

(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
2018 年 11 月 22 日 (22.11.2018)



(10) 国际公布号
WO 2018/209879 A1

(51) 国际专利分类号:
H04N 7/15 (2006.01) H04N 5/268 (2006.01)

(21) 国际申请号: PCT/CN2017/104657

(22) 国际申请日: 2017 年 9 月 29 日 (29.09.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710344454.4 2017 年 5 月 16 日 (16.05.2017) CN

(71) 申请人: 广州视源电子科技有限公司 (GUANGZHOU SHIYUAN ELECTRONICS CO., LTD.) [CN/CN]; 中国广东省广州市黄埔区云埔四路 6 号, Guangdong 510530 (CN)。广州视臻信息科技有限公司 (GUANGZHOU SHIZHEN ELECTRONICS CO., LTD.) [CN/CN]; 中

国广东省广州市经济技术开发区科珠路 192 号, Guangdong 510663 (CN)。

(72) 发明人: 陈双龙 (CHEN, Shuanglong); 中国广东省广州市经济技术开发区科珠路 192 号, Guangdong 510663 (CN)。

(74) 代理人: 广州三环专利商标代理有限公司 (SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路 80 号汇华商贸大厦 1508 室, Guangdong 510070 (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: METHOD AND DEVICE FOR AUTOMATICALLY SELECTING CAMERA IMAGE, AND AUDIO AND VIDEO SYSTEM

(54) 发明名称: 自动选择摄像头画面的方法、装置及音视频系统

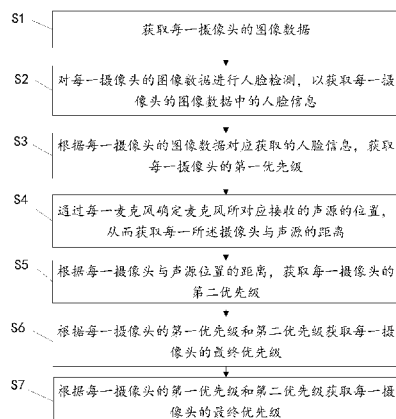


图 2

S1 OBTAIN IMAGE DATA OF EACH CAMERA
S2 PERFORM FACE DETECTION ON THE IMAGE DATA OF EACH CAMERA TO OBTAIN FACE INFORMATION IN THE IMAGE DATA OF EACH CAMERA
S3 OBTAIN A FIRST PRIORITY OF EACH CAMERA ACCORDING TO THE OBTAINED FACE INFORMATION OF THE IMAGE DATA OF EACH CAMERA
S4 DETERMINE THE LOCATION OF A SOUND SOURCE RECEIVED BY A MICROPHONE BY MEANS OF EACH CAMERA TO OBTAIN THE DISTANCE BETWEEN EACH CAMERA AND THE SOUND SOURCE
S5 OBTAIN A SECOND PRIORITY OF EACH CAMERA ACCORDING TO THE DISTANCE BETWEEN EACH CAMERA AND THE SOUND SOURCE
S6 OBTAIN A FINAL PRIORITY OF EACH CAMERA ACCORDING TO THE FIRST PRIORITY AND THE SECOND PRIORITY OF EACH CAMERA
S7 OUTPUT THE IMAGE DATA OF THE CAMERA HAVING THE HIGHEST FINAL PRIORITY FOR DISPLAY

(57) Abstract: Disclosed is a method for automatically selecting a camera image, comprising: obtaining image data of each camera; performing face detection on the image data of each camera to obtain face information in the image data of each camera; obtaining a first priority of each camera according to the face information; obtaining the location of a sound source by means of each microphone to obtain the distance between each camera and the sound source; obtaining a second priority of each camera according to the distance between each camera and the sound source; obtaining a final priority according to the first priority and the second priority of each camera; and outputting the image data of the camera having the highest final priority for display. Correspondingly, the present invention further provides a device for automatically selecting a camera image, an audio and video system and a smart tablet. According to the present invention, during audio and video conferences, the images photographed by the cameras are automatically selected for display, and thus the user experience is improved.

(57) 摘要: 本发明公开了一种自动选择摄像头画面的方法, 包括: 获取每一摄像头的图像数据; 对所述每一摄像头的图像数据进行人脸检测, 以获取所述每一摄像头的图像数据中的人脸信息; 根据人脸信息获取每一所述摄像头的的第一优先级; 通过每一麦克风获取声源的位置, 从而获取每一所述摄像头与声源的距离; 根据所述每一摄像头与声源的距离, 获取每一所述摄像头的第二优先级; 根据所述每一所述摄像头的的第一优先级和第二优先级获取最终优先级; 将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。相应的, 本发明还提供自动选择摄像头画面的装置、音视频系统和智能平板。本发明在音视频会议时, 实现自动选择摄像头的拍摄的画面进行显示, 提升用户体验。

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))。

自动选择摄像头画面的方法、装置及音视频系统

技术领域

本发明涉及通信技术领域，尤其涉及一种自动选择摄像头画面的方法和装置、音视频系统及智能平板。

背景技术

随着科技的发展，大尺寸智能平板（例如 50 寸以上）已经越来越多的应用于教育、会议等领域。在会议领域，智能平板一些品牌存在多个摄像头的配置，如图 1 所示，在智能平板的左侧和右侧各安装有 1 个摄像头 201。在进行远程会议时，通常都是固定（左边或者右边）的摄像头或者手动选择其中一个摄像头来获取图像数据，从而进行远程的音视频会议。

尤其，在多人多方远程音视频会议中，会议中的与会人员可能在会议室的不同的位置，如果单纯选择单个摄像头并不能获取所有与会人员的图像数据。通常的作法，人们根据不同的会议情形来手动切换摄像头以选择对应的会议场景，来满足人们对不同会议显示画面的需求。现有的这种切换摄像头的方式，需要人工干预，用户体验较差。

