



(12) 发明专利申请

(10) 申请公布号 CN 113112577 A

(43) 申请公布日 2021.07.13

(21) 申请号 202110426319.0

(22) 申请日 2021.04.20

(71) 申请人 网易(杭州)网络有限公司  
地址 310052 浙江省杭州市滨江区长河街  
道网商路599号4幢7层

(72) 发明人 石天阳 段颖琳 袁焱 范长杰  
胡志鹏

(74) 专利代理机构 北京超凡宏宇专利代理事务  
所(特殊普通合伙) 11463  
代理人 钟扬飞

(51) Int. Cl.  
G06T 13/40 (2011.01)  
G06K 9/00 (2006.01)  
G06N 3/04 (2006.01)  
G06N 3/08 (2006.01)

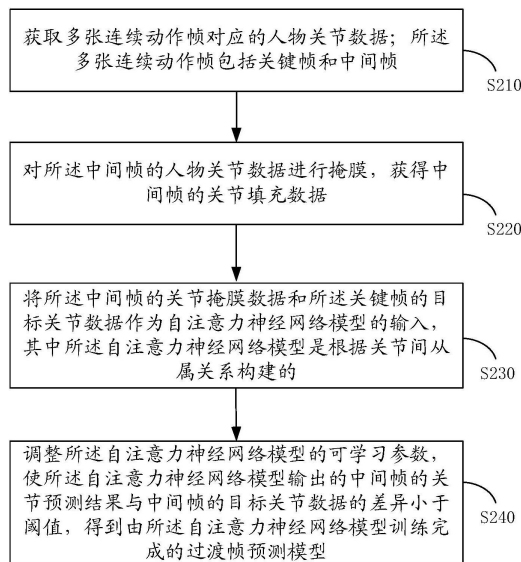
权利要求书3页 说明书11页 附图6页

(54) 发明名称

过渡帧预测模型的训练方法以及过渡帧预测方法

(57) 摘要

本申请提供一种过渡帧预测模型的训练方法以及过渡帧预测方法,该训练方法包括:获取多张连续动作帧对应的目标关节数据,其中多张连续动作帧包括关键帧和中间帧;对中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;将中间帧的关节掩膜数据和关键帧的目标关节数据作为自注意力神经网络模型的输入,其中自注意力神经网络模型是根据关节间从属关系构建的;调整自注意力神经网络模型的可学习参数,使自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到过渡帧预测模型。由此得到的过渡帧预测模型充分考虑了关节之间的联系,利用该模型预测得到的过渡帧,符合人体的运动规律,过渡帧的预测准确性高。



1. 一种过渡帧预测模型的训练方法,其特征在于,包括:

获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

2. 根据权利要求1所述的方法,其特征在于,所述关节间从属关系是指子节点与父节点之间的从属关系。

3. 根据权利要求1所述的方法,其特征在于,所述对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据,包括:

根据所述关键帧的目标关节数据,通过线性插值获得所述中间帧的关节掩膜数据。

4. 根据权利要求1所述的方法,其特征在于,所述对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据,包括:

将所述中间帧的目标关节数据进行置零,获得所述中间帧的关节掩膜数据。

5. 根据权利要求1所述的方法,其特征在于,所述基于关节间从属关系构建的自注意力神经网络模型包括:自注意力模块和前馈神经网络模块;将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,包括:

将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述自注意力模块的输入,获得所述自注意力模块输出的注意力得分矩阵;所述可学习参数为所述自注意力模块的查询权重 $Q$ 、键权重 $K$ 、值权重 $V$ ;

根据用于表征不同关节之间从属关系的关节关系矩阵,将所述注意力得分矩阵与所述关节关系矩阵进行逐元素相乘,获得改进后的注意力得分矩阵;

根据所述改进后的注意力得分矩阵,通过所述前馈神经网络模块计算所述中间帧的关节预测结果。

6. 根据权利要求5所述的方法,其特征在于,所述自注意力模块为多头自注意力模块;将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述自注意力模块的输入,获得所述自注意力模块输出的注意力得分矩阵,包括:

将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述多头自注意力模块的输入,获得所述多头自注意力模块输出的多个注意力得分矩阵;

其中,所述多个注意力得分矩阵对应多个所述改进后的注意力得分矩阵。

7. 根据权利要求6所述的方法,其特征在于,所述根据所述改进后的注意力得分矩阵,通过所述前馈神经网络模块计算所述中间帧的关节预测结果,包括:

根据多个所述改进后的注意力得分矩阵,通过拼接得到注意力得分拼接矩阵;

将由所述中间帧的关节掩膜数据和所述关键帧的目标关节数据构成的矩阵与所述注意力得分拼接矩阵进行求和并归一化,获得特征数据;

将所述特征数据作为所述前馈神经网络模块的输入,获得所述前馈神经网络模块输出

的所述中间帧的关节预测结果。

8. 根据权利要求7所述的方法,其特征在于,所述多头自注意力模块和前馈神经网络模块有6-12组,每组包含一个多头自注意力模块和前馈神经网络模块;前一组前馈神经网络模块的输出作为下一组多头自注意力模块的输入;所述中间帧的关节预测结果是根据最后一组前馈神经网络模块的输出得到的。

9. 根据权利要求1所述的方法,其特征在于,所述调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型,包括:

采用L1范数计算所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据之间的差异;

根据所述差异,采用随机梯度下降法对所述自注意力神经网络模型的可学习参数进行迭代优化,直到所述差异达到最小;

所述差异最小时的自注意力神经网络模型作为所述过渡帧预测模型。

10. 一种过渡帧预测方法,其特征在于,包括:

获取多个关键帧的目标关节数据;

根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据;

将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结果;

其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。

11. 根据权利要求10所述的方法,其特征在于,所述根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据,包括:

根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,通过线性插值获得过渡帧的关节掩膜数据。

12. 根据权利要求10所述的方法,其特征在于,所述根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据,包括:

用零表示过渡帧中人物每个关节的位置数据,获得所述过渡帧的关节掩膜数据。

13. 根据权利要求10所述的方法,其特征在于,所述将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入之前,所述方法还包括:

获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

14. 一种过渡帧预测模型的训练装置,其特征在于,包括:

数据获取模块,用于获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

数据掩膜模块,用于对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

数据输入模块,用于将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

模型训练模块,用于调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

15. 一种过渡帧预测装置,其特征在于,包括:

数据获取模块,用于获取多个关键帧的目标关节数据;

数据填充模块,用于根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据;

数据预测模块,用于将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结果;

其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。

16. 一种电子设备,其特征在于,所述电子设备包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行权利要求1-9任意一项所述的过渡帧预测模型的训练方法,或者,权利要求10-13任意一项所述的过渡帧预测方法。

17. 一种计算机可读存储介质,其特征在于,所述存储介质存储有计算机程序,所述计算机程序可由处理器执行以完成权利要求1-9任意一项所述的过渡帧预测模型的训练方法,或者,权利要求10-13任意一项所述的过渡帧预测方法。

## 过渡帧预测模型的训练方法以及过渡帧预测方法

### 技术领域

[0001] 本申请涉及图像处理技术领域,特别涉及一种过渡帧预测模型的训练方法及装置、过渡帧预测方法及装置、电子设备、计算机可读存储介质。

### 背景技术

[0002] 角色动画是游戏非常重要的组成部分,丰富的角色动画不仅可以充实游戏内容,还可以提升玩家的代入感和游戏性。一般而言,一款大型3D游戏所包含的角色动作数量非常庞大,需要数十人的团队耗费相当长的时间才能够完成全部的制作。

