



## (12) 发明专利

(10) 授权公告号 CN 119625181 B

(45) 授权公告日 2025.09.30

(21) 申请号 202411770211.3

(51) Int.CI.

(22) 申请日 2024.12.04

G06T 17/00 (2006.01)

(65) 同一申请的已公布的文献号

G06T 15/00 (2011.01)

申请公布号 CN 119625181 A

(56) 对比文件

(43) 申请公布日 2025.03.14

CN 115797555 A, 2023.03.14

(73) 专利权人 人工智能与数字经济广东省实验室(广州)

CN 118298127 A, 2024.07.05

地址 510330 广东省广州市海珠区新港东路2429号首层自编051房

审查员 郑诗嘉

(72) 发明人 谭明奎 杨逸凡 张书海 邓泽帅  
刘东 黄子雄 罗海林 周凯  
杨嘉豪

(74) 专利代理机构 广州科粤专利商标代理有限公司 44001

权利要求书2页 说明书12页 附图4页

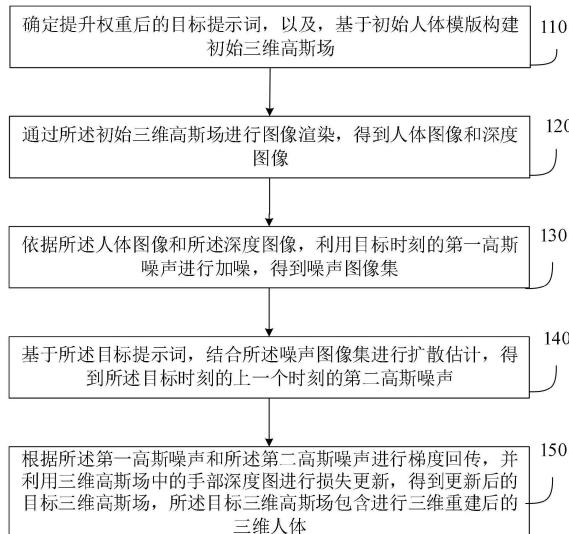
## (54) 发明名称

一种从自然语言描述文本中重建出三维人体的三维重建方法及装置

## (57) 摘要

本申请涉及一种从自然语言描述文本中重建出三维人体的三维重建方法及装置，涉及计算机视觉技术领域，该方法包括：通过人体模版构建的初始三维高斯场进行图像渲染，得到人体图像和深度图像，利用目标时刻的高斯噪声对图像进行加噪，得到噪声图像集，噪声图像集结合提升权重的提示词进行扩散估计，得到上一个时刻的高斯噪声，结合两个时刻的高斯噪声进行梯度回传，并利用三维高斯场中的手部深度图进行损失更新，得到更新后的进行三维重建后的三维人体。本申请利用深度损失，对齐人体模型与局部身体部位深度，从而使得生成的3D人体能精确捕捉到人体部位的复杂特征，恢复精细的人体部位，解决现有的三维人体构建技术无法生成精确的身体部位的问题。

CN 119625181 B



CN

1. 一种从自然语言描述文本中重建出三维人体的三维重建方法,其特征在于,包括:  
确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;  
通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;  
依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪  
声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的  
深度加噪图像;  
基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一  
个时刻的第二高斯噪声;  
根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手  
部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三  
维重建后的三维人体。
2. 根据权利要求1所述的方法,其特征在于,所述确定提升权重后的目标提示词,包括:  
获取预设的基础提示词和相机姿态信息;  
根据所述相机姿态信息确定方向信息,并获取身体部位的细节提示词和权重信息;  
在所述基础提示词的基础上,利用所述方向信息、所述细节提示词和所述权重信息进  
行权重提升,得到目标提示词。
3. 根据权利要求2所述的方法,其特征在于,基于初始人体模版构建初始三维高斯场,  
包括:  
获取初始化姿态的初始人体模版;  
根据所述初始人体模版进行初始化构建,得到三维高斯场。
4. 根据权利要求3所述的方法,其特征在于,所述依据所述人体图像和所述深度图像,  
利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,包括:  
通过噪声采样获取第一高斯噪声,并确定目标时刻;  
通过预设的采样器,利用所述第一高斯噪声和所述目标时刻,对所述人体图像和所述  
深度图像进行加噪,得到在所述目标时刻的噪声图像集。
5. 根据权利要求4所述的方法,其特征在于,根据所述第一高斯噪声和所述第二高斯噪  
声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三  
维高斯场,包括:  
根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算,得到深度损失函数,  
以及,基于所述深度损失函数进行梯度回传更新,得到更新三维高斯场;  
利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行  
对齐,得到相关性损失函数;  
利用所述相关性损失函数对所述更新三维高斯场进行二次更新,得到目标三维高斯  
场。
6. 根据权利要求5所述的方法,其特征在于,所述通过所述初始三维高斯场进行图像渲  
染,得到人体图像和深度图像,包括:  
利用预设的渲染器,按照公式 $I_{gs}, d_{gs} = \text{pyrenderer}(\mathcal{G}, camera\_pose)$ ,从所述相  
机姿态信息和所述初始三维高斯场中渲染出人体的人体图像和深度图像;

其中,  $I_{gs}$  为渲染出的人体图像,  $d_{gs}$  为渲染出的深度图像, pyrenderer 为渲染器,  $\mathcal{G}$  为初始三维高斯场, camera\_pose 为相机姿态信息。

7. 根据权利要求6所述的方法,其特征在于,所述基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声,包括:

利用公式  $\begin{cases} \widehat{\mathcal{E}}_j = UNet(I_{gs}^t, \mathcal{T}_{\ell v}) \\ \widehat{\mathcal{E}}_d = UNet(d_{gs}^t, \mathcal{T}_{\ell v}) \end{cases}$ , 以目标提示词为条件,将人体加噪图像和深度加噪图像分别输入UNet网络进行处理,得到所述目标时刻的上一个时刻的第二高斯噪声;

其中,  $\widehat{\mathcal{E}}_j$  为人体加噪图像对应的第二高斯噪声,  $\widehat{\mathcal{E}}_d$  为深度加噪图像对应的第二高斯噪声,  $\mathcal{T}_{\ell v}$  为目标提示词。

8. 根据权利要求7所述的方法,其特征在于,根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算,得到深度损失函数,以及,基于所述深度损失函数进行梯度回传更新,得到更新三维高斯场,包括:

根据  $\begin{cases} \mathcal{L}_j = MSE(\mathcal{E}_j, \widehat{\mathcal{E}}_j) \\ \mathcal{L}_d = MSE(\mathcal{E}_d, \widehat{\mathcal{E}}_d) \end{cases}$ , 计算均方误差,得到深度损失函数;

根据  $\mathcal{G} \leftarrow \mathcal{G} - \eta \nabla (\mathcal{L}_j + \mathcal{L}_d)$ , 进行梯度回传更新三维高斯场;

其中,  $\mathcal{L}_j$  和  $\mathcal{L}_d$  均为损失函数,  $\eta$  为学习率,在三维高斯场的训练更新过程中持续更新。

9. 根据权利要求7所述的方法,其特征在于,所述利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐,得到相关性损失函数,包括:

根据  $\mathcal{L}_{cr} = 1 - Pearson(d_{gs, hand}, d_{SMPL-X, hand})$ , 对两个手部深度图进行对齐,并使用皮尔森相关性损失;

其中,  $d_{gs, hand}$  为更新三维高斯场中的手部深度图,  $d_{SMPL-X, hand}$  为初始人体模版中的手部深度图。

10. 一种从自然语言描述文本中重建出三维人体的三维重建装置,其特征在于,包括:

提示词获取与模型构建模块,用于确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;

