(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101617360 B
(45)授权公告日 2012.08.22

(21)申请号 200780043560.3

(22)申请日 2007.10.01

(30)优先权数据
10-2006-0096172 2006.09.29 KR

(85)PCT申请进入国家阶段日
2009.05.25

(86)PCT申请的申请数据
PCT/KR2007/004795 2007.10.01

(87)PCT申请的公布数据
WO2008/039038 EN 2008.04.03

(73)专利权人 韩国电子通信研究院
地址 韩国大田市

(72)发明人 白承权 徐廷一 李泰辰 李用主 张仁瑄 刘载铉 张大永 洪镇佑 金镇雄 姜京玉

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 李芳华

(51)Int.Cl.
G10L 19/00(2006.01)

(56)对比文件
US 2005105442 A1,2005.05.19,
CN 1525438 A,2004.09.01,
CN 1787078 A,2006.06.14,
JP 2003032800 A,2003.01.31,
Faller C et al.Parametric Joint-Coding of Audio Sources.《Audio Engineering Society the 102th Convention》.2006,
Herre,J et al.The Teference Model Architecture forMPEG Spatial Audio Coding.《Audio Engineering Society the 118th Convention》.2005,
Faller,C et al.Binaural Cue Coding: A Novel and Efficient Representation of Spatial Audio.《ICASSP,IEEE International Conference on Acoustic,Speech and Signal Processing Proceeding》.2002,(第2期),

审查员 张鑫

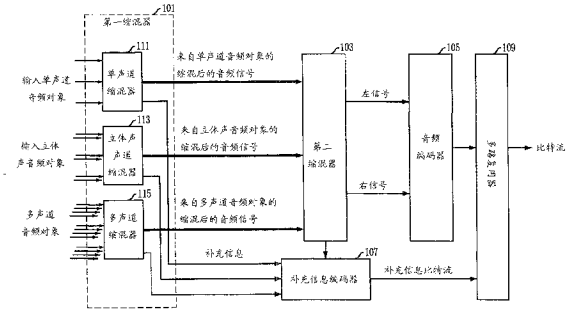
权利要求书 4 页 说明书 10 页 附图 12 页

(54)发明名称

用于编码和解码具有各种声道的多对象音频信号的设备和方法

(57)摘要

提供了一种用于编码和解码多对象音频信号的设备和方法。该设备包括:缩混器,用于将音频信号缩混为一个缩混后的音频信号,并提取用于所述音频信号中的每一个的包括报头信息和空间线索信息的补充信息;编码器,用于对缩混后的音频信号进行编码;和补充信息编码器,用于生成该补充信息作为比特流。该报头信息包括用于所述音频信号中的每一个的标识信息和用于所述音频信号的声道信息。



CN 101617360 B

1. 一种用于编码具有不同声道的多对象音频信号的设备，包括：

缩混部件，用于将音频信号缩混为一个缩混后的音频信号，并提取用于所述音频信号中的每一个的包括报头信息和空间线索信息的补充信息；

编码部件，用于对缩混后的音频信号进行编码；和

补充信息编码部件，用于生成该补充信息作为比特流，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

2. 根据权利要求 1 的设备，其中该声道信息包括：

用于每一声道的音频对象信息；和

用于所述音频信号中的每一声道的音频对象的数目。

3. 根据权利要求 1 的设备，其中该报头信息还包括作为基于空间线索的音频信息的时隙信息或音频信号的预置信息。

4. 根据权利要求 3 的设备，其中该预置信息包括：

用于定义用于所述音频信号的预置模式的预置模式信息；和

用于定义用于支持该预置模式所需的信息的预置模式支持信息，

其中该预置模式信息包括从由卡拉 OK 模式信息、独奏对象提取模式信息、偏好渲染信息、和重放模式设置信息组成的组中所选择的至少一个，并且该预置模式支持信息包括从由音频对象索引信息、每一对象的渲染信息、和每一音频对象的最佳渲染信息组成的组中所选择的至少一个。

5. 根据权利要求 1 的设备，其中该空间线索信息顺序包括用于所述音频信号的单声道和立体声音频对象的空间线索信息和用于多声道音频对象的空间线索信息。

6. 根据权利要求 1 的设备，其中该缩混部件包括：

第一缩混器，用于对于每一声道的所述音频信号进行缩混；和

第二缩混器，用于将来自该第一缩混器的缩混后的信号缩混为一个缩混后的信号。

7. 根据权利要求 6 的设备，其中该第一缩混器包括基本缩混器，用于提取所述音频信号中包括的用于单声道的音频信号的补充信息，并对所述单声道的音频信号进行缩混，

其中按照层叠结构安排 (N-1) 个基本缩混器，N 是所述音频信号的数目，或者一个基本缩混器基于层叠方案执行缩混操作 (N-1) 次，N 是所述音频信号的数目。

8. 根据权利要求 6 的设备，其中该第一缩混器包括基本缩混器，用于提取有关作为所述音频信号中包括的立体声声道的音频信号的左信号和右信号的补充信息，并对所述左信号和右信号进行缩混，

其中对于所述左信号和右信号，按照层叠结构安排 (M-1) 个基本缩混器，M 是左信号或右信号的数目，或者一个基本缩混器基于层叠方案执行缩混操作 2*(M-1) 次。

9. 根据权利要求 6 的设备，其中该第一缩混器包括多声道缩混器，用于提取用于所述音频信号中包括的多声道音频信号的补充信息，并对所述多声道音频信号进行缩混，

其中该多声道缩混器基于运动画面专家组 (MPEG) 环绕方案来对所述多声道音频信号进行缩混，或者该多声道缩混器基于空间音频编码 (SAC) 方案来对所述多声道音频信号进行缩混。

10. 根据权利要求 6 的设备，其中该第二缩混器包括：

第一基本缩混器，用于提取由该第一缩混器缩混为立体声声道的缩混后的信号的左信号和右信号中的每一个的补充信息，并对由该第一缩混器缩混为立体声声道的缩混后的信号的左信号和右信号中的每一个进行缩混；和

第二基本缩混器，用于从由该第一基本缩混器和该第一缩混器缩混的缩混后的信号中提取补充信息，并将由该第一基本缩混器和该第一缩混器缩混的缩混后的信号缩混为立体声声道信号。

11. 根据权利要求 10 的设备，其中该第二基本缩混器基于如下表达的方程来将该第一基本缩混器和该第一缩混器缩混的缩混后的信号缩混为立体声声道信号：

$$\begin{bmatrix} w_b^{11} & w_b^{12} & w_b^{13} \\ w_b^{21} & w_b^{22} & w_b^{23} \end{bmatrix} \begin{bmatrix} s_b^1(f) \\ s_b^2(f) \\ s_b^3(f) \end{bmatrix}$$

其中 w_b^{ij} 表示加权因子；$s_b^j(f)$ 表示该第一基本缩混器和该第一缩混器所缩混的缩混后的信号；并且 b 是表示子带索引，或者该第二基本缩混器是 MPEG 环绕的三到二 (TTT) 盒。

12. 根据权利要求 1 的设备，还包括多路复用部件，用于对来自该编码部件的编码后的音频信号和来自该补充信息编码部件的所生成的补充信息进行多路复用。

13. 一种用于编码具有不同声道的多对象音频信号的方法，包括步骤：

将音频信号缩混为一个缩混后的音频信号，并提取用于所述音频信号中的每一个的包括报头信息和空间线索信息的补充信息；

对缩混后的音频信号进行编码；和

生成该补充信息作为比特流，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

14. 一种用于解码由不同声道组成的多对象音频信号的设备，包括：

输入信号分析部件，用于从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充信息比特流中提取具有报头信息和空间线索信息的补充信息；

音频对象提取部件，用于使用来自该输入信号分析部件的所提取的补充信息，来从所恢复的缩混后的音频信号中恢复每一对象的音频信号；和

输出部件，用于使用所输入的所述音频信号的控制信息，来输出所恢复的每一对象的音频信号作为多对象音频信号，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

15. 根据权利要求 14 的设备，其中该声道信息包括：

用于每一声道的音频对象信息；和

用于所述音频信号中的每一声道的音频对象的数目。

16. 根据权利要求 14 的设备，其中该报头信息还包括作为基于空间线索的音频信息的时隙信息，或者该报头信息还包括用于音频信号的预置信息，

其中该预置信息包括：

用于定义用于所述音频信号的预置模式的预置模式信息；和

用于定义用于支持该预置模式所需的信息的预置模式支持信息，

其中该预置模式信息包括从由卡拉 OK 模式信息、独奏对象提取模式信息、偏好渲染信息、和重放模式设置信息组成的组中所选择的至少一个，并且该预置模式支持信息包括从由音频对象索引信息、每一对象的渲染信息、和每一音频对象的最佳渲染信息组成的组中所选择的至少一个。

17. 根据权利要求 14 的设备，其中该空间线索信息顺序包括用于所述音频信号的单声道和立体声音频对象的空间线索信息和用于多声道音频对象的空间线索信息，并且该控制信息是从由用于所述音频信号的渲染控制信息和输出声道控制信息组成的组中选择的至少一个，并且该输入信号分析部件包括：

解多路复用单元，用于从输入的信号中分离出音频信息比特流和补充信息比特流；

音频恢复单元，用于从该解多路复用单元所分离出的音频信息比特流中恢复缩混后的音频信号；和

补充信息分析单元，用于从该该解多路复用单元所分离出的补充信息比特流中提取补充信息。

18. 一种用于解码由不同声道组成的多对象音频信号的方法，包括步骤：

从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充信息比特流中提取具有报头信息和空间线索信息的补充信息；

使用所提取的补充信息，来从所恢复的缩混后的音频信号中恢复每一对象的音频信号；和

使用所输入的所述音频信号的控制信息，来输出所恢复的每一对象的音频信号作为多对象音频信号，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

19. 一种用于解码由不同声道组成的多对象音频信号的设备，包括：

输入信号分析部件，用于从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充比特流中提取包括报头信息和空间线索信息的补充信息；

补充信息控制部件，用于使用所输入的所述音频信号的控制信息，来控制所提取的补充信息；和

输出部件，用于使用所控制的补充信息，来输出所恢复的缩混后的音频信号作为多对象音频信号，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

20. 一种用于解码由不同声道组成的多对象音频信号的方法，包括步骤：

从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充比特流中提取包括报头信息和空间线索信息的补充信息；

使用所输入的所述音频信号的控制信息，来控制所提取的补充信息；和

使用所控制的补充信息，来输出所恢复的缩混后的音频信号作为多对象音频信号，

其中该报头信息包括：

用于所述音频信号中的每一个的标识信息；和

用于所述音频信号的声道信息。

用于编码和解码具有各种声道的多对象音频信号的设备和方法

技术领域

[0001] 本发明涉及用于编码和解码多对象音频信号的设备和方法；并且，更具体地，涉及用于对具有各种声道的多对象音频信号进行编码和解码以及对形成有各种声道的多对象音频信号进行编码和解码的设备和方法。

