

# CS131 Lecture05: 边缘检测

by: 斯坦福大学计算机科学系

github: [https://github.com/zhaoxiongjun/CS131\\_notes\\_zh-CN](https://github.com/zhaoxiongjun/CS131_notes_zh-CN) (包含中英文版课件及相关课程视频)

## 1 接上一节

### 1.1 线性系统

线性系统（过滤器）形成的新图像，其像素是原始像素选择分组的加权和。使用不同的模式和权重可以放大原始图像中的不同特征。系统  $S$  是一个线性系统，当且仅当它满足系统的叠加性质：

$$S[\alpha f_i[n, m] + \beta f_j[h, m]] = \alpha S[f_i[n, m]] + \beta S[f_j[h, m]]$$

如前一节课中介绍的，对输入图像应用过滤器的过程称为卷积。

### 1.2 LSI（线性位移不变系统）和脉冲响应

移位不变系统是这样一种系统：移动输入也会等量地移动输出。线性位移不变系统（LSI）的特点是对脉冲的响应；这种响应称为脉冲响应。脉冲响应有助于我们理解 LSI 系统对任何给定输入信号的输出。

### 1.3 为什么卷积翻转？

二维互相关交换性的证明：

$$f[n, m] * h[n, m] = \sum_k \sum_l f[k, l] \cdot h[n - k, m - l]$$

let  $N = n - k$ ,  $M = m - l$  so  $k = n - N$  and  $l = m - M$

$$\begin{aligned} &= \sum_k \sum_l f[n - N, m - M] \cdot h[N, M] \\ &= \sum_N \sum_M h[N, M] \cdot f[n - N, m - M] \\ &= h[n, m] * f[n, m] \end{aligned}$$

### 1.3.1 例子

将核  $k$  应用于矩阵  $M$ :

$$M = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, k = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{bmatrix}$$

$$M * k = \begin{bmatrix} x_{3,3} & x_{3,2} & x_{3,1} \\ x_{2,3} & x_{2,2} & x_{2,1} \\ x_{1,3} & x_{1,2} & x_{1,1} \end{bmatrix}$$

这里，内核与卷积匹配。相反，输出等于内核  $k$  沿  $x$  和  $y$  方向翻转。为了解释这个结果，内核在最初的步骤中被翻转，以确保正确的输出形式（这里和后面的 3.1 也是同样的解释）。

