



(12) 发明专利

(10) 授权公告号 CN 102714035 B

(45) 授权公告日 2015. 12. 16

(21) 申请号 201080052486. 3

(22) 申请日 2010. 10. 15

(30) 优先权数据

10171459. 0 2010. 07. 30 EP

61/252, 298 2009. 10. 16 US

61/369, 256 2010. 07. 30 US

(85) PCT国际申请进入国家阶段日

2012. 05. 18

(86) PCT国际申请的申请数据

PCT/EP2010/065503 2010. 10. 15

(87) PCT国际申请的公布数据

W02011/045409 EN 2011. 04. 21

(73) 专利权人 弗兰霍菲尔运输应用研究公司

地址 德国慕尼黑

(72) 发明人 科尔内利娅·法尔克 于尔根·赫莱

莱昂·特伦迪

(74) 专利代理机构 北京康信知识产权代理有限

责任公司 11240

代理人 余刚 吴孟秋

(51) Int. Cl.

G10L 19/008(2013. 01)

(56) 对比文件

CN 101411214 A, 2009. 04. 15,

CN 101529504 A, 2009. 09. 09,

W0 2008/100067 A1, 2008. 08. 21,

Jurgen Herre 等. Technical provisions for limiting perceptible distortions in SAOC. 《90. MPEG Meeting》. 2009, 1-8.

审查员 陈成

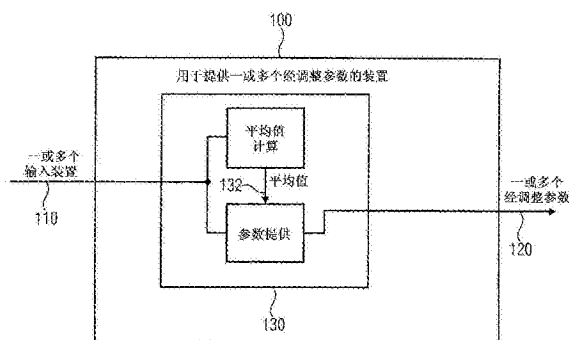
权利要求书3页 说明书22页 附图11页

(54) 发明名称

用以提供一或多个经调整参数的装置及方法

(57) 摘要

一种用以基于下混信号表示形态及与下混信号表示形态相关联的参数侧边信息来提供用于提供上混信号表示形态的一或多个经调整参数的装置,其包含一参数调整器。该参数调整器被配置为接收一或多个参数,及基于此而提供一或多个经调整参数。该参数调整器被配置为依据多个参数值的平均值而提供一或多个经调整参数,使得通过使用非最佳参数所导致的该上混信号表示形态的失真,对偏离最佳参数的参数至少减少大于预定偏差。



1. 一种用以提供一或多个经调整参数 (120、252、352) 的装置 (100、250、350、440、450), 其中所述经调整参数 (120、252、352) 用于提供一上混信号表示形态 (220、430a-430M), 并且所述装置基于一上混信号表示形态 (210、420) 及与所述下混信号表示形态相关联的一参数侧边信息 (212、422) 来提供所述经调整参数 (120、252、352), 其特征在于, 所述装置包含:

一参数调整器, 其被配置为接收一或多个参数 (110、214、337), 及基于此而提供一或多个经调整参数 (120、252、352), 其中所述参数调整器被进一步配置为依据多个参数值 (110、214、337) 的平均值 (132) 而提供一或多个经调整参数, 使得经由使用非最佳参数用以提供所述上混信号表示形态所造成的所述上混信号表示形态的失真, 对偏离最佳参数的一或多个参数至少减少大于一预定偏差。

2. 根据权利要求 1 所述的装置 (100、250、350、440、450), 其中, 所述参数调整器被进一步配置为依据多个参数值的加权平均的平均值而提供一或多个经调整参数。

3. 根据权利要求 1 所述的装置 (100、250、350、440、450), 其中, 所述参数调整器被进一步配置为提供一或多个经调整参数使得所述一或多个经调整参数偏离所述平均值的量小于相对应所接收的参数偏离所述平均值的量。

4. 根据权利要求 1 所述的装置 (100、250、440), 其中, 所述装置被进一步配置为接收描述音频对象对所述上混信号表示形态 (220、430a-430M) 的一或多个声道的贡献的一或多个呈现系数 (214), 及其中, 所述装置被进一步配置为提供一或多个经调整的呈现系数 (252) 作为经调整参数。

5. 根据权利要求 4 所述的装置 (100、250、440), 其中, 所述参数调整器被进一步配置为接收多个呈现系数 (214) 作为输入参数; 及

其中, 所述参数调整器被进一步配置为运算出与多个音频对象相关联的呈现系数的一平均值; 及

其中, 所述参数调整器被进一步配置为提供经调整的呈现系数 (252), 使得经调整的呈现系数偏离与多个音频对象相关联的呈现系数的一平均值的偏差被限制。

6. 根据权利要求 5 所述的装置 (100、250、440), 其中, 所述参数调整器被进一步配置为使得落入于依据呈现系数的平均值所测定的容许区间内的一呈现系数 (214) 维持不变, 及将大于所述容许区间的上边界值的一呈现系数 (214) 选择性地设定为小于或等于所述上边界值的一值, 及

将小于所述容许区间的下边界值的一呈现系数 (214) 选择性地设定为大于或等于所述下边界值的一值。

7. 根据权利要求 5 所述的装置 (100、250、440), 其中, 所述参数调整器被进一步配置为迭代重复地选择所述呈现系数中的一个别者, 其包含于个别迭代重复中与所述呈现系数平均值的最大偏离; 及使得所述呈现系数中的所选择的个别者更接近所述呈现系数平均值, 来将落在依据所述呈现系数平均值所测定的容许区间外侧的呈现系数迭代重复地调整至所述容许区间内部。

8. 根据权利要求 7 所述的装置 (100、250、440), 其中, 所述参数调整器被进一步配置为重复所述呈现系数中的一个别者的迭代重复选择, 及重复所述呈现系数中的所选择的个别者的迭代重复修正, 直至全部呈现系数皆调整至落入适用的容许区间内部为止。

9. 根据权利要求 1 所述的装置 (100、350、450), 其中所述装置被进一步配置为接收一或多个转码系数 (337), 其描述所述下混信号表示形态 (210、420) 的一或多个声道映射至所述上混信号表示形态 (220、430a-430M) 的一或多个声道的映射关系, 及

其中, 所述装置被进一步配置为提供一或多个经调整的转码系数 (352) 作为经调整参数。

10. 根据权利要求 9 所述的装置 (100、350、450), 其中, 所述参数调整器被进一步配置为接收转码系数 (337) 的一时间序列作为输入参数; 及

其中, 所述参数调整器被进一步配置为依据多个转码系数算出一时间均值; 及

其中, 所述参数调整器被进一步配置为提供所述经调整的转码系数 (352), 使得所述经调整的转码系数与所述时间均值的偏差限制。

11. 根据权利要求 10 所述的装置 (100、350、450), 其中, 所述参数调整器被进一步配置为允许落在依据所述时间均值所测定的一容许区间内部的一转码系数 (337) 维持不变, 及

将大于所述容许区间的上边界值的一转码系数选择性地设定为小于或等于所述上边界值的一值, 及

将小于所述容许区间的下边界值的一转码系数选择性地设定为大于或等于所述下边界值的一值。

12. 根据权利要求 10 所述的装置 (100、350、450), 其中, 所述参数调整器被进一步配置为使用所述转码系数 (337) 序列的递归低通滤波而求出所述时间均值。

13. 根据权利要求 1 所述的装置 (100、250、350、440、450), 其中, 所述参数调整器被进一步配置为提供一或多个经调整参数中的一给定者, 使得所述经调整参数中的所述给定者落在容许区间内部, 所述容许区间的边界依据多个输入参数值的平均值 (132) 及一或多个容许参数界定, 以及使得一输入参数与一相对应经调整参数间的偏差为最小化或维持在预定最大容许范围以内。

14. 根据权利要求 13 所述的装置 (100、250、350、440、450), 其中, 所述参数调整器被进一步配置为, 其边界依据多个输入参数值的平均值 (132) 界定的所述容许区间, 将发现落在所述容许区间外部的一输入参数选择性地设定至所述容许区间的一上边界值或一下边界值来获得所述输入参数的经调整版本。

15. 根据权利要求 13 所述的装置 (100、250、350、440、450), 其中, 所述参数调整器被进一步配置为迭代重复地选择所述输入参数中的一个个别者, 其包含于个别迭代重复中与所述平均值 (132) 的最大偏离; 以及将所述输入参数中的所选择的个别者调整至更接近所述平均值, 来迭代重复地将判定为落在其边界依据平均值界定的一容许区间外部的输入参数调整至所述容许区间内部。

16. 根据权利要求 15 所述的装置 (100、350、450), 其中, 所述参数调整器被进一步配置为选择一修正阶大小, 所述修正阶用来将所述输入参数中较为接近所述平均值的所选择的个别者调整至所述输入参数中的所选择的个别者与所述平均值间的差的预定分量。

17. 一种用以提供一上混信号表示形态 (220、430a-430M) 的装置 (200、300、410), 其中所述装置基于一一下混信号表示形态 (210、420) 及一参数侧边信息 (212、422) 来提供所述上混信号表示形态 (220、430a-430M), 其特征在于, 所述装置包含:

根据权利要求 1 所述的用以提供一或多个经调整参数 (120、252、352) 的装置 (100、

250、350、440、450) ;

一信号处理器 (230、330), 其被配置为基于所述下混信号表示形态及所述参数侧边信息而获得所述上混信号表示形态,

其中, 提供一或多个经调整参数的所述装置被进一步配置为调整所述信号处理器的一或多个处理参数 (252、352)。

18. 根据权利要求 17 所述的装置 (200、300、410), 其中, 所述信号处理器 (230) 被进一步配置为依据经调整的呈现系数 (252) 而提供所述上混信号表示形态 (220、430a-430M), 所述经调整的呈现系数描述音频对象对所述上混信号表示形态的一或多个声道的贡献; 及

其中, 用以提供一或多个经调整参数的所述装置 (100、250、440) 被进一步配置为接收多个使用者指定的呈现参数 (214) 作为输入参数, 及基于此而提供由所述信号处理器使用的一或多个经调整的呈现参数 (252)。

19. 根据权利要求 17 所述的装置 (200、300、410), 其中, 用以提供一或多个经调整参数的所述装置 (100、350、450) 被进一步配置为接收一混合矩阵的一或多个混合矩阵元 (337) 作为一或多个输入参数, 及基于此而提供由所述信号处理器 (330) 使用的一或多个经调整的所述混合矩阵的混合矩阵元 (352); 及

其中, 所述信号处理器被进一步配置为依据经调整的所述混合矩阵的混合矩阵元 (352) 而提供所述上混信号表示形态 (220、430a-430M), 其中, 所述混合矩阵描述所述下混信号表示形态的一或多个音频声道信号映射至所述上混信号表示形态的一或多个音频声道信号的映射关系。

20. 根据权利要求 17 所述的装置 (200、300、410), 其中, 所述信号处理器被进一步配置为获得 MPEG 环绕任意下混增益值, 及

其中, 用以提供一或多个经调整参数的所述装置被进一步配置为接收多个任意下混增益值作为输入参数, 及提供多个经调整的任意下混增益值。

21. 一种用以提供一或多个经调整参数的方法, 其中所述经调整参数用于提供一上混信号表示形态, 并且所述方法基于一下混信号表示形态及与所述下混信号表示形态相关联的一参数侧边信息来提供所述经调整参数, 其特征在于, 所述方法包含:

接收一或多个参数; 及

基于此而提供一或多个经调整参数, 其中, 所述一或多个经调整参数依据多个参数值的平均值而提供, 使得经由使用非最佳参数用以提供所述上混信号表示形态所造成的所述上混信号表示形态的失真, 对偏离最佳参数的一或多个参数至少减少大于一预定偏差。

用以提供一或多个经调整参数的装置及方法

技术领域

[0001] 依据本发明的实施例涉及一种用以基于一下混信号表示形态及与该下混信号表示形态相关联的一参数侧边信息来提供用于提供一上混信号表示形态的一或多个经调整参数的装置。

[0002] 依据本发明的另一实施例涉及一种用以基于该下混信号表示形态及该参数侧边信息来提供一上混信号表示形态的装置。

[0003] 依据本发明的另一实施例涉及一种用以基于一下混信号表示形态及与该下混信号表示形态相关联的一参数侧边信息来提供用于提供一上混信号表示形态的一或多个经调整参数的方法。

[0004] 依据本发明的另一实施例涉及一种用以执行该方法的计算机程序。

[0005] 依据本发明的若干实施例涉及一种用于 MPEG SAOC 的失真控制参数限制方案。

背景技术

[0006] 于音频处理、音频传输及音频储存领域,逐渐需要处理多声道内容来改良听觉感受。多声道音频内容的使用给使用者带来显著改进。举例言的,可获得三维空间听觉感受而为用户带来娱乐效果的满足与改善。但多声道音频内容也可用于职业环境,例如用于电话会议应用,原因在于通过使用多声道音频回放可改良发话者的可懂性(易于为人所了解)。

[0007] 但也期望在音频质量与位率需求间获得良好折衷,来避免因多声道应用造成额外过度资源负荷。

[0008] 近来,已经提示用于含有多音频对象的音频场景(audio scene)进行位率有效的传输及/或储存的参数技术,例如双耳线索编码(类别 I)(例如参考参考文献[1])、联合来源编码(例如参考参考文献[2])、及 MPEG 空间音频对象编码(例如参考参考文献[3]、[4]、[5])。

[0009] 若执行极端对象的呈现(rendering),则组合在接收端的使用者互动,此等技术可导致输出信号的低音频质量(例如参考参考文献[6])。

[0010] 此等技术针对听觉上重建期望的输出音频场景而非通过波形匹配。

[0011] 图 8 显示此种系统(此处:MPEG SAOC)的系统概述。图 8 所示 MPEG SAOC 系统 800 包含一 SAOC 编码器 810 及一 SAOC 解码器 820。SAOC 编码器 810 接收多个对象信号 x_1 至 x_N , 其例如可表示为时域信号或时频域信号(例如呈傅立叶型变换的一变换系数集合形式,或呈 QMF 子频带信号形式)。SAOC 编码器 810 典型地也接收下混系数 d_1 至 d_N , 其与对象信号 x_1 至 x_N 相关联。下混系数的分开集合可供下混信号的各声道利用。SAOC 编码器 810 典型地被配置为经由依据相关联的下混系数 d_1 至 d_N 而组合对象信号 x_1 至 x_N 来获得一下混信号声道。典型地,下混声道比对象信号 x_1 至 x_N 少。为了允许(至少近似)于 SAOC 解码器 820 端的对象信号的分离(或分开处理),SAOC 编码器 810 提供该一或多个下混信号(标示为下混声道)812 及一侧边信息 814 二者。侧边信息 814 描述的对象信号 x_1 至 x_N 特性来

