



(12) 发明专利

(10) 授权公告号 CN 108140392 B

(45) 授权公告日 2023.04.18

(21) 申请号 201680058435.9

(22) 申请日 2016.10.07

(65) 同一申请的已公布的文献号
申请公布号 CN 108140392 A

(43) 申请公布日 2018.06.08

(30) 优先权数据
15306589.1 2015.10.08 EP
15306653.5 2015.10.15 EP
62/361,416 2016.07.12 US
62/361,461 2016.07.12 US

(85) PCT国际申请进入国家阶段日
2018.04.04

(86) PCT国际申请的申请数据
PCT/EP2016/073969 2016.10.07

(87) PCT国际申请的公布数据

W02017/060410 EN 2017.04.13

(73) 专利权人 杜比国际公司
地址 荷兰阿姆斯特丹

(72) 发明人 S·科顿 A·克鲁格

(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 宿小猛

(51) Int.Cl.
G10L 19/008 (2006.01)

(56) 对比文件
WO 2015140293 A1, 2015.09.24
CN 104471641 A, 2015.03.25

审查员 张岩

权利要求书2页 说明书17页 附图7页

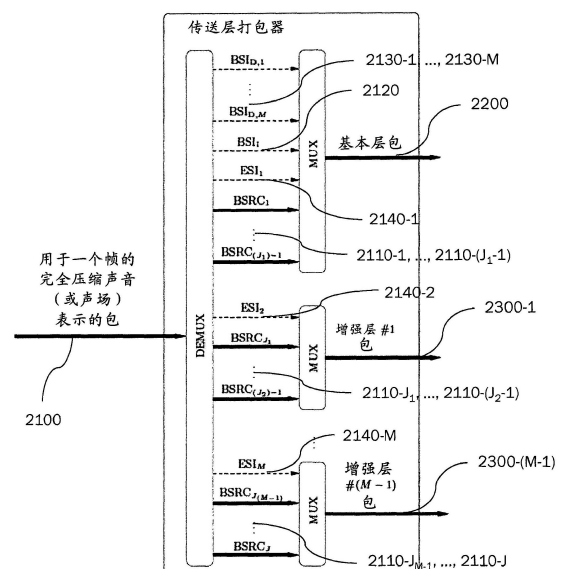
(54) 发明名称

用于压缩声音或声场表示的分层编解码

(57) 摘要

本文件涉及声音或声场的压缩声音表示的分层编码方法。该压缩声音表示包含包括多个分量的基本压缩声音表示、用于将该基本压缩声音表示解码成该声音或声场的基本重构声音表示的基本辅助信息、及包括用于改善该基本重构声音表示的参数的增强辅助信息。该方法包含将该多个分量细分为多个分量组并将该多个组中的每一个分配至多个分级层中的相应一个分级层中,该组的数目对应于该层的数目,且该多个层包括基本层及一个或多个分级增强层,将该基本辅助信息添加至该基本层、及从该增强辅助信息确定增强辅助信息的多个部分并将增强辅助信息的该多个部分中的每一个分配至该多个层中的相应一个分级层,其中增强辅助信息的每一部分包括用于改善可从包括在该相应层及低于该相应层的任何层中的数据得到的重构声音表示的参数。本文件进一步涉及声音或声场的压缩声音表示的解码方法,其中该压缩声音表示被编码

在包括基本层及一个或多个分级增强层的多个分级层中,以及涉及用于压缩声音表示的分层编解码的编码器及解码器。



1. 一种对通过使用分层编码被编码为多个分级层的声音或声场的压缩高阶高保真度立体声响复制HOA声音表示进行解码的方法,该方法包含:

接收包含该压缩HOA表示的位流,该压缩HOA表示对应于包括基本层及一个或多个分级增强层的所述多个分级层,其中该多个层中的至少一个层被分配有该声音或声场的基本压缩声音表示的分量,

基于与该基本层关联的基本辅助信息并基于与该一个或多个分级增强层关联的增强辅助信息解码该压缩HOA表示,其中所述增强辅助信息的每一部分包括用于改善能够从包括在相应层及低于该相应层的任何层中的数据得到的重构声音表示的参数,

其中该基本辅助信息包括与第一个体单声道信号有关的基本独立辅助信息,该第一个体单声道信号将独立于其他单声道信号被解码。

2. 根据权利要求1所述的方法,其中该基本独立辅助信息指示该第一个体单声道信号代表具有入射方向的方向信号。

3. 根据权利要求1至2中任一项所述的方法,其中该基本辅助信息进一步包括与第二个体单声道信号有关的基本依赖性辅助信息,该第二个体单声道信号将依赖于其他单声道信号被解码。

4. 根据权利要求3所述的方法,其中该基本依赖性辅助信息包括在该声场内方向性地分布的基于向量的信号,其中该方向性分布是由向量指定的。

5. 根据权利要求4所述的方法,其中该向量的分量被设定成零且不是压缩向量表示的一部分。

6. 根据权利要求1至5中任一项所述的方法,其中该基本压缩声音表示的分量对应于单声道信号;及

单声道信号代表HOA表示的系数序列或主要声音信号。

7. 根据权利要求1至6中任一项所述的方法,其中该位流包括分别对应于该多个分级层的数据有效载荷。

8. 根据权利要求1至7中任一项所述的方法,其中该增强辅助信息包括与以下中的至少一个有关的参数:空间预测、子带方向信号合成、及参数环境复制。

9. 根据权利要求1至8中任一项所述的方法,其中该增强辅助信息包括允许从方向信号预测该声音或声场的缺失部分的信息。

10. 根据权利要求1至9中任一项所述的方法,进一步包含:

对于每一层确定相应层是否已被有效地接收;及

确定紧接在未被有效地接收的最低层以下的层的层索引。

11. 一种对通过使用分层编码被编码为多个分级层的声音或声场的压缩高阶高保真度立体声响复制HOA声音表示进行解码的设备,该设备包含:

接收器,用于接收包含该压缩HOA表示的位流,该压缩HOA表示对应于包括基本层及一个或多个分级增强层的所述多个分级层,其中该多个层中的至少一个层被分配有该声音或声场的基本压缩声音表示的分量,

解码器,用于基于与该基本层关联的基本辅助信息并基于与该一个或多个分级增强层关联的增强辅助信息解码该压缩HOA表示,其中所述增强辅助信息的每一部分包括用于改善能够从包括在相应层及低于该相应层的任何层中的数据得到的重构声音表示的参数,

其中该基本辅助信息包括与第一个体单声道信号有关的基本独立辅助信息,该第一个体单声道信号将独立于其他单声道信号被解码。

12. 根据权利要求11所述的设备,其中该基本独立辅助信息包括至少指示个体单声道信号以代表具有入射方向的方向信号。

13. 根据权利要求11至12中任一项所述的设备,其中该基本辅助信息进一步包括与第二个体单声道信号有关的基本依赖性辅助信息,该第二个体单声道信号将依赖于其他单声道信号被解码。

14. 根据权利要求13所述的设备,其中该基本依赖性辅助信息包括在该声场内方向性地分布的基于向量的信号,其中该方向性分布是由向量指定的。

15. 根据权利要求14所述的设备,其中该向量的分量被设定成零且不是压缩向量表示的一部分。

16. 根据权利要求11至15中任一项所述的设备,其中该基本压缩声音表示的分量对应于单声道信号;及

单声道信号代表HOA表示的系数序列或主要声音信号。

17. 根据权利要求11至16中任一项所述的设备,其中该位流包括分别对应于该多个分级层的数据有效载荷。

18. 根据权利要求11至17中任一项所述的设备,其中该增强辅助信息包括与以下中的至少一个有关的参数:空间预测、子带方向信号合成、及参数环境复制。

19. 根据权利要求11至18中任一项所述的设备,其中该增强辅助信息包括允许从方向信号预测该声音或声场的缺失部分的信息。

20. 根据权利要求11至19中任一项所述的设备,进一步包含:

对于每一层确定相应层是否已被有效地接收;及

确定紧接在未被有效地接收的最低层以下的层的层索引。

21. 一种设备,包括:

处理器,以及

存储介质,包括软件程序,所述软件程序在由所述处理器执行时使得实行根据权利要求1至10中任一项所述的方法。

22. 一种存储介质,包括软件程序,所述软件程序在由处理器执行时使得实行根据权利要求1至10中任一项所述的方法。

用于压缩声音或声场表示的分层编解码

[0001] 相关申请的交叉引用

[0002] 本申请要求2015年10月8日提交的欧洲专利申请No.15306589.1和2015年10月15日提交的欧洲专利申请No.15306653.5,以及美国专利申请No.62/361461以及62/361416的优先权,这些专利申请通过引用而全文并入这里。

技术领域

[0003] 本文件涉及用于分层音频编解码的方法及设备。本文件特别涉及用于压缩声音(或声场)表示,例如,高阶高保真度立体声响复制(HOA)声音(或声场)表示,的分层音频编解码的方法及设备。

背景技术

[0004] 对于在时间变化条件下传输通道上的声音(或声场)的流传输,分层编解码是一种使被接收的声音表示的质量适合于传输条件,并特别适于避免不期望的信号漏失的方法。

[0005] 针对分层编解码,常将声音(或声场)表示细分为相对小尺寸的高优先度基本层、及具有递减优先度及任意尺寸的附加增强层。典型地将各增强层假设成包含递增信息以补足所有较低层的信息,以改善声音(或声场)表示的质量。用于各个层的传输的错误保护量基于它们的优先度被控制。特别地,基本层设有高错误保护,由于其的小尺寸,这是合理且实惠的。

[0006] 然而,仍需要用于特殊种类的声音或声场的压缩表示(的扩展版本),诸如,压缩HOA声音或声场表示,的分层编解码方案。

[0007] 本文件解决了上述问题。特别描述了用于压缩声音及声场表示的分层编解码的方法及编码器/解码器。

发明内容

[0008] 根据一个方面,描述了声音或声场的压缩声音表示的分层编码方法。该压缩声音表示可包括基本压缩声音表示,其包括多个分量。该多个分量可为补充分量。压缩声音表示可进一步包括用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的基本辅助信息。该压缩声音表示可进一步包括增强辅助信息,其包括用于改善(例如,增强)基本重构声音表示的参数。该方法可包括将该多个分量细分(例如,分组)为多个分量组。该方法可进一步包括将多个组中的每一个分配(例如,添加)至多个分级层中相应的一个分级层。该分配可指示各组与层之间的对应关系。可将被分配给相应层的分量称为包括在该层中。组的数目可对应于(例如,等于)层的数目。该多个层可包括基本层及一个或多个分级增强层。该多个分级层可从基本层、经由第一增强层、第二增强层等、直到总体最高增强层(总体最顶层)被排序。该方法可进一步包括将基本辅助信息添加至基本层(例如,出于传输或储存的目的,例如,将基本辅助信息包括在基本层中,或将基本辅助信息分派至该基本层)。该方法可进一步包括从该增强辅助信息确定增强辅助信息的多个部分。该方法可进一步包括将

增强辅助信息的该多个部分中的每一个分配(例如,添加)至该多个层中的相应的一个。增强辅助信息的各部分可包括用于改善可从包括在(例如,分配或添加至)相应层及低于该相应层的任何层中的数据得到的重构(例如,解压缩)声音表示的参数。分层编码可出于通过传输信道进行传输的目的或出于储存在适当存储介质(诸如,CD、DVD、蓝光光盘™)中的目的而实施。

[0009] 如上文所配置的,所提出的方法使得能够将分层编解码高效地应用于包含多个分量以及具有如上文设定的性质的基本及增强辅助信息(例如,独立基本辅助信息及增强辅助信息)的压缩声音表示。所提出的方法特别确保各层包括用于从包括在直至关注层的任何层中的分量重构声音表示的合适辅助信息。其中“直至关注层的层”理解为,例如,包括基本层、第一增强层、第二增强层等、直到该关注层。因此,不管实际最高可使用层(例如,低于未有效地接收的最低层的层,使得已有效地接收低于最高可使用层的所有层及该最高可使用层自身)如何,即使重构声音表示可能与完全(例如,完整)声音表示不同,解码器仍将能够改善或增强重构声音表示。特别地,不管实际最高可用层如何,解码器仅针对单一层(即,针对最高可使用层)解码增强辅助信息的有效载荷(payload)以改善或增强基于包括在直至实际最高可使用层的层中的所有分量可得到的重构声音表示就足够了。即,针对各时间区间(例如,帧),仅需解码增强辅助信息的单一有效载荷。另一方面,所提出的方法允许充分利用在施用分层编解码时可实现的所需带宽的降低的优点。

