E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP391.4;TP18　　**文献标识码**:A　　**文章编号**:1006-8961(2026)10-3482-20

论文引用格式:Lei L N, Zhu J C, Guo C L, Huang J W, Luo J and Li C Y. 2026. Deep learning methods for portrait relighting. Journal of Image and Graphics, 31(10):3482-3501(雷莉娜, 朱家成, 郭春乐, 黄杰文, 罗俊, 李重仪. 2026. 人像重打光深度学习方法研究进展. 中国图象图形学报, 31(10):3482-3501)[DOI:10.11834/jig.250303]

人像重打光深度学习方法研究进展

雷莉娜[1],朱家成[2],郭春乐[1],黄杰文[2],罗俊[2],李重仪[1*]

1. 南开大学计算机学院,天津　300350; 2. OPPO广东移动通信有限公司,东莞　523860

摘　要:人像重打光技术是计算机视觉领域中的一个重要研究课题,其核心思想是通过调整图像中的光照条件,模拟或改变人物面部的光照效果,从而实现提升图像视觉效果、改善面部特征呈现和增强美学效果的目标。近年来,随着深度学习技术的快速发展,基于深度学习的重打光技术已经成为主流。这类方法能够有效地捕捉光照变化与图像之间的复杂关系,进而输出具有高度真实性和视觉吸引力的重打光效果。本文详细介绍了人像重打光算法的研究背景和应用领域,进一步总结和归纳当前领域内已有的人像数据集和光照数据集。通过对现有基于深度学习的重打光方法进行分类调研与深入分析,本文还综合对比了这些方法的优势与不足,探索其在不同应用场景中的适用性与局限性。最后,结合当前技术发展趋势,本文展望了人像重打光算法在未来研究中的挑战与机遇,提出了可能的研究方向,旨在为进一步提升算法性能、拓宽应用范围以及克服现有技术瓶颈提供参考和思路。

关键词:人像重打光;深度学习;物理模型;图像转换;三维重建

Deep learning methods for portrait relighting

Lei Lina[1], Zhu Jiacheng[2], Guo Chunle[1], Huang Jiewen[2], Luo Jun[2], Li Chongyi[1*]

1. School of Computer Science, Nankai University, Tianjin 300350, China; 2. OPPO Guangdong Mobile Communications Co., LTD., Dongguan 523860, China

Abstract: Portrait relighting is an important research direction bridging computer vision and computer graphics. Its core objective is to simulate the lighting effects under different real-world light sources by adjusting the lighting conditions of the person in an image to optimize the visual representation of images, improve the presentation of facial features, enhance the aesthetic effect and meet the needs of artistic creation. This technology has extensive application value, particularly in scenarios such as postprocessing of portrait photography, virtual reality, augmented reality, film special effects production, digital entertainment, human-computer interaction, intelligent monitoring and video conferencing. Relighting can not only restore or enhance the quality of an image but also add a new emotional atmosphere to it, making it more visually impactful and communicative. Therefore, how to relight human image efficiently, accurately, and naturally has been a key issue that researchers at home and abroad have focused on in recent years. The traditional methods for relighting portraits, such as the spherical harmonics illumination model, the Lambertian reflectance model, the Phong model, and the mixed model of ambient light and directional light, are mostly based on physical modeling ideas. These methods attempt to depict the interrelationships among lighting, materials, and geometry by establishing clear mathematical and physical models. However,

收稿日期:2025-07-02;**修回日期**:2026-01-28;**预印本日期**:2026-02-06

*通信作者:李重仪　lichongyi25@gmail.com

基金项目:国家自然科学基金项目(62306153, U23B2011, 62176130);天津市自然科学基金项目(25ZXRGGX00290,24JCJQJC00020)

Supported by: National Natural Science Foundation of China(62306153, U23B2011, 62176130); Tianjin Municipal Natural Science Foundation(25ZXRGGX00290,24JCJQJC00020)

real-world human faces are highly complex, with variations in skin color, uneven reflection attributes, and fine and irregular geometric structures, which cause traditional physical modeling-based methods encounter challenges such as large modeling errors, high computational complexity, and poor generalization ability when applied to real images. Especially in single-image input scenarios that lack multiperspective or depth information, traditional methods often fail to generate satisfactory relighting effects. With the rapid development of deep learning technology, the image relighting method based on neural networks has gradually emerged and quickly become the mainstream research paradigm. Compared with traditional methods, deep learning methods can automatically learn the complex nonlinear mapping relationship between illumination and facial features in images through large-scale data training, thereby achieving highly realistic and detail-rich relighting effects without explicitly modeling the physical process of illumination. This type of method usually has great flexibility, expressiveness, and scalability and can adapt to various complex lighting conditions and changes in face poses, showing enhanced robustness in practical applications. To sort out the development context and research status of portrait relighting technology comprehensively, this study first starts from the research background, systematically summarizes the development process and key technological breakthroughs in this field, and clarifies its important value in various practical scenarios. This study also collates and compares the current mainstream portrait datasets and illumination datasets, covering synthetic data (such as face data generated on the basis of rendering engines) and real collected data (such as real illumination images obtained using multilight source shooting systems). These datasets provide fundamental support for the training and evaluation of algorithm models. However, problems such as insufficient data scale, imperfect illumination labels, and uneven distribution of races and genders must be addressed urgently. The method review section classifies the current portrait relighting methods based on deep learning into three major categories: 1) the first type is methods combined with physical models, which attempts to establish a connection between the powerful representational ability of neural networks and the physical interpretability of traditional illumination models. For example, by introducing learnable spherical harmonic parameters, normal graphs, and reflection models, it guides the network to learn intermediate representations with physical relevance, making the generated results more realistic and controllable. 2) The second type is methods based on image transformation, which mainly relies on image-to-image mapping networks, such as U-Net structure, generative adversarial network, and self-attention mechanism. It directly learns the mapping relationship between the input image and the target illumination image and has advantages such as end-to-end training, high efficiency, and natural visual effects. However, certain deficiencies in terms of illumination control accuracy and geometric consistency still exist. 3) The third type is methods based on 3D reconstruction, which utilizes face reconstruction technology to first perform 3D reconstruction on the characters in the input image (such as restoring 3DMM parameters, depth maps, or normal maps), then reconstruct the lighting conditions in 3D space, and finally project to generate a 2D image. This approach is usually superior to the pure image conversion method in terms of accuracy and control ability, but it has higher requirements for the quality of the input image and the accuracy of geometric reconstruction, and the computational cost is relatively large. On the basis of a detailed analysis of the above three types of methods, this study also compares and evaluates the advantages and disadvantages of different methods from multiple dimensions such as the realism of the relighting effect, the degree of detail retention, computational complexity, and the generalization ability of the model. Meanwhile, in combination with the requirements of portrait relighting in practical applications, the applicability and robustness of various methods when dealing with complex factors such as different races, postures, the number of light sources and the position of light sources are explored. Finally, considering the current development trends of deep learning and computer graphics, this study looks forward to the future research directions of portrait relighting algorithms. Future research may pay more attention to the lightweight design of the model, real-time processing capabilities, the fusion of multimodal inputs (such as combining semantic information, depth maps, etc.), the improvement of cross-domain transfer capabilities, and more realistic and natural lighting style modeling. Meanwhile, the diversity of the dataset, the accuracy of the labels, and the rationality of the evaluation indicators also become important factors restricting the development of the technology. This study hopes to provide references and inspirations for subsequent research through a systematic review and analysis of portrait relighting technology and promote the further development of this field at the theoretical and application levels. In conclusion, this study aims to provide comprehensive technical references and theoretical support for relevant researchers through a systematic review and

in-depth analysis of the existing research methods in the field of portrait relighting and to promote continuous progress in this field in terms of algorithm performance, application breadth, and visual quality.

Key words: portrait relighting; deep learning; physical model; image transformation; 3D reconstruction

0 引 言

光照是影响图像质量、情感表达和美学效果的关键因素。图像中的光照不仅决定了面部细节的表现，还直接影响人物的气质和视觉效果。在人像摄影中，专业摄影师往往需要在精心布置的环境中花费数小时、利用昂贵的专业设备才能设计出理想的光照条件，使得拍摄的照片能够充分展现人物的面部表情、肤色和细节，同时避免过曝、欠曝或阴影过重等问题。在视频通话和在线会议等场景中，光照条件往往并不理想，可能导致画面模糊、阴影重或面部细节不清晰等问题。因此，如何调整和优化图像中的光照条件，一直是图像处理领域的一个难题。

为了解决光照条件不理想、专业布光成本昂贵的问题，人像重打光(portrait relighting)技术应运而生。该技术旨在通过调整图像中的光照条件，模拟或改变人物面部的光照效果，以达到模拟不同光照场景、改善图像美学效果的目标，如图1所示。

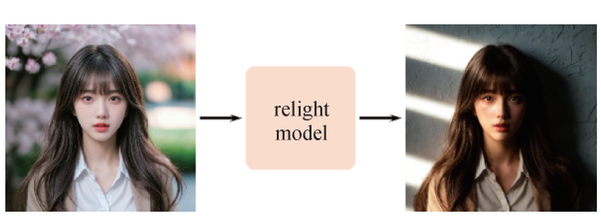


图1 人像重打光任务示例

Fig. 1 Example of portrait relighting task

人像重打光的研究起源于计算机图形学，通过建模物体表面几何形状(法线)、材质(反射率)，根据物理光照模型渲染人像在对应光照下的影像得到重打光图像。但这些方法通常无法精确控制图像中的光照分布和精确捕捉图像中的光照变化和细节，很难在复杂的光照环境下生成自然的光照效果。研究人员尝试通过多种技术手段对光照进行建模和控制(杨克义，2017；宋仪，2021；刘慧，2011)，早期的方法主要依赖于光照模型和手工设计的特征提取技术，而随着卷积神经网络(convolutional neural network，CNN)和生成对抗网络(generative adversarial network，GAN)等深度学习模型的广泛应用，人像重打光技术开始走向深度学习驱动的自动化路径，逐步克服了传统方法在光照建模方面的不足，尤其2015年后，基于深度学习的图像重打光研究得到了快速发展，研究人员通过新的模型架构和训练策略，显著提高了光照生成的质量和效率。

深度学习人像重打光算法相较于传统方法在处理复杂场景和单幅图像时具有更强的灵活性和能力，深度学习人像重打光算法可以分为结合物理模型的方法、基于图像转换的方法和基于三维重建的方法。结合物理模型的方法通常结合了深度学习与传统的物理光照模型的优势，通过估计法线、反射率等物理属性，并利用三维理解进行光照调整，从而提升精度和自然感；基于图像转换的方法通常利用生成对抗网络等生成式模型，通过从输入图像中学习光照模式，并在控制生成过程中实现对光照的编辑与转换，能够生成符合物理规律的图像效果；基于三维重建的方法通常通过对图像中的三维信息进行建模，恢复物体的几何形状和光照特征，从而实现更为精确的重打光。