[0002] 具有各种声道的多对象音频信号是包括其每一个形成有不同声道（例如，单声道、立体声声道、和 5.1 声道）的多个音频对象的音频信号。

[0003] 该工作得到韩国信息通信部 (MIC) 和 / 或韩国信息技术 (IT) 研究院 (IITA) 的信息技术研究和开发规划的部分支持 [2005-S-403-02，“super-intelligentmultimedia anytime-anywhere realistic TV(SmaRTV)technology”]。

背景技术

[0004] 根据现有技术的音频编码和解码技术使得用户能够被动收听音频内容。因此，已存在对于以下对由不同声道组成的多个音频对象进行编码和解码的设备和方法的需求，通过根据用户需求控制由不同声道组成的每一音频对象，使用各种方法合并一个音频上下文，而使得用户能够消费各种音频对象。

[0005] 作为现有技术，介绍了空间音频编码 (SAC)。SAC 是用于将多声道音频信号表达为缩混后的 (down mixed) 单声道信号或缩混后的立体声信号和空间线索、传送并恢复多声道音频信号的技术。基于 SAC，可以以低比特率传送高质量多声道音频信号。

[0006] 然而，SAC 不能编码和解码多声道多对象音频信号 - 例如包括其每一个由诸如单声道、立体声、和 5.1 声道的不同声道组成的各种对象的音频信号，因为尽管音频信号由多个声道组成，但是 SAC 是用于编码和解码单对象音频信号的技术。

[0007] 作为另一现有技术，介绍了双耳线索编码 (BCC)。BCC 可编码和解码多对象信号。然而，BCC 不能编码和解码由除了单声道的各种声道组成的多对象音频信号，因为在 BCC 中音频对象限于形成有单声道的音频对象。

[0008] 如上所述，根据现有技术的音频信号编码和解码技术不能对由各种声道组成的多对象音频信号进行编码和解码，因为它们被设计为对由单声道组成的多对象信号或具有多声道的单对象音频信号进行编码和解码。所以，根据现有技术，用户必须根据音频信号编码和解码技术来被动收听音频上下文。

[0009] 所以，已存在对于以下对由各种声道组成的多个音频对象进行编码和解码的设备和方法的需求，通过根据用户需求控制其每一个具有不同声道的每一音频对象，使用各种方法混合一个音频内容，以便消费各种音频对象。

[0010] 发明内容

[0011] 技术问题

[0012] 本发明的实施例旨在提供一种用于对具有各种声道的多对象音频信号进行编码和解码以及对由各种声道组成的多对象音频信号进行编码和解码的设备和方法。

[0013] 本发明的其他目的和优点可通过以下描述而理解，并通过参考本发明的实施例而变得清楚。而且，对于本发明的领域的技术人员来说显而易见的是，可通过要求保护的部件及其组合来实现本发明的目的和优点。

[0014] 技术方案

[0015] 根据本发明的一个方面，提供了一种用于编码具有不同声道的多对象音频信号的设备，包括：缩混单元，用于将音频信号缩混为一个缩混后的音频信号，并提取用于所述音频信号中的每一个的包括报头信息和空间线索信息的补充信息；编码单元，用于对缩混后的音频信号进行编码；和补充信息编码单元，用于生成该补充信息作为比特流，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；和用于所述音频信号的声道信息。

[0016] 根据本发明的另一方面，提供了一种用于编码具有不同声道的多对象音频信号的方法，包括步骤：将音频信号缩混为一个缩混后的音频信号，并提取用于所述音频信号中的每一个的包括报头信息和空间线索信息的补充信息；对缩混后的音频信号进行编码；和生成该补充信息作为比特流，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；和用于所述音频信号的声道信息。

[0017] 根据本发明的另一方面，提供了一种用于解码由不同声道组成的多对象音频信号的设备，包括：输入信号分析单元，用于从输入的信号中恢复缩混 后的音频信号，并从该输入的信号中包括的补充信息比特流中提取具有报头信息和空间线索信息的补充信息；音频对象提取单元，用于使用来自该输入信号分析单元的所提取的补充信息，来从所恢复的缩混后的音频信号中恢复每一对象的音频信号；和输出单元，用于使用该输入的信号的控制信息，来输出所恢复的每一对象的音频信号作为多对象音频信号，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；和用于所述音频信号的声道信息。

[0018] 根据本发明的另一方面，提供了一种用于解码由不同声道组成的多对象音频信号的方法，包括步骤：从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充信息比特流中提取具有报头信息和空间线索信息的补充信息；使用所提取的补充信息，来从所恢复的缩混后的音频信号中恢复每一对象的音频信号；和使用该输入的信号的控制信息，来输出所恢复的每一对象的音频信号作为多对象音频信号，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；和用于所述音频信号的声道信息。

[0019] 根据本发明的另一方面，提供了一种用于解码由不同声道组成的多对象音频信号的设备，包括：输入信号分析单元，用于从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充比特流中提取包括报头信息和空间线索信息的补充信息；补充信息控制单元，用于使用该输入的信号的控制信息，来控制所提取的补充信息；和输出单元，用于使用所控制的补充信息，来输出所恢复的缩混后的音频信号作为多对象音频信号，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；和用于所述音频信号的声道信息。

[0020] 根据本发明的另一方面，提供了一种用于解码由不同声道组成的多对象音频信号的方法，包括步骤：从输入的信号中恢复缩混后的音频信号，并从该输入的信号中包括的补充比特流中提取包括报头信息和空间线索信息的补充信息；使用该输入的信号的控制信息，来控制所提取的补充信息；和使用所控制的补充信息，来输出所恢复的缩混后的音频信号作为多对象音频信号，其中该报头信息包括：用于所述音频信号中的每一个的标识信息；

和用于所述音频信号的声道信息。

[0021] 有利效果

[0022] 根据本发明实施例的用于对具有各种声道的多对象音频信号进行编码和解码以及对由各种声道组成的多对象音频信号进行编码和解码的设备和方法通过有效编码和解码由不同声道组成的包括各种音频对象的音频内容，而使得用户能够根据其需求来主动消费音频内容。

[0023] 附图说明

[0024] 图 1 是图示了根据本发明示范实施例的用于编码多对象音频信号的设备的图。

[0025] 图 2 是描绘了图 1 中示出的单声道缩混器的图。

[0026] 图 3 是示出了图 1 的立体声声道缩混器的图。

[0027] 图 4 是图 1 的多声道缩混器的图。

[0028] 图 5 是图示了图 1 的第二缩混器的图。

[0029] 图 6 是示出了从图 1 的补充信息编码器生成的补充信息比特流的结构的图。

[0030] 图 7 是图示了图 6 中示出的补充信息比特流的结构的详细图。

[0031] 图 8 是图示了根据本发明另一实施例的图 6 中示出的补充信息比特流的结构的详细图。

[0032] 图 9 是图示了根据本发明实施例的用于解码多对象音频信号的设备的框图。

[0033] 图 10 是图示了根据本发明另一实施例的用于解码多对象音频信号的设备的框图。

[0034] 图 11 是根据本发明实施例的用于使用图 1 的设备来编码多对象音频信号的方法的流程图。

[0035] 图 12 是根据本发明实施例的用于使用图 9 的设备来解码多对象音频信号的方法的流程图。

[0036] 图 13 是根据本发明另一实施例的用于使用图 10 的设备来解码多对象音频信号的方法的流程图。

具体实施方式

[0037] 通过下面阐明的参考附图对实施例进行的以下描述，本发明的优点、特征和方面将变得清楚。

[0038] 图 1 是图示了根据本发明实施例的用于编码多对象音频信号的设备的图。例如，根据本实施例的设备接收多声道音频对象，例如单声道音频对象、立体声声道音频对象、以及 5.1 声道音频对象。

[0039] 如图 1 中所示，根据本实施例的多对象音频编码设备包括第一缩混器 101、第二缩混器 103、音频编码器 105、和补充信息编码器 107、以及多路复用器 109。

[0040] 第一缩混器 101 包括单声道缩混器 111、立体声声道缩混器 113、和多声道缩混器 115。

[0041] 第一缩混器 101 使用输入的音频对象的报头信息而将输入的各种声道多对象音频信号标识为单声道音频对象、立体声声道音频对象、以及多声道音频信号。然后，第一缩混器 101 通过对应声道来对标识的音频信号进行分组。所以，通过声道对多对象音频信号

的不同声道进行分组，而分组后的音频对象由对应缩混器 111、113 和 115 进行缩混。

[0042] 第一缩混器 101 还从输入的音频对象中提取缩混后的音频信号和包括空间线索的补充信息。即，声源由同一声道分组并被输入到第一缩混器 101。单声道缩混器 111 从单声道音频对象中提取缩混后的信号和包括空间线索的补充信息，而立体声声道缩混器 113 从输入的立体声声道音频对象中提取缩混后的信号和包括空间线索的补充信息。多声道缩混器 115 从输入的多声道音频对象（例如，5.1 声道）中提取缩混后的信号和具有空间线索的补充信息。

[0043] 音频编码器 105 对从第二缩混器 103 输出的第二缩混后的信号进行编码。

[0044] 补充编码器 107 使用从第一缩混器 101 输出的补充信息和从第二缩混器 103 输出的补充信息来生成补充信息比特流。这里，将参考图 6 来描述补充比特流中包括的信息。

[0045] 多路复用器 109 通过对来自音频编码器 105 的编码后的信号和从补充编码器 107 生成的补充比特流进行多路复用，而生成要传送到解码设备的比特 流。

[0046] 从第一缩混器 101 输出的第一缩混后的信号是立体声信号或单声道信号。即，从单声道缩混器 111 输出的缩混后的信号是单声道信号，而从其余混音器 113 和 115 输出的缩混后的信号是单声道信号或立体声信号。

[0047] 第二缩混器 103 对从第一缩混器 101 输出的第一缩混后的信号进行缩混，并输出第二缩混后的信号。第二缩混器 103 提取在第二缩混过程中分析的包括空间线索的补充信息。第二缩混后的信号根据模式是单声道信号或立体声信号。

[0048] 补充信息包括用于恢复和控制空间线索和音频信号的报头信息。将参考图 6 来描述补充信息。

[0049] 图 2 是描绘了图 1 中示出的单声道缩混器的图。例如，单声道缩混器 111 接收 N 个单声道音频对象 m1 到 mN。

[0050] 如图 2 中所示，单声道缩混器 111 包括层叠 (cascade) 结构的第一基本缩混器 201a 到 201d。

[0051] 根据单声道音频对象的数目来判断单声道缩混器 111 中包括的第一基本缩混器 201a 到 201b 的数目。即，如果单声道音频对象为 N，则第一基本缩混器 201 的数目为 N-1。如果单声道音频对象为 1，则旁路 (bypassed) 输入信号，而没有基本缩混器。

