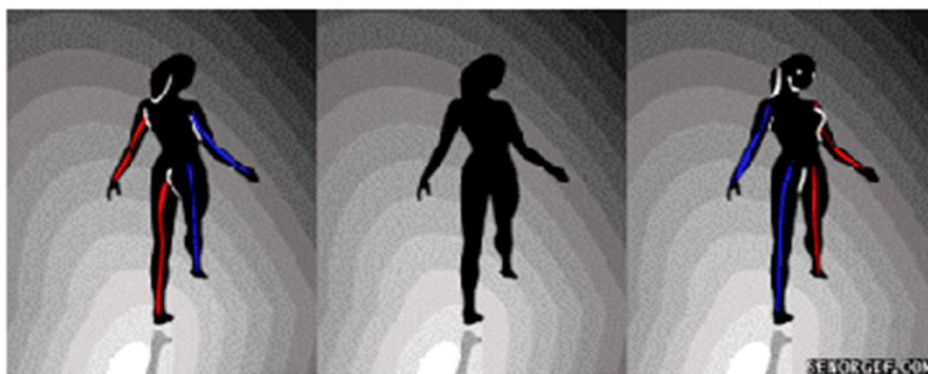




## 视觉系统的有趣现象



# 注意力机制

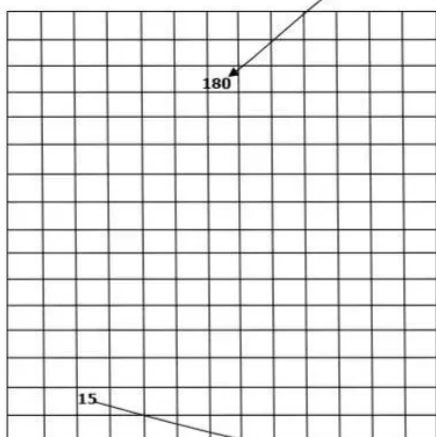
机器学习/深度学习中的概念



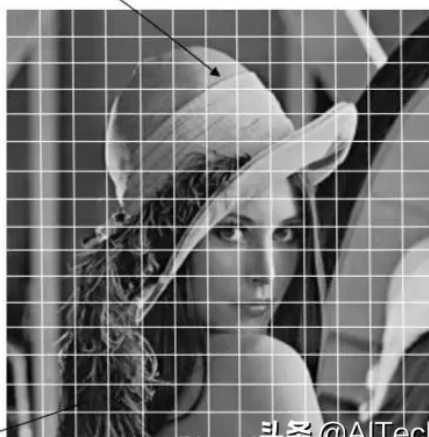
## 什么是数字图像

数字图像的本质

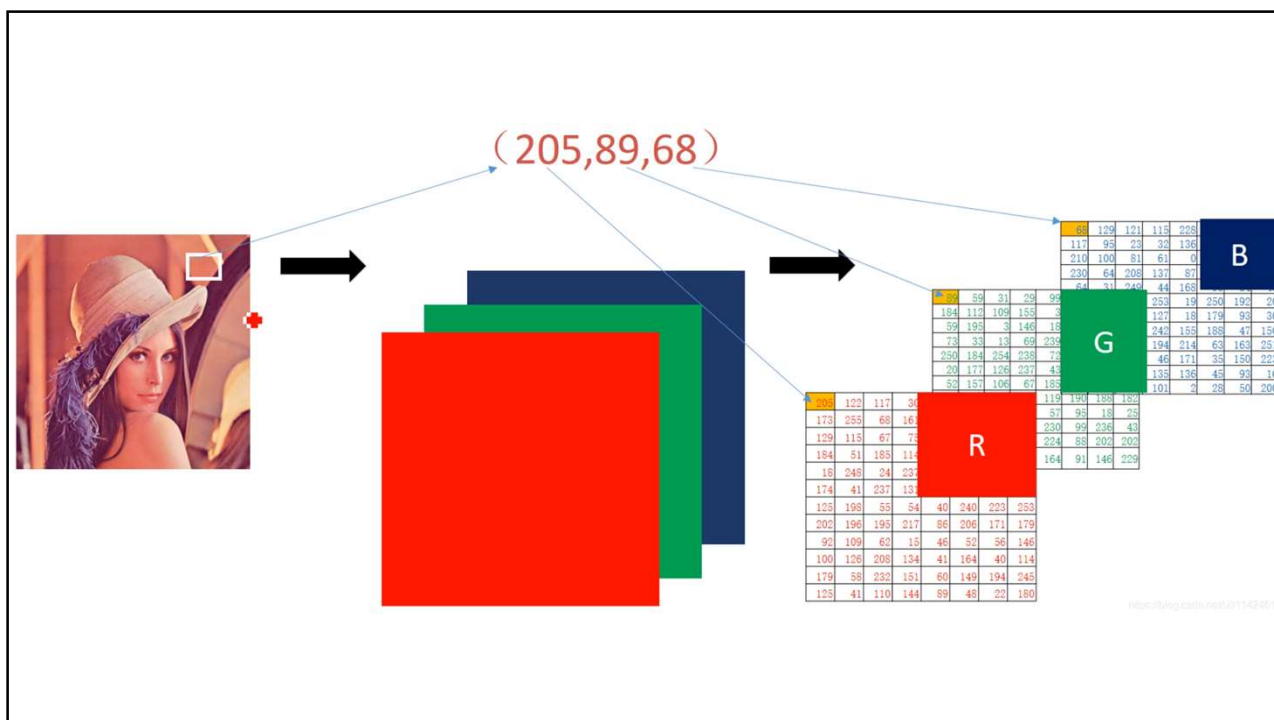
数字图像



模拟图像（照片）



头条 @AITech



## 图像的三个基本属性





## 三个基本属性

图像的三个基本属性

- 分辨率(要求学会区分)
  - 显色分辨率
  - 图象分辨率
- 像数深度(图像深度)
- 正彩色, 伪彩色, 直接色。
  - 略过, 不讲

## 图象分辨率

- 组成一幅图象的像素密度的度量方法:
  - 同样大小的一幅图, 如果像素数目越多, 图象的分辨率越高, 看起来越逼真。
- (DPI, Dots per inch):
  - 如果用300DPI扫描一幅8×10英寸的图像, 就得到一幅2400×3000个像素的图像。分辨率越高, 像素就越多。
- 图像分辨率和显示分辨率两个不同概念。



## 色彩深度

像数深度(图像深度)

- 指**存储每个像素使用的位数**。
- 用来度量图像的分辨率，决定每个像素可能有的颜色数，灰度图像的每个像素可能有的灰度级数。
- 常用的色深一般为：
  - 256色：8位
  - 增强色：16位，64K色
  - 真彩色：24位，16M。

