



(12)发明专利

(10)授权公告号 CN 103854651 B

(45)授权公告日 2017.04.12

(21)申请号 201410084189.7

G10L 21/038(2013.01)

(22)申请日 2010.12.14

(56)对比文件

(65)同一申请的已公布的文献号
申请公布号 CN 103854651 A

KR 20070043651 A,2007.04.25,

CN 1272259 A,2000.11.01,

CN 101540171 A,2009.09.23,

(43)申请公布日 2014.06.11

US 2009226010 A1,2009.09.10,

(30)优先权数据

CN 101292285 A,2008.10.22,

61/286,912 2009.12.16 US

CN 101484936 A,2009.07.15,

(62)分案原申请数据

US 2006140412 A1,2006.06.29,

201080053083.0 2010.12.14

CN 1926610 A,2007.03.07,

CN 1878001 A,2006.12.13,

(73)专利权人 杜比国际公司

CN 1969317 A,2007.05.23,

地址 荷兰阿姆斯特丹

陈茜.“基于SH-Mobile的HE-AAC v2解码设计与实现”.《中国优秀硕士学位论文全文数据库信息科技辑》.2008,(第09期),

(72)发明人 K·克约尔林 R·特辛

审查员 董小东

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 欧阳帆

(51)Int.Cl.

G10L 19/008(2013.01)

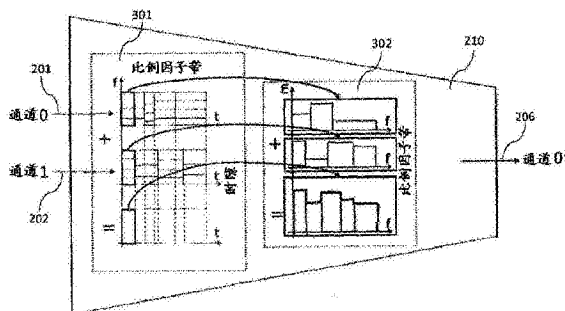
权利要求书5页 说明书17页 附图6页

(54)发明名称

SBR比特流参数缩混

(57)摘要

本发明涉及SBR比特流参数缩混。描述了用于将频带复制参数的第一和第二源集合合并为SBR参数的目标集合的方法和系统。第一和第二源集合分别包括彼此不同的第一和第二频带划分。第一源集合包括与第一频带划分的频带关联的第一能量相关值集合。第二源集合包括与第二频带划分的频带关联的第二能量相关值集合。目标集合包括与基本频带关联的目标能量相关值。该方法包括：将第一和第二频带划分分解为包括基本频带的联合网格；对基本频带分配第一能量相关值集合的第一值；对基本频带分配第二能量相关值集合的第二值；组合第一值和第二值以产生用于基本频带的目标能量相关值。



1. 一种用于将SBR参数的第一源集合和第二源集合合并为SBR参数的目标集合的方法，其中

- 所述第一源集合包括第一起始频率；
- 所述第二源集合包括第二起始频率；
- 所述第一起始频率和第二起始频率不同，并且分别同与SBR参数的所述第一源集合和第二源集合相关联的第一高带信号和第二高带信号的频率下限相关联；以及

其中所述方法包括：

- 比较所述第一起始频率和第二起始频率；
- 选择所述第一起始频率和第二起始频率中较高的起始频率作为所述目标集合的起始频率，其中

-所述第一源集合与第一源通道的第一低带信号相关联，并且包括第一比例因子能量集合；

-所述第二源集合与第二源通道的第二低带信号相关联，并且包括第二比例因子能量集合；

-所述目标集合与根据所述第一低带信号和第二低带信号的时域缩混而获得的目标通道的目标低带信号相关联；以及

-所述目标集合包括目标比例因子能量集合；

其中，所述方法进一步包括：

-通过能量补偿因子对第一缩混系数和第二缩混系数进行加权；其中所述第一缩混系数与所述第一源通道相关联；其中所述第二缩混系数与所述第二源通道相关联；其中所述能量补偿因子与时域缩混期间所述第一低带信号和第二低带信号的相互作用相关联；

-通过加权的第一缩混系数来缩放所述第一比例因子能量集合；

-通过加权的第二缩混系数来缩放所述第二比例因子能量集合；以及

-根据经缩放的第一比例因子能量集合和经缩放的第二比例因子能量集合来确定所述目标比例因子能量集合。

2. 如权利要求1所述的方法，其中

-所述第一源集合包括第一SBR元素头部，所述第一SBR元素头部包括所述第一起始频率；

-所述第二源集合包括第二SBR元素头部，所述第二SBR元素头部包括所述第二起始频率；

其中，所述方法还包括：

-根据所述目标集合的所选择的起始频率，基于所述第一SBR元素头部或者所述第二SBR元素头部来选择所述目标集合的SBR元素头部。

3. 如权利要求1所述的方法，其中所述能量补偿因子同所述目标低带信号的能量与所述第一低带信号和第二低带信号的组合能量的比例相关联。

4. 一种用于将SBR参数的第一源集合和第二源集合合并为SBR参数的目标集合的方法，其中

-所述第一源集合包括第一起始频率；

-所述第二源集合包括第二起始频率；

-所述第一起始频率和第二起始频率不同,并且分别同与SBR参数的所述第一源集合和第二源集合相关联的第一高带信号和第二高带信号的频率下限相关联;以及

其中所述方法包括:

-比较所述第一起始频率和第二起始频率;

-选择所述第一起始频率和第二起始频率中较高的起始频率作为所述目标集合的起始频率,其中

-所述第一源集合包括第一瞬态包络指数;其中所述第一瞬态包络指数标识具有第一起始时间边界的第一瞬态包络;

-所述第二源集合包括第二瞬态包络指数;其中所述第二瞬态包络指数标识具有第二起始时间边界的第二瞬态包络;

-所述目标集合包括多个目标包络,每个目标包络具有起始时间边界;

-所述第一瞬态包络、所述第二瞬态包络和所述多个目标包络分别与第一源信号、第二源信号和目标信号的一个或多个时间间隔相关联;

其中,所述方法进一步包括:

-选择所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界;

-将所述多个目标包络中起始时间边界最接近所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界的包络确定为目标瞬态包络;以及

-设置目标瞬态包络指数,以标识所述目标瞬态包络。

5. 如权利要求4所述的方法,其中所述确定步骤包括将所述多个目标包络中起始时间边界最接近所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界但是不晚于所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界的包络确定为目标瞬态包络。

6. 如权利要求1-5中的任一项所述的方法,其中SBR参数的每个源集合对应于与HE-AAC比特流的通道相关联的SBR参数。

7. 一种用于将SBR参数的N个源集合合并为SBR参数的M个目标集合的方法,其中

-N大于2;

-M小于N;

所述方法包括:

-合并一对源集合以产生中间集合,和与较低声学相关性的源通道相对应的源集合相比,与较高声学相关性的源通道相对应的所述一对源集合更少被合并;以及

-合并所述中间集合与源集合或者另一中间集合,以产生目标集合,

其中,合并步骤是根据权利要求1-6中任一项所述的方法执行的。

8. 一种用于将SBR参数的第一源集合和第二源集合合并为SBR参数的目标集合的设备,其中

-所述第一源集合包括第一起始频率;

-所述第二源集合包括第二起始频率;

-所述第一起始频率和第二起始频率不同,并且分别同与SBR参数的所述第一源集合和第二源集合相关联的第一高带信号和第二高带信号的频率下限相关联;以及

其中所述设备包括:

用于比较所述第一起始频率和第二起始频率的装置；

用于选择所述第一起始频率和第二起始频率中较高的起始频率作为所述目标集合的起始频率的装置,其中

-所述第一源集合与第一源通道的第一低带信号相关联,并且包括第一比例因子能量集合;

-所述第二源集合与第二源通道的第二低带信号相关联,并且包括第二比例因子能量集合;

-所述目标集合与根据所述第一低带信号和第二低带信号的时域缩混而获得的目标通道的目标低带信号相关联;以及

-所述目标集合包括目标比例因子能量集合;

其中,所述设备进一步包括:

用于通过能量补偿因子对第一缩混系数和第二缩混系数进行加权的装置;其中所述第一缩混系数与所述第一源通道相关联;其中所述第二缩混系数与所述第二源通道相关联;其中所述能量补偿因子与时域缩混期间所述第一低带信号和第二低带信号的相互作用相关联;

用于通过加权的第一缩混系数来缩放所述第一比例因子能量集合的装置;

用于通过加权的第二缩混系数来缩放所述第二比例因子能量集合的装置;以及

用于根据经缩放的第一比例因子能量集合和经缩放的第二比例因子能量集合来确定所述目标比例因子能量集合的装置。

9. 如权利要求8所述的设备,其中

-所述第一源集合包括第一SBR元素头部,所述第一SBR元素头部包括所述第一起始频率;

-所述第二源集合包括第二SBR元素头部,所述第二SBR元素头部包括所述第二起始频率;

其中,所述设备还包括:

用于根据所述目标集合的所选择的起始频率,基于所述第一SBR元素头部或者所述第二SBR元素头部来选择所述目标集合的SBR元素头部的装置。

10. 如权利要求8所述的设备,其中所述能量补偿因子同所述目标低带信号的能量与所述第一低带信号和第二低带信号的组合能量的比例相关联。

11. 一种用于将SBR参数的第一源集合和第二源集合合并为SBR参数的目标集合的设备,其中

-所述第一源集合包括第一起始频率;

-所述第二源集合包括第二起始频率;

-所述第一起始频率和第二起始频率不同,并且分别同与SBR参数的所述第一源集合和第二源集合相关联的第一高带信号和第二高带信号的频率下限相关联;以及

其中所述设备包括:

用于比较所述第一起始频率和第二起始频率的装置;

用于选择所述第一起始频率和第二起始频率中较高的起始频率作为所述目标集合的起始频率的装置,其中

-所述第一源集合包括第一瞬态包络指数;其中所述第一瞬态包络指数标识具有第一起始时间边界的第一瞬态包络;

-所述第二源集合包括第二瞬态包络指数;其中所述第二瞬态包络指数标识具有第二起始时间边界的第二瞬态包络;

-所述目标集合包括多个目标包络,每个目标包络具有起始时间边界;

-所述第一瞬态包络、所述第二瞬态包络和所述多个目标包络分别与第一源信号、第二源信号和目标信号的一个或多个时间间隔相关联;

其中,所述设备进一步包括:

用于选择所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界的装置;

用于将所述多个目标包络中起始时间边界最接近所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界的包络确定为目标瞬态包络的装置;以及

用于设置目标瞬态包络指数,以标识所述目标瞬态包络的装置。

12. 如权利要求11所述的设备,其中所述用于确定的装置包括用于将所述多个目标包络中起始时间边界最接近所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界但是不晚于所述第一起始时间边界和第二起始时间边界中较早的一个起始时间边界的包络确定为目标瞬态包络的装置。

13. 如权利要求8-12中的任一项所述的设备,其中SBR参数的每个源集合对应于与HE-AAC比特流的通道相关联的SBR参数。

14. 一种用于将SBR参数的N个源集合合并为SBR参数的M个目标集合的设备,其中

-N大于2;

-M小于N;

所述设备包括:

用于合并一对源集合以产生中间集合的装置,和与较低声学相关性的源通道相对应的源集合相比,与较高声学相关性的源通道相对应的所述一对源集合更少被合并;以及

用于合并所述中间集合与源集合或者另一中间集合以产生目标集合的装置,

其中,合并是根据权利要求1-6中任一项所述的方法执行的。

15. 一种SBR参数合并单元,其配置成根据SBR参数的N个源集合来提供SBR参数的M个目标集合,其中 $N > M \geq 1$,所述SBR参数合并单元包括处理器,所述处理器配置成执行如权利要求1至7中任一项所述的方法步骤。

