



(12) 发明专利

(10) 授权公告号 CN 1756086 B

(45) 授权公告日 2010.05.05

(21) 申请号 200510084143.6

G10L 19/00(2006.01)

(22) 申请日 2005.07.14

(56) 对比文件

(30) 优先权数据

10-2005-0021840 2005.03.16 KR

60/587,626 2004.07.14 US

US 5451942 A, 1995.09.19, 全文.

US 20020165721 A1, 2002.11.07, 全文.

US 6487535 B1, 2002.11.26, 全文.

(73) 专利权人 三星电子株式会社

审查员 詹芊芊

地址 韩国京畿道

(72) 发明人 吴殷美 金美英 金尚煜 金度亨

金重会

(74) 专利代理机构 北京铭硕知识产权代理有限公司

公司 11286

代理人 郭鸿禧 李云霞

(51) Int. Cl.

H03M 7/30(2006.01)

H04S 3/00(2006.01)

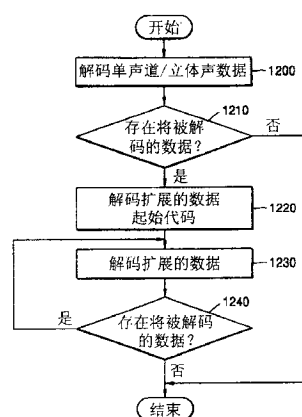
权利要求书 5 页 说明书 9 页 附图 7 页

(54) 发明名称

多通道音频数据编码 / 解码方法和设备

(57) 摘要

提供一种多通道音频数据编码和 / 或解码方法以及设备。该编码方法包括 : 编码单声道和 / 或立体声音频数据 ; 以及编码除了单声道和 / 或立体声音频数据之外的扩展的多通道音频数据。解码方法包括 : 解码单声道和 / 或立体声音频数据 ; 检验除了单声道和 / 或立体声音频数据之外是否存在将被解码的扩展的多通道音频数据 ; 如果存在, 则解码该数据。根据该方法和设备, 仅通过一个比特流, 可根据用户环境来提供单声道、立体声和多通道。另外在多通道音频中, 根据用户终端和网络的状态提供精细可分级 (FGS) 功能。此外, 可实现 BSAC 的性能的增强。可满足对 MPEG 标准化的各种需求。可用于更逼真的数字多媒体广播和基于移动以及基于家庭影院的服务。



1. 一种多通道音频信号编码方法,包括:
编码单声道和 / 或立体声音频数据;
编码除了单声道和 / 或立体声音频数据之外的扩展的多通道音频数据。
2. 如权利要求 1 所述的方法,其中,所述单声道和 / 或立体声音频数据具有分层的比特率。
3. 如权利要求 1 或 2 所述的方法,其中,所述扩展的多通道音频数据包括扩展的通道的类型信息,所述的类型信息至少指示音频通道的配置,并被表达为通道配置索引。
4. 如权利要求 1 或 2 所述的方法,其中,所述扩展的多通道音频数据的编码包括:
编码指示扩展的多通道音频数据的起始的预定的起始代码;
按通道编码扩展的音频数据。
5. 如权利要求 4 所述的方法,其中,所述起始代码包括:
由 32 比特的连续的 0 形成的 zero_code;
由 8 比特的连续的 1 形成的同步字。
6. 如权利要求 4 所述的方法,其中,所述按通道编码扩展的数据的步骤包括:
编码指示音频通道的配置的扩展的通道的类型;
编码扩展的通道音频数据。
7. 如权利要求 6 所述的方法,其中,所述扩展的通道的类型由通道配置索引形成。
8. 如权利要求 6 所述的方法,其中,所述按通道编码扩展的数据的步骤包括:
编码扩展的数据的长度;
编码旁信息。
9. 如权利要求 8 所述的方法,其中,所述的旁信息包括 bsac_header 和 general_header。
10. 如权利要求 6 所述的方法,其中,所述扩展的通道音频数据的编码包括:
编码具有最低比特率的基本层;
编码比特率高于基本层的比特率的增强层,如果有多个增强层,则随增强层数的增加而增加比特率。
11. 一种多通道音频信号编码设备,包括:
单声道 / 立体声编码单元,用于编码单声道和 / 或立体声音频数据;
扩展的数据编码单元,用于编码除了单声道和 / 或立体声音频数据之外的扩展的多通道音频数据。
12. 如权利要求 11 所述的设备,其中,所述单声道 / 立体声编码单元编码具有分层的比特率的单声道和 / 或立体声音频数据。
13. 如权利要求 11 或 12 所述的设备,其中,扩展的数据编码单元的扩展的多通道音频数据包括扩展的通道的类型信息,所述的类型信息至少指示音频通道的配置,并被表达为通道配置索引。
14. 如权利要求 11 或 12 所述的设备,其中,所述扩展的数据编码单元包括:
起始代码编码单元,用于编码指示扩展的多通道音频数据的起始的预定起始代码;
通道编码单元,用于按通道编码扩展的音频数据。
15. 如权利要求 14 所述的设备,其中,所述起始代码编码单元的起始代码包括:

- 由 32 比特的连续的 0 形成的 zero_code ;
由 8 比特的连续的 1 形成的同步字。
16. 如权利要求 14 所述的设备,其中,所述通道编码单元包括:
扩展的通道类型编码单元,用于编码指示音频通道的配置的扩展的通道类型;
扩展的音频编码单元,用于编码扩展的通道音频数据。
17. 如权利要求 16 所述的设备,其中,所述扩展的通道类型由通道配置索引形成。
18. 如权利要求 16 所述的设备,其中,所述通道编码单元包括:
扩展的数据长度编码单元,用于编码扩展的数据的长度;
旁信息编码单元,用于编码旁信息。
19. 如权利要求 18 所述的设备,其中,所述的旁信息包括 bsac_header 和 general_header。
20. 如权利要求 16 所述的设备,其中,所述扩展的音频编码单元包括:
基本层编码单元,用于编码具有最低比特率的基本层;
增强层编码单元,用于编码比特率高于基本层的比特率的增强层,如果存在多个增强层,则随增强层的数量而增加比特率。
21. 一种多通道音频信号解码方法,包括:
解码单声道和 / 或立体声音频数据;
检验除了单声道和 / 立体声音频数据之外是否存在将被解码的扩展的多通道音频数据;
如果存在将被解码的扩展的数据,则解码扩展的多通道音频数据。
22. 如权利要求 21 所述的方法,其中,单声道和 / 或立体声音频数据具有分层的比特率。
23. 如权利要求 21 或 22 所述的方法,其中,所述扩展的多通道音频数据包括扩展的通道类型信息,所述的类型信息至少指示音频通道的配置,并被表达为通道配置索引。
24. 如权利要求 21 或 22 所述的方法,其中,在是否存在扩展的多通道音频数据的检验中,检验指示扩展的多通道音频数据的起始的预定的起始代码的存在,如果存在起始代码,则确定存在扩展的数据。
25. 如权利要求 24 所述的方法,其中,所述的起始代码包括:
由 32 比特的连续的 0 形成的 zero_code ;
由 8 比特的连续的 1 形成的同步字。
26. 如权利要求 21 或 22 所述的方法,其中,在扩展的多通道音频数据的解码中,如果存在将被解码的扩展的数据,则按通道解码扩展的数据。
27. 如权利要求 26 所述的方法,其中,按通道解码扩展的数据的步骤包括:
解码指示音频通道的配置的扩展的通道类型;
解码扩展的通道音频数据。
28. 如权利要求 27 所述的方法,其中,所述的扩展的通道类型由通道配置索引形成。
29. 如权利要求 26 所述的方法,其中,按通道解码扩展的数据的步骤包括:
解码扩展的数据的长度;
解码旁信息。

