



(12) 发明专利

(10) 授权公告号 CN 101151658 B

(45) 授权公告日 2011.07.06

(21) 申请号 200680010200.9

(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

06100245.7 2006.01.11 EP

(85) PCT申请进入国家阶段日

2007.09.28

(86) PCT申请的申请数据

PCT/IB2006/050826 2006.03.16

(87) PCT申请的公布数据

W02006/103586 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

(72) 发明人 G·H·霍索 F·P·迈伯格

A·W·J·乌门

(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 刘红

(51) Int. Cl.

G10L 19/00(2006.01)

(56) 对比文件

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

FALLER C ET AL. Binaural Cue Coding -Part II: Schemes and Applications. 《IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING》. 2003, 第 11 卷 (第 6 期), 520-531.

审查员 张鑫

权利要求书 4 页 说明书 21 页 附图 7 页

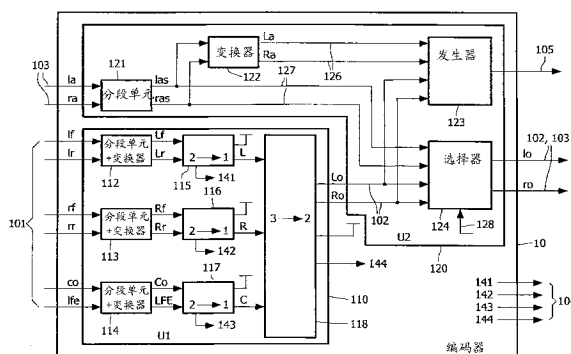
(54) 发明名称

多声道音频编码和解码方法、编码器和解码器

(57) 摘要

多声道音频编码器 (10) 对 N- 声道的音频信号进行编码。第一单元 (110) 产生用于 N- 声道信号的第一编码 M- 声道信号 (N > M), 例如空间立体声的下混合。下混合器 (115、116、117) 产生用于相对于 N- 声道音频信号的信号的第一增强数据。为 N- 声道信号产生第二 M- 声道信号, 例如艺术立体声混合。然后, 处理器 (123) 为相对于第一 M- 声道信号的第二 M- 声道信号产生第二增强数据。第二单元 (120) 产生包含第二 M- 声道信号、第一增强数据和第二增强数据的输出信号。发生器 (123) 能在产生第二增强数据作为绝对增强数据或者作为相对于第二编码 M- 声道信号的相对增强数据之间动态选择。解码器 (20) 能执行逆向操作, 并且能根据接收到的比特流中的指示应用第二增强数据作为绝对的或者相对的增强。

CN 101151658 B



1. 一种多声道音频编码器 (10), 用于编码 N- 声道音频信号, 该多声道音频编码器 (10) 包括:

用于为 N- 声道音频信号产生第一 M- 声道信号的装置 (110), M 小于 N;

用于产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据的装置 (115、116、117、118);

用于为 N- 声道音频信号产生第二 M- 声道信号的装置 (121);

用于产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据的增强装置 (123);

用于产生包含第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号的装置 (120); 并且

其中增强装置 (123) 设置成在产生第二增强数据作为绝对增强数据或者作为相对于第二 M- 声道信号的相对增强数据之间动态地选择。

2. 如权利要求 1 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为响应于 N- 声道信号的特征, 在绝对增强数据和相对增强数据之间进行选择。

3. 如权利要求 1 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为响应于绝对增强数据和相对增强数据的相对特征, 在绝对增强数据和相对增强数据之间进行选择。

4. 如权利要求 3 所述的多声道音频编码器 (10), 其中相对特征是相对于相对增强数据的信号能量的绝对增强数据的信号能量。

5. 如权利要求 1 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为将第二 M- 声道信号划分成信号块, 并个别地为每一个信号块在绝对增强数据和相对增强数据之间进行选择。

6. 如权利要求 5 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为仅基于与一个信号块相关的特征, 为该信号块在绝对增强数据和相对增强数据之间进行选择。

7. 如权利要求 1 所述的多声道音频编码器 (10), 其中在产生增强数据作为绝对增强数据和作为相对增强数据之间切换的切换时间间隔期间, 增强装置 (123) 设置为产生增强数据作为绝对增强数据和相对增强数据的一种组合。

8. 如权利要求 7 所述的多声道音频编码器 (10), 其中, 所述组合包含绝对增强数据和相对增强数据之间的一种插值。

9. 如权利要求 1 所述的多声道音频编码器 (10), 其中用于产生已编码输出信号的装置 (120) 设置为包含指示使用了相对增强数据还是绝对增强数据的数据。

10. 如权利要求 1 所述的多声道音频编码器 (10), 其中第二增强数据包含增强数据的第一部分和增强数据的第二部分, 该第二部分提供比第一部分更高质量的的第一 M- 声道信号表示。

11. 如权利要求 10 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为仅在产生第二部分作为绝对增强数据或者作为相对增强数据之间动态地选择。

12. 如权利要求 10 所述的多声道音频编码器 (10), 其中增强装置 (123) 设置为产生相对于通过将第一部分的增强数据应用于第一 M- 声道信号而产生的参考信号的第二部分的相对数据。

13. 一种多声道音频解码器 (20), 用于解码 N- 声道音频信号, 该多声道音频解码器

(20) 包括：

用于接收已编码音频信号的装置 (210), 该已编码音频信号包括：

用于 N- 声道音频信号的第一 M- 声道信号, 其中 M 小于 N,

用于多声道扩展的第一增强数据, 其中第一增强数据是相对于与第一 M- 声道信号不同的第二 M- 声道信号的；

相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据, 该第二增强数据包含绝对增强数据和相对于第一 M- 声道信号的相对增强数据, 和

指示数据, 指示信号块的第二增强数据是绝对增强数据还是相对增强数据；

用于响应于第一 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号的产生装置 (212) ; 和

用于响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号的装置 (220) ; 并且其中产生装置 (212) 设置为响应于指示数据而将第二增强数据作为绝对增强数据或相对增强数据应用之间进行选择。

14. 如权利要求 13 所述的多声道音频解码器 (20), 其中产生装置 (212) 设置为在时域中将第二增强数据应用于第一 M- 声道信号。

15. 如权利要求 13 所述的多声道音频解码器 (20), 其中产生装置 (212) 设置为在频域中将第二增强数据应用于第一 M- 声道信号。

16. 如权利要求 13 所述的多声道音频解码器 (20), 其中第二增强数据包含增强数据的第一部分和增强数据的第二部分, 该第二部分提供一种质量比第一部分更高的第一 M- 声道信号表示。

17. 如权利要求 16 所述的多声道音频解码器 (20), 其中产生装置 (212) 设置为仅在将第二增强数据的第二部分作为绝对增强数据或者相对增强数据应用之间进行选择。

18. 如权利要求 16 所述的多声道音频解码器 (20), 其中第二增强数据的第二部分包括相对增强数据, 产生装置 (212) 设置为通过将第二部分的相对增强数据应用到通过将第一部分的增强数据应用到第一 M- 声道信号而产生的信号而产生 M- 声道多声道扩展。

19. 一种编码 N- 声道音频信号的方法, 该方法包括：

为 N- 声道音频信号产生第一 M- 声道信号, M 小于 N；

产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据；

为 N- 声道音频信号产生第二 M- 声道信号；

产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据；

产生包括该第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号；并且其中第二增强数据的产生包括在产生第二增强数据作为绝对增强数据或者作为相对于第二 M- 声道信号的相对增强数据之间动态地选择。

20. 一种解码 N- 声道音频信号的方法, 该方法包括：

接收已编码音频信号, 该已编码音频信号包括：

用于 N- 声道音频信号的第一 M- 声道信号, M 小于 N,

用于多声道扩展的第一增强数据, 该第一增强数据是相对于不同于第一 M- 声道信号的第二 M- 声道信号的；

相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据, 该第二增强数据包括绝

对增强数据和相对于第一 M- 声道信号的相对增强数据,和

指示数据,指示信号块的第二增强数据是绝对增强数据还是相对增强数据;

响应于第一 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号;和

响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号;并且其中 M- 声道多声道扩展信号的产生包括响应于指示数据而在将第二增强数据作为绝对增强数据或者相对增强数据应用之间进行选择。

21. 一种用于传送已编码多声道音频信号的发送器(40),该发送器(40)包含根据权利要求1的多声道音频编码器(10)。

22. 一种用于接收多声道音频信号的接收器(50),该接收器(50)包含根据权利要求13的多声道音频解码器(20)。

23. 一种传送系统(70),包括:用于通过传送信道(30)传送已编码多声道音频信号到接收器(50)的发送器(40),该发送器(40)包含根据权利要求1的多声道音频编码器(10),而且该接收器包含根据权利要求13的多声道音频解码器(20)。

24. 一种传送已编码多声道音频信号的方法,该方法包括对 N- 声道音频信号编码,其中该编码包括:

产生用于 N- 声道音频信号的第一 M- 声道信号, M 小于 N;

产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据;

为 N- 声道音频信号产生第二 M- 声道信号;

产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据;

产生包含第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号;并且

其中第二增强数据的产生包含在产生第二增强数据作为绝对增强数据或者相对于第二 M- 声道信号的相对增强数据之间动态选择。

25. 一种接收已编码多声道音频信号的方法,该方法包含对已编码多声道音频信号进行解码,该解码包括:

接收已编码多声道音频信号,该已编码多声道音频信号包含:

用于 N- 声道音频信号的第一 M- 声道信号, M 小于 N,

用于多声道扩展的第一增强数据,该第一增强数据是相对于不同于第一 M- 声道信号的第二 M- 声道信号的;

相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据,该第二增强数据包含绝对增强数据和相对于第一 M- 声道信号的相对增强数据,和

指示数据,指示信号块的第二增强数据是绝对增强数据还是相对增强数据;

响应于第一 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号;和

响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号;并且其中 M- 声道多声道扩展信号的产生包含响应于指示数据而在将第二增强数据作为绝对增强数据或者相对增强数据应用之间进行选择。

26. 一种传送和接收音频信号的方法,该方法包括:

对 N- 声道音频信号进行编码,其中该编码包括:

为 N- 声道音频信号产生第一 M- 声道信号, M 小于 N,

产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据,

为 N- 声道音频信号产生第二 M- 声道信号，

产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据，该第二增强数据的产生包括在产生第二增强数据作为绝对增强数据或者作为相对于第二 M- 声道信号的相对增强数据之间动态选择，

产生包括第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号；

从发送器到接收器传送该已编码输出信号；

在接收器接收该已编码输出信号；

对该已编码输出信号进行解码，其中该解码包括：

响应于第二 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号，该 M- 声道多声道扩展信号的产生包括在将第二增强数据作为绝对增强数据或者相对增强数据应用之间进行选择，和

响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号。

27. 一种多声道音频记录器 (60)，该记录器 (60) 包括根据权利要求 1 的多声道音频编码器 (10)。

28. 一种多声道音频播放器 (60)，该播放器 (60) 包括根据权利要求 13 的多声道音频解码器 (20)。

## 多声道音频编码和解码方法、编码器和解码器

### 技术领域

[0001] 本发明涉及用于多声道信号的音频编码和 / 或解码。

### 背景技术

[0002] 多声道音频信号是具有两个或更多音频声道的音频信号。众所周知的多声道音频信号的例子是两声道立体声音频信号和具有两个前置音频声道、两个后置音频声道、一个中心音频信号和一个附加的低频增强 (LFE) 声道的 5.1 声道音频信号。该 5.1 声道音频信号被用在 DVD (数字化多功能盘) 和 SACD (超级音频压缩盘) 系统中。由于多声道资料的日益普及,多声道资料的高效编码变得更加重要。

[0003] 在音频处理领域,将一定数量的音频声道向另一一定数量的音频声道的转换是为人熟知的。这样的转换可能由于多种原因而进行。例如,音频信号可以被转换为其它格式以提供增强的用户体验。例如,传统的立体声录制品仅仅包含两个声道而现代高级音频系统一般使用五个或六个声道,像在流行的 5.1 环绕声系统中一样。于是,两立体声声道可以转换为五个或六个声道以便发挥高级音频系统的全部优势。

[0004] 声道转换的另一个原因是编码效率。已经发现,例如环绕声音频信号能编码为与参数比特流结合的立体声声道音频信号,该参数比特流描述音频信号的多声道空间特性。解码器能重建具有令人满意的准确程度的环绕声音频信号。这样,可以获得实质的比特速率节省。

[0005] 5.1-2-5.1 多声道音频编码系统是公知的。在这种公知的音频编码系统中,5.1 输入音频信号被编码并表示为两个下混合声道和相关参数。这些下混合信号被整体称为空间下混合。在这种公知系统中,空间下混合形成具有立体声映像的立体声音频信号,即,在质量上,媲美从 5.1 输入声道而来的固定 ITU 下混合。只拥有立体声装备的用户能听这种空间立体声下混合,而拥有 5.1 声道装备的听众能听 5.1 声道再现,其利用这种空间立体声下混合和相关参数制成。5.1 声道装备从空间立体声下混合 (例如,立体声音频信号) 和相关参数中解码 / 重建 5.1 声道音频信号。

