中图法分类号:TP391; TP18　　**文献标识码**: A　　**文章编号**: 1006-8961(XXXX)XX-0001-16

论文引用格式:Liu Congcong, Xue Han, Fang Zhijun. Text-Driven geometry and texture editing of 3D garments[J/OL]. Journal of Image and Graphics, XXXX:1-16. DOI: 10.11834/jig.260486. (刘聪聪, 薛涵, 方志军. 文本驱动的三维服装几何与纹理编辑[J/OL]. 中国图象图形学报, XXXX:1-16. DOI: 10.11834/jig.260486.)[DOI:10.11834/jig.260486]

文本驱动的三维服装几何与纹理编辑

刘聪聪[1],薛涵[1],方志军[1,2]

1. 东华大学信息与智能科学学院,上海　201620; 2. 复旦大学可信具身智能研究院,上海　200438

摘　要:**目的**　可编辑三维服装资产在数字时尚、虚拟试衣、游戏动画和三维内容创作中具有重要应用价值。现有文本驱动服装编辑方法多分别处理几何与纹理,局部编辑时直接生成完整服装表示易引起非目标参数漂移、目标修改不准确和仿真不稳定。为此,本文提出文本驱动的三维服装几何与纹理编辑框架。**方法**　本文构建两阶段顺序式框架。给定穿衣人物图像和文本编辑指令,本框架首先将用户指令拆分为几何与纹理编辑指令。几何编辑分支在参数化缝纫图空间中确定可编辑字段,由语言模型生成字段级编辑计划,再通过确定性程序完成路径、范围及方向检查和非目标变化回滚,经缝合与物理仿真生成三维网格。纹理编辑分支依次接入现有图像编辑模型、服装区域提取和图像引导纹理迁移模块,实现颜色、材质或图案编辑。此外,本文构建了GeometryEdit-10K数据集,并在GeometryEdit-1200上开展定量评价。**结果**　在GeometryEdit-1200上,本文方法的L1范数倒角距离(L1-norm Chamfer distance,CD-L1)为3.501,阈值为1.0的F分数(F-score at a threshold of 1.0,F@1)为0.820;较AIpparel,前者下降48.9%,后者提高23.1%。目标命中率为0.568,非目标漂移率为0.013;较ChatGarment,二者分别提高55.2%和降低97.9%。纹理实验表明,在所测试样例和所采用指标下,本文流程获得了更好的编辑方向一致性。**结论**　本文提出的字段约束局部编辑流程能够改善目标字段执行并减少非目标字段变化,增强局部几何编辑的可控性。通过顺序衔接现有纹理生成与迁移模块,该框架能够形成由服装结构编辑到外观编辑的完整流程。

关键词:三维服装;文本驱动编辑;局部几何编辑;参数化缝纫图;纹理迁移

Text-Driven geometry and texture editing of 3D garments

Liu Congcong[1], Xue Han[1], Fang Zhijun[1,2]

1. School of Information and Intelligent Science, Donghua University, Shanghai 201620, China; 2. Institute of Trusted Embodied Intelligence, Fudan University, Shanghai 200438, China

Abstract: **Objective** Editable three-dimensional (3D) garments are increasingly required in digital fashion, virtual try-on, game animation, and 3D content creation, where users expect to modify both garment structure and surface appearance through intuitive text instructions. Existing text-driven garment methods generally focus on either geometry or texture. Geometry-oriented methods can change garment shapes, but their results are not always designed for subsequent appearance editing; texture-oriented methods usually assume a fixed mesh and therefore do not address structural changes. This separation limits the ability to process an instruction that simultaneously describes sleeve length, neckline shape, silhouette, color, material, or pattern. Local geometry editing is particularly challenging because a request may concern only one garment part, whereas directly regenerating the complete garment representation can modify unrelated parameters, miss the intended target, damage sewing relationships, or produce patterns that cannot be simulated reliably. The objective of this study is therefore to improve the controllability of local garment geometry editing while establishing a complete workflow

收稿日期:2026-08-04;**修回日期**:2026-10-29

© 中国图象图形学报版权所有

that connects structural editing with appearance editing. The main methodological emphasis is placed on reliable field-level geometry modification, whereas the texture branch is implemented by sequentially integrating existing appearance-editing modules. **Method** A sequential two-stage text-driven framework is proposed. Given an image of a clothed person and a natural-language instruction, an instruction parsing module first separates the request into geometry and texture components for sequential processing. The geometry branch operates in the parametric sewing-pattern space. It analyzes the current garment structure and geometry instruction to determine an editable boundary containing fields that are relevant to the requested change and valid for the current garment. Instead of regenerating the complete JavaScript object notation (JSON) representation or an entire sewing pattern, the language model generates only a field-level editing plan within this boundary. A deterministic executor then validates parameter paths, operation types, numerical or categorical value domains, and whether each change follows the direction expressed by the instruction. Valid operations are applied to the source representation, while changes outside the executable field set are rolled back. The verified changes are mapped from the simplified representation to the detailed garment design, supporting local modifications to sleeve length, neckline shape, garment length, trouser silhouette, skirt length, hem shape, and related structural attributes. The edited design is converted into a sewing pattern and physically simulated to obtain the corresponding 3D garment mesh. In the texture branch, the geometrically edited mesh serves as the carrier of the final appearance. An existing image editing model modifies the clothed-person image according to the texture instruction, after which the garment region is extracted to form an appearance reference. An existing image-guided texture transfer module then applies the specified color, material, or pattern to the edited mesh. GeometryEdit-10K is constructed as an extended resource for local garment geometry editing. GeometryEdit-1200, containing 1,200 source-target editing pairs with field-level annotations and simulation verification, is used as the principal quantitative benchmark. Evaluation is conducted at complementary levels. Mesh-level measurements compare predicted and target garments using the L1-norm Chamfer distance (CD-L1) and the F-score at a threshold of 1.0 (F@1). Field-level measurements include the target-hit rate, denoted as Target Hit, which evaluates whether required fields reach their target values or make sufficient progress in the correct direction, and the non-target drift rate, which measures unintended changes to unrelated fields. Texture evaluation uses 50 cases covering color, material, global pattern, and local pattern changes. Each result is rendered from the front, back, left, and right views. Texture quality is evaluated using a contrastive language-image pre-training (CLIP)-based text-image alignment score, denoted as CLIPScore, and Directional CLIP Similarity, which measures whether the visual change agrees with the semantic change specified by the instruction. **Result** On GeometryEdit-1200, the proposed geometry editing method achieves a CD-L1 of 3.501 and an F@1 of 0.820. Compared with AIpparel under the same evaluation protocol, CD-L1 is reduced by 48.9%, while F@1 is improved by 23.1%. These results indicate that the simulated meshes produced by the proposed method are geometrically closer to the target garments. At the JSON field level, the method obtains a Target Hit of 0.568 and a non-target drift rate of 0.013. Compared with ChatGarment, Target Hit is improved by 55.2%, and the non-target drift rate is reduced by 97.9%. The combination of a higher target-hit rate and a lower non-target drift rate shows that the improvement does not arise merely from making fewer changes; rather, the method more frequently executes the required edits while preserving unrelated garment parameters. The mesh-level and field-level results therefore provide complementary evidence of improved local-editing controllability. Ablation experiments further show that removing the editable-boundary constraint reduces target completion and field-selection accuracy, whereas direct full-JSON generation produces a more pronounced decline in both field-level and geometric performance. The texture experiments reveal a trade-off between absolute target-text alignment and editing-direction consistency. InTeX obtains the highest CLIPScore, indicating strong alignment between its edited results and the target texture descriptions. The proposed sequential workflow achieves the highest Directional CLIP Similarity under the tested cases and adopted metrics, indicating that its visual changes more closely follow the intended transition from the source appearance to the target appearance. This result supports the effectiveness of the sequential workflow in executing the requested texture changes, but does not imply superiority over the comparison methods on every texture metric. **Conclusion** The proposed framework combines field-constrained local geometry editing with a sequential appearance-editing workflow. By restricting language-model outputs to field-level plans and using deterministic validation, execution, direction checking, and rollback, the geometry branch improves target-field completion and non-target preservation while retain-

© 中国图象图形学报版权所有

ing compatibility with sewing-pattern generation and physical simulation. The subsequent integration of existing image editing, garment-region extraction, and image-guided texture transfer modules extends the result from structural modification to color, material, and pattern editing. The framework therefore provides a controllable path from a clothed-person image and text instruction to an edited, textured 3D garment. Current limitations remain related to the coverage of parametric garment templates and the quality of appearance references. Extending support to irregular and multilayer garments, improving complex garment combination modeling, and achieving more precise control over localized patterns, fabric materials, and multiview texture consistency constitute important directions for future work.

