

机器视觉测量与建模

Machine vision based surveying and modelling



李明磊

南京航空航天大学电子信息工程学院

E-mail: minglei_li@nuaa.edu.cn

1



4. 特征提取与匹配

4.1 梯度信息提取

4.2 点特征提取

4.3 特征匹配技术

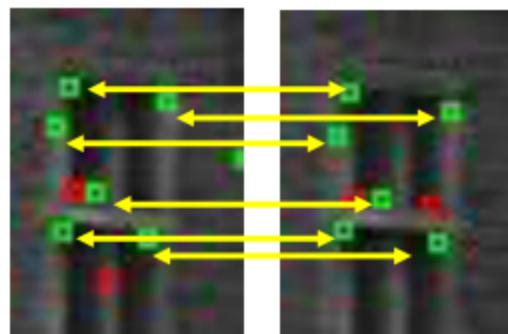
李明磊@nuaa

2

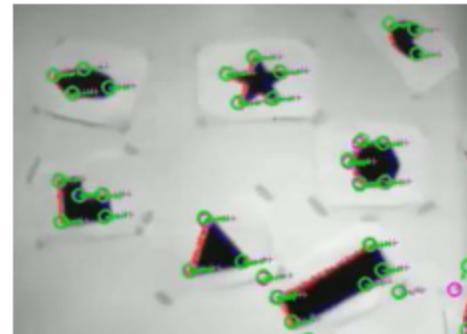
4.2 点特征提取

What is a good feature?

Image-to-image correspondences are key to passive triangulation-based 3D reconstruction



Extract features **independently** and then match by comparing descriptors



Extract features in first images and then try to **find same** feature back in next view

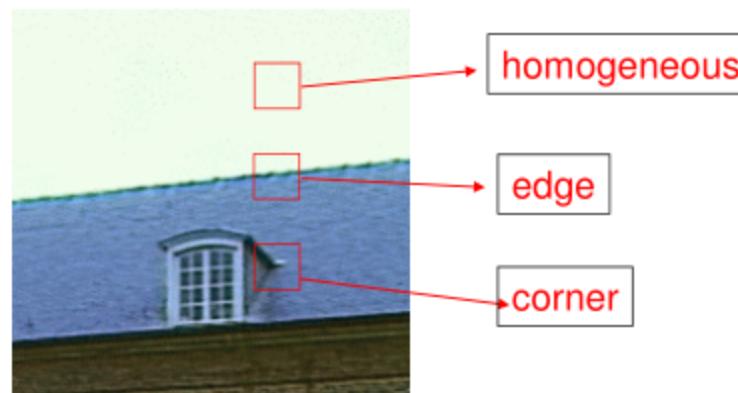
Feature matching vs. tracking

李明磊@nuaa

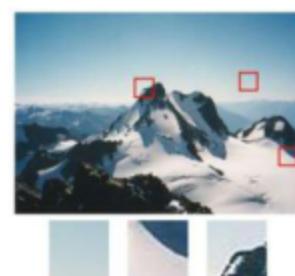
3

4.2 点特征提取

Find points that differ as much as possible from all neighboring points



差异性—可检测
特征点应该呈现出区别于非特征点的明显特征



李明磊@nuaa

重复性—可匹配
对应同一三维点的特征点应该在不同视角中被重复检测到



4



• 特征提取与匹配

- 特征点怎么来?
 - 特征点提取算法**（Harris、SIFT、SURF、ORB等）
 - 特征提取算法还会计算每个特征点周围的图像统计信息
 - 特征描述子**（Descriptor）
 - 描述子能够区别每个关键点——可以用于匹配



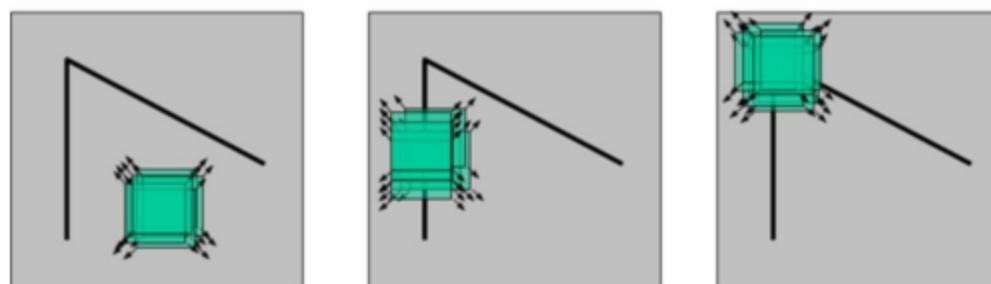
李明磊@nuaa

5



• Harris角点检测

Harris角点是由**Harris**和**Stephens**在1998年提出的。
通过计算像素点在各个方向上的灰度变化来确定角点位置



窗口在不同图像上移动所产生的灰度变化不同

假设 $E(u, v)$ 代表了在以某一点 (x, y) 为中心的窗口像素向 x 方向移动 u 分量，
向 y 方向移动 v 分量后的亮度值变化量的累加值，
所以 $E(u, v)$ 是一种平方差异和（sum squared differences, SSD）：

$$E(u, v) = \sum_{x,y} w(x, y)[I(x + u, y + v) - I(x, y)]^2$$

李明磊@nuaa

6



• Harris角点检测

$E(u, v)$ 是一种平方差异和 (sum squared differences, SSD) :

$$E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2$$

- 这里的 $w(x, y)$ 是一个窗口权函数，最简单直接令 $w(x, y) = 1$ 。
- 但为了更精确， $w(x, y)$ 可以设置为二维高斯权函数
- 如果角点中心位于窗口中心， $E(u, v)$ 会明显比偏离角点中心的窗口的变化值大。可以以此来确定角点的像素级位置。

