

B 榜项目说明文档

1 团队信息

- 团队名称：广告位招租
- B 榜排名：6
- 联系人及联系方式：
 - 微信号：lyz_222222_，手机号：13041208690
 - 微信号：wyx1471593862，手机号：13507525624

2 项目概述

本文的开源代码见：[GitHub 开源链接](#)

2.1 核心任务

将三维网格绑定一套骨骼与蒙皮权重以达到驱动网格的目的。使用三维深度神经网络结构，预测三维人体网络架构相对应的骨骼节点的空间位置以及对应的蒙皮权重。

2.2 相比 A 榜的变化

1. **骨骼数量增多**：在 A 榜中，原本只有 22 个比较粗糙的骨骼。B 榜中额外添加了 30 个手部的骨骼 ($J=52$)。
2. **输入姿态改变**：在 A 榜中，所有需要预测的 mesh 均为 T-pose 或 A-pose 下的姿态，而在 B 榜中还会额外考察选手模型对任意 pose 骨骼预测的能力。

3 解决思路

我们对 A pose/T pose 和其他任意 pose 分别进行训练和预测（骨架采样点均为 4096，蒙皮权重采样点数均为 2048），具体如下：

3.1 A pose/T pose

- **骨架**：训练 baseline 骨架模型 2000 轮，得到预训练的骨架预测模型。
- **蒙皮权重**：我们设计了 Uniseg 模型，在优化原有 encoder 和 decoder 结构的同时，额外融入了对点云进行分段的信息，以提升模型的表达能力。

3.2 任意 pose

- **骨架**：训练 baseline 骨架模型 2000 轮得到预训练的骨架预测模型，然后增加全身骨骼长度和手部骨骼角度的正则约束进行后训练。
- **蒙皮权重**：考虑到角色的动作在变换的过程中，其蒙皮权重保持不变，因此与 A pose/T pose 的蒙皮权重保持统一。

在训练过程中，我们对数据进行了在线增强。具体来说：以 0.8 的概率对数据进行基于动捕关键帧的增强，以 0.2 的概率对数据进行随机动作增强。对于 validation 集，我们全进行基于动捕关键帧的增强。另外，我们利用 jittor 框架对数据的在线增强过程进行了加速。

代码结构说明

B 榜的代码与模型文件组织结构如下所示：

```
project_root/
+-- checkpoints/                                # 存放各类模型训练后的 checkpoint
|   +-- skeleton/
|       +-- AT_pose/
|           +-- mixamo/                        # 基于 Mixamo 数据集的 A/T pose 骨架模型
|               |-- best_model_1425_mixamo.pkl
|               +-- vroid/                     # 基于 Vroid 数据集的 A/T pose 骨架模型
|                   |-- best_model_805_vroid.pkl
|                   |-- any_pose/              # 任意姿态下的骨架预测模型
|                   |-- best_model_1055.pkl
```

```

| | \-- best_model_684_regular.pkl
| \-- skin/
| \-- AT_pose/ # A/T pose 下的蒙皮权重预测模型
| \-- best_model_204.pkl
|
+-- code/ # 完整代码实现
| +-- dataset/ # 数据加载与预处理模块
| +-- launch/ # 训练与预测脚本入口
| | +-- train_skeleton_mixamo.sh # 训练mixamo骨架模型
| | +-- train_skeleton_vroid.sh # 训练vroid骨架模型
| | +-- train_skeleton_pose.sh # 训练任意动作骨架模型
| | +-- train_skeleton_regular.sh # 训练任意动作骨架模型加入正则项后的后训练
| | +-- train_skin_AT.sh # 训练A/T pose蒙皮权重模型
| | +-- train_skin_pose.sh # 训练任意动作的蒙皮权重模型
| | +-- predict_skeleton_mixamo.sh # 预测mixamo骨架
| | +-- predict_skeleton_vroid.sh # 预测vroid骨架
| | +-- predict_skeleton_pose.sh # 预测任意动作骨架
| | +-- predict_skin_AT.sh # 预测A/T pose蒙皮权重
| +-- models/ # 模型结构定义
| | +-- skeleton.py # 骨架预测模型实现
| | +-- skin_unirig_seg.py # 蒙皮权重预测模型实现
| | \-- metrics.py # j2j 等评估指标
| +-- PCT/ # Point Cloud Transformer 模块
| | \-- ... # PCT 相关实现
| +-- train_skeleton.py # 训练骨架模型
| +-- train_skeleton_regular.py # 训练任意动作骨架模型加入正则项后的后训练
| +-- train_skin.py # 训练蒙皮权重模型
| +-- predict_skeleton.py # 预测骨架
| +-- predict_skin_AT.py # 预测A/T pose蒙皮权重
| \-- ... # 其他工具脚本
|
\-- environment.yml # 项目运行环境依赖配置文件

