

主要步骤:

- 1)、尺度空间的生成;
- 2)、检测尺度空间极值点;
- 3)、精确定位极值点;
- 4)、为每个关键点指定方向参数;
- 5)、关键点描述子的生成。

● 尺度空间的生成

尺度空间理论目的是模拟图像数据的多尺度特征, 高斯卷积核是实现尺度变换的唯一线性核, 于是一副二维图像的尺度空间定义为:

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

其中, $G(x, y, \sigma)$ 是尺度可变高斯函数, $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$

(x, y) 是空间坐标, σ 是尺度坐标。 σ 的大小决定图像的平滑程度, 大尺度对应图像的概貌特征, 小尺度对应图像的细节特征。大的 σ 值对应粗糙尺度(低分辨率), 反之, 对应精细尺度(高分辨率)。

构造高斯差分尺度空间 (DOG scale-space):

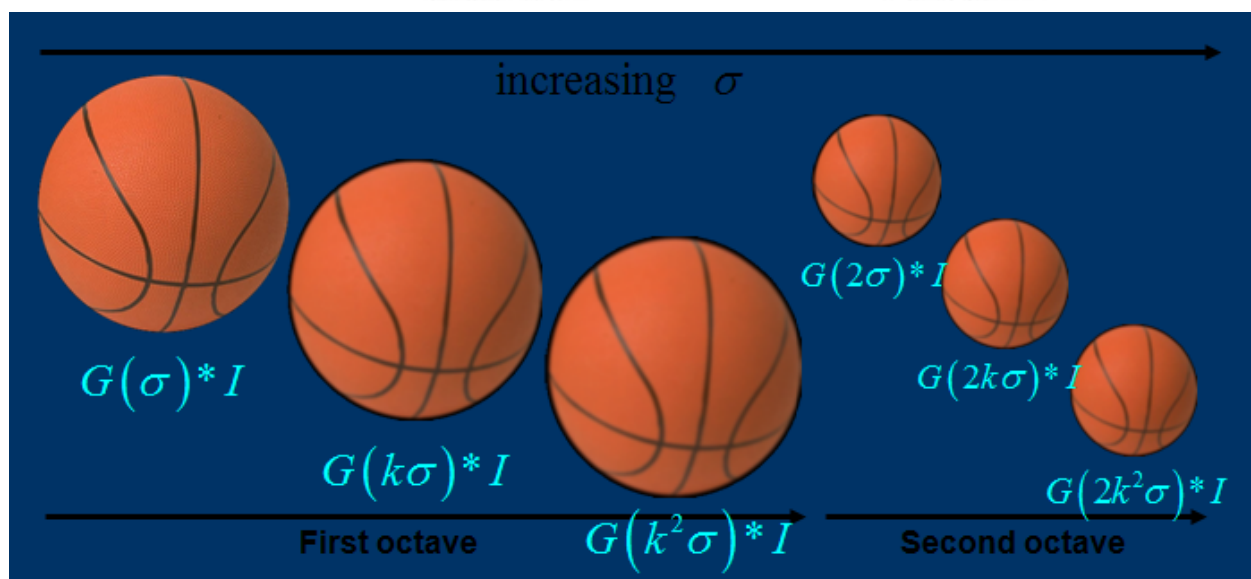
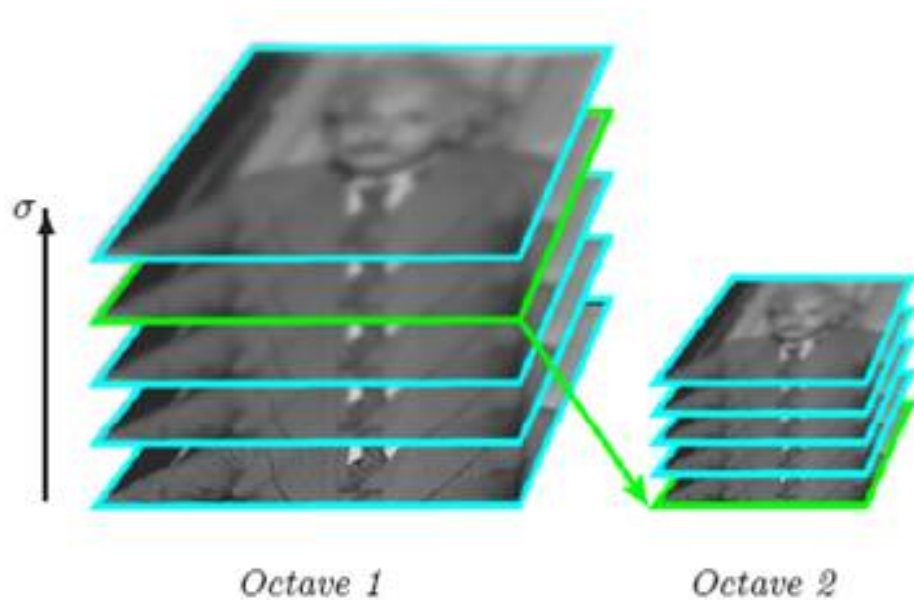
$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma)$$



