

(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
2019年7月4日 (04.07.2019)



(10) 国际公布号
WO 2019/128629 A1

- (51) 国际专利分类号:
H04S 5/00 (2006.01)
- (21) 国际申请号: PCT/CN2018/118764
- (22) 国际申请日: 2018年11月30日 (30.11.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201711432680.4 2017年12月26日 (26.12.2017) CN
- (71) 申请人: 广州酷狗计算机科技有限公司 (GUANGZHOU KUGOU COMPUTER TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省广州市天河区黄埔大道中315号自编1-17, Guangdong 510660 (CN)。
- (72) 发明人: 刘佳泽 (LIU, Jiaze); 中国广东省广州市天河区黄埔大道中315号自编1-17, Guangdong 510660 (CN)。
- (74) 代理人: 北京三高永信知识产权代理有限责任公司 (BEIJING SAN GAO YONG XIN INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市海淀区学院路蓟门里和景园A座1单元102室, Beijing 100088 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,

(54) Title: AUDIO SIGNAL PROCESSING METHOD AND APPARATUS, TERMINAL AND STORAGE MEDIUM

(54) 发明名称: 音频信号的处理方法、装置、终端及存储介质

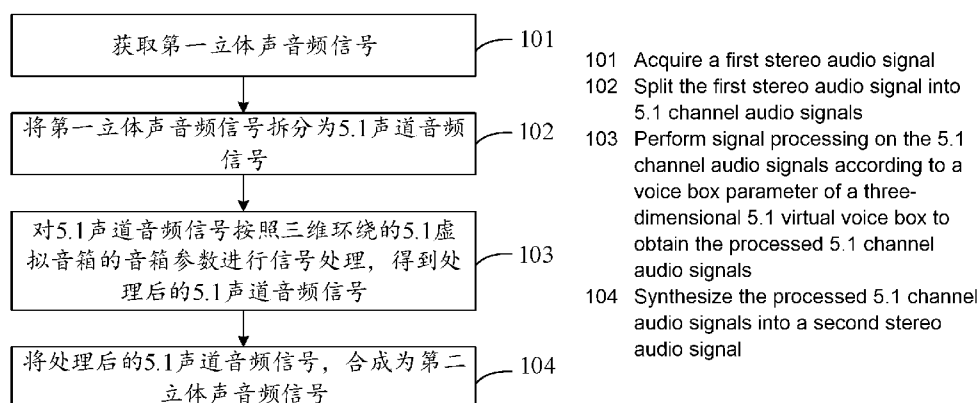


图 1

(57) Abstract: Disclosed in the present application are an audio signal processing method and apparatus, a terminal and a storage medium, belonging to the field of audio processing technologies. The method comprises: acquiring a first stereo audio signal; splitting the first stereo audio signal into 5.1 channel audio signals; performing signal processing on the 5.1 channel audio signals according to a voice box parameter of a three-dimensional 5.1 virtual voice box to obtain the processed 5.1 channel audio signals; and synthesizing the processed 5.1 channel audio signals into a second stereo audio signal. In the present application, a first stereo audio signal is split into 5.1 channel audio signals, and then the 5.1 channel audio signals are processed and synthesized into a second stereo audio signal, and the second stereo audio signal is played back by a binaural audio playback unit, so that the user obtains a stereo effect of a 5.1 channel audio, solving the problem in the related art that playing back binaural audio signals only will result in poor stereo effect, and improving the stereo effect of audio playback.

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：本申请公开了一种音频信号的处理方法、装置、终端及存储介质，属于音频处理技术领域。所述方法包括：获取第一立体声音频信号；将第一立体声音频信号拆分为5.1声道音频信号；对5.1声道音频信号按照三维环绕的5.1虚拟音箱的音箱参数进行信号处理，得到处理后的5.1声道音频信号；将处理后的5.1声道音频信号，合成为第二立体声音频信号。本申请通过将第一立体声音频信号拆分为5.1声道音频信号，再将5.1声道音频信号处理并合成为第二立体声音频信号，通过双声道的音频播放单元播放该第二立体声音频信号使得用户获得5.1声道音频的立体声效果，解决了相关技术中仅播放双声道音频信号所带来的立体效果较差的问题，提高了音频播放的立体效果。

音频信号的处理方法、装置、终端及存储介质

本申请要求于 2017 年 12 月 26 日提交的申请号为 201711432680.4、发明名称为“音频信号的处理方法、装置及终端”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及音频处理技术领域，特别涉及一种音频信号的处理方法、装置、终端及存储介质。

背景技术

相关技术中，音频播放设备通过双声道的耳机或扬声器等音频播放单元播放双声道音频信号使用户获得立体声效果。其中，双声道音频信号为左声道音频信号和右声道音频信号叠加的音频信号，音频播放单元的左声道部分播放左声道音频信号，音频播放单元的右声道部分播放右声道音频信号，用户通过左声道部分播放的左声道音频信号和右声道部分播放的右声道音频信号的相位差获得立体的听感。

相关技术中音频播放单元是通过播放左声道音频信号和右声道音频信号使用户获得立体的听感，由于声音是通过多个方向传播的，仅通过播放两个声道的音频信号的立体效果较差。

发明内容

本申请实施例提供了一种音频信号的处理方法、装置、终端及存储介质，可以解决通过音频播放单元播放左声道音频信号和右声道音频信号时立体效果较差的问题。所述技术方案如下：

一方面，本申请实施例提供了一种音频信号的处理方法，所述方法由终端执行，所述方法包括：

获取第一立体声音频信号；

将所述第一立体声音频信号拆分为 5.1 声道音频信号；

对所述 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信

号处理，得到处理后的 5.1 声道音频信号；

将所述处理后的 5.1 声道音频信号，合成为第二立体声音频信号。

一方面，本申请实施例提供了一种音频信号的处理装置，所述装置应用于终端中，所述装置包括：

获取模块，用于获取第一立体声音频信号；

处理模块，用于将所述第一立体声音频信号拆分为 5.1 声道音频信号；对所述 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信号处理，得到处理后的 5.1 声道音频信号；

合成模块，用于将所述处理后的 5.1 声道音频信号，合成为第二立体声音频信号。

一方面，本申请实施例提供了一种终端，所述终端包括处理器和存储器，所述存储器中存储有至少一条指令，所述指令由所述处理器加载并执行以实现如上所述的音频信号处理方法。

一方面，本申请实施例提供了一种计算机可读存储介质，所述存储介质中存储有至少一条指令，所述指令由处理器加载并执行以实现如上所述的音频信号处理方法。

本申请实施例提供的技术方案带来的有益效果至少包括：

通过将第一立体声音频信号拆分为 5.1 声道音频信号，再将 5.1 声道音频信号处理并合成为第二立体声音频信号，通过双声道的音频播放单元播放该第二立体声音频信号使得用户获得 5.1 声道音频的立体声效果，解决了相关技术中仅播放双声道音频信号所带来的立体效果较差的问题，提高了音频播放的立体效果。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例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。

图 1 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图 2 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图

图；

图 3 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图；

图 4 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图；

图 5 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图；

图 6 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图；

图 7 示出了本申请一个示例性实施例提供的 5.1 声道虚拟音箱的摆放示意图；

图；

图 8 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图；

图；

图 9 示出了本申请一个示例性实施例提供的 HRTF 数据的采集原理图；

图 10 示出了本申请一个示例性实施例提供的音频信号的处理装置的框图；

图 11 示出了本申请一个示例性实施例提供的音频信号的处理装置的框图；

图 12 示出了本申请一个示例性实施例提供的音频信号的处理装置的框图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

图 1，示出了本申请一个示例性的实施例提供的音频信号的处理方法的方法流程图，该方法可具有音频信号处理功能的终端执行，该方法包括：

步骤 101，获取第一立体声音频信号。

终端读取本地存储的第一立体声音频信号，或通过有线或无线网络获取得到服务器上的第一立体声音频信号。

第一立体声音频信号是通过立体声录音设备录制声音得到的，立体声录音设备通常包括位于左侧的第一麦克风和位于右侧的第二麦克风，立体声录音设备通过第一麦克风和第二麦克风分别录制左侧的声音和右侧的声音获得左声道音频信号和右声道音频信号，立体声录音设备将左声道音频信号和右声道音

频信号叠加后获得第一立体声信号。

可选的，终端将接收到的第一立体声音频信号存储在终端的缓存中，第一立体声音频信号记为 X_PCM。

终端将接收到的第一立体声音频信号以左声道音频信号和对应的右声道音频信号的采样对形式存储在一个内置的缓存区域，使用时从该缓存区域获取第一立体声音频信号。

步骤 102，将第一立体声音频信号拆分为 5.1 声道音频信号。

终端通过预置的算法将第一立体声音频信号拆分为 5.1 声道音频信号。其中，5.1 声道音频信号包括前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号。

步骤 103，对 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信号处理，得到处理后的 5.1 声道音频信号。

终端对 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信号处理，得到处理后的 5.1 声道音频信号。

其中，处理后的 5.1 声道音频信号包括处理后的前置左声道信号、处理后的前置右声道信号、处理后的前置中央声道信号、处理后的后置左声道信号和处理后的后置右声道信号。

三维环绕的 5.1 虚拟音箱是终端预置的音频模型，其模拟了现实场景中环绕在用户周围的 5.1 声道音箱的播放效果。

现实场景中，以用户为中心，用户的脸朝向的方向为正前方，5.1 声道的音箱包括位于用户左前方的前置左音箱、位于用户右前方的前置右音箱、位于用户正前方的前置中央音箱、低频音箱（不限定位置）、位于用户左后方的后置左音箱以及位于用户右后方的后置右音箱。

步骤 104，将处理后的 5.1 声道音频信号，合成为第二立体声音频信号。

终端将处理后的 5.1 声道音频信号，合成为第二立体声音频信号。该第二立体声音频信号可通过普通的立体声耳机或 2.0 音箱等播放，用户在听到普通的立体声耳机或 2.0 音箱第二立体声音频信号后会有 5.1 声道立体声效果。

综上所述，本实施例提供的方法，通过将第一立体声音频信号拆分为 5.1 声道音频信号，再将 5.1 声道音频信号处理并合成为第二立体声音频信号，通过双声道的音频播放单元播放该第二立体声音频信号使得用户获得 5.1 声道音频的立体声效果，解决了相关技术中仅播放双声道音频信号所带来的立体效果

较差的问题，提高了音频播放的立体效果。

图 1 实施例中，将第一立体声音频信号拆分为 5.1 声道音频信号分为两个阶段，第一阶段是获取 5.1 声道音频信号中的 5.0 声道音频信号，下述图 2、图 3 以及图 4 的实施例将对从第一立体声音频信号拆分出 5.0 声道音频信号进行阐述；第二阶段是获取 5.1 声道音频信号中的 0.1 声道音频信号，下述图 5 的实施例将对从第一立体声音频信号拆分出 0.1 声道音频信号进行阐述；第三阶段是将 5.0 声道音频信号和 0.1 声道音频信号合成为第二立体声音频信号，下述图 6 和图 8 的实施例提供了对该 5.1 声道音频信号进行处理和合成，得到第二立体声音频信号的方法。

图 2，示出了本申请一个示例性的实施例提供的音频信号的处理方法的方法流程图，该方法可由具有音频信号处理功能的终端执行，该方法可以是图 1 实施例中的步骤 102 和步骤 103 的一种可选的实施方式，该方法包括：

步骤 201，对第一立体声音频信号输入高通滤波器进行滤波，得到第一高频信号。

终端对第一立体声音频信号输入高通滤波器进行滤波，得到第一高频信号。其中，第一高频信号为第一左声道高频信号和第一右声道高频信号的叠加信号。

可选的，终端通过 4 阶的 IIR 高通滤波器对第一立体声滤波，得到第一高频信号。

步骤 202，根据第一高频信号，计算得到左声道高频信号、中央声道高频信号和右声道高频信号。

终端将第一高频信号拆分为左声道高频信号、中央声道高频信号和右声道高频信号。其中，左声道高频信号包含前置左声道信号和后置左声道信号，中央声道高频信号包含前置中央声道信号，右声道高频信号包含前置右声道信号和后置右声道信号。

可选的，终端根据第一高频信号中计算得到中央声道高频信号，将第一左声道高频信号减去中央声道高频信号得到左声道高频信号，将第一右声道高频信号减去中央声道高频信号得到右声道高频信号。

步骤 203，根据左声道高频信号、中央声道高频信号和右声道高频信号，

计算得到 5.1 声道音频信号中的前置左声道信号、前置右声道信号、前置中央声道信号、后置左声道信号和后置右声道信号。

终端根据左声道高频信号中计算得到前置左声道信号和后置左声道信号，根据右声道高频信号中计算得到前置右声道信号和后置右声道信号，根据中央声道高频信号计算得到前置中央声道信号。

可选的，终端提取左声道高频信号中的第一后方/混响信号数据、中央声道高频信号中的第二后方/混响信号数据、右声道高频信号中的第三后方/混响信号数据，根据第一后方/混响信号数据、第二后方/混响信号数据以及第三后方/混响信号数据计算前置左声道信号、后置左声道信号、前置右声道信号、后置右声道信号和前置中央声道信号。

步骤 204，将前置左声道信号、前置右声道信号、前置中央声道信号、后置左声道信号和后置右声道信号分别与对应的音箱参数进行标量相乘，得到处理后的前置左声道信号、处理后的前置右声道信号、处理后的前置中央声道信号、处理后的后置左声道信号和处理后的后置右声道信号。

