



(12) 发明专利

(10) 授权公告号 CN 101151660 B

(45) 授权公告日 2011. 10. 19

(21) 申请号 200680010652. 7

(22) 申请日 2006. 03. 16

(30) 优先权数据

05102515. 3 2005. 03. 30 EP

05103085. 6 2005. 04. 18 EP

(85) PCT申请进入国家阶段日

2007. 09. 29

(86) PCT申请的申请数据

PCT/IB2006/050822 2006. 03. 16

(87) PCT申请的公布数据

WO2006/103584 EN 2006. 10. 05

(73) 专利权人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

专利权人 编码技术股份有限公司

(72) 发明人 G·H·霍索 D·J·布里巴特

E·G·P·舒杰斯 A·C·登布林克

L·F·维勒莫斯 H·普尔恩哈根

K·J·罗登

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 李亚非 谭祐祥

(51) Int. Cl.

G10L 19/00(2006. 01)

H04S 3/00(2006. 01)

(56) 对比文件

WO 03069954 A2, 2003. 08. 21, 全文.

CN 1510662 A, 2004. 07. 07, 全文.

WO 03085645 A1, 2003. 10. 16, 全文.

WO 2004008805 A1, 2004. 01. 22, 全文.

FALLER C. Coding of spatial audio compatible with different playback formats. 《AUDIO ENGINEERING SOCIETY CONVENTION PAPER》. 2004, 1-12.

HERRE J ET AL. MP3 Surround: Efficient and Compatible Coding of Multi-Channel Audio. 《AUDIO ENGINEERING SOCIETY CONVENTION PREPRINT》. 2004, 1-14.

审查员 张鑫

权利要求书 4 页 说明书 12 页 附图 4 页

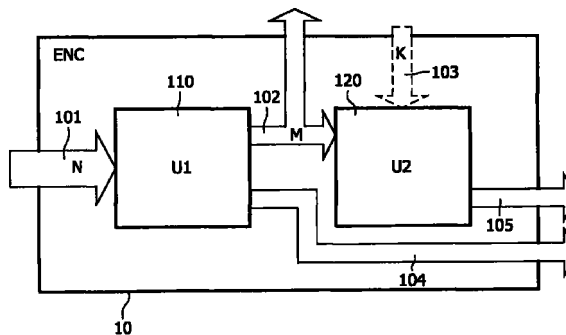
(54) 发明名称

多通道音频编码器、解码器以及相应方法

(57) 摘要

一种多通道音频编码器(10),用于将多通道音频信号(101)编码成空间向下混音(102)和相关参数(104、105),所述多通道音频信号(101)例如为5.1通道音频信号,空间向下混音(102)例如为立体声信号。编码器(10)包括第一和第二单元(110、120)。第一单元(110)将多通道音频信号(101)编码成空间向下混音(102)和参数(104)。这些参数(104)使多通道解码器(20)能够从空间向下混音(102)重建多通道音频信号(203)。第二单元(120)从空间向下混音(102)产生参数(105),这些参数(105)使解码器能够从备选向下混音(103)重建空间向下混音(202),所述备选向下混音(103)例如为录音棚中人工混合的所谓艺术向下混音。这样,解码器(20)能够有效处理这

样的情况,这种情况下接收备选向下混音(103)取代规则的空间向下混音(102)。在解码器(20)中,首先从另一个向下混音(103)和参数(105)重建空间向下混音(202)。接下来,空间向下混音(202)被解码成多通道音频信号(203)。



1. 一种多通道音频编码器 (10), 用于将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 和相关参量数据 (104、105), M 和 N 为整数, $N > M$, $M \geq 1$, 其中多通道音频编码器 (10) 包括:

- 第一单元 (110), 用于将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 和第一相关参量数据 (104), 其中 M 个音频信号 (102) 和第一相关参量数据 (104) 表示所述 N 个音频信号 (101), 以及

- 第二单元 (120), 它耦合到第一单元 (110), 第二单元 (120) 被安排成从所述 M 个音频信号 (102) 产生表示所述 M 个音频信号 (102) 的第二相关参量数据 (105), 所述第二相关参量数据包括使得能够从 K 个另外的音频信号 (103) 重建所述 M 个音频信号 (102) 的修正参数, 所述 K 个另外的音频信号 (103) 是不同于所述 M 个音频信号 (102) 的、所述 N 个音频信号 (101) 的备选向下混音, 其中相关参量数据 (104、105) 包括第一和第二相关参量数据。

2. 根据权利要求 1 的多通道音频编码器 (10), 其中第二单元 (120) 被安排成产生第二相关参量数据 (105), 使得第二相关参量数据 (105) 表示所述 M 个音频信号 (102) 的一个特性。

3. 根据权利要求 1 的多通道音频编码器 (10), 其中第二单元 (120) 被安排成从 M 个音频信号 (102) 并从所述 K 个另外的音频信号 (103) 产生第二相关参量数据 (105), 使得修正参数表示 M 个音频信号 (102) 同所述 K 个另外的音频信号 (103) 之间的一个差异。

4. 根据权利要求 1 的多通道音频编码器 (10), 其中第二单元 (120) 被安排成产生第二相关参量数据 (105), 使得修正参数包括 M 个音频信号 (102) 的特性或者 M 个音频信号 (102) 的特性同 K 个所述另外的音频信号 (103) 的特性之间的差异。

5. 根据权利要求 2 的多通道音频编码器 (10), 其中第二单元 (120) 被安排成产生第二相关参量数据 (105), 使得所述特性包括:

- 至少部分音频信号 (102、103) 的一个能量或功率值; 或者
- 至少部分音频信号 (102、103) 的一个相关性值; 或者
- 至少部分音频信号 (102、103) 的能量或功率值之间的比率。

6. 一种多通道音频解码器 (20), 用于将 K 个音频信号 (103) 及相关参量数据 (104、105) 解码成 N 个音频信号 (203), K 和 N 为整数, $N > K$, $K \geq 1$, 其中 K 个音频信号 (103) 及相关参量数据 (104、105) 表示 N 个音频信号 (203), 且其中多通道音频解码器 (20) 包括:

- 第一单元 (210), 用于从 K 个音频信号 (103) 和相关参量数据 (105) 的至少第一部分重建 M 个另外的音频信号 (102), 所述相关参量数据包括使得能够从 K 个音频信号 (103) 重建所述 M 个另外的音频信号 (202) 的修正参数, 所述 M 个另外的音频信号 (102) 是不同于所述 K 个音频通道 (103) 的、所述 N 个音频通道 (101) 的备选向下混音, M 为整数, $M \geq 1$, 其中相关参量数据 (105) 的第一部分表示所述 M 个另外的音频信号 (202); 以及

- 第二单元 (220), 它耦合到第一单元 (210), 第二单元 (220) 被安排成将所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的至少第二部分解码成 N 个音频信号 (203), 其中所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的第二部分表示 N 个音频信号 (203)。

7. 根据权利要求 6 的多通道音频解码器 (20), 其中相关参量数据 (105) 的第一部分表示所述 M 个另外的音频信号 (202) 的一个特性。

8. 根据权利要求6的多通道音频解码器(20),其中修正参数包括所述M个另外的音频信号(202)的特性或所述M个另外的音频信号(202)的特性同K个音频信号(103)的特性之间的差异。

9. 根据权利要求6的多通道音频解码器(20),其中第一单元(210)被安排成从K个音频信号(103)产生表示K个音频信号(103)的另外的修正参数,其中第一单元(210)进一步被安排成从K个音频信号(103)、包含在相关参量数据(105)的第一部分中的修正参数以及所述另外的修正参数重建所述M个另外的音频信号(202)。

10. 根据权利要求9的多通道音频解码器(20),其中修正参数包括所述M个另外的音频信号(202)的特性,且其中所述另外的修正参数包括K个音频信号(103)的特性,且其中第一单元(210)被安排成从K个音频信号(103)、所述M个另外的音频信号(202)的特性同K个音频信号(103)的特性之间的差异重建所述M个另外的音频信号(202)。

11. 根据权利要求7的多通道音频解码器(20),其中特性包括:

- 至少部分音频信号(103、202)的一个能量或功率值;或者
- 至少部分音频信号(103、202)的一个相关性值;或者
- 至少部分音频信号(103、202)的能量或功率值之间的一个比率。

12. 一种用于将N个音频信号(101)编码成M个音频信号(102)及相关参量数据(104、105)的方法,M和N为整数, $N > M$, $M \geq 1$,其中该方法包括:

- 将N个音频信号(101)编码成M个音频信号(102)和第一相关参量数据(104),其中M个音频信号(102)和第一相关参量数据(104)表示N个音频信号(101);以及
- 从M个音频信号(102)产生表示M个音频信号(102)的第二相关参量数据(105),所述第二相关参量数据包括使得能够从K个另外的音频信号(103)重建所述M个音频信号(102)的修正参数,所述K个另外的音频信号(103)是不同于所述M个音频信号(102)的、所述N个音频信号(101)的备选向下混音,其中相关参量数据(104、105)包括第一和第二相关参量数据。