$L(x, y, \sigma)$, $\sigma = 1.6$ a good tradeoff



$D(x, y, \sigma)$, $\sigma = 1.6$ a good tradeoff



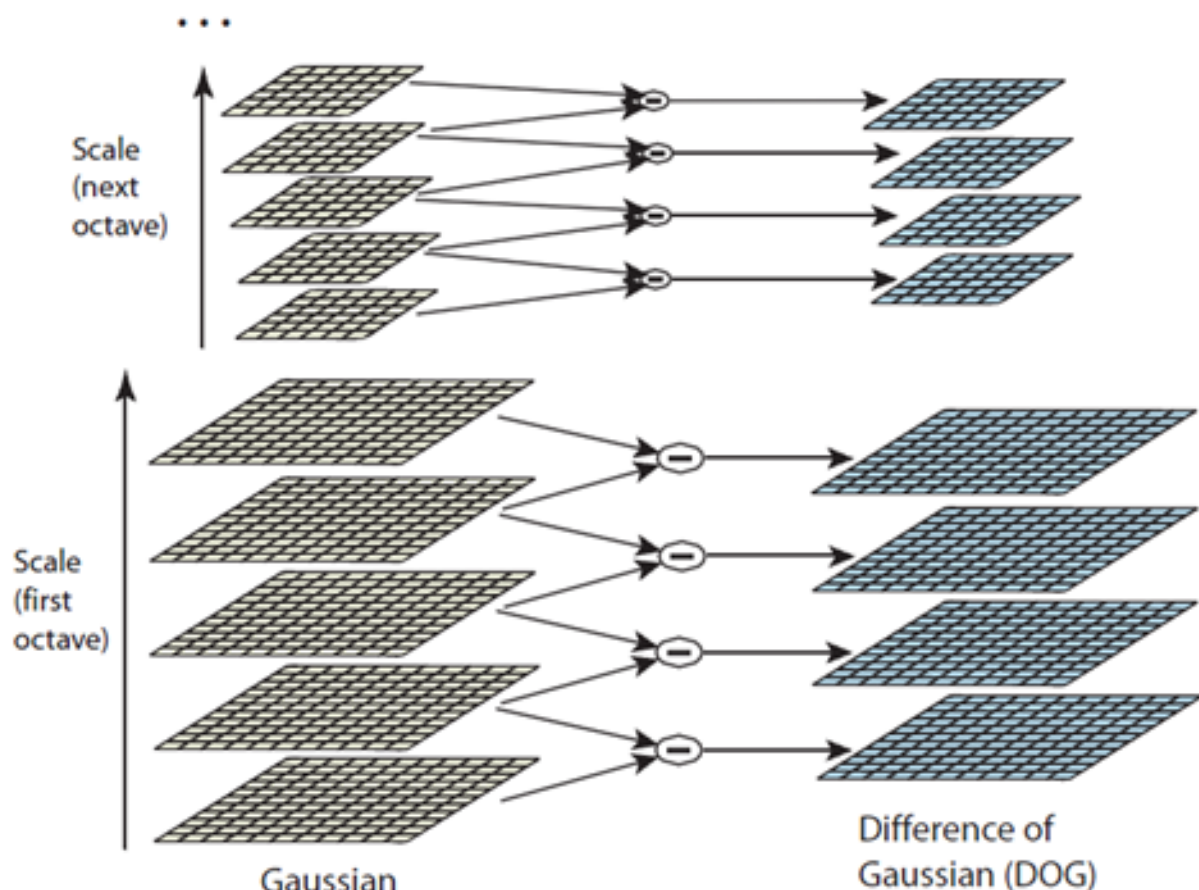
关于尺度空间的理解说明：图中的2是必须的，尺度空间是连续的。

在 Lowe 的论文中，

将第0层的初始尺度定为1.6，图片的初始尺度定为0.5。在检测极值点前对原始图像的高斯

平滑以致图像丢失高频信息，所以 Lowe 建议在建立尺度空间前首先对原始图像长宽扩展

一倍，以保留原始图像信息，增加特征点数量。尺度越大图像越模糊。



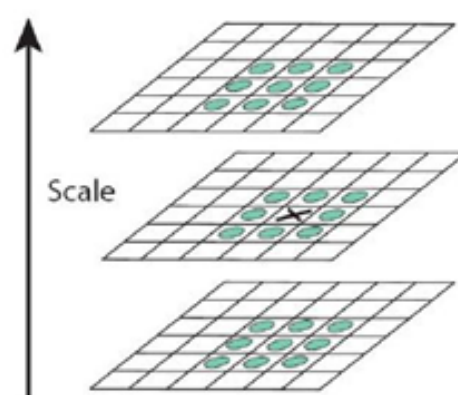
next octave 是由 first octave 降采样得到 (如2)