发明内容

本发明实施例的目的是提供一种自动选择摄像头画面的方法和装置、以及音视频系统，在音视频会议中，能够自动选择摄像头画面，使用户更好地了解会议情况，提升用户体验。

为实现上述目的，本发明实施例提供一种自动选择摄像头画面的方法，包括：

获取每一摄像头的图像数据；其中，获取的所述图像数据为所述每一摄像头

的实时拍摄的画面图像；

对所述每一摄像头的图像数据进行人脸检测，以获取所述每一摄像头的图像数据中的人脸信息；

根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头的第二优先级；

通过每一麦克风确定所述麦克风所对应接收的声源的位置，从而获取每一所述摄像头与所述声源的距离；

根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；

根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级；

将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。

与现有技术相比，本发明实施例提供一种自动选择摄像头画面的方法通过首先获取每一摄像头的图像数据；然后对图像数据进行人脸检测以获取人脸信息，根据所述人脸信息来获取每一所述摄像头的第二优先级；接着，通过每一麦克风来确定声源的位置，根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；最后根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级，将所述最终优先级最高的所述摄像头的图像数据输出以用于显示的技术方案，实现了在远程视频会议中，针对视频会议的需求，基于检测人脸和获取语音位置来自动选择摄像头拍摄的图像进行显示，实现智能实时自动切换摄像头画面，使用户更好地了解会议情况，提升用户体验。

优选的，所述根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级具体包括：

对于每一所述摄像头，将所述第一优先级乘以预置第一权重得到第一乘积值，将所述第二优先级乘以预置第二权重得到第二乘积值；

根据所述第一乘积值和所述第二乘积值的和值得到最终优先级。

作为本发明实施例的优选方案，采用了对第一优先权和的第二优先权分别乘以预置权重值再进行求和的方式来获取最终优先级，能够实现同时考虑人脸检测的识别结果和声源定位结果，同时，该优选方案能够通过调整第一权重和第二权重，来适应用户对会议显示画面的需求，可适应性强，用户体验高。

进一步的，所述根据所述每一摄像头与所述声源位置距离，获取每一所述摄像头的第二优先级包括：

根据所述每一摄像头与所述声源位置距离的大小，对所述每一摄像头进行排序，与所述声源的距离越小的所述摄像头对应获取的第二优先级越高。

作为本发明实施例的进一步方案，设置离声源越近的摄像头的第二优先级越高，使得远程音频时，较多对发言人发言的画面进行显示。

进一步的，所述人脸信息包括人脸个数、人脸面积以及人脸在图像中的位置。

进一步的，所述根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头图像数据的第一优先级具体包括：

根据所述每一摄像头的图像数据对应获取的所述人脸个数、所述人脸面积以及所述人脸在图像中的位置分别获取对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值；

根据所述每一摄像头的图像数据的所述人脸个数分值、所述人脸面积分值和所述人脸在图像中的位置分值的和值的大小，获取所述每一摄像头的所述第一优先级。

本发明实施例还提供一种自动选择摄像头画面的装置，包括：

图像数据获取单元，用于获取每一摄像头的图像数据；其中，获取的所述图像数据为所述每一摄像头的实时拍摄的画面图像；

人脸信息获取单元，用于对所述每一摄像头的图像数据进行人脸检测，以获取所述每一摄像头的图像数据中的人脸信息；

第一优先级获取单元，用于根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头的的第一优先级；

声源距离获取单元，用于通过每一麦克风确定所述麦克风所对应接收的声源的位置，从而获取每一所述摄像头与所述声源的距离；

第二优先级获取单元，用于根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；

最终优先级获取单元，用于根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级；

输出单元，用于将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。

与现有技术相比，本发明实施例提供一种自动选择摄像头画面的装置，首先通过图像数据获取单元获取每一摄像头的图像数据；然后通过人脸信息获取单元对图像数据进行人脸检测以获取人脸信息，并通过第一优先级获取单元根据所述人脸信息来获取每一所述摄像头的的第一优先级；接着通过声源距离获取单元每一麦克风来确定声源的位置，并通过第二优先级获取单元根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；最后通过最终优先级获取单元根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级，并由输出单元将所述最终优先级最高的所述摄像头的图像数据输出以用于显示的技术方案，实现了在远程音视频会议中，针对视频会议的需求，基于检测人脸和获取语音位置来自动选择摄像头拍摄的图像进行显示，

实现智能实时自动切换摄像头画面,使用户更好地了解会议情况,提升用户体验。

进一步的,所述最终优先级获取单元具体用于:

对于每一所述摄像头,将所述第一优先级乘以预置第一权重得到第一乘积值,将所述第二优先级乘以预置第二权重得到第二乘积值;

根据所述第一乘积值和所述第二乘积值的和值得到最终优先级。

进一步的,所述第二优先级获取单元用于根据所述每一摄像头与所述声源位置距离,获取每一所述摄像头的第二优先级时,根据所述每一摄像头与所述声源位置距离的大小,对所述每一摄像头进行排序,与所述声源的距离越小的所述摄像头对应获取的第二优先级越高。

相应的,本发明实施例还提供一种音视频系统,包括本发明实施例提供的一种自动选择摄像头画面的装置,还包括:

至少2个摄像头,所述摄像头分别安装于智能平板的左右两侧或上下两侧,用于实时拍摄画面图像;

麦克风,用于接收声源,并确定接收的所述声源的位置。

与现有技术相比,本发明实施例提供一种音视频系统,一方面通过所述摄像头实时拍摄画面图像,通过所述麦克风接收声源,并确定声源的位置;另一方面通过本发明实施例公开的自动选择摄像头画面的装置,获取所述摄像头实时拍摄的图像数据,对图像数据进行人脸检测获取人脸信息以获取第一优先级,基于所述麦克风所确定的声源位置与摄像头的距离来获取第二优先级,根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级,将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。上述技术方案,实现了在远程视频会议中,针对视频会议的需求,基于检测人脸和获取语音位置来自动选择摄像头拍摄的图像进行显示,实现智能实时自动切换摄像头画面,使用户更好地了解会议情况,提升用户体验。