13. 一种用于将K个音频信号(103)及相关参量数据(104、105)解码成N个音频信号(203)的方法,K和N为整数, $N > K$, $K \geq 1$,其中K个音频信号(103)及相关参量数据(104、105)表示N个音频信号(203),且其中该方法包括:

- 从K个音频信号(103)和相关参量数据(105)的至少第一部分重建M个另外的音频信号(202),所述相关参量数据包括使得能够从K个音频信号(103)重建所述M个另外的音频信号(202)的修正参数,所述M个另外的音频信号(102)是不同于所述K个音频信号(103)的、所述N个音频信号(101)的备选向下混音,M为整数, $M \geq 1$,其中相关参量数据(105)的所述第一部分表示所述M个另外的音频信号(202);以及
- 将所述M个另外的音频信号(202)和相关参量数据(104)的至少第二部分解码成N个音频信号(203),其中所述M个另外的音频信号(202)和相关参量数据(104)的第二部分表示N个音频信号(203)。

14. 一种传输系统(70),包括发送器(40),用于将编码多通道音频信号经由传输信道(30)发送到接收器(50),发送器(40)包括根据权利要求1的多通道音频编码器(10),该多通道音频编码器(10)用以将N个音频信号(101)编码成M个音频信号(102)及相关参量数据(104、105),发送器(40)还包括用于将K个另外的音频信号(103)及相关参量数据

(104、105) 经由传输信道 (30) 发送至接收器 (50) 的装置 (41), 接收器 (50) 包括用于接收所述 K 个另外的音频信号 (103) 和相关参量数据 (104、105) 的装置, 接收器 (50) 还包括根据权利要求 7 的多通道音频解码器 (20), 该多通道音频解码器 (20) 用以将所述 K 个另外的音频信号 (103) 和相关参量数据 (104、105) 解码成 N 个音频信号 (203)。

15. 一种用于发送编码多通道音频信号的发送器 (40), 所述发送器 (40) 包括根据权利要求 1 的多通道音频编码器 (10), 该多通道音频编码器 (10) 用以将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 及相关参量数据 (104、105), 发送器 (40) 还包括用于发送 K 个另外的音频信号 (103) 及相关参量数据 (104、105) 的装置。

16. 一种用于接收编码多通道音频信号的接收器 (50), 所述接收器 (50) 包括用于接收 K 个音频信号 (103) 及相关参量数据 (104、105) 的装置 (51), 接收器 (50) 还包括根据权利要求 6 的多通道音频解码器 (20), 该多通道音频解码器 (20) 用于将 K 个音频信号 (103) 和相关参量数据 (104、105) 解码成 N 个音频信号 (203)。

17. 一种用于发送和接收编码多通道音频信号的方法, 该方法包括将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 和相关参量数据 (104、105), M 和 N 为整数, $N > M$, $M \geq 1$, 其中编码包括:

- 将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 和第一相关参量数据 (104), 其中 M 个音频信号 (102) 和第一相关参量数据 (104) 表示 N 个音频信号 (101); 以及

- 从 M 个音频信号 (102) 产生表示该 M 个音频信号 (102) 的第二相关参量数据 (105), 所述第二相关参量数据包括使得能够从 K 个另外的音频信号 (103) 重建所述 M 个音频信号 (102) 的修正参数, 所述 K 个另外的音频信号 (103) 是不同于所述 M 个音频信号 (102) 的、所述 N 个音频信号 (101) 的备选向下混音, 其中相关参量数据 (104、105) 包括第一和第二相关参量数据,

该方法还包括发送和接收 K 个音频信号 (103) 及相关参量数据 (104、105)、将 K 个音频信号 (103) 和相关参量数据 (104、105) 解码成 N 个音频信号 (203), 该解码包括:

- 从 K 个音频信号 (103) 和相关参量数据 (105) 的至少第一部分重建 M 个另外的音频信号 (202), 其中相关参量数据 (105) 的第一部分表示所述 M 个另外的音频信号 (202) 并且包括所述修正参数; 以及

- 将所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的至少第二部分解码成 N 个音频信号 (203), 其中所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的第二部分表示 N 个音频信号 (203)。

18. 一种发送编码多通道音频信号的方法, 该方法包括将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 及相关参量数据 (104、105), M 和 N 为整数, $N > M$, $M \geq 1$, 其中编码包括:

- 将 N 个音频信号 (101) 编码成 M 个音频信号 (102) 和第一相关参量数据 (104), 其中 M 个音频信号 (102) 和第一相关参量数据 (104) 表示所述 N 个音频信号 (101); 以及

- 从 M 个音频信号 (102) 产生表示该 M 个音频信号 (102) 的第二相关参量数据 (105), 所述第二相关参量数据包括使得能够从 K 个另外的音频信号 (103) 重建所述 M 个音频信号 (102) 的修正参数, 所述 K 个另外的音频信号 (103) 是不同于所述 M 个音频信号 (102) 的、所述 N 个音频信号 (101) 的备选向下混音, 其中相关参量数据 (104、105) 包括第一和第二

相关参量数据,

该方法还包括发送 K 个另外的音频信号 (103) 及相关参量数据 (104、105)。

19. 一种接收编码多通道音频信号的方法,该方法包括接收 K 个音频信号 (103) 及相关参量数据 (104、105),并将 K 个音频信号 (103) 和相关参量数据 (104、105) 解码成 N 个音频信号 (203), K 和 N 为整数, $N > K$, $K \geq 1$, 其中 K 个音频信号 (103) 和相关参量数据 (104、105) 表示 N 个音频信号 (203), 且其中解码包括:

- 从 K 个音频信号 (103) 和相关参量数据 (105) 的至少第一部分重建 M 个另外的音频信号 (202), 所述相关参量数据包括使得能够从 K 个音频信号 (103) 重建所述 M 个另外的音频信号 (202) 的修正参数, 所述 M 个另外的音频信号 (102) 是不同于所述 K 个音频通道 (103) 的、所述 N 个音频通道 (101) 的备选向下混音, M 为整数, $M \geq 1$, 其中相关参量数据 (105) 的第一部分表示所述 M 个另外的音频信号 (202); 以及

- 将所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的至少第二部分解码成 N 个音频信号 (203), 其中所述 M 个另外的音频信号 (202) 和相关参量数据 (104) 的第二部分表示 N 个音频信号 (203)。

20. 一种多通道音频播放器 (60), 包括根据权利要求 6 的多通道音频解码器 (20)。

21. 一种多通道音频记录器 (60), 包括根据权利要求 1 的多通道音频编码器 (10)。

多通道音频编码器、解码器以及相应方法

技术领域

[0001] 本发明涉及一种多通道音频编码器,用于将 N 个音频信号编码成 M 个音频信号及相关参量数据, M 和 N 为整数, $N > M$, $M \geq 1$ 。

[0002] 本发明还涉及一种多通道音频解码器、一种编码多通道音频信号的方法、一种解码多通道音频信号的方法、一种编码的多通道音频信号、一种在其上存储了该编码多通道音频信号的存储介质、一种用于发送和接收编码多通道音频信号的传输系统、一种用于发送编码多通道音频信号的发送器、一种用于接收编码多通道音频信号的接收器、一种发送和接收编码多通道音频信号的方法、一种发送编码多通道音频信号的方法、一种接收编码多通道音频信号的方法、一种多通道音频播放器、一种多通道音频记录器,以及一种用于执行任意前述方法的计算机程序产品。

背景技术

[0003] 一段时间以来,多通道音频信号再现引人关注。多通道音频信号是一种具有两个或两个以上音频通道的音频信号。众所周知的多通道音频信号示例是二通道立体声音频信号和具有两个前音频通道、两个后音频通道、一个中心音频信号以及一个附加的低频增强(LFE)通道的 5.1 通道音频信号。这种 5.1 通道音频信号用于 DVD(数字多功能盘)和 SACD(超级音频压缩盘)系统。由于多通道材料变得更为普及,多音频材料的高效编码变得愈发重要。

[0004] 5.1-2-5.1 多通道音频编码系统是公知的。在这个公知音频编码系统中,一个 5.1 输入音频信号被编码成并表示成两个向下混音(down-mix)通道以及相关参数。该向下混音信号也统称为空间向下混音。在该公知系统中,空间向下混音形成一个具有立体图象的立体声音频信号,从质量上看,它能够同来自 5.1 输入通道的固定 ITU 向下混音相比。只有立体声装置的用户可以听这种空间立体声向下混音,而具有 5.1 通道装置的听众可以听由该空间立体声向下混音及相关参数构成的 5.1 通道再现。5.1 通道装置从空间立体声向下混音(即,立体声音频信号)和相关参数解码/重建该 5.1 通道音频信号。