本文重点介绍基于深度学习的人像重打光技术的主要研究进展，探讨当前存在的技术难题。

1 人像重打光数据集

人像数据的收集需要考虑个人隐私和版权问题，使用真实人物的照片需要得到被摄者的同意。而且在拍摄重打光数据时，很难使被摄对象长时间保持静止，被摄对象的微小动作或面部的改变都会干扰光线效果。使用高速相机可以减少拍摄对象动作带来的影响，为了使成对数据尽可能对齐，拍摄者在拍摄时不得不使用昂贵的高速相机，这就导致成本增加。为了捕捉不同的光照效果，需要精心设计光照场景，这可能涉及到复杂的布光技术和设备。基于上述难点，人像的重打光领域目前并没有质量较高的可以直接公开获取的数据集。为此，研究者们主要通过搭建专用采集设备或改进拍摄方式来构

建配对人像数据集。

1.1 配对人像数据集

现有的数据采集方法大多采用单阶段单光照(one light at a time，OLAT)方法。OLAT方法的核心是一个能够从多个方向照射被摄对象的光照舞台(Debevec等，2000)，光照舞台通常由多个光源组成，每一个光源都能被独立地打开或关闭，这样可以单独捕捉每个光源对物体表面的影响(如图2所示)。数据拍摄过程中，对每个单独的光源照明下的人像使用高速相机进行拍摄，每幅图像反映了场景在特定单一光源照明下的状态，由于高速相机拍摄时间间隔很短，因此相同姿势下每幅图像的人像可以看做是完全静止的。然而由于成本的限制，目前只有少数机构搭建了光照舞台。SunStage(Wang等，2023b)利用智能手机相机和阳光构建了一个简化版的光照舞台，通过让用户在户外拍摄旋转的自拍视频获得在不同角度的阳光下的面部数据，并重建了面部几何、反射率、相机姿态和场景照明参数。还有的工作利用普通的显示器或电视作为光源(Sengupta等，2021)，让被摄对象坐在显示器前，通过屏幕内容被动照明，然后记录显示器上的帧和被摄对象的同步网络摄像头视图，从而得到一组屏幕(照明)图像和对应的被摄对象图像。虽然这种方法更加简单，但由于人像头部的运动，它获得的人像数据对是不对齐的。

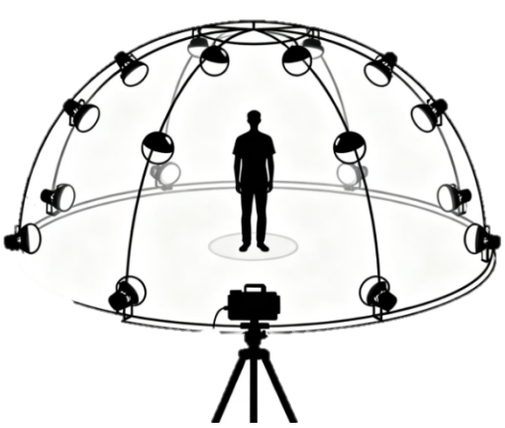


图2 光照舞台示意图

Fig. 2 Example of light stage

以下是目前针对人像重打光的一些可以获取的数据集。

1)MultiPIE数据集(Gross等，2008)。该数据集由美国卡耐基·梅隆大学于2009年发布，是一个多视点人脸图像数据集，由337名被摄者在不同姿态、光照条件和表情下拍摄的面部图像组成，姿势范围捕捉到脸部轮廓的变化，包含15个离散视图，光照条件的变化是用位于房间不同地方的19个手电筒模拟的。数据集包括共计超过750 000幅图像，需付费获取。

2)Dynamic OLAT数据集(Zhang等，2021)。由上海科技大学MARS实验室构建。该数据集是通过一个光舞台设置和一个Phantom Flex4K-GS相机(全局快门，静止的4 K超高速相机，速度为1 000帧/s)捕获的，包含了36位实验参与者共603 288幅图像。数据集需要申请并签署协议获得。

3)ICT-3DRFE(3D relightable facial expression)数据集(Stratou等，2011)。ICT-3DRFE数据集是基于光照舞台构建的，数据集包含了23名不同种族背景的被摄者(17名男性和6名女性)，年龄为22～35岁，包含9种不同的光照条件(从均匀照明到具有特定方向性的照明和模拟实际环境的光照)，每名被摄者有15种表情，数据集还提供了图像对应的漫反射颜色图、镜面强度图和表面法线图。数据集需要申请并签署协议获得。

4)Extended Yale Face Database B数据集(Georghiades等，2001)。该数据集由美国耶鲁大学于2001年发布，是一个灰度图人脸图像数据集，包含28名被摄者，在9种不同的姿势和64种不同的拍摄参数下，总共16 128幅人脸图像。数据集公开，可以直接获取。

5)DPR(deep portrait relighting)数据集(Zhou等，2019)。该数据集借助计算机图形学技术，通过软件模拟人像以及各种复杂光照条件下的场景，生成大量的训练数据。DPR数据集建立在高分辨率Celeb A数据集(Celeb A-HQ)基础上，包含Celeb A数据集中3万幅分辨率为1 024×1 024像素的人脸图像，筛选后得到27 627幅图像。Zhou等人(2019)基于比例图像的概念，通过将源图像与目标和源光照之间的比例相乘，来渲染在目标光照条件下的人脸图像，其中人像法线通过3DDFA(3D dense face alignment)(Zhu等，2019)算法估计得到，球谐光照(Ramamoorthi和Hanrahan，2001)通过SfSNet(shape from shading net)算法(Sengupta等，2018)估计得到，还提出了一种变形方法(Sorkine和Alexa，2007)，通

过最小化变形过程中的刚性偏差，提高人脸法线与图像的对齐精度。对于每一幅图像，从一个光照先验数据集中随机选择5种光照条件生成光照人脸图像，最终得到138 135幅重打光数据。然而，真实人脸和合成人脸之间存在显著的差距，特别是在皮肤反射率和精细几何细节方面。

6）虚拟光照舞台合成数据。为了弥补合成数据和真实数据之间的差距，Yeh等人（2022）提出使用基于高级计算机图形学的虚拟光舞台。数据集包含超过500个不同年龄、性别和种族的主体，每个主体都随机分配了发型、服装和配饰，渲染过程中使用了1 536幅高动态范围图像（high dynamic range imaging，HDR）环境图作为照明，共计30万个配对样本。每个样本包含同一人在相同姿势和表情下，但在两种不同光照条件下的图像对，每幅图像的分辨率为512×512像素，以HDR格式存储。

以上数据的图像分辨率、数量等信息如表1所示。除此之外，还有一些公司，如Google、Adobe、Beeble AI等也构建了自己的人像重打光数据集。这些公司多采用光照舞台的形式，但这些数据集并不公开。真实采集的光舞台数据物理可控、标注可靠，但成本高且多样性受限；合成数据规模大、可控性强，却存在与真实域在材质和几何细节上的显著域差，影响泛化。未来亟需发展低成本高质量采集与HDR/物理标注，并建立跨域统一评测基准以提升真实性与可迁移性。

表1　人像重打光数据集对比

Table 1　Portrait relighting datasets comparison

数据集	发布年份	图像数量/幅	分辨率/像素	人数	光照条件/种
Extended Yale B	2001	16 128	192×168	28	64
ICT-3DRFE	2007	–	1 296×1 944	23	9
MultiPIE	2009	750 000+	640×480	337	19
DPR	2019	138 135	1 024×1 024	27 672	5
Dynamic OLAT	2021	603 288	512×512	36	2 810
虚拟光照舞台合成数据集	2022	300 000+	512×512	500	1 536

注：“–”表示未提供相关信息。

1.2　环境地图数据集

人像重打光算法中，常使用环境地图表示光照，结合OLAT数据集将环境地图投影到光照舞台LED（light-emitting diode）灯的配置上，可生成对应环境光照下的成对人像数据。具体而言，使用光照舞台设备捕获的OLAT图像集中，每幅图像中只有一个LED灯亮起，其他灯关闭，这些图像从不同的角度捕获了人物在单一光源下的外貌。然后将环境光照图调整到低分辨率，每个像素对应一个LED灯，每个LED灯的位置在相机坐标系中都有一个对应的高斯分布。对于每个环境光照图，根据其在LED灯上的投影强度，对OLAT图像集中的相应图像进行加权，然后将这些加权图像相加以生成在该环境光照下的图像。通过这种方式，可以生成在不同环境光照下的人物图像。以下是一些常用的HDR环境地图数据集。

1）Laval Indoor数据集（Gardner等，2019）。该数据集包含2 100+幅高分辨率室内全景图，使用佳能5D Mark III和机器人全景三脚架云台拍摄。每次拍摄都是多重曝光（22档光圈）并且是全HDR，没有任何饱和度。全景图由6次捕获（方位角增量60度）拼接而成，并在各种室内环境中捕获。

2）HDRI Haven。HDRI Haven是一个专业的HDRI（high dynamic range image）贴图素材下载网站，目前收录了1 000多幅免费高清的贴图，包括室内、室外、天空、夜晚、日落和都市等7种类型、1 K/2 K/4 K/8 K/16 K 5种分辨率的HDR图像，图像可以任意使用。

3）HDRI Skies。HDRI Skies提供高质、全球形的HDR天空图，包括2 K分辨率可免费用于商业和非商业用途的图像以及20 K全分辨率可付费使用的图像，除了HDR格式，每幅图像还配有png和jpg格式。

4）HDR Maps。HDR Maps是一个3D素材资源网站，其中的Freebies板块包含100多幅免费高清的

HDRI贴图素材。图像可以在一定限制条件下免费商用。

2 基于深度学习的人像重打光方法

传统的人像重打光算法主要依赖于几何、图像处理和物理光照模型等技术。光度立体法(Woodham，1979；Woodham，1980)通过多幅图像估计物体的表面法线和形状，适用于已知光照条件下的高精度重建和重打光；Retinex理论(Land和McCann，1971；Land，1977)将图像分解为反照率和阴影部分，通过调整反照率和阴影的组合实现重打光，但对复杂光照的建模能力有限；基于物理的光照模型(Blanz和Vetter，2023；Shu等，2017)通过描述光照与物体表面的相互作用利用精确的光照和表面参数来实现重打光。这些传统方法在处理简单光照和已知光照信息时表现较好，但在真实数据上的泛化能力较差，在复杂光照场景和单幅图像条件下存在局限性。随着深度学习的发展，基于深度学习的重打光方法在处理复杂场景和单幅图像时表现出更强的能力和灵活性，图3将基于深度学习人像重打光算法进行了分类。

