



(12)发明专利申请

(10)申请公布号 CN 109821239 A

(43)申请公布日 2019.05.31

(21)申请号 201910127407.3

(22)申请日 2019.02.20

(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网商路  
599号网易大厦

(72)发明人 袁焱 范长杰

(74)专利代理机构 北京同立钧成知识产权代理  
有限公司 11205

代理人 张晓霞 刘芳

(51)Int.Cl.

A63F 13/55(2014.01)

A63F 13/213(2014.01)

G06T 7/70(2017.01)

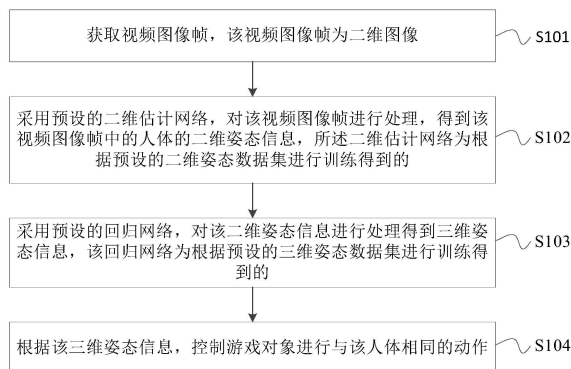
权利要求书2页 说明书11页 附图5页

(54)发明名称

体感游戏的实现方法、装置、设备及存储介  
质

(57)摘要

本发明提供一种体感游戏的实现方法、装  
置、设备及存储介质,该方法可包括获取视频图  
像帧,所述视频图像帧为二维图像;采用预设的  
二维估计网络,对所述视频图像帧进行处理,得  
到所述视频图像帧中人体的二维姿态信息;采用  
预设的回归网络,对所述二维姿态信息进行处理  
得到三维姿态信息;根据所述三维姿态信息,控  
制游戏对象进行与所述人体相同的动作。本发  
明可降低体感游戏的硬件限制,提高体感游戏  
的适用性。



1. 一种体感游戏的实现方法,其特征在于,包括:
  - 获取视频图像帧,所述视频图像帧为二维图像;
  - 采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,所述二维估计网络为根据预设的二维姿态数据集进行训练得到的;
  - 采用预设的回归网络,对所述二维姿态信息进行处理得到三维姿态信息,所述回归网络为根据预设的三维姿态数据集进行训练得到的;
  - 根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作。
2. 根据权利要求1所述的方法,其特征在于,所述采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息之前,所述方法还包括:
  - 对所述视频图像帧进行人体检测,得到至少一个包括所述人体的图像块,以及每个图像块对应的检测得分;
  - 根据所述至少一个包括所述人体的图像块中检测得分最高的一个图像块,从所述视频图像帧中确定目标图像块;
  - 所述采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,包括:
    - 采用所述二维估计网络,对所述目标图像块进行处理,得到所述二维姿态信息。
3. 根据权利要求2所述的方法,其特征在于,所述采用所述二维估计网络,对所述目标图像块进行处理,得到所述二维姿态信息,包括:
  - 采用所述二维估计网络,对所述目标图像块进行处理,得到多个关节点的概率热力图;
  - 根据每个关节点的概率热力图中概率最高的点的像素位置,确定所述每个关节点在所述目标图像块中的像素位置,所述二维姿态信息包括:所述多个关节点在所述目标图像块中的像素位置。
4. 根据权利要求1-3中任一项所述的方法,其特征在于,所述采用预设的回归网络,对所述二维姿态信息进行处理得到三维姿态信息,包括:
  - 对所述二维姿态信息进行处理,使得处理得到的所述二维姿态信息中人体骨骼长度为预设长度;
  - 采用所述回归网络,对所述处理得到的所述二维姿态信息进行处理,得到所述三维姿态信息。
5. 根据权利要求1-3中任一项所述的方法,其特征在于,所述获取视频图像帧包括:
  - 获取输入的视频序列,所述视频序列包括:多个所述视频图像帧。
6. 根据权利要求5所述的方法,其特征在于,所述根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作包括:
  - 对多个所述视频图像帧得到的三维姿态信息,进行滤波处理;
  - 根据滤波处理得到的三维姿态信息,对控制游戏对象进行与所述人体相同的动作。
7. 根据权利要求1-3中任一项所述的方法,其特征在于,所述根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作包括:
  - 获取所述游戏对象,在预设状态下的姿态信息;
  - 根据所述三维姿态信息,和所述预设状态下的姿态信息,得到所述游戏对象的骨骼旋转参数;

根据所述游戏对象的骨骼旋转参数,控制所述游戏对象的骨骼进行旋转,以进行与所述人体相同的动作。

8. 根据权利要求7所述的方法,其特征在于,所述骨骼旋转参数为如下任一类型的参数:欧拉角旋转参数、四元数旋转参数。

9. 一种体感游戏的实现装置,其特征在于,包括:

获取模块,用于获取视频图像帧,所述视频图像帧为二维图像;

第一处理模块,用于采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,所述二维估计网络为根据预设的二维姿态数据集进行训练得到的;

第二处理模块,用于采用预设的回归网络,对所述二维姿态信息进行处理得到三维姿态信息,所述回归网络为根据预设的三维姿态数据集进行训练得到的;

控制模块,用于根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作。

10. 一种游戏设备,其特征在于,包括:存储器及处理器;所述存储器与所述处理器连接;

所述存储器,用于存储程序指令;

所述处理器,用于在程序指令被执行时,实现权利要求1-8中任一项所述的体感游戏的实现方法。

11. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1-8中任一项所述的体感游戏的实现方法。

## 体感游戏的实现方法、装置、设备及存储介质

### 技术领域

[0001] 本发明涉及图像处理技术领域,尤其涉及一种体感游戏的实现方法、装置、设备及存储介质。

### 背景技术

[0002] 随着电子游戏技术的发展,电子游戏的种类也越来越丰富。其中,突破了以往单纯的手柄按键输入的操作方式,可通过肢体动作控制游戏对象的体感游戏,越来越受欢迎。

[0003] 对于体感游戏,可基于三维姿态信息进行游戏对象的驱动,由于三维图像中的人体姿态信息即为三维姿态信息。因而,可根据三维图像,控制游戏对象进行相应的动作。然而,体感摄像头也称三维摄像头,其可得到三维图像,除此之外的大多数摄像头所得到图像实际为二维图像。

[0004] 这使得体感游戏的实现必须基于体感摄像头进行,其适用性较差。

### 发明内容

[0005] 本发明提供一种体感游戏的实现方法、装置、设备及存储介质,以降低体感游戏的硬件限制,提高体感游戏的适用性。

[0006] 第一方面,本发明提供一种体感游戏的实现方法,所述方法包括:

[0007] 获取视频图像帧,所述视频图像帧为二维图像;

[0008] 采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,所述二维估计网络为根据预设的二维姿态数据集进行训练得到的;