30. 如权利要求 29 所述的方法,其中,所述的旁信息包括 bsac_header 和 general_header。

31. 如权利要求 27 所述的方法,其中,扩展的通道音频数据的解码包括:

解码具有最低比特率的基本层;

解码其比特率高于基本层的比特率的增强层,如果存在多个增强层,则随增强层数增加比特率。

32. 一种多通道音频信号解码设备,包括:

单声道 / 立体声解码单元,用于解码单声道和 / 或立体声音频数据;

扩展的数据检验单元,检验除了单声道和 / 或立体声音频数据之外是否存在将被解码的扩展的多通道音频数据;

扩展的数据解码单元,如果存在将被解码的扩展的数据,则解码扩展的多通道音频数据。

33. 如权利要求 32 所述的设备,其中,单声道和 / 立体声音频数据具有分层的比特率。

34. 如权利要求 32 或 33 所述的设备,其中,所述扩展的数据检验单元检验预定起始代码的存在,所述的预定起始代码指示扩展的多通道音频数据的起始,如果存在起始代码,则确定存在扩展的数据。

35. 如权利要求 34 所述的设备,其中,所述的起始代码包括:

由 32 比特的连续的 0 形成的 zero_code;

由 8 比特的连续的 1 形成的同步字。

36. 如权利要求 32 或 33 所述的设备,其中,如果存在将被解码的扩展的数据,则扩展的数据解码单元按通道解码扩展的数据。

37. 如权利要求 36 所述的设备,其中,所述的扩展的数据解码单元包括:

扩展的通道类型解码单元,用于解码指示音频通道的配置的扩展的通道类型;

扩展的通道音频解码单元,用于解码扩展的通道音频数据。

38. 如权利要求 37 所述的设备,其中,所述扩展的通道类型由通道配置索引形成。

39. 如权利要求 37 所述的设备,其中,扩展的数据解码单元包括:

扩展的数据长度解码单元,用于解码扩展的数据的长度;

旁信息解码单元,用于解码旁信息。

40. 如权利要求 39 所述的设备,其中,所述的旁信息包括 bsac_header 和 general_header。

41. 如权利要求 37 所述的设备,其中,所述扩展的通道音频解码单元包括:

基本层解码单元,用于解码具有最低比特率的基本层;

增强层解码单元,用于解码其比特率高于基本层的比特率的增强层,如果存在多个增强层,则随增强层的数量而增加比特率。

42. 一种多通道音频信号编码方法,包括:

编码单声道 / 立体声音频数据的基本层;

编码单声道 / 立体声音频数据的增强层;

编码指示扩展的多通道音频数据的起始的预定起始代码(zero_code,同步字);

对构成扩展的多通道音频数据的至少一个通道数据编码基本层,并对所述的至少一个

通道数据编码增强层。

43. 如权利要求 42 所述的方法,其中,所述的对至少一个通道数据编码基本层的步骤包括:

- 编码通道数据的长度;
- 编码指示通道的类型的通道配置索引(channel_configuration_index);
- 编码旁信息;
- 编码基本层的音频数据。

44. 如权利要求 43 所述的方法,其中,所述的旁信息包括 bsac_header 和 general_header。

45. 一种多通道音频信号解码方法,包括:

- 解码单声道/立体声音频数据的基本层;
- 解码单声道/立体声音频数据的增强层;
- 检验除了单声道/立体声音频数据之外,是否存在将被解码的扩展的多通道音频数据;

如果存在将被解码的扩展的多通道音频数据,则解码预定的起始代码(zero_code,同步字),所述的起始代码指示扩展的多通道音频数据的起始;

对构成扩展的多通道音频数据的至少一个通道数据解码基本层,并对所述的至少一个通道数据解码增强层。

46. 如权利要求 45 所述的方法,其中,所述的对至少一个通道数据解码基本层的步骤包括:

- 解码通道数据的长度;
- 解码指示通道的类型的通道配置索引(channel_configuration_index);
- 解码旁信息(bsac header, general header);
- 解码基本层的音频数据。

47. 一种多通道音频信号编码方法,包括:

- 编码音频数据;
- 生成指示编码音频数据的净荷已经完成的标识代码;
- 生成指示扩展的数据的净荷已经开始的标识代码;
- 生成扩展的数据的类型,所述扩展的数据的类型将音频数据的通道扩展为多通道;
- 编码扩展的数据。

48. 一种多通道音频信号解码方法,包括:

- 解码音频数据;
- 检测指示音频数据的净荷已经完成的标识代码;
- 检测指示扩展的数据的净荷已经开始的标识代码;
- 检测扩展的数据的类型;
- 确定检测到的扩展的数据的类型是否将音频数据的通道扩展为多通道;
- 当确定检测到的类型将音频数据的通道扩展为多通道时,解码扩展的数据。

49. 一种多通道音频信号解码方法,包括:解码音频数据;

确定是否存在还未被解码的数据;

当确定存在还未被解码的数据时,检测指示音频数据的净荷已经被完成的标识代码;
检测指示扩展的数据的净荷已经开始的标识代码;
检测扩展的数据的类型;
确定检测到的扩展的数据的类型是否将音频数据的通道扩展为多通道;
当确定检测到的类型将音频数据的通道扩展为多通道时,解码扩展的数据。

50. 一种多通道音频信号解码方法,包括:

解码分级编码的音频数据;
检测零代码;
检测同步字;
检测扩展的通道的类型;

确定检测到的扩展的通道的类型是否指示扩展的数据,所述扩展的数据将音频数据的通道扩展为多通道;

当确定检测到的扩展的通道的类型指示扩展的数据时,解码扩展的数据。

51. 一种多通道音频信号解码方法,包括:

解码分级编码的音频数据;
确定是否存在还未被解码的数据;
当确定存在还未被解码的数据时,检测零代码;
检测同步字;
检测扩展的通道的类型;

确定检测到的扩展的通道的类型是否指示扩展的数据,所述扩展的数据将音频数据的通道扩展为多通道;

当确定检测到的扩展的通道的类型指示扩展的数据时,解码所述扩展的数据。

多通道音频数据编码 / 解码方法和设备

技术领域

[0001] 本发明涉及音频编码和解码,更具体地讲,涉及多通道音频数据编码和解码方法以及设备。

背景技术

[0002] 从 2003 年起,地面数字多媒体广播 (DMB) 已经使用音频编解码器 MPEG-4 比特分片算术编码 (BSAC)。尽管目前只能提供立体声服务,但是可以预期在将来将包括多通道服务。MPEG-4BASC 应该能够增加压缩效率和改善技术的功能,例如带宽扩展的和空间音频。

[0003] 在传统 BSAC 多通道中,在一个层中交替地编码中央、前左、前右、后左和后右通道。图 1 示出传统 BASC 多通道的结构。BSAC 结构提供精细可分级 (FGS) 功能。即,所有五个通道在一个层中并且可从最后层中剪切数据。关于通道的工具旁信息 (side information) 应该在 general_header (通用头) 中限定。考虑到每个通道中的特性,高性能压缩需要单独的旁信息。

