

# 基于多特征点匹配的三维人体姿态重建

张亚凤, 刘翠响\*, 马杰, 苏亚婷

河北工业大学电子信息工程学院, 天津 300401

**摘要** 由单视角图像重建三维人体姿态时, 深度信息的缺失与姿态的多样性使二维姿态映射到三维姿态时易出现朝向错误、细节姿态处理不佳等问题。为此, 结合立体人体模型的骨骼及顶点分布规律, 以优化模型变形策略为核心, 提出了一种基于多特征点匹配的三维姿态重建方法。所提方法的核心是以优化能量函数的方式将多个二维人体特征点与人体模型三维特征点进行匹配拟合, 从而实现三维姿态的重建, 并通过部分关节点建立朝向约束项来降低深度信息缺失对重建姿态的影响, 同时利用多个头部特征点调整头部姿态降低姿态多样性对头部姿态重建的影响。在公开姿态数据集 MPI-INF-3DHP 与 LSP 上的实验结果表明, 所提方法有效解决了姿态模糊及细节姿态处理不佳的问题, 能够准确重建常见动作下的三维人体姿态。

**关键词** 机器视觉; 图像处理; 三维姿态估计; 人体模型; 特征点匹配拟合; 迭代优化

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP202259.1615003

## Three-dimensional Human Pose Reconstruction Based on Multifeature Point Matching

Zhang Yafeng, Liu Cuixiang\*, Ma Jie, Su Yating

School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China

**Abstract** When reconstructing three-dimensional human posture from a single view image, the lack of depth information and postural diversity cause orientation errors and the poor processing of the postural details when mapping two-dimensional to three-dimensional posture. Therefore, combined with the distribution law of bones, vertices of the stereo human model, and optimization model deformation strategy, we propose a three-dimensional pose reconstruction method using multifeature point matching. The core of the proposed method is to match and fit multiple two-dimensional human feature points with the three-dimensional feature points of the human model by optimizing the energy function to realize the reconstruction of three-dimensional posture. Furthermore, we establish orientation constraints using some joint points to reduce the impact of the lack of depth information on the reconstructed posture. Additionally, multiple head feature points are used to adjust the head pose to reduce the impact of pose diversity on the head pose reconstruction. The experimental results on the public pose datasets, MPI-INF-3DHP and LSP, show that the proposed method effectively solves the problems of pose blur and poorly detailed pose processing, as well as accurately reconstructs the three-dimensional human pose under common actions.

**Key words** machine vision; image processing; three-dimensional pose estimation; human model; feature point matching; iterative optimization

## 1 引言

三维人体姿态重建作为新一代人机交互的关键技术已经成为计算机视觉领域中的研究热点之一。传统的人体姿态重建为多人或者单人二维骨架的重建<sup>[1-2]</sup>, 多用于步态识别、行为识别等领域, 在虚拟现实、增强现实等领域应用较少。现阶段三维人体姿态重建主要

分为两类: 第 1 类是通过估计人体关节的三维坐标从而重建人体三维骨架<sup>[3-4]</sup>, 多用于智能监控领域; 第 2 类是在重建骨架姿态的同时恢复人体表面信息, 多用于虚拟试衣、电影动画制作中。此外, 三维人脸的重建也成为三维重建领域中的热点<sup>[5]</sup>。考虑到人体表面也包含丰富的信息, 本实验组重点研究第 2 类重建方法。目前, 这类方法一般结合人体模型或借助多个相机、扫

收稿日期: 2021-05-17; 修回日期: 2021-06-16; 录用日期: 2021-07-13

基金项目: 河北省自然科学基金(F2020202045)