图像渲染模块,用于通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;

图像加噪模块,用于依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像;

扩散估计模块,用于基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声;

三维高斯场更新模块,用于根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。

## 一种从自然语言描述文本中重建出三维人体的三维重建方法及装置

### 技术领域

[0001] 本申请涉及计算机视觉技术领域,尤其涉及一种从自然语言描述文本中重建出三维人体的三维重建方法及装置。

### 背景技术

[0002] 目前,创建个性化的3D(即三维,或称3d)虚拟化身对于各种应用中的内容生成至关重要,如虚拟试穿、游戏和虚拟现实。用户希望能够自定义角色,从改变他们特定的身份(如衣服、头发等)到应用特定的艺术风格。然而,传统的创建数字化身的方法通常需要熟练的艺术家进行繁琐且耗时的工作,如多视角捕捉、纹理和动画。最近的化身技术进步展示了更用户友好的解决方案,即仅通过自然语言描述自动创建高质量的3D化身。

[0003] 现有的三维人体构建方法倾向于生成局部部分模糊的3D人类,如手和脚,其严重依赖于文本到图像模型,如扩散模型,容易生成不合理的人体部位,原因是文本到图像(T2I)模型通常没有明确模拟人体的细粒度结构,并且常常受到“Janus(多面体)”问题的困扰。例如,扩散模型生成的手图像可能存在多个部位的错误,如手部出现六个手指等,现有技术则倾向于忽视这些错误对象信息对3D人体生成的影响,导致3D人体生成中精确的局部肢体不足。由此可见,现有技术受不准确信息的影响无法生成精确的身体部位。

### 发明内容

[0004] 本申请提供了一种从自然语言描述文本中重建出三维人体的三维重建方法及装置,在关键提示词上使用提示权重有助于匹配生成的内容和文本提示,利用深度损失,对齐人体模型与局部身体部位深度,从而使得生成的3D人体能精确捕捉到人体部位的复杂特征,恢复精细的人体部位,解决现有的三维人体构建技术无法生成精确的身体部位的问题。

[0005] 第一方面,本申请提供了一种从自然语言描述文本中重建出三维人体的三维重建方法,包括:

[0006] 确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;

[0007] 通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;

[0008] 依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像;

[0009] 基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声;

[0010] 根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。

[0011] 可选的,所述确定提升权重后的目标提示词,包括:

- [0012] 获取预设的基础提示词和相机姿态信息；
- [0013] 根据所述相机姿态信息确定方向信息，并获取身体部位的细节提示词和权重信息；
- [0014] 在所述基础提示词的基础上，利用所述方向信息、所述细节提示词和所述权重信息进行权重提升，得到目标提示词。
- [0015] 可选的，基于初始人体模版构建初始三维高斯场，包括：
- [0016] 获取初始化姿态的初始人体模版；
- [0017] 根据所述初始人体模版进行初始化构建，得到三维高斯场。
- [0018] 可选的，所述依据所述人体图像和所述深度图像，利用目标时刻的第一高斯噪声进行加噪，得到噪声图像集，包括：
- [0019] 通过噪声采样获取第一高斯噪声，并确定目标时刻；
- [0020] 通过预设的采样器，利用所述第一高斯噪声和所述目标时刻，对所述人体图像和所述深度图像进行加噪，得到在所述目标时刻的噪声图像集。
- [0021] 可选的，根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传，并利用三维高斯场中的手部深度图进行损失更新，得到更新后的目标三维高斯场，包括：
- [0022] 根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算，得到深度损失函数，以及，基于所述深度损失函数进行梯度回传更新，得到更新三维高斯场；
- [0023] 利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐，得到相关性损失函数；
- [0024] 利用所述相关性损失函数对所述更新三维高斯场进行二次更新，得到目标三维高斯场。
- [0025] 可选的，所述通过所述初始三维高斯场进行图像渲染，得到人体图像和深度图像，包括：
- [0026] 利用预设的渲染器，按照公式 $I_{gs}, d_{gs} = \text{pyrenderer}(\mathcal{G}, camera\_pose)$ ，从所述相机姿态信息和所述初始三维高斯场中渲染出人体的人体图像和深度图像；
- [0027] 其中， $I_{gs}$ 为渲染出的人体图像， $d_{gs}$ 为渲染出的深度图像，pyrenderer为渲染器， $\mathcal{G}$ 为初始三维高斯场，camera\_pose为相机姿态信息。
- [0028] 可选的，所述基于所述目标提示词，结合所述噪声图像集进行扩散估计，得到所述目标时刻的上一个时刻的第二高斯噪声，包括：
- [0029] 利用公式 $\begin{cases} \widehat{\mathcal{E}}_I = UNet(I_{gs}^t, \mathcal{T}_{\ell v}) \\ \widehat{\mathcal{E}}_d = UNet(d_{gs}^t, \mathcal{T}_{\ell v}) \end{cases}$ ，以目标提示词为条件，将人体加噪图像和深度加噪图像分别输入UNet网络进行处理，得到所述目标时刻的上一个时刻的第二高斯噪声；
- [0030] 其中， $\widehat{\mathcal{E}}_I$ 为人体加噪图像对应的第二高斯噪声， $\widehat{\mathcal{E}}_d$ 为深度加噪图像对应的第二高斯噪声， $\mathcal{T}_{\ell v}$ 为目标提示词。
- [0031] 可选的，根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算，得到深度损失函数，以及，基于所述深度损失函数进行梯度回传更新，得到更新三维高斯场，包括：

- [0032] 根据 $\begin{cases} \mathcal{L}_j = MSE(\mathcal{E}_j, \widehat{\mathcal{E}}_j) \\ \mathcal{L}_d = MSE(\mathcal{E}_d, \widehat{\mathcal{E}}_d) \end{cases}$ ,计算均方误差,得到深度损失函数;
- [0033] 根据 $\mathcal{G} \leftarrow \mathcal{G} - \eta \nabla (\mathcal{L}_j + \mathcal{L}_d)$ ,进行梯度回传更新三维高斯场;
- [0034] 其中,  $\mathcal{L}_j$ 和 $\mathcal{L}_d$ 均为损失函数,  $\eta$ 为学习率,在三维高斯场的训练更新过程中持续更新。
- [0035] 可选的,所述利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐,得到相关性损失函数,包括:
- [0036] 根据 $\mathcal{L}_{cr} = 1 - Pearson(d_{gs,hand}, d_{SMPL-X,hand})$ ,对两个手部深度图进行对齐,并使用皮尔森相关性损失;
- [0037] 其中, $d_{gs,hand}$ 为更新三维高斯场中的手部深度图, $d_{SMPL-X,hand}$ 为初始人体模版中的手部深度图。
- [0038] 第二方面,本申请提供了一种从自然语言描述文本中重建出三维人体的三维重建装置,包括:
- [0039] 提示词获取与模型构建模块,用于确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;
- [0040] 图像渲染模块,用于通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;
- [0041] 图像加噪模块,用于依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像;
- [0042] 扩散估计模块,用于基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声;
- [0043] 三维高斯场更新模块,用于根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。
- [0044] 综上,本申请实施例利用初始人体模版构建初始三维高斯场,并获取提升权重后的目标提示词,随后通过初始三维高斯场进行图像渲染,得到人体图像和深度图像,依据人体图像和深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,然后基于目标提示词,结合噪声图像集进行扩散估计,得到目标时刻的上一个时刻的第二高斯噪声,结合第一高斯噪声和第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,该目标三维高斯场包含进行三维重建后的三维人体,由此,本申请实施例先在原提示词的基础上使用提示权重,从而有助于匹配生成的内容和文本提示,再利用深度损失,对齐人体模型与局部身体部位深度,恢复精细的人体部位,从而使得生成的3D人体能精确捕捉到人体部位的复杂特征,解决现有的三维人体构建技术无法生成精确的身体部位的问题。