可选的，终端将所前置左声道信号与虚拟前置左声道音箱的音量 V_1 进行标量相乘，得到处理后的前置左声道信号 X_{FL} ；将前置右声道信号与虚拟前置右声道音箱的音量 V_2 进行标量相乘，得到处理后的前置右声道信号 X_{FR} ；将前置中央声道信号与虚拟前置中央声道音箱的音量 V_3 进行标量相乘，得到处理后的前置中央声道信号 X_{FC} ；将后置左声道信号与虚拟后置左声道音箱的音量 V_4 进行标量相乘，得到处理后的后置左声道信号 X_{RL} ；将所后置右声道信号与虚拟后置右声道音箱的音量 V_5 进行标量相乘，得到处理后的后置右声道信号 X_{RR} 。

综上所述，本实施例提供的方法，通过将第一立体声音频信号滤波得到第一高频信号，根据第一高频信号计算得到左声道高频信号、中央声道高频信号和右声道高频信号，根据左声道高频信号、中央声道高频信号和右声道高频信号计算得到 5.0 声道音频信号，进而得到处理后 5.0 声道音频信号，从而实现了将第一高频信号从第一立体声音频信号中提取并拆分为 5.1 声道音频信号中的 5.0 声道音频信号，并进一步得到处理后的 5.0 声道音频信号。

图 3，示出了本申请一个示例性的实施例提供的音频信号的处理方法的方法流程图，该方法应用于具有音频信号处理功能的终端中，该方法可以是图 2

实施例中的步骤 202 的一种可选的实施方式，该方法包括：

步骤 301，对第一高频信号进行快速傅里叶变换（Fast Fourier transform, FFT），得到高频实数信号和高频虚数信号。

终端对第一高频信号进行快速傅里叶变换后，得到高频实数信号和高频虚数信号。

快速傅里叶变换是将时域的信号转化为频域信号的算法。本实施例中，第一高频信号通过快速傅里叶变换得到高频实数信号和高频虚数信号。其中，高频实数信号包括左声道高频实数信号和右声道高频实数信号，高频虚数信号包括左声道高频虚数信号和右声道高频虚数信号。

步骤 302，根据高频实数信号和高频虚数信号计算向量投影。

终端将高频实数信号中的左声道高频实数信号和右声道高频实数信号相加，得到高频实数和信号。

示例性的，高频实数和信号通过以下公式计算：

$$\text{sumRE} = X_HIPASS_RE_L + X_HIPASS_RE_R$$

其中， $X_HIPASS_RE_L$ 为左声道高频实数信号， $X_HIPASS_RE_R$ 为右声道高频实数信号， sumRE 为高频实数和信号。

终端将高频虚数信号中的左声道高频虚数信号和右声道高频虚数信号相加，得到高频虚数和信号。

示例性的，高频虚数和信号通过以下公式计算：

$$\text{sumIM} = X_HIPASS_IM_L + X_HIPASS_IM_R$$

其中， $X_HIPASS_IM_L$ 为左声道高频虚数信号， $X_HIPASS_IM_R$ 为右声道高频虚数信号， sumIM 为高频虚数和信号。

终端将高频实数信号中的左声道高频实数信号和右声道高频实数信号相减，得到高频实数差信号。

示例性的，高频实数差信号通过以下公式计算：

$$\text{diffRE} = X_HIPASS_RE_L - X_HIPASS_RE_R$$

其中， diffRE 为高频实数差信号。

终端将高频虚数信号中的左声道高频虚数信号和右声道高频虚数信号相减，得到高频虚数差信号。

示例性的，高频虚数差信号通过以下公式计算：

$$\text{diffIM} = X_HIPASS_IM_L - X_HIPASS_IM_R$$

其中，diffIM 为高频虚数差信号。

终端根据高频实数和信号和所述高频虚数和信号，计算得到实数和信号。

示例性的，实数和信号通过以下公式计算：

$$\text{sumSq} = \text{sumRE} * \text{sumRE} + \text{sumIM} * \text{sumIM}$$

其中，sumSq 为实数和信号。

终端根据高频实数差信号和所述高频虚数差信号，计算得到实数差信号。

示例性的，实数差信号通过以下公式计算：

$$\text{diffSq} = \text{diffRE} * \text{diffRE} + \text{diffIM} * \text{diffIM}$$

其中，diffSq 为实数差信号。

终端根据实数和信号和实数差信号，进行向量投影计算，得到向量投影，向量投影代表了三维环绕的 5.1 虚拟音箱中每个虚拟音箱到用户的距离。

可选的，当实数和信号为有效数字时，即当实数和信号不是无穷小或 0 时，向量投影通过以下公式计算：

$$\alpha = 0.5 - \text{SQRT}(\text{diffSq} / \text{sumSq}) * 0.5$$

其中，alpha 为向量投影，SQRT 代表开平方，* 代表标量乘积。

步骤 303，对高频实数信号中的左声道高频实数信号和向量投影的乘积进行快速傅里叶逆变换（Inverse fast Fourier transform, IFFT）和交迭相加（Overlap-Add）后，得到中央声道高频信号。

快速傅里叶逆变换是将频域信号转换为时域信号的算法，本申请中，终端对高频实数信号中的左声道高频实数信号和向量投影的乘积进行快速傅里叶逆变换和交迭相加后，得到中央声道高频信号，其中，交迭相加是一种数学算法，具体可参考 https://en.wikipedia.org/wiki/Overlap-add_method。中央声道高频信号可通过左声道高频实数信号或右声道高频实数信号计算，但是由于第一立体声信号中若只包含一个声道的音频信号，则音频信号大部分集中在左声道，因此中央高频信号通过左声道高频实数计算会更加准确。

步骤 304，将第一高频信号中的左声道高频信号和中央声道信号的差，作为左声道高频信号。

终端将第一高频信号中的左声道高频信号和中央声道信号的差，作为左声道高频信号。

示例性的，左声道高频信号通过以下公式计算：

$$X_PRE_L = X_HIPASS_L - X_PRE_C$$

其中，X_HIPASS_L 为第一高频信号中的左声道高频信号，X_PRE_C 为中央声道信号，X_PRE_L 为左声道高频信号。

步骤 305，将第一高频信号中的右声道高频信号和中央声道信号的差，作为右声道高频信号。

终端将第一高频信号中的右声道高频信号和中央声道信号的差，作为右声道高频信号。

示例性的，右声道高频信号通过以下公式计算：

$$X_PRE_R = X_HIPASS_R - X_PRE_C$$

其中，X_HIPASS_R 为第一高频信号中的右声道高频信号，X_PRE_C 为中央声道信号，X_PRE_R 为右声道高频信号。

步骤 304 和步骤 305 的执行顺序不加以限定，终端可先执行步骤 304 再执行步骤 305，或先执行步骤 305 再执行步骤 304。

综上所述，本实施例提供的方法，通过将第一高频信号进行快速傅里叶变换得到高频实数信号和高频虚数信号，根据高频实数信号和高频虚数信号通过一些列计算得到中央高频信号，进而根据中央高频信号计算得到左声道高频信号和右声道高频信号，从而实现了根据第一高频信号计算得到左声道高频信号、中央声道高频信号和右声道高频信号。

图 4，示出了本申请一个示例性的实施例提供的音频信号的处理方法的方法流程图，该方法可由具有音频信号处理功能的终端执行，该方法可以是图 2 实施例中的步骤 203 的一种可选的实施方式，该方法包括：

在步骤 401 中，对于左声道高频信号、中央声道高频信号和右声道高频信号中的任意一个声道高频信号，根据声道高频信号中的采样点得到至少一个移动窗，每个移动窗包括 n 个采样点，相邻的两个移动窗存在 n/2 个采样点是重叠的。

终端通过移动窗（Moving window）算法对左声道高频信号、中央声道高频信号和右声道高频信号中的任意一个声道高频信号，根据声道高频信号中的采样点得到至少一个移动窗。其中，若每个移动窗的采样点为 n 个，则相邻的两个移动窗之间 n/2 个采样点为重叠的， $n \geq 1$ 。

移动窗是一种类似交迭相加的算法，但只做交迭，不做相加。例如，数据 A 包含 1024 个采样点，若移动步长为 128，重叠长度为 64，那么移动窗每次

输出的信号为：第一次输出 $A[0-128]$ ，第二次输出 $A[64-192]$ ，第三次输出 $A[128-256]$ ，……，其中， A 为移动窗，方括号内为采样点的编号。

步骤 402，计算移动窗中的低相关信号以及低相关信号的起始时间点，低相关信号包括幅度谱的第一衰减包络序列和相位谱的第二衰减包络序列不相等的信号。

终端对第 i 个移动窗中的采样点信号进行快速傅里叶变换，得到快速傅里叶变换后的采样点信号， $i \geq 1$ 。

终端根据预设的移动步长和重叠长度，对左声道高频信号、右声道高频信号和中央声道信号分别进行移动窗和快速傅里叶变换，依次得到左声道高频实数信号和左声道高频虚数信号（记为 FFT_L ）、右声道高频实数信号以及右声道高频虚数信号（记为 FFT_R ）、中央声道实数信号和中央声道虚数信号（记为 FFT_C ）。

终端计算快速傅里叶变换后的采样点信号的幅度谱和相位谱。

终端根据 FFT_L 计算左声道高频信号的幅度谱 AMP_L 以及左声道高频信号的相位谱 PH_L ；根据 FFT_R 计算右声道高频信号的幅度谱 AMP_R 以及左声道高频信号的相位谱 PH_R ；根据 FFT_C 计算中央声道信号的幅度谱 AMP_C 以及中央声道信号的相位谱 PH_C 。

以下将 AMP_L 、 AMP_R 以及 AMP_C 统一记为 $AMP_L/R/C$ ；将 PH_L 、 PH_R 、 PH_C 统一记为 $PH_L/R/C$ 。

终端根据快速傅里叶变换后的采样点信号的幅度谱，计算第 i 个移动窗中的 m 条频率线的第一衰减包络序列；根据快速傅里叶变换后的采样点信号的相位谱，计算第 i 个移动窗中的 m 条频率线的第二衰减包络序列；当 m 条频率线中的第 j 条频率线的衰减包络序列和第二衰减包络序列不同时，确定第 j 条频率线为低相关信号；根据第 i 个移动窗的窗口号和第 j 条频率线的频率线号，确定低相关信号的起始时间点，其中， $m \geq 1$ ， $1 \leq j \leq m$ 。

终端对所有移动窗的 $AMP_L/R/C$ 和 $PH_L/R/C$ 分别计算其所有频率线的衰减包络序列和相关度，其中，计算移动窗之间的衰减包络序列，对应同一个移动窗的幅度谱和相位谱为有效条件。

例如，移动窗 1、移动窗 2、移动窗 3 对应的 0 号频率线的频率谱的衰减包络序列分别为 1.0、0.8、0.6，移动窗 1、移动窗 2、移动窗 3 对应的 0 号频率线的相位谱的衰减包络序列分别为 1.0、0.8、1.0，则认为移动窗 1 的 0 号频

率线和移动窗 2 的 0 号频率线具有高度相关性，移动窗 2 的 0 号频率线和移动窗 3 的 0 号频率线具有低度相关性。

n 个采样点经过快速傅里叶变换后会得到 $n/2+1$ 条频率线，取出低相关度的信号对应的移动窗的窗口号及频率线，通过窗口号可计算出该信号在 X_PRE_L、X_PRE_R 和 X_PRE_C 中的起始时间点。

步骤 403，确定符合后方/混响特征的目标低相关信号。

可选的，终端通过以下方式确定符合后方/混响特征的目标低相关信号：

当低相关信号的甚高频率线的幅度谱能量小于第一阈值且甚高频率线所在窗口的相邻窗口的衰减包络斜率大于第二阈值时，终端确定低相关信号是符合后方/混响特征的目标低相关信号，其中，甚高频 (Very high frequency, VHF) 频率线是指频带由 30Mhz 到 300MHz 的频率线。

可选的，终端确定符合后方/混响特征的目标低相关信号的方法包括但不限于：

当低相关信号的甚高频率线的幅度谱能量小于第一阈值且甚高频率线所在窗口的相邻窗口的衰减速度大于第三阈值时，终端确定低相关信号是符合后方/混响特征的目标低相关信号。

步骤 404，计算目标低相关信号的结束时间点。

可选的，终端通过以下方式计算低相关信号的结束时间点：

终端获取目标低相关信号的幅度谱对应的频率线的能量小于第四阈值的时间点，作为结束时间点。

可选的，终端通过以下方式计算低相关信号的结束时间点：

当目标低相关信号的能量小于下一个低相关信号的能量 $1/n$ 时，终端确定下一个低相关信号的起始时间点作为目标低相关信号的结束时间点。

步骤 405，根据起始时间点和结束时间点提取目标低相关信号，作为声道高频信号中的后方/混响信号数据。

可选的，终端提取位于起始时间点和结束时间点中的声道信号片段；对声道信号片段进行快速傅里叶变换，得到快速傅里叶变换后的信号片段；从快速傅里叶变换后的信号片段中提取目标低相关信号对应的频率线，得到第一部分信号；对第一部分进行快速傅里叶逆变换和交迭相加后，得到声道高频信号中的后方/混响信号数据。

通过上述步骤，终端得到左声道高频信号中的第一后方/混响信号数据、中

央声道高频信号中的第二后方/混响信号数据、右声道高频信号中的第三后方/混响信号数据。

步骤 406, 根据第一后方/混响信号数据、第二后方/混响信号数据、第三后方/混响信号数据计算前置左声道信号、后置左声道信号、前置右声道信号、后置右声道信号和前置中央声道信号。

终端将左声道高频信号和上述步骤中获得的第一后方/混响信号数据的差, 确定为前置左声道信号。

第一后方/混响信号数据是左声道高频信号中包含的音频数据, 是三维环绕的 5.1 虚拟音箱的后置左声道信号包含的音频数据, 而左声道高频信号包含前置左声道信号和部分后置左声道信号, 因此将左声道高频信号减去部分后置左声道信号的部分, 即第一后方/混响信号数据, 即可得到前置左声道信号。

