

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2021年10月7日 (07.10.2021)



(10) 国际公布号  
**WO 2021/197143 A1**

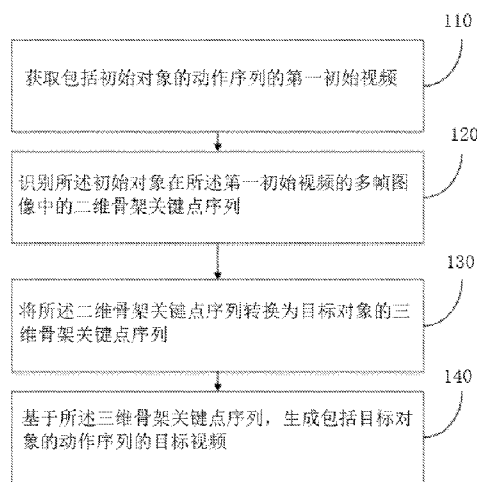
- (51) 国际专利分类号:  
**G06T 7/579** (2017.01)
- (21) 国际申请号: PCT/CN2021/082407
- (22) 国际申请日: 2021年3月23日 (23.03.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
202010243906.1 2020年3月31日 (31.03.2020) CN
- (71) 申请人: 北京市商汤科技开发有限公司(BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) [CN/CN]; 中国北京市海淀区北四环西路58号11层1101-1117室, Beijing 100080 (CN)。
- (72) 发明人: 吴文岩(WU, Wenyan); 中国北京市海淀区北四环西路58号11层1101-1117室, Beijing 100080 (CN)。 朱文韬(ZHU, Wentao); 中国北

京北京市海淀区北四环西路58号11层1101-1117室, Beijing 100080 (CN)。 杨卓谦(YANG, Zhuoqian); 中国北京市海淀区北四环西路58号11层1101-1117室, Beijing 100080 (CN)。

- (74) 代理人: 北京博思佳知识产权代理有限公司(BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街9号嘉华大厦B座409, Beijing 100085 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: MOTION TRANSFER METHOD AND APPARATUS, AND DEVICE AND STORAGE MEDIUM

(54) 发明名称: 动作迁移方法、装置、设备及存储介质



- 110 Obtain a first initial video comprising a motion sequence of an initial object  
120 Identify a two-dimensional framework key point sequence of the initial object in multiple image frames of the first initial video  
130 Convert the two-dimensional framework key point sequence into a three-dimensional framework key point sequence of a target object  
140 On the basis of the three-dimensional framework key point sequence, generate a target video comprising a motion sequence of the target object

图 1

(57) Abstract: The present disclosure provides a motion transfer method and apparatus, and a device and a storage medium. The method comprises: first obtaining a first initial video comprising a motion sequence of an initial object; then identifying a two-dimensional framework key point sequence of the initial object in multiple image frames of the first initial video; then converting the two-dimensional framework key point sequence into a three-dimensional framework key point sequence of a target object; and finally, on the basis of the three-dimensional framework key point sequence, generating a target video comprising a motion sequence of the target object.

(57) 摘要: 本公开提供了一种动作迁移方法、装置、设备及存储介质, 其中首先获取包括初始对象的动作序列的第一初始视频; 之后, 识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列; 再将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列; 最后, 基于所述三维骨架关键点序列, 生成包括目标对象的动作序列的目标视频。



WO 2021/197143 A1

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,  
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,  
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区  
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,  
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,  
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,  
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,  
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,  
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,  
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

## 动作迁移方法、装置、设备及存储介质

### 技术领域

[01]本公开涉及计算机视觉技术领域，具体而言，涉及一种动作迁移方法、装置、设备及存储介质。

### 背景技术

[02]动作迁移是将初始运动视频中初始对象的动作迁移到目标对象上，以形成目标运动视频。由于初始运动视频和目标运动视频存在很大的结构和视角差异，很难在像素级别上实现动作的迁移。尤其在初始对象做出极端动作，或者初始对象和目标对象的结构差异比较大时，迁移到目标对象上的动作准确度较低。

### 发明内容

[03]有鉴于此，本公开至少提供一种动作迁移方法及装置。

[04]第一方面，本公开提供了一种动作迁移方法，包括：获取包括初始对象的动作序列的第一初始视频；识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列；将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列；基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[05]本方面，通过二维骨架关键点序列提取、二维骨架关键点序列到三维骨架关键点序列的重定向以及基于三维骨架关键点序列进行目标对象的动作渲染，实现了动作迁移，避免直接在像素级别上实现动作迁移，能够克服初始视频和目标视频之间存在的结构和视角差异大的问题，尤其在初始对象作出极端动作或初始对象与目标对象的结构差异较大时，提高了动作迁移的准确度。另外，本方面利用二维骨架关键点序列重定向三维骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[06]在一种可能的实施方式中，所述将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列，包括：基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。

[07]本实施方式，利用二维骨架关键点序列正交分解后的动作迁移分量序列，来重定向三维骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[08]在一种可能的实施方式中，在确定所述目标对象的三维骨架关键点序列之前，上述动作迁移方法还包括：获取包括目标对象的第二初始视频；识别所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列；所述基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列，包括：基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；基于所述

目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

[09]本实施方式，将初始对象的二维骨架关键点序列正交分解后的动作迁移分量序列，与目标对象的二维初始骨架关键点序列进行正交分解后的动作迁移分量序列进行融合后，来确定三维骨架关键点序列，能够克服初始对象作出极端动作或初始对象与目标对象的结构差异较大时，动作迁移准确度低的缺陷。

[10]在一种可能的实施方式中，所述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列；所述基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列，包括：基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定所述每一帧图像对应的运动分量信息、对象结构分量信息和拍摄角度分量信息；基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；基于所述第一初始视频的多帧图像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。

[11]本实施方式，动作迁移分量序列可以包括多个正交的分量序列，利用多个正交的分量序列来确定三维骨架关键点序列，能够进一步克服了初始对象作出极端动作或初始对象与目标对象的结构差异较大时，动作迁移准确度低的缺陷。

[12]在一种可能的实施方式中，所述基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频，包括：基于所述三维骨架关键点序列，生成所述目标对象的二维目标骨架关键点序列；基于所述二维目标骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[13]本实施方式，将重建的三维骨架关键点序列重投影得到二维的目标骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[14]在一种可能的实施方式中，所述将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列，包括：采用动作迁移神经网络将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列。

[15]本实施方式，利用训练好的动作迁移神经网络来确定目标对象的三维骨架关键点序列，能够提高了关键点重定向的效率和准确度。

[16]在一种可能的实施方式中，上述动作迁移方法还包括训练所述动作迁移神经网络的步骤：获取包括样本对象的动作序列的样本运动视频；识别所述样本对象在所述样本运动视频的多帧样本图像中的第一样本二维骨架关键点序列；对第一样本二维骨架关键点序列进行肢体比例缩放处理，得到第二样本二维骨架关键点序列；基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列，确定损失函数；基于所述损失函数，调整所述动作迁移神经网络的网络参数。

[17]本实施方式，利用样本对象的第一样本二维骨架关键点序列和对样本对象进行肢体比例缩放后的第二样本二维骨架关键点序列构建损失函数，来训练动作迁移神经网络，能够提高在初始对象与目标对象的结构差异较大时，动作迁移的准确度。并且在训练上述动作迁移神经网络的时候，并未使用真实世界中配对的动作-角色数据，实现了无监督的构建损失函数和训练动作迁移神经网络，有利于提高训练得到的动作迁移神经网络在进行动作迁移时的准确度。

[18]在一种可能的实施方式中，所述基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列，确定损失函数，包括：基于所述第一样本二维骨架关键点序

列，确定所述第一样本动作迁移分量序列；基于所述第二样本二维骨架关键点序列，确定所述第二样本动作迁移分量序列；基于所述第一样本动作迁移分量序列，确定预估三维骨架关键点序列；基于所述第一样本动作迁移分量序列、第二样本动作迁移分量序列和所述预估三维骨架关键点序列，确定所述损失函数。

[19]本实施方式，利用第一样本二维骨架关键点序列正交分解后的第一样本动作迁移分量序列、第二样本二维骨架关键点序列正交分解后的第二样本动作迁移分量序列，以及，基于第一样本动作迁移分量序列重建得到的预估三维骨架关键点序列，来构建损失函数，能够提高在初始对象与目标对象的结构差异较大时，动作迁移的准确度。

[20]在一种可能的实施方式中，所述损失函数包括运动不变损失函数；所述第一样本动作迁移分量序列包括各帧样本图像对应的第一样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息；所述第二样本动作迁移分量序列包括各帧样本图像对应的第二样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息；

[21]所述确定所述损失函数，包括：基于所述各帧样本图像对应的所述第二样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息，确定所述第一样本二维骨架关键点序列中对应的所述第一样本二维骨架关键点对应的第一预估骨架关键点；基于所述各帧样本图像对应的所述第一样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息，确定所述第二样本二维骨架关键点序列中对应的所述第二样本二维骨架关键点对应的第二预估骨架关键点；基于所述第一预估骨架关键点、第二预估骨架关键点、第一样本运动分量信息、第二样本运动分量信息、和所述预估三维骨架关键点序列，确定所述运动不变损失函数。

[22]本实施方式，利用第一样本二维骨架关键点序列和第二样本二维骨架关键点序列正交分解后的信息，对样本对象进行骨架恢复得到第一预估骨架关键点，以及对肢体缩放后的样本对象进行骨架恢复得到第二预估骨架关键点；之后，结合恢复得到的第一预估骨架关键点、第二预估骨架关键点和重建得到的样本对象的预估三维骨架关键点序列能够构建运动不变损失函数，由于样本对象尽管在结构和拍摄视角上存在变化和扰动，但是迁移后的运动信息应该是不变的，因此通过构建运动不变损失函数，并且在训练时，使运动不变损失函数最小，能够提高构建的动作迁移神经网络在进行动作迁移时的准确度。

