

CS131 Lecture07: 特征和拟合/特征描述符

by: 斯坦福大学计算机科学系

github: https://github.com/zhaoxiongjun/CS131_notes_zh-CN (包含中英文版课件及相关课程视频)

1 RANSAC

1.1 目标

随机样本一致性 (ransac) 用于图像中的模型拟合 (例如, 线检测); 它对于对象识别和其他应用非常有用。它通常比纯边缘检测更有效, 这种检测容易受到几个限制: 由于噪声/杂波而包含额外点的边缘、某些边缘部分被遗漏, 以及测量边缘方向上存在噪声。

1.2 动机

RANSAC 的一个主要优点是, 即使在参数数目较高的情况下, 它也相对高效和准确。但是, 应该注意的是, RANSAC 可能会在噪声较大的图像中出现故障或产生不准确的结果。

1.3 通用方法

RANSAC 的算法是, 通过随机抽样一组边上的点, 并多次对这些点应用一条最佳拟合线, 我们很有可能找到一条非常适合这些点的线。以下是“RANSAC 回路”的一般过程:

- 1: 从要与直线匹配的点的总体组中随机选择一组种子点。
- 2: 计算种子组之间的最佳匹配的线。例如, 如果种子组只有两个不同的点, 那么很明显只有一条线穿过这两个点, 这可以相对容易地从点的坐标确定。
- 3: 通过迭代数据集中的每个点并计算其与该线的距离, 找到该线的内点数; 如果小于 (预定的) 阈值, 则计入内点数。否则, 它将被视为异常值。
- 4: 如果内点数量足够大, 则得出该线是“良好”的结论, 并且至少存在一条

线，其中包括在上一步中计算的点。为了进一步改进直线，使用距离在阈值内的所有点，使用最小二乘估计重新计算直线。将此转换保持为最接近数据的直线。

1.4 缺点

RANSAC 最大的缺点是它对高噪声数据的处理效率低下；给定数据集中异常值的分数增加会导致模型拟合所需的样本数量增加（例如，最佳拟合线）。更重要的是，图像的噪声越大，就越不可能认为直线在拟合数据方面足够好。这是一个重要的问题，因为大多数现实世界的问题都有相对较大比例的噪声/异常值。

2 局部不变特征

2.1 动机

局部不变图像特征及其描述符被广泛应用于计算机视觉应用中；它们包括但不限于目标检测、分类、跟踪、运动估计、全景拼接和图像匹配。以前讨论过的方法，如互相关，在许多这样的应用中都是低效和不可靠的。该方法通过在图像中找到局部的、独特的结构（即特征），并使用周围区域（例如，以检测到的特征为中心的小块）描述每个特征。这种图像特征的“局部”表示（而不是“全局”表示，例如互相关）提供了解决上述计算机视觉问题的更强大的方法；这种策略对对象旋转、视点转换和比例变化是不变的。

2.2 通用方法

使用局部不变特征的一般方法如下：

- 1：找到并定义一组独特的关键点。
- 2：定义关键点周围的局部区域。
- 3：提取并规范区域内容。
- 4：从标准化区域（即像素强度函数）计算局部描述符。
- 5：匹配本地描述符。

