



计算机视觉

第2章 视觉特征表达

陈飞

福州大学计算机与大数据学院

chenfei314@fzu.edu.cn



本章内容



2.1 人眼视觉

人眼的构造

人眼视觉特性

2.2 图像表达

数字图像

图像质量

2.3 色彩空间

三基色原理

颜色空间

2.4 距离度量

像素间的距离

图像间的距离

2.5 形状特征表达

形状定义

形状距离

形状表达

2.6 纹理特征表达

纹理结构

纹理距离

纹理表达

2.7 多尺度表达

尺度的概念

金字塔结构

2.1 人眼视觉



■ 电磁辐射成像

- 电磁辐射波包括无线电波、微波、红外线、可见光、紫外线、X射线、 γ 射线。
- 电磁辐射波的波谱范围很广：波长最长的是无线电波为 $3 \times 10^2 \text{m}$ ，其波长是可见光波长的几十亿倍；波长最短的是 γ 射线，波长为 $3 \times 10^{-17} \text{m}$ ，其波长比可见光小几百万倍。

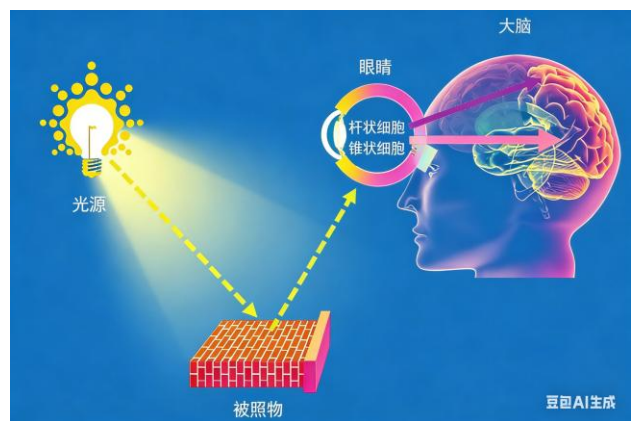
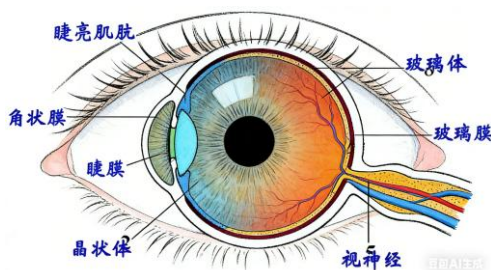
可见光的波长范围通常大约在 380纳米 到 750纳米 之间。

3

2.1 人眼视觉



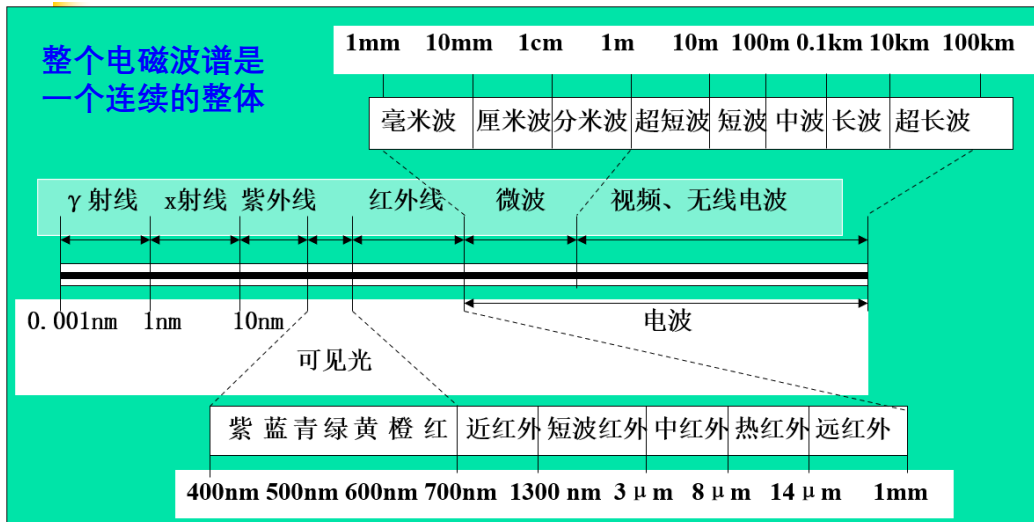
人眼的构造



人眼成像过程：光线 → 角膜 → 瞳孔 → 晶状体 → 视网膜（感光） → 视神经 → 大脑（成像）。

4

2.1 人眼视觉



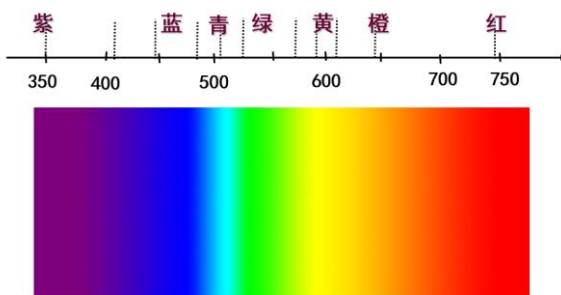
5

可见光成像

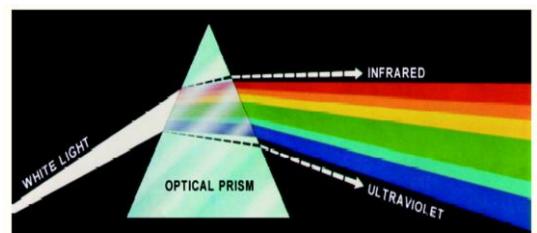


- 太阳的电磁辐射波恰好主要占据整个可见光谱范围（大气光学窗口）。

可见光谱：



大气“过滤”了太阳辐射的紫外和远红外部分，只让能量最集中的可见光波段成为地表的主要辐射来源。



可见光谱

- 可见光成像原理：反射成像。相机接收的是物体反射的太阳光。

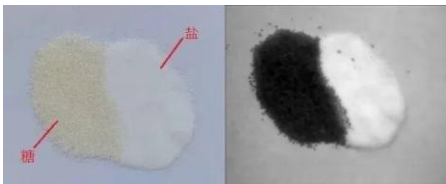
- 特性：最符合人眼习惯，包含丰富的颜色（红绿蓝）和纹理信息。但受光照影响极大（夜晚、阴影下失效）。

6

红外成像



- **近红外**：**反射成像**，但能反映植物叶绿素含量等可见光看不到的信息。
- **热红外**：**辐射成像**。接收的是物体自身发出的热量。因此**不需要外部光源**，可在夜晚工作。
- **特性**：图像反映的是温度分布（通常伪彩色显示）。温度高的物体亮，温度低的物体暗。
- **典型应用**：夜间侦查、体温筛查、工业热故障检测。



短波红外下盐和糖呈现不同的色彩，因为矿物质、人造物质及其他一些地物具有特殊的成分，短波红外能够“看见”这种特有成分。



中波红外



长波红外

7

微波成像



微波属于电磁波谱中**毫米波~米波段**，频率 300MHz~300GHz，波长 1m~1mm。

- **原理**：主动成像。传感器自己发射微波，然后接收地面/接收目标反射回来的信号（雷达）。
- **特性**：**穿透性强**（可穿透云、雨、雾，甚至一定厚度的植被和干沙）。图像反映的是物体的**介电常数和粗糙度**。且因为是主动发射，可**全天时、全天候**工作。
- **典型应用**：合成孔径雷达、遥感卫星、气象雷达。



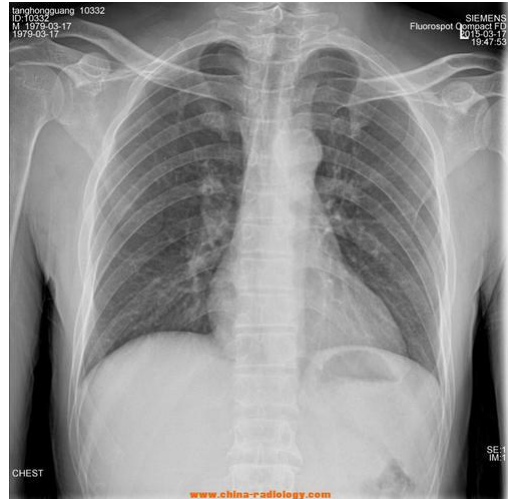
8

X射线与 γ 射线成像



- **原理：**穿透成像。X射线/ γ 射线穿透物体时，会被不同密度的物质不同程度地吸收。
- **特性：**图像是物体的"密度投影"。密度高的区域（如骨骼、金属）吸收射线多，胶片/探测器上呈现为暗区（或亮区，取决于显示方式）。
- **典型应用：**医学CT、安检扫描、工业探伤。

一个区域越亮，说明它吸收的射线越多；越暗，说明它吸收的射线越少（也就是穿透过去的越多）。



9

课堂测试



1、如果你想利用多光谱遥感找出隐藏在茂密树林中的一辆涂有绿色油漆的金属卡车，最不应该单独依赖的是：

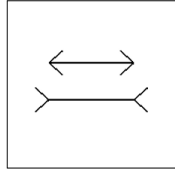
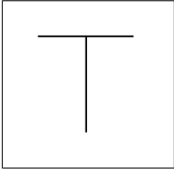
- A 热红外图像（寻找发动机余热） B 激光雷达（LiDAR，获取地面三维形状）
C 高光谱图像（分析物质独特的光谱指纹） D 高分辨率可见光图像（仅依靠颜色）

