



(12) 发明专利申请

(10) 申请公布号 CN 112819951 A

(43) 申请公布日 2021.05.18

(21) 申请号 202110186682.X

(22) 申请日 2021.02.09

(71) 申请人 北京工业大学

地址 100124 北京市朝阳区平乐园100号

(72) 发明人 王瑾 张天意 朱青

(74) 专利代理机构 北京思海天达知识产权代理有限公司 11203

代理人 张慧

(51) Int. Cl.

G06T 17/00 (2006.01)

G06T 7/12 (2017.01)

G06T 7/564 (2017.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

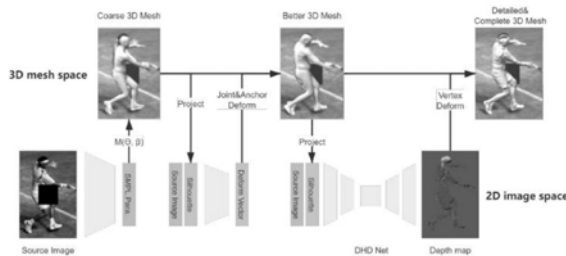
权利要求书2页 说明书6页 附图3页

(54) 发明名称

一种基于深度图修复的带遮挡三维人体重建方法

(57) 摘要

本发明涉及一种基于深度图修复的带遮挡三维人体重建方法,用于解决由单张包含遮挡物的RGB人体图像恢复三维人体模型姿态、体型、表面细节的问题。包括基础模型构建阶段,通过现有SMPL模型构建方法构建基础三维人体模型;体型姿态优化阶段,通过Joint编码器网络,优化三维人体关节位置,通过Anchor编码器网络,优化三维人体体型特征;表面细节优化阶段,提出了DHDNet网络结构,通过重建完整人体深度图像,恢复三维人体表面细节。同时,本发明构建了包含大量人体RGB图像与合成人体深度图像的DepthHuman数据集。达到了仅通过单张包含遮挡物的RGB人体图像重建完整并且包含丰富细节的三维人体模型的效果。



1. 一种基于深度图修复的带遮挡三维人体重建方法,用于对有遮挡的RGB图像进行三维重构,其特征在于包括基础模型构建阶段、体型姿态优化阶段、以及表面细节优化阶段:

所述的基础模型构建阶段,用于获得待重构原始图像的基础三维人体模型,具体如下:

步骤一:使用SMPL模型构建方法,输入单张待重构原始图像,得到待重构原始图像的基础三维人体模型,所述基础三维人体模型包括人体基础关节点、人体基础锚点和所有点;所述的待重构图像为受遮挡的RGB图像;

所述的体型姿态优化阶段,用于优化基础三维人体模型,具体如下:

步骤二:使用两个编码器网络分别获取关节点位移向量和锚点位移向量,所述的两个编码器网络即Joint编码器网络和Anchor编码器网络,

其中,Joint编码器网络用于提取原始图像中的人体姿态特征,即输出关节点位移向量,其输入为原始图像和人体轮廓信息,具体的,人体轮廓信息由上一步得到的三维人体模型投影所得,使用人体轮廓信息对原始图像进行剪裁,将原始图像中人体轮廓之外的颜色信息剔除,得到输入数据;

Anchor编码器网络用于提取原始图像中的人体体型特征,即输出为锚点位移向量,其输入与Joint编码器网络的输入相同;

步骤三:利用关节点位移向量移动基础关节点,得到优化后的关节点;同理,利用锚点位移向量移动基础锚点,得到优化后的锚点,完成体型姿态优化;

表面细节优化阶段,用于得到包含丰富表面细节的三维人体模型,具体如下:

步骤四:构建并训练DHDNet网络,DHDNet网络用于从包含遮挡物的RGB图像生成完整的人体深度信息;

步骤五:利用步骤二得到的人体轮廓信息对待重构原始图像进行剪裁,将裁剪后的图像输入训练好的DHDNet,输出完整的人体深度信息;