相应的，本发明实施例还提供一种智能平板，包括本发明实施例提供的一种音视频系统。

与现有技术相比，本发明实施例提供的一种智能平板，一方面通过所述摄像头实时拍摄画面图像，通过所述麦克风接收声源，并确定声源的位置；另一方面通过本发明实施例公开的自动选择摄像头画面的装置，获取所述摄像头实时拍摄的图像数据，对图像数据进行人脸检测获取人脸信息以获取第一优先级，基于所述麦克风所确定的声源位置与摄像头的距离来获取第二优先级，根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级，将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。上述技术方案，实现了在远程视频会议中，针对视频会议的需求，基于检测人脸和获取语音位置来自动选择摄像头拍摄的图像进行显示，实现智能实时自动切换摄像头画面，使用户更好地了解会议情况，提升用户体验。

附图说明

图 1 是一种配置有两个摄像头的大尺寸智能平板的结构示意图；

图 2 是本发明实施例一提供的一种自动选择摄像头画面的方法的流程示意图；

图 3 是本发明实施例一提供的一种自动选择摄像头画面的方法的步骤 S3 的流程示意图；

图 4 是本发明实施例二提供的一种自动选择摄像头画面的装置的结构示意图；

图 5 是本发明实施例三提供的一种音视频系统的结构示意图；

图 6 是本发明实施例四提供的一种智能平板的结构示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明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发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

参见图 2，图 2 是本发明实施例一提供的一种自动选择摄像头画面的方法的流程示意图，包括：

S1、获取每一摄像头的图像数据；其中，获取的图像数据为每一摄像头的实时拍摄的画面图像；

S2、对每一摄像头的图像数据进行人脸检测，以获取每一摄像头的图像数据中的人脸信息；

S3、根据每一摄像头的图像数据对应获取的人脸信息，获取每一摄像头的第一优先级；

S4、通过每一麦克风确定麦克风所对应接收的声源的位置，从而获取每一摄像头与声源的距离；

S5、根据每一摄像头与声源位置的距离，获取每一摄像头的第二优先级；

S6、根据每一摄像头的第一优先级和第二优先级获取每一摄像头的最终优先级；

S7、将最终优先级最高的摄像头的图像数据输出以用于显示。

其中，步骤 S1 中获取的摄像头实时拍摄的图像数据，会与每一摄像头一一对应。

进一步的，步骤 S6 获取每一摄像头的最终优先级的具体计算过程为：

对于每一摄像头，将第一优先级乘以预置第一权重得到第一乘积值，将第二

优先级乘以预置第二权重得到第二乘积值；

根据第一乘积值和第二乘积值的和值得到最终优先级。

本实施例的步骤 S6 对基于人脸检测获取的第一优先级以及基于声源位置识别获取的第二优先级采取分别乘以对应的权重后相加的方式，能够通过调整预置第一权重和第二权重的来调整第一优先级和第二优先级各占的比重，满足远程会议中的实际需求。另外，对于确定最终优先级的实施方式除上述采用比例值融合相加的优选实施方式，在实际生活中，还可以根据需求，只侧重考虑声源位置所确定的第二优先级作为最终优先级的首要考虑因素，而基于人脸检测所获取的第一优先级作为次要考虑因素，即实现显示画面能够实时显示发言人的画面，当无人发言时，基于人脸检测的结果选择显示的画面。上述实施方式也在本发明的保护范围之内。

步骤 S5 具体包括：

根据每一摄像头与声源位置距离的大小，对每一摄像头进行排序，与声源的距离越小的摄像头对应获取的第二优先级越高。

通过步骤 S5 使离声源越近的摄像头对应获取的第二优先级越高，从而使得离声源越近的摄像头的画面被选择作为显示画面显示的可能性越高。

其中，在步骤 S2 中摄像头的图像进行人脸检测时，通过确定人脸外围矩形框来表示人脸，从而基于人脸外围矩形框来确定人脸个数、人脸面积以及人脸在图像中的位置。这里涉及的具体的的人脸检测方法，为本领域技术人员能从现有技术中获取，所以此处不再做赘述。

步骤 S3 中获取每一摄像头的图像数据中的人脸信息包括获取每一摄像头的图像数据中的人脸个数、人脸面积以及人脸在图像中的位置。对应的，参见图 3，图 3 是本发明实施例一提供的一种自动选择摄像头画面的方法的步骤 S3 的流程示意图，步骤 S3 具体包括以下步骤：

S31、根据每一摄像头的图像数据对应获取的人脸个数、人脸面积以及人脸在图像中的位置分别获取对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值；

优选的，人脸在图像中的位置设定为人脸的中心位置与当前图像的中心的距离，距离越近，则说明人脸越靠近摄像头的图像中心。通常，为了使显示画面能更好展现会议场面，则对应设定：人脸个数越多，对应的人脸个数分值越高；人脸面积越大，则人脸面积分值越高；人脸在图像中的位置越靠近摄像头图像中心，则人脸在图像中的位置分值越高。

进一步优选，为满足实际显示会议画面中对人脸个数、人脸面积以及人脸在图像中的位置的不同考虑，还可以具体通过在根据人脸个数、所述人脸面积以及人脸在图像中的位置确定对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值时，设置对应的权重值来分别调整人脸个数分值、人脸面积分值以及人脸在图像中的位置分值的分值大小，进而影响最终选择显示的摄像头的画面。

S32、根据每一摄像头的图像数据的人脸个数分值、人脸面积分值和人脸在图像中的位置分值的和值的大小，获取每一摄像头的的第一优先级。

具体实施时，本实施例首先获取摄像头实时拍摄的图像数据，然后对每一摄像头拍摄的图像数据进行人脸检测，包括检测每一图像中的人脸个数、人脸面积和人脸在图像中的位置，进而获取人脸个数分值、人脸面积分值和人脸在图像中的位置分值；将每一摄像头的图像获得的上述三个分值的和值作为该摄像头对应的第一优先级；接着，获取麦克风所接收声源的位置，进而获取声源位置与每一摄像头的距离，根据每一摄像头与声源位置的距离来获取对应的第二优先级，其中，与声源位置的距离越小，摄像头的第二优先级越高；最后，对于每一摄像头，将第一优先级和第二优先级分别乘以对应的权重值然后求和得到最终优先级，将最终优先级最高的摄像头的图像数据传输到显示画面。