通信作者: \*liucuix@126.com

描仪等设备实现。Rhodin 等<sup>[6]</sup>用多相机拍摄人体运动视频实现重建; von Marcard 等<sup>[7]</sup>用惯性测量单元采集人体运动信息, 结合多人线性蒙皮模型(SMPL)模型<sup>[8]</sup>实现重建; Mao 等<sup>[9]</sup>使用 RGB-D 传感器, 通过人体旋转获取多视角图像进行重建。虽然这些利用多相机或专业设备的方法能够提取较多信息得到更好的重建效果, 但使用成本高, 预处理工作多, 且使用范围有限。而基于普通二维图像的重建方法仍存在许多待解决的问题, 如 Bogo 等<sup>[10]</sup>提出的 SMPLify 算法虽能从二维图像中重建三维人体, 但难以避免姿态模糊问题, 且未能实现头部姿态的调整; 曾志超等<sup>[11]</sup>使用视频帧间连贯性和面部特征点拟合的方式来解决姿态模糊并调整头部姿态, 但前一帧重建出错时会影响后续重建结果, 造成误差累积, 且仅提取面部特征点的头部姿态重建方式易受人体姿态的影响。在端到端的重建方法中, Omran 等<sup>[12]</sup>利用语义分割将人体分割为 12 部分重建; Kanazawa 等<sup>[13]</sup>利用现有的神经网络通过迭代误差反馈回归的方式实现重建; Kolotouros 等<sup>[14]</sup>结合图卷积

网络模型进行姿态重建, 虽然这些方法在重建效率上有优势, 但由于用于训练的三维人体数据集较少, 这类方法的泛化能力和鲁棒性有待提高。

本文通过剖析现有三维姿态重建算法存在的不足, 提出一种面向普通单视角二维图像的人体姿态重建算法, 即多特征点匹配(MFPM)算法。所提算法结合主流的三维人体模型, 利用提取的多个二维特征点位置信息构建包含多个约束项的能量函数, 改进并优化模型变形策略, 实现了准确的三维人体姿态重建。

## 2 数据预处理

图 1 为所提算法的整体框架, 主要包括数据预处理与姿态重建两部分。数据预处理包括二维图像预处理和标准 SMPL 模型的预处理; 姿态重建主要分为三个阶段, 即朝向优化、整体姿态优化和头部细节姿态优化; 最后通过更新标准 SMPL 模型的参数, 输出人体三维姿态。

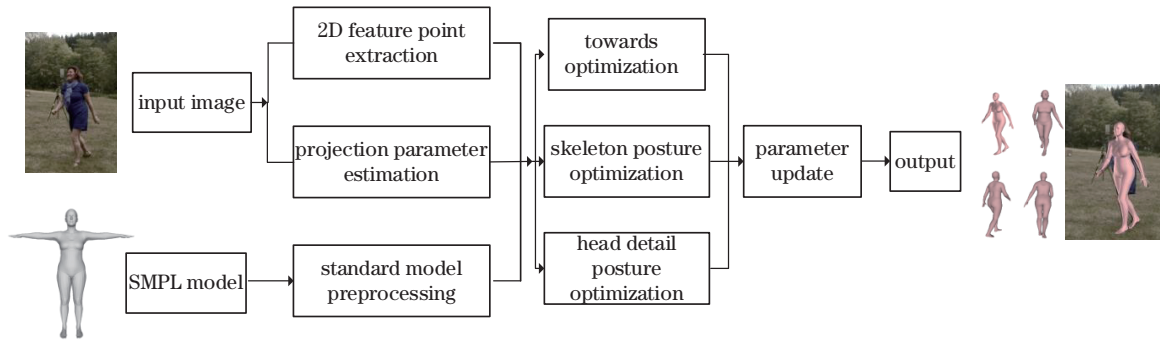


图 1 总体框架图

Fig. 1 Overall framework

### 2.1 SMPL 模型及预处理

三维人体模型通过对人体骨架长度、关节活动范围进行约束, 将重建后的人体姿态约束在合理范围之内。本实验组使用的 SMPL 模型<sup>[8]</sup>是在三维人体扫描数据库 CAESAR 中学习得到的参数化人体模型, 该模型分为女性模型和男性模型, 如 [图 2(a)、(b)] 所示。SMPL 模型由 6890 个顶点构成, 定义了 24 个关节点,

并以骨盆节点为根节点, 根据运动学树定义了各关节之间的父子关系, 如 [图 2(c)] 所示。且该模型通过姿态参数向量  $\theta$  和体形参数向量  $\beta$  控制姿态和体形的变化, 其中  $\theta$  是以轴角表示的关节点之间相对旋转的向量, 维度为  $1 \times 75$ ;  $\beta$  为控制人体高矮胖瘦等体形变化的向量, 维度为  $1 \times 10$ 。