步骤六:利用步骤五得到的人体深度信息,对优化后的基础三维人体模型的所有点进行形变,最终得到包含丰富表面细节的三维人体模型。

2. 根据权利要求1所述的一种基于深度图修复的带遮挡三维人体重建方法,其特征在于:DHDNet中的生成器网络结构以U-Net网络结构为基础,输入为RGB三通道原始图像加一维的遮罩层,输出为一维的深度信息,中间经历了五次由卷积操作和池化操作组成的下采样操作,然后五次由反卷积操作组成的上采样操作,每次上采样会合并下采样路径中同层的特征图,且在神经网络的第四层引入了Shift-Layer,用于更好的得到完整深度信息。

3. 根据权利要求1所述的一种基于深度图修复的带遮挡三维人体重建方法,其特征在于:步骤四进一步包括:

(4.1) 构建训练数据集,包括标记了2d关节点位置的原始图像,原始图像对应的人体轮廓信息,以及原始图像对应的人体深度信息,其中,原始图像的中心区域有四分之一图像边长的正方形遮罩用来模拟图像中的遮挡物;

(4.2) 构建DHDNet网络,DHDNet采用对抗神经网络结构,由生成器与判别器两部分组成;

(4.3) 利用训练数据集训练DHDNet网络。

4. 根据权利要求1所述的一种基于深度图修复的带遮挡三维人体重建方法,其特征在于:DHDNet网络的损失函数如下:

$$L_{\text{final}} = L_{\text{depth}} + \lambda_{\text{gan}} L_{\text{gan}} + \lambda_{\text{photo}} L_{\text{photo}} + \lambda_c L_c + \lambda_s L_s \quad (6)$$

其中,  $\lambda_{\text{gan}}, \lambda_{\text{photo}}, \lambda_c, \lambda_s$  为对应损失函数的权重,  $L_{\text{depth}}$  为MSE损失函数,  $L_{\text{gan}}$  表达式如下:

$$L_{\text{gan}} = \min_G \max_D \mathbb{E}_{x \in X} [\log(D(x))] + \mathbb{E}_{z \in Z} [\log(1 - D(G(z)))] \quad (7)$$

其中G和D表示GAN中的生成器和判别器, X表示Ground Truth图像集, 即训练数据集中原始图像对应的人体深度信息做为真实人体深度图, Z表示输入图像集, 即训练数据集中原始图像, 输入图像为RGB三通道原始图像加一维的遮罩层,

$L_{\text{photo}}$  表示光度损失函数, 具体如下:

$$L_{\text{photo}} = \left\| \rho \sum_{k=1}^9 l_k H_k(n) - I \right\|_2 \quad (8)$$

其中,  $H_k$  表示球谐函数的基,  $l_k$  表示球谐函数系数, I表示输入RGB图像的单通道强度图。

5. 根据权利要求3所述的一种基于深度图修复的带遮挡三维人体重建方法, 其特征在于: 为了恢复高频细节, DHDNet在输出的深度信息以及真实深度信息上使用了VGG特征提取器网络, 并使用  $L_c$  和  $L_s$  两个损失函数进行监督, 其中,  $L_c$  表示内容损失函数:

$$L_c(p, x, l) = \frac{1}{2} \sum_{i,j} (F_{i,j}^l - P_{i,j}^l)^2 \quad (9)$$

其中, p和x分别代表由DHDNet输出的生成图像和depthhuman数据集中的真实人体深度图像,  $P^l$  表示DHDNet输出的生成图像经过VGG特征提取器网络第1层的特征信息,  $F^l$  表示真实人体深度图像经过VGG特征提取器网络第1层的特征信息, i和j代表像素坐标;

$L_s$  表示风格损失函数:

$$L_s(a, x) = \sum_{l=0}^L w_l E_l \quad (10)$$

其中, a表示由DHDNet输出的生成图像, x表示depthhuman数据集中的真实人体深度图像,  $w_l$  表示VGG网络各层贡献度的权重,  $E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (\text{Gram}(F_{i,j}^l) - \text{Gram}(P_{i,j}^l))^2$ , Gram表示Gram矩阵, 使用Gram矩阵计算表达图像特征之间的相关性, i和j代表像素坐标,  $N_1$  表示VGG网络第1层中特征的通道宽度,  $M_1$  表示VGG网络第1层中特征的通道高度。

