



(21) 申请号 202210817641.0
(22) 申请日 2022.07.13
(65) 同一申请的已公布的文献号
申请公布号 CN 114913552 A

G06V 10/44 (2022.01)
G06V 10/82 (2022.01)
G06N 3/04 (2006.01)
G06N 3/08 (2006.01)

审查员 李慧

(43) 申请公布日 2022.08.16
(73) 专利权人 南京理工大学
地址 210094 江苏省南京市玄武区孝陵卫
街200号

(72) 发明人 王康侃 李绍园 尹士豪 李薇

(74) 专利代理机构 青岛锦佳专利代理事务所
(普通合伙) 37283

专利代理师 朱玉建

(51) Int. Cl.
G06V 40/10 (2022.01)
G06V 10/42 (2022.01)

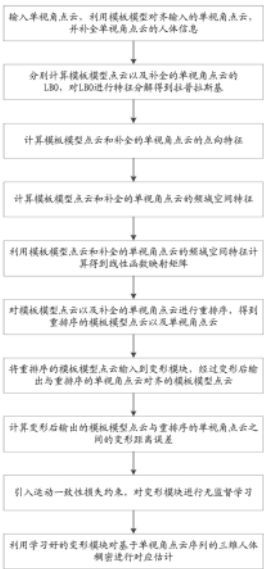
权利要求书3页 说明书7页 附图4页

(54) 发明名称

一种基于单视角点云序列的三维人体稠密
对应估计方法

(57) 摘要

本发明公开了一种基于单视角点云序列的三维人体稠密对应估计方法,其包括如下步骤:使用模板模型对齐输入的单视角点云来补全单视角点云的人体信息;计算模板模型点云和输入点云的LB0,得到拉普拉斯基;通过深度点云特征提取网络得到模板模型点云和输入点云的点云特征;拉普拉斯基与提取的特征矩阵相乘得到它们的频域空间特征;使用频域空间特征计算线性函数映射以表征稠密点对关系;通过点对关系分别对模板模型点云和输入点云进行重排序;将重排序的模板模型通过变形模块得到的变形距离误差作为单帧内部的无监督误差,并引入运动一致性损失函数对变形模块进行无监督学习。本发明能够准确地进行三维人体稠密对应估计。



1. 一种基于单视角点云序列的三维人体稠密对应估计方法,其特征在于,包括如下步骤:

步骤1. 利用模板模型对齐输入的单视角点云,并补全单视角点云的人体信息;

步骤2. 分别计算模板模型点云以及补全的单视角点云的 LB0,通过对LB0进行特征分解,分别得到模板模型点云以及补全的单视角点云的拉普拉斯基;

步骤3. 利用深度点云特征提取网络分别对模板模型点云以及补全的单视角点云进行特征提取,得到模板模型点云和补全的单视角点云的点向特征;

将模板模型点云的点向特征和补全的单视角点云的点向特征分别与对应点云的拉普拉斯基相乘,得到模板模型点云和补全的单视角点云的频域空间特征;

步骤4. 利用模板模型点云和补全的单视角点云的频域空间特征计算得到线性函数映射矩阵,以表示模板模型点云与补全的单视角点云之间的点对关系;

步骤5. 利用步骤4得到的点对关系,对模板模型点云以及补全的单视角点云进行重排序,得到重排序的模板模型点云以及重排序的单视角点云;

步骤6. 将补全的单视角点云的点向特征经过预处理后,与重排序的模板模型点云进行拼接,并输入到变形模块,经过变形后输出与重排序的单视角点云对齐的模板模型点云;

计算变形后输出的模板模型点云与重排序的单视角点云之间的变形距离误差;

将该变形距离误差作为变形模块的无监督学习误差,同时引入运动一致性损失约束,对变形模块进行无监督学习,进而得到学习好的变形模块;

步骤7. 利用变形模块对输入的单视角点云序列进行三维人体稠密对应估计。

2. 根据权利要求1所述的三维人体稠密对应估计方法,其特征在于,

所述步骤1具体为:

步骤1.1. 首先计算单视角点云的二维关节,再通过二维到三维的转换得到三维关节;

步骤1.2. 使用 KEEP IT SMPL 将模板模型通过最小化三维关节的误差,将模板模型变形至与输入的单视角点云对齐,此处对齐是指模板模型点云与单视角点云的躯干对齐;

步骤1.3. 用变形后的模板模型不可见点去补全输入的单视角点云中丢失的人体信息。

3. 根据权利要求1所述的三维人体稠密对应估计方法,其特征在于,

所述步骤2具体为:

步骤2.1. 首先将计算点云上所有顶点的切空间 $\hat{T}_P$;

步骤2.2. 然后将每个点邻域半径 r 以内的所有邻接点映射到该点的切空间 $\hat{T}_P$ 中;

步骤2.3. 计算所有投影后邻接点的 LB0算子;

拉普拉斯矩阵 L 定义为: $L=S^{-1}W$;其中, S 表示面积矩阵, W 表示权重矩阵;

步骤2.4. 将LB0通过特征分解 $W\Phi=S\Phi\Lambda$ 得到拉普拉斯基 Φ 和分解后得到的特征值 Λ 。

4. 根据权利要求3所述的三维人体稠密对应估计方法,其特征在于,

所述步骤2.3中,面积矩阵 S 和权重矩阵 W 的计算过程如下:

首先计算每个点的球形邻域,邻域半径为 r ,在这个邻域中找到其所有邻接点,将它们全部投影到该点的切空间中,并将投影点进行局部三角剖分,再计算点云LB0;

将面积矩阵S定义为:当 $i \neq j$ 时, $s_{ij} = \sum_{k \neq i, j} S_{ijk}$, 当 $i = j$ 时, $s_{ij} = - \sum_{k \neq i} S_{ik}$;

其中, i 是点云的一个顶点, j, k 是 i 的邻接点;

S_{ijk} 是属于 i, j, k 三个点组成的三角区域面积, s_{ij} 表示所有三角区域面积之和;且当 $i = j$ 时, s_{ii} 等于点 i 的所有邻接点所构成的面的面积之和;

