

(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



(10) 国际公布号

WO 2022/222227 A1

(43) 国际公布日

2022 年 10 月 27 日 (27.10.2022)

(51) 国际专利分类号:

G06K 9/32 (2006.01)

G06T 7/246 (2017.01)

G06K 9/62 (2006.01)

G06T 7/277 (2017.01)

(21) 国际申请号:

PCT/CN2021/096757

(22) 国际申请日:

2021 年 5 月 28 日 (28.05.2021)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

202110434628.2

2021 年 4 月 22 日 (22.04.2021) 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: 赵娅琳(ZHAO, Yalin); 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

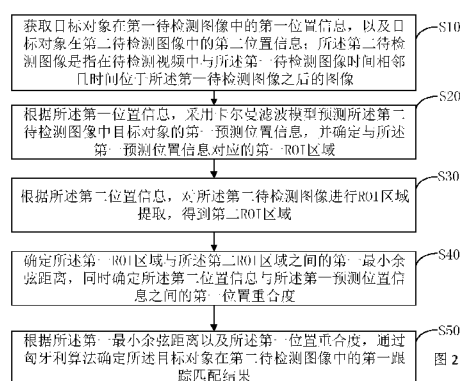
安金融中心 23 楼, Guangdong 518000 (CN)。 陆进(LU, Jin); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 刘玉宇(LIU, Yuyu); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳众鼎专利商标代理事务所(普通合伙)(SHENZHEN ZHONGDING INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区笋岗街道笋岗东路 3012 号中民时代广场 B 座 701, Guangdong 518000 (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,

(54) Title: TARGET DETECTION NETWORK-BASED TARGET TRACKING METHOD AND APPARATUS, DEVICE, AND MEDIUM

(54) 发明名称: 基于目标检测网络的目标跟踪方法、装置、设备及介质



S10 Obtain first position information of target object in first image to undergo detection, and second position information of target object in second image to undergo detection; second image to undergo detection is image in video to undergo detection that is temporally adjacent to first image to undergo detection and is located later in time than first image to undergo detection

S20 Utilize Kalman filter model to predict first predicted position information of target object in second image to undergo detection according to first position information, and determine first ROI region corresponding to first predicted position information

S30 Perform ROI region extraction on second image to undergo detection according to second position information, and obtain second ROI region

S40 Determine first least cosine distance between first ROI region and second ROI region, and simultaneously determine first positional coincidence degree between second position information and first predicted position information

S50 Determine first tracking matching result of target object in second image to undergo detection by means of Hungarian algorithm and according to first least cosine distance and first positional coincidence degree

(57) Abstract: A target detection network-based target tracking method and apparatus, a device, and a medium. The method is performed by means of obtaining first position information of a target object in a first image to undergo detection, and second position information of the target object in a second image to undergo detection (S10); utilizing a Kalman filter model to predict first predicted position information of the target object in the second image to undergo detection according to the first position information, and determining a first ROI region corresponding to the first predicted position information (S20); performing ROI region extraction on the second image to undergo detection according to the second position information, and obtaining a second ROI region (S30); determining a first least cosine distance between the first ROI region and the second ROI region, and simultaneously determining a first positional coincidence degree between the second position information and the first predicted position information (S40); and determining a first tracking matching result by means of the Hungarian algorithm and according to the first least cosine distance and the first positional coincidence degree (S50). The efficiency and accuracy of target tracking are improved.



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, 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: 一种基于目标检测网络的目标跟踪方法、装置、设备及介质。该方法通过获取目标对象在第一待检测图像中的第一位置信息, 以及目标对象在第二待检测图像中的第二位置信息 (S10); 根据第一位置信息, 采用卡尔曼滤波模型预测第二待检测图像中目标对象的第一预测位置信息, 并确定与第一预测位置信息对应的第一ROI区域 (S20); 根据第二位置信息对第二待检测图像进行ROI区域提取得到第二ROI区域 (S30); 确定第一ROI区域与第二ROI区域之间的第一最小余弦距离, 同时确定第二位置信息与第一预测位置信息之间的第一位置重合度 (S40); 根据第一最小余弦距离以及第一位置重合度, 通过匈牙利算法确定第一跟踪匹配结果 (S50)。提高了目标跟踪的准确率以及效率。

基于目标检测网络的目标跟踪方法、装置、设备及介质

本申请要求于 2021 年 4 月 22 日提交中国专利局、申请号为 202110434628.2，发明名称为“基于目标检测网络的目标跟踪方法、装置、设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及检测模型技术领域，尤其涉及一种基于目标检测网络的目标跟踪方法、装置、设备及介质。

背景技术

多目标跟踪技术应用于多个应用领域，例如运动校正，安防监控，无人驾驶领域等。而在安防监控系统中，对目标进行准确的定位与跟踪是常见的任务。

发明人意识到，由于需要跟踪的对象往往同时存在多个，各个跟踪对象之间外观相似度较为接近，现有技术中的安防监控技术无法单从外观特征判断目标身份，造成检测目标和跟踪轨迹错误匹配，影响多目标跟踪的准确性；进一步地，在目标跟踪过程中还会存在跟踪对象之间存在遮挡、尺度变化等情况，且无法判别跟踪对象是因遮挡暂时消失还是已经离开检测区域，导致多目标跟踪的准确率较低。

申请内容

本申请实施例提供一种基于目标检测网络的目标跟踪方法、装置、设备及介质，以解决多目标跟踪的准确率较低的问题。

一种基于目标检测网络的目标跟踪方法，包括：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

一种基于目标检测网络的目标跟踪装置，包括：

第一位置信息获取模块，用于获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

第一位置信息预测模块，用于根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

第一 ROI 区域提取模块，用于根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

第一位置重合度确定模块，用于确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重

合度；

第一跟踪匹配模块，用于根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

一种计算机设备，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：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

上述基于目标检测网络的目标跟踪方法、装置、设备及介质，该方法通过获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

本申请通过引入目标检测网络以及卡尔曼滤波模型，可以较好的将目标检测网络中的浅层特征作为目标跟踪的外观特征，进而通过卡尔曼滤波模型根据目标检测网络确定的外观特征，使得可以通过外观特征预测下一帧待检测图像中目标对象所处位置信息，且计算速度较快，提高了目标跟踪的准确率。

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

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请一实施例中基于目标检测网络的目标跟踪方法的一应用环境示意图；

图 2 是本申请一实施例中基于目标检测网络的目标跟踪方法的一流程图；

图 3 是本申请一实施例中基于目标检测网络的目标跟踪方法的一流程图；

图 4 是本申请一实施例中基于目标检测网络的目标跟踪方法的一流程图；

图 5 是本申请一实施例中基于目标检测网络的目标跟踪装置的一原理框图；

图 6 是本申请一实施例中基于目标检测网络的目标跟踪装置中第一位置信息获取模块的一原理框图；

图 7 是本申请一实施例中基于目标检测网络的目标跟踪装置中目标检测子模块的一原理框图；

图 8 是本申请一实施例中计算机设备的一示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请实施例提供的基于目标检测网络的目标跟踪方法，该基于目标检测网络的目标跟踪方法可应用如图 1 所示的应用环境中。具体地，该基于目标检测网络的目标跟踪方法应用在基于目标检测网络的目标跟踪系统中，该基于目标检测网络的目标跟踪系统包括如图 1 所示的客户端和服务端，客户端与服务端通过网络进行通信，用于解决多目标跟踪的准确率较低的问题。其中，客户端又称为用户端，是指与服务端相对应，为客户提供本地服务的程序。客户端可安装在但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备上。服务端可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在一实施例中，如图 2 所示，提供一种基于目标检测网络的目标跟踪方法，以该方法应用在图 1 中的服务器为例进行说明，包括如下步骤：