[0006] 但是,与原始的立体声信号或者显式产生的立体声信号相比,空间立体声下混合常常被视为降低了质量的。例如,专业的演播室工程师常常发现空间立体声下混合有些呆板枯燥。因为这个原因,常常产生不同于空间立体声下混合的艺术立体声下混合。例如增加额外的混响或者声源,加宽立体声映像,等等。为了使用户能够欣赏艺术立体声下混合,这种艺术下混合而不是空间下混合可以通过传输媒体进行传送或者存储在存储介质上。然而,因为用于从立体声信号中产生 5.1 信号的数据基于原始下混合信号,这一方法严重地影响了 5.1 声道音频信号再生的质量。特别地,输入 5.1 声道音频信号被编码为空间立体声下混合和相关参数。通过用艺术立体声下混合代替空间立体声下混合,空间立体声下混合可能不再可用于系统解码端,并且 5.1 声道音频信号的高质量重建成为不可能。

[0007] 一种改进 5.1 声道音频信号质量的可能途径是包含更多的空间立体声下混合信号的数据。例如,除艺术立体声下混合之外,空间立体声下混合信号可以被包含在同一比特

流中,或者被平行传送。然而,这实质上提高了数据率,因而增加了通信带宽或者存储的需求,并将降低已编码多声道信号的质量 - 数据率比。

[0008] 因此,一种改进的多声道音频的编码 / 解码系统将是有利的,特别是一种允许改进性能、质量和 / 或质量 - 数据率比的系统将是有利的。

### 发明内容

[0009] 因此,本发明致力于更好地减轻、缓和或消除一个或多个上面提到的缺点,无论是单个地还是以某种组合形式。

[0010] 根据本发明的第一方面,提供一种多声道音频编码器,用于编码 N- 声道音频信号,该多声道音频编码器包括:用于为该 N- 声道音频信号产生第一 M- 声道信号的装置, M 小于 N;用于产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据的装置;用于为 N- 声道音频信号产生第二 M- 声道信号的装置;用于产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据的增强装置;用于产生包含第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号的装置;并且其中增强装置设置成可以在产生第二增强数据作为绝对增强数据或者作为相对于第二 M- 声道信号的相对增强数据之间动态地选择。

[0011] 本发明可以允许一种多声道信号的高效编码。特别地,能实现一种增加了质量 - 数据率比的高效编码。本发明可以允许一个 M- 声道信号替代另一个 M- 声道信号,并且对基于对第一 M- 声道信号相关的增强数据的多声道产生的影响得到降低。特别地,在允许基于与空间下混合相关的增强数据在解码器上实现高效多声道重建的同时,一种艺术下混合而不是空间下混合可被传送。增强数据的动态选择允许增强数据大小显著地减小和 / 或可产生信号的质量的提高。

[0012] 绝对增强数据不参照第二 M- 声道信号描述第一 M- 声道信号,而相对增强数据参照第二 M- 声道信号的描述第一 M- 声道信号。

[0013] 用于产生第一和 / 或第二 M- 声道信号的装置可以通过处理 N- 声道信号或者例如通过从内部或外部源中接收 M- 声道信号来产生信号。

[0014] 根据本发明的一个可选特征,增强装置设置为响应于 N- 声道信号的特性,在绝对增强数据和相对增强数据之间进行选择。

[0015] 这可以允许一种高效性能并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。例如,该选择可以通过评测从 N- 声道信号的一个片段的特性导出的一个或更多参数来执行,并且特别地基于从第一和 / 或第二 M- 声道信号(其自身能从 N- 声道信号导出)导出的一个或更多参数。

[0016] 根据本发明的一个可选特征,增强装置设置为响应于绝对增强数据和相对增强数据的相对特性,在绝对增强数据和相对增强数据之间进行选择。

[0017] 这可以允许一种高效性能并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。可替换地或者另外地,其可以允许一种高效的和 / 或低复杂度的实现。

[0018] 根据本发明的一个可选特征,该相对特性是相对于相对增强数据的信号能量的绝对增强数据的信号能量。

[0019] 这可以允许一种高效性能并且特别地可以提供一种提高了质量 - 数据率比的已

编码信号。可替换地或者另外地,其可以允许一种高效的和 / 或低复杂度的实现。特别地,增强装置可以选择具有最低信号 能量的增强数据类型。

[0020] 根据本发明的一个可选特征,该增强装置设置为将第二 M- 声道信号划分成信号块,并个别地为每一个信号块在绝对增强数据和相对增强数据之间进行选择。

[0021] 这可以允许一种高效性能并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。可替换地或者另外地,其可以允许一种高效的和 / 或低复杂度的实现。信号块可以在时域和 / 或频域中划分,而且每一个信号块可以特定地包含一组时间 / 频率片。这种到信号块的划分可以应用于第一 M- 声道信号和 / 或 N- 声道信号。

[0022] 根据本发明的一个可选特征,该增强装置设置为仅基于与一个信号块相关的特性,为该信号块在绝对增强数据和相对增强数据之间进行选择。

[0023] 这可以允许一种高效性能并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。可替换地或者另外地,其可以允许一种高效的和 / 或低复杂度的实现。特别地,增强装置可以选择具有最低信号能量的增强数据类型。

[0024] 根据本发明的一个可选特征,在产生增强数据作为绝对增强数据和作为相对增强数据之间切换的切换时间间隔期间,增强装置设置为将增强数据产生为绝对增强数据和相对增强数据的一种组合。

[0025] 这可以允许改进的切换,并且特别地可以减少与切换相关的人工噪声。可以得到改善的声音质量。当从绝对到相对增强数据和 / 或从相对到绝对增强数据进行切换时,在切换时间间隔期间的组合可以得到应用。该组合可以使用重叠相加技术得到。

[0026] 根据本发明的一个可选特征,该组合包含绝对增强数据和相对增强数据之间的一种插值。

[0027] 这可以允许一种具有高质量的实用且高效的实现。改善的声音质量可以被获得。

[0028] 根据本发明的一个可选特征,用于产生已编码输出信号的装置设置为包含指示使用了相对增强数据还是绝对增强数据的数据。

[0029] 这可以允许一种高效的性能,并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的实现。该指示数据可以特定地包含每一个信号块的一个选择指示。

[0030] 根据本发明的一个可选特征,第二增强数据包含增强数据的第一部分和增强数据的第二部分,该第二部分提供比第一部分更高质量的的第一 M- 声道信号的表示。

[0031] 这可以允许一种高效的性能,并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。第一部分可以比第二部分具有更低的数据率。第二部分可以包括更加准确地允许编码器重建第一 M- 声道信号的数据。

[0032] 根据本发明的一个可选特征,该增强装置设置为仅在产生第二部分作为绝对增强数据或者作为相对增强数据之间动态地选择。

[0033] 这可以允许一种高效的性能,并且特别地可以提供一种提高了质量 - 数据率比的已编码信号。

[0034] 根据本发明的一个可选特征,该增强装置设置为产生相对于通过将第一部分的增强数据应用于第一 M- 声道信号而产生的参考信号的第二部分的相对数据。

[0035] 这可以允许一种高效的性能,并且特别地可以提供一种提高了质量 - 数据率比的



已编码信号。

[0036] 根据本发明的另一个方面,提供一种多声道音频解码器,用于解码 N- 声道音频信号,该多声道音频解码器包括:用于接收已编码音频信号的装置,该已编码音频信号包含用于 N- 声道音频信号的第一 M- 声道信号,其中 M 小于 N,和用于多声道扩展的第一增强数据,其中第一增强数据是相对于与第一 M- 声道信号不同的第二 M- 声道信号的;相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据,该第二增强数据包含绝对增强数据和相对于第一 M- 声道信号的相对增强数据,和指示信号块的第二增强数据是绝对增强数据还是相对增强数据的指示数据;用于响应于第一 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号的产生装置;和用于响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号的装置;并且其中产生装置设置为响应于指示数据而将第二增强数据作为绝对增强数据或相对增强数据应用之间进行选择。

[0037] 本发明可以允许一种高效且高性能的多声道信号的解码。特别地,对于给定的数据率,可以实现信号的具有改善质量的高效解码。本发明可以允许一个 M- 声道信号代替另一个 M- 声道信号,其中减少了对基于相对于第一 M- 声道信号的增强数据的多声道产生的影响。特别地,一种艺术下混合而不是空间下混合可以被传送,同时允许基于与空间下混合相关的增强数据在解码器中实现高效的多声道重建。

[0038] 绝对增强数据不参照第一 M- 声道信号描述第二 M- 声道信号,而相对增强数据参照第一 M- 声道信号描述第二 M- 声道信号。

[0039] 根据本发明的一个可选特征,该产生装置设置为在时域中将第二增强数据应用于第一 M- 声道信号。

[0040] 这可以允许一种高效的性能,并且特别地可以提供一种对于给定的数据率具有提高了的的质量的解码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的实现。

[0041] 根据本发明的一个可选特征,该产生装置设置为在频域中将第二增强数据应用于第一 M- 声道信号。

[0042] 这可以允许一种高效的性能,并且特别地可以提供一种对于给定的数据率具有提高了的的质量的解码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的实现。

[0043] 特别地,在许多实施例中,频域应用可以减少所需频率变换的次数。例如,该频域可以是正交镜像滤波器组 (QMF) 或者改进型离散余弦变换 (MDCT) 域。

[0044] 根据本发明的一个可选特征,该第二增强数据包含增强数据的第一部分和增强数据的第二部分,该第二部分提供一种质量比第一部分更高的第一 M- 声道信号的表示。

[0045] 这可以允许一种高效的性能,并且特别地可以提供一种对于给定的数据率具有提高了的的质量的解码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的实现。该第二部分可包含允许解码器更加准确地重建第一 M- 声道信号的数据。

[0046] 根据本发明的一个可选特征,该产生装置设置为仅在将第二部分的第二增强数据用作绝对增强数据或者相对增强数据之间进行选择。

[0047] 这可以允许一种高效的性能,并且特别地可以提供一种对于给定的数据率具有提高了的的质量的解码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的

实现。

[0048] 根据本发明的一个可选特征,该产生装置设置为通过将第二部分 的相对增强数据应用到通过将第一部分的增强数据应用到第一 M- 声道信号而产生的信号而产生 M- 声道多声道扩展。

[0049] 这可以允许一种高效的性能,并且特别地可以提供一种对于给定的数据率具有提高了的的质量的解码信号。可替换地或者另外地,这可以允许一种高效的和 / 或低复杂度的实现。

[0050] 根据本发明的另一个方面,提供一种编码 N- 声道音频信号的方法,该方法包括:为 N- 声道音频信号产生第一 M- 声道信号, M 小于 N;产生相对于 N- 声道音频信号的第一 M- 声道信号的第一增强数据;为 N- 声道音频信号产生第二 M- 声道信号;产生相对于第一 M- 声道信号的第二 M- 声道信号的第二增强数据;产生包括该第二 M- 声道信号、第一增强数据和第二增强数据的已编码输出信号;并且其中第二增强数据的产生包括在产生第二增强数据作为绝对增强数据或者作为相对于第二 M- 声道信号的相对增强数据之间的动态地选择。

[0051] 根据本发明的另一个方面,提供一种解码 N- 声道音频信号的方法,该方法包括:接收已编码音频信号,其包括:用于 N- 声道音频信号的第一 M- 声道信号, M 小于 N,用于多声道扩展的第一增强数据,该第一增强数据是相对于不同于第一 M- 声道信号的第二 M- 声道信号的;相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据,该第二增强数据包括绝对增强数据和相对于第一 M- 声道信号的相对增强数据,和指示信号块的第二增强数据是绝对增强数据还是相对增强数据的指示数据;响应于第一 M- 声道信号和第二增强数据而产生 M- 声道多声道扩展信号;和响应于 M- 声道多声道扩展信号和第一增强数据而产生 N- 声道解码信号;并且其中 M- 声道多声道扩展信号的产生包括响应于指示数据而在将第二增强数据用作绝对增强数据或者相对增强数据之间的选择。

[0052] 根据本发明的另一个方面,提供一种用于 N- 声道音频信号的已编码多声道音频信号,包括:用于 N- 声道音频信号的 M- 声道信号数据, M 小于 N;用于多声道扩展的第一增强数据,该第一增强数据是相对于不同于第一 M- 声道信号的第二 M- 声道数据的;相对于第二 M- 声道信号的第一 M- 声道信号的第二增强数据,该第二增强数据包括绝对增强数据和相对于第一 M- 声道信号的相对增强数据;和指示信号块的第二增强数据是绝对增强数据还是相对增强数据的指示数据。

[0053] 根据本发明的另一个方面,提供一种其上存储有上述信号的存储媒体。

[0054] 根据本发明的另一个方面,提供一种发送器,用于传送已编码多声道音频信号,该发送器包含上述的多声道音频编码器。

[0055] 根据本发明的另一个方面,提供一种接收器,用于接收多声道音频信号,该接收器包含上述的多声道音频解码器。

[0056] 根据本发明的另一个方面,提供一种传送系统,包括:用于通过传送信道传送已编码多声道音频信号到接收器的发送器,该发送器包含上述的多声道音频编码器,而且该接收器包含上述的多声道音频解码器。

[0057] 根据本发明的另一个方面,提供一种传送已编码多声道音频信号的方法,该方法包括对 N- 声道音频信号的编码,其中编码包括:产生用于 N- 声道音频信号的第一 M- 声道

信号,  $M$  小于  $N$ ; 产生相对于  $N$ - 声道音频信号的第一  $M$ - 声道信号的第一增强数据; 为  $N$ - 声道音频信号产生第二  $M$ - 声道信号; 产生相对于第一  $M$ - 声道信号的第二  $M$ - 声道信号的第二增强数据; 产生包含第二  $M$ - 声道信号、第一增强数据和第二增强数据的已编码输出信号; 并且其中第二增强数据的产生包含在产生第二增强数据作为绝对增强数据或者相对于第二  $M$ - 声道信号的相对增强数据之间的动态选择。