将权重矩阵W定义为:当 $i \neq j$ 时, $w_{ij} = \sum (\cot \alpha_{ij} + \cot \beta_{ij}) / 2$, 当 $i = j$ 时, $w_{ij} = - \sum_{k \neq i} w_{ik}$;

其中, α_{ij} 和 β_{ij} 是边 e_{ij} 的两个角度, 边 e_{ij} 是指点 i 和点 j 所构成的边, 且当 $i = j$ 时, w_{ij} 等于点 i 的所有邻域的所有 w 之和, 即当 $i \neq j$ 时所计算出来的 w_{ij} ; w_{ij} 表示为权重矩阵W中第 i 行第 j 列的权重值, w_{ik} 表示权重矩阵W中第 i 行第 k 列的权重值。

5. 根据权利要求1所述的三维人体稠密对应估计方法, 其特征在于,

所述步骤3具体为:

将点云和与点云对应的拉普拉斯基共同作为输入, 通过网格下采样得到下采样后的点以及拉普拉斯基矩阵, 下采样后的点通过KPConv特征提取网络获得其特征;

将经过KPConv特征提取网络得到的特征, 与下采样后的拉普拉斯基矩阵相乘得到频域空间特征, 此时得到的频域空间特征为点向局部特征;

将该点向局部特征进一步经过一个全局特征提取网络, 得到全局特征。

6. 根据权利要求1所述的三维人体稠密对应估计方法, 其特征在于,

所述步骤4具体为:

定义M表示模板模型点云, N表示补全的单视角点云;

对于模板模型点云M计算出的拉普拉斯基为 Φ , 频域特征表示为 $A = \Phi \Pi_M$, 对于补全的单视角点云N计算出的拉普拉斯基为 Ψ , 频域特征表示为 $B = \Psi \Pi_N$;

其中, Π_M 表示模板模型点云的点向特征, Π_N 表示补全的单视角点云的点向特征;

最后通过最小二乘法获得函数映射矩阵C, 计算公式为:

$$C_{opt} = \operatorname{argmin}_C ||CA - B||^2 + \alpha ||\Lambda_M C - C \Lambda_N||^2;$$

其中, C_{opt} 表示函数映射矩阵C的优化函数, Λ_M 是模板模型点云LBO进行特征分解后得到的特征值, Λ_N 是补全的单视角点云LBO进行特征分解后得到的特征值。

7. 根据权利要求1所述的三维人体稠密对应估计方法, 其特征在于,

所述步骤5具体为:

步骤5.1. 首先将频域空间线性函数映射矩阵与模板模型点云和补全的单视角点云的拉普拉斯基进行矩阵相乘, 通过计算获得点向对应概率矩阵;

步骤5.2. 使用点向对应概率矩阵分别与模板模型点云矩阵、补全的单视角点云矩阵进行矩阵乘法计算, 分别计算得到重排序的模板模型点云和重排序的单视角点云。

8. 根据权利要求1所述的三维人体稠密对应估计方法, 其特征在于,

所述步骤6中, 当输入为连续帧的点云时, 将单帧无监督网络结构表达为时序上的网络结构, 通过帧与帧之间运动变化的几何特性作为时序损失函数进行无监督学习。

9. 根据权利要求8所述的三维人体稠密对应估计方法, 其特征在于,

所述步骤6中, 基于相邻三帧之间对应点之间的运动连续性来构造损失函数;

其中, 运动连续损失函数为: $I_{mc} = \sum_i ||v_{p,i}^{(t-1)} + v_{p,i}^{(t+1)} - 2v_{p,i}^{(t)}||_1$;

其中, i 表示模板模型上的点索引, $\sum_i$ 表示模板模型上所有点的运动连续损失之和;

$|| \cdot ||_1$ 表示深度学习中的L1损失函数;

$v_{p,i}^{(t-1)}$ 、 $v_{p,i}^{(t)}$ 、 $v_{p,i}^{(t+1)}$ 分别是在第t-1帧、第t帧、第t+1帧时i对应点坐标。

10. 根据权利要求9所述的三维人体稠密对应估计方法,其特征在於,

所述步骤6中,变形模块的损失函数 $L = l_d + l_{mc}$;

其中, l_d 为变形几何约束即变形后的模板模型点云与重排序的单视角点云间的变形距离误差; l_{mc} 为时序几何约束即基于相邻三帧之间对应点之间的运动连续性构造的损失函数;

通过最小化该变形模块的损失函数L,实现对变形模块的无监督学习。

一种基于单视角点云序列的三维人体稠密对应估计方法

技术领域

[0001] 本发明属于三维人体重建技术领域,特别涉及一种基于单视角点云序列的三维人体稠密对应估计方法。

背景技术

[0002] 目前在计算机视觉和图形学相关领域中,对应点估计或形状匹配是一个非常重要的问题。计算形状之间的对应点关系是可以帮助解决语义驱动技术的关键问题之一,且各种问题都可以归入形状对应的范畴,于是在近年来成为了计算机视觉领域的一个热门研究方向。形状对应背后的终极目标是识别形状的各个部分并推断它们的用途,因此建立从一个形状到另一个形状的映射在语义上是有意义的。形状点对关系是在三维对齐和时空重建等任务中的一个关键组成部分,也是属性迁移、形状插值和统计建模等各种应用中不可或缺的先决条件。

[0003] 由于表示三维对象、场景的灵活性和效率性问题,三维点云已被广泛应用于各种应用中,例如互动通信、自动驾驶、AR/VR等。在AR/VR中,对应点关系估计是三维重建和运动跟踪等AR/VR关键技术的基石。例如虚拟试衣系统中需要建立三维人体模型,并且对齐人体模型与服装,合成试衣结果,最终着装的模型不能出现错位或运动不一致的情况,便需要应用点对关系的技术。同时它也在影视、医疗、游戏等方面也有着广泛的应用前景。

