中图法分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1006-8961(XXXX)XX-0001-17

论文引用格式:Liu Jingyi, Zhang Jingsong, Ma Jian, Li Kun. CURL: Multi-view 3D Hair Reconstruction via Adaptive Gaussian Ellipsoids under Varying Illumination[J/OL]. Journal of Image and Graphics, XXXX:1-17. DOI: 10.11834/jig.260062.(刘菁薏, 张劲松, 马健, 李坤. 自适应高斯椭球与复杂光照鲁棒的多视图三维头发重建[J/OL]. 中国图象图形学报, XXXX:1-17. DOI: 10.11834/jig.260062.)[DOI:10.11834/jig.260062]

自适应高斯椭球与复杂光照鲁棒的多视图三维头发重建

刘菁薏[1],张劲松[2,1],马健[1*],李坤[1]

1. 天津大学,天津　300350; 2. 福州大学,福州　350108

摘　要:头发的三维重建作为计算机视觉与计算机图形学领域中的重要研究方向,旨在精确建模头发复杂的几何结构与外观特性。现有的头发重建方法大多建立在理想光照假设下,能够在简单发型上取得较好效果,但在面对复杂光照条件以及致密卷发、高卷曲发型等高复杂度发型时,难以保持几何结构的准确性与重建结果的稳定性。本文提出一种基于自适应高斯椭球表示并具备复杂光照鲁棒性的多视图三维头发重建方法,能够在真实复杂场景中实现高精度的头发几何结构建模及细节丰富、外观一致的纹理重建。具体来说,本文提出了一种结合视觉语言模型(Vision-Language Model,VLM)与三维高斯泼溅(3D Gaussian Splatting,3DGS)的重建框架,实现对复杂发型(如致密卷发、高卷曲发型)的精细几何建模与逼真外观恢复。本文在多视图三维头发重建领域引入VLM对多视角图像进行光照质量评估与发型复杂度分析,构建光照-复杂度双感知的闭环引导机制。智能定位低光及逆光视角,动态调整高斯椭球密度,实现稀疏几何区域的自适应填充。本文设计了几何-纹理解耦优化策略,先联合优化几何与颜色,再在固定几何下微调颜色。实验结果表明,本文方法在自建数据和公共数据集上的定性结果均优于现有最先进方法(如Gaussian Haircut、Im2Haircut、DiffLocks),在几何精度、方向一致性和纹理恢复上均取得提升。与Gaussian Haircut相比,本文方法的PSNR(Peak Signal-to-Noise Ratio)提高0.54dB,SSIM(Structural Similarity Index Measure)提高0.004,LPIPS(Learned Perceptual Image Patch Similarity)降低0.01,验证了本文方法在重建质量与视觉一致性方面的有效性。综上,本文方法通过引入视觉语言模型与自适应密度控制的协同机制,提升了复杂光照与高复杂度发型条件下的三维头发几何重建精度与稳定性,实现了几何结构与外观纹理的一致重建,并增强了方法在真实场景中的鲁棒性与泛化能力。本文方法的代码已开源至GitHub平台(https://github.com/Blueekyyy/CURL),同时已在ScienceDB平台完成代码公开(https://doi.org/10.57760/sciencedb.j00240.00223)。

关键词:三维头发重建;视觉语言模型;复杂光照;自适应密度控制;几何-纹理解耦

CURL: Multi-view 3D Hair Reconstruction via Adaptive Gaussian Ellipsoids under Varying Illumination

Liu Jingyi[1], Zhang Jingsong[2,1], Ma Jian[1], Li Kun[1]

1. Tianjin University, Tianjin 300350, China; 2. Fuzhou University, Fuzhou 350108, China

收稿日期:2026-01-28;**修回日期**:2026-07-06

* **通信作者**:马健,男,1993年生,博士,助理研究员,硕士生导师,CSIG会员,主要研究领域为三维视觉。E-mail:jianma@tju.edu.cn;马健 jianma@tju.edu.cn

基金项目:中国国家重点研发计划(2023YFC3082100);天津市杰出青年科学基金(22JCJQJC00040);国家自然科学基金(62501416);天津市自然科学基金(24JCYBJC01300)

Supported by: National Key R&D Program of China (2023YFC3082100), Science Fund for Distinguished Young Scholars of Tianjin (No. 22JCJQJC00040), National Natural Science Foundation of China (62501416), Natural Science Foundation of Tianjin, China (No. 24JCYBJC01300)

Abstract: Hair reconstruction in three dimensions remains a fundamental challenge in computer vision and computer graphics due to the intricate geometric structure and complex appearance of hair. Its highly anisotropic geometry, fine-scale strand organization, and sensitivity to illumination variations make accurate reconstruction from images particularly difficult in real-world scenarios. These challenges are further amplified under non-ideal lighting conditions, such as underexposure or backlighting, where multi-view images often exhibit inconsistent photometric properties. Moreover, complex hairstyles—including curly, wavy, and afro-like patterns—pose additional difficulties, as existing methods frequently suffer from sparse geometric coverage, missing fine details, and unreliable directional cues, resulting in incomplete or visually implausible reconstructions. Recent progress in 3D reconstruction has introduced anisotropic Gaussian representations, most notably 3D Gaussian Splatting (3DGS), which model scenes as collections of oriented Gaussian ellipsoids and enable efficient rasterization-based rendering with adaptive density control. While these representations provide a flexible framework for general scene reconstruction, they are insufficient for hair-specific challenges. In regions affected by poor lighting, sparse geometric signals lead to holes in reconstructed hair volumes, and conventional density control mechanisms lack the adaptability required to handle varying hairstyle complexity. Furthermore, the strong coupling between geometry and texture under complex illumination is often overlooked, limiting photorealistic hair reconstruction. To overcome these limitations, we propose a novel multi-view hair reconstruction pipeline that integrates visual-language models (VLMs) as an intelligent guidance module for 3D Gaussian-based reconstruction. The key insight is that VLMs can provide high-level semantic and photometric understanding across multiple views, enabling informed decisions on where and how Gaussian ellipsoids should be densified. Our method leverages VLMs to guide reconstruction through three core tasks. First, the VLM evaluates the lighting quality of multi-view images using a dedicated prompt, identifying views prone to geometric degradation due to unfavorable illumination. This enables targeted densification only in necessary regions, improving geometric completeness while avoiding redundant computation. Second, for views affected by underexposure or backlighting, we trigger a three-stage adaptive density control strategy that generates denser Gaussian ellipsoids, leading to more reliable 2D orientation maps and more accurate hair direction supervision. Third, the VLM estimates hairstyle complexity from multi-view inputs and dynamically determines the maximum Gaussian density required. Complex hairstyles such as tight curls or afro patterns are assigned higher density to capture fine geometric details, while simpler hairstyles are reconstructed with fewer Gaussians. In addition, we introduce a decoupled geometry–texture optimization strategy to mitigate color biases introduced by lighting enhancement. Geometry and color are jointly optimized using enhanced images to ensure stable geometric reconstruction. Geometry is fixed while color is refined using original images, allowing the recovery of natural hair color and illumination effects. This strategy ensures both geometric precision and visual realism. Our pipeline further incorporates refinement inspired by prior Gaussian-based hair reconstruction methods. Adaptive Gaussian replication and splitting are applied selectively to balance hole filling and over-densification. Training is supervised using a combination of L1, SSIM, mask, orientation, perceptual, and edge losses, enabling accurate modeling of geometry, texture, and fine structural details while maintaining training efficiency. We evaluate our method on both self-built and public datasets, comparing against state-of-the-art approaches such as Gaussian Haircut, Im2Haircut, and DiffLocks. Experimental results demonstrate that our method produces more detailed, coherent, and realistic hair reconstructions, particularly under challenging lighting conditions and for complex hairstyles. Quantitative evaluations show consistent improvements over Gaussian Haircut, with PSNR increased by 0.54dB, SSIM improved by 0.004, and LPIPS reduced by 0.01. Ablation studies further validate the effectiveness of VLM-guided densification, lighting-aware adaptation, and decoupled optimization. In summary, we present a VLM-guided 3D Gaussian hair reconstruction framework that robustly handles non-ideal lighting and complex hairstyles. By integrating semantic analysis, adaptive densification, and decoupled optimization, our method advances the state of the art in photorealistic 3D hair reconstruction and provides a promising direction for future research.

