

(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7月2日 (02.07.2020)

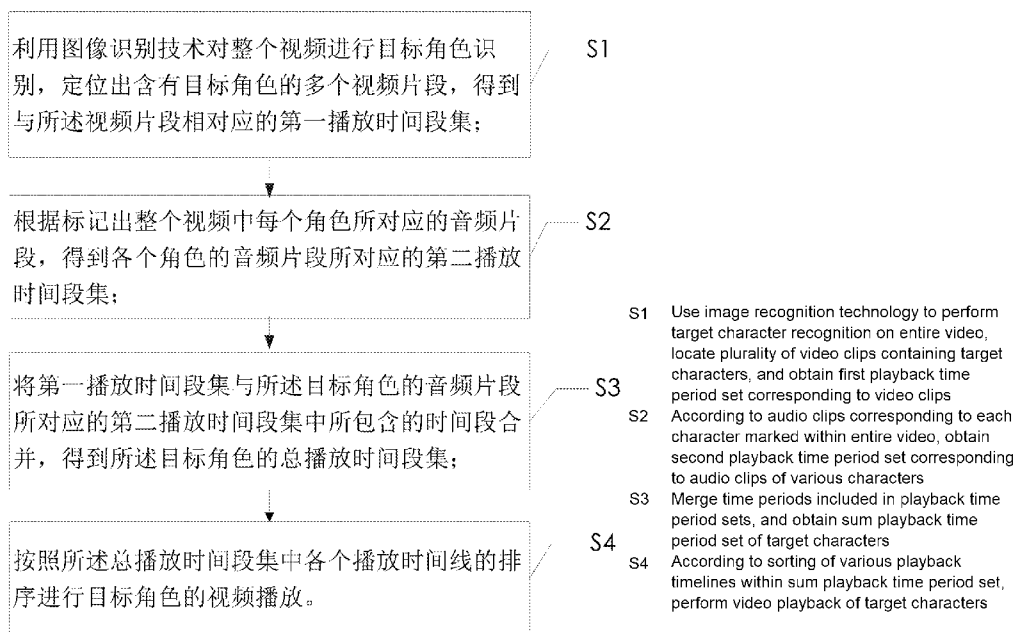


(10) 国际公布号
WO 2020/135643 A1

- (51) 国际专利分类号:
H04N 21/234 (2011.01) *H04N 21/44* (2011.01)
H04N 21/439 (2011.01)
- (21) 国际申请号: PCT/CN2019/128918
- (22) 国际申请日: 2019年12月27日 (27.12.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201811615807.0 2018年12月27日 (27.12.2018) CN
- (71) 申请人: 深圳 TCL 新技术有限公司 (SHENZHEN TCL NEW TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4栋9楼, Guangdong 518052 (CN)。
- (72) 发明人: 鲍舰 (BAO, Jian); 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4栋9楼, Guangdong 518052 (CN)。
- (74) 代理人: 深圳市君胜知识产权代理事务所 (普通合伙) (JOHNSON INTELLECTUAL PROPERTY AGENCY(SHENZHEN)); 中国广东省深圳市南山区粤海街道高新区社区高新南七道20号深圳国家工程实验室大楼A503, Guangdong 518000 (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,

(54) Title: TARGET CHARACTER VIDEO CLIP PLAYBACK METHOD, SYSTEM AND APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 目标角色视频片段播放方法、系统、装置及存储介质



(57) Abstract: Provided are a target character video clip playback method, system and apparatus, and a storage medium. The method comprises: using image recognition technology to perform target character recognition on an entire video, locating a plurality of video clips containing target characters, and obtaining a first playback time period set corresponding to the video clips; according to audio clips corresponding to each character marked within the entire video, obtaining a second playback time period set corresponding to the audio clips of the various characters; merging the time periods included in the playback time period sets, and obtaining a sum playback time period set of the target characters; according to a sorting of various playback timelines within the sum playback time period set, performing video playback of the target characters. The method combines image recognition technology and voiceprint

WO 2020/135643 A1

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))。

matching, thereby preventing misjudgement with respect to target character video clips, and increasing recognition accuracy, so as to make it easy for a user to watch images only containing target characters.

(57) 摘要: 本公开提供了一种目标角色视频片段播放方法、系统、装置及存储介质, 通过利用图像识别技术对整个视频进行目标角色识别, 定位出含有目标角色的多个视频片段, 得到与所述视频片段相对应的第一播放时间段集; 根据标记出整个视频中每个角色所对应的音频片段, 得到各个角色的音频片段所对应的第二播放时间段集; 将播放时间段集所包含的时间段合并, 得到所述目标角色的总播放时间段集; 按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。本公开所提供的方法, 将图像识别技术和声纹匹配相结合, 避免了对目标角色视频片段的误判, 提高了识别准确率, 为用户观看仅仅含有目标角色的图像提供了便利。

目标角色视频片段播放方法、系统、装置及存储介质

优先权

所述 PCT 专利申请要求申请日为 2018 年 12 月 27 日，申请号为 201811615807.0 的中国专利优先权，本专利申请结合了上述专利的技术方案。

