



(12)发明专利

(10)授权公告号 CN 107948729 B

(45)授权公告日 2020.03.27

(21)申请号 201711332691.5

H04N 21/845(2011.01)

(22)申请日 2017.12.13

(56)对比文件

(65)同一申请的已公布的文献号

CN 107392666 A, 2017.11.24, 1-3, 8-10.

申请公布号 CN 107948729 A

CN 104469487 A, 2015.03.25, 1-3, 8-10.

(43)申请公布日 2018.04.20

CN 102163201 A, 2011.08.24, 全文.

(73)专利权人 OPPO广东移动通信有限公司

CN 101547346 A, 2009.09.30, 全文.

地址 523860 广东省东莞市长安镇乌沙海滨路18号

CN 104135705 A, 2014.11.05, 全文.

(72)发明人 董治

CN 104320670 A, 2015.01.28, 全文.

(74)专利代理机构 广州华进联合专利商标代理有限公司 44224

US 2016085762 A1, 2016.03.24, 全文.

代理人 方高明

审查员 郭超

(51)Int.Cl.

H04N 21/439(2011.01)

权利要求书2页 说明书14页 附图9页

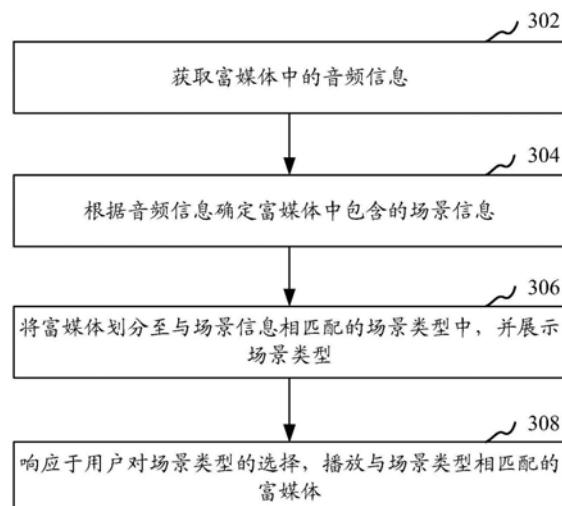
H04N 21/472(2011.01)

(54)发明名称

富媒体处理方法、装置、存储介质和电子设备

(57)摘要

本申请涉及一种富媒体处理方法、装置、存储介质和电子设备。该方法包括：获取富媒体中的音频信息；根据所述音频信息确定所述富媒体中包含的场景信息；将所述富媒体划分至与所述场景信息相匹配的场景类型中，并展示所述场景类型；响应于用户对所述场景类型的选择，播放与所述场景类型相匹配的富媒体。上述富媒体处理方法、装置、存储介质和电子设备可保持提高富媒体处理的灵活性。



1. 一种富媒体处理方法,包括:

获取富媒体中的音频信息;

根据所述音频信息确定所述富媒体中包含的场景信息;

将所述富媒体划分至与所述场景信息相匹配的场景类型中,并展示所述场景类型;

从所述音频信息中提取与所述场景信息相匹配的音频片段,根据所述音频片段形成音频文件;

响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的音频文件。

2. 根据权利要求1所述的方法,其特征在于,所述场景信息包括声音的强弱信息和/或声音所属主题的信息;所述根据所述音频信息确定所述富媒体中包含的场景信息,包括:

对所述音频信息进行音频内容识别;

根据识别出的音频内容判定所述富媒体中的声音信号的强弱;和/或

根据识别出的音频内容判定所述富媒体中的声音所属的主题;

将所述富媒体划分至与所述场景信息相匹配的场景类型中,包括:

将所述富媒体划分至与判定结果相匹配的场景类型中。

3. 根据权利要求2所述的方法,其特征在于,所述将所述富媒体划分至与判定结果相匹配的场景类型中,包括:

将所述富媒体划分至与所述音频内容中信号最强的声音对应的场景类型中;和/或

将所述富媒体划分至与所述主题相匹配的场景类型中。

4. 根据权利要求1所述的方法,其特征在于,在所述将所述富媒体划分至与所述场景信息相匹配的场景类型中,并展示所述场景类型之后,还包括:

对所述富媒体进行视频分离;

根据分离出的视频信息形成视频文件;

所述播放与所述场景类型相匹配的音频文件,包括:播放所述视频文件。

5. 根据权利要求1至4中任一项所述的方法,其特征在于,在所述响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的音频文件之前,还包括:

对所述富媒体设置用于标记所述场景类型的类型标记;

所述响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的音频文件,包括:

接收作用于所述类型标记而触发的播放指令;

根据所述播放指令进入与所述场景类型对应的场景的播放画面,并进行播放。

6. 根据权利要求5所述的方法,其特征在于,所述根据所述播放指令进入与所述场景类型对应的场景的播放画面,并进行播放,包括:

根据所述播放指令获取与所述场景类型对应音频内容在所述音频信息中的起始位置;

根据所述起始位置确定进入的播放画面;

获取本机所处环境的环境音量;

根据所述环境音量和所述场景类型确定所述富媒体的播放音量;

根据所述播放音量对进入的播放画面进行播放。

7. 一种富媒体处理装置,其特征在于,所述装置包括:

音频信息获取模块,用于获取富媒体中的音频信息;

场景信息识别模块,用于根据所述音频信息确定所述富媒体中包含的场景信息;

分类模块,用于将所述富媒体划分至与所述场景信息相匹配的场景类型中,并展示所述场景类型;

音频文件生成模块,用于从所述音频信息中提取与场景信息相匹配的音频片段,根据所述音频片段形成音频文件;

播放模块,用于响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的音频文件。

8.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至6中任一项所述方法的步骤。

9.一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至6中任一项所述方法的步骤。

富媒体处理方法、装置、存储介质和电子设备

技术领域

[0001] 本申请涉及数据处理技术领域,特别是涉及一种富媒体处理方法、装置、存储介质和电子设备。

背景技术

[0002] 随着拍摄功能的普及,越来越多的用户通过带有拍摄功能的终端随时随地的记录周围的场景,或者进行自拍,形成视频。用户通常将拍摄的视频通过诸如即时通信等类型的应用发送给好友或其它用户。

[0003] 用户终端在对接收到的视频或其它具有声音的动画信息的富媒体,或者自身相册中的富媒体进行点击播放时,通常是按照上一次的终端的音量或者既定音量来播放该富媒体。然而,由于在播放之时,终端并不能知道现场的具体环境,因而会存在在安静的环境中以较大的音量的播放该富媒体的情况,对周边的环境造成影响;或还会存在在较为嘈杂的环境中以较小的音量播放该富媒体的情况,又使得难以听清富媒体中的具体声音。

发明内容

[0004] 本申请实施例提供一种富媒体处理方法、装置、存储介质和电子设备,可以提高富媒体处理的灵活性。

[0005] 一种富媒体处理方法,包括:

[0006] 获取富媒体中的音频信息;

[0007] 根据所述音频信息确定所述富媒体中包含的场景信息;

[0008] 将所述富媒体划分至与所述场景信息相匹配的场景类型中,并展示所述场景类型;

[0009] 响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的富媒体。

[0010] 一种富媒体处理装置,所述装置包括:

[0011] 音频信息获取模块,用于获取富媒体中的音频信息;

[0012] 场景信息识别模块,用于根据所述音频信息确定所述富媒体中包含的场景信息;

[0013] 分类模块,用于将所述富媒体划分至与所述场景信息相匹配的场景类型中,并展示所述场景类型;

[0014] 播放模块,用于响应于用户对所述场景类型的选择,播放与所述场景类型相匹配的富媒体。

[0015] 一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现本申请实施例中任一项所述方法的步骤。

[0016] 一种电子设备,包括存储器、处理器及存储在存储器上并可在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本申请实施例中任一项所述方法的步骤。

[0017] 上述的富媒体处理方法,通过获取富媒体中的音频信息,根据所述音频信息确定

所述富媒体中包含的场景信息；将所述富媒体划分至与所述场景信息相匹配的场景类型中，并展示所述场景类型，可使得在进行对富媒体进行播放之前，即可了解富媒体的声音中的场景类型，然后响应于用户对所述场景类型的选择，播放与所述场景类型相匹配的富媒体，提高了富媒体播放的灵活性。

