

# 计算机视觉—特征匹配

申抒含

中国科学院自动化研究所  
模式识别国家重点实验室

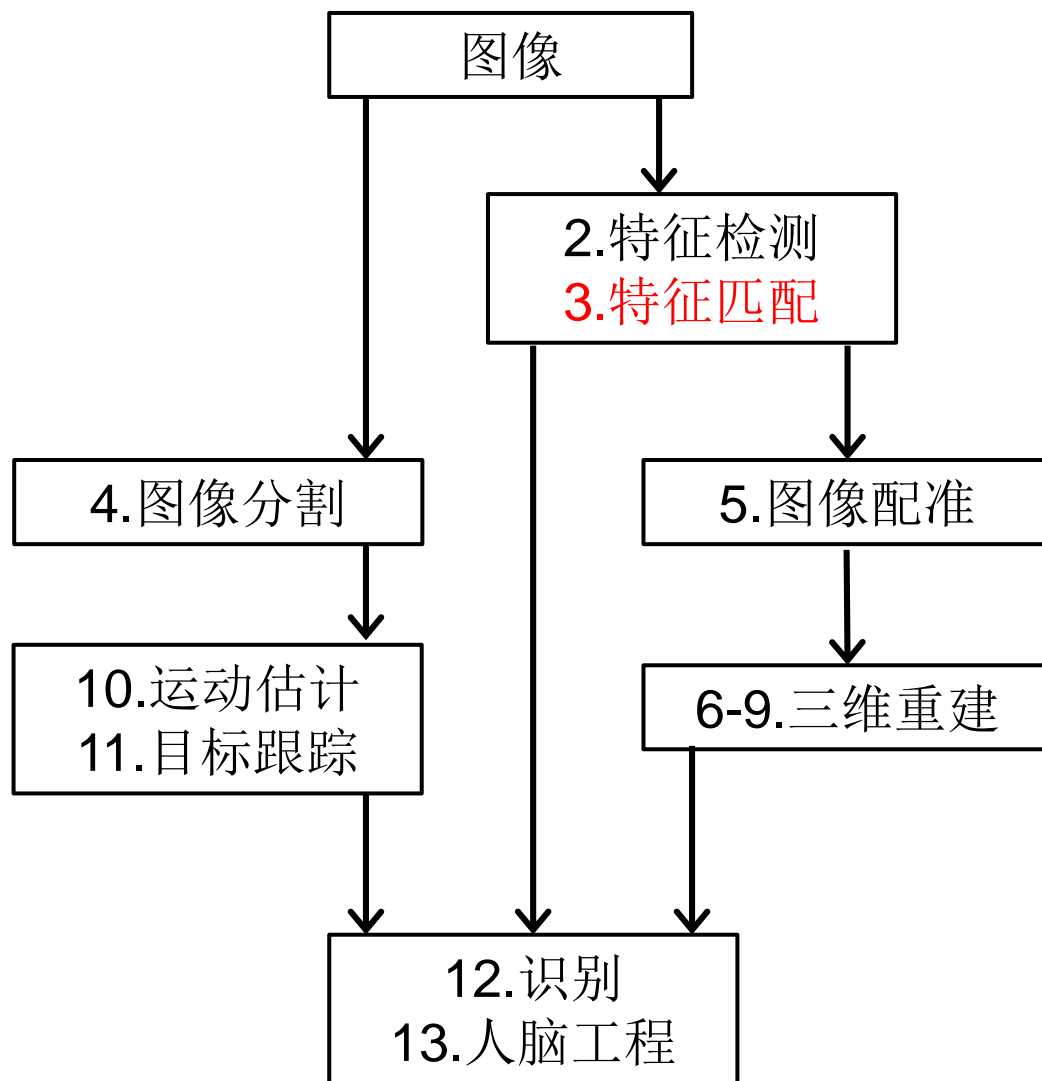


**Robot Vision Group**

National Laboratory of Pattern Recognition

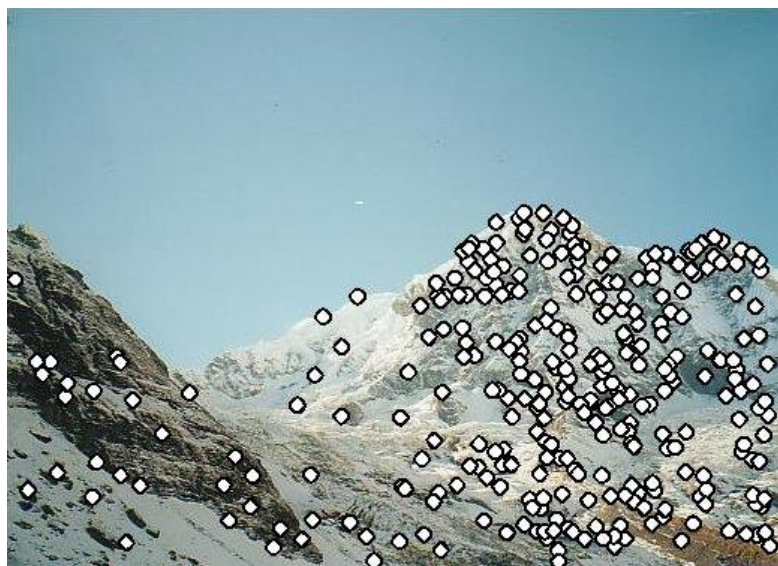
Institute of Automation, Chinese Academy of Sciences

# 计算机视觉课程结构图



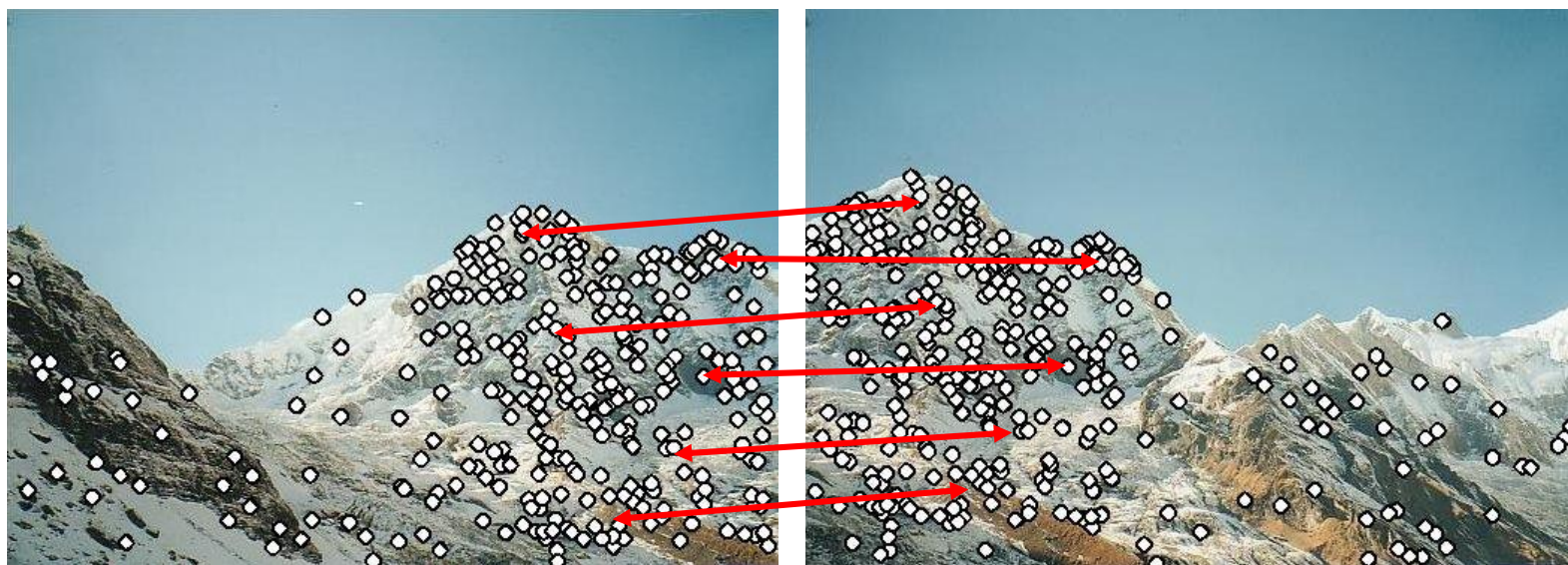
# 什么是特征匹配

上周课学习了如何检测好的特征



# 什么是特征匹配

有了特征以后如何确定它们在不同图像的对对应关系？





# 为什么要匹配特征

## 物体识别



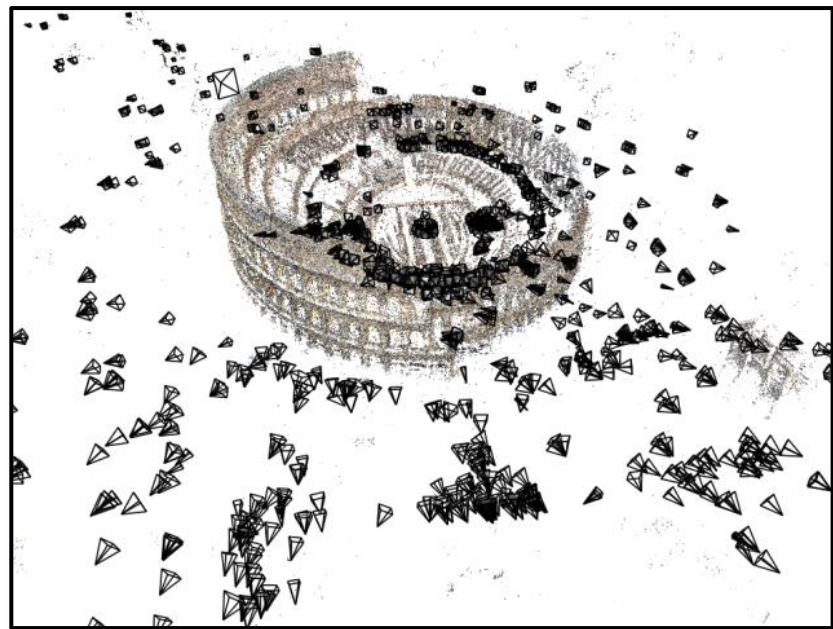
D. Lowe, 2004

# 为什么要匹配特征

## 三维重建



Internet Photos ("Colosseum")



Reconstructed 3D cameras and points

N. Snavely, 2006

# 为什么要匹配特征

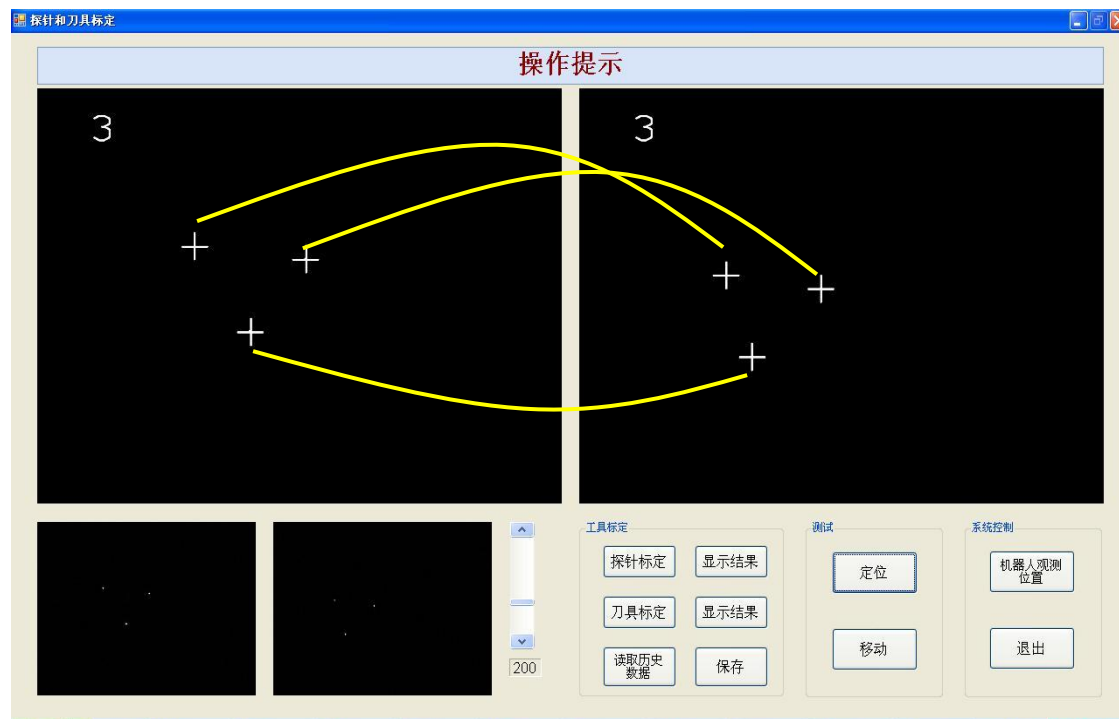
## Sony Aibo

- Recognize charging station
- Communicate with visual cards
- Teach object recognition