从 [图 2(c)] 可以看出, SMPL 模型仅在颈部附近

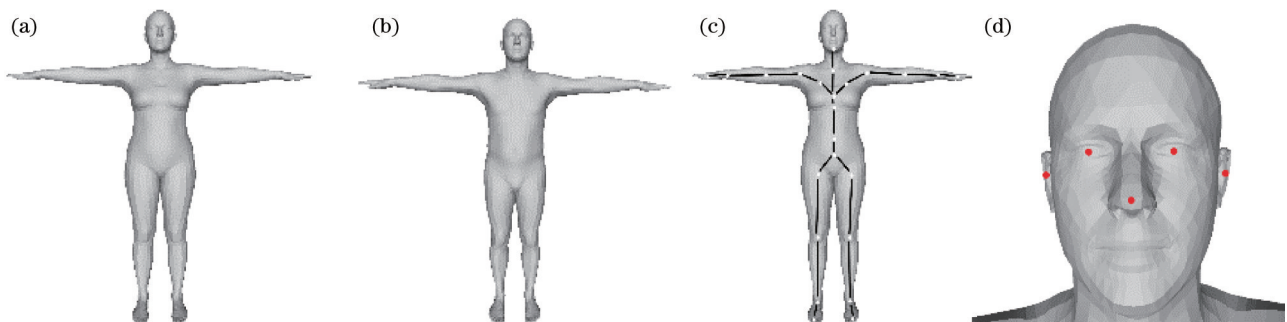


图 2 SMPL 模型。(a) 女性模型; (b) 男性模型; (c) 骨架结构; (d) 头部固定点

Fig. 2 SMPL model. (a) Female model; (b) male model; (c) skeleton structure; (d) head fixed point

定义一个关节点控制头部姿态,这种处理头部姿态的方式易受人体姿态、图片像素、光照等因素的影响,从而导致头部姿态处理失败。文献[11]虽然对标准的 SMPL 模型进行了改进,以在脸部设置固定点的方式调整头部姿态,但脸部特征点分布相对密集,当人体侧对或背对镜头时仍难以重建头部姿态。针对该问题,本实验组对标准的 SMPL 模型进行了处理,通过将标准模型序列化生成三维可视化文件,再借助 MeshLab 三维几何处理系统进行可视化,读取模型在选取的固定特征点处的序号,如[图 2(d)]所示,在模型头部选取多个固定特征点,分别为左右耳、左右眼、鼻子,利用这五个头部特征点优化头部朝向,以此避免单一头部特征点或面部特征点导致的头部姿态重建失败的情况,提升算法对不同头部姿态的适应能力。

## 2.2 二维图像预处理

在二维关键点检测阶段,所使用的算法是 OpenPose,该算法不仅实时性较高,并且具有较高的检测精度<sup>[15]</sup>,同时 OpenPose 能够将身体关节点和头部关键点的检测与定位统一在同一框架下,降低了先检测身体关节点再检测头部关键点方式的冗余度。从二维图像中提取的 Body-25 人体二维骨架模型如图 3 所示,从中可获取到多个关节点的二维位置信息,用于驱动三维模型姿态的变化。

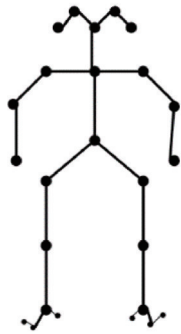


图 3 Body-25 骨架模型

Fig. 3 Body-25 skeleton model

此外,三维姿态的重建离不开二维图像空间与三维空间的相互转换,而传统的利用图像的局部特征如人体部分关节点来计算二者之间的转换参数往往忽略了二维图像整体与三维人体模型空间的映射关系,通常需要假设人体是正对相机平面站立的,如文献[10]和[11],当假设失败时,易导致特征点匹配错误使姿态误差增大。本实验组结合文献[13]中端到端的人体网格恢复(HMR)模型,将二维图像的所有像素信息考虑在内,估计由三维人体模型空间映射到二维图像平面的投影参数,包括缩放尺度  $s$ 、平移向量  $t$  和旋转向量  $R$ ,并由此建立二维图像空间与三维人体模型空间的映射关系。