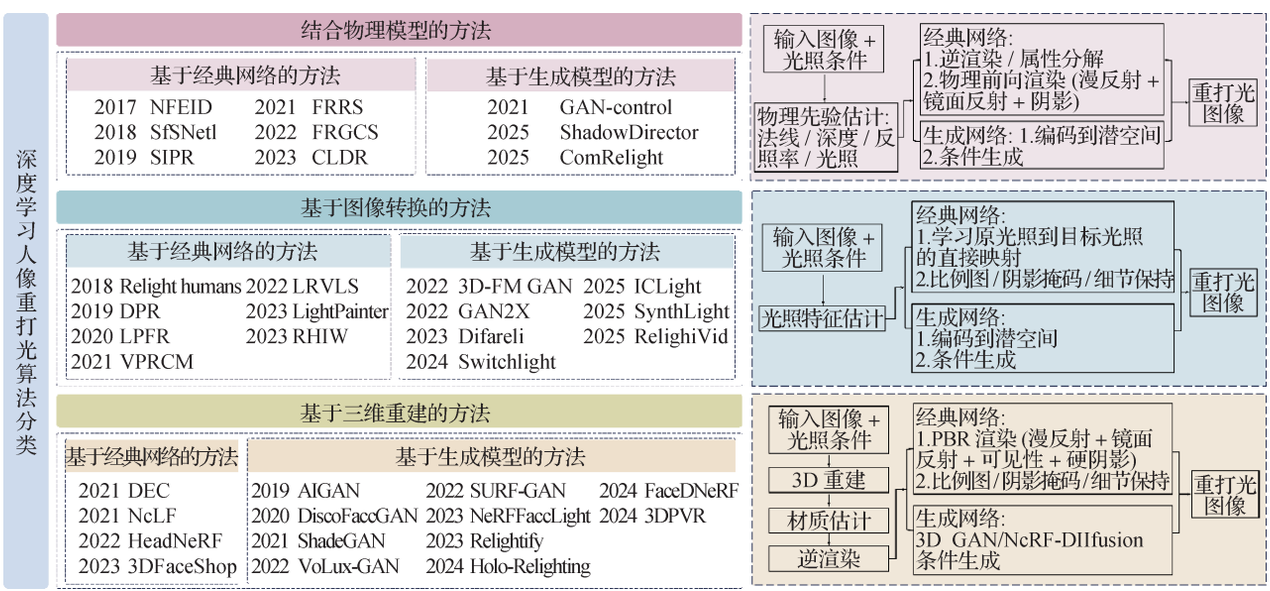


图3 深度学习人像重打光算法分类和流程图

Fig. 3 Classification and flowchart of deep learning algorithms for portrait relighting

2.1 结合物理模型的人像重打光算法

在结合物理模型的方法中，深度学习结合了传统物理光照模型的优势，通过额外的深度、反射率和法线等信息进行光照调整，算法中往往包含法线、反射率等物理性质的估计。这种方法不仅利用二维图像的光照信息，还结合了场景的三维理解，从而使得光照调整更加精确和自然。

2.1.1 基于经典网络的方法

在这类方法中，人像重打光算法通常依赖于卷积神经网络、Transformer等模型，通过对输入图像进行有效的特征提取和处理实现较为精确的光照调整。这些方法一般分为去光照阶段和重打光阶段。去光照阶段估计人像的法线、反射率和光照参数等物理性质；重打光阶段根据这些属性调整光照进行重打光。这类方法结合了深度学习的特征提取能力和传统物理模型的优势，能够较好地应对一些光照变化和物体表面反射特性，尤其是在面对大规模数据时，能够表现出较强的泛化能力。

人像重打光领域的一个主要障碍是缺乏大规模真实世界的反射率和光照的标注数据。Sengupta等人(2018)提出了一个端到端的学习框架SfSNet，利用成对的合成数据和没有真值的真实数据进行训练，通过残差块学习将图像特征分离为法线和反射率，然后结合法线、反射率特征预测光照，算法将光照表示为二阶球谐系数。推理过程中，可以通过调整球谐系数生成在新光照条件下的人像。SfSNet基

于Lambertian模型(Basri和Jacobs, 2003)建模光照,Lambertian模型是一种描述理想漫反射表面(即所有光线入射角相同的表面)反射光照的简单物理模型。Lambertian模型的反射光强度可以表示为

$$\boldsymbol{I} = K_d I_0 (\boldsymbol{N} \cdot \boldsymbol{L}) \tag{1}$$

式中,K_d表示漫反射系数(描述不同材质的物体表面对光的反射程度),I_0是入射光强度,$\boldsymbol{N}$表示表面法线的单位向量,$\boldsymbol{L}$表示光源方向单位向量的反向量,$\boldsymbol{N} \cdot \boldsymbol{L}$表示点积。

尽管Lambertian模型在许多应用中简单有效,但由于它只能建模漫反射光照,因此无法描述与视角有关的具有镜面反射特性的表面,也无法处理由强烈光源产生的硬阴影,在处理复杂的光照条件或细节时面临一定的挑战。

为了解决这一问题,Wang等人(2020)提出了一种新的框架,通过显示估计面部反射率、几何形状、高光和阴影等光照效果实现单图像人像重打光。该算法分为两个阶段:去光照阶段和重打光阶段。去光照阶段的目的是在给定输入图像的情况下,估计面部反照率、法线和面部分割;在重打光阶段,算法使用一个镜面和阴影网络显式预测镜面和阴影,然后训练一个组合网络,以估计的面部反射率、法线、镜面、阴影和目标光照的环境地图作为输入来生成重打光结果。

针对硬阴影和高光等复杂光照,Nestmeyer等人(2020)也提出了一种端到端的网络架构,与Wang等人(2020)方法不同之处是,Nestmeyer等人(2020)的重打光网络在去光照阶段预测的反射率、法线的基础上,通过预测非漫反射残差和可见性图模拟高光和阴影等复杂光照效果。

针对由强烈方向光产生的硬阴影(如鼻子投下的阴影),Hou等人(2022)提出了一种新颖的可微分算法,用于基于光线追踪原理合成硬阴影,该算法利用深度图和目标光照方向生成阴影掩膜(shadow mask)。该方法在合成阴影时直接使用面部几何信息,确保了阴影的形状和边界与面部几何结构一致。

实际应用中,往往要处理复杂的光照条件和多样化的人体表面(如皮肤、衣物、发型和配饰),由于这些表面具有复杂的反射特性,Lambertian模型不能有效地建模这些表面的光照效果。Kim等人(2024)提出了一种结合物理引导架构和自监督预训练框架的方法,称为SwitchLight。SwitchLight基于Cook-Torrance反射模型(Cook和Torrance, 1982)描述光与表面微平面的相互作用,Cook-Torrance模型基于微平面理论,认为物体表面由大量随机分布的微观平面组成,能够同时模拟镜面反射和漫反射,通过将Cook-Torrance模型集成到网络中更精确地反映了光照和物体表面的相互作用,提高了重打光的真实感。

以上人像重打光方法虽然能够实现较为逼真的光照效果,但缺乏直观的用户交互和精确的光照控制。Mei等人(2023)提出了一种基于涂鸦的交互式人像重打光方法LightPainter,允许用户通过简单的涂鸦来轻松地操纵肖像的光照效果,从而实现创意的光照编辑。LightPainter在去光照阶段可以根据肤色条件更精确地恢复几何形状和反射率,重打光模块首先从输入涂鸦和几何形状中补全阴影图,然后细化阴影并神经渲染产生最终图像。

一些方法不仅针对人像的面部,还考虑了人物的衣服、配饰、人体的凹陷区域(如腋下、胯部或衣物褶皱处)等,在以往的算法中,由于忽略了光照遮挡,这些部位往往会产生不自然的光照。Kanamori和Endo(2018)提出了一种基于卷积神经网络的方法,在球谐函数的光照表示下引入了光传输图直接推断光照遮挡,解决了衣物、凹陷区域的不自然光照问题。算法从单幅人体图像中估计反射率、光照传输图和光照信息,通过调整光照条件实现真实的重打光效果。

Tajima等人(2021)也利用光传输图预测光照遮挡,该方法分两阶段实现单图像人体重打光。第1阶段仅进行漫反射的重打光,提取反照率图、光照系数和光照传输图;第2阶段通过学习真实图像与第1阶段重建图像之间的残差来增强非漫反射,从而提高重打光的泛化能力和真实性。

2.1.2 基于生成模型的方法

相较于卷积、Transformer等经典网络结构,生成式模型能够学习高层次的图像特征和模式,进而生成更高质量和细节丰富的图像,还能够通过调整输入噪声或条件信息,生成复杂多样光照条件下的图像,产生自然且符合物理规律的效果,且这种生成过程通常是可以控制的。

基于经典网络的方法通常从显式估计面部属性(如几何形状、法线和光照)开始,然后编辑光照来重打光人像,然而物理属性估计问题的病态性使得算

法产生不自然的结果。Shu 等人(2017)提出一种端到端 GAN,将人像建模为一个紧凑的面部外观流形,对几何属性、反射率等属性增加先验来约束网络从真实人脸图像中预测人脸的解耦表示。这些表示允许语义相关的编辑,通过这种方式可以在保持其他属性不变的情况下,将目标面部的光照潜在变量替换为另一个面部的光照潜在变量,从而实现光照编辑。

针对光照编辑的可控性,Deng 等人(2020)提出了一种名为 DiscoFaceGAN 的方法,用于生成具有高度可解耦和可控的面部图像。该方法将 3D 先验知识嵌入到 GAN 中,通过训练变分自编码器(variational autoencoder,VAE)将身份、表情、姿态和光照等属性的隐变量映射到 3D 形变和渲染过程的参数空间中,从而利用 3D 形变和渲染过程的可解释性指导生成器生成具有特定属性的面部图像。为了进一步增强解耦能力,文章引入了对比学习,通过比较生成图像对之间的差异并惩罚由相同隐变量引起的外观差异,迫使生成器在最终输出中表达每个隐变量的独立影响。网络结构基于 StyleGAN(Karras 等,2019a),通过修改隐变量层并引入新的损失函数进行训练,实现了对生成面部图像光照等属性的精确控制。

Liu 等人(2022)提出了 3D-FM GAN,一个专门用于 3D 可控面部操纵的新型条件 GAN 框架,能够在端到端学习阶段后无需任何调整实现高质面部操纵。该框架通过将输入面部图像和基于物理的 3D 编辑渲染编码到 StyleGAN 的潜在空间中,实现高质量、身份保持的面部操纵。网络结构包括编码器和 StyleGAN 生成器,其中编码器将输入图像和渲染图像的信息编码到潜在空间,StyleGAN 生成器则根据这些信息生成最终的编辑图像。为了优化身份保持和编辑能力之间的权衡,文章提出了多乘法共调制架构,并采用重建训练和解耦训练策略分别增强身份保持和编辑能力。

人像重打光方法通常需要多光照条件下的成对对齐图像,但这在实际应用中往往难以获得。为此,Pan 等人(2022)提出了一种名为 GAN2X 的新方法,能够在无监督的情况下仅使用未配对图像进行训练,从而恢复 2D 图像中的 3D 形状、材质属性和光照条件。该方法采用隐式神经表面表示来建模 3D 表面和材质属性,不仅能够恢复 3D 形状,还能恢复非朗伯材质属性,如高光等。通过可微分体积渲染和 Phong 着色模型,GAN2X 能够从这些内在属性中渲染出图像。

