(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110533685 A
(43)申请公布日 2019.12.03

(21)申请号 201910819029.5

(22)申请日 2019.08.30

(71)申请人 腾讯科技(深圳)有限公司
地址 518000 广东省深圳市南山区高新区科技中一路腾讯大厦35层

(72)发明人 周文

(74)专利代理机构 北京康信知识产权代理有限责任公司 11240
代理人 周婷婷

(51)Int.Cl.
G06T 7/20(2017.01)
G06F 16/51(2019.01)

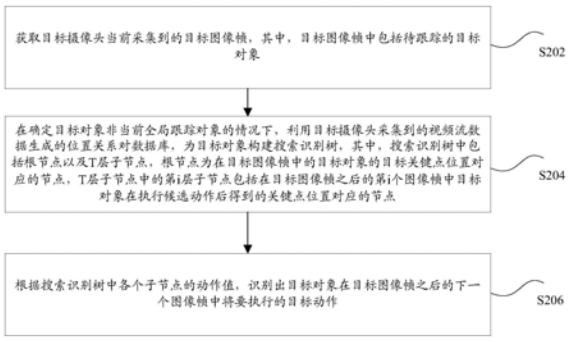
权利要求书4页 说明书16页 附图6页

(54)发明名称

对象跟踪方法和装置、存储介质及电子装置

(57)摘要

本发明公开了一种对象跟踪方法和装置、存储介质及电子装置。其中，该方法包括：获取目标摄像头当前采集到的目标图像帧，目标图像帧中包括待跟踪的目标对象；在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，搜索识别树中包括根节点以及T层子节点；根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作。本发明解决了由于对象被遮挡所导致的跟踪失败的技术问题。



CN 110533685 A

1.一种对象跟踪方法,其特征在于,包括:

获取目标摄像头当前采集到的目标图像帧,其中,所述目标图像帧中包括待跟踪的目标对象;

在确定所述目标对象非当前全局跟踪对象的情况下,利用所述目标摄像头采集到的视频流数据生成的位置关系对数据库,为所述目标对象构建搜索识别树,其中,所述搜索识别树中包括根节点以及T层子节点,所述根节点为在所述目标图像帧中的所述目标对象的目标关键点位置对应的节点,所述T层子节点中的第i层子节点包括在所述目标图像帧之后的第i个图像帧中所述目标对象在执行候选动作后得到的关键点位置对应的节点,所述i小于等于所述T;

根据所述搜索识别树中各个子节点的动作值,识别出所述目标对象在所述目标图像帧之后的下一个图像帧中将要执行的目标动作,其中,每个所述子节点的所述动作值用于指示所述子节点与所述根节点之间的相似度。

2.根据权利要求1所述的方法,其特征在于,所述利用所述目标摄像头采集到的视频流数据生成的位置关系对数据库,为所述目标对象构建搜索识别树包括:

获取所述目标摄像头在所述目标图像帧之后采集到的T个图像帧;

对所述T个图像帧依次执行以下操作;

确定与第i个图像帧对应的待生成的第i层子节点所属的当前父节点;

从所述位置关系对数据库中查找与所述当前父节点相关联的目标位置关系对集合,其中,所述目标位置关系对集合中的每个目标位置关系对中包括的第一关键点位置与所述当前父节点对应的关键点位置的相似度大于第一阈值,所述位置关系对数据库中的每个位置关系对用于记录所述视频流数据中采集时间相邻的两个图像帧中同一个对象执行候选动作前后的关键点位置;

获取每个所述目标位置关系对中与所述第一关键点位置匹配的第二关键点位置,其中,所述第一关键点位置所在图像帧位于所述第二关键点位置所在图像帧之前;

利用所述第二关键点位置生成所述第i层子节点;

在所述第i层子节点未达到第T层子节点的情况下,获取第i+1个图像帧作为所述第i个图像帧。

3.根据权利要求2所述的方法,其特征在于,所述利用所述第二关键点位置生成所述第i层子节点包括:

依次执行以下操作,直至得到所述第i层子节点:

为当前的所述第二关键点位置生成当前对象子节点;

从所述第i个图像帧中提取与当前的所述第二关键点位置对应的第一图像特征;

比对所述第一图像特征和从所述目标图像帧中提取出的与所述目标关键点位置对应的第二图像特征,得到所述当前对象子节点与所述根节点之间的位置相似度和外观相似度;

根据所述位置相似度和所述外观相似度确定所述当前对象子节点与所述根节点之间的对象相似度;

将所述对象相似度作为所述当前对象子节点的动作值,并确定完成所述当前对象子节点的生成;

获取下一个所述第二关键点位置作为所述当前的所述第二关键点位置。

4.根据权利要求3所述的方法，其特征在于，在所述利用所述第二关键点位置生成所述第i层子节点之后，还包括：

利用所述第i层子节点的全部动作值，更新所述当前父节点的动作值；

在所述当前父节点非所述根节点的情况下，确定所述当前父节点所属的上一层父节点；利用与所述当前父节点属于同一层的子节点的全部动作值，更新所述上一层父节点的动作值。

5.根据权利要求1所述的方法，其特征在于，在所述利用所述目标摄像头采集到的视频流数据生成的位置关系对数据库，为所述目标对象构建搜索识别树之前，还包括：

获取所述视频流数据中采集时间相邻的两个图像帧，其中，所述两个图像帧包括第一图像帧和第二图像帧；

在从所述第一图像帧中识别出第一候选对象，并从所述第二图像帧中识别出第二候选对象的情况下，比对所述第一候选对象的关键点位置与所述第二候选对象的关键点位置；

在比对的结果指示所述第一候选对象与所述第二候选对象关联的情况下，确定所述第一候选对象与所述第二候选对象为同一个对象，并将所述第一候选对象的关键点位置与所述第二候选对象的关键点位置，记录为所述位置关系对数据库中的一个位置关系对。

6.根据权利要求1所述的方法，其特征在于，所述根据所述搜索识别树中各个子节点的动作值，识别出所述目标对象在所述目标图像帧之后的下一个图像帧中将要执行的目标动作包括：

在所述目标图像帧之后的下一个图像帧对应的一层子节点中，确定出最大的动作值对应的目标子节点；

将所述目标子节点对应的关键点位置所指示的动作作为所述目标动作。

7.根据权利要求1至6中任一项所述的方法，其特征在于，在所述获取目标摄像头当前采集到的目标图像帧之后，还包括：

对所述目标图像帧进行运动检测；

在检测到所述目标图像帧中包括移动对象的情况下，对所述移动对象进行目标检测得到第一检测结果，并对所述移动对象进行关键点位置检测得到第二检测结果；

在所述第一检测结果和所述第二检测结果指示所述移动对象为所述目标对象的情况下，比对所述目标对象与当前全局跟踪对象；

在比对的结果指示所述目标对象与所述当前全局跟踪对象未匹配的情况下，为所述目标对象分配新的对象标识，并确定为所述目标对象构建所述搜索识别树。

8.一种对象跟踪装置，其特征在于，包括：

第一获取单元，用于获取目标摄像头当前采集到的目标图像帧，其中，所述目标图像帧中包括待跟踪的目标对象；

构建单元，用于在确定所述目标对象非当前全局跟踪对象的情况下，利用所述目标摄像头采集到的视频流数据生成的位置关系对数据库，为所述目标对象构建搜索识别树，其中，所述搜索识别树中包括根节点以及T层子节点，所述根节点为在所述目标图像帧中的所述目标对象的目标关键点位置对应的节点，所述T层子节点中的第i层子节点包括在所述目标图像帧之后的第i个图像帧中所述目标对象在执行候选动作后得到的关键点位置对应的

节点，所述i小于等于所述T；

识别单元，用于根据所述搜索识别树中各个子节点的动作值，识别出所述目标对象在所述目标图像帧之后的下一个图像帧中将要执行的目标动作，其中，每个所述子节点的所述动作值用于指示所述子节点与所述根节点之间的相似度。

9.根据权利要求8所述的装置，其特征在于，所述构建单元包括：

第一获取模块，用于获取所述目标摄像头在所述目标图像帧之后采集到的T个图像帧；

处理模块，用于对所述T个图像帧依次执行以下操作；

确定与第i个图像帧对应的待生成的第i层子节点所属的当前父节点；

从所述位置关系对数据库中查找与所述当前父节点相关联的目标位置关系对集合，其中，所述目标位置关系对集合中的每个目标位置关系对中包括的第一关键点位置与所述当前父节点对应的关键点位置的相似度大于第一阈值，所述位置关系对数据库中的每个位置关系对用于记录所述视频流数据中采集时间相邻的两个图像帧中同一个对象执行候选动作前后的关键点位置；

获取每个所述目标位置关系对中与所述第一关键点位置匹配的第二关键点位置，其中，所述第一关键点位置所在图像帧位于所述第二关键点位置所在图像帧之前；

利用所述第二关键点位置生成所述第i层子节点；

在所述第i层子节点未达到第T层子节点的情况下，获取第i+1个图像帧作为所述第i个图像帧。

10.根据权利要求9所述的装置，其特征在于，所述处理模块通过以下步骤实现利用所述第二关键点位置生成所述第i层子节点：

依次执行以下操作，直至得到所述第i层子节点：

为当前的所述第二关键点位置生成当前对象子节点；

从所述第i个图像帧中提取与当前的所述第二关键点位置对应的第一图像特征；

比对所述第一图像特征和从所述目标图像帧中提取出的与所述目标关键点位置对应的第二图像特征，得到所述当前对象子节点与所述根节点之间的位置相似度和外观相似度；

根据所述位置相似度和所述外观相似度确定所述当前对象子节点与所述根节点之间的对象相似度；

将所述对象相似度作为所述当前对象子节点的动作值，并确定完成所述当前对象子节点的生成；

获取下一个所述第二关键点位置作为所述当前的所述第二关键点位置。

11.根据权利要求10所述的装置，其特征在于，所述处理模块还用于：

在所述利用所述第二关键点位置生成所述第i层子节点之后，利用所述第i层子节点的全部动作值，更新所述当前父节点的动作值；

在所述当前父节点非所述根节点的情况下，确定所述当前父节点所属的上一层父节点；利用与所述当前父节点属于同一层的子节点的全部动作值，更新所述上一层父节点的动作值。

12.根据权利要求8所述的装置，其特征在于，还包括：

第二获取单元，用于在所述利用所述目标摄像头采集到的视频流数据生成的位置关系

对数据库，为所述目标对象构建搜索识别树之前，获取所述视频流数据中采集时间相邻的两个图像帧，其中，所述两个图像帧包括第一图像帧和第二图像帧；

第一比对单元，用于在从所述第一图像帧中识别出第一候选对象，并从所述第二图像帧中识别出第二候选对象的情况下，比对所述第一候选对象的关键点位置与所述第二候选对象的关键点位置；

