

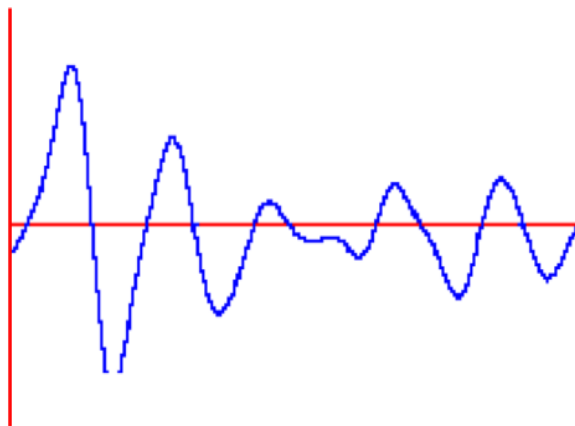


多媒体技术

数字图像处理技术 I

回顾：数字音频处理

- 下图展示了一段音频信号，其中横轴表示时间，纵轴表示幅度。



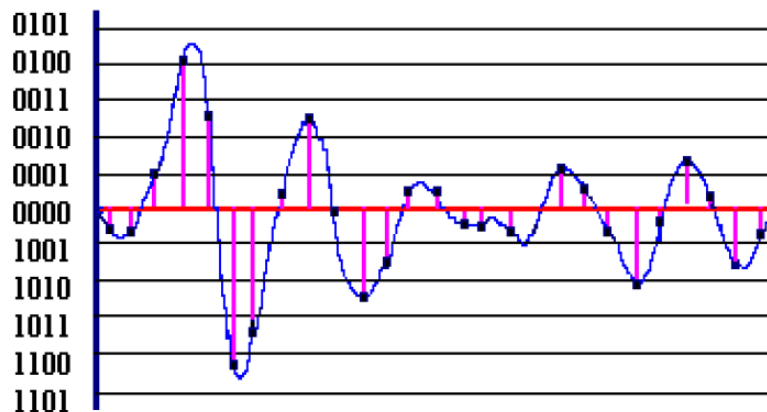
- (1) 如何将这段音频信号转化为数字信号，请简述其中的过程；
- (2) 请简述奈奎斯特采样定律；
- (3) 如果对这段音频采样频率为40kHz，分辨率为16位，立体声，录音时间为10s，符合CD音质的声音文件的大小是多少？

回顾：数字音频处理

- 如何将这段音频信号转化为数字信号，请简述其中的过程。

采样：每隔一段时间就记录一次声音信号的幅度

量化：把采样得到的声音信号幅度转换为数字值，先将整个幅度划分成为有限个小幅度（量化阶距），把落入某个阶距内的样值归为一类，并赋予相同的量化值。



回顾：数字音频处理

- 如何将这段音频信号转化为数字信号，请简述其中的过程。

采样频率：每秒钟所抽取声波幅度样本的次数，采样频率的计算单位是kHz

样本精度：每个采样点能够表示的数据范围，常用的有8位、12位和16位。例如8位量化级表示每个采样点可以表示256个（0~255）不同量化值。

回顾：数字音频处理

- 请简述奈奎斯特采样定律。

采样频率不应低于声音信号最高频率的两倍。

可以这样来理解奈奎斯特理论：声音信号可以看成由许许多多正弦波组成的，一个振幅为 A 、频率为 f 的正弦波至少需要两个采样样本表示，因此，如果一个信号中的最高频率为 F_m ，采样频率最低要选择 $2 \cdot F_m$ 。例如，电话话音的信号频率约为3.4 kHz，采样频率就选为8 kHz。

回顾：数字音频处理

- 如果对这段音频采样频率为40kHz，分辨率为16位，立体声，录音时间为10s，符合CD音质的声音文件的大小是多少？

数据量 Byte = 采样频率Hz

× (采样位数/8)

× 声道数

× 时间s

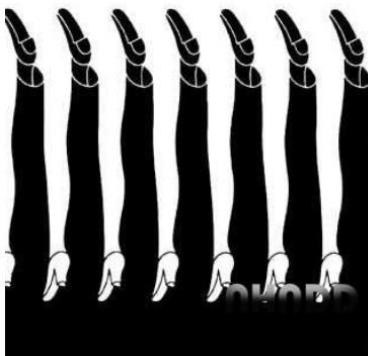
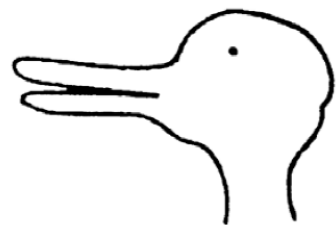
$$40000 \times (16/8) \times 2 \times 10 = 1600000 \text{Byte}$$

