



杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY

篤學勵行

自正求新

CAMALAB
计算机动画与多媒体分析实验室

图像特征表达

谭敏

杭州电子科技大学，计算机学院

- 上周内容回顾
- 图像特征
 - 颜色特征
 - 纹理特征
 - 形状特征
 - 局部特征

上周内容回顾

■ 机器视觉

- 人怎么“看”？
- 机器怎么“看”？

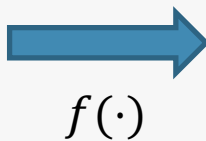
■ 自然语言

- 任务
- 方法：one-hot、BOW、tf-idf

图像特征表达

WHY

- 图像特征表达是理解图像内容的基础
- 把图像从非结构化数据变成结构化
- 使得图像之间的相关性可以有效度量



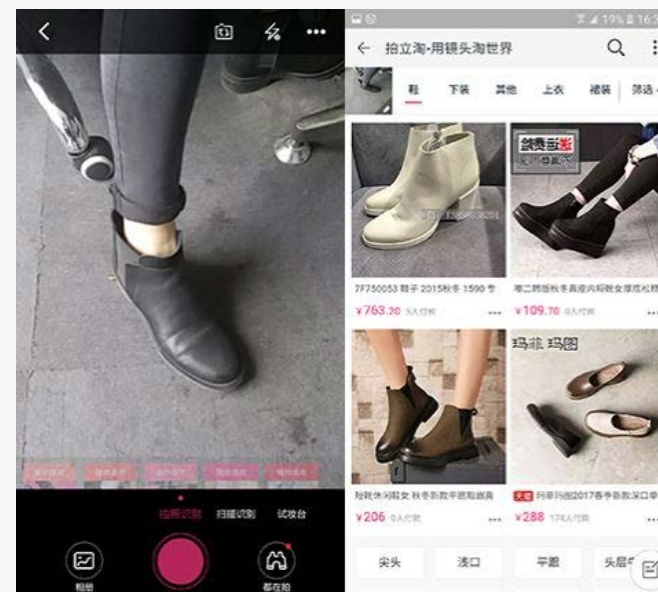
$[0.1, -0.5, \dots, 0.45] \in R^d$

图像特征表达的应用

- 图像相似度计算
 - 图像检索



通用图片搜索



拍照购物

图像特征表达的应用

- 图像相似度计算
 - 图像匹配



行人重识别



人脸验证

图像特征

- 颜色特征
- 纹理特征
- 形状特征
- 局部特征

颜色特征

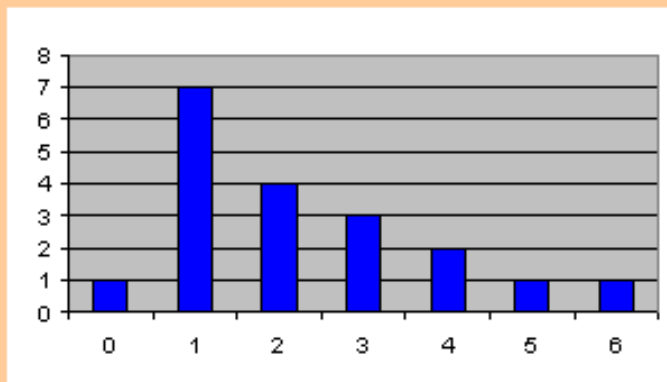
- 颜色特征的特点
 - 颜色与图像中包含的物体或场景关系密切；
 - 颜色特征对图像尺寸、方向、视角等的依赖性小；
- 需要解决的两个问题
 - 选择合适的颜色空间计算颜色特征
 - 选择合适的方法将颜色特征量化
- 表示的主要方法
 - 颜色直方图、主色调、颜色矩、颜色集、聚类

颜色直方图 (COLOR HISTGRAM)

- 颜色直方图描述不同色彩在整幅图像中所占的比例，而并不关心每种色彩所处的空间位置
- 颜色直方图可基于不同的颜色空间。最常用的颜色空间是RGB空间，原因在于大部分的数字图像都是用这种颜色空间表达的。也可以用其他色彩空间，如HSV

4	1	3	2
3	1	1	1
0	1	5	2
1	1	2	2

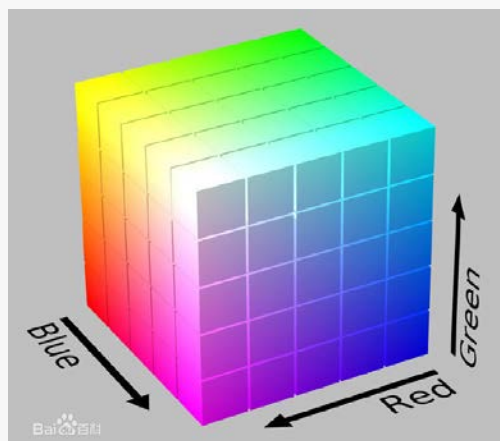
A simple 16x16 image



The histogram

颜色直方图的计算

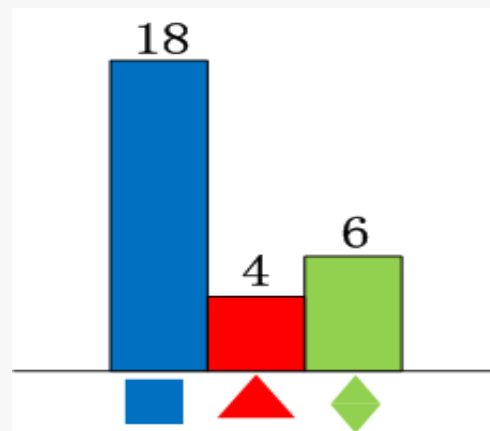
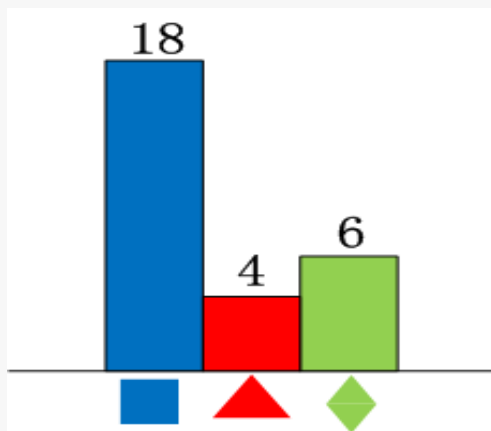
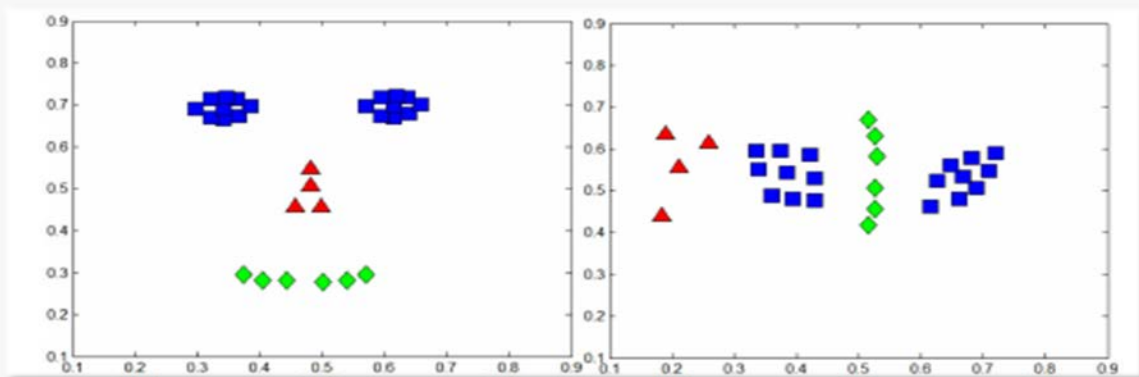
1. 选择颜色空间，并将该颜色空间划分成若干个小的颜色区域，即颜色量化（Color Quantization）。



2. 统计其颜色落入到每个小区间的像素数目可以得到颜色直方图。

颜色直方图的优缺点

- 优点：计算简单，一定场景下能反映图像内容
- 缺点：丢失空间信息



颜色矩(COLOR MOMENT)

- 图像中任何颜色分布均可以用它的矩来表示
- 矩也是图像处理中的常用概念

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{ij}$$

一阶矩 (mean)

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^2 \right)^{\frac{1}{2}}$$

二阶矩 (variance)

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^3 \right)^{\frac{1}{3}}$$

三阶矩 (skewness)

颜色矩(COLOR MOMENT)

- 好处在于无需对特征进行量化。
- 颜色直方图与颜色矩都用于描述一副图像中颜色信息的统计特点，但是颜色矩的特征描述更为紧凑，具有更少的维数。例如采用三维颜色矩只有9维
- 在实际应用中为避免低次矩较弱的分辨能力，常与其它特征结合使用，并且在其他特征使用前起到过滤缩小范围的作用。

颜色聚合向量 (COLOR COHERENCE VECTOR)

- 颜色聚合向量(color coherence vector, CCV)是一种颜色特征，它包含了颜色分布的空间信息。克服了颜色直方图无法表达图像色彩的空间位置的缺点。
- 其核心思想是将属于每一个bin的像素分为两部分：如果该bin内的某些像素所占据的连续区域的面积大于给定的阈值，则该区域内的像素作为**聚合像素**，否则作为**非聚合像素**。

颜色聚合向量计算

- 假设 α_i 和 β_i 分别代表直方图的第 i 个bin中聚合像素和非聚合像素的数量，图像的颜色聚合向量可以表达为

$$CCV = [\alpha_1, \beta_1, \alpha_2, \beta_2, \dots, \alpha_N, \beta_N] \in R^{2N}$$

$$CH = [\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_N + \beta_N] \in R^N$$

- 给定两张图像 I 和 I' 的CCV特征，其 L_1 距离可以表示为

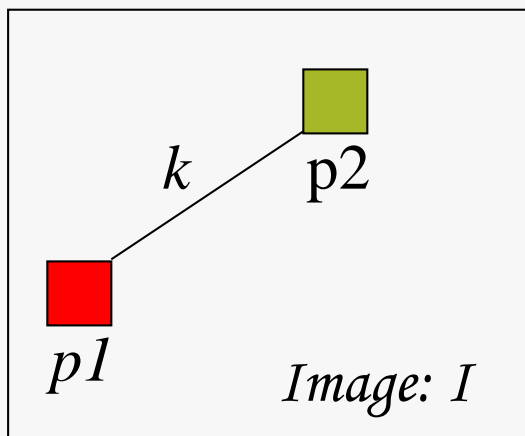
$$\Delta CCV = \sum_i (|\alpha_i - \alpha'_i| + |\beta_i - \beta'_i|)$$

