

(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
2020 年 8 月 20 日 (20.08.2020)



(10) 国际公布号
WO 2020/164277 A1

(51) 国际专利分类号:
H04N 7/18 (2006.01) *G05D 3/12* (2006.01)

(21) 国际申请号: PCT/CN2019/118158

(22) 国际申请日: 2019 年 11 月 13 日 (13.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910111104.2 2019 年 2 月 12 日 (12.02.2019) 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 518000 (CN)。

(72) 发明人: 王路生(WANG, Lusheng); 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)。 陈斌(CHEN, Bin); 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)。 宋晨(SONG, Chen); 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: 深圳中一联合知识产权代理有限公司 (SHENZHEN ZHONGYI UNION INTELLECTUAL PROPERTY AGENCY CO., LTD.);
中国广东省深圳市福田区园岭街道深

(54) Title: MONITORING METHOD AND APPARATUS BASED ON AUDIO AND VIDEO LINKAGE, AND TERMINAL DEVICE AND MEDIUM

(54) 发明名称: 一种基于音视频联动的监控方法、装置、终端设备及介质

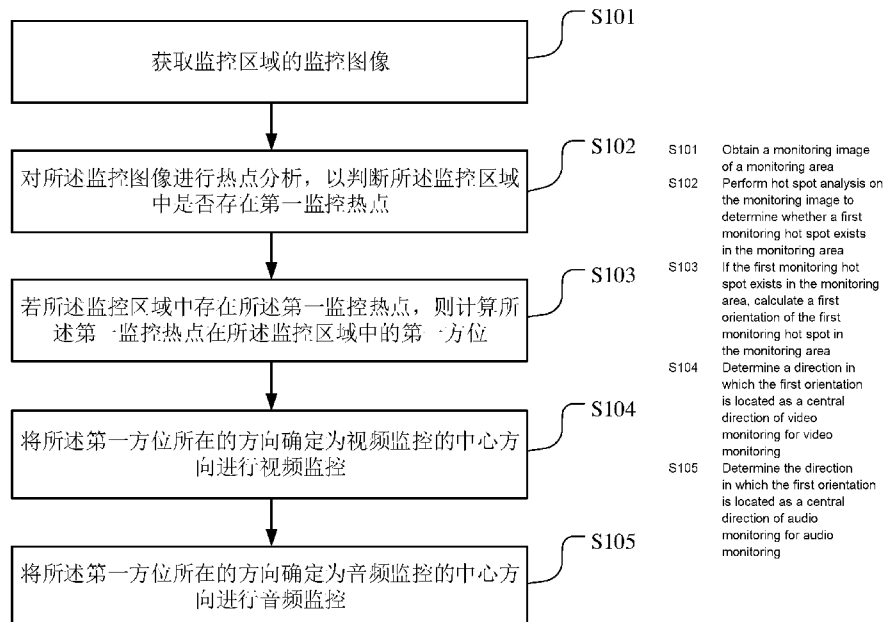


图 1

(57) Abstract: The present application relates to the technical field of security and protection monitoring, and in particular, to a monitoring method and apparatus based on audio and video linkage, and a terminal device and a medium. The monitoring method comprises: obtaining a monitoring image of a monitoring area; performing hot spot analysis on the monitoring image to determine whether a first monitoring hot spot exists in the monitoring area; if the first monitoring hot spot exists in the monitoring area, calculating a first orientation of the first monitoring hot spot in the monitoring area; determining a direction in which the first orientation is located as a

南中路 1014 号报春大厦 9 楼（5 号信箱），
Guangdong 518028 (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, 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))。

central direction of video monitoring for video monitoring; and determining the direction in which the first orientation is located as a central direction of audio monitoring for audio monitoring. According to the present application, hardware costs of a video monitoring device can be reduced when large-scale video monitoring is performed, and the sound quality is facilitated to be improved when audio monitoring is performed.

(57) 摘要：本申请属于安防监控技术领域，尤其涉及一种基于音视频联动的监控方法、装置、终端设备及介质。其中，所述监控方法包括：获取监控区域的监控图像；对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。本申请在进行较大范围的视频监控时能够降低视频监控设备的硬件成本，并且在进行音频监控时有利于提高声音质量。

一种基于音视频联动的监控方法、装置、终端设备及介质

[0001] 本申请要求于2019年02月12日提交中国专利局、申请号为201910111104.2、发明名称为“一种基于音视频联动的监控方法、装置、终端及存储介质”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

[0002] 本申请属于安防监控技术领域，尤其涉及一种基于音视频联动的监控方法、装置、终端设备及介质。

背景技术

[0003] 视频监控是安全防范系统的重要组成部分，传统技术中的视频监控侧重于对监控区域的图像信息进行采集，通常不进行声音采集，而市场上存在的少量配置有声音采集设备的监控终端设备，由于监控区域具备一定的范围，声音采集方向难以对准声源方向，导致采集的声音质量较低，不能起到音频监控的应有作用。

[0004] 另外，由于视频监控的摄像设备存在一定的监控视野范围，导致在监控视范围边缘附近发生的事件难以被摄像设备完整记录，传统技术中通常采用配备多个摄像设备的方式进行多摄像头组合监控，这样提高了监控的硬件成本。

发明概述

技术问题

[0005] 有鉴于此，本申请提供了一种基于音视频联动的监控方法、装置、终端设备及介质，以解决现有技术中音频监控的声音质量不高以及在进行较大范围视频监控时硬件成本高的问题。