[0009] 采用预设的回归网络,对所述二维姿态信息进行处理得到三维姿态信息,所述回归网络为根据预设的三维姿态数据集进行训练得到的;

[0010] 根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作。

[0011] 第二方面,本发明提供一种体感游戏的实现装置,包括:

[0012] 获取模块,用于获取视频图像帧,所述视频图像帧为二维图像;

[0013] 第一处理模块,用于采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,所述二维估计网络为根据预设的二维姿态数据集进行训练得到的;

[0014] 第二处理模块,用于采用预设的回归网络,对所述二维姿态信息进行处理得到三维姿态信息,所述回归网络为根据预设的三维姿态数据集进行训练得到的;

[0015] 控制模块,用于根据所述三维姿态信息,控制游戏对象进行与所述人体相同的动作。

[0016] 第四方面,本发明还提供一种游戏设备,包括:存储器及处理器;所述存储器与所述处理器连接;

[0017] 所述存储器,用于存储程序指令;

[0018] 所述处理器,用于在程序指令被执行时,实现上述第一方面的体感游戏的实现方

法。

[0019] 第五方面,本发明还可提供一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现权实现上述第一方面的体感游戏的实现方法。

[0020] 本发明提供一种体感游戏的实现方法、装置、设备及存储介质,可采用预设的二维估计网络,对获取的视频图像帧进行处理,得到该视频图像帧中人体的二维姿态信息,该视频图像帧为二维图像,该二维估计网络为根据预设的二维姿态数据集进行训练得到的,还采用预设的回归网络,对该二维姿态信息进行处理得到三维姿态信息,该回归网络为根据预设的三维姿态数据集进行训练得到的,并根据该三维姿态信息,控制游戏对象进行与该人体相同的动作。该方法可基于二维图像的视频图像帧,基于二维估计网络得到二维姿态信息,并通过回归网络得到三维姿态信息,并根据该三维姿态信息控制游戏对象进行与人体相同的动作,实现了基于二维图像对游戏对象的控制,无需获取拍摄参数等信息,降低了体感游戏的硬件限制,实现了体感游戏的通用性。

### 附图说明

[0021] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图做一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0022] 图1为本发明实施例提供的一种体感游戏的实现方法的流程图一;

[0023] 图2为本发明实施例提供的一种体感游戏的实现方法的流程图二;

[0024] 图3为本发明实施例提供的用于得到目标图像的检测网络的结构示意图;

[0025] 图4为本发明实施例提供的一种体感游戏的实现方法的流程图三;

[0026] 图5为本发明实施例提供的用于二维姿态估计的二维估计网络的结构示意图;

[0027] 图6为本发明实施例提供的一种体感游戏的实现方法的流程图四;

[0028] 图7为本发明实施例提供的用于三维姿态估计的回归网络的结构示意图;

[0029] 图8为本发明实施例提供的一种体感游戏的实现方法的流程图五;

[0030] 图9为本发明实施例提供的一种体感游戏的实现方法的流程图三;

[0031] 图10为本发明实施例提供的体感游戏的实现装置的结构示意图;

[0032] 图11为本发明一实施例提供的游戏设备的结构示意图。

### 具体实施方式

[0033] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0034] 需要说明的是,本发明实施例各部分及附图中的术语“第一”、“第二”及“第三”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不

排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0035] 本发明下述实施例所涉及的方法流程图仅是示例性说明,不是必须包括所有的内容和步骤,也不是必须按照所描述的顺序执行。例如,有些步骤还可以分解,而有些步骤可以合并或部分合并,因此,实际执行的顺序可根据实际情况改变。

[0036] 本发明下述实施例所涉及的方框图中的功能模块仅仅是功能实体,不一定必须与物理上独立的实体相对应。即,可以采用软件形式实现这些功能实体,或在一个或多个硬件模块或集成电路中实现这些功能实体,或者不同网络和/或处理器和/或微控制器中实现这些功能实体。

[0037] 如下结合多个示例,对本发明实施例所提供的体感游戏的实现方法进行说明。图1为本发明实施例提供的一种体感游戏的实现方法的流程图一。该体感游戏的实现方法可由游戏设备执行,该游戏设备可以为客户端设备或者服务器。客户端设备可以为安装有游戏应用程序的终端设备,该终端设备例如可以为台式电脑、笔记本、个人数字助理(Personal Digital Assistant,简称PDA)、智能手机、平板电脑等任一设备。服务器可以为该游戏应用程序的服务器。其中,该游戏应用程序可以为体感类游戏的应用程序。

[0038] 如图1所示,该方法可包括:

[0039] S101、获取视频图像帧,该视频图像帧为二维图像。

[0040] 该视频图像帧可以为体感摄像头所得到的图像帧,该视频图像帧可以为二维图像,该视频图像帧可不包括摄像头的拍摄参数等信息。若游戏设备为客户端设备,则游戏设备可获取该游戏设备的摄像头所得到的视频图像帧,或者,获取与该游戏设备连接的摄像头即该游戏设备外界的摄像头所得到的视频图像帧。若该游戏设备为服务器,则该游戏设备可从客户端设备获取该视频图像帧。无论何种类型的游戏设备,其仅需获取该视频图像帧,无需额外获取摄像头的拍摄参数等信息。

[0041] 通过执行S101可获取一个或多个视频图像帧。

[0042] S102、采用预设的二维估计网络,对该视频图像帧进行处理,得到该视频图像帧中的人体的二维姿态信息,所述二维估计网络为根据预设的二维姿态数据集进行训练得到的。

[0043] 该二维估计网络可以为根据预设的二维姿态数据集经过预设的训练周期进行训练得到的神经网络。该预设的二维姿态数据集例如可以为开源的人体二维姿态数据集,如上下文公共对象(Common Objects in Context,简称COCO)的二维人体姿态数据集。该预设的训练周期例如可以为350个训练周期(epoch)。

[0044] 该方法中,可将该视频图像帧,输入至该二维估计网络,以通过该二维估计网络,对该视频图像帧进行处理后,输出该视频图像帧中的人体的二维姿态信息。

[0045] 该二维姿态信息可包括:多个关节点在视频图像帧中的像素位置等信息。

[0046] S103、采用预设的回归网络,对该二维姿态信息进行处理得到三维姿态信息,该回归网络为根据预设的三维姿态数据集进行训练得到的。

[0047] 该回归网络可称为三维回归网络。该回归网络可以为根据预设的三维姿态数据集进行预设的训练周期进行训练所得到的神经网络。该预设的三维姿态数据集例如可以为开

源的三维人体姿态数据集,如Human3.6M的三维人体姿态数据集。该预设的训练周期例如可以为200个训练周期(epoch)。

[0048] 该方法中,可将该二维姿态信息,输入至该回归网络,以通过该回归网络,对该二维姿态信息进行处理后,输出该三维姿态信息。

[0049] 由于视频图像帧为二维图像,因此,该视频图像帧所得到的二维姿态信息中,每个关节点在该视频图像帧中的像素位置可包括:该每个关节点在该视频图像帧中两个维度的像素信息。

