

(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
2020 年 10 月 29 日 (29.10.2020)



(10) 国际公布号
WO 2020/216116 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2020/084996

(22) 国际申请日: 2020 年 4 月 16 日 (16.04.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910345010.1 2019年4月26日 (26.04.2019) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 罗镜民 (LUO, Jingmin); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 乔亮 (QIAO, Liang); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 朱晓龙 (ZHU, Xiaolong); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房 (部位: 自编 01-03 和 08-12 单元) (仅限办公用途), Guangdong 510623 (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,

(54) Title: ACTION RECOGNITION METHOD AND APPARATUS, AND HUMAN-MACHINE INTERACTION METHOD AND APPARATUS

(54) 发明名称: 动作识别方法和装置、人机交互方法和装置

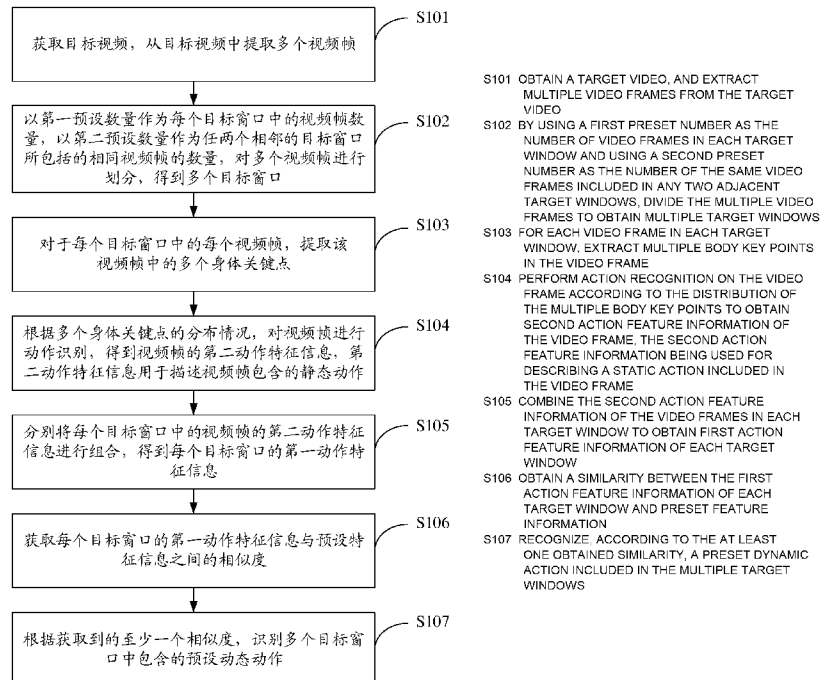


图 1

(57) Abstract: An action recognition method, comprising: extracting multiple target windows from a target video, each of the multiple target windows comprising a first preset number of consecutive video frames; performing action recognition on the video frames in each target window to obtain first action feature information of each target window, the first action feature information being used for describing a dynamic action included in the target window; obtaining a similarity between the first action feature information of each target window and preset feature information, the preset feature information being used for describing a preset dynamic action, and

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, 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, 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))。

the largest similarity among multiple obtained similarities being a first similarity; and if the first similarity is greater than a first preset threshold, and the difference between the first similarity and a second similarity is greater than a second preset threshold, determining that the dynamic action included in a first target window corresponding to the first similarity is the preset dynamic action, the second similarity being a similarity corresponding to the first action feature information of a target window adjacent to the first target window.

(57) 摘要: 一种动作识别方法, 包括: 从目标视频中提取多个目标窗口; 多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧; 分别对每个目标窗口中的视频帧进行动作识别, 得到每个目标窗口的第一动作特征信息, 第一动作特征信息用于描述目标窗口包含的动态动作; 获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度, 预设特征信息用于描述预设动态动作; 获取到的多个相似度中最大的相似度为第一相似度; 及当第一相似度大于第一预设阈值, 且第一相似度与第二相似度的差值大于第二预设阈值时, 确定第一相似度对应的第一目标窗口包含的动态动作为预设动态动作, 第二相似度为与第一目标窗口相邻的目标窗口的第一动作特征信息对应的相似度。

动作识别方法和装置、人机交互方法和装置

本申请要求于 2019 年 04 月 26 日提交中国专利局，申请号为 201910345010.1、发明名称为“动作识别方法、人机交互方法、装置及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机技术领域，特别涉及一种动作识别方法和装置、人机交互方法和装置。

背景技术

动作识别技术是对动作图像进行分析处理，从而理解动作图像中所包含动作的技术，随着计算机技术的发展，动作识别技术广泛应用于视频监控、体感游戏、虚拟现实、互联网医疗等多种领域，已成为一种热门的研究方向。

传统方案中对用户所做动作的识别方式，通常是用户在电子设备的拍摄范围内做出某种动作，电子设备对用户进行拍摄，得到对应的动作图像，进而基于该动作图像来识别用户是否做出了预设动作。然而，传统方案中进行动作识别的方式，主要是简单的对动作图片进行识别，而当动作图片采集的质量不够好时，很容易存在识别误差，因而存在动作识别准确性低的问题。

发明内容

根据本申请的各种实施例提供了一种动作识别方法、装置、电子设备及存储介质，人机交互方法、装置、电子设备及存储介质。

一种动作识别方法，由电子设备执行，所述方法包括：

从目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包

含的动态动作;

获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度, 所述预设特征信息用于描述预设动态动作; 其中, 获取到的多个相似度中最大的相似度为第一相似度; 及

当所述第一相似度大于第一预设阈值, 且所述第一相似度与第二相似度之间的差值大于第二预设阈值时, 确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作, 所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

在其中一个实施例中, 所述从目标视频中提取多个目标窗口, 包括:

从目标视频中提取多个视频帧; 及

以第一预设数量作为每个目标窗口中的视频帧数量, 以第二预设数量作为任两个相邻的目标窗口所包括的相同视频帧的数量, 对所述多个视频帧进行划分, 得到多个目标窗口。

在其中一个实施例中, 所述分别对所述每个目标窗口中的视频帧进行动作识别, 得到所述每个目标窗口的第一动作特征信息, 包括:

对于所述每个目标窗口中的每个视频帧, 提取所述视频帧中的多个身体关键点;

根据所述多个身体关键点的分布情况, 对所述视频帧进行动作识别, 得到所述视频帧的第二动作特征信息, 所述第二动作特征信息用于描述所述视频帧包含的静态动作; 及

分别将所述每个目标窗口中的视频帧的第二动作特征信息进行组合, 得到所述每个目标窗口的第一动作特征信息。

在其中一个实施例中, 所述预设动态动作由至少两个预设身体部位做出; 所述根据所述多个身体关键点的分布情况, 对所述视频帧进行动作识别, 得到所述视频帧的第二动作特征信息, 包括以下至少一项:

根据所述多个身体关键点在所述视频帧中的坐标以及所述多个身体关键点所属的身体部位, 确定所述视频帧中任两个预设身体部位之间的角度, 作为所述第二动作特征信息;

获取所述多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量, 作为所述第二动作特征信息, 所述参考视频帧在

所述视频帧之前，且与所述视频帧相隔第三预设数量的视频帧；及

获取任两个预设身体部位中的参考身体部位的尺寸，以及所述任两个预设身体部位之间的距离，将所述距离与所述参考身体部位的尺寸之间的比例作为第二动作特征信息。

在其中一个实施例中，所述第一动作特征信息为包括 M 个第一动作向量的第一动作矩阵，所述预设特征信息为包括 N 个预设动作向量的预设动作矩阵， M 和 N 为正整数，所述获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，包括：

创建相似度矩阵，所述相似度矩阵的行数为 M 、列数为 N ，或者行数为 N 、列数为 M ；

对于所述相似度矩阵中与第 i 个第一动作向量和第 j 个预设动作向量对应的指定位置，获取第一位置、第二位置和第三位置的相似度中的最大相似度与所述第 i 个第一动作向量与所述第 j 个预设动作向量之间的相似度的和，作为所述指定位置的相似度，所述第一位置为与第 $i-1$ 个第一动作向量和第 j 个预设动作向量对应的位置，所述第二位置为与第 $i-1$ 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置，所述第三位置为与第 i 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置， i 为不小于 1 且不大于 M 的正整数， j 为不小于 1 且不大于 N 的正整数；及

将所述相似度矩阵中与第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度，确定为所述第一动作特征信息与所述预设特征信息之间的相似度。

一种人机交互方法，由电子设备执行，所述方法包括：

通过配置的摄像头拍摄用户的动作视频，作为目标视频；

从所述目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似

度，所述预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度；

当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作，所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度；及

执行与所述预设动态动作对应的交互操作。

在其中一个实施例中，所述执行与所述预设动态动作对应的交互操作，包括以下至少一项：

确定所述用户的动作通过验证，并显示表示验证通过的预设动态效果；
为所述用户分配所述预设动态动作对应的虚拟奖励数据；及
为所述用户分配所述预设动态动作对应的操作权限。

一种动作识别装置，所述装置包括：

提取模块，用于从目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

动作识别模块，用于分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

相似度获取模块，用于获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，所述预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度；及

确定模块，用于当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作，所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

在其中一个实施例中，所述提取模块，包括：

获取单元，用于从目标视频中提取多个视频帧；及

划分单元，用于以第一预设数量作为每个目标窗口中的视频帧数量，以第二预设数量作为任两个相邻的目标窗口包含的相同视频帧的数量，对所述

多个视频帧进行划分，得到多个目标窗口。

在其中一个实施例中，所述动作识别模块，包括：

提取单元，用于对于所述每个目标窗口中的每个视频帧，提取所述视频帧中的多个身体关键点；

动作识别单元，用于根据所述多个身体关键点的分布情况，对所述视频帧进行动作识别，得到所述视频帧的第二动作特征信息，所述第二动作特征信息用于描述所述视频帧包含的静态动作；及

组合单元，用于分别将所述每个目标窗口中的视频帧的第二动作特征信息进行组合，得到所述每个目标窗口的第一动作特征信息。

在其中一个实施例中，所述预设动态动作由至少两个预设身体部位做出；所述动作识别单元还用于执行以下步骤中的至少一项：

根据所述多个身体关键点在所述视频帧中的坐标以及所述多个身体关键点所属的身体部位，确定所述视频帧中任两个预设身体部位之间的角度，作为所述第二动作特征信息；

获取所述多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量，作为所述第二动作特征信息，所述参考视频帧在所述视频帧之前，且与所述视频帧相隔第三预设数量的视频帧；及

获取任两个预设身体部位中的参考身体部位的尺寸，以及所述任两个预设身体部位之间的距离，将所述距离与所述参考身体部位的尺寸之间的比例作为第二动作特征信息。

在其中一个实施例中，所述第一动作特征信息为包括 M 个第一动作向量的第一动作矩阵，所述预设特征信息为包括 N 个预设动作向量的预设动作矩阵， M 和 N 为正整数，所述相似度获取模块，包括：

创建单元，用于创建相似度矩阵，所述相似度矩阵的行数为 M 、列数为 N ，或者行数为 N 、列数为 M ；

相似度获取单元，用于对于所述相似度矩阵中与第 i 个第一动作向量和第 j 个预设动作向量对应的指定位置，获取第一位置、第二位置和第三位置的相似度中的最大相似度与所述第 i 个第一动作向量与所述第 j 个预设动作向量之间的相似度的和，作为所述指定位置的相似度，所述第一位置为与第 $i-1$ 个第一动作向量和第 j 个预设动作向量对应的位置，所述第二位置为与第 $i-1$

个第一动作向量和第 $j-1$ 个预设动作向量对应的位置, 所述第二位置为与第 i 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置, i 为不小于 1 且不大于 M 的正整数, j 为不小于 1 且不大于 N 的正整数; 及

所述相似度获取单元, 还用于将所述相似度矩阵中与第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度, 确定为所述第一动作特征信息与所述预设特征信息之间的相似度。

一种人机交互装置, 所述装置包括:

拍摄模块, 用于通过配置的摄像头拍摄用户的动作视频, 作为目标视频;

提取模块, 用于从所述目标视频中提取多个目标窗口, 所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧;

动作识别模块, 用于分别对所述每个目标窗口中的视频帧进行动作识别, 得到所述每个目标窗口的第一动作特征信息, 所述第一动作特征信息用于描述所述目标窗口包含的动态动作;

相似度获取模块, 用于获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度, 所述预设特征信息用于描述预设动态动作; 其中, 获取到的多个相似度中最大的相似度为第一相似度;

确定模块, 用于当所述第一相似度大于第一预设阈值, 且所述第一相似度与第二相似度之间的差值大于第二预设阈值时, 确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作, 所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度; 及

执行模块, 用于执行与所述预设动态动作对应的交互操作。

在其中一个实施例中, 所述执行模块, 用于执行以下步骤中的至少一项:

确定所述用户的动作通过验证, 并显示表示验证通过的预设动态效果;

为所述用户分配所述预设动态动作对应的虚拟奖励数据; 及

为所述用户分配所述预设动态动作对应的操作权限。

一种电子设备, 所述电子设备包括处理器和存储器, 所述存储器中存储有计算机可读指令, 所述计算机可读指令被所述处理器执行时, 使得所述处理器执行如上述的动作识别方法的步骤。

一种电子设备，所述电子设备包括处理器和存储器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行如上述的人机交互方法的步骤。

一种非易失性的计算机可读存储介质，存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如上述的动作识别方法的步骤。

一种非易失性的计算机可读存储介质，存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如上述的人机交互方法的步骤。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征、目的和优点将从说明书、附图以及权利要求书变得明显。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，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。

图1是本申请实施例提供的一种动作识别方法的流程图；

图2是本申请实施例提供的一种检测身体关键点的流程图；

图3是本申请实施例提供的一种视频帧的身体关键点示意图；

图4是本申请实施例提供的一种获取视频帧的动作特征信息的流程图；

图5是本申请实施例提供的一种相似度矩阵的示意图；

图6是发明实施例提供的一种确定预设动态动作的流程图；

图7是本申请实施例提供的一种动作识别方法的流程图；

图8是本申请实施例提供的一种动作识别装置的结构示意图；

图9是本申请实施例提供的另一种动作识别装置的结构示意图；

图10是本申请实施例提供的一种人机交互装置的结构示意图；

图11是本申请实施例提供的一种电子设备的结构示意图；

图12是本申请实施例提供的一种服务器的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

在对本申请实施例进行详细地解释说明之前，首先对本申请实施例涉及到的概念进行如下解释：

1、静态动作：是指身体部位静止的状态下做出的动作，静态动作没有运动过程，身体部位在做出静态动作时，位置和姿态都不会发生变化。例如，静态动作可以包括抬头动作、倒立动作、或手臂弯曲的动作等。

