



(12) 发明专利

(10) 授权公告号 CN 101379552 B

(45) 授权公告日 2013.06.19

(21) 申请号 200780004508.7

(22) 申请日 2007.02.07

(30) 优先权数据

60/765,747 2006.02.07 US

60/771,471 2006.02.09 US

60/773,337 2006.02.15 US

60/775,775 2006.02.23 US

60/781,750 2006.03.14 US

60/782,519 2006.03.16 US

60/792,329 2006.04.17 US

60/793,653 2006.04.21 US

(85) PCT申请进入国家阶段日

2008.08.05

(86) PCT申请的申请数据

PCT/KR2007/000676 2007.02.07

(87) PCT申请的公布数据

W02007/091849 EN 2007.08.16

(73) 专利权人 LG 电子株式会社

地址 韩国首尔

(72) 发明人 郑亮源 房熙锡 吴贤午 金东秀

林宰显

(74) 专利代理机构 上海专利商标事务所有限公
司 31100

代理人 侯颖嫒

(51) Int. Cl.

G10L 19/008(2013.01)

G10L 19/24(2013.01)

H04S 3/00(2006.01)

(56) 对比文件

US 2003/0035553 A1, 2003.02.20,

J.Breebaart et.al.MPEG Spatial Audio
Coding/MPEG Surround Overview and Current
Status.《Audio Engineering Society 119th
Convention》.2005,全文.

审查员 辛杰

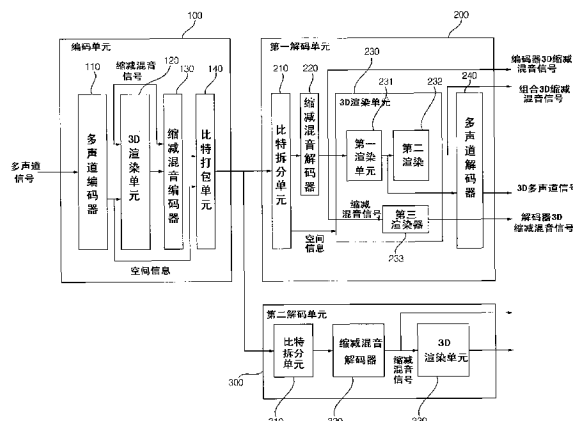
权利要求书2页 说明书21页 附图9页

(54) 发明名称

用于编码/解码信号的装置和方法

(57) 摘要

本发明提供了一种编码方法和装置以及一种解码方法和装置。该解码方法包括:从输入比特流提取任意缩减混音信号和补偿该任意缩减混音信号所必需的补偿信息;利用补偿信息补偿任意缩减混音信号;并通过对经补偿的任意缩减混音信号执行3D渲染操作来生成三维(3D)缩减混音信号。因此,高效率地编码具有3D效果的多声道信号,并根据音频再现环境的特性以最优音质自适应地恢复和再现音频信号是可能的。



1. 一种用于解码来自输入比特流的信号的方法,包括:

从所述输入比特流中提取缩减混音信号;

从所述输入比特流中提取空间信息和缩减混音信息,所述缩减混音信息指示所述缩减混音信号是否是任意缩减混音信号;

如果所述缩减混音信息指示所述缩减混音信号是任意缩减混音信号,则从所述输入比特流提取补偿所述任意缩减混音信号所必需的补偿信息,并且利用所述补偿信息补偿所述任意缩减混音信号;以及

通过使频域的经补偿的任意缩减混音信号、所述空间信息以及滤波器系数相乘来生成三维(3D)缩减混音信号,

其中:

所述任意缩减混音信号是除了由编码装置生成的缩减混音信号以外的缩减混音信号,

所述空间信息是在多声道信号被缩减混音成由编码装置生成的缩减混音信号时被确定的,所述补偿信息包括与由编码装置生成的缩减混音信号和所述任意缩减混音信号之差有关的信息。

2. 如权利要求1所述的方法,其特征在于,所述补偿信息包括与由编码装置生成的缩减混音信号的能级和所述任意缩减混音信号的能级之比有关的缩减混音增益信息。

3. 如权利要求1所述的方法,其特征在于,所述补偿信息以与关于多个声道的空间信息不同的分辨率来量化。

4. 如权利要求1所述的方法,其特征在于,对所述任意缩减混音信号的所述补偿包括:

将所述任意缩减混音信号从第一域转换至第二域;

利用所述补偿信息补偿所述第二域中的所述任意缩减混音信号;以及

将所述经补偿的任意缩减混音信号从所述第二域转换至所述第一域。

5. 如权利要求1所述的方法,其特征在于,所述生成3D缩减混音信号包括:

将所述经补偿的任意缩减混音信号转换至将要执行3D渲染操作的3D渲染域;以及

在所述3D渲染域中对所述经补偿的任意缩减混音信号执行3D渲染操作。

6. 一种用于解码来自输入比特流的信号的装置,包括:

比特拆分单元,其从所述输入比特流提取缩减混音信号、空间信息和缩减混音信息,所述缩减混音信息指示所述缩减混音信号是否是任意缩减混音信号,如果所述缩减混音信息指示所述缩减混音信号是任意缩减混音信号,则所述比特拆分单元从所述输入比特流提取补偿所述任意缩减混音信号所必需的补偿信息;

缩减混音补偿单元,其利用所述补偿信息补偿所述任意缩减混音信号;以及

3D渲染单元,其通过使频域的经补偿的任意缩减混音信号、所述空间信息以及滤波器系数相乘来生成3D缩减混音信号,

其中:

所述任意缩减混音信号是除了由编码装置生成的缩减混音信号以外的缩减混音信号,

所述空间信息是在多声道信号被缩减混音成由编码装置生成的缩减混音信号时被确定的,所述补偿信息包括与由编码装置生成的缩减混音信号和所述任意缩减混音信号之差有关的信息。

7. 如权利要求6所述的装置,其特征在于,所述补偿信息包括与由编码装置生成的缩

减混音信号的能级和所述任意缩减混音信号的能级之比有关的缩减混音增益信息。

8. 如权利要求 6 所述的装置,其特征在于,所述补偿信息以与关于多个声道的空间信息不同的分辨率来量化。

9. 如权利要求 6 所述的装置,其特征在于,所述缩减混音补偿单元包括:

第一域转换单元,用于将所述任意缩减混音信号从第一域转换至第二域;

补偿处理器,用于利用所述补偿信息补偿所述第二域中的所述任意缩减混音信号;以及

第二域转换单元,用于将所述经补偿的任意缩减混音信号从所述第二域转换至所述第一域。

10. 如权利要求 6 所述的装置,其特征在于,所述 3D 渲染单元将所述经补偿的任意缩减混音信号从第三域转换至第四域,在所述第四域中对所述经补偿的任意缩减混音信号执行 3D 渲染操作,并将通过所述 3D 渲染操作获得的信号从所述第四域转换至所述第三域。

用于编码 / 解码信号的装置和方法

技术领域

[0001] 本发明涉及编码 / 解码方法和编码 / 解码装置, 尤其涉及可处理音频信号使得能够产生三维 (3D) 音响效果的编码 / 解码装置, 以及利用该编码 / 解码装置的编码 / 解码方法。

背景技术

[0002] 编码装置将多声道信号缩减混音成具有较少声道的信号, 并将经缩减混音的信号传送到解码装置。然后, 解码装置从经缩减混音的信号恢复多声道信号, 并使用如 5.1 声道扬声器之类的三个或更多扬声器再现所恢复的多声道信号。

[0003] 多声道信号可由诸如耳机之类的 2 声道扬声器来再现。在这种情况下, 为了使用户觉得由 2 声道扬声器输出的声音好像从三个或更多声源再现的一样, 有必要开发出能够编码或解码多声道信号使得能够产生 3D 效果的三维 (3D) 处理技术。

发明内容

[0004] 技术问题

[0005] 本发明提供一种能够通过高效率地处理具有 3D 效果的信号而在各种再现环境中再现多声道信号的编码 / 解码装置以及编码 / 解码方法。

[0006] 技术解决方案

[0007] 根据本发明的一个方面, 提供了一种解码来自输入比特流的信号的解码方法, 该解码方法包括: 从输入比特流提取任意缩减混音信号和补偿该任意缩减混音信号所必需的补偿信息; 利用补偿信息补偿任意缩减混音信号; 并通过对所补偿的任意缩减混音信号执行 3D 渲染操作来生成三维 (3D) 缩减混音信号。

[0008] 根据本发明的另一个方面, 提供了一种解码来自输入比特流的信号的解码方法, 该解码方法包括: 从输入比特流提取任意缩减混音信号和补偿该任意缩减混音信号所需的补偿信息; 将该补偿信息组合至与将用于 3D 渲染操作的滤波器有关的滤波器信息; 并通过使用该组合获取的滤波器信息对任意缩减混音信号执行 3D 渲染操作来生成 3D 缩减混音信号。

[0009] 根据本发明的另一个方面, 提供了一种解码来自输入比特流的信号的解码装置, 该解码装置包括: 比特拆分单元, 其从输入比特流提取任意缩减混音信号和补偿该任意缩减混音信号所必需的补偿信息; 缩减混音补偿单元, 其利用补偿信息补偿任意缩减混音信号; 以及 3D 渲染单元, 其通过对经补偿的任意缩减混音信号执行 3D 渲染操作来生成 3D 缩减混音信号。

[0010] 根据本发明的另一个方面, 提供了一种具有用于执行上述的解码方法中的任一种的计算机程序的计算机可读记录介质。

[0011] 有益效果

[0012] 根据本发明, 可高效地编码具有 3D 效果的多声道信号, 并根据再现环境的特性以

最优音质自适应地恢复和再现音频信号。

[0013] 附图简要说明

[0014] 图 1 是根据本发明的实施例的编码 / 解码装置的框图；

[0015] 图 2 是根据本发明的实施例的编码装置的框图；

[0016] 图 3 是根据本发明的实施例的解码装置的框图；

[0017] 图 4 是根据本发明的另一个实施例的编码装置的框图；

[0018] 图 5 是根据本发明的另一个实施例的解码装置的框图；

[0019] 图 6 是根据本发明的另一个实施例的解码装置的框图；

[0020] 图 7 是根据本发明的实施例的三维 (3D) 渲染装置的框图；

[0021] 图 8 至 11 示出根据本发明的实施例的比特流；

[0022] 图 12 是根据本发明的实施例用于处理任意缩减混音信号的编码 / 解码装置的框图；

[0023] 图 13 是根据本发明的实施例的任意缩减混音信号补偿 / 3D 渲染单元的框图；

[0024] 图 14 是根据本发明的实施例用于处理兼容缩减混音信号的解码装置的框图；

[0025] 图 15 是根据本发明的实施例的缩减混音兼容性处理 / 3D 渲染单元的框图；以及

[0026] 图 16 是根据本发明的实施例用于消除串扰的解码装置的框图。

[0027] 本发明的最佳实施方式

[0028] 下文中将参考示出本发明的示例性实施例的附图更全面地描述本发明。图 1 是根据本发明的实施例的编码 / 解码装置的框图。参考图 1, 编码单元 100 包括多声道编码器 110、三维 (3D) 渲染单元 120、缩减混音编码器 130 以及比特打包单元 140

[0029] 多声道编码器 110 将具有多个声道的多声道信息缩减混音成诸如立体声或单声道信号之类的缩减混音信号, 并生成关于该多声道信号的声道空间信息。需要空间信息是为了从缩减混音信号恢复多声道信号。

[0030] 空间信息的例子包括: 指示一对声道的能级之差的声道电平差 (CLD)、声道预测系数 (CPC)——即用于基于 2 声道信号生成 3 声道信号的预测系数、指示一对声道之间的相关性的声道间相关性 (ICC) 以及一对声道之间的时间间隔的声道时间差 (CTD)。

[0031] 3D 渲染单元 120 基于缩减混音信号生成 3D 缩减混音信号。3D 缩减混音信号可以是具有三个或更多方向性的 2 声道信号, 因此可由诸如耳机之类的 2 声道扬声器再现并具有 3D 效果。换言之, 3D 缩减混音信号可由 2 声道扬声器再现, 使得用户觉得 3D 缩减混音信号好像是从具有三个或更多声道的声源再现的一样。声源的方向可基于分别输入到两耳的两个声音的强度之差、两个声音之间的时间间隔以及两个声音的相位之差中的至少一个来确定。因此, 3D 渲染单元 120 可基于人类如何利用其听觉确定声源的 3D 位置来将缩减混音信号转换成 3D 缩减混音信号。

[0032] 3D 渲染单元 120 可通过利用滤波器滤波缩减混音信号来生成 3D 缩减混音信号。在这种情况下, 可由外部源将滤波器相关信息——如滤波器系数输入到 3D 渲染单元 120。3D 渲染单元 120 可利用由多声道编码器 110 提供的空间信息来基于缩减混音信号生成 3D 缩减混音信号。更具体地, 3D 渲染单元 120 可通过利用空间信息将缩减混音信号转换成想像的多声道信号并滤波该想像的多声道信号来将缩减混音信号转换成 3D 缩减混音信号。

[0033] 3D 渲染单元 120 可通过利用报头相关传递函数 (HRTF) 滤波器滤波缩减混音信号

来生成 3D 缩减混音信号。

[0034] HRTF 是一种传递函数,它描述任意位置的声源和耳膜之间声波的传输,并返回根据声源的方向和高度变化的值。如果利用 HRTF 滤波没有方向性的信号,则可听到该信号好像从某方向再现一样。