Key words: 3D Hair reconstruction; visual language model; complex illumination; adaptive densification; geometry-texture decoupling

论文引用格式: Liu Jingyi, Zhang Jingsong, Ma Jian, Li Kun. CURL: Multi-view 3D Hair Reconstruc-

tion via Adaptive Gaussian Ellipsoids under Varying Illumination[J/OL]. Journal of Image and Graphics. DOI:10.11834/jig.260062(刘菁薏,张劲松,马健,李坤. CURL:自适应高斯椭球与复杂光照鲁棒的多视图三维头发重建[J/OL]. 中国图象图形学报. DOI:10.11834/jig.260062.)

输入图像 CURL 输入图像 CURL

图1高卷曲发型等复杂发型条件下的三维重建结果。

Fig. 1 3D reconstruction results under challenging conditions such as high-curvature hairstyles. .

0 引 言

头发的三维重建作为数字人建模与真实感虚拟内容生成中的关键基础问题,对于推动虚拟现实、增强现实、影视特效、数字娱乐以及人机交互等多个领域的发展具有重要的理论价值与应用意义。由于头发发丝同时具有反射、折射、散射等复杂光学特性以及致密卷发等多样且高度复杂的几何造型,使得在真实场景中头发的三维重建难度增加。早期国内研究主要关注基于图像的人头及头发三维建模,例如(林源等,2011)提出正交图像下的真实感人头重建方法,对头发进行了特征点提取与几何建模;此类方法为头发建模提供了初步思路,但受限于表示能力与输入条件,难以刻画细丝结构与复杂几何。近年来,随着学习型方法与三维表示的发展,现有的方法(Zakharov 等,2024; Sklyarova 等,2025; Rosu 等,2025)虽在一定程度上提升了头发建模质量,但在致密卷发等复杂发型场景下仍难以稳定重建细节,常出现几何缺失和细节丢失的问题。为此,本文提出了一种基于自适应高斯椭球表示、具备复杂光照鲁棒性的多视图三维头发重建方法,能够在复杂光照条件下对复杂发丝实现精细的几何建模与逼真的纹理恢复,重建结果如图1所示。

现有的方法如 Im2Haircut(Sklyarova 等,2025)和 DiffLocks(Rosu 等,2025)主要依赖单视角图像进行头发重建,对于未被观察到的区域通常通过估计或先验推断来恢复,这导致几何信息不够精准;另外,基于多视图输入的三维头发重建方法,代表性工作包括 Gaussian Haircut(Zakharov 等,2024)和 HairGS(Pan 等,2025),均通过对齐和优化多视角的头发曲线或高斯椭球表示以增强重建结果在整体几何一致性和细节表达方面的表现,但在复杂光照和高频卷曲结构下仍存在局限性。特别是在缺乏几何监督信号的区域,这些方法往往依赖稀疏的三维几何信息,导致光栅化后的纹理方向图极度稀疏,从而严重影响最终模型细节的恢复。这些方法普遍忽略了纹理与几何结构之间的深度耦合关系,对复杂光照条件和材质变化的适应能力不足。在使用三维高斯泼溅拟合复杂发型的重建过程中,往往会出现以下几个问题:1)视觉特征不稳定:自然环境中光照变化剧烈,头发表观不稳定及发丝交错缠绕,存在遮挡严重区域。视觉特征极易失真紊乱,复杂光照和发型复杂度的感知困难。导致高斯表征存在不连续性与稀疏性,重建结果出现结构空洞、局部残缺及纹理缺失。2)发丝方向复杂多变:低光、逆光及遮挡区域仅能清晰观测少量特定朝向发丝,大量不同走向的高卷曲致密纤细发丝及内层发丝信息未能有效捕获,存在严重缺失。3)发丝轴向结构不均:真实发丝沿发根、发中、发尾存在直径与形态梯度差异,此类纵向差异难以精准建模,导致单根发丝边界刻画模糊、细节丢失,造成整体发型轮廓松散、体积失真、结构不规整。

本文提出自适应高斯椭球与复杂光照鲁棒的多视图三维重建方法,有效提升复杂光照及复杂发丝重建的精度与真实性。其中,针对视觉特征不稳定的问题,本文将视觉语言模型引入头发重建流程,实现低光、逆光视角下几何空洞的智能定位与精准感知,为后续高斯表征优化提供支撑。此外,针对发丝方向复杂多变的问题,本文提出三阶段自适应密度控制策略,依据发型复杂度动态调控高斯椭球空间密度分布,对光照最差视角进行针对性稠密化;同时

设计几何-纹理解耦优化机制，实现几何结构精确性与渲染真实性的协同优化。最后，针对发丝轴向结构不均的问题，本文基于发丝结构约束构建感知与边缘损失联合的监督机制，对发根、发中、发尾通过语义到结构、全局到局部的递进约束，强化对发丝精细边界与语义层级的引导，提升高复杂度发型重建几何结构的细粒度与保真度。

本文的主要创新：

1）提出自适应高斯椭球与复杂光照鲁棒的多视图三维头发重建方法，为复杂光照、高复杂度发型的精准重建提供了系统性框架。

2）提出基于视觉语言模型（VLM）的高斯稠密化引导方法，该方法由 VLM 智能引导与三阶段自适应密度控制构成，实现了几何空洞的精准感知与针对性稠密化，有效解决光照与发型复杂度感知不足的问题。

3）提出基于发丝结构约束的监督方法，该方法由感知损失与边缘损失联合构成，从发根、发中、发尾到发型整体进行多层次监督，强化了发丝精细边界与层级结构约束，提升了高复杂度发型重建的细粒度与保真度。

4）实验验证表明，本文方法在定性结果中提升了几何精度、方向一致性与纹理恢复效果；定量结果中，PSNR 平均提高 0.54dB、SSIM 平均提高 0.004、LPIPS 平均降低 0.01，充分验证了方法在重建质量与视觉一致性上的有效性。

1 相关工作

1.1 基于3D高斯泼溅表示的重建与头发重建

近年来，以3D高斯泼溅为代表的基于3D高斯椭球的场景表示在实时视图合成与重建方向取得了重要进展。3D高斯泼溅以各向异性高斯基元为基本表示单元，结合可视化感知光栅化渲染策略，大幅提升了训练效率与实时渲染速度，有效改善了场景的高斯点云分布。（Kerbl等，2023）。此类方法后续有多种改进方向——例如更精确的体素/椭球体体渲染（Mai等，2025）、以及引入法线先验或锚定策略来约束高斯方向（Guan等，2025）等，这些工作为本文采用高斯椭球表示头发几何奠定了方法学基础（Kerbl等，2023；Mai等，2025；Guan等，2025）。与此同时，相关综述（杨航等，2023）也系统总结了深度学习驱动的三维重建技术，从体素、点云到网格及多视图重建策略，展示了三维场景重建在视觉感知、几何建模与网络设计方面的整体进展。

专门针对头发的重建研究既有经典的基于物理模拟与多视几何方法（Beeler等，2012；Zhou等，2018），也有近年的深度学习单视图/多视图恢复工作（Zhou等，2018；Zheng等，2024）。这些发型专门化工作通常侧重于发丝的细节建模（线状/卡片状/体积状表示）、数据驱动的形状先验及纹理恢复，但在复杂光照或视角受限时仍容易出现几何稀疏、方向图不可靠、以及局部空洞等问题（Beeler等，2012；Zhou等，2018；Zheng等，2024）。

1.2 视觉–语言模型（VLM）在生成模型和3D视图中的引导和应用

大规模视觉-语言模型（VLM）在图像/视频理解与多模态推理中展现了强大的语义解析与指令式能力（如大型语言与视觉助手（Large Language and Vision Assistant，LLaVA）、视觉指令调优等），特别是在图像复原与增强等挑战性任务中取得了进展（韦炎炎等，2025）。此外，VLM还被广泛用于引导生成模型与视觉评分任务（Liu等，2023）。在与生成模型的结合上，文本/视觉引导的扩散或生成网络（引导语言到图像扩散生成与编辑模型（Guided Language-to-Image Diffusion for Generation and Editing，GLIDE）、DALL·E系列即OpenAI提出的文本生成图像模型）已被证明能以语言条件改善视觉合成的可控性（Nichol等，2021；Ramesh等，2022）。近两年开始出现将VLM扩展到三维场景理解或“三维重建指令调优”的研究方向（例如面向三维重建推理的视觉语言模型（visual-language models for 3D reconstruction reasoning，VLM-3R）、LayoutVLM即基于视觉语言模型的三维布局优化框架），这些工作通过在VLM中注入三维结构或布局指令，使模型能对三维空间关系、视角质量或场景布局给出语义性建议或评分（Fan等，2025；Sun等，2025）。尽管VLM在纯二维图像/视频生成与理解上已有大量工作，但将其直接用于三维重建流程的闭环引导（例如自动定位低质量视角、基于语义估计控制稠密化强度）的研究仍较少且处于起步阶段（Liu等，2023；Fan等，2025）。