颜色相关图 (COLOR CORRELOGRAM)

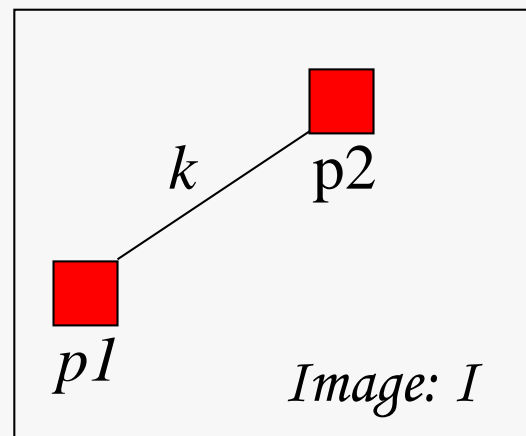
- 颜色相关图 (color correlogram) 是图像颜色分布的另一种表达方式。这种特征不但刻画了某一种颜色的像素数量占整个图像的比例，还反映了不同颜色对之间的空间相关性。
- 颜色相关图比颜色直方图和颜色聚合向量具有更高的检索效率，特别是查询空间关系一致的图像。

颜色相关图计算

- 衡量颜色不同颜色 (bin) i, j 两种不同颜色在所有距离为 k 的像素中出现的概率
- 计算复杂度为 $O(N^2)$, N 为 bin 的个数
- 简化的版本成为颜色自相关图, 仅计算相同颜色在所有距离为 k 的像素中出现的概率



颜色相关图示例

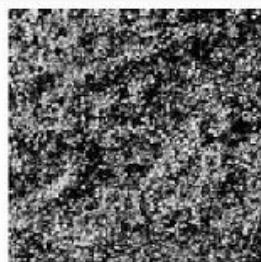


颜色自相关图示例

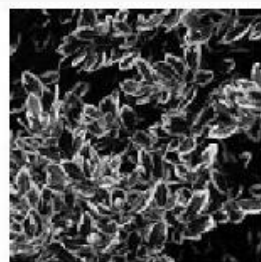
- 图像特征
 - 颜色特征
 - 纹理特征
 - 形状特征
 - 局部特征

纹理特征

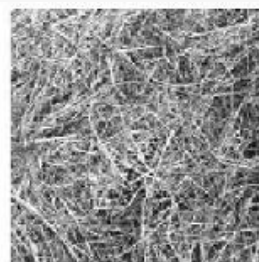
- 纹理是一种普遍存在的视觉现象；
- 尚无一个被广泛接受的定义；
- 纹理是图像灰度或色彩在空间上的变化或重复



Leaves



Leaves



Grass



Brick



Brick



Stone

纹理特征的特点

- 纹理特征是从图像中计算出来的一个值，对区域内部灰度级变化的特征进行量化。
- 是一种全局统计特征，仅利用纹理特征无法获得高层次图像内容。
- 不是基于像素点的特征，需要在包含多个像素点的区域中进行统计计算。
- 具有旋转不变性，且对噪声有较强的抵抗能力。
- 当图像分辨率变化的时候，计算出来的纹理可能会有较大偏差。
- 适用于检索具有粗细、疏密等方面较大差别的纹理图像。

纹理特征

- 灰度共生矩阵
- 局部二进制模式 (LBP)
- HOG梯度直方图

灰度共生矩阵

定义：从灰度为 i 的像素点出发，距离为 (dx, dy) 的另一个像素点的灰度为 j 的概率；

数学表达式：
$$P(i, j | d, \theta) = \# \{ (x, y) | f(x, y) = i, f(x + dx, y + dy) = j; x, y = 0, 1, 2, \dots, N - 1 \}$$


式中， d 是用像素数量表示的相对距离；


θ 一般考虑四个方向，分别为 0° ， 45° ， 90° ， 135° ；

$\#$ 表示集合； $i, j = 0, 1, 2, \dots, L - 1$ ；


(x, y) 为图像中的像素坐标， L 为图像灰度级的数目。


灰度共生矩阵

$$M^{(0^\circ)} = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$


$$M^{(90^\circ)} = \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$


$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 \end{bmatrix}$$

$$M^{(45^\circ)} = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$


$$M^{(135^\circ)} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$


灰度共生矩阵-统计量

- 角二阶矩或能量：它是图像纹理灰度变化均一的度量,反映了图像灰度分布均匀程度和纹理粗细度。

$$ASM = \sum_i \sum_j P(i, j|d, \theta)^2$$

- 对比度：灰度共生矩阵主对角线附近的惯性矩,它度量矩阵的值是如何分布和图像的局部变化,反映了图像的清晰度和纹理的沟纹深浅。

$$CON = \sum_i \sum_j (i - j)^2 P(i, j|d, \theta)$$

- 相关性：度量空间灰度共生矩阵元素在行或列方向上的相似程度,相关值大小反映了图像中局部灰度相关性

$$COR = \sum_i \sum_j (i - \mu_x)(j - \mu_y)P(i, j|d, \theta) / \sigma_x \sigma_y$$

$$\mu_x = \sum_i i \sum_j P(i, j|d, \theta)$$

$$\mu_y = \sum_j j \sum_i P(i, j|d, \theta)$$

$$\sigma_x = \sum_i (i - \mu_x)^2 \sum_j P(i, j|d, \theta)$$

$$\sigma_y = \sum_j (j - \mu_y)^2 \sum_i P(i, j|d, \theta)$$

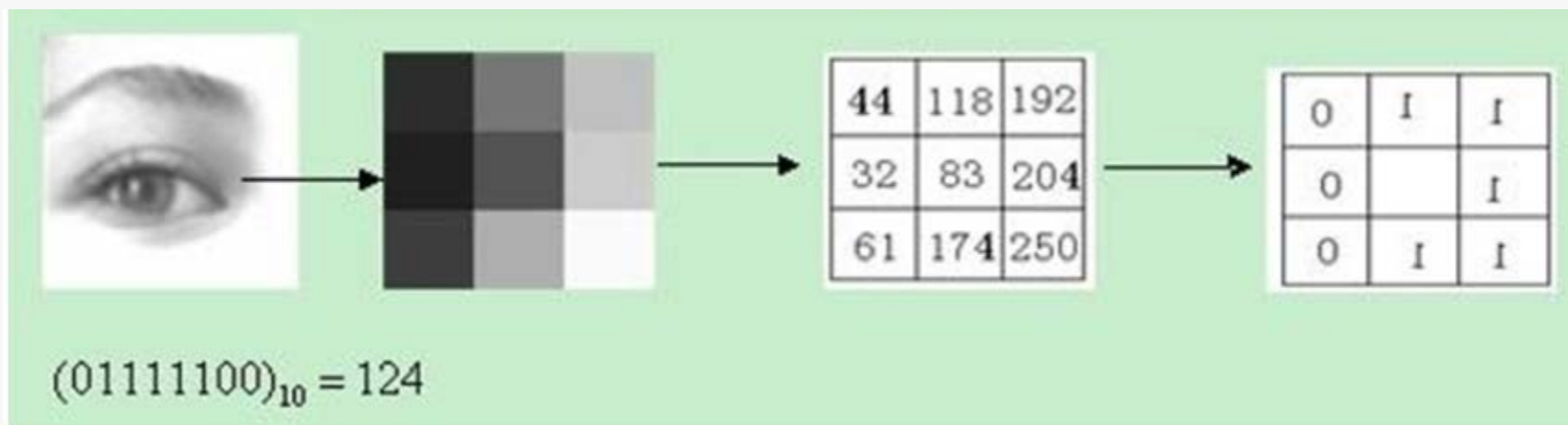
灰度共生矩阵-统计量

- 熵：度量图像纹理的随机性。当共生矩阵中所有值均相等时，它取得最大值；相反，如果共生矩阵中的值非常不均匀时，其值较小。

$$ENT = - \sum_i \sum_j P(i, j | d, \theta) \log P(i, j | d, \theta)$$

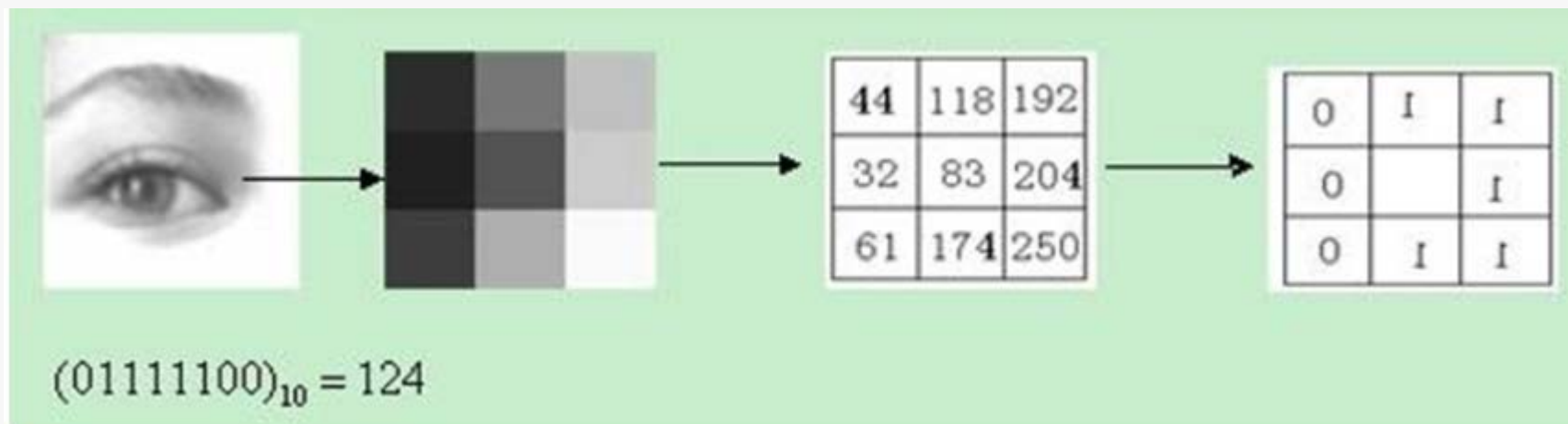
局部二值模式 (LBP)

- LBP是一种3x3模板算子，若周围相似大于中心像素，记为1，否则为0，然后把8位二进制码表达成对应的实数值



- LBP特征向量：把图像分割成M区域（如16x16）。每个区域内的每个像素计算LBP值，然后转变为N个bin的直方图，并归一化。M个区域对应的直方图拼接成MN维的向量作为图像特征

局部二值模式 (LBP)