终端将上述步骤中获得的第一后方/混响信号数据和第二后方/混响信号数据的和, 确定为后置左声道信号。

终端将右声道高频信号和上述步骤中获得第三后方/混响信号数据的差, 确定为前置右声道信号。

第三后方/混响信号数据是右声道高频信号中包含的音频数据, 是三维环绕的 5.1 虚拟音箱的后置右声道信号包含的音频数据, 而右声道高频信号包含前置右声道信号和部分后置右声道信号, 因此将右声道高频信号减去部分后置右声道信号的部分, 即第三后方/混响信号数据, 即可得到前置右声道信号。

终端将上述步骤中获得的第三后方/混响信号数据和第二后方/混响信号数据的和, 确定为后置右声道信号。

终端将中央声道高频信号和上述步骤中获得的第二后方/混响信号数据的差, 确定为前置中央声道信号。

第二后方/混响信号数据为三维环绕的 5.1 虚拟音箱的后置左声道信号包含的音频数据和后置右声道信号包含的音频数据, 中央声道高频信号包括前置中央声道信号和第二后方/混响信号数据, 因此将中央声道高频信号减去第二后方/混响信号数据。

综上所述, 本实施例提供的方法, 通过计算每个声道高频信号中的后方/混响信号数据的起始时间和结束时间提取每个声道高频信号中的后方/混响信号数据, 根据每个声道高频信号中的后方/混响信号数据计算得到前置左声道信号、后置左声道信号、前置右声道信号、后置右声道信号以及前置中央声道信

号，提高了根据左声道高频信号、中央声道高频信号和右声道高频信号计算得到 5.1 声道音频信号的准确度。

图 5，示出了本申请一个示例性的实施例提供的音频信号的处理方法的方法流程图，该方法可由具有音频信号处理功能的终端执行，该方法可以是图 1 实施例中步骤 102 的一个可选的实施例，该方法包括：

步骤 501，对第一立体声音频信号输入低通滤波器进行滤波，得到第一低频信号。

终端对第一立体声音频信号输入低通滤波器进行滤波，得到第一低频信号。其中，第一低频信号为第一左声道低频信号和第一右声道低频信号的叠加信号。

可选的，终端通过 4 阶的 IIR 低通滤波器对第一立体声滤波，得到第一低频信号。

步骤 502，对第一立低频信号和 5.1 虚拟音箱中的低频声道音箱的音量参数进行标量相乘，得到第二低频信号。

终端将第一低频信号和 5.1 虚拟音箱中的低频声道音箱的音量参数进行标量相乘，得到第二低频信号。

示例性的，终端通过以下公式计算第二低频信号：

$$X_LFE_S = X_LFE * V6$$

其中， X_LFE 为第一立体声低频信号， $V6$ 为 5.1 虚拟音箱中的低频声道音箱的音量参数， X_LFE_S 为第二低频信号，是第一左声道低频信号 $X_LFE_S_L$ 和第一右声道低频信号 $X_LFE_S_R$ 的叠加信号，* 代表标量乘法。

步骤 503，对第二低频信号进行单声道转换，得到处理后的低频声道信号。

终端对第二低频信号进行单声道转换，得到处理后的低频声道信号。

示例性的，终端通过以下公式计算处理后的低频声道信号：

$$X_LFE_M = (X_LFE_S_L + X_LFE_S_R) / 2$$

其中， X_LFE_M 为处理后的低频声道信号。

综上所述，本实施例提供的方法，通过将第一立体声音频信号滤波得到第一低频信号，将第一低频信号进行单声道转换，得到 5.1 声道音频信号中的低频声道信号，从而实现了将第一低频信号从第一立体声信号中提取并拆分为 5.1 声道音频信号中的 0.1 声道音频信号。

上述方法实施例将第一立体声音频信号拆分并处理后，得到了 5.1 声道音频信号，分别为前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号，下述图 6 和图 8 的实施例提供了对该 5.1 声道音频信号进行处理和合成，得到第二立体声音频信号的方法，该方法可以是图 1 实施例中步骤 104 的可选的实施例，也可作为单独的实施例，图 6 和图 8 实施例中得到的立体声信号可以是上述方法实施例中的第二立体声信号。

HRTF (Head Related Transfer Function, 头相关变换函数) 处理技术是一种产生立体声环绕音效的处理技术。技术人员可以预先构建 HRTF 数据库，该 HRTF 数据库中记录有 HRTF 数据、HRTF 数据采集点、HRTF 数据采集点相对于参考人头的位置坐标之间的对应关系。HRTF 数据是用于对左声道音频信号和右声道音频信号进行处理的一组参数。

图 6，示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图，该方法可由具有音频信号处理功能的终端执行，该方法可以是图 1 实施例中的步骤 104 的一种可选的实施方式，该方法包括：

步骤 601，获取得到 5.1 声道音频信号；

可选地，该 5.1 声道音频信号是上述图 1 至图 5 实施例中从第一立体声音频信号中拆分出并处理后得到的处理后的 5.1 声道音频信号。或者，该 5.1 声道音频信号是下载或从存储介质上读取到的 5.1 声道音频信号。

该 5.1 声道音频信号包括：前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号。

步骤 602，根据 5.1 虚拟音箱在虚拟环境中的坐标，获取 5.1 虚拟音箱中每个虚拟音箱对应的 HRTF 数据；

可选地，5.1 虚拟音箱包括：前置左声道虚拟音箱 FL、前置右声道虚拟音箱 FR、前置中央声道虚拟音箱 FC、重低音虚拟音箱 LFE、后置左声道虚拟音箱 RL 和后置右声道虚拟音箱 RR。

可选地，该 5.1 虚拟音箱在虚拟环境中具有各自的坐标。该虚拟环境可以是二维平面虚拟环境，也可以是三维虚拟环境平面虚拟环境。

示意性的参考图 7，其示出了一种 5.1 声道虚拟音箱在二维平面虚拟环境中的示意图，假设参考人头处于图 7 中的中心点 70 并朝向中央声道虚拟音箱 FC 所在位置，每个声道与参考人头所在的中心点 70 的距离相等且处于同一平面。

前置中央声道虚拟音箱 FC 处于参考人头的面对方向的正前方。

前置左声道虚拟音箱 FL 和前置右声道虚拟音箱 FR 分别处于前置中央声道 FC 的两侧，分别与参考人头的面对方向呈 30 度夹角，呈对称设置。

后置左声道虚拟音箱 RL 和后置右声道虚拟音箱 RR 分别处于参考人头的面对方向的两侧靠后，分别与参考人头的面对方向呈 100-120 度夹角，呈对称设置。

由于重低音虚拟音箱 LFE 的方向感较弱，重低音虚拟音箱 LFE 的摆放位置没有严格要求，本文中以参考人头的背对方向来举例说明，但本申请不对重低音虚拟音箱 LFE 与参考人头的面对方向的夹角作出限定。

需要说明的一点，上述 5.1 声道虚拟音箱中的每个虚拟音箱与参考人头的面对方向的夹角仅是示例性的，另外，每个虚拟音箱与参考人头之间的距离可以不同。当虚拟环境为三维虚拟环境时，每个虚拟音箱所在的高度也可以不同，每个虚拟音箱的摆放位置的不同都会引起声音信号的不同，本公开对此不作限定。

可选地，以参考人头为原点为二维虚拟环境或三维虚拟环境建立坐标系后，能够得到每个虚拟音箱在虚拟环境中的坐标。

终端内存储有 HRTF 数据库，该 HRTF 数据库包括：至少一个 HRTF 数据采集点和 HRTF 数据之间的对应关系，每个 HRTF 数据采集点具有各自的坐标。

终端根据 5.1 虚拟音箱中的第 i 个虚拟音箱的第 i 坐标，在 HRTF 数据库中查询与第 i 坐标最接近的 HRTF 数据采集点，将与第 i 坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为第 i 个虚拟音箱的 HRTF 数据， $i \geq 1$ 。

步骤 603，根据每个虚拟音箱对应的 HRTF 数据，对 5.1 声道音频信号中的相应声道音频信号进行处理，得到处理后的 5.1 声道音频信号；

可选地，每个 HRTF 数据包括左声道 HRTF 系数和右声道 HRTF 系数。

终端根据第 i 个虚拟音箱对应的 HRTF 数据中的左声道 HRTF 系数，对 5.1 声道音频信号中的第 i 个声道音频信号进行处理，得到处理后的第 i 个声道音频信号对应的左声道分量；

终端根据第 i 个虚拟音箱对应的 HRTF 数据中的右声道 HRTF 系数, 对 5.1 声道音频信号中的第 i 个声道音频信号进行处理, 得到处理后的第 i 个声道音频信号对应的右声道分量。

步骤 604, 将处理后的 5.1 声道音频信号, 合成为立体声音频信号。

需要说明的是, 当本申请实施中的 5.1 声道音频信号是从上述图 1 至图 5 实施例中的第一立体声音频信号中拆分出并处理后得到的处理后的 5.1 声道音频信号时, 该步骤中的立体声音频信号是图 1 实施例中的第二立体声音频信号。

综上所述, 本实施例提供的方法, 通过将 5.1 声道音频信号按照各个 5.1 虚拟音箱的 HRTF 数据进行处理后, 合成得到立体声音频信号, 使得用户只需要普通的立体声耳机或 2.0 音箱也能够播放 5.1 声道音频信号, 且获得较好的播放音质。

图 8, 示出了本申请一个示例性实施例提供的音频信号的处理方法的流程图, 该方法可由具有音频信号处理功能的终端执行, 该方法可以是图 1 实施例中的步骤 104 的一种可选的实施方式, 该方法包括:

步骤 801, 在声学房间中采集一系列以参考人头为球心的至少一条 HRTF 数据, 并记录各条 HRTF 数据对应 HRTF 数据采集点相对于参考人头的位坐标;

参考图 9, 开发人员预先在声学房间 91 (房间四周设置有吸音海绵以减小回声干扰) 中央放置参考人头 92 (模仿真人头部制成), 并将微型全指向性麦克风分别设置在参考人头 92 的左右耳道内。

完成参考人头 92 设置后, 开发人员在以参考人头 92 为球心的球体表面上, 每隔预定距离设置 HRTF 数据采集点, 并在 HRTF 数据采集点处使用扬声器 93 播放预定音频。

由于左右耳道到扬声器 93 的距离不同, 且声波在传输过程中受到折射、绕射和衍射等因素影响, 同一音频达到左右耳道时音频特征不同。因此, 通过分析麦克风采集到的音频与原始音频的差异, 即可得到 HRTF 数据采集点处的 HRTF 数据。其中, 同一 HRTF 数据采集点对应的 HRTF 数据中, 包含左声道对应的左声道 HRTF 系数以及右声道对应的右声道 HRTF 系数。

步骤 802, 根据 HRTF 数据、HRTF 数据采集点的标识和 HRTF 数据采集

点的位置坐标，生成 HRTF 数据库；

可选地，以参考人头 92 为中心点建立坐标系。该坐标系的建立方式与 5.1 声道虚拟音箱的坐标系建立方式是相同的。

当 5.1 声道虚拟音箱对应的虚拟环境是二维虚拟环境时，在采集 HRTF 数据时也可以仅对参考人头 92 所在的水平面建立坐标系，仅采集属于该水平面的 HRTF 数据。比如，在以参考人头 92 为圆心的圆环上，每隔 5° 取一个点作为 HRTF 数据采样点。此时，可以减少终端所需要存储的 HRTF 数据量。

当 5.1 声道虚拟音箱对应的虚拟环境是三维虚拟环境时，在采集 HRTF 数据时可以对以参考人头 92 所在的三维环境建立坐标系，采集与该参考人头 92 为球心的球体表面上的 HRTF 数据。比如，在以参考人头 92 为球心的球体表面上，按照经度方向和纬度方向每隔 5° 取一个点作为 HRTF 数据采样点。

然后，终端根据每个 HRTF 数据采样点的标识、每个 HRTF 数据采样点的 HRTF 数据和每个 HRTF 数据采集点的位置坐标，生成 HRTF 数据库。

需要说明的是，步骤 801 和步骤 802 也可以由其它设备执行和实现。在生成 HRTF 数据库后，再通过网络或存储介质传输到当前终端上。

步骤 803，获取 5.1 声道音频信号；

可选地，终端获取 5.1 声道音频信号。

该 5.1 声道音频信号是上述图 1 至图 5 实施例从第一立体声音频信号中拆分出并处理后得到的处理后的 5.1 声道音频信号。或者，该 5.1 声道音频信号是下载或从存储介质上读取到的 5.1 声道音频信号。

该 5.1 声道音频信号包括：前置左声道信号 X_FL、前置右声道信号 X_FC、前置中央声道信号 X_FC、低频声道信号 X_LFE_M、后置左声道信号 X_RL 和后置右声道信号 X_RR。

步骤 804，获取 HRTF 数据库，HRTF 数据库包括：至少一个 HRTF 数据采集点和 HRTF 数据之间的对应关系，每个 HRTF 数据采集点具有各自的坐标；

终端可以读取存储在本地的 HRTF 数据库，或者，访问存储在网络上的 HRTF 库。

步骤 805，根据 5.1 虚拟音箱中的第 i 个虚拟音箱的第 i 坐标，在 HRTF 数据库中查询与第 i 坐标最接近的 HRTF 数据采集点，将与第 i 坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为第 i 个虚拟音箱的 HRTF 数据。

可选地，终端预先存储有 5.1 虚拟音箱中的各个虚拟音箱的坐标。其中， i

≥ 1 。

终端根据前置左声道虚拟音箱的第一坐标，在 HRTF 数据库中查询与第一坐标最接近的 HRTF 数据采集点，将与第一坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为前置左声道虚拟音箱的 HRTF 数据。