[0050] 而人体的动作并不限于两个维度上进行变化,其动作的变化实际为三维动作的变化,则执行该S103所得到的该三维姿态信息可包括:该多个关节点的像素位置,其中,每个关节点的像素位置可包括:该每个关节点在三个维度的像素信息。每个关节点的像素位置可包括:基于该每个关节点在该视频图像帧中两个维度的像素信息,采用该回归网络进行估计得到的所得到的,该每个关节点在该两个维度的像素信息。该每个关节点的像素位置还可包括:基于该每个关节点在该视频图像帧中两个维度的像素信息,采用该回归网络进行估计得到的该每个关节点在第三个维度的像素信息。

[0051] 也就是说,该S103中实际是根据二维姿态信息中每个关节点的二个维度的像素信息,采用该回归网络进行处理,得到该每个关节点在三个维度的像素位置,从而可得到包括多个关节点在三个维度的像素信息,即该三维姿态信息。

[0052] S104、根据该三维姿态信息,控制游戏对象进行与该人体相同的动作。

[0053] 该方法中,可根据该三维姿态信息,确定该游戏对象的骨骼旋转参数,并根据该骨骼旋转参数,控制该游戏对象的骨骼进行旋转,以控制该游戏对象进行与该人体相同的动作。通过执行本申请实施例提供的方法,可使得游戏对象进行与该人体相同的动作,则该方法还可成为一种动作迁移方法,即将该人体的动作迁移至游戏对象。

[0054] 若执行该方法的游戏设备为客户端设备,则该游戏设备可根据得到的该骨骼旋转参数控制游戏对象进行与人体相同的动作。该游戏对象可以为该客户端设备在应用界面所显示的虚拟对象。

[0055] 若执行该方法的游戏设备为服务器,则该游戏设备可将得到的该骨骼旋转参数发送至客户端设备,使得客户端设备根据该骨骼旋转参数控制游戏对象进行与人体相同的动作。该游戏对象可以为该客户端设备在应用界面所显示的虚拟对象。

[0056] 需要说明的是,对于每个视频图像帧,均可执行上述S101-S103,得到每个图像帧的三维姿态信息,并根据至少一个视频图像帧的三维姿态信息,执行该S104,以控制游戏对象进行与人体相同的动作。

[0057] 本发明实施例提供的体感游戏的实现方法,可采用预设的二维估计网络,对获取的视频图像帧进行处理,得到该视频图像帧中人体的二维姿态信息,该视频图像帧为二维图像,该二维估计网络为根据预设的二维姿态数据集进行训练得到的,还采用预设的回归网络,对该二维姿态信息进行处理得到三维姿态信息,该回归网络为根据预设的三维姿态数据集进行训练得到的,并根据该三维姿态信息,控制游戏对象进行与该人体相同的动作。该方法可基于二维图像的视频图像帧,基于二维估计网络得到二维姿态信息,并通过回归网络得到三维姿态信息,并根据该三维姿态信息控制游戏对象进行与人体相同的动作,实现了基于二维图像对游戏对象的控制,无需获取拍摄参数等信息,降低了体感游戏的硬件

限制,实现了体感游戏的通用性。

[0058] 可选的,在上述图1所示的方法的基础上,本发明实施例还可提供一种体感游戏的实现方法。图2为本发明实施例提供的一种体感游戏的实现方法的流程图二。如图2所示,在上述S102中采用预设的二维估计网络,对该视频图像帧进行处理,得到该视频图像帧中的人体的二维姿态信息之前,该方法还可包括:

[0059] S201、对该视频图像帧进行人体检测,得到至少一个包括该人体的图像块,以及每个图像块对应的检测得分。

[0060] 该方法可根据预设的检测网络,对该视频图像帧进行人体检测,得到至少一个包括该人体的图像块,以及每个图像块对应的检测得分。具体地,可将该视频图像帧输入该检测网络,以通过该检测网络对该视频图像帧进行人体检测。该检测网络例如可以为人体检测网络。在该实施例中,该人体检测可以为人体的轮廓检测。

[0061] 该检测网络可以为根据预设的目标检测数据集进行预设的训练周期进行训练得到的网络。该目标检测数据集可以为人体检测数据集,其例如可以为开源的人体检测数据集,如COCO的人体检测数据集。该预设的训练周期例如可以为50个训练周期(epoch)

[0062] 如下结合示例进行说明。图3为本发明实施例提供的用于得到目标图像的检测网络的结构示意图。该检测网络可以为图3所示的一系列的卷积层和残差模块构成的深度神经网络。如图3所示,该检测网络中卷积(Convolutional)层可包括具有连续的多个卷积层及残差模块的卷积通道、池化(avgpool)函数、全连接(Connected)层及软最大(Softmax)函数。卷积层中依次包括:两个不同尺寸的卷积单元,以及残差(Residual)模块,这两个不同尺寸的卷积单元可包括:尺寸为1X1的卷积单元,和尺寸为3X3的卷积单元。

[0063] 参照图3,例如可将416X416的视频图像帧输入该检测网络中卷积核(Filters)为32,卷积单元的尺寸(Size)为3X3的卷积层进行处理,得到256X256的图像块;将该256X256的图像块输入至卷积核为64,卷积单元的尺寸为3X3或3X2的卷积层进行处理,得到128X128的图像块;还将该128X128的图像块输入至一个卷积通道中,在该一个卷积通道中,由卷积核为32,卷积单元的尺寸为1X1的卷积层进行处理后,将得到的图像块输入至卷积单元的卷积核为64,尺寸为3X3的卷积层进行处理后,输入至残差模块,得到128X128的图像块;还可将该128X128的图像块输入卷积核为128,卷积单元的尺寸为3X3或3X2的卷积层进行处理,得到64X64的图像块;还将该64X64的图像块输入至2个卷积通道中,在每个卷积通道中,由卷积核为64,卷积单元的尺寸为1X1的卷积层进行处理后,将得到的图像块输入至卷积单元的卷积核为128,尺寸为3X3的卷积层进行处理后,输入至残差模块,得到64X64的图像块;还将该64X64的图像块输入至卷积核为256,卷积单元的尺寸为3X3或3X2的卷积层进行处理,得到32X32的图像块;还将32X32的图像块输入至8个卷积通道中,在每个卷积通道中由卷积核为128,卷积单元的尺寸为1X1的卷积层进行处理后,将得到的图像块输入至卷积单元的卷积核为256,尺寸为3X3的卷积层进行处理后,输入至残差模块,得到32X32的图像块;还将该32X32的图像块输入至卷积核为512,卷积单元的尺寸为3X3或3X2的卷积层进行处理,得到16X16的图像块;还将16X16的图像块输入至8个卷积通道中,在每个卷积通道中由卷积核为256,卷积单元的尺寸为1X1的卷积层进行处理后,将得到的图像块输入至卷积单元的卷积核为512尺寸为3X3的卷积层进行处理后,输入至残差模块,得到16X16的图像块;还将该16X16的图像块输入至卷积核为1024,卷积单元的尺寸为3X3或3X2的卷积层进行处理,得到



