



(12) 发明专利

(10) 授权公告号 CN 102160113 B

(45) 授权公告日 2013. 05. 08

(21) 申请号 200880131232. 3

(22) 申请日 2008. 08. 11

(85) PCT申请进入国家阶段日
2011. 03. 22

(86) PCT申请的申请数据
PCT/EP2008/060536 2008. 08. 11

(87) PCT申请的公布数据
W02010/017833 EN 2010. 02. 18

(73) 专利权人 诺基亚公司
地址 芬兰埃斯波

(72) 发明人 M·T·维勒莫 M·T·塔米

(74) 专利代理机构 北京市金杜律师事务所
11256

代理人 吴立明

(51) Int. Cl.
G10L 19/008 (2013. 01)

(56) 对比文件

WO 2004/072956 A1, 2004. 08. 26,
WO 2005/098825 A1, 2005. 10. 20,
CN 1714599 A, 2005. 12. 28,
EP 1821287 A1, 2007. 08. 22,
CN 2006/0190247 A1, 2006. 08. 24,

审查员 张飞弦

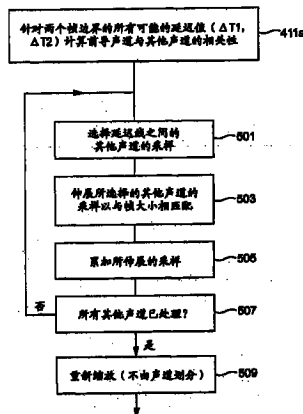
权利要求书4页 说明书15页 附图11页

(54) 发明名称

多声道音频编码器和解码器

(57) 摘要

一种装置,配置用于:确定第一信号与第二信号之间的至少一个时间延迟;根据至少一个时间延迟从第二信号生成第三信号;以及将第一信号与第三信号组合以生成第四信号;将第一信号和第二信号划分为多个时间帧;针对每个时间帧确定与第一信号的时间帧的开始相关联的第一延迟以及与第一信号的时间帧的末尾相关联的第二时间延迟;在定义为开始于时间帧的开始与第一时间延迟的组合处,并且结束于时间帧的末尾与第二时间延迟的组合处的块中,从第二信号中选择至少一个采样;以及将所选择的至少一个采样伸展为等于第一帧的采样的数目。



1. 一种用于对信号进行编码的设备,包括:

用于将第一信号和第二信号的每一个划分为多个频率子带的装置,其中所述第一信号和所述第二信号的每一个被划分为多个时间帧;

用于确定与所述第一信号的所述多个频率子带中的频率子带相关联的时间信号的时间帧和与所述第二信号的所述多个频率子带中的对应频率子带相关联的时间信号的时间帧之间的时间延迟的装置;

用于根据所述时间延迟从与所述第二信号的所述对应频率子带相关联的时间信号的时间帧生成伸展时间信号的装置,其中与所述第二信号的所述对应频率子带相关联的时间信号的时间帧被伸展以适应与所述第一信号的所述对应频率子带相关联的时间信号的时间帧;以及

用于将所述第一信号的频带的时间信号的时间帧与所述伸展时间信号组合以生成所述频带的单音下混频信号的装置。

2. 根据权利要求1所述的设备,进一步包括用于使用以下至少一个来编码所述频带的下混频信号的装置:

MPEG-2AAC, 以及

MPEG-1 层 III (mp3)。

3. 根据权利要求1所述的设备,进一步包括用于将所述第一信号和所述第二信号划分为以下中的至少一个的装置:

多个非重叠时间帧;

多个重叠时间帧; 以及

多个加窗重叠时间帧。

4. 根据权利要求1所述的设备,进一步包括:用于针对每个时间帧确定和与所述第一信号的所述频率子带相关联的时间信号的时间帧的开始相关联的第一时间延迟, 以及和与所述第一信号的所述频率子带相关联的时间信号的时间帧的末尾相关联的第二时间延迟的装置。

5. 根据权利要求4所述的设备,其中与所述第一信号的所述频率子带相关联的时间信号和与所述第二信号的所述对应频率子带相关联的时间信号包括多个采样, 并且所述设备进一步包括:

用于在定义为开始于与所述第二信号的所述对应频率子带相关联的时间信号的时间帧的开始与所述第一时间延迟的组合处, 并且结束于与所述第二信号的所述对应频率子带相关联的时间信号的时间帧的末尾与所述第二时间延迟的组合处的块中, 从与所述第二信号的所述对应频率子带相关联的时间信号的时间帧中选择至少一个采样的装置; 以及

用于将所选择的至少一个采样伸展为等于与所述第一信号的所述频率子带相关联的时间信号的时间帧的采样的数目的装置。

6. 根据权利要求1所述的设备,其中用于确定时间延迟的装置包括:

用于针对与所述第一信号的所述频率子带相关联的时间信号和与所述第二信号的所述对应频率子带相关联的时间信号相关生成相关值的装置; 以及

用于选择具有最高相关值的时间值的装置。

7. 根据权利要求1到6所述的设备,进一步包括用于生成进一步信号的装置,其中所述

进一步信号包括以下中的至少一个：

时间延迟值；以及

与所述第一信号的所述多个频率子带中的频率子带相关联的时间信号的时间帧和与所述第二信号的所述多个频率子带中的对应频率子带相关联的时间信号的时间帧之间的能量差。

8. 根据权利要求 7 所述的设备,进一步包括用于将所述进一步信号与所述单音下混频信号复用以生成编码音频信号的装置。

9. 一种用于对信号进行解码的设备,包括：

用于将第一信号划分为至少第一部分和第二部分的装置；

用于解码所述第一部分以形成第一声道音频信号的装置；

用于从根据第二部分修改的所述第一声道音频信号生成第二声道音频信号的装置,其中所述第二部分包括时间延迟值,并且所述设备包括用于根据对所述第一声道音频信号的时间延迟值、通过应用至少一个时移来生成所述第二声道音频信号的装置,其中所述第二部分包括至少一个第一时间延迟值和至少一个第二时间延迟值,所述第一声道音频信号包括从帧开始时间的第一采样到帧结束时间的末尾采样定义的至少一个帧,并且所述设备进一步包括：

用于在由所述第一声道音频信号的帧开始时间和第一时间延迟值所定义的时刻,将所述第一声道音频信号帧的第一采样复制到所述第二声道音频信号的装置；以及

用于在由所述第一声道音频信号的帧结束时间和第二时间延迟值所定义的时刻,将所述第一声道音频信号的末尾采样复制到所述第二声道音频信号的装置。

10. 根据权利要求 9 所述的设备,其中所述第二部分还包括能量差值,并且其中所述设备进一步包括用于通过根据所述能量差值向所述第一声道音频信号应用增益来生成所述第二声道音频信号的装置。

11. 根据权利要求 9 所述的设备,进一步包括用于将所述第一声道音频信号划分为至少两个频带的装置,其中所述第二声道音频信号的生成通过修改所述第一声道音频信号的每个频带。

12. 根据权利要求 9 所述的设备,进一步包括：

用于复制第一采样时刻与末尾采样时刻之间的任何其他第一声道音频信号帧采样的装置。

13. 根据权利要求 12 所述的设备,进一步包括：

用于重新采样所述第二声道音频信号以与所述第一声道音频信号同步的装置。

14. 一种用于对信号进行编码的方法,包括：

将第一信号和第二信号的每一个划分为多个频率子带,其中所述第一信号和所述第二信号的每一个被划分为多个时间帧；

确定与所述第一信号的所述多个频率子带中的频率子带相关联的时间信号的时间帧和与所述第二信号的所述多个频率子带中的对应频率子带相关联的时间信号的时间帧之间的时间延迟；

根据所述时间延迟从与所述第二信号的所述对应频率子带相关联的时间信号的时间帧生成伸展时间信号,其中与所述第二信号的所述对应频率子带相关联的时间信号的时间

帧被伸展以适应与所述第一信号的所述对应频率子带相关联的时间信号的时间帧；以及

将所述第一信号的频带的时间信号的时间帧与所述伸展时间信号组合以生成所述频带的单音下混频信号。

15. 根据权利要求 14 所述的方法,进一步包括使用以下至少一个来编码所述频带的下混频信号:

MPEG-2AAC, 以及

MPEG-1 层 III (mp3)。

16. 根据权利要求 14 所述的方法,进一步包括:将所述第一信号和所述第二信号划分为以下中的至少一个:

多个非重叠时间帧;

多个重叠时间帧;以及

多个加窗重叠时间帧。

17. 根据权利要求 14 所述的方法,进一步包括:针对每个时间帧确定和与所述第一信号的所述频带子带相关联的时间信号的时间帧的开始相关联的第一时间延迟以及和与所述第一信号的所述频率子带相关联的时间信号的时间帧的末尾相关联的第二时间延迟。

18. 根据权利要求 17 所述的方法,其中与所述第一信号的所述频率子带相关联的时间信号和与所述第二信号的所述对应频率子带相关联的时间信号包括多个采样,并且所述方法进一步包括:

在定义为开始于与所述第二信号的所述对应频率子带相关联的时间信号的时间帧的开始与所述第一时间延迟的组合处,并且结束于与所述第二信号的所述对应频率子带相关联的时间信号的时间帧的末尾与所述第二时间延迟的组合处的块中,从与所述第二信号的所述对应频率子带相关联的时间信号的时间帧中选择至少一个采样;以及

将所选择的至少一个采样伸展为等于与所述第一信号的所述频率子带相关联的时间信号的时间帧的采样的数目。

19. 根据权利要求 14 所述的方法,其中确定所述时间延迟包括:

针对与所述第一信号的所述频率子带相关联的时间信号和与所述第二信号的所述对应频率子带相关联的时间信号相关生成相关值;以及

选择具有最高相关值的时间值。

20. 根据权利要求 14 到 19 中任意一项所述的方法,进一步包括生成进一步信号,其中所述进一步信号包括以下中的至少一个:

时间延迟值;以及

与所述第一信号的所述多个频率子带中的频率子带相关联的时间信号的时间帧和与所述第二信号的所述多个频率子带中的对应频率子带相关联的时间信号的时间帧之间的能量差。

21. 根据权利要求 20 所述的方法,进一步包括:

将所述进一步信号与所述单音下混频信号复用以生成编码音频信号。

22. 一种用于对信号进行解码的方法,包括:

将第一信号划分为至少第一部分和第二部分;

解码所述第一部分以形成第一声道音频信号;以及

从根据所述第二部分修改的第一声道音频信号生成第二声道音频信号,其中所述第二部分包括时间延迟值,并且其中通过应用至少一个时移来生成所述第二声道音频信号取决于对所述第一声道音频信号的时间延迟值,其中所述第二部分包括至少一个第一时间延迟值和至少一个第二时间延迟值,所述第一声道音频信号包括从帧开始时间的第一采样到帧结束时间的末尾采样定义的至少一个帧,并且所述方法进一步包括:

在由所述第一声道音频信号的帧开始时间和第一时间延迟值所定义的时刻,将所述第一声道音频信号帧的第一采样复制到所述第二声道音频信号;以及

在由所述第一声道音频信号的帧结束时间和第二时间延迟值所定义的时刻,将所述第一声道音频信号的末尾采样复制到所述第二声道音频信号。

23. 根据权利要求 22 所述的方法,其中所述第二部分还包括能量差值,并且其中所述方法进一步包括:通过根据所述能量差值向所述第一声道音频信号应用增益来生成所述第二声道音频信号。

24. 根据权利要求 22 所述的方法,进一步包括:将所述第一声道音频信号划分为至少两个频带,其中生成所述第二声道音频信号包括修改所述第一声道音频信号的每个频带。

25. 根据权利要求 22 所述的方法,进一步包括:复制第一采样时刻与末尾采样时刻之间的任何其他第一声道音频信号帧采样;以及

重新采样所述第二声道音频信号以与所述第一声道音频信号同步。

多声道音频编码器和解码器

技术领域

[0001] 本发明涉及用于编码和解码的装置,并且具体地但不仅限于用于音频和语音信号的编码和解码。

背景技术

[0002] 空间音频处理是从音频源发出的音频信号通过不同传播路径到达聆听者的左耳和右耳的效果。由于这种效果,左耳处的信号与到达右耳的相应信号相比通常具有不同的到达时间和信号水平。时间和信号水平之间的差异取决于音频信号为了分别到达左耳和右耳所经过的路径的差异。聆听者的大脑继而解译这些差异,以便给出所接收的音频信号由位于相对于聆听者的特定距离和方向处的音频源生成的感觉。

[0003] 因此,可以将听觉场景视为同时收听位于相对于聆听者的不同位置处的一个或多个音频源生成的音频信号的合成串音。

[0004] 人类大脑可以处理双声道输入信号,以便确定声源的位置和方向的简单事实可以用于编码和合成听觉场景。空间听觉编码的一种典型方法由此可以通过有目的地修改来自一个或多个不同源(声道)的音频信号来试图建模音频场景的显著特征。这可以在被定义为左音频信号和右音频信号的耳机组中使用。左音频信号和右音频信号可以统称为双声道信号。继而可以生成所产生的双声道信号,使得其给出改变位于相对于聆听者的不同位置处的音频源的感觉。双声道信号与立体声信号的区别在于两个方面。首先,双声道信号已经并入了左声道与右声道之间的时间差,其次,双声道信号采用“头影效应”(其中,建模了特定频带的音量的减小)。

