

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2020/048034 A1

(43) 国际公布日

2020 年 3 月 12 日 (12.03.2020)

(51) 国际专利分类号:

H04N 21/43 (2011.01) H04N 21/4402 (2011.01)

H04N 21/439 (2011.01) H04N 21/442 (2011.01)

H04N 21/44 (2011.01)

(21) 国际申请号:

PCT/CN2018/120528

(22) 国际申请日:

2018 年 12 月 12 日 (12.12.2018)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201811043120.4 2018年9月7日 (07.09.2018) CN

(71) 申请人: 深圳创维-RGB电子有限公司(SHENZHEN SKYWORTH-RGB ELECTRONIC CO., LTD.) [CN/CN]; 中国广东省深圳市南山区深南大道创维大厦A座13-16层, Guangdong 518052 (CN)。

(72) 发明人: 赵新科(ZHAO, Xinke); 中国广东省深圳市南山区深南大道创维大厦A座13-16层, Guangdong 518052 (CN)。

(74) 代理人: 北京品源专利代理有限公司(BEYOND ATTORNEYS AT LAW); 中国北京市海淀区莲花池东路39号西金大厦6层, Beijing 100036 (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,

(54) Title: METHOD, APPARATUS, DEVICE, AND STORAGE MEDIUM FOR IMPLEMENTING SOUND AND IMAGE PARITY

(54) 发明名称: 实现声音与图像同位的方法、装置、设备及存储介质

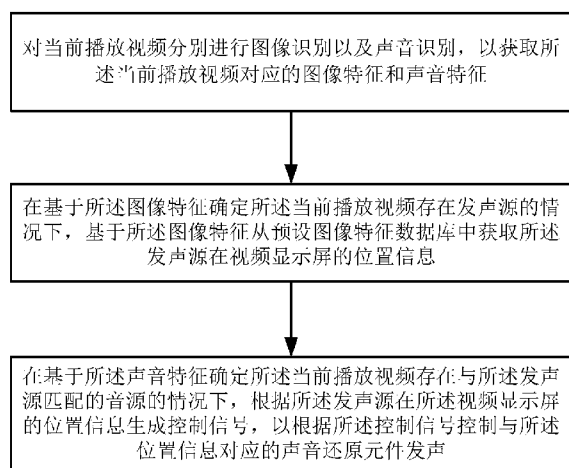


图 1a

- 10 Perform image recognition and sound recognition separately on a video that is currently playing back so as to acquire image features and sound features corresponding to the video that is currently playing back
- 20 When determined on the basis of the image features that a sounding source is present in the video that is currently playing back, acquire position information of the sounding source on a video display screen from a preset image feature database on the basis of the image features
- 30 When determined on the basis of the sound features that a sound source matching the sounding source is present in the video that is currently playing back, generate a control signal according to the position information of the sounding source on the video display screen so as to control a sound reduction element corresponding to the position information to sound according to the control signal

(57) Abstract: Disclosed herein is a method for implementing sound and image parity, the method comprising: performing image recognition and sound recognition separately on a video that is currently playing back so as to acquire image features and sound features corresponding to the video that is currently playing back; when determined on the basis of the image features that a sounding source is present in the video that is currently playing back, then acquiring position information of the sounding source of the video that is currently playing back on a video display screen; when determined on the basis of the sound features that a sound source matching the sounding source is present in the video that is currently playing back, generating a control signal according to the position information of the sounding source on the video display screen so as to control a sound reduction element corresponding to the position information to sound. Also disclosed herein are an apparatus, device, and storage medium for implementing sound and image parity.

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))。

(57) 摘要: 本文公开了一种实现声音与图像同位的方法, 包括: 对当前播放视频分别进行图像识别以及声音识别, 以获取当前播放视频对应的图像特征和声音特征; 在基于所述图像特征确定所述当前播放视频存在发声源的情况下, 则获取当前播放视频的发声源在视频显示屏的位置信息; 在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下, 根据所述发声源在视频显示屏的位置信息生成控制信号, 以控制与所述位置信息对应的声音还原元件发声。本文还公开了一种实现声音与图像同位的装置、设备以及存储介质。

实现声音与图像同位的方法、装置、设备及存储介质

本申请要求在 2018 年 9 月 7 日提交中国专利局、申请号为 201811043120.4 的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开实施例涉及智能电视技术领域，例如涉及一种实现声音与图像同位的方法、装置、设备及存储介质。

背景技术

目前，电子显示类产品，例如大尺寸液晶电视，在播放视频的时候，视频图像是通过显示屏呈现出来，而视频声音则是通过设置在电视其它位置的扬声器发声，由于视频声音与对应的视频图像不在同一个位置播放，导致视频的播放效果不佳，用户观看视频时的临场感不强。

发明内容

本公开提供一种实现声音与图像同位的方法、装置、设备及存储介质，通过所述方法有效实现声音与图像的同位置呈现，提高视频的播放效果。

在一实施例中，本公开实施例提供了一种实现声音与图像同位的方法，所述方法包括：

对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；

在基于所述图像特征确定所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述发声源在视频显示屏的位置信

息；

在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；

其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

在一实施例中，所述对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征，包括：

对所述当前播放视频进行解码，分别得到所述当前播放视频对应的图像数据和声音数据；

基于所述图像数据调用图像识别接口进行图像识别，得到与所述图像数据对应的图像特征，并基于所述声音数据调用声音识别接口进行声音识别，得到与所述声音数据对应的声音特征。

在一实施例中，所述声音还原元件依据所述视频显示屏预先划分的分区独立设置；

其中，所述分区的数量依据显示屏的尺寸进行设定。

在一实施例中，所述声音还原元件包括扬声器。

在一实施例中，所述根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声包括；

将所述声音数据通过杜比全景声（Atmosphere，ATMOS）解码，得到集成音频接口（Integrate Interface of Sound，IIS）音频信号；

根据所述 IIS 音频信号以及所述控制信号控制与所述位置信息对应的功率放大器工作以驱动所述位置信息对应的扬声器发声。在一实施例中，所述基于

所述图像特征确定所述当前播放视频存在发声源，包括：

将所述图像特征与预设图像特征数据库中的图像特征进行相似度匹配；

当匹配的所述相似度达到设定阈值时，则确定所述当前播放视频存在发声源；

其中，所述预设图像特征数据库中的图像特征包括下述至少一项：人体形态特征和动物形态特征。在一实施例中，所述基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源，包括：

将所述声音特征与预先建立的模型发声源的模型特征进行比对；

若存在与所述声音特征一致的模型特征且所述模型特征对应的模型发声源与所述当前播放视频中存在的发声源相同，则确定所述当前播放视频存在与所述发声源匹配的音源。

在一实施例中，所述控制与所述位置信息对应的声音还原元件发声，包括：

当所述声音还原元件发出的声音幅度超过设定上限时，降低所述声音还原元件的发声增益；

当所述声音还原元件发出的声音幅度没有超过所述设定下限时，提高所述声音还原元件的发声增益。

在一实施例中，所述当前播放视频通过按照单位时间内的预设采样次数对播放的视频进行采样而获得。

在一实施例中，所述视频显示屏包括预设尺寸的液晶显示屏。

在一实施例中，本公开实施例提供了一种实现声音与图像同位的装置，所述装置包括：

识别模块，设置为对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；

获取模块，设置为在基于所述图像特征确定所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息；

控制模块，设置为在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；

其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

在一实施例中，本公开实施例提供了一种电子设备，包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述的方法。