$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p$$

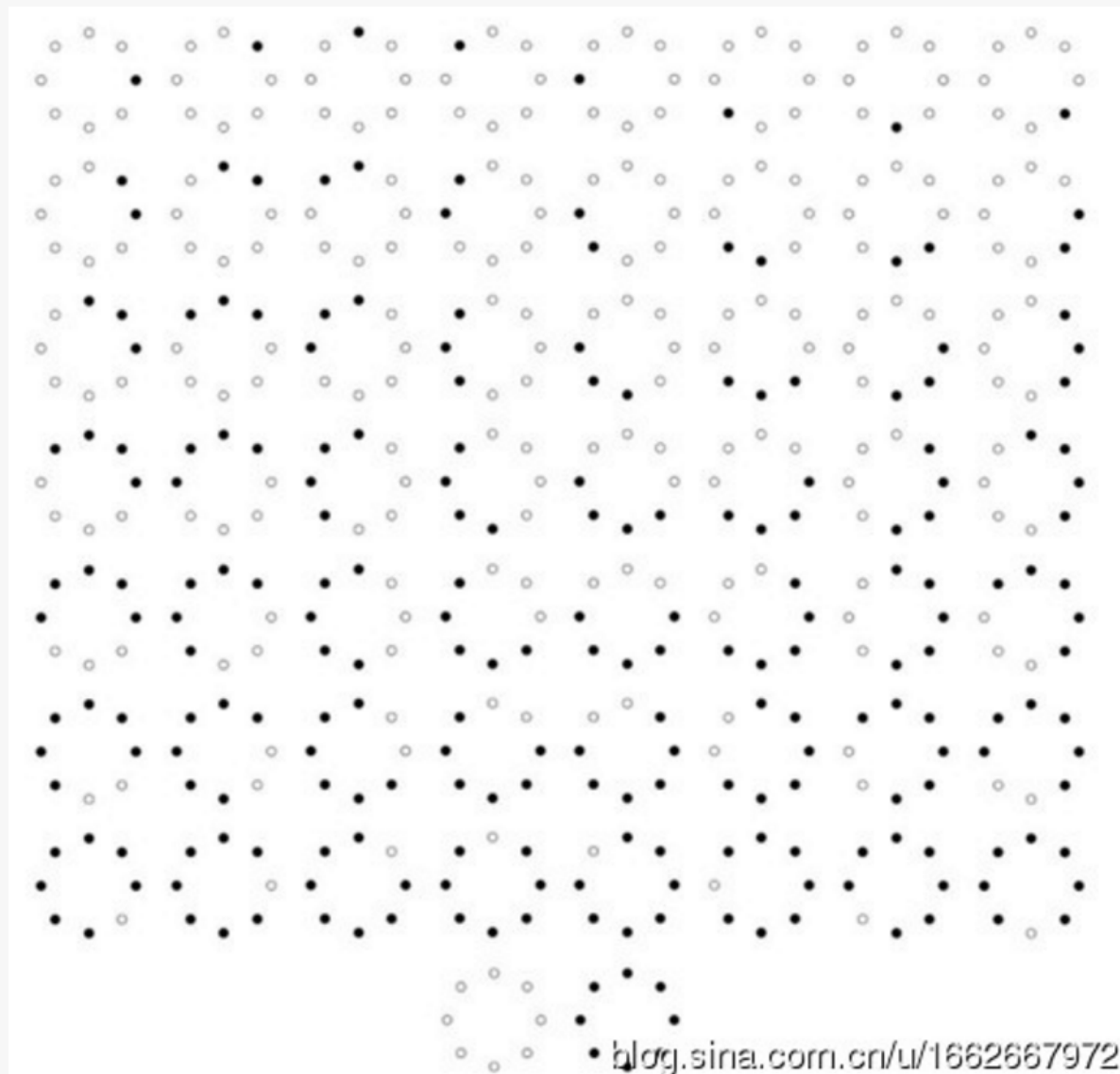
$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

局部二值模式 (LBP)

- N bins ?

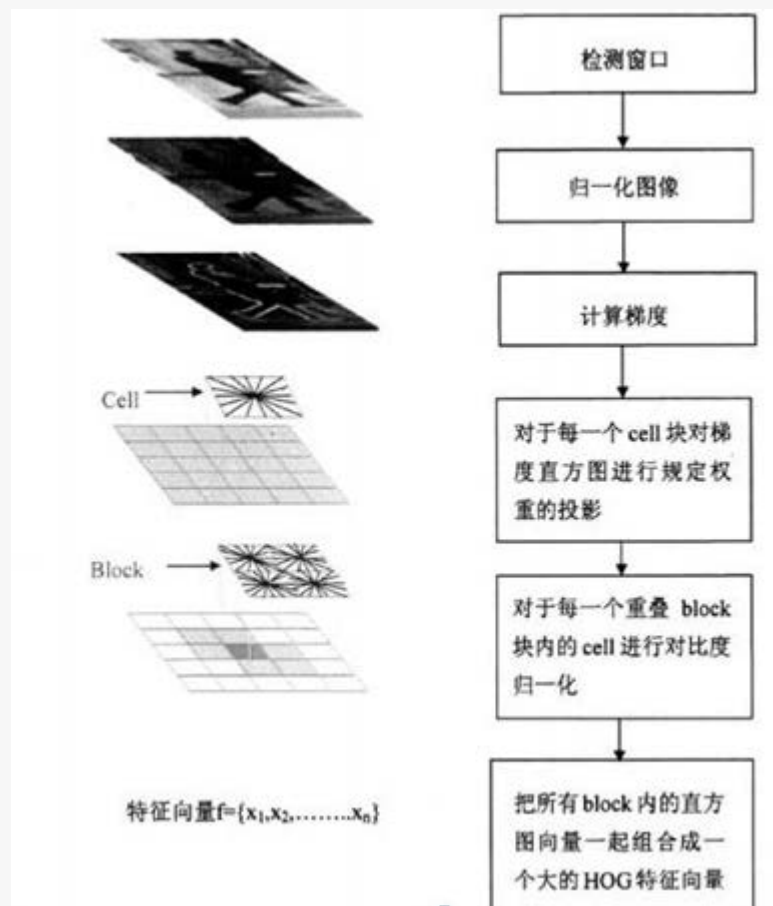
Uniform bins

59 Bins

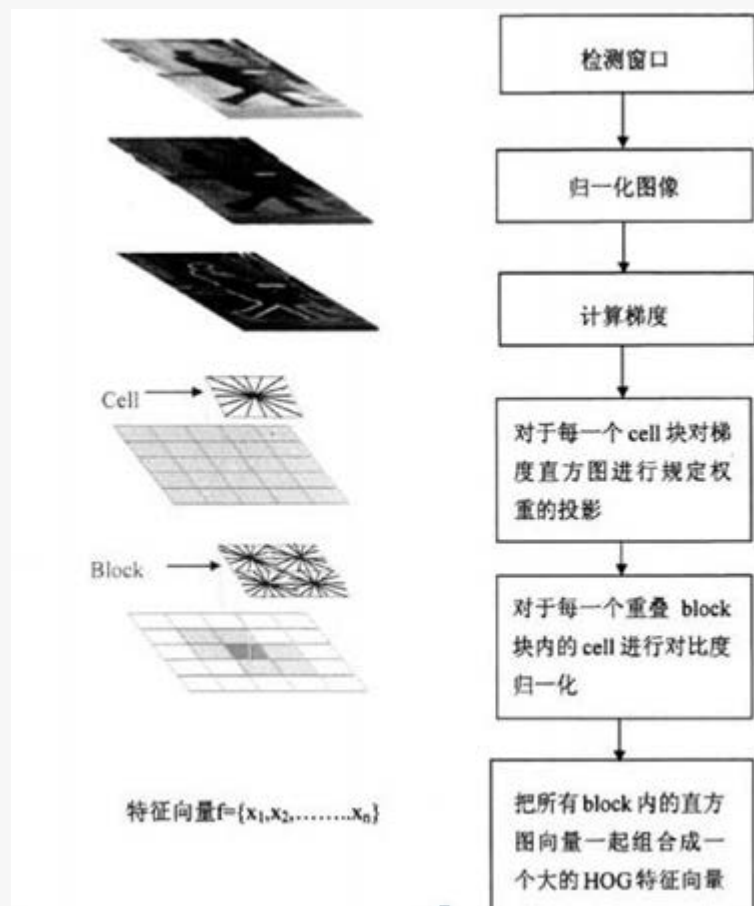


梯度直方图特征 (HOG)

- 深度学习之前目标检测领域最有效的图像特征



梯度直方图特征 (HOG)



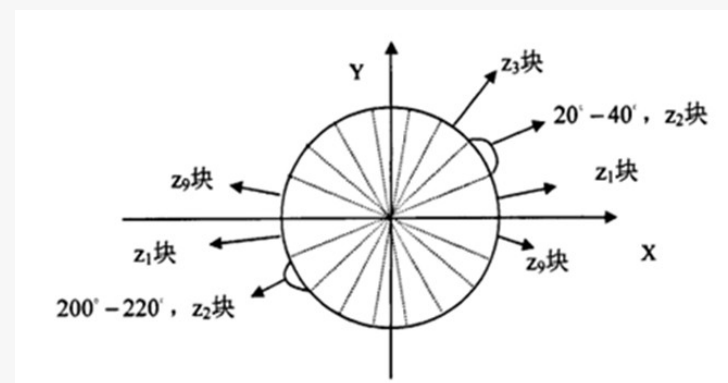
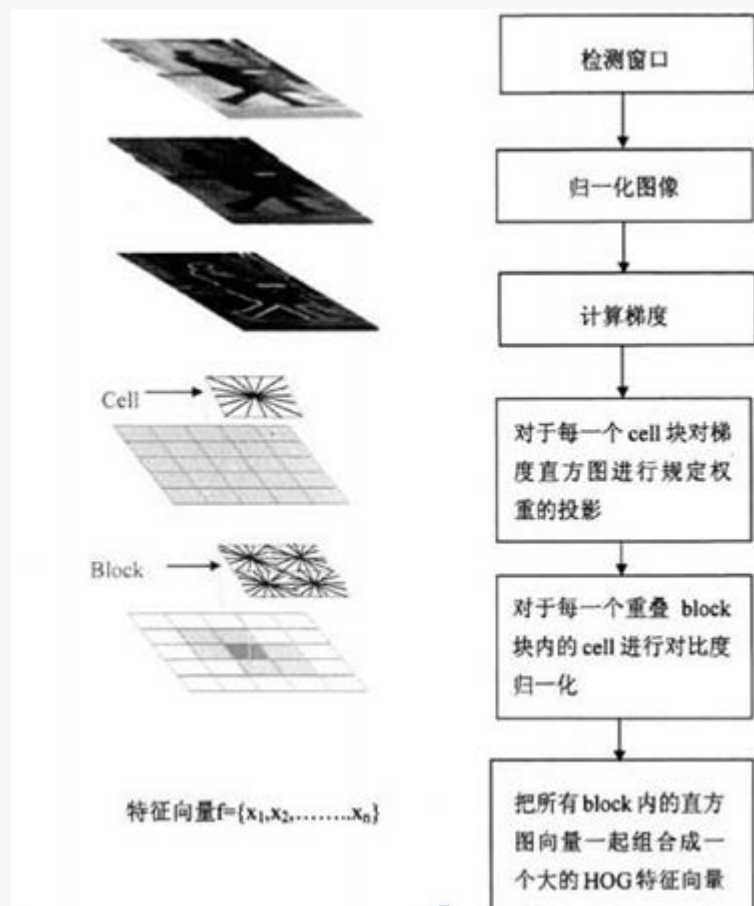
$$G_x = \frac{\partial f}{\partial x} = f(x+1) - f(x)$$

$$G_y = \frac{\partial f}{\partial y} = f(y+1) - f(y)$$

$$\text{梯度值 } G = \sqrt{G_x^2 + G_y^2}$$

$$\text{梯度方向 } \theta = \arctan\left(\frac{G_x}{G_y}\right)$$

梯度直方图特征 (HOG)



