



(12)发明专利

(10)授权公告号 CN 104429102 B

(45)授权公告日 2017.12.15

(21)申请号 201380037326.5

(22)申请日 2013.07.16

(65)同一申请的已公布的文献号
申请公布号 CN 104429102 A

(43)申请公布日 2015.03.18

(30)优先权数据
61/672,280 2012.07.16 US
61/754,416 2013.01.18 US
13/942,657 2013.07.15 US

(85)PCT国际申请进入国家阶段日
2015.01.13

(86)PCT国际申请的申请数据
PCT/US2013/050648 2013.07.16

(87)PCT国际申请的公布数据
W02014/014891 EN 2014.01.23

(73)专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72)发明人 D·森

(74)专利代理机构 北京律盟知识产权代理有限公司 11287
代理人 宋献涛

(51)Int.Cl.
H04S 7/00(2006.01)
H04S 3/00(2006.01)

(56)对比文件
JP H9-244663, 1997.09.16,
CN 1507701 A, 2004.06.23,
EP 2094032 A1, 2009.08.26,
Jérôme Daniel,《Spatial Sound Encoding Including Near Field Effect:Introducing Distance Coding Filters and a Viable,New Ambisonic Format》,《AES 23rd International Conference》,2006,

审查员 李乔

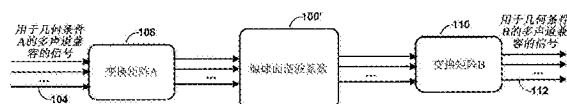
权利要求书2页 说明书21页 附图22页

(54)发明名称

使用3D音频分层译码的扩音器位置补偿

(57)摘要

一般来说,本发明描述用于使用分层三维3D音频译码来补偿扩音器位置的技术。包括一或多个处理器的设备可执行所述技术。所述处理器可经配置以对用于第一扬声器几何条件的第一组音频声道信息执行基于球面波模型的第一变换以产生描述声音场的第一分层要素集合。所述处理器可进一步经配置以在频域中对所述第一分层要素集合执行第二变换以产生用于第二扬声器几何条件的第二组音频声道信息。



1. 一种音频信号处理的方法,其包括:

基于第一扬声器几何条件的坐标对扬声器信道执行平移以产生第一多个虚拟扬声器信道;以及

使用基于球面波模型的第一变换对所述第一多个虚拟扬声器信道执行变换以产生描述声音场的分层要素集合。

2. 根据权利要求1所述的方法,其中所述扬声器信道及所述第一扬声器几何条件的坐标被映射到第二扬声器几何条件,且其中所述第一扬声器几何条件和所述第二扬声器几何条件具有不同半径,和/或其中所述第一扬声器几何条件及第二扬声器几何条件具有不同方位角,和/或其中所述分层要素集合包括球面谐波系数,和/或其中所述扬声器信道和所述第一扬声器几何条件的坐标被映射到所述第二扬声器几何条件以补偿所述第一扬声器几何条件中的要素与所述第二扬声器几何条件中的要素之间的位置差异。

3. 根据权利要求2所述的方法,其中所述第一扬声器几何条件及第二扬声器几何条件具有不同仰角。

4. 根据权利要求1所述的方法,其中对所述扬声器信道执行平移包括对所述扬声器信道执行向量基础振幅平移以产生所述第一多个虚拟扬声器信道,和/或其中所述扬声器信道中的每一者与空间的对应的不同界定区相关联,且其中以音频格式规格及音频格式标准中的一或多者来界定空间的所述不同界定区。

5. 根据权利要求1所述的方法,其进一步包括:

使用第二变换在频域中将所述分层要素集合变换为第二多个虚拟扬声器信道;以及

对所述第二多个虚拟扬声器信道执行平移以产生不同的扬声器信道,其中所述不同的扬声器信道中的每一者与所述空间的对应的不同界定区相关联,且其中对所述第二多个虚拟扬声器信道执行平移包括对所述第二多个虚拟扬声器信道执行向量基础振幅平移以产生所述不同的扬声器信道,和/或其中所述第二多个虚拟扬声器信道中的每一者与所述空间的对应的不同界定区相关联,且其中以音频格式规格及音频格式标准中的一或多者来界定空间的所述不同界定区。

6. 根据权利要求5所述的方法,其中所述扬声器信道与第一空间几何条件相关,且其中所述不同扬声器信道与不同于所述第一空间几何的第二空间几何条件相关。

7. 根据权利要求1所述的方法,其进一步包括:

接收扬声器信道和第一扬声器几何条件的坐标;以及
其中所述第一扬声器几何条件对应于所述信道的位置。

8. 一种音频信号处理的设备,其包括:

用于基于第一扬声器几何条件的坐标对扬声器信道执行平移以产生第一多个虚拟扬声器信道的装置;

用于使用基于球面波模型的第一变换对所述第一多个虚拟扬声器信道执行变换以产生描述声音场的分层要素集合的装置。

9. 根据权利要求8所述的设备,其中所述扬声器信道及所述第一扬声器几何条件的坐标被映射到第二扬声器几何条件,且其中所述第一扬声器几何条件和所述第二扬声器几何条件具有不同半径,和/或其中所述第一扬声器几何条件及第二扬声器几何条件具有不同方位角,和/或其中所述分层要素集合包括球面谐波系数,和/或所述扬声器信道和所述第

一扬声器几何条件的坐标被映射到所述第二扬声器几何条件以补偿所述第一扬声器几何条件中的要素与所述第二扬声器几何条件中的要素之间的位置差异。

10. 根据权利要求9所述的设备, 其中所述第一扬声器几何条件及第二扬声器几何条件具有不同仰角。

11. 根据权利要求9所述的设备, 所述用于对所述扬声器信道执行平移的装置包括用于对所述扬声器信道执行向量基础振幅平移以产生所述第一多个虚拟扬声器信道的装置, 和/或其中所述扬声器信道中的每一者与空间的对应不同界定区相关联, 且其中以音频格式规格及音频格式标准中的一或多者来界定空间的所述不同界定区。

12. 根据权利要求9所述的设备, 其进一步包括:

用于使用第二变换在频域中将所述分层要素集合变换为所述第二多个虚拟扬声器信道的装置; 以及

用于对所述第二多个虚拟扬声器信道执行平移以产生不同扬声器信道的装置, 其中所述不同扬声器信道中的每一者与空间的对应的不同界定区相关联, 且其中所述用于对所述第二多个虚拟扬声器信道执行平移的装置包括用于对所述第二多个虚拟扬声器信道执行向量基础振幅平移以产生所述不同扬声器信道的装置。

13. 根据权利要求12所述的设备, 其中所述第二多个虚拟扬声器信道中的每一者与空间的对应的不同界定区相关联。

14. 根据权利要求13所述的设备, 其中以音频格式规格及音频格式标准中的一或多者来界定空间的所述不同界定区。

15. 根据权利要求12所述的设备, 其中所述扬声器信道与第一空间几何条件相关联, 且其中所述不同扬声器信道与不同于所述第一空间几何条件的第二空间几何条件相关联。

16. 根据权利要求8所述的设备, 其进一步包括用于发送扬声器信道和第一扬声器几何条件的坐标的装置, 且其中所述第一扬声器几何条件对应于所述信道的位置。

使用3D音频分层译码的扩音器位置补偿

[0001] 本申请案主张2012年7月16日申请的第61/672,280号美国临时申请案及2013年1月18日申请的第61/754,416号美国临时申请案的权益。

技术领域

[0002] 本发明涉及空间音频译码。

背景技术

[0003] 存在例如从由NHK(日本广播协会或日本广播公司)开发的5.1家庭影院系统到22.2系统的范围的各种‘环绕声音’格式。常常,这些所谓的环绕声音格式指定将定位扬声器的位置,以使得扬声器可在音频回放系统处最佳地再现声音场。但具有支持环绕声音格式中的一或多者的音频回放系统的人常常未将扬声器准确放置在格式指定的位置处,这常常是因为音频回放系统所处的房间对在何处可放置扬声器具有局限性。虽然某些格式在于何处可定位扬声器方面比其它格式更灵活,但一些格式已经被更广泛地采用,从而导致消费者归因于与到更灵活的格式的升级或转变相关联的高成本而对到这些更灵活的格式的升级或转变犹豫不决。

发明内容

[0004] 本发明描述可用于解决此向后兼容性的缺乏同时还促进到更灵活的环绕声音格式(再次,这些格式在于何处可定位扬声器方面“更灵活”)的转变的方法、系统及设备。本发明中描述的技术可提供用于发送及接收向后兼容音频信号两者的各种方式,其可适应到可提供声音场的二维或三维表示的球面谐波系数(SHC)的变换。通过使得能够将向后兼容音频信号(例如,符合5.1环绕声音格式的音频信号)变换为SHC,所述技术可恢复可映射到几乎任何扬声器几何条件的声音场的三维表示。

[0005] 在一个方面中,一种音频信号处理方法包括:使用基于球面波模型的第一变换将用于第一扬声器几何条件的第一组音频声道信息变换为描述声音场的第一分层要素集合;及使用第二变换在频域中将所述第一分层要素集合变换为用于第二扬声器几何条件的第二组音频声道信息。

[0006] 在另一方面中,一种设备包括一或多个处理器,所述一或多个处理器经配置以:对用于第一扬声器几何条件的第一组音频声道信息执行基于球面波模型的第一变换以产生描述声音场的第一分层要素集合;及在频域中对所述第一分层要素集合执行第二变换以产生用于第二扬声器几何条件的第二组音频声道信息。

[0007] 在另一方面中,一种设备包括:用于使用基于球面波模型的第一变换将用于第一扬声器几何条件的第一组音频声道信息变换为描述声音场的第一分层要素集合的装置;及用于使用第二变换在频域中将所述第一分层要素集合变换为用于第二扬声器几何条件的第二组音频声道信息的装置。

[0008] 在另一方面中,一种在其上存储有指令的非暂时性计算机可读存储媒体,所述指

令在被执行时致使一或多个处理器：使用基于球面波模型的第一变换将用于第一扬声器几何条件的第一组音频声道信息变换为描述声音场的第一分层要素集合；及使用第二变换在频域中将所述第一分层要素集合变换为用于第二扬声器几何条件的第二组音频声道信息。

[0009] 在另一方面中，一种方法包括接收扩音器声道以及第一扬声器几何条件的坐标，其中所述扩音器声道已经被转换为分层要素集合。

[0010] 在另一方面中，一种设备包括一或多个处理器，所述一或多个处理器经配置以接收扩音器声道以及第一扬声器几何条件的坐标，其中所述扩音器声道已经被转换为分层要素集合。

[0011] 在另一方面中，一种设备包括用于接收扩音器声道以及第一扬声器几何条件的坐标的装置，其中所述扩音器声道已经被转换为分层要素集合。

[0012] 在另一方面中，一种非暂时性计算机可读存储媒体包括指令，所述指令在被执行时致使一或多个处理器接收扩音器声道以及第一扬声器几何条件的坐标，其中所述扩音器声道已经被转换为分层要素集合。

[0013] 在另一方面中，一种方法包括传输扩音器声道以及第一扬声器几何条件的坐标，其中所述第一几何条件对应于所述声道的位置。

[0014] 在另一方面中，一种设备包括一或多个处理器，所述一或多个处理器经配置以传输扩音器声道以及第一扬声器几何条件的坐标，其中所述几何条件对应于所述声道的位置。

[0015] 在另一方面中，一种设备包括用于传输扩音器声道以及第一扬声器几何条件的坐标的装置，其中所述几何条件对应于所述声道的位置。

[0016] 在另一方面中，一种非暂时性计算机可读存储媒体在其上存储有指令，所述指令在被执行时致使一或多个处理器传输扩音器声道以及第一扬声器几何条件的坐标，其中所述几何条件对应于所述声道的位置。

[0017] 在附图及以下描述中阐述所述技术的一或多个方面的细节。这些技术的其它特征、目标和优点将从所述描述和图式以及权利要求书而显而易见。

附图说明

[0018] 图1是说明用于使用编解码器的标准化的通用结构的图。

[0019] 图2是说明用于单声道/立体声的向后兼容实例的图。

[0020] 图3是说明在不考虑向后兼容性的情况下的基于场景的译码的实例的图。

[0021] 图4是说明使用向后兼容设计的编码过程的实例的图。

[0022] 图5是说明无法解码基于场景的数据的常规解码器上的解码过程的实例的图。

[0023] 图6是说明使用可处置基于场景的数据的装置的解码过程的实例的图。

[0024] 图7A是说明根据本发明中描述的技术的各种方面的音频信号处理的方法的流程图。

[0025] 图7B是说明可执行本发明中描述的技术的各种方面的设备的框图。

[0026] 图7C是说明根据另一一般配置的用于音频信号处理的设备的框图。

[0027] 图8A是说明根据本发明中描述的技术的各种方面的音频信号处理的方法的流程图。

- [0028] 图8B是说明根据本发明中描述的技术的各种方面的方法的实施方案的流程图。
- [0029] 图9A是说明从SHC到多声道信号的转换的图。
- [0030] 图9B是说明从多声道信号到SHC的转换的图。
- [0031] 图9C是说明从与几何条件A兼容的多声道信号到SHC的第一转换及从SHC到与几何条件B兼容的多声道信号的第二转换的图。
- [0032] 图10A是说明根据一般配置的音频信号处理的方法M400的流程图。
- [0033] 图10B是说明根据一般配置的用于音频信号处理的设备MF400的框图。
- [0034] 图10C是说明根据另一一般配置的用于音频信号处理的设备A400的框图。
- [0035] 图10D是说明可执行本发明中描述的技术的各种方面的系统的实例的图。
- [0036] 图11A是说明可执行本发明中描述的技术的各种方面的另一系统的实例的图。
- [0037] 图11B是说明可由解码器执行的操作序列的图。
- [0038] 图12A是说明根据一般配置的音频信号处理的方法的流程图。
- [0039] 图12B是说明根据一般配置的设备的框图。
- [0040] 图12C是说明根据一般配置的音频信号处理的方法的流程图。
- [0041] 图12D是说明根据一般配置的音频信号处理的方法的流程图。
- [0042] 图13A到13C是说明可执行本发明中描述的技术的各种方面的实例音频回放系统的框图。
- [0043] 图14是说明可执行本发明中描述的技术的各种方面的汽车声音系统的图。

具体实施方式

[0044] 除非通过其上下文明确限制,否则在本文中术语“信号”来指示其一般含义中的任一者,包含如在导线、总线或其它传输媒体上表达的存储器位置的状态(或存储器位置的集合)。除非通过其上下文明确地限制,否则在本文中术语“产生”来指示其一般含义中的任一者,例如计算或以其它方式产生。除非通过其上下文明确限制,否则在本文中术语“计算”来指示其一般含义中的任一者,例如计算、评估、估计,和/或从多个值进行选择。除非通过其上下文明确地限制,否则使用术语“获得”来指示其一般含义中的任一者,例如计算、导出、接收(例如,从外部装置接收),和/或检索(例如,从存储元件的阵列检索)。除非通过其上下文明确地限制,否则使用术语“选择”来指示其一般含义中的任一者,例如识别、指示、应用和/或使用一组两个或两个以上中的至少一者和少于所有者。在本发明的说明书和权利要求书中使用术语“包括”时,并不排除其它元素或操作。术语“基于”(如在“A是基于B”中)用于指示其一般含义中的任一者,包含以下情况(i)“从…导出”(例如,“B是A的前驱体”),(ii)“至少基于”(例如,“A至少基于B”)并且在特定背景中在适当时,(iii)“等于”(例如,“A等于B”)。类似地,使用术语“响应于”来指示其一般含义中的任一者,包含“至少响应于”。

[0045] 对多麦克风音频感测装置的麦克风的“位置”的参考指示所述麦克风的声敏感面的中心的位置,除非上下文另有指示。根据特定上下文,有时使用术语“通道”来指示信号路径且在其它时候指示由此路径载运的信号。除非另有指示,否则使用术语“系列”来指示两个或两个以上项目的序列。使用术语“频率分量”来指示信号的一组频率或频带中的一者,例如(例如,由快速傅里叶变换产生)信号的频域表示的样本或信号的子带(例如,巴克

(Bark) 尺度或梅尔 (mel) 尺度子带)。

[0046] 除非另有指示,否则对具有特定特征的设备的操作的任何揭示内容还明确地希望揭示具有类似特征的方法(且反之亦然),且对根据特定配置的设备的操作的任何揭示内容还明确地希望揭示根据类似配置的方法(且反之亦然)。术语“配置”可参考由其特定上下文指示的方法、设备和/或系统来使用。术语“方法”、“过程”、“程序”和“技术”通用地且可互换地使用,除非特定上下文另有指示。术语“设备”和“装置”也通用地且可互换地使用,除非特定上下文另有指示。术语“元件”和“模块”通常用以指示较大配置的一部分。除非由其上下文明确限制,否则术语“系统”在此用以指示其普通意义中的任一者,包含“交互以用于共同目的的元件群组”。

