



(12) 发明专利

(10) 授权公告号 CN 102789782 B

(45) 授权公告日 2015. 10. 14

(21) 申请号 201210232608. 8

(22) 申请日 2009. 03. 04

(30) 优先权数据

61/033, 590 2008. 03. 04 US

(62) 分案原申请数据

200980116080. 4 2009. 03. 04

(73) 专利权人 弗劳恩霍夫应用研究促进协会

地址 德国慕尼黑

(72) 发明人 马库斯·施内尔 曼弗雷德·卢茨基

马库斯·马特拉斯

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 杨静

(51) Int. Cl.

G10L 19/008(2013. 01)

G10L 21/038(2013. 01)

(56) 对比文件

CN 1875402 A , 2006. 12. 06,

CN 1998046 A , 2007. 07. 11,

US 2003/0125939 A1 , 2003. 07. 03,

US 5463424 A , 1995. 10. 31,

US 5717764 A , 1998. 02. 10,

审查员 贾杨

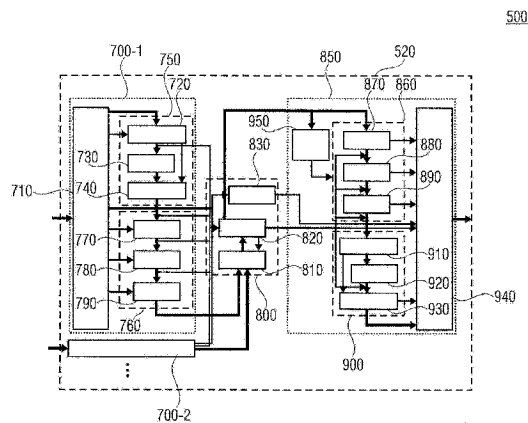
权利要求书1页 说明书27页 附图12页

(54) 发明名称

对输入数据流进行混合以及从中产生输出数据流

(57) 摘要

描述了一种用于对多个输入数据流 (510) 进行混合的设备 (500), 其中输入数据流 (510) 各包括频谱域中的音频数据的帧 (540), 输入数据流 (510) 的帧 (540) 包括针对多个频谱分量的频谱信息。所述设备 (500) 包括: 处理单元 (520), 适于比较多个输入数据流 (510) 的帧 (540); 基于所述比较, 针对输出数据流 (530) 的输出帧 (550) 的频谱分量, 确定多个输入数据流 (510) 的正好一个输入数据流 (510); 通过拷贝确定的输入数据流 (510) 的帧的对应频谱分量的信息中的至少一部分, 以描述输出数据流 (530) 的输出帧 (550) 的频谱分量, 来产生输出数据流。



1. 一种用于对多个输入数据流进行混合的设备 (500), 其中输入数据流 (510) 各包括频谱域中的音频数据的帧, 输入数据流 (510) 的帧 (540) 包括针对多个频谱分量的频谱信息,

所述设备 (500) 包括:

处理单元 (520), 适于基于心理声学模型, 考虑声道间掩蔽, 比较多个输入数据流 (510) 的帧,

其中, 处理单元 (520) 还适于基于所述比较, 针对输出数据流 (530) 的输出帧 (550) 的频谱分量, 确定多个输入数据流 (510) 的正好一个输入数据流 (510);

其中, 处理单元 (520) 还适于通过拷贝确定的输入数据流 (510) 的帧 (540) 的对应频谱分量的信息中的至少一部分, 以描述输出数据流 (530) 的输出帧 (550) 的频谱分量, 来产生输出数据流。

2. 根据权利要求 1 所述的设备 (500), 其中, 处理单元 (520) 适于使得对多个输入数据流 (510) 的帧的比较是基于与两个不同输入数据流 (510) 的帧 (540) 的相同频谱分量相对应的至少两个频谱信息的。

3. 根据权利要求 1 所述的设备, 其中, 设备 (500) 适于使得多个频谱分量中的频谱分量与频率或频带相对应。

4. 根据权利要求 1 所述的设备 (500), 其中, 处理单元 (520) 适于使得产生输出数据流包括: 仅从确定的输入数据流 (510) 的帧拷贝对应频谱分量的信息中的至少一部分, 以描述输出数据流 (530) 的输出帧 (550) 的频谱分量。

5. 根据权利要求 1 所述的设备 (500), 其中, 处理单元 (520) 适于使得产生输出数据流包括: 从确定的输入数据流 (510) 的帧拷贝频谱域中与频谱分量相对应的音频数据。

6. 根据权利要求 1 所述的设备 (500), 其中, 多个输入数据流 (510) 的输入数据流 (510) 各包括关于时间的频谱域中的音频数据的帧序列, 其中, 处理单元 (520) 适于使得对帧 (540) 的比较基于仅与帧序列的公共时间索引相对应的帧。

7. 根据权利要求 1 所述的设备 (500), 其中, 处理单元 (520) 适于使得产生输出数据流 (530) 保持与确定的输入流 (510) 的帧的对应频谱分量的信息中的至少一部分的量化级分布可比的量化级分布。

8. 根据权利要求 1 所述的设备 (500), 其中, 对应频谱分量的信息中的至少一部分包括: 与量化级、感知噪声替换 (PNS) 参数、时间噪声替换 (TNS) 参数或频谱带复制 (SBR) 参数有关的信息。

9. 一种用于对多个输入数据流 (510) 进行混合的方法, 其中, 输入数据流 (510) 各包括频谱域中的音频数据的帧 (540), 输入数据流 (510) 的帧 (540) 包括多个频谱分量,

所述方法包括:

基于心理声学模型, 考虑声道间掩蔽, 比较多个输入数据流 (510) 的帧 (540),

基于所述比较, 针对输出数据流 (530) 的输出帧 (550) 的频谱分量, 确定多个输入数据流 (510) 的正好一个输入数据流 (510);

通过拷贝确定的输入数据流 (510) 的帧的对应频谱分量的信息中的至少一部分, 来描述输出数据流 (530) 的输出帧的频谱分量, 来产生输出数据流 (530)。

对输入数据流进行混合以及从中产生输出数据流

[0001] 本申请是申请日为 2009 年 3 月 4 日的中国专利申请 No. 200980116080.4 (发明名称：“对输入数据流进行混合以及从中产生输出数据流”) 的分案申请。

技术领域

[0002] 根据本发明的实施例涉及对多个输入数据流进行混合以获得输出数据流,并且相应地通过对第一和第二输入数据流进行混合来产生输出数据流。输出数据流例如可以在包括视频会议系统和电话会议系统在内的会议系统领域中使用。

背景技术

[0003] 在许多应用中,要以以下方式来处理多于一个音频信号:根据多个音频信号,要产生一个信号或至少减小数目的信号,这通常称为“混合”。因此,混合音频信号的过程可以称为将多个单独音频信号捆绑为生成信号。例如在创建音乐用于致密光盘(“配音”)时,使用这一过程。在这种情况下,典型地,将不同乐器的不同音频信号与包括声乐表演(歌唱)的一个或多个音频信号混合为歌曲。

[0004] 混合处理扮演重要角色的其他应用领域是视频会议系统和电话会议系统。典型地,这种系统能够通过采用中心服务器来对会议中的多个在空间上分散的参与者进行连接,中心服务器对注册参与者的输入视频和音频数据进行适当混合,并将生成信号返回发送给每个参与者。该生成信号或输出信号包括所有其他会议参与者的音频信号。

[0005] 在现代数字会议系统中,多个部分矛盾的目标和方面互相竞争。必须考虑重构音频信号的质量,以及针对不同类型音频信号(例如话音信号相比于一般音频信号和音乐信号),一些编码和解码技术的实用性和有效性。在设计和实现会议系统时还要考虑的其他方面是可用带宽和延迟问题。

[0006] 例如,在对一方面质量和另一方面带宽进行平衡时,在多数情况下,折衷不可避免。然而,可以通过实现现代编码和解码技术,如 AAC-ELD 技术(AAC = 高级音频编解码;ELD = 增强低延迟),来实现与质量相关的改进。然而,可实现的质量可能以更多的基本问题和方面对采用这种现代技术的系统造成负面影响。

[0007] 仅列出要满足的一个挑战,所有数字信号传输面临所需量化的问题,至少在原理上,该问题在无噪声的模拟系统中的理想环境下是可避免的。由于量化过程,不可避免地将特定量的量化噪声引入要处理的信号。为了对抗可能和可听的失真,可能尝试增加量化级的数目,从而相应提高了量化分辨率。然而,这导致要发送更大数目的信号值,从而导致要发送的数据量增大。换言之,通过减小由量化噪声引入的可能失真来提高质量,在特定环境下将增大要发送的数据量,并且可能最终违反施加于传输系统上的带宽限制。

[0008] 在会议系统的情况下,由于典型地要处理多于一个输入音频信号,因此对质量、可用带宽和其他参数之间的权衡进行改进的挑战甚至更加复杂。因此,在产生由会议系统产生的输出信号或生成信号时,必须考虑多于一个音频信号所施加的边界条件。

[0009] 尤其是考虑到实现具有充分低延迟的会议系统的另一挑战:在不引入参与者认为

不可接受的实质延迟的情况下,实现会议参与者之间的直接通信,这种挑战进一步提升。

[0010] 在会议系统的低延迟实现中,典型地在延迟源的数目方面对其进行限制,另一方面,这可能导致在通过叠加或添加相应信号可以实现音频信号混合的时域之外处理数据的挑战。

[0011] 一般而言,赞成仔细选择适合于会议系统的质量、可用带宽与其他参数之间的权衡,以便应对用于实时混合的处理开销、降低所需的硬件数量、并且保持硬件以及不包括音频质量的合理传输开销的成本。

[0012] 为了减少传输的数据量,现代音频编解码器通常利用高复杂度工具来描述与相应音频信号的频谱分量有关的频谱信息。通过利用这种基于心理声学现象和检查结果的工具,可以实现部分矛盾参数与边界条件(例如,根据传输数据的重构音频信号的质量、计算复杂度、比特率、以及其他参数)之间改进的权衡。

[0013] 例如,针对这种工具的示例是感知噪声替换(PNS)、时间噪声成形(TNS)、以及频谱带复制(SBR),仅列举少数。所有这些技术基于描述具有与不使用这些工具所基于的数据流相比减少比特数目的频谱信息的至少一部分,可以将更多比特分配给频谱中频谱的重要部分。因此,在保持比特率的同时,可以通过使用这样的工具来提高质量感知级。当然,可以选择不同权衡,从而减少音频信号的每帧传输的比特数目,同时保持总音频印象。还可以同样极好地实现这两个极端之间的不同权衡。

[0014] 这些工具也可以使用在电信应用中。然而,当在这种通信情况下存在多于两个参与者时,采用会议系统用于混合多于两个参与者的两个或多个比特流是有利的。类似于上述的情况出现在纯基于音频的或电话会议情况以及视频会议情况。

[0015] 例如,在US 2008/0097764A1中描述了一种在频域中工作的会议系统,该系统执行频域中的实际混合,并从而省略将输入音频信号重变换到时域中的操作。

[0016] 然而,这里所描述的会议系统不考虑如上所述实现以更紧缩的方式对至少一个频谱分量的频谱信息进行描述的工具的可能性。因此,这样的会议系统需要额外的变换步骤来将被提供给会议系统的音频信号至少重构到以下程度:相应音频信号存在于频域中。此外,还需要基于上述附加工具对生成的混合音频信号进行重新变换。然而,这些重新变换和变换步骤要求复杂算法的应用,这会导致计算复杂度的增加,例如,在便携式、能量方面要求严格的应用中,导致增加的能量消耗,并因此导致受限的操作时间。

[0017] 因此,根据本发明的实施例要解决的问题是,实现一种适合于会议系统的质量、可用带宽和其他参数之间的改进权衡,或者实现如上所述会议系统中所需计算复杂度的降低。

发明内容

[0018] 该目的可以通过一种根据权利要求1或12所述的设备、一种根据权利要求10或26所述用于对多个输入数据流进行混合的方法、或者一种根据权利要求11或27所述的计算机程序来实现。

[0019] 根据第一方面,根据本发明的实施例基于以下发现:当混合多个输入数据流时,通过基于比较来确定输入数据流并且将至少部分频谱信息从确定的输入数据流中拷贝到输出数据流中,上述参数与目标之间的改进权衡是可实现的。通过从一个输入数据流中拷贝

至少一部分频谱信息,可以省略重新量化,并从而消除了与重新量化相关联的重新量化噪声。在不可确定主导输入流的频谱信息情况下,在频域中混合相应频谱信息可以由根据本发明的实施例来执行。

[0020] 例如,比较可以基于心理声学模型。比较还可以涉及与来自至少两个不同输入数据流的公共频谱分量(例如,频率或频带)相对应的频谱信息。因此,该比较可以是声道间比较。在比较基于心理声学模型的情况下,从而可以考虑声道间掩蔽来描述比较。

[0021] 根据第二方面,根据本发明的实施例基于以下发现:在混合第一输入数据流和第二输入数据流以产生输出数据流期间所执行的操作复杂度可以通过考虑与相应输入数据流的有效载荷数据相关联的控制值来降低,其中,控制值指示有效载荷数据表示相应音频信号的相应频谱信息或频谱域的至少一部分的方式。在两个输入数据流的控制值相等的情况下,可以省略输出数据流的相应帧处的频谱域的方式的新判定,取而代之,输出流产生可以仅依赖于已经存在的判定,并且统一由输入数据流的编码器来确定,即,采用输入数据流的控制值。根据控制值所指示的方式,甚至能够且优选地避免将相应有效载荷数据重新变换回到表示频谱域的另一方式(例如,具有每时间/空间采样一个频谱值的正常或普通方式)。在后一种情况下,对有效载荷数据直接进行处理以获得输出数据流的相应有效载荷数据,以及等于第一和第二输入数据流的控制值的控制值可以例如通过 PNS 或以下更详细描述类似音频特征,利用意味着“不改变表示频谱域的方式”的“方向性”来产生。

[0022] 在根据本发明实施例的实施例中,控制值仅与至少一个频谱分量有关。此外,在根据本发明的实施例中,当第一输入数据流和第二数据流的数据流的帧与关于两个输入数据流的适当帧序列的公共时间索引相对应时,可以执行这样的操作。

[0023] 在第一和第二数据流的控制值不相等的情况下,根据本发明的实施例可以执行以下步骤:将第一和第二输入数据流之一的一个帧的有效载荷数据进行变换,以获得另一个输入数据流的帧的有效载荷数据的表示。然后可以基于变换后的有效载荷数据和其他两个流的有效载荷数据,产生输出数据流的有效载荷数据。在一些情况下,根据本发明实施例,可以直接执行将一个输入数据流的帧的有效载荷数据变换成其他输入数据流的帧的有效载荷数据的表示,而无需将相应音频信号变换回到普通(plain)频域。

附图说明

[0024] 以下,参考以下附图来描述根据本发明的实施例。

[0025] 图 1 示出了会议系统的框图;

[0026] 图 2 示出了基于一般音频编解码器的会议系统的框图;

[0027] 图 3 示出了使用比特流混合技术,在频域中操作的会议系统的框图;

[0028] 图 4 示出了包括多个帧在内的数据流的示意图;

[0029] 图 5 示意了频谱分量和频谱数据或信息的不同形式;

[0030] 图 6 更详细示出了根据本发明实施例的用于对多个输入数据流进行混合的设备;

[0031] 图 7 示出了根据本发明实施例的图 6 的设备的操作模式;

[0032] 图 8 示出了在会议系统环境中根据本发明另一实施例用于对多个输入数据流进行混合的设备的框图;

[0033] 图 9 示出了根据本发明实施例的用于产生输出数据流的设备的简要框图;

- [0034] 图 10 示出了根据本发明实施例的用于产生输出数据流的设备的更详细框图；
- [0035] 图 11 示出了在会议系统环境中根据本发明另一实施例的用于从多个输入数据流产生输出数据流的设备的框图；
- [0036] 图 12a 示出了用于 PNS 实现的根据本发明实施例的输出数据流产生的操作；
- [0037] 图 12b 示出了用于 SBR 实现的根据本发明实施例的输出数据流产生设备的操作；以及
- [0038] 图 12c 示出了用于 M/S 实现的根据本发明实施例的输出数据流产生设备的操作。

具体实施方式

[0039] 参照图 4 至 12C, 更详细描述根据本发明的不同实施例。然而, 在更详细描述这些实施例之前, 首先参照图 1 至 3, 考虑在会议系统的框架中可能变得重要的挑战和需求, 来给出简要介绍。

[0040] 图 1 示出了会议系统 100 的框图, 会议系统 100 也可以称为多点控制单元 (MCU)。从与其功能相关的描述中变得显而易见, 如图 1 所示, 会议系统 100 是在时域中操作的系统。

[0041] 如图 1 所示, 会议系统 100 适于经由合适数目的输入 110-1, 110-2, 110-3, ... (图 1 中仅示出其中 3 个) 来接收多个输入数据流。每个输入 110 耦合至相应解码器 120。更准确地, 针对第一输入数据流的输入 110-1 耦合至第一解码器 120-1, 而第二输入 110-2 耦合至第二解码器 120-2, 第三输入 110-3 耦合至第三解码器 120-3。

[0042] 会议系统 100 还包括合适数目的加法器 130-1, 130-2, 130-3, ... 图 1 仍示出了其中 3 个。每个加法器与会议系统 100 的输入 110 之一相关联。例如, 第一加法器 130-1 与第一输入 110-1 和对应的解码器 120-1 相关联。

[0043] 每个加法器 130 耦合至所有解码器 120 的输出, 但是输入 110 所耦合的解码器 120 除外。换言之, 第一加法器 130-1 耦合至所有解码器 120, 但第一解码器 120-1 除外。相应地, 第二加法器 130-2 耦合至所有解码器 120, 但第二解码器 120-2 除外。

[0044] 每个加法器 130 还包括耦合至一个编码器 140 的输出。因此, 第一加法器 130-1 输出耦合至第一编码器 140-1。相应地, 第二和第三加法器 130-2、130-3 也分别耦合至第二和第三编码器 140-2、140-3。

[0045] 继而, 每个编码器 140 耦合至相应输出 150。换言之, 例如第一编码器耦合至第一输出 150-1。第二和第三编码器 140-2、140-3 也分别耦合至第二和第三输出 150-2、150-3。