# 特征匹配

有时候匹配是一件很简单的事

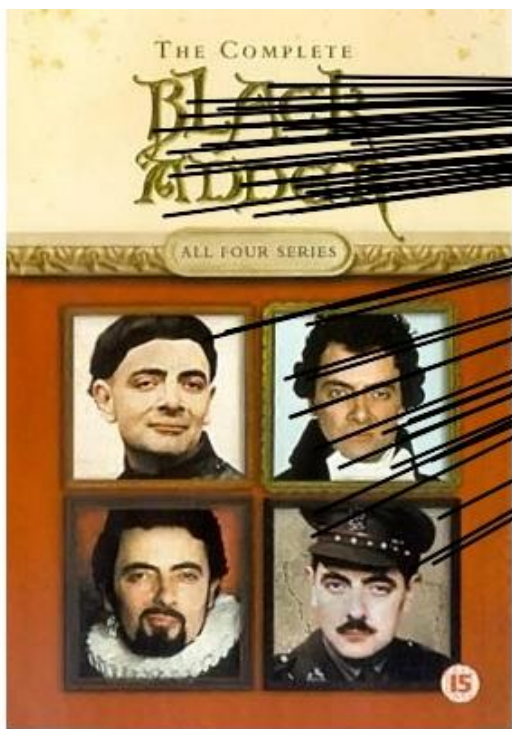


立体相机红外探针匹配



# 特征匹配

但大多数时候很复杂

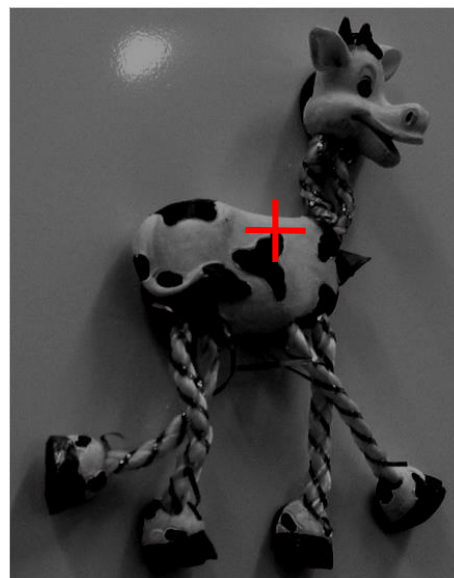
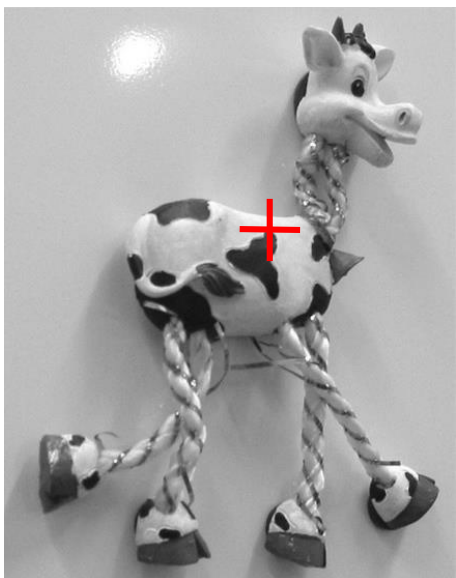


# 特征匹配—基本思路

- 特征准确匹配的要求是待匹配特征对平移、2D/3D旋转、光照、对比度、仿射变换等具有不变性。
- 这需要一个具有不变性的特征检测子：
  - Harris：对平移、2D旋转、光照具有不变性；
  - SIFT：对平移、2D旋转、光照、3D旋转（约60度）、尺度具有不变性；
- 还需要一个具有不变性的特征描述子：
  - 描述子用来记录特征点周围的区域信息；
  - 描述子要有不变性：图像变化时描述子不变；

# 特征描述—基本思路

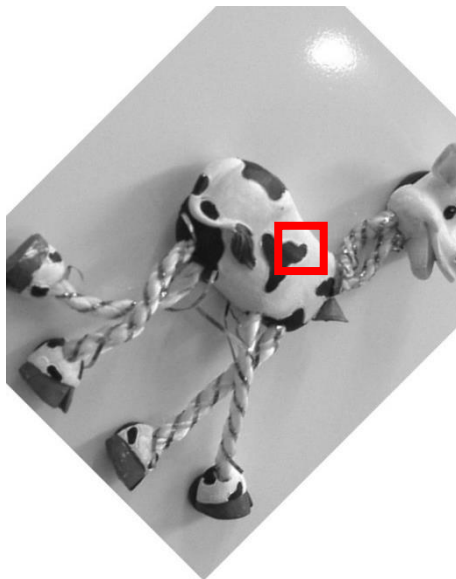
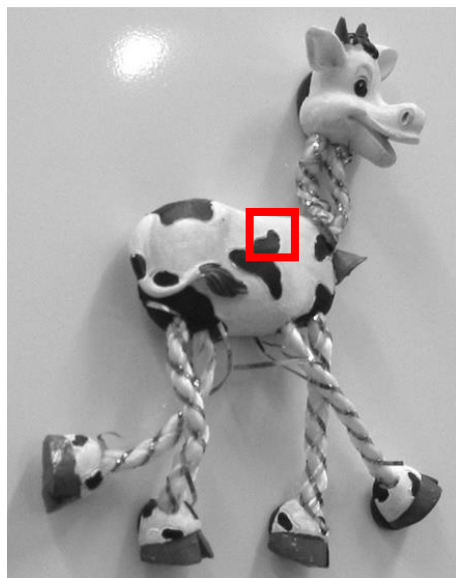
- 最简单的描述子：一个像素点，具有不变性吗？



- 对平移、旋转、仿射、尺度都有不变性；
- 但对光照没有任何不变性。

# 特征描述—基本思路

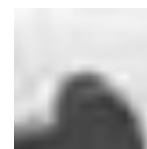
- 另一个简单的描述子：以特征点为中心的一个小方窗区域，具有不变性吗？



$\neq$



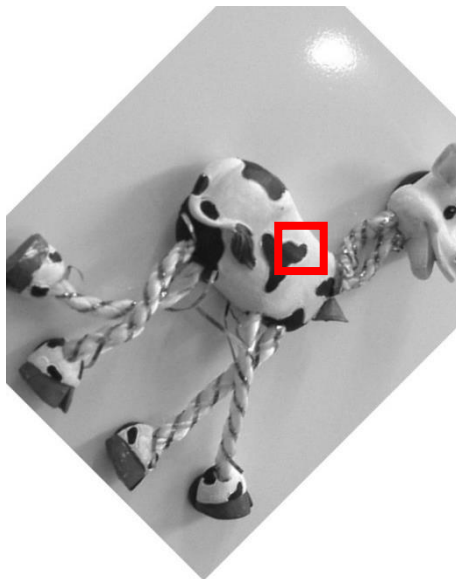
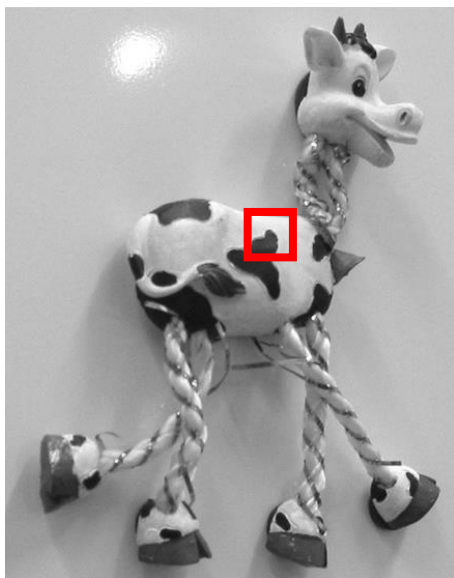
$\neq$





# 特征描述—基本思路

- 另一个简单的描述子：以特征点为中心的一个小方窗区域，具有不变性吗？

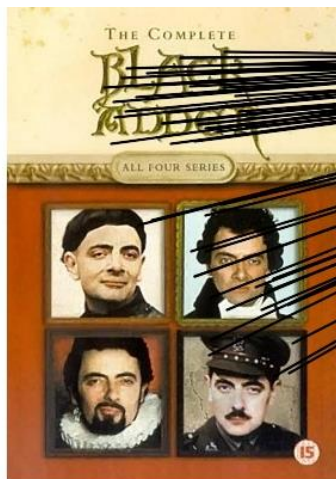


- 对平移、光照（灰度归一化）有不变性；
- 但对旋转、尺度没有不变性。

# 特征描述—基本思路

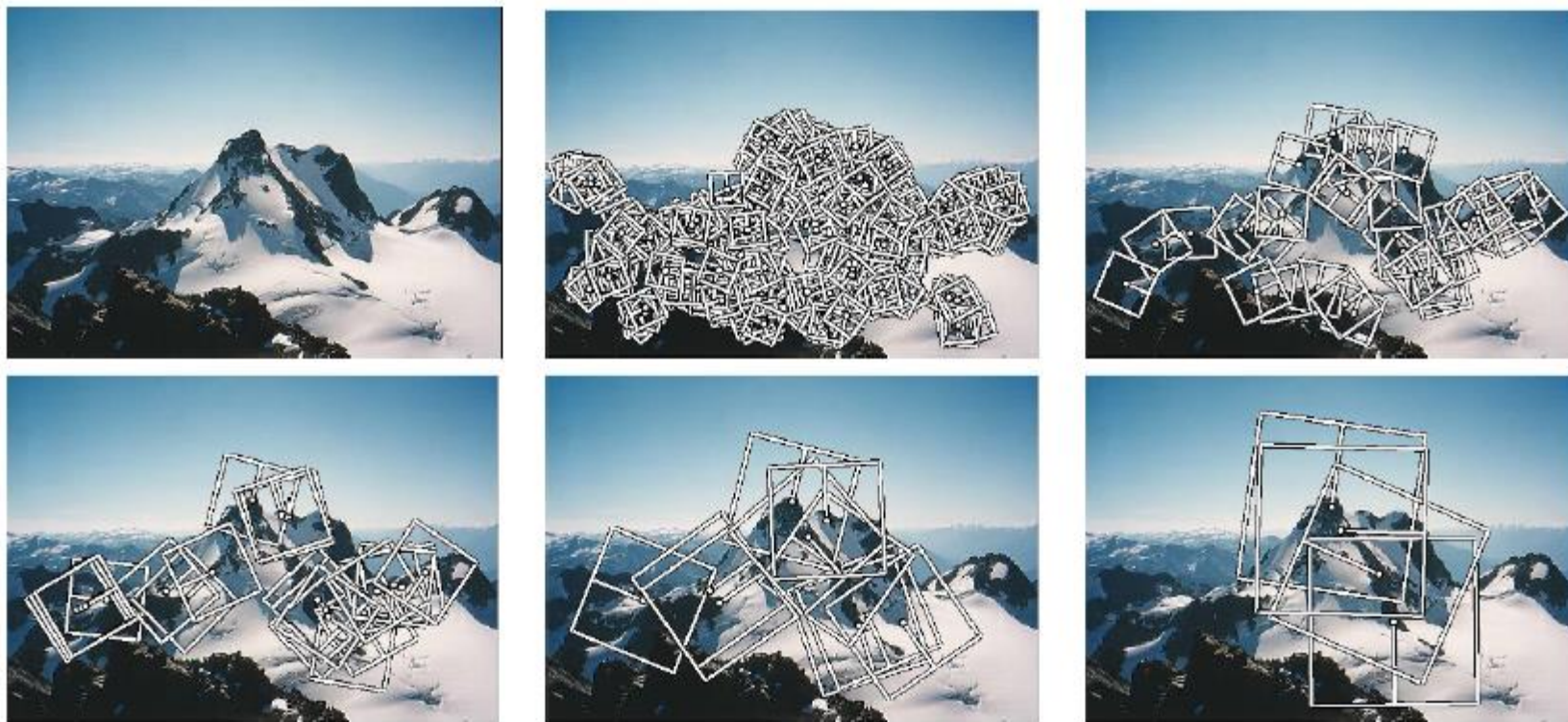
- 我们需要更复杂的描述子:

- MOPS
- SIFT
- PCA-SIFT
- GLOH



# 特征描述—MOPS

MOPS: Multiscale Oriented PatcheS descriptor



在5层图像金字塔上检测Harris角点（上周课内容）  
图中窗口表示了特征点方向和尺度

(Brown, Szeliski and Winder, 2005)

# 特征描述—MOPS

## 计算窗口主方向

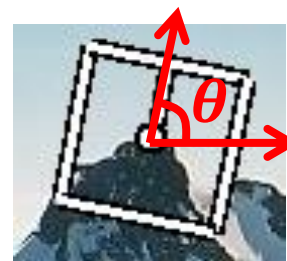
以特征点处高斯平滑后的梯度方向作为窗口的主方向 $\theta$ :

$$[\cos \theta, \sin \theta] = \mathbf{u}/|\mathbf{u}|$$

$$\mathbf{u}_l(x, y) = \nabla_{\sigma_0} P_l(x, y)$$

$P_l(x, y)$ 表示图像的第 $l$ 层高斯金字塔。

每一个特征点的参数:  $(x, y, l, \theta)$

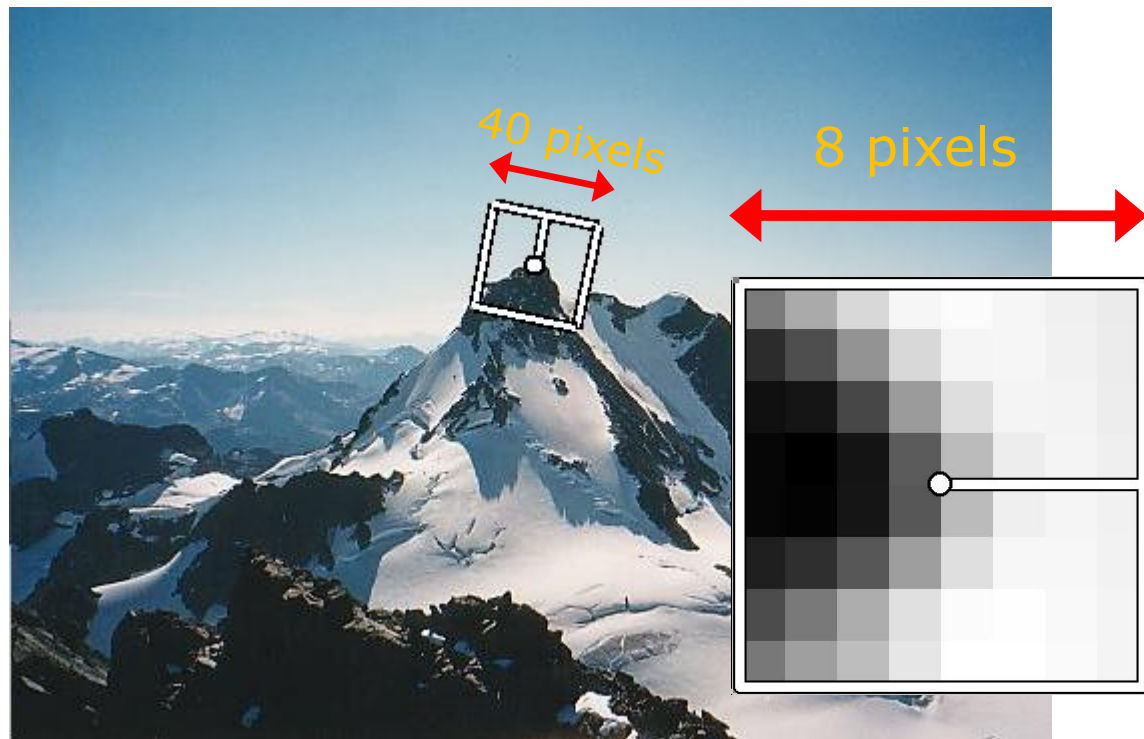




# 特征描述—MOPS

## 构造MOPS描述子

- 以特征点为中心根据主方向取一个 $40 \times 40$ 的窗口；
- 将窗口主方向旋转至水平；
- 窗口下采样至 $8 \times 8$ ；
- 窗口灰度归一化（每一个像素灰度减去均值、除以标准差）；
- 对窗口进行Haar小波变换，将小波系数写成一个64维向量构成描述子。