[0003] 目前,手工制作角色动画一般分为三步:设置关键帧,即在时间轴的特定位置为角色放置特定的动作,该动作越密集,生成的角色动画越流畅;动作插帧,利用线性或运动学模型在关键帧之间进行插值,一般以线性插值为主,所需插值的参数包括角色的位置和各个部位的旋转角;微调,由于自动插帧一般难以满足高质量角色动画的要求,因此还需要制作人员对角色各个部位的运动曲线进行微调。

[0004] 线性插值仅适用于关键帧较为密集的情况,但这样会极大地提高美术制作人员的工作量,如果关键帧数量减少,过渡帧预测的准确性降低,则会使动画中人物的动作缺乏连贯性。

### 发明内容

[0005] 本申请实施例提供了过渡帧预测模型的训练方法,通过该模型预测过渡帧,可以提高过渡帧预测的准确性。

[0006] 本申请实施例提供了一种过渡帧预测模型的训练方法,包括:

[0007] 获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

[0008] 对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

[0009] 将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

[0010] 调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

[0011] 在一实施例中,所述关节间从属关系是指子节点与父节点之间的从属关系。

[0012] 在一实施例中,所述对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据,包括:

[0013] 根据所述关键帧的目标关节数据,通过线性插值获得所述中间帧的关节掩膜数据。

[0014] 在一实施例中,所述对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据,包括:

[0015] 将所述中间帧的目标关节数据进行置零,获得所述中间帧的关节掩膜数据。

[0016] 在一实施例中,所述基于关节间从属关系构建的自注意力神经网络模型包括:自注意力模块和前馈神经网络模块;将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,包括:

[0017] 将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述自注意力模块的输入,获得所述自注意力模块输出的注意力得分矩阵;所述可学习参数为所述自注意力模块的查询权重 $Q$ 、键权重 $K$ 、值权重 $V$ ;

[0018] 根据用于表征不同关节之间从属关系的关节关系矩阵,将所述注意力得分矩阵与所述关节关系矩阵进行逐元素相乘,获得改进后的注意力得分矩阵;

[0019] 根据所述改进后的注意力得分矩阵,通过所述前馈神经网络模块计算所述中间帧的关节预测结果。

[0020] 在一实施例中,所述自注意力模块为多头自注意力模块;将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述自注意力模块的输入,获得所述自注意力模块输出的注意力得分矩阵,包括:

[0021] 将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为所述多头自注意力模块的输入,获得所述多头自注意力模块输出的多个注意力得分矩阵;

[0022] 其中,所述多个注意力得分矩阵对应多个所述改进后的注意力得分矩阵。

[0023] 在一实施例中,所述根据所述改进后的注意力得分矩阵,通过所述前馈神经网络模块计算所述中间帧的关节预测结果,包括:

[0024] 根据多个所述改进后的注意力得分矩阵,通过拼接得到注意力得分拼接矩阵;

[0025] 将由所述中间帧的关节掩膜数据和所述关键帧的目标关节数据构成的矩阵与所述注意力得分拼接矩阵进行求和并归一化,获得特征数据;

[0026] 将所述特征数据作为所述前馈神经网络模块的输入,获得所述前馈神经网络模块输出的所述中间帧的关节预测结果。

[0027] 在一实施例中,所述多头自注意力模块和前馈神经网络模块有6-12组,每组包含一个多头自注意力模块和前馈神经网络模块;前一组前馈神经网络模块的输出作为下一组多头自注意力模块的输入;所述中间帧的关节预测结果是根据最后一组前馈神经网络模块的输出得到的。

[0028] 在一实施例中,所述调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型,包括:

[0029] 采用L1范数计算所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据之间的差异;

[0030] 根据所述差异,采用随机梯度下降法对所述自注意力神经网络模型的可学习参数进行迭代优化,直到所述差异达到最小;

[0031] 所述差异最小时的自注意力神经网络模型作为所述过渡帧预测模型。

[0032] 本申请实施例还提供了一种过渡帧预测方法,包括:

[0033] 获取多个关键帧的目标关节数据;

[0034] 根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数

量,获得过渡帧的关节掩膜数据;

[0035] 将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结果;

[0036] 其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。

[0037] 在一实施例中,所述根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据,包括:

[0038] 根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,通过线性插值获得过渡帧的关节掩膜数据。

[0039] 在一实施例中,所述根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据,包括:

[0040] 用零表示过渡帧中人物每个关节的位置数据,获得所述过渡帧的关节掩膜数据。

[0041] 在一实施例中,所述将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入之前,所述方法还包括:

[0042] 获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

[0043] 对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

[0044] 将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

[0045] 调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

[0046] 本申请实施例提供了一种过渡帧预测模型的训练装置,包括:

[0047] 数据获取模块,用于获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

[0048] 数据掩膜模块,用于对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

[0049] 数据输入模块,用于将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

[0050] 模型训练模块,用于调整所述自注意力神经网络模型的可学习参数,使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

[0051] 本申请实施例提供了一种过渡帧预测装置,包括:

[0052] 数据获取模块,用于获取多个关键帧的目标关节数据;

[0053] 数据填充模块,用于根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据;

[0054] 数据预测模块,用于将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结

果；

[0055] 其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。

[0056] 本申请实施例还提供了一种电子设备,所述电子设备包括:

[0057] 处理器;

[0058] 用于存储处理器可执行指令的存储器;

[0059] 其中,所述处理器被配置为执行上述过渡帧预测模型的训练方法或者上述过渡帧预测方法。

[0060] 本申请实施例还提供了一种计算机可读存储介质,所述存储介质存储有计算机程序,所述计算机程序可由处理器执行以完成上述过渡帧预测模型的训练方法或者上述过渡帧预测方法。

[0061] 本申请上述实施例提供的技术方案,通过对中间帧的目标关节数据进行掩膜得到关节掩膜数据,关节掩膜数据与关键帧的目标关节数据一起作为输入,训练基于关节间从属关系构建的自注意力神经网络模型,使中间帧的关节预测结果与实际的目标关节数据的差异尽可能小,由此得到的过渡帧预测模型充分考虑了关节之间的联系,利用该模型预测得到的过渡帧,符合人体的运动规律,过渡帧的预测准确性高,与传统的LSTM模型相比,将基于自注意力机制的transformer模型应用于动作帧,训练之后在实际运用时效率高、单次运算即可完成、无需迭代。

## 附图说明

[0062] 为了更清楚地说明本申请实施例的技术方案,下面将对本申请实施例中所需要使用的附图作简单地介绍。

[0063] 图1是本申请实施例提供的电子设备的结构示意图;

[0064] 图2是本申请实施例提供的一种过渡帧预测模型的训练方法的流程示意图;

[0065] 图3是本申请实施例提供的人物动捕数据模型示意图;

[0066] 图4是图2对应实施例中步骤S230的细节流程图;

[0067] 图5是图4对应实施例中步骤S233的细节流程图;

[0068] 图6是本申请实施例提供的自注意力神经网络模型的架构示意图;

[0069] 图7是本申请实施例提供的过渡帧预测方法的流程示意图;

[0070] 图8是本申请实施例提供的过渡帧预测方法的架构示意图;

[0071] 图9是本申请一实施例示出的一种过渡帧预测模型的训练装置的框图;