[0010] 在实施例中,基本压缩声音表示的分量可对应于单声道信号(例如,传送信号或单声道传送信号)。该单声道信号可代表HOA表示的系数序列或主要声音信号。可将单声道信号量化。

[0011] 在实施例中,基本辅助信息可包括与其他分量无关地、单独地指定该多个分量中的一或多者的解码(例如,解压缩)的信息。例如,基本辅助信息可代表与个体单声道信号有关、而与其他单声道信号无关的辅助信息。因此,基本辅助信息可称为独立基本辅助信息。

[0012] 在实施例中,增强辅助信息可代表增强辅助信息。增强辅助信息可包括用于改善(例如,增强)可从基本压缩声音表示及基本辅助信息得到的基本重构声音表示的基本压缩声音表示的预测参数。

[0013] 在实施例中,该方法可进一步包括产生用于多个层的数据(例如,分配或添加至各层或另外包括在各层中的数据)的传输的传送流。基本层可具有最高的传输优先度且分级增强层可具有递减的传输优先度。即,传输的优先度可从基本层减小至第一增强层,从第一增强层减小至第二增强层,并依此类推。用于该多个层的数据的传输的错误保护量可根据传输的相应优先度被控制。因此,能确保可靠地传输至少若干较低层,同时在另一方面通过不对于较高层应用过度错误保护而降低总体所需带宽。

[0014] 在实施例中,该方法可进一步包括针对多个层中的每一个,产生包括相应层的数据的传送层包。例如,针对各时间区间(例如,帧),可对多个层中的每一个产生相应传送层包。

[0015] 在实施例中,压缩声音表示可进一步包括用于将基本压缩声音表示解码为基本重构声音表示的附加基本辅助信息。附加基本辅助信息可包括指定多个分量中的一个或多个的依赖于各其他分量的解码的信息。该方法可进一步包括将附加基本辅助信息分解为附加基本辅助信息的多个部分。该方法可进一步包括将附加基本辅助信息的部分添加至基本层

(例如,出于传输或储存的目的,例如,将附加基本辅助信息的一部分包括在基本层中、或将附加基本辅助信息的一部分分派至基本层)。附加基本辅助信息的各部分可对应于相应层并可包括如下信息,该信息指定分配给该相应层中的一或多个分量的(仅)依赖于分配给该相应层及低于该相应层的任何层的各其他分量的解码。即,附加基本辅助信息的各部分指定该附加基本辅助信息的该部分所对应的该相应层中的分量,而没有参考分配给比该相应层更高的层的任何其他分量。

[0016] 在如此配置的情况下,所提出的方法通过将所有部分添加至基本层而避免附加基本辅助信息碎片化。换言之,将附加基本辅助信息的所有部分包括在基本层中。附加基本辅助信息的分解确保对于各层,附加基本辅助信息的一部分可用,而不需要知晓较高层中的分量。因此,不管实际最高可使用层如何,解码器将包括在直至最高可使用层的层中的附加基本辅助信息解码就足够了。

[0017] 在实施例中,附加基本辅助信息可包括指定该多个分量中的一个或多个的依赖于其他分量的解码(例如,解压缩)的信息。例如,附加基本辅助信息可代表依赖于其他单声道信号的、与个体单声道信号有关的辅助信息。因此,附加基本辅助信息可称为依赖性基本辅助信息。

[0018] 在实施例中,压缩声音表示可针对连续时间区间被处理,例如,具有相等尺寸的时间区间。连续时间区间可以是帧。因此,该方法可在帧的基础上操作,即,压缩声音表示可用逐帧方式被编码。压缩声音表示可对于各连续时间区间(例如,对于各帧)可用。即,获得压缩声音表示的压缩操作可在帧的基础上操作。

[0019] 在实施例中,该方法可进一步包括产生配置信息,其为各层指示分配给该层的基本压缩声音表示的分量。因此,解码器能迅速地取得解码所需的信息而无须不必要地解析接收的数据有效载荷。

[0020] 根据另一方面,描述了声音或声场的压缩声音表示的分层编码方法。该压缩声音表示可包括基本压缩声音表示,其包括多个分量。该多个分量可以是补充分量。该压缩声音表示可进一步包括用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的基本辅助信息(例如,独立基本辅助信息)及第三信息(例如,依赖性基本辅助信息)。基本辅助信息可包括与其他分量无关地、单独地指定该多个分量中的一个或多个的解码的信息。附加基本辅助信息可包括指定多个分量中的一个或多个的依赖于各其他分量的解码的信息。该方法可包括将该多个分量细分(例如,分组)为多个分量组。该方法可进一步包括将多个组中的每一个分配(例如,添加)至多个分级层中相应的一个分级层。该分配可指示各组与层之间的对应关系。可将分配给相应层的分量称为包括在该层中。组的数目可对应于(例如,等于)层的数目。该多个层可包括基本层及一个或多个分级增强层。该方法可进一步包括将基本辅助信息添加至基本层(例如,出于传输或储存的目的,例如,将基本辅助信息包括在基本层中,或将基本辅助信息分派至该基本层)。该方法可进一步包括将附加基本辅助信息分解为附加基本辅助信息的多个部分及将附加基本辅助信息的一部分添加至基本层(例如,出于传输或储存的目的,例如,将附加基本辅助信息的一部分包括在基本层中、或将附加基本辅助信息的一部分分派至基本层)。附加基本辅助信息的各部分可对应于相应层并包括如下信息,该信息指定分配给该相应层中的一个或多个分量的依赖于分配给该相应层及低于该相应层的任何层的各其他分量的解码。

[0021] 在如此配置的情况下,所提出的方法确保对于各层,适当的附加基本辅助信息可用于解码包括在直至该相应层的任何层中的分量,而无须有效接收或解码(或一般而言,知晓)任何更高层。在压缩HOA表示的情形中,所提出的方法确保在向量编码模式中,合适的V向量可用于属于直至最高可使用层的层的所有分量。特别地,所提出的方法排除了对应于较高层中的分量的V向量的元素未被显式地标示(signal)的情形。因此,包括在直至最高可使用层的层中的信息对于解码(例如,解压缩)属于直至最高可使用层的层中的任何分量是足够的。因此,即使较高层未被解码器有效地接收,仍可确保较低层的相应重构HOA表示的适当解压缩。另一方面,所提出的方法允许充分利用在应用分层编解码时可实现的所需带宽的降低的优点。

[0022] 此方面的实施例可与上述方面的实施例相关。

[0023] 根据另一方面,描述了声音或声场的压缩声音表示的分层编码方法。该压缩声音表示可已编码在多个分级层中。该多个分级层可包括基本层及一个或多个分级增强层。多个层可已被分配有声音或声场的基本压缩声音表示的分量。换言之,该多个层可包括基本压缩辅助信息的分量。这些分量可在各分量组中分配给各层。该多个分量可以是补充分量。基本层可包括用于解码基本压缩声音表示的基本辅助信息。各层可包含包括用于改善基本重构声音表示的参数的增强辅助信息的一部分,该基本重构声音表示可从包括在相应层及低于该相应层的任何层中的数据得到。该方法可包括接收分别对应于多个分级层的数据有效载荷。该方法可进一步包括确定第一层索引,其指示要用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的多个层中的最高可使用层。该方法可进一步包括使用基本辅助信息从分配给该最高可使用层及低于该最高可使用层的任何层的分量得到基本重构声音表示。该方法可进一步包括确定第二层索引,其指示增强辅助信息的哪一部分应被用于改善(例如,增强)基本重构声音表示。该方法可包括参考第二层索引从该基本重构声音表示得到声音或声场的重构声音表示。

[0024] 在如此配置的情况下,所提出的方法通过在最佳可能程度下使用可用(例如,有效地接收的)信息来确保重构声音表示具有最佳质量。

[0025] 在实施例中,基本压缩声音表示的分量可对应于单声道信号(例如,单声道传送信号)。该单声道信号可代表HOA表示的系数序列或主要声音信号。可将单声道信号量化。

[0026] 在实施例中,基本辅助信息可包括与其他分量无关地、单独地指定该多个分量中的一或多者的解码(例如,解压缩)的信息。例如,基本辅助信息可代表与个体单声道信号有关、而与其他单声道信号无关的辅助信息。因此,基本辅助信息可称为独立基本辅助信息。

[0027] 在实施例中,增强辅助信息可代表增强辅助信息。增强辅助信息可包括用于改善(例如,增强)可从基本压缩声音表示及基本辅助信息得到的基本重构声音表示的基本压缩声音表示的预测参数。

[0028] 在实施例中,该方法可进一步包括对于各层,确定该相应层是否已有效地接收。该方法可进一步包括将该第一层索引确定为紧接在未被有效地接收的最低层之下的层的层索引。

[0029] 在实施例中,确定该第二层索引可涉及确定该第二层索引等于该第一层索引,或将指示在得到重构声音表示时不使用任何增强辅助信息的索引值确定为第二层索引。在后一情形中,该重构声音表示可等于基本重构声音表示。

[0030] 在实施例中,数据有效载荷可针对连续时间区间(例如,相等尺寸的时间区间)被接收及处理。连续时间区间可以是帧。因此,该方法可在帧的基础上操作。该方法可进一步包括,若连续时间区间的压缩声音表示可被彼此独立地解码,则确定第二层索引等于第一层索引。

[0031] 在实施例中,数据有效载荷可针对连续时间区间(例如,相等尺寸的时间区间)被接收及处理。连续时间区间可以是帧。因此,该方法可在帧的基础上操作。该方法可进一步包括,针对连续时间区间中的给定时间区间,若连续时间区间的压缩声音表示不能被彼此独立地解码,则对于各层确定该相应层是否已有效地接收。该方法可进一步包括将给定时间区间的第一层索引确定为在该给定时间区间之前的时间区间的第一层索引及紧接在未被有效地接收的最低层之下的层的层索引中的较小者。

[0032] 在实施例中,该方法可进一步包括,对于给定时间区间,若连续时间区间的压缩声音表示不能被彼此独立地解码,则确定该给定时间区间的第一层索引是否等于在先时间区间的第一层索引。该方法可进一步包括,若该给定时间区间的第一层索引等于在先时间区间的第一层索引,则确定该给定时间区间的第二层索引等于该给定时间区间的第一层索引。该方法可进一步包括,若该给定时间区间的第一层索引不等于在先时间区间的第一层索引,则确定指示在得到重构声音表示时不使用任何增强辅助信息的索引值为该第二层索引。

[0033] 在实施例中,该基本层可包括对应于相应层并包括如下信息的附加基本辅助信息的至少一部分,该信息指定分配给该相应层的分量中的一个或多个分量的依赖于分配给该相应层及低于该相应层的任何层的其他分量的解码。该方法可进一步包括,针对附加基本辅助信息的各部分,通过参考分配给其相应层及低于该相应层的任何层的分量来解码附加基本辅助信息的该部分。该方法可进一步包括通过参考分配给该最高可使用层及在最高可使用层和该相应层之间的任何层的分量来校正附加基本辅助信息的该部分。使用基本辅助信息及从对应于直至该最高可使用层的层的附加基本辅助信息的部分得到的附加基本辅助信息的经校正的部分,可从分配给最高可使用层及低于该最高可使用层的任何层的分量得到基本重构声音表示。

[0034] 在实施例中,附加基本辅助信息可包括指定多个分量中的一个或多个的依赖于其他分量的解码(例如,解压缩)的信息。例如,附加基本辅助信息可代表与个体单声道信号有关的、依赖于其他单声道信号的辅助信息。因此,附加基本辅助信息可称为依赖(dependent)基本辅助信息。