[0004] 图 2 是使用传统的 BSAC 方法的音频编码设备的功能模块的框图。该设备包括心理声学模型单元 200、时间 / 频率映射单元 210、时域噪声修整 (TNS) 单元 220、强度立体声处理单元 230、知觉噪声替换 (PNS) 单元 240、中侧 (M/S) 立体声处理单元 250、量化单元 260 和比特打包单元 270。

[0005] 由于相对于时间人可以觉察的信号之间的差别不是很大,所以时间 / 频率映射单元 210 将时域中的音频信号转换为频域中的信号。然而,在频域中的信号的情况下,相对于人类声学模型,可被人觉察的信号与不能被人觉察的信号之间的差别在每个带宽中都很大。因此,通过改变相对于每个频率带宽的比特数,可增强压缩的效率。

[0006] 心理声学模型单元 200 将被时间 / 频率映射单元从时域转换到频域中的音频信号结合到合适的子带的信号中,并通过使用由每个信号的交互作用而发生的掩蔽现象,计算每个子带的掩蔽阈值。TNS 单元 220 被用于控制每个转换窗口中的量化噪声的时域形状。通过应用频率数据的滤波处理实现 TNS。TNS 单元 220 被随意地应用在编码器中。强度立体声处理单元 230 被设计为用于更有效地处理立体声信号。在该装置中,仅关于相对于两个通道之一的比例因子波段的量化信息被编码,并且相对于剩余的通道仅仅比例因子被发送。在编码器中,不是必须使用单元 230。在当前帧中信号具有强噪声特性的情况下,通过编码与比例因子波段相应的频率分量的每个的能量值,而不编码频率系数的值, PNS 单元 240 可以减少产生的将被使用的比特的量。PNS 单元 240 可确定是否以比例因子波段为单位使用比特。M/S 立体声处理单元 250 也是用于更有效地处理立体声信号的装置。在该装置中,左通道的信号和右通道的信号分别被转为被加的信号和被减的信号,然后这些信号被处理。在编码器中,也不是必须使用 M/S 立体声处理单元。量化单元 260 执行每个波段的频率信号的标量量化,从而使得每个波段的量化噪声的大小小于掩蔽阈值,这样人们就感觉不到噪声。比特打包单元 270 收集在编码设备的每个模式中产生的信息项,并根据产生的适合于可分级的 (scalable) 音频编解码器的句法形成比特流。

[0007] 然而,在图 1 中所示的传统的 BSAC 多通道结构中,不能使用中侧 (M/S) 立体声。这是因为在传统的编码和解码句中,当通道的数目是 2 或更多时,不能使用 M/S 立体声功能。因此,编码效率降低。另外,由于窗口切换和 PNS 对于所有的通道应该使用相同的旁信息,因此,编码效率降低。此外,由于 5 通道全部被交织,需要比单声道音频的内存大 5 倍的内存。

发明内容

[0008] 本发明提供一种多通道音频数据编码方法以及设备,该方法和设备遵守 MPEG 标准并改善传统的多通道 BSAC 方法的性能。

[0009] 本发明还提供一种多通道音频数据解码方法以及设备,该方法和设备遵守 MPEG 标准并改善传统的多通道 BSAC 方法的性能。

[0010] 根据本发明的一方面,提供了一种多通道音频信号编码方法,包括:编码单声道和 / 或立体声音频数据;编码除了单声道和 / 或立体声音频数据之外的扩展的多通道音频数据。单声道和 / 或立体声音频数据可具有分层的比特率。

[0011] 扩展的多通道音频数据可包括扩展的通道的类型信息,该类型信息至少指示音频通道的配置,并被表达为通道配置索引。所述扩展的多通道音频数据的编码可包括:编码指示扩展的多通道音频数据的起始的预定的起始代码 (zero_code, 同步字);以及按通道编码扩展的音频数据。所述起始代码可包括:由 32 比特的连续的 0 形成的 zero_code,以及由 8 比特的连续的 1 形成的同步字。

[0012] 按通道编码扩展的数据的步骤可包括:编码指示音频通道的配置的扩展的通道的类型,以及编码扩展的通道音频数据。扩展的通道的类型可由通道配置索引形成。按通道编码扩展的数据的步骤可包括:编码扩展的数据的长度;以及编码旁信息 (bsac_header, general_header)。

[0013] 扩展的通道音频数据的编码可包括:编码具有最低比特率的基本层;以及编码比特率高于基本层的比特流的增强层,如果有多个增强层,则随增强层的数量而增加比特率。

[0014] 根据本发明的另一方面,提供了一种多通道音频信号编码设备,包括:单声道 / 立体声编码单元,用于编码单声道和 / 或立体声音频数据;以及扩展的数据编码单元,用于编码除了单声道和 / 或立体声音频数据之外的扩展的多通道音频数据。单声道 / 立体声编码单元可编码具有分层的比特率的单声道和 / 或立体声音频数据。

[0015] 扩展的数据编码单元的扩展的多通道音频数据可包括扩展的通道的类型信息,所述的类型信息至少指示音频通道的配置,并被表达为通道配置索引。扩展的数据编码单元可包括:起始代码编码单元,用于编码指示扩展的多通道音频数据的起始的预定起始代码 (zero_code, 同步字);以及通道编码单元,用于按通道编码扩展的音频数据。

[0016] 起始代码编码单元的起始代码可包括:由 32 比特的连续 0 形成的 zero_code,以及由 8 比特的连续 1 形成的同步字。通道编码单元可包括:扩展的通道类型编码单元,用于编码指示音频通道的配置的扩展的通道的类型;以及扩展的音频编码单元,用于编码扩展的通道音频数据。扩展的通道的类型可由通道配置索引形成。通道编码单元可包括:扩展的数据长度编码单元,用于编码扩展的数据的长度;以及旁信息编码单元,用于编码旁信息 (bsac_header, general_header)。

[0017] 扩展的音频编码单元可包括：基本层编码单元，用于编码具有最低比特率的基本层；以及增强层编码单元，用于编码比特流高于基本层的比特率的增强层，如果存在多个增强层，则随增强层的数量而增加比特率。

[0018] 根据本发明的又一方面，提供了一种多通道音频信号解码方法，包括：解码单声道和 / 或立体声音频数据；检验除了单声道和 / 立体声音频数据之外是否存在将被解码的扩展的多通道音频数据；如果存在将被解码的扩展的数据，则解码扩展的多通道音频数据。该单声道和 / 或立体声音频数据可具有分层的比特率。

[0019] 扩展的多通道音频数据可包括扩展的通道的类型信息，所述的类型信息至少指示音频通道的配置，并被表达为通道配置索引。在是否存在扩展的多通道音频数据的检验中，可检验指示扩展的多通道音频数据的起始的预定的起始代码（zero_code、同步字）的存在，如果存在起始代码，则确定存在扩展的数据。所述的起始代码可包括：由 32 比特的连续 0 形成的 zero_code，以及由 8 比特的连续 1 形成的同步字。在扩展的多通道音频数据的解码中，如果存在将被解码的扩展的数据，则可按通道解码扩展的数据。按通道解码扩展的数据的步骤可包括：解码指示音频通道的配置的扩展的通道的类型；以及解码扩展的通道的音频数据。所述的扩展的通道的类型可由通道配置索引形成。