[0005] 然而,演播室工程师往往发现这种空间立体声向下混音很迟钝。这是他们进行艺术立体声向下混音的原因,这种艺术立体声向下混音不同于空间立体声向下混音。例如附加额外的混响和源,使立体图象变宽,等等。为了使用户能够享受这种艺术立体声向下混音,这种艺术立体声向下混音,而不是空间向下混音,可以通过传输介质传输或者存储在存储介质上。但是,这种方法严重影响 5.1 通道音频信号再现的质量。这种输入 5.1 通道音频信号被编码成一个空间立体声向下混音以及相应的参数。通过由艺术立体声向下混音代替空间立体声向下混音,在系统的解码端不再得到空间立体声向下混音,且 5.1 通道音频信号的高质量重建不再可能。

发明内容

[0006] 本发明的一个目的是提供一种开篇段落中阐述的多通道音频编码器,它减轻了上

述问题。在根据本发明的多通道音频编码器中实现了这一目的,其中多通道音频编码器包括:

[0007] - 第一单元,用于将 N 个音频信号编码成 M 个音频信号和第一相关参量数据,其中 M 个音频信号和第一相关参量数据表示 N 个音频信号;以及

[0008] - 第二单元,耦合到第一单元,该第二单元被安排成从这 M 个音频信号产生用以表示这 M 个音频信号的第二相关参量数据,所述第二相关参量数据包括使得能够从 K 个另外的音频信号重建所述 M 个音频信号的修正参数,所述 K 个另外的音频信号是不同于所述 M 个音频信号的、所述 N 个音频信号的备选向下混音,其中相关参量数据包括第一和第二相关参量数据。

[0009] 通过从空间向下混音(即,M 个音频信号)产生表示该空间向下混音的参数,解码器能够重建至少部分该空间向下混音,例如,通过合成一个同该空间向下混音相似的信号。这些参数(即第二相关参量数据)表示该空间向下混音,例如通过空间向下混音信号的一个或多个相关特性。其后,重建的空间向下混音可以同第一相关参量数据(即,常规多通道参数)一起使用,用以解码并重建多通道音频信号(即,N 个音频信号)。本发明基于这样一种认识,即同以备选向下混音为解码基础得到的多通道音频信号相比,这样的多通道音频信号具有更高的质量。此外,当编码器处没有备选向下混音时或者当备选向下混音失真时,解码器仍然能够使用这些参数来重建一个具有良好质量的多通道音频信号。

[0010] 第二单元被安排成产生第二相关参量数据,使得第二相关参量数据包括修正参数,该修正参数使得能够从 K 个另外的音频信号重建 M 个音频信号。这样,解码器能够执行更优的空间向下混音重建。这种重建可以在备选向下混音,即 K 个另外的音频信号(例如艺术向下混音)的基础上完成。解码器可以将所述修正参数施加到备选向下混音信号,以便该向下混音信号更加近似空间向下混音。

[0011] 在根据本发明的多通道音频编码器的一个实施例中,第二单元被安排成从 M 个音频信号和 K 个另外的音频信号产生第二相关参量数据,使得修正参数表示 M 个音频信号和 K 个另外的音频信号之间的差异。在这个实施例中,编码器可使用备选向下混音,且可完成修正参数的有效表示。通过比较空间向下混音和备选向下混音,第二单元能够产生修正参数,该修正参数表示空间向下混音同备选向下混音之间的差异。同前一实施例中的“绝对”修正参数相比,这种“相对”修正参数在编码多通道音频信号中要求更少的空间/位。备选向下混音优选为多通道音频编码器从外部源接收的艺术向下混音。或者,可以在多通道音频编码器内例如从 N 个输入音频信号产生备选向下混音。

[0012] 编码器可包括一个选择器,用于选择备选向下混音或空间向下混音以进行输出。随后,被选的向下混音将成为编码音频信号的一部分。例如,没有备选向下混音时,可以选择空间向下混音。

[0013] 在根据本发明的多通道音频编码器的一个实施例中,第二单元被安排成产生第二相关参量数据,使得修正参数包括 M 个音频信号的特性或 M 个音频信号的特性同 K 个另外的音频信号的特性之间的差异。发明人发现,修正参数优选地包括统计信号特性(之间的差异),例如一个(多个)向下混音信号的方差、协方差和相关性以及标准差。这些统计信号特性使空间向下混音能够得到良好的重建。

[0014] 在根据本发明的多通道音频编码器的一个实施例中,第二单元被安排成产生第二

相关参量数据,使得该特性包括:

[0015] - 至少部分音频信号的一个能量或功率值;或者

[0016] - 至少部分音频信号的一个相关性数值;或者

[0017] - 至少部分音频信号的各能量或功率值之间的一个比值。

[0018] 这些特性或者单独或任意可行组合都能使空间向下混音得到高效和/或高质量的重建。能量或功率值和相关性数值允许得到高质量重建。包括各能量或功率值之间比值的特性十分有效,因为它在编码多通道音频信号/比特流中只需要相对小的空间/少的比特位。

[0019] 修正参数通常被分析为时间和频率的函数(即,对于一组时间/频率片段)。它们可以包含在编码多通道音频信号所包括的参数比特流中。为了进一步提高空间向下混音的重建质量,可以进一步通过空间向下混音的(编码)低频内容扩充参数比特流。

[0020] 在解码器处,从编码多通道音频信号获得修正参数,并使用这些参数从该备选向下混音或从头开始重建空间向下混音。解码器转换该备选向下混音,使所获得的转换后的向下混音信号具有空间向下混音的性质。根据修正参数的表示方法,解码器可采用两种方式工作。如果这些参数表示从备选向下混音到空间向下混音(的所需要的特性)的(相对)转换,则从所发送的参数直接获得转换变量。另一方面,如果所发送的参数表示空间向下混音的(绝对)特性,则解码器首先计算备选向下混音的相应特性。使用这一信息(所发送的参数和所发送的向下混音的求得特性)便可确定转换变量,这些转换变量描述从所发送的向下混音(的特性)到空间向下混音(的特性)的转换。最后,空间参数(即,第一相关参量数据)被应用于重建的空间向下混音,用以解码多通道音频信号。

[0021] 相同的发明构思可用在具有以下发送器和接收器的传输系统中,所述发送器具有1个多通道音频编码器,所述接收器具有1个多通道音频解码器。这种传输系统例如可用于通过传输介质(例如无线电通道、同轴电缆或光纤)传输语音信号或音频信号。这种传输系统还可用于在记录介质(例如磁带、磁盘或光盘或者固态存储器)上记录编码的音频或语音信号。所述发明构思还可有利地用在音频播放器/记录器中,所述音频播放器/记录器是例如具有1个多通道音频解码器/编码器的光盘音频播放器/记录器或硬盘驱动音频播放器/记录器或固态存储音频播放器/记录器。

附图说明

[0022] 通过参考附图对优选实施例的下列描述,本发明的上述目的和特征将变得更加明显,附图中:

[0023] 图1示出根据本发明的多通道音频编码器10的一个实施例的框图;

[0024] 图2示出根据本发明的多通道音频解码器20的一个实施例的框图;

[0025] 图3示出根据本发明的传输系统70的一个实施例的框图;

[0026] 图4示出根据本发明的多通道音频播放器/记录器60的一个实施例的框图;

[0027] 图5示出根据本发明的多通道音频编码器10的另一个实施例的框图;

[0028] 图6示出根据本发明的多通道音频解码器20的另一个实施例的框图。

[0029] 图中,对相同的部分提供了相同的参考数字。

具体实施方式

[0030] 图1示出根据本发明的多通道音频编码器10的一个实施例的框图。这个多通道音频编码器10被安排成将N个音频信号101编码成M个音频信号102和相关参量数据104、105。其中,M和N都是整数,且 $N > M$, $M \geq 1$ 。多通道音频编码器10的一个示例是5.1-2编码器,其中N等于6,即5+1通道,M等于2。这种多通道音频编码器将5.1通道输入音频信号编码成2通道输出音频信号(例如立体声输出音频信号)及相关参数。多通道音频编码器10的其它示例是5.1-1、6.1-2、6.1-1、7.1-2和7.1-1编码器。只要N大于M且M大于或等于1,具有其它N、M值的编码器也是可以的。

[0031] 编码器10包括第一编码单元110和与其耦合的第二编码单元120。第一编码单元110接收N个输入音频信号101,并将这N个音频信号101编码成M个音频信号102和第一相关参量数据104。M个音频信号102和第一相关参量数据104表示N个音频信号101。第一单元110执行的N个音频信号101到M个音频信号102的编码还可以称为向下混音,且M个音频信号102还可以称为空间向下混音102。单元110可以是常规参数多通道音频编码器,它将多通道音频信号101编码成单声道或立体声向下混音音频信号102和相关参数104。相关参数104能使解码器从该单声道或立体声向下混音音频信号102重建多通道音频信号101。应该注意,向下混音102也可以具有两个以上通道。

[0032] 第一单元110将空间向下混音102提供给第二单元120。第二单元120从空间向下混音102生成第二相关参量数据105。第二相关参量数据105表示空间向下混音102,即这些参数105包括空间向下混音102中的某些特征和性质,这些特征和性质使解码器能够例如通过合成1个同空间向下混音102相近的信号来重建至少部分空间向下混音102。该相关参量数据包括第一和第二相关参量数据104和105。

