

(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
2017年11月16日 (16.11.2017)

(10) 国际公布号
WO 2017/193551 A1

- (51) 国际专利分类号:
G10L 19/008 (2013.01)
- (21) 国际申请号: PCT/CN2016/103596
- (22) 国际申请日: 2016年10月27日 (27.10.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610303992.4 2016年5月10日 (10.05.2016) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 张兴涛 (ZHANG, Xingtao); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 刘泽新 (LIU, Zexin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 苗磊 (MIAO, Lei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京龙双利达知识产权代理有限公司 (LONGSUN LEAD IP LTD.); 中国北京市海淀区北清路68号院3号楼101, Beijing 100094 (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, JP, KE, KG, 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,

(54) Title: METHOD FOR ENCODING MULTI-CHANNEL SIGNAL AND ENCODER

(54) 发明名称: 多声道信号的编码方法和编码器

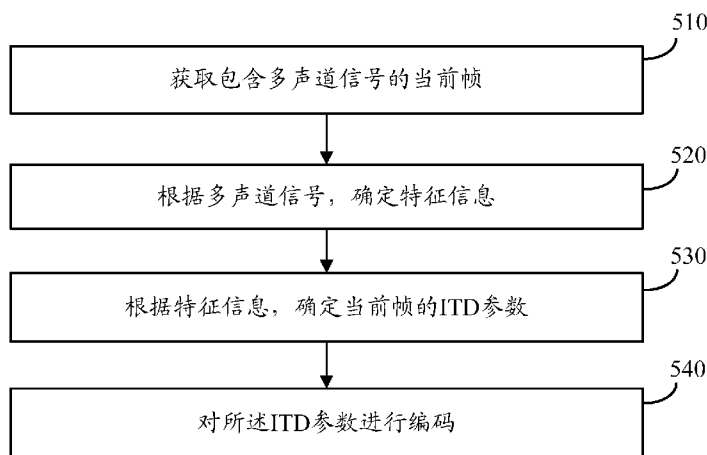


图 5

- 510 Acquire the current frame containing a multi-channel signal
520 Determine characteristic information according to the multi-channel signal
530 Determine an ITD parameter of the current frame according to the characteristic information
540 Encode the ITD parameter

(57) Abstract: A method for encoding a multi-channel signal and an encoder, the method comprising: acquiring the current frame containing a multi-channel signal (510); determining characteristic information according to the multi-channel signal (520), the characteristic information comprising at least one of the frame type and signal type of the current frame, the frame type comprising a speech frame and/or a non-speech frame, the signal type comprising unvoiced sound and/or voiced sound; determining an ITD parameter of the current frame according to the characteristic information (530); and encoding the ITD parameter (540). The present invention can improve the accuracy of ITD parameter extraction.



WO 2017/193551 A1

SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种多声道信号的编码方法和编码器, 该方法包括: 获取包含多声道信号的当前帧 (510); 根据多声道信号, 确定特征信息 (520), 其中, 特征信息包括当前帧的帧类型和信号类型中的至少一种, 帧类型包括语音帧和/或非语音帧, 信号类型包括清音和/或浊音; 根据特征信息, 确定当前帧的ITD参数 (530); 对ITD参数进行编码 (540)。能够提高ITD参数提取的准确性。

多声道信号的编码方法和编码器

本申请要求于 2016 年 05 月 10 日提交中国专利局、申请号为 201610303992.4、发明名称为“多声道信号的编码方法和编码器”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

本发明实施例涉及音频编解码领域，并且更为具体地，涉及一种多声道信号的编码方法和编码器。

10

背景技术

随着生活质量的提高，人们对高质量音频的需求不断增大。相对于单声道音频，立体声音频具有各声源的方位感和分布感，能够提高声音的清晰度、可懂度及临场感，因而备受人们青睐。

15

立体声处理技术主要有和差 (Mid/Sid, MS) 编码、强度立体声 (Intensity Stereo, IS) 编码以及参数立体声 (Parametric Stereo, PS) 编码。

20

MS 编码基于声道间相关性将两路信号作和、差变换，各声道能量主要集中在和声道，使声道间冗余得以去除。在 MS 编码技术中，码率的节省依赖于输入信号的相关性，当左右声道信号的相关性差时，需分别传输左声道信号和右声道信号。IS 编码基于人耳听觉系统对声道的高频成分（例如，大于 2kHz 的成分）的相位差异精细结果不敏感的特性，将左右两路信号的高频分量进行简化处理。但该 IS 编码技术仅对高频成分有效，如将 IS 编码处理扩展到低频，将会引起严重的人为噪声。PS 编码基于双耳听觉模型，在编码端将立体声转换成单声道信号和少量描述空间声场的空间参数（或空间感知参数），如图 1 所示（图 1 中的 x_L 为左声道时域信号， x_R 为右声道时域信号）。解码端得到单声道信号后进一步结合空间参数恢复立体声，如图 2 所示。相对于 MS 编码，PS 编码压缩比高，在保持较好音质的前提下，能够获得更高的编码增益，且可以工作在全音频带宽中，能够很好地还原立体

25

声的空间感知效果。

PS 编码中，空间参数包括声道间相关性（Inter-channel Coherent, IC）、声道间电平差（Inter-channel Level Difference, ILD）、声道间时间差（Inter-channel Time Difference, ITD）以及声道间相位差（Inter-channel Phase Difference, IPD）。IC 描述了声道间的互相关或相干性，该参数决定了声场范围的感知，可以提高音频信号空间感和声响稳定性。ILD 用于分辨立体声源的水平方向角度，描述了声道间的强度差别，该参数将影响整个频谱的频率成分。ITD 和 IPD 为表示声源水平方位的空间参数，描述了声道间的时间和相位的差别，该参数主要影响 2kHz 以下的频率成分。ILD、ITD 和 IPD 能够决定人耳对声源位置的感知，可以有效确定声场位置，对立体声信号的恢复具有重要作用。

在具体的音频编码过程中，可以以帧为单位对立体声进行编码。在对当前帧进行编码时，可以基于当前帧中的多声道信号提取当前帧对应的 ITD 参数。其中，既可以基于时域信号提取当前帧的 ITD 参数，也可以基于频域信号提取当前帧的 ITD 参数。但无论基于那种方式提取当前帧的 ITD 参数，在整个编码过程中，所有帧的 ITD 参数提取方式均保持一致，这样的 ITD 参数提取方式不够灵活。

发明内容

20 本申请提供一种多声道信号的编码方法和编码器，以提高 ITD 参数提取方式的灵活性。

第一方面，提供一种多声道信号的编码方法，包括：获取包含多声道信号的当前帧；根据所述多声道信号，确定特征信息，其中，所述特征信息包括所述当前帧的帧类型和信号类型中的至少一种，所述帧类型包括语音帧和/或非语音帧，所述信号类型包括清音和/或浊音；根据所述特征信息，确定所述当前帧的 ITD 参数；对所述 ITD 参数进行编码。当前帧的 ITD 参数可以表示当前帧中的多声道信号的 ITD 参数。

本方案根据特征信息确定当前帧的 ITD 参数,而非像现有技术那样不考虑当前帧多声道信号的类型或特征,采用固定的方式提取 ITD 参数,因此,本方案能够提高 ITD 参数提取的灵活性。

5 结合第一方面,在第一方面的第一种实现方式中,所述根据所述多声道信号,确定特征信息,包括:根据所述多声道信号,确定所述当前帧的帧类型;所述根据所述特征信息,确定所述当前帧的 ITD 参数,包括:在所述当前帧为非语音帧的情况下,采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数;在所述当前帧为语音帧的情况下,采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数。

10 本方案中,根据当前帧的类型的不同,采用不同的 ITD 参数提取方式,提高了 ITD 参数提取方式的灵活性。

结合第一方面的第一种实现方式,在第一方面的第二种实现方式中,所述采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数,包括:将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

15 具体地,多声道信号可以是以帧为单位进行处理的,通常每帧为 20ms,此外,可以将帧进一步划分为子帧进行处理,例如,将 20ms 的一帧划分成 2 个子帧时,每个子帧为 10ms;将 20ms 的一帧划分成 4 个子帧时,每个子帧为 5ms。当前帧的前一帧可以指的是与当前帧紧邻的前一帧,即当前帧开始点往前推移 20ms 所包含的音频样点。当前帧的前一子帧可以指与当前帧
20 紧邻的前一帧的最后一个子帧。

本方案中,如果当前帧为非语音帧,一般承载的是背景噪声信号,而背景噪声信号的 ITD 参数一般波动较小,可以直接将当前帧的前一帧的 ITD 参数确定为当前帧的 ITD 参数,这样能够提高编码效率。

结合第一方面的第一种或第二种实现方式,在第一方面的第三种实现方
25 式中,所述采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数,包括:根据所述多声道信号,确定所述当前帧的 ITD 参数。

结合第一方面的第三种实现方式，在第一方面的第四种实现方式中，所述根据所述多声道信号，确定所述当前帧的 ITD 参数，包括：根据所述多声道信号，确定所述当前帧的初始 ITD 参数；根据所述当前帧的前一帧或前一子帧的 ITD 参数，对所述当前帧的初始 ITD 参数进行平滑处理，得到所述
5 当前帧的 ITD 参数。

通过平滑处理，能够避免噪声的影响，提高 ITD 参数提取的准确性。

结合第一方面的第四种实现方式，在第一方面的第五种实现方式中，所述根据所述当前帧的前一帧或前一子帧的 ITD 参数，对所述当前帧的初始 ITD 参数进行平滑处理，得到所述当前帧的 ITD 参数，包括：根据 $T_{sm}=w_1*$
10 $T_{sm}^{[-1]} + w_2* T_1$ ，确定所述当前帧的 ITD 参数，其中， T_1 表示所述当前帧的初始 ITD 参数， T_{sm} 表示所述当前帧的 ITD 参数， $T_{sm}^{[-1]}$ 表示所述当前帧的前一帧或前一子帧的 ITD 参数， w_1 和 w_2 表示平滑因子，其中， w_1 和 w_2 的取值范围均为 $[0,1]$ ，且 $w_1+w_2=1$ 。

结合第一方面的第三种实现方式，在第一方面的第六种实现方式中，所
15 述根据所述多声道信号，确定所述当前帧的 ITD 参数，包括：根据所述多声道信号，确定所述当前帧的 K 个子帧的初始 ITD 参数，K 为大于 1 的整数；根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数，对所述每个子帧的初始 ITD 参数进行平滑处理，得到所述每个子帧的 ITD 参数；将所述 K 个子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

20 应理解，上述每个子帧的前一子帧可以指每个子帧紧邻的前一子帧。具体地，对于 K 个子帧中的第 1 个子帧而言，该第 1 个子帧的前一子帧为与当前帧紧邻的前一帧的最后一个子帧，对于 K 个子帧的第 i ($i>2$) 个子帧而言，该第 i 子帧的前一子帧为 K 个子帧中的第 i-1 个子帧。

结合第一方面的第六种实现方式，在第一方面的第七种实现方式中，所
25 述根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数，对所述每个子帧的初始 ITD 参数进行平滑处理，得到所述每个子帧的 ITD 参数，包括：

根据 $T_{sm}(j) = w_1 * T_{sm}(j-1) + w_2 * T(j)$, 确定所述每个子帧的 ITD 参数, 其中, $T(j)$ 表示所述 K 个子帧中的第 j 个子帧的初始 ITD 参数, $T_{sm}(j)$ 表示所述第 j 个子帧的 ITD 参数, $T_{sm}(j-1)$ 表示所述 K 个子帧中的第 $j-1$ 个子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, j 为整数, 且 $1 \leq j \leq K$, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1 + w_2 = 1$ 。

结合第一方面的第五种或第七种实现方式, 在第一方面的第八种实现方式中, 所述平滑因子的取值是基于所述当前帧的信号类型确定的。

根据信号类型确定平滑因子, 能够进一步提高 ITD 参数提取的灵活性。

结合第一方面的第三种实现方式, 在第一方面的第九种实现方式中, 所述根据所述多声道信号, 确定所述当前帧的 ITD 参数, 包括: 根据所述多声道信号, 生成目标频域信号; 对所述目标频域信号进行频时变换, 得到目标时域信号; 根据所述目标时域信号, 确定所述当前帧的 ITD 参数。

在某些实现方式中, 所述目标频域信号的相位与所述多声道信号的 IPD 线性相关。在某些实现方式中, 所述目标频域信号的相位为所述多声道信号的 IPD。应理解, 频域信号可以通过复数表示, 而复数可以通过幅值和相位表示, 目标频域信号的相位可以指表示组成该目标频域信号的复数的相位。

在某些实现方式中, 目标频域信号可以为多声道频域信号的互相关信号。

在某些实现方式中, 所述根据所述目标时域信号, 确定所述当前帧的 ITD 参数, 包括: 从所述目标时域信号的 N 个采样点中选取目标采样点, 所述目标采样点为所述 N 个采样点中的采样值最大的采样点, N 表示所述目标时域信号的采样点的数目; 根据所述目标采样点对应的索引值, 确定所述当前帧的 ITD 参数, 其中, 所述索引值用于指示所述目标采样点在所述 N 个采样点中的排序。可替换地, 所述索引值用于指示所述目标采样点为所述 N 个采样点中的第几个采样点。例如, N 个采样点的索引值的取值范围可以为 $(-N/2, N/2]$, 假设目标采样点为 N 个采样点中的最后一个采样点, 则目标采样点对应的索引值为 $N/2$ 。

结合第一方面的第九种实现方式，在第一方面的第十种实现方式中，所述根据所述多声道信号，生成目标频域信号，包括：根据所述多声道信号，确定所述目标频域信号的幅值；根据所述多声道信号，确定所述当前帧多声道信号的 IPD 参数；根据所述目标频域信号的幅值，以及所述当前帧多声道信号的 IPD 参数，生成所述目标频域信号。

结合第一方面的第十种实现方式，在第一方面的第十一种实现方式中，所述根据所述多声道信号，确定所述目标频域信号的幅值，包括：根据 $A_M(k) = \frac{\min(A_1(k), A_2(k))}{\max(A_1(k), A_2(k))}$ ，确定所述目标频域信号的幅值，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $A_1(k)$ 和 $A_2(k)$ 分别表示所述多声道信号中的任意两个声道的频域信号的幅值， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

结合第一方面的第十种或第十一种实现方式，在第一方面的第十二种实现方式中，所述根据所述目标频域信号的幅值，以及所述当前帧的 IPD 参数（具体指当前帧中的多声道信号的 IPD 参数），生成所述目标频域信号，包括：根据
$$\begin{cases} X_{M_real}(k) = A_M(k) * \cos(IPD(k)) \\ X_{M_image}(k) = A_M(k) * \sin(IPD(k)) \end{cases}$$
，生成所述目标频域信号，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $X_{M_real}(k)$ 表示所述目标频域信号的实部， $X_{M_image}(k)$ 表示所述目标频域信号的虚部， $IPD(k)$ 表示所述 IPD 参数， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