[0058] 根据本发明的另一个方面, 提供一种接收已编码多声道音频信号的方法, 该方法包含对已编码多声道音频信号进行解码, 该解码包括: 接收包含用于  $N$ - 声道音频信号的第一  $M$ - 声道信号的已编码多声道音频信号,  $M$  小于  $N$ , 用于多声道扩展的第一增强数据, 该第一增强数据是相对于不同于第一  $M$ - 声道信号的第二  $M$ - 声道信号的; 相对于第二  $M$ - 声道信号的第一  $M$ - 声道信号的第二增强数据, 该第二增强数据包含绝对增强数据和相对于第一  $M$ - 声道信号的相对增强数据, 和指示信号块的第二增强数据是绝对增强数据还是相对增强数据的指示数据; 响应于第一  $M$ - 声道信号和第二增强数据而产生  $M$ - 声道多声道扩展信号; 和响应于  $M$ - 声道多声道扩展信号和第一增强数据而产生  $N$ - 声道解码信号; 并且其中  $M$ - 声道多声道扩展信号的产生包含响应于指示数据而在将第二增强数据用作绝对增强数据或者相对增强数据之间的选择。

[0059] 根据本发明的另一个方面, 提供一种传送和接收音频信号的方法, 该方法包括: 对  $N$ - 声道音频信号进行编码, 其中编码包括: 为  $N$ - 声道音频信号产生第一  $M$ - 声道信号,  $M$  小于  $N$ , 产生相对于  $N$ - 声道音频信号的第一  $M$ - 声道信号的第一增强数据, 为  $N$ - 声道音频信号产生第二  $M$ - 声道信号, 产生相对于第一  $M$ - 声道信号的第二  $M$ - 声道信号的第二增强数据, 该第二增强数据的产生包括在产生第二增强数据作为绝对增强数据或者作为相对于第二  $M$ - 声道信号的相对增强数据之间的动态选择, 产生包括第二  $M$ - 声道信号、第一增强数据和第二增强数据的已编码输出信号; 从发送器到接收器传送该已编码输出信号; 在接收器接收该已编码输出信号; 对该已编码输出信号进行解码, 其中解码包括: 响应于第二  $M$ - 声道信号和第二增强数据而产生  $M$ - 声道多声道扩展信号, 该  $M$ - 声道多声道扩展信号的产生包括在将第二增强数据作为绝对增强数据或者相对增强数据应用之间的选择, 和响应于  $M$ - 声道多声道扩展信号和第一增强数据而产生  $N$ - 声道解码信号。

[0060] 根据本发明的另一个方面, 提供一种计算机程序产品, 该计算机程序产品可操作用来使处理器执行上述方法的步骤。

[0061] 根据本发明的另一个方面, 提供一种多声道音频记录器, 该记录器包括上述的多声道音频编码器。

[0062] 根据本发明的另一个方面, 提供一种多声道音频播放器 (60), 该播放器包括上述的多声道音频解码器。

[0063] 本发明的这些和其它方面、特征和优点将参照下文所述的实施例进行展示和说明。

#### 附图说明

[0064] 本发明的实施例将仅以举例的方式参照附图进行描述, 其中:

[0065] 图 1 显示了根据本发明一些实施例的多声道音频编码器的方框图;

[0066] 图 2 显示了根据本发明一些实施例的多声道音频解码器的方框图;

- [0067] 图 3 显示了根据本发明一些实施例的传送系统的方框图；
- [0068] 图 4 显示了根据本发明一些实施例的多声道音频播放器 / 记录器 的方框图；
- [0069] 图 5 显示了根据本发明一些实施例的多声道音频编码器的方框图；
- [0070] 图 6 显示了根据本发明一些实施例的增强数据发生器的方框图；
- [0071] 图 7 显示了根据本发明一些实施例的多声道音频解码器的方框图；
- [0072] 图 8 显示了多声道音频解码器的元件的方框图；
- [0073] 图 9 显示了根据本发明一些实施例的多声道音频解码器的元件的方框图；
- [0074] 图 10 显示了根据本发明一些实施例的多声道音频解码器的元件的方框图；和
- [0075] 图 11 显示了根据本发明一些实施例的多声道音频解码器的元件的方框图。

### 具体实施方式

[0076] 下面的描述集中在本发明适用于 5.1 到 2 编码器和 / 或 2 到 5.1 解码器的实施例。但是,应当理解,本发明并不限于这一应用。

[0077] 图 1 显示了根据本发明一些实施例的多声道音频编码器 10 的一个实施例的方框图。多声道音频编码器 10 设置为将 N 个音频信号 101 编码为 M 个音频信号 102 和相关参数数据 104、105。其中, M 和 N 是整数,  $N > M$  且  $M \geq 1$ 。多声道音频编码器 10 的一个例子是 5.1 到 2 编码器, 其中 N 等于 6, 即 5+1 声道, 且 M 等于 2。这种多声道音频编码器将 5.1 声道输入音频信号编码为 2 声道输出音频信号, 例如立体声输出音频信号, 和相关参数。多声道音频编码器 10 的其它例子是 5.1 到 1、6.1 到 2、6.1 到 1、7.1 到 2 和 7.1 到 1 编码器。具有其它 N 和 M 值的编码器也是可能的, 只要 N 大于 M 且 M 大于等于 1。

[0078] 编码器 10 包括第一编码单元 110 和与之相耦接的第二编码单元 120。第一编码单元 110 接收 N 个输入音频信号 101 并将 N 个音频信号 101 编码为 M 个音频信号 102 和第一相关参数数据 104。M 个音频信号 102 和第一相关参数数据 104 代表 N 个音频信号 101。由第一单元 110 执行的从 N 个音频信号 101 到 M 个音频信号 102 的编码也可被称为一种下混合, 而 M 个音频信号 102 也可被称为空间下混合 102。单元 110 可以是传统的参数化多声道音频编码器, 其将多声道音频信号 101 编码为单声或立体声下混合音频信号 102 和相关参数 104。相关参数 104 使得解码器可以从单声或立体声下混合音频信号 102 重建多声道音频信号 101。需要注意的是, 下混合 102 也可以具有多于两个的声道。

[0079] 第一单元 110 将空间下混合 102 提供给第二单元 120。第二单元 120 从空间下混合 102 产生具有第二相关参数数据 105 形式的第二增强数据。第二相关参数数据 105 代表空间下混合 102, 即, 这些参数 105 包含空间下混合 102 的特征或特性, 其使得解码器至少能够重建空间下混合 102 的一部分, 例如, 通过合成类似空间下混合 102 的信号。该相关参数数据包括第一和第二相关参数数据 104 和 105。

[0080] 第二相关参数数据 105 包含修正参数, 该修正参数使得能够从  $K (= M)$  个进一步的音频信号 103 重建空间下混合 102。通过这种方式, 解码器可以完成空间下混合 102 的一种更好的重建。这种重建可以在一种诸如艺术下混合之类的可替换下混合 103, 即, K 个进一步的音频信号 103 的基础上进行。解码器可以将修正参数应用于该可替换下混合信号 103, 以使其更加接近地类似空间下混合 102。

[0081] 第二单元 120 可以在其输入端接收该可替换下混合 103。该可替换下混合 103 可

以从编码器 10 外部的源接收（如图 1 所示）或者，可替换地，该可替换下混合 103 可以在编码器 10 内部产生，例如从 N 个音频信号 101 产生（没有图示）。第二单元 120 可以将空间下混合 102 的至少一部分与可替换下混合 103 相比较，并且产生代表空间下混合 102 和可替换下混合 103 之间的差别的修正参数 105，例如，空间下混合 102 特性和可替换下混合 103 特性之间的差别。在本例中，可替换下混合 103 具体而言是与空间下混合相关联的艺术下混合。

[0082] 在本例中，第二单元 120 可以进一步地产生修正参数作为绝对值，其直接代表空间下混合 102 而没有参照可替换下混合 103。此外，第二单元 120 包含用于为编码器输出信号在相对和绝对修正参数之间进行选择的功能。特别地，这一选择是动态地执行的，并且可以根据信号和 / 或参数数据的特征为个别的信号块进行。

[0083] 另外，第二单元 120 可以包括这样的功能：包含哪个修正参数（绝对的或相对的）被用于已编码信号的不同片段的指示。例如，每一个信号块可以包含一个数据比特，以指示是相对还是绝对参数数据被包含用于该信号块。

[0084] 修正参数 105 优选包括一个或更多统计信号特性（之间的差别），像方差、协方差和相关，或者这些特性的比或下混合信号（之间的差别）的比。需要注意的是，一个信号的方差等于该信号的能量或者功率。这些统计信号特性使空间下混合的良好重建能够获得。

[0085] 图 2 显示了根据本发明一些实施例的多声道音频解码器 20 的一个实施例的方框图。解码器 20 设置为将 K 个音频信号 103 和相关参数数据 104、105 解码为 N 个音频信号 203。其中，K 和 N 是整数， $N > K$  且  $K \geq 1$ 。K 个音频信号 103，即，可替换下混合 103，和相关参数数据 104、105 代表 N 个音频信号 203，即，多声道音频信号 203。多声道音频解码器 20 的一个例子是 2 到 5.1 解码器，其中 N 等于 6，即 5+1 声道，且 K 等于 2。这种多声道音频解码器将 2 声道输入音频信号，例如立体声输入音频信号，和相关参数解码为 5.1 声道输出音频信号。多声道音频解码器 20 的其它例子是 1 到 5.1、2 到 6.1、1 到 6.1、2 到 7.1 和 1 到 7.1 解码器。具有其它 N 和 K 值的解码器也是可能的，只要 N 大于 K 且 K 大于等于 1。

[0086] 多声道音频解码器 20 包括第一单元 210 和与之相耦接的第二单元 220。第一单元 210 接收可替换的下混合 103 和修正参数 105 形式的增强数据，并从可替换的下混合 103 和修正参数 105 中重建 M 个进一步的音频信号 202，即，空间下混合 202 或其近似。其中，M 是整数，且  $M \geq 1$ 。修正参数 105 代表空间下混合 202。第一单元 210 特别地设置为确定修正参数 105 是绝对还是相对修正参数，并相应地使用该参数。特别地，第一单元 210 能基于接收到的比特流中的显式数据确定单独信号块的修正参数 105 是相对还是绝对参数。例如，一个单独的数据比特可以包含在每一个信号块中以指示该参数是该信号块中的绝对还是相对修正参数。

[0087] 第二单元 220 从第一单元 210 中接收空间下混合 202 和修正参数 104。第二单元 220 将空间下混合 202 和修正参数 104 解码为多声道音频信号 203。第二单元 220 可以是传统的参数化多声道音频解码器，其将单声或立体声下混合信号 202 和相关参数 104 解码为多声道音频信号 203。

[0088] 第一单元 210 可以被安排用于确定从输入信号 103 中重建信号 202 是否是必需的或者期望的。当提供给第一单元 210 的是空间下混合信号 202 而不是可替换的下混合 103 时，这样的重建是不可用的。通过从输入信号 103 中产生包含在修正参数 105 中的类似的

或者相同的信号特性,并通过将这些产生的信号特性与修正参数 105 相比较,第一单元 210 能够确定这一状况。如果这种比较显示产生的信号特性等同于或者基本上等同于修正参数 105,那么输入信号 103 就充分地类似空间下混合信号 202,并且第一单元 210 可以转发输入信号 103 到第二单元 220。如果该比较显示产生的信号特性不等同于或者基本上不等同于修正参数 105,则输入信号 103 不能充分地类似空间下混合信号 202,并且第一单元 210 可以从输入信号 103 和修正参数 105 中重建 / 估计空间下混合信号 202。

[0089] 第一单元 210 可以从可替换下混合中产生表示可替换下混合 103 的进一步的修正参数 / 特性。在这种情况下,第一单元 210 可以从可替换下混合 103 和修正参数 105 与进一步的修正参数 (之间的差别) 中重建空间下混合 202。

[0090] 修正参数 105 和进一步的修正参数分别可以包含空间下混合 202 和可替换下混合 103 的统计特性。这些统计特性,例如方差、相关和协方差等等,提供导出它们的信号的良好表示。它们在重建空间下混合 202 中是有用的,例如通过变换可替换下混合,这样其相关特性与包含在修正参数 105 中的特性相符合。

[0091] 图 3 显示了根据本发明一些实施例的传送系统 70 的一个实施例的方框图。传送系统 70 包括发送器 40,用于通过传送声道 30,例如有线或者无线的通信链路,传送已编码多声道音频信号到接收器 50。发送器 40 包括上述的多声道音频编码器 10,用于将多声道音频信号 101 编码为空间下混合 102 和相关参数 104、105。发送器 40 还包括用于通过传送声道 30 传送包含参数 104、105 和空间下混合 102 或可替换下混合 103 的已编码多声道音频信号到接收器 50 的装置 41。接收器 50 包括用于接收已编码多声道音频信号的装置 51 和用于将可替换下混合 103 或空间下混合 102 和相关参数 104、105 解码为多声道音频信号 203 的如上所述的多声道音频解码器 20。