[23]在一种可能的实施方式中，所述损失函数还包括结构不变损失函数；所述确定所述损失函数还包括：从所述第一样本二维骨架关键点序列中，筛选第一时刻对应的样本图像中的第一样本二维骨架关键点以及第二时刻对应的样本图像中的第一样本二维骨架关键点；从所述第二样本二维骨架关键点序列中，筛选所述第二时刻对应的样本图像中的第二样本二维骨架关键点以及所述第一时刻对应的样本图像中的第二样本二维骨架关键点；基于所述第一时刻对应的样本图像中的第一样本二维骨架关键点、所述第二时刻对应的样本图像中的第一样本二维骨架关键点、所述第一时刻对应的样本图像中的第二样本二维骨架关键点、所述第二时刻对应的样本图像中的第二样本二维骨架关键点、和所述预估三维骨架关键点序列，确定所述结构不变损失函数。

[24]本实施方式，利用不同时刻的第一样本二维骨架关键点和第二样本二维骨架关键点，结合重建得到的样本对象的预估三维骨架关键点序列能够构建结构不变损失函数，由于样本对象的结构随着时间的变化存在不变性，因此通过构建结构不变损失函数，并且在训练时，使运动不变损失函数和结构不变损失函数最小，能够提高构建的动作迁移神经网络在进行动作迁移时的准确度。

[25]在一种可能的实施方式中，所述损失函数还包括视角不变损失函数；所述确定所述损失函数还包括：基于所述第一时刻对应的样本图像中的第一样本二维骨架关键点、所

述第二时刻对应的样本图像中的第一样本二维骨架关键点、所述第一时刻和第二时刻对应的样本图像的第一样本角度分量信息、所述第一时刻和第二时刻对应的样本图像的第二样本角度分量信息、和所述预估三维骨架关键点序列，确定所述视角不变损失函数。

[26]本实施方式中，利用不同时刻的第一样本二维骨架关键点和重建得到的样本对象的预估三维骨架关键点序列等能够构建视角不变损失函数，由于样本对象的拍摄视角随着样本对象的运动和结构的变化，存在不变性，因此通过构建视角不变损失函数，并且在训练时，使视角不变损失函数、运动不变损失函数和结构不变损失函数最小，能够提高构建的动作迁移神经网络在进行动作迁移时的准确度。

[27]在一种可能的实施方式中，所述损失函数还包括重建恢复损失函数；所述确定所述损失函数还包括：基于所述第一样本二维骨架关键点序列和所述预估三维骨架关键点序列，确定所述重建恢复损失函数。

[28]本实施方式，利用第一样本二维骨架关键点序列和重建得到的样本对象的预估三维骨架关键点序列能够构建重建恢复损失函数，由于样本对象在进行样本对象恢复时，应该存在不变性，因此通过构建重建恢复损失函数，并且在训练时，使重建恢复损失函数、视角不变损失函数、运动不变损失函数和结构不变损失函数最小，能够提高构建的动作迁移神经网络在进行动作迁移时的准确度。

[29]第二方面，本公开提供了一种动作迁移装置，包括：视频获取模块，用于获取包括初始对象的动作序列的第一初始视频；关键点提取模块，用于识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列；关键点转换模块，用于将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列；图像渲染模块，用于基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[30]在一种可能的实施方式中，所述关键点转换模块在将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列时，用于：基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。

[31]在一种可能的实施方式中，所述视频获取模块还用于获取包括目标对象的第二初始视频；所述关键点提取模块还用于识别所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列；所述关键点转换模块在基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列时，用于：基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

[32]在一种可能的实施方式中，所述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列；所述关键点转换模块在基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列时，用于：基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定所述初始对象的运动分量信息、对象结构分量信息和拍摄角度分量信息；基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；基于所述第一初始视频的多帧图像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。

[33]第三方面，本公开提供了一种电子设备，包括：处理器、存储器和总线，所述存储器存储有所述处理器可执行的机器可读指令，当电子设备运行时，所述处理器与所述存储器之间通过总线通信，所述机器可读指令被所述处理器执行时执行如上述动作迁移方

法的步骤。

[34]第四方面，本公开还提供一种计算机可读存储介质，该计算机可读存储介质上存储有计算机程序，该计算机程序被处理器运行时执行如上述动作迁移方法的步骤。

[35]本公开上述装置、电子设备、和计算机可读存储介质，至少包含与本公开上述方法的任一方面或任一方面的任一实施方式的技术特征实质相同或相似的技术特征，因此关于上述装置、电子设备、和计算机可读存储介质的效果描述，可以参见上述方法内容的效果描述，这里不再赘述。

## 附图说明

[36]为了更清楚地说明本公开实施例的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本公开的某些实施例，因此不应被看作是对范围的限定，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相关的附图。

[37]图 1 示出了本公开实施例提供的一种动作迁移方法的流程图；

[38]图 2 示出了本公开实施例提供的另一种动作迁移方法的流程图；

[39]图 3 示出了本公开实施例提供的一种动作迁移神经网络的训练方法的流程图；

[40]图 4 示出了本公开实施例提供的另一种动作迁移神经网络训练过程中恢复骨架关键点的流程图；

[41]图 5 示出了本公开实施例提供的一种动作迁移装置的结构示意图；

[42]图 6 示出了本公开实施例提供的一种电子设备的结构示意图。

## 具体实施方式

[43]为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，应当理解，本公开中附图仅起到说明和描述的目的，并不用于限定本公开的保护范围。另外，应当理解，示意性的附图并未按实物比例绘制。本公开中使用的流程图示出了根据本公开的一些实施例实现的操作。应该理解，流程图的操作可以不按顺序实现，没有逻辑的上下文关系的步骤可以反转顺序或者同时实施。此外，本领域技术人员在本公开内容的指引下，可以向流程图添加一个或多个其他操作，也可以从流程图中移除一个或多个操作。

[44]另外，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本公开实施例的组件可以以各种不同的配置来布置和设计。因此，以下对在附图中提供的本公开的实施例的详细描述并非旨在限制要求保护的本公开的范围，而是仅仅表示本公开的选定实施例。基于本公开的实施例，本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

[45]需要说明的是，本公开实施例中将会用到术语“包括”，用于指出其后所声明的特征的存在，但并不排除增加其它的特征。

[46]本公开提供了一种动作迁移方法及装置，通过二维骨架关键点序列提取、二维骨架关键点序列到三维骨架关键点序列的重定向以及基于三维骨架关键点序列进行目标对象的动作渲染，实现了动作迁移，避免直接在像素级别上实现动作迁移，能够减轻初始

视频和目标视频之间存在的结构和视角差异大的问题，尤其在初始对象作出极端动作或初始对象与目标对象的结构差异较大时，提高了动作迁移的准确度。另外，本公开利用二维骨架关键点序列重定向三维骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[47]下面通过具体的实施例对本公开的动作迁移方法、装置、设备及存储介质进行说明。

[48]本公开实施例提供了一种动作迁移方法，该方法应用于进行动作迁移的终端设备或服务器等。具体地，如图1所示，本公开实施例提供的动作迁移方法包括如下步骤：

[49]S110、获取包括初始对象的动作序列的第一初始视频。

[50]这里，第一初始视频中包括多帧图像，每帧图像中初始对象可能呈现不同的姿势，这些姿势合并起来为初始对象的动作序列。

[51]S120、识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列。

[52]为了确定初始对象的动作序列，可以从第一初始视频的每帧图像中提取初始对象的二维骨架关键点，多帧图像分别对应的二维骨架关键点形成上述二维骨架关键点序列。示例性的，上述二维骨架关键点可以包括初始对象的各个关节对应的关键点。各个关节对应的关键点组合连接起来，可以得到初始对象的骨架。

[53]在可能的实施方式中，可以利用二维姿态估计神经网络提取每帧图像中初始对象的二维骨架关键点。

[54]上述初始对象可以是真实的人、虚拟的人、动物等，本公开对此不限定。

[55]S130、将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列。

[56]在可能的实施方式中，首先可以基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；之后再基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。

[57]示例性的，上述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列中的至少一种。

[58]其中，运动分量序列表示初始对象的运动，对象结构分量序列表示初始对象的身体形态，拍摄角度分量序列表示摄像机的角度。

[59]在一些实施例中，可以利用如下子步骤形成上述运动分量序列、对象结构分量序列和拍摄角度分量序列：

[60]子步骤一、基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定每一帧图像对应的运动分量信息、对象结构分量信息和拍摄角度分量信息；

[61]子步骤二、基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；

[62]子步骤三、基于所述第一初始视频的多帧图像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；

[63]子步骤四、基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。

[64]上述步骤是将每帧图像对应的二维骨架关键点通过神经网络编码成语义上正交的三个向量，分别得到每帧图像对应的运动分量信息、对象结构分量信息和拍摄角度分量信

息。之后，将多帧图像对应的运动分量信息组合形成运动分量序列，多帧图像对应的对象结构分量信息组合形成对象结构分量序列，多帧图像对应的拍摄角度分量信息组合形成拍摄角度分量序列。

[65]上述三个分量信息中，每个分量信息对于另外两个分量信息存在不变性。

[66]此步骤，利用二维骨架关键点序列正交分解后的动作迁移分量序列，来重定向三维骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度，并且能够进一步降低了初始对象作出极端动作或初始对象与目标对象的结构差异较大时，动作迁移准确度低的缺陷。