[0004] 人体稠密对应点估计可以用于形状插值,在插值或变形中,一个形状逐渐变换成另一个形状,如果满足一定的审美要求,因此形状的渐变具有视觉上的愉悦感,由此用于影视或者游戏方面;在医学领域,可以跟踪病人皮肤上痣的数量和密度的变化(用于癌症预测)。近年来,消费型深度相机得到了快速的发展,如微软公司推出的 Kinect,此类相机除了可以获得 RGB 图像外,还可以获得对应的深度图像。由深度图像便可以获得点云,相对于网格模型的难以获取,点云更容易获得,所以估计点云的稠密对应点更是大势所趋。

[0005] 随着深度学习技术的发展,点云上人体稠密对应点估计性能得到了很大的发展,而点云及其序列的使用,则将进一步提高三维人体稠密对应点估计的精度,为理解人类行为提供更好地帮助,但是点云作为输入也带来了许多的困难和挑战。由于每个相机、扫描仪在自己的空间而不是对象空间中产生一个点云,因此,两个点云(甚至即使它们代表同一对象)之间没有关联,这对下游处理和分析构成了巨大挑战,如运动传输、形状编辑、动态点云压缩、三维对象识别、点云注册、形状检索、表面重建以及三维物体分类等等。

[0006] 目前估计稠密形状对应已经有了相当多的方法,大致可以分为两类:基于手工描述符的方法和数据驱动的方法。基于手工描述符的方法通常预定义手工描述符来计算每个点的独有特征来优化最终对应结果。除此之外,还有Monti等人提出的将点进行分类的方法:输入一个模型后,经过深度学习后将每个点进行分类,也可以做到估计两个模型的对应关系。最近,基于深度学习的方法训练神经网络来估计点对关系,在很大程度上提高了性能。

[0007] 然而,现有的稠密形状估计方法要么查询大量难以获得的注释数据,要么假设输

入数据中的连接信息可用,即多边形网格。由于在更容易获取到的点云中,并没有任何的边连接信息,因此,也更难以计算出准确的特征来预估点对关系。事实上,通常人们只能观察到单一视角的可见点,所以这样的部分点云是没有完整人体信息的。因此,没有边连接信息且在有遮挡或单一视角的情况下,两个点云之间建立对应点关系成为了一个巨大的挑战。

发明内容

[0008] 本发明的目的在于提出一种基于单视角点云序列的三维人体稠密对应估计方法,以便能够准确地估计三维人体稠密对应。

[0009] 本发明为了实现上述目的,采用如下技术方案:

[0010] 一种基于单视角点云序列的三维人体稠密对应估计方法,包括如下步骤:

[0011] 步骤1. 利用模板模型对齐输入的单视角点云,并补全单视角点云的人体信息;

[0012] 步骤2. 分别计算模板模型点云以及补全的单视角点云的 LB0,通过对LB0进行特征分解,分别得到模板模型点云以及补全的单视角点云的拉普拉斯基;

[0013] 步骤3. 利用深度点云特征提取网络分别对模板模型点云以及补全的单视角点云进行特征提取,得到模板模型点云和补全的单视角点云的点向特征;

[0014] 将模板模型点云的点向特征和补全的单视角点云的点向特征分别与对应点云的拉普拉斯基相乘,得到模板模型点云和补全的单视角点云的频域空间特征;

[0015] 步骤4. 利用模板模型点云和补全的单视角点云的频域空间特征计算得到线性函数映射矩阵,以表示模板模型点云与补全的单视角点云之间的点对关系;

[0016] 步骤5. 利用点对关系对模板模型点云以及补全的单视角点云进行重排序,进而得到重排序的模板模型点云以及重排序的单视角点云;

[0017] 步骤6. 将补全的单视角点云的点向特征经过预处理后,与重排序的模板模型点云进行拼接,并输入到变形模块,经过变形后输出与重排序的单视角点云对齐的模板模型点云;

[0018] 计算变形后输出的模板模型点云与重排序的单视角点云之间的变形距离误差;

[0019] 将该变形距离误差作为无监督学习误差,同时引入运动一致性损失约束,对变形模块进行无监督学习,进而得到学习好的变形模块;

[0020] 步骤7. 利用变形模块对输入的单视角点云序列进行三维人体稠密对应估计。

[0021] 本发明具有如下优点:

[0022] 如上所述,本发明述及了一种基于单视角点云序列的三维人体稠密对应估计方法,该方法不仅补全了单视角点云缺失的人体信息,而且采用新的计算LB0方法提高对应精度,引入了变形模块来构建基于数据驱动的无监督方法,解决了目前三维人体对应估计存在缺乏有标签数据集的问题,而且引入运动一致性损失,有效改善了对应结果不稳定或抖动的情况。本发明方法有效地提高了三维人体稠密对应估计结果的精确度。

附图说明

[0023] 图1为本发明实施例中基于单视角点云序列的三维人体稠密对应估计方法的流程图。

[0024] 图2为本发明实施例中单帧网络结构图。

- [0025] 图3为本发明实施例中连续帧结构图。
- [0026] 图4为本发明实施例中点向特征和频域空间特征提取网络结构图。
- [0027] 图5为本发明实施例中全局特征提取网络结构图。
- [0028] 图6为本发明实施例中变形模块的结构图。
- [0029] 图7为本发明实施例中运动一致性示意图。

具体实施方式

[0030] 下面结合附图以及具体实施方式对本发明作进一步详细说明：

[0031] 如图1所示，基于单视角点云序列的三维人体稠密对应估计方法，包括如下步骤：

[0032] 步骤1. 利用模板模型对齐输入的单视角点云，并补全单视角点云的人体信息。

[0033] 本实施例首先利用OPENPOSE计算单视角点云的二维关节像素坐标，再通过二维到三维的转换得到三维关节坐标。具体操作为：

[0034] 步骤1.1. OPENPOSE生成的关节 $v_{ij}=(i, j)$ ，表示二维关节像素坐标，通过彩色图像与深度图像匹配获得该关节像素坐标的深度 d ，得到关节的三维坐标，如公式(1)所示。

[0035] $x=(i-c_x)*d/f_x, y=(j-c_y)*d/f_y$ (1)

[0036] 其中， (i, j) 表示二维关节像素坐标；关节的三维坐标为 $v_{ij}=(x, y, z)$ ， $z=d$ ，常数 c_x, c_y, f_x 以及 f_y 均为相机内参。