问题的解决方案

技术解决方案

[0006] 本申请实施例的第一方面提供了一种基于音视频联动的监控方法，包括：

- [0007] 获取监控区域的监控图像；
- [0008] 对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
- [0009] 若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
- [0010] 将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；
- [0011] 将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0012] 本申请实施例的第二方面提供了一种基于音视频联动的监控装置，包括：
- [0013] 图像获取单元，用于获取监控区域的监控图像；
- [0014] 第一热点分析单元，用于对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
- [0015] 第一方位计算单元，用于若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
- [0016] 视频监控单元，用于将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；
- [0017] 音频监控单元，用于将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0018] 本申请实施例的第三方面提供了一种终端设备，包括存储器、处理器，所述存储器上存储有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：
- [0019] 对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
- [0020] 若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
- [0021] 将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；
- [0022] 将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0023] 本申请实施例的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可读指令，其特征在于，所述计算机可读指令被至少一个

处理器执行时实现如下步骤:

- [0024] 对所述监控图像进行热点分析, 以判断所述监控区域中是否存在第一监控热点;
- [0025] 若所述监控区域中存在所述第一监控热点, 则计算所述第一监控热点在所述监控区域中的第一方位;
- [0026] 将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控;
- [0027] 将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。

发明的有益效果

有益效果

- [0028] 本申请通过获取监控区域的监控图像; 对所述监控图像进行热点分析, 以判断所述监控区域中是否存在第一监控热点; 若所述监控区域中存在所述第一监控热点, 则计算所述第一监控热点在所述监控区域中的第一方位; 将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控; 也即, 通过对监控图像进行热点分析来发现监控区域的监控热点, 并通过调整监控视野的中心方向使监控热点处于监控视野中心, 如此一来在进行较大范围的视频监控时可以无需额外布置多个监控设备, 能够降低视频监控设备的硬件成本; 另一方面, 通过将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控, 可以将音频监控设备的声音采集方向对准监控热点方向, 从而在进行音频监控时有利于提高声音质量。

对附图的简要说明

附图说明

- [0029] 为了更清楚地说明本申请实施例中的技术方案, 下面将对实施例或现有技术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。
- [0030] 图1为本申请实施例中基于音视频联动的监控方法的一个实施例流程图;
- [0031] 图2为本申请实施例中基于音视频联动的监控方法的另一个实施例流程图;
- [0032] 图3为本申请实施例中基于音视频联动的监控装置的一个实施例结构图;

[0033] 图4为本申请实施例中一种终端设备的示意框图。

发明实施例

本发明的实施方式

[0034] 以下描述中，为了说明而不是为了限定，提出了诸如特定系统结构、技术之类的具体细节，以便透彻理解本申请实施例。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本申请。在其它情况中，省略对众所周知的系统、装置、电路以及方法的详细说明，以免不必要的细节妨碍本申请的描述。

[0035] 为了说明本申请所述的技术方案，下面通过具体实施例来进行说明。

[0036] 请参阅图1，本申请实施例中基于音视频联动的监控方法的一个实施例可以包括：

[0037] 在步骤S101中、获取监控区域的监控图像。

[0038] 在本申请实施例中，获取监控终端设备拍摄的监控区域的监控图像，具体的，该监控终端设备可以配置有摄像组件，通过摄像组件对监控区域进行视频监控，采集监控区域的实时画面图像。

[0039] 在一种实现方式中，摄像组件可以设置有云台，通过云台可以调整摄像组件的视频画面拍摄方向。

[0040] 在本申请实施例中，当监控区域的范围大于摄像组件的监控视野时，可以通过云台来控制摄像组件的拍摄方向在监控区域范围内循环移动，实现对范围大于摄像组件的监控视野的监控区域的视频监控。

[0041] 在步骤S102中、对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点。

[0042] 在视频监控过程中，可以实时或每隔指定的时间间隔获取监控区域的监控图像，并对该监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点。

[0043] 在本申请实施例中，第一监控热点是指监控区域中的监控重点，例如，监控热点可以是监控区域中出现的一个目标人物，也可以是监控区域中出现的一个冲突事件。在实际应用中，可以通过预先指定该监控热点的图像特征是实现对

监控区域中是否出现监控热点的判断。

[0044] 例如，以某目标人物A为第一监控热点时，可以预先存储该目标人物A的人脸特征，在视频监控过程中，通过获取监控图像，并对监控图像进行人脸特征识别，当识别出监控图像上出现该目标人物A的人脸特征时，判定监控区域中出现第一监控热点，监控区域中目标人物A的所在位置也即第一监控热点的位置。

[0045] 在一种实现方式中，上述步骤S102可以包括：

[0046] 判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域；

[0047] 若存在所述图像区域，则判定所述监控区域中存在第一监控热点；

[0048] 若不存在所述图像区域，则判定所述监控区域中不存在第一监控热点。

[0049] 在本实施例中，可以预先设置指定大小的滑动窗口，并将该滑动窗口遍历监控图像，在遍历过程中，统计滑动窗口的人群密度，并判断滑动窗口的人群密度是否大于预设的密度阈值，将滑动窗口的人群密度大于预设的密度阈值时，该滑动窗口所在位置对应的图像区域确定第一监控热点的所在区域。

[0050] 需要说明的是，当监控图像中存在两个以上大于预设密度阈值的图像区域时，可以选取其中人群密度最大的一个监控热点作为第一监控热点。

[0051] 在本实施例中，可以通过识别滑动窗口中的人头数量，来确定人群密度。人群密度较大时，表示可能为一个群体事件，故可将其作为监控终端设备。

[0052] 在步骤S103中、若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位。

[0053] 在本申请实施例中，由于监控图像和监控区域存在一定的对应关系，可以根据第一监控热点对应的图像区域在监控图像上的位置来确定第一监控热点在监控区域中的第一方位。

[0054] 在一种实现方式中，上述步骤S103可以包括：

[0055] 根据所述图像区域的中心像素点在所述监控图像上的位置，得到第一坐标；

[0056] 根据预设的坐标方位对应关系，确定所述第一坐标在所述监控区域中的第一方位。

[0057] 在本实施例中，监控图像和监控区域存在的对应关系与摄像组件的实际参数相关，基于摄像组件的实际参数，可以预先建立坐标方位对应关系，该坐标方位

对应关系存储有监控图像上的像素点坐标与实际监控区域方位的对应关系。由此，对于监控图像上与第一监控热点对应的图像区域，可以选取该图像区域的中心像素点，根据坐标方位对应关系，查找与该中心像素点的坐标对应的方位，即可得到所述第一方位。

