

目标跟踪方法和装置、电子设备、存储介质

本申请要求在 2018 年 3 月 6 日提交中国专利局、申请号为 CN201810184967.8、发明名称为“目标跟踪方法和装置、电子设备、程序、存储介质”和申请号为 CN201810182797.X、发明名称为“目标跟踪方法和装置、电子设备、程序、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

本公开涉及计算机视觉技术，尤其是一种目标跟踪方法和装置、电子设备、存储介质。

背景技术

目标跟踪通常指在给定目标物体在视频序列中第一帧的位置信息的前提下，预测目标物体在视频序列中的后续图像帧的位置。目标跟踪技术在视频监控、无人驾驶、人机交互等应用场景中有广泛的应用。由于单帧的检测技术和分割技术耗时相对较高，目标跟踪技术常用来与检测分割技术一起使用，来提高视频处理的速度，保证更好的实时性。因此，实用工程中特定任务上的目标跟踪算法对速度的要求很高。

发明内容

本公开实施例提供一种目标跟踪技术。

根据本公开实施例的一个方面，提供一种目标跟踪方法，包括：

获取第一图像帧的参考帧的特征数据，其中，所述第一图像帧与至少一个第二图像帧具有相同的参考帧；

基于所述参考帧的特征数据，确定所述第一图像帧中跟踪目标的位置。

可选地，所述基于所述参考帧的特征数据，确定所述第一图像帧中跟踪目标的位置，包括：

对所述第一图像帧进行特征提取，得到所述第一图像帧的特征数据；

基于所述参考帧的特征数据和所述第一图像帧的特征数据，得到所述第一图像帧中跟踪目标的位置。

可选地，所述基于所述参考帧的特征数据和所述第一图像帧的特征数据，得到所述第一图像帧中跟踪目标的位置，包括：

连接所述第一图像帧的特征数据与所述参考帧的特征数据，得到连接特征；

基于所述连接特征，得到所述第一图像帧中跟踪目标的位置。

可选地，所述基于所述参考帧的特征数据，得到所述第一图像帧中跟踪目标的位置，包括：

基于所述参考帧的特征数据，获得所述第一图像帧中的跟踪目标相对于所述参考帧中的跟踪目标的相对位置信息；

基于所述相对位置信息和所述参考帧中的跟踪目标的位置信息，得到所述第一图像帧中跟踪目标的位置。

可选地，所述获取第一图像帧的参考帧的特征数据，包括：

从存储器中获取所述第一图像帧的参考帧的特征数据。

可选地，还包括：

基于所述第一图像帧中跟踪目标的位置和所述参考帧中跟踪目标的位置，确定是否将所述第一图像帧作为更新后的参考帧；和/或

若所述第一图像帧与所述参考帧之间的间隔满足预设条件，将所述第一图像帧作为更新后的参考帧。

可选地，所述基于所述第一图像帧中跟踪目标的位置和所述参考帧中跟踪目标的位置，确定是否将所述第一图像帧作为更新后的参考帧，包括：

基于所述第一图像帧中跟踪目标的第一限位框与所述参考框中跟踪目标的第二限位框之间的交并比，确定是否将所述第一图像帧作为更新后的参考帧。

可选地，所述基于所述第一图像帧中跟踪目标的第一限位框与所述参考框中跟踪目标的第二限位框之间的交并比，确定是否将所述第一图像帧作为更新后的参考帧，包括：

响应于所述第一限位框与所述第二限位框之间的交并比小于或等于预设阈值，确定将所述第一图像帧作为更新后的参考帧。

可选地，还包括：

响应于确定将所述第一图像帧作为更新后的参考帧，存储所述第一图像帧的特征数据。

根据本公开实施例的另一个方面，提供一种目标跟踪装置，包括：

特征获取单元，用于获取第一图像帧的参考帧的特征数据，其中，所述第一图像帧与至少一个第二图像帧具有相同的参考帧；

位置确定单元，用于基于所述参考帧的特征数据，确定所述第一图像帧中跟踪目标的位置。

可选地，所述位置确定单元，包括：

特征提取模块，用于对所述第一图像帧进行特征提取，得到所述第一图像帧的特征数据；

位置模块，用于基于所述参考帧的特征数据和所述第一图像帧的特征数据，得到所述第一图像帧中跟踪目标的位置。

可选地，所述位置模块，用于连接所述第一图像帧的特征数据与所述参考帧的特征数据，得到连接特征；

基于所述连接特征，得到所述第一图像帧中跟踪目标的位置。

可选地，所述位置确定单元，用于基于所述参考帧的特征数据，获得所述第一图像帧中的跟踪目标相对于所述参考帧中的跟踪目标的相对位置信息；

基于所述相对位置信息和所述参考帧中的跟踪目标的位置信息，得到所述第一图像帧中跟踪目标的位置。

可选地，所述特征获取单元，用于从存储器中获取所述第一图像帧的参考帧的特征数据。

可选地，还包括：

参考确定单元，用于基于所述第一图像帧中跟踪目标的位置和所述参考帧中跟踪目标的位置，是否将所述第一图像帧确定为更新后的参考帧；和/或

若所述第一图像帧与所述参考帧之间的间隔满足预设条件，将所述第一图像帧确定为更新后的参考帧。

可选地，所述参考确定单元，用于基于所述第一图像帧中跟踪目标的第一限位框与所述参考帧中跟踪目标的第二限位框之间的交并比，确定是否将所述第一图像帧确定为更新后的参考帧。