## 3 三维姿态重建

### 3.1 MFPM 算法能量函数

以从二维图像中提取的多个特征点位置信息为数据基础,根据 SMPL 模型的关节及顶点分布规律,构建了总能量函数来优化 SMPL 模型的变形策略,通过最小化该能量函数实现三维姿态重建。所提算法的能量函数表达式为

$$E(\boldsymbol{\beta}, \boldsymbol{\theta}) = E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}}) + E_B(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}}) + E_H(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{H}_{\text{est}}) + \lambda E_p(\boldsymbol{\beta}, \boldsymbol{\theta}), \quad (1)$$

式中:  $E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$  为朝向约束项;  $E_B(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$  为人体关节位置约束项;  $E_H(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{H}_{\text{est}})$  为头部姿态约束项;  $E_p(\boldsymbol{\beta}, \boldsymbol{\theta})$  为关节旋转与体形约束项。

$E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$  为朝向约束项,由于普通二维图像中不包含深度信息,由二维空间推测三维空间存在模糊问题。尤其是当拍摄的目标人体侧身对向镜头时,左右肩和左右髋部距离较近,重建后的姿态易出现错误朝向。本实验组通过设定两肩和两髋距离的阈值来解决这一问题,当距离小于设定阈值时,将模型朝向翻转,分别计算翻转前后  $E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$  的值,取令该值最小的朝向作为重建人体的朝向。对人体模型手腕、脚腕等末端关节点进行投影变换,调整模型朝向使模型在手腕与脚腕处的投影点与对应二维检测点之间尽可能接近,即

$$E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}}) = \sum_{m \in J(k)} \left\| \Pi_C [JM(\boldsymbol{\beta}, \boldsymbol{\theta})]_m - \mathbf{B}_{\text{est},k} \right\|, \quad (2)$$

式中:  $\mathbf{B}_{\text{est},k}$  为人体手腕与脚腕处的二维位置信息;  $J$  为由三维模型  $M(\boldsymbol{\beta}, \boldsymbol{\theta})$  回归三维关节点的回归矩阵;  $\Pi$  为正交投影变换;  $C$  为投影变换参数。

$E_B(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$  为人体关节位置约束项,与文献[10]相同。该约束项通过约束三维模型关节点的投影点与二维人体之间的骨架姿态距离,驱动模型骨架长度和姿态的变化,确保重建后的人体模型与真实人体具有相同姿态。当三维关节点的投影点与二维关节点距离越近时,该项的值越小;反之越大,即

$$E_B(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}}) = \sum_i \omega_i \rho \left\| \Pi_C [JM(\boldsymbol{\beta}, \boldsymbol{\theta})]_i - \mathbf{B}_{\text{est},i} \right\|, \quad (3)$$

$$\rho(e) = \frac{e^2}{\sigma^2 + e^2}, \quad (4)$$

式中:  $\mathbf{B}_{\text{est},i}$ 、 $\omega_i$  分别为第  $i$  个关节点的位置坐标与置信度;  $\rho(\cdot)$  为杰曼麦克卢尔函数,是用来处理噪点的,其中  $\sigma = 100$ 。

$E_H(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{H}_{\text{est}})$  为头部姿态约束项,该式以模型头部特征点的投影点与二维头部检测点之间的距离作为该项的值,该项仍以重投影误差的方式实现,通过调整模型的姿态向量减小模型的头部固定点的投影点与二维头部特征点之间的距离误差,调整模型头部姿

态,即

$$E_H(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{H}_{\text{est}}) = \sum_j \gamma_j \rho \left\| \Pi_C[\mathbf{M}(\boldsymbol{\beta}, \boldsymbol{\theta})_j] - \mathbf{H}_{\text{est},j} \right\|, \quad (5)$$