结合第一方面的第九种实现方式，在第一方面的第十三种实现方式中，所述根据所述多声道信号，生成目标频域信号，包括：根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，生成所述目标频域信号，其中， $X_M(k)$ 表示所述目标频域信号， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

结合第一方面的第九种实现方式，在第一方面的第十四种实现方式中，所述根据所述多声道信号，生成目标频域信号，包括：根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，确定频域信号 $X_M(k)$ ，其中， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度；对所述频域信号 $X_M(k)$ 的幅值进行归一化处理，得到所述目标频域信号。

结合第一方面的第一种至第十四种实现方式中的任一种，在第一方面的第十五种实现方式中，所述根据所述多声道信号，确定所述当前帧的帧类型，包括：确定所述多声道信号的能量；在所述多声道信号的能量小于或等于预设的能量阈值的情况下，将所述当前帧确定为非语音帧；在所述多声道信号的能量大于所述能量阈值的情况下，将所述当前帧确定为语音帧。

结合第一方面的第一种实现方式，在第一方面的第十六种实现方式中，所述方法还包括：根据所述多声道信号，确定所述当前帧的初始 ITD 参数；所述采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数；所述采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：对所述当前帧的初始 ITD 参数进行调整，得到所述当前帧的 ITD 参数。

结合第一方面的第十六种实现方式，在第一方面的第十七种实现方式中，所述对所述当前帧的初始 ITD 参数进行调整，得到所述当前帧的 ITD 参数，包括：根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，其中， N 为大于 1 的整数。

根据当前帧的前一帧或者前 N 帧的帧类型，以及当前帧的初始 ITD 参数，确定当前帧的 ITD 参数，能够提高 ITD 参数提取的灵活性。

结合第一方面的第十七种实现方式，在第一方面的第十八种实现方式中，所述根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始

ITD 参数，确定所述当前帧的 ITD 参数，包括：在所述当前帧的前一帧或者前 N 帧的帧类型为语音帧的情况下，根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数。

本方案中，如果当前帧的前一帧或者前 N 帧为语音帧，说明当前帧为连续语音帧中的一帧，而连续语音帧之间的 ITD 参数是存在关联的，根据当前帧的前一帧的 ITD 参数和当前帧的初始 ITD 参数，确定当前帧的 ITD 参数，能够提高 ITD 参数提取的灵活性。

结合第一方面的第十八种实现方式，在第一方面的第十九种实现方式中，所述根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，包括：在所述当前帧的前一帧的 ITD 参数不为预设值，且所述当前帧的初始 ITD 参数为预设值的情况下，将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数；否则，将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数。

本方案中，当前帧为连续语音帧中的一帧，连续语音帧的 ITD 参数一般波动较小，将当前帧的前一帧的 ITD 参数确定为当前帧的 ITD 参数，能够避免 ITD 参数计算错误，提高 ITD 参数提取的准确性。

结合第一方面的第十八种实现方式，在第一方面的第二十种实现方式中，所述根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，包括：在所述当前帧的前一帧的 ITD 参数不为预设值，且所述当前帧的初始 ITD 参数为预设值的情况下，如果连续计算出的 ITD 参数为预设值的个数小于预设阈值，将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数；否则，将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数。

结合第一方面的第十六种或第十七种实现方式，在第一方面的第十八种实现方式中，预设值为 0。

第二方面，提供一种编码器，包括能够执行第一方面中的多声道信号的

编码方法的各个步骤的单元。

第三方面，提供一种编码器，包括存储器和处理器，所述存储器用于存储程序，所述处理器用于执行程序，当所述程序被执行时，所述处理器执行第一方面中的方法。

5

附图说明

为了更清楚地说明本发明实施例的技术方案，下面将对本发明实施例中所需要使用的附图作简单地介绍，显而易见地，下面所描述的附图仅仅是本发明的一些实施例，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。

10

图 1 是现有技术中的 PS 编码的流程图。

图 2 是现有技术中的 PS 解码的流程图。

图 3 是现有技术中的基于时域的 ITD 参数提取方法的示例性流程图。

图 4 是现有技术中的基于频域的 ITD 参数提取方法的示例性流程图。

15

图 5 是本发明实施例的多声道信号的编码方法的示意性流程图。

图 6 是本发明实施例的多声道信号的编码方法的示意性流程图。

图 7 是本发明实施例的多声道信号的编码方法的示意性流程图。

图 8 是本发明实施例的多声道信号的编码方法的示意性流程图。

图 9 是当前帧的 ITD 参数的提取方式的示例性流程图。

20

图 10 是当前帧的 ITD 参数的提取方式的示例性流程图。

图 11 是本发明实施例的编码器的示意性结构图。

图 12 是本发明实施例的编码器的示意性结构图。

具体实施方式

25

为了便于理解，先简单介绍多声道的 ILD、ITD、IPD 的含义。以第一个麦克拾取到的信号为第一声道信号，第二个麦克拾取到的信号为第二声道信号为例：

ILD 描述了第一声道信号和第二声道信号之间的强度差别；如果 ILD 大于 0，表示第一声道信号的能量高于第二声道信号的能量；如果 ILD 等于 0，表示第一声道信号的能量等于第二声道信号的能量；如果 ILD 小于 0，表示第一声道信号的能量小于第二声道信号的能量；

5 ITD 描述的第一声道信号和第二声道信号之间的时间差别，即声源到达第一个麦克和第二个麦克的时间差异，如果 ITD 大于 0，表示声源到达第一个麦克的时间早于声源到达第二个麦克的时间；如果 ITD 等于 0，表示声源到达同时到达第一个麦克和第二个麦克；如果 ITD 小于 0，表示声源到达第一个麦克的时间晚于声源到达第二个麦克的时间；

10 IPD 描述了第一声道信号和第二声道信号的相位差别，该参数通常和 ITD 参数结合在一起，以便解码端恢复多声道信号的相位信息。

在现有技术中，ITD 参数的提取方式主要分为基于时域的 ITD 参数提取方式和基于频域的 ITD 参数提取方式，为了便于理解，先结合图 3 和图 4，分别对这两种 ITD 参数提取方式进行介绍。

15 图 3 是基于时域的 ITD 参数提取方法的示例性流程图。图 3 的方法包括：
310、基于左右声道时域信号提取 ITD 参数。

具体而言，可以基于左右声道时域信号，采用时域互相关函数提取 ITD 参数，例如：在 $0 \leq i \leq T_{\max}$ 范围内，计算：

$$c_n(i) = \sum_{j=0}^{Length-1-i} x_R(j) \cdot x_L(j+i) \quad (1)$$

20
$$c_p(i) = \sum_{j=0}^{Length-1-i} x_L(j) \cdot x_R(j+i) \quad (2)$$

如果 $\max_{0 \leq i \leq T_{\max}} (c_n(i)) > \max_{0 \leq i \leq T_{\max}} (c_p(i))$ ，则 T_1 取 $\max(C_n(i))$ 对应的索引值的相反数；否则 T_1 取 $\max(C_p(i))$ 对应的索引值；其中， i 为计算互相关函数的索引值， T_{\max} 对应于不同采样率下 ITD 取值的最大值，Length 为帧长。

320、对 ITD 参数进行量化处理。

25 图 4 是基于频域的 ITD 参数提取方法的示例性流程图。图 4 的方法包括：

410、对左右声道时域信号进行时频变换，得到左右声道频域信号。

具体而言，时频变换可以采用离散傅里叶变换（Discrete Fourier Transformation, DFT）、修正的离散余弦变换（Modified Discrete Cosine Transform, MDCT）等技术，将时域信号变换为频域信号。

5 例如，对于输入的左右声道的时域信号，时频变换可以采用 DFT 变换，具体地，可以采用如下公式进行 DFT 变换。

$$X(k) = \sum_{n=0}^{Length-1} x(n) \cdot e^{-j\frac{2\pi nk}{L}}, 0 \leq k < L \quad (3)$$

其中， n 为时域信号的样点的索引值， k 为频域信号的频点的索引值， L 为时频变换长度。 $x(n)$ 为左声道时域信号或右声道时域信号。

10 420、基于左右声道频域信号提取 ITD 参数。

具体地，可以将频域信号的 L 个频点 (Frequency Bin) 划分为 N 个子带，对于第 b 个子带，其包含的频点为 $A_{b-1} \leq k \leq A_b - 1$ 。在搜索范围 $-T_{\max} \leq j \leq T_{\max}$ ，可以采用如下公式计算幅值：

$$mag(j) = \sum_{k=A_{b-1}}^{A_b-1} X_L(k) * X_R(k) * \exp\left(\frac{2\pi * k * j}{L}\right) \quad (4)$$

15 则第 b 个子带的 ITD 参数可以为 $T(k) = \arg \max_{-T_{\max} \leq j \leq T_{\max}} (mag(j))$ ，即公式 (4) 计算出的最大值对应的样点的索引值。

430、对 ITD 参数进行量化处理。

应理解，图 3 和图 4 描述的是 ITD 参数的大体流程，实际中，可以视情况以帧、子帧或子带为单位提取 ITD 参数，本发明实施例对此不作具体限定。

20 当以帧为单位提取 ITD 参数时，当前帧的 ITD 参数可以为一个 ITD 参数；当以子帧或子带为单位提取 ITD 参数是，当前帧的 ITD 参数可以为多个 ITD 参数，即每个子帧或每个子带对应一个 ITD 参数。

例如，在基于时域的 ITD 参数提取方式实际中，可以以帧或子帧为单位提取 ITD 参数。例如，对于一个 20ms 的当前帧，可以以当前帧（即 20ms）
25 为单位进行时频变换，提取当前帧的 ITD 参数；在当前帧被划分成 2 个子帧

的情况下，可以以子帧（即 10ms）为单位进行时频变换，提取每个子帧对应的 ITD 参数；在当前帧被划分成 4 个子帧的情况下，可以以子帧（即 5ms）为单位进行时频变换，提取每个子帧对应的 ITD 参数。

又如，在基于频域的 ITD 参数提取方式中，可以以帧或子帧为单位提取 ITD 参数。当帧或子帧进一步被划分成多个子带时，还可以以子带为单位提取 ITD 参数。

现有技术中，ITD 参数的提取方式一旦确定，多声道信号所有帧的 ITD 参数提取方式就固定下来，而无法根据实际情况灵活调整。但是多声道信号不同帧有不同的特征，例如，有些帧中包含的是语音信号，有些帧中包含的是背景噪声信号；有些帧中的语音信号是清音，有些帧中的语音信号是浊音；有些帧的能量高，有些帧的能量低。多声道信号不同类型的帧或不同类型的信号可以采用相同或不同的 ITD 参数提取方式，例如，对于背景噪声信号而言，其 ITD 参数在一定时间范围内通常不会发生大的变化，如果按帧重复地计算背景噪声信号的 ITD 参数，会浪费编码资源，降低编码效率。

为了提高 ITD 参数提取的灵活性，下面结合图 5，详细描述本发明实施例的多声道信号的编码方法。

图 5 是本发明实施例的多声道信号的编码方法的示意性流程图。图 5 的多声道信号的编码方法包括：

510、获取包含多声道信号的当前帧。

在一些实施例中，多声道信号可以为多声道时域信号；在一些实施例中，多声道信号可以为多声道频域信号。

520、根据多声道信号，确定特征信息。

本发明实施例对特征信息的具体类型不作限定，在一些实施例中，所述特征信息可用于指示所述多声道信号的特征。在一些实施例中，特征信息可以包括当前帧的帧类型和信号类型中的至少一种，帧类型可以包括语音帧和/或非语音帧；信号类型可以包括清音和/或浊音。在一些实施例中，语音帧

为包含语音信号的帧。在一些实施例中，非语音帧也可称为背景帧。背景帧中的信号例如可以是背景噪声信号。此外，本发明实施例对语音帧和非语音帧的具体命名不作限定，例如，在语音活动检测（Voice Activity Detection, VAD）检测中，包含语音信号的帧可以称之为语音激活帧（或称 active 帧）；

5 非语音帧可以称为语音非激活帧（或称 inactive 帧）。下文以语音帧为语音激活帧，非语音帧为语音非激活帧为例进行举例说明。

本发明实施例对根据多声道信号，确定多声道信号的信号类型的具体方式不作限定。在一些实施例中，当多声道时域信号的过零率（Zero Crossing Rate, ZCR）大于预设阈值时，多声道信号的信号类型为清音（或当前帧为清音帧）；否则，多声道信号的信号类型为浊音（或当前帧为浊音帧）。在另一些实施例中，当多声道信号的相关值（用于描述多声道信号的相关性）大于预设阈值时，多声道信号的信号类型为浊音（或当前帧为浊音帧）；否则，多声道信号的信号类型为清音（或当前帧为清音帧）。

530、根据特征信息，确定当前帧的 ITD 参数。

15 具体地，可以根据当前帧的帧类型，确定当前帧的 ITD 参数。例如，对语音激活帧和语音非激活帧采用不同的 ITD 参数提取方式。或者，可以根据多声道信号的信号类型，确定当前帧的 ITD 参数。例如，对清音信号和浊音信号采用不同的 ITD 参数提取方式。后文会结合具体的例子进行详细描述，此处不再详述。

20 540、对所述 ITD 参数进行编码。

在一些实施例中，图 5 的方法还可以包括：向解码端发送编码后的 ITD 参数。

25 在一些实施例中，步骤 520 可包括：根据所述多声道信号，确定所述当前帧的帧类型；步骤 530 可包括：在所述当前帧为非语音帧的情况下，采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数；在所述当前帧为语音帧的情况下，采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数。

应理解，本发明实施例对根据所述多声道信号，确定所述当前帧的帧类型的方式不作具体限定。例如，可以基于 VAD 确定当前帧的帧类型。

还应理解，本发明实施例对第一 ITD 参数提取方式和第二 ITD 参数提取方式不作具体限定，只要第一 ITD 参数提取方式和第二 ITD 参数提取方式不同即可。

在一些实施例中，第一 ITD 参数提取方式可以是将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

在一些实施例中，第二 ITD 参数提取方式可以是根据所述多声道信号，确定所述当前帧的 ITD 参数。例如，可以采用现有技术中的基于时域、频域的方式提取当前帧的 ITD 参数。或者，可以在现有技术的基础上对提取出的初始 ITD 参数进行平滑处理，得到当前帧的 ITD 参数。或者，可以采用根据本发明实施例的基于混合域（时域和频域）的方式提取 ITD 参数，后文会对基于混合域的 ITD 参数进行详细描述，此处不再赘述。

下文以多声道信号为左右声道信号为例进行详细说明，但本发明实施例不限于此。实际中，本申请中的方案可应用于处理两声道或多声道信号中的任意两个声道，在多声道场景下，下文中的左右声道可以是该多声道中的任意两个声道。

图 6 是本发明实施例的多声道信号的编码方法的示意性流程图。应理解，图 6 示出的处理步骤或操作仅是示例，本发明实施例还可以执行其它操作或者图 6 中的各种操作的变形。此外，图 6 中的各个步骤可以按照与图 6 呈现的不同的顺序来执行，并且有可能并非要执行图 6 中的全部操作。图 6 的方法包括：

