

(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 7 月 14 日 (14.07.2022)



(10) 国际公布号
WO 2022/147703 A1

- (51) 国际专利分类号:
G03B 13/36 (2021.01)
- (21) 国际申请号: PCT/CN2021/070581
- (22) 国际申请日: 2021 年 1 月 7 日 (07.01.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 深圳市大疆创新科技有限公司 (SZ DJI TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。
- (72) 发明人: 滕文猛 (TENG, Wenmeng); 中国广东省深圳市南山区高新区南区粤兴一道9号香

港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 胡涛 (HU, Tao); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 韩守谦 (HAN, Shouqian); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街9号嘉华大厦B座409室, Beijing 100085 (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,

(54) **Title:** FOCUS FOLLOWING METHOD AND APPARATUS, AND PHOTOGRAPHIC DEVICE AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 跟焦方法、装置、拍摄设备及计算机可读存储介质

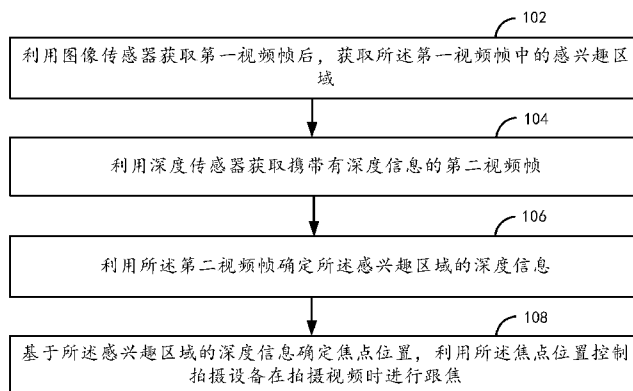


图 1A

- 102 After a first video frame is acquired by using an image sensor, acquire a region of interest from the first video frame
- 104 Acquire, by using a depth sensor, a second video frame carrying depth information
- 106 Determine depth information of the region of interest by using the second video frame
- 108 Determine a focus position on the basis of the depth information of the region of interest, and control, by using the focus position, a photographic device to perform focus following when filming a video

(57) **Abstract:** Provided are a focus following method and apparatus, and a photographic device and a computer-readable storage medium. In the solution of the embodiments of the present application, a first video frame is acquired by using an image sensor, and a region of interest can be automatically acquired from the first video frame; a second video frame carrying depth information is acquired by using a depth sensor; the first video frame and the second video frame include the same picture content, and therefore depth information of the region of interest can be determined by using the depth information of the second video frame; and in this way, the depth information of the region of interest represents depth information of a photographed target, and therefore, a focus position can be found accurately, such that a photographic device can realize automatic focus following when filming a video. By means of the solution, the consistency of video filming can be ensured, without the need for additional operations to be performed by a user, and a focus position can also be accurately found.

(57) **摘要:** 本申请提供一种跟焦方法、装置、拍摄设备及计算机可读存储介质, 本申请实施例方案中, 利用图像传感器获取第一视频帧, 能够自动获取所述第一视频帧中的感兴趣区域; 利用深度传感器获取携带有深度信息的第二视频帧, 由于第一视频帧和第二视频帧包含相同画面内容, 因此能够利用第二视频帧的深度信息来确定所述感兴趣区域的深度信息, 如此, 感兴趣区域的深度信息表征了被摄目标的深度信息, 因此能准确地找到焦点位置, 能够使拍摄设备在拍摄视频时实现自动跟焦, 该方案能够同时保证: 视频拍摄的连贯性、无需用户额外操作, 并且还能准确地找到焦点位置。

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, 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))。

跟焦方法、装置、拍摄设备及计算机可读存储介质

技术领域

5 本申请涉及拍摄技术领域，具体而言，涉及一种跟焦方法、装置、拍摄设备及计算机可读存储介质。

背景技术

10 拍摄设备在拍摄时，需要进行对焦处理，对焦处理的目的是找到准确的焦点位置，使被摄目标在感光器件平面上形成清晰的影像，保证所摄画面上主体影像的清晰。因此，准确地找到焦点位置是决定拍摄质量的关键因素之一。特别是在视频拍摄场景下，用户需要连续地拍摄，如何实时连续地找到焦点位置是亟待解决的技术问题。

发明内容

有鉴于此，本申请提供一种跟焦方法、装置、拍摄设备及计算机可读存储介质，以解决相关技术中拍摄时无法自动跟焦的问题。

第一方面，提供一种跟焦方法，所述方法包括：

15 利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

利用所述第二视频帧确定所述感兴趣区域的深度信息；

20 基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

第二方面，提供一种跟焦装置，所述装置包括处理器、存储器、存储在所述存储器上可被所述处理器执行的计算机程序，所述处理器执行所述计算机程序时实现以下步骤：

利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

25 利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

利用所述第二视频帧确定所述感兴趣区域的深度信息；

基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

30 第三方面，提供一种拍摄设备，包括图像传感器和深度传感器；以及，如第二方面所述的跟焦装置。

第四方面，提供一种计算机可读存储介质，所述可读存储介质上存储有若干计算机指令，所述计算机指令被执行时实现第一方面所述的跟焦方法。

35 应用本申请提供的方案，利用图像传感器获取第一视频帧，能够自动获取所述第一视频帧中的感兴趣区域；利用深度传感器获取携带有深度信息的第二视频帧，由于第一视频帧和第二视频帧包含相同画面内容，因此能够利用第二视频帧的深度信息来

确定所述感兴趣区域的深度信息，如此，感兴趣区域的深度信息表征了被摄目标的深度信息，由于获取到的是整个区域的深度信息而并非单点深度信息，因此能准确地找到焦点位置，能够使拍摄设备在拍摄视频时实现自动跟焦，该方案能够同时保证：视频拍摄的连贯性、无需用户额外操作，并且还能准确地找到焦点位置；由于焦点位置能够准确地找到，因此在视频连续拍摄时可保证跟焦的实时性，不会导致明显的焦点过冲或虚焦的现象。

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 1A 是本申请一个实施例的跟焦方法的示意图。

图 1B 是本申请另一个实施例的跟焦方法的示意图。

图 1C 是本申请一个实施例的图像传感器和深度传感器的示意图。

图 1D 是本申请一个实施例的人脸区域的权重示意图。

图 2 是本申请一个实施例的跟焦装置的示意图。

图 3 是本申请一个实施例的拍摄设备的示意图。

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，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。

拍摄设备的 ISP (Image Signal Processing, 图像信号处理) 系统中通常配置有对焦模块，一些对焦处理方案采用相位检测自动对焦，但此种方案对环境光线要求较高，在夜晚环境下效果较差；另一些对焦处理方案是采用激光方式获取到单点深度信息，基于单点深度信息可能无法获取到用户想要的拍摄主体的深度信息，无法实现准确对焦。因此，实际拍摄时很多时候用户都需要手动对焦。此种方式，在单次的图像拍摄场景下是可接受的，因为在此拍摄场景中，用户不具有连续拍摄的需求，用户操作对焦后再操作拍摄就完成了整个拍摄流程，手动对焦不会对用户造成较多干扰。

然而连续的视频拍摄场景，与单次的图像拍摄场景实际上是有诸多不同的；由于视频拍摄要保证连贯性，让用户中断连贯的拍摄去选取对焦位置，会给用户带来非常差的拍摄体验。特别是在一些专业视频拍摄领域，为了解决保证拍摄连贯性的同时准确对焦，需要有额外的跟焦师负责手动调节焦点，这为用户拍摄高质量的视频带来了较大的成本。

由此可知，在视频拍摄场景下，面临的问题是：一方面需要解放用户双手，无需用户手动操作对焦以保证视频拍摄的连贯性，一方面希望对焦处理只需要一个用户就能够完成以减少拍摄成本，再者，还要同时保证自动跟焦处理的准确性。

基于此，本申请实施例提供的跟焦方案，利用图像传感器获取第一视频帧，能够自动获取所述第一视频帧中的感兴趣区域；利用深度传感器获取携带有深度信息的第

二视频帧，由于第一视频帧和第二视频帧包含相同画面内容，因此能够利用第二视频帧的深度信息来确定所述感兴趣区域的深度信息，如此，感兴趣区域的深度信息表征了被摄目标的深度信息，由于获取到的是整个区域的深度信息而并非单点深度信息，因此能准确地找到焦点位置，能够使拍摄设备在拍摄视频时实现自动跟焦，该方案能够同时保证：视频拍摄的连贯性、无需用户额外操作，并且还能准确地找到焦点位置；由于焦点位置能够准确地找到，因此在视频连续拍摄时可保证跟焦的实时性，不会导致明显的焦点过冲或虚焦的现象。

