



(12) 发明专利

(10) 授权公告号 CN 101632118 B

(45) 授权公告日 2013. 06. 05

(21) 申请号 200780051758. 6

代理人 李芳华

(22) 申请日 2007. 12. 27

(51) Int. Cl.

(30) 优先权数据

H04S 7/00 (2006. 01)

10-2006-0135400 2006. 12. 27 KR

10-2007-0003897 2007. 01. 12 KR

10-2007-0007724 2007. 01. 25 KR

(85) PCT申请进入国家阶段日

2009. 08. 25

(56) 对比文件

US 6021386 A, 2000. 02. 01,

WO 2006/103581 A1, 2006. 10. 05,

WO 2006/108573 A1, 2006. 10. 19,

WO 2006/126856 A2, 2006. 11. 30,

(86) PCT申请的申请数据

审查员 游晓梅

PCT/KR2007/006910 2007. 12. 27

(87) PCT申请的公布数据

W02008/078973 EN 2008. 07. 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

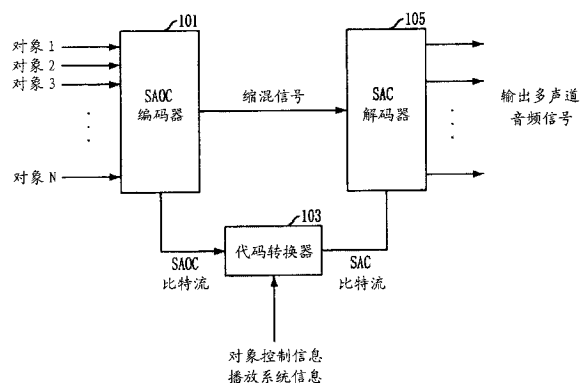
权利要求书1页 说明书14页 附图7页

(54) 发明名称

用于对多对象音频信号进行编码和解码的设备和
方法

(57) 摘要

提供了一种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, 所述设备和方法提供了与传统的空间音频编码 (SAC) 比特流的后向兼容性。所述设备包括: 音频对象编码单元, 用于基于空间线索来编码输入到编码设备的音频对象信号, 并创建用于编码后的音频对象信号的渲染信息, 其中所述渲染信息提供了包括用于音频对象信号的空间线索信息、音频对象信号的声道信息和音频对象信号的标识信息的编码设备, 并且所述编码设备使用在音频信号的编码和解码中。



1. 一种用于编码多对象音频信号的设备,包括:

音频对象编码装置,用于基于空间线索来编码输入到编码设备的音频对象信号,并创建用于编码后的音频对象信号的渲染信息,

其中所述渲染信息包括用于音频对象信号的空间线索信息、音频对象信号的声道信息和音频对象信号的标识信息。

2. 根据权利要求 1 的编码设备,还包括:

音频声道编码装置,用于将输入到编码设备的多声道音频信号变换为音频对象信号,并创建用于多声道音频信号的渲染信息;

预设声场场景创建装置,用于基于输入到编码设备的控制信号来创建包括音频对象信号的声音信息和播放信息的预设信息;和

代表性比特流创建装置,用于创建代表性比特流,该代表性比特流包括从音频对象编码装置输出的渲染信息、从音频声道编码装置输出的渲染信息、和预设信息,

其中从音频声道编码装置输出的渲染信息包括用于多声道音频信号的空间线索信息、多声道音频信号的声道信息、和多声道音频信号的标识信息。

3. 根据权利要求 2 的编码设备,其中所述音频声道编码装置是运动图象专家组 (MPEG) 环绕编码器。

4. 一种用于编码多对象音频信号的方法,包括如下步骤:

基于空间线索来对所输入的音频对象信号进行编码,并创建用于编码后的音频对象信号的渲染信息,

其中所述渲染信息包括用于音频对象信号的空间线索信息、音频对象信号的声道信息和音频对象信号的标识信息。

5. 根据权利要求 4 的编码方法,还包括如下步骤:

将所输入的多声道音频信号变换为音频对象信号,并创建用于多声道音频信号的渲染信息;

基于所输入的控制信号来创建包括音频对象信号的声音信息和播放信息的预设信息;和

创建代表性比特流,该代表性比特流包括从所述对所输入的音频对象信号进行编码的步骤输出的渲染信息、从所述将所输入的多声道音频信号变换为音频对象信号并创建用于多声道音频信号的渲染信息的步骤输出的渲染信息、和预设信息,

其中从所述将所输入的多声道音频信号变换为音频对象信号并创建用于多声道音频信号的渲染信息的步骤输出的渲染信息包括:用于多声道音频信号的空间线索信息、多声道音频信号的声道信息、和多声道音频信号的标识信息。

6. 根据权利要求 5 的编码方法,其中所述将所输入的多声道音频信号变换为音频对象信号并创建用于多声道音频信号的渲染信息的步骤是在运动图象专家组 (MPEG) 环绕编码器中执行的。

用于对多对象音频信号进行编码和解码的设备和方法

技术领域

[0001] 本发明涉及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的设备和方法；且更具体地，涉及一种包括边信息 (side informaion) 比特流转换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的设备和方法，所述边信息比特流转换用于变换边信息比特流、并基于变换后的边信息比特流来恢复具有所期望的输出信号（即，各种声道）的多对象音频信号。

[0002] 具有各种声道的多对象音频信号表示用于针对每个音频对象具有不同声道（例如，单声道、立体声和 5.1 声道）的多对象的音频信号。

[0003] 这项工作受 MIC/IITA 的 IT R&D 规划 [2005-S-403-02, “Development of Super-intelligent Multimedia Anytime-anywhere Realistic TV SmarTV Technology (超智能多媒体任何时间任何地点逼真 TV 智能 TV 技术的开发)”] 支持。

[0004] 背景技术

[0005] 根据传统的音频编码 / 解码技术，用户应该被动地 (inactively) 聆听音频内容。这样，需要开发一种用于对在用于多个音频对象的多声道中的音频信号进行编码和解码的设备和方法，使得可以通过根据用户的需要而控制其每一个具有不同声道的音频对象并以各种方法而组合一个音频内容，来消费各种音频对象。

[0006] 传统的空间音频编码 (SAC) 是一种用于将多声道音频信号代表、传送、和恢复为缩混后的单声道或立体声信号的技术，并且它可以以低比特率来传送高质量的多声道音频信号。

[0007] 然而，由于传统的 SAC 能够对在仅用于一个音频对象的多声道中的信号进行编码和解码，所以它不能编码 / 解码多声道和多对象音频信号，例如，用于多声道（例如，单声道、立体声和 5.1 声道）中的各种对象的音频信号。

[0008] 此外，传统的心理声学编码 (BCC : Binaural Cue Coding) 技术可以编码 / 解码用于多对象的音频信号。然而，由于音频对象的声道限于单声道，所以不能对具有包括单声道的各种声道的多对象音频信号进行编码 / 解码。

[0009] 总之，由于传统的技术仅可以编码 / 解码具有单个声道的多对象音频信号、或者具有多声道的单对象音频信号，所以不能对具有各种声道的多对象音频信号进行编码 / 解码。因此，根据传统的音频编码 / 解码技术，用户应该被动地聆听音频内容。

[0010] 相应地，需要开发一种用于对在用于每个多音频对象的各种声道中的音频信号进行编码和解码的设备和方法，以通过控制根据用户的需要而不同的多声道中的每个音频对象、并根据各种方法而组合一个音频内容来消费各种音频对象。

[0011] 此外，需要一种用于将多对象音频比特流转换为传统的 SAC 比特流并且反之亦然的方法和装置，以提供在多对象音频编码器中创建的边信息比特流与传统的 SAC 编码器 / 解码器的边信息比特流之间的后向兼容性。

[0012] 如上所述，作为用于通过单独控制具有不同声道的多个音频对象并根据各种方法而组合一个音频内容来对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，需要

开发一种可执行比特流转换的多声道和多对象音频编码和解码设备及方法,以提供与传统 SAC 比特流的后向兼容性,并控制具有多声道的每个多音频对象,从而以多样的方法来组合一个音频对象。

发明内容

[0013] 技术问题

[0014] 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、以提供与传统的空间音频编码 (SAC) 比特流的后向兼容性的设备和方法。

[0015] 技术解决方案

[0016] 根据本发明的一方面,提供了一种用于编码多对象音频信号的设备,包括:音频对象编码单元,用于基于空间线索 (spatial cue) 来编码输入到编码设备的音频对象信号,并创建用于编码后的音频对象信号的渲染信息,其中所述渲染信息包括用于音频对象信号的空间线索信息、音频对象信号的声道信息和音频对象信号的标识信息。

[0017] 根据本发明的又一方面,提供了一种用于创建渲染信息来解码多对象音频信号的代码转换设备,包括:第一矩阵单元,用于基于已编码音频对象信号的对象控制信息和播放信息来创建渲染信息,该渲染信息包括用于已编码音频对象信号的功率增益信息和输出位置信息;和渲染单元,用于基于第一矩阵单元所创建的渲染信息和用于从编码设备输入的已编码音频对象信号的渲染信息,来创建用于要从解码设备输出的音频信号的空间线索信息。

[0018] 根据本发明的又一方面,提供了一种用于创建多声道音频信号和用于解码多声道音频信号的渲染信息的代码转换设备,包括:解析单元,用于从用于由编码设备输入的已编码音频信号的渲染信息中分出用于已编码音频对象信号的渲染信息和用于多声道音频信号的渲染信息;第一矩阵单元,用于基于已编码音频对象信号的对象控制信息和播放信息来创建渲染信息,该渲染信息包括用于已编码音频对象信号的功率增益信息和输出位置信息;第二矩阵单元,用于基于由解析单元单独获取的用于已编码多声道音频信号的渲染信息,来创建包括用于多声道音频信号的每个声道的功率增益信息的渲染信息,和渲染单元,用于基于第一矩阵单元所创建的渲染信息、第二矩阵单元所创建的渲染信息、和由解析单元单独获取的用于已编码音频对象信号的渲染信息,来创建用于从解码设备输出的音频信号的空间线索信息。