2、动态动作：是指身体部位运动的状态下做出的动作，动态动作具有运动过程，身体部位在做出动态动作时，位置或姿态可以会发生变化。一个动态动作可以看作是多个静态动作组合得到。

例如，动态动作可以包括挥手动作、摇摆动作、或扩胸运动的动作等。

图 1 是本申请实施例提供的一种动作识别方法的流程图。参见图 1，应用于电子设备，该电子设备可以为手机、平板电脑、计算机、体感游戏机、或门禁设备等多种类型的设备。该方法包括：

S101、获取目标视频，从目标视频中提取多个视频帧。

本申请实施例以识别预设动态动作为例，电子设备确定预设动态动作，当获取到目标视频时，对目标视频进行动作识别，以识别目标视频中包含的预设动态动作。

其中，目标视频可以为电子设备获取到的任意视频。从内容上来讲，目标视频可以为舞蹈类视频、娱乐新闻类视频、体育类视频等。从来源上来讲，目标视频可以为电子设备通过摄像头拍摄的视频，或者从互联网下载的视频，或者其他设备发送的视频等。

需要说明的是，本申请实施例中所提及的“多个”具体指“多于一个”。

在一些实施例中，电子设备接收动作识别指令，根据动作识别指令调用视频数据库，获取用户从视频数据库中选择的视频，作为目标视频，从而对该目标视频进行动作识别。其中，该视频数据库存储有一个或多个视频。

例如，电子设备安装视频监控客户端，该视频监控客户端用于录制视频，将录制的视频存储于视频数据库中。当用户需要核查某些视频中是否包含预

设动态动作时，可以触发动作识别指令，从视频数据库中选择希望核查的视频，从而触发动作识别过程。

在一些实施例中，电子设备接收服务器下发的动作识别指令，该动作识别指令携带预设动态动作，用于指示电子设备识别包含该预设动态动作的视频，则电子设备根据动作识别指令，将本地存储的每个视频分别作为目标视频，从而触发动作识别过程。

其中，该服务器用于进行视频安全检测，可以根据相关规定定期或随机地向电子设备下发携带一些违规的动态动作的动作识别指令。

在一些实施例中，电子设备接收动作识别指令，打开摄像头，通过摄像头拍摄视频，将当前拍摄的视频作为目标视频。例如，电子设备为体感游戏机，当检测到开始进行游戏的操作时，确定接收到动作识别指令，从而根据动作识别指令拍摄用户的视频，作为目标视频，从而触发动作识别过程，识别用户是否做出了预设动态动作，从而根据识别结果进行后续操作。

需要说明的是，当目标视频为实时拍摄的视频时，电子设备会随着视频的拍摄，对当前已拍摄的视频片段进行动作识别，后续再拍摄得到新的视频片段时，再对新拍摄的视频片段进行动作识别。而当目标视频为本地存储的视频时，电子设备可以以目标视频为整体，对该目标视频进行动作识别。

本申请实施例中，目标视频包括按照时间先后顺序依次排列的多个视频帧，每个视频帧中可能会包括一个或多个静态动作，而预设动态动作包括多个静态动作。因此，为了识别目标视频包含的预设动态动作，从目标视频中提取多个视频帧，后续再根据多个视频帧进行处理。

其中，该多个视频帧可以包括目标视频中的所有视频帧，或者仅包括目标视频中的部分连续视频帧。电子设备在从目标视频中提取视频帧时，可以按照目标视频中视频帧的排列顺序依次进行提取，或者从目标视频中提取预设数量的连续视频帧。例如，以目标视频中的任一个视频帧为提取的第一个视频帧，按照该视频帧之后各个视频帧的排列顺序，提取预设数量的连续视频帧。其中，该预设数量可以根据动作识别的复杂度要求确定，或者根据目标视频中的视频帧数量确定。

在一些实施例中，当目标视频为实时拍摄的视频时，电子设备会随着视频的进行，每当拍摄到预设数量的视频帧时，提取该预设数量的视频帧，作

为一个视频片段，对该视频片段进行动作识别，后续再次拍摄得到新的预设数量的视频帧时，继续提取新的预设数量的视频帧，得到另一个视频片段，对另一个视频片段进行动作识别，也即是每次将当前时刻之前的预设数量的视频帧作为输入，进行动作识别。而当目标视频为本地存储的视频时，电子设备从目标视频中提取多个视频帧。

S102、以第一预设数量作为每个目标窗口中的视频帧数量，以第二预设数量作为任两个相邻的目标窗口所包括的相同视频帧的数量，对多个视频帧进行划分，得到多个目标窗口。

每个目标窗口包括第一预设数量的连续视频帧，多个目标窗口构成一个窗口序列，后续电子设备可以以目标窗口为单位，对窗口序列中的每个目标窗口进行动作识别，识别每个目标窗口包含的预设动态动作。

其中，每个目标窗口中包括的视频帧的数量为第一预设数量，且任两个相邻的目标窗口包含的相同视频帧的数量为第二预设数量。其中，该第一预设数量为大于 1 的整数，该第二预设数量为大于 0 的整数。并且，由于第一预设数量不大于第二预设数量时，无法实现相邻目标窗口中包含第二预设数量的相同的视频帧，因此第一预设数量大于第二预设数量。

例如，该第一预设数量为 30，第二预设数量为 29，则每个目标窗口中包括 30 帧视频帧，任两个相邻的目标窗口中包括 29 帧相同的视频帧。

在一些实施例中，电子设备设置预设步长，该预设步长用于规定任两个相邻目标窗口的起始视频帧之间的距离，且该预设步长为第一预设数量与第二预设数量的差值。则电子设备将提取的多个视频帧中的起始视频帧作为第一起始视频帧，从第一起始视频帧开始提取第一预设数量的视频帧，得到一个目标窗口。继续确定与第一起始视频帧之间的距离为预设步长的第二起始视频帧，从该第二起始视频帧开始，提取第一预设数量的视频帧，得到下一个目标窗口，以此类推，即可依次提取出间隔预设步长的目标窗口。

例如，当第一预设数量为 30，预设步长为 2 时，以目标视频中的第一个视频帧为起点，将第 1 帧至第 30 帧作为第一个目标窗口，将第 3 帧至第 32 帧作为第二个目标窗口，以此类推，可以从目标视频中提取多个目标窗口。

需要说明的是，本申请实施例仅是以电子设备提取多个视频帧后按照第一预设数量和第二预设数量进行划分为例，对得到目标窗口的过程进行说明。

而本申请实施例中，还可以采用其他可能实现方式，只需保证在时间维度上对目标视频进行划窗处理，从目标视频中提取多个目标窗口，且每个目标窗口包括第一预设数量的连续视频帧即可。

例如，以第一预设数量作为每个目标窗口中的视频帧的数量，电子设备将提取的多个视频帧中的起始视频帧作为第一起始视频帧，从第一起始视频帧开始提取第一预设数量的视频帧，得到一个目标窗口。此时随机确定一个步长，继续确定与第一起始视频帧之间的距离为该步长的第二起始视频帧，从该第二起始视频帧开始，提取第一预设数量的视频帧，得到下一个目标窗口，后续继续随机确定另一个步长，以此类推，即可依次提取出多个目标窗口，且任两个相邻的目标窗口包括的相同视频帧的数量不固定。

S103、对于每个目标窗口中的每个视频帧，提取该视频帧中的多个身体关键点。

由于目标窗口的动态动作与每个视频帧的静态动作有关，因此为了对目标窗口进行动态动作的识别，可以先对每个视频帧进行特征提取，得到每个视频帧的动作特征信息，再根据每个视频帧的动作特征信息获取目标窗口的动作特征信息。

因此，对于每个目标窗口中的每个视频帧，电子设备对该视频帧进行关键点提取，提取该视频帧中的多个身体关键点，该多个身体关键点用于描述该视频帧中包括的对象。其中，该视频帧中包括的对象可以为人的身体、动物的身体或者其他种类的身体。视频帧中的身体关键点可以为眼睛关键点、耳朵关键点、鼻子关键点、肩膀关键点、手部关键点、脚部关键点等。

在一些实施例中，电子设备将该视频帧输入到关键点检测模型中，基于该关键点检测模型对该视频帧进行关键点检测，可以得到该视频帧中的多个身体关键点。

如图 2 所示，电子设备获取到目标视频后，从目标视频中提取多个目标窗口，获取每个目标窗口中的视频帧集，将视频帧集中的每个视频帧输入到关键点检测模型中，基于该关键点检测模型即可输出每个视频帧中的身体关键点。

可选地，电子设备将视频帧输入至关键点检测模型中，关键点检测模型会根据视频帧生成与多个身体部位对应的热力图，即一个热力图对应一个身

体部位。其中，每个热力图中的每个像素点的像素值代表该像素点在该视频帧中对应的像素点属于对应的身体部位中的关键点的概率，热力图中像素点的像素值越大，表示在该视频帧中对应的像素点越可能是该身体部位上的身体关键点，热力图中像素点的像素值越小，表示在该视频帧中对应的像素点越可能不是该身体部位上的身体关键点。获取热力图中的像素值最大的像素点，将该像素点在该视频帧中对应的像素点确定为该热力图对应的身体部位的身体关键点。

其中，在获取热力图中的像素值最大的像素点时，可以采用 NMS (Non-Maximum Suppression, 非极大值抑制) 算法进行局部最大搜索，抑制不是极大值的像素值，从而获取热力图中的像素值最大的像素点，或者还可以采用其他方式。

且将视频帧输入至关键点检测模型中，关键点检测模型通过热力图确定视频帧中的身体关键点后，可以获取该视频帧中任两个身体关键点构成的向量，通过确定该任两个身体关键点构成的向量是否与预设向量相同，当确定该任两个身体关键点与预设向量相同时，确定该任两个身体关键点属于同一对象上与该预设向量对应的身体部位。

如图 3 所示，该视频帧中包括三个对象，且检测出 6 个身体关键点，将该 6 个身体关键点中的任两个关键点相连，构成多个向量，分别判断该多个向量与预设向量是否相同，从而可以确定该 6 个身体关键点属于的对象。

其中，该关键点检测模型的模型结构可以为 darknet (一种模型结构)，包括特征提取子模型和特征处理子模型，特征提取子模型包括多个网络层，每个网络层采用的卷积核尺寸可以相同也可以不同，且每个网络层的输出作为下一个网络层的输入，该特征提取子模型的多个网络层的输出、输入和卷积核尺寸可以如表 1 所示。特征处理子模型包括两个分支，第一个分支用于获取热力图，第二个分支用于确定身体关键点所属的身体部位。且每个分支包括多个网络层，每个网络层采用的卷积核尺寸可以相同也可以不同，且每个网络层的输出作为下一个网络层的输入，第一个分支的多个网络层的输出、输入和卷积核尺寸可以如表 2 所示，第二个分支的多个网络层的输出、输入和卷积核尺寸可以如表 3 所示。

表 1

输出	输入	卷积核尺寸
输入层		
16	3	3
Maxpooling (池化层)		
32	16	3
Maxpooling (池化层)		
16	32	1
128	16	3
16	128	1
128	16	3
Maxpooling (池化层)		
32	128	1
256	32	3
32	256	1
256	32	3
Maxpooling (池化层)		
64	256	1
512	64	3
64	512	1
512	64	3
256	512	3
128	256	3

表 2

输出	输入	卷积核尺寸
256	256	3
256	256	3
256	256	3
512	256	1

38	512	1
----	-----	---

表 3

输出	输入	卷积核尺寸
256	256	3
256	256	3
256	256	3
512	256	1
19	512	1

在训练关键点检测模型时，电子设备构建初始化的关键点检测模型，获取多个样本视频帧和每个样本视频帧的身体关键点，根据每个样本视频帧和每个样本视频帧中的身体关键点对关键点检测模型进行训练，从而可以使关键点检测模型具备检测身体关键点的能力，得到训练后的关键点检测模型。

在一些实施例中，当训练关键点检测模型时，电子设备可以采用 Coco（Common Objects in Context，上下文中的公共对象）数据集对关键点检测模型进行训练。该关键点检测模型可以为卷积神经网络模型、全卷积神经网络模型、mobilenet 模型（一种用于移动终端的神经网络模型）、Shufflenet 模型（一种用于移动终端的神经网络模型）等，或者还可以为基于 Heatmap（热力图）与 PAF（Part Affinity Fields，部分亲和字段）结合的模式等。

S104、根据多个身体关键点的分布情况，对视频帧进行动作识别，得到视频帧的第二动作特征信息，第二动作特征信息用于描述视频帧包含的静态动作。

电子设备提取视频帧中的多个身体关键点时，即可确定每个身体关键点的坐标，从而确定每个身体关键点在视频帧中的位置，进而确定多个身体关键点的分布情况。电子设备则根据多个身体关键点的分布情况获取用于描述视频帧中包含的静态动作的第二动作特征信息。

根据该多个身体关键点得到第二动作特征信息的方式可以包括步骤 S1041-S1043 中的至少一种：

S1041、根据多个身体关键点在视频帧中的坐标以及多个身体关键点所属的身体部位，确定视频帧中任两个预设身体部位之间的角度，作为第二动作特征信息。

其中，预设动态动作由至少两个预设身体部位做出。因此，电子设备可以根据多个身体关键点在视频帧中的坐标，获取属于任两个预设身体部位中每个预设身体部位的身体关键点，根据该任两个预设身体部位中每个预设身体部位的身体关键点的坐标，确定表示每个预设身体部位的向量，根据确定的任两个预设身体部位的向量确定该任两个预设身体部位之间的角度，将该角度作为第二动作特征信息。