## 附图说明

[0045] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本申请的实施例,并与说明书一起用于解释本申请的原理。

[0046] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,对于本领域普通技术人员而言,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0047] 图1为本申请实施例提供的一种从自然语言描述文本中重建出三维人体的三维重建方法的流程示意图;

[0048] 图2是本申请一个可选实施例提供的一种从自然语言描述文本中重建出三维人体的三维重建方法的步骤流程示意图;

[0049] 图3本申请一个示例提供的现有技术利用文本生成的模型和本实施例文本生成的模型的比对展示图;

[0050] 图4是本申请一个可选示例提供的一种基于文本的三维人体生成算法示意图;

[0051] 图5为本申请实施例提供的一种从自然语言描述文本中重建出三维人体的三维重建装置的结构框图;

[0052] 图6是本申请实施例提供的一种电子设备的结构示意图。

## 具体实施方式

[0053] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本申请的一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0054] 为便于对本申请实施例的理解,下面将结合附图以及具体实施例做进一步的解释说明,实施例并不构成对本申请实施例的限定。

[0055] 图1为本申请实施例提供的一种从自然语言描述文本中重建出三维人体的三维重建方法的流程示意图。如图1所示,本申请实施例提供的从自然语言描述文本中重建出三维人体的三维重建方法具体可以包括如下步骤:

[0056] 步骤110,确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场。

[0057] 在本实施例中,目标提示词是在原提示词的基础上增加了权重和方向信息后,得到的提示词;初始人体模版指的是初始的人体模版模型,其通常包含初始化的人体姿势,如人体姿势可以类似字母A,初始人体模版可以使用SMPL-X人体模板模型进行初始化得到,当然也可以通过其他的人体模版模型进行初始化构建得到,本实施例对此不做限制。

[0058] 在具体实现中,本实施例可以将SMPL-X构建的人体模版进行初始化,构建一个3D高斯场,该3D高斯场即可作为三维人体的三维表达,进行后续的处理。

[0059] 步骤120,通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像。

[0060] 步骤130,依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集。

[0061] 其中,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对

应的深度加噪图像。

[0062] 对步骤120-步骤130进行统一描述：

[0063] 在本实施例中,人体图像指的是人体的RGB(Redgreenblue,三原色)图像。第一高斯噪声可以包含两个高斯噪声,第一高斯噪声为目标时刻的高斯噪声,这两个高斯噪声分别与人体图像和深度图像一一对应,本实施例对此不做限制。

[0064] 在具体实现中,利用3d高斯场进行图像渲染,可以得到不同姿态渲染出的人体的RGB图像和深度图像,随后可以对获取到的两个图像都进行高斯噪声的加噪处理。为实现高斯加噪,本实施例可以先采样高斯噪声,得到两个图像对应的高斯噪声,然后利用图像对应的高斯噪声,在目标时刻对应的时间步上对图像进行加噪,从而得到RGB图像对应的人体加噪图像和深度图像对应的深度加噪图像。

[0065] 需要说明的是,本实施例中目标时刻可以理解为按照实际需求预设的时刻,也可以理解为采样高斯噪声时对应的时刻,目标时刻可以用t表示,本实施例对此不做限制。

[0066] 步骤140,基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声。

[0067] 在本实施例中,第二高斯噪声也可以包含了两个高斯噪声,第二高斯噪声为目标时刻的上一个时刻的高斯噪声,这两个高斯噪声可以分别与人体加噪图像和深度加噪图像一一对应;目标时刻的上一个时刻可以理解为在目标时刻的基础上,上一个时间步的时刻,如目标时刻为t,则目标时刻的上一个时刻可以为t-1,本实施例对此不做限制。

[0068] 在具体实现中,本实施例以目标提示词作为条件,分别利用两个加噪图像进行扩散估计,估计出时间为t-1时刻的噪声,作为第二高斯噪声。

[0069] 步骤150,根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。

[0070] 在具体实现中,本实施例可以两次对3D高斯场的进行更新,从而生成精细的人体部位。具体的,对采样的第一高斯噪声和第二高斯噪声分别进行损失计算,并进行梯度回传,由此进行了初次的3D高斯场更新。随后利用可见性机制,引入相关性损失,如皮尔逊相关性损失等,解决深度值不匹配的问题,将初次更新后的3D高斯场的手部深度图进行对齐,利用损失更新进一步对3D高斯场进行二次更新,以恢复精细人体部位的焦点深度损失,由此生成具有精确拓扑的精细人体部位,如手部等。

[0071] 需要说明的是,本实施例可以根据实际需求,选择人体部位对应的深度图进行对齐处理,从而实现生成精细的人体部位,深度图的对齐并不仅限于手部深度图。

[0072] 可见,本申请实施例通过人体模版构建的初始三维高斯场进行图像渲染,得到人体图像和深度图像,利用目标时刻的高斯噪声对图像进行加噪,得到噪声图像集,噪声图像集结合提升权重的提示词进行扩散估计,得到上一个时刻的高斯噪声,结合两个时刻的高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的进行三维重建后的三维人体。本实施例利用深度损失,对齐人体模型与局部身体部位深度,从而使得生成的3D人体能精确捕捉到人体部位的复杂特征,恢复精细的人体部位,解决现有的三维人体构建技术无法生成精确的身体部位的问题。

[0073] 参照图2,示出了本申请一个可选实施例提供的一种从自然语言描述文本中重建

出三维人体的三维重建方法的步骤流程示意图。该方法具体可以包括如下步骤：

[0074] 步骤210，确定提升权重后的目标提示词，以及，基于初始人体模版构建初始三维高斯场。

[0075] 对步骤210的描述可以参见上述步骤110，本实施例对此不再详述。

[0076] 在一个可选实施例中，本申请实施例确定提升权重后的目标提示词，具体可以包括：获取预设的基础提示词和相机姿态信息；根据所述相机姿态信息确定方向信息，并获取身体部位的细节提示词和权重信息；在所述基础提示词的基础上，利用所述方向信息、所述细节提示词和所述权重信息进行权重提升，得到目标提示词。

[0077] 在相关技术中，现有的三维人体构建方案通常利用文本到图像(T2I)生成工具，并结合3D人体先验知识，从一个文本提示中创建一个3D人类。其中，基于CLIP的方法倾向于利用CLIP模型中的人类语义信息来促进3D人体生成。由于CLIP模型主要关注文本和图像之间高级语义信息的对齐，其在准确生成细粒度细节方面经常挣扎，特别是在文本提示中指定的人体特征和服装方面。考虑到扩散模型在文本到图像合成中的强能力，现有方案多是寻求使用强大的扩散模型来促进3D人体生成，其通常使用得分蒸馏采样(SDS)损失从扩散模型中提取知识，以优化3D模型，如3D高斯散射(3DGS)或神经辐射场(NeRF)。然而，这些方法仍然存在着多种问题，包括但不限于：①现有技术主要倾向于生成局部部分模糊的3D人类，即倾向于生成模糊或不一致的人体部位，如手和头部，主要是由于忽略了关注局部几何和精确视角控制。如图3中的(a)图所示，扩散模型生成的手图像有六个手指，对于这种错误，现有方法倾向于忽视这些错误对象信息对3D人体生成的影响，从而导致3D人体生成中精确的局部肢体不足。尽管基于网格的方法保留了SMPL-X的手结构，但其在模拟头发和配饰的详细拓扑方面存在困难。由此可见，现有方案无法生成精确的身体部位；②未经精心设计的文本提示容易无法控制T2I模型生成高质量且与文本对齐的内容。参照图3，从图3中(b2)图可以看出，尽管文本提示指定了“两条腿”，但StableDiffusion2.0生成的图像描绘了整个身体。内容错位导致3D人体模型中局部身体部位模糊。此外，错位通常导致T2I模型生成的图像视角与实际渲染所用的相机角度不匹配。这种身体部位的多视角信息错位现象可能导致Janus(多面)问题(见图3中(c2)所示)。可见，如何控制T2I模型输出增强3D人体合理性的所需信息仍然是一个重要问题，也即现有技术无法合理控制模型输出增强3D人体。