终端根据前置右声道虚拟音箱的第二坐标，在 HRTF 数据库中查询与第二坐标最接近的 HRTF 数据采集点，将与第二坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为前置右声道虚拟音箱的 HRTF 数据。

终端根据前置中央声道虚拟音箱的第三坐标，在 HRTF 数据库中查询与第三坐标最接近的 HRTF 数据采集点，将与第三坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为前置中央声道虚拟音箱的 HRTF 数据。

终端根据后置左声道虚拟音箱的第四坐标，在 HRTF 数据库中查询与第四坐标最接近的 HRTF 数据采集点，将与第四坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为后置左声道虚拟音箱的 HRTF 数据。

终端根据后置右声道虚拟音箱的第五坐标，在 HRTF 数据库中查询与第五坐标最接近的 HRTF 数据采集点，将与第五坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为后置右声道虚拟音箱的 HRTF 数据。

终端根据低频虚拟音箱的第六坐标，在 HRTF 数据库中查询与第六坐标最接近的 HRTF 数据采集点，将与第六坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为低频虚拟音箱的 HRTF 数据。

其中，“最接近”是指虚拟音箱的坐标和 HRTF 数据采样点的坐标相同或坐标间的距离最短。

步骤 806，对于 5.1 声道音频信号中的第 i 个声道的音频信号，采用第 i 个虚拟音箱对应的 HRTF 数据中的左声道 HRTF 系数进行第一卷积，得到第一卷积后的第 i 个声道的音频信号；

设 5.1 声道音频信号中的第 i 个声道的音频信号为 X_i ，计算 $L_i = X_i * H_{L_i}$ 。其中， $*$ 表示卷积， H_{L_i} 表示第 i 个虚拟音箱对应的 HRTF 数据中的左声道 HRTF 系数。

步骤 807，将第一卷积后的各个声道的音频信号进行叠加，得到立体声音频信号中的左声道信号；

终端将第一卷积后的 6 个声道的音频信号 L_i 进行叠加，得到立体声音频信号中的左声道信号 $L = L_1 + L_2 + L_3 + L_4 + L_5 + L_6$ 。

步骤 808, 对于 5.1 声道音频信号中的第 i 个声道的音频信号, 采用第 i 个虚拟音箱对应的 HRTF 数据中的右声道 HRTF 系数进行第二卷积, 得到第二卷积后的第 i 个声道的音频信号;

设 5.1 声道音频信号中的第 i 个声道的音频信号为 X_i , 计算 $R_i = X_i * H_{R_i}$ 。其中, $*$ 表示卷积, H_{R_i} 表示第 i 个虚拟音箱对应的 HRTF 数据中的右声道 HRTF 系数。

步骤 809, 将第二卷积后的各个声道的音频信号进行叠加, 得到立体声音频信号中的右声道信号;

终端将第二卷积后的 6 个声道的音频信号 R_i 进行叠加, 得到立体声音频信号中的右声道信号 $R = R_1 + R_2 + R_3 + R_4 + R_5 + R_6$ 。

步骤 810, 将左声道信号和右声道信号, 合成为立体声音频信号。

该合成的立体声音频信号可以存储为音频文件, 或者输入播放设备中进行播放。

需要说明的是, 当本申请实施中的 5.1 声道音频信号是从上述图 1 至图 5 实施例中从第一立体声音频信号中拆分出并处理后得到的处理后的 5.1 声道音频信号时, 该步骤中的立体声音频信号是图 1 实施例中的第二立体声音频信号。

综上所述, 本实施例提供的方法, 通过将 5.1 声道音频信号按照各个 5.1 虚拟音箱的 HRTF 数据进行处理后, 合成得到立体声音频信号, 使得用户只需要普通的立体声耳机或 2.0 音箱也能够播放 5.1 声道音频信号, 且获得较好的播放音质。

本实施例提供的方法, 通过将 5.1 声道音频信号按照各个 5.1 虚拟音箱的 HRTF 数据分别进行卷积和叠加, 能够获得具有较好的三维环绕音效的立体声音频信号, 该立体声音频信号在播放时具有较好的三维环绕效果。

图 10, 示出了本申请一个示例性实施例提供的音频信号的处理装置的结构框图, 该装置可以实现成为终端或终端中的一部分。该装置包括:

获取模块 1010, 用于获取第一立体声音频信号;

处理模块 1020, 用于将第一立体声音频信号拆分为 5.1 声道音频信号; 对 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信号处理, 得到处理后的 5.1 声道音频信号;

合成模块 1030, 用于将处理后的 5.1 声道音频信号, 合成为立体声音频信号。

在一个可选的实施例中, 该装置还包括计算模块 1040;

处理模块 1020, 用于对第一立体声音频信号输入高通滤波器进行滤波, 得到第一高频信号;

计算模块 1040, 用于根据第一高频信号, 计算得到左声道高频信号、中央声道高频信号和右声道高频信号; 根据左声道高频信号、中央声道高频信号和右声道高频信号, 计算得到 5.1 声道音频信号中的前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号。

在一个可选的实施例中, 计算模块 1040, 还用于对第一高频信号进行快速傅里叶变换, 得到高频实数信号和高频虚数信号; 根据高频实数信号和高频虚数信号计算向量投影; 对高频实数信号中的左声道高频实数信号和计算向量投影的乘积进行快速傅里叶逆变换, 得到中央声道高频信号; 将第一高频信号中的左声道高频信号和中央声道信号的差, 作为左声道高频信号; 将第一高频信号中的右声道高频信号和中央声道信号的差, 作为右声道高频信号。

计算模块 1040, 还用于将高频实数信号中的左声道高频实数信号和右声道高频实数信号相加, 得到高频实数和信号; 将高频虚数信号中的左声道高频虚数信号和右声道高频虚数信号相加, 得到高频虚数和信号; 将高频实数信号中的左声道高频实数信号和右声道高频实数信号相减, 得到高频实数差信号; 将高频虚数信号中的左声道高频虚数信号和右声道高频虚数信号相减, 得到高频虚数差信号; 根据高频实数和信号和高频虚数和信号, 计算得到实数和信号; 根据高频实数差信号和高频虚数差信号, 计算得到实数差信号; 根据实数和信号和实数差信号, 进行向量投影计算, 得到向量投影。

在一个可选的实施例中,

计算模块 1040, 还用于当实数和信号为有效数字时, 按照如下公式计算向量投影:

$$\alpha = 0.5 - \text{SQRT}(\text{diffSQ} / \text{sumSQ}) * 0.5$$

其中, α 为向量投影, diffSq 为实数差信号, sumSQ 为实数和信号, SQRT 代表开平方, $*$ 代表标量乘法。

在一个可选的实施例中,

处理模块 1020, 还用于提取左声道高频信号中的第一后方/混响信号数据、中央声道高频信号中的第二后方/混响信号数据、右声道高频信号中的第三后方/混响信号数据;

计算模块 1040, 还用于将左声道高频信号和第一后方/混响信号数据的差, 确定为前置左声道信号; 将第一后方/混响信号数据和第二后方/混响信号数据的和, 确定为后置左声道信号; 将右声道高频信号和第三后方/混响信号数据的差, 确定为前置右声道信号; 将第三后方/混响信号数据和第二后方/混响信号数据的和, 确定为后置右声道信号; 将中央声道高频信号和第三后方/混响信号数据的差, 确定为前置中央声道信号。

在一个可选的实施例中, 获取模块 1010, 还用于对于左声道高频信号、中央声道高频信号和右声道高频信号中的任意一个声道高频信号, 根据声道高频信号中的采样点得到至少一个移动窗, 每个移动窗包括 n 个采样点, 相邻的两个移动窗存在 $n/2$ 个采样点是重叠的, $n \geq 1$ 。

计算模块 1040, 还用于计算移动窗中的低相关信号以及低相关信号的起始时间点, 低相关信号包括幅度谱的第一衰减包络序列和相位谱的第二衰减包络序列不相等的信号; 确定符合后方/混响特征的目标低相关信号; 计算目标低相关信号的结束时间点; 根据起始时间点和结束时间点提取目标低相关信号, 作为声道高频信号中的后方/混响信号数据。

在一个可选的实施例中, 计算模块 1040, 还用于计算移动窗中的低相关信号以及低相关信号的起始时间点, 低相关信号包括幅度谱的第一衰减包络序列和相位谱的第二衰减包络序列不相等的信号; 确定符合后方/混响特征的目标低相关信号; 计算目标低相关信号的结束时间点; 根据起始时间点和结束时间点提取目标低相关信号, 作为声道高频信号中的后方/混响信号数据。

计算模块 1040, 还用于对第 i 个移动窗中的采样点信号进行快速傅里叶变换, 得到快速傅里叶变换后的采样点信号; 计算快速傅里叶变换后的采样点信号的幅度谱和相位谱; 根据快速傅里叶变换后的采样点信号的幅度谱, 计算第 i 个移动窗中的 m 条频率线的第一衰减包络序列; 根据快速傅里叶变换后的采样点信号的相位谱, 计算第 i 个移动窗中的 m 条频率线的第二衰减包络序列; 当 m 条频率线中的第 j 条频率线的衰减包络序列和第二衰减包络序列不同时, 确定第 j 条频率线为低相关信号; 根据第 i 个移动窗的窗口号和第 j 条频率线的频率线号, 确定低相关信号的起始时间点, $i \geq 1$, $m \geq 1$, $1 \leq j \leq m$ 。

在一个可选的实施例中，计算模块 1040，还用于当低相关信号的甚高频率线的幅度谱能量小于第一阈值且甚高频率线所在窗口的相邻窗口的衰减包络斜率大于第二阈值时，确定低相关信号是符合后方/混响特征的目标低相关信号；或，当低相关信号的甚高频率线的幅度谱能量小于第一阈值且甚高频率线所在窗口的相邻窗口的衰减速度大于第三阈值时，确定低相关信号是符合后方/混响特征的目标低相关信号。

在一个可选的实施例中，计算模块 1040，还用于获取目标低相关信号的幅度谱对应的频率线的能量小于第四阈值的时间点，作为结束时间点；或，当目标低相关信号的能量小于下一个低相关信号的能量 $1/n$ 时，确定下一个低相关信号的起始时间点作为目标低相关信号的结束时间点。

在一个可选的实施例中，获取模块 1010，还用于提取位于起始时间点和结束时间点中的声道信号片段。

计算模块 1040，还用于对声道信号片段进行快速傅里叶变换，得到快速傅里叶变换后的信号片段；从快速傅里叶变换后的信号片段中提取目标低相关信号对应的频率线，得到第一部分信号；对第一部分信号进行快速傅里叶逆变换和交迭相加后，得到声道高频信号中的后方/混响信号数据。

在一个可选的实施例中，计算模块 1040，还用于将前置左声道信号与虚拟前置左声道音箱的音量进行标量相乘，得到处理后的前置左声道信号；将前置右声道信号与虚拟前置右声道音箱的音量进行标量相乘，得到处理后的前置右声道信号；将前置中央声道信号与虚拟前置中央声道音箱的音量进行标量相乘，得到处理后的前置中央声道信号；将后置左声道信号与虚拟后置左声道音箱的音量进行标量相乘，得到处理后的后置左声道信号；将后置右声道信号与虚拟后置右声道音箱的音量进行标量相乘，得到处理后的后置右声道信号。

在一个可选的实施例中，5.1 声道音频信号包括低频声道信号；

处理模块 1020，还用于对第一立体声音频信号输入低通滤波器进行滤波，得到第一低频信号。

计算模块 1040，还用于对第一低频信号和 5.1 虚拟音箱中的低频声道音箱的音量参数进行标量相乘，得到第二低频信号；将第二低频信号进行单声道转换，得到处理后的低频声道信号。

在一个可选的实施例中，第二低频信号包括：左声道低频信号和右声道低频信号；

计算模块 1040, 还用于将左声道低频信号和右声道低频信号叠加后求平均, 将平均后的音频信号, 作为处理后的低频声道信号。

图 11, 示出了本申请一个示例性实施例提供的音频信号的处理装置的结构框图。该装置可以实现成为终端或终端中的一部分。该装置包括:

第一获取模块 1120, 用于获取 5.1 声道音频信号;

第二获取模块 1140, 用于根据 5.1 虚拟音箱在虚拟环境中的坐标, 获取 5.1 虚拟音箱中每个虚拟音箱对应的头相关变换函数 HRTF 数据;

处理模块 1160, 用于根据每个虚拟音箱对应的 HRTF 数据, 对 5.1 声道音频信号中的相应声道音频信号进行处理, 得到处理后的 5.1 声道音频信号;

合成模块 1180, 用于将处理后的 5.1 声道音频信号, 合成为立体声音频信号。

在一个可选的实施例中, 第二获取模块 1140, 用于获取 HRTF 数据库, HRTF 数据库包括: 至少一个 HRTF 数据采集点和 HRTF 数据之间的对应关系, 每个 HRTF 数据采集点具有各自的坐标; 根据 5.1 虚拟音箱中的第 i 个虚拟音箱的第 i 坐标, 在 HRTF 数据库中查询与第 i 坐标最接近的 HRTF 数据采集点, 将与第 i 坐标最接近的 HRTF 数据采集点的 HRTF 数据确定为第 i 个虚拟音箱的 HRTF 数据, $i \geq 1$ 。

在一个可选的实施例中, 该装置, 还包括:

采集模块 1112, 用于在声学房间中采集一系列以参考人头为球心的至少一条 HRTF 数据, 并记录各条 HRTF 数据对应 HRTF 数据采集点相对于参考人头的坐标;