[0005] 最近,已经结合多声道音频再现而使用空间音频技术。多声道音频再现的目的是提供对包括多个独立音频声道或者声源的多声道音频信号的有效编码。多声道音频信号的编码的最近解决方案集中于参量立体声(PS)和双声道线索编码(BCC)。BCC通常通过将输入音频信号下混频为单(“含量”)声道或者传送该“含量”信号的较小数目的声道来编码多声道音频信号。同时,从输入声道中提取最显著的声道间线索并且将其编码为边信息,该最显著的声道间线索另外也称为空间线索,其描述多声道声像或者音频场景。含量信号和边信息二者形成经编码的参数集,其继而可以作为通信链的一部分来传输,或者存储在存储和转发类型设备中。BCC技术的多数实现通常采用低比特率音频编码方案来进一步编码含量信号。最终,BCC解码器根据所传输或者存储的含量信号和空间线索信息来生成多声道输出信号。通常,使用低比特率听觉音频编码技术(诸如AAC)来附加地编码空间音频编码系统中采用的下混频信号,以进一步减少所需的比特率。

[0006] 迄今为止,其中存在不止两个源的多声道音频编码仅用于家庭影院应用,其中带宽通常不视为主要限制。然而,多声道音频编码可以在多种移动设备上的新出现的多麦克风实现中使用,以帮助挖掘这些多麦克风技术的全部潜能。例如,多麦克风系统可以用于例如通过使得音频在接收机处缩放而在较差的音频环境中产生较好的信噪比,其中接收机具有集中于所接收的信号的特定期源或者方向的能力。这一集中继而可以根据需要由接收机改

进的源而改变。

[0007] 上文提到的多声道系统具有以下固有问题,即,N声道/麦克风源系统在被直接编码时产生的比特流大约需要单声道带宽 N 倍的带宽。

[0008] 该多声道带宽需求对于无线通信系统通常是禁止的。

[0009] 已知通过假设每个声道记录了相同的源信号,但是具有不同的时间延迟和频率相关放大特性,可以建模多声道/多源系统。在用于减少带宽需求的一些解决方案(诸如上文描述的双声道编码解决方案)中,已证实可以将 N 个声道组合到单个声道中,该单个声道是水平(强度)和时间对齐的。然而,这产生了一个问题,即水平和时间对齐针对不同的时间和频率元素是不同的。另外,通常存在若干源信号占用相同的时间-频率位置,每个源信号需要不同的时间和水平对齐。

[0010] 已经提出的一种单独解决方案已经解决了将所有音频源(换言之,随后由麦克风检测的音频信号的原始源)与信号分离并且建模由麦克风定义的原始源和空间的方向和音响效果的问题。然而,这在计算上较困难,并且需要大量的处理功率。另外,这一解决方案需要对所有原始源进行单独编码,并且原始源的数目可能超过原始声道的数目。换言之,所建模的原始源的数目可能大于用于记录音频环境的麦克风声道的数目。

[0011] 因此,目前系统通常仅将多声道系统编码为单个声道或者较小数目的声道,并且将其他声道编码为与最近声道的水平或者强度差值。例如,在两个(左和右)声道系统中,通常,通过平均左声道和右声道来创建单个单声道,并且继而对两个声道系统中的左声道和右声道二者的频带中的信号能量水平进行量化和编码,并且进行存储/发送至接收机。在接收机/解码器处,将单信号复制到两个声道,并且将左声道和右声道的信号水平设置为与两个重新创建的声道二者中每个频带中所接收的能量信息相匹配。

[0012] 这一类型的系统由于编码而产生了并非最优的声像,并且无法产生多声道系统可以产生的音频深度。

发明内容

[0013] 本发明出于以下考虑,即通过将声道之间的时间差以及水平差纳入考虑而期望利用比先前允许的更高的质量来编码多声道信号。

[0014] 本发明的实施方式旨在解决以上问题。

[0015] 根据本发明的第一方面,提供了一种装置,配置用于:确定第一信号与第二信号之间的至少一个时间延迟;根据至少一个时间延迟从第二信号生成第三信号;以及将第一信号与第三信号组合以生成第四信号。

[0016] 由此,本发明的实施方式可以编码音频信号并且产生具有较好定义的声道分离而不需要分离的声道编码。

[0017] 该装置可以进一步配置用于使用以下至少一个来编码第四信号:MPEG-2AAC 和 MPEG-1 层 III(mp3)。

[0018] 该装置可以进一步配置用于将第一信号和第二信号划分为多个频带,并且其中优选地针对每个频带确定至少一个时间延迟。

[0019] 该装置可以进一步配置用于将第一信号和第二信号划分为多个时间帧,并且其中针对每个时间帧确定至少一个时间延迟。

[0020] 该装置可以进一步配置用于将第一信号和第二信号划分为以下中的至少一个：多个非重叠时间帧；多个重叠时间帧；以及多个加窗重叠时间帧。

[0021] 该装置可以进一步配置用于针对每个时间帧确定与第一信号的时间帧的开始相关联的第一时间延迟，以及与第一信号的时间帧的末尾相关联的第二时间延迟。

[0022] 第一帧和第二帧可以包括多个采样，并且该装置可以进一步配置用于：在定义为开始于时间帧的开始与第一时间延迟的组合处，并且结束于时间帧的末尾与第二时间延迟的组合处的块中，从第二信号中选择至少一个采样；以及将所选择的至少一个采样伸展为等于第一帧的采样的数目。

[0023] 该装置可以进一步配置用于通过以下来确定至少一个时间延迟：针对第一信号与第二信号相关生成相关值；以及选择具有最高相关值的时间值。

[0024] 该装置可以进一步配置用于生成第五信号，其中该第五信号包括以下中的至少一个：至少一个时间延迟值；以及第一信号与第二信号之间的能量差。

[0025] 该装置可以进一步配置用于将第五信号与第四信号复用以生成编码音频信号。

[0026] 根据本发明的第二方面，提供了一种装置，配置用于：将第一信号划分为至少第一部分和第二部分；解码第一部分以形成第一声道音频信号；以及从根据第二部分修改的第一声道音频信号生成第二声道音频信号，其中第二部分包括时间延迟值，并且该装置配置用于根据对第一声道音频信号的时间延迟值、通过应用至少一个时移来生成第二声道音频信号。

[0027] 第二部分还可以包括能量差值，并且其中该装置进一步配置用于通过根据能量差值向第一声道音频信号应用增益来生成第二声道音频信号。

[0028] 该装置可以进一步配置用于将第一声道音频信号划分为至少两个频带，其中第二声道音频信号的生成优选地修改第一声道音频信号的每个频带。

[0029] 第二部分可以包括至少一个第一时间延迟值和至少一个第二时间延迟值，第一声道音频信号可以包括从帧开始时间的第一采样到帧结束时间的末尾采样定义的至少一个帧，并且该装置优选地进一步配置用于：在由第一声道音频信号的帧开始时间和第一时间延迟值所定义的时刻，将第一声道音频信号帧的第一采样复制到第二声道音频信号；以及在第一声道音频信号的帧结束时间和第二时间延迟值定义的时刻，将第一声道音频信号的末尾采样复制到第二声道音频信号。

[0030] 该装置可以进一步配置用于复制第一采样时刻与末尾采样时刻之间的任何其他第一声道音频信号帧采样。

[0031] 该装置可以进一步配置用于重新采样第二声道音频信号以与第一声道音频信号同步。

[0032] 一种电子设备可以包括如上所述的装置。

[0033] 一种芯片集可以包括如上所述的装置。

[0034] 一种编码器可以包括如上所述的装置。

[0035] 一种解码器可以包括如上所述的装置。

[0036] 根据本发明的第三方面，提供了一种方法，包括：确定第一信号与第二信号之间的至少一个时间延迟；根据至少一个时间延迟、从第二信号生成第三信号；以及组合第一信号和第三信号以生成第四信号。

[0037] 该方法可以进一步包括：使用以下至少一个来编码第四信号：MPEG-2AAC，以及 MPEG-1 层 III (mp3)。

[0038] 该方法可以进一步包括：将第一信号和第二信号划分为多个频带，并且针对每个频带确定至少一个时间延迟。

[0039] 该方法可以进一步包括：将第一信号和第二信号划分为多个时间帧，并且针对每个时间帧确定至少一个时间延迟。

[0040] 该方法可以进一步包括：将第一信号和第二信号划分为以下中的至少一个：多个非重叠时间帧；多个重叠时间帧；以及多个加窗重叠时间帧。

[0041] 该方法可以进一步包括：针对每个时间帧确定与第一信号的时间帧的开始相关联的第一时间延迟以及与第一信号的时间帧的末尾相关联的第二时间延迟。

[0042] 第一帧和第二帧可以包括多个采样，并且该方法可以进一步包括：在定义为开始于时间帧的开始与第一时间延迟的组合处，并且结束于时间帧的末尾与第二时间延迟的组合处的块中，从第二信号中选择至少一个采样；以及将所选择的至少一个采样伸展为等于第一帧的采样的数目。

[0043] 确定至少一个时间延迟可以包括：针对第一信号与第二信号相关生成相关值；以及选择具有最高相关值的时间值。

[0044] 该方法可以进一步包括生成第五信号，其中该第五信号包括以下中的至少一个：至少一个时间延迟值；以及第一信号与第二信号之间的能量差。

[0045] 该方法可以进一步包括：将第五信号与第四信号复用以生成编码音频信号。

[0046] 根据本发明的第四方面，提供了一种方法，包括：将第一信号划分为至少第一部分和第二部分；解码第一部分以形成第一声道音频信号；以及从根据第二部分修改的第一声道音频信号生成第二声道音频信号，其中第二部分包括时间延迟值，并且其中通过应用至少一个时移来生成第二声道音频信号取决于对第一声道音频信号的时间延迟值。

[0047] 第二部分还可以包括能量差值，并且其中该方法可以进一步包括：通过根据能量差值向第一声道音频信号应用增益来生成第二声道音频信号。

[0048] 该方法可以进一步包括：将第一声道音频信号划分为至少两个频带，其中生成第二声道音频信号可以包括修改第一声道音频信号的每个频带。

[0049] 第二部分可以包括至少一个第一时间延迟值和至少一个第二时间延迟值，第一声道音频信号可以包括从帧开始时间的第一采样到帧结束时间的末尾采样定义的至少一个帧，并且该方法可以进一步包括：在由第一声道音频信号的帧开始时间和第一时间延迟值所定义的时刻，将第一声道音频信号帧的第一采样复制到第二声道音频信号；以及在第一声道音频信号的帧结束时间和第二时间延迟值定义的时刻，将第一声道音频信号的末尾采样复制到第二声道音频信号。

[0050] 该方法可以进一步包括：复制第一采样时刻与末尾采样时刻之间的任何其他第一声道音频信号帧采样。

[0051] 该方法可以进一步包括：重新采样第二声道音频信号以与第一声道音频信号同步。

[0052] 根据本发明的第五方面，提供了一种计算机程序产品，配置用于执行包括以下的方法：确定第一信号与第二信号之间的至少一个时间延迟；根据至少一个时间延迟、从第

二信号生成第三信号；以及组合第一信号和第三信号以生成第四信号。

[0053] 根据本发明的第六方面，提供了一种计算机程序产品，配置用于执行包括以下的方法：将第一信号划分为至少第一部分和第二部分；解码第一部分以形成第一声道音频信号；以及从根据第二部分修改的第一声道音频信号生成第二声道音频信号，其中第二部分包括时间延迟值；并且其中通过应用至少一个时移来生成第二声道音频信号取决于对第一声道音频信号的时间延迟值。

[0054] 根据本发明的第七方面，提供了一种装置，包括：处理装置，用于确定第一信号与第二信号之间的至少一个时间延迟；信号处理装置，用于根据至少一个时间延迟从第二信号生成第三信号；以及组合装置，用于组合第一信号和第三信号以生成第四信号。

[0055] 根据本发明的第八方面，提供了一种装置，包括：处理装置，用于将第一信号划分为至少第一部分和第二部分；解码装置，用于解码第一部分以形成第一声道音频信号；以及信号处理装置，用于从根据第二部分修改的第一声道音频信号生成第二声道音频信号，其中第二部分包括时间延迟值；并且其中信号处理装置配置用于根据对第一声道音频信号的时间延迟值、通过应用至少一个时移来生成第二声道音频信号。