在一实施例中，本公开实施例提供了一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时实现上述的方法。

附图说明

图 1a 为本公开实施例一提供的一种实现声音与图像同位的方法流程示意图；

图 1b 为本公开实施例一提供的另一种实现声音与图像同位的方法流程示意图；

图 2 为本公开实施例一提供的又一种实现声音与图像同位的方法流程示意图；

图 3 为本公开实施例二提供的一种实现声音与图像同位的方法流程示意图；

图 4 为本公开实施例二提供的一种显示屏的分区示意图；

图 5 为本公开实施例二提供的一种根据所述控制信号控制对应功率放大器工作以驱动扬声器发声的流程示意图；

图 6 为本公开实施例三提供的一种实现声音与图像同位的装置的结构示意图；

图 7 为本公开实施例四提供的一种电子设备的结构示意图。

具体实施方式

实施例一

图 1a 为本公开实施例一提供的一种实现声音与图像同位的方法流程示意图。本实施例提供的实现声音与图像同位的方法可应用在具有大尺寸显示屏的电子产品上，例如 65 寸及以上的电视机产品。当显示屏尺寸较小时，由于声音还原系统与视频图像（即发声源）之间的距离较近，声音与图像同位的音响效果无法被突出体现。所述实现声音与图像同位的方法适用于具有明显方向属性的声音特征的视频的播放过程中。在一实施例中，所述具有明显方向属性的声音特征的视频例如是包含有人物且所述人物发出了说话的声音，吵架的声音或者唱歌的声音的视频，或者包含有动物且所述动物发出了叫声的视频，或者包含有物体且所述物体发出了敲打声（例如打铁、电焊等）或者打碎物体声音（例如打碎玻璃或者碗等陶瓷制品的声音）的视频，即具有明显声音特征且有发出声音的源头的视频。在具有声音但该声音没有明显方向属性的视频播放过程中，本实施例提供的实现声音与图像同位的方法无法正常执行。例如对于只包含有背景音乐的视频，由于所述背景音乐没有明显的发声源头，即不具有明显的方向属性，则无法应用本实施例提供的方法达到提升视频播放效果的目的，对于此类视频播放，只当作普通声音进行呈现，不进行声音与图像的同位操作。所述实现声音与图像同位的方法可以由实现声音与图像同位的装置来执行，该装置可由软件和/或硬件实现，一般集成在具有大尺寸显示屏的电子设备中。所述实现声音与图像同位的方法用于提升视频播放效果，提升用户观看视频的临场感，沉浸感。参见图 1a 所示，该方法包括如下步骤：

步骤 10，对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征。

步骤 20, 在基于所述图像特征确定所述当前播放视频存在发声源的情况下, 基于所述图像特征从预设图像特征数据库中获取所述发声源在视频显示屏的位置信息。

步骤 30, 在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下, 根据所述发声源在所述视频显示屏的位置信息生成控制信号, 以根据所述控制信号控制与所述位置信息对应的声音还原元件发声。

其中, 所述预设图像特征数据库依据所述当前播放视频预先构建。

图 1b 为本公开实施例一提供的另一种实现声音与图像同位的方法流程示意图。参见图 1b 所示, 该方法包括:

步骤 110, 对当前播放视频分别进行图像识别以及声音识别, 以获取所述当前播放视频对应的图像特征和声音特征。

其中, 所述图像特征是指当前播放视频中包含的人物、动物或者其他的物体、物品等的特征, 所述物体例如可以是木凳或者木桌, 所述物品例如可以是碗、杯子或者茶壶等陶瓷制品, 只要是当前播放视频中包含的物的特征都属于所述图像特征的范畴。所述声音特征是指当前播放视频中包含的音频特征, 例如唱歌声、说话声、动物叫喊声或者打碎物品的声音等。

所述当前播放视频通过对播放视频进行定期采样获得, 例如每秒中 (即, 单位时间内) 对播放视频采样两次 (即, 预设采样次数), 每次采样得到的视频均为当前播放视频。即: 所述当前播放视频通过按照单位时间内的预设采样次数对播放的视频进行采样而获得。

示例性的, 所述对当前播放视频分别进行图像识别以及声音识别, 以获取所述当前播放视频对应的图像特征和声音特征, 包括:

对当前播放视频进行解码, 分别得到当前播放视频对应的图像数据和声音

数据；

基于所述图像数据调用图像识别接口进行图像识别，得到与所述图像数据对应的图像特征，并基于所述声音数据调用声音识别接口进行声音识别，得到与所述声音数据对应的声音特征。

其中，对当前播放视频进行解码可以利用本领域内成熟的解码算法进行，本实施例中不再赘述。在一实施例中，所述图像识别接口为百度的人脸识别系统，可实现对当前播放视频中的图像特征进行有效识别。在一实施例中，所述声音识别接口为提取声音中的振幅或者频率等能够体现声音特征的信息的程序模块，通过该程序模块可有效提取当前播放视频中的声音特征。

步骤 120，基于所述图像特征判断所述当前播放视频是否存在发声源，若所述当前播放视频存在发声源，则继续执行步骤 130，若所述当前播放视频不存在发声源，则返回执行步骤 110。其中，所述发声源指当前播放视频中发出声音的源头，例如发出声音的人或者物等。

示例性地，所述基于所述图像特征确定所述当前播放视频存在发声源，包括：

将所述图像特征与预设图像特征数据库中的图像特征进行相似度匹配；

当匹配的所述相似度达到设定阈值时，则确定所述当前播放视频存在发声源；

其中，所述预设图像特征数据库中的图像特征包括人体形态特征和/或动物形态特征，还可以包括物体、物品的形态特征，所述人体形态特征可以指人发出声音时的口型姿态，所述动物形态特征可以指动物发出声音时的口型姿态，所述物体、物品的形态特征可以指物体、物品发出声音时的姿态，例如物体之间的敲打、摩擦姿态，物品被打碎时的姿态。所述图像特征的实质是当前播放

视频中存在的发声源的标识。所述预设图像特征数据库中的图像特征可以基于人工智能技术，通过自主学习功能对所述当前播放视频预先进行学习得到，同时对预设图像特征数据库中的图像特征在当前视频播放设备的显示屏的位置进行标记，因此在自主学习的过程中还要加入设置为播放当前视频的电子设备的屏幕尺寸信息，考虑到成本问题，播放当前视频的电子设备可以仅考虑目前市场上比较流行的配置有 65 寸液晶显示屏的智能电视机。

由于当前播放视频可通过对播放视频进行定期采样得到，通过设置所述设定阈值，可以将一些图像特征不明显的图像过滤掉，即将当前次采样得到的视频数据放弃，等待下一次的采样数据，从而可减少实现声音与图像同位的方法对系统资源的占用率，同时可提高确定当前播放视频中是否存在发声源的准确性。

若当前播放视频不存在发声源，则表示当前播放视频不存在具有明显方向属性的声音特征，无法体现声音与图像同位的播放效果，因此不对当前播放视频进行声音与图像的同位操作，直接按照常规的视频播放流程进行播放即可，将视频中的声音通过当前视频播放设备的所有声道进行播放即可。

步骤 130，基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息。

其中，所述视频显示屏是指设置为播放视频的电子设备的显示屏。所述预设图像特征数据库依据所述当前播放视频预先构建，所述预设图像特征数据库中保存有当前播放视频的发声源与其在视频显示屏的位置信息之间的对应关系，通过该对应关系可以查找到所述发声源在视频显示屏的位置信息。

步骤 140，基于所述声音特征判断所述当前播放视频是否存在与所述发声源匹配的音源，若当前播放视频存在与所述发声源匹配的音源，则继续执行步骤

150, 若当前播放视频不存在与所述发声源匹配的音源, 则返回执行步骤 110。

示例性的, 所述基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源, 包括:

将所述声音特征与预先建立的模型发声源的模型特征进行比对;

若存在与所述声音特征一致的模型特征且所述模型特征对应的模型发声源与所述当前播放视频中存在的发声源相同, 则确定所述当前播放视频存在与所述发声源匹配的音源。

所述模型发声源的模型特征根据所述当前播放视频预先构建, 在所述模型发声源中保存有当前播放视频中存在的发声源以及该发声源对应的声音特征, 例如所述发声源为人物, 对应的声音特征为唱歌声, 则表示当前播放视频中的所述发声源对应的人物在唱歌; 若所述发声源为动物, 对应的声音特征为狗叫声, 则表示当前播放视频中所述发声源对应的人物在学狗叫; 若所述发声源为玻璃花瓶, 对应的声音特征为玻璃打碎的声音, 则表示当前播放视频中所述发声源对应玻璃花瓶发出了玻璃打碎的声音。