610、检测当前帧的帧类型。

具体地，可以对当前帧进行 VAD，并根据检测结果判断当前帧是语音激活帧，还是语音非激活帧。

620、判断当前帧是否为语音激活帧。

如果不是语音激活帧（即当前帧为语音非激活帧），可以执行步骤 630；如果是语音激活帧，可以执行步骤 640。

630、采用第一 ITD 参数提取方式确定当前帧的 ITD 参数。

5 在一些实施例中，第一 ITD 参数提取方式可以包括：将当前帧的前一帧或前一子帧的 ITD 参数确定为当前帧的 ITD 参数。

640、采用第二 ITD 参数提取方法确定当前帧的 ITD 参数。

可选地，在一些实施例中，可以采用图 3 描述的方式提取当前帧的 ITD 参数，即在时域提取当前帧的 ITD 参数。

10 可选地，在一些实施例中，可以采用图 4 描述的方式提取当前帧的 ITD 参数，即在频域提取当前帧的 ITD 参数。

可选地，在一些实施例中，可以在混合域提取当前帧的 ITD 参数，下文会结合图 7 和图 8，对根据本发明实施例在混合域提取 ITD 参数的方式进行详细描述，此处不再详述。

15 可选地，在一些实施例中，可以先提取当前帧的初始 ITD 参数；再对当前帧的初始 ITD 参数进行平滑处理，得到当前帧的 ITD 参数。

应理解，本发明实施例对当前帧的初始 ITD 参数的提取方式不作具体限定。

可选地，作为一种实现方式，可以采用图 3 描述的方式提取当前帧的初始 ITD 参数，即在时域提取当前帧的初始 ITD 参数。

20 可选地，作为一种实现方式，可以采用图 4 描述的方式提取当前帧的 ITD 参数，即在频域提取当前帧的初始 ITD 参数。

可选地，作为一种实现方式，可以在混合域提取当前帧的初始 ITD 参数，下文会结合图 7 和图 8，对根据本发明实施例的在混合域提取 ITD 参数的方式进行详细描述。

25 在提取出当前帧的初始 ITD 参数之后，可以采用如下公式进行平滑处理：

$$T_{sm} = w_1 * T_{sm}^{[-1]} + w_2 * T_1 \quad (5)$$

其中, $T_{sm}^{[l]}$ 为当前帧的前一帧的 ITD 参数的平滑值; 对于平滑因子 w_1 、 w_2 , 可以设置为常数, 例如 $w_1=0.75$, $w_2=0.25$, 或者 $w_1=0.8$, $w_2=0.2$, 或者 $w_1=0.9$, $w_2=0.1$ 等; 也可以根据 $T_{sm}^{[l]}$ 和 T_l 的大小关系设置为不同的因子; 也可以结合当前帧中的信号类型, 采用不同的平滑因子进行平滑处理。例如, 5 对于清音帧, 可以采用较小的平滑因子, 对于浊音帧, 可以采用较大的平滑因子。此外, w_1 、 w_2 的关系满足 $w_1+w_2=1$ 。

或者, 如果当前帧划分为 K 个子帧, 每个子帧可以对应一个初始 ITD 参数 (子帧 ITD 参数的提取方式与帧的 ITD 参数提取方式类似, 也可分为基于时域、频域、混合域的提取方式, 为避免重复, 此处赘述), 可以采用 10 如下公式对每个子帧的初始 ITD 参数进行平滑处理:

$$T_{sm}(j) = w_1 * T_{sm}(j-1) + w_2 * T(j) \quad (6)$$

其中, $T_{sm}(j-1)$ 为前一子帧的 ITD 参数的平滑值; 对于平滑因子 w_1 、 w_2 , 可以设置为常数, 例如 $w_1=0.75$, $w_2=0.25$, 或者 $w_1=0.8$, $w_2=0.2$, 或者 $w_1=0.9$, $w_2=0.1$ 等; 也可以根据 $T_{sm}(j-1)$ 和 $T(j)$ 的大小关系设置为不同的因子; 也可 15 以结合当前帧中的信号类型, 采用不同的平滑因子进行平滑处理。例如, 对于清音帧, 可以采用较小的平滑因子, 对于浊音帧, 可以采用较大的平滑因子。此外, w_1 、 w_2 的关系满足 $w_1+w_2=1$ 。

在一些实施例中, 平滑处理可以在编码实现, 也可以在解码端实现。

下面结合图 7 和图 8, 详细描述本发明实施例的基于混合域的 ITD 参数 20 提取方式。图 7 和图 8 描述的 ITD 参数提取方式可用于提取当前帧的 ITD 参数; 此外, 在需要进行平滑处理的实施例中, 图 7 和图 8 描述的 ITD 参数提取方式也可用于提取当前帧的初始 ITD 参数。图 7 和图 8 的 ITD 参数实现方式会在频域构造目标频域信号, 该目标频域信号的相位为多声道信号的 IPD, 这样, 当将目标频域信号转换到时域, 得到目标时域信号时, 当前帧 25 的 ITD 参数会位于该目标时域信号的采样值最大的采样点对应的索引值处。图 7 和图 8 的区别在于目标频域信号的构造方式不同。

图 7 是本发明实施例的多声道信号的编码方法的示意性流程图。在图 7 对应的实施例中，目标频域信号主要是通过逐频点计算单声道频域信号的幅值和左右声道信号的 IPD 而构建出的频域信号。应理解，图 7 示出的处理步骤或操作仅是示例，本发明实施例还可以执行其它操作或者图 7 中的各种操作的变形。此外，图 7 中的各个步骤可以按照与图 7 呈现的不同的顺序来执行，并且有可能并非要执行图 7 中的全部操作。

710、对左右声道的时域信号分别进行时频变换，得到左右声道的频域信号。

具体地，可以采用公式 (7) 和 (8) 对左右声道的时域信号进行离散傅里叶变换 (Discrete Fourier Transformation, DFT):

$$X_L(k) = \sum_{n=0}^{Length-1} x_L(n) \cdot e^{-j\frac{2\pi nk}{L}}, 0 \leq k < L \quad (7)$$

$$X_R(k) = \sum_{n=0}^{Length-1} x_R(n) \cdot e^{-j\frac{2\pi nk}{L}}, 0 \leq k < L \quad (8)$$

其中， $x_L(n)$ 和 $x_R(n)$ 分别为左右声道的时域信号， $Length$ 为帧长或子帧长， k 为频域信号的频点的索引值， L 为时频变换长度。

15 为了提升编码效率，可以采用快速傅里叶变换 (Fast Fourier Transformation, FFT) 进行时频变换，时频变换后得到的频域信号是复数信号，包含了实部和虚部，对于左声道的频域信号，其实部为 $X_{L_real}(k)$ ，虚部为 $X_{L_image}(k)$ ；对于右声道的频域信号，其实部为 $X_{R_real}(k)$ ，虚部为 $X_{R_image}(k)$ ，其中， $0 \leq k \leq \frac{L}{2}$ 。

20 具体地，以左声道的频域信号为例，其实部和虚部的取值可以采用如下计算方式：

$$X_{L_real}(0) = X_L(0), X_{L_image}(0) = 0 \quad (9)$$

$$X_{L_real}\left(\frac{L}{2}\right) = X_L(L-1), X_{L_image}\left(\frac{L}{2}\right) = 0 \quad (10)$$

$$X_{L_real}(k) = X_L(k), X_{L_image}(k) = X_L(L-k), 1 \leq k \leq \frac{L}{2} - 1 \quad (11)$$

或者,

$$X_{L_real}(0) = X_L(0), X_{L_image}(0) = 0 \quad (12)$$

$$X_{L_real}\left(\frac{L}{2}\right) = X_L(1), X_{L_image}\left(\frac{L}{2}\right) = 0 \quad (13)$$

$$X_{L_real}(k) = X_L(2k), X_{L_image}(k) = X_L(2k+1), 1 \leq k \leq \frac{L}{2}-1 \quad (14)$$

5 需要注意的是, 时频变换以后, 对于宽带信号 (WB 信号), 如果时频变换长度为 512, 则得到的频域信号包括 256 个频点, 其中第 256 个频点对应的是 8kHz 的频谱, 第 128 个频点对应的是 4kHz 的频谱, 其它依次类推。

720、对左右声道的频域信号进行频域系数处理, 得到目标频域信号。

10 在一些实施例中, 可以逐频点计算目标频域信号的幅值 $A_M(k)$ 以及声道间相位差 $IPD(k)$, 其中, k 为频点, $0 \leq k \leq L/2$, L 为将左右声道的时域信号变换成左右声道的频域信号时采用的时频变换长度。

具体地, 可以先计算目标频域信号的幅值 $A_M(k)$:

$$A_M(k) = \frac{\min(A_L(k), A_R(k))}{\max(A_L(k), A_R(k))} \quad (15)$$

其中, 左声道频域信号的幅值可以为:

$$15 \quad A_L(k) = \sqrt{X_{L_real}(k) * X_{L_real}(k) + X_{L_image}(k) * X_{L_image}(k)} \quad (16)$$

右声道频域信号的幅值可以为:

$$A_R(k) = \sqrt{X_{R_real}(k) * X_{R_real}(k) + X_{R_image}(k) * X_{R_image}(k)} \quad (17)$$

然后, 可以计算左右声道信号的 $IPD(k)$:

$$IPD(k) = \angle L(k) * R^*(k), \quad k_1 \leq k \leq k_2 \quad (18)$$

20 其中, k 表示频点, $L(k)$ 和 $R(k)$ 分别为左右声道频域信号第 k 个频点值, 该频点值包含实部和虚部, $R^*(k)$ 表示右声道频域信号第 k 个频点值的共轭, $L(k)$ 和 $R(k)$ 的实部和虚部可以基于 $X_L(k)$ 和 $X_R(k)$ 构建, 该公式 (18) 可以进一步整理为:

$$IPD(k) = \frac{A'(k) * A''(k)}{A'(k) * A'(k) + 0.28125 * A''(k) * A''(k)} \quad (19)$$

其中:

$$A'(k) = X_{L_real}(k) * X_{R_real}(k) + X_{L_image}(k) * X_{R_image}(k) \quad (20)$$

$$A''(k) = X_{L_image}(k) * X_{R_real}(k) - X_{L_real}(k) * X_{R_image}(k) \quad (21)$$

接着, 在得到目标频域信号的幅值, 以及左右声道信号的相位差之后,

5 进一步处理得到目标频域信号:

$$X_{M_real}(k) = A_M(k) * \cos(IPD(k)) \quad (22)$$

$$X_{M_image}(k) = A_M(k) * \sin(IPD(k)) \quad (23)$$

在一些实施例中, 在得到目标频域信号的幅值以及左右声道信号的 IPD 后, 可以采用查表法得到目标频域信号, 例如, 设置 sin 函数和 cos 函数表
10 格, 利用查表法得到目标频域信号, 这样可以有效降低算法的计算复杂度。

730、对目标频域信号进行频时变换, 得到目标时域信号。

在一些实施例中, 可以对目标频域信号加窗并进行逆离散傅里叶变换 (Inverse Discrete Fourier Transform, IDFT)。

具体地, 可以先对目标频域信号进行加窗处理:

$$15 \quad X_{M_real_win}(k) = X_{M_real}(k) * \sin\left(\frac{2 * k * \pi}{L}\right) \quad (24)$$

$$X_{M_image_win}(k) = X_{M_image}(k) * \sin\left(\frac{2 * k * \pi}{L}\right) \quad (25)$$

其中, k 为频点, $0 \leq k \leq L/2$, L 为将左右声道的时域信号变换成左右声道的频域信号时采用的时频变换长度。

然后, 对加窗后的信号进行 IDFT 变换, 得到目标时域信号:

$$20 \quad x_{real}(n) = \frac{1}{L/2} \sum_{k=0}^{k=L/2} X_{M_real_win}(k) * \cos\left(\frac{4\pi * k * n}{L}\right) + X_{M_image_win}(k) * \sin\left(\frac{4\pi * k * n}{L}\right) \quad (26)$$

$$x_{image}(n) = \frac{1}{L/2} \sum_{k=0}^{k=L/2} X_{M_image_win}(k) * \cos\left(\frac{4\pi * k * n}{L}\right) - X_{M_real_win}(k) * \sin\left(\frac{4\pi * k * n}{L}\right)$$

(27)

其中, n 为时域信号的采样点的索引值, $0 \leq n < L/2$ 。

在一些实施例中, 步骤 730 可以采用 IDFT 进行频时变换, 也可以采用

逆快速傅里叶变换 (Inverse Fast Fourier Transform, IFFT) 进行频时变换。

在一些实施例中, 可以无需对所有频点进行频时变换, 仅在特定频域范围内进行频时变换, 这样可以有效降低算法的计算复杂度。例如, 可以在频点范围 $[k_3, k_4]$ 内进行频时变换, 其中, $k_3 > 0$, $k_4 < L/2$ 。

5 740、对目标时域信号的幅值进行平滑处理。

具体地, 目标时域信号的幅值可以通过下式表示:

$$A(n) = \sqrt{x_{real}(n) * x_{real}(n) + x_{image}(n) * x_{image}(n)} \quad (28)$$

对目标时域信号的幅值进行平滑处理, 得到幅度平滑值 $A_{sm}(n)$:

$$A_{sm}(n) = w_1 * A_{sm}^{[-1]}(n) + w_2 * A(n) \quad (29)$$

10 其中, $A_{sm}^{[-1]}(n)$ 为当前帧的前一帧/子帧第 n 点的幅度平滑值; w_1 、 w_2 为平滑因子, 可以设置为常数, 也可以随 $A_{sm}^{[-1]}(n)$ 和 $A(n)$ 的大小关系的变化而变化, 同时满足 $w_1 + w_2 = 1$ 。例如, 可以设置 $w_1 = 0.75$, $w_2 = 0.25$, 或者 $w_1 = 0.8$,

$$w_2 = 0.2, \text{ 或者 } w_1 = 0.9, w_2 = 0.1, \text{ 或者 } w_1 = \frac{|A_{sm}^{[-1]}(n)|}{|A_{sm}^{[-1]}(n)| + |A(n)|}, \quad w_2 = \frac{|A(n)|}{|A_{sm}^{[-1]}(n)| + |A(n)|}。$$

15 750、搜索平滑后的目标时域信号的采样值最大的采样点对应的索引值, 得到 ITD 参数。

具体地, 搜索平滑后的时域信号的采样值最大的采样点对应的索引值 $index = \arg \max(A_{sm}(n))$, ITD 参数 = $index$ 。

20 从公式 (22) 和 (23) 可以看出, 频域系数处理以后得到的目标频域信号的相位为第一声道和第二声道的 IPD。进一步地, 由于 IPD 与 ITD 之间存在线性关系, 可以将目标频域信号近似改写成下式:

$$X_{M_real}(k) = A_M(k) * \cos\left(-\frac{2 * k * \pi}{L} ITD\right) \quad (30)$$

$$X_{M_image}(k) = A_M(k) * \sin\left(-\frac{2 * k * \pi}{L} ITD\right) \quad (31)$$