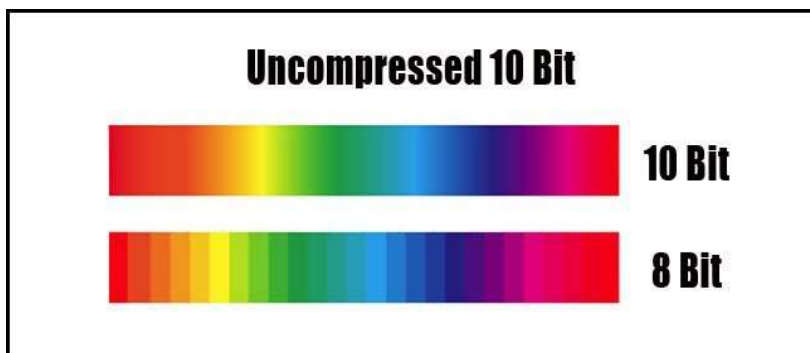
## 真彩色

图像基础

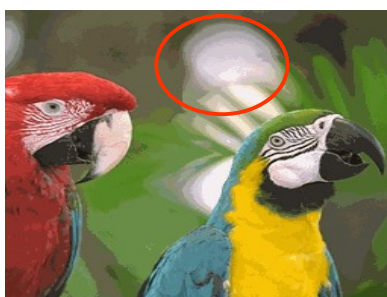
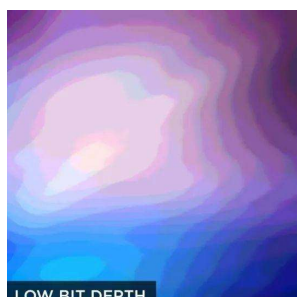
- 真彩色(True Color)
  - 组成一个图像的每个像素值中，有R，G，B三个基色分量，每个基色分量直接决定显示设备的基色强度，这样产生的彩色为“真彩色”。
  - RGB8:8:8,也称为全彩色。

**真彩色：24位，16M**

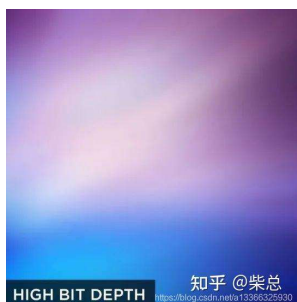




256色标准图像转换成的灰度图



24位标准图像转换成的灰度图





## $\alpha$ 通道(Alpha Channel)

- $\alpha$ 通道(alpha channel):又称为覆盖位(overlay),中断位, 属性位。
  - 除了RGB分量用固定位数表示外, 还增加一位或几位为属性位, 如RGB5:5:5表示一个像素时, 用两个字节共16位表示, RGB各占一位, 剩下一位位属性位, 指定该像素具有的性质。
- $(R,G,B)=(1,0,0)$      $(\alpha,R,G,B)=(0.5,1,0,0)$ ;
- 该像素显示时, 红色强度降低一般。



## 图像的种类

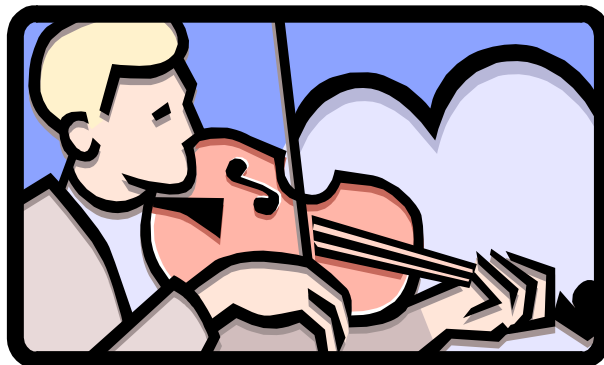
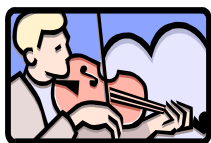
图像分类