## 1.4 卷积和交叉相关性

卷积是一个积分，表示一个函数在另一个函数上移动时的重叠量。我们可以把卷积看作是一种滤波运算。

相关性计算两个输入信号（例如，两个图像块）的相似性度量。当两个信号最匹配时，相关输出达到最大值。相关性可以用来衡量两个信号的相关性。

## 2 哺乳动物边缘检测

### 2.1 Hubel 和 Wiesel 教授

在 Hubel 和 Wiesel[1]进行的一系列实验中，记录了猫大脑中的神经元反应，以观察大脑的每个部分对不同刺激的反应。他们发现猫的神经元在不

同方向的边缘最为兴奋；也就是说，某些神经元与特定方向移动的边缘或边缘的特定方向相关。当这些边中的一条在它的视野中移动时，会导致特定的神经元兴奋地放电。

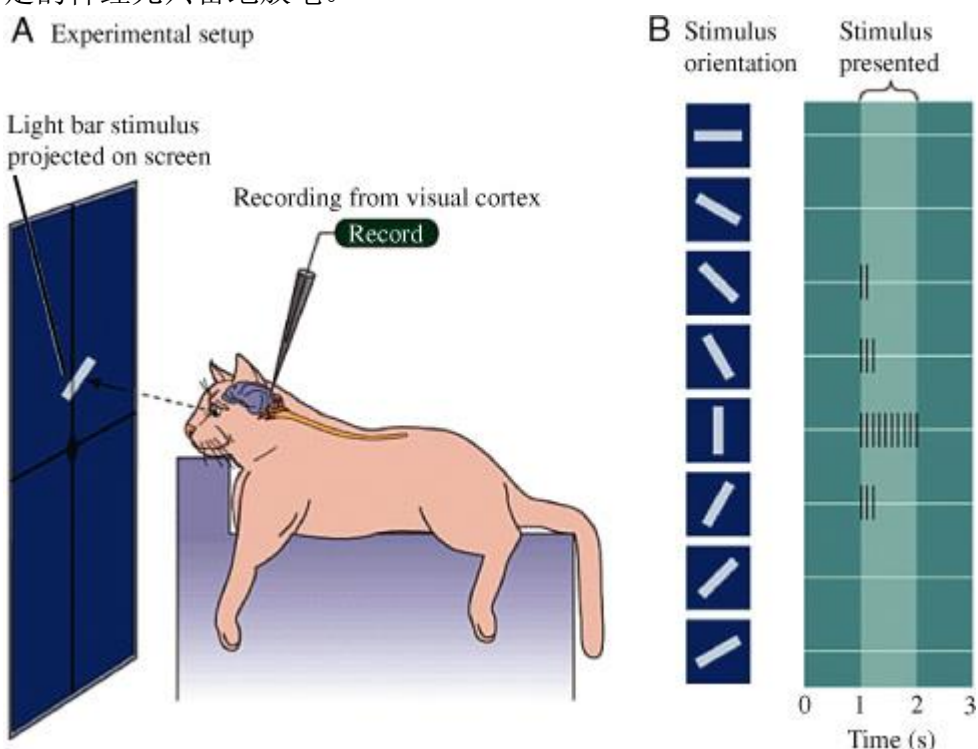


图 1:Hubel 和 Wiesel 的实验。资料来源:【1】; 第 5 讲, 幻灯片 34

## 2.2 Biederman 教授

比德曼研究了人类识别他们所看到的物体的速度。为了测试这一点，他画出了常见的和可识别的物体的轮廓，并把它们分成两半，每个线段只分成两半。然后将这些轮廓展示给参与者，以测试他们是否能够识别原始对象，同时只看到原始轮廓的一半。

令人惊讶的是，在人们识别物体的速度方面，他没有观察到任何差异。他们很容易通过物体边缘的一部分来识别它。这项研究有利于计算机视觉提供一种洞察：即使只显示了原始图像的一部分，理论上一个系统仍然应该能够识别整个对象或场景。

