

人体动态三维语义化重建

清华大学 于涛

GAMES 2019.12.12

报告内容

- 1 研究背景及意义**
- 2 基于骨架跟踪的实时人体动态三维重建**
- 3 实时人体动态三维信息双层语义化重建**
- 4 高精度单视角人体快速动作捕捉算法**
- 5 实时人体动态三维信息多层语义化重建**
- 6 单图像人体三维重建**
- 7 总结与展望**

相关发表文章

1. BodyFusion (ICCV 2017)

Real-time Capture of Human Motion and Surface Geometry Using a Single Depth Camera

2. DoubleFusion (CVPR 2018 Oral, T-PAMI 2019)

Real-time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor

3. SimulCap (CVPR 2019)

Single-View Human Performance Capture with Cloth Simulation

4. DeepHuman (ICCV 2019 Oral)

3D Human Reconstruction from a Single Image

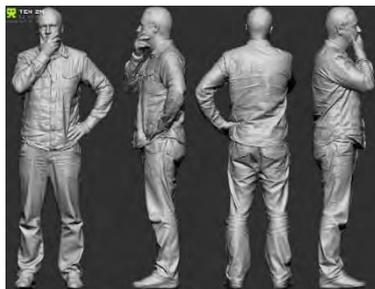
实验室-清华大学刘烨斌研究组: liuyebin.com

个人主页: ytrock.com

1.1 实时人体动态三维重建应用领域



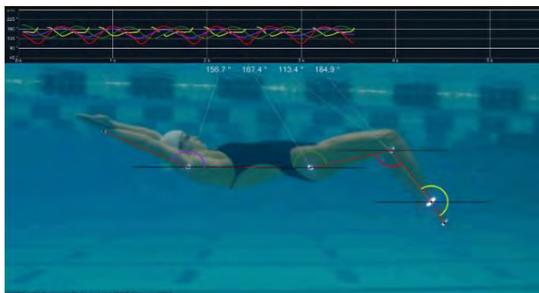
全息通讯



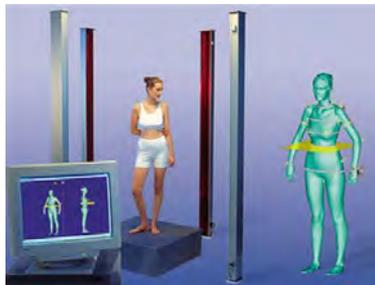
人体建模



影视游戏制作



体育运动分析指导



人体测量



公共安全监测（步态识别）



微软全息通讯系统 2016年

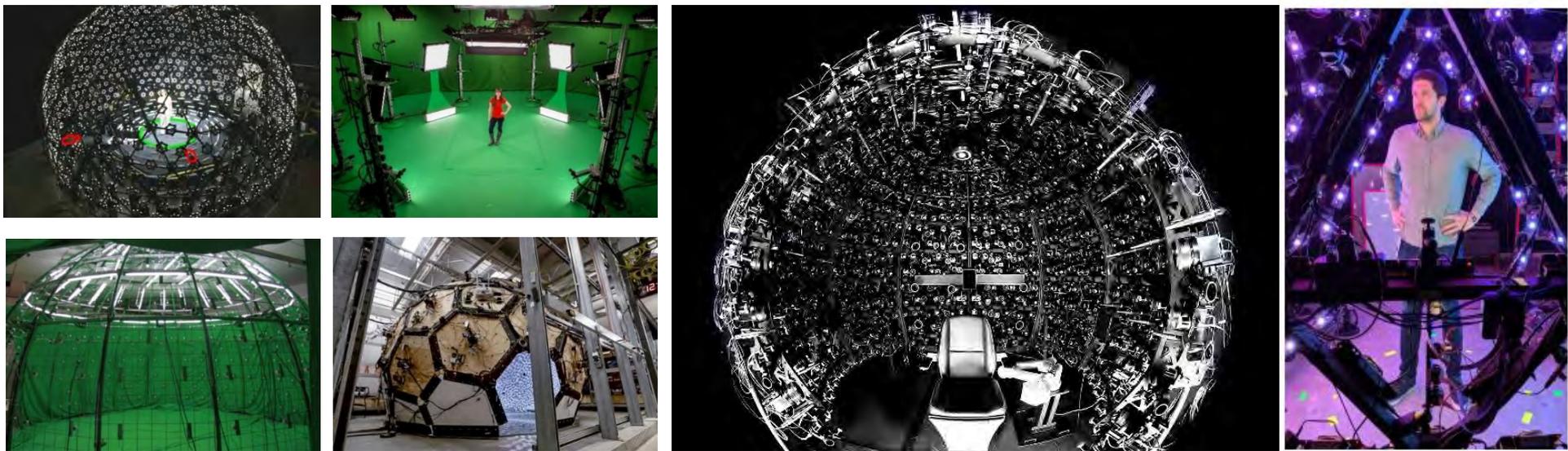
核心：实时人体动态三维重建

实时重建人体三维形状、运动及表面颜色

1.2 研究现状

离线多视点人体三维重建

- 重建精度高，但造价昂贵（多相机+多光照系统）、数据量大且重建效率低



USC LightStage, Microsoft H-Cap, Tsinghua Dome,
CMU Panoptic Studio, Facebook Mugsy, Google The Relightables

1.2 研究现状

实时多视点人体三维重建

- 可以实现实时重建，但多深度相机系统构造复杂，便携性较差



Microsoft Fusion4D



Google Motion2Fusion



Tsinghua UnstructuredFusion

Dou et, al. [2016], Dou et, al. [2017], Xu et, al. [2019]

1.2 研究现状

单视点人体静态三维重建

- 采集过程复杂、重建速度慢（数分钟到数小时）



Newcombe et, al. [2011]

Tong et, al. [2012]

Li et, al. [2013]

Cui et, al. [2013]

Dou et, al. [2015]

Zhu et, al. [2017]

... ..



1.2 研究现状

主动式人体动作捕捉系统

- 精度较高，但穿戴复杂、限制人体运动



(a) Gypsy机电系统



(b) Motion Star电磁系统



(c) MVN惯性传感系统

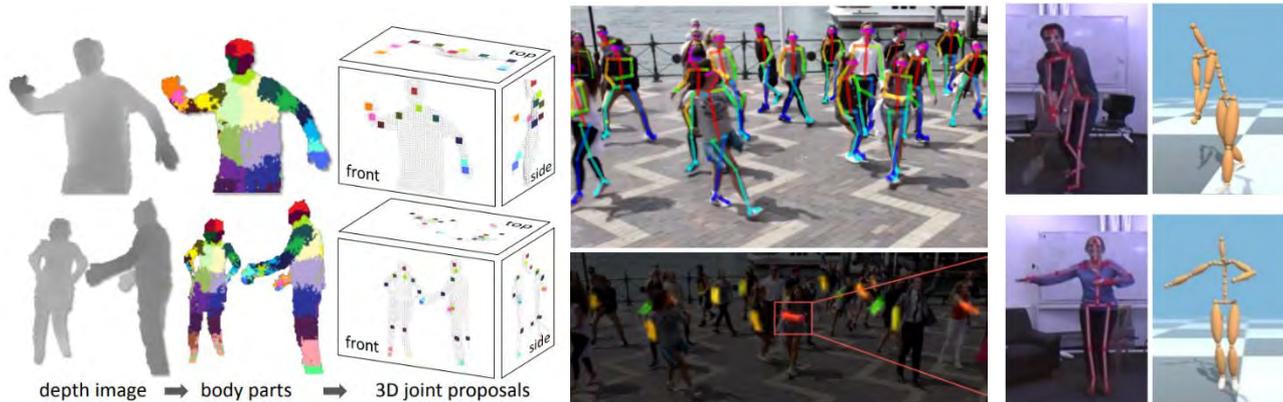


(d) LED光学标记系统

1.2 研究现状

无标记人体动作捕捉

无标记动作捕捉算法（骨架识别/姿态检测）：



- 仅输出人体骨架，无法描述真实外观

无标记动作捕捉算法（姿态检测+骨架跟踪）：

Shotton et, al. [2011]

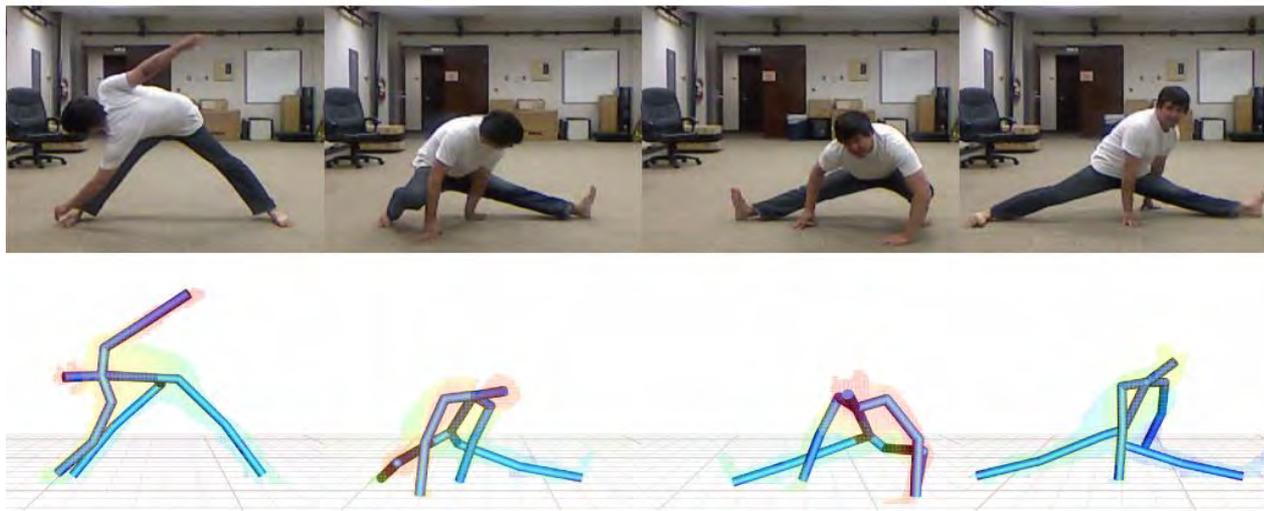
Shih-En et, al. [2016]

Cao et, al. [2017]

Wei et, al. [2012]

Mehta et, al. [2017]

... ..

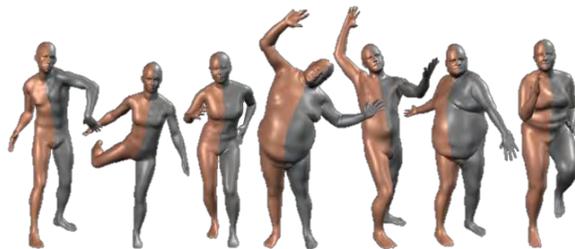


1.2 研究现状

人体动态三维重建 – 参数化人体模型



Anguelov et al. [SCAPE] 2005

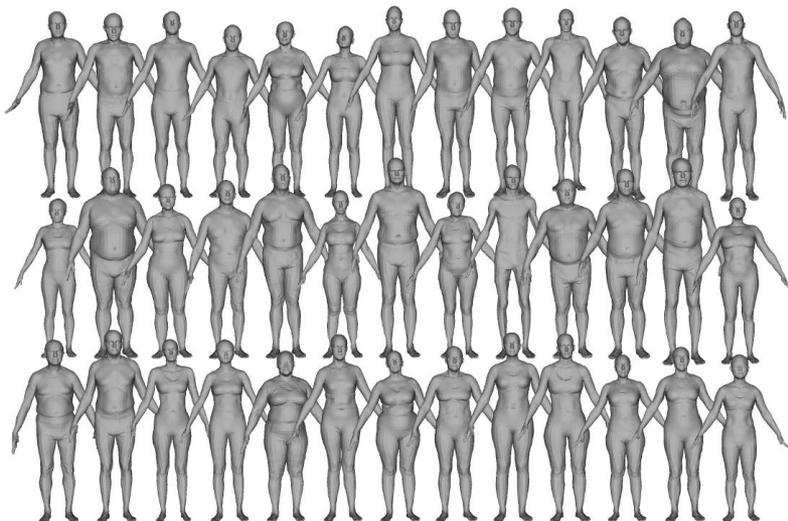


Loper et al. [SMPL] 2015



Pons-Moll et al. [DYNA] 2015

使用大量人体扫描数据，降维出一组基函数，对基函数进行加权以获得不同体形



参数化人体模型SMPL的数学定义：

$$M(\vec{\beta}, \vec{\theta}) = W(T_p(\vec{\beta}, \vec{\theta}), J(\vec{\beta}), \vec{\theta}, \mathcal{W})$$

$$T_p(\vec{\beta}, \vec{\theta}) = \bar{\mathbf{T}} + B_s(\vec{\beta}) + B_p(\vec{\theta})$$

$$B_s(\vec{\beta}; \mathcal{S}) = \sum_{n=1}^{|\vec{\beta}|} \beta_n \mathbf{S}_n$$

$$B_p(\vec{\theta}; \mathcal{P}) = \sum_{n=1}^{9K} (R_n(\vec{\theta}) - R_n(\vec{\theta}^*)) \mathbf{P}_n$$

1.2 研究现状

人体动态三维重建 – 参数化重建

- 人体动态三维参数化重建 (重建鲁棒性高但无法精确描述真实形状)



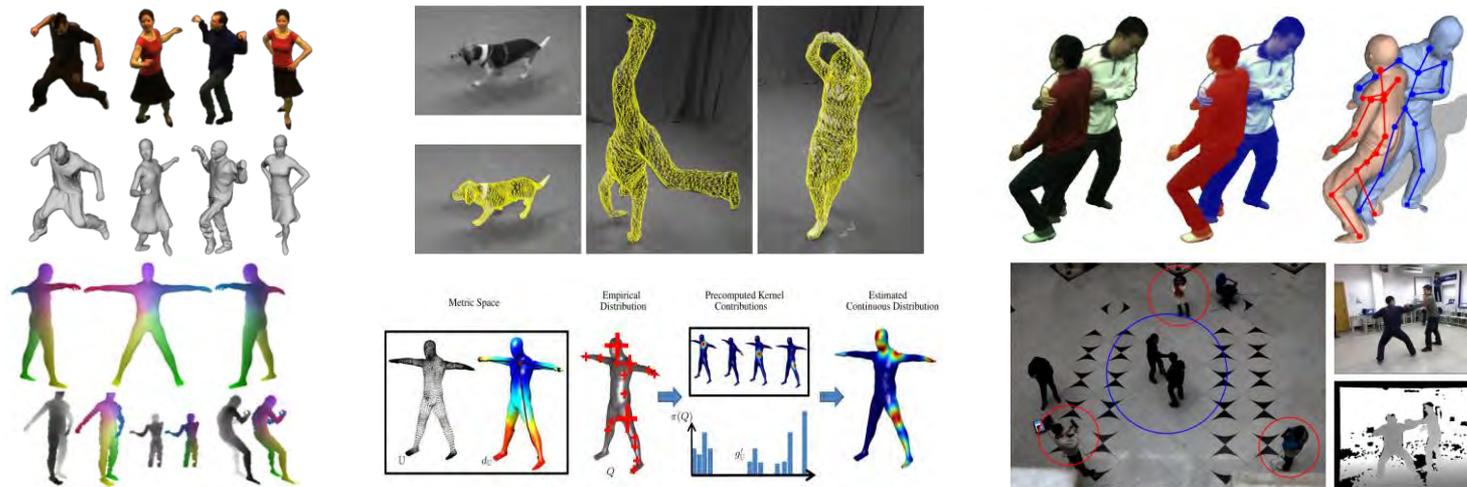
Ye et al. [2014], Chen et al. [2016], Bogo et al. [2015, 2016], Alldieck et al. [2018], Hanbyul Joo et al. [2018], Zhang et al. [2017], Pons-Moll et al. [2017]

1.2 研究现状

人体动态三维重建 – 预扫描模板重建

- 基于预扫描模板的人体动态重建 (需要预先扫描静态人体三维模板)

[预扫描模板+骨架跟踪] Vlasic et al. 2008, Gall et al. 2009, Liu et al. 2011, Taylor et al. 2012, Pons-Moll et al. 2015,



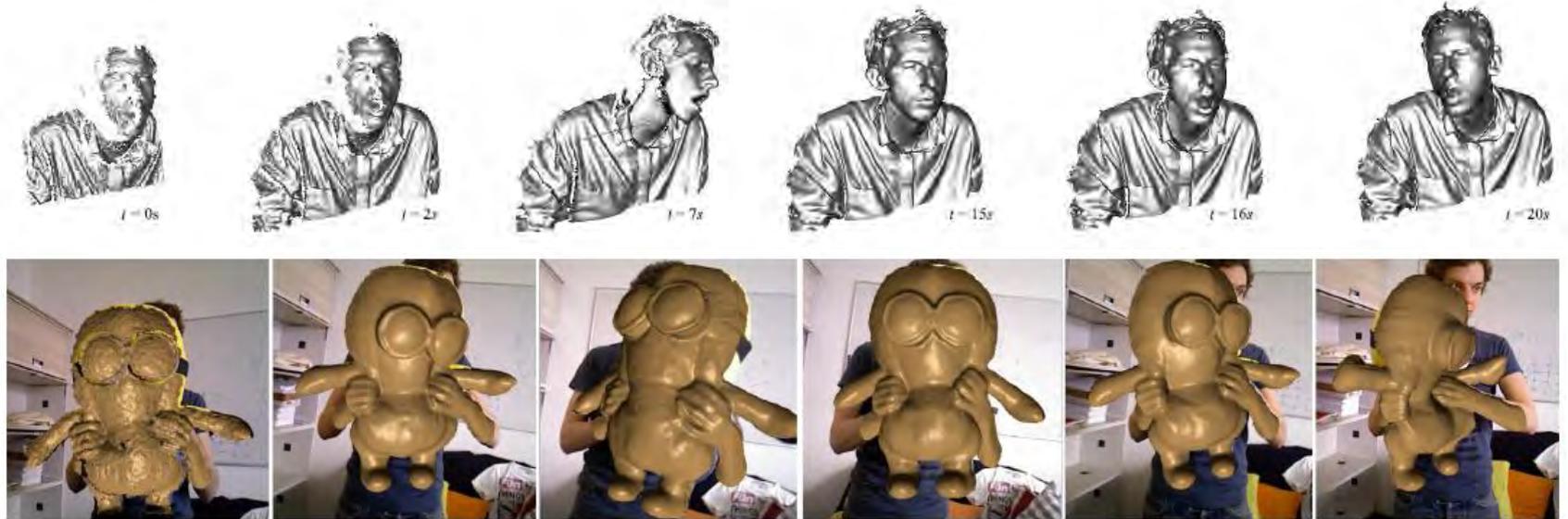
[预扫描模板+表面非刚性跟踪] Sumner et al. 2007, Li et al. 2009, Zollhofer et al. 2014, Guo et al. 2015



1.2 研究现状

单视点实时动态三维重建 – 体融合重建

- 无需参数化模型和预扫描模板，实时重构三维形状及运动
- 无法处理快速人体运动、无法估计人体体形、无法重建衣物动态细节



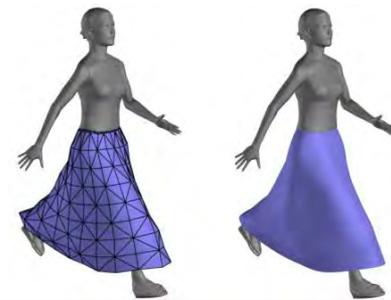
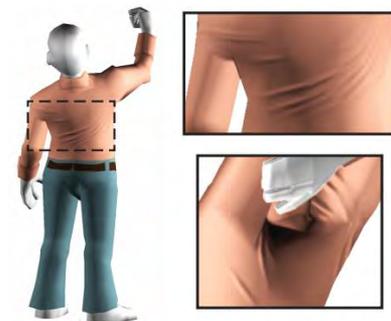
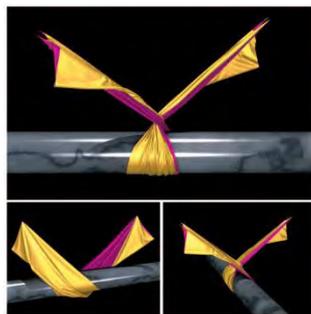
Newcombe et al. 2015, Innmann et al. 2016, Guo et al. 2017, Slavcheva et al. 2017, 2018, Li et al. 2018

1.2 研究现状

衣物动态重建-布料仿真

- 高质量衣物仿真耗时长
- 实时仿真效果需提高

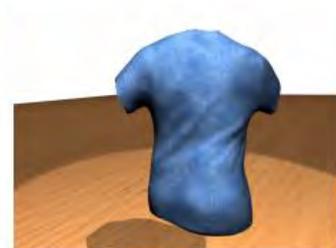
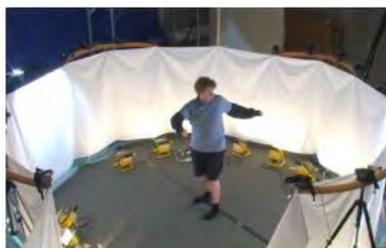
Demetri et al. 1987
Rony et al. 2007
Andrew et al. 2009
Edilson et al. 2010
Guan et al. 2012
Liu et al. 2013
Russel et al. 2015
Jiang et al. 2017, 2018
... ..