[0046] 为了能够更详细地描述图 1 所示的会议系统 100 的操作, 图 1 还示出了第一参与者的会议终端 160。会议终端 160 例如可以是数字电话 (例如 ISDN 电话 (ISDN = 综合业务数字网))、包括 IP 语音基础设施的系统、或类似终端。

[0047] 会议终端 160 包括编码器 170, 编码器 170 耦合至会议系统 100 的第一输入 110-1。会议终端 160 还包括解码器 180, 解码器 180 耦合至会议系统 100 的第一输出 150-1。

[0048] 在其他参与者的站点处还可以有类似的会议终端 160。仅为了简单起见, 图 1 中未示出这些会议终端。还应当注意, 会议系统 100 和会议终端 160 目前不需要在物理上彼此紧邻。会议终端 160 和会议系统 100 可以布置在不同站点, 例如可以仅通过 WAN 技术 (WAN = 广域网) 连接。

[0049] 会议终端 160 还可以包括或连接至附加组件,如麦克风、放大器和扬声器或耳机,以实现以更全面的方式与人类用户交换音频信号。仅为了简单起见,图 1 中未示出这些。

[0050] 如上所述,图 1 中示出的会议系统 100 是在时域中操作的系统。例如,当第一参与者对麦克风(图 1 中未示出)说话时,会议终端 160 的编码器 170 将相应音频信号编码为对应比特流,并将比特流发送至会议系统 100 的第一输入 110-1。

[0051] 在会议系统 100 内,比特流由第一解码器 120-1 解码,并变换回时域。由于第一解码器 120-1 耦合至第二和第三混合器 130-1、130-3,因此通过简单地将重构的音频信号与分别来自第二和第三参与者的其他重构音频信号相加,在时域中对第一参与者产生的音频信号进行混合。

[0052] 对于分别由第二和第三输入 110-2、110-3 接收并由第二和第三解码器 120-2、120-3 处理的、由第二和第三参与者提供的音频信号而言,也同样如此。然后,将第二和第三参与者的这些重构音频信号提供给第一混合器 130-1,第一混合器 130-1 继而将时域中的相加后的音频信号提供给第一编码器 140-1。编码器 140-1 对相加后的音频信号进行重新编码,以形成比特流,并在第一输出 150-1 处将其提供给第一参与者会议终端 160。

[0053] 类似地,第二和第三编码器 140-2、140-3 也对分别从第二和第三加法器 130-2、130-3 接收到的时域中相加后的音频信号进行编码,并分别经由第二和第三输出 150-2、150-3 将编码的数据发送回相应参与者。

[0054] 为了执行实际混合,以非压缩形式对音频信号进行完全解码并相加。此后,可选地,可以通过对相应输出信号进行压缩来执行等级调整,以防止限幅效应(即超过允许值范围)。当单采样值升至允许值范围以上或降至允许值范围以下,使得对应值被截止(限幅)时,可能出现限幅。在 16 比特量化的情况下(如在 CD 情况下所采用的),针对每采样值,-32768 与 32767 之间的整数值范围可用。

[0055] 为了对抗对信号的可能的过度操作或操作不足,采用压缩算法。这些算法限制了特定阈值以上或以下的发展,以将采样值保持在可允许的值范围内。

[0056] 在会议系统(如图 1 所示的会议系统 100)中对音频数据进行编码时,接受一些缺点,从而以最容易可实现的方式,在未编码状态下执行混合。此外,附加地,编码的音频信号的数据速率被限制在所传输频率的较小范围,因为根据奈奎斯特-仙农采样定理,较小的带宽允许较低的采样频率,从而允许较少的数据。奈奎斯特-仙农采样定理指出,采样频率取决于采样信号的带宽并且大小需要(至少)为该带宽的 2 倍。

[0057] 国际电信联盟(ITU)及其电信标准化部门(ITU-T)已经开发了针对多媒体会议系统的多个标准。H. 320 是针对 ISDN 的标准会议协议。H. 323 定义了针对基于分组的网络(TCP/IP)的标准会议系统。H. 324 定义了针对模拟电话网络和无线通信系统的会议系统。

[0058] 在这些标准内,不仅定义了信号的传输,还定义了音频数据的编码和处理。由一个或多个服务器,根据标准 H. 321 的所谓多点控制单元(MCU)来进行会议的管理。多点控制单元还负责多个参与者的视频和音频数据的处理和分发。

[0059] 为了实现这一点,多点控制单元向每个参与者发送包括所有其他参与者的音频数据在内的混合输出或生成信号,并将该信号提供给相应参与者。图 1 不仅示出了会议系统 100 的框图,还示出了这种会议情形下的信号流。

[0060] 在 H. 323 和 H. 320 标准的框架中,定义了 G. 7xx 类别的音频编解码器以在相应会

议系统中操作。标准 G. 711 用于线缆绑定的电话系统中的 ISDN 传输。在 8kHz 的采样频率处, G. 711 标准覆盖 300 至 3400Hz 之间的音频带宽, 需要 8 比特 (量化) 深度的 64Kbit/s 的比特率。该编码由产生仅 0. 125ms 的非常低延迟的、称为 μ 律或 A 律的简单对数编码形成。

[0061] G. 722 标准以 16kHz 的采样频率, 对从 50 至 7000Hz 的较大音频带宽进行编码。因此, 以 1. 5ms 的延迟, 在比特率 48、56 或 64Kbit/s 处, 与频带更窄的 G. 7xx 音频编解码器相比, 该编解码器实现了更好的质量。此外, 存在两个其他改进: G. 722. 1 和 G. 722. 2, 在甚至更低的比特率提供了可比的话音质量。G722. 2 允许以 25ms 的延迟, 在 6. 6kbit/s 与 23. 85kbit/s 之间进行比特率选择。

[0062] G. 729 标准典型用于 IP 电话通信 (也称为 IP 语音通信 (VoIP)) 的情况。该编解码器针对话音进行了优化, 并发送分解的话音参数的集合, 以便随后与误差信号一起进行合成。因此, 与 G. 711 标准相比, G. 729 标准以可比的采样率和音频带宽, 实现了近似 8kbit/s 的明显更好的编码。然而, 这种更加复杂的算法造成了近似 15ms 的延迟。

[0063] 作为缺点, G. 7xx 编解码器是针对话音编码进行优化的, 除了较窄的频率带宽之外, 在对音乐与话音一起或者纯音乐进行编码时, 表现出明显的问题。

[0064] 因此, 尽管在发送和处理话音信号时, 如图 1 所示的会议系统 100 可以用于可接受的质量, 但是在采用针对话音优化的低延迟编解码器时, 不能令人满意地处理一般音频信号。

[0065] 换言之, 采用用于对话音信号进行编码和解码的编解码器来处理一般音频信号, 包括例如具有音乐的音频信号, 在质量方面不能产生满意的结果。如图 1 所示, 通过在会议系统 100 的框架中采用用于对一般音频信号进行编码和解码的音频编解码器, 能够提高质量。然而, 如在图 2 的上下文中更详细阐述的, 在这种会议系统中采用一般音频编解码器可能导致其他不利效果, 例如增大延迟 (仅列出一项)。

[0066] 然而, 在更详细地描述图 2 之前, 应当注意, 在本描述中, 当对象在实施例或附图中出现多于一次, 或者在多个实施例或附图中出现时, 使用相同或相似的参考标记来表示相应对象。除非另外进行显式或隐式表示, 使用相同或相似参考标记来表示的对象可以以相似或相等的方式来实现, 例如在其电路、编程、特征或其他参数方面。因此, 在附图的多个实施例中出现并且使用相同或相似的参考标记来表示的对象可以被实现为具有相同的规格、参数和特征。自然, 也可以实现改变和适配, 例如在边界条件或其他参数根据不同附图或根据不同实施例而改变的情况下。

[0067] 此外, 在以下概括中, 参考标记将用于表示一组或一类对象, 而不是单个对象。在图 1 的框架中已经这样做, 例如在将第一输入表示为 110-1, 将第二输入表示为 110-2, 将第三输入表示为 110-3 时, 已经仅以概括参考标记 110 的方式讨论了这些输入。换言之, 除非另外显式注明, 与使用概括参考标记来表示的对象相关的描述的部分也与具有相应单独参考标记的其他对象相关。

[0068] 由于对于使用相同或相似参考标记来表示的对象而言也是如此, 这些措施有助于缩短描述并以更清楚和简要的方式来描述其中公开的实施例。

[0069] 图 2 示出了另一会议系统 100 以及会议终端 160 的框图, 两者都与图 1 中所示的类似。图 2 中示出的会议系统 100 还包括: 输入 110、解码器 120、加法器 130、编码器 140 和

输出 150, 这些以与图 1 所示的会议系统 100 同样的方式进行互连。图 2 中所示的会议终端 160 也包括编码器 170 和解码器 180。因此, 参照图 1 所示的会议系统 100 的描述。

[0070] 然而, 图 2 所示的会议系统 100, 以及图 2 所示的会议终端 160 适于使用一般音频编解码器 (编码器 - 解码器)。因此, 编码器 140、170 中的每一个包括在量化器 / 编码器 200 之前耦合的时间 / 频率转换器 190 的串联连接。时间 / 频率转换器 190 在图 2 中也示意为“T/F”, 而图 2 中将量化器 / 编码器 200 标记为“Q/C”。

[0071] 解码器 120、180 均包括解码器 / 解量化器 210 (图 2 中称为“Q/C-1”), 与频率 / 时间转换器 220 (图 2 中称为“T/F-1”) 串联连接。仅为了简单起见, 仅在编码器 140-3 和解码器 120-3 的情况下, 将时间 / 频率转换器 190、量化器 / 编码器 200 和解码器 / 解量化器 210 以及频率 / 时间转换器 220 如此标记。然而, 以下描述也涉及其他这种元件。

[0072] 从编码器开始, 如编码器 140 或编码器 170, 转换器 190 将提供给时间 / 频率转换器 190 的音频信号从时域转换至频域或频率相关域。此后, 在时间 / 频率转换器 190 产生的频谱表示中, 对转换后的音频数据进行量化和编码, 以形成比特流, 然后例如在编码器 140 的情况下, 将比特流提供给会议系统 100 的输出 150。

[0073] 对于解码器, 如解码器 120 或解码器 180, 首先对提供给解码器的比特流进行解码和重新量化, 以形成至少一部分音频信号的频谱表示, 然后, 频率 / 时间转换器 220 将频谱表示转换回时域。

[0074] 因此, 时间 / 频率转换器 190, 以及反转元件, 频率 / 时间转换器 220, 分别适于产生向其提供的至少一段音频信号的频谱表示和将频谱表示重新变换为时域中的音频信号的对应部分。

[0075] 再将音频信号从时域转换为频域、以及从频域转换回时域的过程中, 可能出现偏差, 使得重新建立的、重构的、或解码的音频信号可能不同于原始或源音频信号。在量化器编码器 200 和重新编码器 210 的框架中执行的量化和解量化的附加步骤可能加入另外的伪像。换言之, 原始音频信号, 以及重新建立的音频信号, 可能互不相同。

[0076] 例如, 时间 / 频率转换器 190, 以及频率 / 时间转换器 220 可以基于 MDCT (修正离散余弦变换)、MDST (修正离散正弦变换)、基于 FFT 的转换器 (FFT = 快速傅立叶变换) 或另一基于傅立叶的转换器来实现。在量化器 / 编码器 200 和解码器 / 解量化器 210 的框架中的量化和重新量化可以例如基于线性量化、对数量化、或另一更复杂的量化算法 (例如更具体地考虑人类的听觉特性) 来实现。量化器 / 编码器 200 和解码器 / 解量化器 210 的编码器和解码器部分可以例如通过采用霍夫曼编码或霍夫曼解码方案来工作。

[0077] 然而, 在这里描述的不同实施例和系统中, 也可以采用更复杂的时间 / 频率和频率 / 时间转换器 190、220 以及更复杂的量化器 / 编码器和解码器 / 解量化器 200、210, 作为例如 AAC-ELD 编码器 (作为编码器 140、170) 和 AAC-ELD 解码器 (作为解码器 120、180) 的一部分或形成 AAC-ELD 编码器和 AAC-ELD 解码器。

[0078] 不言自明, 在会议系统 100 和会议终端 160 的框架中, 实现相同或至少兼容的编码器 170、140 和解码器 180、120 是可取的。

[0079] 如图 2 所示, 基于一般音频信号编码和解码方案的会议系统 100 还在时域中执行音频信号的实际混合。向加法器 130 提供时域中的重构音频信号, 以执行叠加并向后续编码器 140 的时间 / 频率转换器 190 提供时域中的混合信号。因此, 会议系统再次包括解码

器 120 和编码器 140 的串联连接,这是如图 1 和 2 所示的会议系统 100 典型地被称为“串联编码系统”的原因。

[0080] 串联 (tandem) 编码系统通常表现出高复杂度的缺点。混合的复杂度高度依赖于所采用的解码器和编码器的复杂度,并且在多个音频输入和音频输出信号的情况下可能明显倍增。此外,由于多数编码和解码方案是无损的事实,在图 1 和 2 所示的会议系统 100 中采用的串联编码方案典型地造成对质量的负面影响。

[0081] 作为另一缺点,解码和编码的重复步骤也加大了会议系统的输入 110 和输出 150 之间的总体延迟 (也称为端到端延迟)。根据所使用的解码器和编码器的初始延迟,会议系统 100 本身可能将延迟增大至使得在会议系统的框架中的使用变得没有吸引力 (如果不是恼人或甚至不可能的话) 的程度。通常,50ms 的延迟被认为是会话中参与者可以接受的最大延迟。

[0082] 作为延迟的主要来源,时间 / 频率转换器 190 以及频率 / 时间转换器 220 对会议系统 110 的端到端延迟负责,附加延迟由会议终端 160 施加。其他元件,即量化器 / 编码器 200 和解码器 / 解量化器 210 造成的延迟相对次要,因为与时间 / 频率转换器和频率 / 时间转换器 190、220 相比,这些组件可以在更高得多的频率上操作。多数时间 / 频率转换器和频率 / 时间转换器 190、220 是块操作或帧操作的,这意味着,在许多情况下,必须考虑作为时间量的最小延迟,该最小延迟等于填充具有块的帧的长度的缓冲器或存储器所需的时间。然而该时间受到采样频率的显著影响,采样频率典型地在几 kHz 至几十 kHz 的范围内;而量化器 / 编码器 200 以及解码器 / 解量化器 210 的操作速度主要由基础系统的时钟频率确定。这典型地要更大至少 2、3、4 或更多数量级。

[0083] 因此,在采用一般音频信号编解码器的会议系统中,已经引入了所谓比特流混合技术。例如,比特流混合方法可以基于 MPEG-4AAC-ELD 编解码器来实现,该编解码器提供了避免上述由串联编码引入的至少一些缺陷的可能性。

[0084] 然而,应当注意,原则上,如图 2 所示的会议系统 100 还可以基于与先前描述的 G. 7xx 编解码器族的基于话音的码相比具有类似比特率和明显更大频率带宽的 MPEG-4AAC-ELD 编解码器来实现。这直接还意味着,可以以明显增大的比特率为代价,能够实现针对所有信号类型的明显更好的音频质量。尽管 MPEG-4AAC-ELD 提供了在 G. 7xx 编解码器延迟范围内的延迟,但是,在图 2 所示的会议系统的框架中实现 MPEG-4AAC-ELD 可能不能产生实际会议系统 100。以下将关于图 3 来概述基于前述所谓比特流混合的更实际的系统。

[0085] 应当理解,仅为了简单起见,将主要关注 MPEG-4AAC-ELD 编解码器及其数据流和比特流。然而,在如图 3 中示意和示出的会议系统 100 的环境中,也可以采用其他编码器和解码器。

[0086] 图 3 示出了如在图 2 的上下文中描述的根据比特流混合原理来工作的会议系统 100 以及会议终端 160 的框图。会议系统 100 本身是图 2 中示出的会议系统 100 的简化版本。更精确地,图 2 中的会议系统 100 的解码器 120 已经由图 3 中示出的解码器 / 解量化器 220-1, 220-2, 210-3, ... 所取代。换言之,与图 2 和 3 中示出的会议系统 100 相比,已经去了解码器 120 的频率 / 时间转换器 120。类似地,图 2 的会议系统 100 的编码器 140 已经由量化器 / 编码器 200-1, 200-2, 200-3 所取代。因此,与图 2 和 3 中示出的会议系统 100

相比,已经去除了编码器 140 的时间 / 频率转换器 190。

[0087] 因此,加法器 130 不再在时域中操作,而是由于缺少频率 / 时间转换器 220 和时间 / 频率转换器 190 而在频域或频率相关域中操作。

[0088] 例如,在 MPEG-4AAC-ELD 编解码器的情况下,仅在会议终端 160 中出现的时间 / 频率转换器 190 和频率 / 时间转换器 220 基于 MDCT 变换。因此,在会议系统 100 内,混合器 130 直接在 MDCT 频率表示中的音频信号成分处。

[0089] 由于在图 2 所示的会议系统 100 的情况下,转换器 190、220 表示了延迟的主要来源,因此通过去除这些转换器 190、220,明显减小了延迟。此外,还明显减小了会议系统 100 内由两个转换器 190、220 引入的复杂度。例如,在 MPEG-2AAC 解码器的情况下,在频率 / 时间转换器 220 的框架中执行的逆 MDCT 变换占总复杂度的近似 20%。由于 MPEG-4 转换器也是基于类似变换,因此,通过从会议系统 100 中仅去除频率 / 时间转换器 220,可以去除总复杂度中并非无关紧要的成分。

[0090] 由于在 MDCT 变换的情况下或在类似的基于傅立叶的变换的情况下,这些变换是线性变换,因此可以在 MDCT 域或另一频域中混合音频信号。因此,这些变换具有数学加性属性,即:

$$[0091] \quad f(x+y) = f(x)+f(y), (1)$$

[0092] 以及数学同质性,即:

$$[0093] \quad f(a \cdot x) = a \cdot f(x), (2)$$

[0094] 其中 $f(x)$ 是变换函数, x 和 y 是其适当自变量, a 是实值或复值常数。

[0095] MDCT 变换或另一基于傅立叶的变换的这两种特征允许以在时域中混合相似的方式,在相应频域中进行混合。因此,可以同样地基于频谱值来执行所有计算。不需要将数据变换至时域。