可选地，所述参考确定单元，用于响应于所述第一限位框与所述第二限位框之间的交并比小于或等于预设阈值，确定将所述第一图像帧作为更新后的参考帧。

可选地，所述参考确定单元，还用于响应于确定将所述第一图像帧作为更新后的参考帧，存储所述第一图像帧的特征数据。

根据本公开实施例的另一个方面，提供一种电子设备，包括处理器，所述处理器包括如上所述的目标跟踪装置。

根据本公开实施例的另一个方面，提供一种电子设备，包括：存储器，用于存储可执行指令；

以及处理器，用于与所述存储器通信以执行所述可执行指令从而完成如上所述目标跟

踪方法。

根据本公开实施例的另一个方面，提供一种计算机存储介质，用于存储计算机可读的指令，所述指令被处理器执行时，所述处理器执行如上所述目标跟踪方法。

根据本公开实施例的另一个方面，提供一种计算机程序，包括计算机可读代码，当所述计算机可读代码在设备上运行时，所述设备中的处理器执行如上所述目标跟踪方法。

根据本公开实施例的再一个方面，提供一种计算机程序产品，用于存储计算机可读指令，所述指令被执行时使得计算机执行上述任一可能的实现方式中所述的目标跟踪方法。

在一个可选实施方式中，所述计算机程序产品具体为计算机存储介质，在另一个可选实施方式中，所述计算机程序产品具体为软件产品，例如 SDK 等。

根据本公开实施例还提供了另一种目标跟踪方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品，其中，该方法包括：获取第一图像帧的参考帧的特征数据，其中，所述第一图像帧与至少一个第二图像帧对应相同的参考帧；基于所述参考帧的特征数据，确定所述第一图像帧中跟踪目标的位置。

基于本公开上述实施例提供的一种目标跟踪方法和装置、电子设备、存储介质，获取第一图像帧的参考帧的特征数据；基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置，由于，参考帧对应至少两个图像帧，实现了特征数据复用，减少了特征数据的计算次数，提升了跟踪速度。

下面通过附图和实施例，对本公开的技术方案做进一步的详细描述。

附图说明

构成说明书的一部分的附图描述了本公开的实施例，并且连同描述一起用于解释本公开的原理。

参照附图，根据下面的详细描述，可以更加清楚地理解本公开，其中：

图 1 为本公开实施例提供的目标跟踪方法的流程图。

图 2 为本公开实施例提供的目标跟踪装置的结构示意图。

图 3 为适于用来实现本公开实施例的终端设备或服务器的电子设备的结构示意图。

具体实施方式

现在将参照附图来详细描述本公开的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本公开的范围。

同时，应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本公开及其应用或使用的任何限制。

对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

本公开实施例可以应用于计算机系统/服务器，其可与众多其它通用或专用计算系统环境或配置一起操作。适于与计算机系统/服务器一起使用的众所周知的计算系统、环境和/或配置的例子包括但不限于：个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境，等等。

计算机系统/服务器可以在由计算机系统执行的计算机系统可执行指令（诸如程序模块）的一般语境下描述。通常，程序模块可以包括例程、程序、目标程序、组件、逻辑、数据结构等等，它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施，分布式云计算环境中，任务是由通过通信网络链接的远程处理设备执行的。在分布式云计算环境中，程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

图1为本公开实施例提供的目标跟踪方法的流程图。该方法可以由任意电子设备执行，例如终端设备、服务器、移动设备等等。

步骤110，获取第一图像帧的参考帧的特征数据。

在一个可选示例中，该步骤110可以由处理器调用存储器存储的相应指令执行，也可以由特征获取单元21执行。

为了对跟踪目标实现跟踪，可选地，通过视频等手段采集该跟踪目标的多个图像帧，通过确定这些图像帧中跟踪目标的位置，结合获取图像帧的时间，即可确定该跟踪目标按照时间在这些目标图像中的运动轨迹，实现对跟踪目标的跟踪。

在一些实施例中，为了实现对跟踪目标进行跟踪，需要已知的一个图像帧中跟踪目标的位置，并基于该位置确定当前需要跟踪的跟踪目标；或者，已知跟踪目标为某一特定目标（例如：人脸或人体等），针对确定的跟踪目标采集多个图像帧。

可选地，根据目标跟踪问题的定义，可以基于一个或多个图像中的跟踪目标的已确定的位置，实现对跟踪目标的跟踪，本公开实施例跟踪的目标可以是已知或未知的特定物体或人或其他类型，本公开实施例对此不做限定。

在本公开实施例中，至少两个图像帧可以复用相同的参考帧，其中，可选地，该至少两个图像帧可以为相邻的至少两个图像帧，例如，在本公开实施例中，可以每隔 M 帧更新一次参考帧，则相邻的 $M-1$ 个图像帧可以对应相同的参考帧，其中， M 为大于或等于 2 的预设数值，该预设数值可以根据实际应用需求确定，但本公开实施例对 M 的取值以及复用相同的参考帧的多个图像帧的实现不作限定。

在本公开实施例中，某个图像帧与其参考帧可以相邻或不相邻。例如，第一图像帧与其参考帧之间可以间隔一个或多个图像帧。例如，第一图像帧与其参考帧可以间隔 N 个图像帧，其中， N 为大于或等于 1 的数值，本公开实施例对 N 的数值不作限定。

这样，第一图像帧和其他图像帧（例如一个或多个第二图像帧）可以复用相同的参考帧，其中，第二图像帧可以位于第一图像帧之前或之后，与其他目标跟踪方式中每帧目标图像采用不同的参考帧相比，能够提高目标跟踪的处理效率。