确定单元，用于在比对的结果指示所述第一候选对象与所述第二候选对象关联的情况下，确定所述第一候选对象与所述第二候选对象为同一个对象，并将所述第一候选对象的关键点位置与所述第二候选对象的关键点位置记录为所述位置关系对数据库中的一个位置关系对。

13.根据权利要求8所述的装置，其特征在于，所述识别单元包括：

第一确定模块，用于在所述目标图像帧之后的下一个图像帧对应的一层子节点中，确定出最大的动作值对应的目标子节点；

第二确定模块，用于将所述目标子节点对应的关键点位置所指示的动作作为所述目标动作。

14.根据权利要求8至13中任一项所述的装置，其特征在于，还包括：

第一检测单元，用于在所述获取目标摄像头当前采集到的目标图像帧之后，对所述目标图像帧进行运动检测；

第二检测单元，用于在检测到所述目标图像帧中包括移动对象的情况下，对所述移动对象进行目标检测得到第一检测结果，并对所述移动对象进行关键点位置检测得到第二检测结果；

第二比对单元，用于在所述第一检测结果和所述第二检测结果指示所述移动对象为所述目标对象的情况下，比对所述目标对象与当前全局跟踪对象；

分配单元，用于在比对的结果指示所述目标对象与所述当前全局跟踪对象未匹配的情况下，为所述目标对象分配新的对象标识，并确定为所述目标对象构建所述搜索识别树。

15.一种存储介质，所述存储介质包括存储的程序，其中，所述程序运行时执行上述权利要求1至7任一项中所述的方法。

16.一种电子装置，包括存储器和处理器，其特征在于，所述存储器中存储有计算机程序，所述处理器被设置为通过所述计算机程序执行所述权利要求1至7任一项中所述的方法。

对象跟踪方法和装置、存储介质及电子装置

技术领域

[0001] 本发明涉及计算机领域，具体而言，涉及一种对象跟踪方法和装置、存储介质及电子装置。

背景技术

[0002] 为了对公共区域实现安全防护，通常会在公共区域安装视频监控系统。通过对上述视频监控系统所监控到的图像进行对象识别跟踪，来对公共区域发生的突发事件实现事前智能预警、事中及时告警、事后高效追溯。

[0003] 目前，在基于传统的视频监控系统中的单摄像头对目标对象进行跟踪的过程中，往往需要获取该单摄像头采集到的多帧图像进行目标检测，然后再对上述目标检测的检测结果进行数据关联。然而，在上述数据关联的过程中，受到目标对象被遮挡的影响，很容易使得跟踪轨迹出现中断，从而导致对目标对象跟踪失败的问题。

[0004] 针对上述的问题，目前尚未提出有效的解决方案。

发明内容

[0005] 本发明实施例提供了一种对象跟踪方法和装置、存储介质及电子装置，以至少解决由于对象被遮挡所导致的跟踪失败的技术问题。

[0006] 根据本发明实施例的一个方面，提供了一种对象跟踪方法，包括：获取目标摄像头当前采集到的目标图像帧，其中，上述目标图像帧中包括待跟踪的目标对象；在确定上述目标对象非当前全局跟踪对象的情况下，利用上述目标摄像头采集到的视频流数据生成的位置关系对数据库，为上述目标对象构建搜索识别树，其中，上述搜索识别树中包括根节点以及T层子节点，上述根节点为在上述目标图像帧中的上述目标对象的目标关键点位置对应的节点，上述T层子节点中的第i层子节点包括在上述目标图像帧之后的第i个图像帧中上述目标对象在执行候选动作后得到的关键点位置对应的节点，上述i小于等于上述T；根据上述搜索识别树中各个子节点的动作值，识别出上述目标对象在上述目标图像帧之后的下一个图像帧中将要执行的目标动作，其中，每个上述子节点的上述动作值用于指示上述子节点与上述根节点之间的相似度。

[0007] 根据本发明实施例的另一方面，还提供了一种对象跟踪装置，包括：第一获取单元，用于获取目标摄像头当前采集到的目标图像帧，其中，上述目标图像帧中包括待跟踪的目标对象；构建单元，用于在确定上述目标对象非当前全局跟踪对象的情况下，利用上述目标摄像头采集到的视频流数据生成的位置关系对数据库，为上述目标对象构建搜索识别树，其中，上述搜索识别树中包括根节点以及T层子节点，上述根节点为在上述目标图像帧中的上述目标对象的目标关键点位置对应的节点，上述T层子节点中的第i层子节点包括在上述目标图像帧之后的第i个图像帧中上述目标对象在执行候选动作后得到的关键点位置对应的节点，上述i小于等于上述T；识别单元，用于根据上述搜索识别树中各个子节点的动作值，识别出上述目标对象在上述目标图像帧之后的下一个图像帧中将要执行的目标动

作，其中，每个上述子节点的上述动作值用于指示上述子节点与上述根节点之间的相似度。

[0008] 根据本发明实施例的又一方面，还提供了一种存储介质，该存储介质中存储有计算机程序，其中，该计算机程序被设置为运行时执行上述对象跟踪方法。

[0009] 根据本发明实施例的又一方面，还提供了一种电子装置，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其中，上述处理器通过计算机程序执行上述的对象跟踪方法。

[0010] 在本发明实施例中，在本实施例中，在获取目标摄像头当前采集到的目标图像帧之后，在确定该目标图像帧中包括的待跟踪的目标对象非当前全局跟踪对象的情况下，利用上述目标摄像头采集到的视频流数据生成的位置关系对数据库，为上述目标对象构建搜索识别树，并根据该搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作。也就是说，基于多帧目标图像帧构建的搜索识别树，可以实现根据未来多帧图像帧中对象的动作值，来预测当前被遮挡的对象的动作，从而实现对短时遮挡的对象进行识别，以保证对象跟踪的准确性。进而解决了由于对象被遮挡所导致的跟踪失败的技术问题。

附图说明

[0011] 此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。在附图中：

[0012] 图1是根据本发明实施例的一种可选的对象跟踪方法的网络环境的示意图；

[0013] 图2是根据本发明实施例的一种可选的对象跟踪方法的流程图；

[0014] 图3是根据本发明实施例的一种可选的对象跟踪方法的示意图；

[0015] 图4是根据本发明实施例的另一种可选的对象跟踪方法的示意图；

[0016] 图5是根据本发明实施例的一种可选的对象跟踪方法的构建示意图；

[0017] 图6是根据本发明实施例的另一种可选的对象跟踪方法的构建示意图；

[0018] 图7是根据本发明实施例的又一种可选的对象跟踪方法的构建示意图；

[0019] 图8是根据本发明实施例的又一种可选的对象跟踪方法的构建示意图；

[0020] 图9是根据本发明实施例的又一种可选的对象跟踪方法的构建示意图；

[0021] 图10是根据本发明实施例的一种可选的对象跟踪装置的结构示意图；

[0022] 图11是根据本发明实施例的一种可选的电子装置的结构示意图。

具体实施方式

[0023] 为了使本技术领域的人员更好地理解本发明方案，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分的实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本发明保护的范围。

[0024] 需要说明的是，本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本发明的实施例能够以除了在这里图示或

描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0025] 根据本发明实施例的一个方面，提供了一种对象跟踪方法，可选地，作为一种可选的实施方式，上述对象跟踪方法可以但不限于应用于如图1所示的对象跟踪系统所在的网络环境中。该对象跟踪系统可以包括但不限于：图像采集设备(如摄像头，下文也可称作目标摄像头)102、网络104、用户设备106及服务器108。其中，上述图像采集设备102用于采集指定区域的图像帧，以实现对该区域内出现的对象进行监控跟踪。上述用户设备106中包括人机交互屏幕1062，处理器1064，存储器1066。人机交互屏幕1062用于显示图像采集设备102采集到的图像帧，还用于获取对该图像帧执行的人机交互操作；处理器1064用于响应上述人机交互操作确定待跟踪的目标对象；存储器1066用于存储上述图像帧。服务器108包括：处理模块1082及数据库1084。其中，处理模块1082用于获取一个图像采集设备采集的图像帧，并对该图像帧进行特征提取以识别其中的目标对象，在该目标对象非当前全局跟踪对象的情况下，利用位置关系对数据库，为目标对象构建搜索识别树。并基于该搜索识别数据中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作，从而实现克服相关技术中由于无法识别出图像帧中被遮挡的对象，导致对目标对象跟踪失败的问题。

[0026] 具体过程如以下步骤：如步骤S102，图像采集设备102将采集到的目标图像帧通过网络104发送给服务器108，服务器108将把上述目标图像帧存储到数据库1084中。

[0027] 进一步，服务器108中的处理模块1082将执行步骤S104-S108：获取目标摄像头当前采集到的目标图像帧，并在确定该目标图像帧中的目标对象并非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，并基于该搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作，进而便于准确地确定出目标对象的跟踪轨迹，以避免由于目标对象被其他物体短时遮挡导致跟踪失败的情况。

[0028] 然后如步骤S110-S112，服务器108将上述跟踪轨迹通过网络104发送给用户设备106，并在该用户设备106中显示上述目标对象的跟踪轨迹。

[0029] 需要说明的是，在本实施例中，在获取目标摄像头当前采集到的目标图像帧之后，在确定该目标图像帧中包括的待跟踪的目标对象非当前全局跟踪对象的情况下，利用上述目标摄像头采集到的视频流数据生成的位置关系对数据库，为上述目标对象构建搜索识别树，并根据该搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作。也就是说，基于多帧目标图像帧构建的搜索识别树，可以实现根据未来多帧图像帧中对象的动作值，来预测当前被遮挡的对象的动作，从而实现对短时遮挡的对象进行识别，以克服相关技术中由于图像帧中的对象被遮挡而导致的跟踪失败的问题，进而保证对象跟踪的准确性。

[0030] 可选地，在本实施例中，上述用户设备可以但不限于为手机、平板电脑、笔记本电脑、个人计算机(Personal Computer，简称PC)等支持运行应用客户端的终端设备。上述服务器与用户设备可以但不限于通过网络实现数据交互，上述网络可以包括但不限于无线网

