



中国科学院大学
University of Chinese Academy of Sciences

硕士学位论文

基于深度学习的场景光照估计研究

作者姓名：程大川

指导教师：陈彦云 研究员

中国科学院软件研究所

学位类别：工学硕士

学科专业：计算机图形学

培养单位：中国科学院软件研究所

2019 年 6 月

Deep Scene Illumination Estimation

**A thesis submitted to the
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Master of Engineering
in Computer Graphics**

By

CHENG Dachuan

Supervisor: Professor Chen Yanyun

Institute of Software, Chinese Academy of Sciences

June, 2019

中国科学院大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日 期：

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院大学有关保存和使用学位论文的规定，即中国科学院大学有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

摘要

关键词： 关键词

Abstract

Keywords: keywords

目 录

第1章 绪论	1
1.1 选题的背景及意义	1
1.2 国内外研究现状	2
1.2.1 光照的表示	2
1.2.2 光照估计问题的提出	3
1.2.3 使用光探测物或特定装置估计光照	4
1.2.4 使用额外信息估计光照	4
1.2.5 使用约束模型估计光照	4
1.2.6 用户交互辅助估计光照	4
1.2.7 反向渲染过程	4
1.2.8 光照估计相关数据集	4
1.2.9 利用深度学习工具估计光照	4
第2章 光照估计数据集	5
2.1 引言	5
2.2 高动态范围全景图像	5
2.2.1 全景图像	5
2.2.2 高动态范围全景图像	5
2.2.3 拍摄与合成	5
2.3 光照估计数据集	5
2.4 讨论	5
2.5 本章小结	5
第3章 基于深度学习的光照估计方法	7
3.1 引言	7
3.2 相关工作	7
3.3 问题求解范围	7
3.4 光照分布的球形谐波表示	7
3.5 卷积神经网络结构	7
3.6 损失函数	7
3.7 实验结果与评估	7
3.8 深入研究光照估计网络	7
3.9 讨论	7
3.10 本章小结	7

第4章 光照估计的应用	9
4.1 引言	9
4.2 电脑端应用	9
4.3 移动端应用	9
4.4 本章小结	9
第5章 总结与展望	11
5.1 本文工作总结	11
5.2 未来工作展望	11
附录 A 附录	13
参考文献	15
作者简历及攻读学位期间发表的学术论文与研究成果	17
致谢	19

图形列表

1.1 光照估计的应用	1
1.2 光照表示模型的示例	3

表格列表

符号列表

字符

Symbol	Description	Unit
I_{ij}	identity tensor	1

算子

Symbol	Description
∇	gradient operator

缩写

CFD	Computational Fluid Dynamics
-----	------------------------------

第1章 绪论

1.1 选题的背景及意义

光照估计（又称光照分布估计）是从已知的彩色图像信息中，预测、估计、恢复出整个场景的光照分布。该问题的输入通常是若干张彩色图片，或者是一段视频，有时几何或材质信息也会作为先验知识来辅助估计光照。场景的光照分布是指场景中各个方向的光照的颜色和强度。较为常见的光照分布表示方法包括高动态范围（High Dynamic Range, HDR）全景图、球形谐波（Spherical Harmonics, SH）函数、基于物理的Sun-Sky模型等。其中精度最高、使用比较广泛的是HDR全景图像，在实时渲染领域使用较多的是光照的SH表示。基于物理的光照表示模型则多用于室外场景。

光照估计作为计算机图形学和计算机视觉的基础问题之一，有着广泛的实际应用场景。例如：基于图像的渲染（Image Based Rendering, IBR）、增强现实（Augmented Reality, AR）、电影后期制作、真实感虚实交互等。图1.1展示了光照估计的应用之一。光照估计也与学科中的许多其它问题息息相关。例如：双向反射分布函数（BRDF）估计、场景几何重构、本征信息提取、图像增强，等等。而高质量的光照估计结果，通常能够为这些问题的解决带来很大的帮助。



图 1.1 光照估计的应用之一。使用单张图片估计场景的光照，并利用估计的光照渲染一个新的物体合成到图像中。可以看出使用估计光照渲染后的3D物体，与场景中的已有物体在视觉上较为一致。

Figure 1.1 3D rendering under the predicted illumination. visual effect of the 3D rendering is in line with the original image.