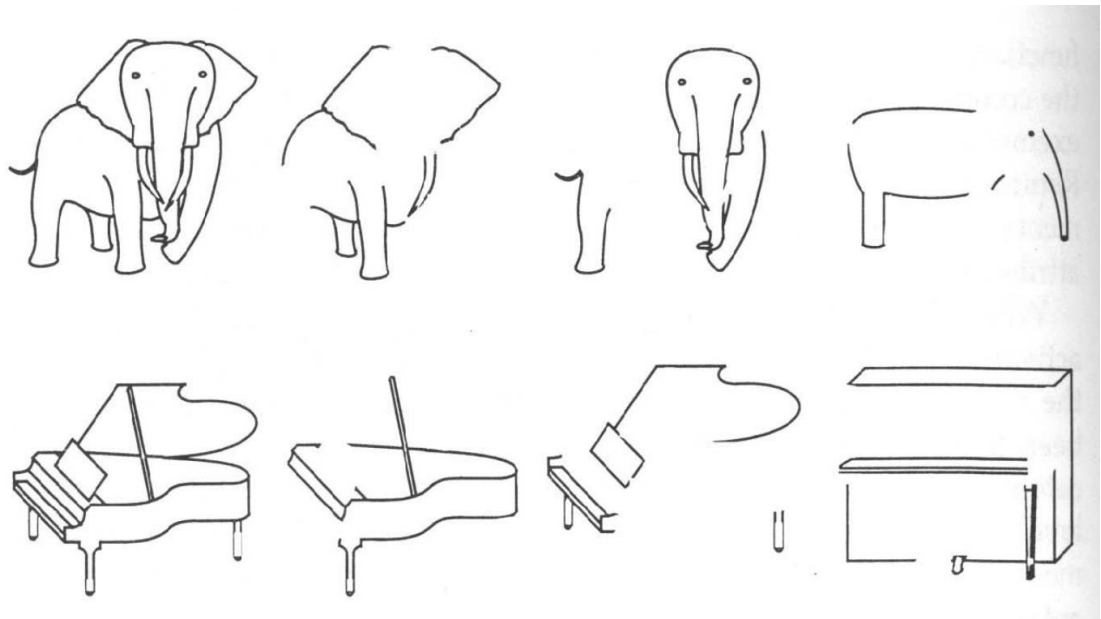


图 2：比德曼轮廓的例子。资料来源：第 5 讲，幻灯片 36

## 2.3 Walther, Chai, Caddigan, Beck & Fei-Fei

在类似的实验中，另一组研究人员对同一幅图像的两种变化——原始彩色图像和该图像的轮廓——进行了人类彩色图像与线条图像识别的对比试验。他们通过不同的层次追踪识别，每一个层次在大脑的视觉皮层上都有不同程度的处理。他们发现，较低层次的人可以通过线条图像更快地识别场景，但是当它向上移动到大脑的各个层次（这就编码了越来越高层次的概念）时，与线条图相比，颜色图在帮助人们识别场景方面更有帮助。这被认为是发生的，因为较低层更好地识别碎片，如边缘，而较高层更好地识别概念（如“人”、“椅子”或“老虎”）。