[67]S140、基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[68]在确定了三维骨架关键点序列之后，可以将该三维骨架关键点序列中的每帧图像对应的三维骨架关键点投影回二维空间，得到目标对象的二维目标骨架关键点，多帧图像对应的二维目标骨架关键点形成二维目标骨架关键点序列。之后，基于所述二维目标骨架关键点序列，生成包括目标对象的动作序列的目标视频。其中，目标对象的动作序列与初始对象的动作序列相对应。

[69]在一些实施例中，在利用二维目标骨架关键点序列，生成包括目标对象的动作序列的目标视频时，可以利用得到每组二维目标骨架关键点，进行动作渲染，得到每帧图像对应的目标对象的姿势，依次将各帧图像中姿势合并起来就能够得到目标对象的动作序列。

[70]示例性的，可以利用视频渲染引擎，基于每帧图像对应的二维目标骨架关键点，生成包括所述目标对象的动作序列的目标视频。

[71]上述，将重建的三维骨架关键点序列重投影得到二维的目标骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[72]示例性的，上述步骤 130，可以利用一个训练好的运动迁移神经网络对二维骨架关键点序列进行正交分解，以及利用分解得到的动作迁移分量序列确定目标对象的三维骨架关键点序列。

[73]上述运动迁移神经网络可以包括三个编码器和一个解码器，其中每个编码器分别用于对二维骨架关键点序列中的每个二维骨架关键点进行分量信息提取，以得到上述运动分量信息、对象结构分量信息和拍摄角度分量信息。在得到上述分量信息之后，利用一个解码器进行解码处理，重建得到目标对象的预估三维骨架关键点，最后将预估三维骨架关键点重投影回二维空间，得到上述三维骨架关键点序列中的一个三维骨架关键点。

[74]应当说明的是，在确定三维骨架关键点的时候，既可以利用编码器直接解码得到的对象结构分量信息和拍摄视角分量信息，也可以利用平均池化后的对象结构分量信息和拍摄视角分量信息来确定。具体地，将包括当前帧图像的连续多帧图像分别对象的二维骨架关键点进行正交分解，得到每帧图像对应的对象结构分量信息和拍摄角度分量信息，之后，对每帧图像对应的对象结构分量信息进行平均池化操作，得到当前帧图像对应的最终的对象结构分量信息；对每帧图像对应的拍摄视角分量信息进行平均池化操作，得到当前帧图像对应的最终的拍摄视角分量信息。最后，利用直接分解得到的运动分量信息、平均池化操作得到的对象结构分量信息和平均池化操作得到的拍摄角度分量信息，确定当前帧图像对应的三维骨架关键点。

[75]上述实施例避免直接在像素级别上实现动作迁移，减少了第一初始视频和目标视频之间存在的结构和视角差异大的问题，尤其在初始对象作出极端动作或初始对象与目标

对象的结构差异较大时，提高了动作迁移的准确度。另外，上述实施例将提取的二维骨架关键点正交分解为运动分量信息、对象结构分量信息和拍摄角度分量信息，进一步减轻了初始对象作出极端动作或初始对象与目标对象的结构差异较大时，动作迁移准确度低的缺陷。

[76]为了进一步减轻初始对象作出极端动作或初始对象与目标对象的结构差异较大时，动作迁移准确度低的缺陷，本公开实施例在确定所述目标对象的三维骨架关键点序列之前，还获取了包括目标对象的第二初始视频，并识别了所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列。

[77]之后，在确定所述目标对象的三维骨架关键点序列时，首先基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；之后，基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；最后，基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

[78]上述确定目标对象的动作迁移分量序列的方法与确定初始对象的动作迁移分量序列的方法相同，同样是首先从第二初始视频的每帧图像中分别提取目标对象的二维骨架关键点，并对每帧图像中的二维骨架关键点进行正交分解，确定了所述目标对象的运动分量信息、对象结构分量信息、和拍摄角度分量信息。最后，利用多帧图像对应的运动分量信息形成运动分量序列，利用多帧图像对应的对象结构分量信息形成对象结构分量序列，利用多帧图像对应的拍摄角度分量信息形成拍摄角度分量序列。

[79]上述实施例，利用融合后的目标动作迁移分量序列，重建目标对象的三维骨架关键点序列，之后再将重建的三维骨架关键点序列重投影得到目标对象的二维目标骨架关键点序列，避免了在动作迁移中使用误差较大的三维关键点估计和重定向，有利于提高动作迁移的准确度。

[80]下面再通过一个具体的实施例对本公开的动作迁移方法进行说明。

[81]如图2所示，本实施例的动作迁移方法包括如下步骤：

[82]步骤一、骨架提取操作。从第一初始视频的每帧图像中提取初始对象的二维骨架关键点，得到初始对象的二维骨架关键点序列；从第二初始视频的每帧图像中提取目标对象的二维骨架关键点，得到目标对象的二维骨架关键点序列。

[83]步骤二、动作迁移处理。分别对初始对象的二维骨架关键点序列中的每个二维骨架关键点和目标对象的二维骨架关键点序列中的每个二维骨架关键点进行编码处理，即进行正交分解，分别得到初始对象的每个二维骨架关键点或每帧图像对应的运动分量信息、对象结构分量信息、和拍摄角度分量信息，以及，目标对象的每个二维骨架关键点或每帧图像对应的运动分量信息、对象结构分量信息、和拍摄角度分量信息。

[84]上述初始对象的多帧图像对应的运动分量信息组成初始对象的运动分量序列，初始对象的多帧图像对应的对象结构分量信息组成初始对象的对象结构分量序列，初始对象的多帧图像对应的拍摄角度分量信息组成初始对象的拍摄角度分量序列。初始对象的运动分量序列、对象结构分量序列和拍摄角度分量序列形成初始对象的动作迁移分量序列。

[85]同样，上述目标对象的多帧图像对应的运动分量信息组成目标对象的运动分量序列，目标对象的多帧图像对应的对象结构分量信息组成目标对象的对象结构分量序列，目标对象的多帧图像对应的拍摄角度分量信息组成目标对象的拍摄角度分量序列。目标对象的运动分量序列、对象结构分量序列和拍摄角度分量序列形成目标对象的动作迁移分量序列。

[86]之后,基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列,确定目标动作迁移分量序列;基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

[87]示例性的,可以是将初始对象的每帧图像对应的运动分量信息、对象结构分量信息、和拍摄角度分量信息,与,目标对象的每帧图像对应的运动分量信息、对象结构分量信息、和拍摄角度分量信息进行重新组合,得到重新组合的目标运动分量信息、目标结构分量信息、和目标角度分量信息。

[88]上述多帧图像对应的目标运动分量信息可以组成目标运动分量序列,多帧图像对应的目标结构分量信息可以组成目标对象结构分量序列,多帧图像对应的目标角度分量信息可以组成目标拍摄角度分量序列。目标运动分量序列、目标对象结构分量序列和目标拍摄角度分量序列形成上述目标动作迁移分量序列。

[89]之后,对目标运动分量信息、目标结构分量信息、和目标角度分量信息进行解码操作,得到目标对象对应于一帧图像在三个预设角度上的三维骨架关键点。多帧图像的三维骨架关键点形成上述三维骨架关键点序列。

[90]最后,分别将每个预设角度上的三维骨架关键点重投回二维空间,分别得到目标对象在每个预设角度上的二维目标骨架关键点。

[91]步骤三、骨架到视频渲染操作。基于每帧图像中目标对象在每个预设角度上的二维目标骨架关键点,确定目标对象在每个预设角度上的目标动作,并基于目标动作生成目标对象在每个预设角度上的目标视频。

[92]上述实施例能够显著提高动作迁移的准确度,并且可以实现任意角度上的动作迁移。同时对于目标对象与初始对象在结构上差异较大、初始对象作为极端动作的情况仍然能够进行准确的动作迁移,取得了较好的视觉效果。

[93]目前,由于运动呈现出复杂的非线性以及很难在真实世界中找到配对的动作-角色数据,因此很难建立准确的动作迁移模型来实现上述动作迁移,致使动作迁移呈现出准确度低的缺陷。为了缓解上述缺陷,本公开还提供了一种动作迁移神经网络的训练方法,该方法既可以应用于上述进行动作迁移处理的终端设备或服务器上,也可以应用于单独进行神经网络训练的终端设备或服务器上。具体地,如图3所示,可以包括如下步骤:

[94]S310、获取包括样本对象的动作序列的样本运动视频。

[95]S320、识别所述样本对象在所述样本运动视频的多帧样本图像中的第一样本二维骨架关键点序列。

[96]这里,从样本运动视频的每帧图像中提取样本对象的第一样本二维骨架关键点,多帧样本图像的第一样本二维骨架关键点形成第一样本二维骨架关键点序列。

[97]上述第一样本二维骨架关键点可以包括样本对象的各个关节对应的关键点。各个关节对应的关键点组合连接起来,可以得到样本对象的骨架。

[98]在具体实施时,可以利用二维姿态估计神经网络提取样本对象的第一样本二维骨架关键点。

[99]上述样本对象可以是真实的人、虚拟的人、动物等,本公开对此不限定。

[100] S330、对第一样本二维骨架关键点序列进行肢体比例缩放处理,得到第二样本二维骨架关键点序列。

[101] 这里,按照预定的缩放比例,对第一样本二维骨架关键点序列中的每个第一样本二维骨架关键点进行肢体比例缩放,得到第二样本二维骨架关键点序列。

[102] 如图 4 所示, 第一样本二维骨架关键点  $x$  进行肢体比例缩放后, 得到第二样本二维骨架关键点  $x'$ 。

[103] S340、基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列, 确定损失函数。基于所述损失函数, 调整所述动作迁移神经网络的网络参数。