随着扩散模型在计算机视觉应用的增多,Ponglertnapakorn 等人(2023)尝试将 Diffusion 模型应用在人像重打光中,提出了一种不依赖于准确物理属性估计的方法 DiFaReli,可以直接在 2D 图像上进行训练,而不需要 3D 人脸扫描、多视图图像或光照真值。DiFaReli 基于条件扩散模型 DDIM(denoising diffusion implicit model)(Song 等,2020)实现人像重打光,可以在单视图条件下改善人像照片的光照效果。该方法首先将输入图像编码为包含光照、形状、相机参数、面部嵌入、阴影标量和背景图像的特征向量,将这些向量作为条件来调制模型的编码和解码过程。推理过程中,可以通过修改特征向量中的光照编码,利用条件 DDIM 解码器生成重打光后的图像。这种方法能够在不依赖于准确的内在属性分解或 3D 和光照真值的情况下,实现高质量的重打光效果。随着预训练扩散模型的应用,Cai 等人(2025)认为,良好训练的扩散模型已经在其中间潜在特征中编码了阴影信息,但这些信息是隐式存在的,因此提出了 Shadow Director,只需要几千幅合成图像和数小时的训练,而不需要昂贵的真实数据,通过在推理时对扩散模型的隐空间特征优化实现对阴影的形状、位置和强度进行直观的控制,同时保持不同风格之间的艺术完整性和同一性。

针对光照舞台数据获取成本高昂且耗时的问题,Yeh 等人(2022)提出可以通过虚拟光舞台生成合成数据进行训练,并通过合成到真实数据的适应来实现高质量的重打光效果,而无需昂贵的光照舞台设备。Yeh 等人(2022)提出了一种基于虚拟光舞台和合成到真实数据适应的肖像重打光方法。算法包括 3 个主要阶段:去光照、GAN 反演和重打光。去光照阶段使用 U-Net 结构从输入图像中去除阴影和高光效果,预测反照率图像;GAN 反演阶段将反照率图像投影到预训练的 3D GAN(EG3D)的潜在空间中,以重建输入图像的 3D 信息;重打光阶段则利用重打光模块根据给定的光照条件处理 GAN 特征,生成嵌入目标光照的 3D 表示(三平面表示),并通过体积渲染生成新的图像。这种方法能够在不依赖昂贵光舞台数据的情况下,实现较高质量的肖像重打光效果。

除了将图像转换为其他多样光照条件下的图像，还有的方法致力于标准化人像面部光照，即将人像面部多样化的光照重打光为均匀的光照。Han等人（2020）提出了一种不对称联合生成对抗网络（attention-based joint generative adversarial network，AJGAN），通过使用AJGAN中的光照标准化GAN将具有非均匀光照的面部图像转换为正面光照的标准化图像，从而消除光照变化对面部外观的影响，达到标准化光照的目的。此外，该方法也能实现人像重打光，通过AJGAN中的重打光GAN可以根据不同的光照标签将正面光照的面部图像重新打光为不同光照条件下的图像。人像光照的标准化可以避免传统方法可能引入的伪影，在保持面部身份的同时实现面部图像的重打光。

由于LightStage数据有限且无法推广到视频，Wang等人（2025）设计了一个由粗到细的重打光框架，利用预训练的扩散模型作为图像先验，再结合球谐光照参数和背景图像，生成精细的重打光结果，并通过光照循环一致性从真实视频中学习时间一致性，能够处理任意身体部位（半身、全身）和动态视频。

总的来说，结合物理模型的方法实现了在复杂光照条件下较为真实的光照效果，通过深度学习端到端的训练方式简化了特征的提取和物理属性的估计，有效地处理了光照编辑问题。然而基于经典网络的方法由于对物理属性的估计是一个病态问题，不只有唯一解，因此这一过程往往容易出错，且法线、反射率等属性的真值获取比较困难，这会导致算法产生错误的光照结果。

2.2 基于图像转换的人像重打光算法

与结合物理模型的方法相比，基于图像转换的方法不依赖于场景的三维信息，不需要成对的反射率、法线等数据，而是专注于从单幅图像中提取光照特征。这种方法通常通过深度学习网络从图像中推断光源的方向、强度及其对物体表面的影响。尽管基于图像转换的方法相较于结合物理模型的方法在处理复杂光照条件时可能存在一定局限性，但它具有实现简单、计算效率高的优势，尤其适用于没有深度、法线等信息的场景。

2.2.1 基于经典网络的方法

结合物理模型的方法通常需要准确的几何和反射率模型，这些面部属性的不准确估计可能导致重打光中出现强烈的伪影，造成结果不真实，这限制了这些方法的实用性。Sun等人（2019）提出了一种基于U-Net的人像重打光系统，该网络以单幅RGB人像图像作为输入，将目标光照输入到网络的瓶颈部分，网络就能够生成在目标光照条件下的重打光图像。与传统的基于物理的渲染方法不同，该模型通过端到端的方式直接从输入图像和目标光照中学习重打光函数，而不显式地估计几何或反射率。

Zhou等人（2019）提出了DPR（deep portrait relighting）方法，首先使用基于物理的重打光方法生成大规模的人像重打光数据集（称为DPR数据集），然后使用该数据集训练网络，以生成重打光人像图像。算法采用Hourglass网络结构，包含编码器和解码器，编码器用于提取面部特征和光照特征，光照特征可以预测输入图像的光照，将预测的光照特征替换为目标光照映射的光照特征后与面部特征结合，即可通过解码器生成重打光图像。

针对处理局部人脸细节、准确去除及合成阴影（尤其是硬阴影问题），Hou等人（2021）提出使用Hourglass网络结构（Newell等，2016）估计源图像和目标图像之间的比例图像，从而在重打光过程中保持面部局部细节，网络仅修改YUV颜色空间中的Y通道，并通过阴影掩码和阴影边界权重来准确处理硬阴影和软阴影。

Futschik等人（2023）提出了一种基于学习的方法，该方法首先从输入图像中提取高光和阴影图，然后通过一个参数化扩散网络，根据用户指定的扩散参数，生成具有不同光照扩散程度的图像。这种方法的灵感来自于专业摄影师使用的柔光器和遮光布，能够在不改变整体场景光照的情况下，软化阴影和高光。通过这种方式，可以在不改变整体场景光照的情况下，软化阴影和高光，使照片更加自然和美观。

传统的光舞台需要昂贵的设备和复杂的设置，通常只能在少数实验室中使用，光照舞台数据集并不公开也限制了重打光算法的研究和应用。Sengupta等人（2021）利用日常使用的显示器作为光源，通过分析显示器上的视频内容对人脸进行重打光。针对显示器获取的数据，他们提出了一个基于U-Net的网络架构，将源图像、源光照和目标光照作为输入，结合一个光编码器估计输入和目标光（显示器图像）的潜在空间编码，并利用获得的编码调节卷积核

的权重。由于头部的运动，这种方法产生的数据是不对齐的，这会造成训练结果模糊，为了解决这一问题，算法设计了循环一致性损失和对抗损失改善模型训练。

针对视频的重打光是一个更加困难的任务，需要对视频的时间一致性建模以产生一致的重打光效果。Zhang等人（2021）提出了一种基于编码器—解码器结构的实时一致的视频肖像重打光方法，该方法通过联合建模语义、时间和光照一致性，实现对动态光照下视频肖像的自然重打光。编码器将输入图像编码为光照和肖像结构的潜在表示，解码器则将这些潜在表示解码为重打光图像。网络引入混合解耦方案，结合多任务训练和对抗训练策略，以实现语义一致性和更真实的重打光效果。

2.2.2　基于生成模型的方法

基于生成模型的方法利用生成模型如生成对抗网络、扩散模型实现图像转换。Shoshan等人（2021）提出了一种通过显式解耦潜在空间实现对生成图像的显式控制的框架，该框架通过对比学习实现潜在空间的显式解耦，使得每个子空间对应一个特定的图像属性。通过训练控制编码器将可解释的输入映射到相应的潜在子空间，从而实现对每个属性的显式控制。算法基于StyleGAN2（Rarras等，2019b）架构，将潜在空间划分为多个子空间，每个子空间对应一个特定的属性（如身份、年龄等），每个属性都由一个独立的多层感知机（multilayer perceptron，MLP）进行控制。

Zhang等人（2025）利用扩散模型进行端到端的光照编辑，他们提出了一种基于物理原理的算法，通过在训练过程中施加一致的光传输（物理原理表明，不同照明条件下物体外观的线性混合与其在混合照明下的外观一致）扩展和稳定基于扩散模型的光照编辑模型的训练。这种方法通过引入一个强物理约束来实现光传输一致性，从而在保持图像细节和固有属性不变的同时，实现精确的照明编辑。

Chaturvedi等人（2025）在物理渲染的头部合成图像上训练一个扩散模型学习不同光照条件的变换，并提出一种多任务训练策略来减轻潜在的模型分布向合成数据漂移的问题，弥补真实图像与合成图像的领域差距。

针对视频重打光，Fang等人（2025）基于预训练的2D图像重打光模型ICLight（imposing consistent light）针对不同的光照条件下对视频中的动态前景主体进行重新照明，通过加入时间层捕获帧与帧之间的时间依赖关系，在保留图像先验的同时增强时间一致性。

基于图像转换的方法不依赖场景的三维信息或复杂的反射率、法线数据，而是通过从单幅图像中提取光照特征，直接学习和生成重打光效果。虽然在复杂光照条件下，基于图像转换的方法可能存在一定局限性，但其计算效率和应用简单性使其在没有深度、法线等信息的场景中具有较大优势。随着GAN和扩散模型的应用，图像转换方法的准确率和可控性不断提升。随着技术的进步，这些方法可能会在更加复杂和多样化的应用场景中展现出更大的潜力。

2.3　基于三维重建的人像重打光算法

相比结合物理模型和图像转换的方法，基于三维重建的方法通过对场景的全面三维建模，能够更精确地描述光照与物体之间的相互作用。这些方法通常需要通过深度学习模型估计物体的三维结构以及光源的精确位置，从而使得重打光效果更加真实且符合物理规律。3D方法的优势在于其能够适应更加复杂和多变的光照条件，但这也意味着它需要更高的计算资源和更为复杂的建模过程。因此，尽管3D方法提供了更强的灵活性和表现力，它在应用时需要在精度与效率之间做出权衡。需要注意的是，基于三维重建的方法往往不是单独针对重打光任务的，还涉及到表情、视角等其他属性的编辑。

2.3.1　基于经典网络的方法

以往的方法要么使用详细的3D扫描数据作为训练数据，无法泛化到不受约束的图像；要么将表情依赖的细节建模为外观图而非几何形状，导致无法真实地进行网格重打光。Feng等人（2021）提出了一种名为DECA（detailed expression capture and animation）的模型，能够从单幅野外图像中回归出3D人脸形状和可动画化的细节，这些细节特定于个体但会随着表情变化而变化。算法引入了一种新颖的细节一致性损失，用于在训练过程中解耦个体特定细节和表情依赖的皱纹，使得可以通过控制表情参数来合成真实感的个体特定皱纹，在保持个体特定细节不变的条件下实现真实的重打光效果。

有的3D重建方法（刘晓楠 等，2024）基于NeRF

(neural radiance fields)(Mildenhall 等, 2022)进行, NeRF是基于神经网络的技术,利用卷积神经网络生成场景中的密度和颜色信息。与传统渲染依赖明确的几何和纹理不同,NeRF是通过网络学习从不同角度和不同视距的图像来重建场景,模型生成的是一个连续的体积数据而不是明确的几何对象。通过体积渲染的方式,NeRF能够生成具有高度真实感的光影效果和细节,尤其在复杂的光照和材质变化方面表现优秀。Sun 等人(2021)提出了 NeLF(neural light-transport field)方法,NeLF包括一个类似U-Net的卷积神经网络用于从输入人像图像中提取特征和预测源光照条件,以及多个多层感知机用于预测每个3D点的体积密度和光传输向量。通过光传输向量与目标环境图的点积,算法能够线性地计算出在新光照条件下的辐射度,并利用体积渲染技术渲染出新视角下的图像,从而实现高质量的重打光和视图合成效果。