本实施例方案可应用于拍摄设备中，由拍摄设备内置的处理器运行本实施例的跟焦方案以实现在视频拍摄时的自动跟焦。在一些例子中，拍摄设备还可以搭载在外部设备中，此处的外部设备可以包括可移动平台等设备，该可移动平台可以包括车辆、无人机或可移动机器人等。

结合图 1A 和图 1B 进行说明，图 1A 和图 1B 分别为本实施例示出的跟焦方法的示意图，可包括如下步骤：

在步骤 102 中，利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

在步骤 104 中，利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

在步骤 106 中，利用所述第二视频帧确定所述感兴趣区域的深度信息；

在步骤 108 中，基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

本实施例的跟焦方案中适用于视频拍摄场景，拍摄设备配置有图像传感器，图像传感器用于连续采集图像，本实施例通过图像传感器获取第一视频帧，该第一视频帧可以是图像传感器采集的原始图像，也可以是对图像传感器采集的原始图像经过一些处理后得到的图像。

本实施例方案中，拍摄设备还配置有深度传感器，以在视频拍摄场景中连续地采集携带有深度信息的图像，本实施例通过深度传感器获取第二视频帧，该第二视频帧可以是深度传感器采集的原始图像，也可以是对深度传感器采集的原始图像经过一些处理后得到的图像。

本实施例中，第一视频帧和第二视频帧包含相同画面内容，利用第一视频帧自动识别出感兴趣区域，并利用所述第二视频帧携带的深度信息来确定感兴趣区域的深度信息。

实际应用中可以通过多种方式令第一视频帧和第二视频帧包含相同画面内容。作为例子，可以在所述拍摄设备中，令所述深度传感器与所述图像传感器平行，使得深度传感器与图像传感器分别采集的图像可以对齐，如图 1C 所示，是一个实施例中拍摄设备内所述深度传感器与所述图像传感器的示意图，图 1C 中，深度传感器以 TOF (Time of flight, 飞行时间) 传感器为例，该深度传感器 TOF 通过一支撑器件 TOF holder 架设与图像传感器 Image sensor 两者处于平行状态，处于平行状态是指深度传感器与水平面的夹角 $X \text{ deg}$ ，与图像传感器 Image sensor 与水平面的夹角相等，因此能够使得所述深度传感器采集的视频帧与所述图像传感器采集的视频帧对应同一画面，即第一视频帧和第二视频帧包含相同画面内容。

当然，深度传感器与图像传感器也可以不具有上述平行状态的限制；在另一些例子中，深度传感器与图像传感器可以处于非平行状态，即可以根据需要配置深度传感器与图像传感器处于其他位置关系，本实施例方案还可以对图像传感器采集的图像与深度传感器采集的图像进行图像对齐处理，以获得所述第一视频帧和第二视频帧。

5 作为例子，深度传感器与图像传感器的安装位置可能没有处于平行状态，图像传感器与深度传感器具有角度偏差，从而图像传感器采集的图像与深度传感器采集的图像具有角度偏差，基于此，本实施例还可以进行图像对齐处理，作为例子，可以包括：将所述图像传感器采集的图像与所述深度传感器采集的图像进行轮廓对齐，例如，根据深度传感器与图像传感器的角度差异，将所述图像传感器采集的图像与所述深度传感器采集的图像的轮廓对齐。

10 在另一些例子中，考虑到深度传感器与图像传感器两者无法完全重合，深度传感器与图像传感器分别对应的视角范围可能无法完全相同，即两者采集的图像中所包含的画面并未完全相同，深度传感器采集的图像所包含的画面，可能存在图像传感器采集的图像所未能包含的画面，图像传感器采集的图像所包含的画面，也可能存在深度传感器采集的图像所未能包含的画面，因此，本实施例将两者采集的图像中不同的画面称之为偏移区域，本实施例通过将所述图像传感器采集的图像与所述深度传感器采集的图像两者的偏移区域进行裁剪，使得裁剪后的第一视频帧和第二视频帧两者包含的画面相同，进而可以便捷地获取到感兴趣区域的深度信息。

15 在一些例子中，所述深度传感器可以包括 3D ToF (TOF, Time-of-Flight) 传感器，通过 3D ToF 传感器采集图像，不会受夜晚等光照较差的影响，使得本实施例的跟焦方案能够适用于夜晚环境；另一方面 3D ToF 传感器能够采集到大范围内的深度信息，从而使本实施例的跟焦方案能够准确地找到焦点位置。

20 本实施例通过对第一视频帧自动识别出感兴趣区域来自动确定拍摄主体，利用所述第二视频帧确定所述感兴趣区域的深度信息，作为例子，该处理过程可以包括：利用所述感兴趣区域在所述第一视频帧中的位置，从所述第二视频帧中确定所述感兴趣区域的深度信息。例如，第一视频帧和第二视频帧包含同一画面，识别出的感兴趣区域，能够确定该感兴趣区域在第一视频帧中的位置，则对应地可以确定该感兴趣区域在第二视频帧中的位置，根据第二视频帧携带的深度信息可以确定出所述感兴趣区域的深度信息。

25 30 在确定了感兴趣区域的深度信息后，可进一步确定焦点位置。感兴趣区域表征了第一视频帧中的拍摄主体所在的区域，从感兴趣区域的深度信息可以确定焦点位置，具体的确定方式实际应用中可以根据需要灵活配置。

35 作为例子，感兴趣区域包含了多个像素点，可以基于所述感兴趣区域的各个目标像素点的深度信息来确定感兴趣区域的平均深度信息，进而利用所述感兴趣区域的平均深度信息来确定焦点位置。

40 可选的，所述感兴趣区域的平均深度信息是基于所述感兴趣区域中各个目标像素点的深度信息以及各个所述目标像素点的预设权重确定的。例如，可以根据感兴趣区域中对应的拍摄主体，预先对感兴趣区域划分为不同的子区域，并对不同的子区域设置权重，在确定焦点位置时，将实际得到的感兴趣区域中各个目标像素点乘以该目标像素点的权重并求平均值，进而得到感兴趣区域的平均深度信息，进一步可确定焦点

位置。

其中，目标像素点可以是感兴趣区域中的所有像素点，也可以是部分像素点，例如，可以通过对感兴趣区域的深度信息进行去噪处理后得到的像素点。实际应用中，感兴趣区域中可能有部分像素点的深度信息不可靠，可以剔除掉部分不可靠的像素点。

5 作为例子，考虑到拍摄主体与拍摄设备的距离应该是在一个合理的范围内，如果第二视频帧中携带的深度信息表征拍摄主体非常近或者非常远，不在合理的范围内，可以认为该采样点的深度信息不可靠，该采样点可能不属于拍摄主体，或者是深度传感器的采集数据有误，基于此，为了防止这些可能错误的像素点的干扰，可以将这些像素点剔除。如何确定这些深度信息可能错误的像素点，在一些例子中，可以预先配置置信度阈值，也即是所述目标像素点是所述感兴趣区域中深度信息大于设定置信度阈值的像素点，深度信息满足置信度阈值的像素点则保留，不满足置信度阈值的像素点则剔除。

15 在一些例子中，所述置信度阈值是基于所述深度传感器的回波强度设定的。实际应用中，不同的深度传感器可能有不同的回波感应能力，正常情况下深度传感器采集到的深度信息也会处于一个合理的范围内，本实施例可以预先基于深度传感器的回波强度来确定置信度阈值，从而使得跟焦处理方案的准确度更为精准。

实际应用中，可以根据需要确定不同拍摄场景下的感兴趣区域。作为例子，若包含有人脸画面，感兴趣区域可以是所述第一视频帧中识别得到的人脸区域；在另一些例子，在未包含有人脸画面时，感兴趣区域还可以是从所述第一视频帧中识别得到的其他动物的脸部区域；在其他拍摄场景中还可以是其他物体的区域；实际实现时，可以利用物体识别算法等方式自动识别出感兴趣区域。在一些例子中，从第一视频帧中可能还会自动识别出多个感兴趣区域，对于识别出的多个感兴趣区域，还可以根据设定的优先级自动保留其中一个优先级最高的，而其他的感兴趣区域剔除，例如，人脸区域的优先级最高，动物脸部区域次之，其他不同类别的物品次之等，实际应用中可以根据需要进行配置，本实施例对此不作限定。

25 在一些例子中，还可能识别到多个人脸，可以根据需要自动选取其中一个人脸所在区域作为感兴趣区域，例如可以基于各个人脸区域的大小确定，或者是基于各个人脸区域的深度信息的高低来确定，例如将人脸所在区域面积较大的作为感兴趣区域，或者是将人脸深度信息较小的作为感兴趣区域等。