1.2 研究现状

衣物动态重建-衣物采集

- 需要特制衣物 (纹理)
- 需要多视点采集系统



Ryan et al. 2007

Bradley et al. 2008

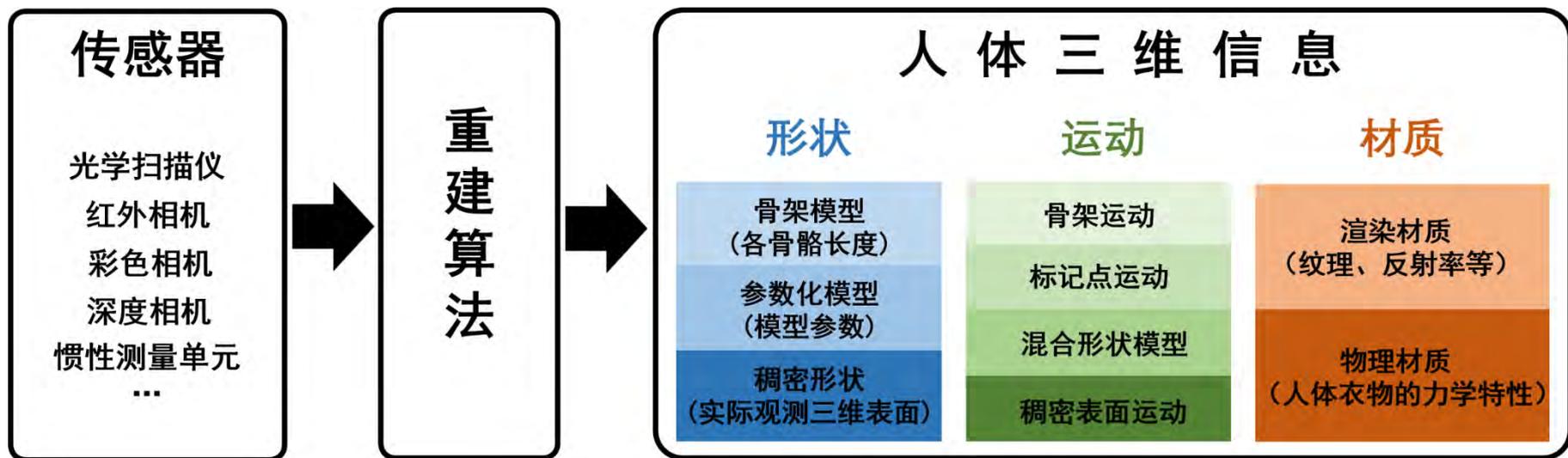
Popa et al. 2009

Pons-Moll et al. 2017

... ..

1.3 人体三维重建系统基本框图

人体三维信息具有多层语义化特性



1.4 人体三维信息

人体形状常用描述方法

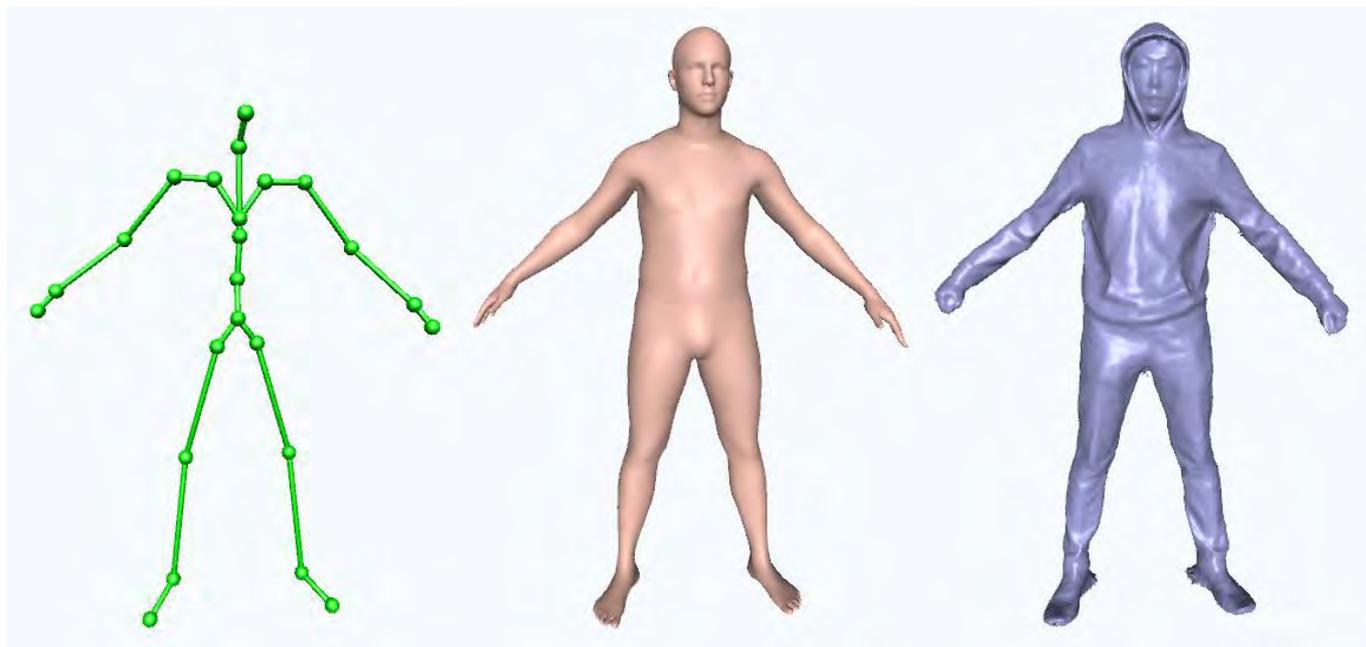
抽象性、语义化特性

形状

骨架模型
(各骨骼长度)

参数化模型
(模型参数)

稠密形状
(实际观测三维表面)



真实性

1.4 人体三维信息

人体运动常用描述方法

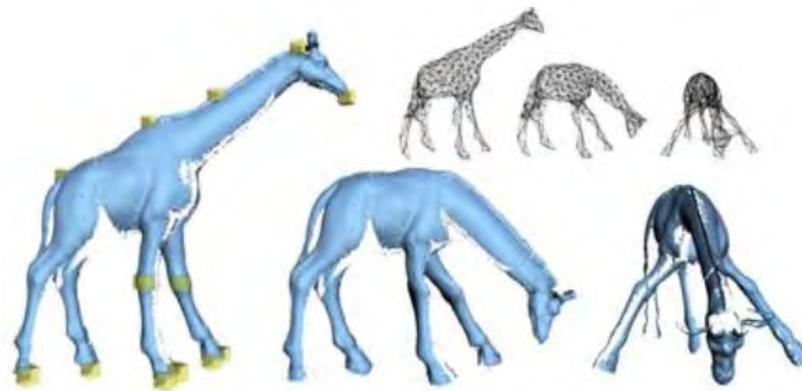
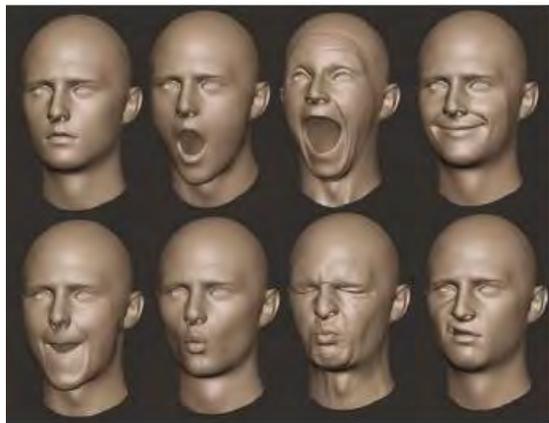
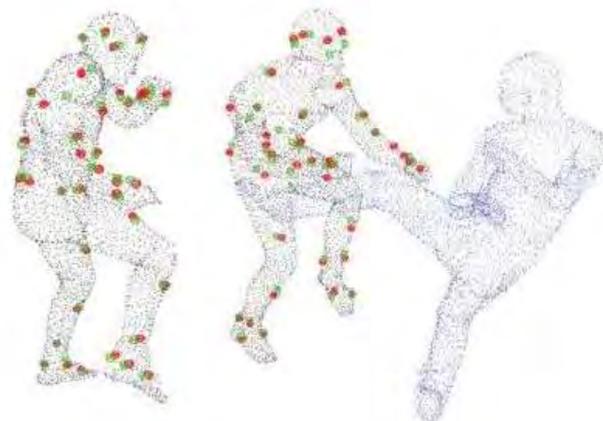
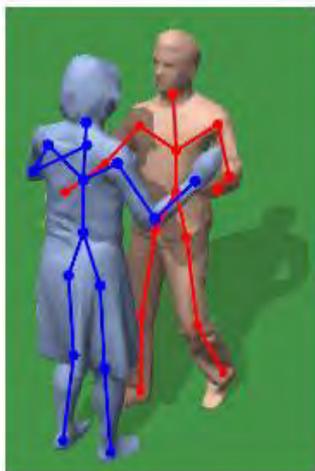
运动

骨架运动

标记点运动

混合形状模型

稠密表面运动



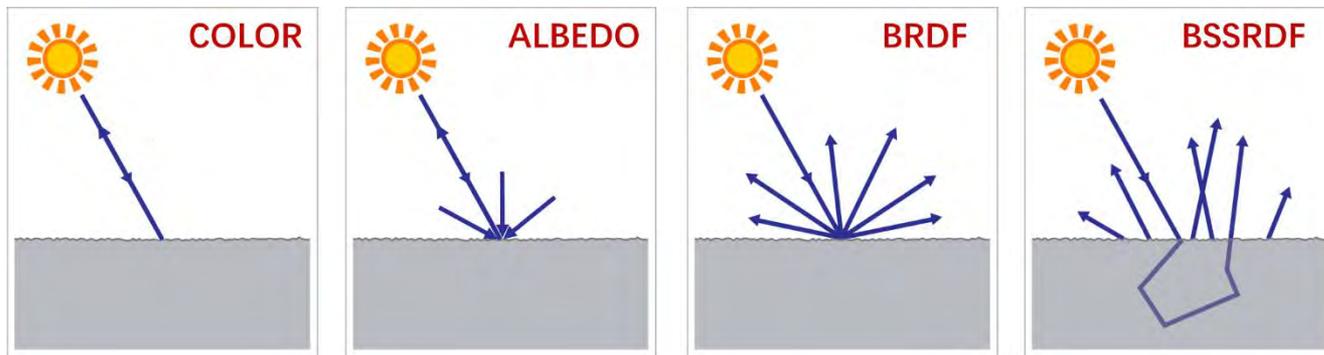
1.4 人体三维信息

人体材质常用描述方法

材质

渲染材质
(纹理、反射率等)

物理材质
(人体衣物的力学特性)



1.5 概述

人体动态三维语义化重建

难点:

1. 无法处理肢体运动

2. 无法重建人体体形

3. 无法处理复杂遮挡

4. 无法重建衣物细节

5. 重建依赖深度信息

研究内容与创新点:

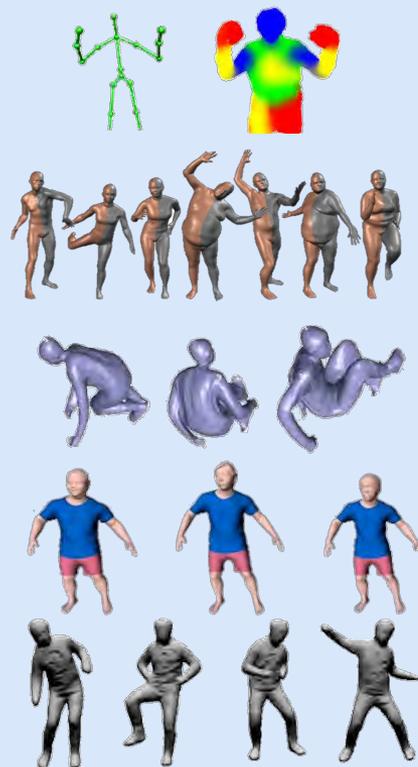
骨架化跟踪绑定重建

参数化双层语义重建

语义化双向混合重建

物理化多层服装重建

单图像人体三维重建



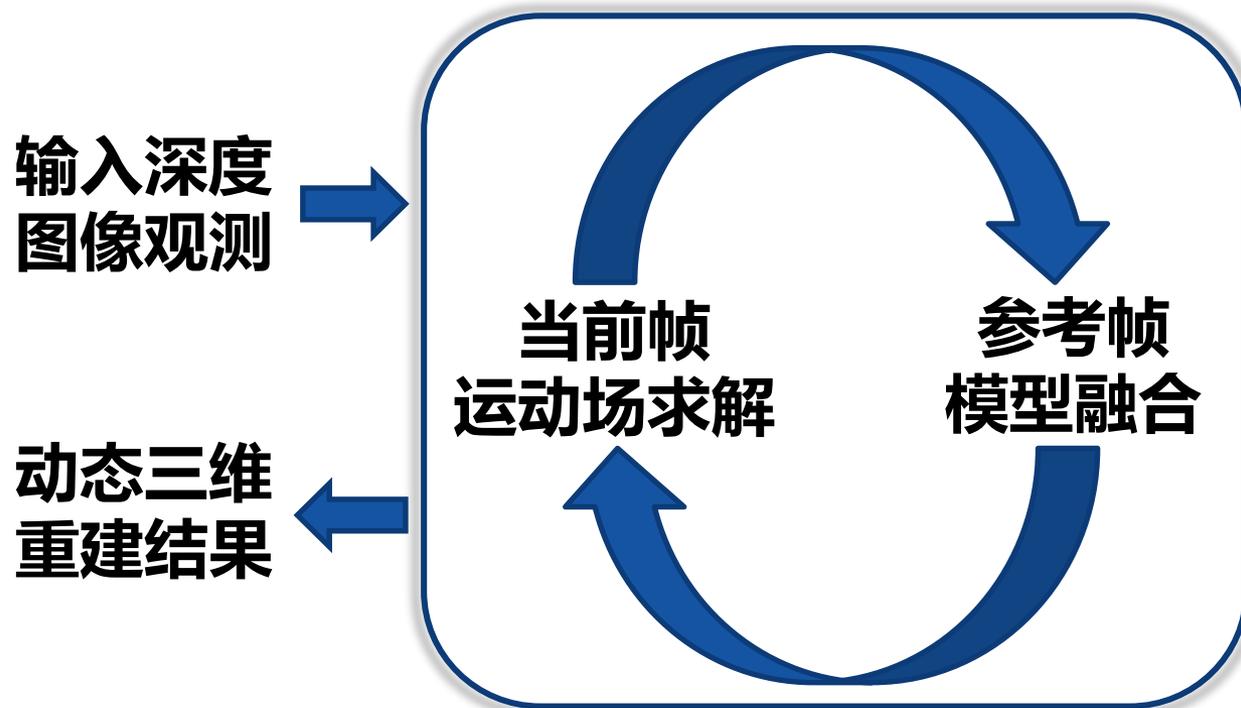
利用人体动态三维信息的多层语义化特性，提升人体动态三维重建的精度、鲁棒性和便捷性。

报告内容

- 1 研究背景及意义
- 2 基于骨架跟踪的实时人体动态三维重建**
- 3 实时人体动态三维信息双层语义化重建
- 4 高精度单视角人体快速动作捕捉算法
- 5 实时人体动态三维信息多层语义化重建
- 6 单图像人体三维重建
- 7 总结与展望

2 基于骨架跟踪的实时人体动态三维重建

2.1 单视点实时动态三维重建基本框架



2 基于骨架跟踪的实时人体动态三维重建

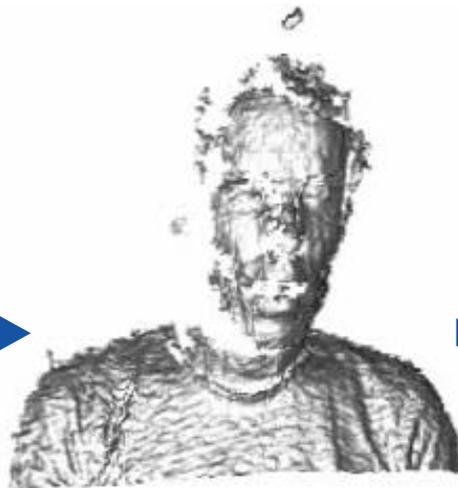
2.1 单视点实时动态三维重建基本框架



参考彩色图像序列



输入深度图像序列



参考帧模型融合



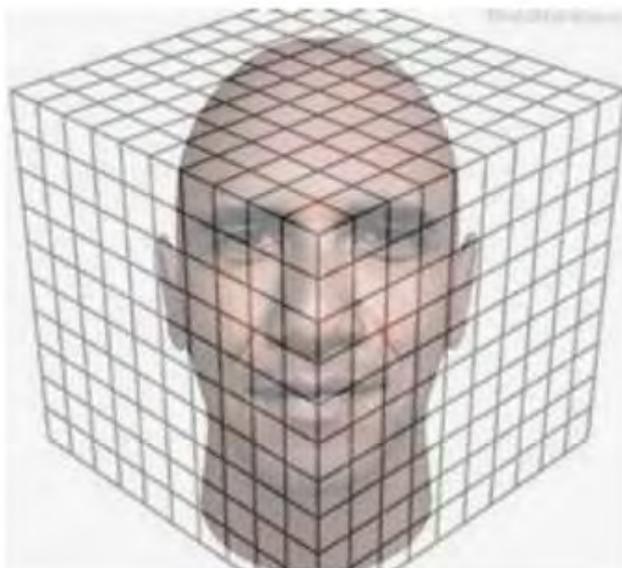
当前帧运动场求解



2 基于骨架跟踪的实时人体动态三维重建

2.1 单视点实时动态三维重建基本框架

符号距离函数



二维切片



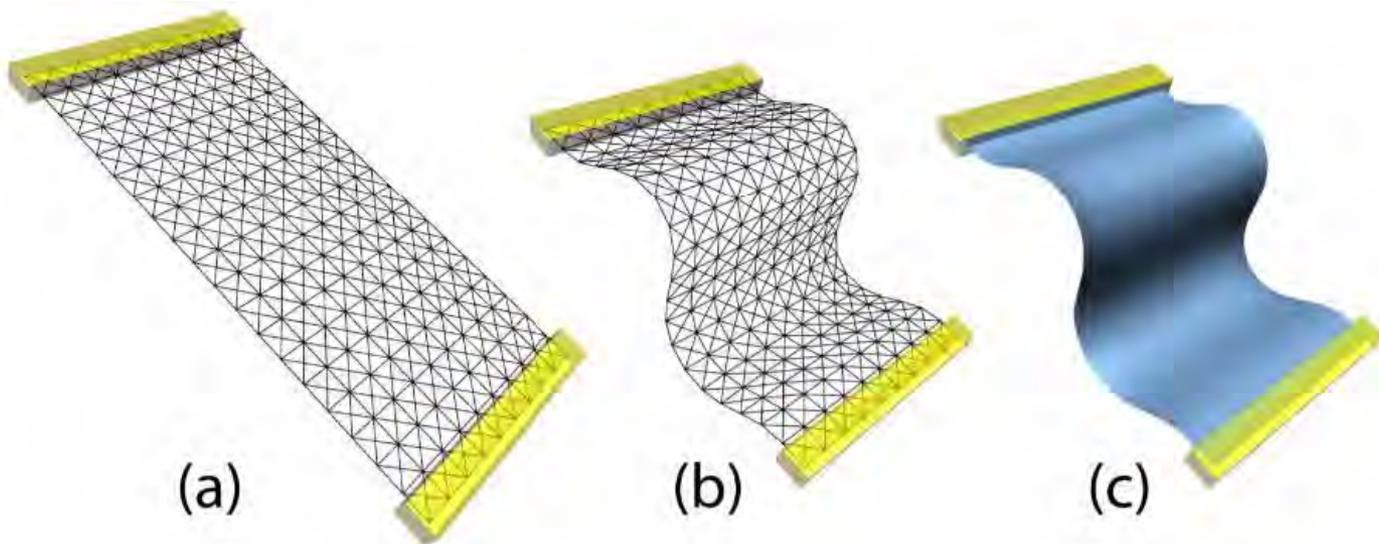
-0.9	-0.4	-0.1	0.2	0.9	1	1	1	1	1
-1	-0.9	-0.2	0.1	0.5	0.9	1	1	1	1
-1	-0.9	-0.3	0.2	0.2	0.8	1	1	1	1
-1	-0.9	-0.4	0.2	0.2	0.8	1	1	1	1
-1	-1	-0.8	-0.1	0.2	0.6	0.8	1	1	1
-1	-0.9	-0.3	-0.2	0.3	0.7	0.9	1	1	1
-1	-0.9	-0.4	-0.1	0.3	0.8	1	1	1	1
-0.9	-0.7	-0.5	0.0	0.4	0.9	1	1	1	1
-0.1	-0.9	-0.1	0.1	0.4	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1