## 一种基于深度图修复的带遮挡三维人体重建方法

### 技术领域：

[0001] 本发明涉及计算机图像处理领域、计算机图形学领域，具体涉及一种基于深度图修复的带遮挡单视图三维人体重建方法。

### 背景技术：

[0002] 近年来，单目相机的三维人体重建技术是一个十分热门的话题，这项技术在VR/AR领域有着广泛的应用场景，例如VR换装，VR聊天室等。同时，在影视制作、游戏研发领域也有巨大的应用前景。现有的单视图三维人体重建技术存在重建模型缺乏细节，受遮挡区域人体表面纹理信息完全丢失的问题。这导致无法通过一张部分人体受到遮挡的输入图像，重建出完整的包含丰富细节的三维人体模型。实际应用场景中，受遮挡的人体图像是十分常见的，现有方法无法满足应用需求。

[0003] 目前，单视图三维人体重建工作已经取得了一些进展，主要分为如下两类：基于模板适应的方法和基于特征匹配的方法。基于模板适应的方法依赖于预先训练好的参数化人体模型，例如SMPL (skinned multi-person linear model) 模型和SCAPE模型。Bogo等人提出的SMPLify使用卷积神经网络以及三维空间人体关节点信息最小化二维空间人体关节点的误差距离。Kanazawa等人提出了一个端到端系统HMR (Human Mesh Recovery)，HMR仅通过二维标注就可以通过一个对抗损失函数有效地控制住重建人体模型的姿态。以HMR为基础，Zhu等人对基础模型进行三阶段的形变操作，使其拥有更多衣服细节，更准确的体型和姿态。Alldieck等人通过UV贴图将人体表面细节展开到二维图像上。从而在UV空间进行法线贴图和顶点位移量的回归，取代在三维网格空间上进行这些操作。基于特征匹配的方法直接从单个RGB图像回归三维拓扑结构，而非从一个标准的人体模型上优化而来。典型的方法有DoubleFusion以及HybridFusion，这两种方法使用单目深度传感器实时捕捉人体运动信息，后者添加了稀疏的IMUs (inertial measurement units) 作为辅助手段。DeepHuman通过体特征转换技术将图像信息融入三维模型中，然后使用一个法线细化网络优化模型表面的细节。

[0004] 综上，现有的单视图三维人体重建方法所得到的模型大部分缺少表面细节，并且现有方法没有考虑输入图像中存在遮挡物的情况，而存在遮挡物在实际应用场景中是十分常见的。

### 发明内容

[0005] 本发明的目的是克服上述现有技术的缺点，提供一种质量高的单视图三维人体重建方法，该方法以单张包含遮挡物的RGB图像作为输入，输出包含丰富细节的完整三维人体模型。

[0006] 本发明的单视图三维人体重建方法，采用从粗糙到细致分阶段优化的策略，构建了一个完整的端到端系统。首先通过现有的基于SMPL模型的方法获取一个低质量的标准三维人体模型，然后根据从输入图像中提取的特征信息分阶段选取人体“关节点”、均匀分布

在人体模型上的若干“锚点”以及四倍细分下三维空间内的“所有点”进行形变,从而获得高质量的重建结果。端到端系统的流程图见图1,对于三阶段操作点的定义见图6,上述方法的步骤如下:

[0007] 包括基础模型构建阶段、体型姿态优化阶段、以及表面细节优化阶段:

[0008] 所述的基础模型构建阶段,用于获得待重构原始图像的基础三维人体模型,具体如下:

[0009] 步骤一:使用SMPL模型构建方法,输入单张待重构原始图像,得到待重构原始图像的基础三维人体模型,所述基础三维人体模型包括人体基础关节点、人体基础锚点和所有点;所述的待重构图像为受遮挡的RGB图像;

[0010] 所述的体型姿态优化阶段,用于优化基础三维人体模型,具体如下:

[0011] 步骤二:使用两个编码器网络分别获取关节点位移向量和锚点位移向量,所述的两个编码器网络即Joint编码器网络和Anchor编码器网络,

[0012] 其中,Joint编码器网络用于提取原始图像中的人体姿态特征,即输出关节点位移向量,其输入为原始图像和人体轮廓信息,具体的,人体轮廓信息由上一步得到的三维人体模型投影所得,使用人体轮廓信息对原始图像进行剪裁,将原始图像中人体轮廓之外的颜色信息剔除,得到输入数据;

[0013] Anchor编码器网络用于提取原始图像中的人体体型特征,即输出为锚点位移向量,其输入与Joint编码器网络的输入相同;

[0014] 步骤三:利用关节点位移向量移动基础关节点,得到优化后的关节点;同理,利用锚点位移向量移动基础锚点,得到优化后的锚点,完成体型姿态优化;

[0015] 表面细节优化阶段,用于得到包含丰富表面细节的三维人体模型,具体如下:

[0016] 步骤四:构建并训练DHDNet网络,DHDNet网络用于从包含遮挡物的RGB图像生成完整的人体深度信息;

[0017] 步骤五:使用训练好的DHDNet,输入待重构原始图像,输出完整的人体深度信息;

[0018] 步骤六:利用步骤五得到的人体深度信息,对优化后的基础三维人体模型的所有点进行形变,最终得到包含丰富表面细节的三维人体模型。

[0019] 有益效果

[0020] 与现有技术相比,本发明通过使用DHDNet建立了受遮挡的RGB图像到完整深度信息的映射关系。DHDNet可以较好的预测出受遮挡区域的深度信息,融入到完整的端到端系统后,可以得到完整且包含丰富细节的三维人体模型。现有方法在输入图像包含遮挡物时,一些方法能做到保证人体模型的完整性,但是往往表面缺乏细节;另一些方法虽然处理理想化输入图像时可以获得包含丰富细节的三维人体模型,但是在有遮挡物的前提下重建结果会出现严重的截断或者凹陷。本方法在此非理想条件下的重建结果是有显著性提升的。

#### 附图说明:

[0021] 图1受遮挡RGB图像到三维人体模型的端到端系统框架图

[0022] 图2DHDNet的生成器部分网络结构图

[0023] 图3DepthHuman数据集

[0024] 图4与现有方法的主观比对结果

[0025] 图5与现有方法的量化比对结果

[0026] 图6三阶段操作点定义,其中,从左至右依次为:“关节点”,“锚点”以及“所有点”

### 具体实施方式:

[0027] 为了能够更清楚地描述本发明的技术内容,下面结合具体实例来进一步的描述:

[0028] 本发明的端到端系统框架图如图1所示,可以看到具体实施过程可以分为三个阶段,基础模型构建阶段,体型姿态优化阶段以及最终的表面细节优化阶段。

[0029] 一、基础模型构建阶段

[0030] 对应前文步骤一,使用HMR提供的预训练模型,输入原始图像可以得到体型、姿态大致准确,表面缺乏细节的标准模型。

[0031] 步骤一的具体操作:HMR是构建SMPL模型的一种现有具体方法,使用HMR的预训练模型构建基础三维人体模型。HMR工作原理如下:首先,输入的RGB图像I经过编码器网络和三维回归模块生成构建SMPL模型所需要的体型参数 $\beta$ 、姿态参数 $\theta$ 以及相机参数 $s, R, T$ 。得到SMPL模型后,三维网格会被送入判别器网络与真实三维模型进行比对,由判别器输出True/False。训练网络模型的损失函数如下:
$$L = \lambda(L_{\text{reproj}} + \mathbb{1}L_{3D}) + L_{\text{adv}} \quad (1)$$

[0032] 其中, $L_{\text{reproj}} = \sum_i \|v_i(\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1$ ,  $\mathbf{x}_i$ 表示第i个二维关键点位置, $v_i$ 表示其是否可见。通过三维标准可以使用 $L_{3D}$ 直接监督编码器网络的训练, $L_{3D}$ 的定义如下:

$$L_{3D} = L_{3D \text{ joints}} + L_{3D \text{ smpl}} \quad (2)$$

$$L_{\text{joints}} = \|(\mathbf{X}_i - \hat{\mathbf{X}}_i)\|_2^2$$

$$L_{\text{smpl}} = \|[\beta_i, \theta_i] - [\hat{\beta}_i, \hat{\theta}_i]\|_2^2.$$