附图说明

- [0018] 为了更清楚地说明本申请实施例或现有技术中的技术方案，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的前提下，还可以根据这些附图获得其他的附图。
- [0019] 图1为一个实施例中富媒体处理方法的应用环境图；
- [0020] 图2为一个实施例中电子设备的内部结构示意图；
- [0021] 图3为一个实施例中富媒体处理方法的流程图；
- [0022] 图4A为一个实施例中对富媒体进行预览的示意图；
- [0023] 图4B为另一个实施例中富媒体进行预览的示意图；
- [0024] 图4C为又一个实施例中富媒体预览的示意图；
- [0025] 图5为一个实施例中对富媒体的进行播放的流程图；
- [0026] 图6为一个实施例中根据播放指令进入与场景类型对应的场景的播放画面，并进行播放的流程图；
- [0027] 图7为另一个实施例中富媒体处理方法的流程图；
- [0028] 图8为一个实施例中富媒体处理装置的结构框图；
- [0029] 图9为另一个实施例中富媒体处理装置的结构框图；
- [0030] 图10为又一个实施例中富媒体处理装置的结构框图；
- [0031] 图11为再一个实施例中富媒体处理装置的结构框图
- [0032] 图12为一个实施例中与电子设备相关的手机的部分结构的框图。

具体实施方式

[0033] 为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

[0034] 图1为一个实施例中富媒体处理方法的应用环境示意图。如图1所示，该应用环境包括电子设备110和服务器120。电子设备110与服务器120通过网络连接，电子设备110包括但不限于手机、掌上游戏机、平板电脑、个人数字助理或穿戴设备等任意一种终端，电子设备可还可为服务器。服务器120可为独立的服务器，还可为多个服务器组成的服务器集群，或者为服务器集群中的某一个或多个子服务器。电子设备110可从该服务器120上获取富媒体，还可获取本机中存储的富媒体，针对该富媒体进行独立的处理，或者与该服务器进行交互，实现对富媒体的处理。

[0035] 在一个实施例中，如图2所示，提供了一种电子设备的内部结构示意图。该电子设备包括通过系统总线连接的处理器、存储器和显示屏。其中，该处理器用于提供计算和控制

能力,支撑整个电子设备的运行。存储器用于存储数据、程序、和/或指令代码等,存储器上存储至少一个计算机程序,该计算机程序可被处理器执行,以实现本申请实施例中提供的适用于电子设备的富媒体处理方法。存储器可包括磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)等非易失性存储介质,或随机存储记忆体(Random-Access-Memory,RAM)等。例如,在一个实施例中,存储器包括非易失性存储介质及内存储器。非易失性存储介质存储有操作系统、数据库和计算机程序。该数据库中存储有用于实现以上各个实施例所提供的一种富媒体处理方法相关的数据,比如可存储有富媒体等。该计算机程序可被处理器所执行,以用于实现以上各个实施例所提供的一种富媒体处理方法。内存储器为非易失性存储介质中的操作系统、数据库和计算机程序提供高速缓存的运行环境。显示屏可以是触摸屏,比如为电容屏或电子屏,用于显示富媒体等可视信息,还可以被用于检测作用于该显示屏的触摸操作,生成相应的指令。

[0036] 本领域技术人员可以理解,图2中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的电子设备的限定,具体的电子设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。如该电子设备还包括通过系统总线连接的网络接口,网络接口可以是以太网卡或无线网卡等,用于与外部的电子设备进行通信,比如可用于同服务器进行通信,以传输的视频等数据。

[0037] 在一个实施例中,如图3所示,提供了一种富媒体处理方法。本实施例主要以该方法应用于如图2所示的电子设备中为例进行说明,该方法包括:

[0038] 步骤302,获取富媒体中的音频信息。

[0039] 富媒体为需要进行分类的富媒体文件。富媒体(Rich Media)是指包含流媒体、声音、Flash、以及Java、Javascript、DHTML等程序设计语言的形式之一或者几种的组合。可选地,本申请中的富媒体是指包含声音信息的富媒体,比如可为视频,或者具有声音信息的gif动画图像等。富媒体可为存储在电子设备自身的存储器中,还可为存储在云端的服务器中。电子设备可从本机的存储器中提取富媒体,还可从云端服务器中获取富媒体。

[0040] 在一个实施例中,电子设备可自动获取富媒体,或者手动获取富媒体。比如,电子设备可接收对富媒体的分类指令,根据该分类指令来获取该富媒体。或者电子设备可对本机中的富媒体进行分析,将符合预设分类条件的视频作为富媒体。比如,可将通过即时通信应用等社交软件接收到的视频作为富媒体,还可将视频的播放时长和/或以及大小在预设范围内的视频,作为富媒体。比如可将播放时长在10分钟之内的视频作为富媒体,或者可将大小在100Mb之内的视频作为富媒体,以减少对富媒体进行处理的工作量。

[0041] 音频信息表示富媒体中所包含的音频信息,即在播放该富媒体时可播放出声音的信息。电子设备可通过预设的音频提取工具对该富媒体进行音频信息提取,或者可调用预设的录音软件,录取该富媒体中的声音,将录取到的声音作为该音频信息。比如,可调用预设的音频提取工具,将富媒体作为该软件的输入,并运行该音频提取工具,以从该富媒体中提取出对应的音频信息。其中,提取的音频信息可为富媒体中的完整的音频信息,或者可为富媒体中的部分音频信息。可选地,可按照该富媒体的视频时长来确定是否对富媒体中的音频信息进行完整地提取。比如当该视频时长超过预设时长时,可提取中部分的音频信息,当小于预设时长时,则提取完整的音频信息。

[0042] 步骤304,根据音频信息确定富媒体中包含的场景信息。

[0043] 场景信息表示用于体现出待处理富媒体中的声音内容、强弱和/或声音主题等信息。声音内容表示播放该待处理富媒体时,听到的声音的具体内容,比如可为鸟叫声、风声、笑声等。声音强弱表示听到的声音的强弱,比如在某个时间段的声音非常大或嘈杂,而在另一个时间段的声音非常的安静。声音主题可根据声音强弱和/或声音内容而划分的主题。比如按照该声音强弱,场景信息可分成嘈杂型或者安静型等场景信息,按照声音内容归属的主题,场景信息可分成音乐型、人物型或自然型等声音主题的场景信息。

[0044] 电子设备预先设置了多种场景信息,并设置了不同的声音的内容、强弱和/或声音类型所属的场景信息,根据对该音频信息中的声音的内容、强弱和/或声音类型的识别,以判定出对应富媒体中包含的场景信息。比如,可按照声音信号的强弱进行分析,富媒体中的场景信息是否是安静型、一般型或嘈杂型等场景信息;或者按照声音的具体内容进行分析,可识别出富媒体中包含的声音所属的主题。比如当音频信息中包含音乐时,判定该富媒体中包含的声音主题属于音乐的场景信息;包含人声时,判定对应富媒体中包含的声音主题属于人物的场景信息;当包含风声、大海声等自然界声音时,判定富媒体的声音主题属于自然型的场景信息等等。

[0045] 步骤306,将富媒体划分至与场景信息相匹配的场景类型中,并展示场景类型。

[0046] 电子设备还设置与不同的场景信息匹配的场景类型。比如,针对安静型、一般型或嘈杂型的场景信息,对应匹配的场景类型即为安静型、一般型或嘈杂型。其中,安静型表示富媒体中的声音信号比较弱或不存在,因此在播放的过程中出现的声音较小,或者不存在;嘈杂型表示富媒体中存在非常激烈的声音,如在安静的环境中播放,则较为容易对他人造成影响;一般型表示介于上述的安静型和嘈杂型之间。

[0047] 比如,当电子设备解析出该富媒体中的音频信息主要为较舒缓的声音时,可判定该富媒体所属的富媒体类型为一般型;当解析出音频信息主要为较为激烈的声音,则判定该富媒体所属的富媒体类型为激烈型;当解析出音频信息中的声音信号均低于一定的预设信号阈值时,判定该富媒体所属的富媒体类型为安静型。可以理解地,富媒体类型的划分方式还可为多种,并不限于上述的几种划分方式。

[0048] 针对将富媒体所划分的场景类型,电子设备可将该场景类型进行展示。比如可展示在对应的富媒体的预览画面上,使得在播放该富媒体之前,即可获知对应富媒体的场景类型。

[0049] 在一个实施例中,电子设备针对不同的场景类型设置了对应的处理方式,并按照对应的处理方式对该富媒体进行处理。比如可对该富媒体设置对应的提示信息并展示,该提示信息用于提示用户该富媒体的场景类型以使得用户可根据具体的现场环境来确定所使用的播放声音,以降低对他人的打扰。

[0050] 步骤308,响应于用户对场景类型的选择,播放与场景类型相匹配的富媒体。

[0051] 用户可对所展示的场景类型进行选择,电子设备对接收到的用户的选择操作进行响应,触发对与用户选择的场景类型的播放指令,并播放与场景类型相匹配的富媒体。比如,选择的场景类型为安静型时,可按照安静型的富媒体进行播放。

