

(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 2020/258499 A1

(43) 国际公布日

2020 年 12 月 30 日 (30.12.2020)

WIPO | PCT

(51) 国际专利分类号:

G06K 9/00 (2006.01)

(21) 国际申请号:

PCT/CN2019/103195

(22) 国际申请日:

2019 年 8 月 29 日 (29.08.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910554748.9 2019年6月25日 (25.06.2019) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN

TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区福田街道福

安社区益田路 5033 号平安金融中心 23

楼, Guangdong 518033 (CN)。

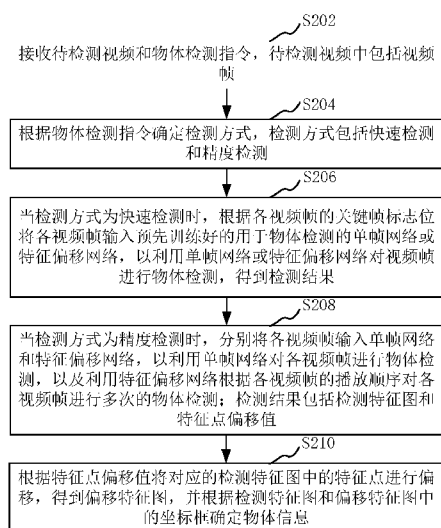
(72) 发明人: 叶明(YE, Ming); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司(ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房(部位: 自编 01-03 和 08-12 单元)(仅限办公用途), Guangdong 510623 (CN)。

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

(54) Title: OBJECT DETECTION METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 物体检测方法、装置、计算机设备和存储介质



S202 RECEIVE A VIDEO TO BE DETECTED AND AN OBJECT DETECTION INSTRUCTION, WHEREIN THE VIDEO TO BE DETECTED COMPRISES A VIDEO FRAME
S204 DETERMINE A DETECTION MODE ACCORDING TO THE OBJECT DETECTION INSTRUCTION, WHEREIN THE DETECTION MODE COMPRISES RAPID DETECTION AND PRECISION DETECTION
S206 WHEN THE DETECTION MODE IS RAPID DETECTION, INPUT, ACCORDING TO A KEY FRAME FLAG BIT OF EACH VIDEO FRAME, EACH VIDEO FRAME INTO A PRE-TRAINED SINGLE-FRAME NETWORK OR FEATURE OFFSET NETWORK USED FOR OBJECT DETECTION, SO AS TO PERFORM OBJECT DETECTION ON EACH VIDEO FRAME BY MEANS OF THE SINGLE-FRAME NETWORK OR THE FEATURE OFFSET NETWORK, AND OBTAIN A DETECTION RESULT
S208 WHEN THE DETECTION MODE IS PRECISION DETECTION, RESPECTIVELY INPUT EACH VIDEO FRAME INTO THE SINGLE-FRAME NETWORK AND THE FEATURE OFFSET NETWORK, SO AS TO PERFORM OBJECT DETECTION ON EACH VIDEO FRAME BY MEANS OF THE SINGLE-FRAME NETWORK, AND TO PERFORM, BY MEANS OF THE FEATURE OFFSET NETWORK, MULTIPLE INSTANCES OF OBJECT DETECTION ON EACH VIDEO FRAME ACCORDING TO A PLAYING SEQUENCE OF EACH VIDEO FRAME, WHEREIN THE DETECTION RESULT COMPRISES A DETECTION FEATURE MAP AND A FEATURE POINT OFFSET VALUE
S210 ACCORDING TO THE FEATURE POINT OFFSET VALUE, OFFSET FEATURE POINTS IN THE CORRESPONDING DETECTION FEATURE MAP TO OBTAIN AN OFFSET FEATURE MAP, AND DETERMINE OBJECT INFORMATION ACCORDING TO A COORDINATE BOX IN THE DETECTION FEATURE MAP AND IN THE OFFSET FEATURE MAP

图 2

(57) Abstract: An object detection method, comprising: receiving a video to be detected and an object detection instruction sent by a terminal, wherein the video to be detected comprises a video frame; determining a detection mode according to the type of the object detection instruction, wherein the detection mode comprises rapid detection and precision detection; when the detection mode is rapid detection, inputting, according to a key frame flag bit of each video frame, each video frame into a pre-trained single-frame network or feature offset network used for object detection, so as to perform object detection on each video frame by means of the single-frame network or the feature offset network, and obtaining a detection result; when the detection mode is precision detection, respectively inputting each video frame into the single-frame network and the feature offset network, so as to perform object detection on each video frame by means of the single-frame network, and to perform, by means of the feature offset network, multiple instances of object detection on each video frame according to a playing sequence of each video frame, wherein the detection result comprises a detection

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

feature map and a feature point offset value; and according to the feature point offset value, offsetting feature points in the corresponding detection feature map to obtain an offset feature map, and determining object information according to a coordinate box in the detection feature map and in the offset feature map.

(57) 摘要: 一种物体检测方法, 包括: 接收终端发送的待检测视频和物体检测指令, 待检测视频中包括视频帧; 根据物体检测指令的类型确定检测方式, 检测方式包括快速检测和精度检测; 当检测方式为快速检测时, 根据各视频帧的关键帧标志位将各视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络, 以利用单帧网络或特征偏移网络对视频帧进行物体检测, 得到检测结果; 当检测方式为精度检测时, 分别将各视频帧输入单帧网络和特征偏移网络, 以利用单帧网络对各视频帧进行物体检测, 以及利用特征偏移网络根据各视频帧的播放顺序对各视频帧进行多次的物体检测; 检测结果包括检测特征图和特征点偏移值; 根据特征点偏移值将对应的检测特征图中的特征点进行偏移, 得到偏移特征图, 并根据检测特征图和偏移特征图中的坐标框确定物体信息。

物体检测方法、装置、计算机设备和存储介质

相关申请的交叉引用

本申请要求于 2019 年 06 月 25 日提交中国专利局，申请号为 2019105547489，申请
5 名称为“物体检测方法、装置、计算机设备和存储介质”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

本申请涉及一种物体检测方法、装置、计算机设备和存储介质。

10

背景技术

物体检测是计算机视觉中的经典问题之一，其任务是用框标出物体的位置并给出物体的类别。因此，物体检测在人脸识别、安全监控以及动态追踪等很多方面都有广泛的应用。

然而，发明人意识到，传统进行物体检测的算法大多采用逐帧检测网络直接对视频帧
15 依次进行检测。虽然在一定程度上保证了检测的精度性，但是在高分辨情况下难以达到实时，
导致检测速度缓慢。

发明内容

根据本申请公开的各种实施例，提供一种物体检测方法、装置、计算机设备和存储介
20 质。

一种物体检测方法，包括：

服务器接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

所述服务器根据所述物体检测指令确定检测方式，所述检测方式包括快速检测和精度
检测；

25 当所述检测方式为所述快速检测时，所述服务器根据各所述视频帧的关键帧标志位将
各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单
帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

当所述检测方式为所述精度检测时，所述服务器分别将各所述视频帧输入所述单帧网
络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所
30 述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述
检测结果包括检测特征图和特征点偏移值；及

所述服务器根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏
移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

一种物体检测装置，包括：

接收模块，用于接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

确定模块，用于根据所述物体检测指令确定检测方式，所述检测方式包括快速检测和精度检测；

5 检测模块，用于当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

10 检测模块还用于当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

偏移模块，用于根据所述特征点偏移值将各所述视频帧输入预先训练好的用于物体检测的单帧网络和特征偏移网络，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