内容提纲

- 人的视觉系统和视觉常识
- 什么是数字图像？
- 颜色模型
- 图像数字化
- 像素空间关系
- 图像输入输出设备

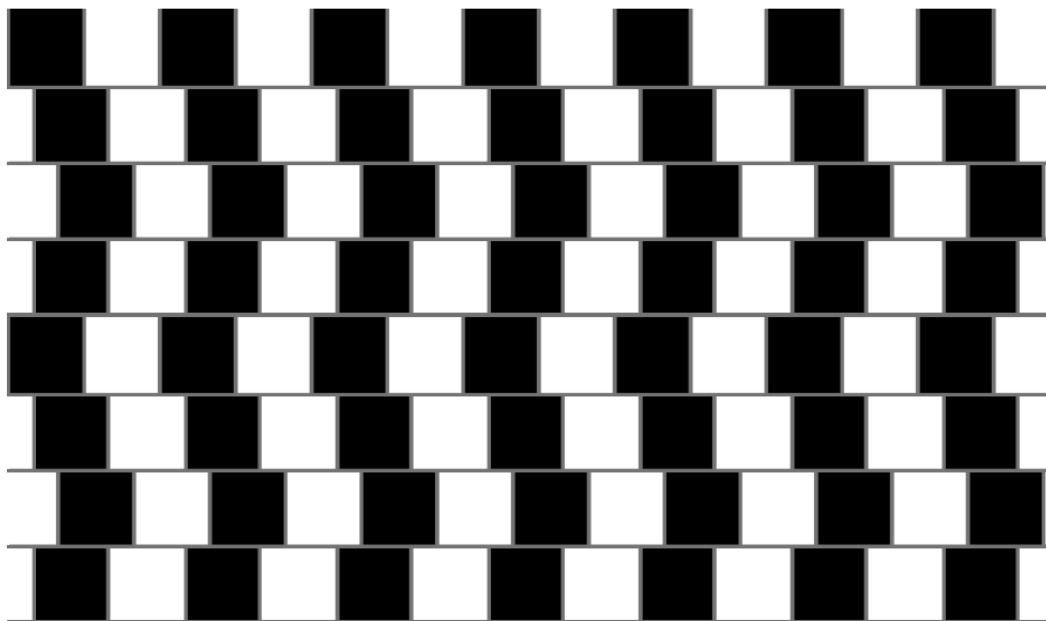
人的视觉系统

- 图像-背景现象



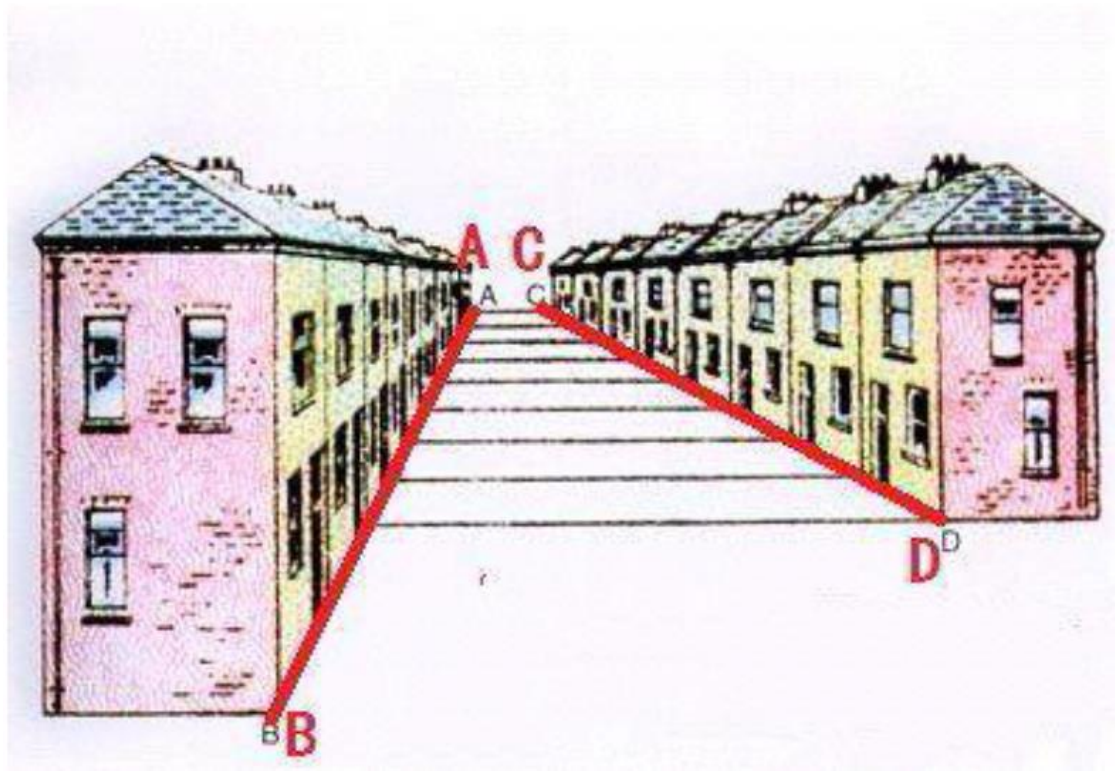
人的视觉系统

- 哪条是直线？



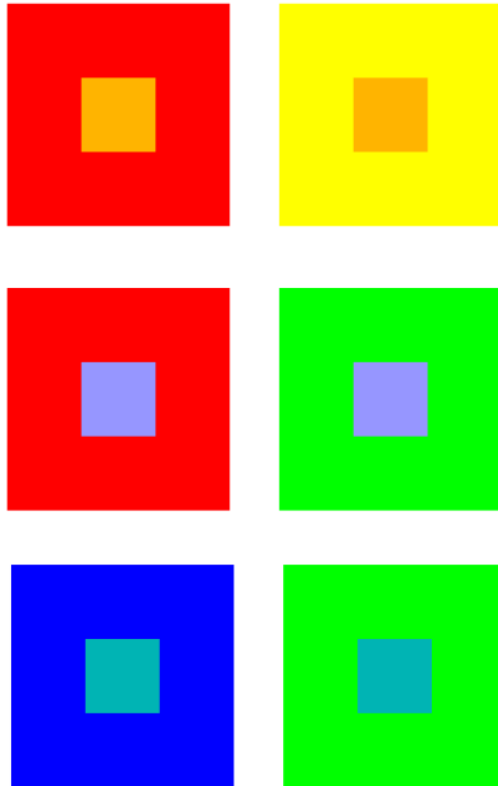
人的视觉系统

- AB和CD谁长？



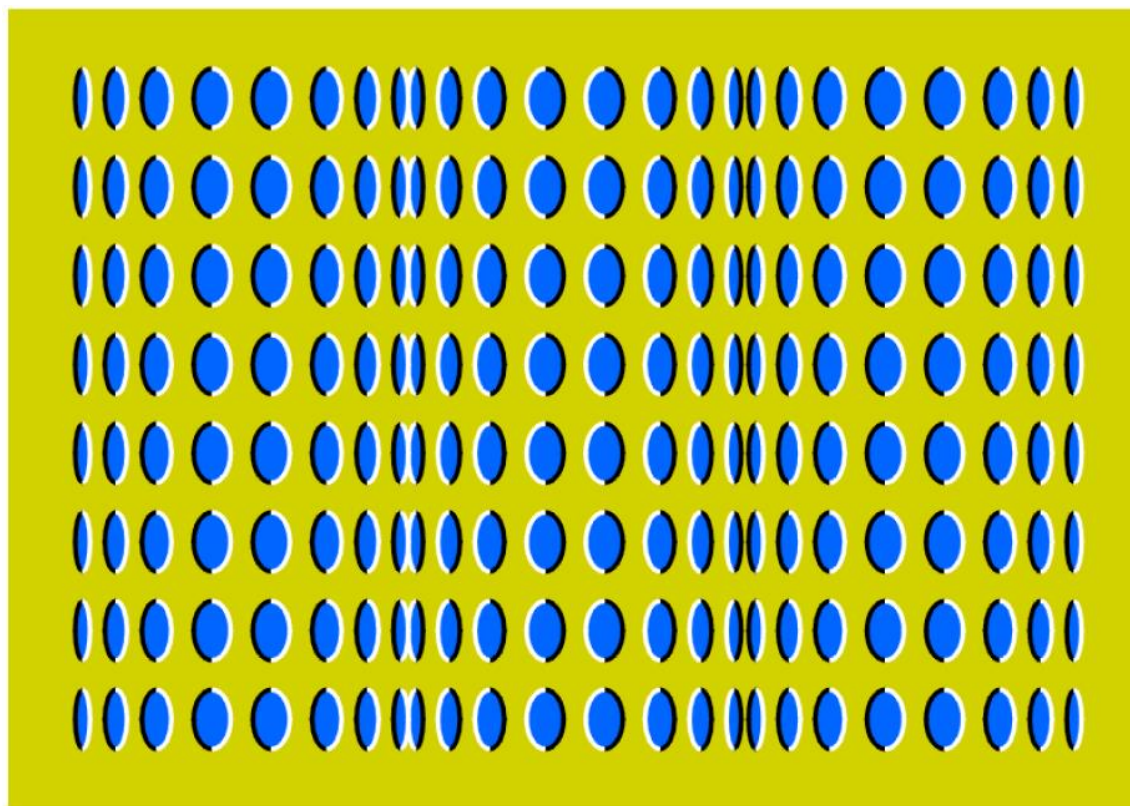
人的视觉系统

- 色彩对比