可选地，第一图像帧的参考帧可以包括一个或多个图像帧，其中，该多个图像帧可以是视频流中的相邻图像帧或者不相邻的图像帧。例如，该多个参考帧可以为视频流中相邻的 K 个参考帧；或者，该多个参考帧中的部分相邻参考帧可以在视频流中相邻，另一部分相邻参考帧可以在视频中不相邻；或者，该多个参考帧中的任意两个相邻参考帧可以在视频流中互不相邻，可选地，多个参考帧之间可以间隔一个或多个图像。在一些实施例中，该多个图像帧中的任意两个相邻图像帧之间间隔相同个数的图像帧，例如间隔 M 个图像帧，其中， M 为大于或等于 2 的预设数值。 M 的数值可以根据应用场景和需求确定，本公开实施例对 M 的数值不作限定。或者，该多个图像帧中的不同的相邻图像帧之间也可以间隔不同个数的图像帧，本公开实施例对此不做限定。

应理解，本文中的“多个参考图像中的相邻图像或相邻参考图像”可以指按照一定排序规则对该多个参考图像进行排序之后相邻的参考图像，其中，该排序规则可以为按照采集时间或检测时间排序，等等，本文中的“视频流中的相邻图像”可以指视频流或视频流片段中的多个图像进行排序之后相邻的图像，该排序可以指视频流中的原本顺序或某种特定排序规则，本公开实施例对排序的具体实现不做限定。

在本公开实施例中，可以通过多种方式获取参考帧的特征数据。在一些实施例中，可以从其他装置处获取参考帧的特征数据，例如，服务器接收终端设备发送的参考帧的特征

数据。再例如，电子设备从数据库中和获取参考帧的特征数据，其中，该数据库可以设置在电子设备本地或其他设备处。在另一些实施例中，可以对参考帧进行特征提取处理，得到参考帧的特征数据。

在一个或多个可选的实施例中，电子设备可以获取存储的参考帧的特征数据，例如，从存储器中获取参考帧的特征数据。可选地，可以在对参考帧进行处理的过程中（例如确定参考帧中的跟踪目标的位置信息）将参考帧的特征数据存储于缓存区中，并从缓存区获取参考帧的特征数据，但本公开实施例不限于此。

在本公开实施例中，可选地，目标图像与其多个参考图像可以相邻。或者，目标图像与其参考图像之间可以间隔至少一个图像帧（称为第二图像）。可选地，目标图像可以与其多个参考图像中的任意参考图像之间间隔一个或多个图像；或者，目标图像可以与多个参考图像中的部分参考图像相邻，与多个参考图像中的另一部分参考图像间隔一个或多个图像，本公开实施例对此不做限定。与采用目标图像的前一帧图像做参考图像的方式相比，本公开实施例提供的技术方案可以提升预测精度，这是由于较早的图像的预测更加准确，用较早的图像来做跟踪的参考图像能有效解决误差累积的问题。

在本公开实施例中，可选地，不同图像可以具有不同的参考图像，或者，多个图像可以复用相同的参考图像，例如，目标图像与至少一个第三图像可以具有相同的参考图像，例如，目标图像的多个参考图像与第三图像的多个参考图像完全相同，或者，目标图像的多个参考图像中的部分参考图像也是第三图像的参考图像，本公开实施例对此不做限定。其中，可选地，该至少一个第三图像可以与该目标图像相邻，但本公开实施例不限于此。此时，可选地，可以将参考图像的特征存储至缓存区，在进行图像的跟踪目标的位置预测时，仅需要调用存储的参考图像的特征，使得位置预测的速度得到提升。

在本公开实施例中，可以通过多种方式获取参考图像的特征。在一些实施例中，可以从其他装置处获取参考图像的特征，例如，服务器接收终端设备发送的参考图像的特征，再例如，电子设备从数据库获取参考图像的特征，其中，该数据库可以设置在电子设备本地或其他设备处。在另一些实施例中，可以在获取到参考图像之后，对参考图像进行特征提取处理，得到参考图像的特征。

在一个或多个可选的实施例中，可以获取电子设备的存储器存储的多个参考图像的特征，例如，从第二缓存区获取目标图像的多个参考图像的特征。可选地，可以将参考图像的特征存储于第二缓存区中，并从第二缓存区获取参考图像的特征，其中，该第二缓存区也可以称为位置模板缓存区（Location Exemplar Buffer），但本公开实施例对第二缓存区的

名称不作限定。

在本公开实施例中，第二缓存区可以用于存储参考图像的信息，可选地，可以存储当前作为参考图像的多个图像的信息，即参考图像列表当前包括的多个参考图像的信息，例如，可以存储参考图像的特征，或者还可以进一步存储参考图像中跟踪目标的位置信息，本公开实施例对此不做限定。

在基于深度回归网络实现的目标跟踪方法中，深度回归网络由参考图像分支和目标图像分支组成，参考图像分支利用参考图像特征提取器可以获得有价值的位置模板（location exemplar）信息，指导目标图像分支对目标物体的位置进行预测。可选地，一般采用利用第 $t-1$ 帧预测第 t 帧的策略。然而，事实上，在预测第 t 帧目标物体位置的时候，第 $t-2$ 帧前面的所有图像已经经过了参考图像分支的特征提取器，有效的位置模板已经获得。因此可以进行位置模板的复用，从而提高跟踪系统的速度。

为了提高跟踪速度，将基于参考帧获得的特征数据存入存储器中，该特征数据可以是对参考帧进行特征提取获得，还可以将获得跟踪目标的最终位置的图像帧作为参考帧存入存储器，以确定后续图像帧中跟踪目标的位置。

步骤 120，基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置。

在一个可选示例中，该步骤 120 可以由处理器调用存储器存储的相应指令执行，也可以由位置确定单元 22 执行。

可选地，跟踪目标的位置可以包括跟踪目标在第一图像帧中的坐标信息，例如，跟踪目标的限位框（bounding box）在第一图像帧中的坐标信息，等等，本公开实施例对此不做限定。