[0033] 第二相关参量数据105可包括多个修正参数,它们使得能够从K个另外的音频信号103重建空间向下混音102。这样,解码器可以对空间向下混音102实施更好的重建。可以在备选向下混音103的基础上完成这种重建,所述备选向下混音103即K个另外的音频信号103,例如艺术向下混音。解码器可将修正参数应用于备选向下混音信号103,使得它与空间向下混音102更加类似。

[0034] 第二单元120可在其输入端接收备选向下混音103。可从编码器10的一个外部源接收该备选向下混音103(如图1所示),或者可在编码器10内,例如从N个音频信号101产生该备选向下混音103(未示出)。第二单元120可将空间向下混音102同该备选向下混音103相比较,并产生修正参数105,用以表示空间向下混音102同备选向下混音103之间的差异,例如空间向下混音102的特性同备选向下混音103的特性之间的差异。在编码的多通道音频信号中,这种表示差异的“相对”修正参数比仅表示空间向下混音102(的一个或多个特性)的“绝对”修正参数所要求的空间/比特位要少。修正参数105优选地包括向下混音信号(之间差异)的一个或多个统计信号特性,例如方差、协方差和相关性,或者这些特性的比(之间的差异)。应该注意,信号的方差等价于该信号的能量或功率。这些统计信号特性使得空间向下混音得以被良好重建。

[0035] 图2示出根据本发明的多通道音频解码器20的一个实施例的框图。解码器20被安排成将K个音频信号103及相关参量数据104、105解码成N个音频信号203。这里,K和N为整数,且 $N > K$, $K \geq 1$ 。该K个音频信号103(即,备选向下混音103)及相关参量数据

104、105 表示 N 个音频信号 203 (即,多通道音频信号 203)。多通道音频解码器 20 的一个示例是 2-5.1 解码器,其中 N 等于 6,即 5+1 通道,且 K 等于 2。这种多通道解码器将 2 通道输入音频信号 (例如立体声输入音频信号) 及相关参数解码成 5.1 通道输出音频信号。多通道音频解码器 20 的其它示例为 1-5.1、2-6.1、1-6.1、2-7.1 和 1-7.1 解码器。只要满足 N 大于 K,且 K 大于等于 1,具有其它 N、K 值的解码器也是可能的。

[0036] 多通道音频解码器 20 包括第一单元 210 和与其耦合的第二单元 220。第一单元 210 接收备选向下混音 103 和修正参数 105,并从该备选向下混音 103 和修正参数 105 重建 M 个另外的音频信号 202,即空间向下混音 202 或其近似信号。这里,M 为整数,且 $M \geq 1$ 。修正参数 105 表示空间向下混音 202。第二单元 220 从第一单元 210 接收空间向下混音 202 并接收修正参数 104。第二单元 220 将空间向下混音 202 和修正参数 104 解码成多通道音频信号 203。第二单元 220 可以是常规参数多通道音频解码器,它将单声道或立体声向下混音音频信号 202 及相关参数 104 解码成多通道音频信号 203。

[0037] 第一单元 210 可被安排成确定是否有必要或希望从输入信号 103 重建信号 202。当空间向下混音信号 202 取代备选向下混音 103 而被提供到第一单元 210 时,无法应用这种重建。第一单元 210 通过从输入信号 103 产生与修正参数 105 中所包括的相近或相同的信号特性,并通过将这些产生的信号特性同修正参数 105 相比较来进行上述确定。如果比较显示,产生的信号特性等价于或基本上等价于修正参数 105,则输入信号 103 与空间向下混音信号 202 充分相似,且第一单元 210 可将输入信号 103 转发至第二单元 220。如果比较显示,所产生的信号特性不等价于或不充分等价于修正参数 105,则输入信号 103 与空间向下混音信号 202 不充分相似,且第一单元 210 可从输入信号 103 和修正参数 105 重建 / 近似空间向下混音信号 202。

[0038] 修正参数 105 可表示备选向下混音 103 同空间向下混音 202 之间的差异,例如统计信号特性中的差异,使得第一单元 210 能够从备选向下混音 103 重建空间向下混音 202。

[0039] 第一单元 210 可从备选向下混音产生表示该备选向下混音 103 的另外的修正参数 / 特性。这种情况下,第一单元 210 可从备选向下混音 103、修正参数 105 和该另外的修正参数 (之间的差异) 重建空间向下混音 202。

[0040] 修正参数 105 和该另外的修正参数可分别包括空间向下混音 202 和备选向下混音 103 的统计特性。从一些信号导出的诸如方差、相关性和协方差等的统计特性为这些信号提供良好表示。它们在重建空间向下混音 202 中是有用的,该重建例如通过转换备选向下混音来完成,这种转换使得其相关特性同修正参数 105 中包含的特性相匹配。

[0041] 图 3 示出根据本发明的传输系统 70 的一个实施例的框图。传输系统 70 包括 1 个发送器 40,用于通过传输信道 30 (例如,有线或无线通信链路) 将编码多通道音频信号发送到接收器 50。发送器 40 包括 1 个如上所述的多通道音频编码器 10,用于将多通道音频信号 101 编码成空间向下混音 102 和相关参数 104、105。发送器 40 还包括装置 41,用于通过传输信道 30 将包括参数 104、105 和空间向下混音 102 或备选向下混音 103 的编码多通道音频信号发送到接收器 50。接收器 50 包括装置 51 和如上所述的多通道音频解码器 20,装置 51 用于接收编码多通道音频信号,多通道音频解码器 20 用于将备选向下混音 103 或空间向下混音 102 及相关参数 104、105 解码成多通道音频信号 203。

[0042] 图 4 示出根据本发明的多通道音频播放器 / 记录器 60 的一个实施例的框图。音

频播放器 / 记录器 60 包括根据本发明的 1 个多通道音频解码器 20 和 / 或 1 个多通道音频编码器 10。音频播放器 / 记录器 60 可具备自身存储器,例如固态存储器或硬盘。音频播放器 / 记录器 60 还可方便可分离式存储装置,例如 (可记录)DVD 盘或 (可记录)CD 盘。可以通过解码器解码所存储的包括备选向下混音 103 和参数 104、105 的编码多通道音频信号,并通过音频播放器 / 记录器 60 播放或再现。编码器 10 可编码多通道音频信号以便在存储装置上存储。

[0043] 图 5 示出根据本发明的多通道音频编码器 10 的另一个实施例的框图。编码器 10 包括第一单元 110 和与其耦合的第二单元 120。第一单元 110 接收 1 个 5.1 多通道音频信号 101,它分别包括左前、左后、右前、右后、中心以及低频增强音频信号 lf、lr、rf、rr、co 和 lfe。第二单元 120 接收分别包括左艺术音频信号 la 和右艺术音频信号 ra 的艺术立体声向下混音 103。多通道音频信号 101 和艺术向下混音 103 为时域音频信号。在第一和第二单元 110 和 120 中,这些信号 101 和 103 被分割并被转换到频 - 时域。

[0044] 第一单元 110 中,在三个阶段导出参量数据 104。在第一阶段,在分割和转换单元 112、113 和 114 中,三对音频信号 lf 和 rf、rf 和 rr 以及 co 和 lfe 分别被分割,而且分割信号分别被转换到频域。得到的分割信号的频域表示分别被示为频域信号 Lf、Lr、Rf、Rr、Co 和 LFE。在第二阶段,分别在向下混音器 115、116、117 中将这三对频域信号 Lf 和 Lr、Rf 和 Rr、Co 和 LFE 向下混音,用以分别产生单声道音频信号 L、R、C 及相关参数 141、142、143。向下混音器 115、116、117 可为常规 MPEG4 参数立体声编码器。最后在第三阶段,在向下混音器 118 中将这三个单声道音频信号 L、R、C 向下混音,以便获得空间立体声向下混音 102 及相关参数 144。空间向下混音 102 包括信号 Lo 和 Ro。

[0045] 参量数据 141、142、143 和 144 被包含在第一相关参量数据 104 中。参量数据 104 和空间向下混音 102 表示 5.1 输入信号 101。

[0046] 在第二单元中,首先在分割单元 121 中将在时域中分别由音频信号 la 和 ra 表示的艺术向下混音信号 103 分割。得到的分割音频信号 127 分别包括信号 las 和 ras。接下来,分割音频信号 127 由转换器 122 转换到频域。得到的频域信号 126 包括信号 La 和 Ra。最后,作为分割艺术向下混音 103 的频域表示的频域信号 126 和分割空间向下混音 102 的频域表示被提供给发生器 123,该发生器 123 产生修正参数 105,修正参数 105 使解码器能够修正 / 转换艺术向下混音 103,从而使其更接近空间向下混音 102。分割时域信号 127 也被提供给选择器 124。这一选择器 124 的另外两个输入是空间立体声向下混音 102 的频域表示和控制信号 128。控制信号 128 确定选择器 124 将艺术向下混音 103 还是空间向下混音 102 作为编码多通道音频信号的一部分输出。无法得到艺术向下混音时可以选择空间向下混音 102。可人工设置或通过检测艺术向下混音 103 的存在自动生成控制信号 128。控制信号 128 可包含在参数比特流中,从而相应的解码器 20 可以按下述方式对其进行利用。