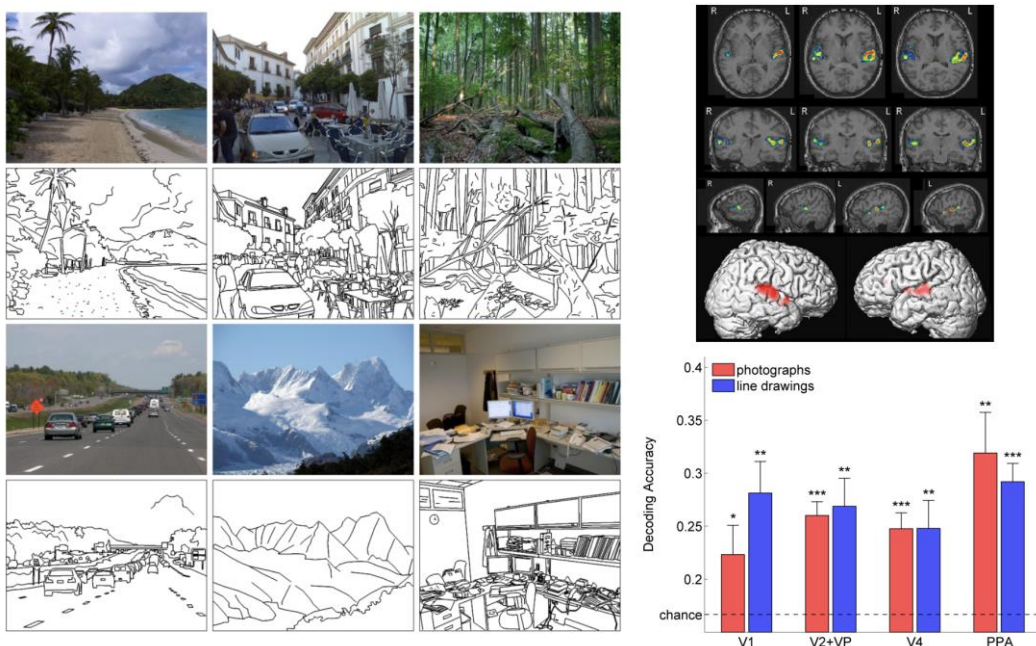


图 3：图像和结果的可视化。资料来源：【2】；第 5 讲，幻灯片 37

## 3 计算机视觉之边缘检测

边缘检测的目标是识别图像中的突然变化部分（不连续的）。直观地说，图像中的大多数语义和形状信息都可以编码到其边缘。

边缘帮助我们提取信息、识别对象以及恢复几何图形和视点。它们是由于表面法向、深度、表面颜色和照明的不连续性而产生的。

### 3.1 一维离散导数的类型

有三种主要类型的导数可以应用于像素。它们的公式和相应的过滤器（中括号里的代表滤波器的值）是：

后向型：

$$\frac{df}{dx} = f(x) - f(x-1) = f'(x) \quad [0, 1, -1]$$

前向型：

$$\frac{df}{dx} = f(x) - f(x+1) = f'(x) \\ [-1, 1, 0]$$

中心型：

$$\frac{df}{dx} = f(x+1) - f(x-1) = f'(x) \\ [1, 0, -1]$$

### 3.2 二维离散导数

梯度向量：

$$\nabla f(x, y) = \begin{bmatrix} f_x \\ f_y \end{bmatrix}$$

梯度幅度：

$$|\nabla f(x, y)| = \sqrt{f_x^2 + f_y^2}$$

梯度方向：

$$\theta = \tan^{-1}\left(\frac{\frac{df}{dy}}{\frac{df}{dx}}\right)$$

### 3.3 举例

矩阵索引处的梯度可以根据扩展到二维的中心离散导数方程，使用相邻像素来近似。一个滤波器如下：

$$\frac{1}{3} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

当重叠在像素  $x[m, n]$  的顶部时，使滤波器的中心位于  $x[m, n]$  处，如下图所示。

$$\begin{bmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & x_{m-1,n-1} & x_{m-1,n} & x_{m-1,n+1} & \dots \\ \dots & x_{m,n-1} & x_{m,n} & x_{m,n+1} & \dots \\ \dots & x_{m+1,n-1} & x_{m+1,n} & x_{m+1,n+1} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

输出：

$$\frac{1}{3} \left( (x_{m-1,n+1} - x_{m-1,n-1}) + (x_{m,n+1} - x_{m,n-1}) + (x_{m+1,n+1} - x_{m+1,n-1}) \right) =$$