[0035] 3D 渲染单元 120 可在例如离散傅立叶变换 (DFT) 域或快速傅立叶变换 (FFT) 域之类的频域中执行 3D 渲染操作。在这种情况下,3D 渲染单元 120 可在 3D 渲染操作前执行 DFT 或 FFT,或者可在 3D 渲染操作后执行逆 DFT (IDFT) 或逆 FFT (IFFT)。

[0036] 3D 渲染单元 120 可在正交镜像滤波器 (QMF)/混合域中执行 3D 渲染操作。在这种情况下,3D 渲染单元 120 可在 3D 渲染操作之前或之后执行 QMF/混合分析和合成操作。

[0037] 3D 渲染单元 120 可在时域中执行 3D 渲染操作。3D 渲染单元 120 可根据所需的音质和编码/解码装置的操作能力确定将在哪个域中执行 3D 渲染操作。

[0038] 缩减混音编码器 130 编码由多声道编码器 110 输出的缩减混音信号或由 3D 渲染单元 120 输出的 3D 缩减混音信号。缩减混音编码器 130 可利用诸如高级音频译码 (AAC) 法、MPEG 层 3 (MP3) 法或比特片算法译码 (BSAC) 法之类的音频编码法来编码由多声道编码器 110 输出的缩减混音信号或由 3D 渲染单元 120 输出的 3D 缩减混音信号。

[0039] 缩减混音编码器 130 可编码非 3D 缩减混音信号或 3D 缩减混音信号。在这种情况下,经编码的非 3D 缩减混音信号和经编码的 3D 缩减混音信号两者都可被包括在待传送的比特流中。

[0040] 比特打包单元 140 基于空间信息以及或者经编码的非 3D 缩减混音信号或者经编码的 3D 缩减混音信号来生成比特流。

[0041] 由比特打包单元 140 生成的比特流可包括空间信息、指示包括在比特流中的缩减混音信号是非 3D 缩减混音信号还是 3D 缩减混音信号的缩减混音标识信息、以及标识由 3D 渲染单元 120 所使用的滤波器的信息 (例如, HRTF 系数信息)。

[0042] 换言之,由比特打包单元 140 生成的比特流可包括还未经 3D 处理的非 3D 缩减混音信号以及通过由编码装置执行的 3D 处理操作获取的编码器 3D 缩减混音信号中的至少一个、以及标识包括在比特流中的缩减混音信号的类型的信息。

[0043] 可按用户选择或根据图 1 所示的编码/解码装置的能力和再现环境的特性确定非 3D 缩减混音信号和编码器 3D 缩减混音信号中的哪一个将被包括在由比特打包单元 140 生成的比特流中。

[0044] HRTF 系数信息可包括由 3D 渲染单元 120 所使用的 HRTF 的反函数的系数。HRTF 系数信息可仅包括由 3D 渲染单元 120 所使用的 HRTF 的系数的简短信息,例如, HRTF 系数的包络信息。如果将包括 HRTF 的反函数的系数的比特流传送到解码装置,则解码装置不需要执行 HRTF 系数转换操作,因此可减少解码装置的计算量。

[0045] 由比特打包单元 140 生成的比特流还可包括关于由基于 HRTF 的滤波导致的信号中的能量变化的信息,即,关于将要滤波的信号的能量与已滤波的信号的能量之差或将要滤波的信号的能量与已滤波的信号的能量之比的信息。

[0046] 由比特打包单元 140 生成的比特流还可包括指示其是否包括 HRTF 系数的信息。如果 HRTF 系数被包括在由比特打包单元 140 生成的比特流中,则该比特流还可包括指示其包括由 3D 渲染单元 120 所使用的 HRTF 的系数还是 HRTF 的反函数的系数的信息。

[0047] 参考图 1, 第一解码单元 200 包括比特拆分单元 210、缩减混音解码器 220、3D 渲染单元 230 以及多声道解码器 240。

[0048] 比特拆分单元 210 从编码单元 100 接收输入比特流, 并从该输入比特流中提取经编码的缩减混音信号和空间信息。缩减混音解码器 220 对经编码的缩减混音信号进行解码。缩减混音解码器 220 可利用诸如 AAC 法、MP3 法或 BSAC 法之类的音频信号解码方法来对经编码的缩减混音信号进行解码。

[0049] 如上所述, 从输入比特流中提取的经编码的缩减混音信号可以是经编码的非 3D 缩减混音信号或经编码的、编码器 3D 缩减混音信号。指示从输入比特流中提取的经编码的缩减混音信号是经编码的非 3D 缩减混音信号还是经编码的、编码器 3D 缩减混音信号的信息可被包括在输入比特流中。

[0050] 如果从输入比特流中提取的经编码的缩减混音信号是编码器 3D 缩减混音信号, 则经编码的缩减混音信号可在由缩减混音解码器 220 解码后容易地再现。

[0051] 另一方面, 如果从输入比特流中提取的经编码的缩减混音信号是非 3D 缩减混音信号, 则经编码的缩减混音信号可由缩减混音解码器 220 解码, 且通过解码获取的缩减混音信号可通过由第三渲染单元 233 执行的 3D 渲染操作转换成解码器 3D 缩减混音信号。解码器 3D 缩减混音信号可被容易地再现。

[0052] 3D 渲染单元 230 包括第一渲染器 231、第二渲染器 232 和第三渲染器 233。第一渲染器 231 通过对由缩减混音解码器 220 提供的编码器 3D 缩减混音信号执行 3D 渲染操作来生成缩减混音信号。例如, 第一渲染器 231 可通过从编码器 3D 缩减混音信号去除 3D 效果来生成非 3D 缩减混音信号。编码器 3D 缩减混音信号的 3D 效果可能不能由第一渲染器 231 完全去除。在这种情况下, 由第一渲染器 231 输出的缩减混音信号可具有相同的 3D 效果。

[0053] 第一渲染器 231 可将由缩减混音解码器 220 提供的 3D 缩减混音信号转换成使用编码单元 100 的 3D 渲染单元 120 所使用的滤波器的逆滤波器来将 3D 效果从其去除的缩减混音信号。关于由 3D 渲染单元 120 所使用的滤波器或由 3D 渲染单元 120 所使用的滤波器的逆滤波器的信息可被包括在输入比特流中。

[0054] 由 3D 渲染单元 120 所使用的滤波器可以是 HRTF 滤波器。在这种情况下, 由编码单元 100 所使用的 HRTF 的系数或 HRTF 的反函数的系数也可被包括在输入比特流中。如果由编码器单元 100 所使用的 HRTF 的系数被包括在输入比特流中, 则 HRTF 系数可被逆转换, 且可在由第一渲染器 231 执行的 3D 渲染操作期间使用该逆转换的结果。如果由编码单元 100 所使用的 HRTF 的反函数的系数被包括在输入比特流中, 则它们可在由第一渲染器 231 执行的 3D 渲染操作期间容易地使用, 而不进行任何逆转换操作。在这种情况下, 可减少第一解码装置 100 的计算量。

[0055] 输入比特流还可包括滤波器信息 (例如, 指示由编码单元 100 所使用的 HRTF 的系数是否被包括在输入比特流中的信息) 以及指示该滤波器信息是否已被逆转换的信息。

[0056] 多声道解码器 240 基于从其去除 3D 效果的缩减混音信号和从输入比特流中提取的空间信息来生成具有三个或更多声道的 3D 多声道信号。

[0057] 第二渲染器 232 可通过对从其去除 3D 效果的缩减混音信号执行 3D 渲染操作来生成具有 3D 效果的 3D 缩减混音信号。换言之, 第一渲染器 231 从由缩减混音解码器 220 提

供的编码器 3D 缩减混音信号中去除 3D 效果。之后,第二渲染器 232 可利用第一解码装置的滤波器,通过对由第一渲染器 231 执行去除而获取的缩减混音信号执行 3D 渲染操作来生成第一解码装置 200 所期望的具有 3D 效果的组合 3D 缩减混音信号。

[0058] 第一解码装置 200 可包括其中结合有执行相同操作的第一、第二和第三渲染器 231、232 和 233 中的两个或更多个的渲染器。

[0059] 由编码单元 100 生成的比特流可被输入到具有与第一解码装置 200 不同的结构的第二解码装置 300。第二解码装置 300 可基于包括在向其输入的比特流中的缩减混音信号生成 3D 缩减混音信号。

[0060] 更具体地,第二解码装置 300 包括比特拆分单元 310、缩减混音解码器 320 以及 3D 渲染单元 330。比特拆分单元 310 从编码单元 100 接收输入比特流,并从该输入比特流中提取经编码的缩减混音信号和空间信息。缩减混音解码器 320 对经编码的缩减混音信号进行解码。3D 渲染单元 330 对经解码的缩减混音信号执行 3D 渲染操作,使得经解码的缩减混音信号可被转换成 3D 缩减混音信号。

[0061] 图 2 是根据本发明的实施例的编码装置的框图。参考图 2,该编码装置包括渲染单元 400 和 420 以及多声道编码器 410。将省略与图 1 的实施例相同的编码过程的详细描述。

[0062] 参考图 2,可将 3D 渲染单元 400 和 420 分别设置在多声道编码器 410 的前面和后面。因此,多声道信号可由 3D 渲染单元 400 进行 3D 渲染,然后,经 3D 渲染的多声道信号可由多声道编码器 410 编码,从而生成经预处理的、编码器 3D 缩减混音信号。或者,多声道信号可由多声道编码器 410 进行缩减混音,然后,经缩减混音的信号可由 3D 渲染单元 420 进行 3D 渲染,从而生成经后处理的、编码器缩减混音信号。

[0063] 指示多声道信号在缩减混音前还是缩减混音后进行 3D 渲染的信息可被包括在待传送的比特流中。

[0064] 3D 渲染单元 400 和 420 两者都可被设置在多声道编码器 410 的前面或后面。

[0065] 图 3 是根据本发明的实施例的解码装置的框图。参考图 3,该解码装置包括 3D 渲染单元 430 和 450 以及多声道解码器 440。将省略与图 1 的实施例相同的解码过程的详细描述。

[0066] 参考图 3,可将 3D 渲染单元 430 和 450 分别设置在多声道解码器 440 的前面和后面。3D 渲染单元 430 可从编码器 3D 缩减混音信号去除 3D 效果,并将通过去除获取的缩减混音信号输入到多声道解码器 430。然后,多声道解码器 430 可解码向其输入的缩减混音信号,从而生成经预处理的 3D 多声道信号。或者,多声道解码器 430 可从经编码的 3D 缩减混音信号恢复多声道信号,且 3D 渲染单元 450 可从所恢复的多声道信号去除 3D 效果,从而生成经后处理的 3D 多声道信号。

[0067] 如果由编码装置提供的编码器 3D 缩减混音信号已经通过执行 3D 渲染操作和随后的缩减混音操作而生成,则编码器 3D 缩减混音信号可通过执行多声道解码操作和随后的 3D 渲染操作而被解码。另一方面,如果编码器 3D 缩减混音信号已经通过执行缩减混音操作和随后的 3D 渲染操作而生成,则编码器 3D 缩减混音信号可通过执行 3D 渲染操作和随后的多声道解码操作而被解码。

[0068] 可从编码装置传送的比特流中提取指示经编码的 3D 缩减混音信号是通过在缩减

混音操作前还是在缩减混音操作后执行 3D 渲染操作而获取的信息。

[0069] 3D 渲染单元 430 和 450 两者都可被设置在多声道解码器 440 的前面或后面。

[0070] 图 4 是根据本发明的另一个实施例的编码装置的框图。参考图 4, 编码装置包括多声道编码器 500、3D 渲染单元 510、缩减混音编码器 520 以及比特打包单元 530。将省略与图 1 的实施例相同的编码过程的详细描述。

[0071] 参考图 4, 多声道编码器 500 基于输入多声道信号生成缩减混音信号和空间信息。3D 渲染单元 510 通过对缩减混音信号执行 3D 渲染操作来生成 3D 缩减混音信号。

[0072] 可按用户选择或根据编码装置的能力、再现环境的特性或所要求的音质确定是否对缩减混音信号执行 3D 渲染操作。

[0073] 缩减混音编码器 520 编码由多声道编码器 500 生成的缩减混音信号或由 3D 渲染单元 510 生成的 3D 缩减混音信号。

[0074] 比特打包单元 530 基于空间信息以及或者经编码的缩减混音信号或者经编码的、编码器 3D 缩减混音信号生成比特流。由比特打包单元 530 生成的比特流可包括指示包括在比特流中的经编码的缩减混音信号是没有 3D 效果的非 3D 缩减混音信号还是具有 3D 效果的编码器 3D 缩减混音信号的缩减混音标识信息。更具体地, 缩减混音标识信息可指示由比特打包单元 530 生成的比特流是否包括非 3D 缩减混音信号、编码器 3D 缩减混音信号或这两者。

[0075] 图 5 是根据本发明的另一个实施例的解码装置的框图。参考图 5, 解码装置包括比特拆分单元 540、缩减混音解码器 550 以及 3D 渲染单元 560。将省略与图 1 的实施例相同的解码过程的详细描述。