本实施例先通过人脸信息来获取第一优先级，然后通过声源位置与摄像头距离来获取第二优先级的顺序只是一种实施示例，调换先后进行的顺序，或并行进行上述两个步骤的实施例都在本实施例的保护范围之内。

与现有技术相比，本实施例基于人脸检测和声源位置的结果来选取所显示的摄像头的拍摄的画面，实现智能自主切换摄像头画面，满足远程音视频会议的显示需求，减少人工操作，更加智能化、自动化；提高用户体验。

参见图 4，图 4 是本发明实施例二提供的一种自动选择摄像头画面的装置的结构示意图，本实施例具体包括以下结构：

图像数据获取单元 11，用于获取每一摄像头的图像数据；其中，获取的图像数据为每一摄像头的实时拍摄的画面图像；

人脸信息获取单元 12，用于对每一摄像头的图像数据进行人脸检测，以获取每一摄像头的图像数据中的人脸信息；

第一优先级获取单元 13，用于根据每一摄像头的图像数据对应获取的人脸信息，获取每一摄像头的第一优先级；

声源距离获取单元 14，用于通过每一麦克风确定麦克风所对应接收的声源的位置，从而获取每一摄像头与声源的距离；

第二优先级获取单元 15，用于根据每一摄像头与声源位置的距离，获取每一摄像头的第二优先级；

最终优先级获取单元 16，用于根据每一摄像头的第一优先级和第二优先级获取每一摄像头的最终优先级；

输出单元 17，用于将最终优先级最高的摄像头的图像数据输出以用于显示。

具体的，最终优先级获取单元 16 用于：

对于每一摄像头，将第一优先级乘以预置第一权重得到第一乘积值，将第二优先级乘以预置第二权重得到第二乘积值；

根据第一乘积值和第二乘积值的和值得到最终优先级。

本实施例的最终优先级获取单元 16 对基于人脸检测获取的第一优先级以及基于声源位置识别获取的第二优先级采取分别乘以对应的权重后相加的方式，能够通过调整预置第一权重和第二权重的来调整第一优先级和第二优先级各占的比重，满足远程会议中的实际需求。另外，对于确定最终优先级的实施方式除上述采用比例值融合相加的优选实施方式，在实际生活中，还可以根据需求，只侧重考虑声源位置所确定的第二优先级作为最终优先级的首要考虑因素，而基于人脸检测所获取的第一优先级作为次要考虑因素，即实现显示画面能够实时显示发言人的画面，当无人发言时，基于人脸检测的结果选择显示的画面。上述实施方式也在本发明的保护范围之内。

第二优先级获取单元 15 用于根据每一摄像头与声源位置距离，获取每一摄像头的第二优先级时，根据每一摄像头与声源位置距离的大小，对每一摄像头进行排序，与声源的距离越小的摄像头对应获取的第二优先级越高。

通过第二优先级获取单元 15 使离声源越近的摄像头对应获取的第二优先级越高，从而使得离声源越近的摄像头的画面被选择作为显示画面显示的可能性越高。

其中，在人脸信息获取单元 12 用于对摄像头的图像进行人脸检测时，通过确定人脸外围矩形框来表示人脸，从而基于人脸外围矩形框来确定人脸个数、人脸面积以及人脸在图像中的位置。这里涉及的具体的的人脸检测方法，为本领域技术人员能从现有技术中获取，所以此处不再做赘述。

人脸信息获取单元 12 获取的每一摄像头的图像数据中的人脸信息包括人脸个数、人脸面积以及人脸在图像中的位置。

第一优先级获取单元 13 具体用于：

根据每一摄像头的图像数据对应获取的人脸个数、人脸面积以及人脸在图像中的位置分别获取对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值；

优选的，人脸在图像中的位置设定为人脸的中心位置与当前图像的中心的位置的距离，距离越近，则说明人脸越靠近摄像头的图像中心。通常，为了使显示画面能更好展现会议场面，则对应设定：人脸个数越多，对应的人脸个数分值越高；人脸面积越大，则人脸面积分值越高；人脸在图像中的位置越靠近摄像头图像中心，则人脸在图像中的位置分值越高。

进一步优选，为满足实际显示会议画面中对人脸个数、人脸面积以及人脸在图像中的位置的不同考虑，还可以具体通过在根据人脸个数、人脸面积以及人脸在图像中的位置确定对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值时，设置对应的权重值来分别调整人脸个数分值、人脸面积分值以及人脸在图像中的位置分值的分值大小，进而影响最终选择显示的摄像头的画面。

根据每一摄像头的图像数据的人脸个数分值、人脸面积分值和人脸在图像中的位置分值的和值的大小，获取每一摄像头的的第一优先级。

具体实施时，本实施例首先通过图像数据获取单元 11 获取摄像头实时拍摄的图像数据，然后通过人脸信息获取单元 12 对每一摄像头拍摄的图像数据进行人脸检测，包括检测每一图像中的人脸个数、人脸面积和人脸在图像中的位置，进而通过第一优先级获取单元 13 获取人脸个数分值、人脸面积分值和人脸在图像中的位置分值，将每一摄像头的图像获得的上述三个分值的和值作为该摄像头对应的第一优先级；接着，通过声源距离获取单元 14 获取麦克风所接收声源的位置，进而获取声源位置与每一摄像头的距离，通过第二优先级获取单元 15 根据每一摄像头与声源位置的距离来获取对应的第二优先级，其中，与声源位置的