允许解码器端的对象专一性处理。

[0012] SAOC 解码器 820 被配置为接收该一或多个下混信号 812 及侧边信息 814。又, SAOC 解码器 820 典型地被配置为接收一使用者互动信息及 / 或一使用者控制信息 822, 其描述期望的呈现设定值。举例言之, 使用者互动信息 / 使用者控制信息 822 可描述一扬声器设定值及提供对象信号 x_1 至 x_N 的该等对象的期望空间配置。

[0013] SAOC 解码器 820 被配置为提供例如多个已解码上混声道信号 $\hat{y}_1$ 至 $\hat{y}_M$ 。上混声道信号例如可与多扬声器呈现配置的个别扬声器相关联。SAOC 解码器 820 可例如包含一对象分离器 820a, 其被配置为基于该一或多个下混信号 812 及侧边信息 814, 重建 (至少近似) 对象信号 x_1 至 x_N , 由此获得已重建对象信号 820b。但已重建对象信号 820b 可能略为偏离原先对象信号 x_1 至 x_N , 例如原因在于由于位率限制, 侧边信息 814 并非相当足够用于完好重建。SAOC 解码器 820 可进一步包含一混合器 820c, 其可被配置为接收已重建对象信号 820b 及使用者互动信息 / 使用者控制信息 822, 及基于此而提供上混声道信号 $\hat{y}_1$ 至 $\hat{y}_M$ 。混合器 820c 可被配置为使用该使用者互动信息 / 使用者控制信息 822 而判定个别已重建对象信号 820b 对上混声道信号 $\hat{y}_1$ 至 $\hat{y}_M$ 的贡献。使用者互动信息 / 使用者控制信息 822 例如可包含呈现参数 (也标示为呈现系数) 其判定个别已重建对象信号 822 对上混声道信号 $\hat{y}_1$ 至 $\hat{y}_M$ 的贡献。

[0014] 但须注意于多个实施例中, 对象的分离于图 8 以对象分离器 820a 指示, 及混合于图 8 以混合器 820c 指示以单一步骤执行。为了实现此项目的, 总参数可经运算, 其描述该一或多个下混信号 812 映射至上混声道信号 $\hat{y}_1$ 至 $\hat{y}_M$ 的直接映射关系。此等参数可基于侧边信息及使用者互动信息 / 使用者控制信息 820 运算。

[0015] 现在参考图 9a、图 9b 及图 9c, 将描述用以基于一下混信号表示形态及对象相关侧边信息来提供一上混信号表示形态的不同的装置。须注意该对象相关侧边信息为与该下混信号相关联的侧边信息的实例。图 9a 显示一种包含 SAOC 解码器 920 的 MPEG SAOC 系统 900 的方块示意图。SAOC 解码器 920 包含一对象解码器 922 及一混合器 / 呈现器 926 作为分开功能方块。对象解码器 922 依据该下混信号表示形态 (例如呈以时域或时频域表示的一或多个下混信号形式) 及该对象相关侧边信息 (例如呈对象元数据 (meta data) 形式) 而提供多个已重建的对象信号 924。混合器 / 呈现器 926 接收与多个 N 个对象相关联的已重建的对象信号 924, 及基于此且系基于该呈现信息而提供一或多个上混声道信号 928。于该 SAOC 解码器 920, 对象信号 924 的提取与混合 / 呈现分开进行, 其允许对象解码功能与混合 / 呈现功能的分离, 但带来相当高的运算复杂度。

[0016] 现在参考图 9b, 将简短讨论另一种 MPEG SAOC 系统 930, 其包含一 SAOC 解码器 950。SAOC 解码器 950 依据该下混信号表示形态 (例如呈一或多个下混信号形式) 及该对象相关侧边信息 (例如呈对象元数据 (meta data) 形式) 而提供多个上混声道信号 958。SAOC 解码器 950 包含对象解码器与混合器 / 呈现器的组合, 其被配置为于联合混合程序获得上混声道信号 958, 而未分开对象解码与混合 / 呈现, 其中用于该联合上混处理的参数取

决于该对象相关侧边信息及该呈现信息。该联合上混处理也依据下混信息,该下混信息被视为该对象相关侧边信息的一部分。

[0017] 综上所述,上混声道信号 928、958 的提供可于一步骤式处理或二步骤式处理执行。

[0018] 现在参考图 9c,将描述一种 MPEG SAOC 系统 960。SAOC 系统 960 包含 SAOC 至 MPEG 环绕转码器 980,而非 SAOC 解码器。

[0019] SAOC 至 MPEG 环绕转码器包含一侧边信息转码器 982,其被配置为接收该对象相关侧边信息(例如呈对象元数据形式)及选择性地,接收一或多个下混信号的信息及呈现信息。该侧边信息转码器也被配置为基于所接收的数据而提供 MPEG 环绕侧边信息(例如呈 MPEG 环绕位流形式)。据此,侧边信息转码器 982 被配置为考虑呈现信息及选择性地,考虑该一或多个下混信号内容的相关信息,而将接收自该对象编码器的一对象相关(参数)侧边信息变换成一声道相关(参数)侧边信息。

[0020] 选择性地,SAOC 至 MPEG 环绕转码器 980 可被配置为操控例如由下混信号表示形态所描述的该一或多个下混信号而获得经操控的下混信号表示形态 988。但可删除下混信号操控器 986,使得 SAOC 至 MPEG 环绕转码器 980 的输出下混信号表示形态 988 与 SAOC 至 MPEG 环绕转码器的输入下混信号表示形态相同。若声道相关的 MPEG 环绕侧边信息 984 不允许基于 SAOC 至 MPEG 环绕转码器 980 的输入下混信号表示形态提供期望的听觉印象(于某些呈现群(rendering constellations)可能为此种情况),则可使用下混信号操控器 986。

[0021] 据此,SAOC 至 MPEG 环绕转码器 980 提供下混信号表示形态 988 及 MPEG 环绕位流 984,使得使用接收 MPEG 环绕位流 984 及下混信号表示形态 988 的 MPEG 环绕解码器,可产生多个上混声道信号,其表示依据输入该 SAOC 至 MPEG 环绕转码器 980 的呈现信息的该等音频对象。

[0022] 综上所述,可使用用以解码 SAOC 编码的音频信号的不同构想。于某些情况下,使用 SAOC 解码器,其依据该下混信号表示形态及对象相关参数侧边信息而提供上混声道信号(例如上混声道信号 928、958)。此种构想的实例可参考图 9a 及图 9b。另外,SAOC 编码的音频信息可经转码来获得一下混信号表示形态(例如下混信号表示形态 988)及一声道相关侧边信息(例如声道相关 MPEG 环绕位流 984),其可由 MPEG 环绕解码器用来提供期望的上混声道信号。

[0023] 于 MPEG SAOC 系统 800,系统概述显示于图 8,一般处理以频率选择方式进行,且于各频带内可描述如下:

[0024] • N 个输入音频对象信号 x_1 至 x_N 经下混作为 SAOC 编码器处理的一部分。用于单声道下混,下混系数标示以 d_1 至 d_N 。此外,SAOC 编码器 810 提取描述该输入音频对象的侧边信息 814。用于 MPEG SAOC,对象功率相对于彼此的关系乃此种侧边信息的最基本形式。

[0025] • 下混信号(或多个信号)812 及侧边信息 814 被传输及/或储存。为了实现此项目的,下混音频信号可使用众所周知的听觉音频编码器压缩,诸如 MPEG-1 层 II 或 III(也称之为“.mp3")、MPEG 高级音频编码(AAC)、或其他音频编码器。

[0026] • 于接收端,SAOC 解码器 820 于构想上尝试使用所传输的侧边信息 814(及当然,一或多个下混信号 812)来重新储存该原先对象信号(“对象分离”)。然后,此等近似的对象信号(也标示为重建的对象信号 820b)使用一呈现矩阵而混合入由 M 个音频输出声道表

示的目标场景（例如可由上混声道信号 y_1 至 y_M 表示）。用于单声道输出，呈现矩阵系数以 r_1 至 r_N 表示。

[0027] 实际上，罕见执行（或甚至未曾执行）对象信号的分离，原因在于分离步骤（以对象分离器 820a 指示）及混合步骤（以混合器 820c 指示）二者组合成单一转码步骤，其经常导致运算复杂度的剧减。

[0028] 已经发现此种方案就传输位率（只需传输数个下混声道加若干侧边信息，而无需传输 N 个分开对象音频信号或分开系统）及运算复杂度（处理复杂度主要涉及输出声道的数目而非音频对象数目）而言极其有效。对于接收端的使用者的额外优点包括选择一呈现设定值的自由度（单声、立体声、环绕、虚拟耳机回放等）及使用者互动的特征结构：呈现矩阵，如此，输出场景可由使用者依据意愿、个人偏好或其他标准而设定且互动改变。举例言之，可以定位共同在一个空间区的谈话者来最大化与其余谈话者间的区别。此种互动性可由设置解码器使用者界面而实现。

[0029] 对各个所传输的声音对象，可调整其相对电平及（用于非单声道呈现）呈现的空间位置。当使用者改变相关联的图形使用者界面（GUI）滑动器位置时可实时发生（例如：对象电平 = +5 分贝，对象位置 = -30 度）。

[0030] 但发现于某些情况下，用以提供上混信号表示形态（例如上混声道信号 y_1 至 y_M ）的参数的解码器端选择造成听觉的降级。

[0031] 有鉴于此种情况，本发明的目的是提供一种构想其允许当提供上混信号表示形态（例如上混声道信号 y_1 至 y_M ）时减少或甚至避免听觉失真。

发明内容

[0032] 此一问题可由下述装置获得解决，该种用以基于一混信号表示形态及与该下混信号表示形态相关联的一参数侧边信息来提供用于提供一上混信号表示形态的一或多个经调整参数的装置。该装置包含一参数调整器，其被配置为接收一或多个参数（于若干实施例可为输入参数），及基于此而提供一或多个经调整参数。该参数调整器被配置为依据多个参数值（于若干实施例可为输入参数值）的平均值而提供一或多个经调整参数，使得经由使用非最佳参数用以提供该上混信号表示形态所造成的该上混信号表示形态的失真，对偏离最佳参数的参数（或输入参数）至少减少大于一预定偏差。

[0033] 依据本发明的此一实施例基于下述构想，多个输入参数值的平均值组成有意义数量，其允许用于参数的调整，该等参数用来基于一混信号表示形态及与该下混信号表示形态相关联的一参数侧边信息而提供一上混信号表示形态，原因在于失真经常因过度偏离此一平均值所造成。平均值的使用允许调整一或多个参数来避免如此过度偏离平均值（偶尔也标示为均值），结果带来避免过度降级音频质量的可能。

[0034] 前文讨论的实施例提供一种保护所呈现的 SAOC 场景的存在声音质量的构想，对该所呈现的 SAOC 场景，全部处理皆可完全于 SAOC 解码器 / 转码器内进行，原因在于 SAOC 解码器 / 转码器包含用以调整参数所需的完整信息。又，前述实施例并未涉及该呈现场景的听觉音频质量的复杂测量值的显式计算，原因在于发现限制参数值与平均值间的偏差典

型地导致良好听觉印象,而参数值与平均值间的重大偏差典型地导致听觉失真。如此,前文讨论的实施例提供一种特别有效的机制,即平均值用来适当调整参数,该等参数被考虑用以提供上混信号表示形态。

[0035] 于优选实施例,该装置的参数调整器被配置为依据属于多个参数值的加权平均的一平均值而提供一或多个经调整的参数。使用加权平均提供高度自由度,原因在于可对不同参数值配置不同的权值。但配置相同的权值对于该等参数值也是可能的。

[0036] 于优选实施例,该装置的参数调整器被配置为提供一或多个经调整的参数,使得该等提供一或多个经调整的参数偏离该平均值小于对应的接收的参数。通过将经调整的参数调整至接近平均值,或甚至经由设定经调整的参数等于平均值,可实现显著失真减少。

[0037] 于优选实施例,该装置被配置为接收描述音频对象对该上混信号表示形态的一或多个声道的贡献的一或多个呈现系数(也标示为呈现参数)。此种情况下,装置优选被配置为提供一或多个经调整的呈现系数作为经调整的参数。已经发现依据多个呈现参数的平均值(其作为输入参数值)而调整呈现参数,带来获得良好适合的经调整的呈现参数的可能,避免过度听觉失真。

[0038] 于优选实施例,参数调整器被配置为接收多个呈现系数作为输入参数。此种情况下,参数调整器被配置为对多个音频对象相关联的呈现系数运算平均。又,参数调整器被配置为提供经调整的呈现系数,使得限制一经调整的呈现系数与对多个音频对象相关联的呈现系数平均间的偏差。依据本发明的此一实施例基于发现若一经调整的呈现系数与对多个音频对象相关联的呈现系数平均间的偏差被限制,则至少对偏离最佳呈现参数达大于一预定偏压的呈现参数而言,经由使用非最佳呈现参数所造成的上混信号表示形态失真典型地减少。如此,一个简单机制即调整呈现系数使得该经调整的呈现系数与对多个音频对象相关联的呈现系数平均间的偏差被限制,则允许避免过度听觉失真。

[0039] 于优选实施例,参数调整器被配置为保持一呈现系数不变,该呈现系数在依据对呈现系数的平均所测定的一容许区间以内;以及将大于该容许区间的上边界值的一呈现系数选择性地设定为小于或等于该上边界值的一值;及将小于该容许区间的下边界值的一呈现系数选择性地设定为大于或等于该下边界值的一值。据此,建立调整呈现系数的一种极为简单的机制,其中此种简单机制仍然允许获得经调整的呈现系数,其避免因使用与平均值有强力差异的非最佳呈现参数所造成的上混信号表示形态的过度失真。

[0040] 于优选实施例,该参数调整器被配置为迭代重复地选择该等呈现系数中的一个别者,其包含于个别迭代重复中与该呈现系数平均值的最大偏离;及使得该等呈现系数中的该选定者更接近该呈现系数平均值。据此,落在依据该呈现系数平均值所测定的容许区间外侧的呈现参数被迭代重复地调整至该容许区间内部。如此,呈现参数系依据平均值而调整,使得使用非最佳呈现参数所造成的上混信号表示形态的失真典型地减低(至少对偏离最佳呈现参数达大于预定偏离的输入呈现参数而言是如此)。

