



(12) 发明专利

(10) 授权公告号 CN 111192592 B

(45) 授权公告日 2023. 09. 15

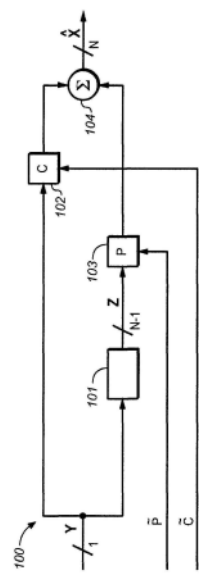
(21) 申请号 202010024100.3
(22) 申请日 2014.10.21
(65) 同一申请的已公布的文献号
 申请公布号 CN 111192592 A
(43) 申请公布日 2020.05.22
(30) 优先权数据
 61/893,770 2013.10.21 US
 61/974,544 2014.04.03 US
 62/037,693 2014.08.15 US
(62) 分案原申请数据
 201480057568.5 2014.10.21
(73) 专利权人 杜比国际公司
 地址 荷兰阿姆斯特丹
(72) 发明人 L·维勒莫斯 H·M·莱托恩
 H·普恩哈根 T·赫冯恩
(74) 专利代理机构 中国贸促会专利商标事务所
 有限公司 11038
 专利代理师 张劲松

(51) Int.Cl.
 G10L 19/005 (2013.01)
 G10L 19/008 (2013.01)
 G10L 19/02 (2013.01)
 G10L 19/16 (2013.01)
 G10L 19/26 (2013.01)
 H04S 5/00 (2006.01)
(56) 对比文件
 CN 101484936 A,2009.07.15
 CN 101529501 A,2009.09.09
 CN 101930742 A,2010.12.29
 CN 103325383 A,2013.09.25
 CN 1969317 A,2007.05.23
 US 2011096932 A1,2011.04.28
 WO 2008131903 A1,2008.11.06
 审查员 陈凭

权利要求书3页 说明书14页 附图10页

(54) 发明名称
 音频信号的参数化重构

(57) 摘要
 本发明公开了音频信号的参数化重构。编码系统(400)将N声道音频信号(X)(其中, $N \geq 3$)编码为单声道下混信号(Y)连同干上混参数和湿上混参数(C,P)。在解码系统(200)中,去相关部分(101)基于下混信号输出(N-1)声道去相关信号(Z);干上混部分(102)根据基于干上混参数而确定的干上混系数(C)线性地映射下混信号;湿上混部分(103)基于湿上混参数并且在已知中间矩阵属于预定义矩阵类的情况下填充所述中间矩阵,通过将中间矩阵乘以预定义矩阵来获得湿上混系数(P),以及根据湿上混系数线性地映射去相关信号;以及组合部分(104)组合来自上混部分的输出以获得与要被重构的信号对应的重构信号(X)。



1. 一种基于单声道下混信号(Y)重构N声道音频信号(X)的方法,所述方法包括:
由参数化重构系统的去相关部分接收单声道下混信号(Y);
处理所述单声道下混信号(Y)以输出去相关信号(Z),其中,所述去相关信号具有N-1个声道,所述处理包括向所述单声道下混信号(Y)应用相应的滤波器;
由所述参数化重构系统的干上混部分接收所述单声道下混信号(Y)和干上混参数(\bar{c}),所述干上混参数(\bar{c})与一组干上混系数(C)的第一部分一致;
基于所述一组干上混系数(C)之间的预定义关系确定所述一组干上混系数(C)的其它部分;
由所述干上混部分输出通过根据所述一组干上混系数(C)线性地映射所述单声道下混信号(Y)而计算的干上混信号(CY);
由所述参数化重构系统的湿上混部分接收所述去相关信号(Z)和一组湿上混参数(\bar{p});
从所述一组湿上混参数(\bar{p})导出一组湿上混系数(P);
由所述湿上混部分输出通过映射所述去相关信号(Z)和所述一组湿上混系数(P)而计算的湿上混信号(PZ);以及
由所述参数化重构系统的组合部分组合所述干上混信号(CY)和所述湿上混信号(PZ)以获得与要被重构的N声道音频信号(X)对应的多维重构信号(\hat{x}),
其中,所述参数化重构系统包括一个或多个处理器。
2. 根据权利要求1所述的方法,包括:
基于接收的湿上混参数填充具有比接收的湿上混参数的数量多的元素的中间矩阵,所述中间矩阵属于预定义矩阵类。
3. 根据权利要求2所述的方法,其中,导出所述一组湿上混系数包括将所述中间矩阵与预定义矩阵相乘,其中,所述一组湿上混系数对应于从所述相乘得到的矩阵并且包括比所述中间矩阵中的元素的数量多的系数。
4. 根据权利要求3所述的方法,其中,所述预定义矩阵类是以下中的一个:
下三角矩阵或上三角矩阵,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素为零;
对称矩阵,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素是相等的;以及
正交矩阵和对角矩阵的乘积,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素之间的已知关系。
5. 根据权利要求2所述的方法,其中,所述湿上混参数包括 $N(N-1)/2$ 个湿上混参数,其中,填充所述中间矩阵包括基于所述 $N(N-1)/2$ 个湿上混参数并且在已知所述中间矩阵属于预定义矩阵类的情况下获得 $(N-1)^2$ 个矩阵元素的值,其中,所述预定义矩阵包括 $N(N-1)$ 个元素,并且其中,所述一组湿上混系数包括 $N(N-1)$ 个系数。
6. 根据权利要求2所述的方法,其中,填充所述中间矩阵包括利用接收的湿上混参数作为所述中间矩阵中的元素。
7. 一种音频解码系统(200),所述音频解码系统(200)包括第一参数化重构部分(100),所述第一参数化重构部分(100)被配置为基于第一单声道下混信号(Y)重构N声道音频信号

(X),所述音频解码系统包括:

参数化重构系统的第一去相关部分,所述第一去相关部分被配置为执行包括以下的操作:

接收单声道下混信号(Y);

处理所述单声道下混信号(Y),所述处理包括向所述单声道下混信号(Y)应用相应的滤波器;以及

输出去相关信号(Z),其中,所述去相关信号具有N-1个声道;

第一干上混部分,所述第一干上混部分被配置为执行包括以下的操作:

接收所述单声道下混信号(Y)和干上混参数(\bar{C}),所述干上混参数(\bar{C})与一组干上混系数(C)的第一部分一致;

基于所述一组干上混系数(C)之间的预定义关系确定所述一组干上混系数(C)的其它部分;以及

输出通过根据所述一组干上混系数(C)线性地映射所述单声道下混信号(Y)而计算的干上混信号(CY);

所述参数化重构系统的第一湿上混部分,所述第一湿上混部分被配置为执行包括以下的操作:

接收所述去相关信号(Z)和一组湿上混参数(\bar{P});

从所述一组湿上混参数(\bar{P})导出一组湿上混系数(P);以及

输出通过映射所述去相关信号(Z)和所述一组湿上混系数(P)而计算的湿上混信号(PZ);以及

组合部分,所述组合部分被配置为执行包括以下的操作:

组合所述干上混信号(CY)和所述湿上混信号(PZ)以获得与要被重构的N声道音频信号(X)对应的多维重构信号(\hat{X}),

其中,所述参数化重构部分包括一个或多个处理器。

8.根据权利要求7所述的系统,其中,所述第一参数化重构部分被配置为执行包括以下的操作:

基于接收的湿上混参数填充具有比接收的湿上混参数的数量多的元素的中间矩阵,所述中间矩阵属于预定义矩阵类。

9.根据权利要求8所述的系统,其中,导出所述一组湿上混系数包括将所述中间矩阵与预定义矩阵相乘,其中,所述一组湿上混系数对应于从所述相乘得到的矩阵并且包括比所述中间矩阵中的元素的数量多的系数。

10.根据权利要求9所述的系统,其中,所述预定义矩阵类是以下中的一个:

下三角矩阵或上三角矩阵,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素为零;

对称矩阵,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素是相等的;以及

正交矩阵和对角矩阵的乘积,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素之间的已知关系。

11. 根据权利要求8所述的系统,其中,所述湿上混参数包括 $N(N-1)/2$ 个湿上混参数,其中,填充所述中间矩阵包括基于所述 $N(N-1)/2$ 个湿上混参数并且在已知所述中间矩阵属于预定义矩阵类的情况下获得 $(N-1)^2$ 个矩阵元素的值,其中,所述预定义矩阵包括 $N(N-1)$ 个元素,并且其中,所述一组湿上混系数包括 $N(N-1)$ 个系数。

12. 根据权利要求8所述的系统,其中,填充所述中间矩阵包括利用接收的湿上混参数作为所述中间矩阵中的元素。

13. 一种非暂时性计算机可读介质,所述非暂时性计算机可读介质存储指令,所述指令在被一个或多个处理器执行时使所述一个或多个处理器执行基于单声道下混信号(Y)重构N声道音频信号(X)的操作,所述操作包括:

由参数化重构系统的去相关部分接收单声道下混信号(Y);

处理所述单声道下混信号(Y)以输出去相关信号(Z),其中,所述去相关信号具有N-1个声道,所述处理包括向所述单声道下混信号(Y)应用相应的滤波器;

由所述参数化重构系统的干上混部分接收所述单声道下混信号(Y)和干上混参数 (\bar{C}) ,所述干上混参数 (\bar{C}) 与一组干上混系数(C)的第一部分一致;

基于所述一组干上混系数(C)之间的预定义关系确定所述一组干上混系数(C)的其它部分;

由所述干上混部分输出通过根据所述一组干上混系数(C)线性地映射所述单声道下混信号(Y)而计算的干上混信号(CY);

由所述参数化重构系统的湿上混部分接收所述去相关信号(Z)和一组湿上混参数 (\bar{P}) ;

从所述一组湿上混参数 (\bar{P}) 导出一组湿上混系数(P);

由所述湿上混部分输出通过映射所述去相关信号(Z)和所述一组湿上混系数(P)而计算的湿上混信号(PZ);以及

由所述参数化重构系统的组合部分组合所述干上混信号(CY)和所述湿上混信号(PZ)以获得与要被重构的N声道音频信号(X)对应的多维重构信号 (\hat{X}) 。

音频信号的参数化重构

[0001] 本申请是基于申请号为201480057568.5、申请日为2014年10月21日、发明名称为“音频信号的参数化重构”的专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求2013年10月21日提交的美国临时专利申请No.61/893,770、2014年4月3日提交的美国临时专利申请No.61/974,544、以及2014年8月15日提交的美国临时专利申请No.62/037,693的优先权,每一专利申请的全部内容特此通过引用并入。