附图说明

[0056] 为了更好地理解本发明，现在通过示例的方式参考附图，其中：

[0057] 图 1 示意性地示出了采用本发明的实施方式的电子设备；

[0058] 图 2 示意性地示出了采用本发明的实施方式的音频编解码器系统；

[0059] 图 3 示意性地示出了图 2 所示的本发明的实施方式中采用的音频编码器；

[0060] 图 4 示出了说明编码多声道信号的本发明的实施方式的操作的流程图；

[0061] 图 5 更加详细地示出了从图 4 所示的多个多声道带块生成下混频信号的操作；

[0062] 图 6 示出了根据本发明的实施方式正在进行编码的信号的示意图；

[0063] 图 7 示意性地示出了根据本发明的实施方式的采样伸展；

[0064] 图 8 示出了本发明的实施方式中采用的帧窗；

[0065] 图 9 示出了根据本发明的实施方式的加窗（重叠和非重叠）以及非重叠组合之间的差异；

[0066] 图 10 示意性地示出了根据本发明的实施方式的在解码器中将单音信号解码到声道中；

[0067] 图 11 示意性地示出了利用重叠和非重叠窗对单音声道的解码；

[0068] 图 12 示出了根据本发明的实施方式的解码器；

[0069] 图 13 示意性地示出了根据本发明的实施方式的声道化合成器；以及

[0070] 图 14 示出了详细描述根据本发明的实施方式的解码器的操作的流程图。

具体实施方式

[0071] 下面进一步详细地描述为音频编解码器提供增强的编码效率和信号保真度适当装置和可能的机制。就这一点而言，首先参考图 1，其示出了示例性装置或者电子设备 10 的示意框图，其可以并入根据本发明的实施方式的编解码器。

[0072] 电子设备 10 例如可以是无线通信系统的移动终端或者用户设备。

[0073] 电子设备 10 包括麦克风 11,其经由模拟数字转换器 14 链接到处理器 21。处理器 21 进一步经由数字模拟转换器 32 链接到扬声器 33。处理器 21 进一步链接至收发机 (TX/RX) 13、用户接口 (UI) 15 和存储器 22。

[0074] 处理器 21 可以配置用于执行各种程序代码。所实现的程序代码可以包括编码代码例程。所实现的程序代码 23 可以进一步包括音频解码代码。所实现的程序代码 23 例如可以存储在存储器 22 中,以便由处理器 21 在需要时获取。存储器 22 可以进一步提供部分 24 以用于存储数据,例如已经根据本发明进行编码的数据。

[0075] 在本发明的实施方式中,编码和解码代码可以在硬件或者固件中实现。

[0076] 用户接口 15 可以使得用户能够例如经由小键盘向电子设备 10 输入命令,和 / 或例如经由显示器从电子设备 10 获得信息。收发机 13 使得能够例如经由无线通信网络与其他电子设备通信。在本发明的一些实施方式中,收发机 13 可以配置用于通过有线连接与其他电子设备通信。

[0077] 还可以理解,电子设备 10 的结构可以按照多种方式进行补充和改变。

[0078] 电子设备 10 的用户可以使用麦克风 11 来输入语音,该语音将要传输至一些其他电子设备,或者将要存储在存储器 22 的数据段 24 中。为此,用户已经经由用户接口 15 激活了相应的应用。该应用可以由处理器 21 运行,使得处理器 21 执行存储在存储器 22 中的编码代码。

[0079] 模拟数字转换器 14 可以将输入的模拟音频信号转换为数字音频信号,并且向处理器 21 提供数字音频信号。

[0080] 处理器 21 继而可以按照相同的方式处理数字音频信号,如参考以下描述所描述的。

[0081] 将产生的比特流提供至收发机 13,以便向另一电子设备传输。备选地,编码的数据可以存储在存储器 22 的数据段 24 中,例如用于随后的传输或者由相同的电子设备 10 随后呈现。

[0082] 电子设备 10 也可以经由收发机 13、从另一电子设备接收具有相应的编码数据的比特流。在这种情况下,处理器 21 可以执行存储在存储器 22 中的解码程序代码。处理器 21 因此可以解码所接收的数据,并且将解码的数据提供至数字模拟转换器 32。数字模拟转换器 32 可以将数字解码数据转换为模拟音频数据,并且将模拟信号输出至扬声器 33。也可以由用户已经经由用户接口 15 调用的应用来触发解码程序代码的执行。

[0083] 所接收的编码数据也可以存储在存储器 22 的数据段 24 中,而不是经由扬声器 33 立即呈现,例如,用于使得能够随后呈现,或者向又一电子设备转发。

[0084] 在本发明的一些实施方式中,可以由耳机组对扬声器 33 进行补充或者将其替换,该耳机组可以与电子设备 10 或者装置例如通过蓝牙协议经由收发机 13 无线地通信,或者使用传统的有线连接来通信。

[0085] 可以理解,图 3、图 12 和图 13 中描述的示意结构以及图 4、图 5 和图 14 中的方法步骤仅呈现了图 1 所示的电子设备实现的完整音频编解码器的操作的一部分。

[0086] 图 2 示出了本发明的实施方式采用的音频编解码器的一般操作。一般音频编码 / 解码系统包括编码器和解码器,如图 2 示意性示出的。示出了具有编码器 104、存储或者媒体声道 106 和解码器 108 的系统 102。

[0087] 编码器 104 压缩产生比特流 112 的输入音频信号 110, 其可以被存储或者通过媒体声道 106 传输。比特流 112 可以在解码器 108 内接收。解码器 108 解压缩比特流 112, 并且产生输出音频信号 114。与输入信号 110 有关的比特流 112 的比特率以及输出音频信号 114 的质量是主要特征, 其定义了编码系统 102 的性能。

[0088] 图 3 示意性地示出了根据本发明第一实施方式的编码器 104。编码器 104 绘出为包括被划分为 N 个声道 $\{C_1, C_2, \dots, C_N\}$ 的输入 302。可以理解, 输入 302 可以布置为接收 N 个声道的音频信号, 或者备选地, 接收来自 N 个单独音频源的 N 个音频信号, 其中 N 是等于或者大于 2 的整数。

[0089] 在图 4 中通过步骤 401 示出了 N 个声道的接收。

[0090] 在下文描述的实施方式中, 并行处理每个声道。然而, 本领域技术人员可以理解, 根据特定实施方式以及相关联的并行 / 串行处理成本 / 益处分析, 可以串行地或者部分串行地和部分并行地处理每个声道。

[0091] 由滤波器组 301 接收 N 个声道。滤波器组 301 包括 N 个滤波器组元件 303 中的多个。每个滤波器组元件 303 接收声道之一并且输出每个声道的一系列频带分量。如图 3 中所看到的, 用于第一声道 C_1 的滤波器组元件是滤波器组元件 $FB_1 303_1$, 其输出 B 个声道带 C_1^1 到 C_1^B 。类似地, 滤波器组元件 $FB_N 303_N$ 输出用于第 N 个声道的一系列 B 个带分量 C_N^1 到 C_N^B 。这些声道中每一个的 B 个带从滤波器组 301 输出, 并且被传送至分隔器和加窗器 305。

[0092] 在本发明的实施方式中, 滤波器组可以是不一致的。在不一致的滤波器组中, 带不一致地分布。例如, 在一些实施方式中, 带针对较低频率可能较窄, 而针对高频率可能较宽。在本发明的一些实施方式中, 带可以重叠。

[0093] 在图 4 中通过步骤 403 示出了向每个声道应用滤波器组以生成针对每个声道的带。

[0094] 分隔器和加窗器 305 接收每个声道带采样值, 并且将声道的每个带分量的采样划分为采样值的块 (另外称为帧)。从分隔器和加窗器向单音块编码器 307 输出这些块或者帧。

[0095] 在本发明的一些实施方式中, 块或者帧在时间上重叠。在这些实施方式中, 可以应用加窗函数, 以使得具有相邻块或者帧的任何重叠部分合计值为 1。

[0096] 图 8 示出了加窗函数的一个示例, 并且可以根据以下公式来数学地描述。

$$[0097] \quad win_tmp = \left[\sin\left(2\pi \frac{\frac{1}{2} + k}{wtl} - \frac{\pi}{2}\right) + 1 \right] / 2, k = 0, \dots, wtl - 1$$

[0098]

$$win(k) = \begin{cases} 0 & , & k = 0, \dots, zl \\ win_tmp(k - (zl + 1)) & , & k = zl + 1, \dots, zl + wtl \\ 1 & , & k = zl + wtl, \dots, wl/2 \\ 1 & , & wl/2 + 1, \dots, wl/2 + ol \\ win_tmp(wl - zl - 1 - (k - (wl/2 + ol + 1))) & , & k = wl/2 + ol + 1, \dots, wl - zl - 1 \\ 0 & , & k = wl - zl, \dots, wl - 1 \end{cases}$$

[0099] 其中 w_{t1} 是窗的正弦部分的长度, z_1 是窗中的前导零的长度, 而 o_1 是窗中间的 1 的长度的半。为了使加窗重叠合计为 1, 以下等式必须成立:

$$[0100] \quad \begin{cases} z_1 + w_{t1} + o_1 = \frac{\text{length}(\text{win})}{2} \\ z_1 = o_1 \end{cases}$$

[0101] 加窗由此使得帧或者块之间的任何重叠能够在加和到一起时等于 1。另外, 加窗使得能够执行后续处理, 其中存在块之间的平滑过渡。

[0102] 然而, 在本发明的一些实施方式中, 不对采样应用加窗, 并且分离器仅将采样划分为块或者帧。

[0103] 在本发明的其他实施方式中, 可以在应用滤波器组之前向信号应用分隔器和加窗器。换言之, 可以在滤波器组 301 之前采用分隔器和加窗器 305, 以使得初始对输入声道信号进行分隔和加窗, 并且继而在进行分隔和加窗之后, 将其馈送至滤波器组, 以生成信号的 B 个带的序列。

[0104] 在图 4 中通过步骤 405 示出了向每个声道的每个带应用分隔和加窗以生成带块的步骤。

[0105] 向单音块编码器 307 传送带块。单块编码器从 N 个声道生成较小数目的下混频声道 N' 。在下文描述的示例中, N' 的值是 1, 然而, 在本发明的实施方式中, 编码器 104 可以生成不止一个下混频声道。在此类实施方式中, 实现将 N 个声道划分为 N' 个组的相似声道的附加步骤, 并且继而针对声道组中的每个组, 可以执行以下过程以产生针对每个声道组的单个单音下混频信号。可以通过将针对用于声道的带的至少一个的声道与相似值进行比较来实现相似声道的选择。然而, 在其他实施方式中, 可以通过任何方便的方式来实现将声道分组为 N' 个声道组。

[0106] 初始将声道 (或者针对特定组的声道) 的带的块 (帧) 分组为带块。换言之, 现在根据音频信号发生的频带来划分音频信号, 而不是根据声道号来划分。

[0107] 在图 4 中通过步骤 407 示出了分组带块的操作。

[0108] 每个带块被馈送至针对该带的前导声道选择器 309 中。由此, 针对第一带, 将声道的第一带 C_x^1 的所有块输入至带 1 前导声道选择器 309_1 , 并且将声道的第 B 个带 C_x^B 输入至带 B 前导声道选择器 309_B 。将其他带信号数据传输至图 3 中未示出的相应的带前导声道选择器, 以便辅助示图的理解。

[0109] 每个带前导声道选择器 309 选择输入声道音频信号之一作为“前导”声道。在本发明的第一实施方式中, 前导声道是固定声道, 例如可以选择声道组的第一声道作为前导声道。在本发明的其他实施方式中, 前导声道可以是任何声道。可以通过以下方式向解码器 108 指示该固定声道选择: 将信息插入到传输中, 或者将信息与音频编码数据流一起编码, 或者在本发明的一些实施方式中, 可以将信息预先确定或者硬接线至编码器 / 解码器因此让二者得知, 而无需在编码 - 解码过程中明确地用信号传送该信息。

[0110] 在本发明的其他实施方式中, 带前导声道选择器 309 对前导声道的选择是动态的, 并且可以根据预定义标准逐块或者逐帧地选择。例如, 前导声道选择器 309 可以选择具有最高能量的声道作为前导声道。在其他实施方式中, 前导声道选择器可以根据心理学建模标准来选择声道。在本发明的其他实施方式中, 前导声道选择器 309 可以通过选择在

与组中的所有其他声道比较时平均具有最小延迟的声道来选择前导声道。换言之,前导声道选择器可以选择具有组中所有声道的最平均特性的声道。

[0111] 前导声道可以由 $C_i^b(i)$ 来表示。

[0112] 在本发明的一些实施方式中,例如,在仅存在两个声道的实施方式中,选择“虚拟”或者“虚构”声道作为前导声道可能更加有效。虚拟或者虚构前导声道不是从麦克风生成的声道或者接收的声道,而是视为这样的另一声道,其延迟平均是两个声道之间的一半,或者所有声道的平均,并且可以视为幅值为 0。

[0113] 在图 4 中通过步骤 409 示出了选择针对每个带块的前导声道的操作。

[0114] 另外,将每个带块传送至带估计器 311,以使得如图 3 所示,将声道组第一带音频信号数据传送至带 1 估计器 311₁,而将声道组第 B 带音频信号数据传送至带 B 估计器 311_B。

