

计算影像学第一次作业

光度立体视觉

Due Sunday, Apr. 19, at 23:59

选题 1：光度立体视觉任务

1、任务要求（共 20 分）

尝试用多种光度立体视觉算法（从基于最小二乘法和朗伯反射模型的经典算法，到基于深度学习的最新算法，具体内容参考课件 handout）计算给定数据集物体的表面法线，编程语言不限。

附件中提供 5 组数据，每组数据中包含 12 张不同光照的图像和 1 张 mask 图像，对应的光照信息在 **lights.txt** 文件中。mask 表明图像中哪些像素在物体上。按照下面任务要求的描述，用几种不同的方法得到物体的**法线图 (normal map)**，MATLAB 中使用 `imshow(uint8(N+1)*128)` 语句来展示 normal map 预测结果，如下图所示。

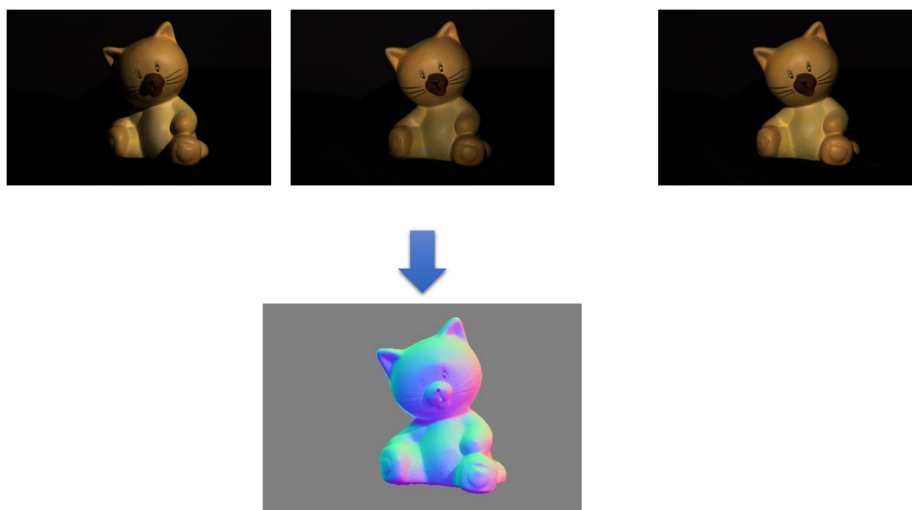


图 1: normal map

整个任务包含如下 3 部分：

1. 使用传统的朗伯反射模型实现基于最小二乘法的 Photometric Stereo, 对附件中的 5 组数据进行测试，得到 normal map。可以参考 github 上的代码，鼓励自己实现算法。（**每组结果各 2 分，共 10 分**）
2. 下载 [DiLiGenT](#) 数据集，使用任务 1 的传统朗伯反射模型的 Photometric Stereo 在该数据集的 **Main Dataset** 上进行测试，得到每组数据的 normal map，并计算你得到的 normal map 与数据集提供的 ground

truth (每组数据下的 Norma_gt.mat 或 normal.txt 文件) 之间的平均 angular error (取所有像素点 angular error 的均值), 观察结果, 对比分析在不同物体上结果有何不同。(4 分)

注: angular error 即为两个向量之间的角度误差, 首先计算两个向量内积的反余弦函数, 再转换成角度的形式。MATLAB 中可用一行代码实现:

```
angular_err=real((180.*acos(dot(vec 1, vec 2)))./pi)
```

3. 实现 Thresholding 算法。在实际应用光度立体视觉算法的时候, 我们一般会用多于 3 张图片来实验。因为当图片数量较多的时候, 我们就可以通过分析某个像素在不同光源下的 intensity, 剔除明显不符合理论模型的观测 (outlier), 而只是保留“靠谱”的观测 (inlier) 来进行计算。最简单的想法就是对于每一个像素点, 将其在不同图片中的 intensity 按照从小到大的顺序排列, 设定一个 threshold (例如在 8 比特图像中, 设定阈值为 $30 \leq B \leq 215$)。我们认为小于设定低于 threshold 下界的像素点, 其拍摄时可能处于阴影中, 认为大于设定 threshold 上界的像素点, 图像在拍摄时可能存在过曝、非朗伯 (如镜面反射) 等情况, 因此这些值都不可靠, 将被舍弃掉。而处在设定 threshold 中间的像素点被认为是更加符合理想情况的观测, 只用它们来做法向量的计算。例如: 对每个像素点, 只对留下的光源进行最小二乘法求解朗伯反射的光度立体视觉。

参考 DiLiGenT 数据集[原文](#)4.1 节中的“A position threshold method”, 编程实现 Thresholding 算法, 并在 DiLiGenT 数据集上进行测试, 对任务 2 中没有应用 threshold 的结果, 分析结果发生了怎样的变化及原因。(6 分)

2、Bonus (共 5 分)

从以下两篇论文任选其一:

1. PS-FCN¹, 官方代码[链接](#)

¹Guanying Chen, Kai Han, and Kwan-Yee K. Wong. PS-FCN: A Flexible Learning Framework for Photometric Stereo. In Proc. of the European Conference on Computer Vision (ECCV), 2018.

2. CNN-PS², 官方代码[链接](#)

完成如下任务：

阅读论文，在附件中的 5 组数据和 DiLiGenT 数据集上进行测试，提交测试结果。与任务 1 中传统朗伯反射模型的结果进行对比，找出问题，并分析传统方法和该方法的结果哪个更合理，如果后者结果比较差，阐述可能的改进方案。（提示：因为附件中的 5 组数据没有 ground truth 的 normal map，因此可以绘制一个球体的 normal map 作为参考，进行定性的对比分析。）

注：附件中的 5 组数据是 uint8 格式的，DiLiGenT 数据集是 uint16 格式的，因此在归一化数据时，前者 $\text{data}/255$ ，而后者 $\text{data}/65535$ 。

²Ikehata, Satoshi. CNN-PS: CNN-based Photometric Stereo for General Non-Convex Surfaces. In Proc. of the European Conference on Computer Vision (ECCV), 2018.

作业提交

1、提交内容

PDF 格式的实验报告，应包含以下内容：

1. 简要描述任务 1 的实现思路和方法，如果参考了网上的代码，请给出链接；
2. 任务 1 计算得到的 5 组数据的 normal map；
3. 任务 2 在 DiLiGenT 数据集上 normal map 和平均 angular error，观察结果，对比分析在不同物体上结果有何不同；
4. 任务 3 在 DiLiGenT 数据集上 normal map 和平均 angular error，对比任务 2 中没有使用 threshold 的结果，观察是否有改进，分析结果的变化及原因；
5. **Bonus** 部分的 normal map 和平均 angular error，对比传统朗伯反射模型的结果，尝试分析问题并提出改进方法（选做）；
6. 任务 1 以及任务 3 的源代码，放在附件中即可，无需写入正文。

2、提交方式

请于 **2020 年 4 月 19 号 23:59** 之前，将以上内容打包发送到

wce@ pku.edu.cn

文件名和邮件名统一为：

学号 + 姓名 + 手机号 + 计算影像学第一次作业