技术领域

本公开涉及视频播放技术领域，尤其涉及的是一种目标角色视频片段播放方法、系统、装置及存储介质。

背景技术

用户在观看电视剧的时候，因为剧情的冗长会选择跳着看，而有一种视频点播的方式为针对自己喜欢的角色看，凡是没有这个角色的一概略过。而现有技术中使用的 AI 图像识别技术只能识别出演员的正面图像，且必须是清晰的，否则就无法识别，然而在实际的影视剧中该角色可能会与对手角色对话，镜头不断切换，实际的情境是目标角色在画面里面，但是由于想要观看的角色的正脸不在画面上，所以 AI 图像识别技术判断当前画面没有目标角色，会自动略过无该喜爱角色正脸的画面，但是这种画面实际是包含目标角色的，因此仅仅基于 AI 图像识别技术是无法得到准确的含有目标角色的剧情。

因此，现有技术有待于进一步的改进。

发明内容

鉴于上述现有技术中的不足之处，本公开的目的在于提供一种目标角色视频片段播放方法、系统、装置及存储介质，克服现有技术中仅仅使用图像识别技术无法准确从整个视频中提取出全部含有目标角色播放片段的缺陷。

本公开提供的第一实施例为一种目标角色视频片段播放方法，其中，包括：

利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；

根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；

将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所

包含的时间段合并，得到所述目标角色的总播放时间段集；

按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。

可选的，所述识别出含有目标角色的多个视频片段还包括：

当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；

若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段。

可选的，所述识别出含有目标角色的多个视频片段还包括：

当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；

若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段。

可选的，所述方法还包括：

对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

若含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中；

可选的，所述方法还包括：

对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

若不含有目标角色和/或含有目标角色的语音特征，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

可选的，所述方法还包括：

根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

其中，所述目标角色包含一个特定角色或包含多个特定角色。

可选的，所述将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集的步骤包括：

获取所述音频片段所对应的播放时间线和第一音频片段集所含音频片段所对应的播放时间线；

将提取出的音频片段与第一音频片段集中所含音频片段所对应的播放时间线重合的部分删除，得到融合后的第一音频片段集。

本公开提供的第二实施例为一种目标角色视频片段播放系统，其中，包括：

第一音频集提取模块，用于利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；

第二音频集获取模块，用于根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；

音频匹配模块，用于将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；

视频播放模块，用于按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。

可选的，所述第一音频集提取模块还包括：

定位检测单元，用于当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段；若含有，则继续对本次视频片段定位。

可选的，所述系统还包括：

延时检测模块，用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

以及，若含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中。

可选的，所述延时检测模块，还用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

若第二预设时间或者第三预设时间内不含有目标角色和/或含有目标角色的语音特征，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

可选的，所述第二音频集获取模块还包括：预标记单元；

所述预标记单元，用于根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

可选的，所述目标角色包含一个特定角色或包含多个特定角色。

本公开所提供的第三实施例为一种目标角色视频片段播放装置，包括：存储器和处理器，所述存储器存储有计算机程序，其中，所述处理器执行所述计算机程序时实现所述方法的步骤。

本公开所提供的第四实施例为一种存储介质，其中，所述存储介质上存储有播放频道的切换的控制程序，所述播放频道的切换的控制程序被处理器执行时实现所述的目标角色视频片段播放方法的步骤。

有益效果，本公开提供了一种目标角色视频片段播放方法、系统、装置及存储介质，通过利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。本公开所提供的方法，将图像识别技术和声纹匹配相结合，避免了对目标角色视频片段的误判，提高了识别准确率，为用户观看仅仅含有目标角色的图像提供了便利。

附图说明

图 1 是本公开所提供的一种目标角色视频片段播放方法的步骤流程图；

图 2 是本公开所提供的所述方法中视频场景切换示意图；

图 3 是本公开所提供的所述方法具体实施例的步骤流程图；

图 4 是本公开所提供的所述系统的原理结构示意图。

具体实施方式

为使本公开的目的、技术方案及优点更加清楚、明确，以下参照附图并举实施例对本公开进一步详细说明。应当理解，此处所描述的具体实施例仅仅用于解释本公开，并不用于限定本公开。

在实施方式和申请专利范围中，除非文中对于冠词有特别限定，否则“一”与“所述”

可泛指单一个或复数个。

另外，若本发明实施例中有涉及“第一”、“第二”等的描述，则该“第一”、“第二”等的描述仅用于描述目的，而不能理解为指示或暗示其相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。另外，各个实施例之间的技术方案可以相互结合，但是必须是以本领域普通技术人员能够实现为基础，当技术方案的结合出现相互矛盾或无法实现时应当认为这种技术方案的结合不存在，也不在本发明要求的保护范围之内。

实施例 1

本公开提供的第一实施例为一种目标角色视频片段播放方法，如图 1，包括：

步骤 S1、利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集。

当用户选定一个视频，发出仅仅观看该视频中指定的目标角色的视频片段后，首先利用图像识别技术对该视频的整个图像信息进行识别，对其中含有目标角色的视频片段进行定位，同时根据定位信息获取与视频片段相对应的音频文件。