[0019] 根据本发明的又一方面,提供了一种用于编码多对象音频信号的方法,包括如下步骤:基于空间线索来对所输入的音频对象信号进行编码,并创建用于编码后的音频对象信号的渲染信息,其中所述渲染信息包括用于音频对象信号的空间线索信息、音频对象信号的声道信息和音频对象信号的标识信息。

[0020] 根据本发明的又一方面,提供了一种用于创建渲染信息来解码多对象音频信号的代码转换方法,包括如下步骤:基于已编码音频对象信号的对象控制信息和播放信息来创建渲染信息,该渲染信息包括用于已编码音频对象信号的功率增益信息和输出位置信息;和基于在所述创建渲染信息的步骤中创建的渲染信息和用于在编码之后输入的已编码音频对象信号的渲染信息,来创建用于要在解码之后输出的音频信号的空间线索信息。

[0021] 根据本发明的又一方面,提供了一种用于创建渲染信息以解码多声道音频信号和

多对象音频信号的代码转换方法,包括如下步骤:从用于在编码之后输入的已编码音频信号的渲染信息中分出用于已编码音频对象信号的渲染信息和用于多声道音频信号的渲染信息;基于用于已编码音频对象信号的对象控制信息和播放信息来创建渲染信息,该渲染信息包括用于已编码音频对象信号的功率增益信息和输出位置信息;基于在所述分出渲染信息的步骤中单独获取的用于已编码多声道音频信号的渲染信息,来创建包括用于多声道音频信号的每个声道的功率增益信息的渲染信息,和基于在所述创建包括功率增益信息和输出位置信息的渲染信息的步骤中创建的渲染信息、在所述创建包括用于多声道音频信号的每个声道的功率增益信息的渲染信息的步骤中创建的渲染信息、和在所述分出渲染信息的步骤中单独获取的用于已编码音频对象信号的渲染信息,来创建用于要在解码之后输出的音频信号的空间线索信息。

[0022] 有利效果

[0023] 通过提供一种能够执行边信息比特流转换的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,本发明可通过有效地编码和解码各种声道中的多对象音频内容而根据用户的需要来主动地消费音频内容。此外,本发明可通过提供与传统上使用的比特流的后向兼容性,来提供与传统编码和解码设备的兼容性。

附图说明

[0024] 图 1 是示出了根据本发明实施例的多对象音频编码器和多对象解码器的框图。

[0025] 图 2 是示出了根据本发明实施例的多对象音频编码器和多对象解码器的框图。

[0026] 图 3 是图示了根据本发明实施例的图 2 的代码转换器 (transcoder) 103 的框图。

[0027] 图 4 图示了根据本发明实施例的由图 2 的比特流格式器 (formatter) 205 创建的代表性空间音频对象编码 (SAOC) 比特流。

[0028] 图 5 示出了根据本发明另一实施例的图 2 的代表性 SAOC 比特流。

[0029] 图 6 是示出了根据本发明另一实施例的图 2 的代码转换器 103 的框图。

[0030] 图 7 是示出了在图 2 的多对象音频编码器和解码器中另外包括音频对象去除器 701 的情况的框图。

[0031] 图 8 是示出了用 MPEG 环绕编码器和解码器来替换图 2 的 SAC 编码器 201 和 SAC 解码器 105 的情况的框图。

具体实施方式

[0032] 根据下文中陈述的参考附图的以下实施例描述,本发明的优点、特征和方面将变明显。下文中将参考附图来详细描述本发明的特定实施例。

[0033] 图 1 是示出了根据本发明实施例的多对象音频编码器和多对象解码器的框图。

[0034] 参考图 1,本发明包括:空间音频对象编码器 (SAOC) 101、代码转换器 103、和空间音频编码 (SAC) 105。

[0035] 根据 SAOC 方法,输入到编码器的信号被编码为音频对象。每个音频对象不由解码器恢复并独立地播放。然而,用于音频对象的信息被渲染 (render) 以形成所期望的音频场景,并输出具有各种声道的多对象音频信号。因此,SAC 解码器需要用于渲染被输入以获取所期望的音频场景的用于音频对象的信息的设备。

[0036] SAOC 编码器 101 是基于空间线索的编码器,并将输入音频信号编码为音频对象。所述音频对象是向 SAOC 编码器 101 输入的单声道或立体声信号。

[0037] SAOC 编码器 101 输出来自多于一个输入音频对象的缩混信号,并通过提取空间线索和边信息来创建 SAOC 比特流。所输出的缩混信号是单声道或立体声信号。SAOC 编码器 101 基于“异类布局 SAOC”或者“Faller”技术来分析所输入的音频对象信号。

[0038] 所提取的 SAOC 比特流包括空间线索和边信息,并且所述边信息包括输入音频对象的空间信息。一般基于频率区域子带单位来分析并提取所述空间线索。

[0039] 所述空间线索是在编码和解码音频信号中使用的信息。它提取自频率区域,并包括用于所输入的两个信号之间的大小差、延迟差和相关性的信息。例如,空间线索包括:用于示出音频信号的功率增益信息的、音频信号之间的声道电平差 (CLD);音频信号之间的声道间电平差 (ICLD);音频信号之间的声道间时间差 (ICTD);用于示出音频信号之间的相关信息的、音频信号之间的相关声道间相关 (ICC);和音频信号之间的虚拟源位置信息,但是不限于这些示例。

[0040] 此外,所述边信息包括用于恢复并控制空间线索和音频信号的信息。所述边信息包括标首信息。所述标首信息包括用于恢复并播放具有各种声道的多对象音频信号的信息,并且可通过定义用于音频对象的声道信息和音频对象的标识 (ID) 来提供用于具有单声道、立体声或多声道的音频对象的解码信息。例如,定义用于每个对象的 ID 和信息,以标识已编码的特定音频对象是单声道音频信号还是立体声音频信号。作为实施例,所述标首信息可包括:空间音频编码 (SAC) 标首信息、音频对象信息和预设信息。

[0041] 代码转换器 103 渲染被输入到 SAOC 编码器 101 的音频对象,并基于从外部输入的控制信号(即,每个对象的声音信息和播放环境信息)来将从 SAOC 编码器 101 提取的 SAOC 比特流变换为 SAC 比特流。

[0042] 也就是说,代码转换器 103 基于所提取的 SAOC 比特流执行渲染,以将输入到 SAOC 编码器 101 的音频对象恢复为具有各种声道的多对象音频信号。可以在参数区域执行基于边信息的渲染。

[0043] 此外,代码转换器 103 将 SAOC 比特流变换为 SAC 比特流。所述代码转换器 103 从 SAOC 比特流获得输入音频对象的信息,并与所期望的音频场景对应地渲染输入音频对象的信息。在渲染过程中,代码转换器 103 预测与所期望的音频场景对应的空间信息,变换并输出已预测的空间信息作为 SAC 边信息比特流。

[0044] 将参考图 3 来详细描述代码转换器 103。

[0045] SAC 解码器 105 是基于空间线索的多声道音频解码器,基于从代码转换器 103 输出的 SAC 比特流而将从 SAOC 编码器 101 输出的缩混信号恢复为每个对象的音频信号,并将每个对象的音频信号恢复为具有各种声道的多对象音频信号。SAC 解码器 105 可以用运动图像专家组 (MPEG) 环绕解码器和心理声学编码 (BCC) 解码器来替换。

[0046] 图 2 是示出了根据本发明实施例的多对象音频编码器和多对象解码器的框图,并示出了输入信号是具有各种声道的多对象音频信号的情况。

[0047] 参考图 2 和图 1,本发明包括:SAOC 编码器 101、代码转换器 103、SAC 解码器 105、SAC 编码器 201、预设音频场景信息 (ASI) 203 和比特流格式器 205。

[0048] 当 SAOC 编码器 101 仅支持单声道或立体声音频对象时,SAC 编码器 201 从所输入

的多声道音频信号输出一个音频对象。所输出的音频对象是缩混后的单声道或立体声信号。此外, SAC 编码器 201 提取空间线索和边信息, 并创建 SAC 比特流。

[0049] SAOC 编码器 101 从包括由 SAC 编码器 201 输出的一个音频对象的多于一个音频对象输出代表性缩混信号, 提取空间线索和边信息, 并创建 SAOC 比特流。

[0050] 预设 ASI 203 将从外部输入的控制信号 (即, 每个对象的声音信息和播放环境信息) 形成为预设 ASI, 并创建包括预设 ASI 的预设 ASI 比特流。将参考图 4 来详细描述预设 ASI。

[0051] 比特流格式器 205 基于由 SAOC 编码器 101 创建的 SAOC 比特流、由 SAC 编码器 201 创建的 SAC 比特流、和由预设 ASI 203 创建的预设 ASI 比特流, 来创建代表性 SAOC 比特流。

[0052] 代码转换器 103 渲染向 SAOC 编码器 101 输入的音频对象, 并基于从外部输入的每个对象的声音信息和播放环境信息来将比特流格式器 205 所创建的代表性 SAOC 比特流变换为代表性 SAC 比特流。代码转换器 103 被包括在 SAC 解码器 105 中并如上所述地起作用。

[0053] SAC 解码器 105 基于从代码转换器 103 输出的 SAC 比特流, 来将从 SAOC 编码器 101 输出的缩混信号恢复为具有各种声道的多对象音频信号。SAC 解码器 105 可以用 MPEG 环绕解码器和 BCC 解码器来替换。

[0054] 图 3 是图示了根据本发明实施例的图 2 的代码转换器 103 的框图。

[0055] 参考图 3, 代码转换器 103 包括: 解析单元 301、渲染单元 303、第二矩阵单元 311、和第一矩阵单元 313, 并将代表性 SAOC 比特流变换为代表性 SAC 比特流。