[0052] 举例来说,如图4A~4C所示,提供了几种对富媒体进行场景类型进行展示的方式。如图4A所示,其中的“安静的富媒体”,即表示富媒体1~6中的场景为安静型的场景类型,该“安静的富媒体”可为一种相册的形式,将具有相同的场景类型的富媒体,集合至同一个相

册中,比如集合至该“安静型的富媒体”的相册中。如图4B所示,为一个实施例中电子设备中的所有的富媒体,可在每个富媒体的缩略图上显示对应富媒体的场景类型,比如在其中的富媒体1~6的缩略图上标注对应富媒体的场景类型为安静型;在富媒体7~8的缩略图上标注对应富媒体的场景类型为安静型;在富媒体9~11的缩略图上标注对应富媒体的场景类型为安静型。或者针对单个富媒体的缩略图展示界面,可进一步在其缩略图上标注体现出声音主题的场景类型,如图4C所示,当该富媒体400为所属的场景类型包含鸟叫和笑声时,可将与该鸟叫和笑声的场景类型对应的“鸟叫”和“笑声”的类型标记设置在对应的缩略图中,如图中的鸟叫标记402和鸟叫标记404。其中,该富媒体400可为一个视频,并可为如图4B中的富媒体7或富媒体8等安静型的富媒体。

[0053] 上述的富媒体处理方法,通过获取富媒体中的音频信息,根据音频信息确定富媒体中包含的场景信息;将富媒体划分至与场景信息相匹配的场景类型中,并展示场景类型,可使得在进行对富媒体进行播放之前,即可了解富媒体的声音中的场景类型。然后响应于用户对场景类型的选择,播放与场景类型相匹配的富媒体,提高了对富媒体播放的灵活性。

[0054] 在一个实施例中,步骤304包括:对音频信息进行音频内容识别;根据识别出的音频内容判定富媒体中的声音信号的强弱;和/或根据识别出的音频内容判定富媒体中的声音所属的主题;步骤306包括:将富媒体划分至与判定结果相匹配的场景类型中。

[0055] 本实施例中,音频内容表示音频中的声音被播放时,所听见的具体内容以及声音信号的强弱。比如,若该音频信息中的声音内容为海浪的声音,则该音频内容为海声;若该音频信息中的声音内容为枪击的声音,则该音频内容为枪声;若该音频信息中的声音内容为笑声,该音频内容则为笑声等。可选地,声音信号的强弱可安静型、一般型以及嘈杂型的声音。电子针对不同的声音信号强度,和/或不同的声音所属的主题,设置了对应相匹配的场景类型。根据识别出的声音强弱或主题,可按照对应关系确定相匹配的场景类型,并将该富媒体划分至对应匹配的场景类型中。

[0056] 在一个实施例中,将富媒体划分至与判定结果相匹配的场景类型中,包括:将富媒体划分至与音频内容中信号最强的声音对应的场景类型中;和/或将富媒体划分至与主题相匹配的场景类型中。

[0057] 电子设备可按照该声音内容中,最强声音信号的强弱来判定该声音信号具体是属于哪一种强度型的场景类型。可选地,可根据声音强弱从小到大设置第一强度、第二强度和第三强度,当待处理富媒体中的最强声音信号超过第三强度时,可将该富媒体划分至对应匹配的嘈杂型的场景类型中;当最强声音信号处于第二强度和第三强度之间时,划分至一般型的场景类型中;当最强声音信号小于第一强度时,划分至安静型的场景类型中。

[0058] 进一步地,电子设备也可对该音频信息进行音频内容识别,以检测该音频信息是否属于预定的几种主题之一,并将富媒体划分至与所属的主题相匹配的场景类型中。可选地,可检测该音频信息中的某一段或几段音频的音频特征是否与预设的几种声音主题对应的音频特征相匹配,若匹配,则判定该音频信息属于对应的主题。比如,该音频信息中,存在时间段处于2分钟至3分钟20秒段的音频的音频特征与某一音乐主题的音频特征相匹配,则判定该段音频包含属于音乐主题,并将该富媒体划分至对应匹配的音乐型的场景类型中。

[0059] 上述方法中,通过音频内容来确定富媒体所属的场景类型,可提高对场景类型确定的准确性。

[0060] 在一个实施例中,在步骤306之后,还包括:从音频信息中提取与场景内容相匹配的音频片段;根据音频片段形成音频文件;播放与场景类型相匹配的富媒体,包括:播放音频文件。

[0061] 在识别出音频信息中包含的音频内容之后,可对与该音频内容对应的音频片段进行提取,将所提取出的音频片段转换成预设格式的音频文件,使得可采用相关的音频播放软件对该音频文件进行播放。

[0062] 可选地,可对属于预定的音频内容的音频片段进行提取,该预定的音频内容可为用户自定义设置的音频内容,使得所形成的音频文件为用户感兴趣的音频文件。

[0063] 在一个实施例中,电子设备可接收对音频内容的提取指令,该提取指令中可包含所选取的音频内容。根据该选取的音频内容,提取与该音频内容相匹配的音频片段,并根据该音频片段形成音频文件。可选地,该提取指令中还可包含从该音频片段的起始时间和截止之间。电子设备可从该音频信息中提取处于该起始时间和截止时间之间的音频片段,根据该音频片段形成音频文件。

[0064] 电子设备响应于用户对分离出的音频文件的选择操作,并根据该选择操作对音频文件进行播放。

[0065] 举例来说,提取指令中可包含的音频内容为音频内容A时,可获取该音频内容A对应的音频片段在音频信息中的起始时间和截止时间,并从该音频信息中提取处于该起始时间和截止时间之间的音频片段,根据该音频片段形成音频文件。当检测到用于对音频文件的点击操作时,对该音频文件进行播放。

[0066] 在一个实施例中,在步骤306之后,还包括:对富媒体进行视频分离;根据分离出的视频信息形成视频文件;播放与场景类型相匹配的富媒体,包括:播放视频文件。

[0067] 可选地,电子设备还可对富媒体进行音频信息和视频信息的分离处理,以分离出其中的视频信息,并根据该分离出的视频信息独立地形成视频文件,使得可在需要完全静音的环境中也可以对该视频文件进行查看。

[0068] 在一个实施例中,电子设备可接收对视频信息的提取指令,该提取指令中可包含所选取的视频内容。根据该选取的视频内容,提取与该视频内容相匹配的视频片段,并根据该视频片段形成视频文件。可选地,该提取指令中还可包含从该视频片段的起始时间和截止时间。电子设备可从该视频信息中提取处于该起始时间和截止时间之间的视频片段,根据该视频片段形成视频文件。当检测到用于对视频文件的点击操作时,对该视频文件进行播放。

[0069] 电子设备响应于用户对分离出的视频文件的选择操作,并根据该选择操作对视频文件进行播放。

[0070] 举例来说,提取指令中可包含的视频内容为视频内容A时,可获取该视频内容A对应的视频片段在视频信息中的起始时间和截止时间,并从该视频信息中提取处于该起始时间和截止时间之间的视频片段,根据该视频片段形成视频文件,使得所形成的视频文件为用户感兴趣的视频文件。

[0071] 在一个实施例中,在步骤308之前,还包括:对富媒体设置用于标记场景类型的类型标记。

[0072] 类型标记用于标记该场景类型,针对每种场景类型,电子设备设置了与该场景类

型对应的类型标记,通过该类型标记来标记对应视频所属的场景类型。比如该类型标记可分为“鸟叫”、“枪声”、“笑声”等。电子设备可将该类型标记设置在富媒体在被播放前或播放过程中的预设展示位置处,使得用户在查看到相应视频标记时,即可获知该富媒体所属的场景类型。

[0073] 可选地,电子设备可在该富媒体的缩略图上的任意位置处加载所设置的类型标记,使得在播放该富媒体之前,即可通过该缩略图上的类型标记获知对应视频的场景类型。或者还可在视频播放的画面中的任意位置加载所设置的类型标记,使得在播放的过程中也可获知对应视频的场景类型。

[0074] 以富媒体为视频进行举例说明,如图4所示,为一个实施例中视频预览的示意图。其中,该缩略图400为某一视频的缩略图,当确定该视频所属的场景类型包含鸟叫和笑声时,可将与该鸟叫和笑声的场景类型对应的“鸟叫”和“笑声”的类型标记设置在对应的缩略图中,如图中的鸟叫标记402和鸟叫标记404。

[0075] 在一个实施例中,如图5所示,步骤308包括:

[0076] 步骤502,接收作用于类型标记而触发的播放指令。

[0077] 可选地,电子设备在设置该类型标记时,还进一步设置了与该类型标记对应的场景的播放指令。该播放指令表示对该类型标记对应的场景的播放指令。电子设备可进一步设置对该场景的播放指令的播放按钮,当侦测到对该播放按钮的点击操作时,触发对应的场景的播放指令。