S10: 获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像。

可以理解地，第一待检测图像以及第二待检测图像均可以为待检测视频中任意连续两帧的图像（如假设第一待检测图像为待检测视频中的第一帧图像，则第二待检测图像为待检测视频中的第二帧图像），待检测视频可以根据具体应用场景进行选取，如在智能监控领域中，该第一待检测图像以及第二待检测图像可以为从监控视频中选取的连续两帧的图像。进一步地，第一位置信息指的是目标对象在第一待检测图像中出现的位置（示例性地，该第一位置信息还可以包括目标对象的运动方向，运动速度等信息），该位置可以在对待检测图像进行目标检测之后，通过与目标对象对应的所有图像块的集合表征；同理，第二位置信息是指目标对象在第二待检测图像中出现的位置；可选地，目标对象可以为一个目标个体，也可以为多个目标个体。

在一实施例中，如图 3 所示，步骤 S10 中，包括：

S101: 获取待检测视频，所述待检测视频中包含多帧待检测图像。

S102: 将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像。

可以理解地, 待检测视频中包含多帧的待检测图像, 由于待检测视频中待检测图像帧速一般为每秒 25 帧以上, 且连续两帧的待检测图像可能较为接近, 目标对象出现变化的可能性不大, 因此在对待检测视频中的各待检测图像进行目标检测时, 可以通过间隔获取待检测图像的方式, 也即可以自待检测视频中间隔 n 帧获取待检测图像, 进而将获取的各待检测图像按照时间顺序形成合成视频; 在该合成视频中可以任意选取一个待检测图像作为第一待检测图像, 而此时第二待检测图像则为合成视频中与被选取的第一待检测图像相邻, 且处于该第一待检测图像之后的待检测图像 (该第二待检测图像在待检测视频中并不是与第一待检测图像相邻且处于第一待检测图像之后的图像, 而是与第一待检测图像间隔 n 帧之后的待检测图像)。其中, n 可以根据具体应用场景进行选取, 示例性地, n 可以取 4, 5 等。进而通过间隔获取待检测图像的方式可以提高对待检测图像进行目标检测的效率, 减少计算机的计算压力。进一步地, 若不考虑计算机的计算压力, 则也可以对待检测视频中所有待检测图像进行目标检测。

S103: 通过目标检测网络对所述第一待检测图像进行目标检测, 得到所述第一位置信息; 同时通过目标检测网络对所述第二待检测图像进行目标检测, 得到所述第二位置信息。

具体地, 在将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像, 根据该第一待检测图像确定第二待检测图像, 以将第一待检测图像以及第二待检测图像依次输入至目标检测网络中, 进而通过目标检测网络对该第一待检测图像进行目标检测, 得到第一位置信息; 通过目标检测网络对第二待检测图像进行目标检测, 得到第二位置信息。

在一具体实施例中, 如图 4 所示, 步骤 S103 中, 通过目标检测网络对所述第一待检测图像进行目标检测, 得到所述第一位置信息, 包括:

S1031: 将所述第一待检测图像输入至所述目标检测网络中的骨干网络, 以对所述第一待检测图像进行下采样处理, 得到与所述第一待检测图像对应的多个待检测特征图层。

可选地, 骨干网络可以采用 Darknet 网络、Resnet 网络等。下采样处理指的是对第一待检测图像进行缩小处理, 也即进行下采样后得到的图像的尺寸小于第一待检测图像的尺寸。

具体地, 在确定第一待检测图像以及第二待检测图像之后, 将第一待检测图像输入至目标检测网络中, 通过目标检测网络中的骨干网络对第一待检测图像进行下采样处理, 在本实施例中, 通过骨干网络对第一待检测图像进行五次下采样处理, 以得到五种不同尺寸的待检测特征图层。示例性地, 如对第一待检测图像进行第一次采样之后得到的待检测特征图层, 该待检测特征图层与第一待检测图像相比, 该待检测特征图层的尺寸为第一待检测图像的尺寸的一半 (也即长度以及宽度均缩小一半), 但该待检测特征图层的通道数为第一待检测图像的通道数的两倍。

进一步地, 由于第一次对第一待检测图像进行下采样得到的待检测特征图层的语义较少, 因此在本实施例中, 通过骨干网络对第一待检测图像进行五次下采样处理之后, 会弃用第一下采样得到的待检测特征图层, 进而本实施例中最终采用的待检测特征图层的总数量为四个。

S1032: 对各所述待检测特征图层依次进行图层处理, 得到与所述第一待检测图像对应的目标特征图层。

在一具体实施方式中, 所述待检测特征图层中包括第一特征图层、第二特征图层、第三特征图层以及第四特征图层; 步骤 S1032 中, 包括:

对所述第四特征图层进行卷积处理, 并对卷积处理后的第四特征图层进行上采样, 得到与第三特征图层具有相同维度的第五特征图层。

可以理解地, 第四特征图层是待检测图像进行五次下采样中最后一次下采样得到的

图层，也即该第四特征图层的尺寸最小；具体地，在将所述第一待检测图像输入至所述目标检测网络中的骨干网络，以对所述第一待检测图像进行下采样处理，得到与所述第一待检测图像对应的多个待检测特征图层之后，对第四特征图层进行卷积处理，如通过卷积核为 3×3 的卷积网络对第四特征图层进行卷积处理，使得第四特征图层与第三特征图层的通道数相同（步骤 S1031 中指出每经过一次下采样之后得到待检测特征图层与下采样之前的待检测特征图层相比，尺寸减半，通道数加倍），并对卷积处理后的第四特征图层进行上采样，使得上采样以及卷积处理后的第四特征图层，与第三特征图层具有相同的尺寸以及相同的通道数，从而得到与第三特征图层具有相同维度的第五特征图层。

将所述第五特征图层与所述第三特征图层进行维度叠加得到第一叠加图层之后，对所述第一叠加图层进行卷积处理，并对卷积处理后的第一叠加图层进行上采样，得到与第二特征图层具有相同维度的第六特征图层。

具体地，在对所述第四特征图层进行卷积处理，并对卷积处理后的第四特征图层进行上采样，得到与第三特征图层具有相同维度的第五特征图层之后，将第五特征图层与第三特征图层进行维度叠加得到第一叠加图层之后，对第一叠加图层进行卷积处理，使得第一叠加图层与第二特征图层的通道数相同，并对卷积后的第一叠加图层进行上采样，使得上采样以及卷积处理后的第一叠加图层，与第二特征图层具有相同的尺寸以及相同的通道数，从而得到与第二特征图层具有相同维度的第六特征图层。

将所述第六特征图层与所述第二特征图层进行维度叠加得到第二叠加图层之后，对所述第二叠加图层进行卷积处理，并对卷积处理后的第二叠加图层进行上采样，得到与第一特征图层具有相同维度的第七特征图层。

具体地，在对所述第一叠加图层进行卷积处理，并对卷积处理后的第一叠加图层进行上采样，得到与第二特征图层具有相同维度的第六特征图层之后，将第六特征图层与第二特征图层进行维度叠加得到第二叠加图层之后，对第二叠加图层进行卷积处理，使得第二叠加图层与第一特征图层的通道数相同，并对卷积后的第一叠加图层进行上采样，使得上采样以及卷积处理后的第一叠加图层，与第一特征图层具有相同的尺寸以及相同的通道数，从而得到与第一特征图层具有相同维度的第七特征图层。

将所述第七特征图层与所述第一特征图层进行维度叠加得到第三叠加图层之后，对所述第三叠加图层进行卷积处理，并对卷积处理后的第三叠加图层进行上采样，得到所述目标特征图层。