8X8的图像块;还将8X8的图像块输入至4个卷积通道中,在每个卷积通道中由卷积核为512,卷积单元的尺寸为1X1的卷积层进行处理后,将得到的图像块输入至卷积单元的卷积核为1024尺寸为3X3的卷积层进行处理后,输入至残差模块,得到8X8的图像块。

[0064] 如上输入至检测网络的视频图像帧例如可以为416X416的二维图像。通过图3所示的检测网络,便可得到至少一个包括该人体的图像块,每个图像块可以为8X8的图像块。

[0065] 在得到8X8的图像块后,还可基于池化参数各卷积通道得到的8X8的图像块,均进行池化处理,并在池化处理后输入至全连接层,由该全连接层在处理后的输入至软最大函数,由该软最大函数进行处理,以确定每个图像块对应的检测得分。

[0066] 通过上述检测网络,可得到至少一个包括人体的图像块以及该每个图像块的检测得分。每个图像块的检测得分可以为该每个图像块的人体检测得分。

[0067] S202、根据该至少一个包括该人体的图像块中检测得分最高的一个图像块,从该视频图像帧中确定目标图像块。

[0068] 该方法中,可根据该检测得分最高的一个图像块为中心,从该视频图像帧确定尺寸为预设尺寸的图像块,确定为该目标图像块。该预设尺寸例如可以为 $\min(1.2w, 1.2h)$ 的矩形框的尺寸,其中,w可以为该得分最高的一个图像块的宽度,该h可以为该得分最高的一个图像块的高度。

[0069] 如上所示的S102中采用预设的二维估计网络,对该视频图像帧进行处理,得到该视频图像帧中的人体的二维姿态信息可包括:

[0070] S203、采用该二维估计网络,对该目标图像块进行处理,得到该二维姿态信息。

[0071] 在得到该目标图像块的情况下,可直接将其输入该二维估计网络,以采用该二维估计网络进行处理,得到该二维姿态信息;也可对该目标图像块进行尺寸重置,并将尺寸重置后的图像块输入该二维估计网络,以采用该二维估计网络进行处理,得到该二维姿态信息。该尺寸重置后的图像块例如可以为384X288的图像块。

[0072] 该方法中从该视频图像帧确定目标图像块,并将该目标图像块该二维估计网络的输入图像块进行二维姿态估计,可针对性地进行二维姿态估计,简化二维姿态估计的计算量,保证了二维姿态估计结果的准确性。

[0073] 可选的,在上述图2所示方法的基础上,本发明实施例还可提供一种体感游戏的实现方法。图4为本发明实施例提供的一种体感游戏的实现方法的流程图三。如图4所示,如上所示的S203中采用该二维估计网络,对该目标图像块进行处理,得到该二维姿态信息,可包括:

[0074] S401、采用该二维估计网络,对该目标图像块进行处理,得到多个关节的概率热力图。

[0075] 该方法中可将该目标图像块输入至该二维估计网络,以采用该二维估计网络,对该目标图像块进行处理,得到多个关节的概率热力图。该多个关节例如可包括如下中的至少两个关节:左眼关节、右眼关节、鼻关节、左耳关节、右耳关节、左肩关节、右肩关节、左肘关节、右肘关节、左腕关节、右腕关节、左臂关节、右关节、左膝关节、右膝关节、左踝关节以及右踝关节。

[0076] 如下结合示例进行说明。图5为本发明实施例提供的用于二维姿态估计的二维估计网络的结构示意图。如图5所示,二维估计网络可包括:全局网络(GlobalNet)和精细网络

(RefineNet)。全局网络中包括:多个卷积层,其中,每个卷积层包括多个卷积模块,相邻两个卷积模块构成一个残差模块。若卷积模块的个数为5,其在,则在每个卷积层中可将第2至第5个卷积模块的残差模块中最后一个残差模块,即最后两个卷积模块构成的残差模块的输出作为该每个卷积层的输入。在该每个卷积层中,可将该最后一个残差模块输出的特征图(feature map)进行1X1的卷积,再进行2倍下采样,随后进行3X3的卷积,最后再进行采样,得到输出尺寸为96X72的图像块,为每个卷积层的检测结构。该方法中,还可将各卷积层的预测结果再进行元素级的相加,得到全局网络输出的关节点热力图。在该全局网络中,还根据预设的损失函数,对每个卷积层所得到的关节点的概率热力图进行处理,以得到每个卷积层所得到的该多个关节点中损失最大的至少一个关节点。该损失函数可以为L2损失函数。该损失最大的至少一个关节点可以为二维姿态估计结果不准确的至少一个关节点。该至少一个关节点例如可以为N个关节点,N可以为8。

[0077] 该方法中还可将每个卷积层得到的该损失最大的至少一个关节点的概率热力图传输至精细网络中的瓶颈(bottleneck)结构,由该瓶颈结构处理后,进行A倍的采样,并将该多个卷积层所得到的关节点的概率热力图通过特征拼接(concatenation)后,依次经过一个瓶颈结构和卷积层后,输出该多个关节点的概率热力图。该多个关节点的概率热力图可以为该二维估计网络输出的概率热力图。

[0078] S402、根据每个关节点的概率热力图中概率最高的点的像素位置,确定该每个关节点在该目标图像块中的像素位置,该二维姿态信息包括:该多个关节点在该目标图像块中的像素位置。

[0079] 在得到每个关节点的概率热力图的情况下,可将每个关节点的概率热力图中概率最高的点的像素位置,确定为该每个关节点在该目标图像块中的像素位置。该概率最高的点的像素位置也可称为响应最高的点。

[0080] 若该每个关节点在该目标图像块中的像素位置为 $[u_i, v_i]$ ,其中,u和v分别为两个维度的像素位置,对于包括17关节点在该目标图像块中的像素位置的二维姿态信息可以为34维的向量,其可表示为 $[u_1, v_1, u_2, v_2, \dots, u_{17}, v_{17}]^T$ 。

[0081] 该实施例中,可采用该二维估计网络,对该目标图像块进行处理,得到多个关节点的概率热力图,并根据每个关节点的概率热力图中概率最高的点的像素位置,确定该每个关节点在所述目标图像块中的像素位置,从而得到包括有该多个关节点在该目标图像块中的像素位置的二维姿态信息,保证了二维姿态估计所得到的二维姿态信息的准确性。

[0082] 可选的,在上述图1-图4中任一所述的方法的基础上,本发明实施例还可提供一种体感游戏的实现方法。图6为本发明实施例提供的一种体感游戏的实现方法的流程图四。如图6所示,在上述S103中采用预设的回归网络,对该二维姿态信息进行处理得到三维姿态信息,可包括:

[0083] S601、对该二维姿态信息进行处理,使得处理得到的该二维姿态信息中人体骨骼长度为预设长度。

[0084] 该方法中,可将该二维姿态信息中一个骨骼的两个关节点在同一维度的像素位置的差值,缩放为预设值,并将所有关节点的像素位置与预设的缩放系数相乘,实现了对该二维姿态信息的归一化处理,使得处理后的该二维姿态信息中人体骨骼长度为预设长度。

[0085] S602、采用该回归网络,对该处理得到的该二维姿态信息进行处理,得到该三维姿

态信息。

[0086] 如下结合示例进行说明。图7为本发明实施例提供的用于三维姿态估计的回归网络的结构示意图。回归网络的输入层,可连接1024维的全连接层,而后连接图7所示的两个处理块(Block),再连接一个51维的全连接层后输出。图7所示的每个处理块可包括依次连接的:1024维的线性函数、归一化(Batch norm)函数、激活函数、丢取(Dropout)函数、1024维的线性函数、归一化(Batch norm)函数、激活函数、丢取函数等。激活函数可以为修正线性单元(Rectified Linear Units,RELU)函数。在每个处理块内,可通过1024维的线性函数对接收到的二维姿态信息进行处理后,由归一化函数进行归一化处理,并由RELU函数进行处理后传输至丢取函数,丢取函数处理后的信息传输至另一个1024维的线性函数进行再次处理后,由另一个归一化函数进行归一化处理,并由另一个RELU函数进行处理后传输至另一个丢取函数。该方法中,可通过总数函数,将该另一个丢取函数的输出以及输入至该每个处理块的二维姿态信息之后,确定为该每个处理后的输出的三维姿态信息。

[0087] 其中, $x$ 、 $y$ 和 $z$ 分别为得到的三维姿态信息中每个关节在三个维度的位置,对于包括17关节,该三维姿态信息可以为51维的向量,其可表示为 $[x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_{17}, y_{17}, z_{17}]^T$ 。

[0088] 该方法中,采用该回归网络,对该处理得到的该二维姿态信息进行处理,得到该三维姿态信息,可避免人体个体差异对三维姿态估计的影响,提高了三维姿态估计的准确性,从而保证对游戏对象的精确控制。

[0089] 在上述任一实施例所示的方法的基础上,本发明实施例还可提供一种体感游戏的实现方法。可选的,如上所示的S101中获取视频图像帧可包括:

[0090] 获取输入的视频序列,该视频序列包括:多个该视频图像帧。

[0091] 在获取该视频序列后,可对该视频序列进行处理,以得到多个视频图像帧,针对每个视频图像帧可执行上述方法以确定三维姿态信息。

[0092] 该方法中,可根据多个视频图像帧所得到的三维姿态信息,控制游戏对象进行与所述人体相同的动作;也可对该多个视频图像帧所得到的三维姿态信息进行处理后,基于处理后的信息,控制游戏对象进行与所述人体相同的动作。

[0093] 图8为本发明实施例提供的一种体感游戏的实现方法的流程图五。如图8所示,如上所示的S104中根据该三维姿态信息,控制游戏对象进行与该人体相同的动作可包括:

[0094] S801、对多个该视频图像帧得到的三维姿态信息,进行滤波处理。

[0095] 该方法中,可对多个该视频图像帧得到的三维姿态信息中,三个维度的信息中,每个维度的信息进行滤波处理。该滤波处理例如可以为高斯滤波,该高斯滤波的高斯核可以为2。

[0096] S802、根据滤波处理得到的三维姿态信息,对控制游戏对象进行与该人体相同的动作。

[0097] 该滤波处理后的三维姿态信息中,每个维度的信息可在时序上平滑变动,可有效避免抖动或不连续问题,保证了对游戏对象的动作控制的连续性,提高用户体验。

[0098] 在上述任一方法的基础上,本发明实施例还可提供一种体感游戏的实现方法。图9为本发明实施例提供的一种体感游戏的实现方法的流程图六。如图9所示,如上所示的S104中根据该三维姿态信息,控制游戏对象进行与该人体相同的动作可包括:

[0099] S901、获取该游戏对象,在预设状态下的姿态信息。

[0100] 该预设状态可以为该游戏对象的初始状态,如T姿态 (pose)。该预设状态下的姿态信息可以为该游戏对象在该游戏场景中的位置信息。

[0101] S902、根据该三维姿态信息,和该预设状态下的姿态信息,得到该游戏对象的骨骼旋转参数。

[0102] 可选的,该骨骼旋转参数为如下任一类型的参数:欧拉角旋转参数、四元数旋转参数。

[0103] 如下以四元数旋转参数为例,该方法中,首先根据胯部的左右骨骼点位置计算整个人体的旋转参数,然后再计算该每个关节点相对旋转参数,根据四元数的乘法,用人体的旋转参数乘以每个关节点的相对旋转参数从而能够得到每个关节点的旋转参数,这里的旋转参数均为四元数旋转参数。其中,该每个关节的相对旋转量可采用下述方法得到:根据该三维姿态信息中每个关节点的三维信息以及该每个关节点的父节点的三维信息,确定该每个关节点与该父节点构成的骨骼相对于该预设状态下该骨骼的旋转角度与旋转轴的信息;并根据每个关节点与该父节点构成的骨骼相对于该预设状态下该骨骼的旋转角度与旋转轴的信息,计算每个关节点的相对旋转量。

[0104] 该游戏对象的骨骼旋转参数包括:多个关节点的旋转参数。

[0105] S903、根据该游戏对象的骨骼旋转参数,控制该游戏对象的骨骼进行旋转,以进行与该人体相同的动作。

[0106] 该方法中,可在针对多个视频图像帧进行处理,得到多个视频图像帧的三维姿态信息,并分别基于该每个视频图像帧的三维姿态信息,确定该每个视频图像帧对应的游戏对象的骨骼旋转参数,从而得到多个视频图像帧对应的游戏对象的骨骼旋转参数,控制该游戏对象的骨骼进行旋转,以进行与该人体相同的动作。

[0107] 该实施例提供的方法,可获取该游戏对象在预设状态下的姿态信息,并根据该三维姿态信息和该预设状态下的姿态信息,得到该游戏对象的骨骼旋转参数,从而根据该游戏对象的骨骼旋转参数,控制该游戏对象的骨骼进行旋转,以进行与该人体相同的动作,使得人体的动作准确迁移至游戏对象,实现了对游戏对象的动作的精准控制,提高了用户体验。

[0108] 下述为本发明装置实施例,可以用于执行本发明上述方法实施例,其实现原理和技术效果类似。

[0109] 图10为本发明实施例提供的体感游戏的实现装置的结构示意图。如图10所示,本实施例的体感游戏的实现装置100可以包括:

[0110] 获取模块101,用于获取视频图像帧,该视频图像帧为二维图像。

[0111] 第一处理模块102,用于采用预设的二维估计网络,对该视频图像帧进行处理,得到该视频图像帧中人体的二维姿态信息,该二维估计网络为根据预设的二维姿态数据集进行训练得到的。

[0112] 第二处理模块103,用于采用预设的回归网络,对该二维姿态信息进行处理得到三维姿态信息,该回归网络为根据预设的三维姿态数据集进行训练得到的。

[0113] 控制模块104,用于根据该维姿态信息,控制游戏对象进行与该人体相同的动作。

[0114] 可选的,体感游戏的实现装置100还包括:

[0115] 检测模块,用于在该第一处理模块102采用预设的二维估计网络,对该视频图像帧

进行处理,得到该视频图像帧中人体的二维姿态信息之前,对该视频图像帧进行人体检测,得到至少一个包括该人体的图像块,以及每个图像块对应的检测得分;根据该至少一个包括该人体的图像块中检测得分最高的一个图像块,从该视频图像帧中确定目标图像块。

[0116] 所述采用预设的二维估计网络,对所述视频图像帧进行处理,得到所述视频图像帧中人体的二维姿态信息,包括:

[0117] 第一处理模块102,具体用于采用该二维估计网络,对该目标图像块进行处理,得到该二维姿态信息。

[0118] 可选的,第一处理模块102,具体用于采用该二维估计网络,对该目标图像块进行处理,得到多个关节点的概率热力图;根据每个关节点的概率热力图中概率最高的点的像素位置,确定该每个关节点在该目标图像块中的像素位置,该二维姿态信息包括:所述多个关节点在该目标图像块中的像素位置。

[0119] 可选的,第二处理模块103,具体用于对该二维姿态信息进行处理,使得处理得到的该二维姿态信息中人体骨骼长度为预设长度;采用该回归网络,对该处理得到的该二维姿态信息进行处理,得到该三维姿态信息。

[0120] 可选的,获取模块101,还用于获取输入的视频序列,该视频序列包括:多个该视频图像帧。

[0121] 可选的,控制模块104,具体用于对多个该视频图像帧得到的三维姿态信息,进行滤波处理;根据滤波处理得到的三维姿态信息,对控制游戏对象进行与该人体相同的动作。

[0122] 可选的,控制模块104,具体用于获取该游戏对象,在预设状态下的姿态信息;根据该三维姿态信息,和该预设状态的姿态信息,得到该游戏对象的骨骼旋转参数;根据该游戏对象的骨骼旋转参数,控制该游戏对象的骨骼进行旋转,以进行与该人体相同的动作。

[0123] 可选的,该骨骼旋转参数为如下任一类型的参数:欧拉角旋转参数、四元数旋转参数。

[0124] 可选的,体感游戏的实现装置100还可执行上述图1-图9中任一方法中的其它操作,具体参加上述,在此不再赘述

[0125] 图11为本发明一实施例提供的游戏设备的结构示意图。如图11所示,本实施例的游戏设备110包括:存储器111和处理器112。其中,存储器111通过总线与处理器112连接。

[0126] 存储器111,用于存储程序指令。

[0127] 处理器112,用于在程序指令被执行时,执行上述图1-图9中任一方法中体感游戏的实现方法。

[0128] 本发明实施例还提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序可被上述图11所述的处理器112执行实现上述体感游戏的实现方法。

[0129] 本发明实施例提供的体感游戏的实现装置、游戏设备和计算机可读存储介质,可执行上述图1-图9中任一方法中的体感游戏的实现方法,其具体实现及有效效果,可参见上述,在此不再赘述。

[0130] 本领域普通技术人员可以理解:实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的计算机程序可以存储于一计算机可读取存储介质中。该程序在执行时,执行包括上述各方法实施例的步骤;而前述的存储介质包括:只读内存(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、磁碟或者光