[104] 在具体实施时, 可以分别对第一样本二维骨架关键点序列中的每个第一样本二维骨架关键点和所述第二样本二维骨架关键点序列中的每个第二样本二维骨架关键点进行正交分解, 利用分解得到信息进行三维骨架关键点序列预估, 和二维的样本骨架关键点恢复, 并利用分解得到的信息、预估的三维骨架关键点序列和恢复的二维的样本骨架关键点构建损失函数。

[105] 这里, 以构建的损失函数取值最小为目标训练动作迁移神经网络。

[106] 本实施方式, 利用样本对象的第一样本二维骨架关键点序列和对样本对象进行肢体比例缩放后的第二样本二维骨架关键点序列构建损失函数, 来训练动作迁移神经网络, 能够提高在初始对象与目标对象的结构差异较大时, 动作迁移的准确度。并且在训练上述动作迁移神经网络的时候, 并未使用真实世界中配对的动作-角色数据, 实现了无监督的构建损失函数和训练动作迁移神经网络, 有利于提高训练得到的动作迁移神经网络在进行动作迁移时的准确度。

[107] 上述动作迁移神经网络具体可以包括三个编码器和一个解码器, 对动作迁移神经网络的训练实质上是对上述三个编码器和一个解码器的训练。

[108] 在一些实施例中, 上述基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列, 确定损失函数, 具体可以利用如下步骤实现:

[109] 步骤一、基于所述第一样本二维骨架关键点序列, 确定所述第一样本动作迁移分量序列。

[110] 对第一样本二维骨架关键点序列中的每个第一样本二维关键点进行正交分解, 得到每帧样本图像对应的第一样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息。多帧样本图像对应的第一样本运动分量信息形成第一样本运动分量序列; 多帧样本图像对应的第一样本结构分量信息形成第一样本结构分量序列; 多帧样本图像对应的第一样本角度分量信息形成第一样本角度分量序列。第一样本运动分量序列、第一样本角度分量序列和第一样本结构分量序列形成上述第一样本动作迁移分量序列。

[111] 这里, 如图 4 所示, 利用动作迁移神经网络中的一个编码器  $E_m$  对一个第一样本二维骨架关键点  $x$  进行处理, 得到第一样本运动分量信息, 利用另一个编码器  $E_s$  对该第一样本二维骨架关键点  $x$  进行处理, 得到第一样本结构分量信息, 利用最后一个编码器  $E_v$  对该第一样本二维骨架关键点进行处理  $x$ , 得到第一样本角度分量信息。

[112] 对当前帧样本图像对应的第一样本结构分量信息和当前帧样本图像相邻的多帧(例如, 64 帧)样本图像对应的第一样本结构分量信息进行均值池化处理, 得到最终的第一样本结构分量信息  $\bar{s}$ 。对当前帧样本图像对应的第一样本角度分量信息和当前帧样本图像相邻的多帧样本图像对应的第一样本结构分量信息进行均值池化处理, 得到最终的第一样本角度分量信息  $\bar{v}$ 。当前帧样本图像对应的第一样本运动分量信息不用进行均值池化处理, 可以直接作为最终的第一样本运动分量信息  $m$ 。

[113] 步骤二、基于所述第二样本二维骨架关键点序列, 确定所述第二样本动作迁移分量序列。

[114] 对第二样本二维骨架关键点序列中的每个第二样本二维关键点进行正交分解,

得到每帧样本图像对应的第二样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息。多帧样本图像对应的第二样本运动分量信息形成第二样本运动分量序列；多帧样本图像对应的第二样本结构分量信息形成第二样本结构分量序列；多帧样本图像对应的第二样本角度分量信息形成第二样本角度分量序列。第二样本运动分量序列、第二样本角度分量序列和第二样本结构分量序列形成上述第二样本动作迁移分量序列。

[115] 这里，如图 4 所示，利用动作迁移神经网络中的一个编码器  $E_m$  对一个第二样本二维骨架关键点  $x'$  进行处理，得到第二样本运动分量信息，利用另一个编码器  $E_s$  对第二样本二维骨架关键点  $x'$  进行处理，得到第二样本结构分量信息，利用最后一个编码器  $E_v$  对第二样本二维骨架关键点  $x'$  进行处理，得到第二样本角度分量信息。

[116] 对当前帧样本图像对应的第二样本结构分量信息和当前帧样本图像相邻的多帧样本图像对应的第二样本结构分量信息进行均值池化处理，得到最终的第二样本结构分量信息  $\bar{s}'$ 。对当前帧样本图像对应的第二样本角度分量信息和当前帧样本图像相邻的多帧样本图像对应的第二样本结构分量信息进行均值池化处理，得到最终的第一样本角度分量信息  $\bar{v}'$ 。当前帧样本图像对应的第二样本运动分量信息不用进行均值池化处理，可以直接作为最终的第二样本运动分量信息  $m'$ 。

[117] 步骤三、基于所述第一样本动作迁移分量序列，确定预估三维骨架关键点序列。

[118] 这里，具体是利用一帧样本图像对应的第一样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息，确定一个预估三维骨架关键点。多帧样本图像对应的预估三维骨架关键点即形成上述预估三维骨架关键点序列。

[119] 这里，具体可以利用一个解码器  $G$  对一帧样本图像的第一样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息进行解码处理，得到重建后的预估三维骨架关键点。

[120] 步骤四、基于所述第一样本动作迁移分量序列、第二样本动作迁移分量序列和所述预估三维骨架关键点序列，确定所述损失函数。

[121] 在具体实施时，可以利用第一样本动作迁移分量序列中的所述第一样本运动分量信息、第一样本结构分量信息、第一样本角度分量信息，第二样本动作迁移分量序列中的第二样本运动分量信息、第二样本结构分量信息、第二样本角度分量信息进行二维的样本骨架关键点恢复，并利用预估三维骨架关键点序列和恢复的二维的样本骨架关键点构建损失函数。

[122] 本实施方式，利用第一样本二维骨架关键点序列正交分解后的第一样本动作迁移分量序列、第二样本二维骨架关键点序列正交分解后的第二样本动作迁移分量序列，以及，基于第一样本动作迁移分量序列重建得到的预估三维骨架关键点序列，来构建损失函数，能够提高在初始对象与目标对象的结构差异较大时，动作迁移的准确度。

[123] 由于样本对象尽管在结构和拍摄视角上存在变化和扰动，但是迁移后的运动信息应该是不变的，因此可以通过构建运动不变损失函数，并且在训练时，使运动不变损失函数最小，来提高构建的动作迁移神经网络在进行动作迁移时的准确度。具体地，可以利用如下步骤构建上述运动不变损失函数：

[124] 步骤一、基于所述第二样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息，确定所述第一样本二维骨架关键点序列中对应的所述第一样本二维骨架关键点对应的第一预估骨架关键点。

[125] 如图 4 所示，具体可以利用如下子步骤实现：利用解码器  $G$  对第二样本运动分

量信息  $m'$ 、第一样本结构分量信息  $\bar{s}$ 、第一样本角度分量信息  $\bar{v}$  进行处理，可以重建得到三维的骨架关键点  $\hat{X}'$ ，之后，利用旋转投影函数  $\Phi(\hat{X}', 0)$  将三维的骨架关键点  $\hat{X}'$  重投影到二维空间，得到第一预估骨架关键点  $\hat{x}'$ 。

[126] 步骤二、基于所述第一样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息，确定所述第二样本二维骨架关键点序列中对应的所述第二样本二维骨架关键点对应的第二预估骨架关键点。

[127] 如图 4 所示，具体可以利用如下子步骤实现：利用解码器 G 对第一样本运动分量信息  $m$ 、第二样本结构分量信息  $\bar{s}'$ 、第二样本角度分量信息  $\bar{v}'$  进行处理，可以重建得到三维的骨架关键点  $\hat{X}''$ ，之后，利用旋转投影函数  $\Phi(\hat{X}'', 0)$  将三维的骨架关键点  $\hat{X}''$  重投影到二维空间，得到第二预估骨架关键点  $\hat{x}''$ 。

[128] 步骤一和步骤二中，生成第一预估骨架关键点  $\hat{x}'$  和第二预估骨架关键点  $\hat{x}''$  的具体公式如下：

$$\begin{aligned}\hat{x}' &= \phi \left[ G(E_m(\mathbf{x}'), \bar{E}_s(\mathbf{x}), \bar{E}_v(\mathbf{x})), 0 \right] \\ \hat{x}'' &= \phi \left[ G(E_m(\mathbf{x}), \bar{E}_s(\mathbf{x}'), \bar{E}_v(\mathbf{x}')), 0 \right]\end{aligned}\quad (1)$$

[129] 式中， $\bar{E}_s$  表示进行对编码器提取的样本结构分量信息进行平均池化操作， $\bar{E}_v$  表示进行对编码器提取的样本角度分量信息进行平均池化操作。

[130] 步骤三、基于所述第一预估骨架关键点、第二预估骨架关键点、第一样本运动分量信息、第二样本运动分量信息、和所述预估三维骨架关键点序列，确定所述运动不变损失函数。

[131] 构建的运动不变损失函数具体可以包括如下三个：

$$\mathcal{L}_{\text{crs}} = \frac{1}{2NT} \left( \frac{1}{2} |\mathbf{x} - \hat{x}'| + \frac{1}{2} |\mathbf{x}' - \hat{x}''| \right) \quad (2)$$

$$\mathcal{L}_{\text{inv}_m}^{(s)} = \frac{1}{MC_m} |E_m(\mathbf{x}) - E_m(\mathbf{x}')| \quad (3)$$

$$\mathcal{L}_{\text{inv}_m}^{(v)} = \frac{1}{KMC_m} \sum_{k=1}^K |E_m(\mathbf{x}) - E_m(\hat{\mathbf{x}}^{(k)})| \quad (4)$$

$$\hat{\mathbf{x}}^{(k)} = \phi \left( \hat{\mathbf{X}}, \frac{k}{K+1} \pi \right), k = 1, 2, \dots, K \quad (5)$$