[0056] 在图 1 中, 代码转换器 103 将 SAOC 比特流变换为 SAC 比特流。

[0057] 解析单元 301 解析由比特流格式器 205 创建的代表性 SAOC 比特流或者由图 1 的 SAOC 编码器 101 创建的 SAOC 比特流, 并划分在所述代表性 SAOC 比特流中包括的 SAOC 比特流和 SAC 比特流。此外, 解析单元 301 从已划分 SAOC 比特流中提取用于输入到 SAOC 编码器 101 的音频对象数目的信息。由于当解析由图 1 的 SAOC 编码器 101 创建的 SAOC 比特流时没有 SAC 比特流, 所以不必划分 SAC 比特流。

[0058] 第二矩阵单元 311 基于由解析单元 301 划分的 SAC 比特流创建第二矩阵。第二矩阵是关于输入到 SAC 编码器 201 的多声道音频信号的行列式。当在代表性 SAOC 比特流中不包括 SAC 比特流时, 即当解析由图 1 的 SAOC 编码器 101 创建的 SAOC 比特流时, 不需要第二矩阵单元 311。

[0059] 第二矩阵示出了输入到 SAC 编码器 201 的多声道音频信号的功率增益值, 并示出在等式 1 中。

[0060]

$$\underbrace{\begin{bmatrix} \mathbf{w}_{ch_1}^b \\ \mathbf{w}_{ch_2}^b \\ \vdots \\ \mathbf{w}_{ch_M}^b \end{bmatrix}}_{\text{矩阵 II}} \left[\mathbf{u}_{SAC}^b(k) \right] = \left[\mathbf{Y}_{SAC}^b(k) \right] = \begin{bmatrix} y_{ch_1}^b(k) \\ y_{ch_2}^b(k) \\ \vdots \\ y_{ch_M}^b(k) \end{bmatrix}$$

[0061] 等式 1

[0062] 通常,在将一个帧划分为子带之后的分析是 SAC 的基本分析过程。

[0063] $u_{SAC}^b(k)$ 是从 SAC 编码器 201 输出的缩混信号; k 是频率系数索引; 并且 b 是子带索引。 $w_{ch_i}^b$ 是从 SAC 比特流获得的多声道信号的空间线索信息, 并用于恢复第 i 声道信号的频率信息, $1 \leq i \leq M$ 。因此, $w_{ch_i}^b$ 可以表达为频率系数的大小信息或相位信息。因此, 在等式 1 的右边项中, $Y_{SAC}^b(k)$ 是等式 1 的结果, 并示出了从 SAC 解码器 105 输出的多声道音频信号。

[0064] $u_{SAC}^b(k)$ 和 $w_{ch_i}^b$ 是向量, 并且 $u_{SAC}^b(k)$ 的转置矩阵的维数是 $w_{ch_i}^b$ 的维数。例如, 这将被描述为等式 2。由于从 SAC 编码器 201 输出的缩混信号是单声道或立体声, 所以 m 是 1 或 2。

$$[0065] \quad \mathbf{w}_{ch_i}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_1^b & w_2^b & \cdots & w_m^b \end{bmatrix} \begin{bmatrix} u_1^b(k) \\ u_2^b(k) \\ \vdots \\ u_m^b(k) \end{bmatrix}$$

[0066] 等式 2

[0067] 如上所述, $w_{ch_i}^b$ 是包括在 SAC 比特流中的空间线索信息。当 $w_{ch_i}^b$ 指示每个声道的子带中的功率增益时, 可以从声道电平差空间线索中预测 $w_{ch_i}^b$ 。当 $w_{ch_i}^b$ 被用于作用于补偿频率系数的相位差的系数时, $w_{ch_i}^b$ 可以从声道时间差空间线索或声道间相干空间线索中预测。

[0068] 作为示例, 将描述 $w_{ch_i}^b$ 被用于作用于补偿频率系数之间的相位差的系数的情况。

[0069] 等式 1 的第二矩阵应表达每个声道的功率增益值, 并且是缩混信号的向量的维数的逆, 使得可以利用从 SAC 编码器 201 输出的缩混信号通过矩阵运算来创建输出信号 $Y_{SAC}^b(k)$ 。

[0070] 当第二矩阵单元 311 创建满足等式 1 和 2 的第二矩阵时, 渲染单元 303 组合所创建的第二矩阵和第一矩阵单元 313 的输出。

[0071] 第一矩阵单元 313 基于控制信号 (例如, 对象控制信息和播放系统信息), 来创建期望多于一个的、输入到 SAOC 编码器 101 的音频对象的输出, 即要被映射到具有各种声道的多对象音频信号的第一矩阵。

[0072] 当输入到 SAOC 编码器 101 的音频对象的数目是 N 时, 从 SAC 编码器 201 输出的缩混信号被认为是一个音频对象并包括在所输入的 N 个音频对象中。相应地, 除了从 SAC 编码器 201 输出的缩混信号之外的每个音频对象可以基于第一矩阵而被映射到从 SAC 解码器 105 输出的声道。

[0073] 当从 SAC 解码器 105 输出的声道的数目是 M 时, 第一矩阵可满足接下来的条件。

[0074]

$$\mathbf{P} \odot \mathbf{W}_{oj}^b = \underbrace{\begin{bmatrix} \mathbf{p}_{1,1}^b & \mathbf{p}_{1,2}^b & \cdots & \mathbf{p}_{1,N-1}^b \\ \mathbf{p}_{2,1}^b & \mathbf{p}_{2,2}^b & \cdots & \mathbf{p}_{2,N-1}^b \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p}_{M,1}^b & \mathbf{p}_{M,2}^b & \cdots & \mathbf{p}_{M,N-1}^b \end{bmatrix}}_{\text{矩阵 I}} \odot \begin{bmatrix} \mathbf{w}_{oj_1}^b \\ \mathbf{w}_{oj_2}^b \\ \vdots \\ \mathbf{w}_{oj_N-1}^b \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{ch_1}^b \\ \mathbf{w}_{ch_2}^b \\ \vdots \\ \mathbf{w}_{ch_M}^b \end{bmatrix}_{SAOC}$$

[0075] 等式 3

[0076] 其中 $\mathbf{w}_{oj_i}^b$ 是用于示出音频对象 i 的子带信号的信息的向量, $1 \leq i \leq N-1$, 并且是可以从 SAOC 比特流获得的空间线索信息。当音频对象 i 是立体声时, $\mathbf{w}_{oj_i}^b$ 是 2×1 矩阵向量。 $\mathbf{p}_{i,j}^b$ 是示出用于将第 j 音频对象映射到第 i 输出声道的功率增益信息或相位信息的第一矩阵的元素向量, 并且可以从由外部输入或者被设置为初始值的控制信息 (例如, 对象控制信息和播放系统信息) 中获得。

[0077] 满足等式 3 的条件的第一矩阵被传送到渲染单元 303, 并且在渲染单元 303 中运算等式 3。

[0078] 将在等式 4 和 5 中详细描述等式 3 的运算符和运算过程 $\odot$ 。

$$[\mathbf{p}_{1,1}^b \quad \mathbf{p}_{1,2}^b \quad \cdots \quad \mathbf{p}_{1,N-1}^b] \odot \begin{bmatrix} \mathbf{w}_{oj_1}^b \\ \mathbf{w}_{oj_2}^b \\ \vdots \\ \mathbf{w}_{oj_N-1}^b \end{bmatrix} = [\mathbf{p}_{1,1}^b \odot \mathbf{w}_{oj_1}^b + \mathbf{p}_{1,1}^b \odot \mathbf{w}_{oj_2}^b \cdots + \mathbf{p}_{1,1}^b \odot \mathbf{w}_{oj_N-1}^b] \text{ 等式 4}$$

[0080]

$$\mathbf{p}_{i,j}^b \odot \mathbf{w}_{oj_i}^b = [\mathbf{p}_{1,i,j}^b \quad \mathbf{p}_{2,i,j}^b \quad \cdots \quad \mathbf{p}_{m,i,j}^b] \odot \begin{bmatrix} \mathbf{w}_{1,oj_i}^b \\ \mathbf{w}_{2,oj_i}^b \\ \vdots \\ \mathbf{w}_{m,oj_i}^b \end{bmatrix}$$

$$= [\mathbf{p}_{1,i,j}^b \times \mathbf{w}_{1,oj_i}^b \quad \mathbf{p}_{2,i,j}^b \times \mathbf{w}_{2,oj_i}^b \quad \cdots \quad \mathbf{p}_{m,i,j}^b \times \mathbf{w}_{m,oj_i}^b] \text{ 等式 5}$$

[0082] 当所输入的音频对象是单声道和立体声时, m 是 2。

[0083] 例如, 当所输入的音频对象的数目是 Y , $m=2$, 并且所输出的声道的数目是 M 时, 第一矩阵的维数是 $M \times Y$, 且 Y 数目个 $\mathbf{p}_{i,j}^b$ 被形成成为 2×1 矩阵。当包括从 SAC 编码器 201 输出的音频对象时, 考虑 $Y = Y-1$ 。作为等式 3 的运算结果, 应该能够表达包括所输出的声道的功率增益向量 $\mathbf{w}_{ch_i}^b$ 的矩阵。所表达的向量的维数是 $M \times 2$, 并反映作为所输出的声道的数目的 M 和作为所输入的音频对象的布局的 2。

[0084] 再次参考图 3, 渲染单元 303 从第一矩阵 313 和第二矩阵 311 接收第一矩阵和第二矩阵。渲染单元 303 获得从由解析单元 301 划分的 SAOC 比特流获得的每个音频对象的空间线索信息 $\mathbf{w}_{oj_i}^b$, 通过组合基于第一矩阵和第二矩阵计算的输出向量来获得所期望的空间线索信息, 并创建包括所期望的空间线索信息的代表性 SAC 比特流。所期望的空间线索指的是与输出多声道音频信号有关的空间线索, 期望由用户从 SAC 解码器 105 输出所述输出多声道音频信号。