# 特征描述—MOPS

补充知识：图像（信号）变换

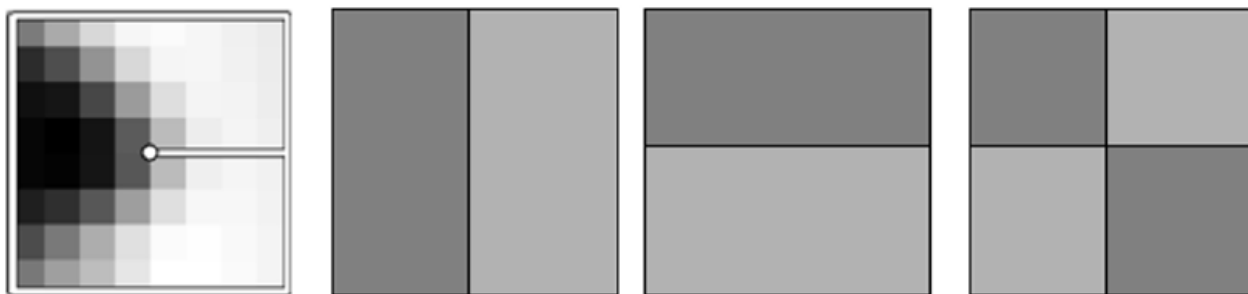
**傅立叶变换**：任意一条在实数域内有意义的曲线都可以分解为若干个正弦曲线的叠加。

**余弦变换**：类似于傅立叶变换，但是只使用实数。（JPEG）

**小波变换**：使用有限宽度基函数进行变换，这些基函数不仅在频率上而且在位置上是变化的。（JPEG2000）

# 特征描述—MOPS

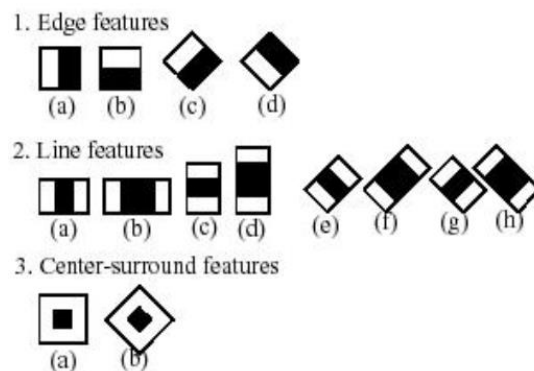
## Haar小波变换



MOPS中使用到的Haar特征中的前三个

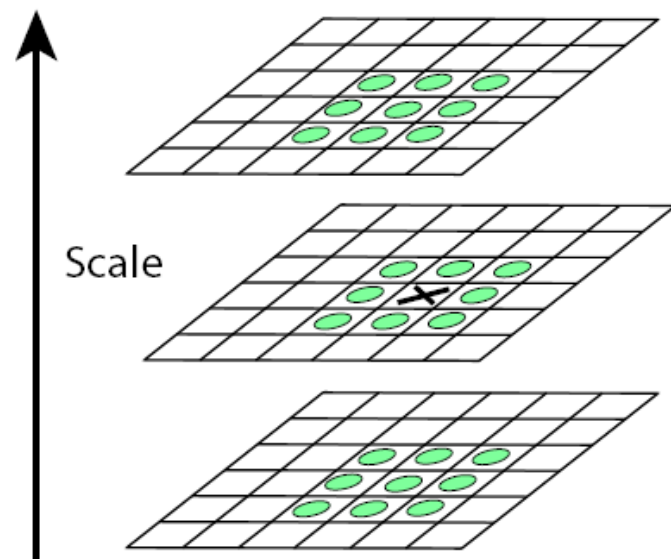
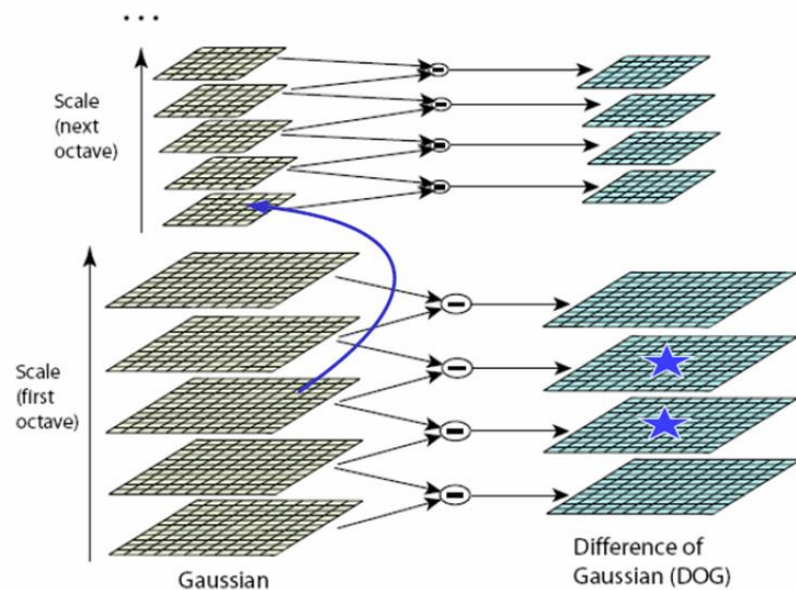
相当于对 $8 \times 8$ 窗口分别在深灰和浅灰区域求和后相减，其值作为这一Haar特征的系数。

在计算机视觉中，Haar特征最早由Viola和Jones在2001年用于人脸检测。



# 特征描述—SIFT

SIFT: Scale Invariant Feature Transform



回忆一下上周课的SIFT特征点检测：高斯尺度空间→高斯差分尺度空间→空间局部极值

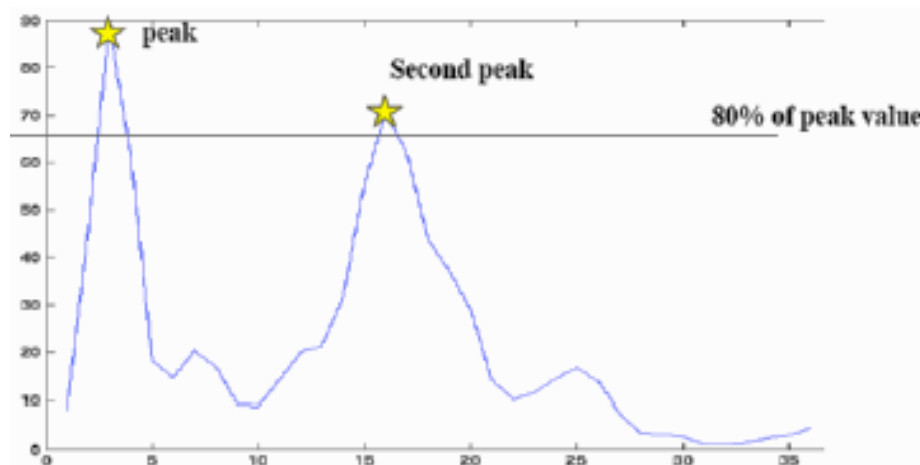
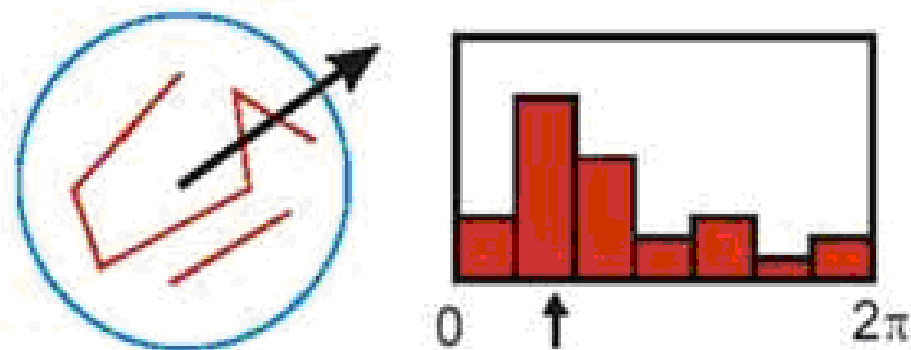
(Lowe, 2004)



# 特征描述—SIFT

## 计算窗口主方向

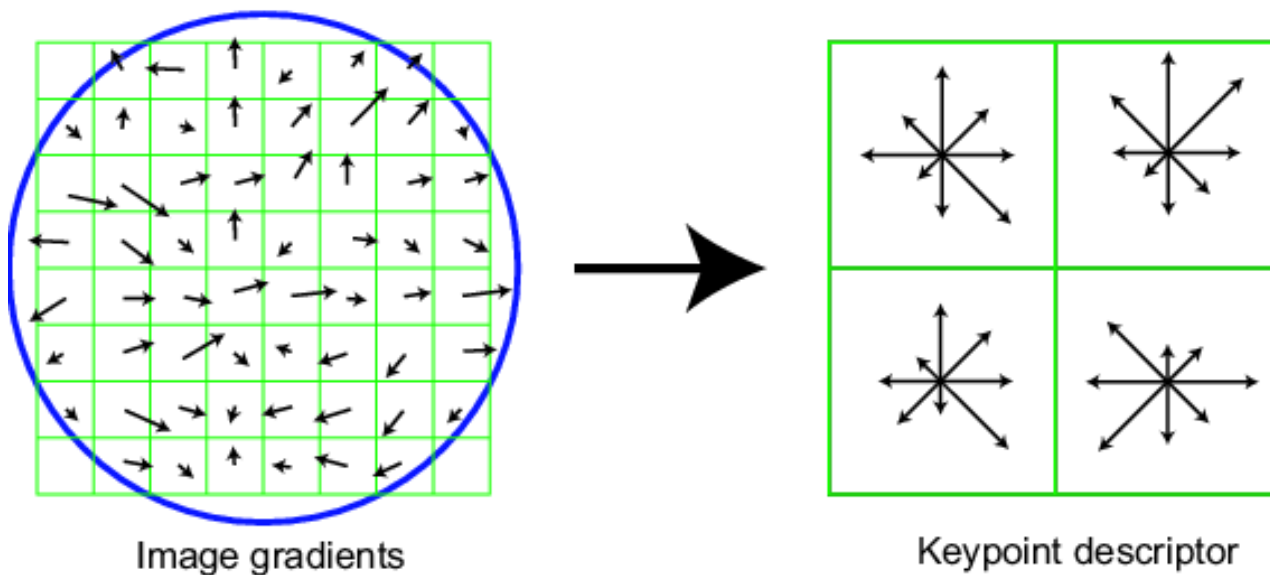
- 将圆周分为36份，每份10度；
- 计算特征点附近区域每个像素梯度方向和幅值；
- 根据每一像素梯度幅值和距中心的高斯距离绘制梯度直方图；
- 将直方图中最大值方向作为特征点主方向；
- 如果直方图其他值达到最大值的80%，则通过抛物线插值确定主方向。



# 特征描述—SIFT

## 梯度直方图

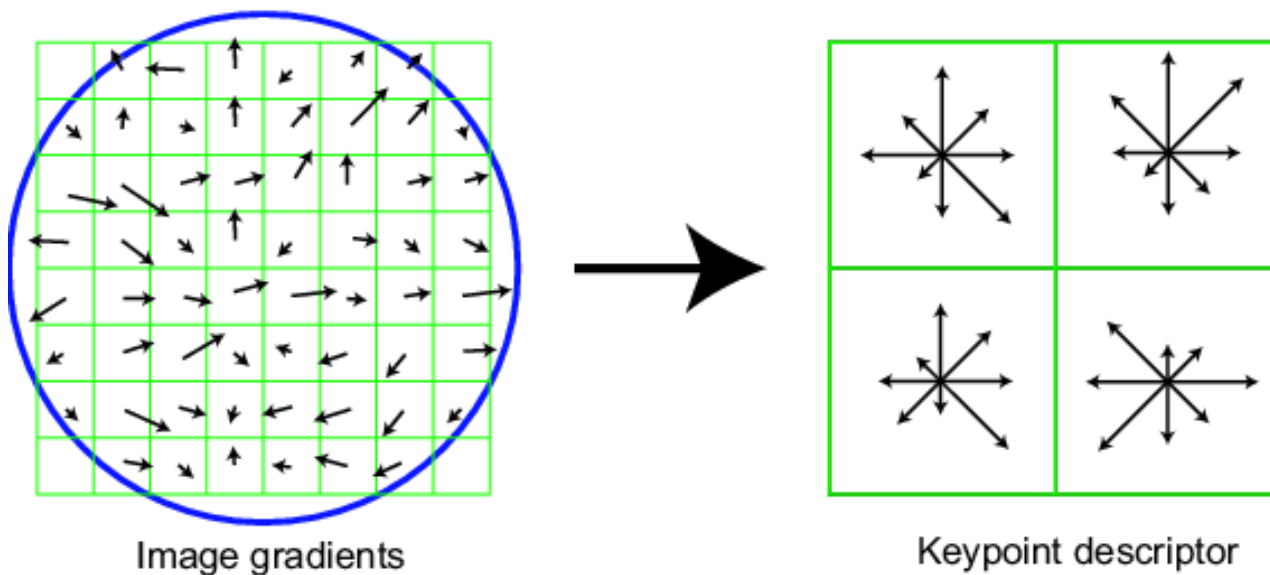
- 计算特征点窗口内每个图像点的梯度方向和梯度幅值；
- 将梯度幅值图进行高斯平滑（避免窗口微移导致描述子剧变）；
- 将窗口划分成子区域（图示为 $2 \times 2$ ），统计每个子区域内的梯度方向直方图（图示中箭头为方向，长度为梯度幅值加权）；
- 梯度幅值的权重为图像点到子区域两个边缘的距离乘积。