三平面(Tri-plane)是一种新颖的3D表示方式,通过将3D信息分解成3个平面的表示,能够高效地压缩存储和计算,特别适用于神经网络处理。在生成、重建和处理具有复杂光照和纹理的3D模型时,Tri-plane提供了一种兼具高效和丰富表达能力的表示方式。Holo-Relighting(Mei 等, 2024)利用预训练的3D GAN(EG3D)从输入肖像中重建几何和外观特征,然后通过一个重打光模块在给定光照条件下处理这些特征,预测出一个嵌入目标光照的3D表示(三平面),最后通过体积渲染生成任意光照的图像。该方法的网络结构分为3个阶段:去阴影阶段用于预测反照率图像;GAN反演阶段将反照率图像投影到EG3D的潜在空间以重建3D信息;重打光阶段则通过一个金字塔结构的网络生成带有目标光照的三平面表示。

2.3.2 基于生成模型的方法

基于生成模型的方法近年来在三维重建和重打光领域取得了显著进展,3D GANs与生俱来的3D表示,极大地消除了从2D图像中估计3D信息进行重打光的需要。这些方法通过优化潜在空间的表示,成功地将光照、材质、姿态等属性与三维结构结合,为重打光任务提供了更加灵活且高效的解决方案,尤其是在处理复杂场景和动态环境时,生成模型展现了强大的泛化能力和表达能力,推动了重打光技术的创新与应用。

由于物理属性估计的病态性,在从2D图像中学习3D形状时常常由于形状和颜色的歧义性而无法准确捕捉3D形状。Pan 等人(2021)提出了 ShadeGAN,认为一个准确的3D形状不仅从不同视角看起来要真实,而且在不同光照条件下也应该产生真实的渲染效果,因此ShadeGAN引入了多光照约束,通过显式建模照明并在不同光照条件下进行渲染来学习更自然精确的3D形状表示。

然而,ShadeGAN方法在光照真实性方面存在不足。Tan 等人(2022)提出了 VoLux-GAN,该框架通过神经隐式场生成具有完整HDR环境光照能力的3D人脸。它从随机采样的潜在向量开始,生成每个3D位置的反照率、体积密度和反射属性,然后通过体积渲染生成低分辨率的反照率、漫反射阴影、镜面反射阴影和神经特征图,最后通过上采样和神经渲染器生成最终的重打光图像。

尽管3D感知GANs在许多方面取得了进展,但在编辑语义属性方面存在困难,且其可解释性和可控性尚未得到充分探索。为此,Kwak 等人(2022)提出了一个3D感知生成对抗网络SURF-GAN,然后将SURF-GAN的先验知识注入到2D StyleGAN中,以实现具有显式3D控制和语义属性编辑能力的人像图像合成。SURF-GAN通过在NeRF网络的每一层学习子空间来捕捉不同的语义属性,如光照、发色等,并允许通过控制参数进行操纵。3D可控StyleGAN则在StyleGAN的基础上,通过引入SURF-GAN的先验知识和潜在空间操作,显式地控制生成图像的光照语义属性,从而实现高质量的光照编辑。Hong等人(2022)也提出了一种可以编辑各种语义属性的参数化头部模型HeadNeRF,HeadNeRF是一种基于NeRF的参数化头部模型,能够实时渲染高保真度的头部图像,并支持直接控制渲染姿态和各种语义属性。HeadNeRF通过提取参考图像和目标图像的潜在编码,并替换目标图像的光照潜在编码,生成目标图像中人物在参考图像的光照条件下的重打光图像。

同样针对高质量的3D肖像生成和显式控制,Tang等人(2024)采用三平面表示来生成3D辐射场,并通过参数化的3DMM(3D morphable model)指导生成过程,实现对姿态、身份、表情和光照的显式控制,通过对比学习和体积混合策略,进一步提高属性解耦和多视角一致性。Jiang 等人(2023)提出了NeRFFaceLight方法,该方法将光照从几何和外观中

解耦出来，通过在预训练的三平面表示基础上，将光照和几何外观分别映射到不同的潜在空间，并通过条件 GAN 实现对光照的隐式控制。Zhang 等人(2023)提出了 FaceDNeRF，结合生成对抗网络和扩散模型，可以从单幅图像重建高质量的人像 NeRF，并支持语义编辑和重照明。FaceDNeRF 利用高分辨率的 3D GAN 反演和预训练的 2D 潜在扩散模型，使用户能够在零样本学习的情况下操纵和构建人像 NeRF，而无需显式的 3D 数据。通过设计的照明和身份保持损失以及多模态预训练，FaceDNeRF 为用户提供了一种仅使用单视图图像、文本提示和显式目标照明来创建和编辑人脸的 NeRF。

在三维重建中，目前已有大量关于从单幅 RGB 图像估计人脸 3D 形状和反射率的研究，但这些方法要么包含场景光照限制无法进行重打光，要么受限于模型的泛化能力。Papantoniou 等人(2023)提出了 Relightify 算法，利用扩散模型作为先验，通过在 UV 空间中进行图像修复来实现从单幅图像中重建可重打光的 3D 人脸。该方法首先使用 3D 形变模型拟合输入图像以获得粗糙的 3D 人脸形状和部分 UV 纹理，然后通过训练好的潜在扩散模型对自遮挡区域和未知的反射率成分进行修复和预测，生成完整的纹理和 BRDF 参数，从而实现逼真的 3D 人脸重建和重打光效果。

针对视频重打光，大多数现有的重打光方法耗时且往往缺乏 3D 感知能力，无法在实时条件下生成逼真且一致的重打光结果，导致视频过渡不自然，出现闪烁等现象。Cai 等人(2024)提出了基于 NeRF 的实时 3D 感知方法，通过双编码器结构分别推断每个视频帧的反照率三平面和阴影三平面，实现高效的反照率和阴影信息捕获；同时引入时间一致性网络，由两个 Transformer 模型和一个卷积神经网络组成，利用交叉注意力机制增强三平面特征的时间一致性，减少闪烁伪影，确保帧间平滑过渡，从而在保持视频数据时间一致性的同时，实现逼真且一致的重打光效果。

2.4 基于深度学习的各类人像重打光算法对比

深度学习人像重打光方法主要包括结合物理模型的方法、基于图像转换的方法和基于三维重建的方法。表 2 对本文提到的方法进行了系统总结。

表 2 基于深度学习的人像重打光方法对比

Table 2 Deep learning based protrait relighting methods comparison

年份	方法	分类	数据	网络结构	光照表示	评价指标
2017	NFEID	物理模型	合成	生成对抗网络	球谐函数	平均方差
2018	SfSNet	物理模型	3D 合成、CelebA	卷积网络	球谐函数	均值 方差 TOP-k%
	Relight humans	图像转换	合成	卷积网络	球谐函数	RMSE SSIM
2019	SIPR	物理模型	OLAT	卷积网络	环境地图	RMSE RMSE-s DSSIM
	DPR	图像转换	DPR、MultiPIE	卷积网络	球谐函数	Si-MSE Si-L2
	AJGAN	三维重建	MultiPIE、CASPEAL、Extended-edYaleB	生成对抗网络	One-Hot 编码	PSNR SSIM RMSE NMI HCD
2020	EMRCM	图像转换	OLAT	卷积网络	环境地图	RMSE PSNR SSIM
	LPFR	图像转换	OLAT	卷积网络	环境地图	LPIPS DSSIM MSD SSIM
	DiscoFaceGAN	三维重建	FFHQ	生成对抗网络	球谐函数	FID PPL

表2　基于深度学习的人像重打光方法对比(续1)

Continued Table 2　Deep learning based protrait relighting methods comparison

年份	方法	分类	数据	网络结构	光照表示	评价指标
	FRRS	物理模型	DPR、ExtendedYaleB	卷积网络	球谐函数	Si-MSE MSE DSSIM
	GAN-control	物理模型	FFHQ	生成对抗网络	光照编码	ID preservation FID
	LSED	图像转换	Desk light stage	卷积网络	光照编码	PSNR RMSE LPIPS
2021	VPRCM	图像转换	动态OLAT	卷积网络	环境地图	RMSE PSNR SSIM
	DECA	三维重建	VGGFace2、BUPTBalanced-face、VoxCeleb2	卷积网络	球谐函数	Median Mean Std
	NeLF	三维重建	FaceScape	卷积网络	环境地图	PSNR SSIM
	ShadeGAN	三维重建	CelebA、BFM	生成对抗网络	预余弦颜色场	FID SIDE MAD
	FRGCS	物理模型	CelebAHQ、MultiPIE、FFHQ	卷积网络	球谐函数 + 主导光照方向	MSE DSSIM LPIPS
	3D-FM GAN	图像转换	FFHQ、合成	生成对抗网络	球谐函数	LM FC FID
	GAN2X	图像转换	CelebA、CelebAHQ、AFHQ猫、LSUN汽车	生成对抗网络	光照方向、环境光系数和漫反射系数	CD SIE MAD
2022	LRVLS	图像转换	合成	卷积网络	环境地图	MAE MSE SSIM LPIPS
	VoLux-GAN	三维重建	CelebA	生成对抗网络	环境地图	ID相似度
	SURF-GAN	三维重建	CelebA、FFHQ	生成对抗网络	光照编码	FID Frames Sec
	HeadNeRF	三维重建	FaceSEIP、FaceScape、FFHQ	卷积网络	光照方向、环境光系数和漫反射系数	L1 PSNR SSIM
	CLDR	物理模型	OLAT	卷积网络	环境地图	MAE MSE SSIM LPIPS
	LightPainter	图像转换	OLAT	卷积网络	环境地图	LPIPS NIQE Deg PSNR SSIM
	RHIW	图像转换	商业收集	卷积网络	光照编码	RMSE SSIM LPIPS
2023	DiFaReli	图像转换	FFHQ、MultiPIE	扩散模型	球谐函数	DSSIM LPIPS MSE
	3DFaceShop	三维重建	FFHQ	卷积网络	–	FID FID-Clip
	NeRFFace-Light	三维重建	FFHQ	生成对抗网络	球谐函数	FID
	Relightify	三维重建	合成	扩散模型	UV映射	PSNR SSIM

表 2　基于深度学习的人像重打光方法对比（续 2）

Continued Table 2　Deep learning based protrait relighting method comparison

时间	方法	分类	数据	网络结构	光照表示	评价指标
2024	SwitcHlight	图像转换	OLAT	卷积网络	环境地图	MAE MSE SSIM LPIPS
	Holo-Relighting	三维重建	OLAT	生成对抗网络	环境地图	LPIPS NIQE Deg PSNR SSIM
	FaceDNeRF	三维重建	FFHQ/AFHQv2/ShapeNet	生成对抗网络 + 扩散模型	球谐函数	ID loss
	3DPVR	三维重建	INSTA	生成对抗网络	球谐函数	LPIPS DISTS Pose ID
2025	ICLight	图像转换	合成/OLAT/3D	扩散模型	环境地图	PSNR SSIM LPIPS
	ShadowDirector	物理模型	合成数据	扩散模型	隐式编码	CVTA CVS CLIPIQA++ NIMA ARNIQA
	ComRelight	物理模型	合成/OLAT	扩散模型	球谐系数/环境地图	L1 PSNR SSIM
	SynthLight	图像转换	合成数据	扩散模型	环境地图	PSNR SSIM LPIPS FaceNet
	RelightVid	图像转换	LightAtlas	扩散模型	环境地图	LPIPS PSNR SSIM