络或有线网络。其中，该无线网络包括：蓝牙、WIFI及其他实现无线通信的网络。上述有线网络可以包括但不限于：广域网、城域网、局域网。上述仅是一种示例，本实施例中对此不作任何限定。

[0031] 可选地，作为一种可选的实施方式，如图2所示，上述对象跟踪方法包括：

[0032] S202，获取目标摄像头当前采集到的目标图像帧，其中，目标图像帧中包括待跟踪的目标对象；

[0033] S204，在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，其中，搜索识别树中包括根节点以及T层子节点，根节点为在目标图像帧中的目标对象的目标关键点位置对应的节点，T层子节点中的第i层子节点包括在目标图像帧之后的第i个图像帧中目标对象在执行候选动作后得到的关键点位置对应的节点，i小于等于T；

[0034] S206，根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作，其中，每个子节点的动作值用于指示子节点与根节点之间的相似度。

[0035] 可选地，在本实施例中，上述对象跟踪方法可以但不限于应用于对象监控平台应用，该对象监控平台应用可以但不限于是基于在建筑内安装的至少两个图像采集设备所采集到的图像，对选定的至少一个目标对象进行实时跟踪定位的平台应用。其中，上述图像采集设备可以但不限于为安装在建筑内的摄像头，如独立的摄像头或其他配置有摄像头的物联网设备等。上述摄像头可以包括但不限于以下至少之一：RGB摄像头，深度摄像头以及激光雷达等其他传感器。上述建筑可以但不限于配置有基于建筑信息模型(Building Information Modeling，简称BIM)构建的地图，如电子地图，在该电子地图中将标记显示物联网中的各个物联网设备所在位置，如上述摄像头所在位置。如图3所示，安装在用户终端上的对象监控平台应用APP-1，用于呈现图像采集设备(下文也可称作摄像头)的设备标识为ID-1的单个摄像头采集到的视频流数据。如呈现采集时间在17:00-18:00之间的图像帧，例如图像帧A为当前采集到的目标图像帧，其中包括待跟踪的对象301。

[0036] 此外，在本实施例中，上述目标对象可以但不限于为图像中识别出的移动对象，如待监控的人或车辆。上述图像帧可以为图像采集设备按照预定周期采集到离散图像中的图像，也可以为图像采集设备实时录制的视频中的图像，也就是说，本实施例中的图像来源可以是图像集合，也可以是视频中的图像帧。本实施例中对此不作限定。此外，上述搜索识别树可以包括但不限于：蒙特卡洛树。是一种应用于人工智能问题中做出最优决策的搜索工具，一般是在组合博弈中的行动规划形式，其结合了随机模拟的一般性和树搜索的准确性。上述仅是示例，本实施例中对此不作任何限定。

[0037] 需要说明的是，上述图2所示对象跟踪方法可以但不限于用于图1所示的服务器108中。在服务器108获取单个图像采集设备(下文也可称作摄像头)102返回的图像帧之后，可基于该单个摄像头采集到的视频流数据生成的位置关系对，构建搜索识别树，以便于基于该搜索识别树来对遮挡的对象所执行的动作进行预测识别，从而避免相关技术中短时遮挡的对象容易跟丢的问题。

[0038] 例如，如图4所示为服务器针对单个摄像头(如摄像头1至摄像头3中的任意一个摄像头)实现的对象跟踪方法。假设目标摄像头为图4所示摄像头1，在摄像头1采集到视频流

数据1之后，将从该视频流数据1中获取当前采集到的目标图像帧，对该目标图像帧进行运动检测，在检测到移动的对象的情况下，将其存入图像处理器(Graphic Processing Controller，简称GPU)池，以便于对该GPU池中的数据进行目标检测和关键点检测，以获取检测结果1。

[0039] 可选地，在本实施例中，上述运动检测可以包括但不限于：采用帧间差、高斯混合模型以及核密度估计等运动检测方法，得到运动前景图像。例如，上述运动检测过程可以包括但不限于：统计上述运动前景图像的像素个数，当该像素个数小于阈值时，即确定运动区域过小，认为没有运动目标，快速返回；而当像素个数超过阈值时，则认为有运动目标，并将所对应的图像帧压入GPU池中的图像队列。

[0040] 可选地，在本实施例中，上述目标检测和关键点位置检测之后，对检测结果和当前跟踪结果进行融合，可以但不限于进行以下相似度计算，如位置相似度和外观相似度。例如，当相似度小于阈值时，则认为摄像头所监控区域进入了新目标，分配新的跟踪对象ID，并为其配置蒙特卡洛树。然后，基于上述蒙特卡洛树，对该跟踪对象ID采用蒙特卡洛树搜索算法进行搜索，以得到与该对象ID对应的跟踪结果。其中，关键点位置可以但不限于为用于确定目标对象一组关键点所在的位置，如以目标对象为目标人物为例，上述关键点位置可以但不限于为用于确定该目标人物在执行当前动作时的16个关键点所在位置。

[0041] 仍以图4所示示例进行说明，假设目标图像帧为第t帧图像帧，通过处理模块实现上述处理过程，得到检测结果(如检测结果1)之后，将融合第t帧之后的未来T帧图像帧基于蒙特卡洛树搜索得到的跟踪结果，来确定当前第t帧的跟踪结果。从而实现即使在第t帧中的对象被短时遮挡的情况下，也可以通过搜索未来T帧图像中检测出的该对象的动作，来被遮挡的对象的动作，进而克服相关技术中由于对象被遮挡所导致的跟踪失败的问题。

[0042] 可选地，在本实施例中，上述搜索识别树可以但不限于是基于目标摄像头采集到的视频流数据得到的位置关系对数据库来构建的蒙特卡洛树。也就是说，在确定为目标图像帧构建搜索识别树(即蒙特卡洛树)的情况下，将基于该目标图像帧中目标对象的动作确定该蒙特卡洛树中的根节点，进一步再根据视频流数据中该目标图像帧之后的T帧图像帧内目标对象的动作集合，确定该蒙特卡洛树中对应的子节点。

[0043] 需要说明的是，上述子节点的构造是从上述动作集合中采样得到的，该动作集合需要有效覆盖目标对象下一步的动作，由于其下一步动作受到摄像头角度的变化而变化，因而具有较大的搜索空间。

[0044] 在本实施例中，可以对上述目标摄像头采集输入的视频流数据的全部图像帧进行运动检测、目标检测和关键点位置检测。然后可以但不限于通过以下方法获取位置关系对数据库：通过二分图匹配算法对上下相邻两帧检测到的对象进行位置匹配，以确定两帧中检测到的对象的关联关系，并建立关键点位置关系对，如(Oi(t-1)，Oj(t))，其中，Oi(t-1)用于表示在第t-1帧检测到的第i个对象的关键点位置，Oj(t)用于表示第t帧检测到的第j个对象的关键点位置，上述关键点位置关系对用于表示上述两个对象为同一个对象，使用相同的对象ID。以此类推，按照关键点位置关系对的关系构建位置关系对数据库，以反映上下相邻两个图像帧中同一个对象的动作预测。从而实现在给定既定对象在上一帧(如第t-1帧)中的关键点位置的情况下，可以通过对位置关系对数据库的检索，得到该既定对象在下一帧(如第t帧)中可能出现的A个关键点位置。进一步，利用上述A个关键点位置生成目标对

象在蒙特卡洛树中第t层的子节点。上述仅是示例，本实施例中对此不做任何限定。

[0045] 进一步，在本实施例中，在生成与第t帧对应的第t层子节点的过程中，还可以包括但不限于：获取这一层子节点中每个对象子节点的动作值。其中，该动作值可以但不限于为根节点与该对象子节点之间的对象相似度。其中，上述对象相似度可以但不限于为根节点与对象子节点对应的关键点位置之间的位置相似度，和关键点位置所指示的对象之间的外观相似度的融合结果。上述仅是一种示例，本实施例中对此不做限定。

[0046] 可选地，在本实施例中，在确定完成第t帧对应的第t层子节点的生成之后，还可以但不限于更新第t层之前的全部节点的动作值。例如，基于第t层的全部子节点的动作值，更新其所属父节点(即第t-1层中的节点)的动作值；进一步，在确定上述父节点非根节点的情况下，基于该父节点所在第t-1层中的全部子节点的动作值，来更新其上一层父节点(如第t-2层中的节点)的动作值，以此类推，直至父节点为根节点时，确定该遍历当前已生成的搜索识别树(即蒙特卡洛树)中的全部节点。上述仅是一种通过节点回溯的方式，来更新动作值的方式，本实施例中对此不做任何限定。

[0047] 具体结合以下示例进行说明：假设本示例是针对单个摄像头采集到的视频流数据所实现的多目标对象跟踪过程，其中为待跟踪的目标对象构建的搜索识别树为蒙特卡洛树。其中，在本示例中。上述蒙特卡洛树可以但不限于将上述目标对象当作智能体，并用目标对象所在目标框的位置以及目标对象的关键点在图像帧中的位置来描述智能体的状态，智能体的状态为当前单摄像头下多目标跟踪的结果。其中，上述智能体可以通过从一系列的有限集合的动作中选择一个动作来改变自己的状态，并计算当前改变后状态与图像变化的吻合度，通过连续决策动作确定该目标对象在蒙特卡洛树中对应的路径，并评估当前路径的最终置信度。通过蒙特卡洛树可以模拟多条路径决策，并从中选择最优的当前决策。由于该方法在长时间尺度上考虑了图像上下文信息，因此能够解决目标对象被遮挡时出现的跟踪丢失的问题。

[0048] 需要说明的是，在本实施例中，上述吻合度可以但不限于是在基于位置关系对数据库确定出A个关键点位置之后，将该A个关键点位置作为目标对象在下一帧中所要执行的A个预测动作对应的关键点位置。从下一帧图像帧中分别确定出上述A个关键点位置对应的图像特征，进一步，通过比对该图像特征与目标图像帧中目标对象对应关键点位置的图像特征，来确定在生成蒙特卡洛树时该A个关键位置分别对应的动作值。