[0085] 如等式 6 地示出了用于基于第一矩阵和第二矩阵来获得所期望的空间线索信息的运算。

$$[0086] \quad pow(p_N) \begin{bmatrix} w_{ch_1}^b \\ w_{ch_2}^b \\ \vdots \\ w_{ch_M}^b \end{bmatrix}_{SAC} + (1 - pow(p_N)) \begin{bmatrix} w_{ch_1}^b \\ w_{ch_2}^b \\ \vdots \\ w_{ch_M}^b \end{bmatrix}_{SAOC} = \begin{bmatrix} w_{ch_1}^b \\ w_{ch_2}^b \\ \vdots \\ w_{ch_M}^b \end{bmatrix} = W_{modified}^b \quad \text{等式 6}$$

[0087] 当创建第一矩阵时没有考虑 P_N , 并且 P_N 示出了从 SAC 编码器 201 输出的音频对象的功率之和与直接输入到 SAOC 编码器 101 的音频对象的功率的比率。

[0088] P_N 可表达为等式 7。

$$[0089] \quad p_N = \frac{\sum_{k=N-1} power(object\#k)}{power(object\#N)} \quad \text{等式 7}$$

[0090] 因此, 当 $w_{ch_j}^b$ 是所输出的声道的功率时, 在音频对象的渲染之后的每个声道的功率比被示出为 $W_{modified}^b$ 。可以从 $W_{modified}^b$ 中重新提取所期望的空间线索参数。例如, 提取 ch_2 和 ch_1 之间的声道电平差 (CLD) 参数如等式 8 所示。

$$[0091] \quad CLD_{ch1/ch2}^b = 20 \log_{10} \frac{w_{ch1}^b}{w_{ch2}^b} = \left[20 \log_{10} \frac{w_{ch1,1}^b}{w_{ch2,1}^b}, 20 \log_{10} \frac{w_{ch1,2}^b}{w_{ch2,2}^b} \right]_{m=2} \quad \text{等式 8}$$

[0092] 当所传送的缩混信号是单声道信号时, CLD 参数如等式 9 所示。

$$[0093] \quad CLD_{ch1/ch2}^b = 10 \log_{10} \frac{(w_{ch1,1}^b)^2 + (w_{ch1,2}^b)^2}{(w_{ch2,1}^b)^2 + (w_{ch2,2}^b)^2} \quad \text{等式 9}$$

[0094] 所输出的声道的功率比被表达为作为空间线索参数的 CLD, 相邻声道之间的空间线索参数被表达为来自给定 $W_{modified}^b$ 信息的各种组合的格式。渲染单元 303 基于霍夫曼编码方法来创建包括从 $W_{modified}^b$ 提取的空间线索 (例如, CLD 参数) 的 SAC 比特流。

[0095] 在由渲染单元 303 创建的 SAC 比特流中包括的空间线索具有根据解码器的特性而不同的分析和提取方法。

[0096] 例如, BCC 解码器可基于一个声道使用等式 8 来提取 $N-1$ 个 CLD 参数。此外, MPEG 环绕解码器可根据 MPEG 环绕的每个声道的比较顺序来提取 CLD 参数。

[0097] 也就是说, 解析单元 301 划分 SAC 比特流和 SAOC 比特流, 并且第二矩阵单元 311 基于解析单元 301 所划分的 SAC 比特流和从 SAC 解码器 105 输出的多声道音频信号来创建第二矩阵, 如等式 1 所示。第一矩阵单元 313 创建与控制信号对应的第一矩阵。解析单元 301 所划分的 SAOC 比特流被传送到渲染单元 303, 并且渲染单元 303 从所传送的 SAOC 比特流中获得对象的信息, 利用第一矩阵执行运算, 组合运算结果和第二矩阵, 创建 $W_{modified}^b$, 从所创建的 $W_{modified}^b$ 中提取空间线索, 并创建代表性 SAC 比特流。

[0098] 也就是说, 从所创建的 $W_{modified}^b$ 中提取的空间线索变成所期望的空间线索。代表

性 SAC 比特流是根据 MPEG 环绕解码器或 BCC 解码器的特性而适当变换的比特流,并且可恢复为具有各种声道的多对象信号。

[0099] 图 4 图示了根据本发明实施例的由图 2 的比特流格式器 205 创建的代表性空间音频对象编码 (SAOC) 比特流。

[0100] 参考图 4,由比特流格式器 205 创建的代表性 SAOC 比特流是通过组合 SAOC 编码器 101 所创建的 SAOC 比特流和 SAC 编码器 201 所创建的 SAC 比特流来创建的,并且所述代表性 SAOC 比特流包括由预设 ASI 203 所创建的预设 ASI 比特流。将参考图 5 来详细描述所述预设 ASI 比特流。

[0101] 用于组合 SAOC 比特流和 SAC 比特流的第一方法是一种用于通过直接多路复用每个比特流来创建一个比特流的方法。在代表性 SAOC 比特流中串联连接 SAOC 比特流和 SAC 比特流 (参见 401)。

[0102] 第二方法是一种用于当存在 SAOC 附属 (ancillary) 数据区时、通过在 SAOC 附属数据区中包括 SAC 比特流信息来创建一个比特流的方法。在代表性 SAOC 比特流中串联连接 SAOC 比特流和附属数据区域,并且附属数据区包括 SAC 比特流 (参见 403)。

[0103] 第三方法是一种用于表达如下区域的方法,所述区域将在 SAOC 比特流和 SAC 比特流中的类似空间线索编码为相同比特流。例如,代表性 SAOC 比特流的标首信息区域包括 SAOC 比特流标首信息和 SAC 比特流标首信息,并且代表性 SAOC 比特流的每个特定区域包括与特定 CLD 有关的 SAOC 比特流和 SAC 比特流 (参见 405)。

[0104] 图 5 示出了根据本发明另一实施例的图 2 的代表性 SAOC 比特流,并示出了代表性 SAOC 比特流包括多个预设 ASI 的情况。

[0105] 参考图 5,代表性 SAOC 比特流包括预设 ASI 区域。所述预设 ASI 区域包括多个预设 ASI,并且预设 ASI 包括音频对象的控制信息和布局信息。

[0106] 当基于代码转换器 103 渲染音频对象时,应该输入每个音频对象的位置信息、控制信息和所输出的播放扬声器布局信息。

[0107] 当没有输入控制信息和播放扬声器布局信息时,在代码转换器 103 中将每个音频对象的控制信息和布局信息设置为默认值。

[0108] 代表性 SAOC 比特流或代表性 SAC 比特流的边信息或标首信息包括被设置为默认值的控制信息和布局信息、或者所输入的音频对象控制信息和布局信息。控制信息可以用两种方式来表达。第一,直接表达每个音频对象的控制信息 (例如,位置和电平) 和扬声器的布局信息。第二,以第一矩阵格式来表达控制信息和扬声器的布局信息,并且可以使用所述信息来替代第一矩阵单元 313 中的第一矩阵。

[0109] 预设 ASI 示出了音频对象控制信息和扬声器的布局信息。也就是说,预设 ASI 包括扬声器的布局信息以及每个音频对象的位置和电平信息,以形成适合于扬声器的布局信息的音频场景。

[0110] 如上所述,直接表达预设 ASI 或者以第一矩阵格式来表达预设 ASI,以便将解析单元 301 所提取的预设 ASI 传送到代表性 SAC 比特流。

[0111] 当直接表达预设 ASI 时,预设 ASI 可包括播放系统的布局 (例如,单声道 / 立体声 / 多声道)、音频对象 ID、音频对象布局 (例如,单声道或立体声)、音频对象位置、范围为从 0 度到 360 度的方位角、范围为 -50 度到 90 度的立体声播放仰角、和音频对象电平信

息 -50dB 至 50dB。

[0112] 当以第一矩阵格式表达预设 ASI 时,形成用于反映预设 ASI 的等式 3 的 P 矩阵,并且将 P 矩阵传送到渲染单元 303。P 矩阵包括作为元素向量的用于将每个音频对象映射到所输出的声道的功率增益信息或相位信息。

[0113] 预设 ASI 可定义关于所输入的同一音频对象的与所期望的播放情景对应的不同音频场景。例如,可根据内容产生器的对象和播放服务来另外传送在立体声或多声道 (5.1、7.1) 播放系统中需要的预设 ASI。

[0114] 图 6 是示出了根据本发明另一实施例的图 2 的代码转换器 103 的框图,并示出了没有从外部输入控制信号的情况。

[0115] 参考图 6,代码转换器 103 包括解析单元 301 和渲染单元 303。代码转换器 103 可接收第二矩阵单元 311、第一矩阵单元 313、预设 ASI 提取单元 601 和矩阵确定单元 603 的帮助。

[0116] 如上所述,当在代码转换器 103 中没有从外部输入的控制信号时,应用预设 ASI。

[0117] 解析单元 301 分开在代表性 SAOC 比特流中包括的 SAOC 比特流和 SAC 比特流,解析在代表性 SAOC 比特流中包括的预设 ASI 比特流,并将所述预设 ASI 比特流传送到预设 ASI 提取单元 601。

[0118] 预设 ASI 提取单元 601 从所解析的预设 ASI 比特流中输出默认的预设 ASI。然而,当存在对于选择预设 ASI 的请求时,输出所请求的预设 ASI。

[0119] 当预设 ASI 提取单元 601 所输出的预设 ASI 是所选择的预设 ASI 时,矩阵确定单元 603 确定所选择的预设 ASI 是否是第一矩阵格式。当所选择的预设 ASI 直接表达该信息时,将预设 ASI 传送到第一矩阵单元 313,并且第一矩阵单元 313 基于预设 ASI 来创建第一矩阵。当所选择的预设 ASI 是第一矩阵时,使用预设 ASI 作为直接输入到渲染单元 303 的信号。

[0120] 图 7 是示出了在图 2 的多对象音频编码器和解码器中另外包括音频对象去除器 701 的情况的框图。

[0121] 参考图 7,音频对象去除器 701 用于从 SAOC 编码器 101 所创建的代表性缩混信号中去除特定的音频对象。音频对象去除器 701 接收 SAOC 编码器 101 所创建的代表性缩混信号和来自代码转换器 103 的代表性 SAOC 比特流信息,并去除特定的音频对象。例如,可通过渲染单元 303 提供被传送到音频对象去除器 701 的代表性 SAOC 比特流信息。