[0047] 环绕声音的演进如今已使得许多输出格式可用于娱乐。此类环绕声音格式的实例包含流行的5.1格式(其包含以下六个声道:左前(FL)、右前(FR)、中央或前方中央、左后或环绕左边、右后或环绕右边及低频效果(LFE)),发展的7.1格式及未来的22.2格式(例如,用于与超高清电视标准一起使用)。进一步的实例包含用于球面谐波阵列的格式。可能需要环绕声音格式以在两个维度中及/或在三个维度中编码音频。

[0048] 可能需要遵循‘创建一次使用多次’基本原理,其中创建一次音频材料(例如,通过内容创建者)且将其编码为可随后解码及再现到不同输出及扬声器设置的格式。

[0049] 将来MPEG编码器的输入是任选地三种可能的格式中的一者:(i) 传统的基于声道的音频,其打算通过预先指定的位置处的扬声器进行播放;(ii) 基于对象的音频,其涉及具有含有其位置坐标(以及其它信息)的相关联元数据的用于单个音频对象的离散脉码调制(PCM)数据;及(iii) 基于场景的音频,其涉及使用球面谐波基底函数的系数(也称为“球面谐波系数”或SHC)来表示声音场。

[0050] 使用第三种基于场景的格式存在众多的优点。然而,使用此格式的一个可能的缺点是对现有的消费者音频系统缺乏向后兼容性。举例来说,大多数现有系统接受5.1声道输入。传统的基于声道的矩阵式音频可通过具有5.1样本作为扩展声道格式的子集来绕过此问题。在位流中,所述5.1样本处于由现有(或“传统的”)系统辨识的位置中,且额外声道可位于含有所有声道样本的帧包的扩展部分中。或者,可从对较高数目的声道的矩阵化操作确定5.1声道数据。

[0051] 在使用SHC时缺乏向后兼容性是归因于SHC不是PCM数据的事实。本发明描述可用于在使用球面谐波基底函数的系数(也称为“球面谐波系数”或SHC)来表示声音场时解决此向后兼容性的缺乏的方法、系统及设备。

[0052] 市场中存在各种‘环绕声音’格式。它们的范围(例如)是从5.1家庭影院系统(其在使起居室享有立体声方面已获得最大成功)到NHK(日本广播协会或日本广播公司)所开发的22.2系统。内容创建者(例如,好莱坞工作室)将希望产生电影的音轨一次,而不花费精力来针对每一扬声器配置对其进行重混(remix)。可能需要提供成为标准化位流的编码及对扬声器几何条件及再现器的位置处的声学条件适合且不可知的后续解码。

[0053] 图1说明使用移动图片专家组(MPEG)编解码器的此类标准化的通用结构,从而提供均匀的收听体验的所述,而不管最终用于重现的特定设置如何。如图1所示,MPEG编码器10编码音频源12以产生音频源12的经编码版本,其中经由传输通道14将音频源12的经编码版本发送到MPEG解码器16。MPEG解码器16解码音频源12的经编码版本以至少部分恢复音频

源12。在图1的实例中，音频源12的经恢复版本被展示为输出18。

[0054] 甚至在引入立体声格式时向后兼容性也是问题，因为需要旧式单声道回放系统保持兼容性。使用矩阵化来保持单声道立体声向后兼容性。立体声‘M-中间’及‘S-边’格式能够通过仅使用M声道保持与具有单声道功能的系统的兼容性。

[0055] 图2是说明可执行简单的 2×2 矩阵操作以解码‘L-左’及‘R-右’声道的具有立体声功能的系统19的图。可通过使用以上矩阵的相反物从L-R信号计算出M-S信号（其恰好相同）。以此方式，旧式单声道播放器20保持功能性，同时立体声播放器22可准确地解码左及右声道。以类似方式，可添加保持向后兼容性的第三声道，其保留单声道播放器20及立体声播放器22的功能性且添加三声道播放器的功能性。

[0056] 一种用于解决基于对象的格式中的向后兼容性的问题的提出的方法是发送降混的5.1声道信号以及对象。在此情形中，旧式5.1系统将播放降混的基于声道的音频，而更高级的再现器将使用5.1音频与个别音频对象的组合，或仅使用所述个别对象，以再现声音场。

[0057] 可能需要使用分层要素集合来表示声音场。分层要素集合是其中要素经排序以使得排序较低的要素的基础集合提供模型化的声音场的完全表示。随着所述集合经扩展以包含较高阶要素，所述表示变得更详细。

[0058] 分层要素集合的一个实例是SHC集合。以下表达式示范使用SHC对声音场的描述或表示：

$$[0059] \quad p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[0060] 此表达式展示声音场的任何点 $\{r_r, \theta_r, \varphi_r\}$ 处的压力 p_i 可通过SHC $A_n^m(k)$ 唯一地表示。此处， $k = \frac{\omega}{c}$ ， c 是声音的速度（ $\sim 343\text{m/s}$ ）， $\{r_r, \theta_r, \varphi_r\}$ 是参考点（或观测点）， $j_n(\cdot)$ 是阶 n 的球面贝塞尔函数，且 $Y_n^m(\theta_r, \varphi_r)$ 是阶 n 及子阶 m 的球面谐波基底函数。可认识到，方括号中的术语是可通过各种时间-频率变换来近似的对信号的频域表示（即， $S(\omega, r_r, \theta_r, \varphi_r)$ ），所述变换例如为离散傅里叶变换（DFT）、离散余弦变换（DCT）或小波变换。分层集合的其它实例包含小波变换系数的集合及多分辨率基底函数的系数的其它集合。

[0061] 除了在频域中之外，以上等式还表示实现对于不同径向距离（或“半径”）的SHC的导出的球面波模型。也就是说，可针对不同半径 r 导出SHC，这意味着SHC适应在距所谓的“甜点”各种及不同距离处定位或其中收听者既定收听的源。SHC可随后用于确定用于具有驻留在不同球面表面上的扬声器的不规则扬声器几何条件的扬声器馈送，且进而潜在地使用不规则扬声器几何条件的扬声器更好地重现声音场。在这方面中，不接收不在与其它扬声器相同的球面表面上的那些扬声器的径向信息（例如，从甜点到扬声器所测得的半径）且随后引入延迟以补偿波头扩散，可使用以上等式导出SHC以更准确地重现不同径向距离处的声音场。

[0062] 可通过各种麦克风阵列配置物理地获取（例如，记录）SHC $A_n^m(k)$ ，或替代地，可从声音场的基于声道或基于对象的描述导出它们。前者表示输入到所提出的编码器的基于场景的音频。举例来说，可使用包括25个系数的四阶表示。

[0063] 可将对应于个别音频对象的声场的系数 $A_n^m(k)$ 表达为

$$[0064] \quad A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^{m*}(\theta_s, \varphi_s),$$

[0065] 其中 i 是 $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ 是阶 n 的球面汉克尔函数(第二种类), 且 $\{r_s, \theta_s, \varphi_s\}$ 是对象的位置。知晓随频率而变的源能量 $g(\omega)$ (例如, 使用时间-频率分析技术, 例如对 PCM 流执行快速傅里叶变换) 允许我们将每一 PCM 对象及其位置转换为 SHC $A_n^m(k)$ 。此外, 可展示(由于以上是线性及正交分解)用于每一对象的 $A_n^m(k)$ 系数累加的。以此方式, 可通过 $A_n^m(k)$ 系数表示众多的 PCM 对象(例如, 作为用于个别对象的系数向量的总和)。基本上, 这些系数含有关于声场的信息(随 3D 坐标而变的压力), 且以上表示在观测点 $\{r_r, \theta_r, \varphi_r\}$ 附近从个别对象到整个声场的表示的变换。所属领域的技术人员将认识到, 以上表达式可以稍微不同的形式出现在文献中。

[0066] 本发明包含可用于将表示声场的完整分层要素集合(例如, SHC 集合, 其原本可能在向后兼容性不是问题的情况下会使用)的子集(例如, 基础集合)转换到多个音频声道(例如, 表示传统的多声道音频格式)的系统、方法及设备的描述。此方法可应用于所要的任何数目的声道以维持向后兼容性。可预期将实施此方法以维持与至少传统的 5.1 环绕/家庭影院能力的兼容性。对于 5.1 格式, 多声道音频声道是左前、中央、右前、左环绕、右环绕及低频效果(LFE)。SHC 的总数可取决于各种因素。对于基于场景的音频, (例如) SHC 的总数可受到记录阵列中的麦克风换能器的数目约束。对于基于声道及对象的音频, SHC 的总数可由可用的带宽确定。

[0067] 可将经编码声道包装到与所要的对应的基于声道的格式相容的包的对应部分中。分层集合的其余部分(例如, 不是所述子集的部分的 SHC) 将不会被转换而是可经编码以在向后兼容的多声道音频旁边进行传输及/或存储)。举例来说, 可将这些经编码位包装到用于帧的包的扩展部分(例如, 用户定义的部分)中。

[0068] 在另一实施例中, 可对多声道信号执行编码或转码操作。举例来说, 可以 AC3 格式(也称为 ATSC A/52 或杜比数字)译码 5.1 声道以保持与在许多消费型装置及机顶盒中的 AC3 解码器的向后兼容性。甚至在此情况下, 分层集合的其余部分(例如, 不是子集的部分的 SHC) 将被单独地编码且在 AC3 包(例如, 辅助数据)的一或多个扩展部分中传输(及/或存储)。可使用的目标格式的其它实例包含杜比 TrueHD、DTS-HD 大师音频, 及 MPEG 环绕。

[0069] 在解码器处, 旧式系统将忽略帧-包的扩展部分, 仅使用多声道音频内容且因此保持功能性。

[0070] 可实施高级再现器以执行反变换以将多声道音频转换到分层集合的原始子集(例如, SHC 的基础集合)。如果声道已经被再编码或转码, 那么可执行解码的中间步骤。将解码包的扩展部分中的位以提取分层集合的其余部分(例如, SHC 的扩展集合)。以此方式, 可恢复完整分层集合(例如, SHC 的集合)以允许发生各种类型的声音场再现。

[0071] 在以下系统图中使用对编码器及解码器结构两者的解释来概述此向后兼容系统的实例。

[0072] 图 3 是说明根据本发明中描述的技术的方面使用基于场景的球面谐波方法来执行编码及解码过程的系统 30 的框图。在此实例中, 编码器 32 产生源球面谐波系数 34 (“SHC 34”) 的描述, 其被传输(及/或存储)且在解码器 40 (展示为“基于场景的解码器 40”)处被解

码以接收用于再现的SHC 34。此类编码可包含一或多个有损或无损译码过程,例如量化(例如,量化为一或多个码簿索引),误差校正、冗余译码等。另外或替代地,此类编码可包含编码为立体混响格式,例如B-格式、G-格式或较高阶立体混响(HOA)。一般来说,编码器32可使用利用冗余度及不相干(对于有损或无损译码)的已知技术来编码SHC 34以产生经编码的SHC 38。编码器32可常常以位流(其可包含经编码的SHC 38以及在解码经编码的SHC 38中可为有用的其它数据)的形式经由传输通道36传输此经编码的SHC 38。解码器40可接收及解码经编码的SHC 38以恢复SHC 34或其略微修改的版本。解码器40可将经恢复的SHC 34输出到球面谐波再现器42,所述球面谐波再现器42可将经恢复的SHC 34再现为一或多个输出音频信号44。不具有基于场景的解码器40的较旧的接收器可不能够解码此类信号,并且因此可不能够播放节目。

[0073] 图4是说明可执行本发明中描述的技术的各种方面的编码器50的图。源SHC 34(例如,与图3中展示的相同)可为由混合工程师在具有基于场景的功能的录音室中混合的源信号。SHC 34还可由麦克风阵列或由环绕扬声器的声波呈现的记录俘获。

[0074] 编码器50可不同地处理SHC 34集合的两个部分。编码器50可将变换矩阵52应用于SHC 34的基础集合(“基础集合34A”)以产生兼容的多声道信号55。再编码器/转码器56可随后将这些信号55(其可在例如FFT域等频域中或在时域中)编码为描述多声道信号的向后兼容的经译码信号59。兼容的译码器可包含多个实例,例如AC3(也称为ATSC A/52或杜比数字)、杜比TrueHD、DTS-HD大师音频、MPEG环绕。此实施方案还有可能包含两个或更多个不同,每一转码器将所述多声道信号译码为不同的相应的格式(例如,AC3转码器及杜比TrueHD转码器),以产生两个不同的向后兼容的位流以用于传输及/或存储。替代地,可完全省略所述译码以仅输出多声道音频信号作为(例如)线性PCM流的集合(其由HDMI标准支持)。

[0075] SHC 34中的剩余一者可表示SHC 34的扩展集合(“扩展集合34B”)。编码器50可调用基于场景的编码器54以编码基础集合34B,这产生位流57。编码器50可随后调用位多路复用器58(“位多路复用器58”)以对向后兼容的位流59及位流57进行多路复用。编码器50可随后经由传输通道(例如,有线及/或无线通道)发送此经多路复用的位流61。

[0076] 图5是说明仅支持标准的非基于场景的解码但其能够恢复根据本发明中描述的技术所形成的向后兼容的位流59的标准解码器70的图。换句话说,在解码器70处,如果接收器较旧且仅支持常规的解码器,那么解码器将仅采用向后兼容的位流59且丢弃扩展位流57,如图5中所展示。在操作中,解码器70接收经多路复用的位流61且调用位多路分用器(“位多路分用器72”)。位多路分用器72对经多路复用的位流61进行多路分用以恢复向后兼容的位流59及扩展位流57。解码器70随后调用向后兼容的解码器74以解码向后兼容的位流59且进而产生输出音频信号75。

[0077] 图6是说明可执行本发明中描述的技术的各种方面的另一解码器80的图。在接收器是新且支持基于场景的解码时,在图6中展示解码过程,其是与图4的编码器互逆的过程。类似于解码器70,解码器80包含位多路分用器72,所述位多路分用器72对经多路复用的位流61进行多路分用以恢复向后兼容的位流59及扩展位流57。然而,解码器80可随后调用转码器82以转码向后兼容的位流59且恢复多声道兼容的信号55。解码器80可随后将反变换矩阵84应用于多声道兼容的信号55以恢复基础集合34A’(其中撇号(‘)表示此基础集合34A’

与基础集合34A相比可能有略微修改)。解码器80还可调用基于场景的解码器86,其可解码扩展位流57以恢复扩展集合34B' (其中再次地,撇号(')表示此扩展集合34B' 与扩展集合34B相比可能有略微修改)。在任何情况下,解码器80可调用球面谐波再现器88以再现基础集合53A' 与扩展集合53B' 的组合以产生输出音频信号90。

[0078] 换句话说,如果可适用,转码器82将向后兼容的位流59转换为多声道信号55。随后,通过反矩阵84处理这些多声道信号55以恢复基础集合34A'。通过基于场景的解码器86恢复扩展集合34B'。SHC 34' 的完整集合被SH再现器88组合及处理。

[0079] 此实施方案的设计可包含选择原始分层集合的将被转换到多声道音频(例如,转换到常规的格式)的子集。可能出现的另一问题是在从基础集合(例如,SHC的基础集合)到多声道音频及回到基础集合的前向及后向转换中产生多少错误。

[0080] 对以上各者的各种解决方案是可能的。在以下论述中,5.1格式将用作典型目标多声道音频格式,且将详细描述实例方法。所述方法可一般化到其它多声道音频格式。

[0081] 由于在5.1格式中五个信号(对应于来自指定位置的全频带音频)是可用的(加上LFE信号,其不具有标准化的位置且可通过对所述五个声道进行低通滤波来确定),所以一种方法是使用五个SHC以转换到5.1格式。此外,由于5.1格式仅能够2D再现,所以可能需要仅使用携带一些水平信息的SHC。举例来说,系数 $A_1^0(k)$ 携带关于水平方向性的极少的信息且可因此被排除出此子集。 $A_2^1(k)$ 的实部或虚部也是如此。这些中的一些依据在实施方案中选择的球面谐波基底函数的定义而变化(在文献中存在各种定义-实数、虚数、复数或组合)。以此方式,可选取五个 $A_n^m(k)$ 系数用于转换。因为系数 $A_0^0(k)$ 携带全向信息,所以可能需要始终使用此系数。类似地,可能需要包含 $A_1^1(k)$ 的实部及 $A_1^{-1}(k)$ 的虚部,因为它们携带重要的水平方向性信息。对于最后两个系数,可能的候选者包含 $A_2^2(k)$ 的实部及虚部。各种其它组合也是可能的。举例来说,可选择基础集合以仅包含三个系数 $A_0^0(k)$ 、 $A_1^1(k)$ 的实部及 $A_1^{-1}(k)$ 的虚部。