[0049] 通过本申请提供的实施例，在获取目标摄像头当前采集到的目标图像帧之后，在确定该目标图像帧中包括的待跟踪的目标对象非当前全局跟踪对象的情况下，利用上述目标摄像头采集到的视频流数据生成的位置关系对数据库，为上述目标对象构建搜索识别树，并根据该搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作。也就是说，基于多帧目标图像帧构建的搜索识别树，可以实现根据未来多帧图像帧中对象的动作值，来预测当前被遮挡的对象的动作，从而实现对短时遮挡的对象进行识别，以保证对象跟踪的准确性。

[0050] 作为一种可选的方案，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树包括：

[0051] S1，获取目标摄像头在目标图像帧之后采集到的T个图像帧；

[0052] S2，对T个图像帧依次执行以下操作；

[0053] S21,确定与第i个图像帧对应的待生成的第i层子节点所属的当前父节点;

[0054] S22,从位置关系对数据库中查找与当前父节点相关联的目标位置关系对集合,其中,目标位置关系对集合中的每个目标位置关系对中包括的第一关键点位置与当前父节点对应的关键点位置的相似度大于第一阈值,位置关系对数据库中的每个位置关系对用于记录视频流数据中采集时间相邻的两个图像帧中同一个对象执行候选动作前后的关键点位置;

[0055] S23,获取每个目标位置关系对中与第一关键点位置匹配的第二关键点位置,其中,第一关键点位置所在图像帧位于第二关键点位置所在图像帧之前;

[0056] S24,利用第二关键点位置生成第i层子节点;

[0057] S25,在第i层子节点未达到第T层子节点的情况下,获取第i+1个图像帧作为第i个图像帧。

[0058] 需要说明的是,在本实施例中,在为目标对象生成对应的搜索识别树(即蒙特卡洛树)中,该树的根节点可以但不限于用于记录目标对象在目标图像帧中所执行动作对应的关键点位置,该树的子节点可以但不限于用于记录基于动作集合,为该目标图像帧之后的每个图像帧预测的A个关键点位置。其中,上述动作集合可以但不限于为基于目标摄像头采集到的视频流数据构建的位置关系对数据库中各个关键点位置对应的动作。

[0059] 如图5所示,假设目标图像帧为目标摄像头采集到的视频流数据中的第t帧图像帧。在确定该第t帧图像帧中检测出的待跟踪的目标对象(如图所示人物对象)非全局跟踪对象的情况下,则利用在第t帧图像帧之后的T个图像帧为其构建蒙特卡洛树,如图5所示在该蒙特卡洛树中至少包括:根节点为节点a,子节点包括节点b至节点g。以第t+2帧为例,第t+2帧为T个图像帧中第2个图像帧,对应的确定第2层子节点,其具体构建过程可以如下:

[0060] 第2层子节点包括节点d至节点g,其中,节点d和节点e所属的当前父节点为第1层子节点中的节点b,节点f和节点g所属的当前父节点为第1层子节点中的节点c。以节点d和节点e为例,从位置关系对数据库中查找与节点b相关联的目标位置关系对集合,其中,目标位置关系对集合中的每个目标位置关系对中的第一关键点位置与上述节点b对应的关键点位置的相似度大于第一阈值。获取上述目标位置关系对集合中每个目标位置关系对中与第一关键点位置对应的第二关键点位置,以根据该第二关键点位置来预测节点b执行后可能执行的动作,其中,第一关键点位置与第二关键点位置为视频流数据中相邻上下两个图像帧中所指示的动作对应的关键点位置,第一关键点位置对应的图像帧在第二关键点位置对应的图像帧之前。

[0061] 进一步,假设与上述节点b对应的关键点位置的相似度大于第一阈值的第一关键点位置包括两个,对应的,将确定出目标关系对集合中包括两个目标关系对,根据两个第二关键点位置来确定出执行两个预测动作,并基于上述两个预测动作执行后对应的关键点位置生成子节点,即节点d和节点e,加入节点b之后。以此类推,可得到节点f和节点g,直到达到预设蒙特卡洛树的深度T。从而实现保证蒙特卡洛树的每一层子节点对应一帧图像帧,以缓存第t帧图像帧之后未来的T个图像帧,进而通过对上述未来的图像帧来确定当前的第t帧图像中目标对象所要执行目标动作或目标动作对应的关键位置等信息。

[0062] 通过本申请提供的实施例,在构建搜索识别树中的子节点的过程中,在确定与第i个图像帧对应的第i层子节点时,可以从位置关系对数据库中查找与第i层子节点所属的父

节点关联的目标位置关系对集合，进一步利用目标位置关系对集合中与第一关键点位置匹配的第二关键位置来预测目标对象在执行父节点对应的动作后可能执行的动作，从而实现对准确而快速地建立与该目标对象对应的蒙特卡洛树，以便于后续搜索确定目标对象在被遮挡时对应的动作。

[0063] 作为一种可选的方案，利用第二关键点位置生成第i层子节点包括：

[0064] S1，依次执行以下操作，直至得到第i层子节点：

[0065] S11，为当前的第二关键点位置生成当前对象子节点；

[0066] S12，从第i个图像帧中提取与当前的第二关键点位置对应的第一图像特征；

[0067] S13，比对第一图像特征和从目标图像帧中提取出的与目标关键点位置对应的第二图像特征，得到当前对象子节点与根节点之间的位置相似度和外观相似度；

[0068] S14，根据位置相似度和外观相似度确定当前对象子节点与根节点之间的对象相似度；

[0069] S15，将对象相似度作为当前对象子节点的动作值，并确定完成当前对象子节点的生成；

[0070] S16，获取下一个第二关键点位置作为当前的第二关键点位置。

[0071] 可选地，在本实施例中，上述位置相似度可以但不限于为关键点位置的位置相似度，确定方式可以如下：从第i个图像帧中提取当前的关键点位置(对应当前对象子节点)周围区域提取图像特征，从目标图像帧中关键点位置(对应根节点)周围区域提取图像特征，然后计算根节点和当前对象子节点之间的重叠系数，以作为二者之间的位置相似度。

[0072] 此外，在本实施例中，上述外观相似度可以但不限于为关键点位置对应的动作所指示的对象的外观相似度。确定方式可以如下：从第i个图像帧中提取当前的关键点位置(对应当前对象子节点)周围区域提取图像特征，从目标图像帧中关键点位置(对应根节点)周围区域提取图像特征，然后通过点积的方式计算外观相似度。

[0073] 如图5所示，假设目标图像帧为目标摄像头采集到的视频流数据中的第t帧图像帧。在确定该第t帧图像帧中检测出的待跟踪的目标对象(如图所示人物对象)非全局跟踪对象的情况下，则利用在第t帧图像帧之后的T个图像帧为其构建蒙特卡洛树，如图5所示在该蒙特卡洛树中至少包括：根节点为节点a，子节点包括节点b至节点g。

[0074] 进一步，如图6所示，假设当前第i层子节点(如图6所示为第t+3帧对应的子节点)的父节点为节点d。根据上述实施例确定当前第i层子节点：新节点j和新节点k。以新节点j为例，其动作值确定过程可以如下：

[0075] 在确定新节点j对应预测动作J的关键点位置之后，从第i层子节点对应的第i个图像帧(如图6所示为第t+3帧)中，提取上述预测动作J的关键点位置对应的第一图像特征，并获取目标图像帧中提取出的与目标关键点位置对应的第二图像特征，通过比对上述第一图像特征和第二图像特征，来获取重叠系数以确定新节点j和根节点a之间的位置相似度α1，并通过比对上述第一图像特征和第二图像特征，通过点积计算以确定新节点j和根节点a之间的外观相似度α2。

[0076] 进一步，对上述位置相似度α1和外观相似度α2进行加权求和，将得到该新节点j与根节点a之间的对象相似度α，将其作为新节点j的动作值。以此类推，计算得到新节点k的动作值，如图6所示带箭头的加粗线用于表示根节点a和新节点j之间的对象相似度(即新节点

j的动作值)，以及根节点a和新节点k之间的对象相似度(即新节点k的动作值)。在计算得到动作值后，确定完成对蒙特卡洛树中第i层子节点的新节点j和新节点k的生成过程。

[0077] 通过本申请提供的实施例，可以利用动作值来表示各个子节点与根节点之间的相似度，其中，上述相似度为基于图像帧提取的关键点位置周围区域的图像特征获取的位置相似度和外观相似度来确定。上述动作值客观地反映了子节点与根节点的吻合度，从而便于后续基于动作值来准确预测目标对象可能执行的动作，进而保证对该目标对象进行跟踪的轨迹的准确性，避免跟丢情况的发生。

[0078] 作为一种可选的方案，在利用第二关键点位置生成第i层子节点之后，还包括：

[0079] S1，利用第i层子节点的全部动作值，更新当前父节点的动作值；

[0080] S2，在当前父节点非根节点的情况下，确定当前父节点所属的上一层父节点；利用与当前父节点属于同一层的子节点的全部动作值，更新上一层父节点的动作值。

[0081] 可选地，在本实施例中，在确定第i层子节点的动作值之后，还将对其上层第i-1层子节点的动作值进行回溯更新，以此类推，直至到根节点。其中，该回溯更新的方式可以包括但不限于：获取第i-1层的既定父节点的访问次数，并获取从属于该既定父节点的第i层子节点的动作值的求和结果，然后根据上述求和结果与访问次数的比值，来更新上述父节点的动作值。以通过回溯更新的方式，使得上层的父节点的动作值可以更精准地反映出与根节点之间的吻合度。

[0082] 如图7-8所示，假设目标图像帧为目标摄像头采集到的视频流数据中的第t帧图像帧。在确定该第t帧图像帧中检测出的待跟踪的目标对象(如图所示人物对象)非全局跟踪对象的情况下，则利用在第t帧图像帧之后的T个图像帧为其构建蒙特卡洛树，如图7所示在该蒙特卡洛树中至少包括：根节点为节点a，子节点包括节点b至节点g，和节点j至节点k。

[0083] 在确定生成每一层子节点之后，都将对当前生成的蒙特卡洛树中的各个子节点的动作值进行回溯更新。如图7所示，假设在生成与第t+1帧对应一层的子节点之后，对根节点a进行动作选择时，可以选择节点b对应的动作B；在生成与第t+2帧对应一层的子节点之后，对节点b进行动作选择时，选择节点d对应的动作D；在生成与第t+3帧对应一层的子节点之后，对节点d进行动作选择时，选择节点k对应的动作K。

[0084] 如图8所示，在根据上述实施例确定完成对蒙特卡洛树中第i层子节点的新节点j和新节点k的生成过程之后，可以基于新节点j和新节点k的动作值进行加权求和得到的求和结果以及节点d的访问次数，确定节点d更新后的动作值。以此类推，更新节点b的动作值，直至上一层父节点为根节点a。

