

# 2022.06.12周报

纪新龙

## 本周计划任务

选取了AR、SFS、重光照方向和光照估计相关的论文，下周继续阅读

## 具体完成情况

阅读了一篇硕士毕业论文

王辰辰. 《基于室内光照估计的增强现实技术》. 硕士, 南京大学, 2021.

<https://doi.org/10.27235/d.cnki.gnjiu.2021.001862>.

## 下周计划任务

继续阅读

## 具体完成情况

这里记录论文里的关键内容。

### 光照估计问题的分类

#### 根据场景分类：室内vs室外

人们根据不同场景将光照估计问题分为室外光照估计和室内光照估计。太阳提供了室外场景中绝大部分的光源能量，因此室外的光照估计问题可以转化为对太阳方向和强度的预测问题 [7–10]。而室内的光照来源相对复杂，透过窗户的阳光、屋顶的日光灯、桌面上的台灯、甚至场景中一些高反射的物体都能成为光源，因此室内光照预测的难度比室外更高。

#### 根据光照的表示方式分类

## 1. 参数化表示

目前主流方法使用球面谐波 [17, 35, 46, 47] 和球面高斯 [39, 48, 49] 对光照信息进行编码，少数工作使用小波近似的方法 [50]。本节主要介绍球谐函数对光照的编码过程。

## 2. 不对光照进行参数化，直接用全景光照贴图来表示光照。

# 球谐函数拟合光照

Sloan 等人 [51] 最早提出基于球谐函数的光照编码方法，把场景的光照信息 存储在球面上离散的位置中。和傅里叶变换十分相似，球谐光照使用多个正交的球面基向量和对应的频域系数来拟合光照，根据频域系数计算出球面上任意位置的光照辐射强度。

虽然使用高阶球谐函数能无限拟合任意的光照贴图。但是在高动态范围 (High-Dynamic Range, HDR) 图像中，像素的最大值和最小值往往差距特别大，在图像中经常出现像素值急剧变化的情况。此时球谐光照的拟合结果会出现振铃现象 (Ringing Artifacts)，如图 2-2中所示，出现与原图不符的黑洞区域。**因此 在实际渲染中，球面谐波往往适用于拟合环境中的低频光照信息。**

原图



球谐函数拟合图

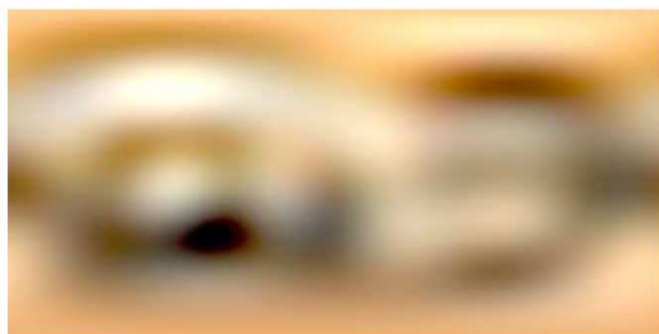


图 2-2: 球谐函数拟合光照中的振铃现象

## 球谐函数的问题

通常情况下，球面谐波更加适用于拟合光照中的低频信息。只能通过提高球面谐波的阶数来拟合高频光照信息。但是随着球谐函数阶数的增加同样会出现图 2-2中展示的振铃效应，导致重构出的环境图极度不稳定。

# 室内光照估计

从一个限制视角的室内二维图片中推断环境的全景光照信息是一个过去几十年一直研究的经典视觉问题。

## 早期研究

早期研究中主要**依赖几何的先验知识**来预测场景中的光照信息，例如

- Barron 等人 [22] 使用深度传感器获得图片中对应的几何信息。
  - 部分工作 [12, 23–26] 输入多视角图片进行三维重建，试图恢复场景中可见的三维信息。
  - Karsch 等人 [16] 需要在图片中手动标记几何信息。
  - Karsch 等人 [5] 则 另辟蹊径，将光照预测问题转为图片匹配问题。建立一个存储室内全景图的数据库，从数据库中找出与输入图最相似的全景图作为预测出的光照环境贴图。
- 以上方法不仅需要额外输入信息，而且只能在某些特定的场景中应用。