15 一种计算机设备，包括存储器和一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行以下步骤：

接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

根据所述物体检测指令确定检测方式，所述检测方式包括快速检测和精度检测；

20 当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

25 当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

30 根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

根据所述物体检测指令确定检测方式，所述检测方式包括快速检测和精度检测；

35 当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

- 5 根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

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。本申请的其它特征和优点将从说明书、附图以及权利要求书变得明显。

10 附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例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的前提下，还可以根据这些附图获得其它的附图。

图 1 为根据一个或多个实施例中物体检测方法的应用场景图。

- 15 图 2 为根据一个或多个实施例中物体检测方法的流程示意图。

图 3 为根据一个或多个实施例中基于检测方式，利用预设的单帧网络和特征偏移网络对各视频帧进行检测，得到检测结果步骤的流程示意图。

图 4 为另一个实施例中基于检测方式，利用预设的单帧网络和特征偏移网络对各视频帧进行检测，得到检测结果步骤的流程示意图。

- 20 图 5 为根据一个或多个实施例中物体检测装置的框图。

图 6 为根据一个或多个实施例中计算机设备的框图。

具体实施方式

- 25 为了使本申请的技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

- 本申请提供的物体检测方法，可以应用于如图 1 所示的应用环境中。该应用环境涉及终端 102 和服务器 104，终端 102 通过网络与服务器 104 进行通信。终端 102 获取到待检测视频和物体检测指令时，可以由终端 102 单独实现上述物体检测方法。也可以通过终端 30 102 将待检测视频和物体检测指令发送给服务器 104，服务器 104 实现物体检测方法。具体地，服务器 104 接收终端 102 发送的待检测视频和物体检测指令，待检测视频中包括视频帧。服务器 104 根据物体检测指令确定检测方式，检测方式包括快速检测和精度检测。当检测方式为快速检测时，服务器 104 根据各视频帧的关键帧标志位将各视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用单帧网络或特征偏移网络对视频帧进行物体检测，得到检测结果；当检测方式为精度检测时，服务器 104 分别将各视频 35 帧输入单帧网络或特征偏移网络，以利用单帧网络或特征偏移网络对视频帧进行物体检测，得到检测结果。

帧输入单帧网络和特征偏移网络，以利用单帧网络对各视频帧进行物体检测，以及利用特征偏移网络根据各视频帧的播放顺序对各视频帧进行多次的物体检测；检测结果包括检测特征图和特征点偏移值；服务器 104 根据特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据检测特征图和偏移特征图中的坐标框确定物体信息。终端 102 可以但不限于是各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备，服务器 104 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。

在一些实施例中，如图 2 所示，提供了一种物体检测方法，以该方法应用于图 1 中的服务器为例进行说明，包括以下步骤：

步骤 S202，接收终端发送的待检测视频和物体检测指令，待检测视频中包括视频帧。

待检测视频是指需要进行物体检测的视频，视频帧则是用于组成视频的视频图像。例如，待检测视频包括但不限于监控设备的监控视频。而物体检测指令是指用于请求服务器进行物体检测的指令。

具体地，当用户有物体检测需求时，通过操作用户终端给服务器下发物体检测指令，以及随物体检测指令向服务器发送本次进行物体检测的待检测视频。

步骤 S204，根据物体检测指令确定检测方式，检测方式包括快速检测和精度检测。

检测方式包括快速检测和精度检测，通过不同的物体检测确定本次进行物体检测的方式。具体地，当用户需要进行物体检测时，通过操作终端发起物体检测任务和选择对应的检测方式，终端根据用户的操作生成对应的物体检测指令。也就是说，若物体检测任务的需求是快速获取检测结果的，则用户在操作终端发起物体检测任务时同步在终端页面选择快速检测方式，当终端接收到用户选择了快速检测方式后，根据用户所选择的快速检测方式生成快速物体检测指令，则生成的物体检测指令为快速物体检测指令。服务器接收到快速物体检测指令时，即可响应该快速物体检测指令确定检测方式为快速检测，则基于快速检测方式进行快速的物体检测。同样的，当用户需要高精度的物体检测结果的，则在下发物体检测指令时同步选择精度检测方式，则生成的物体检测指令为精度物体检测指令。服务器接收到精度物体检测指令时，即可响应该精度物体检测指令进行高精度的物体检测。或者，当用户操作终端发起物体检测任务和选择对应的检测方式时，终端可以直接生成物体检测指令，将用户所选择的方式随物体检测指令发送给服务器，即物体检测指令中包括检测方式。服务器通过对物体检测指令进行判断，确定物体检测指令所包括的检测方式，从而确定物体检测的检测方式。

步骤 S206，当检测方式为快速检测时，根据各视频帧的关键帧标志位将各视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用单帧网络或特征偏移网络对视频帧进行物体检测，得到检测结果。

步骤 S208，当检测方式为精度检测时，分别将各视频帧输入单帧网络和特征偏移网络，以利用单帧网络对各视频帧进行物体检测，以及利用特征偏移网络根据各视频帧的播放顺序对各视频帧进行多次的物体检测；检测结果包括检测特征图和特征点偏移值。

单帧网络和特征偏移网络均为预先根据视频样本训练好，用于物体检测的神经网络，视频样本是指已标注物体的视频帧。单帧网络为传统的逐帧检测算法网络，可以为现有任意一种，例如 faster rcnn 检测网络、yolov3 检测网络等。特征偏移网络是一种近似光流网络的目标检测网络，利用 flownet 作为网络的基础架构。与传统的光流网络相比，特征偏移网络的输出层从计算图像中每个像素点的偏移值更改为只计算特征图中每个像素点的偏移值。单帧网络得到的检测结果是检测特征图，特征偏移网络得到的检测结果是特征点偏移值。基于不同的检测方式将各视频帧输入至预先训练好的单帧网络和特征偏移网络进行物体检测。

具体地，基于不同的检测方式，各视频帧输入至单帧网络和特征偏移网络的形式不一样。当检测方式为快速检测时，需要根据各视频帧对应的关键帧标志位对所有的视频帧进行关键帧和非关键帧的划分。属于关键帧的视频帧利用单帧网络进行检测，属于非关键帧的视频帧利用特征偏移网络进行检测。而当检测方式为精度检测时，所有的视频帧需要同时利用单帧网络和特征偏移网络进行检测，即所有视频帧输入到单帧网络进行检测之外，还需要将视频帧输入到特征偏移网络进行检测。并且，特征偏移网络会根据各视频帧的播放顺序对各视频帧进行多次的物体检测，即根据各视频帧正序的播放顺序检测一次，以及根据各视频帧倒序的播放顺序检测一次。

步骤 S210，根据特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据检测特征图和偏移特征图中的坐标框确定物体信息。

检测特征图是单帧网络对视频帧进行检测输出的特征图，而特征点偏移值是特征偏移网络对视频帧进行检测，得到的该视频帧中的特征点相对于上一帧视频帧中的特征点的偏移值，即相对应的特征点之间的变化程度。也就是通过相邻视频帧之间的特征点的偏移值就能得到视频中物体偏移的程度，即物体的移动变化程度。物体信息是指被检测出的物体所属的类别。

具体地，由于利用网络模型进行物体检测，网络模型输出的特征图上包括被检测出物体的坐标框，则根据该坐标框可直接确定物体信息。即，当视频帧属于关键帧，也就是利用单帧网络对该视频帧进行检测，得到的是该视频帧对应的检测特征图，则根据单帧检测网络输出的检测特征图上的坐标框直接确定物体信息。而当视频帧属于非关键帧，也就是利用特征偏移网络检测对该视频帧进行检测，得到的是该视频帧相对于上一帧视频帧中特征点的变化程度，即特征点偏移值，则进一步根据该视频帧对应的特征点偏移值以及对应的上一帧视频帧的检测结果得到该视频帧的偏移特征图。可以理解，该偏移特征图与检测特征图相同，同样包括物体坐标框，根据偏移特征图即可直接确定该视频帧中物体的物体信息。上一帧视频帧对应的检测结果为上一帧视频帧对应的检测特征图或者偏移特征图。