- 符号距离函数定义在三维空间中（常在空间中一个均匀剖分的立方体内定义）
- 每个小立方体（体素）的中心点为函数采样点，函数值为该点到实际三维表面的距离
- 真实三维表面为符号距离函数的零值面

2 基于骨架跟踪的实时人体动态三维重建

2.1 单视点实时动态三维重建基本框架

嵌入式变形图模型：对真实表面非刚性运动的局部刚性近似



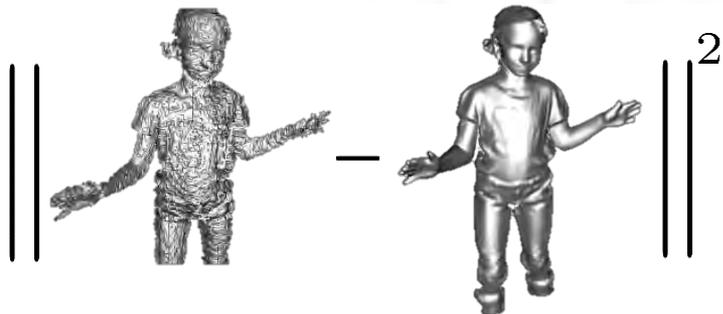
- 在模型表面均匀采集变形节点，每个节点都具有6自由度刚性运动
- 使用变形节点构造嵌入式变形图结构
- 使用嵌入式变形图进行表面非刚性运动的求解或驱动

2 基于骨架跟踪的实时人体动态三维重建

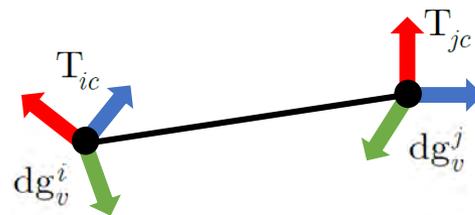
2.1 单视点实时动态三维重建基本框架

表面非刚性运动求解

$$E(\mathcal{W}_t, \mathcal{V}, D_t, \varepsilon) = E_{data}(\mathcal{W}_t, \mathcal{V}, D_t) + \lambda E_{reg}(\mathcal{W}_t, \varepsilon)$$



$$E_{data}(\mathcal{W}, \mathcal{V}, D_t) \equiv \sum_{u \in \Omega} \psi_{data}(\hat{\mathbf{n}}_u^T (\hat{\mathbf{v}}_u - \mathbf{v}l_{\hat{u}}))$$



$$E_{reg}(\mathcal{W}, \varepsilon) \equiv \sum_{i=0}^n \sum_{j \in \varepsilon(i)} \alpha_{ij} \psi_{reg}(T_{ic} dg_v^j - T_{jc} dg_v^i)$$

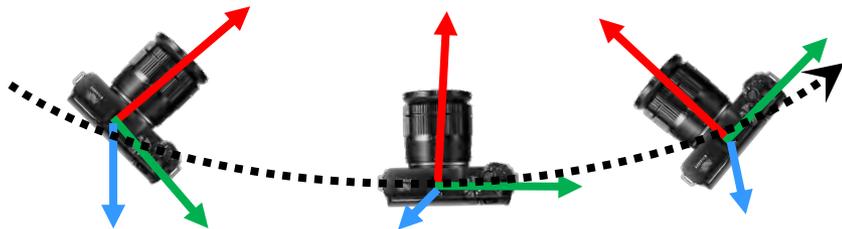
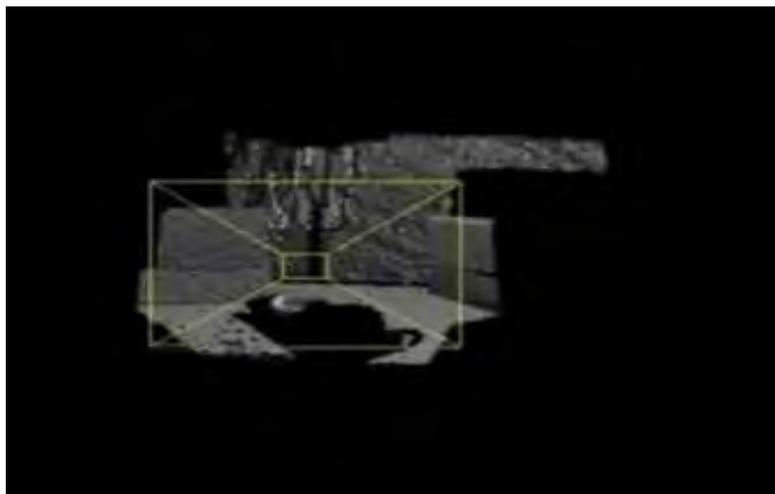
求解：高斯牛顿法（Gauss-Newton）

关键：雅可比矩阵的并行构造和求解

2 基于骨架跟踪的实时人体动态三维重建

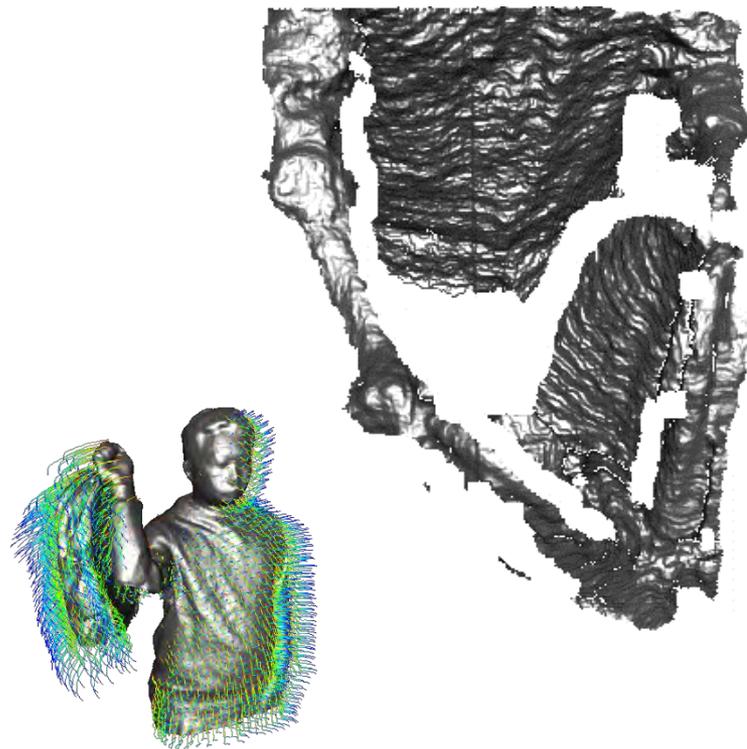
2.2 单视点实时动态三维重建难点

静态三维重建



变量：单个相机位姿
自由度：6

动态三维重建



变量：高维非刚性变形场
自由度： ≈ 6000

2 基于骨架跟踪的实时人体动态三维重建

2.2 单视点实时动态三维重建难点

现有方法重建结果：DynamicFusion (CVPR15最佳论文)