[0072] 图10是本申请一实施例示出的一种过渡帧预测装置的框图。

## 具体实施方式

[0073] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行描述。

[0074] 相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释。同时,在本申请的描述中,术语“第一”、“第二”等仅用于区分描述,而不能理解为指示或暗示相对重要性。

[0075] 图1是本申请实施例提供的电子设备的结构示意图。该电子设备100可以用于执行



本申请实施例提供的过渡帧预测模型的训练方法以及过渡帧预测方法。如图1所示,该电子设备100包括:一个或多个处理器102、一个或多个存储处理器可执行指令的存储器104。其中,所述处理器102被配置为执行本申请下述实施例提供的过渡帧预测模型的训练方法以及过渡帧预测方法。

[0076] 所述处理器102可以是网关,也可以为智能终端,或者是包含中央处理单元(CPU)、图像处理单元(GPU)或者具有数据处理能力和/或指令执行能力的其它形式的处理单元的设备,可以对所述电子设备100中的其它组件的数据进行处理,还可以控制所述电子设备100中的其它组件以执行期望的功能。

[0077] 所述存储器104可以包括一个或多个计算机程序产品,所述计算机程序产品可以包括各种形式的计算机可读存储介质,例如易失性存储器和/或非易失性存储器。所述易失性存储器例如可以包括随机存取存储器(RAM)和/或高速缓冲存储器(cache)等。所述非易失性存储器例如可以包括只读存储器(ROM)、硬盘、闪存等。在所述计算机可读存储介质上可以存储一个或多个计算机程序指令,处理器102可以运行所述程序指令,以实现下文所述的过渡帧预测模型的训练方法以及过渡帧预测方法。在所述计算机可读存储介质中还可以存储各种应用程序和各种数据,例如所述应用程序使用和/或产生的各种数据等。

[0078] 在一实施例中,图1所示电子设备100还可以包括输入装置106、输出装置108以及数据采集装置110,这些组件通过总线系统112和/或其它形式的连接机构(未示出)互连。应当注意,图1所示的电子设备100的组件和结构只是示例性的,而非限制性的,根据需要,所述电子设备100也可以具有其他组件和结构。

[0079] 所述输入装置106可以是用户用来输入指令的装置,并且可以包括键盘、鼠标、麦克风和触摸屏等中的一个或多个。所述输出装置108可以向外部(例如,用户)输出各种信息(例如,图像或声音),并且可以包括显示器、扬声器等中的一个或多个。所述数据采集装置110可以采集对象的图像,并且将所采集的图像存储在所述存储器104中以供其它组件使用。示例性地,该数据采集装置110可以为摄像头。

[0080] 在一实施例中,用于实现本申请实施例的过渡帧预测模型的训练方法以及过渡帧预测方法的示例电子设备100中的各器件可以集成设置,也可以分散设置,诸如将处理器102、存储器104、输入装置106和输出装置108集成设置于一体,而将数据采集装置110分离设置。

[0081] 在一实施例中,用于实现本申请实施例的过渡帧预测模型的训练方法以及过渡帧预测方法的示例电子设备100可以被实现为诸如智能手机、平板电脑、台式电脑、笔记本电脑等智能终端。

[0082] 图2是本申请实施例提供的一种过渡帧预测模型的训练方法的流程示意图。如图2所示,该方法可以由电子设备执行,该方法包括:

[0083] 步骤S210:获取多张连续动作帧对应的目标关节数据;所述多张连续动作帧包括关键帧和中间帧。

[0084] 其中,多张连续动作帧可以是一个动作包含的多张连续的图像,也可以是多个动作,每个动作包含多张连续的图像。举例来说,动作可以是行走、跳舞、跑跳或运动等。以开源数据集Lafan1数据集为例,该数据集包含了5个目标,77个动作序列,共496672个动作帧,动作播放速率为30FPS,即大约4.6小时。本申请实施例中,多张连续动作帧可以是上述数

据集中获取8个动作序列,每个动作序列的长度为50帧,即 $8 \times 50$ 帧图像的目标关节数据作为训练样本。目标可以是具有关节的人物或动物,动作帧中包含的人物或动物的各个关节的位置数据,表示动作帧对应的目标关节数据。

[0085] 关键帧是指人物运动或变化中的关键动作所处的那一帧。关键帧与关键帧之间的动画可以由计算机或人工来创建,叫做过渡帧或者中间帧。本申请为将训练阶段和预测阶段的动作帧进行区分,训练阶段的关键帧与关键帧之间的动作帧称为中间帧。预测阶段的关键帧与关键帧之间的动作帧称为过渡帧。

[0086] 目标关节数据是指图像中人物各个关节的位置坐标 $(x, y, z)$ 和旋转角,其中,旋转角可以用四元数表示。四元数一般可表示为 $a+bi+cj+dk$ ,其中 $a, b, c, d$ 是实数。其中 $i$ 代表X轴与Y轴相交平面中X轴正向向Y轴正向的旋转, $j$ 代表Z轴与X轴相交平面中Z轴正向向X轴正向的旋转, $k$ 代表Y轴与Z轴相交平面中Y轴正向向Z轴正向的旋转。关节数量可以有22个,故一帧图像的目标关节数据有 $22 \times 7 = 154$ 维,多张连续动作帧对应的目标关节数据可以由 $8 \times 50 \times 154$ 维。其中,某个关节的位置坐标和旋转角可以是相对父节点的相对位置坐标和相对旋转角。某个关节的位置坐标和旋转角也可以是世界坐标系下的位置坐标和旋转角。

[0087] 图3是人物动捕数据模型示意图。如图3所示,关节301可以认为是关节302的父节点,关节303可以认为是关节301的父节点,以此类推。关节304可以认为是根节点。人体动捕数据集一般包含人体模型和动作数据两部分,人体模型记录各个关节相对父节点的位置,为静态数据。动作数据记录根节点的运动轨迹和各个关节相对父节点的旋转关系。

[0088] 步骤S220:对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据。

[0089] 其中,掩膜是指采用其他数值替代真实的目标关节数据。进而可以通过模型预测中间帧的目标关节数据,如果关节预测结果与真实的目标关节数据基本一致,可以认为模型训练完毕,之后可以利用训练好的模型来预测中间帧的目标关节数据。

[0090] 中间帧的关节掩膜数据是指去除中间帧真实的目标关节数据,采用现有的中间帧预测方式,预估得到的关节数据。

[0091] 在一实施例中,可以根据所述关键帧的目标关节数据,通过线性插值获得所述中间帧的关节掩膜数据。

[0092] 线性插值(linear interpolation)是一种插值方法,是指以线性函数为插值函数的插值方法。线性插值的公式可以如下所示 $Y = Y_0 + \frac{Y_1 - Y_0}{X_1 - X_0}(X - X_0)$ ,举例来说,第一帧为关键帧,第三帧为关键帧,第一帧和第三帧之间有一帧中间帧,故中间帧的关节掩膜数据 $Y$ 可以根据第一帧目标关节数据 $Y_0$ 和第三帧目标关节数据 $Y_1$ ,通过上述公式计算得到中间帧的关节掩膜数据 $Y$ 。其中, $X_0$ 代表第一帧可以等于1, $X$ 代表第二帧,可以等于2, $X_1$ 代表第三帧可以等于3。

[0093] 举例来说,假设第一帧和第四帧是关键帧,中间帧有两帧,则 $Y_0$ 为第一帧的目标关节数据, $X_0 = 1$ ;  $Y_1$ 为第四帧的目标关节数据, $X_1 = 4$ ,当 $X$ 取2时,计算得到第二帧的关节掩膜数据,当 $X$ 取3时,计算得到第三帧的关节掩膜数据。