[0122] 例如,将描述其中仅将作为 SAC 编码器 201 的缩混信号的音频对象 (对象 #N) 用作 SAC 解码器 105 的输入信号的情况。

[0123] SAOC 编码器 101 根据每个子带来提取所输入的音频对象的每个功率大小作为 CLD 值,并创建包括 CLD 值的 SAOC 比特流。可如下获得用于特定子带 m 的功率信息。

[0124] $P_{m}^{object\#1}, P_{m}^{object\#2}, \dots, P_{m}^{object\#N}$

[0125] 其中 $P_{m}^{object\#N}$ 是由 SAOC 编码器 101 输出的代表性缩混信号的第 m 频带的功率大小。因此, $u(n)$ 是输入到音频对象去除器 701 的代表性缩混信号,并且 $U(f)$ 正将代表性缩混信号变换到频率区域。

[0126] 当 $U^{modified}(f)$ 是音频对象去除器 701 的输出信号 (即, SAC 解码器 105 的输入信号) 时, $U^{modified}(f)$ 对应于 SAC 编码器 201 的缩混信号的音频对象 (对象 #N),并表达为等

式 10。

$$[0127] \quad U^{\text{modified}}(f) = U(f) \times \sqrt{\frac{P_m^{\text{object}\#N}}{\sum_{i=1}^N P_m^{\text{object}\#i}}} \times \delta, \quad A(m+1) \leq f \leq A(m+1)-1 \quad \text{等式 10}$$

[0128] 其中 $A(m)$ 表示第 m 子带的频率区域中的边界； δ 是用于控制电平大小的特定常数；并且 $U(f)$ 是单声道或立体声。

[0129] 下文中将描述 $U(f)$ 是单声道的情况。除了将 $U(f)$ 划分为左声道和右声道并进行处理之外， $U(f)$ 是立体声的情况与 $U(f)$ 是单声道的情况相同。

[0130] $U^{\text{modified}}(f)$ 被认为与作为 SAC 编码器 201 的缩混信号的音频对象（对象 #N）相同。因此，输入到 SAC 解码器 105 的代表性 SAC 比特流是从代表性 SAOC 比特流中排除 SAOC 比特流之后的比特流，并且可同等地用于从 SAC 编码器 201 输出的 SAC 比特流。也就是说，SAC 解码器 105 接收对象 #N，并将对象 #N 恢复为 M 个多声道信号。然而，整个信号的电平是通过代码转换器 103 的渲染单元 303、或者通过利用将等式 10 乘以特定常数而调制对象 #N 的信号电平来控制的。

[0131] 作为实施例，将描述其中仅要从 SAC 解码器 105 的输入信号去除作为 SAC 编码器 201 的缩混信号的对象 #N 的情况。

[0132] 等式 10 与等式 11 相同。

$$[0133] \quad U^{\text{modified}}(f) = U(f) \times \sqrt{\frac{\sum_{i=1}^{N-1} P_m^{\text{object}\#i}}{\sum_{i=1}^N P_m^{\text{object}\#i}}} \times \delta, \quad A(m+1) \leq f \leq A(m+1)-1 \quad \text{等式 11}$$

[0134] 因此，输入到 SAC 解码器 105 的代表性 SAC 比特流是从代表性 SAOC 比特流中排除 SAC 编码器 201 的 SAC 比特流之后的比特流，并被认为是渲染单元 303 的第二矩阵中没有输出。也就是说，代码转换器 103 通过解析代表性 SAOC 比特流块并仅渲染排除用于对象 #N 的音频对象信息之外的剩余信息，来创建代表性 SAC 比特流。

[0135] 因此，在代表性 SAC 比特流中不包括用于对象 #N 的功率增益信息和相关信息。在等式 11 中， δ 是用于控制电平大小的特定常数，正如等式 10 一样，并且可控制整个输出信号电平。

[0136] 音频对象去除器 701 从代表性缩混信号中去除音频对象，并且通过输入到代码转换器 103 的控制信号来确定去除命令。音频对象去除器 701 可以应用时间区域信号和频率区域信号二者。此外，可使用离散傅立叶变换 (DFT) 或正交镜像过滤带 (QMF: Quadrature Mirror Filterbank) 来将代表性缩混信号划分为子带。

[0137] 代码转换器 103 的渲染单元 303 去除 SAOC 比特流或 SAC 比特流，并将其传送到 SAC 解码器 105，并且音频对象去除器 701 去除与传送到 SAC 解码器 105 的比特流对应的音频对象。

[0138] 当代码转换器 103 被包括在 SAC 解码器 105 中时，从代码转换器 103 输出的代表性 SAC 比特流可传送到 SAC 解码器 105，而没有附加的变换过程。附加的变换过程意指诸如

量化或霍夫曼编码方法之类的一般编码过程。

[0139] 考虑了以下情况,即 SAOC 编码器 101 没有连接到 SAC 编码器 201,并且仅控制和恢复除了 SAC 编码器 201 的输出音频对象之外的输入到 SAOC 编码器 101 的音频对象,即对象 #1 至对象 #N-1。

[0140] 图 8 是示出了用 MPEG 环绕编码器和解码器来替换图 2 的 SAC 编码器 201 和 SAC 解码器 105 的情况的框图。

[0141] 参考图 8, SAC 编码器 201 用 MPEG 环绕编码器 (即, MPS 编码器 801) 来替换,并且 SAC 解码器 105 用 MPEG 环绕解码器 (即, MPS 解码器 805) 来替换。此外,当从 SAOC 编码器 101 输出的代表性缩混信号是立体声时,另外需要信号处理单元 803。

[0142] MPS 编码器 801 执行与图 2 的 SAC 编码器 201 相同的功能。也就是说, MPS 编码器 801 输出来自所输入的多声道音频信号的一个音频对象,提取空间线索和边信息,并创建 MPS 比特流。所输出的音频对象是缩混后的单声道或立体声信号。

[0143] 此外, MPS 解码器 805 执行与图 2 的 SAC 解码器 105 相同的功能。也就是说, MPS 解码器 805 基于从代码转换器 103 输出的 SAC 比特流,来恢复从 SAOC 编码器 101 输出的缩混信号或者从信号处理单元 803 输出的代表性重新缩混信号,作为具有各种声道的多对象音频信号。

[0144] 其间,当从 SAOC 编码器 101 输出的缩混信号是立体声时 (即,当 MPS 解码器 805 处理立体声信号时),由于立体声信号的左 / 右处理中的限制,所以信号处理单元 803 需要 MPS 解码器 805。

[0145] 等式 2 示出了缩混信号被概括为普通 SAC 解码器中的 m 个数目的情况,当缩混信号是立体声时,关于已恢复的输出声道 1 的等式 2 与等式 12 相同。

$$[0146] \quad \mathbf{w}_{ch_1}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_{L,ch_1}^b & w_{R,ch_2}^b \end{bmatrix} \begin{bmatrix} u_L^b(k) \\ u_R^b(k) \end{bmatrix} \text{等式 12}$$

[0147] 输出声道的向量应该能够应用到所有缩混信号,但是这在当前 MPS 解码器 805 中是不可能的。如等式 13 所示,这是因为在 MPS 解码器 805 中矩阵值被限制为 0。

$$[0148] \quad \mathbf{w}_{ch_1}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_{L,ch_1}^b & 0 \end{bmatrix} \begin{bmatrix} u_L^b(k) \\ u_R^b(k) \end{bmatrix} \text{等式 13}$$

[0149] 也就是说,由于 $u_R^b(k)$ 元素没有反映在恢复输出声道 1 中,所以不能应用在等式 3、4 和 5 中创建的 $w_{ch_2}^b$ 。因此,不可能进行关于具有立体声以上的布局的信号灵活定位。也就是说,不可能进行立体声信号的左信号和右信号之间的自由渲染。

[0150] 然而,基于信号处理单元 803 来对从 SAOC 编码器 101 输出的代表性缩混信号再次进行缩混,并输出为代表性重新缩混信号。信号处理单元 803 的处理如等式 14 所示。

$$[0151] \quad \begin{bmatrix} \mathbf{w}_{ch_1}^b \\ \mathbf{w}_{ch_2}^b \\ \mathbf{M} \\ \mathbf{w}_{ch_M}^b \end{bmatrix}_{\text{modified}} \times [\mathbf{u}_{stereo}^b(k)] = \begin{bmatrix} y_{ch_1}^b(k) \\ y_{ch_2}^b(k) \\ \mathbf{M} \\ y_{ch_M}^b(k) \end{bmatrix} \quad \text{等式 14}$$

[0152] 当从 SAOC 编码器 101 输出的代表性缩混信号是立体声时,信号处理单元 803 的输出信号如等式 15 所示。

$$[0153] \quad \begin{bmatrix} \mathbf{w}_L^b \\ \mathbf{w}_R^b \end{bmatrix}_{\text{modified}} \times [\mathbf{u}_{stereo}^b(k)] = \begin{bmatrix} y_{ch_L}^b(k) \\ y_{ch_R}^b(k) \end{bmatrix} \quad \text{等式 15}$$

[0154] 其中 $y_{ch_L}^b(k)$ 和 $y_{ch_R}^b(k)$ 是信号处理单元 803 所输出的信号,并被输入到 MPS 解码器 805。由于如等式 15 所示 $y_{ch_L}^b(k)$ 和 $y_{ch_R}^b(k)$ 是用于反映左信号和右信号的渲染的信号,所以尽管如等式 13 所示限制了 MPS 解码器 805,但是 MPS 解码器 805 也可输出其中左信号和右信号被自由渲染的信号。

[0155] 例如,当 MPS 解码器 805 将 w_L^b 、 w_R^b 恢复为 5 个声道时,如下在等式 14 中表达 w_L^b 、 w_R^b 。

$$[0156] \quad \left(\text{例如, } \mathbf{w}_L^b = \mathbf{w}_{ch_Lf}^b + \mathbf{w}_{ch_Ls}^b + \mathbf{w}_{ch_C}^b / \sqrt{2}, \mathbf{w}_R^b = \mathbf{w}_{ch_Rf}^b + \mathbf{w}_{ch_Rs}^b + \mathbf{w}_{ch_C}^b / \sqrt{2} \right)$$