根据泰勒级数展开公式，能够将式中的被减项表示为

$$I(x + u, y + v) \approx I(x, y) + uI_x + vI_y$$

$$\begin{aligned} I_x &= \frac{\partial I}{\partial x} \\ I_y &= \frac{\partial I}{\partial y} \end{aligned}$$

$$E(u, v) = \sum_{x,y} w(x, y) [uI_x + vI_y]^2$$

李明磊@nuaa

7



• Harris角点检测

$$\text{SSD: } E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2$$

$$E(u, v) \approx \sum_{x,y} w(x, y) [uI_x + vI_y]^2$$

$$E(u, v) = \sum_{x,y} w(x, y) (u^2 I_x^2 + 2uv I_x I_y + v^2 I_y^2)$$

$$E(u, v) = w(x, y) [u \quad v] \left(\sum_{x,y} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix}$$

$$E(u, v) = [u \quad v] \mathbf{M} \begin{bmatrix} u \\ v \end{bmatrix} \quad \mathbf{M} = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Harris角点并不需要直接求 $E(u, v)$ 来判断是否是角点，而是可以通过求**矩阵的特征值**来判断。

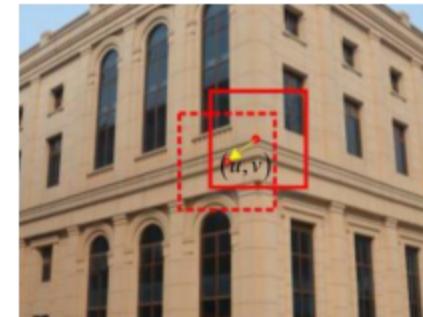
李明磊@nuaa

8



• Harris角点检测

- 当 \mathbf{M} 的特征值
- 一大一小，边缘区域
 - 都很小，平坦区域
 - 都很大，角点区域



假设两个特征值分别为 λ_1 和 λ_2 ，则计算 R

$$R = \det \mathbf{M} - k(\text{trace } \mathbf{M})^2$$

$$\det \mathbf{M} = \lambda_1 \lambda_2$$

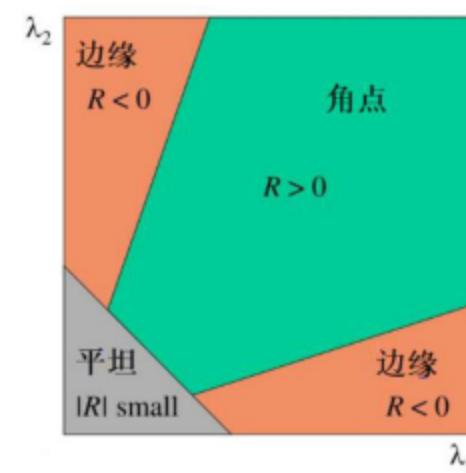
$$\text{trace } \mathbf{M} = \lambda_1 + \lambda_2$$

越小检测越敏感

k 是一个参数，一般其值在0.04到0.06之间

矩阵的特征值反映了两个相互垂直的方向上的变化情况

- R 根据需要设定阈值，过滤掉低响应度的角点
- 在 3×3 或 5×5 的邻域内进行非极大值抑制，局部最大值点即为图像中的角点。



9

李明磊@nuaa



• Harris角点检测

- 流程为：
1. 将原始图像转化为**灰度影像**；
 2. 应用高斯滤波器来**平滑噪声**；
 3. 应用**Sobel算子**寻找灰度图像中每个像素的 x 方向和 y 方向的梯度值；
 4. 对于灰度图像中的每个像素 p ，考虑其周围有一个 3×3 的窗口，并计算**角点强度函数 R** ，称之为**Harris分数**。
 5. 查找超过某个阈值并且是某窗口中局部最大值的像素（非极大值抑制以防止重复特征）

计算图像水平和垂直方向的梯度

$$\frac{\partial L(x, y, \sigma_p)}{\partial x} = \frac{\partial G(x, y, \sigma_p)}{\partial x} * I(x, y), \quad \frac{\partial L(x, y, \sigma_p)}{\partial y} = \frac{\partial G(x, y, \sigma_p)}{\partial y} * I(x, y)$$

计算每个像素位置的Harris矩阵

$$H = G(x, y, \sigma_p) * \begin{bmatrix} \left(\frac{\partial L(x, y, \sigma_p)}{\partial x} \right)^2 & \frac{\partial L(x, y, \sigma_p)}{\partial x} \frac{\partial L(x, y, \sigma_p)}{\partial y} \\ \frac{\partial L(x, y, \sigma_p)}{\partial x} \frac{\partial L(x, y, \sigma_p)}{\partial y} & \left(\frac{\partial L(x, y, \sigma_p)}{\partial y} \right)^2 \end{bmatrix}$$