[0052] 在本实施例中，可基于层叠方法使用一个第一基本缩混器 N-1 次。

[0053] 基本上，第一基本缩混器对两个输入信号进行缩混，生成一个缩混后的单声道信号，并对于该输入信号提取包括空间线索的补充信息。第 1 个第一基本缩混器 201a 生成缩混后的单声道信号，并使用输入到单声道缩混器 111 的两个单声道音频对象来提取包括空间线索的补充信息。第 2 个第一基本缩混器 201b 生成缩混后的单声道信号，并使用从第 1 个第一基本缩混器 201a 输出的缩混后的单声道信号和输入到单声道缩混器 111 的单声道音频对象来提取包括空间线索的补充信息。第 (N-1) 个第一基本缩混器生成缩混后的单声道信号，并使用从第 (N-2) 个基本缩混器（未示出）输出的缩混后的单声道信号和输入到单声道缩混器 111 的单声道音频对象来提取包括空间线索的补充信息。

[0054] 空间线索是用于编码和解码音频信号的信息。空间线索是从频域提取的，并包括有关输入到第一基本缩混器 201 的两个信号之间的幅度差、延迟差、 和相关性的信息。例如，根据本实施例的空间线索包括声道声级差 (CLD)、声道间声级差 (ICLD)、声道间时间差

(ICTD)、声道间相关性 (ICC)、以及音频信号之间的虚拟源位置信息，表示音频信号的功率增益信息。然而，本发明不限于此。

[0055] 补充信息包括用于恢复和控制空间线索和音频信号的报头信息。将参考图 6 来描述补充信息。

[0056] 图 3 是示出了图 1 的立体声声道缩混器的图。例如，该立体声声道缩混器接收 M 个左信号 SL1 到 SLM 以及 M 个右信号 SR1 到 SRM 作为立体声音频对象。

[0057] 输入到立体声声道缩混器 113 的立体声音频对象被划分为左立体声信号和右立体声信号，并且所划分的信号被再次分组。

[0058] 如图 3 中所示，立体声声道缩混器 113 包括多个第一基本缩混器 201。立体声声道缩混器 113 需要 2*(M-1) 个第一基本缩混器 201，以对 M 个左信号和 M 个右信号进行缩混。这里，可在另一实施例中使用一个第一基本缩混器 2*(M-1) 次。

[0059] 如图 3 中所示，用于分析 M 个左信号的 (M-1) 个第一基本缩混器 201la 到 201le 通过分析输入的信号而生成一个混音后的左信号，并提取包括空间线索的补充信息。

[0060] 如图 3 中所示，用于分析 M 个右信号的 (M-1) 个第一基本缩混器 201ra 到 201re 通过分析输入的信号而生成一个混音后的右信号，并提取包括空间线索的补充信息。

[0061] 如图 3 中所示，如果立体声音频对象为 1，则可旁路输入的左信号和右信号。

[0062] 立体声声道缩混器 113 输出立体声缩混信号，并通过生成缩混后的左信号和缩混后的右信号而提取包括空间线索的补充信息。

[0063] 补充信息包括用于恢复和控制空间线索和音频信号的报头信息。将参考图 6 来描述补充信息。

[0064] 图 4 是图 1 的多声道缩混器的图。例如，该多声道缩混器接收 P 个 5.1 声道音频对象。

[0065] 如图 4 中所示，多声道缩混器 115 是采用 MPEG 环绕或空间音频编码 (SAC) 的缩混器。多声道缩混器 115 从多声道音频信号中提取包括空间线 索的补充信息，并将音频信号缩混为单声道缩混后音频信号或立体声缩混后音频信号。

[0066] 即，多声道缩混器 115 从 P 个多声道音频对象中提取空间线索，并传送所提取的空间线索。多声道缩混器 115 还将音频信号缩混为单声道信号或立体声信号。一般来说，多声道音频对象为一。

[0067] 图 5 是图示了图 1 的第二缩混器的图。

[0068] 第二缩混器 103 对从第一缩混器 101 输出的信号再次进行缩混，输出立体声缩混信号，并提取包括空间线索的补充信息。

[0069] 如图 5 所示，第二缩混器 103 包括第一基本缩混器 201f 和 201g 以及第二基本缩混器 501。

[0070] 如果来自立体声声道缩混器 113 和多声道缩混器 115 的缩混后信号是立体声信号，则将对应的缩混后立体声信号分组为左信号和右信号，并且第一基本缩混器 201f 和 201g 对分组后的左信号和分组后的右信号进行缩混。从第一基本缩混器 201f 和 201g 输出的缩混后的单声道信号是左信号和右信号的代表性缩混信号。

[0071] 即，第一基本缩混器 201f 对由立体声声道缩混器 113 缩混和输出的左信号和由多声道缩混器 115 缩混和输出的左信号再次进行缩混，并输出一个缩混后的左信号作为代表

性左信号。然后，第一基本缩混器 201f 提取补充信息。

[0072] 第一基本缩混器 201g 对由立体声声道缩混器 113 缩混和输出的右信号和由多声道缩混器 115 缩混和输出的右信号再次进行缩混，并输出一个代表性右信号。然后，第一基本缩混器 201g 提取补充信息。

[0073] 如图 2 中所示，根据另一实施例，可使用一个第一基本缩混器两次。

[0074] 第二基本缩混器 501 对从单声道缩混器 111 输出的缩混后的单声道信号和从第一基本缩混器 201f 和 201g 输出的左代表性缩混信号和右代表性缩混信号进行缩混，并输出整个缩混后的左信号和右信号。然后，第二基本缩混器 501 提取包括空间线索的补充信息。

[0075] 补充信息包括用于恢复和控制空间线索和音频信号的报头信息。稍后将参考图 6 描述补充信息。

[0076] 第一基本缩混器 201 和第二基本缩混器 501 基于以下方程 1 和方程 2 对输入音频信号进行缩混。

[0077] $$\begin{bmatrix} w_b^{11} & w_b^{12} \end{bmatrix} \begin{bmatrix} s_b^1(f) \\ s_b^2(f) \end{bmatrix}$$ 方程 1

[0078] $$\begin{bmatrix} w_b^{11} & w_b^{12} & w_b^{13} \\ w_b^{21} & w_b^{22} & w_b^{23} \end{bmatrix} \begin{bmatrix} s_b^1(f) \\ s_b^2(f) \\ s_b^3(f) \end{bmatrix}$$ 方程 2

[0079] 在方程 1 和方程 2 中，w_b^{ij} 是用于控制输入音频信号的缩混声级的加权因子。$s_b^j(f)$ 是作为第一基本缩混器 201 和第二基本缩混器 501 的输入音频信号的单声道信号或立体声左和右信号。下标 b 是表示子带的索引，并且每一加权因子 w_b^{ij} 由子带定义。

[0080] 可根据输入音频对象的表达目的而不同地定义加权因子。例如，$s_b^j(f)$ 的加权因子可被定义为相对大的值，以便将单声道信号 $s_b^j(f)$ 编码为主信号。如果方程 1 中的 $w_b^{11} = 0.7$、$w_b^{12} = 0.3$，则缩混后的信号是 $s_b^k(f) = 0.7 s_b^1(f) + 0.3 s_b^2(f)$。即，$s_b^1(f)$ 被缩混为主信号。

[0081] 可根据缩混后信号的表达目的的限制条件来判断加权因子。该限制条件是用于声音场景 (sound scene) 的限制条件。例如，小提琴和吉他的加权因子被设置为 0.7 和 0.3，以便从缩混后的音频信号按照小提琴和吉他比率 0.7 和 0.3 来重放小提琴和吉他的音频信号。基于来自例如系统或用户的外部装置的输入来判断限制条件信息。

[0082] 其间，加权因子必须被反映为空间线索声级信息。例如，如果将 CLD 用作空间线索，则可象用于方程 1 的方程 3 那样预测空间线索信息。

[0083] $$Level_defference_cue = 10\log_{10}\left(\frac{P(w_b^{11}s_b^1)}{P(w_b^{12}s_b^2)}\right)$$ 方程 3

[0084] 在方程 3 中，P() 是功率运算符，而可使用 $\sum_{b=A_b}^{A_{b+1}} |w_b s_b|^2$ 来计算信号功率之和。A_b

和 A_{b+1} 表示子带的边界。

[0085] 第二基本缩混器 501 提取空间线索 -MPEG 环绕的三到二 (TTT) 盒。

[0086] 图 6 是示出了从图 1 的补充信息编码器生成的补充信息比特流的结构的图。

[0087] 如图 6 中所示，补充比特流包括报头信息和空间线索。

[0088] 报头信息包括用于恢复和再现由各种声道组成的多对象音频信号的信息。报头信息还通过定义音频对象的声道信息和对应音频对象的 ID 来提供用于单声道、立体声、多声道音频对象的解码信息。例如，可定义分类 ID 和每一对象的信息，以标识编码后的预定音频对象是单声道音频信号还是立体声音频信号。在实施例中，报头信息包括空间音频编码 (SAC) 报头信息、音频对象信息、和预置信息。

[0089] 在实施例中，SAC 报头信息是在基于空间线索和时隙信息编码音频信号的过程中生成的信息。当第一和第二缩混器 101 和 103 提取补充信息时，第一和第二缩混器 101 和 103 提取 SAC 报头信息。

[0090] 在实施例中，音频对象信息包括用于标识缩混后的音频对象是单声道、立体声还是多声道音频对象的信息和对象 ID 信息。例如，音频对象信息包括有关每一声道的音频对象的数目的信息（单声道音频对象数目、立体声音频对象数目、和多声道音频对象数目）和每一声道的音频对象的索引信息，该索引信息包括音频对象是单声道、立体声还是多声道的 ID 和信息。

[0091] 在本实施例中，预置信息是报头信息的补充信息，并包括所定义的每一对象的控制信息。

[0092] 例如，预置信息包括预置模式信息和预置模式支持信息。预置模式信息例如包括卡拉 OK 模式、诸如提取吉他演奏音频对象和提取钢琴演奏音频对象的独奏对象提取模式、偏好渲染信息、和重放模式设置信息。

[0093] 例如，预置模式支持信息包括用于支持卡拉 OK 模式的嗓音索引信息、用于支持独奏对象提取模式的对应对象索引信息、用于支持偏好渲染的诸如旋转、提升、和速度的用于每一对象的渲染信息、以及用于支持基本立体声和多声道重放模式设置的用于每一音频对象的最佳渲染信息。