[0115] 针对每个带块声道音频信号的带估计器 311 计算或者确定所选择的前导声道 $C_i^b(i)$ (其可以是声道或者虚构声道)与其他声道之间的差异。在所选择的前导声道与其他声道之间所计算的差异的示例包括声道之间的延迟 ΔT 以及声道之间的能量水平 ΔE 。

[0116] 图 6 的部分 (a) 示出了所选择的前导声道 601 与另一声道 602 之间的延迟的计算或者确定,如 ΔT_1 和 ΔT_2 所示。

[0117] 所选择的前导声道 C1 601 与另一声道 C2 602 之间的帧的开始之间的延迟示为 ΔT_1 ,而所选择的前导声道 C1 601 与另一声道 C2 602 之间的帧的结尾之间的延迟示为 ΔT_2 。

[0118] 在本发明的一些实施方式中,延迟时段 ΔT_1 和 ΔT_2 的确定/计算可以通过以下来生成:执行第一声道 C1 601 与第二声道 C2 602 的帧的开始的采样值的窗之间的相关,并且标记具有最高相关值的相关延迟。在本发明的其他实施方式中,延迟时段的确定可以在频域中实现。

[0119] 在本发明的其他实施方式中,通过比较每个声道频率块以及跨单个帧的时域或者频域声道值来确定声道之间的能量差异。

[0120] 在本发明的其他实施方式中,可以确定所选择的前导声道与其他声道之间的差异的其他测量。

[0121] 在图 4 中通过步骤 411 示出了计算前导声道与其他带声道箱之间的差异。

[0122] 通过步骤 411a 示出了确定所选择的前导声道与至少一个其他声道之间的差异(在图 5 所示的示例中是延迟)这一操作。

[0123] 将带估计器 311 的输出传送至带单音下混频器 313 的输入。带单音下混频器 313 接收带差异值(例如针对该帧的声道(例如声道组)的带音频信号和延迟差异),并且生成针对带和帧的单音下混频信号。

[0124] 这在图 4 中由步骤 415 示出,并且参考图 5、图 6 和图 7 进行更加详细的描述。

[0125] 带单音下混频器 313 通过组合来自针对带和帧的声道中每一个的值来生成针对每个带的单音下混频信号。由此,带 1 单音下混频器 313₁ 接收带 1 声道和带 1 估计值,并且产生带 1 单音下混频信号。类似地,带 B 单音下混频器 313_B 接收带 B 声道和带 B 估计差值,并且产生带 B 单音下混频信号。

[0126] 在以下示例中,针对带 1 声道分量和差值生成单音下混频声道信号。然而,可以理

解,以下方法可以在带单音下混频器 313 中实现,以产生任何下混频信号。另外,以下示例描述了迭代过程,以生成针对声道的下混频信号,然而,本领域技术人员可以理解,可以使用并行操作或者结构,其中每个声道基本上同时处理,而不是每个声道单独进行。

[0127] 相关于针对特定其他声道的带和帧信息的单音下混频器使用来自带估计器 311 的延迟信息 ΔT_1 和 ΔT_2 ,以选择将要与前导声道采样组合的其他声道的采样。

[0128] 换言之,单音下混频器选择反映前导声道与当前正在处理的其他声道之间的延迟的延迟线之间的采样。

[0129] 在本发明的一些实施方式中,诸如在非加窗实施方式或者加窗重叠较小的实施方式中,可以选择来自相邻帧的采样以维持信号一致性并且减小失真 (artefact) 生成的可能性。在本发明的一些实施方式中,例如其中延迟超过帧采样限制并且不可能使用来自相邻帧的信息的实施方式中,单音下混频器 313 可以插入零样本采样。

[0130] 在图 5 中,通过步骤 501 示出了选择延迟线之间的采样的操作。

[0131] 单音下混频器 313 继而伸展所选择的采样以适应当前帧大小。可以理解,通过根据延迟值 ΔT_1 和 ΔT_2 选择来自当前其他声道的采样,所选择的当前其他声道中可能存在比前导声道带帧中的采样数目更少或者更多的采样。

[0132] 由此,例如,其中在对当前其他声道应用延迟线之后在其他声道中存在 R 个采样,并且在前导声道帧中存在 S 个采样,则需要对采样的数目进行对齐,以便允许采样值的简单组合下混频。

[0133] 在本发明的第一实施方式中,通过利用 S 因子对信号进行第一上采样,利用适当的低通或者全通滤波器对上采样的信号进行滤波,并且继而利用 R 因子对滤波结果进行下采样,从而伸展 R 个采样长度信号,以形成 S 个采样。

[0134] 该操作可以在图 7 中示出,其中针对该示例,所选择的前导声道帧中的采样的数目是 3, $S = 3$, 并且当前其他声道中的采样的数目是 4, $R = 4$ 。图 7(a) 示出了其他声道采样 701、703、705 和 707, 以及引入的上采样值。在图 7(a) 的示例中,在每个所选择的前导声道帧采样之后,插入另外两个零值采样。由此,在采样 701 之后,存在插入的零值采样 709 和 711, 在采样 703 之后,插入零值采样 713 和 715, 在采样 705 之后,插入零值采样 717 和 719, 在 707 之后,插入零值采样 721 和 723。

[0135] 图 7(b) 示出了对所选择并且上采样添加的采样进行低通滤波的结果,以使得添加的采样现在遵循所选择的前导声道采样的波形。

[0136] 在图 7(c) 中,通过因子 R 对信号进行下采样,其中在该示例中, $R = 4$ 。换言之,从第一采样以及继而每隔四个采样形成下采样的信号。换言之,选择第一、第五和第九采样并且去除其他采样。

[0137] 所产生的信号现在具有将要与所选择的声道带帧采样相组合的正确的采样数目。

[0138] 在本发明的其他实施方式中,信号的伸展可以通过在当前其他声道采样之间线性或者非线性地进行插值来实现。在本发明的其他实施方式中,可以使用上述两种方法的组合。在该混合实施方式中,通过小于 S 的因子对来自延迟线内当前其他声道的采样进行第一上采样,对上采样的采样值进行低通滤波以便使引入的采样值遵循当前其他声道采样,并且通过插值选择新的点。

[0139] 在图 5 的步骤 503 中示出了伸展当前其他声道的采样以与前导声道的帧大小相匹

配。

[0140] 单音下混频器 313 继而将经伸展的采样添加至当前累加总值，以生成新累加总值。在第一次迭代中，将当前累加总值定义为前导声道采样值，而针对每次其他后续迭代，当前累加总值是先前迭代新累加总值。

[0141] 在图 5 中通过步骤 505 示出了生成新累加总值。

[0142] 带单音下混频器 313 继而确定是否已经处理了所有的其他声道。该确定步骤在图 5 中示为步骤 507。如果已经处理了所有的其他声道，则操作通过关键步骤 509，否则操作以另一其他声道开始新的迭代以进行再处理，换言之，操作返回步骤 501。

[0143] 当已经处理了所有声道时，带单音下混频器 313 继而重新缩放累加采样值以生成每带值的平均采样值。换言之，带单音下混频器 313 将累加总值中的每个采样值除以声道的数目，以产生带单音下混频信号。重新缩放累加总值的操作在图 5 中由步骤 509 示出。

[0144] 每个带单音下混频器生成其自身的单音下混频信号。由此，如图 3 所示，带 1 单音下混频器 313_1 产生带 1 单音下混频信号 $M^1(i)$ ，而带 B 单音下混频器 303_b 产生带 B 单音下混频信号 $M^b(i)$ 。将单音下混频信号传送至单音块 315。

[0145] 图 6(b) 和图 6(c) 中示出了针对两个声道系统中的真实和虚拟选择声道生成单音下混频信号的示例。

[0146] 在图 6(b) 中，对两个声道 C1 和 C2 进行下混频以形成单音声道 M。在所选择的前导声道中（在图 6(b) 中是 C1 声道），示出了其中的一个带帧 603。其他声道 C2605 针对相关联的带帧具有延迟值 ΔT_1 和 ΔT_2 。

[0147] 按照以上示出的方法，带下混频器 313 将选择由 ΔT_1 和 ΔT_2 生成的两个延迟线之间的带帧的部分。带下混频器继而将伸展所选择的帧采样以与 C1 的帧大小相匹配。继而将针对 C2 的帧的伸展选择部分添加到帧 C1。在图 6(b) 所示的示例中，在添加帧之前实现缩放。换言之，在将帧值相加到一起之前，带下混频器将每个帧的值除以声道数目，其在该示例中是 2。

[0148] 参考图 6(c)，示出了带单音下混频器的操作的示例，其中所选择的前导声道是虚拟或者虚构前导声道。在该示例中，带帧虚拟声道具有延迟，该延迟是该示例的两个正常声道的带帧（第一声道 C1 带帧 607 和第二声道 C2 的相关联带帧 609）的一半。

[0149] 在该示例中，单音下混频器 313 选择位于由 $+ve \Delta T_1/2651$ 和 $\Delta T_2/2657$ 生成的延迟线内的第一声道 C1 帧的帧采样，并且选择位于由 $-ve \Delta T_1/2653$ 和 $-ve \Delta T_2/2655$ 生成的延迟线之间的第二声道 C2 的帧采样。

[0150] 单音下混频器 313 继而根据虚构或者虚拟前导声道之间的差异而将第一声道 C1 伸展负量（收缩），并且对经收缩的第一声道 C1 值进行重新缩放，在该示例中，这意味着单音下混频器 313 将经收缩的值除以 2。单音下混频器 313 对于第二声道 C2 609 类似地实现相似的过程，其中对帧采样进行伸展并且除以 2。单音下混频器 313 继而将修改的声道值进行组合以形成下混频的单音声道带帧 611。

[0151] 单音块 315 从每个带单音下混频器 313 接收单音下混频带帧信号，并且针对每个声道生成单个单音块信号。

[0152] 可以通过将来自每个单音下混频音频信号的采样相加到一起生成下混频单音块信号。在本发明的一些实施方式中，加权因子可以与每个带相关联，并且将其应用到每个

带单音下混频音频信号,以产生具有带加强或者均衡的单音信号。

[0153] 在图 4 中通过步骤 417 示出了组合带下混频信号以形成单个帧下混频信号的操作。

[0154] 单音块 315 继而可以将帧单音块音频信号输出至块处理器 317。块处理器 317 接收单音块 315 生成的、针对特定帧的所有频带的单音下混频信号,并且将帧进行组合以产生音频下混频信号。

[0155] 在图 4 中通过步骤 419 示出了组合信号的块的可选操作。

[0156] 在本发明的一些实施方式中,块处理器 317 不组合块/帧。

[0157] 在本发明的一些实施方式中,块处理器 317 另外使用已知的音频编解码器对每个帧或者组合的帧单音下混频信号的一部分执行音频编码过程。

[0158] 可以在本发明的实施方式中应用的音频编解码器过程的示例包括:MPEG-2AAC,也称为 ISO/IEC 13818-7:1997;或者 MPEG-1 层 III(mp3),也称为 ISO/IEC 11172-3。然而,可以使用任何适合的音频编解码器来编码单音下混频信号。

[0159] 本领域技术人员可以理解,根据重叠窗、非重叠窗或者信号的分隔的实现,可以按照不同的方式来编码单音声道。参考图 9,示出了以下示例:具有重叠窗的单音声道图 9(a)901,具有非重叠窗的单音声道图 9(b)903,以及存在没有任何加窗或者重叠的信号的信号的分隔的单音声道图 9(c)905。

[0160] 在本发明的实施方式中,当相邻帧之间不存在重叠时(如图 9(c)所示),或者当窗中的重叠合计为 1 时(例如通过使用图 8 所示的窗函数),可以通过利用标准常规的单音音频编解码器编码单音声道来实现编码,并且可以将产生的编码值传送至复用器 319。

[0161] 然而,在本发明的其他实施方式中,当单音声道具有非重叠窗时(如图 9(b)所示),或者当使用具有重叠窗的单音声道但是值相加不为 1 的时候,将帧彼此相继地放置以使得不存在重叠。在一些实施方式中,这由此生成较好质量的信号编码,因为不存在信号与不同延迟的混合。然而,应当注意,这些实施方式将创建要进行编码的更多采样。

[0162] 继而将音频单音编码信号传送至复用器 319。

[0163] 在图 4 中通过步骤 421 示出了编码单音声道的操作。

[0164] 另外,量化器 321 接收针对描述所选择的前导声道与其他声道之间的差异的每个带的每个块(帧)的差值,并且对差异执行量化以生成量化的差异输出,其被传送至复用器 319。在本发明的一些实施方式中,也可以对量化信号实现可变长度编码,其还可以辅助检错或者纠错过程。