Key words: 3D garments; text-driven editing; local geometry editing; parametric sewing patterns; texture transfer

论文引用格式:[DOI:10.11834/jig.260486]

0 引　言

数字服装资产在虚拟试衣、数字时尚、游戏动画和数字内容创作中具有重要应用价值。与一般刚性三维模型不同,服装由二维布片、边界轮廓和缝合关系共同决定,并在人体周围产生复杂的柔性形变。因此,服装编辑不仅需要改变可见外观,还需要保证编辑后的结构能够正确缝合、披挂和仿真。已有三维穿衣人体重建研究表明,服装语义、层次和褶皱等视觉信息对服装形态恢复具有重要作用(普骏程等,2022),但这类方法主要关注从图像中恢复服装表面特征,难以直接满足数字服装设计中对结构化缝纫图和局部编辑的需求。

缝纫图为三维服装建模提供了一种紧凑且可解释的结构表示。它通过二维布片和缝合关系描述服装从平面纸样到三维形态的生成过程,更接近真实制衣流程。GarmentCode(Korosteleva 和 Sorkine-Hornung,2023)将服装表示为参数化程序,并能够生成可仿真的缝纫图和三维服装。近年来,基于大语言模型和视觉语言模型的服装生成方法进一步推动了该方向的发展。ChatGarment(Bian 等,2025)和AIpparel(Nakayama 等,2025)进一步将多模态模型用于结构化服装生成与编辑。这些方法显著降低了数字服装创建和编辑的门槛,但在局部编辑场景中仍存在不足:当用户只希望修改袖长、领口、腰线或裙长等目标部件时,完整预测编辑后的服装表示容易导致非目标区域发生变化,从而影响编辑可控性和仿真稳定性。

除几何结构外,纹理也是数字服装资产的重要组成部分。颜色、材质和表面图案直接决定服装的视觉风格。现有三维纹理生成方法通常在给定几何模型上进行文本或图像引导的纹理合成,例如,TEXTure(Richardson 等,2023)通过多视角迭代绘制生成三维纹理,Hunyuan3D 2.1(Hunyuan3D Team 等,2025)支持图像引导的高保真纹理与基于物理的渲染(physically based rendering,PBR)材质生成。此类方法通常以既定几何为输入,因此对于同时包含结构与外观要求的文本指令,还需将几何编辑结果与后续外观参考生成、三维纹理迁移衔接,使纹理编辑作用于结构修改后的服装网格。

为解决上述问题,本文提出一种文本驱动的三维服装几何与纹理编辑框架。该框架由几何编辑分支和纹理编辑分支构成。给定穿衣人物图像和文本编辑指令,本框架首先通过指令解析模块将用户指令拆分为几何编辑指令和纹理编辑指令。几何编辑分支依据源服装的参数化缝纫图和JavaScript对象表示法(JavaScript object notation,JSON)确定当前可编辑字段。本文使用JSON记录服装部件、参数路径、字段取值和部件间结构关系,使服装几何编辑能够从文本指令进一步落实到具体参数字段上。随后,利用语言模型生成字段级编辑计划,再通过确定性程序完成路径验证、字段执行、参数范围处理、方向一致性检查、非目标字段回滚和详细设计校正,得到编辑后的缝纫图与仿真网格。纹理阶段以该网格为几何载体,连接现有图像编辑模型、服装区域提取和图像引导三维纹理迁移模块,生成目标颜色、材质或图案对应的三维服装结果。

本文主要贡献如下:1)提出一种面向参数化缝纫图的细粒度几何编辑方法。该方法根据当前服装结构和几何编辑指令动态确定可编辑字段,生成字段级编辑计划并回滚非目标变化,实现部件级几何修改并减少无关参数漂移。2)构建具有字段级标注和仿真验证的GeometryEdit-1200评测集,并在此基础上构建包含10383个有效案例的GeometryEdit-

© 中国图象图形学报版权所有

10K扩展数据集，记录源—目标简化JSON、字段变化、缝纫图和三维网格，用于评价局部编辑的目标执行、非目标漂移、几何误差和仿真有效性。3）构建从几何编辑到外观编辑的顺序式系统流程，将编辑后的仿真网格接入现有图像编辑模型、服装区域提取和图像引导纹理迁移模块，并通过纹理实验评价颜色、材质和图案编辑的应用效果。

1 相关工作

1.1 缝纫图表示与三维服装建模

缝纫图通过二维布片、边界轮廓和缝合关系描述服装结构，能够显式表达服装部件并支持编辑与物理仿真。国内研究主要从服装视觉信息表示和穿衣人体重建角度展开，普骏程等人（2022）利用语义、明暗和高频特征增强服装层次与褶皱表达；陈鸿鹄等人（2024b）指出复杂服装形变和高保真细节恢复仍是三维穿衣人体重建中的关键问题。这些工作有助于恢复服装外观和局部细节，但通常不直接提供可编辑的参数化缝纫图表示。

围绕缝纫图的学习式方法主要关注从图像观测恢复布片结构和缝合关系。SewFormer（Liu 等，2023）将单张穿衣人物图像到缝纫图的恢复建模为结构化预测问题，PanelFormer（Chen 等，2024a）则从二维服装图像中恢复布片轮廓和缝合拓扑，DressWild（Tao 等，2026）则研究了野外图像中的缝纫图生成。这些工作表明，缝纫图可作为图像观测与三维服装之间的结构化中间表示。另一方面，GarmentCode（Korosteleva 和 Sorkine-Hornung，2023）将服装设计参数组织为可执行程序，使服装能够从参数空间稳定转换为缝纫图和仿真后三维网格；GarmentX（Guo 等，2027）进一步采用自回归参数化表示生成高保真三维服装，并通过兼容GarmentCode的结构化表示提高服装生成的有效性和可编辑性。

近年来，文本和多模态条件被引入缝纫图生成。DressCode（He 等，2024）采用文本引导的自回归方式生成服装缝纫图。GarmageNet（Li 等，2025b）和Design2GarmentCode（Zhou 等，2025）进一步探索了从设计概念出发的多模态服装建模和程序合成。GarmentDiffusion（Li 等，2025c）利用多模态扩散Transformer从文本、图像和不完整缝纫图等条件生成向量化缝纫图。上述方法扩展了文本和多模态条件下的缝纫图生成方式，但对局部编辑中的目标字段约束和非目标字段保持仍缺少专门设计。

1.2 文本驱动服装编辑

三维服装编辑可在网格、神经表示、三维高斯、程序参数或缝纫图等不同表示空间中进行。Text2Mesh（Michel 等，2022）通过文本提示对给定三维网格进行颜色和局部几何风格化；GarmentDreamer（Li 等，2025a）结合三维高斯表示和生成先验合成具有几何与纹理细节的服装资产。SemanticGarment（Wang 等，2025）通过语义控制三维高斯服装的生成和编辑，GarmentCrafter（Wang 等，2026）从单视角服装图像进行三维服装重建和编辑，并通过渐进式新视角合成提高多视角一致性。这类方法增强了文本对三维服装外观和几何的控制，但其表示通常不显式对应布片及缝合关系。

基于结构化表示的方法更适合可仿真服装编辑。AIpparel（Nakayama 等，2025）将源缝纫图、编辑指令和目标缝纫图组织为编辑样本，使多模态模型能够在缝纫图空间中执行文本指令。ChatGarment（Bian 等，2025）利用视觉语言模型生成结构化JSON，并转换为缝纫图和仿真网格。Garment Particles（Nakayama 等，2026）通过二维—三维对称服装表示组织服装结构，使编辑过程能够同时考虑二维纸样与三维形态之间的对应关系。这些方法验证了语言模型执行服装结构编辑的可行性，但完整生成目标表示仍可能改变非目标部件。相比之下，本文通过字段级约束和非目标变化回滚提高局部编辑的可控性。

1.3 三维服装纹理编辑

三维纹理生成通常利用二维扩散模型和多视角渲染，将文本或图像条件投影到三维表面。TEXTure（Richardson 等，2023）通过多视角迭代绘制实现文本引导三维纹理生成；InTeX（Tang 等，2024）利用深度感知修复框架支持交互式文本进行纹理编辑，使用户能够对三维模型进行局部重绘。这些方法提高了通用三维物体纹理生成质量，但直接用于服装时仍可能受到织物形变、图案定位和多视角一致性的影响。

服装纹理生成需要进一步考虑服装结构、材质属性和二维纹理坐标展开方式。FabricDiffusion（Zhang 等，2024）将野外服装图像中的外观迁移至三维服装。TexGarment（Liu 等，2025）利用三维结构

© 中国图象图形学报版权所有

引导扩散变换器生成具有跨视角一致性的服装纹理，以缓解服装表面复杂变形导致的纹理不连续问题。Hunyuan3D 2.1（Hunyuan3D Team 等，2025）提供图像条件下的高保真三维资产与PBR材质生成能力，为图像引导纹理迁移和渲染展示提供了可用工具。NI-Tex（Shan 等，2026）研究了非等距条件下的图像驱动服装纹理生成。这些方法能够提升服装外观质量，但多数工作将纹理生成作为给定几何后的独立任务，并不直接处理缝纫图级几何编辑。

因此，本文不重新设计纹理生成模型，而是在字段约束的局部几何编辑之后，以编辑后的仿真网格为载体，顺序连接图像编辑模型与图像引导的纹理迁移，使结构修改结果能够进入外观编辑流程。

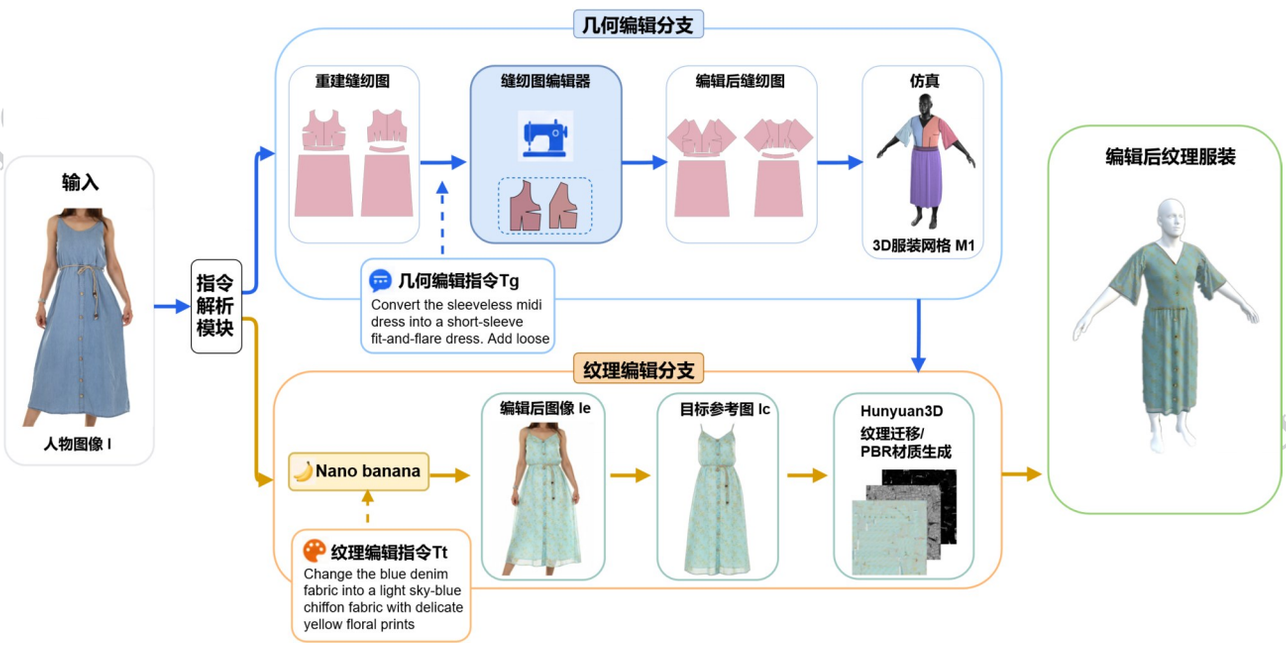
2 方　法

2.1 方法总述

如图1所示，本文提出一种文本驱动的三维服装几何与纹理编辑框架，包括几何编辑和纹理编辑两个分支。给定一张穿衣人物图像和一条文本编辑指令，本框架首先通过指令解析模块将文本指令拆分为几何编辑指令和纹理编辑指令；当指令仅包含其中一类编辑要求时，另一提示置空。其中，几何编