1.3 非理想条件下的鲁棒性处理：图像增强、质量评估与稠密优化的融合

为提高在低光/逆光等非理想拍摄条件下的重

建质量，相关工作主要沿两条路线展开。其一是图像预处理与增强：低光增强和色彩空间改进（如HVI）等方法能在预处理阶段恢复更多可辨信息，从而为后续重建提供更清晰的视觉输入（Yan等，2025；Wu等，2023）。其二是重建端的鲁棒稠密化与优化：例如基于视空间位置或重建误差的自适应复制/分裂策略、法线/深度先验约束、以及多阶段密度控制等，都是为了在重建过程中补偿输入视角缺陷并避免过度稠密化引入噪声（Kerbl等，2023）。总体来看，已有方法多数在预处理到重建的二阶段或被动稠密化上单方向努力，缺乏一种将图像质量评估、语义/光照感知与重建稠密化决策紧密耦合的闭环体系；本文工作正是将水平/垂直强度色彩空间（Horizontal/Vertical-Intensity，HVI）的低光增强、Q-Align即面向视觉质量评分的大型多模态模型微调框架的图像质量评分与VLM驱动的决策机制结合到3D高斯泼溅式重建流水线中，实现预处理

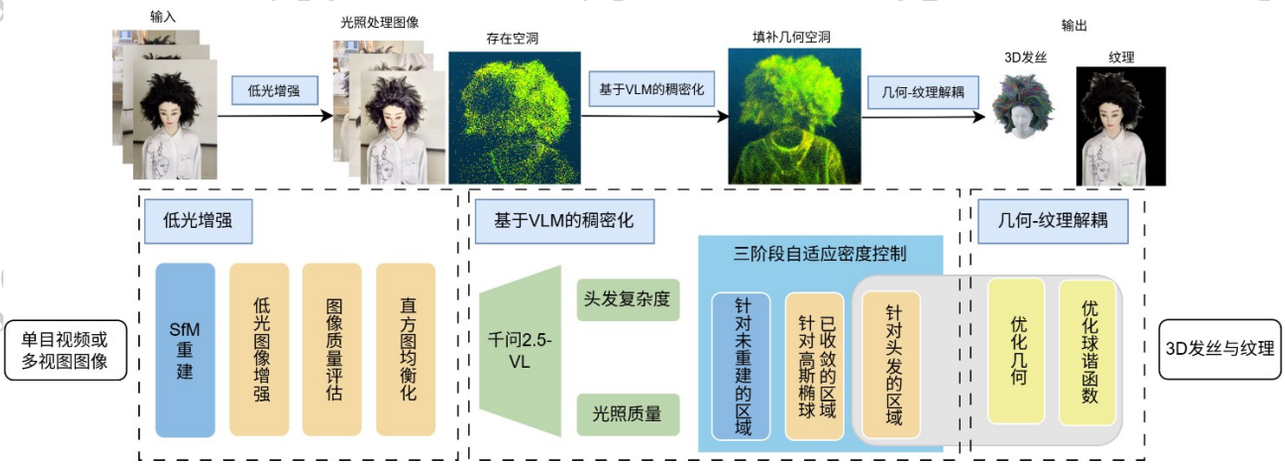


图2　CURL方法整体流程（灰色框内部分受到监督，相关损失函数见第2.3节。方法共包括低光增强、基于VLM的稠密化、几何-纹理解耦三个模块。上面是示例，下面是更详细的流程图）

Fig. 2　Pipeline of the CURL (The components within the gray box are supervised; the corresponding loss functions are detailed in Section 2.3. The proposed method consists of three modules: low-light enhancement, VLM-based densification, and geometry-texture decoupling. The upper part illustrates an overview example, and the lower part presents a more detailed pipeline.)

自适应稠密化的联合优化，从而在复杂光照条件下提升几何完整性与渲染真实性。

针对现有三维高斯泼溅方法在复杂光照和稀疏视角条件下难以稳定重建头发细节的问题，本文提出一种基于视觉语言模型的视角诊断与自适应稠密化策略。该方法利用VLM对光照与语义复杂度的高阶理解，引导高斯椭球的针对性增密，从而有效缓解头发区域的几何空洞与结构退化。

然而，头发本身具有高度复杂的几何结构与各向异性外观属性。在复杂光照条件下易产生强反射、阴影及亮度不一致等现象，对头发外观影响显著，导致多视图之间的视觉特征不稳定，从而增加三维重建的难度。同时，发丝方向复杂多变且轴向结构分布不均，使得发丝几何难以通过高斯表征进行准确建模。但Gaussian Haircut、HairGS（Pan等，2025）及单视图方法Im2Haircut（Sklyarova等，2025）、DiffLocks（Rosu等，2025）普遍缺乏对光照与发型复杂度的联合建模机制，对发丝结构与光照变化的感知能力不足，影响重建结果的几何准确性与纹理一致性。为此，本文提出自适应高斯椭球与复杂光照鲁棒的多视图三维头发重建方法，针对性弥补上述现有方法的不足。构建了面向复杂发型的自适应三维重建框架，将VLM用于闭环引导，通过视觉语言模型对输入场景的发型复杂度与光照质量进行量化评估，据此动态调整高斯椭球数。并将几何优化与颜色优化的监督来源显式分离，这一解耦策略从根本上避免了光照偏差对几何重建的污染，解决了现有方法的核心局限。

2 方 法

2.1 基于视觉语言模型的稠密化

2.1.1 发丝纹理清晰度增强机制

以往三维头发重建方法通常在光照充足且较为理想的条件下进行数据采集，这一假设在复杂真实场景中往往难以满足。特别是在复杂光照环境下，头发区域的纹理细节和方向信息容易被淹没，导致重建结果退化。为此，本文首先对所有目标图像统一进行 HVI（Yan 等，2025）低光增强与直方图均衡化处理，以改善整体亮度分布与局部对比度。利用 Q-Align（Wu 等，2023）对处理后的图像进行逐一质量评判，从光照一致性与视觉可感知质量等方面对增强效果进行衡量，筛选并确保最终输入图像在光照质量上达到较优水准。低光增强即发丝纹理清晰度增强流程如图 2 所示。

表 1 prompt-complexity 提示词示例

Table 1 Prompt-complexity prompt examples

等级	提示词
1	直发
2	自然卷
3	大波浪卷
4	小卷
5	爆炸卷
6	麻花辫
7	鱼骨辫
8	混合发型

2.1.2 基于 VLM 稠密化机理

低光或逆光问题会造成图像质量过低。在图像质量较差的视角，三维高斯泼溅会用较少的高斯椭球去拟合，光栅化后二维椭圆少，用二维椭圆的位置和姿态作为纹理方向图就会很稀疏，方向图二维平面上的椭圆就很少，会导致方向图的监督不精确。当处于这种情况下时，方向图对高斯椭球的优化会产生错误的引导或无引导作用，导致中间重建结果出现空洞，从而使最终的头发几何重建结果质量较低。三维高斯泼溅模型中的高斯椭球数量多时，高斯椭球空间分布越密集，三维高斯椭球光栅化后二维方向图就密集且精确，这使得模型能够对复杂发型的表现程度更好，尤其对于复杂发型（比如致密卷发或者高卷曲发型）的重建精度会更高。为了解决空洞问题，本文提出了基于 VLM 的稠密化，以增强几何细节。对于上百张多视图图像，如何定位空洞的位置，针对选择哪张图像进行稠密化，本文引入视觉语言模型对易产生空洞的图像进行自动识别与定位，并建立稠密化数量与发型复杂度之间的自适应关联机制。由千问 2.5-VL 大模型根据输入的 128 张多视图图像推理出头发复杂度。本文将头发复杂度确定为 8 个量级，每一级包含由不同的描述所形容的发型，从直发到致密卷发、从高卷曲发型到麻花辫，并融合到 prompt-complexity 提示词里，如表 1 所示。同样利用千问 2.5-VL 大模型，由 prompt-light 提示词来评估图像质量，并确定图像中光照质量最差的一帧，如表 2 所示。该提示词包含了多种光照状态并且由细致的光照质量等级组成，并有 18 个评分量级。基于 VLM 稠密化机理如图 2 所示。

表 2 prompt-light 提示词示例

Table 2 Prompt-light prompt examples

等级	提示词	等级	提示词
1	昏暗的	10	柔和的
2	不足的	11	平衡的
3	不均匀的	12	清晰的
4	阴暗的	13	明亮的
5	模糊的	14	自然的
6	刺眼的	15	温暖的
7	不自然的	16	理想的
8	过曝的	17	精致的
9	高对比度的	18	完美的