30 在另一些例子中，还可以根据用户指定的需要跟焦的人脸区域来确定所述第一视频帧中的感兴趣区域。作为例子，用户可以在拍摄前或拍摄过程中指定要跟焦的人脸，根据用户所指定的人脸，在后续连续拍摄过程中可以在第一视频帧中识别出用户要跟焦的人脸所在的区域为感兴趣区域。

35 为了便于用户指定要跟焦的人脸区域，在一些例子中，所述获取所述第一视频帧中的感兴趣区域，可以包括：在拍摄设备的预览画面中突出显示从图像传感器采集的图像中识别出的至少两个人脸区域后，获取用户指定的需跟焦的人脸区域；利用所述需跟焦的人脸区域，获取所述第一视频帧中的人脸区域；基于此，通过突出显示可以供用户快速便捷地指定其想要跟焦的人脸，用户指定的需跟焦的人脸后，接下来进行视频拍摄时，由于用户指定人脸时的视频帧与接下来的视频拍摄时需要进行跟焦的视频帧已经不同，本实施例可以利用用户指定的需跟焦的人脸的人脸特征，从第一视频

帧中确定感兴趣区域。如此，拍摄时在有多个人脸区域的情况下，本实施例通过用户指定的需跟焦的人脸的人脸特征，能够从拍摄的视频帧中自动识别出来作为感兴趣区域。基于此，在视频拍摄时，用户只需要执行一次指定的操作，后续就可以实现全程的自动连续跟焦。

5 以人脸作为感兴趣区域，在根据感兴趣区域的深度信息确定焦点位置时，可以根据需要为人脸区域中的不同区域设定权重，作为例子，如图 1D 所示，是本实施例中一种人脸区域的权重的示意图，所述人脸区域中，眼部区域像素点的权重、嘴部区域像素点的权重、鼻子区域像素点的权重和边缘区域像素点的权重依次递减。实际应用中可以根据需要设置该权重的其他实现方式，例如根据需要划分为多个其他区域，并
10 对不同区域配置不同的权重，本实施例对此不作限定。

实际应用中，在连续拍摄时，图像传感器会持续采集图像，深度传感器也会持续地采集图像，本实施例的跟焦方案可以持续地执行以找出焦点位置。在一些例子中，可以是按照设定时间间隔持续地执行本实施例的跟焦方案；在另一些例子中，可以以图像传感器采集的图像为基准，基于每一帧第一视频帧执行本实施例的跟焦方案；在
15 另一些例子中，还可以是以深度传感器采集的图像为基准，基于每一帧第二视频帧执行本实施例的跟焦方案。

其中，在一些例子中，深度传感器的帧率（FPS，Frames Per Second，每秒传输帧数）与所述图像传感器的帧率相同，在拍摄时深度传感器与图像传感器同时启动，则深度传感器与图像传感器分别采集的图像在时间上可以自动对齐，即图像传感器和深度
20 传感器分别采集的图像的时间戳相同。

在另一些例子中，所述深度传感器的帧率与所述图像传感器的帧率不同，则所述第二视频帧可以是与第一视频帧的采集时刻时间最接近的视频帧，使得第一视频帧和第二视频帧在时间差异最小，以保证利用第二视频帧获取的深度信息与第一视频帧中拍摄主体相匹配，保证跟焦方案的精确度。

25 其中，在深度传感器的帧率与所述图像传感器的帧率不同的情况下，本实施例的跟焦方案，可以是以图像传感器采集图像为基准，基于第一视频帧获取时间最接近的第二视频帧；也可以是以深度传感器采集图像为基准，基于第二视频帧获取时间最接近的第一视频帧。

在深度传感器的帧率与所述图像传感器的帧率不同的情况下，本实施例中可以设置缓存区以缓存图像，以获取到时间差异最小的第一视频帧和第二视频帧，进而提高跟焦准确度。
30

在一些例子中，以图像传感器采集图像为基准，缓存区中缓存的图像包括第二视频帧，即缓存区中缓存有所述深度传感器采集的携带有时间戳的至少一帧图像；在基于第一视频帧获取时间最接近的第二视频帧时，可以从缓存区中根据图像的时间戳
35 获取时间最接近的第二视频帧。

在另一些例子中，以深度传感器采集图像为基准，缓存区中缓存的图像包括第一视频帧，即缓存区中缓存有所述图像传感器采集的携带有时间戳的至少一帧图像；在基于第二视频帧获取时间最接近的第一视频帧时，可以从缓存区中根据图像的时间戳获取时间最接近的第一视频帧。
40

在深度传感器的帧率与所述图像传感器的帧率不同的情况下，为了保证能够精准

地从缓存区中获取到所需的图像，本实施例还基于所述缓存区中缓存的图像的缓存时长或缓存区中缓存的图像的帧数提供了创新的解决方案。

在一些例子中，所述缓存区中缓存的图像的帧数，是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

5 在一些例子中，所述缓存区中缓存的图像的缓存时长，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的差值。

深度传感器的帧率与所述图像传感器的帧率不同的情况下，作为例子：

(1) 在一些例子中，以图像传感器采集图像为基准，缓存区用于存储深度传感器采集的图像；

10 ①假设深度传感器的帧率大于所述图像传感器的帧率，缓存区中需要存储深度传感器采集的至少一帧图像，以在基于第一视频帧获取第二视频帧时，能够找到与第一视频帧时间最接近的第二视频帧。因为在图像传感器采集到一帧图像的时间间隔内，深度传感器已经能够采集到至少一帧图像，因此需要缓存深度传感器采集到的至少一帧图像。

15 为了保证能够获取到与第一视频帧时间最接近的第二视频帧，作为例子，所述缓存区中缓存的图像的帧数，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的商值。

例如，深度传感器的帧率为 40FPS，图像传感器的帧率为 30FPS；深度传感器每隔 1/40 秒采集一帧图像，图像传感器每隔 1/30 秒采集一帧图像，缓存区中缓存的深度传感器采集的图像的帧数，至少要大于 1/30 与 1/40 的商值，即至少要大于 2 帧，即深度传感器采集 2 帧图像的时间内会产生一帧图像传感器采集的图像，因此至少要缓存 2 帧，这样可保证在基于第一视频帧获取第二视频帧时，能够准确找到与第一视频帧时间最接近的第二视频帧。

25 ②假设深度传感器的帧率小于所述图像传感器的帧率，缓存区中需要将深度传感器采集的至少一帧图像缓存一定时间，以在基于第一视频帧获取第二视频帧时，能够找到与第一视频帧时间最接近的第二视频帧。因为在图像传感器采集到一帧图像的时间间隔内，深度传感器还未能够采集到图像，因此需要将深度传感器采集到的图像缓存一定时间。

30 为了保证能够获取到与第一视频帧时间最接近的第二视频帧，作为例子，所述缓存区中缓存的图像的缓存时长，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的差值。

例如，深度传感器的帧率为 30FPS，图像传感器的帧率为 40FPS；深度传感器每隔 1/30 秒采集一帧图像，图像传感器每隔 1/40 秒采集一帧图像，缓存区中缓存的深度传感器采集的图像的时长，至少要大于 1/30 与 1/40 的差值，即至少要大于 1/120 秒，才可保证在基于第一视频帧获取第二视频帧时，能够找到与第一视频帧时间最接近的第二视频帧。

(2) 在另一些例子中，以深度传感器采集图像为基准，缓存区用于存储图像传感器采集的图像，与前述实施例相同的原理：

40 ①假设深度传感器的帧率大于所述图像传感器的帧率，缓存区中需要存储图像传感器采集的多帧图像，以在基于第二视频帧获取第一视频帧时，能够找到与第二视频

帧时间最接近的第一视频帧。因为在深度传感器采集到一帧图像的时间间隔内，图像传感器还未能够采集到图像，因此需要将图像传感器采集到的图像缓存一定时间。

为了保证能够获取到与第二视频帧的采集时刻时间最接近的第一视频帧，作为例子，所述缓存区中缓存的图像的缓存时长，大于或等于所述深度传感器的帧率倒数与
5 所述图像传感器的帧率倒数的差值。

例如，深度传感器的帧率为 40FPS，图像传感器的帧率为 30FPS；深度传感器每隔 1/40 秒采集一帧图像，图像传感器每隔 1/30 采集一帧图像，缓存区中缓存的图像传感器采集的图像的缓存时长，至少要大于 1/30 与 1/40 的差值，即至少要大于 1/120 秒，
10 才可保证在基于第二视频帧获取第一视频帧时，能够找到与第二视频帧采集时刻的时间最接近的第一视频帧。

②假设深度传感器的帧率小于所述图像传感器的帧率，缓存区中需要存储图像传感器采集的多帧图像，以在基于第二视频帧获取第一视频帧时，能够找到与第二视频帧时间最接近的第一视频帧。因为在深度传感器采集到一帧图像的时间间隔内，图像传感器已经能够采集到至少一帧图像，因此需要缓存图像传感器采集到的至少一帧图
15 像。

