



(12) 发明专利

(10) 授权公告号 CN 112689151 B

(45) 授权公告日 2023.04.18

(21) 申请号 202011416720.8

H04N 21/44 (2011.01)

(22) 申请日 2020.12.07

G06V 20/40 (2022.01)

(65) 同一申请的已公布的文献号

(56) 对比文件

申请公布号 CN 112689151 A

CN 110850983 A, 2020.02.28

WO 2019227905 A1, 2019.12.05

(43) 申请公布日 2021.04.20

审查员 杜乾敏

(73) 专利权人 深圳盈天下视觉科技有限公司

地址 518000 广东省深圳市福田区福保街  
道保税区广兰道6号深装总大厦A座  
210-216室

(72) 发明人 刘严 陈权 邓生全

(74) 专利代理机构 深圳中一联合知识产权代理  
有限公司 44414

专利代理师 李红艳

(51) Int. Cl.

H04N 21/2187 (2011.01)

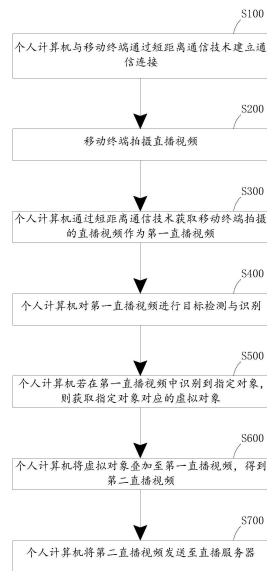
权利要求书2页 说明书12页 附图3页

(54) 发明名称

直播方法、装置、计算机设备和存储介质

(57) 摘要

本申请公开了一种直播方法、装置、计算机设备和存储介质,属于直播技术领域。所述方法应用于个人计算机,包括:通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频;对第一直播视频进行目标检测与识别;若在第一直播视频中识别到指定对象,则获取指定对象对应的虚拟对象;将虚拟对象叠加至第一直播视频,得到第二直播视频;将第二直播视频发送至直播服务器。在本申请中,通过移动终端进行直播视频的拍摄,更便于移动拍摄和移动直播,利用个人计算机存储量大和运算速度快的优势,对移动终端拍摄的直播视频进行处理,即在直播视频中叠加与该直播视频中的指定对象对应的虚拟对象,以增强直播的表现力。



1. 一种直播方法,其特征在于,应用于个人计算机,所述方法包括:  
通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频;  
对所述第一直播视频进行目标检测与识别;  
若在所述第一直播视频中识别到指定对象,则获取所述指定对象对应的虚拟对象;所述指定对象包括:指定物体;  
将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频;  
将所述第二直播视频发送至直播服务器;  
所述若在所述第一直播视频中识别到指定对象,则获取所述指定对象对应的虚拟对象,包括:  
若在所述第一直播视频中识别到指定对象,则根据所述指定对象在所述第一直播视频中的位置,获取所述指定对象对应的虚拟对象,所述在所述第一直播视频中的位置包括所述指定对象在第一直播视频的中间位置、或边缘位置。
2. 如权利要求1所述的方法,其特征在于,所述短距离通信技术包括蓝牙技术、紫蜂技术、无线保真技术、串行总线技术中的至少一种。
3. 如权利要求1所述的方法,其特征在于,所述指定对象为指定物体,所述虚拟对象为增强现实模型。
4. 如权利要求1所述的方法,其特征在于,所述对所述第一直播视频进行目标检测与识别,包括:  
将所述第一直播视频中的每帧视频图像输入目标识别模型,由所述目标识别模型输出所述每帧视频图像中的每个检测框的位置和所述每个检测框内的指定对象的类别;  
将所述每个检测框的位置确定为所述每个检测框内的指定对象的位置。
5. 如权利要求1-4任一所述的方法,其特征在于,所述将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频,包括:  
根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频。
6. 如权利要求5所述的方法,其特征在于,所述根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频,包括:  
根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象与所述第一直播视频叠加显示;  
若检测到针对所显示的所述虚拟对象的调整指令,则根据所述调整指令调整所述虚拟对象在与所述第一直播视频叠加时的尺寸、位置中的至少一个;  
若检测到确认指令,则将叠加有所述虚拟对象的第一直播视频作为第二直播视频。
7. 一种直播装置,其特征在于,应用于个人计算机,所述装置包括:  
第一获取模块,用于通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频;  
检测识别模块,用于对所述第一直播视频进行目标检测与识别;  
第二获取模块,用于若在所述第一直播视频中识别到指定对象,则获取所述指定对象对应的虚拟对象;所述指定对象包括:指定物体;  
叠加模块,用于将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频;

发送模块,用于将所述第二直播视频发送至直播服务器;

所述若在所述第一直播视频中识别到指定对象,则获取所述指定对象对应的虚拟对象,包括:

若在所述第一直播视频中识别到指定对象,则根据所述指定对象在所述第一直播视频中的位置,获取所述指定对象对应的虚拟对象,所述在所述第一直播视频中的位置包括所述指定对象在第一直播视频的中间位置、或边缘位置。

8.一种计算机设备,其特征在于,所述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程序,所述计算机程序被所述处理器执行时实现如权利要求1至6任一项所述的方法。

9.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时实现如权利要求1至6任一项所述的方法。

## 直播方法、装置、计算机设备和存储介质

### 技术领域

[0001] 本申请涉及直播技术领域，特别涉及一种直播方法、装置、计算机设备和存储介质。

### 背景技术

[0002] 随着科技的发展，移动终端（如手机、平板电脑等）的功能越来越多。比如，用户可以使用移动终端的直播功能来拍摄直播视频供其他用户观看，从而丰富了人们的娱乐生活。

[0003] 相关技术中，受限于移动终端的运算能力，用户使用移动终端直播的内容一般局限于移动终端拍摄的图像，表现力较差。

### 发明内容

[0004] 本申请实施例提供了一种直播方法、装置、计算机设备和存储介质，可以结合移动终端较强的移动能力和个人计算机较强的处理能力来进行直播，从而可以使直播内容更加丰富，提升直播表现力。所述技术方案如下：