技术领域

[0004] 本文中公开的发明一般涉及音频信号的编码和解码,并且特别地涉及多声道音频信号从下混信号和相关联的元数据的参数化重构。

背景技术

[0005] 包括多个扬声器的音频回放系统被频繁地用于再现由多声道音频信号所表示的音频场景,其中,多声道音频信号的相应声道在相应的扬声器上被回放。多声道音频信号可能例如已经由多个声换能器被记录或者可能已通过音频制作设备被产生。在许多情形下,对于将音频信号传输到回放设备存在带宽限制,和/或对于将音频信号存储在计算机存储器中或者便携式存储设备上存在有限的空间。存在用于音频信号的参数化编码以便减少所需要的带宽或存储大小的音频编码系统。在编码器侧,这些系统通常将多声道音频信号下混为下混信号(其通常是单声道(一个声道)或立体声(两个声道)下混),并且提取通过比如水平差异(level difference)和互相关的参数描述声道的性质的边信息(side information)。下混和边信息然后被编码,并且被发送到解码器侧。在解码器侧,在边信息的参数的控制下从下混重构(即,近似)多声道音频信号。

[0006] 鉴于可供用于回放多声道音频内容(包括针对终端用户家庭中的这些终端用户的新兴部分)的范围广泛的不同类型的设备和系统,需要新的、替代的方式以高效地对多声道音频内容进行编码,以便减少带宽要求和/或存储所需的存储器大小、和/或便于解码器侧的多声道音频信号的重构。

附图说明

[0007] 在以下,将参照附图且更详细地描述示例实施例,其中:

[0008] 图1是根据示例实施例的用于基于单声道下混信号以及相关联的干(dry)上混参数和湿(wet)上混参数重构多声道音频信号的参数化重构部分的一般化框图;

[0009] 图2是根据示例实施例的包括图1中描绘的参数化重构部分的音频解码系统的一般化框图;

[0010] 图3是根据示例实施例的用于将多声道音频信号编码为单声道下混信号和相关联的元数据的参数化编码部分的一般化框图;

[0011] 图4是根据示例实施例的包括图3中描绘的参数化编码部分的音频编码系统的一

般化框图；

[0012] 图5-11示出根据示例实施例的通过下混声道表示11.1声道音频信号的替代方式；

[0013] 图12-13示出根据示例实施例的通过下混声道表示13.1声道音频信号的替代方式；以及

[0014] 图14-16示出根据示例实施例的通过下混声道表示22.2声道音频信号的替代方式。

[0015] 所有的附图都是示意性的，并且一般仅示出为了阐明本发明所必要的部分，而其它部分则可以被省略或者仅仅被建议。

具体实施方式

[0016] 如本文中所使用的，音频信号可以是纯音频信号、视听信号或多媒体信号的音频部分或者与元数据组合的这些中的任何一个。

[0017] 如本文中所使用的，声道是与预定义/固定的空间位置/方位或未定义的空间位置（诸如“左”或“右”）相关联的音频信号。

[0018] I. 概述

[0019] 根据第一方面，示例实施例提出了用于重构音频信号的音频解码系统以及方法和计算机程序产品。根据第一方面的提出的解码系统、方法和计算机程序产品一般可以共享相同的特征和优点。

[0020] 根据示例实施例，提供了一种用于重构N声道音频信号的方法，其中， $N \geq 3$ 。所述方法包括：对单声道下混信号或携带用于重构更多音频信号的数据的多声道下混信号的声道连同相关联的干上混参数和湿上混参数一起进行接收；将具有多个(N个)声道的第一信号（其被称为干上混信号）计算为所述下混信号的线性映射，其中，作为计算所述干上混信号的一部分，一组干上混系数被应用于所述下混信号；基于所述下混信号产生(N-1)声道去相关信号；将具有多个(N个)声道的另一信号（其被称为湿上混信号）计算为所述去相关信号的线性映射，其中，作为计算所述湿上混信号的一部分，一组湿上混系数被应用于所述去相关信号的声道；以及组合所述干上混信号和湿上混信号以获得与要被重构的N声道音频信号对应的多维重构信号。所述方法进一步包括：基于接收的干上混参数确定所述一组干上混系数；基于接收的湿上混参数并且在已知具有比接收的湿上混参数的数量多的元素的中间矩阵属于预定义矩阵类(class)的情况下，填充所述中间矩阵；以及通过将所述中间矩阵与预定义矩阵相乘来获得所述一组湿上混系数，其中，所述一组湿上混系数对应于从所述相乘得到的矩阵并且包括比所述中间矩阵中的元素的数量多的系数。

[0021] 在该示例实施例中，用于重构N声道音频信号的湿上混系数的数量大于接收的湿上混参数的数量。通过利用预定义矩阵和预定义矩阵类的知晓(knowledge)以从接收的湿上混参数获得湿上混系数，可以减少使得能够重构N声道音频信号所需要的信息量，从而允许减少从编码器侧连同下混信号一起传输的元数据的量。通过减少参数化重构所需要的数据量，可以减少N声道音频信号的参数化表示的传输所需的带宽和/或存储这样的表示所需的存储器大小。

[0022] (N-1)声道去相关信号用于增加收听者所感知到的重构的N声道音频信号的内容的维度。(N-1)声道去相关信号的声道可以具有至少大致与单声道下混信号相同的频谱，或

者可以具有与单声道下混信号的频谱的重新缩放(rescale)/规范化的版本对应的频谱,并且可以连同单声道下混信号一起形成N个至少大致互不相关的声道。为了提供N声道音频信号的声道的忠实重构,去相关信号的声道的每一个优选地具有它被收听者感知为类似于下混信号的这样的性质。因此,尽管可以将互不相关的信号与来自例如白噪声的给定频谱合成,但是去相关信号的声道优选地通过处理下混信号来导出,例如包括将相应的全通滤波器应用于下混信号或者组合下混信号的部分,以便保留下混信号的尽可能多的性质(尤其是局部平稳的性质),包括下混信号的相对更细微的、心理声学制约的性质,诸如音色。

[0023] 组合湿上混信号和干上混信号可以包括将来自湿上混信号的相应声道的音频内容添加到干上混信号的相应的对应声道的音频内容,诸如基于每一个采样或每一个变换系数加性混合(additive mixing)。

[0024] 预定义矩阵类可以与对于该类中的所有矩阵都有效的至少一些矩阵元素的已知性质(诸如矩阵元素中的一些之间的某些关系,或者一些矩阵元素为零)相关联。这些性质的知晓允许基于比中间矩阵中的矩阵元素的全部数量少的湿上混参数来填充中间矩阵。解码器侧至少具有它基于较少的湿上混参数计算所有矩阵元素所需的元素的性质以及这些元素之间的关系的知晓。

[0025] 干上混信号是下混信号的线性映射意指干上混信号是通过将第一线性变换应用于下混信号而获得的。该第一变换将一个声道当作输入并且提供N个声道作为输出,并且干上混系数是定义该第一线性变换的定量性质的系数。

[0026] 湿上混信号是去相关信号的线性映射意指湿上混信号是通过将第二线性变换应用于去相关信号而获得的。该第二变换将N-1个声道当作输入并且提供N个声道作为输出,并且湿上混系数是定义该第二线性变换的定量性质的系数。

[0027] 在示例实施例中,接收所述湿上混参数可以包括接收 $N(N-1)/2$ 个湿上混参数。在本示例实施例中,填充所述中间矩阵可以包括基于接收的 $N(N-1)/2$ 个湿上混参数并且在已知所述中间矩阵属于预定义矩阵类的情况下获得 $(N-1)^2$ 个矩阵元素的值。这可以包括立即将湿上混参数的值作为矩阵元素插入,或者以合适的方式对湿上混参数进行处理以导出矩阵元素的值。在本示例实施例中,所述预定义矩阵可以包括 $N(N-1)$ 个元素,并且所述一组湿上混系数可以包括 $N(N-1)$ 个系数。例如,接收所述湿上混参数可以包括接收至多 $N(N-1)/2$ 个可独立分配的湿上混参数,和/或接收的湿上混参数的数量可以不多于用于重构N声道音频信号的湿上混系数的数量的一半。

[0028] 要理解,当将湿上混信号的声道形成为去相关信号的声道的线性映射时省略来自去相关信号的声道的贡献对应于将具有值零的系数应用于该声道,即,省略来自声道的贡献不影响作为线性映射的部分而应用的系数的数量。

[0029] 在示例实施例中,填充所述中间矩阵可以包括利用接收的湿上混参数作为所述中间矩阵中的元素。由于接收的湿上混参数在没有进行任何进一步处理的情况下被用作中间矩阵中的元素,所以可以降低填充中间矩阵以及获得上混系数所需的计算的复杂度,从而允许N声道音频信号的计算更高效的重构。

[0030] 在示例实施例中,接收所述干上混参数可以包括接收 $(N-1)$ 个干上混参数。在本示例实施例中,所述一组干上混系数可以包括N个系数,并且所述一组干上混系数基于接收的 $(N-1)$ 个干上混参数并且基于所述一组干上混系数中的系数之间的预定义关系而确定。例

如,接收所述干上混参数可以包括接收至多 $(N-1)$ 个可独立分配的干上混参数。例如,所述下混信号可根据预定义规则作为要被重构的 N 声道音频信号的线性映射而获得,并且所述干上混系数之间的预定义关系可以基于所述预定义规则。

[0031] 在示例实施例中,所述预定义矩阵类可以是以下中的一个:下三角矩阵或上三角矩阵,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素为零;对称矩阵,其中,该类中的所有矩阵的已知性质包括(主对角线的任一侧的)预定义矩阵元素是相等的;以及正交矩阵和对角矩阵的乘积,其中,该类中的所有矩阵的已知性质包括预定义矩阵元素之间的已知关系。换句话说,所述预定义矩阵类可以是下三角矩阵类、上三角矩阵类、对称矩阵类、或正交矩阵和对角矩阵的乘积类。以上类中的每一个的共同性质是其维度少于矩阵元素的全部数量。

[0032] 在示例实施例中,所述下混信号可以根据预定义规则作为要被重构的 N 声道音频信号的线性映射而获得。在本示例实施例中,所述预定义规则可以对预定义下混操作进行定义,并且所述预定义矩阵可以基于跨越所述预定义下混操作的核空间的向量。例如,所述预定义矩阵的行或列可以是形成预定义下混操作的核空间的基(例如,正交基)的向量。

[0033] 在示例实施例中,对所述单声道下混信号连同相关联的干上混参数和湿上混参数一起进行接收可以包括对所述下混信号的时间段或时间/频率片(tile)连同与该时间段或时间/频率片相关联的干上混参数和湿上混参数一起进行接收。在本示例实施例中,所述多维重构信号可以对应于要被重构的 N 声道音频信号的时间段或时间/频率片。换句话说,所述 N 声道音频信号的重构在至少一些示例实施例中可以一次一个时间段或时间/频率片地执行。音频编码/解码系统通常例如通过将合适的滤波器组应用于输入的音频信号来将时间-频率空间分成时间/频率片。时间/频率片一般意指时间-频率空间的与时间间隔/段和频率子带对应的一部分。