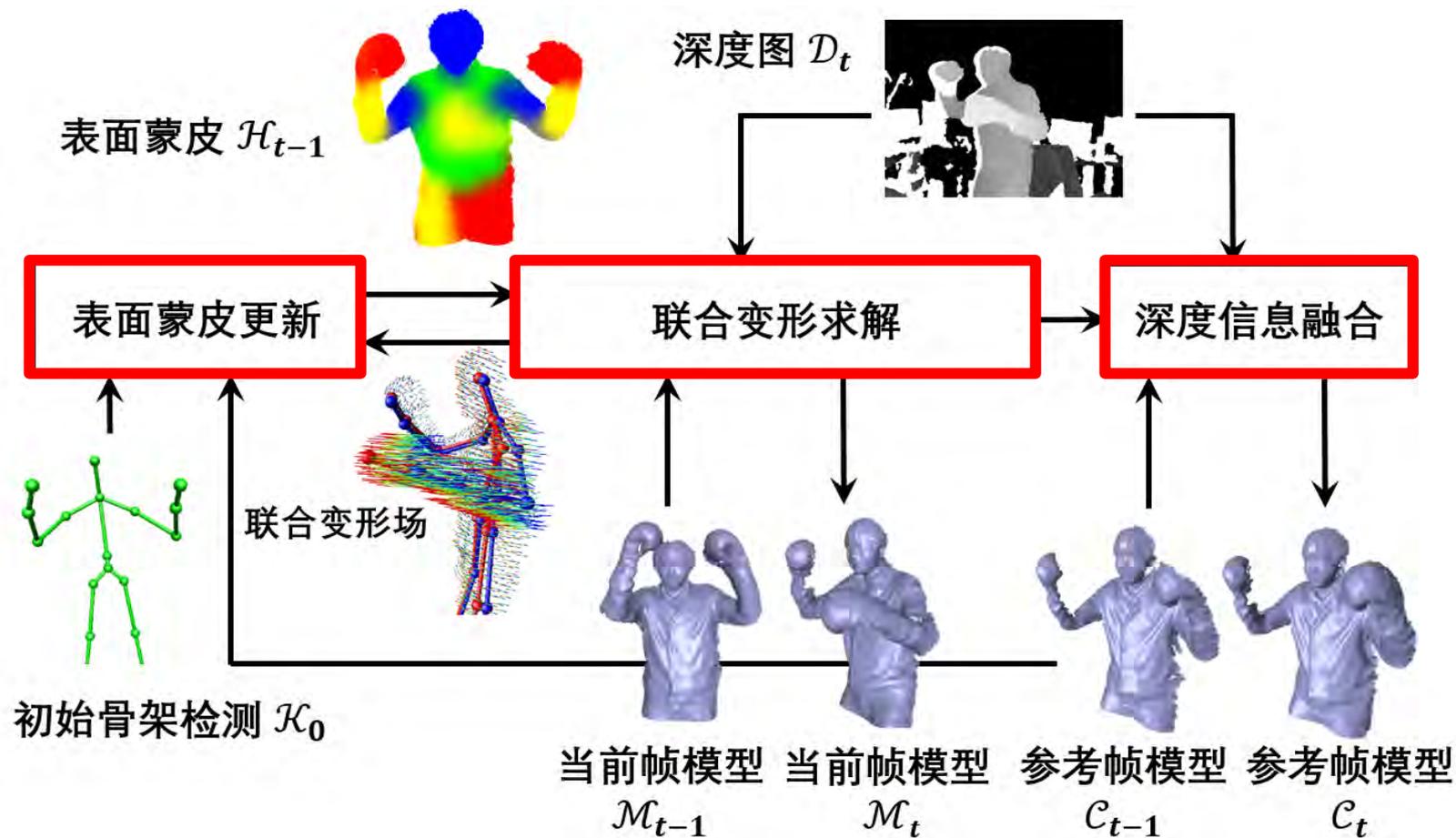
仅能处理慢速运动



无法处理快速运动

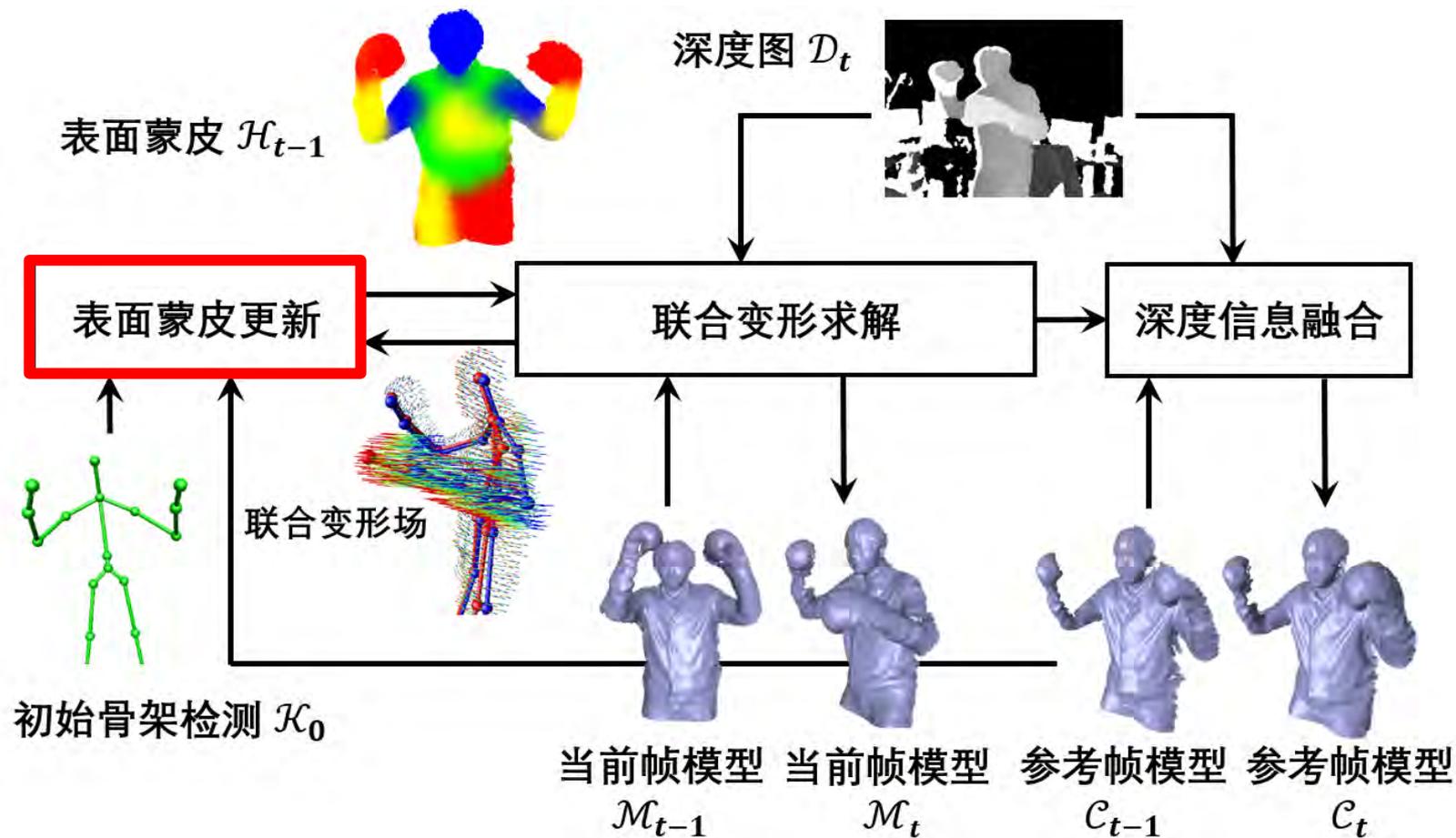
2 基于骨架跟踪的实时人体动态三维重建

2.3 BodyFusion算法框架



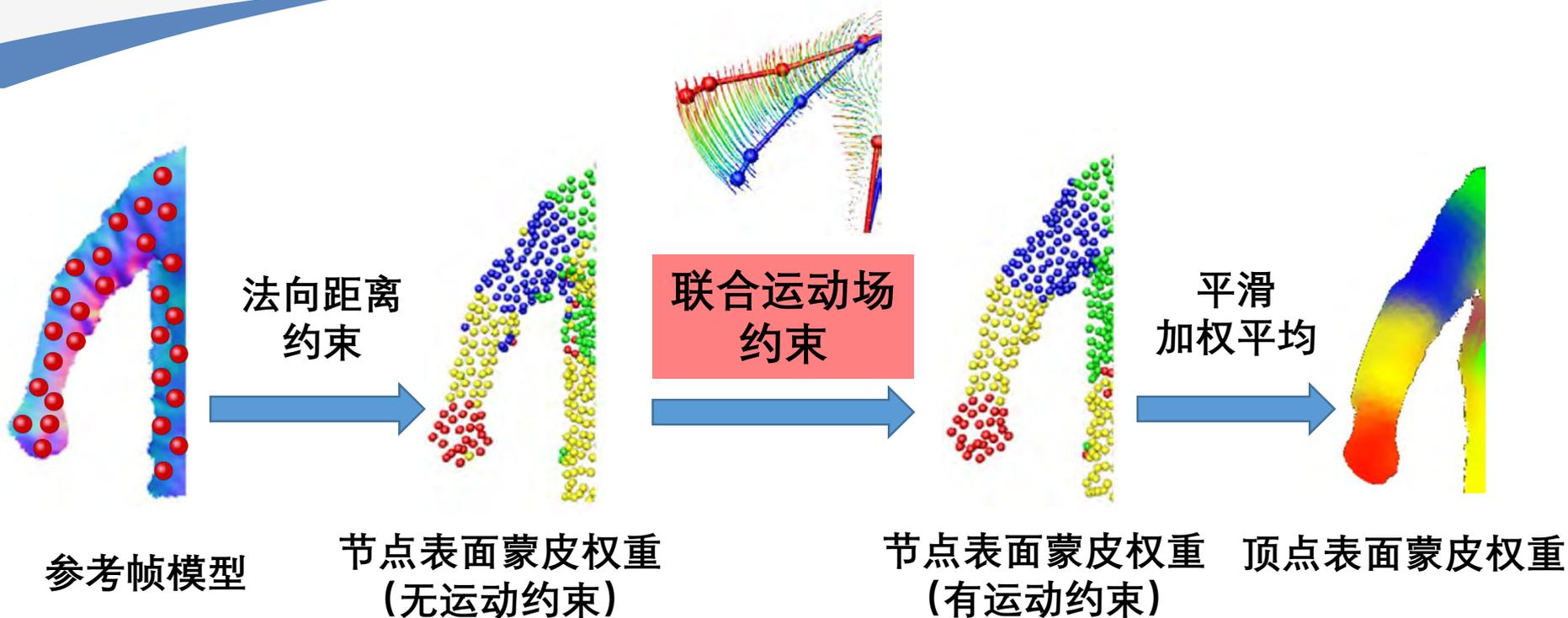
2 基于骨架跟踪的实时人体动态三维重建

2.4 表面蒙皮更新



2 基于骨架跟踪的实时人体动态三维重建

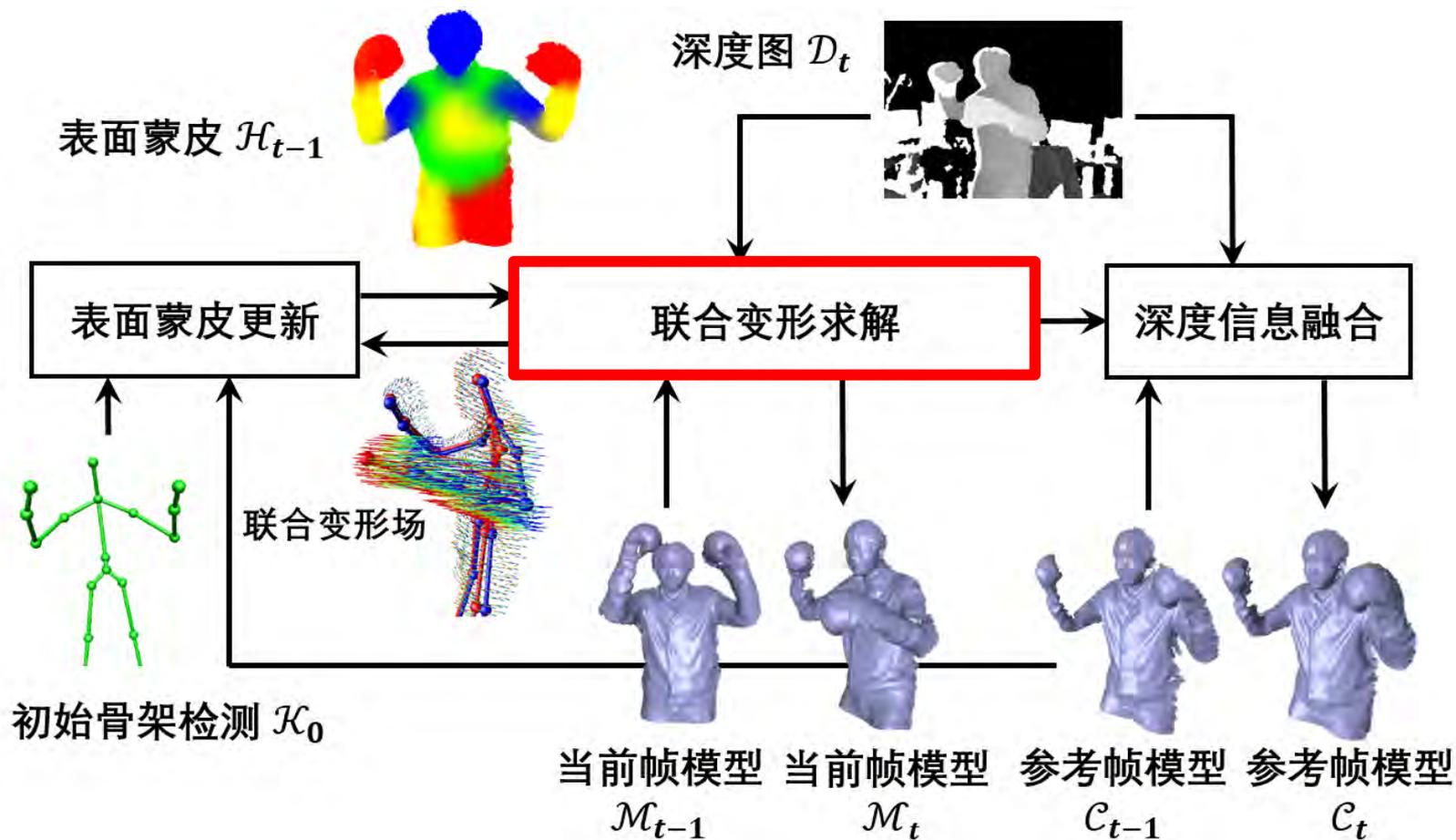
2.4 表面蒙皮更新



使用骨架-非刚性联合运动场约束高效蒙皮更新，单帧计算速度0.6ms

2 基于骨架跟踪的实时人体动态三维重建

2.5 联合变形求解



2 基于骨架跟踪的实时人体动态三维重建

2.5 联合变形求解

联合变形求解能量函数定义

$$E_t = \lambda_n E_{\text{nonrigid}} + \lambda_s E_{\text{skeleton}} + \lambda_g E_{\text{graph}} + \lambda_b E_{\text{binding}}$$

非刚性表面运动跟踪数据能量项
和骨架跟踪数据能量项

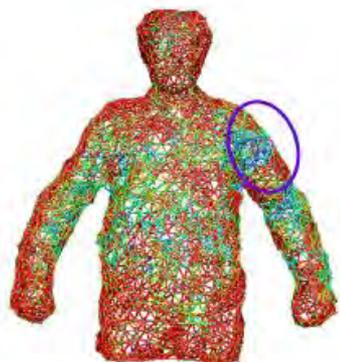
嵌入式变形图模型
正则项（基于表面蒙皮权重）

骨架运动和非刚性表面运动
绑定能量项

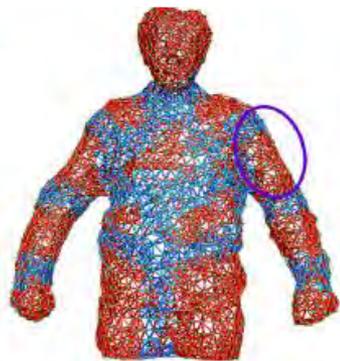
2 基于骨架跟踪的实时人体动态三维重建

2.5 联合变形求解

正则项对比论证



传统方法正则项



本文正则项
(使用表面蒙皮约束)



Using Kinect v1



使用本文正则项



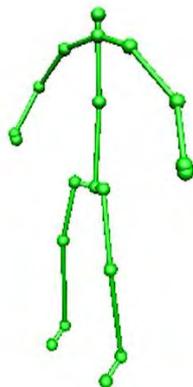
使用传统方法正则项

2 基于骨架跟踪的实时人体动态三维重建

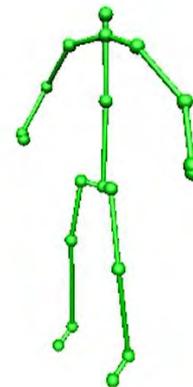
2.5 联合变形求解

绑定能量项对比论证

Using Kinect v2



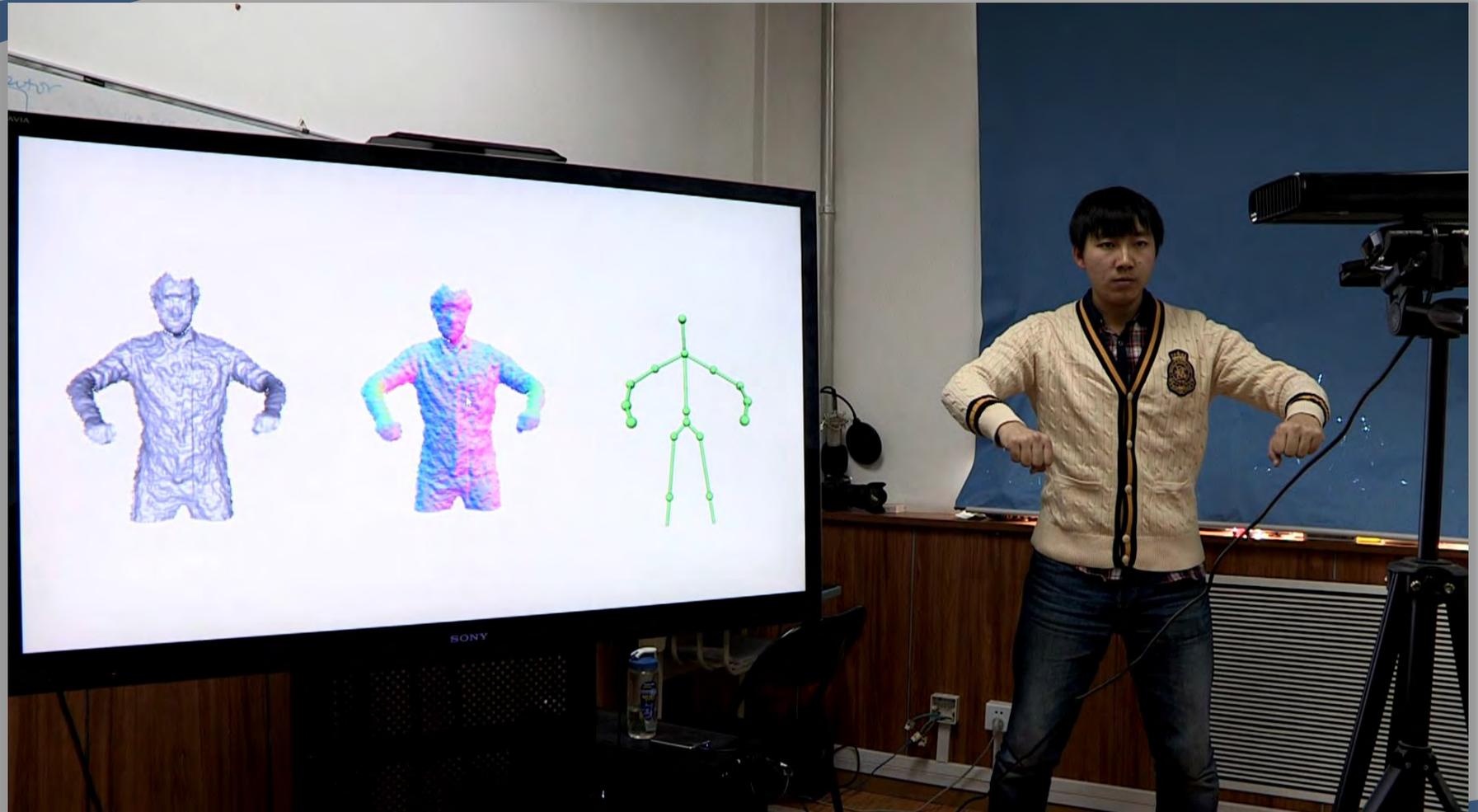
**本文方法BodyFusion
(有绑定能量项)**



**传统方法
(无绑定能量项)**

2 基于骨架跟踪的实时人体动态三维重建

2.6 BodyFusion实时重建结果



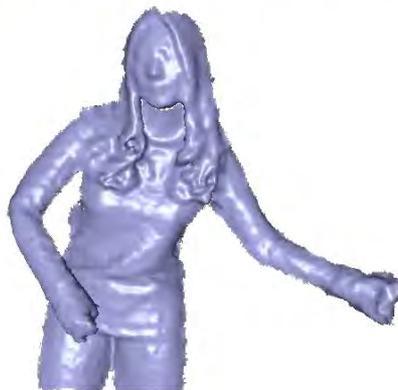
2 基于骨架跟踪的实时人体动态三维重建

2.6 与现有方法对比

Using Kinect v1



本文方法
BodyFusion



DynamicFusion
CVPR15 Best Paper

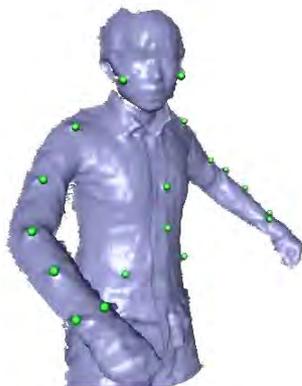
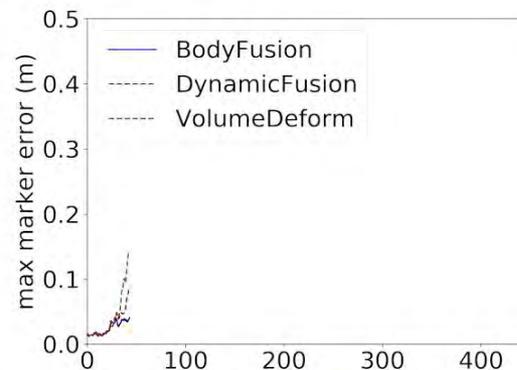


VolumeDeform
ECCV 2016

2 基于骨架跟踪的实时人体动态三维重建

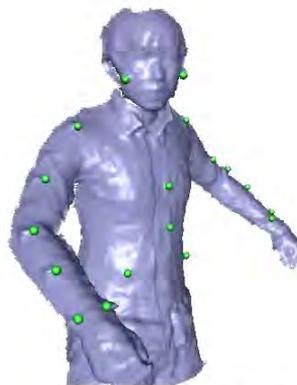
2.7 光学标记点-深度相机联合采集数据集

Using Kinect v1



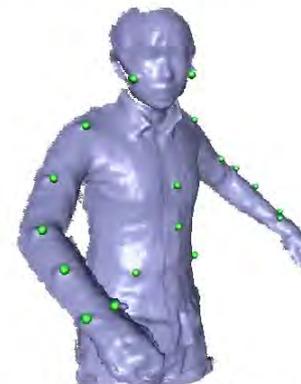
本文方法

BodyFusion(ICC17)



DynamicFusion

CVPR15 Best Paper



VolumeDeform

ECCV 2016

报告内容

- 1 研究背景及意义
- 2 基于骨架跟踪的实时人体动态三维重建
- 3 实时人体动态三维信息双层语义化重建**
- 4 高精度单视角人体快速动作捕捉算法
- 5 实时人体动态三维信息多层语义化重建
- 6 单图像人体三维重建
- 7 总结与展望