[0092] 图 4 显示了根据本发明一些实施例的多声道音频播放器 / 记录器 60 的一个实施例的方框图。根据本发明的一些实施例,音频播放器 / 记录器 60 包括多声道音频解码器 20 和 / 或多声道音频变码器 10。音频播放器 / 记录器 60 可以具有自己的存储器,例如固态存储器或硬盘。音频播放器记录器 60 也可以配备可分离存储装置,例如 (可记录的)DVD 盘片或 (可记录的)CD 盘片。包含可替换下混合 103 和参数 104、105 的所存储的已编码多声道音频信号可以通过解码器 20 解码并通过音频播放器 / 记录器 60 播放或重现。编码器 10 可以为在存储装置上进行存储而编码多声道音频信号。

[0093] 图 5 显示了根据本发明一些实施例的多声道音频编码器 10 的方框图。特别地,图 5 的编码器可以是图 1 的编码器 10。编码器 10 包括第一单元 110 和与之相耦接的第二单元 120。第一单元 110 接收 5.1 多声道音频信号 101,其包含左前、左后、右前、右后、中央和低频增强音频信号,分别记为 lf、lr、rf、rr、co 和 lfe。第二单元 120 接收艺术立体声下混合 103,其包含左艺术和右艺术音频信号,分别记为 la 和 ra。多声道音频信号 101 和艺术下混合 103 是时域音频信号。在第一和第二单元 110 和 120 中,这些信号 101 和 103 被分段并变换到频率 - 时间域。

[0094] 在第一单元 110 中,参数数据 104 用三步导出。在第一步,三对音频信号 lf 与 rf、rf 与 rr、co 与 lfe 被分别分段,并且分段信号在分段和变换单元 112、113 和 114 中分别变换到频域。作为结果产生的分段信号的频域表示分别显示为频域信号 Lf、Lr、Rf、Rr、Co 和 LFE。在第二步,这三对频域信号 Lf 与 Lr、Rf 与 Rr、Co 与 LFE 分别在下混合器 115、116 和

117 中进行下混合,以分别产生单声音频信号 L、R 和 C 和相关参数 141、142 与 143。下混合器 115、116 和 117 可以是传统的 MPEG4 参数化立体声编码器。最后,在第三步,三个单声音频信号 L、R 和 C 在下混合器 118 中进行下混合以获得空间立体声下混合 102 和相关参数 144。该空间下混合 102 包含信号  $L_o$  和  $R_o$ 。

[0095] 参数数据 141、142、143 和 144 以第一相关参数数据 104 的形式包含在第一增强数据中。参数数据 104 和空间下混合 102 代表 5.1 输入信号 101。

[0096] 在第二单元,分别由音频信号  $l_a$  和  $r_a$  在时域内表示的艺术下混合信号 103 在分段单元 121 中进行第一次分段。所得到的分段音频信号 127 分别包含信号  $l_{as}$  和  $r_{as}$ 。接下来,该分段音频信号 127 通过变换器 122 变换到频域。所得到的频域信号 126 包含信号  $L_a$  和  $R_a$ 。最后,频域信号 126 (其是分段艺术下混合 103 的频域表达) 和分段空间下混合 102 的频域表达被提供给发生器 123,其以修正参数 105 的形式产生进一步的 (第二) 增强数据,使得解码器能够修正/变换艺术下混合 103,以便其更加接近地类似空间下混合 102。

[0097] 在该具体的例子中,分段时域信号 127 也馈送到选择器 124。该选择器 124 的另外两个输入是空间立体声下混合 102 的频域表示和控制信号 128。控制信号 128 确定选择器 124 是将艺术下混合 103 还是空间下混合 102 作为已编码多声道音频信号的一部分输出。当艺术下混合不可用时,空间下混合 102 可以被选出。控制信号 128 能手动设置或者通过感测艺术下混合 103 的存在而自动产生。控制信号 128 可以包含于参数比特流之中,以便相应的解码器 20 可以像后面所描述的那样利用它。因此,该特定的例示编码器允许产生包含空间下混合 102 或者艺术下混合 103 的信号。

[0098] 选择器 124 的输出信号 102、103 记为信号  $l_o$  和  $r_o$ 。若艺术立体声下混合 127 将被选择器 124 输出,那么分段时域信号  $l_{as}$  与  $r_{as}$  就在选择器 124 中通过重叠相加组合到信号  $l_o$  与  $r_o$  之中。如果空间立体声下混合 102 由控制信号 128 指示而被输出,那么选择器 124 把信号  $L_o$  和  $R_o$  变换回时域,并通过重叠相加到信号  $l_o$  与  $r_o$  之中将它们合并。时域信号  $l_o$  与  $r_o$  形成 5.1 到 2 编码器 10 的立体声下混合。

[0099] 下面给出发生器 123 的更详细的描述。发生器 123 的功能是确定第二增强数据,特别是修正参数,其描述艺术下混合 103 的变换,从而在某种意义上类似原始的空间下混合 102。

[0100] 一般而言,该变换可以描述为

$$[L_d \ R_d] = [L_a \ R_a \ A_1 \ \dots \ A_N]T \quad (1)$$

[0102] 其中,  $L_a$  和  $R_a$  是包含艺术下混合 103 左右声道时间/频率片的采样的向量,  $L_d$  和  $R_d$  是包含修正艺术下混合左右声道的时间/频率片的采样的向量,  $A_1, \dots, A_N$  包括可选辅助声道的时间/频率片的采样,  $T$  是变换矩阵。注意,任一向量  $\underline{v}$  都定义为列向量。修正艺术下混合是这样的艺术下混合 103,其通过变换而被变换以使其类似原始空间下混合 102。辅助声道  $A_1, \dots, A_N$  在所描述的系统是空间下混合信号或其低频内容。

[0103]  $(N+2) \times 2$  变换矩阵  $T$  描述从艺术下混合 103 和辅助声道向修正艺术下混合的变换。更优选地,变换矩阵  $T$  或其元素包含在修正参数 105 中以使得解码器 20 能重建变换矩阵  $T$  的至少一部分。此后,解码器 20 能将变换矩阵  $T$  应用于艺术下混合 103 以重建空间下混合 102 (如下面所述)。

[0104] 可替换地,修正参数 105 包含空间下混合 102 的信号特性,例如能量或功率值和/

或相关值。因此解码器 20 能够从艺术下混合 103 产生此类信号特性。空间下混合 102 和艺术下混合 103 的信号特性使得解码器 20 能够建立变换矩阵 T (下面将述及) 并将其应用于艺术下混合 103 以重建空间下混合 102 (下面也将述及)。

[0105] 特别地, 发生器 123 设置为产生相对和绝对修正数据两者并为个别的信号块 (或片段) 在该数据间进行选择。因此, 用于已编码信号的修正参数 105 包含用于不同信号块的绝对修正数据和相对修正数据两者。对照绝对修正数据, 相对修正数据相对于艺术下混合 103 描述空间下混合 102。特别地, 相对修正数据可以是差别数据, 其允许艺术下混合采样得到修正以相应于 (更加接近地) 空间下混合采样, 而绝对下混合数据可以直接地相应于空间下混合采样而不参照或依赖艺术下混合采样。

[0106] 将会理解的是, 存在多种途径, 可以修正艺术立体声下混合 103 使其类似原始立体声下混合 102, 这些途径包括:

[0107] I. 波形匹配。

[0108] II. 统计特性匹配:

[0109] a. 左右声道的能量或者功率匹配。

[0110] b. 左右声道的协方差矩阵匹配。

[0111] III. 得到在左右声道能量或者功率匹配约束下的波形的最佳可能的匹配。

[0112] IV. 上述方法 I-III 的混合。

[0113] 为清晰起见, 首先不考虑 (1) 中的辅助声道  $\underline{A}_1, \dots, \underline{A}_N$ , 因此, 变换矩阵 T 可以写成

$$[0114] \quad [\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a] T \quad (2)$$

[0115] 且相对增强数据可以如下述例子所示得到产生:

[0116] I. 波形匹配 (方法 I)

[0117] 艺术下混合 103 和空间下混合 102 的波形匹配可以通过将修正艺术下混合的左右信号表示为艺术立体声下混合 103 的左右信号的线性组合得到:

$$[0118] \quad \underline{L}_d = \alpha_1 \underline{L}_a + \beta_1 \underline{R}_a, \ \underline{R}_d = \alpha_2 \underline{L}_a + \beta_2 \underline{R}_a \quad (3)$$

[0119] 然后, (2) 中的矩阵 T 可以写成:

$$[0120] \quad T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}.$$

[0121] 选择参数  $\alpha_1$ 、 $\alpha_2$ 、 $\beta_1$  和  $\beta_2$  的一种途径是最小化空间下混合信号  $L_s$  和  $R_s$  与其估计 (即, 修正艺术下混合信号  $L_d$  和  $R_d$ ) 之间欧几里得距离的平方, 因此

$$[0122] \quad \min_{\alpha_1, \beta_1} \sum_k \|L_s[k] - L_d[k]\|^2 = \min_{\alpha_1, \beta_1} \sum_k \|L_s[k] - \alpha_1 L_a[k] - \beta_1 R_a[k]\|^2 \quad (4)$$

[0123] 且

$$[0124] \quad \min_{\alpha_2, \beta_2} \sum_k \|R_s[k] - R_d[k]\|^2 = \min_{\alpha_2, \beta_2} \sum_k \|R_s[k] - \alpha_2 L_a[k] - \beta_2 R_a[k]\|^2. \quad (5)$$

[0125] II. 统计特性匹配 (方法 II)

[0126] 方法 II. a: 现在讨论左右信号的能量匹配。分别记为  $L_d$  和  $R_d$  的修正左右艺术下混合信号, 现在可以如下计算:

$$[0127] \quad \underline{L}_d = \alpha \underline{L}_a, \ \underline{R}_d = \beta \underline{R}_a \quad (6)$$

[0128] 这里, 在实参数情况下,  $\alpha$  和  $\beta$  由下式给出:



$$[0129] \quad \alpha = \sqrt{\frac{\sum_k \|L_s[k]\|^2}{\sum_k \|L_a[k]\|^2}}, \beta = \sqrt{\frac{\sum_k \|R_s[k]\|^2}{\sum_k \|R_a[k]\|^2}}, \quad (7)$$

[0130] 因此变换矩阵 T 可以写成

$$[0131] \quad T = \begin{bmatrix} \sqrt{\frac{\sum_k \|L_s[k]\|^2}{\sum_k \|L_a[k]\|^2}} & 0 \\ 0 & \sqrt{\frac{\sum_k \|R_s[k]\|^2}{\sum_k \|R_a[k]\|^2}} \end{bmatrix}. \quad (8)$$

[0132] 利用这些选项,可以确保信号  $L_d$  和  $R_d$  分别具有和信号  $L_s$  和  $R_s$  相同的能量。

[0133] 方法 II. b:为了匹配艺术立体声下混合 103 和空间立体声下混合 102 的协方差矩阵,这些矩阵可以利用特征值分解做如下分解:

$$[0134] \quad C_a = U_a S_a U_a^H, \quad (9)$$

$$[0135] \quad C_0 = U_0 S_0 U_0^H,$$

[0136] 其中艺术立体声下混合 103 的协方差矩阵  $C_a$  由下式给出

$$[0137] \quad C_a = [\underline{L}_a \quad \underline{R}_a]^H [\underline{L}_a \quad \underline{R}_a]. \quad (10)$$

[0138]  $U_a$  是酉矩阵,  $S_a$  是对角矩阵。  $C_0$  是空间立体声下混合 102 的协方差矩阵,  $U_0$  是酉矩阵,  $S_0$  是对角矩阵。当计算

$$[0139] \quad X_{aw} = [\underline{L}_{aw} \quad \underline{R}_{aw}] = [\underline{L}_a \quad \underline{R}_a] U_a S_a^{-1/2} \quad (11)$$

[0140] 时,得到两个互不相关的信号  $\underline{L}_{aw}$  和  $\underline{R}_{aw}$  (由于与矩阵  $U_a$  相乘),两信号具有单位能量 (由于与矩阵  $S_a^{-1/2}$  相乘)。通过计算:

$$[0141] \quad X_d = [\underline{L}_d \quad \underline{R}_d] = [\underline{L}_a \quad \underline{R}_a] U_a S_a^{-1/2} U_r S_0^{1/2} U_0^H, \quad (12)$$

[0142] 首先,  $[\underline{L}_a \quad \underline{R}_a]$  的协方差矩阵变换为等于单位矩阵的协方差矩阵,即,  $[\underline{L}_a \quad \underline{R}_a] U_a S_a^{-1/2}$  的协方差矩阵。应用任意酉矩阵  $U_r$  不会改变协方差结构,而且应用  $S_0^{1/2} U_0^H$  导致等价于空间立体声下混合 102 的协方差结构。

[0143] 定义矩阵  $S_{0w}$  和信号  $\underline{L}_{0w}$  与  $\underline{R}_{0w}$  如下:

$$[0144] \quad S_{0w} = [\underline{L}_{0w} \quad \underline{R}_{0w}] = [\underline{L}_s \quad \underline{R}_s] U_0 S_0^{-1/2} \quad (13)$$

[0145] 矩阵  $U_r$  可以被选择为使得根据最小平方欧几里得距离的最佳可能的波形匹配在信号  $\underline{L}_{0w}$  与  $\underline{L}_{aw}$  和信号  $\underline{R}_{0w}$  与  $\underline{R}_{aw}$  之间得到,其中  $\underline{L}_{aw}$  和  $\underline{R}_{aw}$  由 (11) 给出。利用对  $U_r$  的这个选择,一种在统计方法之内的波形匹配可以得到应用。