[0005] 第一方面，提供了一种直播方法，应用于个人计算机，所述方法包括：

[0006] 通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频；

[0007] 对所述第一直播视频进行目标检测与识别；

[0008] 若在所述第一直播视频中识别到指定对象，则获取所述指定对象对应的虚拟对象；

[0009] 将所述虚拟对象叠加至所述第一直播视频，得到第二直播视频；

[0010] 将所述第二直播视频发送至直播服务器。

[0011] 在本申请中，通过移动终端进行直播视频的拍摄，更便于移动拍摄和移动直播。个人计算机可以通过短距离通信技术获取移动终端拍摄的直播视频，利用个人计算机存储量大和运算速度快的优势，对获取到的直播视频进行处理，即在直播视频中叠加与该直播视频中的指定对象对应的虚拟对象，以使直播视频的内容更加丰富，增强直播的表现力。

[0012] 可选地，所述短距离通信技术包括蓝牙技术、紫蜂技术、无线保真技术、串行总线技术中的至少一种。

[0013] 可选地，所述指定对象为指定场景或指定物体，所述虚拟对象为增强现实模型。

[0014] 可选地，所述对所述第一直播视频进行目标检测与识别，包括：

[0015] 将所述第一直播视频中的每帧视频图像输入目标识别模型，由所述目标识别模型输出所述每帧视频图像中的每个检测框的位置和所述每个检测框内的指定对象的类别；

[0016] 将所述每个检测框的位置确定为所述每个检测框内的指定对象的位置。

[0017] 可选地，所述若在所述第一直播视频中识别到指定对象，则获取所述指定对象对应的虚拟对象，包括：

[0018] 若在所述第一直播视频中识别到指定对象，则根据所述指定对象的类别、所述指

定对象在所述第一直播视频中的位置中的至少一个,获取所述指定对象对应的虚拟对象。

[0019] 可选地,所述将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频,包括:

[0020] 根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频。

[0021] 可选地,所述根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频,包括:

[0022] 根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象与所述第一直播视频叠加显示;

[0023] 若检测到针对所显示的所述虚拟对象的调整指令,则根据所述调整指令调整所述虚拟对象在与所述第一直播视频叠加时的尺寸、位置中的至少一个;

[0024] 若检测到确认指令,则将叠加有所述虚拟对象的第一直播视频作为第二直播视频。

[0025] 第二方面,提供了一种直播装置,应用于个人计算机,所述装置包括:

[0026] 第一获取模块,用于通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频;

[0027] 检测识别模块,用于对所述第一直播视频进行目标检测与识别;

[0028] 第二获取模块,用于若在所述第一直播视频中识别到指定对象,则获取所述指定对象对应的虚拟对象;

[0029] 叠加模块,用于将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频;

[0030] 发送模块,用于将所述第二直播视频发送至直播服务器。

[0031] 可选地,所述短距离通信技术包括蓝牙技术、紫蜂技术、无线保真技术、串行总线技术中的至少一种。

[0032] 可选地,所述指定对象为指定场景或指定物体,所述虚拟对象为增强现实模型。

[0033] 可选地,所述检测识别模块用于:将所述第一直播视频中的每帧视频图像输入目标识别模型,由所述目标识别模型输出所述每帧视频图像中的每个检测框的位置和所述每个检测框内的指定对象的类别;将所述每个检测框的位置确定为所述每个检测框内的指定对象的位置。

[0034] 可选地,该装置还包括:神经网络训练模块,所述神经网络训练模块用于:获取多个训练样本,所述多个训练样本中的每个训练样本包括样本图像和样本标记,所述样本图像中包含指定对象,所述样本标记为样本图像中包含的指定对象的类别;使用所述多个训练样本对神经网络模型进行训练,得到所述目标识别模型。

[0035] 可选地,所述第二获取模块用于:若在所述第一直播视频中识别到指定对象,则根据所述指定对象的类别、所述指定对象在所述第一直播视频中的位置中的至少一个,获取所述指定对象对应的虚拟对象。

[0036] 可选地,所述叠加模块用于:根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象叠加至所述第一直播视频,得到第二直播视频。

[0037] 可选地,所述叠加模块用于:根据所述指定对象在所述第一直播视频中的位置,将所述虚拟对象与所述第一直播视频叠加显示;若检测到针对所显示的所述虚拟对象的调整

指令,则根据所述调整指令调整所述虚拟对象在与所述第一直播视频叠加时的尺寸、位置中的至少一个;若检测到确认指令,则将叠加有所述虚拟对象的第一直播视频作为第二直播视频。

[0038] 第三方面,提供了一种计算机设备,所述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程序,所述计算机程序被所述处理器执行时实现上述的第一方面所述的直播方法。

[0039] 第四方面,提供了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时实现上述的第一方面所述的直播方法。

[0040] 第五方面,提供了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述的第一方面所述的直播方法的步骤。

[0041] 可以理解的是,上述第二方面、第三方面、第四方面、第五方面的有益效果可以参见上述第一方面中的相关描述,在此不再赘述。

## 附图说明

[0042] 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的前提下,还可以根据这些附图获得其他的附图。

[0043] 图1是本申请实施例提供的一种直播系统的示意图;

[0044] 图2是本申请实施例提供的一种直播方法的流程图;

[0045] 图3是本申请实施例提供的一种直播装置的结构示意图;

[0046] 图4是本申请实施例提供的一种计算机设备的结构示意图。

[0047] 其中,各附图标号所代表的含义分别为:

[0048] 12、移动终端;

[0049] 14、个人计算机;

[0050] 16、直播服务器;

[0051] 301、第一获取模块;

[0052] 302、检测识别模块;

[0053] 303、第二获取模块;

[0054] 304、叠加模块;

[0055] 305、发送模块;