2、为了躲避可见光和热红外的联合侦察，士兵在篝火旁铺了一张与草地颜色一致的隔热毯。在以下哪种图像中，这张毯子最有可能暴露？

- A 可见光照片（白天） B 可见光照片（夜晚，无月光）
• C 热红外图像 D 雷达图像（如果毯子很薄且平整）

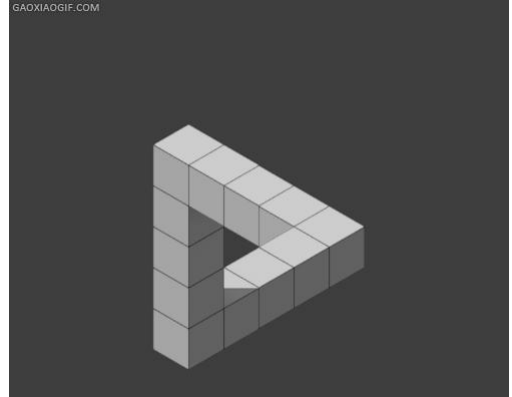
10

视觉错误



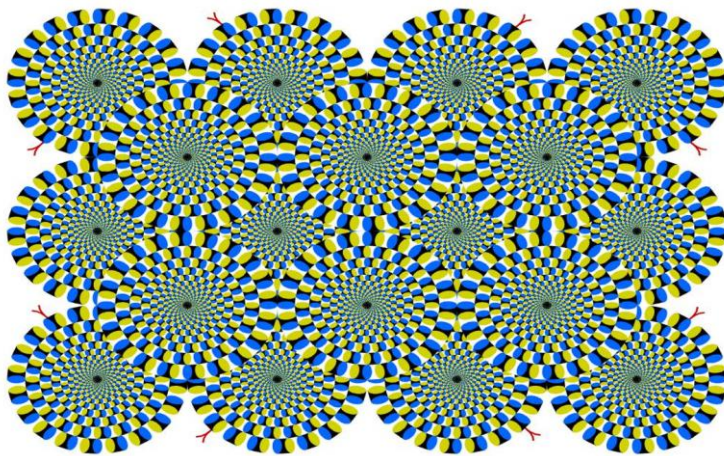
物理现实 ≠ 主观感知

人类视觉系统不是一台被动的摄像机，
而是一个主动的、会根据经验、预期和
上下文进行**解读和推断**的信息处理器。



11

它们真的在动吗？



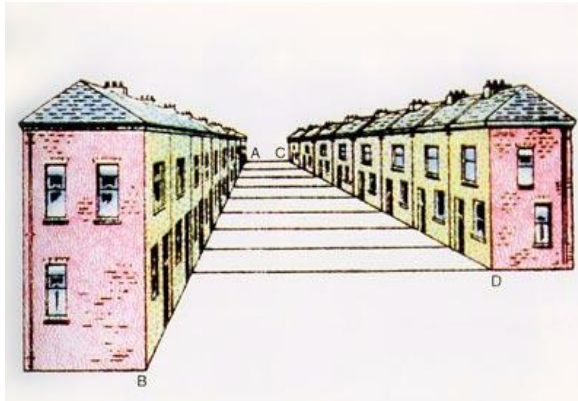
机器视觉与人类视觉的不同！

12

长度与透视

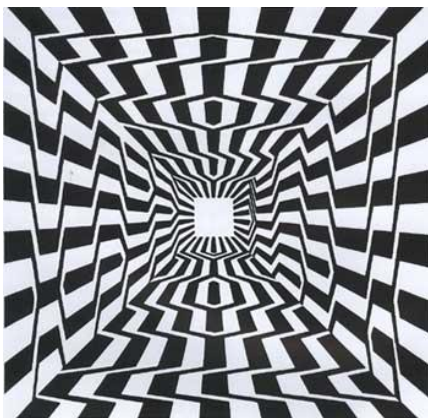


- 长度与透视：线AB和线CD长度完全相等，虽然它们看起来相差很大。



13

曲线正方形：这些是完全的正方形吗？



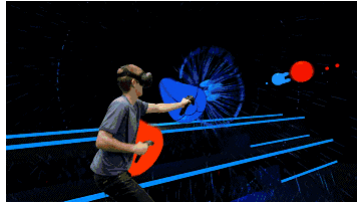
【解析】正方形看起来是变形了，但其实它们的边线都是笔直而彼此平行的。比尔·切斯塞尔创作了这个曲线幻觉的视觉艺术版本。

14

讨论：眼见是否为实？



- 当我们的眼睛“欺骗”了我们，是该相信亲眼所见的事实，还是该相信经过验证的科学解释？这种“错误”究竟是我们感官的缺陷，还是大脑为了适应复杂世界而进化出的高效策略？



- VR（虚拟现实）和AR（增强现实）技术的本质就是在主动“欺骗”我们的视觉。
- 在设计中（如网页UI、建筑空间、服装剪裁），设计师是应该极力避免产生视觉错误，还是应该巧妙地利用视觉错误来达到特定目的（例如让房间看起来更大、让人显瘦）？

15

2.2 图像表达



- 定义：**二维函数 $f(x,y)$** ，其中， x ， y 是空间坐标， $f(x,y)$ 是点 (x,y) 的幅值。
- 灰度图像是一个二维灰度（亮度）函数 $f(x,y)$ 。
- 彩色图像由三个（如RGB，HSV）二维灰度（亮度）函数 $f(x,y)$ 组成。

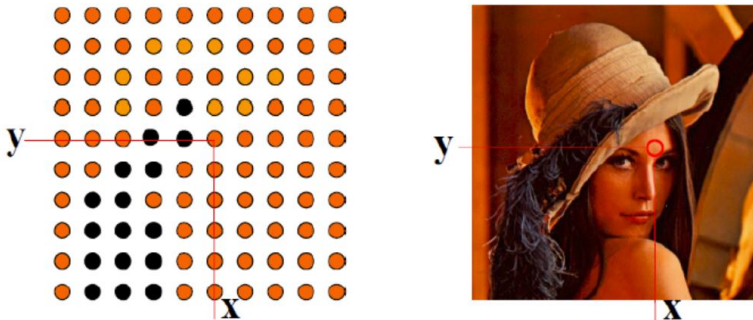


16



2.2.1 数字图像

- 数字图像是由二维像素组成，每一个像素具有一个特定的位置 (x,y) 和幅值 $f(x,y)$ 。
- 数字图像在计算机中通常用**矩阵**来表达。



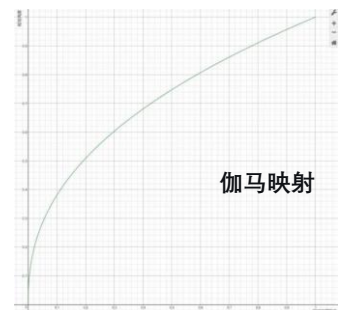
17

2.2.2 图像质量



人眼的非线性感知

- **物理上**：光强从 1 增加到 2（增加了1个单位）。
- **感知上**：感觉亮度只增加了一点点，尤其是在高光区域，这种变化更不明显。
- **主观评价**：观察者是否觉得图像看起来舒服、清晰、自然。这因人而异，不可靠。
- **客观指标**：用数学公式对图像进行量化计算，得到可重复、可比较的数值，但存在局限性。



伽马映射
人眼对于光线的变化感知曲线

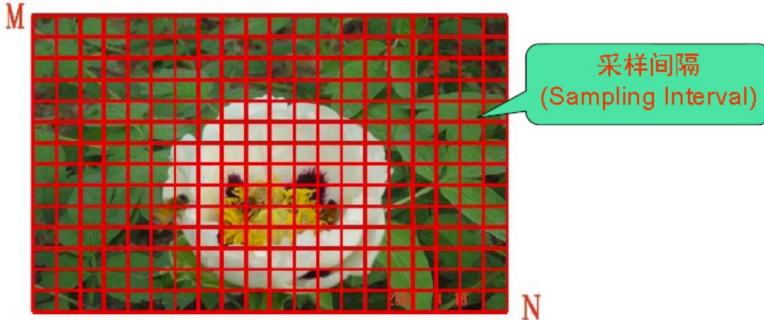
主观感受 vs. 客观指标

18



2.2.2 图像质量

- 图像采样：一个连续图像在**每个采样点处被数字化**。
- 每个采样点对应于数字化图像的一个像素。
- 确定水平和垂直方向上的像素个数N和M。



19

图像采样与图像质量的关系



采样决定了图像细节的上限，是图像质量的基石。

根据**奈奎斯特采样定理**，为了完整复原信号，采样频率必须至少是信号最高频率的两倍。

20

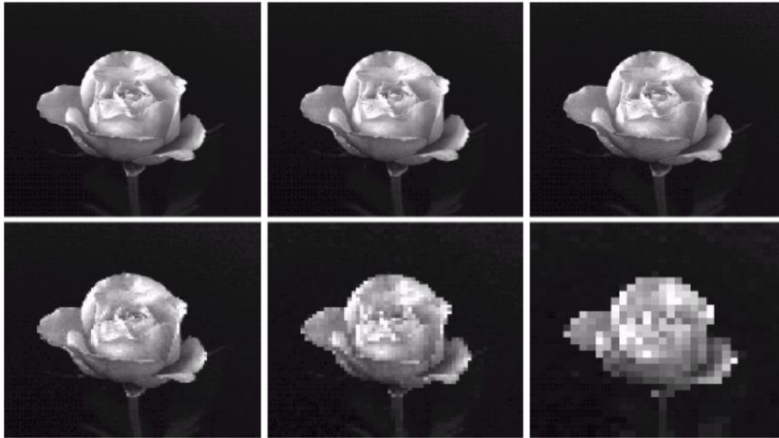
图像质量



1024×1024

512×512

256×256



128×128

64×64

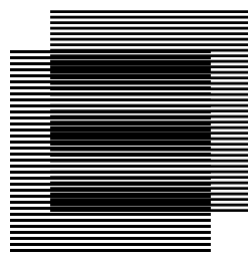
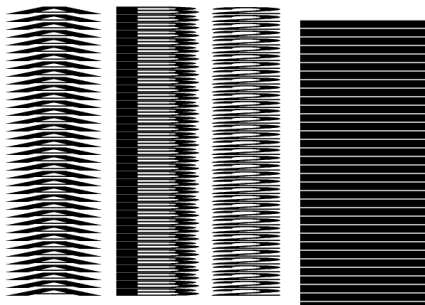
32×32