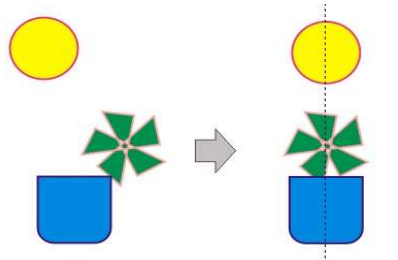
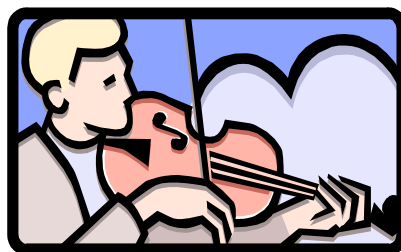
## 矢量图

矢量图，用一系列计算机指令来表示一幅图。



### 矢量图的特点：

- 1)图形缩放后不变形：
  - 矢量图形是以指令集合的形式描述的，图形缩放不会变形。
- 2)占用的存储容量小：
  - 由于矢量图形只是存储的图形指令，所需的存储空间很小。
- 3)灵活性好：
  - 构成矢量图形的各个部分是相对独立的，局部的处理不会影响其他部分。
- 4)表现能力有一定局限：
  - 对于复杂而又不易用数学方法描述的对象，矢量图形就无能为力。显示速度比点位图慢。







## 点位图

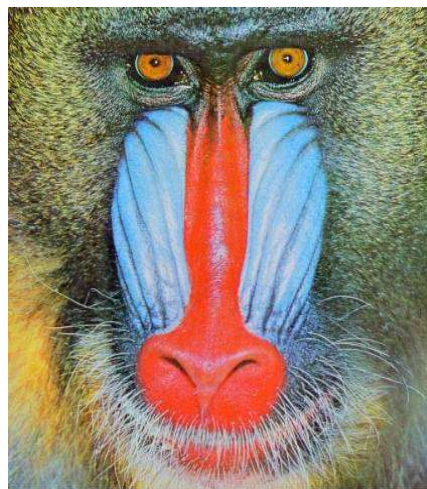
一幅彩色图像可以看成由许多的点组成的。

■ 图像中的单个点称为**像素(Pixel)**，每个像素都有一个值，称为像素值，一个像素值往往用R, G, B三个分量表示。

■ 每种基色的强度是用8位表示的，因此可产生  $2^8 \times 2^8 \times 2^8 = 16\,777\,216$  种颜色

## 点位图

- 点位图
  - 每个像数用若干各二进制位来指定该像素的颜色，亮度和属性。
  - 点位图是对视觉信号进行**直接量化**的媒体形式。
  - 也称：位图、点阵图、图像、像素图。

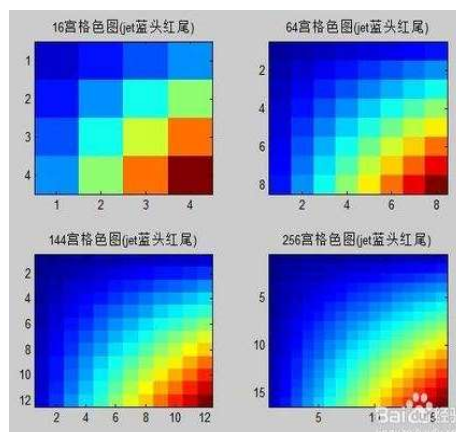




## 点位图文件的大小(知识点)

图像基础

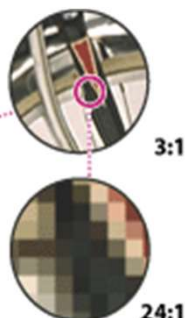
- 计算公式为：
- 文件字节数=图像分辨率X图像（色彩）深度/8



Adobe Creative Suite Icons

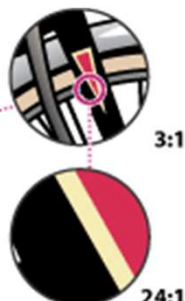
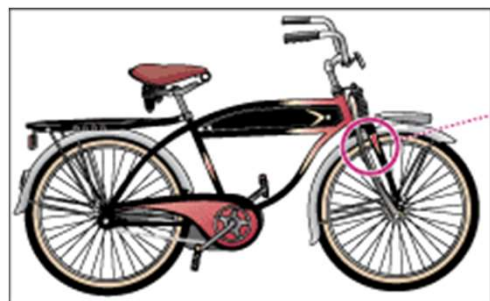


## 点位图与矢量图的比较(知识点)



- 位图图像效果好，放大会失真。
- 1)逼真
  - 现实世界的任何物体，都是由许多微小的单元组成，点阵图由许多微小的像素构成，可以很真实的表现现实世界的任何对象。
- 2)占用的存储容量大
- 3)缺乏灵活性
  - 像素之间无内在联系，如果把图像缩小，再恢复到原来的尺寸时，图像就会变得模糊不清。

## 点位图与矢量图的比较(知识点)



- 矢量图效果差，放大不会失真。
- 1)图形缩放后不变形
  - 以指令集合的形式描的。
- 2)占用的存储容量小
  - 只是存储的图形指令。
- 3)灵活性好
  - 各个部分是相对独立的，局部的处理不会影响其他部分。
- 4)表现能力有一定局限
  - 对于复杂而又不易用数学方法描述的对象，矢量图形就无能为力。显示速度比点位图慢。



## 灰度图

■ 按照灰度等级的数目来划分；只有黑白两种颜色叫单色图像；

- 640×480需要37.5KB的空间；
- (其它的要求计算)





## 关于颜色.....

详细-颜色与色彩空间v课堂版.pptx



## 颜色模型

使用简单方法描述所有颜色的一套规则  
和定义。

- RGB (显示彩色图象用)
- CMY (打印彩色图象用)