生成模块 1114, 用于根据 HRTF 数据、HRTF 数据采集点的标识和 HRTF 数据采集点的位置坐标, 生成 HRTF 数据库。

在一个可选的实施例中, HRTF 数据包括: 左声道 HRTF 系数;

处理模块 1160, 包括:

左声道卷积单元, 用于对于 5.1 声道音频信号中的第 i 个声道的音频信号, 采用第 i 个虚拟音箱对应的 HRTF 数据中的左声道 HRTF 系数进行第一卷积, 得到第一卷积后的第 i 个声道的音频信号;

左声道合成单元, 用于将第一卷积后的各个声道的音频信号进行叠加, 得到立体声音频信号中的左声道信号。

在一个可选的实施例中，HRTF 数据包括：右声道 HRTF 系数；
处理模块 1160，包括：

右声道卷积单元，用于对于 5.1 声道音频信号中的第 i 个声道的音频信号，采用第 i 个虚拟音箱对应的 HRTF 数据中的右声道 HRTF 系数进行第二卷积，得到第二卷积后的第 i 个声道的音频信号；

右声道合成单元，用于将第二卷积后的各个声道的音频信号进行叠加，得到立体声音频信号中的右声道信号。

图 12，示出了本申请一个示例性实施例提供的终端 1200 的结构框图。该终端 1200 可以是：智能手机、平板电脑、MP3 播放器（Moving Picture Experts Group Audio Layer III，动态影像专家压缩标准音频层面 3）、MP4（Moving Picture Experts Group Audio Layer IV，动态影像专家压缩标准音频层面 4）播放器、笔记本电脑或台式电脑。终端 1200 还可能被称为用户设备、便携式终端、膝上型终端、台式终端等其他名称。

通常，终端 1200 包括有：处理器 1201 和存储器 1202。

处理器 1201 可以包括一个或多个处理核心，比如 4 核心处理器、8 核心处理器等。处理器 1201 可以采用 DSP（Digital Signal Processing，数字信号处理）、FPGA（Field-Programmable Gate Array，现场可编程门阵列）、PLA（Programmable Logic Array，可编程逻辑阵列）中的至少一种硬件形式来实现。处理器 1201 也可以包括主处理器和协处理器，主处理器是用于对在唤醒状态下的数据进行处理的处理单元，也称 CPU（Central Processing Unit，中央处理器）；协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中，处理器 1201 可以在集成有 GPU（Graphics Processing Unit，图像处理器），GPU 用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中，处理器 1201 还可以包括 AI（Artificial Intelligence，人工智能）处理器，该 AI 处理器用于处理有关机器学习的计算操作。

存储器 1202 可以包括一个或多个计算机可读存储介质，该计算机可读存储介质可以是非暂态的。存储器 1202 还可包括高速随机存取存储器，以及非易失性存储器，比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中，存储器 1202 中的非暂态的计算机可读存储介质用于存储至少一个指令，该至少一个指令用于被处理器 1201 所执行以实现本申请中各个方法实施例提

供的音频信号的处理方法。

在一些实施例中，终端 1200 还可选包括有：外围设备接口 1203 和至少一个外围设备。处理器 1201、存储器 1202 和外围设备接口 1203 之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口 1203 相连。具体地，外围设备包括：射频电路 1204、触摸显示屏 1205、摄像头 1206、音频电路 1207、定位组件 1208 和电源 1209 中的至少一种。

外围设备接口 1203 可被用于将 I/O (Input /Output, 输入/输出) 相关的至少一个外围设备连接到处理器 1201 和存储器 1202。在一些实施例中，处理器 1201、存储器 1202 和外围设备接口 1203 被集成在同一芯片或电路板上；在一些其他实施例中，处理器 1201、存储器 1202 和外围设备接口 1203 中的任意一个或两个可以在单独的芯片或电路板上实现，本实施例对此不加以限定。

射频电路 1204 用于接收和发射 RF (Radio Frequency, 射频) 信号，也称电磁信号。射频电路 1204 通过电磁信号与通信网络以及其他通信设备进行通信。射频电路 1204 将电信号转换为电磁信号进行发送，或者，将接收到的电磁信号转换为电信号。可选地，射频电路 1204 包括：天线系统、RF 收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路 1204 可以通过至少一种无线通信协议来与其它终端进行通信。该无线通信协议包括但不限于：万维网、城域网、内联网、各代移动通信网络 (2G、3G、4G 及 5G)、无线局域网和/或 WiFi(Wireless Fidelity, 无线保真) 网络。在一些实施例中，射频电路 1204 还可以包括 NFC (Near Field Communication, 近距离无线通信) 有关的电路，本申请对此不加以限定。

显示屏 1205 用于显示 UI (User Interface, 用户界面)。该 UI 可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏 1205 是触摸显示屏时，显示屏 1205 还具有采集在显示屏 1205 的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器 1201 进行处理。此时，显示屏 1205 还可以用于提供虚拟按钮和/或虚拟键盘，也称软按钮和/或软键盘。在一些实施例中，显示屏 1205 可以为一个，设置终端 1200 的前面板；在另一些实施例中，显示屏 1205 可以为至少两个，分别设置在终端 1200 的不同表面或呈折叠设计；在再一些实施例中，显示屏 1205 可以是柔性显示屏，设置在终端 1200 的弯曲表面上或折叠面上。甚至，显示屏 1205 还可以设置成非矩形的不规则图形，也即异形屏。显示屏 1205 可以采用 LCD(Liquid Crystal Display, 液晶显

示屏)、OLED(Organic Light-Emitting Diode,有机发光二极管)等材质制备。

摄像头组件 1206 用于采集图像或视频。可选地,摄像头组件 1206 包括前置摄像头和后置摄像头。通常,前置摄像头设置在终端的前面板,后置摄像头设置在终端的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及 VR (Virtual Reality, 虚拟现实) 拍摄功能或者其它融合拍摄功能。在一些实施例中,摄像头组件 1206 还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

音频电路 1207 可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并将声波转换为电信号输入至处理器 1201 进行处理,或者输入至射频电路 1204 以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在终端 1200 的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器 1201 或射频电路 1204 的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路 1207 还可以包括耳机插孔。

定位组件 1208 用于定位终端 1200 的当前地理位置,以实现导航或 LBS (Location Based Service, 基于位置的服务)。定位组件 1208 可以是基于美国的 GPS (Global Positioning System, 全球定位系统)、中国的北斗系统或俄罗斯的伽利略系统的定位组件。

电源 1209 用于为终端 1200 中的各个组件进行供电。电源 1209 可以是交流电、直流电、一次性电池或可充电电池。当电源 1209 包括可充电电池时,该可充电电池可以是有线充电电池或无线充电电池。有线充电电池是通过有线线路充电的电池,无线充电电池是通过无线线圈充电的电池。该可充电电池还可以用于支持快充技术。

在一些实施例中,终端 1200 还包括有一个或多个传感器 1210。该一个或多个传感器 1210 包括但不限于:加速度传感器 1211、陀螺仪传感器 1212、压力传感器 1213、指纹传感器 1214、光学传感器 1215 以及接近传感器 1216。

加速度传感器 1211 可以检测以终端 1200 建立的坐标系的三个坐标轴上的加速度大小。比如，加速度传感器 1211 可以用于检测重力加速度在三个坐标轴上的分量。处理器 1201 可以根据加速度传感器 1211 采集的重力加速度信号，控制触摸显示屏 1205 以横向视图或纵向视图进行用户界面的显示。加速度传感器 1211 还可以用于游戏或者用户的运动数据的采集。

陀螺仪传感器 1212 可以检测终端 1200 的机体方向及转动角度，陀螺仪传感器 1212 可以与加速度传感器 1211 协同采集用户对终端 1200 的 3D 动作。处理器 1201 根据陀螺仪传感器 1212 采集的数据，可以实现如下功能：动作感应（比如根据用户的倾斜操作来改变 UI）、拍摄时的图像稳定、游戏控制以及惯性导航。

压力传感器 1213 可以设置在终端 1200 的侧边框和/或触摸显示屏 1205 的下层。当压力传感器 1213 设置在终端 1200 的侧边框时，可以检测用户对终端 1200 的握持信号，由处理器 1201 根据压力传感器 1213 采集的握持信号进行左右手识别或快捷操作。当压力传感器 1213 设置在触摸显示屏 1205 的下层时，由处理器 1201 根据用户对触摸显示屏 1205 的压力操作，实现对 UI 界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

指纹传感器 1214 用于采集用户的指纹，由处理器 1201 根据指纹传感器 1214 采集到的指纹识别用户的身份，或者，由指纹传感器 1214 根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时，由处理器 1201 授权该用户执行相关的敏感操作，该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器 1214 可以被设置终端 1200 的正面、背面或侧面。当终端 1200 上设置有物理按键或厂商 Logo 时，指纹传感器 1214 可以与物理按键或厂商 Logo 集成在一起。

光学传感器 1215 用于采集环境光强度。在一个实施例中，处理器 1201 可以根据光学传感器 1215 采集的环境光强度，控制触摸显示屏 1205 的显示亮度。具体地，当环境光强度较高时，调高触摸显示屏 1205 的显示亮度；当环境光强度较低时，调低触摸显示屏 1205 的显示亮度。在另一个实施例中，处理器 1201 还可以根据光学传感器 1215 采集的环境光强度，动态调整摄像头组件 1206 的拍摄参数。

接近传感器 1216，也称距离传感器，通常设置在终端 1200 的前面板。接

近传感器 1216 用于采集用户与终端 1200 的正面之间的距离。在一个实施例中，当接近传感器 1216 检测到用户与终端 1200 的正面之间的距离逐渐变小时，由处理器 1201 控制触摸显示屏 1205 从亮屏状态切换为息屏状态；当接近传感器 1216 检测到用户与终端 1200 的正面之间的距离逐渐变大时，由处理器 1201 控制触摸显示屏 1205 从息屏状态切换为亮屏状态。

本领域技术人员可以理解，图 12 中示出的结构并不构成对终端 1200 的限定，可以包括比图示更多或更少的组件，或者组合某些组件，或者采用不同的组件布置。

本申请还提供一种计算机可读存储介质，所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述方法实施例提供的音频信号的处理方法。

可选地，本申请还提供了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方面所述的音频信号的处理方法。

应当理解的是，在本文中提及的“多个”是指两个或两个以上。“和/或”，描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。字符“/”一般表示前后关联对象是一种“或”的关系。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

以上所述仅为本申请的较佳实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权 利 要 求 书

1、一种音频信号的处理方法，其特征在于，所述方法由终端执行，所述方法包括：

获取第一立体声音频信号；

将所述第一立体声音频信号拆分为 5.1 声道音频信号；

对所述 5.1 声道音频信号按照三维环绕的 5.1 虚拟音箱的音箱参数进行信号处理，得到处理后的 5.1 声道音频信号；

将所述处理后的 5.1 声道音频信号，合成为第二立体声音频信号。

2、根据权利要求 1 所述的方法，其特征在于，所述将所述第一立体声音频信号拆分为 5.1 声道音频信号，包括：

对所述第一立体声音频信号输入高通滤波器进行滤波，得到第一高频信号；

根据所述第一高频信号，计算得到左声道高频信号、中央声道高频信号和右声道高频信号；

根据所述左声道高频信号、中央声道高频信号和右声道高频信号，计算得到所述 5.1 声道音频信号中的前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号。

3、根据权利要求 2 所述的方法，其特征在于，所述根据所述第一高频信号，计算得到左声道高频信号、中央声道高频信号和右声道高频信号，包括：

对所述第一高频信号进行快速傅里叶变换，得到高频实数信号和高频虚数信号；

根据所述高频实数信号和所述高频虚数信号计算向量投影；

对所述高频实数信号中的左声道高频实数信号和所述向量投影的乘积进行快速傅里叶逆变换，得到所述中央声道高频信号；

将所述第一高频信号中的左声道高频信号和所述中央声道信号的差，作为所述左声道高频信号；

将所述第一高频信号中的右声道高频信号和所述中央声道信号的差，作为所述右声道高频信号。

4、根据权利要求2所述的方法，其特征在于，所述根据所述左声道高频信号、中央声道高频信号和右声道高频信号，计算得到所述5.1声道音频信号中的前置左声道信号、前置右声道信号、前置中央声道信号、低频声道信号、后置左声道信号和后置右声道信号，包括：

提取所述左声道高频信号中的第一后方/混响信号数据、所述中央声道高频信号中的第二后方/混响信号数据、所述右声道高频信号中的第三后方/混响信号数据；

将所述左声道高频信号和所述第一后方/混响信号数据的差，确定为所述前置左声道信号；

将所述第一后方/混响信号数据和所述第二后方/混响信号数据的和，确定为所述后置左声道信号；

将所述右声道高频信号和所述第三后方/混响信号数据的差，确定为所述前置右声道信号；

将所述第三后方/混响信号数据和所述第二后方/混响信号数据的和，确定为所述后置右声道信号；

将所述中央声道高频信号和所述第二后方/混响信号数据的差，确定为所述前置中央声道信号。

5、根据权利要求4所述的方法，其特征在于，所述提取所述左声道高频信号中的第一后方/混响信号数据、所述中央声道高频信号中的第二后方/混响信号数据、所述右声道高频信号中的第三后方/混响信号数据，包括：

对于所述左声道高频信号、所述中央声道高频信号和所述右声道高频信号中的任意一个声道高频信号，根据所述声道高频信号中的采样点得到至少一个移动窗，每个移动窗包括 n 个采样点，相邻的两个移动窗存在 $n/2$ 个采样点是重叠的， $n \geq 1$ ；