[0037] 然后使用 KEEP IT SMPL 将模板模型通过最小化三维关节的误差，将模板模型变形至与输入的单视角点云大致对齐，此处的大致对齐即两个点云躯干对齐即可。

[0038] 因为设置了SMPL模型作为限制，能够防止在变形过程中出现畸变。SMPL是一个使用姿态参数 Θ 和形状参数 β 表达的函数 $M(\cdot)$ ，它能表达顶点数为6890的网格模型：

[0039] $M(\beta, \Theta)=W(T(\beta, \Theta), J(\beta), \Theta, \mathcal{W})$ ；

[0040] $T(\beta, \Theta)=T+B_s(\beta)+B_p(\Theta)$ ；

[0041] 其中， $W(T(\beta, \Theta), J(\beta), \Theta, \mathcal{W})$ 代表SMPL函数； $T(\beta, \Theta), J(\beta), \Theta, \mathcal{W}$ 分别为SMPL函数的输入参数， $\mathcal{W}$ 代表SMPL的权重矩阵和 $B_p(\Theta)$ 代表姿态依赖变形函数。

[0042] 该SMPL函数允许模板模型姿态 $T+B_s(\beta)$ 基于关节 $J(\beta)$ 变形，就能够让模板模型变形到与输入的单视角点云大致对齐，然后用变形后的模板模型不可见点补全输入的单视角点云，以补全输入的单视角点云中丢失的人体信息，具体操作为：计算变形后模板模型的法向量，只保留法向量为负的点，直接放在输入的单视角点云后面。

[0043] 也就是，最终补全后点云的点集由两部分组成，前部分为原始输入的点，后部分为补充的点，这样就获得了一个较为完整的人体结构。将模板模型点云和补全的单视角点云将作为之后模块的输入。与直接缺失一大部分的点云相比，本发明通过补全后的点云能够更好地预测出对应关系，利用SMPL 模型作为限制，可以防止在变形过程中出现畸变。

[0044] 步骤2. 分别计算模板模型点云以及补全的单视角点云的 LB0，通过对LB0进行特征分解，分别得到模板模型点云以及补全的单视角点云的拉普拉斯基。

[0045] Laplace-Beltrami算子(LB0)是半正定的，通过拉普拉斯变换将欧式空间扩展到非欧几里得空间，即频域空间。LB0是一个内蕴特征，也就是说，它可以完全用一个黎曼矩阵来表示。因此，它具有流形恒定变形的性质。在密集黎曼流形上，它允许分解相应的特征函数，从而形成一个正交基，也就是拉普拉斯基，它是傅里叶基在非欧几里得域的推广。

[0046] 人体其实是一种黎曼流体,是嵌入在高维空间中的低维流形。对于拥有黎曼流形性质的人体,通常对点云的离散表示进行采样。Belkin等人提出了不依赖于全局网格模型的方法,它通过切空间中的局部三角剖分来计算点云中相应的LB0。

[0047] 然而这个方法有个很致命的缺陷,因为,在计算两个互为邻域的点时,它们各自的切平面不同,所以它们映射后的三角面片的面积会不同而导致求出的 LB0矩阵不对称。在实验时,也曾尝试直接将LB0算子改为对称的,然而结果并不理想。切空间是由微分流形在一点上相关的向量空间,是光滑曲线的切线和欧几里得空间中光滑曲面的切线平面的推广。

[0048] 该步骤2具体为:

[0049] 步骤2.1. 首先将计算点云上所有顶点的切空间 $\hat{T}_P$ 。

[0050] 步骤2.2. 然后将每个点邻域半径 r 以内的所有邻接点映射到该点的切空间 $\hat{T}_P$ 中。

[0051] 步骤2.3. 计算所有投影后邻接点的 LB0算子。

[0052] 拉普拉斯矩阵 L 定义为: $L=S^{-1}W$,其中, S 表示面积矩阵, W 表示权重矩阵。

[0053] 面积矩阵 S 和权重矩阵 W 的计算过程如下:

[0054] 首先计算每个点的球形邻域,邻域半径为 r ,在这个邻域中找到其所有邻接点,将它们全部投影到该点的切空间中,并将投影点进行局部三角剖分,再计算点云 LB0。

[0055] 将面积矩阵 S 定义为:当 $i \neq j$ 时, $s_{ij} = \sum_{k \neq i, j} S_{ijk}$,当 $i = j$ 时, $s_{ij} = - \sum_{k \neq i} S_{ik}$ 。

[0056] 其中, i 是点云的一个顶点, j, k 是 i 的邻接点。

[0057] S_{ijk} 是属于 i, j, k 三个点组成的三角区域面积, s_{ij} 表示所有三角区域面积之和;且当 $i = j$ 时, s_{ii} 等于点 i 的所有邻接点所构成的面的面积之和。

[0058] 将权重矩阵 W 定义为:当 $i \neq j$ 时, $w_{ij} = \sum (\cot \alpha_{ij} + \cot \beta_{ij}) / 2$,当 $i = j$ 时, $w_{ij} = - \sum_{k \neq i} w_{ik}$ 。

[0059] 其中, α_{ij} 和 β_{ij} 是边 e_{ij} 的两个角度,边 e_{ij} 是指点 i 和点 j 所构成的边,且当 $i = j$ 时, w_{ij} 等于点 i 的所有邻域的所有 w 之和,即当 $i \neq j$ 时所计算出来的 w_{ij} , w_{ij} 表示为权重矩阵 W 中第 i 行第 j 列的权重值, w_{ik} 表示权重矩阵 W 中第 i 行第 k 列的权重值。

[0060] 步骤2.4. 将LB0通过特征分解 $W\Phi = S\Phi \Lambda$ 得到拉普拉斯基 Φ 和特征值 Λ 。

[0061] 模板模型点云以及补全的单视角点云分别按照上述步骤2.1至步骤2.4计算,得到模板模型点云以及补全的单视角点云的拉普拉斯基 Φ 和 Ψ 。

[0062] 由于邻域点是可逆的,因此根据邻域信息保留的 LB0也就是对称的。

[0063] 本发明利用新提出的面积矩阵和权重矩阵表示方法,利于得到更优的点云局部几何特征表示,通过将每个点的邻域点映射到该点切空间后进行三角剖分近似得到点云的邻接关系,最后利用计算出的所有面积和权重可以得到两个输入点云的拉普拉斯基 Φ 和 Ψ 。

