

142 | 计算机视觉中的特征提取难在哪里？

2018-08-29 洪亮劫

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 06:09 大小 2.82M



在上一次的分享中，我们聊了计算机视觉的一些最基础的操作和任务，包括像素表达和过滤器这两个视觉问题。我们还简单介绍了边界探测这个任务，了解如何从计算机视觉的角度来对这个任务进行建模。

今天，我们来看计算机视觉基础问题中的另一个核心任务，那就是**特征（Feature）提取**。

特征提取的目的

在深入讨论特征提取之前，我们先来了解一下特征提取的目的，或者说是研究特征提取的必要性。

从大的方面来说，计算机视觉的一部分任务是实现对图像的智能理解。那么，理解图像的语义就是其中一个非常重要的任务。

我们提到的边界检测或者是颜色检测，虽然都是理解图像的任务，但是这些任务并不理解图像中具体的物体，比如哪里是一只动物、哪里是行人等。那么，怎样才能理解图像中具体的物体呢？或者更进一步，整个图像表达了怎样的社会关系或者说是场景关系？例如一张会议室的图像，我们不仅关心里面的陈设和人物，还关心会议室的整体气氛，以及这样的气氛是不是传递出了更复杂的人物之间的社会关系。

那么，如何实现这种更高维度的语义理解呢？这往往需要我们对**底层的一些图像先进行抽象，然后再在抽象出来的特征基础上，进一步来建模。**

除了我们这里提到的对图像本身的理解以外，在很多任务中，我们还需要对图像和其他信息结合起来进行理解和分析。一种常见的形式是图像和一段文字结合起来，对某一个物品或者某一个事件进行描述。例如电商网站的商品信息，一般都有精美的图片和详细的介绍，这些信息组合起来完整地描述了整个商品的信息。

这个时候，我们就要同时理解图像和文字信息。很明显，在这样的任务中，一种比较容易想到的模式是先从图像和文字中分别抽取一定的抽象特征，有了高度概括的图像特征和文字特征之后，我们再在这个基础上进行建模。

从比较小的计算机视觉的任务来说，很多时候，一个任务会涉及到**两个步骤**：把任务抽象为提取特征，然后转换为一个普通的机器学习任务。这个流程的第二步可以是一个监督学习任务，例如回归或者分类；也可以是一个非监督学习任务。需要注意的是，我们这里提到的两个步骤，并不一定是绝对地把建模过程当作两个完全独立的步骤，而是从逻辑上对这两个步骤进行区分。事实上，**在现代的深度学习架构中，这两个步骤往往都在统一的一个架构下进行训练，从而能够得到更好的效果。**

今天，我们就从传统的计算机视觉的角度，来看看特征提取有哪些难点和经典方法。

特征提取的难点及基本思路

图像数据的特征提取为什么有难度呢？原因在于图像信息本身的复杂性。

试想我们有两张人民大会堂的建筑物照片，一张是从地面拍摄的，一张是从空中拍摄的。虽然这两张照片可能在角度、色彩、位置等方面有很多的不同，但是因为这两张照片本身所描

述的对象是一致的，都是人民大会堂，因此我们希望从这两个图片中提取的特征有一些相似性。也就是说，我们需要**找到在诸多变化因素中不变的成分**。

一个经典的思路是从**局部信息**（Local Information）入手，从图像中提取相应的特征。从实际的效果来看，局部特征（Local Feature）比全局特征更加稳固。

回到上面的例子，如何构造一个能够匹配两个图片的普遍的局部特征呢？过程如下：第一，找到一组关键的点或者是像素；第二，在关键点周围定义一个区域；第三，抽取并且归一化这个区域；第四，从归一化后的区域提取“**局部描述子**”（Local Descriptor）。得到局部描述子之后，我们就可以利用它来进行匹配了。

从上面这个流程来看，整体的思路其实就是希望**从局部找到具有代表性的特征，然后把所有因为各种因素造成的特征变化归一化掉**。

当然，这个简单的流程是有一些问题的。比如，如果我们针对两幅不同的图像分别进行上述的流程，那么很有可能最后得到的关键点和局部描述子都不一样。所以我们需要一种更具普适性的方法。

其实从 70 年代开始，就有一大部分的计算机视觉工作是在研究**如何构建局部特征描述子**。在这 30 多年的发展历程中，很多研究工作者提出了不少既有理论基础又有实用价值的特征提取方法。甚至是最近的深度学习热潮，从某种程度上来说也是一个重要的特征提取成果。

在这些研究成果中，比较有代表性的局部描述子包括**SIFT（Scale-invariant feature transform）描述子**和**HOG（Histogram of oriented gradient）描述子**。关于这两个描述子，我在这里不展开介绍它们的细节，因为在深度学习浪潮中，大部分利用描述子来对特征进行提取的方法都被淘汰了，但是这些方法的思路，我们在很多类似的工作中依然可以借鉴。所以，如果你有兴趣继续了解，可以阅读我在文末提供的两个参考文献。

小结

今天我为你讲了计算机视觉中的又一个核心任务：特征提取。帮助你对计算机视觉的一些基本特征提取有一个了解。

一起来回顾下要点：第一，在图像理解任务中，高维度的语义理解，以及理解图像和其他信息的组合形式，都需要特征提取这个关键步骤；第二，特征提取的难点在于图像信息本身的复杂性，一个经典的思路是从局部信息入手，提取局部特征。

最后，给你留一个思考题，除了图像数据以外，还有没有其他形式的数据，也需要我们先对数据的不同形态（例如图像中的颜色、方位、角度等）进行处理，从而识别出相同的物体或者数据个体？

欢迎你给我留言，和我一起讨论。

参考文献

1. David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.
2. N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In CVPR, pages 886-893, 2005.



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 141 | 掌握计算机视觉任务的基础模型和操作

下一篇 143 | 基于深度学习的计算机视觉技术（一）：深度神经网络入门

精选留言 (1)

写留言



咳咳哒

2018-09-30

学习了

展开