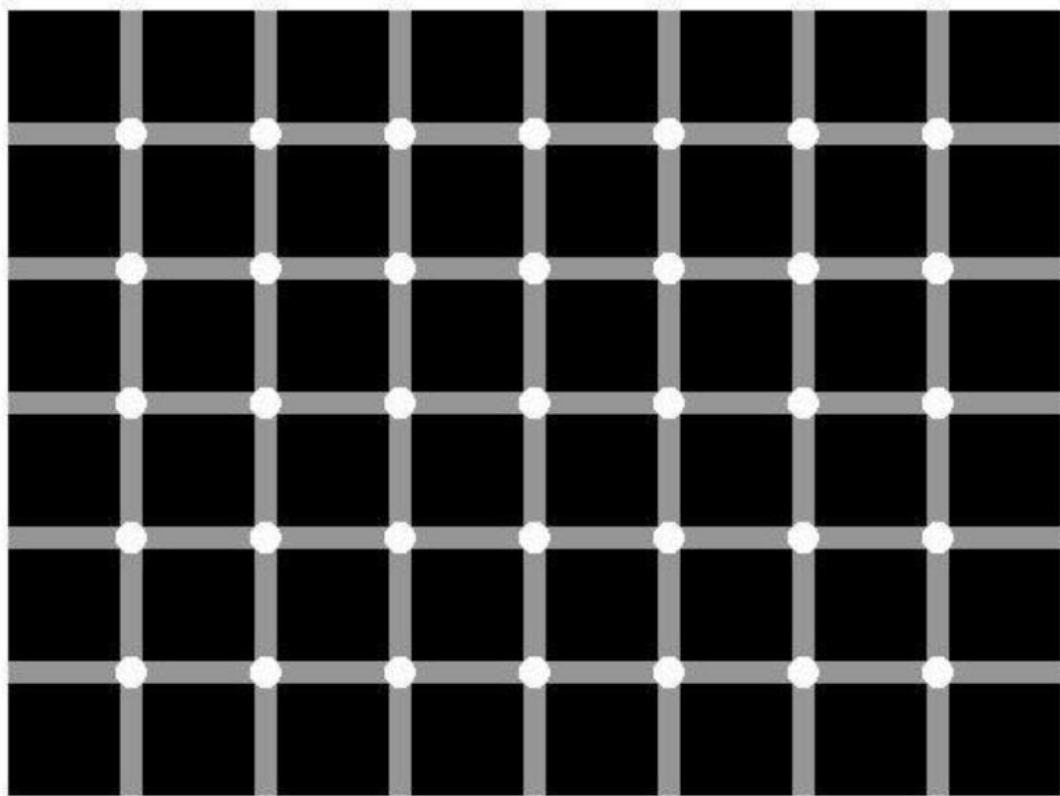
人的视觉系统

- 犯罪嫌疑人心理测试



人的视觉系统

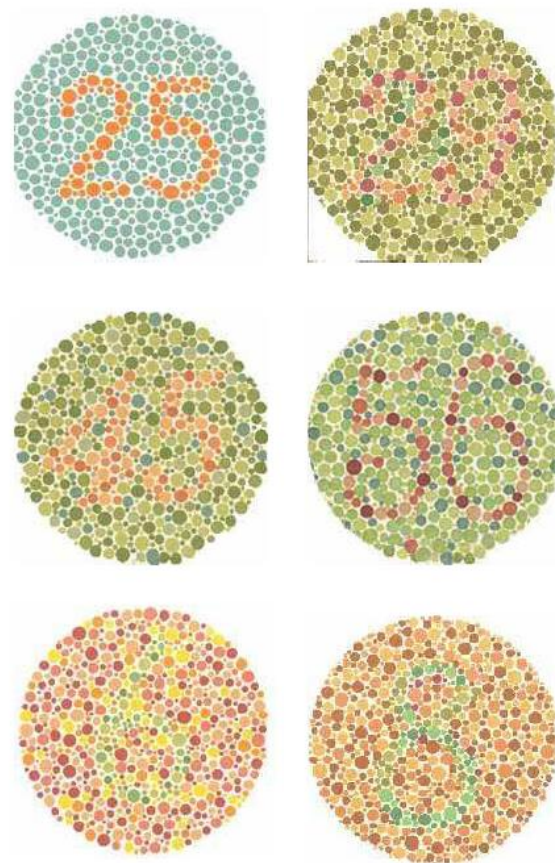
- 寻找黑点



人的视觉系统

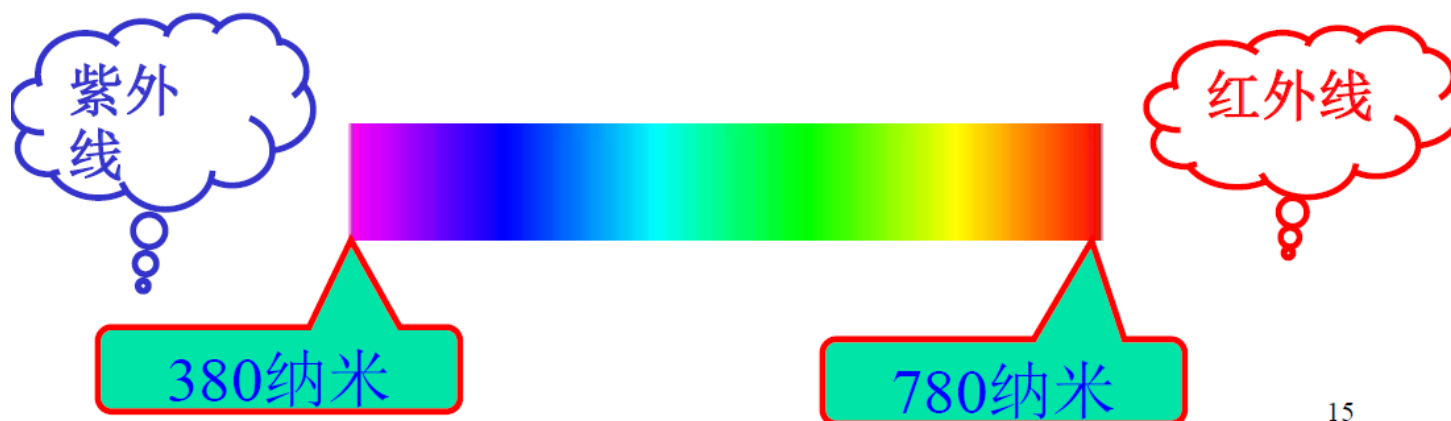
- 红绿色盲

25	Spots
Spots	56
Spots	Spots



视觉的常识

- 颜色是视觉系统对可见光的感知结果；
- 可见光波长在380 nm~780 nm之间；
- 我们看到的大多数光不是一种波长的光，而是由许多不同波长的光组合成的。