[0034] 编码器网络的对抗损失函数如下:

$$\min L_{\text{adv}}(E) = \sum_i \mathbb{E}_{\Theta \sim p_E} [(D_i(E(I)) - 1)^2] \quad (3)$$

[0036] 每个判别器的目标函数如下:

$$\min L(D_i) = \mathbb{E}_{\Theta \sim p_{\text{data}}} [(D_i(\Theta) - 1)^2] + \mathbb{E}_{\Theta \sim p_E} [D_i(E(I))^2]. \quad (4)$$

[0038] 本发明分三阶段选取三维人体模型上的点作为操作点对人体模型进行形变,操作点定义如图6所示,“关节点”为人体模型关节周围的点集合,“锚点”为均匀分布在人体模型上的若干点,“所有点”为对人体模型进行四倍划分后的所有点。该操作点定义参考HMD (Hao Zhu, Xinxin Zuo, Sen Wang, Xun Cao, and Ruigang Yang, “Detailed Human Shape Estimation from a Single Image by Hierarchical Mesh Deformation,” CVPR, 2019.) 中的方法。

[0039] 所述的体型姿态优化阶段,用于优化基础三维人体模型,

[0040] 步骤二的具体操作:为了得到“关节点”和“锚点”级别的形变向量,本发明构建了Joint和Anchor两个编码器神经网络。Joint和Anchor两个编码器神经网络采用简单的编码器网络即可,两个网络的输入相同,其输入为原始图像和人体轮廓信息,具体的,人体轮廓信息由上一步得到的三维人体模型投影所得,使用人体轮廓信息对原始图像进行剪裁,将原始图像中人体轮廓之外的颜色信息剔除,得到输入数据;Joint编码器网络输出关节点位

移向量,Anchor编码器网络输出锚点位移向量。

[0041] 步骤三的具体操作:利用关节点位移向量移动基础关节点,得到优化后的关节点;同理,利用锚点位移向量移动基础锚点,得到优化后的锚点,完成体型姿态优化;在使用关节点位移向量和锚点位移向量形变三维人体模型时,本实施例中采用了拉普拉斯形变(Laplace Deformation)的方法。

[0042] 编码器神经网络的损失函数如下:

$$[0043] \quad L_{joint/anchor} = \| p - \hat{p} \|_2 \quad (5)$$

[0044] 其中的 $p$ 表示由编码器神经网络输出的图像空间中的关键点位移向量, $\hat{p}$ 表示真实位移向量,真实位移向量由基础三维人体模型投影到图像空间得到的关键点位置与数据集中相对应的真实关键点位置计算得到,其中关键点包括关节点和锚点。

[0045] 表面细节优化阶段

[0046] 步骤四的具体操作:通过搜集网络上的公开数据集,包括LSP数据集(Leeds Sports Pose)、LSPET数据集(LSP的补充数据集)、COCO数据集(Common Objects in Context),本发明构建了DepthHuman数据集。其中包含超过18000张标记了2d关节点位置的人体图像。同时,本发明采用现有的深度估计算法,合成了每张RGB图像对应的深度信息,作为真实深度信息的替代;以及包括原始图像对应的人体深度信息,DepthHuman数据集的详细信息如图3所示。

[0047] DHDNet采用对抗神经网络结构,由生成器与判别器两部分组成,其中生成器部分网络结构如图2所示,其中生成器部分网络结构以经典的U-Net网络结构为基础,输入为RGB三通道原始图像加一维的遮罩层,输出为一维的深度信息,值得一提的是本发明引入了Shift-Net的方法,在神经网络的第四层引入了Shift-Layer,从而更好的得到完整深度信息。训练DHDNet使用的损失函数如下:

$$[0048] \quad L_{final} = L_{depth} + \lambda_{gan} L_{gan} + \lambda_{photo} L_{photo} + \lambda_c L_c + \lambda_s L_s \quad (6)$$

