(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
2022 年 2 月 17 日 (17.02.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/033306 A1

(51) 国际专利分类号:
G06T 7/246 (2017.01)

(21) 国际申请号: PCT/CN2021/108893

(22) 国际申请日: 2021 年 7 月 28 日 (28.07.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010805592.X 2020年8月12日 (12.08.2020) CN

(71) 申请人: 深圳市道通智能航空技术股份有限公司 (**AUTEL ROBOTICS CO., LTD.**) [CN/CN]; 中国广东省深圳市南山区西丽街道学苑大道1001号智园B1栋9层, Guangdong 518055 (CN)。

(72) 发明人: 李亚学 (**LI, Yaxue**); 中国广东省深圳市南山区西丽街道学苑大道1001号智园B1栋9层, Guangdong 518055 (CN)。

(74) 代理人: 深圳市六加知识产权代理有限公司 (**LIUJIA CHINA IP LAW OFFICE**); 中国广东省深圳市南山区桃源街道留仙大道塘岭路1号金骐智谷大厦403, Guangdong 518055 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, 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,

(54) **Title:** TARGET TRACKING METHOD AND APPARATUS

(54) **发明名称:** 目标跟踪方法和装置

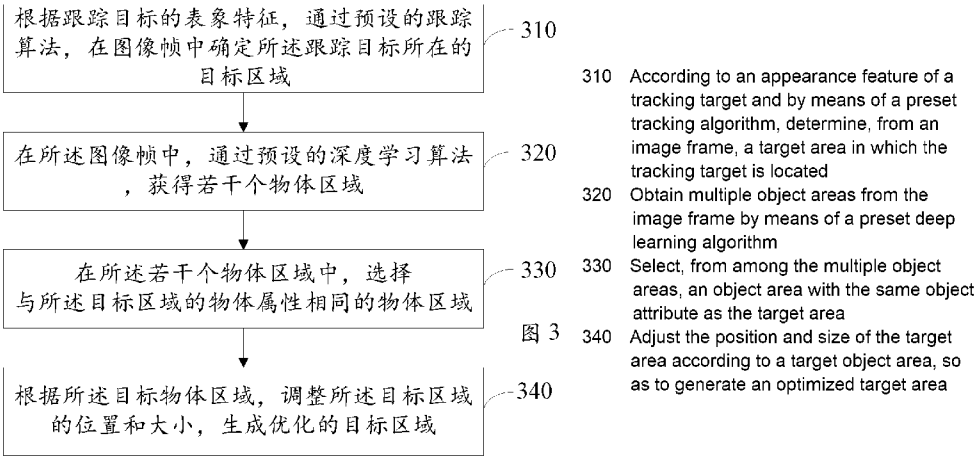

图 3

(57) **Abstract:** The embodiments of the present invention relate to a target tracking method and apparatus. The embodiments comprise: according to an appearance feature of a tracking target and by means of a preset tracking algorithm, determining, from an image frame, a target area in which the tracking target is located; obtaining multiple object areas from the image frame by means of a preset deep learning algorithm; selecting, from among the multiple object areas, an object area with the same object attribute as the target area, and using same as a target object area; and adjusting the position and size of the target area according to the target object area, so as to generate an optimized target area. In the present invention, on the basis of an original tracking algorithm, optimization and adjustment are performed in combination with a detection result of deep learning, such that interference in a complicated environment can be better adapted to and resisted, thereby effectively improving the overall performance of target tracking.

WO 2022/033306 A1

[见续页]

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))。

(57) 摘要: 本发明实施例涉及目标追踪方法和装置。这些实施例包括：根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域；在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域；在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域；根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。本发明在原有的跟踪算法的基础上，结合深度学习的检测结果进行优化和调整，可以更好的适应和抵抗复杂环境下的干扰，有效的提升了目标跟踪的整体性能。

目标跟踪方法和装置

本申请要求于2020年8月12日提交中国专利局、申请号为202010805592X、申请名称为“目标跟踪方法和装置”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】

本发明涉及机器视觉技术领域，尤其涉及一种目标跟踪方法、装置、图像处理芯片以及无人机。

【背景技术】

“目标跟踪”是在给定的视频序列的初始图像帧的目标大小与位置的情况下，预测该目标物体在视频序列后续的图像帧中的大小和位置的技术。其在视频监控、人机交互和多媒体分析等多个领域都有广泛的应用。

但在实际应用过程中，被跟踪目标由于非刚体运动，形状易发生变化，受到光照变换以及背景环境干扰等的情况，经常出现跟踪失败的情况，无法完成目标跟踪的任务。

因此，如何避免视频序列中无关因素对于跟踪目标的干扰，提升目标跟踪的性能，以满足和适应复杂多变的实际应用情况是一个迫切需要解决的问题。

【发明内容】

本发明实施例旨在提供一种目标跟踪方法、装置、图像处理芯片以及无人机，能够解决现有目标跟踪方法存在的缺陷。

为解决上述技术问题，本发明实施例提供以下技术方案：一种目标跟踪方法。该方法包括：

根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域；

在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域；

在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域；

根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。

可选地，在初始图像帧中，通过所述深度学习算法，识别获得若干个可选的物体区域，每一个所述物体区域标记有对应的物体标签；

确定用户选中的物体区域作为跟踪目标。

可选地，所述在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域，具体包括：

选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域；

在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

可选地，所述选取与所述目标区域重叠度最大的候选物体区域作为所述目标物体区域，具体包括：

计算所述候选物体区域与所述目标区域的交集面积和并集面积；

将所述交集面积和所述并集面积的比值作为所述候选物体区域与所述目标物体区域的重叠度。

可选地，所述表象特征包括：梯度方向直方图、局部二值模式和颜色特征。

可选地，所述根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域，具体包括：

设置所述目标物体区域的第一权重和所述目标区域的第二权重；

根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点；所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示，并且

根据所述第一权重和所述第二权重，对所述目标物体区域与所述目标区

域的大小进行加权求和，获得所述优化的目标区域的大小。

可选地，所述优化的目标区域的中心点通过如下算式计算获得：

center_x_opt = λ *center_x_track + (1- λ)*center_x_detect;

center_y_opt = λ *center_y_track + (1- λ)*center_y_detect;

其中，center_x_opt 为优化的目标区域的中心点在图像帧的水平方向上的坐标，center_y_opt 为优化的目标区域的中心点在图像帧的垂直方向上的坐标；center_x_track 为目标区域的中心点在图像帧的水平方向上的坐标，center_y_track 为目标区域的中心点在图像帧的垂直方向上的坐标，center_x_detect 为目标物体区域的中心点在图像帧的水平方向上的坐标，center_y_detect 为目标物体区域的中心点在图像帧的垂直方向上的坐标，λ 为第二权重。