## 颜色模型

使用简单方法描述所有颜色的一套规则和定义。

### ■ RGB (显示彩色图象用)

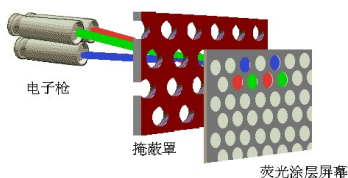
- **有源物体**
- 一个能发出光波的物体称为有源物体，它的颜色由该物体发出的光波决定，使用RGB相加混色模型；

### ■ CMY (打印彩色图象用)

- **无源物体**
- 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CMY相减混色模型。

## RGB模型

相加混色模型

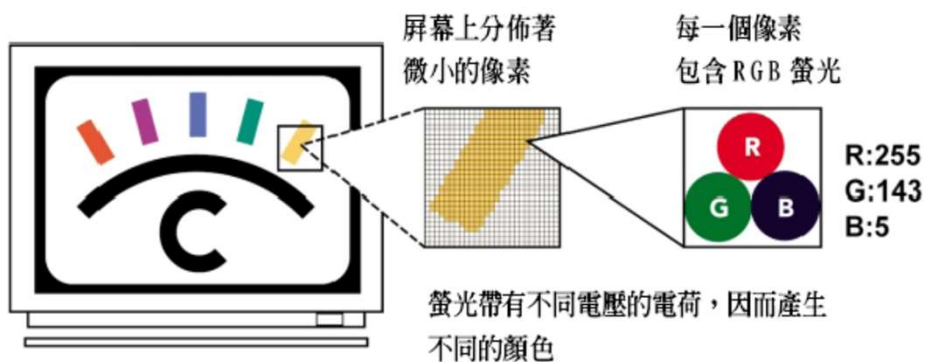


### ■ 有源物体：

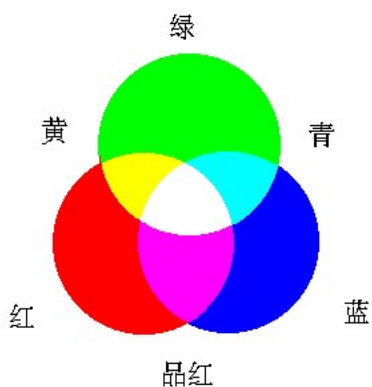
- 一个能发出光的物体，它的颜色由该物体发出的广播决定。
- 电视机和计算机显示器使用的阴极射线管 (cathode ray tube, CRT) 是一个有源物体。



## RGB相加混色模型



只要在屏幕上显示的图像，就是RGB模式表现的。



颜色

$R(\text{红色的百分比}) + G(\text{绿色的百分比}) + B(\text{蓝色的百分比})$ ;





## JPEG压缩编码

### 基本的压缩算法



## 压缩编码

图像基础

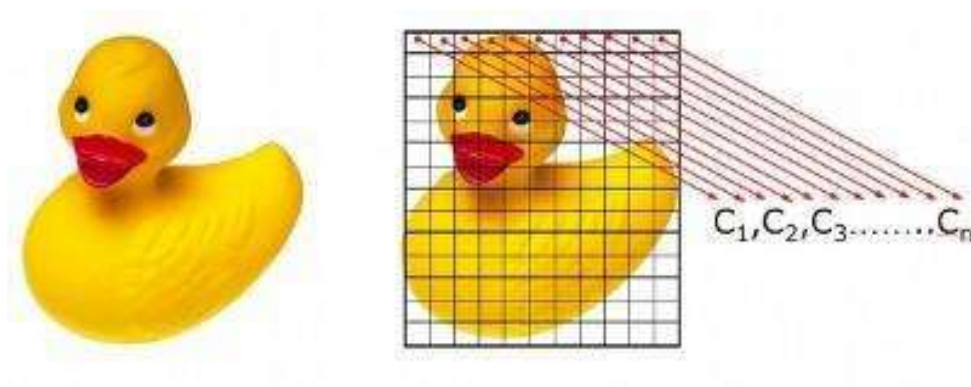
- JPEG专家组开发了两种基本的压缩算法,
  - 一种是采用以离散余弦变换(Discrete Cosine Transform, DCT)为基础的**有损**压缩算法。
    - 利用了人的视觉系统的特性。
  - 一种是采用以预测技术为基础的**无损**压缩算法。
- 为了在保证图像质量的前提下进一步提高压缩比, JPEG 2000(简称JP 2000)标准中将采用小波变换(Wavelet)算法。