可选地，第一图像帧与参考帧间隔至少一个图像帧。

在一些可选的实施例中，可以对第一图像帧进行特征提取，得到第一图像帧的特征数据，并基于第一图像帧的特征数据和参考帧的特征数据，确定第一图像帧中跟踪目标的位置。

可选地，可以通过神经网络对图像帧进行特征提取，例如：通过卷积神经网络中的卷积层对第一图像帧进行卷积操作，得到第一图像帧的特征数据，其中，可选地，该特征数据可以为特征图或其他形式，本公开实施例不限制提取特征数据的方式和获得的特征数据的具体实现形式。

可选地，可以连接（或拼接）第一图像帧的特征数据与参考帧的特征数据，得到连接特征，并基于连接特征或拼接特征，得到第一图像帧中跟踪目标的位置。

这里的连接（或拼接）可以指通道叠加，例如，第一图像帧通过 3 个通道输出 3 个特征图，参考帧通过 3 个通道输出 3 个特征图，经过特征连接可以得到 6 个特征图。

在本公开实施例中，基于连接特征得到跟踪目标位置的方式可以有多种。在一些可选实施例中，可以利用全连接层和/或其他网络层的处理，得到跟踪目标的位置，本公开实施例对此不做限定。

可选地，可以对连接特征进行处理，得到第一图像帧中跟踪目标对应的坐标；

得到的坐标所对应的范围即为跟踪目标在第一图像帧中的位置。

在一些可选的实施例中，可以基于参考帧的特征数据，获得第一图像帧中的跟踪目标相对于参考帧中的跟踪目标的相对位置信息，并基于相对位置信息和参考帧中的跟踪目标的位置信息，得到第一图像帧中跟踪目标的位置。

可选地，基于参考帧的特征数据和第一图像帧的特征数据，可以确定第一图像帧和参考帧之间的相对差异，根据特征数据之间的相对差异，即可确定跟踪目标在第一图像帧和参考帧中的相对位置信息，这里的相对位置信息可以指示跟踪目标在参考帧到第一图像帧所移动的相对距离。由于参考帧中的跟踪目标的位置信息是已确定的，基于已知的参考帧的位置信息和相对位置信息即可确定第一图像帧中跟踪目标的位置。

在其他的目标跟踪方式中，利用第 $t-1$ 帧预测第 t 帧的策略。事实上，在确定第 t 帧目标物体位置的时候，第 $t-2$ 帧前面的所有图像已经经过了参考图像分支的特征提取器，有效的位置模板已经获得。利用深度回归网络前向的这个特性可以提高跟踪系统的速度。

由于相邻两帧之间物体位移并不大，完全可以利用第 $t-2$ 、 $t-3$ 、 $t-K$ 帧或第 t 帧的位置。所以可以令第 $t-1$ 帧根本不经过参考图像分支，这样可以节省很多的时间。考虑实际工业生产中，对于采集的视频帧，如果选用每隔 K 帧更新一次参考图像分支的位置模板，就可以实现在速度上近 $2K/(K+1)$ 的提升。这样大的提升可以在实际应用中进一步提高实时性，降低设备功耗。

在本公开实施例中，多个目标图像帧可以复用相同的参考帧。这样，在对某个目标图像进行跟踪目标的位置预测时，可以调用已保存的参考帧的特征数据，使得位置预测的速度得到提升，并且，不用前一个图像帧做参考帧的策略可以提升精度，这是由于本公开实施例中检测得到的参考帧中跟踪目标的位置的精度较高，较早的图像帧中跟踪目标的位置预测会更加准确，用较早的图像帧来做跟踪的参考帧能有效解决误差累积的现象。

基于本公开上述实施例提供的一种目标跟踪方法，获取第一图像帧的参考帧的特征数据；基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置，由于，第一图像帧与至

少一个第二图像帧具有相同的参考帧，实现了特征数据复用，减少了特征数据的计算次数，提升了跟踪速度。

在一些可选的示例中，步骤 120 可以包括：基于参考帧的特征数据，确定第一图像帧中跟踪目标的多个初始预测位置；

基于多个初始预测位置确定第一图像帧中跟踪目标的最终位置。

可选地，初始预测位置和/或最终位置可以指跟踪目标对应的限位框的位置，例如人脸框、人体框或特定物体或部位的位置，等等，本公开实施例对跟踪目标的位置的具体形式不做限定。

在一些实施例中，可以对多个初始预测位置进行平均处理，例如，数学平均或加权平均等，得到跟踪目标在第一图像帧中的最终位置。在另一些实施例中，可以从多个初始预测位置中选取至少一个初始预测位置，并基于至少一个初始预测位置，得到跟踪目标在第一图像帧中的最终位置。

在一个或多个可选的实施例中，基于多个初始预测位置的置信度，确定第一图像帧中跟踪目标的最终位置。

可选地，基于多个初始预测位置的置信度，从多个初始预测位置中选取至少一个初始预测位置，并基于至少一个初始预测位置，确定最终位置。例如，从多个初始预测位置中选取置信度较高的至少一个初始预测位置。其中，该至少一个初始预测位置可以是多个初始预测位置中置信度最高的预设数量个位置，也可以是多个初始预测位置中置信度大于第二预设阈值的位置，本公开实施例对此不做限定。

置信度表示该预测位置对应于该第一图像帧中跟踪目标的真实位置的概率，在一些实施例中，置信度是通过打分网络获得的，该打分网络为经过训练的神经网络，例如：利用一个二分类网络，当预测位置与真实位置完成重合时，输出真实值为 1，预测值为 0，当预测位置与真实位置不重合时，输出真实值为 0，预测值为 1，等等，本公开实施例对打分网络的具体实现不做限定。