找到Harris角点响应值大于给定阈值且局部最大的位置作为特征点

计算每个像素位置的Harris角点响应值

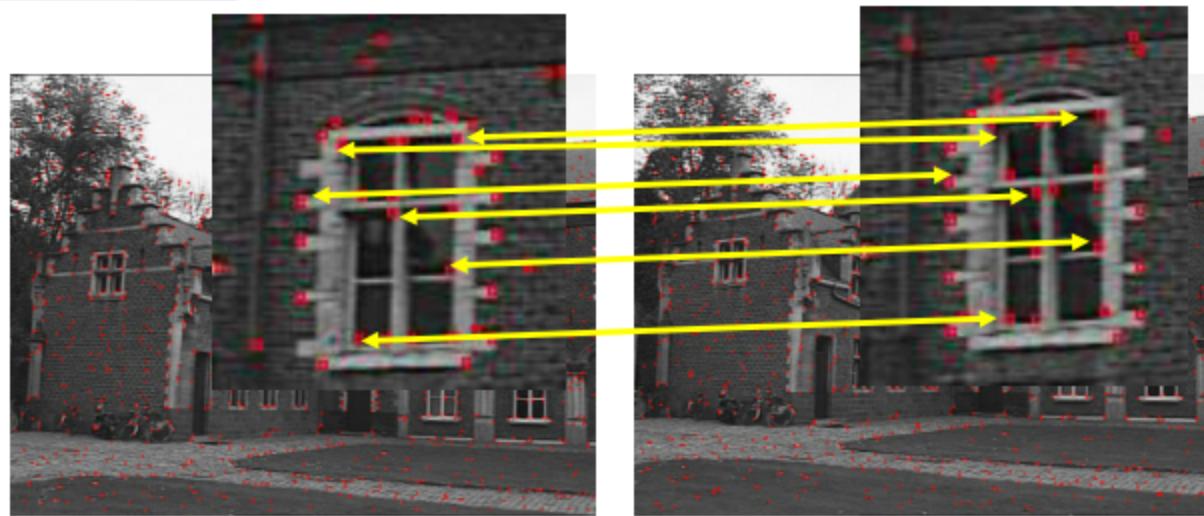
$$C = \det(H) - k \text{trace}(H)^2 = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2$$

李明磊@nuaa

10



- Harris角点检测



- Only use local maxima
- Subpixel accuracy through second order surface fitting
- Select a fix number of strongest features over whole image & over each tile (e.g. 1000/image, 2/tile)

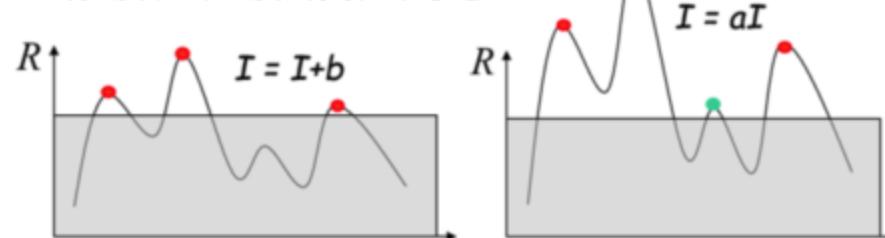
李明磊@nuaa

11



- Harris角点检测

- 对亮度和对比度的变化不敏感



- 不具有尺度不变性



李明磊@nuaa

12



• 特征提取与匹配

- 特征点怎么来?
 - 特征点提取算法** (Harris、 SIFT、 SURF、 ORB等)
- 特征提取算法还会计算每个特征点周围的图像统计信息
 - 特征描述子** (Descriptor)
- 描述子能够区别每个关键点——可以用于匹配



Challenge: Wide baseline matching

- Requirement to cope with larger variations between images
 - Translation, rotation, scaling
 - 透视收缩 Foreshortening
 - 非漫反射 Non-diffuse reflections
 - 光照 Illumination
- } 几何 geometric transformations
- } 光度 photometric changes

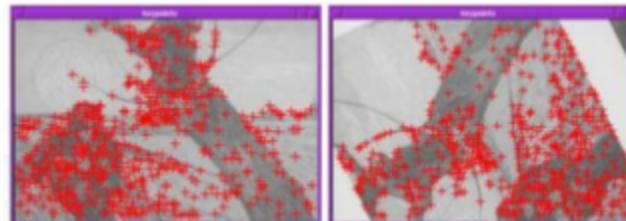




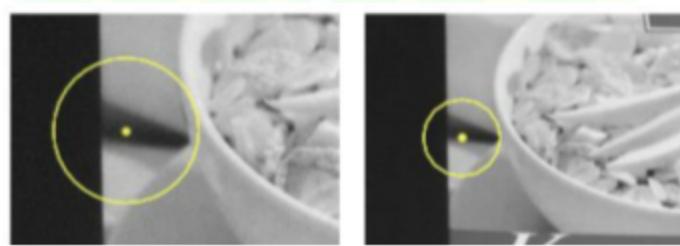
Challenge: Wide baseline matching

Invariant detectors

Rotation invariant



Scale invariant



Affine invariant



李明磊@nuaa

15



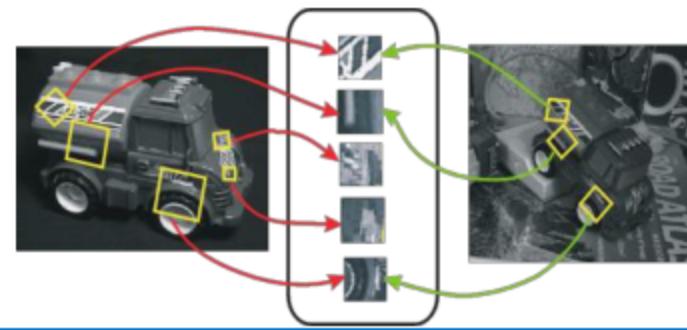
- SIFT特征检测

Lowe's SIFT features (Lowe, ICCV99)

Recover features with **position, orientation and scale**

SIFT, 即尺度不变特征变换 (Scale-invariant feature transform, SIFT)

1999年由David Lowe首先发表于计算机视觉国际会议 (International Conference on Computer Vision, ICCV) , 2004年再次经David Lowe整理完善后发表于International journal of computer vision (IJCV)