[0076] 参考图 5, 比特拆分单元 540 从输入比特流中提取经编码的缩减混音信号、空间信息和缩减混音标识信息。缩减混音标识信息指示经编码的缩减混音信号是不具有 3D 效果的经编码的非 3D 缩减混音信号还是具有 3D 效果的经编码的 3D 缩减混音信号。

[0077] 如果输入比特流包括非 3D 缩减混音信号和 3D 缩减混音信号两者, 则仅可按用户选择或根据解码装置的能力、再现环境的特性或所需的音质来从输入比特流中提取非 3D 缩减混音信号和 3D 缩减混音信号之一。

[0078] 缩减混音解码器 550 对经编码的缩减混音信号进行解码。如果通过由缩减混音解码器 550 执行的解码获取的缩减混音信号是通过执行 3D 渲染操作获取的编码器 3D 缩减混音信号, 则该缩减混音信号可被容易地再现。

[0079] 另一方面, 如果通过由缩减混音解码器 550 执行的解码获取的缩减混音信号是不具有 3D 效果的缩减混音信号, 则 3D 渲染单元 560 可通过对由缩减混音解码器 550 执行解码所获取的缩减混音信号执行 3D 渲染操作来生成解码器 3D 缩减混音信号。

[0080] 图 6 是根据本发明的另一个实施例的解码装置的框图。参考图 6, 解码装置包括比特拆分单元 600、缩减混音解码器 610、第一 3D 渲染单元 620、第二 3D 渲染单元 630 以及滤波器信息存储单元 640。将省略与图 1 的实施例相同的解码过程的详细描述。

[0081] 比特拆分单元 600 从输入比特流中提取经编码的、编码器 3D 缩减混音信号和空间信息。缩减混音解码器 610 对经编码的、编码器 3D 缩减混音信号进行解码。

[0082] 第一 3D 渲染单元 620 使用用于执行 3D 渲染操作的编码装置的滤波器的逆滤波器, 从由缩减混音解码器 610 执行的解码获取的编码器 3D 缩减混音信号去除 3D 效果。第

二渲染单元 630 利用存储在解码装置中的滤波器通过对由第一 3D 渲染单元 620 执行去除而获取的缩减混音信号执行 3D 渲染操作生成具有 3D 效果的组合 3D 缩减混音信号。

[0083] 第二 3D 渲染单元 630 可利用其特性与用于执行 3D 渲染操作的编码单元的滤波器不同的滤波器执行 3D 渲染操作。例如,第二 3D 渲染单元 630 可利用其系数与编码装置所使用的 HRTF 的系数不同的 HRTF 执行 3D 渲染操作。

[0084] 滤波器信息存储单元 640 存储关于用于执行 3D 渲染的滤波器的滤波器信息,例如, HRTF 系数信息。第二 3D 渲染单元 630 可利用存储在滤波器信息存储单元 640 中的滤波器信息生成组合 3D 缩减混音。

[0085] 滤波器信息存储单元 640 可存储分别对应于多个滤波器的多条滤波器信息。在这种情况下,可按用户选择或根据解码装置的能力或所要求的音质来选择多条滤波器信息之一。

[0086] 来自不同种族的人可具有不同的耳结构。因此,针对不同个体优化的 HRTF 系数可彼此不同。图 6 中所示的解码装置可生成针对用户优化的 3D 缩减混音信号。此外,图 6 中所示的解码装置可不管由 3D 缩减混音信号供应商所提供的 HRTF 的类型是什么,而生成与用户所期望的 HRTF 滤波器相对应的具有 3D 效果的 3D 缩减混音信号。

[0087] 图 7 是根据本发明的实施例的 3D 渲染装置的框图。参考图 7,3D 渲染装置包括第一和第二域转换单元 700 和 720 以及 3D 渲染单元 710。为了在预定的域中执行 3D 渲染操作,可将第一和第二域转换单元 700 和 720 分别设置在 3D 渲染单元 710 的前面和后面。

[0088] 参考图 7,输入缩减混音信号可由第一域转换单元 700 转换成频域缩减混音信号。更具体地,第一域转换单元 700 可通过执行 DFT 或 FFT 将输入缩减混音信号转换成 DFT 域缩减混音信号或 FFT 域缩减混音信号。

[0089] 3D 渲染单元 710 通过将空间信息施加于由第一域转换单元 700 提供的频域缩减混音信号来生成多声道信号。之后,3D 渲染单元 710 通过滤波多声道信号来生成 3D 缩减混音信号。

[0090] 由 3D 渲染单元 710 生成的 3D 缩减混音信号由第二域转换单元 720 转换成时域 3D 缩减混音信号。更具体地,第二域转换单元 720 可对由 3D 渲染单元 710 生成的 3D 缩减混音信号执行 IDFT 或 IFFT。

[0091] 在频域 3D 缩减混音信号转换成时域 3D 缩减混音信号期间,可能发生混叠之类的数据丢失或数据失真。

[0092] 为了生成频域中的多声道信号和 3D 缩减混音信号,可将每一个参数频带的空间信息映射到频域,并可将多个滤波器系数转换至频域。

[0093] 3D 渲染单元 710 可通过使第一域转换单元 700 提供的频域缩减混音信号、空间信息和滤波器系数相乘来生成 3D 缩减混音信号。

[0094] 通过使全部在 M 点频域中表示的缩减混音信号、空间信息和多个滤波器系数相乘而获取的时域信号具有 M 个有效信号。为了在 M 点频域中表示缩减混音信号、空间信息和滤波器系数,可执行 M 点 DFT 或 M 点 FFT。

[0095] 有效信号是不一定具有 0 值的信号。例如,可通过经由采样从音频信号中获取 x 个信号来生成总共 x 个有效信号。在该 x 个有效信号中,y 个有效信号被补零。然后,有效信号的数目减少至 (x-y)。之后,具有 a 个有效信号的信号和具有 b 个有效信号的信号被卷

积,从而获取总共 $(a+b-1)$ 个有效信号。

[0096] M 点频域中缩减混音信号、空间信息和滤波器系数的相乘可提供与在时域中卷积缩减混音信号、空间信息和滤波器系数相同的效果。具有 $(3*M-2)$ 个有效信号的信号可通过将 M 点频域中的缩减混音信号、空间信息和滤波器系数转换至时域并卷积该转换的结果来生成。

[0097] 因此,通过使频域中的缩减混音信号、空间信息和滤波器系数相乘并将相乘的结果转换至时域而获取的信号中的有效信号的数目可能与通过卷积时域中的缩减混音信号、空间信息和滤波器系数获取的信号中的有效信号的数目不同。结果,在频域中的 3D 缩减混音信号转换至时域信号期间可发生混叠。

[0098] 为了防止混叠,时域中的缩减混音信号的有效信号的数目、映射到频域的空间信息的有效信号的数目、以及滤波器系数的数目的总和不能大于 M。映射到频域的空间信息的有效信号的数目可根据频域的点的数目确定。换言之,如果对每一参数频带表示的空间信息被映射至 N 点频域,则空间信息的有效信号的数目可以是 N。

[0099] 参考图 7,第一域转换单元 700 包括第一补零单元 701 和第一频域转换单元 702。第三渲染单元 710 包括映射单元 711、时域转换单元 712、第二补零单元 713、第二频域转换单元 714、多声道信号生成单元 715、第三补零单元 716、第三频域转换单元 717 以及 3D 缩减混音信号生成单元 718。

[0100] 第一补零单元 701 对时域中的具有 X 个样本的缩减混音信号执行补零操作使得缩减混音信号的样本数目可从 X 增加至 M。第一频域转换单元 702 将补零缩减混音信号转换成 M 点频域信号。经补零的缩减混音信号具有 M 个样本。在经补零的缩减混音信号的 M 个样本中,仅 X 个样本是有效信号。

[0101] 映射单元 711 将每一个参数频带的空间信息映射至 N 点频域。时域转换单元 712 将通过映射单元 711 执行的映射而获取的空间信息转换至时域。通过时域转换单元 712 执行转换而获取的空间信息具有 N 个样本。

[0102] 第二补零单元 713 对时域中的具有 N 个样本的空间信息执行补零操作,使得空间信息的样本数目可从 N 增加至 M。第二频域转换单元 714 将经补零的空间信息转换成 M 点频域信号。经补零的空间信息具有 N 个样本。在经补零的空间信息的 N 个样本中,仅 N 个样本是有效的。

[0103] 多声道信号生成单元 715 通过使第一频域转换单元 712 提供的缩减混音信号和第二频域转换单元 714 提供的空间信息相乘生成多声道信号。由多声道信号生成单元 715 生成的多声道信号具有 M 个有效信号。另一方面,通过在时域中卷积由第一频域转换单元 712 提供的缩减混音信号和由第二频域转换单元 714 提供的空间信息获取的多声道信号具有 $(X+N-1)$ 个有效信号。

[0104] 第三补零单元 716 可对时域中表示的 Y 滤波器系数执行补零操作,使得样本的数目可增加至 M。第三频域转换单元 717 将补零滤波器系数转换成 M 点频域。经补零的滤波器系数具有 M 个样本。在 M 个样本中,仅 Y 个样本是有效信号。

[0105] 3D 缩减混音信号生成单元 718 通过将多声道信号生成单元 715 所生成的多声道信号与第三频域转换单元 717 提供的多个滤波器系数相乘来生成 3D 缩减混音信号。由 3D 缩

减混音信号生成单元 718 生成的 3D 缩减混音信号具有 M 个有效信号。另一方面,通过在时域中卷积由多声道信号生成单元 715 所生成的多声道信号和由第三频域转换单元 717 所提供的滤波器系数来获取的 3D 缩减混音信号具有 $(X+N+Y-2)$ 个有效信号。

[0106] 通过将第一、第二和第三频域转换单元 702 所使用的 M 点频域设置成满足以下方程式: $M \geq (X+N+Y-2)$ 来防止混叠是可能的。换言之,有可能通过使第一、第二和第三频域转换单元 702、714 和 717 能够执行满足以下方程式: $M \geq (X+N+Y-2)$ 的 M 点 DFT 或 M 点 FFT 来防止混叠。

[0107] 至频域的转换可利用除 DFT 滤波器组、FFT 滤波器组和 QMF 组以外的滤波器组来执行。3D 缩减混音信号的生成可利用 HRTF 滤波器来执行。

[0108] 空间信息的有效信号的数目可利用除上述方法以外的方法来调节,或可利用上述方法中效率最高且需要计算量最少的一种方法来调节。

[0109] 不仅在信号、系数或空间信息从频域转换至时域或反过来的期间,而且在信号、系数或空间信息从 QMF 域转换至混合域或反过来的期间,都可发生混叠。上述防止混叠的方法也可用于在信号、系数或空间信息从 QMF 域转换至混合域或反过来的期间防止混叠发生。

[0110] 用于生成多声道信号或 3D 缩减混音信号的空间信息可改变。作为空间信息改变的结果,在输出信号中可发生作为噪声的信号不连续。

[0111] 输出信号中的噪声可利用平滑法来减少,通过该平滑法可防止空间信息快速变化。

[0112] 例如,当在第一帧和第二帧彼此相邻时施加于第一帧的第一空间信息与施加于第二帧的空间信息不同时,在第一帧和第二帧之间极可能发生不连续。

[0113] 在这种情况下,可利用第一空间信息补偿第二空间信息或利用第二空间信息补偿第一空间信息,使得第一空间信息和第二空间信息之差可减小,从而由第一和第二帧之间的不连续导致的噪声可降低。更具体地,第一空间信息和第二空间信息中的至少一个可用第一空间信息和第二空间信息的平均值来代替,从而降低噪声。

[0114] 噪声也很可能由于一对相邻参数频带之间的不连续而生成。例如,当在第一和第二参数频带彼此相邻时对应于第一参数频带的第三空间信息与对应于第二参数频带的第四空间信息不同时,不连续可能在第一和第二参数频带之间发生。

[0115] 在这种情况下,可利用第四空间信息补偿第三空间信息或利用第三空间信息补偿第四空间信息,使得第三空间信息和第四空间信息之差可减小,并且由第一和第二参数频带之间的不连续导致的噪声可降低。更具体地,第三空间信息和第四空间信息中的至少一个可用第三空间信息和第四空间信息的平均值来代替,从而降低噪声。

[0116] 由一对相邻帧之间或一对相邻参数频带之间的不连续导致的噪声可利用除上述方法以外的方法来降低。

[0117] 更具体地,可将每个帧乘以诸如汉宁 (Hanning) 窗口之类的窗口,并可将“重叠和添加”方案施加于乘法的结果,使得帧之间的改变可减小。或者,施加有多条空间信息的输出信号可被平滑,使得可防止输出信号的多个帧之间的改变。

[0118] 可利用例如 ICC 之类的空间信息将 DFT 域中的声道之间的解相关进行如下调节。

[0119] 可通过使输入到一至二 (OTT) 或二至三 (TTT) 框的信号的系数乘以预定值来调节

解相关的程度。预定值可由以下方程式来限定： $(A+(1-A*A)^{0.5*i})$ ，其中 A 指示施加于 OTT 或 TTT 框的预定频带的 ICC 值，而 i 指示虚部。虚部可以是正的或负的。

[0120] 预定值可附有根据信号的特性的加权因子，信号的特性例如信号的能级、每一频率的信号的能量特性或施加 ICC 值 A 的框的类型。作为引入加权因子的结果，可进一步调节解相关的程度，并可应用帧间平滑或内插法。