[0096] 在一些情况下,可能要满足另一条件。针对所有相关频谱分量,在混合过程中,所有相关频谱数据就其时间索引而言应当相等。如果在变换期间采用所谓块切换技术,使得会议终端 160 的编码器可以根据特定条件在不同块长度之间自由切换,则最终可能不满足上述条件。由于在不同块长度和对应的 MDCT 窗长度之间进行切换,除非要混合的数据已经以相同的窗进行处理,否则块切换可能危及将各个频谱值唯一分配给时域中的采样的可能性。由于在具有分布式会议终端 160 的一般系统中,这可能最终得不到保证,因此可能需要复杂的插值,这继而可能造成附加延迟和复杂度。因此,最终不基于块长度切换来实现比特流混合过程可能是可取的。

[0097] 相反,AAC-ELD 编解码器基于单一块长度,因此能够更容易地保证上述分配或频率数据的同步,从而可以更容易地实现混合。换言之,图 3 中所示的会议系统 100 是一种能够在变换域或频域中执行混合的系统。

[0098] 如上所述,为了消除图 2 所示的会议系统 100 中的转换器 190、200 所引入的附加延迟,在会议终端 160 中使用的编解码器使用具有固定长度和形状的窗。这使得在不将音频流变换回时域的情况下,能够直接实现所描述的混合处理。这种方式能够限制额外引入的算法延迟量。此外,由于解码器中不存在逆变换步骤,编码器中不存在正变换步骤,因此降低了复杂度。

[0099] 然而,同样在如图 3 所示的会议系统 100 的框架中,可能需要在加法器 130 进行混

合之后对音频数据进行重新量化,这可能引入附加的量化噪声。例如,由于提供给会议系统 100 的不同音频信号的不同量化步长,可能造成附加量化噪声。因此,例如在非常低比特率传输的情况下(其中量化步长的数目已经有限),在频域或变换域中混合两个音频信号的过程可能导致所产生的信号中不期望的附加噪声量或其他失真。

[0100] 在以用于对多个输入数据流进行混合的设备的形式来描述根据本发明的第一实施例之前,关于图 4 来简要描述数据流或比特流以及其中包括的数据。

[0101] 图 4 示意性示出了比特流或数据流 250,比特流或数据流 250 包括频谱域中的至少一个(或者更经常多于一个)音频数据帧 260。更准确地,图 4 示出了频谱域中的 3 个音频数据帧 260-1、260-2 和 260-3。此外,数据流 250 还可以包括附加信息或附加信息块 270,例如指示音频数据的编码方式的控制值、其他控制值或与时间索引或其他相关数据有关的信息。自然,图 4 中所示的数据流 250 还可以包括附加帧,或者帧 260 可以包括对于一个声道的音频数据。例如,在立体声音频信号的情况下,每个帧 260 可以例如包括来自左声道、右声道的音频数据,从左和右声道导出的音频数据、或上述数据的任何组合。

[0102] 因此,图 4 示意了数据流 250 可以不仅包括频谱域中的音频数据帧,还包括附加控制信息、控制值、状态值、状态信息、协议相关值(例如校验和)等等。

[0103] 根据如图 1 至 3 的上下文所述的会议系统的具体实现,或者根据如下所述依照本发明实施例的设备的具体实现,具体地,根据关于图 9 至 12C 描述的那些具体实现,指示帧的关联有效载荷数据表示音频信号的频谱域或频谱信息的至少一部分的方式的控制值可以同样包括在帧 260 本身中,或者包括在附加信息的关联块 270 中。在控制值与频谱分量有关的情况下,可以将控制值编码到帧 260 本身中。然而,如果控制值与整个帧有关,则该控制值同样可以包括在附加信息的块 270 中。然而,如上所述,包括控制值的上述位置不需要包括在帧 260 或附加块的块 270 中。在控制值仅与单个或几个频谱分量有关的情况下,该控制值同样可以包括在块 270 中。另一方面,与整个帧 260 有关的控制值也可以包括在帧 260 中。

[0104] 图 5 示意性示出了例如数据流 250 的帧 260 中包括的与频谱分量有关的(频谱)信息。更准确地,图 5 示出了帧 260 的单一声道的频谱域中的信息的简化图。在频谱域中,音频数据帧可以例如以其强度值 I (作为频率 f 的函数)来描述。在离散系统(例如数字系统)中,频率分辨率也是离散的,使得频谱信息典型地仅针对特定频谱分量(如单独频率或窄带或子带)而存在。单独频率或窄带以及子带被称为频谱分量。

[0105] 图 5 示意性示出了针对 6 个单独频率 300-1, ..., 300-6 以及在图 5 所示的情况下包括 4 个单独频率的频带或子带 310 的强度分布。单独频率或对应窄带 300 以及子带或频带 310 形成频谱分量,对于所述频谱分量,帧包括与频谱域中的音频数据有关的信息。

[0106] 与子带 310 有关的信息可以例如是总体强度或平均强度值。除了强度或其他与能量有关的值(如幅度)之外,相应频谱分量本身的能量,或从能量或幅度、相位信息和其他信息导出的另一值也可以包括在帧中,从而被视为与频谱分量有关的信息。

[0107] 在描述了会议系统中所涉及的一些问题以及一些背景之后,对根据本发明第一方面的实施例进行描述,根据实施例,基于比较来确定输入数据流,以便将至少部分频谱信息从所确定的输入数据流拷贝到输出数据流,从而使得能够省略重新量化,并因此消除了与重新量化相关联的重新量化噪声。

[0108] 图 6 示出了用于对多个输入数据流 510 进行混合的设备 500 的框图, 示出了其中的两个输入数据流 510-1、510-2。设备 500 包括适于接收数据流 510 并产生输出数据流 530 的处理单元 520。输入数据流 510-1、510-2 中的每一个分别包括类似于图 5 的上下文中图 4 所示的帧 260 的帧 540-1、540-2, 包括频谱域中的音频数据。这再次通过图 6 所示的坐标系来示意, 在横坐标上, 示出了音频数据的频率 f , 在纵坐标上示出了音频数据的强度 I 。输出数据流 530 还包括输出帧 550, 输出帧 550 包括频谱域中的音频数据, 并同样由对应的坐标系来示意。

[0109] 处理单元 520 适于将多个输入数据流 510 的帧 540-1、540-2 进行比较。如以下更详细概括, 该比较例如可以基于心理声学模型, 考虑掩蔽效应以及人类听力特性的其他属性。基于该比较结果, 处理单元 520 还适于至少针对同时存在于两个帧 540-1、540-2 中的一个频谱分量 (例如图 6 所示的频谱分量 560), 精确确定多个数据流 510 中的一个数据流。然后, 处理单元 520 可以适于产生包括输出帧 550 在内的输出数据流 530, 使得从相应输入数据流 510 的确定的帧 540 拷贝与频谱分量 560 有关的信息。

[0110] 为了更精确, 处理单元 520 适于对多个输入数据流 510 的帧 540 的比较基于至少两个信息段: 强度值是有关的能量值的信息、与两个不同输入数据流 510 的帧 540 中相同频谱分量 560 相对应的信息。

[0111] 为了进一步示意上述, 图 7 示意性示出了与频谱分量 560 相对应的信息段 (强度 I) 的情况, 这里假定频谱分量 560 是第一输入数据流 510-1 的帧 540-1 的频率或窄频带。将该信息与对应的强度值 I 进行比较, 对应的强度值 I 是与第二输入数据流 510-2 的帧 540-2 的频谱分量 560 有关的信息段。例如, 可以基于仅包括一些输入流的混合信号与完整混合信号之间的能量比值的评估来进行比较。例如, 这可以根据以下等式来实现:

$$[0112] \quad E_c = \sum_{n=1}^N E_n \quad (3)$$

[0113] 以及

$$[0114] \quad E_{f^{(n)}} = \sum_{\substack{n=1 \\ n \neq 1}}^N E_f \quad (4)$$

[0115] 根据以下等式来计算比值 $r(n)$:

$$[0116] \quad r(n) = 20 \cdot \log \frac{E_{f^{(n)}}}{E_c} \quad (5)$$

[0117] 其中 n 是输入数据流的索引, N 是所有或有关输入数据流的数目。如果比值 $r(n)$ 足够高, 则可以认为输入数据流 510 的次要声道或次要帧被主要声道或主要帧所掩蔽。因此, 可以处理不相关性减小, 意即仅包括流中完全可察觉的频谱分量, 而丢弃其他流。

[0118] 在等式 (3) 至 (5) 的框架中要考虑的能量值可以例如通过计算相应强度值的平方而从图 6 所示强度值导出。在与频谱分量有关的信息可以包括其他值的情况下, 可以根据帧 510 中包括的信息的形式来执行类似计算。例如, 在复值信息的情况下, 可能必须执行: 计算组成与频谱分量有关的信息的各个值的实部和虚部的模。

[0119] 除了各个频率之外, 针对根据等式 (3) 至 (5) 的心理声学模型的应用, 等式 (3) 和 (4) 中的和值可以包括多于一个频率。换言之, 在等式 (3) 和 (4) 中, 可以使用与多个单独频率相对应的总能量值 (频带的能量) 来代替相应能量值 E_n , 或者更一般而言, 可以使用与

一个或多个频谱分量有关的单一频谱信息或多个频谱信息来代替相应能量值 E_n 。

[0120] 例如,由于 AAC-ELD 以逐频带方式对频谱线进行操作,与人类听觉系统同时处理的频率组相似,可以以类似方式执行不相关性估计或心理声学模型。通过以这种方式来应用心理声学模型,可以在必要时仅去除或取代单一频带的信号的部分。

[0121] 如心理声学实验已经表明,信号被另一信号掩蔽取决于相应信号类型。可以应用最坏情况场景作为不相关性确定的最小阈值。例如,

[0122] 为了使用正弦或另一独特而良好定义的声音来掩蔽噪声,典型地需要 21 至 28dB 的差值。测试已经表明,近似 28.5dB 的阈值得到良好的替代结果。还考虑所考虑的实际频带,可以最终改进该值。

[0123] 因此,在心理声学评估或者基于所考虑的频谱分量的不相关性评估方面,根据等式 (5) 大于 -28.5dB 的值 $r(n)$ 可以被认为不相关。对于不同的频谱分量,可以使用不同的值。因此,对于所考虑的帧,使用 10dB 至 40dB、20dB 至 30dB、或 25dB 至 30dB 的阈值作为输入数据流的心理声学不相关性的指示符可以被认为是有用的。

[0124] 在图 7 所示的情况下,这意味着关于频谱分量 560,确定第一输入数据流 510-1,而关于频谱分量 560 丢弃第二输入数据流 510-2。因此,至少部分将与频谱分量 560 相关的信息段从第一输入数据流 510-1 的帧 540-1 拷贝到输出数据流 530 的输出帧 550。这如图 7 中的箭头 570 所示。同时,如虚线 580 所示,省略与其他输入数据流 510 的帧 540 (即,在图 7 中输入数据流 510-2 的帧 540-2) 的频谱分量 560 相关的信息。

[0125] 换言之,例如可以用作 MCU 或会议系统 100 的设备 500 适于,与其输出帧 550 一起产生输出数据流 530,使得仅从确定的输入数据流 510-1 的帧 540-1 拷贝对输出数据流 530 的输出帧 550 的频谱分量 560 加以描述的对应该频谱分量的信息。自然地,设备 500 还可以适于,从输入数据流拷贝与多于一个频谱分量相关的信息,省略至少关于这些频谱分量的其他输入数据流。此外,设备 500 或其处理单元 520 适于,使得对于不同的频谱分量,确定不同的输入数据流 510。输出数据流 530 的相同输出帧 550 可以包括与来自不同输入数据流 510 的不同频谱分量相关的拷贝频谱信息。自然,实现设备,使得在输入数据流 510 中的帧序列 540 的情况下,在比较和确定期间仅考虑与相似或相同的时间索引相对应的帧 540 是可取的。

[0126] 换言之,图 7 示出了根据实施例的用于对如上所述多个输入数据流进行混合的设备的操作原理。如上所述,在对所有输入流进行解码的情况下,以直接方式进行混合,这包括到时域的逆变换、混合以及再次对信号进行重新编码。

[0127] 图 6 至 8 的实施例基于在相应编解码器的频域中进行的混合。可能的编解码器应当是 AAC-ELD 编解码器,或者具有均匀变换窗的任何其他编解码器。在这样的情况下,不需要时间 / 频率变换以能够混合相应数据。根据本发明实施例的实施例利用以下事实:能够访问所有比特流参数,例如量化步长和其他参数,以及可以使用这些参数来产生混合的输出比特流。

[0128] 图 6 至 8 的实施例利用以下事实:可以通过源频谱线或频谱信息的加权求和来执行与频谱分量相关的频谱线或频谱信息的混合。加权因子可以是 0 或 1,或者原则上,可以是 0 和 1 之间的任何值。0 值意味着将源视为无关并且根本不使用源。诸如频带或缩放因子频带之类的线组可以使用相同加权因子。然而,如上所述,加权因子 (例如,0 和 1 的分

布) 可以针对单个输入数据流 510 的单个帧 540 的频谱分量而变化。此外, 在混合频谱信息时不必专门使用加权因子 0 或 1。可以是以下情况: 不针对输入数据流 510 的帧 540 的总体频谱信息中的单一一个频谱信息, 而是针对多个频谱信息, 相应加权因子可以不同于 0 或 1。

[0129] 一种具体情况是, 将一个源(输入数据流 510)的所有频带或频谱分量设置为因子 1, 将其他源的所有因子设置为 0。在这种情况下, 一个参与者的完整输入比特流相同地被拷贝为最后混合比特流。可以逐帧地计算加权因子, 但是也可以基于较长帧组或帧序列来计算。自然, 甚至在这样的帧序列内或者单个帧内, 如上所述, 加权因子也可以针对不同频谱分量而不同。可以根据心理声学模型的结果来计算或确定加权因子。

[0130] 已经利用等式 (3)、(4) 和 (5) 在上文中描述了心理声学模型的示例。心理声学模型或相应模型计算仅包括一些输入流来产生能量值 E_f 的混合信号与具有能量值 E_c 的完整混合信号之间的能量比值 $r(n)$ 。然后, 将能量比值 $r(n)$ 计算为 E_f 除以 E_c 的对数的 20 倍。

[0131] 如果该比值足够高, 则可以认为主要声道掩蔽了次要声道。因此, 处理不相关性减小, 意味着仅包括完全不可察觉的、具有加权因子 1 的流, 而所有其他流(一个频谱分量的至少一个频谱信息)被丢弃。换言之, 使这些流具有加权因子 0。

[0132] 可以得到以下优点: 由于重新量化步长的数目减少, 串联编码的效应较少出现或不出现。由于每个量化步长表现出减小附加量化噪声的显著风险, 因此可以通过采用用于对多个输入数据流进行混合的设备的形式的根据本发明的实施例来提高音频信号的总体质量。这可以是以下情况, 当如图 6 所示的设备 500 的处理单元 520 适于产生输出数据流 530, 以便维持与确定的输入流或输入流的一部分的帧的量化级分布可比的量化等分布。换言之, 通过拷贝并因此通过重新使用相应数据, 而不对频谱信息进行重新编码, 可以忽略附加量化噪声的引入。

[0133] 此外, 会议系统, 例如具有多于两个参与者采用以上关于图 6 至 8 描述的任何实施例的电视/视频会议系统, 可以提供与时域混合相比复杂度较低的优点, 这是由于可以省略时间-频率变换步骤和重新编码步骤。此外, 与时域中进行混合相比, 由于不存在滤波器组延迟, 这些组件未造成另外的延迟。

[0134] 总之, 例如, 上述实施例可以适于, 使得不对与完全从一个源获取的频谱分量相对应的频带或频谱信息进行重新量化。因此, 仅对被混合的频带或频谱信息进行重新量化, 这减少了附加量化噪声。因此, 上述实施例也可以在不同应用中使用, 例如感知噪声替换(PNS)、时间噪声成形(TNS)、频谱带复制(SBR)、和立体声编码模式。在描述能够处理 PNS 参数、TNS 参数、SBR 参数或立体声编码参数中的至少一个的设备的操作之前, 将参照图 8 对这种实施例进行更详细描述。

[0135] 图 8 示出了用于对多个输入数据流进行混合的设备 500 的示意框图, 设备 500 包括处理单元 520。更准确地, 图 8 示出了一种高灵活性的设备 500, 能够处理输入数据流(比特流)中编码的差异较大的音频信号。因此, 以下将描述的一些组件是可选组件, 不需要在所有情况下都实现。

[0136] 针对处理单元 520 要处理的每个输入数据流或编码的音频比特流, 处理单元 520 包括比特流解码器 700。仅为了简单起见, 图 8 仅示出了两个比特流解码器 700-1、700-2。自然, 根据要处理的输入数据流的数目, 可以实现更多数目或更少数目目的比特流解码器

700(如果例如比特流解码器 700 能够顺序处理多于一个输入数据流)。

[0137] 比特流解码器 700-1 以及其他比特流解码器 700-2,... 均包括比特流读取器 710, 比特流读取器 710 适于接收信号并处理接收的信号, 以及隔离和提取比特流中包括的数据。例如, 比特流读取器 710 可以适于将输入数据与内部时钟同步, 还可以适于将输入比特流分离为合适的帧。

[0138] 比特流解码器 700 还包括: Huffman 解码器 720, 耦合至比特流读取器 710 的输出以从比特流读取器 710 接收隔离的数据。Huffman 解码器 720 的输出耦合至解量化器 730(也称为反量化器)。耦合在 Huffman 解码器 720 之后的解量化器 730 后接缩放器 740。Huffman 解码器 720、解量化器 730 和缩放器 740 形成第一单元 750, 在第一单元 750 的输出处, 相应输入数据流的音频信号的至少一部分在参与者的编码器(图 8 中未示出)所操作的频域或频率相关域中可用。

[0139] 比特流解码器 700 还包括: 第二单元 760, 按数据耦合在第一单元 750 之后。第二单元 760 包括: 立体声解码器 770(M/S 模块), 其后耦合有 PNS 解码器。PNS 解码器 780 按数据后接 TNS 解码器 790, TNS 解码器 790 与 PNS 解码器 780 一起在立体声解码器 770 处形成第二单元 760。

[0140] 除了所描述的音频数据的流程之外, 比特流解码器 700 还包括与控制数据有关的不同模块之间的多个连接。更准确地, 比特流读取器 710 也耦合至 Huffman 解码器 720 以接收合适的控制数据。此外, Huffman 解码器 720 直接耦合至缩放器 740 以向缩放器 740 发送缩放信息。立体声解码器 770、PNS 解码器 780 和 TNS 解码器 790 也均耦合至比特流读取器 710 以接收合适的控制数据。

