

(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 8 日 (08.10.2020)



(10) 国际公布号
WO 2020/199480 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01) **G06K 9/32** (2006.01)

(21) 国际申请号: PCT/CN2019/103164

(22) 国际申请日: 2019 年 8 月 29 日 (29.08.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910264909.0 2019 年 4 月 3 日 (03.04.2019) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 叶明 (YE, Ming); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳中一联合知识产权代理有限公司 (SHENZHEN ZHONGYI UNION INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国广东省深圳市福田区园岭街道深南中路 1014 号报春大厦 9 楼 (5 号信箱), Guangdong 518028 (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, 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,

(54) Title: BODY MOVEMENT RECOGNITION METHOD AND DEVICE

(54) 发明名称: 一种人体动作的识别方法及设备

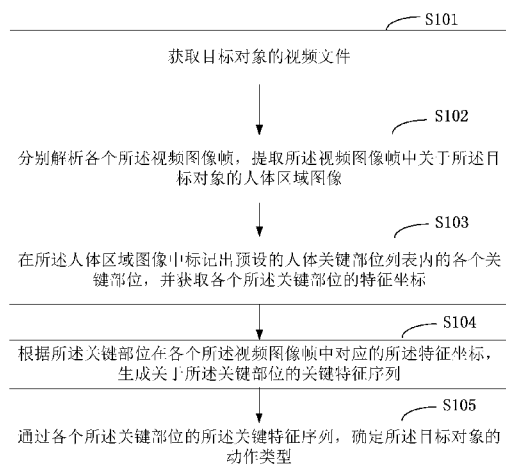


图 1

S101 Obtain a video file of a target object
S102 Parse each video image frame respectively, and extract a body area image about the target object in the video image frame
S103 Mark each key part in a preset body key part list in the body area image, and obtain feature coordinates of each key part
S104 Generate a key feature sequence about the key part according to the corresponding feature coordinates of the key part in each video image frame
S105 Determine the movement type of the target object by means of the key feature sequence of each key part

(57) Abstract: A body movement recognition method and device, applicable to the technical field of image recognition. The method comprises: obtaining a video file of a target object (S101); parsing each video image frame respectively, and extracting a body area image about the target object in the video image frame (S102); marking each key part in a preset body key part list in the body area image, and obtaining feature coordinates of each key part (S103); generating a key feature sequence about the key part according to the corresponding feature coordinates of the key part in each video image frame (S104); and determining the movement type of the target object by means of the key feature sequence of each key part (S105). Determining the movement of a target object according to the changes of multiple key parts further improves the accuracy, thereby improving the image recognition effect and the object behavior analysis efficiency.

(57) 摘要: 一种人体动作的识别方法及设备, 适用于图像识别技术领域, 该方法包括: 获取目标对象的视频文件 (S101); 分别解析各个所述视频图像帧, 提取所述视频图像帧中关于所述目标对象的人体区域图像 (S102); 在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位, 并获取各个所述关键部位的特征坐标 (S103); 根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标, 生成关于所述关键部位的关键特征序列 (S104); 通过各个所述关键部位的所述关键特征序列, 确定所述目标对象的动作类型 (S105)。通过多个关键部位的变动情况确定目标对象的动作, 准确率也进一步提高, 从而提高了图像识别的效果以及对象行为分析的效率。

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,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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))。

一种人体动作的识别方法及设备

本申请申明享有 2019 年 04 月 03 日递交的申请号为 201910264909.0、名称为“一种人体动作的识别方法及设备”中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

5 技术领域

本申请属于图像识别技术领域，尤其涉及一种人体动作的识别方法及设备。

背景技术

随着图像识别技术的不断发展，计算机可以从图像文件以及视频文件中自动识别得到越来越多的信息，例如确定画面中包含的用户的人体动作类型，并基于识别得到动作信息进行对象追踪以及对对象行为分析等操作，因此图像识别技术的准确度以及识别速率，则会直接影响后续步骤的处理效果。现有的人体动作的识别技术，一般是采用卷积神经网络进行识别，然而上述技术需要借助光流信息，需要多次进行时序递归操作，从而识别速度较低，而且准确率也不高，从而降低了图像识别的效果以及后续基于人体动作进行对象行为分析的效率。

15 技术问题

有鉴于此，本申请实施例提供了一种人体动作的识别方法及设备，以解决现有的人体动作的识别方法，识别速度较低，而且准确率也不高，从而降低了图像处理的识别效果以及后续基于人体动作进行对象行为分析的效率的问题。

技术方案

20 本申请实施例的第一方面提供了一种人体动作的识别方法，包括：
获取目标对象的视频文件；所述视频文件包括多个视频图像帧；
分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像；

在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各
25 个所述关键部位的特征坐标；

根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列；

通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

有益效果

本申请实施例通过获取所需要进行动作行为分析的目标用户的视频文件，并对该视频文件的各个视频图像帧进行解析，确定每个视频图像帧中包含的人体区域图像，在人体区域图像中标记出各个关键部位，并根据各个关键部位的特征坐标，确定目标对象的各个部位的变化情况，从而确定目标对象的动作类型，自动识别目标对象的人体动作。与现有的人体动作的识别技术相比，本申请实施例无需依赖神经网络对视频图像进行动作类型的识别，并不借助光流信息，避免了需要进行时序递归而带来的识别时延，从而提高了识别的效率，而且通过定位多个关键部位，通过多个关键部位的变动情况确定目标对象的动作，准确率也进一步提高，从而提高了图像识别的效果以及对象行为分析的效率。

附图说明

图 1 是本申请第一实施例提供的一种人体动作的识别方法的实现流程图；
图 2 是本申请第二实施例提供的一种人体动作的识别方法 S102 具体实现流程图；
图 3 是本申请第三实施例提供的一种人体动作的识别方法 S104 具体实现流程图；
图 4 是本申请第四实施例提供的一种人体动作的识别方法 S103 具体实现流程图；
图 5 是本申请第五实施例提供的一种人体动作的识别方法 S105 具体实现流程图；
图 6 是本申请一实施例提供的一种人体动作的识别设备的结构框图；
图 7 是本申请另一实施例提供的一种终端设备的示意图。

本发明的实施方式

在本申请实施例中，流程的执行主体为终端设备。该终端设备包括但不限于：服务器、计算机、智能手机以及平板电脑等能够执行人体动作的识别操作的设备。图 1 示出了本申请第一实施例提供的人体动作的识别方法的实现流程图，详述如下：

在 S101 中，获取目标对象的视频文件；所述视频文件包括多个视频图像帧。

在本实施例中，管理员可以指定包含目标对象的视频文件作为目标视频文件，在该情况下，终端设备会根据该目标视频文件的文件标识，从视频数据库中下载关于目标对象的视频文件，并对该目标对象的动作行为进行识别。优选地，该终端设备具体为一视频监控设备，会获取当前场景内视频文件；在该情况下，终端设备会将当前场景中拍摄得到的各个对象均识别为目标对象，基于不同拍摄对象的人脸图像，为各个对象配置对象编号，终端设备根据监控过程中生成的视频文件，实时判定各个监控对象的动作类型，若检测到某一目标对象的动作类型在异常动作列表内，则生成警告信息，以通知执行异常动作的监控对象停止该异常行为，实现实时对监控对象的异常动作的警告目的。