[0078] 在一个实施例中,可直接将展示富媒体的场景类型的类型标记设置为该播放按钮,即电子设备可针对富媒体,增加专门用于进入与所判定出的场景类型的播放按钮,并将该类型标记展示在该播放按钮中。如图4所示,则该笑声标记402和鸟叫标记404还可作为对应笑声场景的播放按钮和鸟叫的播放按钮。

[0079] 通过侦测对该类型标记的点击操作,触发与所点击的类型标记对应的场景的播放指令。其中,该类型标记可展示在富媒体的播放之前或播放过程中,当设置在播放过程中时,可通过对该视频标记的点击,实现快速切换至与该视频标记对应的场景。

[0080] 步骤504,识别富媒体中与类型标记对应的场景。

[0081] 在一个实施例中,当电子设备在识别出每个音频片段对应的音频内容之后,可进一步建立根据该音频内容确定的类型标记与该音频片段之间的对应关系。根据该对应关系查询对应的音频片段,将富媒体中,该音频片段所处的时间段内的视频部分作为该类型标记对应的场景。

[0082] 可选地,电子设备也可在识别出每个音频片段对应的音频内容之后,记录该音频片段的起始时间和截止时间。电子设备可根据该播放指令查询对应的音频片段的起始时间和截止时间,将处于该起始时间和截止时间的时间段内的视频部分作为该类型标记对应的场景。

[0083] 在一个实施例中,步骤502和步骤504之间的执行顺序可不做限定,比如还可在步骤502之前即可先执行步骤504,即在播放富媒体之前,即可预先识别出每个类型标记对应的场景,使得根据播放指令,可快速进行对应场景的播放。

[0084] 步骤506,根据播放指令进入与场景类型对应的场景的播放画面,并进行播放。

[0085] 电子设备在确定待播放的场景类型之后,可进入该场景类型对应的播放画面并播

放,以提高播放的灵活性。可选地,可直接进入该场景的起始时间对应的画面,并进行播放。或者比该起始时间早预设时长的时间对应的画面并播放。其中,预设时长可为任意合适的时长,比如为5秒,即根据该播放指令切换至早于对应场景的起始时间的前5秒的画面并播放。

[0086] 同样参考如图4所示,可直接将其上的笑声标记402和鸟叫标记404分别作为播放按钮。当侦测到作用于该鸟叫标记402的点击操作时,可触发对笑声对应的场景的播放指令,并根据该播放指令对对应的笑声的场景进行播放。比如当该笑声的场景对应的时间段为3分0秒至3分8秒时,可直接进入3分0秒的画面并继续播放,或者可进入2分55秒的画面进行播放。

[0087] 上述实施例中,通过接收对类型标记对应的场景的播放指令,进而播放指令进入对场景类型对应的场景的播放画面,从而可快速准确地对视频中的类型标记的场景进行播放,进一步提高了视频播放的灵活性。

[0088] 在一个实施例中,如图6所示,步骤506包括:

[0089] 步骤602,根据播放指令获取与场景类型对应音频内容在音频信息中的起始位置。

[0090] 步骤604,根据起始位置确定进入的播放画面。

[0091] 进入的播放画面可为该起始位置所处的播放画面,还可为早于该起始位置预设时长的播放画面。该预设时长可为任意合适的时长,比如为5秒。举例来说,当起始位置为2分5秒时,可根据该起始位置进入同样为2分5秒时的播放画面,或者进入比该2分5秒早5秒的播放画面,即进入2分0秒的播放画面。

[0092] 步骤606,获取本机所处环境的环境音量;根据环境音量和场景类型确定富媒体的播放音量。

[0093] 步骤608,根据播放音量对进入的播放画面进行播放。

[0094] 环境音量表示电子设备所处环境中的实时的声音的大小。当接收到用户对该富媒体的播放指令时,可调用内置的语音采集装置对环境音量进行检测,提取电子设备所处环境的环境音量。电子设备进一步预设了环境音量、场景类型与播放音量之间的对应关系,该播放音量表示属于该场景类型的富媒体,在该环境音量下的适宜的播放音量。根据该对应关系,查询出与该场景类型和环境音量对应的播放音量,按照该播放音量对富媒体进行播放。或者可将该播放音量提供给用户进行选择,使用户可选择使用所确定的播放音量对富媒体进行播放。当接收到对该播放音量的选择时,再该播放音量对该富媒体进行播放,以进一步提高对富媒体的播放的灵活性。

[0095] 在一个实施例中,该对应关系可通过播放音量对照表来体现,即电子设备中预设了相应的播放音量对照表,该对照表中记录了不同的场景类型在不同环境音量下,对应的播放音量。电子设备可直接从该对照表中查询与该场景类型和环境音量对应的播放音量,可提高对播放音量的确定的速度。

[0096] 在一个实施例中,电子设备可预设播放音量的音量计算模型,并设置不同的场景类型对应的量化数值,将该量化数值与环境音量作为音量计算模型的输入,并运行该音量计算模型,从而输出计算出的播放音量。

[0097] 同样参考如图4所示,当确定视频的类型包括场景类型包含鸟叫和笑声时,当检测到对该播放按钮406的点击操作而触发的播放指令后,可获取环境音量,并获取与该环境音

量与鸟叫和笑声类型对应的播放音量,以提示用户是否使用该播放音量进行播放,当接收到选择该播放音量进行播放时,则按照该播放音量对富媒体进行播放。

[0098] 在一个实施例中,如图7所示,提供了另一种富媒体处理方法,该方法包括:

[0099] 步骤702,获取富媒体中的音频信息;对音频信息进行音频内容识别。

[0100] 可选地,该富媒体可为从服务器上接收到的视频,比如为服务器转发的通过聊天应用等接收到的富媒体。还可为电子设备预先存储的富媒体。电子设备可自动发起对获取的富媒体进行下述的处理,还可根据接收到的对该富媒体的处理指令时,触发对富媒体进行下述的处理。

[0101] 举例来说,电子设备可通过某一聊天应用接收到好友发送过来的视频并下载完毕之后,可将该视频作为富媒体,并自动触发下述的对该视频的处理的过程。

[0102] 电子设备可对该富媒体进行音频提取,以提取出该富媒体中的音频信息,针对提取出的音频信息进行分析。其中,可调用预设的音频提取工具进行音频信息的提取。

[0103] 针对该音频信息,电子设备可按照预设的音频内容识别模型进行识别,将该音频信息作为该内容识别模型的输入,并运行该内容识别算法,以得出该音频信息中所包含的声音内容,以及该声音内容在音频信息中的位置与时长。

[0104] 可选地,一份音频信息包含的声音内容可包括多个,比如可同时包含音乐、鸟叫或枪声等。电子设备可预先对不同的声音内容进行分类,根据该分类形成场景类型。比如,可将声音内容为风声、大海声等自然界声音时,划分为自然型的场景类型;将声音内容为狗叫、猫叫等动物叫声划分为动物声的场景类型。电子设备可针对识别出的声音内容、以及声音内容的数量、每个声音内容在音频信息中占据的比例等来确定对应富媒体所属的场景类型。比如,当声音内容仅为一种时,可直接将该声音内容所属的类型作为该视频的场景类型;当该声音内容包含多种时,可进一步检测每种声音内容在整个音频信息中占据的比例,将超过预设比例的声音内容所属的类型作为该视频的场景类型。其中,该预设比例可为任意设置的合适比例,比如可为10%。

[0105] 步骤704,获取富媒体中的视频帧。

[0106] 视频帧表示组成视频中的播放画面的静态画面。电子设备可进一步对该富媒体进行解析,以获取构成该播放画面的视频帧。其中,可获取该富媒体中的所有的视频帧,还可获取其中的部分视频帧。比如,可按照预设的采样率,以每间隔预设数量进行提取一帧视频帧。预设数量可为固定设置的任意合适的数量,或者可根据视频的播放时长和视频帧的帧数来确定。相同时长的视频,其帧数越多,则间隔数量可越大。比如,该视频播放时长为10分钟,视频帧数为6000,则可每间隔5帧或8帧等来提取一帧视频帧。

[0107] 在一个实施例中,上述步骤702和步骤704之间的执行顺序可不做限定,比如还可同时执行步骤702和步骤704,或者还可先执行步骤704,再执行步骤702。