[0058] 在步骤S104中、将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控。

[0059] 在步骤S105中、将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。

[0060] 在本申请实施例中，由于第一方位为监控区域中的第一监控热点，可以通过调整云台将摄像组件监控的中心方向调整为该第一方向，也即使得第一监控热点位于摄像组件监控视野的中心位置，这样可以使得监控区域中发生监控热点对应的事件时，能够获得较为全面的监控视频信息。

[0061] 同理，在本申请实施例中，由于第一方位为监控区域中的第一监控热点，可以通过调整音频监控设备的音频采集方向，使得音频监控设备的音频采集方向与该第一方向一致，也即使得音频监控设备的音频采集方向对准了第一监控热点，这样可以使得监控区域中发生监控热点对应的事件时，能够获得较高质量的监控音频信息。

[0062] 综上所述，本申请通过获取监控区域的监控图像；对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；也即，通过对监控图像进行热点分析来发现监控区域的监控热点，并通过调整监控视野的中心方向使监控热点处于监控视野中心，如此一来在进行较大范围的视频监控时可以无需额外布置多个监控设备，能够降低视频监控设备的硬件成本；另一方面，通过将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控，可以将音频监控设备的声音采集方向对准监控热点方向，从而在进行音频监控时有利于提高声音质量。

[0063] 请参阅图2，本申请实施例中基于音视频联动的监控方法的另一个实施例可以

包括：

[0064] 步骤S201、获取所述监控区域的监控音频。

[0065] 在本申请实施例中，监控终端设备可以配置有音频采集组件，例如该音频采集组件可以为麦克风。通过音频采集组件采集监控区域的音频信息，获取音频采集组件采集到的监控区域的监控音频。

[0066] 在一种实现方式中，可以以一定的速度调整音频采集组件的音频采集方向，使得该音频采集组件的音频采集方向可以遍历监控区域，以便采集到更多的音频信息。

[0067] 步骤S202、对所述监控音频进行热点分析，以判断所述监控区域中是否存在第二监控热点。

[0068] 在音频监控过程中，可以实时对采集的监控音频进行热点分析，以判断所述监控区域中是否存在第二监控热点。在这里，第二监控热点是指通过声音识别到的监控区域的监控重点，例如，第二监控热点可以为指定的声音（通过声纹特征进行识别），也可以为指定的语音（例如救命或者“help”，通过语音转文字进行语音识别），还可以为声音强度超过一定值的声源。

[0069] 在一种实现方式中，上述步骤S202可以包括：

[0070] 判断所述监控音频中是否存在声音强度大于预设强度阈值的目标声源；

[0071] 若存在所述目标声源，则判定所述监控区域中存在第二监控热点；

[0072] 若不存在所述目标声源，则判定所述监控区域中不存在第二监控热点。

[0073] 监控音频中出现较大声音强度的声音，表示该监控区域中可能出现需重点监控的事件。在本实施例中，可以对采集的不同声源的声音信息进行声音强度的检测，当检测到声音强度大于预设强度阈值的目标声源时，判定监控区域中存在第二监控热点。

[0074] 步骤S203、若所述监控区域中存在所述第二监控热点，则计算所述第二监控热点在所述监控区域中的第二方位。

[0075] 在本申请实施例中，对于监控区域中的第二监控热点，可以通过声音传播时间和声音传播方向的信息进行计算，获得目标声源的方位。

[0076] 在一种实现方式中，音频采集设备可以为麦克风阵列，通过麦克风阵列来采集

监控区域的音频信息，上述步骤S203可以包括：

- [0077] 根据所述麦克风阵列中各麦克风的位置排布，以及，所述麦克风阵列中各麦克风采集到所述目标声源发出的声音的时间差，计算所述目标声源在所述第二监控热点在所述监控区域中的第二方位。
- [0078] 步骤S204、将所述第二方位所在的方向确定为视频监控的中心方向进行视频监控。
- [0079] 步骤S205、将所述第二方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0080] 在本申请实施例中，由于第二方位为监控区域中的第二监控热点，可以通过调整云台将摄像组件监控的中心方向调整为该第二方向，也即使得第二监控热点位于摄像组件监控视野的中心位置，这样可以使得监控区域中发生监控热点对应的事件时，能够获得较为全面的监控视频信息。
- [0081] 同理，在本申请实施例中，由于第二方位为监控区域中的第二监控热点，可以通过调整音频监控设备的音频采集方向，使得音频监控设备的音频采集方向与该第二方向一致，也即使得音频监控设备的音频采集方向对准了第二监控热点，这样可以使得监控区域中发生监控热点对应的事件时，能够获得较高质量的监控音频信息。
- [0082] 在一种实现方式中，上述图1所示实施例和图2所示实施例可以共同实施，以实现音频监控和视频监控的联动。可以使得摄像组件能够获取到较多的视频画面信息，并且，使得音频采集设备获取到较高质量的声音信息。
- [0083] 可选的，在上述图1所示实施例和图2所示实施例合并实施的方案中，若所述监控区域中同时存在所述第一监控热点与所述第二监控热点，且，所述第一监控热点与所述第二监控热点对应的第一方位和第二方位不一致，则可以分别计算所述第一监控热点和第二监控热点的热点度得分；例如，不同的人群密度值对应于一定的热点度分值，不同的声音强度对应于一定的热点度分值，通过将热点度得分较高的监控热点对应的方位确定为目标方位；
- [0084] 将所述目标方位所在的方向确定为视频监控的中心方向进行视频监控；
- [0085] 将所述目标方位所在的方向确定为音频监控的中心方向进行音频监控。