可选地，所述优化的目标区域的大小通过如下算式计算获得：

width_opt = λ * width_track + (1- λ)* width_detect;

height_opt = λ *height_track + (1- λ)*height_detect;

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为第二权重。

为解决上述技术问题，本发明实施例还提供以下技术方案：一种目标跟踪装置，包括：

目标跟踪模块，用于根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域；

深度学习识别模块，用于在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域；

选择模块，用于在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域；

优化模块，用于根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。

为解决上述技术问题，本发明实施例还提供以下技术方案：一种图像处

理芯片，包括：处理器以及与所述处理器通信连接的存储器；所述存储器中存储有计算机程序指令，所述计算机程序指令在被所述处理器调用时，以使所述处理器执行如上所述的目标跟踪方法。

为解决上述技术问题，本发明实施例还提供以下技术方案：一种无人机，包括：无人机主体，安装在所述无人机主体的云台上的图像采集设备以及图像处理芯片；

所述图像采集设备用于连续采集多帧图像；所述图像处理芯片用于接收所述图像采集设备连续采集的多帧图像，并对接收到的所述多帧图像执行如上所述的目标跟踪方法，实现对跟踪目标的跟踪。

与现有技术相比较，本发明实施例的目标跟踪方法在原有的跟踪算法的基础上，结合深度学习的检测结果进行优化和调整，可以更好的适应和抵抗复杂环境下的干扰，有效的提升了目标跟踪的整体性能。

【附图说明】

一个或多个实施例通过与之对应的附图中的图片进行示例性说明，这些示例性说明并不构成对实施例的限定，附图中具有相同参考数字标号的元件表示为类似的元件，除非有特别申明，附图中的图不构成比例限制。

图 1 为本发明实施例的目标跟踪方法的应用场景的示意图；

图 2 为本发明实施例提供的目标跟踪装置的示意图；

图 3 为本发明实施例提供的目标跟踪方法的方法流程图；

图 4 为本发明实施例提供的目标物体区域选择方法的方法流程图；

图 5 为本发明实施例提供的目标跟踪方法的应用实例的示意图；

图 6 为本发明实施例提供的图像处理芯片的结构示意图。

【具体实施方式】

为了便于理解本发明，下面结合附图和具体实施例，对本发明进行更详细的说明。需要说明的是，当元件被表述“固定于”另一个元件，它可以直接在另一个元件上、或者其间可以存在一个或多个居中的元件。当一个元件被表述“连接”另一个元件，它可以是直接连接到另一个元件、或者其间可以存在一个或多个居中的元件。本说明书所使用的术语 “上”、“下”、“内”、“外”、“底

部”等指示的方位或位置关系为基于附图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。此外，术语“第一”、“第二”“第三”等仅用于描述目的，而不能理解为指示或暗示相对重要性。

除非另有定义，本说明书所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本说明书中在本发明的说明书中所使用的术语只是为了描述具体的实施例的目的，不是用于限制本发明。本说明书所使用的术语“和/或”包括一个或多个相关的所列项目的任意的和所有的组合。

此外，下面所描述的本发明不同实施例中所涉及的技术特征只要彼此之间未构成冲突就可以相互结合。

在由一系列连续的图像帧组成的视频序列中，传统的目标跟踪流程包括生成候选样本，特征提取，使用观测模型为候选样本进行评分，更新观测模型适应目标的变化以及融合获得最终的决策结果等步骤。

其中，“特征提取”是指提取具有鉴别性的特征来表示目标的过程。提取获得鉴别性特征是候选样本评分的基础，是决定目标跟踪性能的关键。现有关于目标跟踪方法性能的改进大多集中在如何选取合适的特征。

而本发明实施例提供的目标跟踪方法通过结合和调整优化的方式，在不付出过多的延时或者计算量的情况下，可以克服遮挡，形变，背景杂斑以及尺度变换等的干扰，提升目标跟踪的性能。

图 1 为本发明实施例提供的目标跟踪方法的应用场景。如图 1 所示，在该应用场景中，包括搭载了航拍相机的无人机 10、智能终端 20 以及无线网络 30。

无人机 10 可以是以任何类型的动力驱动的无人飞行载具，包括但不限于四轴无人机、固定翼飞行器以及直升机模型等。其可以根据实际情况的需要，具备相应的体积或者动力，从而提供能够满足使用需要的载重能力、飞行速度以及飞行续航里程等。

无人机 10 上可以搭载有任何类型的图像采集设备，包括运动相机、高清相机或者广角相机。其作为无人机上搭载的其中一种功能模块，可以通过云

台等安装固定支架，安装固定在无人机上，并受控于无人机 10，执行图像采集的任务。

当然，无人机上还可以添加有一种或者多种功能模块，令无人机能够实现相应的功能，例如内置的主控芯片，作为无人机飞行和数据传输等的控制核心或者是图传装置，将采集获得的图像信息上传至与无人机建立连接的设备（如服务器或者智能终端）中。

智能终端 20 可以是任何类型，用以与无人机建立通信连接的智能设备，例如手机、平板电脑或者智能遥控器等。该智能终端 20 可以装配有一种或者多种不同的用户交互装置，用以采集用户指令或者向用户展示和反馈信息。

这些交互装置包括但不限于：按键、显示屏、触摸屏、扬声器以及遥控操作杆。例如，智能终端 20 可以装配有触控显示屏，通过该触控显示屏接收用户对无人机的遥控指令并通过触控显示屏向用户展示由航拍相机获得的图像信息，用户还可以通过遥控触摸屏切换显示屏当前显示的图像信息。

在一些实施例中，无人机 10 与智能终端 20 之间还可以融合现有的图像视觉处理技术，进一步的提供更智能化的服务。例如无人机 10 可以通过航拍相机采集图像，然后由智能终端 20 执行本发明实施例提供的目标跟踪方法，对视频中的特定的人脸进行目标跟踪，最终实现用户与无人机之间的人机交互。

在另一些实施例中，该目标跟踪方法还可以由无人机 10 或者外置的服务器来执行，直接向智能终端 20 提供最终的数据结果。

无线网络 30 可以是基于任何类型的数据传输原理，用于建立两个节点之间的数据传输信道的无线通信网络，例如位于特定信号频段的蓝牙网络、WiFi 网络、无线蜂窝网络或者其结合，实现无人机 10，智能终端 20 和/或服务器之间的数据传输。