[0064] 步骤3. 利用深度点云特征提取网络分别对模板模型点云以及补全的单视角点云进行特征提取,得到模板模型点云和补全的单视角点云的点向特征。

[0065] 将模板模型点云的点向特征和补全的单视角点云的点向特征分别与对应点云的拉普拉斯基相乘,得到模板模型点云和补全的单视角点云的频域空间特征。

[0066] 点云作为形状的离散表示,已经是现今形状表示中最便捷,最容易获取到的人体

表示形式。而点云只有三维坐标,所以需要通过深度学习获取更丰富的特征。

[0067] 特征描述符描述了对象的局部特征,属于谱形状分析方法。随着神经网络的发展,越来越多的学者将卷积神经网络引入到了点云特征提取中来。

[0068] 本发明结合 KPConv进行特征提取,具体的局部特征提取网络见图4。

[0069] 以补全的单视角点云作为输入点云为例:

[0070] 输入点云的大小为 6890×3 ,输入点云的拉普拉斯基大小为 6890×50 ,将输入点云和与输入点云对应的拉普拉斯基共同作为输入,通过网格下采样。

[0071] 下采样得到下采样后的点(大小为 6890×3)以及拉普拉斯基矩阵(大小为 6890×50)。

[0072] 下采样后的点再通过KPConv特征提取网络获得其特征,大小为 6890×50 。

[0073] 将KPConv特征提取网络得到的特征与下采样后的拉普拉斯基矩阵相乘,得到单视角点云的频域空间特征,此时得到的频域空间特征为点向局部特征,大小为 6890×50 。

[0074] 其中,6890为特征点数,50为维数。

[0075] 本实施例中的KPConv特征提取网络为局部特征提取网络,其包括KPConv、ResNet、Strided ResNet、ResNet、最近上采样以及一维卷积。

[0076] 利用KPConv 进行特征提取,通过插值和平滑来计算点云的局部特征。

[0077] 进一步将点向局部特征经过全局特征提取网络进而得到全局特征,如图5示出了全局特征提取网络,其包括 $d \times d$ 维的一维卷积、GN、LeakyReLU、最大池以及平均池。

[0078] 其中,输入的点向局部特征为 $n \times d$ 维,此处 $n=6890$, $d=50$ 。

[0079] 点向局部特征经过一个一维卷积后,分别输入最大池和平均池,两个得到 $1 \times d$ 维的全局特征,两个 $1 \times d$ 维的全局特征拼接后得到 $1 \times 2d$ 维的全局特征。

[0080] 经过全局特征提取网络得到的该全局特征即补全的单视角点云的频域空间特征B,如图2所示,同理,模板模型点云的频域空间特征A也按照上述方法得到,此处不再详细赘述。

[0081] 步骤4. 利用模板模型点云和补全的单视角点云的频域空间特征计算得到线性函数映射矩阵,以表示模板模型点云与补全的单视角点云之间的点对关系Corr。

[0082] 利用非刚性流体在频域空间中的线性特征通过最小二乘法计算函数映射矩阵。

[0083] 函数映射首先是给定两个形状M和N,它们的标量函数为 $f:M \rightarrow R$ 和 $g:N \rightarrow R$,它们之间有一个可逆的映射T,可以进行可逆变换 $g=f \cdot T^{-1}$ 。

[0084] 通过计算拉普拉斯基,可以将欧式空间中复杂的三维对应T转换为频域空间中的线性关系,即 $T_F:F(M,R) \rightarrow F(N,R)$,两个形状之间的关系便可以表示 $CA=B$ 。

[0085] 其中,A和B分别为两个形状在频域空间上的特征,由此获得函数映射矩阵C。

[0086] 通过上述公式可以得到令人满意的点对关系,这样就将复杂的欧式空间对应关系转换为频域空间线性关系,可以通过简单的最小二乘法获得函数映射矩阵。

[0087] 定义M表示模板模型点云,N表示补全的单视角点云。

[0088] 对于模板模型点云M计算出的拉普拉斯基为 Φ ,频域特征表示为 $A=\Phi \Pi_M$,对于补全的单视角点云N计算出的拉普拉斯基为 Ψ ,频域特征表示为 $B=\Psi \Pi_N$ 。

[0089] 其中, Π_M 表示模板模型点云的点向特征, Π_N 表示补全的单视角点云的点向特征。

[0090] 最后通过最小二乘法获得函数映射矩阵C,计算公式为:

[0091] $C_{\text{opt}} = \text{argmin}_c ||CA-B||^2 + \alpha ||\Lambda_M C - C \Lambda_N||^2$ 。

[0092] 其中, C_{opt} 表示函数映射矩阵C的优化函数, Λ_M 是模板模型点云LB0进行特征分解后得到的特征值, Λ_N 是补全的单视角点云LB0进行特征分解后得到的特征值。

[0093] 步骤5. 利用点对关系Corr对模板模型点云以及补全的单视角点云进行重排序, 进而得到重排序的模板模型点云以及重排序的单视角点云。该步骤具体5具体为:

[0094] 步骤5.1. 首先将频域空间线性函数映射矩阵与模板模型点云和补全的单视角点云的拉普拉斯基进行矩阵相乘, 计算获得点向对应概率矩阵P, $P = \Phi C \Psi^T$ 。

[0095] P是 $n \times n$ 的矩阵, n 为输入模板模型点云的点数, P_{ij} 表示第 i 点与第 j 点对应的概率。

[0096] 步骤5.2. 使用点向对应概率矩阵分别与模板模型点云矩阵、补全的单视角点云矩阵进行矩阵乘法计算, 分别计算得到重排序的模板模型点云和重排序的单视角点云。

[0097] 此种点云重排序的方式可避免传统的点对概率方法存在的如下问题:

[0098] 对点云的重排序, 最初使用的是通过点对概率, 计算出最大概率得到软对应点, 再将点云根据软对应重排序, 但是由于这个方式不可导, 会导致梯度不下降。

