

第4章 彩色数字图像基础





第4章 彩色数字图像基础目录

4.1 颜色简介

4.1.1 视觉系统对颜色的感知

4.2.2 图像的颜色模型

4.2 图像简介

4.2.1 图像基本属性

4.2.2 图像的种类

4.2.3 伽玛(γ)校正

4.3 JPEG压缩编码

4.3.1 算法概要

4.3.2 主要步骤

4.3.3 JPEG压缩和编码
举例

4.4 常用图像文件格式

BMP、GIF、JPEG、PNG



第4章 彩色数字图像基础目录

4.1 颜色简介

4.1.1 视觉系统对颜色的感知

4.2.2 图像的颜色模型

4.2 图像简介

4.2.1 图像基本属性

4.2.2 图像的种类

4.2.3 伽玛(γ)校正

4.3 JPEG压缩编码

4.3.1 算法概要

4.3.2 主要步骤

4.3.3 JPEG压缩和编码
举例

4.4 常用图像文件格式

BMP、GIF、JPEG、PNG

4.1.1 什么是颜色？

- 颜色是视觉系统对可见光的感知结果。
- 可见光是波长在380~780 nm之间的电磁波
 - 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。

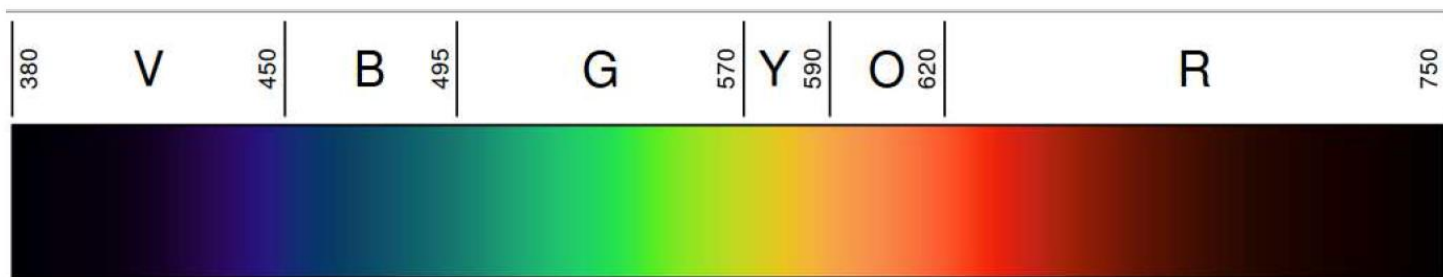


Figure 4.5: Visible light spectrum corresponds to the range of electromagnetic waves that have wavelengths between 400nm and 700nm. (Figure by David Eccles for Wikipedia.)

常见光源的光谱能量分布

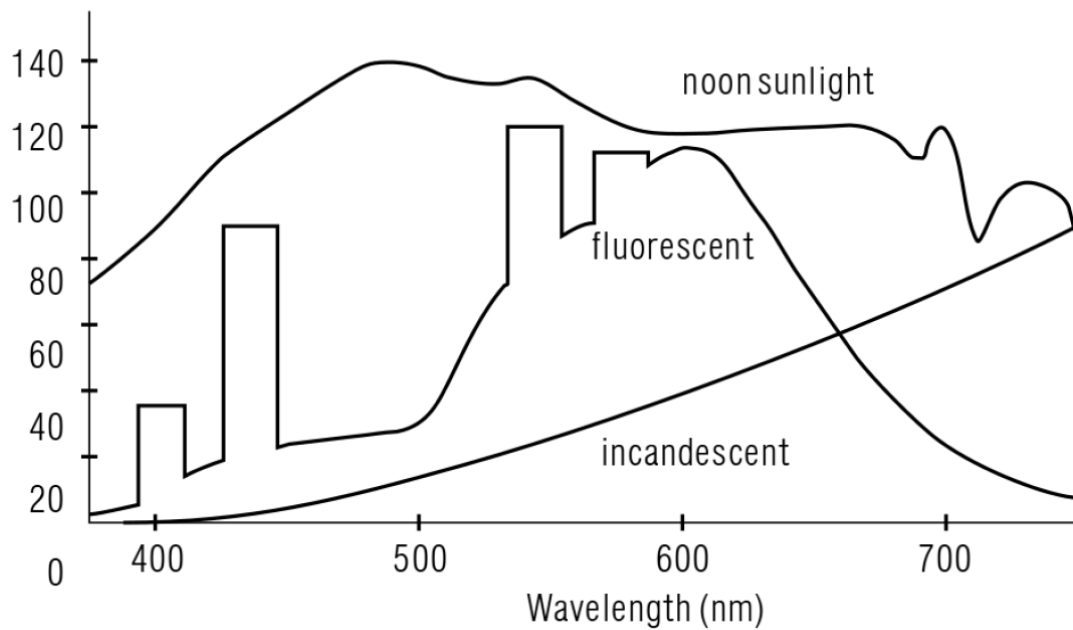


Figure 4.6: The *spectral power distribution* for some common light sources. (Figure from [226]).

频率反射函数

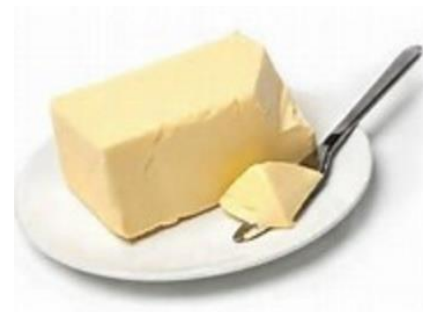
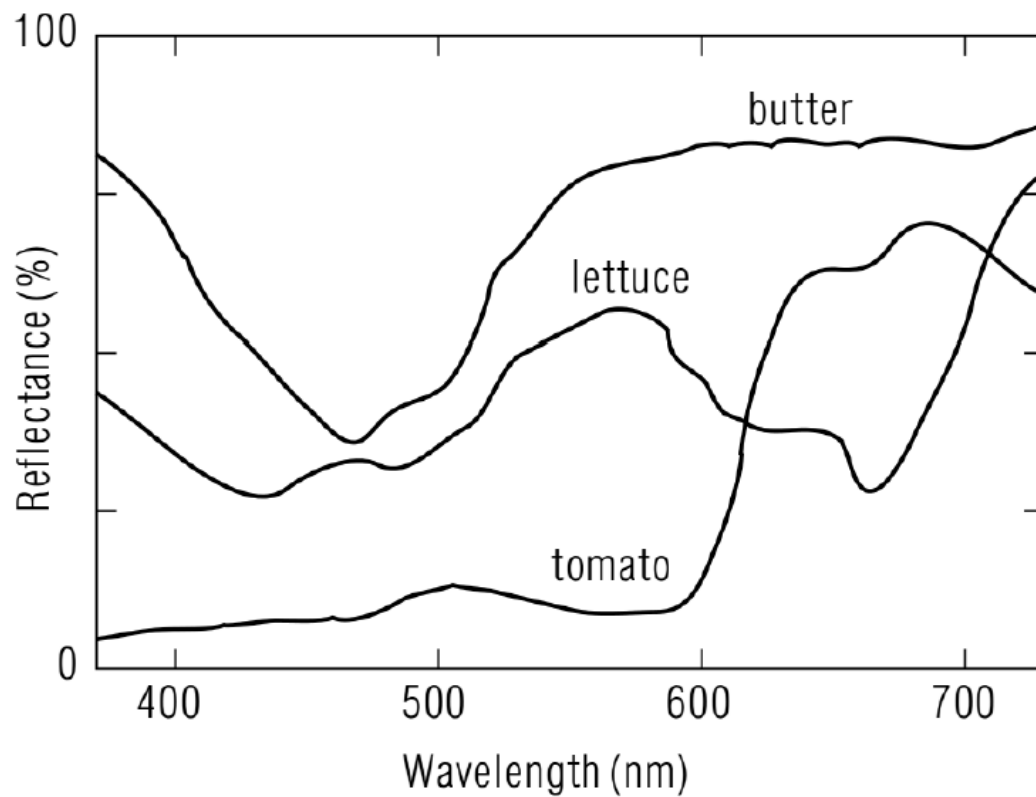


Figure 4.7: The *spectral reflection function* of some common familiar materials (Figure from [226]).

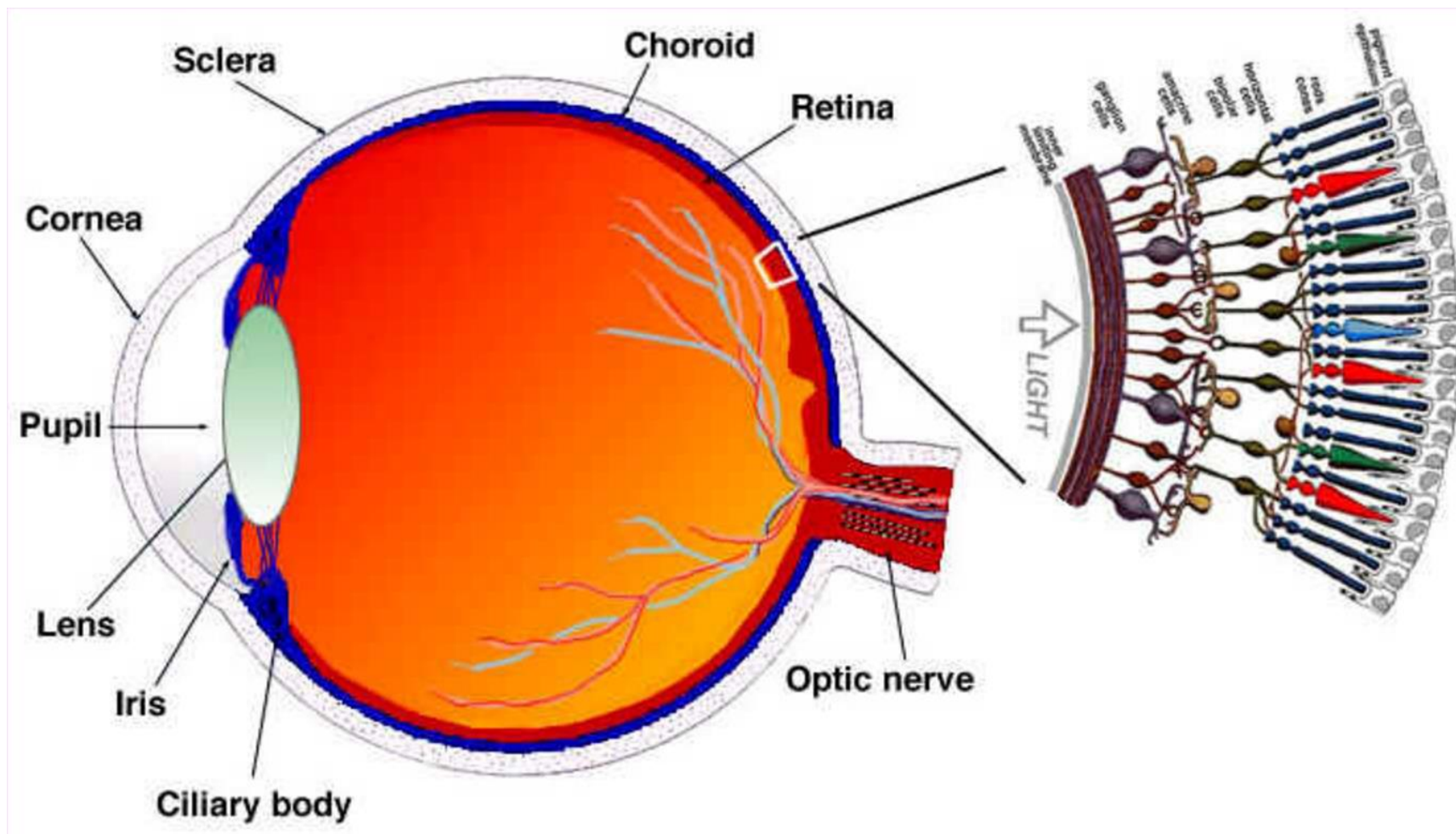


4.1.1 视觉系统对颜色的感知

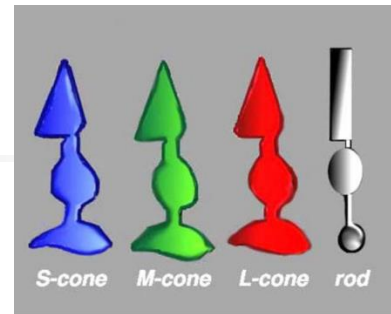
■ 颜色是什么？

- 视觉系统对可见光的感知结果
- 颜色只存在于眼睛和大脑
 - 人的视网膜有对红、绿、蓝颜色敏感程度不同的三种锥体细胞
 - 杆状体细胞在光功率极低的条件下才起作用

4.1.1 视觉系统对颜色的感知



杆状细胞 vs. 锥体细胞



杆状细胞

外观呈杆状

每只眼大约 10^9 个细胞，散布在视网膜上，用于周围视觉

敏感度高-可用于夜视

1种类型，不能分辨色彩

多个杆状细胞与一个双极细胞相连，敏锐度低

椎体细胞

外观呈锥状

每只眼大约 10^6 个细胞，主要分布在中央凹，只能在视网膜中心位置形成图像

敏感度低-需要明亮的光线，通常在白天作用

3种类型，复杂彩色视觉

每个杆状细胞通常与一个双极细胞相连，敏锐度高（用于分辨细节）

4.1.1 视觉系统对颜色的感知

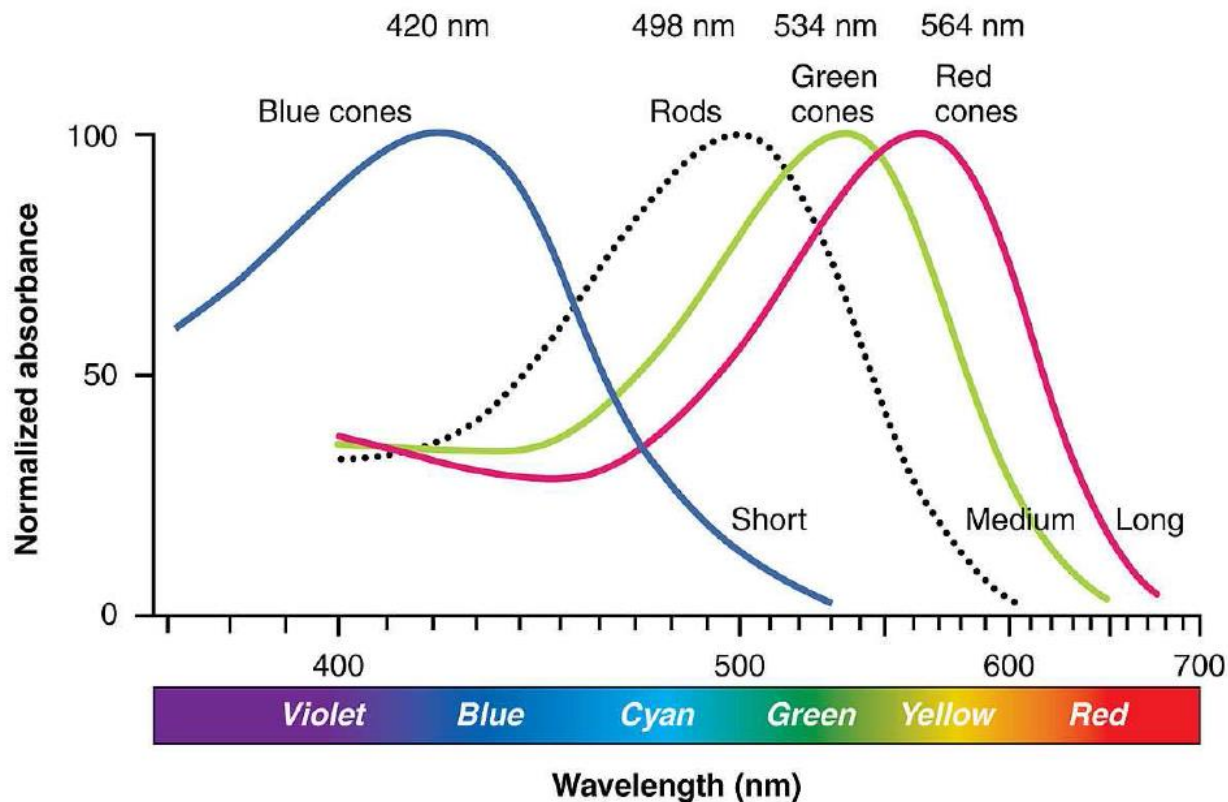


Figure 5.3: The sensitivity of rods and cones as a function of wavelength [20].
(Figure adapted by OpenStax College.)