[0108] 步骤706,根据识别出的音频内容判定富媒体中的声音信号的强弱;和/或根据识别出的音频内容判定富媒体中的声音所属的主题。

[0109] 步骤708,将富媒体划分至与判定结果相匹配的场景类型中,并展示场景类型。

[0110] 可选地,将富媒体划分至与音频内容中信号最强的声音对应的场景类型中;和/或将富媒体划分至与主题相匹配的场景类型中。

[0111] 在一个实施例中,可按照该视频帧和音频内容来共同确定富媒体的场景类型。电

子设备可对于连续提取的每帧视频帧,以及与该视频帧对应的临近的音频片段进行结合分析,确定该视频的场景类型,将该富媒体划分至对应的场景类型中。电子设备可对每帧视频帧进行画面分析,识别出该富媒体中的不同时刻下的画面信息,再结合每个时刻下的音频,确定该视频的场景类型。通过由视频帧和音频信息来共同确认场景类型,可进一步提高对场景类型确定的准确性。

[0112] 比如,当画面中存在大海,且该画面对应的时刻下的音频也属于大海的声音,则可判定该视频的场景类型包含该大海与海声对应的场景类型。

[0113] 步骤710,对富媒体设置用于标记场景类型的类型标记。

[0114] 针对判定出的富媒体的场景类型,可进一步对该视频设置类型标记。比如可在记录该富媒体的类型标记为“海声”,并将该类型标记设置在该富媒体的预览画面上,或者在预览该富媒体中的界面上的任意位置。使得用户通过该类型标记,在不播放该视频的情况下,即可获知该视频的场景类型,初步判定出视频中所包含的声音或画面的信息。比如,可将该类型标记设置在富媒体的预览图或播放过程中的右上角或者左下角等位置,使得可减少对预览画面或播放画面的遮挡。

[0115] 步骤712,接收作用于类型标记而触发的播放指令。根据播放指令进入与场景类型对应的场景的播放画面,并进行播放。

[0116] 可选地,电子设备还可将该类型标记设置成可被触发播放指令的播放按钮。该播放按钮可展示在富媒体的播放之前或播放过程中。当接收到该按钮的点击操作时,触发对该标记对应的场景的播放指令。

[0117] 可选地,该场景可进一步根据上述的视频帧与声音来确定。电子设备在检测出每个类型标记对应的音频信息的音频片段时,可根据该音频片段在整个音频信息中的时间,获取与该时间临近的视频帧,检测该视频帧中的内容是否与该类型标记相匹配,根据检测结果确定与该类型标记匹配的起始的视频帧,将该起始的视频帧作为与该类型标记对应的场景的起始画面,并进入该起始画面进行播放,或者进入该起始画面之前的预设数量的画面进行播放。

[0118] 举例来说,当该类型标记为“海声”时,该海声对应的音频片段在音频信息中处于2分3秒至3分之内,可检测位于2分3秒附近的视频帧是否与海声相匹配,即检测该视频帧内是否存在大海的画面,并进一步检测与该临近的多帧视频帧中,开始出现大海的画面的视频帧,将该视频帧作为对应场景的起始帧画面,比如,该起始帧画面处于1份58秒的位置,则可进入1分58秒并播放,或者早于该1分58秒的位置,如1分55秒,并从1分55秒的画面进行播放。

[0119] 通过结合音频信息与视频帧来确定类型标记对应的场景,使得与标记类型对应的场景更加准确,从而可提高进入该场景的准确性。

[0120] 在一个实施例中,如图8所示,提供了一种富媒体处理装置。该装置包括:

[0121] 音频信息获取模块802,用于获取富媒体中的音频信息。

[0122] 场景信息识别模块804,用于根据音频信息确定富媒体中包含的场景信息;

[0123] 分类模块806,用于将富媒体划分至与场景信息相匹配的场景类型中,并展示场景类型。

[0124] 播放模块808,用于响应于用户对场景类型的选择,播放与场景类型相匹配的富媒

体。

[0125] 在一个实施例中,场景信息包括声音的强弱信息和/或声音所属主题的信息。

[0126] 场景信息识别模块804还用于对音频信息进行音频内容识别;根据识别出的音频内容判定富媒体中的声音信号的强弱;和/或根据识别出的音频内容判定富媒体中的声音所属的主题。

[0127] 分类模块806还用于将富媒体划分至与判定结果相匹配的场景类型中。

[0128] 在一个实施例中,分类模块806还用于将富媒体划分至与音频内容中信号最强的声音对应的场景类型中;和/或将富媒体划分至与主题相匹配的场景类型中。

[0129] 在一个实施例中,如图9所示,提供了另一种富媒体处理装置,该装置还包括:

[0130] 音频文件生成模块810,用于从音频信息中提取与场景信息相匹配的音频片段;根据音频片段形成音频文件。

[0131] 播放模块808还用于播放音频文件。

[0132] 在一个实施例中,如图10所示,提供了又一种富媒体处理装置,该装置还包括:

[0133] 视频文件生成模块812,用于对富媒体进行视频分离;根据分离出的视频信息形成视频文件。

[0134] 播放模块808还用于播放视频文件。

[0135] 在一个实施例中,如图11所示,提供了再一种富媒体处理装置,该装置还包括:

[0136] 类型标记模块814,用于对富媒体设置用于标记场景类型的类型标记。

[0137] 播放模块808还用于接收作用于类型标记而触发的播放指令;根据播放指令进入与场景类型对应的场景的播放画面,并进行播放。

[0138] 在一个实施例中,播放模块808还用于根据播放指令获取与场景类型对应音频内容在音频信息中的起始位置;根据起始位置确定进入的播放画面;获取本机所处环境的环境音量;根据环境音量和场景类型确定富媒体的播放音量;根据播放音量对进入的播放画面进行播放。

[0139] 上述富媒体处理装置中各个模块的划分仅用于举例说明,在其他实施例中,可将富媒体处理装置按照需要划分为不同的模块,以完成上述富媒体处理装置的全部或部分功能。

[0140] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现上述各实施例所提供的富媒体处理方法的步骤。

[0141] 一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,处理器执行计算机程序时实现上述各实施例所提供的富媒体处理方法的步骤。

[0142] 本申请实施例还提供了一种计算机程序产品。一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述各实施例所提供的富媒体处理方法的步骤。

[0143] 本申请实施例还提供了一种电子设备。如图12所示,为了便于说明,仅示出了与本申请实施例相关的部分,具体技术细节未揭示的,请参照本申请实施例方法部分。该电子设备可以为包括手机、平板电脑、PDA(Personal Digital Assistant,个人数字助理)、POS(Point of Sales,销售终端)、车载电脑、穿戴式设备等任意终端设备,以电子设备为手机为例:

[0144] 图12为与本申请实施例提供的电子设备相关的手机的部分结构的框图。参考图

12,手机包括:射频(Radio Frequency,RF)电路1210、存储器1220、输入单元1230、显示单元1240、传感器1250、音频电路1260、无线保真(wireless fidelity,WiFi)模块1270、处理器1280、以及电源1290等部件。本领域技术人员可以理解,图12所示的手机结构并不构成对手机的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0145] 其中,RF电路1210可用于收发信息或通话过程中,信号的接收和发送,可将基站的下行信息接收后,给处理器1280处理;也可以将上行的数据发送给基站。通常,RF电路包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器(Low Noise Amplifier,LNA)、双工器等。此外,RF电路1210还可以通过无线通信与网络和其他设备通信。上述无线通信可以使用任一通信标准或协议,包括但不限于全球移动通讯系统(Global System of Mobile communication,GSM)、通用分组无线服务(General Packet Radio Service,GPRS)、码分多址(Code Division Multiple Access,CDMA)、宽带码分多址(Wideband Code Division Multiple Access,WCDMA)、长期演进(Long Term Evolution,LTE))、电子邮件、短消息服务(Short Messaging Service,SMS)等。

[0146] 存储器1220可用于存储软件程序以及模块,处理器1280通过运行存储在存储器1220的软件程序以及模块,从而执行手机的各种功能应用以及数据处理。存储器1220可主要包括程序存储区和数据存储区,其中,程序存储区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能的应用程序、图像播放功能的应用程序等)等;数据存储区可存储根据手机的使用所创建的数据(比如音频数据、通讯录等)等。此外,存储器1220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、或其他易失性固态存储器件。

[0147] 输入单元1230可用于接收输入的数字或字符信息,以及产生与手机1200的用户设置以及功能控制有关的键信号输入。具体地,输入单元1230可包括触控面板1231以及其他输入设备1232。触控面板1231,也可称为触摸屏,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触控面板1231上或在触控面板1231附近的操作),并根据预先设定的程式驱动相应的连接装置。在一个实施例中,触控面板1231可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给处理器1280,并能接收处理器1280发来的命令并加以执行。此外,可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触控面板1231。除了触控面板1231,输入单元1230还可以包括其他输入设备1232。具体地,其他输入设备1232可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)等中的一种或多种。