[0141] 处理单元 520 还包括混合单元 800, 混合单元 800 继而包括频谱混合器 810, 频谱混合器 810 按输入耦合至比特流解码器 700。频谱混合器 810 可以例如包括一个或多个加法器, 以在频域中执行实际混合。此外, 频谱混合器 810 还可以包括乘法器, 以允许比特流解码器 700 提供的频谱信息的任意线性组合。

[0142] 混合单元 800 还包括: 优化模块 820, 按数据耦合至频谱混合器 810 的输出。然而, 优化模块 820 还耦合至频谱混合器 810 以向频谱混合器 810 提供控制信息。优化模块 820 按数据表示混合单元 800 的输出。

[0143] 混合单元 800 还包括: SBR 混合器 830, 直接耦合至不同比特流解码器 700 的比特流读取器 710 的输出。SBR 混合器 830 的输出形成混合单元 800 的另一输出。

[0144] 处理单元 520 还包括: 比特流编码器 850, 耦合至混合单元 800。比特流编码器 850 包括第三单元 860, 第三单元 860 包括 TNS 编码器 870、PNS 编码器 880 和立体声编码器 890(以所描述的顺序串联耦合)。因此, 第三单元 860 形成比特流解码器 700 的第一单元 750 的反单元。

[0145] 比特流编码器 850 还包括第四单元 900, 第四单元 900 包括缩放器 910、量化器 920 和 Huffman 编码器 930(形成第四单元的输入与其输出之间的串联连接)。因此, 第四单元 900 形成第一单元 750 的反模块。相应地, 缩放器 910 直接耦合至 Huffman 编码器 930, 以向 Huffman 编码器 930 提供相应控制数据。

[0146] 比特流编码器 850 还包括: 比特流写入器 940, 耦合至 Huffman 编码器 930 的输出。此外, 比特流写入器 940 还耦合至 TNS 编码器 870、PNS 编码器 880、立体声编码器 890

和 Huffman 编码器 930, 以从这些模块接收控制数据和信息。比特流写入器 940 的输出形成处理单元 520 的输出和设备 500 的输出。

[0147] 比特流编码器 850 还包括: 心理声学模块 950, 也耦合至混合单元 800 的输出。比特流编码器 850 适于向第三单元 860 的模块提供合适的控制信息, 例如指示在第三单元 860 的单元的框架中, 哪个单元可以用于对混合单元 800 输出的音频信号进行编码。

[0148] 因此, 原则上, 在第二单元 760 的输出直至第三单元 860 的输入处, 如发送方侧使用的编码器所定义的, 可以在频谱域中处理音频信号。然而, 如上所述, 如果例如输入数据流之一的帧的频谱信息是主要的, 则最终可以不需要完整的解码、解量化、解缩放和另外的处理步骤。然后将相应频谱分量的频谱信息的至少一部分拷贝至输出数据流的相应帧的频谱分量。

[0149] 为了允许这种处理, 设备 500 和处理单元 520 包括另外的信号线来进行优化的数据交换。为了在图 8 所示的实施例中允许这种处理, Huffman 解码器 720 的输出, 以及缩放器 740、立体声解码器 770 和 PNS 解码器 780 的输出, 与其他比特流读取器 710 的相应组件一起, 耦合至混合单元 800 的优化模块 820, 以进行相应处理。

[0150] 在相应处理之后, 为了便于比特流编码器 850 内的相应数据流, 还实现了针对优化数据流的对应数据线。更准确地, 优化模块 820 的输出耦合至 PNS 编码器 780 的输入, 立体声编码器 890、第四单元 900 的输入和缩放器 910, 以及 Huffman 编码器 930 的输入。此外, 优化模块 820 的输出还直接耦合至比特流写入器 940。

[0151] 如上所述, 几乎所有上述模块都是可选模块, 不必实现这些可选模块。例如, 在音频数据流仅包括单一声道的情况下, 可以省去立体声编码和解码单元 770、890。相应地, 在没有基于 PNS 的信号要处理的情况下, 也可以省去对应的 PNS 解码器和 PNS 编码器 780、880。在要处理的信号和要输出的信号不基于 TNS 数据的情况下, 还可以省去 TNS 模块 790、870。在第一和第四单元 750、900 内, 还可以最终省去反量化器 730、缩放器 740、量化器 920 以及缩放器 910。Huffman 解码器 720 和 Huffman 编码器 930 可以以不同方式实现, 使用不同算法, 或者完全省略。

[0152] 例如, 如果不存在数据的 SBR 参数, 则最终也可以省略 SBR 混合器 830。此外, 可以以不同方式实现频谱混合器 810, 以与优化模块 820 和心理声学模块 860 协作。因此, 认为这些模块也是可选组件。

[0153] 对于设备 500 以及其中包括的处理单元 520 的操作模式, 比特流读取器 710 首先读取输入的输入数据流并将其分离为合适的信息。在 Huffman 解码之后, 最终, 得到的频谱的信息可以由解量化器 730 重新量化, 并由解缩放器 740 进行合适缩放。

[0154] 此后, 根据输入数据流中包括的控制信息, 在立体声解码器 770 的框架中, 可以将输入数据流中编码的音频信号分解为两个或多个声道的音频信号。如果例如音频信号包括中声道 (M) 和侧声道 (S), 则通过将中声道和侧声道数据相加或相减, 可以获得对应的左声道和右声道数据。在许多实现中, 中声道与左声道和右声道音频数据之和成比例, 而侧声道与左声道 (L) 与右声道 (R) 之差成比例。根据实现方式, 可以考虑因子 1/2 来对上述声道进行相加和 / 或相减, 以防止限幅效应。一般而言, 线性组合可以处理不同声道以产生对应声道。

[0155] 换言之, 在立体声解码器 770 之后, 如果合适, 可以将音频数据分解为两个单独的

声道。自然,立体声解码器 770 还可以执行反解码。如果例如比特流读取器 710 接收的音频信号包括左和右声道,则立体声解码器 770 同样可以计算或确定合适的中声道和侧声道数据。

[0156] 不仅根据设备 500 的实现,还根据提供相应输入数据流的参与者的编码器的实现,相应数据流可以包括 PNS 参数 (PNS = 感知噪声替换)。PNS 基于以下事实:在有限的频率范围或频谱分量(如频带或单独的频率)中,人耳很可能无法将类似噪声的声音与合成产生的噪声区分开。因此,PNS 将音频信号中实际类似噪声的成分替换为指示要合成引入相应频谱分量的噪声电平并忽略实际音频信号的能量值。换言之,PNS 解码器 780 可以在一个或多个频谱分量中,基于输入数据流中包括的 PNS 参数,来产生实际类似噪声的音频信号成分。

[0157] 对于 TNS 解码器 790 和 TNS 编码器 870,可能必须将相应音频信号变换回相对于在发送方侧操作的 TNS 模块而言未修改的版本。时间噪声成形 (TNS) 是用于减小量化噪声导致的预回声伪像的手段,该伪像可能存在于音频信号帧中类似瞬变的信号的情况下。为了对抗这种瞬变,从频谱的低侧、频谱的高侧或者频谱的两侧开始,对频谱信息应用至少一个自适应预测滤波器。可以对预测滤波器的长度以及应用相应滤波器的频率范围进行适配。

[0158] 换言之,TNS 模块的操作基于计算一个或多个自适应 IIR 滤波器 (IIR = 无限冲激响应),并通过编码和发送对预测和实际音频信号之间的差值进行描述的误差信号以及预测滤波器的滤波器系数来进行。因此,可以提高音频质量,同时通过在频域中应用预测滤波器来处理类似瞬变的信号,以减小其余误差信号的幅度(然后,可以使用与以类似的量化噪声对类似瞬变的音频信号进行直接编码相比较少的量化步长来对其余误差信号进行编码),从而维持发送方数据流的比特率。

[0159] 对于 TNS 应用,在一些情况下采用 TNS 解码器 760 的功能来对输入数据流的 TNS 部分进行解码,以达到所使用的编解码器所确定的、频谱域中的“纯”表示是可取的。如果不能基于 TNS 参数中包括的预测滤波器的滤波器系数已经估计出心理声学模型(例如应用于心理声学模块 950 中的心理声学模型)的估计,则应用 TNS 解码器 790 的功能是有用的。在至少一个输入数据流使用 TNS 而另一个不使用 TNS 的情况下,这尤其重要。

[0160] 当处理单元基于输入数据流的帧的比较确定要使用来自使用 TNS 的输入数据流的帧的频谱信息时,TNS 参数可以用于输出数据的帧。如果例如出于不兼容性的原因,输出数据流的接收者不能解码 TNS 数据,则不拷贝误差信号的相应频谱数据以及和另外的 TNS 参数,而根据 TNS 相关数据来处理重构数据以获得频谱域中的信息,并且不使用 TNS 编码器 870 可能是有用的。这再次示意了,不需要实现图 8 中所示的组件或模块的部分,但可选地可以保留。

[0161] 在对 PNS 数据进行比较的至少一个音频输入流的情况下,可以应用类似的策略。如果在针对输入数据流的频谱分量对帧进行的比较中表明一个输入数据流在其当前帧和相应频谱分量方面是主导的,则也可以将相应 PNS 参数(即相应能量值)直接拷贝至输出帧的相应频谱分量。然而,如果接收者不能接受 PNS 参数,则可以通过使用相应能量值指示的合适的能量等级来产生噪声,从而根据相应频谱分量的 PNS 参数来重构频谱信息。然后,可以在频谱域中相应地处理噪声数据。

[0162] 如上所述,发送的数据还包括 SBR 数据,可以在 SBR 混合器 830 中处理该 SBR 数

据。频谱带复制 (SBR) 是一种基于相同频谱的贡献和较低部分来复制音频信号的频谱的一部分的技术。因此不需要传输频谱的较高部分,除了通过采用适当的时间 / 频率栅格以频率相关和时间相关方式描述能量值的 SBR 参数以外。因此,根本不需要传输频谱的较高部分。为了能够进一步提高重构信号的质量,可以在频谱的较高部分中添加附加的噪声贡献和正弦波贡献。

[0163] 更具体地,对于交叉频率 f_x 之上的频率,根据创建了特定数目的子带信号 (例如, 32 个子带信号) 的 QMF 滤波器组 (QMF = 正交镜像滤波器) 来分析音频信号,这些子带信号具有通过等于或与 QMF 滤波器组子带数目 (例如, 32 或 64) 成比例的因子所降低的时间分辨率。因此,可以确定时间 / 频率栅格,时间 / 频率栅格在时间轴上包括两个或多个所谓的包封,并且对于每个包封,包括描述频谱的相应较高部分的 7 到 16 个能量值。

[0164] 此外, SBR 参数可以包括与附加噪声和正弦波相关的信息,然后这些附加噪声和正弦波通过上述时间频率栅格,在强度方面被衰减或确定。

[0165] 在基于 SBR 的输入数据流相对于当前帧是主要输入数据流的情况下,可以执行将相应 SBR 参数连同频谱分量一同拷贝。再次,如果接收方不能对基于 SBR 的信号进行解码,则可以执行到频域中的相应重构,然后根据接收方的需要对重构信号进行编码。由于 SBR 允许两个编码立体声声道,对左声道和右声道单独编码,以及在耦合声道 (C) 方面对左声道和右声道进行编码,因此,根据本发明的实施例,对相应 SBR 参数或至少其部分进行拷贝可以包括根据比较结果和确定结果,将 SBR 参数的 C 元素拷贝至要确定和发送的 SBR 参数的左和右元素,或者反之。

[0166] 此外,由于在本发明的不同实施例中,输入数据流可以分别包括单声道以及包括一个和两个单独声道的立体声音频信号,因此,在产生输出数据流的帧的相对频谱分量的信息的一部分时,可以附加地执行单声道至立体声上混或立体声至单声道下混。

[0167] 如之前的描述所示,对与频谱分量和频谱信息相关的频谱信息和 / 或相应参数 (例如, TNS 参数、SBR 参数、PNS 参数) 进行拷贝的程度可以基于不同数目的要拷贝数据,并且可以确定是否还需要拷贝基本频谱信息或基础频谱信息中的频谱信息。例如,以下操作是可取的:在拷贝 SBR 数据的情况下,拷贝相应数据流的整个帧,以防止针对不同频谱分量的频谱信息的复杂混合。混合这些信息需要实际上可以减少量化噪声的重新量化。

[0168] 在 TNS 参数方面,拷贝相应 TNS 参数以及从主要输入数据流至输出数据流的整个帧的频谱信息,以防止重新量化是可取的。

[0169] 在基于 PNS 的频谱信息的情况下,拷贝各个能量值而不对作为基础频谱分量进行拷贝是可行的方式。此外,在这种情况下,通过仅将相应 PNS 参数从多个输入数据流的帧的主要频谱分量拷贝至输出数据流的输出帧的对应频谱分量,不引入附加的量化噪声。应当注意,同样通过对 PNS 参数形式的能量值的重新量化,不会引入附加量化噪声。

[0170] 如上所述,上述实施例还可以通过以下来实现:在比较多个输入数据流的帧之后,以及在基于所述比较,针对输出数据流的输出帧的频谱分量,确定正好一个数据流为频谱信息的源之后,简单地拷贝与频谱分量相关的频谱信息。

[0171] 在心理声学模块 950 的框架中执行的替换算法检查与生成的信号的基础频谱分量 (例如频带) 相关的每个频谱信息,以标识仅具有单一有效分量的频谱分量。对于这些频带,可以从编码器拷贝输入比特流的相应输入数据流的量化值,而不对指定频谱分量的

相应频谱数据进行重新编码或重新量化。在一些情况下,所有量化数据可以从单一有效输入信号取得,以形成输出比特流或者输出数据流,使得对于设备 500 而言能够实现输入数据流的无损编码。

[0172] 此外,可以省略编码器内如心理声学分析之类的处理步骤。这允许缩短编码过程,从而降低计算复杂度,因为原则上仅将数据从一个比特流拷贝入另一比特流必须在特定情况下执行。

[0173] 例如,在 PNS 的情况下,可以执行替换,因为 PNS 编码的频带的噪声因子可以从输出数据流之一拷贝至输出数据流。可以使用合适的 PNS 参数来替换各个频谱分量,因为 PNS 参数是针对频谱分量指定的,或者换言之,是互相独立的非常好的近似。

[0174] 然而,可能出现以下情况:所描述的算法的两个有力应用可能得到退化的收听体验或者不利的质量降低。因此,将替换限制在各个帧,而不是与各个频谱分量相关的频谱信息,是可取的。在这种操作模式中,可以不变地执行不相关性估计或不相关性确定以及替换分析。然而,在这种操作模式中,仅当有效帧内的所有或至少大量频谱分量可替换时,才执行替换。

[0175] 尽管这可能导致较少次数的替换,在一些情况下,可以改进频谱信息的内在强度,得到甚至略微改进的质量。

[0176] 在下文中,根据本发明第二方面的实施例可以根据以下来描述:考虑那些与相应输入数据流的有效载荷数据相关联的控制值,控制值指示有效载荷数据表示相应音频信号的相对频谱信息或频谱域的至少一部分的方式,其中,在两个输入数据流的控制值相等的情况下,可以避免对输出数据流的相应帧处的频谱域的方式的新判定,取而代之,输出流产生依赖于输入数据流的编码器已经确定的判定。根据以下描述的一些实施例,可以避免将相应有效载荷数据重新变换回到表示频谱域的另一种方式(例如,具有每时间/频谱采样一个频谱值的正常或普通方式)。

[0177] 如上所述,根据本发明的实施例基于在所有输入流被解码的意义上以直接的方式进行的混合,这包括反变换至时域、混合和再次对信号重新编码。根据本发明的实施例基于在相应编解码器的频域中进行的混合。一种可能的编解码器可以是 AAC-ELD 编解码器,或者具有均匀变换窗的任何其他编解码器。在这种情况下,不需要时间/频率变换以能够对相应数据进行混合。此外,可以访问所有比特流参数,如量化步长和其他参数,可以使用这些参数来产生混合的输出比特流。

[0178] 此外,与频谱分量相关的频谱线或频谱信息的混合可以通过源频谱线或频谱信息的加权和来执行。加权因子可以是 0 或 1,或者原则上可以是 0 和 1 之间的任何值。0 值意味着将源视为不相关并且将完全不使用。线组,如频带或缩放因子频带,可以使用相同的加权因子。针对单一输入数据流的单一帧的频谱分量,加权因子(例如 0 和 1 的分布)可以改变。此外,以下描述的实施例在混合频谱信息时,不需要排他地使用 0 或 1 的加权因子。可以有如下情况:在一些情况下,不是针对单一情况,而是针对输入数据流的帧的多个总体频谱信息,相应加权因子可以不同于 0 或 1。一种特殊情况在于,一个源(输入数据流)的所有频带或频谱分量被设置为因子 1,而其他源的所有因子被设置为 0。在这种情况下,一个参与者的完整输入比特流被相同地拷贝为最终混合比特流。可以逐帧计算加权因子,但是可以基于更长的帧组或序列来计算或确定加权因子。自然,如上所述,即使在这种帧序列

之内或者在单一帧之内,加权因子也可以针对不同频谱分量而不同。在一些实施例中,可以根据心理声学模型的结果来计算或确定加权因子。

[0179] 例如,这样的比较可以基于其中仅包括一些输入流的混合信号与完整混合信号之间的能量比值的评估。例如,这可以如关于等式(3)至(5)所描述的来实现。换言之,心理声学模型可以计算仅包括一些输入流以得到能量值 E_f 的混合信号与具有能量值 E_c 的完整混合信号之间的能量比值 $r(n)$ 。然后,将能量比值 $r(n)$ 计算为 E_f 除以 E_c 的对数的20倍。

[0180] 相应地,类似于关于图6至8的上述实施例描述,如果该比值足够高,则可以认为主要声道掩蔽了次要声道。因此,处理不相关性减小,意味着仅包括完全不可察觉的、具有加权因子1的流,而所有其他流(一个频谱分量的至少一个频谱信息)被丢弃。换言之,使这些流具有加权因子0。

[0181] 可以得到以下优点:由于重新量化步长的数目减少,串联编码的效应较少出现或不出现。由于每个量化步长表现出减小附加量化噪声的显著风险,因此可以提高音频信号的总体质量。