$$2^{i-1}(\sigma, k\sigma, k^2\sigma, \dots, k^{n-1}\sigma)$$

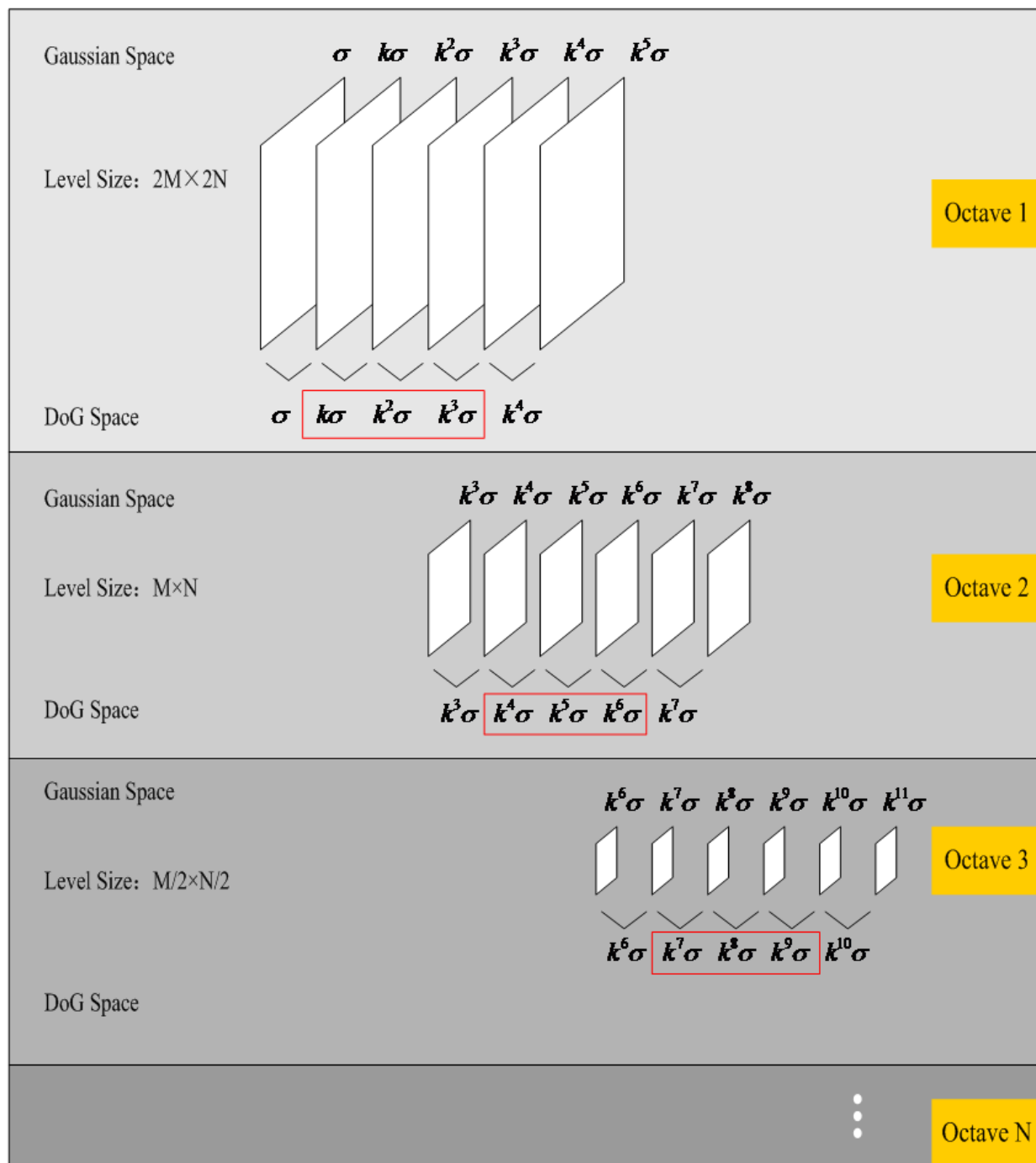
尺度空间的所有取值，s为每组层数，一般为3~5

● 检测尺度空间极值点

为了寻找尺度空间的极值点，每一个采样点要和它所有的相邻点比较，看其是否比它的图像域和尺度域的相邻点大或者小。如图，中间的检测点和它同尺度的8个相邻点和上下相邻尺度对应的 9×2 个点共26个点比较，以确保在尺度空间和二维图像空间都检测到极值点。如果一个点在DOG尺度空间本层以及上下两层的26个领域中是最大或最小值时，就认为该点是图像在该尺度下的一个特征点。



同一组中的相邻尺度（由于k的取值关系，肯定是上下层）之间进行寻找



在极值比较的过程中，每一组图像的首末两层是无法进行极值比较的，为了满足尺度变化的连续性，我们在每一组图像的顶层继续用高斯模糊生成了 3 幅图像，高斯金字塔有每组 $S+3$ 层图像。DOG 金字塔每组有 $S+2$ 层图像。