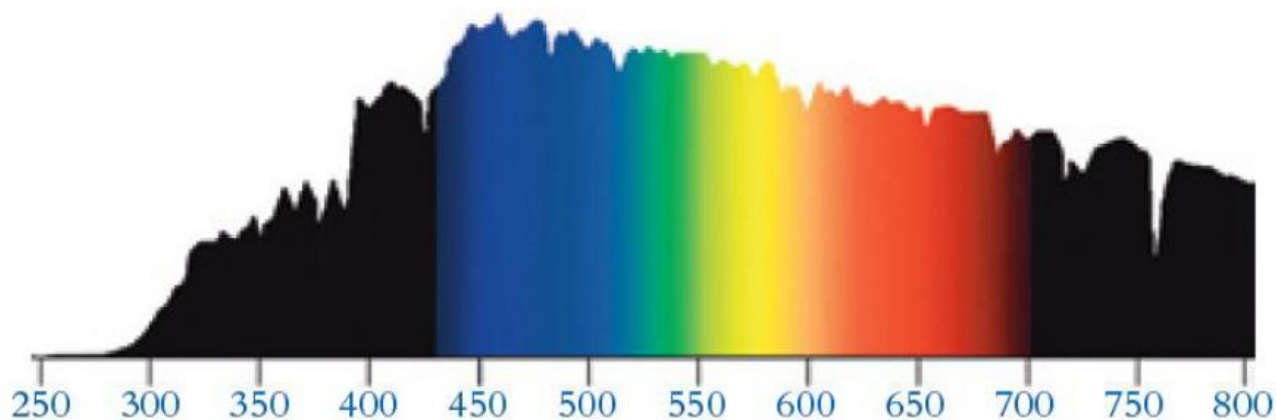


视觉的常识

- 可见光的度量参数

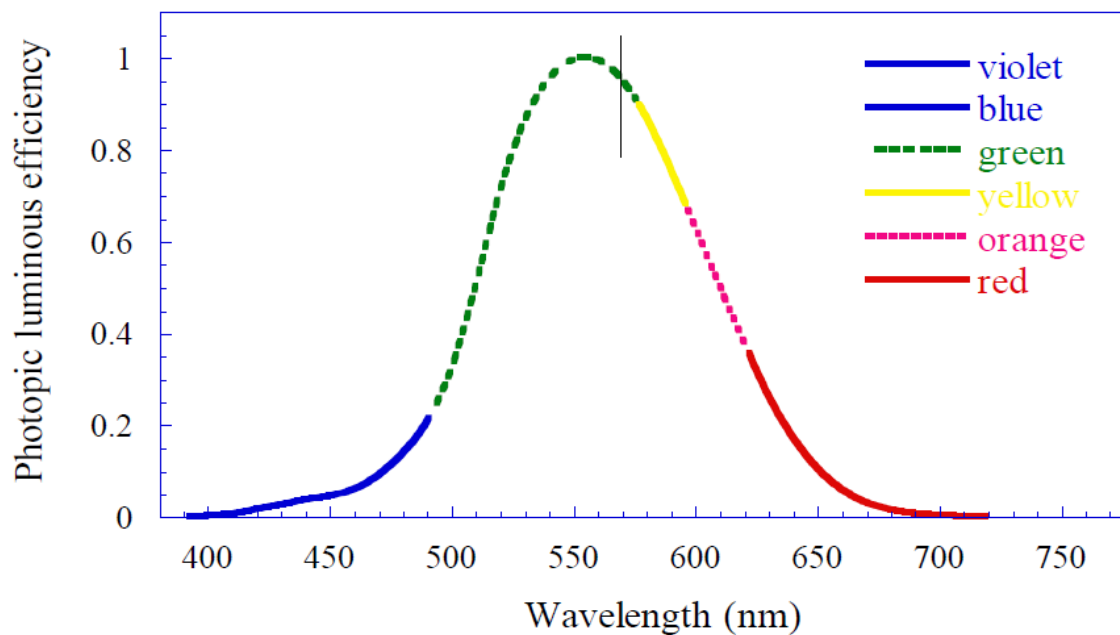
- 波长和强度

- 光谱的功率分布 (spectral power distribution, SPD) 可以完全表征



视觉的常识

- 视觉系统对颜色和亮度的响应特性

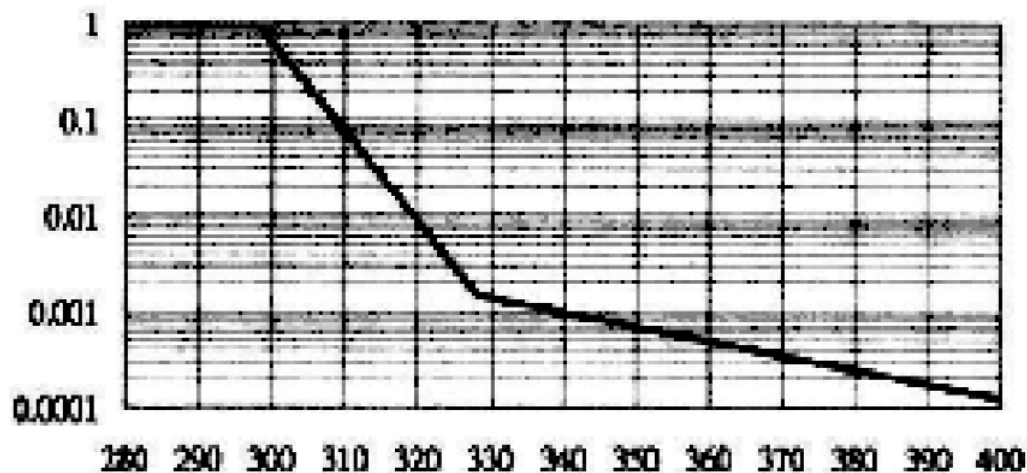


视觉的常识

- 皮肤的红斑作用光谱曲线

- 280nm—298nm是使皮肤变红最主要的紫外线波段

- 涂抹防晒霜相当于形成一个阻碍该波段光线通过的滤波器



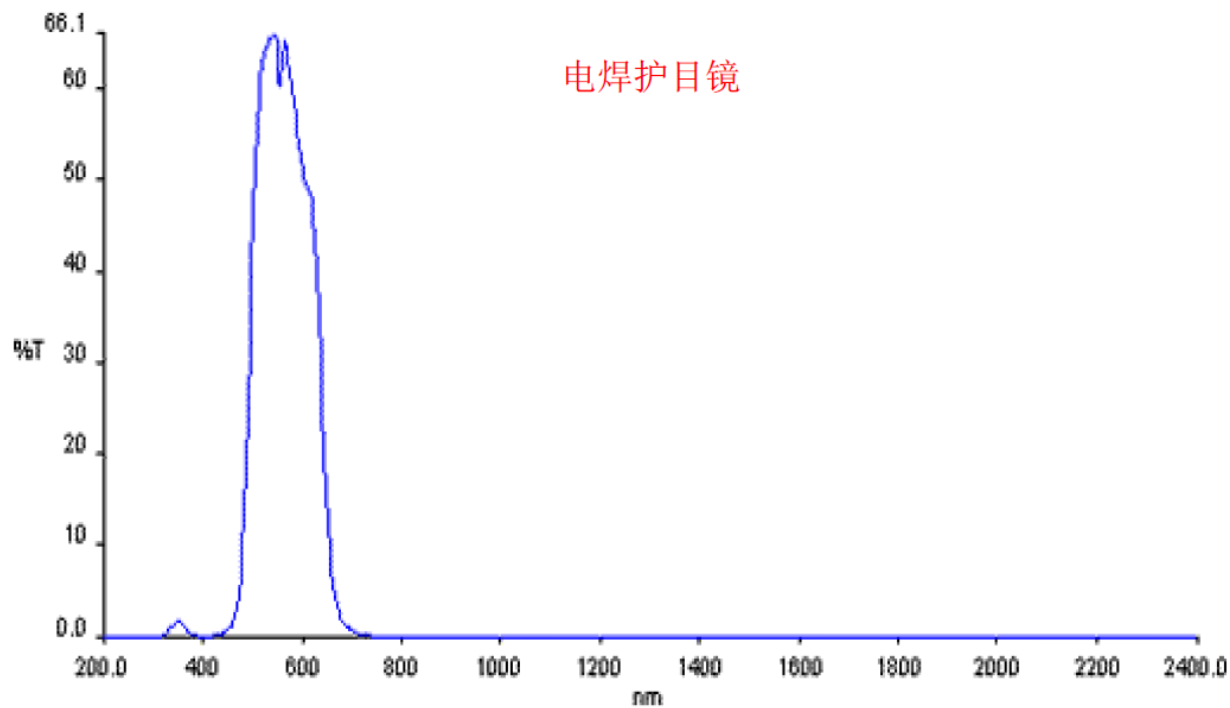
视觉的常识

● 生活中的滤波概念——眼镜

- 太阳镜：可见光透过率约为20%，有绿、灰、茶等多种颜色
- 气焊用护目镜：可见光透过率1%以下，能全部吸收500nm以下波长的光波，仅能有少量红外线通过，供气焊工焊接用
- 电焊用护目镜：可见光透过率0.1%，能全部阻截紫外线，红外线透过<5%
- 蓝色护目镜：吸收500-600nm波段的耀眼眩光，透光率12%以下，适用各种工业高温炉操作人员使用
- 深红色玻璃护目镜：通吸收600nm以下波段的全部光线，适用医务或工业人员操作X-光透视时护目之用
- 3d眼镜：观看3d电影时使用