例如，电子设备获取身体的肘关节相对肩膀的角度、手腕关节相对肘关节的角度，将获取的角度作为第二动作特征信息。

由于一个对象的姿态可以根据多个身体关键点构成的骨架来表示，因此本申请实施例通过将任两个预设身体部位之间的角度作为第二动作特征信息，使第二动作特征信息能够表示多个身体关键点构成的骨架的姿态，也即是该对象的静态动作。

在一些实施例中，当获取到任两个预设身体部位之间的角度后，为了后续能够根据获取到的角度计算预设动态动作和目标窗口中的动态动作的相似度，电子设备可以先对获取到的角度进行编码，从而能够将该角度归一化，通过归一化后的角度来表示第二动作特征信息。

其中，编码的方式可以采用 Onehot (独热编码)，也即是将 360 度的角度划分为多份，例如划分为 4 份：0-90 度、91 度-180 度、181 度-270 度、271 度-360 度，每份均对应于一个编码。当获取的角度属于任一个范围时，将该角度所属的范围对应的编码设置为 1，其余范围对应的编码设置为 0，得到该角度对应的编码。

需要说明的是，当做出预设动态动作的预设身体部位包括三个或三个以上时，电子设备将任两个预设身体部位作为一组，每组预设身体部位可以得到对应的角度，则将多组预设身体部位对应的角度均作为第二动作特征信息。

S1042、获取多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量，作为第二动作特征信息。

其中，参考视频帧在视频帧之前，且与视频帧相隔第三预设数量的视频帧。由于目标窗口中包含一个动态动作，当前视频帧中的身体关键点与参考视频帧中对应的身体关键点之间可能会发生位移，因此，电子设备可确定当前视频帧中的身体关键点与参考视频帧中对应的身体关键点之间的位移量，

该位移量可以表示当前视频帧中的身体关键点相对于参考视频帧中对应的身体关键点的运动幅度，将该位移量作为第二动作特征信息。

在一些实施例中，当获取多个身体关键点后，电子设备获取该身体关键点中的任一个身体关键点的坐标以及参考视频帧中对应的身体关键点的坐标，该坐标包括横坐标和纵坐标，确定该视频帧中的关键点与参考视频帧中对应的关键点的横坐标之间的差值以及纵坐标之间的差值，将横坐标之间的差值与纵坐标之间的差值均作为第二动作特征信息。

例如，当获取到多个身体关键点后，电子设备获取该视频帧中左手手腕关键点与参考视频帧中左手手腕关键点的横坐标之间的差值以及纵坐标之间的差值，该参考视频帧与该视频帧之间间隔5帧。

在一些实施例中，当获取多个身体关键点后，电子设备获取该多个身体关键点中的任一身体关键点的坐标以及参考视频帧中对应的身体关键点的坐标，获取参考视频帧中对应的身体关键点在当前视频帧中对应的位置的坐标，确定该坐标和当前视频帧中的身体关键点的坐标之间的距离，将该距离作为第二动作特征信息。

S1043、获取任两个预设身体部位中的参考身体部位的尺寸，以及该任两个预设身体部位之间的距离，将该距离与参考身体部位的尺寸之间的比例作为第二动作特征信息。

由于预设动态动作由至少两个预设身体部位做出，因此电子设备可以根据多个身体关键点在视频帧中的坐标，获取该视频帧中任两个预设身体部位之间的距离，并且对于这两个预设身体部位，还可以将其中一个身体部位作为参考身体部位，获取该参考身体部位的尺寸，确定这两个预设身体部位之间的距离与该参考身体部位的尺寸的比例，将该比例作为第二动作特征信息。

其中，该参考身体部位可以为身体中尺寸不发生改变的部位，例如该参考身体部位可以为肩膀、脖子等，则参考身体部位的尺寸可以为肩膀的尺寸、脖子的尺寸等。

例如，预设动态动作作为挥手动作，由肩膀和手臂做出，则将肩膀作为参考身体部位，可以根据左手手腕关键点和左肩膀关键点确定左手手腕到肩膀的距离，根据左肩膀关键点和右肩膀关键点可以确定肩膀的宽度，根据左手手腕到肩膀的距离和肩膀的宽度可以确定比例，将该比例作为第二动作特征

信息。

通过将任两个预设身体部位之间的距离与参考身体部位的尺寸的比例作为第二动作特征信息，而不是仅计算任两个预设身体部位之间的距离，考虑到了不同对象尺寸不同的问题，使获取到的动作特征信息可以兼容不同尺寸的对象，避免受到尺寸的影响。

需要说明的是，当做出预设动态动作的预设身体部位包括三个或三个以上时，电子设备将任两个预设身体部位作为一组，每组预设身体部位可以得到对应的角度，则将多组预设身体部位对应的角度均作为第二动作特征信息。

本申请实施例中，可以通过步骤 S1041-S1043 中的任一项可以分别获取第二动作特征信息，该第二动作特征信息可以准确地对视频帧中包含的静态动作进行描述，提高了准确性，便于后续根据视频帧中包含的静态动作描述目标窗口中包含的动态动作。

另外，还可以将上述步骤 S1041-S1043 中的任两项或三项进行组合，从而可以得到组合后的第二动作特征信息，通过组合后的第二动作特征信息描述视频帧中的静态动作，可以更加准确全面地对视频帧中包含的静态动作进行描述，提高了准确性，便于后续根据视频帧中包含的静态动作描述目标窗口中包含的动态动作。

S105、分别将每个目标窗口中的视频帧的第二动作特征信息进行组合，得到每个目标窗口的第一动作特征信息。

其中，该第一动作特征信息用于描述目标窗口包含的动态动作。

每个目标窗口中包括连续的多个视频帧，每个视频帧的第二动作特征信息表示对应的视频帧中包含的静态动作。由于目标窗口的动态动作可以由每个视频帧中的静态动作组合得到，因此将每个目标窗口中的多个视频帧的第二动作特征信息进行组合，得到组合后的第一动作特征信息，该第一动作特征信息可以描述目标窗口中包含的动态动作。

如图 4 所示，电子设备将视频帧集中的每个视频帧输入关键点检测模型，基于该关键点检测模型，确定每个视频帧中的身体关键点集，根据每个视频帧中的身体关键点集确定左上臂与左肩膀的角度 45° 、左手下臂与左上臂的角度 25° ，左手关键点的横坐标位移 0.4、左手关键点的纵坐标位移 0.5、左手关键点到肩膀的距离与肩宽的比值 1.5、右手关键点到肩膀的距离与肩宽

的比值 1.6, 将得到的数据进行组合, 得到该视频帧的第二动作特征信息 [45°, 25°][0.4, 0.5][1.5, 1.6], 将第二动作特征信息采用 onehot 进行编码, 得到编码后的动作特征信息[0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0.4, 0.5, 1.5, 1.6], 将得到的每个视频帧的编码后的动作特征信息进行组合, 可以得到用于表示目标窗口的第一动作特征信息。

S106、获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度。

其中, 预设特征信息用于描述预设动态动作, 可以作为预设动态动作的特征模板, 在后续过程中用于与目标窗口的动作特征信息进行匹配对比。

该预设特征信息的获取过程包括: 获取多个仅包含预设动态动作的视频, 对该多个视频进行动作识别, 得到每个视频的特征信息, 根据得到的每个视频的特征信息进行统计, 得到预设特征信息, 从而该预设特征信息可以表示预设动态动作。其中, 根据多个特征信息进行统计的方式可以为计算总和、计算平均值等。

在本申请实施例中, 通过获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度, 可以确定每个目标窗口的第一动作特征信息与预设特征信息之间的相似程度, 从而根据确定的相似度可以确定目标窗口中的动态动作是否与预设动态动作相同。

在一些实施例中, 电子设备确定第一动作特征信息对应的第一动作矩阵, 确定预设特征信息对应的预设动作矩阵, 再确定第一动作矩阵和预设动作矩阵之间的相似度, 将确定的相似度作为第一动作特征信息和预设特征信息之间的相似度。

其中, 第一动作特征信息为包括 M 个第一动作向量的第一动作矩阵, 预设特征信息为包括 N 个预设动作向量的预设动作矩阵, 且 M 和 N 为正整数。且该第一动作矩阵和该预设动作矩阵均可以数组的形式存储于电子设备中, 或者采用其他存储形式。

例如 M 等于 30, 且每个第一动作向量为由编码后的两个预设身体部位之间的角度、多个身体关键点中任一身体关键点与参考视频帧中对应的身体关键点的横坐标之差、纵坐标之差、任两个身体部位之间的比值组合构成 38 维的向量, 因此该第一动作矩阵为 38*30 的矩阵。

例如 N 等于 20, 且每个预设动作向量均由编码后的两个预设身体部位之间的角度、多个身体关键点中任一关键点的横坐标之差、纵坐标之差、任两个身体部位之间的比值组合构成 38 维的向量, 因此该预设动作矩阵为 38×20 的矩阵。

其中, 确定动作矩阵和预设动作矩阵之间的相似度可以包括步骤 S1061-S1063:

S1061、创建相似度矩阵。

其中, 该相似度矩阵的尺寸由第一动作矩阵包括的第一动作向量的数量 M 和预设动作矩阵包括的预设动作向量的数量 N 确定, 当该相似度矩阵的每行代表一个预设动作向量, 每列代表一个第一动作向量时, 该相似度矩阵的行数为 N, 列数为 M, 而当该相似度矩阵的每行代表一个第一动作向量, 每列代表一个预设动作向量时, 该相似度矩阵的行数为 M, 列数为 N。

S1062、对于相似度矩阵中与第 i 个第一动作向量和第 j 个预设动作向量对应的指定位置, 获取第一位置、第二位置和第三位置的相似度中的最大相似度与第 i 个第一动作向量与第 j 个预设动作向量之间的相似度的和, 作为指定位置的相似度。

其中, 第一位置为与第 i-1 个第一动作向量和第 j 个预设动作向量对应的位置, 也即是指定位置的左侧位置, 第二位置为与第 i-1 个第一动作向量和第 j-1 个预设动作向量对应的位置, 也即是指定位置的左下侧位置, 第三位置为与第 i 个第一动作向量和第 j-1 个预设动作向量对应的位置, 也即是指定位置的下侧位置, i 为不小于 1 且不大于 M 的正整数, j 为不小于 1 且不大于 N 的正整数。因此从第 1 个第一动作向量和第 1 个预设动作向量对应的位置开始确定相似度, 直至确定第 N 个第一动作向量和第 M 个预设动作向量对应的位置的相似度。

相似度矩阵中, 指定位置的相似度用于表示目标视频中的第 i 帧以及第 i 帧之前的动态动作与做出预设动态动作的视频中第 j 帧以及第 j 帧之前的预设动态动作之间的相似度, 可以在一定程度上代表目标视频中到第 i 帧为止的动态动作与预设动态动作是否相同。

需要说明的是, 该相似度可以采用欧式距离、余弦相似度、曼哈顿距离、马氏距离等不同的方式进行表示, 且不同方式的相似度与其所表示的第一动

作特征信息与预设特征信息之间的相似程度的关系不同。例如，欧式距离与相似程度呈负相关关系，欧式距离越大，代表第一动作特征信息与预设特征信息越不相似。而余弦相似度与相似程度呈正相关关系，余弦相似度越大，代表第一动作特征信息与预设特征信息越相似。

则按照上述步骤 S1061-S1062 获取指定位置的相似度时，如果相似度与第一动作特征信息与预设特征信息之间的相似程度呈正相关关系，电子设备则获取第一位置、第二位置和第三位置的相似度中的最大相似度与第 i 个第一动作向量与第 j 个预设动作向量之间的相似度的和，作为指定位置的相似度。如果相似度与第一动作特征信息与预设特征信息之间的相似程度呈负相关关系，则获取第一位置、第二位置和第三位置的相似度中的最小相似度与第 i 个第一动作向量与第 j 个预设动作向量之间的相似度的和，作为指定位置的相似度。

例如，如图 5 所示，该相似度矩阵包括 16 个位置，计算位置 6 的相似度时，获取位置 1、位置 2、位置 5 的相似度分别为 0.4、0.3、0.6，电子设备获取位置 6 对应的第一动作向量和预设特征向量的相似度为 0.1，则位置 6 的相似度为 0.7。电子设备依次确定相似度矩阵中位置 1 到位置 16 的相似度，将位置 16 的相似度确定为第一动作特征信息与预设特征信息之间的相似度。

另外，该相似度矩阵包括 $M*N$ 个位置，电子设备确定该 $M*N$ 个位置中每个位置对应的第一特征向量和预设特征向量之间的相似度，根据确定的每个位置对应的第一特征向量和预设特征向量之间的相似度，以及该位置的左侧位置、左下侧位置和下侧位置的相似度，可以确定每个位置的相似度。

其中，第一特征向量和预设特征向量之间的相似度可以采用多种方式表示，相似度的表示方式包括欧式距离、余弦相似度、曼哈顿距离、马氏距离等。

在一些实施例中，若要确定第 M 个第一特征向量和第 N 个预设特征向量对应的位置的相似度，电子设备要先获取到该位置的左侧位置、左下侧位置以及下侧位置的相似度以及第 M 个第一特征向量和第 N 个预设特征向量之间的相似度。而如果确定左侧位置、左下侧位置以及下侧位置的相似度，仍要确定这些位置的左侧位置、左下侧位置、下侧位置的相似度，以此类推，直至第 1 个第一特征向量与第 1 个预设特征向量对应的位置。在此过程中，相

似度矩阵中的有些位置无需参与计算，因此可以将无需参与计算的位置之外的位置作为预设位置，电子设备只确定该相似度矩阵中预设位置的相似度，而无需确定该相似度矩阵中除预设位置以外的其他位置的相似度，减少了计算量，从而加快了相似度的计算速率。

需要说明的是，当指定位置位于相似度矩阵的边缘时，可能会导致指定位置不存在左侧位置、左下侧位置或下侧位置，因此，当指定位置仅存在左侧位置、左下侧位置和下侧位置中的两个位置时，获取两个位置中最大的相似度，确定该指定位置对应的第一动作向量与预设特征向量的相似度与该最大相似度的和，作为该指定位置的相似度。而当指定位置仅存在左侧位置、左下侧位置和下侧位置中的一个位置时，直接确定该指定位置对应的第一动作向量与预设特征向量的相似度与该一个位置的相似度的和，作为该指定位置的相似度。而当指定位置不存在左侧位置、左下侧位置和下侧位置时，直接将该指定位置对应的第一动作向量和预设特征向量之间的相似度作为该指定位置的相似度。