图1 几何与纹理编辑框架整体流程图

Fig. 1 Overall Workflow of the Geometry and Texture Editing Framework



辑提示词用于描述服装结构和版型变化，包括袖长、领口、衣长、裤长、裤型、裙长、裙摆、开衩和腰头等编辑内容；纹理编辑提示词用于描述服装外观变化，包括颜色、材质、织物纹理、表面图案、局部印花和装饰元素等编辑内容。

整体流程采用“先几何、后纹理”的顺序式设计。第1阶段基于ChatGarment（Bian 等，2025）重建服装的缝纫图和结构化参数，并在参数化缝纫图空间中引入缝纫图样编辑器，通过可编辑字段解析、字段级编辑计划和非目标变化回滚，对袖长、领口、腰线、裤长和裙摆等局部结构进行可控编辑，随后经缝合与物理仿真生成编辑后的三维服装网格。第2阶段以该网格为几何载体，利用图像编辑模型生成目标纹理参考图，并通过Hunyuan3D 2.1（Hunyuan3D Team 等，2025）进行图像引导纹理迁移和PBR材质生成。最终，框架输出编辑后的缝纫图、三维服装网格和带纹理渲染结果，实现从服装结构编辑到外观编辑的完整流程。

2.2 几何编辑

为解决在细粒度局部编辑中，直接预测完整JSON或缝纫图可能造成目标字段修改不足及非目标字段漂移的问题，几何编辑模块不直接自由生成完整目标服装，而是先确定当前服装中允许被修改的字段，再在受约束的字段集合内执行编辑，从而实现对目标区域的精细控制并抑制非目标变化。如图2所示。

© 中国图象图形学报版权所有

在本文中，源服装采用简化JSON和详细设计两种结构化表示。简化JSON保留与语义编辑相关的服装部件层级及叶子字段，用于候选字段检索、编辑计划生成和执行结果检查；详细设计文件包含生成缝纫规格、缝纫图和仿真网格所需的完整参数。设源简化JSON为$\boldsymbol{J}_0$，源详细设计为$\boldsymbol{D}_0$，几何编辑文本为T_g。本文以$\boldsymbol{\Pi}$表示字段级编辑计划，以$\boldsymbol{J}_1$表示确定性执行及一致性检查后的临时JSON，以$\boldsymbol{J}_1$表示完成非目标回滚后的简化JSON，校正后的详细设计、二维缝纫图和仿真网格分别记为$\boldsymbol{D}_1$、$\boldsymbol{P}_1$和$\boldsymbol{M}_1$。

2.2.1　字段级编辑计划生成

缝纫图编辑器首先对源简化JSON $\boldsymbol{J}_0$进行结构解析，识别当前服装实际包含的有效部件和叶子路径，并结合T_g与结构化服装描述定位目标部件。针对不同参数化模板之间的字段差异，候选字段必须对应当前JSON中真实存在的叶子路径；容器节点以及当前服装结构中不存在的字段不进入执行阶段。

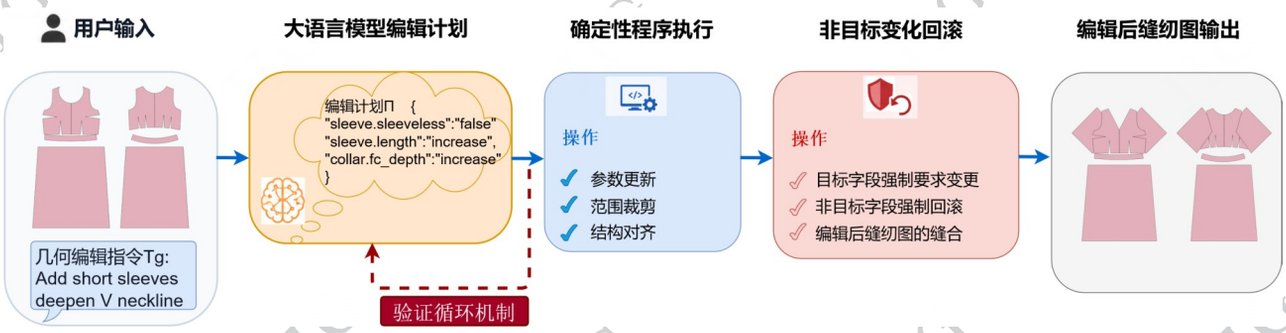
为建立文本语义与服装参数之间的联系，缝纫图编辑器将候选字段的参数路径、当前取值、数据类型、合法范围和几何含义组织为字段语义知识卡片并通过服装参数设计模式记录字段类型、取值范围和枚举约束。二者共同将语言模型的搜索范围缩小至当前服装中与目标部件相关且合法的候选字段。

为形式化描述字段级候选边界，语言模型在候选范围内生成字段级编辑计划，并依次完成路径、类型、取值范围和语义一致性检查。在此基础上，最终可执行路径集合定义为

$$\boldsymbol{E}=\boldsymbol{V}\cap\left(\left[(\boldsymbol{T}\cup\boldsymbol{L})\cap\boldsymbol{U}(\boldsymbol{J}_0)\right]\setminus\boldsymbol{K}\right) \tag{1}$$

图2 几何编辑流程图

Fig. 2 Geometry Editing Workflow



式中，$\boldsymbol{E}$为最终可执行路径集合；$\boldsymbol{V}$为编辑计划$\boldsymbol{\Pi}$中通过路径、类型、取值范围和语义一致性检查的操作路径集合；$\boldsymbol{T}$为文本指令直接对应的目标候选路径集合；$\boldsymbol{L}$为实现特定结构变化时需要同步处理的联动路径集合；$\boldsymbol{U}(\boldsymbol{J}_0)$为源简化JSON $\boldsymbol{J}_0$中实际存在的叶子路径全集，其中$\boldsymbol{U}(\cdot)$为叶子路径提取函数；$\boldsymbol{K}$为锁定、规避或当前参数化模板不支持的路径集合。式(1)将可执行字段限制为当前结构中真实存在、与目标或必要联动相关、通过计划验证且不属于受保护路径的叶子字段。联动路径仅在预定义结构关联被激活或编辑计划确有需要时开放。

字段级编辑计划$\boldsymbol{\Pi}$记录必要路径、可选联动路径、拒绝路径及其字段操作。每项操作包含目标路径、操作类型，并根据指令给出绝对目标值、类别取值或相对变化方向。对于“加长”“缩短”“加深”和“减小”等相对编辑，编辑计划记录相应的增减方向；对于包含明确样式语义的指令，则利用预定义语义锚点将样式描述映射为代表性参数值。

初始计划生成后，系统依次检查字段路径是否存在、目标操作是否完整、字段类型与取值是否合法、变化方向是否符合指令，以及单一编辑意图下是否包含无关字段。未通过验证的计划将根据校验结果进行修正，最多迭代2轮；若修正后仍缺少必要操作，则采用预定义语义锚点进行确定性补全。对于当前参数化模板无法表达的编辑请求，系统不执行相应字段写入。

2.2.2　确定性程序的执行

通过验证的编辑计划不直接作为最终服装表示，而是由确定性程序逐字段应用于$\boldsymbol{J}_0$。这种分工使语言模型主要承担文本语义理解与字段选择，具体参数更新、类型转换、范围处理和结构对齐则由确定性过程完成。

对于布尔型字段，程序仅接受能够明确解析的真值操作；对于类别型字段，目标值必须属于当前设

© 中国图象图形学报版权所有

计模式规定的合法取值集合；对于数值型字段，程序依据绝对目标值、相对增减方向、指令中的数值提示或预定义语义锚点计算候选值。数值候选经过字段级变化幅度控制后，按照设计模式规定的合法范围进行处理。当源字段值本身处于合法范围之外时，程序不强制将其投影至范围内部，而是以源值为基准保持指令要求的变化方向。

字段操作完成后，以$\boldsymbol{J}_0$为结构模板对临时结果进行对齐。该过程不引入源JSON中不存在的新键；缺失字段、类型不一致字段和无效字段均恢复为源值。由此得到与源表示具有相同层级结构的临时结果$\boldsymbol{J}_t$。

系统随后检查数值字段的变化方向。若增大操作产生的字段值未超过源值，或减小操作产生的字段值未低于源值，则在字段允许范围内按照确定性步长进行修正。若计划中的必要字段仍未发生变化，则仅重新执行该字段对应的操作，避免通过扩大编辑范围补偿目标变化不足。绝对值设定、类别设定和布尔设定不参与相对方向修正。