[0121] 如以上参考图 7 所描述的，可通过使用转换至频域的 HRTF 或报头相关脉冲响应 (HRIR) 在频域中生成 3D 缩减混音信号。

[0122] 或者，3D 缩减混音信号可通过在时域中卷积 HRIR 和缩减混音信号来生成。可将频域中生成的 3D 缩减混音信号留在频域中，而不进行逆域变换。

[0123] 为了在时域中卷积 HRIR 和缩减混音信号，可使用有限脉冲响应 (FIR) 滤波器或无限脉冲响应 (IIR) 滤波器。

[0124] 如上所述，根据本发明实施例的编码装置和解码装置可利用涉及使用频域中的 HRTF 或转换至频域的 HRIR 的第一方法、涉及在时域中卷积 HRIR 的第二方法或第一和第二方法的组合来生成 3D 缩减混音信号。

[0125] 图 8 至 11 示出根据本发明的实施例的比特流。

[0126] 参考图 8，比特流包括：包含生成多声道信号所需信息的多声道解码信息字段、包含生成 3D 缩减混音信号所需信息的 3D 渲染信息字段、以及包含利用包含在多声道解码信息字段中的信息和包含在 3D 渲染信息字段中的信息所需的报头信息的报头字段。比特流可仅包括多声道解码信息字段、3D 渲染信息字段和报头字段中的一个或两个。

[0127] 参考图 9，含有解码操作所必需的辅助信息的比特流可包括：包含整个经编码信号的报头信息的特定配置报头字段、以及包含关于多个帧的辅助信息的多个帧数据字段。更具体地，每一个帧数据字段可包括：包含相应帧的报头信息的帧报头字段、以及包含相应帧的空间信息的帧参数数据字段。或者，帧数据字段中的每一个仅可包括帧参数数据字段。

[0128] 帧参数数据字段中的每一个可包括多个模块，每个模块包括标志和参数数据。模块是包括诸如空间信息之类的参数数据以及诸如缩减混音增益和平滑数据之类提高信号音质所必需的其他数据的数据集。

[0129] 如果在没有任何附加标志的情况下接收关于由帧报头字段指定的信息的模块数据，如果由帧报头字段指定的信息被进一步分类，或者如果接收附加标志和数据连同未被帧报头指定的信息，则模块数据可不包括任何标志。

[0130] 关于 3D 缩减混音信号的辅助信息，例如 HRTF 系数信息，可被包括在特定配置报头字段、帧报头字段和帧参数数据字段中的至少一个中。

[0131] 参考图 10，比特流可包括：包含生成多声道信号所必需的信息的多个多声道解码信息字段、以及包含生成 3D 缩减混音信号所必需的信息的多个 3D 渲染信息字段。

[0132] 当接收比特流时，解码装置可使用多声道解码信息字段或 3D 渲染信息字段来执行解码操作并跳过任何未在解码操作中使用的多声道解码信息字段和 3D 渲染信息字段。在这种情况下，可根据要再现的信号的类型确定多声道解码信息字段和 3D 渲染信息字段中的哪些将被用于执行解码操作。

[0133] 换言之，为了生成多声道信号，解码装置可跳过 3D 渲染信息字段，并读取包含在多声道解码信息字段中的信息。另一方面，为了生成 3D 缩减混音信号，解码装置可跳过多

声道解码信息字段,并读取包含在 3D 渲染信息字段中的信息。

[0134] 跳过比特流中的多个字段中的某些的方法如下。

[0135] 首先,关于字段的比特大小的字段长度信息可被包括在比特流中。在这种情况下,可通过跳过对应于字段比特大小的比特数来跳过该字段。可将字段长度信息设置在字段的开头。

[0136] 第二,可将同步字设置在字段的末端或开头。在这种情况下,可通过基于同步字的位置定位字段来跳过该字段。

[0137] 第三,如果提前确定并固定了字段的长度,则可通过跳过对应于该字段的长度的数据量来跳过该字段。可将关于字段长度的固定字段长度信息包括在比特流中或存储在解码装置中。

[0138] 第四,可利用上述字段跳过方法中的两种或更多种的组合来跳过多个字段之一。

[0139] 诸如字段长度信息、同步字或固定字段长度信息之类的字段跳过信息是跳过字段所必需的信息,可将其包括在图 9 所示的特定配置报头字段、帧报头字段和帧参数数据字段之一中,或可将其包括在图 9 所示字段以外的字段中。

[0140] 例如,为了生成多声道信号,解码装置可参照设置在每个 3D 渲染信息字段的开头的字段长度信息、同步字或固定字段长度信息跳过 3D 渲染信息字段,并读取包括在多声道解码信息字段中的信息。

[0141] 另一方面,为了生成 3D 缩减混音信号,解码装置可参照设置在每个多声道解码信息字段的开头的字段长度信息、同步字或固定字段长度信息跳过多声道解码信息字段,并读取包括在 3D 渲染信息字段中的信息。

[0142] 比特流可包括指示包括在该比特流中的数据是生成多声道信号所必需的还是生成 3D 缩减混音信号所必需的信息。

[0143] 然而,即使比特流不包括诸如 CLD 之类的任何空间信息,而仅包括生成 3D 缩减混音信号所必需的数据(例如, HRTF 滤波器系数),也可通过利用生成 3D 缩减混音信号所必需的数据进行解码来再现多声道信号,而不需要空间信息。

[0144] 例如,从缩减混音信号获取作为关于两个声道的空间信息的立体声参数。然后,将立体声参数转换成关于待再现的多个声道的空间信息,并通过将由转换获取的空间信息施加于缩减混音信号来生成多声道信号。

[0145] 另一方面,即使比特流仅包括生成多声道信号所必需的数据,也可再现缩减混音信号而不需要附加解码操作,或可通过利用附加 HRTF 滤波器对缩减混音信号执行 3D 处理来再现 3D 缩减混音信号。

[0146] 如果比特流包括生成多声道信号所必需的数据和生成 3D 缩减混音信号所必需的数据两者,则可允许用户决定是再现多声道信号还是 3D 缩减混音信号。

[0147] 将在下文中参考各自相应的句法详细描述跳过数据的方法。

[0148] 句法 1 指示以帧为单位解码音频信号的方法。

[0149] [句法 1]

[0150]

SpatialFrame()
{
FramingInfo();
bsIndependencyFlag;
OttData();
TttData();
SmgData();
TempShapeData();
if[bsArbitraryDownmix){
ArbitraryDownmixData();
}
if(bsResidualCoding){
ResidualData();
}
}

[0151] 在句法1中,Ottdata()和TttData()是表示从缩减混音信号恢复多声道信号所必需的参数(诸如包括CLD、ICC和CPC的空间信息)的模块,而SmgData()、TempShapeData()、ArbitraryDownmixData()和ResidualData()是表示通过校正编码操作期间可能发生的信号失真来改善音质所必需的信息的模块。

[0152] 例如,如果诸如CLD、ICC或CPC之类的参数和包括在模块ArbitraryDownmixData()中的信息仅在解码操作期间使用,则设置在模块TttData()和ArbitraryDownmixData()之间的模块SmgData()和TempShapeData()是不必需的。因此,跳过模块SmgData()和TempShapeData()是有效率的。

[0153] 在下文中将参考以下的句法2详细描述根据本发明的实施例的跳过模块的方法。

[0154] [句法2]

[0155]

:
TttData() ;
SkipData() {
bsSkipBits ;
}
SmgData() ;
TempShapeData() ;
if[bsArbitraryDownmix) {
ArbitraryDownmixData() ;
}
:

[0156] 参考句法 2, 可将模块 SkipData() 设置在将被跳过的模块之前, 且将被跳过的模块的比特大小在模块 SkipData() 中被指定为 bsSkipBits。

[0157] 换言之, 假设模块 SmgData() 和 TempShapeData() 将被跳过, 且组合的模块 SmgData() 和 TempShapeData() 的比特大小是 150, 则可通过将 bsSkipBits 设置为 150 来跳过模块 SmgData() 和 TempShapeData()。

[0158] 在下文中将参考句法 3 详细描述根据本发明的另一个实施例的跳过模块的方法。

[0159] [句法 3]

[0160]

:
TttData() ;
bsSkipSyncflag ;
SmgData() ;
TempShapeData() ;
bsSkipSyncword ;

if[bsArbitraryDownmix) {
ArbitraryDownmixData() ;
}
:

- [0161] 参考图 3,可通过使用 bsSkipSyncflag 和 bsSkipSyncword 跳过不必要的模块,bsSkipSyncflag 是指示是否使用同步字的标志,bsSkipSyncword 是可被设置在将被跳过的模块的末端的同步字。
- [0162] 更具体地,如果将标志 bsSkipSyncflag 设置成使得同步字可使用,则标志 bsSkipSyncflag 和同步字 bsSkipSyncword 之间的一个或多个模块——即模块 SmgData() 和 TempShapeData() 可被跳过。
- [0163] 参考图 11,比特流可包括:包含再现多声道信号所必需的报头信息的多声道报头字段、包含再现 3D 缩减混音信号所必需的报头信息的 3D 渲染报头字段、以及包含再现多声道信号所必需数据的多个多声道解码信息字段。
- [0164] 为了再现多声道信号,解码装置可跳过 3D 渲染报头字段,并从多声道报头字段和多声道解码信息字段读取数据。
- [0165] 跳过 3D 渲染报头字段的方法与以上参考图 10 所述的字段跳过方法相同,因此,可跳过其详细描述。
- [0166] 为了再现 3D 缩减混音信号,解码装置可从多声道解码信息字段和 3D 渲染报头字段读取数据。例如,解码装置可利用包括在多声道解码信息字段中的缩减混音信号和包括在 3D 缩减混音信号中的 HRTF 系数信息来生成 3D 缩减混音信号。
- [0167] 图 12 是根据本发明的实施例的用于处理任意缩减混音信号的编码 / 解码装置的框图。参考图 12,任意缩减混音信号是除了由包括在编码装置 800 中的多声道编码器 801 生成的缩减混音信号以外的缩减混音信号。将省略与图 1 的实施例相同的过程的详细描述。
- [0168] 参考图 12,编码装置 800 包括多声道编码器 801、空间信息合成单元 802 以及比较单元 803。
- [0169] 多声道编码器 801 将输入多声道信号缩减混音成立体声或单声道缩减混音信号,并生成从缩减混音信号恢复多声道信号所必需的基本空间信息。
- [0170] 比较单元 803 将缩减混音信号与任意缩减混音信号进行比较,并基于比较的结果生成补偿信息。补偿信息是补偿任意缩减混音信号使得任意缩减混音信号可被转换成接近缩减混音信号所必需的。解码装置可利用补偿信息来补偿任意缩减混音信号,并利用经补偿的任意缩减混音信号恢复多声道信号。恢复的多声道信号比从由多声道编码器 801 生成的任意缩减混音信号恢复的多声道信号更类似于原始的输入多声道信号。
- [0171] 补偿信息可以是缩减混音信号和任意缩减混音信号之差。解码装置可通过将缩减混音信号和任意缩减混音信号之差与任意缩减混音信号相加来补偿任意声道缩减混音信号。

[0172] 缩减混音信号和任意缩减混音信号之差可以是指示缩减混音信号和任意缩减混音信号的能级之差的缩减混音增益。

[0173] 可针对每一频带、每一时间 / 时隙、和 / 或每一声道确定缩减混音增益。例如，一部分缩减混音增益可针对每一频带确定，而另一部分缩减混音增益可针对每一时隙确定。

[0174] 缩减混音增益可针对每一参数频带或为任意缩减混音信号优化的每一频带来确定。参数频带是施加有参数类型的空间信息的频率间隔。

[0175] 可将缩减混音信号和任意缩减混音信号的能级之差量化。量化缩减混音信号和任意缩减混音信号的能级之差的量化电平的分辨率可与量化缩减混音信号和任意缩减混音信号之间的 CLD 的量化电平的分辨率相同或不同。此外，缩减混音信号和任意缩减混音信号的能级之差的量化可涉及使用量化缩减混音信号和任意缩减混音信号之间的 CLD 的量化电平的全部或部分。

[0176] 因为缩减混音信号和任意缩减混音信号的能级之差的分辨率一般低于缩减混音信号和任意缩减混音信号之间的 CLD 的分辨率，所以与量化缩减混音信号和任意缩减混音信号之间的 CLD 的量化电平的分辨率相比，量化缩减混音信号和任意缩减混音信号的能级之差的量化电平的分辨率可具有微小的值。

[0177] 补偿任意缩减混音信号的补偿信息可以是包括残差信息的扩展信息，其指定不能利用任意缩减混音信号或缩减混音增益恢复的输入多声道信号的分量。解码装置可利用扩展信息来恢复不能利用任意缩减混音信号或缩减混音增益恢复的输入多声道信号的分量，从而恢复几乎不可与原始的输入多声道信号区分的信号。

[0178] 生成扩展信息的方法如下。

[0179] 多声道编码器 801 可生成与缩减混音信号所缺少的输入多声道信号的分量有关的信息作为第一扩展信息。解码装置可通过将第一扩展信息应用于利用缩减混音信号和基本空间信息生成多声道信号来恢复几乎不可与原始的输入多声道信号区分的信号。