21

摩尔纹



- 摩尔纹是一种高频信号与低频采样之间发生冲突而产生的光学干涉现象。



• 拍摄电子屏幕：比如拍电脑显示器或电视。屏幕本身有像素点阵，相机也有像素点阵，两个点阵打架，就会出现满屏的彩色摩尔纹。

22

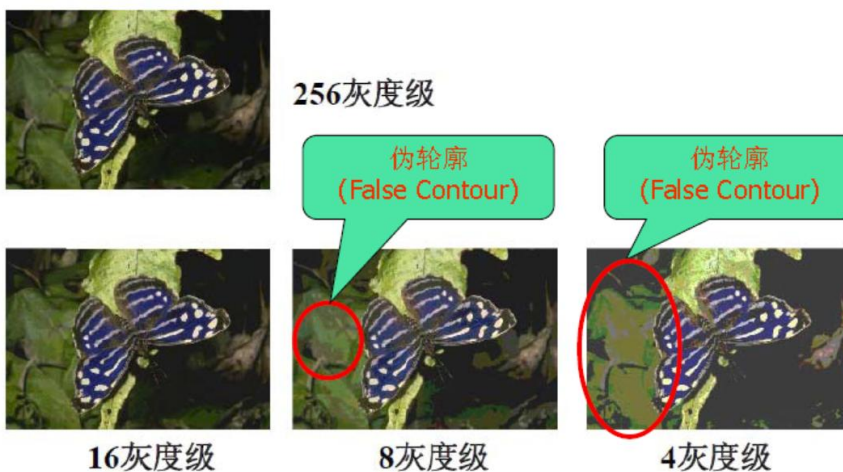
图像量化



- 大部分数字图像处理设备都采用K个等间隔的**量化**方式。
- 对于灰度图像而言，每个像素的亮度用一个数值来表示，该数值范围通常在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(RGB)三元组二维矩阵**来表示。通常三元组的每个数值也是0到255之间，0表示相应的基色在该像素中没有，而255表示相应的基色在该像素中取得最大值。
- **256×256×256种颜色**

23

图像量化与图像质量的关系



24

数字图像的质量：层次



- 灰度级：表示**像素明暗程度的整数量**。
- 像素的量化范围为0~255，就称该图像为256个灰度级的图像。
- 层次：表示图像实际拥有的灰度级的数量。
- 图像数据的实际层次越多，视觉效果就越好。

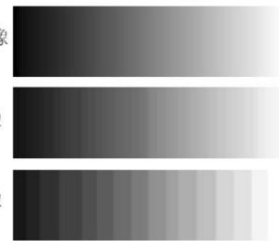
主观上会在边界附近感觉到两条额外的条带：
在亮区一侧感觉到一条更亮的线，在暗区一
侧感觉到一条更暗的线。

人眼的优化（**增强边缘反差**）

256个层次的图像

64个层次的图像

16个层次的图像

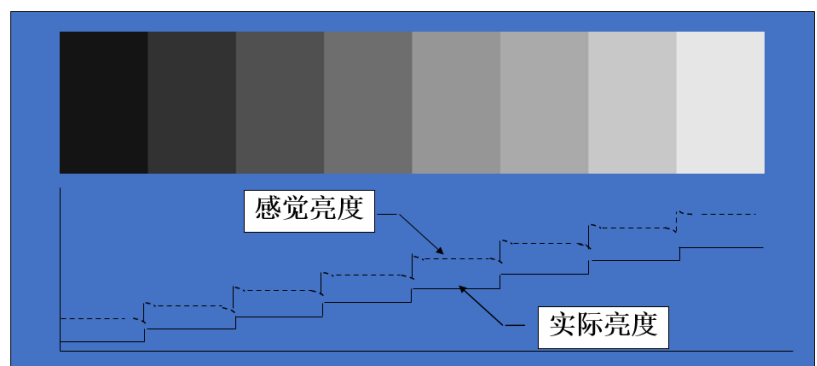


25

马赫带效应



- 人类视觉在亮度有变化的地方会出现**虚幻的亮或暗的条纹**，即在不同亮度区域边界有“欠调”和“过调”现象。



26



数字图像的质量：对比度

- 对比度：是指一幅图像中灰度反差的大小
- 对比度 = **最大亮度 / 最小亮度**



27

数字图像的质量：清晰度



与清晰度相关的因素：

- 亮度
- 对比度
- 尺寸大小
- 颜色饱和度

原图



降低亮度



原图



缩小尺寸



原图



降低颜色饱和度



28

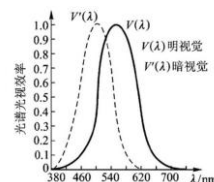
2.3 色彩空间



- 人类视觉可以辨别几千种颜色色调和亮度，相反，**只能辨别几十种灰度层次**。
- 描述彩色光的三个基本量：
 1. **辐射率 —— 纯粹的物理量**：单位面积上、单位立体角内，光源在某个特定方向上辐射出的**功率**。
 2. **光强 —— 与人眼的初步交集**：为了把物理能量（瓦特）变成人眼看到的光，用一个标准曲线对光谱进行加权。这个曲线就是 **CIE 明视觉光度函数**。
 3. **亮度 —— 人眼看到的那一下**：人眼视觉感知的最终结果。描述一个表面看起来有多亮。

辐射率 (Radiance) $\xrightarrow[\text{乘以 } V(\lambda)]{\text{加权}}$ 光强/亮度 (Luminance) $\xrightarrow[\text{非线性}]{\text{伽马编码}}$ 存储的数字值

光强是**点光源**的属性（比如一个发光的灯泡灯丝），而亮度是**扩展光源**或反射表面的属性（比如一张白纸、一个显示器屏幕）



29

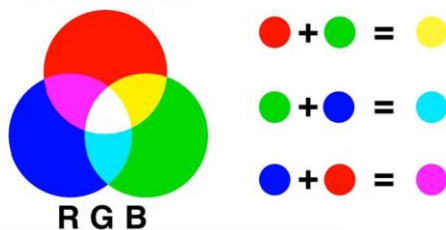
2.3.1 三基色原理



- 实验表明：人眼锥状细胞主要对**三种颜色敏感**：**红、绿和蓝**。这三种颜色被称之为原色(primary colors)。
- 任何颜色都可以看作是三原色融合的结果。

- 红色 + 绿色 = 黄色
- 红色 + 蓝色 = 品红
- 绿色 + 蓝色 = 青色
- 红 + 绿 + 蓝 = 白色

色光三原色



RGB模型：是面向**机器**的，是告诉显示器怎么发光。

30

2.3.2 颜色空间

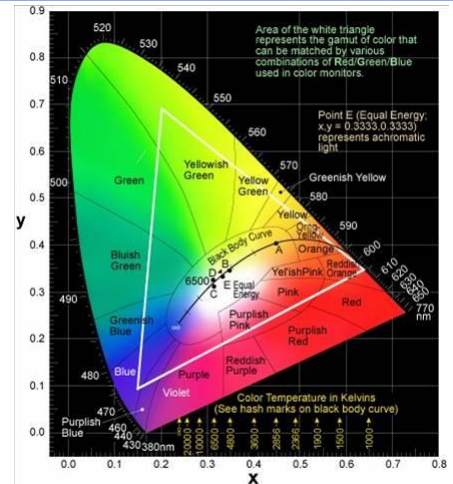


- CIE色度图(第一个标准的色度系统)
- CIE是国际照明委员会的法语缩写。
- 设定每个颜色的三系数之和为1。

$$x = \frac{X}{X+Y+Z}, \quad y = \frac{Y}{X+Y+Z}$$

Y: 代表亮度; X和Z: 代表色度信息。

- 色度图的中心区域有一个白点，表示标准白光。
 - 三角形：三角形的三个顶点分别是显示器红、绿、蓝三颗LED或荧光粉发出的最纯颜色（色度坐标）。
 - 范围：三角形覆盖的马蹄形区域越大，显示器能显示的颜色越多。
- CIE色度图是连接物理光（波长）、生物感知（人眼实验）和工程实践（显示器校准、色彩空间）的桥梁。