视觉的常识

- 生活中的滤波概念——眼镜



内容提纲

- 人的视觉系统和视觉常识
- 什么是数字图像？
- 颜色模型
- 图像数字化
- 像素空间关系
- 图像输入输出设备

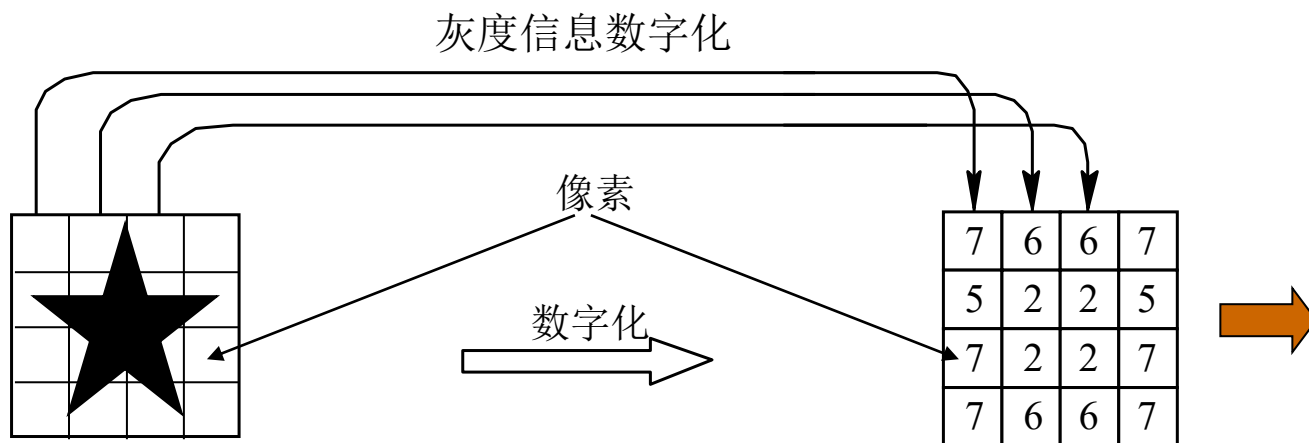
什么是图像？

- 图像是物体在平面坐标上的直观再现，是其所表示物体的信息的直接描述和概况表示。



什么是数字图像？

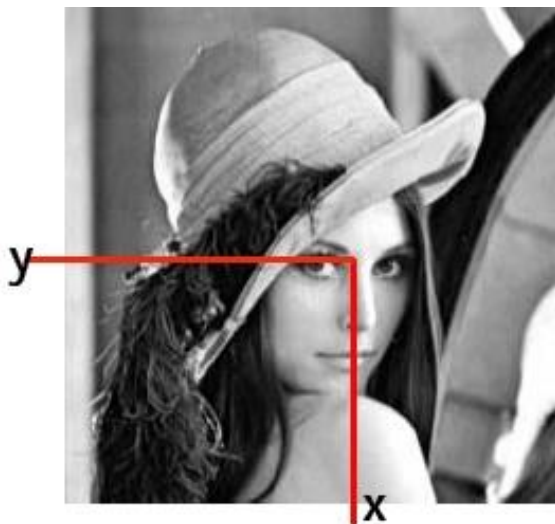
- 像素组成的二维或三维排列，可以用矩阵表示。



用二维矩阵表示

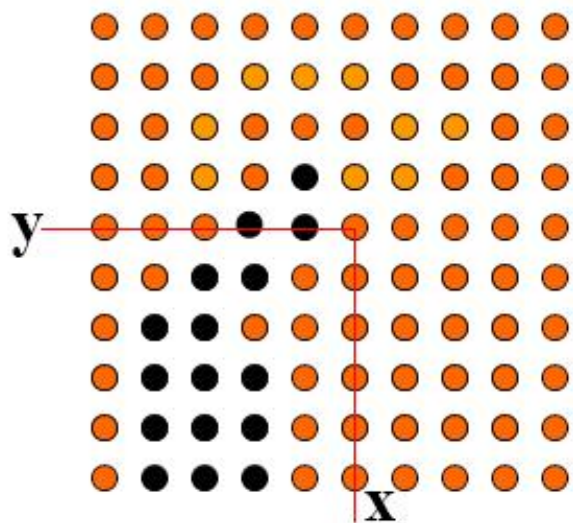
什么是数字图像？

- 图像可以用一个二元函数来表示 $f(x, y)$
 - x, y 分别表示横坐标和纵坐标，函数 f 的值表示亮度（intensity）或者灰度值（gray level）
 - 数字图像（digital image）：当 x, y 和函数 f 的值都是有限、离散的



什么是数字图像？

- 数字图像由若干元素 (elements) 组成，每个元素包括位置和亮度信息。这些像素被称为 picture elements, image elements 和 **pixels (像素)**。



什么是数字图像？

- 对于单色（灰度）图像而言，每个像素的亮度用一个数值来表示，通常数值范围在0到255之间，0表示黑、255表示白，其它值表示处于黑白之间的灰度。
- 彩色图像可以用红、绿、蓝三元组的二维矩阵来表示。通常三元组的每个数值也是在0到255之间，0表示相应的基色在该像素中没有，而255则代表相应的基色在该像素中取得最大值。

数字图像的类型

- 二值图像

- 灰度值只有0和1两个的图像。0代表黑色。1代表白色。

- 灰度图像

- 一般指具有256级灰度值的图像。数据类型一般为8位无符号整数。

- 0代表黑色，255表示白色。其它表示从黑到白的过渡色。

数字图像的类型

- 索引图像

- 包括存放图像数据的二维矩阵，还包括一个颜色索引矩阵（MAP）。两个矩阵建立映射关系。

- RGB彩色图像

- 称为真彩色图像。由R、G、B来表示每个像素的颜色。颜色值直接存放在矩阵中，无需索引。即 $M \times N \times 3$

数字图像处理的 目的

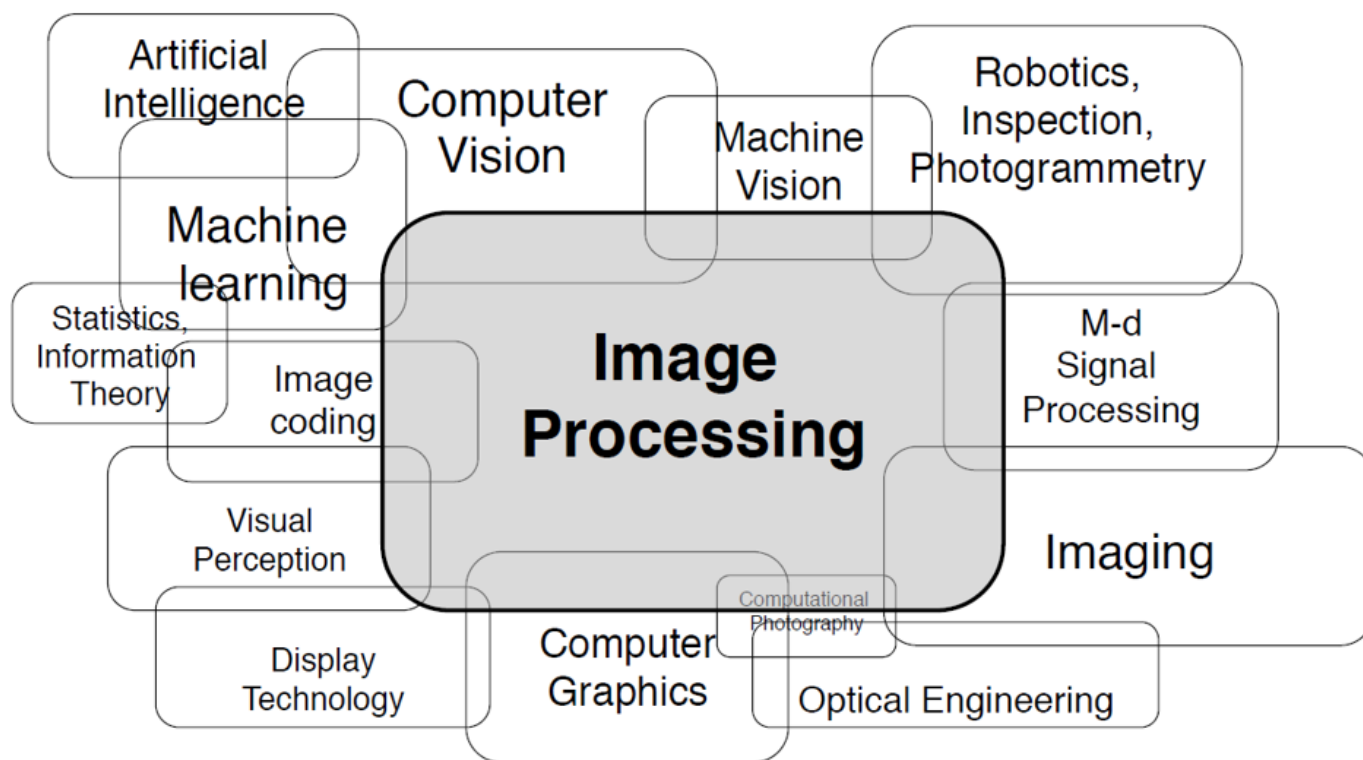
- 目的一：人类理解和分析。
 - 图像在传输过程后的更好复原
 - 空间应用：图像增强、图像复原、医学图像.....
- 作用：通过图像增强、图像复原等技术，增强或复原模糊或损毁的图像。

数字图像处理的 目的

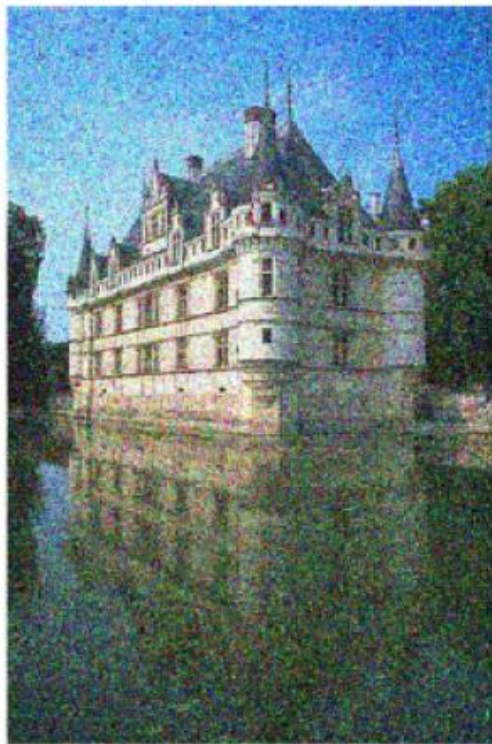
- 目的二：机器识别

- 自动字符识别、人脸识别、指纹识别，生物特征识别.....

