

光流估计入门

Dezeming Family

2023 年 3 月 13 日

DezemingFamily 系列文章和电子书**全部都有免费公开的电子版**，可以很方便地进行修改和重新发布。如果您获得了 DezemingFamily 的系列电子书，可以从我们的网站 [<https://dezeming.top/>] 找到最新的版本。对文章的内容建议和出现的错误也欢迎在网站留言。

目录

一 介绍	1
二 稠密光流估计法	1
三 稀疏光流估计法	2
四 深度学习光流估计法	2
五 入门小结	2
参考文献	2

一 介绍

光流的概念是 Gibson 在 1950 年首先提出来的 [1, 2]。它是空间运动物体在观察成像平面上的像素运动的瞬时速度，是利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性来找到上一帧跟当前帧之间存在的对应关系，从而计算出相邻帧之间物体的运动信息的一种方法。一般而言，光流是由于场景中前景目标本身的移动、相机的运动，或者两者的共同运动所产生的。

一般来说，光流估计有下面几种假设：

- 亮度恒定，三维空间中同一点随着时间的变化，其亮度不会发生改变。这是基本光流法的假定（所有光流法变种都必须满足），用于得到光流法基本方程。
- 小运动，即时间的变化不会引起位置的剧烈变化，这样灰度才能对位置求偏导（换句话说，小运动情况下我们才能用前后帧之间单位位置变化引起的灰度变化去近似灰度对位置的偏导数），这也是光流法不可或缺的假设。
- 空间一致（Lucas-Kanade 光流法特有的假定），一个场景上邻近的点投影到图像上也是邻近点，且邻近点速度一致。注意一般的光流法基本方程约束只有一个，而要求 x, y 方向的速度会有两个未知量。我们假定特征点邻域内做相似运动，就可以连立 n 多个方程求取 x, y 方向的速度（ n 为特征点邻域总点数，包括该特征点）。

研究光流场的目的就是为从图片序列中近似得到不能直接得到的运动场 [2]。运动场其实就是物体在三维真实世界中的运动在图像上的二维速度矢量投影；光流场，是三维物体的运动在二维图像平面上（人的眼睛或者摄像头）亮度模式的投影。通俗的讲，就是通过一个图片序列，把每张图像中每个像素的运动速度和运动方向找出来，找到前后帧之间像素的对应关系，就是光流场。

直观理解是，第 t 帧的时候 A 点的位置是 (x_1, y_1) ，那么我们在第 $t + 1$ 帧的时候 A 点假如位置是 (x_2, y_2) ，那么我们就可以确定 A 点的运动了：

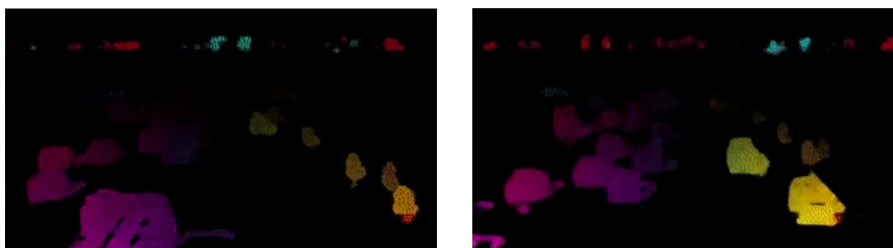
$$(u_1, u_2) = (x_2, y_2) - (x_1, y_1) \tag{1.1}$$