2.2.3　非目标变化回滚与详细设计校正

尽管语言模型的计划已经受到候选边界约束，临时结果$\boldsymbol{J}_t$仍需通过逐路径投影保证非目标字段保持不变。最终简化JSON定义为

$$\boldsymbol{J}_1[p]=\begin{cases}\boldsymbol{J}_t[p], & p\in\boldsymbol{E},\\ \boldsymbol{J}_0[p], & p\in\boldsymbol{U}(\boldsymbol{J}_0)\setminus\boldsymbol{E}.\end{cases}\tag{2}$$

式中，p为叶子字段路径；$\boldsymbol{J}_1[p]$、$\boldsymbol{J}_t[p]$和$\boldsymbol{J}_0[p]$分别表示编辑后简化JSON、临时JSON和源简化JSON在路径p上的字段值；方括号·表示按照叶子路径访问结构化表示中的对应字段。式(2)保留最终可执行路径集合$\boldsymbol{E}$内经过检查的字段值，并将其余路径恢复为源简化JSON中的取值。因此，非目标保持不是对生成结果施加事后软约束，而是对叶子字段实施确定性的投影与回滚。

完成非目标字段回滚后，对源简化JSON和编辑后简化JSON进行逐叶比较，实际变化集合定义为

$$\Delta\boldsymbol{J}=\left\{\left(p,\boldsymbol{J}_0[p],\boldsymbol{J}_1[p]\right)\middle|\,p\in\boldsymbol{U}(\boldsymbol{J}_0),\boldsymbol{J}_0[p]\neq\boldsymbol{J}_1[p]\right\}\tag{3}$$

式中，$\Delta\boldsymbol{J}$为源简化JSON与编辑后简化JSON之间的实际变化集合；集合中的每个三元组依次记录叶子字段路径p、源字段值$\boldsymbol{J}_0[p]$和编辑后字段值$\boldsymbol{J}_1[p]$。式(3)同时记录变化路径、源值和编辑后取值，为简化JSON到详细设计的字段映射提供依据。

随后，将$\boldsymbol{J}_1$映射至完整服装参数空间。由于简化表示向详细设计展开时可能引入额外变化，系统以$\boldsymbol{D}_0$的副本为基准，仅查询$\Delta\boldsymbol{J}$中字段对应的预定义映射关系。只有当简化字段与详细设计字段之间的映射有效，且目标详细字段真实存在时，才将编辑值写入详细设计；缺少有效映射或目标字段不存在时，相应参数保持$\boldsymbol{D}_0$中的源值。由此得到校正后的详细设计$\boldsymbol{D}_1$。最后，由$\boldsymbol{D}_1$生成编辑后的规格文件和二维缝纫图$\boldsymbol{P}_1$，并通过GarmentCodeRC完成缝合与物理仿真，得到编辑后的三维服装网格$\boldsymbol{M}_1$。上述字段约束的局部几何编辑流程见算法1。

算法1 **字段约束的局部几何编辑**

输入：源简化JSON $\boldsymbol{J}_0$、源详细设计$\boldsymbol{D}_0$、几何编辑文本T_g、字段语义知识卡片和服装参数设计模式。**输出：**编辑后简化JSON $\boldsymbol{J}_1$、变化集合$\Delta\boldsymbol{J}$、详细设计$\boldsymbol{D}_1$、二维缝纫图$\boldsymbol{P}_1$和三维网格$\boldsymbol{M}_1$1)解析$\boldsymbol{J}_0$中实际存在的服装部件和叶子路径，依据字段语义知识卡片检索与T_g对应的候选字段。2)建立目标路径、必要联动路径和受保护路径，形成字段级候选边界。3)语言模型在候选边界内生成编辑计划$\boldsymbol{\Pi}$，并检查路径、操作、类型、取值范围及语义一致性。4)根据验证后的计划构造可执行路径集合$\boldsymbol{E}$；若不存在可执行路径，则保持$\boldsymbol{J}_0$不变。5)在$\boldsymbol{J}_0$的副本上逐字段执行编辑计划，依据服装参数设计模式完成类型、范围处理、结构对齐和方向一致性检查，得到$\boldsymbol{J}_t$。6)按式(2)回滚非目标字段，得到$\boldsymbol{J}_1$；按式(3)计算实际变化集合$\Delta\boldsymbol{J}$。7)以$\boldsymbol{D}_0$为基准，仅将具有有效字段映射的变化写入详细设计，得到$\boldsymbol{D}_1$。8)由$\boldsymbol{D}_1$生成$\boldsymbol{P}_1$，并通过GarmentCodeRC仿真得到$\boldsymbol{M}_1$。

2.3　纹理编辑

几何编辑完成后，纹理编辑分支以仿真网格$\boldsymbol{M}_1$为载体，顺序连接图像编辑模型、服装区域提取和图像引导三维纹理迁移模块，将文本描述的颜色、材质和图案变化映射至编辑后的服装表面。该分支的研究重点是建立现有外观生成模块与局部几何编辑结果之间的输入输出衔接，从而形成由结构编辑到外观编辑的完整应用流程。

纹理编辑分支首先根据源人物图像$\boldsymbol{I}$和纹理编辑文本T_t构建目标服装参考图。Nano Banana依据$\boldsymbol{I}$和T_t生成外观编辑后的人物图像$\boldsymbol{I}_e$，服装区域提取模块进一步从$\boldsymbol{I}_e$中获得仅包含目标服装的参考图像

© 中国图象图形学报版权所有

I_c。与仅使用文本条件相比，I_c能够为空间外观生成提供更直接的颜色、材质、图案及大致布局信息。

随后，将编辑后的三维服装网格和参考图像I_c输入Hunyuan3D 2.1的hy3dpaint流程，完成图像引导的三维纹理迁移与PBR材质生成。该阶段最大视角数设置为6，生成分辨率设置为512。输出的带纹理服装网格记为M_2，其几何结构继承自M_1，表面颜色、织物质感和图案等外观信息由I_c引导生成。

最后，使用Blender在固定相机、固定光照和白色背景下渲染M_2得到正面、背面和侧面等视角结果。渲染结果用于定性展示和纹理指标计算；几何编辑指标仍在仿真网格M_1上计算，以避免纹理生成过程影响几何评价。该设计将几何评估与纹理评估解耦：几何编辑分支验证结构编辑是否准确并能够稳定通过仿真，纹理编辑分支验证颜色、材质和图案是否符合用户纹理指令。通过这种顺序式流程，几何变化和纹理变化能够在同一编辑后服装资产上统一表达。

2.4 GeometryEdit-10K数据集

为评估文本驱动的缝纫图服装细粒度局部几何编辑能力，本文构建了源—目标成对的GeometryEdit-10K数据集。该数据集在1200个源服装的基础上扩展而来：首先从1200个已成功生成缝纫图、规格文件并通过GarmentCodeRC仿真的源服装中构建源服装池，保证所有源服装均具有合法的详细设计、简化JSON、缝纫规格、二维缝纫图和三维仿真网格。数据集源服装覆盖上衣、下装和整身服装3类。对于每个源服装生成9条局部几何编辑指令，共得到10800个候选样例；经过目标设计生成、缝合和仿真验证后，最终筛选得到10383个有效编辑案例，总体保留率为96.14%。数据集的最终类别分布如表1所示。

编辑指令采用人工智能（artificial intelligence，AI）辅助生成，但字段选择和目标取值受到参数化服装设计模式的约束。具体而言，指令及其目标修改依据详细设计中实际存在的字段、数值范围和枚举值生成。编辑类型覆盖单字段编辑、双字段编辑、多字段编辑以及整身服装中的局部保持编辑。上衣类编辑包括领口形状、衣长、衣身宽松度、袖长和袖口等变化；下装类编辑包括裤长、裤型、裙长、裙摆体量、开衩和腰头等变化；整身服装编辑包括仅修改上衣、仅修改下装以及同时修改上下装等情况。为保证数据质量，每个候选样例均经过JSON合法性检查、目标字段变化检查、非目标字段保持检查和GarmentCodeRC物理仿真验证。最终得到的GeometryEdit-10K同时包含文本编辑指令、源—目标简化JSON、字段级编辑标注、缝纫图和目标三维网格，可用于评估JSON字段层级编辑准确性、缝纫图一致性、三维几何质量和仿真成功率。

表1 GeometryEdit-10K数据集类别分布

Table 1 Category distribution of GeometryEdit-10K

服装类别	源服装数量	编辑数量	最终样例数	占比
上衣	360	9	3117	30.02%
下装	560	9	4905	47.24%
整身服装	280	9	2361	22.74%
合计	1200	9	10383	100.00%

为进一步说明GeometryEdit-10K与已有服装编辑相关数据集的差异，表2从数据规模、编辑指令、源—目标缝纫图、字段级变化标注和三维网格方面进行比较。Wang等人（2018）构建了包含参数表示、草图表示和披挂三维网格的多模态服装设计数据集，用于草图、参数和三维披挂形态之间的共享空间学习，但该数据集不提供文本编辑指令，也不面向字段级局部编辑评估。AIpparel（Nakayama等，2025）构建的多模态GarmentCodeData数据集（GarmentCodeData-MultiModal，GCD-MM）包含大规模文本、图像、编辑后的缝纫图和编辑指令，能够支持文本指令驱动的缝纫图编辑，但其主要目标是训练多模态生成模型，并未显式提供面向局部可控编辑评估的JSON字段变化标注和源—目标成对仿真网格。相比之下，GeometryEdit-10K在每个编辑样例中同时记录文本指令、源—目标简化JSON、字段级修改、二维缝纫图和三维仿真网格，因而更适合用于评估目标字段是否被正确修改、非目标字段是否保持稳定以及编辑后服装是否能够稳定仿真。