在一些实施例中，将选取的至少一个初始预测位置的平均值作为最终位置，其中，该平均值可以为算数平均值、几何平均值或加权平均值，例如，对选取的至少一个初始预测位置进行加权平均，获得第一图像帧中跟踪目标的最终位置，其中，初始预测位置的权重值与其置信度相关，置信度越大，权重值越大，置信度越小，权重值越小，但本公开实施例不限于此。

通过置信度可以将得到的初始预测位置中偏离较大的初始预测位置删除，相应地，得

到的最终位置更接近该目标图像中的跟踪目标的真实位置，能实现更好的跟踪。

可选地，确定多个初始预测位置中至少一个初始预测位置的置信度，包括：

利用置信度打分网络分别对多个初始预测位置进行处理，得到多个初始预测位置对应的置信度。

其中，置信度打分网络是经过训练获得的，对应输入的初始预测位置训练后的置信度打分网络可输出对应的置信度，可选地，可基于预测置信度和真实置信度获得误差，根据误差利用反向梯度传播方法对打分网络进行训练，训练后的打分网络能更好的对预测位置为真实位置（已知目标位置）的概率进行预测；该打分网络针对需要跟踪的特定物体训练，例如：需要对人脸进行跟踪，就采用人脸图像对该打分网络进行训练。

可选地，在利用置信度打分网络分别对多个初始预测位置进行处理，得到多个初始预测位置对应的置信度之前，还包括：

利用置信度打分网络对样本图像的至少一个初始预测位置进行处理，得到对应至少一个初始预测位置的预测置信度；

基于样本图像的标注位置，确定至少一个初始预测位置的真实置信度；

基于至少一个初始预测位置的预测置信度和真实置信度，训练置信度打分网络。

为了提高置信度网络打分的准确性，需要对置信度打分网络进行训练，训练过程与神经网络的训练过程相同，都是基于已知真实置信度的样本图像对置信度打分网络进行训练，可选地，针对不同的跟踪目标，可采用针对性的训练，例如：对人脸进行跟踪，在训练置信度打分网络过程中，采用已知真实置信度的人脸图像对置信度打分网络进行训练，使得到的置信度打分网络对人脸图像的置信度打分更准确。

在一些实施例中，还可以确定是否将第一图像帧作为更新后的参考帧。

可选地，如果确定不进行参考帧的更新，即确定不将第一图像帧作为更新后的参考帧，则利用相同的参考帧继续对第一图像帧的下一图像帧进行目标跟踪处理。可选地，如果确定进行参考帧的更新，即确定将第一图像帧作为更新后的参考帧，则利用更新后的参考帧，即第一图像帧，对第一图像帧的下一图像帧进行目标跟踪处理。

在一个或多个可选的实施例中，基于第一图像帧中跟踪目标的位置和参考帧中跟踪目标的位置，确定是否将第一图像帧作为更新后的参考帧。

在另一些实施例中，基于第一图像帧与参考帧之间的间隔，确定是否将第一图像帧作为更新后的参考帧。

例如，响应于第一图像帧与参考帧之间的间隔满足预设条件，将第一图像帧作为更新

后的参考帧。可选地，该预设条件可以为达到预设数值，例如，每隔 M 个图像帧更新一次参考帧，但本公开实施例不限于此。

在实际的目标跟踪中，跟踪目标的运动复杂，其速度往往不是恒定不变的。对于第 $t+1$ 帧的图像帧，可以通过对跟踪目标在第 t 帧中的位置和参考帧中的位置进行比对，来动态调整参考帧，以获得更适合的位置模板特征。

可选地，如果利用第 t 帧更新了参考帧，即将第 t 帧作为更新后的参考帧，则第 $t+1$ 帧图像与其参考帧之间可以相邻，而如果未将第 t 帧作为参考帧，则参考帧并未进行更新，此时，第 $t+1$ 帧图像与其参考帧之间可以间隔一个或多个图像帧。例如，图像帧与其参考帧可以间隔 N 个图像帧，其中， N 为大于或等于 2 的预设数值。 N 的数值可以根据应用场景确定，本公开实施例对 N 的数值不作限定。

可选地，基于第一图像帧中跟踪目标的位置与参考帧中跟踪目标的位置之间的距离，确定是否将第一图像帧作为更新后的参考帧。例如，响应于距离达到特定数值，确定将第一图像帧作为更新后的参考帧。再例如，响应于该距离小于特定数值，确定不将第一图像帧作为更新后的参考帧，即不进行参考帧的更新。其中，可以通过多种方式确定两个位置之间的距离，例如，将跟踪目标在第一图像帧中的限位框的中心点或某一点等特定点与跟踪目标在参考帧中的限位框的对应点之间的距离，作为跟踪目标在第一图像帧中的位置与在参考帧中的位置之间的距离，等等，本公开实施例对此不做限定。

可选地，基于第一图像帧中跟踪目标的第一限位框（bounding box）与参考帧中跟踪目标的第二限位框（bounding box）之间的交并比，确定是否将第一图像帧作为更新后的参考帧。

如果一直用同一个参考帧的特征数据对图像帧进行位置预测，当跟踪目标移动位置较大时，将导致的位置不准确，因此，需要对参考帧的特征数据进行更新，何时进行更新，可选地，根据跟踪目标的第一限位框与参考框中跟踪目标的第二限位框之间的交并比是否大于预设值决定的；其中，第一限位框是基于预测的位置对应的至少一个坐标（例如，所有顶点坐标）获得的，而第二限位框是基于参考帧中跟踪目标对应的至少一个坐标（例如，所有顶点坐标）获得的。