盘等各种可以存储程序代码的介质。

[0131] 最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

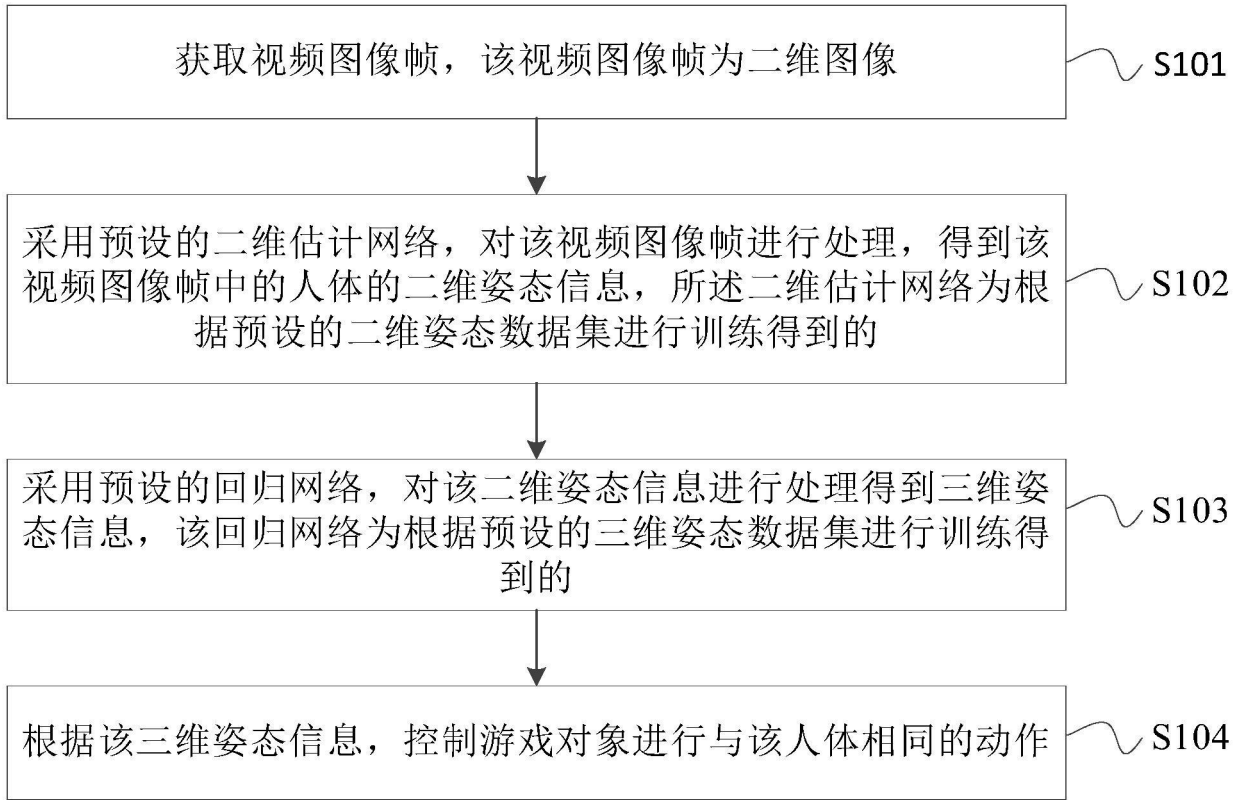


图1

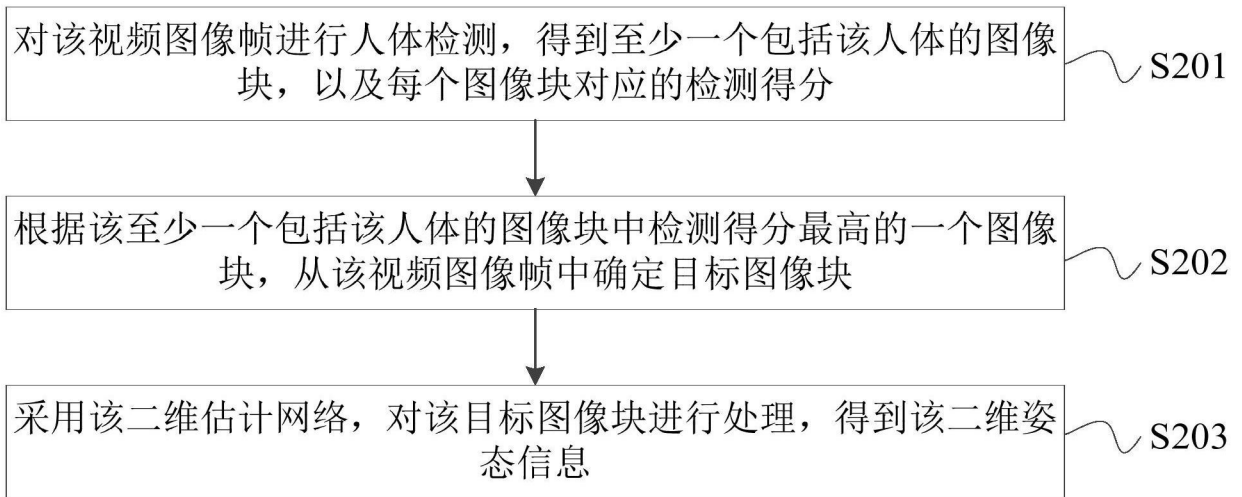


图2

	类型	卷积核	尺寸	输出
	卷积	32	3X3	256X256
	卷积	64	3X3/2	128X128
1X	卷积	32	1X1	
	卷积	64	3X3	
	残差			128X128
	卷积	128	3X3/2	64X64
2X	卷积	64	1X1	
	卷积	128	3X3	
	残差			64X64
	卷积	256	3X3/2	32X32
8X	卷积	128	1X1	
	卷积	256	3X3	
	残差			32X32
	卷积	512	3X3/2	16X16
8X	卷积	256	1X1	
	卷积	512	3X3	
	残差			16X16
	卷积	1024	3X3/2	8X8
4X	卷积	512	1X1	
	卷积	1024	3X3	
	残差			8X8
	池化		全局	
	连接		1000	
	软最大			