16. 一种音频解码器,其配置成解码包括N个音频通道的HE-AAC比特流,所述音频解码器包括:

-AAC解码器,其配置成接收经编码的HE-AAC比特流,并且提供独立的SBR比特流;

-SBR解码器,其配置成根据所述SBR比特流来提供与所述N个音频通道相对应的SBR参数的N个源集合;以及

-根据权利要求15所述的SBR参数合并单元,其配置成根据SBR参数的所述N个源集合来提供SBR参数的M个目标集合,其中 $N > M \geq 1$ 。

17. 如权利要求16所述的音频解码器,其中所述AAC解码器还配置成提供与所述N个音频通道相对应的N个时域低带音频信号;并且其中所述音频解码器还包括:

-时域缩混单元,其配置成根据所述N个时域低带音频信号来提供M个时域低带音频信号;以及

-SBR单元,其配置成根据所述M个低带音频信号和SBR参数的所述M个目标集合来生成M个高带音频信号;

其中所述音频解码器配置成提供分别包括所述M个低带音频信号和所述M个高带音频信号的M个音频信号。

18.一种音频转码器,其配置成根据包括N个音频通道的HE-AAC比特流来提供包括M个音频通道的HE-AAC比特流,其中 $N > M \geq 1$,所述音频转码器包括:

-根据权利要求15所述的SBR参数合并单元。

19.一种电子设备,其配置成根据包括N个音频通道的HE-AAC比特流来呈现与M个通道相对应的M个音频信号,其中 $N > M \geq 1$,所述电子设备包括:

-音频呈现装置,其配置成执行所述M个音频信号的声学呈现;

-接收机,其配置成接收经编码的HE-AAC比特流;以及

-根据权利要求16至17中的任一项所述的音频解码器,其配置成根据所述HE-AAC比特流来提供所述M个音频信号。

SBR比特流参数缩混

[0001] 本分案申请是基于申请号为201080053083.0(国际申请号为PCT/EP2010/069651),申请日为2010年12月14日,发明名称为“SBR比特流参数缩混”的中国专利申请的分案申请。

技术领域

[0002] 本文档涉及音频解码和/或音频转码。特别地,本文档涉及用于根据包括较高数目N个音频通道的比特流而对数目M个音频通道进行有效解码的方案。

背景技术

[0003] 符合高效高级音频编码(HE-AAC)标准的音频解码器通常被设计为解码并且输出要由预定义位置处的独立扬声器再现的多至N个音频数据通道。经HE-AAC编码的比特流通常包括有关与N个音频通道相对应的N个低带信号的数据,以及用于与各低带信号相对应的N个高带信号的重构的经编码的SBR(频带复制)参数。

[0004] 在某些情况下,可期望的是,HE-AAC解码器将输出通道的数目减少为M个通道(M小于N),同时保留来自全部N个通道的音频事件。此类通道减少的一个示例性使用情况是在连接至多通道家庭影院系统时可以回放N个通道的移动设备,但是其在单独使用时限于其内置的单通道或者立体声输出。

[0005] 根据N个输入或者源通道产生M个输出或者目标通道的一种可能方式是经解码的N通道信号的时域缩混(downmix)。在这种系统中,首先对表示N个通道的经编码的比特流进行解码以产生N个时域音频信号,该N个时域音频信号随后在时域中被缩混为与M个通道相对应的M个音频信号。这一解决方案的缺点是首先解码与N个通道相对应的全部N个音频信号以及随后将N个经解码的音频信号缩混为M个经缩混的音频信号所需要的计算和存储器资源的量。

[0006] ETSI技术规范(TS) 126402(3GPP TS26.402)在第6节中描述一种称为“SBR立体声参数到单通道参数的缩混”的方法。通过引用将该文献并入。ETSI技术规范描述了SBR参数合并过程以便根据SBR通道对来导出单SBR通道。然而,所指定的方法限于立体声到单通道的缩混,其中通道被表示为通道对元素(channel pair element,CPE)。

[0007] 鉴于以上所述,需要从任意数目N个通道到任意数目M个通道的低复杂度缩混方案。特别地,需要用于与N个通道相关联的SBR参数到与M个通道相关联的SBR参数的缩混方案,其中该缩混方案保留不同通道的相关高频信息。

发明内容

[0008] 在本文档中,描述了提供用于减少HE-AAC解码器中的输出或者目标通道的数目同时保留来自所有输入或者源通道的音频事件的有效方式的方法和系统。所述方法和系统允许从任意数目N个通道到任意数目M个通道的通道缩混,其中M小于N。与时域中的缩混相比,所述方法和系统可以按照降低的计算复杂度来实现。应当注意,所描述的方法和系统可应

用于使用针对高频再生的SBR的任何多通道解码器。特别地,所描述的方法和系统不限于经HE-AAC编码的比特流。另外,应当注意,针对将第一和第二源通道合并为目标通道而概述了以下方面。这些术语应被理解为“至少第一”和“至少第二”以及“至少目标”通道,并且因此应用于将任意数目N个源通道合并为任意数目M个目标通道。

[0009] 根据一个方面,描述了一种用于将频带复制(SBR)参数的第一和第二源集合合并为SBR参数的目标集合的方法。SBR参数的源集合可对应于与HE-AAC比特流的音频通道相关联的SBR参数。SBR参数的源集合和/或目标集合可与特定音频通道的音频信号的帧的SBR参数相对应。由此,第一源集合可与第一音频通道的第一音频信号相对应,第二源集合可与第二音频通道的第二音频信号相对应,并且目标集合可与目标通道的目标音频信号相对应。源集合和/或目标集合可以包括用于根据各音频信号的低频分量生成各音频信号的高频分量的数据。特别地,SBR参数的集合可以包括与各音频信号的帧的预定义时间间隔内的高频分量的谱包络有关的信息。包括在此类时间间隔内的谱信息通常称为包络。

[0010] 第一和第二源集合(尤其是第一和第二源集合的包络)可以分别包括第一和第二频带划分。这些第一和第二频带划分可以彼此不同。第一源集合可以包括与第一频带划分的频带相关联的第一能量相关值集合;并且第二源集合可以包括与第二频带划分的频带相关联的第二能量相关值集合。目标集合可以包括与基本频带相关联的目标能量相关值。

[0011] 此类能量相关值可以是比例因子能量,并且频带可以是比例因子带。可选地或者附加地,能量相关值可以是噪声基底比例因子能量,并且频带可以是噪声基底比例因子带。

[0012] 该方法可以包括步骤:将第一和第二频带划分分解成包括基本频带的联合网格。第一和第二频带划分可以跨越各音频信号的高频分量的频率范围。该频率范围可以被细分成联合频率网格。联合网格可以与用于确定SBR参数的正交镜像滤波器组(QMF滤波器组)相关联。特别地,QMF滤波器组可以在分析阶段使用,以便将各音频信号的高频分量的谱分割确定成QMF子带。这样的QMF子带可以是联合频率网格的基本频带。

[0013] 应当注意,第一频带划分可以跨越与第二频带划分不同的频率范围。特别地,第一频带划分的起始频率(即,第一频带划分的下限)可以与第二频带划分的起始频率(即,第二频带划分的下限)不同。通常,联合频率网格覆盖第一和第二频带划分的重叠频率范围。特别地,可以不考虑在起始频率中的较高的一个以下的频带或者频带的一个或多个部分。

[0014] 该方法可以包括:对基本频带分配第一能量相关值集合的第一值;和/或对基本频带分配第二能量相关值集合的第二值。第一分配步骤可以执行为使得第一值对应于与包括基本频带的第一频带划分的频带相关联的能量相关值。第二分配步骤可以执行为使得第二值对应于与包括基本频带的第二频带划分的频带相关联的能量相关值。

[0015] 该方法可以包括步骤:对第一和第二值进行组合(例如,相加和/或缩放),以产生用于基本频带的目标能量相关值。另外,可以通过贡献源集合的数目对目标能量相关值进行归一化。通过实例的方式,可以将目标能量相关值除以贡献源集合的数目,以便确定源集合的贡献能量相关值的平均值。

[0016] 已经针对特定基本频带而说明了上述方法。该方法可以包括附加步骤:针对联合网格的所有基本频带重复分配步骤和组合步骤,并从而产生目标集合的目标能量相关值的集合。

[0017] 目标集合可以包括具有预定义目标频带的目标频带划分。通常,这样的目标频带

具有单个相关联的目标能量相关值。为了确定该相关联的目标能量相关值,该方法可以包括步骤:将与包括在目标频带内的基本频带相关联的目标能量相关值的集合进行平均。平均值可以被分配为目标频带的目标能量相关值。

[0018] 第一源集合可以与第一源通道的第一信号相关联;和/或第二源集合可以与第二源通道的第二信号相关联;和/或目标集合可以与目标通道的目标信号相关联。通常,源集合和目标集合与相对应的信号的一定时间间隔相关联。此类时间间隔可以由所谓的包络限定。

[0019] 特别地,目标集合的目标能量相关值可以与目标信号的目标时间间隔相关联;和/或第一源集合的第一能量相关值集合可以与第一信号的第一时间间隔相关联,其中第一时间间隔可以与目标时间间隔重叠。在这种情况下,以上提到的组合步骤可以包括步骤:根据由第一时间间隔与目标时间间隔重叠的长度以及目标时间间隔的长度所给定的比例来缩放第一能量相关值集合的第一值。因此,可以对经缩放的第一值和第二值进行组合(例如,相加),以产生目标能量相关值。

[0020] 另外,第一源集合可以包括第三频带划分;和/或第一源集合可以包括与第三频带划分的频带相关联的第三能量相关值集合;和/或第三能量相关值集合可以与第一低带信号的第三时间间隔相关联,其中第三时间间隔可以与目标时间间隔重叠。应当注意,第三频带划分可以对应于(特别地,其可以等同于)第一频带划分。在这种情况下,该方法还可以包括步骤:将第三频带划分分解成包括基本频带的联合网格;和/或对基本频带分配第三能量相关值集合的第三值。在这种情况下,以上提到的组合步骤可以包括步骤:根据由第三时间间隔和目标时间间隔重叠的长度以及目标时间间隔的长度所给定的比例来缩放第三值。因此,可以对经缩放的第一值、第二值和经缩放的第三值进行组合(例如,相加),以产生目标能量相关值。

[0021] 根据另一方面,描述了一种用于将SBR参数的第一和第二源集合合并为SBR参数的目标集合的方法。第一源集合可以与第一源通道的第一低带信号相关联,并且可以包括第一比例因子能量集合。第二源集合可以与第二源通道的第二低带信号相关联,并且可以包括第二比例因子能量集合。目标集合可以与根据第一和第二低带信号的时域缩混而获得的目标通道的目标低带信号相关联。此外,目标集合可以包括目标比例因子能量集合。

[0022] 该方法可以包括步骤:利用能量补偿因子对第一和第二缩混系数进行加权;其中,第一缩混系数可以与第一源通道相关联;其中,第二缩混系数可以与第二源通道相关联;并且其中能量补偿因子可以与时域缩混期间第一与第二低带信号的相互作用相关联。这样的相互作用可以包括第一和第二低带信号的衰减和/或放大,这可能是由于第一和第二低带信号的同相或者反相行为而造成的。特别地,能量补偿因子可以同目标低带信号的能量与第一和第二低带信号的能量或者第一和第二低带信号的组合能量的比例相关联。