可选地，用户可以将目标对象的人脸信息发送给终端设备。终端设备基于该人脸信息

在视频数据库内的各个视频文件进行人脸查找，将包含该人脸信息的视频文件作为目标视频文件。具体的查找操作可以为：终端设备识别视频数据库内的各个视频文件中每个视频图像帧中的候选人脸，提取候选人脸中关键区域的脸部特征值，将各个候选人脸的脸部特征值与目标人脸的人脸信息进行匹配，若两者匹配度大于预设的匹配阈值，则表示两者对

5 应同一实体人，则将该视频文件识别为包含目标对象的人脸图像。

在本实施例中，视频文件包含多个视频图像帧，每个视频图像帧对应一个帧编号，基于帧编号的正序将各个视频图像帧进行排列并封装，生成视频文件。该帧编号可以根据视频图像帧在视频文件中的播放时间确定。

在 S102 中，分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象

10 的人体区域图像。

在本实施例中，终端设备对视频文件进行解析，分别对视频文件中的各个视频图像帧进行人体识别，并提取各个视频图像帧关于目标对象的人体区域图像。提取人体区域图像的具体方式可以为：终端设备通过人脸识别算法，判断该视频图像帧中是否包含人脸区域图像，若不包含，则表示该视频图像帧不包含人体区域图像；反之，若该视频图像帧包含

15 人脸图像，则基于该人脸图像所在的坐标，对该坐标的区域进行轮廓识别，基于识别得到的轮廓信息提取人脸图像对应的人体区域图像，并根据该人脸图像与目标对象的人脸模板进行匹配，从而判断该人体区域图像是否为目标对象的人体区域图像。

可选地，若目标对象的数量为多个，即需要监控多个对象的行为，则终端设备在确定视频图像帧包含的人脸图像的人体区域图像后，则会将该人脸图像与各个目标对象的人脸

20 模板进行匹配，从而确定该人脸图像多对应的目标对象，并在该人体区域图像上标记关联的目标对象的对象标识，继而可以在视频图像帧中快速确定每个目标对象所对应的人体区域图像，方便对多对象的动作跟踪。

可选地，在本实施例中，终端设备可以根据目标对象的对象标识，获取对象标识关联的对象人体模板。该对象人体模板可以用于表示该目标对象的人体特征，例如体型信息、

25 性别信息和/或发型信息，终端设备可以根据该对象人体模板在视频图像帧中进行滑动框取，计算所框取的候选区域与对象人体模板之间的匹配度，若两者匹配度大于预设的匹配阈值，则识别该候选区域为目标对象的人体区域图像；反之，若两者匹配度小于或等于匹配阈值，则识别该候选区域并非目标对象的人体区域图像，继续进行滑动框取；若视频图像帧中所有候选区域均不包含人体区域图像，则对下一帧的视频图像帧重复执行上述操作，识别目

30 标对象的人体区域图像。

在 S103 中，在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，

并获取各个所述关键部位的特征坐标。

在本实施例中，终端设备存储有一个人体关键部位列表，该人体关键部位列表包含有多个人体关键部位，优选地，人体关键部位列表包含有 17 个关键部位，分别为：鼻子，双眼，双耳，双肩，双腕，双手，双腰，双膝，双脚这 17 个关键部位。通过定位多个人体关键部位，并追踪多个关键部位的运动变化情况，能够提高人体动作识别的准确率。

在本实施例中，终端设备在人体区域图像中标记出各个关键部位，具体的标记方式为：基于人体区域图像的轮廓信息，确定目标对象当前的姿态类型，其中，姿态类型具体为：站立类型、行走类型、平躺类型、正坐类型等，继而根据不同关键部位与姿态类型的对应关系，在人体区域图像上标记出各个关键部位。可选地，该对应关系记录有关键部位与人体区域图像的轮廓中心点的距离值以及相对方向向量，终端设备可以基于该距离值以及相对方向向量定位出各个关键部位，并执行标记操作。

在本实施例中，终端设备基于视频图像帧建立一个图像坐标轴，并根据各个关键部位在视频图像帧上的位置，从而确定各个关键部位的特征坐标。可选地，终端设备可以将视频图像帧的左下角的端点作为坐标原点，也可以将图像中心点作为坐标原点，具体根据管理员或设备的默认设置决定。

在 S104 中，根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列。

在本实施例中，终端设备需要确定各个关键部位的运动轨迹，因此会基于关键部位的部位标识，从各个视频图像帧中提取关于该部位标识对应的特征坐标，并将所有关于该特征部位的特征坐标进行封装，生成关于该特征部位的关键特征序列。其中，该关键特征序列中各个元素的在序列中的次序与所属视频图像帧的帧序号一致，即该关键特征序列中各个元素是具有时序关系的，从而能够通过关键特征序列确定关键部位基于时间的推移而变化的情况。

可选地，若部分视频图像帧中的关键部位因被遮挡而不存在对应的特征坐标，终端设备可以根据视频图像帧的帧序号，在预设的坐标轴上建立关于关键部位的特征曲线，基于帧序号依次连接各个特征坐标，而缺失的视频图像帧对应的特征坐标则可以通过平滑算法进行填补，确定缺失的视频图像帧对应的特征坐标。

在 S105 中，通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

在本实施例中，终端设备根据多个关键部位的关键特征序列，则可以确定不同关键部位的运动轨迹，继而确定该目标对象的动作类型。具体地，终端设备可以根据关键特征序

列确定关键部位的运动方向，继而基于多个关键部位的运动方向，与各个候选动作类型的
关键部位的运动方向一一进行匹配，基于匹配的关键部位的个数，例如选取匹配的关键部
位的个数最大的候选动作类型作为目标对象的动作类型。

可选地，终端设备可以设置有最大的帧数，继而终端设备基于该最大的帧数对关键部
5 位的特征序列进行划分，划分为多个特征子序列，分别确定不同特征子序列的动作类
型，由于拍摄的视频文件的时长较长的情况下，用户可能在该拍摄过程中做出多个动作，
基于此，终端设备为了对不同的动作进行划分以及识别，会设置有最大的帧数，实现单用
户的多动作识别的目的。

以上可以看出，本申请实施例提供的一种人体动作的识别方法通过获取所需要进行动
10 作为行为分析的目标用户的视频文件，并对该视频文件的各个视频图像帧进行解析，确定每
个视频图像帧中包含的人体区域图像，在人体区域图像中标记出各个关键部位，并根据各
个关键部位的特征坐标，确定目标对象的各个部位的变化情况，从而确定目标对象的动作
类型，自动识别目标对象的人体动作。与现有的人体动作的识别技术相比，本申请实施例
15 无需依赖神经网络对视频图像进行动作类型的识别，并不借助光流信息，避免了需要进行
时序递归而带来的识别时延，从而提高了识别的效率，而且通过定位多个关键部位，通过
多个关键部位的变动情况确定目标对象的动作，准确率也进一步提高，从而提高了图像识
别的效果以及对象行为分析的效率。

图2示出了本申请第二实施例提供的一种人体动作的识别方法S102的具体实现流程图。
参见图2，相对于图1所述实施例，本实施例提供的一种人体动作的识别方法S102包括：
20 S1021~S1024，具体详述如下：

进一步地，所述分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标
对象的人体区域图像，包括：

在S1021中，通过轮廓识别算法，获取所述视频图像帧的轮廓曲线，并计算各个所述
轮廓曲线所包围的区域面积。

25 在本实施例中，终端设备通过轮廓识别算法，确定该视频图像帧中的轮廓曲线。具体
识别轮廓线的方式可以为：终端设备计算相邻两个坐标点之间的像素值的差值，若该差值
大于预设的轮廓阈值，则识别该坐标点为轮廓线所在的坐标点，连接所有识别得到的轮廓
线上的坐标点，构成一条连续的轮廓曲线。每一条封闭的轮廓曲线对应一个拍摄对象。