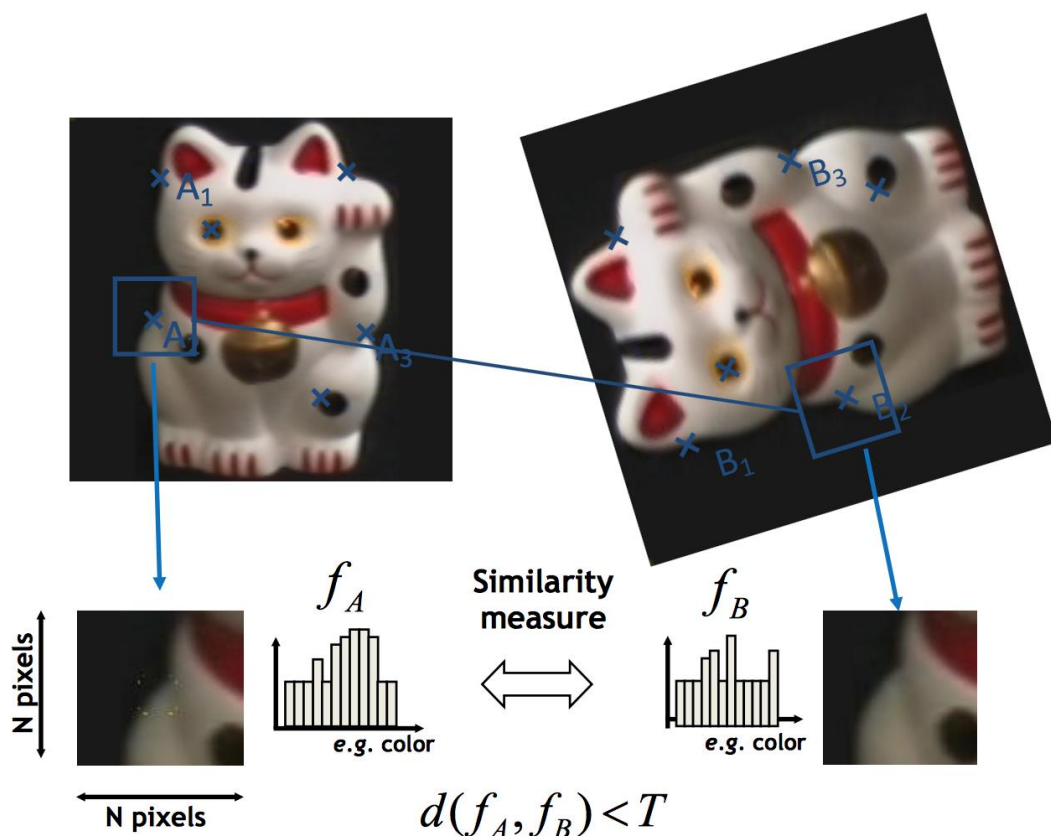


图 1：在相似的图片中识别出局部特征；对标准化区域（小块）计算描述符，并相互比较以量化相似性。

2.3 要求

良好的本地特性应具有以下特性：

- 1：重复性：对于同一物体或场景，在不同的图像条件下，例如照明或视点变化，在两个被比较的图像中都应该可以检测到大量的特征。换句话说，特征检测应该对旋转和视点的变化保持不变，并且能够有力地处理照明变化、噪声和模糊。
- 2：位置：特性应该是局部的，以避免由阻塞和混乱引起的问题。
- 3：数量：应选择足够的特征，以确保依赖于对象检测等描述符的技术的准确性。
- 4：显著性：结果需要包含显示大量变化的“显著”图像特征；这可以确保检测到的特征可以彼此区分，并在不同的图像之间正确匹配。
- 5：效率：新图像中的特征匹配有利于实时应用。

3 关键点定位

3.1 动机

关键点定位的目标是一致和重复地检测特征，允许更精确的定位，并在图像中找到有趣的内容。

3.2 通用方法

我们将寻找角落，因为它们大多数图像中是可重复的和独特的。为了找到角落，我们在各个方向寻找强度的大变化。为了提供上下文，一个“平面”区域在任何方向上都不会发生变化，并且一条边沿着边的方向也不会发生变化。我们将使用 Harris 技术找到这些角落。

3.3 Harris 算子

Harris 算子背后的算法是，如果一个窗口（W）在图像上滑动，则由移动引起的像素值强度变化在角落处最高。这是因为在拐角处的两个方向（x 和 y）都可以观察到像素强度的变化，而在边缘处仅限于一个方向，在平面图像区域可以忽略不计。要计算由移位引起的强度变化 $[u, v]$ ：

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

为了找到角落，我们必须最大化这个函数。然后在过程中使用泰勒展开式得到以下方程：

$$E(u, v) = [u \quad v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

M 定义如下：

$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

这个矩阵表示：

$$M = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

角落既有大的特征值，也有相似的特征值，而边缘有一个显著较大的特征值，而平坦区域有两个较小的特征值。由于特征值的计算非常密集，因此采用另一种方法，使用 CRF（角落反应函数）用于计算每个窗口的分数：