[0041] 于优选实施例,该参数调整器被配置为重复该等呈现系数中的一个别者的迭代重复选择,及重复该等呈现系数中的该选定者的迭代重复修正,直至全部呈现系数皆调整至落入适用的容许区间内部为止。如此,确保于该上混信号表示形态的听觉失真维持够小。

[0042] 于优选实施例,该装置被配置为接收一或多个转码系数,其描述该下混信号表示形态的一或多个声道映射至该上混信号表示形态的一或多个声道的映射关系。此种情况

下,该装置被配置为提供一或多个已调整的转码系数作为经调整的参数。依据本发明的此一实施例基于发现转码参数为极为适合用于依据平均值的调整,原因在于转码系数大为偏离平均值,典型地造成听觉失真。据此,通过依据平均值调整或限制转码参数,可减少因使用非最佳转码参数(至少对偏离最佳转码参数达大于预定偏差的输入转码参数)所引起的上混信号表示形态的失真。

[0043] 于优选实施例,该参数调整器被配置为接收转码系数(也标示为转码参数)的一时间序列作为输入参数。此种情况下,该参数调整器被配置为依据多个转码系数算出一时间均值(也标示为时间平均)。又,该参数调整器被配置为提供该等经调整的转码系数,使得该等经调整的转码系数与该时间均值的偏差限制。再度,提供一种用以避免经由使用非最佳转码参数而造成上混信号表示形态的过度听觉失真的简单机构。

[0044] 于优选实施例,该参数调整器被配置为允许落在依据该时间均值(其构成平均值)所测定的一容许区间内部的一转码系数维持不变。又,该参数调整器被配置为将大于该容许区间的上边界值的一转码系数选择性地设定为小于或等于该上边界值的一值,及将小于该容许区间的下边界值的一转码系数选择性地设定为大于或等于该下边界值的一值。据此,可将转码系数调整至明确界定的容许区间内,其允许减少因使用非最佳转码参数所引起的上混信号表示形态的失真,至少对偏离最佳转码参数达大于预定偏差的输入转码参数尤为如此。当使用时间均值时,容许区间以适应性方式选择。此一构想基于发现转码系数的强时间变化典型地带来听觉失真,因此须限于某种程度。

[0045] 于优选实施例,该参数调整器被配置为使用该转码系数序列的递归低通滤波而算出该时间均值。此种构想显示带来一极为明确界定的时间均值,其将转码系数的长期演化列入考虑。又,发现此种转码系数序列的递归低通滤波可使用低运算强度及存储强度执行,其协助减少内存需求。特别,可获得有意义的时间均值而未长时间储存转码系数历史。

[0046] 于优选实施例,该参数调整器被配置为提供一或多个经调整参数中的一给定者,使得该等经调整参数中的该给定者落在容许区间内部,该容许区间的边界依据多个输入参数值的平均值及一或多个容许参数界定,以及使得一输入参数与一相对应经调整参数间的偏差为最小化或维持在预定最大容许范围以内。已经发现通过限制经调整的参数于容许区间,同时考虑避免输入参数与对应经调整的参数间有过大差异的目的,可获得带来良好听觉印象的经调整的参数。据此,可减少经由使用非最佳转码参数而造成上混信号表示形态的失真而不必损及由该等输入参数所界定期望的听觉设定值。

[0047] 于优选实施例,该参数调整器被配置为,其边界依据多个输入参数值的平均值界定的该容许区间,将发现落在该容许区间外部的一输入参数选择性地设定至该容许区间的一上边界值或一下边界值来获得该输入参数的经调整版本。

[0048] 于另一优选实施例,该参数调整器被配置为迭代重复地选择该等输入参数中的一个别者,其包含于个别迭代重复中与该平均值的最大偏离;以及将该等输入参数中的该选定者调整至更接近该平均值,来迭代重复地将判定为落在其边界依据平均值界定的一容许区间(其边界依据平均值而界定)外部的输入参数调整至该容许区间内部。

[0049] 于优选实施例,该参数调整器被配置为选择一阶大小,该阶用来将该等输入参数中较为接近该平均值的选定者调整至该等输入参数中的该选定者与该平均值间的差的预定分量。

[0050] 依据本发明的另一实施例提供一种用以基于一下混信号表示形态及一参数侧边信息来提供一上混信号表示形态的装置。该装置包含如前文讨论的用以基于一或多个所接收的参数而提供一或多个经调整参数的一装置。该用以提供一上混信号表示形态的装置也包含一信号处理器,其被配置为基于该下混信号表示形态及该参数侧边信息而获得该上混信号表示形态。该用以提供一或多个经调整参数的装置被配置为提供例如输入至该信号处理器的呈现参数的、或于该信号处理器运算的且由该信号处理器施加的转码参数等该信号处理器的一或多个处理参数的经调整版本来获得该上混信号表示形态。

[0051] 此一实施例基于发现大量参数,该等参数由信号处理器施加,及输入信号处理器或甚至于信号处理器计算,及其可基于该平均值而自前文讨论的参数调整获益。已经发现若一参数集合(例如与不同音频对象相关联的一呈现系数集合,或与时间上不同情况相关联的一转码参数值集合)良好平衡,使得此种数值集合的个别值并未包含与平均值的过度大量偏差,则信号处理器典型地提供良好质量的上混信号表示形态,小有失真。如此,经由采用用以提供一或多个经调整的参数的装置组合用以提供上混信号表示形态的装置,可实现本发明构想的效益。

[0052] 于优选实施例,该信号处理器被配置为依据经调整的呈现系数,其描述音频对象对该上混信号表示形态的一或多个声道的贡献而提供该上混信号表示形态。该用以提供一或多个经调整参数的装置被配置为接收多个使用者指定的呈现参数作为输入参数,及基于此而提供由该信号处理器(优选至信号处理器)使用的一或多个经调整的呈现参数。已经发现使用该用以提供一或多个经调整参数的装置所能获得的良好平衡的呈现参数,典型地导致良好听觉印象。

[0053] 于另一实施例,该用以提供一或多个经调整参数的装置被配置为接收一混合矩阵的一或多个混合矩阵元作为该一或多个输入参数,及基于此而提供由该信号处理器使用的一或多个经调整的该混合矩阵的混合矩阵元。此种情况下,该信号处理器被配置为依据经调整的该混合矩阵的混合矩阵元而提供该上混信号表示形态,其中该混合矩阵描述该下混信号表示形态(例如表示呈时域表示形态或时频域表示形态形式)的一或多个音频声道信号映射至该上混信号表示形态的一或多个音频声道信号的映射关系。已经发现混合矩阵元也应良好适应于平均值,例如混合矩阵元的时间变化受限制。

[0054] 依据本发明的另一实施例,该音频处理器被配置为获得 MPEG 环绕任意下混增益值。此种情况下,该用以提供一或多个经调整参数的装置被配置为接收多个任意下混增益值作为输入参数,及提供多个经调整的任意下混增益值。已经发现施加用以提供经调整的参数的装置至任意下混增益值,也导致良好听觉印象且允许限制听觉失真。

[0055] 依据本发明的其他实施例提供一种用以提供一或多个经调整的参数的方法及计算机程序。该方法基于前文讨论的装置的不同发现且可由此处就本发明装置讨论的结构特征及功能中的任一者而扩展延伸。

附图说明

[0056] 图 1 显示依据本发明的实施例一种用以提供一或多个经调整的参数的装置的方块示意图;

[0057] 图 2 显示依据本发明的实施例一种用以提供上混信号表示形态的装置的方块示

意图；

[0058] 图 3 显示依据本发明的另一实施例一种用以提供上混信号表示形态的装置的方块示意图；

[0059] 图 4 显示使用间接控制及直接控制的参数限制方案的方块示意图；

[0060] 图 5a 显示表示收听测试条件的一表；

[0061] 图 5b 显示表示收听测试的音频项目的一表；

[0062] 图 6 显示表示所测试的极端呈现条件的一表；

[0063] 图 7 显示对不同参数限制方案 (PLS), MUSHRA 收听测试结果的一线图表示形态；

[0064] 图 8 显示参考 MPEG SAOC 系统的方块示意图；

[0065] 图 9a 显示使用分开的解码器及混合器的一参考 SAOC 系统的方块示意图；

[0066] 图 9b 显示使用集成型解码器及混合器的一参考 SAOC 系统的方块示意图；

[0067] 图 9c 显示使用 SAOC 至 MPEG 转码器的一参考 SAOC 系统的方块示意图；及

[0068] 图 10 显示一表描述哪些转码系数可由所提示的参数限制方案而修正。

具体实施方式

[0069] 1. 依据图 1, 用以提供一或多个经调整的参数值的装置

[0070] 后文中, 将描述一种用以基于下混信号表示形态及与下混信号表示形态相关联的参数侧边信息来提供用于提供上混信号表示形态的一或多个经调整参数的装置。图 1 显示此种装置 100 的方块示意图。

[0071] 该装置 100 被配置为接收一或多个输入参数 110, 及基于此而提供一或多个经调整的参数 120。装置 100 包含一参数调整器 130, 其被配置为接收一或多个输入参数 110, 及基于此而提供一或多个经调整的参数 120。该参数调整器 130 被配置为依据多个输入参数值的平均值 132 而提供该一或多个经调整的参数 120, 使得至少对偏离最佳参数达大于预定偏差的输入参数 (例如输入参数 110), 经由使用非最佳参数 (例如一或多个输入参数 110) 所造成的上混信号表示形态的失真减少。举例言之, 参数调整器 130 可具有比较该一或多个输入参数 110, 该一或多个经调整的参数 120 是“更接近” (表示造成较少失真) 最佳参数 (其将导致无失真上混信号表示形态) 的效果。

[0072] 为了实现此项目的, 参数调整器 130 实施平均值运算来获得一相关输入参数 110 (例如与一共享时间区间相关联的输入参数, 或与不同时间相关联的相同参数类型的输入参数) 集合的平均值 132 (例如呈时间平均或对象间平均)。有关装置 100 的操作, 须注意基于一或多个输入参数 110 提供一或多个经调整的参数 120 依据平均值 132 实现, 原因在于发现平均值 132 为用以调整参数的有意义数量。更明确言之, 发现 (相对于平均值) 中等参数典型地导致中等失真。

[0073] 进一步细节容后详述。

[0074] 2 依据图 2, 用以提供一种上混信号表示形态的装置

[0075] 后文中, 将描述依据图 2 的用以提供一种上混信号表示形态的装置。图 2 显示可视为音频信号解码器的此种装置 200 的方块示意图。举例言之, 装置 200 可包含 SAOC 解码器或 SAOC 转码器的功能。

[0076] 装置 200 被配置为接收一下混信号表示形态 210 及一参数侧边信息 212。又, 装置

200 被配置为接收使用者指定呈现参数 214。装置被配置为提供一上混信号表示形态 220。

[0077] 下混信号表示形态 210 例如可为一声道音频信号或二声道音频信号的表示形态。下混信号表示形态 210 例如可为时域表示形态或编码表示形态。于若干实施例中,下混信号表示形态 210 可为时频域表示形态,其中该下混信号表示形态 210 的一或多个声道由随后平均值集合表示。

[0078] 上混信号表示形态 220 例如可为呈时域表示形态或时频域表示形态形式的个别音频声道的表示形态。另外,上混信号表示形态 220 可为编码表示形态,包含一下混信号表示形态及一声道相关侧边信息二者,例如 MPEG 环绕侧边信息。

[0079] 使用者指定呈现参数 214 可呈呈现矩阵条目形式提供,该呈现矩阵条目描述多个音频对象对该上混信号表示形态 220 的一或多个声道的期望贡献。另外,使用者指定呈现参数 214 可呈任何其他适当形式提供,例如规定音频对象的期望的呈现位置及呈现体积。

[0080] 装置 200 包含一信号处理器 230,其被配置为基于下混信号表示形态 210 及参数侧边信息 212 而提供上混信号表示形态 220。该信号处理器 230 包含一重新混合功能 232,来基于该下混信号表示形态 210 而提供上混信号表示形态 220。举例言之,重新混合功能 232 可经被配置为线性组合下混信号表示形态 212 的多个声道而获得一上混信号表示形态 220 的声道。于此重新混合中,下混信号表示形态 210 的声道对上混信号表示形态 220 的声道的贡献可经由混合一混合矩阵 G 的矩阵元测定,其中混合矩阵 G 的第一维(例如列数)可由上混信号表示形态 220 的声道数目测定,及其中混合矩阵 G 的第二维(例如行数)可由下混信号表示形态 210 的声道数目测定。

[0081] 举例言之,重新混合处理 232 可用来经以将包含下混信号表示形态 210 的一或多个声道的频谱值的一或多个向量乘以混合矩阵 G,可提供包含与上混信号表示形态 220 的一或多个声道相关联的频谱值的一或多个向量。

[0082] 信号处理器 230 也包含一混合参数运算 236,其提供混合矩阵 G(或相当地,其矩阵元)。混合矩阵元系由混合参数运算 236 依据参数侧边信息 212 及已修正的呈现参数 252 测定。混合矩阵 G 的混合矩阵元例如经提供使得上混信号表示形态 220 的一或多个声道描述音频对象,依据已修正的呈现参数 252 由下混信号表示形态 210 的一或多个声道表示。为了实现此项目的,参数侧边信息 212 由混合参数运算 236 评估,其中该参数侧边信息 212 例如包含,一对象电平差信息 OLD、一对象间相关性信息 IOC、一下混增益信息 DMG、及(选择性地)一下混声道电平差信息 DCLD。该对象电平差信息例如可以逐频带方式,描述多个音频对象间的电平差。同理,该对象间相关性信息例如可以逐频带方式,描述多个音频对象间的相关性。该下混增益信息及该(选择性地)下混声道电平差信息可描述该下混,该下混执行来将来自多个音频对象的音频对象信号组合成该下混信号表示形态的一或多个声道,其中典型地具有比下混信号表示形态 210 的声道更多个音频对象。

[0083] 据此,混合参数运算 236 可评估基于参数侧边信息 212 及已修正的呈现参数 252,如何选择混合矩阵元来获得包含预期的统计性质的一上混信号表示形态 220。

[0084] 信号处理器 230 可选择性地包含侧边信息修正或侧边信息变换 240,其被配置为接收参数侧边信息 212,及提供已修正的侧边信息(例如 MPEG 环绕侧边信息),使得已修正的侧边信息及由重新混合处理 232 所提供的相关联的重新混合下混信号表示形态描述一期望的音频场景。

[0085] 要言之,信号处理器 230 例如可满足 SAOC 解码器 820 的功能,其中该下混信号表示形态 210 扮演该一或多个下混信号 812 的角色,其中该参数侧边信息 212 扮演侧边信息 814 的角色,及其中该上混信号表示形态 220 系相当于输出声道信号 y_1 至 y_M 。