具体地，对所述第二叠加图层进行卷积处理，并对卷积处理后的第二叠加图层进行上采样，得到与第一特征图层具有相同维度的第七特征图层之后，将第七特征图层与第一特征图层进行维度叠加得到第三叠加图层之后，对第三叠加图层进行卷积处理，使得第三叠加图层的通道数加倍，并对卷积处理后的第三叠加图层进行上采样，使得卷积处理后的第三叠加图层的尺寸加倍，从而得到目标特征图层。

S1303: 对所述目标特征图层进行位置特征提取，得到所述第一位置信息。

具体地，在对各所述待检测特征图层依次进行图层处理，得到与所述第一待检测图像对应的目标特征图层之后，对目标特征图层进行位置特征提取，也即在目标特征图层上，提取与目标对象关联的像素框，进而得到目标对象在第一待检测图像上的第一位置信息。

S20: 根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域。

可以理解地，卡尔曼滤波模型是采用卡尔曼滤波的状态估计模型，该卡尔曼滤波模型用于根据目标对象在前一帧待检测图像（如第一待检测图像）上的位置信息（如第一位置信息），对目标对象处于下一帧待检测图像（如第二待检测图像）的位置信息进行预测。进一步地，该卡尔曼滤波模型需要通过待检测视频中前 k 帧的待检测图像对其进行训练，使得卡尔曼滤波模型在步骤 S20 中运动时能够较好的预测出目标对象的位置信息，提高目

标跟踪的准确率。

具体地，在获取第一待检测图像中目标对象所处的第一位置信息之后，根据该第一位置信息，采用卡尔曼滤波模型预测目标对象在第二待检测图像中所处的位置信息，也即第一预测位置信息，并在第二待检测图像中，提取与该第一预测位置信息关联的区域，也即第一 ROI 区域。

S30: 根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域。

具体地，在获取第二待检测图像中目标对象所处的第二位置信息之后，根据该第二位置信息，在第二待检测图像中提取与第二位置信息关联的区域，得到第二 ROI 区域。

S40: 确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度。

可以理解地，第一最小余弦距离用于表征第一 ROI 区域与第二 ROI 区域之间的特征相似度；第一位置重合度用于表征第二位置信息与第一预测位置信息之间的位置相似程度。

具体地，在根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域之后，确定第一 ROI 区域与第二 ROI 区域之间的第一最小余弦距离，该第一最小余弦距离的取值范围可以为 0 至 1，第一最小余弦距离越大，表征第一 ROI 区域与第二 ROI 区域之间的特征相似程度越高；同时确定第二位置信息与第一预测位置信息之间的第一位置重合度，该第一位置重合度的取值范围也可以为 0 至 1，第一位置重合度越高，表征第二位置信息与第一预测位置信息之间的相关程度越大。

在一实施例中，步骤 **S40** 中，所述确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度，包括：

确定所述第二位置信息与所述第一预测位置信息之间的交集位置信息，同时确定所述第二位置信息与所述第一预测位置信息之间的并集位置信息。

可以理解地，交集位置信息指的是第二位置信息与第一预测位置信息之间的共有位置信息；并集位置信息指的是第二位置信息与第一预测位置信息的所有位置信息，也即包括共有位置信息，以及独自拥有的位置信息。

根据所述交集位置信息以及所述并集位置信息，确定所述位置重合度。

具体地，可以根据如下述表达式确定所述位置重合度：

$$C = \frac{|A \cup B|}{|A \cap B|}$$

其中， C 为位置重合度； A 为第二位置信息； B 为第一预测位置信息； $|A \cup B|$ 为并集位置信息； $|A \cap B|$ 为交集位置信息。

S50: 根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

可以理解地，第一跟踪匹配结果可以为表征匹配成功的结果，也即表征第一预测位置信息与第二位置信息相匹配（如第一预测位置信息中包含目标对象的特征，且第二位置信息中也包含目标对象的特征）；也可以为表征匹配失败的结果，也即表征第一预测位置信息与第二位置信息不匹配（如第一预测位置信息中包含目标对象的特征，而第二位置信息中未包含目标对象的特征；亦或者第一预测位置信息中未包含目标对象的特征，而第二位置信息中包含目标对象的特征）。

具体地，在确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度之后，将第一最

小余弦距离以及第一位置重合度作为跟踪代价，通过匈牙利算法确定目标对象在第二待检测图像中的第一跟踪匹配结果。示例性地，如将第一最小余弦距离与预设余弦阈值（如设定为 0.9, 0.95 等）进行比较，在第一最小余弦距离大于或等于预设余弦阈值时，则将第一位置重合度与预设位置重合阈值（如设定为 0.9, 0.95 等）进行比较，在第一位置重合度大于或等于预设位置重合阈值时，则确定第一跟踪匹配结果为匹配成功结果；若第一最小余弦距离小于预设余弦阈值，和/或，第一位置重合度小于预设位置重合阈值，则确定第一跟踪匹配结果为匹配失败结果。

进一步地，在通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果之后，不论该第一跟踪匹配结果是匹配成功结果，亦或者是匹配失败结果，均可以将该第一跟踪匹配结果，以及对应的第一待检测图像以及第二待检测图像作为训练样本，用于训练卡尔曼滤波模型，提高卡尔曼滤波模型的预测准确率。

进一步地，为了提高目标检测网络的准确率，可以采用如下损失函数对目标检测网络进行约束：

$$L = L_1 + \lambda_1 L_2 + \lambda_2 L_3$$

其中， L 为目标检测网络的损失函数； L_1 为目标检测网络的焦点损失函数； L_2 为位置损失函数； L_3 为像素偏移损失函数； λ_1 为位置损失函数的权重（可以取值为 1）； λ_2 为像素偏移损失函数的权重（可以取值为 0.1）。

进一步地， L_1 可以通过下述表达式表征：

$$L_1 = -\frac{1}{N} \sum_m \left(1 - \hat{Y}_m \right)^\alpha \log \left(\hat{Y}_m \right) \Lambda \quad \text{if } Y_m = 1$$

$$L_1 = -\frac{1}{N} \sum_m (1 - Y_m)^\beta \left(\hat{Y}_m \right)^\alpha \log \left(1 - \hat{Y}_m \right) \Lambda \quad \text{otherwise}$$

其中， N 为目标对象中目标个体的总个数； $\hat{Y}_m$ 为第一预测位置信息； Y_m 为第二位置信息； α 以及 β 为目标检测网络的检测参数，可以根据具体地应用场景进行设定；
if $Y_m = 1$ 表征第二位置信息中包含目标对象中第 m 个目标个体；*otherwise* 表征第二位置信息中未包含目标对象中第 m 个目标个体。

进一步地， L_2 可以通过下述表达式表征：

$$L_2 = \frac{|A \cup B|}{|A \cap B|} - \frac{|A_c - |A \cup B||}{|A_c|}$$

其中， A 为第二位置信息； B 为第一预测位置信息； $|A \cup B|$ 为并集位置信息； $|A \cap B|$ 为交集位置信息； $|A_c|$ 为最小闭包面积。

进一步地， L_3 指的是步骤 S1031 中对第一待检测图像进行下采样处理（同理，对第二待检测图像也需要进行下采样处理）过程中，像素的偏移值。

在本实施例中，通过引入目标检测网络以及卡尔曼滤波模型，可以较好的将目标检测网络中的浅层特征作为目标跟踪的外观特征，进而通过卡尔曼滤波模型根据目标检测网络确定的外观特征，使得可以通过外观特征预测下一帧待检测图像中目标对象所处位置信息，且计算速度较快，提高了目标跟踪的准确率。