# 特征描述—SIFT

## 构造SIFT描述子

1. 以特征点为中心根据主方向取一个 $16 \times 16$ 的窗口（图示为 $8 \times 8$ ）；
2. 对梯度幅值图进行高斯平滑；
3. 将窗口划分成 $4 \times 4$ 个子区域（图示为 $2 \times 2$ ）；
4. 计算每个子区域梯度幅值（加权）直方图；
5. 得到一个 $4 \times 4 \times 8 = 128$ 维向量构成描述子。



# 特征描述—PCA-SIFT

## 构造PCA-SIFT描述子

- 以特征点为中心根据主方向取一个 $41 \times 41$ 的窗口；
- 将窗口主方向旋转至水平；
- 计算窗口的水平和垂直梯度，得到两个 $39 \times 39$ 的梯度图；
- 将梯度图写成一个 $39 \times 39 \times 2 = 3042$ 维的向量；
- 使用PCA将得到的3042维向量降到36维构成描述子。



(A1) SIFT:  
4/10 correct



(A2) PCA-SIFT ( $n=20$ ):  
9/10 correct



(B1) SIFT:  
6/10 correct



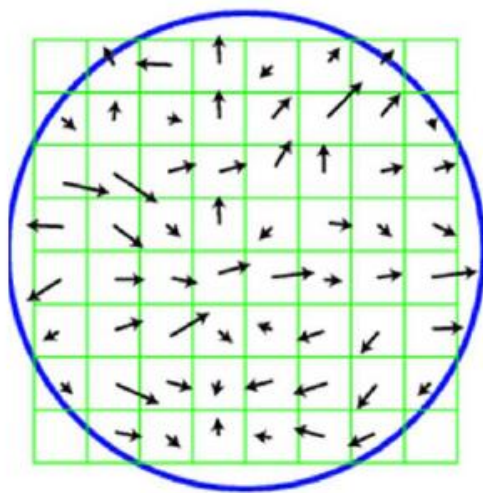
(B2) PCA-SIFT ( $n=20$ ):  
10/10 correct



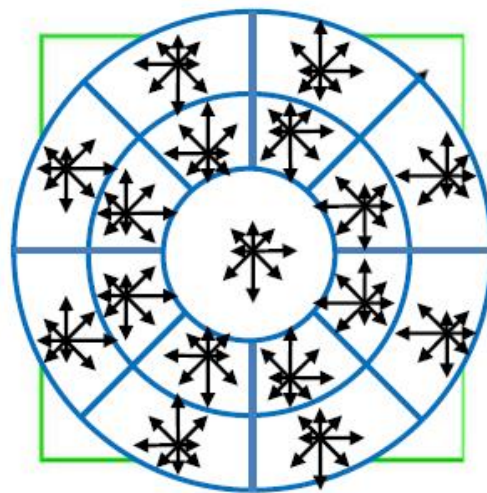
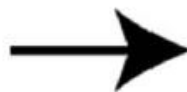
# 特征描述—GLOH

## 构造GLOH描述子

- 使用对数极坐标结构替代SIFT的4象限；
- 空间上取半径6、11和15，角度上分八个区间（除中间区域）；
- 梯度方向分为16份，构成一个272（ $17 \times 16$ ）维梯度直方图；
- 使用PCA将得到的272维向量降到128维构成描述子；
- 这个PCA的协方差矩阵 通过在47,000个图像path上训练得到。



(a) image gradients



(b) keypoint descriptor

(Mikolajczyk and Schmid, 2005)

# 特征描述—GLOH

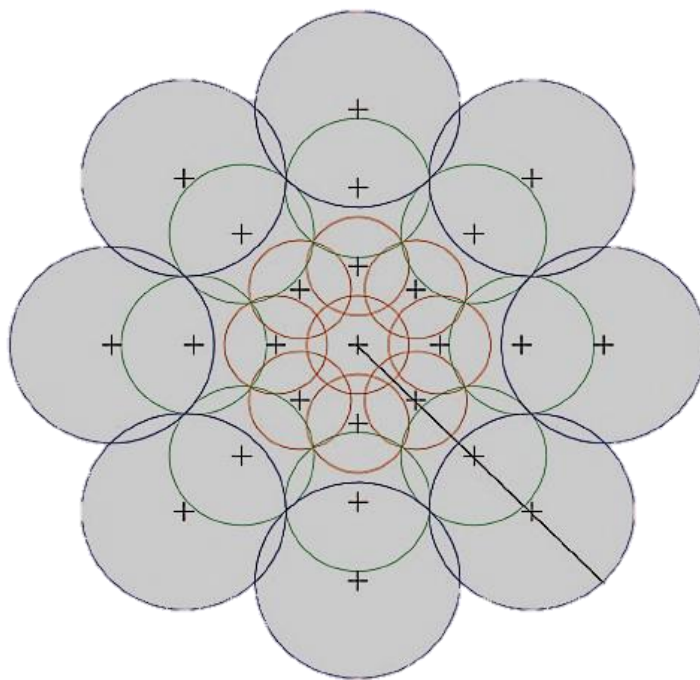


在Mikolajczyk and Schmid. [A performance evaluation of local descriptors](#). IEEE T-PAMI, 2005的评测中，GLOH表现最好，SIFT紧随其后。

# 特征描述—DAISY

DAISY: the descriptor looks like a flower

- DAISY计算速度比SIFT快很多;
- DAISY在每一个图像点上计算描述子（用于稠密匹配），SIFT主要用于特征点（角点）的描述（当然也可以描述任意图像点）。

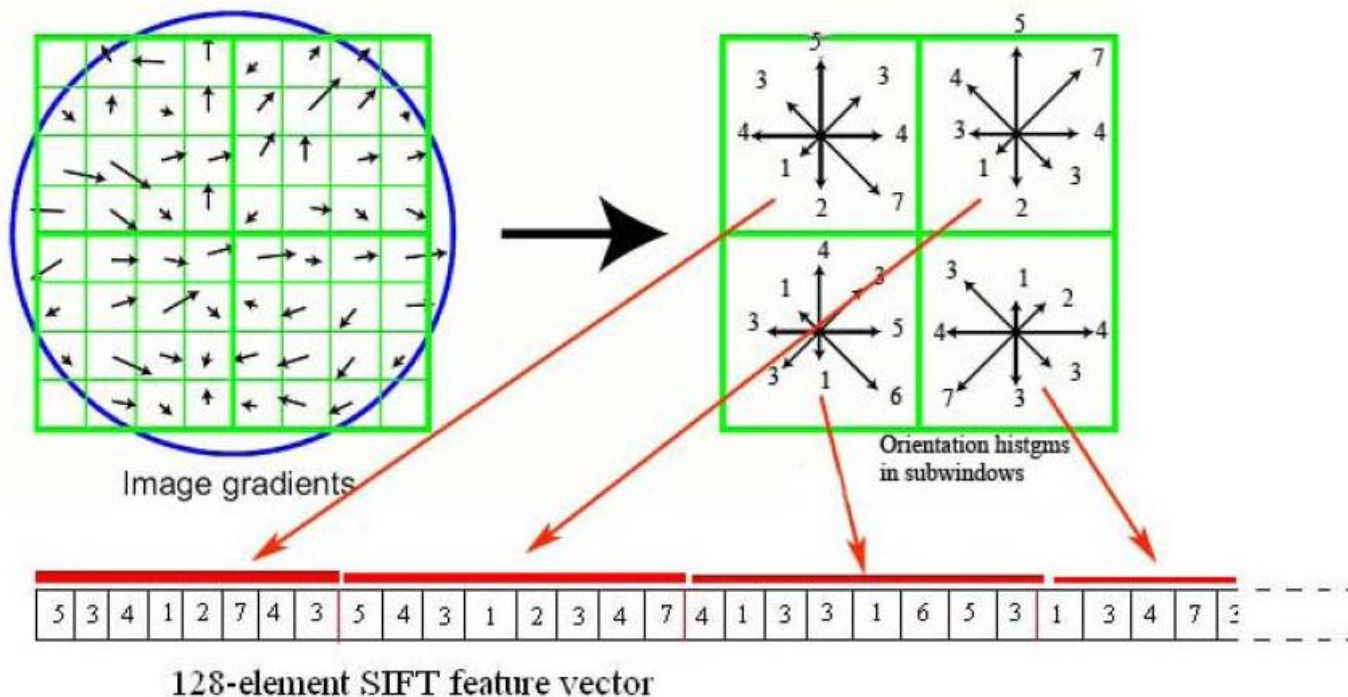


(Tola, Lepetit, and Fua, 2010)

# 特征描述—DAISY

## DAISY的基本思想

SIFT描述子中每一个值对应一个子区域内某一方向的梯度幅值加权和，权值大小与到特征点的距离以及到子区域两个边缘的距离有关。



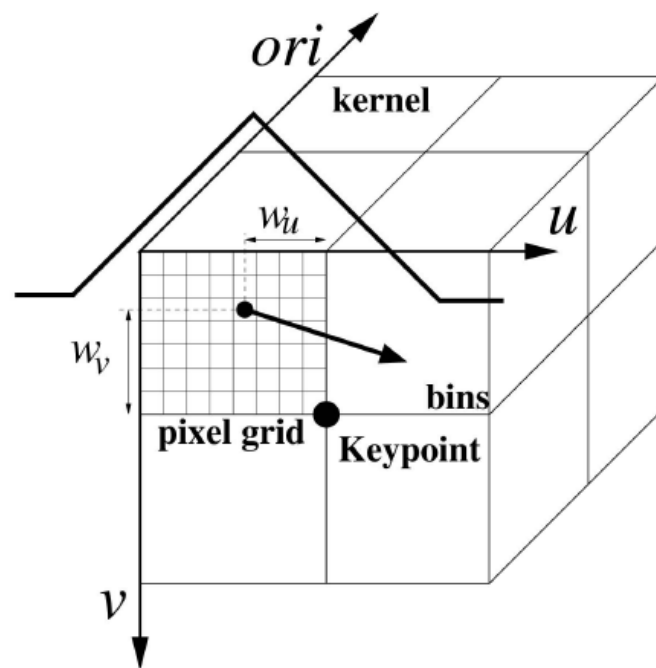


# 特征描述—DAISY

## DAISY的基本思想

可以将SIFT使用的梯度图看作由图像点位置 ( $u, v$ ) 和朝向 ( $ori$ ) 构成的三维空间图，其中每一个空间点的值与梯度幅值、到特征点的距离、以及到子区域两个边缘的距离有关。

**DAISY基本思想：**提前计算好一系列梯度方向图，在构造描述子时可以反复使用，从而提高效率。

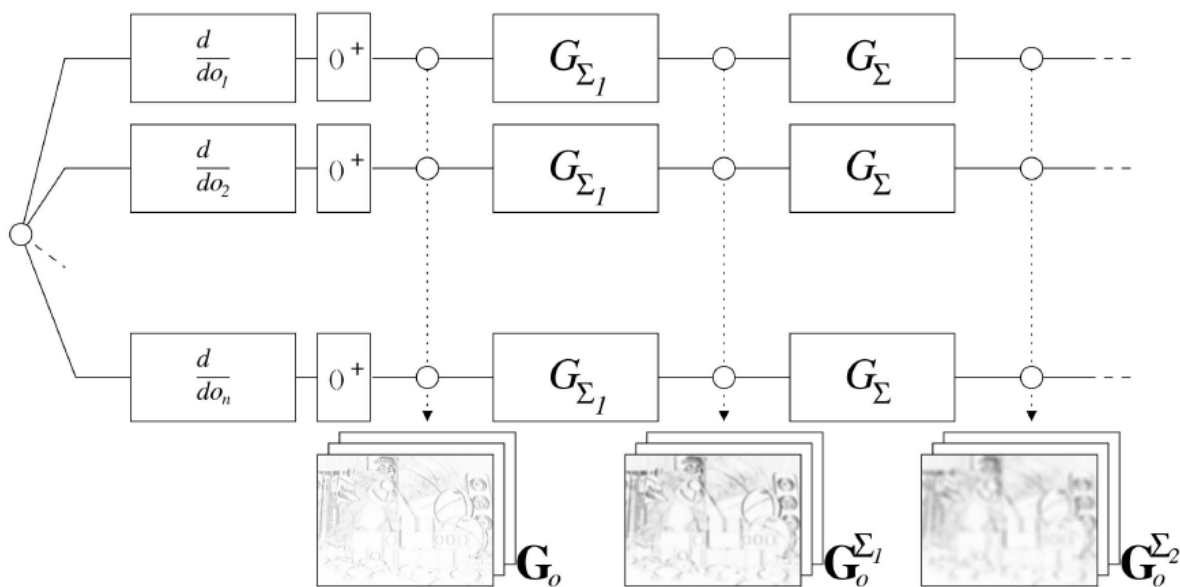
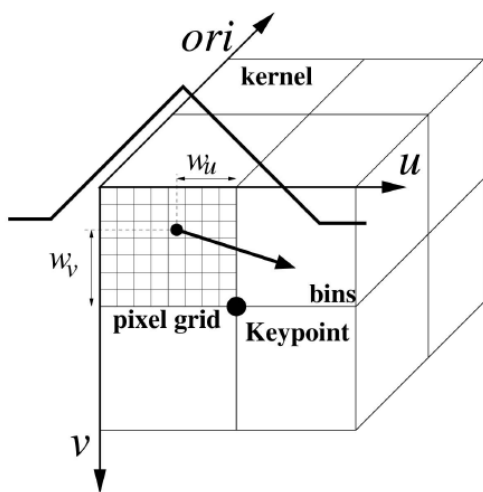




# 特征描述—DAISY

## DAISY的基本思想

计算 $H$ 个方向上的梯度高斯幅值图 $G_o^\Sigma = G_\Sigma * (\frac{\partial I}{\partial o})^+$ ， $o$ 表示梯度方向  
其中 $G_\Sigma$ 为高斯卷积核， $\Sigma$ 控制区域大小， $(a)^+ = \max(a, 0)$