- 红光 → 刺激红色椎体细胞
- 黄光 → 同等程度刺激红色和绿色椎体细胞
- 青色光 → 同等程度刺激蓝色和绿色椎体细胞
- 白光 → 同等程度刺激三类椎体细胞



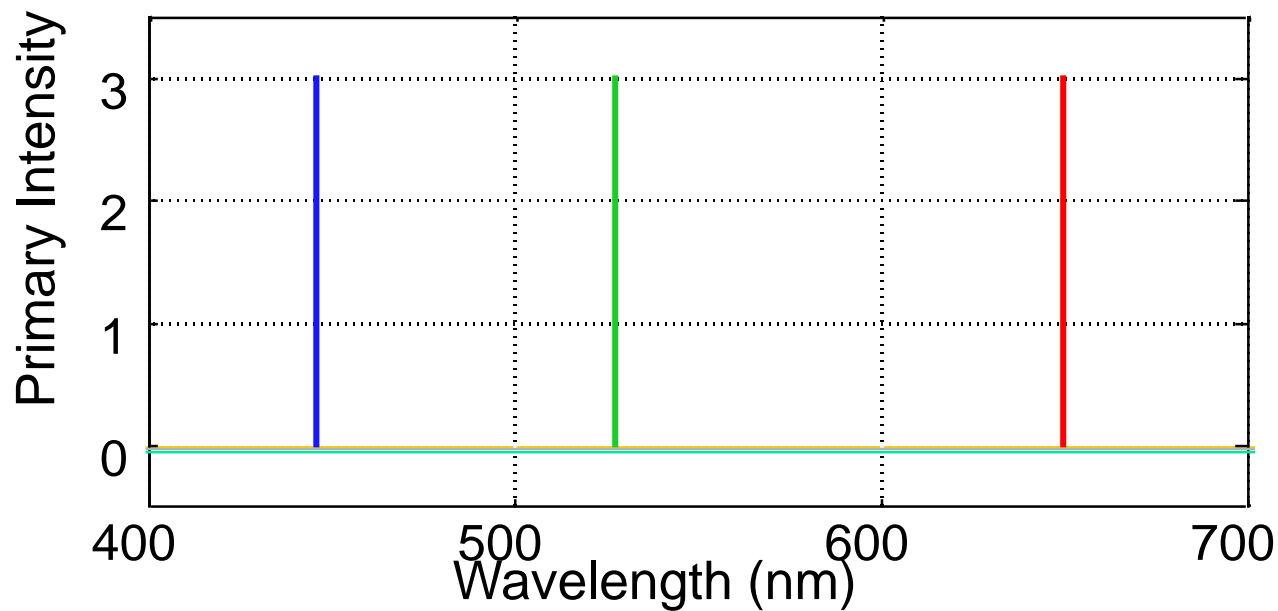
4.1.1 视觉系统对颜色的感知

■ 视觉系统对颜色感知的特性

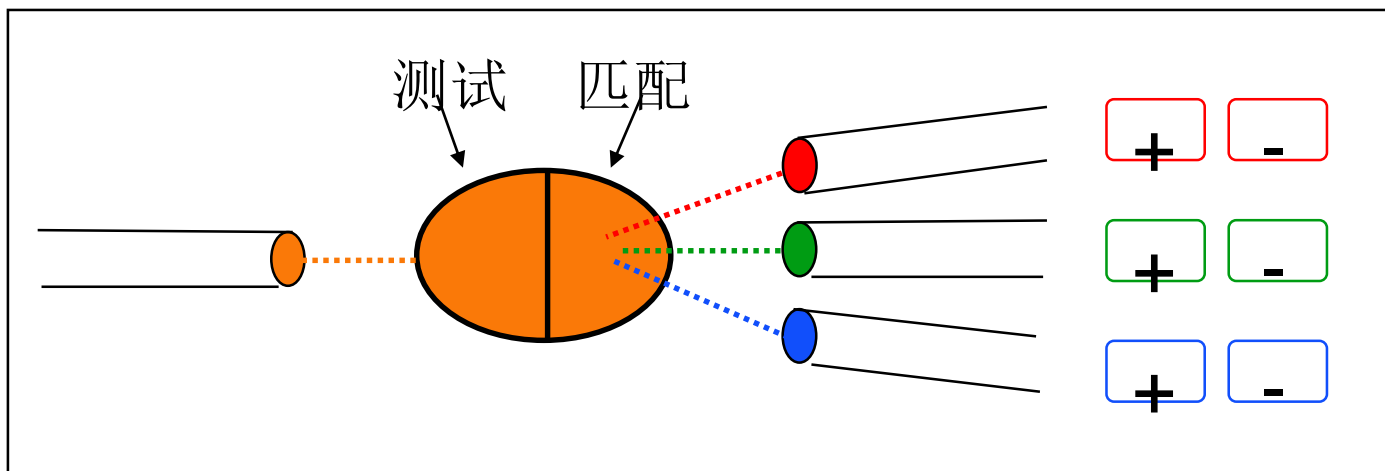
- 眼睛本质上是一个照相机
 - 人的视网膜(human retina)通过神经元感知外部世界的颜色, 每个神经元是一个对颜色敏感的锥体(cone)
- 红、绿和蓝三种锥体细胞对不同频率的光的感知程度不同, 对不同亮度的感知程度也不同
 - 可用数字图像处理技术来降低表示图像的数据量而不使人感到图像质量有明显下降。

RGB 色彩空间

- 选取红、绿、蓝三种基本颜色作为基础色

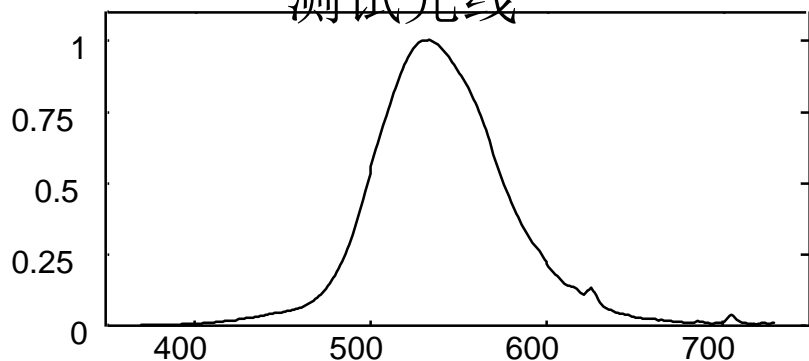


颜色匹配实验



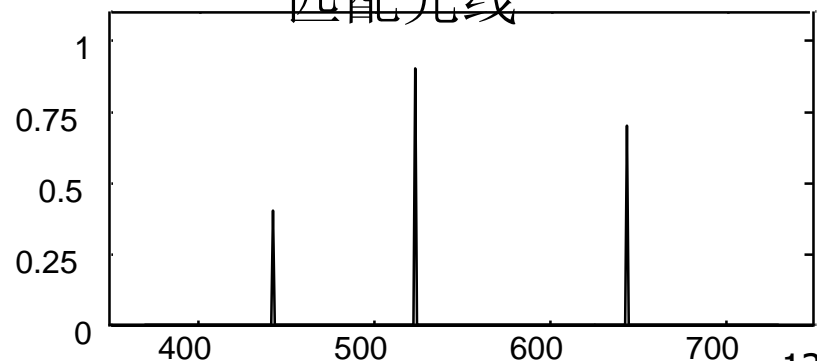
- 设置三种基本颜色的光与测试光线匹配

测试光线



\approx

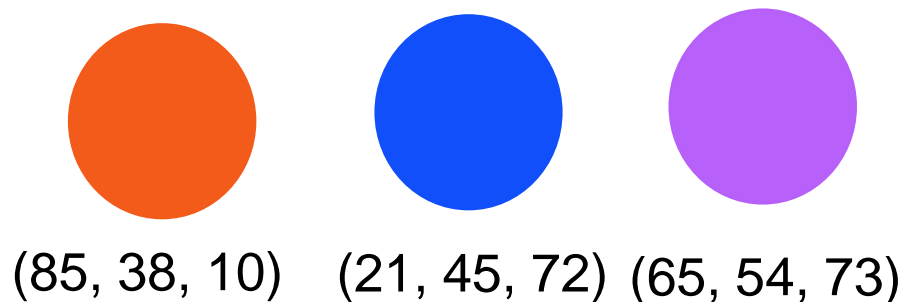
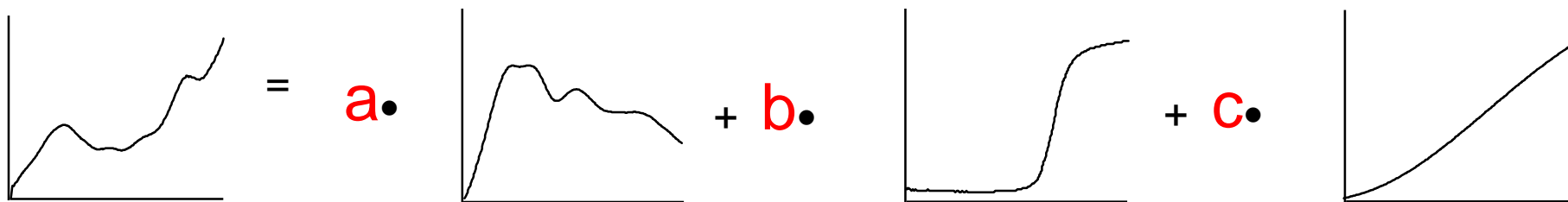
匹配光线



线性色彩空间

■ 三维色彩空间内的颜色可被描述成3种基本颜色（基色）的线性组合

- R, G, B的数值不同，混合得到的颜色就不同，即光波的波长不同





4.1.1 色盲

- 色盲是指缺乏或完全没有辨别色彩的能力。
- 每一种感光色素主要对一种原色光产生兴奋，而对其余两种原色光产生程度不等的反应。如果某一种色素缺乏，则会产生对此种颜色的感觉障碍，表现为色盲或色弱（辨色力弱）。



正常



缺少R-cone,红色盲



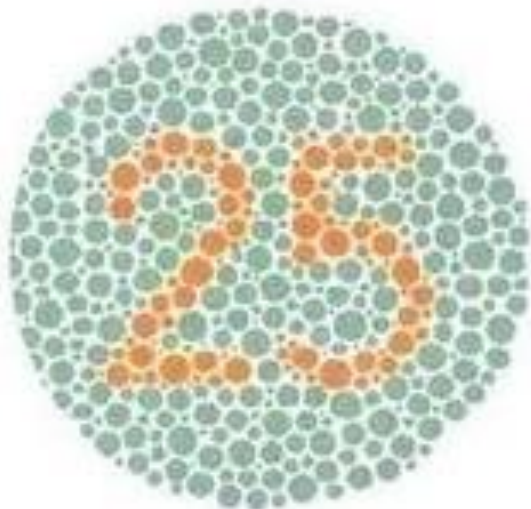
缺少G-cone, 绿色盲



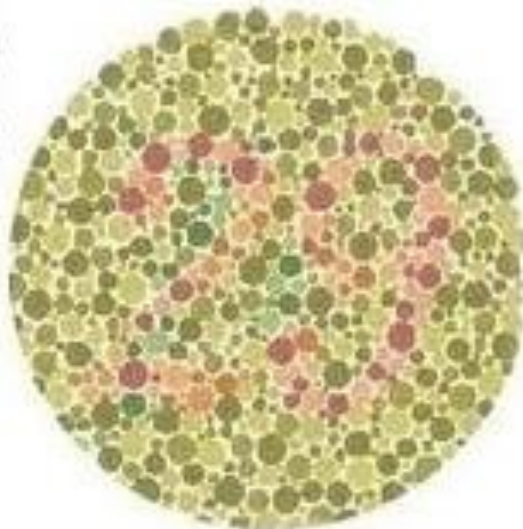
缺少B-cone, 蓝黄色盲

色盲测试

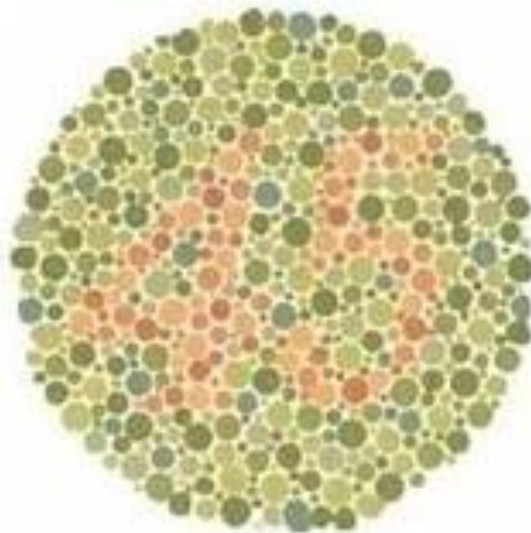
1



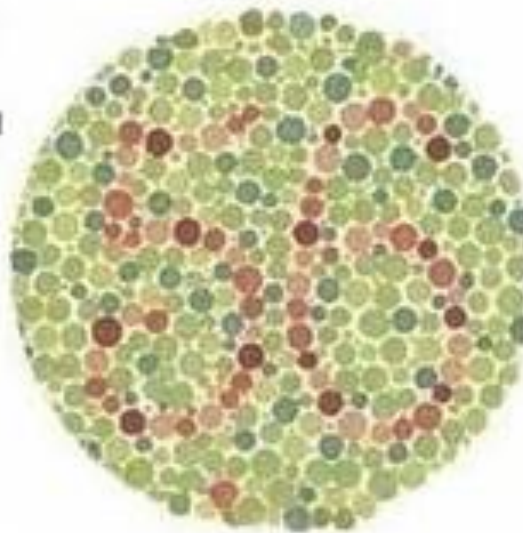
2



3



4



这是什么颜色？



White with gold lace?
Or
Blue with black lace?



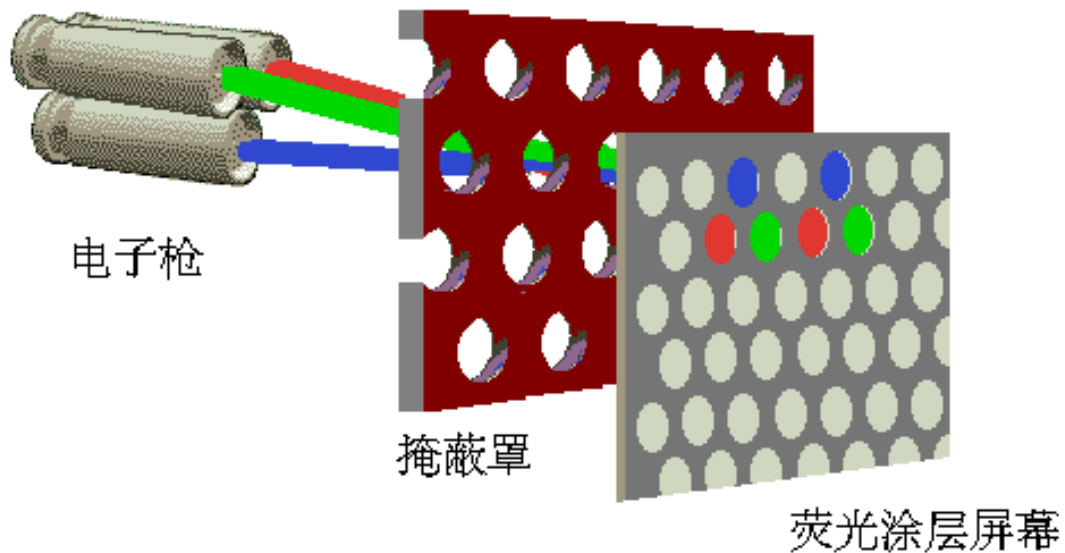
4.1.2 图像的颜色模型

■ RGB相加混色模型(additive color mixture)

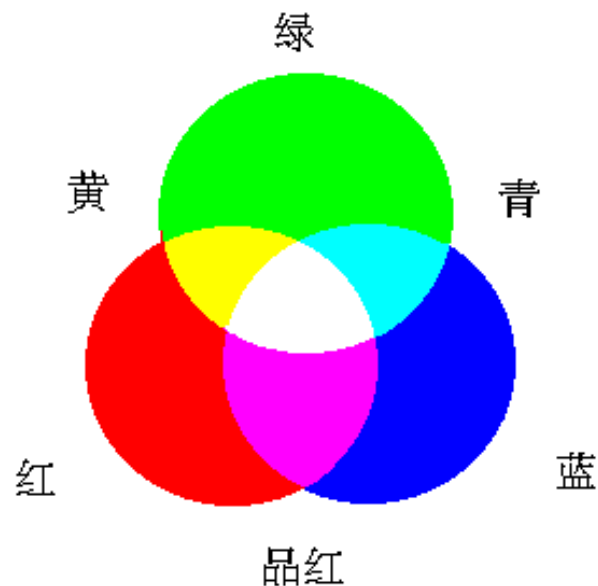
- 一个能发出光波的物体称为有源物体，它的颜色由该物体发出的光波决定
- 显示彩色图像用RGB相加混色模型
 - 组合红、绿和蓝光波来产生特定颜色
 - 相加混色是计算机应用中定义颜色的基本方法
- 任何一种颜色都可用三种基本颜色按不同的比例混合得到

颜色 = R (红的百分比) + G (绿的百分比) + B (蓝的百分比)