[0082] 接下来的步骤是确定可在SHC的基础集合(例如,上文选择的五个系数)与5.1格式中的五个全频带音频信号之间进行转换的可逆矩阵。对可逆性的需要是允许在分辨率的极少损耗或没有损耗的情况下将五个全频带音频信号转换回到SHC的基础集合。

[0083] 一种确定此矩阵的可能的方法是被称为'模式匹配'的操作。此处,通过假定每一扩音器产生一球面波来计算扩音器馈送。在此情形中,通过下式给出归因于第 ℓ 个扩音器而引起的在一定位置 r, θ, φ 处的压力(随频率而变)

$$[0084] \quad P_\ell(\omega, r, \theta, \varphi) = g_\ell(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi i k) h_n^{(2)}(kr_\ell) Y_n^{m*}(\theta_\ell, \varphi_\ell) Y_n^m(\theta, \varphi),$$

[0085] 其中 $\{r_\ell, \theta_\ell, \varphi_\ell\}$ 表示第 ℓ 个扩音器的位置且 $g_\ell(\omega)$ 是第 ℓ 个扬声器的扩音器馈送(在频域中)。因此通过下式给出总压力归因于所有五个扬声器的总压力 P_t

$$[0086] \quad P_t(\omega, r, \theta, \varphi) = \sum_{\ell=1}^5 g_\ell(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi i k) h_n^{(2)}(kr_\ell) Y_n^{m*}(\theta_\ell, \varphi_\ell) Y_n^m(\theta, \varphi)。$$

[0087] 我们还知道通过以下等式给出在五个SHC方面的总压力

$$[0088] \quad P_t(\omega, r, \theta, \varphi) = 4\pi \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta, \varphi)$$

[0089] 使以上两个等式相等允许我们使用变换矩阵来表达在SHC方面的扩音器馈送,如下:

$$[0090] \quad \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ A_2^2(\omega) \\ A_2^{-2}(\omega) \end{bmatrix} = -ik \begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2)Y_0^{0*}(\theta_2, \varphi_2) & \cdot & \cdot & \cdot \\ h_1^{(2)}(kr_1)Y_1^{1*}(\theta_1, \varphi_1) & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ g_4(\omega) \\ g_5(\omega) \end{bmatrix}.$$

[0091] 此表达式展示在五个扩音器馈送与所选择的SHC之间存在直接关系。变换矩阵可依据(例如)哪一SHC用于子集(例如,基础集合)中及使用SH基底函数的哪一定义而变化。以类似方式,可建构用以从所选择的基础集合转换到不同声道格式(例如,7.1、22.2)的变换矩阵

[0092] 虽然以上表达式中的变换矩阵允许从扬声器馈送到SHC的转换,那么我们希望矩阵是可逆的,以使得以SHC开始,我们计算出五个声道馈送且随后在解码器处,我们可任性地转换回到SHC(在存在高级(即,非旧式)再现器时)。

[0093] 可采用操纵以上框架以确保矩阵的可逆性的各种方式。这些包括但不限于:改变扬声器的位置(例如,调整5.1系统的五个扬声器中的一或多者的位置以使得它们仍遵守由ITU-R BS.775-1标准指定的角度公差;换能器的规则间距(例如,遵守T-设计的换能器)通常良好地表现);规则化技术(例如,频率相依规则化);及常常作用以确保满秩及良好定义的本征值的各种其它矩阵操纵技术。最后,可能需要在精神-声学上测试5.1再现以确保在所有操纵之后,经修改的矩阵实际上产生正确及/或可接受的扩音器馈送。只要保留可逆性,确保对SHC的正确解码的相反问题就不是问题。

[0094] 对于一些局部扬声器几何条件(其可指解码器处的扬声器几何条件),以上概述的用以操纵以上框架以确保可逆性的方式可导致大不合意的音频-图像质量。也就是说,当与所俘获的音频相比较时,声音重现可并不总是导致声音的正确定位。为了对此不大合意的图像质量进行校正,可进一步扩充技术以引入可被称作“虚拟扬声器”的概念。可修改以上框架以包含某一形式的平移,例如向量基础振幅平移(VBAP)、基于距离的振幅平移或其它形式的平移,而不是需要将一或多个扩音器再定位或定位于具有由例如上述的ITU-R BS.775-1等标准指定的特定角度公差的空间的特定或所定义的区中。出于说明的目的而聚焦于VBAP上,VBAP实际上可引入可被表征为“虚拟扬声器”的东西。VBAP可修改到一或多个扩音器的馈送以使得这些一或多个扩音器实际上输出显得源自虚拟扬声器的声音,所述虚拟扬声器处于不同于支持所述虚拟扬声器的一或多个扩音器的位置及/或角度中的至少一者的位置及角度中的一或多者处。

[0095] 为进行说明,用于根据SHC确定扩音器馈送的以上等式可如下修改:

$$[0096] \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX \\ M \times N \end{bmatrix} \begin{bmatrix} D \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix}。$$

[0097] 在上文等式中,VBAP矩阵具有M行×N列的大小,其中M表示扬声器的数目(且在以上等式中将等于五),且N表示虚拟扬声器的数目。可依据从收听者的经定义位置到扬声器的位置中的每一者的向量及从收听者的经定义位置到虚拟扬声器的位置中的每一者的向量来计算VBAP矩阵。以上等式中的D矩阵可具有N行×(阶数+1)²列的大小,其中阶数可指代SH函数的阶数。D矩阵可表示以下等式矩阵:

$$[0098] \begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2)Y_0^{0*}(\theta_2, \varphi_2) & \cdot & \cdot & \cdot \\ h_1^{(2)}(kr_1)Y_1^{1*}(\theta_1, \varphi_1) & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}。$$

[0099] 实际上,VBAP矩阵为M×N矩阵,其提供可被称作“增益调整”的调整,所述调整将扬声器的位置及虚拟扬声器的位置考虑在内。以此方式引入平移可导致多声道音频的较好重现,这导致在由局部扬声器几何条件重现时的较好质量图像。此外,通过将VBAP并入到此等式中,技术可克服不与各种标准中所指定的几何条件对准的不佳扬声器几何条件。

[0100] 实际上,所述等式可反转且用于将SHC变换回到用于扩音器的特定几何条件或配置(其可在下文被称为几何条件B)的多声道馈送。也就是说,可反转等式以求解g矩阵。经反转的等式可如下:

$$[0101] \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP^{-1} \\ MATRIX^{-1} \\ M \times N \end{bmatrix} \begin{bmatrix} D^{-1} \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix}。$$

[0102] g矩阵可表示在此实例中用于5.1扬声器配置中的五个扩音器中的每一者的扬声器增益。此配置中所使用的虚拟扬声器位置可对应于5.1多声道格式规格或标准中所定义的位置。可使用任何数目的已知音频定位技术来确定可支持这些虚拟扬声器中的每一者的扩音器的位置,所述技术中的许多技术涉及播放具有特定频率的音调以确定每一扩音器相对于头端单元(例如音频/视频接收器(A/V接收器)、电视、游戏系统、数字视频光盘系统或其它类型的头端系统)的位置。或者,头端单元的用户可手动地指定扩音器中的每一者的位置。在任何情况下,在给定这些已知位置及可能角度的情况下,假定虚拟扩音器的通过VBAP的理想配置,可求解头端单元的增益。

[0103] 在这方面中,所述技术可使得装置或设备能够对第一多个扩音器声道信号执行向量基础振幅平移或其它形式的平移以产生第一多个虚拟扩音器声道信号。这些虚拟扩音器声道信号可表示提供给扩音器的使得这些扩音器能够产生显得源自虚拟扩音器的声音的信号。因此,在对第一多个扩音器声道信号执行第一变换时,所述技术可使得装置或设备能够对第一多个虚拟扩音器声道信号执行第一变换以产生描述声音场的分层要素集合。

[0104] 此外,所述技术可使得设备能够对分层要素集合执行第二变换以产生第二多个虚拟扩音器声道,其中所述第二多个虚拟扩音器声道中的每一者与空间的对应不同区相关联。在一些情况下,所述技术可使得装置能够对所述第二多个虚拟扩音器声道信号执行向量基础振幅平移以产生第二多个扩音器声道信号。

[0105] 虽然以上变换矩阵是从‘模式匹配’准则导出的,但还可从例如压力匹配、能量匹配等其它准则导出替代性变换矩阵。可导出以下矩阵是足够的:允许基础集合(例如,SHC子集)与传统的多声道音频之间的变换,并且还在操纵(其不减少多声道音频的保真度)之后,还可制定也是可逆的略微修改的矩阵。

[0106] 以上部分论述用于5.1兼容系统的设计。可针对不同目标格式相应地调整细节。作为一实例,为了实现对7.1系统的兼容性,根据兼容要求添加两个额外的音频内容声道,且可将另外两个SHC添加到基础集合,使得所述矩阵是可逆的。由于用于7.1系统(例如,杜比TrueHD)的大部分扩音器布置仍在水平平面上,所以对SHC的选择可仍排除具有高度信息的SHC。以此方式,水平平面信号再现将受益于所述再现系统中的添加的扩音器声道。在包含具有高度多样性的扩音器的系统(例如,9.1、11.1及22.2系统)中,可能需要在基础集合中包含具有高度信息的SHC。

[0107] 对于较低数目的声道(比如,立体声及单声道),许多现有技术中的现有的5.1解决方案应足以覆盖以维持内容信息。这些情况被视为不重要的,且未在本发明中进一步论述。

[0108] 以上内容因此表示用以在分层要素集合(例如,SHC集合)与多个音频声道之间转换的无损机制。只要多声道音频信号不遭受进一步的译码噪声,就不招致错误。在它们遭受译码噪声的情况下,到SHC的转换可招致错误。然而,有可能通过监视系数的值来考虑这些错误并采取适当的动作以减少其影响。这些方法可考虑到SHC的特性,包含SHC表示中的固有冗余。

[0109] 虽然我们已经一般化到多声道,当前市场中的主要重点是针对5.1声道,因为它是‘最小公分母’以确保例如机顶盒等旧式消费者音频系统的功能性。

[0110] 本文中描述的方法提供对在声音场的基于SHC的表示的使用中的潜在缺点的解决方案。在没有此解决方案的情况下,归因于由于不能够在数百万旧式回放系统中具有功能性而导致的重大缺点,可能永远部署不了基于SHC的表示。

[0111] 图7A是说明根据一般配置的音频信号处理的方法M100的流程图,所述方法包含与本发明中描述的技术的各种方面一致的任务T100、T200及T300。任务T100将声音场的描述(例如,一组SHC)划分为基础要素集合(例如,图4的实例中展示的基础集合34A),及扩展要素集合(例如,扩展集合34B)。任务T200对基础集合34A执行例如变换矩阵52等可逆变换以产生多个声道信号55,其中所述多个声道信号55中的每一者与空间的对应不同区相关联。任务T300产生包,所述包包含描述所述多个声道信号55的第一部分及描述扩展集合34B的第二部分(例如,辅助数据部分)。

[0112] 图7B是说明根据与本发明中描述的技术的各种方面一致的一般配置的设备MF100的框图。设备MF100包含用于产生包含基础要素集合(例如,图4的实例中展示的基础集合34A)及扩展要素集合34B(如本文例如参考任务T100所描述)的对声音场的描述的装置F100。设备MF100还包含用于对基础集合34A执行例如变换矩阵52等可逆变换以产生多个声道信号55的装置F200,其中所述多个声道信号55中的每一者与空间的对应不同区相关联

(如本文例如参考任务T200所描述)。设备MF100还包含用于产生包的装置F300,所述包包含描述所述多个声道信号55的第一部分及描述扩展要素集合34B的第二部分(如本文例如参考任务T300所描述)。

[0113] 图7C是根据与本发明中描述的技术的各种方面一致的另一一般配置的用于音频信号处理的设备A100的框图。设备A100包含编码器100,所述编码器经配置以产生包含基础要素集合(例如,图4的实例中展示的基础集合34A)及扩展要素集合34B(如本文例如参考任务T100所描述)的对声音场的描述。设备A100还包含变换模块200,所述变换模块经配置以对基础集合34A执行例如变换矩阵52等可逆变换以产生多个声道信号55,其中所述多个声道信号55中的每一者与空间的对应不同区相关联(如本文例如参考任务T200所描述)。设备A100还包含打包器300,所述打包器经配置以产生包,所述包包含描述所述多个声道信号55的第一部分及描述扩展要素集合34B的第二部分(如本文例如参考任务T300所描述)。

[0114] 图8A是说明根据一般配置的音频信号处理的方法M100的流程图,所述方法包含表示本发明中描述的技术的一个实例的任务T400及T500。任务T400将包划分为:描述多个声道信号(例如,图5和6的实例中展示的信号55)的第一部分,每一声道信号与空间的对应不同区相关联;及描述扩展要素集合(例如,图5的实例中展示的基础集合34A)的第二部分。任务T500对多个声道信号55执行例如反变换矩阵84等反变换以恢复基础要素集合34A'。在此方法中,基础集合34A'包括描述声音场的分层要素集合的较低阶部分(例如,一组SHC),且扩展要素集合34B'包括所述分层集合的较高阶部分。

[0115] 图8B是说明包含任务T505及T605的方法M100的实施方案M300的流程图。对于多个音频信号(例如,音频对象)中的每一者,任务T505将信号及用于所述信号的空间信息编码到描述声音场的对应的分层要素集合中。任务T605组合所述多个分层集合以产生将在任务T100中处理的的声音场的描述。举例来说,可实施任务T605以添加所述多个分层集合(例如,执行系数向量添加)以产生对组合声音场的描述。用于一个对象的分层要素集合(例如,SHC向量)可比用于对象中的另一者的分层要素集合具有更高阶(例如,更长的长度)。举例来说,可使用比后台中的对象(例如,音效)更高阶的集合来表示前台中的对象(例如,男主角的话音)。

[0116] 本文中揭示的原理还可用于实施用以补偿基于声道的音频方案中的扩音器几何条件中的差异的系统、方法及设备。举例来说,通常专业音频工程师/能手使用呈一定几何条件(“几何条件A”)的扬声器来混合音频。可能需要产生用于一定替代性扩音器几何条件(“几何条件B”)的扩音器馈送。本文中揭示的技术(例如,参考扩音器馈送与SHC之间的变换矩阵)可用于将扩音器馈送从几何条件A转换为SHC且随后将其重新再现为扩音器几何条件B。在一个实例中,几何条件B是任意所要的几何条件。在另一实例中,几何条件B是标准化的几何条件(例如,如标准文献(例如,ITU-R BS.775-1标准)中所指定)。也就是说,此标准化的几何条件可界定每一扬声器将位于的空间的位置或区。由标准界定的空间的这些区可被称为空间的界定区。此方法可用于不仅补偿几何条件A与B之间的扬声器中的一或多个者相对于收听者的距离(半径)上的差异,并且还补偿一或多个扩音器相对于收听者的方位角及/或仰角上差异。可在编码器处及/或解码器处执行此转换。

[0117] 图9A是说明根据本发明中描述的技术的各种方面通过应用变换矩阵102从SHC100到与特定几何条件兼容的多声道信号104的如上文所描述的转换的图。

[0118] 图9B是说明根据本发明中描述的技术的各种方面通过应用变换矩阵106(其可为变换矩阵102的倒置形式)从与特定几何条件兼容的多声道信号104恢复SHC 100'的如上文所描述的转换的图。

[0119] 图9C是说明根据本发明中描述的技术的各种方面通过应用如上文所描述的变换矩阵A108从与几何条件A兼容的多声道信号104恢复SHC 100'的第一转换及通过应用变换矩阵110从SHC 100'到与几何条件B兼容的多声道信号112的第二转换的图。应注意,如图9C中所说明的实施方案可经扩展以包含从SHC到与其它几何条件兼容的多声道信号的一或多个额外转换。

[0120] 在基础情况下,几何条件A及B中的声道的数目是相同的。应注意,对于所述几何条件转换应用,有可能放宽上文所描述的约束条件以确保变换矩阵的可逆性。进一步的实施方案包含其中几何条件A中的声道的数目比几何条件B中的声道的数目多或少的系统、方法及设备。

[0121] 图10A是说明包含与本发明中描述的技术的各种方面一致的任务T600及T700的根据一般配置的音频信号处理的方法M400的流程图。任务T600对第一多个声道信号(例如,信号104)执行第一变换(例如,图9C中展示的变换矩阵A 108)以产生描述声音场的分层要素集合(例如,经恢复的SHC 100'),其中第一多个声道信号104中的每一者与空间的对应不同区相关联(例如,如参考图9B及9C所描述)。任务T700对分层要素集合100'执行第二变换(例如,变换矩阵110)以产生第二多个声道信号112,其中第二多个声道信号112中的每一者与空间的对应不同区相关联(例如,如本文参考任务T200及图4、9A及9C所描述)。