如何得到第 $t + 1$ 帧的时候 A 点的位置，这就需要依赖光流估计方法了。

光流估计可以笼统地分为两大类，一是稠密光流估计，二是稀疏光流估计。

二 稠密光流估计法

稠密光流估计，比如下图，会跟踪多个点集或者块的移动：



传统算法主要有：

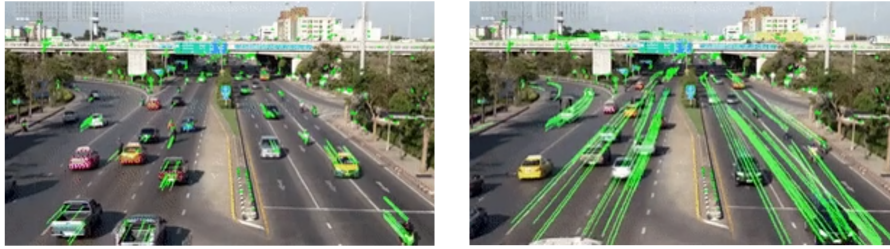
- Horn-Schunck 方法：Horn 和 Schunck 于 1981 年提出。此方法是首次使用亮度恒定假设和推导出基本的亮度恒定方程的方法之一。OpenCV 中的 `cvCalcOpticalFlowHS` 即为该方法的实现。
- 块匹配方法，对像素的集合进行处理而非单个像素，返回的“速度图像”通常比输入图像分辨率低。OpenCV 中的 `cvCalcOpticalFlowBM` 即为该方法的实现。

这两种光流方法现在并不常用，而且它们不支持图像金字塔匹配而不能用于跟踪大幅度的运动。

实际上计算稠密光流并不容易，比如一张白纸的运动，上一帧中的白色的像素在下一帧中仍然为白色，因此很难寻找该白纸上任一点的前后帧之间的对应点。

三 稀疏光流估计法

稀疏光流需要指定一组点进行跟踪：



这组点最好具有某种明显的特性，例如角点（关键点），那么跟踪就会相对稳定和可靠。稀疏跟踪的计算开销比稠密跟踪小得多。

常见的稀疏光流跟踪方法：

- Lucas-Kanade(LK) 光流法：于 1981 年提出，最初是用于求稠密光流的，由于算法易于应用在输入图像的一组点上，而成为求稀疏光流的一种重要方法。
- 金字塔 LK 光流法：LK 光流法有个“小运动”的假设，因此较大运动会将点移出小窗口，造成算法无法再找到这些点。金字塔 LK 光流法可以解决这个问题，即从金字塔的最高层（细节最少）开始向金字塔的低层（丰富的细节）进行跟踪，该方法允许小窗口捕获较大的运动。

使用 python 的 OpenCV 时，调用 `cv2.calcOpticalFlowPyrLK` 就可以进行稀疏光流估计。从该函数名即可看出它是使用了金字塔 LK 算法：Pyr 表示金字塔，LK 表示 Lucas-Kanade。

四 深度学习光流估计法

2015 年，FlowNet 的作者首先使用 CNN 解决光流估计问题，取得了较好的结果，并且在 CVPR2017 上发表改进版本 FlowNet2.0, 成为当时 State-of-the-art 的方法。截止到现在，FlowNet 和 FlowNet2.0 依然和深度学习光流估计算法中引用率最高的论文 [3]。关于深度学习方面的光流估计论文可以参考目录 [4]，给出了很详细的论文列表。

五 入门小结

我们知道，要想找到前后帧的对应点，就需要前后帧目标的特征。因此光流估计与目标检测、轮廓检测、边缘检测、特征点检测等都有很多关联。好的稳定的检测算法能够提高光流估计的准确性。一些常见的计算机视觉算法，比如图像拼接、图像配准，也会需要依赖特征点，对于一些缺少纹理的目标，人眼可能很容易看出前后帧之间的关联，但对计算机来说却不会很容易，也有不少文章提出了一些关于缺少纹理的目标的跟踪和配准功能，但也都有局限性。

本栏目下，我们不但会介绍传统的光流估计方案，还会介绍使用深度学习解决光流估计的方法。

参考文献

- [1] <https://blog.csdn.net/justremind/article/details/23662273>
- [2] https://blog.csdn.net/Sunshine_in_Moon/article/details/45797599
- [3] <https://zhuanlan.zhihu.com/p/74460341>
- [4] <https://github.com/hzwer/Awesome-Optical-Flow/blob/main/README.md>