3 实 验

为直观展示几何编辑与纹理编辑的协同效果，如图3所示，图3给出了完整流程的定性结果。给定穿衣人物图像和文本编辑指令，本文所提出的框架

© 中国图象图形学报版权所有

首先在缝纫图空间中修改目标服装结构，再将编辑后的纹理外观迁移到仿真三维服装网格上。可以看到，该流程能够同时表达袖长、裙长、样式等几何变化，以及颜色、材质和图案等纹理变化，生成具有目标结构和目标外观的三维服装结果。

表2　服装编辑相关数据集比较

Table 2　Comparison of garment editing related datasets

数据集	规模	编辑指令	成对缝纫图	字段变化标注	三维网格
Wang等(2018)	8K		✔		✔
AIpparel(2025)	120K	✔	✔		
本文	10K	✔	✔	✔	✔

注：✔表示数据集包含相应信息，　表示数据集不包含相应信息；K表示千，即10^3。

3.1　几何编辑结果对比

为评估几何编辑分支的局部编辑能力，本文构建了包含1200个源—目标编辑样例的GeometryEdit-1200数据集。该实验从仿真三维网格、缝纫图结构和简化JSON字段3个层级评估编辑结果，重点考察方法是否能够准确修改目标服装部件，并保持非目标区域稳定。

3.1.1　基线方法与实现细节

本文将所提出方法与两个文本驱动服装编辑基线进行比较：ChatGarment(Bian 等，2025)和AIpparel(Nakayama 等，2025)。其中，ChatGarment根据源服装和编辑指令直接预测编辑后的结构化服装表示，可作为完整目标表示生成式编辑的代表方法；AIpparel面向多模态缝纫图生成与编辑，能够根据文本指令输出编辑后的缝纫图规格，可作为缝纫图层级编辑的代表方法。为保证比较公平，所有方法均使用相同的原服装和编辑指令。对于网格层级与缝纫图层级评估，本文将各方法生成的缝纫图规格输入相同的GarmentCodeRC仿真流程，并在缝纫图与仿真后三维网格上计算几何指标。对于JSON字段层级评估，由于AIpparel直接输出缝纫图规格，难以可靠转换为本文字段级指标所需的ChatGarment风格简化JSON，因此该层级仅比较ChatGarment和本文所提出方法。

3.1.2　评价指标

几何编辑效果从网格层级、缝纫图层级和JSON字段层级进行评价。网格层级将预测网格与目标网格进行比较，报告L1范数倒角距离(L1-norm Chamfer distance，CD-L1)和阈值为1.0的F分数(F-score at a threshold of 1.0，F@1)。上述网格指标在共同仿真成功的样例子集上计算，同时在全部样例上报告仿真成功率。

图3 几何与纹理编辑定性结果图

Fig. 3 Qualitative Results of Geometry and Texture Editing

缝纫图层级遵循AIpparel(Nakayama 等，2025)的设置，报告准确率和布片L2距离(panel L2 distance，Panel L2)。准确率由布片数量准确率和匹配布片的边准确率相乘得到，前者用于判断预测结果与目标结果的布片数量是否一致，后者用于评价匹配布片的边结构准确性。Panel L2用于衡量匹配预测布片和目标布片之间的平均顶点距离。上述缝纫图层级指标在共同缝合成功的样例子集上计算。

三维网格和二维缝纫图指标能够反映整体几何形状及布片结构差异，但不能直接判断文本指令要求修改的字段是否得到执行，也不能量化非目标字段发生的额外变化。因此，本文进一步在简化JSON的叶子字段上评价目标执行和非目标保持情况。

设源简化JSON、目标简化JSON和预测简化JSON分别为$\boldsymbol{S}$、$\boldsymbol{T}$和$\boldsymbol{P}$。每个样例包含必要路径集合$\boldsymbol{R}$和可联动路径集合$\boldsymbol{O}$。当标注中未提供必要路径时，以源JSON和目标JSON之间实际发生变化的叶子路径作为必要路径。$\boldsymbol{U}$表示为$\boldsymbol{S}$、$\boldsymbol{T}$和$\boldsymbol{P}$展开后的叶子路径并集，预测结果相对于源JSON发生变化的路径集合定义为

$$\boldsymbol{\Delta P}=\left\{p\in \boldsymbol{U}\,\middle|\,\boldsymbol{P}[p]\neq \boldsymbol{S}[p]\right\} \tag{4}$$

式中，$\boldsymbol{\Delta P}$为预测结果相对于源简化JSON发生变化的叶子路径集合；对于数值字段，当预测值与源值之差的绝对值大于数值比较阈值ε时，判定该字段发生变化，其中$\varepsilon=10^{-6}$。

允许发生变化的路径集合为

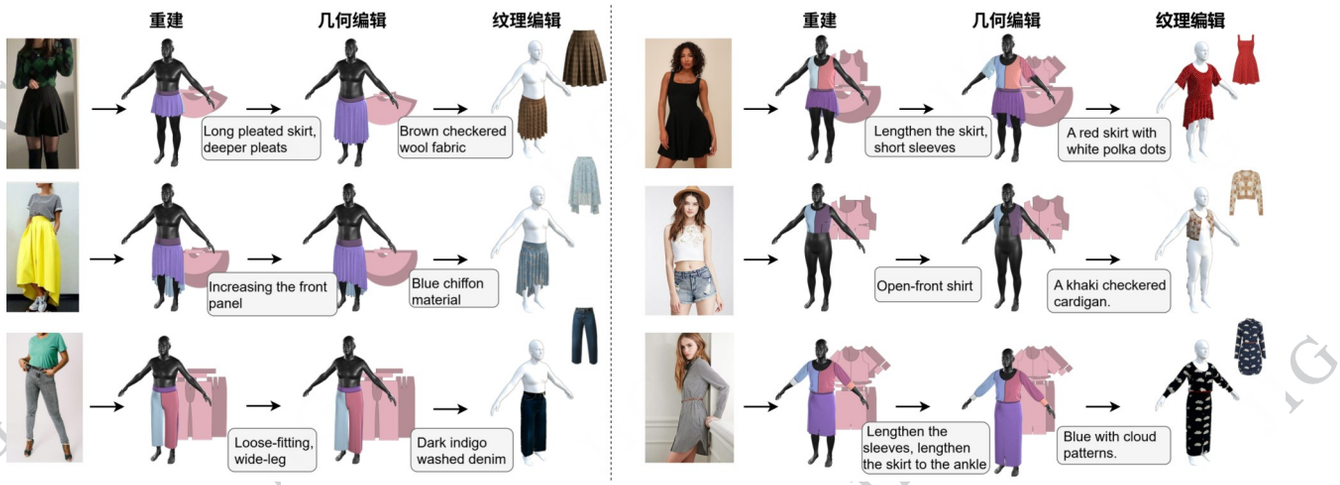
$$\boldsymbol{A}=\boldsymbol{R}\cup\boldsymbol{O} \tag{5}$$

式中，$\boldsymbol{A}$表示允许发生变化的路径集合。目标命中率不仅判断必要字段是否变化，还判断字段值是否满足目标编辑要求。对于连续型和整数型字段，本文结合字段的绝对容差与源值到目标值的变化幅度设置目标容差，定义为

$$\tau_p=\max\left(\tau_0(p),0.35\left|\boldsymbol{T}[p]-\boldsymbol{S}[p]\right|\right) \tag{6}$$

式中，

© 中国图象图形学报版权所有



τ_p为字段p的目标容差；$\tau_0(p)$为字段p对应的绝对容差，连续型字段取0.08，整数型字段取1；0.35为根据目标变化幅度设置相对容差的比例系数。对于相对增减操作，编辑进度定义为

$$\rho_p=\frac{\left|\boldsymbol{P}[p]-S[p]\right|}{\left|\boldsymbol{T}[p]-S[p]\right|} \tag{7}$$

式中，ρ_p为字段p的编辑进度；分子表示预测值相对于源值的实际变化幅度，分母表示目标值相对于源值的目标变化幅度。绝对值设定操作要求预测值满足$\left|\boldsymbol{P}[p]-\boldsymbol{T}[p]\right|\leqslant\tau_p$。相对增减操作首先要求预测值相对于源值沿计划规定的方向变化；在方向正确的前提下，当预测值落入目标容差，或编辑进度满足$0.40\leqslant\rho_p\leqslant1.75$时，将该必要字段判定为命中。当$\boldsymbol{T}[p]=S[p]$时，不计算式(7)，而按照目标容差判断。对于布尔型、枚举型和其他类别型字段，预测值必须与目标值一致且不同于源值；对于删除或置空操作，预测值必须为空。源、目标或预测JSON中缺失的必要字段均判定为未命中。

按照上述条件统计满足目标取值、变化方向或编辑进度要求的必要路径，单个案例的目标命中率定义为

$$r_{\mathrm{h}}=\frac{|\boldsymbol{H}|}{|\boldsymbol{R}|} \tag{8}$$

式中，$\boldsymbol{H}$为满足上述命中条件的必要路径集合；$\boldsymbol{R}$为该案例的必要路径集合；$|\cdot|$表示集合中元素的数量。当必要路径集合为空时，按照实际评价实现将目标命中率记为1。为评价预测变化区域与标注允许区域之间的一致性，本文进一步报告字段级精确率、字段级召回率和编辑F1值。字段级精确率表示预测发生变化的路径中属于允许变化区域的比例；字段级召回率表示必要路径中被预测结果实际修改的比例；编辑F1值为字段级精确率和字段级召回率的调和平均值。最后，非目标漂移率用于衡量允许区域之外的字段变化，定义为

$$r_{\mathrm{d}}=\frac{|\boldsymbol{\Delta P}\backslash\boldsymbol{A}|}{|\boldsymbol{U}\backslash\boldsymbol{A}|} \tag{9}$$