[0157] 如上所述,当 MPS 解码器 805 由于 MPEG 环绕的限制而难以处理立体声信号时,信号处理单元 803 通过基于从代码转换器 103 传送的对象位置信息而再次执行缩混,来输出代表性的重新缩混信号。例如,可以由渲染单元 303 来提供传送到信号处理单元 803 的对象位置信息。根据如上所述的类似方法,渲染单元 303 可基于代表性 SAOC 比特流来创建代表性 MPS 比特流,其包括相对于输入到 SAOC 编码器 101 和 MPS 编码器 801 的音频信号的、用于要由 MPS 解码器 805 输出的音频信号的左信号和右信号中的每个的空间线索信息。

[0158] MPS 解码器 805 可通过利用信号处理单元 803 而进行操作,来执行与图 2 的 SAC 解码器 105 相同的功能。

[0159] MPS 解码器 805 将从信号处理单元 803 输出的代表性重新缩混信号恢复为所期望的输出,即具有各种声道的多对象信号。

[0160] 利用图 2 的 SAC 解码器 105 或信号处理单元 803 进行操作的 MPS 解码器 805 的解码方法包括如下步骤:接收多声道和多对象缩混信号以及多声道多对象边信息信号;将多声道多对象缩混信号变换为多声道缩混信号;将多声道和多对象信息信号变换为多声道信息信号;基于变换后的多声道缩混信号和多声道信息信号来合成音频信号。

[0161] 变换多声道缩混信号的步骤包括:基于从多声道和多对象信息信号获得的对象相关信息来从多声道多对象缩混信号去除对象信息的步骤。变换多声道缩混信号的步骤包括:基于从多声道多对象信息信号获得的对象相关信息来控制来自多声道多对象缩混信号的对象信息的步骤。

[0162] 在包括变换多声道缩混信号的步骤的解码方法中,可通过对象控制信息来控制对象相关信息。这里,可通过解码系统信息来控制对象相关信息。

[0163] 尽管上面结合设备描述了根据本发明的编码和解码过程,但是在设备中包括的每个组成元件也可以用在处理角度中所需要的每个构成元件来替换。在这个情况下,显然的是,根据本发明的编码和解码过程可结合方法来理解。

[0164] 上面描述的本发明的技术可实现为程序并存储在计算机可读记录介质(诸如,CD-ROM、RAM、ROM、软盘、硬盘和磁光盘)中。由于本发明技术领域中的技术人员可容易地实现所述处理,所以这里将不再提供进一步的描述。

[0165] 尽管已经结合特定的优选实施例来描述了本发明,但对于本领域的技术人员显然的是,可以做出各种改变和修改,而不脱离接下来的权利要求所限定的本发明的范围。

[0166] 工业实用性

[0167] 本发明可通过有效地编码和解码具有各种声道的多对象音频内容,而根据用户需求来主动地消费音频内容,并通过提供与传统上使用的比特流的后向兼容性来提供与传统的编码和解码设备的兼容性。