## 不依赖几何先验的深度学习方法

随着深度学习在越来越多视觉任务中的出色表现，人们开始根据不同的光照表示方法提出基于机器学习的光照预测方法。

- 直接预测环境光照贴图的方法（端到端）

- Gardner17

Gardner 等人 [27] (Gardner17) 提出了一种**端到端**的卷积神经网络，根据一张 室内限定视角 (Field-of-View, FoV) 图片预测出对应的全局光照贴图，而且不需要输入额外的场景几何信息和材质信息。

Gardner17预测出的光源位置、强度都与实际相差过大。

没有考虑深度信息，意味着绘制不出物体在场景中移动时光影的变化。

- Neural Illumination

Neural Illu预测了输入图的法向和深度信息。

Neural Illu 通过输入像素坐标解决了 Gardner17 中光照贴图不能随着物体的 移动而变化的问题。但是每当物体 移动时，Neural Illu 都需要重新输入像素坐标。相近位置预测出的光照贴图可 能出现极大的偏差，导致光照随着虚拟物体移动发生突变，降低渲染的真实感。

- Lighthouse

Lighthouse比Neural Illu多解决了光照突变问题。

要求输入相近基线上的 图片对 (narrow-baseline stereo image pair)、对应的深度信息和相机外参。根据输 入图片和深度信息进行双目重建 [34]，得到场景中已知的几何信息。后面的做法和 Neural Illu 相似，通过一个 GAN 网络对立方体补全。最后只需要输入坐标就可以通过光线追踪得到对应的 光照贴图。

**Lighthouse 在 InteriorNet [45] 上有十分优异的表现，预测出和真实场景相似 的光照结果。同时解决了光照随着坐标移动而突变的问题。然而 Lighthouse 比 其他方法需要输入更多信息。**

- 预测环境光照参数的方法（非端到端）

直接预测光照贴图最直观的缺陷是网络输出的参数量过多，导致网络难以 建立局部图像和全局光照之间的联系。为了解决这个 问题，人们通过球谐函数或球面高斯来拟合光照。

- Garon 等人

提出使用 6 阶球谐函数 (Spherical Harmonic, SH) 来拟合环境 光照。

6 阶球面谐波很难拟合出场景中的高频信息，因此预测出的光照贴图鲜有高光出现。

- Gardner19

不同于球谐函数，球面高斯能够很好的拟合光照贴图中的高频信息。Gardner等人提出，使用**三个球面高斯**拟合场景中的高频光照，统一的颜色表示剩余的低频光照。

尽管 Gardner19 对场景中高频光照的拟合效果优于球面谐波，但是 Gardner19 中球面高斯可能出现在球面上任意位置，因此网络

- 本文提出的MSGLight

球面上均匀放置**128个球面高斯**，固定它们的带宽（远离高斯中心时亮度衰减的速度）和位置，仅让它们的振幅（即中心亮度）可以被调节。

## MSGLight神经网络

在固定球面高斯位置和带宽的 MSGLight 中，室内的光照预测任务转换成了对  $N$  ( $=128$ ) 个球面高斯 (RGB) 振幅参数的预测，这显然比预测不定位置球面高斯更加简单稳定。