[0035] 根据另一方面,描述声音或声场的压缩声音表示的解码方法。该压缩声音表示可已编码在多个分级层中。该多个分级层可包括基本层及一或多个分级增强层。多个层可被分配有声音或声场的基本压缩声音表示的分量。换言之,该多个层可包括基本压缩辅助信息的分量。这些分量可在各分量组中分配给各层。该多个分量可以是补充分量。基本层可包括用于解码基本压缩声音表示的基本辅助信息。该基本层可进一步包括对应于相应层并包括如下信息的附加基本辅助信息的至少一部分,该信息指定分配给该相应层的分量中的一个或多个分量的依赖于分配给该相应层及低于该相应层的任何层的其他分量的解码。该方法可包括接收分别对应于多个分级层的数据有效载荷。该方法可进一步包括确定第一层索引,其指示要用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的多个层中

的最高可使用层。该方法可进一步包括,针对附加基本辅助信息的各部分,通过参考分配给其相应层及低于该相应层的任何层的分量来解码附加基本辅助信息的该部分。该方法可进一步包括,针对附加基本辅助信息的各部分,通过参考分配给最高可使用层及该最高可使用层和该相应层之间的任何层的分量来校正附加基本辅助信息的该部分。通过使用基本辅助信息及从对应于直至该最高可使用层的层的附加基本辅助信息的部分得到的附加基本辅助信息的经校正的部分,可从分配给最高可使用层及低于该最高可使用层的任何层的分量得到基本重构声音表示。该方法可进一步包含确定第二层索引,其等于第一层索引或指示在解码期间省略增强辅助信息。

[0036] 在如此配置的情况下,所提出的方法确保最终用于解码基本压缩声音表示的附加基本辅助信息不包括冗余元素,从而使基本压缩声音表示的实际解码更高效地呈现。

[0037] 此方面的实施例可与前述方面的实施例相关。

[0038] 根据另一方面,描述了用于声音或声场的压缩声音表示的分层编码的编码器。该压缩声音表示可包括基本压缩声音表示,其包括多个分量。该多个分量可以是补充分量。压缩声音表示可进一步包括用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的基本辅助信息。该压缩声音表示可进一步包括增强辅助信息,其包括用于改善(例如,增强)基本重构声音表示的参数。该编码器可包括被配置成实施根据上文第一个提及的方面的及上文第二个提及的方面的方法的方法步骤的部分或全部的处理器。

[0039] 根据另一方面,描述了用于解码声音或声场的压缩声音表示的解码器。该压缩声音表示可已编码在多个分级层中。该多个分级层可包括基本层及一个或多个分级增强层。多个层可被分配有声音或声场的基本压缩声音表示的分量。换言之,该多个层可包括该基本压缩辅助信息的分量。这些分量可在各分量组中分配给相应层。该多个分量可以是补充分量。基本层可包括用于解码基本压缩声音表示的基本辅助信息。各层可包含包括用于改善(例如,增强)基本重构声音表示的参数的增强辅助信息的部分,该基本重构声音表示可从包括在该相应层及低于该相应层的任何层中的数据得到。该解码器可包括被配置成实施根据上文第三个提及的方面的及上文第四个提及的方面的方法的方法步骤的部分或全部的处理器。

[0040] 根据其他方面,方法、设备及系统涉及解码声音或声场的压缩高阶高保真度立体声响复制(HOA)声音表示。该设备可具有接收器,该接收器被配置成或该方法可接收包含对应于包括基本层及一个或多个分级增强层的多个分级层的压缩HOA表示的位流。该多个层被分配有声音或声场的基本压缩声音表示的分量,该分量在各分量组中被分配给各层。该设备可具有解码器,该解码器被配置成或该方法可基于与基本层关联的基本辅助信息以及基于与该一个或多个分级增强层关联的增强辅助信息来解码该压缩HOA表示。该基本辅助信息可包括与第一个体单声道信号有关的基本独立辅助信息,该第一个体单声道信号将独立于其他单声道信号被解码。该一个或多个分级增强层中的每一个可包括包含用于改善基本重构声音表示的参数的增强辅助信息的一部分,该基本重构声音表示可从包括在该相应层及低于该相应层的任何层中的数据得到。

[0041] 该基本独立辅助信息可指示该第一个体单声道信号代表具有入射方向的方向信号。该基本辅助信息可进一步包括与第二个体单声道信号有关的基本依赖性辅助信息,该第二个体单声道信号将依赖于其他单声道信号被解码。该基本依赖性辅助信息可包括在该

声场内方向性地分布的基于向量的信号,其中该方向性分布是由向量指定的。该向量的分量被设定成零且不是压缩向量表示的一部分。

[0042] 该基本压缩声音表示的分量可对应于代表HOA表示的系数序列或主要声音信号的单声道信号。该位流包括分别对应于多个分级层的数据有效载荷。该增强辅助信息可包括与下列中的至少一者有关的参数:空间预测、子带方向信号合成、及参数环境复制。该增强辅助信息可包括允许从方向信号预测该声音或声场的缺失部分的信息。对于每一层,可进一步确定该相应层是否已被有效地接收以及确定紧接在未被有效地接收的最低层以下的层的层索引。

[0043] 根据另一方面,描述了软件程序。该软件程序可适于在处理器上执行,且当在计算装置上实行时适于实施概述于本文件中的方法步骤的部分或全部。

[0044] 根据还另一方面,描述了存储介质。该存储介质可包含软件程序,该软件程序适于在处理器上执行且当在计算装置上实行时适于实施概述于本文件中的方法步骤的部分或全部。

[0045] 如技术人员将理解的,关于上述方面或其实施例中的任一个的描述也应用于各个其他方面或其实施例。为了简明,已省略了对于每一方面或实施例重复这样的叙述。

[0046] 包括如本文件所概述的它们的优选实施例的方法及设备可单独地使用,或者与此文件中公开的其他方法及系统组合使用。此外,在本文件中概述的方法及设备的所有方面可任意地组合。特别地,权利要求的特征可以任意方式彼此组合。

[0047] 方法步骤及设备特征可用许多方式互换。特别地,如技术人员将理解的,所公开的方法的细节可实现为适用于执行该方法的步骤的部分或全部的设备,反之亦然。

附图说明

[0048] 下文参考附图以示例性方式来解释本发明,其中:

[0049] 图1是示出根据本公开的实施例的分层编码方法的示例的流程图;

[0050] 图2是示意地示出根据本公开的实施例的编码器级的示例的框图;

[0051] 图3是示出根据本公开的实施例的解码已编码为多个分级层的声音或声场的压缩声音表示的方法的示例的流程图;

[0052] 图4A及图4B是示意地示出根据本公开的实施例的解码器级的示例的框图;

[0053] 图5是示意地示出根据本公开的实施例的编码器的硬件实现的示例的框图;及

[0054] 图6是示意地示出根据本公开的实施例的解码器的硬件实现的示例的框图。

具体实施方式

[0055] 首先,将描述根据本公开的方法及编码器/解码器可应用于的压缩声音(或声场)表示(为了简洁而在下文被称为压缩声音表示)。通常,完全压缩声音(或声场)表示(为了简洁而在下文被称为完全压缩声音表示)可包含下列三个分量(例如,由其组成):基本压缩声音(声场)表示(为了简洁而在下文被称为基本压缩声音表示)、基本辅助信息、以及增强辅助信息。

[0056] 基本压缩声音表示自身包含数个分量(例如,补充分量)(例如,由其组成)。基本压缩声音表示可占据完全压缩声音表示的特别的最大百分比。基本压缩声音表示可由代表原

始HOA表示的系数序列或主要声音信号的单声道传送信号组成。

[0057] 基本辅助信息是解码基本压缩声音表示所需的,并可假定其尺寸远小于基本压缩声音表示。其最大部分可由不相交部分构成,各不相交部分指定基本压缩声音表示的仅一个特定分量的解压缩。基本辅助信息可包含可被认为是独立基本辅助信息的第一部分以及可被认为是附加基本辅助信息的第二部分。

[0058] 第一及第二部分(独立基本辅助信息及附加基本辅助信息)二者可指定基本压缩声音表示的特定分量的解压缩。第二部分是可选的并可省略。在此情形中,压缩声音表示可被称为包含第一部分(例如,基本辅助信息)。

[0059] 第一部分(例如,基本辅助信息)可包含与其他(补充)分量无关地描述基本压缩声音表示的个体(补充)分量的辅助信息。特别地,第一部分(例如,基本辅助信息)可单独地指定多个分量中的一个或多个的解码,而与其他分量无关。因此,第一部分可称为独立基本辅助信息。

[0060] 第二(可选)部分可包含也被认为是附加基本辅助信息的辅助信息,可依赖于其他(补充)分量描述基本压缩声音表示的个体(补充)分量。此第二部分也可称为依赖性基本辅助信息。该依赖性可特别具有下列性质:

[0061] 一用于基本压缩声音表示的各个体(补充)分量的依赖性基本辅助信息可在基本压缩声音表示中不包含其他特定(补充)分量时最大程度地保持。

[0062] 一在将附加特定(补充)分量添加至基本压缩声音表示的情形中,用于所考虑的个体(补充)分量的依赖性基本辅助信息可变为原始依赖性基本辅助信息的子集,因此减少其尺寸。

[0063] 增强辅助信息也是可选的。其可用于改善或增强(例如,参数化地改善或增强)基本压缩声音表示。也可假设其尺寸远小于基本压缩声音表示的尺寸。

[0064] 因此,在实施例中,压缩声音表示可包含其包含多个分量的基本压缩声音表示、用于将基本压缩声音表示解码(例如,解压缩)为声音或声场的基本重构声音表示的基本辅助信息、及包括用于改善或增强(例如,参数化地改善或增强)基本重构声音表示的参数的增强辅助信息。压缩声音表示可进一步包含用于将基本压缩声音表示解码(例如,解压缩)为基本重构声音表示的附加基本辅助信息,其可包括指定多个分量中的一个或多个的依赖于各其他分量的解码的信息。

[0065] 此种种类的完全压缩声音表示的一个示例是由MPEG-H 3D音频标准(参考文件1)的初级版本、第12章及附件C.5所指定的压缩高阶高保真度立体声响复制(HOA)声场表示所给出的。即,压缩声音表示可对应于声音或声场的压缩HOA声音(或声场)表示。

[0066] 针对此示例,基本压缩声场表示(基本压缩声音表示)可包含数个分量(例如,可由数个分量标识)。该分量可是(例如,对应于)单声道信号。单声道信号可以是量化单声道信号。单声道信号可代表环境HOA声场分量的系数序列或主要声音信号。

[0067] 基本辅助信息可尤其对于这些单声道信号中的每一个描述其如何在空间上对声场做出贡献。例如,基本辅助信息可将主要声音信号指定为纯方向信号,意指具有特定入射方向的通用平面波。作为替代,基本辅助信息可将单声道信号指定为具有特定索引的原始HOA表示的系数序列。如上文所指示的,基本辅助信息可进一步分为第一部分及第二部分。

[0068] 第一部分是与特定个体单声道信号有关的辅助信息(例如,独立基本辅助信息)。

此独立基本辅助信息与其他单声道信号的存在无关。例如,此种辅助信息可指定单声道信号以代表具有特定入射方向的方向信号(例如,意指通用平面波)。作为替代地,可将单声道信号指定为具有特定索引的原始HOA表示的系数序列。第一部分可称为独立基本辅助信息。通常,第一部分(例如,基本辅助信息)可单独地指定多个单声道信号中的一个或多个的解码,而与其他单声道信号无关。