- [0086] 在本申请实施例中，通过热点度得分的比较，从第一监控热点和第二监控热点中选取更重要的监控热点作为监控的中心点，得到目标方位，并将目标方位所在的方向确定为视频监控的中心方向进行视频监控，以及音频监控的中心方向进行音频监控，实现对监控区域中最重要的事件进行重点监控。
- [0087] 综上所述，本申请通过获取监控区域的监控图像；对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；也即，通过对监控图像进行热点分析来发现监控区域的监控热点，并通过调整监控视野的中心方向使监控热点处于监控视野中心，如此一来在进行较大范围的视频监控时可以无需额外布置多个监控设备，能够降低视频监控设备的硬件成本；另一方面，通过将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控，可以将音频监控设备的声音采集方向对准监控热点方向，从而在进行音频监控时有利于提高声音质量。
- [0088] 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。
- [0089] 对应于上文实施例所述的基于音视频联动的监控方法，图3示出了本申请实施例提供的基于音视频联动的监控装置的一个实施例结构图。
- [0090] 本实施例中，基于音视频联动的监控装置3可以包括：图像获取单元31，第一热点分析单元32，第一方位计算单元33，视频监控单元34和音频监控单元35。
- [0091] 图像获取单元31，用于获取监控区域的监控图像；
- [0092] 第一热点分析单元32，用于对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
- [0093] 第一方位计算单元33，用于若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
- [0094] 视频监控单元34，用于将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；

- [0095] 音频监控单元35, 用于将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0096] 可选的, 基于音视频联动的监控装置3还可以包括:
- [0097] 密度判断单元, 用于判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域;
- [0098] 第一热点分析单元32具体用于, 若密度判断单元判定存在所述图像区域, 则判定所述监控区域中存在第一监控热点; 以及, 若密度判断单元判定不存在所述图像区域, 则判定所述监控区域中不存在第一监控热点。
- [0099] 可选的, 基于音视频联动的监控装置3还可以包括:
- [0100] 坐标获取单元, 用于根据所述图像区域的中心像素点在所述监控图像上的位置, 得到第一坐标;
- [0101] 第一方位计算单元33具体用于, 根据预设的坐标方位对应关系, 确定所述第一坐标在所述监控区域中的第一方位。
- [0102] 可选的, 基于音视频联动的监控装置3还可以包括:
- [0103] 音频获取单元, 用于获取所述监控区域的监控音频;
- [0104] 第二热点分析单元, 用于对所述监控音频进行热点分析, 以判断所述监控区域中是否存在第二监控热点;
- [0105] 第二方位计算单元, 用于若所述监控区域中存在所述第二监控热点, 则计算所述第二监控热点在所述监控区域中的第二方位;
- [0106] 视频监控单元34还用于, 将所述第二方位所在的方向确定为视频监控的中心方向进行视频监控;
- [0107] 音频监控单元35还用于, 将所述第二方位所在的方向确定为音频监控的中心方向进行音频监控。
- [0108] 可选的, 基于音视频联动的监控装置3还可以包括:
- [0109] 强度判断单元, 用于判断所述监控音频中是否存在声音强度大于预设强度阈值的目标声源;
- [0110] 第二热点分析单元具体用于, 若强度判断单元判定存在所述目标声源, 则判定所述监控区域中存在第二监控热点; 以及, 若强度判断单元判定不存在所述目

标声源，则判定所述监控区域中不存在第二监控热点。

[0111] 可选的，音频获取单元具体用于，利用麦克风阵列获取所述监控区域的监控音频；

[0112] 相应的，第二方位计算单元具体用于，根据所述麦克风阵列中各麦克风的位置排布，以及，所述麦克风阵列中各麦克风采集到所述目标声源发出的声音的时间差，计算所述目标声源在所述第二监控热点在所述监控区域中的第二方位。

[0113] 可选的，基于音视频联动的监控装置3还可以包括：

[0114] 热点度计算单元，用于若所述监控区域中同时存在所述第一监控热点与所述第二监控热点，且，所述第一监控热点与所述第二监控热点对应的第一方位和第二方位不一致，则分别计算所述第一监控热点和第二监控热点的热点度得分。

[0115] 目标方位确定单元，用于将热点度得分较高的监控热点对应的方位确定为目标方位；

[0116] 相应的，视频监控单元34还用于，将所述目标方位所在的方向确定为视频监控的中心方向进行视频监控；音频监控单元35还用于，将所述目标方位所在的方向确定为音频监控的中心方向进行音频监控。

[0117] 综上所述，本申请通过获取监控区域的监控图像；对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；也即，通过对监控图像进行热点分析来发现监控区域的监控热点，并通过调整监控视野的中心方向使监控热点处于监控视野中心，如此一来在进行较大范围的视频监控时可以无需额外布置多个监控设备，能够降低视频监控设备的硬件成本；另一方面，通过将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控，可以将音频监控设备的声音采集方向对准监控热点方向，从而在进行音频监控时有利于提高声音质量。

[0118] 所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的装置、模块和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

- [0119] 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。
- [0120] 图4示出了本申请实施例提供的一种终端设备的示意框图，为了便于说明，仅示出了与本申请实施例相关的部分。
- [0121] 在本实施例中，所述终端设备4可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。该终端设备4可包括：处理器40、存储器41以及存储在所述存储器41中并可在所述处理器40上运行的计算机可读指令42，例如执行上述的基于音视频联动的监控方法的计算机可读指令。所述处理器40执行所述计算机可读指令42时实现上述各个基于音视频联动的监控方法实施例中的步骤，例如图1所示的步骤S101至步骤S105。或者，所述处理器40执行所述计算机可读指令42时实现上述各装置实施例中各单元的功能，例如图3所示单元31至35的功能。
- [0122] 示例性的，所述计算机可读指令42可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器41中，并由所述处理器40执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机可读指令段，该指令段用于描述所述计算机可读指令42在所述终端设备4中的执行过程。
- [0123] 所述处理器40可以是中央处理单元（Central Processing Unit，CPU），还可以是其它通用处理器、数字信号处理器（Digital Signal Processor，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现场可编程门阵列（Field-Programmable Gate Array，FPGA）或者其它可编程逻辑器件、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该处理器也可以是任何常规的处理器等。
- [0124] 所述存储器41可以是所述终端设备4的内部存储单元，例如终端设备4的硬盘或内存。所述存储器41也可以是所述终端设备4的外部存储设备，例如所述终端设备4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，所述存储器41还可以既包括所述终端设备4的内部存储单元也包括外部存储设备。所述存储器