从有限的图像信息估计出整个场景的光照分布是一个复杂的问题。首先，图像的视野范围比较有限，例如一张视场角（FOV）为60°的照片所拍摄到的区域，在其对应的全景图中占比不足6%。此外，一幅图片是光照分布、场景几何结构、物体材质、摄像机参数等多个单位之间的复杂交互结果（公式 1.1）。

$$Image = ComplexInteraction(Light, Geometry, Material, Camera) \quad (1.1)$$

通过公式1.1可以看出,在其它三个信息未知的情况下,从图像(Image)反推出光照(Light)是一个严重的不适定(ill-posed)问题。不仅如此,在不同的条件下拍摄的彩色图像可能存在很多误差。例如图像中的过曝光/欠曝光区域、相机畸变、不正确的白平衡等。这些都会对光照估计造成一定程度的干扰,增加光照估计的难度。

研究者在光照估计问题上开展了一系列的研究。Debevec[1]首次提出可以通过拍摄不同曝光下的镜面球体,来获取一张高动态范围的全景图片。*Related limited methods.*不难看出,这些方法或依赖特殊的探针、或依赖特殊的拍摄设备、或依赖额外的辅助信息与模型假设。均具有一定的局限性。

近年来,深度学习在多种计算机视觉问题上大放异彩,用于分割、检测、标识、分类的神经网络层出不穷。一些研究者尝试将深度学习应用在光照估计问题当中。*Related learning based methods.*不过,目前已有的深度学习方法也有一定的局限性。训练一个鲁棒的神经网络往往需要大量的数据。目前用于光照估计问题的数据集比较有限,主要包括:大规模的低动态范围全景数据集(SUN360等)和中小规模特定场景的高动态范围全景数据集(Laval Indoor等)。这些数据集在规模和质量上很难同时到达训练深度神经网络的要求。

在这样的背景下,本文构建了一个用于光照估计的数据集。数据集具有一定的规模和质量,它不仅能被用来训练更加鲁棒的光照估计网络,也可以被应用到其它多种相关的深度学习问题当中。在已有数据集和本文构建的数据集基础上,本文深入探索了基于深度学习网络的光照估计方法,对网络结构,损失函数,光照表示等多个模块进行了细致的研究。

1.2 国内外研究现状

1.2.1 光照的表示

光照的表示对光照估计问题至关重要。场景的光照分布有着多种多样的表示方法。其中高动态范围(High Dynamic Range,简称HDR)全景图是一种被广泛使用而且精度较高的光照表示方式。与普通的8位三通道图像不同,HDR图像的颜色值范围可以从0取到非常大。这意味着HDR图像可以更细致地表示每个像素位置的真实亮度值。因此无论是HDR全景图表示的光照,还是使用其渲染的物体,都能更接近真实值。Reinhard在*High Dynamic Range Imaging*[2]中对HDR图像进行了详细的介绍。

研究者们也提出了一些小巧、高效的光照近似模型，代价是牺牲一部分精度。Ramamoorthi[3]提出使用球形谐波函数来表示场景的光照。这种方法表示的光照，不仅参数规模较小，而且其在渲染物体时非常高效。他在文章中指出，该模型仅使用9个系数时，在漫反射物体上渲染结果的平均误差也不超过1%。不过由于球形谐波函数本身的特性，这种方式不能很好的保留光照分布的高频细节。尽管如此，它的简洁和高效使得它成为实时渲染中的重要方式之一[4, 5]。许多光照估计方法中采用了这种模型来表示场景的光照分布。图1.2展示了不同的几种光照表示示例。



图 1.2 几种不同的光照表示模型对比

Figure 1.2 Comparison of different representations of illumination

在室外场景中，光线的主要来源是太阳光、光在空气中的散射折射、以及光在地面的反射。因此基于物理的光照表示模型更加适合表示室外场景。Richard[6]提出了一个参数化的物理模型来表示天空中的发光度（Luminance）分布。之后的许多工作 [7–10] 考虑了空气散射和折射、大气浑浊度的影响，将该模型扩展为多种不同类型颜色模型。Hosek [11, 12]基于这些工作，考虑并加入了多种因素，整理出一个可调节太阳位置、大气浑浊度、地面反照率的物理参数模型。这类模型的参数量较小，精度也很可观，因此成为室外场景光照估计问题中常用的光照表示方法。