式中:  $\mathbf{H}_{\text{est},j}$ 、 $\gamma_j$  分别为第  $j$  个二维头部特征点的位置坐标与误差权重。

为避免重建后的人体姿态违背运动学规律,本实验组增加了文献[10]中的关节旋转与体形约束项  $E_p(\boldsymbol{\beta}, \boldsymbol{\theta})$ 。

$$E_p(\boldsymbol{\beta}, \boldsymbol{\theta}) = \lambda_a E'_a(\boldsymbol{\theta}) + \lambda_\theta E_\theta(\boldsymbol{\theta}) + \lambda_{\text{sp}} E_{\text{sp}}(\boldsymbol{\theta}; \boldsymbol{\beta}) + \lambda_\beta E_\beta(\boldsymbol{\beta}), \quad (6)$$

式中:  $E'_a(\boldsymbol{\theta})$  对关节旋转角度进行约束,避免出现极端姿态;  $E_\theta$  为姿态先验项,进一步排除不符合运动学规律的姿态;  $E_{\text{sp}}$  为肢体互穿误差项,避免人体出现四肢穿过身体的情况;  $E_\beta$  避免人体出现异常体形;  $\lambda_a$ 、 $\lambda_\theta$ 、 $\lambda_{\text{sp}}$ 、 $\lambda_\beta$  为对应项的权重。

### 3.2 能量函数优化求解

本实验组利用 OpenDR<sup>[16]</sup> 计算三维特征点的投影,结合 Chumpy 使用狗腿法优化能量函数,优化过程如下:

#### 1) 初始位姿求解。

由二维图像回归得到投影变换参数  $\mathbf{C}$  后,对预处理后的模型进行投影变换,通过最小化式  $\sum_n \left\| \Pi_C[\mathbf{JM}(\boldsymbol{\beta}, \boldsymbol{\theta})]_n - \mathbf{B}_{\text{est},n} \right\|$  使模型 4 个躯干关节点投影后的位置与对应二维关节尽可能接近,得到模板模型初始位姿。

#### 2) 模型朝向优化。

为避免错误的朝向对后续关节拟合产生影响,在姿态与体形优化前对人体朝向进行估计。通过设定阈值,计算左右两肩、左右髋部之间的距离,当这些距离均小于设定阈值时,将式(1)估计的朝向翻转 180°,计算翻转前后的朝向误差值  $E_o(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}})$ ,取使该误差项最小的人体朝向作为人体模型朝向,在此基础上进行后续姿态优化。

#### 3) 三维骨架姿态优化。

在确定模型初始位姿及朝向后,通过调整姿态参数与形状参数向量,使  $E_B(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{B}_{\text{est}}) + \lambda E_p(\boldsymbol{\beta}, \boldsymbol{\theta})$  最小,实现多个二维关节点对三维模型关节点的驱动,完成整体姿态与体形重建,实现二维特征点对三维模型特征点的驱动,使模型与真实人体姿态一致。为避免陷入局部极小值,采用与文献[10]相同的优化方式,在优化过程中逐渐减小  $\lambda$  的值。

#### 4) 头部姿态优化。

经过上述优化后,模型整体姿态已经与目标人体相似,此时,通过调整模型参数向量,最小化  $E_H(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{C}, \mathbf{H}_{\text{est}})$  的值,继续优化头部姿态,实现对头部细节姿态的调整。

## 4 实验结果与分析

将所提算法与 SMPLify 算法<sup>[10]</sup>、HMR<sup>[13]</sup> 和 CMR<sup>[14]</sup> 算法进行了对比,在 LSP 数据集上进行定性比较,在 MPI-INF-3DHP 数据库<sup>[17]</sup> 的测试集上进行定量比较,其中 SMPLify 为基于优化的方法,HMR 和 CMR 均为基于回归的端到端的方法。本实验基于 Ubuntu18.04 操作系统,必要的依赖包及版本号为 OpenDR 0.78、Chumpy 0.69、opencv-python 4.2.0.32。

### 4.1 定性分析

二维姿态数据集 LSP 包含丰富的运动姿态,如打棒球、网球、排球等动作。为验证所提算法的普适性,从中选取不同性别和姿态的运动图进行实验,与其他算法进行定性比较,结果如图 4 所示,图中对出现误差较大的位置进行了标记。