对目标频域信号进行频时变换之后, 目标时域信号的采样值最大的采样点对应的索引值会在 ITD 处。

图 8 是本发明实施例的多声道信号的编码方法的示意性流程图。在图 8 对应的实施例中，目标频域信号主要是基于左右声道信号中的一个声道的信号和另一个声道的信号的共轭构建出的频域信号。应理解，图 8 示出的处理步骤或操作仅是示例 本发明实施例还可以执行其它操作或者图 8 中的各种操作的变形。此外，图 8 中的各个步骤可以按照与图 8 呈现的不同的顺序来执行，并且有可能并非要执行图 8 中的全部操作。此外，图 8 中的各个步骤与图 7 中的各个步骤对应，不同之处在于步骤 820 的处理方式不同于步骤 720 的处理方式，其他步骤可以参照图 7，此处不再详述。

810、对左右声道的时域信号分别进行时频变换，得到左右声道的频域信号。

820、通过左右声道信号中一个声道的频域信号和另一声道的频域信号的共轭相乘，得到目标频域信号。

应理解，一个声道的频域信号和另一个声道的频域信号的共轭相乘，得到的频域信号的相位为这两个声道的 IPD。

具体地，目标频域信号 $X_M(k)$ 可以通过下式计算得到：

$$X_M(k) = L(k) * R^*(k) \quad (32)$$

其中， $L(k)$ 和 $R(k)$ 分别为左右声道频域信号第 k 个频点值，该频点值包含实部和虚部， $R^*(k)$ 表示右声道频域信号第 k 个频点值的共轭， $L(k)$ 和 $R(k)$ 的实部和虚部可以基于 $X_L(k)$ 和 $X_R(k)$ 构建。

或者 $X_M(k) = R(k) * L^*(k)$ (33)

其中， $R(k)$ 为右声道的频域信号第 k 个频点值， $L^*(k)$ 为左声道的频域信号第 k 个频点值的共轭， $0 \leq k \leq L/2$ 。

在一些实施例中，在得到 $X_M(k)$ 之后，可以进一步对 $X_M(k)$ 进行归一化处理，得到目标频域信号。

具体地，可以先计算 $X_M(k)$ 的最大幅值：

$$A_{\max} = \max_{0 \leq k \leq \frac{L}{2}} (\sqrt{X_{M_real}(k) * X_{M_real}(k) + X_{M_image}(k) * X_{M_image}(k)}) \quad (34)$$

然后对 $X_M(k)$ 的幅值进行归一化处理:

$$X_{M_real_nor}(k) = \frac{X_{M_real}(k)}{A_{\max}} \quad (35)$$

$$X_{M_image_nor}(k) = \frac{X_{M_image}(k)}{A_{\max}} \quad (36)$$

5 830、对目标频域信号进行频时变换，得到目标时域信号。

840、对目标时域信号的幅值进行平滑处理。

850、搜索平滑后的目标时域信号的采样值最大的采样点对应的索引值，得到 ITD 参数。

继续参见图 6，图 6 中的步骤 610 的实现方式可以有多种，例如，可以在时域检测当前帧的帧类型；也可以在频域检测当前帧的帧类型。

在一些实施例中，可以采用 VAD 检测算法检测当前帧的帧类型。具体地，可以基于当前帧中的信号的能量检测当前帧的帧类型。下面结合图 9 对基于能量的帧类型检测方式进行举例说明。

图 9 是当前帧的 ITD 参数的提取方式的示例性流程图。图 9 主要是基于当前帧中的信号的能量对当前帧进行 VAD 检测，以确定当前帧是语音激活帧，还是语音非激活帧。应理解，图 9 示出的处理步骤或操作仅是示例 本发明实施例还可以执行其它操作或者图 9 中的各种操作的变形。此外，图 9 中的各个步骤可以按照与图 9 呈现的不同的顺序来执行，并且有可能并非要执行图 9 中的全部操作。

20 910、对左右声道的时域信号进行时频变换。

具体地，可以对左右声道的时域信号分别进行快速傅里叶变换 (Fast Fourier Transformation, FFT) 变换，得到左右声道的频域信号:

$$X_L(k) = \sum_{n=0}^{Length-1} x_L(n) \cdot e^{-j\frac{2\pi nk}{L}}, 0 \leq k < L \quad (37)$$

$$X_R(k) = \sum_{n=0}^{Length-1} x_R(n) \cdot e^{-j\frac{2\pi nk}{L}}, 0 \leq k < L \quad (38)$$

其中, $x_L(n)$ 和 $x_R(n)$ 分别表示左右声道的时域信号, k 为频域信号的频点的索引值, $Length$ 为帧长, L 为时频变换长度。

FFT变换后得到的复数信号包含实部和虚部, 对于左声道频域信号实部为 $X_{L_real}(k)$, 虚部为 $X_{L_image}(k)$; 对于右声道信号实部为 $X_{R_real}(k)$, 虚部为 $X_{R_image}(k)$, 其中, $0 \leq k \leq L/2$ 。

在一些实施例中, $X_{L_real}(k)$ 、 $X_{L_image}(k)$ 可以采用如下公式描述的取值方式 ($X_{R_real}(k)$ 、 $X_{R_image}(k)$ 的取值方式同理, 此处不再赘述):

$$X_{L_real}(0) = X_L(0), X_{L_image}(0) = 0 \quad (39)$$

$$X_{L_real}\left(\frac{L}{2}\right) = X_L(L-1), X_{L_image}\left(\frac{L}{2}\right) = 0 \quad (40)$$

$$X_{L_real}(k) = X_L(k), X_{L_image}(k) = X_L(L-k), 1 \leq k \leq \frac{L}{2}-1 \quad (41)$$

或者:

$$X_{L_real}(0) = X_L(0), X_{L_image}(0) = 0 \quad (42)$$

$$X_{L_real}\left(\frac{L}{2}\right) = X_L(1), X_{L_image}\left(\frac{L}{2}\right) = 0 \quad (43)$$

$$X_{L_real}(k) = X_L(2*k), X_{L_image}(k) = X_L(2*k+1), 1 \leq k \leq \frac{L}{2}-1 \quad (44)$$

920、计算当前帧/子帧的能量。

具体地, 可以根据如下公式计算当前帧/子帧的能量:

$$E_{tot} = \sum_{k=0}^{L/2} X_{L_real}(k)^2 + X_{L_image}(k)^2 + X_{R_real}(k)^2 + X_{R_image}(k)^2 \quad (45)$$

然后, 可以判断当前帧/子帧的能量 E_{tot} 是否大于预设的VAD阈值 E_{VAD} ;

20 其中, E_{VAD} 可以设置为固定值, 也可以根据当前帧/子帧能量自适应调整。

如果 $E_{tot} \leq E_{VAD}$, 可以执行步骤930; 如果 $E_{tot} \geq E_{VAD}$, 可以执行步骤940。

930、采用第一ITD参数提取方式提取当前帧/子帧的ITD参数。

第一ITD参数提取方式可以是: 保持当前帧的前一帧/子帧的ITD值。

940、采用第二 ITD 参数提取方式提取当前帧/子帧的 ITD 参数。

第二 ITD 参数提取的实现方式可以参照图 6 的步骤 640。下面仍以基于混合域的 ITD 参数提取方式进行举例说明。

Step 1, 可以结合当前帧/子帧的能量进行频域系数处理。

5 具体地, 假设当前帧/子帧第 k 个频点的能量为 $E(k)$, 如果 $E(k)*L \leq E_{tot}$, 可以将目标频域信号当前频点设置为 0; 否则, 可以计算目标频域信号当前频点的幅值以及 IPD, 采用类似图 7 描述的方式, 处理得到目标频域信号, 其中, L 为将左右声道的时域信号变换成左右声道的频域信号时采用的时频变换长度。

10 具体地, 可以采用如下公式计算幅值 $A_M(k)$:

$$A_M(k) = \frac{\min(A_L(k), A_R(k))}{\max(A_L(k), A_R(k))} \quad (46)$$

其中, 左声道频域信号在第 k 个频点的幅值为:

$$A_L(k) = \sqrt{X_{L_real}(k) * X_{L_real}(k) + X_{L_image}(k) * X_{L_image}(k)};$$

右声道频域信号在第 k 个频点的幅值为:

$$15 \quad A_R(k) = \sqrt{X_{R_real}(k) * X_{R_real}(k) + X_{R_image}(k) * X_{R_image}(k)}。$$

其次, 可以采用如下公式计算左右声道信号的声道间相位差 $IPD(k)$:

$$IPD(k) = \angle L(k) * R^*(k) \quad (47)$$

其中, $L(k)$ 和 $R(k)$ 分别为左右声道频域信号第 k 个频点值, 该频点值包含实部和虚部, $R^*(k)$ 表示右声道频域信号第 k 个频点值的共轭, $L(k)$ 和 $R(k)$ 的实部和虚部可以基于 $X_L(k)$ 和 $X_R(k)$ 构建。

然后, 构建目标频域信号, 使得该目标频域信号的相位与左右声道信号的 IPD 线性相关。

具体地, 可以采用如下公式构建目标频域信号:

$$\begin{aligned} X_{M_real}(k) &= A_M(k) * \cos(IPD(k)) \\ X_{M_image}(k) &= A_M(k) * \sin(IPD(k)) \end{aligned} \quad (48)$$

25 Step 2, 对目标频域信号进行频时变换, 得到目标时域信号。

在一些实施例中，可以先对目标频域信号进行加窗及 IDFT 变换，得到目标时域信号。

具体地，可以采用如下公式对目标频域信号进行加窗处理：

$$\begin{aligned} X_{M_real_win}(k) &= X_{M_real}(k) * \sin\left(\frac{2 * k * \pi}{L}\right) & 0 \leq k \leq \frac{L}{2} \\ X_{M_image_win}(k) &= X_{M_image}(k) * \sin\left(\frac{2 * k * \pi}{L}\right) & 0 \leq k \leq \frac{L}{2} \end{aligned} \quad (49)$$

5 进一步地，可以采用如下公式对加窗后的频域信号进行 IDFT 变换，得到目标时域信号：

$$\begin{aligned} x_{real}(n) &= \frac{1}{L/2} \sum_{k=0}^{k=L/2} X_{M_real_win}(k) * \cos\left(\frac{4\pi * k * n}{L}\right) + X_{M_image_win}(k) * \sin\left(\frac{4\pi * k * n}{L}\right) \\ x_{image}(n) &= \frac{1}{L/2} \sum_{k=0}^{k=L/2} X_{M_image_win}(k) * \cos\left(\frac{4\pi * k * n}{L}\right) - X_{M_real_win}(k) * \sin\left(\frac{4\pi * k * n}{L}\right) \end{aligned} \quad (50)$$

其中， $0 \leq n < L/2$ 。

10 此外，在一些实施例中，还可以对目标时域信号的幅值 $A(n)$ 进行平滑处理，得到幅度平滑值 $A_{sm}(n)$ 。例如，可以结合信号类型采用不同的平滑因子进行平滑处理。例如，对于清音帧，采用较小的平滑因子，对于浊音帧，采用较大的平滑因子。

具体地，可以采用如下公式计算目标时域信号幅值 $A(n)$ ：

$$A(n) = \sqrt{x_{real}(n) * x_{real}(n) + x_{image}(n) * x_{image}(n)} \quad (51)$$

15 然后可以采用如下公式对 $A(n)$ 进行平滑处理，得到幅度平滑值 $A_{sm}(n)$ ：

$$A_{sm}(n) = w_1 * A_{sm}^{[-1]}(n) + w_2 * A(n) \quad (52)$$

其中， $A_{sm}^{[-1]}(n)$ 表示当前帧的前一帧/子帧第 n 点的幅度平滑值；对于平滑因子 w_1 、 w_2 可以设置为常数，也可以根据 $A_{sm}^{[-1]}(n)$ 和 $A(n)$ 的大小关系设置为不同的因子， w_1 、 w_2 满足 $w_1 + w_2 = 1$ 。例如，可以设置 $w_1 = 0.75$ ， $w_2 = 0.25$ ，或者

20 $w_1 = 0.8$ ， $w_2 = 0.2$ ，或者 $w_1 = 0.9$ ， $w_2 = 0.1$ ，或者 $w_1 = \frac{|A_{sm}^{[-1]}(n)|}{|A_{sm}^{[-1]}(n)| + |A(n)|}$ ，

$$w_2 = \frac{|A(n)|}{|A_{sm}^{[-1]}(n)| + |A(n)|}。$$

Step 3, 根据目标时域信号的采样值最大的采样点对应的索引值, 确定当前帧或当前子帧的 ITD 参数。

具体地, 可以将目标频域信号的采样值最大的采样点对应的索引值确定为当前帧或当前子帧的 ITD 参数。例如, 可以搜索平滑后时域信号的采样值最大的采样点对应的索引值 $index = \arg(\max(A_{sm}(n)))$, 得到 ITD 参数: $ITD = index$ 。

或者, 可以对目标频域信号的采样值最大的采样点对应的索引值进行变换 (如归一化、等比例缩放等), 将变换后的值确定为当前帧或当前子帧的 ITD 参数。

在上文描述的实现方式中, 如果当前帧为非语音帧, 可以将当前帧的前一帧或前一子帧的 ITD 参数确定为当前帧的 ITD 参数, 但本发明实施例不限于此。例如, 如果当前帧为非语音帧, 可以在时域、频域或混合域提取当前帧的 ITD 参数; 如果当前帧为语音帧, 且当前帧的前一帧也是语音激活帧 (即当前帧为连续语音帧中的一帧), 由于连续语音帧的 ITD 参数一般不会有大的波动, 如果当前帧的前一帧的 ITD 参数不为预设值, 而当前帧的 ITD 参数的计算结果为预设值 (预设值例如可以为 0), 这可能是当前帧的 ITD 参数计算错误引起的, 因此, 可以考虑将当前帧的前一帧或前一子帧的 ITD 参数确定为当前帧的 ITD 参数。下面结合图 10 对这种实现方式进行详细描述。

图 10 是当前帧的 ITD 参数的提取方式的示例性流程图。应理解, 图 10 示出的处理步骤或操作仅是示例, 本发明实施例还可以执行其它操作或者图 10 中的各种操作的变形。此外, 图 10 中的各个步骤可以按照与图 10 呈现的不同的顺序来执行, 并且有可能并非要执行图 10 中的全部操作。

1010、对左右声道的时域信号进行时频变换。

该步骤与步骤 910 类似, 可以参照步骤 910, 为避免重复, 此处不再详述。

1020、确定当前帧是否为语音激活帧。

具体地，可以基于左右声道的频域信号进行 VAD 检测。如果当前帧为语音非激活帧，执行步骤 1030；如果当前帧为语音激活帧，执行步骤 1040。

1030、采用第一 ITD 参数提取方式提取当前帧的 ITD 参数。

具体地，可以基于左右声道频域系数，根据频域互相关算法计算当前帧的 ITD 参数。频域互相关算法可以通过如下公式实现：

$$ITD = \arg \max(IFTT(\frac{L(k)R^*(k)}{|L(k)R^*(k)|})) \quad (53)$$

其中， $L(k)$ 和 $R(k)$ 分别为左右声道频域信号第 k 个频点值，该频点值包含实部和虚部， $R^*(k)$ 表示右声道频域信号第 k 个频点值的共轭， $L(k)$ 和 $R(k)$ 的实部和虚部可以基于 $X_L(k)$ 和 $X_R(k)$ 构建。