在一实施例中，步骤 S50 之后，也即所述通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果之后，包括：

在所述第一跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一。

可以理解地，匹配失败总次数指的是第一跟踪匹配结果为匹配失败结果的总次数。

在预设检测时间内，所述匹配失败总次数小于预设失败阈值时，获取所述待检测视频中第三待检测图像，以及目标对象在所述第三待检测图像中的第三位置信息；所述第三待检测图像是指与所述第二待检测图像时间相邻且时间位于所述第二待检测图像之后的图像。

可选地，预设失败阈值可以为 3 次，4 次等。预设检测时间可以为 2 分钟，5 分钟等。可以理解地，第二待检测图像以及第三待检测图像可以为待检测视频中任意连续两帧的图像，如假设第一待检测图像为待检测视频中的第一帧图像，则第二待检测图像为待检测视频中的第二帧图像，第三待检测图像为待检测视频中的第三帧图像。

在预设检测时间内，所述匹配失败总次数小于预设失败阈值时，表征可能是由于该目标对象可能暂时被遮挡导致的匹配错误，进而继续对后续帧数对应的待检测图像进行目标跟踪，进而获取待检测视频中的第三待检测图像；并将第三待检测图像输入至目标检测网络中，通过目标检测网络对第三待检测图像进行目标检测，得到目标对象在第三待检测图像中的第三位置信息。

根据所述第二位置信息，采用卡尔曼滤波模型预测所述第三待检测图像中目标对象的第二预测位置信息，并确定与所述第二预测位置信息对应的第三 ROI 区域。

具体地，在获取第二待检测图像中目标对象所处的第二位置信息之后，根据该第二位置信息，采用卡尔曼滤波模型预测目标对象在第三待检测图像中所处的位置信息，也即第二预测位置信息，并在第三待检测图像中，提取与该第二预测位置信息关联的区域，也即第三 ROI 区域。

根据所述第三位置信息，对所述第三待检测图像进行 ROI 区域提取，得到第四 ROI 区域。

具体地，在获取第三待检测图像中目标对象所处的第三位置信息之后，根据该第三位置信息，在第三待检测图像中提取与第三位置信息关联的区域，得到第四 ROI 区域。

确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离，同时确定所述第二位置信息与所述预测位置信息之间的第二位置重合度。

具体地，在根据所述第三位置信息，对所述第三待检测图像进行 ROI 区域提取，得到第四 ROI 区域之后，确定第三 ROI 区域与第四 ROI 区域之间的第二最小余弦距离，该第二最小余弦距离的取值范围可以为 0 至 1，第二最小余弦距离越大，表征第三 ROI 区域与第四 ROI 区域之间的特征相似程度越高；同时确定第三位置信息与第二预测位置信息之间的第二位置重合度，该第二位置重合度的取值范围也可以为 0 至 1，第二位置重合度越高，表征第三位置信息与第二预测位置信息之间的相关程度越大。

根据所述第二最小余弦距离以及所述第二位置重合度，通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果。

可以理解地，第二跟踪匹配结果可以为表征匹配成功的结果，也即表征第二预测位置

信息与第三位置信息相匹配（如第二预测位置信息中包含目标对象的特征，且第三位置信息中也包含目标对象的特征）；也可以为表征匹配失败的结果，也即表征第二预测位置信息与第三位置信息不匹配（如第二预测位置信息中包含目标对象的特征，而第三位置信息中未包含目标对象的特征；亦或者第二预测位置信息中未包含目标对象的特征，而第三位置信息中包含目标对象的特征）。

具体地，在确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离，同时确定所述第三位置信息与所述第二预测位置信息之间的第二位置重合度之后，将第二最小余弦距离以及第二位置重合度作为跟踪代价，通过匈牙利算法确定目标对象在第三待检测图像中的第二跟踪匹配结果。示例性地，如将第二最小余弦距离与预设余弦阈值（如设定为 0.9, 0.95 等）进行比较，在第二最小余弦距离大于或等于预设余弦阈值时，则将第二位置重合度与预设位置重合阈值（如设定为 0.9, 0.95 等）进行比较，在第二位置重合度大于或等于预设位置重合阈值时，则确定第二跟踪匹配结果为匹配成功结果；若第二最小余弦距离小于预设余弦阈值，和/或，第二位置重合度小于预设位置重合阈值，则确定第二跟踪匹配结果为匹配失败结果。

在一实施例中，所述通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果之后，包括：

在所述第二跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

具体地，在通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果之后，若第二跟踪匹配结果为匹配失败结果时，则将匹配失败总次数累加一。

在预设检测时间内，所述匹配失败总次数大于或等于所述预设失败阈值时，删除与所述目标对象关联的跟踪 ID，并确认对所述目标对象跟踪结束。

可以理解地，跟踪 ID 是在对目标对象进行目标跟踪之前分配至每个目标对象的唯一 ID，若目标对象中包含多个目标个体，则可以为每一个目标个体分配一个跟踪 ID。

在预设检测时间内，匹配失败总次数大于或等于预设失败阈值时，可以理解地，在预设检测时间内，也即表征连续三次目标跟踪对应的跟踪匹配结果均为匹配失败，进而表征该目标对象并不是在短时间内被暂时的遮挡导致匹配失败，而是该目标对象离开检测区域，进而导致匹配失败，因此可以删除与目标对象关联的跟踪 ID，在后续目标跟踪中，不需要继续对该目标对象进行跟踪，也即确认对目标对象跟踪结束，可以降低计算机的计算复杂度。进一步地，在预设检测时间内，匹配失败总次数大于或等于预设失败阈值时，可以将该匹配失败总次数清零。

在本实施例中，通过判定是否在预设检测时间内，匹配失败总次数大于或等于预设失败阈值，进而判定目标对象是短时间内被暂时遮挡，亦或者目标对象不在检测区域内，提高目标跟踪的准确率。

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，而不对本申请实施例的实施过程构成任何限定。

在一实施例中，提供一种基于目标检测网络的目标跟踪装置，该基于目标检测网络的目标跟踪装置与上述实施例中基于目标检测网络的目标跟踪方法一一对应。如图 5 所示，该基于目标检测网络的目标跟踪装置包括第一位置信息获取模块 10、第一位置信息预测模块 20、第一 ROI 区域提取模块 30、第一位置重合度确定模块 40 和第一跟踪匹配模块 50。各功能模块详细说明如下：

第一位置信息获取模块 10，用于获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

第一位置信息预测模块 20，用于根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应

的第一 ROI 区域;

第一 ROI 区域提取模块 30, 用于根据所述第二位置信息, 对所述第二待检测图像进行 ROI 区域提取, 得到第二 ROI 区域;

第一位置重合度确定模块 40, 用于确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离, 同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度;

第一跟踪匹配模块 50, 用于根据所述第一最小余弦距离以及所述第一位置重合度, 通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

优选地, 如图 6 所示, 第一位置信息获取模块 10 包括:

待检测视频获取子模块 101, 用于获取待检测视频, 所述待检测视频中包含多帧待检测图像;

待检测图像记录子模块 102, 用于将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像;

目标检测子模块 103, 用于通过目标检测网络对被选取的所述第一待检测图像进行目标检测, 得到所述第一位置信息; 同时通过目标检测网络对所述第二待检测图像进行目标检测, 得到所述第二位置信息。

优选地, 如图 7 所示, 所述目标检测子模块 103, 包括:

下采样处理单元 1031, 用于将所述第一待检测图像输入至所述目标检测网络中的骨干网络, 以对所述第一待检测图像进行下采样处理, 得到与所述第一待检测图像对应的多个待检测特征图层;

图层处理单元 1032, 用于对各所述待检测特征图层依次进行图层处理, 得到与所述第一待检测图像对应的目标特征图层;

