

141 | 掌握计算机视觉任务的基础模型和操作

2018-08-27 洪亮劫

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 06:12 大小 2.84M



今天，我们来聊一聊计算机视觉的一些最基础的操作和任务，包括像素表达、过滤器和边界探测。基于这些内容，我们一起讨论利用计算机来处理视觉问题的核心思路。很多时候，越是基础的内容就越重要，因为只有掌握了基础的思路，我们才能在今后复杂的任务中灵活应用。

像素表达

我们在上一次的分享中谈到了计算机视觉任务中一个非常重要的步骤，那就是把现实世界的信号通过感知仪器（Sensing Device）收集起来，然后在计算机系统中加以表达。那么，在所有的表达中，最基础的就是“**像素表达**”（Pixel）。我们这里就展开说一说这种表达的思路。

把图像信息利用像素来进行表达是一种非常直观简单的表达方式。

对于黑白图像来说，图像就被转换为了 0 或者 1 的**二元矩阵**。这个矩阵的每一个元素就是一个像素，0 代表黑，1 则代表白。

对于灰度图像来说，每一个像素，或者说是矩阵的每一个元素，代表灰度的“**强度**”（Intensity），从 0 到 255，0 代表黑，255 代表白。

对于彩色的图像来说，我们一般要先选择一种模型来表示不同的颜色。一种较为流行的表达方式是 RGB（红、绿、蓝）模型。在这样的模型中，任何一个彩色图像都能够转化成为 RGB 这三种颜色表达的叠加。具体来说，就是 RGB 分别代表三种不同的“通道”（Channel）。每一种通道都是原始图像在这个通道，也就是这个原始颜色下的表达。每一个通道都是一个矩阵像素表达。每一个像素代表着从 0 到 255 的值。换句话说，**一个彩色图像在 RGB 模型下，是一个“张量”（Tensor），也就是三个矩阵叠加在一起的结果。**

针对像素，你需要建立一种概念，那就是**像素本身是对真实世界中的“采样”（Sample）**。每一个像素是一个整数，整个像素表达并不是一个“连续”（Continuous）表达。因此，在把世界上的连续信号采样到离散像素的这一过程中，难免会有失真。而不同的“分辨率”，会带来失真程度不同的像素表达。

过滤器

既然已经把图像表达成为了像素，也就是某种矩阵的形式，那么我们就可以利用线性代数等工具在这个矩阵上进行变换，从而能够对图像进行某种操作。这就是“过滤器”（Filter）的一个基本思想。

很多计算机视觉的操作本质上都是过滤器操作。除了把过滤器想成某种线性代数变换之外，更普遍的一种思路是把在矩阵上的操作想成某种函数的操作。因此，我们也可以认为过滤器是函数在某一个特定区间内的操作。

举一个最简单的过滤器的例子，就是“**移动平均**”（Moving Average）。这个过滤器的本质就是针对每一个像素点，计算它周围 9 个像素点的平均值。如果我们针对每一个像素进行这样的操作，就会得到一个新的矩阵。然后我们把这个矩阵当作新的像素表达进行视觉化，就会发现是在原有图像基础上进行了“柔化”处理。反过来，如果我们需要对某一个图像进行柔化处理，就需要对其进行“移动平均”过滤操作。

有了这个直观的例子，你一定能够想到，很多我们熟知的图像特效处理，其实都对应着某种过滤器操作。

这里，我们提及一种比较特殊的过滤处理，那就是“**卷积**”（Convolution）。这个概念我们在深度学习中经常接触到。

刚才我们说到“移动平均”这个过滤器。如果我们把图像看作是一个函数 F 在某一个区域的取值，那么，“移动平均”这个过滤器是针对函数在某一点的取值（也就是某一个像素的取值），通过利用同样的函数 F 在周围的取值，从而得到一个新的计算值。

那卷积操作的思想是怎样的呢？卷积是针对 F 在某一个点的取值，除了需要利用 F 在周围的点以外，还需要利用另外一个函数，这里称作是 H 的取值。也就是说，我们要利用 H 来针对 F 进行操作。

边界探测

除了通过过滤器对图像进行简单操作之外，还有一些图像的基本操作蕴含了计算机视觉的基本原理。我们这里也稍微做一些介绍。

例如我们通常需要了解图像的边界。有研究表明，图像的边界对于人类认知图像的内涵有着特殊的意义。因此从一个完整的图像中找到不同物体的边界是一个很有现实意义的任务，并且通常被称作是“**边界探测**”（Edge Detection）。

那么，怎么来认识图像中的物体边界呢？我们先从直观上来想一想，在图像中，“边界”都有什么特征？一般来说，如果我们遇到了色彩、景深、照明的突然变化，或者是其他某种图像特质上的突然变化，我们就有可能遇到了边界。现在的问题是，在像素或者函数表达的情况下，如何来描述和检测这些“突然变化”？

在数学分析中我们学过，描述函数值变化的概念叫“导数”或者“梯度”（Gradient）。梯度大小（Gradient Magnitude）和梯度方向（Gradient Direction）都包含了函数变化的重要信息。

虽然梯度从数学的角度来说刻画了函数的变化，但是这对于设计一个实际的边界探测器依然是不够的。一个好的边界探测器需要真正能够探测到边界，也就是要尽可能少地出现错误，能够对边界进行定位（Location）的同时还需要尽量使边界平滑和链接。

在真实的边界探测中，我们往往先让图像变得更加平滑，比如利用“**高斯柔化**”（Gaussian Blur）的办法，然后在这个基础上计算**梯度大小和方向**，有了这些之后再进行一些后期处理。

小结

今天我和分享了计算机视觉的一些基本数学模型和操作。

一起来回顾下要点：第一，像素是对真实世界的采样，我们分别讲了对黑白、灰度和彩色图像的像素表达；第二，我们讲了在像素表达之上的过滤器，介绍了移动平均和卷积两种方式；第三，我们简要提及了利用函数的梯度计算来进行边界检测的任务。

最后，给你留一个思考题，从边界检测的任务中，如何知道检测到的像素是真正属于边界呢？

欢迎你给我留言，和我一起讨论。



AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 140 | 什么是计算机视觉？

下一篇 142 | 计算机视觉中的特征提取都有哪些？

精选留言 (1)

写留言



医用地狗精...

2018-08-27



既然边界可以是横竖斜，那么可以用个过滤器来测试两边的骤然变化。

展开 ∨