● 精确定位极值点

通过拟和三维二次函数以精确确定关键点的位置和尺度(达到亚像素精度),同时去除低对比度的关键点和不稳定的边缘响应点(因为DoG算子会产生较强的边缘响应),以增强匹配稳定性、提高抗噪声能力。

①空间尺度函数

$$D(x, y, \sigma) = D(x, y, \sigma) + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (1)$$

$$\text{求导,并令其为0,得到精确的位置}\hat{x}, \hat{x} = - \frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (2)$$

②在已经检测到的特征点中,要去掉低对比度的特征点和不稳定的边缘响应点。去除低对比度的点:把公式(2)代入公式(1),只取前两项可得:

$$D(\hat{x}) = D(x, y, \sigma) + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x}$$

若 $|D(\hat{x})| \geq 0.03$, 该特征点就保留下来, 否则丢弃。

③边缘响应的去除, 一个定义不好的高斯差分算子的极值在横跨边缘的地方有较大的主曲率, 而在垂直边缘的方向有较小的主曲率。主曲率通过一个2x2的Hessian矩阵H求出:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

导数由采样点相邻差估计得到。

D的主曲率和H的特征值成正比, 令 α 为最大特征值, β 为最小的特征值, 则

$$\text{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$

$$\text{Det}(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta.$$

令 $\alpha = r\beta$, 则:

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r},$$

If ratio > (r+1)2/(r), throw it out (SIFT uses r=10)

表示DOG金字塔中某一尺度的图像x方向求导两次

通过拟和三维二次函数以精确确定关键点的位置和尺度（达到亚像素精度）？

● 为每个关键点指定方向参数

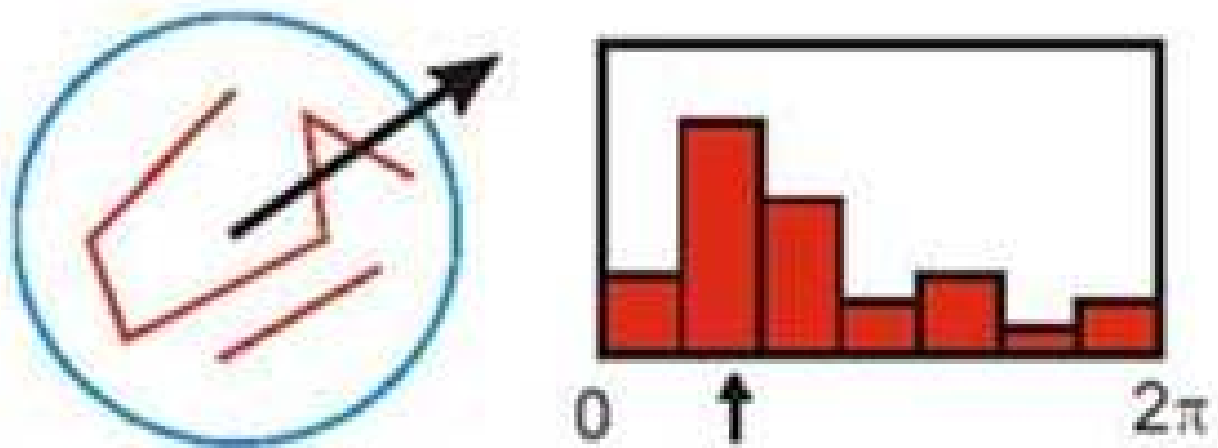
利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数，使算子具备旋转不变性。

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$
$$\theta(x, y) = \arctan 2((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y)))$$

