

CS131 Lecture04: 像素和过滤器

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 图像采样和量化

1.1 图像类型

- 二进制图像：包含黑色（0）或白色（1）的像素。
- 灰度图像：比黑白图像有更大的强度范围。每个像素都是灰色阴影，像素值介于 0（黑色）和 255（白色）之间。
- 彩色图像：有多个颜色通道；每个彩色图像可以用不同的颜色模型（例如，RGB、LAB、HSV）表示。例如，RGB 模型中的图像由红色、绿色和蓝色通道组成。通道中的每个像素的强度值都在 0-255 之间。请注意，此范围取决于颜色模型的选择。三维张量通常表示彩色图像（宽度 x 长度 x 3），其中用 3 个通道可以表示的颜色模型，比如有 RGB（红-绿-蓝）、LAB（亮度-A-B）和 HSV（色调饱和度值）。

1.2 采样和分辨率

图像是样本：它们不是连续的；它们由一定大小和密度的离散像素组成。这可能导致误差（或颗粒度），因为像素强度只能用一定的分辨率来测量，而且必须近似。

分辨率是一个采样参数，以每英寸点数(dpi)定义。屏幕的标准 dpi 值为 72 dpi。

像素是量化的（即，所有像素（或一个像素的通道）都有一个值属于一组数值范围（通常为[0, 255]）。由于精度有限，量化和采样会丢失信息。

2 图像直方图

柱状图测量图像中亮度的频率：特定像素值在图像中出现的次数。



图 1：不同像素密度的图示。摘自随附的演讲幻灯片。（幻灯片 14）

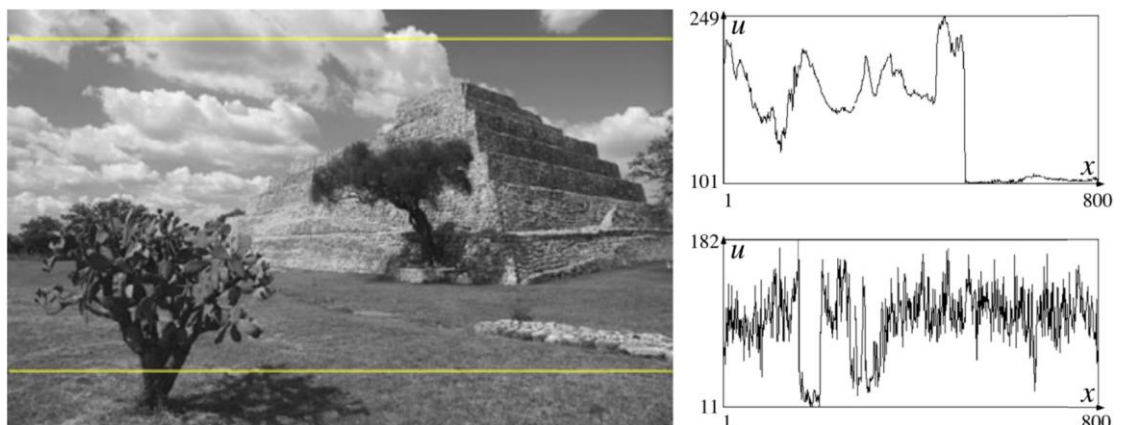


图 2：图像在两个垂直位置采样，采样一片天空，采样一片草地。相应的柱状图显示在右侧。改编自随附的演讲幻灯片（幻灯片 23）

柱状图帮助我们检测图像中的特定特征，例如：

- 天空：平滑的颜色表示图像的一致性，与天空的图像一致。
- 草地：锯齿状的柱状图显示了各种各样的颜色，与草地的阴影一致。
- 人脸：人脸的颜色组成将显示在柱状图中。

图像直方图测量图像中某些灰度强度的频率。柱状图可以用来对事物的外观提供可量化的描述；它可以用作分类器的输入。

3 用函数表示图像

我们在计算机视觉中处理的大多数图像都是数字的，这意味着它们是被拍摄场景的离散表示。这种离散化是通过将二维空间采样到一个规则的网格上实现的，最终将图像表示为一个整数值矩阵。

在处理图像时，我们可以将图像矩阵想象为无限高和无限宽。然而，显示的图像只是这个无限矩阵的有限子集。利用这种图像的定义，我们可以把它们写成矩阵中的坐标。

$$\begin{bmatrix} \ddots & & & & \vdots & & \ddots \\ \dots & f[-1, 1] & f[0, 1] & f[1, 1] & \dots & & \\ & f[-1, 0] & f[0, 0] & f[0, 1] & \dots & & \\ & f[-1, -1] & f[0, -1] & f[1, -1] & \dots & & \\ \ddots & & \vdots & & \ddots & & \end{bmatrix}$$

$$f[n, m] \longrightarrow \boxed{\mathcal{S}} \longrightarrow g[n, m]$$

图 3: f 到 g 映射的图形表示

图像也可以作为函数 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^N$ 来处理。当我们这样做时， $f[n, m]$ ，其中 $f[n, m]$ 表示位置 (m, n) 处像素的强度。注意，我们使用方括号而不是典型的括号来表示离散函数。