李明磊@nuaa

16



- SIFT特征检测

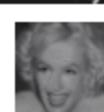
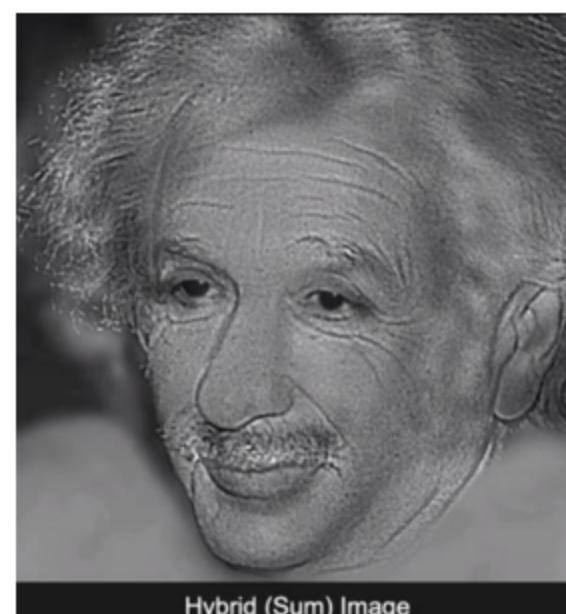
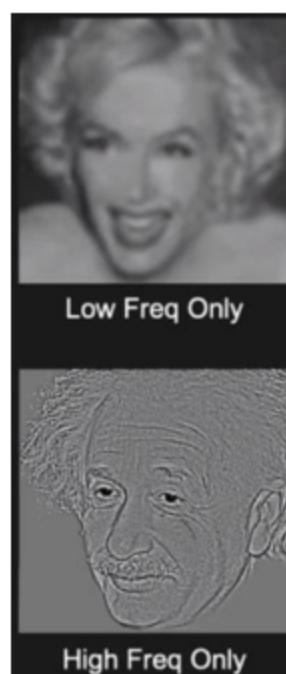
SIFT特征是图像的局部特征，

- 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。

- 1) 生成图像的**尺度空间**
- 2) 检测尺度空间**变化极值点**
- 3) 计算极值点的**精确定位**
- 4) 为每个关键点指定**方向参数**
- 5) 计算生成关键点**描述子**



尺度空间(scale space)思想最早是由Iijima于1962年提出的，后经Witkin和Koenderink等人的推广逐渐得到关注，在计算机视觉领域使用广泛。



• SIFT特征检测

首先了解高斯平滑

$L(x, y, \sigma)$ 定义为原始图像 $I(x, y)$ 与一个可变尺度的2维高斯函数 $G(x, y, \sigma)$ 卷积运算

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp \frac{-(x^2 + y^2)}{2\sigma^2}$$

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

高斯模板矩阵的值，与原图像做卷积，即可获得原图像的平滑(高斯模糊)图像。为了确保模板矩阵中的元素在[0,1]之间，需将模板矩阵归一化。

5*5 的高斯模板($\sigma=0.6$)

6.58573e-006	0.000424781	0.00170354	0.000424781	6.58573e-006
0.000424781	0.0273984	0.109878	0.0273984	0.000424781
0.00170354	0.109878	0.440655	0.109878	0.00170354
0.000424781	0.0273984	0.109878	0.0273984	0.000424781
6.58573e-006	0.000424781	0.00170354	0.000424781	6.58573e-006



二维高斯模糊效果图

李明磊@nuaa

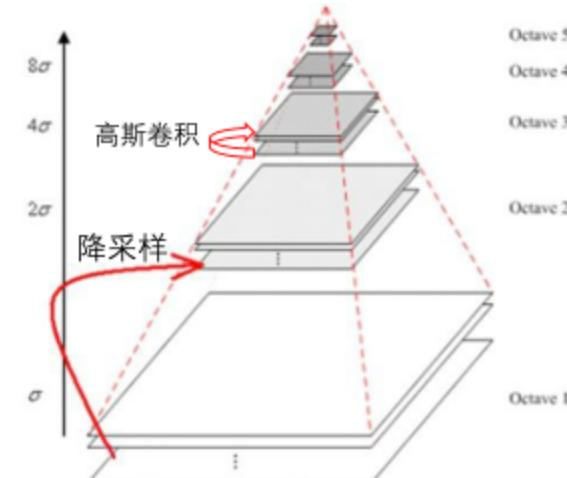
19

• SIFT特征检测

1) 尺度空间

图像的金字塔模型：

- 将原始图像不断降阶采样，得到一系列大小不一的图像，由大到小，从下到上构成的塔状模型。
- 原图像为金字塔的第一层，每次降采样所得到的新图像为金字塔的上一层(每层一张图像)，每个金字塔共n层。