[0056] 40、计算机设备;

[0057] 41、存储器;

[0058] 42、计算机程序;

[0059] 43、处理器。

## 具体实施方式

[0060] 为使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0061] 应当理解的是,本申请提及的“多个”是指两个或两个以上。在本申请的描述中,除非另有说明,“/”表示或的意思,比如,A/B可以表示A或B;本文中的“和/或”仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,比如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,为了便于清楚描述本申请的技术方案,采用了“第一”、“第二”等字样对功能和作用基本相同的相同项或相似项进行区分。本领域技术人员可以理解“第一”、“第二”等字样并不对数量和执行次序进行限定,并且“第一”、“第二”等字样也并不限定一定不同。

[0062] 在对本申请实施例进行详细地解释说明之前,先对本申请实施例的应用场景予以说明。

[0063] 相关技术中,使用移动终端(如手机、平板电脑等)进行直播时,通常由移动终端拍摄直播视频,并将直播视频上传到直播服务器,供其他用户观看。在该过程中,受限于移动终端的运算能力,移动终端无法对直播视频进行进一步的处理,导致直播的内容仅局限于移动终端拍摄的图像,表现力较差。

[0064] 为此,本申请实施例提供了一种直播方法,可以结合移动终端较强的移动能力和个人计算机(Personal Computer,PC)较强的处理能力来进行直播,从而可以使直播内容更加丰富,提升直播表现力。

[0065] 下面对本申请实施例涉及的系统架构予以说明。

[0066] 图1是本申请实施例提供的一种直播系统的示意图。参见图1,该直播系统包括移动终端12、个人计算机14和直播服务器16。

[0067] 移动终端12可以是移动能力较强的终端,如移动终端12可以是手机或平板电脑等可以自由移动的终端。移动终端12中可以设置有摄像头,用于拍摄直播视频。

[0068] 个人计算机14可以是处理能力较强的设备,如个人计算机14可以是台式机、一体机或笔记本电脑等。如在图1所示的实施例中,个人计算机14是包括主机、显示器和键盘的台式机。

[0069] 个人计算机14相较于移动终端12具有较强的处理能力;移动终端12相较于个人计算机14具有较强的移动能力。移动终端12与个人计算机14之间可以通过短距离通信技术进行通信连接,使个人计算机14可以通过此通信连接获取移动终端12拍摄的直播视频。

[0070] 这里的短距离通信技术包括蓝牙(Bluetooth)技术、紫蜂(ZigBee)技术、无线保真(Wireless-Fidelity,wifi)技术、串行总线(Universal Serial Bus,USB)技术中的至少一种。其中,蓝牙技术、紫蜂技术和无线保真技术属于短距离无线通信技术。串行总线技术属于短距离有线通信技术。移动终端12与个人计算机14通过短距离通信技术进行通信连接,以使移动终端12与个人计算机14之间可以进行信息传输。这里的信息不仅可以包括视频图像信息,还可以包括控制信息。

[0071] 直播服务器16也叫直播伺服器。个人计算机14与直播服务器16之间可以通过有线网络或无线网络进行通信连接,个人计算机14可以通过此通信连接向直播服务器16发送直播视频。直播服务器16接收到该直播视频后,可以将该直播视频发送给其他终端,以供其他终端的用户观看。

[0072] 下面对本申请实施例提供的直播方法进行详细地解释说明。

[0073] 图2是本申请实施例提供的一种直播方法的流程图。参见图2,该方法包括以下步

骤:

[0074] S100,个人计算机与移动终端通过短距离通信技术建立通信连接。

[0075] 示例地,个人计算机和移动终端均可开启蓝牙功能,从而使个人计算机和移动终端通过蓝牙技术建立通信连接。或者,个人计算机和移动终端均开启紫蜂功能,从而使个人计算机和移动终端通过紫蜂技术建立通信连接。或者,个人计算机和移动终端均开启无线保真功能,从而使个人计算机和移动终端通过无线保真技术建立通信连接。或者,个人计算机和移动终端之间通过串行总线建立通信连接。当然,个人计算机和移动终端也可以通过其他短距离通信技术建立通信连接,本申请实施例对此不作限定。

[0076] S200,移动终端拍摄直播视频。

[0077] 移动终端在拍摄直播视频之前,可以先开启移动终端的摄像功能。一种可能的方式中,移动终端可以在检测到摄像开启指令时开启移动终端的摄像功能。或者,移动终端可以在接收到个人计算机发送的摄像开启消息时开启移动终端的摄像功能。一般地,个人计算机可以通过WDM(Windows Driver Model,视窗驱动程序模块)向移动终端发送摄像开启消息。

[0078] 摄像开启指令用于指示开启移动终端的摄像功能。该摄像开启指令可以是由用户在移动终端上触发的,用户可以通过点击操作、滑动操作、语音操作、手势操作、体感操作等操作进行触发。

[0079] 摄像开启消息用于指示开启移动终端的摄像功能。摄像开启消息可以是个人计算机在检测到摄像开启指令时向移动终端发送的。该摄像开启指令可以是由用户在个人计算机上触发的,用户可以通过点击操作、滑动操作、语音操作、手势操作、体感操作等操作进行触发。

[0080] S300,个人计算机通过短距离通信技术获取移动终端拍摄的直播视频作为第一直播视频。

[0081] 在本申请实施例中,为便于描述和区分,将由移动终端所拍摄的,且传输至个人计算机的直播视频称为第一直播视频。第一直播视频中的“第一”是用于与下述的第二直播视频中的“第二”进行区别。

[0082] 具体地,移动终端拍摄直播视频后,可以直接将所拍摄的直播视频通过短距离通信技术传输至个人计算机。这种情况下,个人计算机也就通过短距离通信技术获取到了移动终端拍摄的直播视频,可以将获取到的直播视频称为第一直播视频。