41用于存储所述计算机可读指令以及所述终端设备4所需的其它指令和数据。所述存储器41还可以用于暂时地存储已经输出或者将要输出的数据。

[0125] 在本申请各个实施例中的各功能单元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。

[0126] 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取存储介质中，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。

[0127] 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。

权利要求书

- [权利要求 1] 一种基于音视频联动的监控方法，其特征在于，包括：
- 获取监控区域的监控图像；
- 对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
- 若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
- 将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；
- 将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [权利要求 2] 根据权利要求1所述的基于音视频联动的监控方法，其特征在于，所述对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点包括：
- 判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域；
- 若存在所述图像区域，则判定所述监控区域中存在第一监控热点；
- 若不存在所述图像区域，则判定所述监控区域中不存在第一监控热点。
- [权利要求 3] 根据权利要求2所述的基于音视频联动的监控方法，其特征在于，所述若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位，包括：
- 根据所述图像区域的中心像素点在所述监控图像上的位置，得到第一坐标；
- 根据预设的坐标方位对应关系，确定所述第一坐标在所述监控区域中的第一方位。
- [权利要求 4] 根据权利要求1至3任一项所述的基于音视频联动的监控方法，其特征在于，所述监控方法还包括：

获取所述监控区域的监控音频；

对所述监控音频进行热点分析，以判断所述监控区域中是否存在第二监控热点；

若所述监控区域中存在所述第二监控热点，则计算所述第二监控热点在所述监控区域中的第二方位；

将所述第二方位所在的方向确定为视频监控的中心方向进行视频监控；

将所述第二方位所在的方向确定为音频监控的中心方向进行音频监控。

[权利要求 5]

根据权利要求4所述的基于音视频联动的监控方法，其特征在于，所述对所述监控音频进行热点分析，以判断所述监控区域中是否存在第二监控热点，包括：

判断所述监控音频中是否存在声音强度大于预设强度阈值的目标声源；

若存在所述目标声源，则判定所述监控区域中存在第二监控热点；

若不存在所述目标声源，则判定所述监控区域中不存在第二监控热点。

[权利要求 6]

根据权利要求5所述的基于音视频联动的监控方法，其特征在于，所述获取所述监控区域的监控音频包括：

利用麦克风阵列获取所述监控区域的监控音频；

相应的，所述若所述监控区域中存在所述第二监控热点，则计算所述第二监控热点在所述监控区域中的第二方位包括：

根据所述麦克风阵列中各麦克风的位置排布，以及，所述麦克风阵列中各麦克风采集到所述目标声源发出的声音的时间差，计算所述目标声源在所述第二监控热点在所述监控区域中的第二方位。

[权利要求 7]

根据权利要求4所述的基于音视频联动的监控方法，其特征在于，所述监控方法还包括：

若所述监控区域中同时存在所述第一监控热点与所述第二监控热点，

且，所述第一监控热点与所述第二监控热点对应的第一方位和第二方位不一致，则分别计算所述第一监控热点和第二监控热点的热点度得分；

将热点度得分较高的监控热点对应的方位确定为目标方位；

将所述目标方位所在的方向确定为视频监控的中心方向进行视频监控；

将所述目标方位所在的方向确定为音频监控的中心方向进行音频监控。

[权利要求 8]

一种基于音视频联动的监控装置，其特征在于，包括：

图像获取单元，用于获取监控区域的监控图像；

第一热点分析单元，用于对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；

第一方位计算单元，用于若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；

视频监控单元，用于将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；

音频监控单元，用于将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。

[权利要求 9]

如权利要求8所述的基于音视频联动的监控装置，其特征在于，第一热点分析单元，包括：

密度判断单元，用于判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域；

判断分析单元，用于若密度判断单元判定存在所述图像区域，则判定所述监控区域中存在第一监控热点；以及，若密度判断单元判定不存在所述图像区域，则判定所述监控区域中不存在第一监控热点。

[权利要求 10]

如权利要求9所述的基于音视频联动的监控装置，其特征在于，第一方位计算单元，包括：

坐标获取单元，用于根据所述图像区域的中心像素点在所述监控图像

上的位置，得到第一坐标；

坐标计算单元，用于根据预设的坐标方位对应关系，确定所述第一坐标在所述监控区域中的第一方位。

[权利要求 11] 如权利要求8至10任意一项所述的基于音视频联动的监控装置，其特征在于，还包括：

音频获取单元，用于获取所述监控区域的监控音频；

第二热点分析单元，用于对所述监控音频进行热点分析，以判断所述监控区域中是否存在第二监控热点；

第二方位计算单元，用于若所述监控区域中存在所述第二监控热点，则计算所述第二监控热点在所述监控区域中的第二方位；

所述视频监控单元，还用于将所述第二方位所在的方向确定为视频监控的中心方向进行视频监控；

所述音频监控单元，还用于将所述第二方位所在的方向确定为音频监控的中心方向进行音频监控。

[权利要求 12] 如权利要求11所述的基于音视频联动的监控装置，其特征在于，所述第二热点分析单元，包括：

强度判断单元，用于判断所述监控音频中是否存在声音强度大于预设强度阈值的目标声源；

强度分析单元，用于若强度判断单元判定存在所述目标声源，则判定所述监控区域中存在第二监控热点；以及，若强度判断单元判定不存在所述目标声源，则判定所述监控区域中不存在第二监控热点。