[0086] 另外,信号处理器 230 可包含分开解码器及混合器 920 的功能,其中该下混信号表示形态 210 可扮演一或多个下混信号的角色,其中该参数侧边信息 212 可扮演对象元数据的角色,及其中该上混信号表示形态 220 可扮演一或多个输出声道信号 928 的角色。

[0087] 另外,信号处理器 230 可包含集成解码器及混合器 950 的功能,其中该下混信号表示形态 210 可扮演一或多个下混信号的角色,其中该参数侧边信息 212 可扮演对象元数据的角色,及其中该上混信号表示形态 220 可扮演一或多个输出声道信号 958 的角色。

[0088] 另外,信号处理器 230 可包含 MPEG 环绕转码器 980 的功能,其中该下混信号表示形态 210 可扮演一或多个下混信号的角色,其中该参数侧边信息 212 可扮演对象元数据的角色,及其中该上混信号表示形态当与 MPEG 环绕侧边信息 984 组合时可相当于该一或多个下混信号 988。

[0089] 总而言之,已修正呈现参数 252 可扮演使用者互动 / 控制信息 822 或呈现信息的角色。

[0090] 装置 200 也包含用以提供经调整的呈现参数的装置 250。用以提供经调整的呈现参数的装置 250 接收使用者指定的呈现参数 214,及基于此而提供已修正呈现参数 252。装置 250 典型地被配置为计算与不同音频对象相关联的多个使用者指定的呈现参数的平均值而获得平均值。又,装置 250 被配置为依据该平均值执行呈现参数限制,来经由限制该使用者指定的呈现参数 214 而获得已修正呈现参数 252。已修正呈现参数 252 所受限的容许区间典型地依据该平均值测定,因而避免已修正呈现参数 252 与平均值间有强烈偏差,即使使用者指定的呈现参数 214 中的一者或多者包含此种与平均值的强烈偏差亦如此。由此方式,典型地避免上混信号表示形态 220 内部的过度失真,原因在于包含有限的对象间偏差的已修正呈现参数 252 将导致具有低失真的上混信号表示形态,同时与不同音频对象相关联的呈现参数间的重大差异典型地将导致听觉假象 (audible artifacts)。

[0091] 此处须注意用以提供经调整的呈现参数的装置 250 可包含与用以提供一或多个经调整参数的装置 100 相同的总体功能,其中该使用者指定的呈现参数 214 可扮演一或多个输入参数 110 的角色,及其中该已修正呈现参数 252 可扮演一或多个经调整参数 120 的角色。

[0092] 有关提供已修正呈现参数 252 的细节将参考图 4 讨论如下。

[0093] 3 依据图 3,用以提供上混信号表示形态的装置

[0094] 后文中,依据本发明的另一实施例的用以提供上混信号表示形态的装置将参考图 3 作说明,该图显示此种装置 300 的方块示意图。

[0095] 装置 300 典型地接收与装置 200 同类型输入信号,及提供相同类型输出信号,因此相同组件符号用于此处来描述相同的或相当的信号。要言之,装置 300 接收一下混信号表示形态 210、参数侧边信息 212 及使用者指定的呈现参数 214;及装置 300 基于此而提供一上混信号表示形态 220。

[0096] 装置 300 包含一信号处理器 330,其功能可实质上相当于信号处理器 230。信号处

理器 330 包含一重新混合功能 332, 其与信号处理器 230 的重新混合功能 232 相同, 在于其基于下混信号表示形态提供重新混合的音频声道信号。但重新混合 332 使用经调整的混合矩阵, 而非直接得自混合参数运算的一混合矩阵。

[0097] 信号处理器 330 也包含一混合参数运算 336, 其功能上可与信号处理器 230 的混合参数运算 236 的功能相同。据此, 混合参数运算 336 接收参数侧边信息 212 及使用者指定的呈现参数 214, 及基于此而提供一混合矩阵 G (或相当地, 混合矩阵 G 的混合矩阵元, 也标示以 337)。

[0098] 信号处理器 330 选择性地也包含一侧边信息修正 338, 其功能与侧边信息修正 240 相同。

[0099] 此外, 装置 300 包含用以提供经调整的混合矩阵元的装置 350。装置 350 可为或可为信号处理器 330 的一部分。装置 350 被配置为接收由混合参数运算 336 所提供的混合矩阵 337, G (或相当地, 其混合矩阵元), 及基于此而提供经调整的混合矩阵 352 G' (或相当地, 其经调整的混合矩阵元)。举例言之, 每一频带及每个音频帧可提供一混合矩阵元集合及一个经调整的混合矩阵元集合。换言之, 若选用逐帧处理, 则对下混信号表示形态 210 的每个音频帧, 混合矩阵 G 及经调整的混合矩阵 G' 可更新一次。又并非必要并不同频带有多个混合矩阵 G 及经调整的混合矩阵 G' 。

[0100] 但装置 350 被配置为基于由混合参数运算 336 所提供的混合矩阵 337 的混合矩阵元而提供经调整的混合矩阵 352 的经调整的混合矩阵元。举例言之, 处理可以对混合矩阵 (或经调整的混合矩阵) 的每个位置个别进行, 使得一给定混合矩阵位置的经调整的混合矩阵元序列可取决于位在相同混合矩阵位置的混合矩阵 337 的混合矩阵元序列, 但与位在不同混合矩阵位置的混合矩阵元不相干。

[0101] 用以提供经调整的混合矩阵元的装置 350 被配置为依据基于混合矩阵 337 而运算的一或多个平均值 (例如一或多个矩阵位置个别平均值) 而提供该经调整的混合矩阵 352 的一或多个经调整的混合矩阵元。用以提供经调整的混合矩阵 352 的经调整的混合矩阵元的装置 350 优选被配置为计算在一给定混合矩阵位置随时间的经过, 混合矩阵元的平均值。如此, 对一给定混合矩阵位置, 平均值 (优选地, 但非必要地, 时间平均值, 例如浮动平均或准无限脉冲响应平均值, 或经由众所周知用于时间平均的递归低通滤波或类似数算运算所得的平均值) 可基于该给定混合矩阵位置的混合矩阵元序列运算。举例言之, 描述下混信号表示形态 210 的一给定声道对上混信号表示形态 220 的一给定声道的贡献的混合矩阵元序列 (该等混合矩阵元系与多个音频帧相关联) 可用来获得此种平均值 (也标示为均值), 该平均值可为有限脉冲响应平均值或 (准) 无限脉冲响应平均值 (例如使用众所周知用于时间平均的递归低通滤波或类似数算运算所得)。该给定混合矩阵位置的一目前经调整的混合矩阵元 (描述下混信号表示形态 210 的一给定声道对上混信号表示形态 220 的一给定声道的贡献) 可被装置 350 限制一容许区间, 该容许区间依据与该给定混合矩阵位置相关联的平均值界定。

[0102] 据此, 避免混合矩阵元的过度时间起伏波动, 原因在于经调整的混合矩阵元受限于例如由在相同混合矩阵位置的先前混合矩阵元的平均 (有限脉冲响应平均或 (准) 无限脉冲响应平均) 所测定的容许区间。已经发现此种该经调整的混合矩阵 352 的经调整的混合矩阵元的限制典型地带来由使用非最佳参数 (例如非最佳使用者指定的呈现参数) 所导

致上混信号 220 的失真限制,至少若该非最佳使用者指定的呈现参数偏离最佳使用者指定的呈现参数达多于一个预定偏离时为如此。

[0103] 此处须注意用以提供经调整的混合矩阵元的装置 350 可包含与用以提供一或多个经调整的参数的装置 100 相同的整个功能,其中该混合矩阵 337 的混合矩阵元呈扮演一或多个输入参数 110 的角色,及其中该经调整的混合矩阵 352 的经调整的混合矩阵元可扮演一或多个经调整的参数 120 的角色。

[0104] 4 依据图 4 的参数限制方案

[0105] 后文中,依据本发明的参数限制方案将参考图 4 作说明,该图显示此种参数限制方案的示意表示形态。

[0106] 图 4 显示参数限制方案组合 SAOC 解码器 410 的应用。但参数限制方案可组合不同类型音频解码器或音频转码器,例如 SAOC 转码器施用。

[0107] SAOC 解码器 410 接收下混 420 及 SAOC 位流 422。又,SAOC 解码器提供一或多个输出声道 430a 至 430M。

[0108] 于第一实施例,标示为 (a),参数限制方案实施间接控制。参数限制方案 440 接收一输入呈现矩阵 R ,例如使用者指定的呈现矩阵,及基于此而提供一经调整的呈现矩阵 $\tilde{R}$ 予 SAOC 解码器。此种情况下,SAOC 解码器如前述使用经调整的呈现矩阵 $\tilde{R}$ 用于混合矩阵 G 的导算。参数限制方案 440 也接收参数 Λ_R 、 Λ_{R+} ,其可决定容许区间边界。

[0109] 另外或此外,可施加第二参数限制方案 450。第二参数限制方案接收转码参数 T ,及基于此而提供经调整的转码参数 $\tilde{T}$ 。转码参数 T 可于 SAOC 解码器 410 运算,而经调整的转码参数 $\tilde{T}$ 可由 SAOC 解码器 410 施用。举例言之,转码参数 T 可相当于如前文讨论的混合矩阵 G 的混合矩阵元,而经调整的转码参数 $\tilde{T}$ 可相当于经调整的混合矩阵 G' 的经调整的混合矩阵元。

[0110] 参数限制方案 450 也接收一或多个参数 Λ_T 、 Λ_{T+} ,其可决定容许区间边界。

[0111] 4.1 概述

[0112] 后文中,将概述用于失真控制的参数限制方案。

[0113] 一般性 SAOC 处理以时 / 频选择方式进行,容后详述。

[0114] SAOC 编码器提取若干输入音频对象信号的心理声学特性(例如对象功率关系及相关性),及然后,下混之成为一单声道或立体声道组合(例如可标示为下混信号表示形态)。此种下混信号及所提取的侧边信息使用众所周知的听觉音频编码器,以压缩格式传输(或储存)。在接收端,SAOC 解码器于构想上尝试使用所传输的侧边信息(例如对象电平差信息 OLD、对象间相关性信息 IOC 下混增益信息 DMG、及下混声道电平差信息 DCLD)来回复原先对象信号(即分开的下混物件)。此等近似对象信号然后使用呈现矩阵(其中该呈现矩阵典型地述不同音频对象对上混信号表示形态的不同声道的贡献)混合入一目标场景。呈现矩阵由对各个所传输的音频对象及上混设定扬声器规定的相对呈现系数 RC (或对象增益)组成。此等对象增益判定全部分开的 / 呈现的对象的空间位置。实际上,罕见执行(或甚至未曾执行)对象信号的分离,原因在于分离及混合二者组合成单一组合处理步骤,其经常导致运算复杂度的剧减。单一组合处理步骤例如可使用转码系数执行,其描述

分开对象的对象分离与混合的组合。

[0115] 已经发现就传输位率（只要求传输一或二下混声道加若干侧边信息而非个别对象音频信号数目）及运算复杂度（处理复杂度主要涉及输出声道数目而非音频对象数目）两方面而言，此一方案极为有效。

[0116] SAOC 解码器（于参数字准）将对象增益及其他侧边信息直接变换成转码系数（TC），其施加至该下混信号来形成已呈现的输出音频场景的对应信号（或进一步解码操作的前处理下混信号，即典型地多声道 MPEG 环绕呈现）。

[0117] 已经发现经由施加失真控制措施或 DCM 可改良所呈现的输出音频场景的主观听觉音频质量，如非预公开的 US 61/173,456 所述。此项改良可由接受目标呈现场景的温和动态修正而实现。呈现信息的修正具有时间及频率变量本质，在特定情况下可能导致不自然的音色及时间波动假象。

[0118] 参考文献 [6] 所述失真控制措施（DCM）的替代之道中，依据本发明的实施例使用多项参数限制方案，其聚焦在音频假象（音色、时间波动等）的减少及同时保有天然声音质量。

[0119] 此处所提示的参数限制方案构想并未使用心理声学算法，基于心理声学模型调整基于计算得的失真测量值的呈现系数（RC）。反而所提示的参数限制方案构想显示低度运算及结构复杂度，因此具有整合入 SAOC 技术的吸引力。虽言如此，其也可优异地组合参考文献 [6] 所述方案来由彼此互补而实现更佳的总体输出质量。

[0120] 在总 SAOC 系统中，参数限制方案可以两种方式整合入 SAOC 解码器处理连锁。举例言之，参数限制方案可放在前端通过控制呈现系数（RC）R 而用于 SAOC 输出信号的间接（外部）修正，于图 4 显示为替代之道（a）。另外，在特性转码系数（TC）T 施加至下混信号前，系数 T 直接（内部）于 SAOC 解码器后端修正，于图 4 显示为替代之道（b）。

[0121] 4.2 间接控制

[0122] 后文中，将讨论间接控制构想的进一步细节。

[0123] 间接控制方法的基本假说考虑失真电平与 RC 偏离其对象平均值的偏差间的关系。此点基于观察到相较于其他对象，由 RC 施加更特定衰减/增强至一个特定对象，由 SAOC 解码器/转码器执行所传输的下混信号的更积极修正。换言之：“对象增益”值相对于彼此的偏差愈高，则发生无法接受的失真机率愈高（假设相同下混系数）。发现可通过检验 RC 与跨全部对象的 RC 平均值（例如平均呈现值）的偏差测试。

[0124] 未丧失普适性，后文描述基于考虑对全部对象具有统一下混增益的单声道下混的组态。对非平凡的下混情况（带有不同的及/或动态的对象增益），算法可经适当修正。此外，RC 假设为频率不变来简化记法（notation）。

[0125] 基于带有对象指针 i 的系数 R(i) 表示的使用者指定的呈现状况，PLS 通过产生实际上由 SAOC 呈现引擎所使用的修正 RC 值 $\tilde{R}(i)$ 而避免极端呈现值。其可呈如下函数导算

$$[0126] \quad \tilde{R}(i) = F_{\lambda}\{R(i), \Lambda\},$$

[0127] 此处为 PLS 控制参数（即临界值）。PLS 控制参数可视为容许参数。

[0128] 呈现系数 R(i) 与平均呈现值 $\bar{R}$ （例如算术平均）的偏差 $R_d(i)$ 可获得为

$$[0129] \quad R_d(i) = \frac{R(i)}{\bar{R}},$$

[0130] 此处

$$[0131] \quad \bar{R} = \frac{1}{N_{ob}} \sum_{i=1}^{N_{ob}} R(i).$$

[0132] 据此, $R_d(i)$ 为呈现系数 $R(i)$ 与平均呈现值 $\bar{R}$ 间的比。平均呈现值 $\bar{R}$ 为对具有音频对象指针 i 的音频对象求取平均所得呈现系数 $R(i)$ 的平均值。