距离越小，摄像头的第二优先级越高；最后，通过最终优先级获取单元 16 对于每一摄像头，将第一优先级和第二优先级分别乘以对应的权重值然后求和得到最终优先级，有输出单元 17 将最终优先级最高的摄像头的图像数据输出以用于显示。

本实施例二先通过人脸信息来获取第一优先级，然后通过声源位置与摄像头距离来获取第二优先级的顺序只是一种实施示例，调换先后进行的顺序，或并行进行上述两个步骤的实施例都在本实施例的保护范围之内。

与现有技术相比，本实施例基于人脸检测和声源位置的结果来选取所显示的摄像头的拍摄的画面，实现智能自主切换摄像头画面，满足远程音视频会议的显示需求，减少人工操作，更加智能化、自动化；提高用户体验。

本发明实施例三还提供一种音视频系统，参见图 5，图 5 为本发明实施例三的结构示意图，其中，本实施例三包括本发明实施例二提供的一种自动选择摄像头画面的装置 1，具体可以见上述本发明实施例二所述的装置内容，此处不做赘述。另外，本实施例三还包括以下结构：

2 个摄像头 201，摄像头 201 分别安装于智能平板的左右两侧或上下两侧，用于实时拍摄画面图像；优选的，采用图 1 中 2 个摄像头 201 安装在智能平板的左右两侧的实施方式；此处，图 1 所示的 2 个摄像头 201 在智能平板的安装位置仅仅为一种实施方式，基于本发明实施例的原理，仅仅只是调整摄像头 201 在智能平板上的安装位置或增加摄像头 201 的数量的实施方式，也在本发明的保护范围之内。

麦克风 202，用于接收声源，并能确定接收的声源的位置。

具体实施时，本实施例首先通过图像数据获取单元 11 获取摄像头 201 的图像数据，其中，摄像头的图像为摄像头 201 实时拍摄的画面图像然后通过人脸信

息获取单元 12 对每一摄像头 201 拍摄的图像数据进行人脸检测，包括检测每一图像中的人脸个数、人脸面积和人脸在图像中的位置，进而通过第一优先级获取单元 13 获取人脸个数分值、人脸面积分值和人脸在图像中的位置分值，将每一摄像头 201 的图像获得的上述三个分值的和值作为摄像头 201 对应的第一优先级；接着，通过声源距离获取单元 14 获取声源的位置，进而获取声源位置与每一摄像头 201 的距离，其中，声源的位置为麦克风 202 接收声源，并确定所接收的声源位置所获得；然后，通过第二优先级获取单元 15 根据每一摄像头 201 与声源位置的距离来获取对应的第二优先级，其中，与声源位置的距离越小，摄像头 201 的第二优先级越高；最后，通过最终优先级获取单元 16 对于每一摄像头 201，将第一优先级和第二优先级分别乘以对应的权重值然后求和得到最终优先级，由输出单元 17 将最终优先级最高的摄像头 201 的图像数据传输到显示画面。

与现有技术相比，本实施例一种音视频系统能够基于人脸检测和声源位置的结果来选取所显示的摄像头的拍摄的画面，实现智能自主切换摄像头画面，满足远程视频会议的显示需求，减少人工操作，更加智能化、自动化；提高用户体验。

本发明实施例四提供一种智能平板，参见图 6，图 6 为本发明实施例四的结构示意图，其中，本实施例四包括本发明实施例三提供的一种音视频系统，具体可以见上述本发明实施例三所述的音视频系统装置的内容，此处不做赘述。

具体实施时，本实施例首先通过图像数据获取单元 11 获取摄像头 201 的图像数据，其中，摄像头的图像为摄像头 201 实时拍摄的画面图像然后通过人脸信息获取单元 12 对每一摄像头 201 拍摄的图像数据进行人脸检测，包括检测每一图像中的人脸个数、人脸面积和人脸在图像中的位置，进而通过第一优先级获取单元 13 获取人脸个数分值、人脸面积分值和人脸在图像中的位置分值，将每一摄像头 201 的图像获得的上述三个分值的和值作为摄像头 201 对应的第一优先

级；接着，通过声源距离获取单元 14 获取声源的位置，进而获取声源位置与每一摄像头 201 的距离，其中，声源的位置为麦克风 202 接收声源，并确定所接收的声源位置所获得；然后，通过第二优先级获取单元 15 根据每一摄像头 201 与声源位置的距离来获取对应的第二优先级，其中，与声源位置的距离越小，摄像头 201 的第二优先级越高；最后，通过最终优先级获取单元 16 对于每一摄像头 201，将第一优先级和第二优先级分别乘以对应的权重值然后求和得到最终优先级，由输出单元 17 将最终优先级最高的摄像头 201 的图像数据传输到显示画面。

与现有技术相比，本实施例一种智能平板能够基于人脸检测和声源位置的结果来选取所显示的摄像头的拍摄的画面，实现智能自主切换摄像头画面，满足远程音视频会议的显示需求，减少人工操作，更加智能化、自动化；提高用户体验。

以上所述是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也视为本发明的保护范围。

权 利 要 求 书

1、一种自动选择摄像头画面的方法，其特征在于，包括：

获取每一摄像头的图像数据；其中，获取的所述图像数据为所述每一摄像头实时拍摄的画面图像；

对所述每一摄像头的图像数据进行人脸检测，以获取所述每一摄像头的图像数据中的人脸信息；

根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头的的第一优先级；

通过每一麦克风确定所述麦克风所对应接收的声源的位置，从而获取每一所述摄像头与所述声源的距离；

根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；

根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级；

将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。

2、如权利要求 1 所述的一种自动选择摄像头画面的方法，其特征在于，所述根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级具体包括：