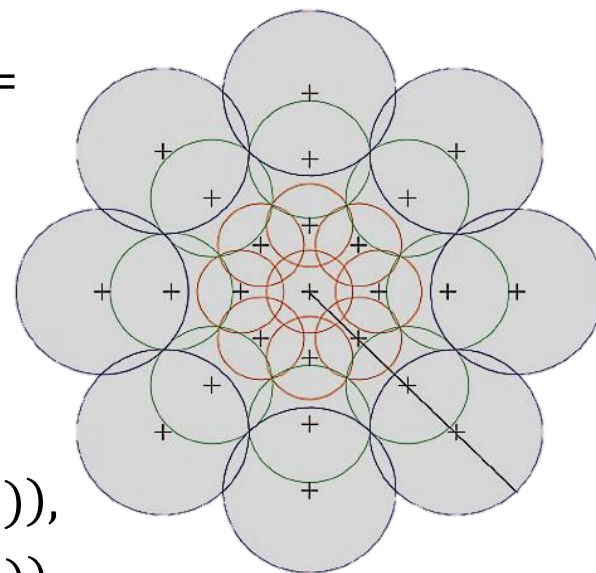


# 特征描述—DAISY

## 构造DAISY描述子

1. 在每一个图像点 $(u,v)$ 计算描述向量 $\mathbf{h}_{\Sigma}(u, v) = [G_1^{\Sigma}, G_2^{\Sigma}, \dots, G_H^{\Sigma}]^T$ ;
2. 将 $\mathbf{h}_{\Sigma}(u, v)$ 归一化为 $\overline{\mathbf{h}}_{\Sigma}(u, v)$ ;
3. 构成DAISY描述子:

$$D(u, v) = [\overline{\mathbf{h}}_{\Sigma_1}^T(u, v), \overline{\mathbf{h}}_{\Sigma_1}^T(l_1(u, v, R_1)), \dots, \overline{\mathbf{h}}_{\Sigma_1}^T(l_T(u, v, R_1)), \overline{\mathbf{h}}_{\Sigma_2}^T(l_1(u, v, R_2)), \dots, \overline{\mathbf{h}}_{\Sigma_2}^T(l_T(u, v, R_2)), \dots, \overline{\mathbf{h}}_{\Sigma_Q}^T(l_1(u, v, R_Q)), \dots, \overline{\mathbf{h}}_{\Sigma_Q}^T(l_T(u, v, R_Q))]^T$$



参数设置:  $Q=3, T=8, H=8$ , 描述子维数:  $(T \times Q + 1) \times H$

# 特征描述—DAISY

## DAISY的计算效率

- 高斯卷积计算效率高
- 大尺度高斯卷积可以通过小尺度卷积加快计算

$$\mathbf{G}_o^{\Sigma_2} = G_{\Sigma_2} * \left( \frac{\partial \mathbf{I}}{\partial o} \right)^+ = G_{\Sigma} * G_{\Sigma_1} * \left( \frac{\partial \mathbf{I}}{\partial o} \right)^+ = G_{\Sigma} * \mathbf{G}_o^{\Sigma_1},$$

$$\text{with } \Sigma = \sqrt{\Sigma_2^2 - \Sigma_1^2}.$$

Image Size	DAISY	SIFT
800x600	3.8	252
1024x768	6.5	432
1280x960	9.8	651

# 特征描述—DAISY

## DAISY的应用



NCC

SIFT

DAISY

# 特征描述—DAISY

## DAISY的应用

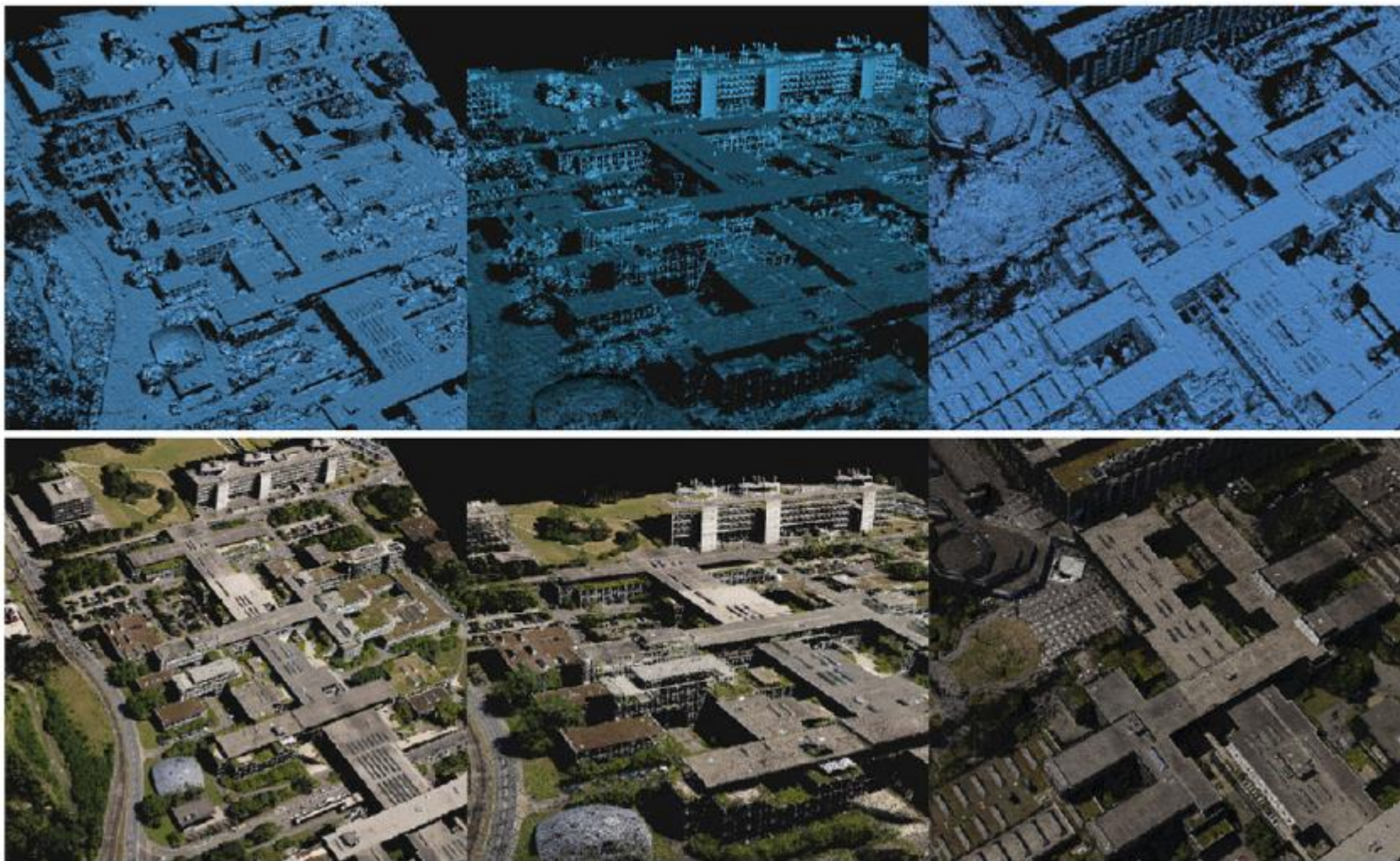


Tola, Strecha, and Fua, 2013



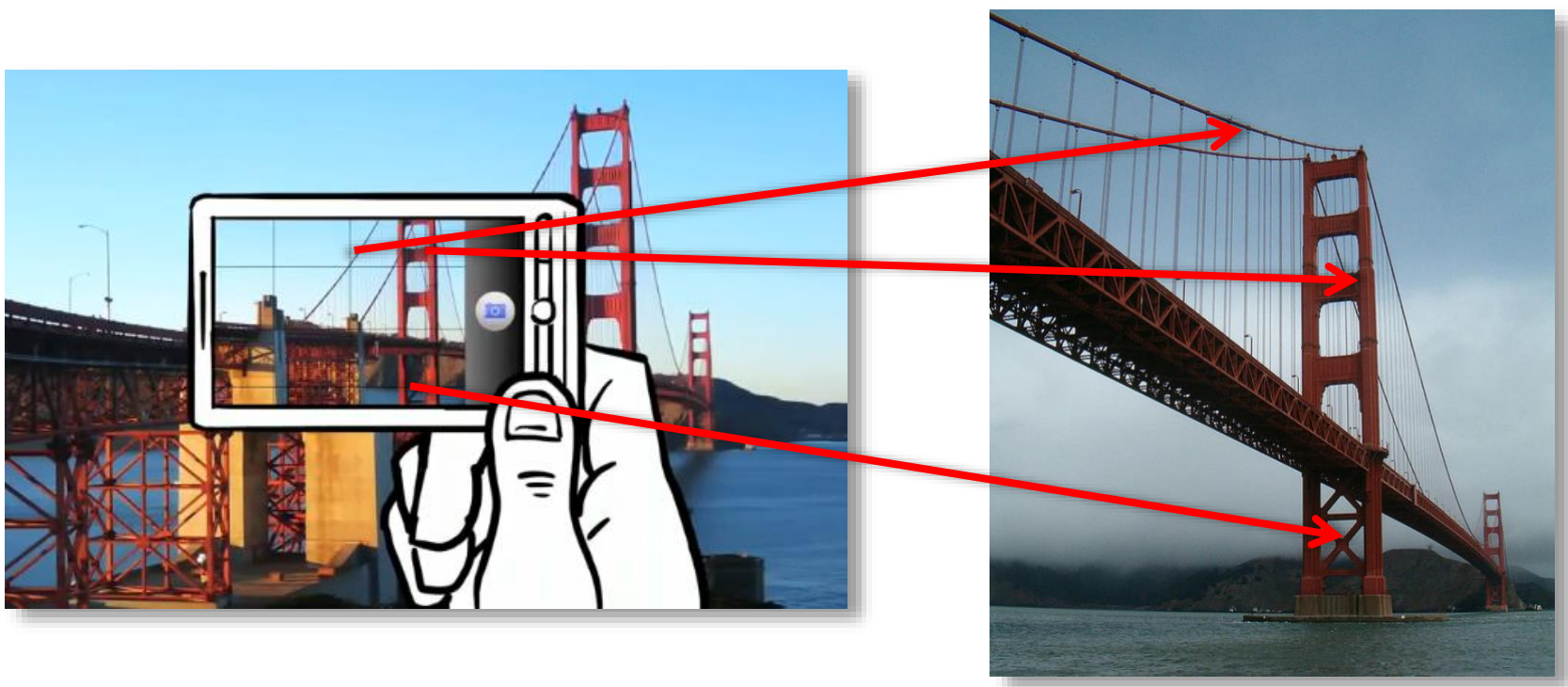
# 特征描述—DAISY

## DAISY的应用



# 特征匹配

对于图像 $I_1$ 中的一个特征点，图像 $I_2$ 中那个特征点与它匹配？



# 特征匹配

对于图像 $I_1$ 中的一个特征点，图像 $I_2$ 中那个特征点与它匹配？

特征匹配的基本假设：可以在特征空间通过特征描述子之间的欧氏距离判断匹配程度，距离越小匹配度越高。

特征匹配基本思路：

- 1) 在特征空间定义特征描述子之间的某种距离度量函数；
- 2) 找出 $I_2$ 中与 $I_1$ 特征点距离最小的作为匹配点。

常用距离度量函数：**SSD** (Sum of Square Differences )  
两个描述子对应值差的平方和

# 特征匹配

用于定量衡量匹配质量的指标：

TP: True Positive, 被判定为匹配, 事实上也是匹配;

FN: False Negative, 被判定为不匹配, 但事实上匹配;

FP: False Positive, 被判定为匹配, 但事实上不匹配;

TN: True Negative, 被判定为不匹配, 事实上也是不匹配。

$TPR = TP / (TP + FN)$ , 真阳性率;

$FPR = FP / (FP + TN)$ , 假阳性率;

$PPV = TP / (TP + FP)$ , 阳性预测值;

$ACC = (TP + TN) / (TP + FN + FP + TN)$ , 预测精度。

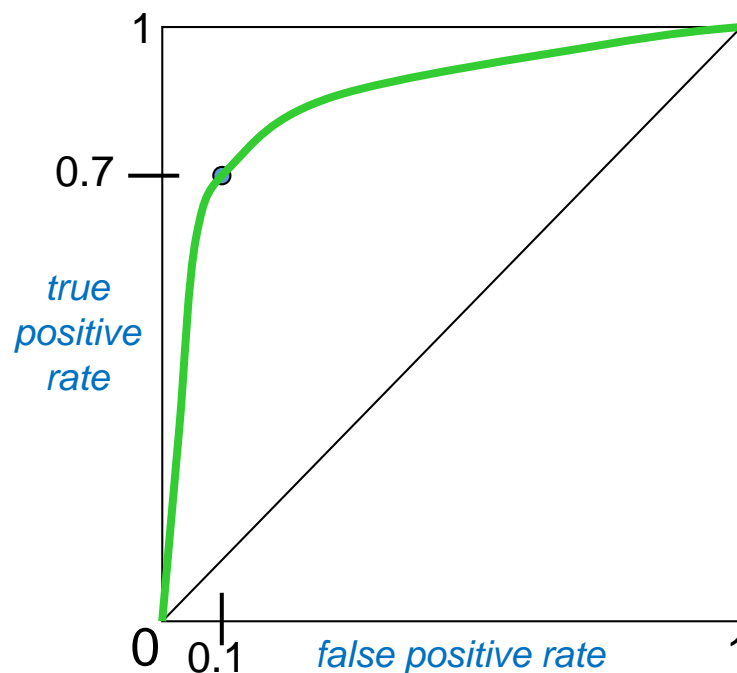


# 特征匹配

用于定量衡量匹配质量的指标—ROC曲线

ROC: Receiver Operating Characteristic curve

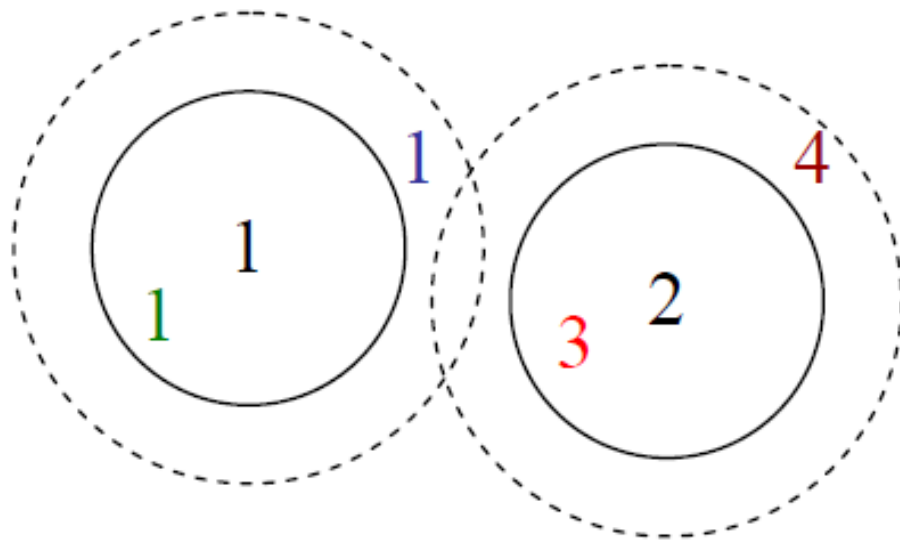
- 阈值最高时，ROC曲线在左下角；
- 阈值最低时，ROC曲线在右上角；
- ROC曲线越靠近左上角（ROC曲线下面积越大）性能越好。