具体的，所述目标角色可以包含一个特定角色也可以是多个特定角色。所述对视频片段进行定位为在检测到目标角色的图像帧处开始定位，在检测到目标角色消失的播放时间点结束定位，从而得到目标角色出现的时间段，定位出视频片段，当再次识别到目标角色时，开始定位计时，并同时获取相对应的音频信息，生成音频文件。从而当对整个视频进行识别后，得到对目标角色所在整个视频中出现的视频片段对应的播放时间段定位点和与视频片段相对应的音频片段。全部的音频片段组成第一音频片段集。

步骤 S2、根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集。

因为每个角色的语音特征不同，比如：音色和声纹等特征，因此根据每个角色的语音特征的不同，对整个视频中的每个角色所发声音的音频片段进行标记，得到含有各个角色音频片段的第二音频片段集。

步骤 S3、将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集。

根据目标角色的声纹特征将第二音频片段集中含有的与目标角色相关的音频片段提取出来，并将提取出来的音频片段添加到第一音频片段集中。

为了避免有目标角色的声音却没有目标角色画面在步骤 S1 中被漏定位，则本步骤中所述将提取出的音频片段融合到第一音频片段集中的步骤包括：

获取所述音频片段所对应的播放时间线和第一音频片段集所含音频片段所对应的播放时间线；

将提取出的音频片段与第一音频片段集中所含音频片段所对应的播放时间线重合的部分删除，得到融合后的第一音频片段集。

也即是通过将全部含有目标角色声音的音频片段所对应的播放时间段与第一音频片段集的播放时间段相整合，得到整个视频中既包含目标角色图像也包含目标角色声音的全部的视频片段。

步骤 S4、按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。

根据整合后的第一音频片段集中所含音频片段对应的时间线进行视频播放。

具体的，在步骤 S1 中为了避免目标角色短暂离开画面，导致定位出的视频场景画面缺失，所述识别出含有目标角色的多个视频片段还包括：

当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；

若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段；

若含有，则继续对本次视频片段定位。

也即是，当出现目标角色从有到无的切换时，对时间进行定时，判断在第一预设时间内目标角色是否再次出现，若无出现，则结束对本次视频片段的定位及音频文件的提取，若有出现，则持续对视频片段的定位，不间断本次音频片段的提取。

较佳的，为了避免目标角色定位与语音特征识别相结合后，仍然可能导致的视频片段之间缺失图像帧，所述方法还包括：

启动定时器，对识别或者判断时间进行计时，同时对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

若含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中；

若不含有，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

也即是对视频片段接下来时间内播放片段中是否含有目标角色的画面或者语音进

行再次筛查，若有画面，则将接下来的视频片段所对应的时间线添加到相连接的视频片段时间线上，若不含有目标角色的画面，但含有目标角色的音频信息，则同样将其所对应的时间线添加到视频片段时间线上，若没有目标角色的画面和声音，则略过本次视频片段后的第二预设时间段所对应的视频片段。

为了实现对整个视频中各个角色的语音信息进行准确标记，可以通过人工标记的方法，也可以使用语音特征自动匹配实现。所述方法还包括：

根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

本步骤可以先通过人工标记的方式，将目标角色的音频片段进行标记，然后利用语音识别装置根据人工标记出的音频片段所含有的语音特征对将整个目标角色的音频片段进行完全标记。

具体的，根据人工标记出的音频片段对所含有的语音特征进行识别时，可以通过人工标记出的音频片段中含有语音特征与视频中其他音频片段所含有的音频信号进行匹配，当匹配度超出 80%，则判断为匹配成，若低于 80%则认为不匹配。

下面结合图 3 所述，以具体实施例对本公开所提供的方法做更进一步的说明。

若某一个完整视频里面有若干个角色 A，B，C 等等，目前我们的需求是向用户提供只看目标角色 A 的影视精简版。

步骤 H1、先用图像识别技术对整个视频进行目标角色 A 图像的识别，提炼出如下信息：画面只有目标角色 A 的部分，记录只有目标角色 A 画面的播放时间段，将各个播放时间段组合成播放时间集合。较佳的，选择使用 AI 图像识别技术。

步骤 H2、对整个视频中的每一个角色人工标记其对应的音频片段的播放时间段位置，得到各个角色的音频片段所对应的播放时间段集；

步骤 H3、用将步骤 H1 得到的播放时间段集与步骤 H2 得到的播放时间段集中时间重复的部分删除后，根据时间线拼接在一起，得到目标角色 A 的总播放时间段集。

步骤 H4、根据目标角色 A 的音频文件所对应的总播放时间段集，按照时间从前到后的顺序进行播放。

为了避免部分未识别出一个场景中目标角色 A 和不含有目标角色 A 的声音的图像画面丢失，导致图像画面不连贯，在具体实施例中还包括：对视频片段接下来的一段时间内是否有 A 的画面和/或声音出现进行识别，下面结合图 2 对其进行说明：

用户选择仅仅观看目标角色 A 的场景，第 N 帧有：主角 A 的图像，主角 A 的声音：

配角 B 的图像，配角 B 的声音。而第 N+j 帧中，主角 A 的图像缺失，还有 Y 声音（疑似 A，但是不确定），有配角 B 的图像，主角 B 的声音。提取 Y 声音，与 AvoiceSet 中的声音进行匹配，有 80% 的声音特征在阈值 X 以内我们认为 Y 声音为主角 A 发出的，因此第 N+j 帧虽然没有 A 的图像，但是我们认为因为 A 的声纹存在，所以我们认为 A 存在，所以第 N+j 帧我们保留此帧。