[0099] 步骤 6. 将补全的单视角点云的点向特征经过预处理后, 与重排序的模板模型点云进行拼接, 并输入到MLP变形模块, 经过变形后输出与重排序的单视角点云对齐的模板模型点云。

[0100] 其中, 预处理过程包括最大池化以及复制 n 维操作, 如图2和图6所示。

[0101] 计算变形后输出的模板模型点云与重排序的单视角点云之间的变形距离误差 l_d 。

[0102] 将该变形距离误差 l_d 作为无监督学习误差, 同时引入运动一致性损失约束, 对变形模块进行无监督学习, 得到学习好的变形模块。因此, 变形模块的损失函数 $L = l_d + l_{mc}$ 。

[0103] 其中, l_d 为变形几何约束即变形后的模板模型点云与重排序的单视角点云间的变形距离误差; l_{mc} 为时序几何约束即基于相邻三帧之间对应点之间的运动连续性构造的损失函数。

[0104] 如图7所示, 本发明实施例在对变形模块进行无监督学习时, 通过运动分析, 还引入了运动一致性损失约束, 进一步优化变形模块的模型参数。

[0105] 通过引入运动一致性约束, 使得当模型输入连续帧点云序列时, 变形模块能更好的预测出变形点云与输入点云序列之间的稠密对应, 使模型表现稳定和防止抖动发生。

[0106] 具体的, 当输入为连续帧的点云时, 将单帧无监督网络结构表达为时序上的网络结构, 通过帧与帧之间运动变化的几何特性, 作为时序损失函数进行无监督学习。

[0107] 如图7所示, 本发明基于相邻三帧之间对应点之间的运动连续性来构造损失函数。

[0108] 这个损失函数将同一个点看作在微小时刻内的运动距离为相同的, 因此认为对应到模板模型点云同一个点的三个帧的点应该是同一个点在不同时间帧所处的位置。

[0109] 而它在连续三帧之间的运动应该时连续线性的, 所以, 运动连续损失函数为:

[0110] $l_{mc} = \sum_i ||v_{p,i}^{(t-1)} + v_{p,i}^{(t+1)} - 2v_{p,i}^{(t)}||_1$ 。

[0111] 其中, i 表示模板模型上的点索引; $\sum_i$ 表示模板模型上所有点的运动连续损失之和; $v_{p,i}^{(t-1)}$ 、 $v_{p,i}^{(t)}$ 、 $v_{p,i}^{(t+1)}$ 分别是在第 $t-1$ 帧、第 t 帧、第 $t+1$ 帧时 i 对应点坐标。

[0112] 此时在运动状态下假设第 $t-1$ 帧和第 t 帧运动距离 $d_i^{t,t-1}$ 与第 $t+1$ 帧和第 t 帧运动距离 $d_i^{t,t+1}$ 相同, 也就是假设时间足够短的情况下帧与帧之间同一个点的运动距离是接近相同的。

[0113] 其中, $d_i^{t,t+1} = ||v_{p,i}^{(t+1)} - v_{p,i}^{(t)}||_1$, $d_i^{t,t-1} = ||v_{p,i}^{(t)} - v_{p,i}^{(t-1)}||_1$, $|| \cdot ||_1$ 表示深度学习中的L1损失函数。

[0114] 在图7显示了连续三帧点云运动一致性示意。

[0115] 本发明通过运动一致性损失约束进一步优化模型参数,使得模型最终输入是连续帧点云序列时,变形模块能更好地预测出变形点云与输入点云序列之间的稠密对应。

[0116] 通过最小化该变形模块的损失函数L,实现对变形模块的无监督学习。

[0117] 步骤7. 利用变形模块对输入的单视角点云序列进行三维人体稠密对应估计。

[0118] 具体的,将单视角点云序列逐帧输入,各帧单视角点云分别按照上述步骤1至5的顺序进行处理,最后输入到变形模型,进行三维人体稠密对应估计。

[0119] 本发明实施例通过变形模块构建上述无监督约束,有效地解决了点云对应标签缺少的问题,并且通过连续运动约束,很好地减少了对应结果的不稳定性。

[0120] 当然,以上说明仅仅为本发明的较佳实施例,本发明并不限于列举上述实施例,应当说明的是,任何熟悉本领域的技术人员在本说明书的教导下,所做出的所有等同替代、明显变形形式,均落在本说明书的实质范围之内,理应受到本发明的保护。