式中，$\boldsymbol{\Delta P}\backslash\boldsymbol{A}$为预测变化路径中不属于允许变化区域的路径集合；$\boldsymbol{U}\backslash\boldsymbol{A}$为全部叶子路径中不属于允许变化区域的路径集合。非目标漂移率越低，表示无关服装参数保持得越好。上述JSON指标均先在单个案例上计算，再对评价案例进行宏平均。目标命中率考察必要字段的目标值、变化方向和编辑进度，字段级精确率、召回率和编辑F1值则考察变化路径的定位情况，二者从不同角度评价字段级局部编辑结果。

3.1.3 结果比较

如表3所示，AIpparel相较于ChatGarment获得了更高的缝纫图和网格精度，说明其是具有竞争力的缝纫图编辑基线。本文所提出方法在所有指标上取得最佳性能，包括最低的CD-L1、最高的F@1、最高的仿真成功率、最高的缝纫图准确率以及最低的Panel L2。这些结果表明，细粒度参数编辑能够同时提高几何精度和仿真稳定性。

如表4所示，表4给出了GeometryEdit-1200上的JSON层级定量结果。与ChatGarment相比，本文方法取得了更高的目标命中率和编辑F1值，非目标漂移率由0.621降至0.013。这说明本文方法不仅能够在三维网格和缝纫图层级获得更好的几何结果，也能够在结构化参数层级更准确地完成用户请求的

© 中国图象图形学报版权所有

表3　GeometryEdit-1200上的网格与缝纫图层级定量比较

Table 3　Quantitative comparison of mesh-level and sewing-pattern-level metrics on GeometryEdit-1200

方法	CD-L1↓	F@1↑	仿真成功率↑	准确率↑	Panel L2↓
ChatGarment	12.722	0.455	0.979	0.516	10.116
AIpparel	6.850	0.666	0.855	0.875	3.158
本文	**3.501**	**0.820**	**0.989**	**0.956**	**2.409**

注：加粗字体表示各列最优结果；“↑”和“↓”分别表示指标值越大越优、值越小越优。

部件级修改，并保持无关服装的字段稳定。该结果进一步表明，JSON层级指标能够从字段变化角度揭示局部编辑中的目标命中和非目标漂移问题，为可控服装编辑提供比整体几何误差更细粒度的评价依据。

表4　GeometryEdit-1200上的JSON层级定量比较

Table 4　Quantitative comparison of JSON-level metrics on GeometryEdit-1200

方法	目标命中率↑	编辑F1↑	非目标漂移率↓
ChatGarment	0.366	0.193	0.621
本文	**0.568**	**0.875**	**0.013**

注：加粗字体表示各列最优结果；“↑”和“↓”分别表示指标值越大越优、值越小越优。

如图4所示，图4展示了ChatGarment、AIpparel和本文方法的定性比较。ChatGarment能够响应编辑指令，但部分结果改变了非目标区域或产生了不稳定结构。AIpparel在部分样例中能够生成合理的缝纫图，但在复杂服装或精确部件级编辑中仍存在不足。相比之下，本文方法能够更准确地编辑目标服装部件，并生成与目标几何更一致的仿真服装。

3.2　纹理编辑结果对比

本文构建纹理评测集用于考察纹理处理流程。该评测集包含50个纹理编辑案例，编辑类型包括全局颜色变化、材质变化、全局图案变化和局部图案添加。同一案例中，各方法使用相同的源三维服装网格，以控制服装几何、轮廓、接缝和褶皱等结构因素。每个案例提供源纹理描述、目标纹理描述和纹理编辑指令，并为源结果、InTeX、TEXTure和本文流程生成正面、背面、左侧和右侧四视角渲染图。

对比方法包括TEXTure（Richardson 等，2023）和InTeX（Tang 等，2024）。其中，TEXTure用于文本引导的三维纹理生成，InTeX通过交互式文本引导重绘实现纹理编辑。本文流程首先利用图像编辑模型生成纹理编辑后的服装参考图，再使用Hunyuan3D 2.1（Hunyuan3D Team 等，2025）将编辑后的纹理迁移到服装网格上。

在定量评估中，本文采用两种基于对比语言—图像预训练（contrastive language-image pre-training，CLIP）的指标。每件带纹理的服装由四个红绿蓝（red，green and blue，RGB）视图表示。$\boldsymbol{I}^{s}_{i,v}$和$\boldsymbol{I}^{e}_{i,v,m}$分别表示第$i$个样例、第$v$个视角下的源RGB图像以及方法$m$的编辑后RGB图像。设$T^{s}_{i}$和$T^{t}_{i}$分别表示源纹理描述和目标纹理描述。本文使用CLIP图像编码器$E_{\mathrm{I}}(\cdot)$和文本编码器$E_{\mathrm{T}}(\cdot)$提取经L2范数归一化的嵌入。

图4 几何编辑定性比较图

Fig. 4 Qualitative Comparison of Geometry Editing

注：图中红色边框为目标编辑区域，蓝色边框为非目标编辑区域

首先，本文采用CLIPScore（Hessel 等，2021）来衡量编辑结果与目标纹理描述之间的文本—图像一致性，定义为

$$c_{i,v,m}=2.5\times\max\left(\cos\left(E_{\mathrm{I}}\left(\boldsymbol{I}^{e}_{i,v,m}\right),E_{\mathrm{T}}\left(T^{t}_{i}\right)\right),0\right) \tag{10}$$

式中，$c_{i,v,m}$为方法m在第i个样例、第v个视角上的文本—图像一致性得分；2.5为固定缩放系数，$\max(\cdot,0)$用于将负余弦相似度截断为0。

其次，按照CLIP引导编辑（Gal 等，2022）中使用的定向CLIP公式计算定向CLIP相似度。它衡量从源渲染图到编辑后渲染图的视觉变化是否与从源纹理提示到目标纹理提示的语义变化一致。图像方向定义为

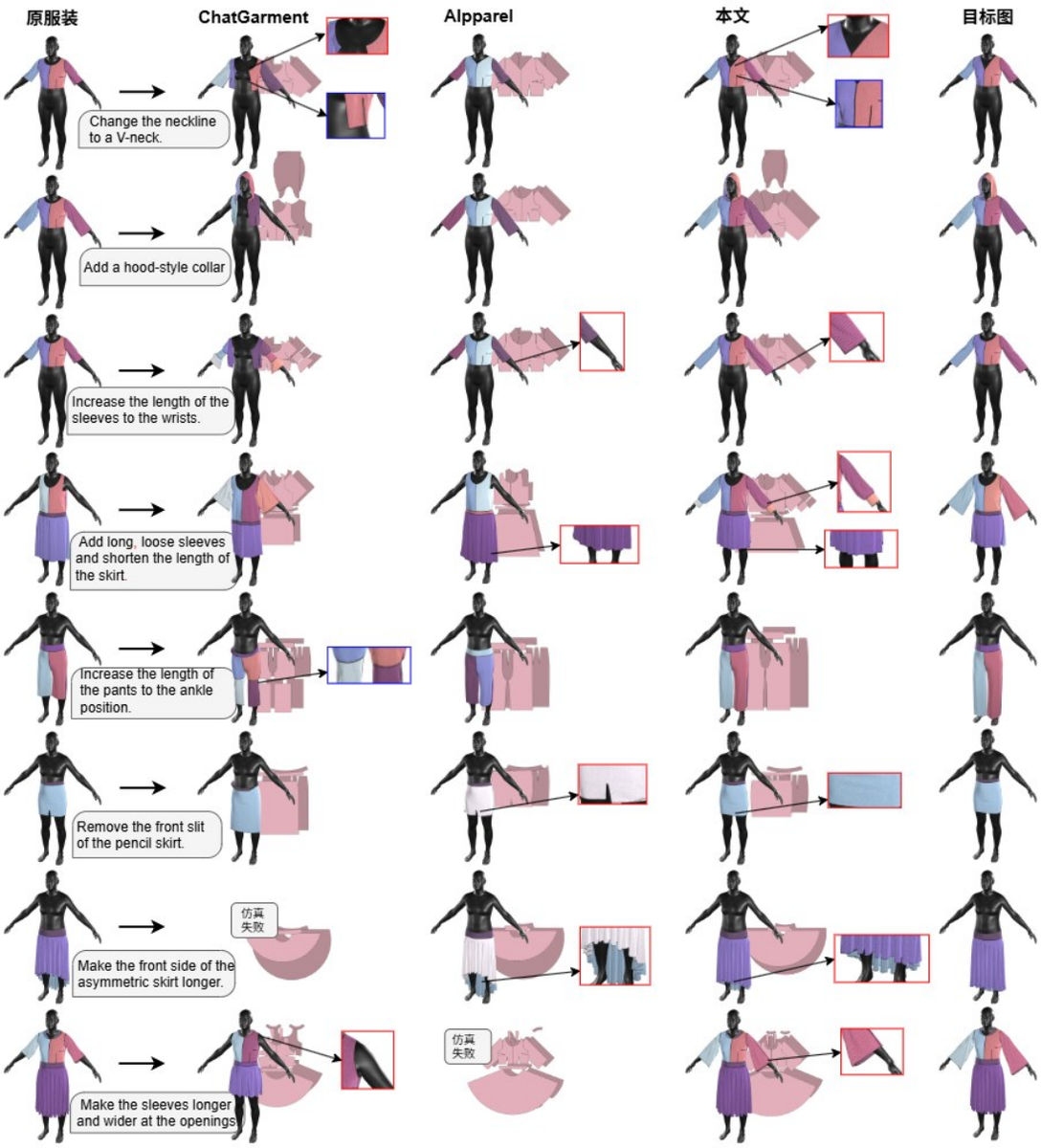
$$\Delta\boldsymbol{I}_{i,v,m}=E_{\mathrm{I}}\left(\boldsymbol{I}^{e}_{i,v,m}\right)-E_{\mathrm{I}}\left(\boldsymbol{I}^{s}_{i,v}\right) \tag{11}$$