图 2 为本发明实施例提供的目标跟踪装置的结构框图。该目标跟踪装置可以由任何合适类型的电子计算平台所执行，例如内置在无人机中的图像处理芯片，与无人机建立无线通信连接的服务器或者智能终端。在本实施例中，以功能模块的方式描述该目标跟踪装置的组成。

本领域技术人员可以理解的是，图 2 所示的功能模块可以根据实际情况的需要，选择性的通过软件、硬件或者软件和硬件相结合的方式来实现。例

如，可以通过处理器调用存储器中存储的相关软件应用程序予以实现。

如图 2 所示，该目标跟踪装置 200 包括：目标跟踪模块 210，深度学习识别模块 220，选择模块 230 以及优化模块 240。

其中，目标跟踪模块 210 用于根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域。

“表象特征”是指在传统的目标跟踪方法中使用的，一些手工设计的鉴别性特征。其具有运算速度快，时延小等的特点。具体的，所述表象特征包括梯度方向直方图、局部二值模式和颜色特征。该目标跟踪模块 210 是执行传统的目标跟踪方法的功能模块，具体可以选择使用任何合适类型的跟踪算法。

深度学习识别模块 220 用于在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域。

“深度学习”是一种利用深度神经网络，通过样本数据训练而进行图像识别的方式。通过深度学习的方式，可以在图像帧中识别得到多个不同的物体区域。每个物体区域表示某个特定的物体。

选择模块 230 用于在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域。

“物体属性”是指用于判断物体区域是否与目标区域属于同一个物体的衡量标准。其具体可以由一种或者多种指标或者条件组成，基本上可以确定或者选择出属于同一个目标物体的物体区域。

在一些实施例中，所述选择模块 230 具体用于选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域；并且在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

优化模块 240 用于根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。

“目标物体区域”是基于深度学习方式获得的区域，对于复杂环境的干扰因素会有比较好的抵抗能力，不容易受到干扰。因此，在传统跟踪算法的基础上，引入目标物体区域作为参考，可以使得跟踪结果的效果得到优化，

生成优化的目标区域。

在一些实施例中，优化模块 240 具体用于：设置所述目标物体区域的第一权重和所述目标区域的第二权重；

根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点；并且根据所述第一权重和所述第二权重，对所述目标物体区域与所述目标区域的大小进行加权求和，获得所述优化的目标区域的大小。

其中，所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示。

具体的，一方面，所述优化的目标区域的中心点可以通过如下算式计算获得：

center_x_opt = λ *center_x_track + (1- λ)*center_x_detect;

center_y_opt = λ *center_y_track + (1- λ)*center_y_detect;

其中，center_x_opt 为优化的目标区域的中心点在图像帧的水平方向上的坐标，center_y_opt 为优化的目标区域的中心点在图像帧的垂直方向上的坐标；center_x_track 为目标区域的中心点在图像帧的水平方向上的坐标，center_y_track 为目标区域的中心点在图像帧的垂直方向上的坐标，center_x_detect 为目标物体区域的中心点在图像帧的水平方向上的坐标，center_y_detect 为目标物体区域的中心点在图像帧的垂直方向上的坐标，λ 为第二权重。

另一方面，所述优化的目标区域的大小则可以通过如下算式计算获得：

width_opt = λ * width_track + (1- λ)* width_detect;

height_opt = λ *height_track + (1- λ)*height_detect;

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为第二权重。

虽然，图 1 所示的应用场景中以应用在无人机搭载的图像采集设备为例。

但是，本领域技术人员可以理解的是，该目标跟踪方法还可以在其它类型的场景和设备中使用，以提高目标跟踪算法的性能。本发明实施例公开的目标跟踪方法并不限于在图 1 所示的无人机上应用。

在一些实施例中，如图 2 所示，该目标跟踪装置 200 还可以包括标记模块 250。该标记模块 250 用于在初始图像帧中，通过所述深度学习算法，识别获得若干个可选的物体区域，并确定用户选中的物体区域作为跟踪目标。其中，每一个所述物体区域标记有对应的物体标签。

其中，物体标签”（label）是深度学习算法输出的，用于标记物体区域所对应的物体。物体标签的具体形式取决于使用的深度学习算法及其训练数据。

基于深度学习算法的特点，在图像帧中可能识别出多个具有相同物体标签的物体区域。因此，可以据此作为候选的物体提供给用户选择想要进行跟踪的目标物体。

图 3 为本发明实施例提供的目标跟踪方法的方法流程图。如图 3 所示，该目标跟踪方法可以包括如下步骤：

310、根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域。

其中，“图像帧”是指视频序列中，正在被处理的某一帧图像。跟踪算法是以连续的图像帧组成的视频序列为处理对象，逐帧的预测追踪目标在图像中的位置。

跟踪算法可以采用现有技术中任何类型的，以表象特征作为追踪目标的鉴别性特征的快速跟踪算法。具体的，该表象特征包括但不限于梯度方向直方图（HOG）、局部二值模式（LBP）和颜色特征等。

不同的表象特征具有其自身独有的优势和缺陷，例如，梯度方向直方图容易受到目标发生非刚体运动（如作为追踪目标的人由站立变为蹲下的运动）和遮挡等情况的干扰。而颜色特征容易受到光照环境变化的影响。

目标区域是指采用“跟踪算法”计算后输出的，在图像帧中具有特定大小的，将跟踪目标包含在内的矩形框。其具体可以通过任何类型的跟踪算法计算获得。

320、在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域。

其中，“深度学习算法”可以是任何类型的，利用样本数据实现神经网络模型的图像处理方式。通过深度学习算法，可以具有较高置信度的获得在图像帧中存在的多个物体。

与目标区域的输出形式相类似的，深度学习算法输出的结果也是包含有可识别物体的矩形框。当然，深度学习算法还会输出每个物体区域对应的物体标签，用于标记物体区域所对应的目标物体（如人脸，飞机等）。

330、在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域。

其中，深度学习算法通常会在图像帧中识别出多于一个的物体区域。因此，需要采用合适的筛选方式，选择其中一个最为可靠的物体区域引入到跟踪结果中。

具体使用的筛选方式，取决于实际使用的物体属性。技术人员可以根据实际情况的需要，选择使用一种或者多种的衡量标准作为实际使用的物体属性。

在一些实施例中，具体可以采用如图 4 所示的步骤，从多个物体区域中选择获得作为调整和优化参考的目标物体区域：

331、选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域。

其中，“物体标签”（label）是深度学习算法输出的，用于标记物体区域所对应的物体。物体标签的具体形式取决于使用的深度学习算法及其训练数据。

332、在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

其中，由于深度学习算法的特点，在图像帧中可能识别出多个具有相同物体标签的物体区域。因此，在经过了物体标签的筛选以后，仍然有可能存在不止一个的候选物体区域。

此时，可以以候选区域与上一帧图像帧的跟踪结果的重叠度作为衡量标准，进一步的选择出可以作为调整参考和标准的目标物体区域。

在本实施例中，通过物体标签和重叠度两个衡量标准用于判断物体属性，可以有把握的从深度学习算法的输出结果中找到真实的跟踪目标，并据此作

为调整和优化跟踪结果的参考依据。

340、根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。

其中，“调整”是指采用任何合适类型的函数映射方式，将目标物体区域和目标区域进行整合，通过调整位置和大小的方式，生成并输出一个优化结果，即优化的目标区域。调整的形式包括改变和优化表示目标区域的矩形框在图像帧中的位置和大小。