从图 4 可以看出,在前三列动作中,其他算法在头部姿态、关节定位上出现了不同程度的姿态误差,如第 1 列中,原图中人体头部朝向身后左下方,SMPLify 重建后的人体头部朝向身体前方,且未调整肩部姿态;HMR 中头部姿态虽进行了细微调整,但仍未朝向身后左下方,且右肩高度比实际肩膀高度低;CMR 不仅未重建头部姿态,且右臂未合理弯曲。

在图 4 第 4~7 列人体未正对镜头的动作中,其他算法在头部姿态与四肢关节的定位上仍存在误差,如第 5 列中,SMPLify 算法出现了严重的姿态翻转,HMR 在头部与左脚出现误差,CMR 同样仍未能重建准确的头部姿态,并且左右腿出现了姿态翻转。而所提算法无论在人体正对、侧对、背对镜头时,均能更准确地定位四肢关节位置,避免朝向错误,并且能够精准地刻画头部姿态,重建后的姿态与真实人体更相符。

### 4.2 定量分析

在 MPI-INF-3DHP 的测试集上与其他算法进行定量对比,从数据层面给出说明,验证所提算法的有效性。

MPI-INF-3DHP 数据库为大型公开三维人体姿态估计数据库,包含多种不同人体姿态,共标记了 17 个关节点的三维位置信息,评价实验结果时从数据集标定的关节中选取与 SMPL 模型共有的 14 个关节点计算重建姿态误差。

为消除缩放、旋转及平移对实验结果的影响,利用普氏对齐将算法重建的人体关节点坐标与数据集真值坐标对齐,以对齐后的关节点与真值之间的欧氏距离作为重建姿态误差。

$$\arg \min_{s_A, \mathbf{R}_A, \mathbf{t}_A} \left\| s_A \mathbf{R}_A \mathbf{J}_{\text{pre}} + \mathbf{t}_A - \mathbf{V}_{\text{GT3D}} \right\|, \quad (7)$$

$$\mathbf{J}_{\text{align}} = s_A \mathbf{R}_A \mathbf{J}_{\text{pre}} + \mathbf{t}_A, \quad (8)$$

$$E = \sum_{i=1}^n \left\| \mathbf{J}_{\text{align},i} - \mathbf{V}_{\text{GT3D},i} \right\|, \quad (9)$$

式中:  $s_A$  为缩放尺度;  $\mathbf{R}_A$  为旋转矩阵;  $\mathbf{t}_A$  为平移向量;



图4 LSP重建结果对比

Fig. 4 Comparison of reconstruction results on LSP dataset

$J_{\text{pre}}$  为重建人体的关节点;  $V_{\text{GT3D}}$  为数据集标定真值;  $J_{\text{align}}$  为对齐后的关节点估计值;  $E$  为重建姿态误差, 其中  $n = 14$ 。

图5为MPI-INF-3DHP数据集上的部分实验结果图。从图中可以看出: 两肩过近时SMPLify重建的姿态出现错误朝向, 且未重建头部姿态; HMR与CMR在人体关节点定位时出现误差, 如HMR在第1行的手腕、脚踝处出现误差, CMR在第1行左臂和第2行右臂过度弯曲, 在第3行动作中左右腿关节点识别错误, 导致腿部姿态翻转, 此外这两种算法都不能准确重建头部姿态。而所提算法不仅能够恢复正确的人体朝向, 且对关节点的定位更准确。这是由于所提算法在优化过程中不仅增加了对人体朝向的判断和头部姿态的约束, 并且通过优化关节重投影误差损失使模型关节点更逼近真实人体关节点位置, 在这一优化过程中SMPL模型的骨架长度也随之调整, 从而保证了关节定位的精度, 提高了姿态的准确率。

为进一步验证所提算法的有效性, 在MPI-INF-3DHP测试集的不同序列(TS1~TS4)上选取多种姿

态进行实验, 为突出姿态重建效果, 取各动作下的最大重建姿态误差进行对比, 结果如表1所示。从表1可以看出, 对常见的人体动作, 与其他算法相比, 所提MFPM算法重建后的姿态与真实人体更加接近, 误差更小。