位置特征提取单元 1033, 用于对所述目标特征图层进行位置特征提取, 得到所述第一位置信息。

优选地, 所述待检测特征图层中包括第一特征图层、第二特征图层、第三特征图层以及第四特征图层; 所述图层处理单元, 包括:

第一图层处理子单元, 用于对所述第四特征图层进行卷积处理, 并对卷积处理后的第四特征图层进行上采样, 得到与第三特征图层具有相同维度的第五特征图层;

第二图层处理子单元, 用于将所述第五特征图层与所述第三特征图层进行维度叠加得到第一叠加图层之后, 对所述第一叠加图层进行卷积处理, 并对卷积处理后的第一叠加图层进行上采样, 得到与第二特征图层具有相同维度的第六特征图层;

第三图层处理子单元, 用于将所述第六特征图层与所述第二特征图层进行维度叠加得到第二叠加图层之后, 对所述第二叠加图层进行卷积处理, 并对卷积处理后的第二叠加图层进行上采样, 得到与第一特征图层具有相同维度的第七特征图层;

第四图层处理子单元, 用于将所述第七特征图层与所述第一特征图层进行维度叠加得到第三叠加图层之后, 对所述第三叠加图层进行卷积处理, 并对卷积处理后的第三叠加图层进行上采样, 得到所述目标特征图层。

优选地, 第一位置重合度确定模块 40, 包括:

交并位置信息确定子模块, 用于确定所述第二位置信息与所述预测位置信息之间的交集位置信息, 同时确定所述第二位置信息与所述预测位置信息之间的并集位置信息;

位置重合度确定子模块, 用于根据所述交集位置信息以及所述并集位置信息, 确定所述位置重合度。

优选地, 基于目标检测网络的目标跟踪装置还包括:

第一匹配失败总次数累加模块, 用于在所述第一跟踪匹配结果为匹配失败结果时, 将匹配失败总次数累加一;

第二位置信息获取模块，用于在所述匹配失败总次数小于预设失败阈值时，获取所述待检测视频中第三待检测图像，以及目标对象在所述第三待检测图像中的第三位置信息；所述第三待检测图像是指与所述第二待检测图像时间相邻且时间位于所述第二待检测图像之后的图像；

第二位置信息预测模块，用于根据所述第二位置信息，采用卡尔曼滤波模型预测所述第三待检测图像中目标对象的第二预测位置信息，并确定与所述第二预测位置信息对应的第三 ROI 区域；

第二 ROI 区域提取模块，用于根据所述第三位置信息，对所述第三待检测图像进行 ROI 区域提取，得到第四 ROI 区域；

第二位置重合度确定模块，用于确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离，同时确定所述第二位置信息与所述预测位置信息之间的第二位置重合度；

第二跟踪匹配模块，用于根据所述第二最小余弦距离以及所述第二位置重合度，通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果。

优选地，基于目标检测网络的目标跟踪装置，包括：

第二匹配失败总次数累加模块，用于在所述第二跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

跟踪结束确认模块，用于在所述匹配失败总次数大于或等于所述预设失败阈值时，删除与所述目标对象关联的跟踪 ID，并确认对所述目标对象跟踪结束。

关于基于目标检测网络的目标跟踪装置的具体限定可以参见上文中对于基于目标检测网络的目标跟踪方法的限定，在此不再赘述。上述基于目标检测网络的目标跟踪装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 8 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括可读存储介质、内存储器。该可读存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为可读存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储上述实施例中基于目标检测网络的目标跟踪方法所使用到的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种基于目标检测网络的目标跟踪方法。本实施例所提供的可读存储介质包括非易失性可读存储介质和易失性可读存储介质。

在一个实施例中，提供了一种计算机设备，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：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

在一实施例中，提供了一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

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或者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。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

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。

1. 一种基于目标检测网络的目标跟踪方法，其中，包括：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

2. 如权利要求 1 所述的基于目标检测网络的目标跟踪方法，其中，所述获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息，包括：

获取待检测视频，所述待检测视频中包含多帧待检测图像；

将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像；

通过目标检测网络对所述第一待检测图像进行目标检测，得到所述第一位置信息；同时通过目标检测网络对所述第二待检测图像进行目标检测，得到所述第二位置信息。

3. 如权利要求 2 所述的基于目标检测网络的目标跟踪方法，其中，所述通过目标检测网络对被选取的所述第一待检测图像进行目标检测，得到所述第一位置信息，包括：

将所述第一待检测图像输入至所述目标检测网络中的骨干网络，以对所述第一待检测图像进行下采样处理，得到与所述第一待检测图像对应的多个待检测特征图层；

对各所述待检测特征图层依次进行图层处理，得到与所述第一待检测图像对应的目标特征图层；

对所述目标特征图层进行位置特征提取，得到所述第一位置信息。

4. 如权利要求 3 所述的基于目标检测网络的目标跟踪方法，其中，所述待检测特征图层中包括第一特征图层、第二特征图层、第三特征图层以及第四特征图层；所述对各所述待检测特征图层依次进行图层处理，得到与所述第一待检测图像对应的目标特征图层，包括：

对所述第四特征图层进行卷积处理，并对卷积处理后的第四特征图层进行上采样，得到与第三特征图层具有相同维度的第五特征图层；

将所述第五特征图层与所述第三特征图层进行维度叠加得到第一叠加图层之后，对所述第一叠加图层进行卷积处理，并对卷积处理后的第一叠加图层进行上采样，得到与第二特征图层具有相同维度的第六特征图层；

将所述第六特征图层与所述第二特征图层进行维度叠加得到第二叠加图层之后，对所述第二叠加图层进行卷积处理，并对卷积处理后的第二叠加图层进行上采样，得到与第一特征图层具有相同维度的第七特征图层；

将所述第七特征图层与所述第一特征图层进行维度叠加得到第三叠加图层之后，对所述第三叠加图层进行卷积处理，并对卷积处理后的第三叠加图层进行上采样，得到所述目标特征图层。

5. 如权利要求 1 所述的基于目标检测网络的目标跟踪方法，其中，所述确定所述第二位置信息与所述第一预测位置信息之间的位置重合度，包括：

确定所述第二位置信息与所述第一预测位置信息之间的交集位置信息，同时确定所述第二位置信息与所述第一预测位置信息之间的并集位置信息；

根据所述交集位置信息以及所述并集位置信息，确定所述位置重合度。

6. 如权利要求 1 所述的基于目标检测网络的目标跟踪方法，其中，所述通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果之后，包括：

在所述第一跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

在预设检测时间内，所述匹配失败总次数小于预设失败阈值时，获取所述待检测视频中第三待检测图像，以及目标对象在所述第三待检测图像中的第三位置信息；所述第三待检测图像是指与所述第二待检测图像时间相邻且时间位于所述第二待检测图像之后的图像；

根据所述第二位置信息，采用卡尔曼滤波模型预测所述第三待检测图像中目标对象的第二预测位置信息，并确定与所述第二预测位置信息对应的第三 ROI 区域；

根据所述第三位置信息，对所述第三待检测图像进行 ROI 区域提取，得到第四 ROI 区域；

确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离，同时确定所述第二位置信息与所述预测位置信息之间的第二位置重合度；

根据所述第二最小余弦距离以及所述第二位置重合度，通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果。

7. 如权利要求 6 所述的基于目标检测网络的目标跟踪方法，其中，所述通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果之后，包括：