[0094] 而且，补充信息中包括的空间线索包括输入的多对象音频信号的每一对象的空间线索信息。

[0095] 可根据设计者的选择按照各种方式来形成该补充信息的格式。

[0096] 图 7 是图示了图 6 中示出的补充信息比特流的结构的详细图。即，图 7 示出了用于由单声道和立体声声道组成的多对象音频信号的补充信息。

[0097] 如图 7 中所示，报头信息包括有关每一声道的音频对象的数目的信息，例如单声道音频对象的数目和立体声音频对象的数目。报头信息还包括有关 每一声道的音频对象的索引信息，包括有关 ID 的信息和有关音频对象是单声道、立体声、还是多声道的信息。而且，补充比特流包括空间线索。作为示例，CDL 或 ICC 被用作图 7 中示出的实施例中的空间线索的示例。

[0098] 如图 7 中所示，补充信息包括与单声道和立体声对象中的每一个对应的例如 CLD 或 ICC 的空间线索。即，与输入音频对象对应的空间线索信息包括所有补充信息。

[0099] 图 8 是图示了根据本发明另一实施例的图 6 中示出的补充信息比特流的结构的详

细图。即，图8示出了用于由单声道、立体声、和多声道组成的多对象音频信号的补充信息。

[0100] 如图 8 中所示，报头信息包括有关每一声道的音频对象的数目的信息，例如单声道音频对象的数目、立体声音频对象的数目、和多声道音频对象的数目。报头信息还包括每一声道的音频对象的索引信息，诸如 ID 以及音频对象是单声道、立体声、还是多声道。而且，补充比特流包括空间线索。作为空间线索的示例，在图 8 的示例中使用 CLD 和 ICC。

[0101] 通过对多声道对象的空间线索与单声道和立体声对象的空间线索进行层叠多路复用，可将多声道对象的空间线索表达为一个补充比特流。由单声道缩混器 111、立体声声道缩混器 113、和第二缩混器 103 提取的空间线索是用于图 8 的单声道和立体声音频对象的空间线索。而且，图 8 的多声道音频对象的空间线索是由多声道缩混器 115 提取的空间线索。

[0102] 图 9 是图示了根据本发明实施例的用于解码多对象音频信号的设备的框图。

[0103] 根据本实施例的多对象音频信号解码设备通过从图 1 中示出的多对象音频信号编码设备生成的音频比特流中提取空间线索信息并使用提取的空间线索预测每一声道信息，来恢复由各种声道组成的多对象音频信号，该多对象音频信号是包括单声道音频对象、立体声音频对象、和多声道音频对象的音频信号。

[0104] 如图 9 中所示，根据本实施例的多对象音频信号解码设备包括解多路复用器 (DEMUX)901、音频解码器 903、补充信息分析器 905、音频对象提取器 907、和渲染处理器 909。

[0105] 例如，解多路复用器 901 从图 1 的多对象音频信号编码设备生成的音频比特流中分离出音频信息比特流和补充信息比特流。

[0106] 音频解码器 903 从来自解多路复用器 901 的分离的音频信息比特流中恢复缩混后的音频信号。

[0107] 补充信息分析器 905 从来自解多路复用器 901 的补充比特流中提取包括每一音频对象的空间线索信息的补充信息。

[0108] 音频对象提取器 907 使用来自补充信息分析器 905 的提取的补充信息的报头信息，来从缩混后的音频信号恢复每一对象的音频信号。由于报头信息包括诸如单声道音频对象的数目、立体声音频对象的数目、和多声道音频对象的数目的有关每一声道的音频对象的数目的信息以及诸如 ID 和音频对象是单声道音频对象、立体声音频对象、还是多声道音频对象的每一音频对象的索引信息，所以音频对象提取器 907 可基于从补充信息分析器 905 提取的补充信息的报头信息和空间线索信息，来从音频解码器 903 输出的缩混后的音频信号中恢复每一对象的音频信号。

[0109] 渲染处理器 909 对于从音频对象提取器 907 输出的所恢复的音频对象中的每一个从外部装置接收诸如空间音频对象的位置和尺寸的渲染控制信息和诸如 5.1 或 7.1 声道或立体声的输出声道控制信息。基于该渲染控制信息和该输出声道控制信息，渲染处理器 909 安排所恢复的每一对象的音频信号并输出该音频信号。

[0110] 图 10 是图示了根据本发明另一实施例的用于解码多对象音频信号的设备的框图。与用于渲染根据每一对象恢复的音频信号的图 9 的解码设备不同，图 10 中所示的根据另一实施例的多对象音频信号解码设备通过控制补充信息并根据控制的补充信息渲染音频对象，来恢复音频信号。

[0111] 如图 10 中所示，根据另一实施例的多对象音频信号解码设备包括解多路复用器 901、音频解码器 903、补充信息分析器 905、补充信息控制器 1001、和 SAC 解码器 1003。

[0112] 图 10 的解多路复用器 901、音频解码器 903、和补充信息分析器 905 与图 9 的解多路复用器 901、音频解码器、和补充信息分析器 905 相同。

[0113] 补充信息控制器 1001 对于从音频解码器 903 恢复的缩混后的音频信号从外部装置接收诸如空间音频对象的位置和尺寸的渲染控制信息和诸如 5.1 或 7.1 声道和立体声的输出声道控制信息，并根据外部输入信号控制从补充信息分析器 905 提取的诸如每一音频对象的信号幅度和相关性信息的补充信息。

[0114] SAC 解码器 1003 使用来自补充信息控制器 1001 的受控制的补充信息，来从音频解码器 903 恢复的缩混后的音频信号中恢复多声道多对象音频信号。SAC 解码器 1003 使用来自补充信息控制器 1001 的受控制的补充信息的报头信息，来从缩混后的音频信号中恢复每一对象的音频信号。由于报头信息包括诸如单声道音频对象的数目、立体声音频对象的数目、和多声道音频对象的数目的有关每一声道的音频对象的数目的信息以及诸如 ID 和音频对象是单声道音频对象、立体声音频对象、还是多声道音频对象的每一音频对象的索引信息，所以 SAC 解码器 103 可基于从补充信息控制器 1001 控制的补充信息的报头信息和空间线索信息，来从音频解码器 903 输出的缩混后的音频信号中恢复每一对象的音频信号。

[0115] 图 11 是根据本发明实施例的用于使用图 1 的设备来编码多对象音频信号的方法的流程图。

[0116] 参考图 11，在步骤 S1101 中，输入的各种声道的多对象音频信号基于输入音频对象的报头信息被分类为单声道音频信号、立体声音频信号、和多声道音频信号，并通过每一声道分组。

[0117] 在步骤 S1103 中，对同一声道所分组的声源进行缩混，并提取包括空间线索的补充信息。即，从输入的单声道音频对象中提取缩混后的信号和包括空间线索的补充信息，从输入的立体声音频对象中提取缩混后的信号和包括空间线索的补充信息，并从例如 5.1 声道的输入的多声道音频对象中提取缩混后的信号和包括空间线索的补充信息。

[0118] 在步骤 S1103 中输出的第一缩混后的信号是立体声信号或单声道信号。即，从输入的单声道音频对象输出的缩混后的信号是单声道信号，而从输入的立体声音频对象或输入的多声道音频对象输出的缩混后的信号是单声道信号或立体声信号。

[0119] 然后，在步骤 S1105 中，第一缩混后的信号被再次缩混，并且提取包括空间线索的补充信息。这里，第二缩混后的信号根据模式可以是单声道信号或立体声信号。

[0120] 然后，在步骤 S1107 中，对步骤 S1105 中输出的第二缩混后的信号进行编码。

[0121] 在步骤 S1109 中，使用在步骤 S1103 中输出的补充信息和在步骤 S1105 中输出的补充信息，来生成补充信息比特流。

[0122] 在步骤 S1111 中，通过对从步骤 S1107 生成的补充信息比特流进行多路 复用，而生成要传送到解码设备的比特流。

[0123] 图 12 是根据本发明实施例的用于使用图 9 的设备来解码多对象音频信号的方法的流程图。

[0124] 参考图 12，在步骤 S1201 中，从步骤 S1111 生成的音频比特流中分离出音频信息比

特流和补充信息比特流。

[0125] 在步骤 S1203 中，从分离的音频信息比特流中恢复缩混后的音频信号。

[0126] 在步骤 S1205 中，从分离的比特流中提取包括每一音频对象的空间线索信息的补充信息。

[0127] 在步骤 S1207 中，使用提取的补充信息的报头信息来从缩混后的音频信号中恢复每一对象的音频信号。由于报头信息包括诸如单声道音频对象的数目、立体声音频对象的数目、和多声道音频对象的数目的有关每一声道的音频对象的数目的信息以及诸如 ID 和音频对象是单声道音频对象、立体声音频对象、还是多声道音频对象的每一音频对象的索引信息，所以可基于在步骤 S1205 中提取的所提取的补充信息的报头信息和空间线索信息，来从步骤 S1203 中输出的缩混后的音频信号中恢复每一对象的音频信号。

[0128] 在步骤 S1207 中，从外部装置接收例如空间音频对象的位置和尺寸的用于每一恢复的音频对象的渲染控制信息和例如 5.1 或 7.1 声道或立体声的输出声道控制信息，并安排每一恢复的对象的音频信号，并输出多对象音频信号。

[0129] 图 13 是根据本发明另一实施例的用于使用图 10 的设备来解码多对象音频信号的方法的流程图。

[0130] 在步骤 S1301 中，从步骤 S1111 中生成的音频比特流中分离音频信息比特流和补充信息比特流。

[0131] 在步骤 S1303 中，从分离的音频信息比特流中恢复缩混后的音频信号。

[0132] 在步骤 S1305 中，从分离的补充比特流中提取包括每一音频对象的空间线索信息的补充信息。

[0133] 在步骤 S1307 中，从外部设备接收例如空间音频对象的位置和尺寸的用于每一恢复的音频对象的渲染控制信息和例如 5.1 或 7.1 声道和立体声的输出声道控制信息，并根据外部输入信号控制从步骤 S1305 提取的补充信息，其中所提取的补充信息例如包括有关每一音频对象的信号幅度的信息和相关性信息。

[0134] 在步骤 S1309 中，使用受控制的补充信息来从来自步骤 S1303 的缩混后 的音频信号中恢复各种声道的多对象音频信号。使用受控制的补充信息的报头信息来从缩混后的音频信号中恢复每一对象的音频信号。由于报头信息包括诸如单声道音频对象的数目、立体声音频对象的数目、和多声道音频对象的数目的有关每一声道的音频对象的数目的信息以及诸如 ID 和音频对象是单声道音频对象、立体声音频对象、还是多声道音频对象的每一音频对象的索引信息，所以可基于来自步骤 S1307 的受控制的补充信息的报头信息和空间线索信息，来从步骤 S1303 输出的缩混后的音频信号中恢复每一对象的音频信号。