式中为(x,y)处梯度的模值和方向公式。其中L所用的尺度为每个关键点各自所在的尺度。

至此，图像的关键点已检测完毕，每个关键点有三个信息：**位置**、**所处尺度**、**方向**。由此可以确定一个SIFT特征区域。

直方图中的峰值就是主方向，其他的达到最大值80%的方向可作为辅助方向

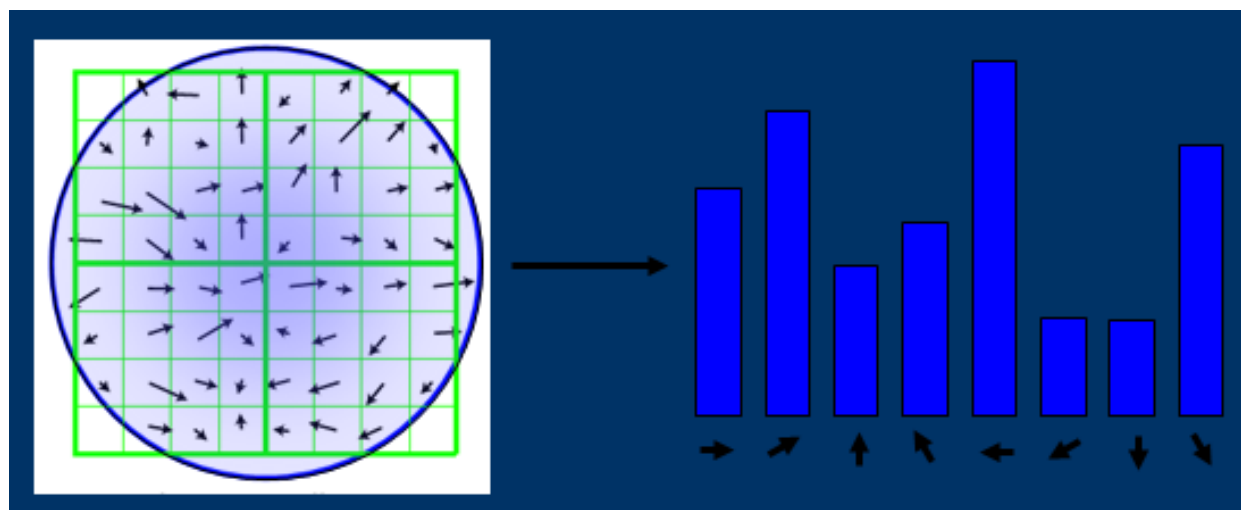


Identify peak and assign orientation and sum of magnitude to key point

The user may choose a threshold to exclude key points based on their assigned sum of

magnitudes.

利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数，使算子具备旋转不变性。以关键点为中心的邻域窗口内采样，并用直方图统计邻域像素的梯度方向。梯度直方图的范围是0~360度，其中每10度一个柱，总共36个柱。随着中心点越远的领域其对直方图的贡献也响应减小. Lowe论文中还提到要使用高斯函数对直方图进行平滑，减少突变的影响。