注："–"表示无相关信息。

基于图像转换的方法利用深度学习模型学习图像之间的转换关系，具有较高的实时性和灵活性，适合实时应用，但生成效果受限于训练数据，且缺乏可解释性。结合物理模型的方法通过模拟光照与物体交互过程，生成高度真实且符合物理规律的效果，对比图 4 中的两种算法可以发现，相对于基于图像转换的方法 DPR（Zhou 等，2019），结合物理模型的 FRRS（face relighting with realistic shadows）（Hou 等，2021）算法能更好地建模人像面部的阴影，重打光效果更加真实。然而结合物理模型的方法面临物理属性估计的难题，容易产生误差。基于三维重建的方法通过构建三维模型并进行光照渲染，能够精确模拟不同光照条件下的效果，生成真实感强且符合物理规律的图像，但计算开销大且建模复杂，实时性差，例如 NeRFFaceLight（Jiang 等，2023）算法对一幅图像进行重打光需要 2～3 min。

本文对现有算法进行了系统的可视化对比分析，如图 4 和图 5 所示。受限于算法开源性不足的客观条件，实验选取了具有代表性的 FRRS、DPR、FRGCS（face relighting with geometrically consistent shadows）（Hou 等，2022）、DiFaReli（Ponglertnapakorn 等，2023）、NeRFFaceLight 及 ICLight（Zhang 等，2025）等 6 种典型方法进行评测。实验结果表明，不同方法在复杂光照条件下的性能差异主要源于以下因素：

首先，在固定光源方向场景中（图 4），传统卷积架构方法（FRRS、DPR）与物理驱动方法（FRGCS）表现出光照不连续现象，其根本原因在于网络架构对高频光照细节的捕捉能力不足，当输入光照条件超出训练分布时，无法有效解耦材质反射率与光照参数的耦合关系。

其次，生成式方法（DiFaReli、NeRFFaceLight）虽然能产生连续的光照变化，但其潜在空间采样偏差导致身份特征退化问题（图 5）。这种现象可归因于生成模型在隐式表征过程中，光照条件与人脸身份信息在潜在空间存在不可忽视的维度纠缠。

值得注意的是，扩散模型 ICLight 通过显式光照条件约束，显著降低了生成式方法的身份失真风险。但其在仰视场景仍出现纹理偏移，这揭示了一个重要现象：扩散模型的生成先验会与显式物理约束产生竞争效应——当视角变化引起光照路径复杂度超过临界阈值时，条件扩散模型倾向于优先保持光照

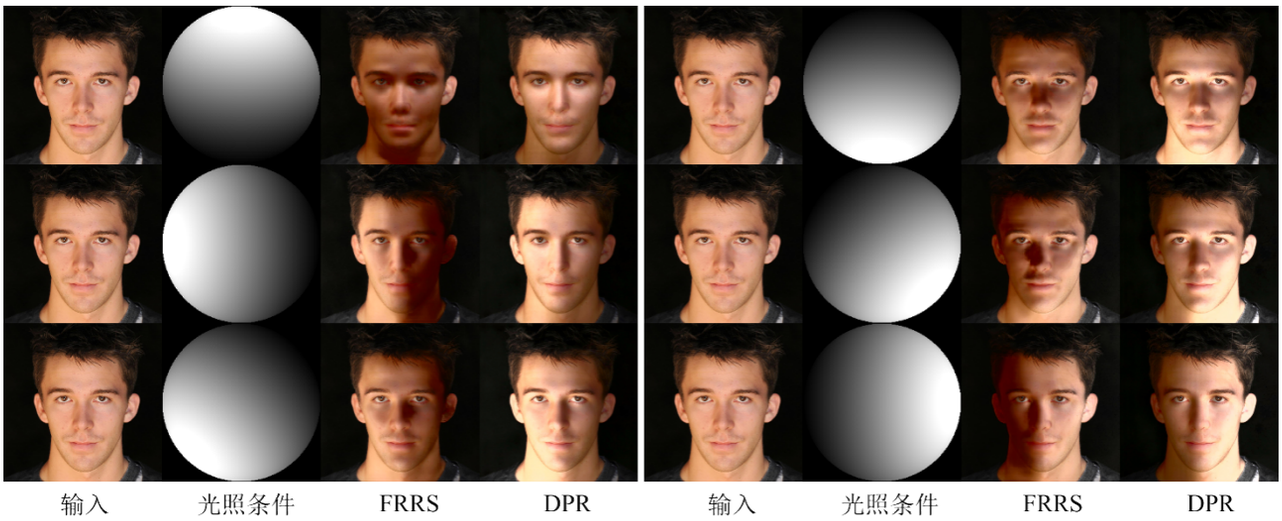


图4　FRRS(Hou等,2021)和DPR(Zhou等,2019)算法的结果

Fig. 4　The results of FRRS(Hou et al., 2021) and DPR(Zhou et al., 2019) algorithms

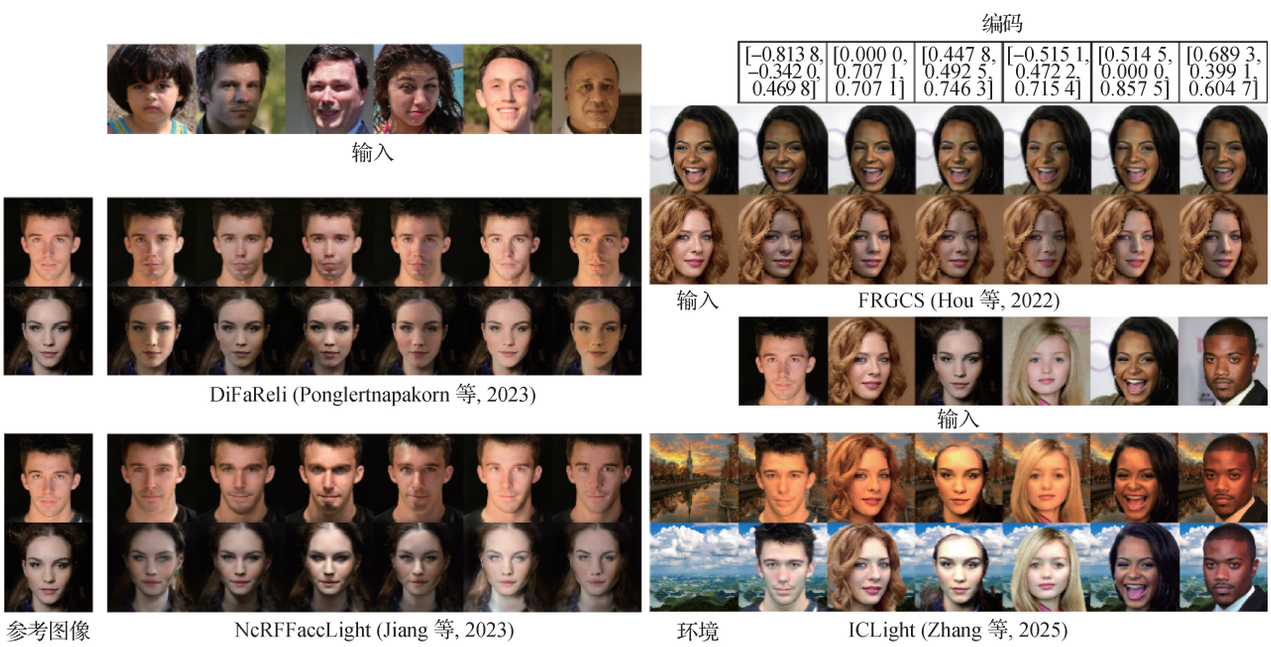


图5　DiFaReli、NeRFFaceLight、FRGCS和ICLight算法在对应光照条件(参考图像、光照编码和环境地图)下的结果

Fig. 5　The results of the DiFaReli, NeRFFaceLight, FRGCS and ICLight algorithms under corresponding lighting conditions (reference images, light codes and environment maps)

连续性而非几何一致性。

为了对不同方法的性能提供全面和整体的评估,对这些算法的效果进行了定量分析,人像重打光主要从光照一致性与真实感、伪影与清晰度、身份与细节保持、整体观感美学等方面评估:其中,NIQE(natural image quality evaluator)(Zhang等,2015)衡量自然度与失真,MUSIQ(multi-scale image quality Transformer)(Ke等,2021)预测综合感知质量/美学,CLIPIQA(CLIP-based image quality assessment)(Wang等,2023a)用CLIP先验评估“像不像好照片”的观感质量。如表3所示,DPR算法由于网络结构比较简单,且训练使用的是合成数据,因此在3个指标上表现不佳;而同样基于图像转换的ICLight算法由于应用生成能力强大的扩散模型,且训练数据包括真实数据和3D合成数据,因此达到了目前人像重打光的最佳效果。结合物理模型的FRGCS、DiFaReli和FRRS算法效果较好,但与基于三维重建的NeRFFaceLight算法还存在差距。

表 3　深度学习人像重打光算法定量分析

Table 3　Quantitative analysis of deep learning-based portrait relighting algorithm

算法	NIQE ↓	MUSIQ ↑	CLIPIQA ↑
DPR	5.27	59.24	0.47
FRGCS	5.83	61.03	0.52
DiFaReli	5.01	64.63	0.63
FRRS	6.56	61.04	0.66
ICLight	3.32	65.35	0.55
NeRFFaceLight	5.12	65.26	0.61

注：“↑”表示值越高越优。

3　挑战与展望

重打光技术已经成为当前计算机视觉和图像处理领域的研究热点。在未来的研究中，如何应对数据匮乏、复杂场景、光照控制和实时处理等方面的挑战，将是推动该领域发展的关键。

1）数据匮乏问题。高质量成对对齐数据集是重打光算法研究的基础。数据集需要包含多样化的人脸特征（如不同的年龄、性别、种族等），以及在不同光照条件下的图像，并附有准确的光照标注信息，以便于算法学习和验证。然而，虽然目前有的机构拍摄了光照舞台数据，但因为隐私、成本等问题，许多高质量数据集尚未开放，这限制了研究人员的进一步探索。合成数据虽然能够提供丰富的光照变化和精确的标注信息，但与真实场景可能存在差距。针对数据匮乏问题可以考虑从以下几个方向解决：（1）构建高质量的合成与真实混合数据集，通过域适应技术提升合成数据的真实感与泛化能力；（2）发展基于少量样本或弱监督学习的重打光算法，降低对大规模标注数据的依赖；（3）推动跨机构数据共享协议，在保障隐私的前提下实现资源共建共享。