[0165] 在图 4 中通过步骤 413 示出了实现不同值的量化的操作。

[0166] 复用器 319 接收编码单音声道信号以及经量化和编码的不同信号,并且对信号进行复用以形成编码音频信号比特流 112。

[0167] 在图 4 中通过步骤 423 示出了信号的复用以形成比特流。

[0168] 可以理解,通过对差异(例如强度和时差二者)进行编码,来自下混频声道的多声道成像效应比简单的强度差异以及先前使用的下混频声道方法更加显著,并且比所使用的非下混频多声道编码方法更加有效地进行编码。

[0169] 参考图 12 和图 13,示出了根据本发明的实施方式的解码器。相关于图 14 所示的流程图进一步描述此类解码器的操作。解码器 108 包括解复用器和解码器 1201,其接收经

编码的信号。解复用器和解码器 1201 可以从编码比特流 112 中分离单音编码音频信号（或者在编码不止一个单音声道的实施方式中是单音编码音频信号）和量化差值（例如所选择的前导声道之间的时间延迟以及强度差异分量）。

[0170] 虽然所示出和描述的本发明的实施方式仅具有单个单音音频流，但是可以理解，可以采用下文描述的装置和过程来生成不止一个下混频音频声道 - 其中针对每个下混频（或者单音）音频声道独立地采用下文描述的操作。

[0171] 在图 14 中通过步骤 1401 示出了比特流的接收和解复用。

[0172] 继而解复用器和解码器 1201 可以使用来自编码器 104 内使用的编解码器的解码器算法部分来解码单音声道音频信号。

[0173] 在图 14 中，通过步骤 1403 示出了解码信号的编码单音部分，以生成解码单音声道信号估计。

[0174] 继而将解码单音或者下混频声道信号 \hat{M} 传送至滤波器组 1203。

[0175] 接收单音（下混频）声道音频信号的滤波器组 1203 使用滤波器组 1203 来执行滤波，以生成或者将单音信号分离成与在编码器内使用的频带等效的频带。

[0176] 由此，滤波器组 1203 输出 B 个带，即下混频信号 \hat{M}^1 到 \hat{M}^B 。继而将这些下混频信号频带分量传送至帧格式化器 1205。

[0177] 在图 14 中通过步骤 1405 示出了将下混频音频信号滤波为带。

[0178] 帧格式化器 1205 从滤波器组 1203 接收带划分下混频音频信号，并且进一步根据帧，执行帧格式化过程，将划分为带的单音音频信号进行划分。帧划分在长度上通常与编码器中采用的相似。在本发明的一些实施方式中，帧格式化器就帧开始指示符（其可能已经在编码器中被插入到比特流中）来检查下混频音频信号，并且使用帧指示符将带划分下混频音频信号划分为帧。在本发明的其他实施方式中，帧格式化器 1205 可以通过采样的数目进行计数并且当已经到达预定采样数目时选择新的帧来将音频信号划分为帧。

[0179] 将下混频带的帧传送至声道合成器 1207。

[0180] 在图 14 中通过步骤 1407 示出了将带分成帧的操作。

[0181] 声道合成器 1207 可以从帧格式化器接收下混频音频信号的帧，并且另外从解复用器和解码器 1201 接收差异数据（延迟和强度差值）。

[0182] 声道合成器 1207 可以针对根据下混频音频声道的帧以及差异数据而重构的每个声道来合成帧。在图 13 中，更详细地示出了声道合成器的操作。

[0183] 如图 13 所示，声道合成器 1207 包括采样重新伸展器 1303，其接收针对每个带以及差异信息的下混频音频信号的帧，该差异信息例如可以是时间延迟 ΔT 和强度差异 ΔE 。

[0184] 采样重伸展器 1303 根据延迟信息，通过采样重新缩放或者“重新伸展”下混频音频信号来重新生成原始声道带帧的近似值。该过程可以视为类似于在编码器内实现的用于在编码期间伸展采样的过程，但是按照相反顺序使用因子。由此，使用图 7 所示的示例，其中在编码器中，将所选择的 4 个采样伸展至 3 个采样，而在解码器中，将来自解码器帧的 3 个采样重新伸展为形成 4 个采样。在本发明的实施方式中，这可以通过插值或者通过添加附加采样值并且滤波，以及继而在需要时丢弃采样，或者通过以上的组合来实现。

[0185] 在本发明的实施方式中，其中存在前导和尾部窗采样，延迟通常将不会延伸通过窗区域。例如，在 44.1 千赫采样系统中，延迟通常在 -25 到 +25 个采样之间。在本发明的

一些实施方式中,当采样选择器用于选择延伸超过当前帧或者窗的采样时,采样选择器提供附加零值采样。

[0186] 重新伸展器 1303 的输出由此针对每个合成的声道 (1 到 N) 产生表示频率块 (1 到 B) 的采样值的帧。继而将每个合成的声道频率块帧输入到带组合器 1305。

[0187] 重新伸展器的操作的示例可以在图 10 中示出。图 10 示出了下混频音频声道频带帧 1001 的帧。如图 10 所示,将下混频音频声道频带帧 1001 在不进行修改的情况下复制到第一声道频带帧 1003。换言之,第一声道 C1 是编码器中所选择的前导声道,并且由此具有为 0 的 ΔT_1 和 ΔT_2 值。

[0188] 重新伸展器从非零 ΔT_1 和 ΔT_2 值重新伸展下混频音频声道频带帧 1001 的帧,以形成第二声道 C2 频带帧 1005 的帧。

[0189] 在图 14 中通过步骤 1411 示出了根据延迟值来重新伸展所选择的采样的操作。

[0190] 带组合器 1305 接收重新伸展的下混频音频声道频带帧,并且组合所有的频带,以便产生针对第一声道的估计声道值 $\tilde{C}_1(i)$ 直到针对第 N 个合成声道的估计声道值 $\tilde{C}_N(i)$ 。

[0191] 在本发明的一些实施方式中,根据缩放因子来修改每个频带内的采样值,以均衡编码器中应用的加权因子。换言之,均衡在编码过程期间设置的重点。

[0192] 在图 14 中通过步骤 1413 示出了组合针对每个合成声道帧的频带的操作。

[0193] 另外,将每个声道帧的输出传送至水平调节器 1307。水平调节器 1307 根据差异强度值 ΔE 而向该值应用增益,以使得针对每个声道的输出水平与针对原始声道的每个帧的能量水平近似相同。

[0194] 在图 14 中通过步骤 1415 示出了调节针对每个合成声道帧的水平 (应用增益)。

[0195] 另外,将每个水平调节器 1307 的输出输入到帧重新组合器 1309。帧重新组合器组合针对每个声道的每个帧,以便产生针对每个合成声道的一致输出比特流。

[0196] 图 11 示出了帧组合的两个示例。在第一示例 1101 中,存在具有重叠窗的声道要进行组合,而在 1103 中,存在具有非重叠窗的声道要进行组合。可以通过简单地将重叠加和到一起生成这些值以产生估计声道音频信号。通过声道合成器 1207 来输出该估计声道信号。

[0197] 在本发明的一些实施方式中,在合成帧上实现的延迟可能在相邻帧之间突然改变,并且导致失真,其中采样值的组合也会突然改变。在本发明的一些实施方式中,帧重新组合器 1309 还包括中值滤波器,以辅助防止在组合的信号采样值中出现失真。在本发明的其他实施方式中,可以采用其他滤波配置,或者可以使用信号插值来防止失真。

[0198] 在图 14 中通过步骤 1417 示出了对帧进行组合以生成声道比特流。

[0199] 上文描述的本发明的实施方式在单独的编码器 104 和解码器 108 装置的方面描述了编解码器,以便辅助所涉及的过程的理解。然而,可以理解,装置、结构和操作可以实现为单个编码器-解码器装置/结构/操作。另外,在本发明的一些实施方式中,编码器和解码器可以共享某些和/或全部的共同元件。

[0200] 虽然上述示例描述了在电子设备 610 内的编解码器内操作的本发明的实施方式,但是可以理解,以下描述的本发明可以实现为任何可变速率/自适应速率音频 (或者语音) 编解码器的一部分。由此,例如,本发明的实施方式可以在音频编解码器中实现,该音频编解码器可以实现固定或者有线通信路径上的音频编码。

[0201] 由此,用户设备可以包括音频编解码器,诸如本发明的以上实施方式所述的那些。

[0202] 应当理解,术语用户设备旨在覆盖任何类型的无线用户设备,诸如移动电话、便携式数据处理设备或者便携式 web 浏览器。

[0203] 另外,公共陆地移动网络 (PLMN) 的元件也可以包括上文所述的音频编解码器。

[0204] 一般而言,本发明的各种实施方式可以在硬件或者专用电路、软件、逻辑或者其任意组合中实现。例如,某些方面可以在硬件中实现,而其他方面可以在固件或者软件中实现,固件或者软件可以由控制器、微处理器或者其他计算设备执行,但是本发明不限于此。虽然可以将本发明的各个方面示出和描述为框图、流程图,或者使用某些其他图形表示,但是可以理解,作为非限制性的示例,本文描述的这些框、装置、系统、技术或者方法可以在硬件、软件、固件、专用电路或者逻辑、通用硬件或者控制器或者其他计算设备或者其某些组合中实现。

[0205] 本发明的实施方式可以由移动设备的数据处理器(诸如在处理器实体中)可执行的计算机软件或者硬件,或者由软件和硬件的组合来实现。另外,在此方面,应当理解,图中所示的逻辑流程的任何框可以表示程序步骤或者互连的逻辑电路、框和功能,或者程序步骤与逻辑电路、框和功能的组合。

[0206] 存储器可以是适于本地技术环境的任何类型,并且可以使用任何适当的数据存储技术来实现,诸如基于半导体的存储器设备、磁存储器设备和系统、光存储器设备和系统、固定存储器和可移动存储器。数据处理器可以是适于本地技术环境的任何类型,并且作为非限制性的示例,可以包括以下中的一个或多个:通用计算机、专用计算机、微处理器、数字信号处理器 (DSP) 和基于多核处理器架构的处理器。

[0207] 本发明的实施方式可以在各种组件中实践,诸如集成电路模块。集成电路的设计大体上是高度自动化的过程。复杂和强大的软件工具可用于将逻辑水平设计转换为准备进行刻蚀并且在半导体衬底上形成的半导体电路设计。

[0208] 诸如加利福尼亚山景城的 Synopsys 公司以及加利福尼亚圣何塞的 Cadence Design 提供的程序使用建立好的设计规则以及预先存储的设计模块库自动地对导体进行布线,以及在半导体芯片上定位组件。一旦已经完成用于半导体电路的设计,可以将所产生的标准化电子格式(例如,Opus、GDSII 等)的设计传输至半导体制造厂或者用于制造的“芯片厂”。

[0209] 以上描述已经通过示例性和非限制性示例的方式提供了本发明的示例性实施方式的全面和有益描述。然而,当结合附图和所附权利要求进行阅读时,考虑以上描述,各种修改和调整可能对本领域技术人员变得易见。然而,本发明的教导的所有此类和相似修改仍然落入所附权利要求限定的本发明的范围内。