关键点描述子的生成步骤

旋转主方向

将坐标轴旋转为关键点方向，以确保旋转不变性

生成描述子

对于一个关键点产生128个数据，即最终形成128维的SIFT特征向量

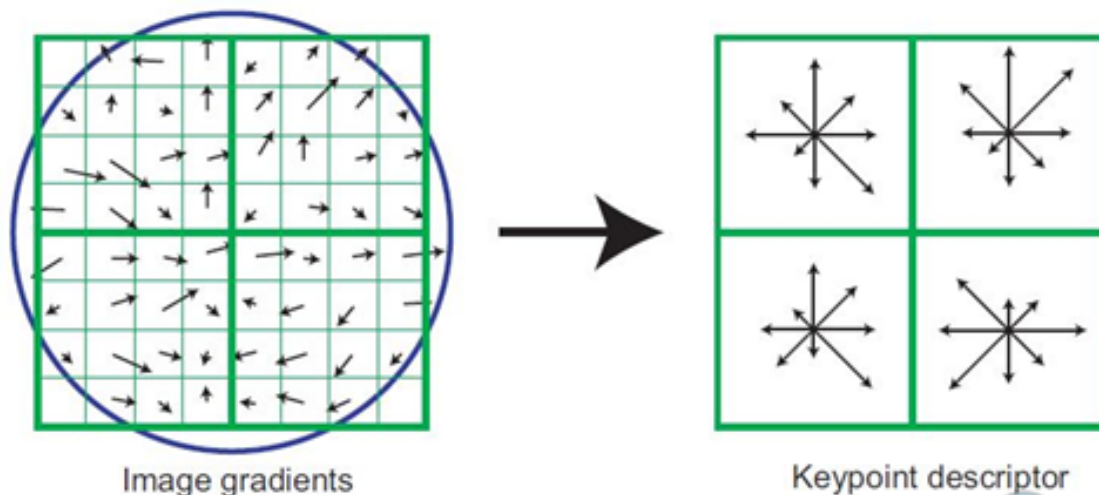
归一化处理

将特征向量的长度归一化，则可以进一步去除光照变化的影响。

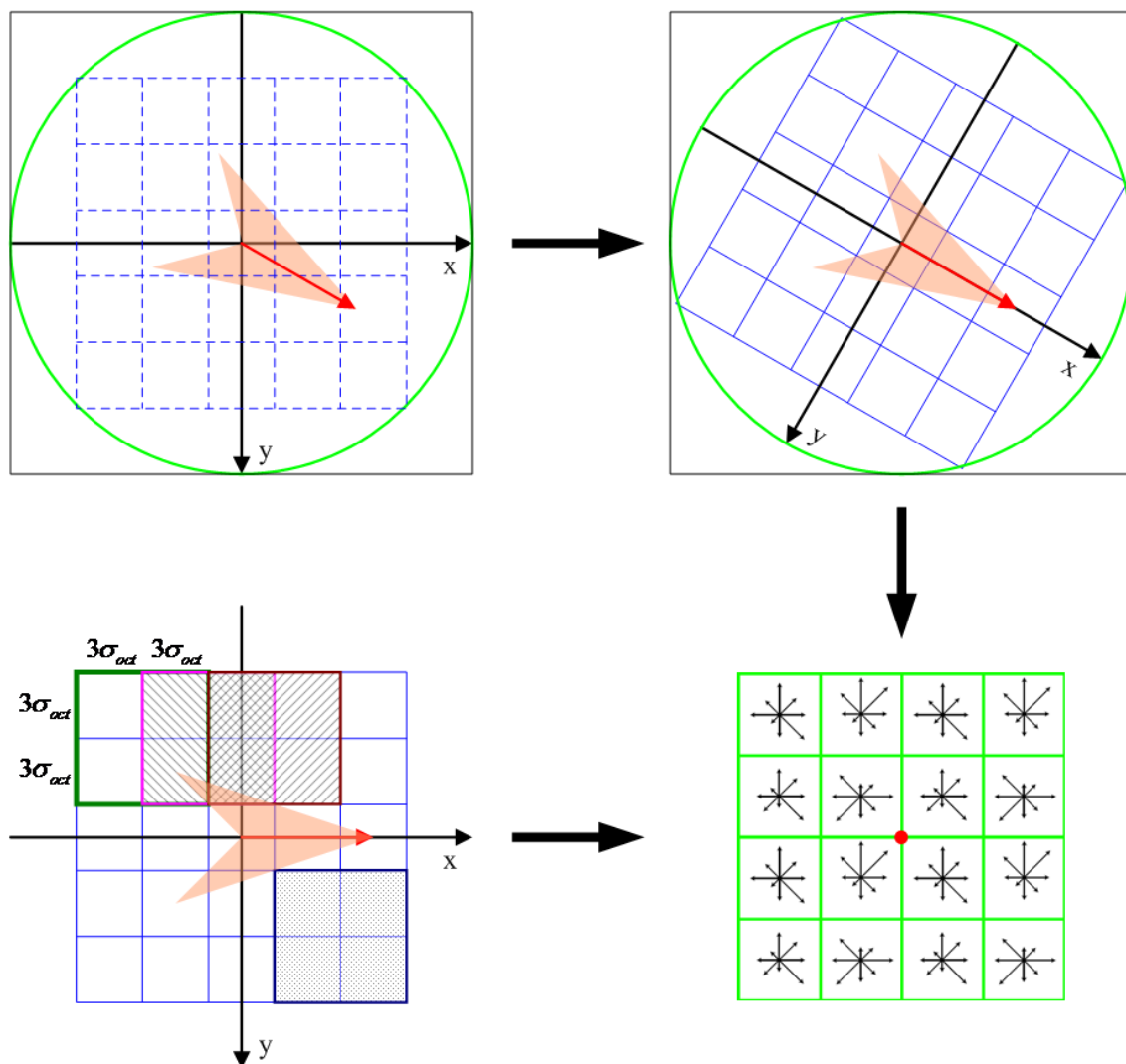
通过对关键点周围图像区域分块，计算块内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象，具有唯一性。

● 关键点描述子的生成

首先将坐标轴旋转为关键点的方向，以确保旋转不变性。以特征点为中心取 8×8 的邻域作为采样窗口，将采样点与特征点的相对方向通过高斯加权后归入包含8个方向直方图，最后获得 $2 \times 2 \times 8$ 的32维特征描述子。示意图如下：



每一个小格都代表了特征点邻域所在的尺度空间的一个像素，箭头方向代表了像素梯度方向，箭头长度代表该像素的幅值。然后在 4×4 的窗口内计算8个方向的梯度方向直方图。绘制每个梯度方向的累加可形成一个种子点。