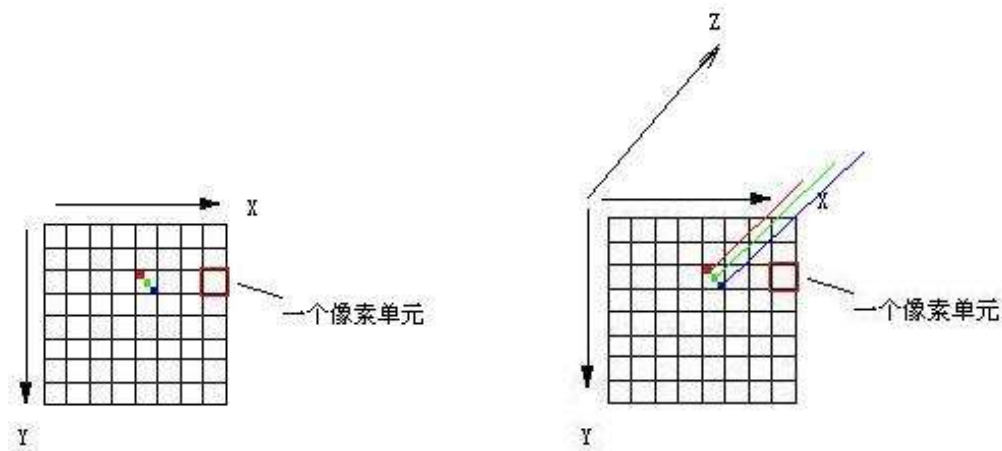




Think

怎么存储数字图像？





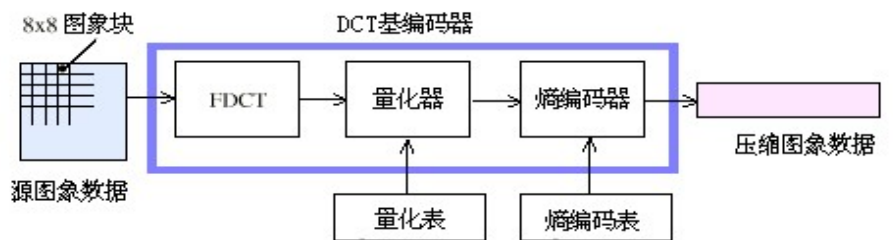
平面的图像，可以理解为除了水平 X 和垂直 Y 以外，还有一个色彩值的 Z 的三维的系统。

## JPEG算法的主要计算步骤

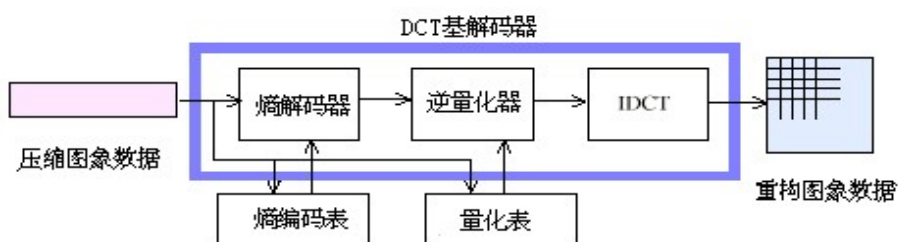
图像基础

### ■ 主要步骤：

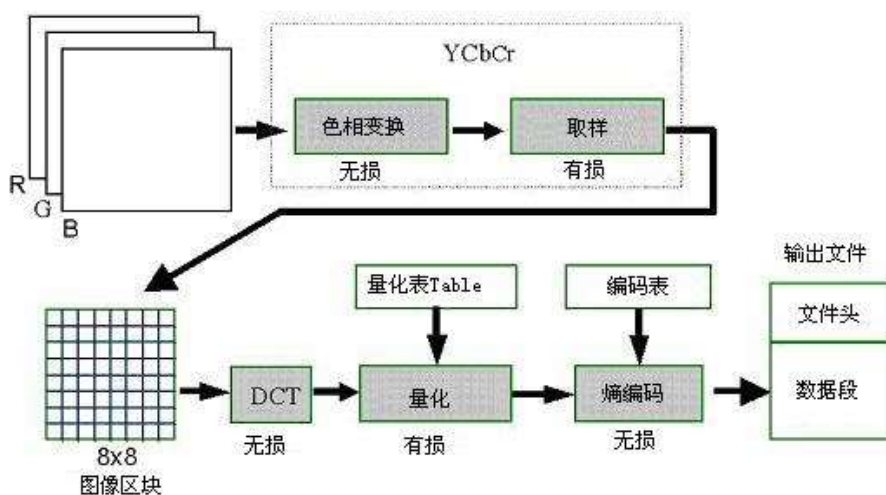
- 1) 正向离散余弦变换(FDCT)。
- 2) 量化(Quantization)。
- 3) Z字形编码(Zigzag Scan)。
- 4) 使用差分脉冲编码调制(Differential Pulse Code Modulation, DPCM)对直流系数(DC)进行编码。
- 5) 使用行程长度编码(Run-length Encoding, RLE)对交流系数(AC)进行编码。
- 6) 熵编码(Entropy Coding)。
- 哪个步骤是有损压缩？每步的原理和作用？



(a) DCT基压缩编码步骤



(b) DCT基解压缩步骤



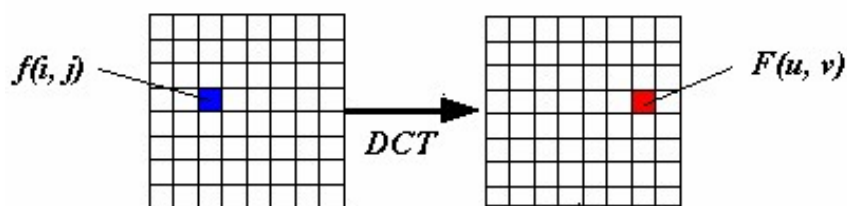


## 1. 正向离散余弦变换

图像基础

### ■ 正向离散余弦变换(FDCT)变换说明。

- (1) 对每个单独的彩色图像分量图像分成 $8 \times 8$ 的图像块，并作为两维离散余弦变换DCT的输入。
  - 通过DCT变换，把能量集中在少数几个系数上。