# 特征匹配—单阈值法

单阈值法：给定一个阈值，返回所有匹配距离小于阈值的特征点作为匹配点。

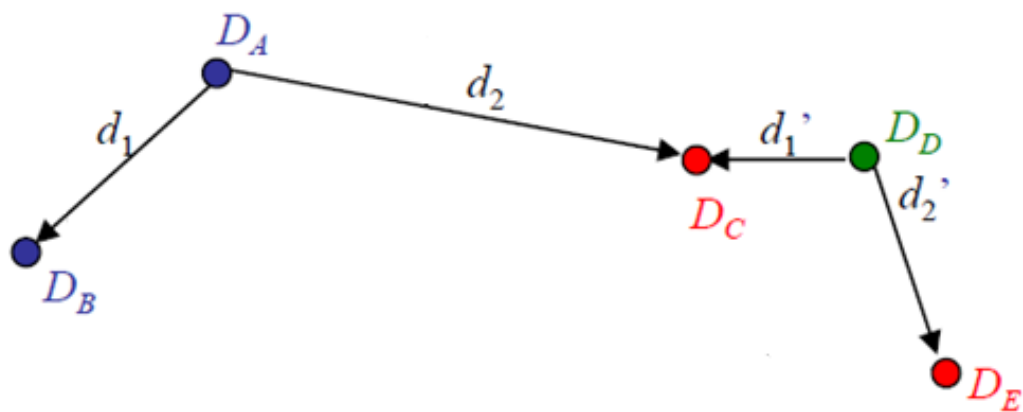


左图中数字表示类别

阈值高时（实线），容易产生False Negative;  
阈值低时（虚线），容易产生False Positive。

# 特征匹配—最近邻法

最近邻法：给定一个较低阈值，返回阈值内匹配距离最小的特征点作为匹配点。

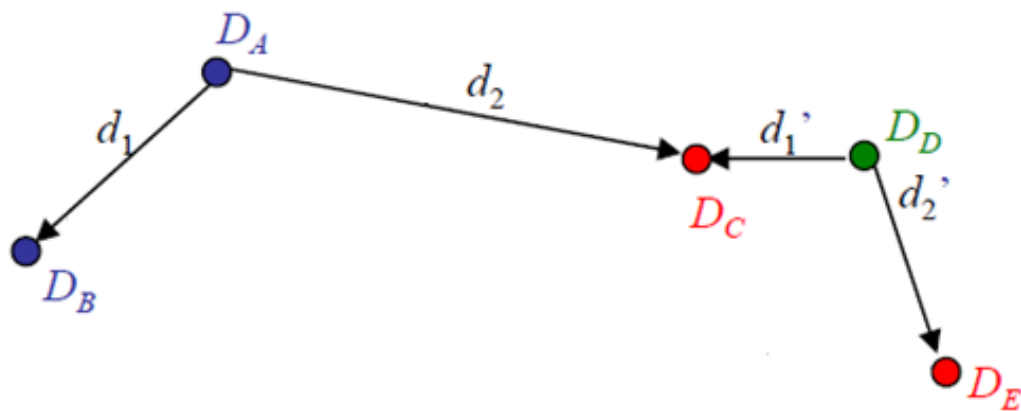


左图中颜色表示类别

使用最近邻法， $D_A$ 与 $D_B$ 正确匹配，但 $D_D$ 与 $D_C$ 错误匹配。

# 特征匹配—最近邻比值法

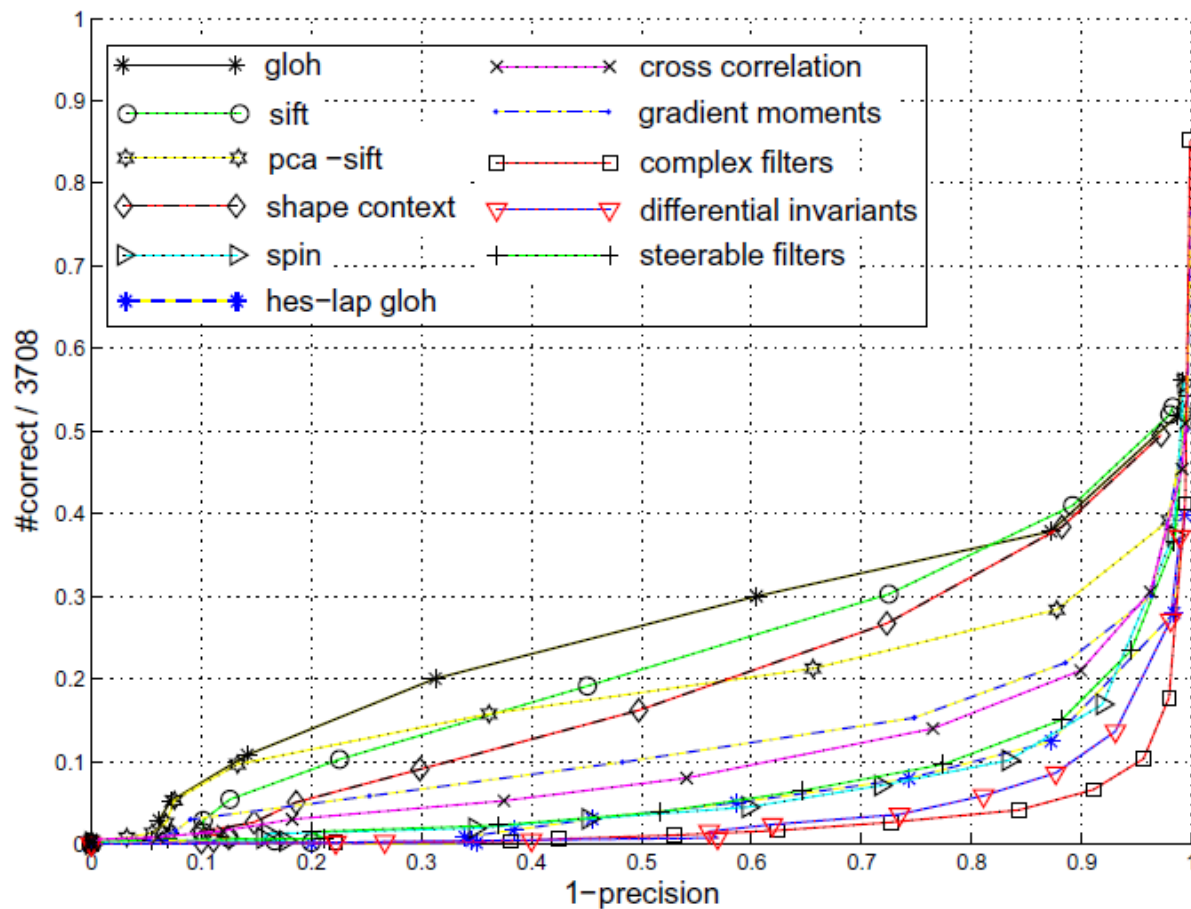
最近邻比值法：当距离最近的特征点和距离次近特征点间距离比值小于一定阈值时，将距离最近的特征点作为匹配点，否则无匹配点。



左图中颜色表示类别

使用最近邻比值法， $d_1/d_2$ 较小， $D_A$ 与 $D_B$ 正确匹配； $d_1'/d_2'$ 较大，判定 $D_C$ 不是 $D_D$ 的匹配点。

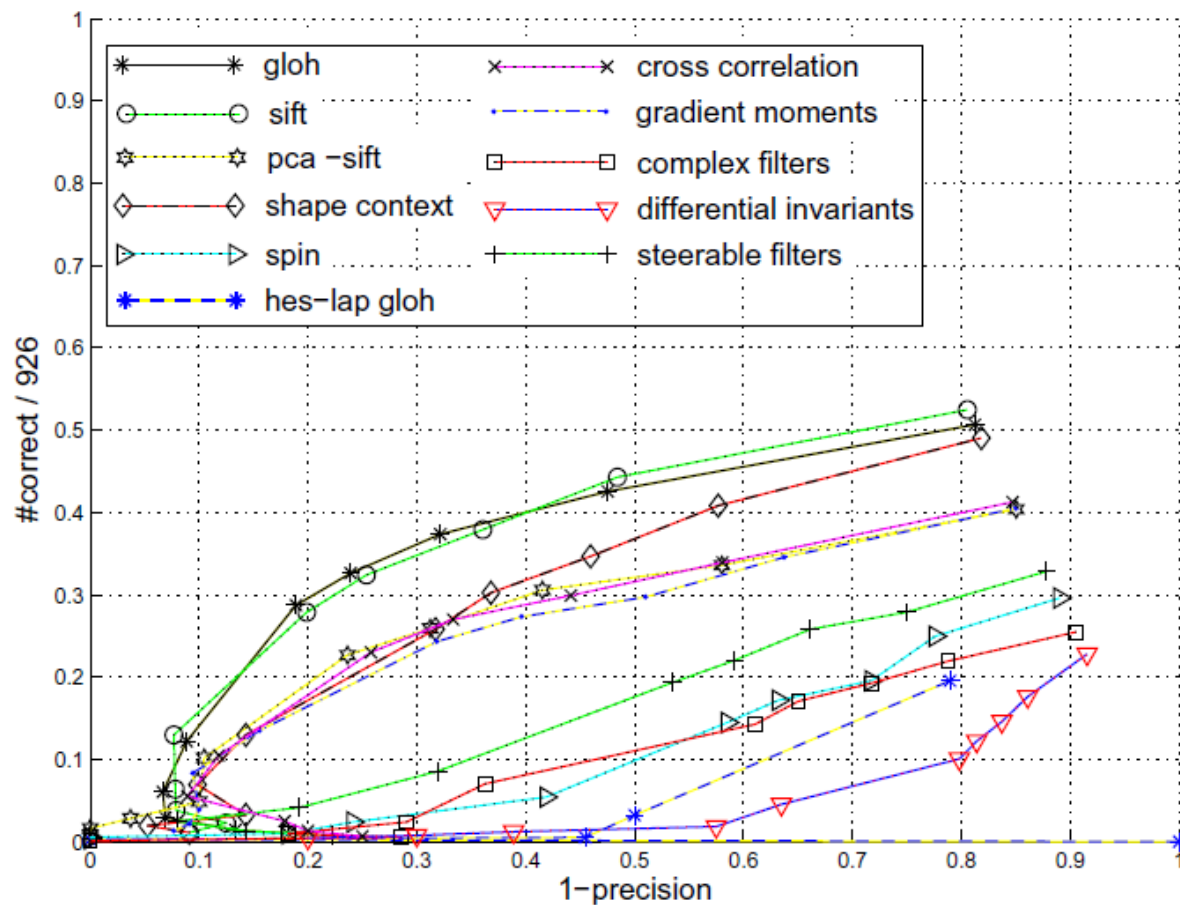
# 特征匹配—不同阈值方法ROC比较



单阈值法

(Mikolajczyk and Schmid, 2005)

# 特征匹配—不同阈值方法ROC比较

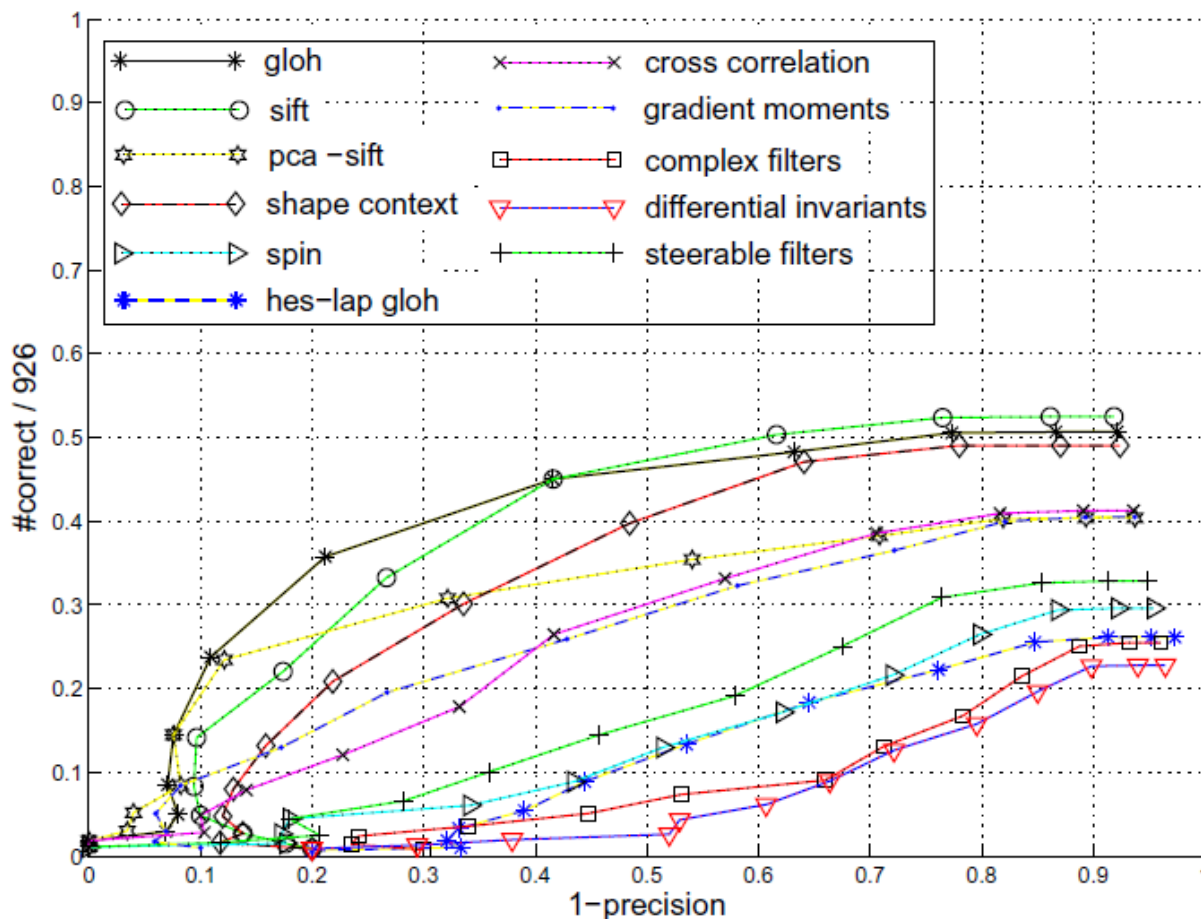


最近邻法

(Mikolajczyk and Schmid, 2005)