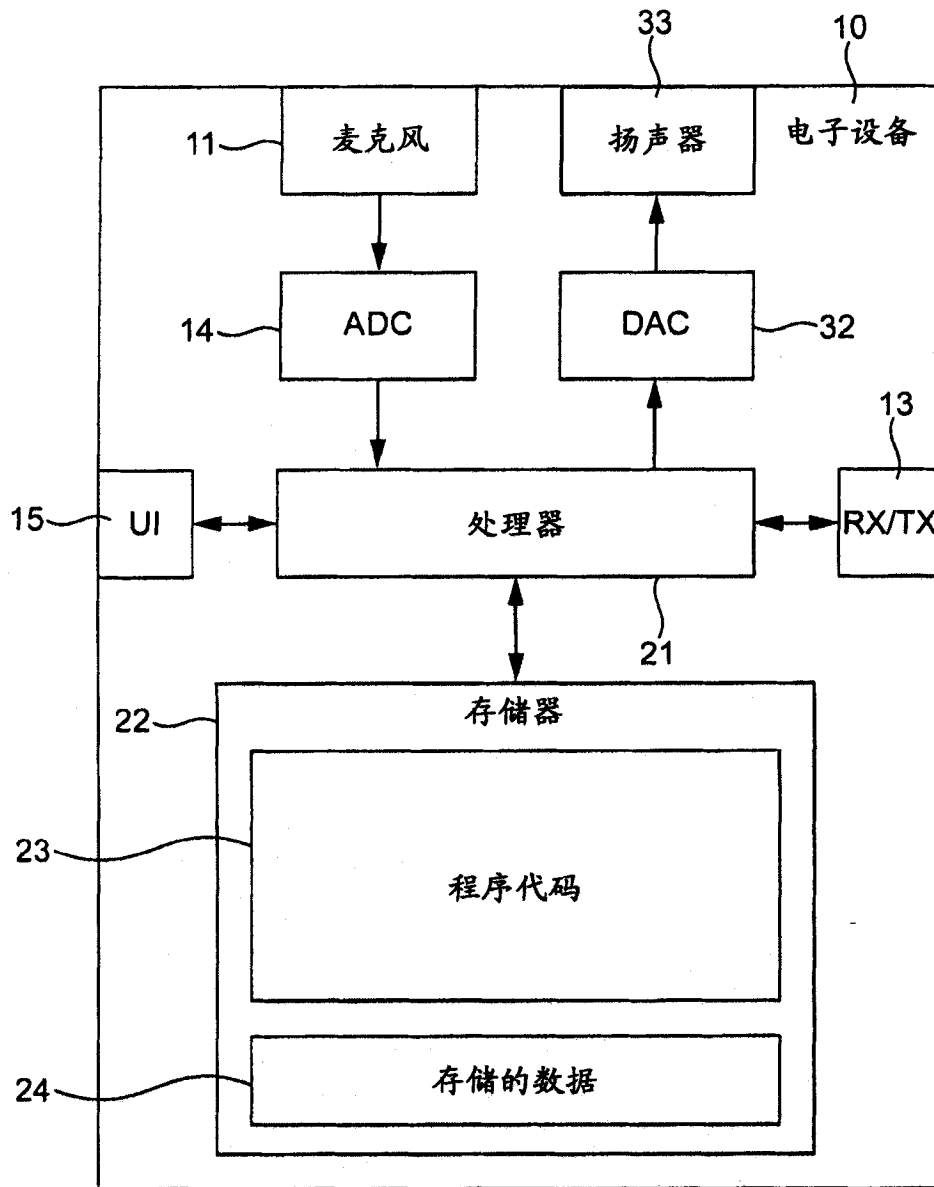


图 1

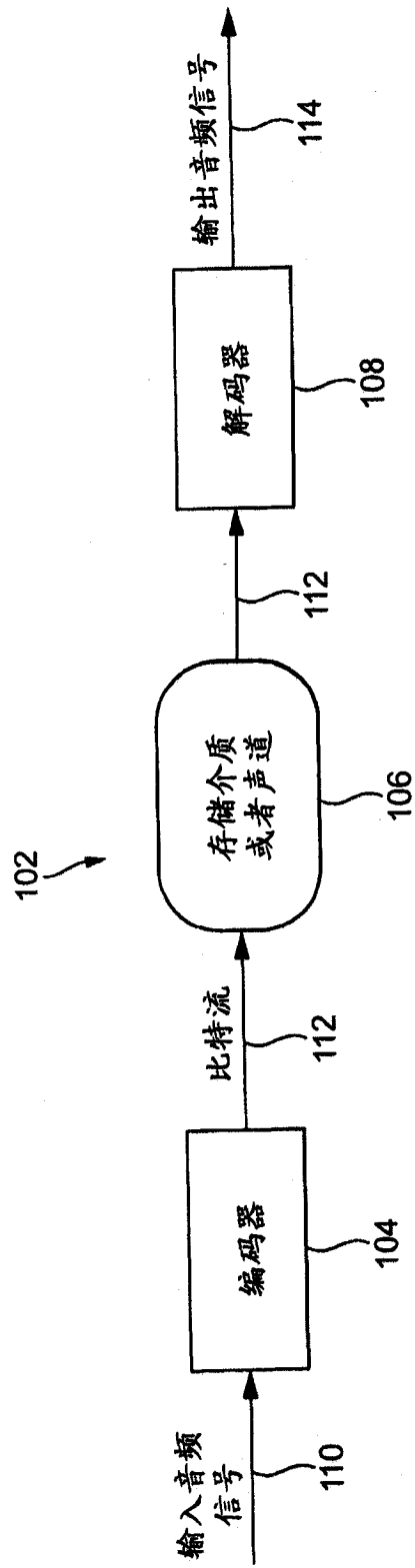


图 2

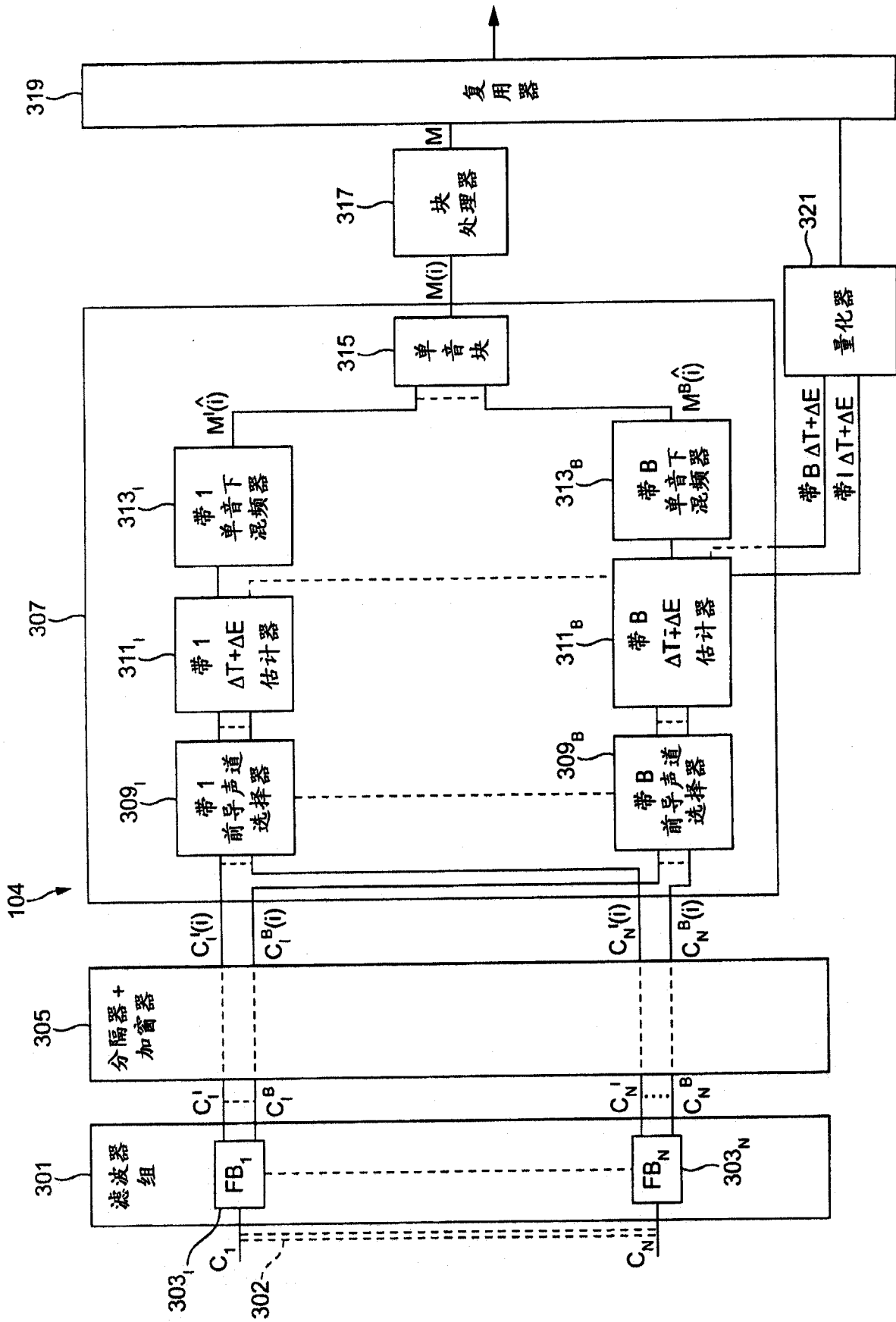


图 3

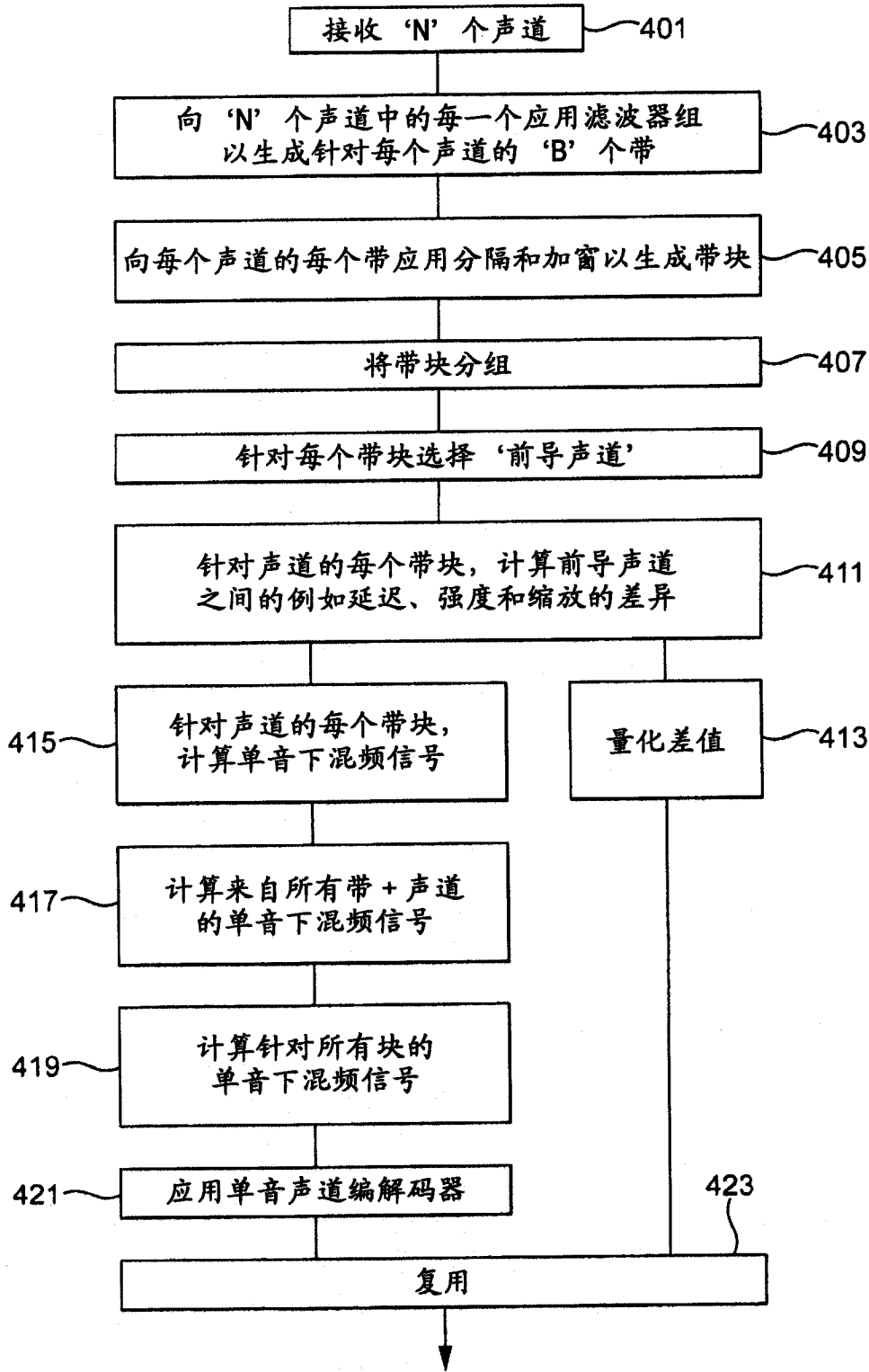


图 4

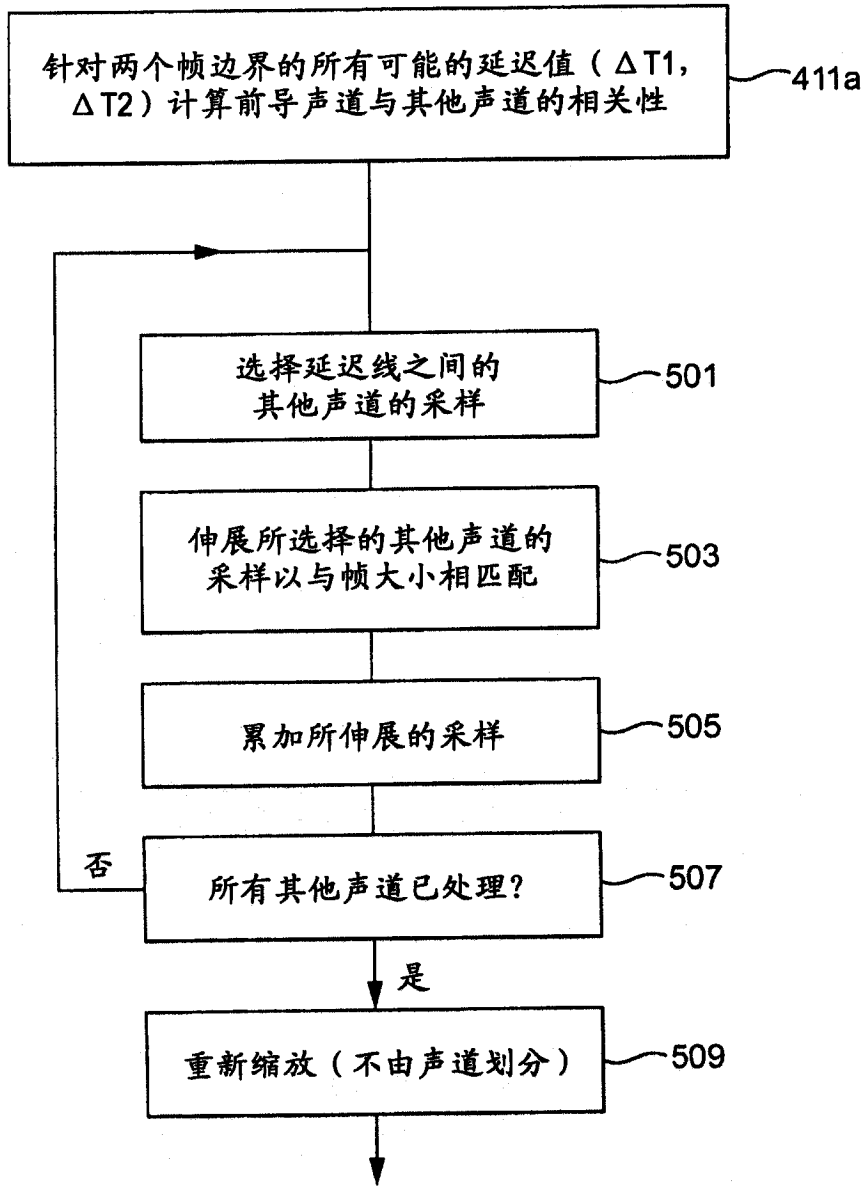