也就是说，无论视频帧是关键帧还是非关键帧，都有对应的检测特征图。关键帧对应的检测特征图利用单帧网络可以直接得到，而非关键帧对应的检测特征图是利用特征偏

移网络得到特征点偏移值后,根据对应上一帧视频帧的检测特征图和本视频帧的特征点偏移值计算得到的本视频帧的偏移特征图。即,当得到检测特征图和特征点偏移值时,根据检测特征图和特征点偏移值计算得到属于非关键帧的视频帧对应的偏移特征图。也就是说,根据检测特征图和偏移特征图中的坐标框直接确定视频帧中的物体信息。

5 根据上一帧视频帧的检测特征图和本视频帧的特征点偏移值得到本视频帧的偏移特征图的步骤,具体包括:基于特征点偏移值,对对应的上一帧视频帧的检测特征图或是偏移特征图进行偏移。若对应的上一帧视频帧为关键帧,则是检测特征图,对应的上一视频帧是非关键帧,则是偏移特征图。经过偏移得到的特征图即为本视频帧的偏移特征图,该偏移特征图即为本视频帧的检测特征图。也就是说,当特征点偏移值为向右平移 2 个单位,10 则将检测特征图中所对应的特征点向右平移 2 个单位。

上述物体检测方法中,当接收到待检测视频和物体检测指令后,根据物体检测指令确定检测方式,从而保证根据用户需求使用准确的检测方式。然后基于检测方式,且利用预设的单帧网络和特征偏移网络对待检测视频中的各视频帧进行检测,根据检测得到的检测特征图和特征点偏移值确定物体的信息,从而保证上述方法不仅能够通过单帧网络保证15 检测的精度,还能利用特征偏移网络得到能够表明各视频帧变化程度的特征点偏移值,不仅利用了帧间信息还提高了检测速度。

在一些实施例中,当检测方式为快速检测时,根据各视频帧的关键帧标志位将各视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络,以利用单帧网络或特征偏移网络对视频帧进行物体检测,得到检测结果的步骤,具体包括:依次将待检测视频中的各视频帧作为当前帧,获取关键帧标志位。基于获取的关键帧标志位,从单帧网络和特征偏移网络中确定当前帧对应的检测网络。利用当前帧对应的检测网络对当前帧进行物体20 检测,得到当前帧对应的检测结果。根据检测结果更改关键帧标志位,返回获取关键帧标志位的步骤,直至视频帧均得到对应的检测结果。

依次将待检测视频中的各视频帧作为当前帧,即为从各视频帧中确定当前帧时需按照视频帧的顺序依次确定,例如,若待检测视频中的视频帧按播放顺序为视频帧 1、视频帧 2 和视频帧 3,则首先将视频帧 1 作为当前帧,得到视频帧 1 的检测结果后在将视频帧25 2 作为当前帧,得到视频帧 2 的检测结果后在将视频帧 3 作为当前帧。

关键帧标志位则是指用于表示当前获取的视频帧是否为关键帧的标志,即用于表示当前帧是否为关键帧的标志。关键帧标志位包括 true(正确)和 false(错误),若关键帧标志位为 true,则表示当前获取的视频帧为关键帧,若关键帧标志位为 false,则表示30 当前获取的视频帧为非关键帧。

具体地,当确定一帧视频帧作为当前帧后,获取当前的关键帧标志位。基于关键帧标志位为当前帧选择对应的检测网络,检测网络为单帧网络和特征偏移网络中任意一种。若获取的关键帧标志位为 true,则表示当前帧为关键帧,关键帧对应的检测网络为单帧35 网络,则选择单帧网络作为当前帧的检测网络。而若关键帧标志位为 false,则表示当前

帧为非关键帧，非关键帧对应的检测网络为特征偏移网络，则选择特征偏移网络作为当前帧的检测网络。获取到当前帧对应的检测网络后，利用对应的检测网络对当前帧进行检测。也就是说，当前帧对应的检测网络为单帧网络，则利用单帧网络对当前帧进行检测。当前帧对应的检测网络为特征偏移网络，则利用特征偏移网络对当前帧进行检测。然后，根据检测得到的检测结果更改关键帧标志位，并且重新获取关键帧标志位，作为下个当前帧的关键帧标志位，直到所有的视频帧都有对应的检测结果为止。

由于是依次将待检测视频中的各视频帧作为当前帧，也就是说第一次作为当前帧的视频帧为待检测视频中的第一帧视频帧。而第一帧视频帧作为第一个当前帧是本次物体检测的开始，所以第一帧视频帧对应的关键帧标志位为预设的初始值。在本实施例中，由于特征偏移网络的检测需要当前帧以及对应的上一帧视频帧进行偏移检测，而第一帧视频帧是没有对应的上一帧视频帧的。因此，关键帧标志位为的初始值为 true，即待检测视频中的第一帧视频帧一定被作为关键帧。

在一些实施例中，如图 3 所示的流程图，当检测方式为快速检测时，基于检测方式，利用预设的单帧网络和特征偏移网络对各视频帧进行检测，得到检测结果具体包括以下步骤：

步骤 S302，确定当前帧。

步骤 S304，获取关键帧标志位。

步骤 S306，基于获取的关键帧标志位，从单帧网络和特征偏移网络中确定当前帧对应的检测网络。

步骤 S308，利用当前帧对应的检测网络对当前帧进行检测，得到当前帧对应的检测结果。

步骤 S310，根据检测结果更改关键帧标志位，返回步骤 S302，从各视频帧中确定当前帧，直至视频帧均得到对应的检测结果。

具体地，以待检测视频含有三帧视频帧为例进行说明，首先确定视频帧 1 为当前帧，获取当前的关键帧标志位。而关键帧标志位为 true，则表示视频帧 1 为关键帧。那么，视频帧 1 对应的检测网络则为单帧网络，则选择单帧网络作为视频帧 1 的检测网络。利用单帧网络对视频帧 1 进行检测，得到视频帧 1 对应的检测结果。并且，根据视频帧 1 的检测结果更改关键帧标志位。然后，按顺序确定视频帧 2 为当前帧，获取当前的关键帧标志位，而该关键帧标志位已经根据视频帧 1 的检测结果进行更改了。可能该关键帧标志位没有变，还是 true，也可能已经变为 false 了，假设经过视频帧 1 的检测结果将该关键帧标志位改为了 false，则表示视频帧 2 为非关键帧。那么，视频帧 2 对应的检测网络为特征偏移网络，则选择特征偏移网络作为视频帧 2 的检测网络。利用特征偏移网络对视频帧 2 进行检测，得到视频帧 2 对应的检测结果。同样的，根据视频帧 2 的检测结果再次更改关键帧标志位。最后，确定视频帧 3 为当前帧，具体检测过程同视频帧 1 和视频帧 2 一致，在此不再赘述。

在本实施例中，利用视频帧之间信息冗余较大的特点，通过划分关键帧和非关键帧，使用单帧网络对关键帧进行检测。利用单帧网络保证检测精度的同时下，采用特征偏移网络对非关键帧进行偏移检测，只需要快速检测确定连续帧之间物体的变化程度就能得到对应的特征图，不仅利用了帧间信息，还提升了检测速度。

5 在一些实施例中，基于获取的关键帧标志位，从单帧网络和特征偏移网络中确定当前帧对应的检测网络，具体包括：当根据关键帧标志位确定当前帧为关键帧时，将单帧网络作为当前帧对应的检测网络。当根据关键帧标志位确定当前帧为非关键帧时，将特征偏移网络作为当前帧对应的检测网络。