31

色彩模型(color model)



- 色彩模型(color model), 或称为色彩空间(color space): 是用标准和普遍的方式来指定颜色的特征表达。
- 色彩模型需要包含一个坐标系统, 以及在该坐标系的一个子空间, 使得每一种颜色都可以表达为该子空间内一个单独的点。
- 图像处理和计算机视觉领域常用的色彩模型:
 - RGB
 - CMY和CMYK
 - HSI
 - CIE-LUV

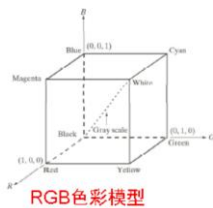
32

RGB模型



RGB模型：是一个三维直角坐标空间，每一维分别代表一种原色，即红(R)、绿(G)和蓝(B)。

灰度 (强度) 图像： $I = (R + G + B) / 3$



全色图像

R分量图像

G分量图像

B分量图像

33

CMY模型



• CMY分别是青色(cyan)、深红色(magenta)和黄色(yellow)的简称。这三种颜色被称为二次色。

- 青色：从白光里吸收（即减去）红色光得到。
- 深红色：从白光里吸收（即减去）绿色光得到。
- 黄色：从白光里吸收（即减去）蓝色光得到。

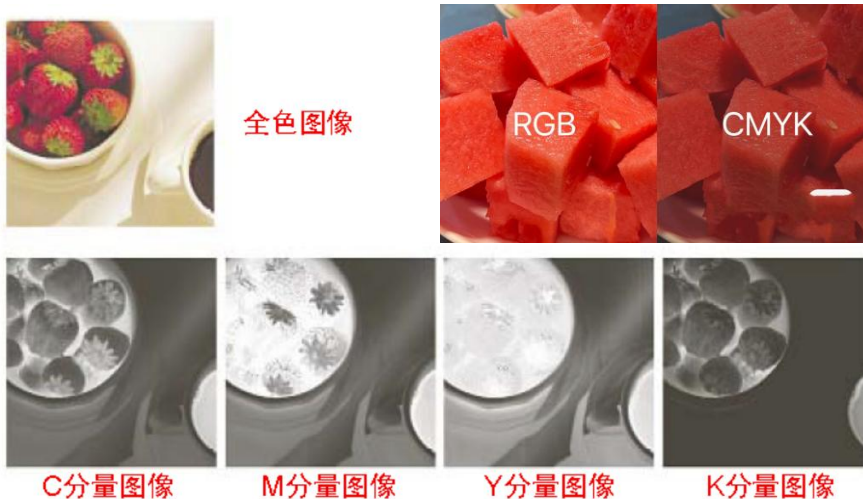
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 255 \\ 255 \\ 255 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- 大多数在纸上沉积彩色颜料的设备，如彩色打印机和复印机，要求输入CMY数据。
- 等量的C、M和Y可以产生黑色，但实际中，这种组合打印出来的黑色是不纯的。

CMYK是工程上的解决方案，为了解决偏色、提高对比度、降低成本而加入了独立的黑色（K）

34

CMYK模型示例



35

HSI模型



- HSI: 色调(Hue)、饱和度(Saturation)和亮度(Brightness)。
- **色调**: 表示观察者接收的**主要颜色**，是混合光波中与主波长有关的属性。
- **饱和度**: 描述了色调被**白光稀释（混合）**的程度，即表达了色调被白光稀释后的纯度。饱和度越高，说明混合的白光越少，即颜色越纯。
- **亮度**: 是一个主观描述子，用来**描述对于颜色的感觉**，其实它是不可测量的。它体现了单色图像中，强度(Intensity)的概念。

HSI模型是面向人的。让我们能像画画一样去调色：**先画素描（亮度I），再决定涂什么颜色（色调H），最后决定颜料挤多少（饱和度S）。**

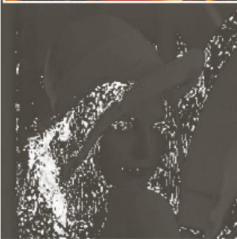
36



HSI模型示例



全色图像



H分量图像



S分量图像



I分量图像

37

RGB模型转换到HSI模型



- 给定一副RGB彩色图像，首先将R、G和B值归一化到[0,1]范围。
- H组件由下式得到：

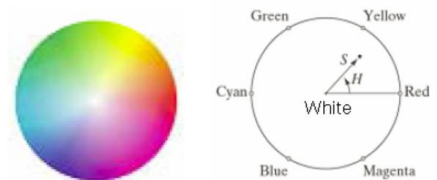
$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad \theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{\left[\frac{1}{4}(R-G)^2 + (R-B)(G-B) \right]^{\frac{1}{2}}} \right\}$$

- H值代表了在HSI空间，该颜色的色调与红色调之间的旋转角度。
- S组件由下式得到：

$$S = 1 - \frac{3}{R+G+B} [\min(R,G,B)]$$

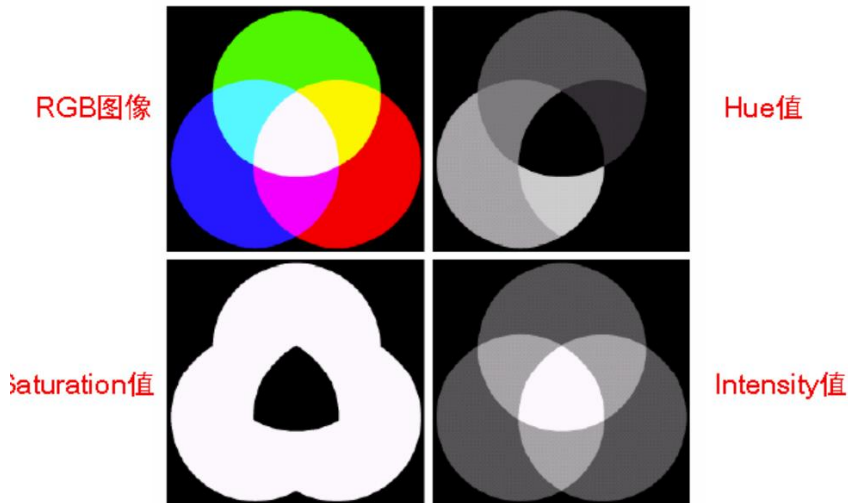
- I组件由下式得到：

$$I = (R + G + B) / 3$$



38

RGB模型转换到HSI模型：示例

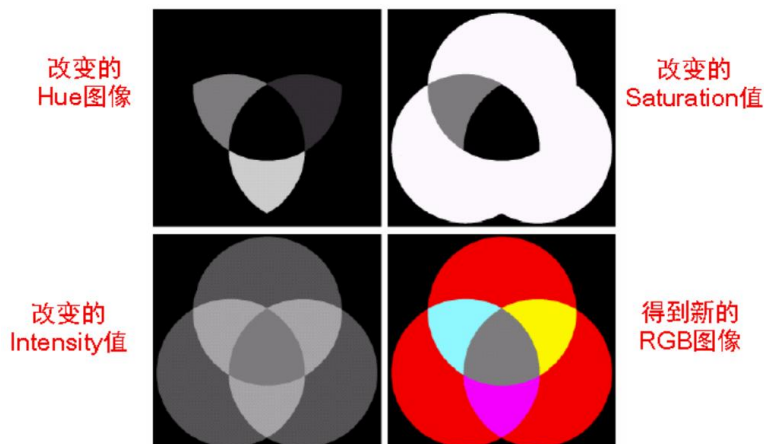


39

HSI模型转换到RGB模型：示例



- 改变上一示例中的HSI值，起到改变图像颜色特征的目的

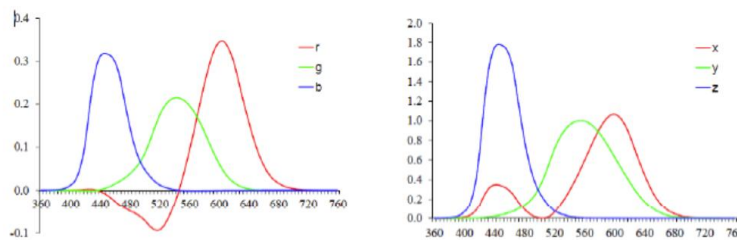


40

CIE-XYZ模型



- CIE在用红绿蓝三原色做颜色标定的时候，发现用红、绿、蓝三种光无法调配出所有的颜色（有些颜色需要把其中一种光加到目标颜色上才能匹配，即出现了负值）。
- Y组件近似于指定光强信息，X和Z指定颜色信息。**CIE-XYZ** 是一种**数学定义的绝对标准**。它完全不依赖于任何设备。只要是人眼（标准观察者）看到的颜色，都能在XYZ色度图上找到一个唯一的坐标点。



41

RGB模型和CIE-XYZ模型的转换



- 给定RGB组件，XYZ组件可以由下式得到：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.2 \\ 0.17697 & 0.8124 & 0.01063 \\ 0 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