进一步的，同时启动定时器 T_{max} ，从第 N+j 帧开始，到 T_{max} 时间内都无主角 A 的画面，且没有收到 A 的声音，我们就认为 A 已经真正离开了画面，从此刻开始，将画面切断，重新开始聆听下一个 A 的声音或者 A 的画面出现。

实施例 2

本公开提供的第二实施例为一种目标角色视频片段播放系统，如图 4 所示，包括：
第一音频集提取模块 410，用于利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；其功能如步骤 S1 所述。

第二音频集获取模块 420，用于根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；其功能如步骤 S2 所述。

音频匹配模块 430，用于将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；其功能如步骤 S3 所述。

视频播放模块 440，用于按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。其功能如步骤 S4 所述。

可选的，所述第一音频集提取模块 410 还包括：

定位检测单元，用于当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段；若含有，则继续对本次视频片段定位。

具体的，为了避免定位出的视频片段中图像信息的缺失，还包括：

延时检测模块，用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别；若含有，则将第二预设时间段内的画面添加到所述视频片段中；若不含有，则对所述画面所对应的语音文件所含有的语音特征与目标角色的语音特征进行匹配；若相似度超出预设阈值，则将预设时间段内的画面添加到所述视频片段中。

较佳的，为了实现目标角色所在图像的完整性，所述系统还包括：

延时检测模块，用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

以及，若含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中；

若不含有，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

由于每个角色的语音特征信息是已知，则可以先在音频文件中对其特征进行标记，然后系统自动根据标记出的语音特征信息，将整个视频所对应的语音片段中所有的和目标角色相似的语音信息进行标记。因此所述第二音频集获取模块还包括：预标记单元；

所述预标记单元，用于根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

实施例 3

本公开所提供的第三实施例为一种目标角色视频片段播放装置，包括：存储器和处理器，所述存储器存储有计算机程序，所述处理器执行所述计算机程序时实现所述方法的步骤。

所述目标角色视频片段播放装置可以包括以下一个或多个组件：处理器，存储器，电源组件，多媒体组件，音频组件，输入/输出（I/O）的接口，传感器组件，以及通信组件。

处理器通常控制装置的整体操作，诸如与显示，电话呼叫，数据通信，相机操作和记录操作相关联的操作。处理器可以完成上述的方法的全部或部分步骤。此外，处理器可以包括一个或多个模块，便于处理器和其他组件之间的交互。例如，处理器可以包括多媒体模块，以方便多媒体组件和处理器之间的交互。所述处理器在一些实施例中可以是一中央处理器（Central Processing Unit, CPU），微处理器或其他数据处理芯片，用于运行所述存储器中存储的程序代码或处理数据，例如执行所述目标角色视频片段播放方法等。

存储器被配置为存储各种类型的数据以支持在设备的操作。这些数据的示例包括用于在装置上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。存储器可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编

程只读存储器 (EPROM), 可编程只读存储器 (PROM), 只读存储器 (ROM), 磁存储器, 快闪存储器, 磁盘或光盘。

所述存储器在一些实施例中可以是播放装置的内部存储单元, 例如智能电视的硬盘或内存。所述存储器在另一些实施例中也可以是所述播放装置的外部存储设备, 例如所述智能电视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, 所述存储器还可以既包括所述播放装置的内部存储单元也包括外部存储设备。所述存储器用于存储安装于所述播放装置的应用软件及各类数据, 例如所述安装智能电视的程序代码等。所述存储器还可以用于暂时地存储已经输出或者将要输出的数据。在一实施例中, 存储器上存储有目标角色视频片段播放的控制程序, 该基于目标角色视频片段播放的控制程序可被处理器所执行, 从而实现本申请中目标角色视频片段播放方法。

电源组件为装置的各种组件提供电力。电源组件可以包括电源管理系统, 一个或多个电源, 及其他与为装置生成、管理和分配电力相关联的组件。

多媒体组件包括在所述装置和用户之间的提供一个输出接口的屏幕。在一些实施例中, 屏幕可以包括液晶显示器 (LCD) 和触摸面板 (TP)。如果屏幕包括触摸面板, 屏幕可以被实现为触摸屏, 以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界, 而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中, 多媒体组件包括一个前置摄像头和/或后置摄像头。当设备处于操作模式, 如拍摄模式或视频模式时, 前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件被配置为输出和/或输入音频信号。例如, 音频组件包括一个麦克风 (MIC), 当装置处于操作模式, 如呼叫模式、记录模式和语音识别模式时, 麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器或经由通信组件发送。在一些实施例中, 音频组件还包括一个扬声器, 用于输出音频信号。

I/O 接口为处理组件和外围接口模块之间提供接口, 上述外围接口模块可以是键盘, 点击轮, 按钮等。这些按钮可包括但不限于: 主页按钮、音量按钮、启动按钮和锁定按钮。