当从当前播放视频中识别出的声音特征以及确定出的发声源均与所述模型发声源中的模型特征以及对应的发声源匹配时, 则表示当前播放视频中存在具有明显方向属性的声音特征, 即音源。

其中, 所述声音特征是指当前播放视频中包含的音频特征, 例如唱歌声、说话声、动物叫喊声或者打碎物品的声音等。

步骤 150, 根据所述发声源在视频显示屏的位置信息生成控制信号, 以控制与所述位置信息对应的声音还原元件发声。

其中, 所述声音还原元件包括扬声器, 所述声音还原元件依据所述视频显示屏预先划分的分区独立设置;

所述分区的数量依据显示屏幕的大小进行设定。

在一实施例中，可以将所述视频显示屏预先划分为特定数量的小区域，每个小区域由独立的声音还原元件进行发声，通过确定所述发声源在哪个小区域，进而控制对应小区域的声音还原元件发声，从而实现了声音与图像同位的目的，给观看视频的用户一种所述声音是由发声源发出的临场效果。例如，假设当前播放视频内容为“皇后娘娘在训斥奴才”，通过本实施例提供的实现声音与图像同位的方法，让观看者感觉训斥奴才的话正好是从皇后娘娘的嘴部发出来的，给观看者一种较强的临场感体验，提高观看者的沉浸感。假设当前播放视频内容为“鸟飞过树林发出悦耳动听的鸟叫声”，通过本实施例提供的实现声音与图像同位的方法，让观看者感觉鸟叫声是从显示屏上某个位置处的鸟发出来的，给观看者一种较强的临场感体验。

本实施例提供的实现声音与图像同位的方法，通过图像识别以及声音识别识别当前播放视频中的发声源以及对应的声音特征，当当前播放视频存在具有明显方向属性的声音特征时，通过获取当前播放视频中的发声源在视频显示屏的位置信息，根据所述位置信息控制发声源处的声音还原元件发声，从而实现了声音与图像的同位，给人一种视频声音从对应的发声源处发出的感觉，提高了观看者的临场感、沉浸感。

在上述技术方案的基础上，本实施例提供了另一种实现声音与图像同位的方法流程示意图，参见图 2 所示，所述方法包括：

步骤 210，视频开始播放。

步骤 220，对正在播放的视频进行视频采样。

考虑到系统资源的占用问题以及视频的帧频，本实施例采用每秒 2 次的采样频率对正在播放的视频进行视频采样，在尽量减少对系统资源的占用的同时

保证不对实现声音与图像同位的方法造成影响，不遗落视频中任何一个具有明显方向属性的发声源。

步骤 230，对采样得到的视频进行视频解码。

在一实施例中，对采样得到的视频进行视频解码是为了分别得到视频中的图像数据和声音数据。

其中，对采样得到的视频进行视频解码可以利用本领域内成熟的解码算法进行，本实施例中不再赘述。

步骤 231，得到视频中的声音数据。

步骤 240，得到视频中的图像数据。

步骤 250，根据图像数据进行图像识别，得到图像特征。

在一实施例中，可以通过调用图像识别接口进行图像识别操作，示例性地，所述图像识别接口为百度的人脸识别系统，可实现对当前播放视频中的图像特征进行有效识别。

步骤 260，将所述图像特征与图像数据库中的图像特征进行匹配。

其中，所述图像数据库依据当前播放视频预先构建，其中保存有当前播放视频中存在的发声源的图像特征。

步骤 270，确认是否获得匹配数据，若获得匹配数据，则执行步骤 280，若没有获得匹配数据，则放弃当前的采样数据，进行下次采样。

其中，确认是否获得匹配数据的实质是判断图像数据库中是否存在与所述图像特征匹配的数据，若图像数据库中存在与所述图像特征匹配的数据，则执行步骤 280 的根据声音数据进行声音识别，得到声音特征；若图像数据库中不存在与所述图像特征匹配的数据，则放弃当前的采样数据，进行下次采样。

步骤 280，根据声音数据进行声音识别，得到声音特征。

步骤 290，将所述声音特征与声音数据库中的声音特征进行匹配。

其中，所述声音数据库依据当前播放视频预先构建，其中保存有当前播放视频中的发声源发出的声音的特征数据。

步骤 2100，确认是否获得匹配数据，若获得匹配数据，则执行步骤 2110，若没有获得匹配数据，则放弃当前的采样数据，进行下次采样。

其中，确认是否获得匹配数据的实质是判断所述声音数据库中是否存在与所述声音特征匹配的数据。

在一实施例中，所述声场控制信息为控制发声源在当前显示屏的位置处的扬声器发声的控制信息。

步骤 2110，根据发声源在视频显示屏的位置信息输出声场控制信息，以控制相应的声场发声。

通过对采样得到的视频数据进行图像识别，实现了确定所述视频数据中是否存在具有明显方向属性的发声源的目的，当存在发声源时，则继续对视频数据进行声音特征识别，当具有与所述发声源匹配的声音特征时，则控制所述发声源在显示屏的位置处的喇叭发声，实现了声音与图像的同位，提高了视频的播放效果，给观看者带来较强的临场感体验。

实施例二

图 3 为本公开实施例二提供的一种实现声音与图像同位的方法流程示意图。在上述实施例的基础上，本实施例对所述发声源声音还原的实现过程进行了说明。参见图 3 所示，所述方法包括：

步骤 310，对当前播放视频进行解码，分别得到当前播放视频对应的图像数据和声音数据。

步骤 320, 基于所述图像数据调用图像识别接口进行图像识别, 得到与所述图像数据对应的图像特征, 并基于所述声音数据调用声音识别接口进行声音识别, 得到与所述声音数据对应的声音特征。

步骤 330, 基于所述图像特征判断所述当前播放视频是否存在发声源, 若所述当前播放视频存在发声源, 则继续执行步骤 340, 若所述当前播放视频不存在发声源, 则返回执行步骤 310。

步骤 340, 基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息。

步骤 350, 基于所述声音特征判断所述当前播放视频是否存在与所述发声源匹配的音源, 若存在与所述发声源匹配的音源, 则继续执行步骤 360, 若不存在与所述发声源匹配的音源, 则返回执行步骤 310。

步骤 360, 根据所述发声源在视频显示屏的位置信息生成控制信号。

在一实施例中, 设置为播放所述视频的电子设备的显示屏能够分区发声是本公开实施例提供的实现声音与图像同位的方法实现的基本前提, 只有在显示屏相应位置处安装了发声元件才能实现具有临场感的声音效果。但是由于音源是具备区域大小属性的, 因此不可能实现图像与声音的绝对同位。若采用虚拟声音算法将声音虚拟在显示屏上发声, 则要实时进行视频图像识别与声场虚拟应用, 会占用较多的中央处理单元 (Central Processing Unit, CPU) 资源, 为了节约系统资源同时又能体现声音与图像同位的效果, 本实施例将所述显示屏预先划分为特定数量的分区, 且为每个分区虚拟出独立的声场, 通过为每个分区配置独立的扬声器实现为每个分区虚拟出独立的声场。在一实施例中, 所述视频显示屏包括预设尺寸的液晶显示屏。示例性的, 所述预设尺寸可以为 65 寸 (即 6.5 尺)。图 4 是一种显示屏的分区示意图, 如图 4 所示, 为了节约系统

资源同时又能体现声像同位的效果，将 65 寸及以上的显示屏划分为面积相等的 6 个分区，分别对应 6 个虚拟声场，所述 6 个虚拟声场分别为声场 1、声场 2、声场 3、声场 4、声场 5 和声场 6；每个虚拟声场由独立的扬声器实现，对应的 6 个扬声器分别安装在显示屏的左上、左中、左下、右上、右中和右下六个方位。声场 1 和声场 2 的两个扬声器通过第一功率放大器驱动来还原声场 1 和声场 2 两个声场的声音；声场 3 和声场 4 的两个扬声器通过第二功率放大器驱动来还原声场 3 和声场 4 两个声场的声音；声场 5 和声场 6 对应的两个扬声器通过第三功率放大器驱动来还原声场 5 和声场 6 两个声场的声音。