[权利要求 13] 如权利要求12所述的基于音视频联动的监控装置，其特征在于，音频获取单元，具体用于：

利用麦克风阵列获取所述监控区域的监控音频；

相应的，所述第二方位计算单元具体用于，根据所述麦克风阵列中各麦克风的位置排布，以及，所述麦克风阵列中各麦克风采集到所述目标声源发出的声音的时间差，计算所述目标声源在所述第二监控热点在所述监控区域中的第二方位。

- [权利要求 14] 如权利要求11所述的基于音视频联动的监控装置，其特征在于，
热点度计算单元，用于若所述监控区域中同时存在所述第一监控热点与所述第二监控热点，且，所述第一监控热点与所述第二监控热点对应的第一方位和第二方位不一致，则分别计算所述第一监控热点和第二监控热点的热点度得分；
目标方位确定单元，用于将热点度得分较高的监控热点对应的方位确定为目标方位；
相应的，视频监控单元还用于，将所述目标方位所在的方向确定为视频监控的中心方向进行视频监控；音频监控单元还用于，将所述目标方位所在的方向确定为音频监控的中心方向进行音频监控。
- [权利要求 15] 一种终端设备，其特征在于，所述终端设备包括存储器、处理器，所述存储器上存储有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：
获取监控区域的监控图像；
对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；
若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；
将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；
将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。
- [权利要求 16] 根据权利要求15所述的终端设备，其特征在于，所述对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点包括：
判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域；
若存在所述图像区域，则判定所述监控区域中存在第一监控热点；
若不存在所述图像区域，则判定所述监控区域中不存在第一监控热点。

。

[权利要求 17] 根据权利要求16所述的终端设备，其特征在于，所述若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位，包括：

根据所述图像区域的中心像素点在所述监控图像上的位置，得到第一坐标；

根据预设的坐标方位对应关系，确定所述第一坐标在所述监控区域中的第一方位。

[权利要求 18] 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可读指令，其特征在于，所述计算机可读指令被至少一个处理器执行时实现如下步骤：

获取监控区域的监控图像；

对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点；

若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位；

将所述第一方位所在的方向确定为视频监控的中心方向进行视频监控；

将所述第一方位所在的方向确定为音频监控的中心方向进行音频监控。

[权利要求 19] 根据权利要求18所述的计算机可读存储介质，其特征在于，所述对所述监控图像进行热点分析，以判断所述监控区域中是否存在第一监控热点包括：

判断所述监控图像上是否存在人群密度大于预设密度阈值的图像区域；

若存在所述图像区域，则判定所述监控区域中存在第一监控热点；

若不存在所述图像区域，则判定所述监控区域中不存在第一监控热点。

- [权利要求 20] 根据权利要求19所述的计算机可读存储介质，其特征在于，所述若所述监控区域中存在所述第一监控热点，则计算所述第一监控热点在所述监控区域中的第一方位，包括：
- 根据所述图像区域的中心像素点在所述监控图像上的位置，得到第一坐标；
- 根据预设的坐标方位对应关系，确定所述第一坐标在所述监控区域中的第一方位。