[0078] 为解决上述问题，本实施例提出了从自然语言描述文本中重建出三维人体的三维重建方法，主要从三个方面，解决现有技术所存在的问题。首先，本实施例提出了新的文本驱动3D人体生成的焦点范式。利用该焦点范式，专注于3D人体的关键时刻和文本提示，以规范局部几何和视角方向。具体而言，本实施例预设了精细的基础提示词，在基础提示词的基础上，根据相机姿态加入方向信息，随后结合细节的身体部位提示词，并引入权重信息，使用提示词权重提升技术，从而得到目标提示词。本申请实施例在原提示词的基础上引入了具有细粒度视角控制的焦点视角依赖提示，以促进局部身体部位的3D一致性。

[0079] 进一步的，本实施例提出了用于恢复精细人体部位的焦点深度损失。与现有方法使用粗略控制如人体骨架来引导扩散模型不同，本实施例专注于对齐目标3D人体模型(即3D高斯)和SMPL-X中的局部身体部位的深度，其中包含精细人体肢体的先验信息。通过设计一个可见性机制，并引入了皮尔逊相关损失来解决深度值不匹配问题。焦点深度损失使本实施例能够生成具有精确拓扑的精细人体部位，如手。关于深度损失部分，可参见后续步骤

260-步骤280,本实施例在此先不详述。

[0080] 进一步的,本申请提出了用于解决局部身体部位3D不一致问题的焦点视角依赖SDS。为确保生成内容与文本提示的匹配,本实施例设计了一个具有细粒度视角控制的焦点视角依赖提示,利用了焦点视角依赖的提示,强调文本提示中关于人体的关键词,并使用提示加权机制。此外,为了实现3D一致性生成而不引入多个面孔,通过在原始文本提示中添加与视角相关的文本关于局部身体部位和全局身体,对后续的SDS损失进行了更精确的视角控制,从而利用焦点视角依赖SDS,解决了局部身体部位3D不一致问题。

[0081] 示例性的,基础提示词可以为:“一个穿着连帽衫的男孩”,将该基础提示词记作 $\mathcal{T}$ 。在此基础上,根据相机姿态加入方向信息,例如方向信息可以包含但不限于:前方、侧方以及后方等,得到新的提示词,如新的提示词可以为“一个穿着连帽衫的男孩,前方/侧方/后方”,表示为: $\mathcal{T}_{\ell v} = \mathcal{T} + \text{方向信息}$ 。随后加上细节的身体部位提示词,并使用提示词权重提升技术(用“ $(\cdot)^{++}$ ”表示),通过提示词权重提升技术提升特定身体部位的权重。例如,需要增加模型对手部细节的关注度时,提示词权重提升技术用“(手部) $^{++}$ ”表示,如“一个穿着连帽衫的男孩的(手部) $^{++}$ ,前方/侧方/后方”,表示为 $\mathcal{T}_{\ell v} = \mathcal{T} + (\text{手部})^{++} + \text{方向信息}$ ,由此得到目标提示词,目标提示词可以使用 $\mathcal{T}_{\ell v}$ 表示。

[0082] 在一个可选实施例中,本申请实施例基于初始人体模版构建初始三维高斯场,具体可以包括:获取初始化姿态的初始人体模版;根据所述初始人体模版进行初始化构建,得到三维高斯场。

[0083] 在具体实现中,本实施例可以使用SMPL-X人体模板模型初始化一个类似字母A的姿势的人体模版模型。将该模板用于初始化一个3D高斯场,该3D高斯场可以用G表示,该3D高斯场即可作为本实施例中需要训练的三维表达:

[0084]  $\mathcal{G} = \text{SMPL-X}(A\text{-pose})$

[0085] 本实施例可以通过3D高斯场渲染出多视角的人体图片。

[0086] 步骤220,通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像。

[0087] 对步骤220的描述可以参见上述步骤120,本实施例对此不再详述。

[0088] 可选的,上述通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像,具体可以包括:利用预设的渲染器,按照公式 $I_{gs}, d_{gs} = \text{pyrenderer}(\mathcal{G}, camera\_pose)$ ,从所述相机姿态信息和所述初始三维高斯场中渲染出人体的人体图像和深度图像;其中, $I_{gs}$ 为渲染出的人体图像, $d_{gs}$ 为渲染出的深度图像,pyrenderer为渲染器, $\mathcal{G}$ 为初始三维高斯场,camera\_pose为相机姿态信息。

[0089] 在具体实现中,本实施例可以预设一个渲染器,如pyrenderer渲染器,通过渲染器,从不同的采样相机姿态中渲染出人体的RGB图像 $I_{gs}$ 和深度图像 $d_{gs}$ 。考虑到现有文生图模型很难合成类似手部等精致的身体部位,本实施例在生成局部身体部分的同时添加视图方向提供了更精确的视图控制,为确保能够渲染出不同视图方向,本实施例利用公式:

[0090]  $I_{gs}, d_{gs} = \text{pyrenderer}(\mathcal{G}, camera\_pose)$ ,实现图像渲染。

[0091] 步骤230,通过噪声采样获取第一高斯噪声,并确定目标时刻。

[0092] 步骤240,通过预设的采样器,利用所述第一高斯噪声和所述目标时刻,对所述人体图像和所述深度图像进行加噪,得到在所述目标时刻的噪声图像集。

[0093] 其中,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像。

[0094] 对步骤230-步骤240进行统一描述:

[0095] 在具体实现中,本实施例预设了采样器(简称DDIM采样器),通过采样器对人体图像和深度图像进行加噪。具体而言,在目标时刻,通过DDIM采样器,结合采样到的第一高斯噪声,分别对人体图像和深度图像进行加噪,得到人体图像对应的人体加噪图像和深度图像对应的深度加噪图像。

[0096] 为实现对人体图像和深度图像的加噪处理,本实施例设置了公式实现加噪。其中,

图像加噪处理的公式可以表达为: $\begin{cases} I_{gs}^t = DDIM(I_{gs}, \mathcal{E}_j, t) \\ d_{gs}^t = DDIM(d_{gs}, \mathcal{E}_d, t) \end{cases}$ 。本实施例中,RGB图像表

示为 $I_{gs}$ ,深度图像表示为 $d_{gs}$ ,采样高斯噪声(即第一高斯噪声)包含 $\mathcal{E}_j$ 和 $\mathcal{E}_d$ ,在时间步(即目标时刻)t上加噪,其中, $\mathcal{E}_j$ 用于对RGB图像 $I_{gs}$ 进行加噪,得到人体加噪图像 $I_{gs}^t$ ; $\mathcal{E}_d$ 用于对深度图像 $d_{gs}$ 进行加噪,得到深度噪声图像 $d_{gs}^t$ 。