音源是指具备声音信息的视频信号，所述音源可通过对视频进行解码获取，从视频中解码出的音源，即声音数据通过解码，可以从声音数据中分离出多个方向的声音数据，声音解码的方式有很多，例如 ATMOS 解码、数字化影院系统（Digital Theater System, DTS）解码等，但只有采用 ATMOS 解码才能将双声道的声音解码成 8 声道的声音。本实施例中，获取解码出来的 6 个方向的声音，即声场 1、声场 2、声场 3、声场 4、声场 5 和声场 6 六个方向的声音信号。由于这 6 个方向的声音信号都是调制在一个 IIS 信号中，因此，可以将上述第一功率放大器、第二功率放大器和第三功率放大器的驱动功能连接到同一个 IIS 信号进行解码。对音源进行解码并根据所述控制信号控制与所述位置信息对应的功率放大器工作以驱动扬声器发声的流程示意图可参见图 5 所示，所述方法包括：

步骤 510，获取音源。

在一实施例中，可以通过对视频进行解码获取视频中的音源。

步骤 520，通过 ATMOS 芯片对所述音源解码。

其中，所述 ATMOS 芯片配置在播放所述视频的电子设备中，所述 IIS 音频

信号包含有对所述第一功率放大器、第二功率放大器和第三功率放大器进行控制的控制逻辑。

步骤 530，得到 IIS 音频信号。

步骤 540，将所述声场控制信息发送至所述 IIS 音频信号。

其中，所述声场控制信息为控制所述发声源在当前显示屏的位置处的声场发声的控制信息，通过所述位置处的功率放大器驱动所述位置处的扬声器工作。

将所述声场控制信息发送至所述 IIS 音频信号的目的是将所述声场控制信息编码至所述 IIS 音频信号。IIS 音频信号是一种数字信号，其内部调制了视频中多个方向的声音信号，而声场控制信息中包含有要被触发的声场位置信息，当 IIS 音频信号被还原成模拟信号时，根据所述声场控制信息来选择还原 IIS 音频信号中哪个方向的声音，因此可将所述声场控制信息编码至所述 IIS 音频信号一起被还原为模型信号。

在一实施例中，对视频中的声音进行还原的过程为：对所述 IIS 音频信号进行解码，获取与上述声场 1、声场 2、声场 3、声场 4、声场 5 和声场 6，6 个方向对应的声音，并利用解码得到声场控制信息控制与所述位置处的功率放大器工作以驱动对应的扬声器发声，从而实现还原所述位置处所在方向的声音。例如所述声场控制信息为触发声场 3 的控制信息，则在还原 IIS 音频信号时，就只还原声场 3 所在方向的声音信号，其它声场区不输送声音信号。

步骤 370，将所述声音数据通过 ATMOS 解码，得到 IIS 音频信号。

步骤 380，根据所述 IIS 音频信号以及所述控制信号控制与所述位置信息对应的功率放大器工作以驱动对应的扬声器发声。

例如，通过对当前播放视频进行图像识别，识别到当前播放视频中的发声源位于当前显示屏的声场 3 的区域，则控制声场 3 的扬声器发声，同时关闭其

他声场的扬声器，只保留声场 3 的扬声器发声，呈现一种声音的位置与发出声音的人的位置基本一致，给观看者带来较强的临场感体验。

在一实施例中，所述控制与所述位置信息对应的声音还原元件发声，包括：

当所述声音还原元件发出的声音幅度超过设定上限时，则降低所述声音还原元件的发声增益，当所述声音还原元件发出的声音幅度没有超过所述设定下限时，提高所述声音还原元件的发声增益。

当视频中没有具有明显方向属性的发声源时，视频声音是通过上述 6 个声场的扬声器共同发声来呈现的，而如果突然进入具有明显方向属性的发声源的视频画面时，则需仅通过与所述发声源位置对应的扬声器来发声，会导致视频声音的幅度突然变化，给观看者带来不好的体验。因此，通过采用声音幅度动态调整技术，即通过专业音效算法，将声音幅度控制在一个设定的范围，当声音幅度低于设定下限时，则提升与所述发声源位置对应的扬声器的增益；当声音幅度超过设定上限时，则降低与所述发声源位置对应的扬声器的增益，从而实现任何时刻视频的音量都在设定范围内。

本实施例提供的一种实现声音与图像同位的方法，通过将播放视频的电子设备的显示屏预先划分为特定数量的分区，并为每个分区配置独立的扬声器，以为每个分区虚拟出独立的声场，实现了节约系统资源同时又能体现声音与图像同位的播放效果的目的。

实施例三

图 6 为本公开实施例三提供的一种实现声音与图像同位的装置的结构示意图；参见图 6 所示，所述装置包括：识别模块 610、获取模块 620 和控制模块 630；

其中，识别模块 610，设置为对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；获取模块 620，设置为在基于所述图像特征获取所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息；控制模块 630，设置为在基于所述声音特征获取所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

在一实施例中，识别模块 610 是设置为对当前播放视频进行解码，分别得到当前播放视频对应的图像数据和声音数据；基于所述图像数据调用图像识别接口进行图像识别，得到与所述图像数据对应的图像特征，并基于所述声音数据调用声音识别接口进行声音识别，得到与所述声音数据对应的声音特征。

在一实施例中，所述声音还原元件依据所述视频显示屏幕预先划分的分区独立设置；

其中，所述分区的数量依据显示屏幕的尺寸进行设定。

在一实施例中，控制模块 630 是设置为；将所述声音数据通过 ATMOS 解码，得到 IIS 音频信号；根据所述 IIS 音频信号以及所述控制信号控制与所述位置信息对应的功率放大器工作以驱动所述位置信息对应的扬声器发声。

在一实施例中，获取模块 620 包括发声源确定子模块 640；发声源确定子模块 640 设置为基于所述图像特征确定所述当前播放视频存在发声源。

在一实施例中，发声源确定子模块 640 包括：

匹配单元，设置为将所述图像特征与预设图像特征数据库中的图像特征进

行相似度匹配；

确定单元，设置为当匹配的所述相似度达到设定阈值时，则确定所述当前播放视频存在发声源；

其中，所述预设图像特征数据库中的图像特征包括人体形态特征和/或动物形态特征。

在一实施例中，控制模块 630 包括音源确定子模块 650；所述音源确定子模块 650 设置为：基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源。

在一实施例中，音源确定子模块 650 是设置为：将所述声音特征与预先建立的模型发声源的模型特征进行比对；若存在与所述声音特征一致的模型特征且所述模型特征对应的模型发声源与所述当前播放视频中存在的发声源相同，则确定所述当前播放视频存在与所述发声源匹配的音源。

在一实施例中，所述控制模块 630 还设置为：当所述声音还原元件发出的声音幅度超过设定上限时，降低所述声音还原元件的发声增益，当所述声音还原元件发出的声音幅度没有超过所述设定下限时，提高所述声音还原元件的发声增益。

本实施例提供的实现声音与图像同位的装置，通过图像识别以及声音识别识别当前播放视频中的发声源以及对应的声音特征，当当前播放视频存在具有明显方向属性的声音特征时，通过获取当前播放视频中的发声源在视频显示屏的位置信息，根据所述位置信息控制发声源处的声音还原元件发声，从而实现了声音与图像的同位，给人一种视频声音从对应的发声源处发出的感觉，提高了观看者的临场感、沉浸感。

