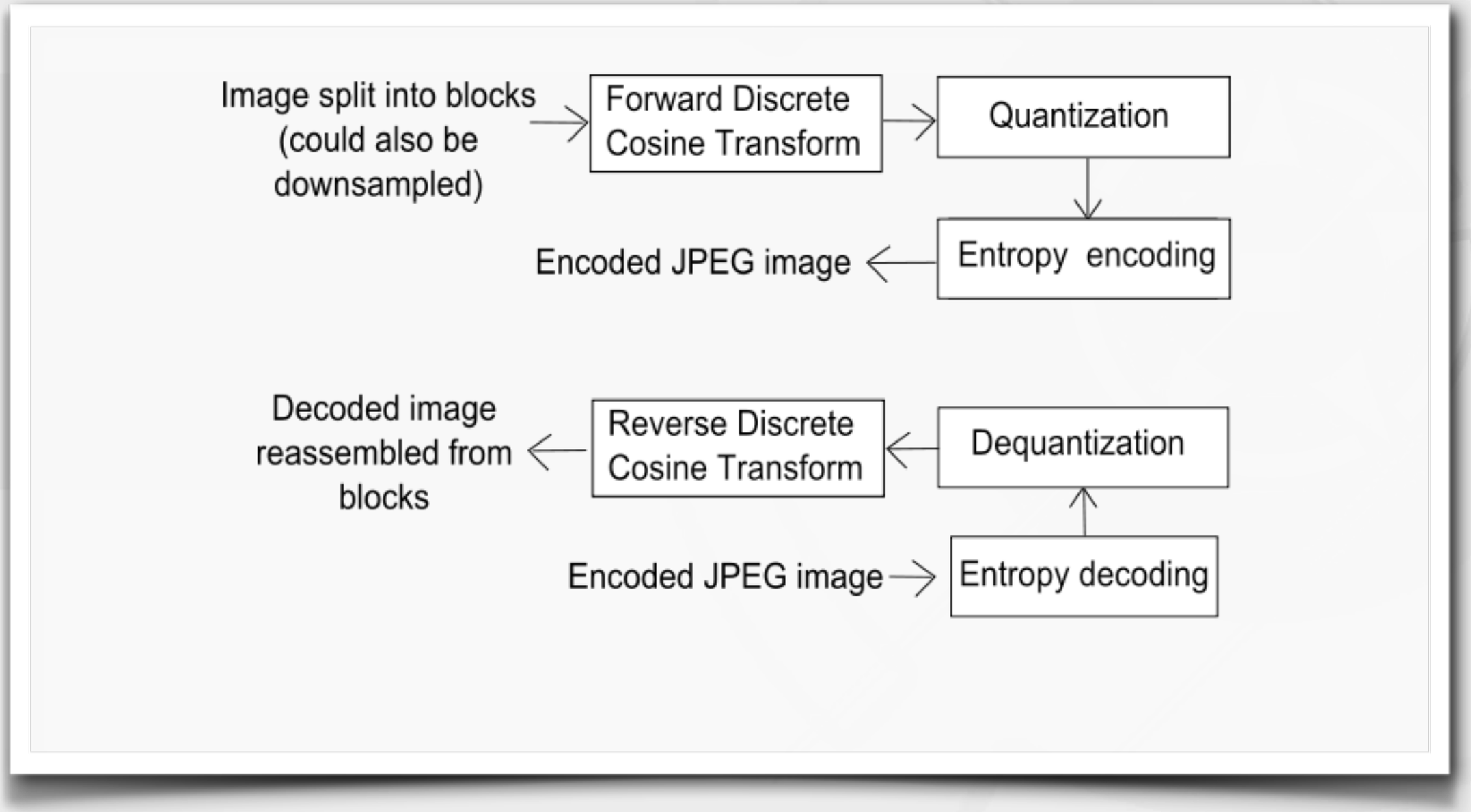
$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120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

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

RGB Image → YCbCr Color Model → 8x8 image blocks → DCT → frequency-domain representation → Quantization



JPEG compression: implementation



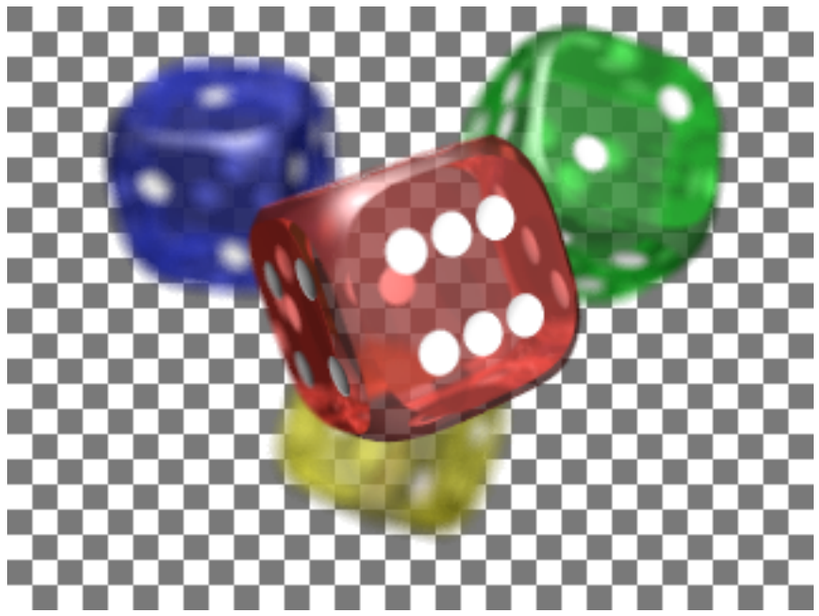
Common image formats - GIF

- Graphics Interchange Format
  - –UNISYS Corporation and Compuserve
  - –Lempel-Ziv-Welch compression method
  - –GIF87／GIF89a

  - –Features
    - Only support 8-bit (256) color image
    - Support several animation effects
    - Support interlaced image coding

Common image formats - PNG

- Portable Network Graphics
  - –motivation： Compuserv owns the LZW coding patent for GIF images
  - –open source
  - –Transparent
  - –PNG64



Common image formats - JPEG

- Lossy to lossless editing

Common image formats - TIFF (6.0)

- Tagged Image File Format
  - –flexible and adaptable
  - –handling images and data within a single file
  - –header tags: size, definition, image-data arrangement, applied image compression
  - –defining the image's geometry.

Common image formats - TIFF (6.0)

- a TIFF can be a container file
  - –compressed JPEG and RLE
  - –lossless compression
- include a vector-based Clipping path (outlines, cropping, image frames)

DNG: Digital Negative (数字负片)

- a royalty free **RAW** image format
- design by Adobe
- based on TIFF/EP
- mandates use of metadata

Summary – Essential factors of image storage

- Resolution
- Compression rate
  - –1bpp,2bpp, …
  - –Compression methods
- Color representation
  - –RGB, YUV, Lab …

Image converting tools

- ACDSEE
- imagemagik (Linux)
- XnView
  - http://perso.orange.fr/pierre.g/