相当于像素  $(m, n)$  在水平 ( $n$ ) 方向上的梯度近似值。该滤波器检测水平边缘，还需要一个单独的内核来检测垂直边缘。

## 4 简单边缘探测器

### 4.1 特征化边缘

对边缘进行特征化（即对其进行适当的特征化以便能够识别）是检测边缘的重要第一步。为了我们的目的，我们将把边缘定义为图像强度函数中一个快

速变化的地方。如果我们沿着水平扫描线绘制强度函数，我们可以看到边缘对应于导数的极值部分。因此，注意到沿着这个图的急剧变化的地方可能会给我们带来边缘。

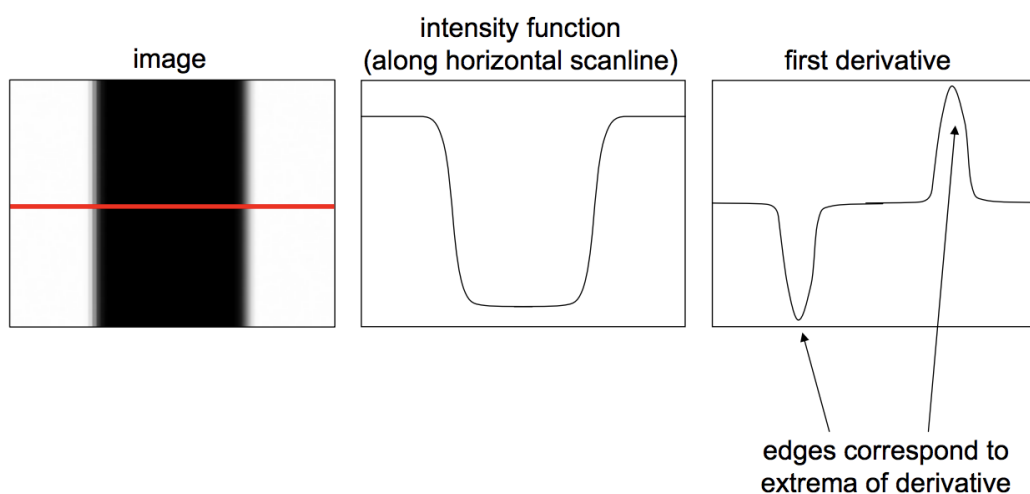


图 4：图像的强度函数和一阶导数。资料来源：第 5 讲，幻灯片 66

## 4.2 图像梯度

图像的梯度定义如下：

$$\nabla f = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right],$$

其方向定义为：

$$\theta = \tan^{-1} \left( \frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}} \right).$$