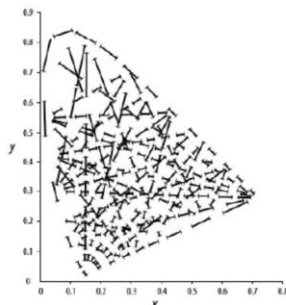
- 给定XYZ模型，RGB可以由上述矩阵的逆矩阵得到。

42

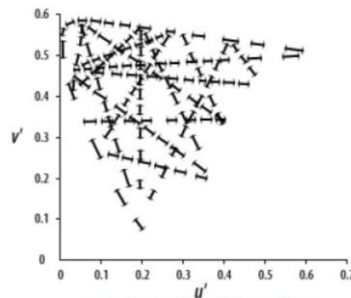
CIE-LUV模型



- XYZ模型存在着很大的分布不均匀性，即相同比例的颜色差，在CIE色图中表现为不同的长度。
- 因此，CIE做了改进，提出了LUV模型，实现均匀的颜色分布。



XYZ模型分布不均匀性



LUV模型分布均匀性

43

投票



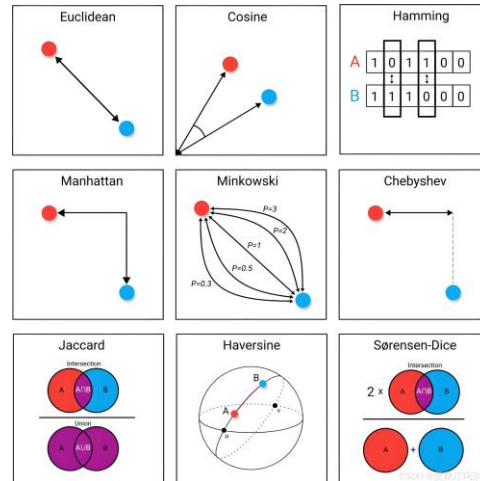
- 背景：美术老师给同学们展示了一条裙子，它在普通的室内白光下看起来是蓝色与黑色相间的。在学校的文艺汇演上，这条裙子被搬上了舞台。为了营造氛围，舞台灯光师只打开了纯红色的灯光（只有红色波长，无绿无蓝）。
- 投票问题：在全红光的照射下，这条裙子原本蓝色的部分和黑色的部分，分别会呈现什么颜色？
 - 蓝色部分看起来是黑色的，黑色部分看起来也是黑色的。
 - 蓝色部分看起来是红色的，黑色部分看起来是黑色的。
 - 蓝色部分看起来是黑色的，黑色部分看起来是红色的。
 - 蓝色部分看起来是紫色的，黑色部分看起来是暗红色的。

44



2.4 距离度量

- 像素间的一些基本关系
- 相邻像素：
 - 4邻域
 - D邻域
 - 8邻域
- 连通性
 - 4联通
 - 8联通
- 距离测量

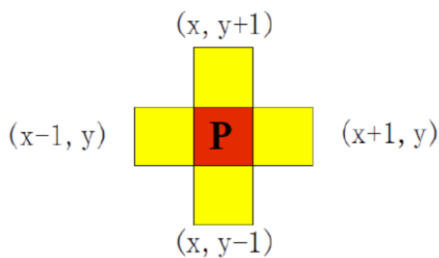


45

相邻像素：4邻域



- 4邻域：像素 $p=(x, y)$ 的4邻域是：
 $(x+1, y)$; $(x-1, y)$; $(x, y-1)$; $(x, y+1)$



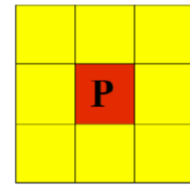
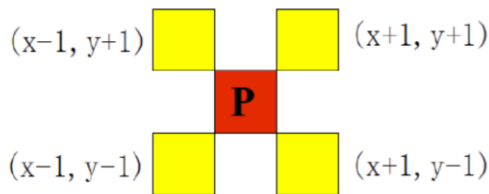
46

相邻像素：D邻域



- D邻域：像素 $p=(x,y)$ 的D邻域是其对角上的点：
- $(x+1,y+1)$; $(x+1,y-1)$; $(x-1,y+1)$; $(x-1,y-1)$

相邻像素：8邻域



47

2.4.1 像素间的距离

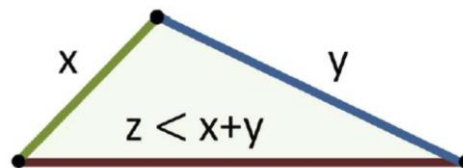


- 像素间的距离的定义

对于像素 p , q 和 r , 如果测量 D 满足以下三条特性:

- (1) 同一性: $D(p, q) \geq 0$, 而且 $D(p, q) = 0$, 当且仅当 $p=q$ 。
- (2) 对称性: $D(p, q) = D(q, p)$
- (3) 三角不等性: $D(p, r) \leq D(p, q) + D(q, r)$

则称 D 是距离测量。



48

像素间的距离：欧式距离



- 像素 $p=(x, y)$ 和 $q=(s, t)$ 间的欧式距离(Euclidean Distance) 定义如下：

$$D_E = \sqrt{(x-s)^2 + (y-t)^2}$$

- 优点：直观。
- 缺点：平方根的计算费时，且数值不为整数。

49

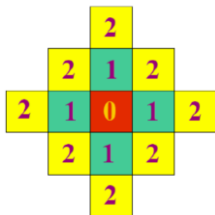
像素间的距离：城市距离 (D_4 距离)



- 像素 $p=(x, y)$ 和 $q=(s, t)$ 间的城市距离定义如下：

$$D_4 = |x-s| + |y-t|$$

- 含义：在数字栅格中，如果只允许横向和纵向移动，城市距离表示从起点移动到终点所需的最少的步数。



到某像素的城市距离小于或者等于某个值的那些像素形成了一个菱形。

50

像素间的距离：棋盘距离 (D_8 距离)



- 像素 $p=(x, y)$ 和 $q=(s, t)$ 间的棋盘距离定义如下：

$$D_8 = \max\{|x-s|, |y-t|\}$$

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

- 含义：在数字栅格中，如果允许横向、纵向和对角线移动，棋盘距离表示从国王在棋盘上从一处移动到另一处所需的步数。
- 棋盘距离特性：到某像素的棋盘距离小于或者等于某个值的那些像素形成了一个正方形。

51

向量的长度（范数）



向量的长度（也称为范数）是衡量向量大小的标量值。对于一个 n 维向量 $\mathbf{v} = (v_1, v_2, \dots, v_n)$ ，其长度通常通过欧几里得范数（2-范数）计算，公式为：

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \quad \|\mathbf{v}\|_0 = \text{非零元素的个数}$$

- 1-范数（曼哈顿范数）：

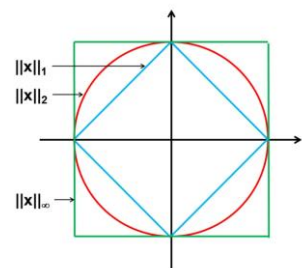
$$\|\mathbf{v}\|_1 = |v_1| + |v_2| + \dots + |v_n|$$

2. p -范数：

$$\|\mathbf{v}\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{1/p}$$

3. 无穷范数：

$$\|\mathbf{v}\|_\infty = \max(|v_1|, |v_2|, \dots, |v_n|)$$



52

范数正则化



1-范数 (列和范数) :

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|$$

(列元素绝对值之和的最大值)

无穷范数 (行和范数) :

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

(行元素绝对值之和的最大值)

2-范数 (谱范数) :

$$\|A\|_2 = \sigma_{\max}(A)$$

(A 的最大奇异值)

Frobenius 范数 (F-范数)

Frobenius 范数是矩阵元素的平方和的平方根, 类似于向量的 2-范数。对于矩阵 $A \in \mathbb{R}^{m \times n}$, 其定义为:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$$

53

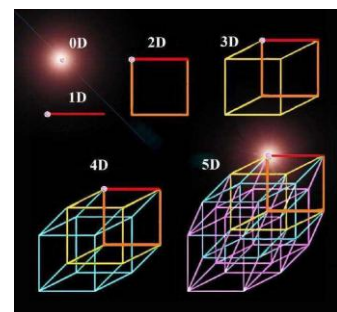
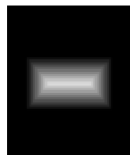
2.4.2 图像间的距离



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



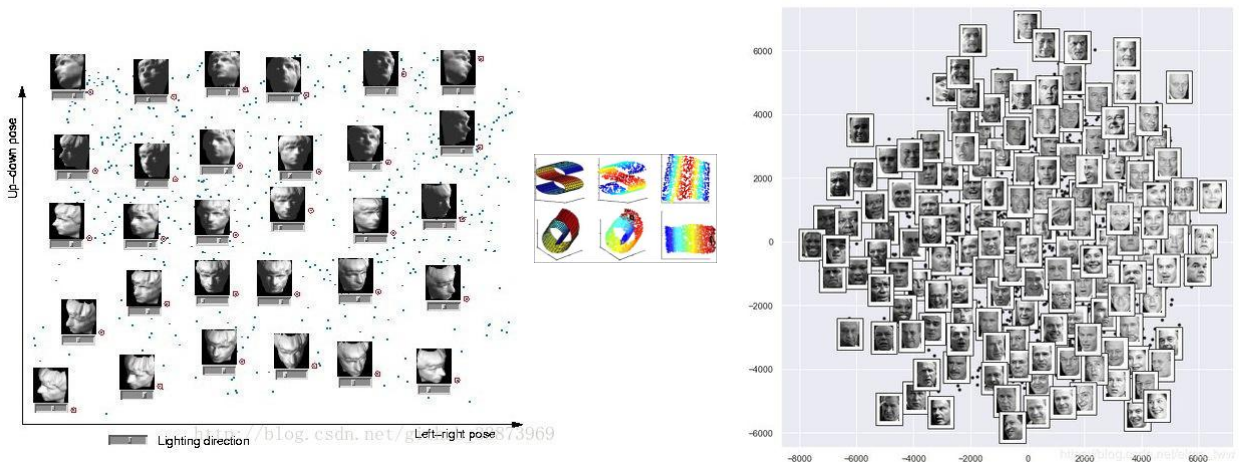
0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	0
0	1	2	2	2	2	1	0
0	1	2	3	3	2	1	0
0	1	2	2	2	2	1	0
0	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0