上述产品可执行本公开任意实施例所提供的方法，具备执行方法相应的功

能模块和效果。未在本实施例中详尽描述的技术细节，可参见本公开任意实施例所提供的方法。

本公开的技术方案实现了声音与图像的同位置呈现，使观看视频的用户感觉到视频声音的位置与视频中发出声音的对象的位置基本一致，提高了视频的播放效果，提升了用户体验。

实施例四

图 7 为本公开实施例四提供的一种电子设备的结构示意图。如图 7 所示，该电子设备包括：处理器 770、存储器 771 及存储在存储器 771 上并可在处理器 770 上运行的计算机程序；其中，处理器 770 的数量可以是一个或多个，图 7 中以一个处理器 770 为例；处理器 770 执行所述计算机程序时实现如上述实施例所述的实现声音与图像同位的方法。如图 7 所示，所述电子设备还可以包括输入装置 772 和输出装置 773。处理器 770、存储器 771、输入装置 772 和输出装置 773 可以通过总线或其他方式连接，图 7 中以通过总线连接为例。

存储器 771 作为一种计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块，如本公开实施例中实现声音与图像同位的方法对应的程序指令/模块（例如，实现声音与图像同位的装置中的识别模块 610、获取模块 620、和控制模块 630 等）。处理器 770 通过运行存储在存储器 771 中的软件程序、指令以及模块，从而执行电子设备的多种功能应用以及数据处理，即实现上述的实现声音与图像同位的方法。

存储器 771 可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序；存储数据区可存储根据终端的使用所创建的数据等。此外，存储器 771 可以包括高速随机存取存储器，还可以包

括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实例中，存储器 771 可以包括相对于处理器 770 远程设置的存储器，这些远程存储器可以通过网络连接至电子设备/存储介质。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入装置 772 可以设置为接收输入的数字或字符信息，以及产生与电子设备的用户设置以及功能控制有关的键信号输入。输出装置 773 可以包括显示屏等显示设备。

实施例五

本公开实施例五还提供一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行一种实现声音与图像同位的方法，该方法包括：

对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；

在基于所述图像特征确定所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息；

在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；

其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

当然，本公开实施例所提供的一种包含计算机可执行指令的存储介质，其计算机可执行指令不限于如上所述的方法操作，还可以执行本公开任意实施例

所提供的实现声音与图像同位的方法中的相关操作。

通过以上关于实施方式的描述，所属领域的技术人员可以清楚地了解到，本公开可借助软件及通用硬件来实现，也可以通过硬件实现。基于这样的理解，本公开的技术方案本质上或者说对相关技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在计算机可读存储介质中，如计算机的软盘、只读存储器（Read-Only Memory, ROM）、随机存取存储器（Random Access Memory, RAM）、闪存（FLASH）、硬盘或光盘等，包括多个指令用以使得一台计算机设备（可以是个人计算机，存储介质，或者网络设备）执行本公开一个或多个实施例所述的方法。

22
权 利 要 求 书

1、一种实现声音与图像同位的方法，包括：

对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；

在基于所述图像特征确定所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述发声源在视频显示屏的位置信息；

在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；

其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

2、根据权利要求 1 所述的方法，其中，所述对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征，包括：

对所述当前播放视频进行解码，分别得到所述当前播放视频对应的图像数据和声音数据；

基于所述图像数据调用图像识别接口进行图像识别，得到与所述图像数据对应的图像特征，并基于所述声音数据调用声音识别接口进行声音识别，得到与所述声音数据对应的声音特征。

3、根据权利要求 2 所述的方法，其中，所述声音还原元件依据所述视频显示屏预先划分的分区独立设置；

其中，所述分区的数量依据显示屏的尺寸进行设定。

4、根据权利要求 1 或 3 所述的方法，其中，所述声音还原元件包括扬声器。

5、根据权利要求 4 所述的方法，其中，所述根据所述发声源在所述视频显

示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声包括；

将所述声音数据通过杜比全景声 ATMOS 解码，得到集成音频接口 IIS 音频信号；

根据所述 IIS 音频信号以及所述控制信号控制与所述位置信息对应的功率放大器工作以驱动所述位置信息对应的扬声器发声。

6、根据权利要求 1 所述的方法，其中，所述基于所述图像特征确定所述当前播放视频存在发声源，包括：

将所述图像特征与预设图像特征数据库中的图像特征进行相似度匹配；

当匹配的所述相似度达到设定阈值时，则确定所述当前播放视频存在发声源；

其中，所述预设图像特征数据库中的图像特征包括下述至少一项：人体形态特征和动物形态特征。

7、根据权利要求 1 所述的方法，其中，所述基于所述声音特征切断所述当前播放视频存在与所述发声源匹配的音源，包括：

将所述声音特征与预先建立的模型发声源的模型特征进行比对；

若存在与所述声音特征一致的模型特征且所述模型特征对应的模型发声源与所述当前播放视频中存在的发声源相同，则确定所述当前播放视频存在与所述发声源匹配的音源。

8、根据权利要求 1 所述的方法，其中，所述控制与所述位置信息对应的声音还原元件发声，包括：

当所述声音还原元件发出的声音幅度超过设定上限时，降低所述声音还原元件的发声增益；

当所述声音还原元件发出的声音幅度没有超过所述设定下限时，提高所述声音还原元件的发声增益。

9、根据权利要求 1-8 任一项所述的方法，其中，所述当前播放视频通过按照单位时间内的预设采样次数对播放的视频进行采样而获得。

10、根据权利要求 1-8 任一项所述的方法，其中，所述视频显示屏包括预设尺寸的液晶显示屏。

11、一种实现声音与图像同位的装置，包括：

识别模块，设置为对当前播放视频分别进行图像识别以及声音识别，以获取所述当前播放视频对应的图像特征和声音特征；

获取模块，设置为在基于所述图像特征确定所述当前播放视频存在发声源的情况下，基于所述图像特征从预设图像特征数据库中获取所述当前播放视频的发声源在视频显示屏的位置信息；

控制模块，设置为在基于所述声音特征确定所述当前播放视频存在与所述发声源匹配的音源的情况下，根据所述发声源在所述视频显示屏的位置信息生成控制信号，以根据所述控制信号控制与所述位置信息对应的声音还原元件发声；

其中，所述预设图像特征数据库依据所述当前播放视频预先构建。

12、根据权利要求 11 所述的装置，其中，所述获取模块包括发声源确定子模块；所述发声源确定子模块设置为：

将所述图像特征与预设图像特征数据库中的图像特征进行相似度匹配；

当匹配的所述相似度达到设定阈值时，则确定所述当前播放视频存在发声源；

其中，所述预设图像特征数据库中的图像特征包括下述至少一项：人体形

态特征和动物形态特征。

13、根据权利要求 11 所述的装置，其中，所述控制模块包括音源确定子模块；所述音源确定子模块设置为：

将所述声音特征与预先建立的模型发声源的模型特征进行比对；

若存在与所述声音特征一致的模型特征且所述模型特征对应的模型发声源与所述当前播放视频中存在的发声源相同，则确定所述当前播放视频存在与所述发声源匹配的音源。

14、一种电子设备，包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现如权利要求 1-10 中任一项所述的方法。

15、一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时实现如权利要求 1-10 中任一项所述的方法。