可选地，响应于第一限位框与第二限位框之间的交并比小于或等于预设阈值，确定将第一图像帧作为更新后的参考帧。

可选地，响应于第一限位框与第二限位框之间的交并比达到预设阈值，确定不进行参考帧的更新。

交并比(intersection-over-union)表示两个区域的交集与并集之比,在本公开实施例中指第一限位框和第二限位框的交集和第一限位框和第二限位框的并集之比;交并比越小说明跟踪目标从参考帧到第一图像帧运动的距离越大,此时,再继续基于当前参考帧作为参考帧将导致预测的位置不准确,因此,当交并比小于设定值时,将该第一图像帧作为参考帧。

可选地,参考帧对应的特征数据的获得,可以是通过对第一图像帧进行特征提取,获得更新后的参考帧的特征数据;或者,确定跟踪目标在第一图像帧中的位置坐标,对位置坐标执行特征提取,得到更新后的参考帧的特征数据。

如果预测的位置与参考帧中的位置的绝对距离较近(第一限位框与第二限位框之间的交并比大于或等于设定阈值),则参考帧无需更新,可以继续用目前缓存的参考帧的特征数据来指导下一帧图像帧。若预测的位置与参考帧中的位置的绝对距离较远(第一限位框与第二限位框之间的交并比小于设定阈值),则可以利用第一图像帧的位置提取特征来更新参考帧的特征数据;通过这种动态调整机制,可以充分利用先前计算过的特征数据(如:中间特征图),从而减少了很多不必要的特征提取过程。

例如:若利用目标跟踪方法跟踪在视频片段内绝对位置基本不变的物体,可以一直利用第一帧作为参考帧,极大的提高了每帧视频帧中跟踪目标的位置速度,进而提高了跟踪效率。

对于特定的问题(跟踪特定的目标,如:人脸跟踪),在应用中可以采用针对性的参考帧更新策略对目标跟踪方法的参考帧进行调整,有着非常强的推广价值。

在一个或多个可选的实施例中,还包括:响应于确定将第一图像帧作为更新后的参考帧,存储第一图像帧的特征数据。

本公开实施例中可以通过缓存的方式保存生成的特征数据,为了节省存储空间,可以每次将更新后的参考帧的特征数据替换前一个参考帧的特征数据进行保存,在需要进行位置预测时,只需从缓存中调用当前存储的特征数据即可;可选地,还可以按更新时间保存至少一个参考帧的特征数据;存储至少一个特征数据(例如,所有包括首次生成的特征数据和更新后的特征数据),这样,在需要进行位置预测时,调用最新存储的特征数据作为参考帧的特征数据。

用户在使用手机、相机进行人脸及一般物体的检测时,常见的思路是通过与目标跟踪算法结合来提高速度,降低设备功耗。本公开实施例在深度回归网络目标跟踪技术上使用特征缓存和特征复用机制,可以大幅提升传统回归跟踪器的速度,可以在特定任务(人脸

跟踪)上实现 1500fps 的高速率跟踪。这在实际产品应用中非常关键,大幅提高了目标跟踪算法的前向速度。

本公开实施例的核心是提出基于特征缓存器的框架,实现高精度的实用高速目标跟踪方法。本公开通过对回归网络结构的分析,提出通过特征缓存对回归网络中间特征图的维护和复用机制。本公开实施例提出利用位置模板缓存来存储计算得到的中间特征(参考帧分支计算得到的位置模板),在之后的跟踪过程中复用之前缓存的位置模板,从而能够快速利用多个参考帧模板对物体进行。本公开实施例提出的方法可以应用于多类目标跟踪场景中,对基于深度回归网络的目标跟踪算法提升运行速度到近原来的两倍,大幅提高了跟踪实时性,降低了设备功耗。

本公开实施例提出的目标跟踪算法,在保证精度下,对传统深度回归网络目标跟踪系统实现了近一倍的速度提升。在特定任务(人脸跟踪)上,本公开提出的基于特征复用的回归框架显著提升了跟踪器的精度-速度平衡曲线,一方面可以实现在保持精度的条件下提升速度,另一方面也可以在速度不变的情况下提高网络的表达能力,提高目标跟踪器的精度。

对于人脸跟踪上在一个长序列视频(例如,7289 帧)上做实验结果如下。表 1 为该实验中应用不同跟踪方法实现跟踪,获得的指标值对比表。如表 1 所示,实验中采取动态更新参考帧位置模板特征的策略。可以看到,在保证精度不变的情况下速度提高近一倍。事实上,这种不确定用前一帧来做参考帧的策略可以小幅提升精度,这是由于检测帧的精度较高,较早的帧更加准确,用较早的帧来做跟踪的参考帧能有效解决误差累积的现象。

表 1 不同跟踪方法实现跟踪的指标值对比表

指标 跟踪方法	速度 (ms)	鲁棒性([0.2:0.5])	精度 ([0.5:0.95])
16 通道跟踪器	0.4015	0.771	0.568
16 通道+特征复用	0.2944	0.772	0.579
8 通道跟踪器	0.2716	0.768	0.540
8 通道+特征复用	0.1808	0.768	0.545

本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执

行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

图2为本公开实施例提供的目标跟踪装置的结构示意图。该装置可用于实现本公开上述各方法实施例。如图2所示，该装置包括：

特征获取单元21，用于获取第一图像帧的参考帧的特征数据。

为了对跟踪目标实现跟踪，可选地，通过视频等手段采集该跟踪目标的多个图像帧，通过确定这些图像帧中跟踪目标的位置，结合获取图像帧的时间，即可确定该跟踪目标按照时间在这些目标图像中的运动轨迹，实现对跟踪目标的跟踪。