此外, 人体作为统一的整体, 各部位之间互相关联, 为验证头部姿态约束项对整体姿态误差的影响, 在MPI-INF-3DHP测试集的室外数据集TS5和TS6上选取4种动作序列, 在每个动作序列中选取了20张姿态图进行实验, 计算MFPM算法去掉头部姿态约束项前后的平均姿态误差, 结果如表2所示。从表2可以看出, 增加头部姿态约束项后, 整体的姿态误差均有不同程度的降低。

通过分析表2中受头部姿态约束影响较明显的TS5 Forward动作与TS6 Running动作发现, 这两个动作的头部特征点可见数较多, 头部无严重遮挡, 即式(5)中的 $\gamma$ 略大, 因此头部约束项对整体姿态优化的影响略大。

将受头部姿态影响较明显的姿态图列于图6前4



图 5 重建效果对比

Fig. 5 Comparison of reconstruction effects

表 1 最大重建误差对比

Table 1 Comparison of Maximum reconstruction error

unit: mm

Method	TS1		TS2		TS3		TS4	
	Forward	Rolling	Chatting	Sitting	Walking	Working	Boxing	Batting
SMPLify	1107.94	1171.40	1489.51	916.03	1365.89	1051.75	1439.05	1387.29
HMR	1094.27	844.10	1111.24	1138.04	1141.70	992.76	1233.73	1476.62
CMR	834.59	789.15	825.41	678.21	1038.60	536.68	959.45	982.22
MFPM	830.74	782.48	720.13	536.23	916.30	528.07	825.05	1114.93

表 2 头部姿态约束对重建误差的影响

Table 2 Influence of head posture constraint on reconstruction error

unit: mm

Head constraint	TS5		TS6	
	Forward	Crouch	Turn right	Running
×	468.65	461.02	564.32	577.95
✓	390.60	411.59	543.02	457.87

列中,其中前 2 列来自 TS5 Forward 动作,第 3、4 列来自 TS6 Running 动作。从图中可以看出,增加头部姿态约束不仅能够对头部细节进行调整,也能够对肩膀、人体朝向等进一步约束。同时也列举了受头部姿态约束影响较小的 Crouch 和 Trun right 动作的实验结果

图,如图 6 后 3 列所示。从图中可以看出,在这 3 列中,头部特征点可见数略少,在没有头部姿态约束时,所提算法已经比较准确地估计了整体姿态,增加头部约束后,只对头部细节姿态进行了微调,对整体姿态的影响不明显。

实验结果表明,头部姿态约束能够与身体关节相辅相成,进一步纠正估计错误的朝向,且对上半身姿态求精有一定的贡献,在一定程度上提高了重建姿态的准确率。

在 MPI-INF-3DHP 测试集的 6 个序列上进行了实验,这 6 个序列为不同性别人体在不同场景下的运动姿态,每张图片中包含一人,不同算法在各序列上的平均重建误差如表 3 所示。



图 6 头部姿态约束对整体姿态的影响

Fig. 6 Influence of head constraint on overall posture

表 3 不同算法在 MPI-INF-3DHP 数据库上的重建误差

Table 3 Errors of different algorithms on MPI-INF-3DHP

Method	database					
	TS1	TS2	TS3	TS4	TS5	TS6
SMPLify	854.45	1052.70	1104.81	1007.72	867.67	968.41
HMR	825.47	935.51	1018.16	1045.06	866.71	1196.18
CMR	607.35	735.41	776.69	797.47	699.94	868.28
MFPM	600.93	657.96	671.17	683.86	618.97	756.72

从表 3 可以看出:与现有的其他算法相比,所提算法的姿态误差最小,说明所提算法能够更准确地重建人体姿态;且对 6 个序列不同场景和性别的人体姿态重建时,所提 MFPM 算法的表现较稳定。

## 5 结 论

结合 SMPL 模型,提出了一种基于多特征点匹配的三维人体姿态重建算法。所提算法利用从图像中提取的多个二维特征点与 SMPL 模型三维特征点进行匹配拟合来驱动模型姿态的变化,使用部分肢体关节建立朝向约束项来优化初始的模型朝向,有效避免了深度信息缺失造成的姿态歧义;又提取多个头部特征点构建头部姿态约束项来重建头部细节姿态,有效解决了单一头部特征点或仅面部特征点在复杂姿态下对头部细节姿态处理不佳的问题,且由多个头部特征点建立的头部姿态约束项能够进一步优化整体姿态,提高了算法的稳定性。在公开数据集 LSP 与 MPI-INF-3DHP 上的实验结果表明,所提算法在多数常见动作下可较准确地重建人体的三维姿态,并且能够更好地刻画人体细节姿态。