4.1.2 图像的颜色模型



彩色显像产生颜色的原理



RGB相加混色

4.1.2 图像的颜色模型

➤ 彩色图像

- 一幅彩色图像可看成是由多个点组成, 单个点称为像素(pixel), 每个像素都有一个值, 称为像素值, 它表示特定颜色的强度
- 一个像素值通常用R, G, B三个分量表示。如果每个像素的每个颜色分量“1”和“0”表示, 即每种颜色的强度是100%或0%, 每个像素显示的颜色是8种颜色之一



R	G	B	颜色
0	0	0	黑
0	0	1	蓝
0	1	0	绿
0	1	1	青
1	0	0	红
1	0	1	品红
1	1	0	黄
1	1	1	白



RGB 图像

10 111	200 36	128 12	126 36	200 14	12 36
17 36	200 111	36 14	14 126	36 17	72 111
12 17	36 111	126 200	17 36	111 14	200 36
14 14	17 36	126 12	12 36	126 17	17 111
126 17	200 111	111 14	14 126	36 17	72 126
36 12	72 126	12 200	17 36	106 14	155 36

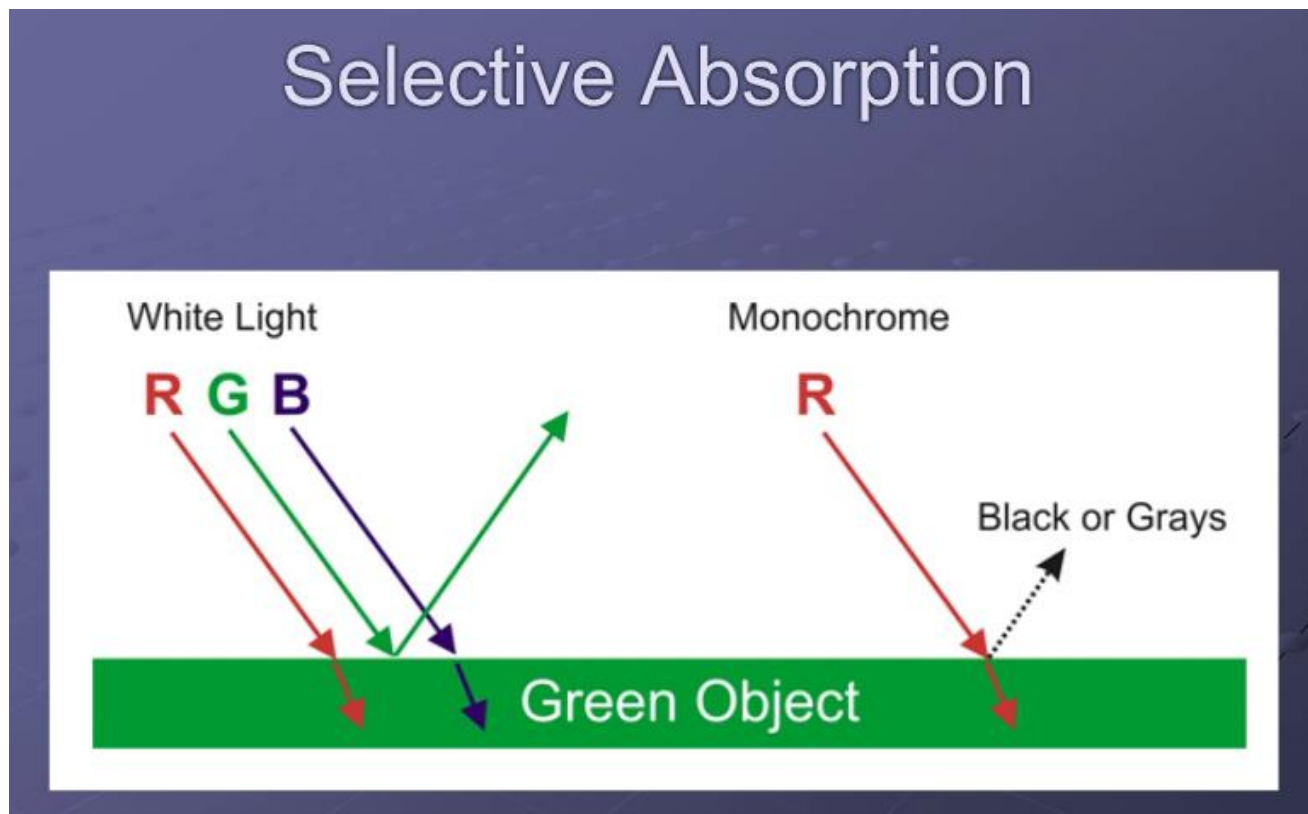


4.1.2 图像的颜色模型

■ CMY相减混色模型

- 一个不发光波的物体称为无源物体，它的颜色由该物体吸收或者反射哪些光波决定
 - 用彩色墨水或颜料进行混合，绘制的图画是一种无源物体，用这种方法生成的颜色称为相减色
- 打印彩色图像用CMY相减混色模型
 - 用三种基本颜色即青色(cyan)、品红(magenta)和黄色(yellow)的颜料按一定比例混合得到颜色
 - 从理论上说，任何一种颜色都可以用青色(cyan)、品红(magenta)和黄色(yellow)混合得到
- 相减混色减少了为视觉系统识别颜色所需要的反射光

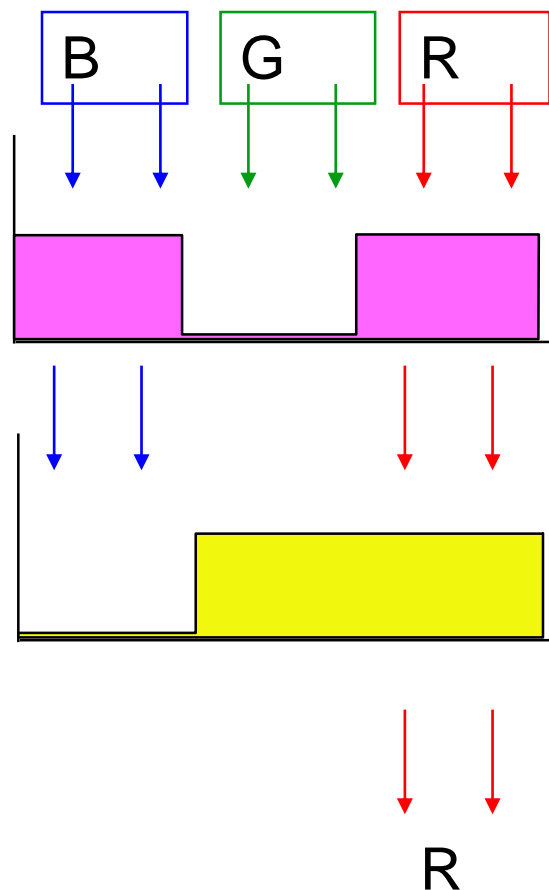
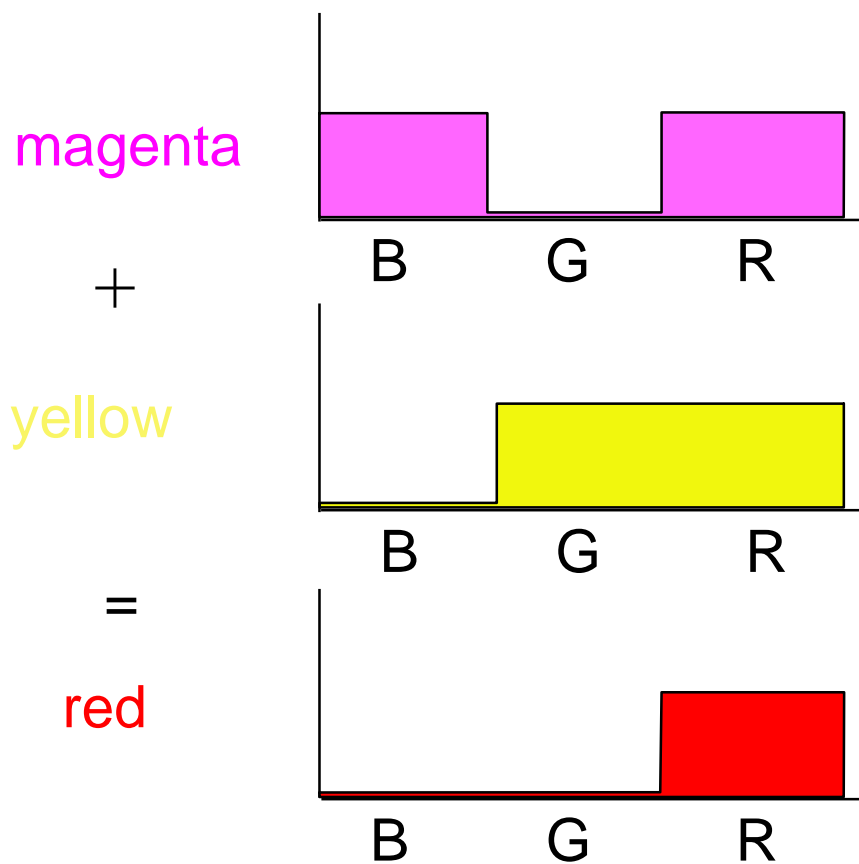
4.1.2 图像的颜色模型



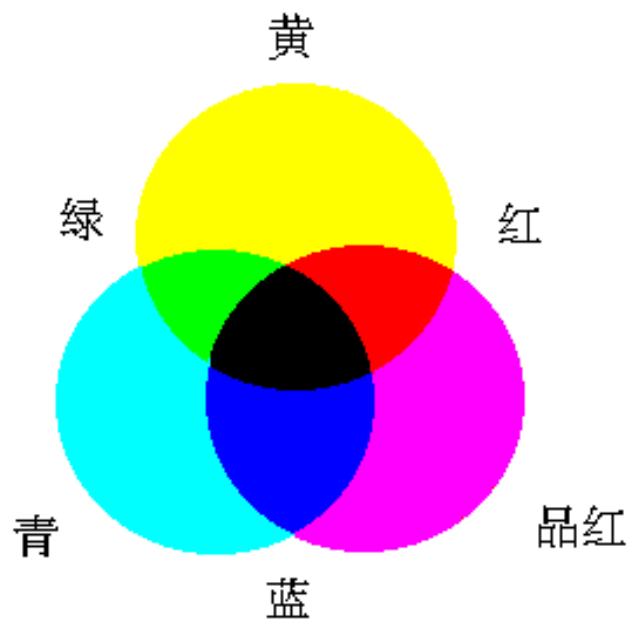
CMY相减混色模型示例

■ CMY相减混色模型

品红 + 黄色 = 红色



4.1.2 图像的颜色模型

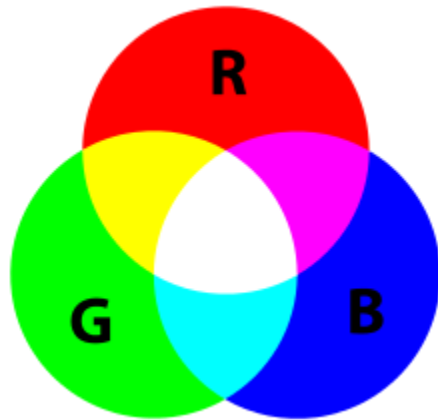


相减混色

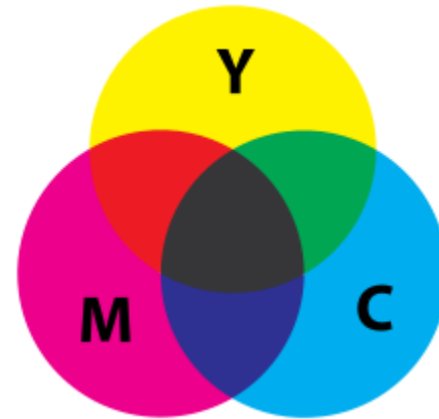
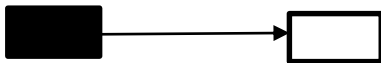
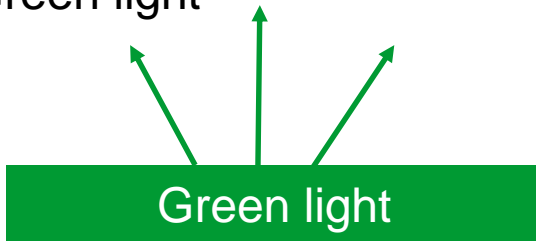
相减色

C(青色)	M(品红)	Y(黄色)	相减色
0	0	0	白
0	0	1	黄
0	1	0	品红
0	1	1	红
1	0	0	青
1	0	1	绿
1	1	0	蓝
1	1	1	黑

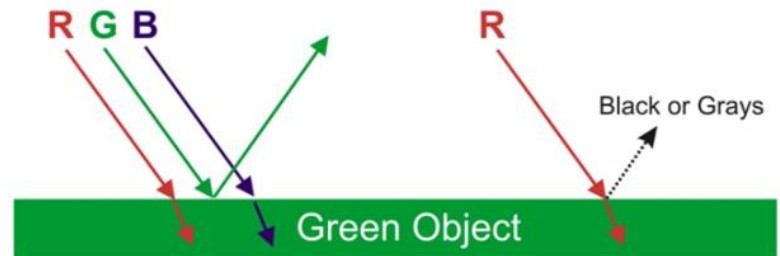
相加混色 vs. 相减混色



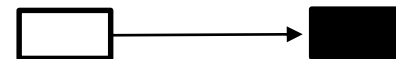
Green light



White Light



Monochrome



4.1.2 图像的颜色模型

■ 相加色与相减色是互补色

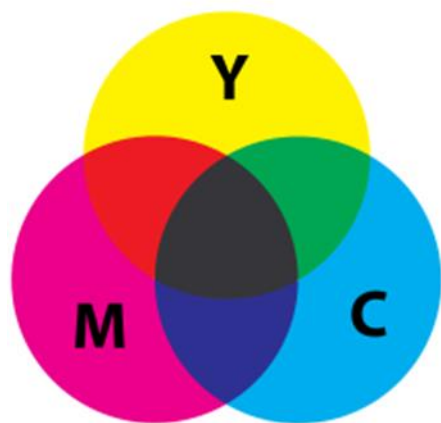
- 相加混色和相减混色之间成对出现互补色。
- 利用它们之间的关系，可把显示的颜色转换成打印的颜色
- 在RGB中的颜色值为1的地方，在CMY对应的位置上，其颜色值为0。例如，RGB为0：1：0时，对应CMY为1：0：1

相加色与相减色的关系

相加混色(RGB)	相减混色(CMY)	生成的颜色
000	111	黑
001	110	蓝
010	101	绿
011	100	青
100	011	红
101	010	品红
110	001	黄
111	000	白

CMYK

- 由于彩色墨水和颜料的化学特性,等量三基色得到的黑色不是真正的黑色,因此在印刷术中常加一种真正的黑色(Black ink)

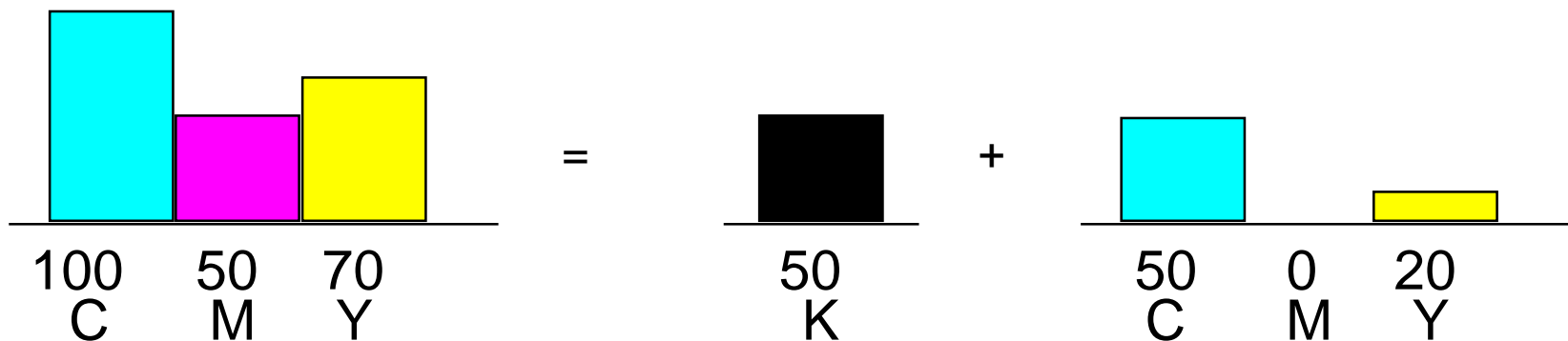


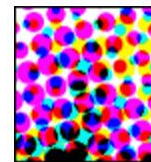
Color printing typically uses ink of four colors: cyan, magenta, yellow and key (black).