## 1. 正向离散余弦变换

图像基础

### ■ (2) DCT变换使用下式计算

- JPEG标准并**没有规定**FDCT和IDCT的具体算法。

- 傅氏变换
- Walsh-Hadamard沃尔什哈达玛变换
- 正弦变换
- 余弦变换（应用最广）
- 斜变换
- 哈尔变换
- K-L变换



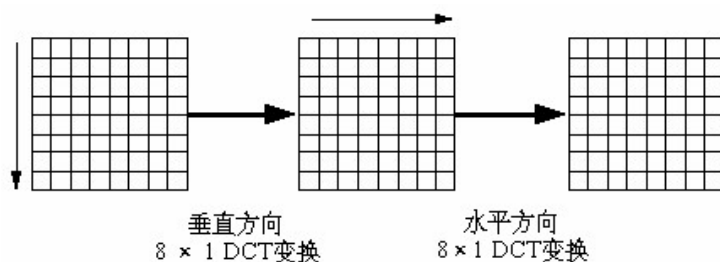
## 余弦型变换(DCT)

图像基础

- (3) 在计算两维的DCT变换时，用下面的计算式把两维的DCT变换变成一维的DCT变换，

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

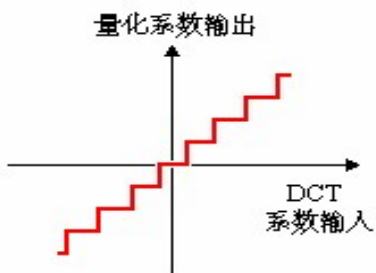
$$G(i, v) = \frac{1}{2} C(v) \left[ \sum_{j=0}^7 f(i, j) \cos \frac{(2j+1)v\pi}{16} \right]$$



## 2. 量化

图像基础

- 对经过FDCT变换后的频率系数进行量化。
  - 目的是减小非“0”系数的幅度以及增加“0”值系数的数目。
  - 量化是图像质量下降的最主要原因。
  - JPEG算法使用均匀量化器进行量化，步距是按照系数所在的位置和每种颜色分量的色调值来确定。





## 图像基础

- 人眼对色差信号不如亮度信号敏感。
  - 1) 量化表分为亮度量化表和色度量化表，对**色度信号**作更大程度的压缩。
  - 2) 量化结果一般是频率低的分量系数大，频率高的分量系数小且大多为0。

色度量化值

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

由于人眼对低频分量的图像比对高频分量的图像更敏感

左上角的量化步距要比右下角的量化步距小。

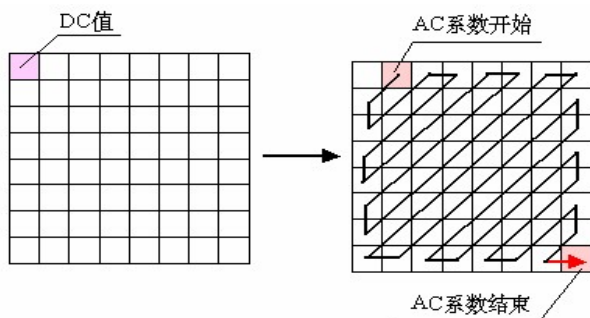
亮度量化值

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	72
24	35	55	64	81	104	113	99
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

- 3) 量化表可由用户规定，但JPEG给出了参考值。

## 图像基础

- Z字形编排
  - 系数重新编排：**增加连续的“0”系数的个数**(“0”游程长度)
  - 方法：
    - 按照Z字形的式样编排。把一个 $8 \times 8$ 的矩阵变成一个 $1 \times 64$ 的矢量，频率较低的系数放在矢量的顶部。



[illegible]



## 4. 直流系数的编码

图像基础

### ■ 直流系数的编码

- 经过DCT变换之后得到的DC直流系数特点:
  - 一是系数的**数值比较大**，二是相邻8 \* 8图像块的DC系数值**变化不大**。
- 根据这个特点: 使用什么编码调制技术?
- **差分脉冲调制编码(DPCM)**技术, 对相邻图像块之间量化DC系数的差值(Delta)进行编码,  
$$\Delta = DC(0, 0)_k - DC(0, 0)_{k-1} \dots\dots (5-5)$$

## 5. 交流系数的编码

图像基础

### ■ 交流系数的编码

- AC系数的特点: 包含有**许多“0”系数, 并且许多“0”是连续的**,
- 根据这个特点: 使用什么编码调制技术?
- 使用非常简单和直观的**游程长度编码(RLE)**对它们进行编码。
  - JPEG使用了**1个字节**的高4位来表示连续“0”的个数, 低4位来表示编码下一个非“0”系数所需要的位数, 跟在它后面的是量化AC系数的数值。



## 6. 熵编码

图像基础

### ■ 熵编码

- 对DPCM编码后的直流DC系数和RLE编码后的交流AC系数作进一步的压缩。
- 在JPEG有损压缩算法中，使用**霍夫曼编码器**来减少熵。
  - 使用霍夫曼编码器的理由是可以使用很简单的查表(Lookup Table)方法进行编码。

## 图像文件格式

图像基础

### ■ BMP

### ■ GIF

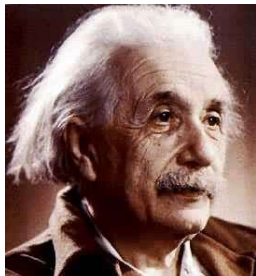
- LZW压缩算法来存储图像数据，定义了允许用户为图像设置背景的透明(Transparency)属性。
- GIF文件格式可在一个文件中存放多幅彩色图形/图像。