- 1.物理距离: 图像像素矩阵的数值差异。
- 2.感知距离: 人类大脑处理视觉线索后感受到的差异。
- 3.语义距离: 图像在概念层面的相似度。

54

降维



55

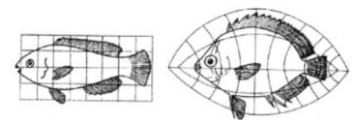
2.5 形状特征表达



形状特征：用来表达物体形状的一种描述。

构建形状特征的途径：

1. **基于边缘信息**，构建关于物体轮廓的表达。
例如，形状上下文(Shape Context)描述，傅立叶算子描述等。
2. **基于区域内像素的统计信息**，构建物体形状的整体表达，
例如，矩(Moments)，凸包(Convex Hull)等。



形状特征描述的要求：

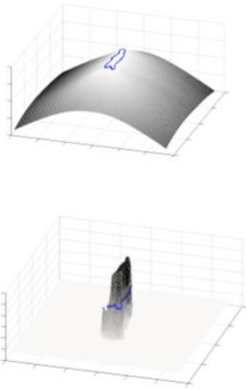
1. **具备一定的局部描述能力**：适用于物体部分被遮挡的情况。
2. **对平移、旋转、尺度变换等具有一定鲁棒性**。

56

2.5 形状特征表达



1. 形状概率表示vs符号距离函数



12个样本形状



形状概率表示

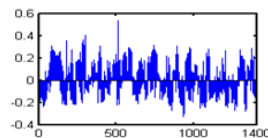
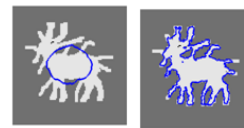
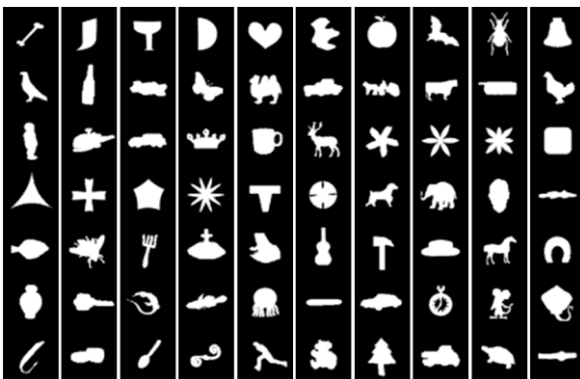


符号距离函数



57

2.5.1 形状定义



58

形状上下文描述



- 形状上下文(shape context)描述包含了两个层次：
- 形状表达：物体的**外部轮廓和内部轮廓上的一组离散的点**。
- 形状上下文：描述了每一个点与其它点之间的关系。这些关系构成了对于物体形状特征的描述。

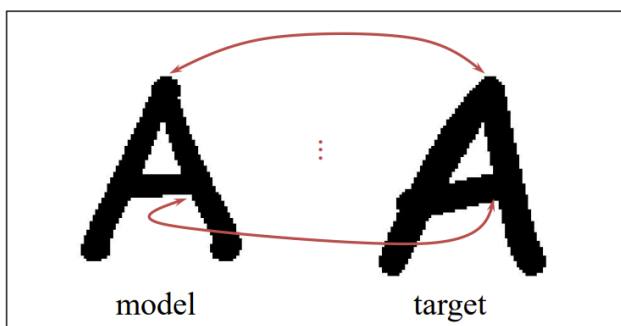


59

2.5.2 形状距离



- 动机：如何度量两个二值图像中的目标相似性

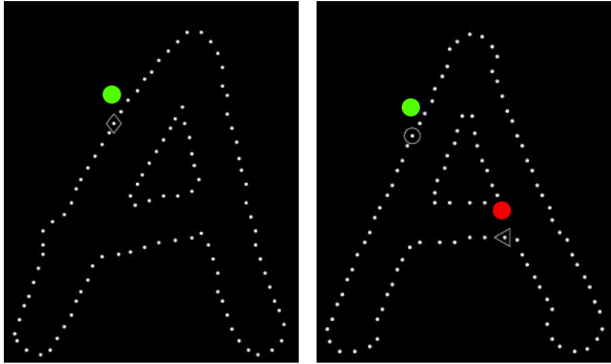


60

2.5.2 形状距离



- 提取二值图像中的物体轮廓，比较轮廓点的相似性，以及匹配关系

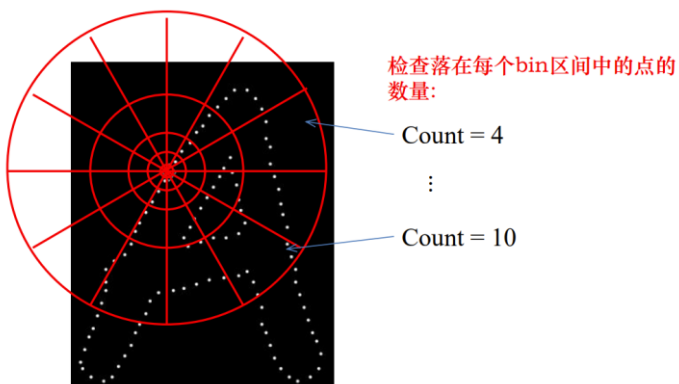


61

2.5.2 形状距离



- 以某一个轮廓点为参考原点，统计其他轮廓点的分布

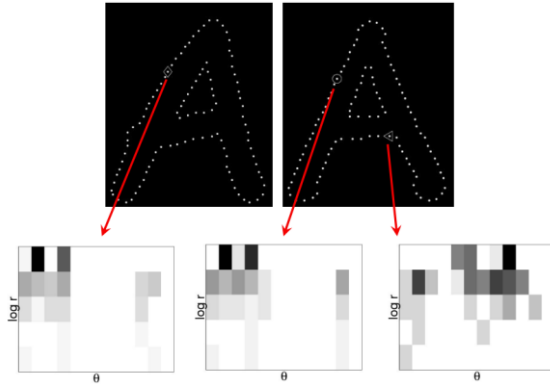


62



2.5.2 形状距离

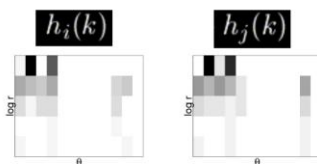
- 极坐标空间划分涉及两个变量 r 和 θ ，轮廓点分布的直方图可以用二维矩阵表示



- 对**平移变换**具有不变性
- 通过选择合适的半径，可对**缩放变换**具有不变性
- 通过选择局部切向作为bin编号的参考方向，可对**旋转**具有不变性
- 可容忍小的仿射畸变

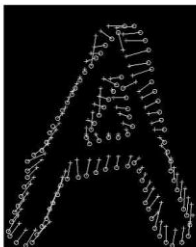
63

2.5.2 形状距离



基于Chi Squared distance, 计算两个轮廓的距离: :

$$C_{ij} = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$



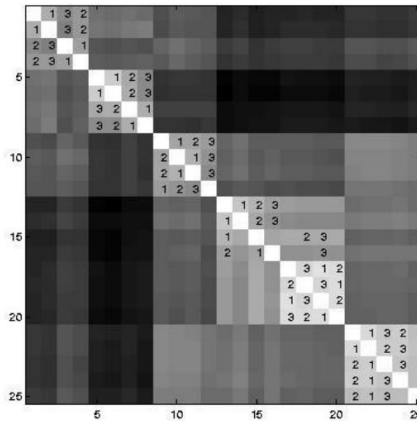
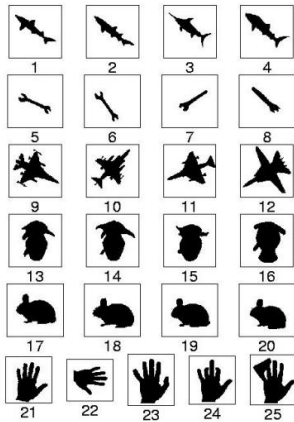
- 利用匹配代价 $C_{i,j}$, 通过求解线性分配问题, 得到两组轮廓点的对应关系:

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)})$$

二分图匹配算法: 匈牙利算法

64

2.5.2 形状距离



混淆矩阵图

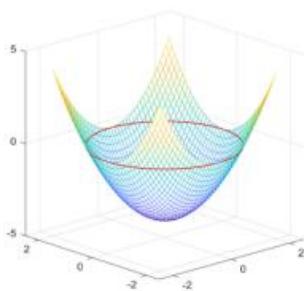
65

2.5.3 形状表达



• 水平集

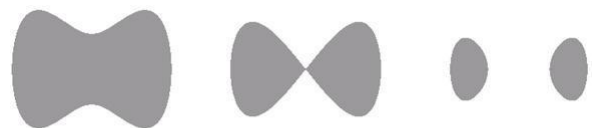
用一个高一维的函数的某个“水平面”（通常是零平面）的截线，来隐式地表示一个低一维的形状。