[0047] 选择器 124 的输出信号 102、103 被示为信号 lo 和 ro。如果选择器 124 输出艺术立体声向下混音 127,则在选择器 124 中通过重叠 - 附加将分割时域信号 las 和 ras 组合成信号 lo 和 ro。如果按控制信号 128 指示的那样输出空间立体声向下混音 102,则选择器 124 将信号 Lo 和 Ro 转换回时域,并通过重叠 - 附加将它们组合成信号 lo 和 ro。时域信号 lo 和 ro 形成 5.1-2 编码器 10 的立体声向下混音。

[0048] 下面是对发生器 123 更为详细的阐述。发生器 123 的功能是确定对艺术向下混音

103 的转换进行描述的修正参数,使得从某种意义上说,它类似于初始空间向下混音 102。通常,这一转换可描述为

$$[0049] \quad [\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a \ \underline{A}_1 \ \cdots \ \underline{A}_N]T \quad (1)$$

[0050] 其中 $\underline{L}_a$ 和 $\underline{R}_a$ 为向量,包括艺术向下混音 103 左右通道的时 / 频片段样本,且其中 $\underline{L}_d$ 和 $\underline{R}_d$ 为向量,包括修正艺术向下混音左右通道的时 / 频片段采样,其中 $\underline{A}_1 \ \cdots \ \underline{A}_N$ 包括可选辅助通道时 / 频片段的样本,且其中 T 为转换矩阵。注意任何向量 $\underline{v}$ 都定义为列向量。修正的艺术向下混音是经过转换的艺术向下混音 103,使得它同初始空间向下混音 102 相似。辅助通道 $\underline{A}_1 \ \cdots \ \underline{A}_N$ 例如可以是艺术向下混音信号的去相关版本,或者包含空间向下混音信号的低频成分。后一种情况下,这一低频成分可以包含在参数 105 中。 $(N+2) \times 2$ 转换矩阵 T 描述自艺术向下混音 103 和辅助通道到修正艺术向下混音的转换。转换矩阵 T 或其元素优选地包含在修正参数 105 中,使得解码器 20 能够重建至少部分转换矩阵 T。此后,解码器 20 可(如下所述)将转换矩阵 T 应用于艺术向下混音 103,用以重建空间向下混音 102。

[0051] 或者,修正参数 105 包括空间向下混音 102 的信号特性,例如能量或功率值和 / 或相关性数值。随后,解码器 20 可从艺术向下混音 103 产生这些信号特性。空间向下混音 102 和艺术向下混音 103 的信号特性使解码器 20 能够建立转换矩阵 T(下面阐述),并能够将该矩阵应用到艺术向下混音 103,用以重建空间向下混音 102(同样在下面阐述)。

[0052] 还有几种使艺术立体声向下混音 103 同初始立体声向下混音 102 相似的可能:

[0053] I. 波形的匹配。

[0054] II. 统计特性的匹配:

[0055] 左右通道能量或功率的匹配。

[0056] 左右通道协方差矩阵的匹配。

[0057] III. 在左右通道能量或功率匹配的约束下获得波形的最佳可能匹配。

[0058] IV. 混合上述方法 I - III。

[0059] 下面,不考虑 (1) 的辅助通道 $\underline{A}_1 \ \cdots \ \underline{A}_N$,由此可将转换矩阵 T 写成

$$[0060] \quad [\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a]T \quad (2)$$

[0061] I. 波形匹配(方法 I)

[0062] 通过将修正艺术向下混音的左右信号表示成艺术立体声向下混音 103 的左右信号的线性组合,可以获得艺术向下混音 103 同空间向下混音 102 的波形匹配:

$$[0063] \quad \underline{L}_d = \alpha_1 \underline{L}_a + \beta_1 \underline{R}_a, \underline{R}_d = \alpha_2 \underline{L}_a + \beta_2 \underline{R}_a \quad (3)$$

[0064] 则, (2) 的矩阵 T 可写成:

$$[0065] \quad T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}$$

[0066] 一种选择参数 α_1 、 α_2 、 β_1 、 β_2 的方式是最小化空间向下混音信号 L_0 、 R_0 同其估计(即,修正艺术向下混音信号 L_d 、 R_d) 之间的欧几里德距离的平方,因此

$$[0067] \quad \min_{\alpha_1, \beta_1} \sum_k \|L_0[k] - L_d[k]\|^2 = \min_{\alpha_1, \beta_1} \sum_k \|L_0[k] - \alpha_1 L_a[k] - \beta_1 R_a[k]\|^2 \quad (4)$$

[0068] 且

$$[0069] \quad \min_{\alpha_2, \beta_2} \sum_k \|R_0[k] - R_d[k]\|^2 = \min_{\alpha_2, \beta_2} \sum_k \|R_0[k] - \alpha_2 L_a[k] - \beta_2 R_a[k]\|^2 \quad (5)$$

[0070] II. 统计特性的匹配 (方法 II)

[0071] 方法 II . a : 现在讨论左右通道能量的匹配。现将分别由 L_d 和 R_d 表示的修正左右艺术向下混音信号计算为 :

$$[0072] \quad L_d = \alpha L_a, R_d = \beta R_a \quad (6)$$

[0073] 其中, 实参数的情况下, α 和 β 由下式给出

$$[0074] \quad \alpha = \sqrt{\frac{\sum_k \|L_0[k]\|^2}{\sum_k \|L_a[k]\|^2}}, \beta = \sqrt{\frac{\sum_k \|R_0[k]\|^2}{\sum_k \|R_a[k]\|^2}} \quad (7)$$

[0075] 从而, 可将转换矩阵 T 写成 :

$$[0076] \quad T = \begin{bmatrix} \sqrt{\frac{\sum_k \|L_0[k]\|^2}{\sum_k \|L_a[k]\|^2}} & 0 \\ 0 & \sqrt{\frac{\sum_k \|R_0[k]\|^2}{\sum_k \|R_a[k]\|^2}} \end{bmatrix} \quad (8)$$

[0077] 这些选择能够保证信号 L_d 和 R_d 分别同信号 L_0 和 R_0 具有相同的能量。

[0078] 方法 II . b : 为了匹配艺术立体声向下混音 103 和空间立体声向下混音 102 的协方差矩阵, 可按式通过特征值分解对这些矩阵进行分解 :

$$[0079] \quad C_a = U_a S_a U_a^H, C_0 = U_0 S_0 U_0^H \quad (9)$$

[0080] 其中, 艺术立体声向下混音 103 的协方差矩阵 C_a 由下式给出

$$[0081] \quad C_a = [L_a \ R_a]^H [L_a \ R_a] \quad (10)$$

[0082] U_a 为酉矩阵, S_a 为对角阵。 C_0 为空间立体声向下混音 102 的协方差矩阵, U_0 为酉矩阵, S_0 为对角阵。计算

$$[0083] \quad X_{aw} = [L_{aw} \ R_{aw}] = [L_a \ R_a] U_a S_a^{-1/2} \quad (11)$$

[0084] 时, 获得了两个互不相关信号 L_{aw} 和 R_{aw} (由于乘以矩阵 U_a), 它们具有单位能量 (由于乘以矩阵 $S_a^{-1/2}$)。通过计算

$$[0085] \quad X_d = [L_d \ R_d] = [L_a \ R_a] U_a S_a^{-1/2} U_r S_0^{1/2} U_0^H \quad (12)$$

[0086] 首先将 $[L_a \ R_a]$ 的协方差矩阵转换成等于单位阵的协方差矩阵, 即 $[L_a \ R_a] U_a S_a^{-1/2}$ 的协方差矩阵。应用任何酉矩阵 U_r 都不会改变协方差结构, 应用 $S_0^{-1/2} U_0^H$ 产生一个协方差结构, 该结构同空间立体声向下混音 102 的协方差结构等价。

[0087] 将矩阵 S_{0w} 和信号 L_{0w} 和 R_{0w} 定义为

$$[0088] \quad S_{0w} = [L_{0w} \ R_{0w}] = [L_0 \ R_0] U_0 S_0^{-1/2} \quad (13)$$

[0089] 可以这样选择矩阵 U_r , 使得在最小欧几里德距离平方意义上, 在信号 L_{0w} 同 L_{aw} 、信号 R_{0w} 同 R_{aw} 间获得最佳可能波形匹配, 其中 L_{aw} 同 R_{aw} 由 (11) 给出。通过 U_r 的这一选择, 可在统计方法范围内使用波形匹配。