1040、采用第二 ITD 参数提取方式提取当前帧的 ITD 参数。

具体地，可以基于左右声道频域信号，结合当前帧的前一帧的 ITD 参数和/或计算出的连续为零的 ITD 参数的个数，对当前帧计算得到的 ITD 参数进行调整。

可选地，作为一种实现方式，通过 VAD 检测获知当前帧为连续语音帧（即当前帧的前一帧或前几帧均为语音帧）时，若当前帧的前一帧的 ITD 参数不为预设值（预设值例如可以是 0），而当前帧的 ITD 参数为预设值时，可以将当前帧的前一帧的 ITD 参数作为当前帧的 ITD 参数；否则，可以将当前帧的初始 ITD 参数确定为当前帧的 ITD 参数。

可选地，作为另一种实现方式，通过 VAD 检测获知当前帧为连续语音帧（即当前帧的前一帧或前几帧均为语音帧）时，若当前帧的前一帧的 ITD 参数不为预设值（预设值例如可以是 0），而当前帧的 ITD 参数为预设值，且连续计算得到的 ITD 参数（包括当前帧的 ITD 参数在内的）为预设值的个数小于预设阈值时，将当前帧的前一帧的 ITD 参数作为当前帧的 ITD 参数，并增加 ITD 参数连续为预设值的计数值；否则，可以将当前帧的初始 ITD 参数确定为当前帧的 ITD 参数。

上文结合图 5 至图 10, 详细描述了根据本发明实施例的多声道信号的编码方法。下文结合图 11 至图 12, 详细描述根据本发明实施例的编码器。应理解, 图 11 或图 12 对应的编码器能够执行图 5 至图 10 中的各个步骤, 为避免重复, 此处不再详述。

5 图 11 是本发明实施例的编码器的示意性结构图。图 11 的编码器 1100 包括:

获取单元 1110, 用于获取包含多声道信号的当前帧;

10 第一确定单元 1120, 用于根据所述多声道信号, 确定特征信息, 其中, 所述特征信息包括所述当前帧的帧类型和信号类型中的至少一种, 所述帧类型包括语音帧和/或非语音帧, 所述信号类型包括清音和/或浊音;

第二确定单元 1130, 用于根据所述特征信息, 确定所述当前帧的声道间时间差 ITD 参数;

编码单元 1140, 用于对所述 ITD 参数进行编码。

15 可选地, 作为一个实施例, 所述第一确定单元 1110 具体用于根据所述多声道信号, 确定所述当前帧的帧类型; 所述第二确定单元 1120 具体用于在所述当前帧为非语音帧的情况下, 采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数; 在所述当前帧为语音帧的情况下, 采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数。

20 可选地, 作为一个实施例, 所述第二确定单元 1120 具体用于将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

可选地, 作为一个实施例, 所述第二确定单元 1120 具体用于根据所述多声道信号, 确定所述当前帧的 ITD 参数。

25 所述第二确定单元 1120 具体用于根据所述多声道信号, 生成目标频域信号; 对所述目标频域信号进行频时变换, 得到目标时域信号; 根据所述目标时域信号, 确定所述当前帧的 ITD 参数。

所述第二确定单元 1120 具体用于根据所述多声道信号, 确定所述目标

频域信号的幅值；根据所述多声道信号，确定所述当前帧多声道信号的 IPD 参数；根据所述目标频域信号的幅值，以及所述当前帧多声道信号的 IPD 参数，生成所述目标频域信号。

所述第二确定单元 1120 具体用于根据 $A_M(k) = \frac{\min(A_1(k), A_2(k))}{\max(A_1(k), A_2(k))}$ ，确定所述

- 5 目标频域信号的幅值，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $A_1(k)$ 和 $A_2(k)$ 分别表示所述多声道信号中的任意两个声道的频域信号的幅值， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

所述第二确定单元 1120 具体用于根据 $\begin{cases} X_{M_real}(k) = A_M(k) * \cos(IPD(k)) \\ X_{M_image}(k) = A_M(k) * \sin(IPD(k)) \end{cases}$ ，

- 10 生成所述目标频域信号，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $X_{M_real}(k)$ 表示所述目标频域信号的实部， $X_{M_image}(k)$ 表示所述目标频域信号的虚部， $IPD(k)$ 表示所述 IPD 参数， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

- 所述第二确定单元 1120 具体用于根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，生成所述目
15 标频域信号，其中， $X_M(k)$ 表示所述目标频域信号， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

- 所述第二确定单元 1120 具体用于根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，确定频域信
20 号 $X_M(k)$ ，其中， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度；对所述频域信号 $X_M(k)$ 的幅值进行归一化处理，得到所述目标频域信号。

- 可选地，作为一个实施例，所述第二确定单元 1120 具体用于根据所述
25 多声道信号，确定所述当前帧的初始 ITD 参数；根据所述当前帧的前一帧或

前一子帧的 ITD 参数, 对所述当前帧的初始 ITD 参数进行平滑处理, 得到所述当前帧的 ITD 参数。

5 可选地, 作为一个实施例, 所述第二确定单元 1120 具体用于根据 $T_{sm}=w_1 * T_{sm}^{[l-1]} + w_2 * T_1$, 确定所述当前帧的 ITD 参数, 其中, T_1 表示所述当前帧的初始 ITD 参数, T_{sm} 表示所述当前帧的 ITD 参数, $T_{sm}^{[l-1]}$ 表示所述当前帧的前一帧或前一子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1+w_2=1$ 。

10 可选地, 作为一个实施例, 所述第二确定单元 1120 具体用于根据所述多声道信号, 确定所述当前帧的 K 个子帧的初始 ITD 参数, K 为大于 1 的整数; 根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数, 对所述每个子帧的初始 ITD 参数进行平滑处理, 得到所述每个子帧的 ITD 参数; 将所述 K 个子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

15 可选地, 作为一个实施例, 所述第二确定单元 1120 具体用于根据 $T_{sm}(j)=w_1 * T_{sm}(j-1) + w_2 * T(j)$, 确定所述每个子帧的 ITD 参数, 其中, $T(j)$ 表示所述 K 个子帧中的第 j 个子帧的初始 ITD 参数, $T_{sm}(j)$ 表示所述第 j 个子帧的 ITD 参数, $T_{sm}(j-1)$ 表示所述 K 个子帧中的第 $j-1$ 个子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, j 为整数, 且 $1 \leq j \leq K$, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1+w_2=1$ 。

20 可选地, 作为一个实施例, 所述平滑因子的取值是基于所述当前帧的信号类型确定的。

可选地, 作为一个实施例, 所述第一确定单元 1110 具体用于确定所述多声道信号的能量; 在所述多声道信号的能量小于或等于预设的能量阈值的情况下, 将所述当前帧确定为非语音帧; 在所述多声道信号的能量大于所述能量阈值的情况下, 将所述当前帧确定为语音帧。

25 可选地, 作为一个实施例, 所述编码器还包括: 第三确定单元, 用于根据所述多声道信号, 确定所述当前帧的初始 ITD 参数; 所述第二确定单元

1120 具体用于将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数；
对所述当前帧的初始 ITD 参数进行调整，得到所述当前帧的 ITD 参数。

5 可选地，作为一个实施例，所述第二确定单元 1120 具体用于根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始 ITD 参数，
确定所述当前帧的 ITD 参数，其中，N 为大于 1 的整数。

可选地，作为一个实施例，所述第二确定单元 1120 具体用于在所述当前帧的前一帧或者前 N 帧的帧类型为语音帧的情况下，根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数。

10 可选地，作为一个实施例，所述第二确定单元 1120 具体用于在所述当前帧的前一帧的 ITD 参数不为预设值，且所述当前帧的初始 ITD 参数为预设值的情况下，将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数；否则，可以将当前帧的初始 ITD 参数确定为当前帧的 ITD 参数。

15 图 12 是本发明实施例的编码器的示意性结构图。图 12 的编码器 1200 包括：

存储器 1210，用于存储程序；

20 处理器 1220，用于执行存储器 1210 中的程序，当所述程序被执行时，所述处理器 1220 获取包含多声道信号的当前帧；根据所述多声道信号，确定特征信息，其中，所述特征信息包括所述当前帧的帧类型和信号类型中的至少一种，所述帧类型包括语音帧和/或非语音帧，所述信号类型包括清音和/或浊音；根据所述特征信息，确定所述当前帧的声道间时间差 ITD 参数；对所述 ITD 参数进行编码。

25 可选地，作为一个实施例，所述处理器 1220 具体用于根据所述多声道信号，确定所述当前帧的帧类型；在所述当前帧为非语音帧的情况下，采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数；在所述当前帧为语音帧的情况下，采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于根据所述多声道信号，确定所述当前帧的 ITD 参数。

- 5 所述处理器 1220 具体用于根据所述多声道信号，生成目标频域信号；对所述目标频域信号进行频时变换，得到目标时域信号；根据所述目标时域信号，确定所述当前帧的 ITD 参数。

所述处理器 1220 具体用于根据所述多声道信号，确定所述目标频域信号的幅值；根据所述多声道信号，确定所述当前帧多声道信号的 IPD 参数；
10 根据所述目标频域信号的幅值，以及所述当前帧多声道信号的 IPD 参数，生成所述目标频域信号。

所述处理器 1220 具体用于根据 $A_M(k) = \frac{\min(A_1(k), A_2(k))}{\max(A_1(k), A_2(k))}$ ，确定所述目标频域信号的幅值，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $A_1(k)$ 和 $A_2(k)$ 分别表示所述多声道信号中的任意两个声道的频域信号的幅值， k 表示频点，
15 $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

所述处理器 1220 具体用于根据 $\begin{cases} X_{M_real}(k) = A_M(k) * \cos(IPD(k)) \\ X_{M_image}(k) = A_M(k) * \sin(IPD(k)) \end{cases}$ ，生成所述目标频域信号，其中， $A_M(k)$ 表示所述目标频域信号的幅值， $X_{M_real}(k)$ 表示所述目标频域信号的实部， $X_{M_image}(k)$ 表示所述目标频域信号的虚部，
20 $IPD(k)$ 表示所述 IPD 参数， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度。

所述处理器 1220 具体用于根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，生成所述目标频域信号，其中， $X_M(k)$ 表示所述目标频域信号， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ， L 表示将所述多声道信号从时域变换至频
25

域时采用的时频变换长度。

所述处理器 1220 具体用于根据 $X_M(k) = X_1(k) * X_2^*(k)$ ，确定频域信号 $X_M(k)$ ，其中， $X_1(k)$ 表示所述多声道信号中的第一声道的频域信号， $X_2^*(k)$ 表示所述多声道信号中的第二声道的频域信号的共轭， k 表示频点， $0 \leq k \leq L/2$ ，
5 L 表示将所述多声道信号从时域变换至频域时采用的时频变换长度；对所述频域信号 $X_M(k)$ 的幅值进行归一化处理，得到所述目标频域信号。

可选地，作为一个实施例，所述处理器 1220 具体用于根据所述多声道信号，确定所述当前帧的初始 ITD 参数；根据所述当前帧的前一帧或前一子帧的 ITD 参数，对所述当前帧的初始 ITD 参数进行平滑处理，得到所述当
10 前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于根据 $T_{sm} = w_1 * T_{sm}^{[-1]} + w_2 * T_1$ ，确定所述当前帧的 ITD 参数，其中， T_1 表示所述当前帧的初始 ITD 参数， T_{sm} 表示所述当前帧的 ITD 参数， $T_{sm}^{[-1]}$ 表示所述当前帧的前一帧或前一子帧的 ITD 参数， w_1 和 w_2 表示平滑因子，其中， w_1 和 w_2 的取值范围均
15 为 $[0,1]$ ，且 $w_1 + w_2 = 1$ 。

可选地，作为一个实施例，所述处理器 1220 具体用于根据所述多声道信号，确定所述当前帧的 K 个子帧的初始 ITD 参数， K 为大于 1 的整数；根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数，对所述每个子帧的初始 ITD 参数进行平滑处理，得到所述每个子帧的 ITD 参数；将所述 K
20 个子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于根据 $T_{sm}(j) = w_1 * T_{sm}(j-1) + w_2 * T(j)$ ，确定所述每个子帧的 ITD 参数，其中， $T(j)$ 表示所述 K 个子帧中的第 j 个子帧的初始 ITD 参数， $T_{sm}(j)$ 表示所述第 j 个子帧的 ITD 参数， $T_{sm}(j-1)$ 表示所述 K 个子帧中的第 $j-1$ 个子帧的 ITD 参数， w_1 和 w_2 表
25 示平滑因子， j 为整数，且 $1 \leq j \leq K$ ，其中， w_1 和 w_2 的取值范围均为 $[0,1]$ ，且 $w_1 + w_2 = 1$ 。

可选地，作为一个实施例，所述平滑因子的取值是基于所述当前帧的信号类型确定的。

可选地，作为一个实施例，所述处理器 1220 具体用于确定所述多声道信号的能量；在所述多声道信号的能量小于或等于预设的能量阈值的情况下，
5 将所述当前帧确定为非语音帧；在所述多声道信号的能量大于所述能量阈值的情况下，将所述当前帧确定为语音帧。

可选地，作为一个实施例，所述处理器 1220 还用于根据所述多声道信号，确定所述当前帧的初始 ITD 参数；所述处理器 1220 具体用于将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数；对所述当前帧的初始
10 ITD 参数进行调整，得到所述当前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，其中，N 为大于 1 的整数。

可选地，作为一个实施例，所述处理器 1220 具体用于在所述当前帧的
15 前一帧或者前 N 帧的帧类型为语音帧的情况下，根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数。

可选地，作为一个实施例，所述处理器 1220 具体用于在所述当前帧的前一帧的 ITD 参数不为预设值，且所述当前帧的初始 ITD 参数为预设值的情况下，将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数；
20 否则，可以将当前帧的初始 ITD 参数确定为当前帧的 ITD 参数。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案
25 的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描

述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

10 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

15 另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用
20 时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备）执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、磁碟或者光盘等各种可
25 以存储程序代码的介质。

以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限

于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应所述以权利要求的保护范围为准。

权利要求

1、一种多声道信号的编码方法，其特征在于，包括：

获取包含多声道信号的当前帧；

5 根据所述多声道信号，确定特征信息，其中，所述特征信息包括所述当前帧的帧类型和信号类型中的至少一种，所述帧类型包括语音帧和/或非语音帧，所述信号类型包括清音和/或浊音；

根据所述特征信息，确定所述当前帧的声道间时间差 ITD 参数；

对所述 ITD 参数进行编码。

10 2、如权利要求 1 所述的方法，其特征在于，所述根据所述多声道信号，确定特征信息，包括：

根据所述多声道信号，确定所述当前帧的帧类型；

所述根据所述特征信息，确定所述当前帧的 ITD 参数，包括：

15 在所述当前帧为非语音帧的情况下，采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数；

在所述当前帧为语音帧的情况下，采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数。

3、如权利要求 2 所述的方法，其特征在于，所述采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：

20 将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

4、如权利要求 2 或 3 所述的方法，其特征在于，所述采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：