为了让尺度体现其连续性，高斯金字塔在简单降采样的基础上加上了高斯滤波。

高斯金字塔的构建分为两步：



李明磊@nuaa

20



- SIFT特征检测

1) 尺度空间

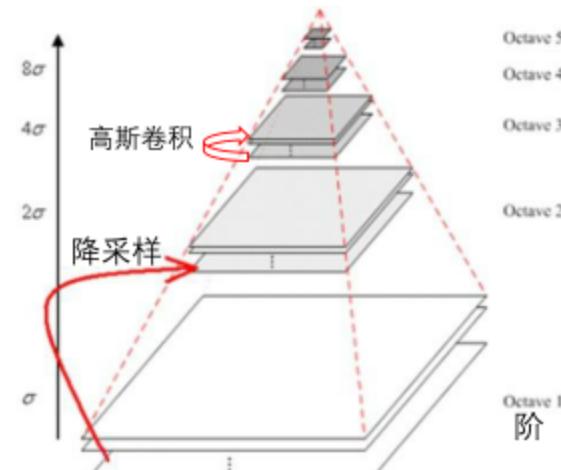
高斯金字塔的构建分为两步：

高斯平滑 → 降采样

将图像金字塔每层的一张图像使用不同参数做高斯模糊，Octave表示一幅图像可产生的图像阶数（组）

$$\sigma(s) = \sigma_0 2^{s/S}$$

σ 表示尺度空间坐标， s 表示 sub-level 层坐标， σ_0 表示初始尺度， S 是“每”阶的层数（一般为 3~5）



σ 是尺度空间因子，值越小表示图像被平滑的越少，相应的尺度也就越小。
大尺度对应于图像的概貌特征，小尺度对应于图像的细节特征。

李明磊@nuaa

21



- SIFT特征检测

1) 尺度空间

检测特征点较好的算子是高斯拉普拉斯 $\Delta^2 G$ (Laplacian of Gaussian, LoG)。

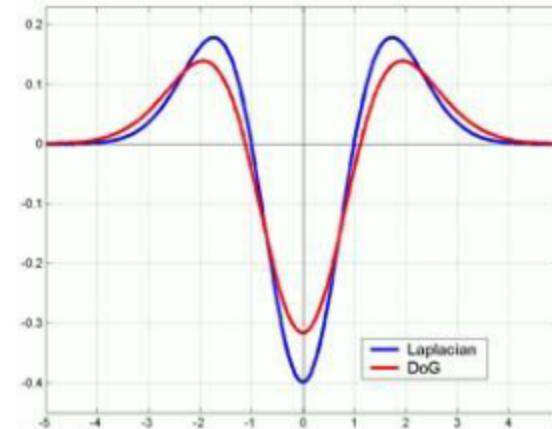
LoG 运用到影像中的特征检测时，运算量过大，

使用差分高斯 (Difference of Gaussina, DoG) 来近似计算 LoG

DOG图像描绘的是目标的轮廓。

先对不同尺度的高斯核（比例因子为 k ）进行差分，然后将差分结果和影像进行卷积生成高斯差分尺度空间 $D(x, y, \sigma)$

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



高斯拉普拉斯和高斯差分的比较

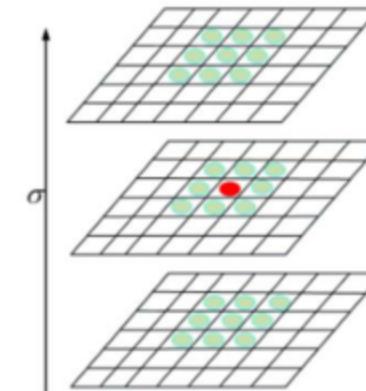
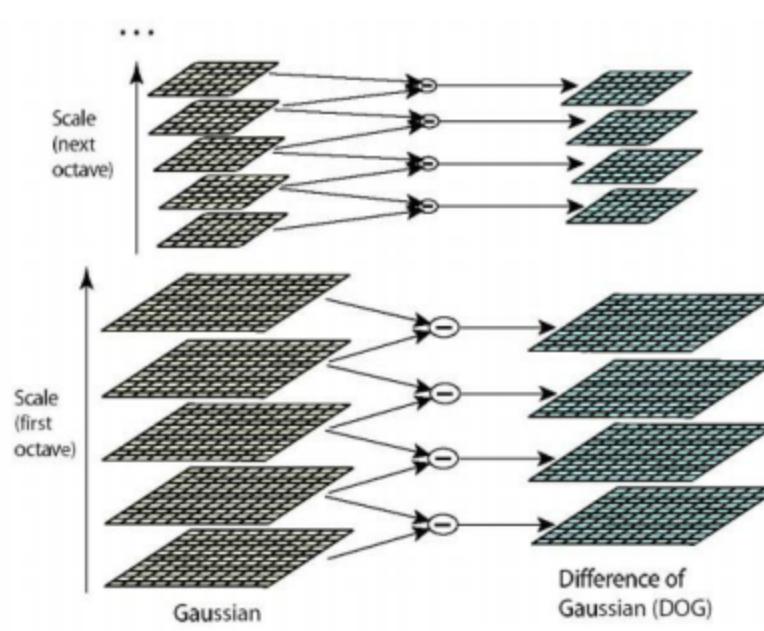
李明磊@nuaa