[0094] 为简化计算,在另一实施例中,可以将所述中间帧的目标关节数据进行置零,获得

所述中间帧的关节掩膜数据。

[0095] 也就是说,将中间帧的目标关节数据中每个关节的位置坐标和旋转角都用0来替代,作为中间帧的关节掩膜数据。

[0096] 步骤S230:将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为基于关节间从属关系构建的自注意力神经网络模型的输入。

[0097] 多帧连续的动作帧因中间帧进行掩膜,以置零为例,此时模型的输入数据可以表示为  $\tilde{C} = [C1, C2, 0, C4, \dots, Cn]$ 。C1、C2……Cn表示第一帧到第n帧的目标关节数据,其中,用0表示中间帧的目标关节数据。

[0098] 其中,自注意力神经网络模型是指将自注意力机制应用于transformer神经网络中。其中,第一个关节的运动带动第二个关节的运动,则表示这两个关节之间存在从属关系,第一个关节可以称为父节点,第二个关节可以称为子节点,所以关节间从属关系是指父节点与子节点之间的从属关系。基于关节间从属关系构建的自注意力神经网络模型是指根据不同关节之间的从属关系可以形成节点关系图,在图模型的框架下进一步增强自注意力神经网络模型的性能。

[0099] 自注意力神经网络模型可以包括自注意力模块和前馈神经网络模块。如图4所示,上述步骤S230具体包括:步骤S231,将中间帧的关节掩膜数据和关键帧的目标关节数据作为自注意力模块的输入,即可获得自注意力模块输出的注意力得分矩阵。注意力得分矩阵用于表征不同帧之间的相关性。

[0100] 自注意力模块进行如下运算:

$$[0101] \quad Q = Q^Q X$$

$$[0102] \quad K = W^K X$$

$$[0103] \quad V = W^V X \quad (1)$$

[0104] 其中,公式(1)中的X代表输入的中间帧的关节掩膜数据和关键帧的目标关节数据, $W^Q, W^K, W^V$ 分别为可学习参数,Q,K,V为注意力机制下的得分。根据上述结果,即可得到该模块的输出Z(即注意力得分矩阵):

$$[0105] \quad Z = \text{softmax}\left(\frac{QK^T}{\sqrt{n}}\right)V \quad (2)$$

[0106] 上述自注意力模块可以进一步改进为如下形式:

$$[0107] \quad Z = (M \odot \text{softmax}\left(\frac{QK^T}{\sqrt{n}}\right))V \quad (3)$$

[0108] 其中, $\sqrt{n}$ 为归一化参数,n表示序列长度。M是指用于表征关节之间从属关系的矩阵, $\odot$ 表示两个矩阵逐元素相乘,公式(3)的Z可以认为是改进后的注意力得分矩阵。

[0109] 上述注意力得分  $\text{softmax}\left(\frac{QK^T}{\sqrt{n}}\right)$ 并未考虑关节之间的关系,因此本申请引入了基于关节间从属关系的矩阵M,以便该模块可以对关节的从属关系进行建模。上述步骤S230还包括:步骤S232,根据用于表征不同关节之间从属关系的关节关系矩阵M,将注意力得分矩阵Z与所述关节关系矩阵M进行逐元素相乘,获得改进后的注意力得分矩阵。M与注意力得分矩阵相同,即为J\*J的矩阵(J为关节数量),若第p个关节和第q个关节为从属关系,则M

$(p, q) = 1$ , 其余为0。例如22个关节点,  $M$ 为22\*22的矩阵, 假设第一个关节点和第三个关节点具有从属关系, 第一个关节点是第三个关节点的父节点或第三个关节点是第一个关节点的父节点, 则 $M(1, 3) = 1$ 或 $M(3, 1) = 1$ 等。

[0110] 上述步骤S230还包括: 步骤S233, 根据改进后的注意力得分矩阵, 通过前馈神经网络模块计算所述中间帧的关节预测结果。本申请采用简单的线性映射实现前馈神经网络模块的功能, 具体地, 假设 $X$ 表示某个时刻的特征, 则输出 $Y$ 服从如下关系:

$$[0111] \quad Y = \text{norm}(X + f(W^T X + b)) \quad (4)$$

[0112] 其中,  $X$ 可以是上述注意力得分 $Z$ 。 $f$ 为非线性激活层, 如ReLUd(激活函数)等, 而 $\text{norm}$ 为归一化层,  $W^T$ 为可学习的参数,  $Y$ 为输出的预测结果, 预测结果包括中间帧的关节预测结果。此外, 由于深度模型具有更好地拟合特性, 因此本申请采用多个上述模块以搭建更深层的模型, 在具体实践中, 神经网络深度取6~12较为合适, 优选为6。

[0113] 步骤S240: 调整所述自注意力神经网络模型的可学习参数, 使所述自注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值, 得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

[0114] 参照上文实施例, 可学习参数 $W^Q, W^K, W^V$ 以及 $W^T$ 可以认为是需要调整的自注意力神经网络模型的参数。

[0115] 中间帧的关节预测结果是指中间帧的目标关节数据的预测值。中间帧的关节预测结果可以通过自注意力神经网络模型利用上述公式(1)(2)(3)(4)计算得到。通过不断调整自注意力神经网络模型的参数, 可以改变关节预测结果, 最终使关节预测结果与中间帧真实的目标关节数据的差异小于阈值, 优化完成的自注意力神经网络模型, 可以称为过渡帧预测模型。

[0116] 在一实施例中, 可以采用L1范数度量自注意力神经网络模型 $G$ 所预测的结果与真值之间的距离, 具体损失函数可表示如下:

$$[0117] \quad \mathcal{L} = l1(X, G(\tilde{X})) \quad (5)$$

[0118] 在一实施例中, 根据损失函数, 可以采用随机梯度下降法对自注意力神经网络注意力模型 $G$ 的可学习参数进行迭代优化, 直到差异达到最小; 所述差异最小时的自注意力神经网络模型作为所述过渡帧预测模型。

[0119] 具体地, 对于给定的随机批次数据 $X$ , 掩膜后的数据可表示为 $\tilde{X} = [X_1, X_2, 0, X_4, \dots, X_n]$ , 将 $\tilde{X}$ 作为自注意力神经网络模型的输入, 计算模型输出结果 $G(\tilde{X})$ 与真实值 $X$ 之间的距离, 求解下列目标函数:  $\min_G \mathcal{L}$ ; 即目标函数最小时的自注意力神经网络模型即为训练得到的过渡帧预测模型。

[0120] 在一实施例中, 上述实施例中的注意力模块可以是多头注意力模块。多头是指公式(1)执行了多次, 每一次算一个头, 每次之间的参数 $W^Q, W^K, W^V$ 不共享, 即每次的参数是不一样的。上述步骤S231具体包括将中间帧的关节掩膜数据和关键帧的目标关节数据作为多头自注意力模块的输入, 获得所述多头自注意力模块输出的多个注意力得分矩阵; 进而执行上述步骤S232即可获得多个注意力得分矩阵对应的多个改进后的注意力得分矩阵。即对于每个“头”都进行公式(1)(2)(3)的运算, 可以得到多个改进后的注意力得分矩阵 $Z_1, Z_2, Z_3 \dots$ 。