当我们把一个图像当作一个函数时，它是在一个范围有限的矩形上定义的。例如，下面的函数 f 返回位于 a 和 b 水平和 c 和 d 垂直之间的图像中单个像素的（灰度）强度。

$$f: [a, b] \times [c, d] \rightarrow [0, 255] \quad (\text{Grayscale Pixel Intensity})$$

集合 $[a, b] \times [c, d]$ 的值被称为域，包含对函数 f 有效输入的所有值，而 $[0, 255]$ 是定义可能输出集的范围。

图像也可以作为函数映射 $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ 来处理。例如，给定像素的 RGB 强度可以写成下面的函数 g ：

$$g[x, y] = \begin{bmatrix} r[x, y] \\ g[x, y] \\ b[x, y] \end{bmatrix} \quad (\text{Color Pixel Intensity})$$

4 线性系统（过滤器）

术语过滤是指形成新图像的过程，其像素值是原始像素值的变换。通常，应用过滤器的目的是提取有用信息（例如边缘检测）或调整图像的视觉特性（例如去噪）。

滤波器是系统的例子，它是将输入函数 $f[m, n]$ 转换为输出（或响应）函数 $g[m, n]$ 的单位，其中 m, n 是独立变量。处理图像时， m, n 表示图像中的空间位置。通常， S 被称为“系统操作员”，它将输出集 $g[m, n]$ 的一个成员映射到可能的输入集 $f[m, n]$ 的一个成员。当使用符号表示 S 时，我们可以写：

$$\begin{aligned} S[g] &= f \\ S\{f[m, n]\} &= g[m, n] \\ f[m, n] &\xrightarrow{S} g[m, n] \end{aligned}$$

4.1 过滤器举例

平均移动

一个直观的过滤器例子是平均移动。此过滤器将一个像素的值设置为其相邻像素的平均值（例如，应用 3×3 过滤器时，求都是半径为 3×3 的 9 个像素得平均值）。在数学上，我们可以把它表示为：

$$g[m, n] = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 f[m-i, n-j] \quad (\text{Weighted Average})$$

这个加权平均滤波器可以平滑图像的锐利边缘，产生模糊或平滑的效果。

图像分割

我们还可以使用过滤器来执行基于简单阈值系统的基本图像分割。在这种情况下，根据像素是否满足阈值 t ，过滤器将像素值设置为极高或极低值。表示如下：

$$g[m, n] = \begin{cases} 255 & f[m, n] \geq t \\ 0 & \text{otherwise} \end{cases} \quad (\text{Threshold})$$

这个基本的图像分割过滤器将图像像素分为亮区和暗区的二进制分类，这取决于 $f[m, n]$ 是否大于等于 t 。

4.2 系统属性

在讨论特定的系统时，描述它们的属性是很有用的。以下是系统可能拥有的属性列表。但是，并非所有系统都具有所有（或任何）这些属性。换句话说，这些是单个系统的潜在特征，而不是一般系统的特征。

振幅特性

- 可加性：一个系统是可加的，如果满足方程：

$$\mathcal{S}[f_i[m, n] + f_j[m, n]] = \mathcal{S}[f_i[m, n]] + \mathcal{S}[f_j[m, n]]$$

- 同质性：一个系统是同质的，如果满足方程：

$$\mathcal{S}[\alpha f_i[n, m]] = \alpha \mathcal{S}[f_i[n, m]]$$

- 叠加性：一个系统具有叠加性，如果满足方程：

$$\mathcal{S}[\alpha f_i[n, m] + \beta f_j[n, m]] = \alpha \mathcal{S}[f_i[n, m]] + \beta \mathcal{S}[f_j[n, m]]$$

- 稳定性：一个系统具有稳定性，如果满足方程（对于相同的 c ）：

$$|f[n, m]| \leq k \implies |g[n, m]| \leq ck$$

- 可逆性：一个系统具有可逆性，如果满足方程：

$$\mathcal{S}^{-1}[\mathcal{S}[f[n, m]]] = f[n, m]$$

空间属性

- 因果关系：一个系统具有因果关系，如果对于 $m < m_0$ 且 $n < n_0$ 有：

$$f[m, n] = 0 \implies g[m, n] = 0$$

- 移不变性：一个系统具移不变性，如果满足方程：

$$f[m - m_0, n - n_0] \xrightarrow{S} g[m - m_0, n - n_0]$$

4.3 线性系统

线性系统是一个满足叠加性质的系统。当我们使用线性系统进行过滤时，我们创建一个新的图像，其像素是原始像素值的加权和，对每个像素使用相同的权重集。线性移动不变量系统是一个线性系统，也是一个移动不变量。

线性系统也有所谓的脉冲响应。为了确定一个系统的脉冲响应，考虑前 $\delta_2 [m, n]$ 。定义如下的函数：

$$\delta_2[m, n] = \begin{cases} 1 & m = 0 \text{ and } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