[0182] 类似于图6至8的上述实施例,以下描述的实施例可以与会议系统(例如具有多于两个参与者的电话/视频会议系统)一同使用,并且与时域混合相比,这种会议系统可以提供较低复杂度的优点,因为可以省略时间-视频变换步骤和重新编码步骤。此外,与在时域中进行混合相比,由于不存在滤波器组延迟,这些组件未造成另外的延迟。

[0183] 图9示出了根据本发明实施例的用于对输入数据流进行混合的设备500的简要框图。采用了图6至8的实施例中的多数附图标记,以便容易理解并避免重复描述。其他附图标记大于1000,以便表示与以上图6至8的实施例相比,以不同方式定义的相同功能,附加功能或备选功能,但是具有可比较的相应元件的一般功能。

[0184] 基于第一输入数据流510-1、和第二输入数据流510-2,包括在设备1500中的处理单元1520适于产生输出数据流1530。第一和第二输入数据流510分别包括帧541-1、540-2,相应地帧541-1、541-2分别包括控制值1541-1、1541-2,控制值1541-1、1541-2分别指示帧540的有效载荷数据表示音频信号的频谱域或频谱信息的至少一部分的方式。

[0185] 输出数据流530还包括具有控制值555的输出帧1550,控制值555以类似的方式指示输出帧550的有效载荷数据表示在输出数据流530中编码的音频信号的频谱域中的频谱信息的方式。

[0186] 设备1500的处理器单元1520适于,将第一输入数据流510-1的帧540-1的控制值1541-1与第二输入数据流510-2的帧540-2的控制值1542-2进行比较,以获得比较结果。基于该比较结果,处理器单元1520还适于产生包括输出帧550在内的输出数据流530,使得当比较结果指示第一和第二输入数据510的帧540的控制值1541等同或相等时,输出帧550包括等于两个输入数据流510的帧540的控制值1545的值作为控制值1550。通过在频谱域中进行的处理(即,不访问时域),针对帧540的相同控制值1545,从帧540的对应有效载荷数据中导出输出帧550中包括的有效载荷数据。

[0187] 例如,如果控制值1545指示一个或多个频谱分量(例如,PNS数据)的频谱信息的专用编码,并且两个输入数据流的相应控制值1545相同,则可以通过直接对频谱域中的对应有效载荷数据进行处理(即,不离开频谱域的代表类型),来获得输出帧550中与相同频谱分量相对应的对应频谱信息。如下所述,在基于PNS频谱表示的情况下,这可以通过以

下来实现：对相应 PNS 数据进行求和，可选地，通过归一化处理来完成。即，均不将两个输入数据流的 PNS 数据转换回到具有每频谱采样一个值的普通表示中。

[0188] 图 10 示出了主要针对处理单元 1520 的内部结构不同于图 9 的设备 1500 的更详细框图。为了更具体，处理单元 1520 包括比较器 1560，比较器 1560 耦合至针对第一和第二输入数据流 510 的适当输入，并且适于将它们的相关帧 540 的控制值 1545 进行比较。此外，将输入数据流提供至针对两个输入数据流 510 中的每一个的可选变换器 1570-1、1570-2。比较器 1560 还耦合至可选变换器 1570，以向可选变换器 1570 提供比较结果。

[0189] 处理单元 1520 还包括混合器 1580，混合器 1580 按输入耦合至可选变换器 1570，或者在没有实现一个或多个变换器 1570 的情况下，耦合至输入数据流 510 的对应输入。混合器 1580 的输出耦合至可选归一化器 1590，如果实现了归一化器 1590，则归一化器 1590 与处理器单元 1520 的输出和设备 1500 的输出耦合，以提供输出数据流 530。

[0190] 如上所述，比较器 1560 适于将两个输入数据流 510 的帧 540 的控制值进行比较。比较器 1560 向变换器 1570（如果实现）提供对相应帧 540 的控制值 1545 是否相同加以指示的信号。如果表示比较结果的信息指示两个控制值 1545 至少相对于一个频谱分量是相同或相等的，则变换器 1570 不对帧 540 中所包括的相应有效载荷数据进行变换。

[0191] 然后包括在输入数据流 510 的帧 540 中的有效载荷数据由混合器 1580 进行混合，并且向归一化器 1590（如果实现）输出，以执行归一化步骤，从而确保生成的值高于或低于允许的值范围。以下在图 12a 至 12c 的上下文中更详细描述对有效载荷数据进行混合的示例。

[0192] 可以将归一化器 1590 实现为适于根据有效载荷数据的相应值分别对有效载荷数据进行重新量化的量化器，备选地，归一化器 1590 还适于根据其具体实现仅改变指示量化步骤的分布缩放因子，或最小或最大量化级的绝对值。

[0193] 在比较器 1560 指示控制值 1545 至少相对于一个或多个频谱分量是不同的情况下，比较器 1560 可以向变换器 1570 中的一个或两个变换器 1570 提供相应控制信号，该相应控制信号指示相应变换器 1570 将输入数据流 510 中的至少一个的有效载荷数据转换成其他输入数据流的有效载荷数据。在这种情况下，变换器可以适于同时改变变换后的帧的控制值，使得混合器 1580 能够产生输出数据流 530 的输出帧 550，该输出帧 550 具有等于两个输入数据流中没有变换的帧 540 的控制值的控制值 1555，或者具有两个帧 540 的有效载荷数据的公共值。

[0194] 以下在分别针对不同应用（例如，PNS 实现、SBR 实现和 M/S 实现）的图 12a 至 12c 的上下文中描述更详细示例。

[0195] 应当指出，图 9 至图 12C 的实施例不限于图 9、10 和即将描述的图 11 中所示的两个输入数据流 1510-1、1510-2。而是，这些实施例可以适于对包括多于两个输入数据流 510 的多个输入数据流进行处理。在这种情况下，例如，比较器 1560 可以适于将适当数目的输入数据流 510 进行比较，以及将输入数据流 510 中包括的帧 540 进行比较。此外，根据具体实现，也可以实现适当数目的变换器 1570。混合器 1580 连同可选归一化器 1590 一起可以最终适于增加数目的要处理数据流。

[0196] 在仅多于两个输入数据流 510 的情况下，比较器 1560 可以适于将输入数据流 510 的所有相关控制值进行比较，以判断是否通过一个或多个可选实现的变换器 1570 来执行

变换步骤。备选或附加地,比较器 1560 还可以适于,在比较结果指示至有效载荷数据的表示的通常方式的变换是可实现的情况下,确定要由变换器 1570 变换的输入数据流集合。例如,除非所涉及的有效载荷数据的不同表示需要特定表示,否则比较器例如可以适于,以使总复杂度最小化的方式来激活变换器 1570。例如,这可以基于存储在比较器 1560 内或者以不同方式可用于比较器 1560 的复杂度值的预定估计。

[0197] 此外,应当注意,例如当可选地根据需要通过混合器 1580 执行到频域的变换时,变换器 1570 最终是可省略的。备选或附加地,变换器 1570 的功能也可以合并到混合器 1580 中。

[0198] 此外,应当注意,帧 540 可以包括多于一个控制值,例如,感知噪声替换 (PNS)、时间噪声成形 (TNS)、和立体声编码模式。在描述能够处理 PNS 参数、TNS 参数或立体声编码参数中的至少一个的设备的操作之前,参照图 11,图 11 与图 8 相同,但是其中用附图标记 1500 和 1520 分别代替 500 和 520,以便示出图 8 已经示出了用于根据第一和第二输入数据流产生输出数据流的实施例,其中,处理单元 520 和 1520 还可以分别适于执行关于图 9 和 10 描述的功能。具体地,在处理单元 1520 内,包括频谱混合器 810、优化模块 820 和 SBR 混合器 830 的混合单元 800 执行关于图 9 和 10 提出的上述功能。如上所述,包括在输入数据流的帧中的控制值同样可以是 PNS 参数、SBR 参数或与立体声编码相关的控制数据,换言之,M/S 参数。在相应控制值相等或等同的情况下,混合单元 800 可以对有效载荷数据进行处理,来产生要进一步处理以包括在输出数据流的输出帧中的对应有效载荷数据。关于这一点,如上所述,由于 SBR 允许针对两个编码立体声声道,来分别对左声道和右声道进行编码,以及在耦合声道 (C) 方面对左声道和右声道进行编码,因此,根据本发明的实施例,对相应 SBR 参数或至少其部分进行拷贝可以包括根据比较结果和确定结果,将 SBR 参数的 C 元素拷贝至要确定和发送的 SBR 参数的左和右元素,或者反之。类似的,对与频谱分量相关的频谱信息和 / 或相应参数 (例如,TNS 参数、SBR 参数、PNS 参数) 进行处理的程度可以基于不同数目的要处理数据,并且可以确定是否还需要对基本频谱信息或基础频谱信息中的频谱信息进行解码。例如,在拷贝 SBR 数据的情况下,对相应数据流的整个帧进行处理,以防止针对不同频谱分量的频谱信息的复杂混合是可取的。混合这些信息需要实际上可以减少量化噪声的重新量化。在 TNS 参数方面,将相应 TNS 参数以及整个帧的频谱信息从主要输入数据流分解至输出数据流,以防止重新量化是可取的。在基于 PNS 的频谱信息的情况下,对各个能量值进行处理而不对基础频谱分量进行拷贝是可行的方式。此外,在这种情况下,通过仅处理从多个输入数据流的帧的主要频谱分量至输出数据流的输出帧的对应频谱分量的相应 PNS 参数,不引入附加的量化噪声。应当注意,同样通过对 PNS 参数形式的能量值的重新量化,不会引入附加量化噪声。

[0199] 关于图 12A 至 12C,将更详细描述基于相应控制值的比较对有效载荷数据进行混合的三种不同模式。图 12a 示出了根据本发明实施例的设备 500 的基于 PNS 的实现的示例,而图 12b 示出了设备 500 的类似 SBR 实现,图 12c 示出了设备 500 的 M/S 实现。

[0200] 图 12a 示出了具有第一和第二输入数据流 510-1、510-2 的示例,第一和第二输入数据流 510-1、510-2 分别具有适当输入帧 540-1、540-2 以及相应控制值 545-1、545-2。如图 11a 中的箭头所示,输入数据流 510 的帧 540 的控制值 545 指示,没有在频谱信息方面间接描述频谱分量,但是在噪声源的能量值方面 (换言之,通过适当 PNS 参数) 描述频谱分

量。更具体地,图 12a 示出了第一 PNS 参数 2000-1 和包括 PNS 参数 2000-2 的第二输入数据流 510-2 的帧 540-2。

[0201] 关于图 12a,由于假定两个输入数据流 510 的两个帧 540 的控制值 1545 指示特定频谱分量要其相应 PNS 参数 2000 来代替,如上所述,处理单元 1520 和设备 1500 能够对两个 PNS 参数 2000-1、2000-2 进行混合,以获得要包括在输出数据流 530 中的输出帧 550 的 PNS 参数 2000-3。输出帧 550 的相应控制值 1555 实质上还指示,相应频谱分量要由混合后的 PNS 参数 2000-3 来代替。通过示出 PNS 参数 2000-2,在图 12a 中示出了该混合过程,PNS 参数 2000-2 是相应帧 540-1、540-2 的组组合的 PNS 参数 2000-1、2000-2。

[0202] 然而,PNS 参数 2000-3(也被称作 PNS 输出参数)的确定也可以基于线性组合根据以下等式来实现:

$$[0203] \quad PNS = \sum_{i=1}^N a_i \cdot PNS(i) \quad (6)$$

[0204] 其中 PNS(i) 是输入数据流 i 的相应 PNS 参数,N 是要混合的输入数据流的数目, a_i 是适当的加权因子。根据具体实现,可以选择加权因子 a_i 等于:

$$[0205] \quad a_1 = \dots = a_N \quad (7)$$

[0206] 图 12a 所示的直接实现可以是在所有加权参数 a_i 等于 1 的情况下的实现,换言之,

$$[0207] \quad a_1 = \dots = a_N = 1 \quad (8)$$

[0208] 在要省略图 10 所示的归一化器 1590 的情况下,同样可以将加权因子定义为等于 $1/N$,使得以下等式成立:

$$[0209] \quad a_1 = \dots = a_N = \frac{1}{N} \quad (9)$$

[0210] 这里参数 N 是要混合的输入数据流的数目,提供至设备 1500 的输入数据流的数目是类似的数目。为了简要起见,应当注意,同样可以在加权因子 a_i 的方面实现不同的归一化。

[0211] 换言之,在激活的 PNS 工具在参与方侧的情况下,噪声能量因子代替适当的缩放因子,以及频谱分量(例如,频谱带)中的量化数据。除了该因子以外,PNS 工具不会将其其他数据提供到输出数据流中。在混合 PNS 频谱分量的情况下,会出现两种不同的情况。

[0212] 如上所述,当相关输入数据流的所有帧 540 的相应频谱分量分别按照 PNS 参数来表达时,由于频率分量(例如,频带)的 PNS 相关描述的频率数据可以直接从噪声能量因子(PNS 参数)中导出,因此可以通过简单加入相应值来混合适当因子。然后,混合后的 PNS 参数在接收方侧上的 PNS 解码器内产生要与其他频谱分量的纯频谱值混合的相同频率分辨率。在混合期间使用归一化过程的情况下,能够有助于按照加权因子 a_i 实现类似的归一化因子。例如,当利用与 $1/N$ 成比例的因子进行归一化时,可以根据等式(9)来选择加权因子 a_i 。

[0213] 在至少一个输入数据流 510 的控制值 1545 相对于频谱分量而不同的情况下,并且如果由于低能量级而不能丢弃相应输入数据流,对于图 11 所示的 PNS 解码器,以下是可取的:基于 PNS 参数产生频谱信息或频谱数据以及在混合单元的频谱混合器 810 的框架中混合相应数据,而不在优化模块 820 的框架中混合 PNS 参数。

[0214] 由于 PNS 频谱分量彼此无关,并且相对于输出数据流以及输入数据流的全局定义

参数无关,因此可以基于频带方式改变混合方法的选择。在这样的情况下,基于 PNS 的混合是不可能的,考虑在频谱域中进行混合之后通过 PNS 编码器 1880 对相应频谱分量进行重新编码是可取的。

[0215] 图 12b 示出了根据本发明实施例的实施例的操作原理的另一示例。为了更加精确,图 12b 示出了具有适当帧 540-1、540-2 及其控制值 1545-1、1545-2 的两个输入数据流 510-1、510-2 的情况。帧 540 包括针对在所谓的交叉频率 f_x 之上的频谱分量的 SBR 数据。控制值 1545 包括是否完全使用 SBR 参数的信息,以及与实际帧栅格或时间 / 频率栅格相关的信息。

[0216] 如上所述,SBR 工具通过复制以不同方式编码的频谱的较低部分,来在交叉频率 f_x 之上的较高频谱带中复制频谱的一部分。SBR 工具确定每个 SBR 帧的时隙的数目,每个 SBR 帧等同于还包括其他频谱信息在内的输入数据流 510 的帧 540。时隙以较小等间隔频带或频谱分量分开 SBR 工具的频率范围。SBR 帧中这些频带的数目可以在编码之前由发送方或 SBR 工具来确定。在 MPEG-4AAC-ELC 的情况下,时隙的数目固定为 16。

[0217] 现在,时隙包括在所谓的包封中,使得每个包封包括形成相应组的至少两个或多个时隙。每个包封归因于 SBR 频率数据的数目。在帧栅格或时间 / 频率栅格中,存储了各个单独包封的时隙的数目和单位长度。

[0218] 各个单独包封的频率分辨率确定针对包封计算并相对于包封存储多少个 SBR 能量数据。SBR 工具仅在高和低分辨率之间是不同的,其中,包括高分辨率的包封包括具有低分辨率的包封所包括的那么多值的二倍。包括高或低分辨率的包封的频率或频谱分量的数目取决于诸如比特率、采样频率等编码器其他参数。

[0219] 在 MPEG-4AAC ELC 的上下文中,SBR 工具针对具有高分辨率的包封通常使用 16 到 14 个值。

[0220] 由于相对于频率具有适当数目能量值的帧 540 的动态划分,可以考虑瞬变。在帧中存在瞬变的情况下,SBR 编码器以适当数目的包封划分相应帧。该分布在于 ACC ELD 编解码器一起使用的 SBR 工具的情况下是标准化的,并且取决于时隙单位中瞬变 transpose 的位置。在一些情况下,生成的栅格帧或时间 / 频率栅格在存在瞬变的情况下包括三个包封。第一包封,即开始包封,包括帧的开始直至接收到时隙索引从 0 至 transpose-1 的瞬变的时隙。第二包封包括两个时隙的长度,这两个时隙包围时隙索引从 transpose 至 transpose+2 的包封。第三包封包括所有后续时隙(时隙索引为 transpose+3 至 16)。

[0221] 然而,包封的长度是两个时隙。因此,包括接近帧边界的瞬变的帧可能最后仅包括两个包封。在帧中不存在瞬变的情况下,时隙在等长包封上分布。

[0222] 图 12b 示出了帧 540 中的这样的时间 / 频率栅格或帧栅格。在控制值 1545 指示在两个帧 540-1、540-2 中存在相同的 SBR 时间栅格或时间 / 频率栅格的情况下,可以类似于在以上等式 (6) 至 (9) 的上下文中描述的方法,对相应 SBR 数据进行拷贝。换言之,在这样的情况下,SBR 混合工具或如图 11 所示的 SBR 混合器 830 可以将相应输入帧的时间 / 频率栅格或帧栅格拷贝至输出帧 550,并且类似于等式 (6) 至 (9) 计算相应能量值。换言之,帧栅格的 SBR 能量数据可以通过对相应数据进行简单求和、以及可选地通过对相应数据进行归一化,来混合。

[0223] 图 12c 示出了根据本发明实施例的操作模式的另一示例。为了更加精确,图 12c

示出了 M/S 实现。图 12c 再次示出了两个输入数据流 520 和两个帧 540 以及关联的控制值 545, 关联的控制值 545 指示至少针对有效载荷数据帧 540 的至少一个频谱分量表示有效载荷数据帧 540 的方式。

[0224] 每个帧 540 包括两个声道 (第一声道 2020 和第二声道 2030) 的音频数据或频谱信息。根据相应帧 540 的控制值 1545, 例如第一声道 2020 可以是左声道或中间声道, 而第二声道 2030 可以是立体声的右声道或侧声道。编码模式中的第一编码模式通常被称作 LR 模式, 而第二模式通常被称作 M/S 模式。