[0023] 通过实例的方式,在合并N个源通道, $N \geq 2$ 的情况下,为了获得M个目标通道, $M < N$ 并且 $M \geq 1$,能量补偿因子 f_{comp} 可以由以下给出:

$$[0024] \quad f_{comp} = \sqrt{\frac{\sum_{chout=0}^{M-1} \sum_n x_{dmx}^2[chout][n]}{\sum_{chin=0}^{N-1} \sum_n (c_{chin} \cdot x_{in}[chin][n])^2}}$$

[0025] 其中, $x_{in}[chin][n]$ 是源通道 $chin$ 中的低带时域信号, c_{chin} 是用于源通道 $chin$ 的缩混系数, $x_{dmx}[chout][n]$ 是目标通道 $chout$ 的低带时域信号, 并且 $n=0, \dots, 1023$ 是时域信号的帧内的信号样本的样本指数。应当注意, 可以基于时域信号的帧内的信号样本的子集来确定 f_{comp} 。由此, 可以例如使用帧的每第 P 个样本而在整个样本子集上计算上述总和, 其中 P 是整数, 即 $n=0, P, 2P, 3P, \dots$ 。

[0026] 该方法还可以包括步骤: 通过第一加权缩混系数来缩放第一比例因子能量集合; 和/或通过第二加权缩混系数来缩放第二能量集合。可以根据经缩放的第一比例因子能量集合和经缩放的第二比例因子能量集合来确定目标比例因子能量集合。特别地, 可以根据本档中描述的方法中的任一种来确定目标比例因子能量集合。

[0027] 根据另一方面, 描述了一种用于将SBR参数的第一和第二源集合合并为SBR参数的目标集合的方法。第一源集合可以包括第一起始频率。第二源集合可以包括第二起始频率。第一和第二起始频率可以不同, 并且它们可以分别同与SBR参数的第一和第二源集合相关联的第一和第二高带信号的频率下限相关联。特别地, 第一和第二起始频率可以与第一和第二频带划分的下限相关联。

[0028] 该方法可以包括步骤: 比较第一和第二起始频率; 和/或选择第一和第二起始频率中较高或较低的作为目标集合的起始频率。一般来说, 可以基于贡献源集合(例如, 第一和第二源集合)的起始频率的水平来选择目标集合的起始频率。

[0029] 起始频率选择可以用于确定目标集合的SBR元素(element)头部。第一源集合可以包括具有第一起始频率的第一SBR元素头部。第二源集合可以包括具有第二起始频率的第二SBR元素头部。在这样的情况下, 该方法可以包括步骤: 根据目标集合的所选择的起始频率, 基于第一或者第二SBR元素头部来选择目标集合的SBR元素头部。特别地, 可以选择包括较高或者较低起始频率的SBR元素头部作为用于确定目标集合的SBR元素头部的的基础。

[0030] 还可以将起始频率选择限制于具有特殊属性的源集合, 例如, 起始频率选择可以排他地或者优先地考虑某些源通道。特别地, 起始频率选择可以优先考虑展示出与目标通道的目标集合的期望关系相似的彼此关系的源通道的源集合。

[0031] 通过实例的方式, 如果目标集合是通道对元素, 并且源集合中的至少一个包括通道对元素, 则可以从包括通道对元素的源集合的一个中选择目标集合的SBR元素头部。如果目标集合是通道对元素, 并且没有源集合包括通道对元素, 则可以选择包括最高或者最低起始频率的源集合的SBR元素头部作为用于目标集合的SBR元素头部的的基础。如果目标集合是单个通道元素, 并且源集合中的至少一个是单个通道元素, 则可以选择目标集合的SBR元素头部作为包括单个通道元素的源集合中的一个的SBR元素头部。如果目标集合是单个通道元素, 并且所有的源集合是通道对元素, 则可以使用包括最高或者最低起始频率的源集合的SBR元素头部作为用于目标集合的SBR元素的基础。

[0032] 根据另一方面,描述了一种用于将SBR参数的第一和第二源集合合并为SBR参数的目标集合的方法。第一源集合可以包括第一瞬态包络指数;其中第一瞬态包络指数标识具有第一起始时间边界的第一瞬态包络。第二源集合可以包括第二瞬态包络指数;其中第二瞬态包络指数标识具有第二起始时间边界的第二瞬态包络。目标集合可以包括多个目标包络,每个目标包络具有起始时间边界。

[0033] 如上所述,包络(即,尤其是第一瞬态包络、第二瞬态包络和多个目标包络)可以分别与相应的音频信号(即,尤其是第一源信号、第二源信号和目标信号)的一个或多个时间间隔相关联。特别地,包络可以与各音频信号的帧内的一个或多个时间间隔相关联。瞬态包络指数可以用于标识包括与声学瞬变有关的信息的包络。

[0034] 该方法可以包括步骤:选择第一和第二起始时间边界中较早的一个;和/或将多个目标包络中起始时间边界最接近第一和第二起始时间边界中较早的一个的包络确定为目标瞬态包络;和/或设置目标瞬态包络指数以标识目标瞬态包络。在一个实施方案中,该方法可以包括步骤:将多个目标包络中起始时间边界最接近第一和第二起始时间边界中较早的一个并且不晚于第一和第二起始时间边界中较早的一个的包络确定为目标瞬态包络。

[0035] 根据另一方面,描述了一种用于将SBR参数的N个源集合合并为SBR参数的M个目标集合的方法。N可以大于2,并且M可以小于N。该方法可以包括步骤:合并一对源集合,以产生中间集合;和/或将中间集合与源集合或者另一中间集合合并,以产生目标集合。由此,该方法可以包括后续的合并步骤,并从而提供用于将SBR参数的N个源集合合并为SBR参数的M个目标集合的层级式方法。可以根据本文档中描述的方法和方面的任一种来执行合并步骤。在一个实施方案中,与较高声学相关性的源通道相对应的源集合比与较低声学相关性的源通道相对应的源集合的合并较少。

[0036] 根据另一方面,描述了一种软件程序。该软件程序可以适于在处理器上执行,并且当在计算设备上实施时用于执行本文档中描述的任何方法步骤。

[0037] 根据另一方面,描述了一种存储介质。该存储介质可以包括软件程序,该软件程序适于在处理器上执行,并且当在计算设备上实施时用于执行本文档中描述的任何方法步骤。

[0038] 根据另一方面,描述了一种计算机程序产品。该计算机程序可以包括可执行指令,该可执行指令当在计算机上执行时,用于执行本文档中描述的任何方法步骤。

[0039] 根据另一方面,描述了一种SBR参数合并单元。该SBR合并单元可以配置成根据SBR参数的N个源集合提供SBR参数的M个目标集合,其中, $N \geq M \geq 1$ 。SBR参数合并单元可以包括处理器,该处理器配置成执行本文档中描述的任何方面和方法步骤。

[0040] 根据另一方面,描述了一种音频解码器,该音频解码器配置成解码包括N个音频通道的HE-AAC比特流。该音频解码器可以包括:AAC解码器,其配置成接收经编码的HE-AAC比特流,并且提供单独的SBR比特流;和/或SBR解码器,其配置成根据SBR比特流提供与N个音频通道相对应的SBR参数的N个源集合;和/或SBR参数合并单元,如上所述,其配置成根据SBR参数的N个源集合来提供SBR参数的M个目标集合,其中, $N \geq M \geq 1$ 。

[0041] AAC解码器可以配置成提供与N个音频通道相对应的N个时域低带音频信号。音频解码器可以包括:时域缩混单元,其配置成根据N个时域低带音频信号提供M个时域低带音频信号;和/或SBR单元,其配置成根据M个低带音频信号和SBR参数的M个目标集合而生成M

个高带音频信号。从而,音频解码器可以配置成提供分别包括M个低带音频信号和M个高带音频信号的M个音频信号。

[0042] 根据另一方面,描述了一种音频转码器,其配置成根据包括N个音频通道的HE-AAC比特流提供包括M个音频通道的HE-AAC比特流,其中 $N > M \geq 1$ 。该音频转码器可以包括如上所述的SBR参数合并单元。

[0043] 根据另一方面,描述了一种电子设备,其配置成根据包括N个音频通道的HE-AAC比特流呈现与M个通道相对应的M个音频信号,其中 $N > M \geq 1$ 。该电子设备例如可以是媒体播放器、机顶盒或者智能电话。该电子设备可以包括:音频呈现装置,其配置成执行M个音频信号的声学呈现;和/或接收机,其配置成接收经编码的HE-AAC比特流;和/或音频解码器,其配置成根据本文档中描述的任何方面而根据HE-AAC比特流提供M个音频信号。

[0044] 应当注意,可以对本文档中描述的实施方案和方面进行任意的组合。特别地,应当注意,在系统的上下文中描述的方面和特征也可应用于相应的方法的上下文,反之亦然。另外,应当注意,本文档的公开也覆盖了由从属权利要求中的向后引用而明确给出的权利要求组合以外的权利要求组合,即,权利要求及其技术特征可以按照任何顺序和任何形式进行组合。

附图说明

[0045] 现在将参考附图,通过示意性实例并且不限制本发明的范围或者精神的方式描述本发明,在附图中:

[0046] 图1例示了用于将N个通道HE-AAC比特流缩混为立体声音频信号的系统的示例性框图;

[0047] 图2例示了具有5个输入通道和2个输出通道的SBR参数合并单元的示例性框图;

[0048] 图3示出了具有2个输入通道和1个输出通道的SBR参数合并单元的示例性框图;

[0049] 图4例示了在图3的SBR参数合并单元内执行的包络时间边界的示例性合并;

[0050] 图5a、图5b、图5c和图5d例示了用于根据2个源通道来确定目标通道的比例因子能量的示例性过程;以及

[0051] 图6例示了利用缩混系数对源通道进行的示例性加权方案。

具体实施方式

[0052] HE-AAC解码器可以分为AAC核心解码器和频带复制(SBR)算法,所述AAC核心解码器对经编码的音频信号的低带进行解码,所述频带复制算法使用经解码的低带信号和在比特流中传送的参数化信息而再产生音频信号的高带。通常,SBR算法需要比AAC核心解码器更多的计算资源。这是由于高频重构(即,频带复制)的分析和合成阶段使用的滤波器组而造成的。通过实例的方式,在典型的实施方案中,AAC解码需要的计算资源大约是HE-AAC比特流的解码所需要的总计算资源的1/3,其中用于SBR参数的解码和执行高频重构所需要的计算资源大约是HE-AAC比特流的解码所需要的总计算资源的2/3。

[0053] 解码器可以接收表示N个通道音频信号的HE-AAC比特流。然而,由于各种原因,例如音频呈现设备的限制,解码器可能需要提供仅包括M个音频通道(M小于N)的输出信号。在一种可选使用场景中,转码器可以接收表示N个通道音频信号的输入HE-AAC比特流,并且可

以提供表示M个通道音频信号的输出HE-AAC比特流。

[0054] 鉴于使用SBR参数的音频信号的高频分量或者高带的重构的高计算复杂度,可能有益的是,在经编码的域中执行从N个通道到M个通道的缩混,在这之后可选地,解码经缩混的比特流以及生成与M个通道相对应的M个高带音频信号。在下文中,将描述允许将N个输入或者源通道的SBR参数合并为M个输出或者目标通道的SBR参数的方法。SBR参数的合并执行为使得保留与特定音频事件有关的信息。

[0055] 所提出的方法可以包括步骤:解码用于N个输入通道的SBR参数,从而提供与N个源通道相对应的N个SBR参数集合。随后,执行SBR参数的合并的步骤,以获得与M个目标通道相对应的M个SBR参数集合。为了提供M个通道输出信号,该方法可以包括步骤:利用后续的时域缩混解码用于全部N个输入通道的经AAC编码的低带信号,以获得M个输出通道。另外,可以使用根据经AAC编码的低带信号获得的M个缩混通道以及在上述SBR合并步骤中获得的相应的新SBR参数集合来执行M个通道的谱带重构。