在预设检测时间内，所述第二跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

在所述匹配失败总次数大于或等于所述预设失败阈值时，删除与所述目标对象关联的跟踪 ID，并确认对所述目标对象跟踪结束。

8. 一种基于目标检测网络的目标跟踪装置，其中，包括：

第一位置信息获取模块，用于获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

第一位置信息预测模块，用于根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

第一 ROI 区域提取模块，用于根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

第一位置重合度确定模块，用于确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

第一跟踪匹配模块，用于根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

9. 一种计算机设备，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，其中，所述处理器执行所述计算机可读指令时实现如下步骤：

获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息；所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像；

根据所述第一位置信息，采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息，并确定与所述第一预测位置信息对应的第一 ROI 区域；

根据所述第二位置信息，对所述第二待检测图像进行 ROI 区域提取，得到第二 ROI 区域；

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离，同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度；

根据所述第一最小余弦距离以及所述第一位置重合度，通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

10.如权利要求 9 所述的计算机设备，其中，所述获取目标对象在第一待检测图像中的第一位置信息，以及目标对象在第二待检测图像中的第二位置信息，包括：

获取待检测视频，所述待检测视频中包含多帧待检测图像；

将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像；

通过目标检测网络对所述第一待检测图像进行目标检测，得到所述第一位置信息；同时通过目标检测网络对所述第二待检测图像进行目标检测，得到所述第二位置信息。

11.如权利要求 10 所述的计算机设备，其中，所述通过目标检测网络对被选取的所述第一待检测图像进行目标检测，得到所述第一位置信息，包括：

将所述第一待检测图像输入至所述目标检测网络中的骨干网络，以对所述第一待检测图像进行下采样处理，得到与所述第一待检测图像对应的多个待检测特征图层；

对各所述待检测特征图层依次进行图层处理，得到与所述第一待检测图像对应的目标特征图层；

对所述目标特征图层进行位置特征提取，得到所述第一位置信息。

12.如权利要求 11 所述的计算机设备，其中，所述待检测特征图层中包括第一特征图层、第二特征图层、第三特征图层以及第四特征图层；所述对各所述待检测特征图层依次进行图层处理，得到与所述第一待检测图像对应的目标特征图层，包括：

对所述第四特征图层进行卷积处理，并对卷积处理后的第四特征图层进行上采样，得到与第三特征图层具有相同维度的第五特征图层；

将所述第五特征图层与所述第三特征图层进行维度叠加得到第一叠加图层之后，对所述第一叠加图层进行卷积处理，并对卷积处理后的第一叠加图层进行上采样，得到与第二特征图层具有相同维度的第六特征图层；

将所述第六特征图层与所述第二特征图层进行维度叠加得到第二叠加图层之后，对所述第二叠加图层进行卷积处理，并对卷积处理后的第二叠加图层进行上采样，得到与第一特征图层具有相同维度的第七特征图层；

将所述第七特征图层与所述第一特征图层进行维度叠加得到第三叠加图层之后，对所述第三叠加图层进行卷积处理，并对卷积处理后的第三叠加图层进行上采样，得到所述目标特征图层。

13.如权利要求 9 所述的计算机设备，其中，所述确定所述第二位置信息与所述第一预测位置信息之间的位置重合度，包括：

确定所述第二位置信息与所述第一预测位置信息之间的交集位置信息，同时确定所述第二位置信息与所述第一预测位置信息之间的并集位置信息；

根据所述交集位置信息以及所述并集位置信息，确定所述位置重合度。

14.如权利要求 9 所述的计算机设备，其中，所述通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果之后，包括：

在所述第一跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

在预设检测时间内，所述匹配失败总次数小于预设失败阈值时，获取所述待检测视频中第三待检测图像，以及目标对象在所述第三待检测图像中的第三位置信息；所述第三待检测图像是指与所述第二待检测图像时间相邻且时间位于所述第二待检测图像之后的图像；

根据所述第二位置信息,采用卡尔曼滤波模型预测所述第三待检测图像中目标对象的第二预测位置信息,并确定与所述第二预测位置信息对应的第三 ROI 区域;

根据所述第三位置信息,对所述第三待检测图像进行 ROI 区域提取,得到第四 ROI 区域;

确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离,同时确定所述第二位置信息与所述预测位置信息之间的第二位置重合度;

根据所述第二最小余弦距离以及所述第二位置重合度,通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果。

15.一个或多个存储有计算机可读指令的可读存储介质,其中,所述计算机可读指令被一个或多个处理器执行时,使得所述一个或多个处理器执行如下步骤:

获取目标对象在第一待检测图像中的第一位置信息,以及目标对象在第二待检测图像中的第二位置信息;所述第二待检测图像是指在待检测视频中与所述第一待检测图像时间相邻且时间位于所述第一待检测图像之后的图像;

根据所述第一位置信息,采用卡尔曼滤波模型预测所述第二待检测图像中目标对象的第一预测位置信息,并确定与所述第一预测位置信息对应的第一 ROI 区域;

根据所述第二位置信息,对所述第二待检测图像进行 ROI 区域提取,得到第二 ROI 区域;

确定所述第一 ROI 区域与所述第二 ROI 区域之间的第一最小余弦距离,同时确定所述第二位置信息与所述第一预测位置信息之间的第一位置重合度;

根据所述第一最小余弦距离以及所述第一位置重合度,通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果。

16.如权利要求 15 所述的可读存储介质,其中,所述获取目标对象在第一待检测图像中的第一位置信息,以及目标对象在第二待检测图像中的第二位置信息,包括:

获取待检测视频,所述待检测视频中包含多帧待检测图像;

将所述待检测视频中任意一帧所述待检测图像记录为所述第一待检测图像;

通过目标检测网络对所述第一待检测图像进行目标检测,得到所述第一位置信息;同时通过目标检测网络对所述第二待检测图像进行目标检测,得到所述第二位置信息。

17.如权利要求 16 所述的可读存储介质,其中,所述通过目标检测网络对被选取的所述第一待检测图像进行目标检测,得到所述第一位置信息,包括:

将所述第一待检测图像输入至所述目标检测网络中的骨干网络,以对所述第一待检测图像进行下采样处理,得到与所述第一待检测图像对应的多个待检测特征图层;

对各所述待检测特征图层依次进行图层处理,得到与所述第一待检测图像对应的目标特征图层;

对所述目标特征图层进行位置特征提取,得到所述第一位置信息。

18.如权利要求 17 所述的可读存储介质,其中,所述待检测特征图层中包括第一特征图层、第二特征图层、第三特征图层以及第四特征图层;所述对各所述待检测特征图层依次进行图层处理,得到与所述第一待检测图像对应的目标特征图层,包括:

对所述第四特征图层进行卷积处理,并对卷积处理后的第四特征图层进行上采样,得到与第三特征图层具有相同维度的第五特征图层;

将所述第五特征图层与所述第三特征图层进行维度叠加得到第一叠加图层之后,对所述第一叠加图层进行卷积处理,并对卷积处理后的第一叠加图层进行上采样,得到与第二特征图层具有相同维度的第六特征图层;

将所述第六特征图层与所述第二特征图层进行维度叠加得到第二叠加图层之后,对所述第二叠加图层进行卷积处理,并对卷积处理后的第二叠加图层进行上采样,得到与第一特征图层具有相同维度的第七特征图层;

将所述第七特征图层与所述第一特征图层进行维度叠加得到第三叠加图层之后，对所述第三叠加图层进行卷积处理，并对卷积处理后的第三叠加图层进行上采样，得到所述目标特征图层。