[0135] 根据本发明的上述方法可被实施为程序并存储在计算机可读记录介质上。计算机可读记录介质是可存储其后可由计算机系统读取的数据的任何数据储存装置。计算机可读记录介质包括只读存储器 (ROM)、随机存取存储器 (RAM)、CD-ROM、软盘、硬盘和磁光盘。

[0136] 尽管已针对某些优选实施例而描述了本发明，但是本领域技术人员将理解，可进行各种改变和变型，而不脱离以下权利要求中限定的本发明的精神和范围。

[0137] 工业实用性

[0138] 根据本发明实施例的用于编码和解码多对象音频信号的设备和方法通过有效编码和解码由各种声道组成的各种对象的音频内容，使得用户能够根据需求来主动消费音频内容。

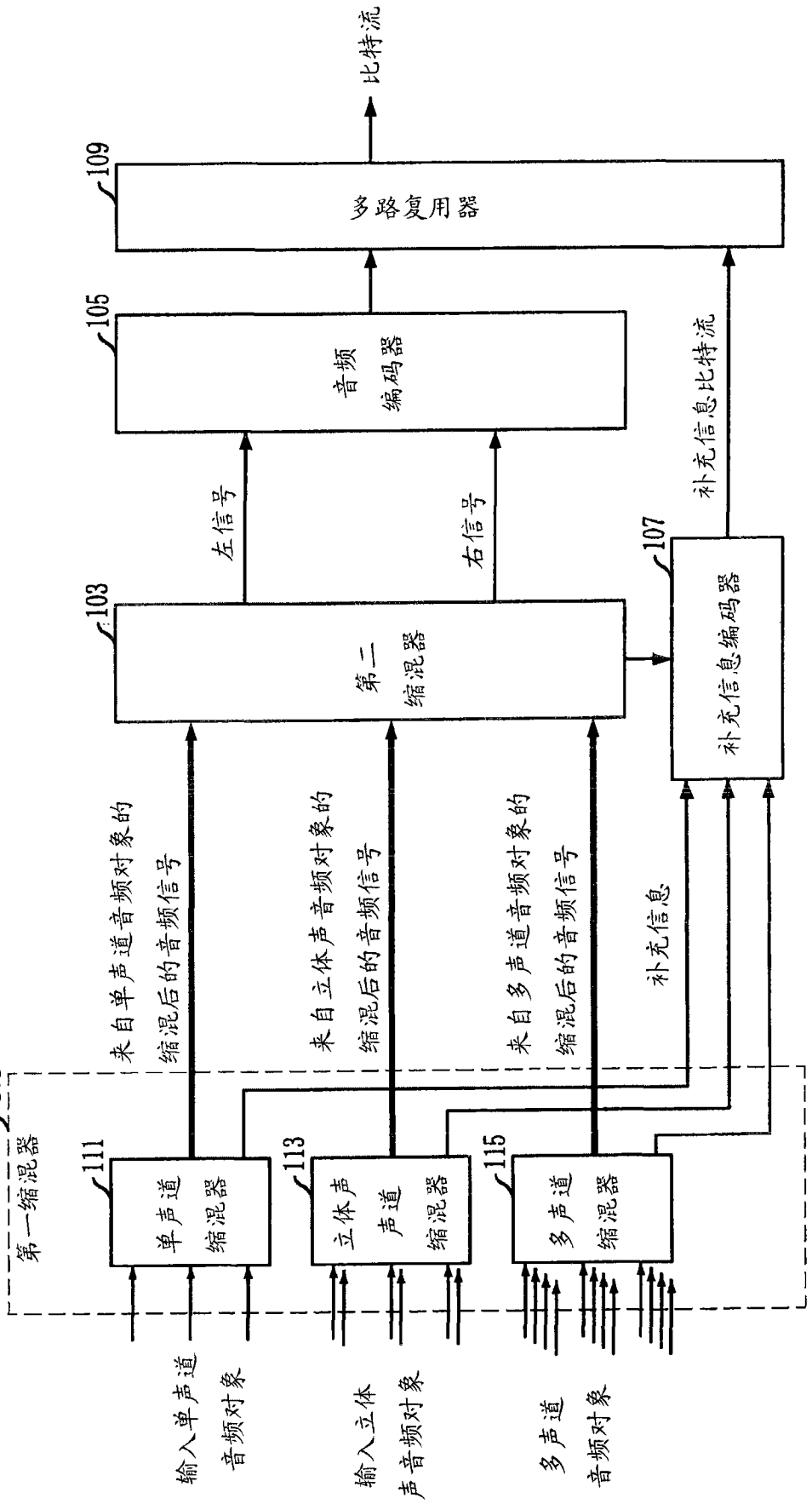

比特流
109
多路复用器
105
音频
编码器
左信号
右信号
103
第二
缩混器
107
补充信息编码器
补充信息比特流
来自单声道音频对象的
缩混后的音频信号
来自立体声音频对象的
缩混后的音频信号
来自多声道音频对象的
缩混后的音频信号
补充信息
101
第一缩混器
111
单声道
缩混器
113
立体声
声道
缩混器
115
多声道
缩混器
输入单声道
音频对象
输入立体
声音频对象
多声道
音频对象

图 1

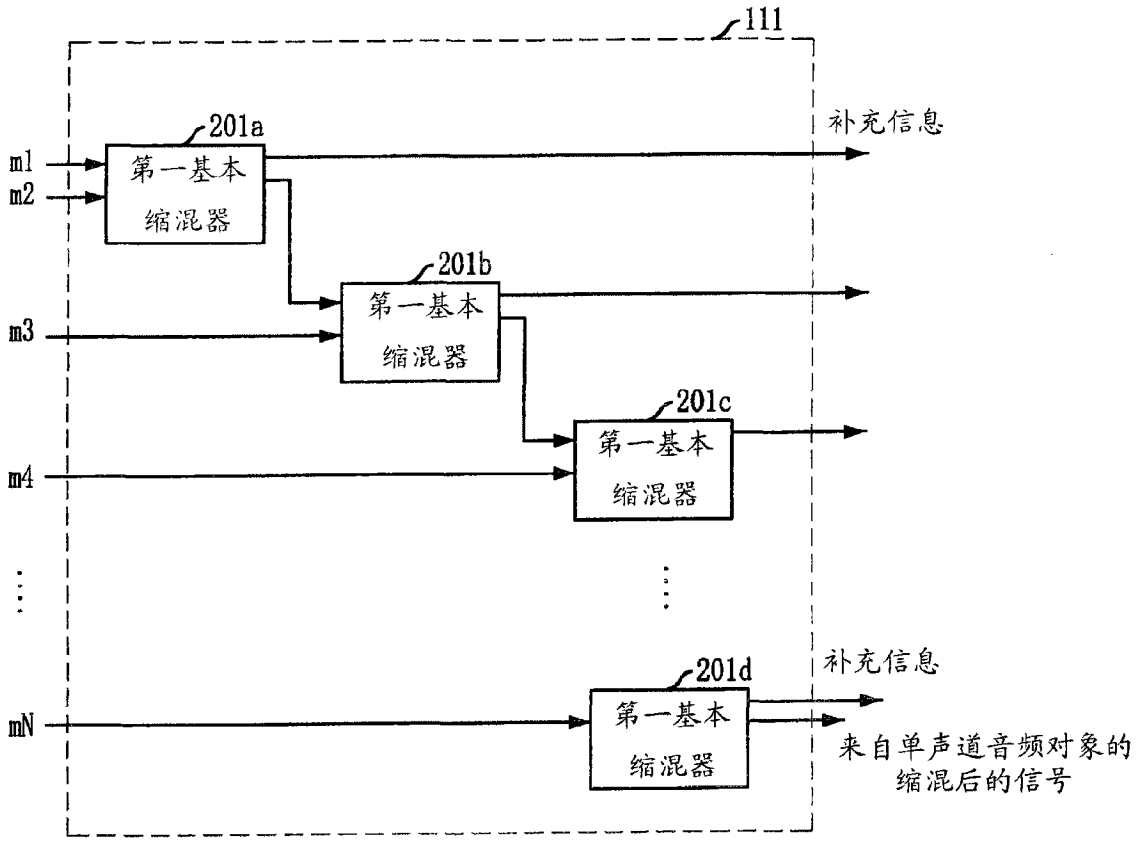

111
201a
第一基本
缩混器
m1
m2
补充信息
201b
第一基本
缩混器
m3
201c
第一基本
缩混器
m4
201d
第一基本
缩混器
mN
补充信息
来自单声道音频对象的
缩混后的信号