在本实施例中，终端设备在视频图像帧上标记出所有轮廓曲线，并将轮廓曲线和/或视
30 频图像帧的边界之间所围成的区域进行积分，从而能够得到关于各个轮廓曲线对应的区域
面积，由于一条轮廓曲线对应一个拍摄对象，基于区域面积，可以确定被拍摄对象的缩放

比例，从而能够选取合适的窗口在提取人体区域图像，提高人体区域图像提取的准确性。

在 S1022 中，根据各个所述区域面积，生成所述视频图像帧的人体识别窗口。

在本实施例中，由于不同的缩放比例，人体识别窗口的尺寸也需要随之调整，基于此，终端设备可以根据各个拍摄对象的区域面积，计算出视频图像帧对应的缩放比例，并查询该缩放比例关联的人体识别窗口尺寸，继而生成与视频图像帧匹配的人体识别窗口。

可选地，在本实施例中，终端设备采用的是 yolov3 的人体识别算法，而 yolov3 需要配置 3 个人体识别窗口。基于此，终端设备根据各个轮廓曲线所围成的区域面积，生成区域面积的分布情况，选取分布密度最大的三个区域面积作为特征面积，并基于三个特征面积生成与之对应的人体识别窗口，即三个 feature map。

在 S1023 中，基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像。

在本实施例中，终端设备在生成与视频图像帧的缩放比例对应的人体识别窗口后，可以通过人体识别窗口在视频图像帧上进行滑动框取，将每一次框取的区域图像作为候选区域图像。若存在多个尺寸的人体识别窗口，则创建与人体识别窗口数量对应的并发线程，并复制该多个视频图像帧，通过多条并发线程分别控制人体识别窗口在不同的视频图像帧上进行滑动框取，即不同尺寸的人体识别窗口的滑动框取操作是相互独立、互不影响的，生成不同尺寸的候选区域图像。

在 S1024 中，分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

在本实施例中，终端设备计算该候选区域图像与标准人体模板之间的重合率，若两者之间的重合率越高，则表示该区域图像所对应的拍摄对象与目标对象的相似度越高，因此可以识别该候选区域为人体的区域图像；反之，若两者之间的重合率越低，则表示该区域图像的形态与目标对象的相似度较低，识别为非人体区域图像。由于视频图像帧中可以包含多个不同用户，因此终端设备会将所有重合率超过预设的重合率阈值的候选区域均识别为人体的区域图像，在该情况下，终端设备可以定位各个人体区域图像的人脸图像，从而将人体图像与目标对象的标准人脸进行匹配，从而选取与标准人脸相匹配的人体的区域图像作为目标对象的人体的区域图像。

在本申请实施例中，通过获取视频图像帧中的轮廓曲线，从而基于各个轮廓曲线的区域面积，确定视频图像帧的缩放比例，并生成与之对应的人体识别窗口进行人体区域图像的识别操作，从而能够提高识别的准确率。

图 3 示出了本申请第三实施例提供的一种人体动作的识别方法 S104 的具体实现流程图。

参见图 3, 相对于图 1 所述的实施例, 本实施例提供的一种人体动作的识别方法 S104 包括: S1041~S1045, 具体详述如下:

进一步地, 所述根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标, 生成关于所述关键部位的关键特征序列, 包括:

5 在 S1041 中, 获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标, 并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值。

在本实施例中, 终端设备需要进行人体关键部位追踪, 若检测到两个相邻图像帧中相同关键部位的位移过大, 则标识两个关键部位属于不同的人体, 从而能够快速进行重追踪, 并且提高动作识别的准确率。基于此, 终端设备会获取帧数相邻的两个视频图像帧中相同
10 关键部位的第一特征坐标以及第二特征坐标, 将两个特征坐标导入到欧氏距离计算公式, 计算出两个坐标点之间的距离值, 即图像距离值。该图像距离值具体指在视频图像帧上两个坐标点之间的距离, 并非该关键部位在实际场景下的移动距离, 因此需要对该图像距离值进行数值转换。

在 S1042 中, 计算所述人体区域图像的图像面积, 并基于所述图像面积确定所述目标
15 对象与拍摄模块之间的拍摄焦距。

在本实施例中, 终端设备获取人体区域图像在视频图像帧中所占据的面积, 即图像面积。终端设备设置有标准的人体面积以及该面积所对应的标准拍摄焦距。终端设备可以计算当前的图像面积与标准的人体面积之间的比例, 确定缩放比例, 基于所述缩放比例以及标准拍摄焦距, 计算该目标对象与拍摄模型之间的实际拍摄焦距, 即上述的拍摄焦距。

20 在 S1043 中, 将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型, 计算两个所述视频图像帧中所述关键部分的实际移动距离; 所述距离转换模型具体为:

$$Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中, $Dist$ 为所述实际移动距离; $StandardDist$ 为所述图像距离值; $FigDist$ 为所述拍
25 摄焦距; $BaseDist$ 为预设的基准焦距; $ActFrame$ 为所述拍摄帧率; $BaseFrame$ 为所述基准帧率。

在本实施例中, 终端设备该视频图像帧对应的拍摄焦距以及两个关键部位的图像距离值以及该视频文件的拍摄帧率导入到距离转换模型内, 从而能够计算关键部位在场景中的实际移动距离。

在 S1044 中, 将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标。

在本实施例中, 终端设备若检测到实际移动距离大于或等于预设的距离阈值, 则表示该关键部位移动距离超过了正常的移动距离, 此时会识别两个视频图像帧中该关键部位属于不同的目标对象, 此时会判定上述两个特征坐标为非关联的特征坐标; 反之, 若该实际移动距离值小于预设的距离阈值, 则表示两个视频图像帧中该关键部位属于同一目标对象, 此时会判定上述两个特征坐标为关联的特征坐标, 实现对目标对象的追踪的目的, 避免在追踪用户 A 的运动轨迹的情况下, 切换到追踪用户 B 的运动轨迹, 提高了动作识别的准确率。

在 S1045, 根据所有所述互为关联的特征坐标生成关于所述关键部位的所述关键特征序列。

在本实施例中, 终端设备将所有非关联的特征坐标进行过滤, 将互为关联的特征坐标进行封装, 生成关于关键部位的关键特征序列。

在本申请实施例中, 通过计算不同帧数下关键部位的实际移动距离, 从而能够对异常的特征坐标点进行过滤, 提高了动作识别的准确性。

图 4 示出了本申请第四实施例提供的一种人体动作的识别方法 S103 的具体实现流程图。参见图 4, 相对于图 1-图 3 所述的实施例, 本实施例提供的一种人体动作的识别方法中 S103 包括: S1031~S1032, 具体详述如下:

进一步地, 所述在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位, 并获取各个所述关键部位的特征坐标, 包括:

在 S1031 中, 对所述人体区域图像进行人脸识别, 确定所述人体区域图像的人脸部位坐标。

在本实施例中, 终端设备对人体区域图像进行人脸识别, 获取该人体区域图像中包含的人脸区域图像, 并将人体区域图像的中心坐标作为人脸部位坐标。具体地, 识别人脸的方式可以为: 终端设备对人体区域图像进行灰度处理, 提取人体区域图像中的各个轮廓线, 根据轮廓线的形状, 选取与人脸曲线相匹配的轮廓线, 并将匹配的轮廓线所围成的区域识别为人脸区域图像, 并获取人脸区域图像的人脸部位坐标。