2.1.3 3D 高斯泼溅重建

受到因子化 3DGS（factorized 3D Gaussian splatting，F-3DGS）（Sun 等，2024）和三维高斯泼溅（3D Gaussian Splatting，3DGS，Kerbl 等，2023）的启发，初始化方案对于实现三维高斯泼溅的高重建质量至关重要，尤其对于复杂光照条件下的头发和场景而言。原始三维高斯泼溅利用运动恢复结构技术（Schönberger 等，2016），在训练开始时初始化三维高斯椭球。为了尽量满足三维高斯泼溅模型中高斯椭球的均匀分布，本文在原有框架之前增加了三维高斯泼溅非结构化高斯重建过程，此重建过程中的三

维高斯泼溅模型只是多视图对应的图像场景重建，未考虑高斯椭球与头发区域发丝方向对齐。本文将三维高斯泼溅重建步骤的训练迭代次数设定为30000次，这点与原始3DGS的做法相同。本文还启动了深度估计、反锯齿和曝光控制三种机制（Kerbl等，2023），以便在重建三维高斯泼溅模型时使非结构化高斯重建结果更加精确。

2.1.4　三阶段自适应密度控制

原始3DGS（Kerbl 等，2023）中提出的自适应密度控制方法主要关注“重建不足”和“重建过度”的区域，这两类区域通常对应较大的视图空间位置梯度。尽管该方法包含自适应密度控制机制，其稠密化过程主要服务于三维高斯泼溅模型的非结构化重建，难以在低质量输入视角或头发区域内有效保持高斯椭球分布的空间密集性。原始三维高斯泼溅在训练中针对未重建空间的高斯椭球进行稠密化，同时将优化过程固定为500到15000次循环之间处理，本文称之为基础场景稠密化阶段。而在此之后，本文关注已重建空间中已匹配的高斯椭球，这部分高斯椭球已经收敛并充分拟合场景，其参数变化小。本文的优化过程固定在15000到30000次循环之间，本文称之为光照缺陷靶向稠密化阶段。本文在间隔100次迭代后进行一次高斯椭球复制和分裂的同时，移除高斯椭球密度控制修剪方法以保持椭球数量不减少。本文关注低精度的三维高斯泼溅模型，并根据头发分割标签单独针对头发区域进行稠密化，本文称之为发丝区域精细稠密化阶段。该阶段的自适应密度控制是为了解决引导高精度的三维高斯泼溅模型时的拟合问题。本文根据最大高斯椭球数量值max_points，来限制高斯椭球增加的总密度数量。该值依据发型复杂度推理来确定。在循环到光照质量最差相机视角时，才应用三阶段自适应密度控制，这样可以确保高斯椭球的复制和分裂是限定在由于图像质量差而导致的空洞模型视角处，从而得以填补空洞，使得模型能够完整的构建出。三阶段自适应密度控制如图2所示。

基于VLM的稠密化过程如算法1所述。利用视觉语言模型对输入的多视图图像逐一进行分析，分别评估每个视角的发型复杂度max_complexity和光照质量min_quality，并记录光照质量最差的视角worst_frame。根据发型复杂度通过转化函数确定高斯椭球的最大稠密化数量max_points（每种发型对应的经验值），并初始化三维高斯非结构化模型G。三阶段自适应密度控制过程包括：（1）执行原始三维高斯泼溅的自适应稠密化控制策略，重建出多视图图像场景对应的三维高斯泼溅模型G；（2）针对当前视角为光照质量最差视角且高斯椭球数量小于max_points/2时，触发第二阶段稠密化，模型G的高斯椭球数量大于等于max_points；（3）根据头发分割标签统计头发区域高斯椭球数量，在模型G的高斯椭球总密度数小于densify_points时执行第三阶段稠密化，模型G的高斯椭球数量大于等于densify_points（椭球总数的经验值）。本算法输出优化后的模型G，此时获得了高斯椭球按发丝对齐排列的三维高斯泼溅结构化模型，模型G会渲染出粗级发丝拟合和精细发丝拟合重建过程的真值。

算法1：基于VLM的稠密化算法

输入：多视图图像Images，视觉语言模型VLM

输出：优化后的三维高斯泼溅结构化模型G

1）max_complexity ← 0，min_quality ← INF，worst_frame ← NULL

2）for each image ∈ Images

3）complex_level ← VLM. Evaluate（image，ComplexityPrompt）

4）quality_level ← VLM. Evaluate（image，LightingQualityPrompt）

5）max_complexity ← max（max_complexity，complex_level）

6）if quality_level < min_quality

7）min_quality，worst_frame ← quality，image

8）max_points ← ComplexityToPoints（max_complexity）

9）初始化高斯模型G

10）DensifyAndPrune_Stage1（G）

11）if current_frame == worst_frame and G_sum < max_points / 2

12）DensifyAndPrune_Stage2（G）

13）label_sum ← CountGaussians（G_sum，hair_segmentation_label >= 0. 5）

14）if G_sum + label_sum <= 1_300_000

15）DensifyAndPrune_Stage3（G）

16）return G

由此，本文促进了三维高斯泼溅模型中高斯椭球的均匀分布，解决了由低光或逆光引起的三维高

斯泼溅模型中高斯椭球分布的空洞问题。

2.1.5 几何-纹理解耦

在Gaussian Haircut方法(Zakharov等,2024)的基础上,本文执行几何重建与几何优化,包括发丝粗级拟合和发丝精细拟合。在发丝粗级拟合和发丝精细拟合训练之间,本文进一步引入了一个独立的中间优化步骤,单独采用边缘损失进行额外的几何细化优化,针对头发几何结构进行更高精度的局部调整,提升重建结果的边界一致性与细节表现。使用低光增强后的图像$\boldsymbol{I}^{\text{enh}}$进行训练时,图像发丝纹理清晰度更高,能够获取较为准确的几何细节,但低光增强同时会对颜色分布引入系统性偏差,影响最终渲染结果的准确性。为此本文提出几何-纹理解耦策略,将优化变量显式分离为几何参数集合与颜色参数集合:

$$\mathcal{G}=\{\mu,\boldsymbol{R},\boldsymbol{S},\alpha,\boldsymbol{d}\},\mathcal{C}=\{sh\} \tag{1}$$

式中μ、$\boldsymbol{R}$、$\boldsymbol{S}$、α分别表示高斯基元的位置、旋转、缩放、透明度,$\boldsymbol{d}$为发丝线段方向,sh为球谐函数系数。在几何与颜色联合优化过程中,以低光增强图像为监督信号,对几何与颜色参数进行联合优化:

$$\{\mathcal{G}^{*},\mathcal{C}^{*}\}=\arg\min_{\mathcal{G},\mathcal{C}}\mathrm{L}_{\text{hair}}(\hat{\boldsymbol{I}}(\mathcal{G},\mathcal{C}),\boldsymbol{I}^{\text{enh}}) \tag{2}$$

式中L_{hair}为总损失函数,参考2.2节。几何与颜色信息均参与梯度更新,使模型能够从低光增强图像中充分学习发丝的几何空间结构与局部细节。在颜色解耦优化过程中,几何参数固定的前提下,以原始图像$\boldsymbol{I}^{\text{raw}}$为监督信号,单独对球谐函数系数进行优化:

$$\mathcal{C}^{**}=\arg\min_{c}\mathrm{L}_{\text{hair}}(\hat{\boldsymbol{I}}(\mathcal{G}^{*},\mathcal{C}),\boldsymbol{I}^{\text{raw}}),\ \text{s.t.}\nabla_{\mathcal{G}}=0 \tag{3}$$

相机参数、高斯模型的几何属性及发丝线段方向的优化均被关闭,仅更新球谐函数系数sh。该优化过程在几何与颜色联合优化基础上继续训练,以确保颜色修正建立在精确几何的约束之上。最终渲染结果由几何与颜色联合优化过程所得几何参数与颜色解耦优化过程所得颜色参数联合决定:

$$\widehat{\boldsymbol{I}^{final}}=\mathcal{R}(\mathcal{G}^{*},\mathcal{C}^{**}) \tag{4}$$

通过上述解耦策略,本文将几何精度的来源(低光增强图像)与颜色保真度的来源(原始图像)显式分离,使渲染结果在保留发丝几何结构的同时,尽可能与真实图像保持一致的光照与颜色表现。几何-纹理解耦流程如图2所示。