# 特征匹配—不同阈值方法ROC比较

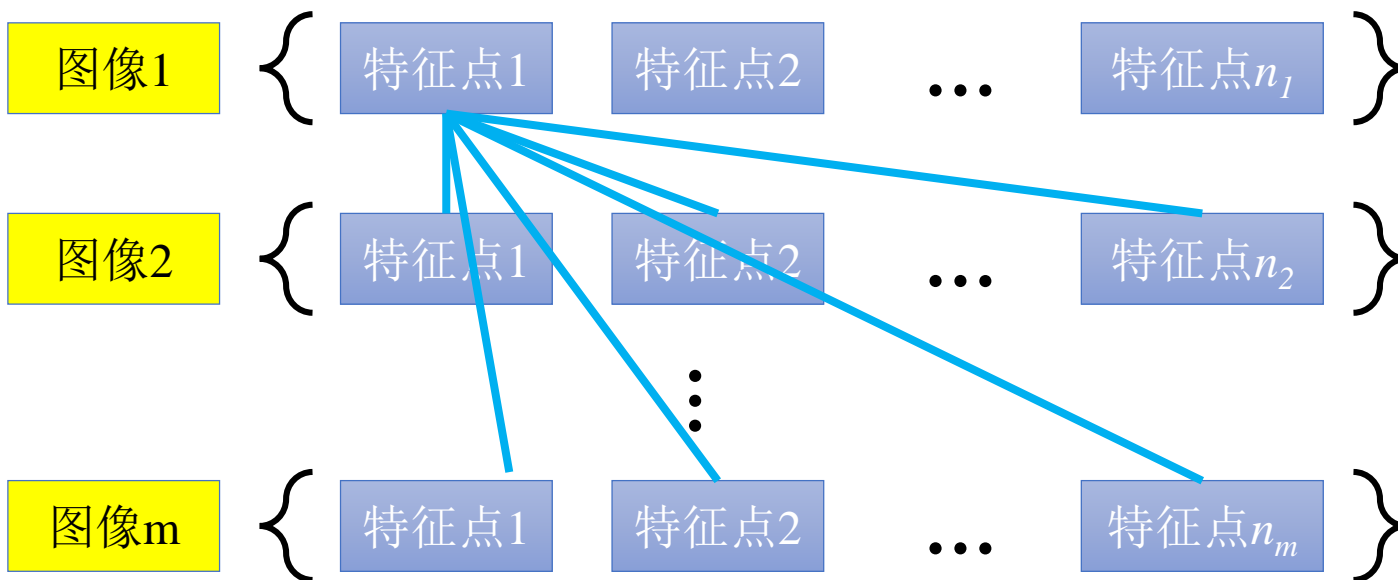


最近邻比值法

(Mikolajczyk and Schmid, 2005)

# 特征匹配

定义了匹配距离度量和匹配点判断方法，下一步是如何进行有效的匹配。



最简单的方法：穷举法

# 特征匹配—Hashing

哈希表（Hash table，也叫散列表），是根据关键字（Key value）而直接访问在内存存储位置的数据结构。Hashing算法通过一个函数将高维特征映射为哈希表中的一个键值，来加快查找速度，这个映射函数称做哈希函数。

特征匹配中常用哈希方法：局部敏感哈希（Locality Sensitive Hashing, LSH）

LSH的基本思想：将特征空间中的两个相邻特征点通过哈希函数映射后，这两个数据点在哈希表中仍然相邻的概率很大，而不相邻的特征点被映射到同一个键值的概率很小。

# 特征匹配—Hashing

LSH哈希函数应满足的条件：

- 1) 如果 $d(x,y) \leq d_1$ ，则 $h(x) = h(y)$ 的概率至少为 $p_1$ ；
- 2) 如果 $d(x,y) \geq d_2$ ，则 $h(x) = h(y)$ 的概率至多为 $p_2$ ；

$d(x,y)$ 表示 $x$ 和 $y$ 之间的距离， $d_1 < d_2$ ， $h(x)$ 表示哈希函数；

满足以上两个条件的哈希函数称为 $(d_1, d_2, p_1, p_2)$ -sensitive。

LSH把特征描述子从实数空间映射到二进制空间，然后通过  
在二进制空间搜索最近邻来完成原空间的近似最近邻匹配。

# 特征匹配—Hashing

在特征匹配中，对于每一个特征描述子  $D \in R^n$ ，给定LSH哈希函数集合为  $H = \{h_i(x): R^n \rightarrow \{0,1\}\}$ ，从  $H$  中挑选  $M$  个哈希函数  $h_i(x)$ ，利用它们把  $D$  映射为  $M$  维二进制描述子  $B = b_1 b_1 \dots b_M$ ，其中第  $i$  维为  $b_i = h_i(D)$ 。进行描述子匹配时，直接计算二进制描述子  $B$  之间的海明距离即可。

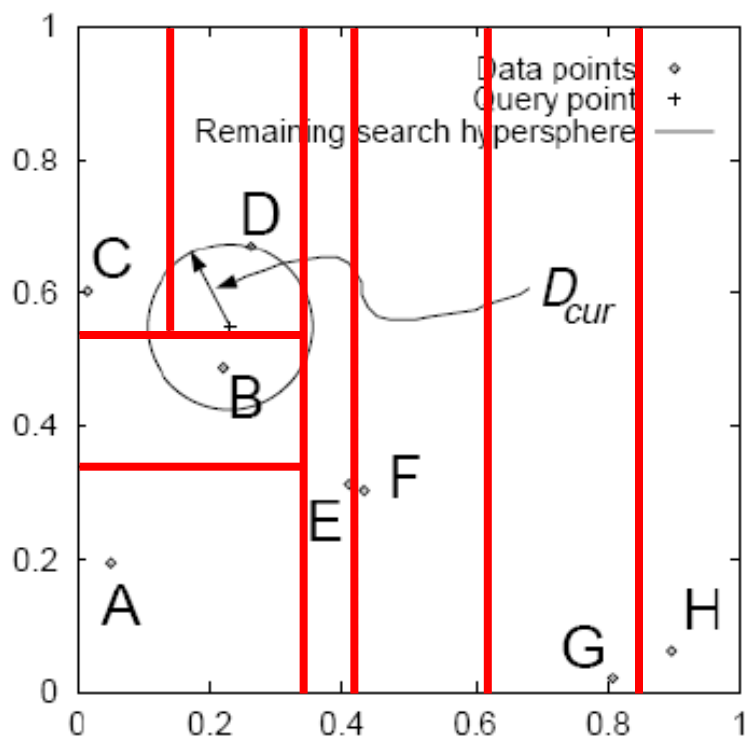
两个长度相等的字符串的海明距离是在相同位置上不同的字符的个数，如10101和00110的海明距离为3。

由于二进制空间中海明距离的计算仅涉及到每个比特的异或操作，而目前的处理器通常都有直接异或操作的处理器指令，因此相比于在实数空间中计算欧氏距离，在二进制空间中计算海明距离在运算速度上提升几十倍。

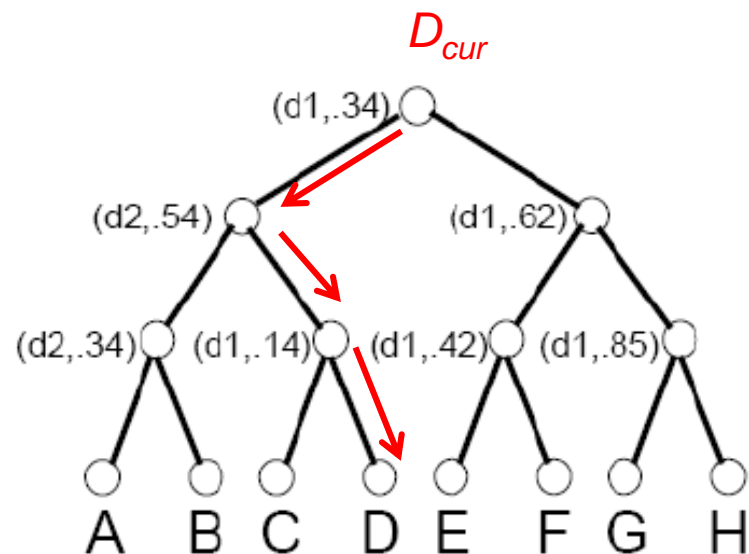


# 特征匹配—KD树

Kd-树（K-dimension tree）是对数据点在k维空间中划分的一种数据结构，是一种平衡二叉树。



(a)

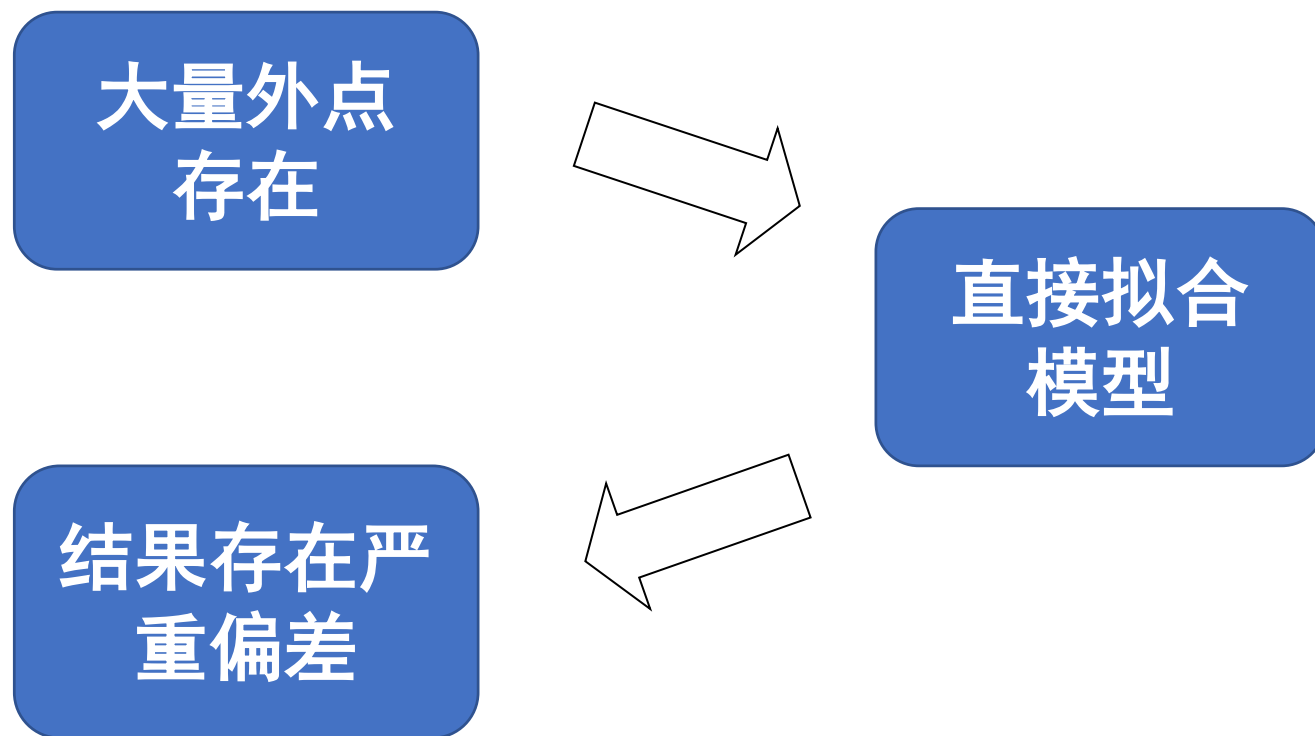


Best Bin First (BBF)搜索

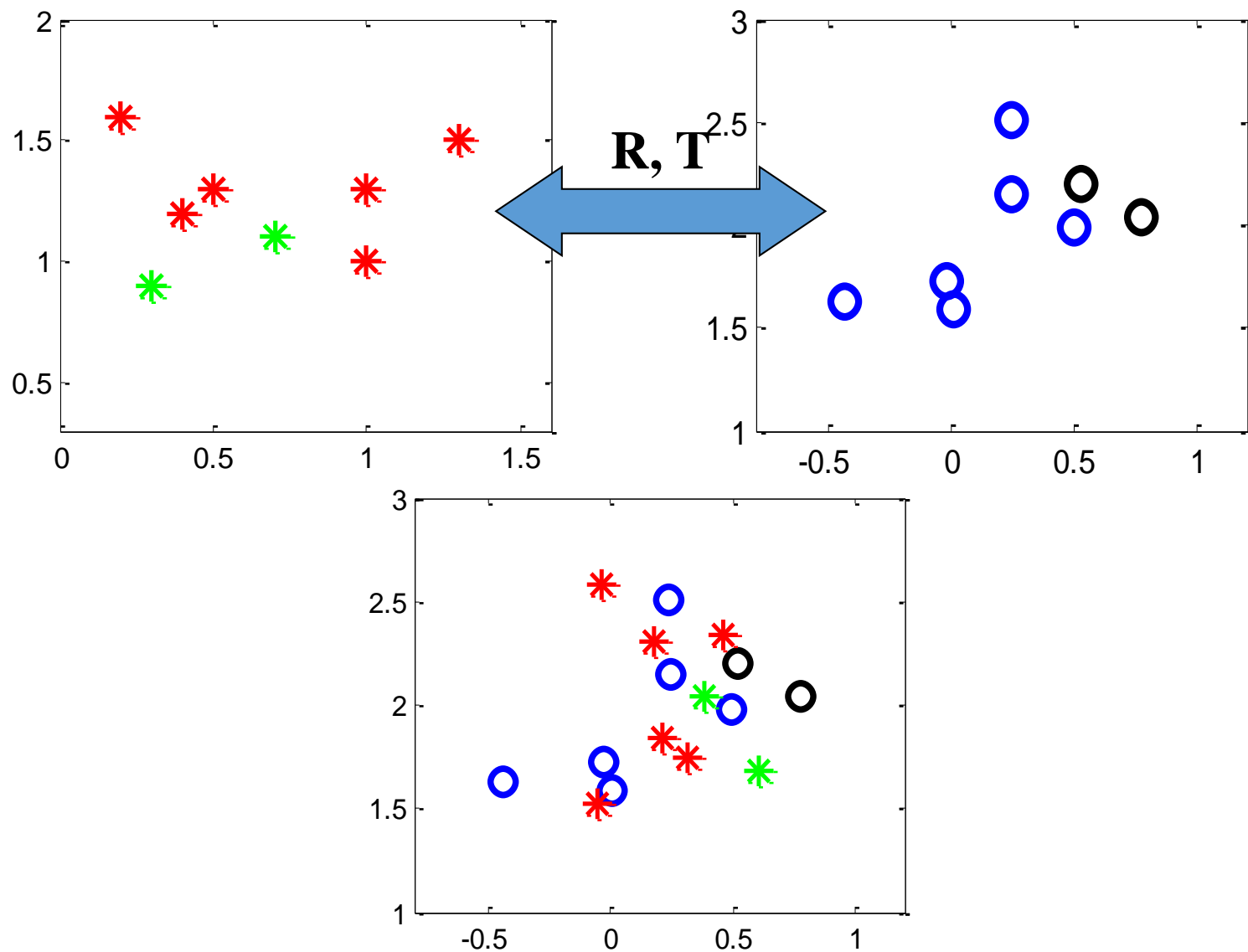
(b)