[0069] 第二部分是与特定个体单声道信号有关的辅助信息(例如,附加基本辅助信息)。此辅助信息依赖于其他单声道信号的存在。若将单声道信号指定成基于向量的信号(见,例如,参考文件1,第12.4.2.4.4节),则可使用此种辅助信息。这些信号在声场内方向性地分布,其中该方向性分布可由向量指定。在某一模式中(见例如,CodedVVecLength=1),此向量的特定分量被隐含地设定为零且不是压缩向量表示的一部分。这些分量是具有与原始HOA表示的系数序列的索引相等的索引的分量,且是基本压缩声音表示的一部分。这意味着若将向量的各分量编码,它们的总数可依赖于基本压缩声音表示。特别地,该总数可依赖于原始HOA表示所包含的系数序列。

[0070] 若在基本压缩声音表示中没有包含原始HOA表示的系数序列,用于各基于向量的信号的依赖性基本辅助信息由所有向量分量组成并具有其最大尺寸。在将具有某些索引的原始HOA表示的系数序列添加至基本压缩声音表示的情形中,将具有那些索引的向量分量从用于各基于向量的信号的辅助信息移除,从而减少用于基于向量的信号的依赖性基本辅助信息的尺寸。

[0071] 增强辅助信息(例如,增强辅助信息)可包含与(宽带)空间预测有关的参数(见参考文件1,第12.4.2.4.3节)和/或与子带方向信号合成及参数环境复制有关的参数。

[0072] 与(宽带)空间预测有关的参数可用于从方向信号(线性地)预测声场的缺失部分。

[0073] 子带方向信号合成及参数环境复制是最近通过修订而导入MPEG-H 3D音频标准中的压缩工具[见参考文件2,第1节]。这两个工具允许附加单声道信号的频率相关参数预测被空间分布,以补充空间上不完整或不足的压缩HOA表示。该预测可基于基本压缩声音表示的系数序列。

[0074] 重要地应指出,上文提及的对于声场的补充贡献在压缩HOA表示内不是由附加的量化信号表示,而是由具有相对较小尺寸的额外辅助信息表示。因此,所提及的这两个工具特别适合于HOA表示的以低数据率的压缩。

[0075] 具有上文提及的结构的一个或多个单声道信号的压缩表示的第二示例可包含用于直至某一高频的不相交的频带的编码频谱信息,其可被视为是基本压缩表示;指定(例如,由编码频带的数目及宽度指定)编码频谱信息的基本辅助信息;及包含频谱频带复制(SBR)的参数(例如,由其组成)的增强辅助信息,其描述了如何从基本压缩表示参数地重构用于在基本压缩表示中未考虑的较高频带的频谱信息。

[0076] 本公开提出了用于具有上文提及的结构完全压缩声音(或声场)表示的分层编解码方法。

[0077] 从对于连续时间区间提供压缩表示(为数据包或等同的帧有效载荷的形式)的意义上而言,该压缩可是基于帧的。时间区间可具有相等的或不同的尺寸。可假设这些数据包包含有效性标志、指示它们的尺寸的值、以及实际压缩表示数据。在下文中,不是作为限制,将假设压缩是基于帧的。另外,不是作为限制并且除非另外指示,将关注单一帧的处理,且

因此将省略帧索引。

[0078] 所考虑的完全压缩声音(或声场)表示的每个帧有效载荷被假定为包含J个数据包(或帧有效载荷),每个数据包(或帧有效载荷)用于基本压缩声音表示的一个分量,这些分量用 $BSRC_j$, $j=1, \dots, J$ 来表示。此外,它被假定为包含用 BSI_I 表示的具有独立基本辅助信息(基本辅助信息)的包,该包独立于其他分量指定基本压缩声音表示的特定分量 $BSRC_j$ 。可选地,它另外还被假定为包含用 BSI_D 表示的具有依赖性基本辅助信息(附加基本辅助信息)的包,该包依赖于其他分量指定基本压缩声音表示的特定分量 $BSRC_j$ 。

[0079] 两个数据包 BSI_I 和 BSI_D 中包含的信息可以可选地被分组到基本辅助信息的一个单个数据包 BSI 中。单个数据包 BSI 可尤其被称为包含J个部分,该J个部分中的每一个指定基本压缩声音表示中的一个特定分量 $BSRC_j$ 。这些部分中的每一个继而可被称为包含独立辅助信息的一部分,并且可选地包含依赖辅助信息的一部分。

[0080] 最后,其可包括由 ESI 表示的增强辅助信息有效载荷(增强辅助信息),其描述了如何改善或增强从完全基本压缩表示重构的声音(或声场)。

[0081] 所提出的分层编解码方案解决了使得既能够实现压缩部分(包括对数据包进行打包以用于传输)、又能够实现接收器和解压缩部分的所需步骤。下面将详细地描述每个部分。

[0082] 首先,将描述压缩及打包(例如,用于传输)。特别地,将描述在分层编解码的情形中的完全压缩声音(或声场)表示的分量及元素。

[0083] 图1示意地示出了用于压缩及打包的方法(例如,编码方法、或声音或声场的压缩声音表示的分层编码方法)的示例的流程图。个体有效载荷至基本层及(M-1)个增强层的分配(例如,分派)可通过传送层打包器完成。图2示意地示出了个体有效载荷的分配/分派的示例的框图。

[0084] 如上文所指示的,例如,完全压缩声音表示2100可涉及包含基本压缩声音表示的压缩HOA表示。完全压缩声音表示2100可包含多个分量(例如,单声道信号)2110-1, ... 2110-J、独立基本辅助信息(基本辅助信息)2120、可选的增强辅助信息(增强辅助信息)2140、及可选的依赖性基本辅助信息(附加基本辅助信息)2130。基本辅助信息2120可以是用于将基本压缩声音表示解码为声音或声场的基本重构声音表示的信息。基本辅助信息2120可包括如下信息,其与其他分量无关地单独指定一个或多个分量(例如,单声道信号)的解码。增强辅助信息2140可包括用于改善(例如,增强)基本重构声音表示的参数。附加基本辅助信息2130可以是用于将基本压缩声音表示解码为基本重构声音表示的(另外的)信息,并可包括指定多个分量中的一个或多个分量的依赖于各其他分量的解码的信息。

[0085] 图2示出了存在包括一个基本层(基本层)及一个或多个(分级)增强层的多个分级层的基本假设。例如,总共可有M个层,即,一个基本层及M-1个增强层。多个分级层具有递增的层索引。层索引的最低值(例如,层索引1)对应于基本层。进一步应理解,这些层从基本层、经由增强层、直至总体最高增强层(即,总体最高层)被排序。

[0086] 所提出的方法可在帧的基础上实施(即,逐帧地实施)。特别地,压缩声音表示2100可对于连续时间区间(例如,相等尺寸的时间区间)被压缩。各时间区间可与帧对应。以下描述的步骤可对各连续时间区间(例如,帧)实施。

[0087] 在图1中的S1010,将多个分量2110细分为多个分量组。然后将多个组中的每一个

分配(例如,添加或分派)给多个分级层中相应的一个层。其中,组的数目对应于层的数目。例如,组的数目可等于层的数目,使得每层有一个分量组。如上文所指示的,多个层可包括基本层及一个或多个(例如, $M-1$)分级增强层。

[0088] 换言之,将基本压缩声音表示细分为待分配给各个层的部分。不失一般性地,分组可由 $M+1$ 个 $J_m, m=0, \dots, M$ 描述,其中 $J_0=1$ 且 $J_M=J+1$,使得对于 $J_{m-1} \leq j < J_m$,分量 $BSRC_j$ 被分配给第 m 层。

[0089] 在S1020,将分量组分配给它们的相应层。在S1030,将基本辅助信息2120添加(例如,分派)至基本层(即,多个分级层中的最低层)。

[0090] 即,由于其的小尺寸,提出了将完全基本辅助信息(基本辅助信息及可选的附加基本辅助信息)包括至基本层中以避免其不必要的碎片化。

[0091] 若所考虑的压缩声音表示包含依赖性基本辅助信息(附加基本辅助信息),该方法可进一步包含(未示出于图1中)将附加基本辅助信息分解为附加基本辅助信息的多个部分2130-1, \dots , 2130-M中。然后,可将附加基本辅助信息的一部分添加(例如,分配)至基本层。换言之,可将附加基本辅助信息的一部分包括在基本层中。附加基本辅助信息的各部分可对应于相应层并可包括如下信息,该信息指定分配给该相应层的一个或多个分量的依赖于分配给该相应层及低于该相应层的任何层的其他分量的解码。

[0092] 因此,在独立基本辅助信息 BSI_I (基本辅助信息)2120对于分配保持不变的同时,依赖性基本辅助信息必需对于分层编解码被特殊处理,以另一方面允许在接收器侧正确解码,另一方面减少待传输的依赖性基本辅助信息的尺寸。提出将依赖性基本辅助信息分解为由 $BSI_{D,m}, m=1, \dots, M$ 指示的 M 个部分(部分),其中假定对于所考虑的压缩声音表示存在可选的依赖性基本辅助信息,第 m 部分包含用于分配给第 m 层的基本压缩声音表示的各分量 $BSRC_j, J_{m-1} \leq j < J_m$ 的依赖性基本辅助信息。在相应的依赖辅助信息不存在的情形中,对于压缩声音表示,假设部分 $BSI_{D,m}$ 是空的。依赖性基本辅助信息的各部分 $BSI_{D,m}$ 可依赖于包含在直至第 m 层的所有层(即,包含在所有层中 $j=1, \dots, m$)中的所有分量 $BSRC_j, 1 \leq j < J_m$ 。

[0093] 若独立基本辅助信息包 BSI_I 的尺寸小得可忽略,将其保持为整体并将其添加(分配)至基本层是合理的。可选地,也可对提供包 $BSI_{I,m}, m=1, \dots, M$ 的独立基本辅助信息实行与用于依赖性基本辅助信息的分解相似的分解。通过将独立基本辅助信息的一部分添加(分配)至具有基本压缩声音表示的对应分量的层,可有用地减少基本层的尺寸。

[0094] 在S1040,可确定增强辅助信息的多个部分2140-1, \dots , 2140-M。增强辅助信息的各部分可包括用于改善(例如,增强)可从包括在该相应层及低于该相应层的任何层中的数据得到的重构声音表示的参数。

[0095] 实施此步骤的原因是在分层编解码的情况下,重要的是意识到需要对于每个层额外地计算增强辅助信息,因为它的意图是增强初步解压缩的声音(或声场),不过这依赖于可用于解压缩的层。特别地,用于给定最高可解码层(最高可使用层)的初步解压缩声音(或声场)依赖于包括在最高可解码层及低于该最高可解码层的任何层中的分量。因此,压缩需要提供用 $ESI_m, m=1, \dots, M$ 指示的 M 个个体增强辅助信息数据包(增强辅助信息的一部分),其中,计算第 m 数据包 ESI_m 中的增强辅助信息以便增强从基本层和具有低于 m 的索引的增强层中包含的所有数据(例如,包含在第 m 层以及低于该底 m 层的所有层中的所有数据)获得的聲音(或声场)表示。

[0096] 在S1050,将增强辅助信息的多个部分2140-1, ..., 2140-M分配(例如,添加或分派)至多个层。将增强辅助信息的多个部分中的每一个分配给多个层中的相应一层。例如,多个层中的每一个包括增强辅助信息的相应部分。

[0097] 基本和/或增强辅助信息至相应层的分配可在由编码方法产生的配置信息中被指示。换言之,可将基本和/或增强辅助信息与各层的间的对应关系指示在配置信息中。另外,配置信息可为每一层指示分配给(例如,包括在)该层的基本压缩声音表示的分量。附加基本辅助信息的一部分被包括在基本层中,仍可对应于与基本层不同的层。

[0098] 总之,在压缩级,提供由FRAME指示的具有以下组成的帧数据包:

[0099] $FRAME = [BSRC_1 \dots BSRC_J \ BSI_1 \ BSI_{D,1} \dots BSU_{D,M} \ USI_1 \dots ESI_M]$ (1)

[0100] 另外,可将包 BSI_1 及 $BSI_{D,m}$,其中 $m=1, \dots, M$,组合成单一包 BSI ,在此情况下,由FRAME指示的帧数据包将具有以下组成:

[0101] $FRAME = [BSRC_1 \ BSRC_2 \dots BSRC_J \ BSI \ ESI_1 \ ESI_2 \dots ESI_M]$ (2)

[0102] 具有帧数据包的各有效载荷的次序通常可是任意的。

[0103] 然后,各数据包可在有效载荷内分组,其被定义为包含有效性标志、指示它们的尺寸的值以及实际压缩表示数据的特殊数据包。有效载荷的使用允许在接收器侧进行简单解复用,提供了能丢弃陈旧的有效载荷而无须对其进行解析的优点。一种可能的分组由以下给出:

[0104] 一将各 $BSRC_j$ 包, $j=1, \dots, J$ 分配(例如,分派)给标记为 \overline{BP}_j 的各有效载荷。

[0105] 一将第 m 个增强辅助信息数据包 ESI_m 及第 m 个依赖辅助信息数据包 $BSI_{D,m}$ 分配(例如,分派)给由 \overline{EP}_m , $m=1, \dots, M$ 指示的一个增强有效载荷。

[0106] 一将独立基本辅助信息 BSI_1 包分配给由 \overline{BSIP} 指示的单独的辅助信息有效载荷。

[0107] 可选地,若独立基本辅助信息的尺寸大,可将其分量 $BSI_{1,m}$, $m=1, \dots, M$ 中的每第 m 个分量分配(例如,分派)给增强有效载荷 \overline{EP}_m 。在此情形中,辅助信息有效载荷 \overline{BSIP} 是空的并可被忽略。

[0108] 另一选项是将所有依赖性基本辅助信息数据包 $BSI_{D,m}$ 分配至辅助信息有效载荷 \overline{BSIP} 中,这在依赖性基本辅助信息的尺寸小的情况下是合理的。

[0109] 最后,可提供由FRAME标记的具有以下组成的帧数据包:

[0110] $FRAME = [\overline{BP}_1 \dots \overline{BP}_J \ \overline{BSIP} \ \overline{EP}_1 \dots \overline{EP}_M]$ (3)

[0111] 具有帧数据包的各有效载荷的次序通常可是任意的。

[0112] 该方法可进一步包含(未显示于图1中)对于多个层中的每一个产生传送层包(例如,基本层包2200及 $M-1$ 个增强层包2300-1, ..., 2300-($M-1$)),其包括该相应层的数据(例如,用于基本层的分量、基本辅助信息及增强辅助信息、或用于该一个或多个增强层的分量及增强辅助信息)。

[0113] 用于不同层的传送层包可具有不同的传输优先度。因此,该方法可进一步包含(未显示于图1中)产生用于多个层的数据的传输的传送流,其中该基本层具有最高传输优先度,且分级增强层具有减小的传输优先度。其中,较高的传输优先度可对应于较大程度的错误保护,反之亦然。

[0114] 除非步骤需要某些其他步骤作为先决条件,否则前述步骤可以按任何次序执行,并且图1所示的示例性次序被理解为是非限制性的。

[0115] 图3示出了用于解码或解压缩(解包)的声音(或声场)的压缩声音表示的解码方法。对应接收器及解压缩级的示例被示意性地描绘在图4A及图4B的框图中。

[0116] 遵循上文,可将压缩声音表示编码在多个分级层中。多个层可被分配有(例如,可包括)基本压缩声音表示的分量,该分量在各分量组中被分配给各层。基本层可包括用于解码基本压缩声音表示的基本辅助信息。各层可包括其包括用于改善基本重构声音表示的参数的增强辅助信息的上文提及的部分中的一个,该基本重构声音表示可从包括在该相应层及低于该相应层的任何层中的数据得到。

[0117] 所提出的方法可在帧的基础上实施(亦即,以逐帧方式实施)。特别地,声音或声场的复原表示可针对连续时间区间(例如,尺寸相等的时间区间)产生。例如,时间区间可是帧。以下描述的步骤可对于各连续时间区间(例如,帧)实施。

[0118] 在S3010,接收对应于多个层的数据有效载荷(例如,传送层包)。数据有效载荷可作为包含声音或声场的压缩HOA表示的位流中的一部分被接收,该表示对应于多个分级层。该分级层包括基本层及一个或多个分级增强层。多个层被分配有声音或声场的基本压缩声音表示的分量。分量在各分量组中分配给各层。

[0119] 各层包可被复用以提供被接收的完全压缩声音表示的帧包。可由下式指示接收的帧包:

$$[0120] \quad [BSI_1 \quad BSI_{D,1} \quad \dots \quad BSI_{D,M} \quad ESI_1 \quad BSRC_1 \quad \dots \quad BSRC_{(J_1)-1} \quad \dots \quad ESI_M \quad BSRC_{J(M-1)} \quad \dots \quad BSRC_J] \quad (4)$$

[0121] 在将包 BSI_1 及 $BSI_{D,m}$,其中 $m=1, \dots, M$ 组合为单一包BSI的替代情况中,各层包可被复用以提供被接收的完全压缩声音表示的帧包,其由下式指示:

$$[0122] \quad [BSI \quad ESI_1 \quad BSRC_1 \quad \dots \quad BSRC_{(J_1)-1} \quad \dots \quad ESI_M \quad BSRC_{J(M-1)} \quad \dots \quad BSRC_J] \quad (5)$$

[0123] 就有效载荷而言,接收的帧包可由下式给出:

$$[0124] \quad \text{FRAME} = [\overline{BP}_1 \quad \dots \quad \overline{BP}_J \quad \overline{BSIP} \quad \overline{EP}_1 \quad \dots \quad \overline{EP}_M] \quad (6)$$

[0125] 然后,可将接收的帧包传至解压缩器或解码器4100。若单个层的传输已无错误,则至少所包含的增强辅助信息有效载荷 \overline{EP}_m (例如,对应于增强辅助信息的一部分)部分的有效性标志被设定成“真(true)”。在由于单个层的传输而导致错误的情况下,至少此层中的增强辅助信息有效载荷内的有效性标志被设定成“伪(false)”。因此,可从所包含的增强辅助信息有效载荷的有效性(例如,从其有效性标志)确定层包的有效性。

[0126] 在解压缩器4100中,可将接收的帧包解复用。针对此目的,可利用与各有效载荷的尺寸有关的信息以避免不必要地解析各有效载荷的数据。

[0127] 在S3020,指示最高层(例如,最高可使用层或最高可解码层)的第一层索引被多个层中确定以用于将基本压缩声音表示解码成声音或声场的基本重构声音表示。

[0128] 此外,在S3020,可选择将用于基本声音表示的解压缩的最高层(最高可使用层)的值(例如,层索引) N_B 。要实际用于基本声音表示的解压缩的最高增强层由 $N_B - 1$ 给定。因为各层正好包含一个增强辅助信息有效载荷(增强辅助信息的一部分),可基于增强辅助信息有效载荷确定包含层是否有效(例如,是否被有效地接收)。因此,该选择可使用所有的增强辅助

信息有效载荷 $ESI_m, m=1, \dots, M$ (或对应地, $\overline{EP}_m, m=1, \dots, M$)来完成。

[0129] 在S3030,得到基本重构声音表示。使用基本辅助信息(或一般地,使用基本辅助信息),可从分配给由第一层索引所指示的最高可使用层及低于此最高可使用层的任何层的分量得到基本重构声音表示。

[0130] 基本压缩声音表示分量 $BSRC_1, \dots, BSRC_J$ 的有效载荷可连同(全部)基本辅助信息有效载荷(例如, BSI 或 BSI_1 及 $BSI_{D,m}, m=1, \dots, M$)及值 N_B 一起被提供至基本表示解压缩处理单元4200。基本表示解压缩处理单元4200(描绘于图4A及4B中)仅使用包含在最低的 N_B 个层(即,基本层及 N_B-1 个增强层(即,直至由第一层索引所指示的层的层))内的那些基本压缩声音表示分量来重构基本声音(或声场)表示。作为替代地,可仅将包含在最低的 N_B 个层中的基本压缩声音表示分量的有效载荷连同相应的基本辅助信息有效载荷一起提供至基本表示解压缩处理单元4200。

[0131] 关于基本压缩声音(或声场)表示的哪些分量被包含在各层中的所需信息被假定由解压缩器4100已从具有配置信息的数据包得知,假设该数据包在帧数据包之前被发送及接收。

[0132] 为了提供依赖辅助信息数据包 $BSI_{D,m}, m=1, \dots, N_B$ 及增强辅助信息数据包 ESI_{N_E} ,可将所有增强有效载荷连同值 N_E 及值 N_B 一起输入至解压缩器4100的部分解析器4400(见图4B)。解析器可丢弃将不用于实际解压缩的所有有效载荷及数据包。若 N_E 的值等于零,则假定所有增强辅助信息数据包均是空的。

[0133] 若基本层包括对应于相应层的至少一个依赖性基本辅助信息有效载荷(附加基本辅助信息的一部分),各个依赖性基本辅助信息有效载荷(例如, $BSI_{D,m}, m=1, \dots, N_B$ (附加基本辅助信息的一部分))的解码可包括(i)通过参考被分配给其相应层及低于该相应层的任何层的分量来解码附加基本辅助信息的该部分(初步解码),及(ii)通过参考分配给该最高可使用层及在最高可使用层和该相应层之间的任何层的分量来校正附加基本辅助信息的该部分(校正)。其中,对应于相应层的附加基本辅助信息包括如下信息,该信息指定分配给该相应层的分量中的一个或多个分量的依赖于分配给该相应层及低于该相应层的任何层的其他分量的解码。

[0134] 然后,使用基本辅助信息及从对应于直至该最高可使用层的层的附加基本辅助信息的一部分得到的附加基本辅助信息的校正部分,可从分配给最高可使用层及低于该最高可使用层的任何层的分量得到(例如,产生)基本重构声音表示。

[0135] 特别地,各有效载荷 $BSI_{D,m}, m=1, \dots, N_B$ 的初步解码可涉及利用其对于包含于前 m 个层中的前 J_m-1 个基本压缩声音表示分量 $BSRC_1, \dots, BSRC_{(J_m)-1}$ 的依赖性,该依赖性在编码级被假定。

[0136] 各有效载荷 $BSI_{D,m}, m=1, \dots, N_B$ 的依次校正可涉及考虑基本声音分量最终被从包含在前 $N_B > m$ 个层中的前 $J_{N_B}-1$ 个基本压缩声音表示分量 $BSRC_1, \dots, BSRC_{(J_{N_B})-1}$ (其比假设用于初步解码的分量更多)重构。因此,校正可通过丢弃陈旧信息而完成,该丢弃是由于依赖性基本辅助信息的最初假设性质而成为可能,该最初假设性质为如果将某些补充分量添加至基本压缩声音表示,用于各个体(补充)分量的依赖性基本辅助信息变为原始信息的子集。

[0137] 在S3040,可确定第二层索引。第二层索引可指示应被用于改善(例如,增强)基本重构声音表示的增强辅助信息的一个或多个部分。

[0138] 除了第一层索引之外,还可确定待用于解压缩的增强辅助信息有效载荷(第二增强信息的部分)的索引(第二层索引) N_E 。第二层索引 N_E 可总是等于第一层索引 N_B 或等于零。增强可总是根据从最高可使用层得到的基本声音表示而完成,或根本不进行。

[0139] 在S3050,参考第二层索引从基本重构声音表示得到(例如,产生)声音或声场的重构声音表示。

[0140] 即,重构声音表示是通过(参数地)改善或增强基本重构声音表示(诸如,通过使用由第二层索引指示的增强辅助信息(增强辅助信息的部分))而得到。如下文进一步指示的,第二层索引可指示在此阶段根本不使用任何增强辅助信息。然后,重构声音表示会对应于基本重构声音表示。