根据所述多声道信号，确定所述当前帧的 ITD 参数。

25 5、如权利要求 4 所述的方法，其特征在于，所述根据所述多声道信号，确定所述当前帧的 ITD 参数，包括：

根据所述多声道信号，确定所述当前帧的初始 ITD 参数；

根据所述当前帧的前一帧或前一子帧的 ITD 参数,对所述当前帧的初始 ITD 参数进行平滑处理,得到所述当前帧的 ITD 参数。

- 6、如权利要求 5 所述的方法,其特征在于,所述根据所述当前帧的前一帧或前一子帧的 ITD 参数,对所述当前帧的初始 ITD 参数进行平滑处理,得到所述当前帧的 ITD 参数,包括:

根据 $T_{sm}=w_1 * T_{sm}^{[-1]} + w_2 * T_1$, 确定所述当前帧的 ITD 参数, 其中, T_1 表示所述当前帧的初始 ITD 参数, T_{sm} 表示所述当前帧的 ITD 参数, $T_{sm}^{[-1]}$ 表示所述当前帧的前一帧或前一子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1+w_2=1$ 。

- 7、如权利要求 4 所述的方法,其特征在于,所述根据所述多声道信号,确定所述当前帧的 ITD 参数,包括:

根据所述多声道信号,确定所述当前帧的 K 个子帧的初始 ITD 参数, K 为大于 1 的整数;

- 根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数,对所述每个子帧的初始 ITD 参数进行平滑处理,得到所述每个子帧的 ITD 参数;

将所述 K 个子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

- 8、如权利要求 7 所述的方法,其特征在于,所述根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数,对所述每个子帧的初始 ITD 参数进行平滑处理,得到所述每个子帧的 ITD 参数,包括:

- 根据 $T_{sm}(j)=w_1 * T_{sm}(j-1) + w_2 * T(j)$, 确定所述每个子帧的 ITD 参数, 其中, $T(j)$ 表示所述 K 个子帧中的第 j 个子帧的初始 ITD 参数, $T_{sm}(j)$ 表示所述第 j 个子帧的 ITD 参数, $T_{sm}(j-1)$ 表示所述 K 个子帧中的第 j-1 个子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, j 为整数, 且 $1 \leq j \leq K$, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1+w_2=1$ 。

- 9、如权利要求 6 或 8 所述的方法,其特征在于,所述平滑因子的取值是基于所述当前帧的信号类型确定的。

10、如权利要求 2 所述的方法，其特征在于，所述方法还包括：

根据所述多声道信号，确定所述当前帧的初始 ITD 参数；

所述采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：

将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数；

5 所述采用第二 ITD 参数提取方式确定所述当前帧的 ITD 参数，包括：

对所述当前帧的初始 ITD 参数进行调整，得到所述当前帧的 ITD 参数。

11、如权利要求 10 所述的方法，其特征在于，所述对所述当前帧的初始 ITD 参数进行调整，得到所述当前帧的 ITD 参数，包括：

10 根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，其中，N 为大于 1 的整数。

12、如权利要求 11 所述的方法，其特征在于，所述根据所述当前帧的前一帧或者前 N 帧的帧类型，以及所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，包括：

15 在所述当前帧的前一帧或者前 N 帧的帧类型为语音帧的情况下，根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数。

13、如权利要求 12 所述的方法，其特征在于，所述根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数，确定所述当前帧的 ITD 参数，包括：

20 在所述当前帧的前一帧的 ITD 参数不为预设值，且所述当前帧的初始 ITD 参数为预设值的情况下，将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数；否则，将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数。

14、一种编码器，其特征在于，包括：

25 获取单元，用于获取包含多声道信号的当前帧；

第一确定单元，用于根据所述多声道信号，确定特征信息，其中，所述

特征信息包括所述当前帧的帧类型和信号类型中的至少一种，所述帧类型包括语音帧和/或非语音帧，所述信号类型包括清音和/或浊音；

第二确定单元，用于根据所述特征信息，确定所述当前帧的声道间时间差 ITD 参数；

5 编码单元，用于对所述 ITD 参数进行编码。

15、如权利要求 14 所述的编码器，其特征在于，所述第一确定单元具体用于根据所述多声道信号，确定所述当前帧的帧类型；所述第二确定单元具体用于在所述当前帧为非语音帧的情况下，采用第一 ITD 参数提取方式确定所述当前帧的 ITD 参数；在所述当前帧为语音帧的情况下，采用第二 ITD
10 参数提取方式确定所述当前帧的 ITD 参数。

16、如权利要求 15 所述的编码器，其特征在于，所述第二确定单元具体用于将所述当前帧的前一帧或前一子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

17、如权利要求 15 或 16 所述的编码器，其特征在于，所述第二确定单元
15 具体用于根据所述多声道信号，确定所述当前帧的 ITD 参数。

18、如权利要求 17 所述的编码器，其特征在于，所述第二确定单元具体用于根据所述多声道信号，确定所述当前帧的初始 ITD 参数；根据所述当前帧的前一帧或前一子帧的 ITD 参数，对所述当前帧的初始 ITD 参数进行平滑处理，得到所述当前帧的 ITD 参数。

20 19、如权利要求 18 所述的编码器，其特征在于，所述第二确定单元具体用于根据 $T_{sm}=w_1 * T_{sm}^{[-1]} + w_2 * T_1$ ，确定所述当前帧的 ITD 参数，其中， T_1 表示所述当前帧的初始 ITD 参数， T_{sm} 表示所述当前帧的 ITD 参数， $T_{sm}^{[-1]}$ 表示所述当前帧的前一帧或前一子帧的 ITD 参数， w_1 和 w_2 表示平滑因子，其中， w_1 和 w_2 的取值范围均为 $[0,1]$ ，且 $w_1+w_2=1$ 。

25 20、如权利要求 17 所述的编码器，其特征在于，所述第二确定单元具体用于根据所述多声道信号，确定所述当前帧的 K 个子帧的初始 ITD 参数，

K 为大于 1 的整数；根据所述 K 个子帧中的每个子帧的前一子帧的 ITD 参数,对所述每个子帧的初始 ITD 参数进行平滑处理,得到所述每个子帧的 ITD 参数；将所述 K 个子帧的 ITD 参数确定为所述当前帧的 ITD 参数。

21、如权利要求 20 所述的编码器,其特征在于,所述第二确定单元具体用于根据 $T_{sm}(j) = w_1 * T_{sm}(j-1) + w_2 * T(j)$, 确定所述每个子帧的 ITD 参数,其中, $T(j)$ 表示所述 K 个子帧中的第 j 个子帧的初始 ITD 参数, $T_{sm}(j)$ 表示所述第 j 个子帧的 ITD 参数, $T_{sm}(j-1)$ 表示所述 K 个子帧中的第 j-1 个子帧的 ITD 参数, w_1 和 w_2 表示平滑因子, j 为整数, 且 $1 \leq j \leq K$, 其中, w_1 和 w_2 的取值范围均为 $[0,1]$, 且 $w_1 + w_2 = 1$ 。

22、如权利要求 19 或 21 所述的编码器,其特征在于,所述平滑因子的取值是基于所述当前帧的信号类型确定的。

23、如权利要求 15 所述的编码器,其特征在于,所述编码器还包括:

第三确定单元,用于根据所述多声道信号,确定所述当前帧的初始 ITD 参数;

所述第二确定单元具体用于将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数;对所述当前帧的初始 ITD 参数进行调整,得到所述当前帧的 ITD 参数。

24、如权利要求 23 所述的编码器,其特征在于,所述第二确定单元具体用于根据所述当前帧的前一帧或者前 N 帧的帧类型,以及所述当前帧的初始 ITD 参数,确定所述当前帧的 ITD 参数,其中, N 为大于 1 的整数。

25、如权利要求 24 所述的编码器,其特征在于,所述第二确定单元具体用于在所述当前帧的前一帧或者前 N 帧的帧类型为语音帧的情况下,根据所述当前帧的前一帧的 ITD 参数和所述当前帧的初始 ITD 参数,确定所述当前帧的 ITD 参数。

26、如权利要求 25 所述的编码器,其特征在于,所述第二确定单元具体用于在所述当前帧的前一帧的 ITD 参数不为预设值,且所述当前帧的初始

ITD 参数为预设值的情况下, 将所述当前帧的前一帧的 ITD 参数确定为所述当前帧的 ITD 参数; 否则, 将所述当前帧的初始 ITD 参数确定为所述当前帧的 ITD 参数。