[0097] 步骤250,基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声。

[0098] 对步骤250的描述可以参见上述步骤140,本实施例对此不再详述。

[0099] 可选的,上述基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声,具体可以包括:利用公式 $\begin{cases} \widehat{\mathcal{E}}_j = UNet(I_{gs}^t, \mathcal{T}_{\ell v}) \\ \widehat{\mathcal{E}}_d = UNet(d_{gs}^t, \mathcal{T}_{\ell v}) \end{cases}$ ,以目标提示词为条件,将人体加噪图像和深度加噪图像分别输入UNet网络进行处理,得到所述目标时刻的上一个时刻的第二高斯噪声;其中,  $\widehat{\mathcal{E}}_j$ 为人体加噪图像对应的第二高斯噪声,  $\widehat{\mathcal{E}}_d$ 为深度加噪图像对应的第二高斯噪声,  $\mathcal{T}_{\ell v}$ 为目标提示词。

[0100] 在本实施例中,可以利用StableDiffusion的UNet实现第二高斯噪声的噪声估计,以 $T_{1v}$ 作为条件(condition),分别将人体加噪图像和深度加噪图像作为输入,使用StableDiffusion的UNet估算t-1时刻的高斯噪声。

[0101] 步骤260,根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算,得到深度损失函数,以及,基于所述深度损失函数进行梯度回传更新,得到更新三维高斯场。

[0102] 在具体实现中,对3d高斯场的更新可以包含两次更新过程,第一次更新可以使用梯度回传更新。具体的,本实施例可以先对第一高斯噪声和第二高斯噪声,分别做MSE loss(深度损失),计算均方误差,并进行梯度回传,更新3D高斯场,直到3D高斯场渲染的图片与stablediffusion生成的图片分布足够接近。

[0103] 在一个可选的实施方式中,本申请实施例根据所述第一高斯噪声和所述第二高斯

噪声进行均方误差计算,得到深度损失函数,以及,基于所述深度损失函数进行梯度回传更新,得到更新三维高斯场,具体可以包括:根据

$$\begin{cases} \mathcal{L}_j = MSE(\mathcal{E}_j, \widehat{\mathcal{E}}_j) \\ \mathcal{L}_d = MSE(\mathcal{E}_d, \widehat{\mathcal{E}}_d) \end{cases}$$

计算均方误差,得到

深度损失函数;根据  $\mathcal{G} \leftarrow \mathcal{G} - \eta \nabla (\mathcal{L}_j + \mathcal{L}_d)$ , 进行梯度回传更新三维高斯场;其中,  $\mathcal{L}_j$  和  $\mathcal{L}_d$  均为损失函数,  $\eta$  为学习率, 在三维高斯场的训练更新过程中持续更新。训练过程持续进行, 直到3D高斯场渲染的图片分布与Stable Diffusion生成的图片分布足够接近。

[0104] 步骤270,利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐,得到相关性损失函数。

[0105] 在具体实现中,对3d高斯场的第二次更新主要可以涉及到深度图对齐和相关性损失函数更新。具体的,在对齐手部深度图时,本实例从3D高斯场中提取手部的深度图,将其与SMPL-X的手部深度图进行对齐,计算皮尔森相关性loss,作为相关性损失函数。

[0106] 在具体实现中,为实现手部深度图的对齐,使得3D高斯场渲染的图片与 stable diffusion 生成的图片分布足够接近,本实施例通过公式  $\mathcal{L}_{cr} = 1 - Pearson(d_{gs,hand}, d_{SMPL-X,hand})$ , 实现深度图的对齐。

[0107] 在一个可选实施例中,本申请实施例利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐,得到相关性损失函数,具体可以包括:根据  $\mathcal{L}_{cr} = 1 - Pearson(d_{gs,hand}, d_{SMPL-X,hand})$ , 对两个手部深度图进行对齐,并使用皮尔森相关性损失;其中,  $d_{gs,hand}$  为更新三维高斯场中的手部深度图,  $d_{SMPL-X,hand}$  为初始人体模版中的手部深度图。

[0108] 步骤280,利用所述相关性损失函数对所述更新三维高斯场进行二次更新,得到目标三维高斯场。

[0109] 在本实施例中,完成两个模型的手部深度图对齐后,即可利用对齐得到的皮尔森相关性loss,进一步更新3D高斯场。

[0110] 在实际实现中,为实现进一步更新3D高斯场,本实施例通过公式:  $\mathbf{G} \leftarrow \mathbf{G} - \eta \nabla \mathcal{L}_{cr}$ , 在已更新的3D高斯场的基础上进一步更新3D高斯场,得到目标三维高斯场。

[0111] 示例性的,结合图3的对比图: (b1) 和 (b2), 以及 (c1) 和 (c2), 可见,本实施例通过加入权重,进行权重提升,相比于现有的T2I方案能够更加准确的生成人体部位,并且参照图4所示的基于文本的三维人体生成算法示意图,可见,广泛的实验和消融研究表明,本申请提供的三维重建算法(或称:从文本描述生成精细一致肢体细节的三维穿衣人重建算法),在各种提示词下表现出优越性,生成的3D人体精确捕捉了人体部位的复杂特征,在文本驱动的3D人体生成方面具备有效性。

[0112] 综上,本申请实施例通过确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场,通过初始三维高斯场进行图像渲染,得到人体图像和深度图像,随后通过预设的采样器,利用采样到的第一高斯噪声,在目标时刻对人体图像和深度图像进行加噪,得到噪声图像集,以基于目标提示词,结合噪声图像集进行扩散估计,得到目标时刻的上一个时刻的第二高斯噪声,根据第一高斯噪声和第二高斯噪声进行均方误差计算,得

到深度损失函数,以及,基于深度损失函数进行梯度回传更新,得到更新三维高斯场,从而实现三维高斯场的第一次更新,随后利用更新三维高斯场中的手部深度图和人体模版模型中的手部深度图进行对齐,得到相关性损失函数,利用相关性损失函数对更新三维高斯场进行二次更新,得到目标三维高斯场。由此,本申请实施例为了在优化3D模型时施加更精确的控制,提出了Focal Gaussian,专注于在使用固定扩散模型优化基于3DGS的人体模型时关注人体的关键部位,即聚焦于人体关键部位。具体的,为了生成精细的人体部位,本实施例利用SMPL-X的几何先验,通过焦点深度损失来对齐SMPL-X手的深度与相应的3DGS(三维高斯场)深度。考虑到这两个深度图的值不匹配,本实施例引入了皮尔逊相关损失进行规范,利用焦点深度损失,通过对齐3D人体模型与SMPL-X的局部身体部位深度来恢复精细的人体部位,从而解决现有方法难以生成精细的身体部位问题。此外,为了解决Janus问题,本实施例提出了焦点视角依赖的SDS损失,专注于关键组件,并在局部身体部位上使用细粒度视角控制。具体来说,为了实现3D一致性生成而不引入多个面孔,本实施例通过在原始文本提示中添加与视角相关的文本关于局部身体部位和全局身体,对SDS损失进行了更精确的视角控制,并且为了确保生成内容与文本提示的匹配,本实施例设计了一个焦点视角依赖的提示,强调文本提示中关于人体的关键词,并使用提示加权机制,利用焦点视角相关的提示词,强调并在局部身体部位的关键词上施加视角控制,以解决现有技术受限于“Janus”的问题。

[0113] 需要说明的是,对于方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本申请实施例并不受所描述的动作顺序的限制,因为依据本申请实施例,某些步骤可以采用其他顺序或者同时进行。

[0114] 如图5所示,本申请实施例还提供了一种从自然语言描述文本中重建出三维人体的三维重建装置500,包括:

[0115] 提示词获取与模型构建模块510,用于确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;

[0116] 图像渲染模块520,用于通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;

[0117] 图像加噪模块530,用于依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像;

[0118] 扩散估计模块540,用于基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声;