[0085] 通过本申请提供的实施例，通过从属于父节点的子节点的动作值，来对父节点的动作值进行进一步更新，从而实现将下一步预测的动作对应的动作值反映在父节点中，也就是说将预测的可能性体现在父节点中，以便于在蒙特卡洛树中从上向下进行动作决策时，保证决策的准确性，进而达到即使目标对象出现短时遮挡的情况下，也可以基于上述路径确定跟踪轨迹的目的，避免跟踪失败的情况。

[0086] 作为一种可选的方案，在利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树之前，还包括：

[0087] S1，获取视频流数据中采集时间相邻的两个图像帧，其中，两个图像帧包括第一图像帧和第二图像帧；

[0088] S2,在从第一图像帧中识别出第一候选对象,并从第二图像帧中识别出第二候选对象的情况下,比对第一候选对象的关键点位置与第二候选对象的关键点位置;

[0089] S3,在比对的结果指示第一候选对象与第二候选对象关联的情况下,确定第一候选对象与第二候选对象为同一个对象,并将第一候选对象的关键点位置与第二候选对象的关键点位置记录为位置关系对数据库中的一个位置关系对。

[0090] 具体结合以下示例进行说明:对视频流数据中的连续两帧图像帧进行运动检测、目标检测和关键点位置检测,以识别出两帧图像帧中的候选对象。

[0091] 例如,获取采集时间相邻的两个图像帧分别为第t帧图像帧和第t+1帧图像帧。然后从上述第t帧图像帧中识别出第一候选对象,从第t帧图像帧中识别出第二候选对象,对上述两个候选对象的关键点位置进行比对。如通过二分图匹配算法对上述两个候选对象进行关键点位置匹配,以确定关联关系。若确定二者关联的情况下,表示第一候选对象与第二候选对象为同一个对象。则进一步可以通过以下方式构建二者之间的位置关系对:(Oi(t),Oj(t+1))。其中,Oi(t)用于表示上述对象为第t帧检测到的第i个对象的关键点位置,Oj(t)用于表示上述对象为第+1帧检测到的第j个对象的关键点位置。

[0092] 通过本申请提供的实施例,对视频流数据进行识别分类处理,以确定连续两帧图像帧中存在关联关系的对象,并据此建立该视频流数据对应的位置关系对数据库,从而实现基于该位置关系对数据库来获取预测的动作,进而保证所构建的搜索识别树的真实性和准确性。

[0093] 作为一种可选的方案,根据搜索识别树中各个子节点的动作值,识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作包括:

[0094] S1,在目标图像帧之后的下一个图像帧对应的一层子节点中,确定出最大的动作值对应的目标子节点;

[0095] S2,将目标子节点对应的关键点位置所指示的动作作为目标动作。

[0096] 如图9所示,假设目标图像帧为目标摄像头采集到的视频流数据中的第t帧图像帧。在确定该第t帧图像帧中检测出的待跟踪的目标对象(如图所示人物对象)非全局跟踪对象的情况下,则利用在第t帧图像帧之后的T个图像帧为其构建蒙特卡洛树,如图9所示在该蒙特卡洛树中至少包括:根节点为节点a,子节点包括节点b至节点g,和节点j至节点k。

[0097] 进一步,假设根据上述实施例确定完成对蒙特卡洛树中第i层子节点的新节点j和新节点k的生成过程之后,还可以对上述蒙特卡洛树中各个子节点的动作值进行回溯更新。

[0098] 然后,如图9所示,可以基于更新后的动作值来进行决策,如在第t+1帧对应的子节点中,从更新后的动作值中,决策选择动作值最大的节点b对应的动作B作为目标动作,进一步在第t+2帧对应的子节点中,从更新后的动作值中,决策选择动作值最大的节点e对应的动作E作为目标动作,在第t+3帧对应的子节点中,从更新后的动作值中,决策选择动作值最大的节点k对应的动作K作为目标动作。

[0099] 通过本申请提供的实施例,基于更新后的动作值进行决策,选择动作值最大的节点对应的动作作为目标动作,从而实现将预测的可能性体现在父节点中,以便于在蒙特卡洛树中从上向下进行动作决策时,保证决策的准确性,进而达到即使目标对象出现短时遮挡的情况下,也可以基于上述路径确定跟踪轨迹的目的,避免跟踪失败的情况。

[0100] 作为一种可选的方案,在获取目标摄像头当前采集到的目标图像帧之后,还包括:

[0101] S1,对目标图像帧进行运动检测;

[0102] S2,在检测到目标图像帧中包括移动对象的情况下,对移动对象进行目标检测得到第一检测结果,并对移动对象进行关键点位置检测得到第二检测结果;

[0103] S3,在第一检测结果和第二检测结果指示移动对象为目标对象的情况下,比对目标对象与当前全局跟踪对象;

[0104] S4,在比对的结果指示目标对象与当前全局跟踪对象未匹配的情况下,为目标对象分配新的对象标识,并确定为目标对象构建搜索识别树。

[0105] 可选地,在本实施例中,上述运动检测可以包括但不限于:采用帧间差、高斯混合模型以及核密度估计等运动检测方法,得到运动前景图像。例如,上述运动检测过程可以包括但不限于:统计上述运动前景图像的像素个数,当该像素个数小于阈值时,即确定运动区域过小,认为没有运动目标,快速返回;而当像素个数超过阈值时,则认为有运动目标,并将所对应的图像帧压入GPU池中的图像队列。

[0106] 可选地,在本实施例中,上述目标检测和关键点位置检测之后,对检测结果和当前跟踪结果进行融合,可以但不限于进行以下相似度计算,如位置相似度和外观相似度。例如,当相似度小于阈值时,则认为摄像头所监控区域进入了新目标,分配新的跟踪对象ID,并为其配置蒙特卡洛树。然后,基于上述蒙特卡洛树,对该跟踪对象ID采用蒙特卡洛树搜索算法进行搜索,以得到与该对象ID对应的跟踪结果。其中,关键点位置可以但不限于为用于确定目标对象一组关键点所在的位置,如以目标对象为目标人物为例,上述关键点位置可以但不限于为用于确定该目标人物在执行当前动作时的16个关键点所在位置。

[0107] 通过本申请提供的实施例,通过运动检测、目标检测和关键点检测,来确定视频流数据中包括目标对象的图像帧,以便于快速基于上述图像帧来进行搜索定位,从而达到提高确定目标对象的跟踪轨迹的效率的效果。

[0108] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0109] 根据本发明实施例的另一个方面,还提供了一种用于实施上述对象跟踪方法的对象跟踪装置。如图10所示,该装置包括:

[0110] 1)第一获取单元1002,用于获取目标摄像头当前采集到的目标图像帧,其中,目标图像帧中包括待跟踪的目标对象;

[0111] 2)构建单元1004,用于在确定目标对象非当前全局跟踪对象的情况下,利用目标摄像头采集到的视频流数据生成的位置关系对数据库,为目标对象构建搜索识别树,其中,搜索识别树中包括根节点以及T层子节点,根节点为在目标图像帧中的目标对象的目标关键点位置对应的节点,T层子节点中的第i层子节点包括在目标图像帧之后的第i个图像帧中目标对象在执行候选动作后得到的关键点位置对应的节点,i小于等于T;

[0112] 3)识别单元1006,用于根据搜索识别树中各个子节点的动作值,识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作,其中,每个子节点的动作值用于指示子节点与根节点之间的相似度。

[0113] 可选地，在本实施例中，上述对象跟踪方法可以但不限于应用于对象监控平台应用，该对象监控平台应用可以但不限于是基于在建筑内安装的至少两个图像采集设备所采集到的图像，对选定的至少一个目标对象进行实时跟踪定位的平台应用。其中，上述图像采集设备可以但不限于为安装在建筑内的摄像头，如独立的摄像头或其他配置有摄像头的物联网设备等。上述摄像头可以包括但不限于以下至少之一：RGB摄像头，深度摄像头以及激光雷达等其他传感器。上述建筑可以但不限于配置有基于建筑信息模型(Building Information Modeling，简称BIM)构建的地图，如电子地图，在该电子地图中将标记显示物联网中的各个物联网设备所在位置，如上述摄像头所在位置。如图3所示，安装在用户终端上的对象监控平台应用APP-1，用于呈现图像采集设备(下文也可称作摄像头)的设备标识为ID-1的单个摄像头采集到的视频流数据。如呈现采集时间在17:00-18:00之间的图像帧，例如图像帧A为当前采集到的目标图像帧，其中包括待跟踪的对象301。

[0114] 此外，在本实施例中，上述目标对象可以但不限于为图像中识别出的移动对象，如待监控的人或车辆。上述图像帧可以为图像采集设备按照预定周期采集到离散图像中的图像，也可以为图像采集设备实时录制的视频中的图像，也就是说，本实施例中的图像来源可以是图像集合，也可以是视频中的图像帧。本实施例中对此不作限定。此外，上述搜索识别树可以包括但不限于：蒙特卡洛树。是一种应用于人工智能问题中做出最优决策的搜索工具，一般是在组合博弈中的行动规划形式，其结合了随机模拟的一般性和树搜索的准确性。上述仅是示例，本实施例中对此不作任何限定。

[0115] 需要说明的是，上述图2所示对象跟踪装置可以但不限于用于图1所示的服务器108中。在服务器108获取单个图像采集设备(下文也可称作摄像头)102返回的图像帧之后，可基于该单个摄像头采集到的视频流数据生成的位置关系对，构建搜索识别树，以便于基于该搜索识别树来对遮挡的对象所执行的动作进行预测识别，从而避免相关技术中短时遮挡的对象容易跟丢的问题。

[0116] 作为一种可选的方案，构建单元1004包括：

[0117] 1)第一获取模块，用于获取目标摄像头在目标图像帧之后采集到的T个图像帧；

[0118] 2)处理模块，用于对T个图像帧依次执行以下操作；

[0119] S1，确定与第i个图像帧对应的待生成的第i层子节点所属的当前父节点；

[0120] S2，从位置关系对数据库中查找与当前父节点相关联的目标位置关系对集合，其中，目标位置关系对集合中的每个目标位置关系对中包括的第一关键点位置与当前父节点对应的关键点位置的相似度大于第一阈值，位置关系对数据库中的每个位置关系对用于记录视频流数据中采集时间相邻的两个图像帧中同一个对象执行候选动作前后的关键点位置；