### ■ JPEG

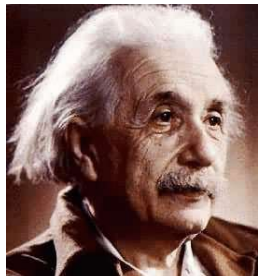
### ■ PNG

- PNG是20世纪90年代中期开始开发的图像文件存储格式
  - 企图替代GIF和TIFF文件格式，同时增加一些GIF文件格式所不具备的特性。
- PNG使用从LZ77派生的无损数据压缩算法。

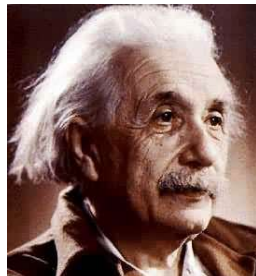




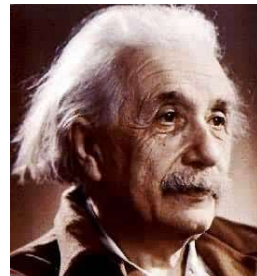
Einstein.bmp  
颜色: 真彩色  
大小: 239.1 KB  
压缩比: 1.0



Einstein.gif  
颜色: 256  
大小: 61.4 KB  
压缩比: 1.3



Einstein.jpg  
颜色: 真彩色  
大小: 11.1 KB  
压缩比: 21.4



Einstein.tif  
颜色: 真彩色  
大小: 238.4 KB  
压缩比: 1.0



## OpenCV

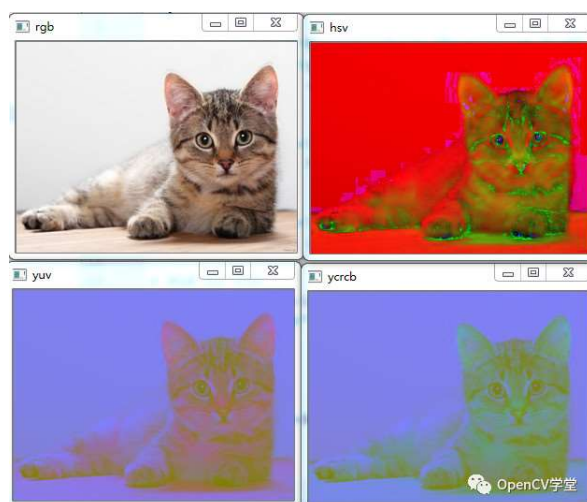
跨平台计算机视觉和机器学习软件库



# 可选作业1 OpenCV

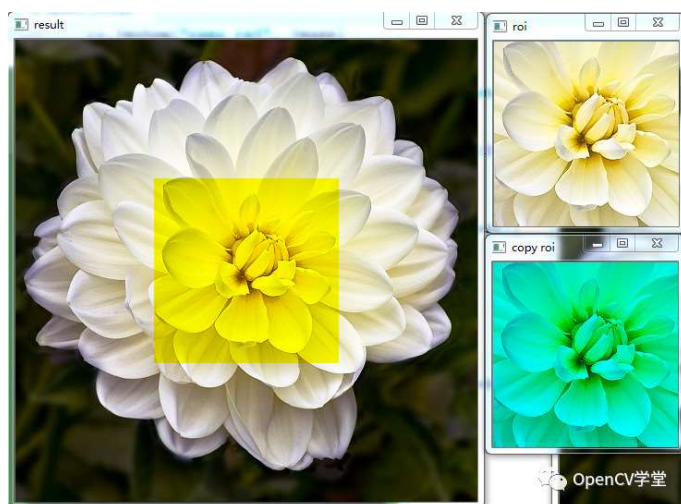
用OpenCV演示一些图片处理相关的程序

## 色彩空间转换





## ROI提取

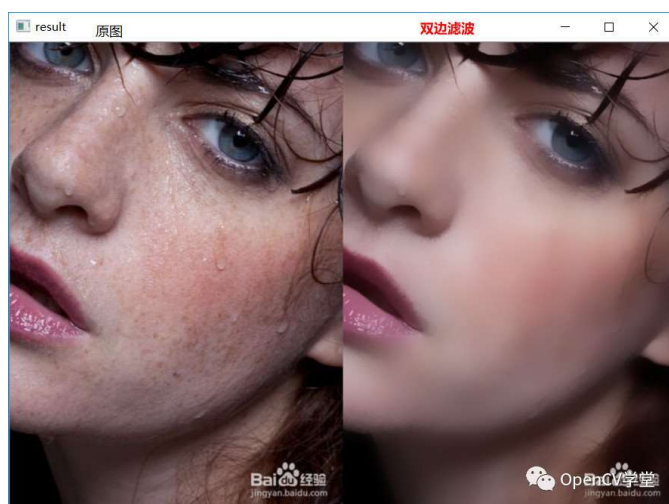