为了保证能够获取到与第二视频帧时间最接近的第一视频帧，作为例子，所述缓存区中缓存的图像的帧数，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的商值。

例如，深度传感器的帧率为 30FPS，图像传感器的帧率为 40FPS；深度传感器每隔 1/30 秒采集一帧图像，图像传感器每隔 1/40 秒采集一帧图像，缓存区中缓存的图像传感器采集的图像的帧数，至少要大于 1/30 与 1/40 的商值，即至少要大于 2 帧，即图像传感器在采集两帧图像的
20 时间内会产生一帧深度传感器采集的图像，因此至少要缓存 2 帧，才可保证在基于第二视频帧获取第一视频帧时，能够找到与第二视频帧时间最接近的第一视频帧。

实际应用中，考虑到跟焦方案的执行需要一定时间，其中可能还涉及其他的图像处理，缓存区中缓存的图像还可以包括其他图像，缓存的图像的帧数及时长还可以根据
25 需要结合其他因素进行配置，例如进一步缓存更多帧数的图像，或者是将图像的缓存时间进一步延长，本实施例对此不作限定。

实际应用中，在一例子中，图像传感器的分辨率与深度传感器的分辨率相同，可以便捷地利用第二视频帧的深度信息确定感兴趣区域的深度信息。
30

在另一些例子中，图像传感器的分辨率与深度传感器的分辨率不同，可以对分辨率较低的图像进行插值处理以使两者的分辨率相同，之后再利用第二视频帧的深度信息确定感兴趣区域的深度信息。

实际应用中，通常是所述图像传感器采集的图像的第一分辨率高于所述深度传感器采集的图像的第二分辨率，所述第二视频帧是对所述深度传感器采集的图像进行插
35 值处理得到的与所述第一分辨率相同的视频帧。其中，插值处理的实现过程可以根据需要灵活配置，本实施例对此不作限定。

在一些例子中，本实施例方案还可以同时对所述深度传感器采集的图像进行与图像传感器采集的图像的图像对齐处理时，还可以同时对所述深度传感器采集的图像进
40 行插值处理。图像对齐处理的实施例可见于前述描述，在此不再赘述。

本实施例中，可以基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦；作为例子，可以将深度信息作为物距代入至高斯成像公式 $1/f=1/u+1/v$ ，其中， u 为物距， v 为像距， f 为焦距，进而可以得出像距，根据像距得到焦点位置，计算得的焦点位置，可以基于此控制拍摄设备在拍摄下一视频帧时的焦点位置，从而实现控制拍摄设备在拍摄视频时进行跟焦。

在一些例子中，可以将计算得到的焦点位置直接作为拍摄设备在拍摄下一视频帧时的焦点位置。

在另一些例子中，考虑到利用本实施例的跟焦方法得到的焦点位置的置信度也有可能无法满足拍摄要求，而拍摄设备中配置有其他跟焦模式，基于此，可以综合考虑各种不同跟焦模式获得的结果，将本实施例的跟焦方案与其他跟焦模式的结果进行融合。融合的处理方式可以有多种：

作为例子，一种可选的方式是，在利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，基于所述感兴趣区域的深度信息的整体置信度，将所述焦点位置与基于其他跟焦模式确定的焦点位置进行融合，利用融合结果控制拍摄设备在拍摄视频时进行跟焦。作为例子，感兴趣区域的深度信息的整体置信度，可以是基于感兴趣区域中目标像素点的深度信息的置信度的平均值而确定的。

在另一些例子中，在利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦时，可以基于所述感兴趣区域的深度信息的整体置信度，将所述感兴趣区域的深度信息与基于其他跟焦模式确定的深度信息进行融合，利用融合结果确定出的焦点位置控制拍摄设备在拍摄视频时进行跟焦。

当然，实际应用中，若利用本实施例的跟焦方案获得的感兴趣区域的深度信息的整体置信度较低，也可以终止流程，采用其他方式获得焦点位置。

上述方法实施例可以通过软件实现，也可以通过硬件或者软硬件结合的方式实现。以软件实现为例，作为一个逻辑意义上的装置，是通过其所在处理器将存储器中对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言，如图 2 所示，为实施本实施例图像处理方法的跟焦装置 200 的一种硬件结构图，除了图 2 所示的处理器 201、以及存储器 202 之外，实施例中用于实施本跟焦方法的装置，通常根据该跟焦装置的实际功能，还可以包括其他硬件，对此不再赘述。

本实施例中，所述处理器 201 执行所述计算机程序时实现以下步骤：

利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

利用所述第二视频帧确定所述感兴趣区域的深度信息；

基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

在一些例子中，在所述拍摄设备中，所述深度传感器与所述图像传感器平行。

在一些例子中，所述深度传感器包括：3D ToF 传感器。

在一些例子中，所述利用所述第二视频帧确定所述感兴趣区域的深度信息，包括：

利用所述感兴趣区域在所述第一视频帧中的位置，从所述第二视频帧中确定所述感兴趣区域的深度信息。

在一些例子中,所述焦点位置是利用所述感兴趣区域的平均深度信息确定的,所述感兴趣区域的平均深度信息是基于所述感兴趣区域的各个目标像素点的深度信息确定的。

5 在一些例子中,所述感兴趣区域的平均深度信息是基于所述感兴趣区域中各个目标像素点的深度信息以及各个所述目标像素点的预设权重确定的。

在一些例子中,所述目标像素点是所述感兴趣区域中深度信息大于设定置信度阈值的像素点。

在一些例子中,所述置信度阈值是基于所述深度传感器的回波强度设定的。

在一些例子中,所述感兴趣区域包括:从所述第一视频帧中识别得到的人脸区域。

10 在一些例子中,所述人脸区域中,眼部区域像素点的权重、嘴部区域像素点的权重、鼻子区域像素点的权重和边缘区域像素点的权重依次递减。

在一些例子中,所述感兴趣区域是基于用户指定的需跟焦的人脸区域确定的。

在一些例子中,所述获取所述第一视频帧中的感兴趣区域:

15 在拍摄设备的预览画面中突出显示从图像传感器采集的图像中识别出的至少两个个人脸区域后,获取用户指定的需跟焦的人脸区域;

利用所述需跟焦的人脸区域,获取所述第一视频帧中的人脸区域。

在一些例子中,所述深度传感器的帧率与所述图像传感器的帧率不同;

所述第二视频帧是与第一视频帧的采集时刻时间最接近的视频帧。

20 在一些例子中,所述第二视频帧是从缓存区中获取的,所述缓存区中缓存有所述深度传感器采集的携带有时间戳的至少一帧图像。

在一些例子中,所述第一视频帧是从缓存区中获取的,所述缓存区中缓存有所述图像传感器采集的携带有时间戳的至少一帧图像。

在一些例子中,所述缓存区中缓存的图像的帧数,是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

25 在一些例子中,所述缓存区中缓存的图像的缓存时长,是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

在一些例子中,所述缓存区中缓存的图像的帧数,大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的商值。

30 在一些例子中,所述缓存区中缓存的图像的缓存时长,大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的差值。

在一些例子中,所述图像传感器采集的图像的第一分辨率高于所述深度传感器采集的图像的第二分辨率;

所述第二视频帧是对所述深度传感器采集的图像进行插值处理得到的与所述第一分辨率相同的视频帧。

35 在一些例子中,所述第一视频帧与第二视频帧是经过图像对齐处理后得到的。

在一些例子中,所述图像对齐处理,包括如下任一:将所述图像传感器采集的图像与所述深度传感器采集的图像进行轮廓对齐,或,将所述图像传感器采集的图像与所述深度传感器采集的图像两者的偏移区域进行裁剪处理。

40 在一些例子中,所述偏移区域,是基于所述图像传感器的安装位置和所述深度传感器的安装位置的位置差异确定的。

在一些例子中，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

基于所述感兴趣区域的深度信息的整体置信度，将所述焦点位置与基于其他跟焦模式确定的焦点位置进行融合，利用融合结果控制拍摄设备在拍摄视频时进行跟焦。

5 在一些例子中，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

基于所述感兴趣区域的深度信息的整体置信度，将所述感兴趣区域的深度信息与基于其他跟焦模式确定的深度信息进行融合，利用融合结果确定出的焦点位置控制拍摄设备在拍摄视频时进行跟焦。

10 如图3所示，是申请实施例还提供一种拍摄设备300，包括：图像传感器301和深度传感器302，以及任一实施例所述的跟焦装置200。

本说明书实施例还提供一种计算机可读存储介质，所述可读存储介质上存储有若干计算机指令，所述计算机指令被执行时实任一实施例所述跟焦方法的步骤。

15 本说明书实施例可采用在一个或多个其中包含有程序代码的存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。计算机可用存储介质包括永久性和非永久性、可移动和非可移动媒体，可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括但不限于：相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。