```

4 关键模块说明

关节位置预测模型

本模块基于 Point Cloud Transformer 和 MLP 架构，实现从点云到 52 个关节位置的预测。

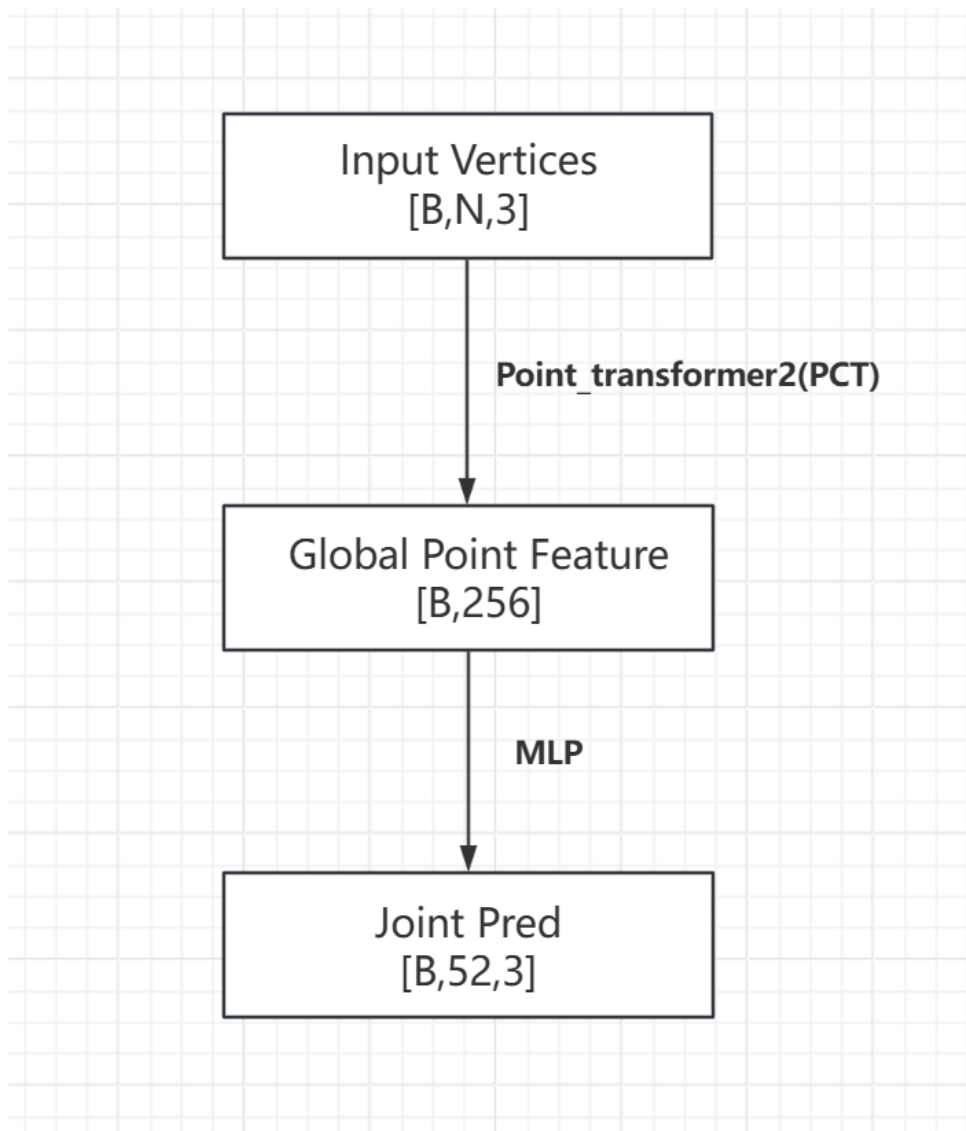


图 1: 关节位置预测模型示意图

蒙皮权重预测模型

采用 Uniseg 模型，在 encoder-decoder 结构基础上增加点云分段信息，以预测每个顶点对应 52 个关节的蒙皮权重，实现骨骼驱动网格的高精度蒙皮效果。