[0121] S3，获取每个目标位置关系对中与第一关键点位置匹配的第二关键点位置，其中，第一关键点位置所在图像帧位于第二关键点位置所在图像帧之前；

[0122] S4，利用第二关键点位置生成第i层子节点；

[0123] S5，在第i层子节点未达到第T层子节点的情况下，获取第i+1个图像帧作为第i个图像帧。

[0124] 具体实施例可以参考上述实施例，这里不再赘述。

[0125] 作为一种可选的方案，处理模块通过以下步骤实现利用第二关键点位置生成第i

层子节点：

[0126] 依次执行以下操作，直至得到第i层子节点：

[0127] S11，为当前的第二关键点位置生成当前对象子节点；

[0128] S12，从第i个图像帧中提取与当前的第二关键点位置对应的第一图像特征；

[0129] S13，比对第一图像特征和从目标图像帧中提取出的与目标关键点位置对应的第二图像特征，得到当前对象子节点与根节点之间的位置相似度和外观相似度；

[0130] S14，根据位置相似度和外观相似度确定当前对象子节点与根节点之间的对象相似度；

[0131] S15，将对象相似度作为当前对象子节点的动作值，并确定完成当前对象子节点的生成；

[0132] S16，获取下一个第二关键点位置作为当前的第二关键点位置。

[0133] 具体实施例可以参考上述实施例，这里不再赘述。

[0134] 作为一种可选的方案，处理模块还用于：

[0135] S1，在利用第二关键点位置生成第i层子节点之后，利用第i层子节点的全部动作值，更新当前父节点的动作值；

[0136] S2，在当前父节点非根节点的情况下，确定当前父节点所属的上一层父节点；利用与当前父节点属于同一层的子节点的全部动作值，更新上一层父节点的动作值。

[0137] 具体实施例可以参考上述实施例，这里不再赘述。

[0138] 作为一种可选的方案，还包括：

[0139] 1)第二获取单元，用于在利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树之前，获取视频流数据中采集时间相邻的两个图像帧，其中，两个图像帧包括第一图像帧和第二图像帧；

[0140] 2)第一比对单元，用于在从第一图像帧中识别出第一候选对象，并从第二图像帧中识别出第二候选对象的情况下，比对第一候选对象的关键点位置与第二候选对象的关键点位置；

[0141] 3)确定单元，用于在比对的结果指示第一候选对象与第二候选对象关联的情况下，确定第一候选对象与第二候选对象为同一个对象，并将第一候选对象的关键点位置与第二候选对象的关键点位置记录为位置关系对数据库中的一个位置关系对。

[0142] 具体实施例可以参考上述实施例，这里不再赘述。

[0143] 作为一种可选的方案，识别单元1006包括：

[0144] 1)第一确定模块，用于在目标图像帧之后的下一个图像帧对应的一层子节点中，确定出最大的动作值对应的目标子节点；

[0145] 2)第二确定模块，用于将目标子节点对应的关键点位置所指示的动作作为目标动作。

[0146] 具体实施例可以参考上述实施例，这里不再赘述。

[0147] 作为一种可选的方案，还包括：

[0148] 1)第一检测单元，用于在获取目标摄像头当前采集到的目标图像帧之后，对目标图像帧进行运动检测；

[0149] 2)第二检测单元，用于在检测到目标图像帧中包括移动对象的情况下，对移动对

象进行目标检测得到第一检测结果，并对移动对象进行关键点位置检测得到第二检测结果；

[0150] 3)第二比对单元，用于在第一检测结果和第二检测结果指示移动对象为目标对象的情况下，比对目标对象与当前全局跟踪对象；

[0151] 4)分配单元，用于在比对的结果指示目标对象与当前全局跟踪对象未匹配的情况下，为目标对象分配新的对象标识，并确定为目标对象构建搜索识别树。

[0152] 具体实施例可以参考上述实施例，这里不再赘述。

[0153] 根据本发明实施例的又一个方面，还提供了一种用于实施上述对象跟踪方法的电子装置，如图11所示，该电子装置包括存储器1102和处理器1104，该存储器1102中存储有计算机程序，该处理器1104被设置为通过计算机程序执行上述任一项方法实施例中的步骤。

[0154] 可选地，在本实施例中，上述电子装置可以位于计算机网络的多个网络设备中的至少一个网络设备。

[0155] 可选地，在本实施例中，上述处理器可以被设置为通过计算机程序执行以下步骤：

[0156] S1，获取目标摄像头当前采集到的目标图像帧，其中，目标图像帧中包括待跟踪的目标对象；

[0157] S2，在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，其中，搜索识别树中包括根节点以及T层子节点，根节点为在目标图像帧中的目标对象的目标关键点位置对应的节点，T层子节点中的第i层子节点包括在目标图像帧之后的第i个图像帧中目标对象在执行候选动作后得到的关键点位置对应的节点，i小于等于T；

[0158] S3，根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作，其中，每个子节点的动作值用于指示子节点与根节点之间的相似度。

[0159] 可选地，本领域普通技术人员可以理解，图11所示的结构仅为示意，电子装置也可以是智能手机(如Android手机、iOS手机等)、平板电脑、掌上电脑以及移动互联网设备(Mobile Internet Devices，MID)、PAD等终端设备。图11其并不对上述电子装置的结构造成限定。例如，电子装置还可包括比图11中所示更多或者更少的组件(如网络接口等)，或者具有与图11所示不同的配置。

[0160] 其中，存储器1102可用于存储软件程序以及模块，如本发明实施例中的对象跟踪方法和装置对应的程序指令/模块，处理器1104通过运行存储在存储器1102内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述的对象跟踪方法。存储器1102可包括高速随机存储器，还可以包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器1102可进一步包括相对于处理器1104远程设置的存储器，这些远程存储器可以通过网络连接至终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。其中，存储器1102具体可以但不限于用于存储目标图像帧、搜索识别树中各个节点的内容等信息。作为一种示例，如图11所示，上述存储器1102中可以但不限于包括上述对象跟踪装置中的第一获取单元1002、构建单元1004及识别单元1006。此外，还可以包括但不限于上述对象跟踪装置中的其他模块单元，本示例中不再赘述。

[0161] 可选地，上述的传输装置1106用于经由一个网络接收或者发送数据。上述的网络具体实例可包括有线网络及无线网络。在一个实例中，传输装置1106包括一个网络适配器(Network Interface Controller,NIC)，其可通过网线与其他网络设备与路由器相连从而可与互联网或局域网进行通讯。在一个实例中，传输装置1106为射频(Radio Frequency，RF)模块，其用于通过无线方式与互联网进行通讯。

[0162] 此外，上述电子装置还包括：显示器1108，用于显示上述目标图像帧等信息；和连接总线1110，用于连接上述电子装置中的各个模块部件。

[0163] 根据本发明的实施例的又一方面，还提供了一种存储介质，该存储介质中存储有计算机程序，其中，该计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

[0164] 可选地，在本实施例中，上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

[0165] S1，获取目标摄像头当前采集到的目标图像帧，其中，目标图像帧中包括待跟踪的目标对象；

[0166] S2，在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，其中，搜索识别树中包括根节点以及T层子节点，根节点为在目标图像帧中的目标对象的目标关键点位置对应的节点，T层子节点中的第i层子节点包括在目标图像帧之后的第i个图像帧中目标对象在执行候选动作后得到的关键点位置对应的节点，i小于等于T；

[0167] S3，根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作，其中，每个子节点的动作值用于指示子节点与根节点之间的相似度。

[0168] 可选地，在本实施例中，本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令终端设备相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：闪存盘、只读存储器(Read-Only Memory，ROM)、随机存取器(Random Access Memory,RAM)、磁盘或光盘等。

[0169] 上述本发明实施例序号仅仅为了描述，不代表实施例的优劣。

[0170] 上述实施例中的集成的单元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时，可以存储在上述计算机可读取的存储介质中。基于这样的理解，本发明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全部或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存储介质中，包括若干指令用以使得一台或多台计算机设备(可为个人计算机、服务器或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。

[0171] 在本发明的上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其他实施例的相关描述。

[0172] 在本申请所提供的几个实施例中，应该理解到，所揭露的客户端，可通过其它的方式实现。其中，以上所描述的装置实施例仅仅是示意性的，例如所述单元的划分，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，例如多个单元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接口，单元或模块的间接耦合或通信连

接,可以是电性或其它的形式。

[0173] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元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。