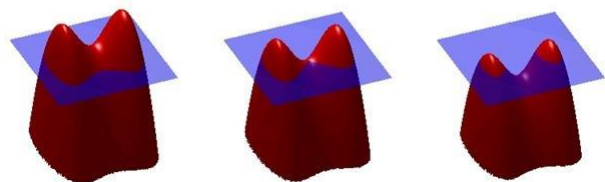
$$z = f(x, y)$$

$$\{(x, y) \mid z = f(x, y) = 0\}$$

稳定的数值计算与亚像素精度



自然地处理拓扑结构变化



66

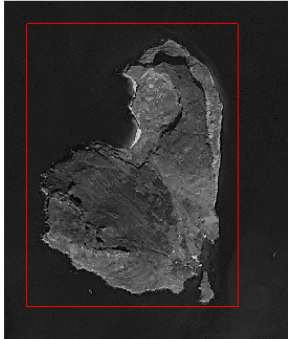
水平集方法



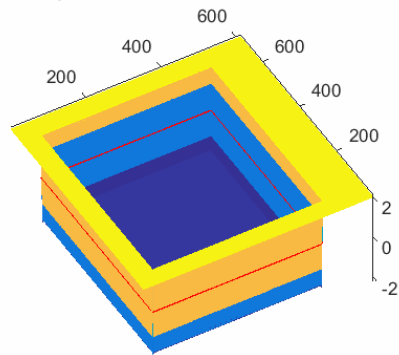
偏微分方程:

$$\frac{\partial \phi}{\partial t} = F|\nabla \phi|$$

contour result , iteration number=1



phi result , iteration number=1



67

2.6 纹理特征表达



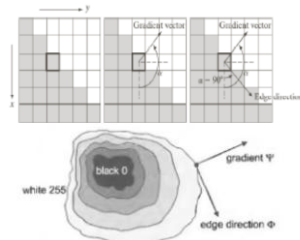
• 边缘特征的表达

以单个像素为单位，用图像函数在该像素邻域范围内的特性来计算。每个边缘点由一个具有**位置(x-y坐标)**、**幅值(magnitude)**和**方向(direction)**的向量来表达。

- **位置**: 表示边缘点的空间位置。位置的确定是一个边缘点定位的过程。
- **幅值**: 表示边缘点上亮度变化的程度。
- **方向**: 表示边缘点上边缘的方向。

梯度方向: 即亮度变化的方向(边缘法向)

边缘方向: 梯度方向旋转90度的方向。

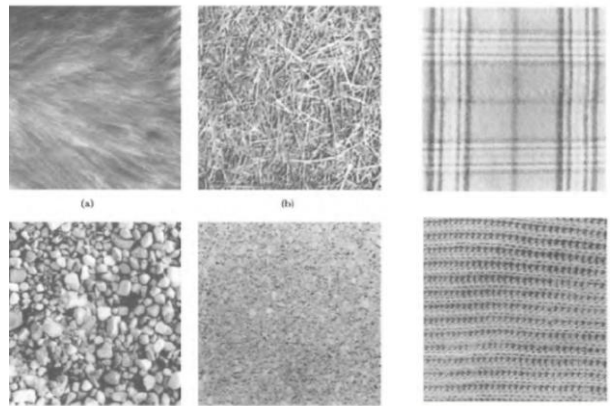


68

2.6.1 纹理的概念



- **纹理(texture)**: 是物体表面的一种属性。它是由**区域内**的一组元素构成, 这些元素在**亮度、颜色、边缘方向、形状、尺寸、空间联系(结构)**等属性上表现出一定的内在关系, 例如相似性、周期性重复出现等。
- 这些元素被称为**纹理基元(texture primitive)**。



69

纹理结构



- 纹理基元通常是一个**具有某种颜色或者区域属性的像素集合**。
例如, 茅草纹理中, 一个基元对应于茎秆; 在狗毛纹理中, 一个基元对应于一簇绒毛; 花格纺织纹理中, 一个基元对应于一个纺织格子或者编制条纹。
- 由纹理基元组成的纹理特征, 是**对整个区域的一个整体特征描述**, 而不是针对某个特征点及其邻域的特征描述。
- 不同基元在**颜色、大小、形状、空间关系**等方面的不同组合, 可以形成不同的纹理特征。

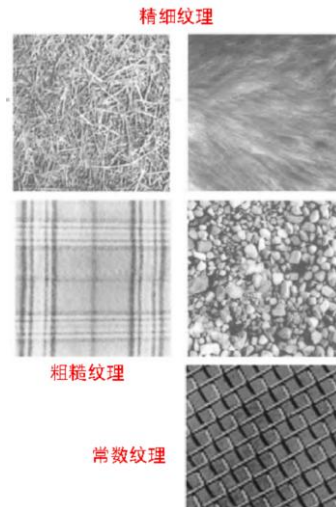


70



2.6.2 纹理的种类

- **精细纹理**：如果纹理基元小、并且相邻基元之间的色调等特征相差很大，则产生精细纹理。
- **粗糙纹理**：如果纹理基元大，包含了很多像素，则产生粗糙纹理。
- **常数纹理**：每一个或者每组基元的集合在属性上呈现恒定的、或近似周期性的变化。即这些属性按照一定的频率重复出现。



71

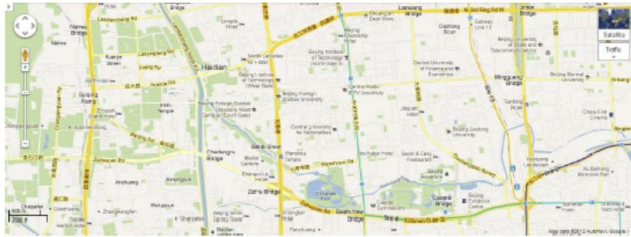
2.6.3 纹理表达



- 由于纹理特征是由**区域内基元**在各方面的属性（例如亮度、颜色、边缘方向、空间排列关系等）体现出来的，所以纹理特征可以由这些属性的组合来表达。这意味着，在提取出上述部分基本特征后（通常由滤波器来提取这些基本特征），重点在于**如何利用这些基本特征来构建纹理描述**。
- 作为一个区域的整体描述，纹理特征也可以用一些整体测量来描述：
 - (a) **粗糙度**：主要描述基元的大小、基元之间的空间间隔等。
 - (b) **方向性**：主要描述纹理基元自身的方向(orientation)以及它们的空间排列方向。

72

2.7多尺度表达



尺度为1:500

细节信息少，但
是全局信息多



尺度为1:200

细节信息多，但
是全局信息少

73

2.7.1 尺度的概念



- **尺度(scale)**: 它是一种测量标准，该标准决定了一个测量单位所能代表的信号尺寸。
- 对于图像信号而言，尺度决定了在一个观测方式下得到的图像中，**一个像素所代表的物体大小**。
- **尺度与分辨率的概念**很接近。但是，尺度是针对人的观测方式而言，是由观测者主动设定；而分辨率既包含了主观设定的概念，又包含了客观决定的能力，即针对图像采集系统固有的采样能力而言。

74

图像的多尺度表达



- 构建**图像的多尺度表达**的原因：
 1. **物体在不同的观测尺度下会有不同的表现方式**。因此，给定一幅图像，构建它的多尺度表达，可以展现出该图像在不同尺度下的特性，从而为选取一个合适的尺寸来提取具有尺度不变性的特征提供基础。
 2. **任意给定两幅图像，起初我们并不能确定这两幅图像之间的尺度关系**。因此，构建它们各自的多尺度表达，为提取各自的特征，从而进行匹配，提供了基础。
- 构建一幅图像的多尺度表达的同时，也就构建了一个**尺度空间 (Scale-space)**。该空间由2维的空间轴和1维的尺度轴构成。

75

图像多尺度表达的构建方法



- 研究表明，**高斯核 (Gaussian kernel)**，即高斯低通滤波器，可以生成图像的尺度空间，称为**高斯尺度空间 (Gaussian scalespace)**。
- 用于生成尺度空间的高斯核可以表达为：

$$g(\sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$



只有标准偏差 σ 一个参数， σ 代表了尺度空间中的尺度坐标。

76

图像多尺度表达的构建方法



• 高斯核作为的尺度空间生成器的原因：

1. **保持图像信息的完整性**。当使用高斯核来平滑图像时，没有任何附加的、人造的结构产生，即得到的粗糙尺度图像只是其细尺度图像的简化表达。该特性被称为**因果性(causality)**。

2. **Semi-group(半群)特性：高斯核的可叠加性、尺度空间的连续性**。连续n次使用小尺寸高斯滤波器平滑一幅图像，与使用一次大尺寸高斯滤波器平滑的结果一样，它们的尺寸关系为：

$$g(\sigma_1) * \dots * g(\sigma_n) * f(x, y) = g(\sigma_1 + \dots + \sigma_n) * f(x, y)$$