图 5

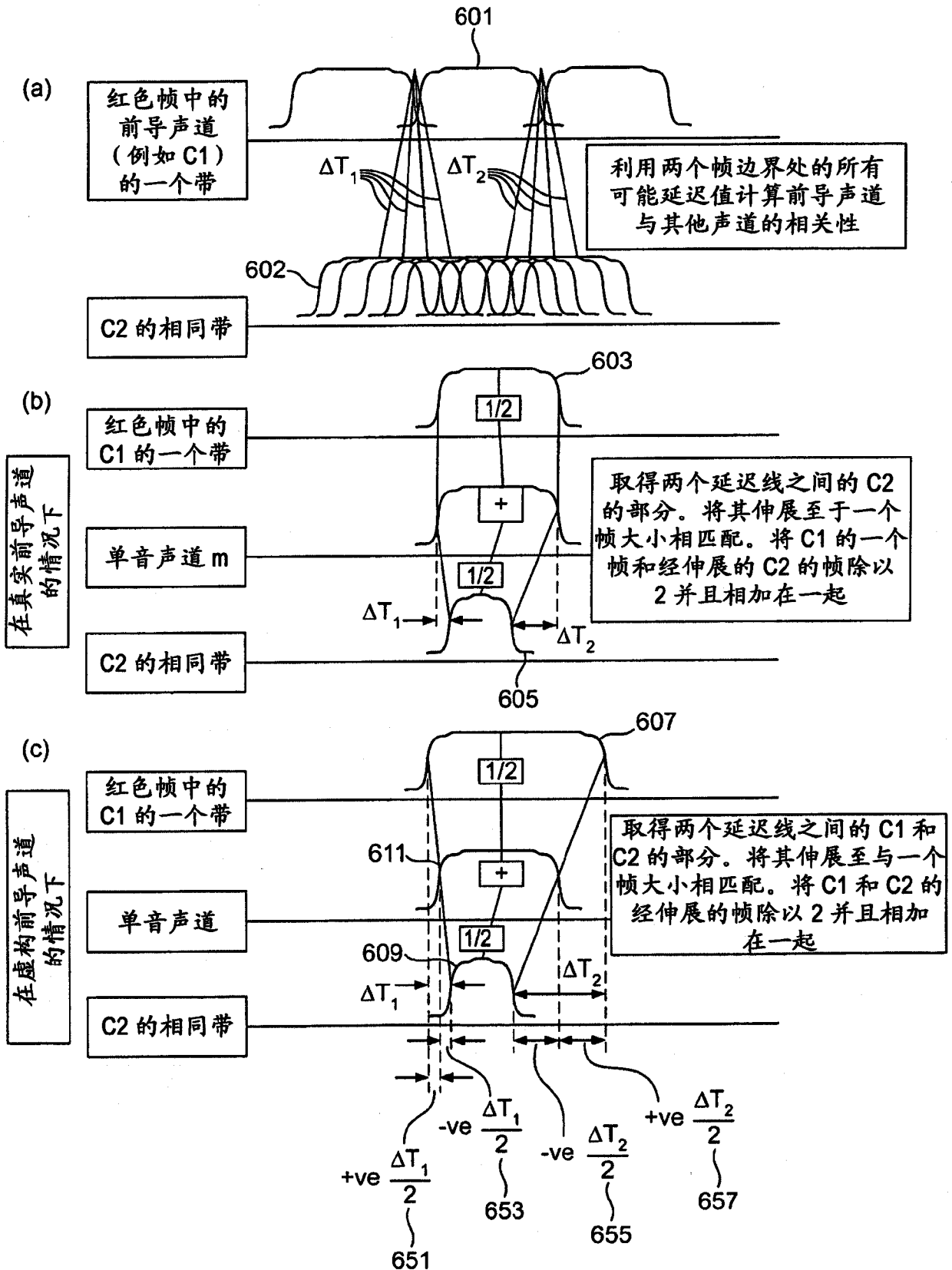


图 6

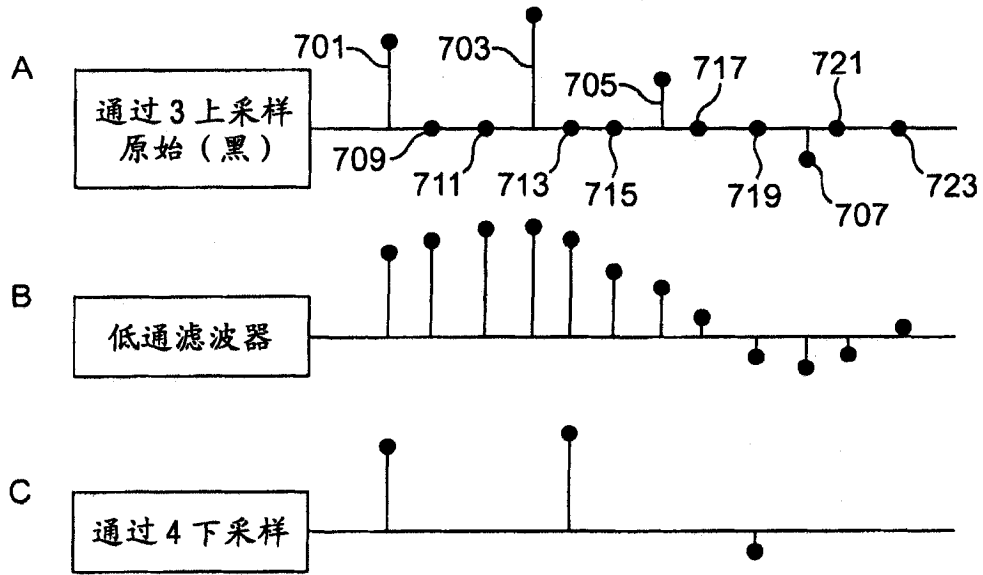


图7

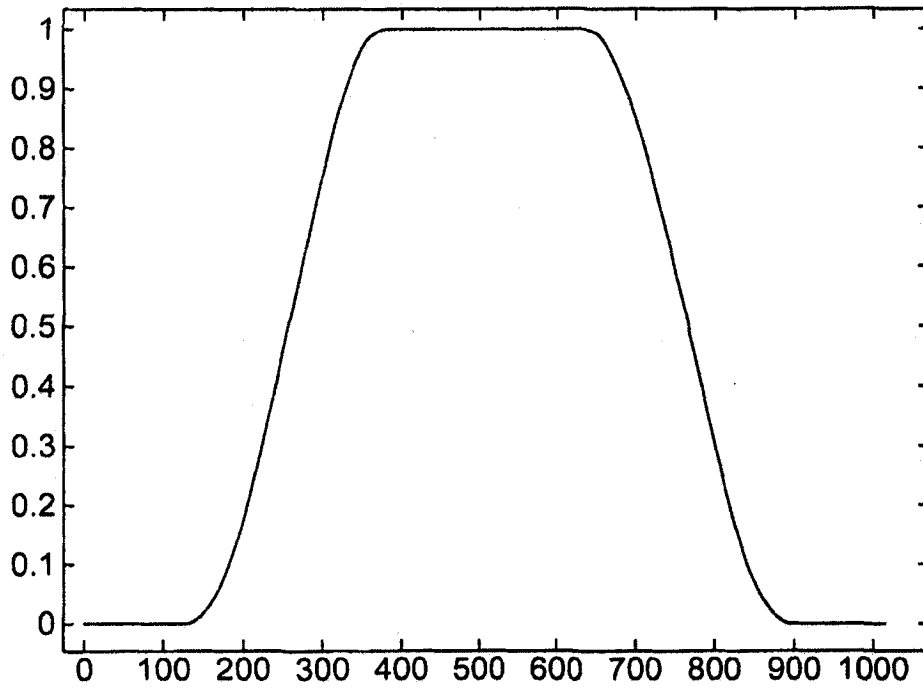


图8

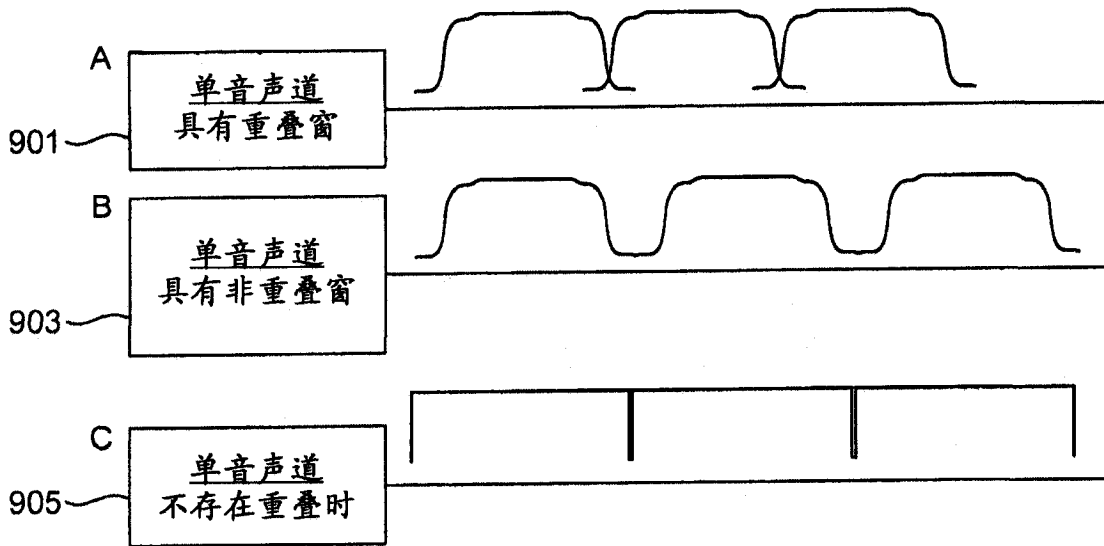


图 9