2.2 基于发丝结构约束的损失函数设计

本文引入了感知损失和边缘损失两种方法,以进一步提升头发几何重建精度并增强其结构细节。感知损失与边缘损失均应用于整个训练过程。

对于感知损失,采用VGG网络中间层特征构建感知损失,使用其特征图进行感知损失计算,并使用MSE损失函数计算渲染图像和真实图像在特征空间中的差异,如公式(5)和(6)。式中$\boldsymbol{F}_{\text{input}}$和$\boldsymbol{F}_{\text{target}}$分别是渲染图像和目标图像在所选VGG层的特征图,$\Phi(\cdot)$是通过VGG网络前向传播得到的特征,$N$是特征图的元素总数。

$$\boldsymbol{F}_{\text{input}}=\phi(\boldsymbol{I}_{\text{input}}),\boldsymbol{F}_{\text{target}}=\phi(\boldsymbol{I}_{\text{target}}) \tag{5}$$

$$\mathrm{L}_{\text{perceptual}}=\frac{1}{N}\sum_{i=1}^{n}(\boldsymbol{F}_{\text{input},i}-\boldsymbol{F}_{\text{target},i})^{2} \tag{6}$$

对于边缘损失,采用Canny算子提取图像边缘,其计算式为公式(7)和(8)。式中,$\boldsymbol{I}_{\text{pred}}$是渲染图像,$\boldsymbol{I}_{\text{gt}}$是目标图像,$\boldsymbol{E}_{\text{pred}}$和$\boldsymbol{E}_{\text{gt}}$分别是从渲染图像和目标图像中得到的Canny边缘图。$\boldsymbol{E}_{\text{pred},i,j,k}$和$\boldsymbol{E}_{\text{gt},i,j,k}$分别是渲染边缘图和目标边缘图的第$i$个通道、第$j$个高度和第$k$个宽度位置的值,$C\times H\times W$是图像的总像素数。

$$\boldsymbol{E}_{\text{pred}}=\text{Canny}(\boldsymbol{I}_{\text{pred}}),\boldsymbol{E}_{\text{gt}}=\text{Canny}(\boldsymbol{I}_{\text{gt}}) \tag{7}$$

$$\mathrm{L}_{\text{Edge}}=\frac{1}{C\times H\times W}\sum_{i=1}^{C}\sum_{j=1}^{H}\sum_{k=1}^{W}|\boldsymbol{E}_{\text{pred},i,j,k}-\boldsymbol{E}_{\text{gt},i,j,k}| \tag{8}$$

此外,遵循3DGS(Kerbl等,2023)的方法,训练中均包含L_1损失函数(Isola等,2017)和L_{ssim}损失函数(Wang等,2004),以便衡量渲染出来的图像与原始输入图像之间的差异,并引导三维高斯模型的优化。遵循Gaussian Haircut(Zakharov等,2024)的做法,L_{mask}用于将头发剪影和头发分割与真值进行匹配,$\mathrm{L}_{\text{orient}}$用于监督头发方向。本文最终得到的损失函数$\mathrm{L}_{\text{hair}}$可以写成如下形式:

$$\mathrm{L}_{\text{hair}}=\lambda_1\mathrm{L}_1+\lambda_2\mathrm{L}_{\text{ssim}}+\lambda_3\mathrm{L}_{\text{mask}}+\lambda_4\mathrm{L}_{\text{orient}}+\lambda_5\mathrm{L}_{\text{perceptual}}+\lambda_6\mathrm{L}_{\text{edge}} \tag{9}$$

式中$\lambda_1=0.8$,$\lambda_2=0.2$,$\lambda_3=\lambda_4=0.1$,$\lambda_5=\lambda_6=0.2$。取值依据为根据经验,加上小范围搜索,在经验值附近搜索得到最优取值。

3 实验及结果分析

3.1 实验设置

3.1.1 实施细节

本算法架构基于PyTorch实现,并在单个NVIDIA RTX 4090 GPU上进行训练。优化器采用

Adam，批大小为1，初始学习率和权重衰减均遵循3DGS（Kerbl等，2023）和Gaussian Haircut（Zakharov等，2024）的设置。模型在非结构化和结构化训练30000轮次，发丝粗级拟合训练40000轮次，发丝精细拟合训练50000轮次，在单个NVIDIA RTX 4090上完成训练大约需要13小时。本文完全沿用3DGS的对比和评估方法，并给出了几何的定性对比和纹理的定性与定量对比。类似现象在三维头发重建及三维重建领域中也被广泛观察到，例如Gaussian Haircut（Zakharov等，2024）工作中，几何的定性视觉效果显著改善。

3.1.2　对比方法

本文将本文方法与两类最先进的头发重建方法进行了对比评估。在多视图方法中，本文选择了结合经典头发曲线与对齐式三维高斯泼溅表示的多视图头发重建方法Gaussian Haircut（Zakharov等，2024）作为基线。这种方法采用多视图图像来重建头发的三维结构，通过对齐和优化头发曲线来恢复精细的发型细节。本文特别关注其对不同视角和不同光照条件下头发结构的建模能力，评估其在真实世界场景中的表现。

在单视图方法中，本文选择了两种具有代表性的最新方法进行比较：一是Im2Haircut（Sklyarova等，2025），该方法通过结合合成与真实数据训练的发型先验，利用高斯点渲染技术恢复头发的细节和内部结构。这种方法特别注重在单张图像上恢复真实感较强的头发纹理和形态。二是DiffLocks（Rosu等，2025），该方法使用单图像条件扩散Transformer模型进行训练。通过这种方式，DiffLocks在生成头发细节和逼真的发型结构方面表现出了出色的性能，尤其是在处理复杂发型和细节上具有优势。

3.1.3　数据集

本文使用真实数据集对本文的方法和基线方法进行了全面的评估。为了进行定性分析，本文在多个数据集上进行了测试，包括自建数据和公共数据集（Sklyarova等，2023）。自建数据是室内室外测试数据，由普通智能手机在室内室外复杂自然光照环境下拍摄，拍摄者手持手机面向模特头部，在同一水平面绕模特头部完整拍摄一圈（视频分辨率1080*1920，每秒30帧，时长约20秒）。自建数据涵盖不同发型，包括短波浪卷发、致密卷发、高卷曲

发型等多种发型，从中选择3组作为独立测试样例，平均照度范围约为300－500 lux。公共数据集则涵盖了六种发型室内捕捉视频，涵盖了从短发到卷发等不同类型的发型，平均照度范围约为20000 lux。通过这两组数据集，能够更加全面地评估本文方法的适应性和鲁棒性，针对室内和室外等低光场景来做训练，验证本文方法是否能够保持高效和准确的发型重建效果。这些实验为本文定性和定量评估提供了有力支持，并展示了本文方法在实际应用中的潜力。

3.2　定性评估

由于三维头发重建更关注结构合理性与视觉真实感，单一像素指标难以全面反映方法性能。因此，本文同时结合定量与定性结果进行综合评估。与Gaussian Haircut（Zakharov等，2024）方法相比，本文主要依靠从低光增强后的图像学习几何细节来

提升其几何形状。本文首先通过原始三维高斯泼溅进行非结构化训练，采用自适应密度控制来填补空洞，然后通过结构化训练进行针对头发区域的自适应密度控制来提升低精度3D高斯泼溅模型的精度。本文方法在最终的发型质量上取得了提升，如定性评估图3与图4所示。Gaussian Haircut倾向于创建与图像整体形状相符的头发，但对于高复杂度的头发而言则缺乏精细的细节和卷曲度，发尾部分的头发丝无法重建出正确方向以及长度。相比之下，本文方法能够为各种高复杂度的头发生成高质量且逼真的头发丝。