[0141] 针对此目的,重构基本声音表示连同所有的增强辅助信息有效载荷 ESI_1, \dots, ESI_M 、基本辅助信息有效载荷(例如, BSI 或 BSI_I 及 $BSI_{D,m}, m=1, \dots, M$)、及值 N_E 一起被提供至增强表示解压缩处理单元4300(描绘于图4A及4B中),其仅使用增强辅助信息有效载荷 ESI_{N_E} 来计算最终增强声音(或声场)表示2100',并丢弃所有其他增强辅助信息有效载荷。作为替代,不是所有的增强辅助信息有效载荷,而是仅将增强辅助信息有效载荷 ESI_{N_E} 提供至增强表示解压缩处理单元4300。若 N_E 的值等于零,则丢弃所有增强辅助信息有效载荷(或替代地,不提供增强辅助信息有效载荷),且重构的最终增强声音表示2100'等于重构基本声音表示。增强辅助信息有效载荷 ESI_{N_E} 可已由部分解析器4400得到。

[0142] 图3也一般性地示出了基于与基本层关联的基本辅助信息并基于与一个或多个分级增强层关联的增强辅助信息来解码压缩HOA表示。

[0143] 除非步骤需要某些其他步骤作为先决条件,否则前述步骤可以按任何次序执行,并且图3所示的示例性次序被理解为是非限制性的。

[0144] 其次,将描述步骤S3020及S3040的用于解压缩的层选择(第一及第二层索引的选择)的细节。

[0145] 确定第一层索引可涉及为各层确定相应层是否已被有效地接收。确定第一层索引可进一步涉及将第一层索引确定为紧接在未被有效地接收的最低层之下的层的层索引。层是否已被有效地接收可通过评估该层的增强辅助信息有效载荷是否已被有效地接收而被确定。此继而可通过评估增强辅助信息有效载荷内的有效性标志而完成。

[0146] 确定该第二层索引通常可涉及确定该第二层索引等于该第一层索引,或将指示在获得重构声音表示时不使用任何增强辅助信息的索引值确定为第二层索引(例如,索引值0)。

[0147] 在所有帧数据包可被彼此独立地解压缩的情形中,可将要被实际用于基本声音表示的解压缩的最高层(最高可使用层)的编号 N_B 及要用于解压缩的增强辅助信息有效载荷的索引 N_E 设定成有效增强辅助信息有效载荷的最高数目 L ,其自身可通过评估增强辅助信息有效载荷内的有效性标志而确定。通过利用各增强辅助信息有效载荷的尺寸的知识,可避免为了确定有效载荷的有效性而对有效载荷的实际数据对复杂解析。

[0148] 即,若用于连续时间区间的压缩声音表示可被独立地解码,则可将第二层索引确

定成等于第一层索引。在此情形中,重构基本声音表示可基于最高可使用层的增强辅助信息有效载荷被增强。

[0149] 在使用具有帧间相关性的差分解压缩的情形中,必需另外考虑来自先前帧的决定。应指出,关于差分解压缩,独立帧数据包通常以规则的时间间隔被传输,以允许从这些时间点开始解压缩,其中值 N_B 及 N_E 的确定变为与帧无关,并如上文所述地实行。

[0150] 为了详细地解释所提出的依赖于帧的决定,将第 k 帧的有效增强辅助信息有效载荷的最高数目(例如,层索引)标记为 $L(k)$,将要被选择并用于基本声音表示的解压缩的最高层编号(例如,层索引)标记为 $N_B(k)$,并将要用于解压缩的增强辅助信息有效载荷的数目(例如,层索引)标记为 $N_E(k)$ 。

[0151] 由此,由 $N_B(k)$ 标记的要用于基本声音表示的解压缩的最高层编号可根据下式计算

$$[0152] \quad N_B(k) = \min(N_B(k-1), L(k)). \quad (7)$$

[0153] 通过选择 $N_B(k)$ 不大于 $N_B(k-1)$ 及 $L(k)$,确保基本声音表示的差分解压缩所需的所有信息是可获得的。

[0154] 即,若连续时间区间(例如,帧)的压缩声音表示不能相互独立地解码,确定第一层索引可包含对于每一层确定相应层是否已被有效地接收,并将对于给定时间区间的第一层索引确定为在该给定时间区间之前的时间区间的第一层索引和紧接在未被有效地接收的最低层之下的层的层索引中的较小者。

[0155] 要用于解压缩的增强辅助信息有效载荷的数目 $N_E(k)$ 可根据下式被确定:

$$[0156] \quad N_E(k) = \begin{cases} N_B(k) & \text{若 } N_B(k) = N_B(k-1) \\ 0 & \text{否则} \end{cases} \quad (8)$$

[0157] 其中,将 $N_E(k)$ 选择为0指示重构基本声音表示将不使用增强辅助信息被改善或增强。

[0158] 此特别意味着只要要用于基本声音表示的解压缩的最高层编号 $N_B(k)$ 不改变,就选择相同的对应增强层编号。然而,在 $N_B(k)$ 改变的情形中,通过将 $N_E(k)$ 设定为零而禁用增强。由于假定的增强辅助信息的差分解压缩,其的根据 $N_B(k)$ 的改变是不可能的,这是因为会需要在先前帧处的对应增强辅助信息层的解压缩,而其被假设为未实行。

[0159] 即,若用于连续时间区间(例如,帧)的压缩声音表示不能相互独立地解码,确定第二层索引可包含确定给定时间区间的第一层索引是否等于在先时间区间的第一层索引。若给定时间区间的第一层索引等于在先时间区间的第一层索引,可将给定时间区间的第二层索引确定(例如,选择)成等于给定时间区间的第一层索引。另一方面,若给定时间区间的第一层索引不等于在先时间区间的第一层索引,可将指示在得到重构声音表示时不使用任何增强辅助信息的索引值确定(例如,选择)为第二层索引。

[0160] 作为替代,若在解压缩时将具有多达 $N_E(k)$ 的数目的所有增强辅助信息有效载荷并行地解压缩,则式(4)中的选择规则可由下式替代:

$$[0161] \quad N_E(k) = N_B(k) \quad (9)$$

[0162] 最终指出,对于差分解压缩,最高使用层的编号 N_B 可仅在独立帧数据包时增加,然而可能在每个帧减少。

[0163] 应理解所提出的压缩声音表示的分层编码方法可由用于压缩声音表示的分层编

码的编码器来实现。此种编码器可包含适于实行上述各步骤的各单元。此种编码器5000的示例示意地描绘于图5中。例如,此种编码器5000可包含适用于实施上文提及的S1010的分量细分单元5010、适用于实施上文提及的S1020的分量分配单元5020、适用于实施上文提及的S1030的基本辅助信息分配单元5030、适用于实施上文提及的S1040的增强辅助信息分区单元5040、及适用于实施上文提及的S1050的增强辅助信息分配单元5050。还应理解,此种编码器的各单元可由计算装置的处理器5100具现,其适用于实施由所述各单元中的每一个实行的处理,即适用于实施上文提及的步骤的部分或全部,以及所提出的编码方法的任何进一步的步骤。编码器或计算装置可进一步包含可由处理器5100存取的存储器5200。

[0164] 应理解,所提出的对编码在多个分级层中的压缩声音表示进行解码的方法可由用于对编码在多个分级层中的压缩声音表示进行解码的解码器来实现。此种解码器可包含适用于实行上述各步骤的各单元。此种解码器6000的示例示意地描绘于图6中。例如,此种解码器6000可包含适用于实施上文提及的S3010的接收单元6010、适用于实施上文提及的S3020的第一层索引确定单元6020、适用于实施上文提及的S3030的基本重构单元6030、适用于实施上文提及的S3040的第二层索引确定单元6040、及适用于实施上文提及的S3050的增强重构单元6050。还应理解,此种解码器的各单元可由计算装置的处理器6100具现,其适用于实施由所述各单元中的每一个实行的处理,即适用于实行上文提及的步骤的部分或全部,以及所提出的解码方法的任何进一步的步骤。解码器或计算装置可进一步包含可由处理器6100存取的存储器6200。

[0165] 应注意到说明书及附图仅描述了所提出的方法及设备的原理。因此应理解,本领域的技术人员将能够构想出具现本发明的原理并包括在其精神及范围内的各种配置,即使它们未被显式地描述或示出于本文中。此外,原则上将本文陈述的所有示例明确地视为仅供教学目的使用,以协助阅读者理解由本发明人提供的所提出方法及设备的原理及概念以进一步发展本技术,并应被解释为不局限于这样的具体陈述的示例及条件。此外,在文中陈述了本发明的原理、实施方面、及实施例的叙述以及其特定示例意图涵盖其等同物。

[0166] 本文件中描述的方法及设备可实现为软件、固件、和/或硬件。某些组件可,例如,实现为在数字信号处理器或微处理器上运作的软件。其他组件可,例如,实现为硬件和/或专用集成电路。在所描述的方法及设备中遇到的信号可储存在介质中,诸如,随机存取存储器或光学存储介质。它们可经由网络,诸如,无线网络、卫星网络、无线网络或有线网络,例如,因特网,被传递。

[0167] 参考文件1:ISO/IEC JTC1/SC29/WG11 23008-3:2015(E)。Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3:3D audio,February 2015。

[0168] 参考文件2:ISO/IEC JTC1/SC29/WG11 23008-3:2015/PDAM3。Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3:3D audio,AMENDMENT 3:MPEG-H 3D Audio Phase 2,July 2015。