在 S1032 中, 基于所述人脸部位与各个所述关键部位的位置关系, 在所述人体区域图像中标记各个所述关键部位。

在本实施例中, 终端设备以人脸部位坐标为基准坐标, 以及预设的人脸部位与各个关键部位的位置关系, 能够定位出各个关键部位所在的位置, 并于人体区域图像上进行标

记，该位置关系包含为一个距离向量，即包含以人脸部位坐标为起点、以关键部位为终点所构成的向量。

在本申请实施例中，终端设备通过识别人脸部位坐标，从而能够定位出各个关键部位，提高了关键部位识别的准确性。

5 图 5 示出了本申请第五实施例提供的一种人体动作的识别方法 S105 的具体实现流程图。参见图 5，相对于图 1 至图 3 所述实施例，本实施例提供的一种人体动作的识别方法 S105 包括：S1051~S1052，具体详述如下：

进一步地，所述通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型，包括：

10 在 S1051 中，在预设的坐标轴内标记各个所述关键特征序列的特征坐标，生成关于各个所述关键部位的部位变化曲线。

在本实施例中，终端设备根据各个关键特征序列中各个特征坐标的坐标值以及对应的视频图像帧的帧数，在预设的坐标轴上标记出各个各个特征坐标，并连接各个特征坐标，生成关于关键部位的部位变化曲线。该坐标轴可以以视频图像帧为基础建立的坐标轴，横
15 轴标对应视频图像帧的长，纵坐标对应视频图像帧的宽。

在 S1052 中，将所述部位变化曲线与预设动作库内的各个候选动作的标准动作曲线进行匹配，基于匹配结果确定所述目标对象的动作类型。

在本实施例中，终端设备根据所有关键部位的部位变化曲线与预设动作库中各个候选动作的标准动作曲线进行匹配，计算两个变化曲线的重合率，选取重合率最高的一个候选
20 动作为目标对象的动作类型。

在本申请实施例中，通过绘制关键部位的部位变化曲线从而能够直观地确定目标对象的动作类型，提高了动作类型的准确性。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不应对本申请实施例的实施过程构成任何限定。

25 图 6 示出了本申请一实施例提供的一种人体动作的识别设备的结构框图，该人体动作的识别设备包括的各单元用于执行图 1 对应的实施例中的各步骤。具体请参阅图 1 与图 1 所对应的实施例中的相关描述。为了便于说明，仅示出了与本实施例相关的部分。

参见图 6，所述人体动作的识别设备包括：

30 视频文件获取单元 61，用于获取目标对象的视频文件；所述视频文件包括多个视频图像帧；

人体区域图像提取单元 62，用于分别解析各个所述视频图像帧，提取所述视频图像帧

中关于所述目标对象的人体区域图像；

关键部位识别单元 63，用于在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标；

关键特征序列生成单元 64，用于根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列；

动作类型识别单元 65，用于通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

可选地，所述人体区域图像提取单元 62 包括：

轮廓曲线获取单元，用于通过轮廓识别算法，获取所述视频图像帧的轮廓曲线，并计算各个所述轮廓曲线所包围的区域面积；

人体识别窗口生成单元，用于根据各个所述区域面积，生成所述视频图像帧的人体识别窗口；

候选区域图像提取单元，用于基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像；

人体区域图像匹配单元，用于分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

可选地，所述关键特征序列生成单元 64 包括：

图像距离值计算单元，用于获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标，并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值；

拍摄焦距确定单元，用于计算所述人体区域图像的图像面积，并基于所述图像面积确定所述目标对象与拍摄模块之间的拍摄焦距；

实际移动距离计算单元，用于将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型，计算两个所述视频图像帧中所述关键部分的实际移动距离；所述距离转换模型具体为：

$$Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中， $Dist$ 为所述实际移动距离； $StandardDist$ 为所述图像距离值； $FigDist$ 为所述拍摄焦距； $BaseDist$ 为预设的基准焦距； $ActFrame$ 为所述拍摄帧率； $BaseFrame$ 为所述基准

帧率；

关联坐标识别单元，用于将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标；

关联坐标封装单元，用于根据所有所述互为关联的特征坐标生成关于所述关键部位的
5 所述关键特征序列。

可选地，所述关键部位识别单元 63 包括：

人脸识别单元，用于对所述人体区域图像进行人脸识别，确定所述人体区域图像的人脸部位坐标；

关键部位标记单元，用于基于所述人脸部位与各个所述关键部位的位置关系，在所述
10 人体区域图像中标记各个所述关键部位。

可选地，所述动作类型识别单元 65 包括：

部位变化曲线生成单元，用于在预设的坐标轴内标记各个所述关键特征序列的特征坐标，生成关于各个所述关键部位的部位变化曲线；

候选动作选取单元，用于将所述部位变化曲线与预设动作库内的各个候选动作的标准
15 动作曲线进行匹配，基于匹配结果确定所述目标对象的动作类型。

因此，本申请实施例提供的人体动作的识别设备同样可以无需依赖神经网络对视频图像进行动作类型的识别，并不借助光流信息，避免了需要进行时序递归而带来的识别时延，从而提高了识别的效率，而且通过定位多个关键部位，通过多个关键部位的变动情况确定目标对象的动作，准确率也进一步提高，从而提高了图像识别的效果以及对象行为分析的效率。
20

图 7 是本申请另一实施例提供的一种终端设备的示意图。如图 7 所示，该实施例的终端设备 7 包括：处理器 70、存储器 71 以及存储在所述存储器 71 中并可在所述处理器 70 上运行的计算机可读指令 72，例如人体动作的识别程序。所述处理器 70 执行所述计算机可读指令 72 时实现上述各个人体动作的识别方法实施例中的步骤，例如图 1 所示的 S101
25 至 S105。或者，所述处理器 70 执行所述计算机可读指令 72 时实现上述各装置实施例中各单元的功能，例如图 6 所示模块 61 至 65 功能。

示例性的，所述计算机可读指令 72 可以被分割成一个或多个单元，所述一个或者多个单元被存储在所述存储器 71 中，并由所述处理器 70 执行，以完成本申请。所述一个或多个单元可以是能够完成特定功能的一系列计算机可读指令指令段，该指令段用于描述所述
30 计算机可读指令 72 在所述终端设备 7 中的执行过程。例如，所述计算机可读指令 72 可以被分割视频文件获取单元、人体区域图像提取单元、关键部位识别单元、关键特征序列生

成单元以及动作类型识别单元，各单元具体功能如上所述。

所述终端设备 7 可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述终端设备可包括，但不仅限于，处理器 70、存储器 71。本领域技术人员可以理解，图 7 仅仅是终端设备 7 的示例，并不构成对终端设备 7 的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件，例如所述终端设备还可以包括输入输出设备、网络接入设备、总线等。

所称处理器 70 可以是中央处理单元（Central Processing Unit，CPU），还可以是其他通用处理器、数字信号处理器（Digital Signal Processor，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现成可编程门阵列（Field-Programmable Gate Array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

所述存储器 71 可以是所述终端设备 7 的内部存储单元，例如终端设备 7 的硬盘或内存。所述存储器 71 也可以是所述终端设备 7 的外部存储设备，例如所述终端设备 7 上配备的插接式硬盘，智能存储卡（Smart Media Card，SMC），安全数字（Secure Digital，SD）卡，闪存卡（Flash Card）等。进一步地，所述存储器 71 还可以既包括所述终端设备 7 的内部存储单元也包括外部存储设备。所述存储器 71 用于存储所述计算机可读指令以及所述终端设备所需的其他程序和数据。所述存储器 71 还可以用于暂时地存储已经输出或者将要输出的数据。