[0020] 按通道解码扩展的数据的步骤可包括：解码扩展的数据的长度；以及解码旁信息（bsac_header, general_header）。扩展的通道音频数据的解码可包括：解码具有最低比特率的基本层；以及解码比特率高于基本层的比特率的增强层，如果存在多个增强层，则随增强层的数量而增加比特率。

[0021] 根据本发明的再一方面，提供了一种多通道音频信号解码设备，包括：单声道 / 立体声解码单元，用于解码单声道和 / 或立体声音频数据；扩展的数据检验单元，检验除了单声道和 / 或立体声音频数据之外是否存在将被解码的扩展的多通道音频数据；以及扩展的数据解码单元，如果存在将被解码的数据，则解码扩展的多通道音频数据。单声道和 / 立体声音频数据具有分层的比特率。所述扩展的数据检验单元可通过预定起始代码（zero_code、同步字）的存在进行检验，所述的预定起始代码指示扩展的多通道音频数据的起始。如果存在起始代码，则确定存在扩展的数据。所述的起始代码可包括：由 32 比特的连续 0 形成的 zero_code；以及由 8 比特的连续 1 形成的同步字。如果存在将被解码的数据，则扩展的数据解码单元按通道解码扩展的数据。所述的扩展的数据解码单元可包括：扩展的通道类型解码单元，用于解码指示音频通道的配置的扩展的通道的类型；以及扩展的通道音频解码单元，用于解码扩展的通道音频数据。所述扩展的通道的类型可由通道配置索引形成。

[0022] 扩展的数据解码单元可包括：扩展的数据长度解码单元，用于解码扩展的数据的长度；以及旁信息解码单元，用于解码旁信息（bsac_header, general_header）。所述扩展的通道音频解码单元可包括：基本层解码单元，用于解码具有最低比特率的基本层；增强层解码单元用于解码比特流高于基本层的比特率的增强层，如果存在多个增强层，则随增强层的数量而增加比特率。

[0023] 根据本发明的又一方面，提供了一种多通道音频信号编码方法，包括：编码单声道 / 立体声音频数据的基本层；编码单声道 / 立体声音频数据的增强层；编码指示扩展的多通道音频数据的起始的预定起始代码（zero_code、同步字）；对构成扩展的多通道音频数据

的至少一个通道数据编码基本层,并对所述的至少一个通道数据编码增强层。

[0024] 所述的对至少一个通道数据的基本层的编码可包括:编码通道数据的长度;编码指示通道的类型的通道配置索引(channel_configuration_index);编码旁信息(bsac_header, general_header);以及编码基本层的音频数据。

[0025] 根据本发明的又一方面,提供了一种多通道音频信号解码方法,包括:解码单声道/立体声音频数据的基本层;解码单声道/立体声音频数据的增强层;检验除了单声道/立体声音频数据之外,是否存在将被解码的扩展的多通道音频数据;如果存在将被解码的扩展的多通道音频数据,则解码预定的起始代码(zero_code、同步字),所述的起始代码指示扩展的多通道音频数据的起始;对构成扩展的多通道音频数据的至少一个通道数据解码基本层,并对所述的至少一个通道数据解码增强层。

[0026] 对至少一个通道数据的基本层的解码包括:解码通道数据的长度;解码指示通道的类型的通道配置索引(channel_configuration_index);解码旁信息(bsac_header, general_header);解码基本层的音频数据。

[0027] 根据本发明的另一方面,提供了一种其上包含有计算程序的计算机可读记录介质,该计算机程序用于执行多通道音频数据编码和解码方法。

附图说明

[0028] 通过下面参照附图对本发明实施例的详细描述,本发明的上述和其他特点和优点将会变得更清楚,其中:

[0029] 图1示出传统的比特分片算术编码(BSAC)多通道的结构;

[0030] 图2是使用传统的BSAC方法的音频编码设备的功能模块的框图;

[0031] 图3是根据本发明的多通道音频数据编码设备的结构的框图;

[0032] 图4是扩展的数据编码单元的详细框图;

[0033] 图5是扩展的音频编码单元的详细框图;

[0034] 图6示出根据本发明的用于多通道音频数据编码的基本数据结构;

[0035] 图7是由根据本发明的多通道音频数据编码方法执行的操作的流程图;

[0036] 图8是用于扩展的通道的音频数据编码的详细的流程图;

[0037] 图9是多通道音频解码设备的结构的框图;

[0038] 图10是图9的扩展的数据解码单元的框图;

[0039] 图11是图9的扩展的通道音频解码单元的框图;

[0040] 图12是根据本发明的多通道音频解码方法的操作的流程图;

[0041] 图13是用于图12中的操作1230的扩展的通道的音频数据解码的详细流程图;

[0042] 图14是示出表示图12中的操作1200至1240的实施例的Basc_raw_data_block()的句法;

[0043] 图15示出表示每个扩展的音频通道解码的实施例的extended_bsac_raw_data_block()的句法;

[0044] 图16示出操作1100的extended_bsac_raw_element()的例子的句法;

[0045] 图17示出通过使用根据本发明的多通道音频信号编码和/或解码方法及设备测量声音质量的测试结果。

具体实施方式

[0046] 现在,将参照附图更全面地描述本发明的多通道音频编码和 / 或解码设备及方法,其中表示了发明的示例性实施例。

[0047] 首先,将解释多通道音频编码设备和方法。图 3 是根据本发明的多通道音频数据编码设备的结构的框图。该设备包括单声道 / 立体声编码单元 300 和扩展的数据编码单元 350。

[0048] 单声道 / 立体声编码单元 300 编码单声道或立体声音频数据。优选地,单声道 / 立体声编码单元 300 编码具有分层的比特率的单声道或立体声音频数据。详细地说,根据 ISO/IEC 14496-3 以比特分片算术编码 (BSAC) 方法来编码单声道或立体声音频数据是更可取的。由于 BSAC 方法的音频编码是公知技术,所以这里将省略对该方法的解释。

[0049] 扩展的数据编码单元 350 除了编码单声道或立体声音频数据外还编码扩展的多通道音频数据。

[0050] 优选地,扩展的多通道音频数据至少包括指示音频通道的配置的扩展的通道的类型信息,扩展的通道类型信息被表达为通道配置索引 (channel_configuration_index)。优选地,通道配置索引具有指示音频输出通道配置的 3 比特字段,如表 1 中所示。因此,通道配置索引指示与通道相应的每个扬声器的特性。

[0051] 表 1

[0052]

索引	通道到扬声器的映射	通道数 (nch)
0	前中扬声器	1
1	左、右前扬声器	2
2	后环绕扬声器	1
3	左环绕、右环绕后扬声器	2
4	前低频效果扬声器	1
5	左、右外前扬声器	2
6-7	保留	-

[0053] 图 4 是扩展的数据编码单元 350 的详细框图,包括起始代码编码单元 400 和通道编码单元 450。起始代码编码单元 400 编码预定的起始代码,所述的起始代码指示扩展的多通道音频数据的起始。起始代码形成有 zero_code 和同步字。zero_code 是由指示立体声音频数据的算术解码完成的 32 比特的连续 0 形成。同步字由指示扩展的多通道音频数据的起始的 8 比特的连续 1 形成。该比特串是 1111 1111。

[0054] 通道编码单元 450 编码每个通道中的扩展的音频数据,并形成有扩展的通道长度编码单元 452、扩展的通道类型编码单元 454、旁信息编码单元 456 和扩展的音频编码单元 458。