[0034] 根据示例实施例,提供了一种音频解码系统,所述音频解码系统包括第一参数化重构部分,所述第一参数化重构部分被配置为基于第一单声道下混信号以及相关联的干上混参数和湿上混参数重构 N 声道音频信号,其中, $N \geq 3$ 。所述第一参数化重构部分包括第一去相关部分,所述第一去相关部分被配置为接收所述第一下混信号并且基于此而输出第一 $(N-1)$ 声道去相关信号。所述第一参数化重构部分还包括第一干上混部分,所述第一干上混部分被配置为:接收干上混参数和下混信号;基于所述干上混参数确定第一组干上混系数;以及输出通过根据所述第一组干上混系数线性地映射所述第一下混信号而计算的第一干上混信号。换句话说,通过将所述单声道下混信号乘以相应系数来获得第一干上混信号的声道,所述相应系数可以是干上混系数本身,或者可以是可经由干上混系数控制的系数。所述第一参数化重构部分进一步包括第一湿上混部分,所述第一湿上混部分被配置为:接收湿上混参数和第一去相关信号;基于接收的湿上混参数并且在已知具有比接收的湿上混参数的数量多的元素的第一中间矩阵属于第一预定义矩阵类的情况下(即,通过利用已知为对于预定义矩阵类中的所有矩阵成立的某些矩阵元素的性质),填充所述第一中间矩阵;通过将所述第一中间矩阵与第一预定义矩阵相乘来获得第一组湿上混系数,其中,所述第一组湿上混系数对应于从所述相乘得到的矩阵并且包括比所述第一中间矩阵中的元素的数量多的系数;以及输出通过根据所述第一组湿上混系数线性地映射所述第一去相关信号(即,通过利用湿上混系数形成去相关信号的声道的线性组合)而计算的第一湿上混信号。

所述第一参数化重构部分还包括第一组合部分,所述第一组合部分被配置为接收所述第一干上混信号和第一湿上混信号,并且组合这些信号以获得与要被重构的N维音频信号对应的第一多维重构信号。

[0035] 在示例实施例中,所述音频解码系统可以进一步包括第二参数化重构部分,所述第二参数化重构部分可独立于第一参数化重构部分操作,并且被配置为基于第二单声道下混信号以及相关联的干上混参数和湿上混参数重构 N_2 声道音频信号,其中, $N_2 \geq 2$ 。 $N_2 = 2$ 或 $N_2 \geq 3$ 例如可以成立。在本示例实施例中,所述第二参数化重构部分可以包括第二去相关部分、第二干上混部分、第二湿上混部分以及第二组合部分,并且所述第二参数化重构部分的所述部分可以类似于所述第一参数化重构部分的对应部分被配置。在本示例实施例中,所述第二湿上混部分可以被配置为利用属于第二预定义矩阵类的第二中间矩阵和第二预定义矩阵。所述第二预定义矩阵类和第二预定义矩阵可以分别与第一预定义矩阵类和第一预定义矩阵不同或相等。

[0036] 在示例实施例中,所述音频解码系统可以适于基于多个下混声道以及相关联的干上混参数和湿上混参数重构多声道音频信号。在本示例实施例中,所述音频解码系统可以包括:多个重构部分,所述多个重构部分包括参数化重构部分,所述参数化重构部分可操作为基于相应的下混声道以及相应的相关联的干上混参数和湿上混参数独立地重构相应的多组音频信号声道;和控制部分,所述控制部分被配置为接收信令,所述信令指示与多声道音频信号的声道到由相应的下混声道所表示的、并且对于下混声道中的至少一些由相应的相关联的干上混参数和湿上混参数所表示的多组声道的划分对应的所述多声道音频信号的编码格式。在本示例实施例中,所述编码格式可以进一步对应于用于基于相应的湿上混参数获得与相应的多组声道中的至少一些相关联的湿上混系数的一组预定义矩阵。可选地,所述编码格式可以进一步对应于指示相应的中间矩阵基于相应的多组湿上混参数而将被如何填充的一组预定义矩阵类。

[0037] 在本示例实施例中,所述解码系统可以被配置为响应于接收的指示第一编码格式的信令而使用所述多个重构部分的第一子集来重构所述多声道音频信号。在本示例实施例中,所述解码系统可以被配置为响应于接收的指示第二编码格式的信令而使用所述多个重构部分的第二子集来重构所述多声道音频信号,并且所述重构部分的第一子集和第二子集中的至少一个可以包括所述第一参数化重构部分。

[0038] 根据多声道音频信号的音频内容的组成、用于从编码器侧到解码器侧的传输的可用带宽、收听者所感知的所需的回放质量和/或在解码器侧重构的音频信号的所需的保真度,最适合的编码格式在不同的应用和/或时段之间可以不同。通过对多声道音频信号支持多种编码格式,本示例实施例中的音频解码系统允许编码器侧利用更特别适合于当前情况的编码格式。

[0039] 在示例实施例中,所述多个重构部分可以包括单声道重构部分,所述单声道重构部分可操作为基于其中至多单个音频声道已被编码的下混声道独立地重构单个音频声道。在本示例实施例中,所述重构部分的第一子集和第二子集中的至少一个可以包括所述单声道重构部分。所述多声道音频信号的一些声道对于收听者所感知到的多声道音频信号的总体印象可能是特别重要的。通过利用单声道重构部分来单独地将例如这样的声道编码在它自己的下混声道中,而其它声道则在其它下混声道中被一起参数化编码,可以增加重构的

多声道音频信号的保真度。在一些示例实施例中,多声道音频信号的一个声道的音频内容可以具有与多声道音频信号的其它声道的音频内容不同的类型,并且可以通过利用以下的编码格式来增加重构的多声道音频信号的保真度:在该编码格式中,该声道被单独地编码在它自己的下混声道中。

[0040] 在示例实施例中,所述第一编码格式可以对应于从比第二编码格式数量少的下混声道重构所述多声道音频信号。通过利用较少数量的下混声道,可以减少从编码器侧到解码器侧的传输所需的带宽。通过利用较多数量的下混声道,可以增加重构的多声道音频信号的保真度和/或感知的音频质量。

[0041] 根据第二方面,示例实施例提出了用于对多声道音频信号进行编码的音频编码系统以及方法和计算机程序产品。根据第二方面的提出的编码系统、方法和计算机程序产品一般可以共享相同的特征和优点。而且,以上对于根据第一方面的解码系统、方法和计算机程序产品的特征呈现的优点对于根据第二方面的编码系统、方法和计算机程序产品的对应特征一般可以是有效的。

[0042] 根据示例实施例,提供了一种用于将N声道音频信号编码为单声道下混信号和元数据的方法,所述元数据适合于所述音频信号从下混信号和基于所述下混信号而确定的(N-1)声道去相关信号的参数化重构,其中, $N \geq 3$ 。所述方法包括:接收所述音频信号;根据预定义规则将单声道下混信号计算为所述音频信号的线性映射;以及确定一组干上混系数以便定义近似所述音频信号的下混信号的线性映射(例如,在仅下混信号可供用于重构的假设下经由最小均方误差近似)。所述方法进一步包括基于接收的所述音频信号的协方差和通过所述下混信号的线性映射近似的所述音频信号的协方差之间的差确定中间矩阵,其中,所述中间矩阵在被乘以预定义矩阵时对应于一组湿上混系数,所述一组湿上混系数定义作为所述音频信号的参数化重构的一部分的所述去相关信号的线性映射,并且其中,所述一组湿上混系数包括比所述中间矩阵中的元素的数量多的系数。所述方法进一步包括将下混信号连同可从其导出所述一组干上混系数的干上混参数以及湿上混参数一起输出,其中,所述中间矩阵具有比输出的湿上混参数的数量多的元素,并且其中,假如所述中间矩阵属于预定义矩阵类,则所述中间矩阵由输出的湿上混参数唯一地定义。

[0043] 解码器侧的音频信号的参数化重构副本包括作为一个贡献的通过下混信号的线性映射形成的干上混信号、以及作为另一贡献的通过去相关信号的线性映射形成的湿上混信号。所述一组干上混系数定义下混信号的线性映射,而所述一组湿上混系数定义去相关信号的线性映射。通过输出比湿上混系数的数量少的并且基于预定义矩阵和预定义矩阵类可从其导出湿上混系数的湿上混参数,可以减少被发送到解码器侧以使得能够重构N声道音频信号的信息量。通过减少参数化重构所需要的数据量,可以减少N声道音频信号的参数化表示的传输所需的带宽和/或存储这样的表示所需的存储器大小。

[0044] 所述中间矩阵可以基于接收的音频信号的协方差和通过下混信号的线性映射近似的音频信号的协方差之间的差(例如对于补充通过下混信号的线性映射近似的音频信号的协方差的、通过去相关信号的线性映射获得的信号的协方差)而确定。

[0045] 在示例实施例中,确定所述中间矩阵可以包括确定中间矩阵使得通过由所述一组湿上混系数定义的所述去相关信号的线性映射获得的信号的协方差近似于接收的所述音频信号的协方差和通过所述下混信号的线性映射近似的所述音频信号的协方差之间的差,

或者与该差基本上一致。换句话说,所述中间矩阵可以被确定为使得作为通过下混信号的线性映射形成的干上混信号与通过去相关信号的线性映射形成的湿上混信号的和而获得的音频信号的重构副本完全地或至少近似地恢复接收的音频信号的协方差。

[0046] 在示例实施例中,输出所述湿上混参数可以包括输出至多 $N(N-1)/2$ 个可独立分配的湿上混参数。在本示例实施例中,所述中间矩阵可以具有 $(N-1)^2$ 个矩阵元素,并且假如所述中间矩阵属于预定义矩阵类,则所述中间矩阵可以由输出的湿上混参数唯一地定义。在本示例实施例中,所述一组湿上混系数可以包括 $N(N-1)$ 个系数。

[0047] 在示例实施例中,所述一组干上混系数可以包括 N 个系数。在本示例实施例中,输出所述干上混参数可以包括输出至多 $N-1$ 个干上混参数,并且所述一组干上混系数可使用所述预定义规则从所述 $N-1$ 个干上混参数导出。

[0048] 在示例实施例中,确定的一组干上混系数可以定义与所述音频信号的最小均方误差近似对应的所述下混信号的线性映射,即,在一组下混信号的线性映射当中,确定的一组干上混系数可以定义最小均方意义上最佳近似音频信号的线性映射。