[0133] 有限偏差 $\tilde{R}_d(i)$ 限于某个容许 Λ 范围为

$$[0134] \quad \tilde{R}_d(i) = \Lambda \text{ 对 } R_d(i) > \Lambda,$$

$$[0135] \quad \tilde{R}_d(i) = \frac{1}{\Lambda} \text{ 对 } R_d(i) < \frac{1}{\Lambda}。$$

[0136] 注意如此对应于相对于参考值例如 $\bar{R}$ 进行的 RC 限制运算, 其是自输入 RC 动态运算而非特定预定值。

[0137] 对所述 PLS 办法, 最佳解可以最小限问题公式化, 对此给定 RC $R(i)$ 与经修正 (经限制的) $\tilde{R}(i)$ 值间的差为最小化

$$[0138] \quad \|\tilde{R}(i) - R(i)\| \rightarrow \min.$$

[0139] 后文中, 将描述用来提供经调整的呈现系数 $\tilde{R}(i)$ 的若干算法解, 其中该经调整的呈现系数 $\tilde{R}(i)$ 可视为经调整的参数。

[0140] 以下二算法解基于位在容许范围以外的该等呈现值的偏差, 即

$$[0141] \quad R_{d, \text{out}}(i) = R_d(i) \text{ 对 } R_d(i) > \Lambda, \text{ 或 } R_d(i) < \frac{1}{\Lambda}。$$

[0142] 4.2.1 步骤式解

[0143] 可采用简单而快速的一步骤式解来由下述限制容许范围以外的全部呈现值:

$$[0144] \quad \tilde{R}_d(i) = \Lambda \bar{R} \text{ 对 } R_d(i) > \Lambda,$$

$$[0145] \quad \tilde{R}_d(i) = \frac{\bar{R}}{\Lambda} \text{ 对 } R_d(i) < \frac{1}{\Lambda}。$$

[0146] 相反地, 在容许范围以内的呈现值可维持不受影响, 使得对此等呈现值 $\tilde{R}(i)$,

$$[0147] \quad \tilde{R}(i) = R(i)。$$

[0148] 4.2.2 迭代重复解

[0149] 另一项可采用的直接方法其中该等具有相关联的偏差的超出范围的呈现值 $R_{d, \text{out}}(i)$ 逐渐受限制。此项算法的迭代重复中, 最大呈现偏差 $R_{d, \text{max}}$ 定义为

$$[0150] \quad R_{d, \text{max}} = \max \{R_{d, \text{out}}(i)\} \text{ 对 } R_d > \Lambda,$$

[0151] $R_{d, \max} = \min \{R_{d, \text{out}}(i)\}$ 对 $R_d < \frac{1}{\Lambda}$

[0152] 对应的呈现系数限制使得

[0153] $\tilde{R}(i) = (1 - \lambda)R(i) + \lambda\tilde{R}$, $\lambda \in (0, 1)$.

[0154] 此项处理可执行直至全部值皆在容许区以内或具有预定迭代重复次数。

[0155] 据此,于各次迭代重复,选定一呈现系数 $R(i_{\max})$,其导数 $R_{d, \text{out}}(i_{\max})$ (例如得自平均值 $\tilde{R}$) 具有最大值 $R_{d, \max}$ 。换言之,选定呈现系数 $R(i_{\max})$,其包含于个别迭代重复得自呈现系数平均 $\tilde{R}$ 的一最大导数 (导数值 $R_{d, \text{out}}$ 表示)。此外,使用前述 $R(i)$ 与 $\tilde{R}$ 的线性组合,该选定的呈现系数 $R(i_{\max})$ 调整至更接近呈现系数的平均。于迭代重复程序的各步骤,可进行自平均值具有最大导数的呈现系数的新颖选择,使得于迭代重复算法的不同步骤可修正不同呈现系数。换言之, $i_{\max}$ 典型地于每次迭代重复时更新。又,平均值可选择性地对迭代重复算法的每个步骤,考虑前一个已修正的呈现系数重新运算。

[0156] 4.3 直接控制

[0157] 直接控制方法的潜在假说考虑失真电平与 TC 偏离其时间均值的偏差间的关系。此点基于观察到比较其他对象,更特定的衰减 / 增强施加至一特定对象,由 SAOC 解码器 / 转码器执行由 TC 对所传输的下混信号的更积极修正。换言之:若 TC 值异常地大,则获得结论 SAOC 算法试图通过施加强力增强而将具有小功率的一对象信号修正成由其他具大功率的对象信号主控的一输出信号。相反地,若 TC 值异常地小,则获得结论 SAOC 算法试图通过施加强力衰减而将具有大功率的一对象信号修正成由其他具小功率的对象信号主控的一输出信号。两种情况下,在 SAOC 的输出端有产生无法接受地低信号质量的高风险。如此,中心思想是防止 TC 大为偏离平均值。

[0158] 此种 PLS 可视为时间及频率变量,原因在于其包含与 SAOC 信号参数 (例如 OLD、IOC) 及转码 / 解码处理的试探性元素的全部相依性。

[0159] 并未丧失一般性,后文描述基于考虑单声道上混的组态。

[0160] 基于 SAOC 输出信号 TC $T(k)$ 具有频率指针 k , PLS 通过以修正的 TC 值置换 TC 极值 (例如在容许区间以外的转码系数),及然后由实际 SAOC 呈现方法使用之来防止 TC 的极值。已修正 TC 值 $\tilde{T}(k)$ 可以如下函数导算:

[0161] $\tilde{T}(k) = F_T(T(k), \Lambda)$,

[0162] 此处 Λ 为 PLS 控制参数 (即临界值)。PLS 控制参数可视为容许参数。

[0163] 因 TC 为时间变量,故应用递归低通滤波器来计算均值:

[0164] $\tilde{T}_n(k) = \mu T_n(k) + (1 - \mu)\tilde{T}_{n-1}(k)$.

[0165] 均值 $\tilde{T}$ 被视为平均值,其中个别转码值的加权由施加递归低通滤波而导入。

[0166] 此处, n 表示 TC 的时间指标,而 $\mu \in (0, 1]$ 为平均参数。已修正 TC 值 $\tilde{T}(k)$ 的容许范围定义为:

[0167]
$$\frac{\tilde{T}(k)}{\Lambda} \leq \tilde{T}(k) \leq \Lambda \tilde{T}(k).$$

[0168] 注意如此是与 TC 限制运算相对应,其相对于参考值进行运算,其是自 TC 而非特定预定值由动态运算。

[0169] 对所述 PLS 办法,最佳解可调配为最小限解,对该最小限解,给定 TC $T(k)$ 与已修正 (已限制) TC $\tilde{T}(k)$ 值间的差为最小化:

[0170]
$$\|\tilde{T}(k) - T(k)\| \rightarrow \min.$$

[0171] 后文中,将描述此一问题的可能的解算法。

[0172] 4.3.1 解算法

[0173] 已修正 TC 值 $\tilde{T}(k)$ 可获得为:

[0174]
$$\tilde{T}(k) = \Lambda \tilde{T}(k) \text{ 对 } T(k) > \Lambda,$$

[0175]
$$\tilde{T}(k) = \frac{T(k)}{\Lambda} \text{ 对 } T(k) < \frac{1}{\Lambda}.$$

[0176] 4.3.2 转码系数实例

[0177] 前文讨论的用于转码系数的参数限制方案可应用至不同转码系数,其例如用于前文讨论的 SAOC 解码器及 SAOC 转码器。

[0178] 举例言之,用于转码系数的参数限制方案可应用至混合矩阵 G 的限制参数,其系用于装置 300 的信号处理器 330。此种情况下,在混合矩阵 G 的一给定矩阵位置的混合矩阵元可取代转码系数 $T(k)$,其中 k 为频率指针。混合矩阵 G' 的对应混合矩阵元可与经调整的转码系数 $\tilde{T}(k)$ 相对应。转码参数限制方案例如可个别施加至混合矩阵的不同矩阵位置。

举例言之,若混合矩阵 G 包含混合矩阵元 g_{11} 、 g_{12} 、 g_{21} 及 g_{22} ,及经调整的混合矩阵 G' 包含混合矩阵元 g_{11}' 、 g_{12}' 、 g_{21}' 及 g_{22}' ,经调整的混合矩阵元 $g_{11}'(n_0)$ 可自一序列 $g_{11}(1)$ 至 $g_{11}(n_0)$ 导算出。相当导算可用于经调整的混合矩阵 G' 的其他混合矩阵元 g_{12}' 、 g_{21}' 及 g_{22}' 。

[0179] 图 10 的表提供对全部 SAOC 运算模式,由所提示的参数限制方案可修正、例如可限制的一转码系数列表。图 10 的表显示不同 SAOC 模式于第一栏 1010。图 10 的表进一步显示可由所提示的参数限制方案修正 (例如限制) 的参数于第二栏 1020。第三栏 1030 显示参考文献 [8] 的 MPEG SAOC FCD 文件的相对应子类别的参考文献。要言之,图 10 的表显示使用参考文献 [8] 的 MPEG SAOC FCD 文件的相对应子类别的参考文献,对全部 SAOC 运算模式,由所提示的参数限制方案可修正 (例如可限制) 的一转码系数列表。

[0180] 4.4 参数限制方案用于限制相对偏差的通式

[0181] 存在有前文讨论的 PLS 的一通式。此式可以如下最小化问题形式对通用参数变量 X_i 表示为:

$$[0182] \quad \begin{cases} \frac{\bar{X}_i}{\Lambda} \leq \tilde{X}_i \leq \Lambda \bar{X}_i, \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases}$$

[0183] 此处, 初步给定 X_i 值, “参考”值 $\bar{X}_i$ 可估算为已修正的 $\tilde{X}_i$ 变量的函数为 $\bar{X}_i = F(\tilde{X}_i)$ 。

[0184] 前文中, 参数变数 X_i 例如可与 $R(i)$ 或 $T(i)$ 相同。同理, 经调整的参数变数 $\tilde{X}_i$ 可与经调整的呈现系数 $\tilde{R}(i)$ 或经调整的转码系数 $\tilde{T}(i)$ 相同。变数 X_i 、 $\tilde{X}_i$ 例如可相于混合矩阵元 $g_{mn}(i)$ 及 $g'_{mn}(i)$ 。

[0185] 后文将讨论两种解算法。

[0186] 大致上, 用以对此种最小限问题获得正确解的分析办法系需要运算。但虽言如此, 仍有简单快速的替代的道可提供次最佳结果, 而仍然的用于 PLS 目的。其中两种简单办法说明于此处。

[0187] 4.4.1 步骤式解

[0188] 步骤式解系基于假设 $\tilde{X}_i \approx F(X_i)$ 限制全部在容许范围以外的全部数值系在其外侧,

$$[0189] \quad \tilde{X}_i = \Lambda \bar{X}_i \text{ 对 } X_i > \Lambda,$$

$$[0190] \quad \tilde{X}_i = \frac{\bar{X}_i}{\Lambda} \text{ 对 } X_i < \frac{1}{\Lambda}.$$

[0191] 容许范围以内的数值 (可视为容许区间) 例如可维持不变。

[0192] 4.4.2 重复迭代解

[0193] 于各步骤, 重复迭代解修正一个所选超出范围的值 X_{i*} 至 $\tilde{X}_{i*}$

$$[0194] \quad \tilde{X}_{i*} = (1-\lambda) X_{i*} + \lambda \bar{X}_i \text{ 其中 } \lambda \in (0, 1)$$

[0195] 例如, 处理指标 i^* 可使用下列条件选择:

$$[0196] \quad X_{i*} = \max\left(\frac{X_i}{\bar{X}_i}\right) \text{ 及 } \frac{X_i}{\bar{X}_i} > \Lambda, \text{ 或}$$

$$[0197] \quad X_{i*} = \min\left(\frac{X_i}{\bar{X}_i}\right) \text{ 及 } \frac{X_i}{\bar{X}_i} < \frac{1}{\Lambda}.$$

[0198] 重复迭代次数可设定为某一值或自该算法隐含地导出。

[0199] 须注意全部此等方法皆可应用于如前述限制 RC 及 TC。

[0200] 4.5 通用线性公式

[0201] 对前文讨论的 PLS 存在有通用线性公式。前一章节中, 通用参数 X_i 的偏差描述为

比 $\frac{X_i}{\bar{X}}$ 。相反地,也可定义为 $\|X_i - \bar{X}_i\|$, 结果导致对通用参数变量 $\tilde{X}_i$ 如下的最小化问题:

$$[0202] \quad \begin{cases} (\bar{X}_i - \Lambda_{x-}) \leq \tilde{X}_i \leq (\bar{X}_i + \Lambda_{x+}), \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases}$$

[0203] 此处, 初步给定 X_i 值, 及“参考”值 $\bar{X}_i$ 可估算为已修正的 $\tilde{X}_i$ 变量的函数为 $\bar{X}_i = F(\tilde{X}_i)$ 。

[0204] 后文中, 将描述此一问题的两个解算法。

[0205] 一般而言, 获得此种最小化问题的正确解的分析办法通常具有运算需求。虽言如此, 仍有简单且快速的替代之道来提供非最佳解而仍然适用于 PLS 目的。其中两种简单办法描述于此处:

[0206] 4.5.1 一步骤式解

[0207] 一步骤式解基于假设: $\tilde{X}_i \approx F(X_i)$ 限制在容许范围以外的全部值皆落入其内定义为:

$$[0208] \quad \tilde{X}_i = \min(\max(X_i, \bar{X}_i - \Lambda_{x-}), \bar{X}_i + \Lambda_{x+}).$$

[0209] 4.5.2 重复迭代解

[0210] 于各步骤, 若 X_{i*} 在容许范围以外, 则重复迭代解修正一个所选的值 X_{i*} 至 $\tilde{X}_{i*}$:

$$[0211] \quad X_{i*} > \bar{X}_{i*} \text{ 及 } \|X_{i*} - \bar{X}_{i*}\| > \Lambda_{x+} \Rightarrow \tilde{X}_{i*} = X_{i*} - S,$$

$$[0212] \quad X_{i*} < \bar{X}_{i*} \text{ 及 } \|X_{i*} - \bar{X}_{i*}\| < \Lambda_{x-} \Rightarrow \tilde{X}_{i*} = X_{i*} + S.$$

[0213] 举例言之, 处理指数 i^* 可使用如下条件选定: $\|X_{i*} - \bar{X}_{i*}\| \geq \|X_i - \bar{X}_i\|$ 及修正阶大小值为 $S = \lambda \|X_{i*} - \bar{X}_{i*}\|$, 具有 $\lambda \in (0, 1)$ 。迭代重复次数可设定为某个值或暗示地自该算法导出。