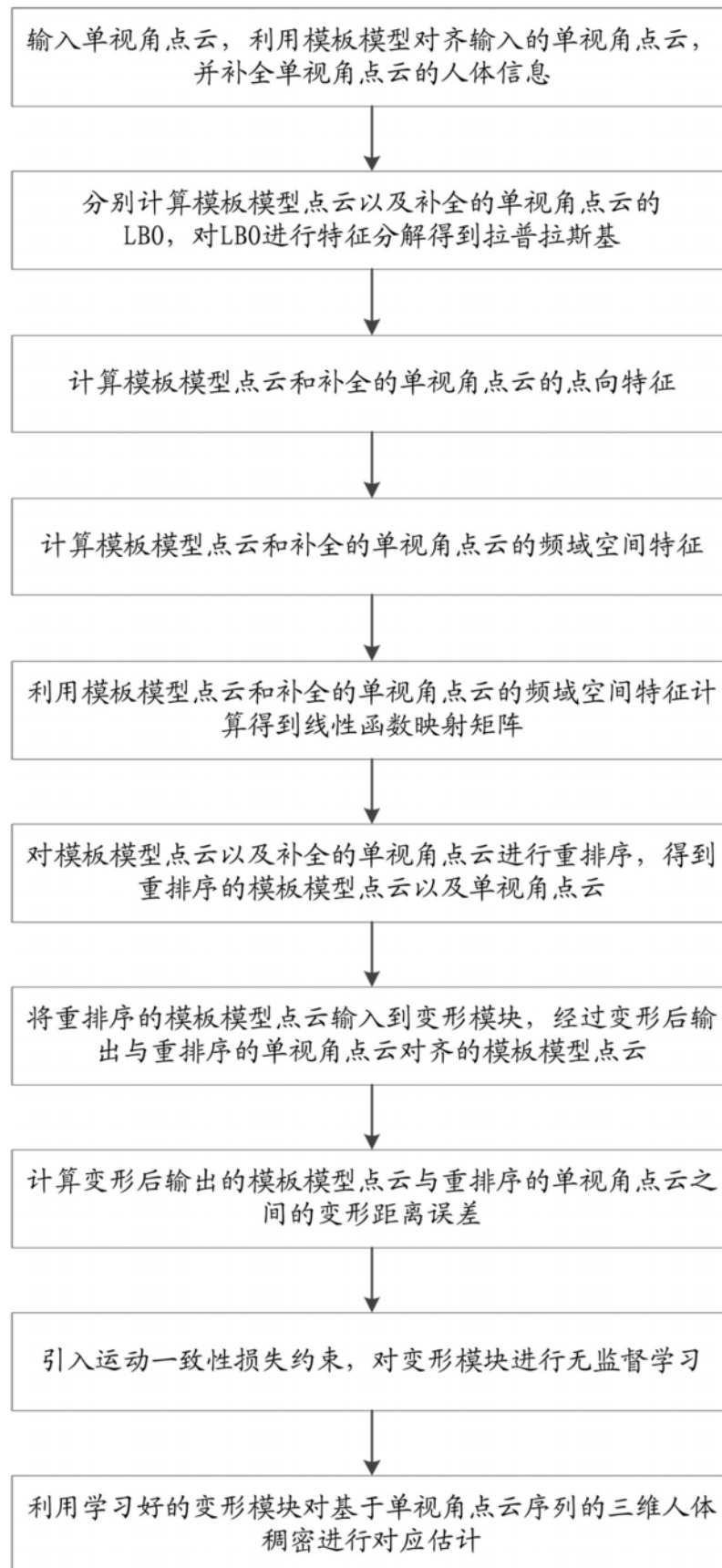


图1

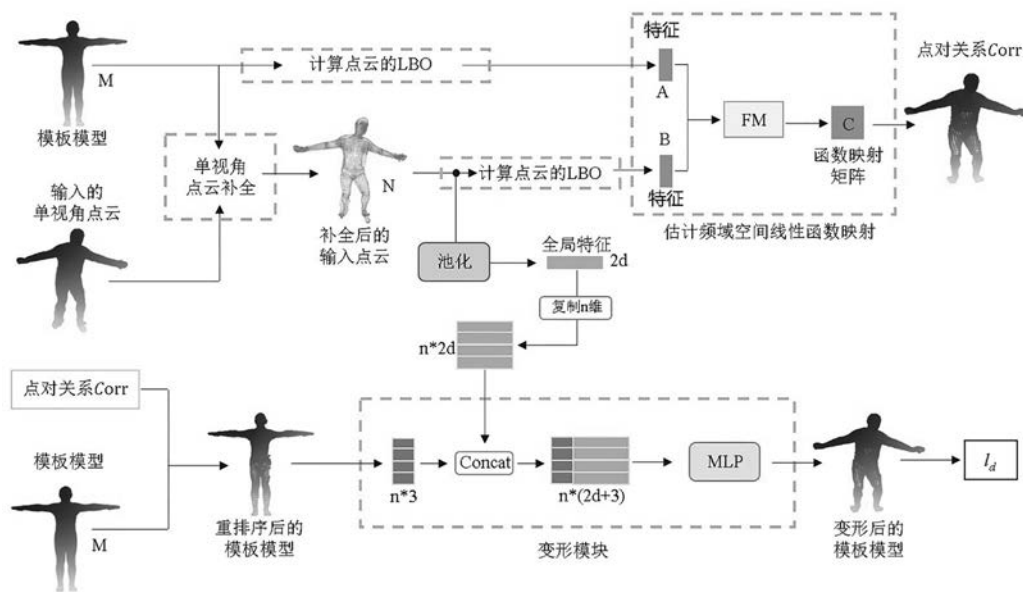


图2

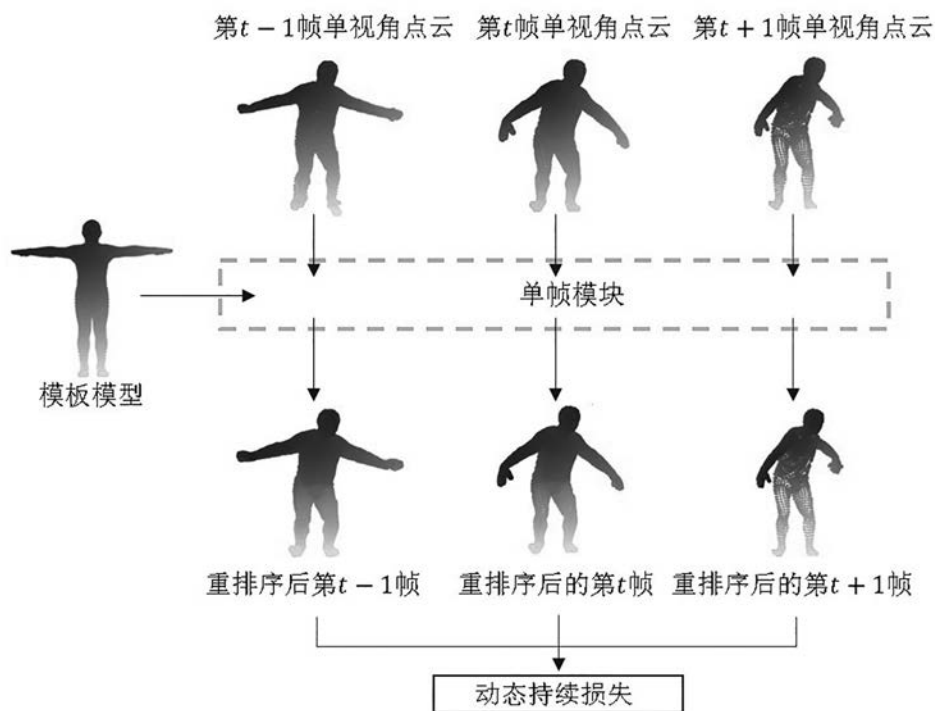


图3

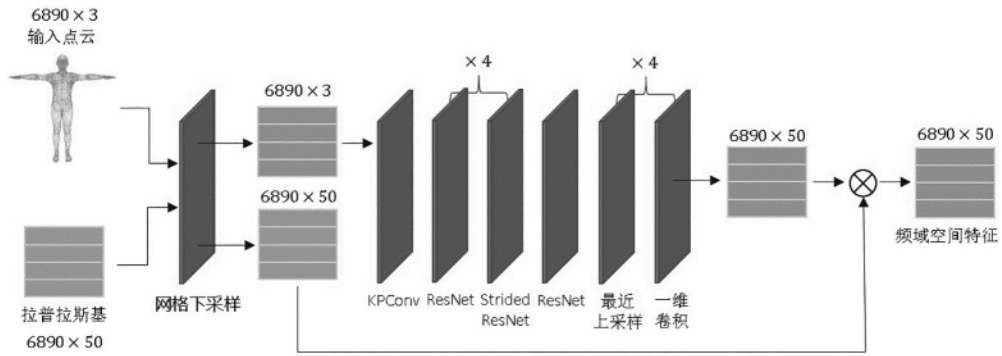