[132] 式中，N 表示样本运动视频的帧数，T 表示一个第一样本二维骨架关键点对应的关节的数量，M 表示一个预设的数值， $C_m$  表示第一样本运动分量信息对应的编码长度，K 表示样本对象旋转的数量， $\hat{X}$  表示一个预估三维骨架关键点， $\mathcal{L}_{\text{crs}}$ 、 $\mathcal{L}_{\text{inv}_m}^{(s)}$ 、

$\mathcal{L}_{inv\_m}^{(v)}$  表示三个运动不变损失函数。

[133] 本公开实施例中，利用第一样本二维骨架关键点序列和第二样本二维骨架关键点序列正交分解后的信息，对样本对象进行骨架恢复得到第一预估骨架关键点，以及对肢体缩放后的样本对象进行骨架恢复得到第二预估骨架关键点；之后，结合恢复得到的第一预估骨架关键点、第二预估骨架关键点和重建得到的样本对象的预估三维骨架关键点序列能够构建运动不变损失函数。

[134] 由于样本对象的结构随着时间的变化存在不变性，因此可以通过构建结构不变损失函数，并且在训练时，使运动不变损失函数和结构不变损失函数最小，来提高构建的动作迁移神经网络在进行动作迁移时的准确度。具体地，可以利用如下步骤构建上述结构不变损失函数：

[135] 步骤一、从所述第一样本二维骨架关键点序列中，筛选所述样本对象在第一时刻的第一样本二维骨架关键点、所述样本对象在第二时刻的第一样本二维骨架关键点。

[136] 从所述第二样本二维骨架关键点序列中，筛选所述样本对象在第二时刻的第二样本二维骨架关键点、和所述样本对象在第一时刻的第二样本二维骨架关键点。

[137] 上述第一样本二维骨架关键点是从样本运动视频中第一时刻 t1 和第二时刻 t2 对应的样本图像中分别提取的样本对象的二维骨架关键点，是未经过肢体比例缩放的样本对象的骨架关键点。上述第二样本二维骨架关键点是在样本运动视频中第一时刻 t1 和第二时刻 t2 对应的样本图像中分别提取的样本对象的骨架关键点进行肢体比例缩放后的关键点。

[138] 步骤二、基于所述样本对象在第一时刻的第一样本二维骨架关键点、所述样本对象在第二时刻的第一样本二维骨架关键点、所述样本对象在第二时刻的第二样本二维骨架关键点、所述样本对象在第一时刻的第二样本二维骨架关键点、和所述预估三维骨架关键点序列，确定所述结构不变损失函数。

[139] 在具体实施时，构建的结构不变损失函数包括如下两个：

$$\mathcal{L}_{trip\_s} = \frac{1}{2M} \sum_{t_1, t_2} [\tau(\mathbf{s}_{t_1}, \mathbf{s}_{t_2}, \mathbf{s}'_{t_2}) + \tau(\mathbf{s}'_{t_1}, \mathbf{s}'_{t_2}, \mathbf{s}_{t_2})] \quad (6)$$

$$\mathcal{L}_{inv\_s} = \frac{1}{KC_b} \sum_{i=1}^K |\bar{E}_s(\mathbf{x}) - \bar{E}_s(\hat{\mathbf{x}}^{(k)})| \quad (7)$$

$$\tau(\mathbf{s}_{t_1}, \mathbf{s}_{t_2}, \mathbf{s}'_{t_2}) = \max \{0, s(\mathbf{s}_{t_1}, \mathbf{s}'_{t_2}) - s(\mathbf{s}_{t_1}, \mathbf{s}_{t_2}) + m\} \quad (8)$$

[140] 式中，St1 表示从时刻 t1 的第一样本二维骨架关键点中直接提取的样本结构分量信息，St2 表示从时刻 t2 的第一样本二维骨架关键点中直接提取的样本结构分量信息，St2' 表示从时刻 t2 的第二样本二维骨架关键点中直接提取的样本结构分量信息，St1' 表示从时刻 t1 的第二样本二维骨架关键点中直接提取的样本结构分量信息，Cb 表示第一样本结构分量信息对应的编码长度，m 是一个预设的数值，s() 表示余弦相似函数，

$\mathcal{L}_{trip\_s}$ 、 $\mathcal{L}_{inv\_s}$  表示两个结构不变损失函数。

[141] 本公开实施例中，利用不同时刻的第一样本二维骨架关键点和第二样本二维骨架关键点，结合重建得到的样本对象的预估三维骨架关键点序列能够构建结构不变损失函数。

[142] 由于样本对象的拍摄视角随着样本对象的运动和结构的变化，存在不变性，因此可以通过构建视角不变损失函数，并且在训练时，使视角不变损失函数、运动不变损失函数和结构不变损失函数最小，能够提高构建的动作迁移神经网络在进动作迁移时的准确度。具体地，可以利用如下步骤构建视角不变损失函数：

[143] 基于所述样本对象在第一时刻的第一样本二维骨架关键点、所述样本对象在第二时刻的第一样本二维骨架关键点、第一样本角度分量信息、第二样本角度分量信息、和所述预估三维骨架关键点序列，确定所述视角不变损失函数。

[144] 构建的视角不变损失函数具体包括如下两个：

$$\mathcal{L}_{\text{rip}} = \frac{1}{2MK} \sum_{k, t_1, t_2} \left[ \tau(\mathbf{v}_{t_1}, \mathbf{v}_{t_2}, \mathbf{v}_{t_2}^{(k)}) + \tau(\mathbf{v}_{t_1}^{(k)}, \mathbf{v}_{t_2}^{(k)}, \mathbf{v}_{t_2}) \right] \quad (9)$$

$$\mathcal{L}_{\text{inv}_v} = \frac{1}{C_v} |\bar{E}_v(\mathbf{x}) - \bar{E}_v(\mathbf{x}')| \quad (10)$$

$$\mathbf{v}^{(k)} = \bar{E}_v(\hat{\mathbf{x}}^{(k)}), t_1 \neq t_2 \quad (11)$$

[145] 式中， $\mathbf{v}_{t_1}$  表示从时刻  $t_1$  的第一样本二维骨架关键点中直接提取的样本角度分量信息， $\mathbf{v}_{t_2}$  表示从时刻  $t_2$  的第一样本二维骨架关键点中直接提取的样本角度分量信息， $C_v$  表示第一样本角度分量信息对应的编码长度， $\mathcal{L}_{\text{rip}}$ 、 $\mathcal{L}_{\text{inv}_v}$  表示两个视角不变损失函数。

[146] 由于样本对象在进行样本对象恢复时，应该存在不变性，因此可以通过构建重建恢复损失函数，并且在训练时，使重建恢复损失函数、视角不变损失函数、运动不变损失函数和结构不变损失函数最小，来提高构建的动作迁移神经网络在进动作迁移时的准确度。具体的，可以利用如下步骤构建重建恢复损失函数：

[147] 基于所述第一样本二维骨架关键点序列和所述预估三维骨架关键点序列，确定所述重建恢复损失函数。

[148] 构建的重建恢复损失函数具体包括如下两个：

$$\mathcal{L}_{\text{rec}} = \frac{1}{2NT} \left| \mathbf{x} - \phi(\hat{\mathbf{X}}, 0) \right| \quad (12)$$

$$\mathcal{L}_{\text{adv}} = \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{\mathbf{x} \sim p_{\mathbf{x}}} \left[ \frac{1}{2} \log D(\mathbf{x}) + \frac{1}{2} \log(1 - D(\hat{\mathbf{x}}^{(k)})) \right] \quad (13)$$

[149] 式中， $D$  表示一个时序上的卷积网络， $\mathbb{E}_{\mathbf{x} \sim p_{\mathbf{x}}}$  表示  $\mathbf{x}$  取自样本的概率分布，

再对后面的函数，即  $\frac{1}{2} \log D(\mathbf{x}) + \frac{1}{2} \log(1 - D(\hat{\mathbf{x}}^{(k)}))$  求期望， $\mathcal{L}_{\text{rec}}$ 、 $\mathcal{L}_{\text{adv}}$  表示两个重建恢复损失函数。

[150] 通过上面的实施例构建了重建恢复损失函数、视角不变损失函数、运动不变损失函数和结构不变损失函数，在具体实施时，可以利用如下公式对上述损失函数进行融合，得到目标损失函数：

$$\begin{aligned} \mathcal{L} = & \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} + \lambda_{\text{crs}} \mathcal{L}_{\text{crs}} + \lambda_{\text{adv}} \mathcal{L}_{\text{adv}} + \lambda_{\text{trip}} (\mathcal{L}_{\text{trip}_s} + \mathcal{L}_{\text{trip}_v}) \\ & + \lambda_{\text{inv}} (\mathcal{L}_{\text{inv}_m}^{(s)} + \mathcal{L}_{\text{inv}_m}^{(v)} + \mathcal{L}_{\text{inv}_s} + \mathcal{L}_{\text{inv}_v}) \end{aligned} \quad (14)$$

[151] 式中， $\lambda_{\text{rec}}$ 、 $\lambda_{\text{crs}}$ 、 $\lambda_{\text{adv}}$ 、 $\lambda_{\text{trip}}$ 、 $\lambda_{\text{inv}}$  均表示预设的权重。

[152] 在训练动作迁移神经网络的时候，以上述目标损失函数取值最小即可。

[153] 对应于上述动作迁移方法，本公开还提供了一种动作迁移装置，该装置应用于进行动作迁移的终端设备或服务器上，并且各个模块能够实现与上述方法中相同的方法步骤以及取得相同的有益效果，因此对于其中相同的部分，本公开不再进行赘述。