在实际操作中，将一系列连续的图像帧序列中各个图像帧的优化的目标区域联系起来即可实现对跟踪目标的跟踪，确定跟踪目标在图像帧序列中的位置、尺寸等的变动情况。

在本实施例中，在原有的跟踪算法的基础上，通过合理的筛选方式，结合了深度学习的检测结果进行优化和调整，可以更好的抵抗复杂环境对于跟踪结果的干扰，有效的提升了目标跟踪的整体性能。

以下结合图 5 所示的目标跟踪方法在视频序列中的实际应用示意图，详细描述目标跟踪的具体过程。

如图 5 所示，假设视频序列由 n 帧连续的图像帧所组成（本发明实施例提供的目标跟踪方法依次在第 2 帧至第 n 帧图像帧中执行，预测跟踪目标在各个图像帧中的位置。）跟踪结果、物体区域以及目标区域都可以采用包含目标物体的最小外接矩形框来表示。

首先，在初始图像帧中，根据用户指令，确定所述跟踪目标所在的目标区域。

其中，初始图像帧是指进行目标跟踪的起点。为陈述简便，在本实施例中，初始图像帧为视频序列中的第 1 帧图像帧。当然，根据用户的实际需求，也可以在视频序列中随意选择任意一帧图像帧作为初始图像帧，用以作为目标跟踪的起点。

如图 5 所示，在实际使用过程中，为了便于与用户交互和后续处理，可以使用深度学习算法在初始图像帧中进行识别检测，获得若干个可选的物体区域(每一个所述物体区域标记有对应的物体标签,图 5 中以 L1 至 L4 中表示）作为可选的追踪目标。用户可以根据自己的需要，发出相应的用户选择指令，在这些可选的物体区域中选择其中的一个作为追踪目标（如 L4）。

然后，在后续的图像帧中，在通过深度学习算法检测获得多个的物体区域中，选出与跟踪目标的物体属性相同的目标物体区域。

具体的，对于检测获得的物体区域，首先选择与初始图像帧的追踪目标的物体标签相同的物体区域作为候选物体区域。然后，使用与上一帧图像帧的目标区域 D 的重叠度最高的物体区域作为目标物体区域。

其中，两个矩形框之间的重叠度可以通过物体区域与目标区域之间的交并比来表示。如图 5 所示，“交并比”（IoU）是指两个感兴趣区域的交集与并集之间的比值。其计算过程具体为：计算所述候选物体区域与所述目标区域的交集面积和并集面积，并将所述交集面积和所述并集面积的比值作为所述候选物体区域与所述目标物体区域的重叠度。

最后，将选择出的目标物体区域和跟踪算法检测获得的目标区域进行结合，即可获得在图像帧中的优化的目标区域。

只要依次在后续的图像帧中重复执行该上述步骤，即可获得在一系列连续的图像帧组成的视频序列中的目标追踪结果。总结而言，完整的目标追踪方法可以通过如下算式（1）表示：

$$\begin{cases} \boldsymbol{Bbox_{track}^{0_opt} = Bbox_{detect}^{choice_j}} & \boldsymbol{i = 0} \\ \boldsymbol{Bbox_{track}^{i_opt} = opt\left(Bbox_{track}^{i-1}, max_iou\left(Bbox_{track}^{i-1_opt}, \{Bbox_{detect}^{j}\}\right)\right)} & \boldsymbol{i > 0} \end{cases} \quad (1)$$

其中，Bbox 表示追踪目标，下标 track 和 detect 分别表示跟踪算法获得的目标区域和深度学习算法获得的物体区域。物体区域可能包含多个，分别使用$\{\boldsymbol{Bbox_{detect}^{j}}\}$表示（j 为物体区域的序号）。

在 i=0 时，表示在初始图像帧中选择的跟踪目标；i>0 时，表示在后续图像帧的跟踪结果。其中，$\boldsymbol{Bbox_{track}^{i}}$表示在第 i 帧图像帧的目标区域，$\boldsymbol{Bbox_{track}^{i_opt}}$表示第 i 帧图像帧的优化跟踪结果，$\boldsymbol{max_iou\left(Bbox_{track}^{i-1_opt}, \{Bbox_{detect}^{j}\}\right)}$表示目标物体区域（即与上一帧的目标区域的物体标签相同并且重叠度最大的物体区域）。

而对于追踪算法和深度学习算法的具体整合方法为：首先，设置所述目标物体区域的第一权重和所述目标区域的第二权重。然后，根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点并且根据所述第一权重和所述第二权重，对所述目标物体区域与所述目标区域的大小进行加权求和，获得所述优化的目标区域的大小。

其中，所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示，目标物体区域和目标区域的第一权重和第二权重可以由技术人员根据实际情况预先设置，是一个可以通过实验方式或者经验确定的常数值。在本实施例中，可以将目标区域的置信度作为跟踪算法所占的权重（第二权重）。相应地，深度学习算法所占的权重（即第一权重）为 1 减去目标区域的置信度。

具体而言，所述优化的目标区域的位置通过如下算式（2）和（3）计算获得：

$$center_x_opt = \lambda * center_x_track + (1-\lambda) * center_x_detect \quad (2);$$

$$center_y_opt = \lambda * center_y_track + (1-\lambda) * center_y_detect \quad (3);$$

其中，λ 为置信度。而且，由于在本实施例中，优化的目标区域、物体区域以及目标区域都是采用矩形框来表示。因此，优化的目标区域、物体区域以及目标区域的位置可以由矩形框的中心的在图像帧中的位置坐标来表示。亦即，优化的目标区域的位置可以表示为（center_x_opt，center_y_opt），物体区域的位置可以表示为（center_x_detect，center_y_detect），目标区域的位置可以表示为（center_x_track，center_y_track）。