22



特征点是由**DOG空间的局部极值点**组成的。

为了寻找DoG函数的极值点，每一个像素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。



和其**图像域、尺度域**的所有相邻点进行比较

$$8 + 9 \times 2 = 26$$

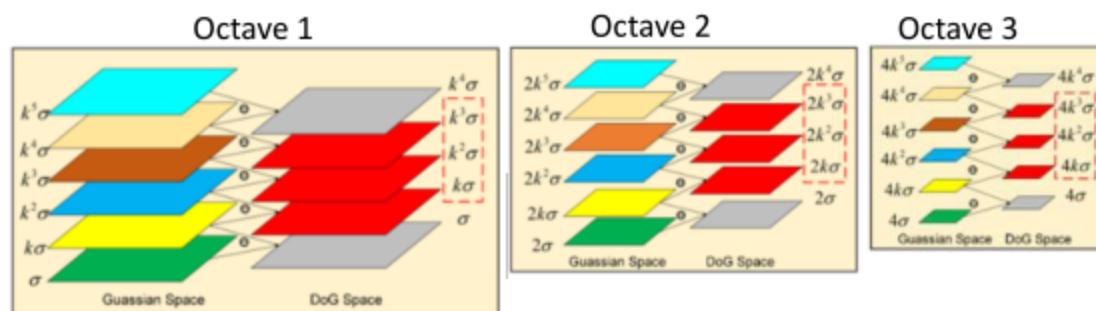
确保在尺度空间和二维图像空间都检测到极值点

李明磊@nuaa

23



举例



阶数 $O = 3$

每阶有效差分数 $S = 3$

每阶层数 $N = S + 3$

高斯空间 $2^{o-1}(\sigma, k\sigma, \dots, k^{S+2}\sigma), k = 2^{\frac{1}{S}}$

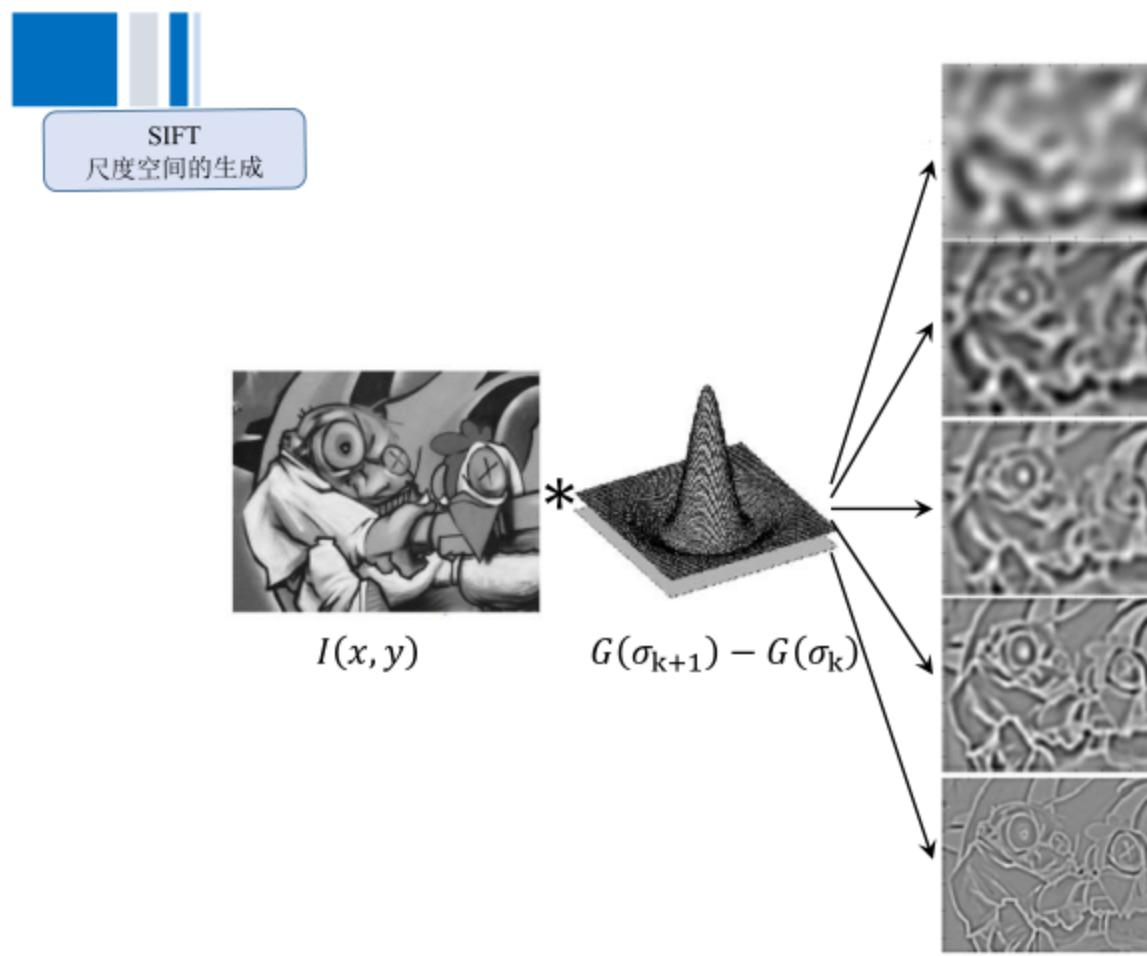
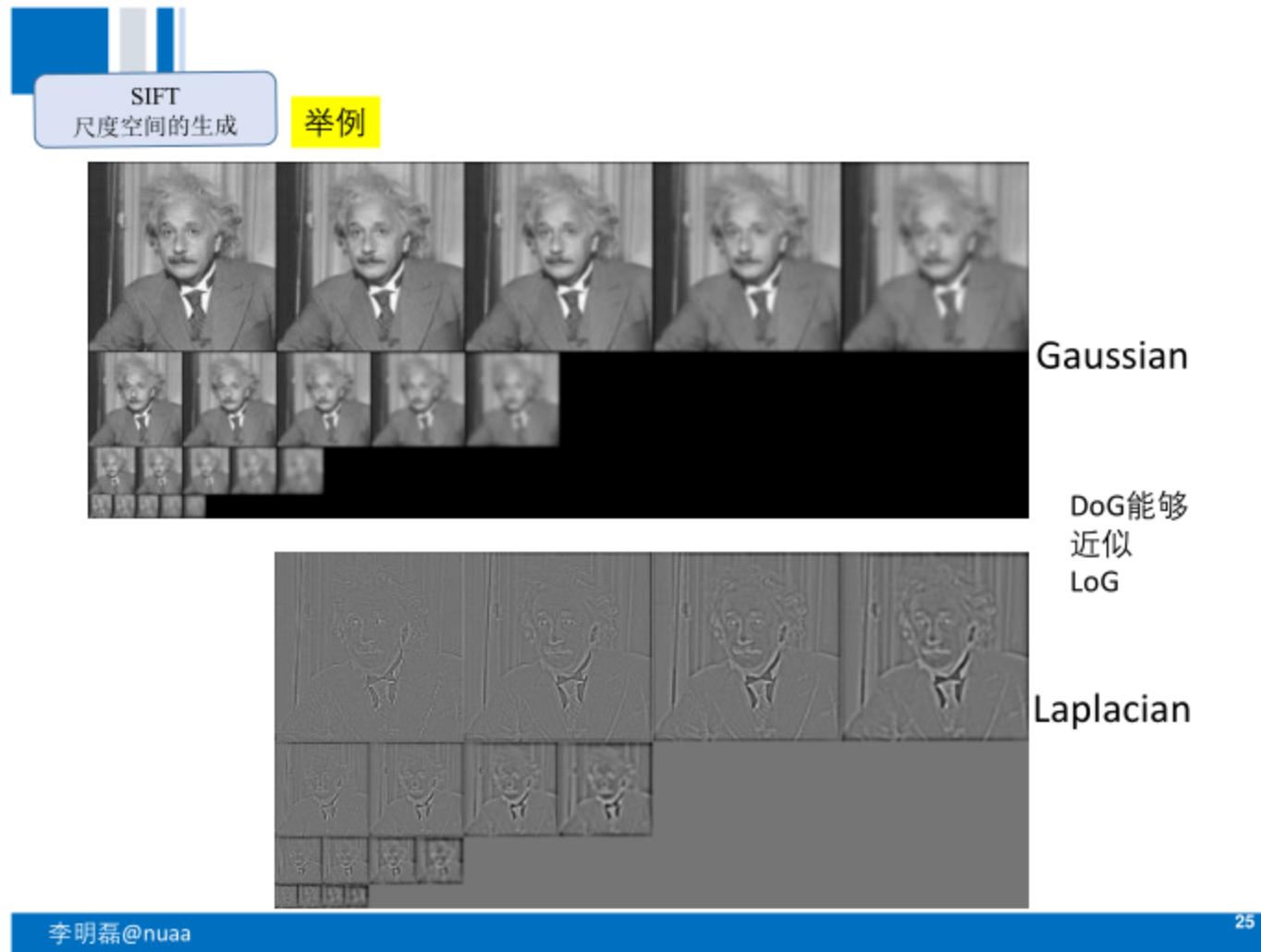
高斯差分 $2^{o-1}(\sigma, k\sigma, \dots, k^{S+1}\sigma), k = 2^{\frac{1}{S}}$