[0225] 在有时被称作联合立体声的 M/S 模式中, 将中间声道 (M) 定义为与左声道 (L) 和右声道 (R) 之和成比例。通常, 在定义中包括附加因子 $1/2$, 使得中间声道在时域和频域中包括两个立体声声道的平均值。

[0226] 典型地, 将侧声道定义为与两个立体声声道的差值成比例, 即, 与左声道 (L) 和右声道 (R) 的差值成比例。有时还包括附加因子 $1/2$, 使得侧声道实质上表示立体声信号的两个声道之间的偏差值, 或者与中间声道的偏差。相应地, 可以通过对中间声道和侧声道进行求和来重构左声道, 而可以通过从中间声道中减去侧声道来获得右声道。

[0227] 在对于帧 540-1、540-2 使用相同立体声编码 (L/R 或 M/S) 的情况下, 可以省略帧所包括的声道的重新变换, 允许在相应 L/R 或 M/S 编码域中进行直接混合。

[0228] 在这种情况下, 可以直接在频域中再次执行混合, 产生包括在输出数据流 530 中具有相应控制值 1555 的帧 550, 相应控制值 1555 具有等于两个帧 540 的控制值 1545-1、1545-2 的值。输出帧 550 相应地包括从输入数据流的帧的第一和第二声道导出的两个声道 2020-3、2030-3。

[0229] 在两个帧 540 的控制值 1545-1、1545-2 不相等的情况下, 基于上述过程将帧中的一个变换成其他标识是可取的。相应地, 可以将输出帧 550 的控制值 1555 设置为指示变换后的帧的值。

[0230] 根据本发明的实施例, 控制值 1545、1555 能够分别指示整个帧 540、550 的表示, 或者相应控制值是针对频率分量而指定的。而在第一种情况下, 通过特定方法之一在整个帧上对声道 2020、2030 进行编码, 在第二种情况下, 原则上, 可以以不同方式对针对频谱分量的每个频谱信息进行编码。自然, 也可以通过控制值 1545 之一对频谱分量的子组进行描述。

[0231] 附加地, 可以在心理声学模块 950 的框架中执行的替换算法, 检查与生成的信号的基础频谱分量 (例如频带) 相关的每个频谱信息, 以标识仅具有单一有效分量的频谱分量。对于这些频带, 可以从编码器拷贝输入比特流的相应输入数据流的量化值, 而不对指定频谱分量的相应频谱数据进行重新编码或重新量化。在一些情况下, 所有量化数据可以从单一有效输入信号取得, 以形成输出比特流或者输出数据流, 使得对于设备 500 而言能够实现输入数据流的无损编码。

[0232] 此外, 可以省略编码器内如心理声学分析之类的处理步骤。这允许缩短编码过程, 从而降低计算复杂度, 因为原则上仅将数据从一个比特流拷贝入另一比特流必须在特定情况下执行。

[0233] 例如, 在 PNS 的情况下, 可以执行替换, 因为 PNS 编码的频带的噪声因子可以从输出数据流之一拷贝至输出数据流。可以使用合适的 PNS 参数来替换各个频谱分量, 因为 PNS

参数是针对频谱分量指定的,或者换言之,是互相独立的非常好的近似。

[0234] 然而,可能出现以下情况:所描述的算法的两个有力应用可能得到退化的收听体验或者不利的质量降低。因此,将替换限制在各个帧,而不是与各个频谱分量相关的频谱信息,是可取的。在这种操作模式中,可以不变地执行不相关性估计或不相关性确定以及替换分析。然而,在这种操作模式中,仅当有效帧内的所有或至少大量频谱分量可替换时,才执行替换。

[0235] 尽管这可能导致较少次数的替换,在一些情况下,可以改进频谱信息的内在强度,得到甚至略微改进的质量。

[0236] 自然,上述实施例针对它们实现方式而不同。尽管在前述实施例中,将 Huffman 解码和编码描述为单一熵编码方案,但是也可以使用其他熵编码方案。此外,目前不需要实现熵编码器或熵解码器。相应地,尽管先前实施例的描述主要关注 ACC-ELD 编解码器,但是也可以使用其他编解码器来提供输入数据流和对参与者侧的输出数据流进行解码。例如,可以采用基于例如无块长度切换的单一窗口的任何编解码器。

[0237] 如图 8 和 11 所示的实施例的先前描述也已经示出,其中描述的模块不是强制必需的。例如,根据本发明实施例的设备可以简单地通过对帧的频谱信息进行操作来实现。

[0238] 还应当注意,根据图 6 至 12C 描述的实施例可以以非常不同的方式来实现。例如,可以基于分立的电学和电子器件(如电阻器、晶体管、电感器等等)来实现用于对多个输入数据流进行混合的设备 500/1500 及其处理单元 520/1520。此外,还可以仅基于集成电路,例如以 SOC(SOC = 片上系统)、如 CPU(CPU = 中央处理单元)、GPU(GPU = 图形处理单元)之类的处理器以及如专用集成电路(ASIC)之类的其他集成电路(IC),来实现根据本发明的实施例。

[0239] 还应当注意,在实现根据本发明实施例的设备的整个过程中,作为离散实现的一部分或作为集成电路的一部分的电学设备可以用于不同目的和不同功能。自然,还可以使用基于集成电路和分立电路的电路的组合来实现根据本发明的实施例。

[0240] 基于处理器,根据本发明的实施例还可以基于计算机程序、软件程序或在处理器上执行的程序来实现。

[0241] 换言之,根据本发明方法的实施例的特定实现要求,本发明方法的实施例可以以硬件或软件来实现。可以使用数字存储介质,尤其是盘、CD 或 DVD 来执行实现,所述数字存储介质上存储有电子可读信号,与可编程计算机或处理器协作,以执行本发明方法的实施例。因此,一般地,本发明的实施例是一种计算机程序产品,具有在机器可读载体上存储的程序代码,当计算机程序产品在计算机或处理器上运行时,所述程序代码操作于执行本发明方法的实施例。再换言之,因此,本发明方法的实施例是一种具有程序代码的计算机程序,当计算机程序在计算机或处理器上运行时,程序代码用于执行本发明方法的至少一个实施例。可以由计算机、芯片卡、智能卡、专用集成电路、片上系统(SOC)或集成电路(IC)来形成处理器。

[0242] 参考标记列表

[0243] 100 会议系统

[0244] 110 输入

[0245] 120 解码器

- [0246] 130 加法器
- [0247] 140 编码器
- [0248] 150 输出
- [0249] 160 会议终端
- [0250] 170 编码器
- [0251] 180 解码器
- [0252] 190 时间 / 频率转换器
- [0253] 200 量化器 / 编码器
- [0254] 210 解码器 / 解量化器
- [0255] 220 频率 / 时间转换器
- [0256] 250 数据流
- [0257] 260 帧
- [0258] 270 附加信息块
- [0259] 300 频率
- [0260] 310 频带
- [0261] 500 设备
- [0262] 510 输入数据流
- [0263] 520 处理单元
- [0264] 530 输出数据流
- [0265] 540 帧
- [0266] 550 输出帧
- [0267] 560 频谱分量
- [0268] 570 箭头
- [0269] 580 虚线
- [0270] 700 比特流解码器
- [0271] 710 比特流读取器
- [0272] 720 Huffman 编码器
- [0273] 730 解量化器
- [0274] 740 缩放器
- [0275] 750 第一单元
- [0276] 760 第二单元
- [0277] 770 立体声解码器
- [0278] 780 PNS 解码器
- [0279] 790 TNS 解码器
- [0280] 800 混合单元
- [0281] 810 频谱混合器
- [0282] 820 优化模块
- [0283] 830 SBR 混合器
- [0284] 850 比特流编码器

- [0285] 860 第三单元
- [0286] 870TNS 编码器
- [0287] 880PNS 编码器
- [0288] 890 立体声编码器
- [0289] 900 第四单元
- [0290] 910 缩放器
- [0291] 920 量化器
- [0292] 930Huffman 编码器
- [0293] 940 比特流写入器
- [0294] 950 心理声学模块
- [0295] 1500 设备
- [0296] 1520 处理单元
- [0297] 1545 控制值
- [0298] 1550 输出帧
- [0299] 1555 控制值