CMY + 黑色

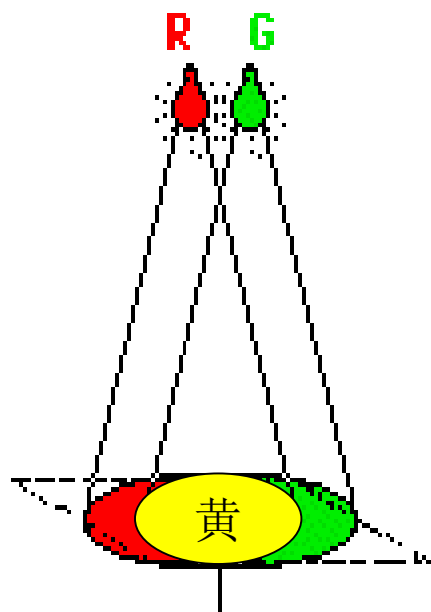
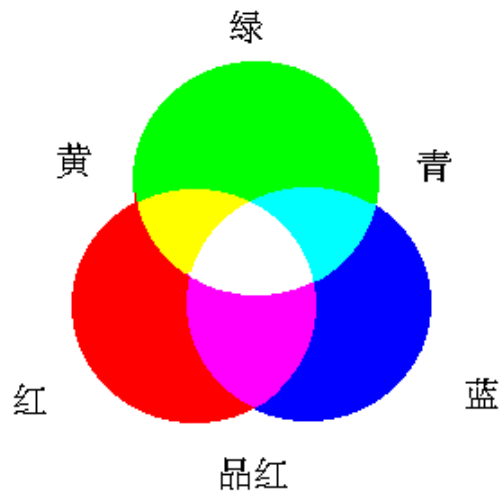
$$C + M + Y = K \text{ (黑色)}$$

- 用三种颜料表示黑色很昂贵
- $C+M+Y$ = 深褐色，而非黑色
- 用黑色代替 $C+M+Y$ 很锐、对比度更强

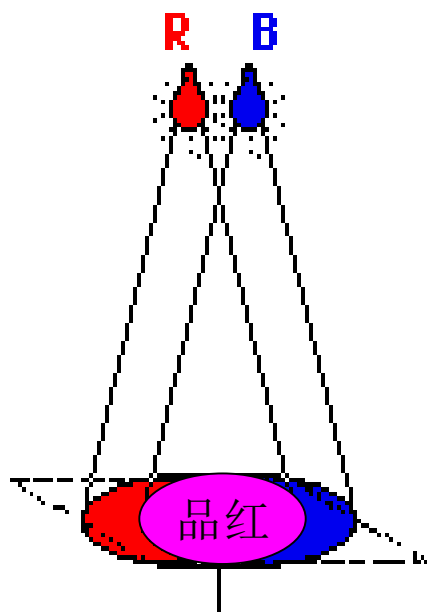




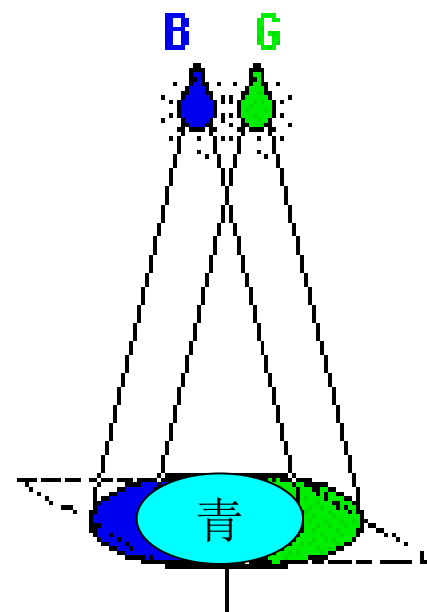
Quiz



Paper Appears ????



Paper Appears ????



Paper Appears ????



第4章 彩色数字图像基础目录

4.1 颜色简介

4.1.1 视觉系统对颜色的感知

4.2.2 图像的颜色模型

4.2 图像简介

4.2.1 图像基本属性

4.2.2 图像的种类

4.2.3 伽玛(γ)校正

4.3 JPEG压缩编码

4.3.1 算法概要

4.3.2 主要步骤

4.3.3 JPEG压缩和编码
举例

4.4 常用图像文件格式

BMP、GIF、JPEG、PNG

4.2.1 图像的三个基本属性

■ 图像分辨率(resolution)

➤ 屏幕分辨率(screen resolution)



iPhone X

超视网膜高清显示屏

5.8 英寸 (对角线)

OLED 全面屏

Multi-Touch 显示屏

HDR 显示^[1]

2436 x 1125 像素
分辨率, 458 ppi

1200 万像素广角及长焦双镜头摄像头

- 衡量显示设备再现图像时所能达到的精细程度的度量方法。也称显示分辨率

- 用“水平像素数×垂直像素数”表示, 如640×480表示显示屏分成480行, 每行显示640个像素, 整个显示屏含有307200个显像点

- ◆ 常见的屏幕分辨率: 640×480, 800×600, 1024×768, 1280×1024, 1920×1080, 3840×2160

- 传统电视屏幕的宽高比为4:3; 高清晰度电视屏幕的宽高比为16:9

➤ 图像分辨率(image resolution)

- 图像精细程度的度量方法。对同样尺寸的一幅图, 如果像素数目越多, 则说明图像的分辨率越高, 看起来就越逼真。相反, 图像显得越粗糙

- 图像分辨率也称空间分辨率(spatial resolution)和像素分辨率(pixel resolution)

4.2.1 图像的三个基本属性

➤ 在图像显示应用中的图像分辨率表示法

- (1) 物理尺寸：每毫米线数(或行数)
- (2) 行列像素：像素/行 \times 行/幅，如640像素/行 \times 480行/幅
- (3) 像素总数：如数码相机上标的500万像素
- (4) 单位长度上的像素：如像素每英寸(pixels per inch, PPI)
- (5) 线对(line pair)数：以黑白相邻的两条线为一对，如5对线

➤ 在图像数字化和打印应用中的图像分辨率表示法

- 通常用多少点每英寸(dots per inch, DPI)表示，如300 DPI
- 分辨率越高，图像质量就越高，像素就越多，要求存储容量

佳能 (Canon) E568 彩色喷墨一体机



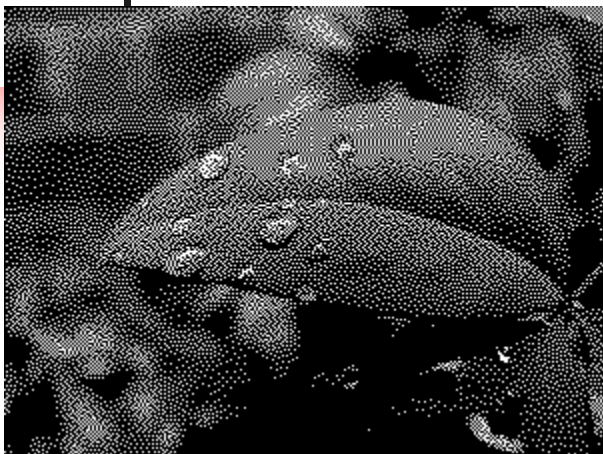
分辨率	4800 (水平)*1 x 1200 (垂直) dpi
打印速度	文档： 彩色：约 5.7 ipm文档：

扫描类型	平板式
扫描分辨率	光学分辨率：1200 x 2400dpi插值分辨率：25 - 19200
扫描速度	线性扫描速度：灰度：1.2 毫秒/线 (300dpi)彩色：3.5
色彩深度	灰度：16位/8位 彩色：48位/24位 (RGB 各 16位/8位)
其他功能	最大扫描尺寸：A4, LTR (216 × 297毫米)
扫描方式	CIS

4.2.1 图像的三个基本属性

■ 像素深度与阿尔法(α)通道

- 存储每个像素所用的位数
 - 例如，用R，G，B三个分量表示的彩色图像，若每个分量用8位表示，那么一个像素共用24位表示，就说像素深度为24位
- 像素深度决定彩色图像的每个像素可能有的颜色数，或者确定灰度图像的每个像素可能有的灰度级数
 - 例如，像素深度为24位时，每个像素可以是 $2^{24}=16\ 777\ 216$ 种颜色中的一种
- 像素深度越深，表达的颜色数目就越多，所占用的存储空间也越大。相反，如果像素深度太浅，则影响图像的质量，图像看起来让人觉得很粗糙和很不自然
- 由于受到设备和人眼分辨率的限制，不一定要追求特别深的像素深度



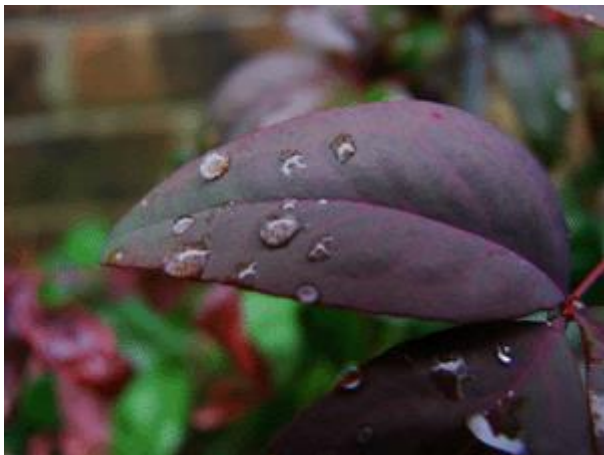
1 bit (2 colors) with dithering



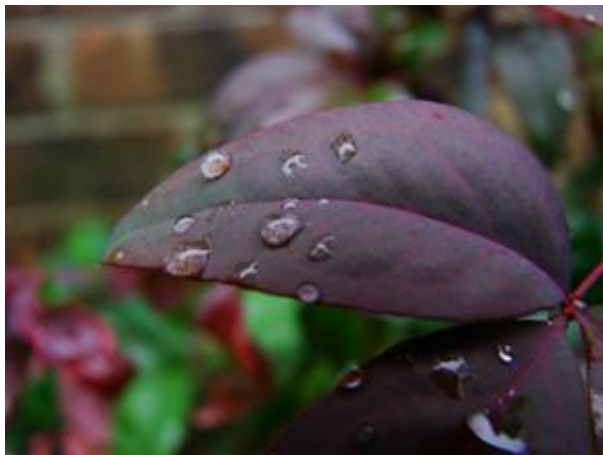
2 bits (4 colors)



4 bits (16 colors)



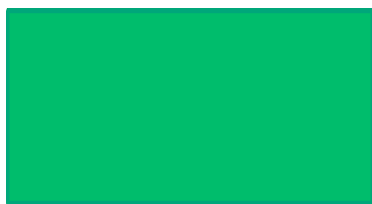
8 bits (256 colors)



24 bits (16,777,216 colors, "truecolor")

4.2.1 图像的三个基本属性

■ 阿尔法(α)通道

 像素用32位表示的图像表示法中的高8位，
低8位是颜色通道，红色、绿色和蓝色分量各占
8位的通道

- 用于表示像素在对象中的透明度
 - 例如，用两幅图像A和B混合成一幅新图像，新图像(New)的像素为： $\text{New pixel} = (\alpha)(\text{pixel A color}) + (\alpha)(\text{pixel B color})$
 - 又如，一个像素(A, R, G, B)的四个分量都用规一化的数值表示时，
 - ◆ 像素值为(1, 1, 0, 0)时显示红色，表示红色强度为1
 - ◆ 像素值为(0.5, 1, 0, 0)时，使用 α 通道中的预乘数0.5与R, G, B相乘，其结果为(0.5,0.5,0,0)，表示红色强度为0.5

4.2.1 图像的三个基本属性

■ 真彩色、伪彩色与直接色

➤ 真彩色(true color)

- 每个像素的颜色值用红(R)、绿(G)和蓝(B)表示的颜色
- 通常用24位表示，其颜色数 $2^{24}=16\,777\,216$ 种。也称24位颜色(24-bit color)或全彩色(full color)

➤ 伪彩色(pseudo color)

- 不是物体固有的而是人为的颜色
- 将像素值当作彩色查找表(color look-up table, CLUT)的表项入口地址，查找显示图像时要使用的R, G, B值，用查找出的R, G, B值产生的彩色
- 使用查找得到的R, G, B数值显示的彩色是真的，但不是图像本身真正的颜色，它没有完全反映原图的颜色

➤ 直接色(direct color)

- 每个像素值由R, G, B分量构成，每个分量作为单独的索引值对它做变换，用变换后的R, G, B值产生的颜色

4.2.2 图像的种类

■ 矢量图(vector graphics)

➤ 根据数学规则描述而生成的图

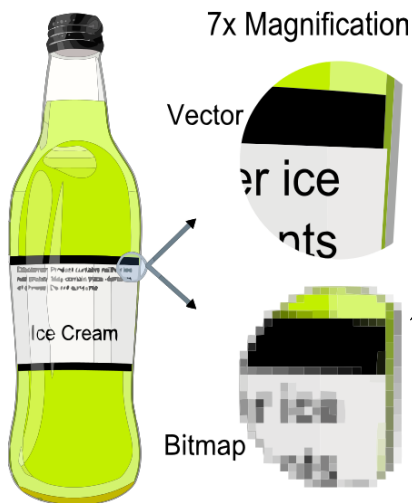
- 一幅图用数学描述的点、线、弧、曲线、多边形和其他几何实体和几何位置来表示，创建的图是对象的集合而不是点或像素模式的图
- 绘制和显示矢量图的软件通常称为绘图程序(draw programs)；存放矢量图的存储格式称为矢量图格式；存储的数据主要是绘制图形的数学描述

优点

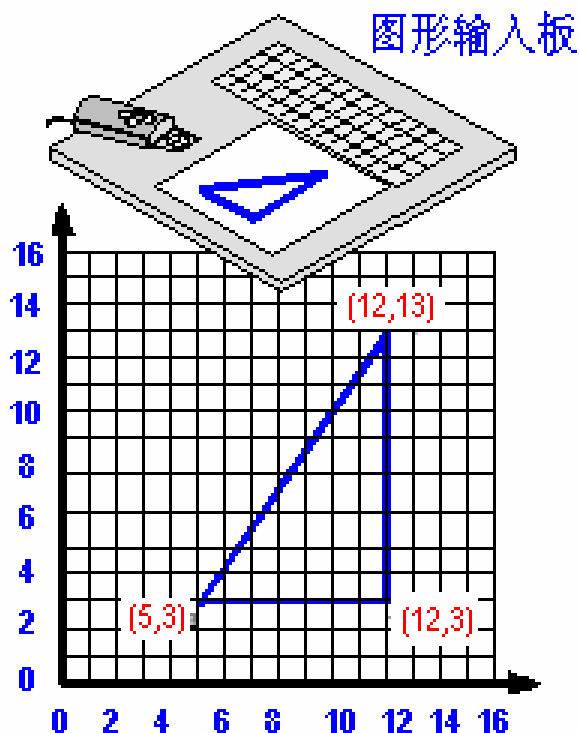
- 目标图像的移动、缩小或放大、旋转、拷贝、属性(如线条变宽变细、颜色)变更都很容易做到
- 相同或类似的图可以把它们当作图的构造块，并把它们存到图库中，这样不仅可加速矢量图的生成，而且可减小矢量图的文件大小

➤ 局限性

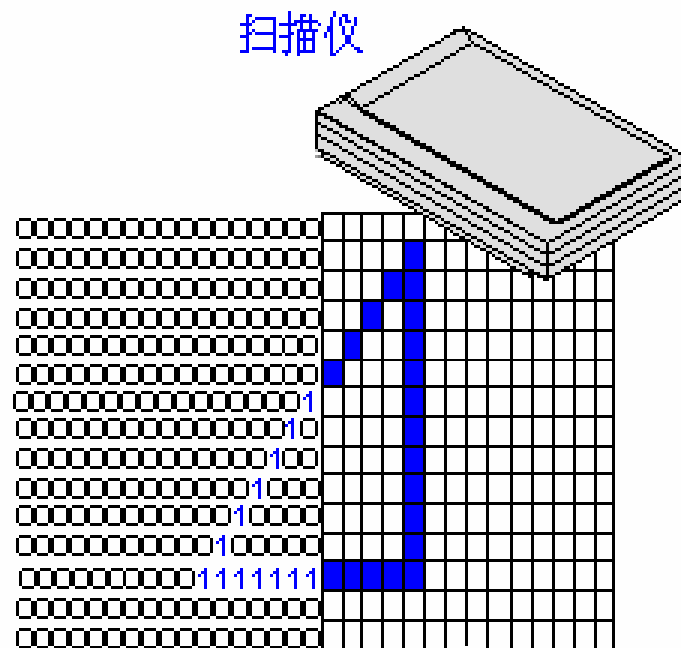
- 很难用数学方法来描述真实世界的彩照，这就要用位图法表示



4.2.2 图像的种类



(a) 矢量图



(b) 位图

矢量图与位图

4.2.2 图像的种类

■ 位图(bitmap, bitmapped image)

➤ 用像素值阵列表示的图

- 对位图进行操作时，只能对图中的像素进行操作，而不能把位图中的物体作为独立实体进行操作。也称光栅图(raster graphics)
- 画位图或编辑位图的软件称为画图程序(paint programs); 存放位图的格式称为位图格式; 存储的内容是描述像素的数值