3 DoubleFusion 实时人体动态三维信息双层语义化重建

3.1 难点与思路

难点： 现有实时单视点人体动态三维重建方法

无法实时重建人体内部体形、无法处理人体快速运动

思路： 人体内部体形是人体三维信息中非常强的一种先验信息



单视点实时动态三维重建
可重建稠密非刚性形状

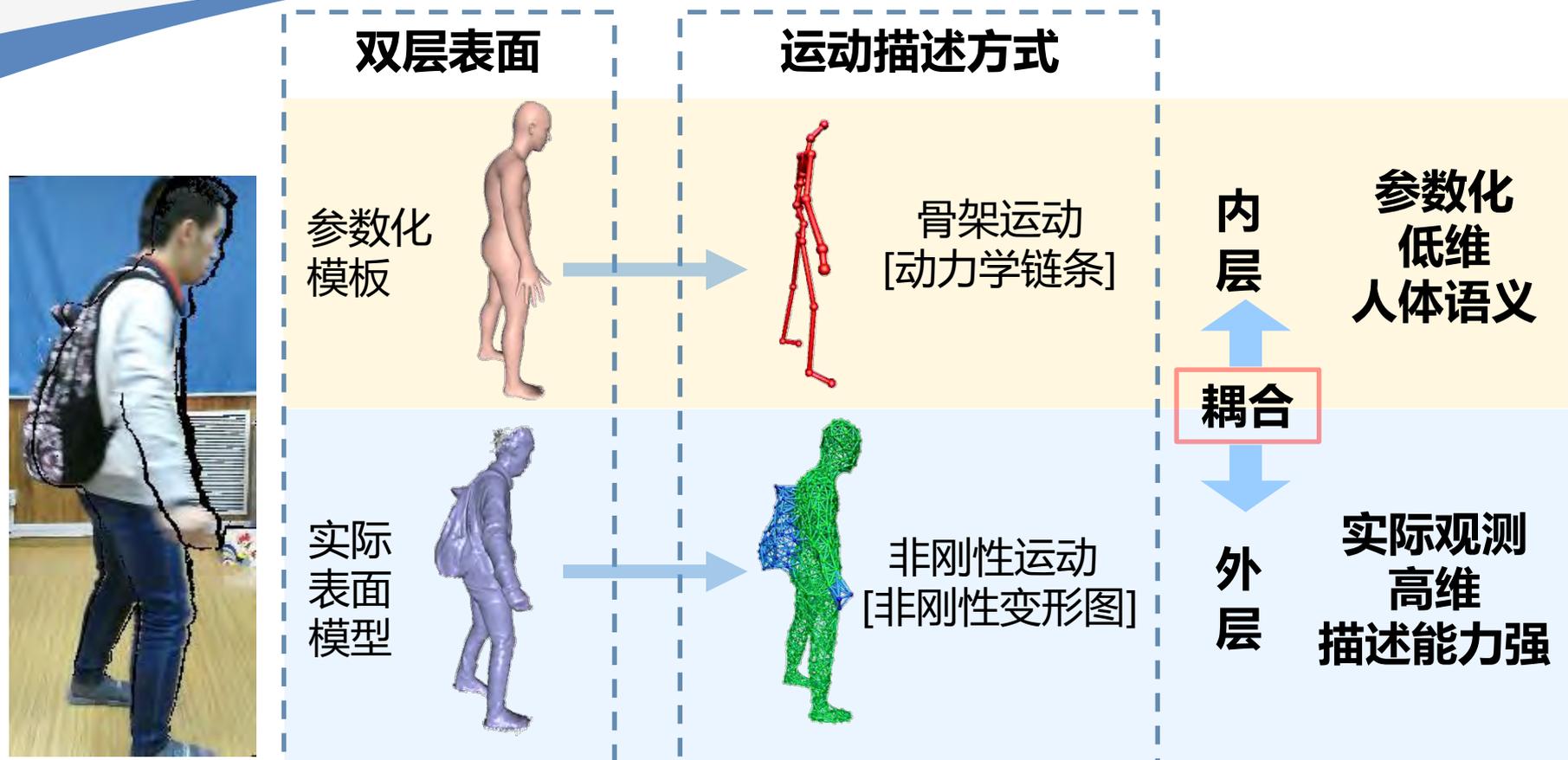


参数化人体体形
可实现参数化鲁棒重建

DoubleFusion

3 实时人体动态三维信息双层语义化重建

3.2 双层人体三维信息表示方法

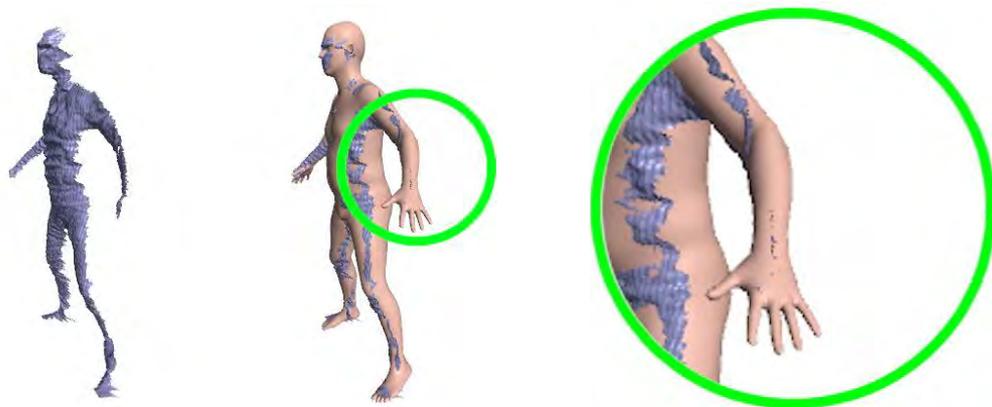


双层表面实现**高效时空全局映射**，提高跟踪和重建精度

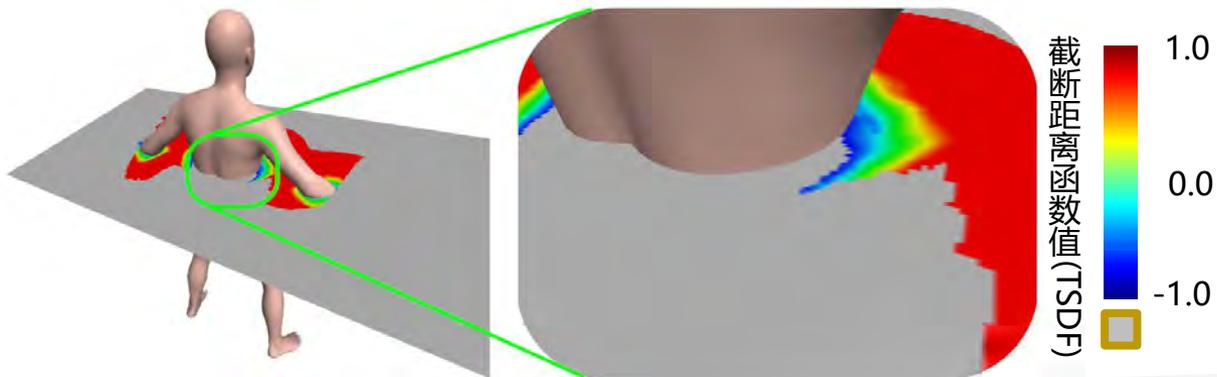
3 实时人体动态三维信息双层语义化重建

3.3 双层表面融合重建方法

内层参数化模板提高了运动跟踪及表面融合精度

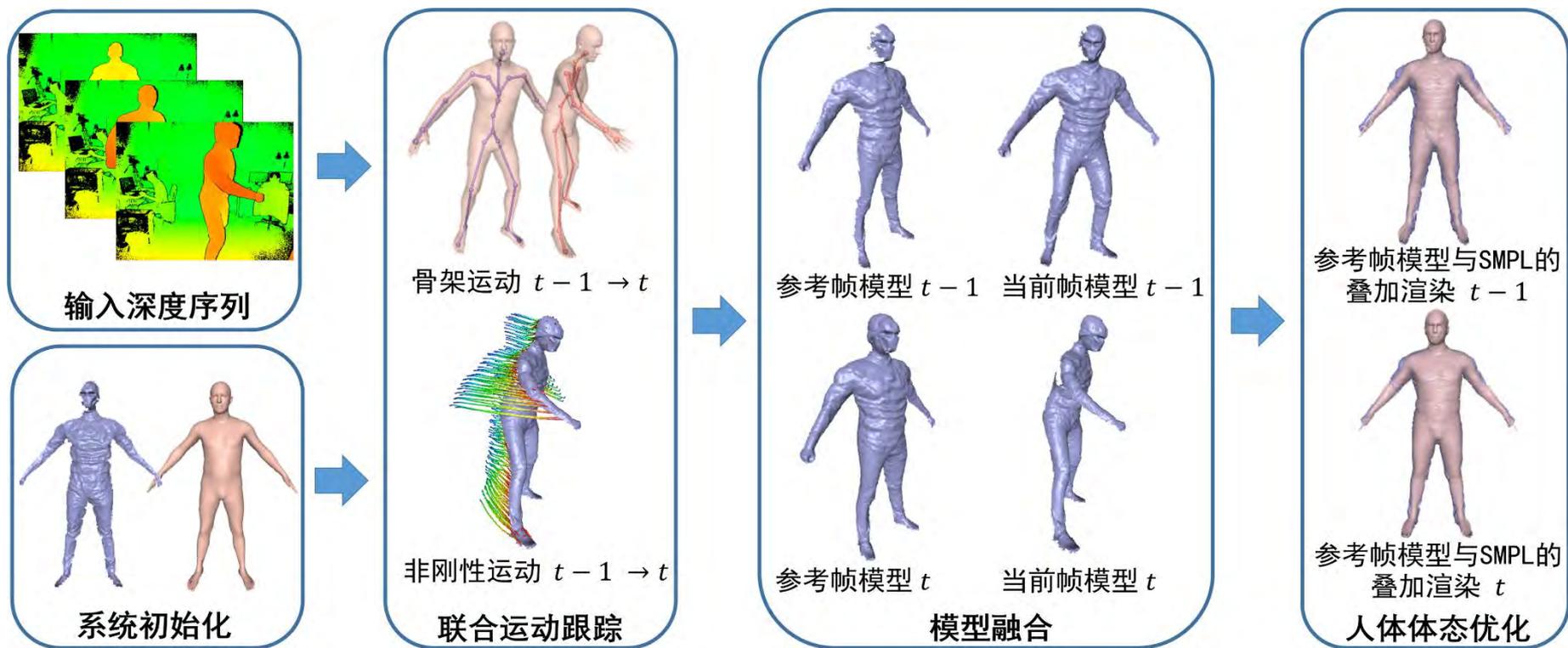


外层融合表面提升了内层参数化人体估计精度



3 实时人体动态三维信息双层语义化重建

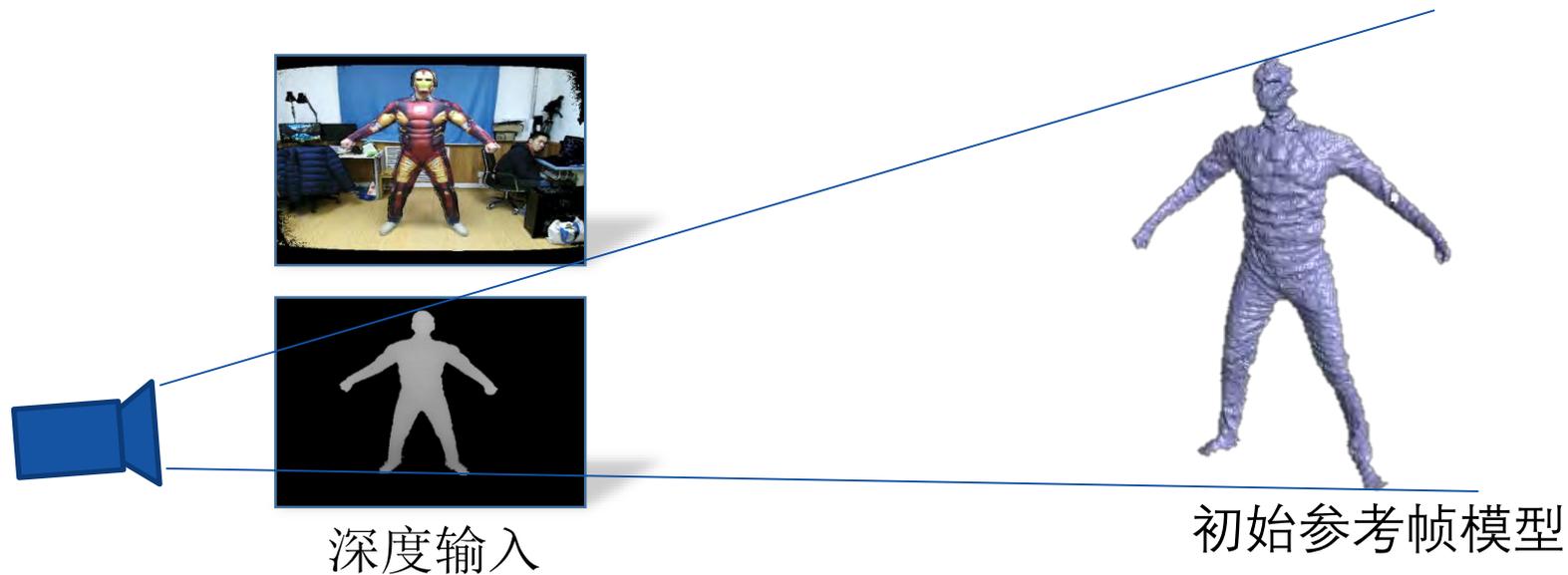
3.4 算法框图



3 实时人体动态三维信息双层语义化重建

3.5 系统初始化

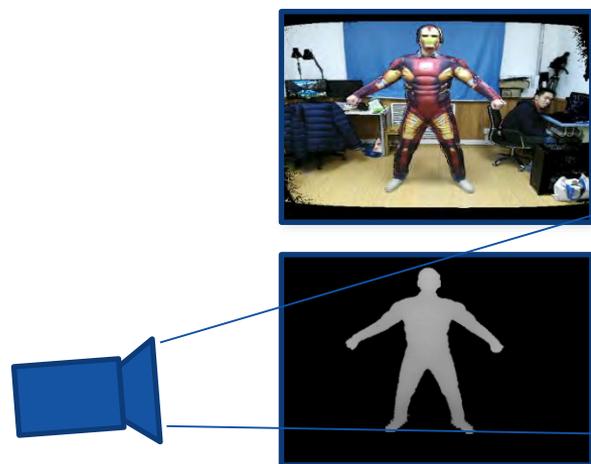
- 初始化符号距离函数



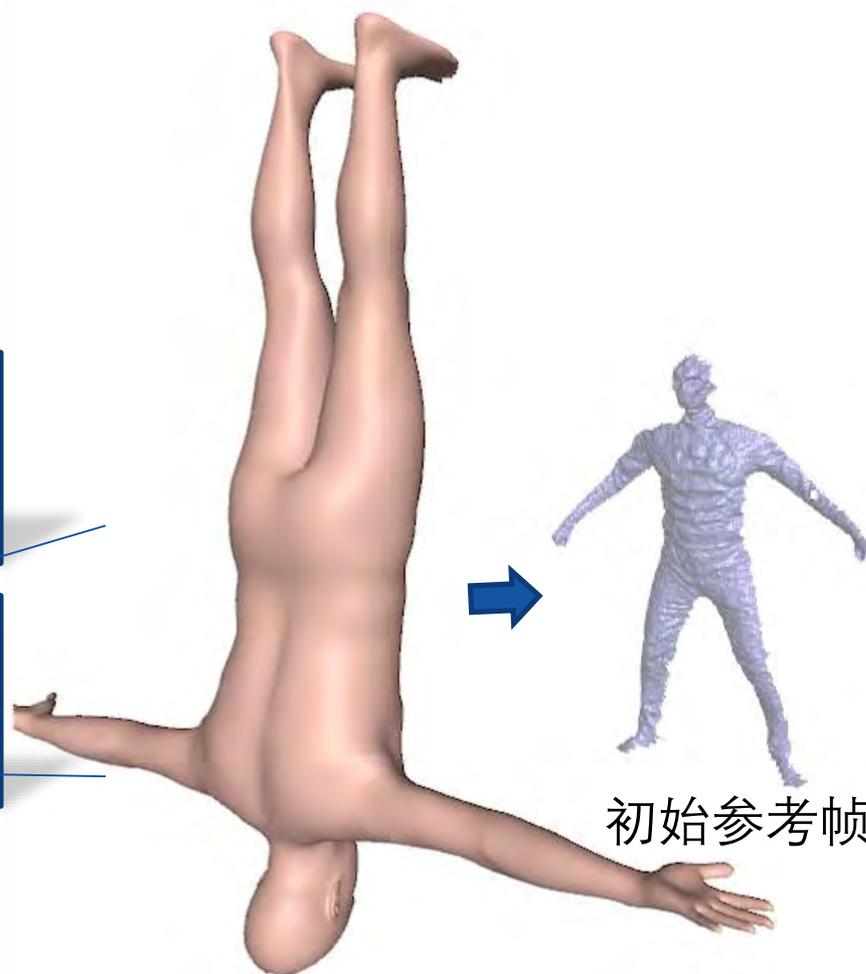
3 实时人体动态三维信息双层语义化重建

3.5 系统初始化

- 初始化符号距离函数
- 初始化人体参数化体形



深度输入



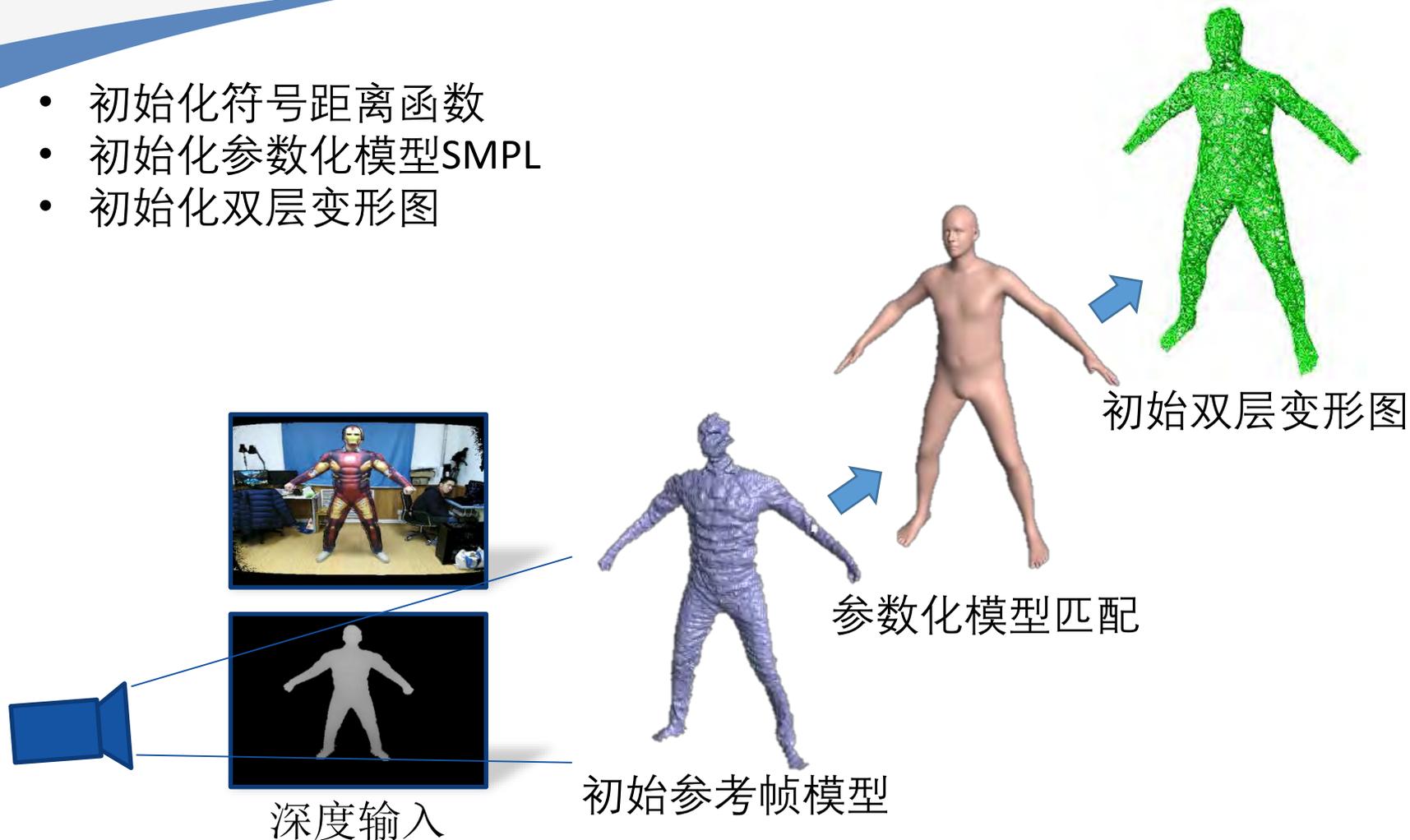
初始参考帧模型

参数化模型初始化过程

3 实时人体动态三维信息双层语义化重建

3.5 系统初始化

- 初始化符号距离函数
- 初始化参数化模型SMPL
- 初始化双层变形图



3 实时人体动态三维信息双层语义化重建

3.6 联合运动跟踪 – 能量函数

$$E_{mot} = \lambda_{data} E_{data} + \lambda_{bind} E_{bind} + \lambda_{reg} E_{reg} + \lambda_{prior} E_{prior}$$

双层人体表面模型



输入深度观测

骨架运动



非刚性运动

非刚性运动



局部刚性假设

骨架运动



人体自然运动分布

双层人体表面与深度观测构造时空全局映射，提高了跟踪重建精度

3 实时人体动态三维信息双层语义化重建

3.7 参考帧模型融合

深度观测序列



参考帧模型融合过程

右侧视角

正面视角

左侧视角



3 实时人体动态三维信息双层语义化重建

3.8 基于符号距离函数的人体体态优化

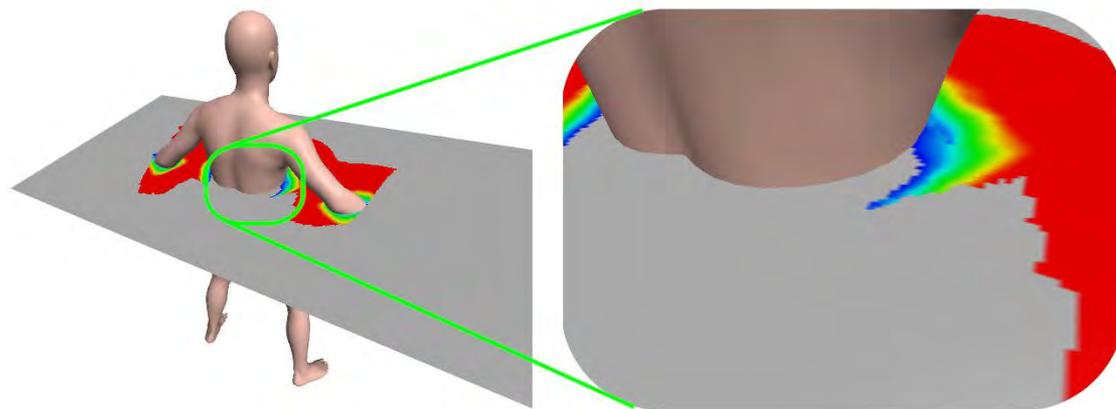
$$E_{shape} = \lambda_{data} E_{sdata} + \lambda_{sreg} E_{sreg} + \lambda_{pri} E_{pri}$$