20 对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中所述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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

25 需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

30 以上对本发明实施例所提供的方法和装置进行了详细介绍，本文中应用了具体个例对本发明的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本发明的方法及其核心思想；同时，对于本领域的一般技术人员，依据本发明的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对

本发明的限制。

权利要求书

1、一种跟焦方法，其特征在于，所述方法包括：

利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

5 利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

利用所述第二视频帧确定所述感兴趣区域的深度信息；

基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

10 2、根据权利要求1所述的方法，其特征在于，在所述拍摄设备中，所述深度传感器与所述图像传感器平行。

3、根据权利要求1所述的方法，其特征在于，所述深度传感器包括：3D ToF 传感器。

4、根据权利要求2所述的方法，其特征在于，所述利用所述第二视频帧确定所述感兴趣区域的深度信息，包括：

15 利用所述感兴趣区域在所述第一视频帧中的位置，从所述第二视频帧中确定所述感兴趣区域的深度信息。

5、根据权利要求1所述的方法，其特征在于，所述焦点位置是利用所述感兴趣区域的平均深度信息确定的，所述感兴趣区域的平均深度信息是基于所述感兴趣区域的各个目标像素点的深度信息确定的。

20 6、根据权利要求5所述的方法，其特征在于，所述感兴趣区域的平均深度信息是基于所述感兴趣区域中各个目标像素点的深度信息以及各个所述目标像素点的预设权重确定的。

7、根据权利要求5或6所述的方法，其特征在于，所述目标像素点是所述感兴趣区域中深度信息大于设定置信度阈值的像素点。

25 8、根据权利要求7所述的方法，其特征在于，所述置信度阈值是基于所述深度传感器的回波强度设定的。

9、根据权利要求6所述的方法，其特征在于，所述感兴趣区域包括：从所述第一视频帧中识别得到的人脸区域。

30 10、根据权利要求9所述的方法，其特征在于，所述人脸区域中，眼部区域像素点的权重、嘴部区域像素点的权重、鼻子区域像素点的权重和边缘区域像素点的权重依次递减。

11、根据权利要求 1 所述的方法，其特征在于，所述感兴趣区域是基于用户指定的需跟焦的人脸区域确定的。

12、根据权利要求 11 所述的方法，其特征在于，所述获取所述第一视频帧中的感兴趣区域：

5 在拍摄设备的预览画面中突出显示从图像传感器采集的图像中识别出的至少两个人脸区域后，获取用户指定的需跟焦的人脸区域；

利用所述需跟焦的人脸区域，获取所述第一视频帧中的人脸区域。

13、根据权利要求 1 所述的方法，其特征在于，所述深度传感器的帧率与所述图像传感器的帧率不同；

10 所述第二视频帧是与第一视频帧的采集时刻时间最接近的视频帧。

14、根据权利要求 13 所述的方法，其特征在于，所述第二视频帧是从缓存区中获取的，所述缓存区中缓存有所述深度传感器采集的携带有时间戳的至少一帧图像。

15、根据权利要求 13 所述的方法，其特征在于，所述第一视频帧是从缓存区中获取的，所述缓存区中缓存有所述图像传感器采集的携带有时间戳的至少一帧图像。

15 16、根据权利要求 14 或 15 所述的方法，其特征在于，所述缓存区中缓存的图像的帧数，是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

17、根据权利要求 14 或 15 所述的方法，其特征在于，所述缓存区中缓存的图像的缓存时长，是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

20 18、根据权利要求 16 所述的方法，其特征在于，所述缓存区中缓存的图像的帧数，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的商值。

19、根据权利要求 17 所述的方法，其特征在于，所述缓存区中缓存的图像的缓存时长，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的差值。

20、根据权利要求 1 所述的方法，其特征在于，所述图像传感器采集的图像的第一分辨率高于所述深度传感器采集的图像的第二分辨率；

25 所述第二视频帧是对所述深度传感器采集的图像进行插值处理得到的与所述第一分辨率相同的视频帧。

21、根据权利要求 1 所述的方法，其特征在于，所述第一视频帧与第二视频帧是经过图像对齐处理后得到的。

22、根据权利要求 21 所述的方法，其特征在于，所述图像对齐处理，包括如下任一：
30 一：将所述图像传感器采集的图像与所述深度传感器采集的图像进行轮廓对齐，或，将所述图像传感器采集的图像与所述深度传感器采集的图像两者的偏移区域进行裁剪

处理。

23、根据权利要求 22 所述的方法，其特征在于，所述偏移区域，是基于所述图像传感器的安装位置和所述深度传感器的安装位置的位置差异确定的。

5 24、根据权利要求 1 所述的方法，其特征在于，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

基于所述感兴趣区域的深度信息的整体置信度，将所述焦点位置与基于其他跟焦模式确定的焦点位置进行融合，利用融合结果控制拍摄设备在拍摄视频时进行跟焦。

25、根据权利要求 1 所述的方法，其特征在于，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

10 基于所述感兴趣区域的深度信息的整体置信度，将所述感兴趣区域的深度信息与基于其他跟焦模式确定的深度信息进行融合，利用融合结果确定出的焦点位置控制拍摄设备在拍摄视频时进行跟焦。

15 26、一种跟焦装置，其特征在于，所述装置包括处理器、存储器、存储在所述存储器上可被所述处理器执行的计算机程序，所述处理器执行所述计算机程序时实现以下步骤：

利用图像传感器获取第一视频帧后，获取所述第一视频帧中的感兴趣区域；

利用深度传感器获取携带有深度信息的第二视频帧；其中，所述第一视频帧和第二视频帧包含相同画面内容；

利用所述第二视频帧确定所述感兴趣区域的深度信息；

20 基于所述感兴趣区域的深度信息确定焦点位置，利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦。

27、根据权利要求 26 所述的装置，其特征在于，在所述拍摄设备中，所述深度传感器与所述图像传感器平行。

25 28、根据权利要求 27 所述的装置，其特征在于，所述深度传感器包括：3D ToF 传感器。

29、根据权利要求 27 所述的装置，其特征在于，所述利用所述第二视频帧确定所述感兴趣区域的深度信息，包括：

利用所述感兴趣区域在所述第一视频帧中的位置，从所述第二视频帧中确定所述感兴趣区域的深度信息。

30 30、根据权利要求 26 所述的装置，其特征在于，所述焦点位置是利用所述感兴趣区域的平均深度信息确定的，所述感兴趣区域的平均深度信息是基于所述感兴趣区域

的各个目标像素点的深度信息确定的。

31、根据权利要求 30 所述的装置，其特征在于，所述感兴趣区域的平均深度信息是基于所述感兴趣区域中各个目标像素点的深度信息以及各个所述目标像素点的预设权重确定的。

5 32、根据权利要求 30 或 31 所述的装置，其特征在于，所述目标像素点是所述感兴趣区域中深度信息大于设定置信度阈值的像素点。

33、根据权利要求 32 所述的装置，其特征在于，所述置信度阈值是基于所述深度传感器的回波强度设定的。

10 34、根据权利要求 31 所述的装置，其特征在于，所述感兴趣区域包括：从所述第一视频帧中识别得到的人脸区域。

35、根据权利要求 34 所述的装置，其特征在于，所述人脸区域中，眼部区域像素点的权重、嘴部区域像素点的权重、鼻子区域像素点的权重和边缘区域像素点的权重依次递减。

15 36、根据权利要求 26 所述的装置，其特征在于，所述感兴趣区域是基于用户指定的需跟焦的人脸区域确定的。

37、根据权利要求 36 所述的装置，其特征在于，所述获取所述第一视频帧中的感兴趣区域：

在拍摄设备的预览画面中突出显示从图像传感器采集的图像中识别出的至少两个人脸区域后，获取用户指定的需跟焦的人脸区域；

20 利用所述需跟焦的人脸区域，获取所述第一视频帧中的人脸区域。

38、根据权利要求 26 所述的装置，其特征在于，所述深度传感器的帧率与所述图像传感器的帧率不同；

所述第二视频帧是与第一视频帧的采集时刻时间最接近的视频帧。

25 39、根据权利要求 38 所述的装置，其特征在于，所述第二视频帧是从缓存区中获取的，所述缓存区中缓存有所述深度传感器采集的携带有时间戳的至少一帧图像。

40、根据权利要求 38 所述的装置，其特征在于，所述第一视频帧是从缓存区中获取的，所述缓存区中缓存有所述图像传感器采集的携带有时间戳的至少一帧图像。

41、根据权利要求 39 或 40 所述的装置，其特征在于，所述缓存区中缓存的图像的帧数，是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