[0122] 图10B是说明根据一般配置的用于音频信号处理的设备MF400的框图。设备MF400包含用于对第一多个声道信号(例如,信号104)执行第一变换(例如,图9C的实例中展示的变换矩阵A 108)以产生描述声音场的分层要素集合(例如,经恢复的SHC 100')的装置F600,其中第一多个声道信号104中的每一者与空间的对应不同区相关联(如本文例如参考任务T600所描述)。设备MF100还包含用于对分层要素集合100'执行第二变换(例如,变换矩阵B 110)以产生第二多个声道信号112的装置F700,其中第二多个声道信号112中的每一者与空间的对应不同区相关联(例如,如本文例如参考任务T200及T700所描述)。

[0123] 图10C是说明根据与本发明中描述的技术一致的另一一般配置的用于音频信号处理的设备A400的框图。设备A400包含第一变换模块600,其经配置以对第一多个声道信号(例如,信号104)执行第一变换(例如,变换矩阵A 108)以产生描述声音场的分层要素集合(例如,经恢复的SHC 100'),其中第一多个声道信号104中的每一者与空间的对应不同区相关联(如本文例如参考任务T600所描述)。设备A100还包含第二变换模块250,其经配置以对分层要素集合100'执行第二变换(例如,变换矩阵B 110)以产生第二多个声道信号112,其中第二多个声道信号112中的每一者与空间的对应不同区相关联(例如,如本文例如参考任务T200及T600所描述)。第二变换模块250可实现为(例如)变换模块200的一实施方案。

[0124] 图10D是说明包含编码器122的系统120的实例的图,所述编码器接收输入声道123(例如,PCM流集合,每一PCM流对应于不同声道)且产生对应的经编码信号125以用于经由传输通道126进行传输(且/或存储到存储媒体(例如,DVD盘)(虽然为了便于说明的目的而未展示))。此系统120还包含解码器124,所述解码器接收经编码信号125且根据特定扩音器几何条件而产生扩音器馈送127的对应集合。在一个实例中,编码器122经实施以执行如图9C

中所说明的程序,其中输入声道对应于几何条件A且经编码信号125描述对应于几何条件B的多声道信号。在另一实例中,解码器124知晓几何条件A且经实施以执行如图9C中所说明的程序。

[0125] 图11A是说明另一系统130的实例的图,所述另一系统包含编码器132,所述编码器接收对应于几何条件A的输入声道133的集合且产生对应的经编码信号135以用于与对应几何条件A的描述(例如,空间中的扬声器的坐标的描述)一起经由传输通道136进行传输(且/或用于存储到存储媒体,例如DVD盘)。此系统130还包含解码器134,所述解码器接收经编码信号135及几何条件A描述且根据不同扩音器几何条件B而产生扩音器馈送137的对应集合。

[0126] 图11B是使用从多声道信号140到SHC 142的第一转换(通过应用如上文所描述的变换矩阵A 144及从SHC 142到与几何条件B兼容的多声道信号148的第二转换(通过应用变换矩阵B 146)的可由解码器134执行的操作序列的图说明,所述第一转换是根据几何条件A的描述141自适应的(例如,通过第一变换模块600的对应实施方案)。所述第二转换对于特定几何条件B可为固定的,或也可根据所要的几何条件B(例如,如提供到第二变换模块250的对应实施方案)的描述(为了便于说明的目的而在图11B的实例中未展示)而为自适应的。

[0127] 图12A是说明包含任务T800及T900的根据一般配置的音频信号处理的方法M500的流程图。任务T800使用第一变换(例如图11B的实例中展示的变换矩阵A 144)将来自第一扬声器几何条件的第一组音频声道信息(例如,信号140)变换为描述声音场的第一分层要素集合(例如,SHC 142)。任务T900使用第二变换(例如,变换矩阵B 146)将第一分层要素集合144变换为用于第二扬声器几何条件的第二组音频声道信息148。所述第一和第二几何条件可具有(例如)不同半径、方位角及/或仰角。

[0128] 图12B是说明根据一般配置的设备A500的框图。设备A500包含处理器150,所述处理器经配置以对来自第一扬声器几何条件的第一组音频声道信息(例如,信号140)执行到描述声音场的第一分层要素集合(例如,SHC 144)的第一变换(例如,图11B的实例中展示的变换矩阵A 144)。设备A500还包含经配置以存储第一组音频声道信息的存储器152。

[0129] 图12C是说明根据一般配置的音频信号处理的方法M600的流程图,所述方法接收扩音器声道(例如,图11B的实例中展示的信号140)以及第一扬声器几何条件的坐标(例如,描述141),其中扩音器声道已经被转换为分层要素集合(例如,SHC 144)。

[0130] 图12D是说明根据一般配置的音频信号处理的方法M700的流程图,所述方法传输扩音器声道(例如,图11B的实例中展示的信号140)以及第一扬声器几何条件的坐标(例如,描述141),其中所述第一几何条件对应于所述声道的位置。

[0131] 图13A到13C是说明可执行本发明中描述的技术的各种方面的实例音频回放系统200A到200C的框图。在图13A的实例中,音频回放系统200A包含音频源装置212、头端装置214、左前扬声器216A、右前扬声器216B、中央扬声器216C、左边环绕声音扬声器216D及右边环绕声音扬声器216E。虽然展示为包含专用扬声器216A到216E(“扬声器216”),但可在其中使用包含扬声器的其它装置来取代专用扬声器216的情况下执行所述技术。

[0132] 音频源装置212可表示能够产生源音频数据的任何类型的装置。举例来说,音频源装置212可表示电视机(包含所谓的“智能电视”或“smarTV”(其具有因特网接入的特征及/或其执行能够支持应用的执行的操作系统)、数字机顶盒(STB)、数字视频光盘(DVD)播放器、高清晰度光盘播放器、游戏系统、多媒体播放器、流式传输多媒体播放器、记录播放器、

桌上型计算机、膝上型计算机、平板电脑 (tablet) 或平板计算机 (slate computer)、蜂窝式电话 (包含所谓的“智能电话”), 或能够产生或以其它方式提供源音频数据的任何其它类型的装置或组件。在一些情况下, 例如在音频源装置212表示电视、桌上型计算机、膝上型计算机、平板电脑 (tablet) 或平板计算机 (slate computer) 或蜂窝式电话的情况下, 音频源装置212可包含显示器。

[0133] 头端装置214表示能够处理 (或, 换句话说, 再现) 由音频源装置212产生或以其它方式提供的源音频数据的任何装置。在一些情况下, 头端装置214可与音频源装置212集成以形成单一装置, 例如, 以使得音频源装置212在头端装置214的内部或是其部分。为了说明, 在音频源装置212表示电视、桌上型计算机、膝上型计算机、平板 (slate) 或平板 (tablet) 计算机、游戏系统、移动电话或高清晰度光盘播放器 (提供几个实例) 时, 音频源装置212可与头端装置214集成。也就是说, 头端装置214可为例如电视、桌上型计算机、膝上型计算机、平板 (slate) 或平板 (tablet) 计算机、游戏系统、蜂窝式电话或高清晰度光盘播放器或其类似者等多种装置中的任一者。头端装置214在未与音频源装置212集成时可表示提供若干接口的音频/视频接收器 (其通常被称为“A/V接收器”), 通过所述若干接口经由有线或无线连接与音频源装置212及扬声器216进行通信。

[0134] 扬声器216中的每一者可表示具有一或多个换能器的扬声器。通常, 左前扬声器216A类似于右前扬声器216B或几乎与其相同, 同时环绕左边扬声器216D类似于环绕右边扬声器216E或几乎与其相同。扬声器216可提供借以与头端装置214进行通信的有线及/或 (在一些情况下) 无线接口。扬声器216可被有源供电或无源供电, 其中在被无源供电时, 头端装置214可驱动扬声器216中的每一者。

[0135] 在典型多声道声音系统 (其也可被称作“多声道环绕声音系统”或“环绕声音系统”) 中, 可表示头端装置214的一个实例的A/V接收器处理源音频数据以适应专用的左前、前方中央、右前、左后 (其也可被称作“环绕左边”)) 及右后 (其也可被称作“环绕右边”) 扬声器216的放置。A/V接收器常常提供到这些扬声器中的每一者的专用有线连接以便提供更好的音频质量、向所述扬声器供电并减少干扰。A/V接收器可经配置以将适当的声道提供到扬声器216中的适当一者。

[0136] 存在复制声音的一级或区域且进而更好地呈现更沉浸的声音体验的若干不同环绕声音格式。在5.1环绕声音系统中, A/V接收器再现音频的五个声道, 其包含中央声道、左声道、右声道、右后声道及左后声道。形成5.1的“.1”的额外声道针对于低音炮或巴斯声道。其它环绕声音格式包含7.1环绕声音格式 (其添加额外的左后及右后声道) 及22.2环绕声音格式 (其除了额外的前及后声道之外在不同高度处添加额外声道及另一低音炮或巴斯声道)。

[0137] 在5.1环绕声音格式的背景中, A/V接收器可再现所述五个扬声器216的这些五个声道及低音炮的巴斯声道 (在图13A或13B的实例中未展示)。A/V接收器可再现所述信号以改变所述信号的音量水平及其它特性以便在其中环绕声音系统操作的特定房间中充分复制声音场。也就是说, 原始环绕声音音频信号可能已经被俘获及处理以适应给定房间, 例如15×15英尺房间。A/V接收器可处理此信号以适应环绕声音系统在其中操作的房间。A/V接收器可执行此再现以产生较好的声音级且进而提供更好或更沉浸的收听体验。

[0138] 在图13B的实例中, 扬声器216布置在由虚线矩形指示的矩形扬声器几何条件218

中。此扬声器几何条件可类似于由上述各种音频标准中的一或多个者指定的扬声器几何条件或几乎与其相同。给定与标准化的扬声器几何条件的相似性,头端装置214可不变换音频信号220或另外以上文所描述的方式将所述音频信号转换为SHC,但可仅经由扬声器216回放这些音频信号220。

[0139] 然而,头端装置214可为可配置的,以甚至在扬声器几何条件218类似于但不等同于上述标准中的一者中指定的几何条件时执行此变换,以便潜在地产生更好地再现既定声音场的扬声器馈送。在这方面中,虽然类似于那些扬声器几何条件,但头端装置214可仍执行本发明中上文所描述的技术以更好地重现声音场。

[0140] 在图13B的实例中,系统200B与系统200A类似之处在于系统200B还包含音频源装置212、头端装置214及扬声器216。然而,并非具有以矩形扬声器几何条件218布置的扬声器216,系统200B具有以不规则扬声器几何条件222布置的扬声器216。不规则扬声器几何条件222可表示不对称扬声器几何条件的一个实例。

[0141] 由于此不规则扬声器几何条件222,用户可与头端装置214介接以输入扬声器216中的每一者的位置,以使得头端装置214能够指定所述不规则扬声器几何条件222。给定扬声器216的不规则扬声器几何条件222,头端装置214可随后执行上文所描述的技术以将输入音频信号220变换到SHC,且随后将SHC变换到可最佳地重现声音场的扬声器馈送。

[0142] 在图13C的实例中,系统200C与系统200A及200B的类似之处在于系统200C还包含音频源装置212、头端装置214及扬声器216。然而,并非具有以矩形扬声器几何条件218布置的扬声器216,系统200C具有以多平面扬声器几何条件226布置的扬声器216。多平面扬声器几何条件226可表示不对称多平面扬声器几何条件的一个实例,其中至少一个扬声器不驻留在与其它扬声器216中的两者或更多者相同的平面上,例如,在图13C的实例中的平面228。如图13C的实例中所展示,右边环绕扬声器216E具有从平面228到扬声器216E的位置的垂直移位230。剩余的扬声器216A到216D各自位于对于扬声器216A到216D中的每一者可为共同的平面228上。然而,扬声器216E驻留在与扬声器216A到216D不同的平面上并且因此扬声器216驻留在两个或更多个或(换句话说)多个平面上。

[0143] 由于此多平面扬声器几何条件228,用户可与头端装置214介接以输入扬声器216中的每一者的位置,以使得头端装置214能够指定所述多平面扬声器几何条件226。给定扬声器216的多平面扬声器几何条件226,头端装置214可随后执行上文所描述的技术以将输入音频信号220变换到SHC,且随后将SHC变换到可最佳地重现声音场的扬声器馈送。

[0144] 图14是说明可执行本发明中描述的技术的各种方面的汽车声音系统250的图。如图14的实例中所展示,汽车声音系统250包含可实质上类似于在图13A到13C的实例中展示的上述音频源装置212的音频源装置252。汽车声音系统250还可包含头端装置254(“H/E装置254”),其可实质上类似于上文所描述的头端装置214。虽然展示为位于汽车251的前仪表板上,但音频源装置252及头端装置254中的一或两者可位于汽车251内的任何地方,包含(作为实例)汽车的地板、顶板或后舱。

[0145] 汽车声音系统250进一步包含前扬声器256A、驾驶员侧扬声器256B、乘客侧扬声器256C、后扬声器256D、周围扬声器256E及低音炮258。虽然未个别地指示,但在图14的实例中的每一圆圈及或扬声器状物体表示单独的或个别的扬声器。然而,虽然操作为各自接收其自身的扬声器馈送的单独的扬声器,但所述扬声器中的一或多个者可结合另一扬声器而操作

以提供位于所述扬声器的两个协作者之间某处的可被称为虚拟扬声器的东西。

[0146] 在这方面中,前扬声器256A中的一或多者可表示中央扬声器,其类似于图13A到13C的实例中展示的中央扬声器216C。前扬声器256A中的一或多者还可表示左前扬声器,其类似于左前扬声器216A,同时前扬声器256A中的一或多者在一些情况下可表示右前扬声器,其类似于右前扬声器216B。在一些情况下,驾驶员侧扬声器256B中的一或多者可表示类似于右前扬声器216B的右前扬声器。在一些情况下,前扬声器256A及驾驶员侧扬声器256B两者中的一或多者可表示类似于左前扬声器216A的左前扬声器。同样,在一些情况下,乘客侧扬声器256C中的一或多者可表示类似于右前扬声器216B的右前扬声器。在一些情况下,前扬声器256A及乘客侧扬声器256C两者中的一或多者可表示类似于右前扬声器216B的右前扬声器。

[0147] 此外,驾驶员侧扬声器256B中的一或多者在一些情况下可表示类似于环绕左扬声器216D的环绕左扬声器。在一些情况下,后扬声器256D中的一或多者可表示类似于环绕左扬声器216D的环绕左扬声器。在一些情况下,驾驶员侧扬声器256B及后扬声器256D两者中的一或多者可表示类似于环绕左扬声器216D的环绕左扬声器。同样,乘客侧扬声器256C中的一或多者在一些情况下可表示类似于环绕右扬声器216E的环绕右扬声器。在一些情况下,后扬声器256D中的一或多者可表示类似于环绕右扬声器216E的环绕右扬声器。在一些情况下,乘客侧扬声器256C及后扬声器256D两者中的一或多者可表示类似于环绕右扬声器216E的环绕右扬声器。

[0148] 周围扬声器256E可表示安装在汽车251的地板中、汽车251的顶板中或汽车251的任何其它可能的内部空间(包含汽车251内的座椅、任何控制台或其它舱)中的扬声器。低音炮258表示经设计以重现低频效果的扬声器。

[0149] 头端装置254可执行上文所描述的技术的各种方面以变换来自音频源装置252的可使用扩展集合扩充的向后兼容的信号,以恢复表示声音场(常常表示声音场的三维表示,如上所述)的SHC。由于可被表征为声音场的全面表示的东西,头端装置254可随后变换所述SHC以产生用于扬声器256A到256E中的每一者的个别馈送。头端装置254可以此方式产生扬声器馈送以使得在经由扬声器256A到256E播放时,与使用符合一标准的标准化的扬声器馈送重现声音场(作为一个实例)相比,可更好地再现声音场(尤其在给定相对大数目的扬声器256A到256E的情况下,与通常具有至多10到16个扬声器的特征的一般汽车声音系统相比)

[0150] 本文中揭示的方法和设备可一般应用于任何收发及/或音频感测应用中,包含来自远场源的信号分量的此类应用及/或感测的移动或另外便携的实例。举例来说,本文中揭示的配置的范围包含驻留在经配置以使用码分多址(CDMA)空中接口的无线电话通信系统中的通信装置。尽管如此,所属领域的技术人员将理解,具有如本文所描述的特征的方法及设备可驻留在采用所属领域的技术人员已知的广泛范围的技术的各种通信系统中的任一者中,例如采用经由有线及/或无线(例如,CDMA、TDMA、FDMA及/或TD-SCDMA)传输通道的IP语音(VoIP)的系统。