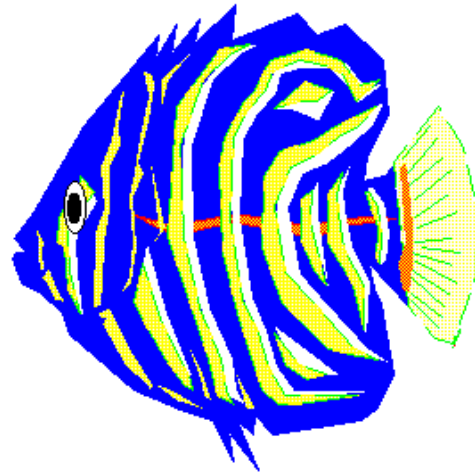
➤ 特性

- 位图的获取通常用扫描仪、数码相机、摄像机、录像机、视像光盘和相关的数字化设备
- 位图文件占据的存储空间比较大
- 影响位图文件大小的因素
 - ◆ 图像分辨率：分辨率越高，表示组成一幅图的像素就越多，图像文件就越大
 - ◆ 像素深度：像素深度越深，表达单个像素的颜色和亮度的位数越多，图像文件就越大

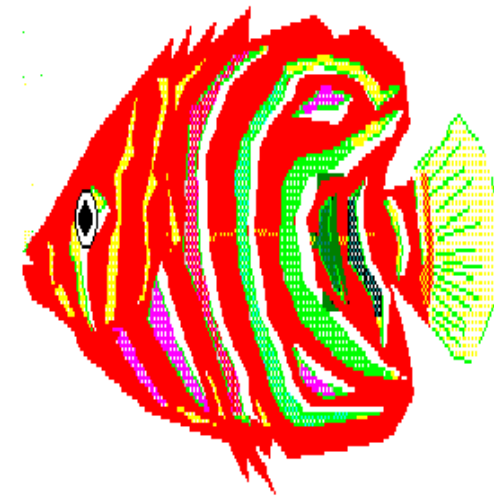
4.2.2 图像的种类

■ 图形图像(graphical image)

- 也称矢量图像(vector based image)
- 表示矢量图的位图，即用矢量图表示的图形转换成用像素表示的图像
- 用像素值阵列表示的位图边沿较光滑；用矢量图转换成的位图边沿较粗糙



位图



图形图像

4.2.2 图像的种类

- 灰度图(gray-scale image或intensity image)
 - 只有明暗不同的像素而没有彩色像素组成的图像
 - 只有黑白两种颜色的图像称为单色图像(monochrome/bit image), 每个像素的像素值用一位存储, 其值是“0”或“1”
 - 用一个字节表示一个像素的灰度图(256级灰度) 所示
 - 一幅 640×480 的灰度图像需要占据300 KB的存储空间



标准单色图



标准灰度图

4.2.2 图像的种类

■ 彩色图像(color image)

- 每个像素包含颜色信息的图像。
- 可按照颜色的数目划分
 - 256色图像：每个像素的R、G和B值用一个字节来表示，一幅 640×480 的彩色图像需要300 KB的存储空间
 - 真彩色图像：每个像素的R，G，B分量分别用一个字节表示，一幅 640×480 的真彩色图像需要900 KB的存储空间



256色标准图像



24位标准图像

4.2.3 伽马(γ)校正

■ γ 校正(gamma correction)

➤ 为补偿显示设备非线性的显示特性而采用的校正技术

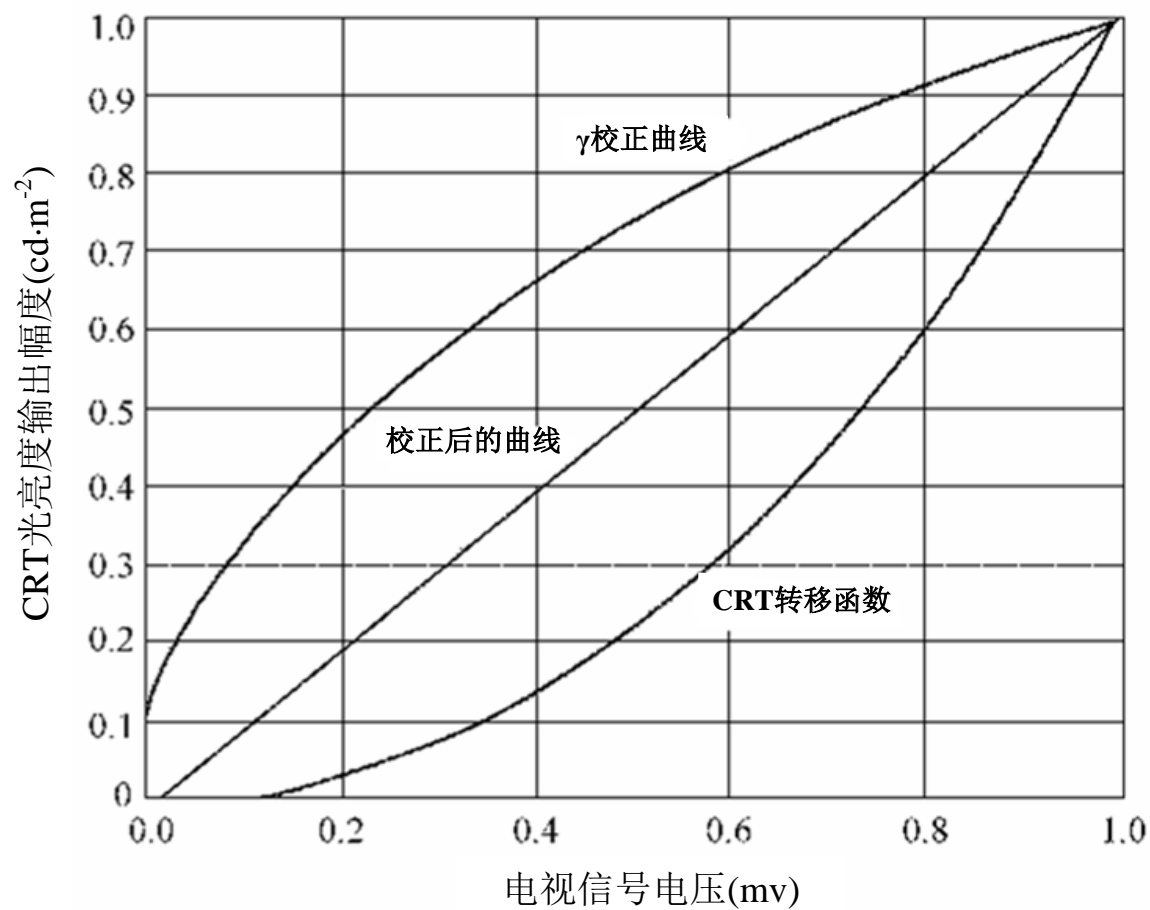
- 显示设备产生的光亮度与输入给它的电压之间呈现的关系为

$$\text{光亮度} = (\text{输入电压})^\gamma$$

其中， γ 是幂函数的指数，是一个常数，用来衡量非线性部件的转换特性

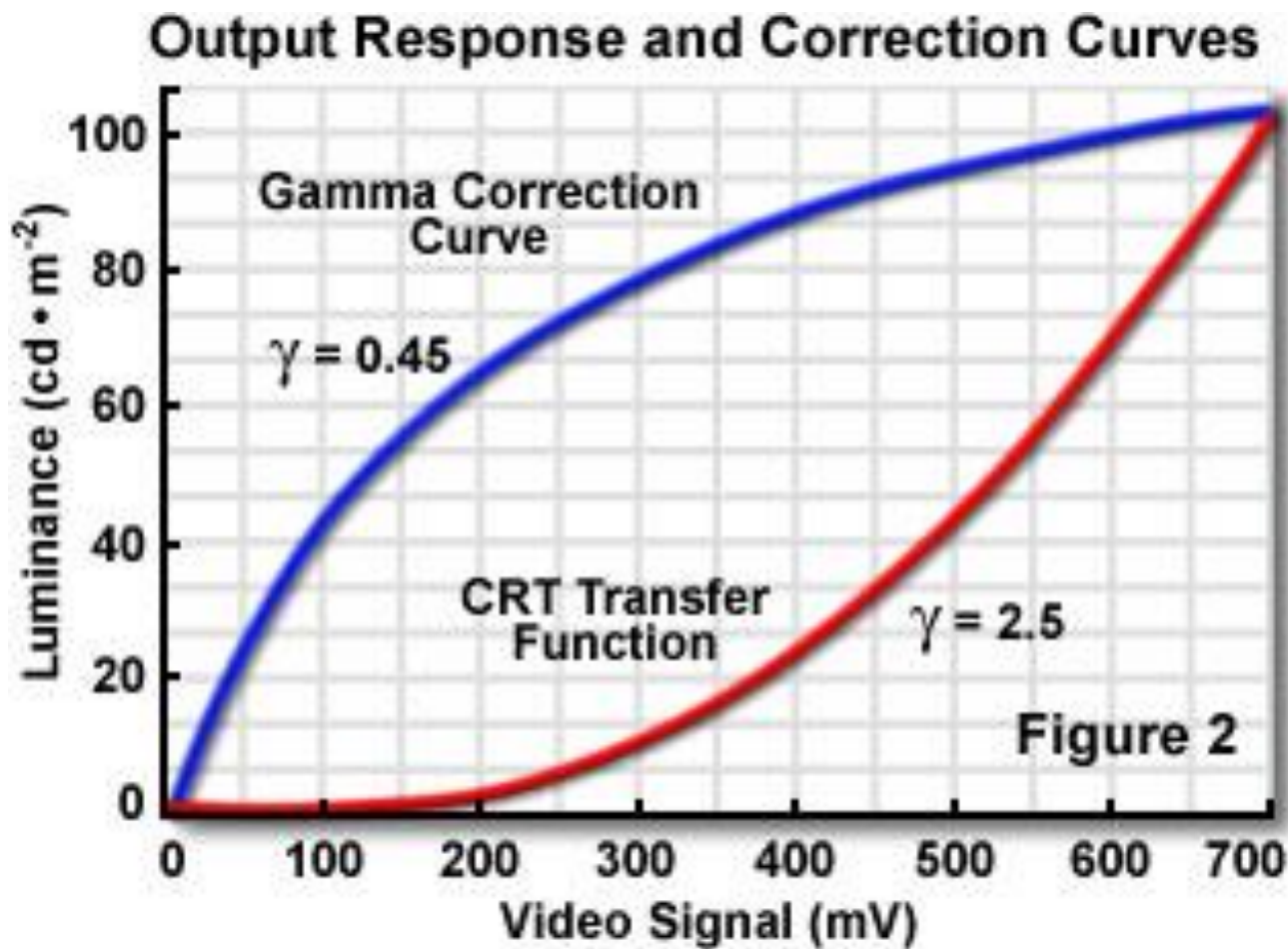
- 在把输入电压送到显示设备之前对它做一次变换，使输入到显示设备的电压 $V' = (\text{输入电压})^\gamma$ ，这个变换称为校正
- 显示设备不同， γ 的值也有所不同。阴极射线管(CRT)典型的值在2.25~2.45之间

4.2.3 伽马(γ)校正

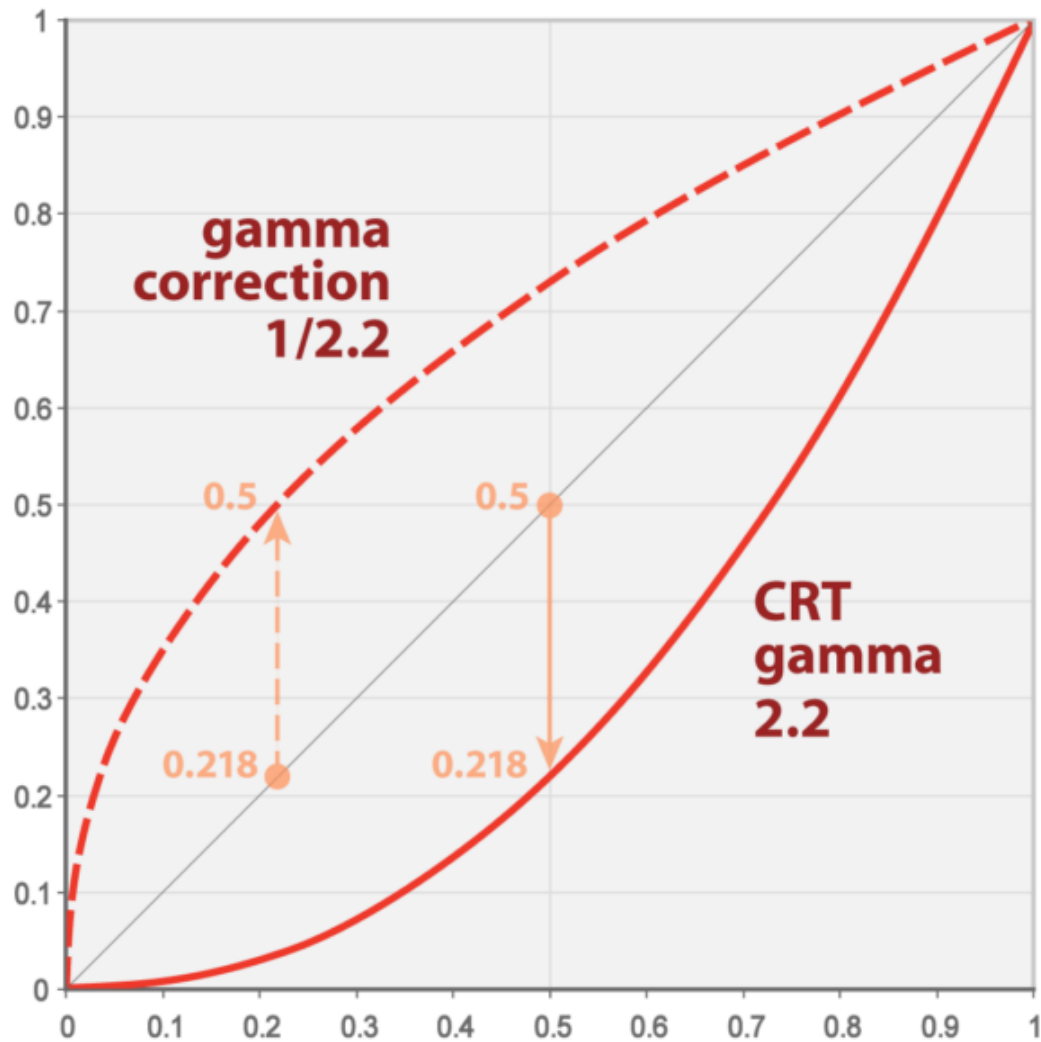


γ 校正

4.2.3 伽马(γ)校正



4.2.3 伽马(γ)校正





第4章 彩色数字图像基础目录

4.1 颜色简介

4.1.1 视觉系统对颜色的感知

4.2.2 图像的颜色模型

4.2 图像简介

4.2.1 图像基本属性

4.2.2 图像的种类

4.2.3 伽玛(γ)校正

4.3 JPEG压缩编码

4.3.1 算法概要

4.3.2 主要步骤

4.3.3 JPEG压缩和编码
举例

4.4 常用图像文件格式

BMP、GIF、JPEG、PNG

4.3.1 JPEG压缩编码

■ JPEG是什么

- Joint Photographic Experts Group联合图像专家组
 - 由ISO和IEC联合组成的，负责制定静态的数字图像数据压缩编码标准
- JPEG标准
 - 用于压缩灰度图像和彩色图像。
 - ◆ 有损压缩算法：以离散余弦变换(DCT)为基础
 - ◆ 采用以预测技术为基础的无损压缩算法
- JPEG格式
 - 存放使用JPEG压缩的图像文件交换格式
 - ◆ 图像文件的后缀是.JPG或.JFF(也称JFIF)





4.3.1 JPEG压缩编码

■ JPEG标准的文档

- 标准号：ISO/IEC IS 10918-1或ITU-T Recommendation T.81
- 标准名：Information technology — Digital compression and coding of continuous-tone still images (信息技术 — 连续色调静态图像的数字压缩和编码)