所述优化的目标区域的大小则可以通过如下算式（4）和（5）计算获得：

$$width_opt = \lambda * width_track + (1-\lambda) * width_detect \quad (4);$$

$$height_opt = \lambda * height_track + (1-\lambda) * height_detect \quad (5);$$

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为置信度。

如上所描述的，本发明实施例提供的目标追踪方法在原有的跟踪算法的基础上，结合深度学习的检测结果进行优化和调整，可以更好的适应和抵抗复杂环境下的干扰，有效的提升了目标跟踪的整体性能。

本发明实施例还提供了一种非易失性计算机存储介质，所述计算机存储介质存储有至少一可执行指令，该计算机可执行指令可执行上述任意方法实

施例中的目标跟踪方法。

图 6 示出了本发明实施例的图像处理芯片的结构示意图，本发明具体实施例并不对图像处理芯片的具体实现做限定。

如图 6 所示，该图像处理芯片可以包括：处理器(processor)602、通信接口(Communications Interface)604、存储器(memory)606、以及通信总线 608。

其中：处理器 602、通信接口 604、以及存储器 606 通过通信总线 608 完成相互间的通信。通信接口 604，用于与其它设备比如客户端或其它服务器等的网元通信。处理器 602，用于执行程序 610，具体可以执行上述目标追踪方法实施例中的相关步骤。

具体地，程序 610 可以包括程序代码，该程序代码包括计算机操作指令。

处理器 602 可能是中央处理器 CPU，或者是特定集成电路 ASIC（Application Specific Integrated Circuit），或者是被配置成实施本发明实施例的一个或多个集成电路。网络切片设备包括的一个或多个处理器，可以是同一类型的处理器，如一个或多个 CPU；也可以是不同类型的处理器，如一个或多个 CPU 以及一个或多个 ASIC。

存储器 606，用于存放程序 610。存储器 606 可能包含高速 RAM 存储器，也可能还包括非易失性存储器（non-volatile memory），例如至少一个磁盘存储器。

程序 610 可以用于使得处理器 602 执行如下步骤：

首先，根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域。其次，在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域。再次，在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域。最后，根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域。

在一些实施例中，程序 610 还用于使得处理器 602 在执行通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域的步骤之前，执行如下步骤：首先，在初始图像帧中，通过所述深度学习算法，识别获得若干个可选的物体区域，每一个所述物体区域标记有对应的物体标签。然后，确定用户选中的物体区域作为跟踪目标。

在一些实施例中，程序 610 可以用于使得处理器 602 在执行所述在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域的步骤时，具体用于：

选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域；在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

在一些实施例中，程序 610 可以用于使得处理器 602 在执行所述选取与所述目标区域重叠度最大的候选物体区域作为所述目标物体区域的步骤时，具体用于：计算所述候选物体区域与所述跟踪结果的交集面积和并集面积；将所述交集面积和所述并集面积的比值，作为所述候选物体区域与所述跟踪结果的重叠度。

在一些实施例中，程序 610 可以用于使得处理器 602 在执行所述根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域的步骤时，具体用于：

设置所述目标物体区域的第一权重和所述目标区域的第二权重；根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点；所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示，并且根据所述第一权重和所述第二权重，对所述目标物体区域与所述目标区域的大小进行加权求和，获得所述优化的目标区域的大小。

具体而言，所述处理器 602 可以通过如下算式计算获得所述优化的目标区域的中心点：

center_x_opt = λ *center_x_track + (1- λ)*center_x_detect;

center_y_opt = λ *center_y_track + (1- λ)*center_y_detect;

其中，center_x_opt 为优化的目标区域的中心点在图像帧的水平方向上的坐标，center_y_opt 为优化的目标区域的中心点在图像帧的垂直方向上的坐标；center_x_track 为目标区域的中心点在图像帧的水平方向上的坐标，center_y_track 为目标区域的中心点在图像帧的垂直方向上的坐标，

center_x_detect 为目标物体区域的中心点在图像帧的水平方向上的坐标，center_y_detect 为目标物体区域的中心点在图像帧的垂直方向上的坐标，λ 为第二权重。

所述处理器 602 还可以通过如下算式计算获得所述优化的目标区域的大小：

width_opt = λ * width_track + (1- λ)* width_detect;

height_opt = λ *height_track + (1- λ)*height_detect;

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为第二权重。

本领域技术人员应该还可以进一步意识到，结合本文中所公开的实施例描述的示例性的目标追踪方法的各个步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。

本领域技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。所述的计算机软件可存储于计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体或随机存储记忆体等。

基于本发明实施例提供的图像处理芯片，本发明实施例还提供了一种无人机。其中，该无人机包括：无人机主体、安装在所述无人机主体的云台上的图像采集设备以及图像采集芯片。

所述图像采集设备用于连续采集多帧图像；所述图像处理芯片用于接收所述图像采集设备连续采集的多帧图像，并对接收到的所述多帧图像执行如下所述的步骤：

首先，根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域。其次，在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域。再次，在所述若干个物体区域中，选择与

所述目标区域的物体属性相同的物体区域作为目标物体区域。最后，根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域。

无人机的基于图像采集芯片在多个连续图像帧中确定的优化的目标区域，实现无人机对跟踪目标的跟踪。

在一些实施例中，所述图像处理芯片在图像帧中确定所述跟踪目标所在的目标区域的步骤之前，还执行如下步骤：

首先，在初始图像帧中，通过所述深度学习算法，识别获得若干个可选的物体区域，每一个所述物体区域标记有对应的物体标签。然后，确定用户选中的物体区域作为跟踪目标。

在一些实施例中，所述图像处理芯片在执行所述在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域的步骤时，具体用于：

选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域；在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

在一些实施例中，所述图像处理芯片在执行所述选取与所述目标区域重叠度最大的候选物体区域作为所述目标物体区域的步骤时，具体用于：计算所述候选物体区域与所述跟踪结果的交集面积和并集面积；将所述交集面积和所述并集面积的比值，作为所述候选物体区域与所述跟踪结果的重叠度。

在一些实施例中，所述图像处理芯片在执行所述根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域的步骤时，具体用于：

设置所述目标物体区域的第一权重和所述目标区域的第二权重；根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点；所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示，并且根据所述第一权重和所述第二权重，对所述目标物体区

域与所述目标区域的大小进行加权求和，获得所述优化的目标区域的大小。

具体而言，所述图像处理芯片可以通过如下算式计算获得所述优化的目标区域的中心点：

center_x_opt = λ *center_x_track + (1- λ)*center_x_detect;

center_y_opt = λ *center_y_track + (1- λ)*center_y_detect;