数字图像处理的 目的



数字图像处理举例



数字图像处理举例



数字图像处理举例



数字图像处理举例



数字图像处理举例



内容提纲

- 人的视觉系统和视觉常识
- 什么是数字图像？
- 颜色模型
- 图像数字化
- 像素空间关系
- 图像输入输出设备

各种颜色模型

- RGB相加颜色模型

- 红色 + 绿色 + 蓝色

- CMY相减颜色模型

- 青色 + 品红 + 黄色

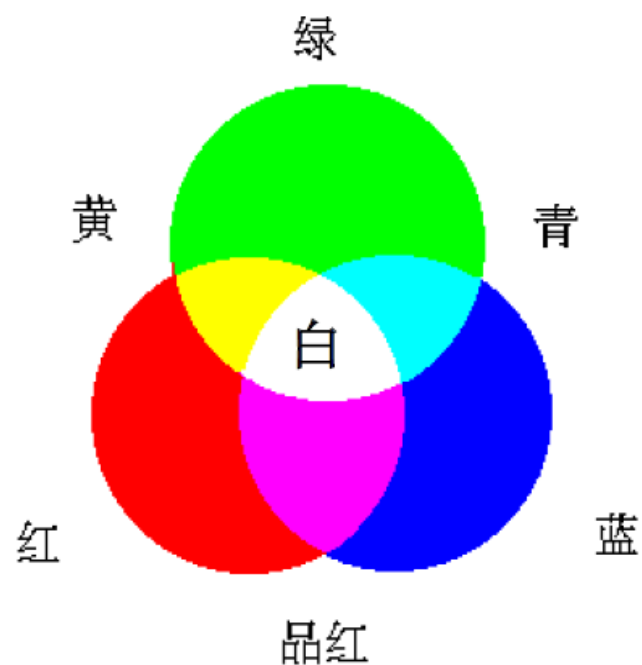
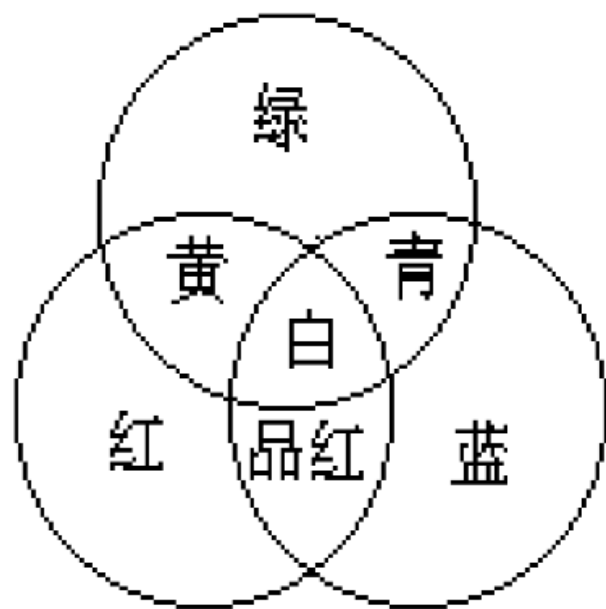
- HSL颜色模型

- 色调 (hue)、饱和度 (saturation)、亮度 (luminance)

- YUV颜色模型

- Y表示亮度、UV用来表示色差，UV是构成彩色的两个分量

RGB相加颜色模型

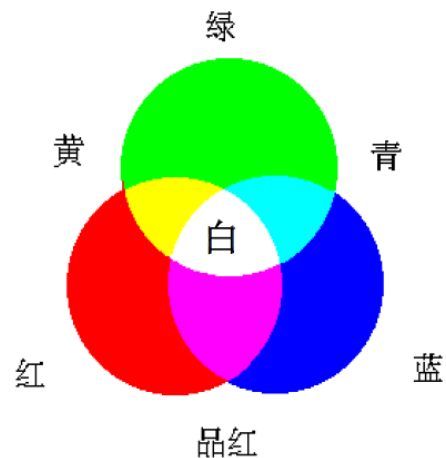
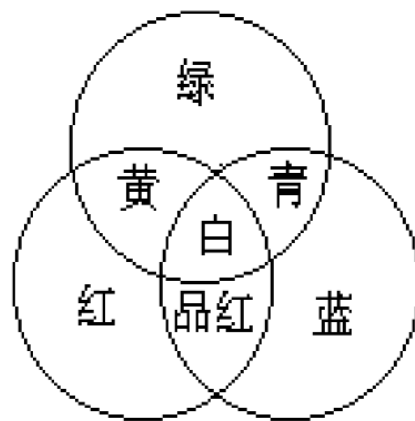


CMY相减颜色模型

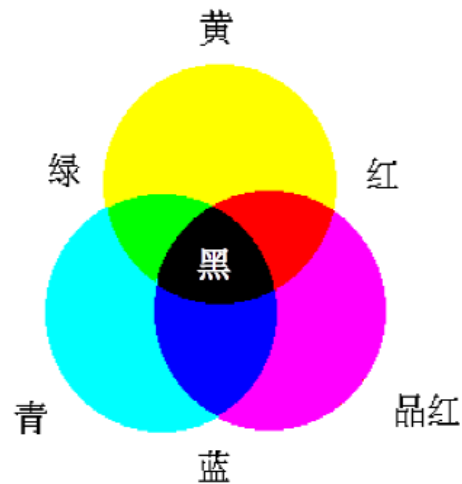
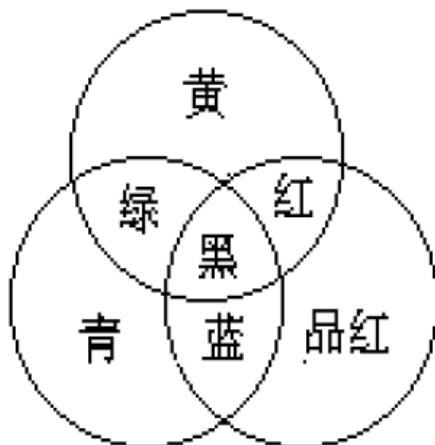
- 用彩色墨水或颜料进行混合，这样得到的颜色称为相减色。
- 这三种颜色是青色(Cyan)、品红(Magenta)和黄色(Yellow)，通常写成CMY，称为CMY模型。
- 用这种方法产生的颜色之所以称为相减色，是因为它减少了为视觉系统识别颜色所需要的反射光。

CMY相减颜色模型

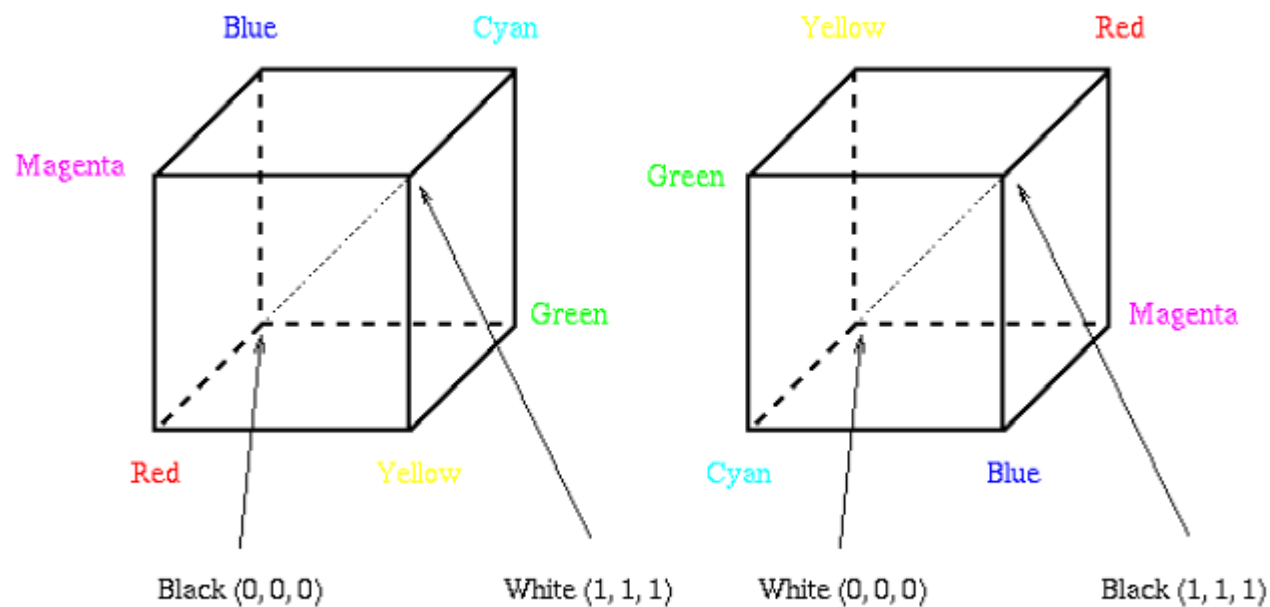
RGB



CMY



RGB & CMY立方体



The RGB Cube

The CMY Cube

RGB & CMY存在的问题

- 以RGB或CMY格式进行图像处理并不是最有效率的一种方法。因为在RGB或CMY格式下，对一个像素的每一步操作都需要涉及所有红、绿、蓝三个通道。这比其它色彩格式需要更多的存储能力和数据处理带宽。
- 为解决这一问题，许多广播标准，如欧洲的PAL和北美的NTSC电视系统，都采用亮度和色差视频信号。