[0180] 或者，多声道编码器 801 可利用缩减混音信号和基本空间信息来恢复多声道信号，并生成所恢复的多声道信号和原始输入的多声道信号之差作为第一扩展信息。

[0181] 比较单元 803 可生成与任意缩减混音信号所缺少的缩减混音信号的分量——即不能利用缩减混音增益补偿的缩减混音信号的分量——有关的信息作为第二扩展信息。解码装置可利用任意缩减混音信号和第二扩展信息恢复几乎不能与缩减混音信号区分的信号。

[0182] 除上述方法外，扩展信息还可利用各种残差译码方法来生成。

[0183] 缩减混音增益和扩展信息两者都可被用作补偿信息。更具体地，可针对缩减混音信号的整个频带获取缩减混音增益和扩展信息两者，并可将它们一起用作补偿信息。或者，可将缩减混音增益用作针对缩减混音信号的一部分频带的补偿信息，并将扩展信息用作缩减混音信号的另一部分频带的补偿信息。例如，可将扩展信息用作缩减混音信号的低频带的补偿信息，并将缩减混音增益用作缩减混音信号的高频带的补偿信息。

[0184] 还可将与除缩减混音信号的低频带以外的、诸如显著影响音质的峰值或凹口之类的部分缩减混音信号有关的扩展信息用作补偿信息。

[0185] 空间信息合成单元 802 合成基本空间信息（例如，CLD、CPC、ICC 和 CTD）和补偿信息，从而生成空间信息。换言之，传送至解码装置的空间信息可包括基本空间信息、缩减混音增益及第一和第二扩展信息。

[0186] 空间信息可连同任意缩减混音信号一起被包括在比特流中,且可将比特流传送至解码装置。

[0187] 扩展信息和任意缩减混音信号可利用诸如 AAC 法、MP3 法或 BSAC 法之类的音频编码方法来编码。扩展信息和任意缩减混音信号可利用相同的音频编码方法或不同的音频编码方法来编码。

[0188] 如果扩展信息和任意缩减混音信号利用相同的音频编码方法编码,则解码装置可利用单一音频解码方法来解码扩展信息和任意缩减混音信号两者。在这种情况下,因为任意缩减混音信号总是能够被解码,所以扩展信息也总是能够被解码。然而,因为任意缩减混音信号一般被输入至解码装置作为脉冲编码调制 (PCM) 信号,用于编码任意缩减混音信号的音频编解码器的类型可能不能容易地识别,因此,用于编码扩展信息的音频编解码器的类型可能也不能容易地识别。

[0189] 因此,与用于编码任意缩减混音信号和扩展信息的音频编解码器的类型有关的音频编解码器信息可被插入比特流中。

[0190] 更具体地,可将音频编解码器信息插入比特流的特定配置报头字段。在这种情况下,解码装置可从比特流的特定配置报头字段提取音频编解码器信息,并使用所提取的音频编解码器信息解码任意缩减混音信号和扩展信息。

[0191] 另一方面,如果任意缩减混音信号和扩展信息利用不同的编码方法来编码,则扩展信息可能不能够被解码。在这种情况下,因为不能识别扩展信息的末端,所以不能执行进一步的解码操作。

[0192] 为了解决该问题,可将与分别用于编码任意缩减混音信号和扩展信息的音频编解码器的类型有关的音频编解码器信息插入比特流的特定配置报头字段。然后,解码装置可从比特流的特定配置报头字段读取音频编解码器信息,并使用读取信息来解码扩展信息。如果解码装置不包括任何可解码扩展信息的解码单元,则可能不能进一步进行扩展信息的解码,且可读取紧接扩展信息之后的信息。

[0193] 与用于编码扩展信息的音频编解码器的类型有关的音频编解码器信息可由包括在比特流的特定配置报头字段中的句法元素表示。例如,音频编解码信息可由 4 比特句法元素 bsResidualCodecType 来表示,如以下的表 1 中指示的。

[0194] 表 1

	bsResidualCodecType	编解码器
[0195]	0	AAC
	1	MP3
	2	BSAC
	3...15	保留

[0196] 扩展信息不仅可包括残差信息,还可包括声道扩充信息。声道扩充信息是将通过利用空间信息解码获取的多声道信号扩充成具有更多声道的多声道信号所必需的信息。例如,声道扩充信息可以是 5.1 声道信号或 7.1 声道信号扩充成 9.1 声道信号所必需的信息。

[0197] 可将扩展信息包括在比特流中,并可将比特流传送至解码装置。然后,解码装置可补偿缩减混音信号,或利用扩展信息扩充多声道信号。然而,解码装置可跳过扩展信息,而不是从比特流中提取扩展信息。例如,在利用包括在比特流中的 3D 缩减混音信号生成多声道信号或利用包括在比特流中的缩减混音信号生成 3D 缩减混音信号的情况下,解码装置可跳过扩展信息。

[0198] 跳过包括在比特流中的扩展信息的方法可与以上参考图 10 所述的字段跳过方法之一相同。

[0199] 例如,扩展信息可利用附连到包括扩展信息的比特流的开头且指示扩展信息的比特大小的比特大小信息、附连到包括扩展信息的字段的开头或末端的同步字、以及指示扩展信息的固定比特大小的固定比特性大小信息中的至少一个来跳过。可将比特大小信息、同步字和固定比特大小信息全部包括在比特流中。还可将固定比特大小信息存储在解码装置中。

[0200] 参考图 12,解码单元 810 包括缩减混音补偿单元 811、3D 渲染单元 815 以及多声道解码器 816。

[0201] 缩减混音补偿单元 811 利用包括在空间信息中的补偿信息——例如利用缩减混音增益或扩展信息来补偿任意缩减混音信号。

[0202] 3D 渲染单元 815 通过对经补偿的缩减混音信号执行 3D 渲染操作来生成解码器 3D 缩减混音信号。多声道解码器 816 利用经补偿的缩减混音信号和包括在空间信息中的基本空间信息来生成 3D 多声道信号。

[0203] 缩减混音补偿单元 811 可按以下方式补偿任意缩减混音信号。

[0204] 如果补偿信息是缩减混音增益,则缩减混音补偿单元 811 利用缩减混音增益补偿任意缩减混音信号的能级,使得任意缩减混音信号可被转换成类似于缩减混音信号的信号。

[0205] 如果补偿信息是第二扩展信息,则缩减混音补偿单元 811 可利用第二扩展信息补偿任意缩减混音信号所缺少的分量。

[0206] 多声道解码器 816 可通过将前置矩阵 M1、混音矩阵 M2 和后置矩阵 M3 顺序地施加于缩减混音矩阵信号来生成多声道信号。在这种情况下,第二扩展信息可用于在将混音矩阵 M2 施加于缩减混音信号期间补偿缩减混音信号。换言之,第二扩展信息可用于补偿已经施加有前置矩阵 M1 的缩减混音信号。

[0207] 如上所述,可通过将扩展信息应用于生成多声道信号来选择性地补偿多个声道中的每一个。例如,如果将扩展信息施加于混音矩阵 M2 的中心声道,则可由扩展信息补偿缩减混音信号的左声道和右声道分量。如果将扩展信息施加于混音矩阵 M2 的左声道,则可由扩展信息补偿缩减混音信号的左声道分量。

[0208] 缩减混音增益和扩展信息两者都可用作补偿信息。例如,可利用扩展信息补偿任意缩减混音信号的低频带,且可利用缩减混音增益补偿任意缩减混音信号的高频带。此外,还可利用扩展信息补偿除任意缩减混音信号的低频带以外的、诸如可显著影响音质的峰值或凹口之类的部分任意缩减混音信号。与将由扩展信息补偿的部分有关的信息可被包括在比特流中。指示包括在比特流中的缩减混音信号是否是任意缩减混音信号的信息以及指示比特流是否包括补偿信息的信息可被包括在比特流中。

[0209] 为了防止由编码单元 800 生成的缩减混音信号被削波,可将缩减混音信号除以预定增益。预定增益可具有静态值或动态值。

[0210] 缩减混音补偿单元 811 可通过利用预定增益补偿为防止削波而削弱的缩减混音信号来恢复原始缩减混音信号。

[0211] 可容易地再现由缩减混音补偿单元 811 补偿的任意缩减混音信号。或者,还待补偿的任意缩减混音信号可被输入至 3D 渲染单元 815,并可由 3D 渲染单元 815 转换成解码器 3D 缩减混音信号。

[0212] 参考图 12,缩减混音补偿单元 811 包括第一域转换器 812、补偿处理器 813 和第二域转换器 814。

[0213] 第一域转换器 812 将任意缩减混音信号的域转换成预定域。补偿处理器 813 利用补偿信息——例如,缩减混音增益或扩展信息——来补偿预定域中的任意缩减混音信号。

[0214] 任意缩减混音信号的补偿可在 QMF/ 混合域中进行。为此,第一域转换器 812 可对任意缩减混音信号执行 QMF/ 混合分析。第一域转换器 812 可将任意缩减混音信号的域转换成除 QMF/ 混合域以外的域,例如,诸如 DFT 或 FFT 域之类的频域。任意缩减混音信号的补偿也可在除 QMF/ 混合域以外的域中进行,例如,频域或时域。

[0215] 第二域转换器 814 将经补偿的任意缩减混音信号的域转换成与原始任意缩减混音信号相同的域。更具体地,第二域转换器 814 通过反向地执行由第一域转换器 812 所执行的域转换操作来将经补偿的任意缩减混音信号的域转换成与原始任意缩减混音信号相同的域。

[0216] 例如,第二域转换器 814 可通过对经补偿的任意缩减混音信号执行 QMF/ 混合合成将经补偿的任意缩减混音信号转换成时域信号。同样,第二域转换器 814 可对经补偿的任意缩减混音信号执行 IDFT 或 IFFT。

[0217] 类似于图 7 所示的 3D 渲染单元 710,3D 渲染单元 815 可对在频域、QMF/ 混合域或时域中的、经补偿的任意缩减混音信号执行 3D 渲染操作。为此,该 3D 渲染单元 815 可包括域转换器(未示出)。域转换器将经补偿的任意缩减混音信号的域转换成将要执行 3D 渲染操作的域,或转换通过 3D 渲染操作获取的信号域。

[0218] 其中补偿处理器 813 补偿任意缩减混音信号的域可与其中 3D 渲染单元 815 对经补偿的任意缩减混音信号执行 3D 渲染操作的域相同或不同。

[0219] 图 13 是根据本发明的实施例的缩减混音补偿 /3D 渲染单元 820 的框图。参考图 13,缩减混音补偿 /3D 渲染单元 820 包括第一域转换器 821、第二域转换器 822、补偿 /3D 渲染处理器 823 以及第三域转换器 824。

[0220] 缩减混音补偿 /3D 渲染单元 820 可在单个域中对任意缩减混音信号执行补偿操作和 3D 渲染操作,从而减少解码装置的计算量。

[0221] 更具体地,第一域转换器 821 将任意缩减混音信号的域转换成其中将执行补偿操作和 3D 渲染操作的第一域。第二域转换器 822 转换空间信息,其包括生成多声道信号所必需的基本空间信息和补偿任意缩减混音信号所必需的补偿信息,使得空间信息变得可适用于第一域。补偿信息可包括缩减混音增益和扩展信息中的至少一个。

[0222] 例如,第二域转换器 822 可将 QMF/ 混合域中对应于参数频带的补偿信息映射到频带,使得补偿信息可变得容易地适用于频域。

[0223] 第一域可以是诸如 DFT 或 FFT 之类的频域、QMF/ 混合域或时域。或者,第一域可以是除本文中陈述的域以外的域。

[0224] 在补偿信息的转换期间,可发生时延。为了解决该问题,第二域转换器 822 可执行时延补偿操作,使得补偿信息的域和第一域之间的时延能够被补偿。

[0225] 补偿 /3D 渲染处理器 823 利用经转换的空间信息对第一域中的任意缩减混音信号执行补偿操作,然后对通过补偿操作获取的信号执行 3D 渲染操作。补偿 /3D 渲染处理器 823 可按与本文陈述的不同的顺序执行补偿操作和 3D 渲染操作。

[0226] 补偿 /3D 渲染处理器 823 可同时对任意缩减混音信号执行补偿操作和 3D 渲染操作。例如,补偿 /3D 渲染处理器 823 可通过使用新滤波器系数对第一域中的任意缩减混音信号执行 3D 渲染操作来生成经补偿的 3D 缩减混音信号,该新滤波器系数是补偿信息和通常在 3D 渲染操作中使用的现有滤波器系数的组合。

[0227] 第三域转换器 824 将补偿 /3D 渲染处理器 823 所生成的 3D 缩减混音信号的域转换成频域。

[0228] 图 14 是根据本发明的实施例用于处理兼容缩减混音信号的解码装置 900 的框图。参考图 14,解码装置 900 包括第一多声道解码器 910、缩减混音兼容性处理单元 920、第二多声道解码器 930 和 3D 渲染单元 940。将省略与图 1 的实施例相同的解码过程的详细描述。

[0229] 兼容缩减混音信号是可由两个或更多多声道解码器解码的缩减混音信号。换言之,兼容缩减混音信号是最初针对预定多声道解码器优化、然后能够通过兼容性处理操作转换成针对除该预定多声道解码器以外的多声道解码器优化的信号的缩减混音信号。