其中，center_x_opt 为优化的目标区域的中心点在图像帧的水平方向上的坐标，center_y_opt 为优化的目标区域的中心点在图像帧的垂直方向上的坐标；center_x_track 为目标区域的中心点在图像帧的水平方向上的坐标，center_y_track 为目标区域的中心点在图像帧的垂直方向上的坐标，center_x_detect 为目标物体区域的中心点在图像帧的水平方向上的坐标，center_y_detect 为目标物体区域的中心点在图像帧的垂直方向上的坐标，λ 为第二权重。

所述图像处理芯片还可以通过如下算式计算获得所述优化的目标区域的大小：

width_opt = λ * width_track + (1- λ)* width_detect;

height_opt = λ *height_track + (1- λ)*height_detect;

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为第二权重。

最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；在本发明的思路下，以上实施例或者不同实施例中的技术特征之间也可以进行组合，步骤可以以任意顺序实现，并存在如上所述的本发明的不同方面的许多其它变化，为了简明，它们没有在细节中提供；尽管参照前述实施例对本发明进行了详细的说明，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；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权利要求书

1、一种目标跟踪方法，其特征在于，包括：

根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域；

在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域；

在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域；

根据所述目标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域。

2、根据权利要求 1 所述的方法，其特征在于，在根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域的步骤之前，所述方法还包括：

在初始图像帧中，通过所述深度学习算法，识别获得若干个可选的物体区域，每一个所述物体区域标记有对应的物体标签；

确定用户选中的物体区域作为跟踪目标。

3、根据权利要求 2 所述的方法，其特征在于，所述在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域，具体包括：

筛选选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域；

在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域。

4、根据权利要求 3 所述的方法，其特征在于，所述选取与所述目标区域重叠度最大的候选物体区域作为所述目标物体区域，具体包括：

计算所述候选物体区域与所述跟踪结果的交集面积和并集面积；

将所述交集面积和所述并集面积的比值，作为所述候选物体区域与所述跟踪结果的重叠度。

5、根据权利要求 1 所述的方法，其特征在于，所述表象特征包括：梯度方向直方图、局部二值模式和颜色特征。

6、根据权利要求 1-5 任一项所述的方法，其特征在于，所述根据所述目

标物体区域，通过调整所述目标区域的位置和大小的方式，生成优化的目标区域，具体包括：

设置所述目标物体区域的第一权重和所述目标区域的第二权重；

根据所述第一权重和所述第二权重，对所述目标物体区域的中心点和所述目标区域的中心点进行加权求和，获得所述优化的目标区域的中心点；所述目标物体区域的位置由所述目标物体区域的中心点表示，所述目标区域的位置由所述目标区域的中心点表示，所述优化的目标区域的位置由所述优化的目标区域的中心点表示，并且

根据所述第一权重和所述第二权重，对所述目标物体区域与所述目标区域的大小进行加权求和，获得所述优化的目标区域的大小。

7、根据权利要求 6 所述的方法，其特征在于，所述优化的目标区域的中心点通过如下算式计算获得：

center_x_opt = λ *center_x_track + (1- λ)*center_x_detect；

center_y_opt = λ *center_y_track + (1- λ)*center_y_detect；

其中，center_x_opt 为优化的目标区域的中心点在图像帧的水平方向上的坐标，center_y_opt 为优化的目标区域的中心点在图像帧的垂直方向上的坐标；center_x_track 为目标区域的中心点在图像帧的水平方向上的坐标，center_y_track 为目标区域的中心点在图像帧的垂直方向上的坐标，center_x_detect 为目标物体区域的中心点在图像帧的水平方向上的坐标，center_y_detect 为目标物体区域的中心点在图像帧的垂直方向上的坐标，λ 为第二权重。

8、根据权利要求 6 所述的方法，其特征在于，所述优化的目标区域的大小通过如下算式计算获得：

width_opt = λ * width_track + (1- λ)* width_detect；

height_opt = λ *height_track + (1- λ)*height_detect；

其中，width_opt 为优化的目标区域的宽度，width_track 为目标区域的宽度，width_detect 为目标物体区域的宽度，height_opt 为优化的目标区域的高度，height_track 为目标区域的高度，height_detect 为目标物体区域的高度，λ 为第二权重。

9、一种目标跟踪装置，其特征在于，包括：

目标跟踪模块，用于根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域；

深度学习识别模块，用于在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域；

选择模块，用于在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域作为目标物体区域；

优化模块，用于根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域。

10、一种图像处理芯片，其特征在于，包括：处理器以及与所述处理器通信连接的存储器；

所述存储器中存储有计算机程序指令，所述计算机程序指令在被所述处理器调用时，以使所述处理器执行如权利要求 1-8 任一项所述的目标跟踪方法。

11、一种无人机，其特征在于，包括：无人机主体，安装在所述无人机主体的云台上的图像采集设备以及图像处理芯片；

所述图像采集设备用于连续采集多帧图像；所述图像处理芯片用于接收所述图像采集设备连续采集的多帧图像，并对接收到的所述多帧图像执行如权利要求 1-8 任一项所述的目标跟踪方法，实现对跟踪目标的跟踪。

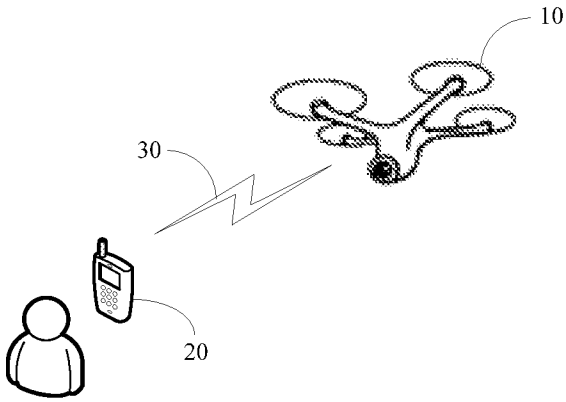
10
30
20

图 1

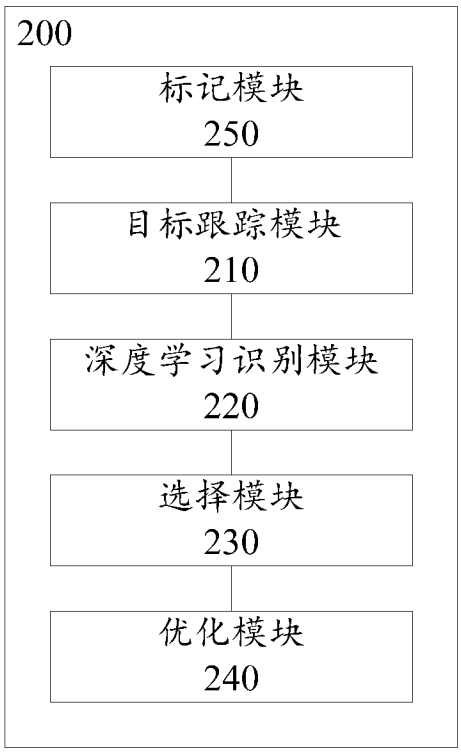
200
标记模块
250
目标跟踪模块
210
深度学习识别模块
220
选择模块
230
优化模块
240