[0174] 另外,在本发明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。

[0175] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

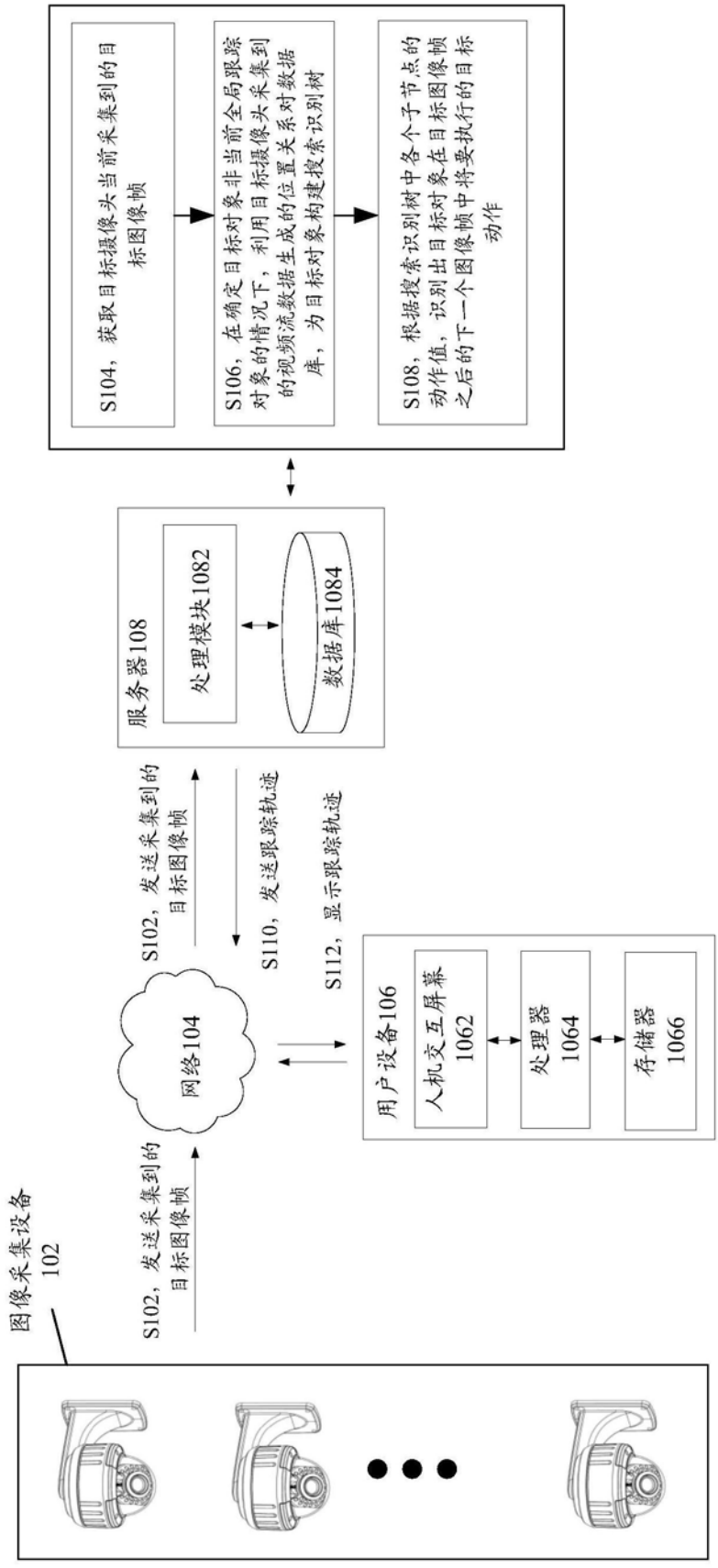

S104，获取目标摄像头当前采集到的目标图像帧
S106，在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树
S108，根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作
服务器108
处理模块1082
数据库1084
S102，发送采集到的目标图像帧
S110，发送跟踪轨迹
S112，显示跟踪轨迹
网络104
用户设备106
人机交互屏幕1062
处理器1064
存储器1066
S102，发送采集到的目标图像帧
图像采集设备102

图1

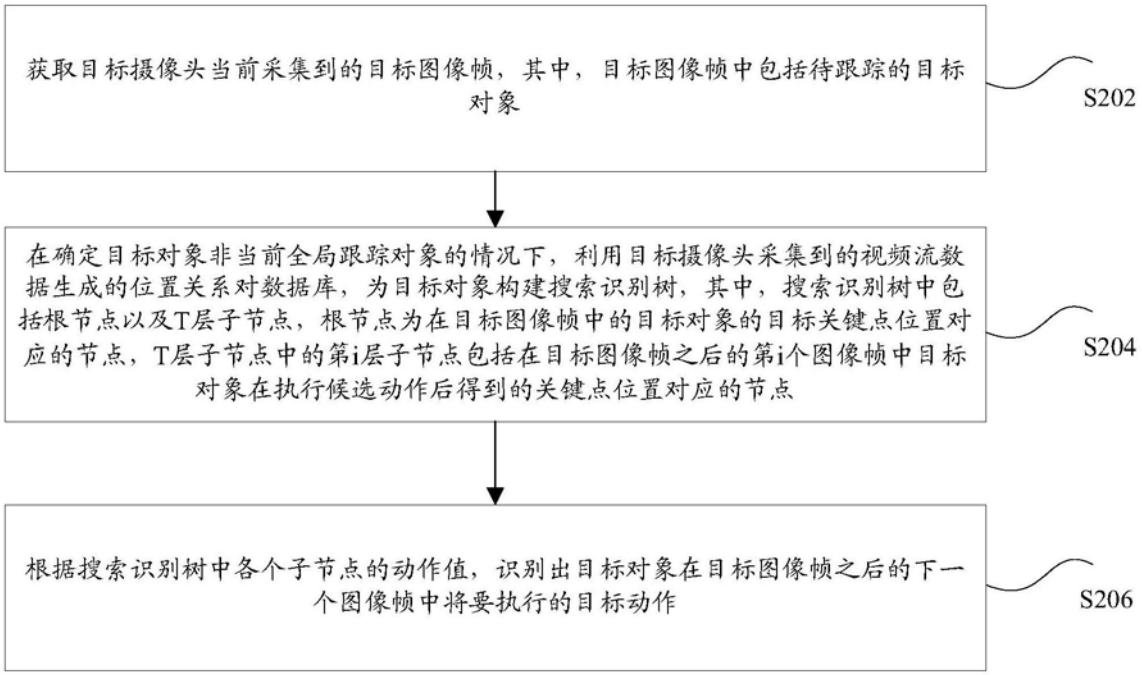

获取目标摄像头当前采集到的目标图像帧，其中，目标图像帧中包括待跟踪的目标对象
S202
在确定目标对象非当前全局跟踪对象的情况下，利用目标摄像头采集到的视频流数据生成的位置关系对数据库，为目标对象构建搜索识别树，其中，搜索识别树中包括根节点以及T层子节点，根节点为在目标图像帧中的目标对象的目标关键点位置对应的节点，T层子节点中的第i层子节点包括在目标图像帧之后的第i个图像帧中目标对象在执行候选动作后得到的关键点位置对应的节点
S204
根据搜索识别树中各个子节点的动作值，识别出目标对象在目标图像帧之后的下一个图像帧中将要执行的目标动作
S206

图2

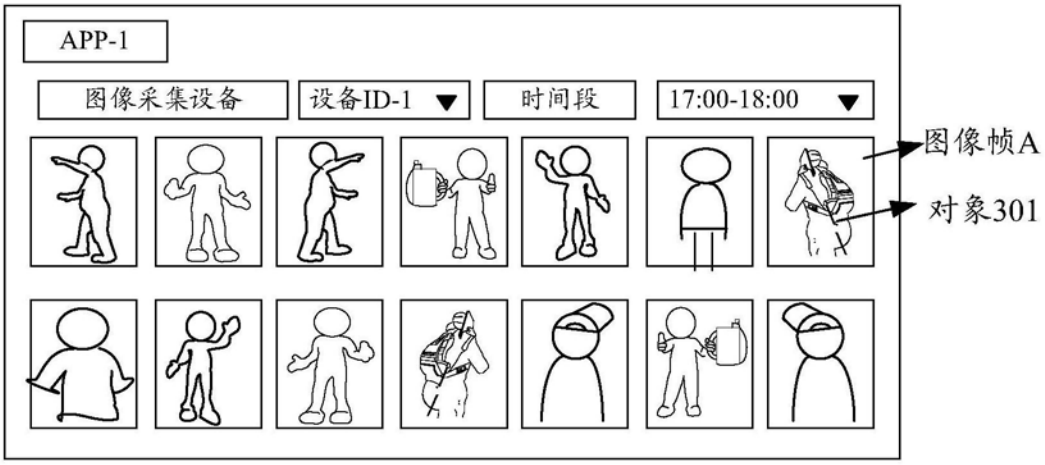

APP-1
图像采集设备
设备ID-1
时间段
17:00-18:00
图像帧A
对象301

图3

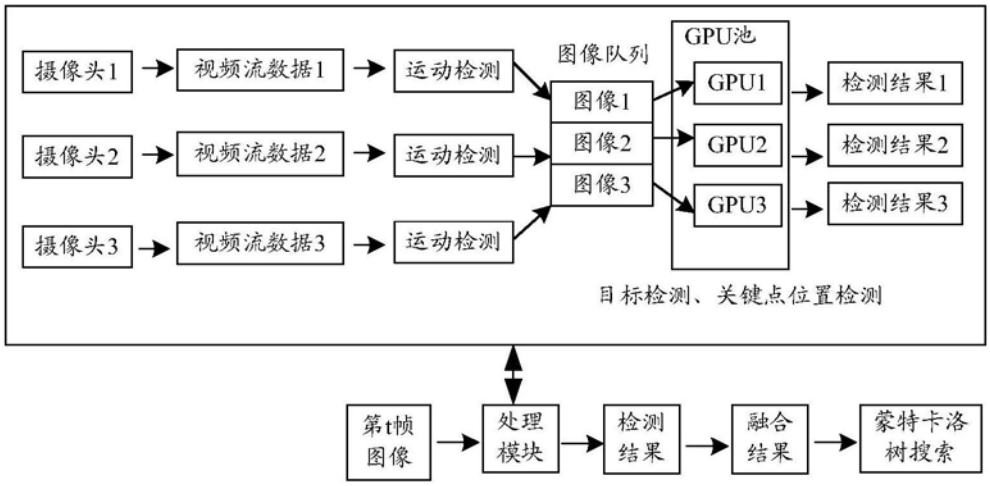

摄像头1
视频流数据1
运动检测
摄像头2
视频流数据2
运动检测
摄像头3
视频流数据3
运动检测
图像队列
图像1
图像2
图像3
GPU池
GPU1
GPU2
GPU3
检测结果1
检测结果2
检测结果3
目标检测、关键点位置检测
第t帧图像
处理模块
检测结果
融合结果
蒙特卡洛树搜索