[0056] 图1中示出了示例性的HE-AAC解码器100,其根据表示N个音频通道的输入HE-AAC比特流101来提供与两个输出或者目标通道相对应的两个输出音频信号107、108。AAC解码器110执行将HE-AAC比特流101解码成包括N个音频信号的低频分量的N个音频信号103(也称为低带音频信号103)。在时域缩混单元113内,将N个低带音频信号103缩混为两个低带音频信号106。AAC解码器还提供包括用于N个音频通道的SBR参数的SBR比特流102。在SBR解码器111内解码SBR比特流102,以产生N个SBR参数集合104,一个SBR参数集合104用于N个音频通道中的每一个。可以根据通过引用并入本文的ISO/IEC14496-3子部分4.4.2.8和4.5.2.8来执行参数提取和解码。在SBR参数合并单元112中,N个SBR参数集合104被合并为两个SBR参数集合105。最后,在SBR单元114中执行两个输出音频信号107、108的频带复制或者高频重构。SBR单元114使用低带音频信号106以及经合并的SBR参数的集合105生成两个音频信号的高频分量,并且提供包括各低频分量和高频分量的两个音频信号107、108作为输出。

[0057] 图2例示了示例性SBR参数合并单元112的框图。所例示的SBR参数合并单元112具有用于将输入处的5个SBR参数集合201、202、203、204、205合并为输出处的2个SBR参数集合208、209的层级式结构。SBR参数合并单元112包括“2到1”SBR参数合并单元210、211、212、213,它们将输入处的2个SBR参数集合201、202合并为输出处的1个SBR参数集合206。“2到1”SBR参数合并单元210、211、212、213将被称为“基本合并单元”。通过使用层级式组织的基本合并单元210,有可能提供灵活且适应的SBR参数合并单元112,其可操作用于将输入处的任意数目N个SBR参数集合201合并为输出处的任意数目M个SBR参数集合208。通过添加或者移除基本合并单元210,总的SBR参数合并单元112可以适用于改变输入通道的数目N和/或改变输出通道的数目M。

[0058] 图2例示了示例性SBR参数合并单元112的框图,所例示的SBR参数合并单元112具有层级式结构,该层级式结构用于将输入处的5个SBR参数集合201、202、204、204、205合并为输出处的2个SBR参数集合208、209。SBR参数合并单元112包括“2到1”SBR参数合并单元210、211、212、213,它们将输入处的2个SBR参数集合201、202合并为输出处的1个SBR参数集合206。“2到1”SBR参数合并单元210、211、212、213将被称为“基本合并单元”。通过使用层级式组织的基本合并单元210,有可能提供灵活并且自适应的SBR参数合并单元112,其可操作用于将输入处的任意数目N个SBR参数集合201合并为输出处的任意数目M个SBR参数集合

208。通过添加或者移除基本合并单元210,总体的SBR参数合并单元112可以适用于改变输入通道的数目N和/或改变输出通道的数目M。

[0059] 图2例示了SBR参数合并单元112的实例,该SBR参数合并单元112将5.1输入信号的SBR参数合并为立体声输出信号的SBR参数。5.1信号包括5个全范围通道,称为左(L)、右(R)、环绕左(LS)、环绕右(RS)和中心(C)通道,以及低频效果(LFE)通道。在所例示的实例中,没有考虑LFE通道。通常,仅在LFE通道也可供作为输出通道之一时才保留此类LFE通道的内容。

[0060] 在所例示的实施方案中,将与C通道相对应的SBR参数集合201在第一基本合并单元210中与LS通道的SBR参数集合202合并,并且在第二基本合并单元211中与RS通道的SBR参数集合203合并。这分别产生了两个经合并的SBR参数集合206、207。这些经合并的SBR参数集合206、207可以称为SBR参数的中间集合。随后,在基本合并单元212中将经合并的SBR参数集合206与L通道的SBR参数集合204合并,以产生与立体声输出信号的左通道(L')相对应的经合并的SBR参数集合208。在基本合并单元213中将经合并的SBR参数集合207与R通道的SBR参数集合205合并,以产生与立体声输出信号的右通道(R')相对应的经合并的SBR参数集合209。

[0061] 所例示的层级式合并方案仅是用于合并输入处的多个SBR参数集合的一种可能性。SBR参数集合也可以按照不同的顺序进行合并。然而,应当注意,通常基本合并单元210内的每个合并步骤导致包括在SBR参数集合内的信息的稀释。因此,可以优选的是,使较高声学重要性或者较高声学相关性的通道比相对较低声学重要性或者声学相关性的通道经历较低数目的合并步骤。通过实例的方式,可以使L和R通道经历比C通道较少的合并步骤。作为另一实例,在C通道在其中传送对话的电影声带具有高声学重要性的情况下,可以使C通道比L和R通道经历较少的合并步骤。

[0062] 在可选实施方案中,SBR参数合并单元112可以实施为总体矩阵,直接将输入处的N个SBR参数集合201合并为输出处的M个SBR参数集合208。

[0063] 在下文中,将描述在基本合并单元210中将2个SBR参数集合201、202合并为1个经合并的SBR参数集合206。可以通过考虑输入处的不止两个SBR参数集合而推广所描述的方法和系统。

[0064] 在图3中,示出了示例性基本合并单元210的框图。基本合并单元210根据2个SBR参数集合201、202(也称为源集合)而提供经合并的SBR参数集合206(也称为目标集合)。所例示的基本合并单元210通常逐帧地执行SBR参数的合并,即,合并与各个输入通道相对应的输入信号的帧的SBR参数,以提供输出通道的输出信号的对应帧的SBR参数。为了便于例示,SBR参数集合201、202、206是指在下文中的单个帧的SBR参数集合。

[0065] 通过实例的方式,输入信号的帧可以包括覆盖输出信号采样率为2048个样本的标称长度的包络集合。例如,如果QMF滤波器组具有64个子带的频率分辨率,则2048的帧长度将与每个子带中的32个QMF子带样本相对应。另外,可以引入附加单元,即“时隙”,其按照2个子带-样本的间隔尺寸(granularity)组合子带样本。换言之,帧可以包括与16个时隙相对应的32个QMF子带样本(每QMF子带)。

[0066] 所例示的基本合并单元210包括包络时间边界确定单元301,其根据两个源集合201、202的包络时间边界来确定目标集合206的包络时间边界。参照图4更加详细地描述包

络时间边界确定单元301。随后,在比例因子能量确定单元302中根据源集合201、202的比例因子能量来确定目标集合206的比例因子能量。参照图5a、图5b、图5c和图5d更加详细地描述比例因子能量确定单元302。

[0067] 除了包络时间边界参数和比例因子能量的合并以外,SBR参数合并单元112或者基本合并单元210还可以执行其他SBR参数的合并。可以根据通过引用并入本文的ETSI TS126402,6.1节来合并SBR参数“反向滤波水平”。可以根据通过引用并入本文的ETSI TS126402,6.2节来合并SBR参数“附加谐波”。

[0068] 另外,可能需要SBR参数“每包络的频率分辨率”。该参数包括参数bs_freq_res,其是用于选择两个频率表中的一个的二进制开关。值bs_freq_res==0选择低分辨率表,而bs_freq_res==1选择高分辨率表。通常,两个表都通过选择频带的子集而根据主频率表得出。主频率表的频率分辨率由参数bs_freq_scale确定。值bs_freq_scale==0是每频带具有一个QMF子带的最细分辨率。参数bs_freq_scale的较高值造成每倍频程8-12个频带的较粗分辨率。关于该SBR参数的细节可以参见ISO/IEC14496-3子部分4.6.18.3.2,通过引用将其并入本文。通常,参数bs_freq_scale包括在SBR元素头部内。下文处理SBR元素头部的合并。针对经合并的通道,参数bs_freq_res可以设置为1,从而指示要使用具有细分分辨率的表。

[0069] 可以根据以下过程合并参数“SBR元素头部”:

[0070] 1) 可以确定所有源通道元素的起始/终止频率。在SBR参数合并单元112的情况下,可能的源通道是通道201、202、203、204、205。

[0071] 2) 选择具有最高起始频率的源通道元素的头部作为用于其一部分的目标通道元素的头部。在目标通道元素208的情况下,考虑源通道元素201、202和204的头部。在目标通道元素209的情况下,考虑源通道元素201、203和205的头部。应当注意,在可选实施方案中,可以是有益的是,选择具有最低起始频率的源通道元素的头部作为用于其一部分的目标通道元素的头部。

[0072] 3) 可以对目标通道头部选择进行进一步限制,以便于匹配目标通道元素的通道元素类型。

[0073] 如果目标通道元素是CPE(通道对元素),则选择作为混合的一部分的具有最高起始频率的源CPE的头部作为用于目标通道元素的头部。如果不存在源CPE,则选择具有最高起始频率的源SCE(单通道元素)的头部,并且用于构造用于目标通道元素的CPE头部。

[0074] 如果目标通道元素是SCE,则选择作为混合的一部分的具有最高起始频率的源SCE的头部作为用于目标通道元素的头部。如果不存在源SCE,则选择具有最高起始频率的源CPE的头部,并且用于构造用于目标通道元素的SCE头部。

[0075] 应当注意,通常,第一和第二源集合201,202的起始和终止频率是不同的。起始/终止频率通常在各源集合201、202的SBR元素头部内限定。音频通道的起始频率也称为交叉频率,指定低频分量的最大频率和/或高频分量的最小频率。当合并一定数目的音频通道时,可以有益的是,保证经合并的高频分量不与经合并的低频分量发生干扰。原因在于这样的事实,即经AAC编码的低频分量通常比经SBR编码的高频分量包括更多的相关声学信息。因此,应避免由经合并的SBR参数造成的低频信号分量与高频信号分量的干扰。这可以通过选择目标集合206的起始频率或者作为贡献于目标集合206的源集合201、202的最大起始频率的目标通道而得到保证。特别地,可以通过选择如上所述的目标集合206的SBR元素头部,而

避免上文提到的经合并的低频分量与经合并的高频分量之间的干扰的风险。

[0076] 在下文中,描述了与时间边界相关的SBR参数的合并。应当注意,虽然以下描述涉及包络时间边界的合并,但是其还可以应用于噪声包络时间边界。另外,参考ETSI TS126402,6.4节,通过引用将其并入本文,其中描述了一种用于合并噪声包络时间边界的方案。

[0077] HE-AAC允许在帧内限定多达5个包络。这些包络指定帧的特定时间间隔内经编码的音频信号的高频分量的谱包络。可以根据一定时间网格,沿时间轴限定不同包络的时间边界。通常,帧的长度(例如,24ms)被细分为多个时隙(例如,16个时隙),每个时隙限定用于包络的可能的时间边界。可以根据ETSI TS126402,6.3节对源集合201、202的包络时间边界进行合并,通过引用将其并入本文。

[0078] 图4例示了由两个源集合201、202限定的谱包络。谱包络被表示为时间/频率图上的并排显示(tile),其中时间 t_{401} 表示帧的长度,并且频率 f_{402} 表示各音频信号的高频分量的频率。所例示的实例中的源集合201指定具有中间时间边界415、416、417的四个包络411、412、413、414。所例示的实例中的源集合202指定具有中间时间边界425、426、427的四个包络421、422、423、424。中间时间边界是用于之后的包络的起始时间边界以及之前包络的终止时间边界。另外,图4示出了第一包络的起始时间边界403和最后包络的终止时间边界404。