[0148] 显示单元1240可用于显示由用户输入的信息或提供给用户的信息以及手机的各种菜单。显示单元1240可包括显示面板1241。在一个实施例中,可以采用液晶显示器(Liquid Crystal Display,LCD)、有机发光二极管(Organic Light-Emitting Diode,OLED)等形式来配置显示面板1241。在一个实施例中,触控面板1231可覆盖显示面板1241,当触控面板1231检测到在其上或附近的触摸操作后,传送给处理器1280以确定触摸事件的类型,随后处理器1280根据触摸事件的类型在显示面板1241上提供相应的视觉输出。虽然在图12中,触控面板1231与显示面板1241是作为两个独立的部件来实现手机的输入和输入

功能,但是在某些实施例中,可以将触控面板1231与显示面板1241集成而实现手机的输入和输出功能。

[0149] 手机1200还可包括至少一种传感器1250,比如光传感器、运动传感器以及其他传感器。具体地,光传感器可包括环境光传感器及接近传感器,其中,环境光传感器可根据环境光线的明暗来调节显示面板1241的亮度,接近传感器可在手机移动到耳边时,关闭显示面板1241和/或背光。运动传感器可包括加速度传感器,通过加速度传感器可检测各个方向上加速度的大小,静止时可检测出重力的大小及方向,可用于识别手机姿态的应用(比如横竖屏切换)、振动识别相关功能(比如计步器、敲击)等;此外,手机还可配置陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器等。

[0150] 音频电路1260、扬声器1261和传声器1262可提供用户与手机之间的音频接口。音频电路1260可将接收到的音频数据转换后的电信号,传输到扬声器1261,由扬声器1261转换为声音信号输出;另一方面,传声器1262将收集的声音信号转换为电信号,由音频电路1260接收后转换为音频数据,再将音频数据输出处理器1280处理后,经RF电路1210可以发送给另一手机,或者将音频数据输出至存储器1220以便后续处理。

[0151] WiFi属于短距离无线传输技术,手机通过WiFi模块1270可以帮助用户收发电子邮件、浏览网页和访问流式媒体等,它为用户提供了无线的宽带互联网访问。虽然图12示出了WiFi模块1270,但是可以理解的是,其并不属于手机1200的必须构成,可以根据需要而省略。

[0152] 处理器1280是手机的控制中心,利用各种接口和线路连接整个手机的各个部分,通过运行或执行存储在存储器1220内的软件程序和/或模块,以及调用存储在存储器1220内的数据,执行手机的各种功能和处理数据,从而对手机进行整体监控。在一个实施例中,处理器1280可包括一个或多个处理单元。在一个实施例中,处理器1280可集成应用处理器和调制解调处理器,其中,应用处理器主要处理操作系统、用户界面和应用程序等;调制解调处理器主要处理无线通信。可以理解的是,上述调制解调处理器也可以不集成到处理器1280中。

[0153] 手机1200还包括给各个部件供电的电源1290(比如电池),优选的,电源可以通过电源管理系统与处理器1280逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

[0154] 在一个实施例中,手机1200还可以包括摄像头、蓝牙模块等。

[0155] 在本申请实施例中,该移动终端所包括的处理器1280执行存储在存储器上的计算机程序时实现上述所描述的富媒体处理方法的步骤。

[0156] 本申请所使用的对存储器、存储、数据库或其它介质的任何引用可包括非易失性和/或易失性存储器。合适的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),它用作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(DDR SDRAM)、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)。