具体地，若获取的关键帧标志位为 true，则表示当前帧为关键帧，关键帧对应的检测网络为单帧网络，则选择单帧网络作为当前帧的检测网络。而若关键帧标志位为 false，则表示当前帧为非关键帧，非关键帧对应的检测网络为特征偏移网络，则选择特征偏移网络作为当前帧的检测网络。

10 在一些实施例中，利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果的步骤，具体包括：若当前帧对应的检测网络为单帧网络，则利用单帧网络对当前帧进行单帧的物体检测，得到的检测特征图作为当前帧对应的检测结果。若当前帧对应的检测网络为特征偏移网络，则利用特征偏移网络对当前帧以及当前帧对应的上一帧视频帧进行特征点偏移值的检测，得到的特征点偏移值作为当前帧对应的检测结果。

具体地，当视频帧对应的检测网络为单帧网络时，即可直接将该视频帧输入至单帧网络中，利用单帧网络直接对该视频帧进行单帧检测，得到的检测结果为该视频帧的检测特征图。而，当视频帧对应的检测网络为特征偏移网络时，需要检测该视频帧的偏移值，因此除了将该视频帧输入至特征偏移网络之外，还有将该视频帧对应的上一帧视频帧输入至特征偏移网络一起进行偏移检测。即利用特征偏移网络得到的检测结果为该视频帧中的特征点相对上一帧视频帧中的特征点的偏移值。

在一些实施例中，根据检测结果更改关键帧标志位的步骤，具体包括：

25 当检测结果为检测特征图时，更改关键帧标志位。或者，

当检测结果为特征点偏移值时，基于特征点偏移值计算得到物体偏移均值，并将物体偏移均值累加至预设的物体偏移总值。若物体偏移总值不小于门限值，则将物体偏移总值清零，以及更改关键帧标志位。

具体地，当检测结果为检测特征图时，直接将关键帧标志位进行更改。例如关键帧标志位为 true，则更改为 false。

30 当检测结果为特征点偏移值时，基于该特征点偏移值计算得到物体偏移均值，然后将物体偏移均值累加到物体偏移总值中，最后根据物体偏移总值决定是否更改标志位。即，当物体偏移总值大于等于预设的门限值时，则更改关键帧标志位，并同时将物体偏移总值清零。反之，物体偏移总值小于预设的门限值时，不更改关键帧标志位也不清零。预设的物体偏移总值的初始值为 0，通过一次次将计算得到的物体偏移均值进行累加，使得物体

偏移总值达到门限值时才会再次清除到 0，而物体偏移总值一旦清零就意味着需要更改关键帧标志位。

在一些实施例中，计算物体偏移均值具体包括：根据视频帧的特征点偏移值，对对应上一帧视频帧对应的特征图进行偏移，得到视频帧对应的偏移特征图，由于上一帧视频帧中的特征图包括坐标框，即根据特征点偏移值得到的偏移特征图也有对应的偏移的坐标框。因此，确定该偏移特征图中的坐标框，获取该坐标框内所有特征点对应的特征点偏移值，基于坐标框内特征点对应的特征点偏移值做均值运算，得到的均值为物体偏移均值。上一帧视频帧为关键帧时，上一帧视频帧对应的特征图为检测特征图。上一帧视频帧为非关键帧时，上一帧视频帧对应的特征图为偏移特征图。

在一些实施例中，如图 4 所示，当检测方式为精度检测时，检测结果中的特征点偏移值包括特征点正偏移值和特征点反偏移值。则，当检测方式为精度检测时，分别将各视频帧输入单帧网络和特征偏移网络，以利用单帧网络对各视频帧进行物体检测，以及利用特征偏移网络根据各视频帧的播放顺序对各视频帧进行多次的物体检测的步骤，具体包括以下步骤：

步骤 S402，利用单帧网络对各视频帧进行单帧的物体检测，得到各视频帧对应的检测特征图。

当确定检测方式为精度检测时，表示需要进行高精度的物体检测。则无需如快速检测一般将各视频帧进行关键帧、非关键帧的划分后，再输入到对应的网络进行检测。

具体地，将待检测视频中的各视频帧均输入到单帧网络中进行检测，得到各视频帧对应的检测特征图。例如，假设待检测视频帧包括视频帧 1、视频帧 2 和视频帧 3，则将视频帧 1、视频帧 2 和视频帧 3 均输入到单帧网络中，分别得到视频帧 1 对应的检测特征图、视频帧 2 对应的检测特征图和视频帧 3 对应的检测特征图。

步骤 S404，利用特征偏移网络对各视频帧进行特征点正序偏移值的检测，得到各视频帧对应的特征点正偏移值。

步骤 S406，利用特征偏移网络对各视频帧进行特征点倒序偏移值的检测，得到各视频帧对应的特征点反偏移值。

特征点正偏移值是指视频帧中的特征点相对上一帧视频帧中特征点的偏移，而特征点反偏移值则是指视频帧中的特征点相对下一帧视频帧中特征点的偏移。正序偏移检测是指特征偏移网络接收正序的视频帧，倒序偏移检测是指特征偏移网络接收倒序的视频帧。正序即是按照待检测视频的正播放顺序，例如视频帧 1、视频帧 2 和视频帧 3。而倒序即是按照待检测视频的反播放顺序，例如视频帧 3、视频帧 2 和视频帧 1。

具体地，将待检测视频中的各视频帧均输入到特征偏移网络中进行检测。而为了充分利用连续视频帧的上下文信息，除了将待检测视频中的各视频帧按照正序输入至特征偏移网络中进行正序偏移检测之外，还将待检测视频中的各视频帧倒序输入至特征偏移网络中进行反序偏移检测。通过正序偏移检测可以得到每一帧视频帧相对上一帧视频帧的特征

点偏移值，即特征点正偏移值，以视频帧 1、视频帧 2 和视频帧 3 为例，即根据视频帧 1 和视频帧 2 计算得到视频帧 2 相对于视频帧 1 的特征点正偏移值。同样的，通过倒序偏移检测可以得到每一帧视频帧相对下一帧视频帧的特征点偏移值，即特征点反偏移值，以视频帧 1、视频帧 2 和视频帧 3 为例，即根据视频帧 2 和视频帧 3 计算得到视频帧 2 相对于视频帧 3 的特征点反偏移值。

在本实施例中，通过对待检测视频中的各视频帧进行三遍不同的物体检测，虽然进行三次检测使得检测速度有所下降，但是充分利用了连续视频帧的上下文信息，通过更多的信息进行检测从而提高物体检测的精度，尤其在运动比较模糊的情况下运用精度检测，能够提高检测的准确率。

在一些实施例中，当检测方式为精度检测，且检测结果中的特征点偏移值包括特征点正偏移值和特征点反偏移值时，则根据特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据检测特征图和偏移特征图中的坐标框确定物体信息的步骤，具体包括：根据视频帧对应的特征点正偏移值，将视频帧对应的上一帧视频帧的检测特征图中的特征点进行偏移，生成视频帧对应的正偏移特征图。根据视频帧对应的特征点反偏移值，将视频帧对应的下一帧视频帧的检测特征图中的特征点进行偏移，生成视频帧对应的反偏移特征图。将视频帧的检测特征图、正偏移特征图以及反偏移特征图进行特征融合，生成最终的特征图，根据最终的特征图中的坐标框确定物体信息。