[0230] 参考图 14,假设输入的兼容缩减混音信号针对第一多声道解码器 910 优化。为了使第二多声道解码器 930 解码输入的兼容缩减混音信号,缩减混音兼容性处理单元 920 可对输入的兼容缩减混音信号执行兼容性处理操作,使得输入的兼容缩减混音信号可被转换成针对第二多声道解码器 930 优化的信号。第一多声道解码器 910 通过解码输入的兼容缩减混音信号来生成第一多声道信号。第一多声道解码器 910 可通过仅仅使用输入的兼容缩减混音信号而不需要空间信息进行解码来生成多声道信号。

[0231] 第二多声道解码器 930 利用由缩减混音兼容性处理单元 920 执行的兼容性处理操作获取的缩减混音信号生成第二多声道信号。3D 渲染单元 940 可通过对由缩减混音兼容性处理单元 920 执行的兼容性处理操作获取的缩减混音信号执行 3D 渲染操作来生成解码器 3D 缩减混音信号。

[0232] 可利用诸如逆矩阵之类的兼容性信息,将针对预定多声道解码器优化的兼容缩减混音信号转换成针对除预定多声道解码器以外的多声道解码器优化的缩减混音信号。例如当存在利用不同编码方法的第一和第二多声道编码器与利用不同编码 / 解码方法的第一和第二多声道解码器时,编码装置可将矩阵施加于第一多声道编码器生成的缩减混音信号,从而生成针对第二多声道解码器优化的兼容缩减混音信号。然后,解码装置可将逆矩阵施加于由编码装置所生成的兼容缩减混音信号,从而生成针对第一多声道解码器优化的兼容缩减混音信号。

[0233] 参考图 14,缩减混音兼容性处理单元 920 可利用逆矩阵对输入的兼容缩减混音信号执行兼容性处理操作,从而生成针对第二多声道解码器 930 优化的缩减混音信号。

[0234] 与缩减混音兼容性处理单元 920 所使用的逆矩阵有关的信息可提前存储在解码

装置 900 中,或可被包括在编码装置传送的比特流中。此外,指示包括在输入比特流中的缩减混音信号是任意缩减混音信号还是兼容缩减混音信号的信息可被包括在输入比特流中。

[0235] 参考图 14,缩减混音兼容性处理单元 920 包括第一域转换器 921、兼容性处理器 922 和第二域转换器 923。

[0236] 第一域转换器 921 将输入的兼容缩减混音信号的域转换成预定域,且兼容性处理器 922 利用诸如逆矩阵之类的兼容性信息执行兼容性处理操作,使得在预定域中的输入兼容缩减混音信号可被转换成针对第二多声道解码器 930 优化的信号。

[0237] 兼容性处理器 922 可在 QMF/混合域中执行兼容性处理操作。为此,第一域转换器 921 可对输入的兼容缩减混音信号执行 QMF/混合分析。同样,第一域转换器 921 可将输入的兼容缩减混音信号的域转换成除 QMF/混合域以外的域,例如,诸如 DFT 或 FFT 域之类的频域,且兼容性处理器 922 可在除 QMF/混合域以外的域——如频域或时域中执行兼容性处理操作。

[0238] 第二域转换器 923 转换由兼容性处理操作获取的兼容缩减混音信号的域。更具体地,第二域转换器 923 可通过反向地执行由第一域转换器 921 所执行的域转换操作来将通过兼容性处理操作获取的兼容性缩减混音信号的域转换成与原始输入兼容缩减混音信号相同的域。

[0239] 例如,第二域转换器 923 可通过对由兼容性处理操作获取的兼容缩减混音信号执行 QMF/混合域合成来将由兼容性处理操作获取的兼容缩减混音信号转换成时域信号。或者,第二域转换器 923 可对由兼容处理操作获取的兼容缩减混音信号执行 IDFT 或 IFFT。

[0240] 3D 渲染单元 940 可对在频域、QMF/混合域或时域中的、由兼容性处理操作获取的兼容缩减混音信号执行 3D 渲染操作。为此,该 3D 渲染单元 940 可包括域转换器(未示出)。域转换器将输入的兼容缩减混音信号的域转换成其中将执行 3D 渲染操作的域,或转换由 3D 渲染操作获取的信号的域。

[0241] 其中兼容性处理器 922 执行兼容性处理操作的域可与其中 3D 渲染单元 940 执行 3D 渲染操作的域相同或不同。

[0242] 图 15 是根据本发明的实施例的缩减混音兼容性处理/3D 渲染单元 950 的框图。参考图 15,缩减混音兼容性处理/3D 渲染单元 950 包括第一域转换器 951、第二域转换器 952、兼容性/3D 渲染处理器 953 以及第三域转换器 954。

[0243] 缩减混音兼容性处理/3D 渲染单元 950 在单个域中执行兼容性处理操作和 3D 渲染操作,从而减少解码装置的计算量。

[0244] 第一域转换器 951 将输入的兼容缩减混音信号转换至其中将执行兼容性处理操作和 3D 渲染操作的第一域。第二域转换器 952 转换空间信息和兼容性信息,例如逆矩阵,使得空间信息和兼容性信息能变得适用于第一域。

[0245] 例如,第二域转换器 952 可将 QMF/混合域中对应于参数频带的逆矩阵映射到频域,使得逆矩阵可容易地适用于频域。

[0246] 第一域可以是诸如 DFT 或 FFT 域之类的频域、QMF/混合域或时域。或者,第一域可以是除本文中陈述的域以外的域。

[0247] 在空间信息和兼容性信息的转换期间,可发生时间延迟。

[0248] 为了解决该问题,第二域转换器 952 可执行时延补偿操作,使得空间信息和补偿

信息的域与第一域之间的时延能够被补偿。

[0249] 兼容性 /3D 渲染处理器 953 利用经转换的兼容性信息对第一域中的输入兼容缩减混音信号执行兼容性处理操作,然后对通过兼容性处理操作获取的兼容缩减混音信号执行 3D 渲染操作。兼容性 /3D 渲染处理器 953 可按与本文陈述的不同的顺序执行兼容性处理操作和 3D 渲染操作。

[0250] 兼容性 /3D 渲染处理器 953 可同时对输入的兼容缩减混音信号执行兼容性处理操作和 3D 渲染操作。例如,兼容性 /3D 渲染处理器 953 可通过使用新滤波器系数对第一域中的输入兼容缩减混音信号执行 3D 渲染操作来生成 3D 缩减混音信号,该新滤波器系数是兼容性信息和通常在 3D 渲染操作中使用的现有滤波器系数的组合。

[0251] 第三域转换器 954 将兼容性 /3D 渲染处理器 953 所生成的 3D 缩减混音信号的域转换成频域。

[0252] 图 16 是根据本发明的实施例用于消除串扰的解码装置的框图。参考图 16,解码装置包括比特拆分单元 960、缩减混音解码器 970、3D 渲染单元 980 以及串扰消除单元 990。将省略与图 1 的实施例相同的解码过程的详细描述。

[0253] 由 3D 渲染单元 980 输出的 3D 缩减混音信号可由耳机再现。然而,当 3D 缩减混音信号由远离用户的扬声器再现时,声道间串扰很可能发生。

[0254] 因此,解码装置可包括对 3D 缩减混音信号执行串扰消除操作的串扰消除单元 990。

[0255] 解码装置可执行声场处理操作。

[0256] 声场处理操作中使用的声场信息,即,标识其中将再现 3D 缩减混音信号的空間的信息,可被包括在由编码装置传送的输入比特流中,或可由解码装置选择。

[0257] 输入比特流可包括混响时间信息。可根据混响时间信息控制在声场处理操作中使用的滤波器。

[0258] 对于早前部分和后面的混响部分可有差别地执行声场处理操作。例如,早前部分可利用 FIR 滤波器来处理,而后面的混响部分可利用 IIR 滤波器处理。

[0259] 更具体地,可通过使用 FIR 滤波器在时域中执行卷积操作或通过时在域中执行乘法操作、并将乘法操作的结果转换至时域来对早前部分执行声场处理操作。声场处理操作可在时域中对后面的混响部分执行。

[0260] 可将本发明实现为写在计算机可读记录介质上的计算机可读代码。计算机可读记录介质可以是其中数据以计算机可读方式存储的任何类型的记录设备。计算机可读记录介质的例子包括 ROM、RAM、CD-ROM、磁带、软盘、光数据存储器和载波(例如,通过因特网的数据传输)。可将计算机可读记录介质分布在连接至网络的多个计算机系统上,使得计算机可读代码能以分散的方式向其写入或从其执行。实现本发明所需的功能程序、代码和代码段可由本领域的普通技术人员容易地解释。

[0261] 如上所述,根据本发明,高效率地编码具有 3D 效果的多声道信号,并根据再现环境的特性以最优音质自适应地恢复和再现音频信号是可能的。

[0262] 工业实用性

[0263] 其他实现在以下权利要求的范围内。例如,根据本发明的编组、数据译码和熵译码可被应用于各种应用领域和各种产品。应用本发明的一个方面的存储数据的存储介质在本发明的范围内。