避免显式搜寻对应点，实现了人体体态优化的实时求解，求解时间3ms

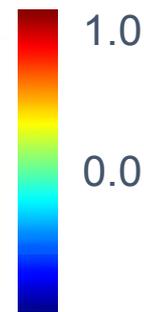
参数化模型
↓ 匹配
符号距离函数

人体体态参数
↓ 正则
时域平滑先验

人体姿态参数
↓ 正则
自然运动先验



符号距离函数值



1.0 真实三维表面区域

-1.0 尚未观测区域

3 实时人体动态三维信息双层语义化重建

3.9 DoubleFusion实时重建结果



模型实时更新结果

3 实时人体动态三维信息双层语义化重建

3.9 重建结果 - 单视点采集自由视点渲染



Iron Man

Kicking

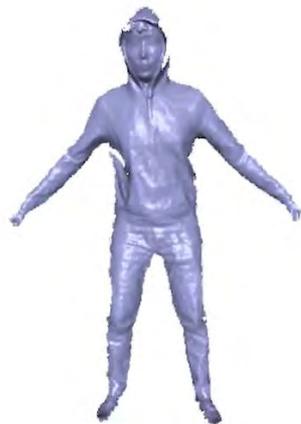
3 实时人体动态三维信息双层语义化重建

3.9 与现有方法对比

对比序列1



本文方法
DoubleFusion



DynamicFusion

对比序列2



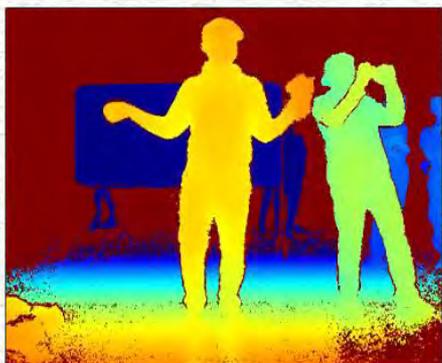
本文方法
DoubleFusion



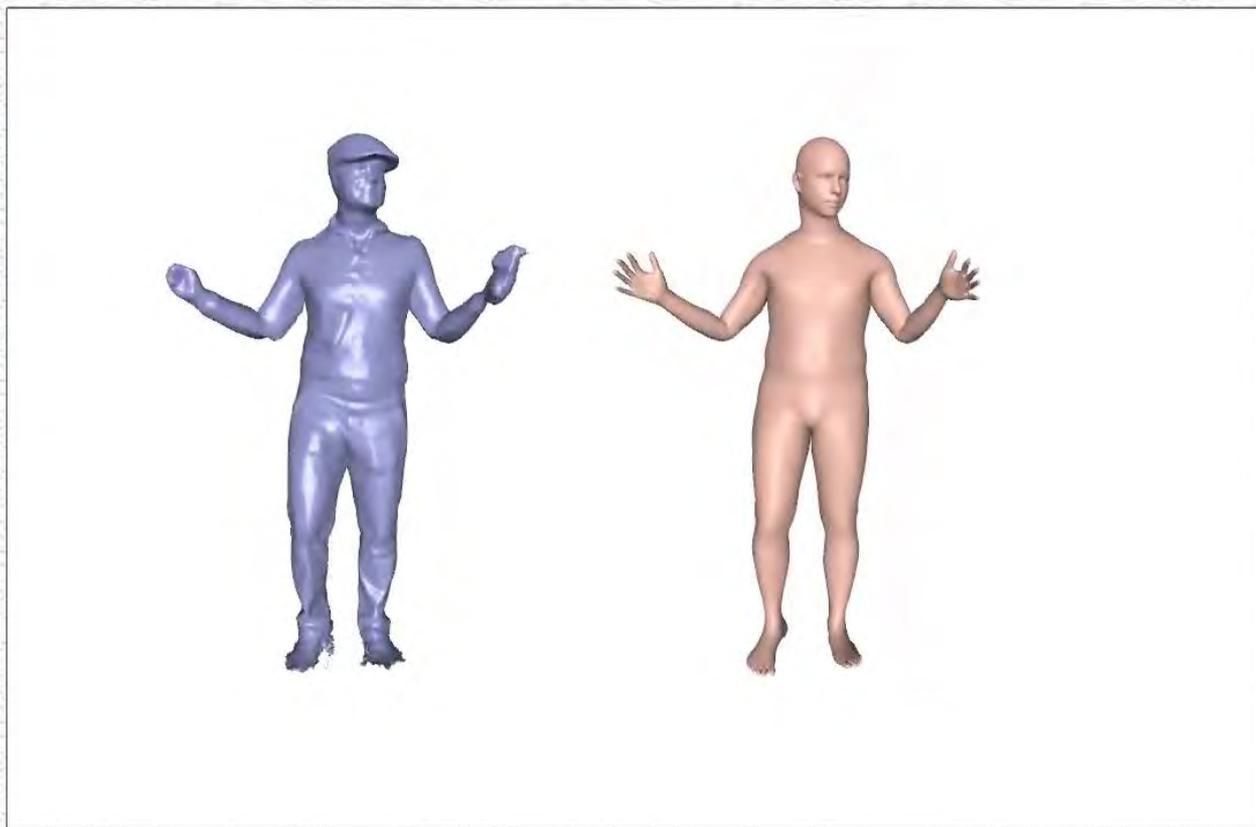
DynamicFusion

3 实时人体动态三维信息双层语义化重建

3.10 CVPR 2018 现场 Live Demo



o color image



3 DoubleFusion 实时人体动态三维信息双层语义化重建

3.11 DoubleFusion软件

DoubleFusionSoftwareV1.4 is available

http://www.liuyebin.com/doublefusion/doublefusion_software.htm

please email:

liuyebin@mail.tsinghua.edu.cn

ytrack@126.com

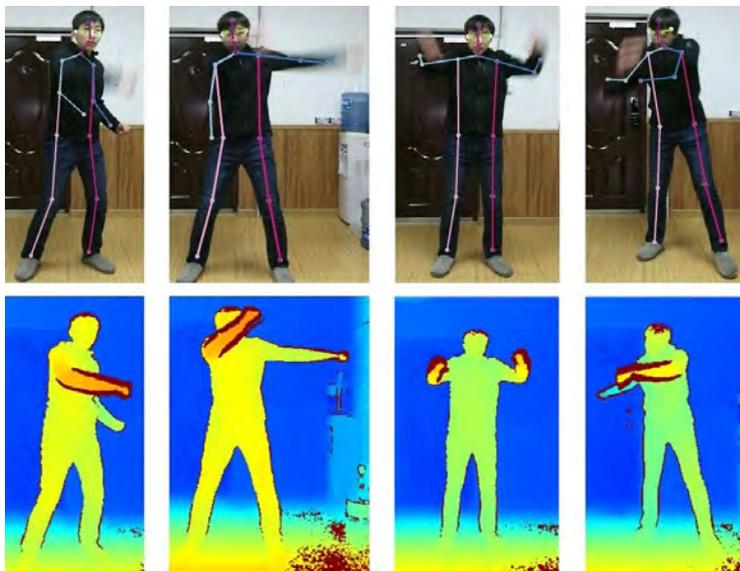
报告内容

- 1 研究背景及意义
- 2 基于骨架跟踪的实时人体动态三维重建
- 3 实时人体动态三维信息双层语义化重建
- 4 高精度单视角人体快速动作捕捉算法**
- 5 实时人体动态三维信息多层语义化重建
- 6 单图像人体三维重建
- 7 总结与展望

4 HybTrack 高精度单视角人体快速动作捕捉算法

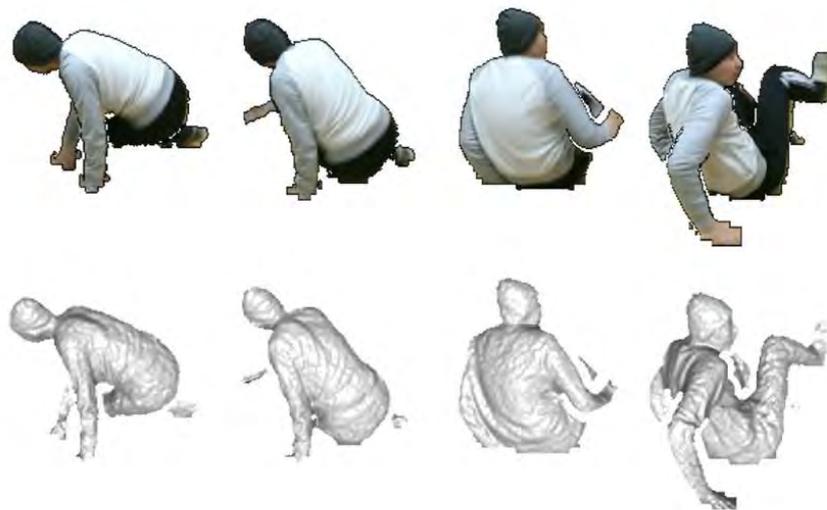
4.1 难点与思路

快速运动对相机成像质量影响较大



提出了基于语义的混合运动捕捉算法
使用姿态检测器抑制跟踪误差累积

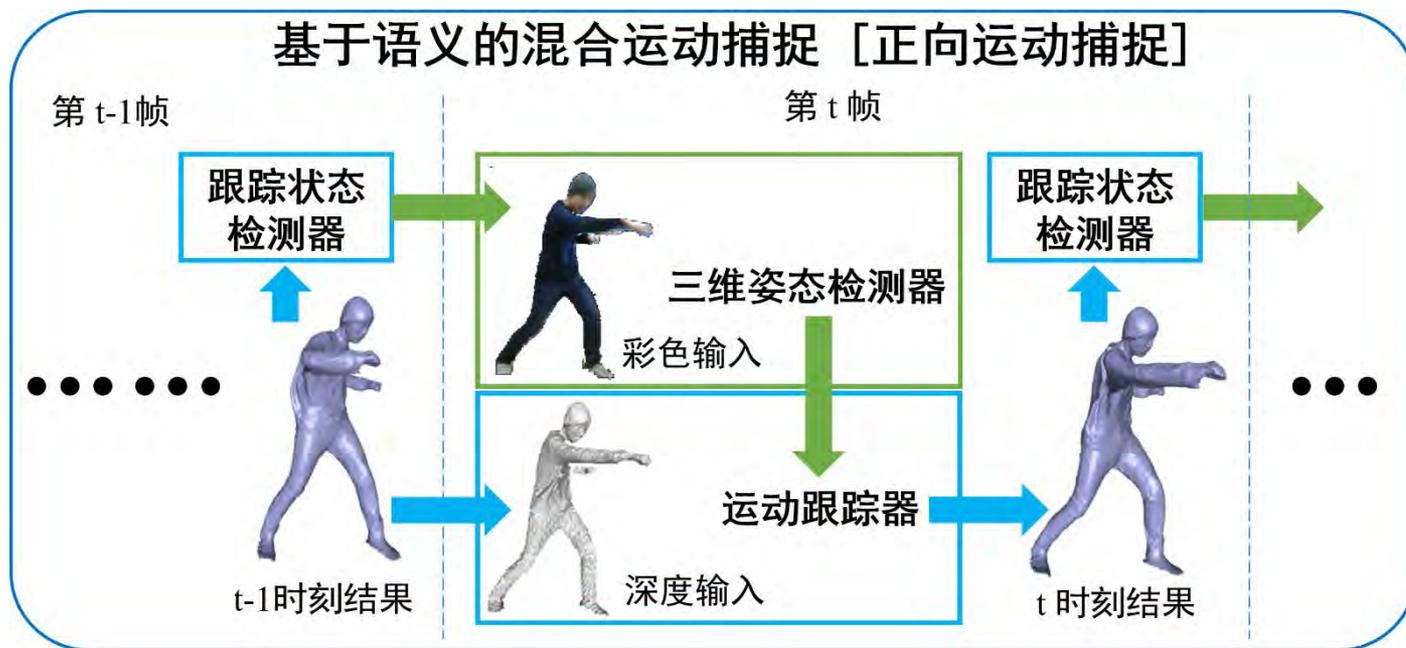
单视角人体运动自遮挡严重



提出了基于语义的双向运动混合算法
进行严重遮挡区域运动的精确鲁棒估计

4 高精度单视角人体快速动作捕捉算法

4.2 基于语义的混合运动捕捉

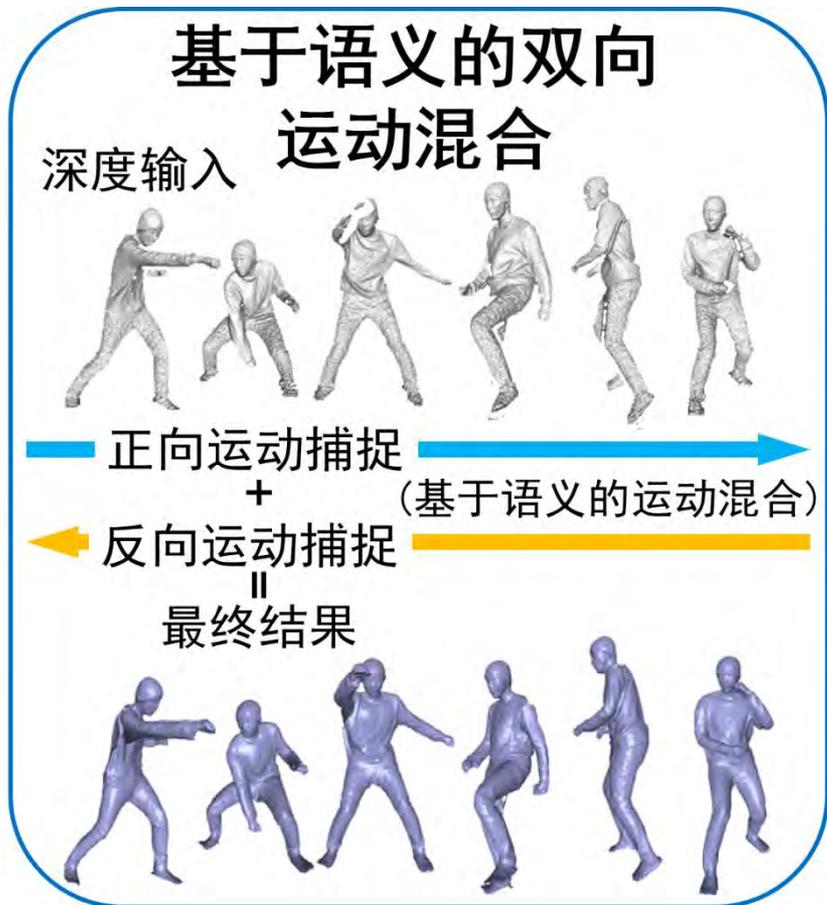


$$E_{track} = \lambda_{data} E_{data} + \lambda_{bind} E_{bind} + \lambda_{reg} E_{reg} + \lambda_{prior} E_{prior} + \lambda_{detect} E_{detect}$$

$$E_{detect} = \|\mathcal{M}_b^T \cdot (\theta_d - \theta_t)\|_2^2$$

4 高精度单视角人体快速动作捕捉算法

4.3 基于语义的双向运动混合



目标：在复杂自遮挡的情况下实现鲁棒的动作捕捉性能，并且对于完全被遮挡的身体部位给出合理的运动估计结果

观察：反向运动跟踪可以在采集对象被遮挡部分突然出现的情况下产生更为稳定的动作捕捉效果

基于语义的双向运动混合公式：

$$\theta_{final_j} = F \left(\text{Slerp} \left(\frac{L_{i_b}}{L_{i_f} + L_{i_b}} Q(\theta_{f_j}), \frac{L_{i_f}}{L_{i_f} + L_{i_b}} Q(\theta_{b_j}) \right) \right)$$

Slerp() 表示球面线性插值

4 高精度单视角人体快速动作捕捉算法

4.4 基于语义的双向运动混合算法数值论证

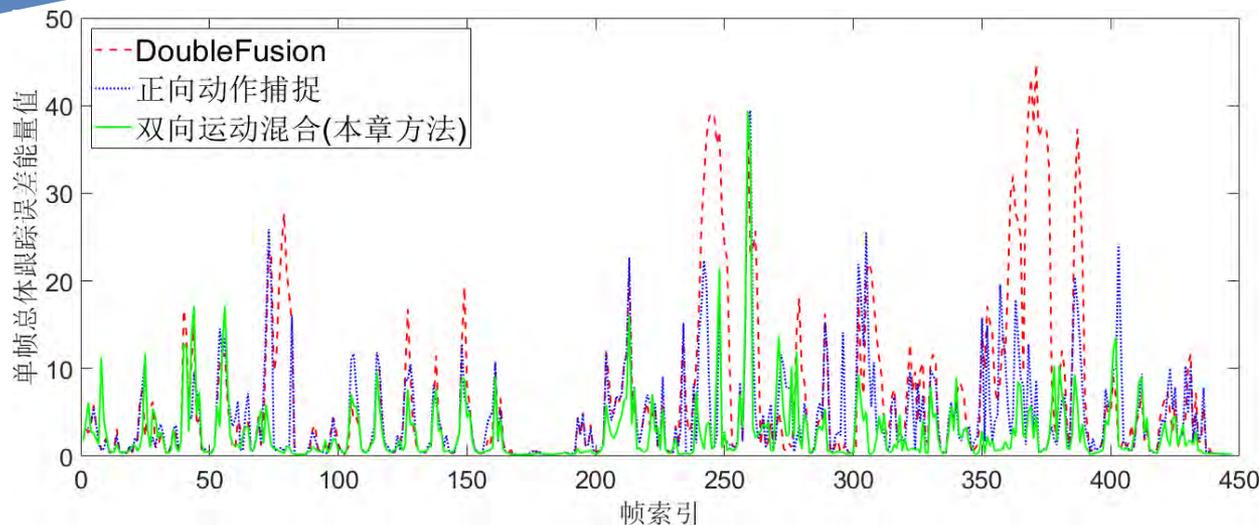


表 6 序列全局平均跟踪误差

	Fast Motion	Kicking	Basketball	Hand Stand
DoubleFusion	6.7357	3.2065	3.5232	6.2772
正向动作捕捉	4.4647	2.1403	2.3201	3.9986
双向运动混合	2.7119	1.3798	1.794	3.0436

4 高精度单视角人体快速动作捕捉算法

4.5 重建结果 – 单视角复杂自遮挡

Challenging Motion with Occlusion



深度观测

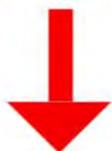


重建结果

4 高精度单视角人体快速动作捕捉算法

4.5 重建结果 – 快速运动及复杂自遮挡

Fast Motion



彩色图像



深度观测



重建结果

4 高精度单视角人体快速动作捕捉算法

4.5 重建结果 – 快速运动及复杂自遮挡

Challenging Motion with Occlusion



彩色图像



深度观测



重建结果

4 高精度单视角人体快速动作捕捉算法

4.5 与现有方法对比



深度观测



稠密表面重建
DoubleFusion



稠密表面重建
本文结果



深度&人体姿态
本文结果



深度&人体姿态
HMR

4 高精度单视角人体快速动作捕捉算法

4.5 与现有方法对比



深度观测



稠密表面重建
DoubleFusion



**稠密表面重建
本文结果**



**深度&人体姿态
本文结果**



深度&人体姿态
HMR

报告内容

- 1 研究背景及意义
- 2 基于骨架跟踪的实时人体动态三维重建
- 3 实时人体动态三维信息双层语义化重建
- 4 高精度单视角人体快速动作捕捉算法
- 5 实时人体动态三维信息多层语义化重建
- 6 单图像人体三维重建
- 7 总结与展望

5 SimulCap 实时人体动态三维信息多层语义化重建

5.1 难点与思路

难点：现有方法重建过程中不区分人体和衣物，使用统一框架对二者进行重建，
因此：现有方法无法在单视角被遮挡区域重建出有效衣物动态细节



正面观测（输入）

背面观测（参考）

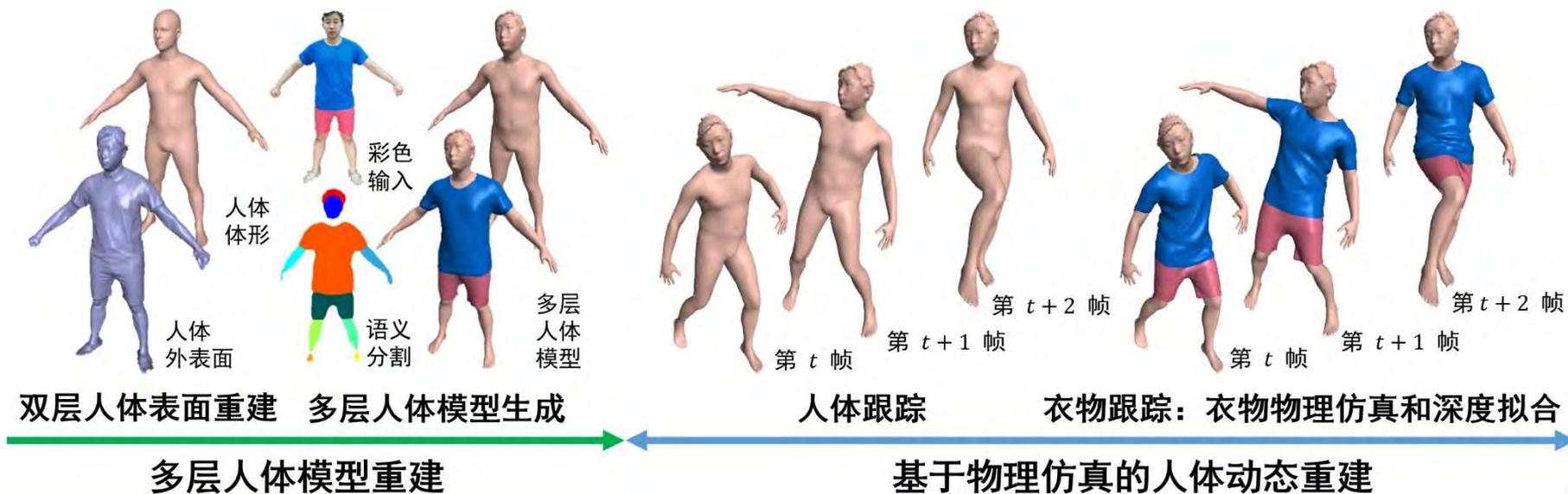
DoubleFusion

多层非刚性跟踪

思路：使用衣物物理仿真模型代替原有非刚性变形模型，在被遮挡区域产生动态效果
实现衣物动态三维的单视点采集->全视点动态重建

5 实时人体动态三维信息多层语义化重建

5.2 算法框架



5 实时人体动态三维信息多层语义化重建

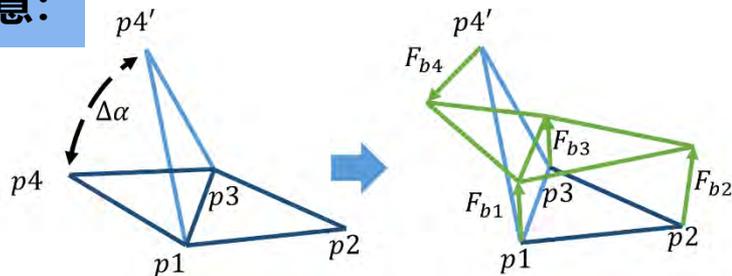
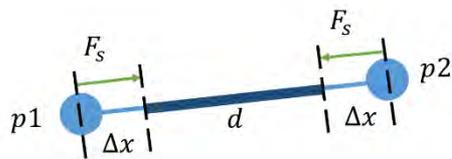
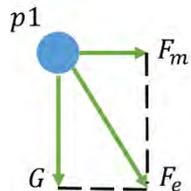
5.3 基于物理仿真的人体动态重建