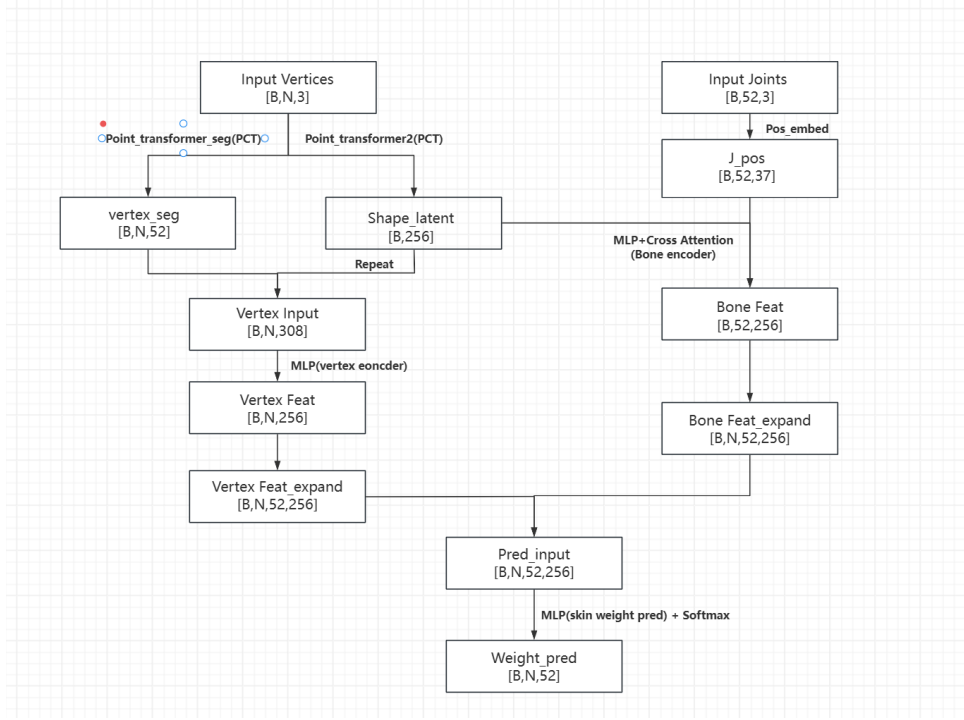


图 2: 蒙皮权重预测模型示意图

正则损失中的骨长约束和方向约束：

$$\mathcal{L}_{\text{bone}} = \sum_{(i,j) \in \mathcal{E}} \left(\|\hat{J}_i - \hat{J}_j\|_2 - \|J_i - J_j\|_2 \right)^2 \quad (1)$$

其中 $(i, j) \in \mathcal{E}$ 表示骨架中存在连接关系的关节对， J_i 和 \hat{J}_i 分别为 GT 与预测的第 i 个关节坐标。该约束保证预测骨架的骨长比例与真实解剖结构一致。

$$\mathcal{L}_{\text{dir}} = \sum_{(i,j) \in \mathcal{E}} \left(1 - \frac{\langle \hat{J}_i - \hat{J}_j, J_i - J_j \rangle}{\|\hat{J}_i - \hat{J}_j\|_2 \cdot \|J_i - J_j\|_2} \right) \quad (2)$$

该方向约束保证预测的骨骼方向与真实方向一致，特别适用于手部关节等方向限制较强的结构。

最终损失函数为：

$$\mathcal{L} = \mathcal{L}_{\text{pred}} + \lambda_b \mathcal{L}_{\text{bone}} + \lambda_d \mathcal{L}_{\text{dir}} \quad (3)$$

环境配置步骤

1. 配置环境：

```
conda env create -f environment.yml
```

2. 启动虚拟环境:

```
conda activate jittor_comp_human
```

运行步骤

首先将 B 榜数据集下载到”code/”下，然后将”code/”文件夹下的所有”.txt”文件剪切到”code/data/”文件夹下。

在项目根目录下执行:

训练流程

1. 第一步:

```
bash code/launch/train_skeleton_mixamo.sh
```

训练 2000epoch 得到 baseline 对 mixamo 的骨架预测模型

2. 第二步:

```
bash code/launch/train_skeleton_vroid.sh
```

训练 2000epoch 得到 baseline 对 vroid 的骨架预测模型

3. 第三步:

```
bash code/launch/train_skin_AT.sh
```

训练蒙皮权重预测模型 400 轮得到预测 A/T pose 的蒙皮权重模型

4. 第四步:

```
bash code/launch/train_skeleton_pose.sh
```

训练 2000epoch 得到 baseline 对任意动作的骨架预测模型

5. 第五步:

```
bash code/launch/train_skeleton_regular.sh
```

加入全身骨骼长度和手部骨骼方向的正则项对任意动作的骨架预测模型进行后训练, 2000 轮。

预测流程

依次执行:

```
predict_skeleton_mixamo.sh # 预测mixamo骨架
predict_skeleton_vroid.sh  # 预测vroid骨架
predict_skeleton_pose.sh   # 预测任意动作骨架
predict_skin_AT.sh         # 预测A/T pose蒙皮权重
```

checkpoint 说明

训练得到的各阶段模型及 loss 对应情况如下:

skeleton/AT_pose/mixamo/checkpoint.pkl	# 基于 Mixamo 数据集的 A/T pose 骨架模型
skeleton/AT_pose/vroid/checkpoint.pkl	# 基于 Vroid 数据集的 A/T pose 骨架模型
skeleton/any_pose/checkpoint.pkl	# 任意姿态下的骨架预测模型
skin/AT_pose/checkpoint.pkl	# A/T pose 下的蒙皮权重预测模型

其他补充说明

1. 注意 batchsize 的设置要与代码中一致。
2. 上述训练和预测均在 Ubuntu 22.04 + NVIDIA 4090 + CUDA 12.4 环境下进行。