[0049] 根据示例实施例,提供了一种音频编码系统,所述音频编码系统包括参数化编码部分,所述参数化编码部分被配置为将 N 声道音频信号编码为单声道下混信号和元数据,所述元数据适合于所述音频信号从下混信号和基于所述下混信号而确定的 $(N-1)$ 声道去相关信号的参数化重构,其中, $N \geq 3$ 。所述参数化编码部分包括:下混部分,所述下混部分被配置为接收所述音频信号,并且根据预定义规则将单声道下混信号计算为所述音频信号的线性映射;以及第一分析部分,所述第一分析部分被配置为确定一组干上混系数以便定义近似所述音频信号的下混信号的线性映射。所述参数化编码部分进一步包括第二分析部分,所述第二分析部分被配置为基于接收的所述音频信号的协方差和通过所述下混信号的线性映射近似的所述音频信号的协方差之间的差确定中间矩阵,其中,所述中间矩阵在被乘以预定义矩阵时对应于一组湿上混系数,所述一组湿上混系数定义作为所述音频信号的参数化重构的一部分的所述去相关信号的线性映射,其中,所述一组湿上混系数包括比所述中间矩阵中的元素的数量多的系数。所述参数化编码部分被进一步配置为将下混信号连同可从其导出所述一组干上混系数的干上混参数以及湿上混参数一起输出,其中,所述中间矩阵具有比输出的湿上混参数的数量多的元素,并且其中,假如所述中间矩阵属于预定义矩阵类,则所述中间矩阵由输出的湿上混参数唯一地定义。

[0050] 在示例实施例中,所述音频编码系统可以被配置为提供多个下混声道以及相关干上混参数和湿上混参数的形式的多声道音频信号。在本示例实施例中,所述音频编码系统可以包括:多个编码部分,所述多个编码部分包括参数化编码部分,所述参数化编码部分可操作为基于相应的多组音频信号声道独立地计算相应的下混声道和相应的相关联的上混参数。在本示例实施例中,所述音频编码系统可以进一步包括控制部分,所述控制部分被配置为确定与所述多声道音频信号的声道到要由相应的下混声道所表示的、并且对于下混声道中的至少一些要由相应的相关联的干上混参数和湿下混参数所表示的多组声道的划分对应的所述多声道音频信号的编码格式。在本示例实施例中,所述编码格式可以进一步对应于用于计算所述相应的下混声道中的至少一些的一组预定义规则。在本示例实施例中,所述音频编码系统可以被配置为响应于确定的编码格式为第一编码格式而使用所述多个编码部分的第一子集来对所述多声道音频信号进行编码。在本示例实施例中,所

述音频编码系统可以被配置为响应于确定的编码格式为第二编码格式而使用所述多个编码部分的第二子集来对所述多声道音频信号进行编码,并且所述编码部分的第一子集和第二子集中的至少一个可以包括所述第一参数化编码部分。在本示例实施例中,所述控制部分可以例如基于用于将多声道音频信号的编码版本传输到解码器侧的可用带宽、基于多声道音频信号的声道的音频内容和/或基于指示期望的编码格式的输入信号来确定编码格式。

[0051] 在示例实施例中,所述多个编码部分可以包括单声道编码部分,所述单声道编码部分可操作为在下混声道中独立地对至多单个音频声道进行编码,并且所述编码部分的第一子集和第二子集中的至少一个可以包括所述单声道编码部分。

[0052] 根据示例实施例,提供了一种计算机程序产品,所述计算机程序产品包括具有用于执行所述第一方面和第二方面的方法中的任何一个的指令的计算机可读介质。

[0053] 根据示例实施例,在所述第一方面和第二方面的方法、编码系统、解码系统和计算机程序产品中的任何一个中, $N=3$ 或 $N=4$ 可以成立。

[0054] 进一步的示例实施例在从属权利要求中被定义。注意,示例实施例包括特征的所有组合,即使在互不相同的权利要求中被记载。

[0055] II. 示例实施例

[0056] 在将参照图3和图4描述的编码器侧,单声道下混信号 Y 根据以下方程被计算为 N 声道音频信号 $X=[x_1 \cdots x_N]^T$ 的线性映射:

$$[0057] \quad Y = [d_1 \cdots d_N] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \sum_{n=1}^N d_n x_n = DX, \quad (1)$$

[0058] 其中, d_n ($n=1, \cdots, N$) 是由下混矩阵 D 表示的下混系数。在将参照图1和图2描述的解码器侧, N 声道音频信号的参数化重构根据以下方程执行:

$$[0059] \quad \hat{X} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} Y + \begin{bmatrix} p_{11} & \cdots & p_{1,N-1} \\ p_{21} & \cdots & p_{2,N-1} \\ \vdots & \ddots & \vdots \\ p_{N,1} & \cdots & p_{N,N-1} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_{N-1} \end{bmatrix} = CY + PZ, \quad (2)$$

[0060] 其中, c_n ($n=1, \cdots, N$) 是由矩阵干上混矩阵 C 表示的干上混系数, $p_{n,k}$ ($n=1, \cdots, N, k=1, \cdots, N-1$) 是由湿上混矩阵 P 表示的湿上混系数,并且 z_k ($k=1, \cdots, N-1$) 是基于下混信号 Y 而产生的 $(N-1)$ 声道去相关信号 Z 的声道。如果每个音频信号的声道被表示为行,则原始音频信号 X 的协方差矩阵可以被表达为 $R=XX^T$,并且重构的音频信号 \hat{X} 的协方差矩阵可以被表达为 $\hat{R}=\hat{X}\hat{X}^T$ 。要注意,如果例如音频信号被表示为包括复值变换系数的行,则可以例如考虑 XX^* (其中, X^* 是矩阵 X 的复共轭转置)的实数部分,而不是 XX^T 。

[0061] 为了提供原始音频信号 X 的忠实重构,对于由方程(2)给出的重构来说可能有利的是恢复(reinstate)全协方差,即,可能有利的是利用干上混矩阵 C 和湿上混矩阵 P 使得

$$[0062] \quad R=\hat{R}. \quad (3)$$

[0063] 一种方法是首先通过对以下正规方程(normal equation)进行求解来找到给出最小二乘意义上的最佳可能的“干”上混 $\hat{X}_0=CY$ 的干上混矩阵 C :

[0064] $CYY^T = XY^T$. (4)

[0065] 对于 $\hat{X}_0 = CY$, 通过矩阵C求解方程(4), 以下方程成立:

[0066] $R = \hat{X}_0 \hat{X}_0^T + (\hat{X}_0 - X)(\hat{X}_0 - X)^T = R_0 + \Delta R$. (5)

[0067] 假定去相关信号Z的声道是互不相关的, 并且全部都具有等于单声道下混信号Y的能量的相同能量 $\|Y\|^2$, 则可以根据以下方程来对正定缺失 (missing) 协方差 ΔR 进行因子分解:

[0068] $\Delta R = PP^T \|Y\|^2$. (6)

[0069] 可以通过利用求解方程(4)的干上混矩阵C和求解方程(6)的湿上混矩阵P来根据方程(3)恢复全协方差。方程(1)和(4)隐含对于非退化下混矩阵D而言, $DCYY^T = YY^T$, 并且从而

[0070] $\sum_{n=1}^N d_n c_n = DC = 1$, (7)

[0071] 方程(5)和(7)隐含 $D(X_0 - X) = DCY - Y = 0$ 并且

[0072] $D \Delta R = 0$. (8)

[0073] 因此, 缺失协方差 ΔR 具有秩N-1, 并且实际上可以通过利用具有N-1个互不相关的声道的去相关信号Z来提供。方程(6)和(8)隐含 $DP = 0$, 使得求解方程(6)的湿上混矩阵P的列可以从跨越下混矩阵D的核空间的向量构造。用于找到合适的湿上混矩阵P的计算因此可以被移至该较低维数的空间。

[0074] 令V是包含下混矩阵D的核空间(即, 向量v的线性空间, 其中 $Dv = 0$) 的正交基的、大小为N(N-1)的矩阵。对于N=2、N=3和N=4的这样的预定义矩阵V的示例分别是:

[0075] $\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$ 和 $\frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$. (9)

[0076] 在由V给出的基中, 缺失协方差可以被表达为 $R_v = V^T (\Delta R) V$ 。为了找到求解方程(6)的湿上混矩阵P, 因此可以首先通过对 $R_v = HH^T$ 进行求解来找到矩阵H, 并然后按照 $P = VH / \|Y\|$ 获得P, 其中, $\|Y\|$ 是单声道下混信号Y的能量的平方根。可以按照 $P = VH0 / \|Y\|$ 获得其它合适的上混矩阵P, 其中, 0是正交矩阵。可替代地, 可以通过单声道下混信号Y的能量 $\|Y\|^2$ 来重新缩放缺失协方差 R_v , 并且改为对以下方程进行求解:

[0077] $\frac{R_v}{\|Y\|^2} = H_R H_R^T$, (10)

[0078] 其中, $H = H_R \|Y\|$, 并且按照以下方程获得P:

[0079] $P = V G_R$. (11)

[0080] 当 H_R 的项被量化并且期望的输出具有静音 (silent) 声道时, 如以上所述的预定义矩阵V的性质可能是不方便的。作为示例, 对于N=3, 对于(9)的第二个矩阵更好的选择将是:

[0081] $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix}$. (12)

[0082] 幸运的是, 只要矩阵V的列是线性独立的, 就可以丢弃这些列成对正交的要求。对

于 $\Delta R = VR_V V^T$ 的期望的解 R_V 然后通过 $R_R = W^T (\Delta R) W$ 与 $= V (V^T V)^{-1}$ (V 的伪逆) 来获得。

[0083] 矩阵 R_V 是大小为 $(N-1)^2$ 的正半定矩阵, 并且存在找到对于方程 (10) 的解、得到维数为 $N(N-1)/2$ 的相应矩阵类 (即, 在所述相应矩阵类中, 矩阵由 $N(N-1)/2$ 个矩阵元素唯一地定义) 内的解的若干方法。可以例如通过利用以下来获得解:

[0084] a. Cholesky 因子分解, 得到下三角 H_R ;

[0085] b. 正平方根, 得到对称正半定 H_R ; 或

[0086] c. 极分解 (polar), 得到形式 $J_R = OA$ 的 H_R , 其中, O 是正交的, 并且 Λ 是对角的。

[0087] 而且, 存在选项 a) 和 b) 的规范化版本, 在这些版本中, H_R 可以被表达为 $H_R = \Lambda H_0$, 其中, Λ 是对角的, 并且 H_0 的全部对角元素都等于一。以上的替代方案 a、b 和 c 提供了不同矩阵类 (即, 下三角矩阵、对称矩阵以及对角矩阵和正交矩阵的乘积) 中的解 H_R 。如果 H_R 所属于的矩阵类在解码器侧是已知的, 即, 如果已知 H_R 属于例如根据以上替代方案 a、b 和 c 中的任何一个的预定义矩阵类, 则可以仅基于 H_R 的 $N(N-1)/2$ 个元素来填充 H_R 。如果同样矩阵 V 在解码器侧是已知的, 例如, 如果已知 V 是 (9) 中给出的矩阵中的一个, 则然后可以经由方程 (11) 来获得根据方程 (2) 进行重构所需要的湿上混矩阵 P 。