梯度直方图特征 (HOG)

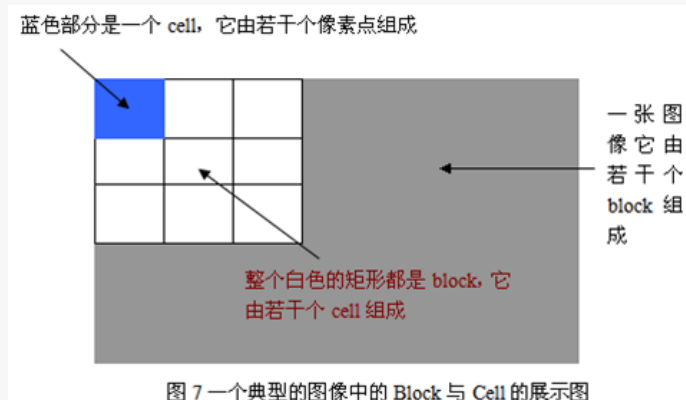
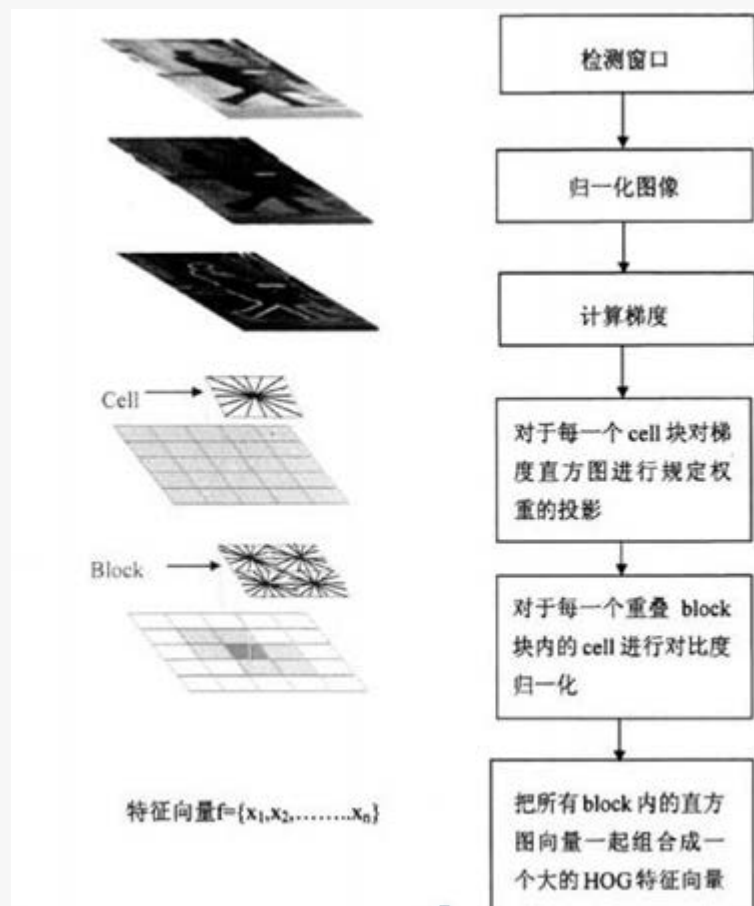
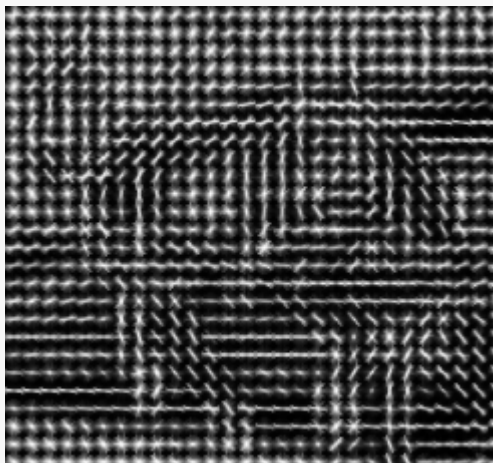
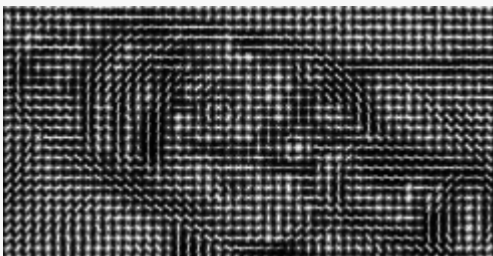
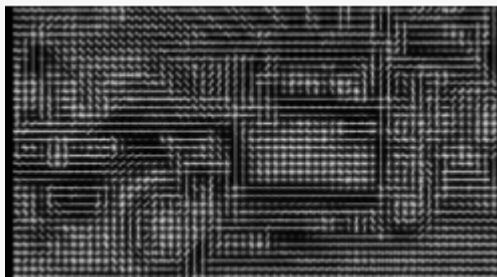


图 7 一个典型的图像中的 Block 与 Cell 的展示图

- a) $L2-norm, v \leftarrow v / \sqrt{\|v\|_2^2 + \epsilon^2}$ (ϵ 是个很小的常值, 避免分母为 0)
- b) $L2-Hys$, (方法同上, 只是限制 v 的最大值到 0.2, 然后再重新标准化)
- c) $L1-norm, v \leftarrow v / (\|v\|_1 + \epsilon)$
- d) $L1-sqrt, v \leftarrow \sqrt{v / (\|v\|_1 + \epsilon)}$

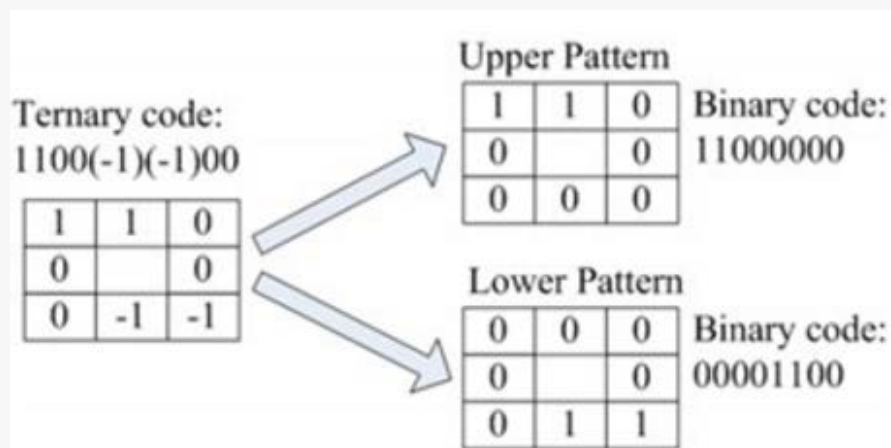
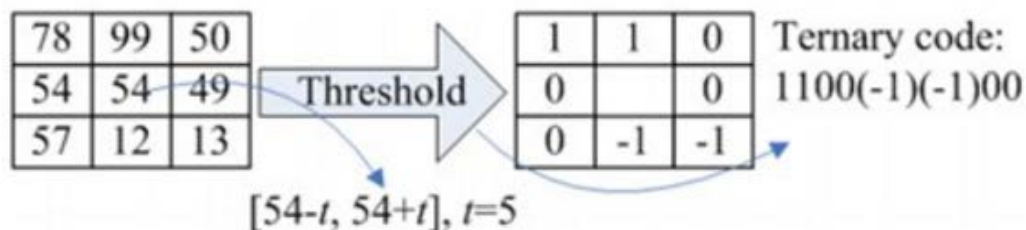
HOG特征可视化



局部三值模式 (LTP)

- 三值编码

$$s'(u, i_c, t) = \begin{cases} 1, & u \geq i_c + t \\ 0, & |u - i_c| < t \\ -1, & u \leq i_c - t \end{cases}$$



- 图像特征
 - 颜色特征
 - 纹理特征
 - 形状特征
 - 局部特征

形状特征

- 形状特征的表达必须以对图像中**物体或区域的分割**为基础；
- 形状特征必须满足对**变换、旋转和缩放**无关；
- 形状特征表示方法：**基于边界和基于区域**。

形状特征

- 基于区域：用物体区域的整体性质来表示
 - 四叉树、包围区域
 - 骨架
- 基于边界：用物体或区域的外边界来表示
 - 边界的长度，直径，曲率；
 - 链码描述；
 - 傅立叶描述符；
 - 基于内角的表示法；