[0090] 从 (12) 可以看到, 转换矩阵 T 可由下式给出

$$[0091] \quad T = U_a S_a^{-1/2} U_r S_0^{1/2} U_0^H \quad (14)$$

[0092] III. 在能量约束下获得最佳波形匹配 (方法III)

[0093] 假设 (3) 可通过在如下能量约束下最小化 (4)、(5) 来获得参数 α_1 、 α_2 、 β_1 、 β_2

$$[0094] \quad \sum_k \|L_0[k]\|^2 = \sum_k \|L_d[k]\|^2, \sum_k \|R_0[k]\|^2 = \sum_k \|R_d[k]\|^2 \quad (15)$$

[0095] IV. 混合方法 (方法IV)

[0096] 就混合不同的方法而言,可能的组合是混合方法 II . a 和 II . b, 或者混合方法 II . a 和 III。可按下述进行:

[0097] 如果通过方法 II . b/ III 获得的 L_0 同 L_d 之间、 R_0 同 R_d 之间的波形匹配都良好, 使用方法 II . b/ III。

[0098] 如果这一波形匹配较差, 使用方法 II . a。

[0099] 通过混合它们的变换矩阵, 保证两种方法之间的一个逐步转换, 作为这一波形匹配质量的函数。

[0100] 这在数学上可表示如下:

[0101] 通过 (3) 和 (2), 可将转换矩阵 T 写成下面的通用形式

$$[0102] \quad T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} \quad (16)$$

[0103] 通过两个向量 T_L 和 T_R , 可将这一矩阵重写成下式形式

$$[0104] \quad T = [T_L \ T_R], T_L = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}, T_R = \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} \quad (17)$$

[0105] 采用方法 II . b 或方法 III 得到的 L_0 同 L_d 之间的波形匹配质量由 γ_L 表示。它定义为

$$[0106] \quad \gamma_L = \max \left(0, \frac{\sum_k L_0[k] L_d^*[k]}{\sum_k \|L_0[k]\| \|L_d[k]\|} \right) \quad (18)$$

[0107] 采用方法 II . b 或方法 III 得到的 R_0 同 R_d 之间的波形匹配质量由 γ_R 表示。它定义为

$$[0108] \quad \gamma_R = \max \left(0, \frac{\sum_k R_0[k] R_d^*[k]}{\sum_k \|R_0[k]\| \|R_d[k]\|} \right) \quad (19)$$

[0109] γ_L 和 γ_R 都在 0 和 1 之间。左通道的混合系数 δ_L 和右通道的混合系数 δ_R 可按下式定义

[0110]

$$\delta_L = \begin{cases} 1 & \gamma_L > \mu_{L,\max} \\ 0 & \gamma_L < \mu_{L,\min} \\ \frac{1}{2} - \frac{1}{2} \cos\left(\pi \frac{(\gamma_L - \mu_{L,\min})}{(\mu_{L,\max} - \mu_{L,\min})}\right) & \text{其它} \end{cases}$$

[0111]

$$\delta_R = \begin{cases} 1 & \gamma_R > \mu_{R,\max} \\ 0 & \gamma_R < \mu_{R,\min} \\ \frac{1}{2} - \frac{1}{2} \cos\left(\pi \frac{(\gamma_R - \mu_{R,\min})}{(\mu_{R,\max} - \mu_{R,\min})}\right) & \text{其它} \end{cases} \quad (20)$$

[0112] 其中 $\mu_{L,\min}$ 、 $\mu_{L,\max}$ 、 $\mu_{R,\min}$ 、 $\mu_{R,\max}$ 是 0 到 1 之间的值， $\mu_{L,\min} < \mu_{L,\max}$ ，且 $\mu_{R,\min} < \mu_{R,\max}$ 。式 (20) 确保混合系数 δ_L 和 δ_R 在 0 到 1 之间。

[0113] 将方法 II . a、II . b 和 III 的转换矩阵 T 分别定义为 (8) 给出的 T_e 、(14) 给出的 T_a 和 T_{ce} 。类似 (17) 中 T 的分解，每个转换矩阵都可按下式分为两个向量：

$$[0114] \quad T_a = [\underline{T}_{a,L} \quad \underline{T}_{a,R}], T_e = [\underline{T}_{e,L} \quad \underline{T}_{e,R}], T_{ce} = [\underline{T}_{ce,L} \quad \underline{T}_{ce,R}] \quad (21)$$

[0115] 用于混合方法 II . a 和方法 II . b 的转换矩阵 T 可获得为

$$[0116] \quad T = [\underline{T}_L \quad \underline{T}_R] = [\delta_L \underline{T}_{a,L} + (1 - \delta_L) \underline{T}_{e,L} \quad \delta_R \underline{T}_{a,R} + (1 - \delta_R) \underline{T}_{e,R}] \quad (22)$$

[0117] 用于混合混合方法 II . a 和方法 III 的转换矩阵 T 可获得为

[0118]

$$T = [\underline{T}_L \quad \underline{T}_R] = [\delta_L \underline{T}_{ce,L} + (1 - \delta_L) \underline{T}_{e,L} \quad \delta_R \underline{T}_{ce,R} + (1 - \delta_R) \underline{T}_{e,R}] \quad (23)$$

[0119] 转换矩阵 T 的元素可以是实值的或复值的。可如下将这些元素编码到修正参数中：可将转换矩阵 T 中实的、正的元素像 MPEG4 参数立体声中使用的 IID 参数那样按对数量化。可以对参数的数值设定一个上限，以避免小信号的过分放大。这一上限可以是固定的，也可以是自动生成的左通道同艺术左通道之间相关性以及自动生成的右通道同艺术右通道之间相关性的函数。对于 T 的复元素，可通过 IID 参数对其幅值进行量化，并可线性量化其相位。可采用取元素绝对值的对数来对 T 中实的、可能为负的元素进行编码，同时保证负、正数值之间的区别。

[0120] 图 6 示出根据本发明的多通道音频解码器 20 的另一个实施例的框图。解码器 20 包括第 1 个一单元 210 和 1 个与其耦合的第二单元 220。第一单元 210 接收向下混音信号 l_o 、 r_o 和修正参数 105 作为输入。向下混音信号 l_o 和 r_o 可以是空间向下混音 102 或艺术向下混音 103 的一部分。第一单元 210 包括分割和变换单元 211 以及向下混音修正单元 212。在分割和变换单元 211 中，分别对向下混音信号 l_o 和 r_o 进行分割，并将分割信号变换到频域。分别将得到的分割向下混音信号的频域表示用频域信号 L_o 和 R_o 表示。接下来，频域信号 L_o 和 R_o 在向下混音修正单元 212 中进行处理。这一向下混音修正单元 212 的功能是修正输入向下混音，使其同空间向下混音 202 相似，即从艺术向下混音 103 和修正参数 105 重建空间向下混音 202。如果空间向下混音 102 被解码器 20 接收，则向下混音修正单

元 212 不必修正向下混音信号 L_o 和 R_o , 且这些向下混音信号 L_o 和 R_o 可作为空间向下混音 202 的向下混音信号 L_d 和 R_d 简单传送到第二单元 220。控制信号 217 可指示是否需要修正输入向下混音, 即输入向下混音是空间向下混音还是备选向下混音。可例如通过分析输入向下混音及相关参数 105 在解码器 20 内产生控制信号 217, 其中输入向下混音及相关参数 105 可以描述期望空间向下混音的信号特性。如果输入向下混音匹配期望的信号特性, 控制信号 217 可以被设置来指示不需要修正的情况。或者, 可人工设置控制信号 217, 或将其设置作为编码多通道音频信号的一部分例如在参数组 105 中来接收。

[0121] 如果编码器 20 接收艺术向下混音 103, 且控制信号 217 指示接收的向下混音信号 L_o 和 R_o 将被向下混音修正单元 212 修正, 则解码器可以有两种运行方式, 这依赖于所发送的参数的表示。如果参数表示从发送的向下混音到空间向下混音 (的所需特性) 的 (相对) 转换, 则直接从所发送的参数获得转换变量。通过这些转换变量就直接构成了变换矩阵 T 。

[0122] 另一方面, 如果所发送的参数表示空间向下混音的 (绝对) 特性, 则解码器首先计算实际发送的向下混音的相应特性。使用这一信息 (发送的参数和所发送的向下混音的求得特性), 转换变量随后被确定, 它描述从所发送的向下混音 (的特性) 向空间向下混音 (的特性) 的转换。更具体地, 可通过前述方法 II . a 或 (轻微修正的) II . b 确定转换矩阵 T 。

[0123] 如果在参量数据中仅发送 (绝对) 能量, 则使用方法 II . a。所发送的 (绝对) 参数 E_{L_o} 和 E_{R_o} 分别表示空间向下混音左右信号的能量,

[0124] 并由下式给出

$$[0125] \quad E_{L_o} = \sum_k \|L_o[k]\|^2, E_{R_o} = \sum_k \|R_o[k]\|^2 \quad (24)$$