[0157] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并

不能因此而理解为对本申请专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

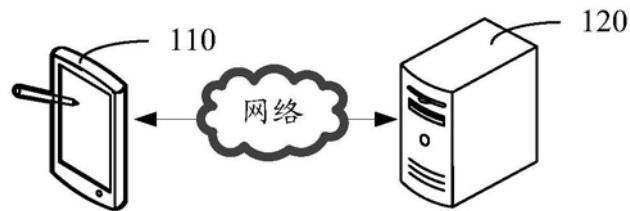


图1

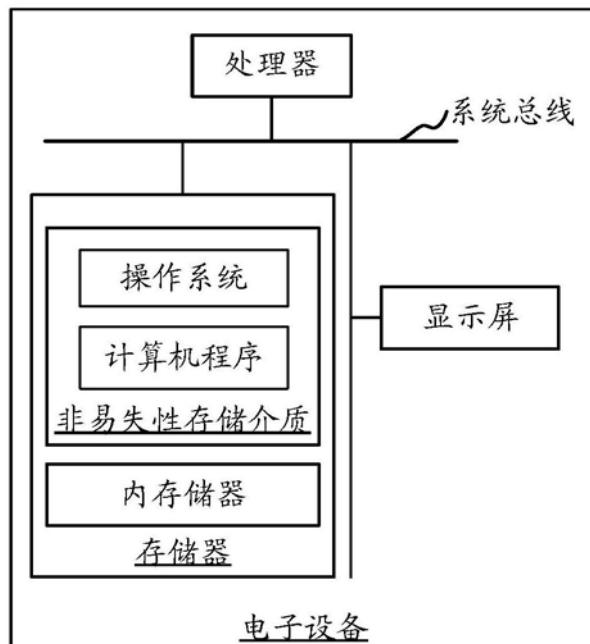


图2

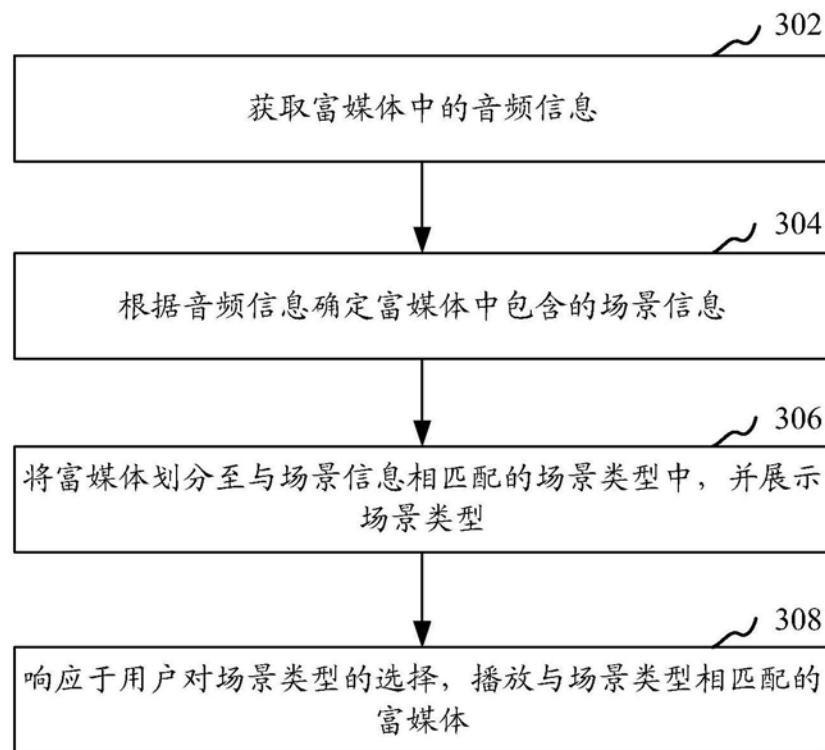


图3

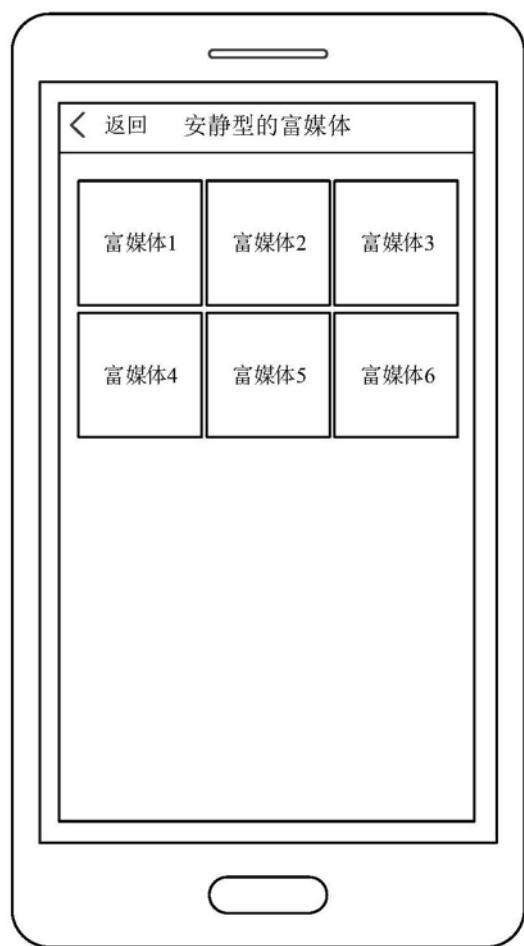


图4A

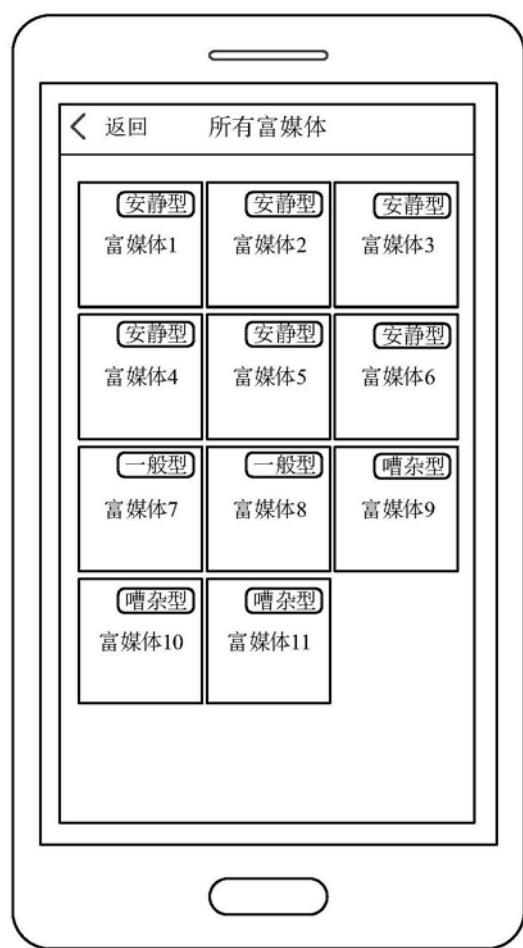


图4B

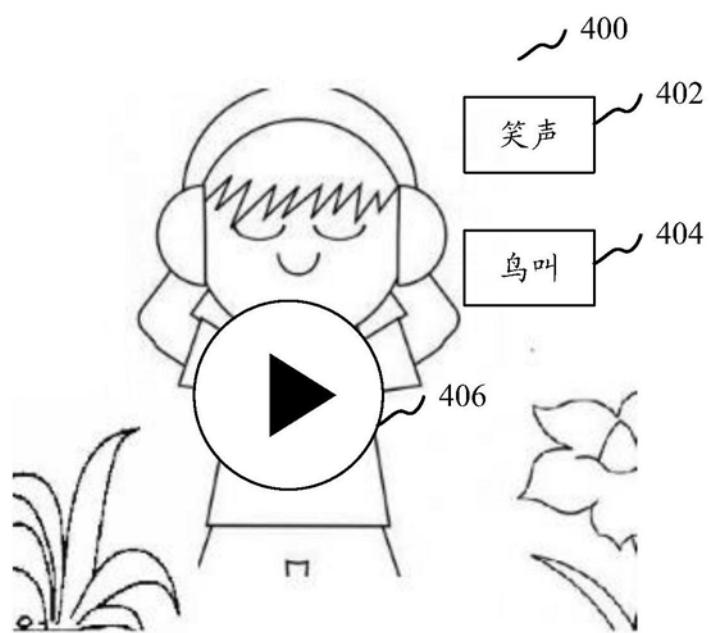