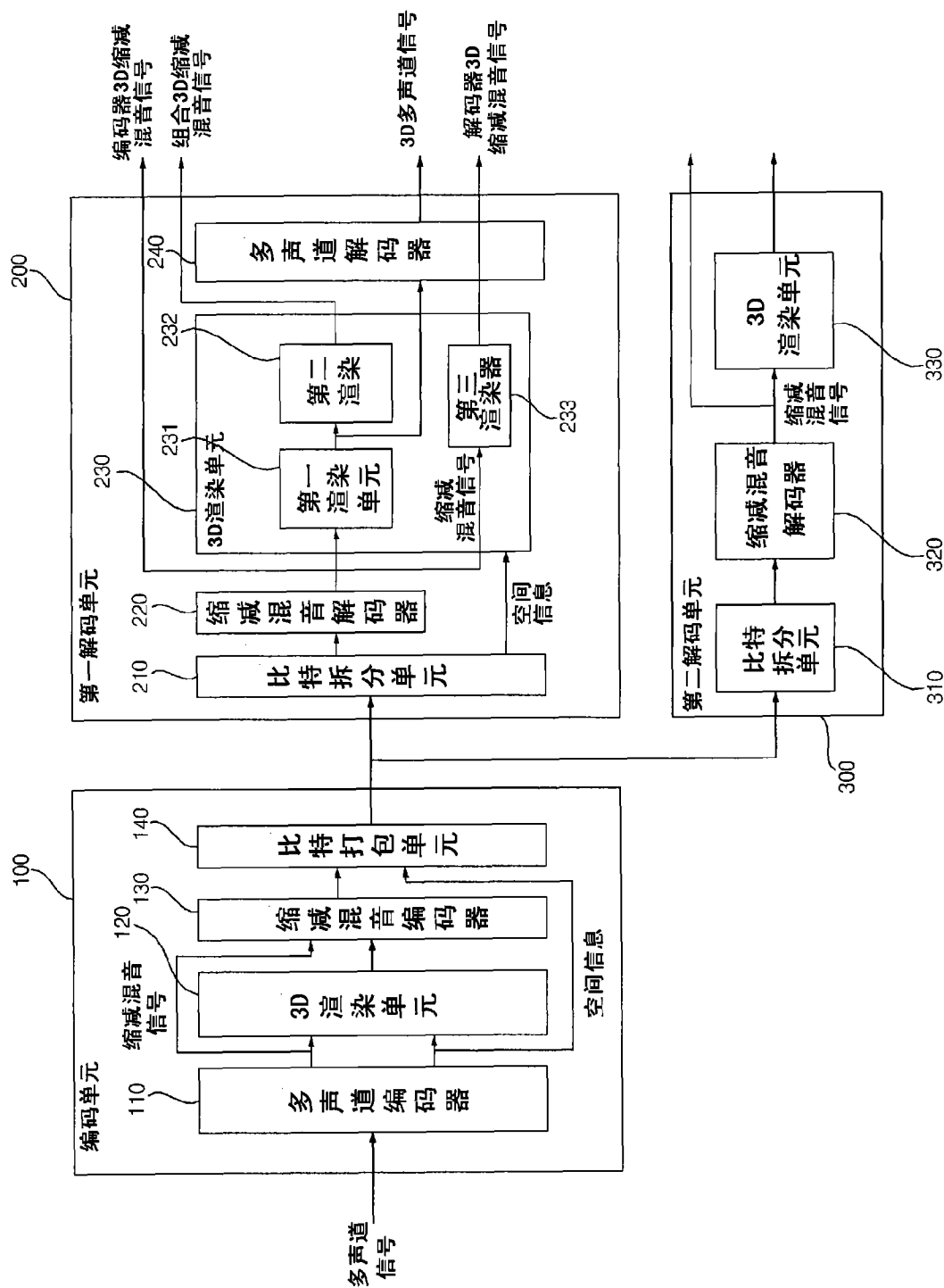


图 1

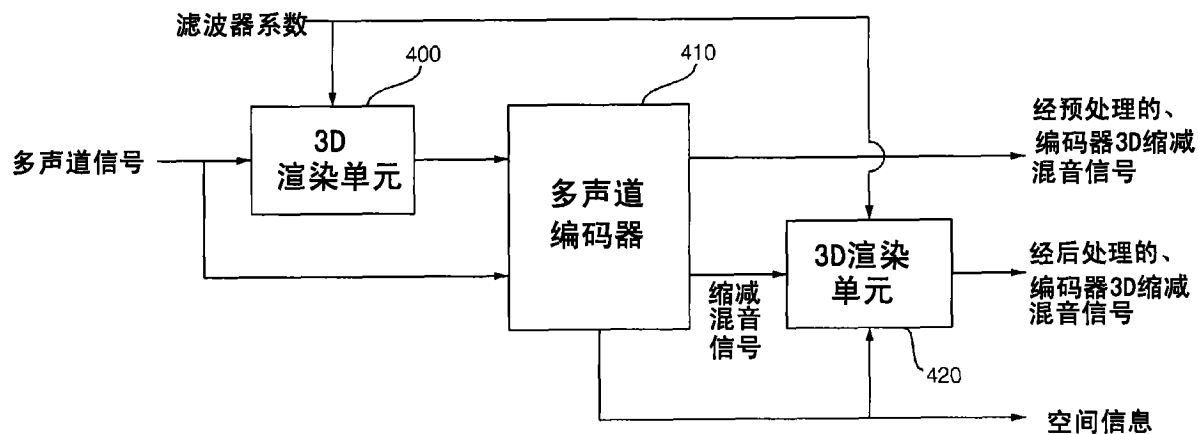


图 2

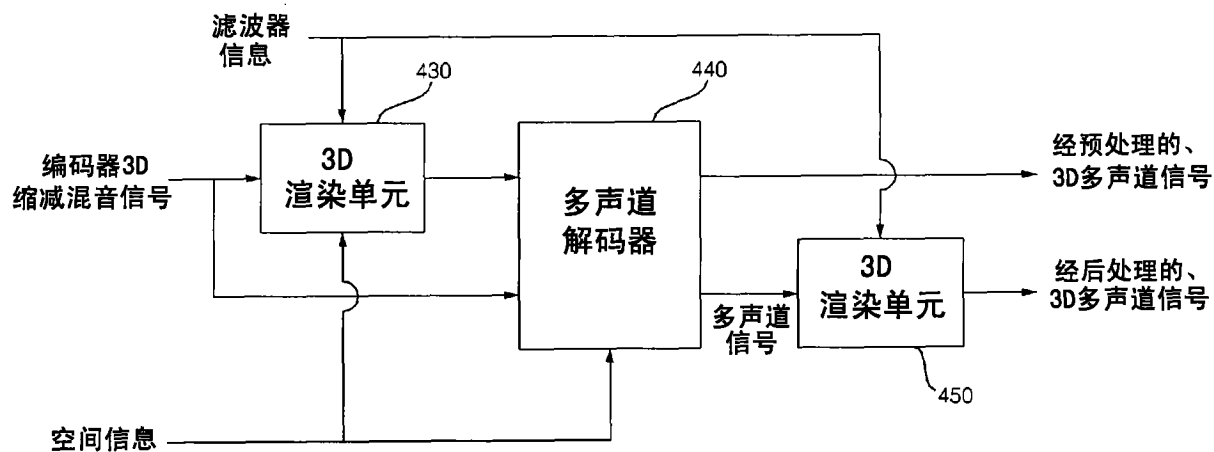


图 3

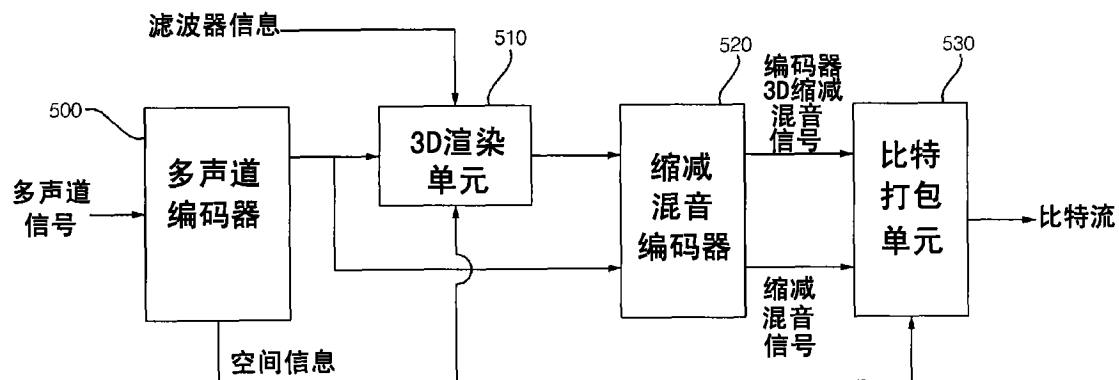


图 4

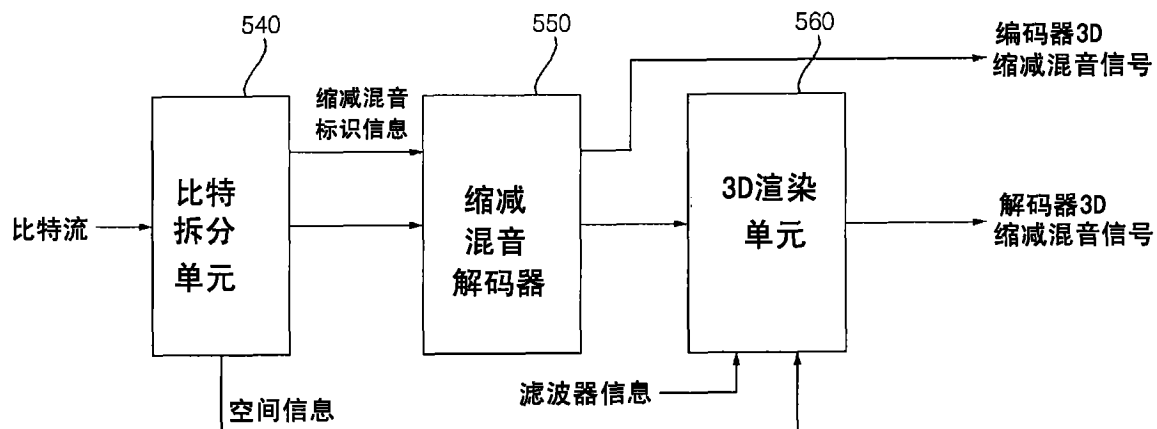


图 5

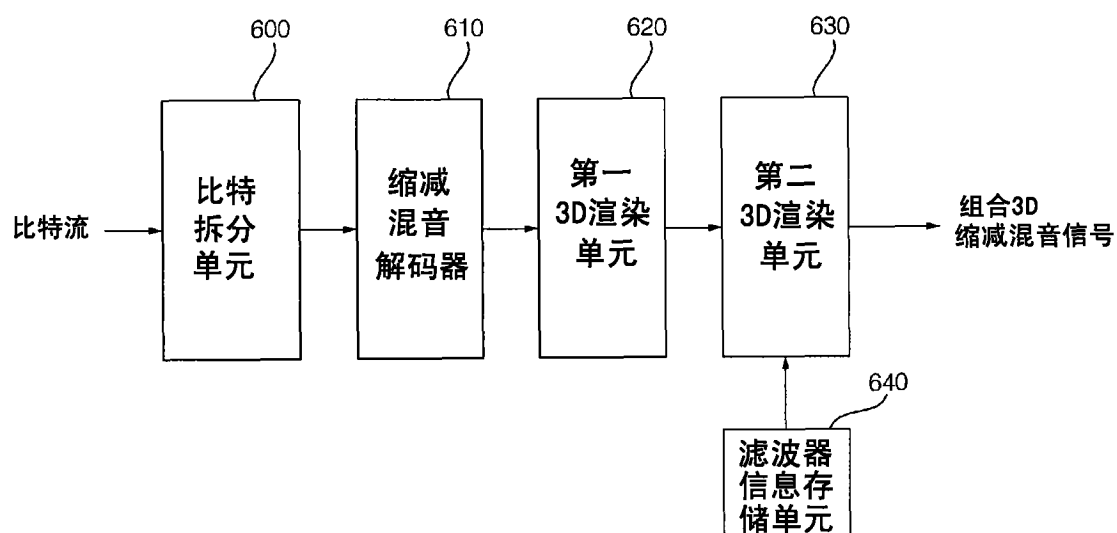


图 6

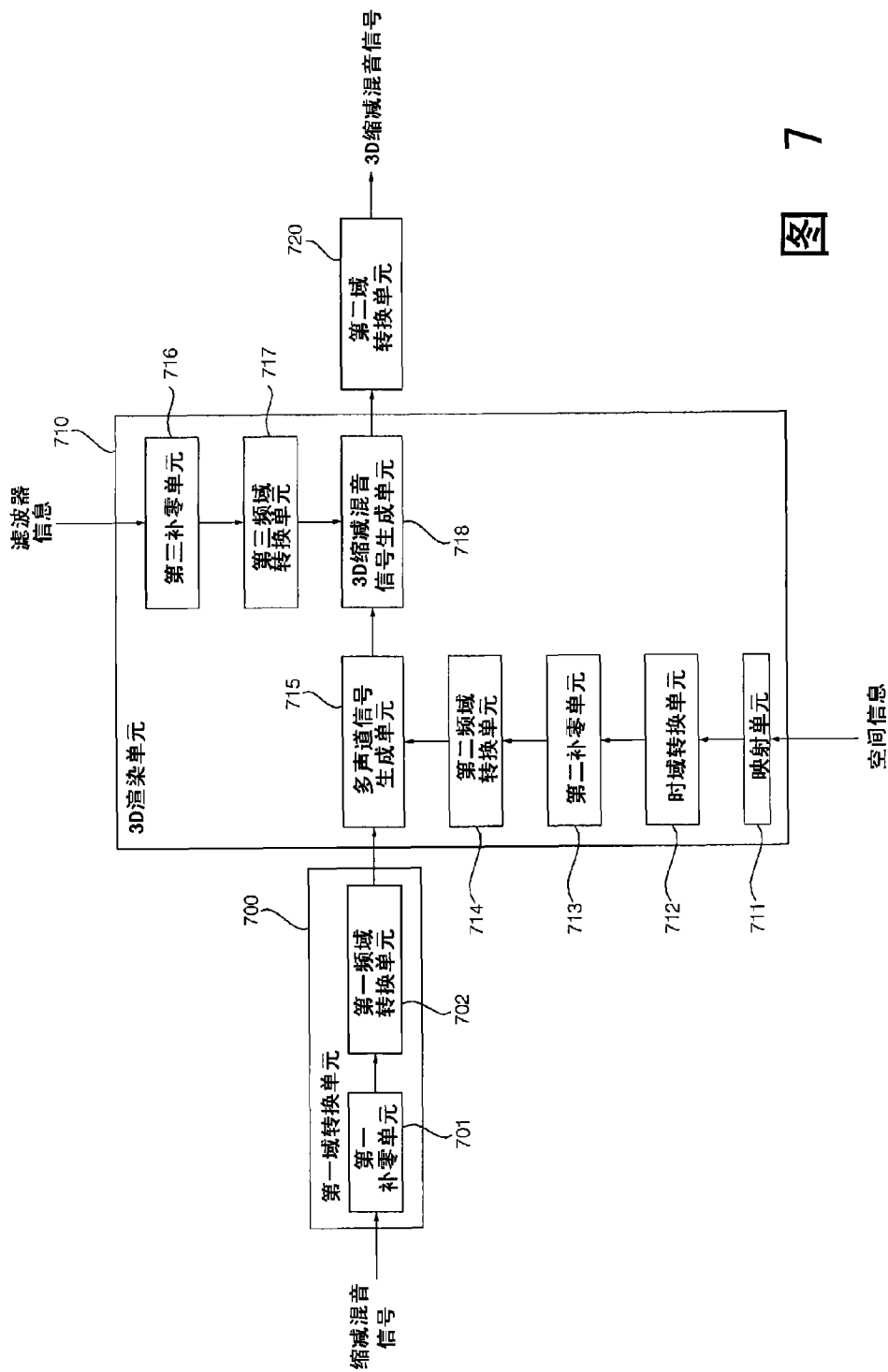


图 7

图 8



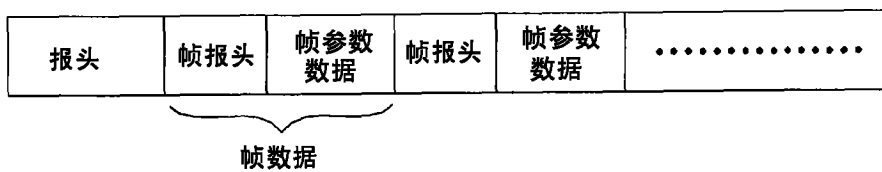