[0146] 从 (12) 可以看出,变换矩阵 T 由下式给出:

$$[0147] \quad T = U_a S_a^{-1/2} U_r S_0^{1/2} U_0^H. \quad (14)$$

[0148] III. 能量约束下的最佳波形匹配 (方法 III)

[0149] 假设 (3) 中参数  $\alpha_1$ 、 $\alpha_2$ 、 $\beta_1$  和  $\beta_2$  可以通过在能量约束

$$[0150] \quad \sum_k \|L_s[k]\|^2 = \sum_k \|L_d[k]\|^2, \quad \sum_k \|R_s[k]\|^2 = \sum_k \|R_d[k]\|^2 \quad (15)$$

[0151] 下最小化 (4) 和 (5) 得到。

[0152] IV. 混合方法 (方法 IV)

[0153] 关于混合不同的方法,可能的组合包括混合方法 II. a 和 II. b, 或者混合方法 II. a 和 III。过程可以如下:

[0154] a) 若使用方法 II. b/III 时获得的  $L_s$  与  $L_d$  之间和  $R_s$  与  $R_d$  之间的波形匹配良好, 则使用方法 II. b/III。

[0155] b) 若该波形匹配差, 使用方法 II. a。

[0156] c) 通过混合其变换矩阵, 作为这种波形匹配的质量的函数, 保证两种方法之间的逐渐过渡。

[0157] 这可以做如下的数学表达:

[0158] 利用 (3) 和 (2), 变换矩阵 T 可以写成如下的一般形式:

$$[0159] \quad T = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}. \quad (16)$$

[0160] 该矩阵可以利用两个向量  $T_L$  和  $T_R$  改写如下:

$$[0161] \quad T = [T_L \quad T_R], \quad T_L = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}, \quad T_R = \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix}. \quad (17)$$

[0162]  $L_s$  与  $L_d$  之间波形匹配的质量可以通过使用方法 II. b 或方法 III 得到, 记为  $\gamma_L$ 。其定义为

$$[0163] \quad \gamma_L = \max \left( 0, \frac{\sum_k L_s[k] L_d^*[k]}{\sum_k |L_s[k]| |L_d[k]|} \right). \quad (18)$$

[0164]  $R_s$  与  $R_d$  之间波形匹配的质量可以通过使用方法 II. b 或方法 III 得到, 记为  $\gamma_R$ 。其定义为

$$[0165] \quad \gamma_R = \max \left( 0, \frac{\sum_k R_s[k] R_d^*[k]}{\sum_k |R_s[k]| |R_d[k]|} \right). \quad (19)$$

[0166]  $\gamma_L$  和  $\gamma_R$  都在 0-1 之间。左声道的混合系数  $\delta_L$  和右声道的混合系数  $\delta_R$  可以如下定义:

[0167]

$$[0167] \quad \delta_L = \begin{cases} 1 & \gamma_L > \mu_{L,\max} \\ 0 & \gamma_L < \mu_{L,\min} \\ \frac{1}{2} - \frac{1}{2} \cos \left( \pi \frac{(\gamma_L - \mu_{L,\min})}{(\mu_{L,\max} - \mu_{L,\min})} \right) & \text{其它} \end{cases}, \quad (20)$$

[0168]

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

[0169] 其中,  $\mu_{L,\min}$ 、 $\mu_{L,\max}$ 、 $\mu_{R,\min}$  和  $\mu_{R,\max}$  在 0-1 间取值,  $\mu_{L,\min} < \mu_{L,\max}$  且  $\mu_{R,\min} < \mu_{R,\max}$ 。等式 (20) 保证混合系数  $\delta_L$  和  $\delta_R$  在 0-1 间取值。

[0170] 分别定义方法 II. a、II. b 和 III 的变换矩阵 T 为由 (8) 给出的  $T_e$ 、由 (14) 给出的  $T_a$  和  $T_{ce}$ 。每一个变换矩阵都可以分成类似 (17) 中 T 的划分的两个向量如下:

$$[0171] \quad T_a = [T_{a,L} \ T_{a,R}], T_e = [T_{e,L} \ T_{e,R}], T_{ce} = [T_{ce,L} \ T_{ce,R}]. \quad (21)$$

[0172] 混合方法 II. a 和方法 II. b 的变换矩阵 T 可如下得到:

$$[0173] \quad T = [T_L \ T_R] = [\delta_L T_{a,L} + (1 - \delta_L) T_{e,L} \ \delta_R T_{a,R} + (1 - \delta_R) T_{e,R}]. \quad (22)$$

[0174] 混合方法 II. a 和方法 III 的变换矩阵 T 可如下得到:

$$[0175] \quad T = [T_L \ T_R] = [\delta_L T_{ce,L} + (1 - \delta_L) T_{e,L} \ \delta_R T_{ce,R} + (1 - \delta_R) T_{e,R}]. \quad (23)$$

[0176] 现在, 考虑对应于两个增强层声道的两个辅助声道, 上述式 (1) 可以改写为

$$[0177] \quad [L_d \ R_d] = [L_a \ R_a \ L_{enh} \ R_{enh}] T'. \quad (24)$$

[0178] 其中,  $L_a$ 、 $R_a$  (同前) 分别包含艺术下混合的左右声道的时间 / 频率片的采样,  $L_d$ 、 $R_d$  分别包含修正艺术下混合的左右声道的时间 / 频率片的采样, 且  $L_{enh}$ 、 $R_{enh}$  分别包含增强层信号的时间 / 频率片的采样。因此,  $4 \times 2$  变换矩阵  $T'$  描述从艺术下混合和增强层信号到修正艺术下混合的变换。联系式 (1), 这里仅使用的两个辅助声道是增强层信号  $L_{enh}$  和  $R_{enh}$ 。

[0179] 在该特定的示例系统中, 第二增强层可以包含两种不同类型的数据:

[0180] 第一种类型的数据包含式 (1) 的矩阵 T 中所包含的参数。在本例中, 对全部信号带宽计算这些参数, 并变换艺术立体声下混合以使得在某种意义上其类似于空间下混合。这样, 这种类型的参数可提供一种修正艺术下混合, 其更加接近地类似于原始的空间下混合但不 (必要) 允许解码器准确地产生空间下混合。在每一个时间 / 频率片上仅需四个参数, 即仅需 T 的值 ( $T_{11}$ 、 $T_{12}$ 、 $T_{21}$  和  $T_{22}$ )。这些参数可以绝对地或差分地编码, 且编码器 10 可以明确地在绝对和差分编码之间动态地切换。

[0181] 第二种类型的数据对应于实际的空间下混合, 且在该特定例子中是空间下混合的带宽受限版本的一种表示。特别地, 这种类型的数据表示空间下混合的低频部分 (例如, 低于 1.7kHz 的频率)。这使得在解码器非常准确地重建空间下混合的这一部分成为可能, 而不是 (像使用矩阵 T 那样) 只产生一种具有相同特性, 例如统计特性, 的信号。这种类型的数据可以绝对地或相对地编码为艺术下混合。特别地, 这种类型的数据能被差分地编码。举例来说, 变换矩阵 T 被用于艺术下混合 (例如, 见式 (26)), 且这种信号与空间下混合的差可得到编码。

[0182] 这样, 在一些实施例中, 第二增强数据被分成增强数据的第一和第二部分, 其中第一部分不如第二部分准确地描述空间下混合。典型地, 第二增强数据的第一部分的相应数据率低于第二部分的数据率。第二增强数据的第二部分的增强数据可仅仅与下混合的一部分相联系, 并特别地可以仅仅与低频部分相联系。

[0183] 在一些实施例中,发生器 123 可设置成为第二增强数据的第一部分和第二部分在绝对和相对数据之间进行选择,无论个别地还是一起。在其它实施例中,发生器 123 可以仅仅为数据的某一部分在绝对和相对数据之间进行选择。特别地,此后的实施例将被描述为其中第二增强数据的第一部分包含 T 的参数,而第二部分包含空间下混合的一种低频表示,并且绝对和相对数据之间的动态选择仅仅应用于第二增强数据的第二部分。

[0184] 举例来说,在这些实施例中,在第一部分的增强数据得到应用之后(即,作为相对于修正艺术下混合的差分值),用于第二增强数据的第二部分的相对数据能产生为相对于艺术下混合的差分值。

[0185] 下面,实施例如下文描述,其中发生器 123 仅仅在第二增强数据的第二部分的相对和绝对数据之间进行选择。

[0186] 本例中,用于第二增强数据第一和第二部分的一部分的绝对增强数据可为相关的时间/频率片导出,通过设定:

$$\begin{aligned} [0187] \quad \underline{L}_{\text{enh}} &= \underline{L}_s, \\ [0188] \quad \underline{R}_{\text{enh}} &= \underline{R}_s, \end{aligned} \quad (25)$$

$$[0189] \quad T' = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

[0190] 其中,  $\underline{L}_s$ 、 $\underline{R}_s$  分别包含空间立体声下混合的左右声道的时间/频率片的采样。因此,在这个特定的例子中,绝对增强数据简单地对应于空间下混合 102 的实际的时间/频率片采样,其可替代相应的艺术下混合 103 的时间/频率片采样。

[0191] 进一步地,对第二增强数据的第一和第二部分的一部分,相关的时间/频率片的相对增强数据可作为差分数据导出,通过设定:

$$\begin{aligned} [0192] \quad \underline{L}_{\text{enh}} &= \underline{L}_s - T_{11} \underline{L}_a - T_{21} \underline{R}_a, \\ [0193] \quad \underline{R}_{\text{enh}} &= \underline{R}_s - T_{12} \underline{L}_a - T_{22} \underline{R}_a, \end{aligned} \quad (26)$$

$$[0194] \quad T' = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

[0195] 这里,参数  $T_{11}$ 、 $T_{12}$ 、 $T_{21}$  和  $T_{22}$  构成式 (2) 中的矩阵 T:

$$[0196] \quad T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}. \quad (27)$$

[0197] 通过这种方式,发生器 123 能为艺术下混合 103 产生绝对增强数据和相对增强数据,允许解码器产生修正艺术下混合,其更加接近地类似于用来产生多声道增强数据的空间下混合 102。

[0198] 进一步地,发生器 123 设置为在绝对增强数据和相对增强数据之间进行选择。在这个特定的例子中,这一选择对于个别的信号块(例如,个别的片段)并且基于这些信号块内的信号的特征执行。特别地,发生器 123 能评价所给信号块的绝对增强数据和相对增强数据的特征,并能确定哪一个数据将要包含在给定信号块的增强层中。另外,发生器 123 可

以包含哪一个数据被选择的指示,从而允许解码器正确地应用接收到的增强数据。

[0199] 在一些实施例中,发生器 123 能评价编码以确定是绝对增强数据还是相对增强数据能被最有效地编码(例如,对于给定精度具有最小数量的比特)。强制的方法(brute force approach)可以对所有的两种类型的增强数据进行准确编码,并且比较所编码数据的大小。但是,在一些实施例中,这会是一种复杂的方法,且在例示的编码器 10 中,发生器 123 评价相对于相对增强数据的信号能量的绝对增强数据的信号能量,并且基于对两者的比较选择将包含哪一种类型的数据。

[0200] 特别地,对于音频编码器,就比特速率而言将信号编码为具有最小可能能量常常是有益的。因此,发生器 123 选择具有最低信号能量的类型的增强数据。特别地,当

[0201]  $\| \underline{L}_s - T_{11} \underline{L}_a - T_{21} \underline{R}_a \|^2 + \| \underline{R}_s - T_{12} \underline{L}_a - T_{22} \underline{R}_a \|^2 < \| \underline{L}_s \|^2 + \| \underline{R}_s \|^2$  (28) 时,选择相对增强数据,否则选择绝对增强数据。

[0202] 在不同的增强数据之间进行切换的一个问题是会产生一些明显的人工噪声。在例示的编码器 10 中,发生器 123 还包括在不同的增强数据之间逐渐切换的功能。因此,代替从一个信号块的增强数据的一种类型向下一个信号块的另一种类型的直接切换,切换从一批数据到另一批数据逐渐的进行。

[0203] 这样,在一个时间间隔(其可以具有小于或大于一个信号块的持续时间)期间,发生器 123 产生增强数据作为绝对增强数据和相对增强数据的一个组合。该组合例如可以通过在不同类型的数据之间插值实现,或者可以使用重叠和相加技术。

[0204] 作为一个具体的例子,取代在不同类型的增强数据之间的突然切换:

[0205]  $\underline{L}_{enh} = \underline{L}_s - T_{11} \underline{L}_a - T_{21} \underline{R}_a$ ,  $\underline{R}_{enh} = \underline{R}_s - T_{12} \underline{L}_a - T_{22} \underline{R}_a$  或者  $\underline{L}_{enh} = \underline{L}_s$ ,  $\underline{R}_{enh} = \underline{R}_s$

[0206] 所传送的增强数据可如下产生:

[0207]  $\underline{L}_{enh} = \underline{L}_s - \alpha T_{11} \underline{L}_a - \alpha T_{21} \underline{R}_a$ ,  $\underline{R}_{enh} = \underline{R}_s - \alpha T_{12} \underline{L}_a - \alpha T_{22} \underline{R}_a$ , (29)

[0208] 其中,第 k 个数据帧中  $\alpha$  的值可如下确定:

[0209]

$$\alpha_k = \begin{cases} \max(0, \alpha_{k-1} - \delta), & \text{若当前帧是绝对编码,} \\ \min(1, \alpha_{k-1} + \delta), & \text{若当前帧是差分编码,} \end{cases} \quad (30)$$