另外，在本申请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。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程 ROM（PROM）、电可编程 ROM（EPROM）、电可擦除可编程 ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM（SRAM）、动态 RAM（DRAM）、同步 DRAM（SDRAM）、双数据率 SDRAM（DDRSDRAM）、增强型 SDRAM（ESDRAM）、同步链路（Synchlink）DRAM（SLDRAM）、存储器总线（Rambus）直接 RAM（RDRAM）、

直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，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，或者对其中部分技术特征进行等同替换；而这些修改或者
5 替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权利要求书

1.一种人体动作的识别方法，其特征在于，包括：

获取目标对象的视频文件；所述视频文件包括多个视频图像帧；

分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域
5 图像；

在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标；

根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列；

10 通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

2. 根据权利要求1所述的识别方法，其特征在于，所述分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像，包括：

通过轮廓识别算法，获取所述视频图像帧的轮廓曲线，并计算各个所述轮廓曲线所包围的区域面积；

15 根据各个所述区域面积，生成所述视频图像帧的人体识别窗口；

基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像；

分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

3. 根据权利要求1所述的识别方法，其特征在于，所述根据所述关键部位在各个所述
20 视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列，包括：

获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标，并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值；

计算所述人体区域图像的图像面积，并基于所述图像面积确定所述目标对象与拍摄模块之间的拍摄焦距；

25 将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型，计算两个所述视频图像帧中所述关键部分的实际移动距离；所述距离转换模型具体为：

$$Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中， $Dist$ 为所述实际移动距离； $StandardDist$ 为所述图像距离值； $FigDist$ 为所述拍摄焦距； $BaseDist$ 为预设的基准焦距； $ActFrame$ 为所述拍摄帧率； $BaseFrame$ 为所述基准

帧率;

将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标;

根据所有所述互为关联的特征坐标生成关于所述关键部位的所述关键特征序列。

- 5 4. 根据权利要求 1-3 任一项所述的识别方法, 其特征在于, 所述在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位, 并获取各个所述关键部位的特征坐标, 包括:

对所述人体区域图像进行人脸识别, 确定所述人体区域图像的人脸部位坐标;

- 10 基于所述人脸部位与各个所述关键部位的位置关系, 在所述人体区域图像中标记各个所述关键部位。

5. 根据权利要求 1-3 任一项所述的识别方法, 其特征在于, 所述通过各个所述关键部位的所述关键特征序列, 确定所述目标对象的动作类型, 包括:

在预设的坐标轴内标记各个所述关键特征序列的特征坐标, 生成关于各个所述关键部位的部位变化曲线;

- 15 将所述部位变化曲线与预设动作库内的各个候选动作的标准动作曲线进行匹配, 基于匹配结果确定所述目标对象的动作类型。

6. 一种人体动作的识别设备, 其特征在于, 包括:

视频文件获取单元, 用于获取目标对象的视频文件; 所述视频文件包括多个视频图像帧;

- 20 人体区域图像提取单元, 用于分别解析各个所述视频图像帧, 提取所述视频图像帧中关于所述目标对象的人体区域图像;

关键部位识别单元, 用于在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位, 并获取各个所述关键部位的特征坐标;

- 25 关键特征序列生成单元, 用于根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标, 生成关于所述关键部位的关键特征序列;

动作类型识别单元, 用于通过各个所述关键部位的所述关键特征序列, 确定所述目标对象的动作类型。

7. 根据权利要求 6 所述的识别设备, 其特征在于, 所述人体区域图像提取单元包括:

- 30 轮廓曲线获取单元, 用于通过轮廓识别算法, 获取所述视频图像帧的轮廓曲线, 并计算各个所述轮廓曲线所包围的区域面积;

人体识别窗口生成单元, 用于根据各个所述区域面积, 生成所述视频图像帧的人体识

别窗口；

候选区域图像提取单元，用于基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像；

5 人体区域图像匹配单元，用于分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

8. 根据权利要求 6 所述的识别设备，其特征在于，所述关键特征序列生成单元包括：

10 图像距离值计算单元，用于获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标，并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值；

拍摄焦距确定单元，用于计算所述人体区域图像的图像面积，并基于所述图像面积确定所述目标对象与拍摄模块之间的拍摄焦距；

15 实际移动距离计算单元，用于将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型，计算两个所述视频图像帧中所述关键部分的实际移动距离；所述距离转换模型具体为：

$$Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中， $Dist$ 为所述实际移动距离； $StandardDist$ 为所述图像距离值； $FigDist$ 为所述拍摄焦距； $BaseDist$ 为预设的基准焦距； $ActFrame$ 为所述拍摄帧率； $BaseFrame$ 为所述基准帧率；

20 关联坐标识别单元，用于将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标；

关联坐标封装单元，用于根据所有所述互为关联的特征坐标生成关于所述关键部位的所述关键特征序列。

25 9. 根据权利要求 6-8 任一项所述的识别设备，其特征在于，所述关键部位识别单元包括：

人脸识别单元，用于对所述人体区域图像进行人脸识别，确定所述人体区域图像的人脸部位坐标；

关键部位标记单元，用于基于所述人脸部位与各个所述关键部位的位置关系，在所述人体区域图像中标记各个所述关键部位。

10. 根据权利要求 6-8 任一项所述的识别设备，其特征在于，所述动作类型识别单元包括：

部位变化曲线生成单元，用于在预设的坐标轴内标记各个所述关键特征序列的特征坐标，生成关于各个所述关键部位的部位变化曲线；

5 候选动作选取单元，用于将所述部位变化曲线与预设动作库内的各个候选动作的标准动作曲线进行匹配，基于匹配结果确定所述目标对象的动作类型。

11. 一种终端设备，其特征在于，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现如下步骤：

10 获取目标对象的视频文件；所述视频文件包括多个视频图像帧；

分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像；

在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标；

15 根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列；

通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

12. 根据权利要求 11 所述的终端设备，其特征在于，所述分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像，包括：

20 通过轮廓识别算法，获取所述视频图像帧的轮廓曲线，并计算各个所述轮廓曲线所包围的区域面积；

根据各个所述区域面积，生成所述视频图像帧的人体识别窗口；

基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像；

25 分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

13. 根据权利要求 11 所述的终端设备，其特征在于，所述根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列，包括：

获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标，并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值；

30 计算所述人体区域图像的图像面积，并基于所述图像面积确定所述目标对象与拍摄模块之间的拍摄焦距；

将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型，计算两个所述视频图像帧中所述关键部分的实际移动距离；所述距离转换模型具体为：

$$Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中，*Dist* 为所述实际移动距离；*StandardDist* 为所述图像距离值；*FigDist* 为所述拍摄焦距；*BaseDist* 为预设的基准焦距；*ActFrame* 为所述拍摄帧率；*BaseFrame* 为所述基准帧率；

将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标；

根据所有所述互为关联的特征坐标生成关于所述关键部位的所述关键特征序列。

10 14. 根据权利要求 11-13 任一项所述的终端设备，其特征在于，所述在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标，包括：

对所述人体区域图像进行人脸识别，确定所述人体区域图像的人脸部位坐标；

15 基于所述人脸部位与各个所述关键部位的位置关系，在所述人体区域图像中标记各个所述关键部位。

15. 根据权利要求 11-13 任一项所述的终端设备，其特征在于，所述通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型，包括：