式中，$\Delta\boldsymbol{I}_{i,v,m}$为方法$m$在第$i$个样例、第$v$个视角上的图像变化方向，由编辑后图像嵌入与源图像嵌入之差得到。与图像方向相对应，文本方向定义为

$$\Delta\boldsymbol{T}_{i}=E_{\mathrm{T}}\left(T^{t}_{i}\right)-E_{\mathrm{T}}\left(T^{s}_{i}\right) \tag{12}$$

式中，$\Delta\boldsymbol{T}_{i}$为第i个样例的文本语义变化方向，由目标纹理描述嵌入与源纹理描述嵌入之差得到。根据图像方向与文本方向之间的夹角，定向CLIP相似度

© 中国图象图形学报版权所有



定义为

$$d_{i,v,m}=\cos\left(\Delta \boldsymbol{I}_{i,v,m},\Delta \boldsymbol{T}_i\right) \tag{13}$$

式中，$d_{i,v,m}$为方法m在第i个样例、第v个视角上的定向CLIP相似度。对每种方法，报告50个样例和4个视角上的均值。

如表5所示，InTeX获得了最高的CLIPScore，表明其编辑结果与目标纹理文本之间具有较强的图文对齐能力。本文流程的CLIPScore略低于InTeX，但仍保持在接近水平，说明其生成纹理同样能够较好地响应目标指令。本文流程的定向CLIP相似度取得了最高分数，优于InTeX和TEXTure，说明从源服装到编辑后服装的视觉变化与预期纹理编辑更一致。对于纹理编辑任务而言，目标不仅是生成与目标提示匹配的图像，还需要从源外观出发执行正确编辑，因此方向一致性指标具有重要意义。

图5给出了代表性纹理编辑结果。在图示样例中，InTeX和TEXTure能够产生一定的纹理变化，但部分结果存在颜色偏差、纹理伪影或局部图案定位

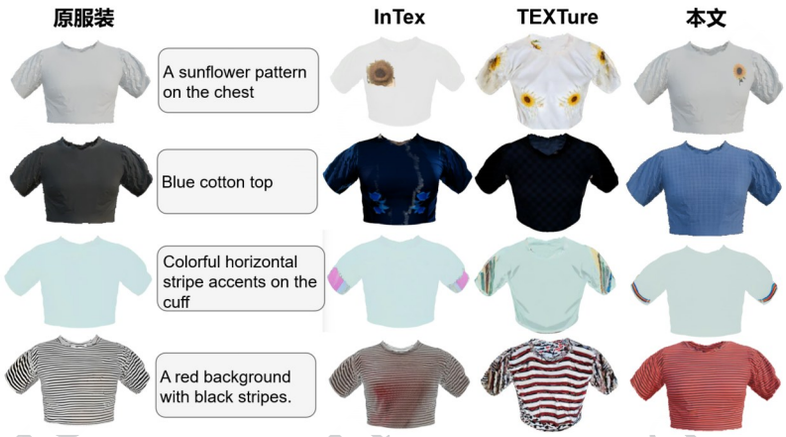
表5　纹理编辑定量比较

Table 5　Quantitative comparison of texture editing

方法	CLIPScore ↑	定向CLIP相似度 ↑
InTeX	**0.7279**	0.1190
TEXTure	0.7220	0.0931
本文	0.7232	**0.1588**

注：加粗字体表示各列最优结果；“↑”表示指标值越大越优。

© 中国图象图形学报版权所有



不准确等现象。本文流程较好地响应了相应纹理指令，包括在胸前生成向日葵图案、生成蓝色棉质外观、小花图案和红黑条纹。上述结果表明，纹理编辑后的服装参考图能够为目标外观向编辑后三维服装网格的迁移提供引导。

3.3 消融实验

为分析可编辑边界解析和字段级确定性编辑计划对局部几何编辑性能的影响，本文在构建的1200个样例实验集上进行实验。消融实验比较3种设置：完整方法、无边界约束变体和直接JSON生成变体。其中，无边界约束变体去除可编辑字段边界限制，使编辑计划能够在更宽的参数范围内生成和执行；直接JSON生成变体则不生成字段级编辑计划，而是由语言模型直接输出编辑后的简化JSON结果。通过这两类变体，可以分别验证“可编辑边界限定”和“字段级执行计划”在局部缝纫图编辑中的作用。

如表6所示，完整方法在CD-L1、F@1、准确率、Panel L2、目标命中率和编辑F1等主要指标上均取得最佳结果，说明可编辑边界解析和字段级执行计划能够共同提升局部几何编辑的准确性和稳定性。去除可编辑边界后，CD-L1升高至3.738，准确率、目标命中率和编辑F1分别降至0.936、0.552和0.844，这表明缺少边界约束时，系统对指令相关字段的定位能力下降，容易将语义相关但并非目标的字段纳入编辑范围，从而削弱局部编辑效果。

直接JSON生成变体的性能下降更加明显。虽然该变体的非目标漂移率最低，但其目标命中率仅为0.228，编辑F1仅为0.376，且CD-L1升高至4.254，Panel L2升高至3.104。这说明直接JSON生成变体并非实现了更好的局部保持，而是由于许多必要编辑没有被有效执行，导致字段变化较少、漂移率偏低。换言之，低漂移率需要与目标命中率和编辑F1共同分析，单独较低的漂移率并不能说明编辑质量更好。相比之下，完整方法在保持较高仿真成功率的同时，兼顾了目标字段修改、非目标区域保持和最终几何质量，表现出更好的综合性能。

图5 纹理编辑定性比较图

Fig. 5 Qualitative Comparison of Texture Editing

如图6所示，进一步展示了代表性局部几何编辑任务中的定性结果。图中左侧包含上衣结构编辑案例；右侧包含下装和整身服装编辑案例。可以看到，无边界约束变体虽然能够对指令产生响应，但由于缺少明确的可编辑路径限制，容易引起非目标部位变化。例如在上衣编辑中，袖子、衣身或领口附近可能出现额外形变。直接JSON生成变体具有较高自由度，但缺少字段级操作约束和确定性后处理，容易出现编辑幅度不足、目标结构不完整或局部形态不准确的问题。例如未能充分完成裤脚加长、铅笔裙加长和裙摆褶皱增强。相比之下，完整方法能够更准确地完成目标部件变化：在兜帽式领型和V形领口编辑中能够集中修改领口区域；在铅笔裙加长和裙摆褶皱增强任务中能够更好地控制下装形态，同时保持非目标上衣区域稳定。定量和定性结果共同表明，可编辑边界解析能够减少无关字段扰动，字段级执行计划能够提高目标操作的执行稳定性，两者对于可靠的局部几何编辑均具有必要作用。

4 结 论

针对文本驱动三维服装编辑中几何与纹理处理割裂、局部编辑可控性不足等问题，本文提出一种文本驱动的三维服装几何与纹理编辑框架。该框架以

© 中国图象图形学报版权所有

表6　几何编辑消融结果

Table 6　Geometric editing ablation results

方法	CD-L1↓	*F*@1↑	仿真成功率↑	准确率↑	Panel L2↓	目标命中率↑	编辑F1↑	非目标漂移率↓
无边界约束	3.738	0.811	0.973	0.936	2.562	0.552	0.844	0.023
直接JSON生成	4.254	0.786	0.960	0.915	3.104	0.228	0.376	**0.004**
本文	**3.228**	**0.816**	**0.989**	**0.957**	**2.445**	**0.566**	**0.873**	0.013

注：加粗字体表示各列最优结果；“↑”和“↓”分别表示指标值越大越优、值越小越优。

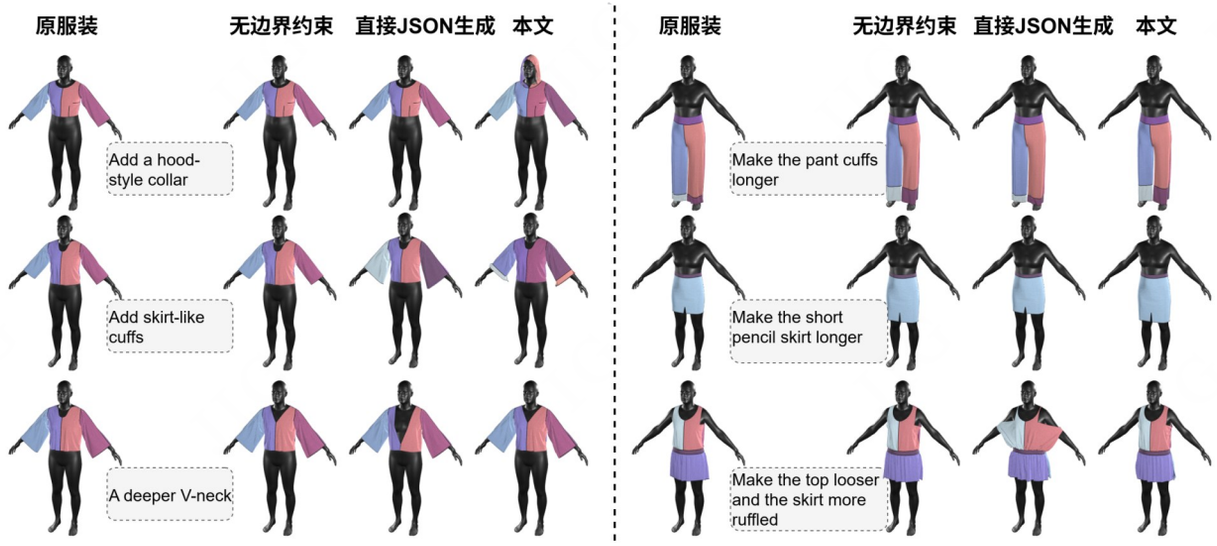


图6　几何编辑消融实验定性比较图

Fig. 6　Qualitative Comparison of Geometry Editing Ablation Experiments

参数化缝纫图作为几何编辑载体，以图像引导纹理迁移作为外观编辑手段，使文本指令能够作用于服装结构和表面纹理，形成从穿衣人物图像到编辑后带纹理三维服装资产的完整流程。