[0126] 在解码器处计算所发送的向下混音的能量 E_{DL_o} 和 E_{DR_o} 。使用这些变量, 我们可按下列式计算 (7) 中的参数 α 和 β

$$[0127] \quad \alpha = \sqrt{\frac{E_{L_o}}{E_{DL_o}}}, \beta = \sqrt{\frac{E_{R_o}}{E_{DR_o}}} \quad (25)$$

[0128] 转换矩阵 T 由下式给出

$$[0129] \quad T = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \quad (26)$$

[0130] 如果 (绝对) 能量和 (绝对) 相关性都被发送, 则使用方法 II . b。所发送的 (绝对) 能量参数 E_{L_o} 和 E_{R_o} 分别表示空间向下混音的左右信号的能量, 并由 (24) 给出。这些能量以及所发送的空间向下混音的左右信号之间的相关性 $\rho_{L_o R_o}$, 可被用来按下式确定空间向下混音的协方差矩阵 C_o 。

$$[0131] \quad C_o = \begin{bmatrix} E_{L_o} & \rho_{L_o R_o}^* \sqrt{E_{L_o} E_{R_o}} \\ \rho_{L_o R_o} \sqrt{E_{L_o} E_{R_o}} & E_{R_o} \end{bmatrix} \quad (27)$$

[0132] 在解码器处计算所发送的向下混音的协方差矩阵 C_a 。对两个协方差矩阵进行 (9) 给出的特征值分析, 除了任意酉矩阵 U_r 外, 我们可以通过 (14) 计算转换矩阵 T 。由于空间

向下混音的波形无法得到,所以不能按前述方式选择这个矩阵。例如,现在可以将转换矩阵 T 选择为尽可能接近对角结构的形式。

[0133] 使用辅助信号时,同样构成了它们。如果不修正所接收的向下混音,转换矩阵 T 等于单位阵,且不使用辅助通道。使用式 (1) 计算输出信号 L_d 和 R_d 。应该注意,图 5 和图 6 将类似 L_d 和 R_d 的向量分别示为 L_d 和 R_d 。

[0134] 第二单元 220 是常规 2-5.1 多通道解码器,它将重建的空间向下混音 202 和相关参量数据 104 解码成一个 5.1 通道输出信号 203。正如前面阐述的那样,参量数据 104 包括参量数据 141、142、143 和 144。第二单元 220 执行编码器 10 中第一单元 110 的相反处理。第二单元 220 包括 1 个向上混音器 221,它将立体声向下混音 202 及相关参数 144 转换成 3 个单声道音频信号 L 、 R 和 C 。接下来,分别在去相关器 222、225、228 中去相关各单声道音频信号 L 、 R 和 C 。然后,混音矩阵 223 将单声道音频信号 L 、其去相关对应部分及相关参数 141 转换成信号 L_f 和 L_r 。类似地,混音矩阵 226 将单声道音频信号 R 、其去相关对应部分及相关参数 142 转换成信号 R_f 和 R_r ,且混音矩阵 229 将单声道音频信号 C 、其去相关对应部分及相关参数 143 转换成信号 C_o 和 LFE 。最后,3 对分割的频域信号 L_f 和 L_r 、 R_f 和 R_r 、 C_o 和 LFE 在逆转换器 224、227、230 中分别被转换到时域,并通过重叠-附加进行组合,以分别获得三对输出信号 l_f 和 l_r 、 r_f 和 r_r ,以及 c_o 和 lfe 。输出信号 l_f 、 l_r 、 r_f 、 r_r 、 c_o 和 lfe 形成解码的多通道音频信号 203。

[0135] 可通过数字硬件,或通过由数字信号处理器或通用微处理器执行的软件来实现多通道音频编码器 10 和多通道音频解码器 20。

[0136] 本发明的范畴不限于明确公开的实施例所给出的范围。本发明体现在每个新特征和每个新特征组合中。任何参考符号不限制权利要求的范围。词“包括”不排除出现权利要求列出以外的其它元素或步骤。在元素之前使用词“一个”不排除出现多个这种元素。