[0083] 可选地,如果移动终端的操作系统为Android系统,则移动终端可以通过Android调试桥(adb)将直播视频发送给个人计算机。如果移动终端的操作系统为IOS系统,则可以通过Airplay无线技术将直播视频发送给个人计算机。

[0084] 值得注意的是,本申请实施例中,个人计算机通过短距离通信技术获取到了移动终端拍摄的直播视频,相当于个人计算机将移动终端的摄像头作为了自己的一个虚拟摄像头来拍摄直播视频,此时个人计算机可以以使用自身的普通摄像头的方式来使用移动终端的摄像头,也即,可以指示移动终端对直播视频的拍摄。

[0085] S400,个人计算机对第一直播视频进行目标检测与识别。

[0086] 目标检测包括检测指定对象在第一直播视频中的位置。目标识别包括识别第一直播视频中指定对象的类型。

[0087] 具体地,步骤S400的操作可以为:个人计算机将第一直播视频中的每帧视频图像输入目标识别模型,由目标识别模型输出每帧视频图像中的每个检测框的位置和每个检测框内的指定对象的类别;将每个检测框的位置确定为这个检测框内的指定对象的位置。

[0088] 目标识别模型可以是预先训练好的能够确定视频图像中出现的指定对象所在的检测框,且能够确定检测到的指定对象的类别的模型。也即,目标识别模型用于对视频图像进行目标检测与识别。

[0089] 检测框用于指示视频图像中存在指定对象的区域,该指定对象完全处于该检测框内。因而可以将某个检测框的位置确定为这个检测框内的指定对象的位置。一般的,检测框可以呈矩形。

[0090] 指定对象可以是预先指定的一些对象,如指定对象可以是指定场景或指定物体。这里的指定场景的类别可以是天空或草原,也可以是白天或夜晚。这里的指定物体的类别可以是人、电线杆或汽车等。

[0091] 进一步地,在个人计算机将第一直播视频中的每帧视频图像输入目标识别模型之前,个人计算机还可以先训练得到目标识别模型。具体地,个人计算机可以获取多个训练样本,使用该多个训练样本对神经网络模型进行训练,得到目标识别模型。

[0092] 该多个训练样本可以是预先设置的。该多个训练样本中的每个训练样本包括样本图像和样本标记,样本图像中包含指定对象,样本标记为样本图像中包含的指定对象的类别。也即,该多个训练样本中的每个训练样本中的输入数据为包含有指定对象的样本图像、样本标记为指定对象的类别。

[0093] 该神经网络模型可以包括多个网络层,该多个网络层中包括输入层、多个隐含层和输出层。输入层负责接收输入数据;输出层负责输出处理后的数据;多个隐含层位于输入层与输出层之间,负责处理数据,多个隐含层对于外部是不可见的。比如,该神经网络模型可以为深度神经网络等,且可以是深度神经网络中的卷积神经网络等。

[0094] 其中,个人计算机使用多个训练样本对神经网络模型进行训练时,对于该多个训练样本中的每个训练样本,可以将这个训练样本中的输入数据输入神经网络模型,获得输出数据;通过损失函数确定该输出数据与这个训练样本中的样本标记之间的损失值;根据该损失值调整该神经网络模型中的参数。在使用该多个训练样本中的每个训练样本对该神经网络模型中的参数进行调整后,参数调整完成的该神经网络模型即为目标识别模型。

[0095] 其中,个人计算机根据该损失值调整该神经网络模型中的参数的操作可以参考相关技术,本申请实施例对此不进行详细阐述。

[0096] 比如,个人计算机可以通过公式 $\hat{w} = w - \alpha * dw$ 来对该神经网络模型中的任意一个参数进行调整。其中, $\hat{w}$ 是调整后的参数。 $w$ 是调整前的参数。 $\alpha$ 是学习率, $\alpha$ 可以预先设置,如 $\alpha$ 可以为0.001、0.000001等,本申请实施例对此不作唯一限定。 $dw$ 是该损失函数关于 $w$ 的偏导数,可以根据该损失值求得。

[0097] 值得注意的是,个人计算机在使用目标识别模型的过程中,还可以对目标识别模型进行在线更新。具体地,个人计算机将第一直播视频中的某帧视频图像输入目标识别模型后,若目标识别模型输出这帧视频图像中的检测框的位置以及该检测框内的指定对象的类别,则说明这帧视频图像中包含有指定对象,此时可以将这帧视频图像作为一个训练样本中的样本图像,将这帧视频图像中包含的指定对象的类别作为这个训练样本中的样本标

记,然后使用这个训练样本对目标识别模型进行训练,实现目标识别模型的在线更新。

[0098] 在本申请实施例中,目标识别模型是一个深度学习的神经网络模型,利用深度学习的训练样本越多,识别结果越精准的特点,个人计算机可以根据在同一场景下进行直播时拍摄到的直播视频的视频图像来持续在线更新目标识别模型,如此在这一场景下的直播体验会持续提升,有利于长期直播。

[0099] 在本申请实施例中,个人计算机可以基于机器学习框架训练神经网络模型以及使用训练得到的目标识别模型。这里的机器学习框架例如可以是TensorFlow、Caffe或Apache Singa等,本申请实施例对此不作限定。

[0100] S500,个人计算机若在第一直播视频中识别到指定对象,则获取指定对象对应的虚拟对象。

[0101] 虚拟对象可以是增强现实模型,如可以是图片、视频、三维模型等,当然,虚拟对象也可以是其他类型的虚拟对象,本申请实施例对此不作限定。不同的指定对象可以对应不同的虚拟对象。个人计算机内可以预先存储有多个虚拟对象,该多个虚拟对象均是可以与视频图像叠加显示的虚拟对象。

[0102] 具体地,步骤S500的操作可以为:个人计算机若在第一直播视频中识别到指定对象,则根据指定对象的类别、指定对象在第一直播视频中的位置中的至少一个,获取指定对象对应的虚拟对象。