19.如权利要求 15 所述的 readable 存储介质，其中，所述确定所述第二位置信息与所述第一预测位置信息之间的位置重合度，包括：

确定所述第二位置信息与所述第一预测位置信息之间的交集位置信息，同时确定所述第二位置信息与所述第一预测位置信息之间的并集位置信息；

根据所述交集位置信息以及所述并集位置信息，确定所述位置重合度。

20.如权利要求 15 所述的 readable 存储介质，其中，所述通过匈牙利算法确定所述目标对象在第二待检测图像中的第一跟踪匹配结果之后，包括：

在所述第一跟踪匹配结果为匹配失败结果时，将匹配失败总次数累加一；

在预设检测时间内，所述匹配失败总次数小于预设失败阈值时，获取所述待检测视频中第三待检测图像，以及目标对象在所述第三待检测图像中的第三位置信息；所述第三待检测图像是指与所述第二待检测图像时间相邻且时间位于所述第二待检测图像之后的图像；

根据所述第二位置信息，采用卡尔曼滤波模型预测所述第三待检测图像中目标对象的第二预测位置信息，并确定与所述第二预测位置信息对应的第三 ROI 区域；

根据所述第三位置信息，对所述第三待检测图像进行 ROI 区域提取，得到第四 ROI 区域；

确定所述第三 ROI 区域与所述第四 ROI 区域之间的第二最小余弦距离，同时确定所述第二位置信息与所述预测位置信息之间的第二位置重合度；

根据所述第二最小余弦距离以及所述第二位置重合度，通过匈牙利算法确定所述目标对象在第三待检测图像中的第二跟踪匹配结果。

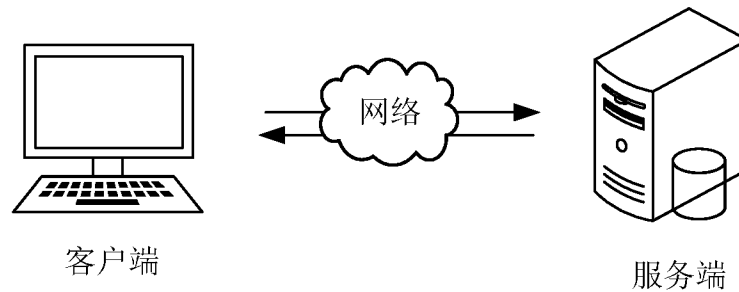


图 1

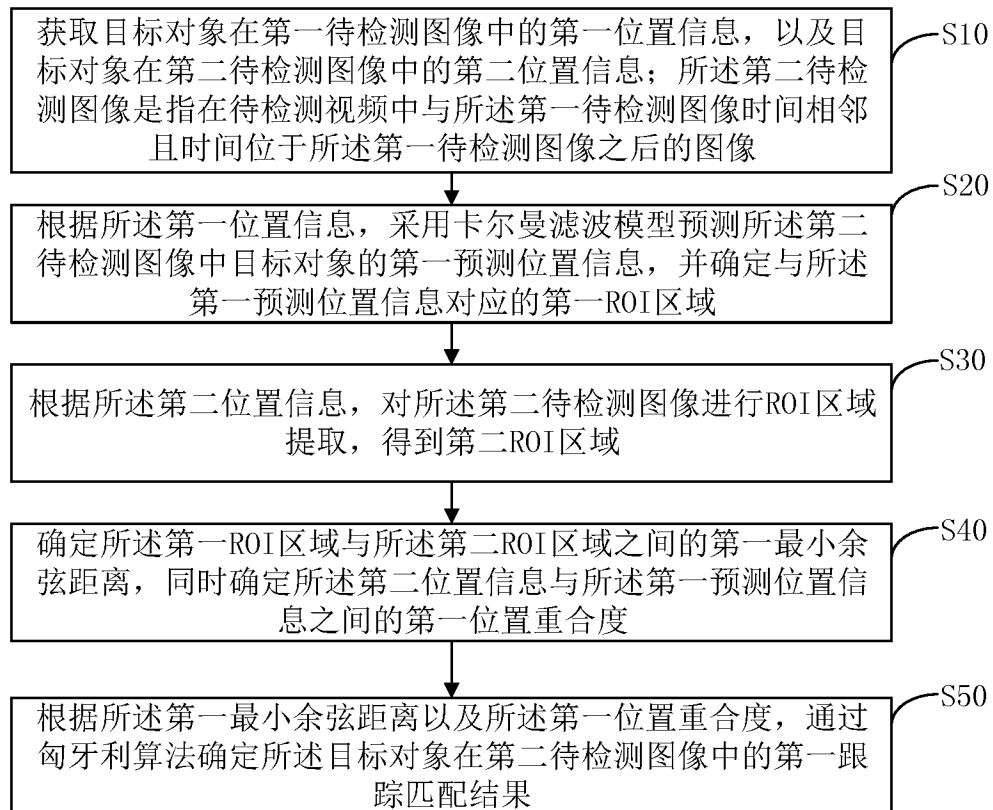


图 2

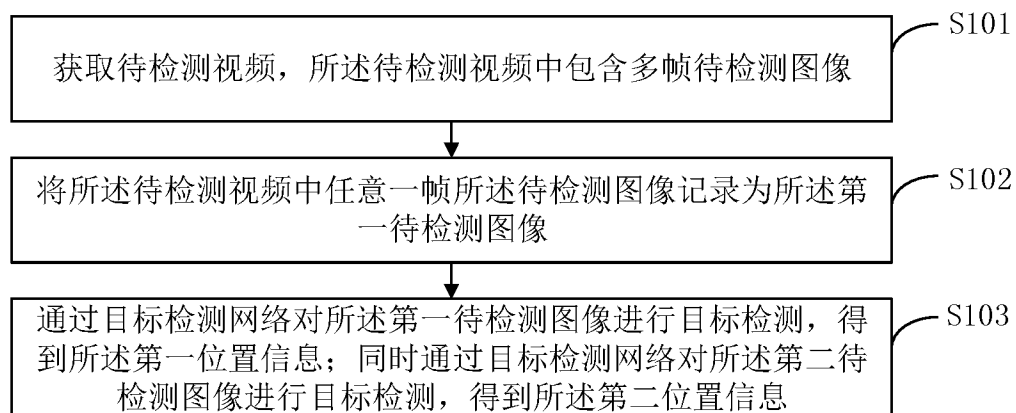


图 3

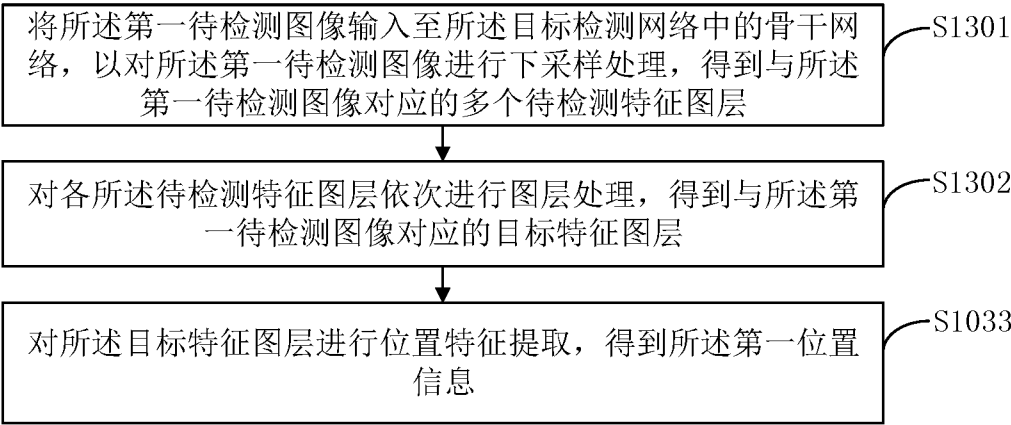


图 4

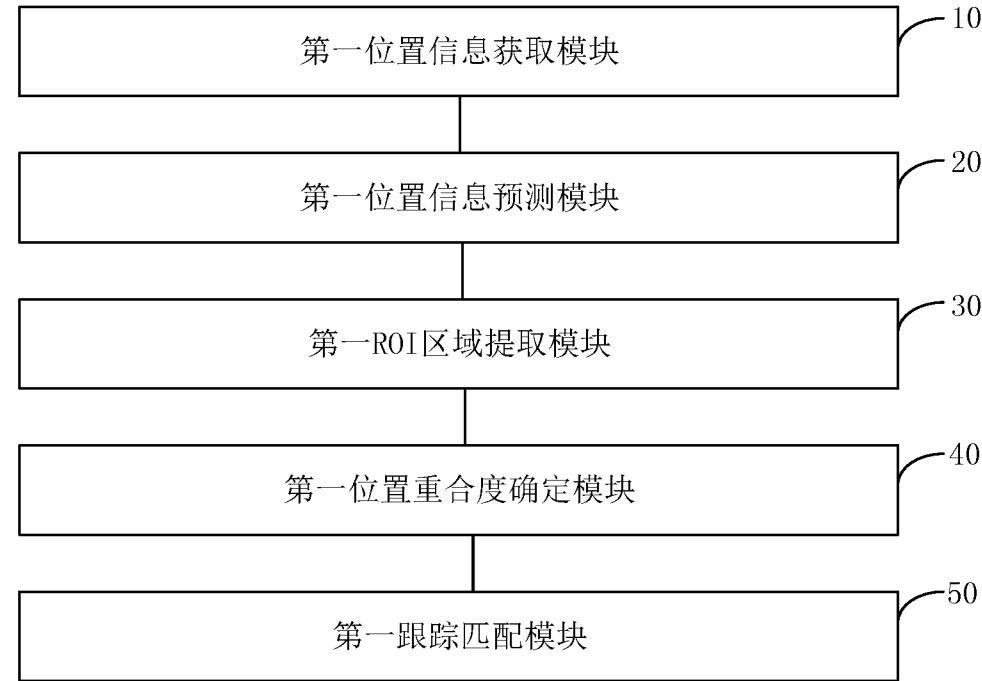


图 5



图 6

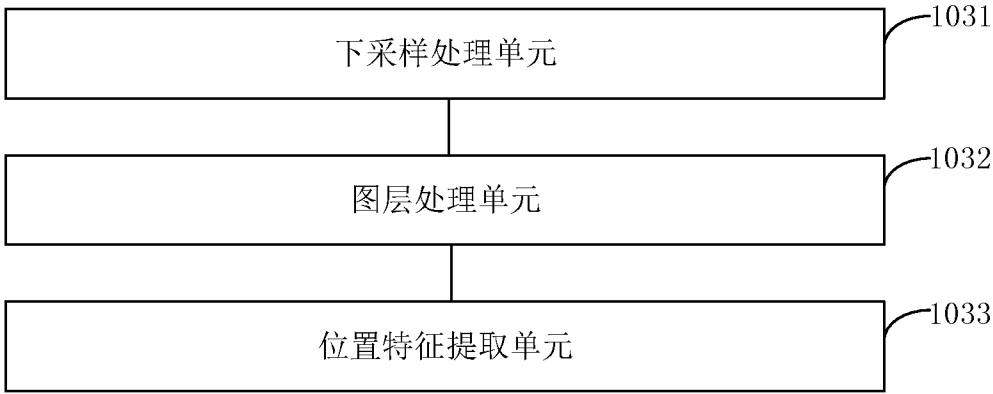


图 7

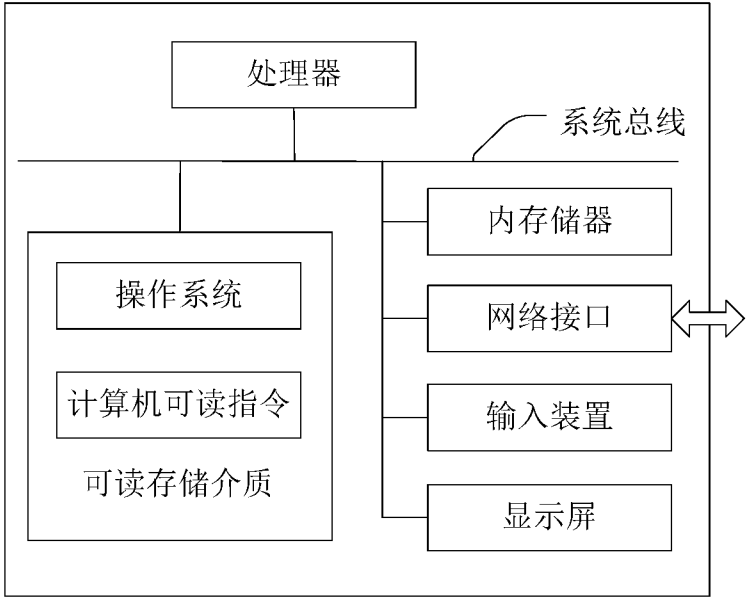


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/096757

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/32(2006.01)i; G06K 9/62(2006.01)i; G06T 7/246(2017.01)i; G06T 7/277(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)

G06K; 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)

WPI, EPODOC, CNPAT, CNKI, IEEE: 目标, 跟踪, 位置, 图像, 预测, 区域, 距离, 重合度, object, target, track, position, image, prediction, ROI, region, distance, coincidence, degree

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 110866428 A (HANGZHOU HIKVISION DIGITAL TECHNOLOGY CO., LTD.) 06 March 2020 (2020-03-06) claims 1-10	1-3, 5, 8-11, 13, 15-17, 19
A	CN 111127513 A (BEIJING JIAOTONG UNIVERSITY) 08 May 2020 (2020-05-08) entire document	1-20