[0151] 明确地预期且在此揭示本文中揭示的通信装置(例如,智能电话、平板计算机)可适于在包交换(例如,经布置以根据例如VoIP等协议携带音频传输的有线及/或无线网络)及/或电路交换的网络中使用。还明确地预期且在此揭示本文中揭示的通信装置可适于在

窄带译码系统(例如,编码约四或五千赫兹的音频频率范围的系统)中使用及/或在宽带译码系统(例如,编码大于五千赫兹的音频频率的系统)(包含整个频带宽带译码系统及分裂频带宽带译码系统)中使用。

[0152] 提供对所描述的配置的前述呈现以使得所属领域的技术人员能够制作或使用本文中揭示的方法及其它结构。本文中所展示和描述的流程图、框图和其它结构仅为实例,且这些结构的其它变型也在本发明的范围内。对这些配置的各种修改均为可能的,且本文中所呈现的一般原理还可应用于其它配置。因此,本发明无意限于以上展示的配置而是将被赋予与本文中(包含在形成原始揭示内容的部分的所申请的所附权利要求书中)以任何方式所揭示的原理及新颖特征一致的最广范围。

[0153] 所属领域的技术人员将了解,可使用多种不同技术和技法中的任一者来表示信息和信号。例如,可通过电压、电流、电磁波、磁场或磁粒子、光场或光粒子或其任何组合来表示在整个以上描述中参考的数据、指令、命令、信息、信号、位及符号。

[0154] 如本文所揭示的配置的实施方案的重要设计要求可包含使处理延迟及/或计算复杂度(通常以每秒百万指令或MIPS进行测量)最小化,尤其对于计算密集的应用,例如压缩音频或视听的信息的回放(例如,根据压缩格式而编码的文件或流,例如本文中识别的实例中的一者)或对于宽带通信的应用(例如,在高于八千赫兹(例如,12kHz、16kHz、44.1kHz、48kHz或192kHz)的取样率下的话音通信)。

[0155] 多麦克风处理系统的目标可包含:实现总体噪声减少中的十dB到十二dB;在所要的扬声器的移动期间保留话音水平及色彩;获得噪声已移动到后台中而不是积极的噪声移除的感知;话语的反混响;及/或启用后处理选项以用于更积极的噪声减少。

[0156] 如本文所揭示的设备(例如,设备A100、MF100)可以被视为适合于既定应用的硬件与软件及/或与固件的任何组合实施。举例来说,可将此设备的元件制造为(例如)驻留在相同芯片上或芯片组中的两个或更多个芯片当中的电子及/或光学装置。此装置的一个实例是逻辑元件(例如,晶体管或逻辑门)的固定或可编程阵列,及可实施为一或多个所述阵列的这些元件中的任一者。设备的任何两个或更多个或甚至所有元件可实施在相同阵列内。此阵列或此些阵列可实施在一或多个芯片内(例如,包含两个或更多个芯片的芯片组内)。

[0157] 本文中揭示的设备的各个实施方案的一或多个元件还可完全或部分实施为一或多个指令集,所述一或多个指令集经布置以在逻辑元件的一或多个固定或可编程阵列上执行,例如微处理器、嵌入式处理器、IP核心、数字信号处理器、FPGA、(现场可编程门阵列)、ASSP(专用标准产品)及ASIC(专用集成电路)。如本文所揭示的设备的实施方案的各个元件中的任一者还可体现为一或多个计算机(例如,包含经编程以执行指令的一或多个集合或序列的一或多个阵列的机器,也称为“处理器”),且这些元件中的任何两者或更多者或甚至全部可实施在相同的此类计算机内。

[0158] 可将如本文所揭示的处理器或其它用于处理的装置制造为(例如)驻留在相同芯片上或芯片组中的两个或更多个芯片当中的一或多个电子及/或光学装置。此装置的一个实例是逻辑元件(例如,晶体管或逻辑门)的固定或可编程阵列,及可实施为一或多个所述阵列的这些元件中的任一者。此阵列或此些阵列可实施在一或多个芯片内(例如,包含两个或更多个芯片的芯片组内)。所述阵列的实例包含逻辑元件的固定或可编程阵列,例如微处理器、嵌入式处理器、IP核心、DSP、FPGA、ASSP及ASIC。如本文所揭示的处理器或其它用于处

理的装置还可体现为一或多个计算机(例如,包含经编程以执行指令的一或多个集合或序列的一或多个阵列的机器)或其它处理器。如本文所描述的处理器有可能将用于执行任务或执行不直接与如本文所描述的音频译码程序相关的其它指令集,例如,与其中嵌入有处理器的装置或系统(例如,音频感测装置)的另一操作相关的任务。如本文所揭示的方法的部分还可能由音频感测装置的处理器执行,且方法的另一部分在一或多个其它处理器的控制下执行。

[0159] 此外,所属领域的技术人员将了解,结合本文所揭示的配置而描述的各种说明性模块、逻辑块、电路及测试和其它操作可实施为电子硬件、计算机软件或两者的组合。可使用通用处理器、数字信号处理器(DSP)、ASIC或ASSP、FPGA或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件、或其经设计以产生如本文所揭示的配置的任何组合来实施或执行所述模块、逻辑块、电路及操作。举例来说,可将此配置至少部分实施为硬连线电路、实施为制造到专用集成电路中的电路配置,或实施为加载到非易失性存储器的固件程序或作为机器可读代码从数据存储媒体加载或加载到数据存储媒体中的软件程序,所述代码是可由例如通用处理器或其它数字信号处理单元等逻辑元件阵列执行的指令。通用处理器可为微处理器,但在替代方案中,处理器可为任何常规的处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP与微处理器的组合、多个微处理器的组合、一或多个微处理器与DSP核心的联合,或任何其它此类配置。软件模块可驻留在非暂时性存储媒体中,所述非暂时性存储媒体例如为RAM(随机存取存储器)、ROM(只读存储器)、非易失性RAM(NVRAM)(例如,快闪RAM、可擦除可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM))、寄存器、硬盘、可装卸式磁盘或CD-ROM;或驻留在此项技术中已知的任何其它形式的存储媒体中。说明性存储媒体耦合到处理器,使得处理器可从存储媒体读取信息和将信息写入到存储媒体。在替代方案中,储存媒体可与处理器成一体式。处理器和存储媒体可驻留在ASIC中。ASIC可驻留在用户终端中。在替代方案中,处理器和存储媒体可作为离散组件驻留在用户终端中。

[0160] 应注意,本文中揭示的各种方法(例如,方法M100、M200、M300)可由例如处理器等逻辑元件阵列执行,且如本文所描述的设备的各种元件可实施为经设计以在此阵列上执行的模块。如本文所用,术语“模块”或“子模块”可指以软件、硬件或固件形式包含计算机指令(例如,逻辑表达式)的任何方法、设备、装置、单元或计算机可读数据存储媒体。应理解,可将多个模块或系统组合为一个模块或系统,且可将一个模块或系统分离成多个模块或系统以执行相同功能。在实施于软件或其它计算机可执行指令中时,过程的元件基本上是用以例如使用例程、程序、对象、组件、数据结构及类似者执行相关任务的代码段。术语“软件”应理解为包括源代码、汇编语言代码、机器代码、二进制代码、固件、宏码、微码、可由逻辑元件阵列执行的任何一或多个指令集或序列以及此类实例的任何组合。所述程序或代码段可存储于处理器可读存储媒体中或由体现在传输媒体或通信链路中的载波中的计算机数据信号传输。

[0161] 本文中揭示的方法、方案及技术的实施方案还可有形地体现(例如,在如本文所列举的一或多个计算机可读媒体中)为可由包含逻辑元件阵列(例如,处理器、微处理器、微控制器或其它有限状态机)的机器读取及/或执行的一或多个指令集。术语“计算机可读媒体”可包含可存储或传递信息的任何媒体,包含易失性、非易失性、可装卸式及非可装卸式媒

体。计算机可读媒体的实例包含电子电路、半导体存储器装置、ROM、快闪存储器、可擦除ROM (EROM)、软盘或其它磁性存储装置、CD-ROM/DVD或其它光学存储装置、硬盘、光纤媒体、射频 (RF) 链路,或可用于存储所要的信息且可对其存取的任何其它媒体。计算机数据信号可包含可经由传输媒体传播的任何信号,所述传输媒体例如为电子网络通道、光纤、空气、电磁、RF链路等。可经由例如因特网或内联网等计算机网络下载代码段。在任何情况下,不应将本发明的范围解释为被所述实施例限制。

[0162] 本文中所述的方法的任务中的每一者可直接以硬件、以由处理器执行的软件模块,或以所述两者的组合体现。在如本文所揭示的方法的实施方案的典型应用中,逻辑元件(例如,逻辑门)的阵列经配置以执行所述方法的各种任务中的一者、一者以上或甚至全部。还可将所述任务中的一或多个者(可能全部)实施为体现在计算机程序产品(例如,一或多个数据存储媒体,例如磁盘、快闪或其它非易失性存储器卡、半导体存储器芯片等)中的代码(例如,一或多个指令集),所述计算机程序产品可由包含逻辑元件阵列(例如,处理器、微处理器、微控制器或其它有限状态机)的机器(例如,计算机)读取及/或执行。如本文所揭示的方法的实施方案的任务还可由一个以上所述阵列或机器执行。在这些或其它实施方案中,所述任务可在用于无线通信的装置(例如,蜂窝式电话或具有此类通信能力的其它装置)内执行。此装置可经配置以与电路交换及/或包交换网络通信(例如,使用一或多个协议,例如VoIP)。举例来说,此装置可包含经配置以接收及/或传输经编码帧的RF电路。

[0163] 明确地揭示的是,本文中揭示的各种方法可由例如手持机、耳机或便携式数字助理(PDA)等便携式通信装置执行,且本文中描述的各种设备可包含在此装置内。典型的实时(例如,在线)应用是使用此移动装置进行的电话交谈。

[0164] 因此,在一或多个示范性实施例中,本文中描述的操作可在硬件、软件、固体或其任何组合中实施。如果实施于软件中,那么可将此类操作作为一或多个指令或代码存储在计算机可读媒体上或经由计算机可读媒体进行传输。术语“计算机可读媒体”包含计算机可读存储媒体及通信(例如,传输)媒体两者。举例来说而非限制,计算机可读存储媒体可包括存储元件阵列,例如半导体存储器(其可包含(不限于)动态或静态RAM、ROM、EEPROM及/或快闪RAM),或铁磁阻双聚合物或相变存储器;CD-ROM或其它光盘存储装置;和/或磁盘存储器或其它磁性存储装置。此类存储媒体可以可由计算机存取的指令或数据结构的形式存储信息。通信媒体可包括可用于携带呈指令或数据结构的形式所要的程序代码及可由计算机存取的任何媒体,包含促进将计算机程序从一个地方传递到另一地方的任何媒体。而且,将任何连接恰当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或无线技术(例如红外线、无线电及/或微波)从网站、服务器或其它远程源传输软件,那么所述同轴电缆、光纤电缆、双绞线、DSL或无线技术(例如红外线、无线电及/或微波)包含在媒体的定义中。如本文中所使用的磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘及蓝光Disc™(蓝光光盘协会,环球城,加州),其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。以上各者的组合也应包含在计算机可读媒体的范围内。

[0165] 如本文所描述的声学信号处理设备(例如,设备A100或MF100)可并入到接受话语输入以便控制某些操作或可另外受益于所要的噪声与后台噪声的分离的电子装置中,例如通信装置。许多应用可受益于从源自多个方向的后台声音促进或分离明晰的所要的声音。

此类应用可包含并入例如话音辨识及检测、话语增强及分离、话音激活控制及类似者等能力的电子或计算装置中的人机接口。可能需要实施在仅提供有限的处理能力的装置中合适的此声学信号处理设备。

[0166] 可将本文中描述的模块、元件及装置的各种实施方案的元件制造为(例如)驻留在相同芯片上或芯片组中的两个或更多个芯片当中的电子及/或光学装置。此装置的一个实例是逻辑元件(例如,晶体管或门)的固定或可编程阵列。本文中描述的设备的各种实施方案的一或多个元件还可完全或部分实施为经布置以在逻辑元件的一或多个固定或可编程阵列上执行的一或多个指令集,所述阵列例如为微处理器、嵌入式处理器、IP核心、数字信号处理器、FPGA、ASSP及ASIC。

[0167] 如本文所描述的设备的实施方案的一或多个元件有可能将用于执行任务或执行不直接与设备的操作相关的其它指令集,例如,与其中嵌入有所述设备的装置或系统的另一操作相关的任务。此设备的实施方案的一或多个元件还有可能具有共同结构(例如,处理器用于执行对应于不同时间的不同元件的代码的部分,执行指令集以执行对应于不同时间的不同元件的任务,或电子及/或光学装置的布置执行用于不同时间的不同元件的操作)。