传感器组件包括一个或多个传感器, 用于为装置提供各个方面的状态评估。例如传感器组件可以检测到设备的打开/关闭状态, 组件的相对定位, 例如所述组件为装置的显

示器和小键盘，传感器组件还可以检测装置或装置一个组件的位置改变，用户与装置接触的存在或不存在，装置方位或加速/减速和装置的温度变化。传感器组件可以包括接近传感器，被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件还可以包括光传感器，如 CMOS 或 CCD 图像传感器，用于在成像应用中使用。在一些实施例中，该传感器组件还可以包括加速度传感器，陀螺仪传感器，磁传感器，压力传感器或温度传感器。

通信组件被配置为便于装置和其他设备之间有线或无线方式的通信。装置可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信部件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信部件还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

在示例性实施例中，装置可以被一个或多个应用专用集成电路（ASIC）、数字信号处理器（DSP）、数字信号处理设备（DSPD）、可编程逻辑器件（PLD）、现场可编程门阵列（FPGA）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

实施例 4

本公开提供的第四实施例为一种存储介质，其中，所述存储介质上存储有播放频道的切换的控制程序，所述播放频道的切换的控制程序被处理器执行时实现所述的目标角色视频片段播放方法的步骤。

存储介质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；存储数据区可存储工业设备管理方法使用或接收的数据等。此外，存储介质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。

上述计算机设备可执行本公开实施例所提供的方法，具备执行该方法相应的功能模块和得到相应的有益效果。未在本实施例中详尽描述的技术细节，可参见本公开实施例所提供的方法。

有益效果，本公开提供了一种目标角色视频片段播放方法、系统、装置及存储介质，通过利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；将第一播

放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。本公开所提供的方法，将图像识别技术和声纹匹配相结合，避免了对目标角色视频片段的误判，提高了识别准确率，为用户观看仅仅含有目标角色的图像提供了便利。

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，是可以通过计算机程序来指令相关硬件（如处理器，控制器等）来完成，所述的程序可存储于一计算机可读的存储介质中，所述程序在执行时可包括如上述各方法实施例的流程。其中所述的存储介质可为存储器、磁碟、光盘等。

可以理解的是，对本领域普通技术人员来说，可以根据本公开的技术方案及其发明构思加以等同替换或改变，而所有这些改变或替换都应属于本公开所附的权利要求的保护范围。

1、一种目标角色视频片段播放方法，其中，包括：

利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；

根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；

将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；

按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。

2、根据权利要求1所述的目标角色视频片段播放方法，其中，所述识别出含有目标角色的多个视频片段还包括：

当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；

若含有，则继续对本次视频片段定位。

3、根据权利要求1所述的目标角色视频片段播放方法，其中，所述识别出含有目标角色的多个视频片段还包括：

当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；

若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段。

4、根据权利要求1所述的目标角色视频片段播放方法，其中，所述方法还包括：

对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征进行判断；

若第二预设时间或者第三预设时间内含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中。

5、根据权利要求1所述的目标角色视频片段播放方法，其中，所述方法还包括：

对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征进行判断；

若第二预设时间或者第三预设时间内不含有目标角色和/或含有目标角色的语音特

征，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

6、根据权利要求2所述的目标角色视频片段播放方法，其中，所述方法还包括：

根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

7、根据权利要求1-6任一项所述的目标角色视频片段播放方法，其中，所述目标角色包含一个特定角色或包含多个特定角色。

8、根据权利要求1所述的目标角色视频片段播放方法，其中，所述将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集的步骤包括：

获取所述音频片段所对应的播放时间线和第一音频片段集所含音频片段所对应的播放时间线；

将提取出的音频片段与第一音频片段集中所含音频片段所对应的播放时间线重合的部分删除，得到融合后的第一音频片段集。

9、一种目标角色视频片段播放系统，其中，包括：

第一音频集提取模块，用于利用图像识别技术对整个视频进行目标角色识别，定位出含有目标角色的多个视频片段，得到与所述视频片段相对应的第一播放时间段集；

第二音频集获取模块，用于根据标记出整个视频中每个角色所对应的音频片段，得到各个角色的音频片段所对应的第二播放时间段集；

音频匹配模块，用于将第一播放时间段集与所述目标角色的音频片段所对应的第二播放时间段集中所包含的时间段合并，得到所述目标角色的总播放时间段集；

视频播放模块，用于按照所述总播放时间段集中各个播放时间线的排序进行目标角色的视频播放。

10、根据权利要求9所述的目标角色视频片段播放系统，其中，所述第一音频集提取模块还包括：

定位检测单元，用于当检测到视频画面从有目标角色切换到无目标角色，则持续检测在接下来的第一预设时间内是否含有目标角色；若不含有，则结束本次视频片段定位，并记录本次视频片段所对应的播放时间段；若含有，则继续对本次视频片段定位。

11、根据权利要求9所述的目标角色视频片段播放系统，其中，还包括：

延时检测模块，用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含

有目标角色的语音特征判断；若含有目标角色和/或含有目标角色的语音特征，则将第二预设时间段和/或第三预设时间段内对应的播放时间段添加到所述总播放时间段集中。

12、根据权利要求9所述的目标角色视频片段播放系统，其中，所述延时检测模块，还用于对各个视频片段接下来的第二预设时间段内画面中是否含有目标角色进行识别和/或对各个音频片段接下来的第三预设时间段内所对应的音频是否含有目标角色的语音特征判断；