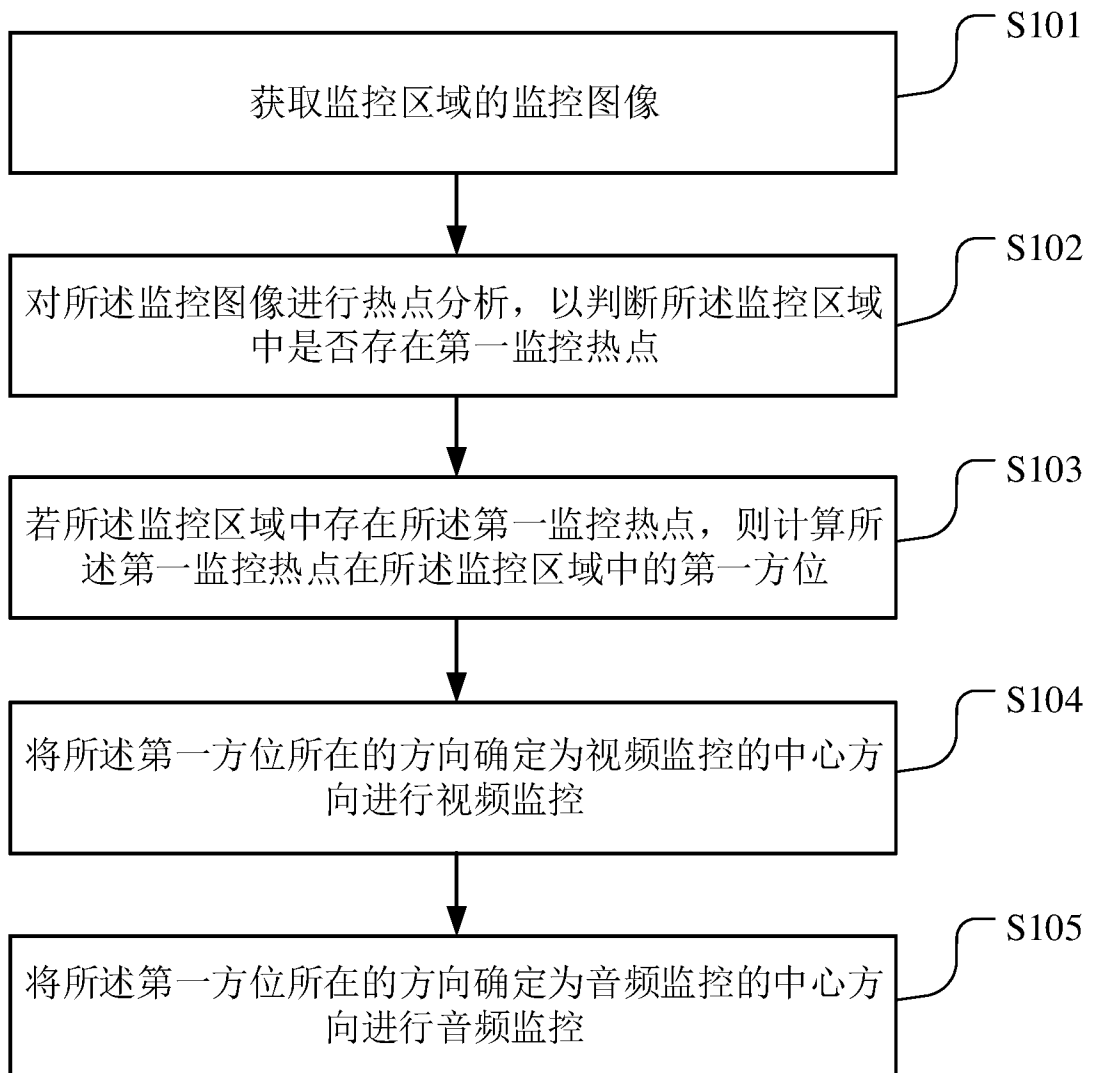


图 1

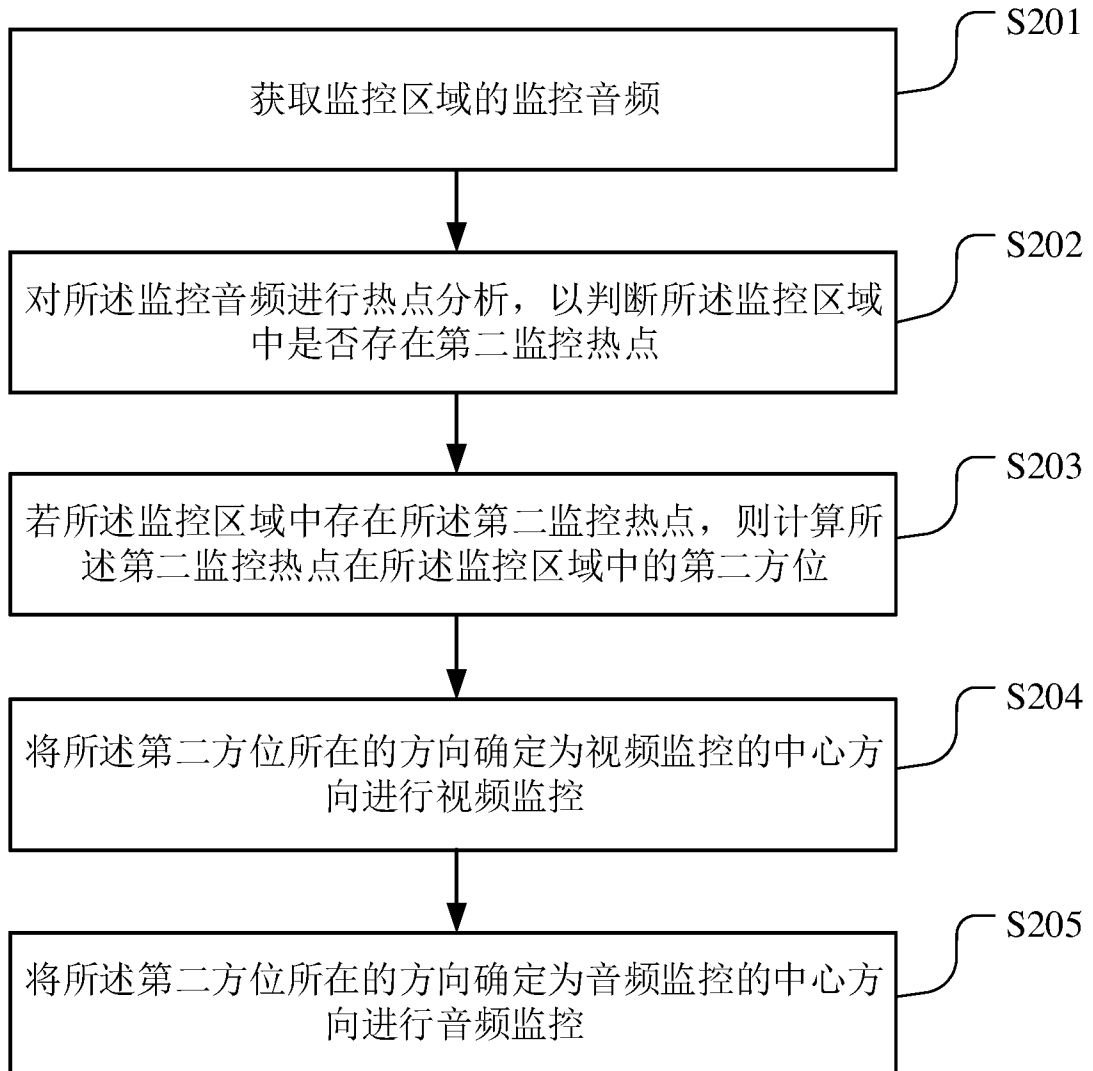


图 2

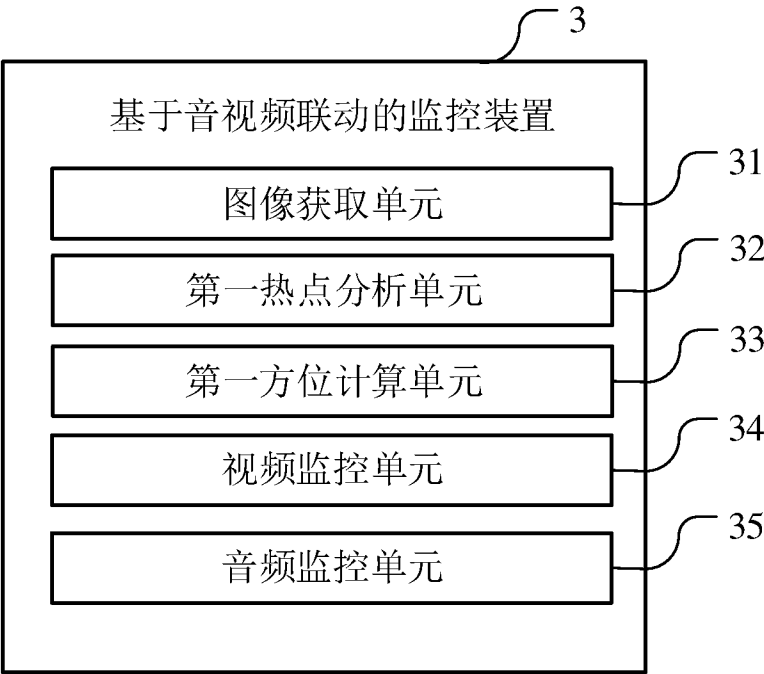


图 3

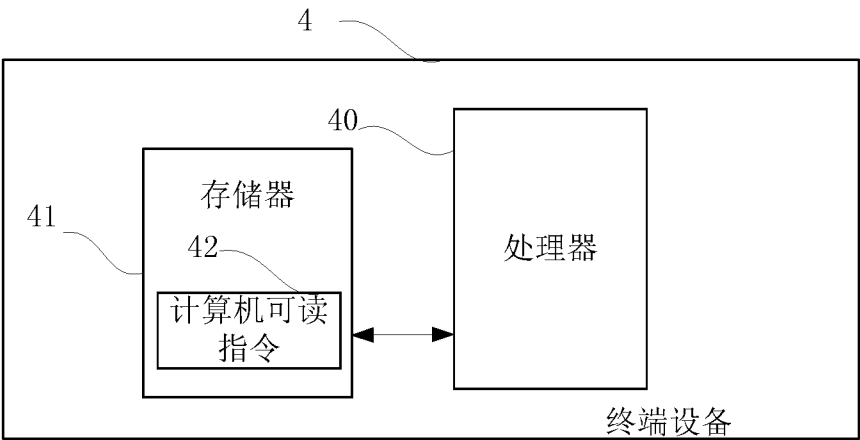


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/118158

A. CLASSIFICATION OF SUBJECT MATTER

H04N 7/18(2006.01)i; G05D 3/12(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)

H04N; G05D

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: 监控, 视频, 音频, 摄像机, 摄像头, 图像, 目标, 方位, 位置, 调整, 方向, 跟踪, monitor+, video, volume, camera, target, locat+, adjus+, mov+, rotat+, trac+, direction

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109922311 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 21 June 2019 (2019-06-21) claims 1-10, and description, paragraphs [0027]-[0112]	1-20
X	CN 106303211 A (XIAOMI TECHNOLOGY CO., LTD.) 04 January 2017 (2017-01-04) description, paragraphs [0139]-[0408]	1-20
X	CN 102833476 A (GOERTEK INC.) 19 December 2012 (2012-12-19) description, paragraphs [0075]-[0113]	1-20
A	CN 106341665 A (ZHEJIANG UNIVIEW TECHNOLOGIES CO., LTD.) 18 January 2017 (2017-01-18) entire document	1-20
A	CN 101511000 A (SUN YAT-SEN UNIVERSITY) 19 August 2009 (2009-08-19) entire document	1-20
A	US 2008252722 A1 (WANG, Y.K. et al.) 16 October 2008 (2008-10-16) 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

16 January 2020

Date of mailing of the international search report

27 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
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/CN2019/118158