基于区域的表达

- 四叉树
 - 目标节点（黑）
 - 背景节点（白）
 - 混合节点（灰）

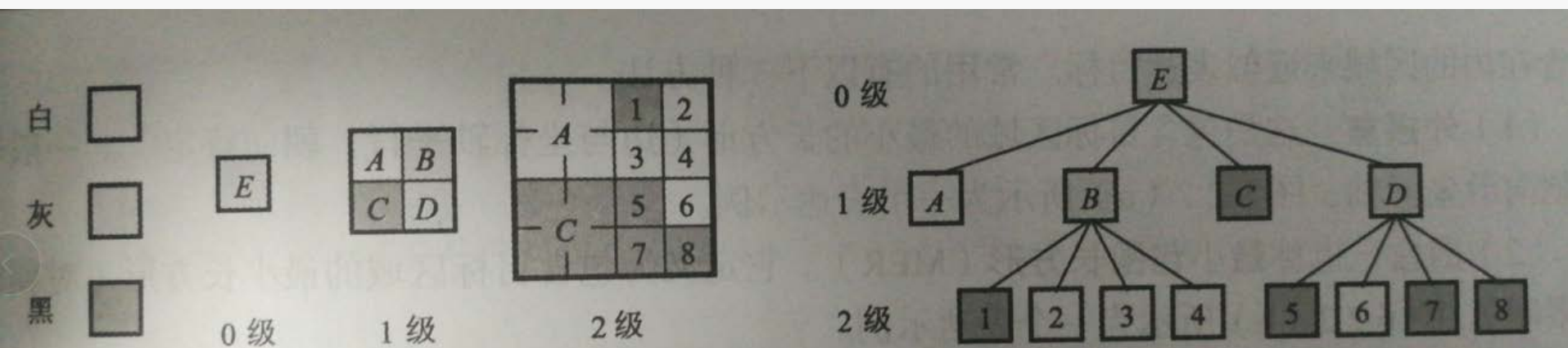
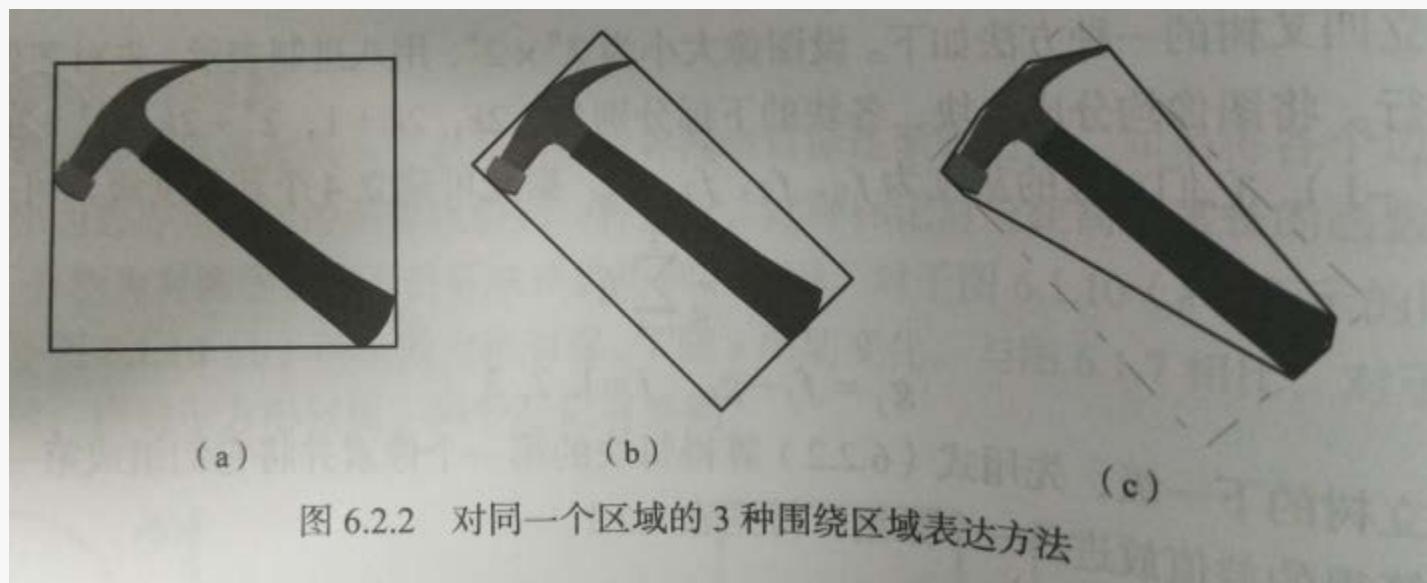


图 6.2.1 四叉树表达图示

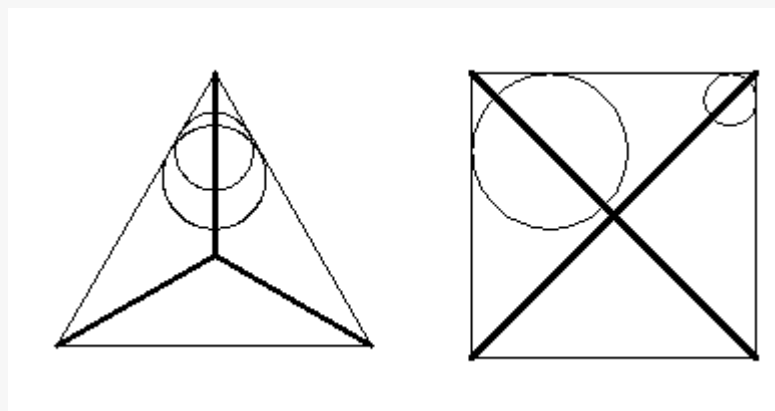
基于区域的表达

- 包围区域
 - 外接盒、最小包围矩形、凸包



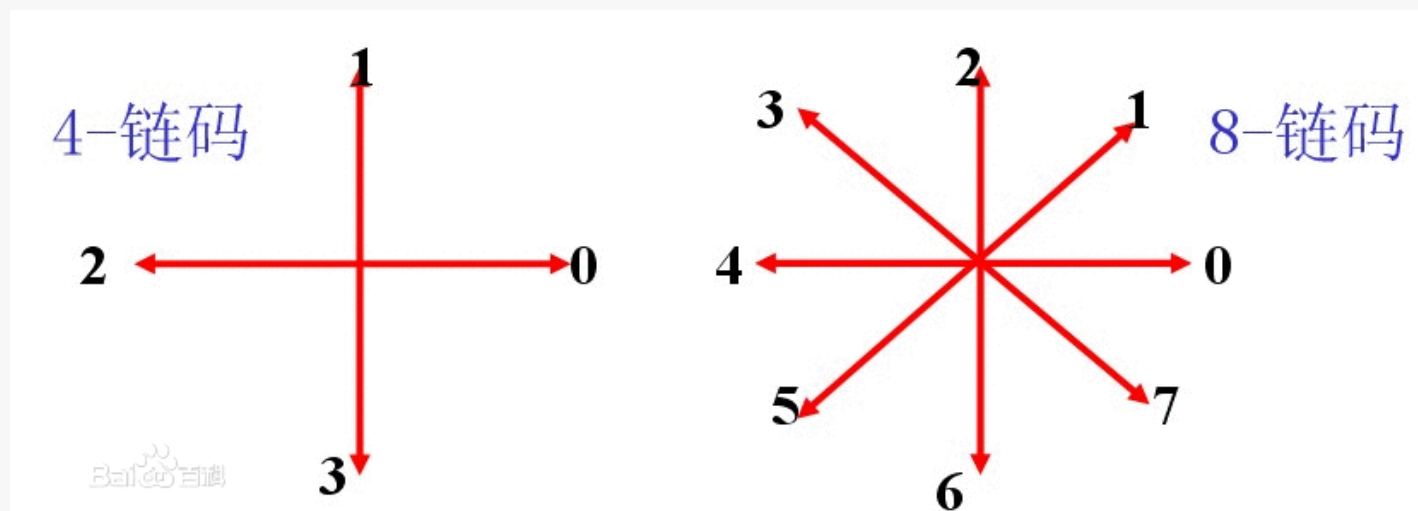
基于区域的表达

- 骨架：也称为中轴。区域内任意像素到边界最小的点**大于等于2个**，则该像素点是骨架上的一个点



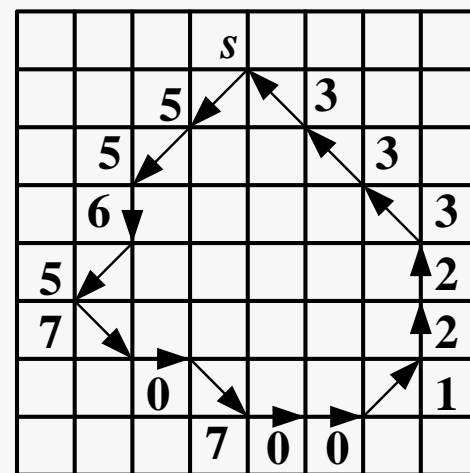
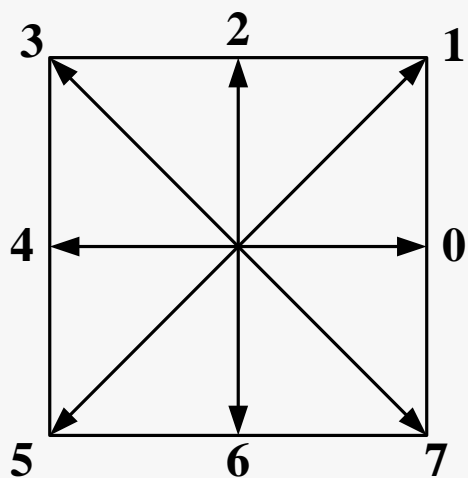
基于边界的表达

- 链码：采用曲线起始点的坐标和斜率(方向)来表示曲线



基于边界的表达

- 链码具有**方向性**
- 8-链码分成偶数码（水平或垂直，长度为1）和奇数码（对角线，长度为 $\sqrt{2}$ ）

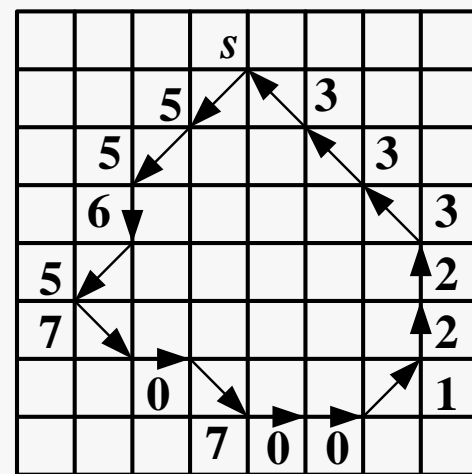
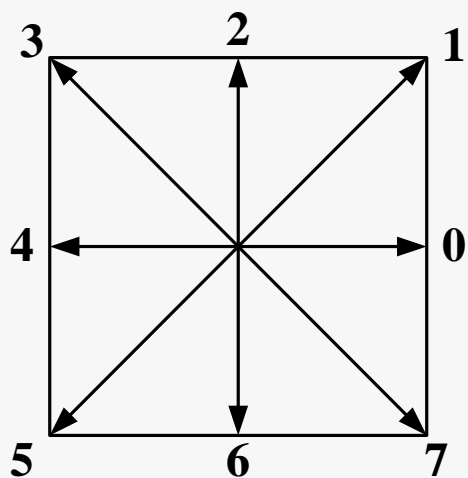


八链码:

556570700122333

基于边界的表达

- 链码具有**方向性**
- 8-链码分成偶数码（水平或垂直，长度为1）和奇数码（对角线，长度为 $\sqrt{2}$ ）



八链码:

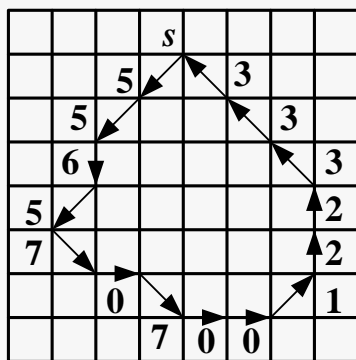
556570700122333

基于边界的表达

- 链码归一化

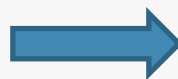
- 给定一个从任意点开始得到的链码，把它看作一个由各方向数构成的自然数。将这些方向数依一个方向循环，以使它

如何解决旋转不变性？



原始:

556570700122333



归一化后:

001223335565707

- 图像特征
 - 颜色特征
 - 纹理特征
 - 形状特征
 - 局部特征

局部特征&局部特征匹配

- 局部特征
 - SIFT
 - SURF
 - ORB
- 局部特征匹配
 - SIFT特征匹配
- 局部特征 → 全局特征
 - 视觉词袋或空间金字塔

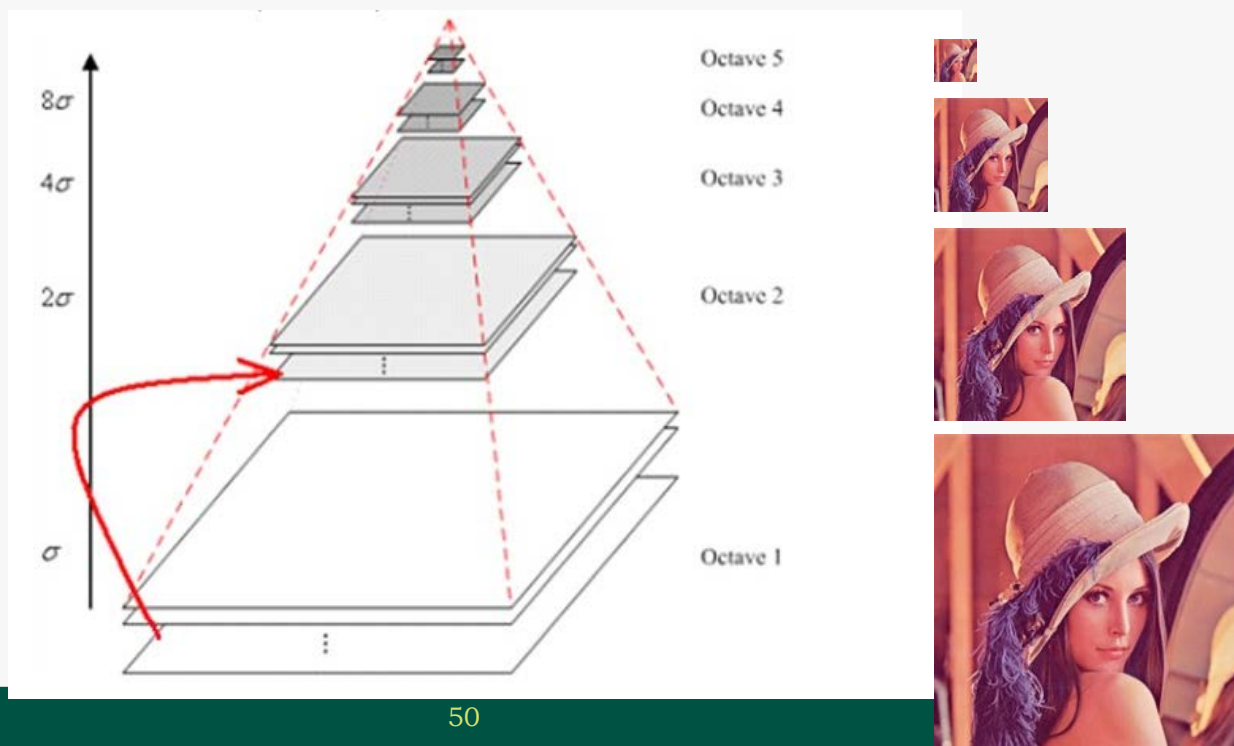
SIFT描述子

- 尺度不变特征变换 (Scale Invariant Feature Transformation)
 - 检测图片中的兴趣点，每个兴趣点形成128-D的特征



SIFT描述子

- 金字塔表示的图像尺度空间,总用有n层（octave），每一层中有一组相同尺度的图像



SIFT描述子

- 每一组内的尺度图片使用的是不同参数的高斯平滑模板

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

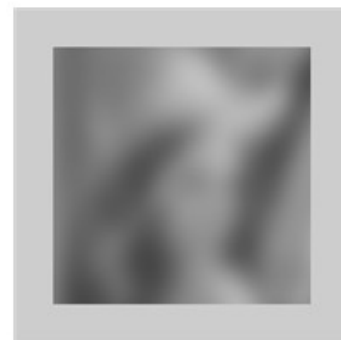
$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m/2)^2 + (y-n/2)^2}{2\sigma^2}}$$



a) 原图



b) $\sigma = 0.6$



c) $\sigma = 10.0$

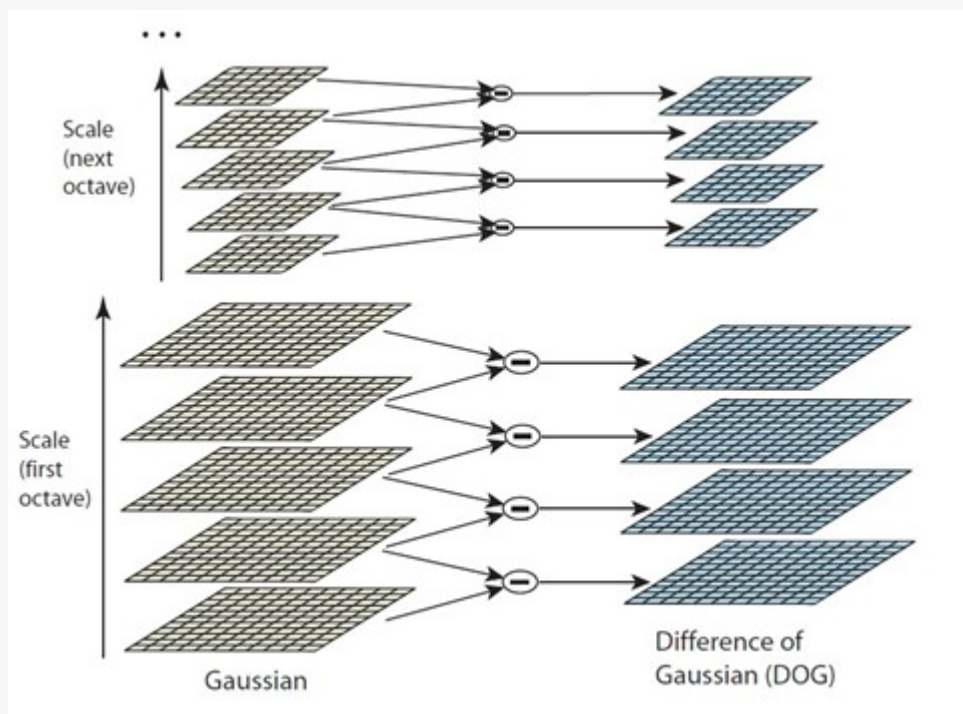


d) $\sigma = 10.0$ (边缘处理)

SIFT描述子

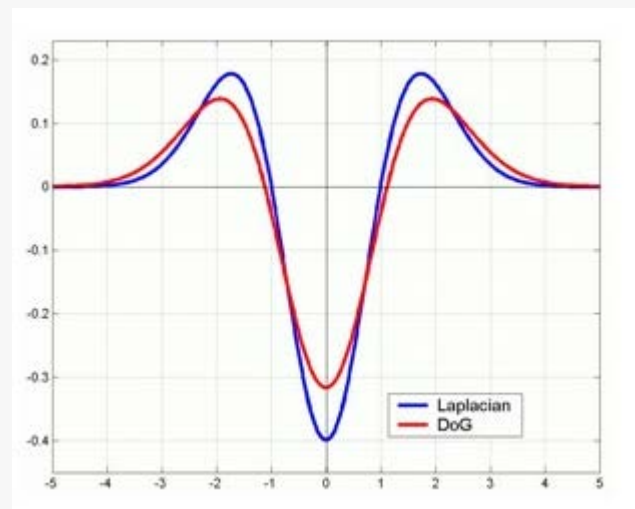
- 高斯差分 (DoG)

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$



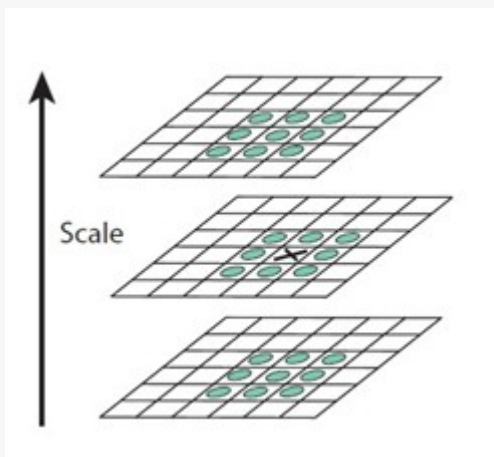
SIFT描述子

- Why DoG?
 - 目标是检测图像中的“极值”作为关键点
 - 之前的研究证明拉普拉斯高斯函数（LoG）的极值能产生相比于梯度、Harris角点等更稳定的特征
 - 高斯差分近似等价于LoG