[0103] 作为一种示例,个人计算机中可以预先存储有指定对象类别与虚拟对象之间的对应关系,个人计算机可以根据在第一直播视频中识别到的指定对象的类别,从此对应关系中获取对应的虚拟对象。

[0104] 例如,当指定对象是一个指定场景,如天空时,该指定对象对应的虚拟对象可以是呈飞机或热气球形态的增强现实模型。当指定对象是一个指定物体,如人时,该指定对象对应的虚拟对象可以是呈羽翼形态的增强现实模型。

[0105] 作为另一种示例,个人计算机中可以预先存储有指定对象位置与虚拟对象之间的对应关系,个人计算机可以根据在第一直播视频中识别到的指定对象在第一直播视频中的位置,从此对应关系中获取对应的虚拟对象。

[0106] 例如,当指定对象在第一直播视频的中间位置,则虚拟对象可以是呈皇冠形态的增强现实模型。若指定对象在第一直播视频的边缘位置,则虚拟对象可以是呈花瓣形态的增强现实模型。

[0107] 作为又一种示例,个人计算机中可以预先存储有指定对象类别、指定对象位置与虚拟对象之间的对应关系,个人计算机可以根据在第一直播视频中识别到的指定对象的类别和该指定对象在第一直播视频中的位置,从此对应关系中获取对应的虚拟对象。

[0108] 例如,当指定对象是一个指定物体,如汽车时,若汽车在第一直播视频的中间位置,则该指定对象对应的虚拟对象可以是呈白色羽翼形态的增强现实模型。若汽车在第一直播视频的边缘位置,则该指定对象对应的虚拟对象可以是呈金色羽翼形态的增强现实模型。

[0109] S600,个人计算机将虚拟对象叠加至第一直播视频,得到第二直播视频。

[0110] 将虚拟对象叠加至第一直播视频,是指在第一直播视频上覆盖虚拟对象。当虚拟对象不透明时,第一直播视频被虚拟对象所覆盖的部分在第二直播视频上无法显现。当虚

拟对象具有一定的透明度时,第一直播视频被虚拟对象所覆盖的部分在第二直播视频上与该虚拟对象一起显现,且第一直播视频被虚拟对象所覆盖的部分与该虚拟对象的颜色深浅程度由虚拟对象的透明度决定。

[0111] 将虚拟对象叠加至第一直播视频,即可完成对第一直播视频的增强现实处理。换句话说,第一直播视频是个人计算机所获取但未进行处理的,由移动终端拍摄的直播视频。第二直播视频是个人计算机对第一直播视频进行增强现实处理后所得的直播视频。

[0112] 具体地,步骤S600的操作可以为:个人计算机根据指定对象在第一直播视频中的位置,将虚拟对象叠加至第一直播视频,得到第二直播视频。

[0113] 个人计算机在将虚拟对象叠加至第一直播视频时,需要根据指定对象在第一直播视频中的位置得到虚拟对象的位置。

[0114] 一种可能的方式中,个人计算机可以根据指定对象在第一直播视频中的位置,自动将虚拟对象叠加至第一直播视频,得到第二直播视频。

[0115] 这种方式中,虚拟对象在第一直播视频中的叠加位置可以根据指定对象在第一直播视频中的位置自动确定。比如,虚拟对象在第一直播视频中的叠加位置可以是指定对象在第一直播视频中的位置,或者,虚拟对象在第一直播视频中的叠加位置可以是指定对象在第一直播视频中的位置的周侧位置。

[0116] 另一种可能的方式中,个人计算机可以根据指定对象在第一直播视频中的位置,将虚拟对象与第一直播视频叠加显示;若检测到针对所显示的虚拟对象的调整指令,则根据调整指令调整虚拟对象在与第一直播视频叠加时的尺寸、位置中的至少一个;若检测到确认指令,则将叠加有虚拟对象的第一直播视频作为第二直播视频。

[0117] 在本申请实施例中,个人计算机不仅可以包括用于发送、接收和处理信息的主机,还可以包括与主机连接、用于显示图像的显示器,以及与主机连接、用于输入指令的键盘/鼠标。

[0118] 个人计算机在第一直播视频中检测到指定对象,并获取指定对象对应的虚拟对象后,个人计算机可以根据指定对象在第一直播视频中的位置,将虚拟对象与第一直播视频叠加,并将叠加后的画面通过与主机连接的显示器显示。比如,主机可以通过HDMI (High Definition Multimedia Interface,高清晰度多媒体接口)将叠加后的画面传输至显示器进行显示。

[0119] 可选地,个人计算机可以通过键盘或鼠标等输入设备检测调整指令,该调整指令用于对所显示的虚拟对象的尺寸、位置等进行调整。或者,该调整指令也可以是由用户在个人计算机上触发的,用户可以通过点击操作、滑动操作、语音操作、手势操作、体感操作等操作进行触发。

[0120] 例如,第一直播视频中的指定对象为人,对应该指定对象的虚拟对象为呈羽翼形态的增强现实模型。此时,虚拟对象与第一直播视频叠加显示时,个人计算机通过输入设备获取调整指令后即可调整该虚拟对象(呈羽翼形态的增强现实模型)的尺寸或/和位置。如可以放大羽翼、缩小羽翼,或/和,调整羽翼相对于人的位置等。

[0121] 个人计算机对所显示的虚拟对象调整后,若检测到确认指令,则确定已完成对所显示的虚拟对象的调整,也即,已经得到满足用户需求的直播视频,因而此时可以将当前所显示的叠加有虚拟对象的第一直播视频作为第二直播视频。该确认指令也可以是由用户在

个人计算机上触发的,用户可以通过点击操作、滑动操作、语音操作、手势操作、体感操作等操作进行触发。

[0122] S700,个人计算机将第二直播视频发送至直播服务器。

[0123] 个人计算机在得到第二直播视频后,将第二直播视频发送至直播服务器。直播服务器接收到第二直播视频后,可以将第二直播视频发送给其他终端,以供其他终端的用户观看。