[154] 如图 5 所示，本公开提供的一种动作装置可以包括：

[155] 视频获取模块 510，用于获取包括初始对象的动作序列的第一初始视频。

[156] 关键点提取模块 520，用于识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列。

[157] 关键点转换模块 530，用于将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列。

[158] 图像渲染模块 540，用于基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[159] 在一些实施例中，所述关键点转换模块 530 在将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列时，用于：基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。

[160] 在一些实施例中，所述视频获取模块 510 还用于获取包括目标对象的第二初始视频；所述关键点提取模块 520 还用于识别所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列；所述关键点转换模块 530 在基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列时，用于：基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

[161] 在一些实施例中，所述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列；所述关键点转换模块 530 在基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列时，用于：基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定每一帧图像对应的运动分量信息、对象结构分量信息和拍摄角度分量信息；基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；基于所述第一初始视频的多帧图

像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。

[162] 本公开实施例公开了一种电子设备，如图 6 所示，包括：处理器 601、存储器 602 和总线 603，所述存储器 602 存储有所述处理器 601 可执行的机器可读指令，当电子设备运行时，所述处理器 601 与所述存储器 602 之间通过总线 603 通信。

[163] 所述机器可读指令被所述处理器 601 执行时执行以下动作迁移方法的步骤：获取包括初始对象的动作序列的第一初始视频；识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列；将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列；基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

[164] 除此之外，机器可读指令被处理器 61 执行时，还可以执行上述方法部分描述的任一实施方式中的方法内容，这里不再赘述。

[165] 本公开实施例还提供的一种对应于上述方法及装置的计算机程序产品，包括存储了程序代码的计算机可读存储介质，程序代码包括的指令可用于执行前面方法实施例中的方法，具体实现可参见方法实施例，在此不再赘述。

[166] 上文对各个实施例的描述倾向于强调各个实施例之间的不同之处，其相同或相似之处可以相互参考，为了简洁，本文不再赘述。

[167] 所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统 and 装置的具体工作过程，可以参考方法实施例中的对应过程，本公开中不再赘述。在本公开所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，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，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，又例如，多个模块或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口，装置或模块的间接耦合或通信连接，可以是电性，机械或其它的形式。

[168] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[169] 另外，在本公开各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

[170] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个处理器可执行的非易失的计算机可读取存储介质中。基于这样的理解，本公开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，或者网络设备等）执行本公开各个实施例所述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、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[171] 以上仅为本公开的具体实施方式，但本公开的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本公开揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本公开的保护范围之内。因此，本公开的保护范围应以权利要求的保护范围为准。

## 权利要求书

- 1、一种动作迁移方法，其特征在于，包括：  
获取包括初始对象的动作序列的第一初始视频；  
识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列；  
将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列；  
基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。
- 2、根据权利要求1所述的动作迁移方法，其特征在于，所述将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列，包括：  
基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；  
基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。
- 3、根据权利要求2所述的动作迁移方法，其特征在于，  
在确定所述目标对象的三维骨架关键点序列之前，还包括：  
获取包括所述目标对象的第二初始视频；  
识别所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列；  
所述基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列，包括：  
基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；  
基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；  
基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。
- 4、根据权利要求2所述的动作迁移方法，其特征在于，所述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列；  
所述基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列，包括：  
基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定每一帧图像对应的运动分量信息、对象结构分量信息和拍摄角度分量信息；  
基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；  
基于所述第一初始视频的多帧图像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；  
基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。
- 5、根据权利要求1所述的动作迁移方法，其特征在于，所述基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频，包括：  
基于所述三维骨架关键点序列，生成所述目标对象的二维目标骨架关键点序列；  
基于所述二维目标骨架关键点序列，生成包括所述目标对象的动作序列的所述目标视频。
- 6、根据权利要求1至5任一项所述的动作迁移方法，其特征在于，所述将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列，包括：  
采用动作迁移神经网络将所述二维骨架关键点序列转换为所述目标对象的所述三维骨架关键点序列。
- 7、根据权利要求6所述的动作迁移方法，其特征在于，还包括训练所述动作迁移神经网络的步骤：  
获取包括样本对象的动作序列的样本运动视频；  
识别所述样本对象在所述样本运动视频的多帧样本图像中的第一样本二维骨架关

键点序列;

对第一样本二维骨架关键点序列进行肢体比例缩放处理,得到第二样本二维骨架关键点序列;

基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列,确定损失函数;

基于所述损失函数,调整所述动作迁移神经网络的网络参数。

8、根据权利要求7所述的动作迁移方法,其特征在于,所述基于所述第一样本二维骨架关键点序列和所述第二样本二维骨架关键点序列,确定损失函数,包括:

基于所述第一样本二维骨架关键点序列,确定所述第一样本动作迁移分量序列;

基于所述第二样本二维骨架关键点序列,确定所述第二样本动作迁移分量序列;

基于所述第一样本动作迁移分量序列,确定预估三维骨架关键点序列;

基于所述第一样本动作迁移分量序列、所述第二样本动作迁移分量序列和所述预估三维骨架关键点序列,确定所述损失函数。

9、根据权利要求8所述的动作迁移方法,其特征在于,所述损失函数包括运动不变损失函数;所述第一样本动作迁移分量序列包括各帧样本图像对应的第一样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息;所述第二样本动作迁移分量序列包括各帧样本图像对应的第二样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息;

所述确定所述损失函数,包括:

基于所述各帧样本图像对应的所述第二样本运动分量信息、第一样本结构分量信息和第一样本角度分量信息,确定所述第一样本二维骨架关键点序列中对应的所述第一样本二维骨架关键点对应的第一预估骨架关键点;

基于所述各帧样本图像对应的所述第一样本运动分量信息、第二样本结构分量信息和第二样本角度分量信息,确定所述第二样本二维骨架关键点序列中对应的所述第二样本二维骨架关键点对应的第二预估骨架关键点;

基于所述第一预估骨架关键点、所述第二预估骨架关键点、所述第一样本运动分量信息、所述第二样本运动分量信息、和所述预估三维骨架关键点序列,确定所述运动不变损失函数。

10、根据权利要求9所述的动作迁移方法,其特征在于,所述损失函数还包括结构不变损失函数;

所述确定所述损失函数还包括:

从所述第一样本二维骨架关键点序列中,筛选第一时刻对应的样本图像中的第一样本二维骨架关键点以及第二时刻对应的样本图像中的第一样本二维骨架关键点;

从所述第二样本二维骨架关键点序列中,筛选所述第二时刻对应的样本图像中的第二样本二维骨架关键点以及所述第一时刻对应的样本图像中的第二样本二维骨架关键点;

基于所述第一时刻对应的样本图像中的第一样本二维骨架关键点、所述第二时刻对应的样本图像中的第一样本二维骨架关键点、所述第二时刻对应的样本图像中的第二样本二维骨架关键点、所述第一时刻对应的样本图像中的第二样本二维骨架关键点、和所述预估三维骨架关键点序列,确定所述结构不变损失函数。

11、根据权利要求10所述的动作迁移方法,其特征在于,所述损失函数还包括视角不变损失函数;

所述确定所述损失函数还包括:

基于所述第一时刻对应的样本图像中的第一样本二维骨架关键点、所述第二时刻对应的样本图像中的第一样本二维骨架关键点、所述第一时刻和第二时刻对应的

样本图像的第一样本角度分量信息、所述第一时刻和第二时刻对应的样本图像的第二样本角度分量信息、和所述预估三维骨架关键点序列，确定所述视角不变损失函数。

12、根据权利要求 11 所述的动作迁移方法，其特征在于，所述损失函数还包括重建恢复损失函数；

所述确定所述损失函数还包括：

基于所述第一样本二维骨架关键点序列和所述预估三维骨架关键点序列，确定所述重建恢复损失函数。

13、一种动作迁移装置，其特征在于，包括：

视频获取模块，用于获取包括初始对象的动作序列的第一初始视频；

关键点提取模块，用于识别所述初始对象在所述第一初始视频的多帧图像中的二维骨架关键点序列；

关键点转换模块，用于将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列；

图像渲染模块，用于基于所述三维骨架关键点序列，生成包括目标对象的动作序列的目标视频。

14、根据权利要求 13 所述的动作迁移装置，其特征在于，所述关键点转换模块在将所述二维骨架关键点序列转换为目标对象的三维骨架关键点序列时，用于：

基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列；

基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列。

15、根据权利要求 14 所述的动作迁移装置，其特征在于，所述视频获取模块还用于获取包括目标对象的第二初始视频；

所述关键点提取模块还用于识别所述目标对象在所述第二初始视频的多帧图像中的二维骨架关键点序列；

所述关键点转换模块在基于所述初始对象的动作迁移分量序列，确定所述目标对象的三维骨架关键点序列时，用于：

基于所述目标对象的二维骨架关键点序列，确定所述目标对象的动作迁移分量序列；

基于所述初始对象的动作迁移分量序列和所述目标对象的动作迁移分量序列，确定目标动作迁移分量序列；

基于所述目标动作迁移分量序列确定所述目标对象的三维骨架关键点序列。

16、根据权利要求 14 所述的动作迁移装置，其特征在于，所述初始对象的动作迁移分量序列包括运动分量序列、对象结构分量序列和拍摄角度分量序列；

所述关键点转换模块在基于所述二维骨架关键点序列，确定所述初始对象的动作迁移分量序列时，用于：

基于所述第一初始视频的多帧图像中每一帧图像对应的二维骨架关键点，分别确定所述初始对象的运动分量信息、对象结构分量信息和拍摄角度分量信息；

基于所述第一初始视频的多帧图像中每一帧图像对应的运动分量信息，确定所述运动分量序列；

基于所述第一初始视频的多帧图像中每一帧图像对应的对象结构分量信息，确定所述对象结构分量序列；

基于所述第一初始视频的多帧图像中每一帧图像对应的拍摄角度分量信息，确定所述拍摄角度分量序列。

17、一种电子设备，其特征在于，包括：处理器、存储介质和总线，所述存储介质存储有所述处理器可执行的机器可读指令，当电子设备运行时，所述处理器与所述存储

介质之间通过总线通信，所述处理器执行所述机器可读指令，以执行如权利要求 1~12 任一所述的动作迁移方法。

18、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有计算机程序，所述计算机程序被处理器运行时执行如权利要求 1~12 任一所述的动作迁移方法。