■ JPEG算法概要

- 利用视觉系统特性，使用变换、量化和熵编码相结合的方法，以去掉或减少视觉的冗余信息和数据本身的冗余信息

JPEG示例

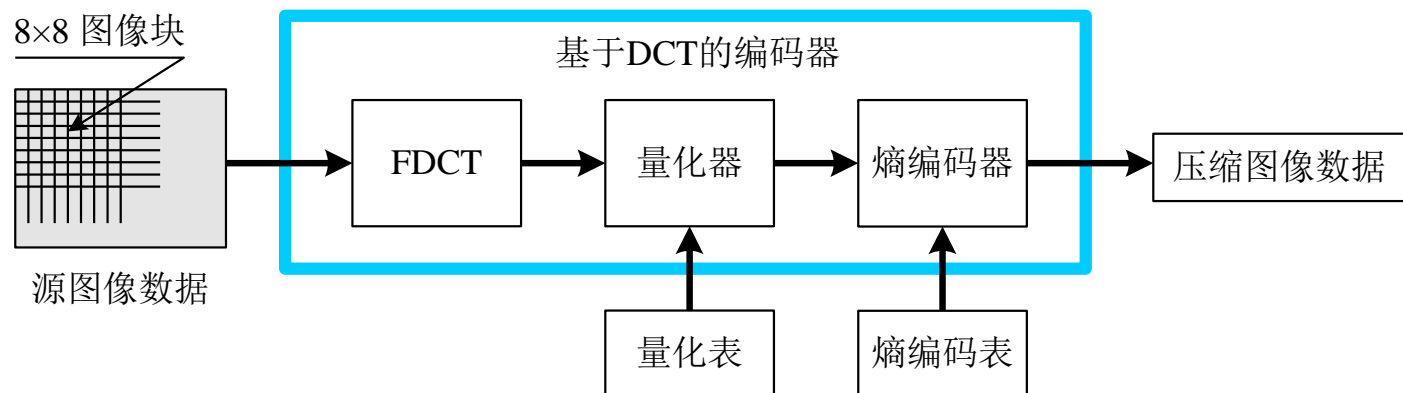




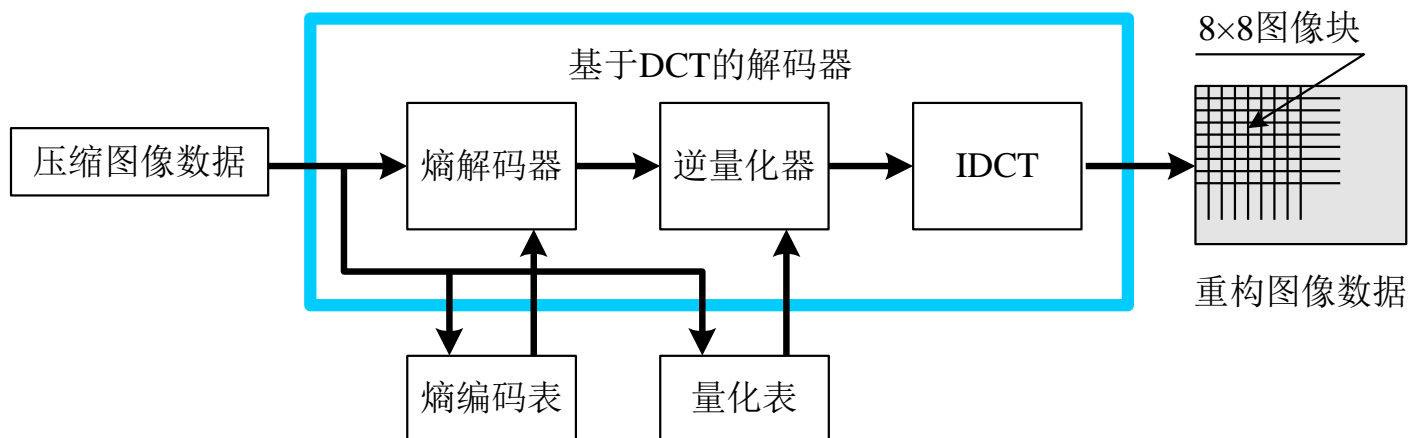
4.3.1 JPEG压缩编码

- JPEG标准的压缩算法步骤：
 - 使用正向离散余弦变换(FDCT)把空间域表示的图变换成频率域表示的图
 - 使用加权函数对DCT系数进行量化，加权函数对人的视觉系统是最佳的
 - 使用霍夫曼编码器对量化系数进行编码
- JPEG算法与颜色空间无关
 - RGB和YUV之间的变换不包含在JPEG算法中
 - JPEG算法处理单独的图像彩色分量，因此可压缩来自不同颜色空间的数据，如RGB, YCbCr, CMYK

4.3.2 JPEG压缩编码

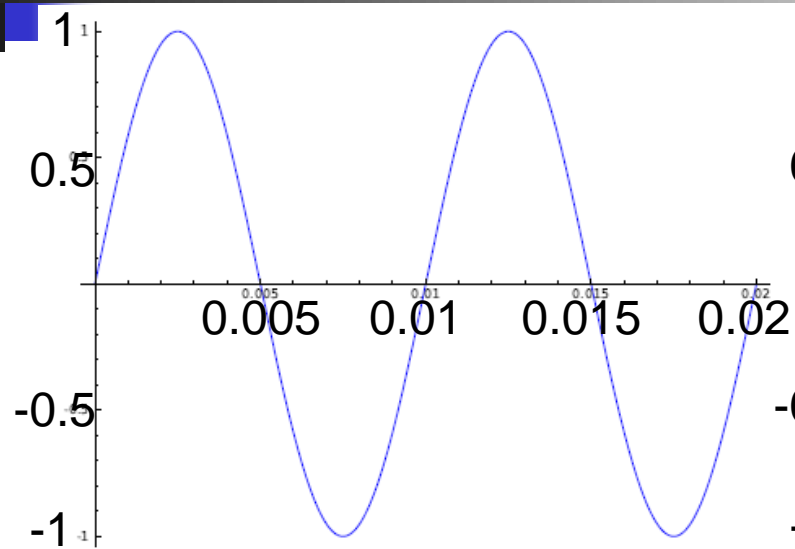


(a) JPEG压缩算法框图

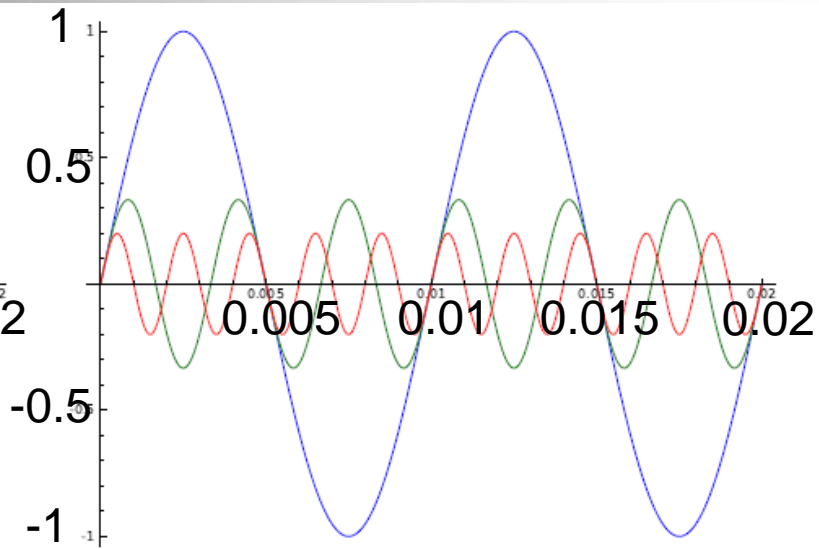


(b) JPEG解压缩算法框图

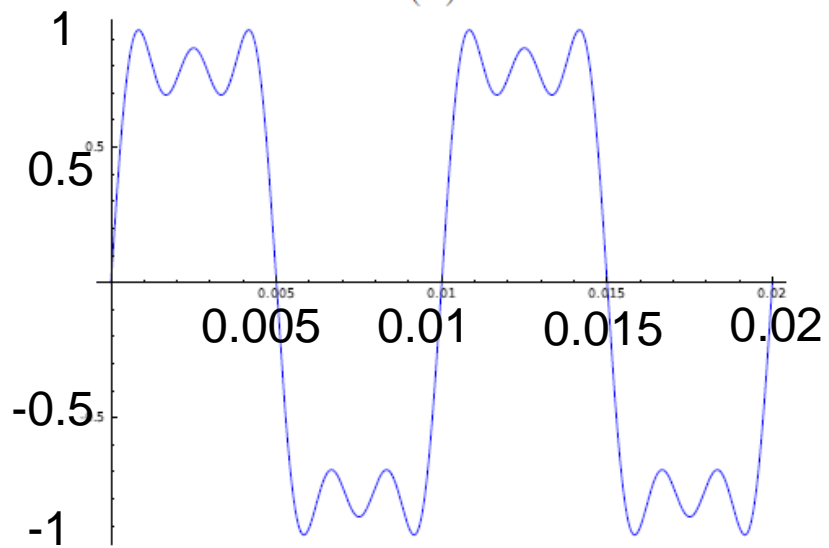
声波示例



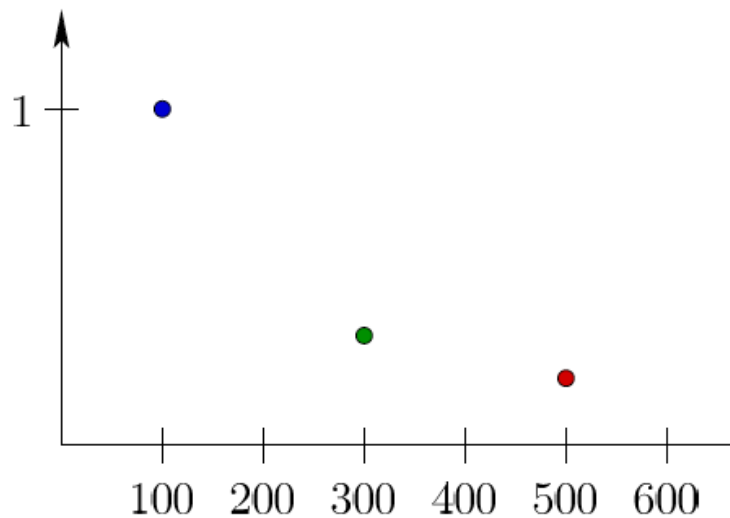
(a)



(b)

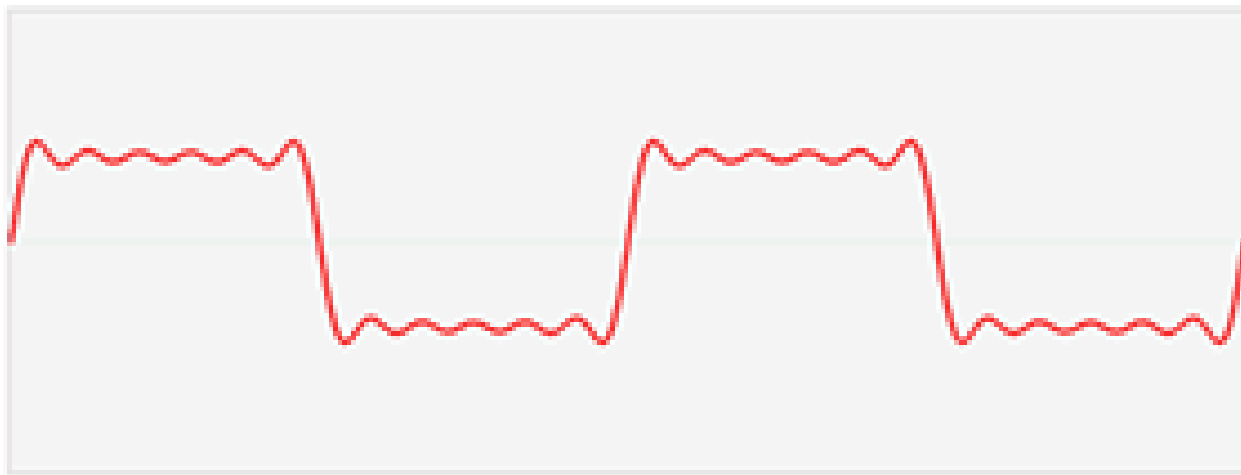


(c)



(d)

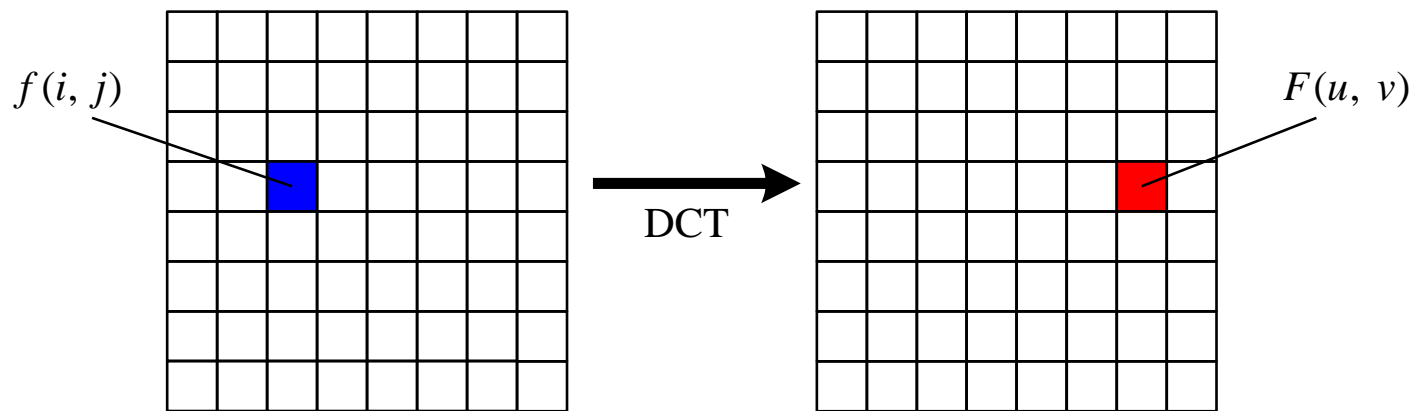
声波示例



4.3.2 JPEG压缩编码

■ 离散余弦变换(Discrete Cosine Transform, DCT)

- 用余弦函数的离散值构成的变换矩阵对信号的一系列样本值进行运算的数学变换
- 可将能量集中到频率较低的系数上



4.3.2 JPEG压缩编码

- DCT变换使用下式计算

$$\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{A}^T$$

- 逆变换使用下式计算

$$\mathbf{X} = \mathbf{A}^T \mathbf{Y} \mathbf{A}$$

其中,

$$A_{ij} = C_i \cos \frac{(2j+1)i\pi}{2N} \quad \text{where } C_i = \sqrt{\frac{1}{N}} \ (i=0), \ C_i = \sqrt{\frac{2}{N}} \ (i>0)$$

4.3.2 JPEG压缩编码

- DCT变换使用下式计算

$$F(u, v) = \frac{1}{4} C(u) C(v) \left[\sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \right]$$

- 逆变换使用下式计算

$$F(i, j) = \frac{1}{4} C(u) C(v) \left[\sum_{u=0}^7 \sum_{v=0}^7 f(u, v) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \right]$$

其中, $C(u), C(v) = 1/\sqrt{2}$ 当 $u, v=0$;
 $C(u), C(v) = 1$ 其他

DCT变换示例



126	159	178	181
98	151	181	181
80	137	176	156
75	114	88	68

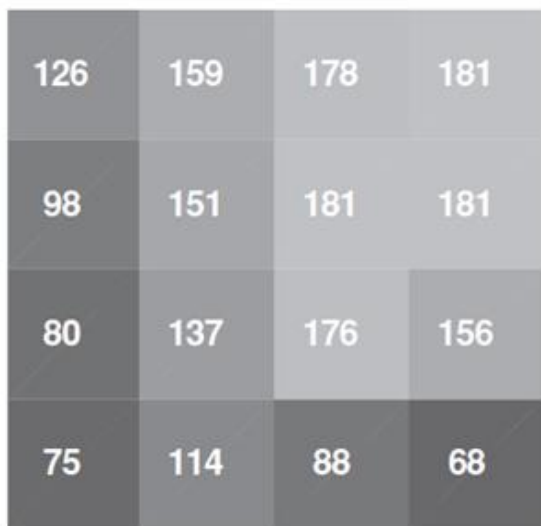
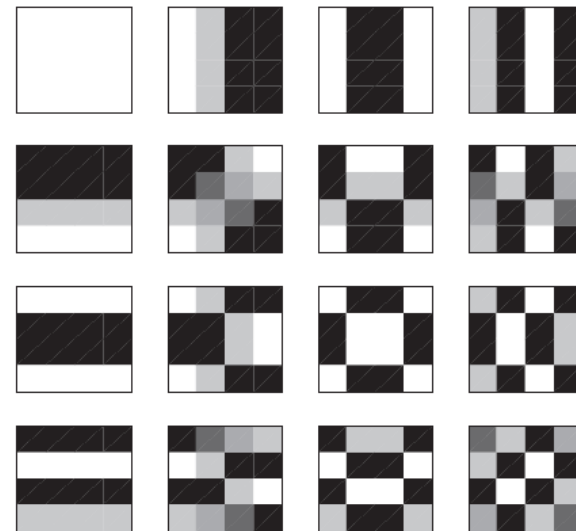
Original block

$$A = \begin{bmatrix} \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) \\ \sqrt{\frac{1}{2}} \cos\left(\frac{\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{3\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{5\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{7\pi}{8}\right) \\ \sqrt{\frac{1}{2}} \cos\left(\frac{2\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{6\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{10\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{14\pi}{8}\right) \\ \sqrt{\frac{1}{2}} \cos\left(\frac{3\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{9\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{15\pi}{8}\right) & \sqrt{\frac{1}{2}} \cos\left(\frac{21\pi}{8}\right) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.653 & 0.271 & -0.271 & -0.653 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.271 & -0.653 & 0.653 & -0.271 \end{bmatrix}$$