[0079] 包络时间边界确定单元301可操作用于根据源集合201、202的包络411、412、413、414、421、422、423、424的时间结构来提供目标集合206的包络的时间结构,即,起始时间边界和终止时间边界。为此,源集合201、202的时间结构(即,起始时间边界和终止时间边界)重叠,如图4所示。由于两个源集合201、202的包络的这一重叠,获得了包括用于目标集合206的7个时间间隔的时间结构,其中,这些时间间隔由时间边界[403,425]、[425,415]、[415,416]、[416,426]、[426,417]、[417,427]和[427,404]限定。这些时间间隔可以理解为目标集合206的各包络的时间间隔。如果所获得的目标集合206的时间间隔的数目没有超过所允许的包络的最大数目,则可以维持所获得的时间边界。所允许的包络的最大数目可能受到下层编码方案的影响。在HE-AAC的情况下,每帧所允许的包络的最大数目被固定为5。

[0080] 然而,如果超过了所允许的时间间隔的数目,则需要对目标集合206的时间间隔的一定数目进行合并。这可以通过将小于两个时隙的所有时间间隔与直接在前或者在后的时间间隔进行合并来完成。这可以通过起始于时间轴401的开始(由起始时间边界403指示),并且移除与相应的起始时间边界比2更加接近的所有终止时间边界来实现。在所例示的实例中,终止时间边界426将被移除,从而创建具有时间边界[416,417]的新时间间隔。如果在这一操作之后,仍然存在比所允许的包络最大数目(例如,5)更多的时间间隔,则可以进一步减少时间间隔的数目。这可以通过以下来实现:起始于时间轴401的结尾(由终止时间边界404指示),并且向时间轴401的开始(由附注标记403指示)搜索小于4个时隙的时间间隔,并且移除该时间间隔的起始时间边界。这一搜索操作可以继续,直到达到与所允许的包络的最大数目相对应的时间间隔的数目。在所例示的实例中,起始时间边界417将被移除,从而创建具有时间边界[416,427]的新时间间隔。

[0081] 通过使用上述合并时间间隔的过程,可以保证目标集合206的时间间隔的数目不超过所允许的包络的最大数目。在以上实例中,时隙的数目是16,并且所允许的包络的最大

数目是5。目标集合206的包络的平均时间间隔应当不小于 $16/5=3.2$ 个时隙,这可以通过将时间间隔与逐步增加的阈值进行合并(如上所述)来实现。通常,可以规定时间间隔的平均长度应当至少是每帧的时隙的数目与所允许的包络的最大数目的比例。

[0082] 作为包络时间边界确定单元301的输出,获得了目标集合206的谱包络的时间间隔,其由时间边界403、425、415、416、427、404限定。已经减少了时间间隔的数目,以使得时间间隔的数目不超过谱包络的最大允许数目。

[0083] 上述的确定目标集合206的包络的时间间隔的过程可以推广至任意数目个源集合201。在这种情况下,源集合201的所有时间边界将重叠,如图4所示并如上所述。使用后续的时间间隔的合并过程,可以确定目标集合206的包络的时间间隔的预定数目。

[0084] 应当注意,帧的包络可以被标记为瞬态谱包络,从而指示在帧内的特定时间间隔中存在音频信号中的瞬态。通常,每帧和每通道的瞬态谱包络的数目限制于1。瞬态谱包络通常由指数 l_A 标记,其指示谱包络的数目。如果所允许的谱包络的最大数目是5,则指数 l_A 可以例如采用值0, ..., 4中的任何一个值。源集合的瞬态包络指数可以按照如下进行合并:

[0085] i. 针对每个源集合201、202,确定当前帧的瞬态包络指数 l_A 是否指示存在瞬态,即, $l_A \neq -1$ 。

[0086] ii. 针对每个 $l_A \neq -1$,确定该包络的起始时间边界。

[0087] iii. 如果不同源集合201、202中存在瞬态,并且因而确定多个起始时间边界,则可以选择最小起始时间边界(即,最早的一个)。

[0088] iv. 在目标集合206内,标识最接近已经在步骤i-iii中确定的起始时间边界的时间边界。

[0089] v. 选择目标集合206的、其起始时间边界与在步骤iv中标识的边界相对应的时间间隔或者包络,作为经合并的通道的瞬态包络 l_A 。

[0090] 如果假设在图4所例示的实例中,源集合201包络瞬态包络414,并且源集合202包络瞬态包络423,则步骤iii选择起始时间边界426。随后,在步骤iv中,确定目标集合206的、最接近起始时间边界426的起始时间边界416,并且通过将瞬态包络指数 l_A 设置为2,将时间间隔[416, 427]标记为瞬态包络。通过应用以上方法,易于将瞬态移至较早的可能时间间隔。例如,由于较早瞬态的时间遮蔽效应,相对于选择较晚的起始时间边界,这可能具有音质方面的益处。另外,以上的方法通常保证目标集合206的瞬态包络覆盖源集合201、203的瞬态包络414、423的时隙中的很多时隙。然而,应当注意,作为另外的或者可选的限制,可以选择目标集合206的瞬态包络,以使得其起始时间边界不晚于源集合201、202的瞬态包络414、423的起始时间边界中的任何一个。

[0091] 以上用于根据源集合201、202的一个或多个瞬态包络指数来确定目标集合206的瞬态包络指数的过程可以推广至任意数目的源集合的任意数目的瞬态包络指数。为了这个目的,针对任意数目的瞬态包络指数执行方法步骤ii、iii、iv和v。

[0092] 在下文中,描述了在比例因子能量确定单元302内合并两个源集合201、202的谱包络。谱包络包括一个或多个比例因子带以及针对比因子带中的每一个的比例因子。换言之,谱包络指定谱包络的时间间隔内的各通道的高带信号的谱能量分布。

[0093] 如上所述,已经在包络时间边界确定单元301中确定了目标集合206的谱包络的时间间隔。比例因子能量确定单元302可操作用于根据源集合201、202的谱包络确定目标集合

206的谱包络的比例因子带以及相关连的比例因子。

[0094] 图5a例示了用于合并包括在两个源集合201、202的谱包络内的比例因子能量的基本原理。在包络时间边界确定单元301中,已经确定了目标集合206的包络532的时间边界403、425。该包络532跨越由各时间边界403、425限定的时间间隔503。将时间间隔503应用于源集合201、202的谱包络,从而指定贡献于目标集合的谱包络532的源集合201、202的谱包络。在所例示的实例中,可以看出,源集合201的谱包络411落在时间间隔503之内,因此贡献于目标集合206的谱包络532。另外,可以看出,源集合202的谱包络421落在时间间隔503之内,因此贡献于目标集合206的谱包络532。

[0095] 应当注意,通常,源集合201的一个或多个谱包络411可落在目标集合206的谱包络532的时间间隔503之内。因此,源集合201的不止一个谱包络411可以贡献于目标集合206的谱包络532。在较晚的阶段将描述多个贡献谱包络的这一方面。为了易于例示,将在第一阶段描述源集合201、202的两个谱包络的合并。这些谱包络称为第一源包络512和第二源包络522,并且分别与源集合201、202的谱包络411、421相关联。在一个实施方案中,第一和第二源包络512,522可以分别与源集合201、202的谱包络411、421相对应。

[0096] 另外,应当注意,贡献源包络411、421的起始频率可以不同。如上所述,通常将目标集合206的起始频率选择为贡献源集合201、202的最大起始频率。在一个实施方案中,可以将目标集合206的起始频率选择为贡献于SBR参数合并单元112的最终目标集合208的所有源集合201、202、204的最大起始频率(如上在SBR元素头部的合并的上下文中所述)。因此,不是源集合201、202的谱包络411、421的完整频率范围都可以贡献于目标集合206的谱包络532(也称为目标包络532)。这在图5b中例示,其中示出了源集合201、202的谱包络411、421。在所例示的实例中,谱包络411具有起始频率551,其低于谱包络421的起始频率552。如果将较高的起始频率552选择作为目标包络532的起始频率553,则可以截去谱包络411。这是由于这样的事实,即较低起始频率551与较高起始频率552之间的频率范围中的比例因子带通常将不贡献于目标包络532。由此,可以通过在合并过程期间忽略较低起始频率551与较高起始频率552之间的频率范围而实现谱包络411的“截去”。

[0097] 通常,可以规定,可以截去贡献于目标包络532的源包络512、522,以使得其频率范围与目标包络532的频率范围相对应。特别地,可以截去位于目标包络532的起始频率以下且在终止频率以上的频带或者频带的一个或多个部分。在下文中,假设已经如上所述截去了贡献源包络512、522,从而使得其起始和/或终止频率与目标包络532的起始和/或终止频率相对应。

[0098] 通常,第一源包络512的比例因子带划分不与第二源包络522的比例因子带划分相对应。换言之,针对不同的源包络512、522,具有恒定能量的频带(即,具有恒定比例因子能量的频带)不同。这在图5a中例示,其中第一源包络512的边界频率513、514不同于第二源包络522的边界频率523、524、525。另外,第一源包络512中的比例因子带的数目(在所例示的实例中是3)可以不同于第二源包络522中的比例因子带的数目(在所例示的实例中是4)。另外,源包络512、522可以包括取决于频率的不同水平的能量。比例因子能量确定单元302可操作于根据贡献源包络512、522来确定目标包络532,其中,目标包络532包括一个或多个比例因子带以及各比例因子能量。

[0099] 在下文中,将描述与源包络512、522的比例因子带相对应的比例因子能量的合并。

基本构思是提供多个源包络512、522与目标包络532之间的联合频率网格。可以通过基于SBR的编解码器中使用的分析/合成滤波器组的QMF(正交镜像滤波器)子带来提供这样的联合频率网格。使用联合频率网格,例如,QMF子带,将与相同的QMF子带相对应的贡献源包络的比例因子相加,以便提供目标包络的相应QMF子带的累积比例因子能量。最后,可以将累积比例因子能量除以贡献源集合的数目,以便提供平均比例因子,作为目标包络的相应QMF子带的比例因子能量。

[0100] 在图5c和图5d中示出了比例因子能量的该合并过程。图5c例示了与源包络512相关联的多个比例因子能量515、516和517,以及与源包络522相关联的比例因子能量526、527、528和529。针对混合成目标包络的每个源包络512、522,执行以下步骤。针对一定的比例因子带511描述这些步骤。特别地,针对比例因子带511内的一定QMF子带541描述这些步骤。应当针对位于目标包络532的频率范围内的所有QMF子带541执行这些步骤。

[0101] 在第一步骤中,可以通过相应的、针对与源集合201相对应的通道的经过能量补偿的缩混系数来对每个比例因子带511的比例因子能量517进行缩放。将在后面的阶段描述经能量补偿的缩混系数的确定。

[0102] 如上所述,每个源比例因子带511被分解为QMF子带541,即,比例因子带511被分解为联合频率网格。为比例因子带511的每个QMF子带541分配各比例因子带511的比例因子能量517。换言之,为QMF子带541分配其所位于的比例因子带511的比例因子能量517。在下文中,将QMF子带541的网格上的比例因子带511和相应的比例因子能量517的表示称为“QMF表示”。

[0103] 在下面步骤中,向目标通道的相应目标QMF表示添加源QMF表示。在图5c所例示的实例中,向目标包络532的相应的QMF子带543的比例因子能量533添加源集合201的QMF子带541的比例因子能量517。以类似的方式,向目标包络532的相应的QMF子带543的比例因子能量533添加源集合202的QMF子带542的比例因子能量529。最后,可以将累积比例因子能量533除以贡献源集合201、202的数目,以产生平均比例因子能量533。