具体地，基于视频帧的特征点正偏移值，对对应的上一帧视频帧的检测特征图进行偏移，得到视频帧对应的正偏移特征图。例如，基于视频帧 2 的特征点正偏移值对视频帧 1 的检测特征图进行偏移，得到的特征图为视频帧 2 对应的正偏移特征图。相同的，基于视频帧的特征点反偏移值，对对应的下一帧视频帧的检测特征图进行偏移，得到视频帧对应的反偏移特征图。例如，基于视频帧 2 的特征点反偏移值对视频帧 3 的检测特征图进行偏移，得到的特征图为视频帧 2 对应的反偏移特征图。而确定物体信息时，即将视频帧对应的三张特征图进行特征融合，即将视频帧对应的检测特征图、正偏移特征图以及反偏移特征图进行特征融合，生成最终的特征图。特征融合包括但不限于取平均值或者加权平均的方法。例如，采用平均值即三张特征图相加除以 3，加权平均则是根据预设的权重分别相乘后相加即可。由于待检测视频中的第一帧视频帧没有对应的上一帧视频帧，以及最后一帧视频帧没有对应的下一帧视频帧，因此，第一帧视频帧无法根据上一帧视频帧的检测特征图生成没有正偏移特征图，即第一帧视频帧没有正偏移特征图。而最后一帧视频帧无法根据下一帧视频帧的检测特征图生成反偏移特征图，即最后一帧视频帧没有反偏移特征图。

在一些实施例中，关键帧标志位根据预设的等间隔规则进行设定，等间隔规则是指包括关键帧划分的规则，例如，若等间隔规则规定每隔 2 帧取一帧视频帧为关键帧，则相隔的两帧视频帧为非关键帧。也就是说，每 3 帧中有两帧非关键帧和一帧关键帧。在待检测视频帧的第一帧一定为关键帧的基础上，则可以简单表示为“关键帧、非关键帧、非关键

帧、关键帧……”。同样的，当根据等间隔规则确定好关键帧和非关键帧后，同样是关键帧输入至单帧网络进行单帧检测，得到检测特征图。而为非关键帧的视频帧则同样将该非关键帧以及对应的上一帧视频帧输入至特征偏移网络进行偏移检测，从而得到非关键帧的特征点偏移值。也就是说，当接收到待检测视频后，根据预设的等间隔规则预先将待检测视频中的各视频帧进行关键帧划分，确定各视频帧是关键帧还是非关键帧，后续直接调用对应的网络进行检测，无需对关键帧标志位设定初始值，也无需进行关键标志位的更改。在本实施例中，利用等间隔循环的方法可以适用于偏静止的视频场景，可以进一步提升检测速度。

应该理解的是，虽然图 2-4 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2-4 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

在一些实施例中，如图 5 所示，提供了一种物体检测装置，包括：接收模块 502、确定模块 504、检测模块 506 和偏移模块 508，

接收模块 502，用于接收终端发送的待检测视频和物体检测指令，待检测视频中包括视频帧。

确定模块 504，用于根据物体检测指令确定检测方式，检测方式包括快速检测和精度检测。

检测模块 506，用于当检测方式为快速检测时，根据各视频帧的关键帧标志位将各视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果。

检测模块 506 还用于当检测方式为精度检测时，分别将各视频帧输入单帧网络和特征偏移网络，以利用单帧网络对各视频帧进行物体检测，以及利用特征偏移网络根据各视频帧的播放顺序对各视频帧进行多次的物体检测；检测结果包括检测特征图和特征点偏移值。

偏移模块 508，用于根据特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据检测特征图和偏移特征图中的坐标框确定物体信息。

在一些实施例中，检测模块 506 还用于依次将待检测视频中的各视频帧作为当前帧，获取关键帧标志位；基于获取的关键帧标志位，从单帧网络和特征偏移网络中确定当前帧对应的检测网络；利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果；根据检测结果更改关键帧标志位，返回获取关键帧标志位的步骤，直至各视频帧均得到对应的检测结果。

在一些实施例中，检测模块 506 还用于当根据关键帧标志位确定当前帧为关键帧时，将单帧网络作为当前帧对应的检测网络；当根据关键帧标志位确定当前帧为非关键帧时，将特征偏移网络作为当前帧对应的检测网络。

在一些实施例中，检测模块 506 还用于若当前帧对应的检测网络为单帧网络，则利用单帧网络对当前帧进行单帧的物体检测，得到的检测特征图作为当前帧对应的检测结果；若当前帧对应的检测网络为特征偏移网络，则利用特征偏移网络对当前帧以及当前帧对应的上一帧视频帧进行特征点偏移值的检测，得到的特征点偏移值作为当前帧对应的检测结果。

在一些实施例中，检测模块 506 还用于当检测结果为检测特征图时，更改关键帧标志位，并返回获取关键帧标志位的步骤，直至各视频帧均得到对应的检测结果；或

当检测结果为特征点偏移值时，基于特征点偏移值计算得到物体偏移均值，并将物体偏移均值累加至预设的物体偏移总值；若物体偏移总值不小于门限值，则将物体偏移总值清零，以及更改关键帧标志位，并返回获取关键帧标志位的步骤，直至各视频帧均得到对应的检测结果。

在一些实施例中，检测模块 506 还用于利用单帧网络对各视频帧进行单帧的物体检测，得到各视频帧对应的检测特征图；利用特征偏移网络对各视频帧进行特征点正序偏移值的检测，得到各视频帧对应的特征点正偏移值；利用特征偏移网络对各视频帧进行特征点倒序偏移值的检测，得到各视频帧对应的特征点反偏移值。

在一些实施例中，偏移模块 508 还用于根据视频帧对应的特征点正偏移值，将视频帧对应的上一帧视频帧的检测特征图中的特征点进行偏移，生成视频帧对应的正偏移特征图；根据视频帧对应的特征点反偏移值，将视频帧对应的下一帧视频帧的检测特征图中的特征点进行偏移，生成视频帧对应的反偏移特征图；将视频帧的检测特征图、正偏移特征图以及反偏移特征图进行特征融合，生成最终的特征图；根据最终的特征图中的坐标框确定物体信息。

关于物体检测装置的具体限定可以参见上文中对于物体检测方法的限定，在此不再赘述。上述物体检测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。

在一些实施例中，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 6 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。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非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性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物体检测方法。

本领域技术人员可以理解，图 6 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

一种计算机设备，包括存储器和一个或多个处理器，存储器中存储有计算机可读指令，计算机可读指令被处理器执行时实现本申请任意一个实施例中提供的物体检测方法的步骤。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器实现本申请任意一个实施例中提供的物体检测方法的步骤。

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，是可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一非易失性计算机可读取存储介质中，该计算机程序在执行时，可包括如上述各方法的实施例的流程。本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程 ROM（PROM）、电可编程 ROM（EPROM）、电可擦除可编程 ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM（SRAM）、动态 RAM（DRAM）、同步 DRAM（SDRAM）、双数据率 SDRAM（DDRSDRAM）、增强型 SDRAM（ESDRAM）、同步链路（Synchlink） DRAM（SLDRAM）、存储器总线（Rambus）直接 RAM（RDRAM）、直接存储器总线动态 RAM（DRDRAM）、以及存储器总线动态 RAM（RDRAM）等。

以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求书

1、一种物体检测方法，包括：

服务器接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

所述服务器根据所述物体检测指令的类型确定检测方式，所述检测方式包括快速检测和精度检测；

5 当所述检测方式为所述快速检测时，所述服务器根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

当所述检测方式为所述精度检测时，所述服务器分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述
10 检测结果包括检测特征图和特征点偏移值；及

所述服务器根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

2、根据权利要求 1 所述的方法，其特征在于，所述服务器根据各所述视频帧的关键
15 帧标志位将各所述视频帧输入至预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对各所述视频帧进行物体检测，得到检测结果，包括：