光照分布也有很多其它的表示方式。Ng[13]提出使用小波函数来近似光照分布。LeGendre[14]提出了一个实用的框架，能够使用LED灯台精确地重构出各个方向的场景光照。近期Weber[15]结合深度学习，使用自编码器（auto-encoder）对光照分布进行建模，为光照估计问题提供了新的思路。

1.2.2 光照估计问题的提出

对于光照估计来说，估计出高精度的HDR全景图通常较难。本节介绍了国内外研究中主要的光照估计方法。光照估计是一个复杂的问题，研究者们常常采用这些方法中的某种或多种组合，来实现不同的应用需求。

1.2.3 使用光探测物或特定装置估计光照

光照估计是一个复杂的问题，研究者们常常采用多种方法的组合来解决这个问题。使用光探测物（light probe）来估计光照就是这些方法之一。常见的光照探测物包括：镜面球体，漫反射球体，镜面反射/漫反射混合球体，人眼，人脸等等。Debevec[1]建立了一个基于光照模型的虚拟物体插入框架。其中所涉及光照的就是使用镜面反射球体作为光探测物来估计的。该工作首次提出可以通过拍摄不同曝光下的镜面球体，来获取一张高动态范围的全景图片。在之后的工作中，Reinhard[2]和Debevec[16]指出，使用一个漫反射球体，或者漫反射与镜面反射表面混合的球体，也可以达到同样的效果。值得一提的是，Debevec在其工作[1]中，将远距离场景、近距离场景、以及待插入的虚拟物体分离开，并假设插入的虚拟物体并不会造成远距离场景的变化。之后的虚拟物体插入、增强现实等工作都遵从了这个假设。

上述的光探测物大多是已知物体几何和表面材质、且预先放置在场景中的规则物体。还有一些特殊的“物体”也可以作为光探测物，例如，人眼和人脸。Tsumra[17]在假设人眼是规则球体的前提下，利用眼球上光线的反射，估计场景的光照。Nishino[18]分析了眼球的大致结构，并利用包含眼球的照片估计场景的光照分布。不过该方法需要较为清晰的相机，而且其文中也指出，该方法没有考虑瞳孔以及虹膜颜色等对光照估计会产生较大影响的因素。

1.2.4 使用额外信息估计光照

一些工作使用额外的输入信息或先验知识来辅助估计光照。这是大部分光照估计问题的思路之一。[]使用额外的depth信息 []使用已知的几何信息 []采用对同一幅场景拍摄多张照片的方式

使用额外的信息来辅助光照估计确实能够提升估计的效果。但是此类方法通常需要使用额外的输入设备（例如深度相机），或者繁琐的获取步骤（多次拍摄）。这难免限制了它们的应用场景。

1.2.5 使用约束模型估计光照

1.2.6 用户交互辅助估计光照

1.2.7 反向渲染过程

1.2.8 光照估计相关数据集

1.2.9 利用深度学习工具估计光照

深度学习与光照估计的结合是主流趋势之一。

第2章 光照估计数据集

2.1 引言

2.2 高动态范围全景图像

2.2.1 全景图像

2.2.2 高动态范围全景图像

2.2.3 拍摄与合成

2.3 光照估计数据集

2.4 讨论

2.5 本章小结

第3章 基于深度学习的光照估计方法

- 3.1 引言
- 3.2 相关工作
- 3.3 问题求解范围
- 3.4 光照分布的球形谐波表示
- 3.5 卷积神经网络结构
- 3.6 损失函数
- 3.7 实验结果与评估
- 3.8 深入研究光照估计网络
- 3.9 讨论
- 3.10 本章小结