[0104] 应当注意,由于在单元301中的包络时间边界确定过程期间移除了起始/终止时间边界,所以目标包络532的时间间隔503可能覆盖第一和/或第二源集合201,202的若干包络。以上已经指示了源集合201的多个贡献包络411的这一方面。在下文中,将描述可以如何在比例因子能量确定单元302中考虑此类多个源包络。一般的构思是根据其部分贡献考虑源集合201的每个贡献源包络。源集合的源包络可以仅与目标包络的时间间隔部分重叠。换言之,目标包络的时间间隔可以跨越源集合的若干包络,以使得源集合的每个包络仅覆盖目标包络的时间间隔的一部分时间。可以通过根据源集合的贡献包络的比例因子能量对目标包络的时间间隔贡献的时间分数来对其进行缩放,而将此类部分贡献纳入考虑。如果时间轴被细分为时隙,则可以根据重叠时隙(即,各源包络与目标包络的重叠时隙)与包括在目标包络的时间间隔中的时隙的数目的比例来执行比例因子能量的缩放。

[0105] 部分贡献可以在图4中例示。目标集合206的时间间隔[416,427]包括第一源集合201的源包络413、414和第二源集合202的源包络422、423。在这种情况下,贡献于目标集合206的目标包络531的第一和第二源集合201,202的所有源包络413、414、422、423应当被考虑用于比例因子能量的合并。不同的源包络413、414、422、423的比例因子带内的比例因子能量应当部分地根据由贡献包络413、414、422、423和目标包络的时间间隔[416,427]的重

叠时隙的数目以及目标包络的时间间隔 [416, 427] 的时隙的数目给出的比例来贡献。可以在如上所述的用于合并比例因子能量的过程中使用考虑源包络 413、414、422、423 对目标包络的部分贡献的这一方面。特别地, 可以将贡献源包络 413、414、422、423 的经缩放的比例因子能量相加, 以确定目标包络 532 的 QMF 子带 543 的累积比例因子能量 533。

[0106] 作为以上过程的结果, 获得了用于目标包络 532 的目标比例因子带。根据贡献源包络 512 的数目, 包括在源包络 512 内的比例因子带 511 的数目和比例因子带 511 之间的频率边界 513 的位置, 用于目标包络 532 的比例因子带的数目可相对较高。例如, 由于下层编码方案的限制和/或由于预定的比例因子带划分或者结构原因, 减少目标包络 532 内的比例因子带的数目可以是有益的。

[0107] 通过实例的方式, 如果目标集合 206 使用源集合 201、202 中的一个的 SBR 元素头部, 则可以使用各源集合 201、202 的比例因子带结构。如已经在用于合并多个源集合的 SBR 元素头部的方法的上下文中所述的, 目标集合的 SBR 元素头部可以对应于或者可以基于源集合之一的 SBR 元素头部。除了指定包括在各 SBR 参数集合内的谱包络的起始和/或终止频率之外, SBR 元素头部还可以指定谱包络的比例因子带结构。该比例因子带结构可以用于在如上所述的比例因子能量合并过程中确定的目标包络。在下文中, 描述了可以如何将根据合并过程获得的比例因子带结构 (也称为第一比例因子带结构) 转换为预定比例因子带结构 (例如, 由目标集合 206 的 SBR 元素头部而给出的结构, 也称为第二比例因子带结构) 的方法。

[0108] 为了从第一比例因子带结构转换为第二比例因子带结构, 可以使用以下过程, 该过程参照图 5d 描述。该过程针对第二比例因子带结构的特定比例因子带而进行描述, 并且应当针对第二比例因子带结构的所有比例因子带执行。该过程依赖于频率网格, 例如 QMF 子带 543。

[0109] 在第一步骤中, 将第二比例因子带结构的比例因子带中的所有 QMF 子带 543 的比例因子能量 533 加和。如上所述, 可以通过已经在 SBR 元素头部的合并过程期间选择的 SBR 元素头部来确定目标比例因子带划分 (即, 第二比例因子带结构)。

[0110] 将第一步骤中计算的 QMF 子带能量的总和除以已经加和的 QMF 子带的数目。换言之, 确定第二比例因子带结构的比例因子带的平均比例因子能量 534。结果是各比例因子带的目标比例因子能量 534。针对第二比例因子带结构的其他比例因子带重复这一过程。

[0111] 总之, 已经描述了用于确定目标包络 532 的目标比例因子带结构中的比例因子能量的过程。通过使用上述用于目标集合 206 的所有目标包络 532 的合并过程, 可以获得目标集合 206 的包络的经合并的比例因子能量的完整集合。所描述的过程可以推广至任意数目的源集合 201。在这种情况下, 任意数目的源包络可以贡献于目标包络 532。使用联合频率网格 (例如, QMF 子带) 来分解贡献源包络, 并且将相应的 QMF 子带的源比例因子能量加和, 以确定相应的 QMF 子带的目标比例因子能量。可以利用贡献源集合的数目来对目标比例因子能量进行归一化。如果源集合的源包络仅部分贡献, 则可以根据如上所述的方法来缩放比例因子能量。另外, 可以通过经能量补偿的缩混因子对比例因子能量进行加权。最后, 可以将所确定的比例因子能量和比例因子带结构转换为预定的比例因子带结构。

[0112] 应当注意, 源集合 201、202 可以指定噪声基底水平。可以按照与比例因子能量相似的方式来合并不同源通道的此类噪声基底水平。在这种情况下, 比例因子能量与噪声基底水平相对应, 并且包络时间边界与噪声基底边界相对应。然而, 应当注意, 用于噪声的时间

间隔的数目通常低于包络的数目。在一个实施方案中,使用起始边界、终止边界和中间边界,可以在帧内仅限定两个噪声时间间隔。在此类噪声时间间隔中,可以指定一个或多个噪声基底水平和相应的频带结构(或者噪声基底比例因子带结构)。可以使用参照图4描述的过程来合并多个源集合201的起始边界、终止边界和/或中间边界。可以使用参照图5a-5d描述的过程来合并多个源集合201的一个或多个噪声基底水平。

[0113] 然而,应当注意,通常不通过经能量补偿的缩混系数来缩放噪声基底水平。但是,可以缩放贡献源噪声基底水平和/或目标噪声基底水平,以便微调经合并的音频通道的主观音频质量。

[0114] 在比例因子能量合并方法的上下文中,已经指示了对源通道应用缩混系数可以是有益的。通常对低带信号应用此类缩混系数,以便对经缩混的通道提供削波保护。图6示出了对相应的音频通道的低带信号应用缩混系数。可以看出,利用缩混系数 c_0 对C通道进行加权或者缩放,利用缩混系数 c_1 对R和L通道进行加权,并且利用缩混系数 c_2 对LS和RS通道进行加权。在从5个通道缩混至2个通道的上下文中,缩混系数可以指定如下: $c_0=0.7/scale$, $c_1=1.0/scale$, $c_2=0.5/scale$,其中 $scale=0.7+1.0+0.5=2.2$ 。这些系数值对应于国际电信联盟(ITU)对5.1通道信号的缩混的建议。如果缩混小于5个通道(例如,仅有左、右和中心通道),也可以使用这些系数。

[0115] 以与低带信号相似的方式,利用缩混系数对源通道或者源集合201、202的比例因子能量进行加权可以是有益的。这可能对于维持音频信号的低频分量与高频分量之间的比例是重要的。特别地,这可能对于维持低频分量与高频分量的能量的比例是重要的。在该上下文中,图6例示了5个输入通道缩混至2个输出通道的单个步骤。对输入通道直接应用缩混系数。在一种可选实施方案中,可以使用如图2所示的层级式缩混,从而将在输入通道201、202、203、204、205上直接应用缩混系数。

[0116] 然而,应当注意,时域中的源通道可以是同相或者反相的,以使得可以根据相位关系来放大或者衰减时域中的经缩混的目标通道。为了在合并比例因子能量时将这一效果纳入考虑,可以将以上缩混系数乘以能量补偿因子,该能量补偿因子考虑了贡献源通道的音频信号的同相和/或反相行为。特别地,能量补偿因子考虑相对于贡献低带音频信号所产生的经缩混的低带音频信号的衰减或放大。针对音频信号的给定帧,可以根据以下公式计算能量补偿因子:

$$[0117] \quad f_{comp} = \sqrt{\frac{\sum_{chout=0}^{M-1} \sum_{n=0}^{1023} x_{dmx}^2[chout][n]}{\sum_{chin=0}^{N-1} \sum_{n=0}^{1023} (c_{chin} \cdot x_{in}[chin][n])^2}}$$

[0118] 其中, f_{comp} 是用于缩混系数的补偿因子, $x_{in}[chin][n]$ 是源通道 $chin$ (通道 in)中的低带时域信号, c_{chin} 是用于通道 $chin$ 的缩混系数(例如,图6的 c_0, c_1, c_2), $x_{dmx}[chout][n]$ 是目标通道 $chout$ (通道 out)的低带时域信号,并且 $n=0, \dots, 1023$ 是帧内样本的样本指数。该公式计算一个帧的可用样本的能量。特别地,该公式确定目标通道的能量与源通道的能量之间的比例,其中通过源通道的各缩混系数对源通道进行加权。在很多情况下,具有较低精

度(例如,仅使用可用样本的一部分)的能量估计可以足以确定适当的能量补偿因子。

[0119] 使用能量补偿因子,可以维持不同音频通道的音频信号的低频分量与高频分量之间的能量均衡。这可以通过考虑源通道的信号对缩混通道的经缩混信号的正和/或负贡献来实现。应当注意,在根据N个输入通道提供M个输出通道的缩混系统中,有可能提供用于整个系统的单个能量补偿因子。可选地或者附加地,可以确定多个能量补偿因子。通过实例的方式,可以针对M个经缩混的输出通道中的每一个确定专用的能量补偿因子。这可以通过仅考虑贡献于各输出通道的输入通道来进行。在另一实例中,可以针对每个基本合并单元210来确定专用能量补偿因子。

[0120] 可以将已经用于产生AAC解码器输出的时域缩混的缩混系数c(例如,以上指定的 c_0, c_1 和 c_2)乘以该能量补偿因子 f_{comp} ,以产生经能量补偿的缩混系数。在合并源集合201、202的比例因子能量之前,可以利用如上所述的各个经能量补偿的缩混系数对比例因子能量517进行加权或者缩放。鉴于已经针对时域信号限定缩混系数c的事实,应当利用各源通道的经能量补偿的缩混系数的平方值(即, $(f_{comp} * c_{chin})^2$)来缩放比例因子能量517。由此,应当注意, $(f_{comp})^2$ 的计算可以是充分的。通常,由于可以省略用于 f_{comp} 的确定的平方根运算,所以这应当是更加有效的。

[0121] 通常,如上所述对缩混系数c进行缩放或者归一化,以使得其总和达到常数值,例如,1。在缩放至值1的情况下,经缩放的缩混系数的范围限于 $[0.01; 1]$ 。然而,鉴于缩混系数用于指定不同源通道的相对加权的事实,可以为归一化选择不同的常数值。因此,可以根据常数归一化值来增加或者降低上述极限值,条件是维持缩混系数之间的相对比例。

[0122] 应当注意,在可选实施方案中,可以对低带缩混信号应用能量补偿。这是由于应用能量补偿因子以维持高带信号与低带信号之间的均衡的事实。该均衡也可以通过对缩混信号的缩混阶段应用逆能量补偿因子来维持。在这样的实施方案中,用于比例因子能量的缩混系数将保持不变,即,其将不经过任何缩混补偿。

[0123] 在本文档中,已经描述了用于缩混SBR参数的方法和系统。所描述的方法和系统允许实施一般合并过程以便根据N个通道的SBR参数产生用于M个通道的SBR参数,其中 $M < N$ 。特别地,所述方法和系统允许利用不同的起始/终止频率来合并通道的SBR参数。另外,所述方法和系统允许利用不同的比例因子带划分来合并通道的SBR参数。另外,已经描述了用于瞬态包络信息的精确合并的方案。另外,描述了层级式合并过程,这使得有可能适应地处理多个通道配置。另外,已经描述了适应性能量补偿方案,其抑制或者提高SBR能量,以便使经缩混的信号的重构的高带信号的能量与经缩混的信号的低带信号的能量相匹配。通过使用这样的补偿方案,可以在经编码的域中直接补偿在时域中的缩混阶段期间的不同音频通道的同相和/或反相行为。