图4C

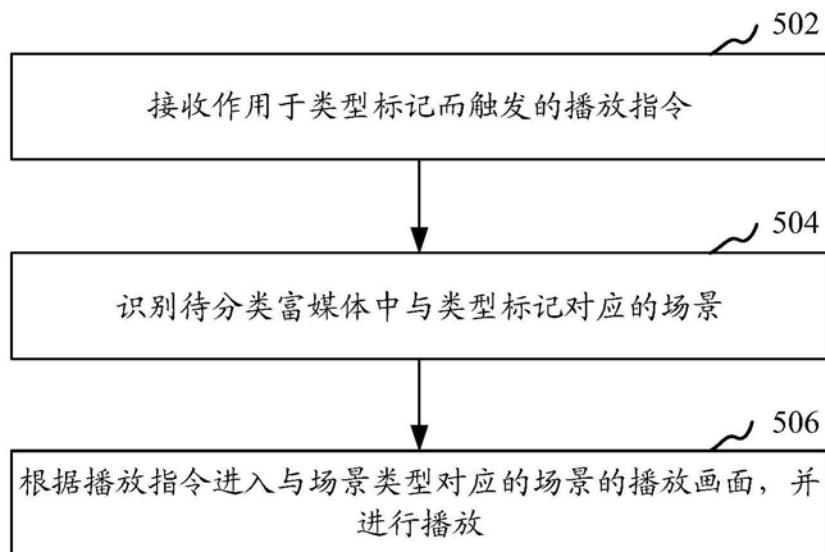


图5

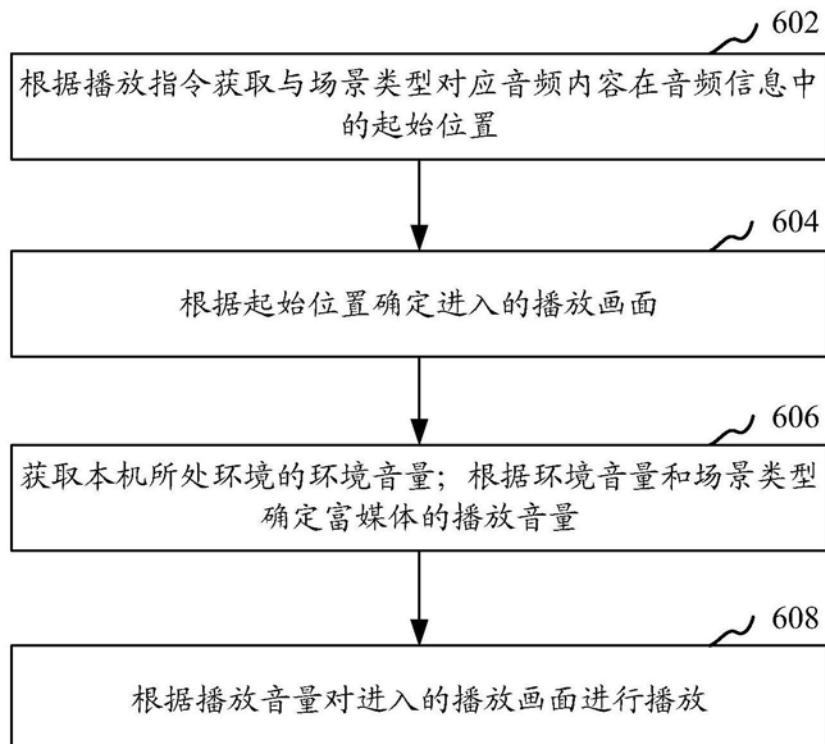


图6

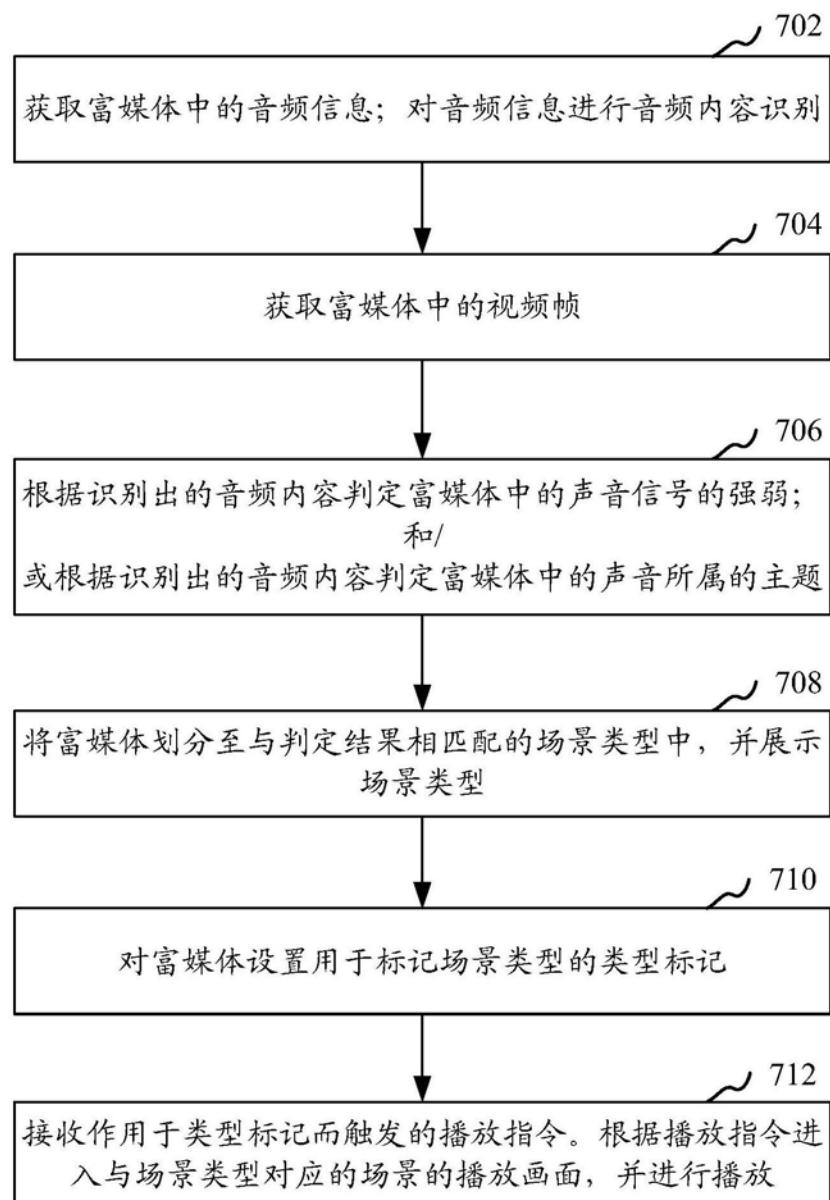


图7

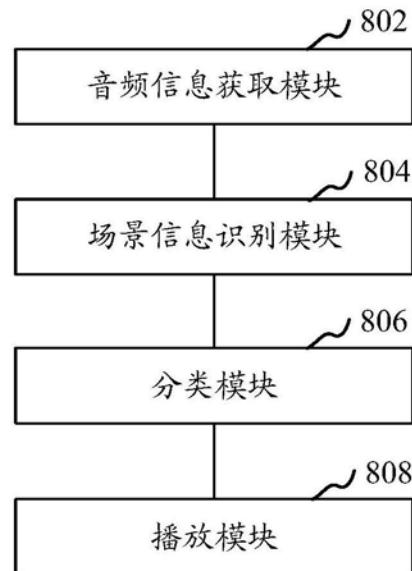


图8

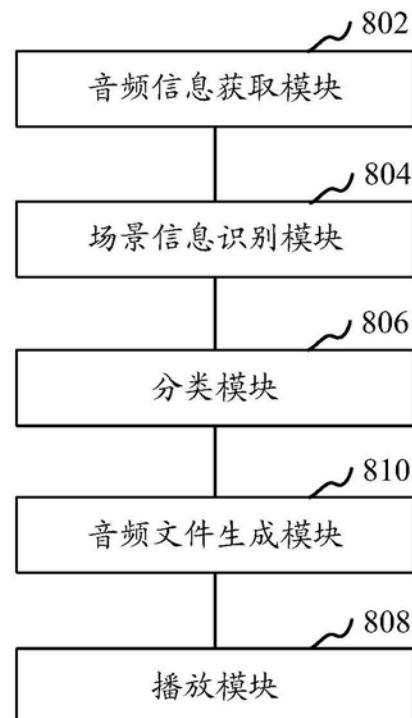


图9

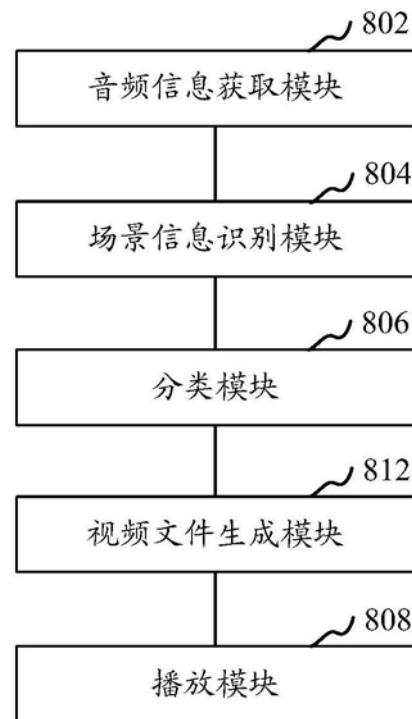


图10

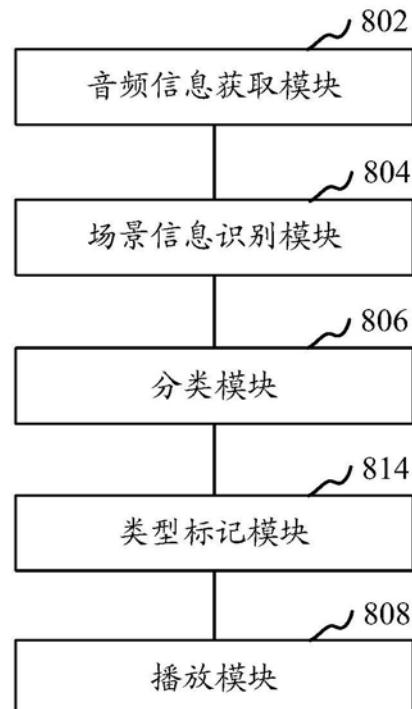


图11

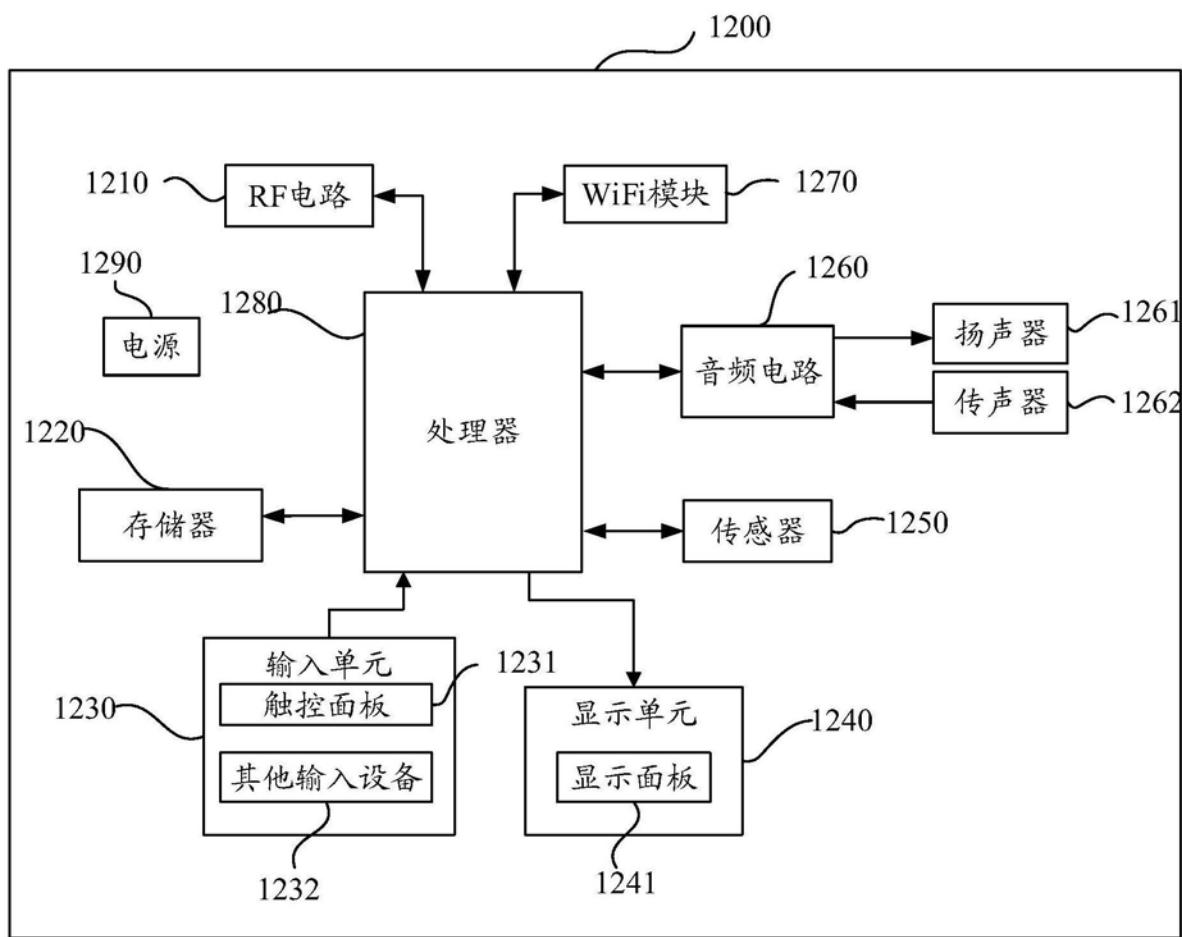


图12