计算所述移动窗中的低相关信号以及所述低相关信号的起始时间点，所述低相关信号包括幅度谱的第一衰减包络序列和相位谱的第二衰减包络序列不相等的信号；

确定符合后方/混响特征的目标低相关信号；

计算所述目标低相关信号的结束时间点；

根据所述起始时间点和所述结束时间点提取所述目标低相关信号，作为所

述声道高频信号中的后方/混响信号数据。

6、根据权利要求5所述的方法，其特征在于，所述计算所述移动窗中的低相关信号以及所述低相关信号的起始时间点，包括：

对第*i*个移动窗中的采样点信号进行快速傅里叶变换，得到快速傅里叶变换后的采样点信号， $i \geq 1$ ；

计算所述快速傅里叶变换后的采样点信号的幅度谱和相位谱；

根据所述快速傅里叶变换后的采样点信号的幅度谱，计算所述第*i*个移动窗中的*m*条频率线的第一衰减包络序列， $m \geq 1$ ；

根据所述快速傅里叶变换后的采样点信号的相位谱，计算所述第*i*个移动窗中的*m*条频率线的第二衰减包络序列；

当所述*m*条频率线中的第*j*条频率线的所述衰减包络序列和所述第二衰减包络序列不同时，确定所述第*j*条频率线为所述低相关信号， $1 \leq j \leq m$ ；

根据所述第*i*个移动窗的窗口号和所述第*j*条频率线的频率线号，确定所述低相关信号的起始时间点。

7、根据权利要求1所述的方法，其特征在于，所述5.1声道音频信号包括低频声道信号；

所述将所述第一立体声音频信号拆分为5.1声道音频信号，包括：

对所述第一立体声音频信号输入低通滤波器进行滤波，得到第一低频信号；

所述对所述5.1声道音频信号按照三维环绕的5.1虚拟音箱的音箱参数进行信号处理，得到处理后的5.1声道音频信号，包括：

对所述第一低频信号和所述5.1虚拟音箱中的低频声道音箱的音量参数进行标量相乘，得到第二低频信号；

将所述第二低频信号进行单声道转换，得到处理后的低频声道信号。

8、一种音频信号的处理装置，其特征在于，所述装置应用于终端中，所述装置包括：

获取模块，用于获取第一立体声音频信号；

处理模块，用于将所述第一立体声音频信号拆分为5.1声道音频信号；对所述5.1声道音频信号按照三维环绕的5.1虚拟音箱的音箱参数进行信号处理，得

到处理后的 5.1 声道音频信号；

合成模块，用于将所述处理后的 5.1 声道音频信号，合成为第二立体声音频信号。

9、一种终端，其特征在于，所述终端包括处理器和存储器，所述存储器中存储有至少一条指令，所述指令由所述处理器加载并执行以实现如权利要求 1 至 7 任一所述的音频信号处理方法。