[0119] 三维高斯场更新模块550,用于根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。

[0120] 可选的,所述提示词获取与模型构建模块510,包括:

[0121] 提示词和姿态获取子模块,用于获取预设的基础提示词和相机姿态信息;

[0122] 方向信息获取子模块,用于根据所述相机姿态信息确定方向信息,并获取身体部位的细节提示词和权重信息;

[0123] 权重提升子模块,用于在所述基础提示词的基础上,利用所述方向信息、所述细节

提示词和所述权重信息进行权重提升,得到目标提示词;

[0124] 人体模板获取子模块,用于获取初始化姿态的初始人体模版;

[0125] 初始化构建子模块,用于根据所述初始人体模版进行初始化构建,得到三维高斯场。

[0126] 可选的,所述图像渲染模块520,具体用于:利用预设的渲染器,按照公式 $I_{gs}, d_{gs} = pyrenderer(\mathcal{G}, camera\_pose)$ ,从所述相机姿态信息和所述初始三维高斯场中渲染出人体的人体图像和深度图像;其中, $I_{gs}$ 为渲染出的人体图像, $d_{gs}$ 为渲染出的深度图像,pyrenderer为渲染器, $\mathcal{G}$ 为初始三维高斯场,camera\_pose为相机姿态信息。

[0127] 可选的,所述图像加噪模块530,包括:

[0128] 采样子模块,用于通过噪声采样获取第一高斯噪声,并确定目标时刻;

[0129] 夹在子模块,用于通过预设的采样器,利用所述第一高斯噪声和所述目标时刻,对所述人体图像和所述深度图像进行加噪,得到在所述目标时刻的噪声图像集。

[0130] 可选的,所述扩散估计模块540,具体用于:利用公式 $\begin{cases} \widehat{\mathcal{E}}_j = UNet(I_{gs}^t, \mathcal{T}_{\ell v}) \\ \widehat{\mathcal{E}}_d = UNet(d_{gs}^t, \mathcal{T}_{\ell v}) \end{cases}$ ,以

目标提示词为条件,将人体加噪图像和深度加噪图像分别输入UNet网络进行处理,得到所述目标时刻的上一个时刻的第二高斯噪声;其中, $\widehat{\mathcal{E}}_j$ 为人体加噪图像对应的第二高斯噪声, $\widehat{\mathcal{E}}_d$ 为深度加噪图像对应的第二高斯噪声, $\mathcal{T}_{\ell v}$ 为目标提示词。

[0131] 可选的,所述三维高斯场更新模块550,包括:

[0132] 第一更新子模块,用于根据所述第一高斯噪声和所述第二高斯噪声进行均方误差计算,得到深度损失函数,以及,基于所述深度损失函数进行梯度回传更新,得到更新三维高斯场;

[0133] 对齐子模块,用于利用所述更新三维高斯场中的手部深度图和所述初始人体模版中的手部深度图进行对齐,得到相关性损失函数;

[0134] 第二更新子模块,用于利用所述相关性损失函数对所述更新三维高斯场进行二次更新,得到目标三维高斯场。

[0135] 可选的,所述第一更新子模块,具体用于:根据 $\begin{cases} \mathcal{L}_j = MSE(\mathcal{E}_j, \widehat{\mathcal{E}}_j) \\ \mathcal{L}_d = MSE(\mathcal{E}_d, \widehat{\mathcal{E}}_d) \end{cases}$ ,计算均方误

差,得到深度损失函数;根据 $\mathcal{G} \leftarrow \mathcal{G} - \eta \nabla (\mathcal{L}_j + \mathcal{L}_d)$ ,进行梯度回传更新三维高斯场;其中, $\mathcal{L}_j$ 和 $\mathcal{L}_d$ 均为损失函数, $\eta$ 为学习率,在三维高斯场的训练更新过程中持续更新。

[0136] 可选的,所述对齐子模块,具体用于:根据 $\mathcal{L}_{cr} = 1 - Pearson(d_{gs, hand}, d_{SMPL-X, hand})$ ,对两个手部深度图进行对齐,并使用皮尔森相关性损失;其中, $d_{gs, hand}$ 为更新三维高斯场中的手部深度图, $d_{SMPL-X, hand}$ 为初始人体模版中的手部深度图。

[0137] 需要说明的是,本申请实施例提供的从自然语言描述文本中重建出三维人体的三维重建装置可执行本申请任意实施例所提供的从自然语言描述文本中重建出三维人体的三维重建方法,具备执行方法相应的功能和有益效果。

[0138] 在具体实现中,上述从自然语言描述文本中重建出三维人体的三维重建装置可以集成在设备中,使得该设备可以利用提升权重的提示词,在三维人体的局部身体部位施加视角控制,精确捕捉人体部位的复杂特征,作为电子设备,实现生成精确的身体部位。该电子设备可以是由两个或多个物理实体构成,也可以是一个物理实体构成,如电子设备可以是个人计算机(Personal Computer, PC)、电脑、服务器等,本申请实施例对此不作具体限制。

[0139] 如图6所示,本申请实施例提供了一种电子设备,包括处理器111、通信接口112、存储器113和通信总线114,其中,处理器111,通信接口112,存储器113通过通信总线114完成相互间的通信;存储器113,用于存放计算机程序;处理器111,用于执行存储器113上所存放的程序时,实现前述任意一个方法实施例提供的从自然语言描述文本中重建出三维人体的三维重建方法的步骤。示例性的,从自然语言描述文本中重建出三维人体的三维重建方法的步骤可以包括如下步骤:确定提升权重后的目标提示词,以及,基于初始人体模版构建初始三维高斯场;通过所述初始三维高斯场进行图像渲染,得到人体图像和深度图像;依据所述人体图像和所述深度图像,利用目标时刻的第一高斯噪声进行加噪,得到噪声图像集,所述噪声图像集包含所述人体图像对应的人体加噪图像和所述深度图像对应的深度加噪图像;基于所述目标提示词,结合所述噪声图像集进行扩散估计,得到所述目标时刻的上一个时刻的第二高斯噪声;根据所述第一高斯噪声和所述第二高斯噪声进行梯度回传,并利用三维高斯场中的手部深度图进行损失更新,得到更新后的目标三维高斯场,所述目标三维高斯场包含进行三维重建后的三维人体。

[0140] 本申请实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现如前述任意一个方法实施例提供的从自然语言描述文本中重建出三维人体的三维重建方法的步骤。

[0141] 需要说明的是,在本文中,诸如“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0142] 以上所述仅是本申请的具体实施方式,使本领域技术人员能够理解或实现本申请。对这些实施例的多种修改对本领域的技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本申请的精神或范围的情况下,在其它实施例中实现。因此,本申请将不会被限制于本文所示的这些实施例,而是要符合与本文所申请的原理和新颖特点相一致的最宽的范围。