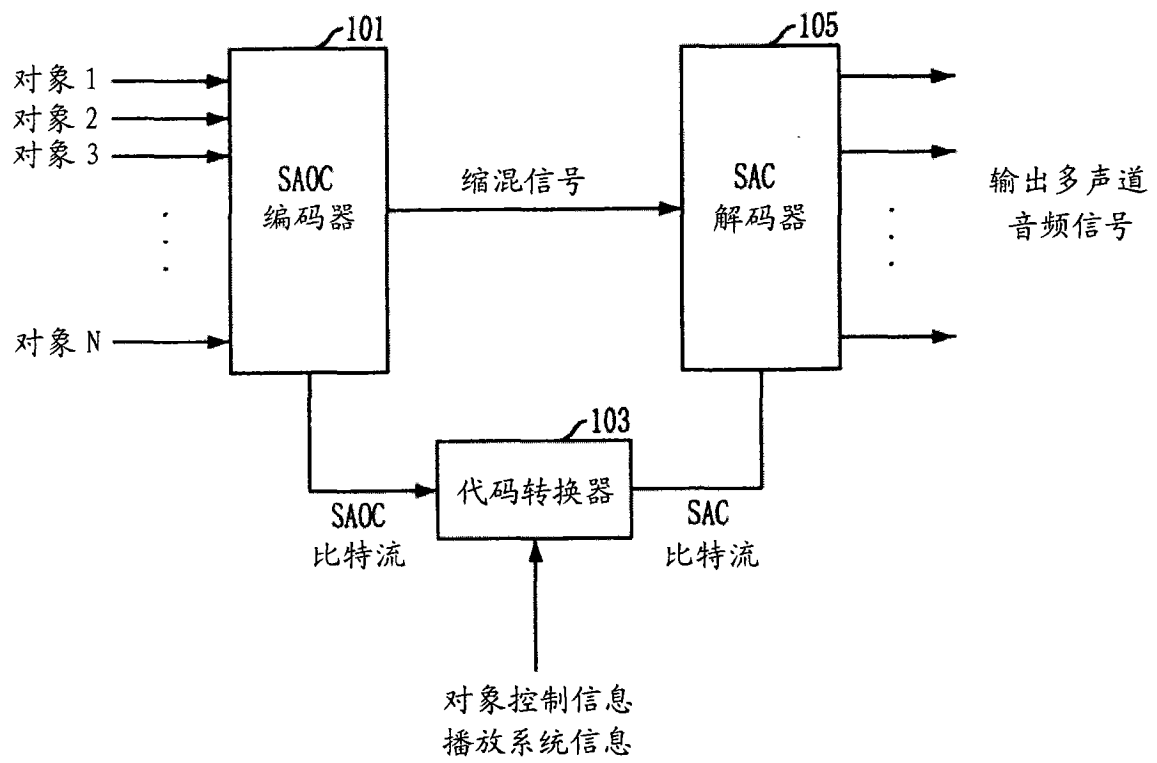


图 1

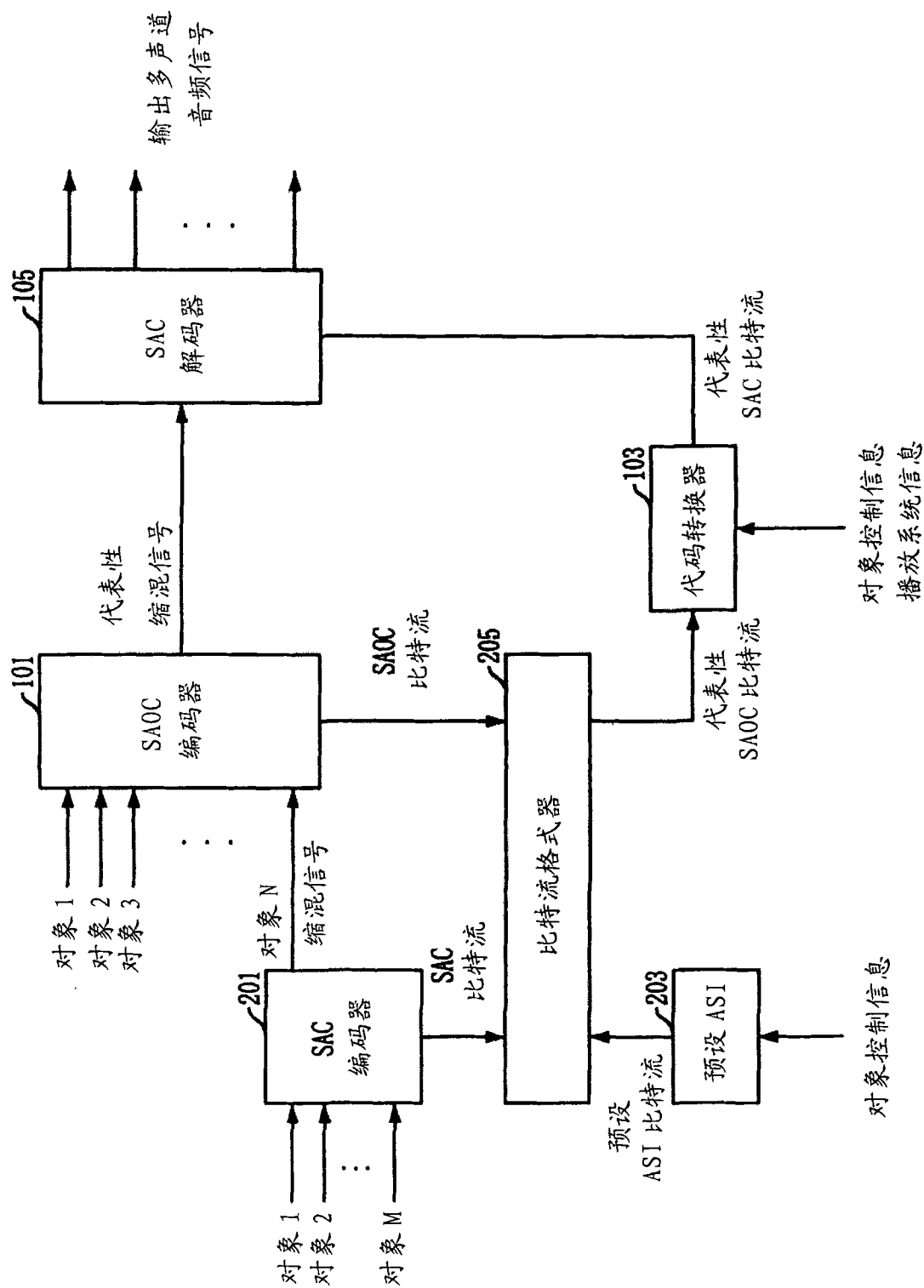


图 2

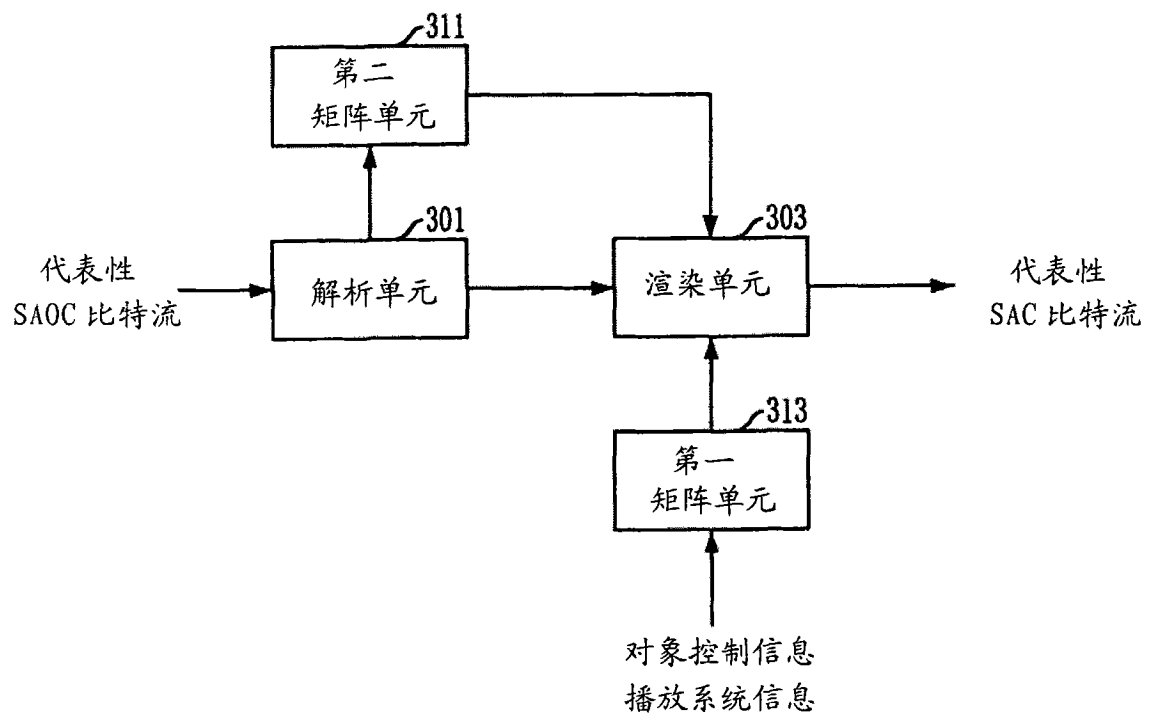


图 3

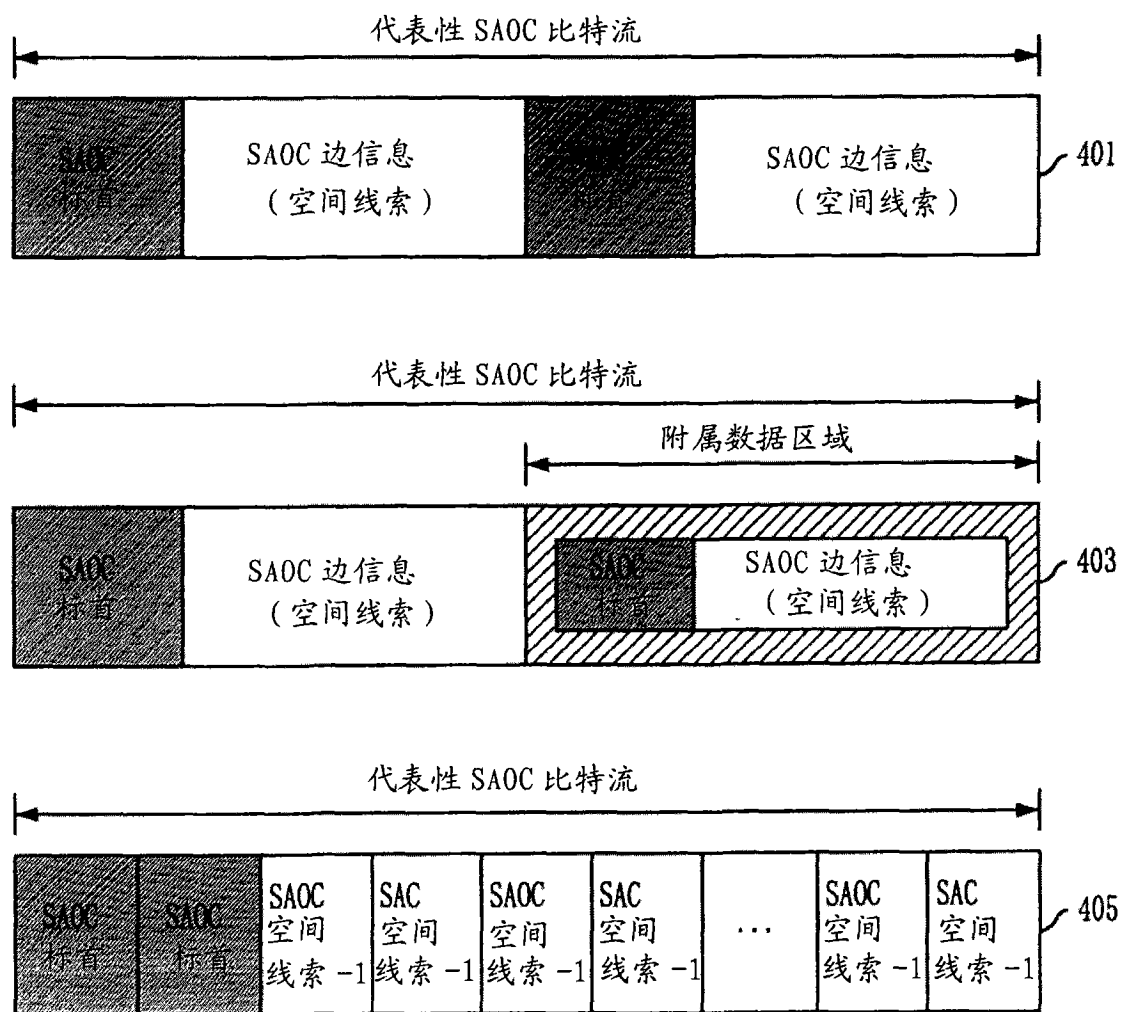


图 4