图 2

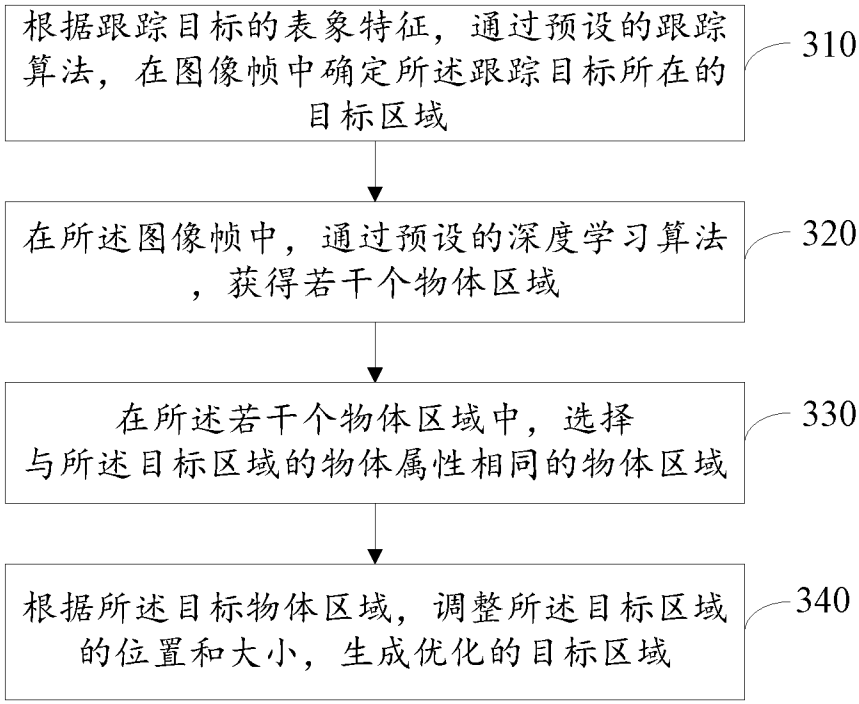
根据跟踪目标的表象特征，通过预设的跟踪算法，在图像帧中确定所述跟踪目标所在的目标区域
310
在所述图像帧中，通过预设的深度学习算法，获得若干个物体区域
320
在所述若干个物体区域中，选择与所述目标区域的物体属性相同的物体区域
330
根据所述目标物体区域，调整所述目标区域的位置和大小，生成优化的目标区域
340

图 3

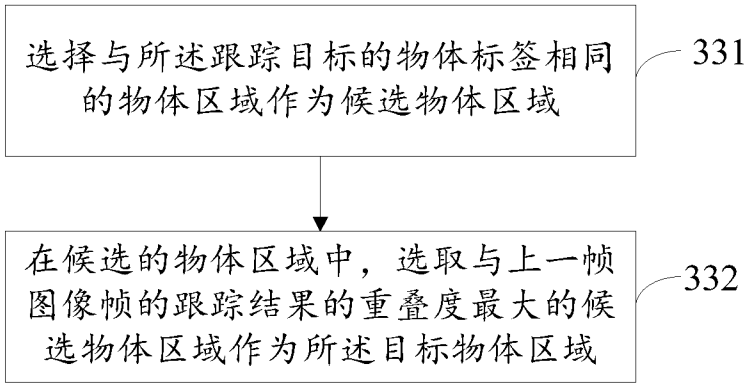
选择与所述跟踪目标的物体标签相同的物体区域作为候选物体区域
331
在候选的物体区域中，选取与上一帧图像帧的跟踪结果的重叠度最大的候选物体区域作为所述目标物体区域
332

图 4

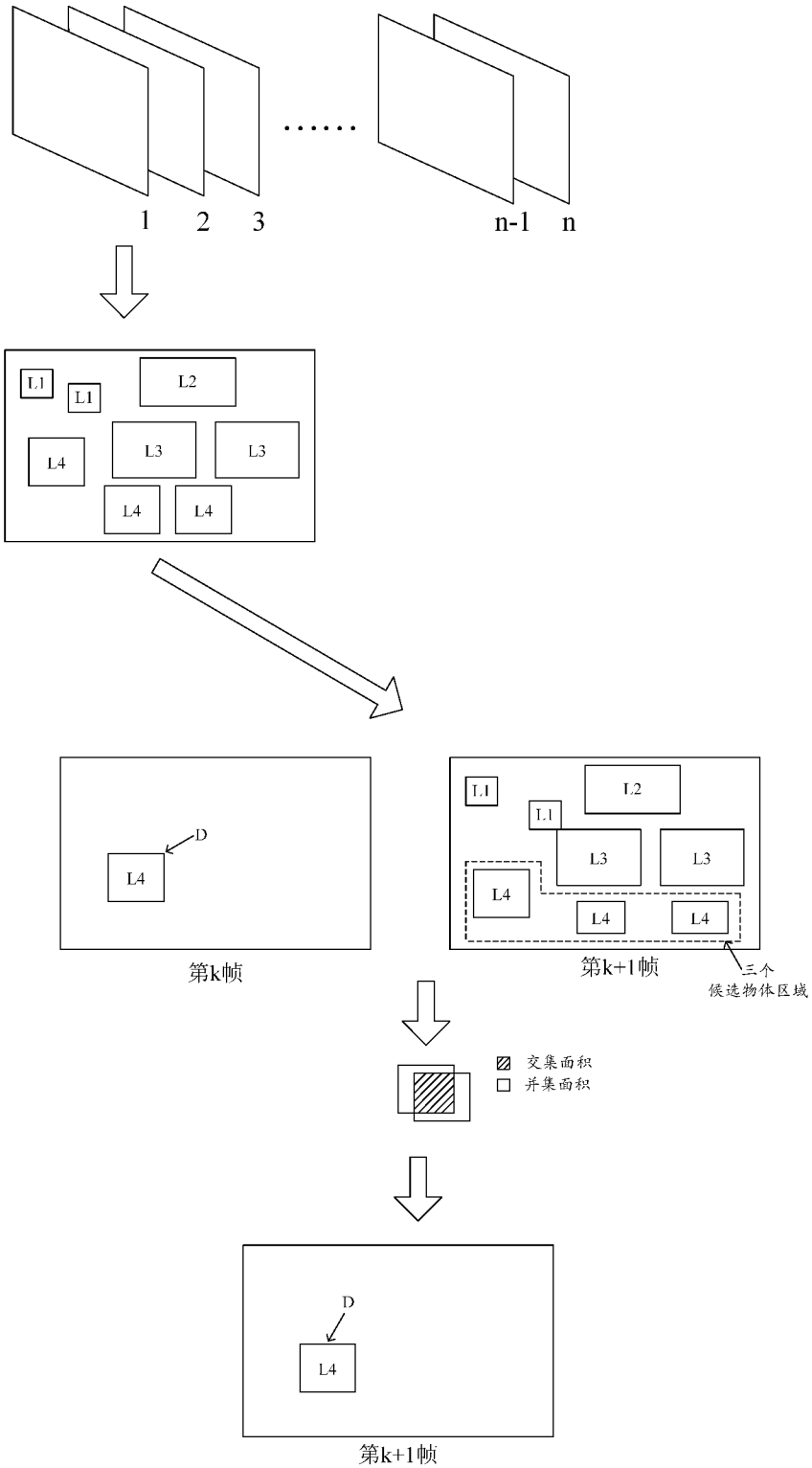
1 2 3 ..... n-1 n
L1
L1
L2
L4
L3
L3
L4
L4
D
L4
第k帧
L1
L1
L2
L3
L3
L4
L4
L4
第k+1帧
三个
候选物体区域
交集面积
并集面积
D
L4
第k+1帧