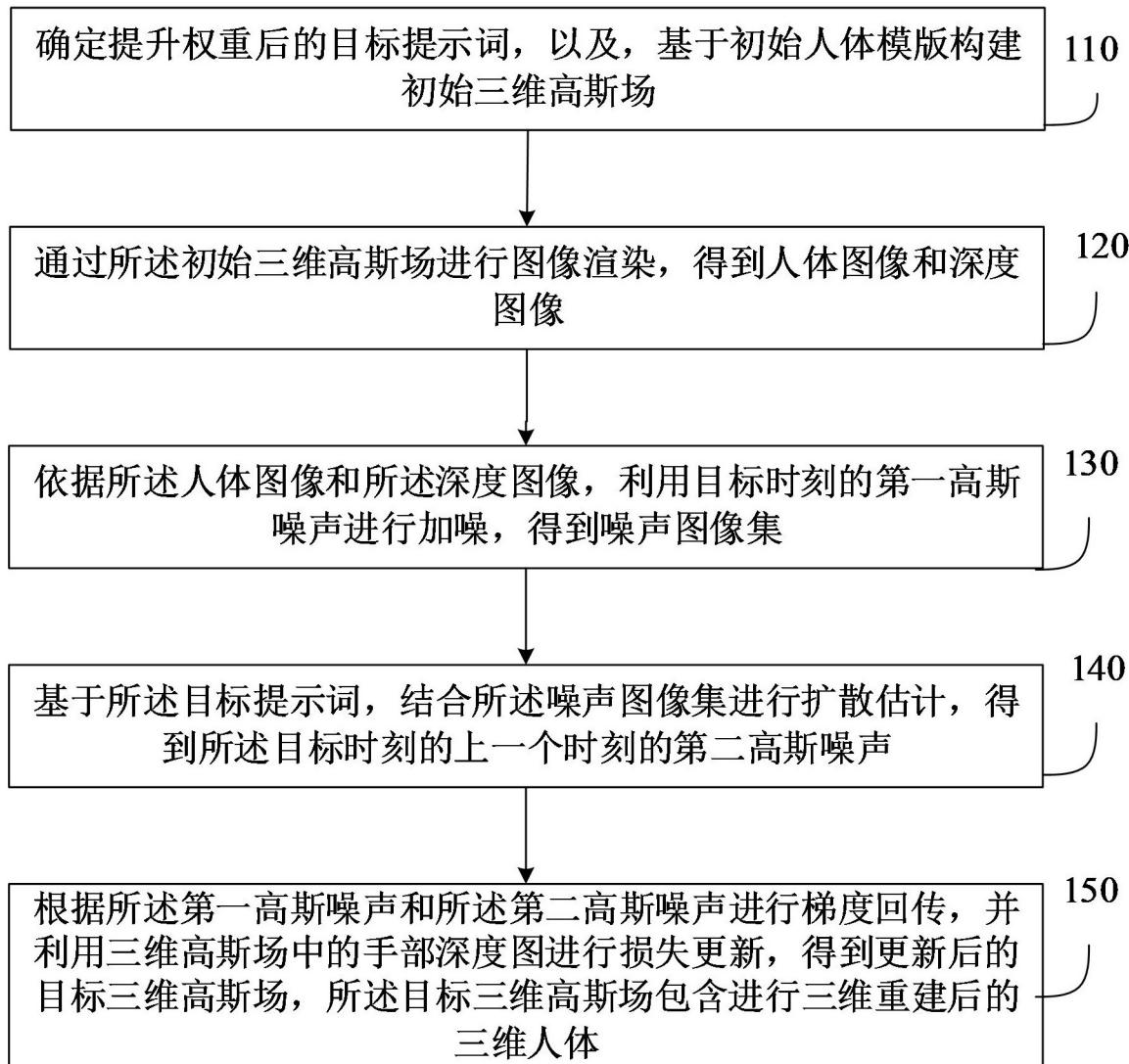


图1

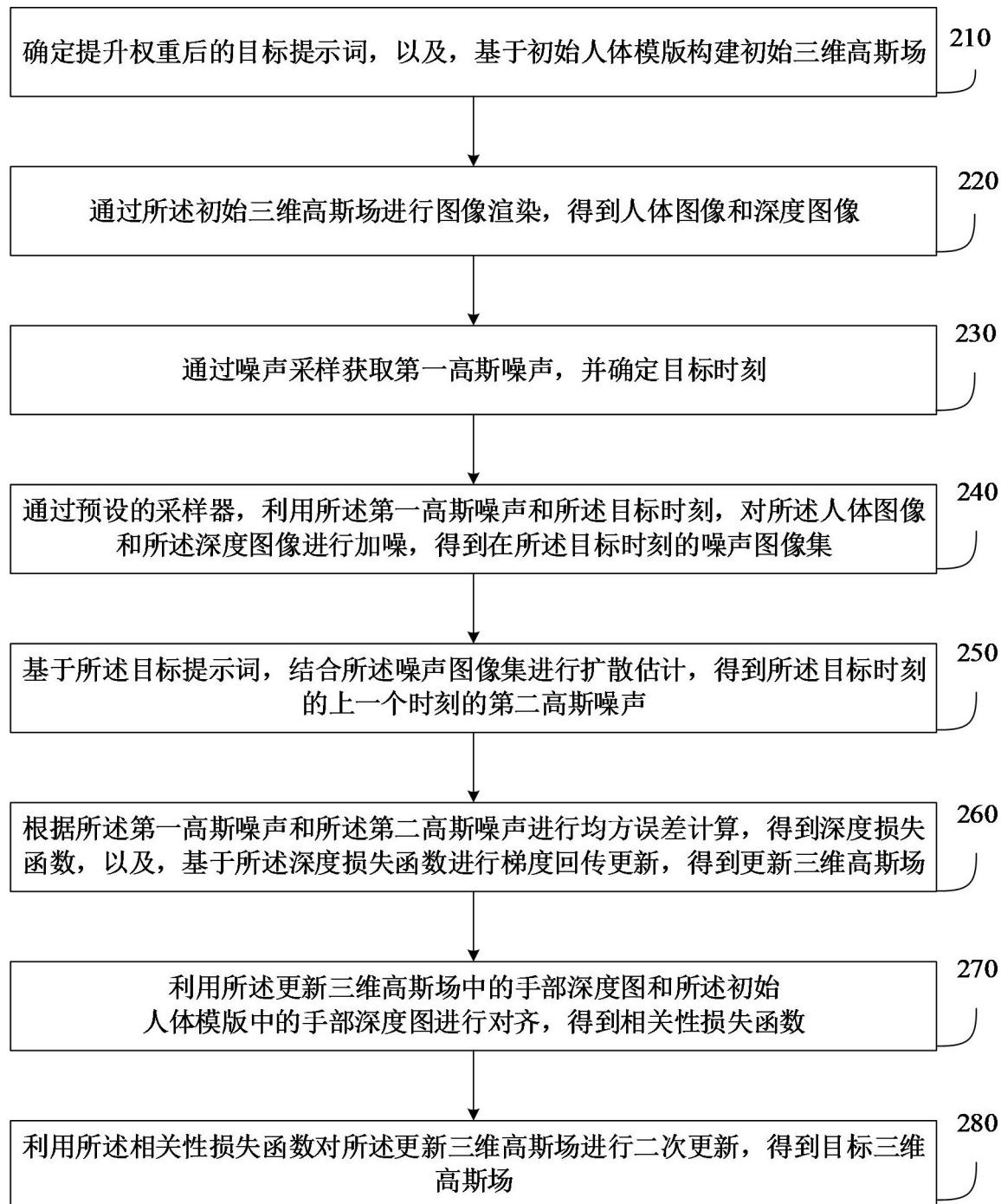


图2

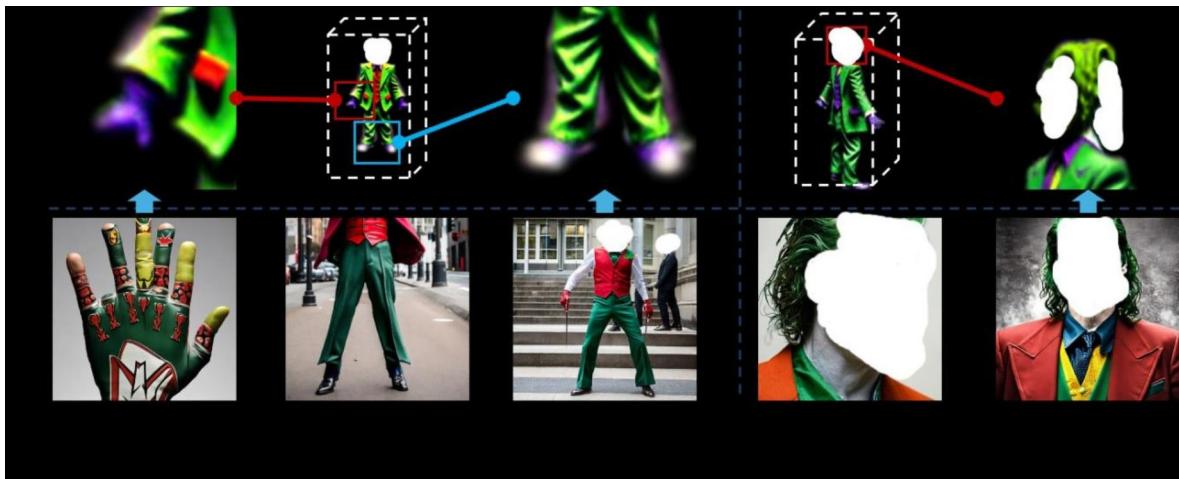


图3