[0121] 如图5所示,上述步骤S233具体包括:步骤S501,根据多个所述改进后的注意力得分矩阵,通过拼接得到注意力得分拼接矩阵。

[0122] 具体可以利用下式对多个改进后的注意力得分矩阵 $Z_1, Z_2, Z_3 \dots$ 进行拼接

$$[0123] \quad \hat{Z} = W_0^T [Z_1, Z_2, Z_3, \dots] \quad (6)$$

[0124] 其中, $W_0^T$ 为可学习参数。 $\hat{Z}$ 为注意力得分拼接矩阵。

[0125] 步骤S502,将由所述中间帧的关节掩膜数据和所述关键帧的目标关节数据构成的矩阵与所述注意力得分拼接矩阵进行求和并归一化,获得特征数据。

[0126] 考虑到上述操作可能造成信息损失,本申请利用残差学习的方式解决该问题,因此最终该前馈神经网络的输入表示如下:

$$[0127] \quad Out = norm(X + \hat{Z}) \quad (7)$$

[0128] 其中, $X$ 表示公式(1)输入的 $X$ (即中间帧的关节掩膜数据和关键帧的目标关节数据构成的矩阵), $\hat{Z}$ 为注意力得分拼接矩阵; $norm$ 表示归一化,归一化可以是基于层的归一化,即在每一层特征上做归一化(例如前述的154个维度)。也可以是基于批次的归一化,即对每一批输入数据上做归一化(例如前述8个维度。将输入 $X$ 和注意力得分拼接矩阵 $\hat{Z}$ 先求和,再进行归一化,得到的结果可以称为特征数据,用于表征输入数据的特征。

[0129] 步骤S503,将所述特征数据作为所述前馈神经网络模块的输入,获得所述前馈神经网络模块输出的所述中间帧的关节预测结果。

[0130] 在一实施例中,所述多头自注意力模块和前馈神经网络模块可以有6-12组,优选为6,每组包含一个多头自注意力模块和前馈神经网络模块;前一组前馈神经网络模块的输出作为下一组多头自注意力模块的输入;由此重复6-12次,故所述中间帧最终的关节预测结果可以是根据最后一组前馈神经网络模块的输出得到的。由此得到的中间帧的关节预测结果的准确性更高。

[0131] 图6是自注意力神经网络模型的架构示意图, $\oplus$ 表示求和并归一化。如图6所示,基于图的自注意力模块是指基于关节之间从属关系改进的自注意力模块。基于图的自注意力模块是多头注意力模块,图6中画出2个进行示例。每个基于图的注意力模块可以进行上述公式(1)(2)(3)的运算,得到 $Z_1, Z_2, Z_3 \dots$ 。

[0132] 图6中拼接模块可以执行上述公式(6)  $\hat{Z} = W_0^T [Z_1, Z_2, Z_3, \dots]$ 对每个“基于图的注意力模块”的计算结果进行整合。拼接模块的输出表示如下(即图中601 $\oplus$ 所涉及的运算):  $Out = norm(X + \hat{Z})$ 。OUT作为前馈神经网络模型的输入,前馈神经网络模块的输出包含关键帧的目标关节数据以及中间帧的关节预测结果。 $N_x$ 可以是6-12,优选为6,通过优化基于图的自注意力模块、拼接模块以及前馈神经网络模块的参数,可以使计算得到的关节预测结果与中间帧的目标关节数据的差异最小化,从而得到训练完成的过渡帧预测模型。

[0133] 图7是本申请实施例提供的过渡帧预测方法的流程示意图。如图7所示,该方法包括以下步骤S710-步骤S730。

[0134] 步骤S710:获取多个关键帧的目标关节数据。

[0135] 目标关节数据是指图像中人物的22个关节的位置坐标和旋转角。关键帧至少有2张,即至少获取2张关键帧的目标关节数据。

[0136] 步骤S720:根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据。

[0137] 过渡帧数量是指关键帧之间包含的过渡帧的数量。关节掩膜数据是相对实际的目标关节数据而言的,是初步设定的关节数据。由于在预测阶段,只有关键帧,过渡帧的目标关节数据未知,故可以按照所需的过渡帧数量,初步生成指定帧数的关节掩膜数据。在一实施例中,可以用零表示过渡帧中人物每个关节的位置数据,获得过渡帧的关节掩膜数据。

[0138] 在一实施例中,可以根据多个关键帧的目标关节数据以及多个关键帧之间预设的过渡帧数量,通过线性插值获得过渡帧的关节掩膜数据。其中,线性插值获得过渡帧的关节掩膜数据的方式可以参照上文通过线性插值获得中间帧的关节掩膜数据的方式。在此不再赘述。

[0139] 步骤S730:将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结果。

[0140] 其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。在一实施例中,过渡帧预测模型可以采用本申请实施例提供的过渡帧预测模型训练方法训练得到,具体训练过程可以参照上文图2-图6对应实施例的描述。

[0141] 如图8所示,关键帧的目标关节数据以及过渡帧的关节掩膜数据作为过渡帧预测模型的输入,过渡帧预测模型输出关键帧的目标关节数据以及过渡帧的关节预测结果。过渡帧的关节预测结果是指过渡帧中每个关节的坐标位置和旋转角的预测值,从而实现自动化的动作插帧。该模型可以在一次前向传播的情况下就直接预测出全部的过渡动作,无需迭代求解,因此具有非常高的计算效率,且方便并行求解结合关节之间的联系进行建模,进一步提高了过渡帧预测的准确性。

[0142] 采用本申请实施例提供的技术方案可以预测较长时间的过渡帧,对于动画制作人员而言,大幅降低动画制作人员的工作量。目前游戏中主要采用线性插值生成过渡帧,该动画往往较为生硬,利用本申请则可以进一步生成更加流畅的过渡动画,大幅改善玩家的游戏体验。

[0143] 下述为本申请装置实施例,可以用于执行本申请上述过渡帧预测模型的训练方法实施例。对于本申请装置实施例中未披露的细节,请参照本申请过渡帧预测模型的训练方法实施例。

[0144] 图9为本申请一实施例示出的一种过渡帧预测模型的训练装置的框图。如图9所示,该装置包括:数据获取模块910、数据掩膜模块920、数据输入模块930以及模型训练模块940。

[0145] 数据获取模块910,用于获取多张连续动作帧对应的目标关节数据,其中所述多张连续动作帧包括关键帧和中间帧;

[0146] 数据掩膜模块920,用于对所述中间帧的目标关节数据进行掩膜,获得中间帧的关节掩膜数据;

[0147] 数据输入模块930,用于将所述中间帧的关节掩膜数据和所述关键帧的目标关节数据作为自注意力神经网络模型的输入,其中所述自注意力神经网络模型是根据关节间从属关系构建的;

[0148] 模型训练模块940,用于调整所述自注意力神经网络模型的可学习参数,使所述自

注意力神经网络模型输出的中间帧的关节预测结果与中间帧的目标关节数据的差异小于阈值,得到由所述自注意力神经网络模型训练完成的过渡帧预测模型。

[0149] 上述装置中各个模块的功能和作用的实现过程具体详见上述过渡帧预测模型的训练方法中对应步骤的实现过程,在此不再赘述。

[0150] 下述为本申请另一装置实施例,可以用于执行本申请上述过渡帧预测方法实施例。对于本申请装置实施例中未披露的细节,请参照本申请过渡帧预测方法实施例。

[0151] 图10为本申请一实施例示出的一种过渡帧预测装置的框图。如图10所示,该装置包括:数据获取模块1010、数据填充模块1020以及数据预测模块1030。

[0152] 数据获取模块1010,用于获取多个关键帧的目标关节数据;