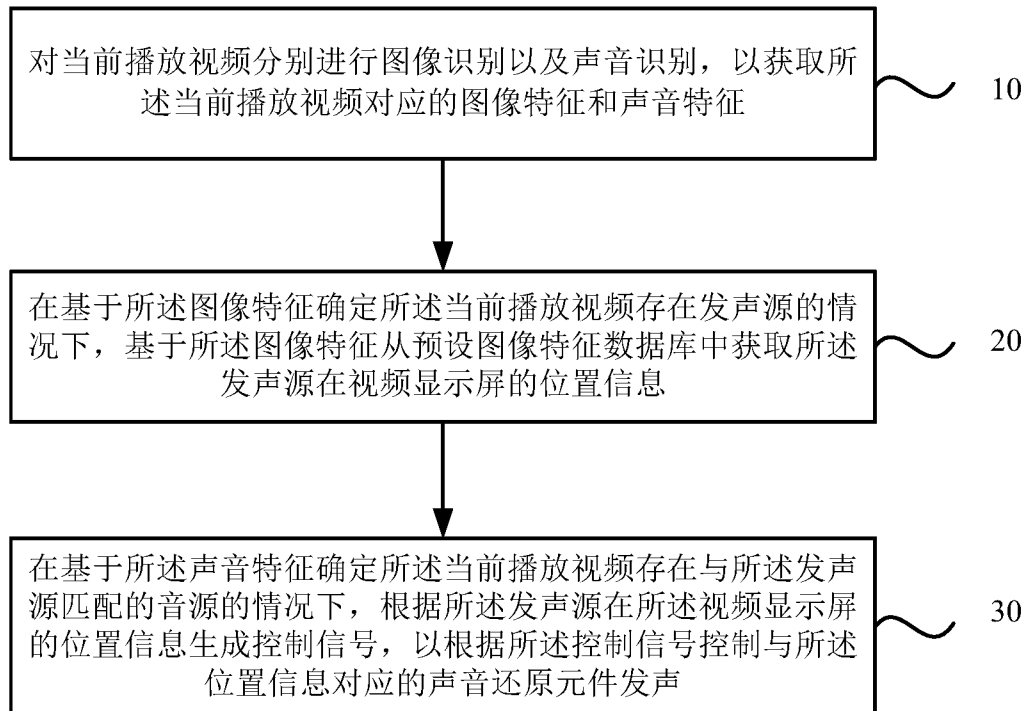


图 1a

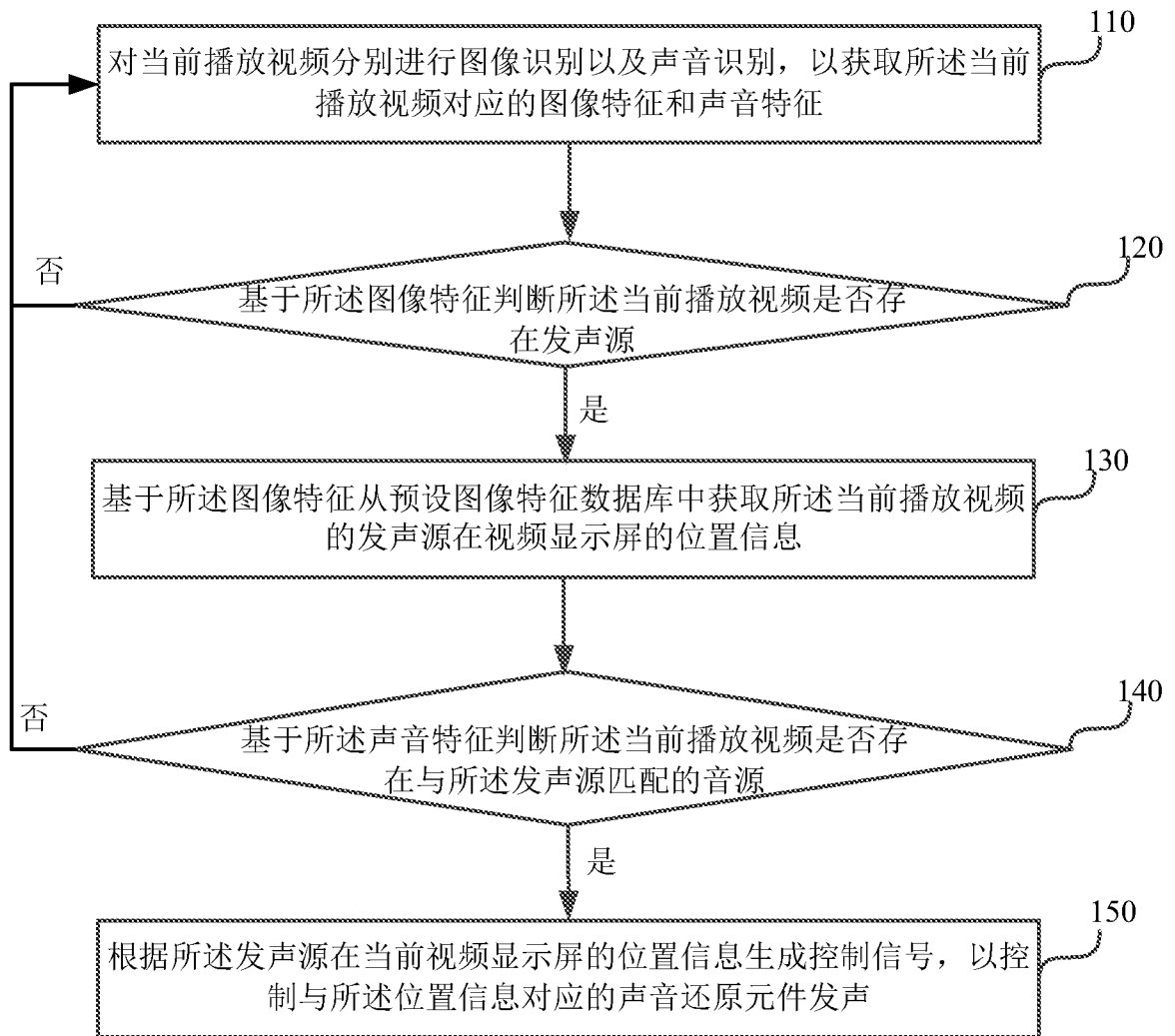


图 1b

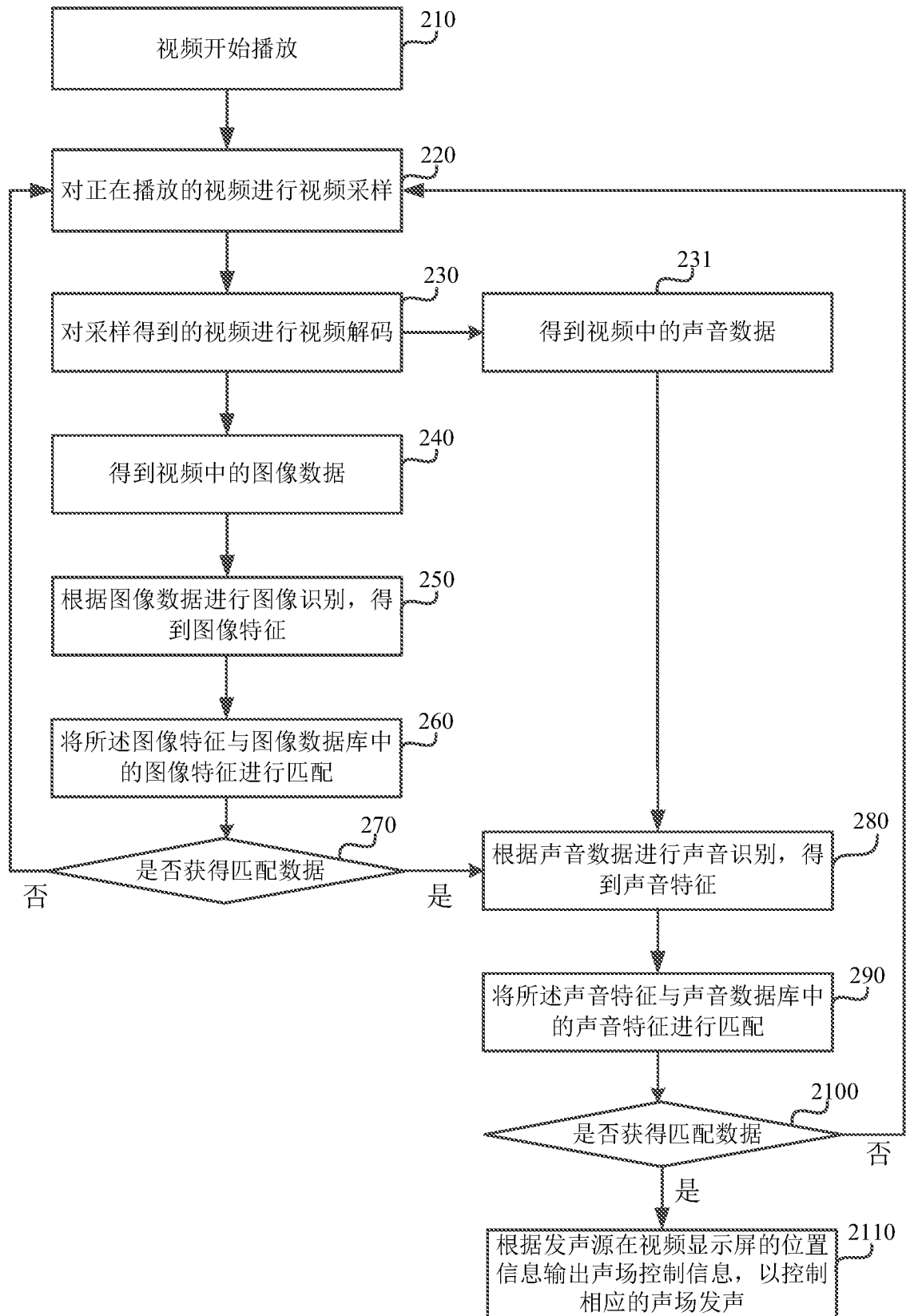


图 2

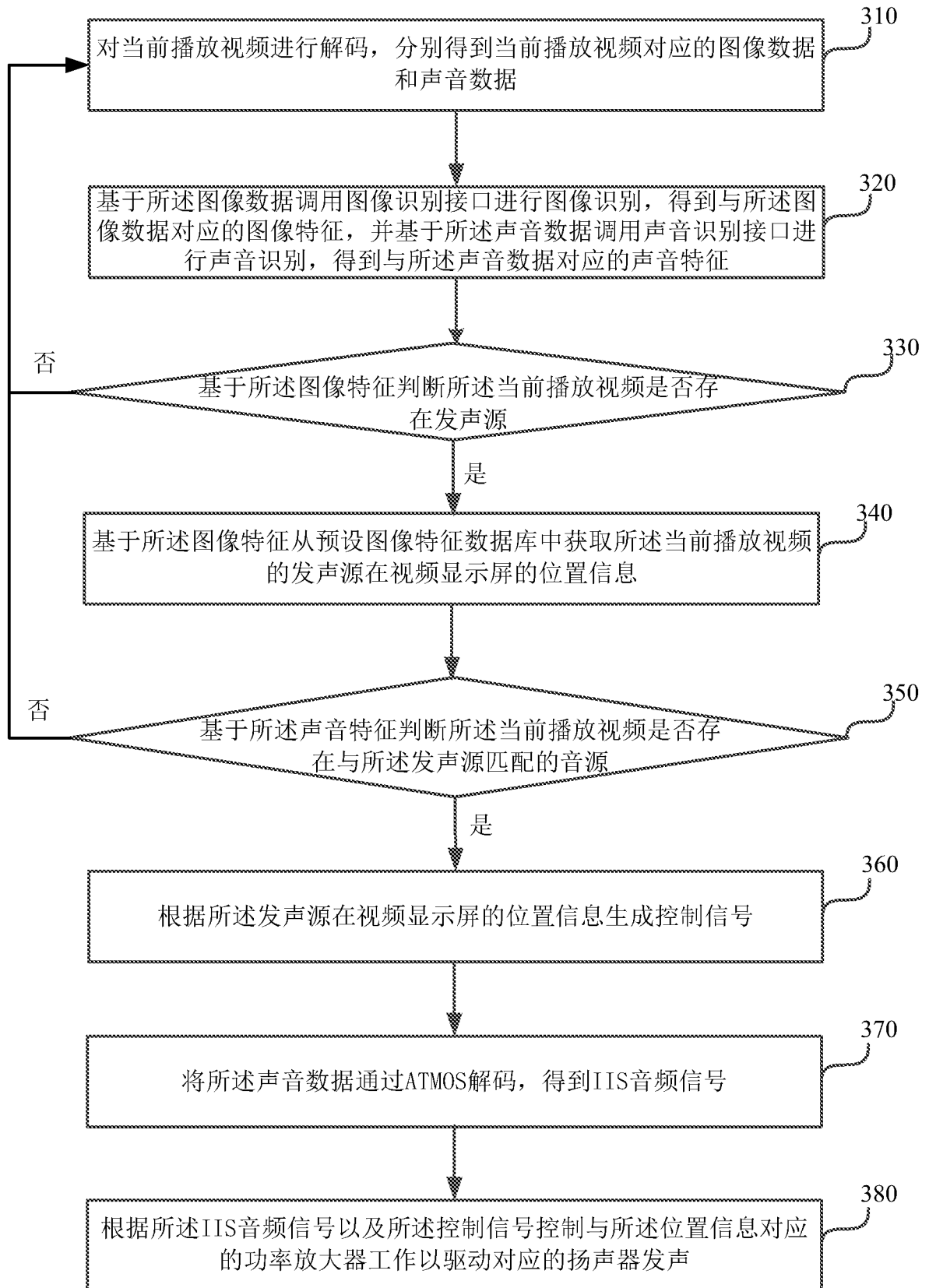


图 3



图 4

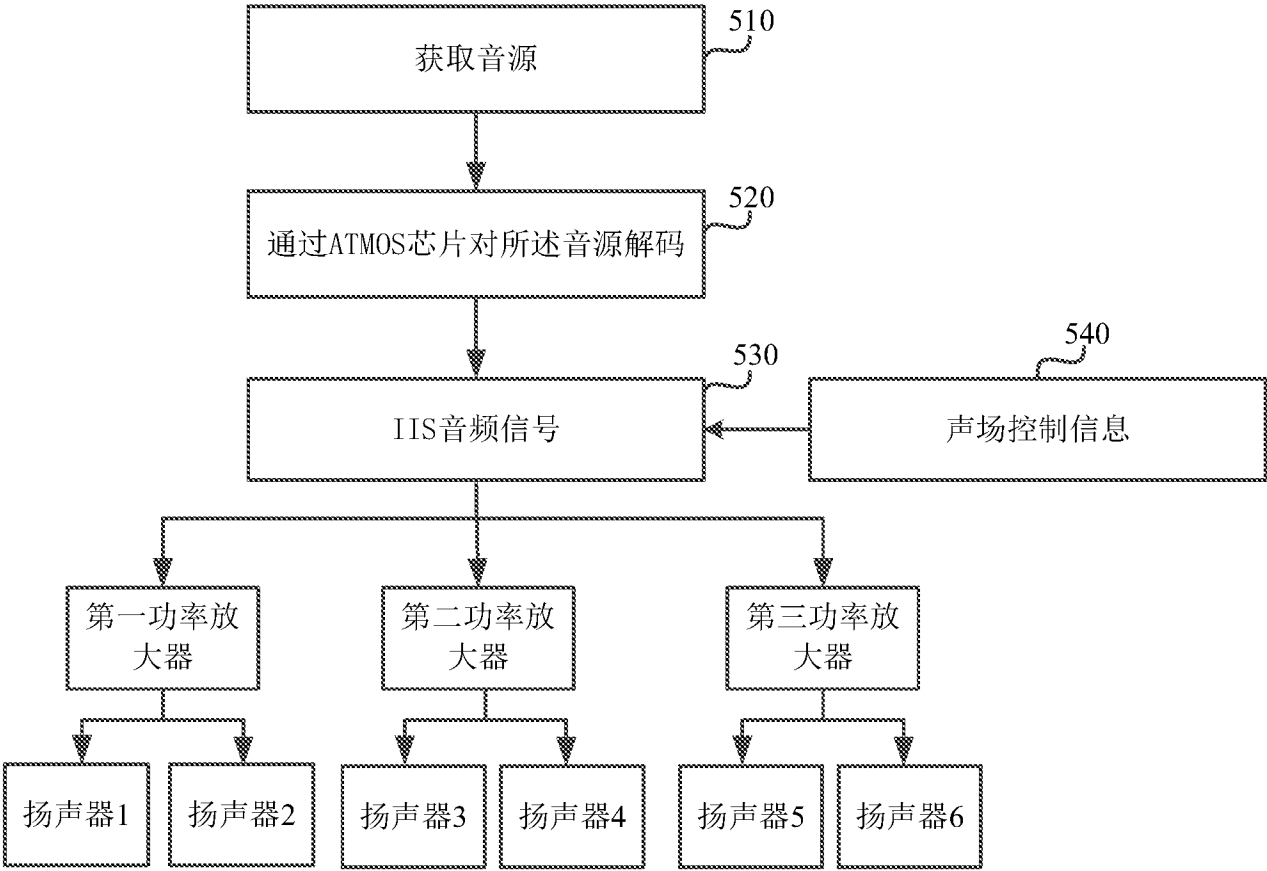


图 5

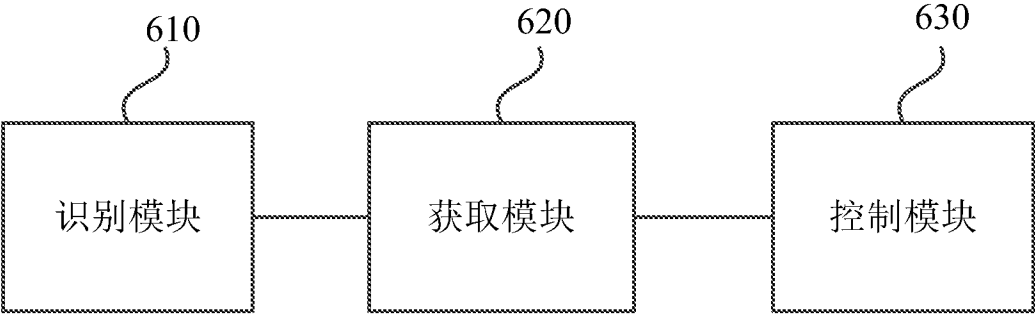


图 6

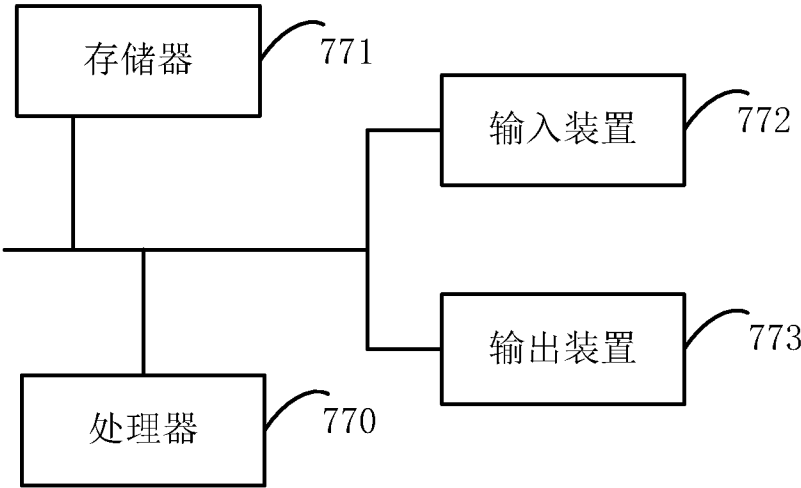


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/120528

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/43(2011.01)i; H04N 21/439(2011.01)i; H04N 21/44(2011.01)i; H04N 21/4402(2011.01)i; H04N 21/442(2011.01)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: 声源, 音源, 位置, 坐标, 识别, 图像, 视频, 声音, 发声; WPI, EPODOC, IEEE: sound, source, acoustic, position, coordinate, identify, distinguish, image, picture, video

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104036789 A (BEIJING ZHIGU RUITUO TECHNOLOGY SERVICE CO., LTD.) 10 September 2014 (2014-09-10) description, paragraphs 0002, 0029-0074, and 0122	1-15
A	CN 105979470 A (BEIJING QIYI CENTURY SCIENCE & TECHNOLOGY CO., LTD.) 28 September 2016 (2016-09-28) entire document	1-15
A	CN 101370107 A (SONY CORPORATION) 18 February 2009 (2009-02-18) entire document	1-15
A	CN 104270552 A (HUAWEI TECHNOLOGIES CO., LTD.) 07 January 2015 (2015-01-07) entire document	1-15
A	US 6829018 B2 (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 07 December 2004 (2004-12-07) entire document	1-15

☐ 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

24 May 2019

Date of mailing of the international search report