对于每一所述摄像头，将所述第一优先级乘以预置第一权重得到第一乘积值，将所述第二优先级乘以预置第二权重得到第二乘积值；

根据所述第一乘积值和所述第二乘积值的和值得到最终优先级。

3、如权利要求 1 所述的一种自动选择摄像头画面的方法，其特征在于，所述根据所述每一摄像头与所述声源位置距离，获取每一所述摄像头的第二优先级包括：

根据所述每一摄像头与所述声源位置距离的大小，对所述每一摄像头进行排序，与所述声源的距离越小的所述摄像头对应获取的第二优先级越高。

4、如权利要求 1 所述的一种自动选择摄像头画面的方法，其特征在于，所述人脸信息包括获取人脸个数、人脸面积以及人脸在图像中的位置。

5、如权利要求 4 所述的一种自动选择摄像头画面的方法，其特征在于，所述根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头图像数据的第一优先级具体包括：

根据所述每一摄像头的图像数据对应获取的所述人脸个数、所述人脸面积以及所述人脸在图像中的位置分别获取对应的人脸个数分值、人脸面积分值以及人脸在图像中的位置分值；

根据所述每一摄像头的图像数据的所述人脸个数分值、所述人脸面积分值和所述人脸在图像中的位置分值的和值的大小，获取所述每一摄像头的所述第一优先级。

6、一种自动选择摄像头画面的装置，其特征在于，包括：

图像数据获取单元，用于获取每一摄像头的图像数据；其中，获取的所述图

像数据为所述每一摄像头的实时拍摄的画面图像；

人脸信息获取单元，用于对所述每一摄像头的图像数据进行人脸检测，以获取所述每一摄像头的图像数据中的人脸信息；

第一优先级获取单元，用于根据所述每一摄像头的图像数据对应获取的所述人脸信息，获取每一所述摄像头的第一优先级；

声源距离获取单元，用于通过每一麦克风确定所述麦克风所对应接收的声源的位置，从而获取每一所述摄像头与所述声源的距离；

第二优先级获取单元，用于根据所述每一摄像头与所述声源位置的距离，获取每一所述摄像头的第二优先级；

最终优先级获取单元，用于根据所述每一摄像头的所述第一优先级和所述第二优先级获取所述每一摄像头的最终优先级；

输出单元，用于将所述最终优先级最高的所述摄像头的图像数据输出以用于显示。

7、如权利要求 6 所述的一种自动选择摄像头画面的装置，其特征在于，所述最终优先级获取单元具体用于：

对于每一所述摄像头，将所述第一优先级乘以预置第一权重得到第一乘积值，将所述第二优先级乘以预置第二权重得到第二乘积值；

根据所述第一乘积值和所述第二乘积值的和值得到最终优先级。

8、如权利要求 6 所述的一种自动选择摄像头画面的装置，其特征在于，所述第二优先级获取单元用于根据所述每一摄像头与所述声源位置距离，获取每一所述摄像头的第二优先级时，根据所述每一摄像头与所述声源位置距离的大小，