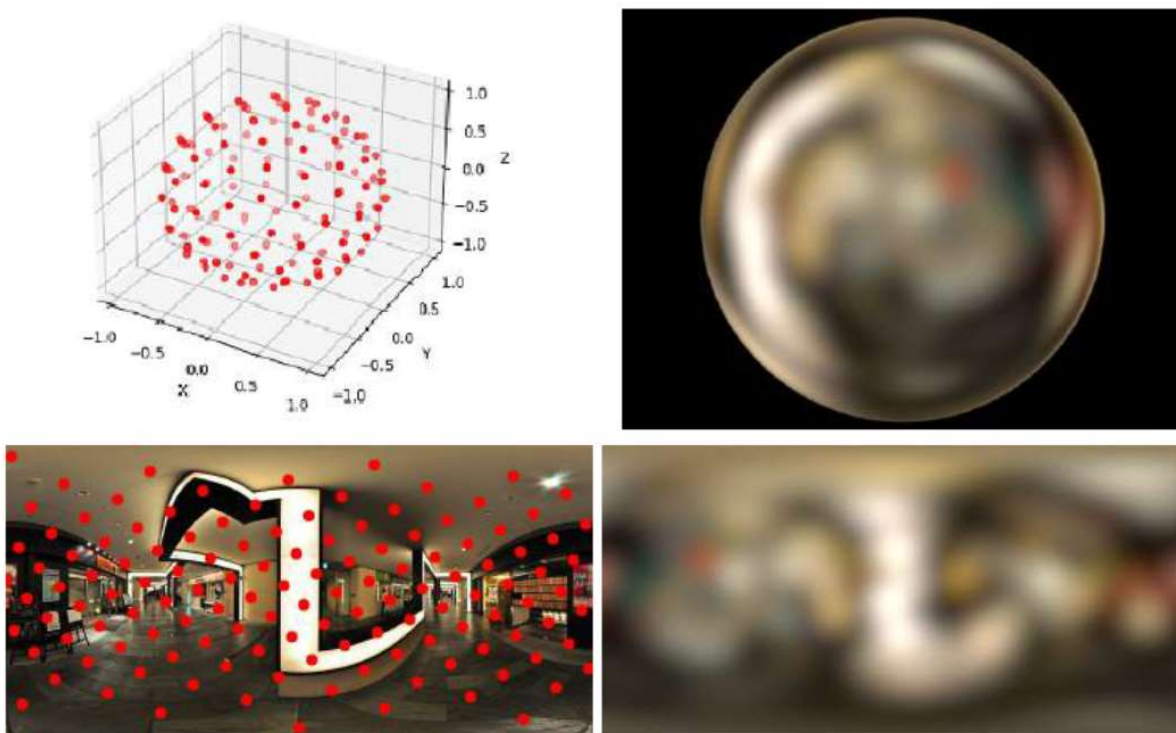


图 3-2: 左列分别表示 MSGLight 在三维和二维上的可视化结果，右列为使用 MSGLight 拟合全局光照信息的可视化结果

## MSGLight数据集的制作

从全景图中采样室内限定视角 (Field-of-View, FoV) 图片，作为输入。使用非负最小二乘法优化出对应的 MSGLight 振幅权重参数，作为输出。

## 损失函数

我们对球面高斯的振幅计算带掩码的 L2 损失并传播，同时使用绘制损失函数 (Render Loss) 来提高网络预测光照贴图在不同材质物体上的绘制效果，将上述两个损失函数混合使用作为网络完整的损失函数。如果使用传统的 L2 函数计算 MSGLight 的损失，网络倾向于预测十分平滑的光照效果，丢失了全景图中绝大部分的高频光信息。为了解决过度平滑的问题，我们只对光照强度前 20% 的球面高斯振幅参数计算 L2 损失。

## 同时预测深度

我们参考 MSGLight 对于全局光照信息编码的方式，对全局深度信息同样使用 128 个球面高斯进行编码。和光照信息中 RGB 三通道都需要编码不同，在深度图中只需要编码深度信息。

预测 MSGLight 每个球面高斯深度的方法，来解决虚实融合中的光照局部性问题。如果我们把球面高斯到原点的距离设为场景中对应区域的深度，则球面高斯可以看成有准确坐标的点光源。随着模型在场景中的移动，每个球面高斯相对于模型的方向向量可以通过彼此的坐标实时计算，因此模型产生的阴影方向可以随着模型的移动正确的变化，同时可以计算每个光源到物体处的光源衰减程度。

## 不同光照参数化方法示意图

下面这张图展示了“预测环境光照参数的方法”中球面谐波函数、球面高斯函数的预测效果。

输入都是一张Fov图，却能预测出全景的光照图。3SG表示使用三个球面高斯，这三个球面高斯位置可变，由神经网络预测。5SG表示使用5个球面高斯，这三个球面高斯位置可变，由神经网络预测。对应 Gardner19。