若第二预设时间或者第三预设时间内不含有目标角色和/或含有目标角色的语音特征，则略过接下来的第二预设时间段和第三预设时间段所对应的播放时间段。

13、根据权利要求9所述的目标角色视频片段播放系统，其中，所述第二音频集获取模块还包括：预标记单元；

所述预标记单元，用于根据各个角色的语音特征，选择整个视频的多个音频片段对角色语音进行标记，并根据标记出的语音特征对整个视频内的每个音频片段所对应的角色均进行标记。

14、根据权利要求9所述的目标角色视频片段播放系统，其中，所述目标角色包含一个特定角色或包含多个特定角色。

15、一种目标角色视频片段播放装置，包括：存储器和处理器，所述存储器存储有计算机程序，其中，所述处理器执行所述计算机程序时实现权利要求1至8中任一项所述方法的步骤。

16、一种存储介质，其中，所述存储介质上存储有播放频道的切换的控制程序，所述播放频道的切换的控制程序被处理器执行时实现如权利要求1至8中任一项所述的目标角色视频片段播放方法的步骤。

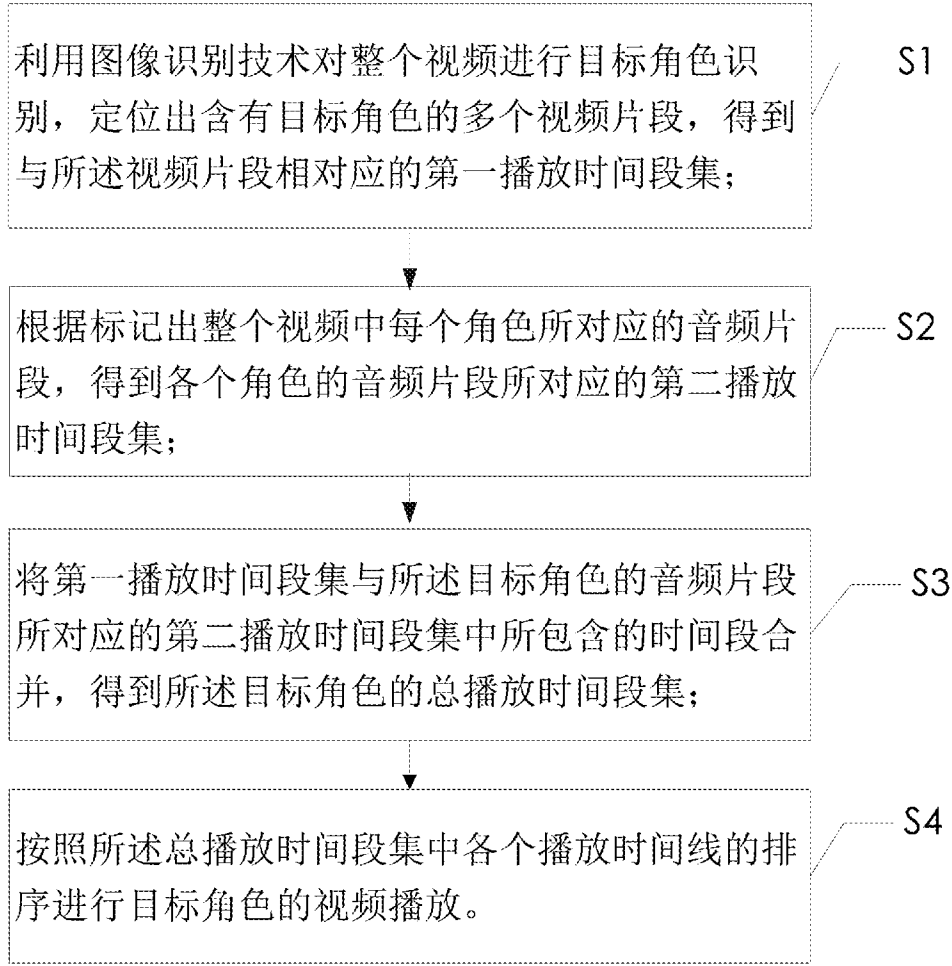


图 1

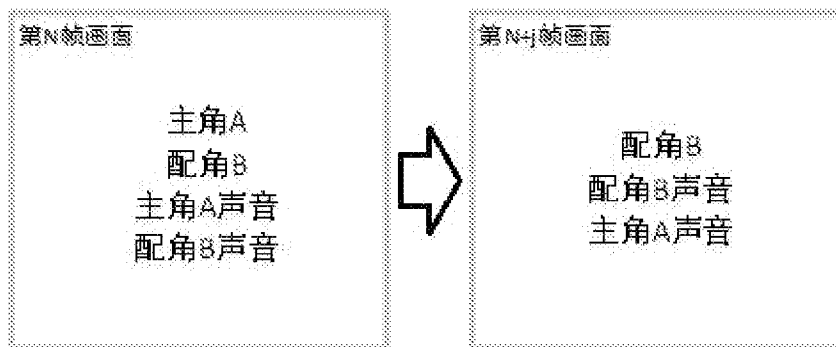


图 2

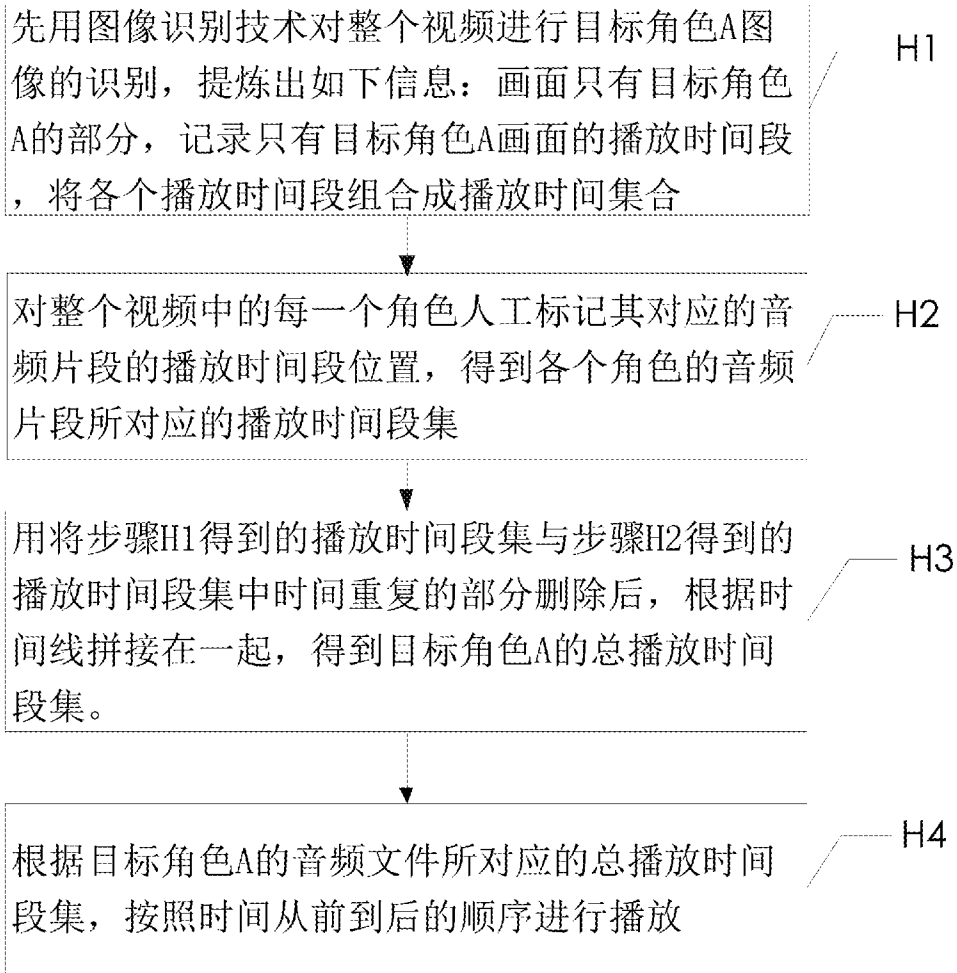


图 3

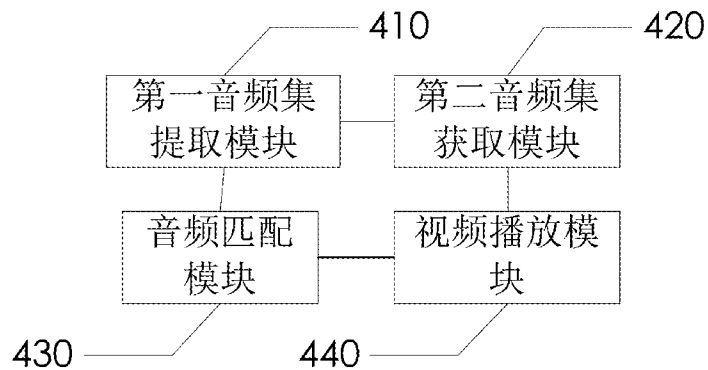


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/128918

A. CLASSIFICATION OF SUBJECT MATTER		
H04N 21/234(2011.01)i; H04N 21/439(2011.01)i; H04N 21/44(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, WPI, EPODOC, IEEE: 视频, 播放, 目标, 特定, 角色, 演员, 人物, 识别, 音频, 声音, 片段, 时间, 阈值, video, play, target, character, role, actor, voice, audio, clip, segment, fragment, time, threshold		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108337532 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 27 July 2018 (2018-07-27) claims 20-21, and description, paragraphs 0170-0205 and 0234-0246	1, 4-9, 11-16
Y	CN 108337532 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 27 July 2018 (2018-07-27) claims 20-21, and description, paragraphs 0170-0205 and 0234-0246	2-3, 10
Y	CN 106021496 A (HISENSE GROUP CO., LTD.) 12 October 2016 (2016-10-12) description, paragraphs 0117-0118	2-3, 10
X	CN 105100892 A (NUBIA TECHNOLOGY CO., LTD.) 25 November 2015 (2015-11-25) description, paragraphs 0135-0144	1, 4-9, 11-16
Y	CN 105100892 A (NUBIA TECHNOLOGY CO., LTD.) 25 November 2015 (2015-11-25) description, paragraphs 0135-0144	2-3, 10
A	CN 108271069 A (BEIJING QIYI CENTURY SCIENCE AND TECHNOLOGY CO., LTD.) 10 July 2018 (2018-07-10) entire document	1-16
A	CN 104796781 A (XIAOMI TECHNOLOGY CO., LTD.) 22 July 2015 (2015-07-22) entire document	1-16
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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		Date of mailing of the international search report
16 March 2020		27 March 2020
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088 China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/128918