[0124] 在本文档中描述的用于缩混的方法和系统可以实现为软件、固件和/或硬件。某些组件例如可以实现为在数字信号处理器或者微处理器上运行的软件。其他组件例如可以实现为硬件和/或专用集成电路。在所描述的方法和系统中遇到的信号可以存储在介质上,如随机存取存储器或者光存储介质。其可以经由网络进行传送,网络如无线网络、卫星网络、无线网络或者有线网络(例如互联网)。使用本文档中描述的方法和系统的典型设备是便携式电子设备或者用于存储和/或呈现音频信号的其他消费者设备。所述方法和系统也可以在计算机系统(例如互联网web服务器)上使用,其存储和提供音频信号(例如,音乐信

号)以用于下载。

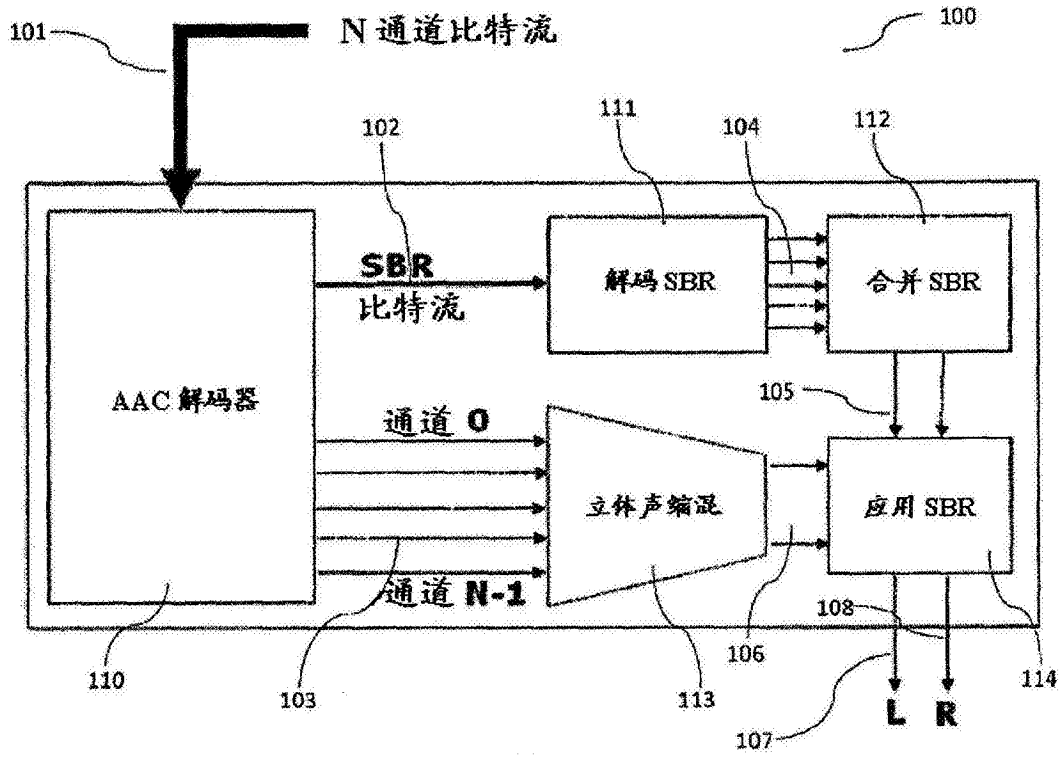


图1