图 9

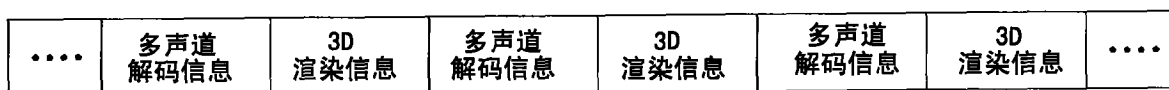


图 10

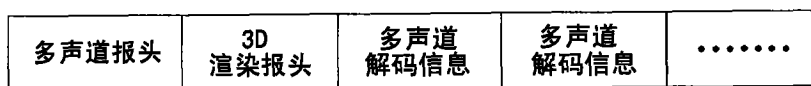


图 11

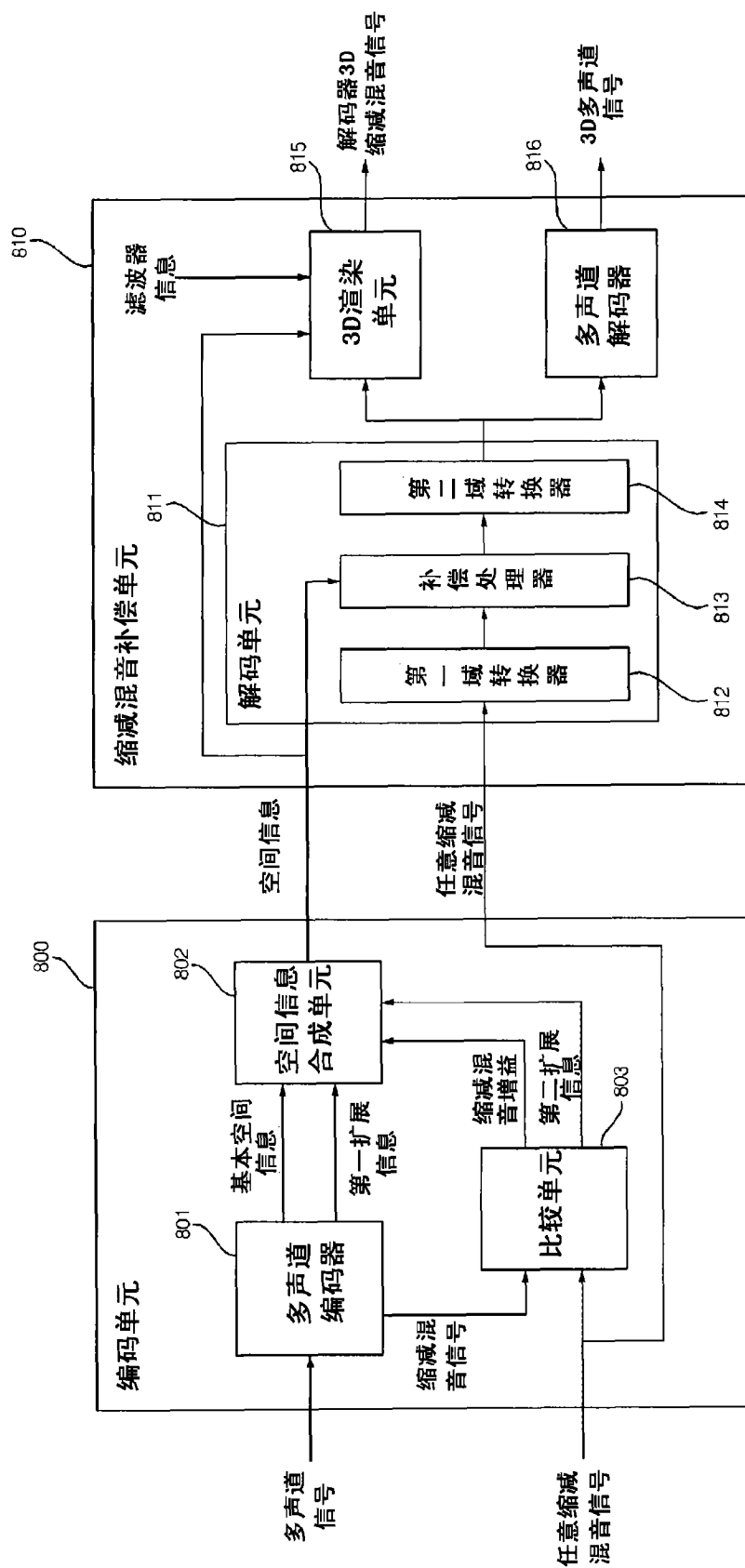


图 12

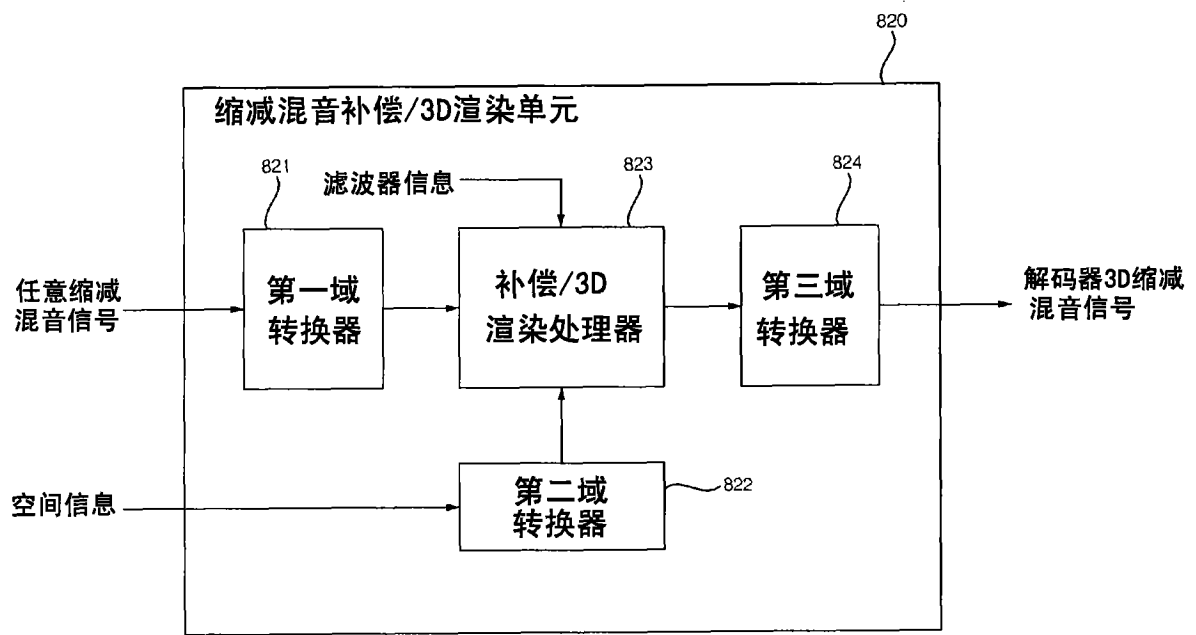


图 13

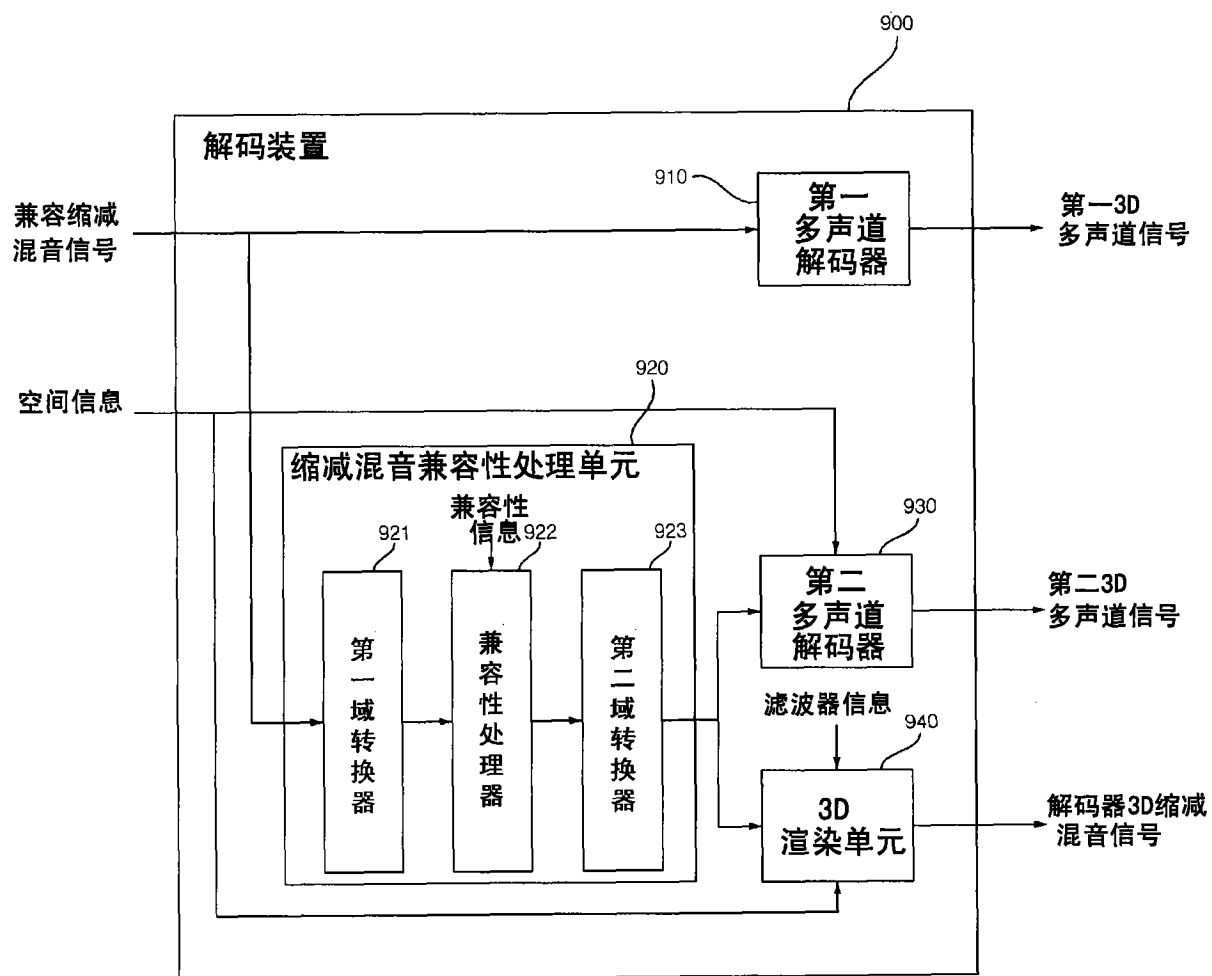


图 14

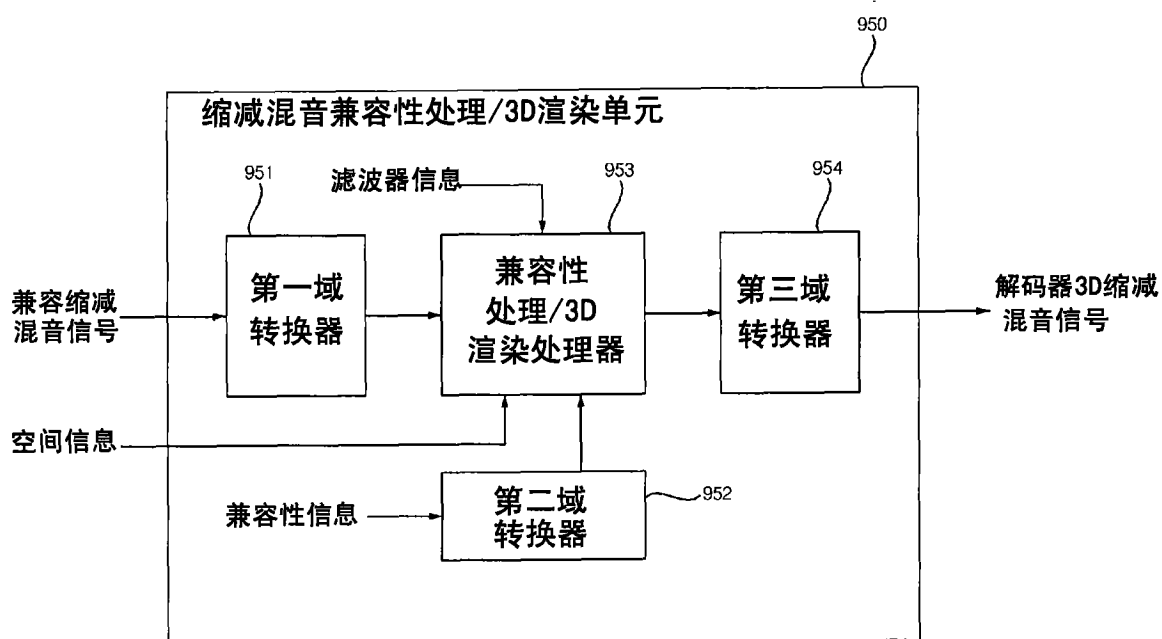


图 15

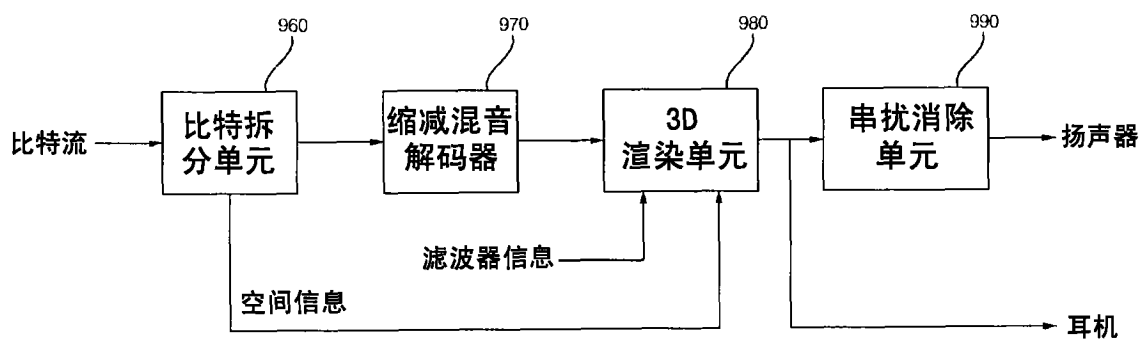


图 16