[0088] 图3是根据示例实施例的参数化编码部分300的一般化框图。该参数化编码部分300被配置为将 N 声道音频信号 X 编码为单声道下混信号 Y 和适合于根据方程 (2) 的音频信号 X 的参数化重构的元数据。参数化编码部分300包括下混部分301, 该下混部分301接收音频信号 X , 并且根据预定义规则将单声道下混信号 Y 计算为音频信号 X 的线性映射。在本示例实施例中, 下混部分301根据方程 (1) 计算下混信号 Y , 其中, 下混矩阵 D 是预定义的并且对应于预定义规则。第一分析部分302确定干上混矩阵 C 所表示的一组干上混系数, 以便定义近似音频信号 X 的下混信号 Y 的线性映射。该下混信号 Y 的线性映射在方程 (2) 中由 CY 表示。在本示例实施例中, 根据方程 (4) 来确定 N 个干上混系数 C , 使得下混信号 Y 的线性映射 CY 对应于音频信号 X 的最小均方近似。第二分析部分303基于接收的音频信号 X 的协方差矩阵和通过下混信号 Y 的线性映射 CY 近似的音频信号的协方差矩阵之间的差来确定中间矩阵 H_R 。在本示例实施例中, 协方差矩阵是分别由第一处理部分304和第二处理部分305计算的, 并然后被提供给第二分析部分303。在本示例实施例中, 中间矩阵 H_R 根据上述对方程 (10) 进行求解的方法 b 确定, 从而得到对称的中间矩阵 H_R 。如方程 (1) 和 (11) 中所指示的, 中间矩阵 H_R 在被乘以预定义矩阵 V 时经由一组湿上混参数 P 来定义作为解码器侧的音频信号 X 的参数化重构的一部分的、去相关信号 Z 的线性映射 PZ 。在本示例实施例中, 对于情况 $N=3$, 中间矩阵 V 是 (9) 中的第二个矩阵, 并且对于情况 $N=4$, 是 (9) 中的第三个矩阵。参数化编码部分300将下混信号 Y 连同干上混参数 \bar{C} 以及湿上混参数 \bar{P} 一起输出。在本示例实施例中, N 个干上混系数 C 中的 $N-1$ 个是干上混参数 \bar{C} , 而剩余的一个干上混系数可经由方程 (7) 从干上混参数 \bar{C} 导出 (如果预定义下混矩阵 D 已知的话)。由于中间矩阵 H_R 属于对阵矩阵类, 所以它由它的 $(N-1)^2$ 个元素中的 $N(N-1)/2$ 个唯一地定义。在本示例实施例中, 中间矩阵 H_R 的元素中的 $N(N-1)/2$ 个因此是湿上混参数 \bar{P} , 在已知中间矩阵 H_R 是对称的情况下, 可从湿上混参数 \bar{P} 导出中间矩阵 H_R 的其余部分。

[0089] 图4是根据示例实施例的、包括参照图3描述的参数化编码部分300的音频编码系统400的一般化框图。在本示例实施例中, 例如由一个或多个声换能器401记录的或者由音

频制作设备401产生的音频内容是以N声道音频信号X的形式提供的。正交镜像滤波器(QMF)分析部分402将音频信号X逐个时间段地变换到QMF域中以供时间/频率片的形式的音频信号X的参数化编码部分300的处理。由参数化编码部分300输出的下混信号Y被QMF合成部分403从QMF域变换回去,并且被变换部分404变换到修正离散余弦变换(MDCT)域中。量化部分405和406分别对干上混参数 \bar{C} 和湿上混参数 \bar{P} 进行量化。例如,可以利用0.1或0.2(无量纲)的步长大小的均匀量化,接着进行哈夫曼编码的形式的熵编码。具有步长大小0.2的较粗略的量化可以例如被利用以节省传输带宽,而具有步长大小0.1的较精细的量化可以例如被利用以改善解码器侧的重构的保真度。MDCT变换的下混信号Y以及量化的干上混参数 \bar{C} 和湿上混参数 \bar{P} 然后被复用器407组合成比特流B,以供传输到解码器侧。音频编码系统400还可以包括核心编码器(图4中未示出),该核心编码器被配置为在下混信号Y被提供给复用器407之前使用感知音频编解码器(诸如Dolby Digital或MPEG AAC)对下混信号Y进行编码。

[0090] 图1是根据示例实施例的、被配置为基于单声道下混信号Y以及相关联的干上混参数 \bar{C} 和湿上混参数 \bar{P} 来重构N声道音频信号X的参数化重构部分100的一般化框图。该参数化重构部分100适于根据方程(2)(即,使用干上混参数C和湿上混参数P)执行重构。然而,代替接收干上混参数C和湿上混参数P本身,可从其导出干上混参数C和湿上混参数P的干上混参数 \bar{C} 和湿上混参数 \bar{P} 被接收。去相关部分101接收下混信号Y,并且基于此而输出(N-1)声道去相关信号 $Z = [z_1 \cdots z_{N-1}]^T$ 。在本示例实施例中,通过对下混信号Y进行处理(包括将相应的全通滤波器应用于下混信号Y)来导出去相关信号Z的声道,以便提供与下混信号Y不相关的、并且具有在频谱上类似于下混信号Y而且也被收听者感知为类似于下混信号Y的音频内容的音频内容的声道。(N-1)声道去相关信号Z用于增加收听者所感知到的N声道音频信号X的重构版本 \hat{X} 的维度。在本示例实施例中,去相关信号Z的声道具有至少大致与单声道下混信号Y的频谱相同的频谱,并且连同单声道下混信号Y一起形成N个至少大致互不相关的声道。干上混部分102接收干上混参数 \bar{C} 和下混信号Y。在本示例实施例中,干上混参数 \bar{C} 与N个干上混系数C中的头N-1个一致,而剩余的干上混系数基于由方程(7)给出的干上混系数C之间的预定义关系来确定。干上混部分102输出通过根据所述一组干上混系数C线性地映射下混信号Y而计算的并且由方程(2)中的CY表示的干上混信号。湿上混部分103接收湿上混参数 \bar{P} 和去相关信号Z。在本示例实施例中,湿上混参数 \bar{P} 是根据方程(10)在编码器侧确定的中间矩阵 H_R 的 $N(N-1)/2$ 个元素。在本示例实施例中,在已知中间矩阵 H_R 属于预定义矩阵类(即,它是对称的)并且利用该矩阵的元素之间的对应关系的情况下,湿上混部分103填充中间矩阵 H_R 的剩余元素。湿上混部分103然后通过利用方程(11)(即,通过将中间矩阵 H_R 乘以预定义矩阵V(即,对于情况 $N=3$, (9)中的第二个矩阵,以及对于情况 $N=4$, (9)中的第三个矩阵))来获得一组湿上混系数P。因此, $N(N-1)$ 个湿上混系数P从接收的 $N(N-1)/2$ 个可独立分配的湿上混参数 \bar{P} 导出。湿上混部分103输出通过根据所述一组湿上混系数P线性地映射去相关信号Z而计算的并且由方程(2)中的PZ表示的湿上混信号。组合部分104接收干上混信号CY和湿上混信号PZ,并且组合这些信号以获得与要被重构的N声道音频信号X对应的第一多维重构信号 \hat{X} 。在本示例实施例中,组合部分104通过根据方程(2)将干上混信号CY的相应声道的音频内容与湿上混信号PZ的相应声道进行组合来获得重构信号 \hat{X} 的相应声道。

[0091] 图2是根据示例实施例的音频解码系统200的一般化框图。该音频解码系统200包括参照图1描述的参数化重构部分100。接收部分201(例如,包括解复用器)接收从参照图4描述的音频编码系统400传输的比特流B,并且从比特流B提取下混信号Y以及相关联的干上混参数 \bar{C} 和湿上混参数 \bar{P} 。在下混信号Y使用感知音频编解码器(诸如Dolby Digital或MPEG AAC)被编码在比特流B中的情况下,音频解码系统200可以包括核心解码器(图2中未示出),该核心解码器被配置为当下混信号Y被从比特流B提取时对该下混信号Y进行解码。变换部分202通过执行逆MDCT来变换下混信号Y,并且QMF分析部分203将下混信号Y变换到QMF域中,以供时间/频率片的形式的下混信号Y的参数化重构部分100的处理。去量化部分204和205在将干上混参数 \bar{C} 和湿上混参数 \bar{P} 供给到参数化重构部分100之前将干上混参数 \bar{C} 和湿上混参数 \bar{P} 例如从熵编码格式去量化。如参照图4描述的,量化可能已经被以两个不同的步长大小(例如,0.1或0.2)中的一个执行。所利用的实际步长大小可以是预定义的,或者可以例如经由比特流B从编码器侧用信号通知给音频解码系统200。在一些示例实施例中,干上混系数C和湿上混系数P可以分别从已经在相应的去量化部分204和205中的干上混参数 \bar{C} 和湿上混参数 \bar{P} 导出,该去量化部分204和205可以可选地被认为分别是干上混部分102和湿上混部分103的一部分。在本示例实施例中,由参数化重构部分100输出的重构音频信号 \hat{x} 在被作为音频解码系统200的输出提供以供在多扬声器系统207上回放之前被QMF合成部分206从QMF域变换回去。

[0092] 图5-11示出根据示例实施例的通过下混声道表示11.1声道音频信号的替代方式。在本示例实施例中,11.1声道音频信号包括以下声道:左(L)、右(R)、中心(C)、低频效果(LFE)、左侧(LS)、右侧(RS)、左后(LB)、右后(RB)、顶部左前(TFL)、顶部右前(TFR)、顶部左后(TBL)和顶部右后(TBR),这些在图5-11中由大写字母指示。表示11.1声道音频信号的替代方式对应于替代地将声道划分为多组声道,每一组由单个下混信号(可选地由相关联的湿上混参数和干上混参数)表示。多组声道中的每一组到其相应的单声道下混信号(和元数据)的编码可以独立地并且并行地执行。类似地,相应的多组声道从其相应的单声道下混信号的重构可以独立地并且并行地执行。

[0093] 要理解,在参照图5-11(以及以下还参照图13-16)描述的示例实施例中,没有一个重构声道可以包括来自多于一个的下混声道以及从该单个下混信号导出的任何去相关信号的贡献,即,来自多个下混声道的贡献在参数化重构期间不被组合/混合。

[0094] 在图5中,声道LS、TBL和LB形成由单个下混声道 I_s (及其相关联的元数据)所表示的声道组501。参照图3描述的参数化编码部分300可以以 $N=3$ 被利用,以通过单个下混声道 I_s 以及相关联的干上混参数和湿上混参数来表示三个音频声道LS、TBL和LB。假定预定义矩阵 V 和中间矩阵 H_R 的预定义矩阵类(两者都与在参数化编码部分300中执行的编码相关联)在解码器侧是已知的,则参照图1描述的参数化重构部分100可以被利用以从下混信号 I_s 以及相关联的干上混参数和湿上混参数重构三个声道LS、TBL和LB。类似地,声道RS、TBR和RB形成由单个下混声道 r_s 所表示的声道组502,并且参数化编码部分300的另一实例可以与第一编码部分并行地被利用以通过单个下混声道 r_s 以及相关联的干上混参数和湿上混参数表示三个声道RS、TBR和RB。而且,假定预定义矩阵 V 和中间矩阵 H_R 所属于的预定义矩阵类(两者都与参数化编码部分300的第二实例相关联)在解码器侧是已知的,则参数化重构部