30 42、根据权利要求 39 或 40 所述的装置，其特征在于，所述缓存区中缓存的图像的缓存时长，是基于所述深度传感器的帧率与所述图像传感器的帧率的差异确定的。

43、根据权利要求 41 所述的装置，其特征在于，所述缓存区中缓存的图像的帧数，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的商值。

44、根据权利要求 42 所述的装置，其特征在于，所述缓存区中缓存的图像的缓存时长，大于或等于所述深度传感器的帧率倒数与所述图像传感器的帧率倒数的差值。

5 45、根据权利要求 26 所述的装置，其特征在于，所述图像传感器采集的图像的第一分辨率高于所述深度传感器采集的图像的第二分辨率；

所述第二视频帧是对所述深度传感器采集的图像进行插值处理得到的与所述第一分辨率相同的视频帧。

10 46、根据权利要求 45 所述的装置，其特征在于，在对所述深度传感器采集的图像进行插值处理时，还对所述深度传感器采集的图像进行与所述第一视频帧的图像对齐处理。

47、根据权利要求 46 所述的装置，其特征在于，所述图像对齐处理，包括如下任一：将所述深度传感器采集的图像与所述第一视频帧的轮廓对齐，或，将所述深度传感器采集的图像与所述第一视频帧两者的偏移区域进行裁剪。

15 48、根据权利要求 47 所述的装置，其特征在于，所述偏移区域，是基于所述图像传感器的安装位置和所述深度传感器的安装位置的位置差异确定的。

49、根据权利要求 26 所述的装置，其特征在于，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

20 基于所述感兴趣区域的深度信息的整体置信度，将所述焦点位置与基于其他跟焦模式确定的焦点位置进行融合，利用融合结果控制拍摄设备在拍摄视频时进行跟焦。

50、根据权利要求 26 所述的装置，其特征在于，所述利用所述焦点位置控制拍摄设备在拍摄视频时进行跟焦，包括：

25 基于所述感兴趣区域的深度信息的整体置信度，将所述感兴趣区域的深度信息与基于其他跟焦模式确定的深度信息进行融合，利用融合结果确定出的焦点位置控制拍摄设备在拍摄视频时进行跟焦。