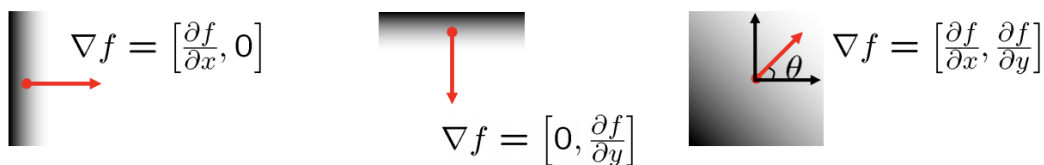


图 5：梯度向量方向。资料来源：第 5 讲，幻灯片 67

梯度向量指向强度增加最快的方向。例如，在垂直边缘，强度的最快变化发生在 X 方向。



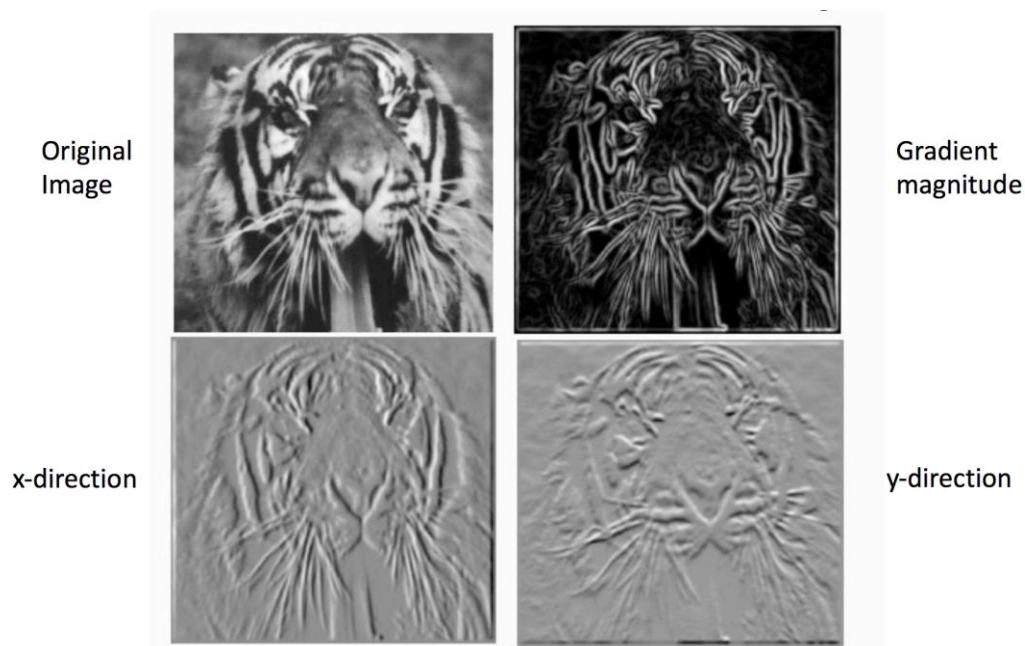


图 6：应用于老虎图像的渐变。资料来源：第 5 讲，幻灯片 68

### 4.3 噪音的影响

如果图像中存在过多的噪声，偏导数将无法有效识别边缘。

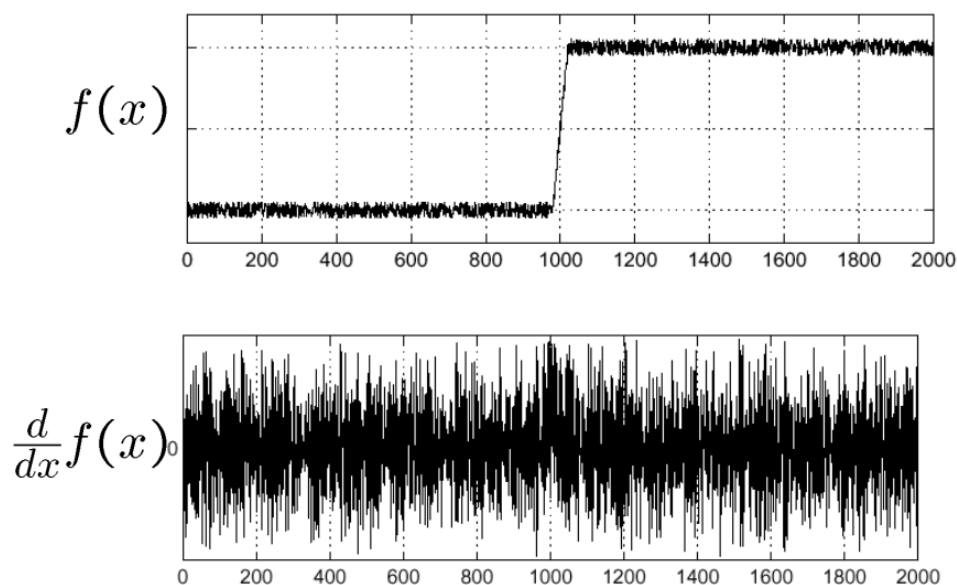


图 7：噪声图像中边缘的导数。资料来源：Steve Seitz；第 5 讲，幻灯片 70

为了减少噪声，图像必须先平滑。这是一个重新计算像素值的过程，以便它们更接近它们的邻居。平滑是通过将图像与滤波器（如高斯核）卷积来实现的。

当然，平滑图像时要记住一些问题。图像平滑可以消除噪声，但也会使边缘



变模糊；使用大的过滤器可能会导致丢失图像的边缘和更丰富的细节。

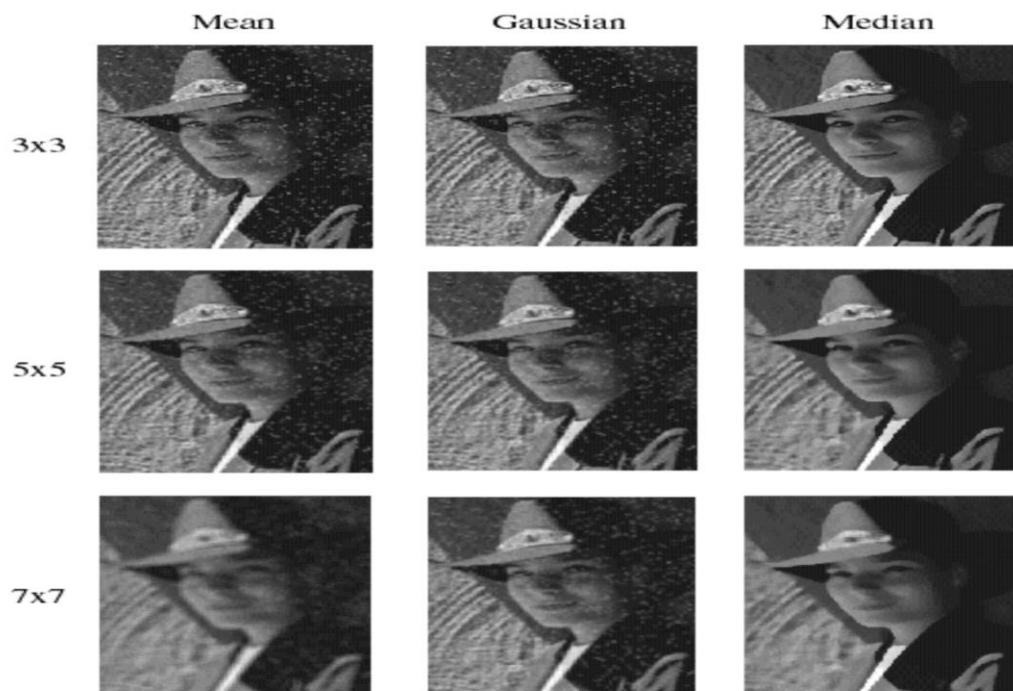


图 8：使用不同的过滤器和过滤器尺寸进行平滑处理比对。资料来源：Steve Seitz；第 5 讲，幻灯片 75

图像平滑有助于边缘检测。在对图像  $f$  进行平滑处理后，通过核  $g$  计算  $f * \frac{dg}{dx}$ ，寻找峰值。