图 2

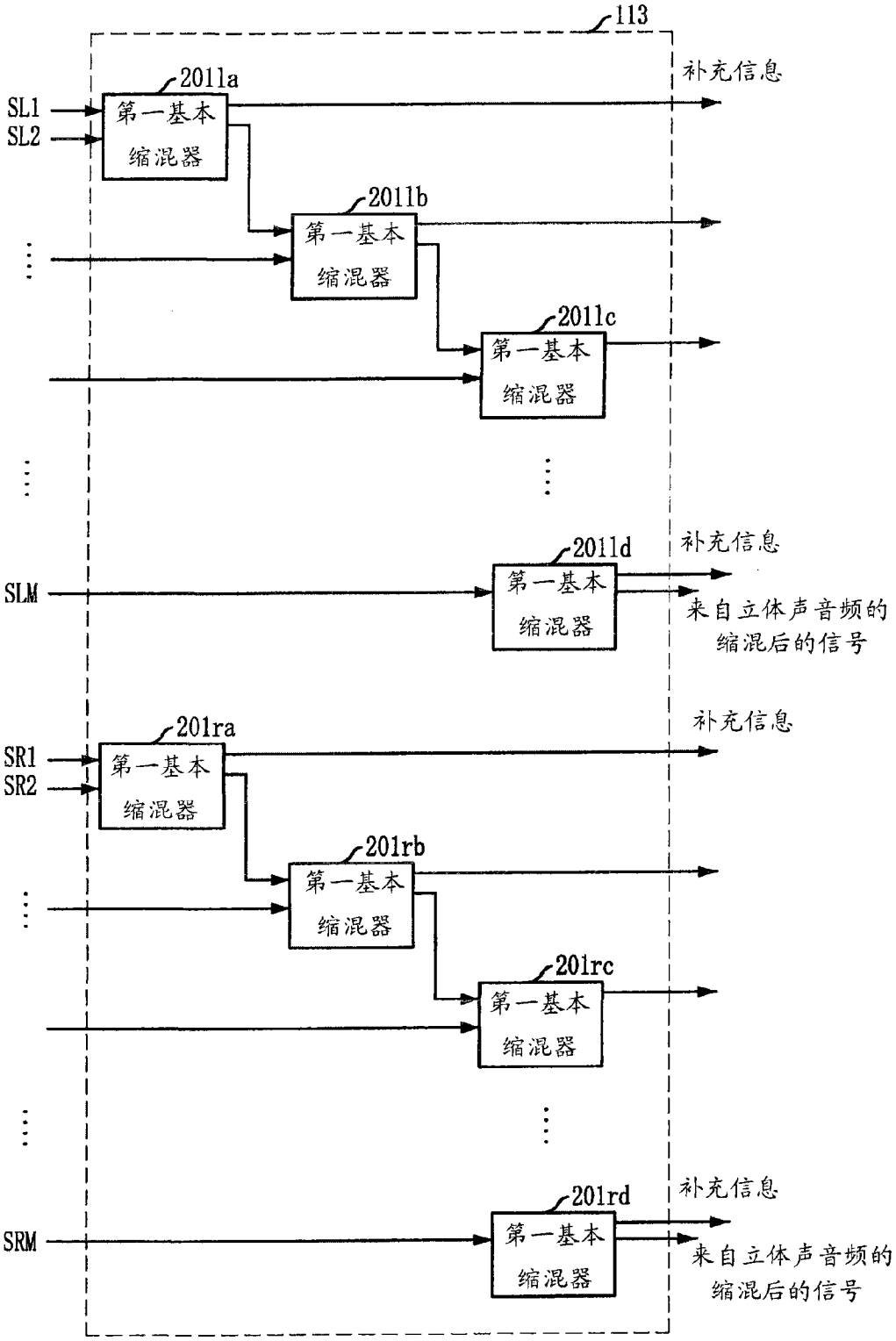

113
2011a
第一基本
缩混器
SL1
SL2
补充信息
2011b
第一基本
缩混器
2011c
第一基本
缩混器
2011d
第一基本
缩混器
SLM
补充信息
来自立体声音频的
缩混后的信号
201ra
第一基本
缩混器
SR1
SR2
补充信息
201rb
第一基本
缩混器
201rc
第一基本
缩混器
201rd
第一基本
缩混器
SRM
补充信息
来自立体声音频的
缩混后的信号

图 3

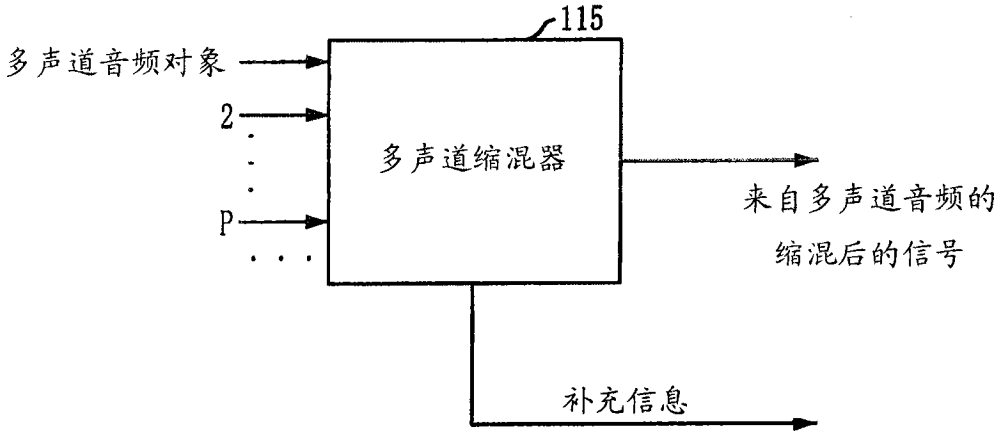

115
多声道音频对象
2
p
多声道缩混器
来自多声道音频的
缩混后的信号
补充信息

图 4

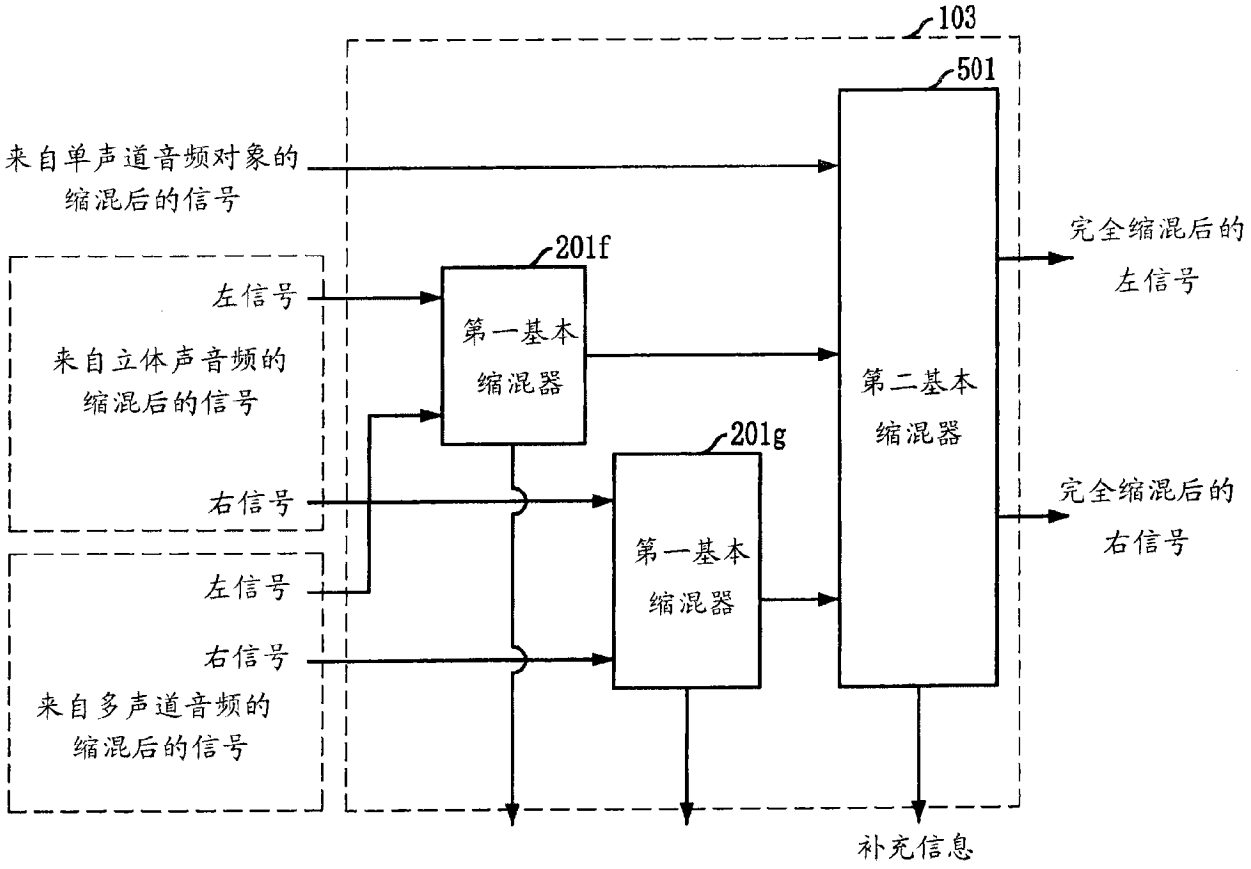

103
501
来自单声道音频对象的
缩混后的信号
201f
左信号
来自立体声音频的
缩混后的信号
第一基本
缩混器
201g
右信号
第一基本
缩混器
第二基本
缩混器
完全缩混后的
左信号
完全缩混后的
右信号
左信号
右信号
来自多声道音频的
缩混后的信号
补充信息

图 5

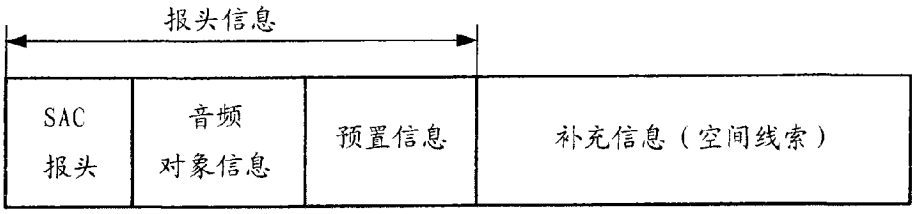


图 6

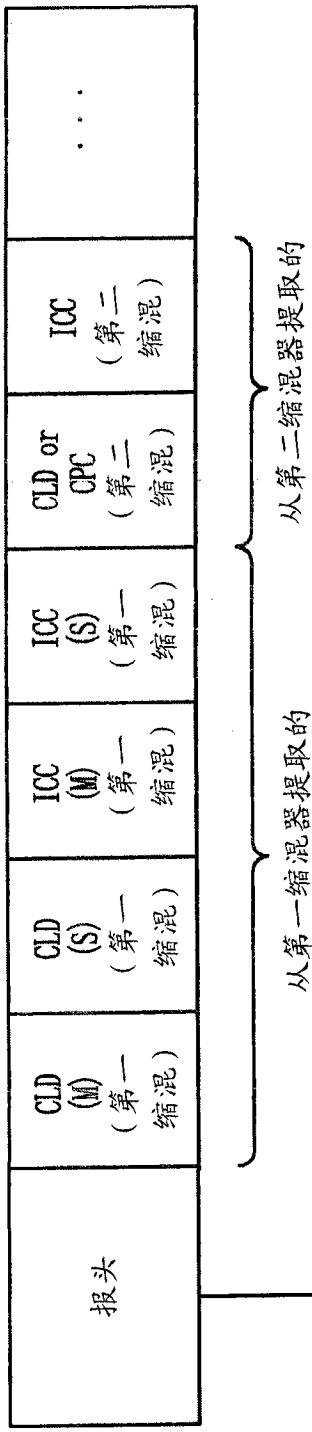

报头
CLD (M) (第一缩混)
CLD (S) (第一缩混)
ICC (M) (第一缩混)
ICC (S) (第一缩混)
CLD or CPC (第二缩混)
ICC (第二缩混)
. . .
从第一缩混器提取的空间线索
从第二缩混器提取的空间线索
声道的音频对象信息（ID、声道信息）
每一声道的音频对象的数目