在纹理恢复方面，本文方法同样展现出相较于Gaussian Haircut（Zakharov等，2024）的优势，如图5所示。尽管本方法在训练过程中引入了低光增强

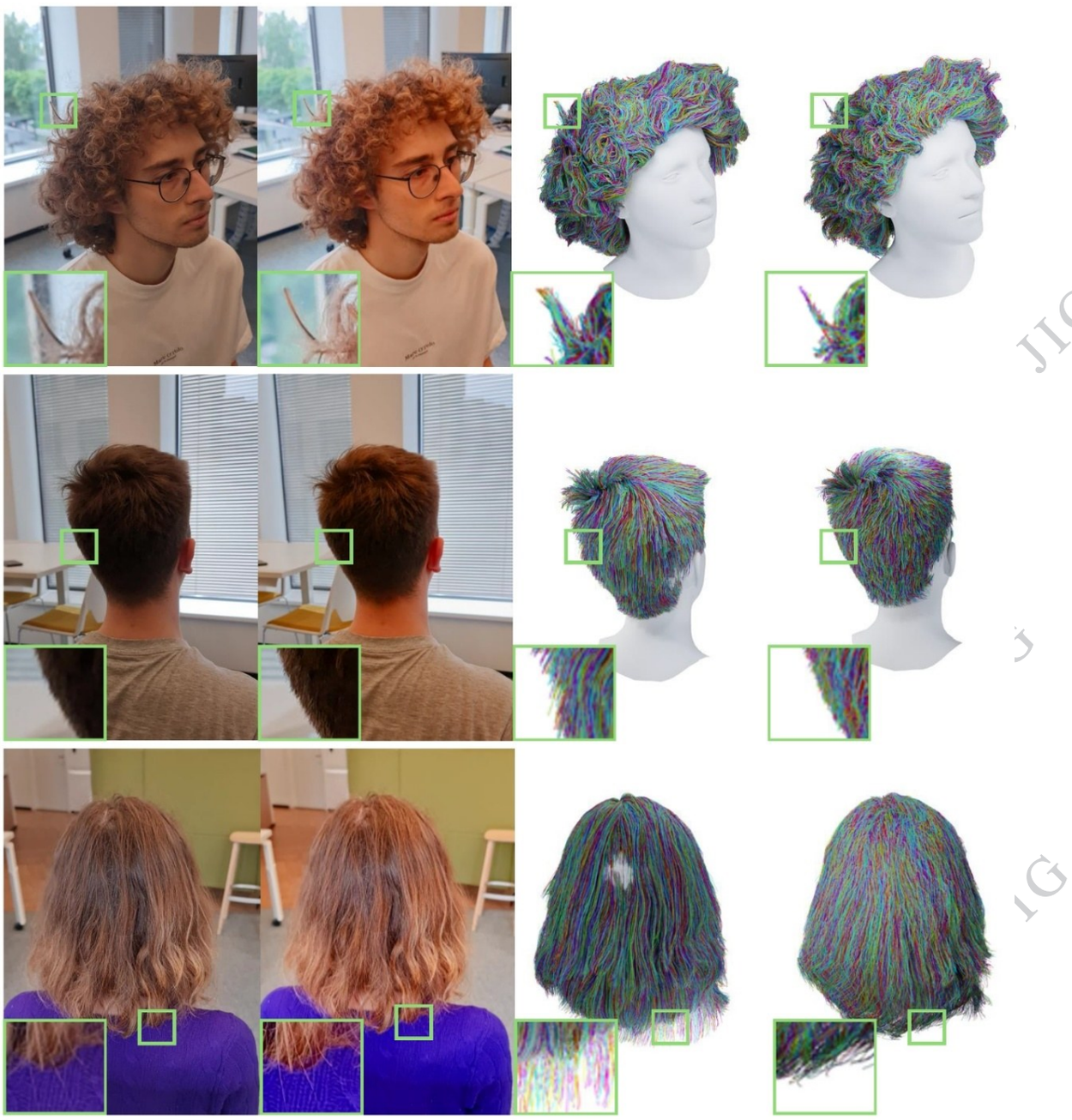
以增强几何学习的稳定性，但通过所提出的几何-纹理解耦优化策略，最终恢复的纹理在光照一致性与视觉逼真度上均与原始图像高度一致。相比之下，Gaussian Haircut在纹理重建上对光照变化的适应能力有限，而本文方法能够在保持真实感同时提升整体渲染质量。综合来看，本方法在纹理细节的稳定性与视觉一致性方面均优于Gaussian Haircut。

本文方法与当前最先进的单视图方法Im2Haircut（Sklyarova等，2025）和DiffLocks（Rosu等，2025）在定性评估方面进行了对比，结果如图6所示。Im2Haircut（Sklyarova等，2025）在正视图中能够较好地还原头发的整体形状与轮廓，然而在后视图中，头发表现为直发，偏离了原始发型的方向；而Dif-

fLocks(Rosu 等,2025)虽然在先验模型推理中呈现出一定的卷曲度,但其卷曲方向与原图严重不符,且不同发型的卷曲度差异不大,甚至有些原本应该偏直的头发也被表现得过于卷曲。相比之下,本文方法不仅能够保持与原图发型轮廓的一致性,还能够精准地重建原图的头发方向和卷曲度,从而达到理想的重建效果。受篇幅所限,更多详细结果请参见补充材料。

3.3 定量评估

由于真实头发几何真值难以获取,现有主流头发重建方法普遍缺乏统一的定量对比基准。为提升实验评价的权威性与可比性,本文引入3DGS领域广泛采用的量化指标开展定量评估,选取PSNR(PeakSignal-to-Noise Ratio)、LPIPS(Learned Perceptual Image Patch Similarity)和SSIM(Structural Similarity Index Measure)作为核心定量指标。PSNR衡量渲染图像与原始输入图像之间的像素级重建误差,以均方误差(MSE)为基础,反映整体光度一致性。更高的PSNR表示更低的像素差异和更好的重建质量。SSIM从亮度、对比度与结构三个方面评估

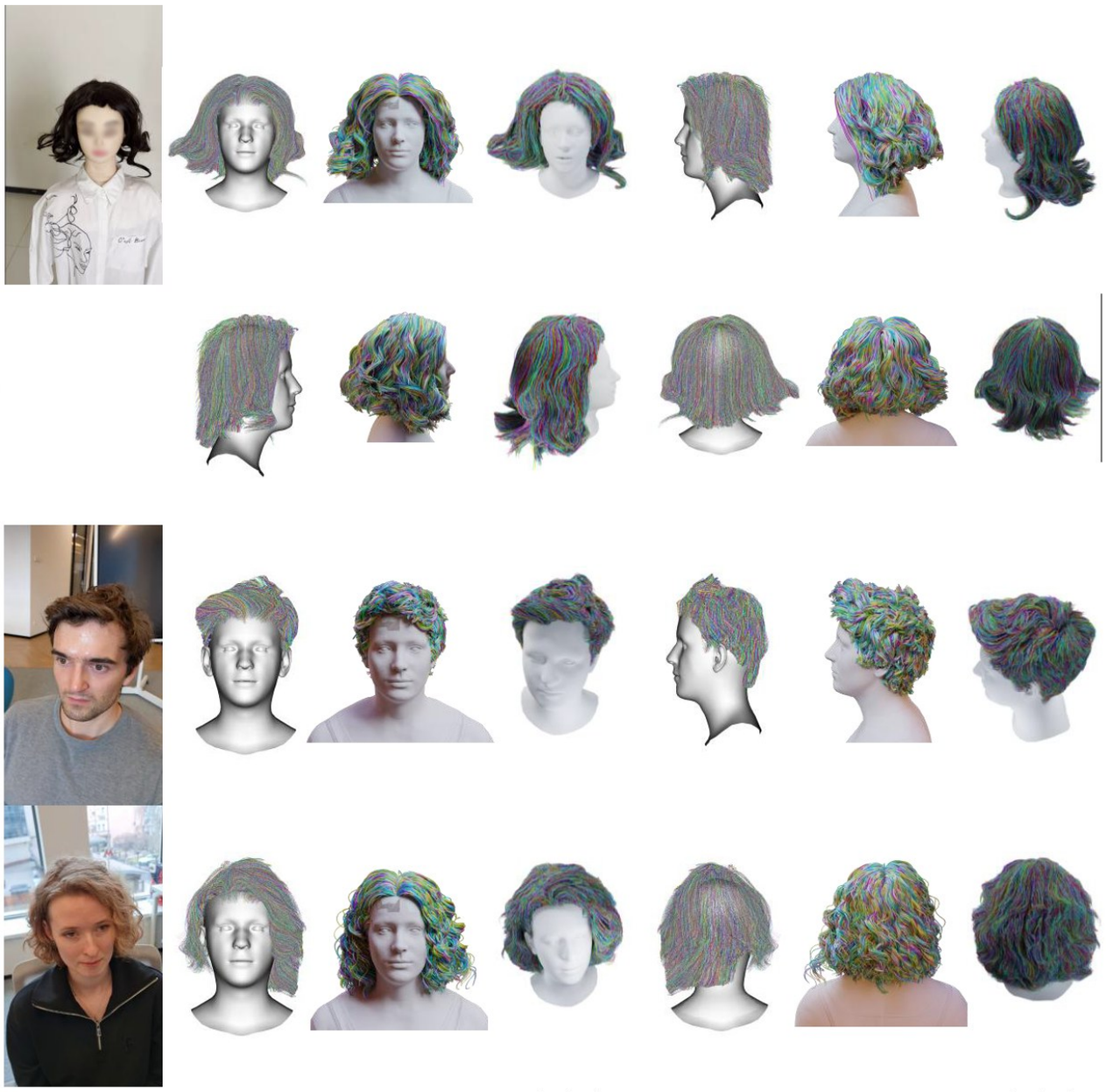


图像的结构相似性，相比像素误差更能体现感知一致性。更高的SSIM代表更好的结构保真度。LPIPS使用深度网络的特征差异来衡量图像的感知距离，更符合人类视觉对纹理与细节差异的判断。更低的LPIPS表示更接近人眼主观感受的图像质量。在无真值场景下，上述指标可从数值保真度、结构完整性与视觉感知一致性多个维度综合评价头发重建质量，为方法有效性提供客观量化依据。本文比较结果展现在表3。