所述服务器依次将待检测视频中的各所述视频帧作为当前帧，获取关键帧标志位；

所述服务器基于获取的关键帧标志位，从所述单帧网络和特征偏移网络中确定当前帧对应的检测网络；

20 所述服务器利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果；及

所述服务器根据所述检测结果更改所述关键帧标志位，返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

3、根据权利要求 2 所述的方法，其特征在于，所述服务器基于获取的关键帧标志位，
25 从所述单帧网络和特征偏移网络中确定当前帧对应的检测网络，包括：

当所述服务器根据所述关键帧标志位确定当前帧为关键帧时，将所述单帧网络作为当前帧对应的检测网络；及

当所述服务器根据所述关键帧标志位确定当前帧为非关键帧时，将所述特征偏移网络作为当前帧对应的检测网络。

30 4、根据权利要求 2 所述的方法，其特征在于，所述服务器利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果，包括：

若当前帧对应的检测网络为所述单帧网络，所述服务器则利用所述单帧网络对所述当前帧进行单帧的物体检测，得到的检测特征图作为当前帧对应的检测结果；及

若当前帧对应的检测网络为所述特征偏移网络，所述服务器则利用所述特征偏移网络

对当前帧以及当前帧对应的上一帧视频帧进行特征点偏移值的检测，得到的特征点偏移值作为当前帧对应的检测结果。

5 5、根据权利要求 2 所述的方法，其特征在于，所述服务器根据所述检测结果更改所述关键帧标志位，返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果，包括：

当所述检测结果为检测特征图时，所述服务器更改所述关键帧标志位，并返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果；

或

10 当所述检测结果为特征点偏移值时，所述服务器基于所述特征点偏移值计算得到物体偏移均值，并将所述物体偏移均值累加至预设的物体偏移总值；及

若所述物体偏移总值不小于门限值，所述服务器则将所述物体偏移总值清零，以及更改所述关键帧标志位，并返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

15 6、根据权利要求 1 所述的方法，其特征在于，所述特征点偏移值包括特征点正偏移值和特征点反偏移值；

所述服务器分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络和所述特征偏移网络分别对各所述视频帧进行物体检测，得到检测结果，包括：

当所述服务器确定所述检测方式为所述精度检测时，利用所述单帧网络对各所述视频帧进行单帧的物体检测，得到各所述视频帧对应的检测特征图；

20 所述服务器利用特征偏移网络对各所述视频帧进行特征点正序偏移值的检测，得到各所述视频帧对应的特征点正偏移值；及

所述服务器利用特征偏移网络对各所述视频帧进行特征点倒序偏移值的检测，得到各所述视频帧对应的特征点反偏移值。

25 7、根据权利要求 6 所述的方法，其特征在于，所述服务器根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息，包括：

所述服务器根据所述视频帧对应的特征点正偏移值，将所述视频帧对应的上一帧视频帧的检测特征图中的特征点进行偏移，生成所述视频帧对应的正偏移特征图；

30 所述服务器根据所述视频帧对应的特征点反偏移值，将所述视频帧对应的下一帧视频帧的检测特征图中的特征点进行偏移，生成所述视频帧对应的反偏移特征图；

所述服务器将所述视频帧的检测特征图、正偏移特征图以及反偏移特征图进行特征融合，生成最终的特征图；及

根据所述最终的特征图中的坐标框确定物体信息。

8、一种物体检测装置，包括：

35 接收模块，用于接收终端发送的待检测视频和物体检测指令，所述待检测视频中包

括视频帧；

确定模块，用于根据所述物体检测指令的类型确定检测方式，所述检测方式包括快速检测和精度检测；

5 检测模块，用于当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

10 检测模块还用于当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

偏移模块，用于根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

9、根据权利要求8所述的装置，其特征在于，所述检测模块还用于

依次将所述待检测视频中的各所述视频帧作为当前帧，获取关键帧标志位；

15 基于获取的关键帧标志位，从所述单帧网络和特征偏移网络中确定当前帧对应的检测网络；

利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果；及根据所述检测结果更改所述关键帧标志位，返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

20 10、根据权利要求9所述的装置，其特征在于，所述检测模块还用于

当根据所述关键帧标志位确定当前帧为关键帧时，将所述单帧网络作为当前帧对应的检测网络；及

当根据所述关键帧标志位确定当前帧为非关键帧时，将所述特征偏移网络作为当前帧对应的检测网络。

25 11、一种计算机设备，包括存储器及一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述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以下步骤：

接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

30 根据所述物体检测指令的类型确定检测方式，所述检测方式包括快速检测和精度检测；

当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

35 当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移

网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

5 12、根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

依次将所述待检测视频中的各所述视频帧作为当前帧，获取关键帧标志位；

基于获取的关键帧标志位，从所述单帧网络和特征偏移网络中确定当前帧对应的检测网络；

10 利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果；及根据所述检测结果更改所述关键帧标志位，返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

13、根据权利要求 12 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

15 当根据所述关键帧标志位确定当前帧为关键帧时，将所述单帧网络作为当前帧对应的检测网络；及

当根据所述关键帧标志位确定当前帧为非关键帧时，将所述特征偏移网络作为当前帧对应的检测网络。

20 14、根据权利要求 12 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

若当前帧对应的检测网络为所述单帧网络，则利用所述单帧网络对所述当前帧进行单帧的物体检测，得到的检测特征图作为当前帧对应的检测结果；及

25 若当前帧对应的检测网络为所述特征偏移网络，则利用所述特征偏移网络对当前帧以及当前帧对应的上一帧视频帧进行特征点偏移值的检测，得到的特征点偏移值作为当前帧对应的检测结果。

15、根据权利要求 12 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

当所述检测结果为检测特征图时，更改所述关键帧标志位，并返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果；

30 或

当所述检测结果为特征点偏移值时，基于所述特征点偏移值计算得到物体偏移均值，并将所述物体偏移均值累加至预设的物体偏移总值；及

若所述物体偏移总值不小于门限值，则将所述物体偏移总值清零，以及更改所述关键帧标志位，并返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

35 16、一个或多个存储有计算机可读指令的非易失性计算机可读存储介质，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以下步骤：

接收终端发送的待检测视频和物体检测指令，所述待检测视频中包括视频帧；

根据所述物体检测指令的类型确定检测方式，所述检测方式包括快速检测和精度检测；

5 当所述检测方式为所述快速检测时，根据各所述视频帧的关键帧标志位将各所述视频帧输入预先训练好的用于物体检测的单帧网络或特征偏移网络，以利用所述单帧网络或特征偏移网络对所述视频帧进行物体检测，得到检测结果；

10 当所述检测方式为所述精度检测时，分别将各所述视频帧输入所述单帧网络和所述特征偏移网络，以利用所述单帧网络对各所述视频帧进行物体检测，以及利用所述特征偏移网络根据各所述视频帧的播放顺序对各所述视频帧进行多次的物体检测；所述检测结果包括检测特征图和特征点偏移值；及

根据所述特征点偏移值将对应的检测特征图中的特征点进行偏移，得到偏移特征图，并根据所述检测特征图和偏移特征图中的坐标框确定物体信息。

15 17、根据权利要求 16 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

依次将所述待检测视频中的各所述视频帧作为当前帧，获取关键帧标志位；

基于获取的关键帧标志位，从所述单帧网络和特征偏移网络中确定当前帧对应的检测网络；

20 利用当前帧对应的检测网络对当前帧进行物体检测，得到当前帧对应的检测结果；及根据所述检测结果更改所述关键帧标志位，返回获取关键帧标志位的步骤，直至各所述视频帧均得到对应的检测结果。