有效差分 $2^{o-1}(k\sigma, \dots, k^S\sigma), k = 2^{\frac{1}{S}}$

上一阶图像中的第一张是由其下面一阶图像倒数第三张降采样得到
→ 保证尺度变化的连续性

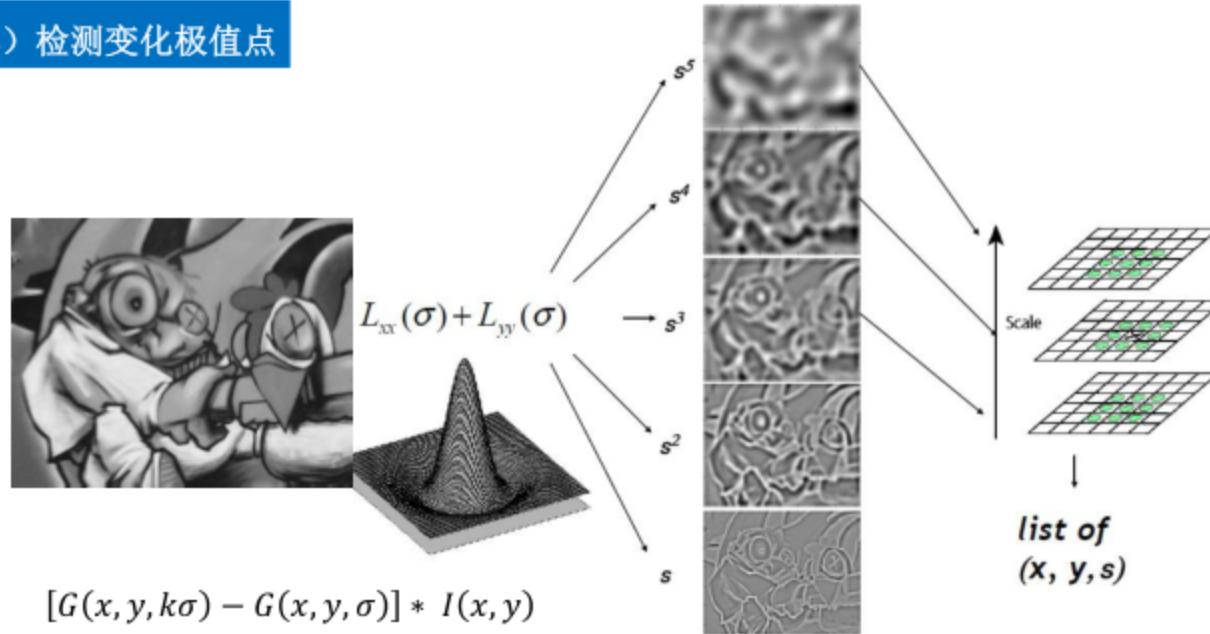
李明磊@nuaa

24





2) 检测变化极值点



DOG图像描绘的是目标的轮廓/边缘/梯度信息。

S是每阶的层数（一般为3~5）



3) 极值点精确定位

确定亚像素特征点位置，同时去除低对比度的关键点和不稳定的边缘响应点

通过尺度空间DoG函数进行曲线拟合寻找极值点，本质是去掉DoG局部曲率非常不对称的点。

$$\begin{aligned}
 f(x) &= f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(x_0) (x - x_0) \\
 \text{DoG值} \\
 f(\delta x) &= \nabla f(x_0)^T \delta x + \frac{1}{2} \delta x^T \nabla^2 f(x_0) \delta x \quad \rightarrow \boxed{\text{泰勒展开, 令 } \delta x = x - x_0} \\
 \frac{\partial f(\delta x)}{\partial \delta x} &= \nabla f^T(x_0) + \nabla^2 f(x_0) \delta x = \theta \quad \rightarrow \boxed{\text{求导数, 令等于0}} \\
 \delta x &= -\nabla^2 f(x_0)^{-1} \nabla f^T(x_0) \quad \Rightarrow \quad f(x) = f(x_0) + \frac{1}{2} \nabla f(x_0)^T (x - x_0) \\
 &\quad \boxed{|f(x)| \geq 0.04} \quad \text{则保留}
 \end{aligned}$$