[0049] 其中, $\lambda_{gan}$ , $\lambda_{photo}$ , $\lambda_c$ , $\lambda_s$ 为对应损失函数的权重。 $L_{depth}$ 为MSE损失函数。 $L_{gan}$ 表达式如下:

$$[0050] \quad L_{gan} = \min_G \max_D \mathbb{E}_{x \in X} [\log(D(x))] + \mathbb{E}_{z \in Z} [\log(1 - D(G(z)))] \quad (7)$$

[0051] 其中G和D表示GAN中的生成器和判别器,X表示Ground Truth图像集,即训练数据集中原始图像对应的人体深度信息做为真实人体深度图,Z表示输入图像集,即训练数据集中原始图像,输入图像为RGB三通道原始图像加一维的遮罩层。

[0052]  $L_{photo}$ 表示光度损失函数,在朗伯表面的假设前提下,DHDNet使用球谐函数作为全局光照的表现方法,其中 $\rho$ 在(Sean Bell,Kavita Bala,and Noah Snavely,“Intrinsic images in the wild”ACM Transactions on Graphics vol.33,no.4,pp.159,2014.)中有详细计算过程。 $H_k$ 表示球谐函数的基, $l_k$ 表示球谐函数系数。

$$[0053] \quad L_{photo} = \left\| \rho \sum_{k=1}^9 l_k H_k(n) - I \right\|_2 \quad (8)$$

[0054] 为了恢复高频细节,DHDNet在输出的深度信息以及真实深度信息上使用了VGG特

征提取器网络,并使用 $L_c$ 和 $L_s$ 两个损失函数进行监督。其中, $L_c$ 表示内容损失函数:

$$[0055] \quad L_c(p, x, l) = \frac{1}{2} \sum_{i,j} (F_{i,j}^l - P_{i,j}^l)^2 \quad (9)$$

[0056] 其中, $p$ 和 $x$ 分别代表由DHDNet输出的生成图像和depthhuman数据集中的真实人体深度图像, $P^l$ 表示DHDNet输出的生成图像经过VGG特征提取器网络第1层的特征信息, $F^l$ 表示真实人体深度图像经过VGG特征提取器网络第1层的特征信息, $i$ 和 $j$ 代表像素坐标。

[0057]  $L_s$ 表示风格损失函数:

$$[0058] \quad L_s(a, x) = \sum_{l=0}^L w_l E_l \quad (10)$$

[0059] 其中, $a$ 表示由DHDNet输出的生成图像, $x$ 表示depthhuman数据集中的真实人体深度图像, $w_l$ 表示VGG网络各层贡献度的权重, $E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (\text{Gram}(F_{i,j}^l) - \text{Gram}(P_{i,j}^l))^2$ 。 $\text{Gram}$ 表示Gram矩阵,使用Gram矩阵计算表达图像特征之间的相关性。 $i$ 和 $j$ 代表像素坐标, $N_l$ 表示VGG网络第1层特征的通道宽度, $M_l$ 表示VGG网络第1层特征的通道高度。

[0060] 步骤五、使用步骤二中优化得到的三维人体模型投影出人体轮廓信息,利用得到的人体轮廓信息对待重构原始图像进行剪裁,将裁剪后的待重构原始图像中人体轮廓之外的颜色信息剔除,之后将待重构原始图像输入训练好的DHDNet,得到完整的人体深度信息。

[0061] 步骤六、此步骤为常规操作,具体为:对现阶段SMPL模型进行四倍划分,得到四倍顶点数量的精细模型。深度信息可以简单转换为对模型上三维空间点的位移操作。在经过这个阶段的形变之后,可以得到包含丰富衣服褶皱,发型纹理等表面细节。

[0062] 重建质量评价:

[0063] 现在分别对SYN,RECON数据集采用本发明提供的方法与当前主流的几种算法进行比较评价。这些算法分别是HMR (Angjoo Kanazawa, Michael J. Black, David W. Jacobs, and Jitendra Malik, "End-to-End Recovery of Human Shape and Pose," CVPR, pp.7122-7131, 2018.), HMD (Hao Zhu, Xinxin Zuo, Sen Wang, Xun Cao, and Ruigang Yang, "Detailed Human Shape Estimation from a Single Image by Hierarchical Mesh Deformation," CVPR, 2019.), SMPLify (Federica Bogo, Angjoo Kanazawa, Christoph Lassner, Peter Gehler, Javier Romero, and Michael J. Black, "Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image," in European Conference on Computer Vision, 2016, vol.9909LNCS, pp.561-578.), BodyNet (Varol, Gul, et al. "BodyNet: Volumetric inference of 3d human body shapes." Proceedings of the European Conference on Computer Vision (ECCV). 2018.)。

[0064] 本发明的算法与现有方法的量化对比结果如图5所示。重建模型点坐标与真实模型点坐标间的误差是衡量三维重建结果的客观标准,这个值越小说明重建质量越好:

$$[0065] \quad D_{3d\_err} = \frac{1}{N} \sum_{i=1}^N \|v_i - \hat{v}_i\|_2 \quad (11)$$

[0066] 其中, $N$ 表示SMPL模型的顶点数, $v_i$ 表示重建模型上的顶点, $\hat{v}_i$ 表示真实人体模型上的顶点。 $D_{3d\_err\_visi}$ 表示相机可见区域的三维空间点误差,IoU (Intersection over