HSL 颜色模型

- HSL颜色模型：色调 (hue)、饱和度 (saturation)和亮度 (luminance)
 - 亮度：在一个影像中所包含的光线数量，亮度显示色彩有多亮或多暗。



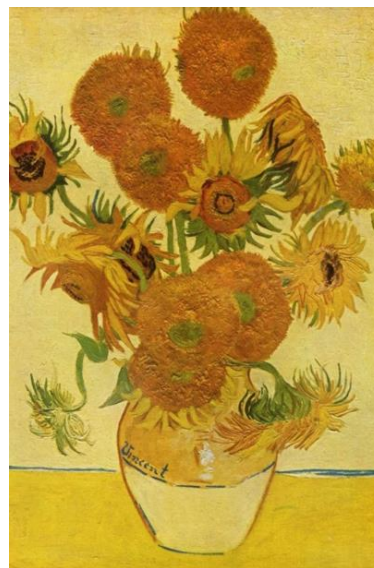
HSL 颜色模型

- HSL颜色模型：色调 (hue)、饱和度 (saturation)和亮度 (luminance)
 - 饱和度：是指色彩的纯度，一种颜色的饱和度越高，它就越鲜艳；反之，一种颜色的饱和度越低，它就越接近于灰色。

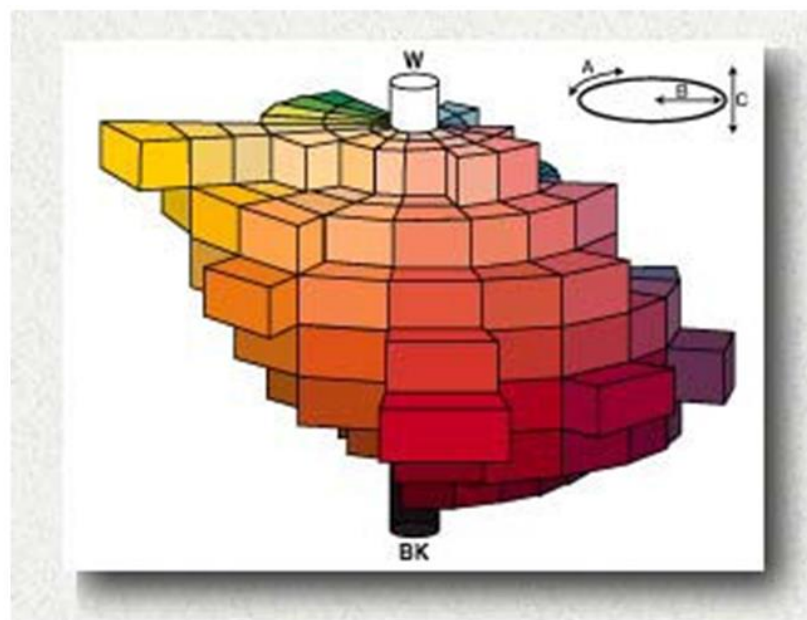
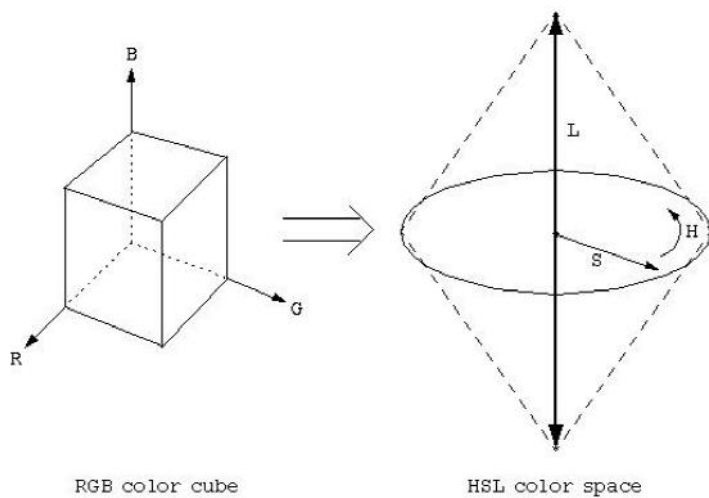


HSL 颜色模型

- HSL颜色模型：色调 (hue)、饱和度 (saturation)和亮度 (luminance)
 - 色调：对一张图像整体颜色的概况评价。一张图像虽然用了多种颜色，但是总体有一种倾向，或是偏蓝或是偏红，或是偏暖或是偏冷。



HSL 颜色模型

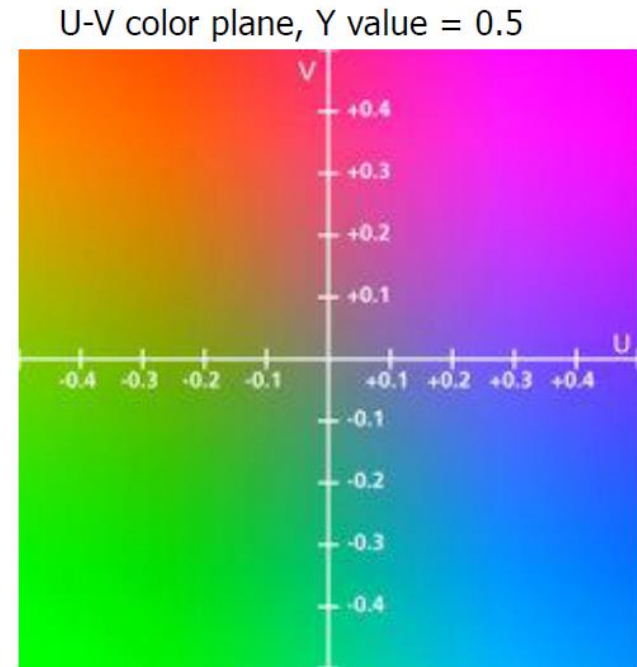


A: 色调
B: 亮度
C: 饱和度

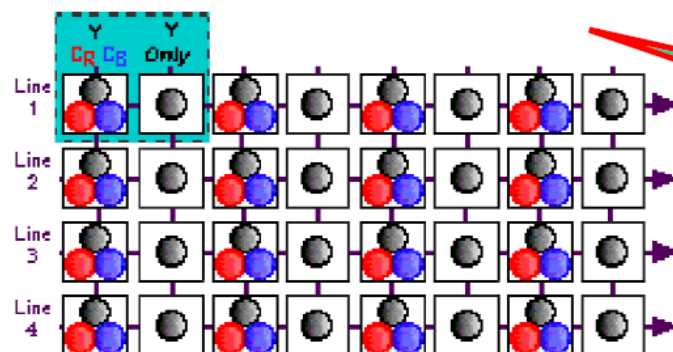
YUV 颜色模型

- YUV 颜色模型

- Y: 亮度;
- U: 色度, chrominance;
- V: 浓度, chroma。



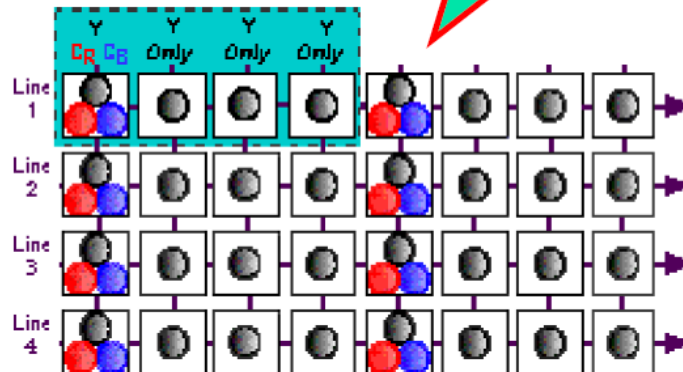
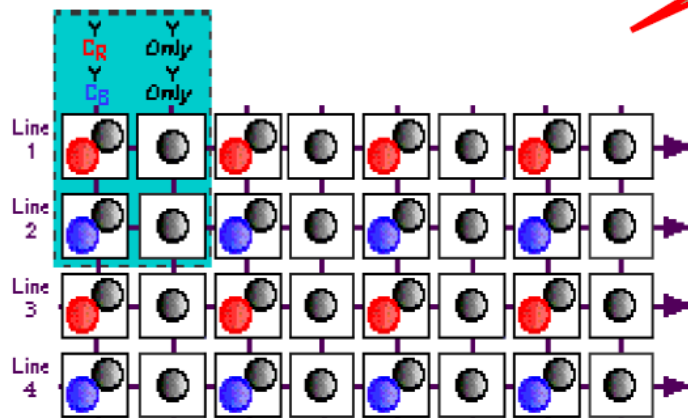
YUV 的不同格式