分100的另一实例可以与第一参数化重构部分并行地被利用以从下混信号rs以及相关联的干上混参数和湿上混参数重构三个声道RS、TBR和RB。另一声道组503仅包括由下混声道I所表示的两个声道L和TFL。这两个声道到下混声道I以及相关联的湿上混参数和干上混参数的编码可以分别由与参照图3和图1描述的编码部分和重构部分类似的编码部分和重构部分执行,但是是针对 $N=2$ 。另一声道组504仅包括由下混声道Ife所表示的单个声道LFE。在该情况下,不需要下混,并且下混声道Ife可以是声道LFE本身,可选地被变换到MDCT域中和/或使用感知音频编解码器被编码。

[0095] 在图5-11中被利用以表示11.1声道音频信号的下混声道的总数有所变化。例如,图5中所示的示例利用6个下混声道,而图7中的示例利用10个下混声道。不同的下混配置可以适合于不同的情形,例如取决于用于传输下混信号和相关联的上混参数的可用带宽、和/或对11.1声道音频信号的重构应当达到的忠实程度的要求。

[0096] 根据示例实施例,参照图4描述的音频编码系统400可以包括多个参数化编码部分,该参数化编码部分包括参照图3描述的参数化编码部分300。音频编码系统400可以包括控制部分(图4中未示出),该控制部分被配置为从与图5-11中所示的11.1声道音频信号的相应划分对应的编码格式的集合确定/选择用于11.1声道音频信号的编码格式。该编码格式进一步对应于用于计算相应的下混声道的一组预定义规则(其中的至少一些可以一致)、用于中间矩阵 H_R 的一组预定义矩阵类(其中的至少一些可以一致)、以及用于基于相应的相关联的湿上混参数来获得与相应的多组声道中的至少一些相关联的湿上混系数的一组预定义矩阵 V (其中的至少一些可以一致)。根据本示例实施例,音频编码系统被配置为使用所述多个编码部分的适合于确定的编码格式的子集来对11.1声道音频信号进行编码。如果例如确定的编码格式对应于图1中所示的11.1声道的划分,则编码系统可以利用被配置用于通过相应的单个下混声道表示相应的多组3个声道的2个编码部分、被配置用于通过相应的单个下混声道表示相应的多组2个声道的2个编码部分、以及被配置用于将相应的单个声道表示为相应的单个下混声道的2个编码部分。所有的下混信号以及相关联的湿上混参数和干上混参数可以被编码在同一个比特流B中,以供传输到解码器侧。要注意,伴随下混声道的元数据(即,湿上混参数和湿上混参数)的紧凑格式可以被编码部分中的一些利用,而在至少一些示例实施例中,其它元数据格式可以被利用。例如,编码部分中的一些可以输出全部数量的湿上混系数和干上混系数,而不是湿上混参数和干上混参数。还设想以下实施例:在这些实施例中,一些声道被编码以供利用少于 $N-1$ 个去相关声道(或者甚至根本不利用去相关)进行重构,并且在这些实施例中用于参数化重构的元数据因此可以采取不同的形式。

[0097] 根据示例实施例,参照图2描述的音频解码系统200可以包括对应的多个重构部分,该重构部分包括参照图1描述的用于重构由相应的下混信号所表示的11.1声道音频信号的相应的多组声道的参数化重构部分100。音频解码系统200可以包括被配置为从编码器侧接收指示确定的编码格式的信号的信令的控制部分(图2中未示出),并且音频解码系统200可以利用所述多个重构部分的适当子集以从接收的下混信号以及相关联的干上混参数和湿上混参数重构11.1声道音频信号。

[0098] 图12-13示出根据示例实施例的通过下混声道表示13.1声道音频信号的替代方式。13.1声道音频信号包括以下声道:左屏幕(LSCRN)、左宽(LW)、右屏幕(RSCRN)、右宽(RW)、中心(C)、低频效果(LFE)、左侧(LS)、右侧(RS)、左后(LB)、右后(RB)、顶部左前(TFL)、

顶部右前 (TFR)、顶部左后 (TBL) 和顶部右后 (TBR)。将相应的声道组编码为相应的下混声道可以由如以上参照图5-11描述的独立并行地操作的相应的编码部分执行。类似地,基于相应的下混声道和相关联的上混参数对相应的声道组的重构可以由独立并行地操作的相应的重构部分执行。

[0099] 图14-16示出根据示例实施例的通过下混声道表示22.2声道音频信号的替代方式。22.2声道音频信号包括以下声道:低频效果1 (LFE1)、低频效果2 (LFE2)、底部前中 (BFC)、中心 (C)、顶部前中 (TFC)、左宽 (LW)、底部左前 (BFL)、左 (L)、顶部左前 (TFL)、顶侧左 (TSL)、顶部左后 (TBL)、左侧 (LS)、左后 (LB)、顶部中心 (TC)、顶部中后 (TBC)、中后 (CB)、底部右前 (BFR)、右 (R)、右宽 (RW)、顶部右前 (TFR)、顶侧右 (TSR)、顶部右后 (TBR)、右侧 (RS) 和右后 (RB)。图16中所示的22.2声道音频信号的划分包括声道组1601,其包括四个声道。参照图3描述的、但是以 $N=4$ 实现的参数化编码部分300可以被利用以将这些声道编码为下混信号以及相关联的湿上混参数和干上混参数。类似地,参照图1描述的、但是以 $N=4$ 实现的参数化重构部分100可以被利用以从下混信号以及相关联的湿上混参数和干上混参数重构这些声道。

[0100] III. 等同、扩展、替代和其它

[0101] 在研究以上描述之后,本公开的进一步的实施例对于本领域技术人员将变得清楚。即使目前的描述和附图公开了实施例和示例,但本公开也不限于这些具体示例。在不脱离由随附权利要求限定的本公开的范围的情况下,可以进行许多修改和变型。在权利要求中出现的任何附图标记都不应被理解为限制它们的范围。

[0102] 另外,对公开的实施例的变型可以由技术人员在实施本公开时从附图、公开和所附权利要求的研究来理解和实现。在权利要求中,词语“包括”不排除其它元件或步骤,并且不定冠词“一个”不排除多个。仅有的某些措施在互不相同的从属权利要求中被记载的事实并不表明这些措施的组合不能被用于获利。

[0103] 在上文中公开的设备和方法可以被实现为软件、固件、硬件或其组合。在硬件实现中,在以上描述中提及的功能单元之间的任务的划分不一定对应于划分成物理单元;相反,一个物理组件可以具有多个功能,并且一个任务可以由若干物理组件合作执行。某些组件或全部组件可以被实现为由数字信号处理器或微处理器执行的软件,或者被实现为硬件或专用集成电路。这样的软件可以分发给计算机可读介质上,该计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。如本领域技术人员公知的,术语计算机存储介质包括以存储信息(诸如计算机可读指令、数据结构、程序模块或其它数据)的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质两者。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪速存储器或其它存储器技术、CD-ROM、数字多功能盘(DVD)或其它光盘存储、磁盒、磁带、磁盘存储或其它磁存储设备、或者可以被用于存储期望信息并且可以被计算机访问的任何其它介质。此外,技术人员公知的是,通信介质通常包含计算机可读指令、数据结构、程序模块、或调制数据信号(诸如载波或其它输送机制)中的其它数据,并且包括任何信息递送介质。