[0210] 其中,  $\alpha_k$  表示在第 k 个帧中  $\alpha$  的值,而  $\delta$  是适配速度。在许多情形下,  $\delta = 0.33$  的值能提供可靠的没有人工噪声的编码。式 (29) 给出的信号  $\underline{L}_{enh}$  和  $\underline{R}_{enh}$  可使用参数插值或重叠相加技术获得,并且编码和添加到比特流。另外,关于差分或绝对增强数据的决策被包含在比特流中,从而使解码器导出与编码器中所使用的相同的  $\alpha$  值成为可能。

[0211] 应当理解,尽管本说明书集中在使用具有 M- 声道的每一个个别(声道内)编码的差分和绝对模式上,其它实施例可以使用不同的编码方式。例如,对  $M = 2$ , 当执行立体声信号的(声道间)编码时,下一个步骤可以应用诸如 M/S 编码(中部/边缘编码,因此对和信号和差分信号编码)之类。在许多实施例中,这将是一个优势,在个别的声道的(声道内)编码的差分和绝对模式中都是如此。

[0212] 变换矩阵  $T'$  的元素可以是实数值或复数值。这些元素可以按如下方式编码到修正参数中:变换矩阵  $T$  的为正实数的那些元素可以被对数量化,象 MPEG4 参数化立体声中的 IID 参数那样。可以为参数值设置一个上限,以避免小信号过度放大。这一上限可以是固定的或者是自动产生的左声道与艺术左声道之间的相关和自动产生的右声道与艺术右声道

之间的相关的函数。在变换矩阵  $T'$  的为复数值的元素中,其幅值可使用 IID 参数量化,而相位可线性量化。实数值的且可能是负值的  $T'$  的元素可以通过取该元素的绝对值的对数进行编码,同时保证负值与正值之间的区别。

[0213] 图 6 更加详细地图解了图 5 中的发生器 123 的一个例子。在该例中,发生器 123 包括信号块处理器 145,其接收频域的空间和艺术下混合 102、126,并将这些信号划分为信号块。每个信号块都对应于预定持续时间的一个时间间隔。在一些实施例中,信号块可以可替换地或者另外地在频域中得到划分,例如变换子声道可以被一起分组到不同信号块中。

[0214] 信号块处理器 145 与绝对增强数据处理器 146 相耦接,如前所述,绝对增强数据处理器 146 为个别的信号块产生绝对增强数据。另外,信号块处理器 145 与相对增强数据处理器 147 相耦接,如前所述,相对增强数据处理器 147 为个别的信号块产生相对增强数据。相对和绝对增强数据基于信号块内的信号特征而确定,并且特别地,用于给定的时间 / 频率片组的增强数据可仅基于该时间 / 频率片组而确定。

[0215] 绝对增强数据处理器 146 与第一信号能量处理器 148 相耦接,如前所述,第一信号能量处理器 148 确定每一个信号块的绝对增强数据的信号能量。类似地,相对增强数据处理器 147 与第二信号能量处理器 149 相耦接,如前所述,第二信号能量处理器 149 确定每一个信号块的相对增强数据的信号能量。

[0216] 第一和第二信号能量处理器 148、149 与选择处理器 150 相耦接,选择处理器 150 根据哪种类型具有最低的信号能量为每一个信号块选择绝对或者相对增强数据。

[0217] 选择处理器 150 被馈送给增强数据处理器 151,增强数据处理器 151 进一步地与增强数据处理器 146 和相对增强数据处理器 147 相耦接。选择处理器 151 接收指示哪一种类型的增强数据被选择的控制信号,并相应地产生与所选择的增强数据一样的增强数据。进一步地,选择处理器 151 设置为执行一种逐渐切换,包括切换时间间隔期间在绝对和相对参数之间插值。

[0218] 选择处理器 151 与编码处理器 152 相耦接,该编码处理器依照给定的协议对增强数据进行编码。另外,编码处理器 152 通过例如为每一个信号块设置一个比特以指示数据类型,对指示每一个信号块中哪一种类型的数据被选择的数据进行编码。来自编码处理器 152 的已编码数据被包含在由编码器 10 产生的已编码比特流中。

[0219] 图 7 显示了根据本发明一些实施例的多声道音频解码器的另一个实施例的方框图,特别地,其可以是图 2 中的音频解码器 20。

[0220] 解码器 20 包括第一单元 210 和与之相耦接的第二单元 220。第一单元 210 接收下混合信号  $l_0$  与  $r_0$  和修正参数 105 作为输入。例如,该输入可以作为来自图 1 或图 5 中的编码器 10 的单个比特流接收。下混合信号  $l_0$  与  $r_0$  可以是空间下混合 102 或者艺术下混合 103 的一部分。

[0221] 第一单元 210 包括分段和变换单元 211 和下混合修正单元 212。在分段和变换单元 211 中,下混合信号  $l_0$  与  $r_0$  分别被分段,且这些分段信号被变换到频域。所得到的分段下混合信号的频域表示分别显示为频域信号  $L_0$  与  $R_0$ 。接下来,频域信号  $L_0$  与  $R_0$  在下混合修正单元 212 中得到处理。下混合修正单元 212 的功能是修正输入的下混合,以使得其类似于空间下混合 202,即,从艺术下混合 103 和修正参数 105 中重建空间下混合 202。

[0222] 若空间下混合 102 被解码器 20 接收到,则下混合修正单元 212 不必修正下混合信

号  $L_o$  与  $R_o$ , 并且这些下混合信号  $L_o$  与  $R_o$  简单地传递到第二单元 220, 作为空间下混合 202 的下混合信号  $L_d$  和  $R_d$ 。控制信号 217 可指示是否需要修正输入的下混合, 即, 输入的下混合是空间下混合还是可替换下混合。控制信号 217 可在解码器 20 内部产生, 例如, 通过分析输入的下混合和相关参数 105, 该相关参数 105 可以描述期望的空间下混合的信号特性。若输入的下混合与期望的信号特性相匹配, 控制信号 217 则可设为指示不需要修正。可替换地, 控制信号 217 可以手动设置, 或者其设置也可以作为已编码多声道音频信号的一部分接收, 例如在参数集 105 中。

[0223] 如果编码器 20 接收艺术下混合 103, 并且控制信号 217 指示所接收的下混合信号  $L_o$  与  $R_o$  需通过下混合修正单元 212 进行修正, 则解码器可以两种方式操作, 这取决于所接收的修正参数的表示。若该参数表示从艺术下混合到空间下混合的相对变换 (即, 若该参数是相对增强数据), 那么变换变量可通过将修正参数应用于艺术下混合直接获得, 即编码器中所执行操作的逆向操作。在不同的实施例中, 这一方式可能仅应用于第二增强数据的第二部分。

[0224] 另一方面, 若该传送的参数表示空间下混合的绝对特性, 那么解码器可用空间下混合采样直接代替艺术下混合采样。例如, 若第二增强数据的第二部分简单地存在于空间下混合的时间 / 频率片采样内, 则解码器可用这些直接代替相应的艺术下混合的时间 / 频率片采样。将会理解的是, 解码器也可能首先计算实际传送的艺术下混合的相应特性。然后, 用此信息 (所传送的参数和所传送艺术下混合的计算的特性), 描述从所传送艺术下混合 (的特性) 到空间下混合 (的特性) 变换的变换变量被确定。更具体地, 变换矩阵  $T$  可利用如前所述的方法 II. a 或者 (稍加修正的) II. b 得到确定。

[0225] 如果绝对能量在第二增强数据的第一部分中进行传送, 则可以使用方法 II. a。所传送的 (绝对) 参数  $E_{L_s}$  与  $E_{R_s}$  分别表示空间下混合的左右信号的能量, 且由下式给出:

$$[0226] \quad E_{L_o} = \sum_k \|L_s[k]\|^2, \quad E_{R_o} = \sum_k \|R_s[k]\|^2. \quad (31)$$

[0227] 所传送的下混合的能量  $E_{DL_s}$  与  $E_{DR_s}$  在解码器进行计算。利用这些变量, 我们可以求得 (7) 中的参数  $\alpha$  与  $\beta$  如下

$$[0228] \quad \alpha = \sqrt{\frac{E_{L_s}}{E_{DL_s}}}, \beta = \sqrt{\frac{E_{R_s}}{E_{DR_s}}}. \quad (32)$$

[0229] 变换矩阵由下式给出:

$$[0230] \quad T = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}. \quad (33)$$

[0231] 特别地, 下混合修正单元 212 包括这样的功能: 从所接收的比特流中提取艺术下混合和修正参数 105。艺术下混合被分成信号块 (对应于解码器所用的信号块)。对每一个信号块, 下混合修正单元 212 评价所接收的比特流的数据指示以确定式相对还是绝对第二增强数据被提供用于该信号块的第一和第二部分。然后, 下混合修正单元 212 应用第二增强数据的第一和第二部分作为绝对增强数据或相对增强数据以响应该指示数据。

[0232] 已经发现, 当变换矩阵元素  $T_{12}$  和  $T_{21}$  设为零时, 可以得到低的复杂度和高的性能。下面, 在这一约束下, 对下混合修正单元 212 的一些特别实现进行说明。但是, 应当理解, 这些实现可以很容易地扩展到  $T_{12}$  和 / 或  $T_{21}$  不为零的情况。

[0233] 在没有第二增强数据的第二部分的增强数据传送用于艺术下混合信号的情况下，第一单元 210 可如图 8 所示实现。时域立体声下混合声道  $l_0$  与  $r_0$  被首先分段并通过 QMF 变换变换到频域，得到表示艺术立体声下混合的时间 / 频率片的信号  $L_a$  与  $R_a$ 。接着，这些信号使用变换矩阵  $T$  进行变换，得到信号  $T_{11}L_a$  与  $T_{22}R_a$ 。

[0234] 应当理解，增强数据可以在时域和 / 或频域中产生和应用。因此，将已编码时域增强数据 ( $L_{enh}, R_{enh}$ ) 包含在比特流中是可能的。但是，在一些应用中，包含已编码频域增强数据而不是时域增强数据是有利的。例如，在许多编码器中，增强数据在频域中为时间 / 频率片产生，并且为了产生时域信号，编码器需要一种频域到时域的变换。进一步地，为了应用这种增强数据，解码器将该数据从时域转变到频域。那么，通过在频域中包括增强数据，可以简化域转换。

[0235] 在一些实施例中，各种不同的时频转换可用于产生艺术下混合及增强数据。例如，艺术下混合的编码可使用 QMF 变换，而增强数据使用 MDCT 变换。在这种情况下，增强数据将包含在 (MDCT) 频域中，且在两种频域间的直接变换可如图 9 所示由下混合修正单元 212 执行。

[0236] 在本例中，变换矩阵  $T^*$  可以简单地是式 (2) 中的变换矩阵  $T$ 。但是，为了减少切换人工噪声， $T^*$  可以对应于式 (2) 中的变换矩阵  $T$  但为逐渐切换加以修正。特别地，矩阵  $T^*$  可以包含式 (30) 所确定的因子  $\alpha$ ，其中关于绝对还是相对增强数据的判断从比特流中取出。这一方案应用于那些存在第二增强数据的第二部分的增强层数据的信号块 / 频段；否则，可以使用图 8 的方法。

[0237] 如果增强数据 ( $L_{enh}, R_{enh}$ ) 在时域中提供，可以使用类似图 9 的方法，如图 10 所示。但是，在这种情况下，频域到频域的变换被时域到频域的变换所代替，特别地，当 QMF 变换被用于艺术下混合的编码时，变换可以替换为时域到 QMF 域的变换。因此，在这个例子中，增强数据在频域中使用。

[0238] 在许多实施例中，可以使用一种时域增强数据的解码器实现，它在第一单元 210 中仅使用一次时域到频域变换。

[0239] 特别地，可以使用下述的差分增强数据参数：

$$[0240] \quad \underline{L}_{enh} = \frac{T_{22}\underline{L}_s - T_{21}\underline{R}_s}{\det(T)} - \underline{L}_a,$$

$$[0241] \quad \underline{R}_{enh} = \frac{-T_{12}\underline{L}_s + T_{11}\underline{R}_s}{\det(T)} - \underline{R}_a, \quad (34)$$

$$[0242] \quad T' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

[0243] 如果由式 (27) 给出的矩阵  $T$  是非奇异的 (因此其逆存在) 的话。现在式 (1) 可以改写为：

$$[0244] \quad [\underline{L}_d \ \underline{R}_d] = [\underline{L}_a \ \underline{R}_a \ \underline{L}_{enh} \ \underline{R}_{enh}]T' T. \quad (35)$$

[0245] 图 11 说明了基于式 (34) 和 (35) 的一种下混合修正单元 212 用于时域增强数据的高效实现。为清晰起见，矩阵  $T$  中  $T_{12}$  与  $T_{21}$  都设为零。与图 10 的实现相比，图 11 的实现



只需一个时域到 QMF 域的变换。

[0246] 这样,如上所述,下混合修正单元 212 产生信号 202,其非常接近地类似用于多声道增强数据的空间下混合。这可被第二单元 220 有效地用于将两声道音频信号扩展为完全的环境声多声道信号。更进一步,通过为每一个信号块动态地和灵活地选择最适当的增强数据类型(相对的或绝对的),可以实现一种更加有效的编码,并且可以实现一种在质量-数据率之比方面有改进的多声道编码/解码。