[0124] 在本申请实施例中,通过移动终端进行直播视频的拍摄,更便于移动拍摄和移动直播,且可以利用移动终端更新换代频率高、镜头工艺和拍摄算法发展迅速的优势,提高直播视频的拍摄质量。个人计算机可以通过短距离通信技术获取移动终端拍摄的直播视频,利用个人计算机存储量大和运算速度快的优势,对获取到的直播视频进行处理,即在直播视频中叠加与该直播视频中的指定对象对应的虚拟对象,以增强直播的表现力。

[0125] 并且,本申请实施例提供的直播方法通过移动终端及个人计算机即可实现,软件升级和功能扩展都很方便,可以兼容更多应用场景,比如可以在常规直播场景下接入视频会议系统等。

[0126] 另外,本申请实施例中,个人计算机可以通过目标识别模型进行指定对象的检测和识别,并自动调用指定对象对应的虚拟对象来进行叠加,如此可以减少直播过程对工作人员的依赖程度,降低工作人员的负担。

[0127] 图3是本申请实施例提供的一种直播装置的结构示意图。该装置应用于个人计算机,该个人计算机可以是上文图1实施例所述的个人计算机14。参见图3,该装置可以包括:第一获取模块301、检测识别模块302、第二获取模块303、叠加模块304和发送模块305。

[0128] 第一获取模块301,用于通过短距离通信技术获取移动终端12拍摄的直播视频作为第一直播视频;

[0129] 检测识别模块302,用于对第一直播视频进行目标检测与识别;

[0130] 第二获取模块303,用于若在第一直播视频中识别到指定对象,则获取指定对象对应的虚拟对象;

[0131] 叠加模块304,用于将虚拟对象叠加至第一直播视频,得到第二直播视频;

[0132] 发送模块305,用于将第二直播视频发送至直播服务器。

[0133] 可选地,短距离通信技术包括蓝牙技术、紫蜂技术、无线保真技术、串行总线技术中的至少一种。

[0134] 可选地,指定对象为指定场景或指定物体,虚拟对象为增强现实模型。

[0135] 可选地,检测识别模块302用于:

[0136] 将第一直播视频中的每帧视频图像输入目标识别模型,由目标识别模型输出每帧视频图像中的每个检测框的位置和每个检测框内的指定对象的类别;将每个检测框的位置确定为每个检测框内的指定对象的位置。

[0137] 可选地,该装置还包括:神经网络训练模块,神经网络训练模块用于:

[0138] 获取多个训练样本,多个训练样本中的每个训练样本包括样本图像和样本标记,样本图像中包含指定对象,样本标记为样本图像中包含的指定对象的类别;使用多个训练样本对神经网络模型进行训练,得到目标识别模型。

[0139] 可选地,第二获取模块303用于:

[0140] 若在第一直播视频中识别到指定对象,则根据指定对象的类别、指定对象在第一直播视频中的位置中的至少一个,获取指定对象对应的虚拟对象。

[0141] 可选地,叠加模块304用于:

[0142] 根据指定对象在第一直播视频中的位置,将虚拟对象叠加至第一直播视频,得到第二直播视频。

[0143] 可选地,叠加模块304用于:

[0144] 根据指定对象在第一直播视频中的位置,将虚拟对象与第一直播视频叠加显示;若检测到针对所显示的虚拟对象的调整指令,则根据调整指令调整虚拟对象在与第一直播视频叠加时的尺寸、位置中的至少一个;若检测到确认指令,则将叠加有虚拟对象的第一直播视频作为第二直播视频。

[0145] 在本申请实施例中,通过移动终端进行直播视频的拍摄,更便于移动拍摄和移动直播,且可以利用移动终端更新换代频率高、镜头工艺和拍摄算法发展迅速的优势,提高直播视频的拍摄质量。个人计算机可以通过短距离通信技术获取移动终端拍摄的直播视频,利用个人计算机存储量大和运算速度快的优势,对获取到的直播视频进行处理,即在直播视频中叠加与该直播视频中的指定对象对应的虚拟对象,以增强直播的表现力。

[0146] 需要说明的是:上述实施例提供的直播装置在直播时,仅以上述各功能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模块完成,即将装置的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。

[0147] 上述实施例中的各功能单元、模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请实施例的保护范围。

[0148] 上述实施例提供的直播装置与直播方法实施例属于同一构思,上述实施例中单元、模块的具体工作过程及带来的技术效果,可参见方法实施例部分,此处不再赘述。

[0149] 图4为本申请实施例提供的一种计算机设备40的结构示意图。如图4所示,计算机设备40包括:处理器43、存储器41以及存储在存储器41中并可在处理器43上运行的计算机程序42,处理器43执行计算机程序42时实现上述实施例中的直播方法中的步骤。

[0150] 计算机设备40可以是一个通用计算机设备或一个专用计算机设备。在具体实现中,计算机设备40可以是台式机、一体机或笔记本电脑,本申请实施例不限定计算机设备40的类型。本领域技术人员可以理解,图仅仅是计算机设备40的举例,并不构成对计算机设备40的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,比如还可以包括输入输出设备、网络接入设备等。

[0151] 处理器43可以是中央处理单元(Central Processing Unit,CPU),处理器43还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者也可以是任何常规的处理器。

[0152] 存储器41在一些实施例中可以是计算机设备40的内部存储单元,比如计算机设备40的硬盘或内存。存储器41在另一些实施例中也可以是计算机设备40的外部存储设备,比

如计算机设备40上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,存储器41还可以既包括计算机设备40的内部存储单元也包括外部存储设备。存储器41用于存储操作系统、应用程序、引导装载程序(BootLoader)、数据以及其他程序等,比如计算机程序的程序代码等。存储器41还可以用于暂时地存储已经输出或者将要输出的数据。

[0153] 本申请实施例还提供了一种计算机设备,该计算机设备包括:至少一个处理器、存储器以及存储在该存储器中并可在该至少一个处理器上运行的计算机程序,该处理器执行该计算机程序时实现上述任意各个方法实施例中的步骤。