S1063、将相似度矩阵中与第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度，确定为第一动作特征信息与预设特征信息之间的相似度。

通过步骤 S1062 获取相似度矩阵中每个位置的相似度后，电子设备将第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度确定为第一动作特征信息与预设特征信息之间的相似度。

例如，计算第一动作特征信息与预设特征信息之间的相似度的代码如下：

```
int DTWDistance ( s: array[1...n], t: array[1...m] ) {  
    DTW: =array[0...n, 0...m]  
    for i: = 1 to n  
        DTW[1, n]: =infinity  
    for i: = 1 to m  
        DTW[1, m]: =infinity  
    DTW[0, 0]: =0  
    for i: =1 to n  
        for j: =1 to m
```

```

cost: =d ( s[i], t[j] )
DTW[i, i]: =cost+minimum ( DTW[i-1, j], //insertion
                           DTW[i, j-1], // deletion
                           DTW[i-1, j] ) //match
return DTW[n, m]
}

```

本申请实施例构建用于描述目标窗口中包含的动态动作的第一动作矩阵，创建该第一动作矩阵与预设动作矩阵的相似度矩阵，可以根据相似度矩阵确定该第一动作矩阵与预设动作矩阵的相似度，该相似度可以作为对目标窗口的打分，衡量目标窗口中包含的动态动作与预设动态动作的相似程度。利用动态规划的思想，采用自底向上的方式计算相似度，快速地获取到了目标窗口中包含的动态动作与预设动态动作的相似度，且提高了准确性。

S107、根据获取到的至少一个相似度，识别多个目标窗口中包含的预设动态动作。

从目标视频中提取到多个目标窗口，电子设备则可以获取到该多个目标窗口中每个目标窗口的第一动作特征信息所对应的一个相似度，此时，获取多个相似度中最大的相似度作为第一相似度。当第一相似度大于第一预设阈值时，表示第一目标窗口包含的动态动作与预设动态动作的相似程度较高，可以认为第一相似度对应的第一目标窗口包含的动态动作为预设动态动作。而当第一相似度不大于第一预设阈值时，可以认为多个目标窗口包含的动态动作均不是预设动态动作。

当获取多个相似度中最大的第一相似度时，电子设备可以按照从小到大的顺序将多个相似度进行排列，获取最后一个相似度，该最后一个相似度即为多个相似度中最大的第一相似度。或者电子设备按照从大到小的顺序将多个相似度进行排列，获取第一个相似度，该第一个相似度即为多个相似度中最大的第一相似度。

例如，当获取的相似度为 0.4、0.5、0.6、0.7、0.8 且第一预设阈值为 0.7 时，获取最大的相似度 0.8，且该相似度 0.8 大于第一预设阈值，则确定该相似度 0.8 对应的第一目标窗口包含的动态动作为预设动态动作。

本申请实施例中，当第一相似度大于第一预设阈值时，获取与第一相似

度对应的第一目标窗口相邻的第二目标窗口，确定第二目标窗口的第一动作特征信息所对应的第二相似度与第一相似度之间的差值，当差值大于第二预设阈值时，确定第一目标窗口包含的动态动作为预设动态动作。

通过确定第二目标窗口对应的第二相似度与第一相似度之间的差值是否大于第二预设阈值，电子设备可以确定第一目标窗口中包含的动态动作与第二目标窗口包含的动态动作之间的差别是否足够大，当第一目标窗口中包含的动态动作与第二目标窗口包含的动态动作差别足够大时，表示第二目标窗口中包含的动态动作与预设动态动作区别较大，即第一目标窗口中包含的动态动作已经结束，可以确定该第一目标窗口中的动态动作为预设动态动作。而当第一目标窗口中包含的动态动作与第二目标窗口包含的动态动作差别不够大时，表示第二目标窗口中包含的动态动作与预设动态动作区别较小，即第一目标窗口中包含的动态动作还未结束，第一目标窗口中包含的动态动作与第二目标窗口中包含的动态动作可能为连续的动作，此时不会确定该第一目标窗口中的动态动作为预设动态动作。

如图 6 所示，获取多个目标窗口的第一动作特征信息，获取每个目标窗口中的第一动作特征信息与预设动作特征信息之间的相似度，获取的相似度结果为[0.4、0.45、0.6、0.8、0.6、0.45]，确定该 0.8 是否为局部最大的相似度，且该 0.8 大于第一预设阈值，且 0.8 与 0.6 之间的差值大于第二预设阈值，因此可以确定用户完成了预设动态动作。

需要说明的第一点是，获取到多个目标窗口，且多个目标窗口的相似度中最大的第一相似度大于第一预设阈值时，获取该第一相似度对应的第一目标窗口相邻的第二目标窗口。当该第一目标窗口相邻的第二目标窗口存在一个时，电子设备确定该第一目标窗口和第二目标窗口之间的差值，当该差值大于第二预设阈值时，确定第一目标窗口包含的动态动作为预设动态动作。而当该第一目标窗口相邻的第二目标窗口存在两个时，电子设备确定该第一目标窗口与该两个第二目标窗口中每个第二目标窗口之间的差值，得到两个差值，当该两个差值均大于第二预设阈值时，电子设备确定第一目标窗口包含的动态动作为预设动态动作。

需要说明的第二点是，本申请实施例中仅是以目标视频中仅识别一个对象的动态动作进行说明，在另一实施例中，当目标视频中包括多个对象时，

通过关键点检测模型可以确定每个视频帧中的身体关键点，且可以确定每个身体关键点所属的对象，因此可以根据每个对象的身体关键点分别执行步骤S104-S107，从而可以确定每个对象的动态动作是否为预设动态动作。

需要说明的第三点是，本申请实施例中的动作识别方法仅是以执行主体为电子设备为例进行说明，在另一实施例中，该动作识别方法的执行主体还可以为服务器，电子设备将目标视频发送给服务器，服务器执行步骤S101-S107，确定目标窗口中包含的动态动作是否为预设动态动作，将执行结果发送给电子设备。

其中，该服务器可以为一台服务器，或者由若干台服务器组成的服务器集群，或者是一个云计算服务中心。

本申请实施例提供的动作识别方法，通过从目标视频中提取多个目标窗口，根据每个目标窗口中的视频帧，获取每个目标窗口的第一动作特征信息，第一动作特征信息用于描述目标窗口包含的动态动作，获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，预设特征信息用于描述预设动态动作，当获取到的多个相似度中最大的第一相似度大于第一预设阈值，且第一相似度与相邻目标窗口对应的第二相似度之间的差值大于第二预设阈值时，表示第一相似度对应的第一目标窗口中的动态动作与相邻目标窗口中的动态动作差别足够大，不会受到相邻目标窗口中动态动作的影响，因此确定第一目标窗口包含的动态动作为预设动态动作。本申请实施例提供了一种识别预设动态动作的方案，能够以目标视频中的目标窗口为单位，对目标窗口中的视频帧进行动作识别，从而识别目标窗口中包含的预设动态动作，突破了相关技术中只能识别静态动作的限制，扩展了应用范围，提高了灵活性。而且通过计算相邻目标窗口的相似度之间的差值的方式确定相邻目标窗口中动作是否相互影响，避免由于相邻目标窗口之间的影响而造成识别错误的问题，提高了准确率。而且动态动作的识别过程中不会受到动态动作频率的限制，可以识别出任意频率的预设动态动作，提高了识别动态动作的鲁棒性。

另外，利用动态规划的思想，采用自底向上的方式计算相似度，快速地获取到了目标窗口中包含的动态动作与预设动态动作的相似度，且提高了准确性。

相关技术中，如果只有静态的动作图像，则仅能够根据静态的动作图像

识别出用户在某一时刻做出的静态动作，但无法识别出用户做出的完整的动态动作。因此采用这种识别方式无法自然地与用户进行人机交互，交互效果非常生硬。而若要识别出用户的动态动作，需要其他传感设备的辅助，也会增加硬件成本，使用不便。而本申请实施例中，可以根据动态的动作视频识别出用户做出的预设动态动作，而不仅是某一时刻的静态动作，因此采用这种识别方式可以自然地与用户进行人机交互，提升了交互效果。而且无需额外增加传感设备，节省了硬件设备，操作非常便捷。

图 7 是本申请实施例提供的一种动作识别方法的流程图，如图 7 所示，该方法包括：

S701、从目标视频中提取 4 个目标窗口，每个目标窗口包括连续的多个视频帧。

S702、确定目标窗口中每个视频帧的第二动作特征信息：

左手上臂与左肩膀的角度 45° 、左手下臂与左手上臂的角度 25° 、相隔 5 个视频帧的左手关键点的横坐标位移 0.4、纵坐标位移 0.5，左手关键点到肩膀的距离与肩宽的比值 1.5、右手关键点到肩膀的距离与肩宽的比值 1.6。则组合得到的第二动作特征信息为 $[45^\circ, 25^\circ][0.4, 0.5][1.5, 1.6]$ 。

S703、将组合得到的第二动作特征信息进行独热编码，得到用于表示该视频帧的动作特征信息 $[0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0.4, 0.5, 1.5, 1.6]$ ，将每个视频帧的动作特征信息进行组合，得到用于描述目标窗口的第一动作矩阵 $M = \{ [0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0.4, 0.5, 1.5, 1.6][...] \dots \}$ 。

S704、确定每个目标窗口的第一动作矩阵和预设特征矩阵的相似度，得到相似度序列 $[0.4, 0.45, 0.8, 0.6]$ 。

S705、第 3 个目标窗口的相似度最大且大于第一预设阈值，表示第 3 个目标窗口包含的动态动作作为预设动态动作。

本公开实施例还提供了一种人机交互方法，该方法可以包括以下步骤：

1、电子设备通过配置的摄像头拍摄用户的动作视频，作为目标视频。

其中，该电子设备上设置有摄像头，电子设备通过该摄像头拍摄用户的

动作视频。

该电子设备中预先设置有预设动态动作，当用户位于电子设备的摄像头的拍摄范围内时，电子设备可以通过摄像头拍摄用户的动作，后续可以确定用户的动作是否为预设动态动作。

例如，电子设备可以显示提示信息，该提示信息可以包括预设动态动作的画面，以及提示用户做出预设动态动作的信息。

2、从目标视频中提取多个目标窗口，多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧。

3、分别对每个目标窗口中的视频帧进行动作识别，得到每个目标窗口的第一动作特征信息，第一动作特征信息用于描述目标窗口包含的动态动作。

4、获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度。

5、当第一相似度大于第一预设阈值，且第一相似度与第二相似度之间的差值大于第二预设阈值时，确定第一相似度对应的第一目标窗口包含的动态动作为预设动态动作，第二相似度为与第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

其中，步骤 2-5 与上述步骤 S101-S107 类似，在此不再赘述。

6、电子设备执行与预设动态动作对应的交互操作。

当确定第一目标窗口包含的动态动作为预设动态动作时，表示用户做出了预设动态动作，此时，电子设备可以执行与该预设动态动作对应的交互操作。

在一些实施例中，电子设备执行与预设动态动作对应的交互操作，包括以下至少一项：

S601、电子设备确定用户的动作通过验证，并显示表示验证通过的预设动态效果。

例如，电子设备为体感游戏机，体感游戏机获取用户所做的动作视频，当用户的动作与预设动态动作匹配时，确定用户的动作通过验证，体感游戏机显示动作完成的预设动态效果，从而指示用户做对了预设动态动作。该预设动态效果可以为撒花效果、显示的预设动态动作逐渐消失等显示效果。

S602、电子设备为用户分配预设动态动作对应的虚拟奖励数据。

电子设备基于用户标识进行登录，用户标识用于确定唯一的用户，可以为用户账号、电话号码等。且电子设备还会为用户标识管理账户，账户中可以存储为用户分配的各种数据。

因此，当电子设备确定第一目标窗口包含的动态动作为预设动态动作时，表示用户做出了预设动态动作，此时可以为用户分配预设动态动作对应的虚拟奖励数据，也即是向用户标识的账户中添加虚拟奖励数据。

其中，不同的预设动态动作对应的虚拟奖励数据可以相同，也可以不同。该虚拟奖励数据可以包括虚拟礼物、虚拟货币、虚拟分数等。

例如，电子设备为体感游戏机，体感游戏机获取用户所做的动作视频，当确定用户的动作与预设动态动作匹配时，为用户增加奖励分数，随着用户做对的动态动作越来越多，得到的奖励分数也越来越多。

S603、电子设备为用户分配预设动态动作对应的操作权限。

电子设备基于用户标识进行登录，用户标识用于确定唯一的用户，可以为用户账号、电话号码等。且电子设备还会为用户标识设置操作权限，只有为用户标识分配了操作权限，才会允许用户执行相应的操作。

因此，当电子设备确定第一目标窗口包含的动态动作为预设动态动作时，表示用户做出了预设动态动作，此时可以为用户分配预设动态动作对应的操作权限。

例如，电子设备为门禁设备，门禁设备获取用户所做的动作视频，当确定用户的动作与预设动态动作相同时，确定用户的动作通过验证，则门禁设备开启，允许用户进入。

例如，电子设备为体感游戏机，体感游戏机获取用户所做的动作视频，当确定用户的动作与预设动态动作相同时，确定用户的动作通过验证，则游戏过关，进入下一关游戏。

本申请实施例提供的人机交互方法，能够识别用户做出的动态动作，根据用户做出的动态动作执行相应的交互操作，与用户进行交互，提供了一种新型准确的交互方式。

图 8 是本申请实施例提供的一种动作识别装置的结构示意图。如图 8 所

示，该装置包括：

提取模块 801，用于从目标视频中提取多个目标窗口，多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧。

动作识别模块 802，用于分别对每个目标窗口中的视频帧进行动作识别，得到每个目标窗口的第一动作特征信息，第一动作特征信息用于描述目标窗口包含的动态动作。

相似度获取模块 803，用于获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度。

确定模块 804，用于当第一相似度大于第一预设阈值，且第一相似度与第二相似度之间的差值大于第二预设阈值时，确定第一相似度对应的第一目标窗口包含的动态动作为预设动态动作，第二相似度为与第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