围绕局部几何编辑的可控性问题，本文设计了基于可编辑边界确定、字段级编辑计划、确定性执行和非目标变化回滚的图样编辑器。该编辑器将袖长、领口、裤型、裙长、裙摆等部件级修改约束在当前服装结构支持的参数范围内，减少完整生成式编辑带来的无关结构变化。在纹理编辑阶段，本文将编辑后的仿真网格顺序接入现有图像编辑模型、服装区域提取和图像引导纹理迁移模块，完成颜色、材质和图案编辑。

实验结果表明，本文所提出框架在几何编辑中能够同时改善三维网格精度、缝纫图结构一致性和JSON字段级编辑可控性；消融实验进一步说明，可编辑边界解析和字段级执行计划对减少非目标漂移、提高目标部件命中率具有重要作用。纹理实验表明，在所测试样例和所采用指标下，本文流程取得了更高的定向CLIP相似度，表现出更好的编辑方向一致性。总体而言，本文所提出方法为虚拟试衣、数字时尚和三维内容创作中的可仿真服装编辑提供了一种结构可控、外观可编辑的实现途径。目前该方法仍受到参数化服装模板和纹理参考图质量的限制。对于多层叠穿服装、非规则服装结构以及极细粒度局部纹理控制，仍可能出现结构表达不足或纹理迁移不够精确的情况。未来工作将进一步扩展服装模板覆盖范围，增强复杂服装组合建模能力，并提高局部图案、织物材质和多视角纹理一致性的控制精度。

参考文献（References）

Bian S, Xu C, Xiu Y, Grigorev A, Liu Z, Lu C, et al. 2025. ChatGarment: Garment estimation, generation and editing via large language models//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE:

© 中国图象图形学报版权所有

2924-2934［DOI:10.1109/CVPR52734.2025.00278］

Chen C H, Su J W, Hu M C, Yao C Y and Chu H K. 2024a. PanelFormer: Sewing pattern reconstruction from 2D garment images//Proceedings of 2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE: 443-452［DOI:10.1109/WACV57701.2024.00051］

Chen H H, Tao Y F and Zhang J Y. 2024b. A survey on 3D clothed human reconstruction: from traditional techniques to high-fidelity models. Journal of Image and Graphics, 29(9): 2566-2595（陈鸿鹄,陶云帆,张举勇. 2024b. 三维穿衣人体重建综述——从传统方法到高保真模型. 中国图象图形学报,29(9): 2566-2595）［DOI:10.11834/jig.230646］

Gal R, Patashnik O, Maron H, Bermano A H, Chechik G and Cohen-Or D. 2022. StyleGAN-NADA: CLIP-guided domain adaptation of image generators. ACM Transactions on Graphics (TOG), 41(4): #141［DOI:10.1145/3528223.3530164］

Guo J, Chen J, Chen W, Sun Z, Li L, Zhao B, et al. 2027. GarmentX: Autoregressive parametric representations for 3D garment generation. Pattern Recognition, 183: #114937［DOI:10.1016/j.patcog.2026.114937］

He K, Yao K, Zhang Q, Yu J, Liu L and Xu L. 2024. DressCode: Autoregressively sewing and generating garments from text guidance. ACM Transactions on Graphics (TOG), 43(4): #72［DOI:10.1145/3658147］

Hessel J, Holtzman A, Forbes M, Le Bras R and Choi Y. 2021. CLIPScore: A reference-free evaluation metric for image captioning//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics: 7514-7528［DOI:10.18653/v1/2021.emnlp-main.595］

Hunyuan3D Team, Yang S, Yang M, Feng Y, Huang X, Zhang S, et al. 2025. Hunyuan3D 2.1: From images to high-fidelity 3D assets with production-ready PBR material[EB/OL]. [2026-07-03]. https://arxiv.org/pdf/2506.15442.pdf

Korosteleva M and Sorkine-Hornung O. 2023. GarmentCode: Programming parametric sewing patterns. ACM Transactions on Graphics (TOG), 42(6): #199［DOI:10.1145/3618351］

Li B, Li X, Jiang Y, Xie T, Gao F, Wang H, et al. 2025a. GarmentDreamer: 3DGS guided garment synthesis with diverse geometry and texture details//Proceedings of 2025 International Conference on 3D Vision. Singapore, Singapore: IEEE: 1416-1426［DOI:10.1109/3DV66043.2025.00134］

Li S, Liu R, Liu C, Wang Z, He G, Li Y L, et al. 2025b. GarmageNet: A multimodal generative framework for sewing pattern design and generic garment modeling. ACM Transactions on Graphics (TOG), 44(6): #216［DOI:10.1145/3763271］

Li X, Yao Q and Wang Y. 2025c. GarmentDiffusion: 3D garment sewing pattern generation with multimodal diffusion transformers//Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. Montreal, Canada: International Joint Conferences on Artificial Intelligence Organization: 1458-1466［DOI:10.24963/ijcai.2025/163］

Liu J, Wu J, Gao X, Hu J, Xiong B, Liu X, et al. 2025. TexGarment: Consistent garment UV texture generation via efficient 3D structure-guided diffusion transformer//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 26566-26575［DOI: 10.1109/CVPR52734.2025.02474］

Liu L, Xu X, Lin Z, Liang J and Yan S. 2023. Towards garment sewing pattern reconstruction from a single image. ACM Transactions on Graphics (TOG), 42(6): #200［DOI:10.1145/3618319］

Michel O, Bar-On R, Liu R, Benaim S and Hanocka R. 2022. Text2Mesh: Text-driven neural stylization for meshes//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 13482-13492［DOI:10.1109/CVPR52688.2022.01313］

Nakayama K, Ackermann J, Kesdogan T L, Zheng Y, Korosteleva M, Sorkine-Hornung O, et al. 2025. AIpparel: A multimodal foundation model for digital garments//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 8138-8149［DOI:10.1109/CVPR52734.2025.00762］

Nakayama K, Shen I C, Liu R, Wang Y, Wetzstein G and Igarashi T. 2026. Garment Particles: A 2D-3D symmetric garment representation for generation and editing//Proceedings of ACM SIGGRAPH 2026 Conference Papers. Los Angeles, USA: ACM: #178［DOI:10.1145/3799902.3811102］

Pu J C, Liu L, Fu X D, Liu L J and Huang Q S. 2022. Clothing visual representation for 3D human reconstruction. Journal of Computer-Aided Design and Computer Graphics, 34(3): 352-363（普骏程,刘骊,付晓东,刘利军,黄青松. 2022. 三维人体重建中的服装视觉信息表示. 计算机辅助设计与图形学学报,34(3): 352-363）［DOI:10.3724/SP.J.1089.2022.18921］

Richardson E, Metzer G, Alaluf Y, Giryes R and Cohen-Or D. 2023. TEXTure: Text-guided texturing of 3D shapes//Proceedings of ACM SIGGRAPH 2023 Conference Papers. Los Angeles, USA: ACM: #54［DOI:10.1145/3588432.3591503］

Shan H, Li M, Yang H, Zheng K, Zheng S, Fu Y, et al. 2026. NITex: Non-isometric image-based garment texture generation//Proceedings of 2026 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Denver, USA: IEEE: 19884-19893

Tang J, Lu R, Chen X, Wen X, Zeng G and Liu Z. 2024. InTeX: Interactive text-to-texture synthesis via unified depth-aware inpainting [EB/OL]. [2026-07-03]. https://arxiv.org/pdf/2403.11878.pdf

Tao Z, Jiang Y, Chen Y, Xie T, Wang H, Wu Y, et al. 2026. DressWild: Feed-forward pose-agnostic garment sewing pattern generation from in-the-wild images[EB/OL]. [2026-07-03].

© 中国图象图形学报版权所有

https://arxiv.org/pdf/2602.16502.pdf

Wang R, Cheng Z, Lin Z, Ling J, Liu Y, An Y, et al. 2025. Semantic-Garment: Semantic-controlled generation and editing of 3D Gaussian garments//Proceedings of the 33rd ACM International Conference on Multimedia. Dublin, Ireland: ACM: 9793-9802 [DOI:10.1145/3746027.3755136]

Wang T Y, Ceylan D, Popović J and Mitra N J. 2018. Learning a shared shape space for multimodal garment design. ACM Transactions on Graphics (TOG), 37(6): #203 [DOI:10.1145/3272127.3275074]

Wang Y, Zhang C, Frazão G, Yang J, Ichim A E, Beeler T, et al. 2026. GarmentCrafter: Progressive novel view synthesis for single-view 3D garment reconstruction and editing//Proceedings of 2026 International Conference on 3D Vision. Vancouver, Canada: IEEE: 977-988 [DOI:10.1109/3DV69130.2026.00098]

Zhang C, Wang Y, Vicente F, Wu C, Yang J, Beeler T, et al. 2024. FabricDiffusion: High-fidelity texture transfer for 3D garments generation from in-the-wild images//Proceedings of ACM SIGGRAPH Asia 2024 Conference Papers. Tokyo, Japan: ACM: #46 [DOI:10.1145/3680528.3687637]

Zhou F, Liu R, Liu C, He G, Li Y L, Jin X, et al. 2025. Design2GarmentCode: Turning design concepts to tangible garments through program synthesis//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 23712-23722 [DOI:10.1109/CVPR52734.2025.02208]

作者简介

刘聪聪，女，硕士研究生，主要研究方向为AIGC+虚拟试衣。E-mail: 2242816@mail.dhu.edu.cn

方志军，通信作者，男，二级教授，博士生导师，主要研究方向为人工智能、机器视觉、数据智能。E-mail: zjfang@dhu.edu.cn

薛涵，男，讲师，主要研究方向为生成式人工智能、图像与视频生成及编辑、数字人。E-mail: xuehan@dhu.edu.cn

© 中国图象图形学报版权所有