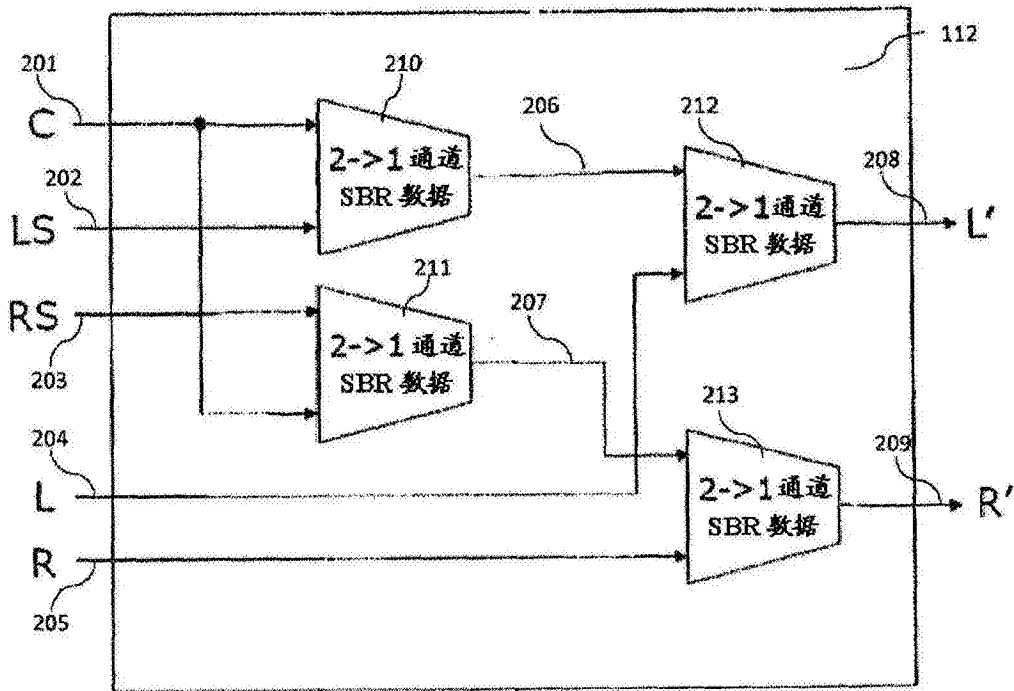


图2

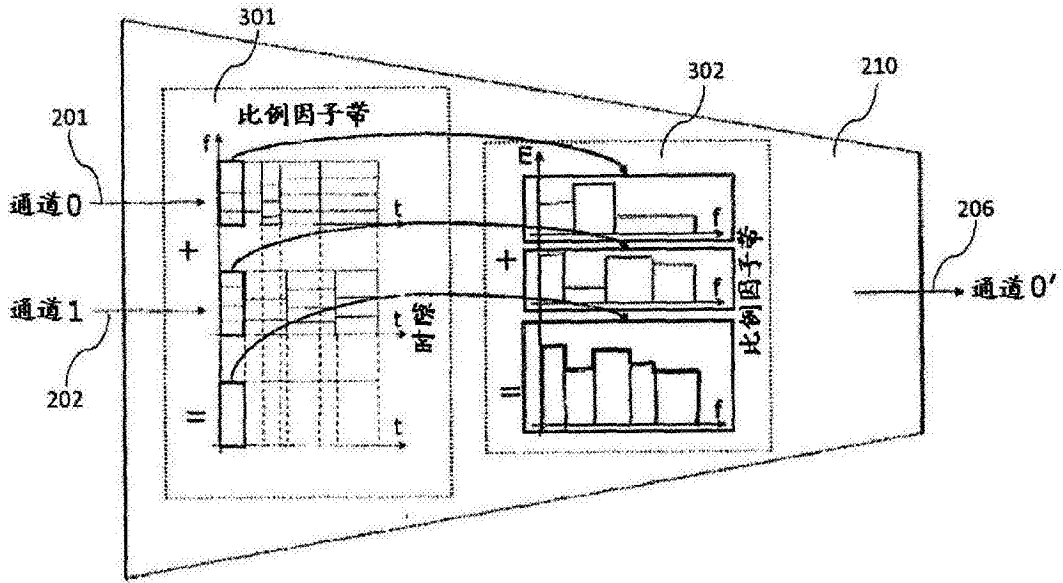


图3

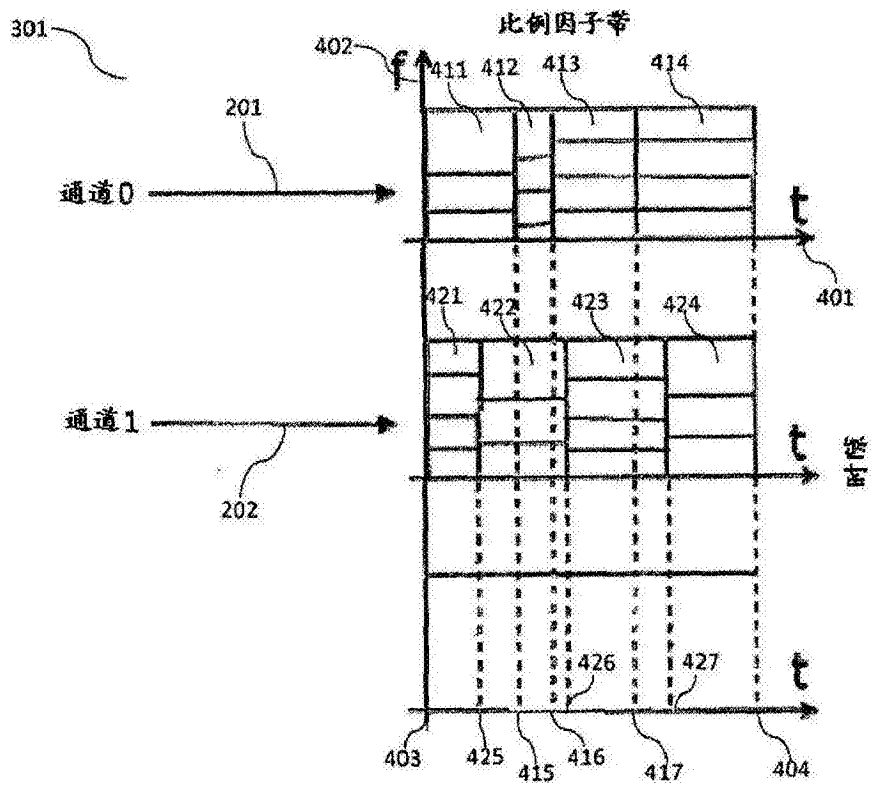


图4

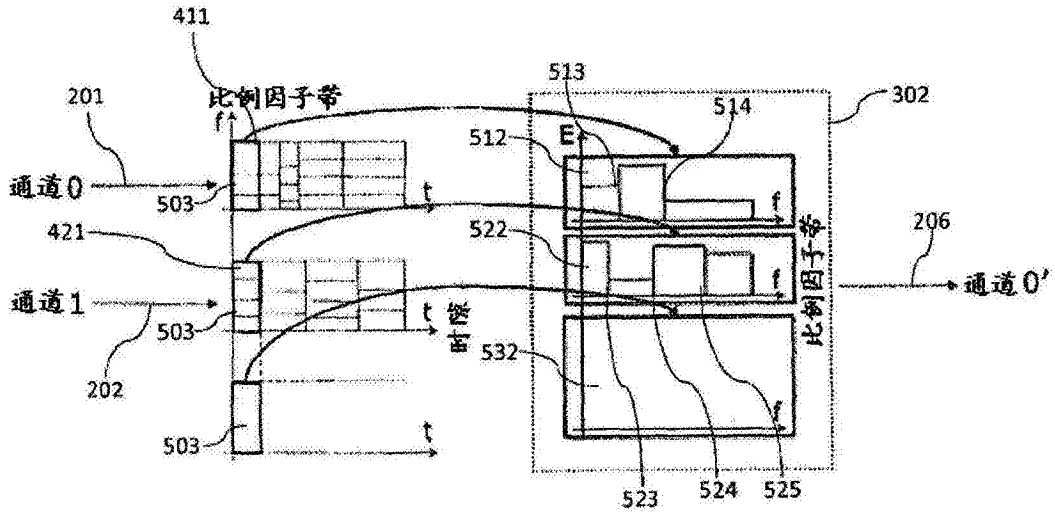


图5a

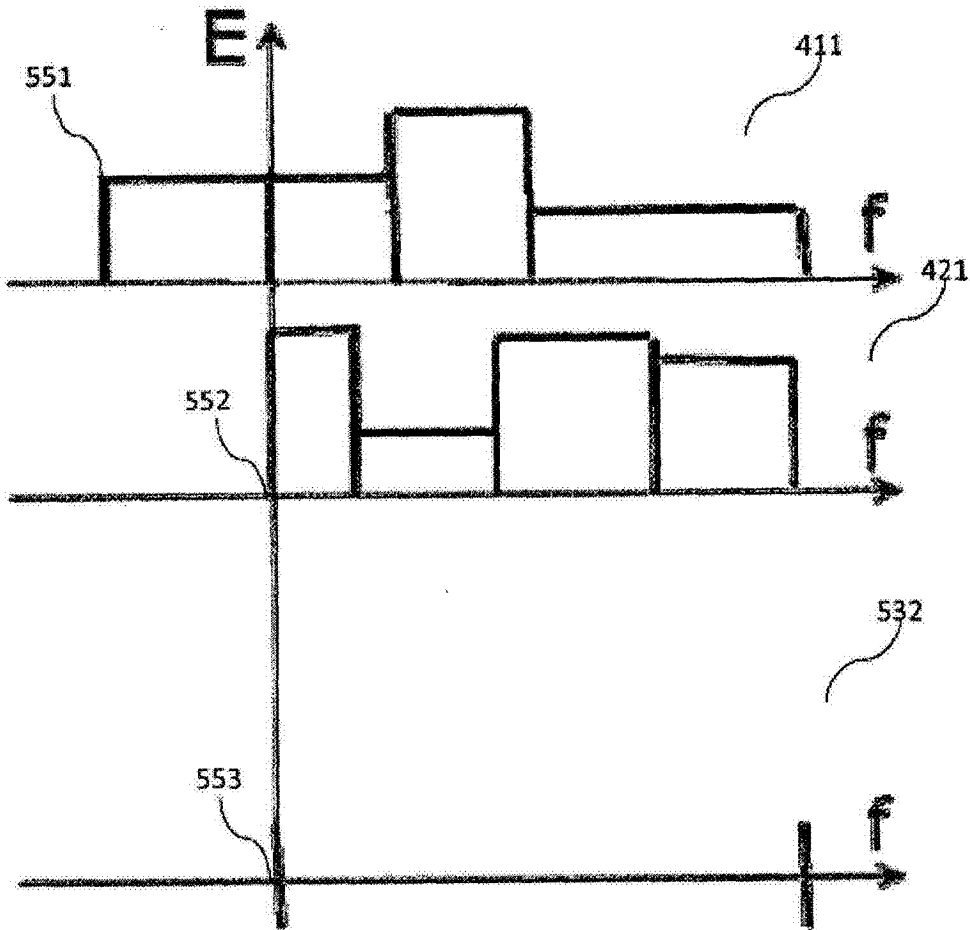


图5b

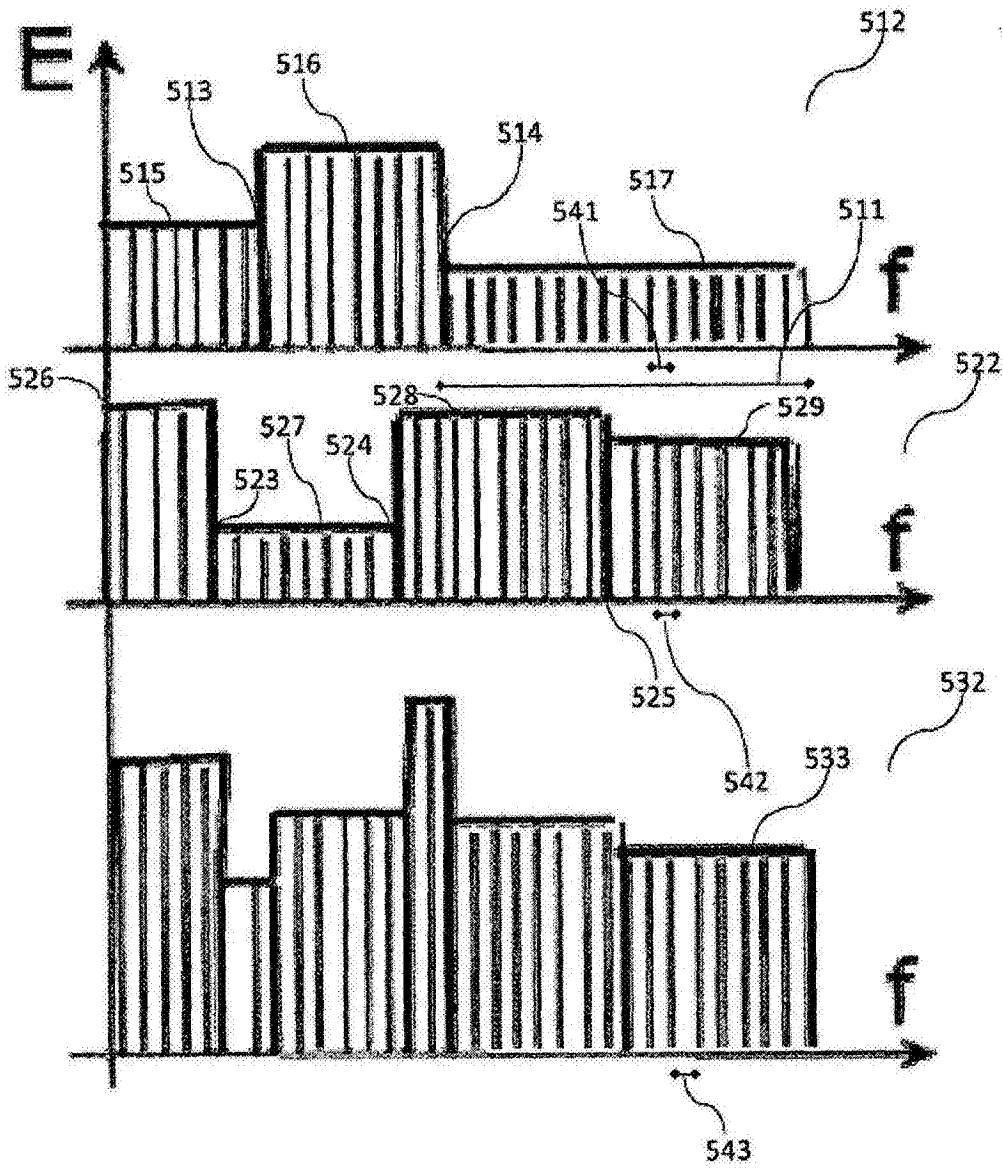


图5c

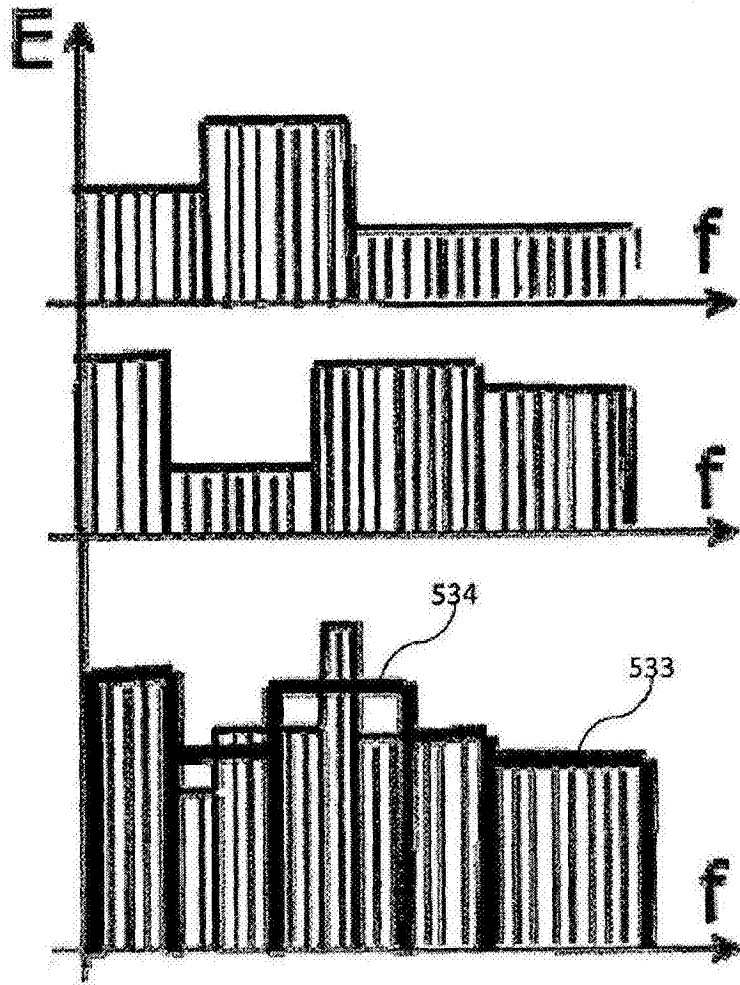


图5d

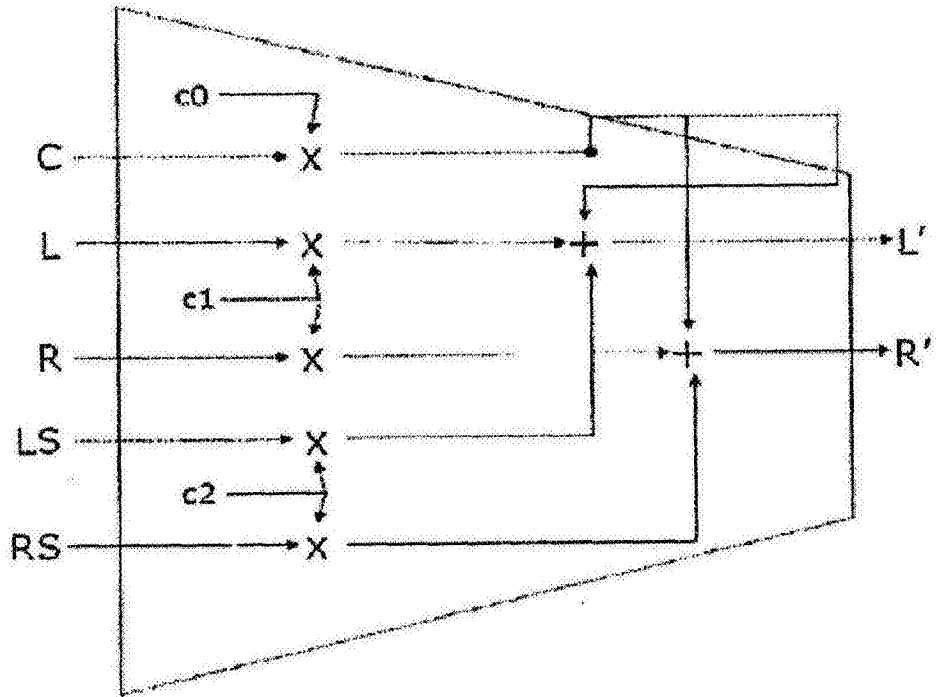


图6