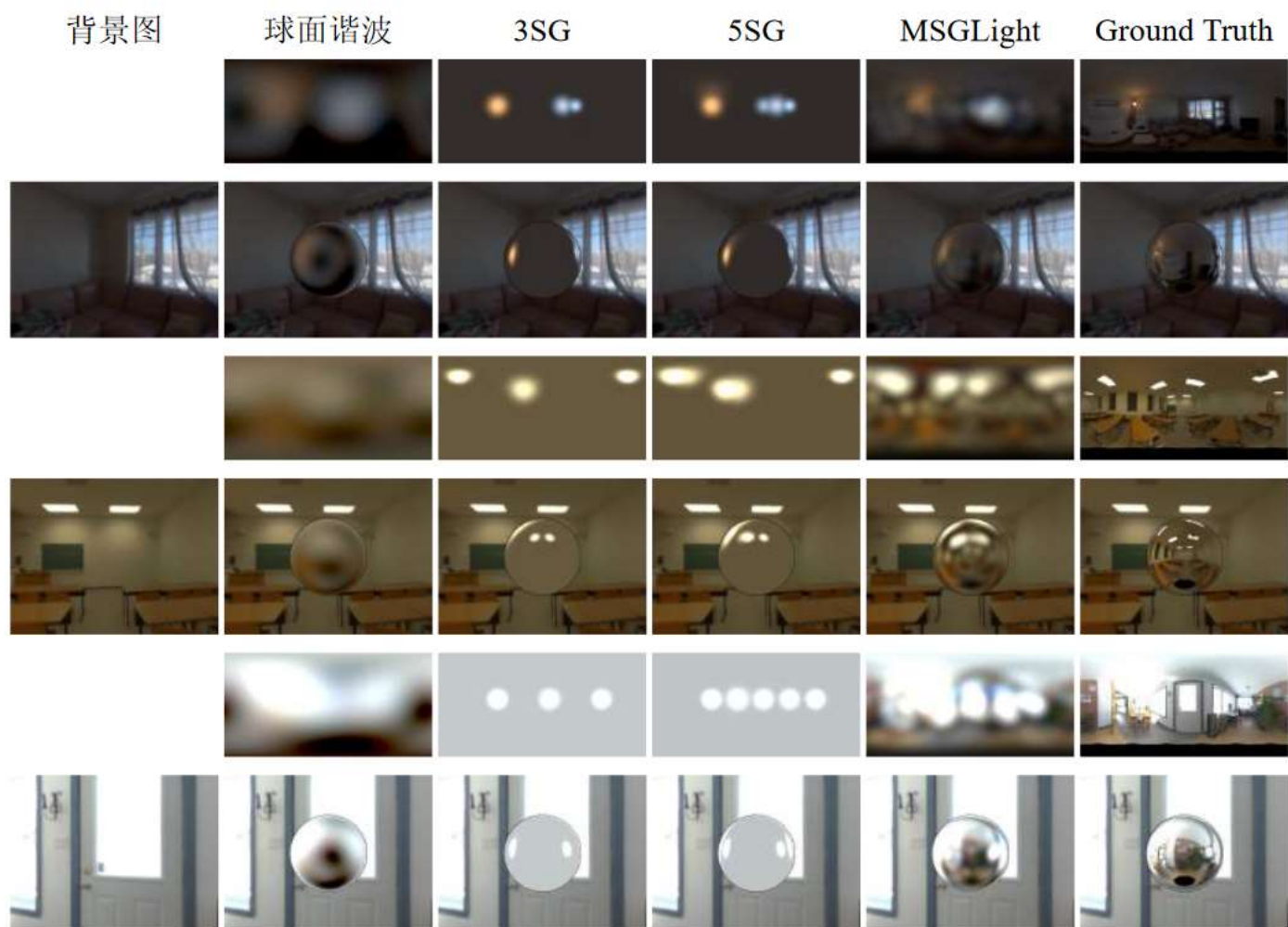


图 4-4: 不同方法对光照贴图编码效果的可视化对比