在一些实施例中，为了实现对跟踪目标进行跟踪，需要已知的一个图像帧中跟踪目标的位置，并基于该位置确定当前需要跟踪的跟踪目标；或者，已知跟踪目标为某一特定目标（例如：人脸或人体等），针对确定的跟踪目标采集多个图像帧。

可选地，根据目标跟踪问题的定义，可以基于一个或多个图像中的跟踪目标的已确定的位置，实现对跟踪目标的跟踪，本公开实施例跟踪的目标可以是已知或未知的特定物体或人或其他类型，本公开实施例对此不做限定。

在本公开实施例中，至少两个图像帧可以复用相同的参考帧，其中，可选地，该至少两个图像帧可以为相邻的至少两个图像帧，例如，在本公开实施例中，可以每隔M帧更新一次参考帧，则相邻的M-1个图像帧可以对应相同的参考帧，其中，M为大于或等于2的预设数值，该预设数值可以根据实际应用需求确定，但本公开实施例对M的取值以及复用相同的参考帧的多个图像帧的实现不作限定。

在一个或多个可选的实施例中，特征获取单元21，可以获取存储的参考帧的特征数据，例如，从存储器中获取第一图像帧的参考帧的特征数据。可选地，可以将参考帧的特征数据存储在缓存区中，并从缓存区获取参考帧的特征数据，但本公开实施例不限于此。

位置确定单元22，用于基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置。其中，可选地，第一图像帧与参考帧间隔至少一个图像帧。

在一些可选的实施例中，位置确定单元22，可以包括：

特征提取模块，用于对第一图像帧进行特征提取，得到第一图像帧的特征数据；

位置模块，用于基于参考帧的特征数据和第一图像帧的特征数据，得到第一图像帧中跟踪目标的位置。

可选地，可以通过神经网络对图像帧进行特征提取，例如：通过卷积神经网络中的卷积层对第一图像帧进行卷积操作，得到第一图像帧的特征数据，其中，可选地，该特征数

据可以为特征图或其他形式，本公开实施例不限制提取特征数据的方式和获得的特征数据的具体实现形式。

可选地，位置模块，可以用于连接（或拼接）第一图像帧的特征数据与参考帧的特征数据，并得到连接特征或拼接特征；基于连接特征或拼接特征，得到第一图像帧中跟踪目标的位置。

在一些可选的实施例中，位置确定单元 22，用于基于参考帧的特征数据，获得第一图像帧中的跟踪目标相对于参考帧中的跟踪目标的相对位置信息；

基于相对位置信息和参考帧中的跟踪目标的位置信息，得到第一图像帧中跟踪目标的位置。

可选地，基于参考帧的特征数据和第一图像帧的特征数据，可以确定第一图像帧和参考帧之间的相对差异，根据特征数据之间的相对差异，即可确定跟踪目标在第一图像帧和参考帧中的相对位置信息，这里的相对位置信息可以指示跟踪目标在参考帧到第一图像帧所移动的相对距离。由于参考帧中的跟踪目标的位置信息是已确定的，基于已知的参考帧的位置信息和相对位置信息即可确定第一图像帧中跟踪目标的位置。

基于本公开上述实施例提供的一种目标跟踪装置，获取第一图像帧的参考帧的特征数据；基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置，由于，第一图像帧与至少一个第二图像帧具有相同的参考帧，实现了特征数据复用，减少了特征数据的计算次数，提升了跟踪速度。

在一个或多个可选的实施例中，还包括：

参考确定单元，用于基于第一图像帧中跟踪目标的位置和参考帧中跟踪目标的位置，是否将第一图像帧作为更新后的参考帧。

可选地，若第一图像帧与参考帧之间的间隔满足预设条件，将第一图像帧作为更新后的参考帧。

在实际的目标跟踪中，跟踪目标的运动复杂，其速度往往不是恒定不变的。对于第 t 帧的图像帧，可以通过对参考帧得到的绝对位置和参考帧绝对位置进行比对来动态调整参考帧，以获得更适合的位置模板特征。

可选地，如果利用第 t 帧更新了参考帧，即将第 t 帧作为更新后的参考帧，则第 $t+1$ 帧图像与其参考帧之间可以相邻，而如果未将第 t 帧作为参考帧，则参考帧并未进行更新，此时，第 $t+1$ 帧图像与其参考帧之间可以间隔一个或多个图像帧。例如，图像帧与其参考帧可以间隔 N 个图像帧，其中， N 为大于或等于 2 的预设数值。 N 的数值可以根据应用场

景确定，本公开实施例对 N 的数值不作限定。

可选地，参考确定单元，用于基于第一图像帧中跟踪目标的第一限位框（bounding box）与参考框中跟踪目标的第二限位框（bounding box）之间的交并比，是否将第一图像帧作为更新后的参考帧。

可选地，参考确定单元，用于响应于第一限位框与第二限位框之间的交并比小于或等于预设阈值，将第一图像帧作为更新后的参考帧。

可选地，参考确定单元，还用于响应于第一图像帧作为参考帧，存储参考帧的特征数据。

根据本公开实施例的另一个方面，提供一种电子设备，包括处理器，处理器包括本公开上述任一实施例的目标跟踪装置。

根据本公开实施例的另一个方面，提供一种电子设备，包括：存储器，用于存储可执行指令；

以及处理器，用于与存储器通信以执行可执行指令从而完成本公开提供的目标跟踪方法上述任一实施例。

根据本公开实施例的另一个方面，提供一种计算机存储介质，用于存储计算机可读的指令，指令被处理器执行时，该处理器执行本公开提供的目标跟踪方法上述任一实施例。

根据本公开实施例的另一个方面，提供一种计算机程序，包括计算机可读代码，当计算机可读代码在设备上运行时，设备中的处理器执行本公开提供的目标跟踪方法。

根据本公开实施例的再一个方面，提供一种计算机程序产品，用于存储计算机可读指令，所述指令被执行时使得计算机执行上述任一可能的实现方式中所述的目标跟踪方法。

在一个或多个可选实施方式中，本公开实施例还提供了一种计算机程序产品，用于存储计算机可读指令，所述指令被执行时使得计算机执行上述任一实施例中所述的目标跟踪方法。

该计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选例子中，所述计算机程序产品具体体现为计算机存储介质，在另一个可选例子中，所述计算机程序产品具体体现为软件产品，例如软件开发包（Software Development Kit，SDK）等等。