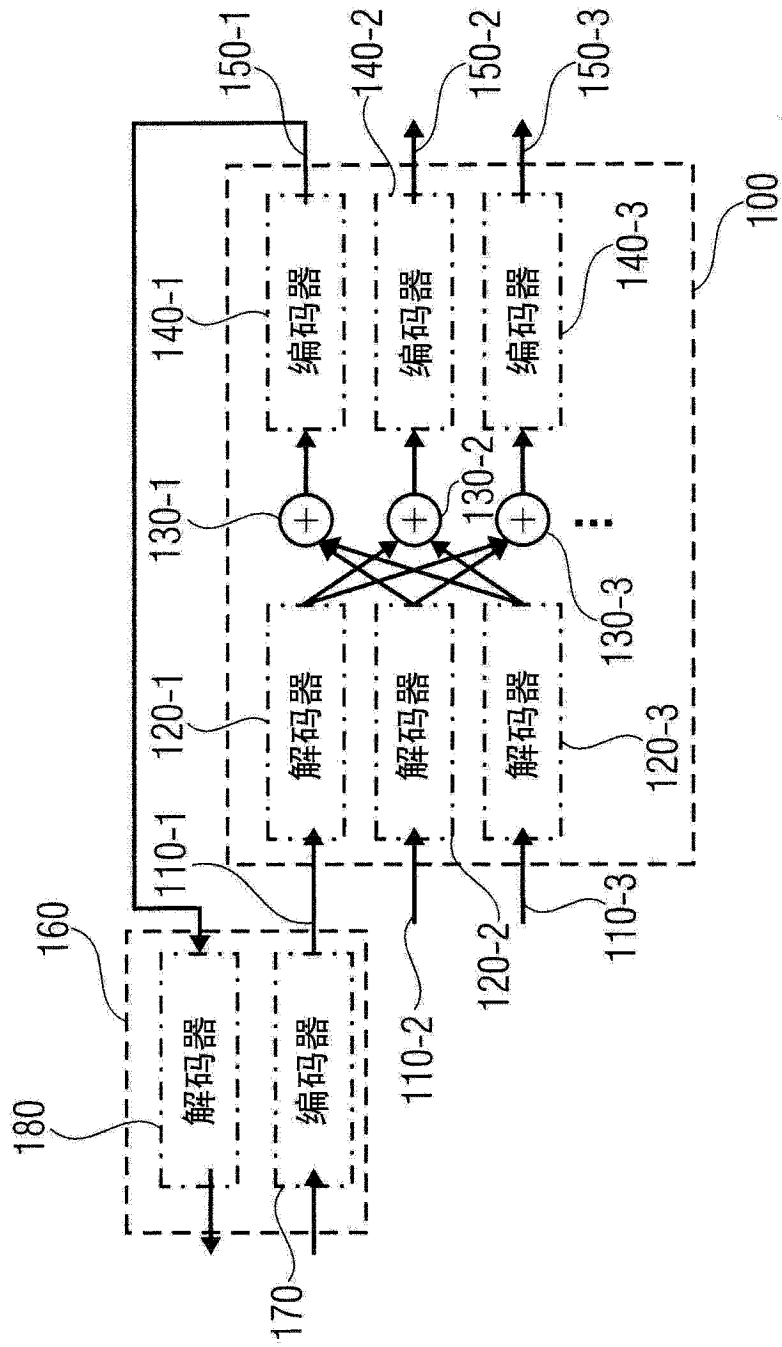


图 1

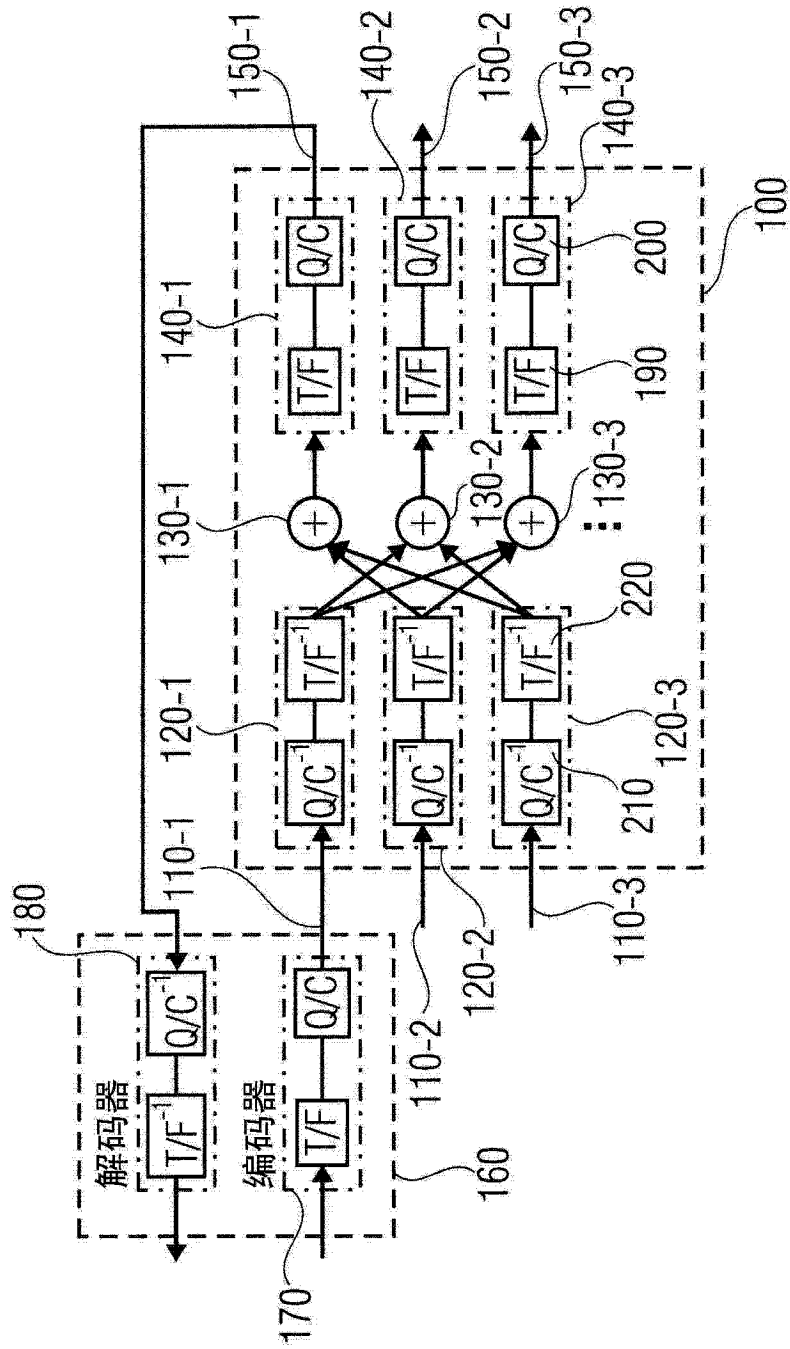


图 2

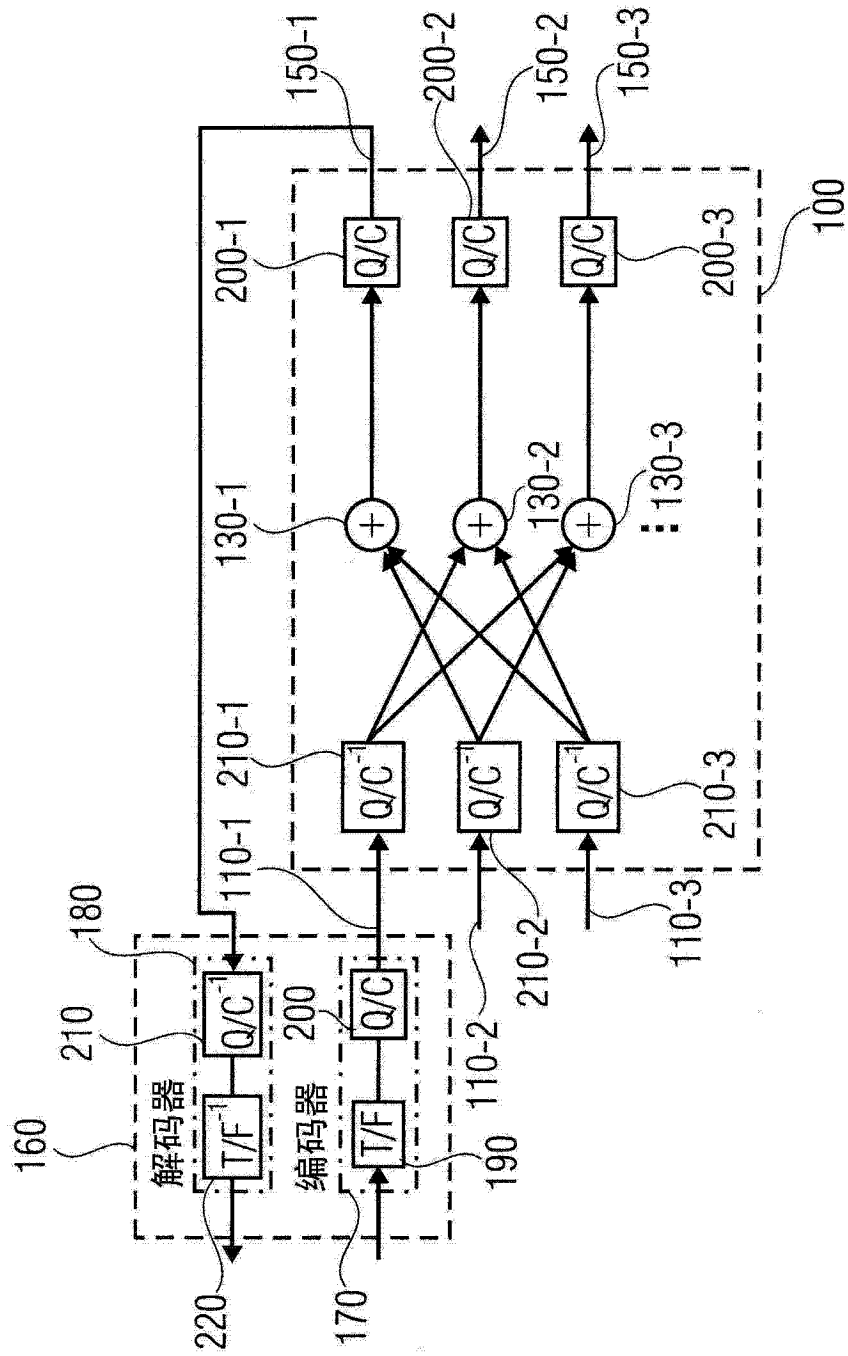


图 3

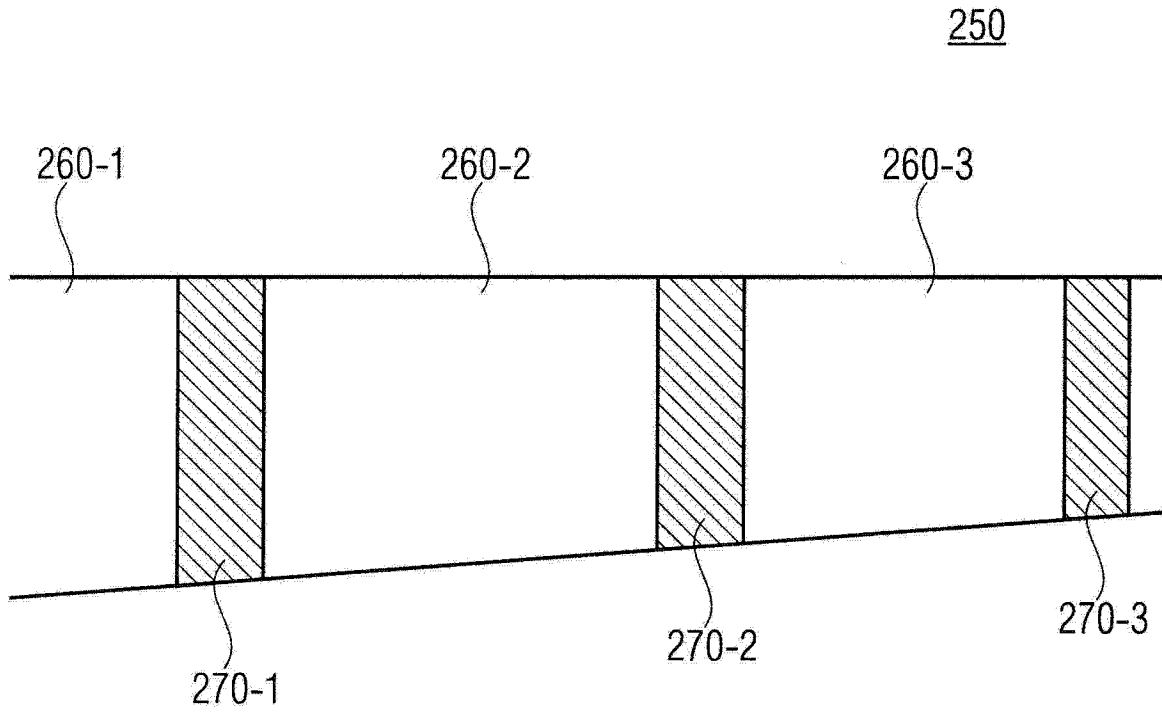


图 4

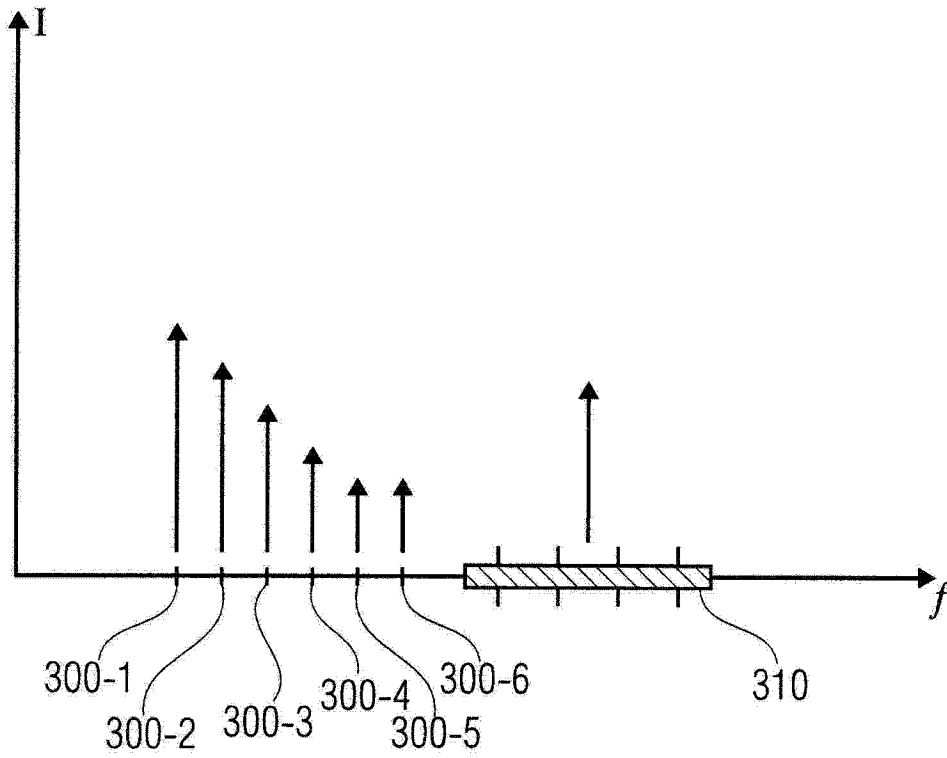


图 5

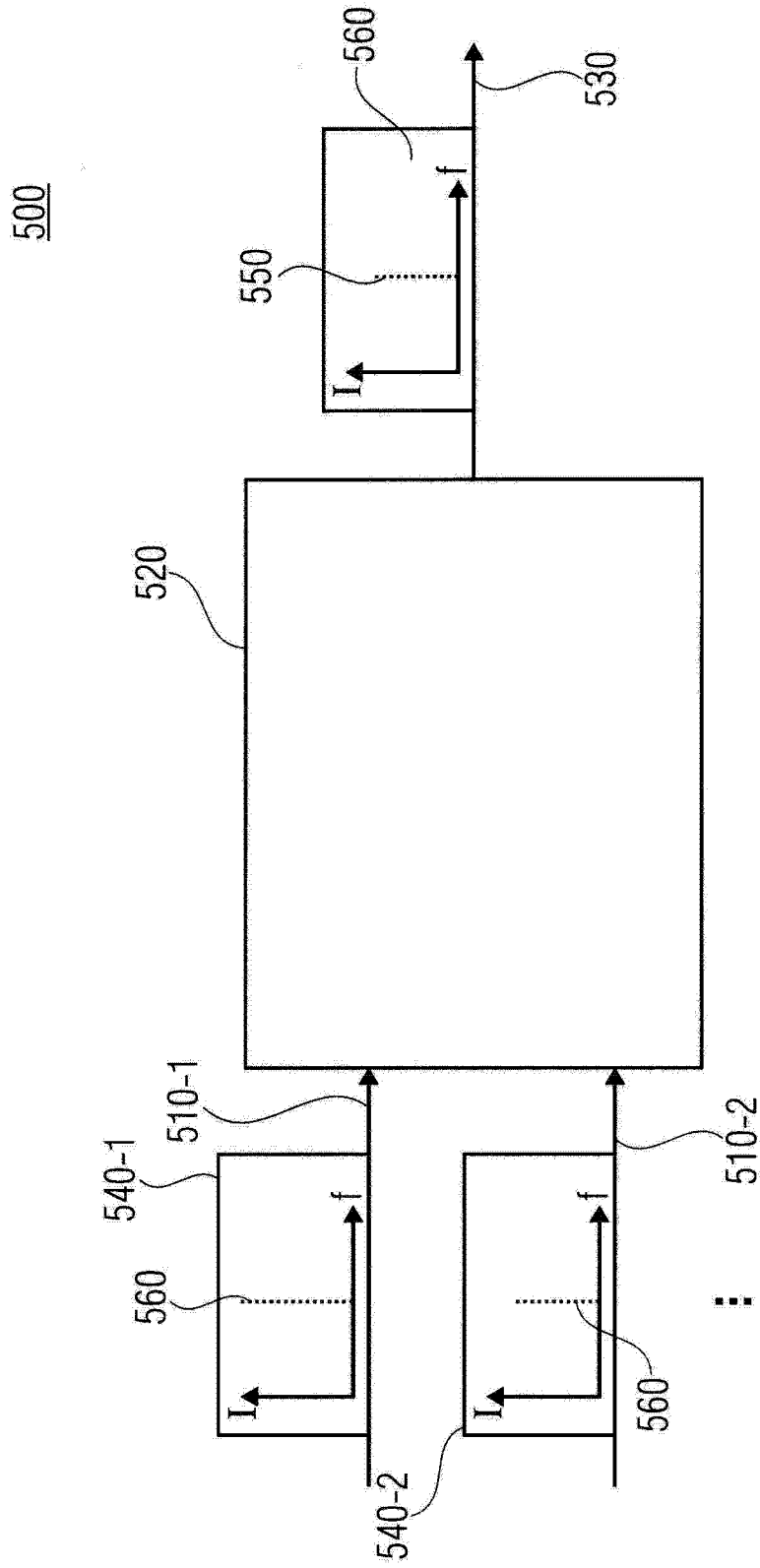


图 6

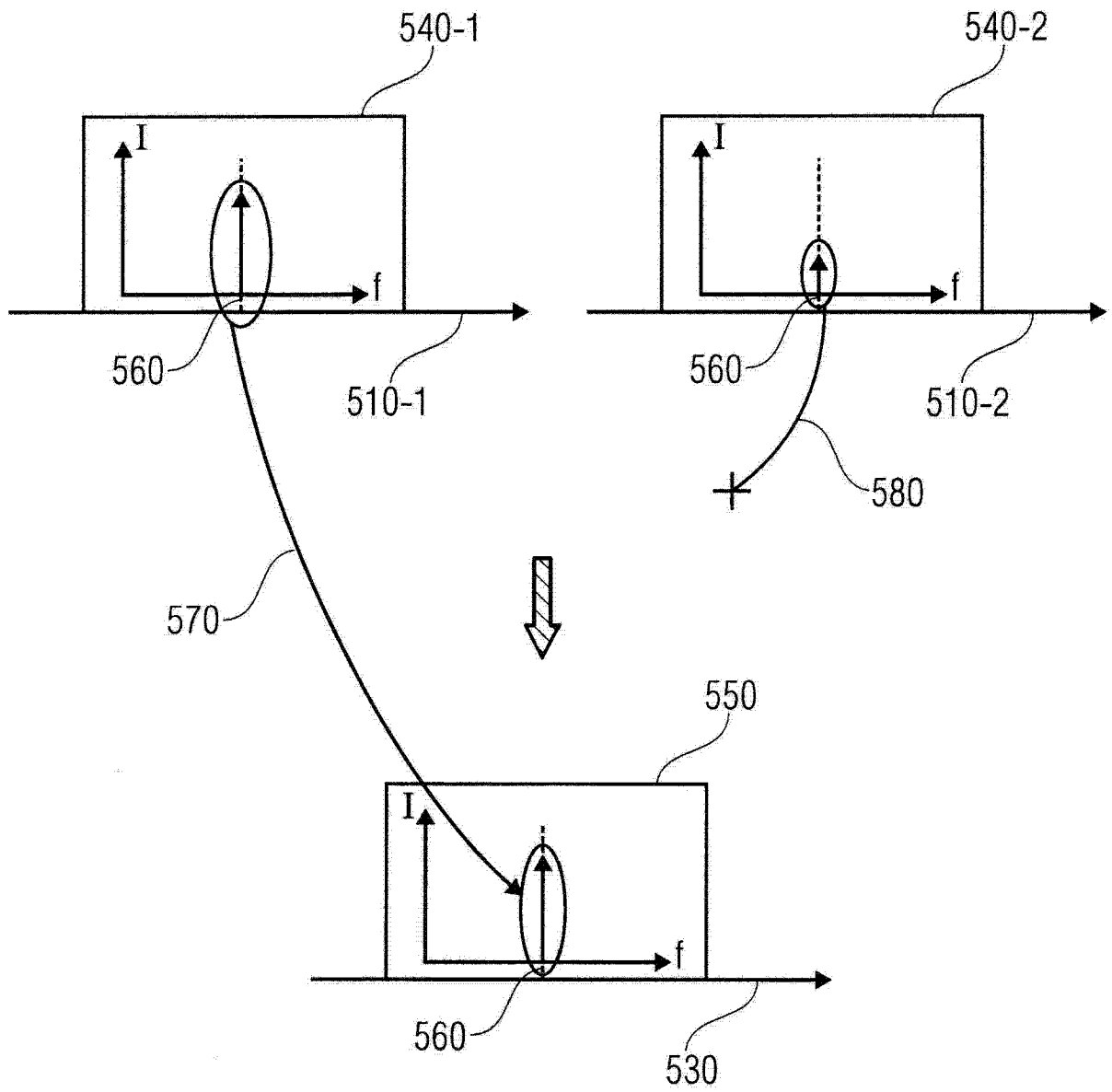


图 7

500