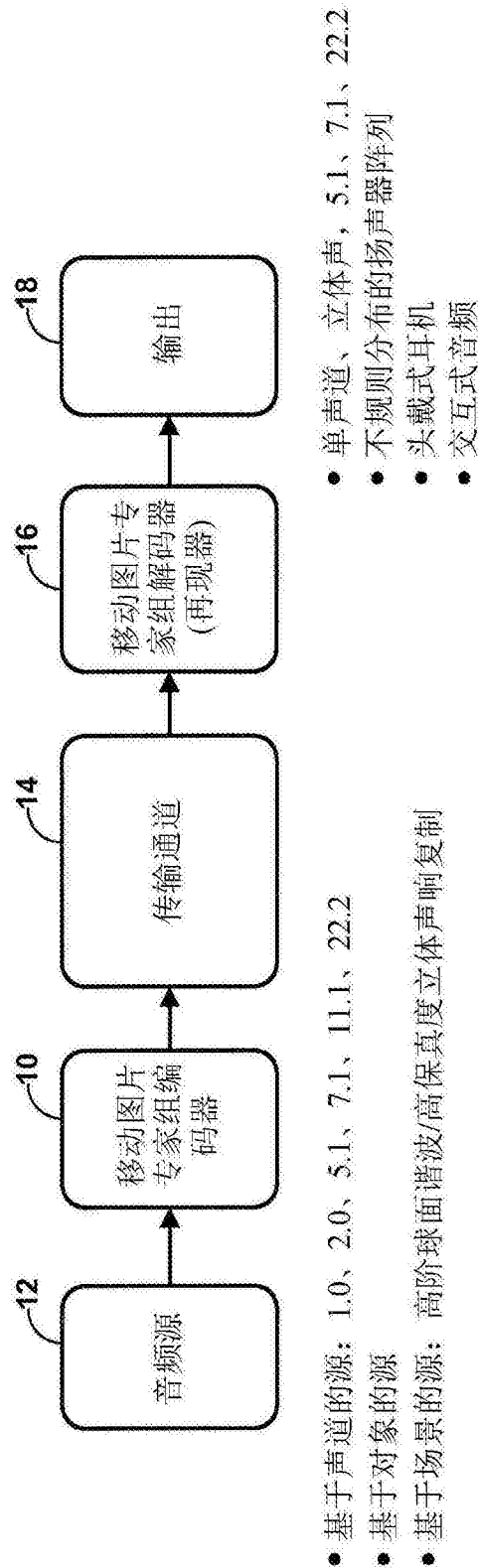


图1

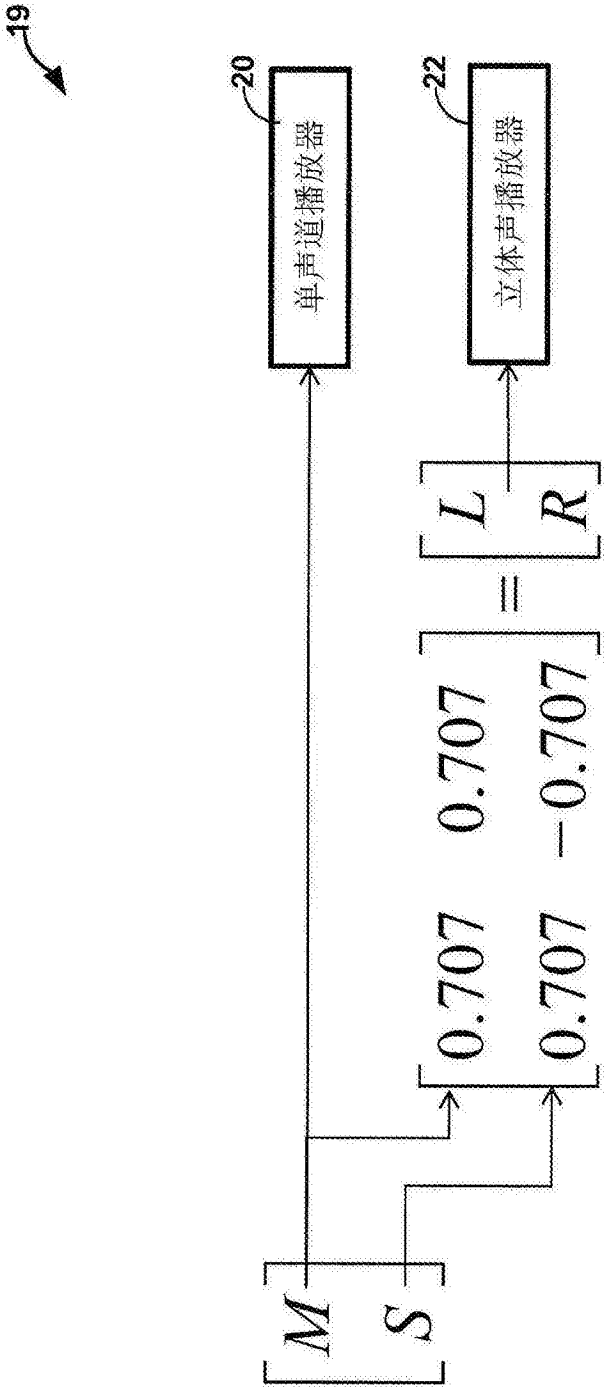


图2

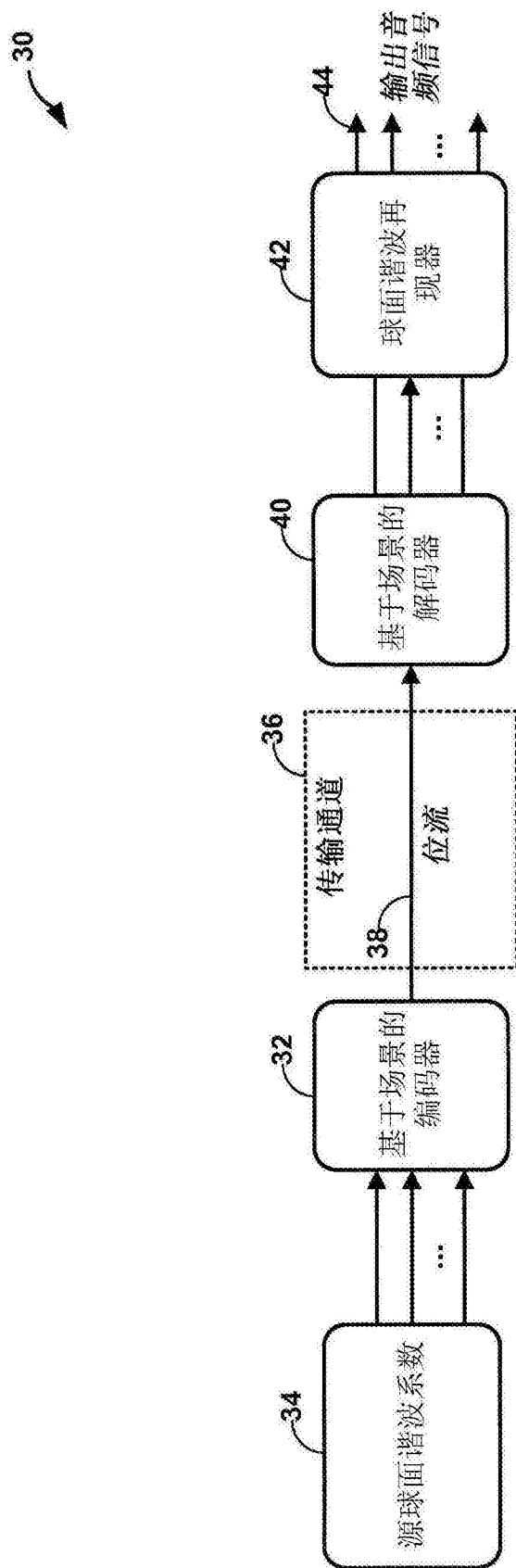


图3

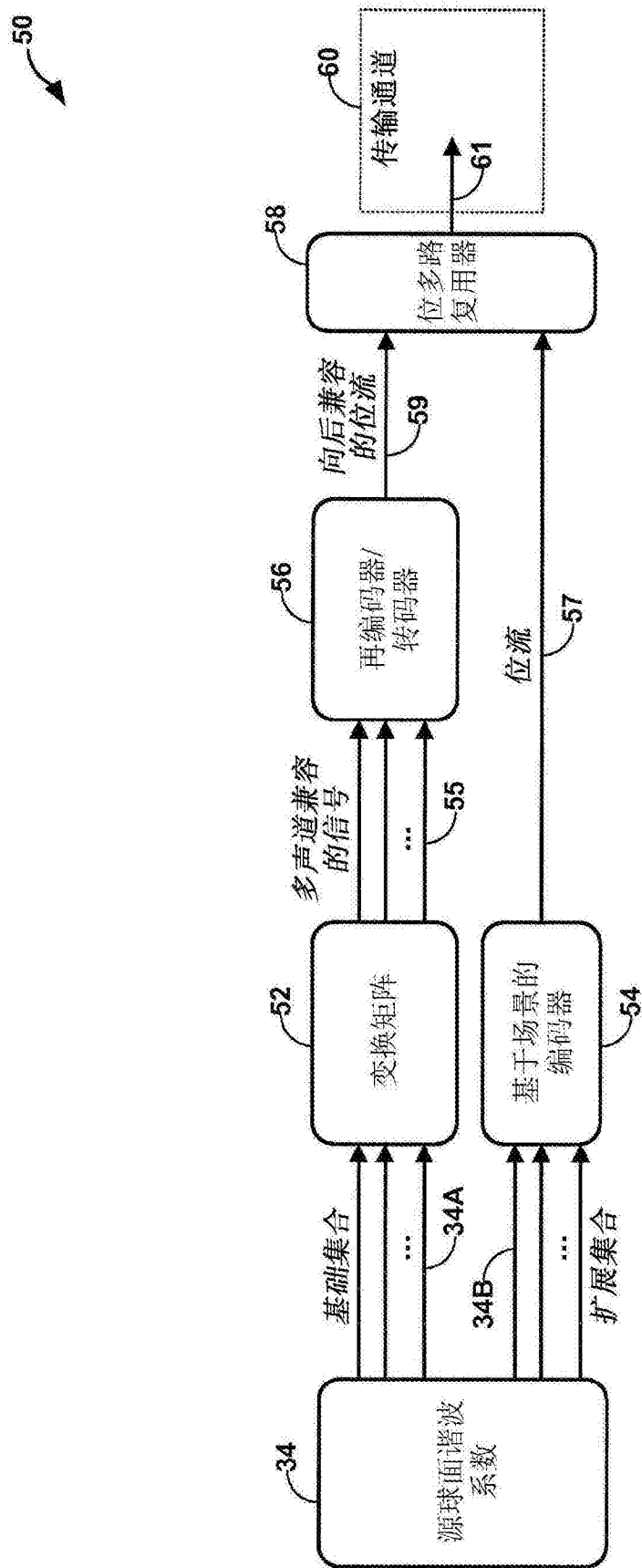


图4

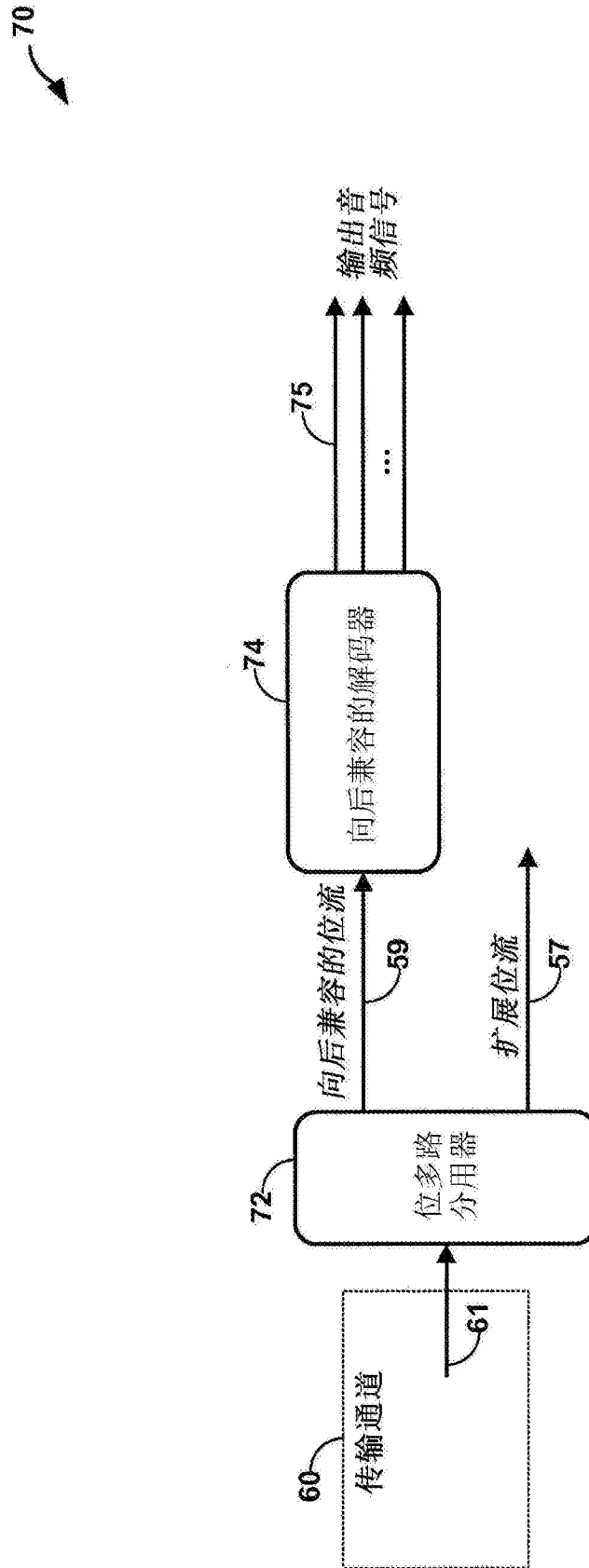


图5

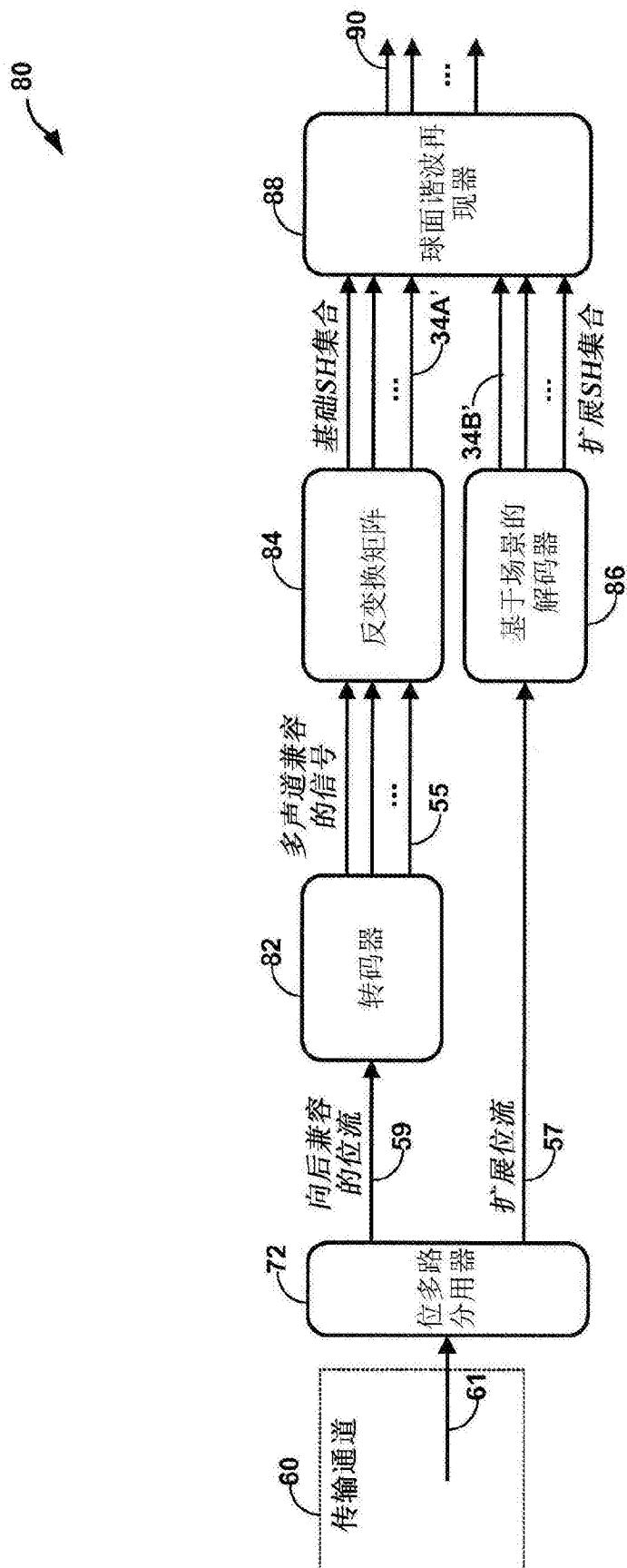


图6

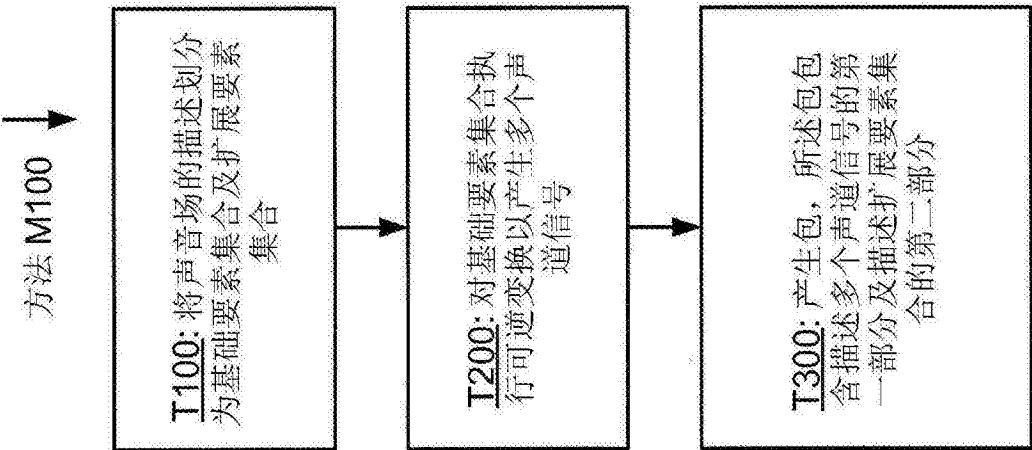


图7A

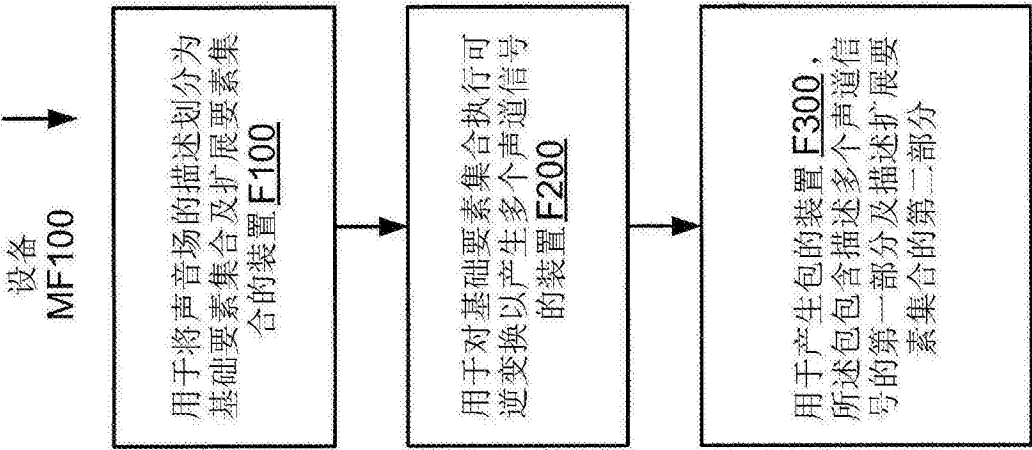


图7B

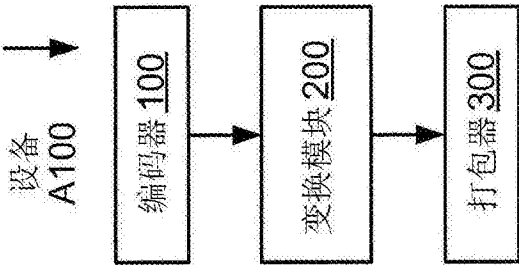


图7C