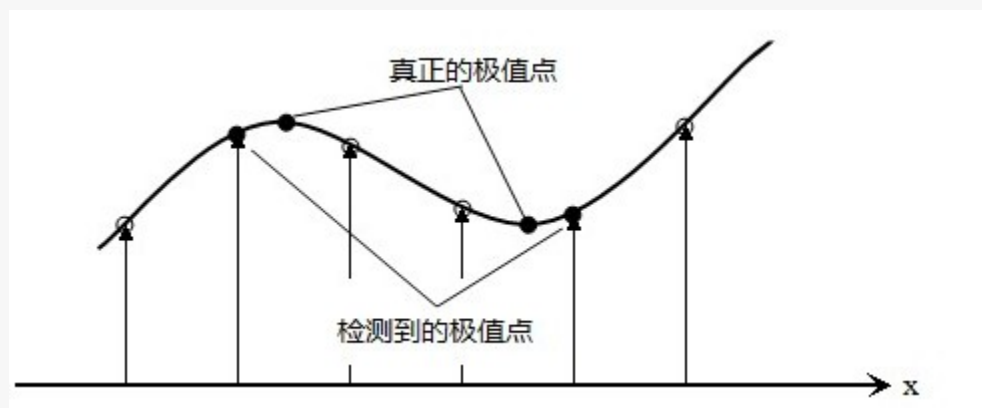
SIFT描述子

- 极值点检测
 - 得到统一尺度的DoG图后，一张图内的像素和它相邻的8个像素，以及相邻的两张DoG图中的对应9个像素进行比较，判断是否是极大值或极小值点



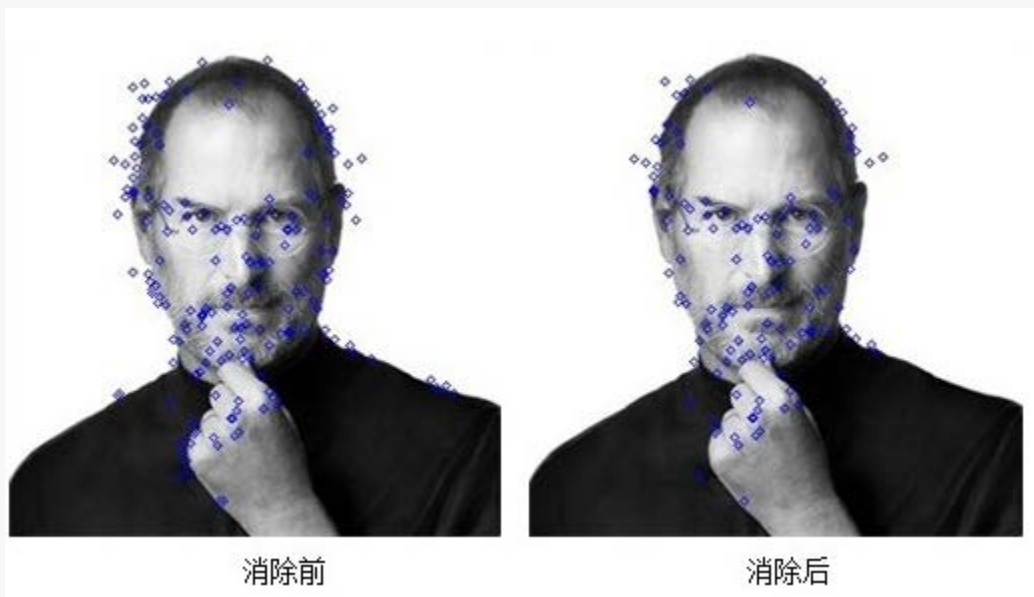
SIFT描述子

- 把极值点从尺度空间映射会原始空间
 - 子像素差值



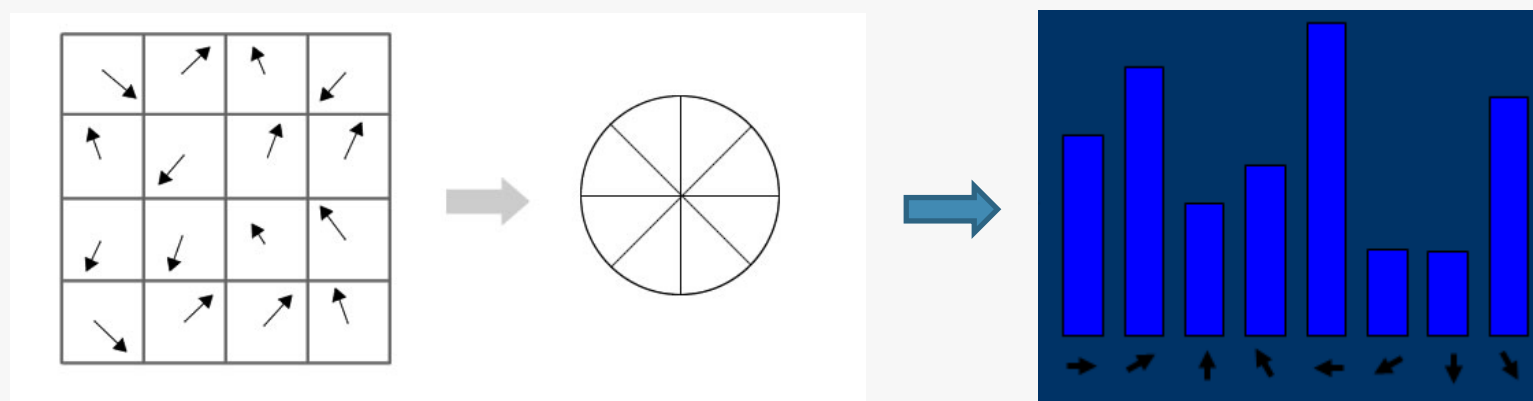
SIFT描述子

- 后处理，消除边缘响应
 - Harris角点法



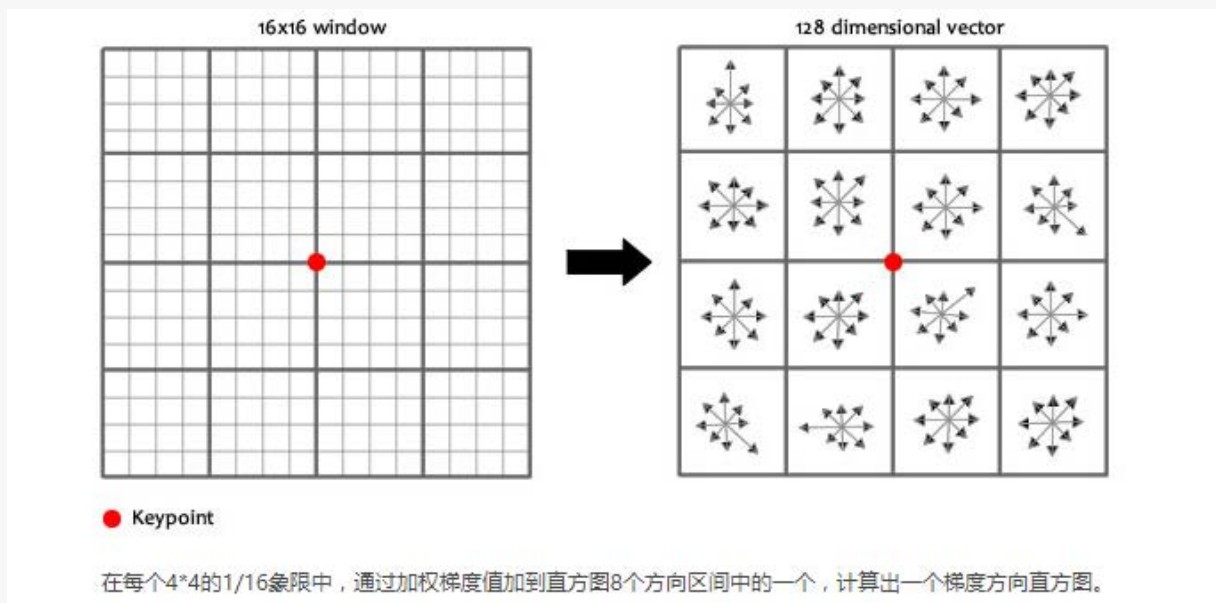
SIFT描述子

- 关键点特征
 - DoG图像上计算4*4像素上梯度在8个方向(bin)上的直方图



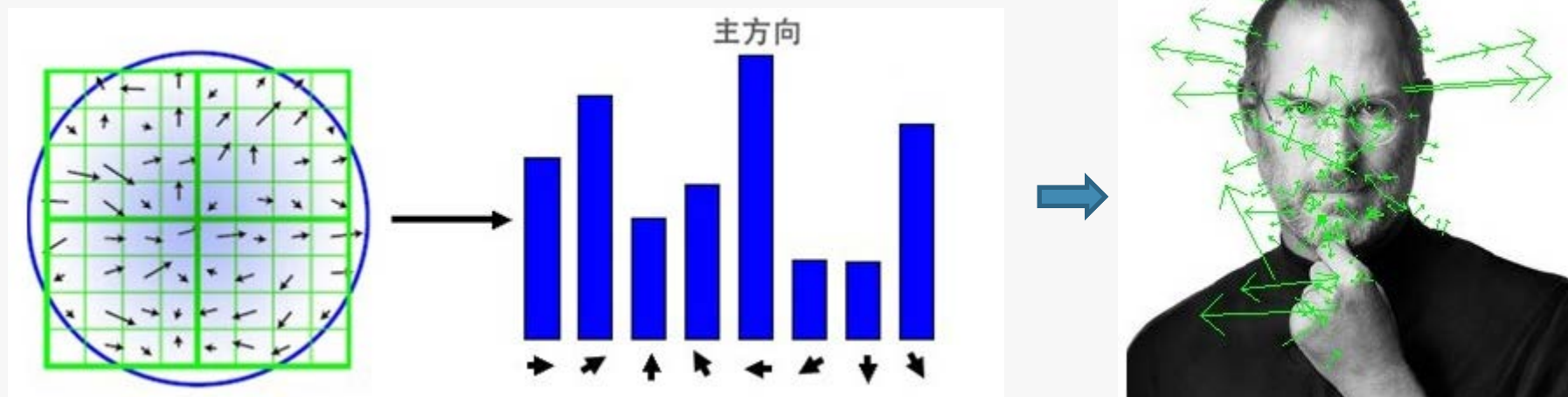
SIFT描述子

- 关键点特征
 - 每个关键点周围 16×16 个像素按照上述计算梯度直方图，最终形成 $8 \times 16 = 128D$ 的特征



SIFT描述子

- 主方向可视化
 - 挑选直方图中的最大的方向进行可视化



后SIFT时代的局部特征

- 局部特征抽取可以分成两步
 - 关键点检测
 - ORB
 - SURF
 - 关键点特征表达
 - BRIEF
 - PCA-SIFT
 - GLOH

后SIFT时代的局部特征

- 局部特征抽取可以分成两步

- 关键点检测

- ORB
- SURF

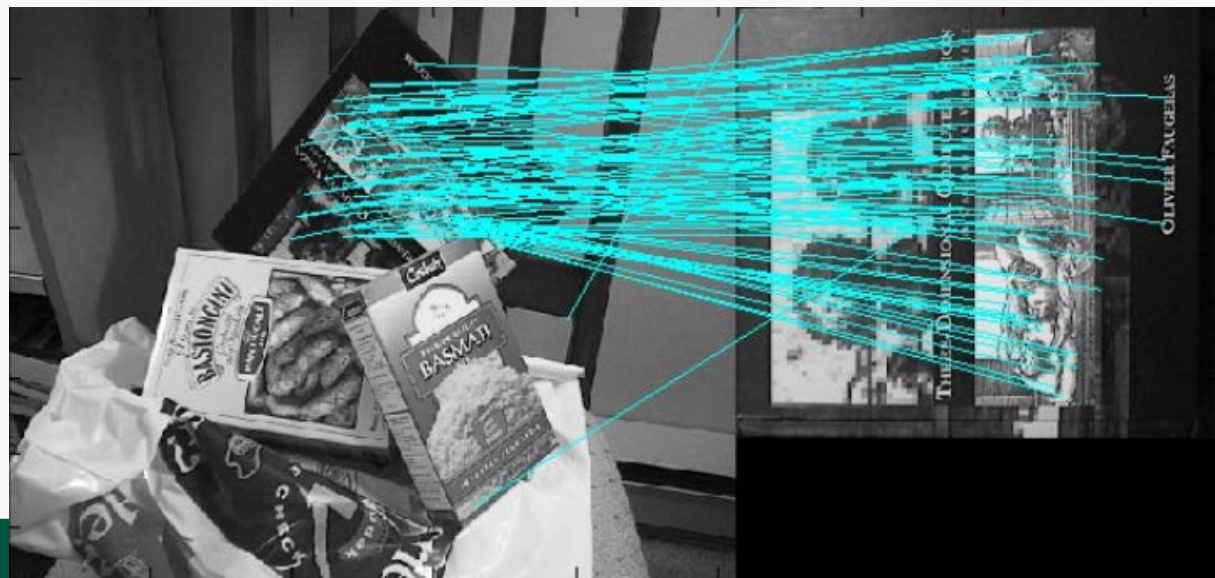
- 关键点特征表达

更快、更高、更强

- BRIEF