对所述每一摄像头进行排序，与所述声源的距离越小的所述摄像头对应获取的第二优先级越高。

9、一种音视频系统，其特征在于，包括如权利要求 6~8 任一项所述的一种自动选择摄像头画面的装置，还包括：

至少 2 个摄像头，所述摄像头分别安装于智能平板的左右两侧或上下两侧，用于实时拍摄画面图像；

麦克风，用于接收声源，并确定接收的所述声源的位置。

10、一种智能平板，其特征在于，包括权利要求 9 所述的音视频系统。



图 1

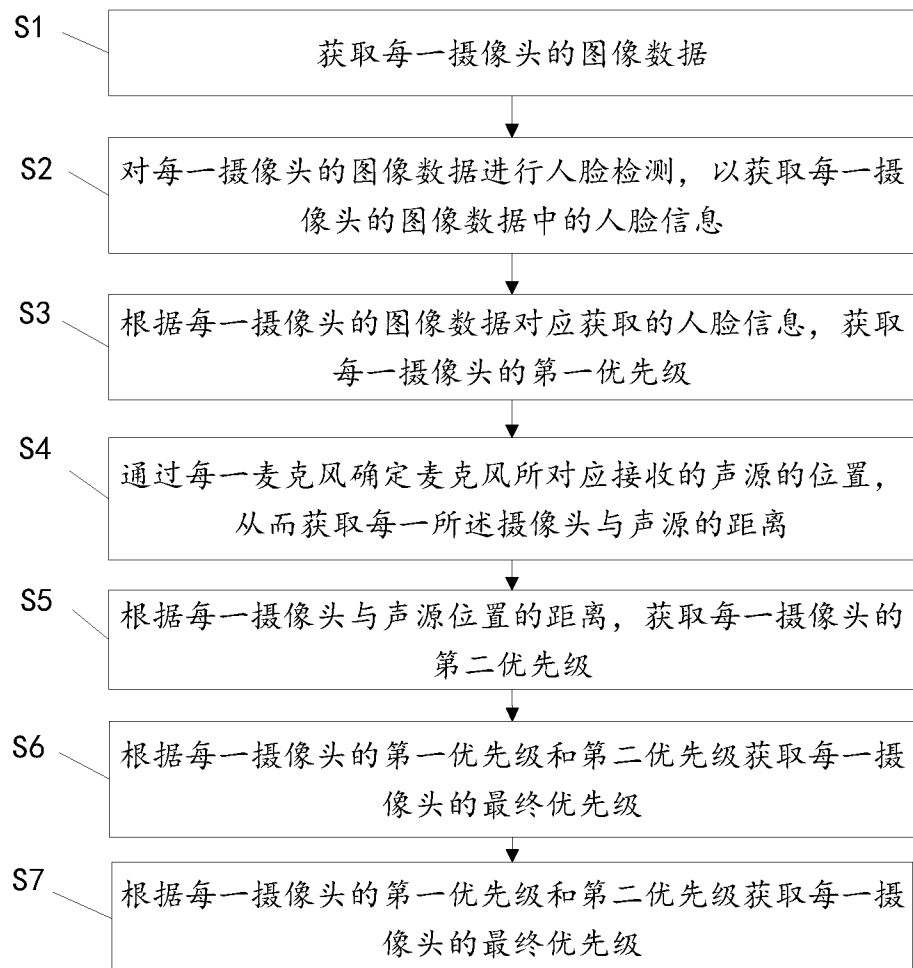


图 2

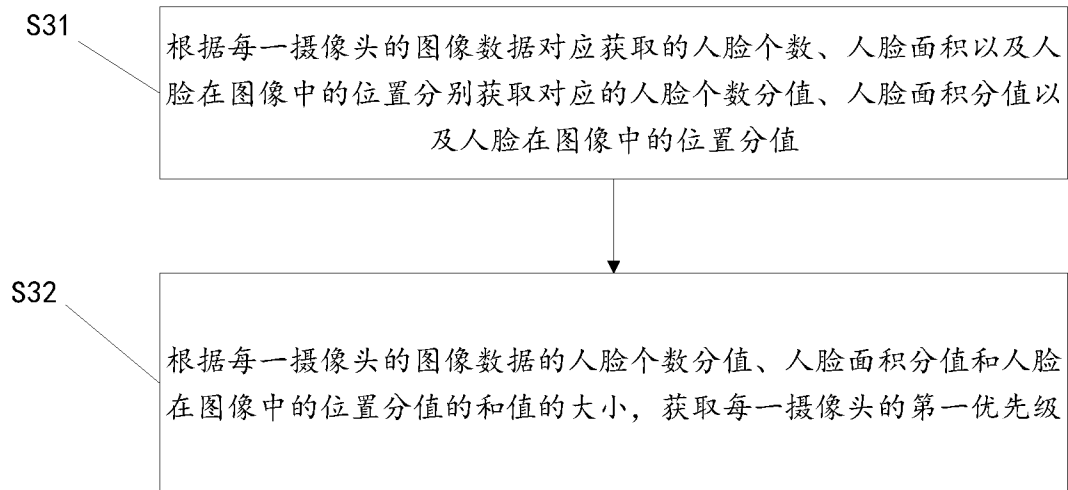


图 3

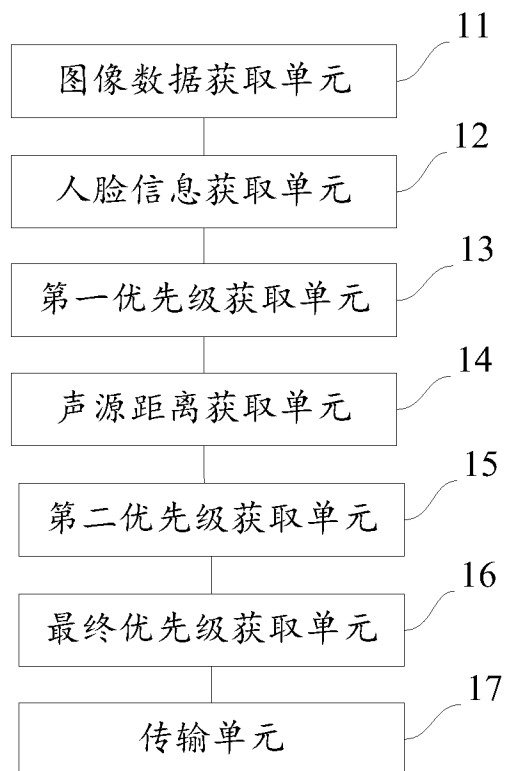


图 4

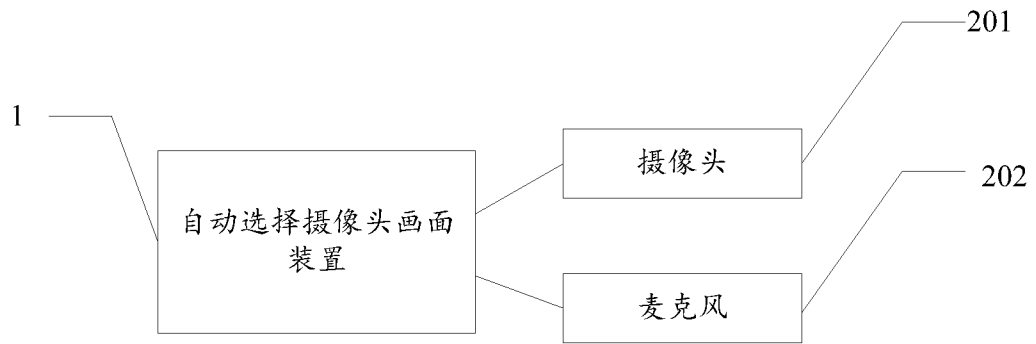


图 5

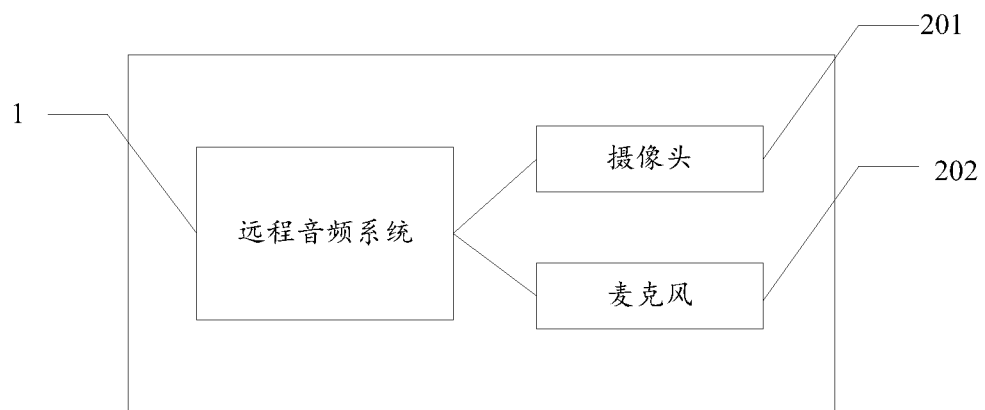


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/104657

A. CLASSIFICATION OF SUBJECT MATTER

H04N 7/15 (2006.01) i; H04N 5/268 (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