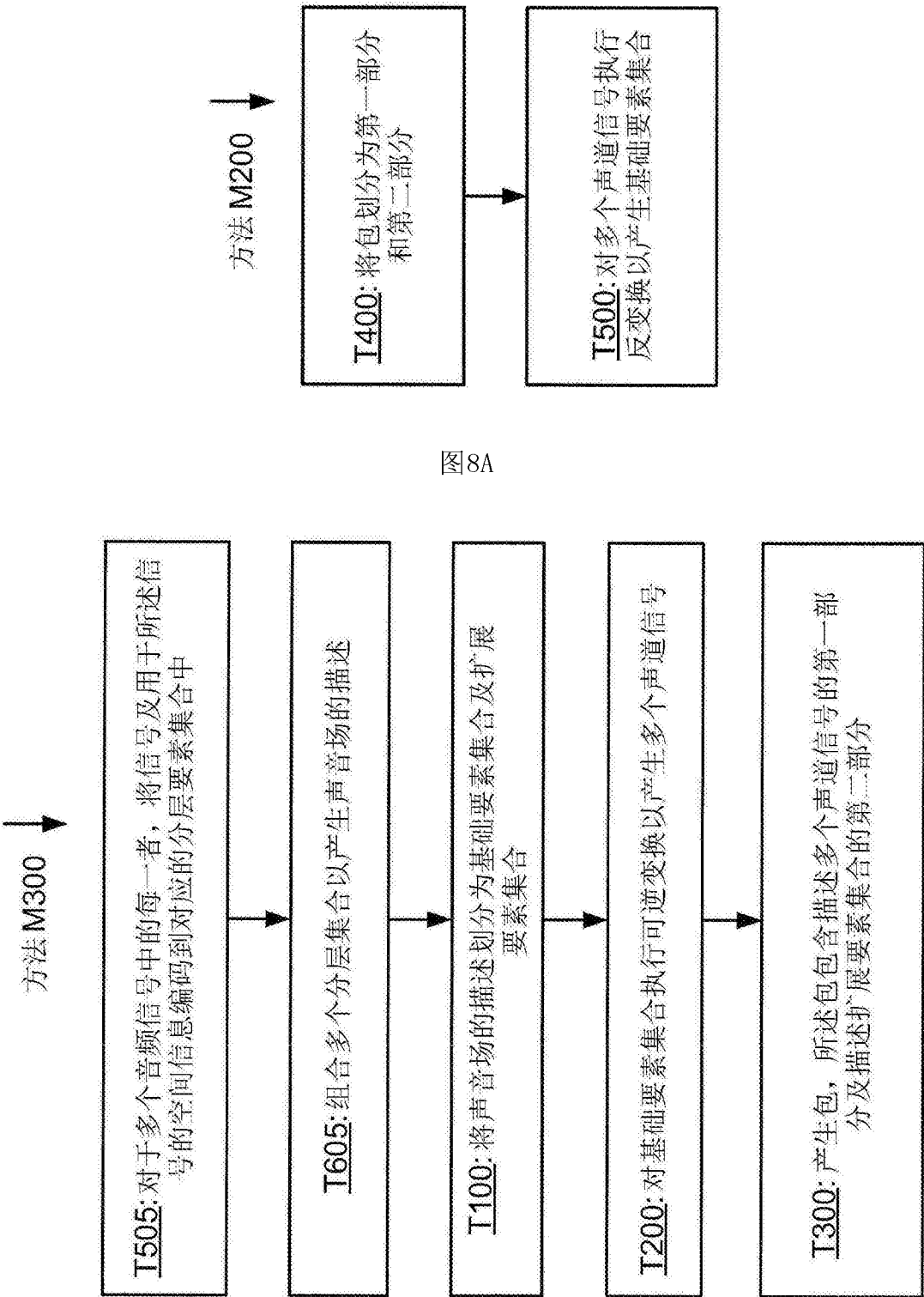


图8A

图8B

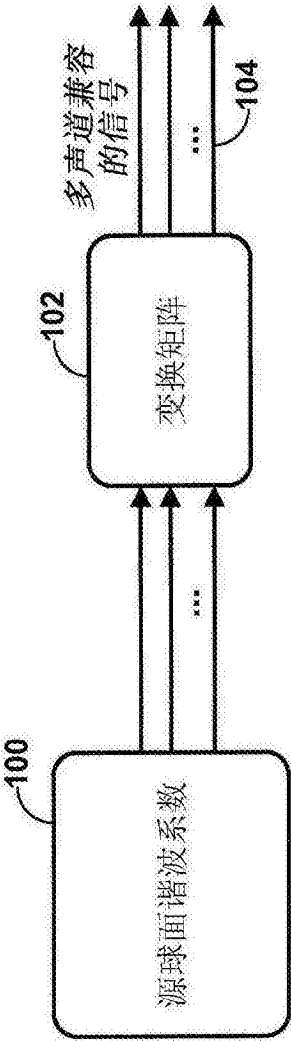


图9A

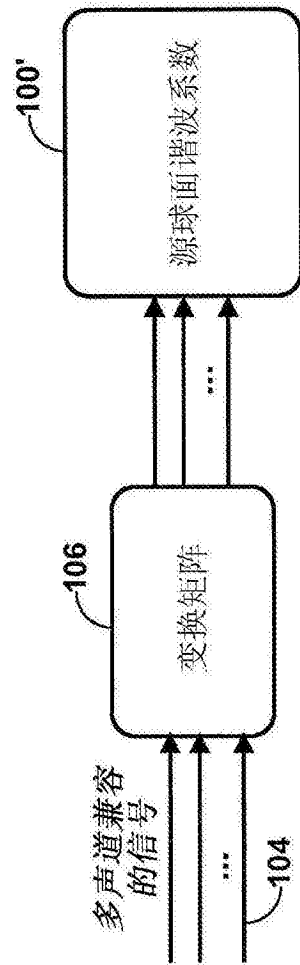


图9B

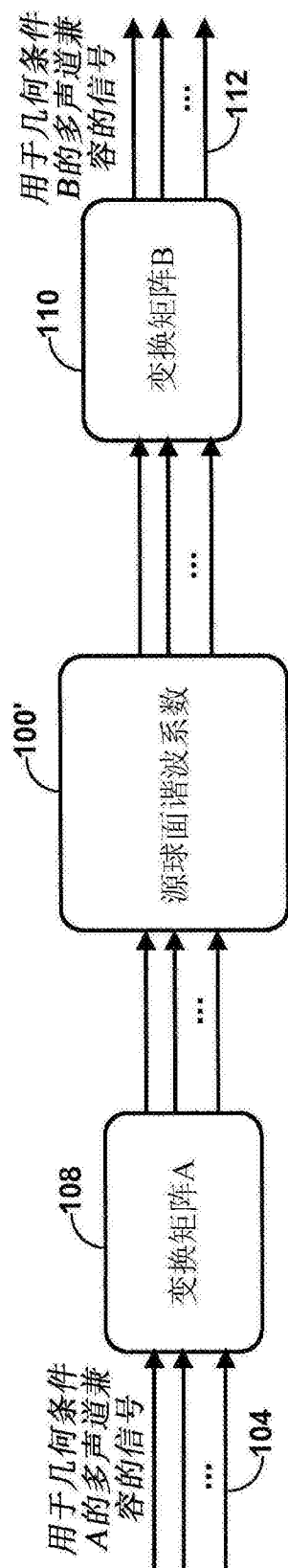


图9C

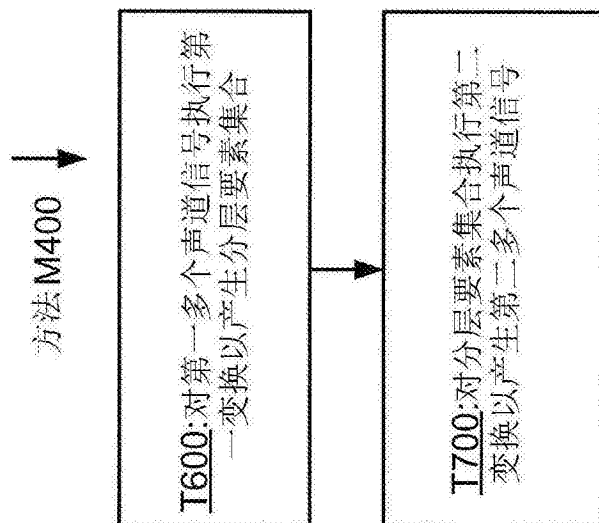


图10A

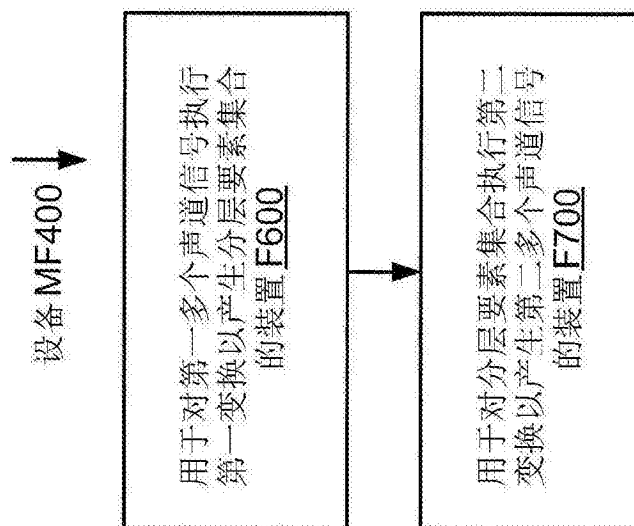


图10B

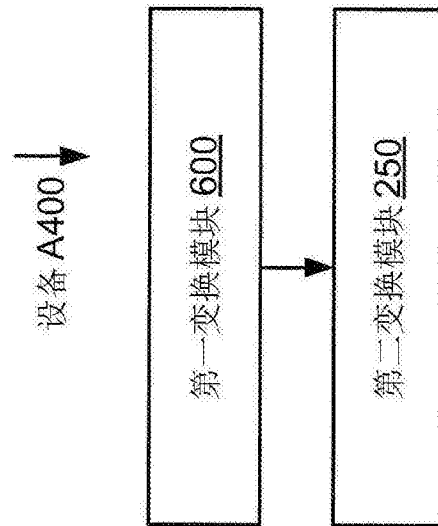


图10C

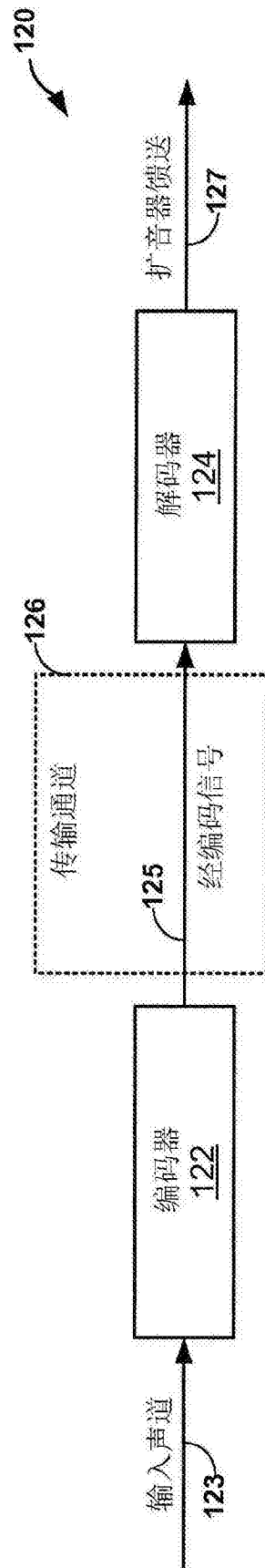


图10D

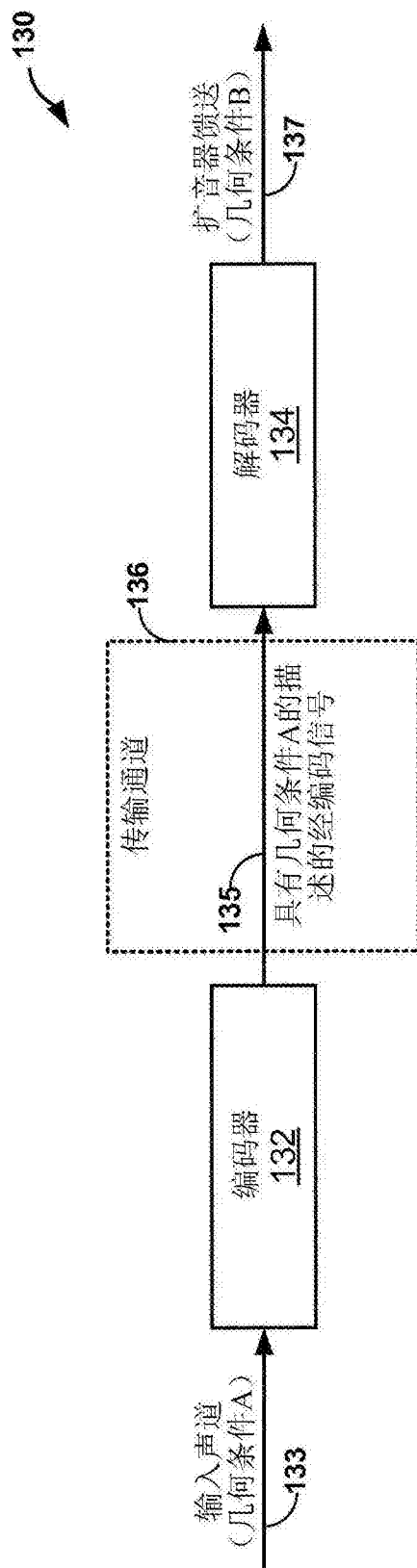


图11A