[0055] 扩展的通道长度编码单元 452 编码扩展的数据的长度。当执行算术解码时,扩展的数据长度信息被使用。

[0056] 扩展的通道类型编码单元 454 编码指示音频通道的配置的扩展的通道类型。旁信息编码单元 456 编码旁信息 (bsac_header、general_header)。旁信息 (bsac_header、general_header) 与当以 BSAC 方法编码单声道或立体声音频数据时使用的旁信息相同。扩展的音频编码单元 458 编码扩展的通道音频数据。

[0057] 图 5 是扩展的音频编码单元 458 的详细框图。该扩展的音频编码单元 458 包括基本层编码单元 500 和增强层编码单元 550。基本层编码单元 500 编码具有最低比特率的基本层。增强层编码单元 550 编码其比特率比基本层的比特率高的增强层,如果有多个层,则比特率随层数而增加。

[0058] 本发明使用在传统的立体声比特率中扩展的通道的方法。通道配置索引被指定给每个通道元素,并且指示修改关于当编码音频时每个可用工具的旁信息的可能性。由于在窗口、M/S、和 PNS 信息的每个通道元素中存在通用头 (general header),所以可修改需要修改的所有工具。

[0059] 图 6 示出根据本发明的用于多通道音频数据编码的基本数据结构。图 7 是根据本发明的多通道音频数据编码方法的操作的流程图。参照图 3 和图 7,将解释根据本发明的多通道音频编码方法和设备的操作。

[0060] 首先,在操作 700 中,在单声道 / 立体声编码单元 300 中编码单声道或立体声音频数据。然后,在扩展的数据编码单元 350 中,除了单声道或立体声数据之外,扩展的多通道音频数据被编码。优选地,如上所述,单声道或立体声数据具有分层的比特率。此外,扩展的多通道音频数据包括如上所述的扩展的通道类型信息,该信息至少指示音频通道的配置,并被表达为通道配置索引。

[0061] 现在将更详细地解释扩展的多通道音频数据的编码。单声道或立体声音频数据被编码,然后在操作 710 中检验将被编码的数据是否存在。如果将被编码的数据存在,则在操作 720 中,在起始代码编码单元 400 中编码指示扩展的多通道音频数据的起始的预定起始代码 (zero_code、同步字)。该起始代码与上述编码设备中的起始代码相同。然后,通过通道编码单元 450 编码用于每个通道的扩展的音频数据。这里,首先在操作 730 中编码用于一个通道的扩展的音频数据,当该通道的编码完成时,在操作 740 中检验是否存在用于另一个通道的将被编码的音频数据。如果用于另一通道的将被编码的音频数据存在,则编码用于该通道的音频数据。对所有扩展的通道执行该处理。

[0062] 图 8 是操作 730 中的用于扩展的通道的音频数据编码的详细流程图。在操作 800 中,在扩展的数据长度编码单元 452 中编码扩展的数据的长度。此外,在操作 820 中,在扩展的通道类型编码单元 454 中编码指示音频通道的配置的扩展的通道类型。在操作 840 中,在旁信息编码单元 456 中编码旁信息 (bsac_header、general_header)。然后,在操作 860 中,在扩展的音频编码单元 458 中编码扩展的通道音频数据。

[0063] 在操作 860 中的扩展的通道音频数据的编码中,首先,在基本层编码单元 500 中编

码具有最低比特率的基本层中的音频数据,然后在增强层编码单元 550 中编码增强层的音频数据。增强层具有比基本层的比特率高的比特率。当多个增强层存在时,比特率随着增强层的数量而增加。

[0064] 同时,将解释根据本发明的多通道音频解码设备和方法。基本上,多通道音频解码以与编码操作相反的顺序执行操作。

[0065] 图 9 是多通道解码设备的结构的框图。该设备包括单声道 / 立体声解码单元 900、扩展的数据检验单元 920 和扩展的数据解码单元 940。

[0066] 单声道 / 立体声解码单元 900 解码单声道或立体声音频数据。优选地,单声道或立体声音频数据具有分层的比特率,并根据 ISO/IEC 14496-3 以 BASC 方法被解码。

[0067] 扩展的数据检验单元 920 检验除了单声道或立体声音频数据之外是否存在将被解码的扩展的多通道音频数据。扩展的数据检验单元 920 检验指示扩展的多通道音频数据的起始的预定起始代码 (zero_code、同步字) 的存在,如果存在起始代码,则确定存在扩展的数据。该起始代码由 zero_code 和同步字形成。zero_code 由指示立体声音频数据的算术解码完成的 32 比特的连续的 0 形成。同步字由指示扩展的多通道音频数据的起始的 8 比特的连续 1 形成。该比特串是 1111 1111。

[0068] 如果存在将被解码的扩展的数据,则扩展的数据解码单元 940 解码扩展的多通道音频数据。此外,优选地,当执行解码时,扩展的数据解码单元 940 按通道解码扩展的数据。

[0069] 图 10 是图 9 的扩展的数据解码单元 940 的框图,包括扩展的数据长度解码单元 1000、扩展的通道类型解码单元 1020、旁信息解码单元 1040 和扩展的通道音频解码单元 1060。

[0070] 扩展的数据长度解码单元 1000 解码扩展的数据的长度信息。扩展的通道类型解码单元 1020 解码指示音频通道的配置的扩展的通道类型。优选地,扩展的通道类型信息被表达为通道配置索引 (channel_configuration_index)。通道配置索引定义通道被映射到扬声器时的通道数,并如表 1 所示具有指示音频输出通道配置的 3 比特字段。

[0071] 旁信息解码单元 1040 解码旁信息。旁信息对解码音频数据是必需的,所述旁信息是除了音频数据之外的信息,如 bsac header 和 general header。基本上,该旁信息 (bsac_header 和 general_header) 与 BSAC 方法中解码单声道或立体声音频数据所需的旁信息相同。

[0072] 扩展的通道音频解码单元 1060 解码扩展的音频数据。图 11 是图 10 的扩展的通道音频解码单元 1060 的框图,该通道音频解码单元包括基本层解码单元 1100 和增强层解码单元 1150。基本层解码单元 1100 解码具有最低比特率的基本层。增强层解码单元解码其比特率比基本层的比特率高的增强层。如果存在多个层,则增加随增加的层数而增加的比特率。

[0073] 图 12 是由根据本发明的多通道音频解码方法执行的操作的流程图。参照图 12,将解释根据本发明的多通道音频数据解码方法和设备的操作。

[0074] 首先,在操作 1200 中,通过单声道 / 立体声解码单元 900 解码单声道或立体声音频数据。然后,在操作 1210 中,由扩展的数据检验单元 920 检验除了单声道 / 立体声音频数据之外是否存在扩展的多通道音频数据。在操作 1220 中,通过解码预定起始代码 (zero_code、同步字) 并检验起始代码的存在来确定扩展的多通道音频数据的存在,所述的预定

起始代码指示扩展的多通道音频数据的起始。如果存在起始代码,则确定存在扩展的数据。即,如果存在 zero_code,则表明单声道或立体声音频数据的解码被完成,如果其后存在同步字,则表明存在将被解码的多通道音频数据。