$$\theta = \det(M) - \alpha \text{trace}(M)^2$$

其中 α 是一个范围在.04 – .06 的常数

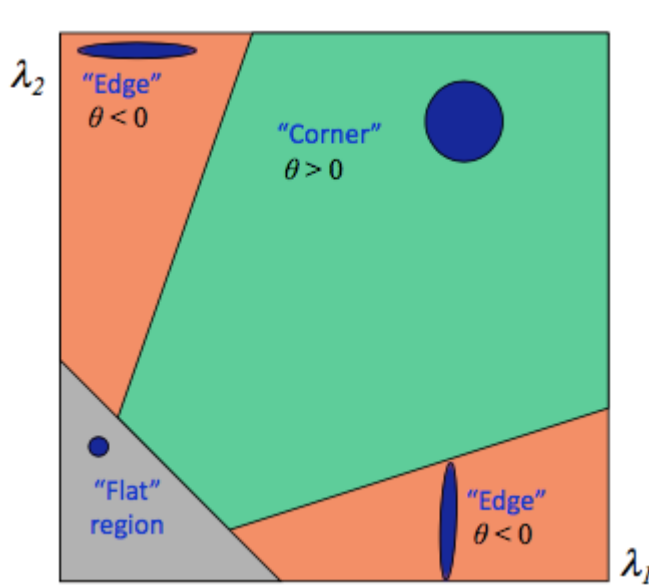


图 2: 来自角响应函数 CRF 的 θ 阈值的可视化表示角、边或平面区域
这不是旋转不变量。为了考虑旋转不变性，我们将使用高斯平滑，其中高斯已经执行加权:

$$M = g(\sigma) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

下图是 Harris 算子识别的一个关键点:



图 3: 图像上 Harris 关键点检测的示例

4 规模不变的关键点检测

4.1 动机

早些时候，我们用 Harris 算子寻找关键点或角落。这个探测器使用小窗户以保持良好的位置。因为 Harris 算子使用这个小窗口，如果是动画，那么现在的窗口不认为相同的整体具有较小的边缘。

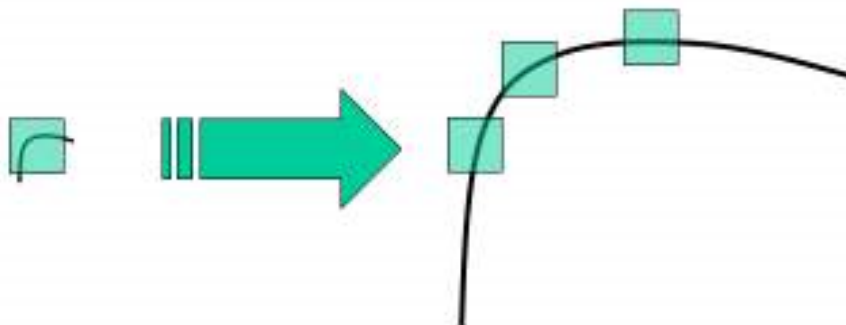


图 4:Harris 探测器窗口的比例增加。

在上面的图片中，我们可以看到右边的三个窗口不再看到 x 和 y 方向的锐利梯度。因此，这三个窗口都是分类的边缘。为了解决这个问题，我们需要规范化检测器的规模。

4.2 解决办法

我们设计了一个尺度不变的函数，这意味着无论尺度如何，对应区域的函数值都是相同的。我们可以用一个圆来表示这个可伸缩函数。圆上的点是圆半径区域大小的函数。

4.3 通用方法

我们可以找到一个函数的局部最大值。相对于局部最大值，区域大小应相同，而不考虑比例。这也意味着区域大小与图像比例是同一变量。一个“好”函数产生一个单独的局部最大值。一般来说，我们应该使用一个对强烈对比反应良好的函数。

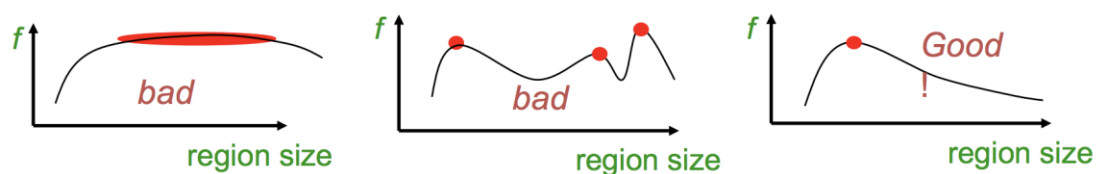


图 5：查找局部最大值的不同函数示例

函数定义为： $f = \text{kernel} * \text{image}$ 。两个这样的核包括拉普拉斯核和高斯核（DoG）。

$$L = \sigma^2(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

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

这两个内核都是尺度不变和旋转不变的。