在预设的坐标轴内标记各个所述关键特征序列的特征坐标，生成关于各个所述关键部位的部位变化曲线；

20 将所述部位变化曲线与预设动作库内的各个候选动作的标准动作曲线进行匹配，基于匹配结果确定所述目标对象的动作类型。

16. 一种计算机非易失性可读存储介质，所述计算机非易失性可读存储介质存储有计算机可读指令，其特征在于，所述计算机可读指令被处理器执行时实现如下步骤：

获取目标对象的视频文件；所述视频文件包括多个视频图像帧；

25 分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像；

在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标；

根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键

部位的关键特征序列；

通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型。

17. 根据权利要求 16 所述的计算机非易失性可读存储介质，其特征在于，所述分别解析各个所述视频图像帧，提取所述视频图像帧中关于所述目标对象的人体区域图像，包括：

5 通过轮廓识别算法，获取所述视频图像帧的轮廓曲线，并计算各个所述轮廓曲线所包围的区域面积；

根据各个所述区域面积，生成所述视频图像帧的人体识别窗口；

基于所述人体识别窗口在所述视频图像帧上进行滑动框取，生成多个候选区域图像；

10 分别计算各个所述候选区域图像与标准人体模板之间的重合率，并选取所述重合率大于预设重合率阈值的所述候选区域图像作为所述人体区域图像。

18. 根据权利要求 16 所述的计算机非易失性可读存储介质，其特征在于，所述根据所述关键部位在各个所述视频图像帧中对应的所述特征坐标，生成关于所述关键部位的关键特征序列，包括：

15 获取帧数相邻的两个所述视频图像帧内同一所述关键部位的第一特征坐标以及第二特征坐标，并计算所述第一特征坐标与所述第二特征坐标之间的图像距离值；

计算所述人体区域图像的图像面积，并基于所述图像面积确定所述目标对象与拍摄模块之间的拍摄焦距；

将所述拍摄焦距、所述图像距离值以及所述视频文件的拍摄帧率导入到距离转换模型，计算两个所述视频图像帧中所述关键部分的实际移动距离；所述距离转换模型具体为：

$$20 \quad Dist = StandardDist \cdot \left(1 + \frac{FigDist - BaseDist}{BaseDist}\right) \cdot \left(1 + \frac{BaseFrame - ActFrame}{BaseFrame}\right)$$

其中， $Dist$ 为所述实际移动距离； $StandardDist$ 为所述图像距离值； $FigDist$ 为所述拍摄焦距； $BaseDist$ 为预设的基准焦距； $ActFrame$ 为所述拍摄帧率； $BaseFrame$ 为所述基准帧率；

25 将所述实际移动距离小于预设的距离阈值的两个所述特征坐标识别为互为关联的特征坐标；

根据所有所述互为关联的特征坐标生成关于所述关键部位的所述关键特征序列。

19. 根据权利要求 16-18 任一项所述的计算机非易失性可读存储介质，其特征在于，所述在所述人体区域图像中标记出预设的人体关键部位列表内的各个关键部位，并获取各个所述关键部位的特征坐标，包括：

对所述人体区域图像进行人脸识别，确定所述人体区域图像的人脸部位坐标；

基于所述人脸部位与各个所述关键部位的位置关系，在所述人体区域图像中标记各个所述关键部位。

20. 如权利要求 16-18 任一项所述的计算机非易失性可读存储介质，其特征在于，所述通过各个所述关键部位的所述关键特征序列，确定所述目标对象的动作类型，包括：