[0075] 如果通过起始代码确定存在将被解码的扩展的数据,则在操作 1230 中通过扩展的数据解码单元 940 解码扩展的多通道音频数据。

[0076] 如图 14 所示,以句法 (Bsac_raw_data_block()) 表达了操作 1200 至 1230 的实施例。

[0077] 参照图 14, Bsac_raw_data_block() 是包含编码的音频数据、相关的信息和其他数据的原始数据块,并且主要由 bsac_base_element() 和几个 basc_layer_element() 形成。Bsac_raw_data_block() 是用于确定 bsac 比特流是否具有扩展部分的模块。优选地,如上所述单声道或立体声数据具有分层的比特率。此外,扩展的多通道音频数据包括扩展的通道在上面描述的类型信息,该类型信息至少指示音频通道的配置并被表达为通道配置索引。

[0078] 在关于一个通道的扩展的音频数据在操作 1230 中被解码之后,在操作 1240 中检验是否存在将被解码的用于另一通道的音频数据。如果存在将被解码的用于另一通道的音频数据,则解码用于另外的通道的音频数据。通过对所有扩展的通道执行该处理,解码所有扩展的通道音频数据。

[0079] 图 15 中示出表示每个音频通道的解码的实施例的句法 (extended_bsac_raw_block())。

[0080] 参照图 15, extended_bsac_raw_block() 是包括与多通道扩展的数据相应的编码的音频数据以及与该音频数据相关的信息的原始数据块。extended_bsac_raw_block() 主要由 extended_bsac_element() 和几个 bsac_layer_element() 形成。

[0081] 图 13 是操作 1230 的用于扩展的通道的音频数据解码的详细流程图。在操作 1300 中,在扩展的数据长度解码单元 1000 中解码扩展的数据的长度。此外,在操作 1320 中,在扩展的通道类型解码单元 1020 中解码指示音频通道的配置的扩展的通道类型。在操作 1340 中,在旁信息解码单元 1040 中解码旁信息 (bsac header、general header)。解码操作 1300 至 1340 的执行顺序无关紧要。然后,在操作 1360 中,在扩展的通道音频解码单元 1060 中解码扩展的通道音频数据。

[0082] 在操作 1360 的扩展的音频数据的解码中,具有最低比特率的基本层的音频数据首先在基本层解码单元 1100 中被解码,然后,增强层的音频数据在增强层解码单元 1150 中被解码。增强层的比特率高于基本层的比特率,如果存在多个增强层,则比特率随着增强层数的增加而增加。图 16 表示操作 1230 的句法 (extended_bsac_raw_data_block()) 的实施例。

[0083] 参照图 16, extended_bsac_raw_data_block() 是基本层比特流的句法的元素,包含与 BSAC 扩展的部分相应的编码的音频数据以及与音频数据相关的信息。

[0084] 本发明也可被实现为计算机 (包括具有信息处理功能的所有设备) 可读记录介质上的计算机可读代码。计算机可读记录介质是能够存储可由计算机系统随后读取的数据的任何数据存储装置。计算机可读记录介质的例子包括只读存储器 (ROM)、随机存取存储器 (RAM)、CD-ROM、磁带、软盘和光学数据存储装置。

[0085] 虽然已经参照本发明的示例性实施例具体表示和描述了本发明,但是本领域的普通技术人员应该理解,在不脱离由权利要求限定的本发明的精神和范围的情况下,可在其中作出形式和细节的各种改变。这些优选实施例应该被理解为只是描述性的,而不是为了限制的目的。因此,本发明的范围不是由本发明的详细描述来限定,而是所附权利要求限定。该范围内的所有不同应该认为包括在本发明中。

[0086] 根据本发明的多通道音频编码和 / 或解码设备以及方法,多通道数据交错需要的内存比使用传统的 BSAC 方法所需的内存少 20%。这是因为当使用根据本发明的多通道方法时,正被添加的通道元素被顺序地处理,因此同时进行的内存使用的数量相对小,而在传统的多通道方法中,整个多通道的所有数据应该被装载到内存上。

[0087] 图 17 表示通过使用根据本发明的多通道音频信号编码和 / 或解码方法测量声音质量的结果。

[0088] 收听试验条件如下。使用窗口切换和 M/S 立体声工具并且在前和后通道元件的每个中控制比特率。参与试验的四个音频专家以及传统的 BSAC 相比的相对声音质量 (-2~+2) 被测量。对于测试项,用于 MPEG-2 NBC 的共 46 项被选择。

[0089] 根据本发明的多通道音频编码和 / 或解码方法以及设备,通过仅仅一个比特流,可根据用户环境提供单声道、立体声和多通道音频。此外,在多通道音频中,根据用户终端和网络的状态来提供 FGS 功能。而且,可实现多通道 BSAC 的性能的增强,例如,高的声音质量、低的复杂性和可分级性。详细地说,可满足对 MPEG 标准化 (传统 BSAC 的兼容性、维持 FGS 功能和最小的修改) 的各种需求,此外,该方法和设备可被用于更逼真的数字多媒体广播和基于移动影院的和基于家庭影院的服务中。

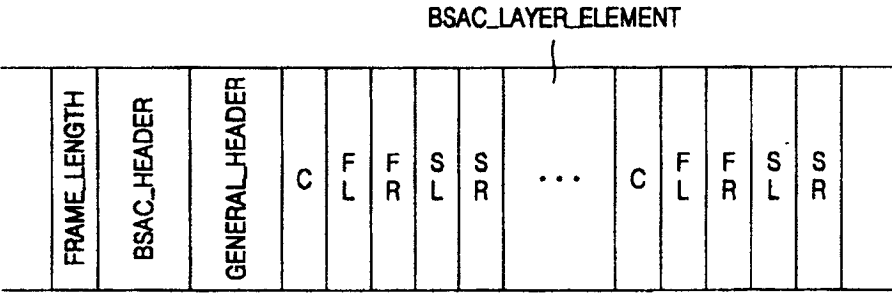
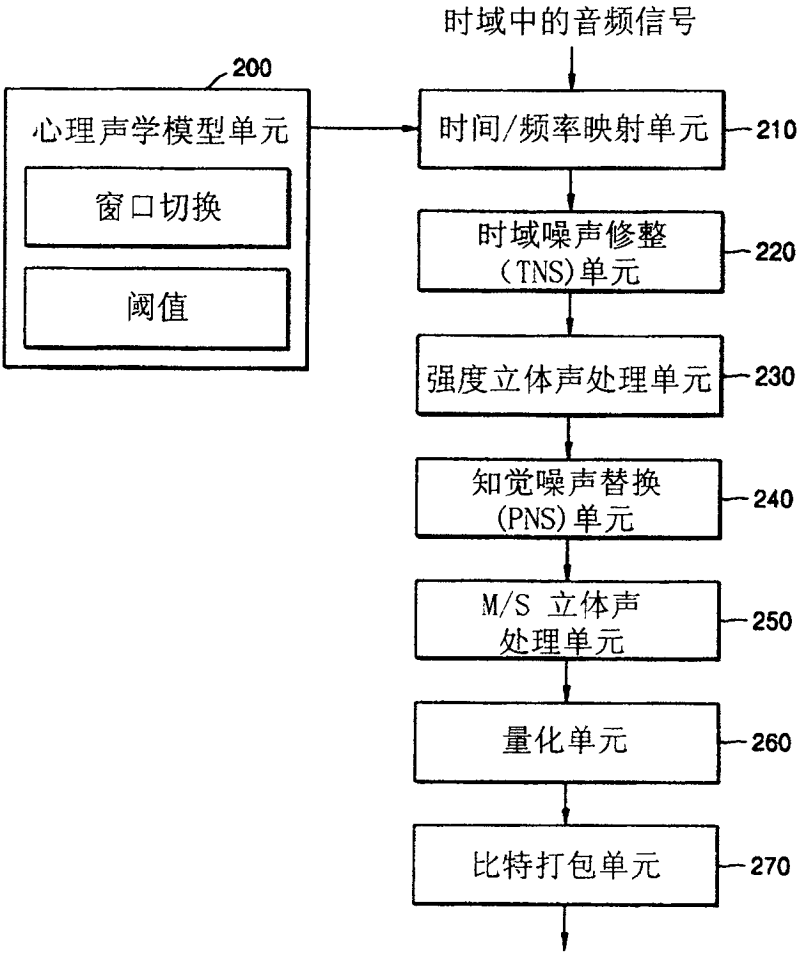