## 4.4 高斯模糊

高斯模糊是用高斯函数对图像进行模糊以降低图像噪声的结果。它是一个低通滤波器，这意味着它可以衰减高频信号。通常使用高斯模糊作为一个初步步骤。

一维定义：

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

二维定义：

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

## 5 设计一个好的边缘检测器

好边缘探测器必须具有下面特性：

- 良好的检测结果：  
它必须将检测假阳性（通常由噪声引起的假阳性边缘）和假阴性（丢失真实边缘，可能由平滑等原因引起）的概率降到最低。如果它检测到某个边缘，它应该是一个边缘。
- 检测位置精确  
检测到的边缘必须尽可能接近原始图像中的实际边缘。探测器必须确定边缘出现的位置并精确定位边缘的准确位置；在确定每个边缘中涉及的像素时，探测器也必须保持一致。
- 无声反应（Silent response）  
它必须最小化真实边缘周围的局部最大值（每个真实边缘点只返回一个点）。返回存在的一个非常特定的边，而不是将一个边拆分，检测到多个边。换句话说，只有实际边界的边缘被捕获；其他可能性被抑制。

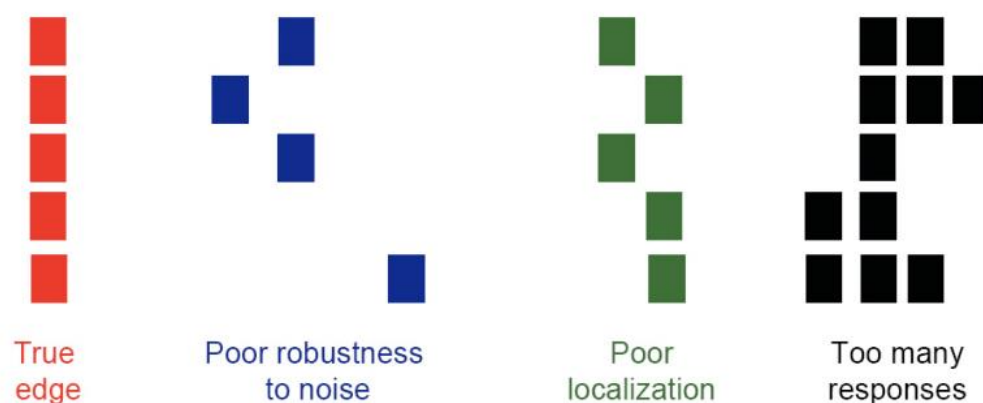


图 9：差的边缘检测器的样本。资料来源：第 5 讲，第 84 页幻灯片