根据本公开实施例还提供了另一种目标跟踪方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品，其中，该方法包括：获取第一图像帧的参考帧的特征数据，其中，参考帧与第一图像帧之间间隔至少一个第二图像帧；基于参考帧的

特征数据，确定第一图像帧中跟踪目标的位置。

在一些实施例中，该目标跟踪指示可以具体为调用指令，第一装置可以通过调用的方式指示第二装置执行目标跟踪，相应地，响应于接收到调用指令，第二装置可以执行上述目标跟踪方法中的任意实施例中的步骤和/或流程。

应理解，本公开实施例中的“第一”、“第二”等术语仅仅是为了区分，而不应理解成对本公开实施例的限定。

还应理解，在本公开中，“多个”可以指两个或两个以上，“至少一个”可以指一个、两个或两个以上。

还应理解，对于本公开中提及的任一部件、数据或结构，在没有明确限定或者在前后文给出相反启示的情况下，一般可以理解为一个或多个。

还应理解，本公开对各个实施例的描述着重强调各个实施例之间的不同之处，其相同或相似之处可以相互参考，为了简洁，不再一一赘述。

本公开实施例还提供了一种电子设备，例如可以是移动终端、个人计算机（PC）、平板电脑、服务器等。下面参考图 3，其示出了适于用来实现本申请实施例的终端设备或服务器的电子设备 300 的结构示意图：如图 3 所示，计算机系统 300 包括一个或多个处理器、通信部等，所述一个或多个处理器例如：一个或多个中央处理单元（CPU）301，和/或一个或多个图像处理器（GPU）313 等，处理器可以根据存储在只读存储器（ROM）302 中的可执行指令或者从存储部分 308 加载到随机访问存储器（RAM）303 中的可执行指令而执行各种适当的动作和处理。通信部 312 可包括但不限于网卡，所述网卡可包括但不限于 IB（Infiniband）网卡。

处理器可与只读存储器 302 和/或随机访问存储器 330 中通信以执行可执行指令，通过总线 304 与通信部 312 相连、并经通信部 312 与其他目标设备通信，从而完成本申请实施例提供的任一项方法对应的操作，例如，获取第一图像帧的参考帧的特征数据，其中，第一图像帧与至少一个第二图像帧具有相同的参考帧；基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置。

此外，在 RAM 303 中，还可存储有装置操作所需的各种程序和数据。CPU301、ROM302 以及 RAM303 通过总线 304 彼此相连。在有 RAM303 的情况下，ROM302 为可选模块。RAM303 存储可执行指令，或在运行时向 ROM302 中写入可执行指令，可执行指令使处理器 301 执行上述通信方法对应的操作。输入/输出（I/O）接口 305 也连接至总线 304。通信部 312 可以集成设置，也可以设置为具有多个子模块（例如多个 IB 网卡），并在总线链接

上。

以下部件连接至 I/O 接口 305：包括键盘、鼠标等的输入部分 306；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 307；包括硬盘等的存储部分 308；以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 309。通信部分 309 经由诸如因特网的网络执行通信处理。驱动器 310 也根据需要连接至 I/O 接口 305。可拆卸介质 311，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 310 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 308。

需要说明的，如图 3 所示的架构仅为一种可选实现方式，在具体实践过程中，可根据实际需要对上述图 3 的部件数量和类型进行选择、删减、增加或替换；在不同功能部件设置上，也可采用分离设置或集成设置等实现方式，例如 GPU 和 CPU 可分离设置或者可将 GPU 集成在 CPU 上，通信部可分离设置，也可集成设置在 CPU 或 GPU 上，等等。这些可替换的实施方式均落入本公开公开的保护范围。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括有形地包含在机器可读介质上的计算机程序，计算机程序包含用于执行流程图所示的方法的程序代码，程序代码可包括对应执行本申请实施例提供的方法步骤对应的指令，例如，获取第一图像帧的参考帧的特征数据，其中，第一图像帧与至少一个第二图像帧具有相同的参考帧；基于参考帧的特征数据，确定第一图像帧中跟踪目标的位置。在这样的实施例中，该计算机程序可以通过通信部分 309 从网络上被下载和安装，和/或从可拆卸介质 311 被安装。在该计算机程序被中央处理单元（CPU）301 执行时，执行本申请的方法中限定的上述功能。

本说明书中各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其它实施例的不同之处，各个实施例之间相同或相似的部分相互参见即可。对于系统实施例而言，由于其与方法实施例基本对应，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

可能以许多方式来实现本公开的方法和装置。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法和装置。用于所述方法的步骤的上述顺序仅是为了进行说明，本公开的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施例中，还可将本公开实施为记录在记录介质中的程序，这些程序包括用于实现根据本公开的方法的机器可读指令。因而，本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

本公开的描述是为了示例和描述起见而给出的，而并不是无遗漏的或者将本公开限于所公开的形式。很多修改和变化对于本领域的普通技术人员而言是显然的。选择和描述实施例是为了更好说明本公开的原理和实际应用，并且使本领域的普通技术人员能够理解本公开从而设计适于特定用途的带有各种修改的各种实施例。