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; IEEE: 自动, 选择, 切换, 摄像, 声源, 音频, 声音, 距离, 位置, 方位, 图像, 人脸, 面部, 优先级, 显示, 对应, camera, switch+, image, face, priority, position, distance, sound, audio, display+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103237178 A (BEIJING XIAOMI TECHNOLOGY CO.) 07 August 2013 (07.08.2013), claims 1-3, description, paragraphs [0056]-[0114], and figures 1-3	1-10
A	CN 1917623 A (SONY CORPORATION) 21 February 2007 (21.02.2007), entire document	1-10
A	CN 104038725 A (HUAWEI DEVICE CO., LTD.) 10 September 2014 (10.09.2014), entire document	1-10
A	CN 101685153 A (SHENZHEN HUAWEI COMMUNICATION TECHNOLOGY CO., LTD.) 31 March 2010 (31.03.2010), entire document	1-10
A	US 2006136962 A1 (FUNAI ELECTRIC CO., LTD.) 22 June 2006 (22.06.2006), entire document	1-10
A	US 2012050601 A1 (SAMSUNG ELECTRONICS CO., LTD.) 01 March 2012 (01.03.2012), entire document	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 08 January 2018	Date of mailing of the international search report 26 January 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer WU, Qian Telephone No. (86-10) 62413079

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/104657

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103237178 A	07 August 2013	CN 103237178 B	16 June 2017
CN 1917623 A	21 February 2007	US 2007040903 A1	22 February 2007
		US 8044990 B2	25 October 2011
		JP 2007053513 A	01 March 2007
		JP 4356663 B2	04 November 2009
		CN 100525433 C	05 August 2009
CN 104038725 A	10 September 2014	None	
CN 101685153 A	31 March 2010	CN 101685153 B	16 November 2011
US 2006136962 A1	22 June 2006	US 7673313 B2	02 March 2010
		JP 2006180117 A	06 July 2006
US 2012050601 A1	01 March 2012	KR 20120019703 A	07 March 2012
		US 8797442 B2	05 August 2014

国际检索报告

国际申请号

PCT/CN2017/104657

A. 主题的分类 H04N 7/15(2006.01)i; H04N 5/268(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 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;IEEE:自动, 选择, 切换, 摄像, 声源, 音频, 声音, 距离, 位置, 方位, 图像, 人脸, 面部, 优先级, 显示, 对应, camera, switch+, image, face, priority, position, distance, sound, audio, display+																							
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>X</td> <td>CN 103237178 A (北京小米科技有限责任公司) 2013年 8月 7日 (2013 - 08 - 07) 权利要求1-3, 说明书第[0056]-[0114]段, 附图1-3</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 1917623 A (索尼株式会社) 2007年 2月 21日 (2007 - 02 - 21) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 104038725 A (华为终端有限公司) 2014年 9月 10日 (2014 - 09 - 10) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 101685153 A (深圳华为通信技术有限公司) 2010年 3月 31日 (2010 - 03 - 31) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2006136962 A1 (FUNAI ELECTRIC CO., LTD.) 2006年 6月 22日 (2006 - 06 - 22) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2012050601 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2012年 3月 1日 (2012 - 03 - 01) 全文</td> <td>1-10</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	X	CN 103237178 A (北京小米科技有限责任公司) 2013年 8月 7日 (2013 - 08 - 07) 权利要求1-3, 说明书第[0056]-[0114]段, 附图1-3	1-10	A	CN 1917623 A (索尼株式会社) 2007年 2月 21日 (2007 - 02 - 21) 全文	1-10	A	CN 104038725 A (华为终端有限公司) 2014年 9月 10日 (2014 - 09 - 10) 全文	1-10	A	CN 101685153 A (深圳华为通信技术有限公司) 2010年 3月 31日 (2010 - 03 - 31) 全文	1-10	A	US 2006136962 A1 (FUNAI ELECTRIC CO., LTD.) 2006年 6月 22日 (2006 - 06 - 22) 全文	1-10	A	US 2012050601 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2012年 3月 1日 (2012 - 03 - 01) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	CN 103237178 A (北京小米科技有限责任公司) 2013年 8月 7日 (2013 - 08 - 07) 权利要求1-3, 说明书第[0056]-[0114]段, 附图1-3	1-10																					
A	CN 1917623 A (索尼株式会社) 2007年 2月 21日 (2007 - 02 - 21) 全文	1-10																					
A	CN 104038725 A (华为终端有限公司) 2014年 9月 10日 (2014 - 09 - 10) 全文	1-10																					
A	CN 101685153 A (深圳华为通信技术有限公司) 2010年 3月 31日 (2010 - 03 - 31) 全文	1-10																					
A	US 2006136962 A1 (FUNAI ELECTRIC CO., LTD.) 2006年 6月 22日 (2006 - 06 - 22) 全文	1-10																					
A	US 2012050601 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2012年 3月 1日 (2012 - 03 - 01) 全文	1-10																					
☐ 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 2018年 1月 8日		国际检索报告邮寄日期 2018年 1月 26日																					
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(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 吴倩 电话号码 (86-10)62413079																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/104657

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	103237178	A	2013年 8月 7日	CN	103237178	B	2017年 6月 16日
CN	1917623	A	2007年 2月 21日	US	2007040903	A1	2007年 2月 22日
				US	8044990	B2	2011年 10月 25日
				JP	2007053513	A	2007年 3月 1日
				JP	4356663	B2	2009年 11月 4日
				CN	100525433	C	2009年 8月 5日
CN	104038725	A	2014年 9月 10日	无			
CN	101685153	A	2010年 3月 31日	CN	101685153	B	2011年 11月 16日
US	2006136962	A1	2006年 6月 22日	US	7673313	B2	2010年 3月 2日
				JP	2006180117	A	2006年 7月 6日
US	2012050601	A1	2012年 3月 1日	KR	20120019703	A	2012年 3月 7日
				US	8797442	B2	2014年 8月 5日