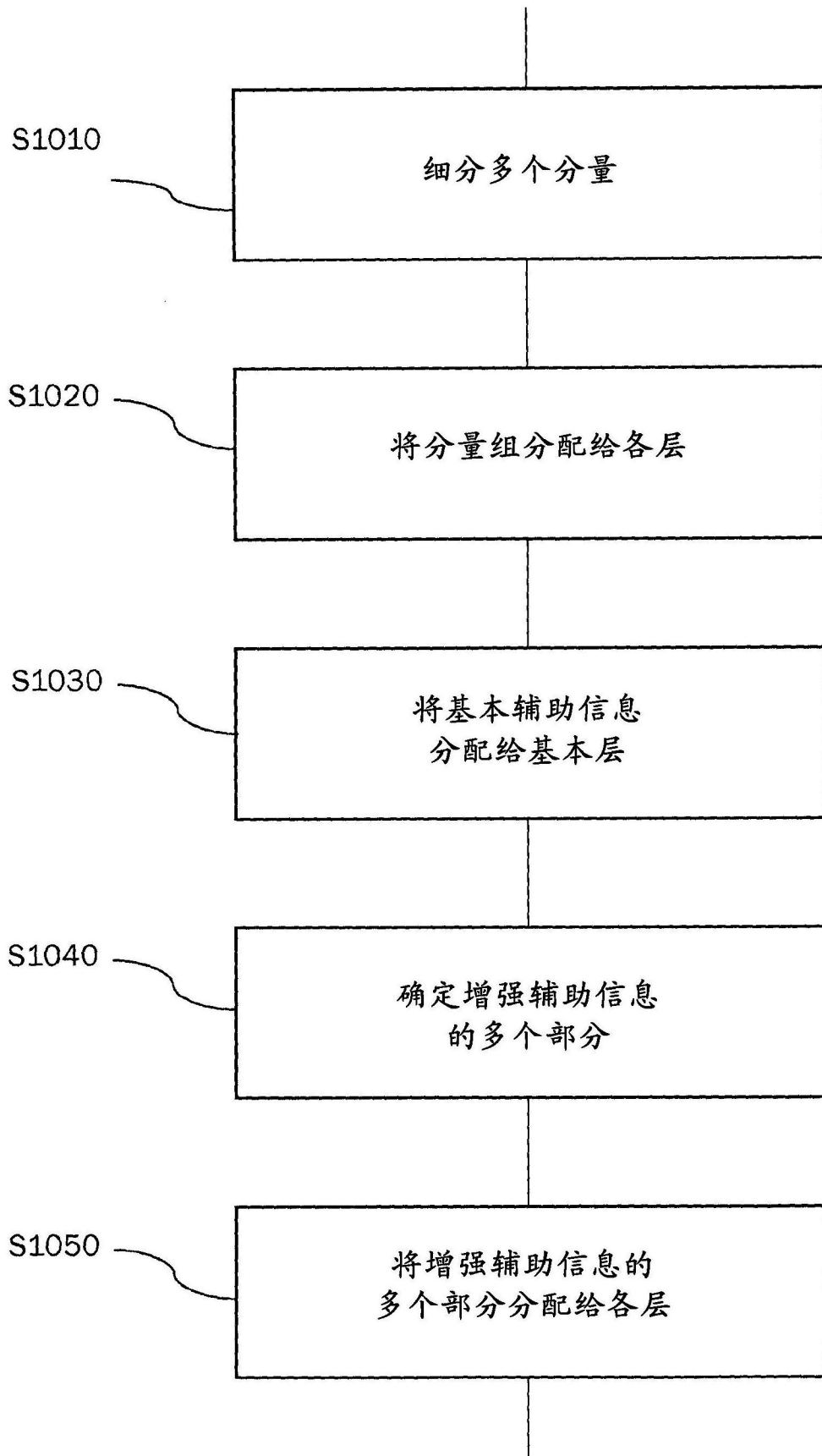


图1

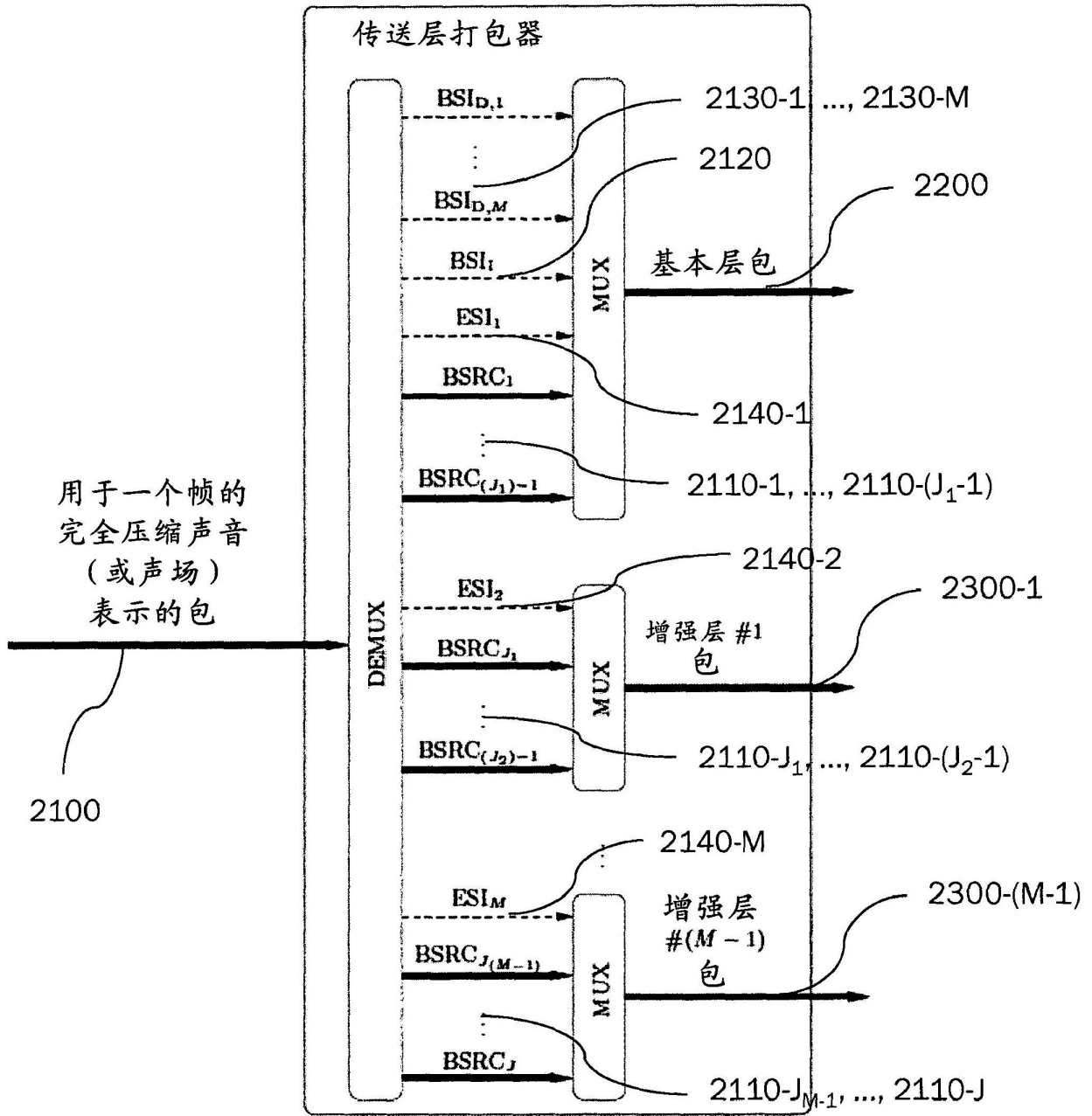


图2

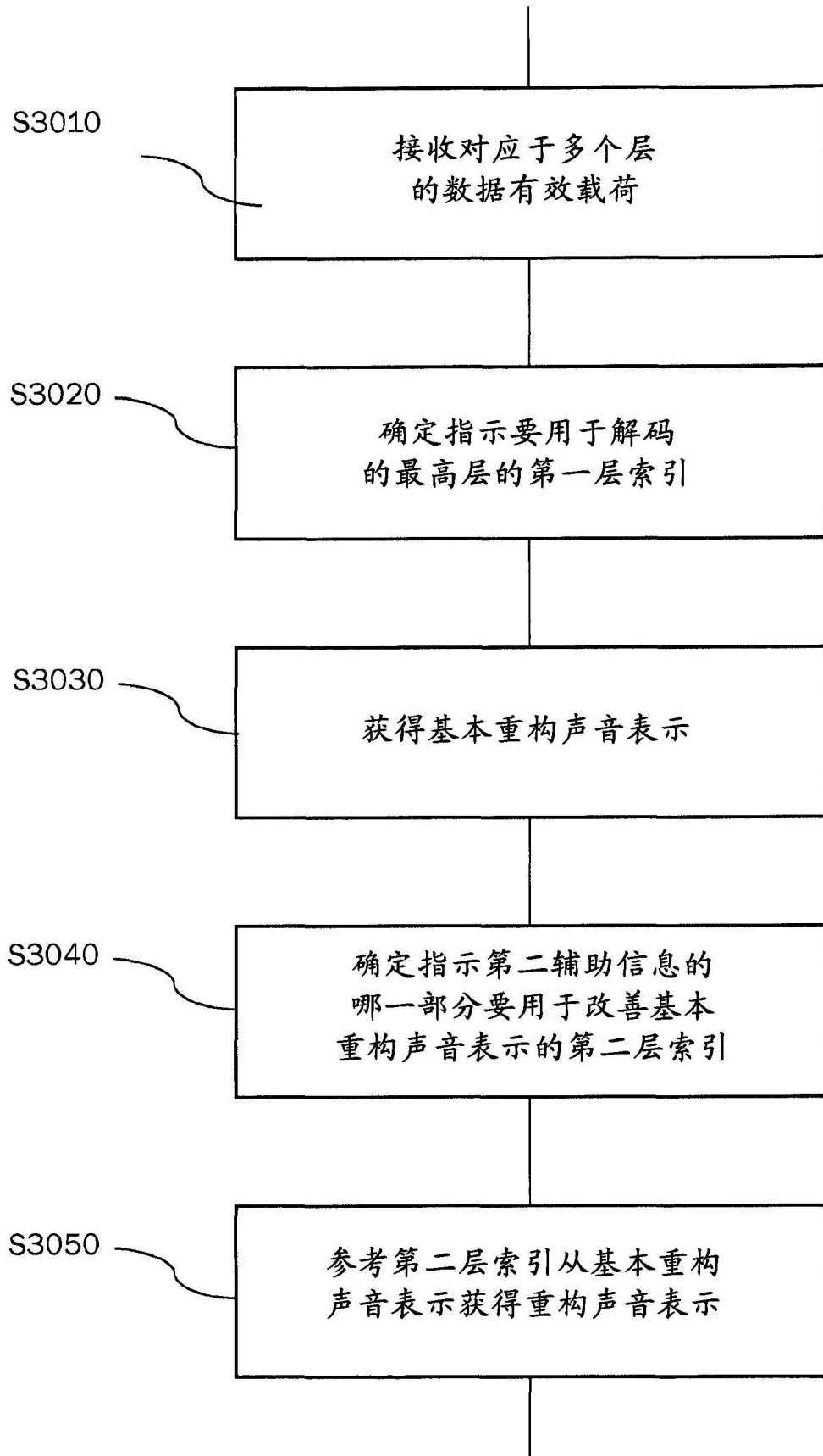


图3

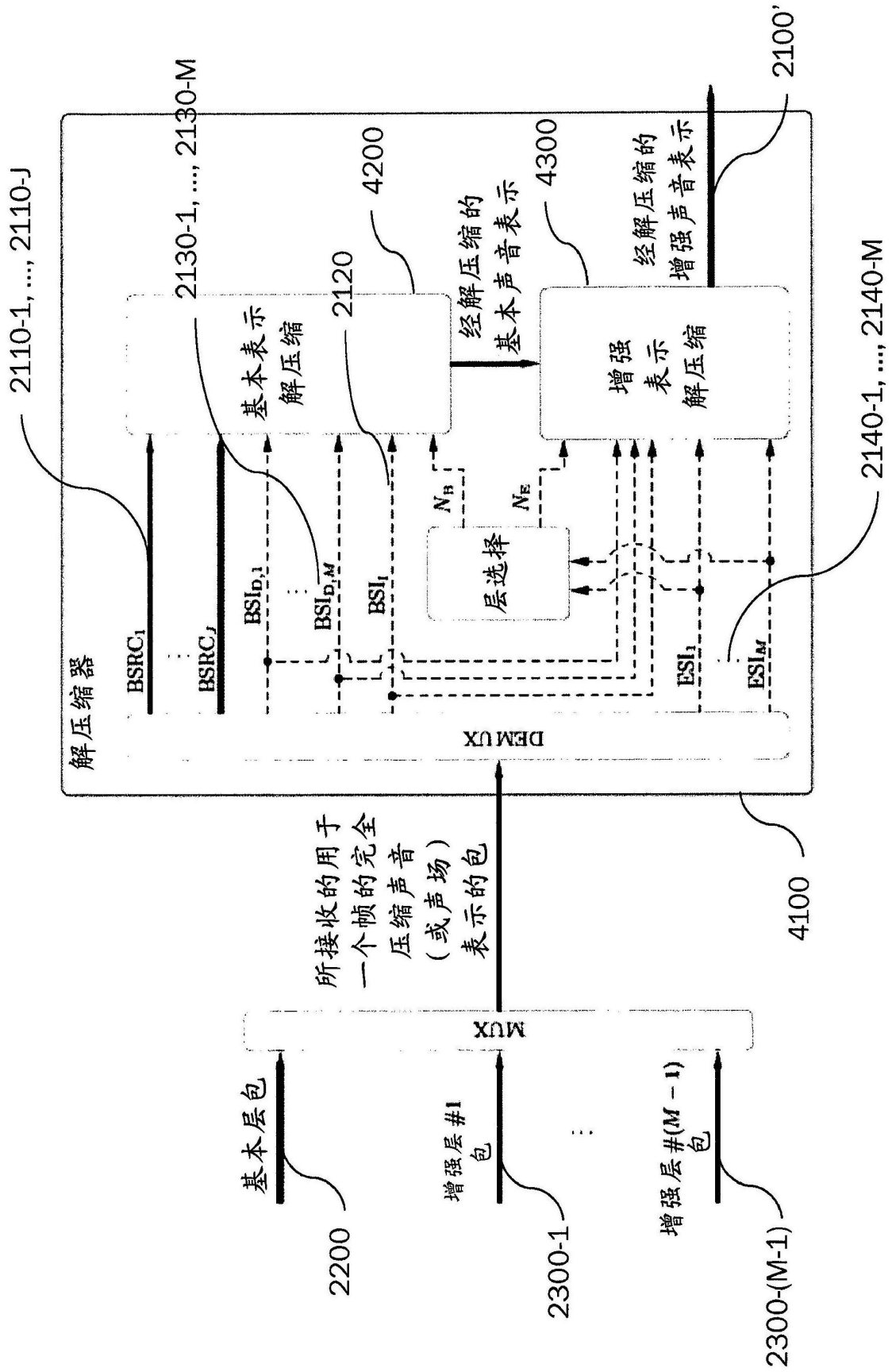