[0214] 此一算法提供使用容许范围的弹性方式, 即其动态地改变 (取决于 X_{i*})。

[0215] 须注意全部此等方法皆可应用于如前述限制 RC 及 TC。

[0216] 另外, 可使用如下算法:

[0217] 若 $X_{i*} > \bar{X}_{i*}$ 并且 $\|X_{i*} - \bar{X}_{i*}\| > \Lambda_{x+}$, 则

$$[0218] \quad \tilde{X}_{i*} = X_{i*} - S, \text{ 以及}$$

[0219] 若 $X_{i*} < \bar{X}_{i*}$ 并且 $\|X_{i*} - \bar{X}_{i*}\| > \Lambda_{x-}$, 则

$$[0220] \quad \tilde{X}_{i*} = X_{i*} + S$$

[0221] 此一算法版本使用固定（静态）容许范围 Λ_x , Λ_{x+} 。

[0222] 4.6 额外备注

[0223] 须注意全部此等方法皆可应用于限制呈现系数及转码系数,说明如前。

[0224] 5 参数限制方案应用至多声道下混 / 上混情况

[0225] 考虑下混 / 上混声道的任一种组合,单声道下混 / 单声道上混情况的单一 TC PLS(例如直接控制)扩充至 TC 矩阵。结果,直接控制可个别地应用至各个 TC。多声道上混情况用于 RC PLS(例如间接控制)例如可于单多重单声道办法实现,此处全部个别呈现系数皆独立处理。

[0226] 6 收听测试结果

[0227] 6.1 测试设计及项目

[0228] 已经进行主观收听测试来评估所提示的失真控制测量 (DCM) 构想的听觉性能,且与常规 SAOC 参考模型 (SAO CRM) 解码处理比较。

[0229] 测试设计包括所提示的参数限制方案及其组合的直接及间接控制办法。常规(未由参数限制方案 PLS 处理的)SAOC 解码器的输出信号包括于该测试来验证 SAOC 的基准线性能。此外,与下混信号相对应的微不足道的呈现情况用于收听测试作为比较目的。

[0230] 图 5a 的表描述收听测试条件。

[0231] 已经自提案 (CfP) 收听测试材料中选出四项代表极端呈现状况的典型及最关键性假象类型用于目前收听测试。

[0232] 图 5b 的表描述收听测试的音频项目。

[0233] 依据图 6 的表的呈现对象增益已经应用于所考虑的上混情况。

[0234] 因所提示的 PLS 使用常规 SAOC 位流及下混信号运算(无需 SAOC 编码器端的任何 PLS 相关活性)且未转接残余信息,故无核心编码器应用至相对应 SAOC 下混信号。

[0235] 对全部测试项目及所考虑的呈现条件,PLS 的通用设定值取作为:

[0236] $\Lambda_{\{R, R+\}} = \Lambda_{\{T, T+\}} = 6$.

[0237] 6.2 测试方法

[0238] 本收听测试于设计来允许高质量收听的隔音收听室内进行。使用耳机 (STAX SR λ Pro, 附有湖人 (Lake-People) D/A- 转换器及 STAX SRM 监视器) 进行回放。

[0239] 测试方法遵照空间音频验证测试所用程序,基于“隐藏参考及基准的多重刺激”(MUSHRA)法用于中间质量音频的主观评估 [7]。测试方法据此修正来评估所提示的 DCM 构想的听觉性能。依据所采用的测试方法,指示收听者依据下列收听测试指示而比较全部测试条件:

[0240] 对各项音频请:

[0241] • 首先研读期望的混音说明,您作为系统使用者,想要实现:

[0242] 项目「BlackCoffee」:混音中有轻柔喇叭小节

[0243] 项目「Fanta4」:混音中有强鼓声

[0244] 项目「LovePop」:混音中有轻柔弦乐小节

[0245] 项目「试音」:轻音乐及强噪音

[0246] • 然使用一个公共等级描述二者来分级信号

[0247] - 实现期望的混音目标

[0248] - 全场景音质（考虑失真、假象、不自然…）

[0249] 共有九位收听者参考各项测试。全部个体皆视为经验老练的收听者。

[0250] 测试条件对各个测试项目及各个收听者自动随机分配。以自 0 至 100 范围的分数由基于计算机的 MUSHRA 程序记录主观反应。允许接受测试各项目间的瞬间切换。

[0251] 6.3 收听测试结果

[0252] 以图解验证所得收听测试结果的简短概述可参考附录。此等作图显示对全部收听者对每个项目的平均 MUSHRA 分级及对全部评估项目的统计均值连同相关 95% 信赖区间。

[0253] 基于所进行收听测试结果可做出下列观察：对全部所进行收听测试结果，所得 MUSHRA 分数证实就总统计均值而言，所提示的 PLS 功能提供比较常规 SAOC RM 系统更佳的性能。须注意由常规 SAOC 解码器（对所考虑的极端呈现条件，显示强音频假象）所产生的全部项目质量分级，比较丝毫也未满足期望的呈现情况的下混相同呈现设定值的质量仅略高。因此，可获得结论：所提示的 PLS 结果导致对全部所考虑的收听测试情况，主观信号质量皆有显著改良。也可获得结论：最具展望的限制系统由 RC 及 TC PLS 的组合所组成。

[0254] 有关收听测试结果的细节可参考图 7 的图解表示形态。

[0255] 7 替代实施例

[0256] 虽然于装置上下文已经说明若干方面，但显然此等方面也表示相对应方法的描述，此处一方块或一装置与一方法步骤或一方法步骤的一特征相对应。同理，于一方法步骤上下文所描述的方面也表示相对应方块或项目或相对应装置的特征的描述。部分或全部方法步骤可由（或使用）硬件装置，例如微处理器、可程序计算机或电子电路执行。若干实施例中，最重要方法步骤中的某一者或多者可由此种装置执行。

[0257] 本发明的编码音频信号可储存于数字储存介质或可经由传输介质诸如无线传输介质或有线传输介质诸如因特网传输。

[0258] 依据某些实施要求，本发明的实施例可于硬件或于软件实施。实施的执行可使用有可电子式读取的控制信号储存其上的数字储存介质例如软盘、DVD、蓝光盘、CD、ROM、PROM、EPROM、EEPROM 或闪存，该等介质与可程序规划计算机系统协力合作（或可协力合作）因而执行个别方法。因此，数字储存介质可为计算机可读取式。

[0259] 依据本发明的若干实施例包含具有可电子式读取的控制信号于其上的数据载体，其与可程序规划计算机系统可协力合作因而执行此处所述方法中的一者。

[0260] 一般而言，本发明的实施例可实施为带有程序代码的计算机程序产品，该程序代码可操作当该计算机程序产品于计算机上运行时用于执行该等方法中的一者。程序代码例如可储存于机器可读取载体上。

[0261] 其他实施例包含用以执行此处所述方法中的一者的储存在机器可读取载体上的计算机程序。

[0262] 换言之，因而本发明方法的实施例为一种具有程序代码的计算机程序，当该计算机程序产品于计算机上运行时用以执行此处所述方法中的一者。

[0263] 因而本发明方法的又一实施例为一种数据载体（或数字储存介质，或计算机可读取介质）包含用以执行该等方法中的一者的计算机程序记录于其上。该数据载体或数字储存介质或记录介质典型地为有实体及 / 或非瞬时的。

[0264] 因此，本发明方法的又一实施例为一种数据流或一序列信号表示用以执行此处所

述方法中的一者的计算机程序。该数据流或该序列信号例如可被配置为经由数据通信连接,例如经由因特网传输。

[0265] 又一实施例包含一种处理装置,例如计算机或可程序逻辑装置被配置为或调整适用于执行此处所述方法中的一者。

[0266] 又一实施例包含一种计算机,其上安装用以执行此处所述方法中的一者的计算机程序。

[0267] 于若干实施例,可编程逻辑装置(例如现场可编程门阵列)可用来执行此处所述方法的部分或全部函数。于若干实施例,现场可编程门阵列可与微处理器协力合作来执行此处所述方法中的一者。大致上,该方法优选由硬件装置执行。

[0268] 前述实施例仅供举例说明本发明的原理。须了解本领域技术人员显然易知此处所述配置及细节的修正及变化。因此意图本发明只受随附的权利要求书的范围所限,而非受通过此处实施例的描述及解说所呈现的特定细节所限。

[0269] 8 结论

[0270] 依据本发明的实施例提供用于音频解码器的失真控制的参数限制方案。依据本发明的若干实施例聚焦在空间音频对象编码(SAOC),其提供用以选择期望的回放设定值(例如单声道、立体声、5.1等)的使用者界面手段以及经由依据个人偏好或其他标准而控制呈现矩阵的期望输出呈现场景的交互式实时修正。但一般而言调整所提示的方法用于参数技术为直接任务。

[0271] 由于基于下混/分离/混合参数办法,所呈现的音频输出信号的主观质量系取决于呈现参数设定值。选用由使用者选择呈现设定值有使用者选择不当对象呈现选项的风险,诸如总体声音场景内部的对象之极端增益操控。

[0272] 对商业产品而言,绝对无法接受在使用者界面上产生任何设定的不佳音质及/或音频假象。为了控制所产生的SAOC音频输出信号的过度降级,已经描述若干运算措施,其基于运算所呈现的场景的听觉质量测量值,及依据此测量值(及其他信息),修正实际施加呈现系数(例如请见参考文献[6])。

[0273] 本发明提供替代构想用来保护所呈现的SAOC场景的主观音质:

[0274] • 全部处理全然在SAOC解码器/转码器内部进行,及

[0275] • 未涉及所呈现的音频场景的听觉音质的复杂测量值的显式(explicit)计算

[0276] 如此此等构想可以结构简单而又极端有效方式在SAOC解码器/转码器内部实施。因所提示的失真控制机制(DCM)针对SAOC解码器特有的限制参数,即呈现系数(RC)及转码系数(TC),故于全文说明中称作为参数限制方案(PLS)。

[0277] 但参数限制方案也可应用于任一种不同的音频解码器。

[0278] 9 参考文献

[0279] [1]C.Faller and F.Baumgarte," Binaural Cue Coding-Part II :Schemes and applications", IEEE Trans.on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.

[0280] [2]C.Faller, " Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006, Preprint 6752.

[0281] [3]J.Herre, S.Disch, J.Hilpert, O.Hellmuth: " From SAC To SAOC-Recent Developments in Parametric Coding of Spatial Audio ", 22nd Regional UK AES

Conference, Cambridge, UK, April 2007.

[0282] [4] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding(SAOC)-The Upcoming MPEG Standard on Parametric Object Based Audio Coding" ,124th AES Convention, Amsterdam 2008, Preprint 7377.

[0283] [5] ISO/IEC, " MPEG audio technologies-Part 2:Spatial Audio Object Coding (SAOC), " ISO/IEC JTC1/SC29/WG11(MPEG)FCD 23003-2.

[0284] [6] US patent application 61/173,456, METHODS, APPARATUS, AND COMPUTER PROGRAMS FOR DISTORTION AVOIDING AUDIO SIGNAL PROCESSING

[0285] [7] EBU Technical recommendation: " MUSHRA-EBU Method for Subjective Listening Tests of Intermediate Audio Quality" , Doc. B/AIM022, October 1999.

[0286] [8] ISO/IEC JTC1/SC29/WG11(MPEG), Document N10843, "Study on ISO/IEC 23003-2:200x Spatial Audio Object Coding(SAOC)", 89th MPEG Meeting, London, UK, July 2009.