图3

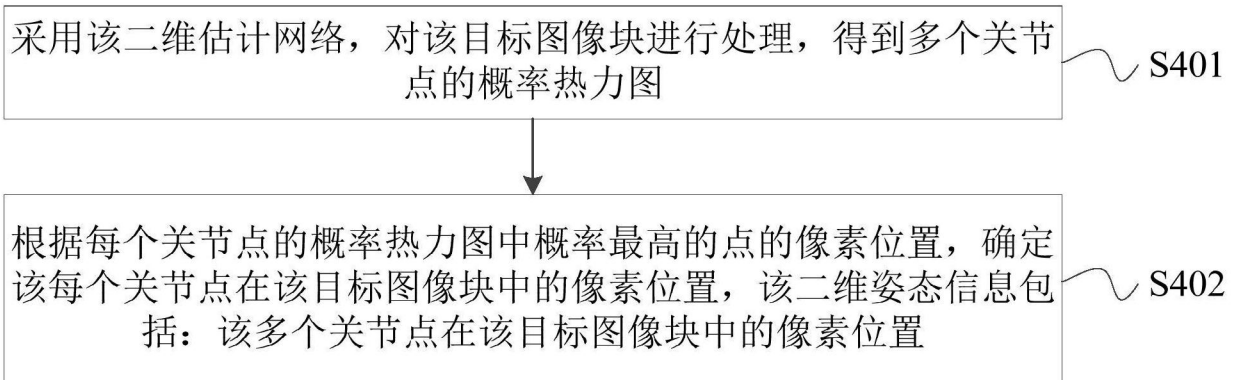


图4



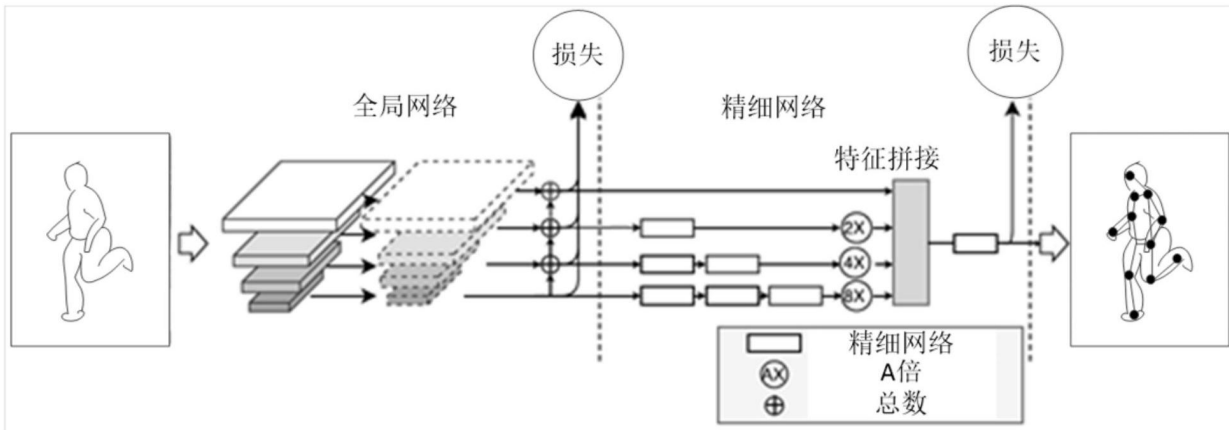


图5

对该二维姿态信息进行处理，使得处理得到的该二维姿态信息中人体骨骼长度为预设长度

S601

采用该回归网络，对该处理得到的该二维姿态信息进行处理，得到该三维姿态信息

S602

图6

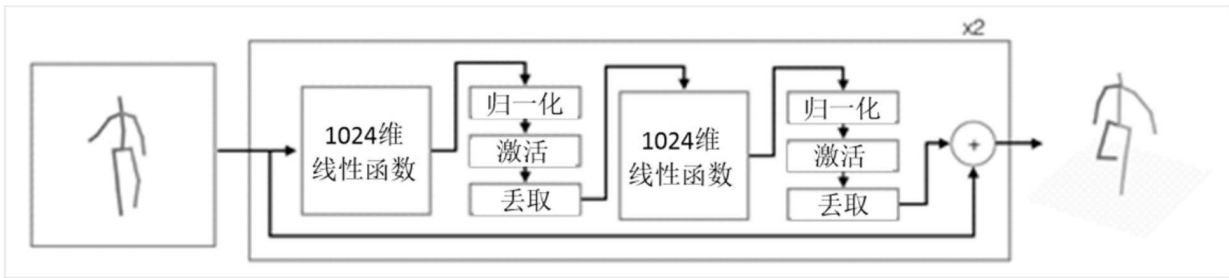


图7

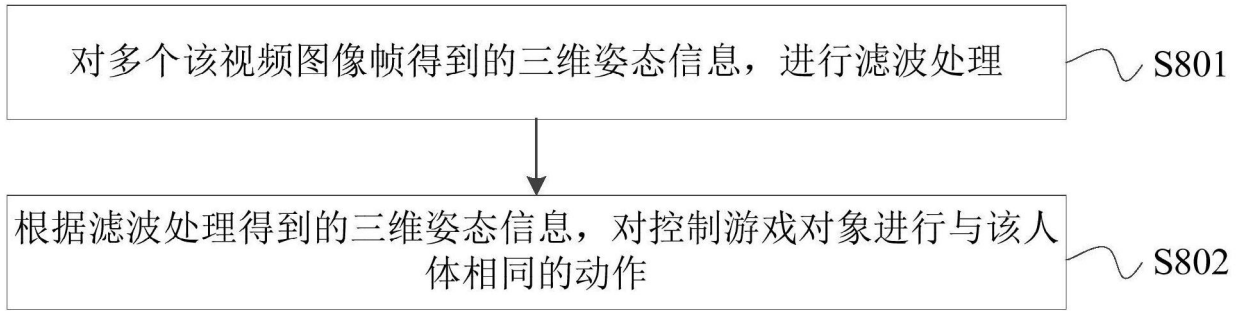


图8

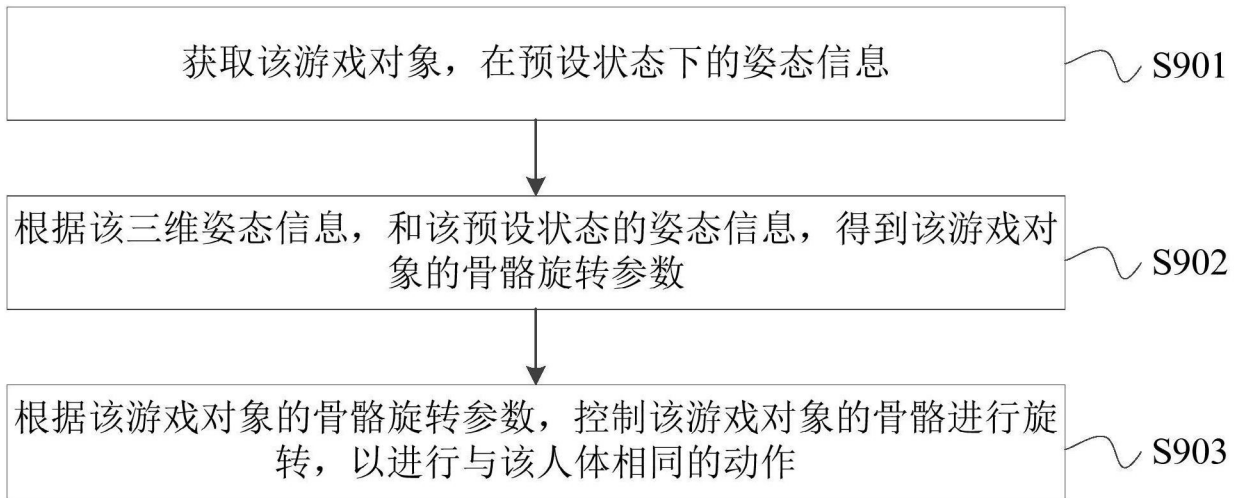


图9

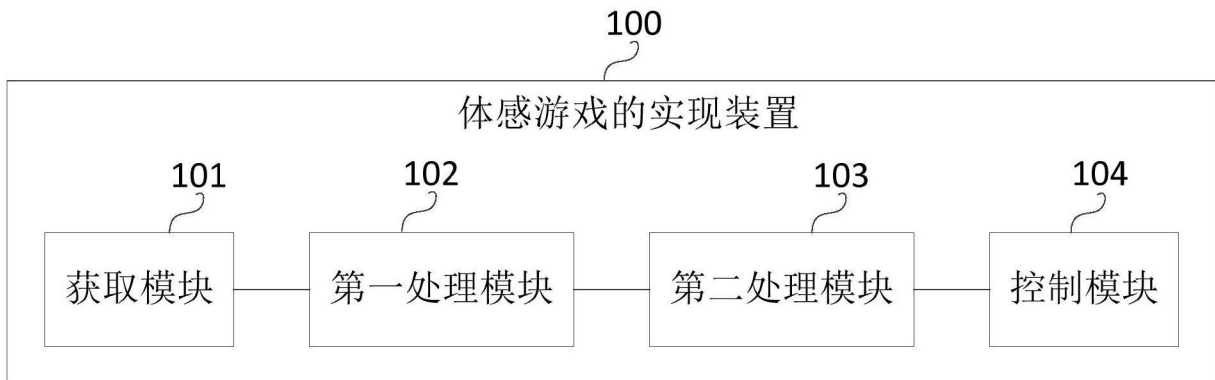


图10

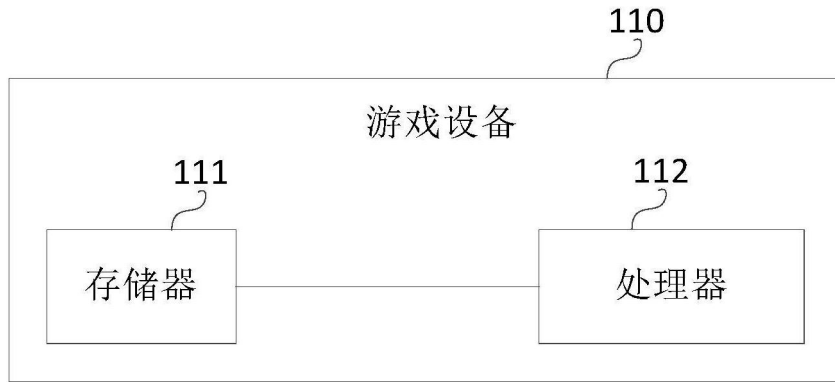


图11