本申请实施例提供了一种识别预设动态动作的方案，能够以目标视频中的目标窗口为单位，对目标窗口中的视频帧进行动作识别，从而识别目标窗口中包含的预设动态动作，突破了相关技术中只能识别静态动作的限制，扩展了应用范围，提高了灵活性。而且通过计算相邻目标窗口的相似度之间的差值的方式确定相邻目标窗口中动作是否相互影响，避免由于相邻目标窗口之间的影响而造成识别错误的问题，提高了准确率。

在一些实施例中，如图 9 所示，提取模块 801，包括：

获取单元 8011，用于从目标视频中提取多个视频帧。

划分单元 8012，用于以第一预设数量作为每个目标窗口中的视频帧数量，以第二预设数量作为任两个相邻的目标窗口包含的相同视频帧的数量，对多个视频帧进行划分，得到多个目标窗口。

在一些实施例中，如图 9 所示，动作识别模块 802，包括：

提取单元 8021，用于对于每个目标窗口中的每个视频帧，提取视频帧中的多个身体关键点。

动作识别单元 8022，用于根据多个身体关键点的分布情况，对视频帧进行动作识别，得到视频帧的第二动作特征信息，第二动作特征信息用于描述视频帧包含的静态动作。

组合单元 8023, 用于分别将每个目标窗口中的视频帧的第二动作特征信息进行组合, 得到每个目标窗口的第一动作特征信息。

预设动态动作由至少两个预设身体部位做出; 动作识别单元 8022, 还用于执行以下至少一项:

根据多个身体关键点在视频帧中的坐标以及多个身体关键点所属的身体部位, 确定视频帧中任两个预设身体部位之间的角度, 作为第二动作特征信息。

获取多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量, 作为第二动作特征信息, 参考视频帧在视频帧之前, 且与视频帧相隔第三预设数量的视频帧。

获取任两个预设身体部位中的参考身体部位的尺寸, 以及任两个预设身体部位之间的距离, 将该距离与参考身体部位的尺寸之间的比例作为第二动作特征信息。

在一些实施例中, 第一动作特征信息为包括 M 个第一动作向量的第一动作矩阵, 预设特征信息为包括 N 个预设动作向量的预设动作矩阵, M 和 N 为正整数, 相似度获取模块 803, 包括:

创建单元 8033, 用于创建相似度矩阵, 相似度矩阵的行数为 M 、列数为 N , 或者行数为 N 、列数为 M 。

相似度获取单元 8034, 用于对于相似度矩阵中与第 i 个第一动作向量和第 j 个预设动作向量对应的指定位置, 获取第一位置、第二位置和第三位置的相似度中的最大相似度与第 i 个第一动作向量与第 j 个预设动作向量之间的相似度的和, 作为指定位置的相似度, 第一位置为与第 $i-1$ 个第一动作向量和第 j 个预设动作向量对应的位置, 第二位置为与第 $i-1$ 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置, 第三位置为与第 i 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置, i 为不小于 1 且不大于 M 的正整数, j 为不小于 1 且不大于 N 的正整数。

相似度获取单元 8034, 还用于将相似度矩阵中与第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度, 确定为第一动作特征信息与预设特征信息之间的相似度。

需要说明的是: 上述实施例提供的动作识别装置在识别动作时, 仅以上

述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将电子设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的动作识别装置的实施例与动作识别方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 10 是本申请实施例提供的一种人机交互装置的结构示意图。如图 10 所示，该装置包括：

拍摄模块 1001，用于通过配置的摄像头拍摄用户的动作视频，作为目标视频。

提取模块 1002，用于从目标视频中提取多个目标窗口，多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧。

动作识别模块 1003，用于分别对每个目标窗口中的视频帧进行动作识别，得到每个目标窗口的第一动作特征信息，第一动作特征信息用于描述目标窗口包含的动态动作。

相似度获取模块 1004，用于获取每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度。。

确定模块 1005，用于当第一相似度大于第一预设阈值，且第一相似度与第二相似度之间的差值大于第二预设阈值时，确定第一相似度对应的第一目标窗口包含的动态动作为预设动态动作，第二相似度为与第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

执行模块 1006，用于执行与预设动态动作对应的交互操作。

本申请实施例提供的人机交互装置，能够识别用户做出的动态动作，根据用户做出的动态动作执行相应的交互操作，与用户进行交互，提供了一种新型准确的交互方式。

在一些实施例中，执行模块，用于执行以下至少一项：

确定用户的动作通过验证，并显示表示验证通过的预设动态效果。

为用户分配预设动态动作对应的虚拟奖励数据。

为用户分配预设动态动作对应的操作权限。

需要说明的是：上述实施例提供的人机交互装置在进行人机交互时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将电子设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的人机交互装置的实施例与人机交互方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 11 是本申请实施例提供的一种电子设备的结构示意图。该电子设备 1100 可以是：体感游戏机、门禁设备、智能手机、平板电脑、MP3 播放器（Moving Picture Experts Group Audio Layer III，动态影像专家压缩标准音频层面 3）、MP4（Moving Picture Experts Group Audio Layer IV，动态影像专家压缩标准音频层面 4）播放器、笔记本电脑、台式电脑、头戴式设备，或其他任意智能电子设备。电子设备 1100 还可能被称为用户设备、便携式电子设备、膝上型电子设备、台式电子设备等其他名称。

通常，电子设备 1100 包括有：处理器 1101 和存储器 1102。

处理器 1101 可以包括一个或多个处理核心，比如 4 核心处理器、8 核心处理器等。处理器 1101 可以采用 DSP（Digital Signal Processing，数字信号处理）、FPGA（Field - Programmable Gate Array，现场可编程门阵列）、PLA（Programmable Logic Array，可编程逻辑阵列）中的至少一种硬件形式来实现。处理器 1101 也可以包括主处理器和协处理器，主处理器是用于对在唤醒状态下的数据进行处理的处理器，也称 CPU（Central Processing Unit，中央处理器）；协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中，处理器 1101 可以在集成有 GPU（Graphics Processing Unit，图像处理器），GPU 用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中，处理器 1101 还可以包括 AI（Artificial Intelligence，人工智能）处理器，该 AI 处理器用于处理有关机器学习的计算操作。

存储器 1102 可以包括一个或多个计算机可读存储介质，该计算机可读存储介质可以是非暂态的。存储器 1102 还可包括高速随机存取存储器，以及非易失性存储器，比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中，存储器 1102 中的非暂态的计算机可读存储介质用于存储至少一个指

令,该至少一个指令用于被处理器 1101 所具有以实现本申请中方法实施例提供的动作识别方法。

在一些实施例中,电子设备 1100 还可选包括有:外围设备接口 1103 和至少一个外围设备。处理器 1101、存储器 1102 和外围设备接口 1103 之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口 1103 相连。具体地,外围设备包括:射频电路 1104、触摸显示屏 1105、摄像头 1106、音频电路 1107、定位组件 1108 和电源 1109 中的至少一种。

外围设备接口 1103 可被用于将 I/O (Input /Output, 输入/输出) 相关的至少一个外围设备连接到处理器 1101 和存储器 1102。在一些实施例中,处理器 1101、存储器 1102 和外围设备接口 1103 被集成在同一芯片或电路板上;在一些其他实施例中,处理器 1101、存储器 1102 和外围设备接口 1103 中的任意一个或两个可以在单独的芯片或电路板上实现,本实施例对此不加以限定。

射频电路 1104 用于接收和发射 RF (Radio Frequency, 射频) 信号,也称电磁信号。射频电路 1104 通过电磁信号与通信网络以及其他通信设备进行通信。射频电路 1104 将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频电路 1104 包括:天线系统、RF 收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路 1104 可以通过至少一种无线通信协议来与其它电子设备进行通信。该无线通信协议包括但不限于:城域网、各代移动通信网络(2G、3G、4G 及 8G)、无线局域网和/或 WiFi(Wireless Fidelity, 无线保真)网络。在一些实施例中,射频电路 1104 还可以包括 NFC (Near Field Communication, 近距离无线通信) 有关的电路,本申请对此不加以限定。

显示屏 1105 用于显示 UI (UserInterface, 用户界面)。该 UI 可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏 1105 是触摸显示屏时,显示屏 1105 还具有采集在显示屏 1105 的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器 1101 进行处理。此时,显示屏 1105 还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏 1105 可以为一个,设置电子设备 1100 的前面板;在

另一些实施例中，显示屏 1105 可以为至少两个，分别设置在电子设备 1100 的不同表面或呈折叠设计；在再一些实施例中，显示屏 1105 可以是柔性显示屏，设置在电子设备 1100 的弯曲表面上或折叠面上。甚至，显示屏 1105 还可以设置成非矩形的不规则图形，也即异形屏。显示屏 1105 可以采用 LCD(Liquid Crystal Display, 液晶显示屏)、OLED(Organic Light-Emitting Diode, 有机发光二极管)等材质制备。

摄像头组件 1106 用于采集图像或视频。可选地，摄像头组件 1106 包括前置摄像头和后置摄像头。通常，前置摄像头设置在电子设备的前面板，后置摄像头设置在电子设备的背面。在一些实施例中，后置摄像头为至少两个，分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种，以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及 VR (Virtual Reality, 虚拟现实) 拍摄功能或者其它融合拍摄功能。在一些实施例中，摄像头组件 1106 还可以包括闪光灯。闪光灯可以是单色温闪光灯，也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合，可以用于不同色温下的光线补偿。

音频电路 1107 可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波，并将声波转换为电信号输入至处理器 1101 进行处理，或者输入至射频电路 1104 以实现语音通信。出于立体声采集或降噪的目的，麦克风可以为多个，分别设置在电子设备 1100 的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器 1101 或射频电路 1104 的电信号转换为声波。扬声器可以是传统的薄膜扬声器，也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时，不仅可以将电信号转换为人类可听见的声波，也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中，音频电路 1107 还可以包括耳机插孔。

定位组件 1108 用于定位电子设备 1100 的当前地理位置，以实现导航或 LBS (Location Based Service, 基于位置的服务)。定位组件 1108 可以是基于美国的 GPS (Global Positioning System, 全球定位系统)、中国的北斗系统、俄罗斯的格雷纳斯系统或欧盟的伽利略系统的定位组件。

电源 1109 用于为电子设备 1100 中的各个组件进行供电。电源 1109 可以是交流电、直流电、一次性电池或可充电电池。当电源 1109 包括可充电电池

时, 该可充电电池可以支持有线充电或无线充电。该可充电电池还可以用于支持快充技术。

在一些实施例中, 电子设备 1100 还包括有一个或多个传感器 1110。该一个或多个传感器 1110 包括但不限于: 加速度传感器 1111、陀螺仪传感器 1112、压力传感器 1113、指纹传感器 1114、光学传感器 1115 以及接近传感器 1116。

加速度传感器 1111 可以检测以电子设备 1100 建立的坐标系的三个坐标轴上的加速度大小。比如, 加速度传感器 1111 可以用于检测重力加速度在三个坐标轴上的分量。处理器 1101 可以根据加速度传感器 1111 采集的重力加速度信号, 控制触摸显示屏 1105 以横向视图或纵向视图进行用户界面的显示。加速度传感器 1111 还可以用于游戏或者用户的运动数据的采集。

陀螺仪传感器 1112 可以检测电子设备 1100 的机体方向及转动角度, 陀螺仪传感器 1112 可以与加速度传感器 1111 协同采集用户对电子设备 1100 的 3D 动作。处理器 1101 根据陀螺仪传感器 1112 采集的数据, 可以实现如下功能: 动作感应 (比如根据用户的倾斜操作来改变 UI)、拍摄时的图像稳定、游戏控制以及惯性导航。

压力传感器 1113 可以设置在电子设备 1100 的侧边框和/或触摸显示屏 1105 的下层。当压力传感器 1113 设置在电子设备 1100 的侧边框时, 可以检测用户对电子设备 1100 的握持信号, 由处理器 1101 根据压力传感器 1113 采集的握持信号进行左右手识别或快捷操作。当压力传感器 1113 设置在触摸显示屏 1105 的下层时, 由处理器 1101 根据用户对触摸显示屏 1105 的压力操作, 实现对 UI 界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

指纹传感器 1114 用于采集用户的指纹, 由处理器 1101 根据指纹传感器 1114 采集到的指纹识别用户的身份, 或者, 由指纹传感器 1114 根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时, 由处理器 1101 授权该用户具有相关的敏感操作, 该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器 1114 可以被设置电子设备 1100 的正面、背面或侧面。当电子设备 1100 上设置有物理按键或厂商 Logo 时, 指纹传感器 1114 可以与物理按键或厂商标志集成在一起。

光学传感器 1115 用于采集环境光强度。在一个实施例中, 处理器 1101

可以根据光学传感器 1115 采集的环境光强度，控制触摸显示屏 1105 的显示亮度。具体地，当环境光强度较高时，调高触摸显示屏 1105 的显示亮度；当环境光强度较低时，调低触摸显示屏 1105 的显示亮度。在另一个实施例中，处理器 1101 还可以根据光学传感器 1115 采集的环境光强度，动态调整摄像头组件 1106 的拍摄参数。

接近传感器 1116，也称距离传感器，通常设置在电子设备 1100 的前面板。接近传感器 1116 用于采集用户与电子设备 1100 的正面之间的距离。在一个实施例中，当接近传感器 1116 检测到用户与电子设备 1100 的正面之间的距离逐渐变小时，由处理器 1101 控制触摸显示屏 1105 从亮屏状态切换为息屏状态；当接近传感器 1116 检测到用户与电子设备 1100 的正面之间的距离逐渐变大时，由处理器 1101 控制触摸显示屏 1105 从息屏状态切换为亮屏状态。

本领域技术人员可以理解，图 11 中示出的结构并不构成对电子设备 1100 的限定，可以包括比图示更多或更少的组件，或者组合某些组件，或者采用不同的组件布置。

图 12 是本申请实施例提供的一种服务器的结构示意图，该服务器 1200 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上处理器（central processing units, CPU）1201 和一个或一个以上的存储器 1202，其中，所述存储器 1202 中存储有至少一条指令，所述至少一条指令由所述处理器 1201 加载并执行以实现上述各个方法实施例提供的方法。当然，该服务器还可以具有有线或无线网络接口、键盘以及输入输出接口等部件，以便进行输入输出，该服务器还可以包括其他用于实现设备功能的部件，在此不做赘述。