18、根据权利要求 17 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

25 当根据所述关键帧标志位确定当前帧为关键帧时，将所述单帧网络作为当前帧对应的检测网络；及

当根据所述关键帧标志位确定当前帧为非关键帧时，将所述特征偏移网络作为当前帧对应的检测网络。

19、根据权利要求 17 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

30 若当前帧对应的检测网络为所述单帧网络，则利用所述单帧网络对所述当前帧进行单帧的物体检测，得到的检测特征图作为当前帧对应的检测结果；及

若当前帧对应的检测网络为所述特征偏移网络，则利用所述特征偏移网络对当前帧以及当前帧对应的上一帧视频帧进行特征点偏移值的检测，得到的特征点偏移值作为当前帧对应的检测结果。

35 20、根据权利要求 17 所述的存储介质，其特征在于，所述计算机可读指令被所述处

理器执行时还执行以下步骤:

当所述检测结果为检测特征图时,更改所述关键帧标志位,并返回获取关键帧标志位的步骤,直至各所述视频帧均得到对应的检测结果;

或

- 5 当所述检测结果为特征点偏移值时,基于所述特征点偏移值计算得到物体偏移均值,并将所述物体偏移均值累加至预设的物体偏移总值;及

若所述物体偏移总值不小于门限值,则将所述物体偏移总值清零,以及更改所述关键帧标志位,并返回获取关键帧标志位的步骤,直至各所述视频帧均得到对应的检测结果。

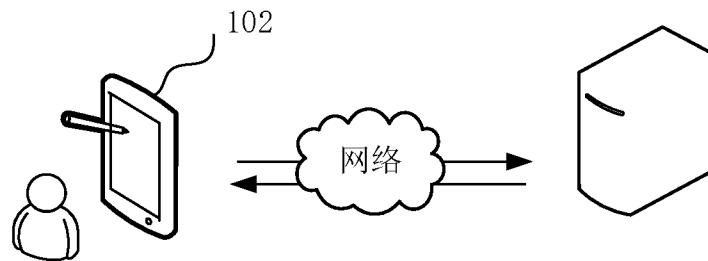


图 1

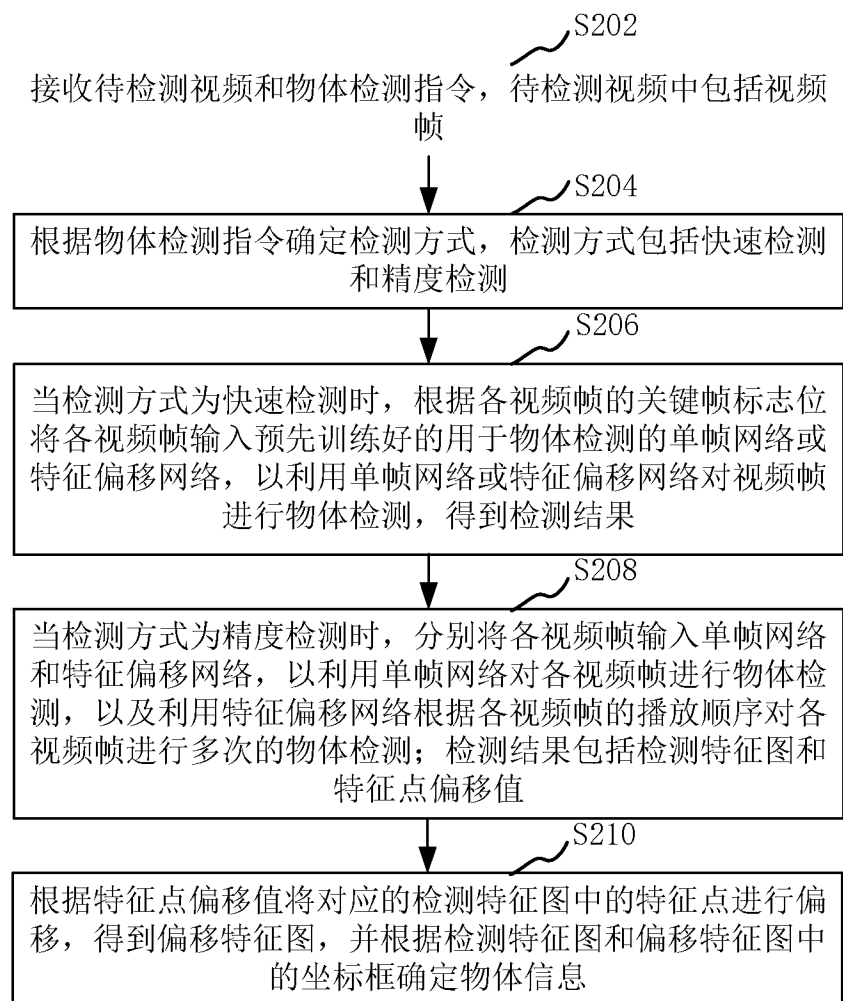


图 2

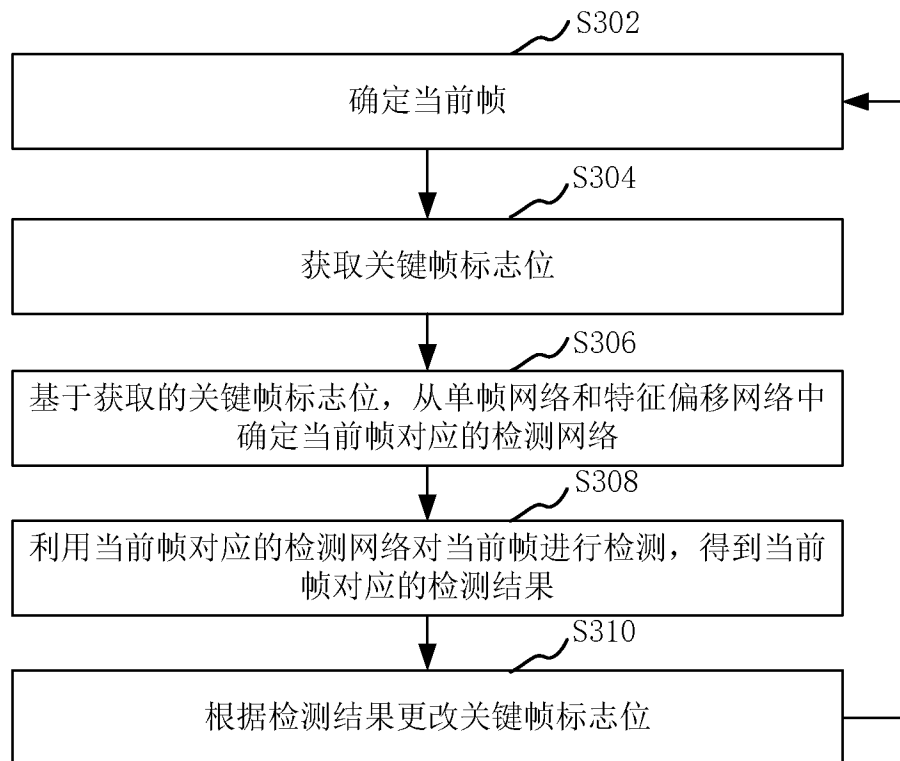


图 3

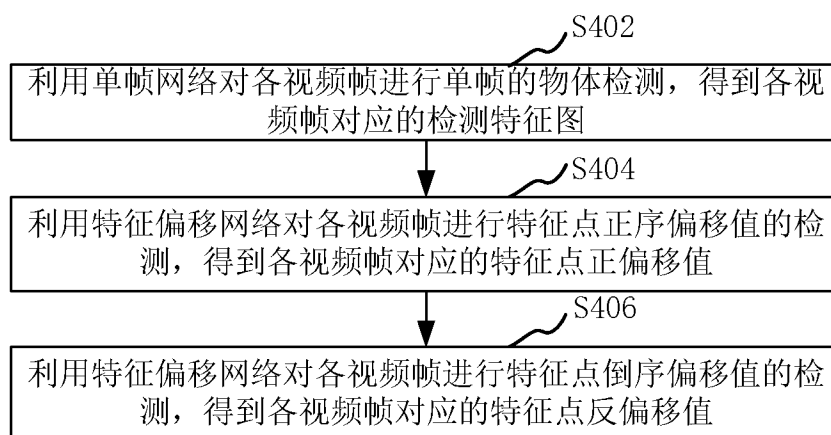


图 4

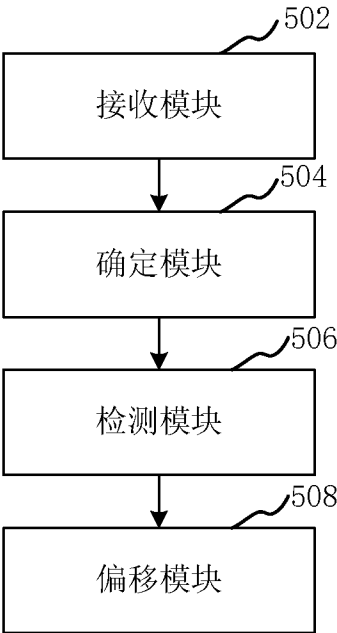


图 5

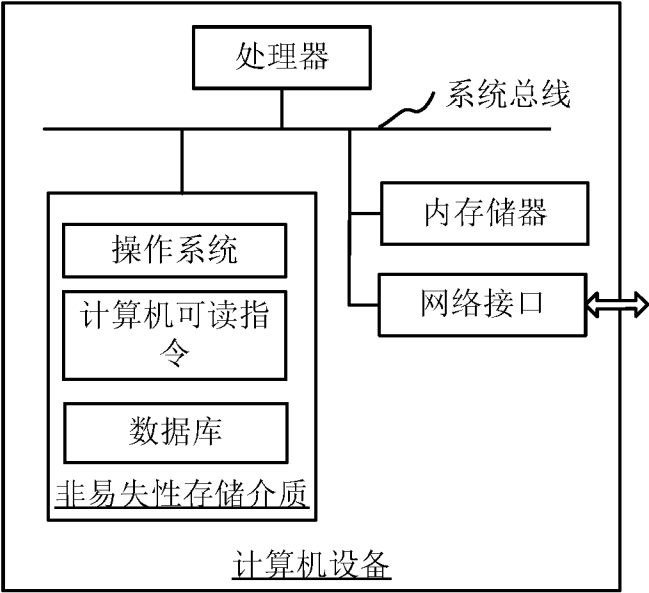


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/103195

A. CLASSIFICATION OF SUBJECT MATTER G06K 9/00(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC														
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS, CNTXT, CNKI, VEN, USTXT, EPTXT, WOTXT: 检测, 速度, 视频帧, 单帧, 特征, 偏移, 精度, detect, speed, video frame, single frame, feature, offset, precision														
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 5px;"> <thead> <tr> <th style="width: 10%;">Category*</th> <th style="width: 70%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width: 20%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">A</td> <td>CN 108256506 A (BEIJING SENSETIME SCIENCE TECHNOLOGY DEVELOPMENT CO., LTD.) 06 July 2018 (2018-07-06) description, paragraphs 4-51, figures 1-3</td> <td style="text-align: center;">1-20</td> </tr> <tr> <td style="text-align: center;">A</td> <td>CN 107578011 A (COLD AND ARID REGIONS ENVIRONMENTAL AND ENGINEERING RESEARCH INSTITUTE, CHINESE ACADEMY OF SCIENCES) 12 January 2018 (2018-01-12) description, paragraphs 28-91</td> <td style="text-align: center;">1-20</td> </tr> <tr> <td style="text-align: center;">A</td> <td>US 9449216 B1 (A9.COM, INC.) 20 September 2016 (2016-09-20) entire document</td> <td style="text-align: center;">1-20</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	CN 108256506 A (BEIJING SENSETIME SCIENCE TECHNOLOGY DEVELOPMENT CO., LTD.) 06 July 2018 (2018-07-06) description, paragraphs 4-51, figures 1-3	1-20	A	CN 107578011 A (COLD AND ARID REGIONS ENVIRONMENTAL AND ENGINEERING RESEARCH INSTITUTE, CHINESE ACADEMY OF SCIENCES) 12 January 2018 (2018-01-12) description, paragraphs 28-91	1-20	A	US 9449216 B1 (A9.COM, INC.) 20 September 2016 (2016-09-20) entire document	1-20
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
A	CN 108256506 A (BEIJING SENSETIME SCIENCE TECHNOLOGY DEVELOPMENT CO., LTD.) 06 July 2018 (2018-07-06) description, paragraphs 4-51, figures 1-3	1-20												