11 June 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (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/CN2018/120528

Patent document cited in search report				Publication date (day/month/year)		Patent family member(s)		Publication date (day/month/year)	
CN	104036789	A	10 September 2014	CN	104036789	B		02 February 2018	
				US	2016330512	A1		10 November 2016	
				US	9749684	B2		29 August 2017	
				WO	2015101247	A1		09 July 2015	
CN	105979470	A	28 September 2016	CN	105979470	B		16 April 2019	
CN	101370107	A	18 February 2009	JP	2007110582	A		26 April 2007	
				KR	20070042104	A		20 April 2007	
				CN	101370108	B		16 June 2010	
				JP	5067595	B2		07 November 2012	
				CN	101370108	A		18 February 2009	
				CN	101009792	A		01 August 2007	
				US	8483414	B2		09 July 2013	
				CN	100477762	C		08 April 2009	
				CN	101370107	B		19 October 2011	
				US	2007104341	A1		10 May 2007	
CN	104270552	A	07 January 2015	CN	106576132	A		19 April 2017	
				KR	20160119218	A		12 October 2016	
				WO	2016029806	A1		03 March 2016	
				US	2016065791	A1		03 March 2016	
US	6829018	B2	07 December 2004	US	2003053680	A1		20 March 2003	

国际检索报告

国际申请号

PCT/CN2018/120528

A. 主题的分类

H04N 21/43 (2011.01) i; H04N 21/439 (2011.01) i; H04N 21/44 (2011.01) i; H04N 21/4402 (2011.01) i; H04N 21/442 (2011.01) 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: 声源, 音源, 位置, 坐标, 识别, 图像, 视频, 声音, 发声; WPI, EP0D0C, IEEE: sound, source, acoustic, position, coordinate, identify, distinguish, image, picture, video

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