服务器 1200 可以用于执行上述动作识别方法中或上述人机交互方法中电子设备所执行的步骤。

本申请实施例还提供了一种电子设备，电子设备包括处理器和存储器，存储器中存储有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行如上述的动作识别方法的步骤。

本申请实施例还提供了一种电子设备，电子设备包括处理器和存储器，存储器中存储有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行如上述的人机交互方法的步骤。

本申请实施例还提供了一种非易失性的计算机可读存储介质，存储有计算机可读指令，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行如上述的动作识别方法的步骤。

本申请实施例还提供了一种非易失性的计算机可读存储介质，存储有计算机可读指令，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行如上述的人机交互方法的步骤。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

以上所述仅为本申请的较佳实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权利要求书

1、一种动作识别方法，由电子设备执行，所述方法包括：

从目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，所述预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度；及

当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作，所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

2、根据权利要求1所述的方法，其特征在于，所述从目标视频中提取多个目标窗口，包括：

从目标视频中提取多个视频帧；及

以第一预设数量作为每个目标窗口中的视频帧数量，以第二预设数量作为任两个相邻的目标窗口所包括的相同视频帧的数量，对所述多个视频帧进行划分，得到多个目标窗口。

3、根据权利要求1所述的方法，其特征在于，所述分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，包括：

对于所述每个目标窗口中的每个视频帧，提取所述视频帧中的多个身体关键点；

根据所述多个身体关键点的分布情况，对所述视频帧进行动作识别，得到所述视频帧的第二动作特征信息，所述第二动作特征信息用于描述所述视频帧包含的静态动作；及

分别将所述每个目标窗口中的视频帧的第二动作特征信息进行组合，得到所述每个目标窗口的第一动作特征信息。

4、根据权利要求3所述的方法，其特征在于，所述预设动态动作由至少两个预设身体部位做出；所述根据所述多个身体关键点的分布情况，对所述视频帧进行动作识别，得到所述视频帧的第二动作特征信息，包括以下至少一项：

根据所述多个身体关键点在所述视频帧中的坐标以及所述多个身体关键点所属的身体部位，确定所述视频帧中任两个预设身体部位之间的角度，作为所述第二动作特征信息；

获取所述多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量，作为所述第二动作特征信息，所述参考视频帧在所述视频帧之前，且与所述视频帧相隔第三预设数量的视频帧；及

获取任两个预设身体部位中的参考身体部位的尺寸，以及所述任两个预设身体部位之间的距离，将所述距离与所述参考身体部位的尺寸之间的比例作为第二动作特征信息。

5、根据权利要求1所述的方法，其特征在于，所述第一动作特征信息为包括M个第一动作向量的第一动作矩阵，所述预设特征信息为包括N个预设动作向量的预设动作矩阵，M和N为正整数，所述获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，包括：

创建相似度矩阵，所述相似度矩阵的行数为M、列数为N，或者行数为N、列数为M；

对于所述相似度矩阵中与第i个第一动作向量和第j个预设动作向量对应的指定位置，获取第一位置、第二位置和第三位置的相似度中的最大相似度与所述第i个第一动作向量与所述第j个预设动作向量之间的相似度的和，作为所述指定位置的相似度，所述第一位置为与第i-1个第一动作向量和第j个预设动作向量对应的位置，所述第二位置为与第i-1个第一动作向量和第j-1个预设动作向量对应的位置，所述第三位置为与第i个第一动作向量和第j-1个预设动作向量对应的位置，i为不小于1且不大于M的正整数，j为不小于1且不大于N的正整数；及

将所述相似度矩阵中与第M个第一动作向量和第N个预设动作向量对应的位置的相似度，确定为所述第一动作特征信息与所述预设特征信息之间的相似度。

6、一种人机交互方法，由电子设备执行，所述方法包括：

通过配置的摄像头拍摄用户的动作视频，作为目标视频；

从所述目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，所述预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度；

当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应的第一目标窗口包含的动态动作作为所述预设动态动作，所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度；及

执行与所述预设动态动作对应的交互操作。

7、根据权利要求6所述的方法，其特征在于，所述执行与所述预设动态动作对应的交互操作，包括以下至少一项：

确定所述用户的动作通过验证，并显示表示验证通过的预设动态效果；

为所述用户分配所述预设动态动作对应的虚拟奖励数据；及

为所述用户分配所述预设动态动作对应的操作权限。

8、一种动作识别装置，所述装置包括：

提取模块，用于从目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

动作识别模块，用于分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

相似度获取模块，用于获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，所述预设特征信息用于描述预设动态动作；其中，

获取到的多个相似度中最大的相似度为第一相似度；及

确定模块，用于当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应的第一目标窗口包含的动态动作为所述预设动态动作，所述第二相似度为与所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度。

9、根据权利要求8所述的装置，其特征在于，所述提取模块，包括：

获取单元，用于从目标视频中提取多个视频帧；及

划分单元，用于以第一预设数量作为每个目标窗口中的视频帧数量，以第二预设数量作为任两个相邻的目标窗口所包括的相同视频帧的数量，对所述多个视频帧进行划分，得到多个目标窗口。

10、根据权利要求8所述的装置，其特征在于，所述动作识别模块，包括：

提取单元，用于对于所述每个目标窗口中的每个视频帧，提取所述视频帧中的多个身体关键点；

动作识别单元，用于根据所述多个身体关键点的分布情况，对所述视频帧进行动作识别，得到所述视频帧的第二动作特征信息，所述第二动作特征信息用于描述所述视频帧包含的静态动作；及

组合单元，用于分别将所述每个目标窗口中的视频帧的第二动作特征信息进行组合，得到所述每个目标窗口的第一动作特征信息。

11、根据权利要求10所述的装置，其特征在于，所述预设动态动作由至少两个预设身体部位做出；所述动作识别单元还用于执行以下步骤中的至少一项：

根据所述多个身体关键点在所述视频帧中的坐标以及所述多个身体关键点所属的身体部位，确定所述视频帧中任两个预设身体部位之间的角度，作为所述第二动作特征信息；

获取所述多个身体关键点中至少一个身体关键点与参考视频帧中对应的身体关键点之间的位移量，作为所述第二动作特征信息，所述参考视频帧在所述视频帧之前，且与所述视频帧相隔第三预设数量的视频帧；及

获取任两个预设身体部位中的参考身体部位的尺寸，以及所述任两个预设身体部位之间的距离，将所述距离与所述参考身体部位的尺寸之间的比例

作为第二动作特征信息。

12、根据权利要求 8 所述的装置，其特征在于，所述第一动作特征信息为包括 M 个第一动作向量的第一动作矩阵，所述预设特征信息为包括 N 个预设动作向量的预设动作矩阵， M 和 N 为正整数，所述相似度获取模块，包括：

创建单元，用于创建相似度矩阵，所述相似度矩阵的行数为 M 、列数为 N ，或者行数为 N 、列数为 M ；

相似度获取单元，用于对于所述相似度矩阵中与第 i 个第一动作向量和第 j 个预设动作向量对应的指定位置，获取第一位置、第二位置和第三位置的相似度中的最大相似度与所述第 i 个第一动作向量与所述第 j 个预设动作向量之间的相似度的和，作为所述指定位置的相似度，所述第一位置为与第 $i-1$ 个第一动作向量和第 j 个预设动作向量对应的位置，所述第二位置为与第 $i-1$ 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置，所述第三位置为与第 i 个第一动作向量和第 $j-1$ 个预设动作向量对应的位置， i 为不小于 1 且不大于 M 的正整数， j 为不小于 1 且不大于 N 的正整数；及

所述相似度获取单元，还用于将所述相似度矩阵中与第 M 个第一动作向量和第 N 个预设动作向量对应的位置的相似度，确定为所述第一动作特征信息与所述预设特征信息之间的相似度。

13、一种人机交互装置，所述装置包括：

拍摄模块，用于通过配置的摄像头拍摄用户的动作视频，作为目标视频；

提取模块，用于从所述目标视频中提取多个目标窗口，所述多个目标窗口中的每个目标窗口均包括第一预设数量的连续的视频帧；

动作识别模块，用于分别对所述每个目标窗口中的视频帧进行动作识别，得到所述每个目标窗口的第一动作特征信息，所述第一动作特征信息用于描述所述目标窗口包含的动态动作；

相似度获取模块，用于获取所述每个目标窗口的第一动作特征信息与预设特征信息之间的相似度，所述预设特征信息用于描述预设动态动作；其中，获取到的多个相似度中最大的相似度为第一相似度；

确定模块，用于当所述第一相似度大于第一预设阈值，且所述第一相似度与第二相似度之间的差值大于第二预设阈值时，确定所述第一相似度对应

的第一目标窗口包含的动态动作为所述预设动态动作，所述第二相似度为与
所述第一目标窗口相邻的目标窗口的第一动作特征信息所对应的相似度；及
执行模块，用于执行与所述预设动态动作对应的交互操作。

14、根据权利要求 13 所述的装置，其特征在于，所述执行模块还用于执行以下步骤中的至少一项：

确定所述用户的动作通过验证，并显示表示验证通过的预设动态效果；
为所述用户分配所述预设动态动作对应的虚拟奖励数据；及
为所述用户分配所述预设动态动作对应的操作权限。

15、一种电子设备，所述电子设备包括处理器和存储器，所述存储器中
存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所
述处理器执行如权利要求 1 至 5 任一项所述的动作识别方法的步骤。

16、一种电子设备，所述电子设备包括处理器和存储器，所述存储器中
存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所
述处理器执行如权利要求 6 至 7 任一项所述的人机交互方法的步骤。

17、一种非易失性的计算机可读存储介质，存储有计算机可读指令，所
述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器
执行如权利要求 1 至 5 任一项所述的动作识别方法的步骤。

18、一种非易失性的计算机可读存储介质，存储有计算机可读指令，所
述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器
执行如权利要求 6 至 7 任一项所述的人机交互方法的步骤。