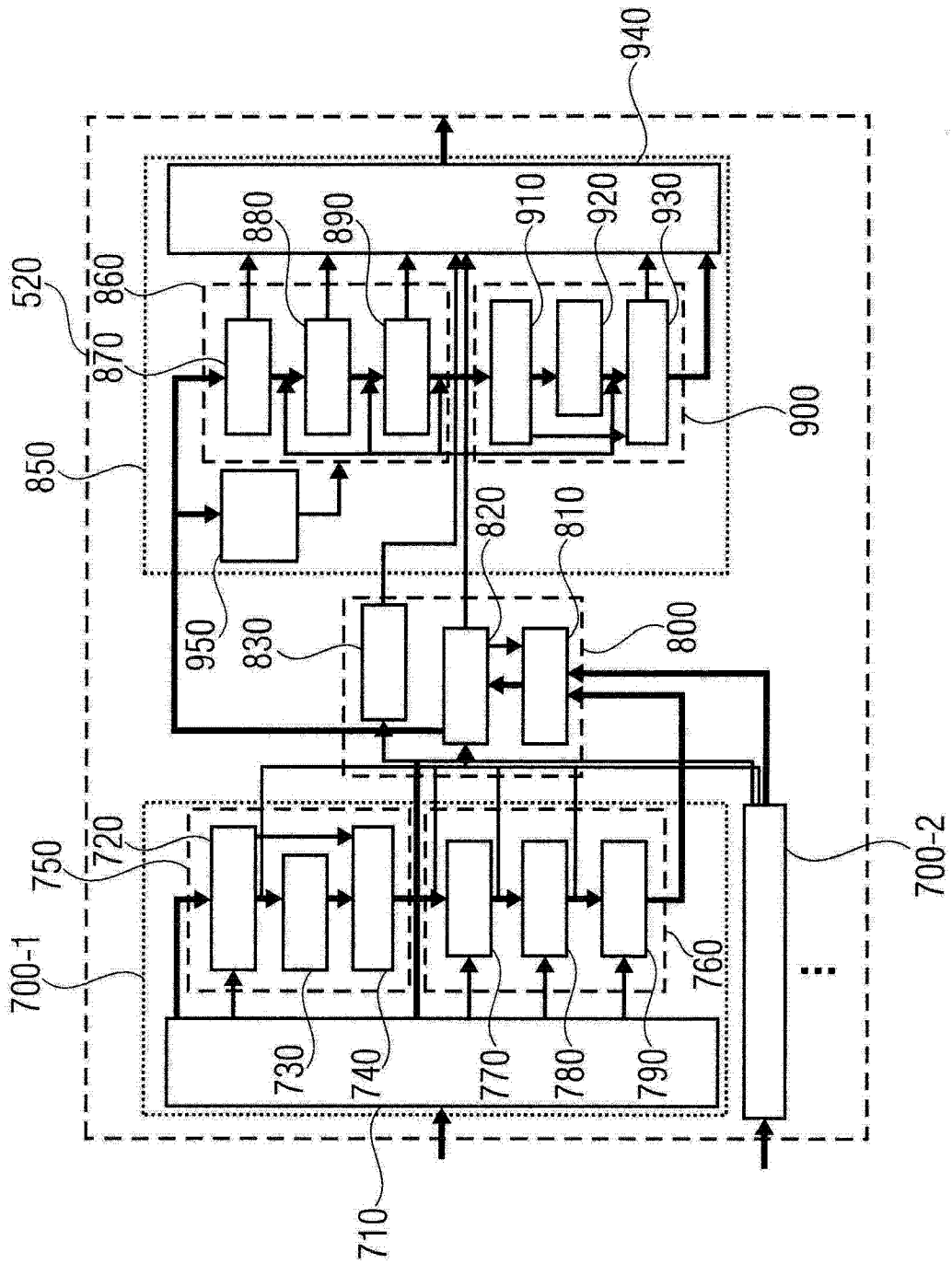


图 8

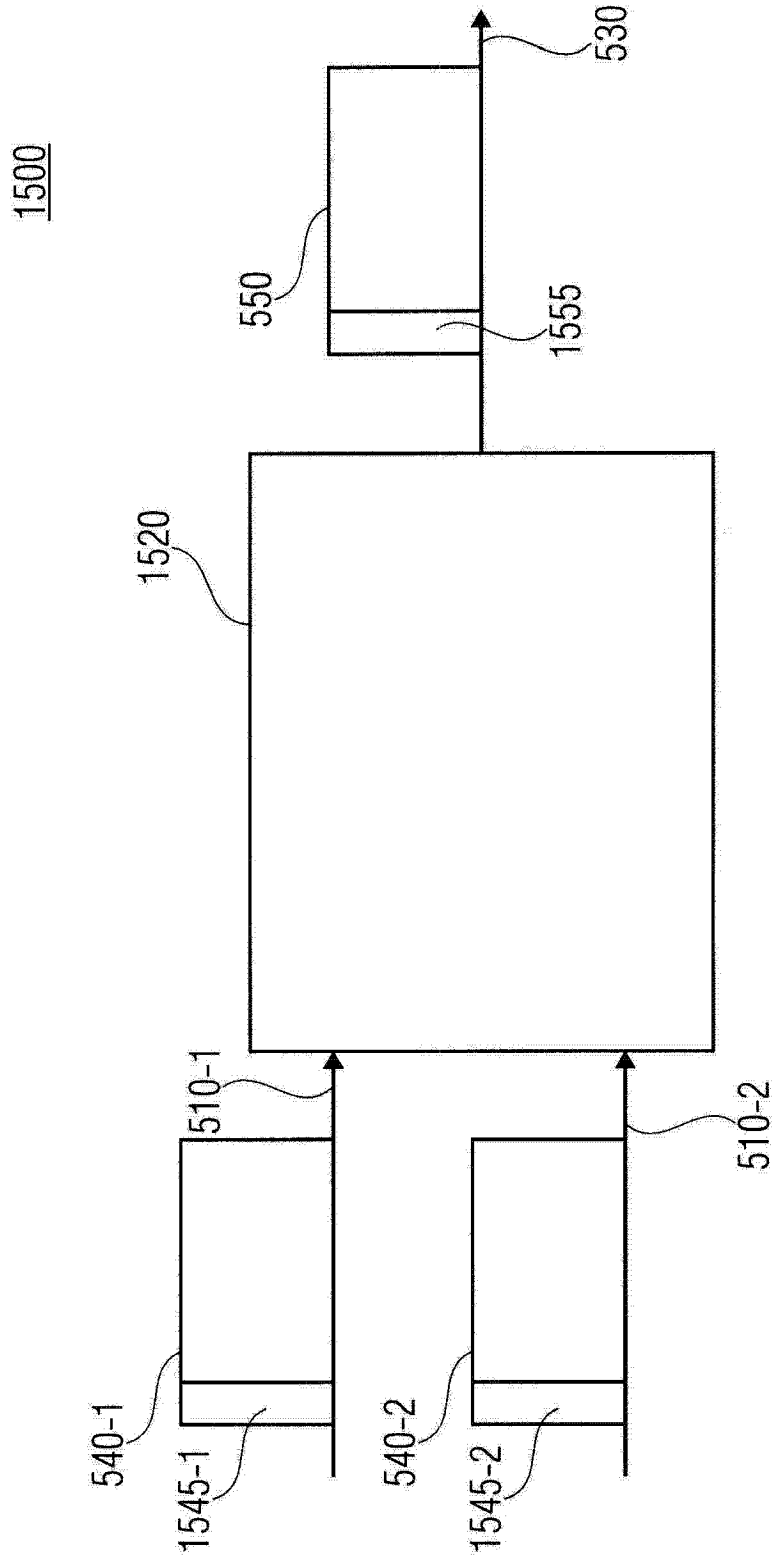


图 9

1500

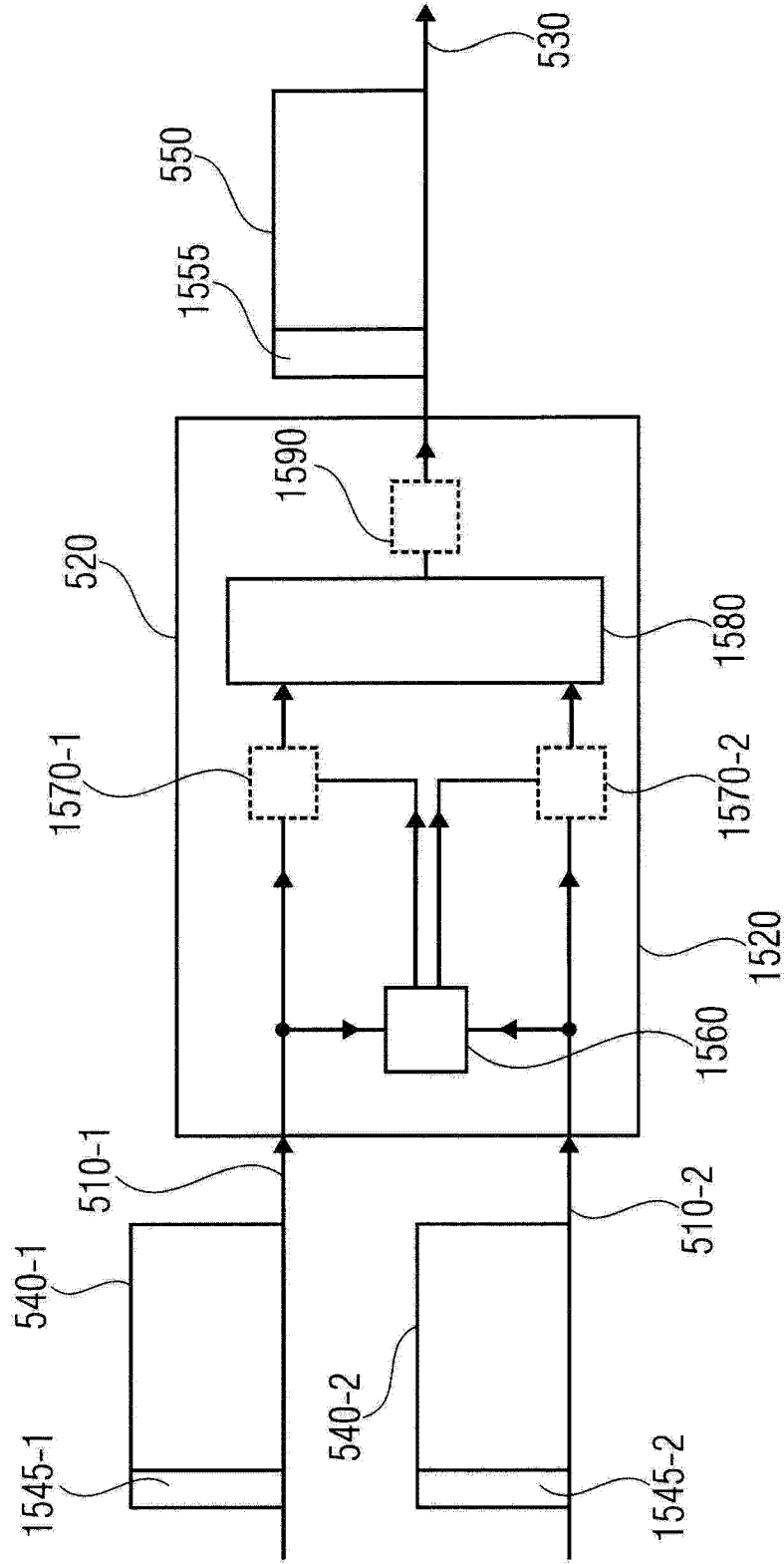


图 10

1500

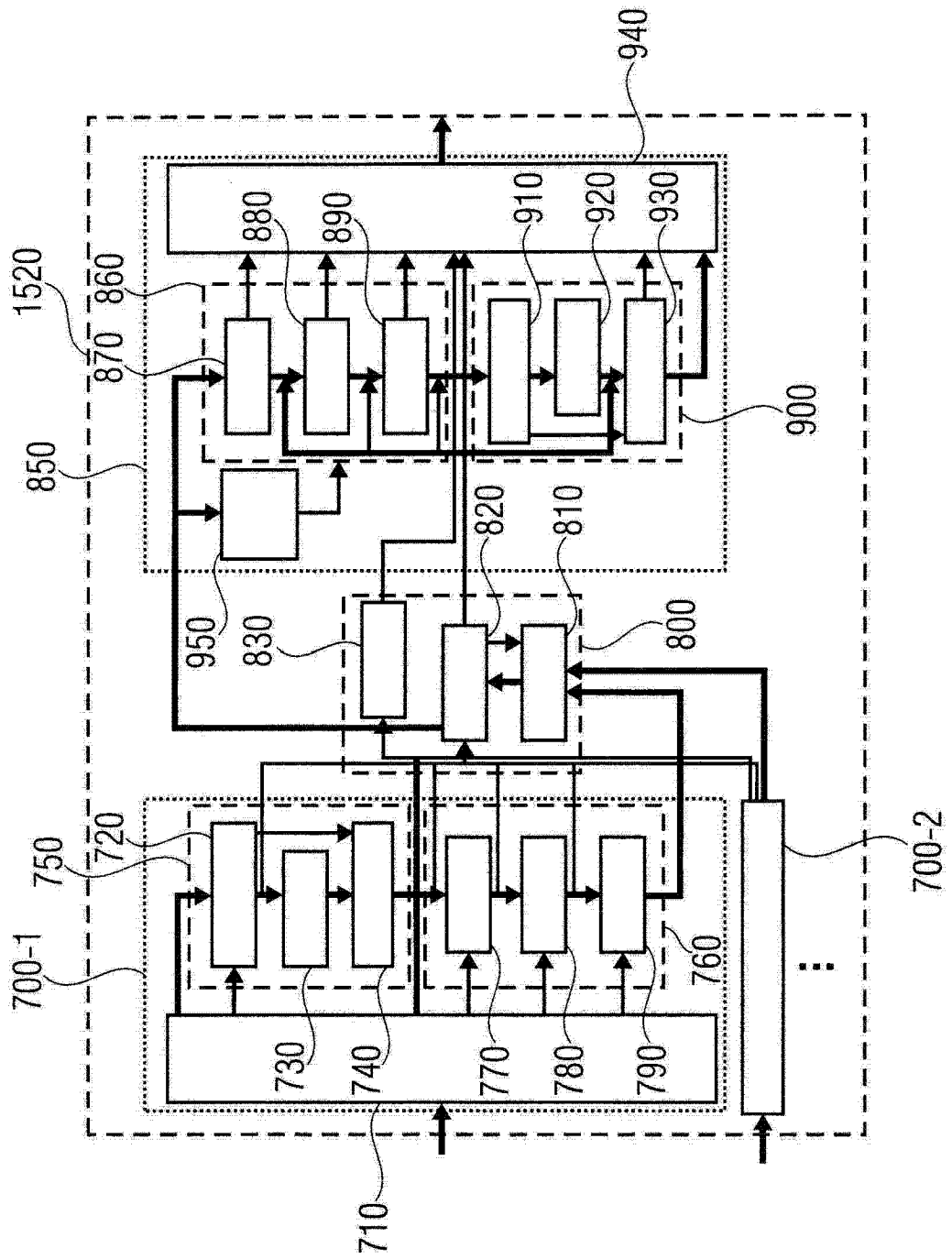


图 11

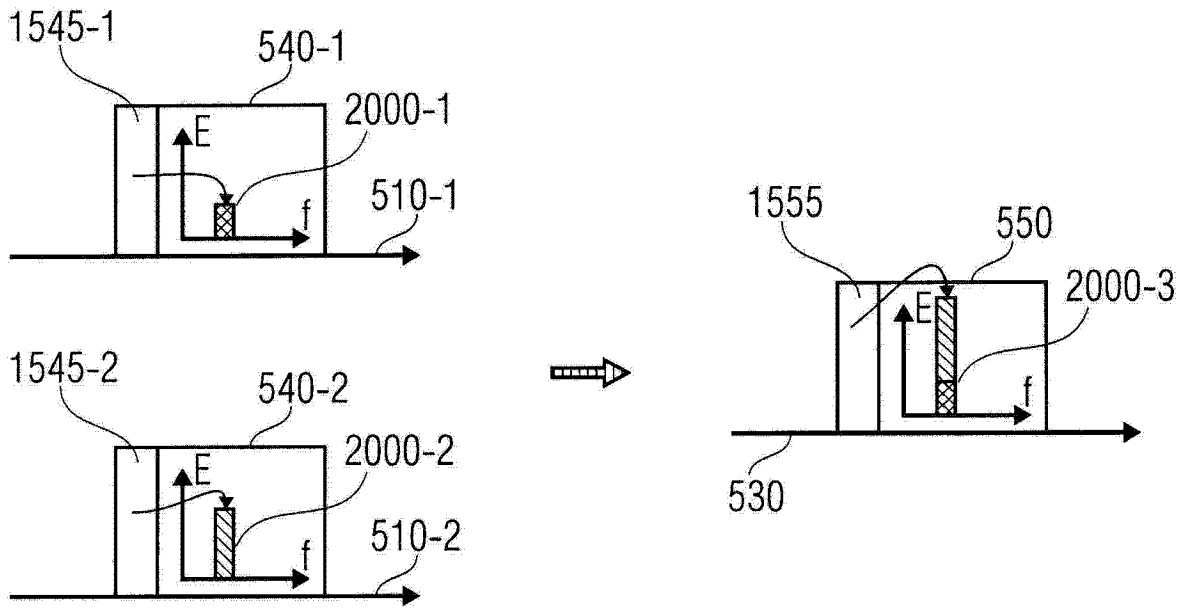


图 12A

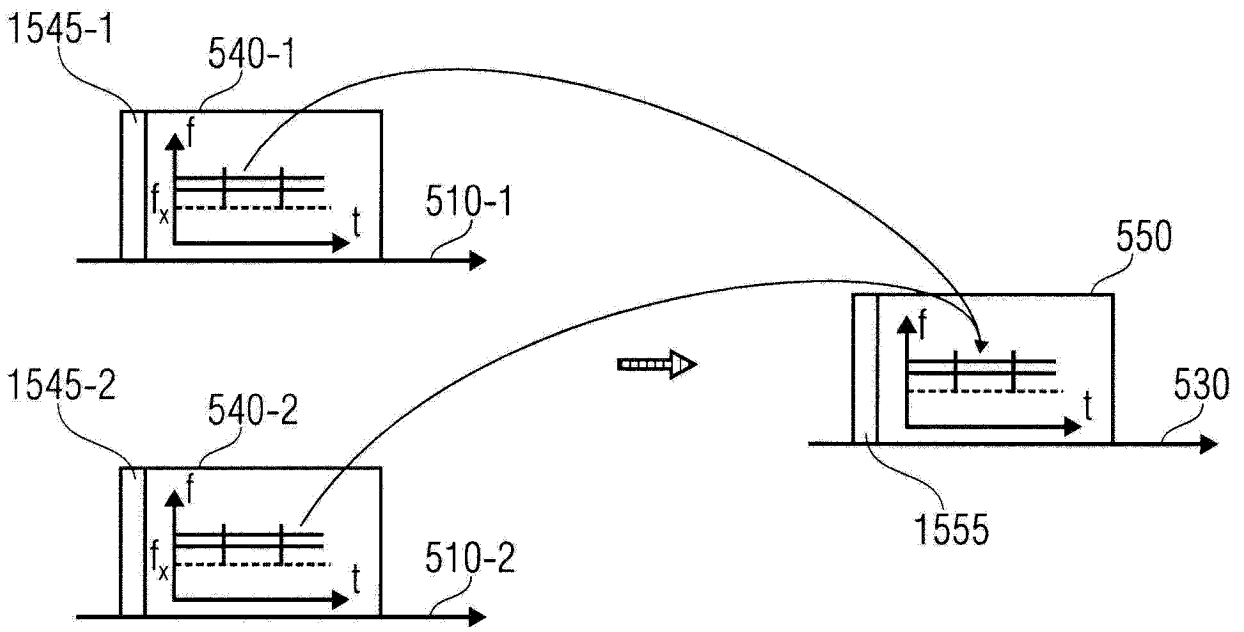


图 12B

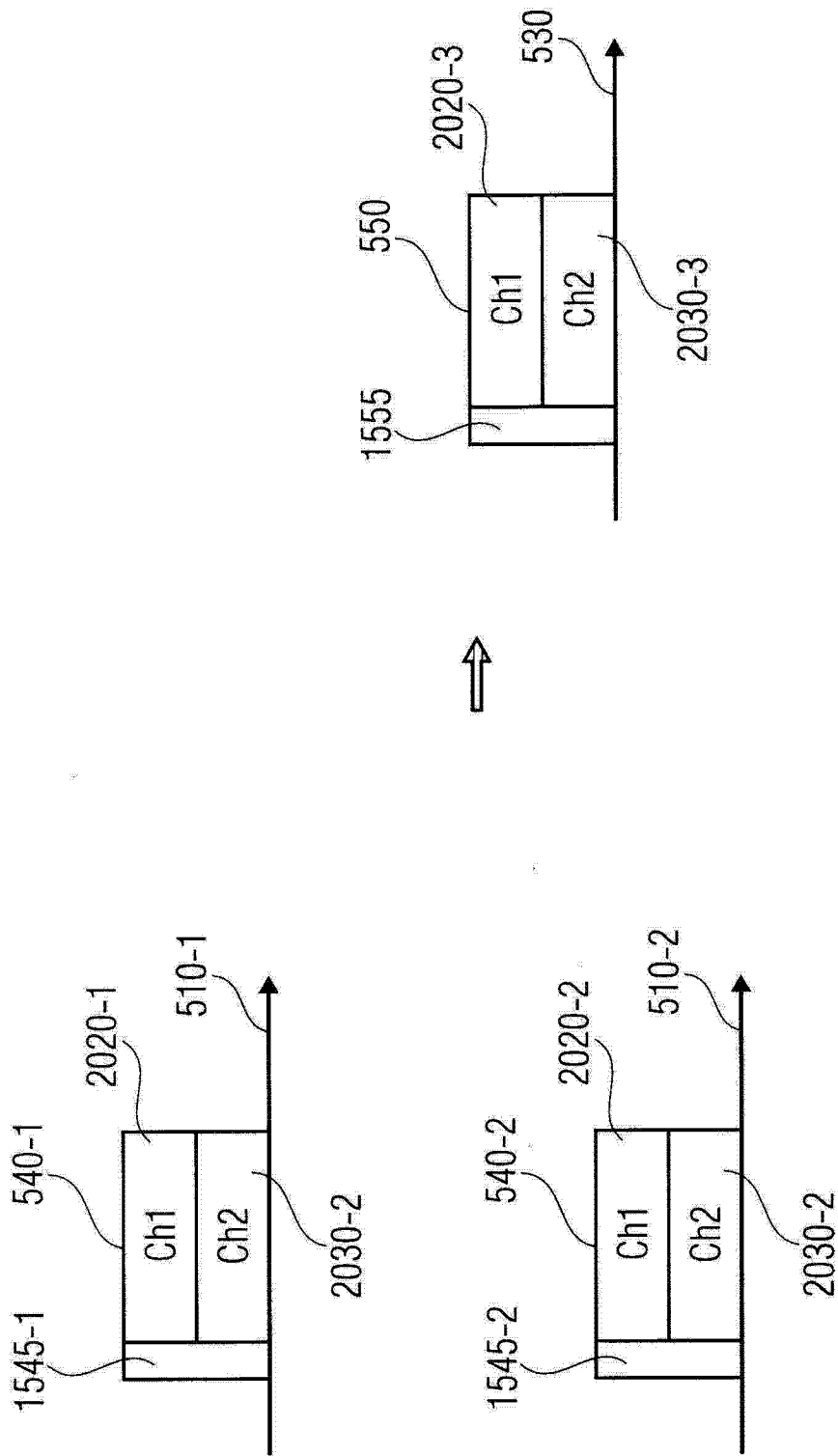


图 12C