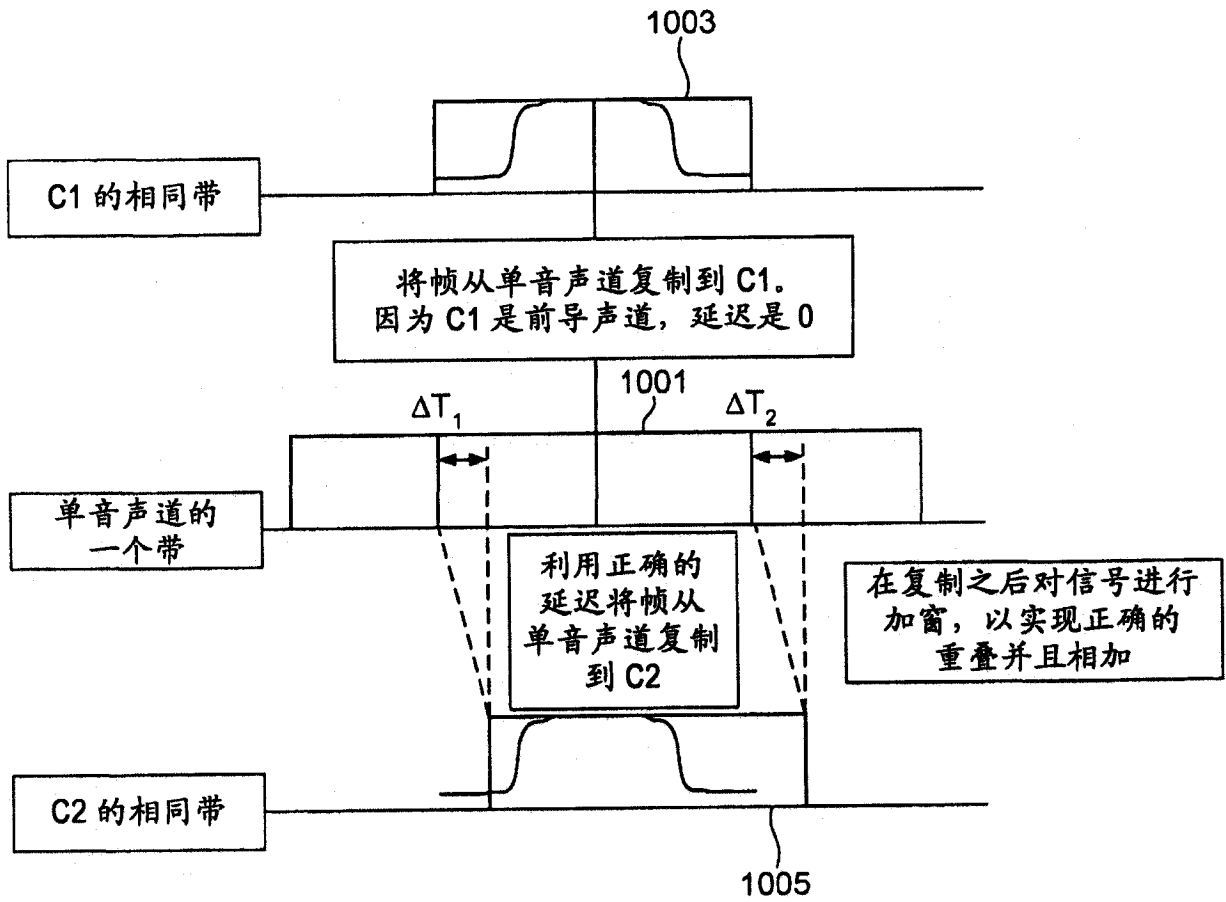


图 10

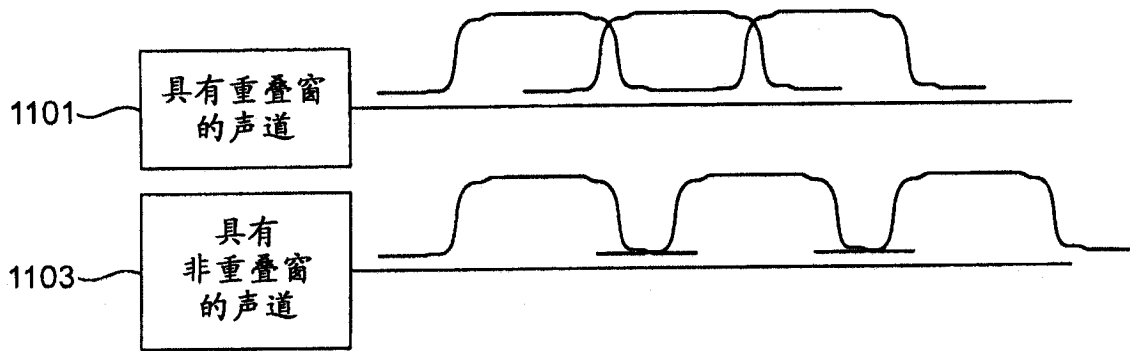


图 11

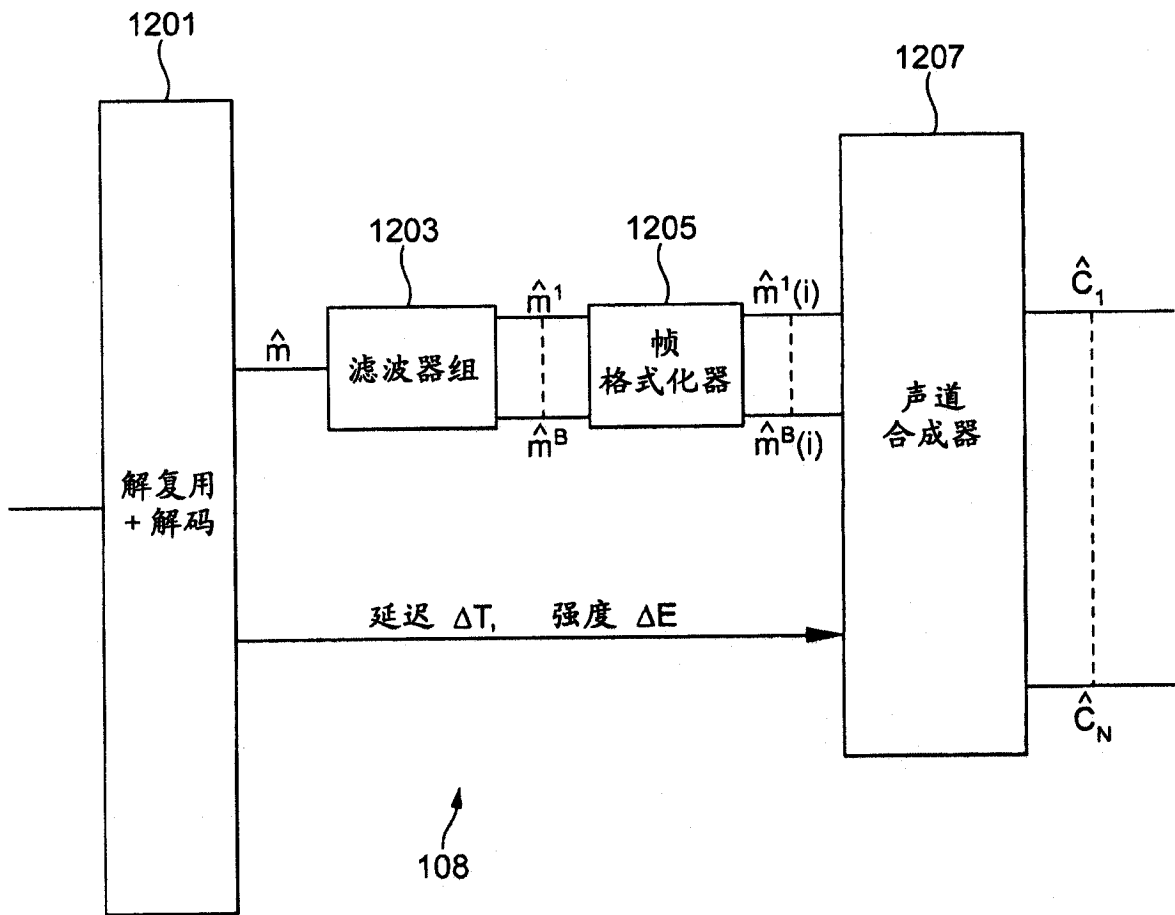


图 12

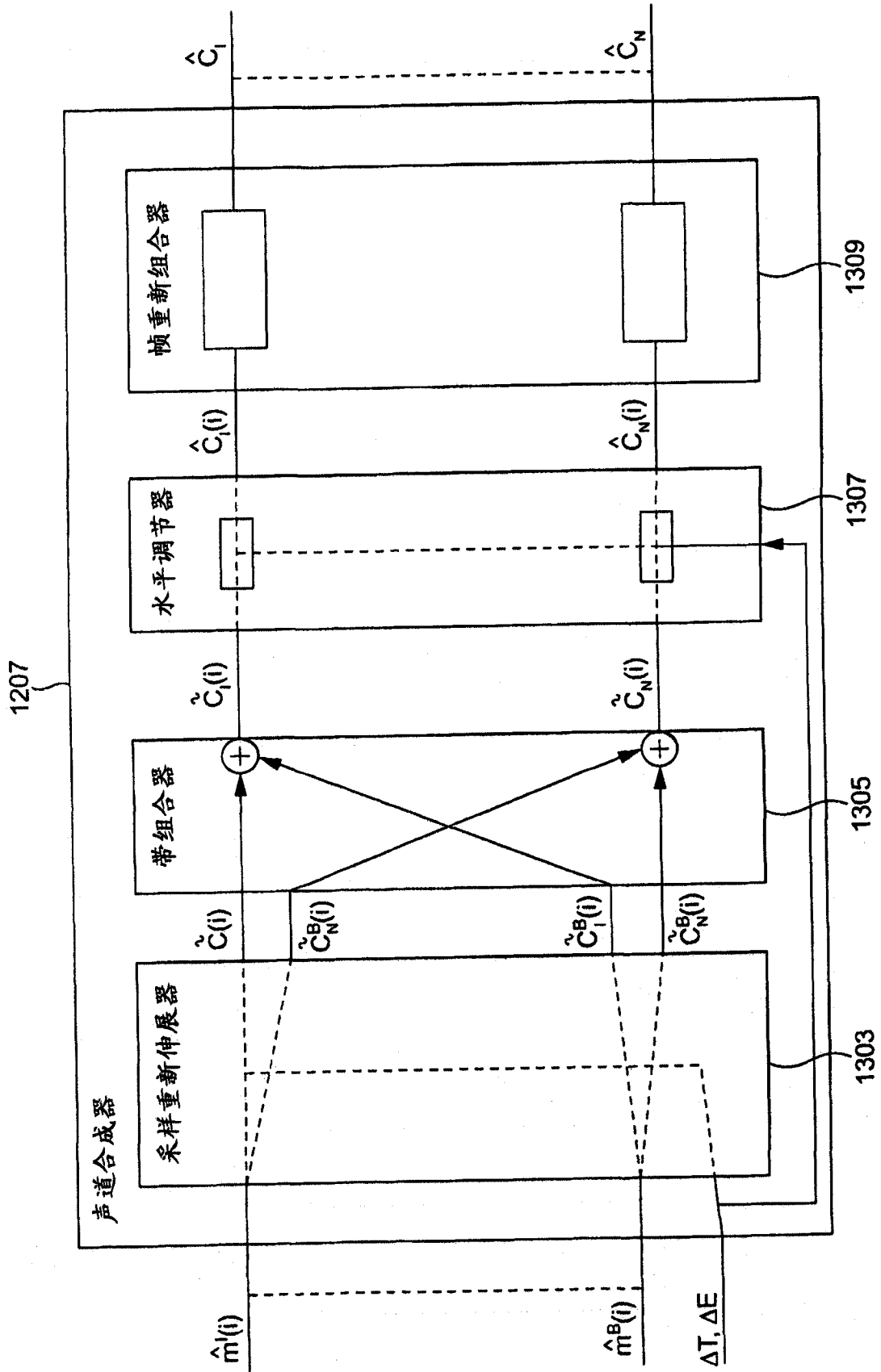


图 13

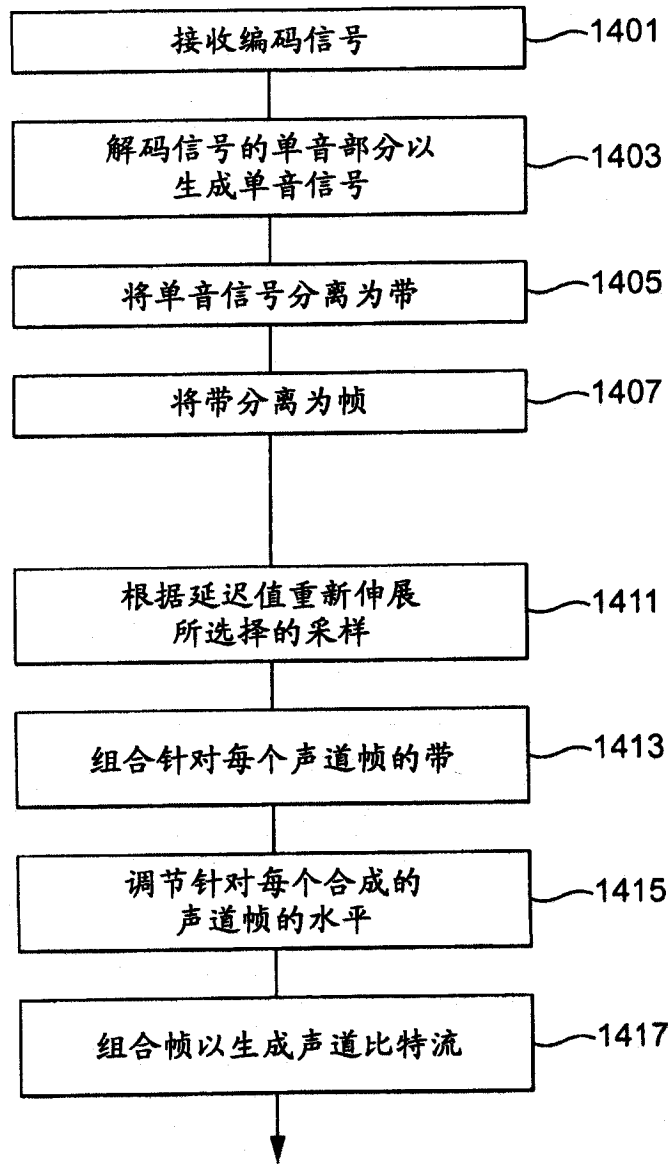


图 14