图 7

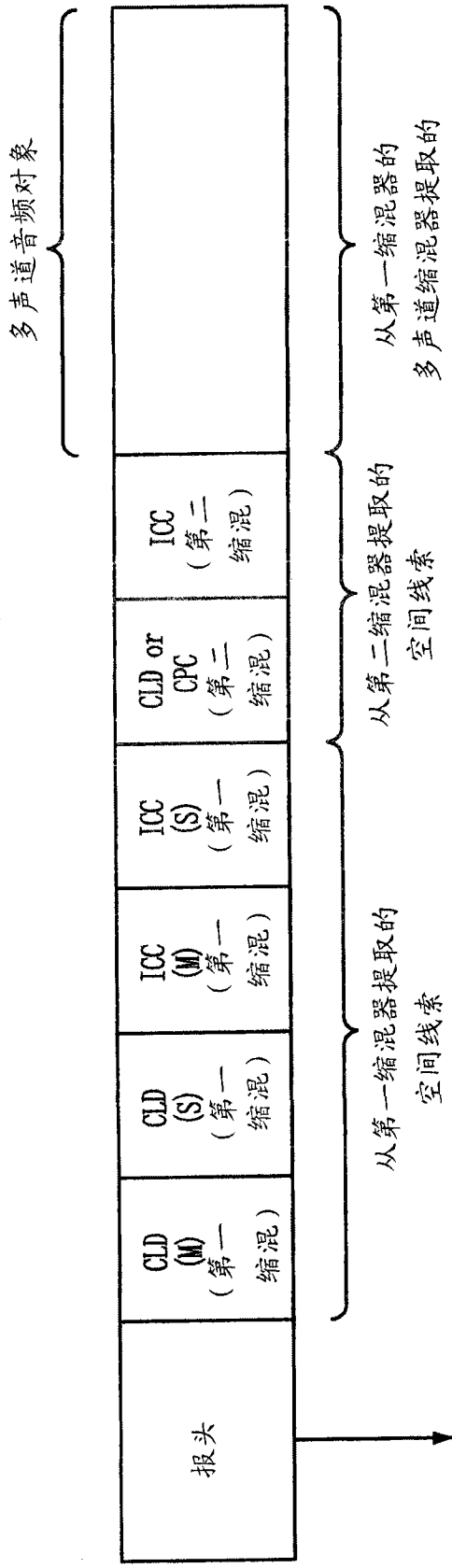

多声道音频对象
报头
CLD (M) (第一缩混)
CLD (S) (第一缩混)
ICC (M) (第一缩混)
ICC (S) (第一缩混)
CLD or CPC (第二缩混)
ICC (第二缩混)
从第一缩混器提取的空间线索
从第二缩混器提取的空间线索
从第一缩混器的多声道缩混器提取的空间线索
声道的音频对象信息（ID、声道信息）
声道的音频对象的数目

图 8

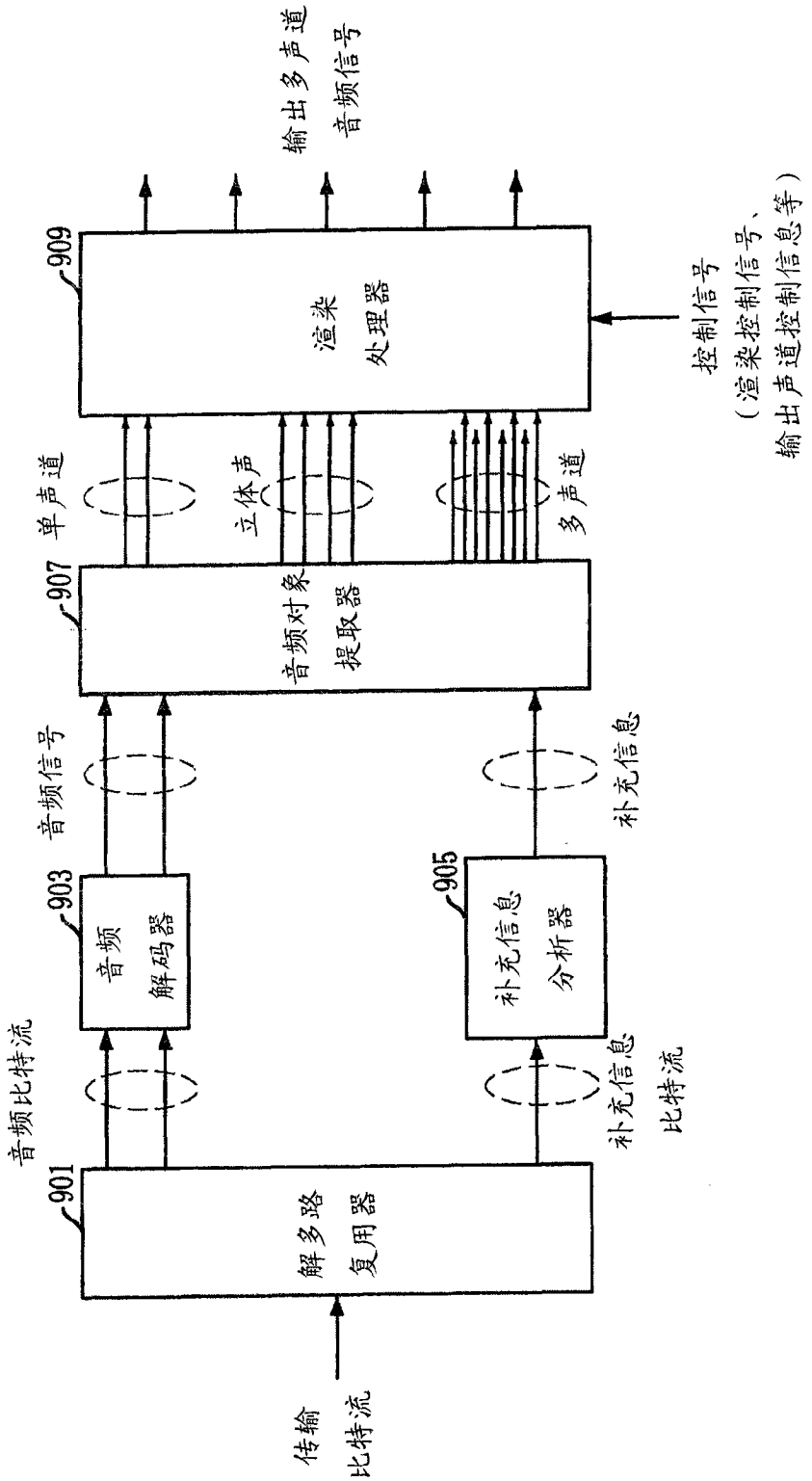

传输比特流
解多路复用器
901
音频比特流
音频解码器
903
音频信号
补充信息分析器
905
补充信息
比特流
补充信息
音频对象提取器
907
单声道
立体声
多声道
渲染处理器
909
控制信号
（渲染控制信号、输出声道控制信息等）
输出多声道音频信号

图 9

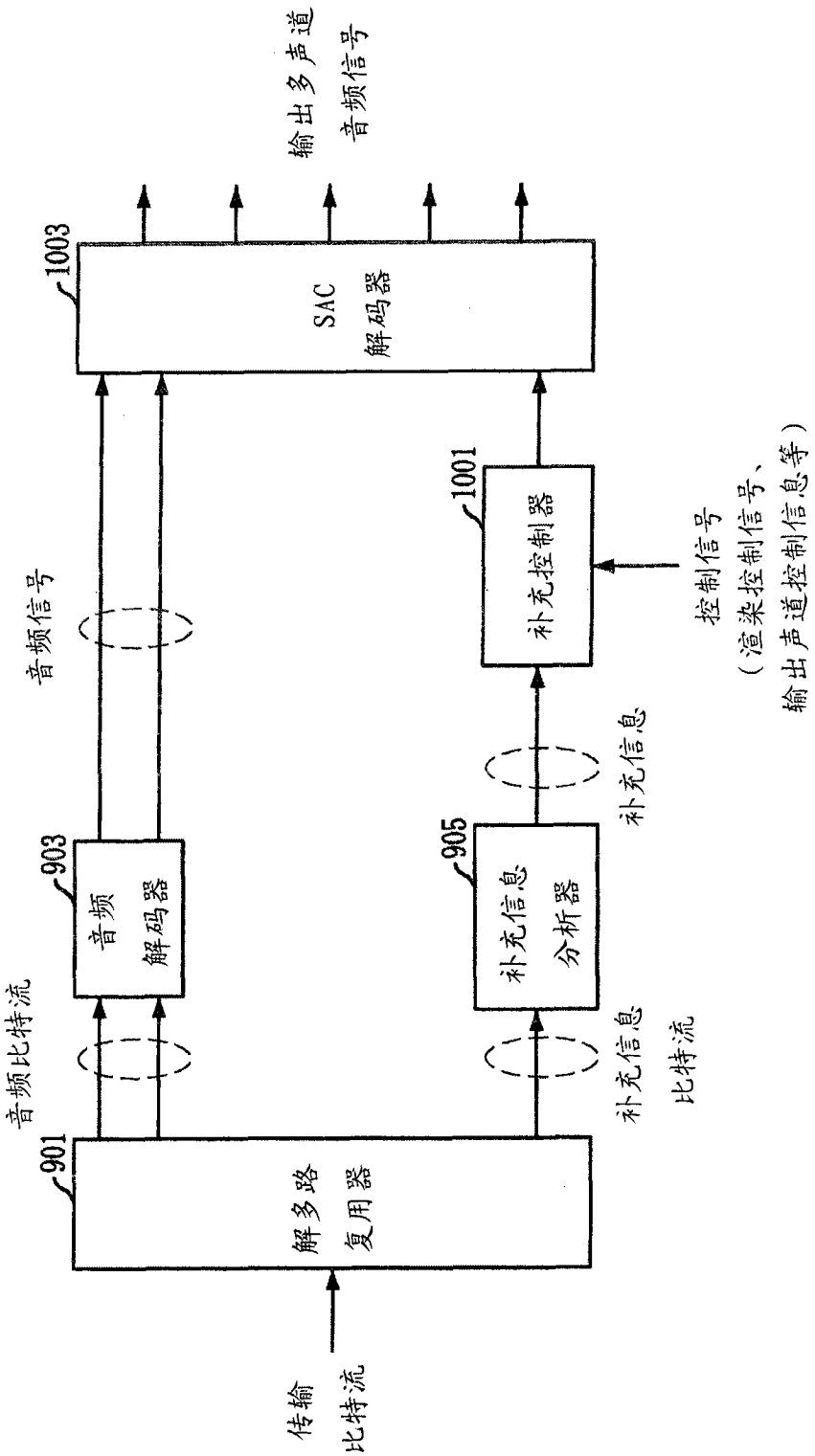

传输
比特流
解多路
复用器
901
音频比特流
音频
解码器
903
音频信号
补充信息
分析器
905
补充信息
比特流
补充控制器
1001
控制信号
(渲染控制信号、
输出声道控制信息等)
SAC
解码器
1003
输出多声道
音频信号