每个直方图有8方向的梯度方向，每一个描述符包含一个位于关键点附近的四个直方图

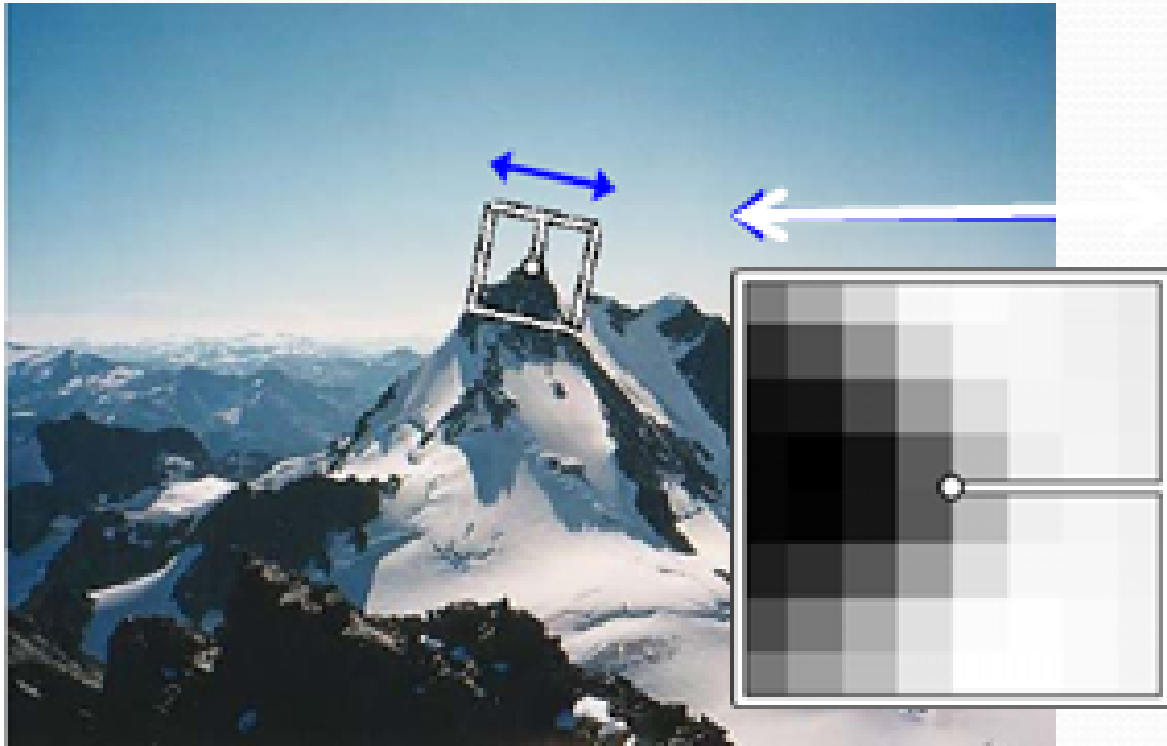
数组. 这就导致了SIFT的特征向量有128维.（先是一个4×4的来计算出一个直

方图，

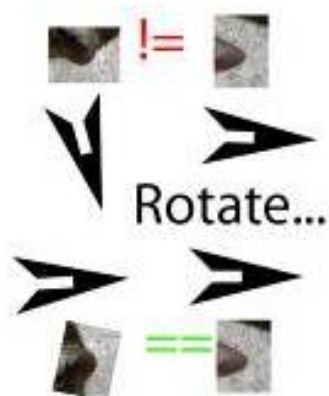
每个直方图有8个方向。所以是4×4×8=128维）将这个向量归一化之后，就进

一步

去除了光照的影响。



旋转为主方向



● 128维关键点描述子生成步骤

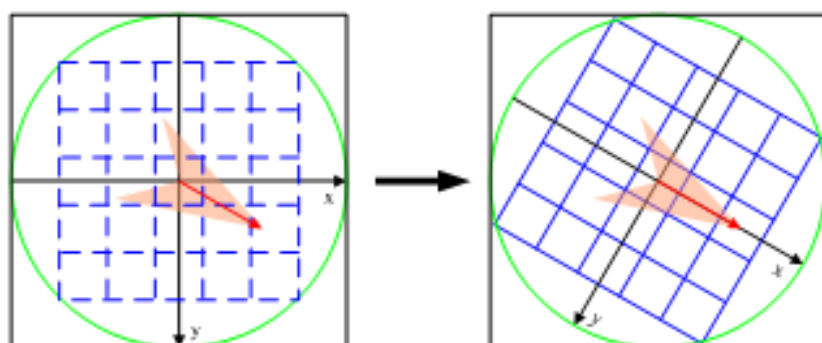
1. 确定计算描述子所需的图像区域

描述子梯度方向直方图由关键点所在尺度的模糊图像计算产生。图像区域的半径通过下式计算：

$$radius = \frac{3\sigma_{oct} \times \sqrt{2} \times (d+1) + 1}{2}$$

σ_{oct} 是关键点所在组 (octave) 的组内尺度, $d=4$

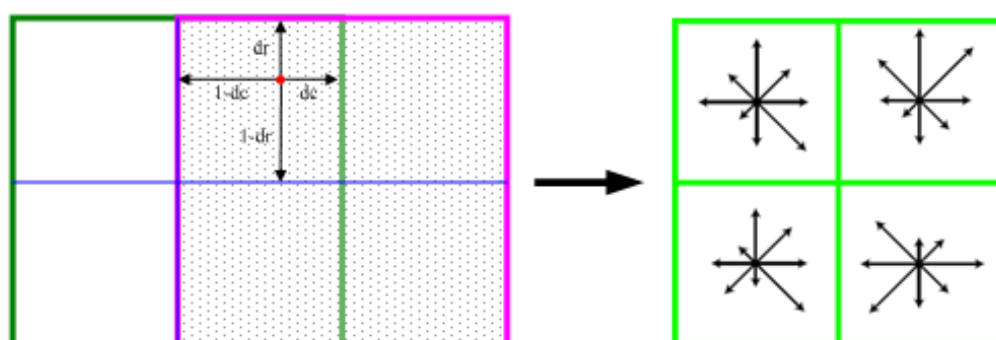
2. 将坐标移至关键点主方向



那么旋转角度后新坐标为：

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \times \begin{pmatrix} x \\ y \end{pmatrix}$$