由于Im2Haircut（Sklyarova等，2025）和DiffLocks（Rosu等，2025）未提供纹理重建结果，本文在纹理质量方面仅与Gaussian Haircut（Zakharov等，2024）方法进行对比。与Gaussian Haircut相比，本文方法的PSNR提高了0.54dB，SSIM提高了0.004，LPIPS降低了0.01。这些结果表明，本文方法在图像纹理质量上优于Gaussian Haircut，特别是在真实场景数据集中，表明本文算法能够更好地保持图像的光度一致性和结构保真度，同时减少感知上的细节差异。相比之下Gaussian Haircut在这些指标上表现较弱，尤其是在真实场景下。这些定量结果充分验证了本文方法在多种标准

3.4 消融实验

本文评估了方法中关键设计选择的影响如图7和图8所示。对于几何重建而言，移除低光增强和稠密化后，生成的发丝在细节上不足，导致头发的丰富度和真实度下降。发丝的结构会变得更加粗糙，缺乏真实的镂空度和细腻度。如果不进行边缘处理，发丝的轮廓会变得模糊，缺乏清晰的细节。在缺少感知损失约束的情况下，重建结果在发丝整体结构方面出现不同程度的退化，尤其在发丝密集区域与复杂发型区域，容易出现发丝走向不一致的问题。在复杂卷曲区域与发丝交错区域，该问题表现更为明显，影响后续几何优化的效果。如果不含VLM智能引导，在复杂光照条件下的结构化模型对应区域会因缺乏引导而导致直发变多，光照质量最差视角的发丝密度和方向连贯性显著下降，模型在复杂光照与高复杂度发型场景下的重建精度、完整性与鲁



棒性均出现显著下降，复杂发型几何拟合不足。而相比之下，当所有组件完整结合时，本文方法能够在细化和边缘处理方面都达到最优的效果，从而实现更加逼真和自然的发丝重建。对于纹理恢复而言，移除纹理恢复步骤后，头发的纹理和光泽度与原图不符，导致头发的质感不够自然，且整体效果显得生硬，影响整体的清晰度和真实感。当所有组件完整结合时，本文方法能够有效地恢复纹理，从而使发丝更加自然、真实和细腻，光照与原图相符。因此，完整的方法能够实现更精确的几何重建精度和更高质量的纹理恢复，确保最终效果的逼真度。

本文对方法中关键设计的定量指标进行了评估，如表4所示。在不含低光增强与稠密化的情况下，模型渲染效果有所下降，PSNR从33.344dB降至32.163dB，下降1.181，SSIM从0.908降至0.904，同时LPIPS从0.089上升至0.093，表明纹理图像结构一致性和感知质量均有所退化。该下降幅度相对较小，说明在缺少低光增强与稠密化策略的情况下，模型具备一定的几何重建性能，但整体性能低于完整模型。这进一步验证了低光增强与稠密化策略能够提升复杂光照条件下的几何重建精度与发丝精细度，从而有效改善最终渲染质量。当去除边缘损失时，模型性能出现一定程度下降，PSNR从33.344dB降至33.328dB，SSIM从0.908降至0.907，LPIPS从

表3　本文方法与Gaussian Haircut在自建数据和公共数据集(Sklyarova等,2023)上的定量比较

Table 3　Quantitative comparison of CURL and Gaussian Haircut on self-built data and public dataset.

	发型	Gaussian Haircut			CURL		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
自建数据	curly_0	31.9196dB	0.8992	0.0993	**33.2780dB**	**0.9074**	**0.0899**
	curly_1	27.5743dB	0.8863	0.0796	**28.6927dB**	**0.9019**	**0.0715**
	curly_3	31.7389dB	**0.9403**	0.0699	**32.3197dB**	0.9369	0.0752
公共数据集	jenya	30.9684dB	0.9028	0.0740	**31.0920dB**	**0.9068**	**0.0641**
	ksyusha	25.0173dB	0.7947	0.1417	**25.9699dB**	**0.8071**	**0.1201**
	person	26.8145dB	0.8321	0.1350	**27.2134dB**	**0.8417**	**0.1212**
	evgenia	31.2495dB	**0.9108**	0.0745	**31.2613dB**	0.9082	**0.0636**
	igor	**30.7001dB**	**0.9224**	0.0805	30.5147dB	0.9136	**0.0782**
	nastya	24.9644dB	**0.7401**	0.1772	**25.4628dB**	0.7391	**0.1529**
	平均值	28.9941dB	0.8699	0.1035	**29.5338dB**	**0.8736**	**0.0930**

注:加粗字体表示各行最优结果

评估指标下的优势,并体现了其更高的渲染质量。本文方法在个别样本上指标略有下降,主要原因在于在直发场景中,由于其几何结构相对简单、曲率变化较小且高频细节较少,优化过程更容易在较早阶段达到稳定状态。此时高斯分布的调整空间相对有限,在部分指标上表现为轻微波动。该现象主要源于直发场景本身的结构特性,而非方法退化,对方法整体性能影响较小。在复杂发型(如高卷曲发型、致密卷发)以及复杂光照条件下,本文方法在几何结构恢复与视觉质量方面均表现出优势,验证了方法的有效性与鲁棒性。(a)输入图像 (b)低光增强后 (c)不含光照处理和 (d)不含边缘损失 (e)不含感知损失 (f)不含VLM智 (g)CURL稠密化 能引导

图7　几何重建的消融实验。(a)输入图像;(b)低光增强图(辅助参考);(c)不含光照处理和稠密化;(d)不含边缘损失;(e)不含感知损失;(f)不含VLM智能引导;(g)本文方法CURL重建结果。绿框区域为局部放大细节。

Fig. 7　Ablation experiments for geometric reconstruction. (a) Input image; (b) low-light enhanced image (reference only); (c) w/o illumination processing and densification; (d) w/o edge loss; (e) w/o perceptual loss; (f) w/o VLM-guided densification; (g) the proposed CURL reconstruction results. Green boxes show zoomed-in regions for detail comparison.

0.089上升至0.091,整体下降幅度较小,表现为发丝边界区域细节减弱和局部结构略显模糊,说明边缘损失在三维发丝几何建模过程中能够提供有效的结构约束,增强发丝边界的清晰度和连续性,提升渲染图像的细节表现能力,验证了边缘损失在精细发丝重建中的重要作用。在移除感知损失后,模型的纹理恢复性能略有下降。相比完整方法,PSNR下降了0.072dB,SSIM下降了0.001,LPIPS略有上升。这表明感知损失能够在一定程度上提升纹理恢复质量,使发丝几何模型重建结果在感知层面更加接近真实图像。在去除VLM智能引导后,模型PSNR由33.344dB降至33.259dB,LPIPS由0.089升至0.090,SSIM保持0.908不变。实验结果表明,缺少VLM对光照质量与发型复杂度的感知引导,模型

表4 消融实验定量评估

Table 4　Quantitative evaluation of ablation experiments

方法	PSNR ↑	SSIM ↑	LPIPS ↓
不含低光增强与稠密化	32.163dB	0.904	0.093
不含边缘损失	33.328dB	0.907	0.091
不含感知损失	33.272dB	0.907	0.090
不含VLM智能引导	33.259dB	**0.908**	0.090
不含纹理恢复	15.054dB	0.786	0.172
本文方法	**33.344dB**	**0.908**	**0.089**

注：加粗字体表示各列最优结果

无法精准定位低光、逆光等劣质视角，也难以根据发型复杂度自适应分配高斯密度。这会导致光照缺陷区域高斯分布不足、复杂发型几何拟合不充分，进而使重建细节与结构连贯性略有下降，验证了VLM智能引导在提升复杂光照鲁棒性与发型适应性方面的关键作用。移除几何-纹理解耦优化方式对模型性能影响最为显著，PSNR从33.344dB大幅下降至15.054dB，SSIM从0.908降至0.786，LPIPS从0.089上升至0.172，表明渲染质量出现退化，纹理图像结构一致性与感知质量均显著下降，说明缺乏纹理恢复机制时，模型难以准确恢复外观信息。因此，验证了几何-纹理解耦优化提升纹理还原能力的技术优势。