[0154] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质存储有计算机程序,该计算机程序被处理器执行时实现可实现上述各个方法实施例中的步骤。

[0155] 本申请实施例提供了一种计算机程序产品,当其在计算机上运行时,使得计算机执行上述各个方法实施例中的步骤。

[0156] 集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读存储介质中。基于这样的理解,本申请实现上述方法实施例中的全部或部分流程,可以通过计算机程序来指令相关的硬件来完成,该计算机程序可存储于一计算机可读存储介质中,该计算机程序在被处理器执行时,可实现上述各个方法实施例的步骤。其中,该计算机程序包括计算机程序代码,该计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。该计算机可读介质至少可以包括:能够将计算机程序代码携带到拍照装置/终端设备的任何实体或装置、记录介质、计算机存储器、ROM(Read-Only Memory,只读存储器)、RAM(Random Access Memory,随机存取存储器)、CD-ROM(Compact Disc Read-Only Memory,只读光盘)、磁带、软盘和光数据存储设备等。本申请提到的计算机可读存储介质可以为非易失性存储介质,换句话说,可以是非瞬时性存储介质。

[0157] 应当理解的是,实现上述实施例的全部或部分步骤可以通过软件、硬件、固件或者其任意结合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。该计算机程序产品包括一个或多个计算机指令。该计算机指令可以存储在上述计算机可读存储介质中。

[0158] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述或记载的部分,可以参见其它实施例的相关描述。

[0159] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。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。专业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申请的范围。

[0160] 在本申请所提供的实施例中,应该理解到,所揭露的装置/计算机设备和方法,可以通过其它的方式实现。例如,以上所描述的装置/计算机设备实施例仅仅是示意性的,例如,模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口,装

置或单元的间接耦合或通讯连接,可以是电性,机械或其它的形式。

[0161] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0162] 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。