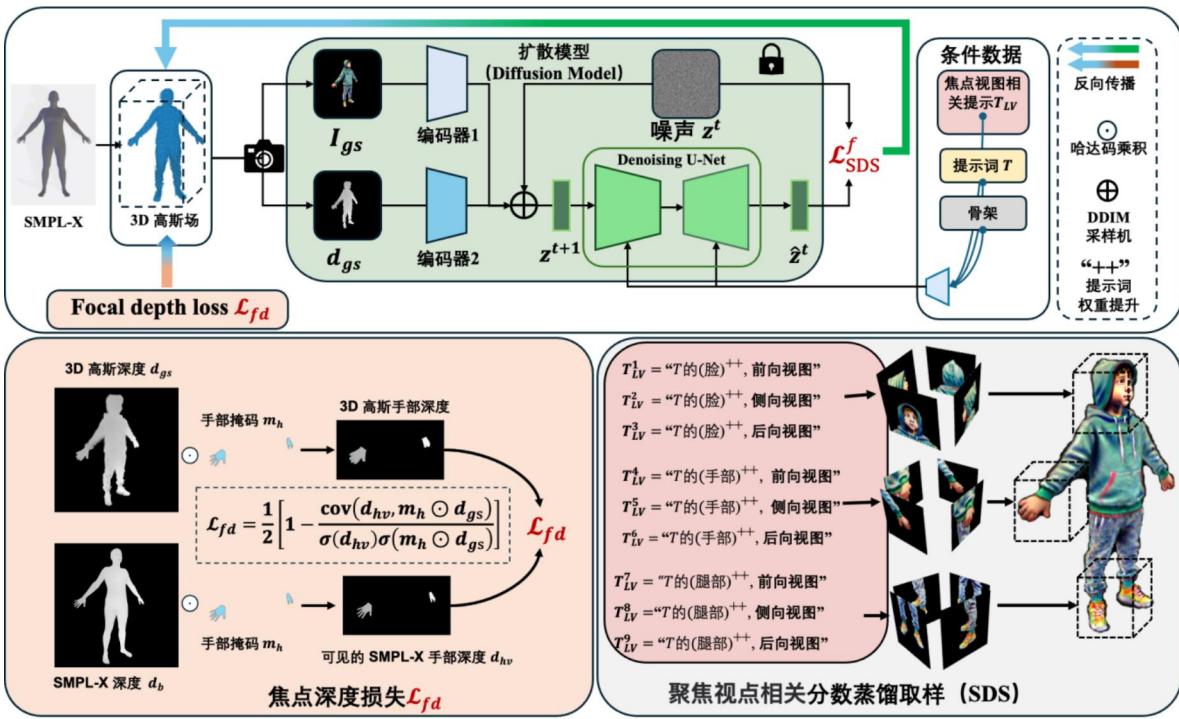


图4

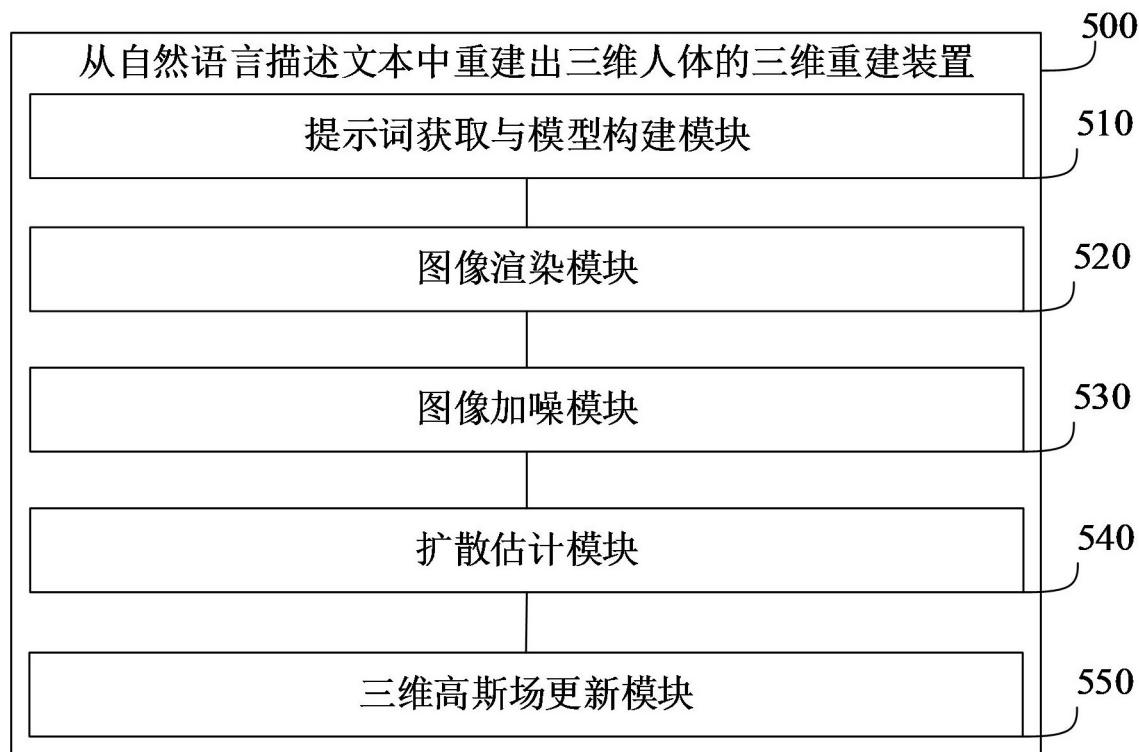


图5

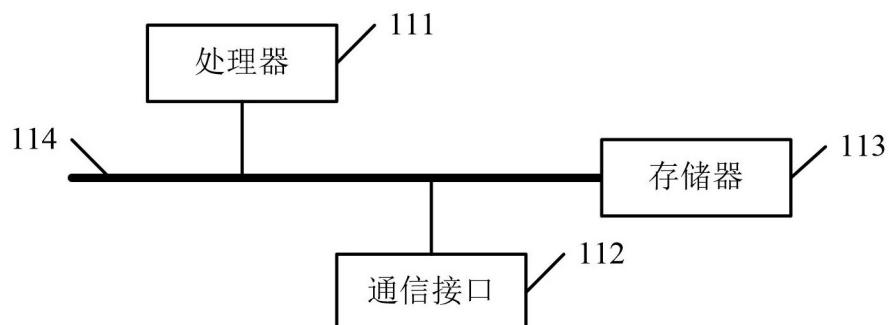


图6