A	CN 110796686 A (ZHEJIANG DAHUA TECHNOLOGY CO., LTD.) 14 February 2020 (2020-02-14) entire document	1-20
A	US 9811916 B1 (AMAZON TECHNOLOGIES, INC.) 07 November 2017 (2017-11-07) 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

22 December 2021

Date of mailing of the international search report

19 January 2022

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/096757

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	110866428	A	06 March 2020	None	
CN	111127513	A	08 May 2020	None	
CN	110796686	A	14 February 2020	None	
US	9811916	B1	07 November 2017	None	

国际检索报告

国际申请号

PCT/CN2021/096757

A. 主题的分类

G06K 9/32(2006.01)i; G06K 9/62(2006.01)i; G06T 7/246(2017.01)i; G06T 7/277(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)

G06K; 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))

WPI, EPDOC, CNPAT, CNKI, IEEE: 目标, 跟踪, 位置, 图像, 预测, 区域, 距离, 重合度, object, target, track, position, image, prediction, ROI, region, distance, coincidence, degree

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
X	CN 110866428 A (杭州海康威视数字技术股份有限公司) 2020年3月6日 (2020 - 03 - 06) 权利要求1-10	1-3、5、8-11、 13、15-17、19
A	CN 111127513 A (北京交通大学) 2020年5月8日 (2020 - 05 - 08) 全文	1-20
A	CN 110796686 A (浙江大华技术股份有限公司) 2020年2月14日 (2020 - 02 - 14) 全文	1-20
A	US 9811916 B1 (AMAZON TECHNOLOGIES, INC.) 2017年11月7日 (2017 - 11 - 07) 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

2021年12月22日

国际检索报告邮寄日期

2022年1月19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

张涛

电话号码 86-(10)-53961356

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/096757

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	110866428	A	2020年3月6日	无	
CN	111127513	A	2020年5月8日	无	
CN	110796686	A	2020年2月14日	无	
US	9811916	B1	2017年11月7日	无	