YUV422

YUV420

YUV411



四种颜色模型

- RGB相加颜色模型
- CMY相减颜色模型
- HSL颜色模型
- YUV颜色模型

颜色空间变换

- 从RGB到CMY

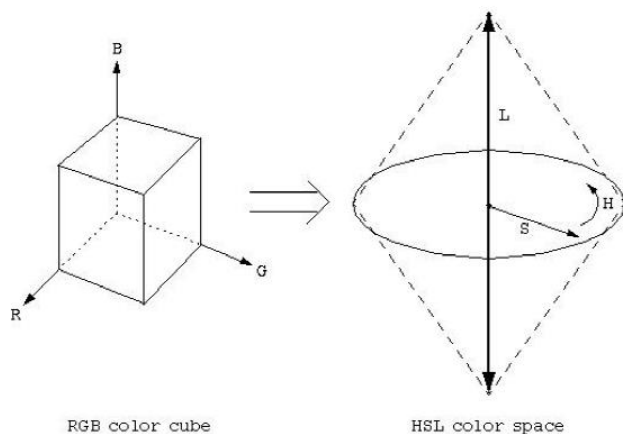
$$C = 255 - R$$

$$M = 255 - G$$

$$Y = 255 - B$$

颜色空间变换

● 从RGB到HSL



$$h = \begin{cases} 0^\circ & \text{if } \max = \min \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, & \text{if } \max = r \text{ and } g \geq b \\ 60^\circ \times \frac{g-b}{\max-\min} + 360^\circ, & \text{if } \max = r \text{ and } g < b \\ 60^\circ \times \frac{b-r}{\max-\min} + 120^\circ, & \text{if } \max = g \\ 60^\circ \times \frac{r-g}{\max-\min} + 240^\circ, & \text{if } \max = b \end{cases}$$

$$l = \frac{1}{2}(\max + \min)$$

$$s = \begin{cases} 0 & \text{if } l = 0 \text{ or } \max = \min \\ \frac{\max-\min}{\max+\min} = \frac{\max-\min}{2l}, & \text{if } 0 < l \leq \frac{1}{2} \\ \frac{\max-\min}{2-(\max+\min)} = \frac{\max-\min}{2-2l}, & \text{if } l > \frac{1}{2} \end{cases}$$

颜色空间变换

- 从RGB到YUV

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

$$U = -0.147R - 0.289G + 0.436B$$

$$V = 0.615R - 0.515G - 0.100B$$

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

真色彩 & 伪色彩

- **真彩色**是指在组成一幅彩色图像的几个像素值中，有R，G，B三个基色分量，每个基色分量直接决定显示设备的基色强度，这样产生的彩色称为真彩色。
- **伪彩色**的含义是，每个像素的颜色不是由每个基色分量的数值直接决定，而是把像素值当作彩色查找表(color look-up table, CLUT)的表项入口地址，去查找一个显示图像时使用的R，G，B强度值。

关于彩色习惯说法

- 8 位色可显示 $2^8=256$ 种颜色，称之为**伪彩色**
- 16位色可显示 $2^{16}=65536$ 种颜色，称之为**增强色**
- 24位色可显示 $2^{24}=16800000$ 种颜色，称之为**真彩色**

为什么要使用伪彩色

- 65536色以下的显示方式中一般采用伪彩色。标准的调色板是在256K色谱（即25万6千种色彩）中按色调均匀地选取16种或256种色彩。
- 一般应用中，有的图像往往偏向于某一种或几种色调，此时如果采用标准调色板，则色彩失真较多。因此，同一幅图像，采用不同的调色板显示可能会出现不同的色彩效果。

矢量图 & 位图

- 矢量图（vector -based image）是用一系列计算机指令来描述和记录一幅图，这幅图可分解为一系列子图如点、线、面等。
- 位图（bit-mapped image）是用像素点来描述或映射的图，也即位映射图。位图在内存中也就是一组计算机内存地址位（bit）组成，这些位定义图像中每个像素点的颜色和亮度。

矢量图 & 位图

- 矢量图：不受分辨率的影响。可以任意放大或缩小图形而不会影响出图的清晰度，可以按最高分辨率显示到输出设备上。难以表现色彩层次丰富的逼真图像效果！
- 位图：缩放和旋转容易失真，同时文件容量较大！但只要有足够多的不同色彩的像素，就可以制作出色彩丰富的图象，逼真地表现自然界的景象。

图像 & 图形

- 图形是指用计算机绘制工具绘制的画面，包括直线、曲线，圆/圆弧，方框等成分。图形一般按各个成分的参数形式存储，可以对各个成分进行移动、缩放、旋转和扭曲等变换，可以在绘图仪上将各个成分输出。
- 图像是由输入设备捕捉的实际场景或以数字化形式存储的任意画面。图像可以用位图或矢量图形式存储。

最普及的图像文件格式 **BMP**

- 位图文件(Bitmap-File, BMP)格式是Windows采用的图像文件存储格式，在Windows环境下运行的所有图像处理软件都支持这种格式。
- 位图文件的特点是包含的图像信息较丰富，几乎不进行压缩，但由此导致了它与生俱来的缺点--占用磁盘空间过大。所以，BMP在单机上比较流行。现在BMP（Window位图）图像主要被用在PC机运行Window时的墙纸。

和扫描仪一起出现的 **TIFF**

- TIFF文件：TIFF（Tag Image File Format）文件是由Aldus和Microsoft公司为扫描仪和桌上出版系统研制开发的一种较为通用的图像文件格式，在苹果公司的Mac计算机中广泛应用。
- TIFF格式灵活易变，它又定义了四类不同的格式：TIFF-B适用于二值图像；TIFF-G适用于黑白灰度图像；TIFF-P适用于带调色板的彩色图像；TIFF-R适用于RGB真彩图像。TIFF支持多种编码方法，其中包括RGB无压缩、RLE压缩及JPEG压缩等。

网络流行的格式之一 GIF

- GIF文件：GIF（Graphics Interchange Format）是CompuServe公司在1987年开发的图像文件格式。GIF文件的数据是经过压缩的，它采用了可变长度等压缩算法。
- GIF文件的图像深度从1 bit到8 bit，也即GIF最多支持256种色彩的图像。GIF格式的另一个特点是其在一个GIF文件中可以存多幅彩色图像，如果把存于一个文件中的多幅图像数据逐幅读出并显示到屏幕上，就可构成一种最简单的动画。

网络流行的格式之二 JPEG

- JPEG文件：JPEG（Joint Photographic Experts Group）是由CCITT（国际电报电话咨询委员会）和ISO（国际标准化组织）联合组成的一个图像专家组。该专家组制定的第一个压缩静态数字图像的国际标准，其标准名称为“连续色调静态图像的数字压缩和编码（Digital Compression and Coding of Continuous -tone Still Image）”，简称为JPEG算法。采用JPEG压缩编码算法压缩的图像，其压缩比约为5:1至50:1，甚至更高。

网络流行的格式之三 PNG

- PNG是可携式网络图像Portable Network Graphics 这三个单词第一个字母的的缩写。PNG文件格式是一种新兴的网络图像格式，结合了GIF和JPEG格式的一些优点，允许连续读出和写入图像数据并使用无损压缩。在Linux世界里比较常见，是一种十分有前途的文件格式。
- 1994年，Unysis公司宣布gif拥有专利的压缩方法，要求开发gif软件的作者须缴交一定费用，促使免费的png图像格式的诞生。

图像文件的格式转换工具

- 格式工厂
- PhotoShop
- ACDSee
- CorelDRAW 12
- 光影魔术手
- 美图秀秀
- 画图