- PCA-SIFT
- GLOH

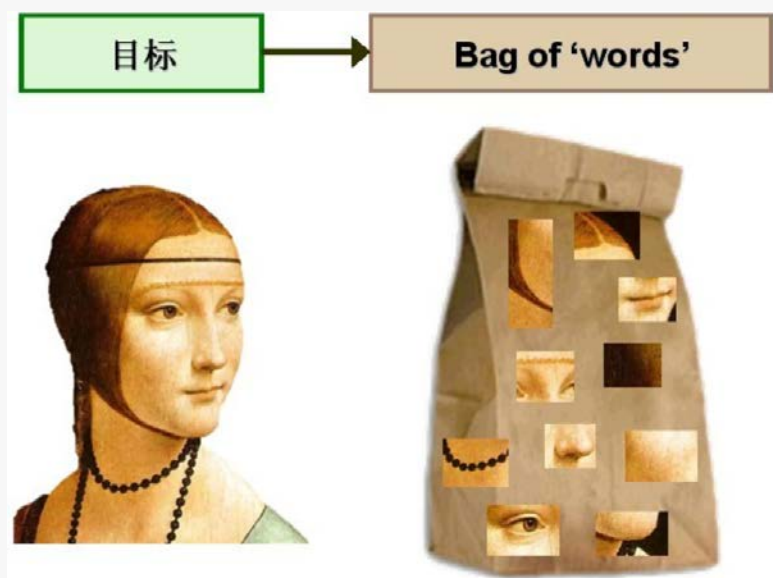
SIFT特征匹配

- 对图像 I_1 中的任意SIFT特征点 d_1 ，在 I_2 的特征点中寻找距离最小的两个特征点 d_2, d_3 ，若 $\frac{Dis(d_1, d_2)}{Dis(d_1, d_3)} < T$ ，则匹配，否则不匹配



视觉词袋

- 把图像内 $N \times 128$ 的局部特征编码成 $1 \times D$ 的词袋 (Bag-of-word) 特征



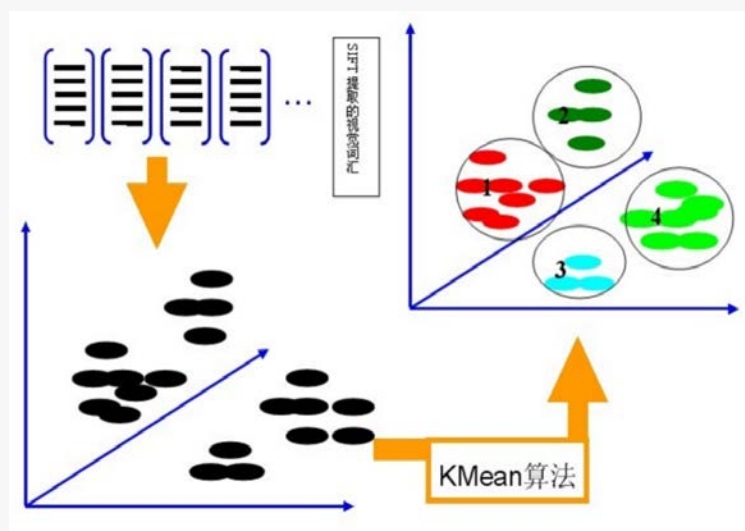
视觉词袋

- Step0: 对一个图像数据集中的图片都提取局部特征（如SIFT）



视觉词袋

- Step1:构建视觉“词汇”表
 - 对step0中提取的局部特征进行聚类，选取聚类中心作为视觉单词



视觉词袋

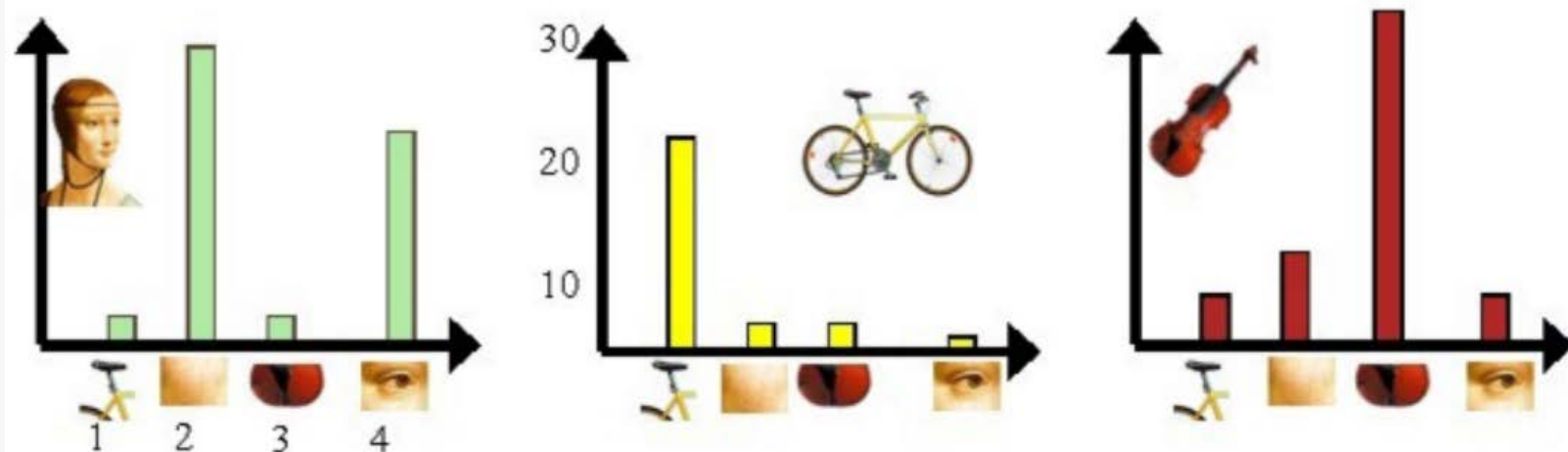
- Step2:使用视觉词汇表对一张图像编码
 - 统计一张图像上所有局部特征在词汇表上的分布（计算和所有聚类中心的距离，找到距离最近的聚类中心点）



视觉词袋

- Step2:使用视觉词汇表对一张图像编码
 - 统计一张图像上所有局部特征在词汇表上的分布（计算和所有聚类中心的距离，

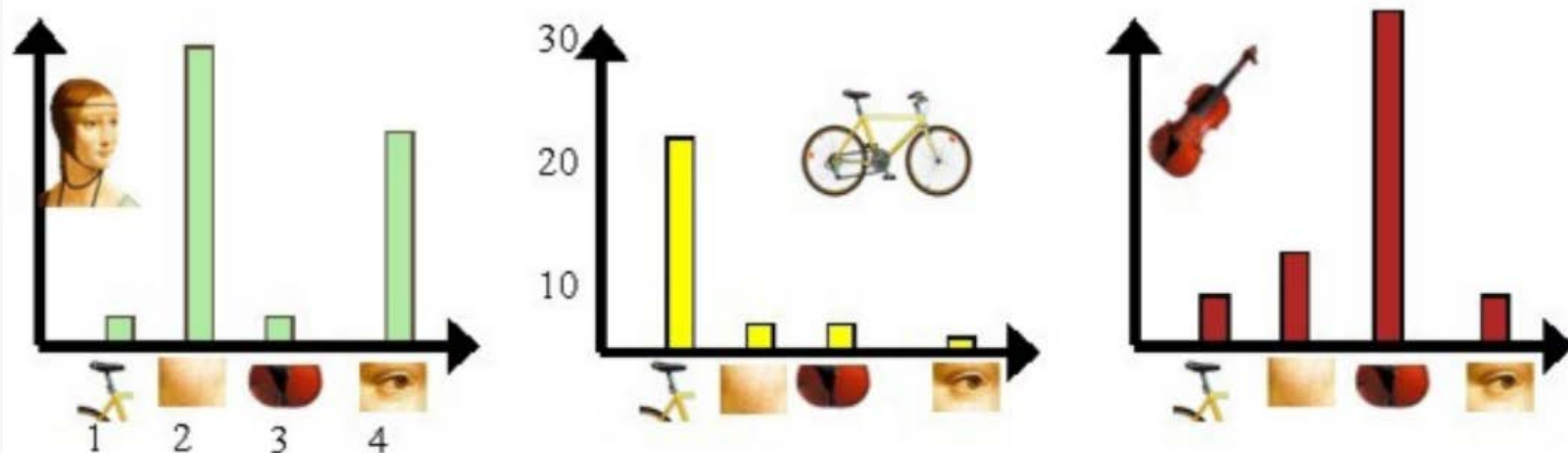
视觉词袋特征存在什么缺陷？



视觉词袋

- Step2:使用视觉词汇表对一张图像编码
 - 统计一张图像上所有局部特征在词汇表上的分布（计算和所有聚类中心的距离，

空间信息丢失！



空间金字塔

- 对图像建立多尺度描述，每个尺度上分别进行视觉单词表达