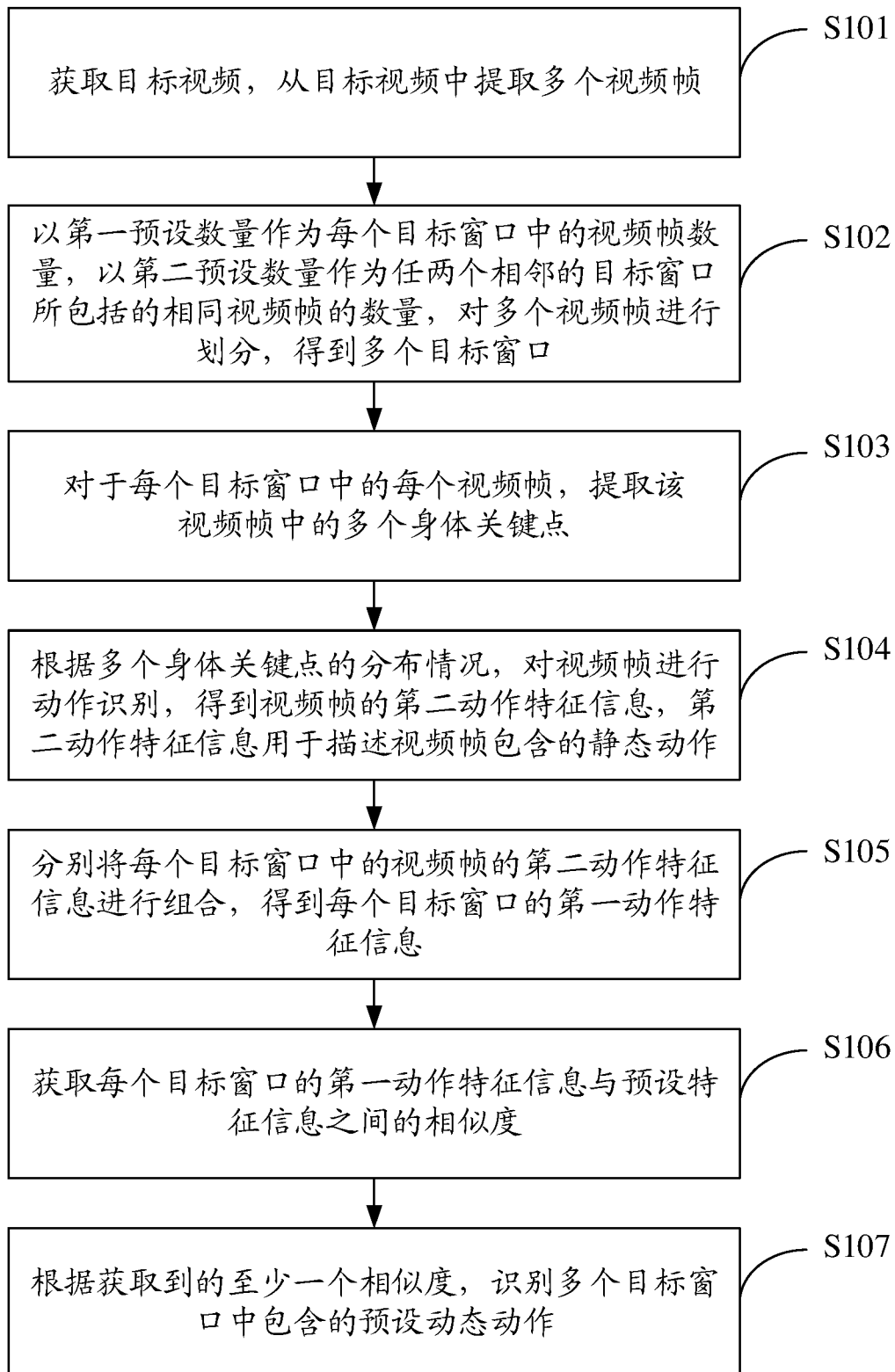


图 1

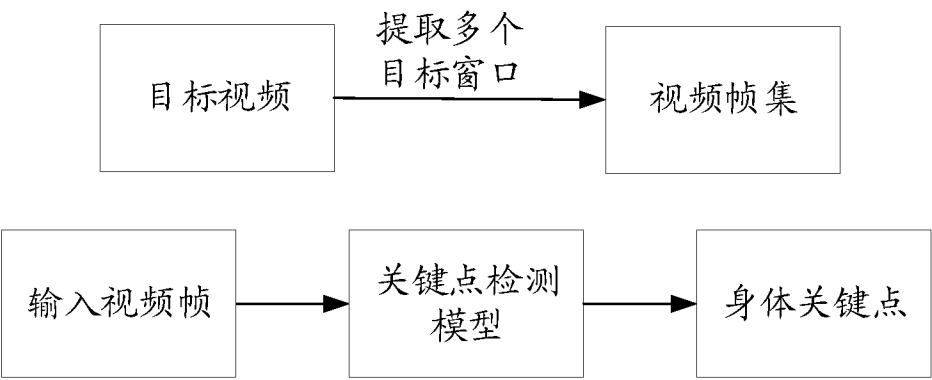


图 2

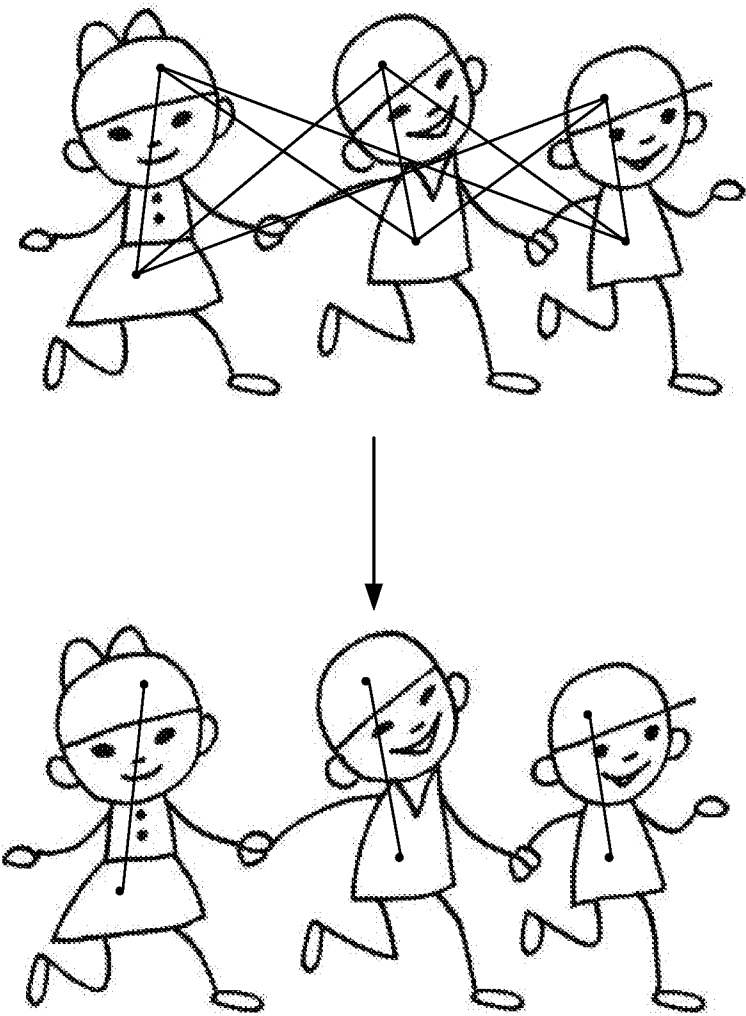


图 3

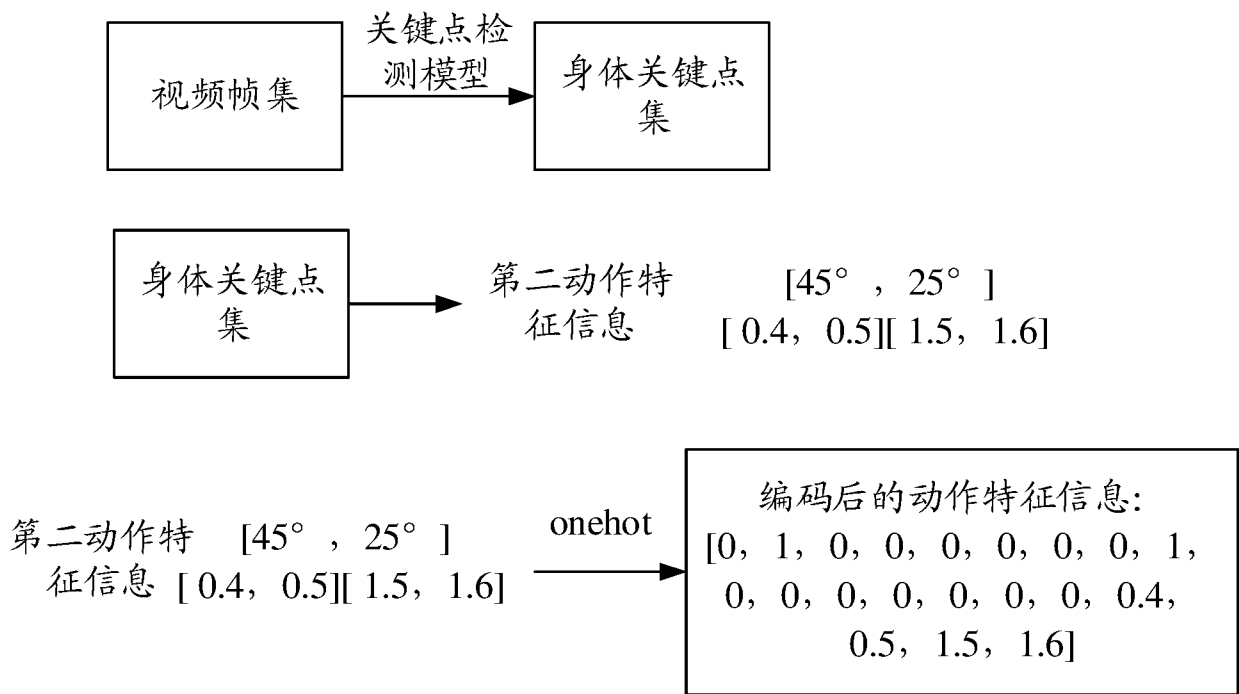


图 4

13	14	15	16
9	10	11	12
5	6	7	8
1	2	3	4

图 5

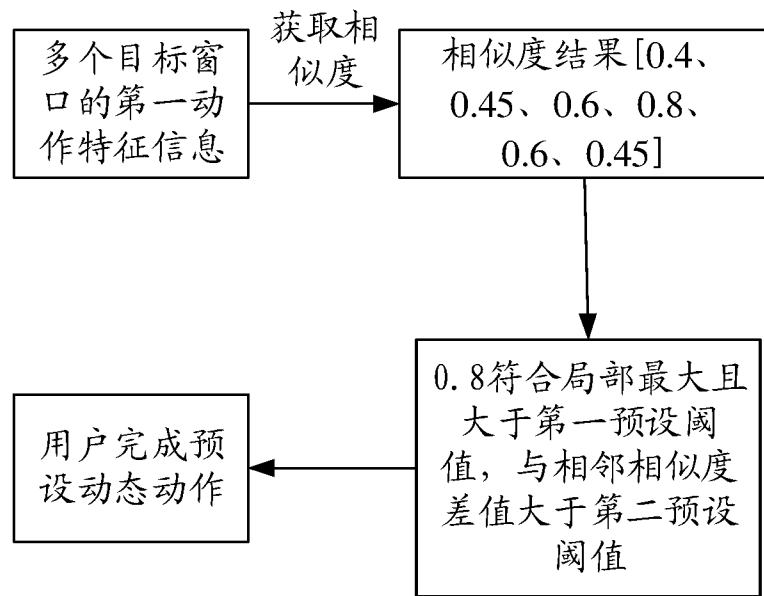


图 6

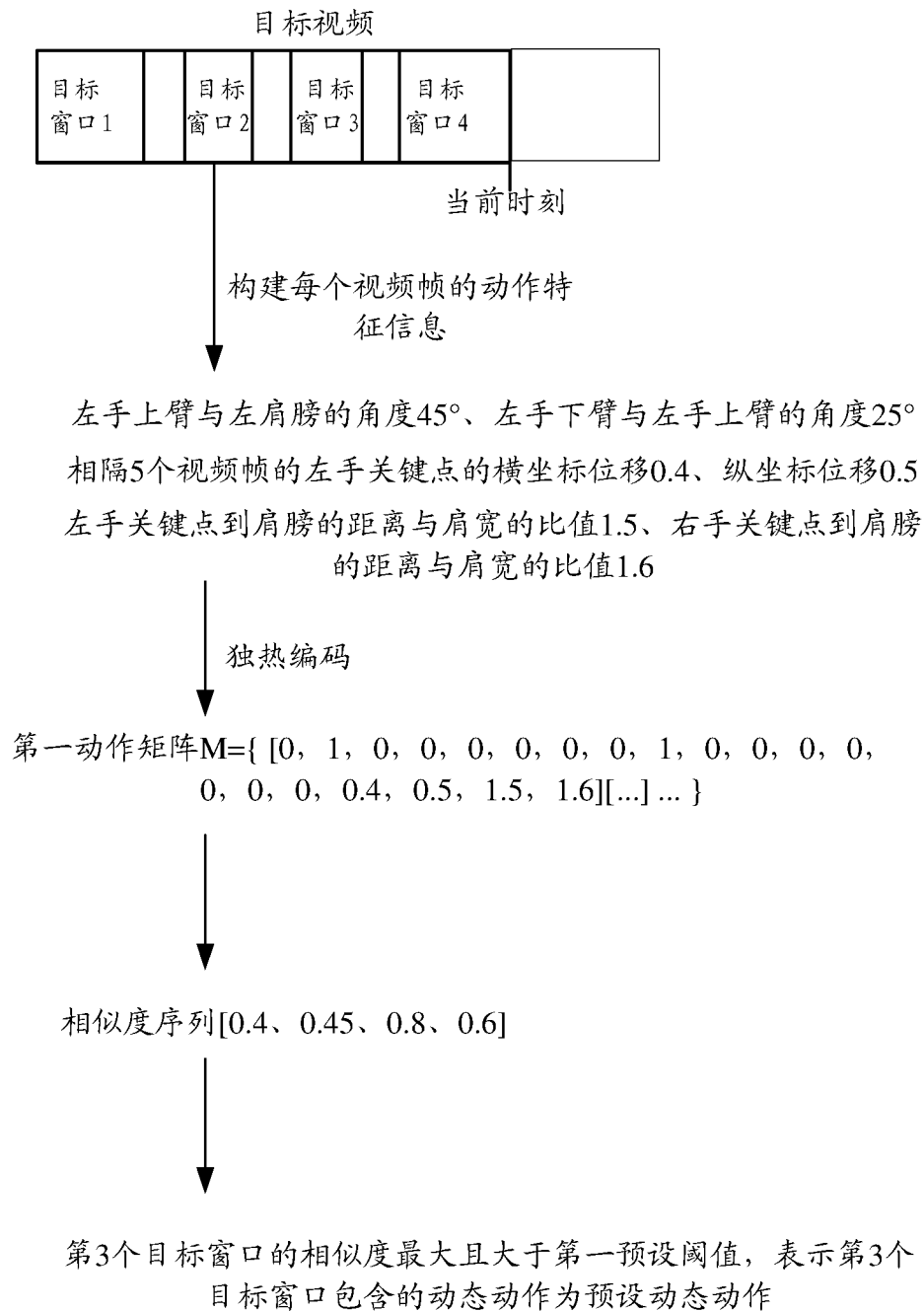


图 7

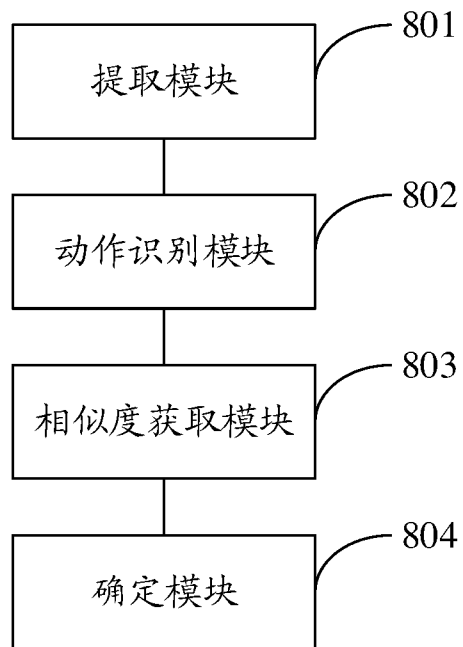


图 8

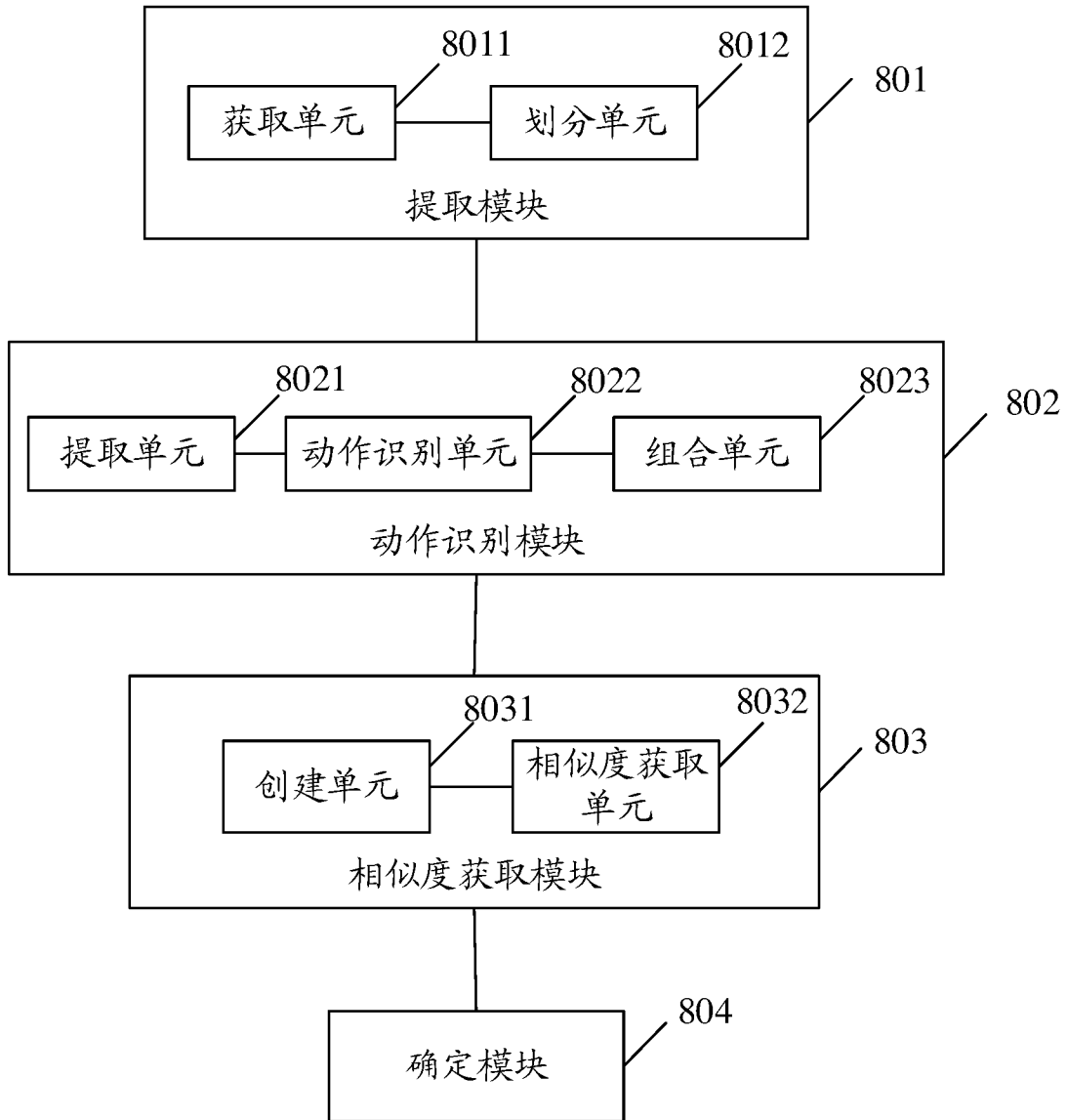


图 9

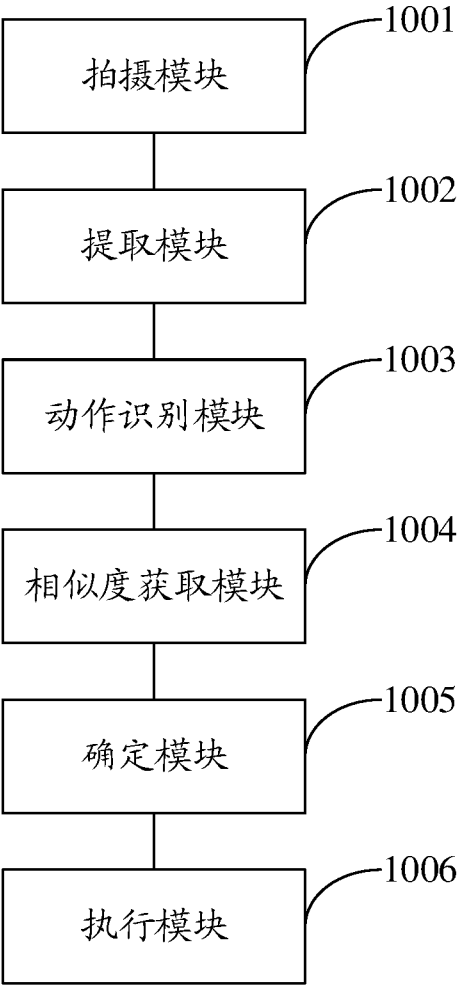


图 10

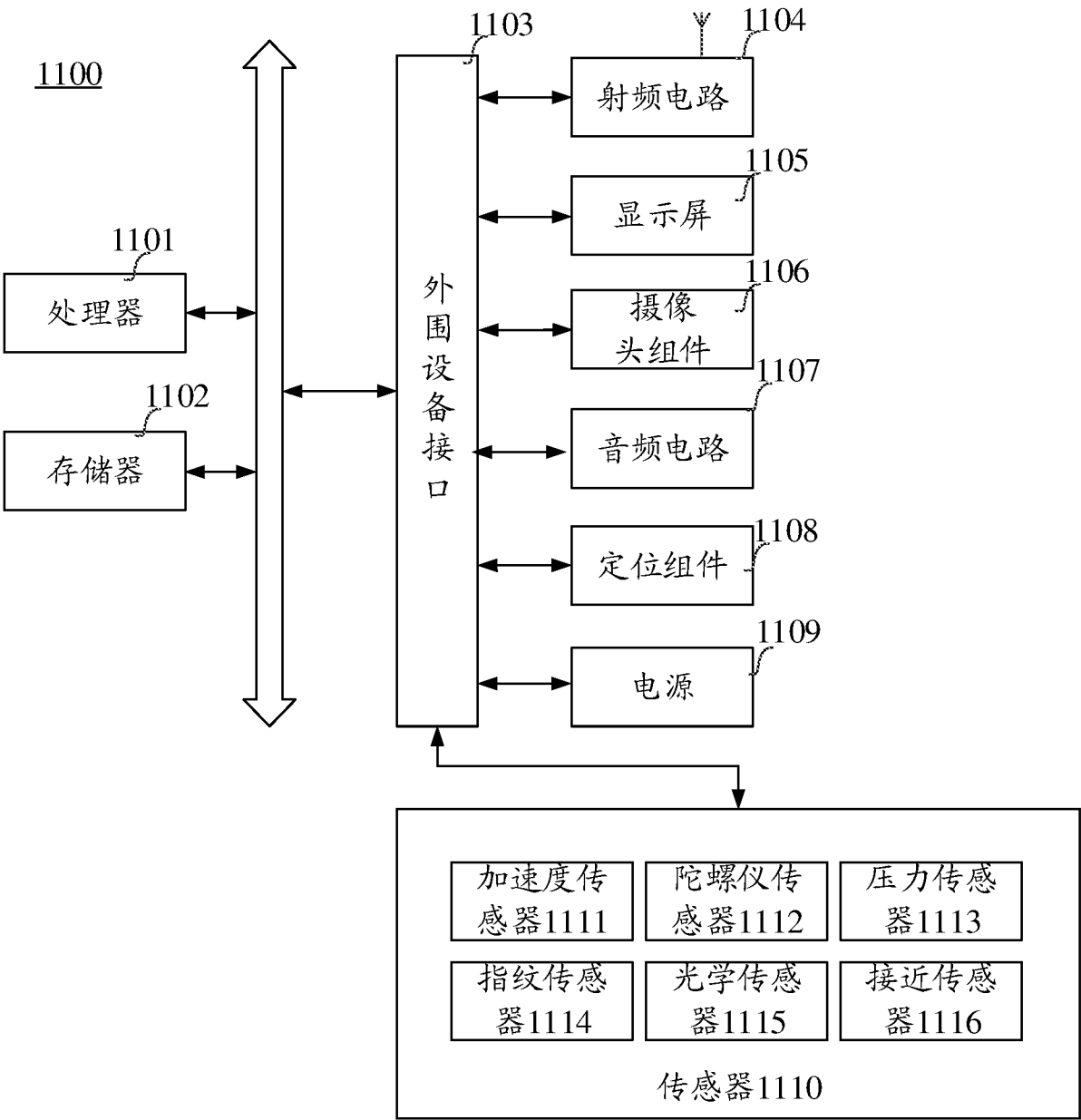


图 11

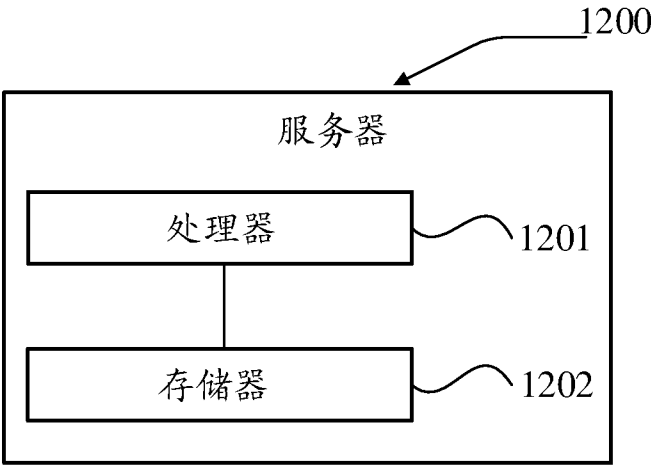


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/084996

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

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)

CNTEXT; CNABS; DWPI; CNKI; SIPOABS: 动作, 运动, 手势, 姿势, 姿态, 步态, 识别, 检测, 匹配, 辨识, 静态, 动态, 交互, 窗口, 划窗, 连续, 帧, 相似度, 阈值, action, motion, hand, gesture, posture, walking, recognition, detection, match, distinguish, static, dynamic, alternation, window, divide, continuous, frame, similarity, threshold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 108304757 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 20 July 2018 (2018-07-20) description, paragraphs [0034]-[0120]	1-4, 6-11, 13-18
Y	CN 101400001 A (TSINGHUA UNIVERSITY) 01 April 2009 (2009-04-01) description page 6, line 8 to page 13, line 1	1-4, 6-11, 13-18
Y	CN 107786848 A (ZHOU, Yanhong) 09 March 2018 (2018-03-09) claims 1-10	1, 2, 6-9, 13-18
PX	CN 110059661 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 26 July 2019 (2019-07-26) entire document	1-18
A	CN 109035299 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 18 December 2018 (2018-12-18) entire document	1-18
A	US 8027542 B2 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 27 September 2011 (2011-09-27) entire document	1-18



Further documents are listed in the continuation of Box C.



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

10 June 2020

Date of mailing of the international search report

01 July 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/084996

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108304757	A	20 July 2018	WO	2018228218	A1	20 December 2018
CN	101400001	A	01 April 2009	CN	101400001	B	02 June 2010
CN	107786848	A	09 March 2018	None			
CN	110059661	A	26 July 2019	None			
CN	109035299	A	18 December 2018	WO	2019237516	A1	19 December 2019
US	8027542	B2	27 September 2011	US	2008310734	A1	18 December 2008

A. 主题的分类 G06K 9/00 (2006.01) i 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																							
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 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)) CNTXT; CNABS; DWPI; CNKI; SIPOABS: 动作, 运动, 手势, 姿势, 姿态, 步态, 识别, 检测, 匹配, 辨识, 静态, 动态, 交互, 窗口, 划窗, 连续, 帧, 相似度, 阈值, action, motion, hand, gesture, posture, walking, recognition, detection, match, distinguish, static, dynamic, alternation, window, divide, continuous, frame, similarity, threshold																							