2）挑战性场景。重打光技术在各种场景都有广泛的应用，但在面对复杂场景时存在多重挑战，这些挑战主要来源于以下几个方面：

（1）复杂光照条件：复杂光照条件下的场景是重打光技术中的一大挑战，源于多光源、非均匀光照、硬阴影和镜面反射等因素对光照效果的影响。这些因素不仅会显著改变图像的亮度和纹理分布，还可能导致光照方向和强度的剧烈波动，增加了光照估计和模拟的难度。未来可以考虑引入基于物理渲染的混合建模方式，将可微渲染器与神经网络结合，实现对真实光照传播规律的建模。除此之外，可采用多任务学习框架，将光照估计、几何重建与阴影检测等任务联合训练，从而增强模型在复杂光照环境下的稳定性与一致性。

（2）动态变化的场景：动态变化的场景是指在同一视频帧中，不同物体或区域因表情变化、头部运动等原因而导致的光照变化。在这些场景中，光照效果随时间或运动而变化，例如人物表情的变化可能会改变面部的光照分布，头部运动则可能导致光源的相对位置发生变化。现有的重打光算法通常依赖于静态光照模型，这使得它们在面对动态变化时，难以实时捕捉和适应光照变化。未来可以考虑结合时序建模以捕捉跨帧光照变化规律，或开发实时跟踪与调整机制，实现对动态光照的持续估计和调控。

（3）非标准人脸图像：非标准人脸图像，如部分遮挡、模糊和低分辨率等，是重打光算法面临的挑战之一，其复杂性源于图像信息的缺失和质量下降。针对这个问题可以通过借助超分辨率与修复技术，在重打光前对图像进行预处理；也可融合3D人脸重建模块，提升在人脸信息不完整情况下的几何估计能力。

3）光照控制。在人像重打光算法中，一个关键挑战是如何提供既灵活又专业的调节手段，确保用户能够根据需求调整光照效果，同时保持高效且直观的操作体验。

（1）灵活的光照编辑：在实际应用中，要求重打光算法具备灵活的光照调整和编辑功能，让用户能够轻松地调整光照效果，如改变光照方向和强度、增加或减少光源等，同时避免过于复杂的操作，现有的算法大多利用球谐光照或环境地图作为输入光照条件，这些光照控制过于单一；GAN-control（Shoshan等，2021）等方法使用光照编码控制光照，这种非显式的方法更加限制了用户对光照的编辑。未来可以引入基于图形界面的交互式光照编辑系统，结合神经渲染实现“所见即所得”的光照调节；或构建可解释的光照潜空间，使用户能通过图像拖拽或自然语言高效控制光照效果。

（2）智能的光照优化：如果算法可以根据输入的图像内容（例如，面部特征、背景环境等）智能地推荐合适的光照设置，将极大地降低用户操作难度，提

高效果的专业性。然而目前的重打光算法并不能进行智能的光照优化，未来可结合大模型中的图文对齐能力，构建“图像—语义—光照”的映射模型，实现基于内容的光照自动推荐。此外，可探索与人脸美学模型结合，实现符合美学原则的光照优化。

（3）人像身份信息的保持：在重打光过程中，光照与人脸的几何结构是相互关联的。现有的许多算法尤其是生成模型，在改变人像光照的同时，也会造成身份信息的丢失，图6展示了DPR、FRRS、DiFaReli和ICLight算法在对应参考光照条件下身份信息的丢失。目前已经有方法在训练中引入身份保持损失，未来可以考虑采用结构先验引导建模，确保几何结构不被破坏；也可借助3D人脸模型或中间表征稳定面部结构，实现真实感与一致性的统一。

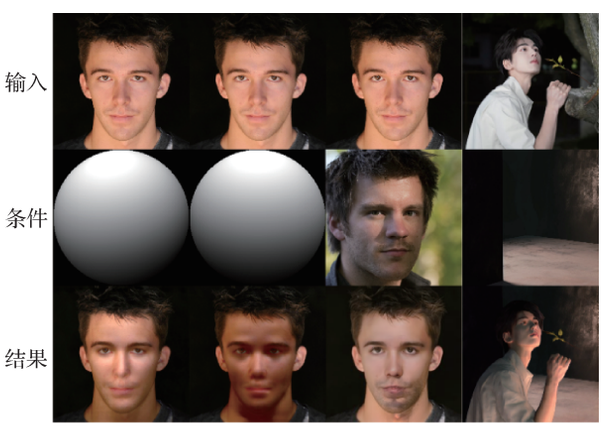


图6　DPR、FRRS、DiFaReli和ICLight造成了人像面部扭曲

Fig. 6　DPR, FRRS, DiFaReli and ICLight caused distortion of human facial features

4）实时处理。随着深度学习模型的广泛应用，重打光算法的性能和重打光质量显著提升，但参数量的增长和计算复杂度的增加给算法的实时性应用带来了阻碍，限制了算法在视频通话等动态场景的实时应用。未来可通过模型压缩（如蒸馏、剪枝、量化）与轻量化网络设计提升处理速度，有望实现实时重打光效果，满足视频会议、直播等动态场景需求。

5）大模型驱动的人像重打光正在形成新范式。以大规模预训练模型为代表的技术在图像生成与编辑中展现出强大能力，凭借丰富的特征表征与跨模态理解，即便在弱监督或零样本条件下也能学习光照与人脸几何的复杂关联，在数据有限或标注不完整时仍可实现高质量的光照重建与编辑；多模态模型进一步使“文本驱动的光照编辑”成为可能，用户可用自然语言调节光照方向、强度与风格，显著提升交互性与可控性。

4　结　语

本文系统梳理了人像重打光数据集与方法，将方法归纳为3类：结合物理模型的方法、基于图像转换的方法和基于三维重建的方法，并在统一协议下给出定量与定性对比。展望未来，亟需更高质量与更大规模的数据、兼顾物理与学习的混合范式、轻量高效的推理方案，以及面向应用的统一评价基准。

参考文献（References）

Basri R and Jacobs D W. 2003. Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(2): 218-233 [DOI: 10.1109/TPAMI.2003.1177153]

Blanz V and Vetter T. 2023. A morphable model for the synthesis of 3D faces//Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: Association for Computing Machinery, 2023: 187-194 [DOI: 10.1145/311535.311556]

Cai H M, Huang T W, Gehlot S, Feng B Y, Shah S, Su G M, et al. 2025. Parametric shadow control for portrait generation in text-to-image diffusion models [EB/OL]. [2025-04-07]. https://arxiv.org/pdf/2503.21943.pdf

Cai Z Q, Jiang K W, Chen S Y, Lai Y K, Fu H B, Shi B X, et al. 2024. Real-time 3D-aware portrait video relighting//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 6221-6231 [DOI: 10.1109/CVPR52733.2024.00595]

Chaturvedi S, Ren M, Hold-Geoffroy Y, Liu J Y, Dorsey J and Shu Z X. 2025. SynthLight: portrait relighting with diffusion model by learning to re-render synthetic faces//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 369-379 [DOI: 10.1109/CVPR52734.2025.00043]

Cook R L and Torrance K E. 1982. A reflectance model for computer graphics. ACM Transactions on Graphics (TOG), 1(1): 7-24 [DOI: 10.1145/357290.357293]

Debevec P, Hawkins T, Tchou C, Duiker H P, Sarokin W and Sagar M. 2000. Acquiring the reflectance field of a human face//Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: Association for Comput-

ing Machinery,: 145-156[DOI: 10.1145/344779.344855]

Deng Y, Yang J L, Chen D, Wen F and Tong X. 2020. Disentangled and controllable face image generation via 3D imitative-contrastive learning//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 5153-5162[DOI: 10.1109/CVPR42600.2020.00520]

Fang Y, Sun Z Y, Zhang S Z, Wu T, Xu Y H, Zhang P, et al. 2025. RelightVid: Temporal-consistent diffusion model for video relighting[EB/OL].[2025-01-27]. https://arxiv.org/pdf/2501.16330.pdf

Feng Y, Feng H W, Black M J and Bolkart T. 2021. Learning an animatable detailed 3D face model from in-the-wild images. ACM Transactions on Graphics (ToG), 40(4): #88[DOI: 10.1145/3450626.3459936]

Futschik D, Ritland K, Vecore J, Fanello S, Orts-Escolano S, Curless B, et al. 2023. Controllable light diffusion for portraits//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 8412-8421[DOI: 10.1109/CVPR52729.2023.00813]

Gardner M A, Hold-Geoffroy Y, Sunkavalli K, Gagné C and Lalonde J F. 2019. Deep parametric indoor lighting estimation//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 7174-7182[DOI: 10.1109/ICCV.2019.00727]

Georghiades A S, Belhumeur P N and Kriegman D J. 2001. From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6): 643-660[DOI: 10.1109/34.927464]

Gross R, Matthews I, Cohn J, Kanade T and Baker S. 2008. Multi-PIE//Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition. Amsterdam, the Netherlands: IEEE: 1-8[DOI: 10.1109/AFGR.2008.4813399]

Han X J, Yang H Y, Xing G Y and Liu Y L. 2020. Asymmetric joint GANs for normalizing face illumination from a single image. IEEE Transactions on Multimedia, 22(6): 1619-1633[DOI: 10.1109/TMM.2019.2945197]

Hong Y, Peng B, Xiao H Y, Liu L G and Zhang J Y. 2022. HeadNeRF: a realtime NeRF-based parametric head model//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 20342-20352[DOI: 10.1109/CVPR52688.2022.01973]

Hou A, Sarkis M, Bi N, Tong Y Y and Liu X M. 2022. Face relighting with geometrically consistent shadows//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 4207-4216[DOI: 10.1109/CVPR52688.2022.00418]

Hou A, Zhang Z, Sarkis M, Bi N, Tong Y Y and Liu X M. 2021. Towards high fidelity face relighting with realistic shadows//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 14714-14723[DOI: 10.1109/CVPR46437.2021.01448]

Jiang K W, Chen S Y, Fu H B and Gao L. 2023. NeRFFaceLighting: implicit and disentangled face lighting representation leveraging generative prior in neural radiance fields. ACM Transactions on Graphics, 42(3): 1-18[DOI: 10.1145/3597300]

Kanamori Y and Endo Y. 2018. Relighting humans: occlusion-aware inverse rendering for full-body human images. ACM Transactions on Graphics (TOG), 37(6): #270[DOI: 10.1145/3272127.3275104]

Karras T, Laine S and Aila T. 2019. A style-based generator architecture for generative adversarial networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 4396-4405[DOI: 10.1109/CVPR.2019.00453]

Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J and Aila T. 2019. Analyzing and improving the image quality of StyleGAN//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 8107-8116[DOI: 10.1109/CVPR42600.2020.00813]

Ke J J, Wang Q F, Wang Y L, Milanfar P and Yang F. 2021. MUSIQ: multi-scale image quality transformer//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 5128-5137[DOI: 10.1109/ICCV48922.2021.00510]

Kim H, Jang M J, Yoon W, Lee J, Na D and Woo S. 2024. SwitchLight: co-design of physics-driven architecture and pre-training framework for human portrait relighting//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 25096-25106[DOI: 10.1109/CVPR52733.2024.02371]