[0153] 数据填充模块1020,用于根据多个所述关键帧的目标关节数据以及多个所述关键帧之间预设的过渡帧数量,获得过渡帧的关节掩膜数据;

[0154] 数据预测模块1030,用于将多个所述关键帧的目标关节数据以及所述过渡帧的关节掩膜数据作为过渡帧预测模型的输入,获得所述过渡帧预测模型输出的过渡帧的关节预测结果;

[0155] 其中,所述过渡帧预测模型是由基于关节间从属关系构建的自注意力神经网络模型训练得到的。

[0156] 上述装置中各个模块的功能和作用的实现过程具体详见上述过渡帧预测方法中对应步骤的实现过程,在此不再赘述。

[0157] 在本申请所提供的几个实施例中,所揭露的装置和方法,也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本申请的多个实施例的装置、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分,模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现方式中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0158] 另外,在本申请各个实施例中的各功能模块可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或两个以上模块集成形成一个独立的部分。

[0159] 功能如果以软件功能模块的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本申请各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

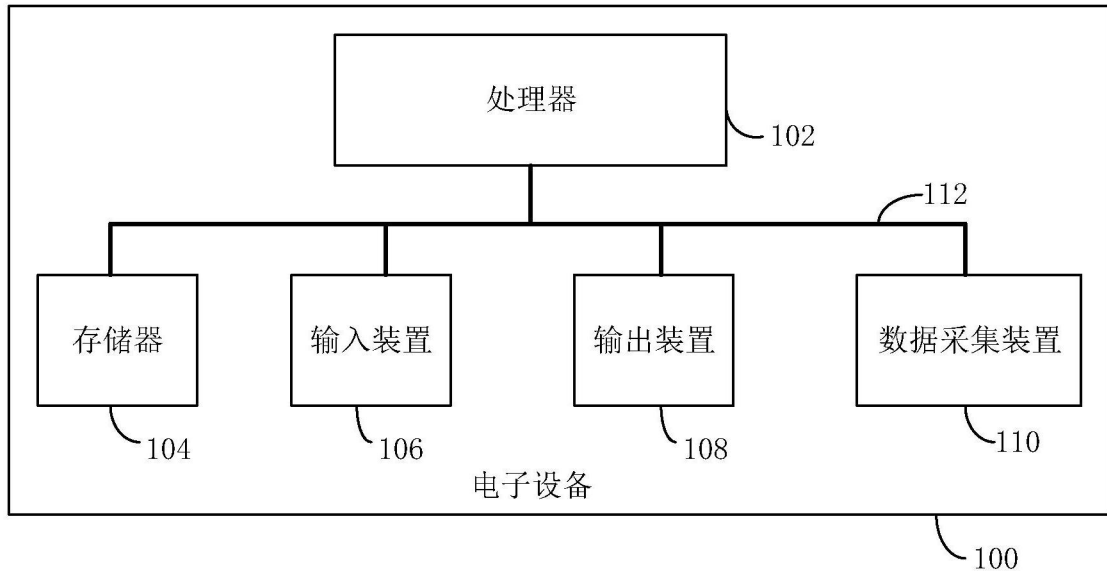


图1