## 参 考 文 献

- [1] 闫芬婷,王鹏,吕志刚,等.基于视频的实时多人姿态估计方法[J].激光与光电子学进展,2020,57(2):021006.  
Yan F T, Wang P, Lü Z G, et al. Real-time multi-person video-based pose estimation[J]. Laser & Optoelectronics Progress, 2020, 57(2): 021006.
- [2] 张显坤,张荣芬,刘宇红.基于二次生成对抗的人体姿态估计[J].激光与光电子学进展,2020,57(20):201509.  
Zhang X K, Zhang R F, Liu Y H. Human pose estimation based on secondary generation adversary[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201509.
- [3] Dabral R, Mundhada A, Kusupati U, et al. Learning 3D human pose from structure and motion[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11213: 679-696.
- [4] Kocabas M, Karagoz S, Akbas E. Self-supervised learning of 3D human pose using multi-view geometry [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 1077-1086.
- [5] 黄硕,胡勇,巩彩兰,等.基于梯度光图像的高精度三维人脸重建算法[J].光学学报,2020,40(4):0410001.  
Huang S, Hu Y, Gong C L, et al. High precision three-dimensional face reconstruction algorithm based on gradient light image[J]. Acta Optica Sinica, 2020, 40(4): 0410001.
- [6] Rhodin H, Robertini N, Casas D, et al. General automatic human shape and motion capture using volumetric contour cues[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in

- computer science. Cham: Springer, 2016, 9909: 509-526.
- [7] von Marcard T, Rosenhahn B, Black M J, et al. Sparse inertial poser: automatic 3D human pose estimation from sparse IMUs[J]. *Computer Graphics Forum*, 2017, 36(2): 349-360.
- [8] Loper M, Mahmood N, Romero J, et al. SMPL: a skinned multi-person linear model[J]. *ACM Transactions on Graphics*, 2015, 34(6): 1-16.
- [9] Mao A H, Zhang H, Liu Y X, et al. Easy and fast reconstruction of a 3D avatar with an RGB-D sensor[J]. *Sensors*, 2017, 17(5): 1113.
- [10] Bogo F, Kanazawa A, Lassner C, et al. Keep it SMPL: automatic estimation of 3D human pose and shape from a single image[M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9909: 561-578.
- [11] 曾志超, 李桂清, 邹敬仪, 等. 三维人体模型姿态与形状重构[J]. *计算机辅助设计与图形学学报*, 2019, 31(9): 1485-1493.
- Zeng Z C, Li G Q, Zou X Y, et al. Human model pose and shape reconstruction[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2019, 31(9): 1485-1493.
- [12] Omran M, Lassner C, Pons-Moll G, et al. Neural body fitting: unifying deep learning and model based human pose and shape estimation[C] // 2018 International Conference on 3D Vision (3DV), September 5-8, 2018, Verona, Italy. New York: IEEE Press, 2018: 484-494.
- [13] Kanazawa A, Black M J, Jacobs D W, et al. End-to-end recovery of human shape and pose[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7122-7131.
- [14] Kolotouros N, Pavlakos G, Daniilidis K. Convolutional mesh regression for single-image human shape reconstruction[C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 4496-4505.
- [15] Cao Z, Simon T, Wei S H, et al. Realtime multi-person 2D pose estimation using part affinity fields[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1302-1310.
- [16] Loper M M, Black M J. OpenDR: an approximate differentiable renderer[M] // Fleet D, Pajdla T, Schiele B, et al. *Computer vision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8695: 154-169.
- [17] Mehta D, Rhodin H, Casas D, et al. Monocular 3D human pose estimation in the wild using improved CNN supervision [C] // 2017 International Conference on 3D Vision (3DV), October 10-12, 2017, Qingdao, China. New York: IEEE Press, 2017: 506-516.