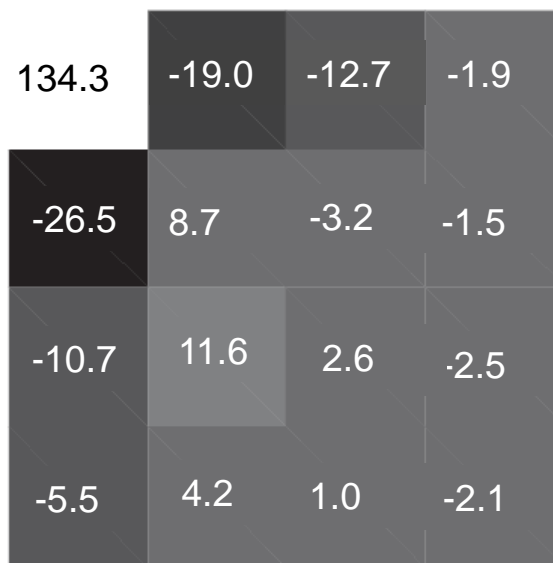
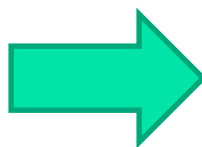
DCT变换示例

$$Y = AXA^T$$

$$Y_{xy} = \frac{1}{4} C_x C_y \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{ij} \cos \frac{(2j+1)y\pi}{2N} \cos \frac{(2i+1)x\pi}{2N}$$



Original block



DCT coefficients

DCT变换示例

126	159	178	181
98	151	181	181
80	137	176	156
75	114	88	68

Original block

134	134	134	134
134	134	134	134
134	134	134	134
134	134	134	134

1 coefficient

169	169	169	169
149	149	149	149
120	120	120	120
100	100	100	100

2 coefficients

DCT变换示例

126	159	178	181
98	151	181	181
80	137	176	156
75	114	88	68

Original block

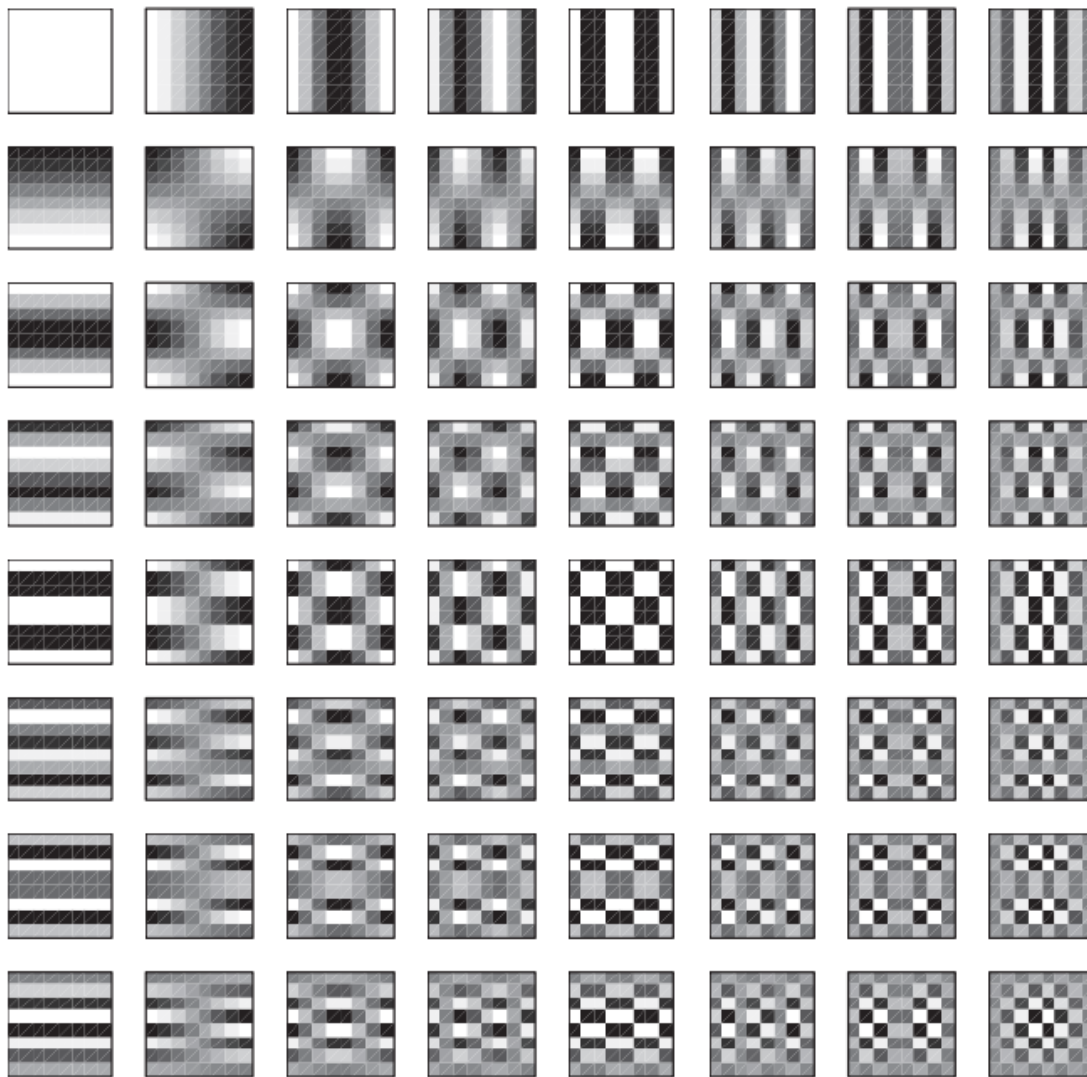
144	159	179	194
124	138	159	173
95	110	130	145
75	89	110	124

3 coefficients

146	179	187	165
95	146	179	175
66	117	150	146
76	109	117	96

5 coefficients

二维DCT变化系数



4.3.2 JPEG压缩编码

➤ 二维DCT的计算

- 将二维DCT变成一维DCT
- 当计算精度足够高时，DCT变换不会损失图像质量

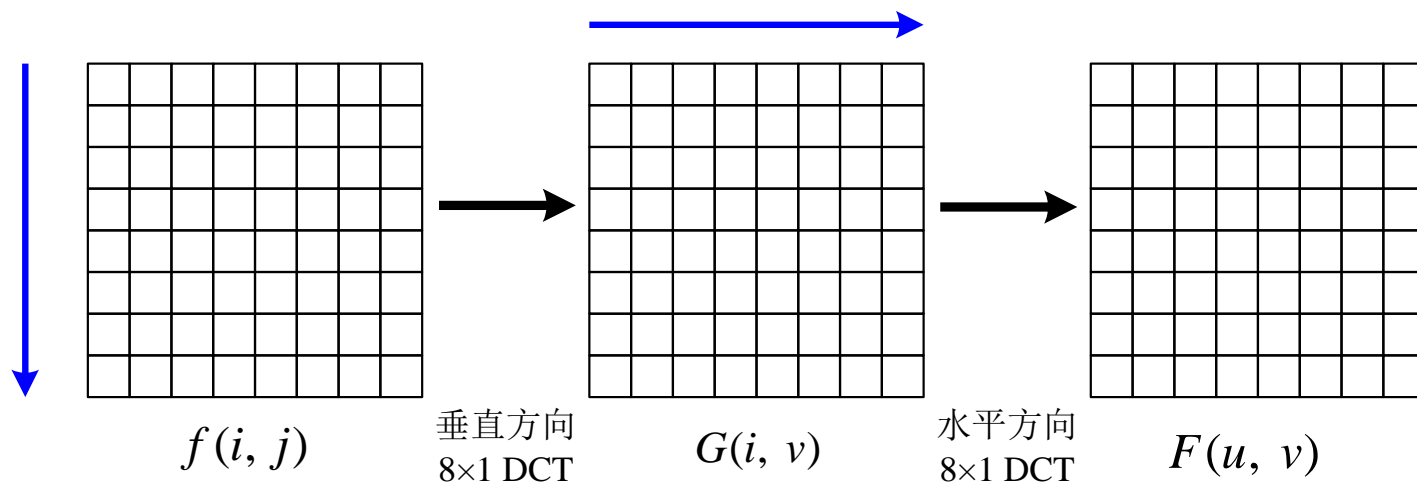


图4-11 二维DCT变换方法

4.3.2 JPEG压缩编码

■ 量化

- 对FDCT变换后的(频率的)系数进行量化
- 量化目的是降低非“0”系数的幅度以及增加“0”值系数的数目
- 量化是造成图像质量下降的最主要原因

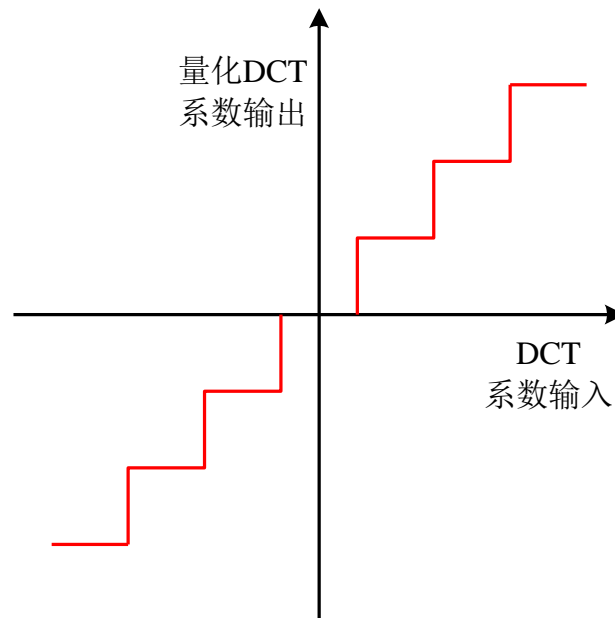


图4-12 均匀量化器

$$\hat{F}(u, v) = \text{round}\left(\frac{F(u, v)}{Q(u, v)}\right)$$

量化表

亮度量化表

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

色差量化表

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



4.3.2 JPEG压缩编码

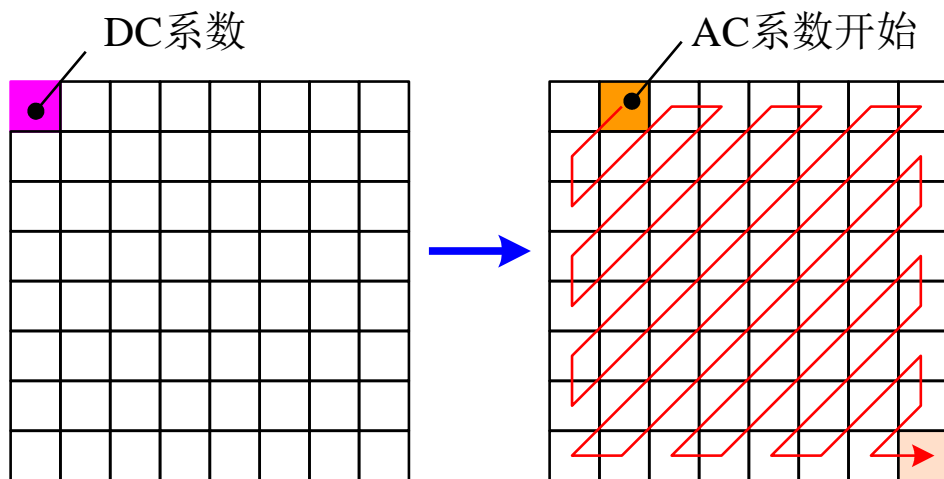
➤ 量化步距

- 按照系数所在的位置和每种颜色分量的色调值来确定的
- 因为人眼对亮度信号比对色差信号更敏感，因此使用两种量化表：亮度量化表和色差量化表
 - ◆ 由于人眼对低频分量的图像比对高频分量的图像更敏感，因此表中的左上角的量化步距要比右下角的量化步距小

4.3.2 JPEG压缩编码

■ Z字形编排(Zig-zag scan)

- 为增加连续的“0”值系数的个数，即“0”的游程长度，
- DCT系数的序号:序号小的位置表示频率较低，把一个 8×8 的矩阵变成一个 1×64 的矢量

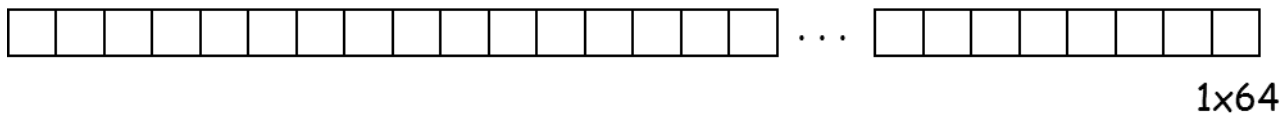
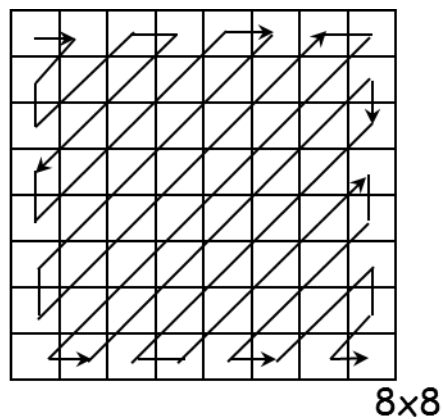


量化DCT系数的编排

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

DCT系数序号

Zig-zag Scan

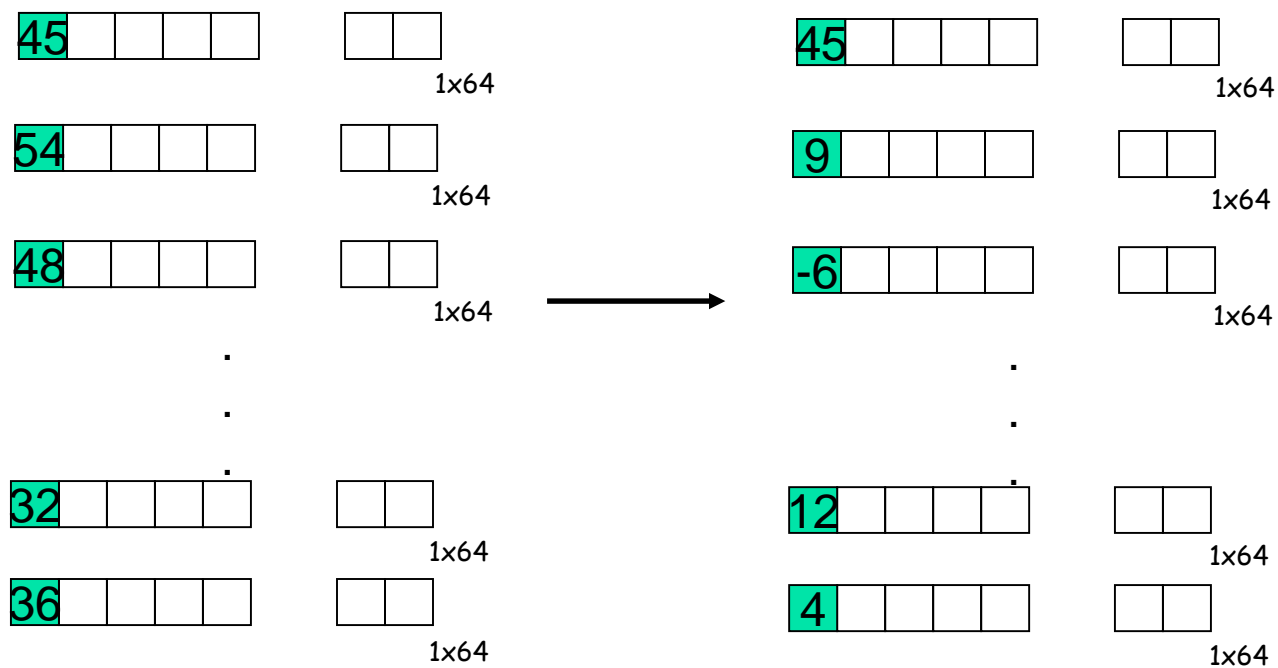


4.3.2 JPEG压缩编码

■ 熵编码

- 用于进一步压缩采用DPCM编码后的DC系数差值和RLE编码后的AC系数
- 先将 8×8 图像块的DC系数和63个AC系数用中间符号表示，中间符号由两个符号组成
 - 一个符号是表示数据大小的可变长度码(variable-length code, VLC)，用的代码是霍夫曼码
 - 一个符号是直接表达实际幅度的可变长度整数(variable-length integer, VLI)
 - 由于DC系数和AC系数的统计特性不同，因此在熵编码时需要对DC系数和AC系数分别处理

DC系数的DPCM编码



熵编码:DC系数

- DC系数中间符号: SSSS(size) DIFF(Amplitude)

SIZE	Value	Code
0	0	---
1	-1,1	0,1
2	-3, -2, 2,3	00,01,10,11
3	-7,..., -4, 4,..., 7	000,..., 011, 100,...111
4	-15,..., -8, 8,..., 15	0000,..., 0111, 1000,..., 1111
.		.
.		.
11	-2047,..., -1024, 1024,... 2047	...