图4

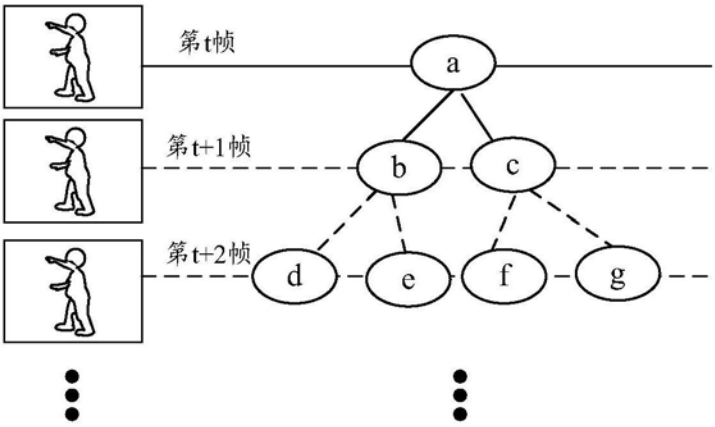

第t帧
第t+1帧
第t+2帧
a
b
c
d
e
f
g

图5

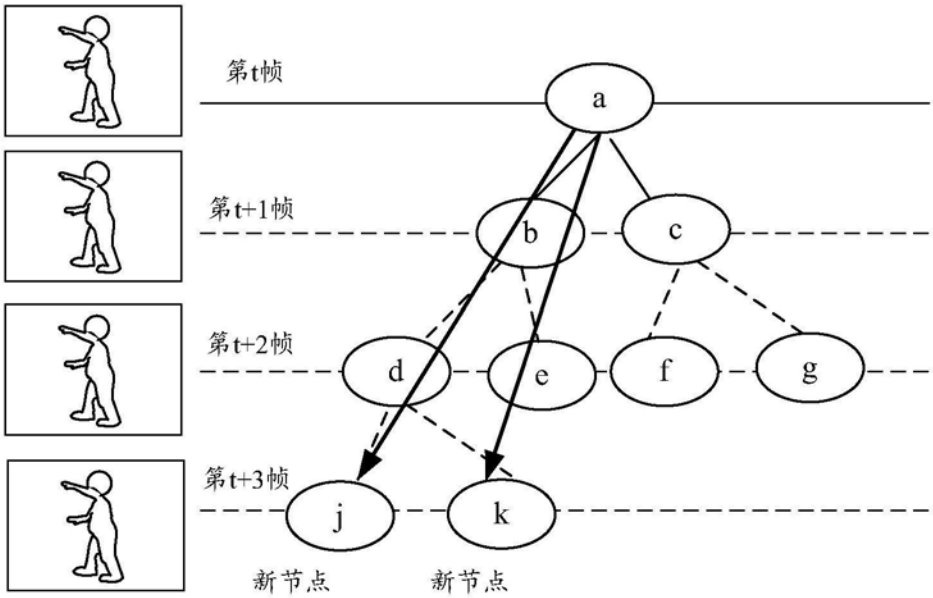

第t帧
a
第t+1帧
b
c
第t+2帧
d
e
f
g
第t+3帧
j
k
新节点
新节点

图6

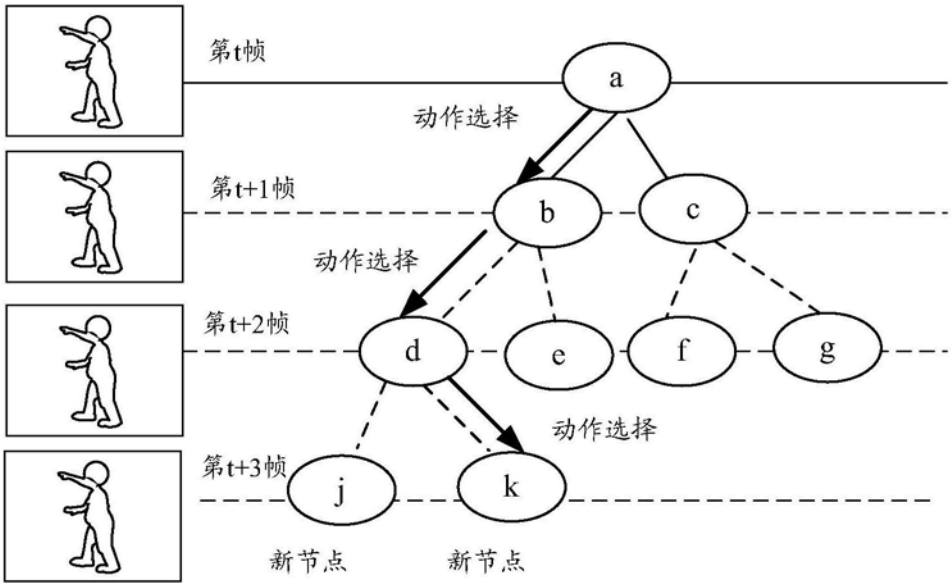

第t帧
a
动作选择
第t+1帧
b
c
动作选择
第t+2帧
d
e
f
g
动作选择
第t+3帧
j
k
新节点
新节点

图7

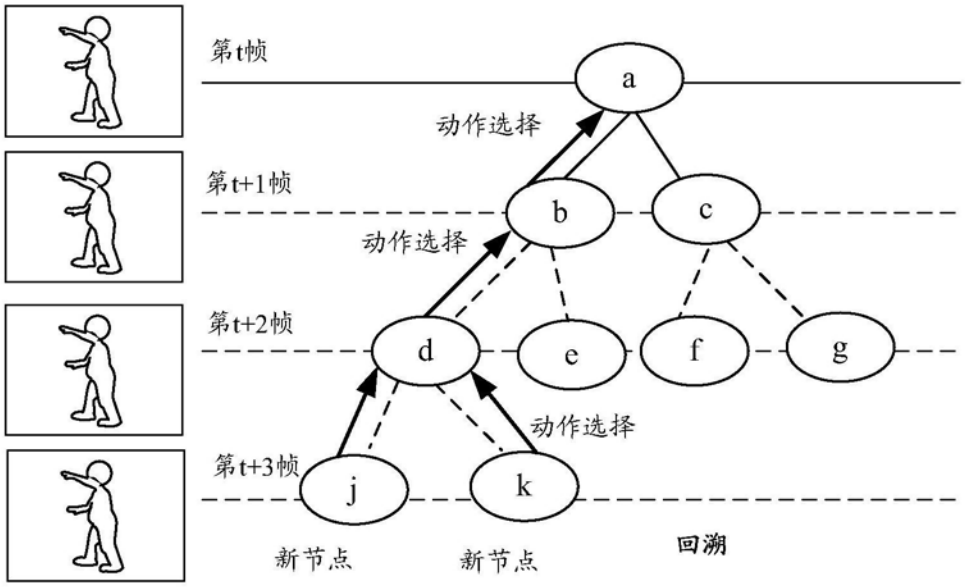

第t帧
a
动作选择
第t+1帧
b
c
动作选择
第t+2帧
d
e
f
g
动作选择
第t+3帧
j
k
新节点
新节点
回溯

图8

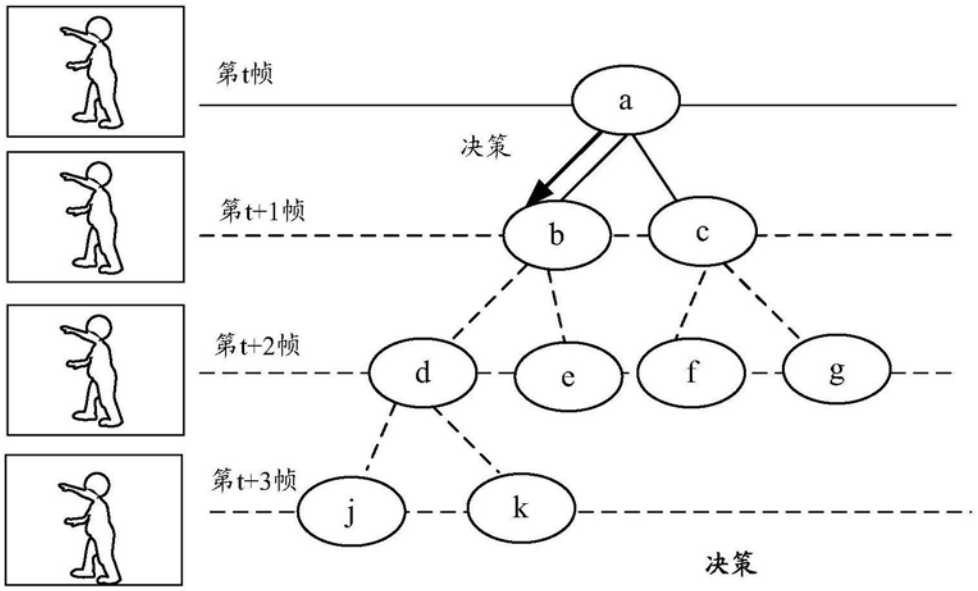

第t帧
a
决策
第t+1帧
b
c
第t+2帧
d
e
f
g
第t+3帧
j
k
决策

图9

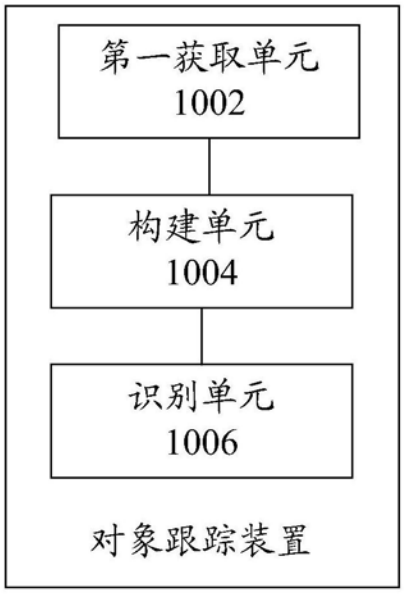

第一获取单元
1002
构建单元
1004
识别单元
1006
对象跟踪装置

图10

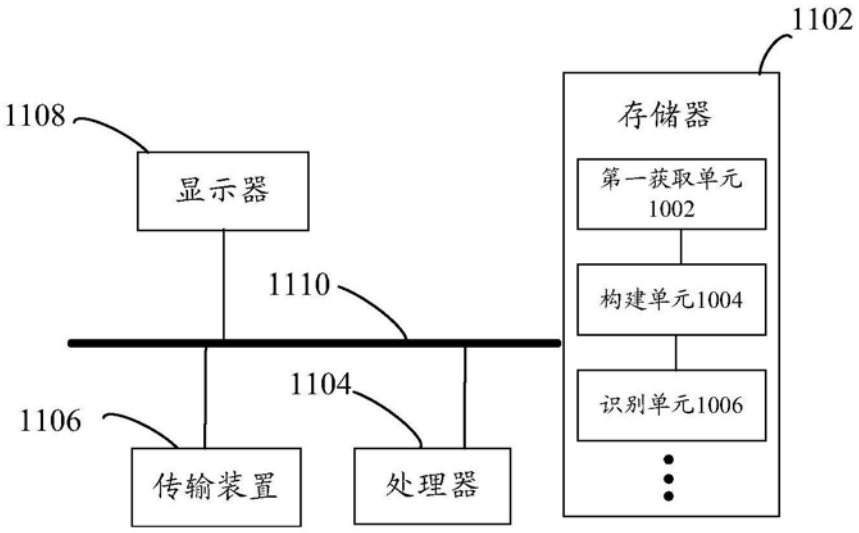

1102
存储器
第一获取单元
1002
构建单元1004
识别单元1006
1108
显示器
1110
1104
1106
传输装置
处理器

图11