[0247] 第二单元 220 可以是传统的 2 到 5.1 多声道解码器,其将所重建的空间下混合 202 和相关的参数数据 104 解码为 5.1 声道输出信号 203。如前所述,参数数据 104 包含参数化数据 141、142、143 和 144。第二单元 220 执行编码器 10 中第一单元 110 的逆处理。第二单元 220 包含向上混合器 221,其将立体声下混合 202 和相关参数 144 转换为三个单声音频信号 L、R 和 C。然后,单声音频信号 L、R 和 C 中的每一个分别在去相关器 222、225 和 228 中进行去相关。此后,混合矩阵 223 将单声音频信号 L、其去相关对应部分和相关参数 141 变换为信号 Lf 和 Lr。类似地,混合矩阵 226 将单声音频信号 R、其去相关对应部分和相关参数 142 变换为信号 Rf 和 Rr,混合矩阵 229 将单声音频信号 C、其去相关对应部分和相关参数 143 变换为信号 Co 和 LFE。最后,三对分段频域信号 Lf 和 Lr、Rf 和 Rr、Co 和 LFE 分别变换到时域,并分别在逆变换器 224、227 和 230 中通过重叠相加来组合,以分别获得三对输出信号 lf 和 lr、rf 和 rr、co 和 lfe。输出信号 lf、lr、rf、rr、co 和 lfe 形成已解码多声道音频信号 203。

[0248] 多声道音频编码器 10 和多声道音频解码器 20 可以通过数字化硬件或通过软件实现,该软件可由数字信号处理器或通用微处理器执行。

[0249] 应当理解,为了清楚,上面的说明参照多种功能单元和处理器对本发明的实施例进行描述。然而,显而易见,功能在不同功能单元或处理器间的任何适当分布都可以在不减损本发明的情况下得到应用。例如,被描述为由分离的处理器或控制器执行的功能可以由同一个处理器或控制器执行。因此,对特定功能单元的参照只能看成是参考提供所描述功能的适当装置,而不能看成是指示严格的逻辑或物理结构或组织。

[0250] 本发明可以以任何合适的形式实现,包括硬件、软件、固件或它们的任何结合。可选地,本发明可以至少部分地作为计算机软件实现,该计算机软件可以在一个或更多数据处理器和/或数字信号处理器上运行。本发明的实施例的元件和组件可以以任何合适的方式物理地、功能地和逻辑地实现。毫无疑问地,功能可以在单个单元中、在多个单元中或者作为其它功能单元的一部分而得到实现。像这样,本发明可以在单个单元中实现或者物理和功能地分布在不同的单元和处理器中而得以实现。

[0251] 尽管本发明结合一些实施例进行说明,却并非企图将本发明限制于在此给出的特定形式。相反,本发明的范围仅受所附权利要求限制。另外,尽管一个特征看起来被描述为与特定的实施例相联系,本领域的技术人员将认识到所描述的实施例的各种特征可以根据本发明进行组合。在权利要求书中,术语包括并不排除其它元素或步骤的存在。

[0252] 此外,尽管多个装置、元件或方法步骤被个别列出,它们可以由单个的单元或处理器等实现。另外,尽管个别特征被包含在不同的权利要求中,它们可能会被有利地组合,且其被不同的权利要求所包含并不意味着这些特征的组合不可行和/或没有好处。一个特征在一类权利要求中的包含并不意味着其限制于该类别而是表明该特征在适当时可以同

样适用于其它权利要求类别。此外,特征在权利要求中的顺序并不意味着特征必须以特定的顺序工作,并且特别地,方法权利要求中个别步骤的顺序并不意味着该步骤必须以该顺序执行。相反,这些步骤可以任何合适的顺序执行。另外,单数引用并不排除多个。因此,对“一个”、“一”、“第一”、“第二”等的引用并不排除多个。权利要求中的参照符号仅仅作为一种澄清实例而提供,不能解释为对权利要求的范围的任何方式的限制。