图 1



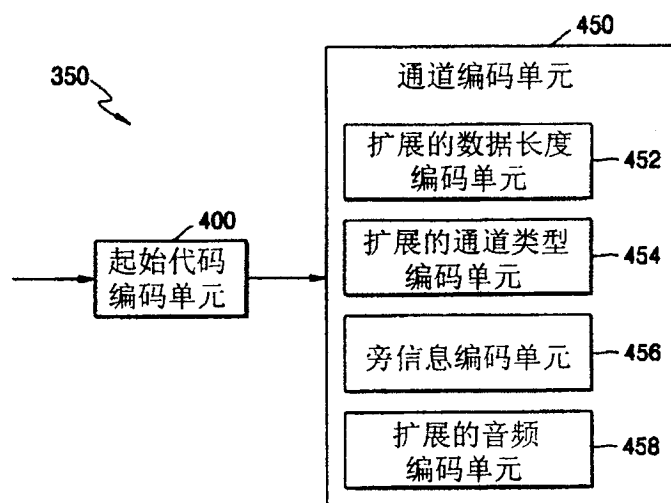


图 4

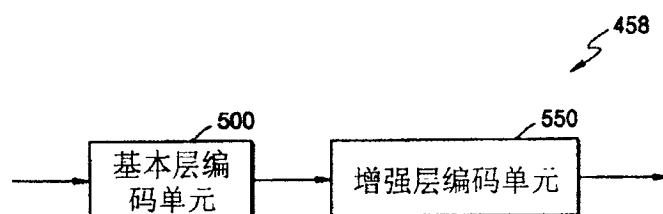


图 5

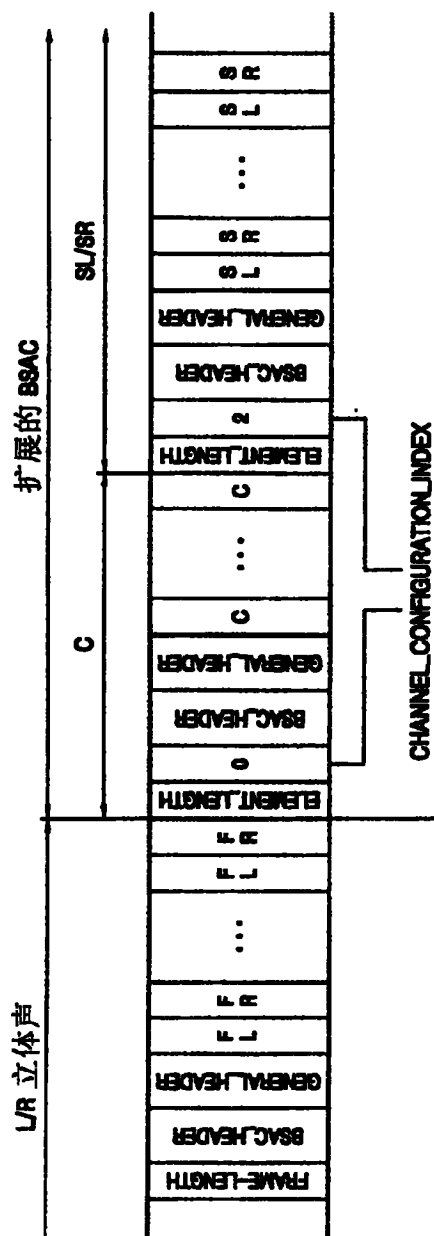


图 6

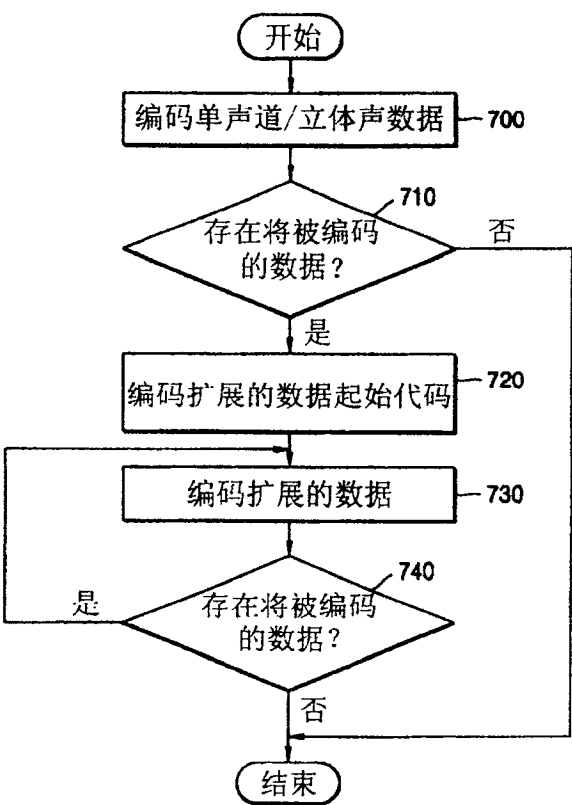


图 7

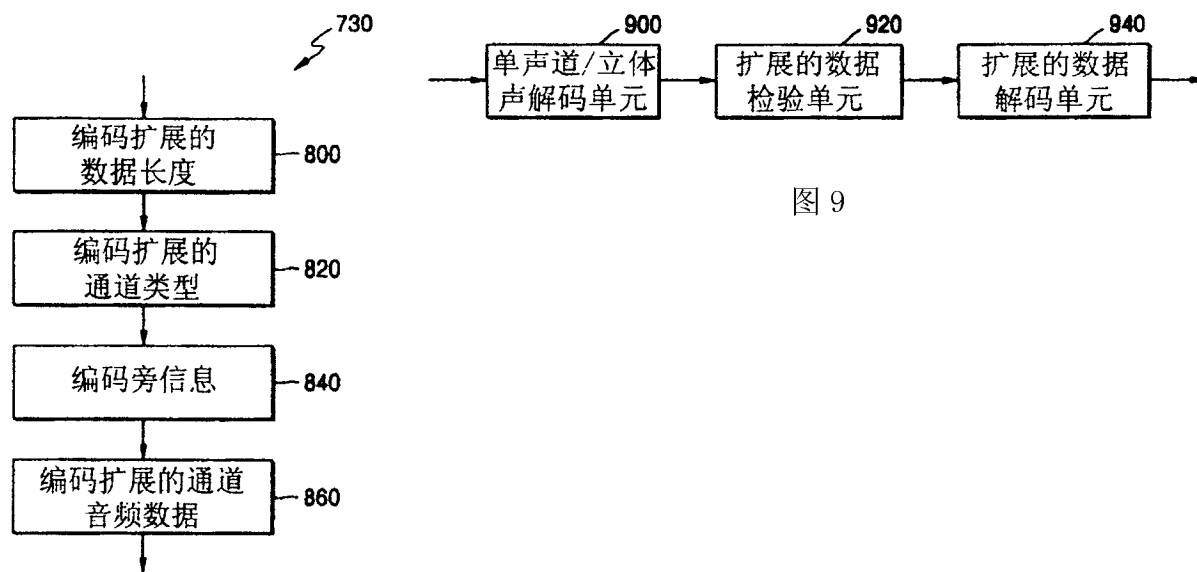


图 9

图 8

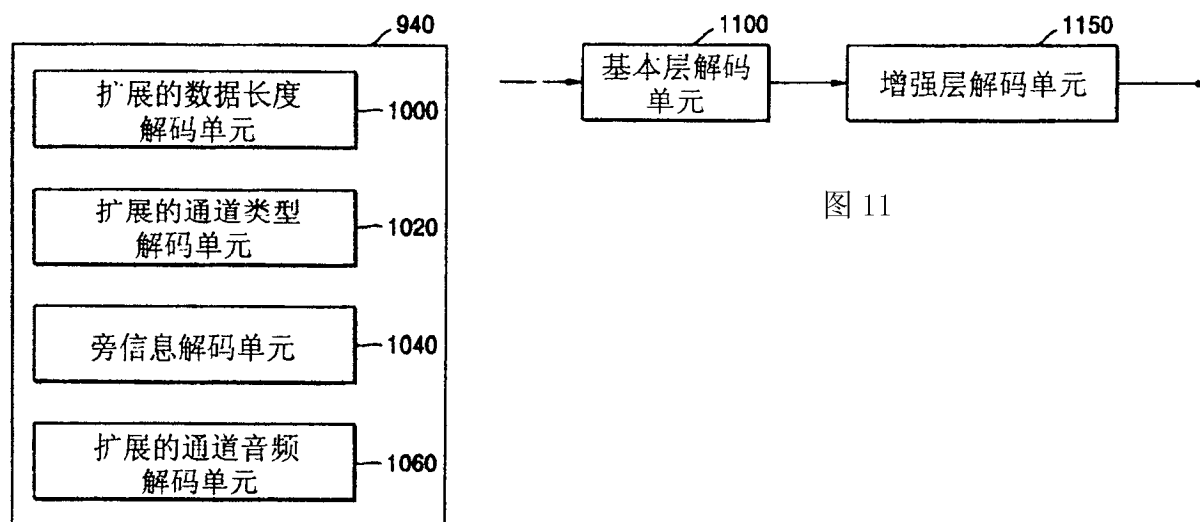


图 11

图 10

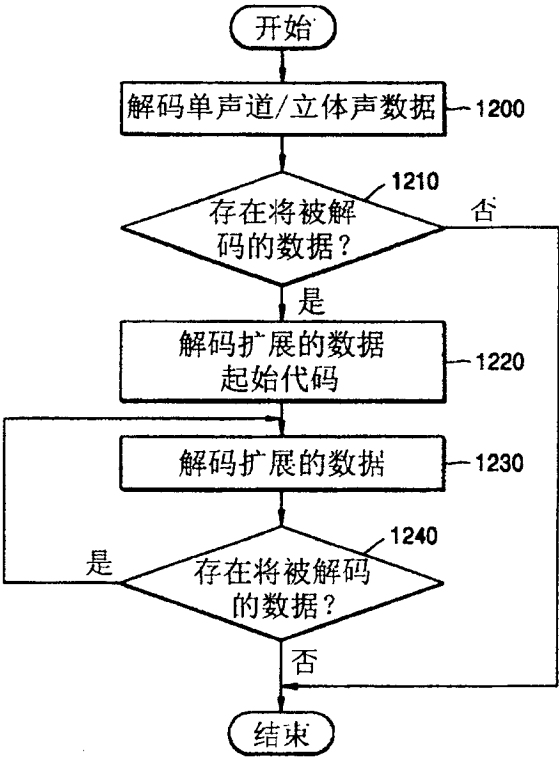


图 12

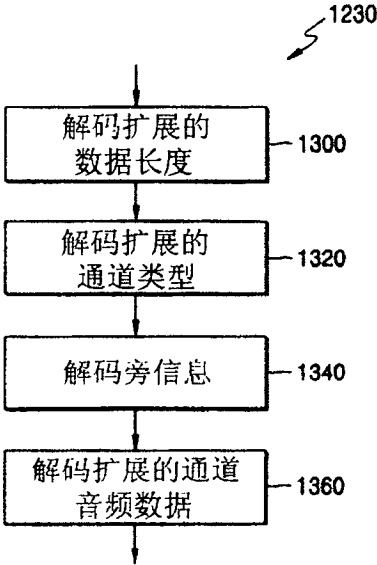


图 13

bsac_raw_data_block() 的句法

Syntax	No. of bits	Mnemonic
Bsac_raw_data_block() { bsac_base_element(); layer=slayer_size; while(data_available() && layer<(top_layer+slayer_size)) { bsac_layer_element(layer); layer++; } byte_alignment(); if (data_available()) { Zero_code Syncword while(data_available()) extended_bsac_raw_data_block(); } }		