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108337532	A	27 July 2018	WO	2019157977	A1	22 August 2019
CN	106021496	A	12 October 2016	None			
CN	105100892	A	25 November 2015	CN	105100892	B	15 May 2018
CN	108271069	A	10 July 2018	None			
CN	104796781	A	22 July 2015	CN	104796781	B	18 January 2019
US	2015082349	A1	19 March 2015	CA	2924065	A1	19 March 2015
				MX	349609	B	04 August 2017
				CA	2924065	C	15 May 2018
				WO	2015038749	A1	19 March 2015
				BR	112016006860	A2	01 August 2017
				MX	2016003315	A	09 May 2016
				US	9888279	B2	06 February 2018

国际检索报告

国际申请号

PCT/CN2019/128918

<p>A. 主题的分类</p> <p>H04N 21/234(2011.01)i; H04N 21/439(2011.01)i; H04N 21/44(2011.01)i</p> <p>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</p>																										
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, CNKI, WPI, EPDOC, IEEE: 视频, 播放, 目标, 特定, 角色, 演员, 人物, 识别, 音频, 声音, 片段, 时间, 阈值, video, play, target, character, role, actor, voice, audio, clip, segment, fragment, time, threshold</p>																										
<p>C. 相关文件</p> <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 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段</td> <td>1, 4-9, 11-16</td> </tr> <tr> <td>Y</td> <td>CN 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段</td> <td>2-3, 10</td> </tr> <tr> <td>Y</td> <td>CN 106021496 A (海信集团有限公司) 2016年 10月 12日 (2016 - 10 - 12) 说明书第0117-0118段</td> <td>2-3, 10</td> </tr> <tr> <td>X</td> <td>CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段</td> <td>1, 4-9, 11-16</td> </tr> <tr> <td>Y</td> <td>CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段</td> <td>2-3, 10</td> </tr> <tr> <td>A</td> <td>CN 108271069 A (北京奇艺世纪科技有限公司) 2018年 7月 10日 (2018 - 07 - 10) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 104796781 A (小米科技有限责任公司) 2015年 7月 22日 (2015 - 07 - 22) 全文</td> <td>1-16</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 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段	1, 4-9, 11-16	Y	CN 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段	2-3, 10	Y	CN 106021496 A (海信集团有限公司) 2016年 10月 12日 (2016 - 10 - 12) 说明书第0117-0118段	2-3, 10	X	CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段	1, 4-9, 11-16	Y	CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段	2-3, 10	A	CN 108271069 A (北京奇艺世纪科技有限公司) 2018年 7月 10日 (2018 - 07 - 10) 全文	1-16	A	CN 104796781 A (小米科技有限责任公司) 2015年 7月 22日 (2015 - 07 - 22) 全文	1-16
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
X	CN 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段	1, 4-9, 11-16																								
Y	CN 108337532 A (腾讯科技深圳有限公司) 2018年 7月 27日 (2018 - 07 - 27) 权利要求20-21, 说明书第0170-0205、0234-0246段	2-3, 10																								
Y	CN 106021496 A (海信集团有限公司) 2016年 10月 12日 (2016 - 10 - 12) 说明书第0117-0118段	2-3, 10																								
X	CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段	1, 4-9, 11-16																								
Y	CN 105100892 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第0135-0144段	2-3, 10																								
A	CN 108271069 A (北京奇艺世纪科技有限公司) 2018年 7月 10日 (2018 - 07 - 10) 全文	1-16																								
A	CN 104796781 A (小米科技有限责任公司) 2015年 7月 22日 (2015 - 07 - 22) 全文	1-16																								
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 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</p>																										
<p>国际检索实际完成的日期</p> <p>2020年 3月 16日</p>		<p>国际检索报告邮寄日期</p> <p>2020年 3月 27日</p>																								
<p>ISA/CN的名称和邮寄地址</p> <p>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</p>		<p>授权官员</p> <p>龙玄耀</p> <p>电话号码 86-(10)-53961698</p>																								

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
A	US 2015082349 A1 (ARRIS ENTERPRISES, INC.) 2015年 3月 19日 (2015 - 03 - 19) 全文	1-16

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/128918

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	108337532	A	2018年 7月 27日	WO	2019157977	A1	2019年 8月 22日
CN	106021496	A	2016年 10月 12日	无			
CN	105100892	A	2015年 11月 25日	CN	105100892	B	2018年 5月 15日
CN	108271069	A	2018年 7月 10日	无			
CN	104796781	A	2015年 7月 22日	CN	104796781	B	2019年 1月 18日
US	2015082349	A1	2015年 3月 19日	CA	2924065	A1	2015年 3月 19日
				MX	349609	B	2017年 8月 4日
				CA	2924065	C	2018年 5月 15日
				WO	2015038749	A1	2015年 3月 19日
				BR	112016006860	A2	2017年 8月 1日
				MX	2016003315	A	2016年 5月 9日
US	9888279	B2	2018年 2月 6日				