图4A

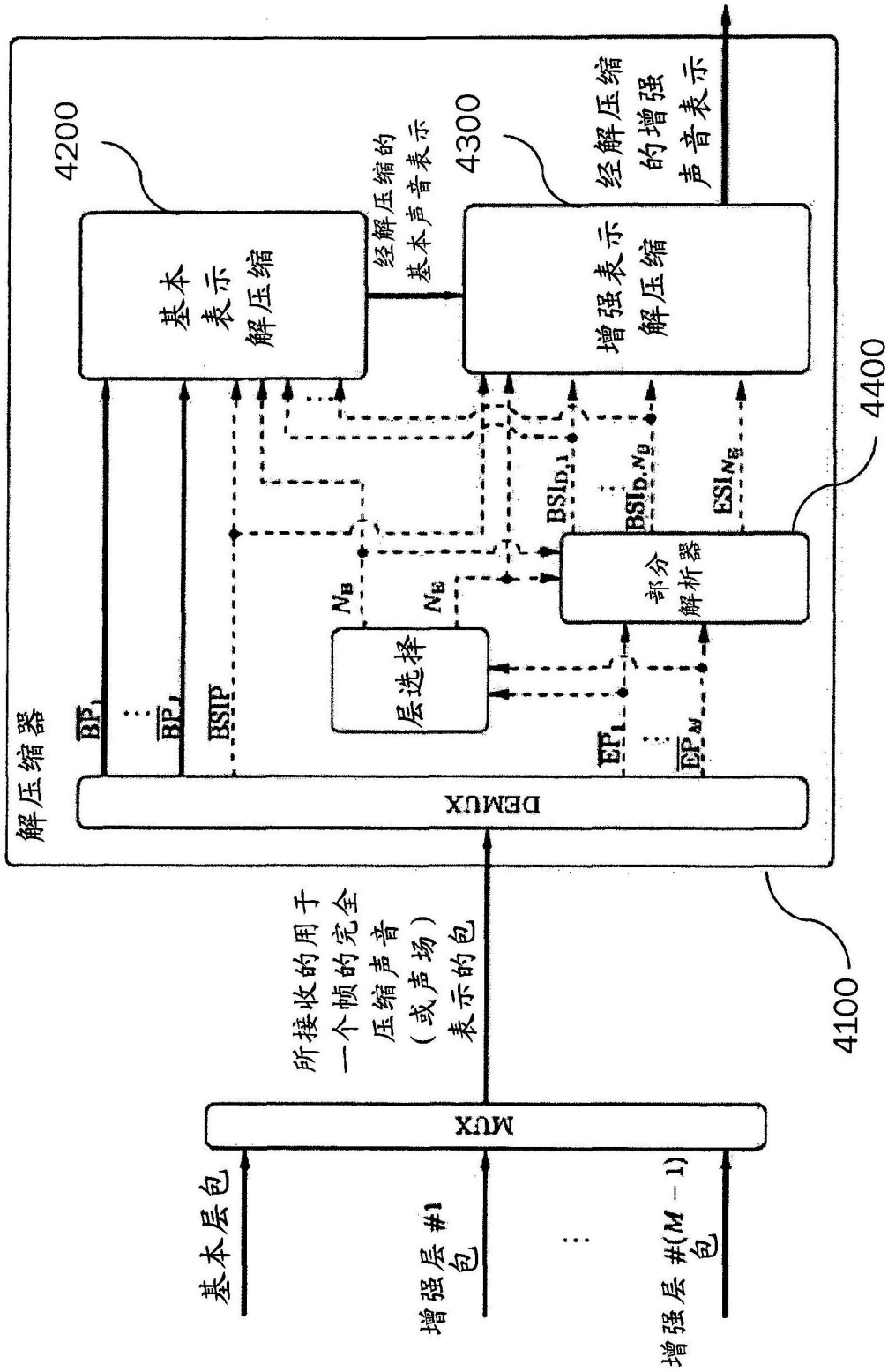


图4B

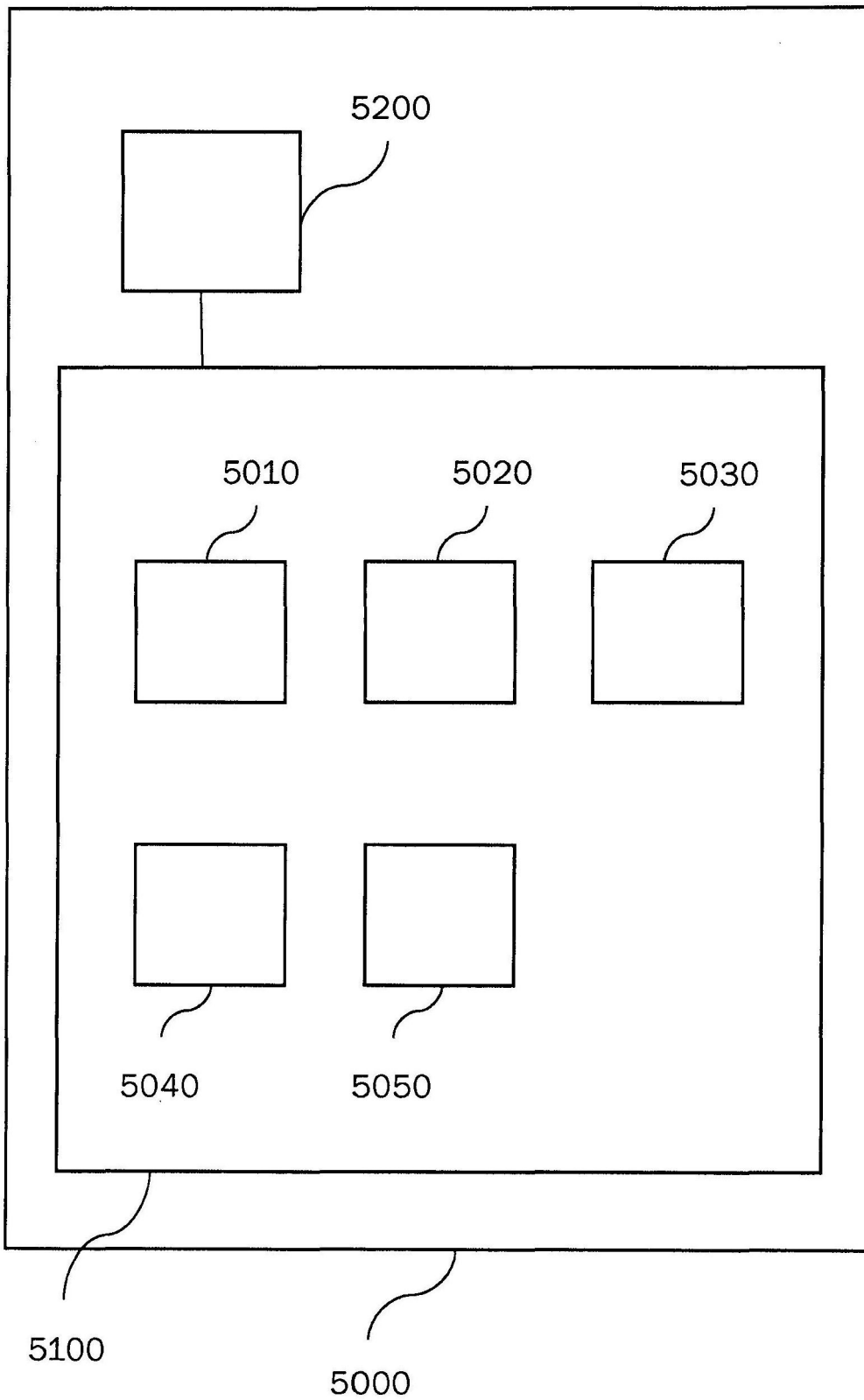


图5

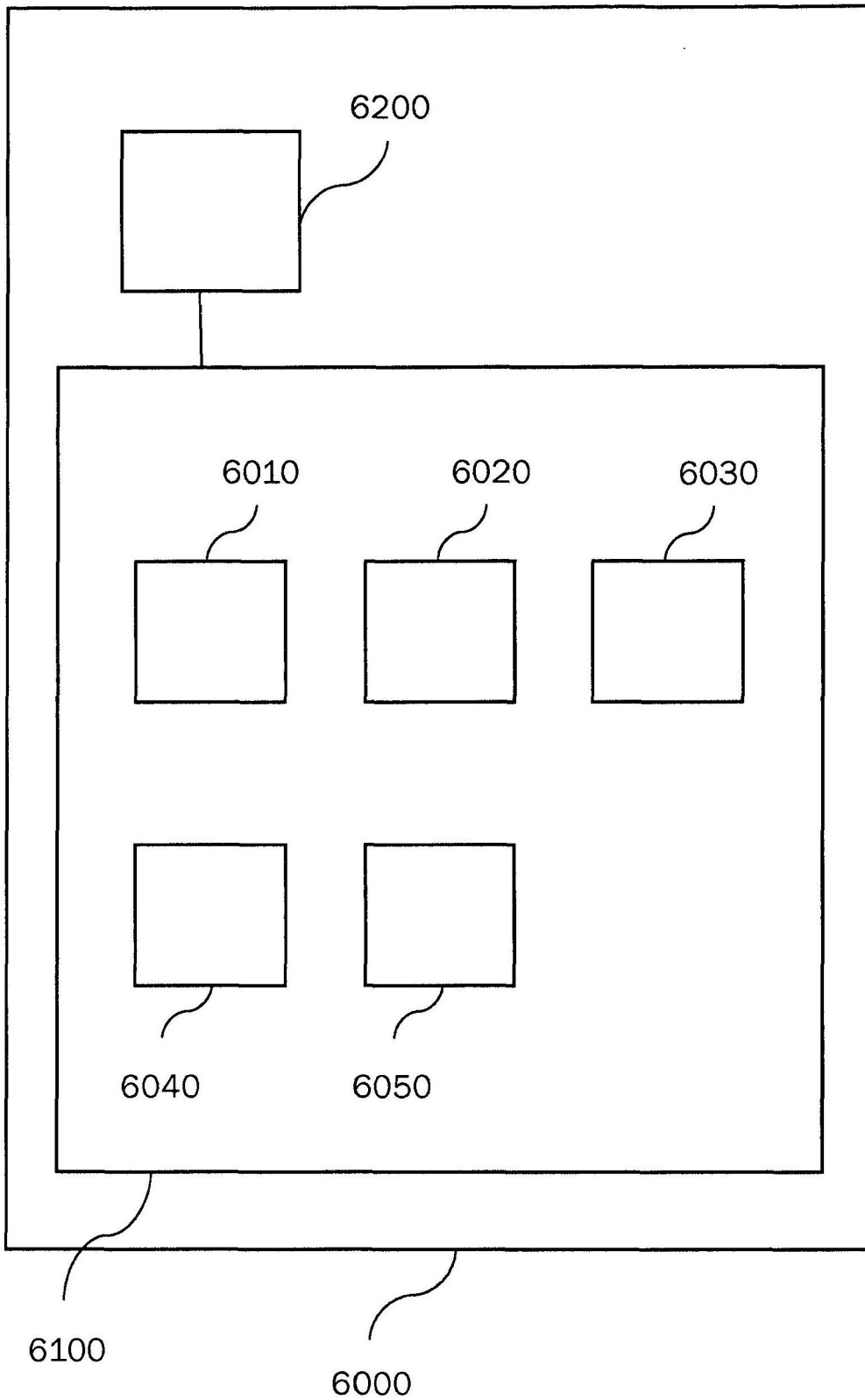


图6