- 人体姿态跟踪（构造穿透能量项防止衣物与人体相互穿透）

$$E_{skel} = \lambda_{data} E_{data} + \lambda_{inter} E_{inter} + \lambda_{pri} E_{pri}$$

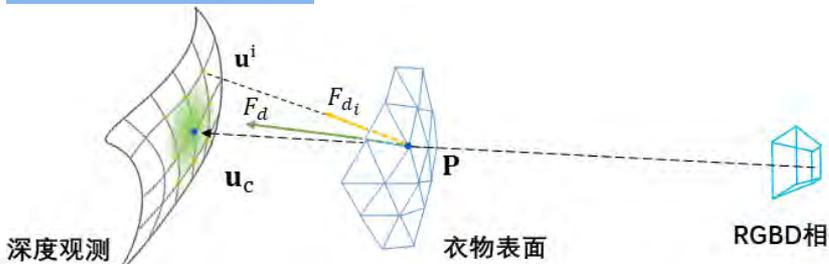
- 衣物跟踪 = 衣物物理仿真（质点-弹簧系统） + 迭代深度拟合

衣物顶点所受外力合力、拉伸力和剪切力示意：



■ 初始状态
■ 当前状态
→ 力
■ 更新状态

深度拟合外力：



3D边界平滑权重



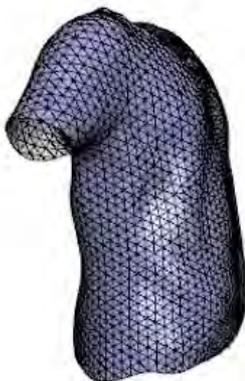
5 实时人体动态三维信息多层语义化重建

5.3 迭代深度拟合论证

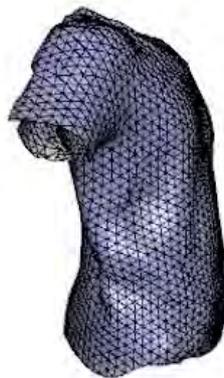
迭代深度拟合可以保证物理仿真框架的求解稳定性
3D平滑过渡可以在可见区域边界处产生平滑的动态重建结果



深度观测
(输入)



迭代深度拟合
+
3D平滑过渡



仅迭代深度拟合



直接深度匹配

5 实时人体动态三维信息多层语义化重建

5.3 3D平滑加权论证

迭代深度拟合使重建结果与深度观测更为匹配，提高了重建精度



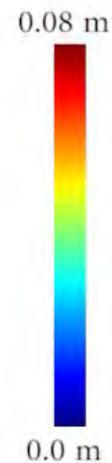
深度观测
(输入)



使用
迭代深度拟合



未使用
迭代深度拟合



5 实时人体动态三维信息多层语义化重建

5.4 实时重建结果演示



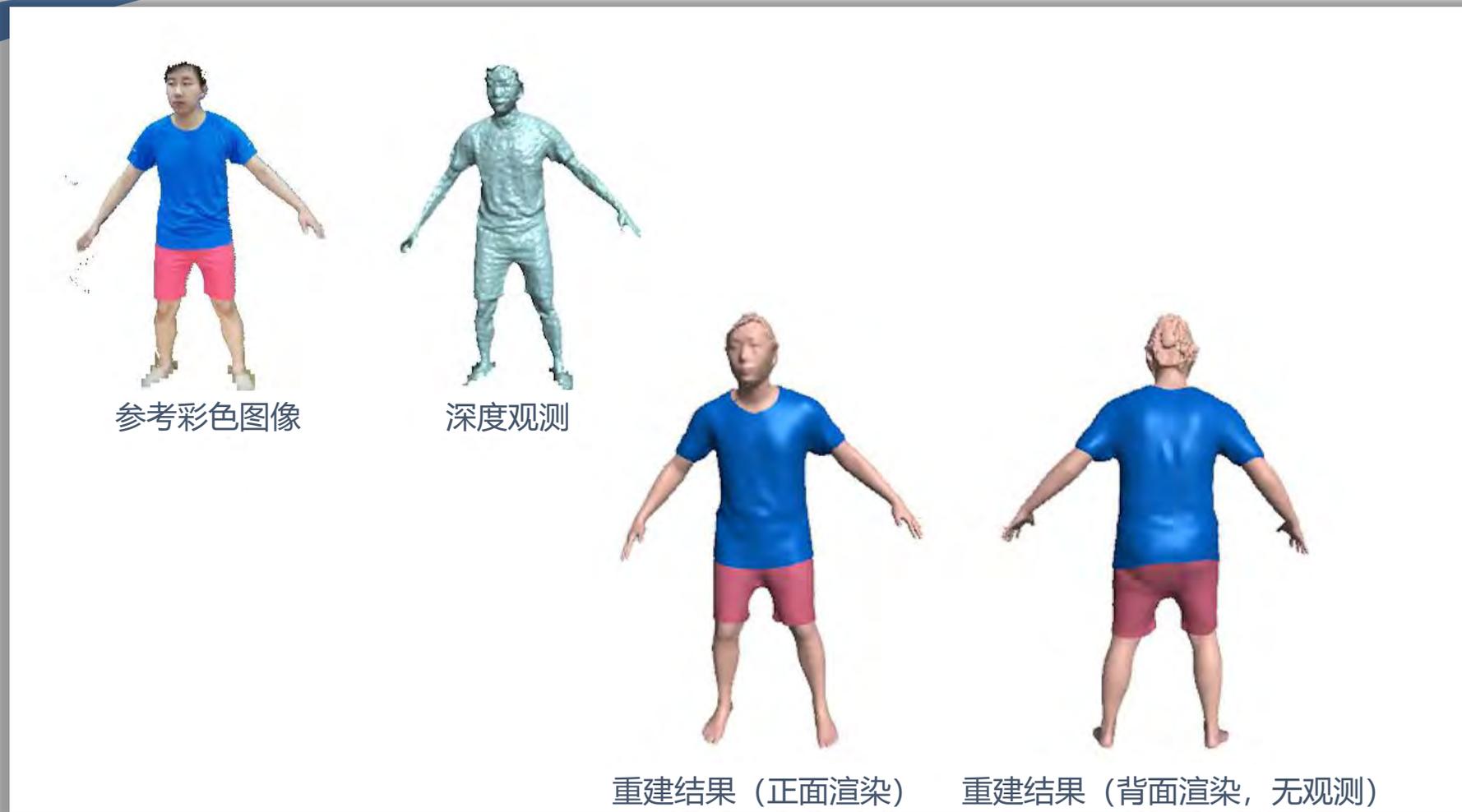
1. 双层人体表面重建

2. 多层人体表面分割及重建

3. 物理驱动的多层人体动态重建

5 实时人体动态三维信息多层语义化重建

5.4 重建结果



5 实时人体动态三维信息多层语义化重建

5.4 重建结果



参考彩色图像



深度观测



重建结果 (正面渲染)



重建结果 (背面渲染, 无观测)

5 实时人体动态三维信息多层语义化重建

5.4 重建结果



参考彩色图像



深度观测



重建结果（正面渲染）



重建结果（背面渲染，无观测）

5 实时人体动态三维信息多层语义化重建

5.4 与现有方法对比

相对于现有方法，基于物理驱动的多层人体三维语义化重建算法可以重建出：

- 衣物局部的复杂形变
- 衣物与身体的交互运动
- 衣物的动态细节



深度观测
(输入)



本文方法
SimulCap



DoubleFusion 多层非刚性跟踪



5 实时人体动态三维信息多层语义化重建

5.6 应用 – 动态衣物重定向



目标体形1



源重建结果



目标体形2

5 实时人体动态三维信息多层语义化重建

5.6 SimulCap 局限性分析

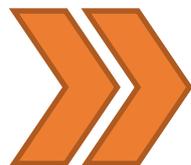
- 仍需手动指定布料材质进行仿真，在单视角下估计布料物理材质仍是十分具有挑战性的问题；
- SimulCap中布料仿真的真实性仍有待提升，尤其是针对多层衣物、较厚的针织衣物以及具有特定裁剪模式的衣物，需要实现更加先进的布料仿真器进行仿真；
- 现有方法没有考虑人体自身的软组织运动(Soft tissue motion)，因此人体重建的真实性仍有待提高；
- 无法处理比较复杂的衣物-人体交互（比如手抓衣物、穿脱衣物等动作）。

报告内容

- 1 研究背景及意义
- 2 基于骨架跟踪的实时人体动态三维重建
- 3 实时人体动态三维信息双层语义化重建
- 4 高精度单视角人体快速动作捕捉算法
- 5 实时人体动态三维信息多层语义化重建
- 6 单图像人体三维重建
- 7 总结与展望

6 DeepHuman 单图像人体三维重建

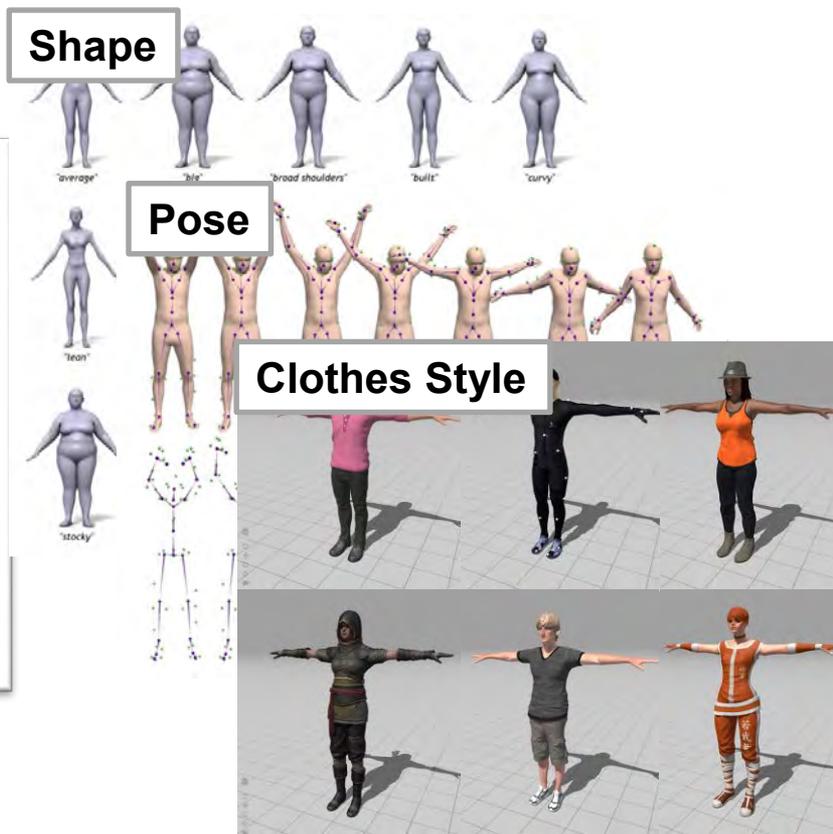
6.1 目标与挑战



**Clothed Body
Reconstruction**
真实稠密形状

6 单图像人体三维重建 DeepHuman

6.1 目标与挑战



6 单图像人体三维重建 DeepHuman

6.1 目标与挑战

单图像3D人体重建

处理不同姿态和衣物

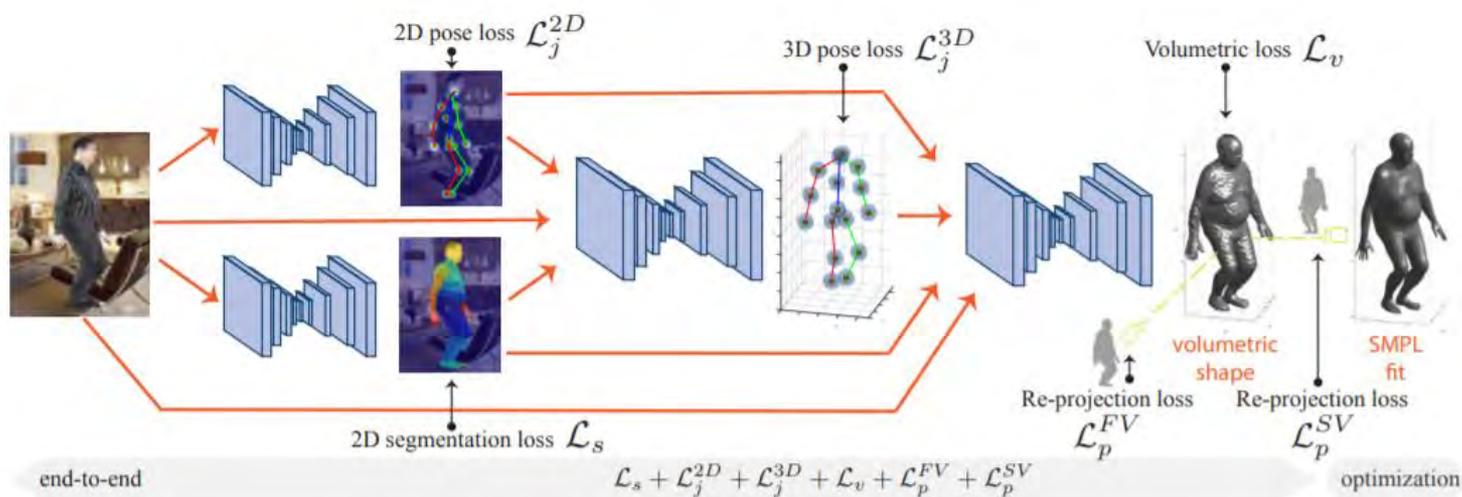
重建真实表面细节

6 单图像人体三维重建 DeepHuman

6.2 相关工作

- **BodyNet (ECCV'18)**

- 缺少表面细节
- 无法保证肢体完整性

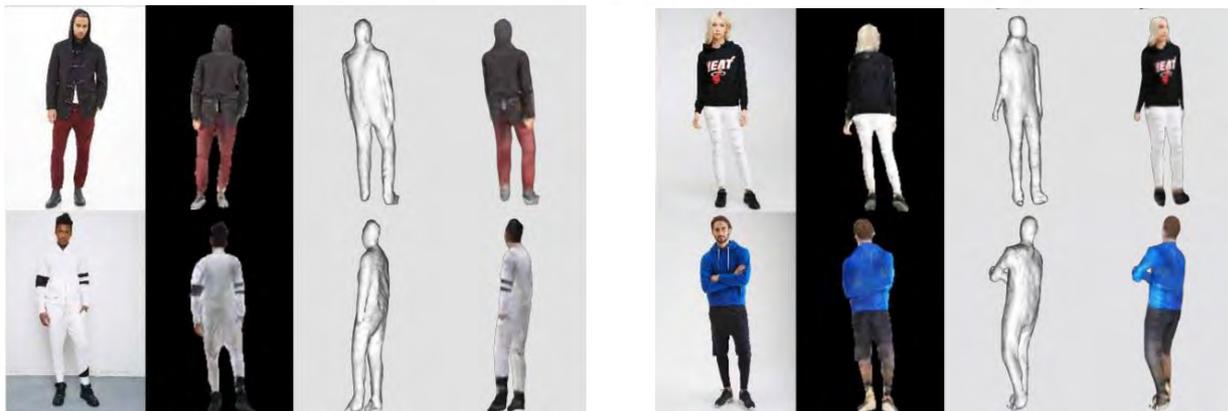
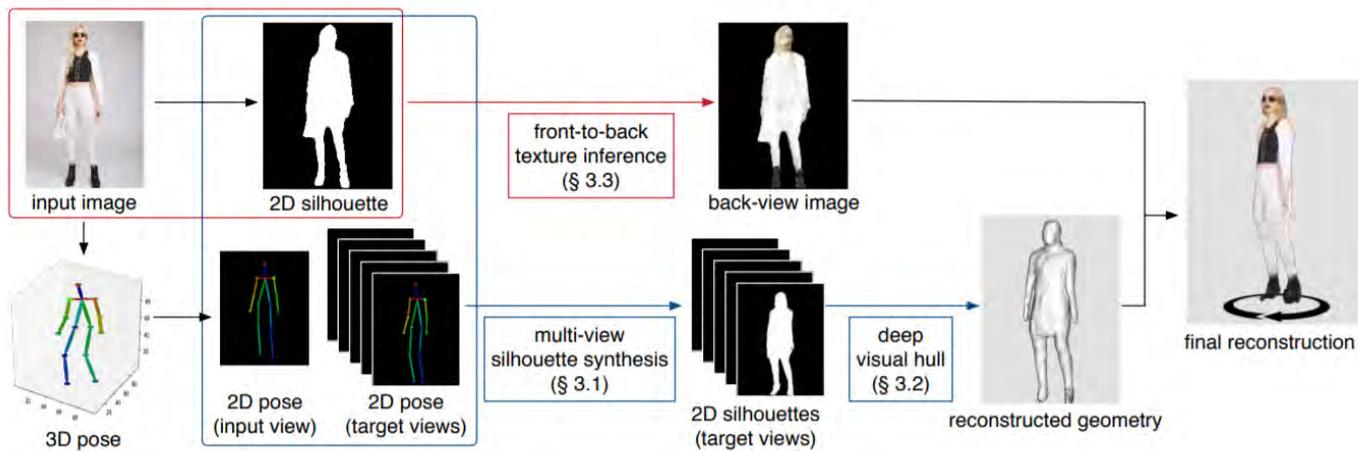


6 单图像人体三维重建 DeepHuman

6.2 相关工作

- SiCloPe (CVPR'19)

- 缺少表面细节
- 无法保证不同视角下重建的一致性

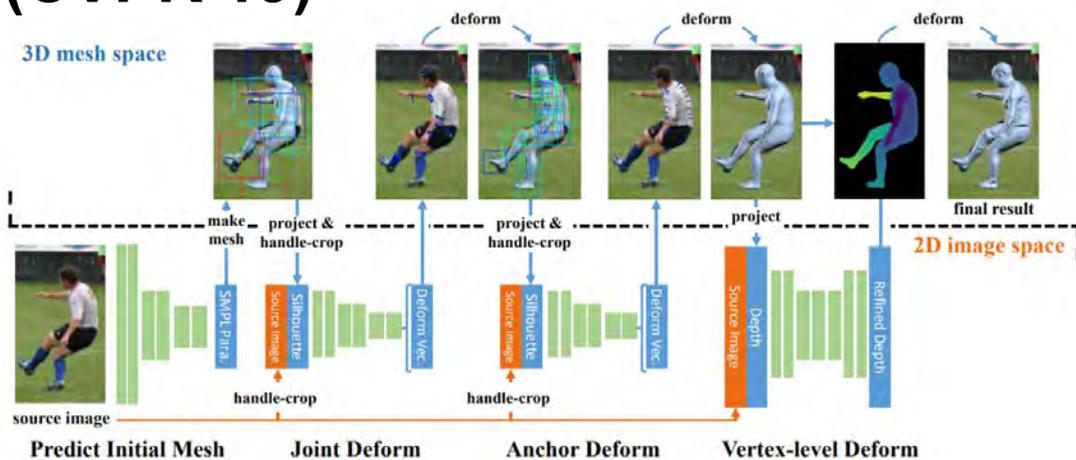


6 单图像人体三维重建 DeepHuman

6.2 相关工作

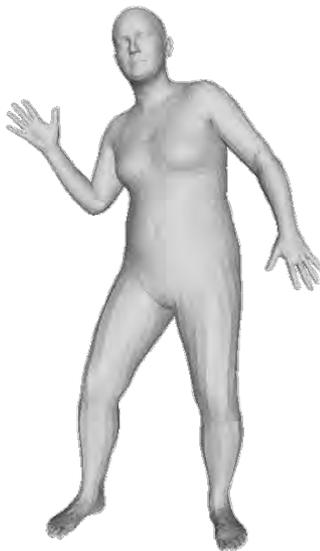
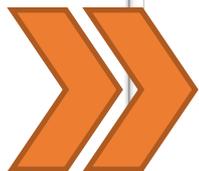
- HMD (CVPR'19)

• 无法重建宽松衣物



6 单图像人体三维重建 DeepHuman

6.3 核心思想



SMPL估计

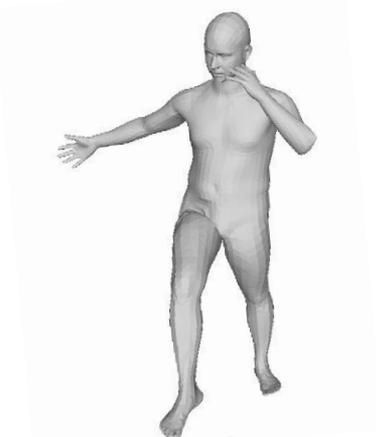
- 提供三维人体姿态和形状先验
- 解决RGB深度歧义性
- 使网络专注于重建不同衣物及表面细节

6 单图像人体三维重建 DeepHuman

6.4 输入输出



输入单张RGB图像



SMPL估计



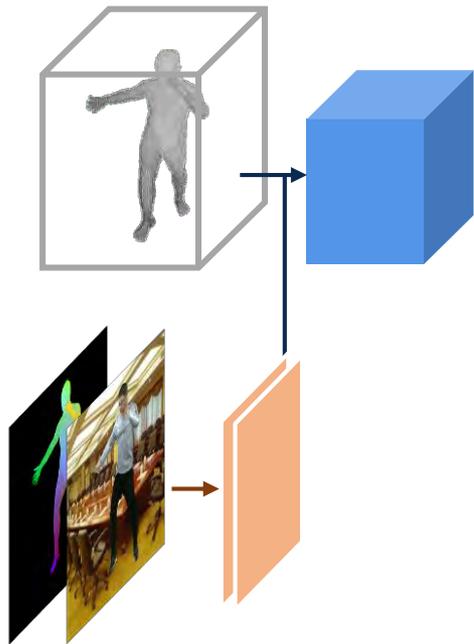
HOW?