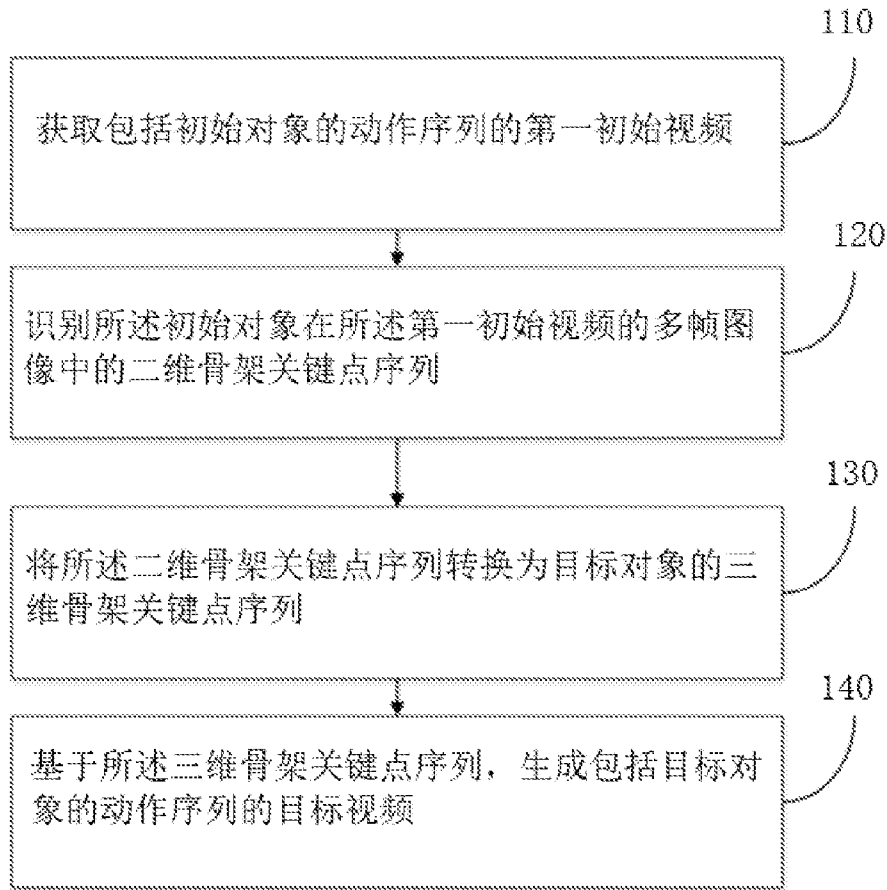


图 1

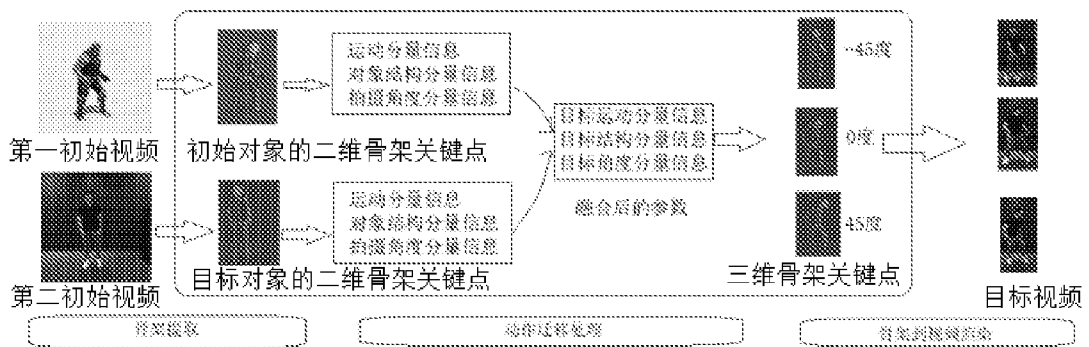


图 2

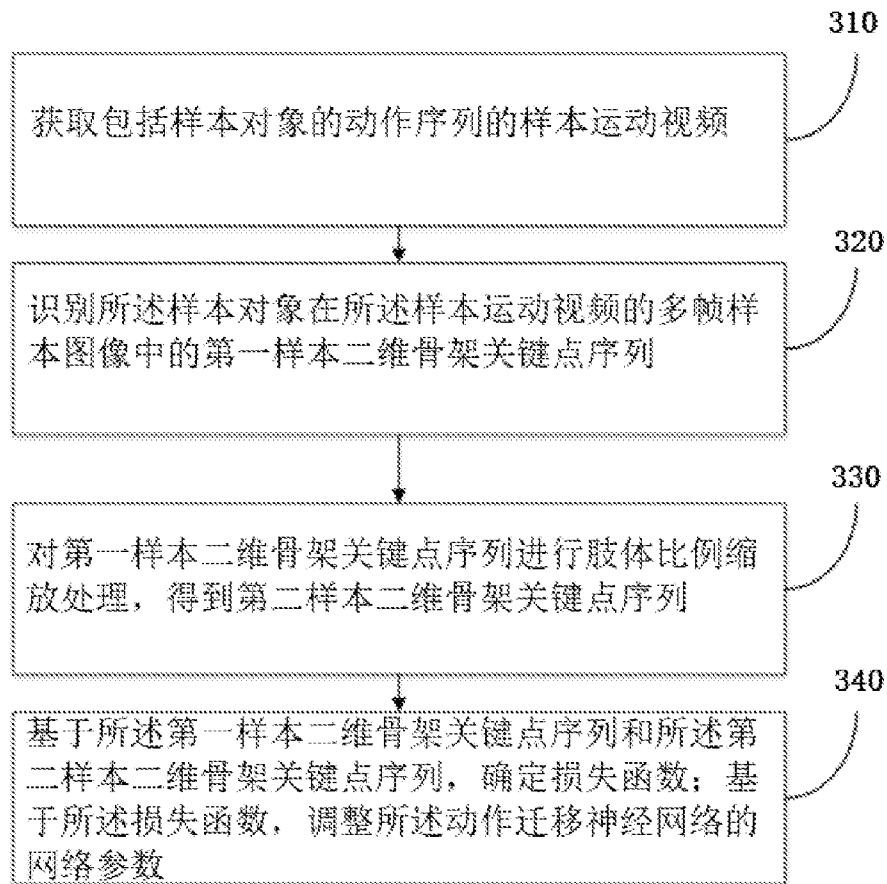


图 3

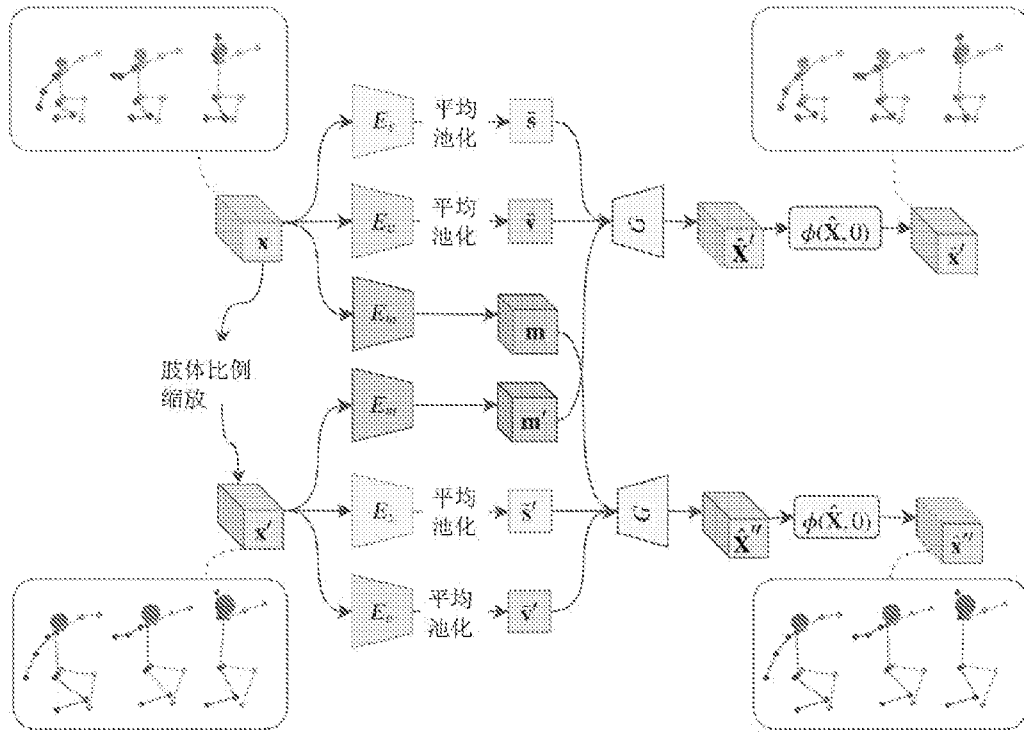


图 4

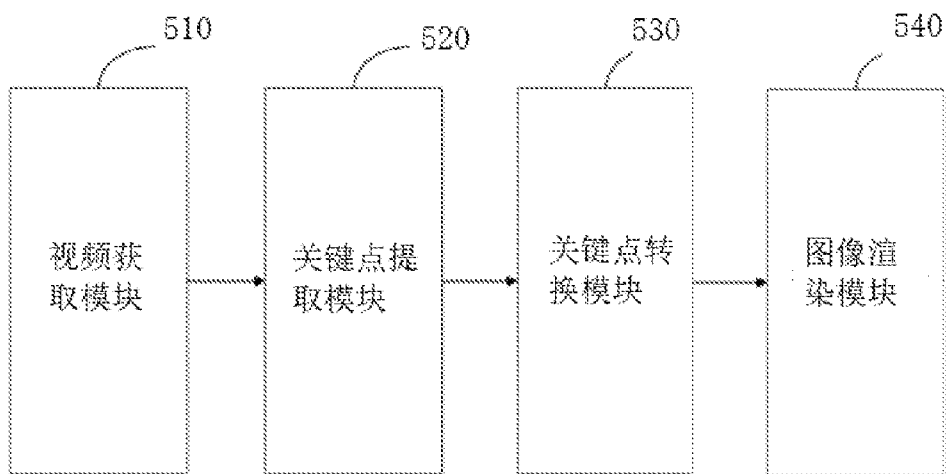


图 5

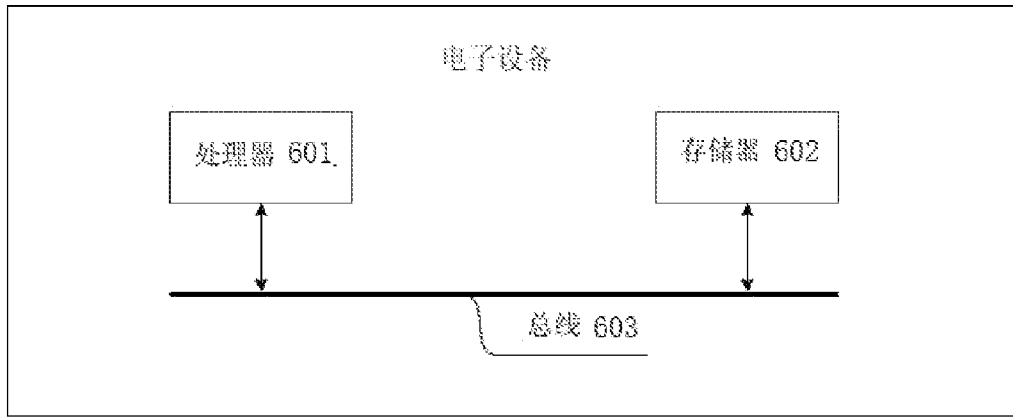


图 6

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/082407

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
G06T 7/579(2017.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
G06T		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 动作, 姿势, 姿态, 迁移, 重定向, 关键点, 关节, 骨骼, 三维, 二维, 帧, motion, pose, transfer, retarget, keypoint, joint, skeleton, three dimension, two dimension, frame		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111462209 A (BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) 28 July 2020 (2020-07-28) entire document	1-18
PX	YANG, Zhuoqian et al. "TransMoMo: Invariance-Driven Unsupervised Video Motion Retargeting" <i>2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition</i> , 19 June 2020 (2020-06-19), ISSN: 2575-7075, sections 3-4	1-18
PX	CN 111540055 A (GUANGZHOU HUYA TECHNOLOGY CO., LTD.) 14 August 2020 (2020-08-14) description paragraphs 37-61	1, 13, 17-18
X	ABERMAN, Kfir et al. "Learning Character-Agnostic Motion for Motion Retargeting in 2D" <i>ACM Transactions on Graphics</i> , Vol. 38, No. 4, 12 July 2019 (2019-07-12), section 1, 3-4, 6	1-8, 13-18
X	CN 109821239 A (NETEASE (HANGZHOU) NETWORK CO., LTD.) 31 May 2019 (2019-05-31) description, paragraphs [0038]-[0057]	1, 13, 17-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
07 June 2021		23 June 2021
Name and mailing address of the ISA/CN		Authorized officer
<b>China National Intellectual Property Administration (ISA/CN)</b> <b>No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088</b> <b>China</b>		
Facsimile No. (86-10)62019451		Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2021/082407**

<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 109978975 A (SHENZHEN SENSITIVE TECHNOLOGY CO., LTD.) 05 July 2019 (2019-07-05) description paragraphs 89-153	1-18
A	CN 108510577 A (INSTITUTE OF SOFTWARE, CHINESE ACADEMY OF SCIENCES) 07 September 2018 (2018-09-07) entire document	1-18
A	US 2019295305 A1 (ADOBE INC.) 26 September 2019 (2019-09-26) entire document	1-18

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No. <b>PCT/CN2021/082407</b>
---

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111462209	A	28 July 2020	None			
CN	111540055	A	14 August 2020	None			
CN	109821239	A	31 May 2019	None			
CN	109978975	A	05 July 2019	None			
CN	108510577	A	07 September 2018	CN	108510577	B	23 March 2021
US	2019295305	A1	26 September 2019	US	10546408	B2	28 January 2020

<p><b>A. 主题的分类</b></p> <p>G06T 7/579 (2017.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																				
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06T</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 动作, 姿势, 姿态, 迁移, 重定向, 关键点, 关节, 骨骼, 三维, 二维, 帧, motion, pose, transfer, retarget, keypoint, joint, skeleton, three dimension, two dimension, frame</p>																				