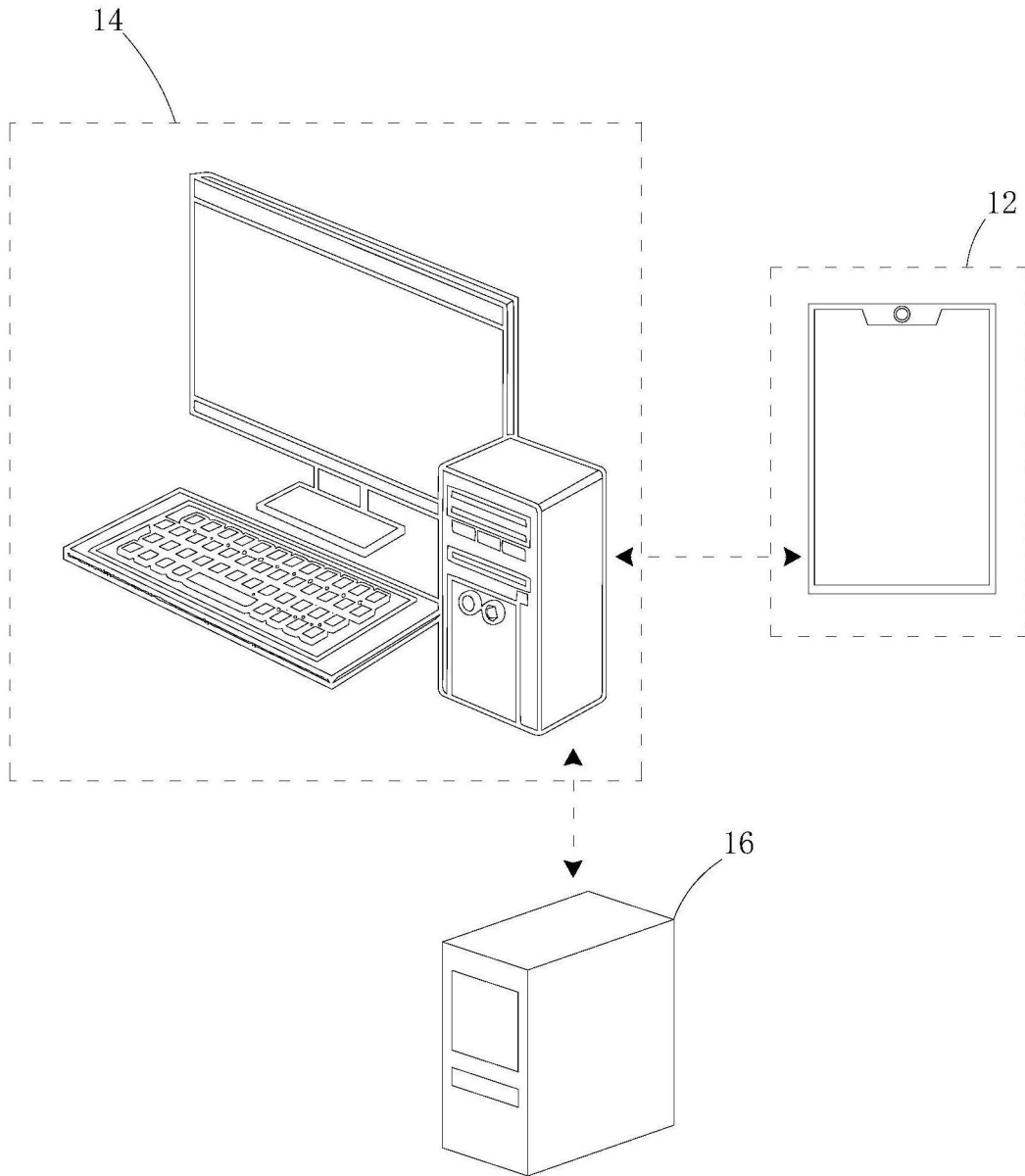


图1

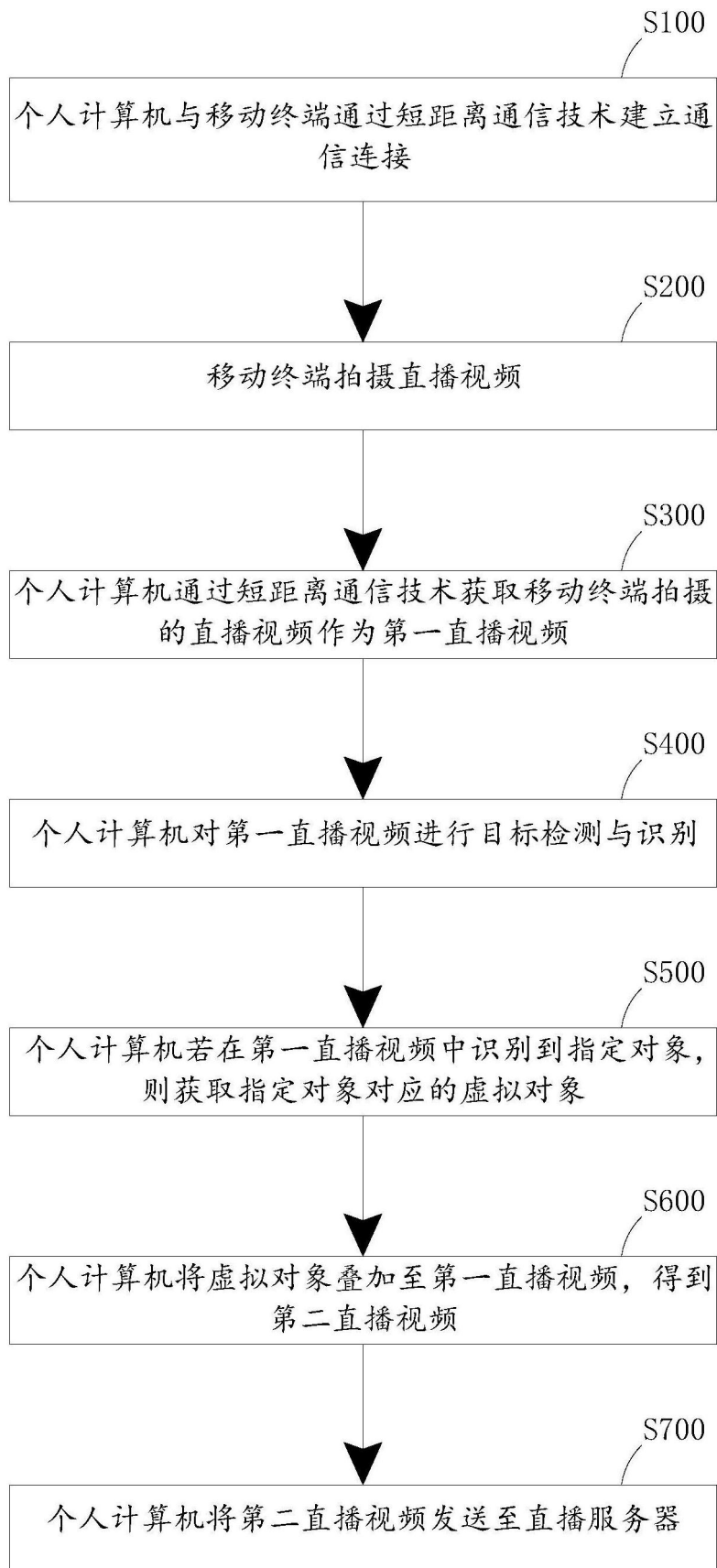


图2

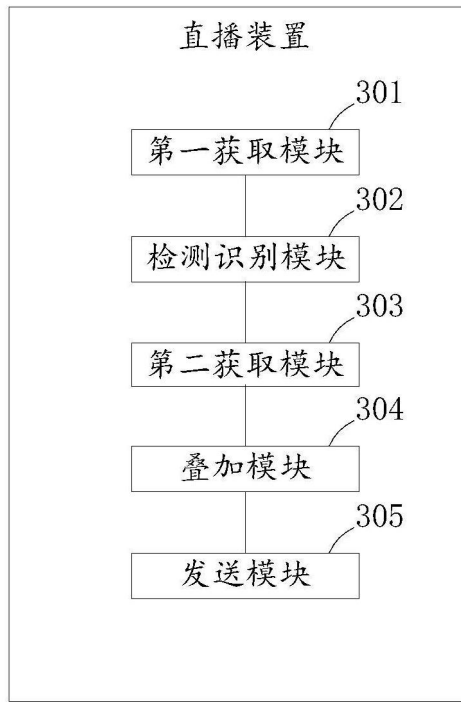


图3

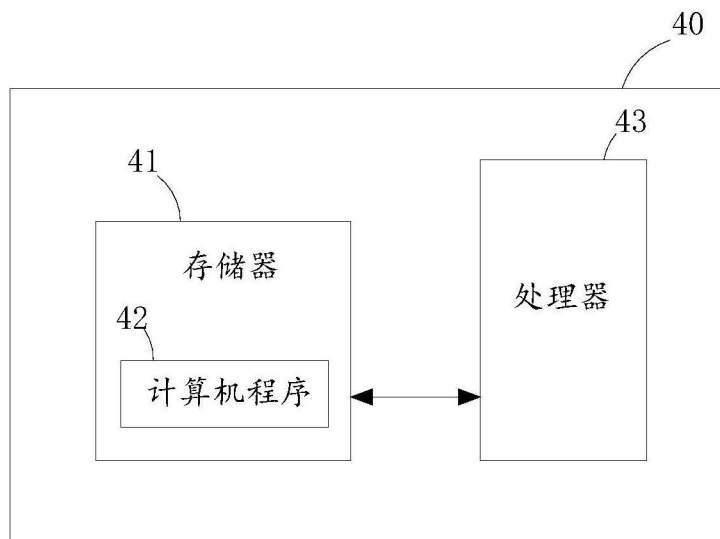


图4