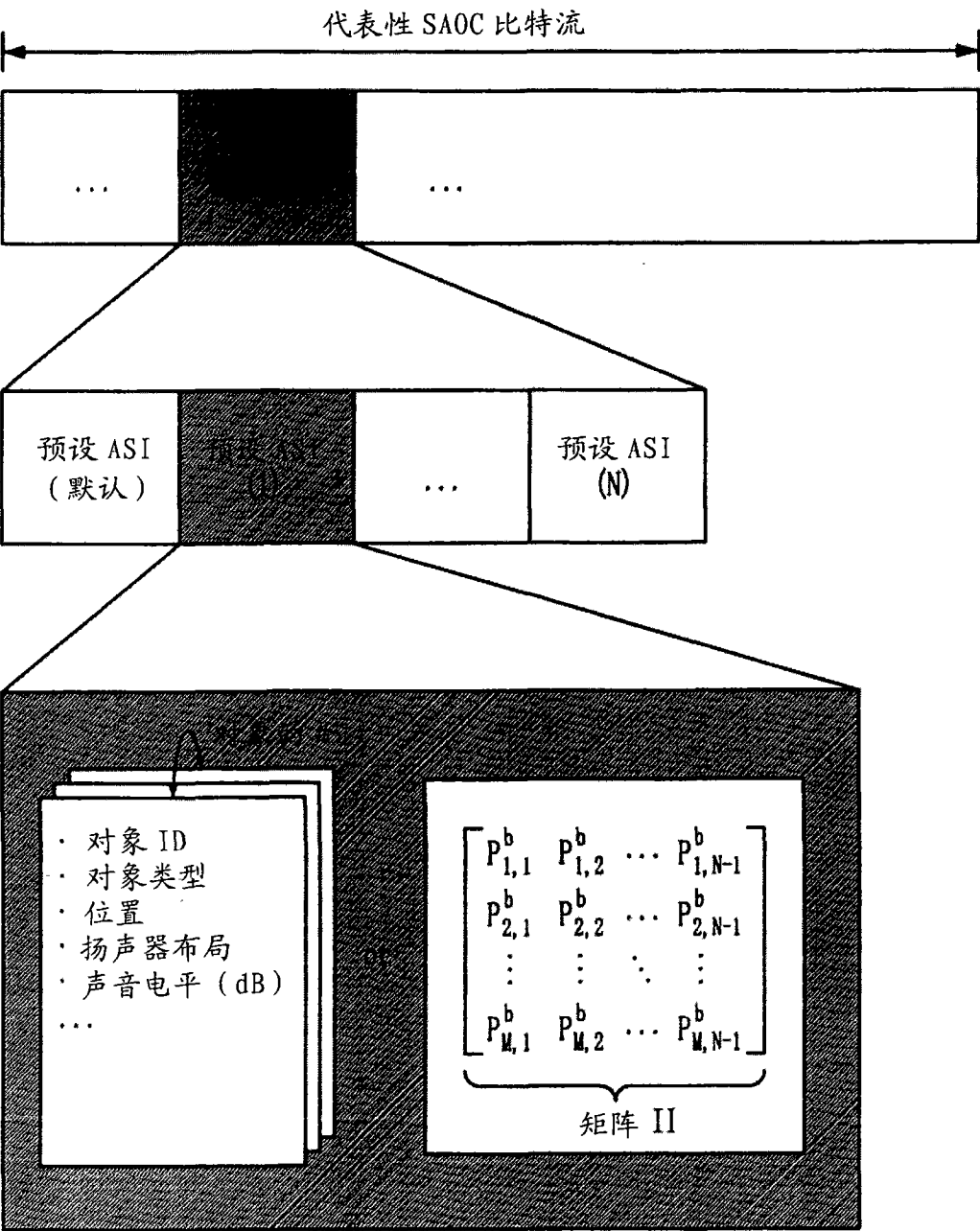


图 5

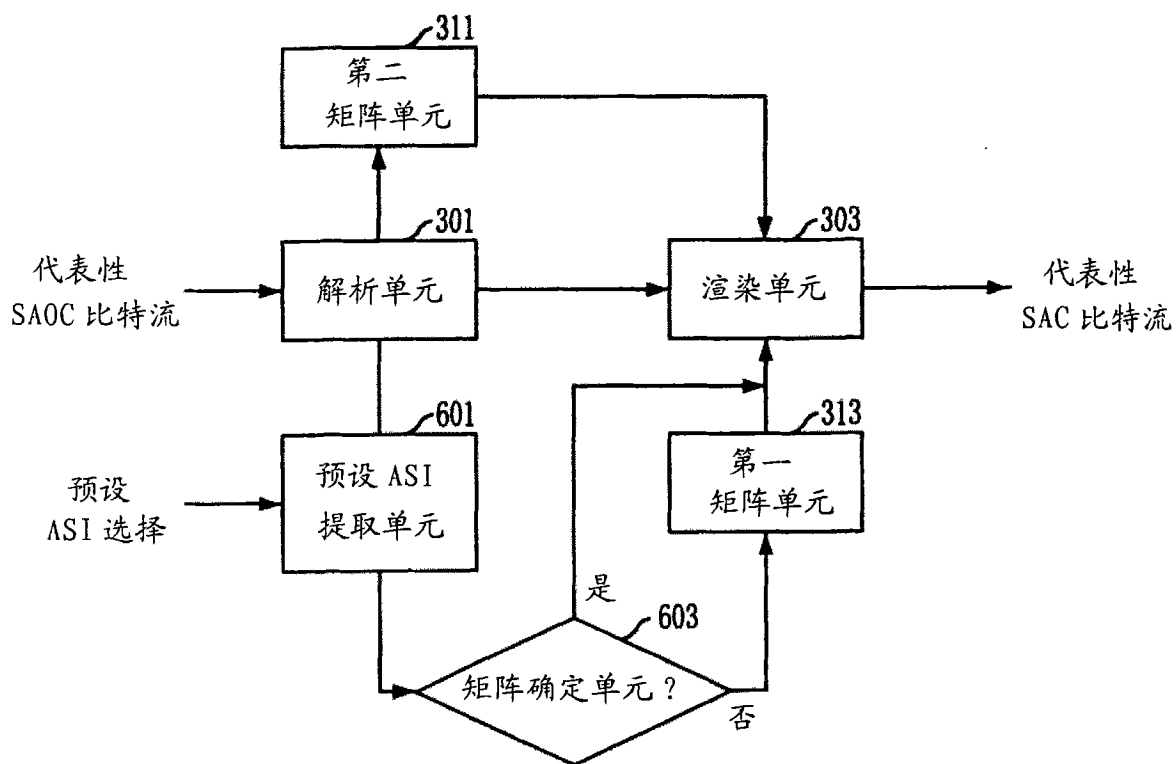


图 6

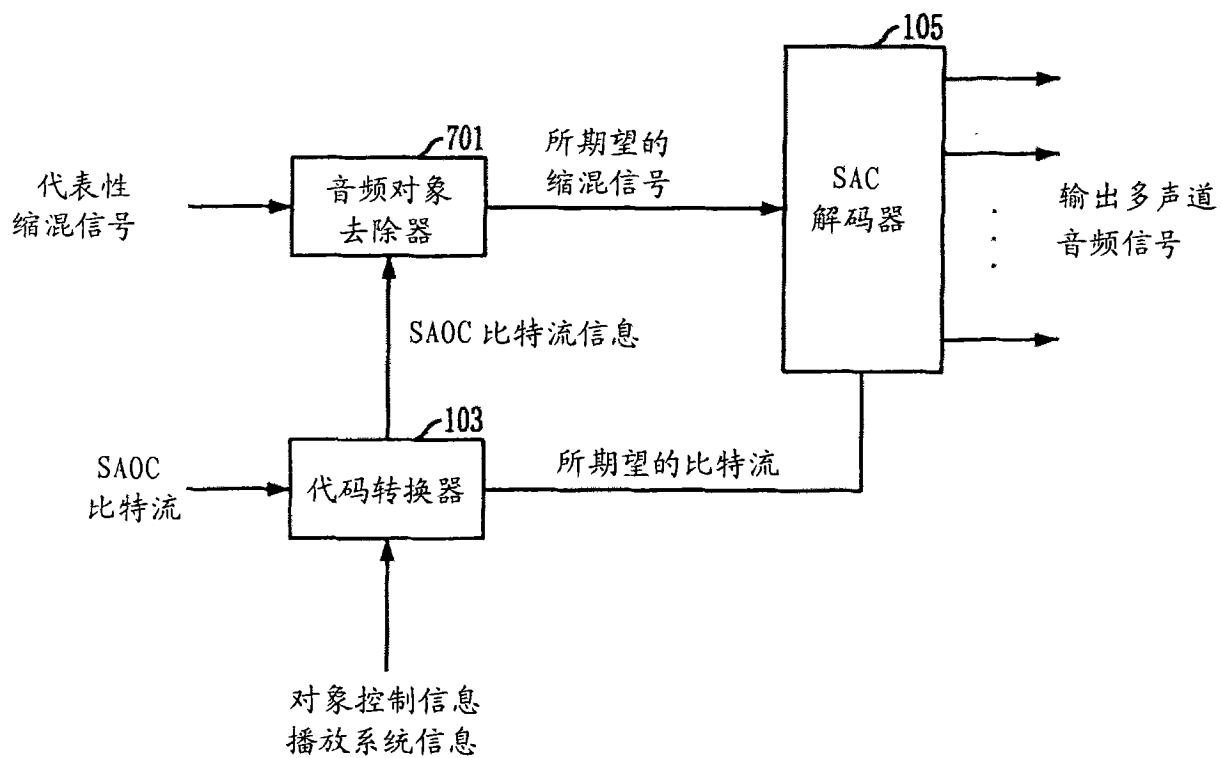


图 7

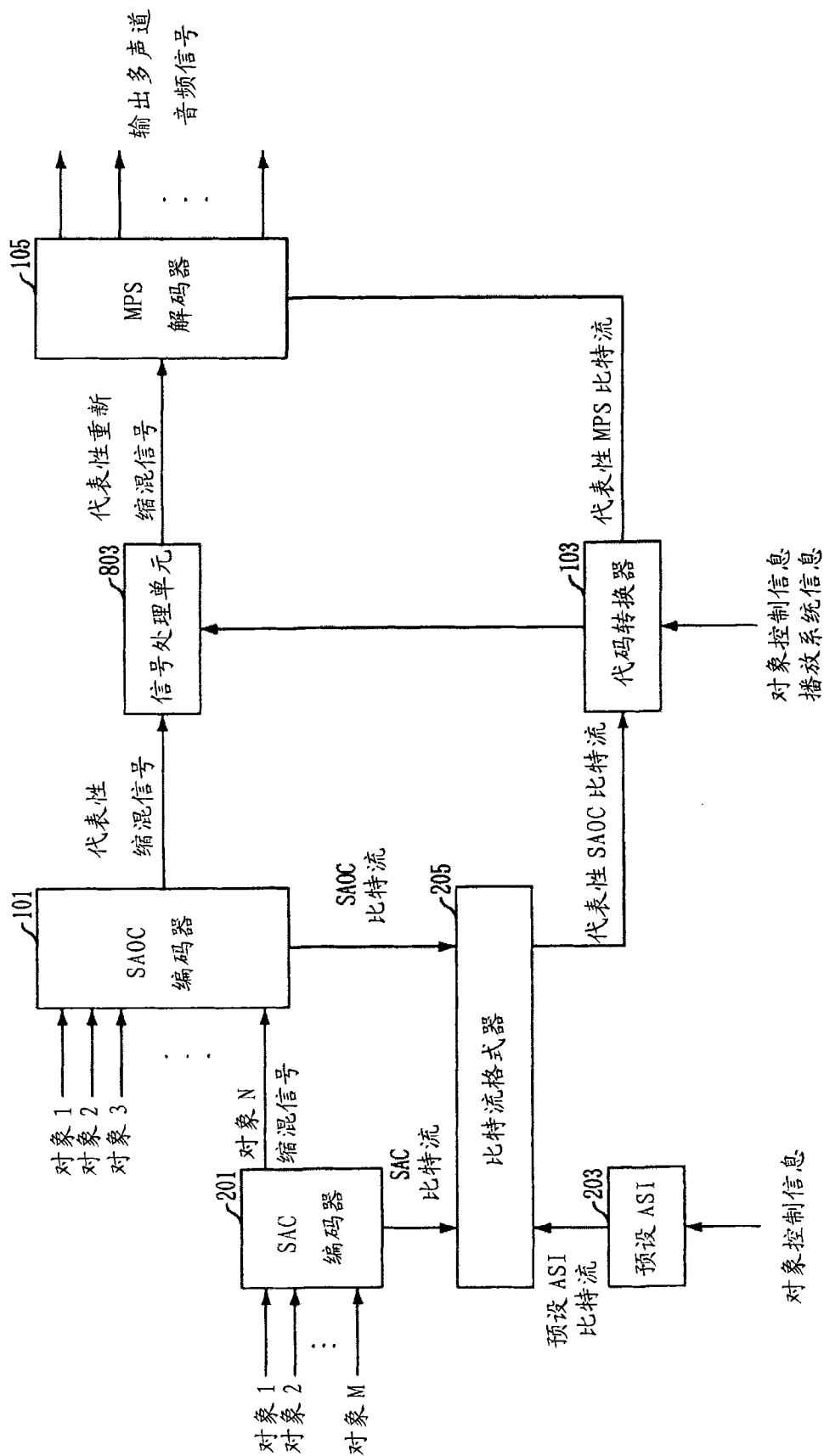


图 8