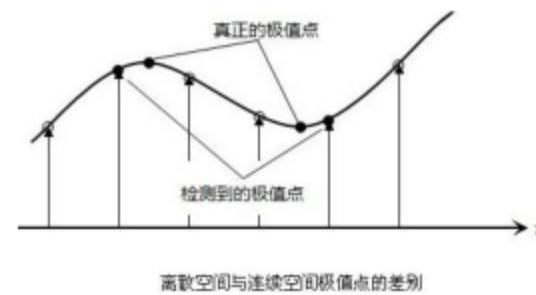


3) 极值点精确定位

在边缘梯度的方向上主曲率值比较大，而沿着边缘方向则主曲率值较小。候选特征点的DoG函数 $D(x)$ 的主曲率与 2×2 的Hessian矩阵 H 的特征值成正比。

(1) 确定亚像素特征点位置，拟合三维二次函数来精确确定关键点的位置和尺度

为了提高关键点的稳定性，需要对尺度空间DoG函数进行曲线拟合。利用DoG函数在尺度空间的Taylor展开式，求导并让方程等于零，可以得到极值点的偏移量



(2) 去除低对比度的关键点和不稳定的边缘响应点

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

D_{xx} 表示DOG金字塔中某一尺度的图像 x 方向求导两次

$$\frac{\text{trace}(H)^2}{\text{Det}(H)} = \frac{(\lambda_1 + \lambda_2)^2}{\lambda_1 \lambda_2} \quad \frac{\text{trace}(H)^2}{\text{Det}(H)} < T \quad \text{则保留, 反之剔除}$$

Lowe论文中建议阈值 T 为1.2

李明磊@nuaa

29



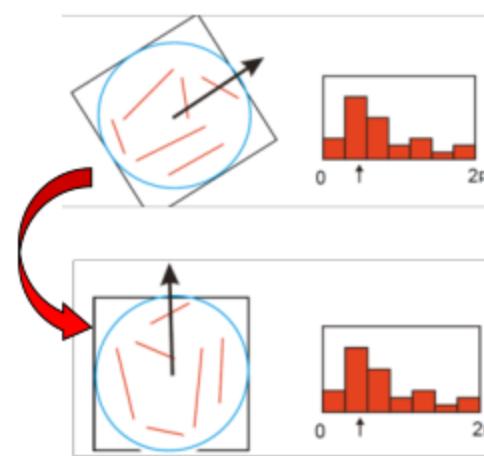
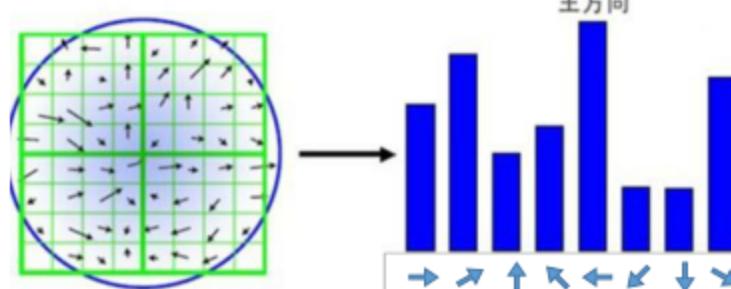
4) 指定方向参数

通过统计梯度直方图的方向，确定**主方向**，使算法具有旋转不变性

找到了特征点的尺度空间位置 $L(x, y, \sigma)$ ，以特征点为中心、以 $3 \times 1.5\sigma$ 为半径的区域影像的幅值 $m(x, y)$ 和幅角 $\theta(x, y)$ 计算方向：

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2}$$

$$\theta(x, y) = \alpha \cdot \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}$$



李明磊@nuaa

30

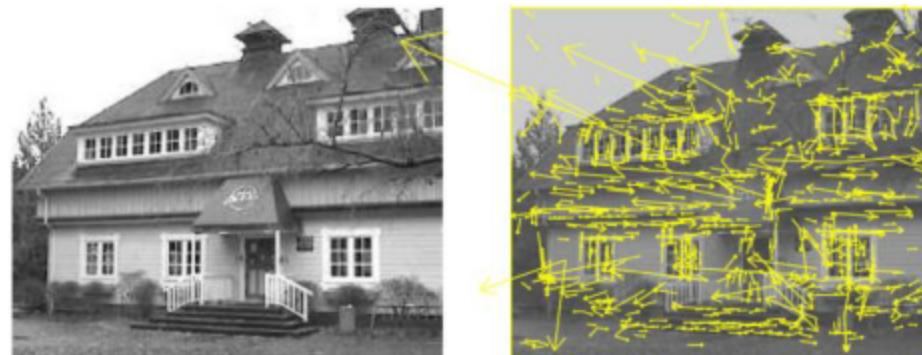


4) 指定方向参数



- 梯度方向从0到360度，每36度一个柱共10个柱，或每45度一个柱共8个柱
- 纵轴是梯度方向对应梯度幅值的累加，取峰值的方向就是主方向

得到特征点的主方向后，对于每一个关键点，都拥有
位置、尺度、方向三个信息 (x, y, σ, θ) 。



5) 关键点描述子

- 描述符：用一组向量将这个关键点描述出来，使其不随各种变化而改变，比如光照变化、视角变化等等。
- 这个描述子不但包括关键点，也包含关键点周围对其有贡献的像素点，并且描述符应该有较高的独特性，以便于提高特征点正确匹配的概率。
 - 将坐标轴旋转为特征点的主方向
 - 以主方向为中心取 16×16 的窗口，求取每个像素的梯度幅值与梯度方向，在每个 4×4 像素的小块上绘制 8 个方向的梯度直方图，每个特征点由 16 个种子点（小块）组成。这样一个关键点就可以产生 128 维的 SIFT 特征向量。