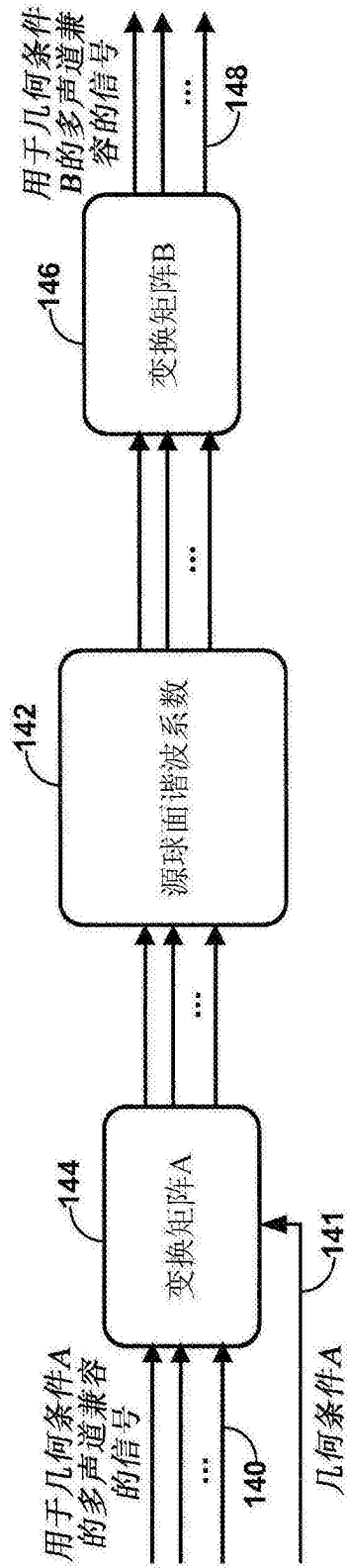


图11B

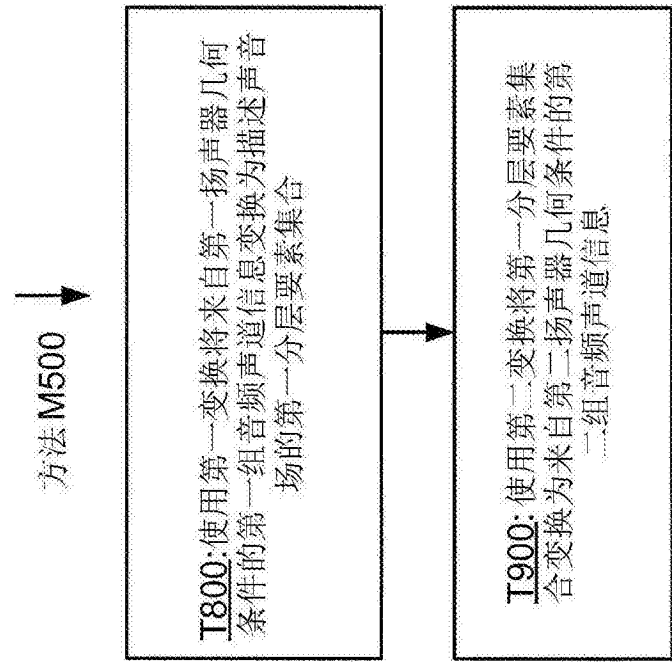


图12A

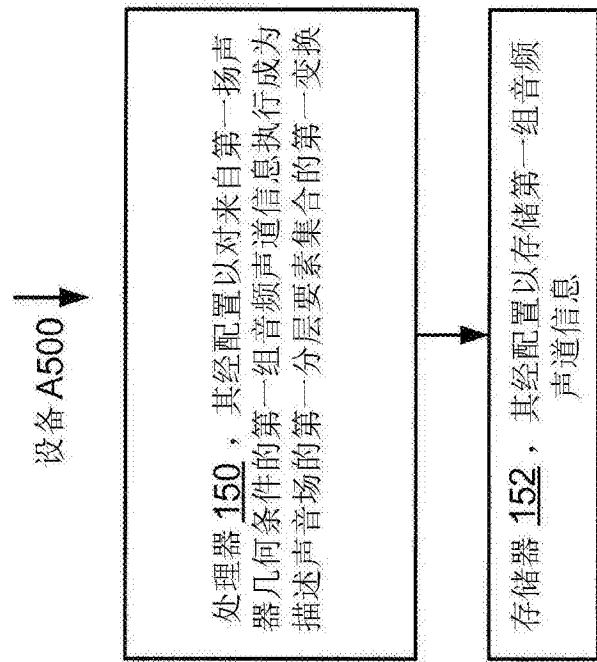


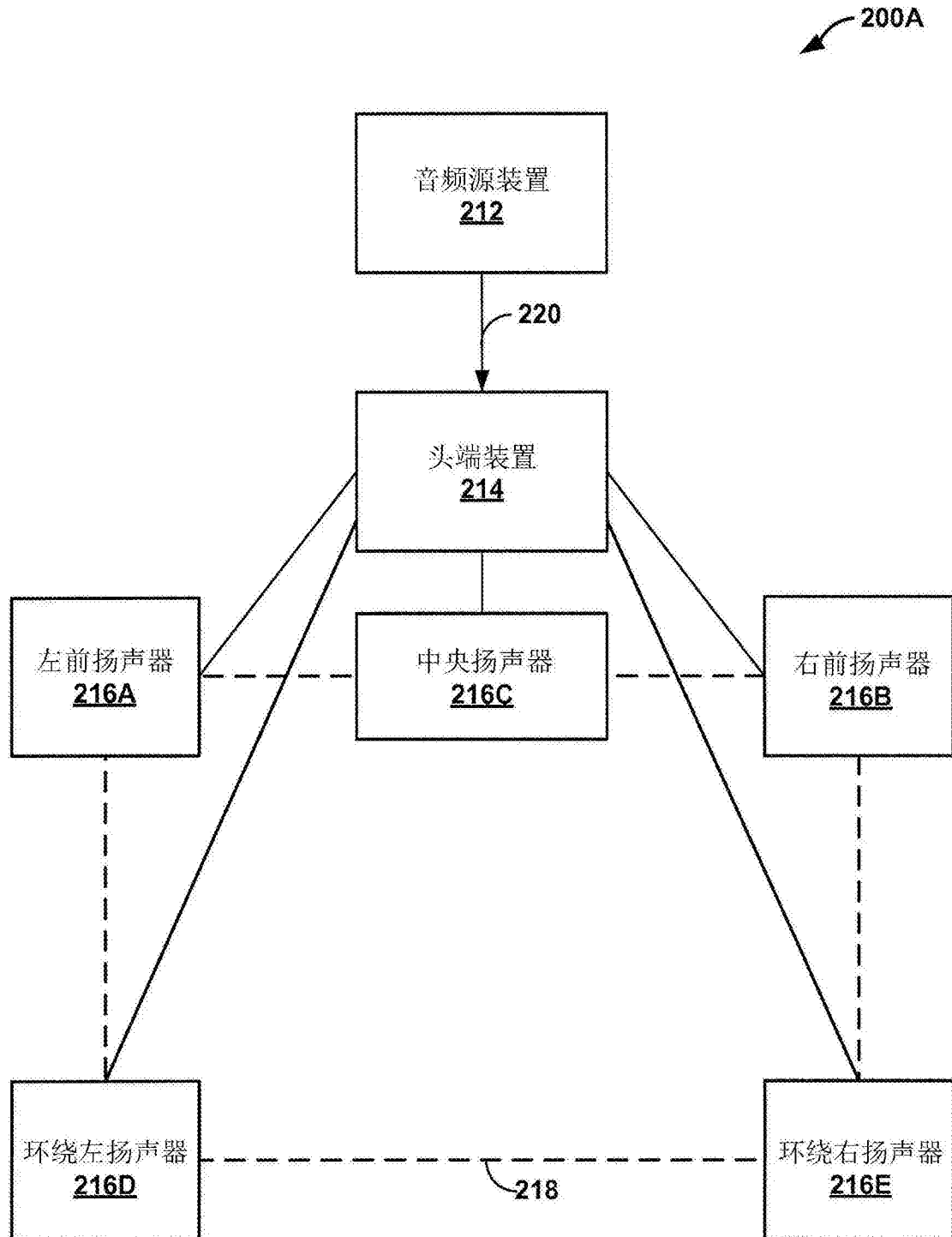
图12B

M600:接收扩音器声道以及第一扬声器几
何条件的坐标,其中扩音器声道已经被变
换为分层要素集合

图12C

M700:传输扩音器声道以及第一扬声器几
何条件的坐标,其中所述第一几何条件对
应于声道的位置

图12D



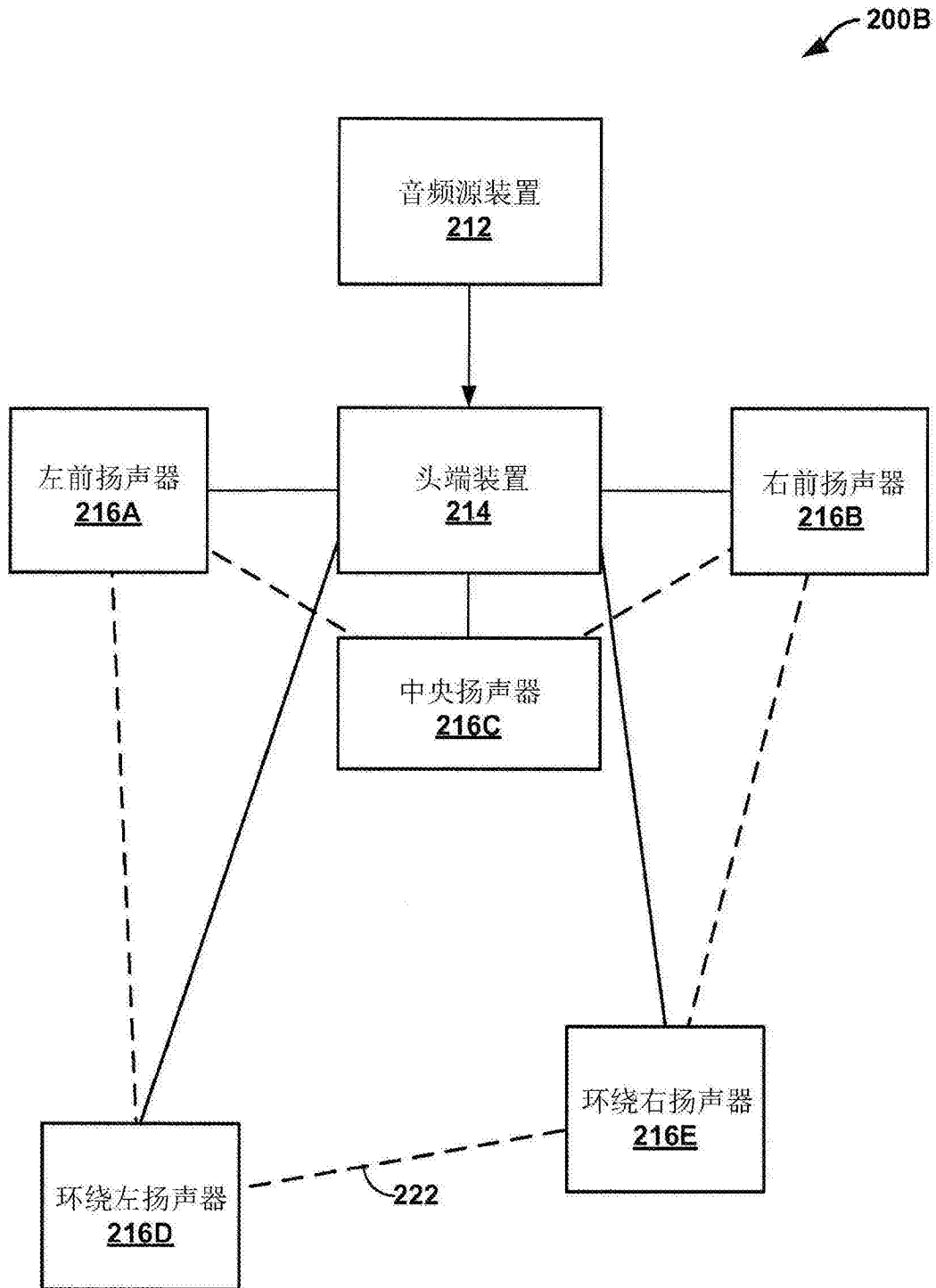


图13B

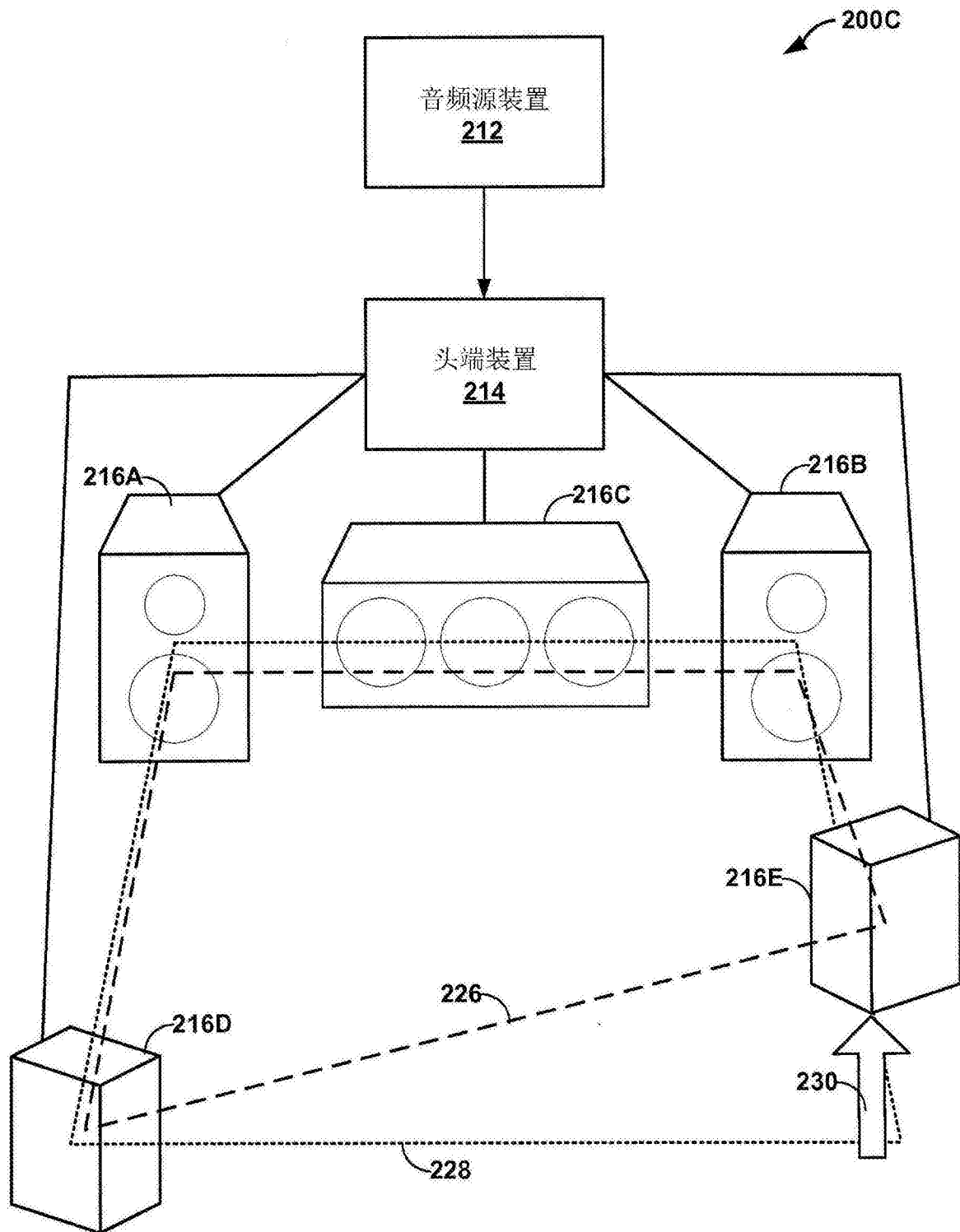


图13C

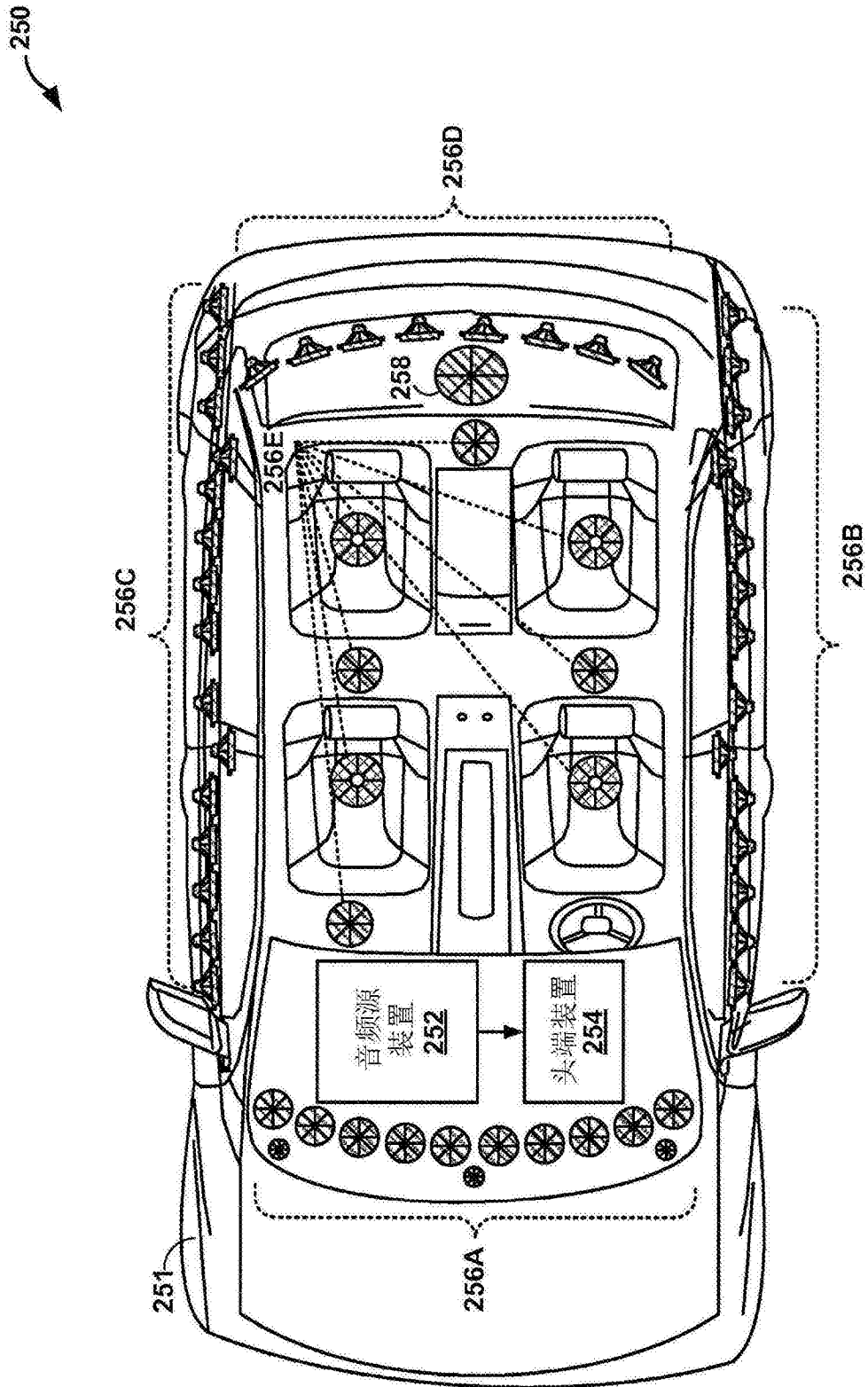


图14