Size_and_Value Table

熵编码:DC系数

- DC系数中间符号: SSSS(size) DIFF(Amplitude)

SIZE	Code Length	Code
0	2	00
1	3	010
2	3	011
3	3	100
4	3	101
5	3	110
6	4	1110
7	5	11110
8	6	111110
9	7	1111110
10	8	11111110
11	9	111111110

Example: If a DC component is 40 and the previous DC component is 48. The difference is -8. Therefore it is coded as:

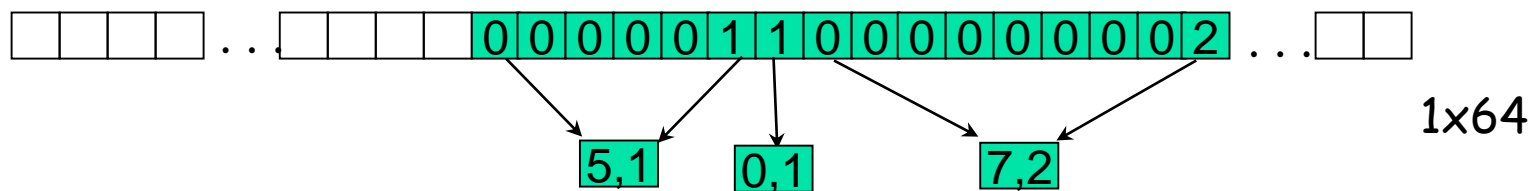
1010111

0111: The value for representing -8 (see Size_and_Value table)

101: The size from the same table reads 4. The corresponding code from the table at left is 101.

Huffman Table for DC component SIZE field

AC系数的RLE编码



熵编码: AC系数

- AC系数中间符号: RRRR(Run-length)SSSS(size) AC系数 (Amplitude)

Run/ SIZE	Code Length	Code
0/0	4	1010
0/1	2	00
0/2	2	01
0/3	3	100
0/4	4	1011
0/5	5	11010
0/6	7	1111000
0/7	8	11111000
0/8	10	1111110110
0/9	16	111111110000010
0/A	16	111111110000011

Run/ SIZE	Code Length	Code
1/1	4	1100
1/2	5	11011
1/3	7	1111001
1/4	9	111110110
1/5	11	11111110110
1/6	16	111111110000100
1/7	16	111111110000101
1/8	16	111111110000110
1/9	16	111111110000111
1/A	16	111111110001000
2/1	5	11010
... 15/A	More	Such rows

熵编码: AC系数

40	12	0	0	0	0	0	0
10	-7	-4	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
0	0	0	0	0	0	0	0

Example: Consider encoding the AC components by arranging them in a zig-zag order \rightarrow 12, 10, 1, -7
2 0s, -4, 56 zeros

12: read as zero 0s, 12: $(0/4)12 \rightarrow$ 10111100

1011: The code for (0/4 from AC code table)

1100: The code for 12 from the
Size_and_Value table.

10: $(0/4)10 \rightarrow$ 10111010

1: $(0/1)1 \rightarrow$ 001

-7: $(0/3)-7 \rightarrow$ 100000

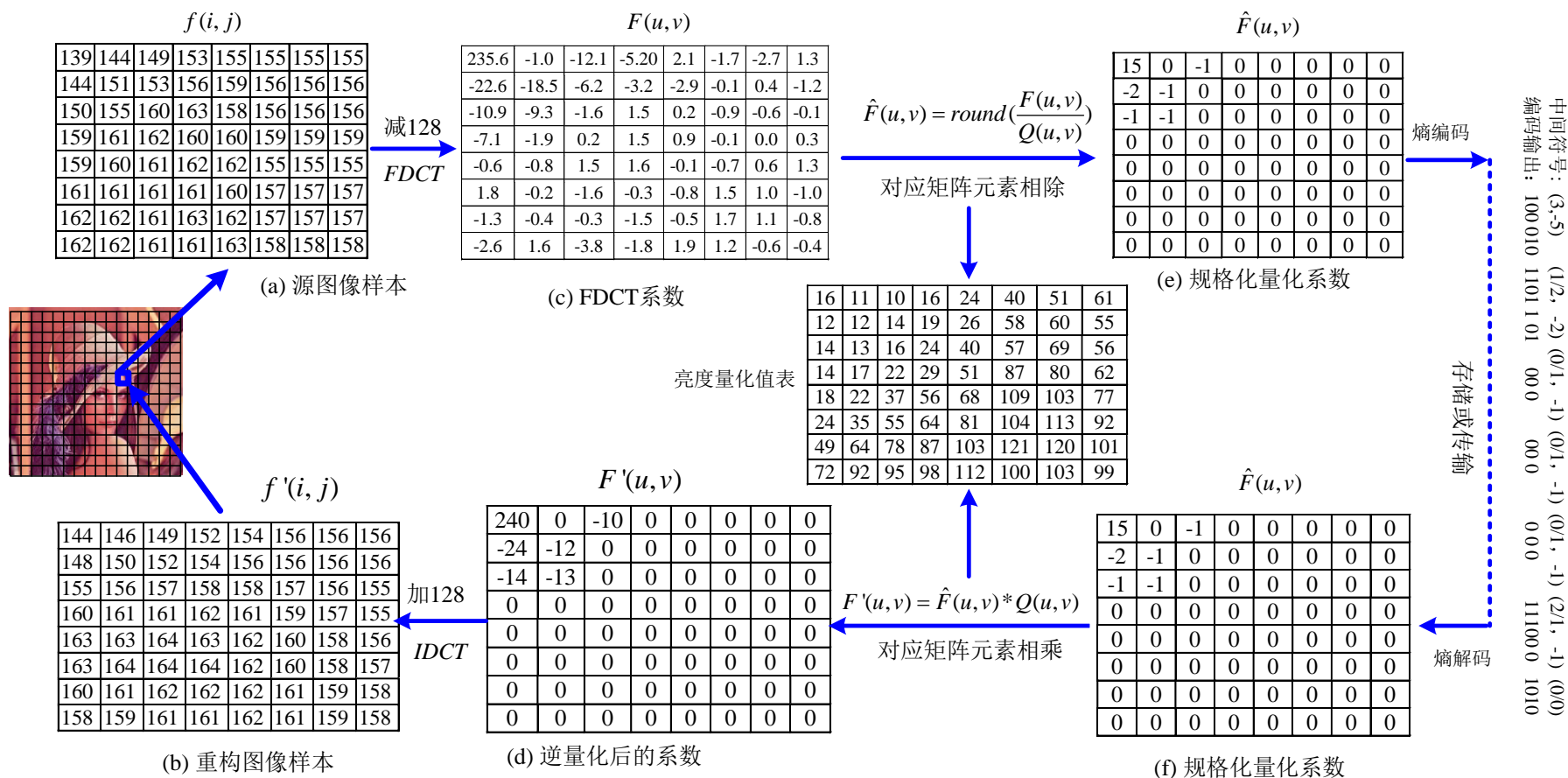
2 0s, -4: $(2/3)-4 \rightarrow$ 1111110111011

1111110111: The 10-bit code for 2/3

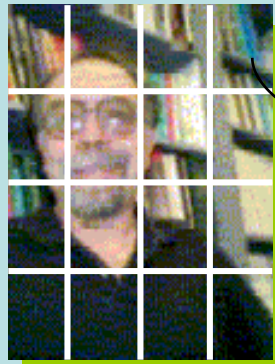
011: representation of -4 from Size_and_Value
table.

56 0s: $(0,0) \rightarrow$ 1010 (Rest of the components are
zeros therefore we simply put the EOB to
signify this fact)

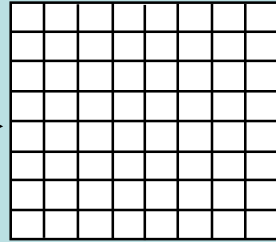
4.3.3 JPEG压缩编码示例



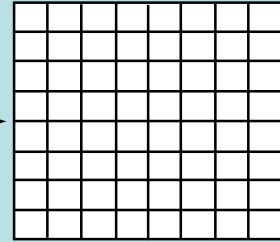
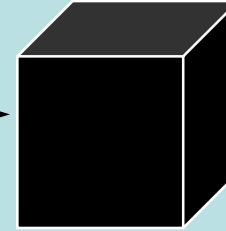
通过DCT变换，把能量集中在少数几个系数上。



8x8



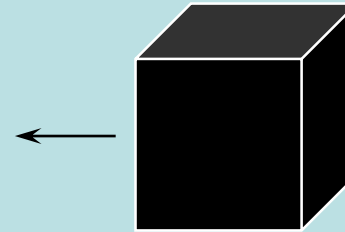
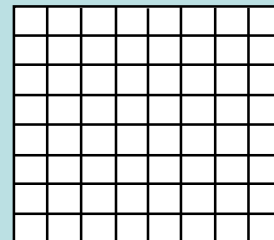
DCT



2D-〉1D，目的是为了增加连续的“0”系数的个数，就是“0”的游程长度

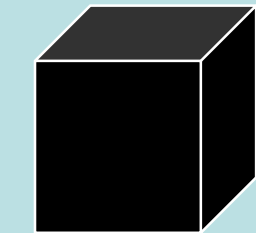


Zig-zag

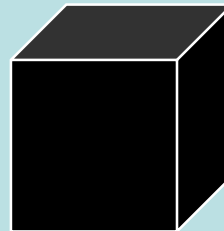
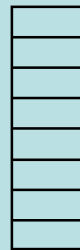


Quantize

对经过FDCT变换后的频率系数进行量化。量化是图像质量下降的最主要原因



Run-length Code



Huffman Code

011010001011101...



第4章 彩色数字图像基础目录

4.1 颜色简介

4.1.1 视觉系统对颜色的感知

4.2.2 图像的颜色模型

4.2 图像简介

4.2.1 图像基本属性

4.2.2 图像的种类

4.2.3 伽玛(γ)校正

4.3 JPEG压缩编码

4.3.1 算法概要

4.3.2 主要步骤

4.3.3 JPEG压缩和编码
举例

4.4 常用图像文件格式

BMP、GIF、JPEG、PNG

■ 文件格式

- 存储文本、图形或者图像数据的数据结构。

■ 图像文件格式

- 存储图像数据的数据结构。

4.4 图像文件格式

■ BMP文件格式 (Bitmap-File)

- Windows采用的文件格式
- 可没有任何压缩，或用RLE行程编码进行压缩

BMP图像文件结构

组成部分的名称	数据结构的名称
位图文件头(bitmap-file header)	BITMAPFILEHEADER
位图信息头(bitmap-information header)	BITMAPINFOHEADER
彩色表(color table)	RGBQUAD
图像数据阵列字节(byte)	BYTE

4.4 图像文件格式

■ GIF (Graphics Interchange Format) 格式——图形文件交换格式

- CompuServe公司开发的图像文件存储格式



- 图像的相关信息以数据块(block)为单位
- 在一个文件中可存放多幅彩色图形/图像，并可像幻灯片那样显示或像动画那样演示
- 采用LZW压缩算法来压缩图像数据
- 用户可为图像设置透明(transparency)的背景



4.4 图像文件格式

■ JPEG (Joint Photographic Experts Group) 格式

- JPEG委员会在制定JPEG标准时定义了许多标记(marker)，用来区分和识别图像数据及其相关信息
- 广泛使用的JPEG文件格式是JPEG文件交换格式(JPEG File Interchange Format, JFIF)
- 由于JFIF文件格式直接使用JPEG标准为应用程序定义的许多标记，因此JFIF格式就成了事实上的JPEG文件交换格式标准

4.4 图像文件格式

- **PNG (Portable Network Graphic Format) 格式——便携网络图形格式**
 - 一种位图文件(bitmap file)存储格式
 - 用来存储灰度图像时，灰度图像的深度可多达16位，存储彩色图像时，彩色图像的深度可多达48位，并且还可存储多达16位的 α 通道数据
 - PNG使用从LZ77派生的数据无损压缩算法

4.4 图像文件格式

➤ PNG企图代替GIF和TIFF, 保留的GIF文件格式特性



A PNG
image with
an 8-bit
transparency
channel

- 使用彩色查找表或称调色板, 支持256种颜色的彩色图像
- 流式读/写性能(streamability)
- 逐次逼近显示(progressive display)
- 透明性(transparency)
- 辅助信息(ancillary information): 可用来在图像文件中存储一些文本注释信息
- 独立于计算机软硬件环境
- 使用无损压缩。

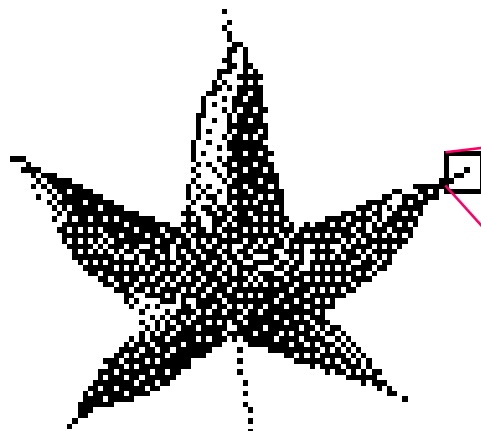


4.4 图像文件格式

- PNG增加的GIF文件格式所没有的特性：
 - 每个像素为48位的真彩色图像
 - 每个像素为16位的灰度图像
 - 可为灰度图和真彩色图添加 α 通道
 - 添加图像的 γ 信息
 - 使用循环冗余码(cyclic redundancy code, CRC)检测损害的文件
 - 加快图像显示的逐次逼近显示方式。
 - 标准的读/写工具包

黑白图象的表示

Black and
white image

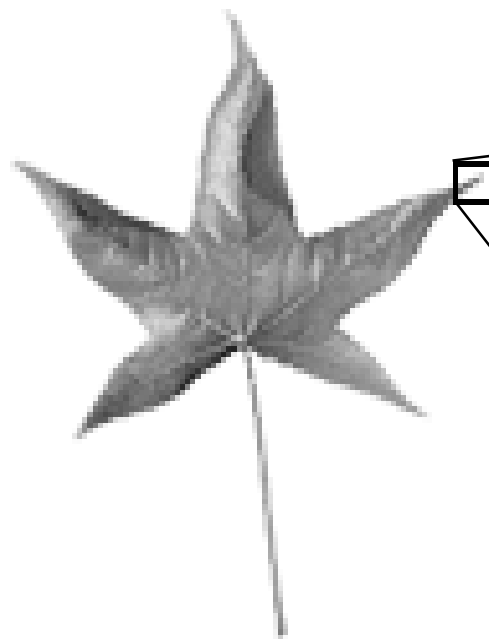


1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	1
1	1	0	0	1	1	1	1
0	0	1	1	1	1	1	1
0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

像素深度是多少？

灰度图象的表示

Gray scale
image



255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	240	213	255	255
255	255	238	176	133	206	255	255
200	174	109	128	224	255	255	255
115	108	184	245	255	255	255	255
175	224	255	255	255	255	255	255
255	255	255	255	255	255	255	255

像素深度是多少？

彩色图象的表示



R component

153	156	159	170	150	151	175	176
150	154	159	166	156	158	177	178
147	153	158	162	156	168	180	188
168	175	175	174	177	182	187	183
225	225	219	217	216	218	223	227
225	224	221	220	214	215	222	225
240	233	226	223	219	220	224	229
233	231	229	226	220	220	227	230

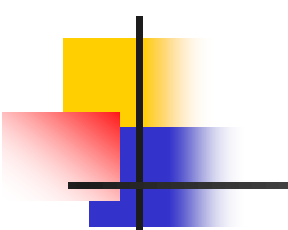
G component

178	176	176	176	176	205	216	226
179	178	175	180	177	200	223	231
174	175	178	184	181	189	217	224
208	203	208	196	193	197	216	225
212	210	215	202	192	196	207	218
211	212	212	210	198	194	207	210
224	227	224	214	197	196	213	220
228	231	233	220	202	197	210	217

B component

180	177	187	190	190	220	225	231
182	184	179	188	192	217	239	233
182	185	190	191	194	207	229	235
219	215	218	198	198	205	220	237
211	214	218	202	192	195	212	234
217	214	213	210	194	192	214	238
222	225	226	214	197	200	216	230
228	230	230	225	200	206	212	220

像素深度是多少？



假如有一副彩色图像，尺寸为8"x 6"，分辨率为300 DPI（Dot Per Inch），每个像素有512种可能的颜色，请问未经压缩的图像大小为多少字节？

Solution: $8 \times 6 \times 300 \times 300 \times 9/8 = 4860000$ Bytes

作业题-JPEG编码

前一个 8×8 图像块的
DC系数为20，求

- 1) Z字形扫描后的
 1×64 矢量;
- 2) DC系数的中间符号
及编码输出;
- 3) AC系数的中间符号
及编码输出。

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-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

注:可使用教材上的表4-8, 表4-9, 表4-11, 表4-12或者此ppt第74,75,77的码表。



作业

- JPEG编码作业题
- 教科书
 - 课后题 3.7, 4.5, 6.3
- 提交时间：12月5号（下周三）



END

第4章 彩色数字图像基础