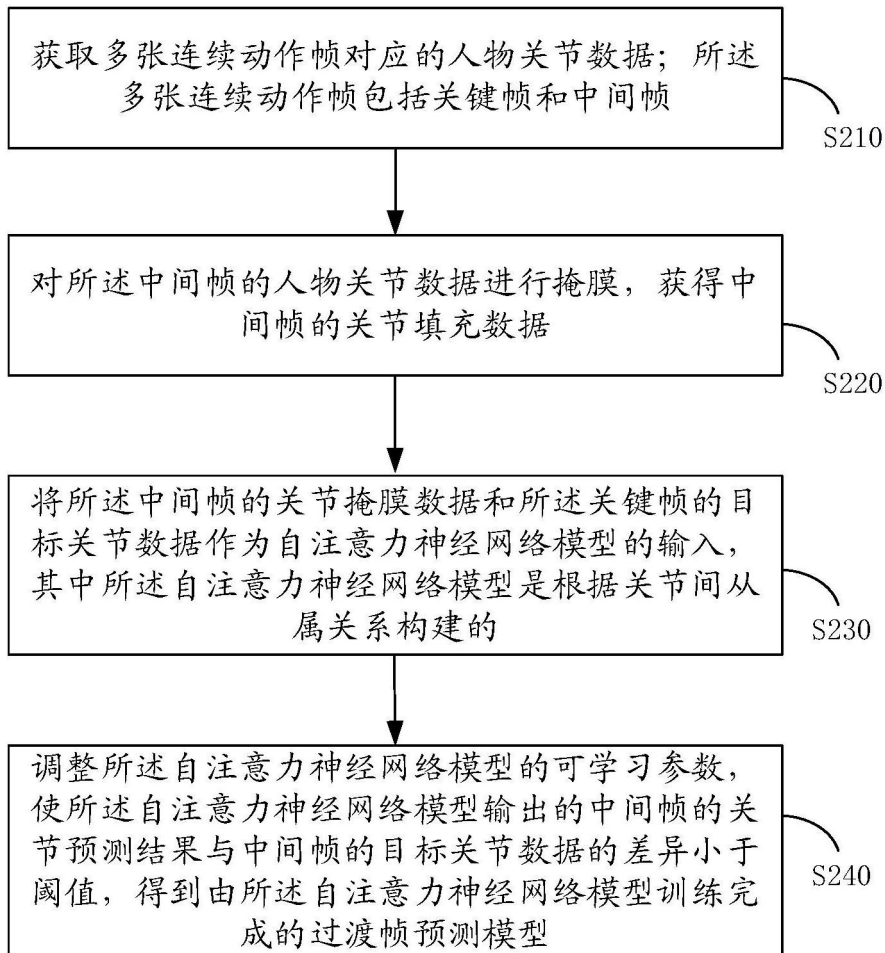


图2



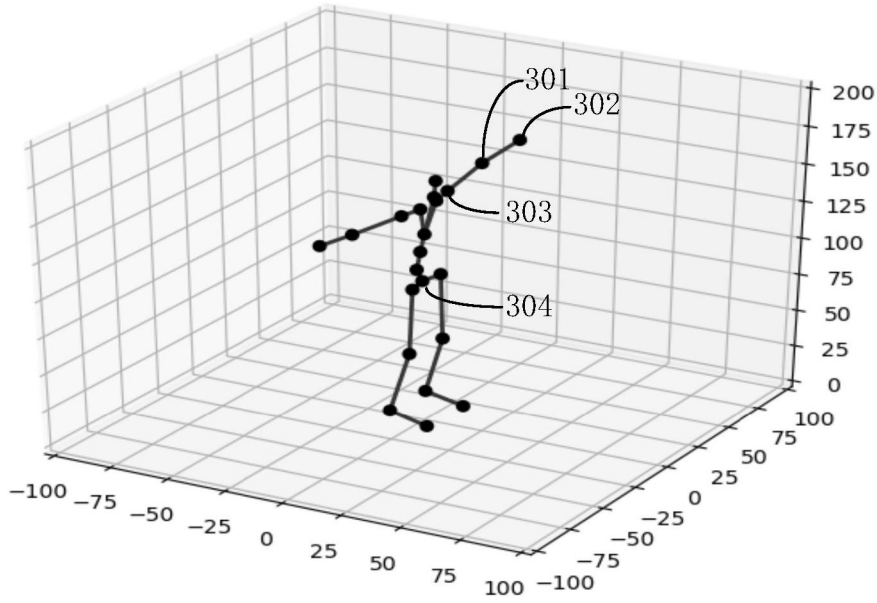


图3

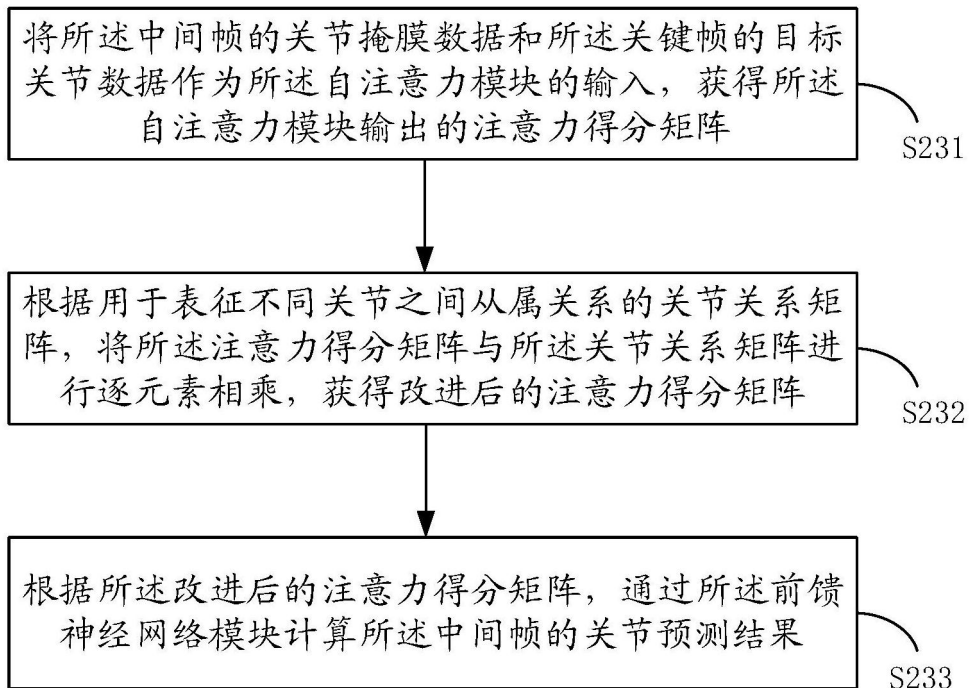


图4

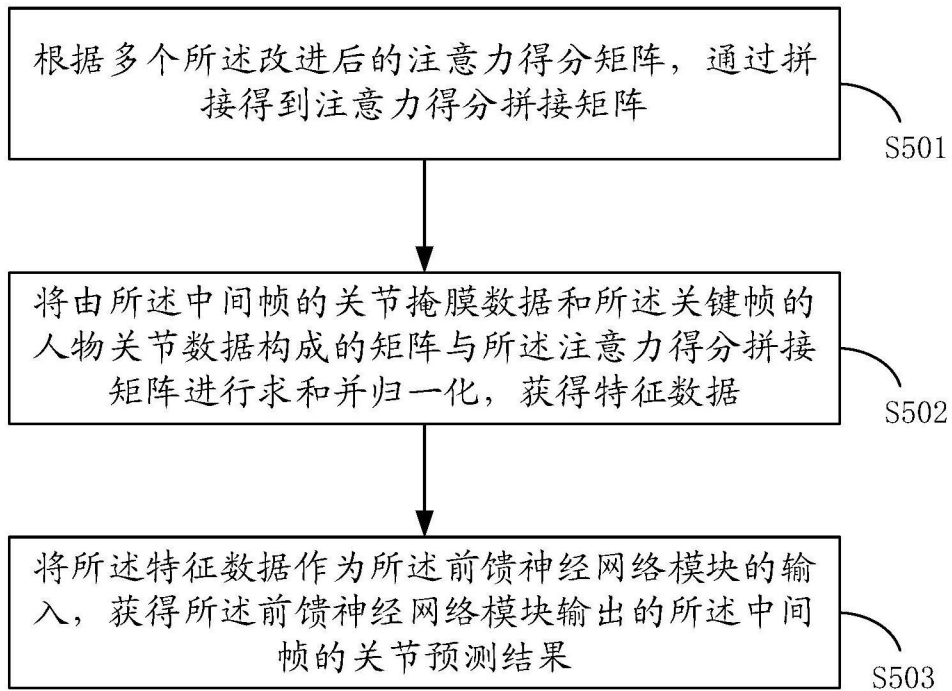


图5

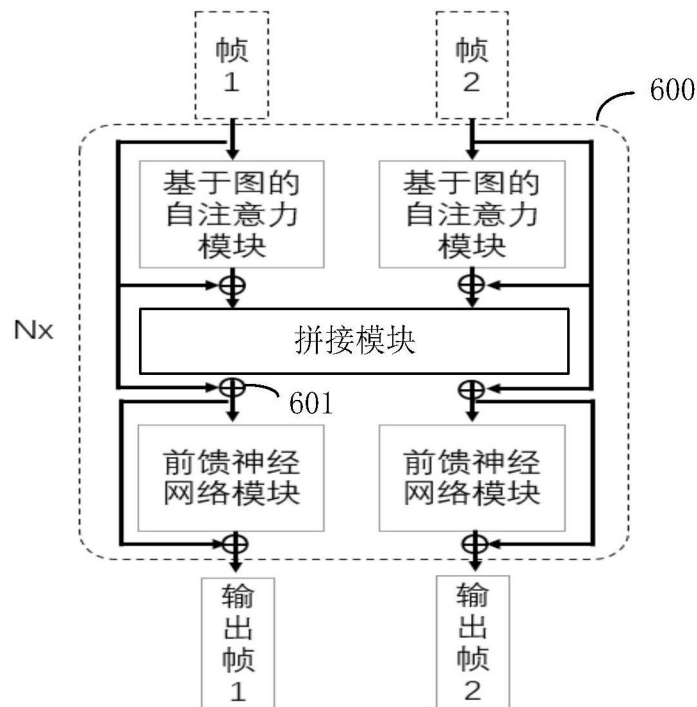


图6

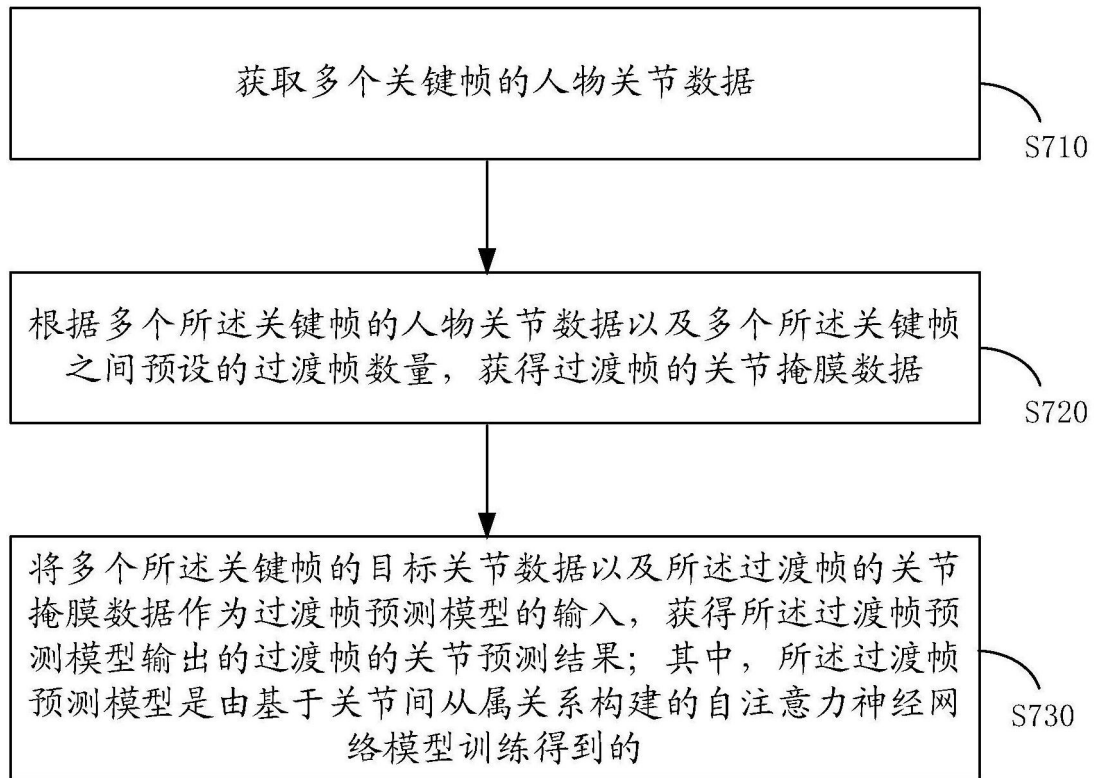


图7

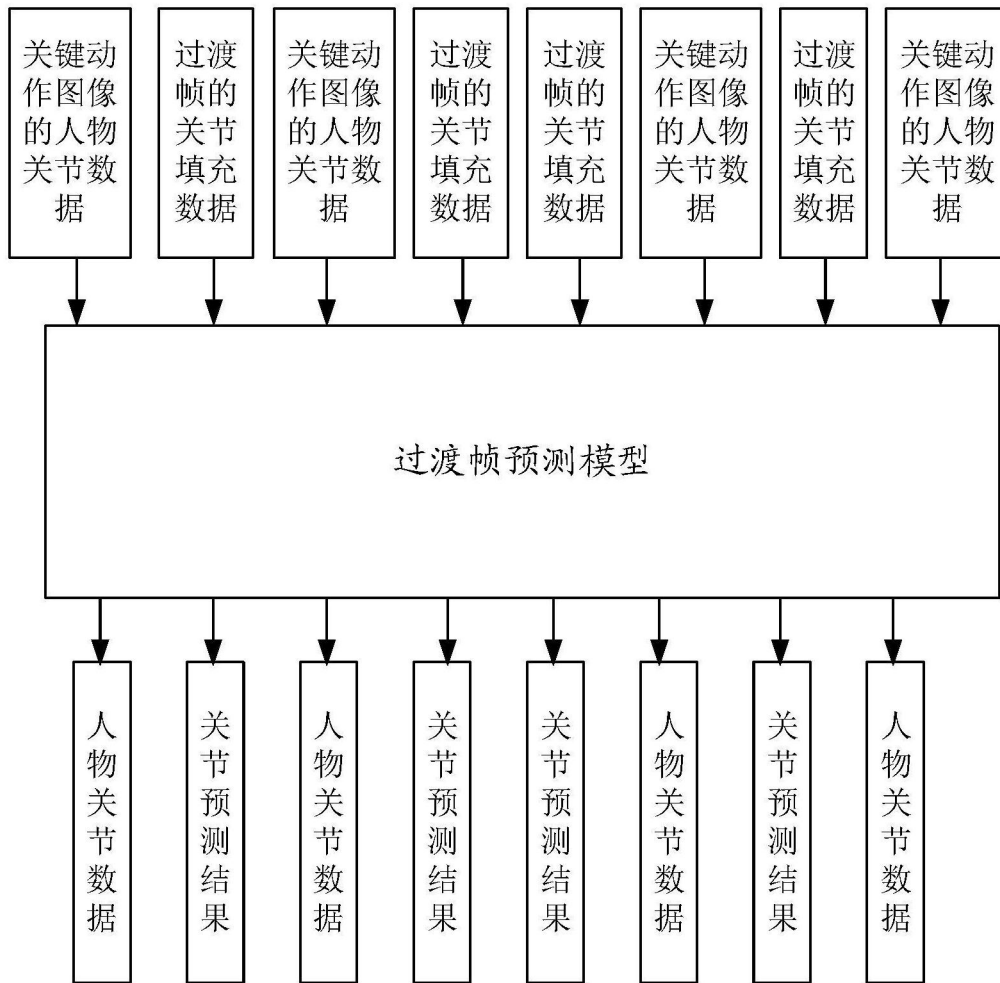


图8

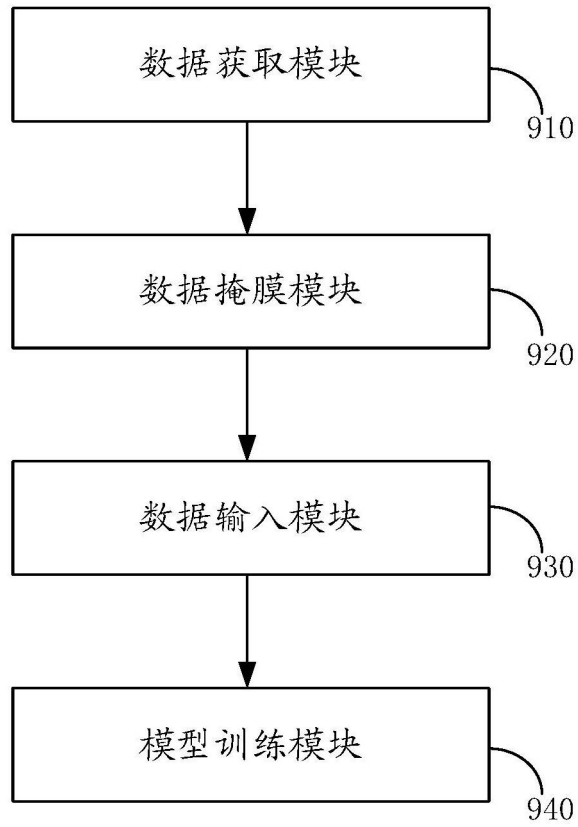


图9

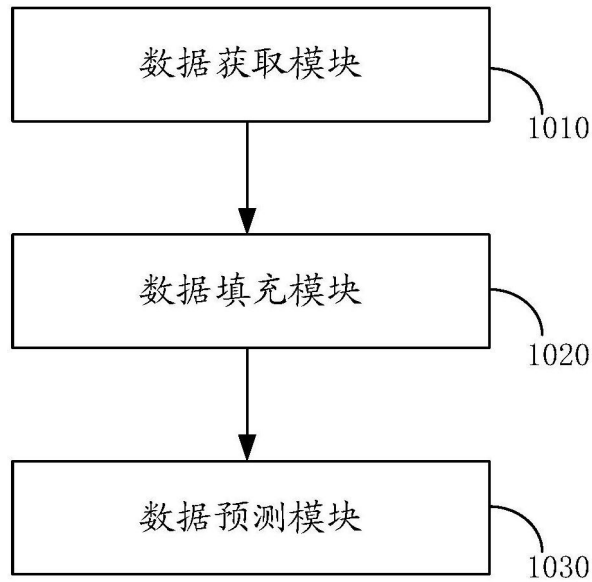


图10