10、一种计算机可读存储介质，其特征在于，所述存储介质中存储有至少一条指令，所述指令由处理器加载并执行以实现如权利要求 1 至 7 任一所述的音频信号处理方法。

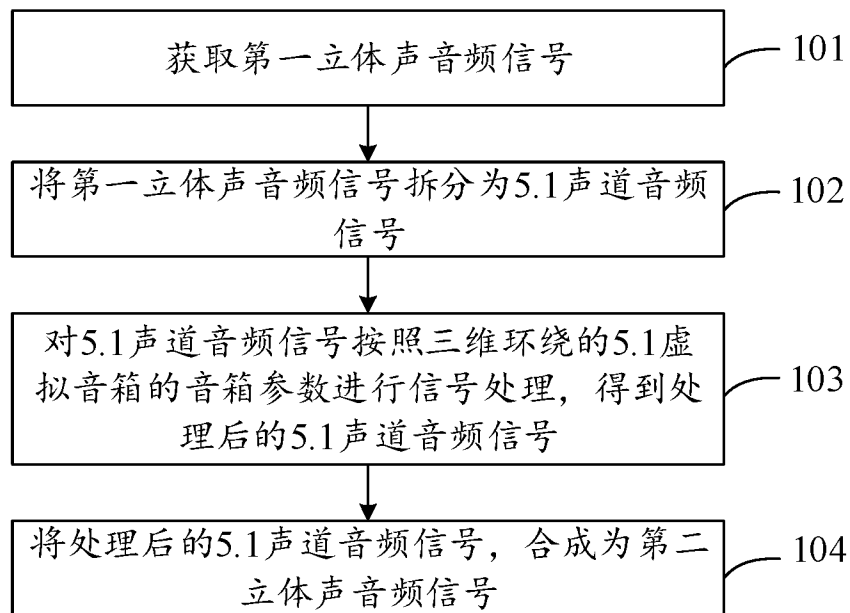


图 1

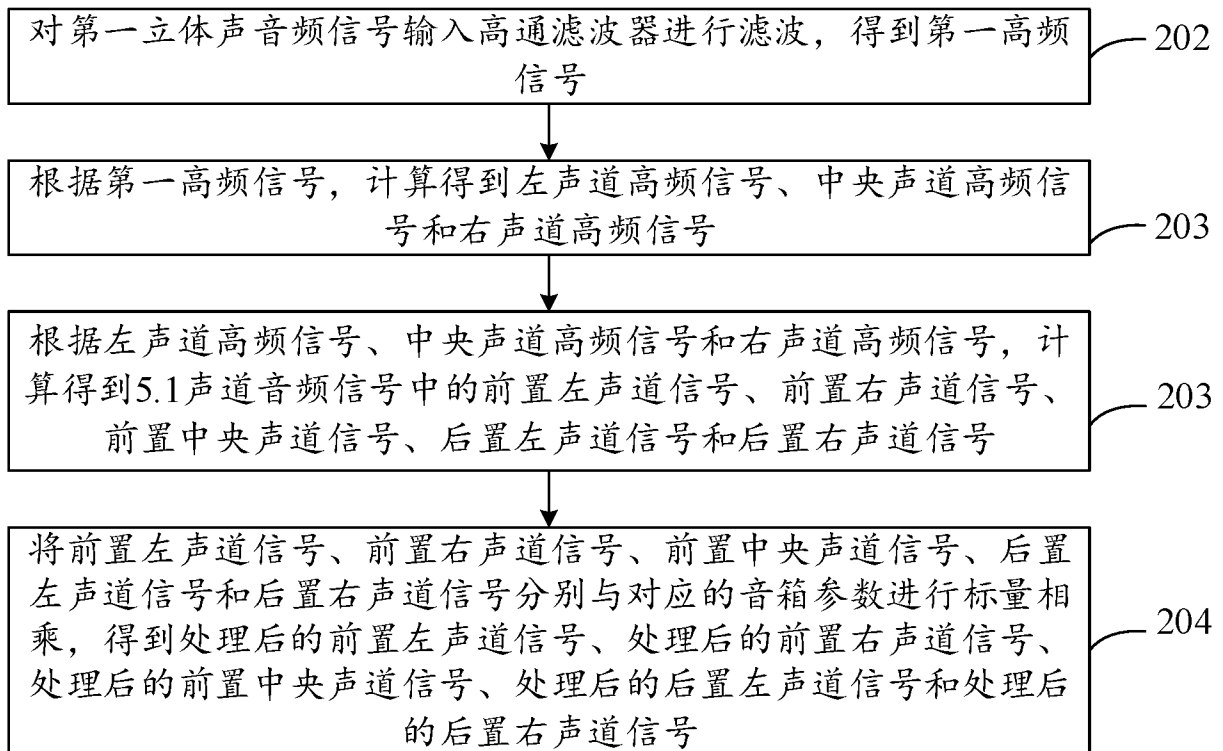


图 2

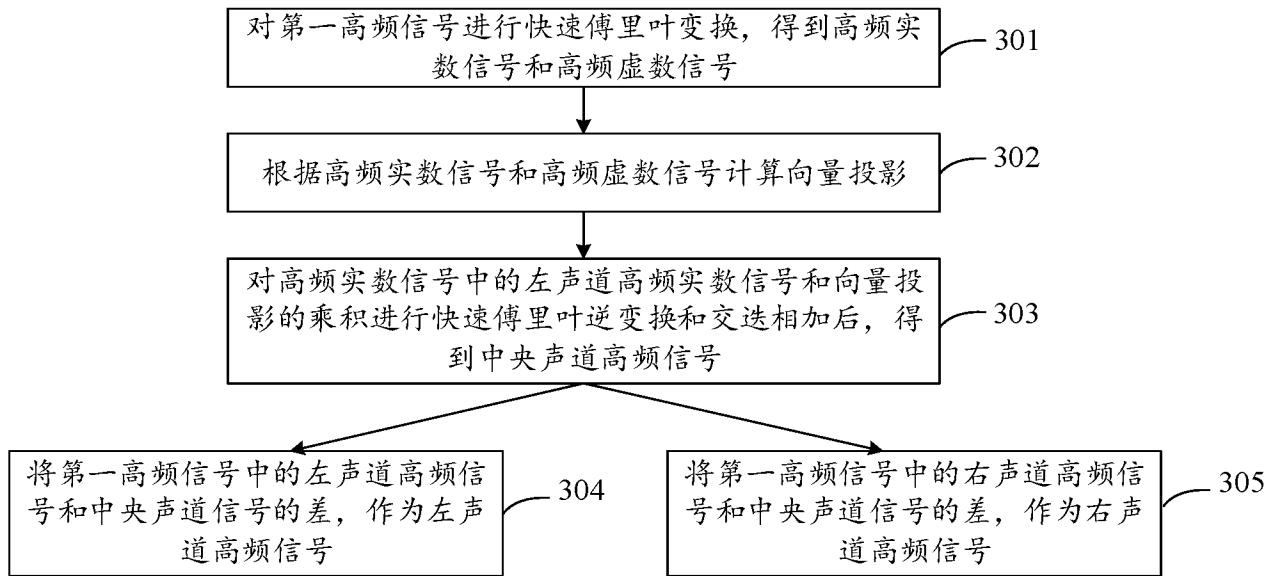


图 3

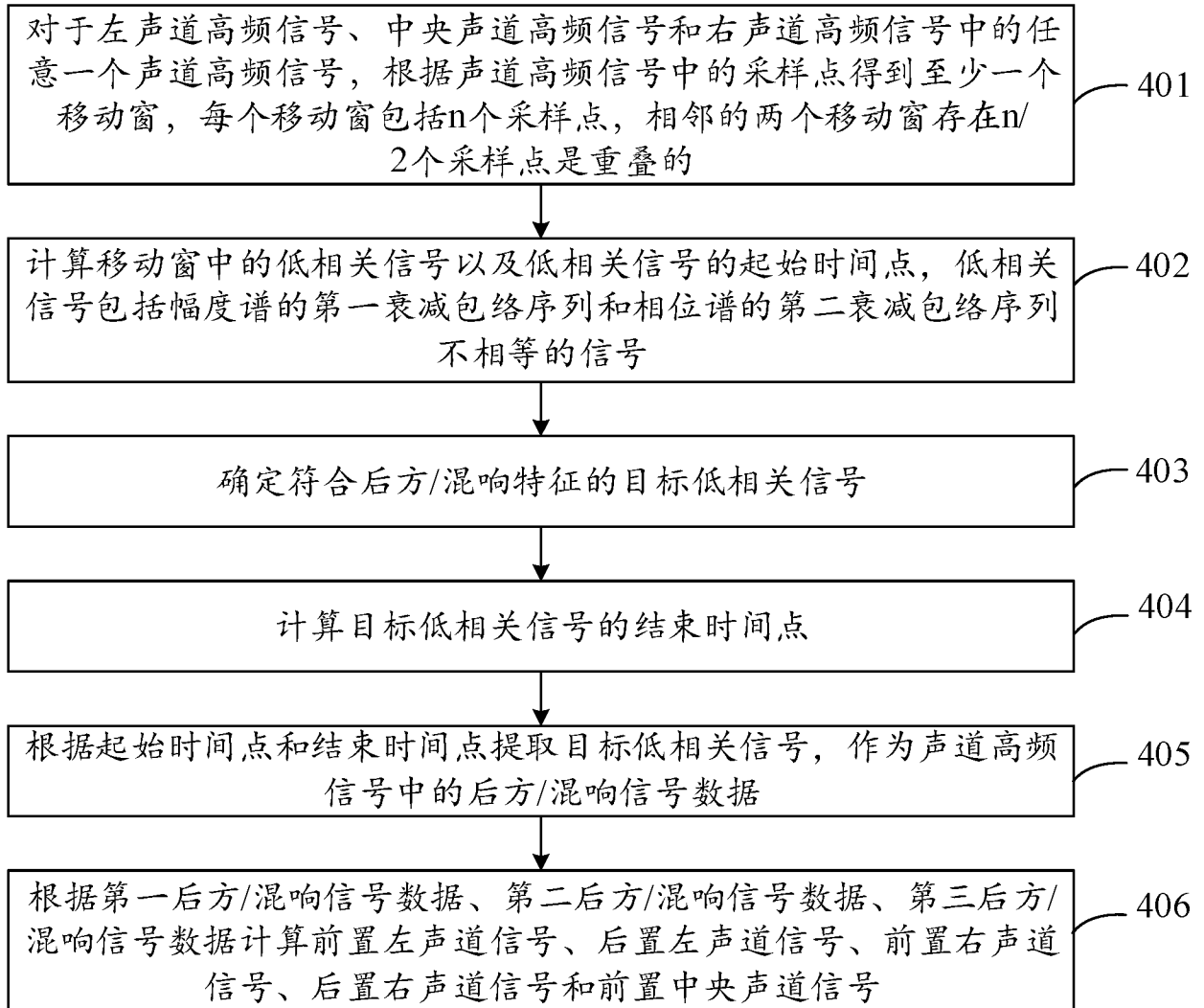


图 4

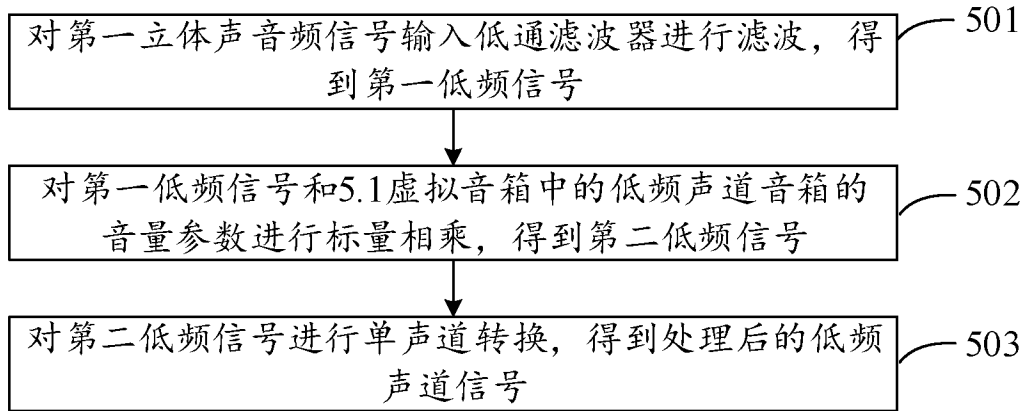


图 5

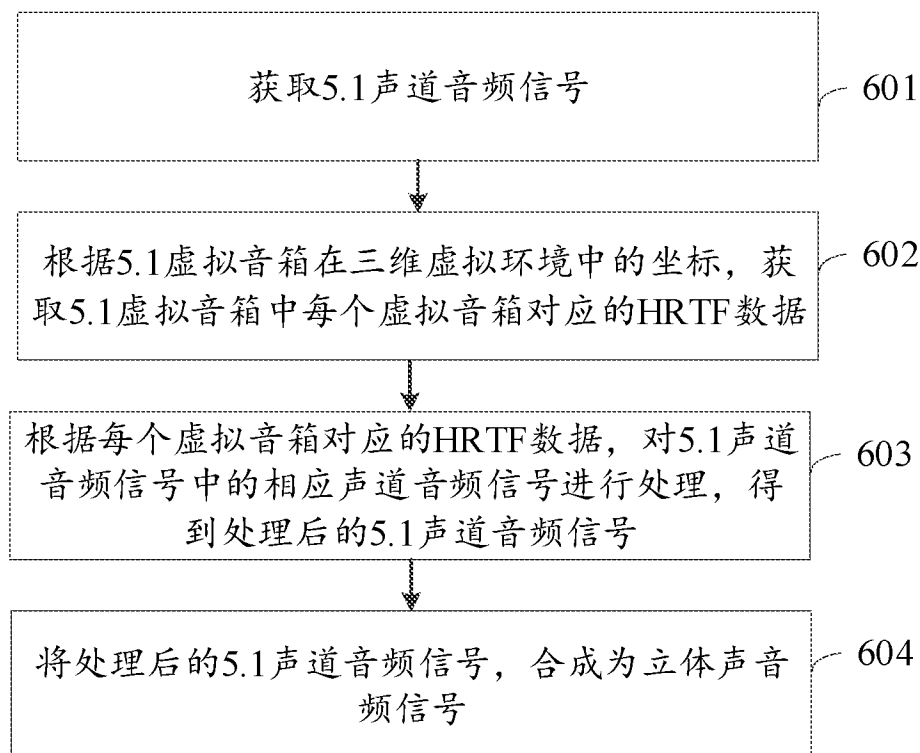


图 6

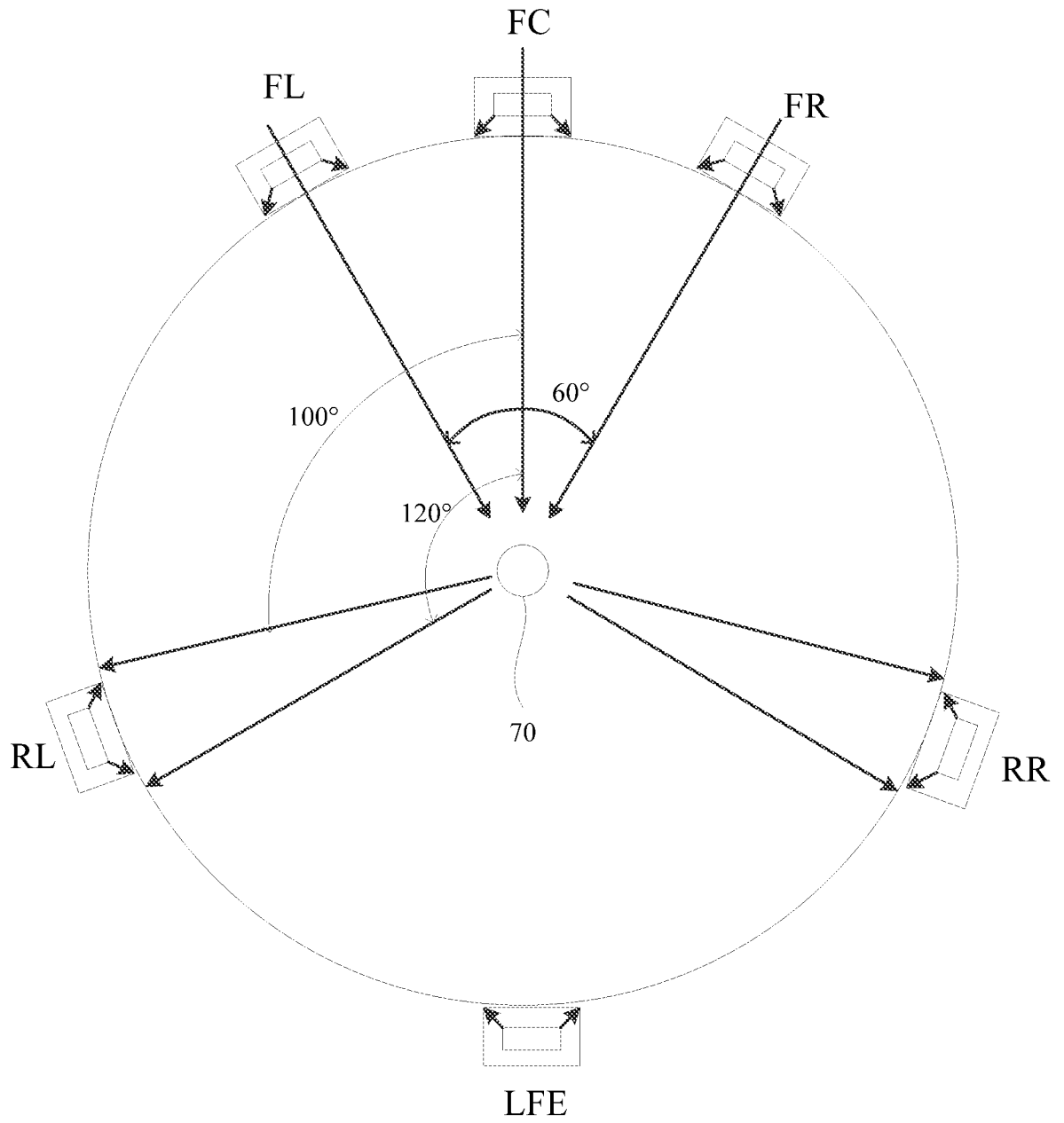


图 7

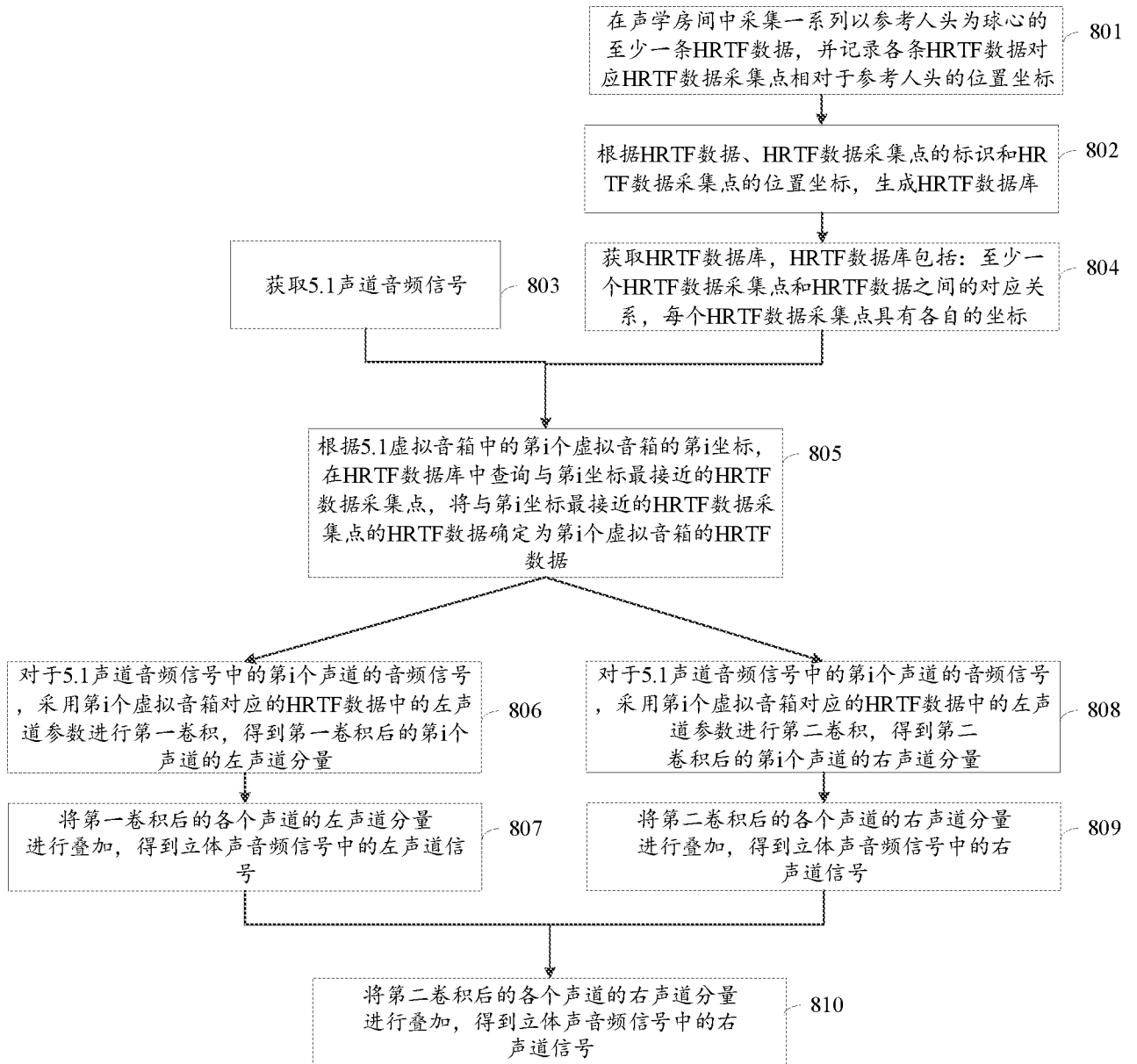


图 8

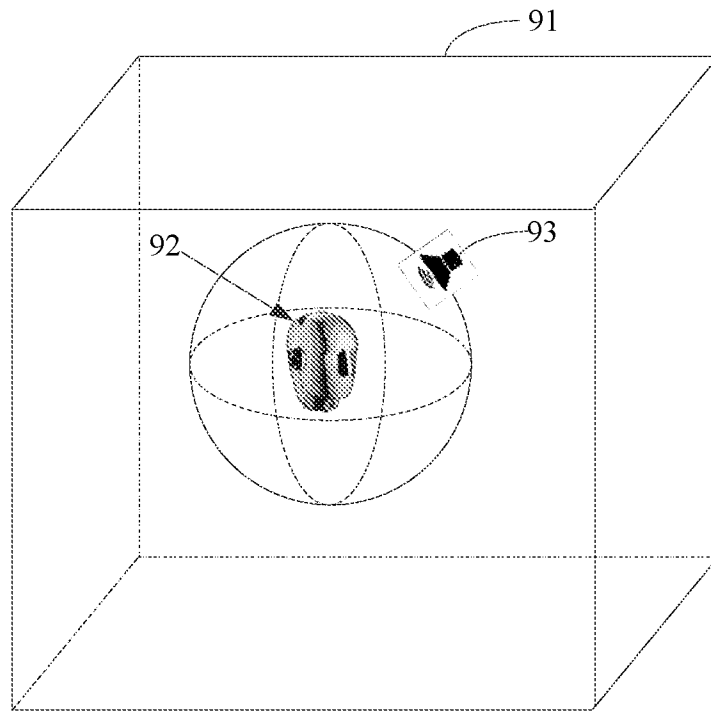


图 9

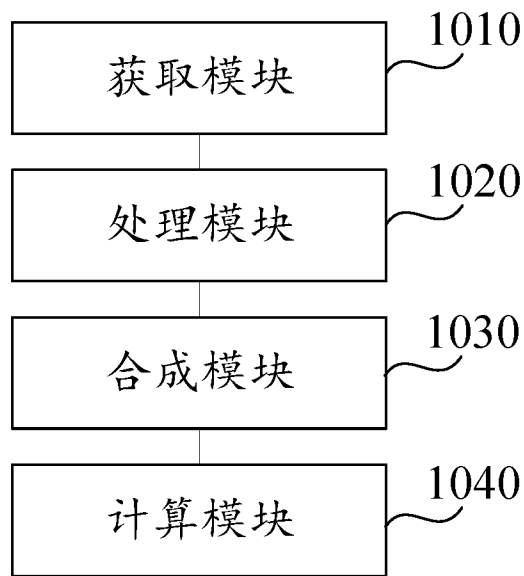


图 10

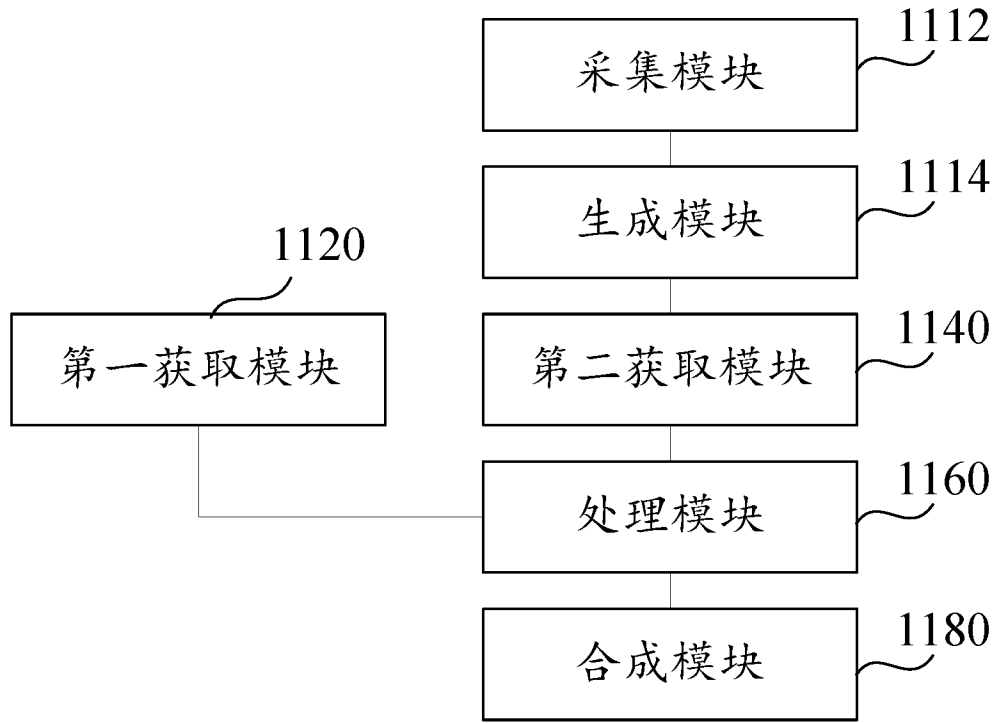


图 11

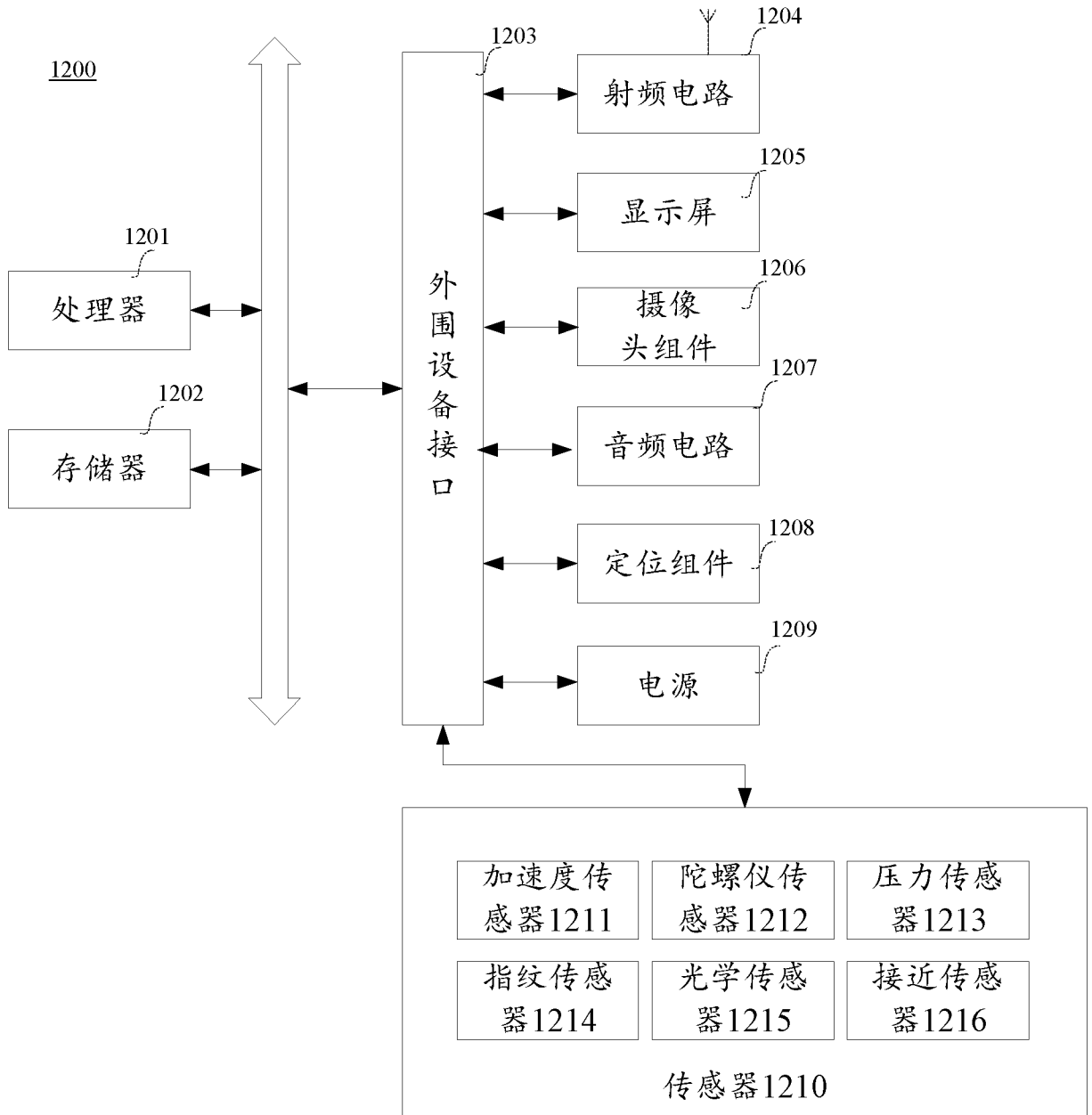


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/118764

A. CLASSIFICATION OF SUBJECT MATTER		
H04S 5/00(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)		
H04S; H04R		
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)		
CNABS; CNTXT; CNKI; VEN; USTXT; WOTXT; EPTXT: 5.1声道, 环绕声, 双声道, 两声道, 二声道, 左声道, 右声道, 高频, 低频, 滤波, 傅里叶, 付立叶, 傅立叶, 付里叶, 虚拟, 5.1, channel?, surround, two, double, dual, left, right, binaural, high frequency, high-frequency, low frequency, low-frequency, filter+, fourier, FFT, DFT, virtual		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 108156575 A (GUANGZHOU KUGOU TECHNOLOGY CO., LTD.) 12 June 2018 (2018-06-12) claims 1-10	1-10
Y	WANG, Chao. "多通道立体声虚拟重建技术研究 (The Study of Virtual Multichannel Surround Sound Reproduction Technology)" 中国优秀硕士学位论文全文数据库 (Chinese Master's Theses Full-Text Database), 31 December 2011 (2011-12-31), pp. 55-57	1, 7-10
A	WANG, Chao. "多通道立体声虚拟重建技术研究 (The Study of Virtual Multichannel Surround Sound Reproduction Technology)" 中国优秀硕士学位论文全文数据库 (Chinese Master's Theses Full-Text Database), 31 December 2011 (2011-12-31), pp. 55-57	2-6
Y	CN 101902679 A (BYD COMPANY LIMITED) 01 December 2010 (2010-12-01) description, paragraphs [0007]-[0029], and figure 1	1, 7-10
A	CN 101902679 A (BYD COMPANY LIMITED) 01 December 2010 (2010-12-01) description, paragraphs [0007]-[0029], and figure 1	2-6
<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
23 January 2019		20 February 2019
Name and mailing address of the ISA/CN		Authorized officer
National Intellectual Property Administration, PRC (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/118764**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 7243073 B2 (VIA TECHNOLOGIES, INC.) 10 July 2007 (2007-07-10) description, column 2, line 34 to column 3, line 11, and figure 4	1, 7-10
A	US 7243073 B2 (VIA TECHNOLOGIES, INC.) 10 July 2007 (2007-07-10) description, column 2, line 34 to column 3, line 11, and figure 4	2-6
A	CN 104581602 A (CHANGZHOU HEARING WORKSHOP INTELLIGENT TECHNOLOGY CO., LTD.) 29 April 2015 (2015-04-29) entire document	1-10
A	CN 107040862 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 11 August 2017 (2017-08-11) entire document	1-10
A	YI, Zhao et al. "Multi-channel Audio Signal Retrieval Based on Multi-factor Data Mining with Tensor Decomposition" <i>Proceedings of the 19th International Conference on Digital Signal Processing</i> , 23 August 2014 (2014-08-23), entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/118764

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108156575	A	12 June 2018	None			
CN	101902679	A	01 December 2010	CN	101902679	B	24 July 2013
US	7243073	B2	10 July 2007	US	2004037425	A1	26 February 2004
				TW	1236307	B	11 July 2005
CN	104581602	A	29 April 2015	None			
CN	107040862	A	11 August 2017	None			

国际检索报告

国际申请号

PCT/CN2018/118764