脉冲响应 r 简单的表示：

$$r = \mathcal{S}[\delta_2]$$

一个简单的线性移动不变量系统是一个基于 δ 函数的移动特性移动图像像素的系统。

$$f[m, n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f[i, j] \delta_2[m - i, n - j]$$

然后，我们可以用叠加性质将任何线性移位不变系统写为该移位系统的加权和。

$$\alpha_1 \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f[i, j] \delta_{2,1}[m - i, n - j] + \alpha_{2,2} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f[i, j] \delta_{2,3}[m - i, n - j] + \dots$$

然后我们将线性移位不变系统的滤波器 h 定义为：

$$h[m, n] = \alpha_1 \delta_{2,1}[m - i, n - j] + \alpha_2 \delta_{2,2}[m - i, n - j] + \dots$$

脉冲响应（对于所有线性系统） $\delta_2[n, m]$ 函数：在一个具有 1 的特定值的像素处返回响应， $h[n, m]$ ：位移 δ 函数返回位移响应线性位移不变系统 (LSI)
示例：移动平均滤波器是脉冲响应的总和。

- 系统满足叠加性

- 有一个脉冲响应: $S[\delta_2[n, m]] = \delta_2[n, m]$
- 离散卷积: $f[n, m] * h[n, m]$ (原始函数对脉冲响应的移位形式的乘法)

5 卷积和关联

5.1 卷积

卷积最简单的方法是使用相邻像素的信息过滤目标像素。一个很好的例子是移动平均值，或者前面讨论过的框式过滤器。

卷积用*表示，例如: $f[n, m] * h[n, m]$ 表示一个函数和移位的脉冲响应相乘。

卷积允许我们通过简单地考虑系统的脉冲响应来计算任何输入信号系统的输出。为了理解这意味着什么，我们必须首先理解如何将信号分解成一组脉冲函数。脉冲(delta)函数 $\delta[n]$ 在 $n=0$ 时定义为1，在其他地方定义为0。如下图所示，任何信号都可以分解为脉冲函数的加权和。

一般来说，任意信号 x 可以写为:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n - k].$$

系统的脉冲响应 $h[n]$ 定义为将脉冲函数传递到系统中所产生的输出。当系统为线性时，缩放脉冲函数会导致相同数量的脉冲响应缩放。此外，当系统是移动不变的，移动脉冲函数移动脉冲响应相同的数量。下图说明了由3个分量组成的信号的这些特性。

更一般地说，对于任意输入信号 $x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n - k]$ 传递到一个线性的、移动不变的系统中，输出为 $y[n] = \sum_{k=-\infty}^{\infty} x[k] h[n - k]$ ，信号 x 与脉冲响应 h 的卷积。

卷积也可以在二维空间中进行。例如，当一个任意的二维信号 $x[n, m] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x[i, j] \delta[n - i, m - j]$ 传递到一个线性的、移动不变的系统中，输出为 $y[n, m] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x[i, j] h[n - i, m - j]$ 信号 x 与脉冲响应 h 的二维卷积。

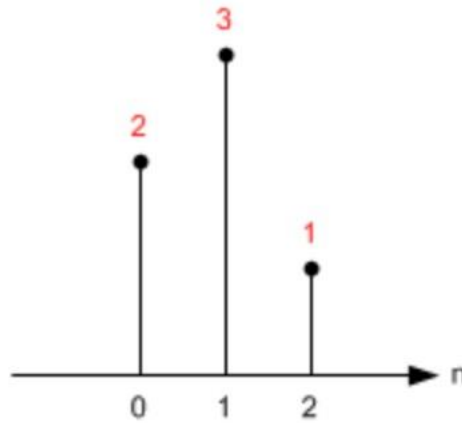


图 4：将信号分解为脉冲函数的示例。（摘自 <http://www.songho.ca/dsp/covolcution/covolcution.html>）

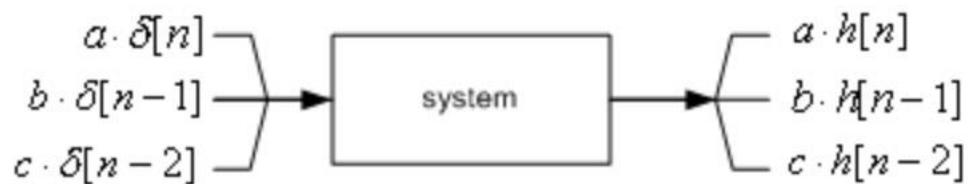


图 5：通过系统发送脉冲函数以创建脉冲响应函数。改编自 <http://www.songho.ca/dsp/covolcution/covolcution.html>

5.2 关联

互相关与卷积相同，只是滤波核没有翻转。二维互相关表示为：

$$r[k, l] = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f[m + k, n + l] g[m, n]$$

它可以通过使用包含目标特性的内核在图像中查找已知特性。

感谢

感谢 Niebles 教授、Ranjay Krishna 和其他教职员编写了宝贵的随附演讲幻灯片并提供了这样一个信息丰富的演讲。