# 特征匹配—鲁棒匹配

得到大量匹配点之后，如何拟合模型（如配准、拼接）？



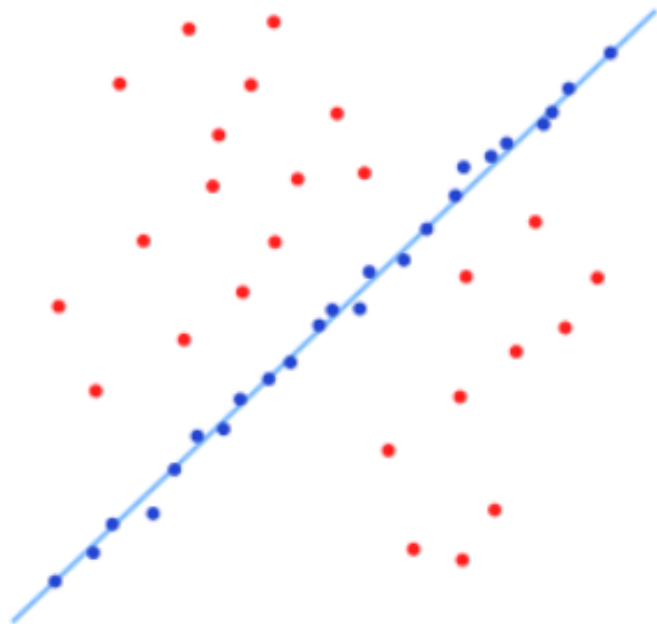
# 特征匹配—鲁棒匹配



# 特征匹配—RANSAC

得到大量匹配点之后，如何拟合模型（如配准、拼接）？

解决方法：RANSAC（**RAN**dom **SA**mples **C**onsensus）



(Fischler and Bolles, 1981)

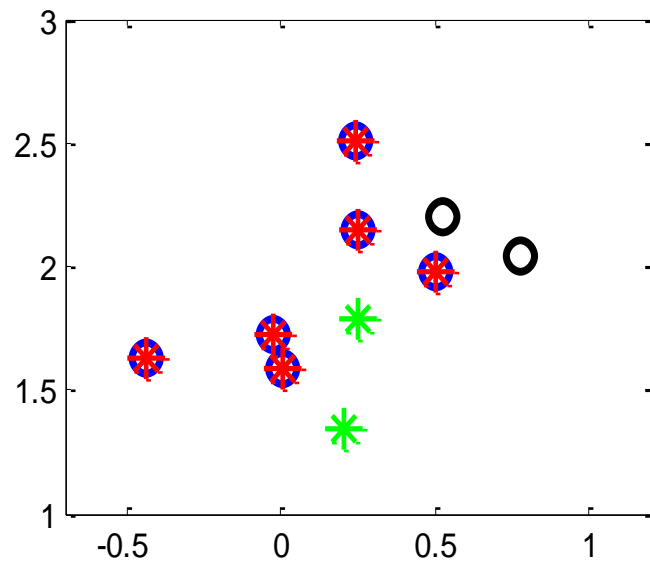
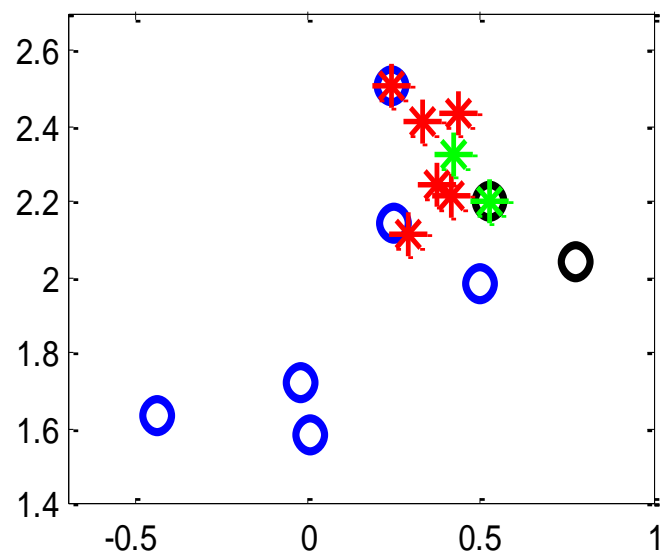
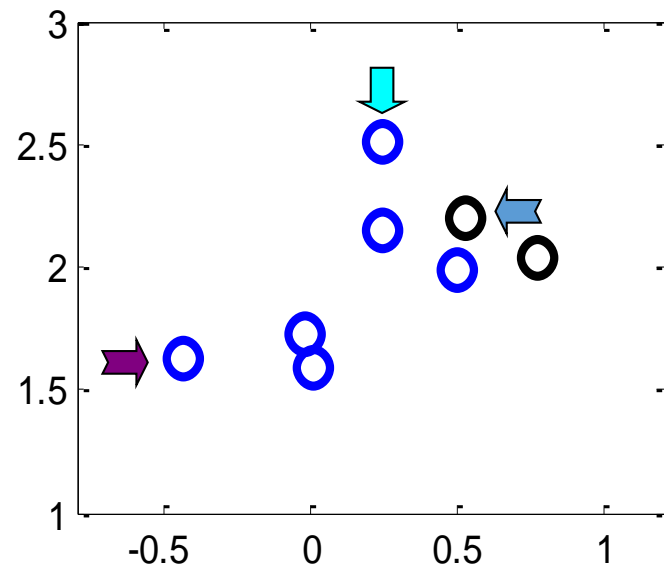
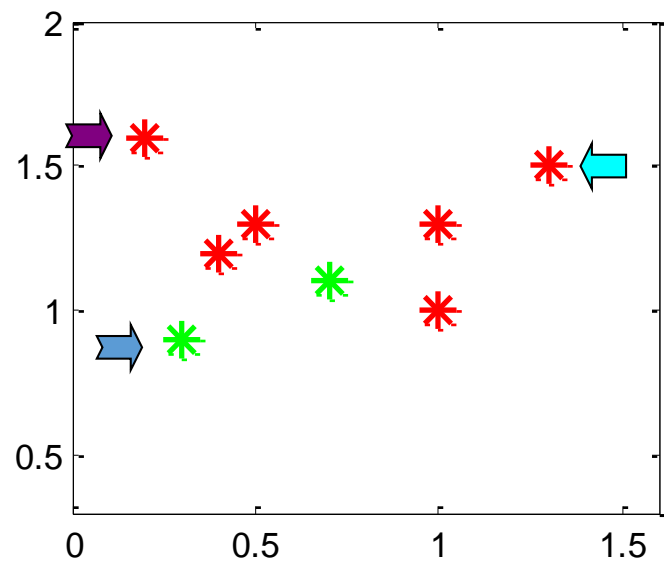
# 特征匹配—RANSAC

## RANSAC计算流程:

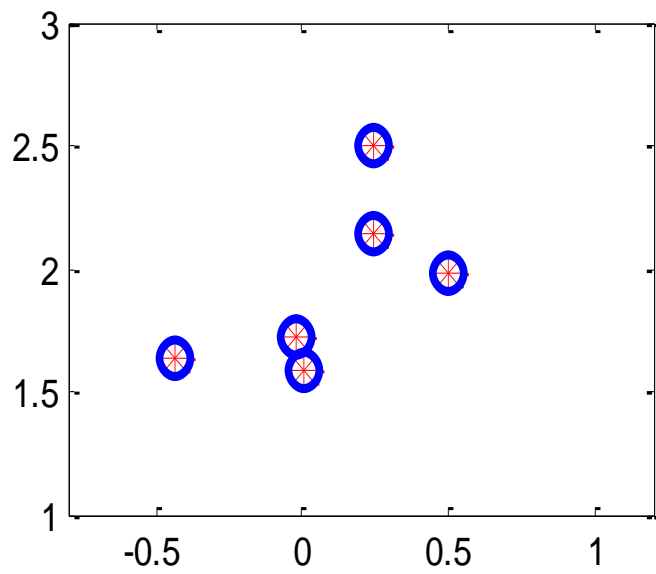
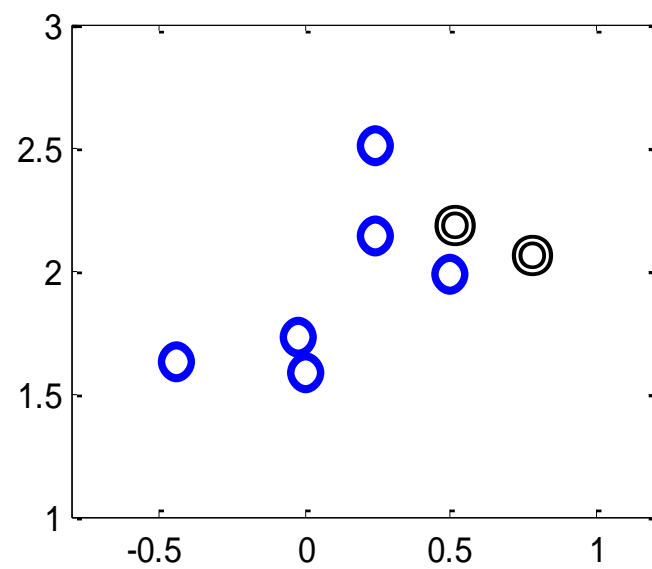
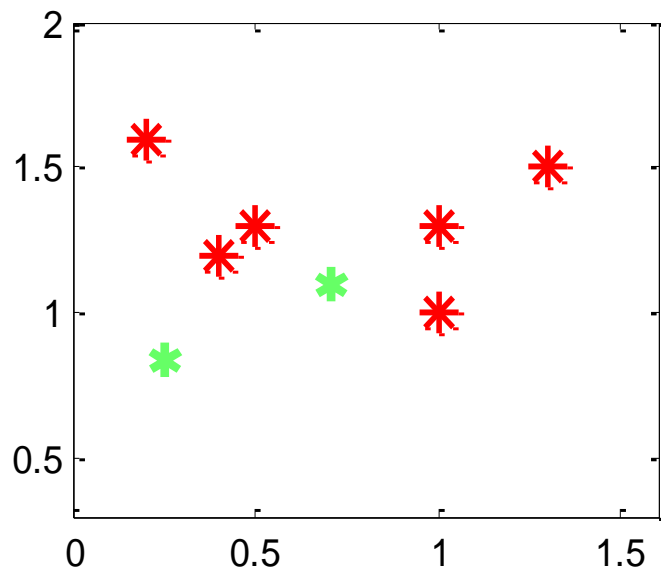
1. 确定求解模型 $M$ ，所需要的最小数据点的个数 $n$ 。由 $n$ 个数据点组成的子集称为模型 $M$ 的一个样本；
2. 从数据点集 $D$ 中随机地抽取一个样本 $J$ ，由该样本计算模型的一个实例 $M(J)$ ，确定与 $M(J)$ 之间几何距离 $<$ 阈值 $t$ 的数据点所构成的集合，记为 $S(M(J))$ ，称为实例 $M(J)$ 的一致集；
3. 如果在一致集 $S(M(J))$ 中数据点的个数 $\#S(M(J)) >$  阈值 $T$ ，则用 $S(M(J))$ 重新估计模型 $M$ ，并输出结果；如果 $\#S(M(J)) <$  阈值 $T$ ，返回到步骤2；
4. 经过 $K$ 次随机抽样，选择最大的一致集 $S(M(J))$ ，用 $S(M(J))$ 重新估计模型 $M$ ，并输出结果。



# 特征匹配—RANSAC



# 特征匹配—RANSAC



# 总结—本节课内容回顾

特征描述子:

- MOPS
- SIFT
- PCA-SIFT
- GLOH
- DAISY

快速匹配:

- LSH
- KD Tree

特征匹配方法:

- 单阈值
- 最近邻
- 最近邻比值

鲁棒匹配:

- RANSAC

# 总结—参考文献

特征匹配参考文献:

- Brown, M., Szeliski, R., and Winder, S. **Multi-image matching using multi-scale oriented patches**. CVPR 2005.
- Lowe, D. G. **Distinctive image features from scale-invariant keypoints**. International Journal of Computer Vision, 60(2):91–110, 2004.
- Mikolajczyk, K. and Schmid, C. **A performance evaluation of local descriptors**. IEEE T-PAMI, 27(10):1615–1630, 2005.
- Tola, E., Lepetit, V., and Fua, P. **DAISY: An efficient dense descriptor applied to wide-baseline stereo**. IEEE T-PAMI, 32(5):815–830, 2010.