第4章 光照估计的应用

- 4.1 引言
- 4.2 电脑端应用
- 4.3 移动端应用
- 4.4 本章小结

第5章 总结与展望

5.1 本文工作总结

5.2 未来工作展望

附录 A 附录

参考文献

- [1] DEBEVEC P. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography[C]//Proceedings of the 25th annual conference on Computer graphics and interactive techniques. ACM, 1998: 189-198.
- [2] REINHARD E, WARD G, PATTANAIK S, et al. High dynamic range imaging: Acquisition, display, and image-based lighting[M]. Elsevier, 2005.
- [3] RAMAMOORTHY R, HANRAHAN P. An efficient representation for irradiance environment maps[C]//Proceedings of the 28th annual conference on Computer graphics and interactive techniques. ACM, 2001: 497-500.
- [4] GREEN R. Spherical harmonic lighting: The gritty details[C]//Archives of the Game Developers Conference: volume 56. 2003: 4.
- [5] SLOAN P P. Stupid spherical harmonics (sh) tricks[C]//Game developers conference: volume 9. 2008: 42.
- [6] PEREZ R, SEALS R, MICHALSKY J. All-weather model for sky luminance distribution—preliminary configuration and validation[J]. Solar energy, 1993, 50(3):235-245.
- [7] NISHITA T, DOBASHI Y, NAKAMAE E. Display of clouds taking into account multiple anisotropic scattering and sky light[C]//Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. ACM, 1996: 379-386.
- [8] SIRAI T N T, NAKAMAE K T E. Display of the earth taking into account atmospheric scattering[C]//Siggraph: volume 93. Citeseer, 1993: 175.
- [9] PREETHAM S. A practical analytic model for daylight[J]. 1999.
- [10] RAAB M, SEIBERT D, KELLER A. Unbiased global illumination with participating media [M]//Monte Carlo and Quasi-Monte Carlo Methods 2006. Springer, 2008: 591-605.
- [11] HOSEK L, WILKIE A. An analytic model for full spectral sky-dome radiance[J]. ACM Transactions on Graphics (TOG), 2012, 31(4):95.
- [12] HOŠEKHOŠEK L, WILKIE A. Adding a solar-radiance function to the hošek-wilkie skylight model[J]. IEEE computer graphics and applications, 2013, 33(3):44-52.
- [13] NG R, RAMAMOORTHY R, HANRAHAN P. All-frequency shadows using non-linear wavelet lighting approximation[C]//ACM Transactions on Graphics (TOG): volume 22. ACM, 2003: 376-381.
- [14] LEGENDRE C, YU X, LIU D, et al. Practical multispectral lighting reproduction[J]. ACM Transactions on Graphics (TOG), 2016, 35(4):32.
- [15] WEBER H, PRÉVOST D, LALONDE J F. Learning to estimate indoor lighting from 3d objects[C]//2018 International Conference on 3D Vision (3DV). IEEE, 2018: 199-207.

- [16] DEBEVEC P, GRAHAM P, BUSCH J, et al. A single-shot light probe[C]//ACM SIGGRAPH 2012 Talks. ACM, 2012: 10.
- [17] TSUMURA N, DANG M N, MAKINO T, et al. Estimating the directions to light sources using images of eye for reconstructing 3d human face[C]//Color and Imaging Conference: volume 2003. Society for Imaging Science and Technology, 2003: 77-81.
- [18] NISHINO K, NAYAR S K. Eyes for relighting[J]. ACM Transactions on Graphics (TOG), 2004, 23(3):704-711.

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历

作者基本情况

程大川，男，汉族人，1994年11月11日生，山东菏泽人

2012年2月–2016年6月 山东科技大学信息学院软件工程工学学士

2016年9月–2019年6月 中国科学院大学软件研究所计算机应用技术工学硕士

联系方式

通讯地址：北京市海淀区中关村南四街4号中国科学院软件研究所

计算机科学国家重点实验室

邮政编码：100190

电子邮箱：chengdc@ios.ac.cn

已发表(或正式接受)的学术论文:

[1] Cheng, Dachuan, Jian Shi, Yanyun Chen, Xiaoming Deng, and Xiaopeng Zhang. "Learning Scene Illumination by Pairwise Photos from Rear and Front Mobile Cameras." In Computer Graphics Forum, vol. 37, no. 7, pp. 213-221. 2018.

致 谢

致谢