51、一种拍摄设备，其特征在于，包括图像传感器和深度传感器；以及，如权利要求 26 至 50 任意一项所述跟焦装置。

52、一种计算机可读存储介质，其特征在于，所述可读存储介质上存储有若干计算机指令，所述计算机指令被执行时实现权利要求 1 至 25 任一项所述方法的步骤。

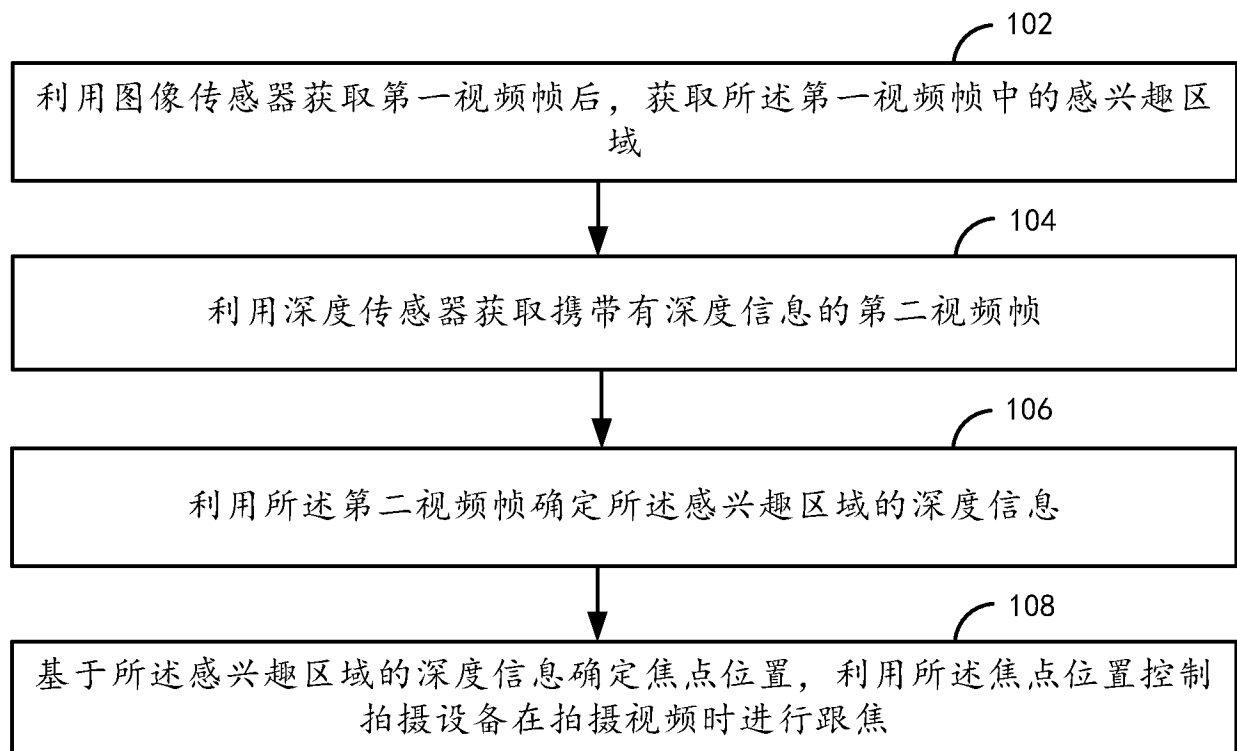


图 1A

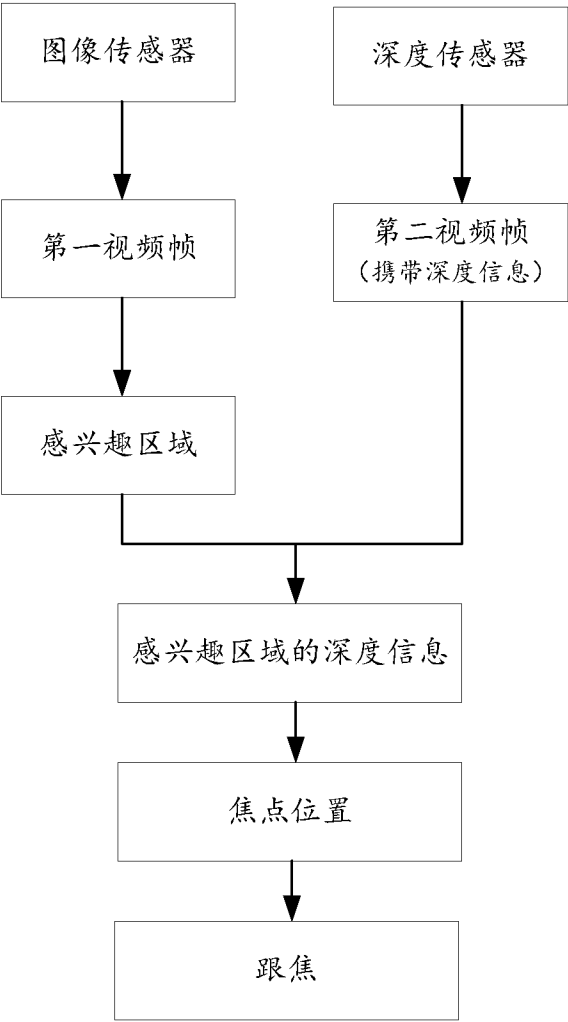


图 1B

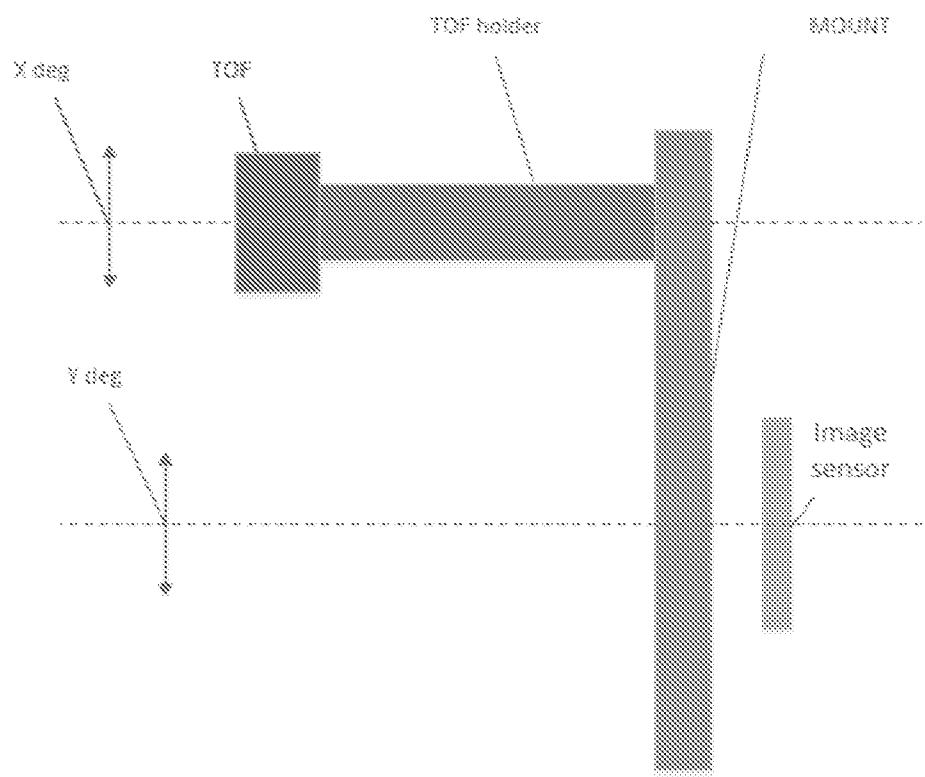


图 1C

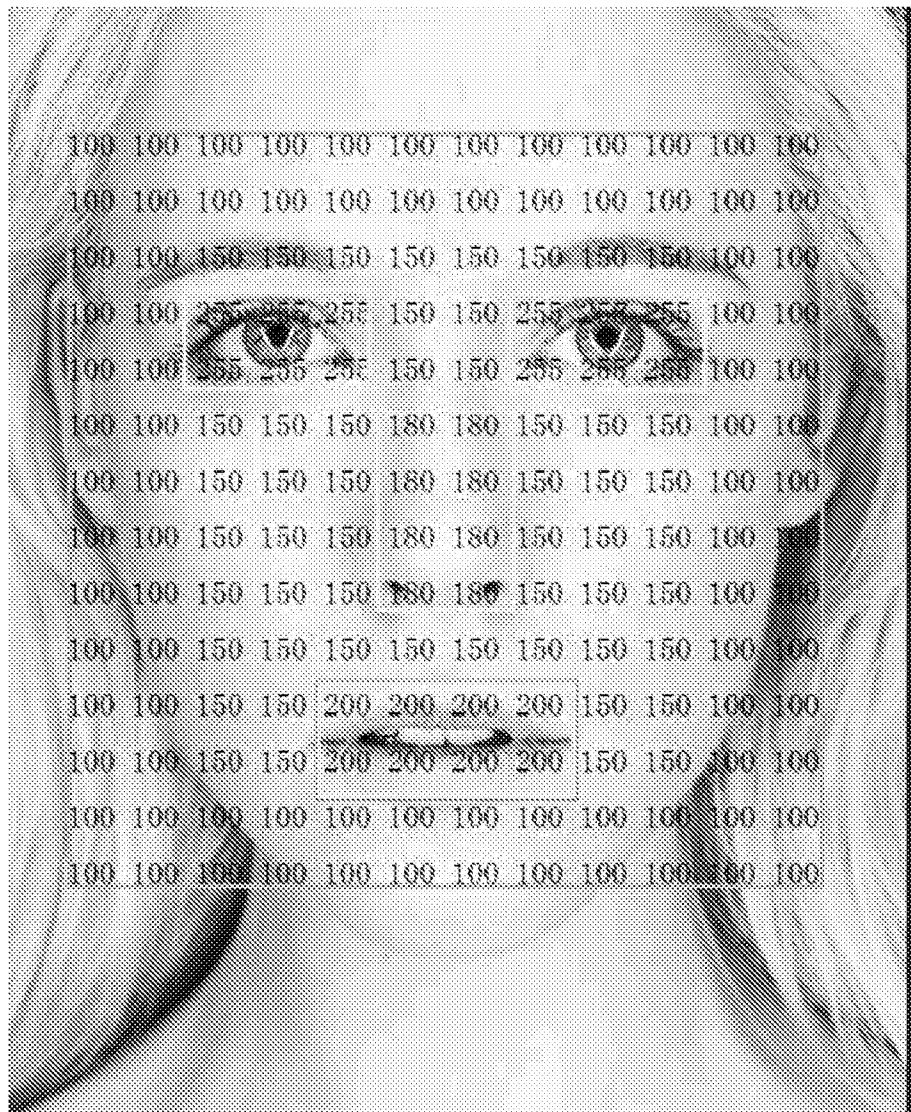


图 1D

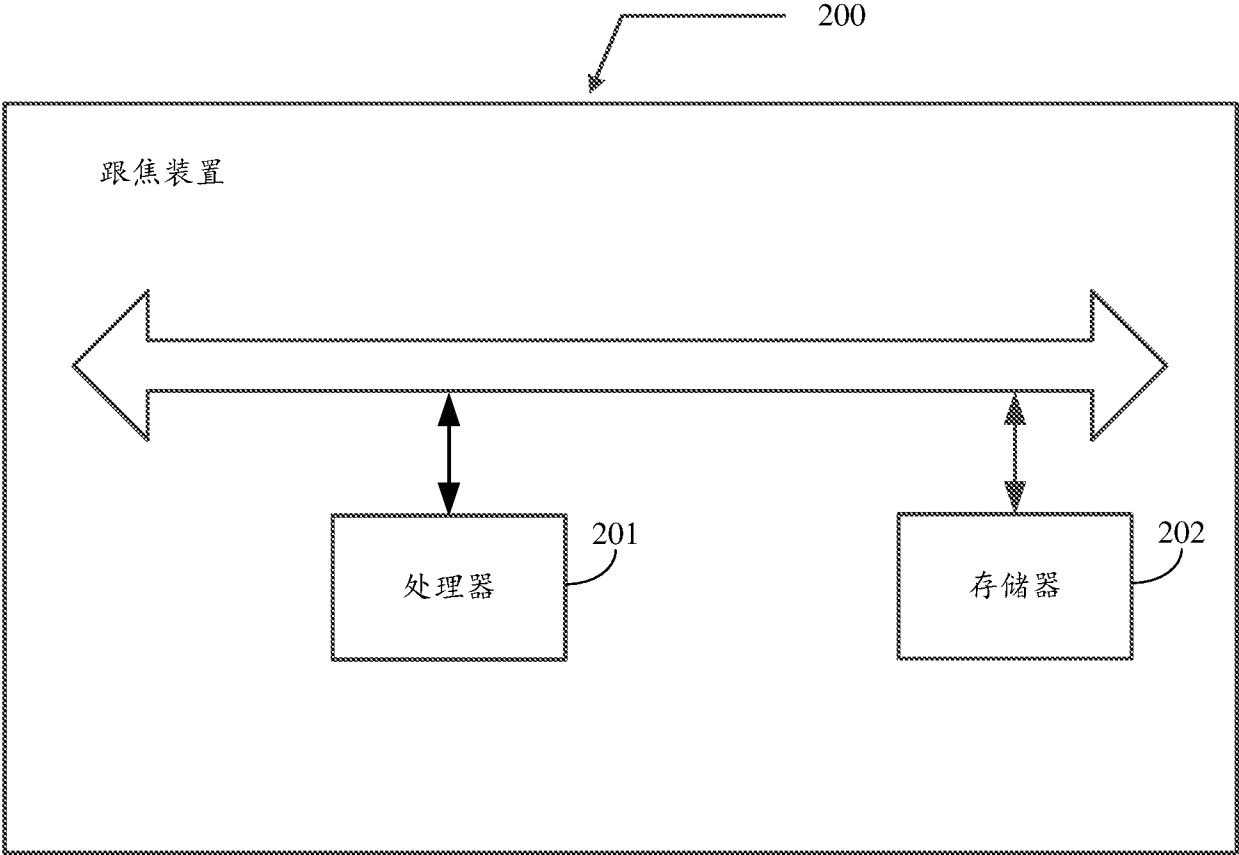


图 2

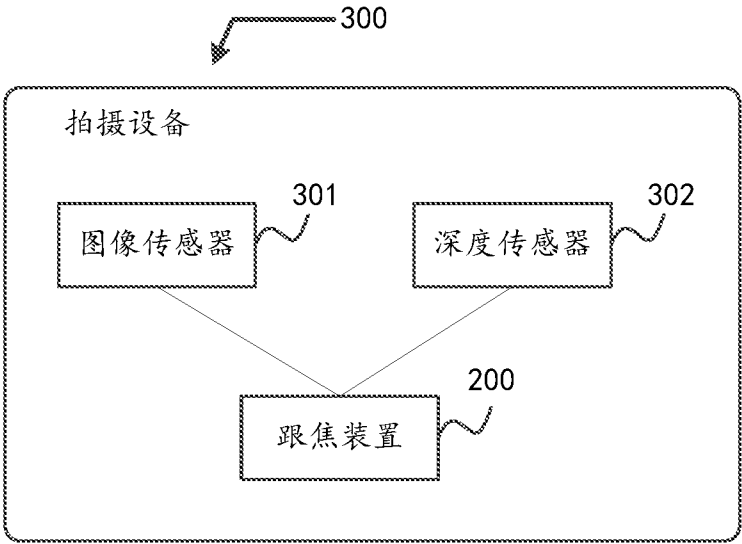


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/070581

A. CLASSIFICATION OF SUBJECT MATTER G03B 13/36(2021.01)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) G03B 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) CNPAT; CNKI; WPI; EPODOC: 跟焦, 变焦, 对焦, 深度, 景深, 距离, 图像, 帧, 感兴趣区域, 焦点, ROI, TOF, distance, focus +, follow, automatic, depth																				
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 111226154 A (SHENZHEN DAJIANG INNOVATION TECHNOLOGY CO., LTD.) 02 June 2020 (2020-06-02) description, paragraphs [0056] and [0060]-[0067]</td> <td>1-52</td> </tr> <tr> <td>Y</td> <td>CN 109696667 A (HRG INTERNATIONAL INSTITUTE (HEFEI) OF RESEARCH AND INNOVATION) 30 April 2019 (2019-04-30) description, paragraph [0034]</td> <td>1-52</td> </tr> <tr> <td>A</td> <td>CN 105264436 A (CINEMA CONTROL LAB INC.) 20 January 2016 (2016-01-20) entire document</td> <td>1-52</td> </tr> <tr> <td>A</td> <td>CN 110381261 A (CHONGQING UNISINSIGHT TECHNOLOGY CO., LTD.) 25 October 2019 (2019-10-25) entire document</td> <td>1-52</td> </tr> <tr> <td>A</td> <td>US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 28 December 2017 (2017-12-28) entire document</td> <td>1-52</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	Y	CN 111226154 A (SHENZHEN DAJIANG INNOVATION TECHNOLOGY CO., LTD.) 02 June 2020 (2020-06-02) description, paragraphs [0056] and [0060]-[0067]	1-52	Y	CN 109696667 A (HRG INTERNATIONAL INSTITUTE (HEFEI) OF RESEARCH AND INNOVATION) 30 April 2019 (2019-04-30) description, paragraph [0034]	1-52	A	CN 105264436 A (CINEMA CONTROL LAB INC.) 20 January 2016 (2016-01-20) entire document	1-52	A	CN 110381261 A (CHONGQING UNISINSIGHT TECHNOLOGY CO., LTD.) 25 October 2019 (2019-10-25) entire document	1-52	A	US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 28 December 2017 (2017-12-28) entire document	1-52
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																		
Y	CN 111226154 A (SHENZHEN DAJIANG INNOVATION TECHNOLOGY CO., LTD.) 02 June 2020 (2020-06-02) description, paragraphs [0056] and [0060]-[0067]	1-52																		
Y	CN 109696667 A (HRG INTERNATIONAL INSTITUTE (HEFEI) OF RESEARCH AND INNOVATION) 30 April 2019 (2019-04-30) description, paragraph [0034]	1-52																		
A	CN 105264436 A (CINEMA CONTROL LAB INC.) 20 January 2016 (2016-01-20) entire document	1-52																		
A	CN 110381261 A (CHONGQING UNISINSIGHT TECHNOLOGY CO., LTD.) 25 October 2019 (2019-10-25) entire document	1-52																		
A	US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 28 December 2017 (2017-12-28) entire document	1-52																		
☐ 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 26 September 2021		Date of mailing of the international search report 11 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
Information on patent family members

International application No.

PCT/CN2021/070581

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111226154	A	02 June 2020	WO	2020061857	A1	02 April 2020
				US	2020267323	A1	20 August 2020
CN	109696667	A	30 April 2019	None			