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109922311	A	21 June 2019	None	
CN	106303211	A	04 January 2017	None	
CN	102833476	A	19 December 2012	None	
CN	106341665	A	18 January 2017	None	
CN	101511000	A	19 August 2009	None	
US	2008252722	A1	16 October 2008	None	

国际检索报告

国际申请号

PCT/CN2019/118158

A. 主题的分类 H04N 7/18(2006.01)i; G05D 3/12(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) H04N; G05D 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:监控, 视频, 音频, 摄像机, 摄像头, 图像, 目标, 方位, 位置, 调整, 方向, 跟踪, monitor+, video, volume, camera, target, locat+, adjust+, mov+, rotat+, trac+, direction		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 109922311 A (平安科技深圳有限公司) 2019年 6月 21日 (2019 - 06 - 21) 权利要求1-10, 说明书第[0027]-[0112]段	1-20
X	CN 106303211 A (小米科技有限责任公司) 2017年 1月 4日 (2017 - 01 - 04) 说明书第[0139]-[0408]段	1-20
X	CN 102833476 A (歌尔声学股份有限公司) 2012年 12月 19日 (2012 - 12 - 19) 说明书第[0075]-[0113]段	1-20
A	CN 106341665 A (浙江宇视科技有限公司) 2017年 1月 18日 (2017 - 01 - 18) 全文	1-20
A	CN 101511000 A (中山大学) 2009年 8月 19日 (2009 - 08 - 19) 全文	1-20
A	US 2008252722 A1 (WANG, Yuan-Kai等) 2008年 10月 16日 (2008 - 10 - 16) 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 1月 16日		国际检索报告邮寄日期 2020年 2月 27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 全红红 电话号码 86-(10)-53961595

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/118158

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	109922311	A	2019年 6月 21日	无	
CN	106303211	A	2017年 1月 4日	无	
CN	102833476	A	2012年 12月 19日	无	
CN	106341665	A	2017年 1月 18日	无	
CN	101511000	A	2009年 8月 19日	无	
US	2008252722	A1	2008年 10月 16日	无	