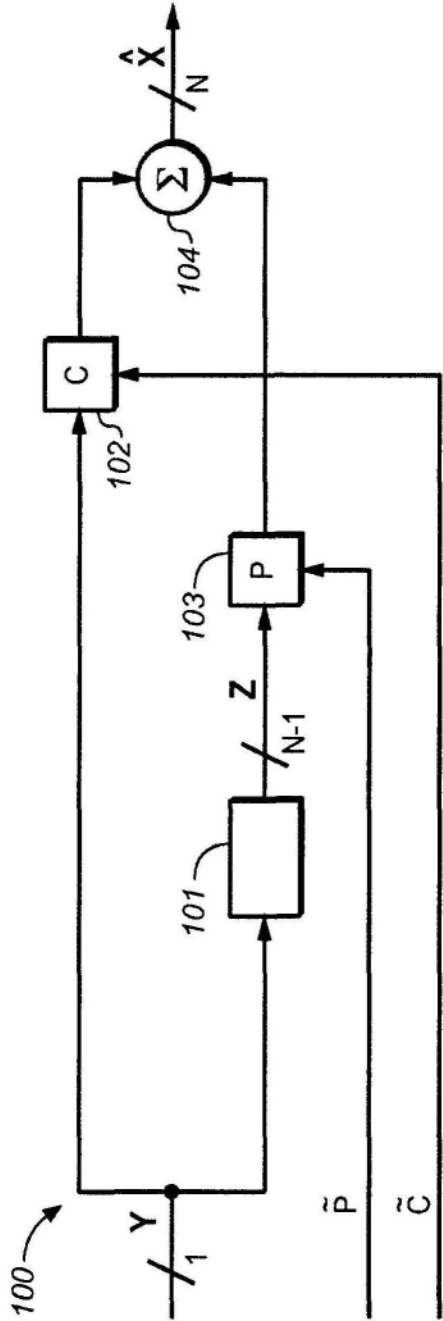


图1

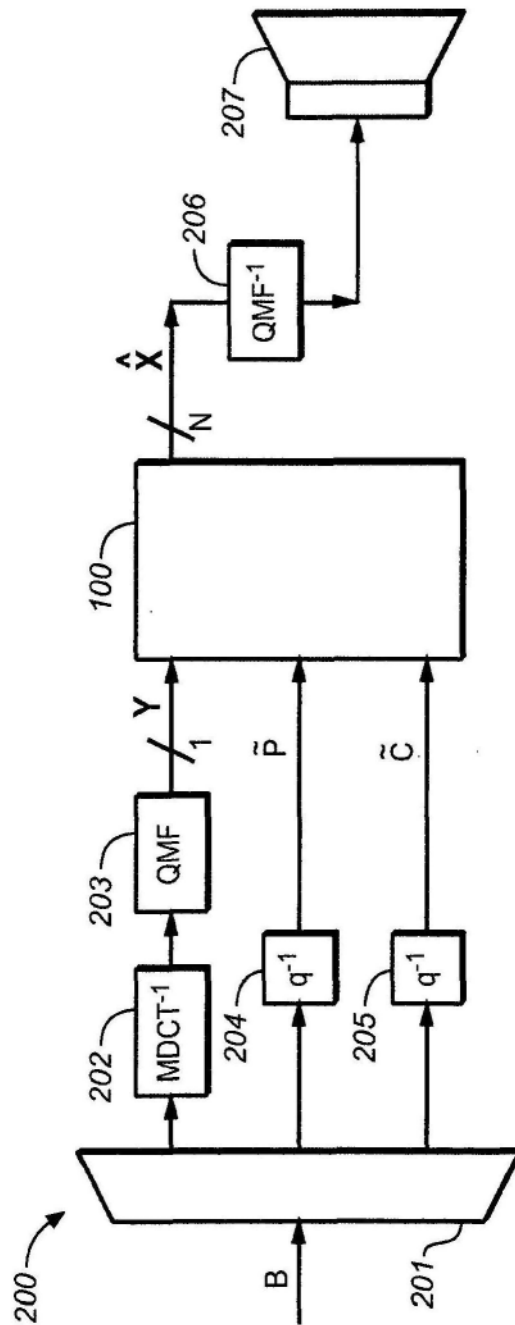


图2

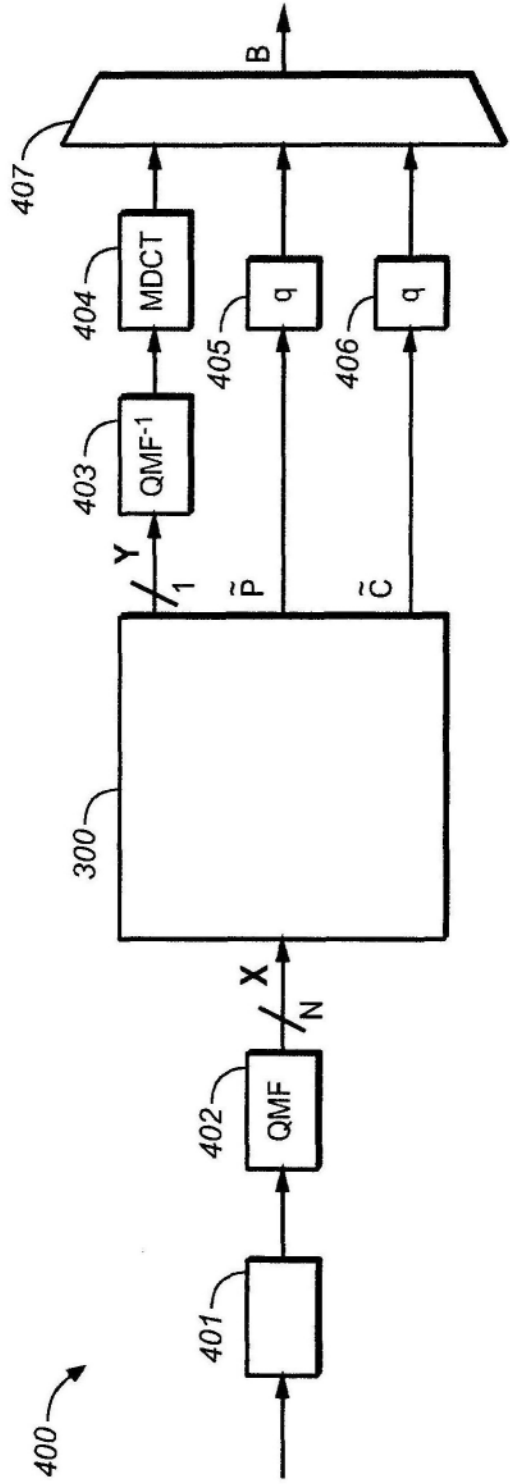


图4

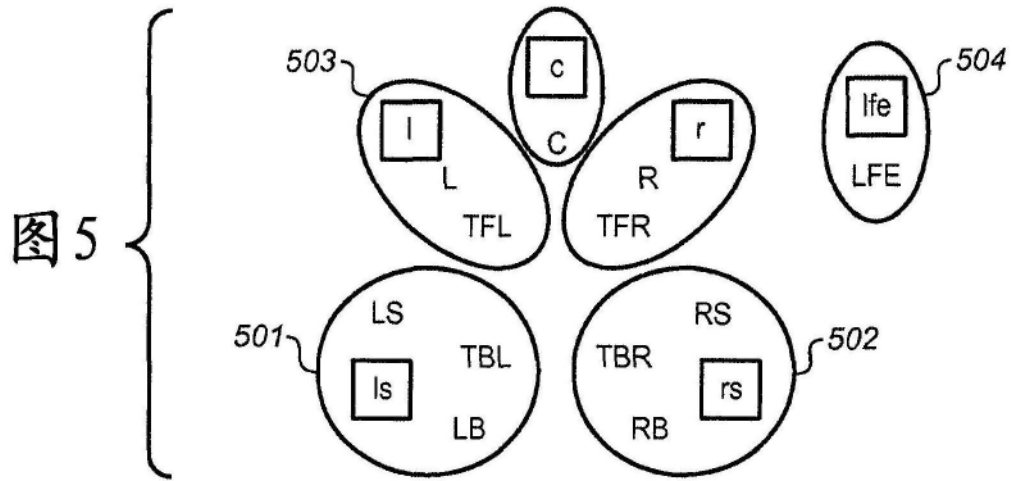


图5

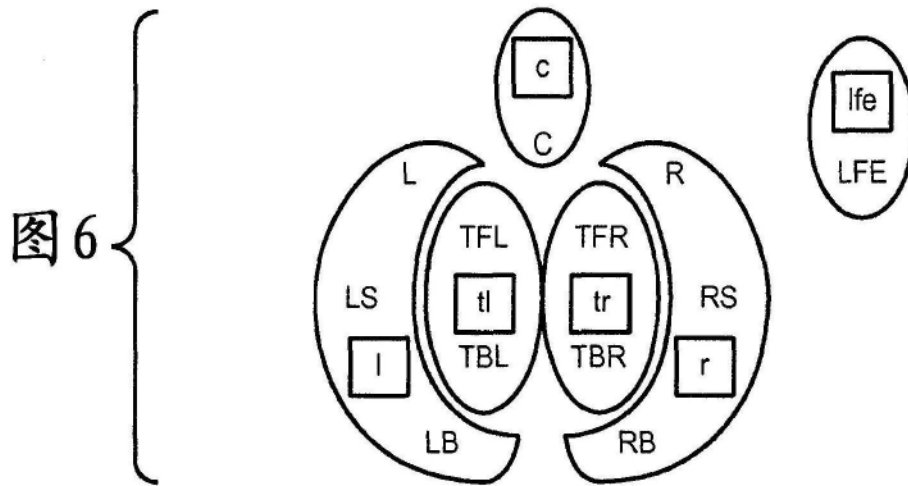


图6

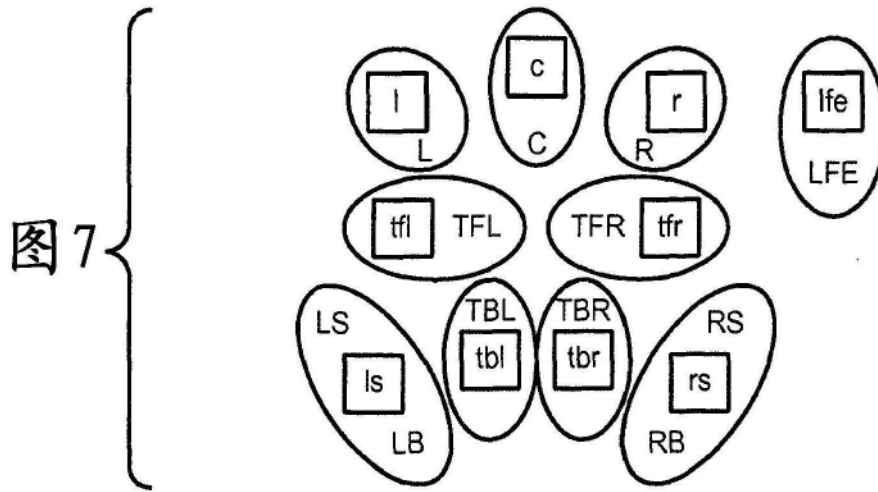


图7

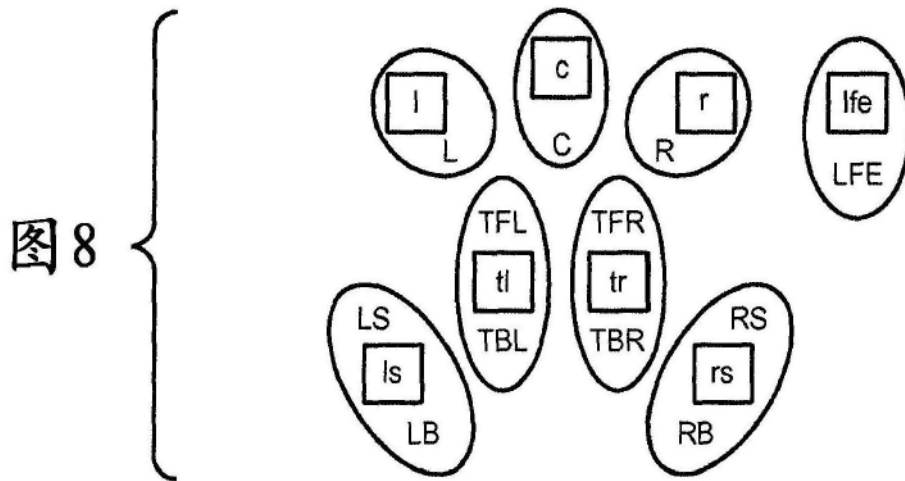


图8

图 9

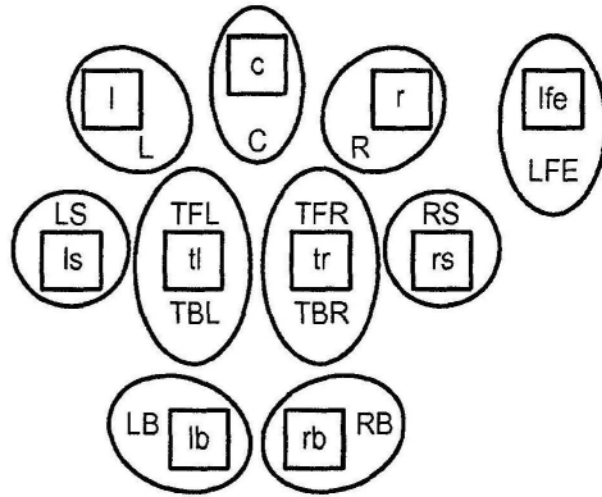