CN	105264436	A	20 January 2016	US	2018176456	A1	21 June 2018
				JP	2016522598	A	28 July 2016
				EP	2987026	A1	24 February 2016
				CA	2908719	A1	09 October 2014
				US	2016050360	A1	18 February 2016
				WO	2014161092	A1	09 October 2014
				IN	201510285	P1	05 February 2016
				HK	1220512	A1	17 April 2020
CN	110381261	A	25 October 2019	None			
US	2017374354	A1	28 December 2017	CN	111371986	A	03 July 2020
				CN	107409205	A	28 November 2017
				EP	3108653	A1	28 December 2016
				US	2020195908	A1	18 June 2020
				WO	2016145602	A1	22 September 2016

国际检索报告

国际申请号

PCT/CN2021/070581

A. 主题的分类 G03B 13/36 (2021.01) 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) G03B 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)) CNPAT; CNKI; WPI; EPDOC: 跟焦, 变焦, 对焦, 深度, 景深, 距离, 图像, 帧, 感兴趣区域, 焦点, ROI, TOF, distance, focus+, follow, automatic, depth																				
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>Y</td> <td>CN 111226154 A (深圳市大疆创新科技有限公司) 2020年 6月 2日 (2020 - 06 - 02) 说明书[0056]、[0060]–[0067]段</td> <td>1–52</td> </tr> <tr> <td>Y</td> <td>CN 109696667 A (哈工大机器人合肥国际创新研究院) 2019年 4月 30日 (2019 - 04 - 30) 说明书[0034]段</td> <td>1–52</td> </tr> <tr> <td>A</td> <td>CN 105264436 A (安德拉运动技术股份有限公司) 2016年 1月 20日 (2016 - 01 - 20) 全文</td> <td>1–52</td> </tr> <tr> <td>A</td> <td>CN 110381261 A (重庆紫光华山智安科技有限公司) 2019年 10月 25日 (2019 - 10 - 25) 全文</td> <td>1–52</td> </tr> <tr> <td>A</td> <td>US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 2017年 12月 28日 (2017 - 12 - 28) 全文</td> <td>1–52</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	Y	CN 111226154 A (深圳市大疆创新科技有限公司) 2020年 6月 2日 (2020 - 06 - 02) 说明书[0056]、[0060]–[0067]段	1–52	Y	CN 109696667 A (哈工大机器人合肥国际创新研究院) 2019年 4月 30日 (2019 - 04 - 30) 说明书[0034]段	1–52	A	CN 105264436 A (安德拉运动技术股份有限公司) 2016年 1月 20日 (2016 - 01 - 20) 全文	1–52	A	CN 110381261 A (重庆紫光华山智安科技有限公司) 2019年 10月 25日 (2019 - 10 - 25) 全文	1–52	A	US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 2017年 12月 28日 (2017 - 12 - 28) 全文	1–52
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
Y	CN 111226154 A (深圳市大疆创新科技有限公司) 2020年 6月 2日 (2020 - 06 - 02) 说明书[0056]、[0060]–[0067]段	1–52																		
Y	CN 109696667 A (哈工大机器人合肥国际创新研究院) 2019年 4月 30日 (2019 - 04 - 30) 说明书[0034]段	1–52																		
A	CN 105264436 A (安德拉运动技术股份有限公司) 2016年 1月 20日 (2016 - 01 - 20) 全文	1–52																		
A	CN 110381261 A (重庆紫光华山智安科技有限公司) 2019年 10月 25日 (2019 - 10 - 25) 全文	1–52																		
A	US 2017374354 A1 (SZ DJI TECHNOLOGY CO., LTD.) 2017年 12月 28日 (2017 - 12 - 28) 全文	1–52																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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 9月 26日		国际检索报告邮寄日期 2021年 10月 11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 李迪 电话号码 86-(10)-53961300																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/070581

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	111226154	A	2020年 6月 2日	WO	2020061857	A1	2020年 4月 2日
				US	2020267323	A1	2020年 8月 20日
CN	109696667	A	2019年 4月 30日	无			
CN	105264436	A	2016年 1月 20日	US	2018176456	A1	2018年 6月 21日
				JP	2016522598	A	2016年 7月 28日
				EP	2987026	A1	2016年 2月 24日
				CA	2908719	A1	2014年 10月 9日
				US	2016050360	A1	2016年 2月 18日
				WO	2014161092	A1	2014年 10月 9日
				IN	201510285	P1	2016年 2月 5日
				HK	1220512	A1	2020年 4月 17日
CN	110381261	A	2019年 10月 25日	无			
US	2017374354	A1	2017年 12月 28日	CN	111371986	A	2020年 7月 3日
				CN	107409205	A	2017年 11月 28日
				EP	3108653	A1	2016年 12月 28日
				US	2020195908	A1	2020年 6月 18日
				WO	2016145602	A1	2016年 9月 22日