77

图像多尺度表达的构建方法



• 给定一幅图像 $f(x, y)$ ，该图像构成**第0层尺度**（即**最精细的尺度**）；然后，高斯核与第0层图像卷积，得到第1层尺度图像；接着，第1层图像再与高斯核卷积，得到第2层尺度图像，从而形成一组尺度逐级粗糙的图像，最终构成原始图像的多尺度表达。

• 多尺度表达定义为**以空间位置 x 和 y 以及尺度 σ 为自变量的函数**：

$$L_0(x, y, 0) = f(x, y)$$

$$L_1(x, y, \sigma) = g(\sigma) * L_0(x, y, 0)$$

$$L_2(x, y, 2\sigma) = g(\sigma) * L_1(x, y, \sigma)$$

$$\vdots$$

$$L_n(x, y, n\sigma) = g(\sigma) * L_{n-1}(x, y, (n-1)\sigma)$$

$$L_0(x, y, 0) = f(x, y)$$

$$L_1(x, y, \sigma_1) = g(\sigma_1) * L_0(x, y, 0)$$

$$L_2(x, y, \sigma_2) = g(\sigma_2) * L_1(x, y, \sigma_1)$$

$$\vdots$$

$$L_n(x, y, \sigma_n) = g(\sigma_n) * L_{n-1}(x, y, \sigma_{n-1})$$

78

图像多尺度表达的构建方法



- 空间域二维高斯核的离散化
- 1. **高斯函数的可分离性**。将二维高斯核与图像的卷积变为两个一维高斯核分别与图像沿着x轴和y轴的卷积。

$$g(x, \sigma) = \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \right], \quad g(y, \sigma) = \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) \right]$$

$$g(x, y, \sigma) = g(x, \sigma) \times g(y, \sigma)$$

$$L_x(x, y, \sigma) = g(x, \sigma) * f(x, y) \quad \text{沿着x轴卷积}$$

$$L(x, y, \sigma) = g(y, \sigma) * L_x(x, y, \sigma) \quad \text{沿着y轴卷积}$$

因此，只需要在空间域设计一维离散高斯核即可。

79

图像多尺度表达的构建方法



- 空间域二维高斯核的离散化
- 2. 根据高斯分布，计算**高斯核的离散值**，即高斯滤波器每一项系数。下表是**标准偏差 $\sigma=1$ 的高斯分布**函数值。为了保证图像的均匀灰度区域不受影响，平滑滤波器要求所有系数之和为1。因此，标准偏差为1的一维高斯核(大小为5)可以表达为： $g(x)=[0.05, 0.25, 0.4, 0.25, 0.05]$
- 相对应的**5×5二维高斯核**为： $g(x, y)=g(x)^T g(y)$ 。

x	-2	-1	0	1	2
$g(x; \sigma=1)$	0.054	0.242	0.398	0.242	0.054

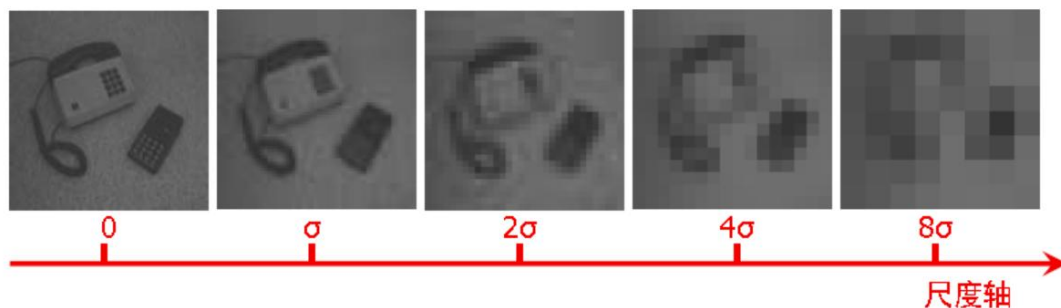
80



图像多尺度表达：示例

一副图像的多尺度表达：**每幅图像代表了相应尺度下的表达。**

该示例中，每一级尺度图像都是由二维高斯核与上一级尺度图像的卷积得到。每一级所用的高斯核的标准偏差 σ 不同，这些标准偏差形成一个等比数列。



81

相对尺度的概念和尺度因子



- **相对尺度**：给定一副原图像，它自身所在尺度为尺度坐标的零点。进行卷积的**高斯核的标准偏差 σ** 则代表了所得到的每一幅粗尺度图像在该尺度空间中的尺度坐标。
- **σ 也代表了每一幅粗尺度图像的相对分辨率**。因为 σ 越大，滤波器的尺寸就越大（通过比较 σ 分别为1和2的一维高斯核可以看出），卷积得到的图像就越模糊，因此分辨率就越低。
- **尺度因子(Scale Factor)**：两个相邻级别尺度之间的倍数，即两个相邻级别所用的高斯核的标准偏差的倍数： $s = \sigma_n / \sigma_{n-1}$

x	-4	-3	-2	-1	0	1	2	3	4
$g(x; \sigma=1)$	0.001	0.004	0.054	0.242	0.399	0.242	0.054	0.004	0.001
$g(x; \sigma=2)$	0.027	0.065	0.121	0.176	0.199	0.176	0.121	0.065	0.027

82

2.7.2 金字塔结构

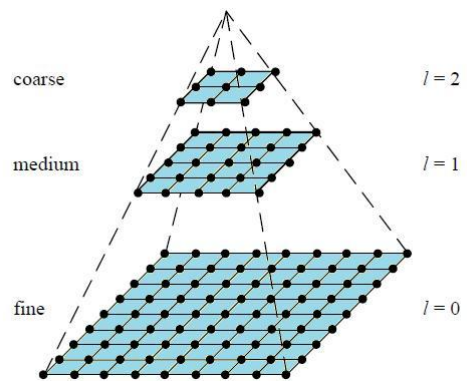


为了加快处理速度，对粗糙尺度图像进行子采样(sub-sampling)，从而得到一种分辨率和采样频率都在降低的多尺度表达，称为**高斯金字塔(Gaussian Pyramid)**。

• 高斯金字塔的**计算特点**：

1. 每一级与其相邻级别的尺度因子相同。
2. 采样频率和尺度因子要保持一致，从而避免在采样过程中产生**混淆(aliasing)**。

例如：如果相邻两级的尺度因子 $s=2$ （即两个相邻级别的高斯核的标准偏差倍数为2），它意味着这两个相邻级别的图像分辨率倍数也为2。因此，相应的采样频率也应该2，即从细一级尺度图像中的每两个像素采样一个像素。



83

高斯金字塔：示例



- 高斯金字塔的**缺点**：
- 经过子采样，**导致不同尺度级别图像的大小不一样**，需要确定相邻级别图像之间的点与点之间的位置关系。这将会提高理论分析的复杂度。
- 因此，不经过子采样，维持每一个尺度级别图像有同样的大小，可以保持理论分析和实际计算的直接联系。



84

拉普拉斯金字塔



- 两个相邻尺度图像之差，构成一个带通金字塔，称为**拉普拉斯金字塔 (Laplacian Pyramid)**。

计算步骤：

1. **将粗尺度图像插值到相邻的细尺度**。插值计算为一个相反的卷积计算。它是通过在两个像素中间插入新像素，达到把低分辨率图像扩展成高分辨率图像的效果。如果使用的高斯核的 $\sigma=1$ ，尺度因子 $s=2$ ，插值的计算公式如下：

$$L_{l \rightarrow l-1}(x, y; \sigma_l) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) L_l(x, y; \sigma_l)$$

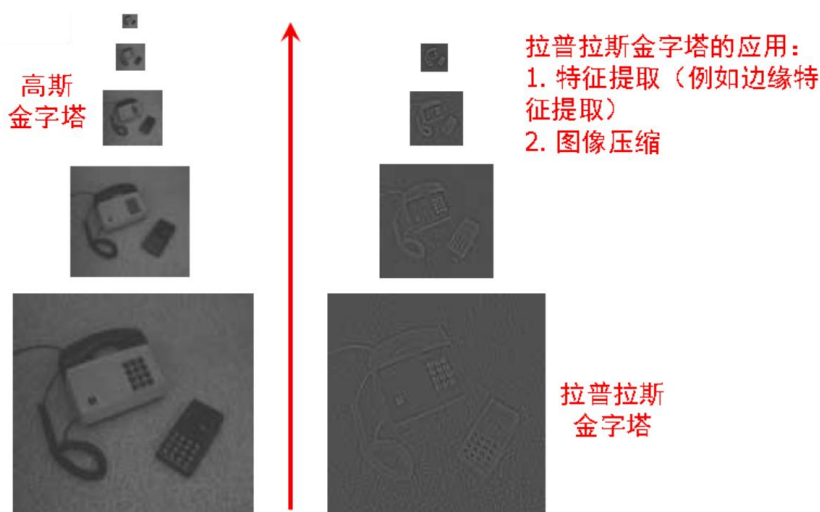
2. 计算差值：

$$D_{l-1}(x, y; \sigma_{l-1}) = L_{l-1}(x, y; \sigma_{l-1}) - L_{l \rightarrow l-1}(x, y; \sigma_l)$$

拉普拉斯金字塔的计算过程，类似于**DoG (Difference of Gaussians)**高斯差分算子。

85

拉普拉斯金字塔：示例



86



Thank you!