图9

图 10

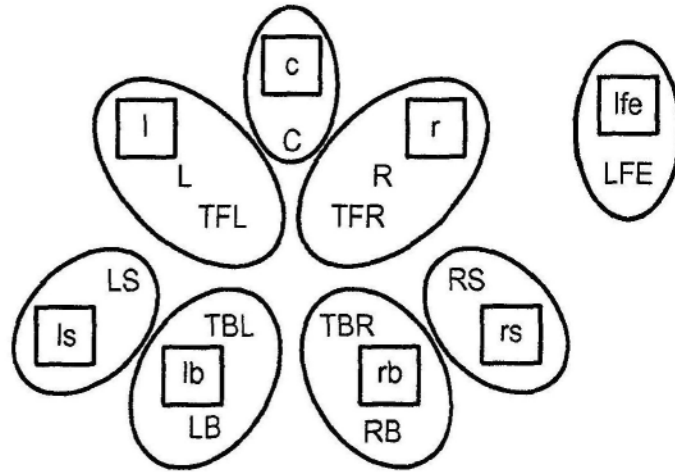


图10

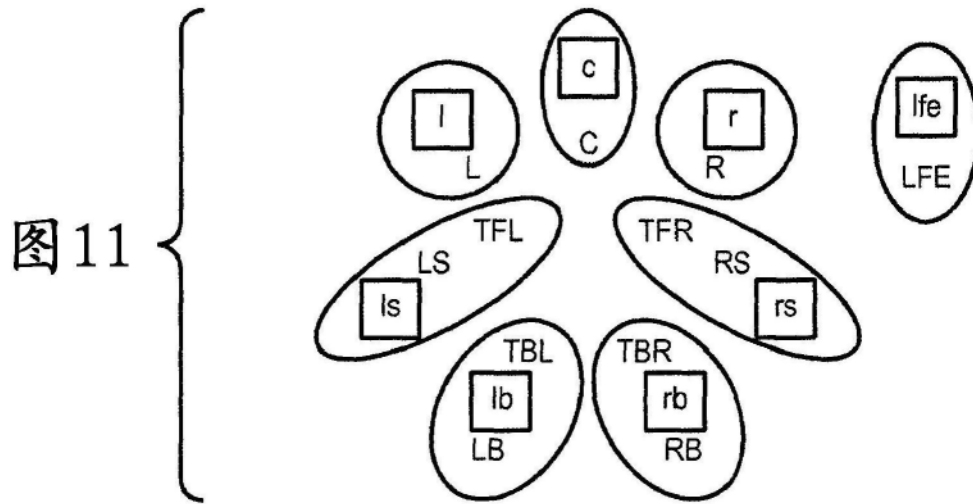


图11

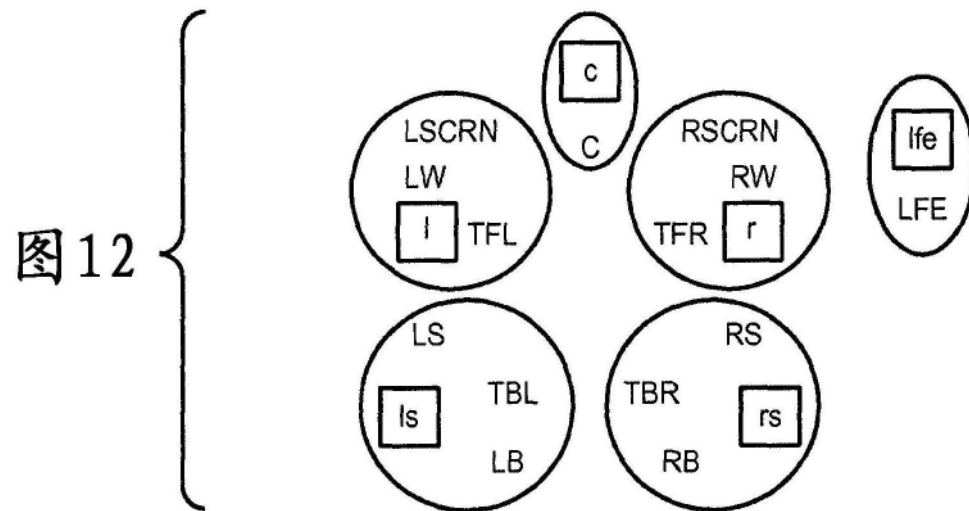


图12

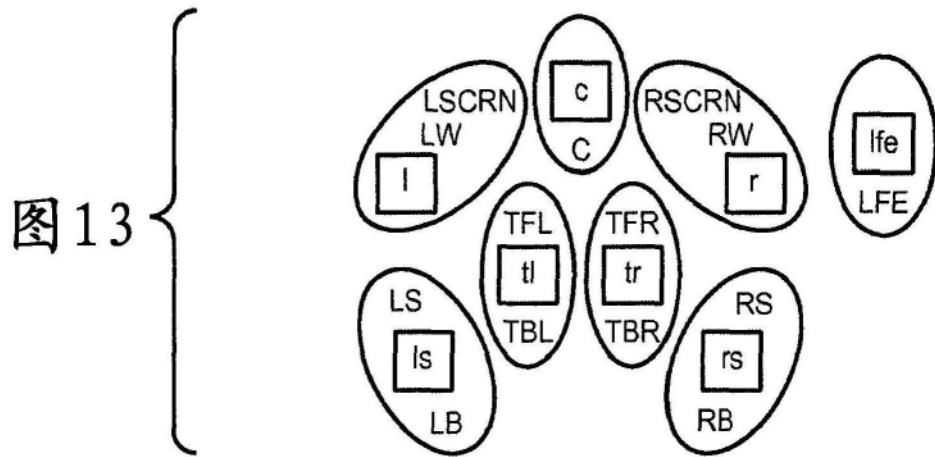


图13

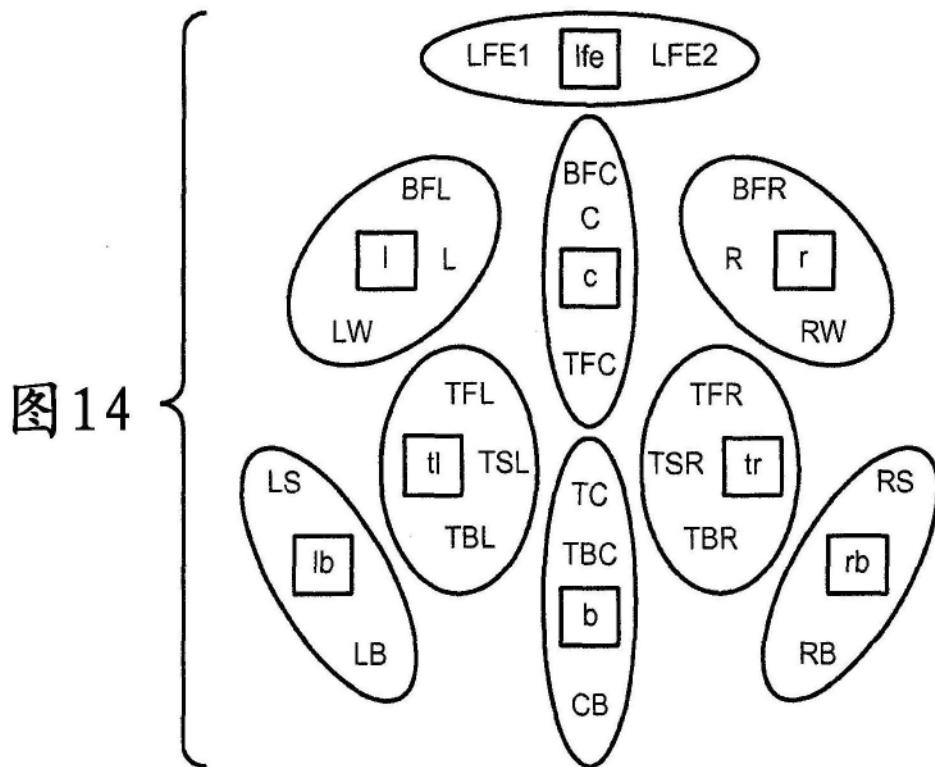


图14

图 15

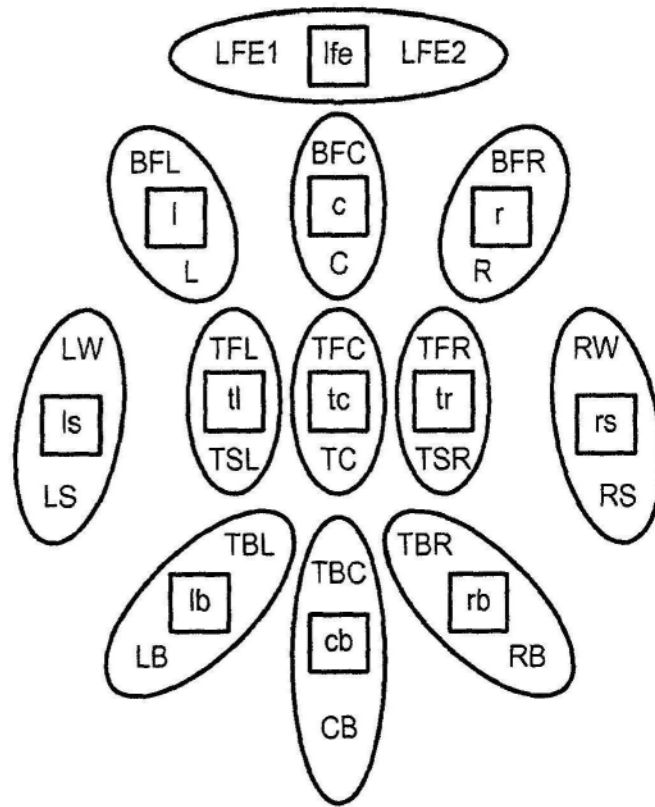


图15

图 16

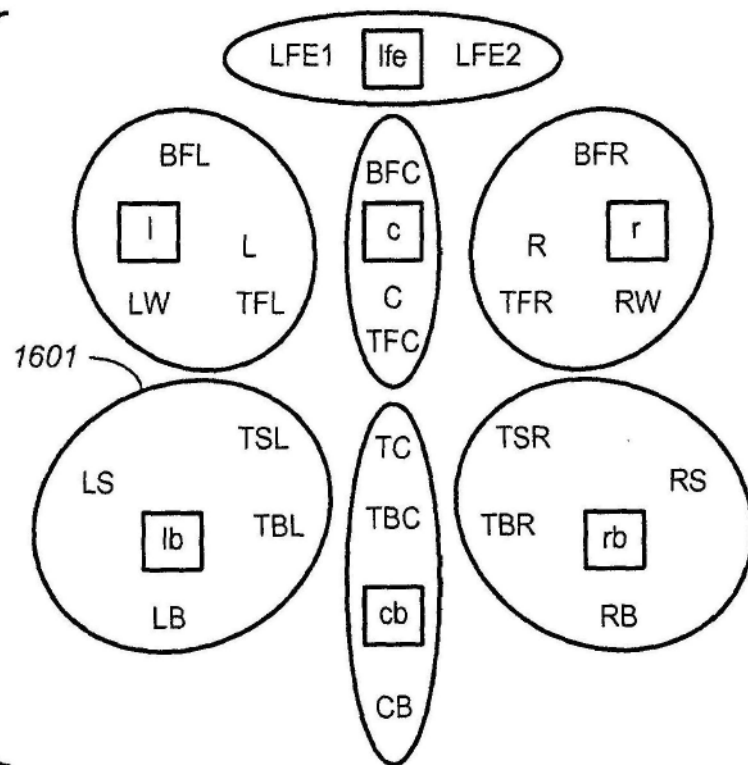


图16