在预设的坐标轴内标记各个所述关键特征序列的特征坐标，生成关于各个所述关键部位的部位变化曲线；

将所述部位变化曲线与预设动作库内的各个候选动作的标准动作曲线进行匹配，基于匹配结果确定所述目标对象的动作类型。

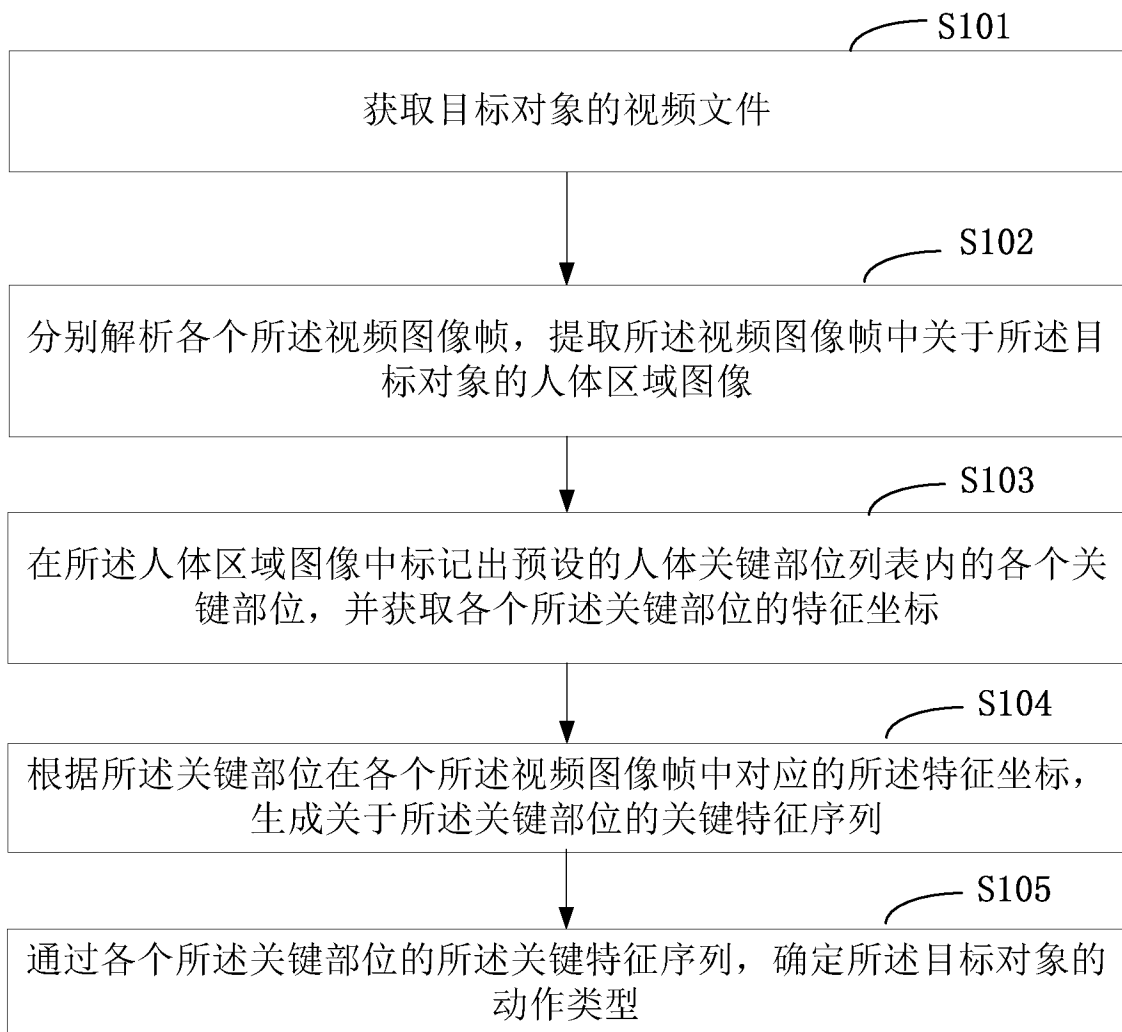


图 1

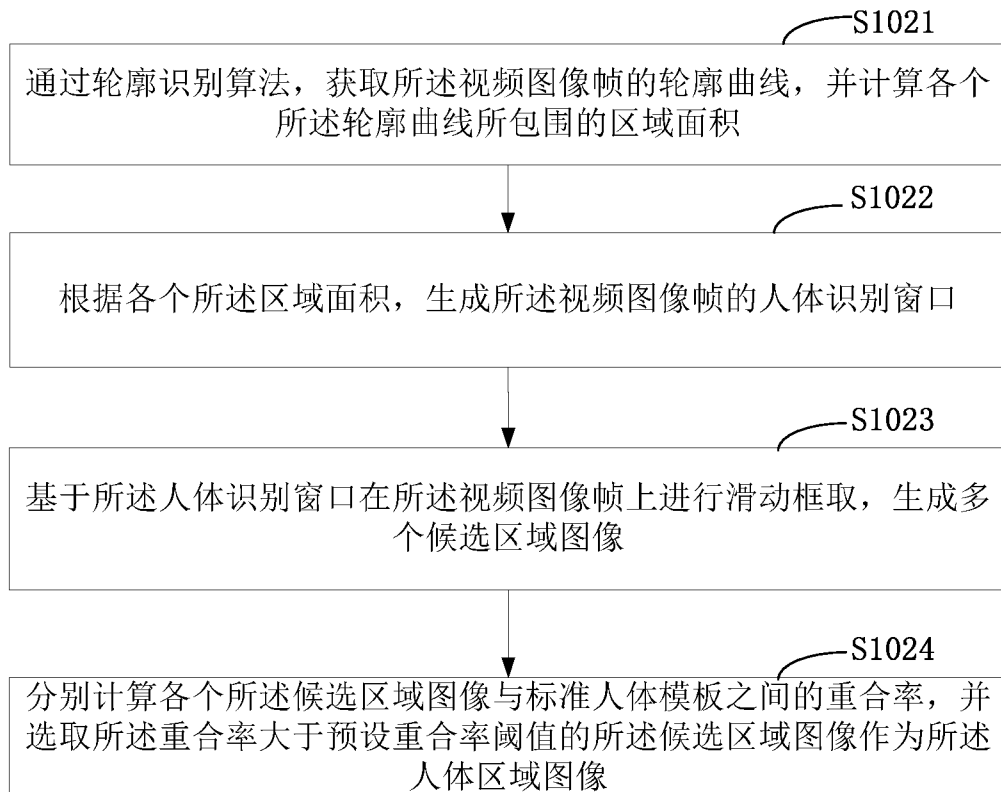


图 2

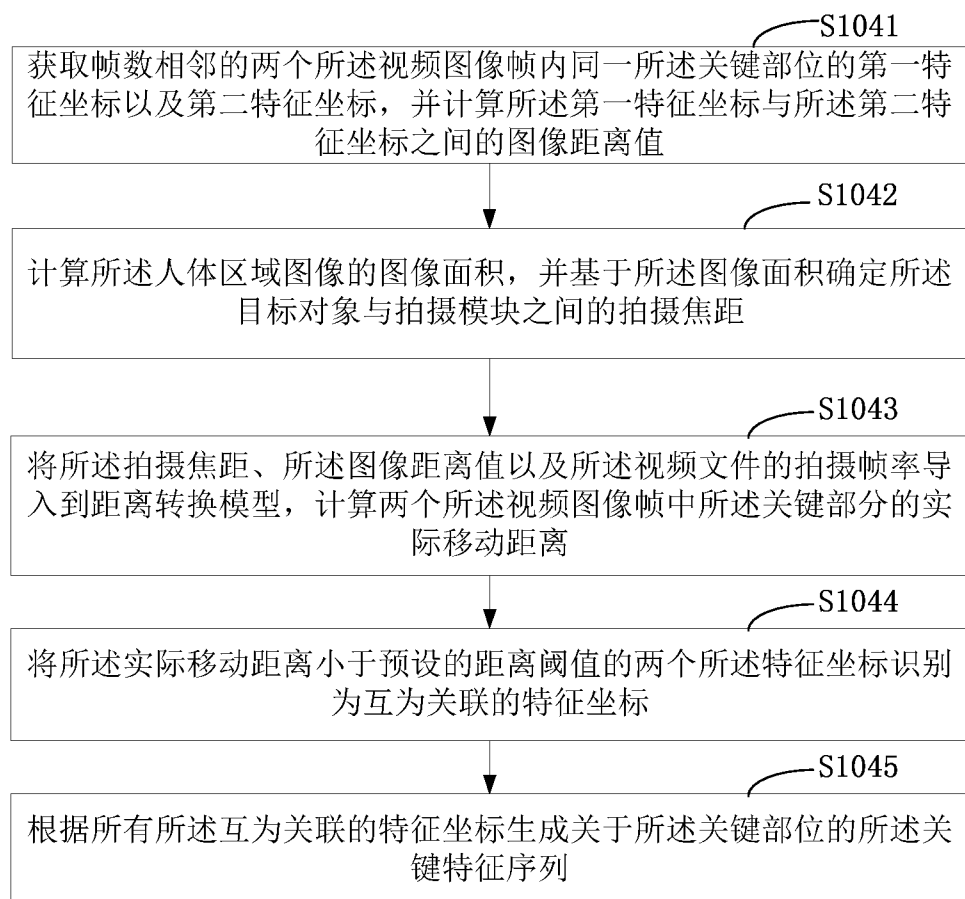


图 3

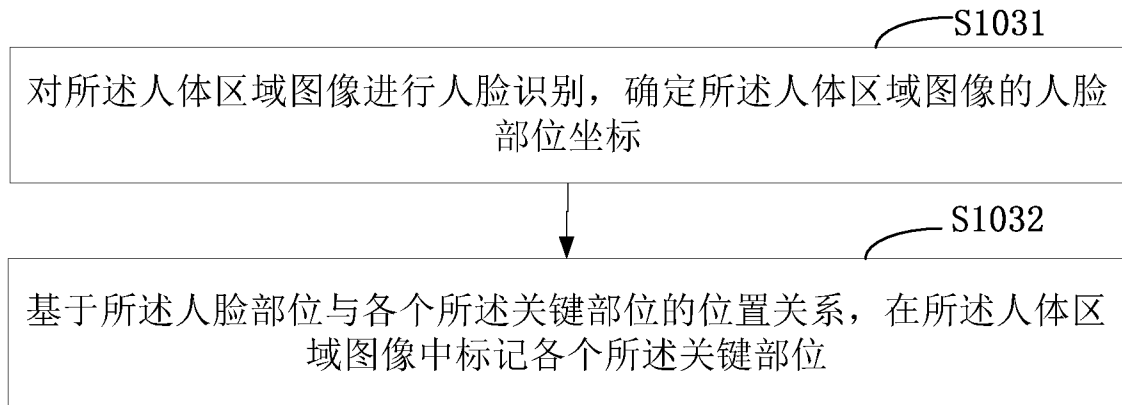


图 4

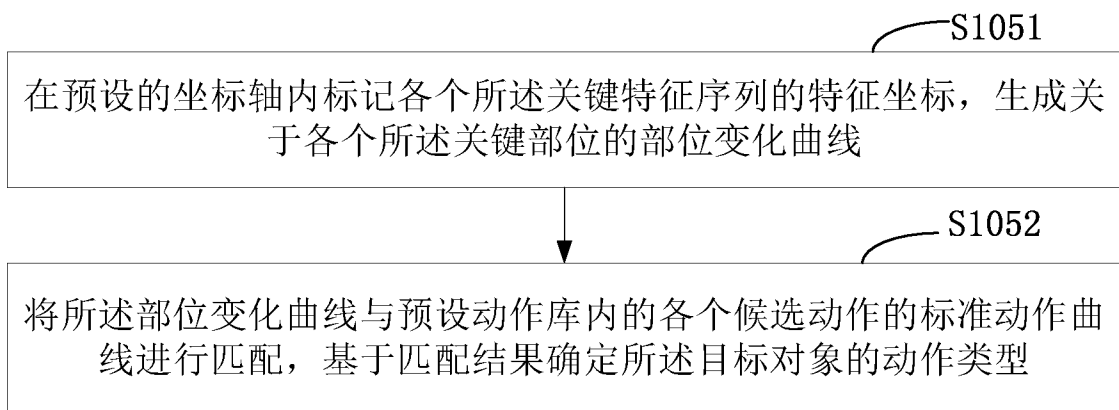


图 5

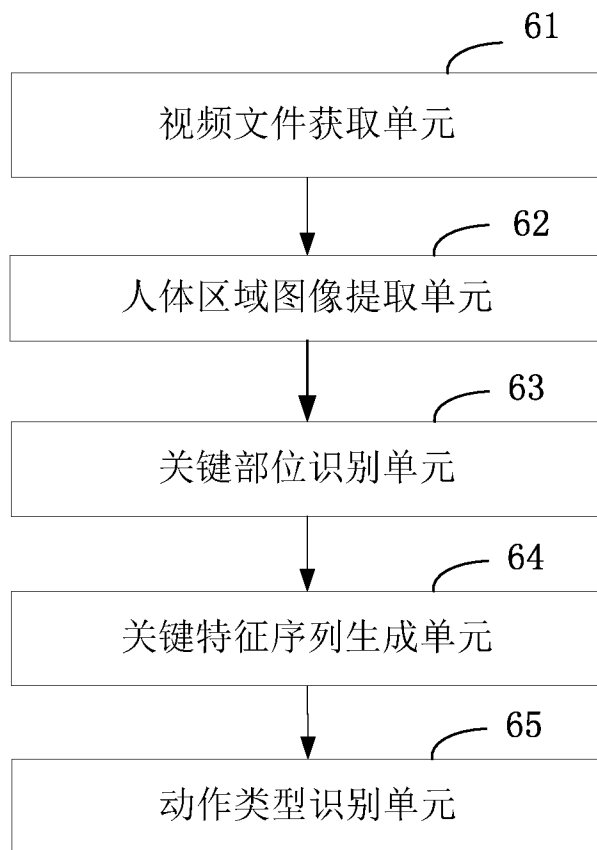


图 6

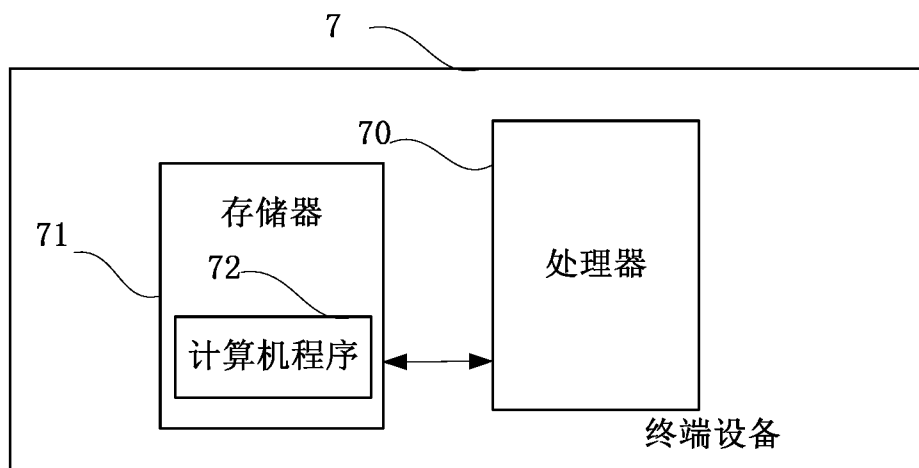


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/103164

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i; G06K 9/32(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)

WPI, EPODOC, CNPAT, CNKI: 人体, 动作, 识别, 视频, 图像, 序列, 部位, 特征, 匹配, 坐标, 目标, 对象, human, body, identify, recognition, video, image, characteristic, match, coordinate, target, object

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110147717 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 20 August 2019 (2019-08-20) description, paragraphs [0031]-[0049]	1-20
X	CN 105930767 A (NANJING HUAJIE IMI TECHNOLOGY CO., LTD.) 07 September 2016 (2016-09-07) description, paragraphs [0055]-[0089]	1, 2, 4-7, 9-12, 14-17, 19, 20
A	CN 109325456 A (PCI-SUNTEK TECHNOLOGY CO., LTD.) 12 February 2019 (2019-02-12) entire document	1-20
A	CN 109460702 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 12 March 2019 (2019-03-12) entire document	1-20
A	CN 107335192 A (SHENZHEN ORBBEC CO., LTD.) 10 November 2017 (2017-11-10) entire document	1-20
A	US 9911198 B2 (CANON KABUSHIKI KAISHA) 06 March 2018 (2018-03-06) entire document	1-20



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

23 December 2019

Date of mailing of the international search report

31 December 2019