3.5　局限性

经过基于VLM的稠密化处理后，头发几何结构的表现确实得到了改善，然而发尾部分仍未能准确拟合其弯曲度。这主要是由于结构化训练得到的低精度3D高斯泼溅模型在引导发丝粗级和精细拟合高精度3D高斯泼溅模型时存在不匹配的问题。

4　结　论

本文提出VLM闭环引导的自适应高斯头发重建框架，为复杂光照与高复杂度发型三维重建提供了新思路。通过结合视觉语言模型与3D高斯椭球稠密化策略，引入三阶段的自适应密度控制和光照质量评估，本文有效地解决了传统方法中存在的几何稀疏与方向图不精确的问题，从而实现了更为逼真和自然的头发重建效果。实验结果表明，本文方法在定性与定量评估中均优于现有的最先进方法，尤其在处理复杂发型（如致密卷发、高卷曲发型）时，能够更好地保持头发的细节和方向一致性。此外，本文所提出的几何-纹理解耦的优化策略，有效地平衡了几何精确性与颜色真实性，确保了最终渲染的自然度。尽管如此，本文方法仍存在一定的局限性，例如在某些复杂发尾的弯曲度拟合上仍有待提升。未来的工作可以进一步探索更精细的优化策略，以及更加鲁棒的光照适应机制，增加动态头发、极端遮挡、移动端部署等方向，从而在更加复杂和多变的环境中，实现更高水平的头发重建。综上所述，本文的研究为头发重建领域提供了新的思路和方法，不仅提升了重建质量，也为未来相关研究的深入提供了重要的参考价值。

参考文献(References)

Beeler T, Hahn F, Bradley D, Bickel B, Beardsley P, Ghosh A, et al. 2012. High-quality single-shot capture of facial geometry. ACM Transactions on Graphics, 31(4): 1-9 [DOI: 10.1145/2185520.2185529]

Fan L, Xu P, Zhang Z and Wang X. 2025. VLM-3R: visual-language models for 3D reconstruction reasoning//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE(to appear)

Guan S, Li S and Wu Z. 2025. Normal-anchored gaussian representations for stable 3D reconstruction//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Paris: IEEE (to appear)

Isola P, Zhu J Y, Zhou T and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 1125-1134 [DOI: 10.1109/CVPR.2017.632]

Kerbl B, Kopanas G, Leimkühler T and Drettakis G. 2023. 3D gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics, 42(4): 139:1 - 139:14 [DOI: 10.1145/3592433]

Lin Y, Lin Q, Tang F, Tang L and Wang S J. 2011. Realistic 3D head reconstruction based on feature points matching between photo images and 3D generic model. Journal of Image and Graphics, 16(10): 1876-1882 (林源, 林茜, 汤锋, 唐亮, 王生进. 2011. 匹配图像与3维模型特征点的真实感3维头重建. 中国图象图形学报, 16(10): 1876-1882) [DOI: 10.11834/jig.20111012]

Liu H T, Li C Y, Wu Q Y and Lee Y J. 2023. Visual instruction tuning [EB/OL]. [2023-04-17]. https://arxiv.org/abs/2304.08485 [DOI: 10.48550/arXiv.2304.08485]

Mai A, Hedman P, Kopanas G, Verbin D, Futschik D, Xu Q, et al. 2025. EVER: exact volumetric ellipsoid rendering for real-time view synthesis//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Honolulu: IEEE: 4930-4939 [DOI: 10.48550/arXiv.2410.01804]

Mai S, Zhang H and Sun J. 2025. Ellipsoid volume rendering for improved gaussian splatting//Proceedings of the European Conference on Computer Vision (ECCV). Milan: Springer(to appear)

Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2021. NeRF: representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1): 99 - 106 [DOI: 10.1145/3503250]

Nichol A, Dhariwal P, Ramesh A, Shyam P, Mishkin P, McGrew B, et al. 2022. GLIDE: towards photorealistic image generation and editing with text-guided diffusion models//Proceedings of the 39th International Conference on Machine Learning (ICML). Baltimore: PMLR: 16784-16804 [DOI: 10.48550/arXiv.2112.10741]

Pan Y M, Nießner M and Kirschstein T. 2025. HairGS: hair strand reconstruction based on 3D gaussian splatting//Proceedings of the 36th British Machine Vision Conference (BMVC 2025). Sheffield: BMVA [DOI: 10.48550/arXiv.2509.07774]

Ramesh A, Dhariwal P, Nichol A, Chu C and Chen M. 2022. Hierarchical text-conditional image generation with CLIP latents [EB/OL]. [2022-04-13]. https://arxiv.org/abs/2204.06125 [DOI: 10.48550/arXiv.2204.06125]

Rosu R A, Wu K Y, Feng Y, Zheng Y Y and Black M J. 2025. DiffLocks: generating 3D hair from a single image using diffusion models//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 10847-10857 [DOI: 10.48550/arXiv.2505.06166]

Schönberger J L and Frahm J M. 2016. Structure-from-motion revisited//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 4104-4113 [DOI: 10.1109/CVPR.2016.445]

Sklyarova V, Chelishev J, Dogaru A, Medvedev I, Lempitsky V and Zakharov E. 2023. Neural haircut: prior-guided strand-based hair reconstruction//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Paris: IEEE: 19762-19773 [DOI: 10.1109/ICCV51070.2023.01810]

Sklyarova V, Zakharov E, Prinzler M, Becherini G, Black M J and Thies J. 2025. Im2Haircut: single-view strand-based hair reconstruction for human avatars//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Honolulu: IEEE: 10656-10665 [DOI: 10.48550/arXiv.2509.01469]

Sun F Y, Liu W Y, Gu S Y, Lim D, Bhat G, Tombari F, et al. 2025. LayoutVLM: differentiable optimization of 3D layout via vision-language models//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 29469-29478 [DOI: 10.48550/arXiv.2412.02193]

Sun X Y, Lee J C, Rho D, Ko J H, Ali U and Park E B. 2024. F-3DGS: factorized coordinates and representations for 3D gaussian splatting//Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne: ACM: 7957-7965 [DOI: 10.1145/3664647.3681116]

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]

Wei Y Y, Mao T Y, Li B A, Wang F, Li F and Zhang Z, et al. 2025. Visual and large multimodal models promote image restoration and enhancement: research progress. Journal of Image and Graphics, 30(5): 1197-1219 (韦炎炎, 毛天一, 李柏昂, 王飞, 李锋, 张召, 等. 2025. 视觉模型及多模态大模型推进图像复原增强研究进展. 中国图象图形学报, 30(5): 1197-1219) [DOI: 10.11834/

jig.240436］

Wu H N, Zhang Z C, Zhang W X, Chen C F, Liao L and Li C Y, et al. 2024. Q-Align: teaching large multimodal models for visual scoring via discrete text-defined levels//Proceedings of the 41st International Conference on Machine Learning (ICML). Vienna: PMLR［DOI: 10.48550/arXiv.2312.17090］

Yan Q S, Feng Y X, Zhang C, Pang G S, Shi K B and Wu P, et al. 2025. HVI: a new color space for low-light image enhancement//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 5678-5687［DOI: 10.1109/CVPR52734.2025.00533］

Yang H, Chen R, An S P, Wei H and Zhang H. 2023. The growth of image-related three dimensional reconstruction techniques in deep learning-driven era: a critical summary. Journal of Image and Graphics, 28(8): 2396-2409 (杨航, 陈瑞, 安仕鹏, 魏豪, 张衡. 2023. 深度学习背景下的图像三维重建技术进展综述. 中国图象图形学报, 28(8): 2396-2409)［DOI: 10.11834/jig.220376］

Zakharov E, Sklyarova V, Black M J, Nam G, Thies J and Hilliges O. 2024. Human hair reconstruction with strand-aligned 3D Gaussians//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer: 409-425［DOI: 10.1007/978-3-031-72640-8_23］

Zhou Y, Hu L W, Xing J, Chen W K, Kung H W and Tong X, et al. 2018. HairNet: single-view hair reconstruction using convolutional neural networks//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer: 235-251［DOI: 10.1007/978-3-030-01252-6_15］

Zheng Y J, Qiu Y D, Jin L Y, Ma C Y, Huang H B and Zhang D, et al. 2024. Towards unified 3D hair reconstruction from single-view portraits//Proceedings of ACM SIGGRAPH Asia 2024 Conference Papers. Tokyo: ACM［DOI: 10.1145/3680528.3687597］