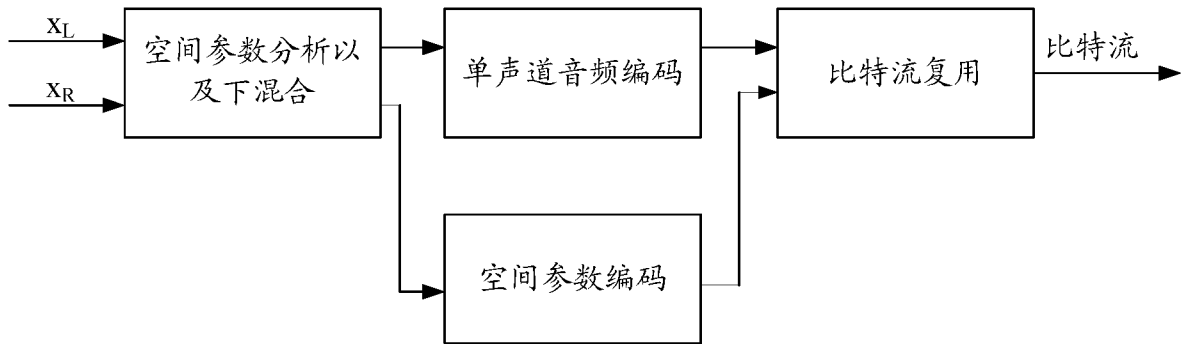


图 1

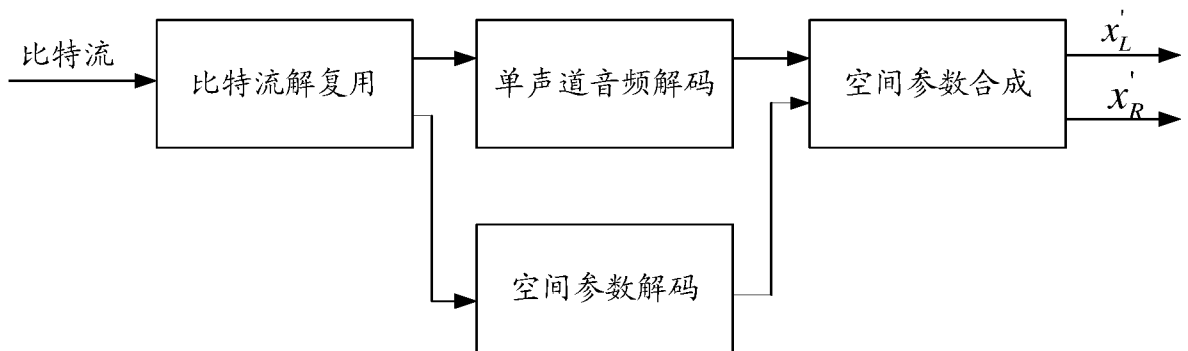


图 2

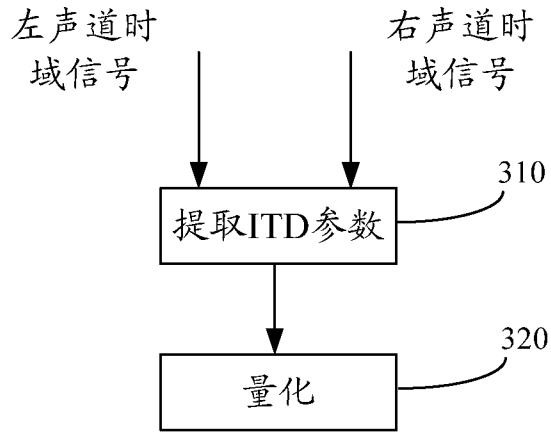


图 3

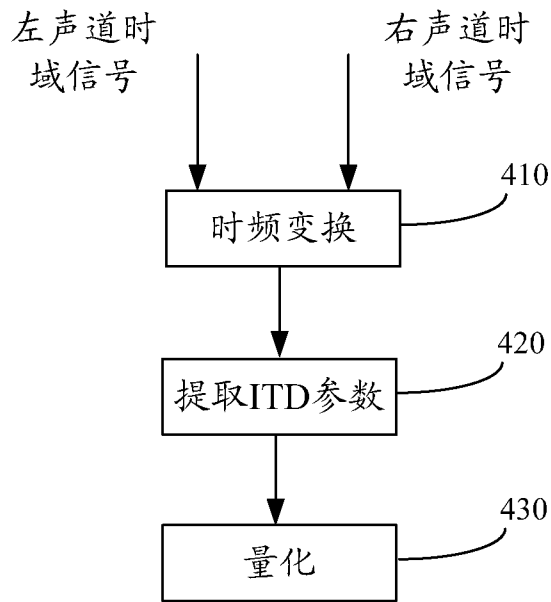


图 4

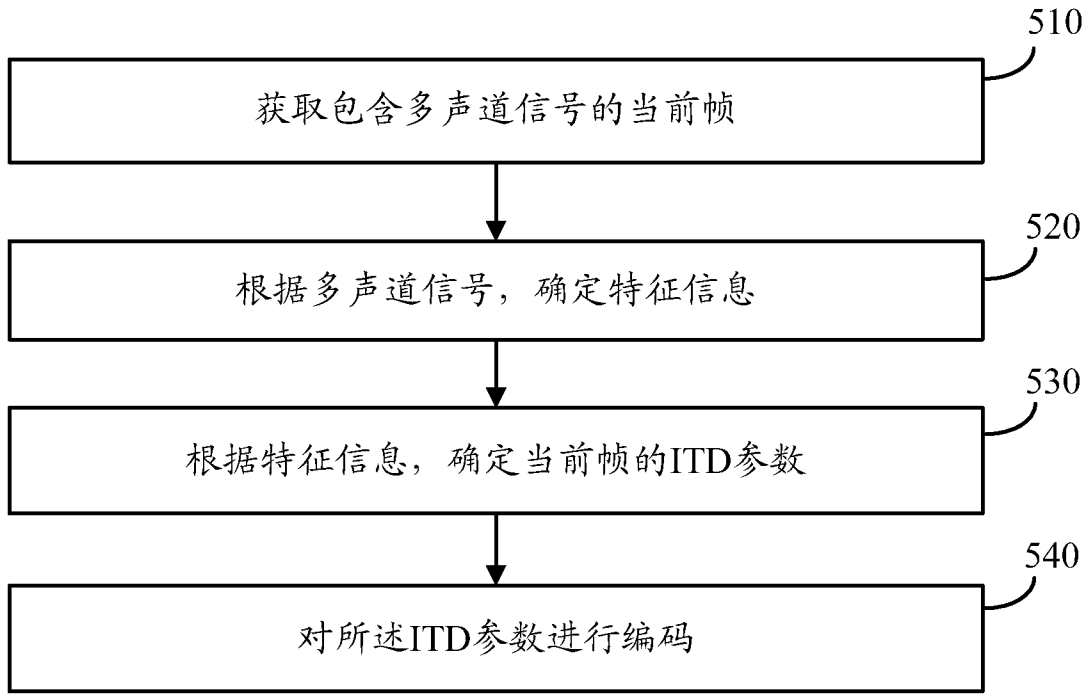


图 5

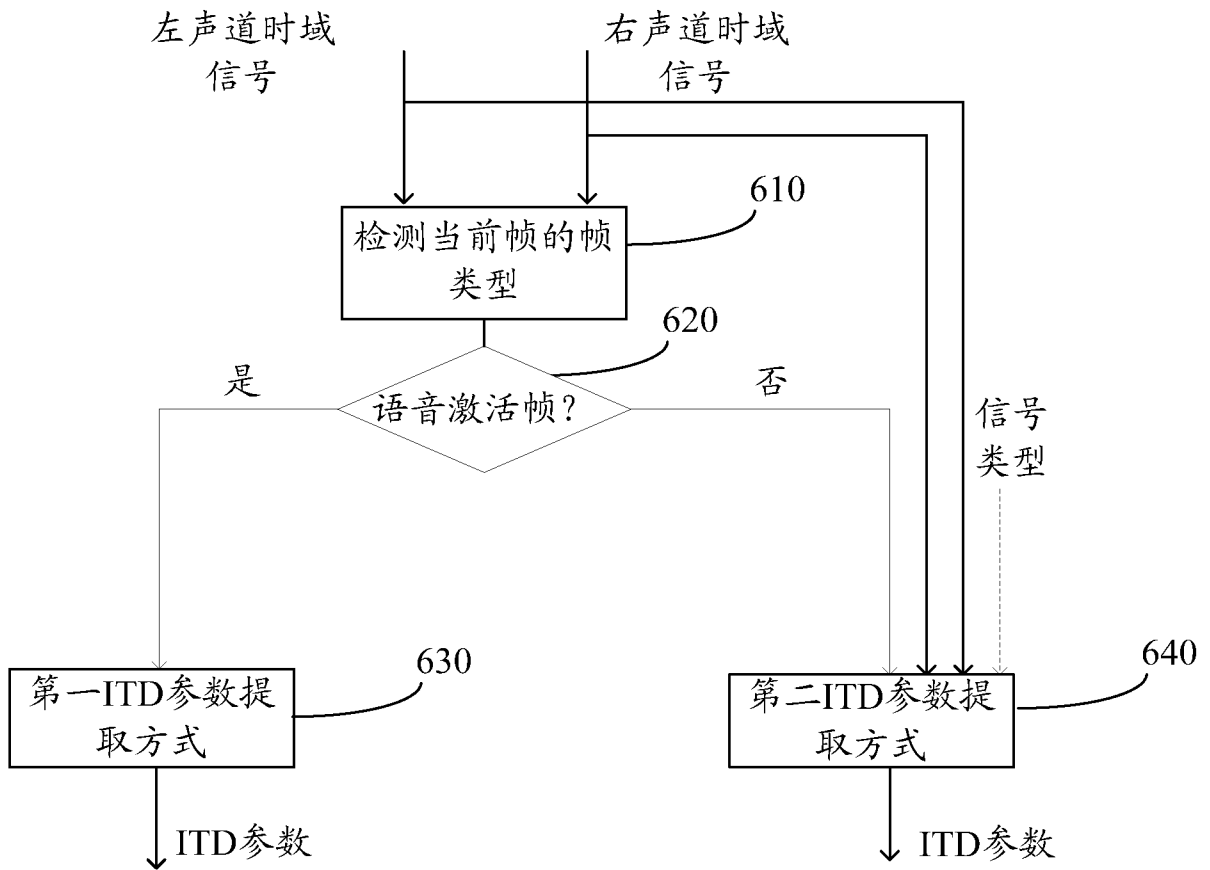


图 6

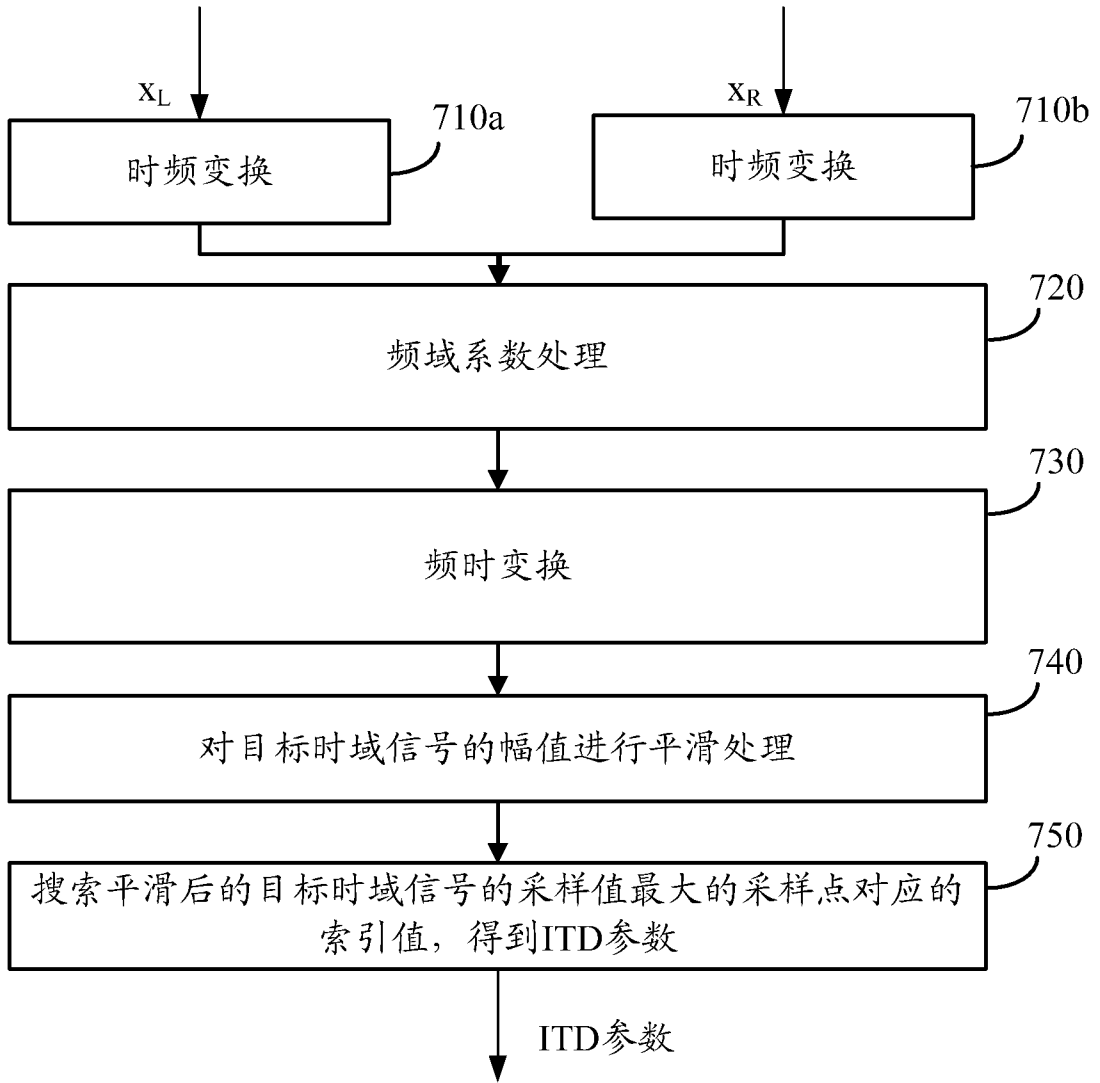


图 7

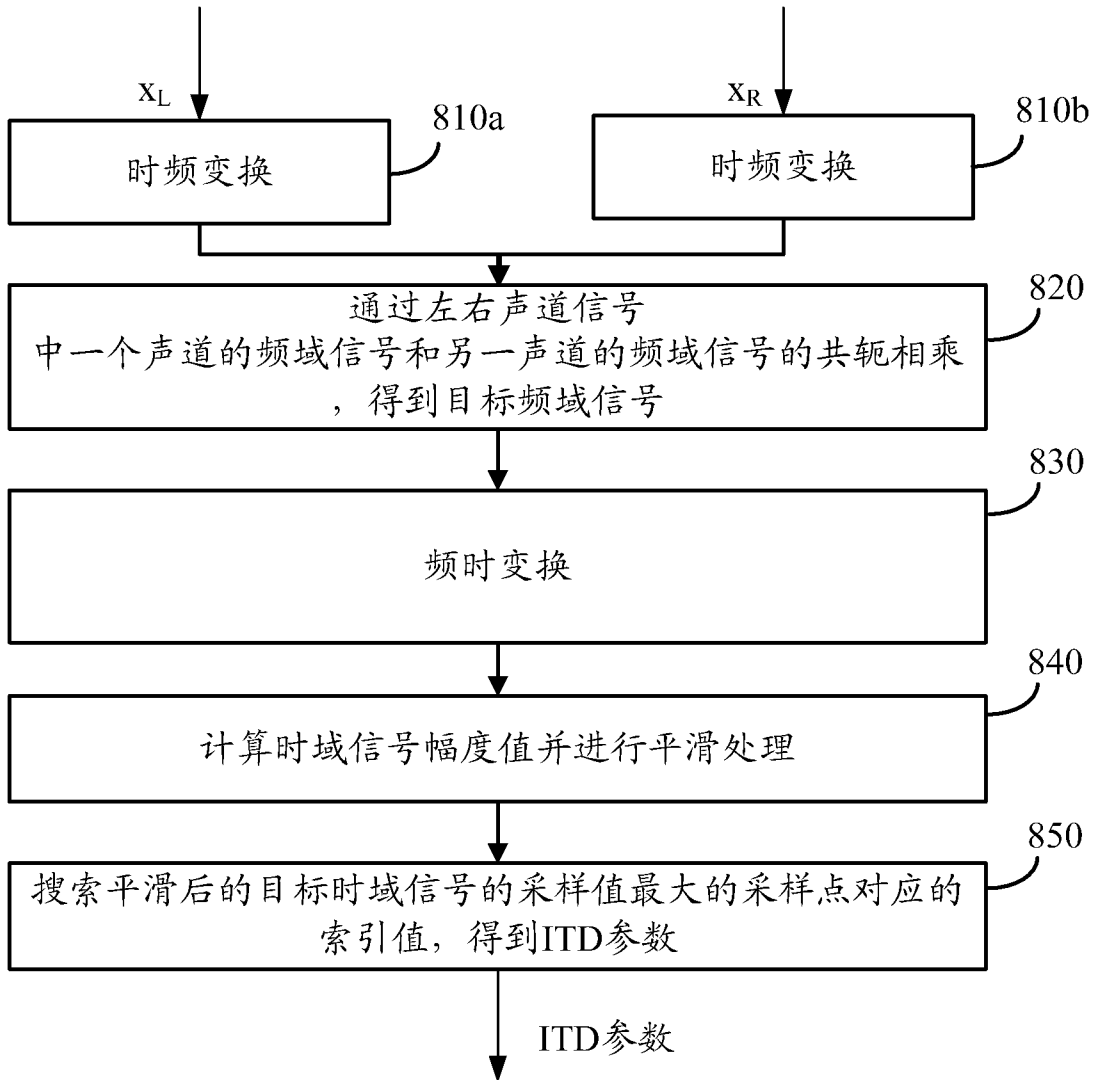


图 8

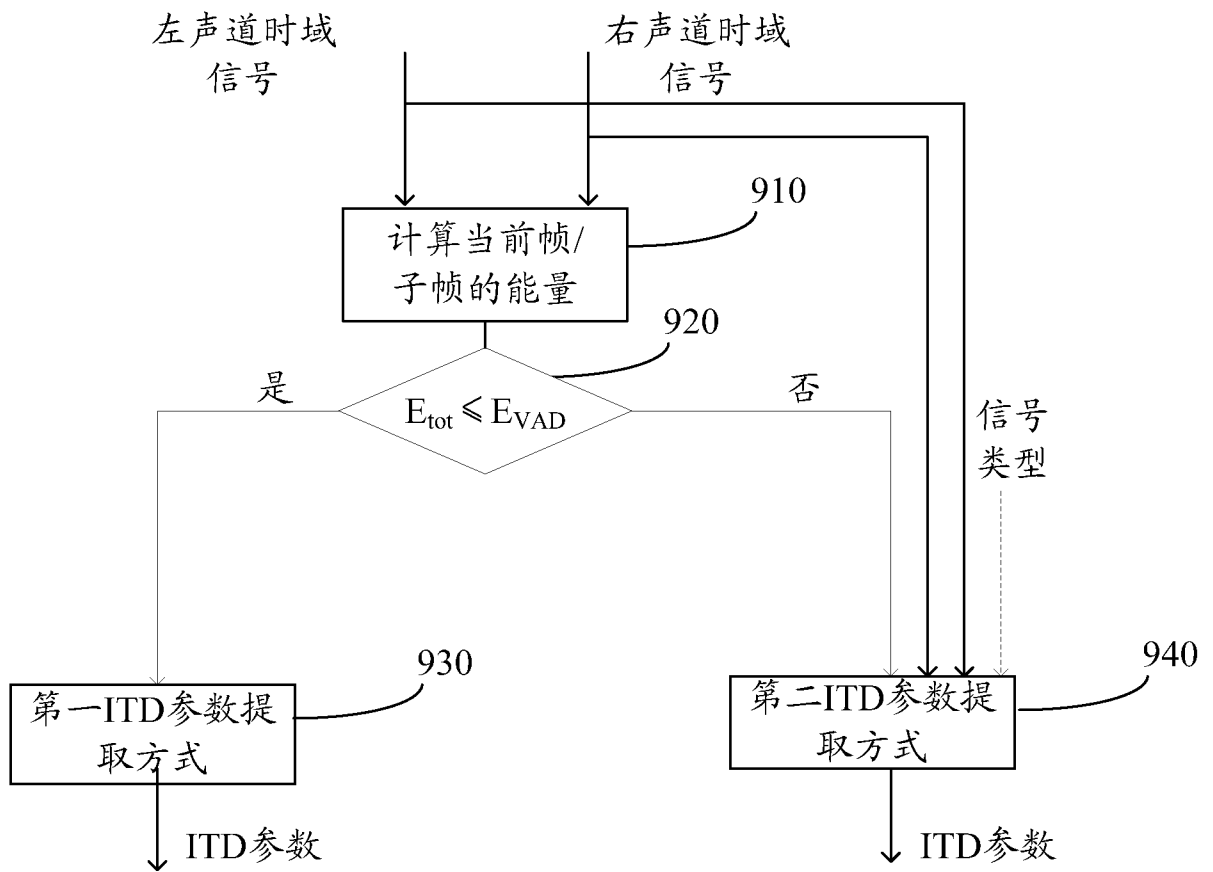


图 9

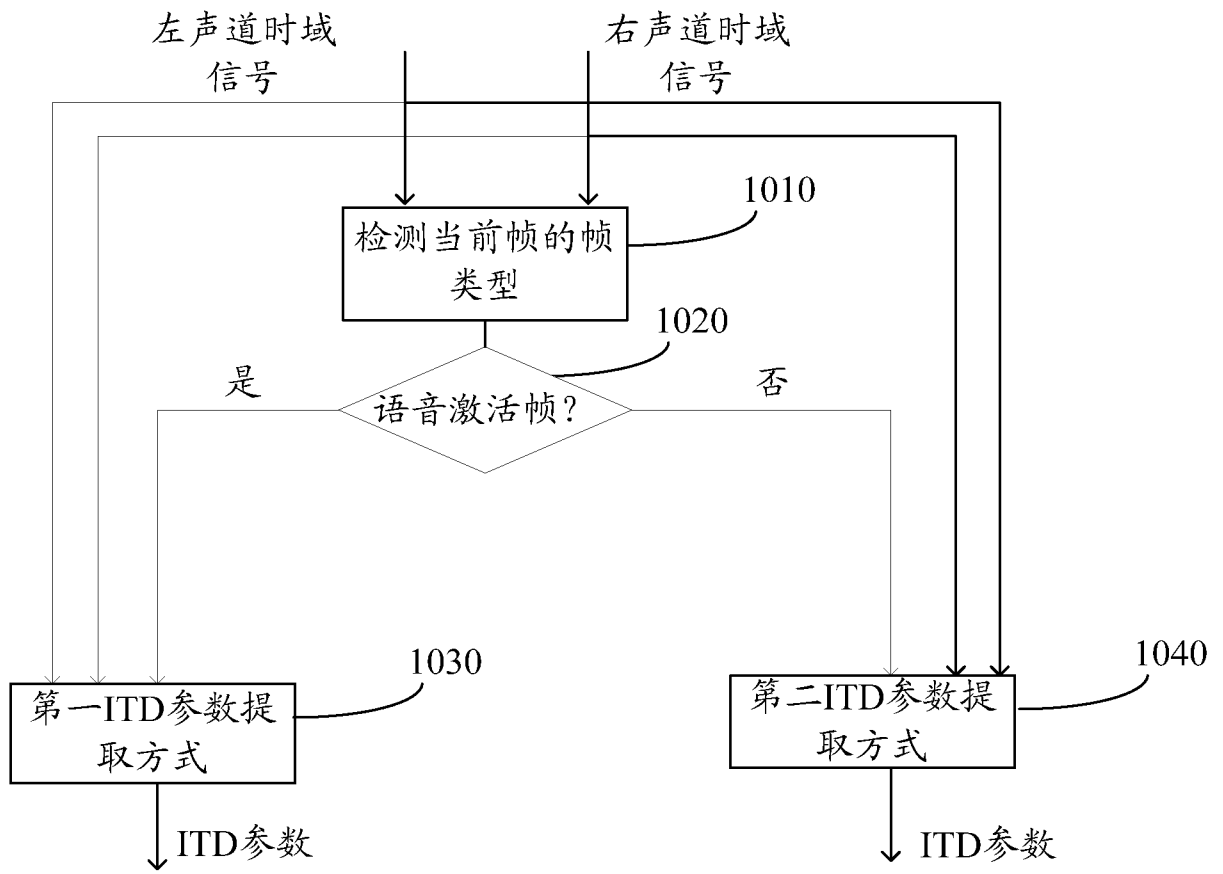


图 10

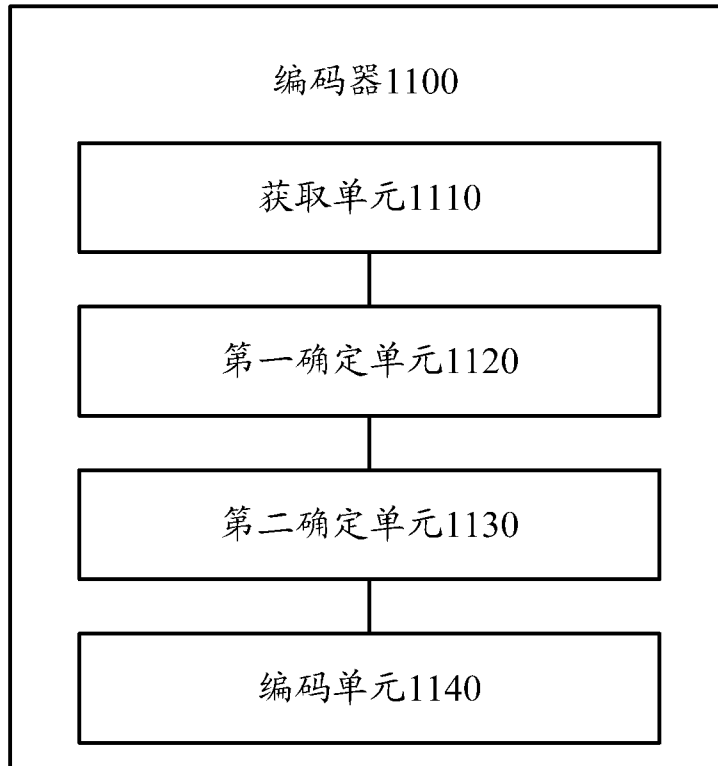


图 11



图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/103596

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/008 (2013.01) 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)

G10L 19/-; G10L 21/-

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, WPI, EPODOC: non-speech, unvoiced sound, time difference, multi w channel, stereo, speece, voice, non, surd, sonant, parameter, time, difference, ITD

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103295577 A (SHENZHEN RISING SOURCE TECHNOLOGY CO., LTD.), 11 September 2013 (11.09.2013), description, paragraphs [0078]-[0090], and figures 5-7	1, 14
A	CN 103295577 A (SHENZHEN RISING SOURCE TECHNOLOGY CO., LTD.), 11 September 2013 (11.09.2013), description, paragraphs [0078]-[0090], and figures 5-7	2-13, 15-26
A	CN 1428953 A (XI'AN DATANG TELECOM CO., LTD.), 09 July 2003 (09.07.2003), the whole document	1-26
A	CN 103180899 A (PANASONIC CORPORATION), 26 June 2013 (26.06.2013), the whole document	1-26
A	CN 101147191 A (PANASONIC CORPORATION), 19 March 2008 (19.03.2008), the whole document	1-26
A	CN 101517637 A (ROYAL DUTCH PHILIPS ELECTRONICS LTD.), 26 August 2009 (26.08.2009), the whole document	1-26
A	US 9129593 B2 (NOKIA TECHNOLOGIES OY), 08 September 2015 (08.09.2015), the whole document	1-26

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

<p>Date of the actual completion of the international search</p> <p style="text-align: center;">16 December 2016 (16.12.2016)</p>	<p>Date of mailing of the international search report</p> <p style="text-align: center;">05 January 2017 (05.01.2017)</p>
<p>Name and mailing address of the ISA/CN:</p> <p>State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451</p>	<p>Authorized officer</p> <p style="text-align: center;">WANG, Xinning</p> <p>Telephone No.: (86-10) 62413706</p>

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/103596

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2011123031 A1 (NOKIA CORP.), 26 May 2011 (26.05.2011), the whole document	1-26

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/103596

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103295577 A	11 September 2013	CN 103295577 B	02 September 2015
CN 1428953 A	09 July 2003	CN 1284319 C	08 November 2006
CN 103180899 A	26 June 2013	US 2013223633 A1	29 August 2013
		JP 5753540 B2	22 July 2015
		WO 2012066727 A1	24 May 2012
		CN 103180899 B	22 July 2015
CN 101147191 A	19 March 2008	US 2009055172 A1	26 February 2009
		EP 1858006 A1	21 November 2007
		WO 2006104017 A1	05 October 2006
		JP 4887288 B2	29 February 2012
		CN 101147191 B	13 July 2011
		US 8768691 B2	01 July 2014
CN 101517637 A	26 August 2009	WO 2008035275 A2	27 March 2008
		RU 2009114741 A	27 October 2010
		US 2009326960 A1	31 December 2009
		EP 2067138 B1	23 February 2011
		JP 2010503887 A	04 February 2010
		DE 602007012730 D1	07 April 2011
		KR 20090080945 A	27 July 2009
		JP 5281575 B2	04 September 2013
		US 8271290 B2	18 September 2012
		CN 101517637 B	15 August 2012
		RU 2460155 C2	27 August 2012
		EP 2067138 A2	10 June 2009
		KR 101396140 B1	20 May 2014
		BR PI0716854 A2	01 October 2013
		AT 499677 T	15 March 2011
		MX 2009002795 A	01 April 2009
		ES 2360740 T3	08 June 2011
US 9129593 B2	08 September 2015	TW 201126509 A	01 August 2011
		EP 2427881 A1	14 March 2012
		US 2011123031 A1	26 May 2011
		TW I508058 B	11 November 2015
		GB 2470059 A	10 November 2010
		WO 2010128386 A1	11 November 2010
US 2011123031 A1	26 May 2011	TW 201126509 A	01 August 2011
		US 9129593 B2	08 September 2015
		EP 2427881 A1	14 March 2012
		TW I508058 B	11 November 2015
		GB 2470059 A	10 November 2010
		WO 2010128386 A1	11 November 2010

国际检索报告

国际申请号

PCT/CN2016/103596