真实稠密表面输出

6 单图像人体三维重建 DeepHuman

6.4 网络结构



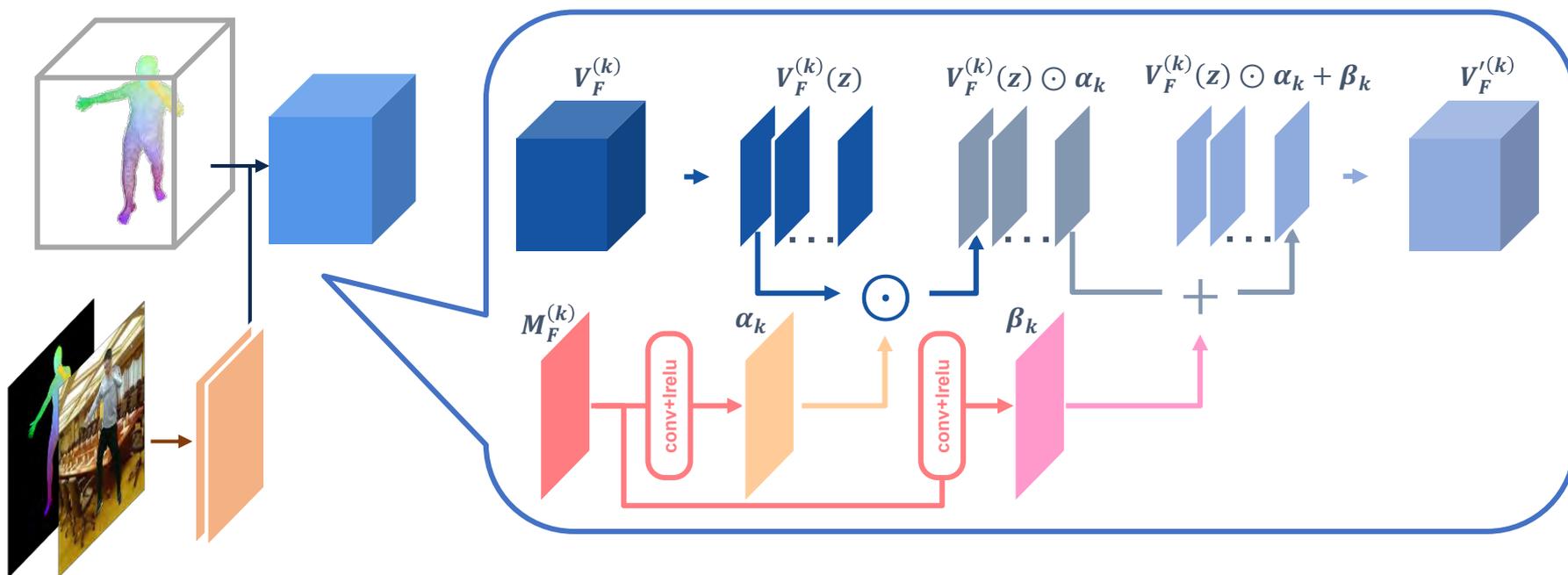
6 单图像人体三维重建 DeepHuman

6.4 网络结构



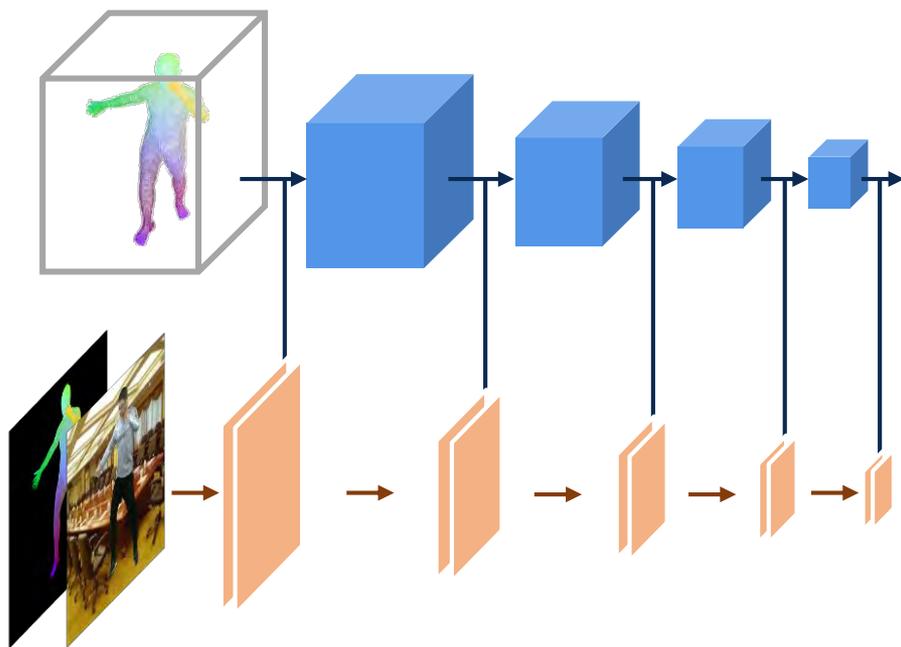
6 单图像人体三维重建 DeepHuman

6.4 网络结构



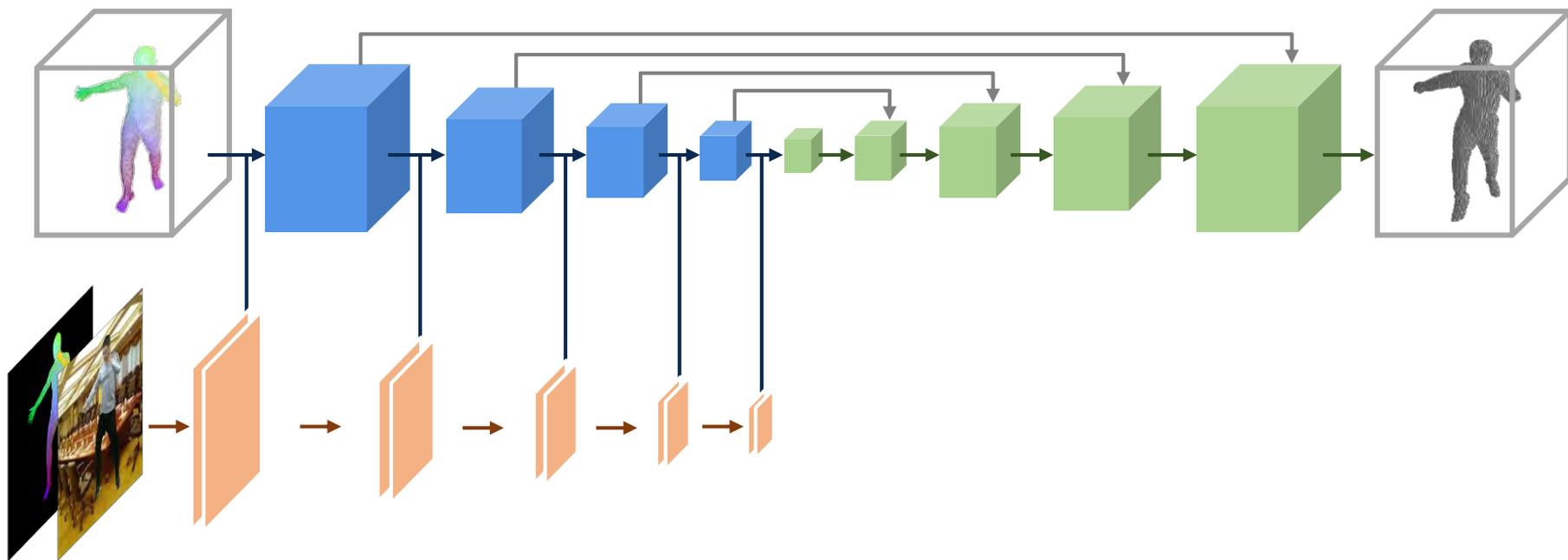
6 单图像人体三维重建 DeepHuman

6.4 网络结构



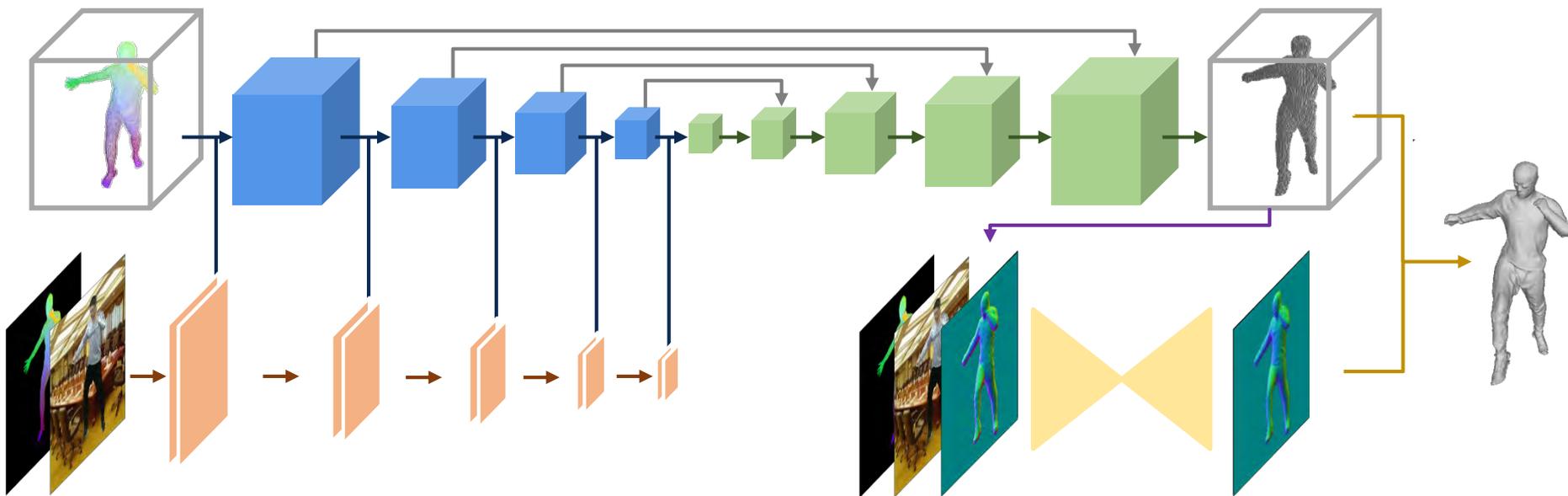
6 单图像人体三维重建 DeepHuman

6.4 网络结构



6 单图像人体三维重建 DeepHuman

6.4 网络结构



Binary Entropy Loss + Silhouette Loss (front view & side view) + L1 Normal Loss

6 单图像人体三维重建 DeepHuman

6.5 THUMAN数据集

- 目前仍缺少开源的大规模真实人体扫描数据集



DoubleFusion (CVPR'18)

THUMAN

A 3D real-world human model dataset

6 单图像人体三维重建 DeepHuman

6.6 重建结果



6 单图像人体三维重建 DeepHuman

6.6 重建结果 – 单RGB视频人体动态三维重建结果（每帧单独重建）



6 单图像人体三维重建 DeepHuman

6.6 与现有方法对比

- Comparison against BodyNet (ECCV'18)



Input

Ours

BodyNet



Input

Ours

BodyNet

6 单图像人体三维重建 DeepHuman

6.6 与现有方法对比

- Comparison against SiCloPe (CVPR'19)



Input



Ours



SiCloPe



Input



Ours



SiCloPe

6 单图像人体三维重建 DeepHuman

6.6 与现有方法对比

- Comparison against HMD (CVPR'19)



Input



Ours



HMD



Input



Ours



HMD

6 单图像人体三维重建 DeepHuman

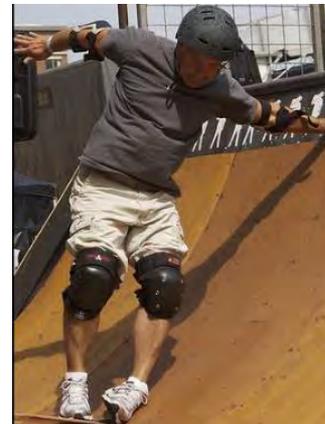
6.6 代码和THUMAN数据集已开源

- Code & data available at:



单图像人体三维重建相关文章:

- PIFu: Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization
- Tex2Shape: Detailed Full Human Body Geometry From a Single Image
- Moulding Humans: Non-Parametric 3D Human Shape Estimation From Single Images
- 3DPeople: Modeling the Geometry of Dressed Humans



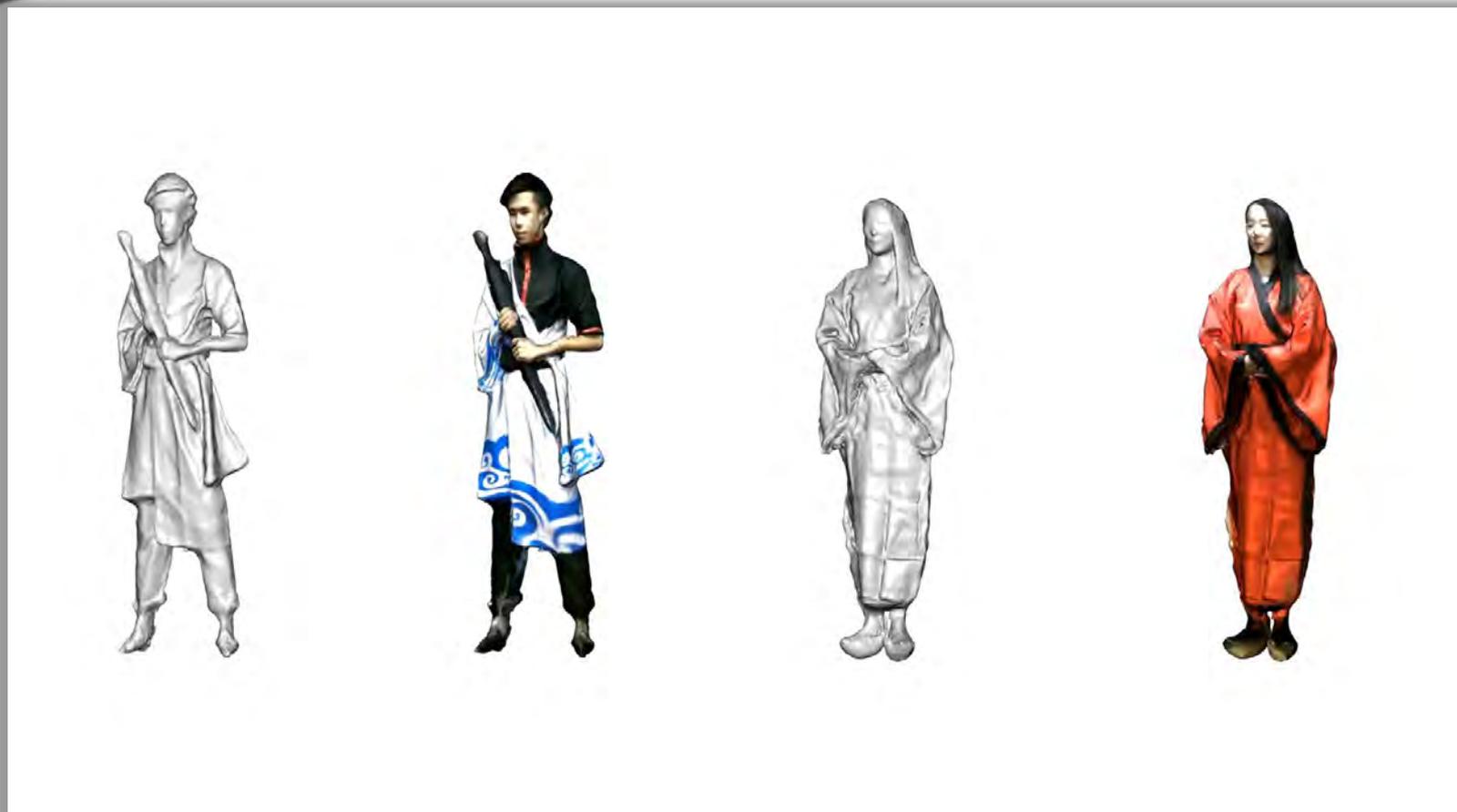
6 单图像人体三维重建 DeepHuman

6.6 3D Portraits in Seconds - THUMAN 2.0 ?



6 单图像人体三维重建 DeepHuman

6.6 3D Portraits in Seconds - THUMAN 2.0 ?



报告内容

- 1 研究背景及意义**
- 2 基于骨架跟踪的实时人体动态三维重建**
- 3 实时人体动态三维信息双层语义化重建**
- 4 高精度单视角人体快速动作捕捉算法**
- 5 实时人体动态三维信息多层语义化重建**
- 6 单图像人体三维重建**
- 7 总结与展望**

7.1 总结

无论是对传统方法还是统计学习方法，合理利用人体动态三维信息的多层语义化特性，可以提升人体动态三维重建的精度、鲁棒性和便捷性。

7.2 未来工作展望

1. 全人体动态重建（包括人脸和人手的实时动态三维重建）
2. 场景剧烈拓扑变化情况下的动态重建（穿脱衣物）
3. 神经渲染技术 **Neural Rendering**
4. 深度模板编码技术 **Codec Avatar**



谢谢大家!

清华大学刘焯斌研究组: liuyebin.com

个人主页: ytrock.com