X	CN 104036789 A (北京智谷睿拓技术服务有限公司) 2014年 9月 10日 (2014 - 09 - 10) 说明书第0002、0029-0074、0122段	1-15
A	CN 105979470 A (北京奇艺世纪科技有限公司) 2016年 9月 28日 (2016 - 09 - 28) 全文	1-15
A	CN 101370107 A (索尼株式会社) 2009年 2月 18日 (2009 - 02 - 18) 全文	1-15
A	CN 104270552 A (华为技术有限公司) 2015年 1月 7日 (2015 - 01 - 07) 全文	1-15
A	US 6829018 B2 (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 2004年 12月 7日 (2004 - 12 - 07) 全文	1-15

☐ 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

2019年 5月 24日

国际检索报告邮寄日期

2019年 6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

授权官员

刘昶

传真号 (86-10)62019451

电话号码 53961699

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/120528

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	104036789	A	2014年 9月 10日	CN	104036789	B	2018年 2月 2日		
				US	2016330512	A1	2016年 11月 10日		
				US	9749684	B2	2017年 8月 29日		
				WO	2015101247	A1	2015年 7月 9日		
CN	105979470	A	2016年 9月 28日	CN	105979470	B	2019年 4月 16日		
CN	101370107	A	2009年 2月 18日	JP	2007110582	A	2007年 4月 26日		
				KR	20070042104	A	2007年 4月 20日		
				CN	101370108	B	2010年 6月 16日		
				JP	5067595	B2	2012年 11月 7日		
				CN	101370108	A	2009年 2月 18日		
				CN	101009792	A	2007年 8月 1日		
				US	8483414	B2	2013年 7月 9日		
				CN	100477762	C	2009年 4月 8日		
				CN	101370107	B	2011年 10月 19日		
				US	2007104341	A1	2007年 5月 10日		
CN	104270552	A	2015年 1月 7日	CN	106576132	A	2017年 4月 19日		
				KR	20160119218	A	2016年 10月 12日		
				WO	2016029806	A1	2016年 3月 3日		
				US	2016065791	A1	2016年 3月 3日		
US	6829018	B2	2004年 12月 7日	US	2003053680	A1	2003年 3月 20日		