图4

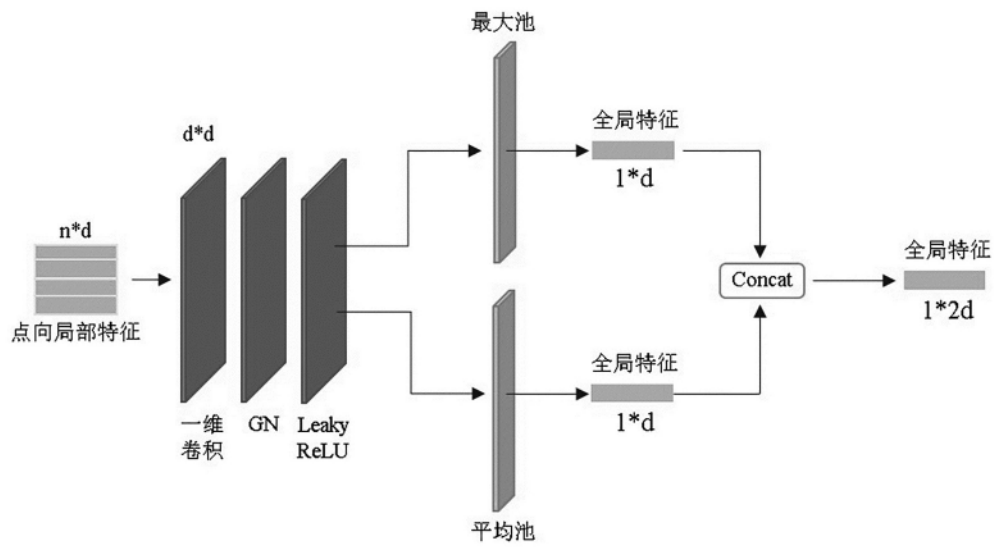


图5

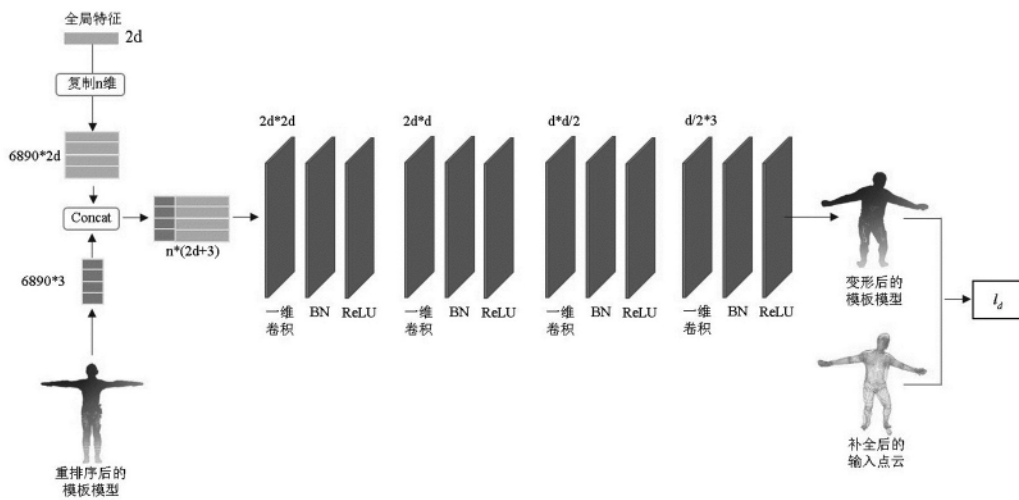


图6

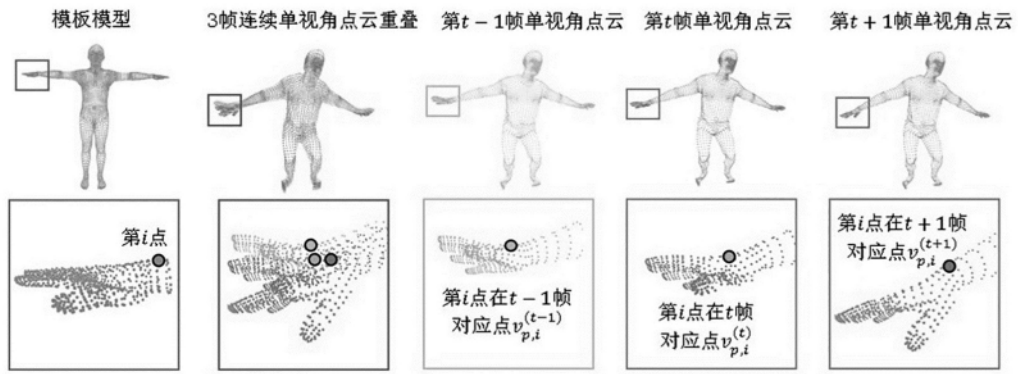


图7