Union) 表示重建模型与真实模型投影出的轮廓信息吻合度, 这个指标越大说明重建质量越高。

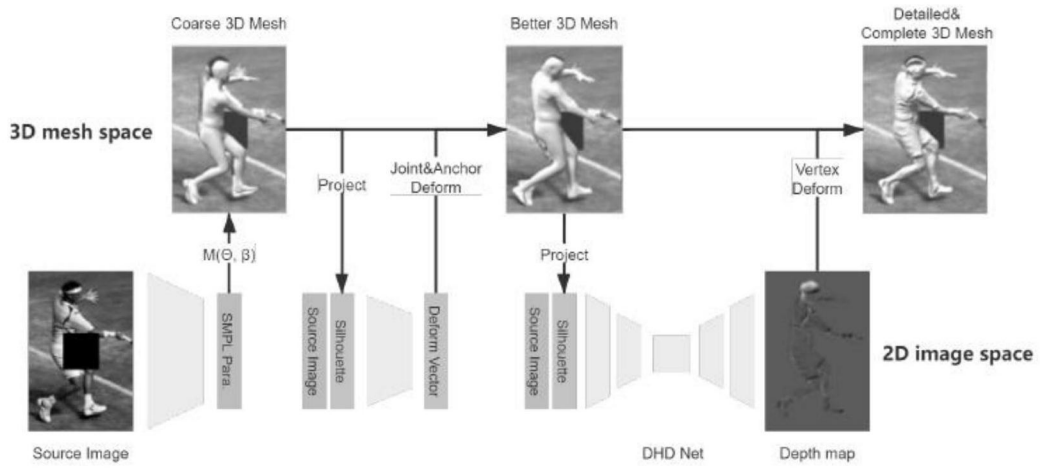


图1

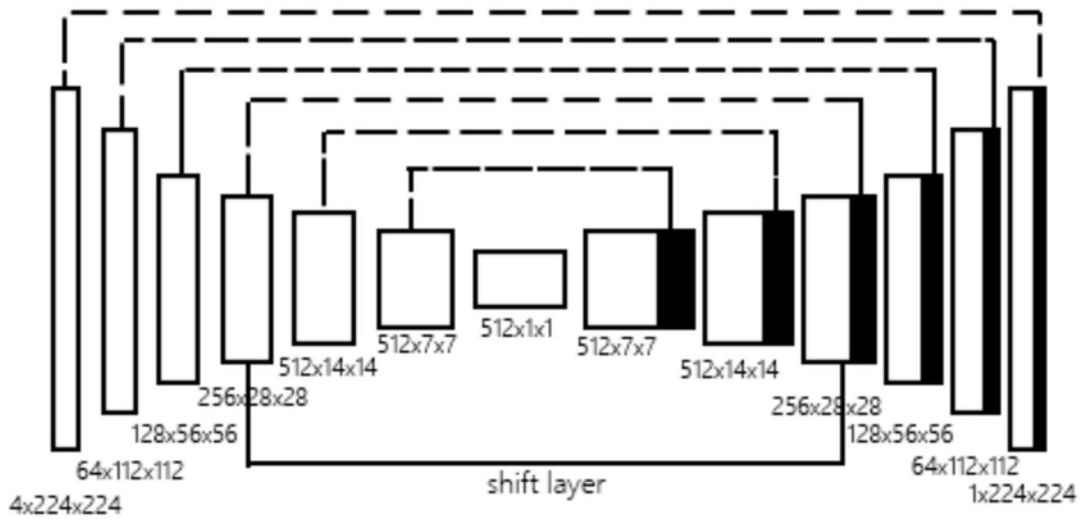


图2

Data Source	LSP	LSPET	MPII	COCO	TOTAL
training num	987	5376	8035	4004	18402
testing num	703	0	1996	606	3305

图3

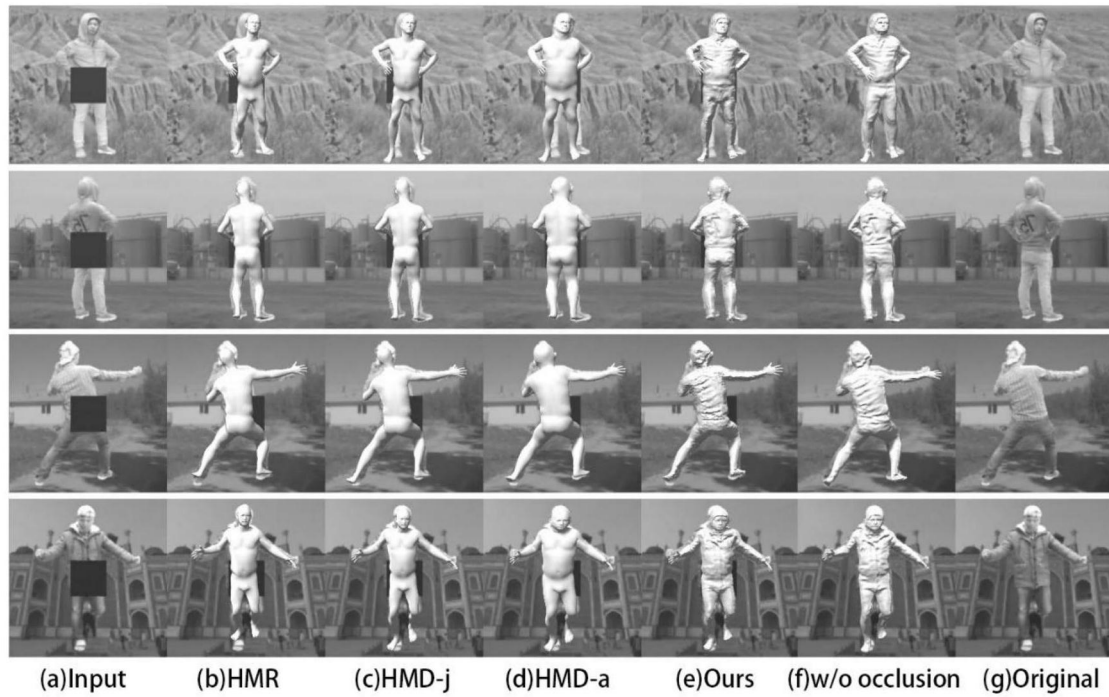


图4

方法名称	SYN 数据集			RECON 数据集		
	3d_err	3d_err_visi	iou	3d_err	3d_err_visi	iou
SMPLify	77.310	75.670	66.200	61.840	60.690	69.900
BodyNet	69.410	61.550	65.200	52.750	51.050	68.500
HMR	67.993	62.467	66.901	60.047	51.501	70.358
HMD-j	63.876	59.591	69.977	58.918	51.047	73.231
HMD-a	61.224	58.373	<b>73.799</b>	53.215	49.814	<b>79.114</b>
本文方法	<b>58.503</b>	<b>54.317</b>	-	<b>50.592</b>	50.617	-

图5



图6