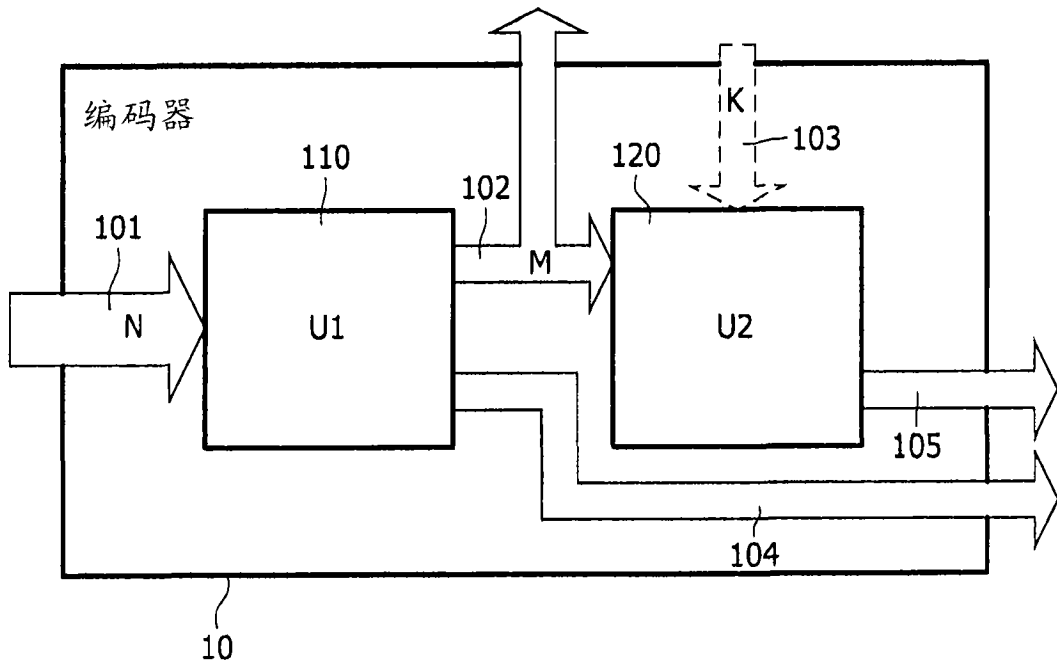


图 1

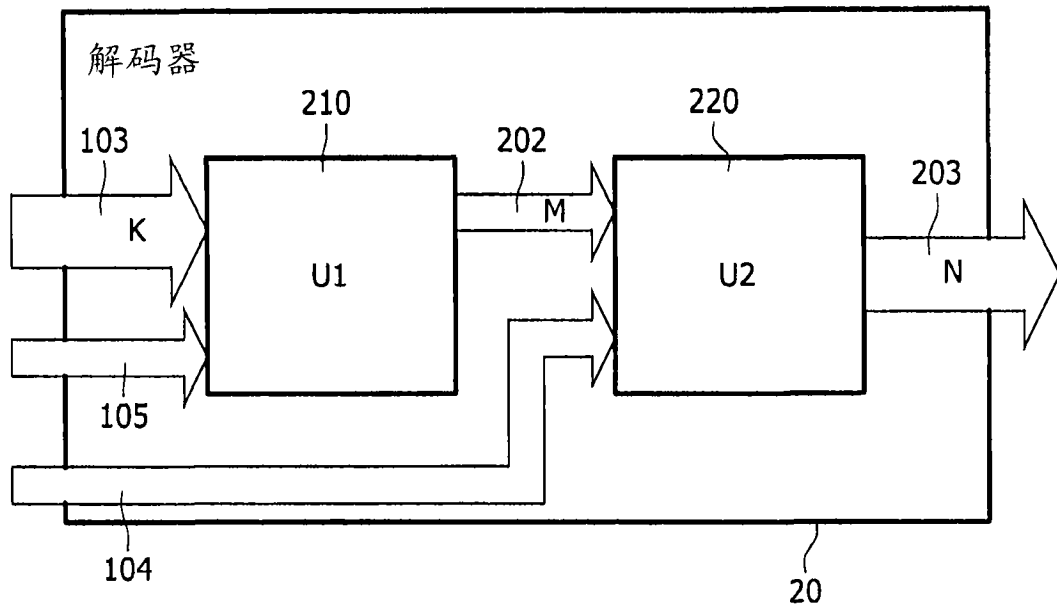


图 2

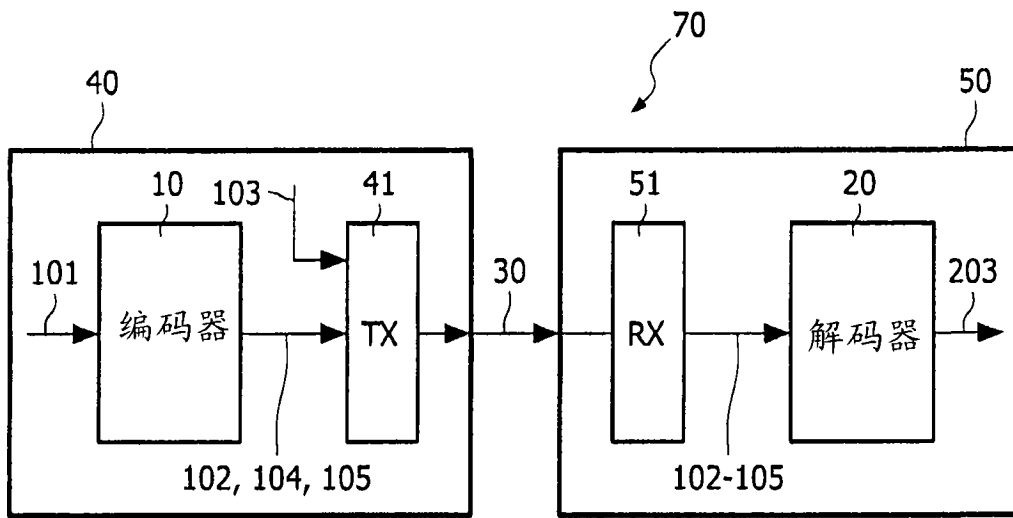


图 3

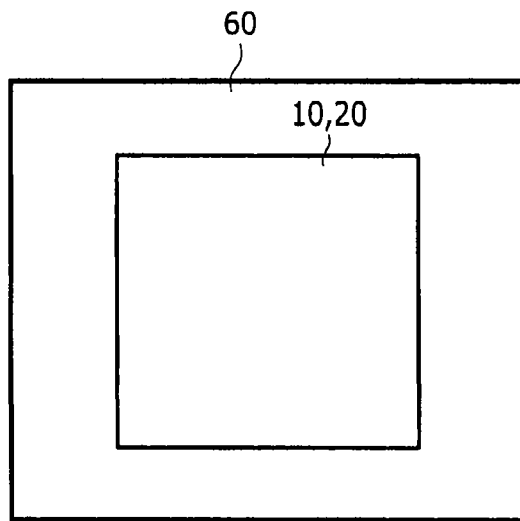


图 4

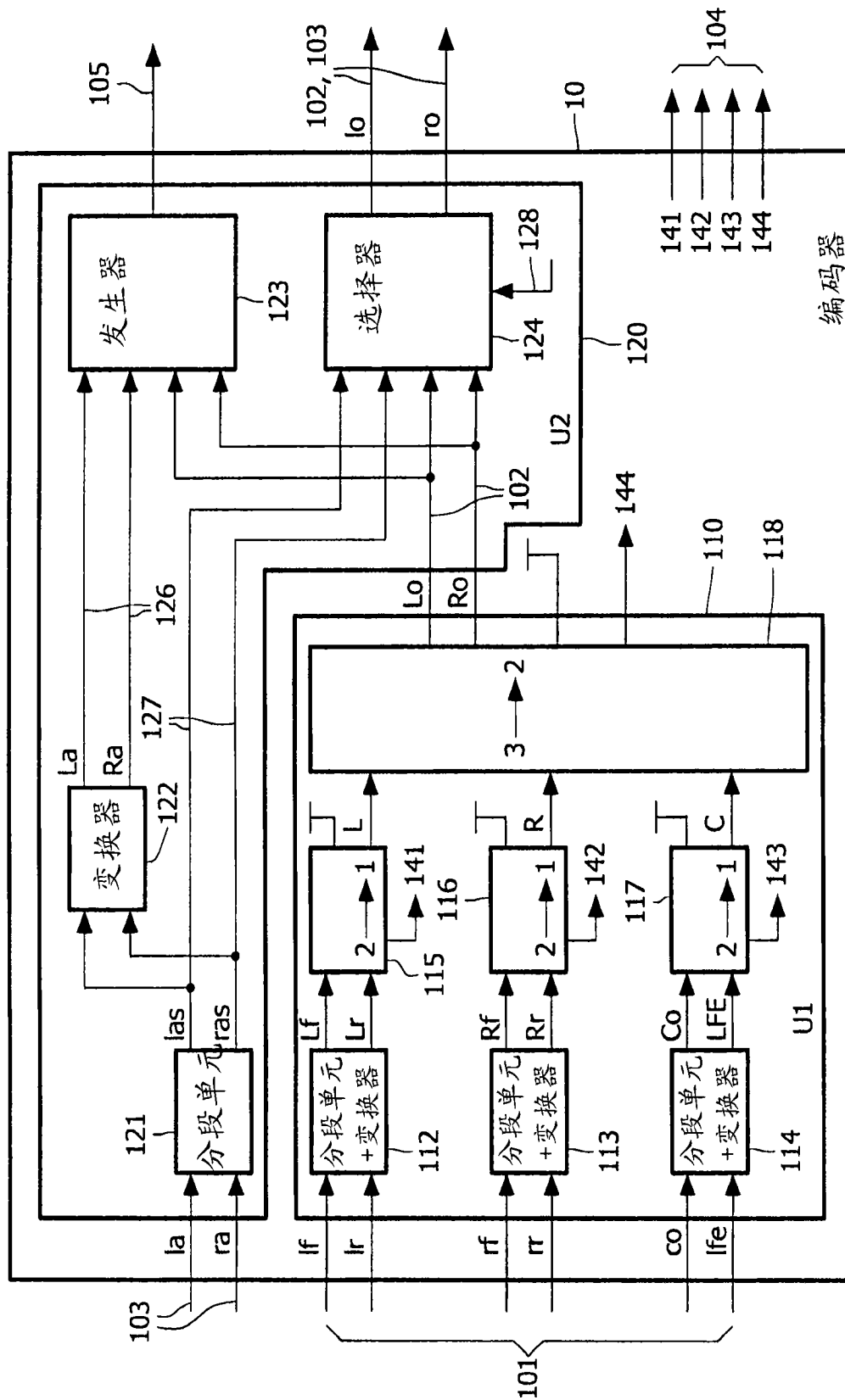


图 5

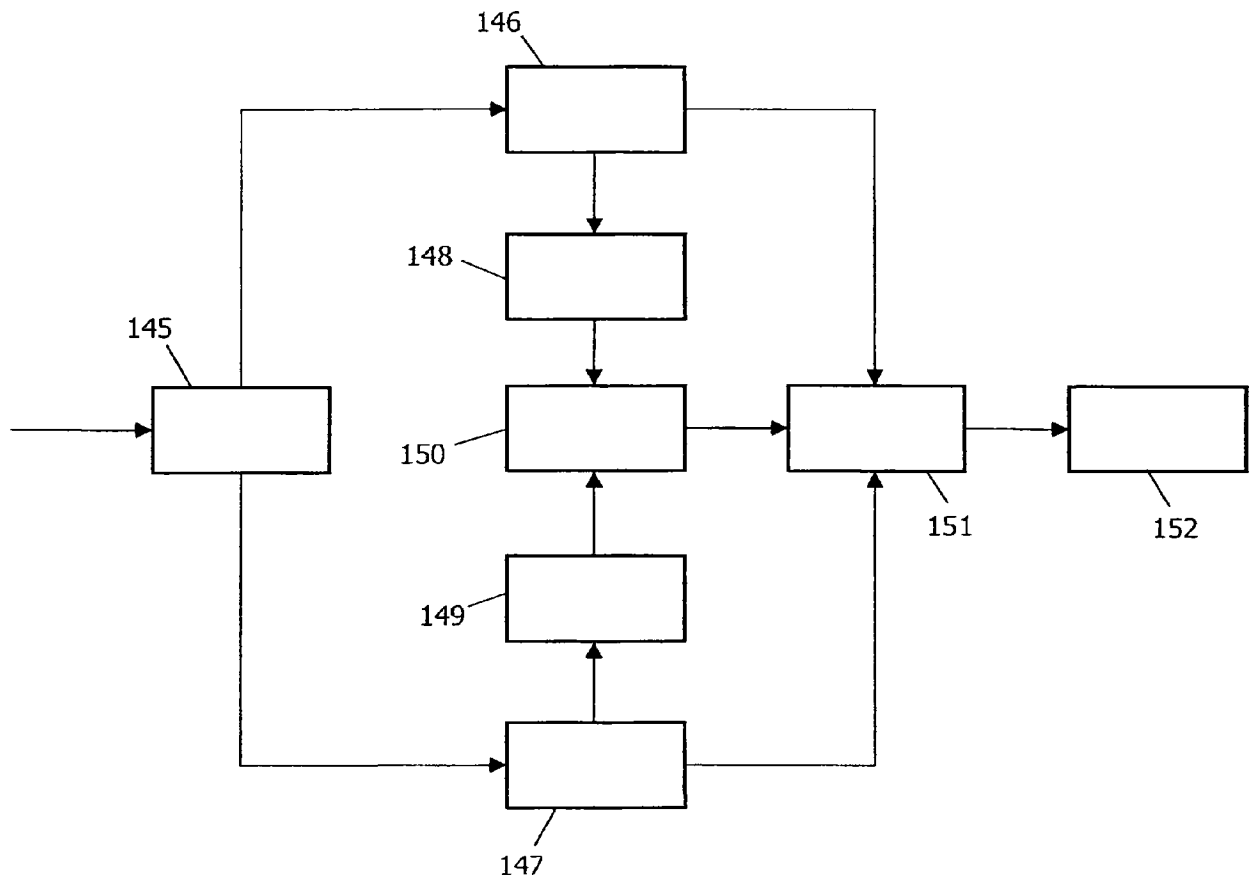


图 6

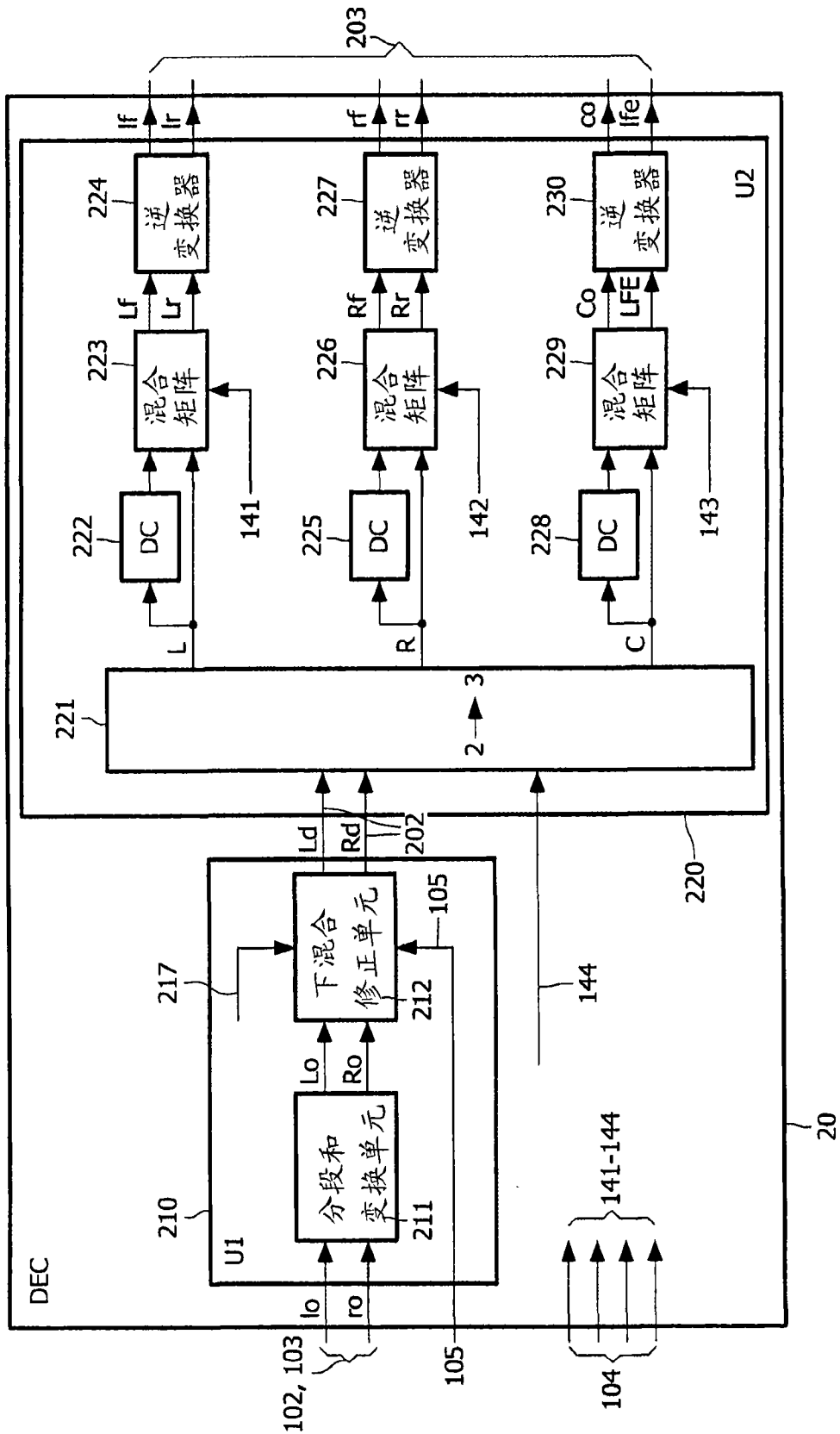


图 7

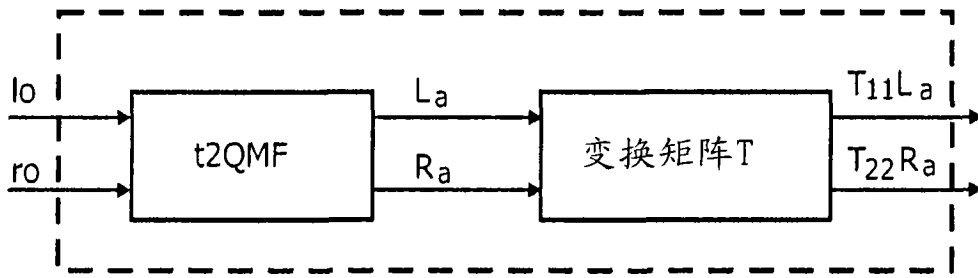


图 8

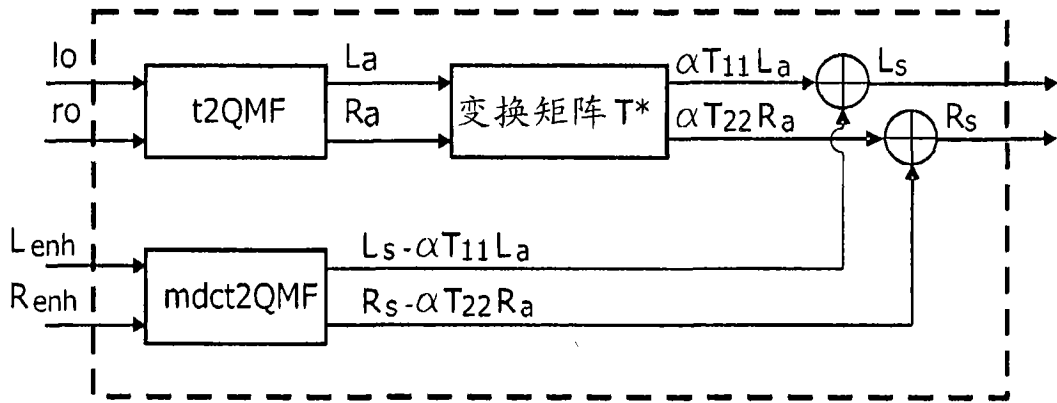


图 9



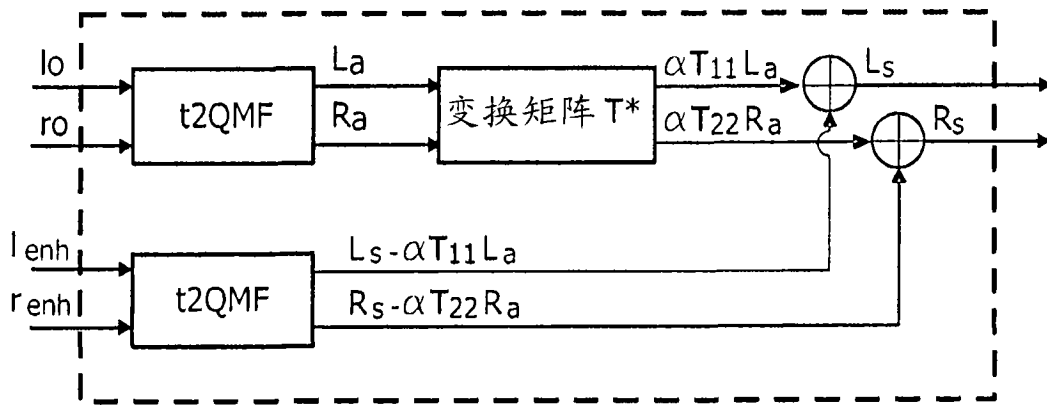


图 10

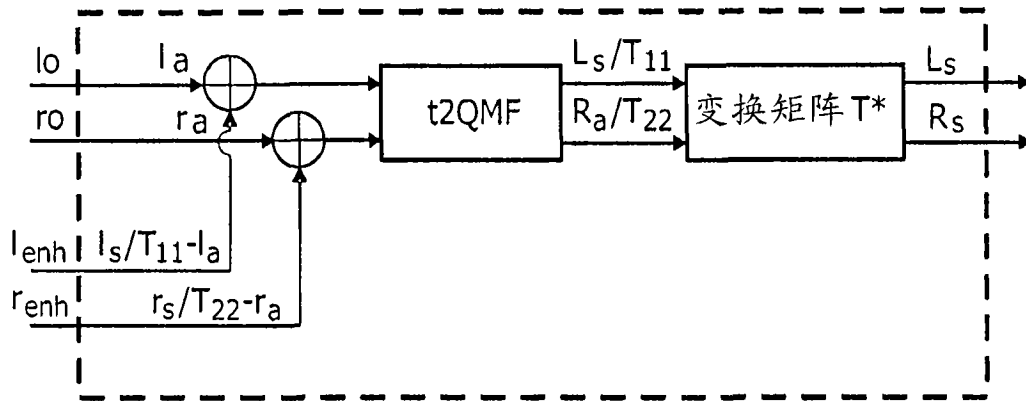


图 11