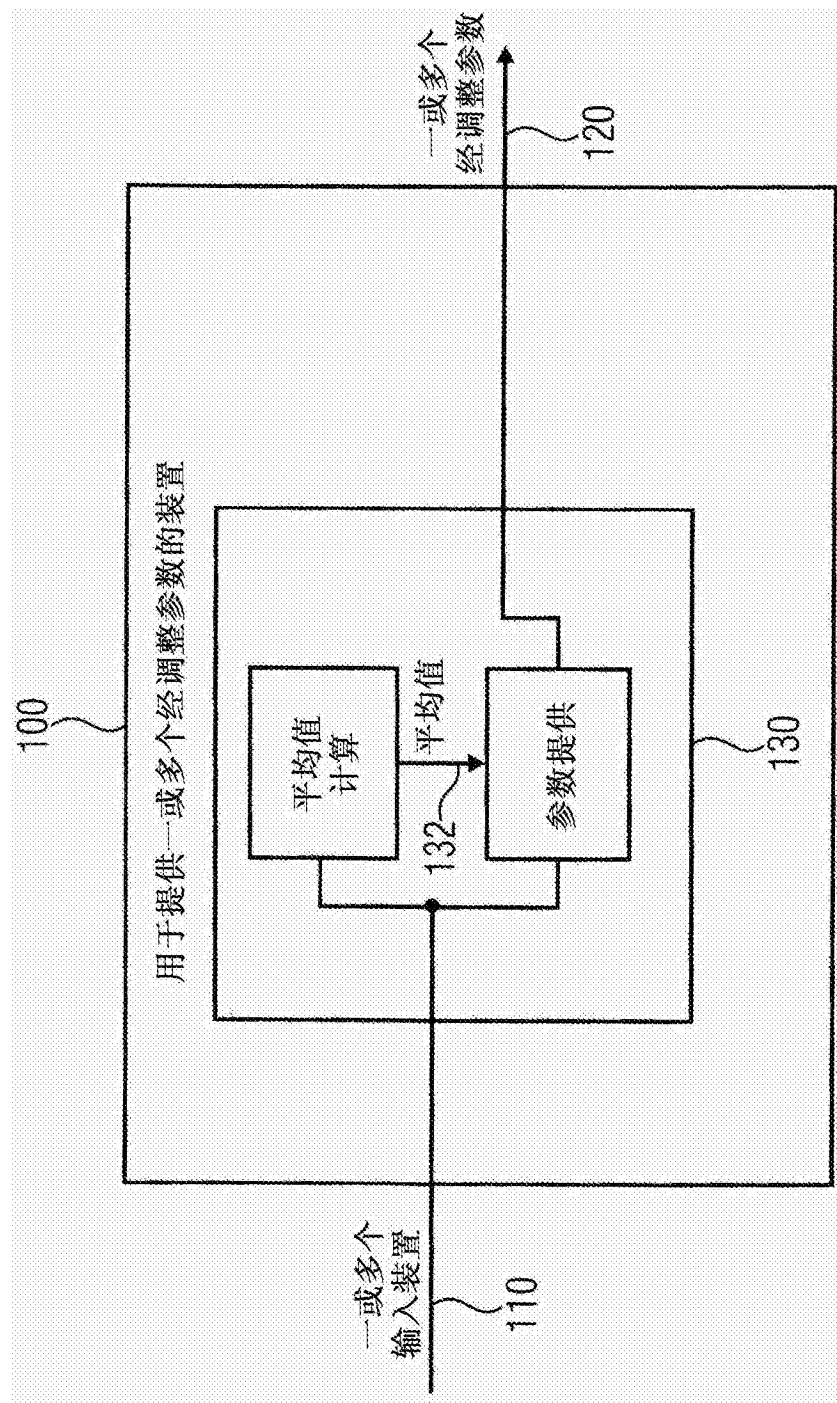


图 1

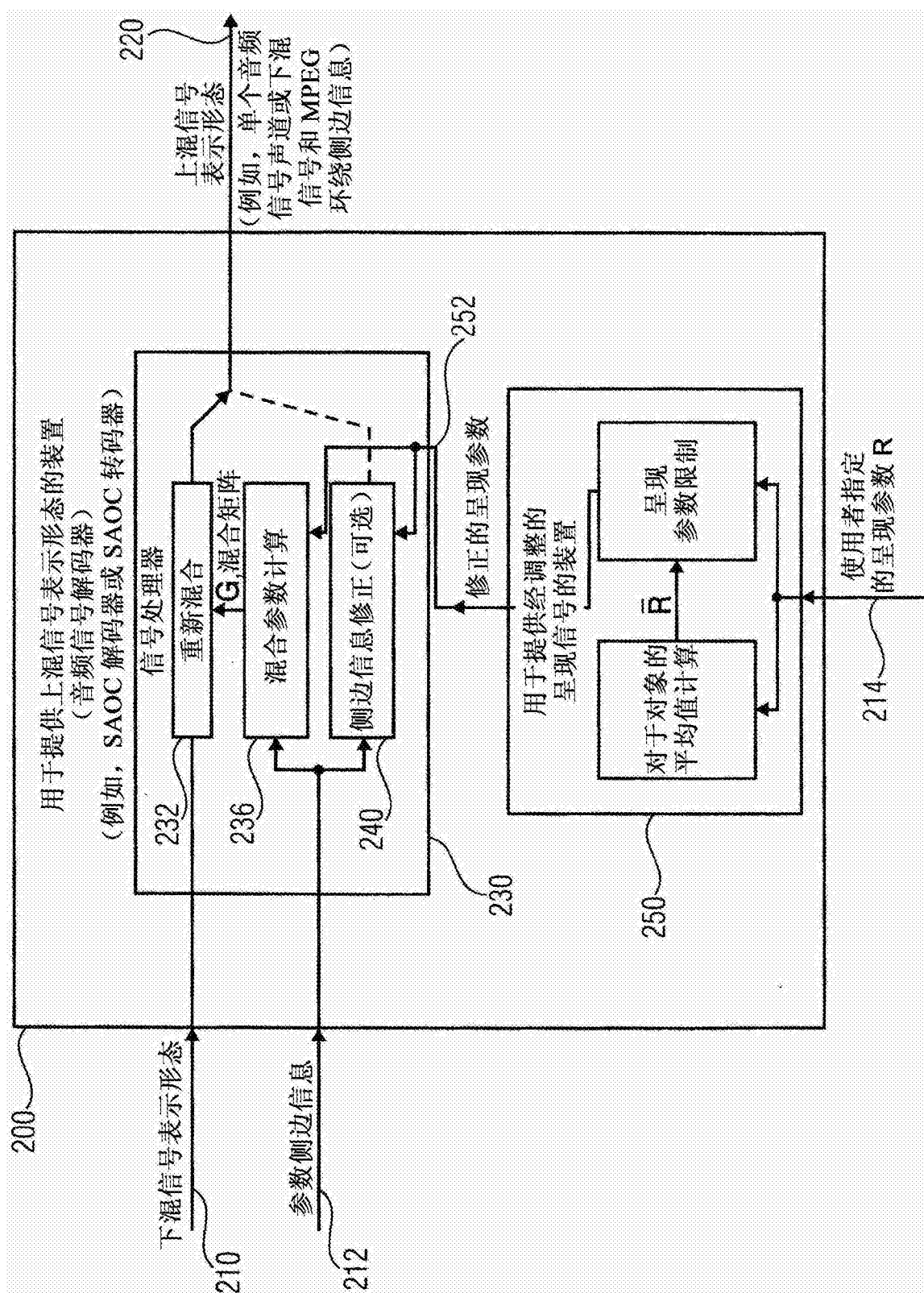


图 2

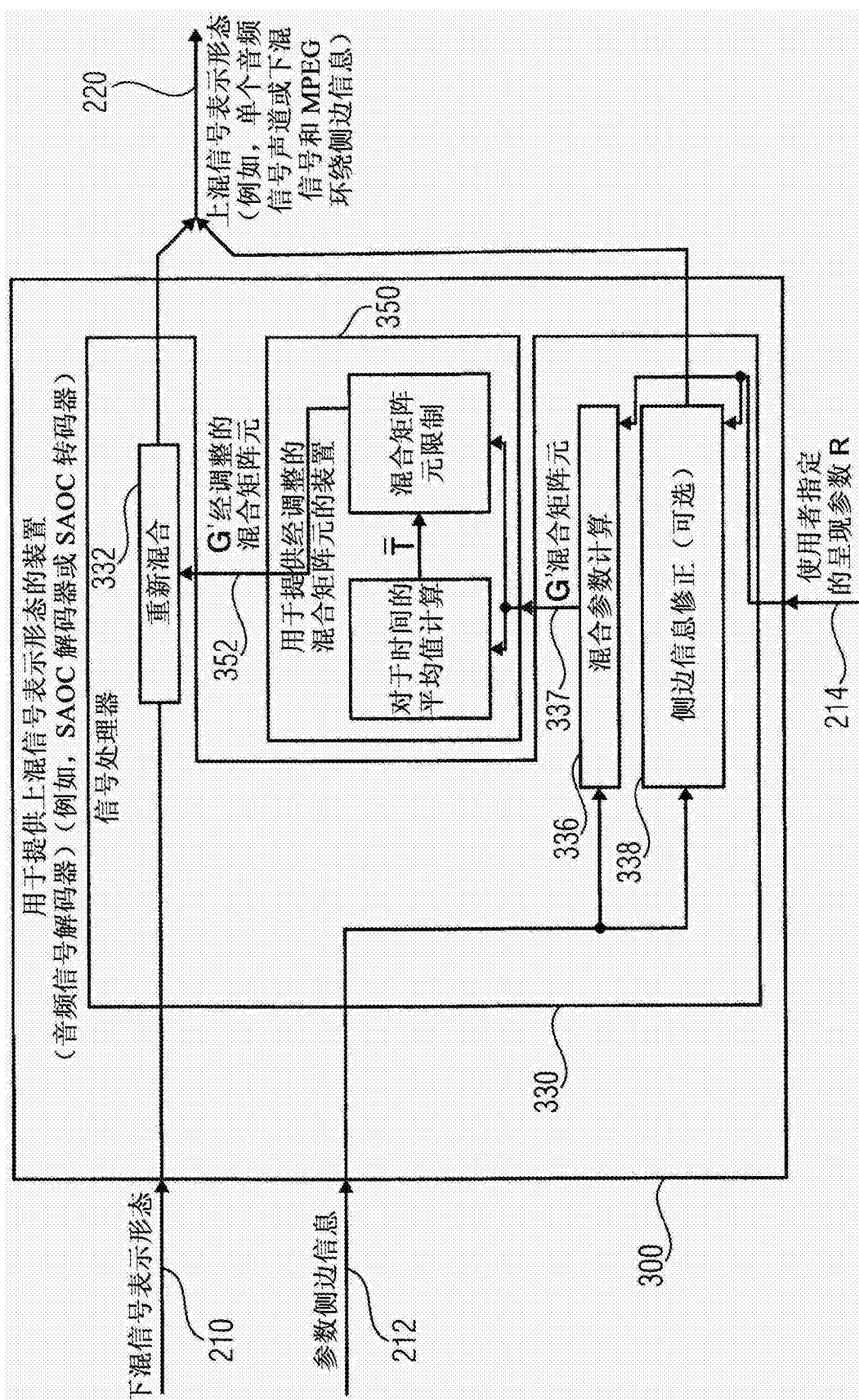


图 3

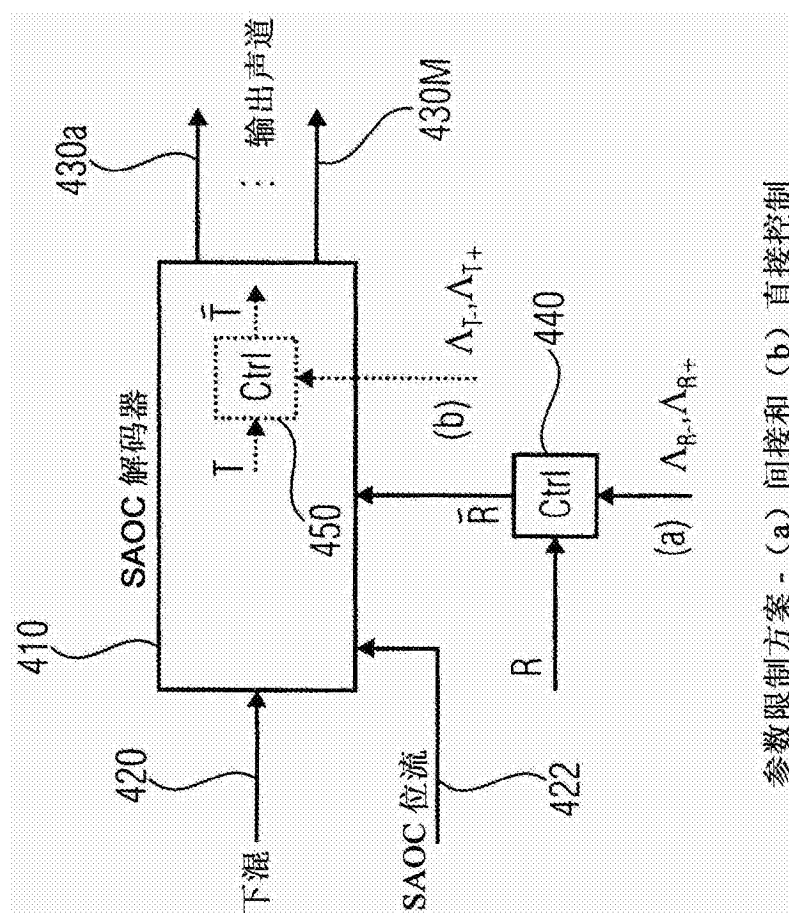


图 4

表 - 收听测试条件

编码器名	描述
"DMX"	常规 SAOC 解码器的平凡下混类似呈现信号
"noLim"	常规 (未由 PLS 处理) SAOC 解码器的输出
"TC"	通过 TC 直接控制 PLS 的 SAOC 解码器的输出
"R0"	通过 TC 间接控制 PLS 的 SAOC 解码器的输出
"R0_TC_1"	通过组合 TC/RC 控制 PLS 的 SAOC 解码器的输出

图 5A

表 - 收听测试的音频项目

收听项目	呈现类型和矩阵	经时
"BlackCoffee"	混音中有轻柔喇叭小节	7 秒
"Fanta4"	混音中有强鼓声	7 秒
"LovePop"	混音中有轻柔弦乐小节	7 秒
"试音"	轻音乐及强噪音	13 秒

图 5B

表 - 测试的极端呈现条件

收听项目	对象 #	音频对象描述	呈现	系数
"BlackCoffee"	1	铜管 2	0.0001	~-80dB
	2	铜管 1	0.0001	~-80dB
	3	风琴	3.1623	~10dB
	4	鼓 & 低音乐器	3.1623	~10dB
	5	打击乐器	3.1623	~10dB
"Fanta4"	1	人声	0.2512	~-12dB
	2	低音乐器	0.2512	~-12dB
	3	敲击乐器	2.0	~6dB
"LovePop"	1	鼓	2.5119	~8dB
	2	低音乐器	2.5119	~8dB
	3	电吉他	2.5119	~8dB
	4	非电吉他	2.5119	~8dB
	5	弦乐器	0.0001	~-80dB
"试音"	1	BG 人声	0.1778	~-15dB
	2	低音乐器	0.1778	~-15dB
	3	鼓混响	0.1778	~-15dB
	4	大鼓	0.1778	~-15dB
	5	主音吉他	0.1778	~-15dB
	6	双主唱	3.1623	~10dB
	7	主唱	3.1623	~10dB
	8	架子鼓	0.1778	~-15dB