Kwak J G, Li Y M, Yoon D, Kim D, Han D and Ko H. 2022. Injecting 3D perception of controllable NeRF-GAN into StyleGAN for editable portrait image synthesis//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 236-253[DOI: 10.1007/978-3-031-19790-1_15]

Land E H. 1977. The Retinex theory of color vision. Scientific American, 237(6): 108-129[DOI: 10.1038/scientificamerican1277-108]

Land E H and McCann J J. 1971. Lightness and Retinex theory. Journal of the Optical Society of America, 61(1): 1-11[DOI: 10.1364/JOSA.61.000001]

Liu H. 2011. The relighting technology of live action in different viewpoints. Science Technology and Engineering, 11(32): 8036-8039 (刘慧. 2011. 多视角下运动人体的重光照技术. 科学技术与工程, 11(32): 8036-8039)[DOI: 10.3969/j.issn.1671-1815.2011.32.044]

Liu X N, Chen C Y, Hu X J and Yu H Y. 2024. Virtual viewpoint image synthesis using neural radiance fields with depth information super-

vision. Journal of Image and Graphics, 29(7): 2035-2045 (刘晓楠, 陈纯毅, 胡小娟, 于海洋. 2024. 带深度信息监督的神经辐射场虚拟视点画面合成. 中国图象图形学报, 29(7): 2035-2045) [DOI: 10.11834/jig.221188]

Liu Y C, Shu Z X, Li Y J, Lin Z, Zhang R and Kung S Y. 2022. 3D-FM GAN: towards 3D-controllable face manipulation//Proceedings of the 17th European Conference on Computer Vision — ECCV 2022. Tel Aviv, Israel: Springer: 107-125 [DOI: 10.1007/978-3-031-19784-0_7]

Mei Y Q, Zeng Y, Zhang H, Shu Z X, Zhang X E, Bi S, et al. 2024. Holo-Relighting: controllable volumetric portrait relighting from a single image//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 4263-4273 [DOI: 10.1109/CVPR52733.2024.00408]

Mei Y Q, Zhang H, Zhang X E, Zhang J M, Shu Z X, Wang Y L, et al. 2023. LightPainter: interactive portrait relighting with freehand scribble//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 195-205 [DOI: 10.1109/CVPR52729.2023.00027]

Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2022. NeRF: representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1): 99-106 [DOI: 10.1145/3503250]

Nestmeyer T, Lalonde J F, Matthews I and Lehrmann A. 2020. Learning physics-guided face relighting under directional light//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 5123-5132 [DOI: 10.1109/CVPR42600.2020.00517]

Newell A, Yang K Y and Deng J. 2016. Stacked hourglass networks for human pose estimation//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 483-499 [DOI: 10.1007/978-3-319-46484-8_29]

Pan X G, Tewari A, Liu L J and Theobalt C. 2022. GAN2X: non-Lambertian inverse rendering of image GANs//Proceedings of 2022 International Conference on 3D Vision (3DV). Prague, Czech Republic: IEEE: 711-721 [DOI: 10.1109/3DV57658.2022.00081]

Pan X G, Xu X D, Loy C C, Theobalt C and Dai B. 2021. A shading-guided generative implicit model for shape-accurate 3D-aware image synthesis//Proceedings of the 35th International Conference on Neural Information Processing Systems. Red Hook, USA: Curran Associates Inc.: #1530

Papantoniou F P, Lattas A, Moschoglou S and Zafeiriou S. 2023. Relightify: relightable 3D faces from a single image via diffusion models//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 8772-8783 [DOI: 10.1109/ICCV51070.2023.00809]

Ponglertnapakorn P, Tritrong N and Suwajanakorn S. 2023. DiFaReli: diffusion face relighting//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 22589-22600 [DOI: 10.1109/ICCV51070.2023.02070]

Ramamoorthi R and Hanrahan P. 2001. An efficient representation for irradiance environment maps//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM: 497-500 [DOI: 10.1145/383259.383317]

Sengupta S, Curless B, Kemelmacher-Shlizerman I and Seitz S M. 2021. A light stage on every desk//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 2400-2409 [DOI: 10.1109/ICCV48922.2021.00242]

Sengupta S, Kanazawa A, Castillo C D and Jacobs D W. 2018. SfSNet: learning shape, reflectance and illuminance of faces 'in the Wild'//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 6296-6305 [DOI: 10.1109/CVPR.2018.00659]

Shoshan A, Bhonker N, Kviatkovsky I and Medioni G. 2021. GAN-control: explicitly controllable GANs//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 14063-14073 [DOI: 10.1109/ICCV48922.2021.01382]

Shu Z X, Yumer E, Hadap S, Sunkavalli K, Shechtman E and Samaras D. 2017. Neural face editing with intrinsic image disentangling//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5444-5453 [DOI: 10.1109/CVPR.2017.578]

Song J M, Meng C L and Ermon S. 2020. Denoising diffusion implicit models//Proceedings of the 9th International Conference on Learning Representations. Austria: OpenReview.net

Song Y. 2021. Research on Image Relighting Based on Deep Learning. Changchun: Jilin University (宋仪. 2021. 基于深度学习的图像重光照研究. 长春: 吉林大学) [DOI: 10.27162/d.cnki.gjlin.2021.006044]

Sorkine O and Alexa M. 2007. As-rigid-as-possible surface modeling//Proceedings of the 5th Eurographics Symposium on Geometry Processing. Barcelona, Spain: Eurographics Association: 109-116

Stratou G, Ghosh A, Debevec P and Morency L P. 2011. Effect of illumination on automatic expression recognition: a novel 3D relightable facial database//Proceedings of 2011 IEEE International Conference on Automatic Face and Gesture Recognition. Santa Barbara, USA: IEEE: 611-618 [DOI: 10.1109/FG.2011.5771467]

Sun T C, Barron J T, Tsai Y T, Xu Z X, Yu X M, Fyffe G, et al. 2019. Single image portrait relighting. ACM Transactions on Graphics (TOG), 38(4): #79 [DOI: 10.1145/3306346.3323008]

Sun T C, Lin K E, Bi S, Xu Z X and Ramamoorthi R. 2021. NeLF: neural light-transport field for portrait view synthesis and relighting//Proceedings of the 32nd Eurographics Symposium on Rendering, EGSR 2021 — Digital Library Only Track. Saarbrücken, Germany: Eurographics Association: 155-166

Tajima D, Kanamori Y and Endo Y. 2021. Relighting humans in the wild: monocular full-body human relighting with domain adapta-

tion. Computer Graphics Forum, 40(7): 205-216 [DOI: 10.1111/cgf.14414]

Tan F T, Fanello S, Meka A, Orts-Escolano S, Tang D H, Pandey R, et al. 2022. VoLux-GAN: a generative model for 3D face synthesis with HDRI relighting//Proceedings of 2022 ACM SIGGRAPH Conference Proceedings. Vancouver, Canada: Association for Computing Machinery: #58 [DOI: 10.1145/3528233.3530751]

Tang J S, Zhang B, Yang B X, Zhang T, Chen D, Ma L Z, et al. 2024. 3DFaceShop: explicitly controllable 3D-aware portrait generation. IEEE Transactions on Visualization and Computer Graphics, 30(9): 6020-6037 [DOI: 10.1109/TVCG.2023.3323578]

Wang J Y, Chan K C K and Loy C C. 2023a. Exploring CLIP for assessing the look and feel of images//Proceedings of 2023 AAAI Conference on Artificial Intelligence. Washington, USA: AAAI Press: 2555-2563 [DOI: 10.1609/aaai.v37i2.25353]

Wang J Y, Liu J Y, Sun X, Singh K K, Shu Z X, Zhang H, et al. 2025. Comprehensive relighting: generalizable and consistent monocular human relighting and harmonization//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 380-390 [DOI: 10.1109/CVPR52734.2025.00044]

Wang Y F, Holynski A, Zhang X M and Zhang X E. 2023b. SunStage: portrait reconstruction and relighting using the sun as a light stage//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 20792-20802 [DOI: 10.1109/CVPR52729.2023.01992]

Wang Z B, Yu X, Lu M, Wang Q, Qian C and Xu F. 2020. Single image portrait relighting via explicit multiple reflectance channel modeling. ACM Transactions on Graphics (ToG), 39(6): #220 [DOI: 10.1145/3414685.3417824]

Woodham R J. 1979. Photometric stereo: a reflectance map technique for determining surface orientation from image intensity//Proceedings of SPIE 0155, Image Understanding Systems and Industrial Applications I. San Diego, USA: SPIE: 136-143

Woodham R J. 1980. Photometric method for determining surface orientation from multiple images. Optical Engineering, 19(1): #191139 [DOI: 10.1117/12.7972479]

Yang K Y. 2017. Research on Algorithm of Re-light Illumination Based on Single Face Image. Shenyang: Shenyang University of Technology (杨克义. 2017. 基于单幅图像人脸重光照的算法研究. 沈阳: 沈阳工业大学)

Yeh Y Y, Nagano K, Khamis S, Kautz J, Liu M Y and Wang T C. 2022. Learning to relight portrait images via a virtual light stage and synthetic-to-real adaptation. ACM Transactions on Graphics (TOG), 41(6): #231 [DOI: 10.1145/3550454.3555442]

Zhang H, Xu Y B, Dai T Y, Tai Y W and Tang C K. 2023. FaceDNeRF: semantics-driven face reconstruction, prompt editing and relighting with diffusion models//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #2428

Zhang L, Zhang L and Bovik A C. 2015. A feature-enriched completely blind image quality evaluator. IEEE Transactions on Image Processing, 24(8): 2579-2591 [DOI: 10.1109/TIP.2015.2426416]

Zhang L M, Rao A Y and Agrawala M. 2025. Scaling in-the-wild training for diffusion-based illumination harmonization and editing by imposing consistent light transport//Proceedings of the 13th International Conference on Learning Representations. Singapore, Singapore: OpenReview.net

Zhang L W, Zhang Q X, Wu M Y, Yu J Y and Xu L. 2021. Neural video portrait relighting in real-time via consistency modeling//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 782-792 [DOI: 10.1109/ICCV48922.2021.00084]

Zhou H, Hadap S, Sunkavalli K and Jacobs D. 2019. Deep single-image portrait relighting//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 7193-7201 [DOI: 10.1109/ICCV.2019.00729]

Zhu X Y, Liu X M, Lei Z and Li S Z. 2019. Face alignment in full pose range: a 3D total solution. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(1): 78-92 [DOI: 10.1109/TPAMI.2017.2778152]

作者简介

雷莉娜，女，硕士研究生，主要研究方向为计算机视觉和图像视频生成。E-mail: leilina@mail.nankai.edu.cn

李重仪，通信作者，男，教授，主要研究方向为计算机视觉、图像处理和计算成像。E-mail: lichongyi25@gmail.com

朱家成，男，硕士，主要研究方向为计算机视觉和图像处理。E-mail: zhugucheng@oppo.com

郭春乐，男，副教授，主要研究方向为计算机视觉、图像增强与复原、底层视觉。E-mail: guochunle@nankai.edu.cn

黄杰文，男，博士，主要研究方向为计算机视觉和图像处理。E-mail: huangjiewen@oppo.com

罗俊，男，硕士，主要研究方向为计算机视觉和图像处理。E-mail: jluo@oppo.com