C. 相关文件 <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>Y</td> <td>CN 108304757 A (腾讯科技深圳有限公司) 2018年 7月 20日 (2018 - 07 - 20) 说明书第[0034]-[0120]段</td> <td>1-4, 6-11, 13-18</td> </tr> <tr> <td>Y</td> <td>CN 101400001 A (清华大学) 2009年 4月 1日 (2009 - 04 - 01) 说明书第6页第8行-第13页第1行</td> <td>1-4, 6-11, 13-18</td> </tr> <tr> <td>Y</td> <td>CN 107786848 A (周燕红) 2018年 3月 9日 (2018 - 03 - 09) 权利要求1-10</td> <td>1, 2, 6-9, 13-18</td> </tr> <tr> <td>PX</td> <td>CN 110059661 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 109035299 A (平安科技深圳有限公司) 2018年 12月 18日 (2018 - 12 - 18) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>US 8027542 B2 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 2011年 9月 27日 (2011 - 09 - 27) 全文</td> <td>1-18</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	Y	CN 108304757 A (腾讯科技深圳有限公司) 2018年 7月 20日 (2018 - 07 - 20) 说明书第[0034]-[0120]段	1-4, 6-11, 13-18	Y	CN 101400001 A (清华大学) 2009年 4月 1日 (2009 - 04 - 01) 说明书第6页第8行-第13页第1行	1-4, 6-11, 13-18	Y	CN 107786848 A (周燕红) 2018年 3月 9日 (2018 - 03 - 09) 权利要求1-10	1, 2, 6-9, 13-18	PX	CN 110059661 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文	1-18	A	CN 109035299 A (平安科技深圳有限公司) 2018年 12月 18日 (2018 - 12 - 18) 全文	1-18	A	US 8027542 B2 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 2011年 9月 27日 (2011 - 09 - 27) 全文	1-18
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
Y	CN 108304757 A (腾讯科技深圳有限公司) 2018年 7月 20日 (2018 - 07 - 20) 说明书第[0034]-[0120]段	1-4, 6-11, 13-18																					
Y	CN 101400001 A (清华大学) 2009年 4月 1日 (2009 - 04 - 01) 说明书第6页第8行-第13页第1行	1-4, 6-11, 13-18																					
Y	CN 107786848 A (周燕红) 2018年 3月 9日 (2018 - 03 - 09) 权利要求1-10	1, 2, 6-9, 13-18																					
PX	CN 110059661 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文	1-18																					
A	CN 109035299 A (平安科技深圳有限公司) 2018年 12月 18日 (2018 - 12 - 18) 全文	1-18																					
A	US 8027542 B2 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 2011年 9月 27日 (2011 - 09 - 27) 全文	1-18																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “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 </td> <td> “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 “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 “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 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “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	“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 “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 “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 “&” 同族专利的文件																			
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “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	“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 “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 “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 “&” 同族专利的文件																						
国际检索实际完成的日期 2020年 6月 10日		国际检索报告邮寄日期 2020年 7月 1日																					
ISA/CN的名称和邮寄地址 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 传真号 (86-10)62019451		授权官员 钟阳雪 电话号码 86-010-62411634																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/084996

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	108304757	A	2018年 7月 20日	WO	2018228218	A1	2018年 12月 20日
CN	101400001	A	2009年 4月 1日	CN	101400001	B	2010年 6月 2日
CN	107786848	A	2018年 3月 9日	无			
CN	110059661	A	2019年 7月 26日	无			
CN	109035299	A	2018年 12月 18日	WO	2019237516	A1	2019年 12月 19日
US	8027542	B2	2011年 9月 27日	US	2008310734	A1	2008年 12月 18日