A	CN 107578011 A (COLD AND ARID REGIONS ENVIRONMENTAL AND ENGINEERING RESEARCH INSTITUTE, CHINESE ACADEMY OF SCIENCES) 12 January 2018 (2018-01-12) description, paragraphs 28-91	1-20												
A	US 9449216 B1 (A9.COM, INC.) 20 September 2016 (2016-09-20) entire document	1-20												
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.														
<table style="width: 100%;"> <tr> <td style="width: 50%; vertical-align: top;"> * 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 </td> <td style="width: 50%; vertical-align: top;"> "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 </td> </tr> </table>			* 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										
* 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 09 March 2020		Date of mailing of the international search report 17 March 2020												
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/CN2019/103195

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108256506	A	06 July 2018	None			
CN	107578011	A	12 January 2018	None			
US	9449216	B1	20 September 2016	US	2016358015	A1	08 December 2016

国际检索报告

国际申请号

PCT/CN2019/103195

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类														
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, CNTXT, CNKI, VEN, USTXT, EPTXT, WOTXT: 检测, 速度, 视频帧, 单帧, 特征, 偏移, 精度, detect, speed, video frame, single frame, feature, offset, precision														
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>CN 108256506 A (北京市商汤科技开发有限公司) 2018年 7月 6日 (2018 - 07 - 06) 说明书第4-51段, 附图1-3</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 107578011 A (中国科学院寒区旱区环境与工程研究所) 2018年 1月 12日 (2018 - 01 - 12) 说明书第28-91段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 9449216 B1 (A9 COM INC) 2016年 9月 20日 (2016 - 09 - 20) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	CN 108256506 A (北京市商汤科技开发有限公司) 2018年 7月 6日 (2018 - 07 - 06) 说明书第4-51段, 附图1-3	1-20	A	CN 107578011 A (中国科学院寒区旱区环境与工程研究所) 2018年 1月 12日 (2018 - 01 - 12) 说明书第28-91段	1-20	A	US 9449216 B1 (A9 COM INC) 2016年 9月 20日 (2016 - 09 - 20) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
A	CN 108256506 A (北京市商汤科技开发有限公司) 2018年 7月 6日 (2018 - 07 - 06) 说明书第4-51段, 附图1-3	1-20												
A	CN 107578011 A (中国科学院寒区旱区环境与工程研究所) 2018年 1月 12日 (2018 - 01 - 12) 说明书第28-91段	1-20												
A	US 9449216 B1 (A9 COM INC) 2016年 9月 20日 (2016 - 09 - 20) 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 2020年 3月 9日		国际检索报告邮寄日期 2020年 3月 17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 杨越松 电话号码 86-(20)-28950385												

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/103195

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	108256506	A	2018年 7月 6日	无	
CN	107578011	A	2018年 1月 12日	无	
US	9449216	B1	2016年 9月 20日	US 2016358015 A1	2016年 12月 8日