<p><b>C. 相关文件</b></p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 111462209 A (北京市商汤科技开发有限公司) 2020年 7月 28日 (2020 - 07 - 28) 全文</td> <td>1-18</td> </tr> <tr> <td>PX</td> <td>YANG, Zhuoqian 等. "TransMoMo: Invariance-Driven Unsupervised Video Motion Retargeting" 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020年 6月 19日 (2020 - 06 - 19), ISSN: 2575-7075, 第3-4节</td> <td>1-18</td> </tr> <tr> <td>PX</td> <td>CN 111540055 A (广州虎牙科技有限公司) 2020年 8月 14日 (2020 - 08 - 14) 说明书第37-61段</td> <td>1、13、17-18</td> </tr> <tr> <td>X</td> <td>ABERMAN, Kfir 等. "Learning Character-Agnostic Motion for Motion Retargeting in 2D" ACM Transactions on Graphics, 第38卷, 第4期, 2019年 7月 12日 (2019 - 07 - 12), 第1、3-4、6节</td> <td>1-8、13-18</td> </tr> <tr> <td>X</td> <td>CN 109821239 A (网易杭州网络有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第38-57段</td> <td>1、13、17-18</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 111462209 A (北京市商汤科技开发有限公司) 2020年 7月 28日 (2020 - 07 - 28) 全文	1-18	PX	YANG, Zhuoqian 等. "TransMoMo: Invariance-Driven Unsupervised Video Motion Retargeting" 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020年 6月 19日 (2020 - 06 - 19), ISSN: 2575-7075, 第3-4节	1-18	PX	CN 111540055 A (广州虎牙科技有限公司) 2020年 8月 14日 (2020 - 08 - 14) 说明书第37-61段	1、13、17-18	X	ABERMAN, Kfir 等. "Learning Character-Agnostic Motion for Motion Retargeting in 2D" ACM Transactions on Graphics, 第38卷, 第4期, 2019年 7月 12日 (2019 - 07 - 12), 第1、3-4、6节	1-8、13-18	X	CN 109821239 A (网易杭州网络有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第38-57段	1、13、17-18
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 111462209 A (北京市商汤科技开发有限公司) 2020年 7月 28日 (2020 - 07 - 28) 全文	1-18																		
PX	YANG, Zhuoqian 等. "TransMoMo: Invariance-Driven Unsupervised Video Motion Retargeting" 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020年 6月 19日 (2020 - 06 - 19), ISSN: 2575-7075, 第3-4节	1-18																		
PX	CN 111540055 A (广州虎牙科技有限公司) 2020年 8月 14日 (2020 - 08 - 14) 说明书第37-61段	1、13、17-18																		
X	ABERMAN, Kfir 等. "Learning Character-Agnostic Motion for Motion Retargeting in 2D" ACM Transactions on Graphics, 第38卷, 第4期, 2019年 7月 12日 (2019 - 07 - 12), 第1、3-4、6节	1-8、13-18																		
X	CN 109821239 A (网易杭州网络有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第38-57段	1、13、17-18																		
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>"A" 认为不特别相关的表示了现有技术一般状态的文件</p> <p>"E" 在国际申请日的当天或之后公布的在先申请或专利</p> <p>"L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>"O" 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>"P" 公布日先于国际申请日但迟于所要求的优先权日的文件</p> <p>"T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>"X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>"Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>"&amp;" 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2021年 6月 7日</p>		<p>国际检索报告邮寄日期</p> <p>2021年 6月 23日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>谢宜瑾</p> <p>电话号码 86-(20)-28950396</p>																		

C. 相关文件		
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 109978975 A (深圳市商汤科技有限公司) 2019年 7月 5日 (2019 - 07 - 05) 说明书第89-153段	1-18
A	CN 108510577 A (中国科学院软件研究所) 2018年 9月 7日 (2018 - 09 - 07) 全文	1-18
A	US 2019295305 A1 (ADOBE INC) 2019年 9月 26日 (2019 - 09 - 26) 全文	1-18

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2021/082407

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	111462209	A	2020年 7月 28日	无	
CN	111540055	A	2020年 8月 14日	无	
CN	109821239	A	2019年 5月 31日	无	
CN	109978975	A	2019年 7月 5日	无	
CN	108510577	A	2018年 9月 7日	CN	108510577 B 2021年 3月 23日
US	2019295305	A1	2019年 9月 26日	US	10546408 B2 2020年 1月 28日