<p>A. 主题的分类</p> <p>H04S 5/00(2006.01)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>H04S; H04R</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>CNABS;CNTXT;CNKI;VEN;USTXT;WOTXT;EPTXT;IEEE: 5.1声道, 环绕声, 双声道, 两声道, 二声道, 左声道, 右声道, 高频, 低频, 滤波, 傅里叶, 付立叶, 傅立叶, 付里叶, 虚拟, 5.1, channel?, surround, two, double, dual, left, right, binaural, higt frequency, high-frequency, low frequency, low-frequency, filter+, fourier, FFT, DFT, virtual</p>																							
<p>G. 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>PX</td> <td>CN 108156575 A (广州酷狗计算机科技有限公司) 2018年 6月 12日 (2018 - 06 - 12) 权利要求1-10</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页</td> <td>1、7-10</td> </tr> <tr> <td>A</td> <td>王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页</td> <td>2-6</td> </tr> <tr> <td>Y</td> <td>CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1</td> <td>1、7-10</td> </tr> <tr> <td>A</td> <td>CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1</td> <td>2-6</td> </tr> <tr> <td>Y</td> <td>US 7243073 B2 (VIA TECH INC) 2007年 7月 10日 (2007 - 07 - 10) 说明书第2栏第34行-第3栏第11行, 图4</td> <td>1、7-10</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	PX	CN 108156575 A (广州酷狗计算机科技有限公司) 2018年 6月 12日 (2018 - 06 - 12) 权利要求1-10	1-10	Y	王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页	1、7-10	A	王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页	2-6	Y	CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1	1、7-10	A	CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1	2-6	Y	US 7243073 B2 (VIA TECH INC) 2007年 7月 10日 (2007 - 07 - 10) 说明书第2栏第34行-第3栏第11行, 图4	1、7-10
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 108156575 A (广州酷狗计算机科技有限公司) 2018年 6月 12日 (2018 - 06 - 12) 权利要求1-10	1-10																					
Y	王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页	1、7-10																					
A	王超. "多通道立体声虚拟重建技术研究" 《中国优秀硕士学位论文全文数据库》, 2011年 12月 31日 (2011 - 12 - 31), 第55-57页	2-6																					
Y	CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1	1、7-10																					
A	CN 101902679 A (比亚迪股份有限公司) 2010年 12月 1日 (2010 - 12 - 01) 说明书第[0007]-[0029]段, 图1	2-6																					
Y	US 7243073 B2 (VIA TECH INC) 2007年 7月 10日 (2007 - 07 - 10) 说明书第2栏第34行-第3栏第11行, 图4	1、7-10																					
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																							
<p>* 引用文件的具体类型:</p> <p>"A" 认为不特别相关的表示了现有技术一般状态的文件</p> <p>"E" 在国际申请日的当天或之后公布的在先申请或专利</p> <p>"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)</p> <p>"O" 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&" 同族专利的文件</p>																							
<p>国际检索实际完成的日期</p> <p>2019年 1月 23日</p>		<p>国际检索报告邮寄日期</p> <p>2019年 2月 20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</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>田珊</p> <p>电话号码 86-(010)-62089397</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 7243073 B2 (VIA TECH INC) 2007年 7月 10日 (2007 - 07 - 10) 说明书第2栏第34行-第3栏第11行, 图4	2-6
A	CN 104581602 A (常州听觉工坊智能科技有限公司) 2015年 4月 29日 (2015 - 04 - 29) 全文	1-10
A	CN 107040862 A (腾讯科技深圳有限公司) 2017年 8月 11日 (2017 - 08 - 11) 全文	1-10
A	Yi Zhao 等. "Multi-channel audio signal retrieval based on multi-factor data mining with tensor decomposition" Proceedings of the 19th International Conference on Digital Signal Processing, 2014年 8月 23日 (2014 - 08 - 23), 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/118764

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	108156575	A	2018年 6月 12日	无			
CN	101902679	A	2010年 12月 1日	CN	101902679	B	2013年 7月 24日
US	7243073	B2	2007年 7月 10日	US	2004037425	A1	2004年 2月 26日
				TW	I236307	B	2005年 7月 11日
CN	104581602	A	2015年 4月 29日	无			
CN	107040862	A	2017年 8月 11日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)