## 不规则ROI提取

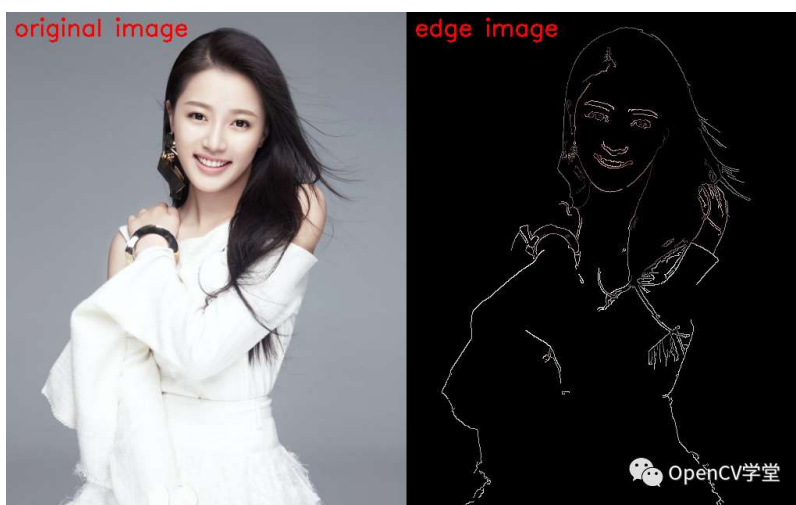




## 图像边缘保留滤波



## 图像梯度与边缘





## 图像锐化增强

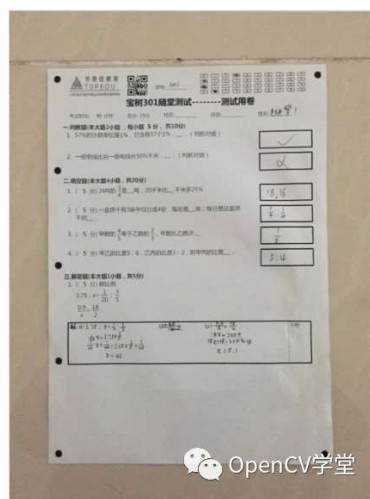


## 二值图像分析



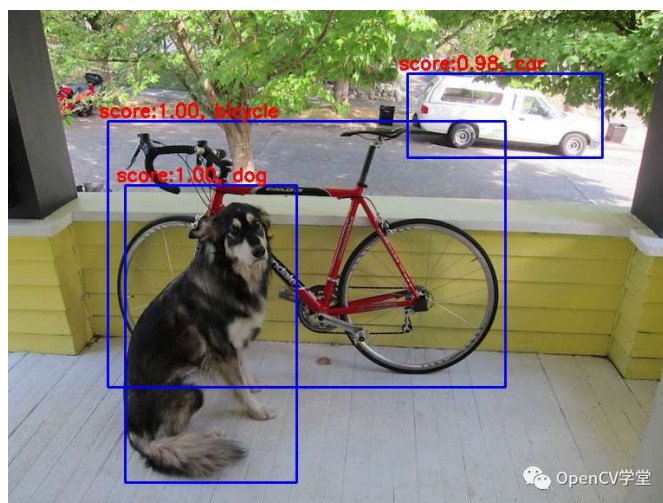


## 二值图像分析



OpenCV学堂

## 视频分析与跟踪 DNN模块

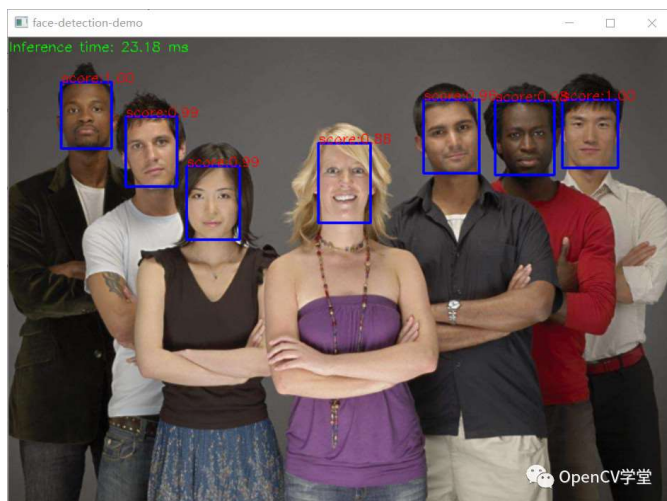


OpenCV学堂

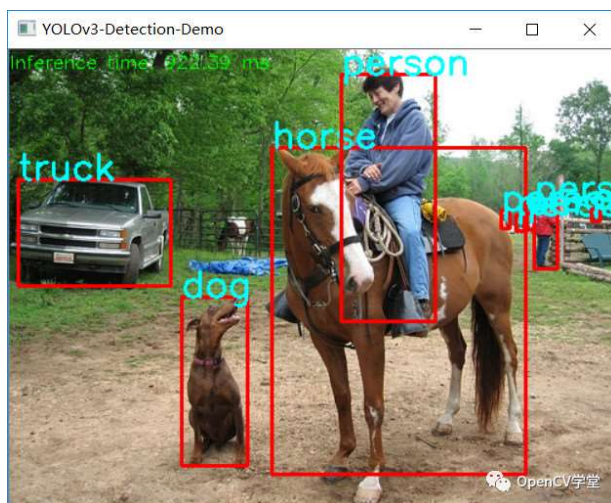




## 视频分析与跟踪 DNN模块



## 视频分析与跟踪 DNN模块





## 本章 小结

### 主要内容

多媒体概念    要点    未来展望

- 数字图像
- 图像的三个基本属性
  - 1 分辨率
  - 2 像数深度(图像深度)
  - 3 正彩色, 伪彩色, 直接色 (略过)
- 图像的种类
  - 矢量图
  - 点位图: 文件字节数=图像分辨率X图像 (色彩) 深度/8
- JPEG压缩编码
- OpenCV





## 一些专题

多媒体概念    要点    未来展望

- 显示技术
- 打印技术
- 您好，色彩！