图 10

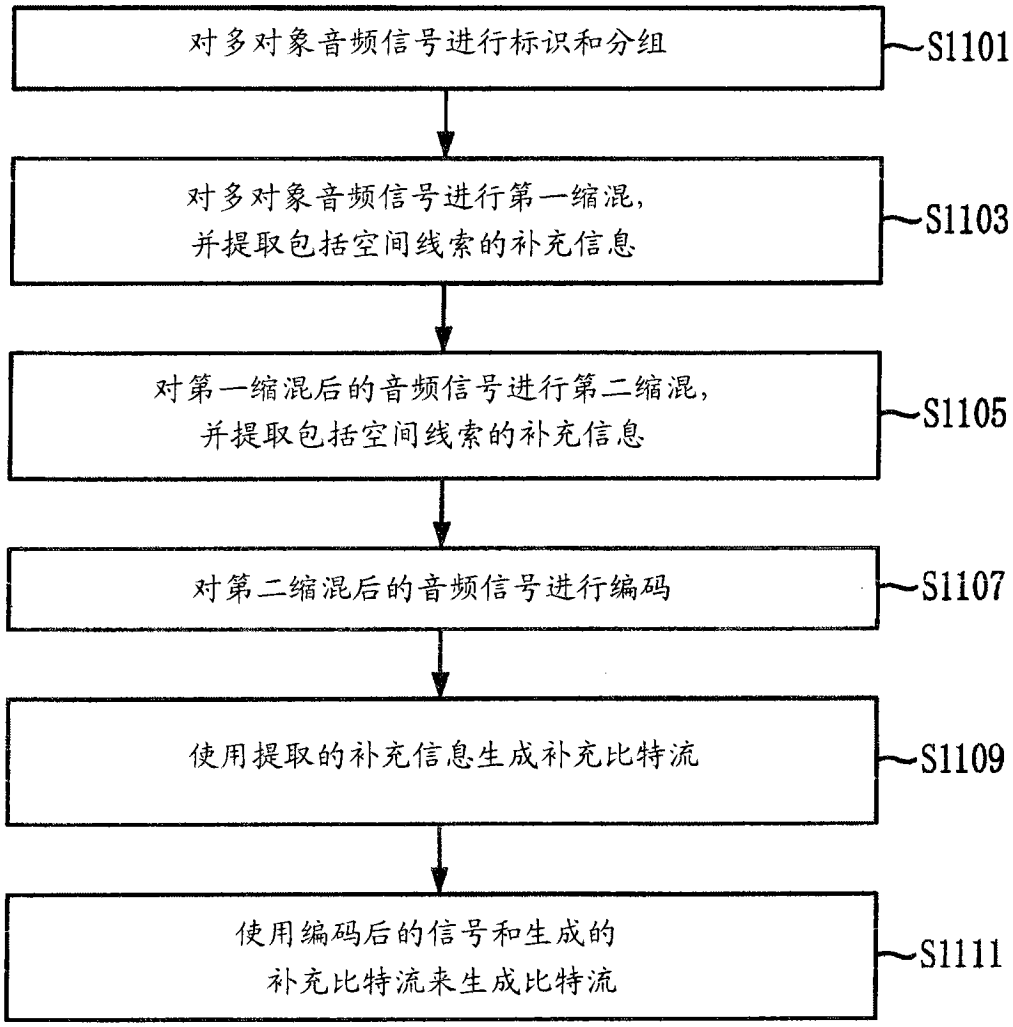

对多对象音频信号进行标识和分组
S1101
对多对象音频信号进行第一缩混，
并提取包括空间线索的补充信息
S1103
对第一缩混后的音频信号进行第二缩混，
并提取包括空间线索的补充信息
S1105
对第二缩混后的音频信号进行编码
S1107
使用提取的补充信息生成补充比特流
S1109
使用编码后的信号和生成的
补充比特流来生成比特流
S1111

图 11

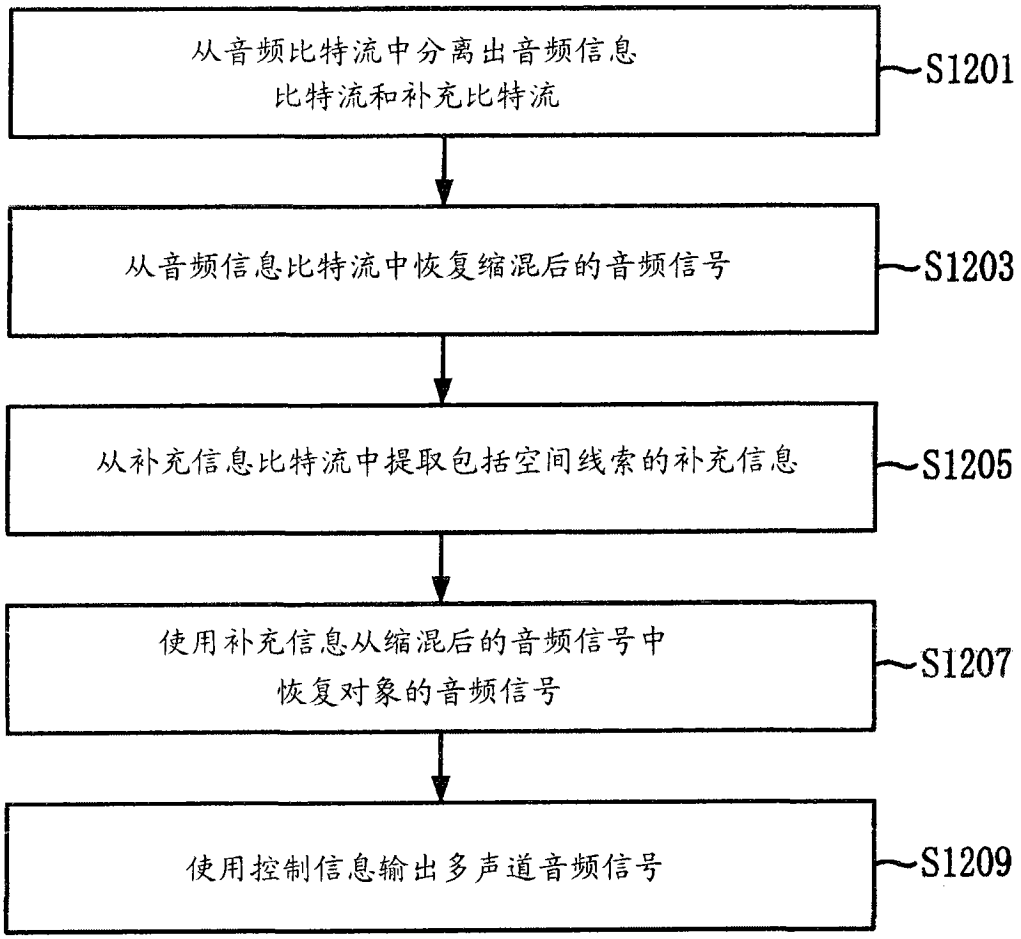

从音频比特流中分离出音频信息
比特流和补充比特流
S1201
从音频信息比特流中恢复缩混后的音频信号
S1203
从补充信息比特流中提取包括空间线索的补充信息
S1205
使用补充信息从缩混后的音频信号中
恢复对象的音频信号
S1207
使用控制信息输出多声道音频信号
S1209

图 12

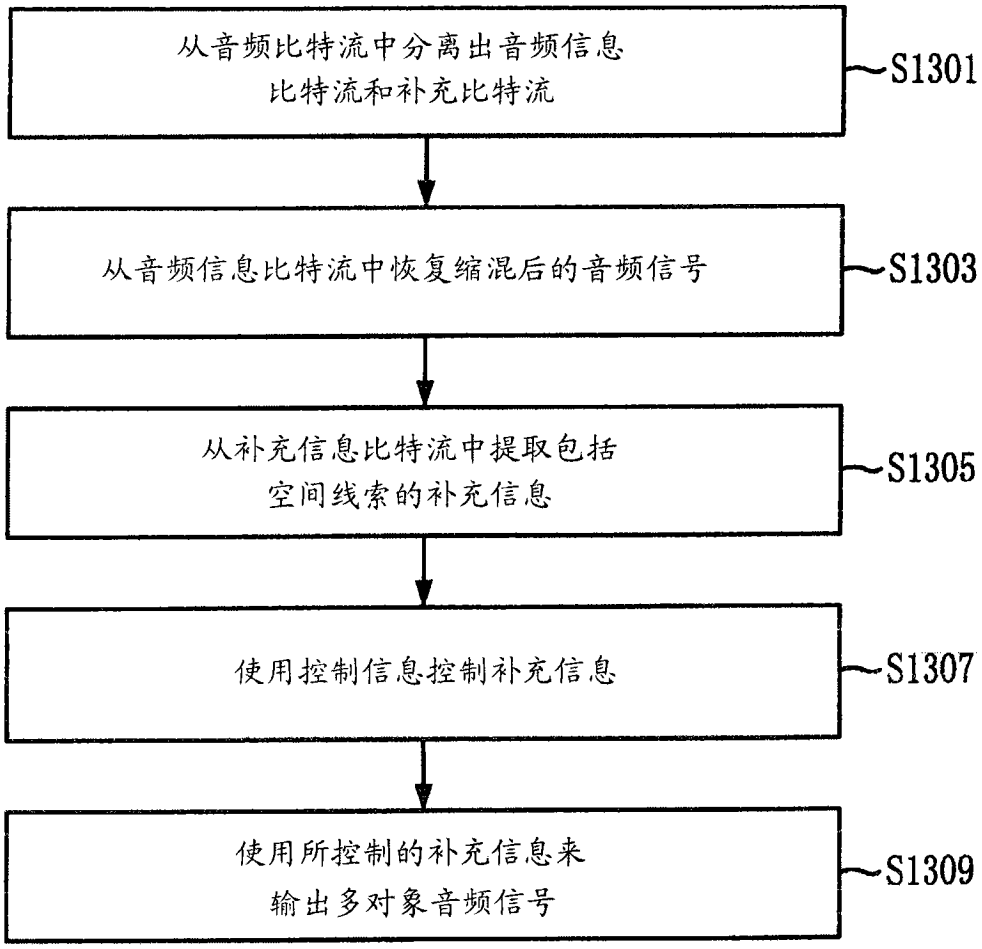

从音频比特流中分离出音频信息
比特流和补充比特流
S1301
从音频信息比特流中恢复缩混后的音频信号
S1303
从补充信息比特流中提取包括
空间线索的补充信息
S1305
使用控制信息控制补充信息
S1307
使用所控制的补充信息来
输出多对象音频信号
S1309

图 13