Zero_code	32	bslbf
Syncword	8	bslbf

图 14

extended_bsac_raw_data_block() 的句法

Syntax	No. of bits	Mnemonic
<pre> extended_bsac_raw_data_block() { extended_bsac_base_element(); layer=slayer_size; while(data_available() && layer<(top_layer+slayer_size)) { bsac_layer_element(layer); layer++; } byte_alignment(); } </pre>		

图 15

extended_bsac_base_element() 的句法

Syntax	No. of bits	Mnemonic
<pre> extended_bsac_base_element() { element_length 11 uimbf channel_configuration_index 3 uimbf reserved_bit 1 uimbf bsac_header(); general_header(); byte_alignment(); for (slayer = 0; slayer < slayer_size; slayer++) bsac_layer_element(slayer); } </pre>		

图 16

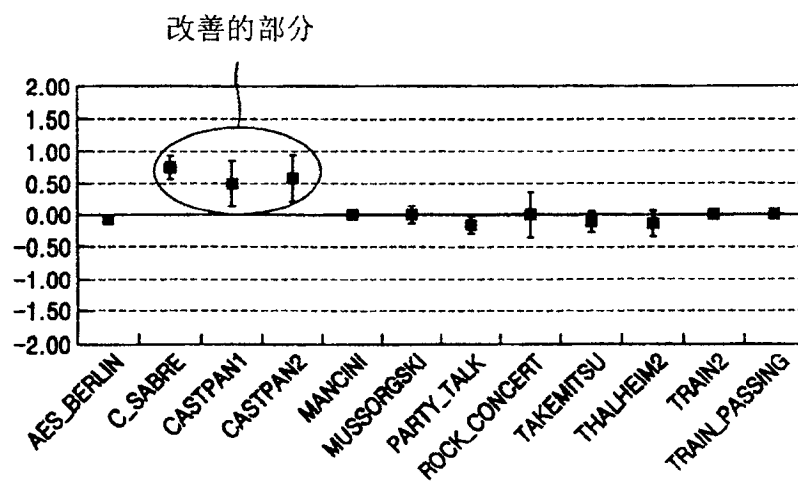


图 17