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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/103164

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	110147717	A	20 August 2019	None	
CN	105930767	A	07 September 2016	None	
CN	109325456	A	12 February 2019	None	
CN	109460702	A	12 March 2019	None	
CN	107335192	A	10 November 2017	None	
US	9911198	B2	06 March 2018	None	

国际检索报告

国际申请号

PCT/CN2019/103164

A. 主题的分类

G06K 9/00(2006.01)i; G06K 9/32(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))

WPI, EPDOC, CNPAT, CNKI: 人体, 动作, 识别, 视频, 图像, 序列, 部位, 特征, 匹配, 坐标, 目标, 对象,
human, body, identify, recognition, video, image, characteristic, match, coordinate, target, object

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
PX	CN 110147717 A (平安科技深圳有限公司) 2019年 8月 20日 (2019 - 08 - 20) 说明书第[0031]-[0049]段	1-20
X	CN 105930767 A (南京华捷艾米软件科技有限公司) 2016年 9月 7日 (2016 - 09 - 07) 说明书第[0055]-[0089]段	1-2、4-7、9-12、 14-17、19-20
A	CN 109325456 A (佳都新太科技股份有限公司) 2019年 2月 12日 (2019 - 02 - 12) 全文	1-20
A	CN 109460702 A (华南理工大学) 2019年 3月 12日 (2019 - 03 - 12) 全文	1-20
A	CN 107335192 A (深圳奥比中光科技有限公司) 2017年 11月 10日 (2017 - 11 - 10) 全文	1-20
A	US 9911198 B2 (CANON KABUSHIKI KAISHA) 2018年 3月 6日 (2018 - 03 - 06) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 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

2019年 12月 23日

国际检索报告邮寄日期

2019年 12月 31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

受权官员

刘芬

电话号码 86-(10)-53961434

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/103164

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	110147717	A	2019年 8月 20日	无	
CN	105930767	A	2016年 9月 7日	无	
CN	109325456	A	2019年 2月 12日	无	
CN	109460702	A	2019年 3月 12日	无	
CN	107335192	A	2017年 11月 10日	无	
US	9911198	B2	2018年 3月 6日	无	