图 6

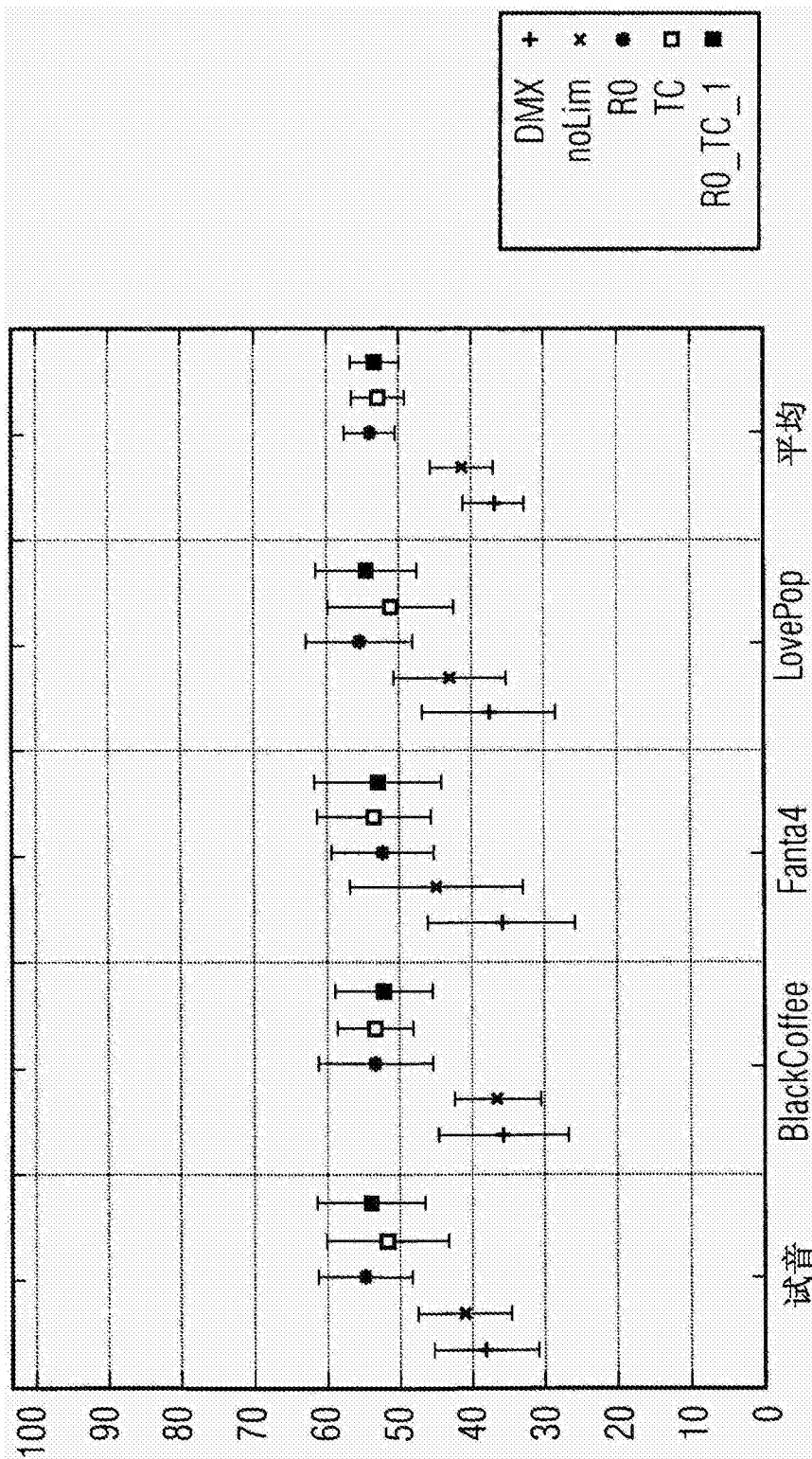
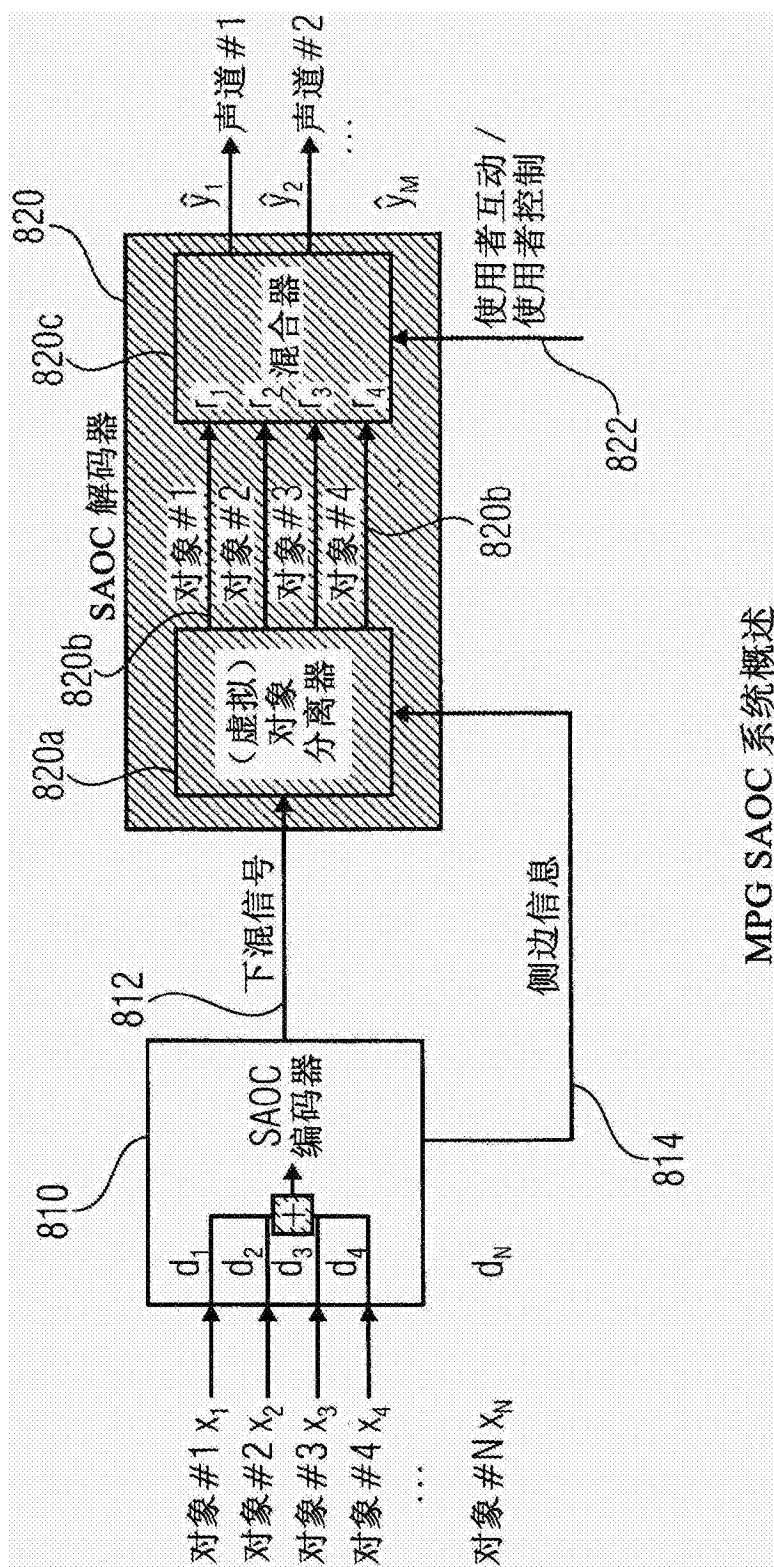


图 - 对于 PLS 的 MUSHRA 收听测试结果

图 7



MPG SAOC 系统概述

图 8

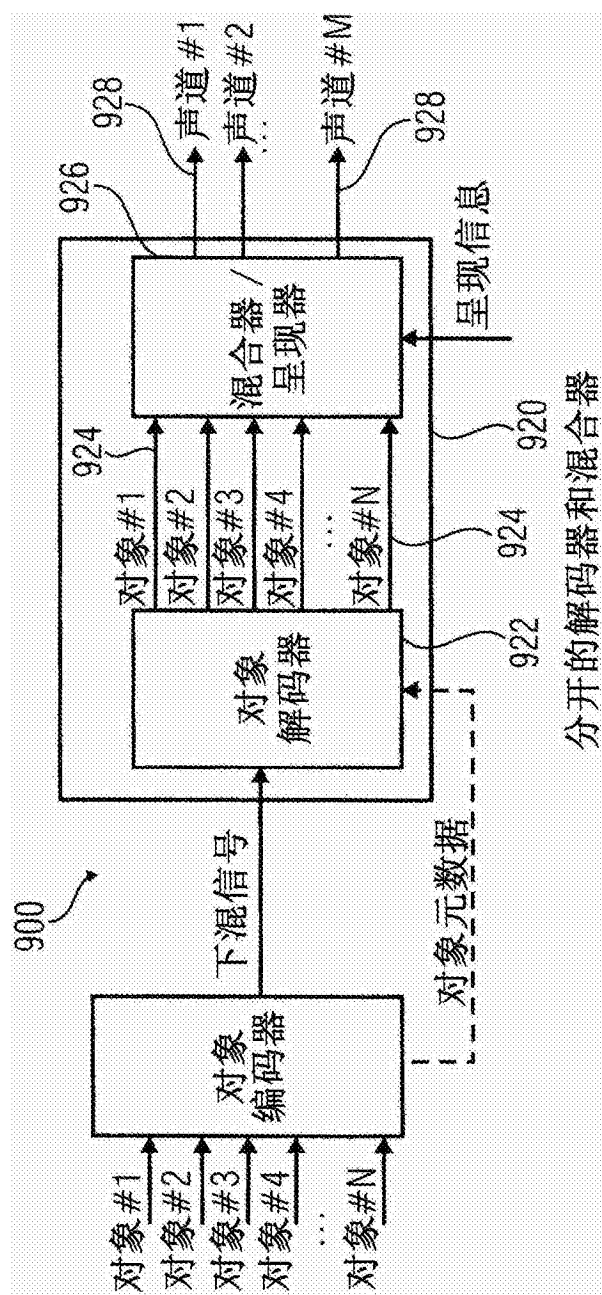


图 9A

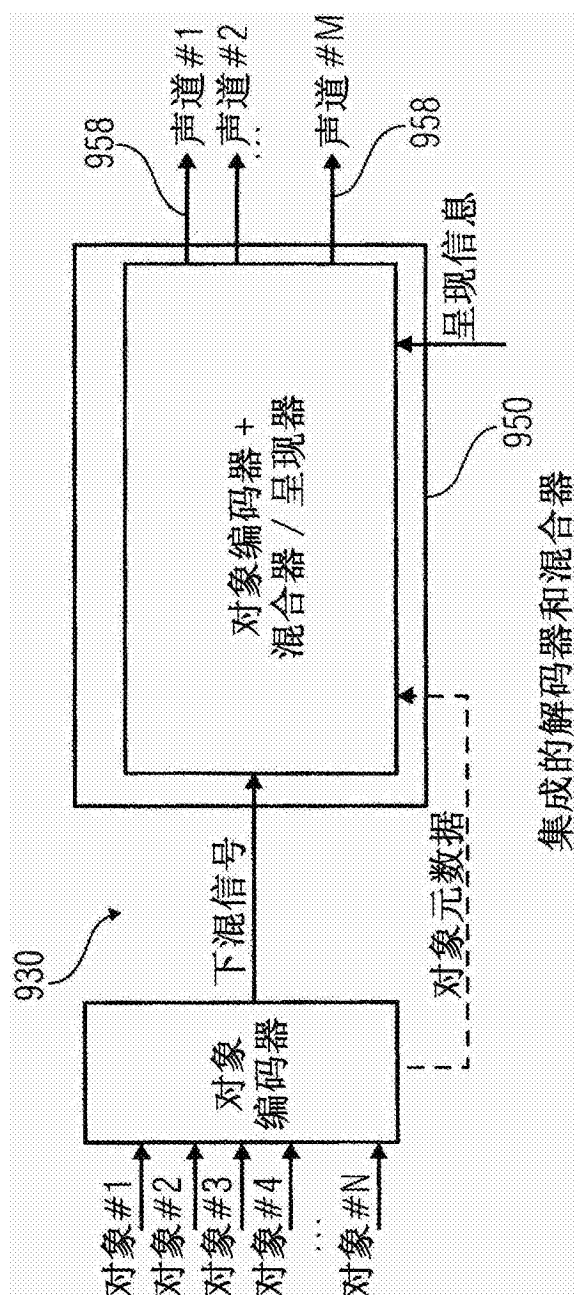


图 9B

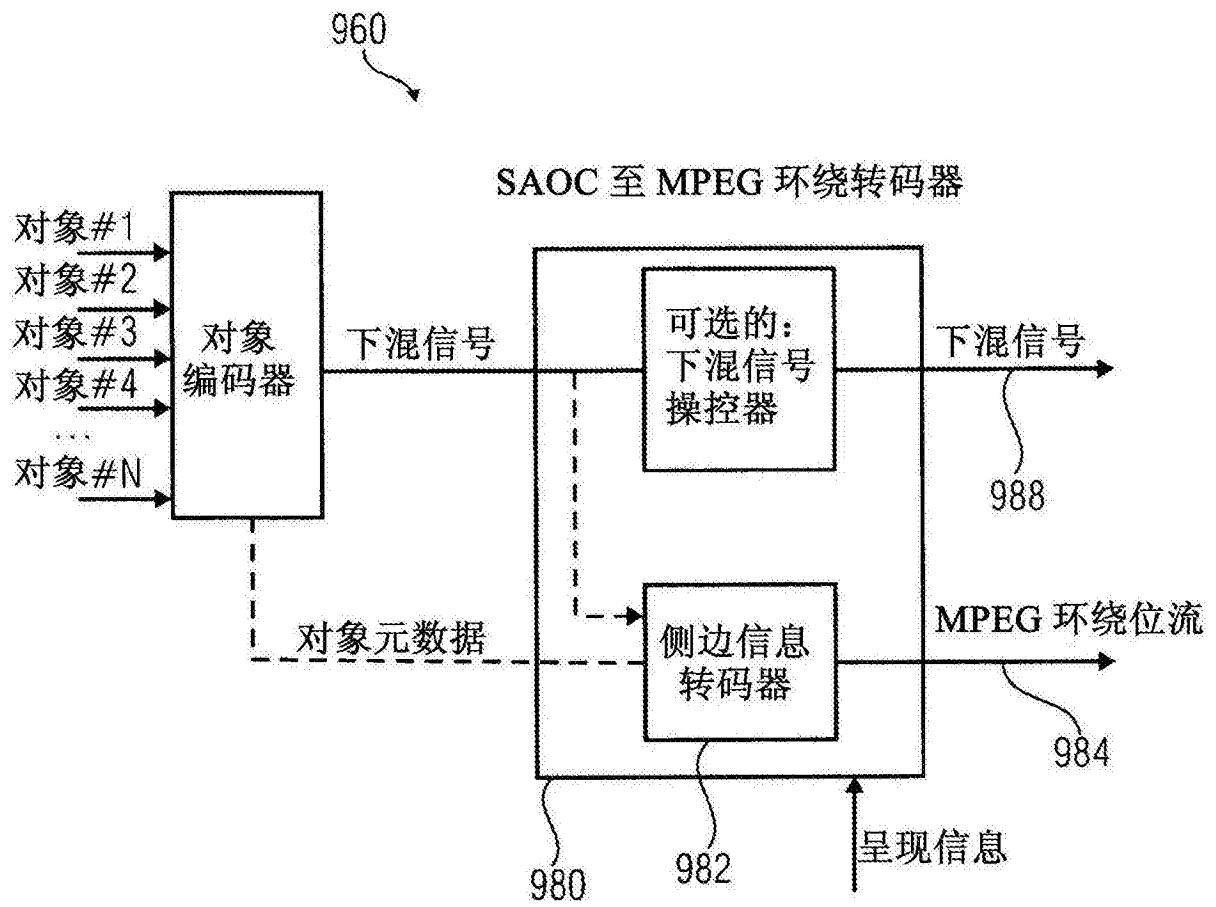


图 9C

1010			1020		1030
SAOC 模式			修正的系数		参考
"x-1-1"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.3
"x-1-2"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.2
"x-1-b"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.1
"x-1-5"转码模式			任意下混增益系数($ADG^{l,m}$)		6.5.2.4
"x-2-1"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.6
"x-2-2"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.5
"x-2-b"解码模式			混合矩阵系数($G^{l,m}$)		6.6.1.4
"x-2-5"转码模式			预测矩阵系数(C_3)		6.5.3.3.1

图 10