<p>A. 主题的分类</p> <p>G10L 19/008(2013.01)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>G10L 19/-; G10L21/-</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, WPI, EPODOC: 多声道, 立体声, 语音, 非语音, 清音, 浊音, 参数, 时间差, multi w channel, stereo, speece, voice, non, surd, sonant, parameter, time, difference, ITD</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 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7</td> <td>1, 14</td> </tr> <tr> <td>A</td> <td>CN 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7</td> <td>2-13, 15-26</td> </tr> <tr> <td>A</td> <td>CN 1428953 A (西安大唐电信有限公司) 2003年 7月 9日 (2003 - 07 - 09) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 103180899 A (松下电器产业株式会社) 2013年 6月 26日 (2013 - 06 - 26) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 101147191 A (松下电器产业株式会社) 2008年 3月 19日 (2008 - 03 - 19) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 101517637 A (皇家飞利浦电子股份有限公司) 2009年 8月 26日 (2009 - 08 - 26) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>US 9129593 B2 (NOKIA TECHNOLOGIES OY) 2015年 9月 8日 (2015 - 09 - 08) 全文</td> <td>1-26</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 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7	1, 14	A	CN 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7	2-13, 15-26	A	CN 1428953 A (西安大唐电信有限公司) 2003年 7月 9日 (2003 - 07 - 09) 全文	1-26	A	CN 103180899 A (松下电器产业株式会社) 2013年 6月 26日 (2013 - 06 - 26) 全文	1-26	A	CN 101147191 A (松下电器产业株式会社) 2008年 3月 19日 (2008 - 03 - 19) 全文	1-26	A	CN 101517637 A (皇家飞利浦电子股份有限公司) 2009年 8月 26日 (2009 - 08 - 26) 全文	1-26	A	US 9129593 B2 (NOKIA TECHNOLOGIES OY) 2015年 9月 8日 (2015 - 09 - 08) 全文	1-26
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
X	CN 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7	1, 14																								
A	CN 103295577 A (深圳广晟信源技术有限公司) 2013年 9月 11日 (2013 - 09 - 11) 说明书第[0078]-[0090]段、附图5-7	2-13, 15-26																								
A	CN 1428953 A (西安大唐电信有限公司) 2003年 7月 9日 (2003 - 07 - 09) 全文	1-26																								
A	CN 103180899 A (松下电器产业株式会社) 2013年 6月 26日 (2013 - 06 - 26) 全文	1-26																								
A	CN 101147191 A (松下电器产业株式会社) 2008年 3月 19日 (2008 - 03 - 19) 全文	1-26																								
A	CN 101517637 A (皇家飞利浦电子股份有限公司) 2009年 8月 26日 (2009 - 08 - 26) 全文	1-26																								
A	US 9129593 B2 (NOKIA TECHNOLOGIES OY) 2015年 9月 8日 (2015 - 09 - 08) 全文	1-26																								
<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>2016年 12月 16日</p>		<p>国际检索报告邮寄日期</p> <p>2017年 1月 5日</p>																								
<p>ISA/CN的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>王馨宁</p> <p>电话号码 (86-10)62413706</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 2011123031 A1 (NOKIA CORP.) 2011年 5月 26日 (2011 - 05 - 26) 全文	1-26

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/103596

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	103295577	A	2013年 9月 11日	CN	103295577	B	2015年 9月 2日
CN	1428953	A	2003年 7月 9日	CN	1284319	C	2006年 11月 8日
CN	103180899	A	2013年 6月 26日	US	2013223633	A1	2013年 8月 29日
				JP	5753540	B2	2015年 7月 22日
				WO	2012066727	A1	2012年 5月 24日
				CN	103180899	B	2015年 7月 22日
CN	101147191	A	2008年 3月 19日	US	2009055172	A1	2009年 2月 26日
				EP	1858006	A1	2007年 11月 21日
				WO	2006104017	A1	2006年 10月 5日
				JP	4887288	B2	2012年 2月 29日
				CN	101147191	B	2011年 7月 13日
				US	8768691	B2	2014年 7月 1日
CN	101517637	A	2009年 8月 26日	WO	2008035275	A2	2008年 3月 27日
				RU	2009114741	A	2010年 10月 27日
				US	2009326960	A1	2009年 12月 31日
				EP	2067138	B1	2011年 2月 23日
				JP	2010503887	A	2010年 2月 4日
				DE	602007012730	D1	2011年 4月 7日
				KR	20090080945	A	2009年 7月 27日
				JP	5281575	B2	2013年 9月 4日
				US	8271290	B2	2012年 9月 18日
				CN	101517637	B	2012年 8月 15日
				RU	2460155	C2	2012年 8月 27日
				EP	2067138	A2	2009年 6月 10日
				KR	101396140	B1	2014年 5月 20日
				BR	PI0716854	A2	2013年 10月 1日
				AT	499677	T	2011年 3月 15日
				MX	2009002795	A	2009年 4月 1日
				ES	2360740	T3	2011年 6月 8日
US	9129593	B2	2015年 9月 8日	TW	201126509	A	2011年 8月 1日
				EP	2427881	A1	2012年 3月 14日
				US	2011123031	A1	2011年 5月 26日
				TW	I508058	B	2015年 11月 11日
				GB	2470059	A	2010年 11月 10日
				WO	2010128386	A1	2010年 11月 11日
US	2011123031	A1	2011年 5月 26日	TW	201126509	A	2011年 8月 1日
				US	9129593	B2	2015年 9月 8日
				EP	2427881	A1	2012年 3月 14日
				TW	I508058	B	2015年 11月 11日
				GB	2470059	A	2010年 11月 10日
				WO	2010128386	A1	2010年 11月 11日

表 PCT/ISA/210 (同族专利附件) (2009年7月)