图 5

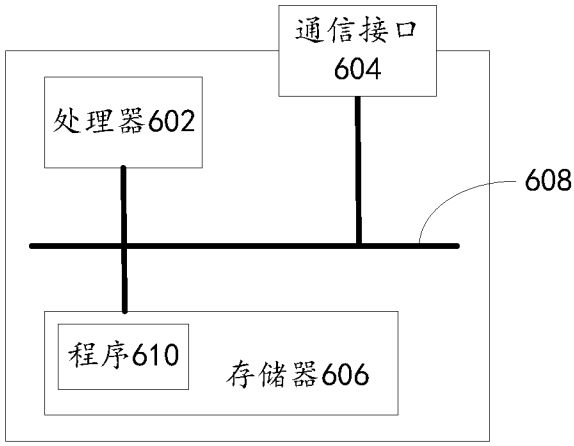

图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2021/108893

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/246(2017.01)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)

G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; CNKI; SIPOABS; DWPI; USTXT; WOTXT; EPTXT: 深度, 卷积, 神经网络, 目标, 跟踪, 特征, 梯度方向直方图, 局部二值模式, 颜色, 混合, 融合, 属性, 重叠度, 重合度, 交并比, depth, convolution, neural network, object, tracking, feature, HOG, LBP, color, mix, fusion, attribute, overlap ratio, IOU

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112037255 A (SHENZHEN AUTEL INTELLIGENT AVIATION TECHNOLOGY CO., LTD.) 04 December 2020 (2020-12-04) claims 1-11, description, paragraphs [0005]-[0180], figures 1-6	1-11
X	CN 109993769 A (ANHUI CREARO TECHNOLOGY CO., LTD.) 09 July 2019 (2019-07-09) description, paragraphs [0005]-[0066], figure 1	1-11
A	卢宪凯 (LU, Xiankai). "基于深度学习的目标跟踪研究 (Research on Object Tracking Based on Deep Learning)" *中国博士学位论文全文数据库 信息科技辑 (Electronic Technology & Information Science, China Doctoral Dissertations Full-Text Database),* No. 01, 15 January 2020 (2020-01-15), ISSN: 1674-022X, full text, pages 30-59	1-11
A	CN 109785385 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 21 May 2019 (2019-05-21) entire document	1-11
A	CN 109409283 A (SHENZHEN JINRUN DEFENSE TECHNOLOGY CO., LTD.) 01 March 2019 (2019-03-01) entire document	1-11

☑ 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	Date of mailing of the international search report
09 October 2021	**20 October 2021**
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

International application No.
PCT/CN2021/108893

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 111098815 A (WUHAN ZHONGXIANG TECHNOLOGY CO., LTD.) 05 May 2020 (2020-05-05) entire document	1-11

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2021/108893

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 112037255 A	04 December 2020	None	
CN 109993769 A	09 July 2019	None	
CN 109785385 A	21 May 2019	CN 109785385 B	29 January 2021
CN 109409283 A	01 March 2019	None	
CN 111098815 A	05 May 2020	CN 111098815 B	12 January 2021

国际检索报告

国际申请号
PCT/CN2021/108893

A. 主题的分类

G06T 7/246(2017.01)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)

G06T

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））

CNABS;CNTXT;CNKI;SIPOABS;DWPI;USTXT;WOTXT;EPTXT：深度，卷积，神经网络，目标，跟踪，特征，梯度方向直方图，局部二值模式，颜色，混合，融合，属性，重叠度，重合度，交并比，depth，convolution，neural network，object，tracking，feature，HOG，LBP，color，mix，fusion，attribute，overlap ratio，IOU

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 112037255 A (深圳市道通智能航空技术有限公司) 2020年 12月 4日 (2020-12-04) 权利要求第1-11项，说明书第[0005]-[0180]段，图1-6	1-11
X	CN 109993769 A (安徽创世科技股份有限公司) 2019年 7月 9日 (2019-07-09) 说明书第[0005]-[0066]段，图1	1-11
A	卢宪凯. "基于深度学习的目标跟踪研究" 中国博士学位论文全文数据库 信息科技辑，第01期，2020年 1月 15日 (2020-01-15)， ISSN：1674-022X, 正文第30-59页	1-11
A	CN 109785385 A (中国科学院自动化研究所) 2019年 5月 21日 (2019-05-21) 全文	1-11
A	CN 109409283 A (深圳市锦润防务科技有限公司) 2019年 3月 1日 (2019-03-01) 全文	1-11
A	CN 111098815 A (武汉市众向科技有限公司) 2020年 5月 5日 (2020-05-05) 全文	1-11

☐ 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	国际检索报告邮寄日期
2021年 10月 9日	2021年 10月 20日

ISA/CN的名称和邮寄地址	受权官员
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	李慧
传真号 (86-10)62019451	电话号码 （86-512）88995679

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2021/108893

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	112037255	A	2020年 12月 4日	无			
CN	109993769	A	2019年 7月 9日	无			
CN	109785385	A	2019年 5月 21日	CN	109785385	B	2021年 1月 29日
CN	109409283	A	2019年 3月 1日	无			
CN	111098815	A	2020年 5月 5日	CN	111098815	B	2021年 1月 12日