3. 在图像半径区域内对每个像素点求其梯度幅值和方向，然后对每个梯度幅值乘以高斯权重参数，生成方向直方图。



$$weight = |grad(I_o(x, y))| \times \exp\left(-\frac{x_k^2 + y_k^2}{2\sigma_w}\right) \times (1-d_r) \times (1-d_c) \times (1-d_o)$$

x_k : 该点与关键点的列距离;

y_k : 该点与关键点的行距离;

σ_w : 等于描述子窗口宽度 $3\sigma \times$ 直方图列数 (取4) 的一半;

4. 在窗口宽度为2X2的区域内计算8个方向的梯度方向直方图，绘制每个梯度方向的累加值，即可形成一个种子点。然后再在下一个2X2的区域内进行直方图统计，形成下一个种子点，共生成16个种子点。

5. 描述子向量元素门限化及门限化后的描述子向量规范化。

描述子向量元素门限化:

方向直方图每个方向上梯度幅值限制在一定门限值以下（门限一般取0.2）。

描述子向量元素规范化:

$W = (w_1, w_2, \dots, w_{128})$ 为得到的128描述子向量,

$L = (l_1, l_2, \dots, l_{128})$ 为规范化后的向量

$$l_j = w_j / \sqrt{\sum_{i=1}^{128} w_i^2} \quad j = 1, 2, \dots, 128$$

基本概念及一些补充

什么是局部特征？

- 局部特征从总体上说是图像或在视觉领域中一些有别于其周围的地方
- 局部特征通常是描述一块区域，使其能具有高可区分度
- 局部特征的好坏直接会决定着后面分类、识别是否会得到一个好的结果

局部特征需具备的特性

- 重复性
- 可区分性
- 准确性
- 数量以及效率
- 不变性

局部特征提取算法-sift

• SIFT算法由D. G. Lowe 1999年提出，2004年完善总结。后来Y. Ke将其描述子部分用PCA代替直方图的方式，对其进行改进。

• SIFT算法是一种提取局部特征的算法，在尺度空间寻找极值点，提取位置，尺度，旋转不变量

• SIFT特征是图像的局部特征，其对旋转、尺度缩放、亮度变化保持不变性，对视角变化、仿射变换、噪声也保持一定程度的稳定性。

- 独特性好，信息量丰富，适用于在海量特征数据库中进行快速、准确的匹配。

- 多量性，即使少数的几个物体也可以产生大量SIFT特征向量。
- 可扩展性，可以很方便的与其他形式的特征向量进行联合。

尺度空间理论

- 尺度空间理论目的是模拟图像数据的多尺度特征
- 其基本思想是在视觉信息图像信息处理模型中引入一个被视为尺度的参数，通过连续变化尺度参数获得不同尺度下的视觉处理信息，然后综合这些信息以深入地挖掘图像的本质特征。

描述子生成的细节

- 以极值点为中心点，并且以此点所处于的高斯尺度sigma值作为半径因子。对于远离中心点的梯度值降低对其所处区域的直方图的贡献，防止一些突变的影响。
- 每个极值点对其进行三线性插值，这样可以把此极值点的贡献均衡的分到直方图中相邻的柱子上

归一化处理

- 在求出 $4 \times 4 \times 8$ 的128维特征向量后，此时SIFT特征向量已经去除了尺度变化、旋转等几何变形因素的影响。而图像的对比度变化相当于每个像素点乘上一个因子，光照变化是每个像素点加上一个值，但这些对图像归一化的梯度没有影响。因此将特征向量的长度归一化，则可以进一步去除光照变化的影响。
- 对于一些非线性的光照变化，SIFT并不具备不变性，但由于这类变化影响的主要是梯度的幅值变化，对梯度的方向影响较小，因此作者通过限制梯度幅值的值来减少这类变化造成的影响。

PCA-SIFT算法

- PCA-SIFT与标准SIFT有相同的亚像素位置，尺度和主方向。但在第4步计算描述子的设计，采用的主成分分析的技术。
- 下面介绍一下其特征描述子计算的部分：
 - 用特征点周围的 41×41 的像斑计算它的主元，并用PCA-SIFT将原来的 $2 \times 39 \times 39$ 维的向量降成20维，以达到更精确的表示方式。
 - 它的主要步骤为，对每一个关键点：在关键点周围提取一个 41×41 的像斑于给定的尺度，旋转到它的主方向；计算 39×39 水平和垂直的梯度，形成一个大小为3042的矢量；用预先计算好的投影矩阵 $n \times 3042$ 与此矢量相乘；这样生成一个大小为n的PCA-SIFT描述子。