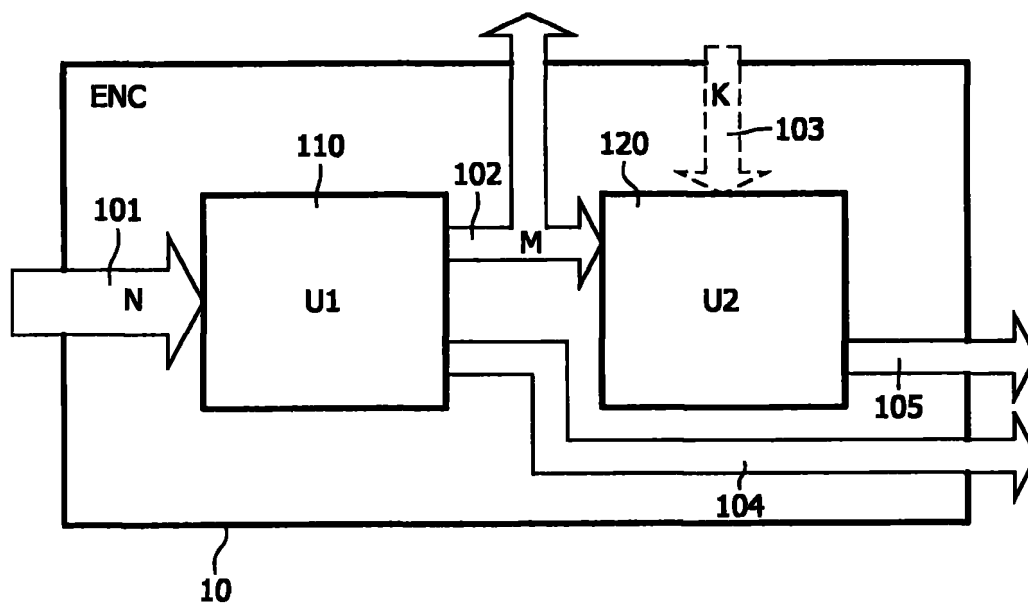


图 1

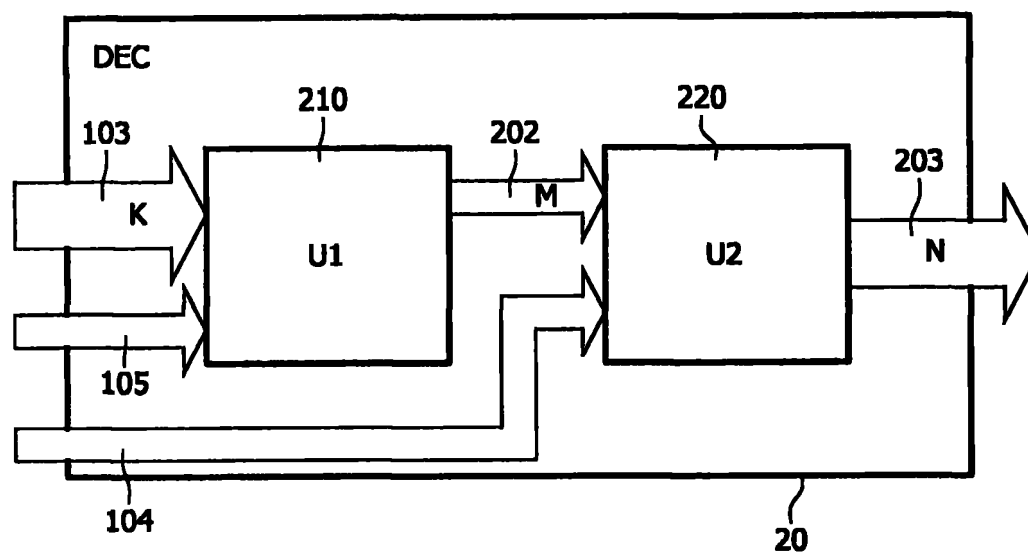


图 2

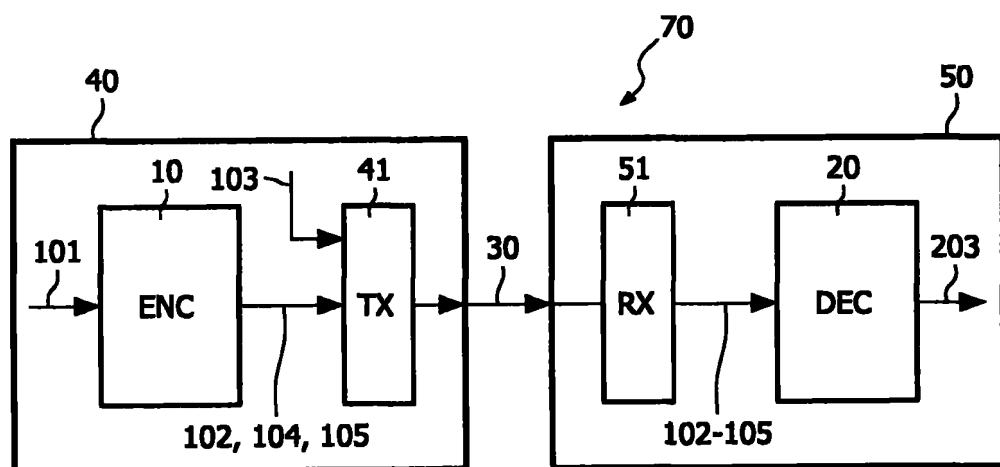


图 3

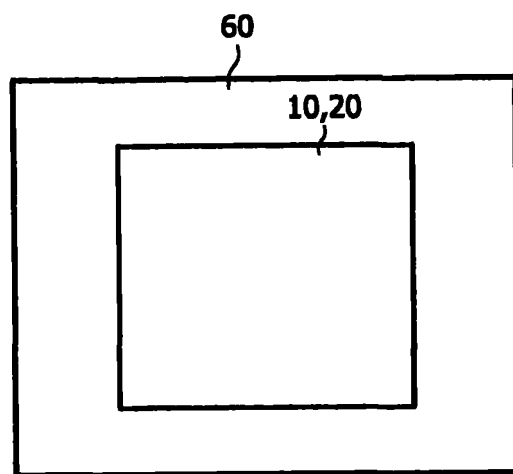
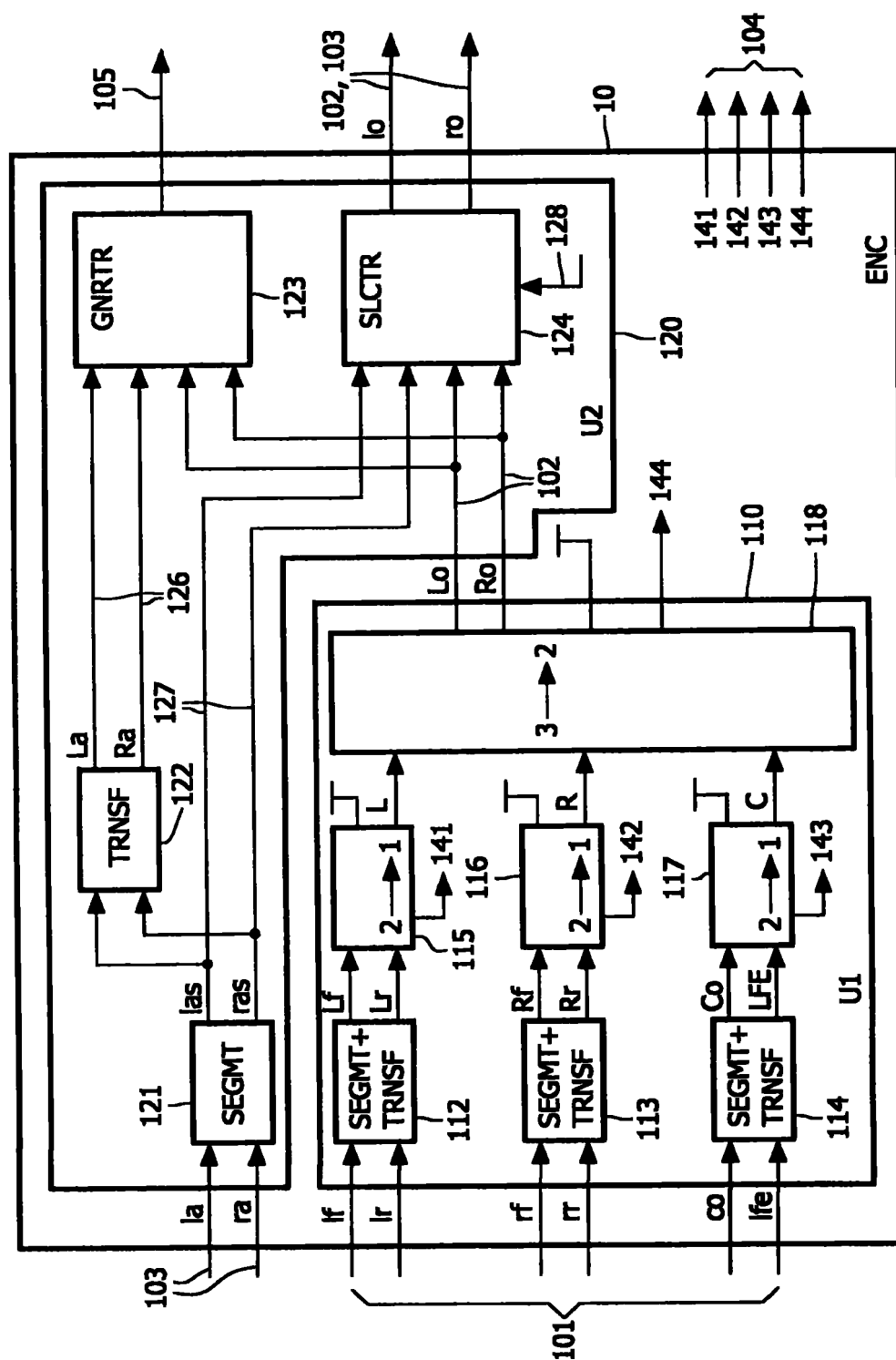


图 4



5

14

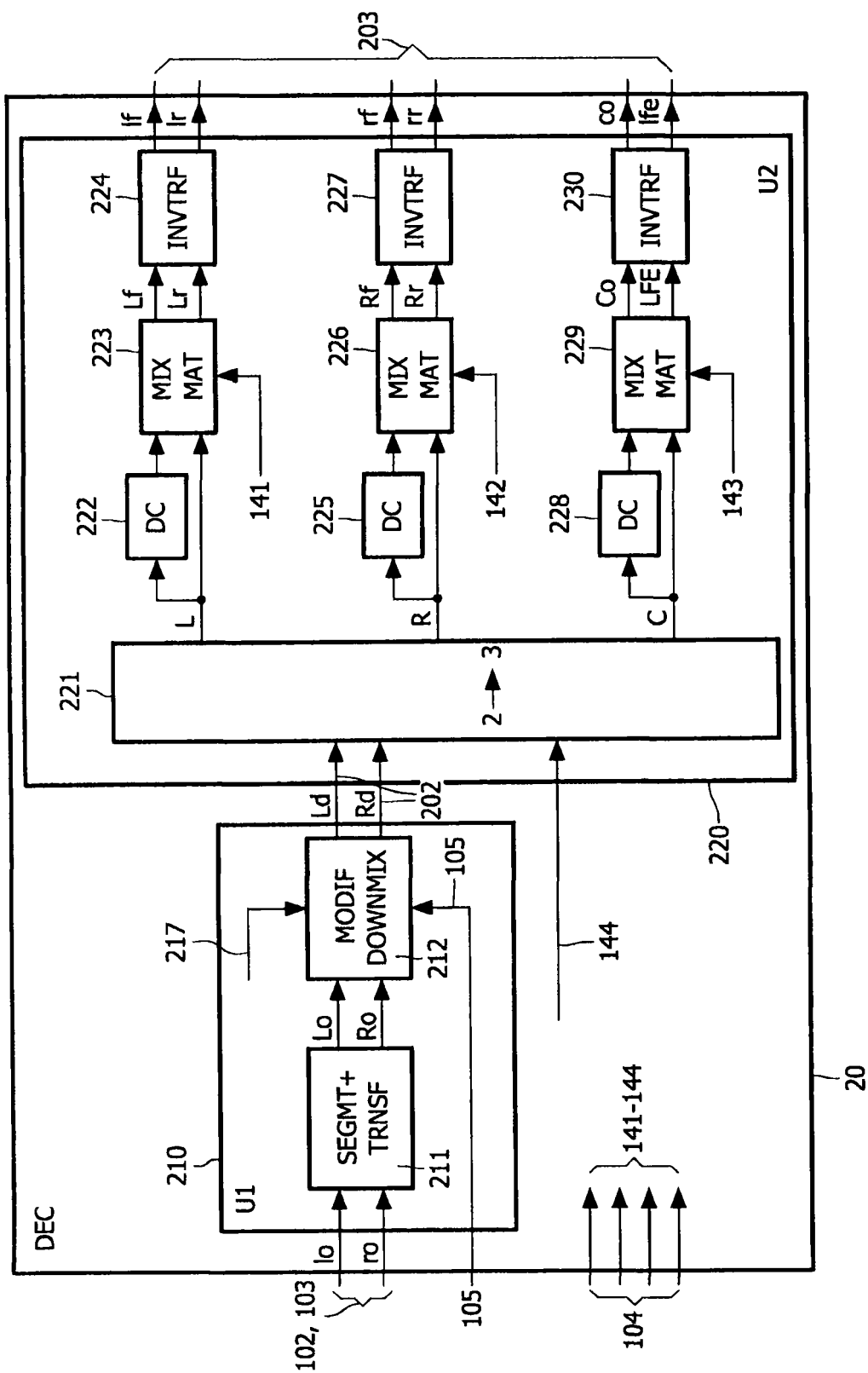


图 6