(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117767898 A
(43)申请公布日 2024.03.26

(21)申请号 202311785926.1

(22)申请日 2014.09.08

(30)优先权数据
61/877,230 2013.09.12 US
61/891,324 2013.10.15 US
61/938,043 2014.02.10 US

(62)分案原申请数据
201480050047.7 2014.09.08

(71)申请人 杜比实验室特许公司
地址 美国加利福尼亚
申请人 杜比国际公司

(72)发明人 J·科喷斯 J·瑞德米勒
K·克约尔林 A·格罗舍尔
H·霍伊里奇 A·J·西费尔特

(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 宿小猛

(51)Int.Cl.
*H03G 7/00*(2006.01)
*H03G 9/00*(2006.01)
*G10L 19/16*(2013.01)

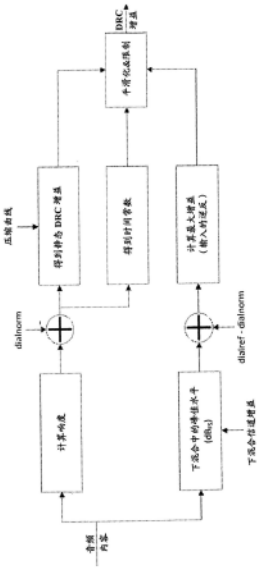
权利要求书2页 说明书29页 附图13页

(54)发明名称

用于各种回放环境的动态范围控制

(57)摘要

本公开涉及用于各种回放环境的动态范围控制。在音频编码器中,对在源音频格式中接收的音频内容,基于缺省动态范围压缩(DRC)曲线生成缺省增益,并且,对非缺省增益简档生成非缺省增益。基于缺省增益和非缺省增益,生成差分增益。生成包含音频内容、缺省DRC曲线和差分增益的音频信号。在音频解码器中,从音频信号识别缺省DRC曲线和差分增益。基于缺省DRC曲线重新生成缺省增益。基于重新生成的缺省增益和差分增益的组合,对从音频信号提取的音频内容执行动作。



CN 117767898 A

1.一种用于音频信号的动态范围控制(DRC)的方法,所述方法包括:

通过在具有不同于基准信道配置的对应特定扬声器配置的特定回放环境中操作的音频解码器接收用于基准信道配置的音频信号,所述音频信号包括用于所述基准信道配置中的每个信道的音频采样数据、以及编码器生成的DRC元数据,所述编码器生成的DRC元数据包括用于多个回放环境的DRC增益,所述DRC增益包括用于具有对应特定扬声器配置的特定回放环境的一组DRC增益和用于所述基准信道配置的一组DRC增益;

从用于所述多个回放环境的所述DRC增益中选择用于具有对应特定扬声器配置的特定回放环境的所述一组DRC增益;

将用于具有对应特定扬声器配置的特定回放环境的所述一组DRC增益作为被应用于所述音频采样数据的总增益的一部分而应用,以生成用于与所述特定回放环境对应的特定扬声器配置中的每个信道的输出音频采样数据,

其中,与所述特定回放环境对应的特定扬声器配置具有与所述基准信道配置数量不同的音频信道。

2.根据权利要求1所述的方法,其中,所述总增益包括以下中的一个或多个:

与下混合有关的增益,

与恢复原始动态范围有关的增益,音频采样数据的输入动态范围是从所述原始动态范围转换得到的,

与增益限制有关的增益,

与增益平滑化有关的增益,或者

与对话响度归一化有关的增益。

3.根据权利要求1所述的方法,其中,所述总增益包括要被部分地/单个地施加的增益、要被串行施加的增益、要被并行施加的增益、或者要被部分串行部分并行施加的增益。

4.根据权利要求1所述的方法,其中,与所述特定回放环境对应的特定扬声器配置是两信道配置。

5.一种用于音频信号的动态范围控制的音频信号处理设备,其中所述音频信号处理设备:

通过在具有不同于基准信道配置的对应特定扬声器配置的特定回放环境中操作的音频解码器接收用于基准信道配置的音频信号,所述音频信号包括用于所述基准信道配置中的每个信道的音频采样数据、以及编码器生成的DRC元数据,所述编码器生成的DRC元数据包括用于多个回放环境的DRC增益,所述DRC增益包括用于具有对应特定扬声器配置的特定回放环境的一组DRC增益和用于所述基准信道配置的一组DRC增益;

从用于所述多个回放环境的所述DRC增益中选择用于具有对应特定扬声器配置的特定回放环境的所述一组DRC增益;以及

将用于具有对应特定扬声器配置的特定回放环境的所述一组DRC增益作为被应用于所述音频采样数据的总增益的一部分而应用,以生成用于与所述特定回放环境对应的特定扬声器配置中的每个信道的输出音频采样数据,

其中,与所述特定回放环境对应的特定扬声器配置具有与所述基准信道配置数量不同的音频信道。

6.一种用于音频信号的动态范围控制(DRC)的设备,包括:

一个或多个处理器,和

非暂时性计算机可读存储介质,存储有软件指令,所述软件指令在被所述一个或多个处理器执行时使得执行根据权利要求1-4中的任一项所述的方法。

7.一种处理音频内容的方法,包括:

接收音频信号,所述音频信号包括音频内容以及一组或多组差分增益;

对特定回放环境中的增益简档识别所述一组或多组差分增益之中的特定组的差分增益;

至少部分地基于与所述特定组的差分增益有关的缺省动态范围压缩曲线生成一组缺省增益;以及

至少部分地基于所述一组缺省增益与所述特定组的差分增益的组合,使用根据从所述音频信号提取的音频内容的一个或多个部分中的每一个确定的响度水平,对于所述音频内容的所述一个或多个部分执行一个或多个动作,

保持瞬时响度水平的直方图,所述直方图被根据所述音频内容中的在所述音频内容的所述一个或多个部分之前的时间间隔计算的瞬时响度水平占据;

确定特定响度水平是否在直方图的高概率区域中低于阈值,特定响度水平是从所述音频内容的在所述一个或多个部分中的一部分计算的;和

响应于确定特定响度水平在所述直方图的所述高概率区域中低于所述阈值,执行以下过程:

确定根据其计算所述特定响度水平的所述音频内容的所述一部分出现了响度过渡;以及

缩短在平滑化中使用的时间常数以加速响度过渡。

8.一种非暂时性计算机可读存储介质,存储有软件指令,所述软件指令在被一个或多个处理器执行时使得执行根据权利要求1-4或7中的任一项所述的方法。

9.一种包括用于执行根据权利要求1-4或7中的任一项所述的方法的部件的装置。

# 用于各种回放环境的动态范围控制

[0001] 本发明专利申请是申请号为201910279663.4、申请日为2014年9月8日、发明名称为“用于各种回放环境的动态范围控制”的发明专利申请的分案申请,该申请号为201910279663.4的发明专利申请是申请号为201480050047.7、申请日为2014年9月8日、发明名称为“用于各种回放环境的动态范围控制”的发明专利申请的分案申请。

[0002] (对相关申请的交叉引用)

[0003] 本申请要求在2013年9月12日提交的美国临时专利申请No.61/877230、在2013年10月15日提交的美国临时专利申请No.61/891324和在2014年2月10日提交的美国临时专利申请No.61/938043作为优先权,在这里通过引用并入其全部内容。

## 技术领域

[0004] 本发明总体涉及处理音频信号,更特别地,涉及可被用于在各种回放环境中的任一种中向音频信号施加动态范围控制和其它类型的音频处理动作的技术。

## 背景技术

[0005] 媒体消费者装置的日益普及对用于在这些装置处回放的媒体内容的创建者和分配者以及装置的设计人员和制造商提出新的机会和挑战。许多消费者装置能够回放广泛的媒体内容类型和格式,包括常常与用于HDTV、Blu-ray或DVD的高质量、宽带宽和宽动态范围音频内容相关的那些。媒体处理装置可被用于任意地在它们自身的内部声学换能器或诸如耳机的外部换能器上回放这种类型的音频内容;但是,它们一般不能跨着各种媒体格式和内容类型用一致的响度和清晰度(intelligibility)再现该内容。

[0006] 在本部分中描述的方法是可追寻的方法,但未必是以前构想或追寻的方法。因此,除非另外指示,否则,不应仅通过包含于本部分中而假定在本部分中描述的方法中的任一个为现有技术。类似地,除非另外指示,否则,关于一个或更多个方法识别的问题不应基于本部分而假定在任何现有技术中已被识别。

## 附图说明

[0007] 在附图中作为例子而不是限制示出本发明,并且,在这些附图中,类似的附图标记指的是类似的要素,其中,

[0008] 图1A和图1B分别示出示例性音频解码器和示例性音频编码器;

[0009] 图2A和图2B示出示例性动态范围压缩曲线;

[0010] 图3示出组合DRC和限制增益的确定/计算的示例性处理逻辑;

[0011] 图4示出增益的示例性差分编码;

[0012] 图5示出包含音频编码器和音频解码器的示例性编解码器系统;

[0013] 图5A示出示例性音频解码器;

[0014] 图6A～6D示出示例性处理流程;以及

[0015] 图7示出可在上面实现这里描述的计算机或计算装置的示例性硬件平台。

## 具体实施方式

[0016] 这里描述涉及在各种回放环境中的任一种向音频信号施加动态范围控制处理和其它类型的音频处理动作的示例性实施例。在以下的描述中,出于解释的目的,为了使得能够彻底地理解本发明,阐述了大量的详细细节。但是,显然可以在没有这些特定细节的情况下实现本发明。在其它情况下,为了避免不必要地混淆、掩盖本发明或者使其变模糊,不以详尽的细节描述公知的结构和装置。

[0017] 这里根据以下的提纲描述示例性实施例:

[0018] 1.一般概要

[0019] 2.动态范围控制

[0020] 3.音频解码器

[0021] 4.音频编码器

[0022] 5.动态范围压缩曲线

[0023] 6.DRC增益、增益限制和增益平滑化

[0024] 7.输入平滑化和增益平滑化

[0025] 8.多个频带上的DRC

[0026] 9.响度域中的音量调整

[0027] 10.具有差分增益的增益简档

[0028] 11.与增益有关的附加动作

[0029] 12.特定和广带(或宽带)响度水平

[0030] 13.用于信道的各单个子集的各单个增益

[0031] 14.听觉场景分析

[0032] 15.响度水平过渡

[0033] 16.复位

[0034] 17.编码器提供的增益

[0035] 18.示例性系统和处理流程

[0036] 19.实现机构—硬件概要

[0037] 20.等同物、扩展、替代和杂项

[0038] 1.一般概要

[0039] 本概要给出本发明的实施例的一些方面的基本描述。应当注意,本概要不是实施例的各方面的广延或详尽的总结。并且,应当注意,本概要不应被理解为识别实施例的任何特别重要的方面或要素,也不应被理解为特别界定实施例的任何范围,也不应被理解为一般地界定本发明。本概要仅以概括或简化的形式给出与示例性实施例有关的一些概念,并且应被理解为以下给出的示例性实施例的更详细描述的概念性序言。注意,虽然这里讨论各单独的实施例,但是这里讨论的实施例和/或部分实施例的任意组合可被组合以形成其它的实施例。

[0040] 在一些方法中,编码器假定音频内容是出于动态范围控制的目的对特定环境编码的,并且对特定环境确定诸如用于动态范围控制等的增益的音频处理参数。根据这些方法由编码器确定的增益一般在一些时间间隔等上通过一些时间常数(例如,在指数衰减的函数等中)被平滑化。另外,对确保信号不大于假定环境的修剪水平的增益限制,可能已加入

根据这些方法由编码器确定的增益。因此,根据这些方法通过编码器用音频信息编码成音频信号的增益是许多不同影响的结果并且是不可逆的。根据这些方法接收增益的解码器可能不能区分增益的哪个部分用于动态范围控制、增益的哪个部分用于增益平滑化、增益的哪个部分用于增益限制,等等。

[0041] 根据这里描述的技术,音频编码器不假定只需要支持音频解码器处的特定回放环境。在实施例中,音频编码器传送具有音频内容的编码音频信号,从该编码音频信号可确定正确的响度水平(例如,没有修剪等)。音频编码器也向音频解码器传送一个或更多个动态范围压缩曲线。一个或更多个动态范围压缩曲线中的任一个可以是基于标准的、专有的、定制的、内容提供商特有的,等等。基准响度水平、起音时间(attack time)、释音时间(release time)等可作为一个或更多个动态范围压缩曲线的一部分或者与其结合被音频编码器传送。基准响度水平、起音时间、释音时间等中的任一个可以是基于标准的、专有的、定制的、内容提供商特有的,等等。

[0042] 在一些实施例中,音频编码器实现听觉场景分析(ASA)技术,并且使用ASA技术以检测音频内容中的听觉事件,并且,向音频解码器传送描述检测的听觉事件的一个或更多个ASA参数。

[0043] 在一些实施例中,音频编码器可也被配置为检测音频内容中的复位事件,并且以与音频内容时间同步的方式向诸如音频解码器等的下游装置传送复位事件的指示。

[0044] 在一些实施例中,音频编码器可被配置为计算音频内容的各单个部分(例如,音频数据块、音频数据帧等)的一组或更多组增益(例如,DRC增益等),并且,用音频内容的各单个部分将多组增益编码到编码的音频信号中。在一些实施例中,由音频编码器生成的多组增益与一个或更多个不同增益简档(例如,在表1等中示出的那些)对应。在一些实施例中,可以使用Huffman编码、差分编码等以将多组增益编码成音频数据帧的成分、分支等或者从中读取多组增益。这些成分、分支等可被称为音频数据帧中的子帧。不同组的增益可与不同组的子帧对应。各组增益或者各组子帧可包含两个或更多个时间成分(例如,子帧等)。在一些实施例中,这里描述的音频编码器中的位流格式化器(formatter)可用一个或更多个for循环将一组或更多组增益一起作为差分数据代码写入到音频数据帧中的一组或更多组子帧中;相应地,这里描述的音频解码器中的位流剖析器可从音频数据帧中的一组或更多组子帧读取被编码为差分数据代码的一组或更多组增益中的任一个。

[0045] 在一些实施例中,音频编码器确定要被编码成编码音频信号的音频内容中的对话响度水平,并且,用音频内容向音频编码器传送对话响度水平。

[0046] 在一些实施例中,音频编码器向下游受体音频解码器发送用于回放环境或情形中的缺省增益简档的缺省动态压缩曲线。在一些实施例中,音频编码器假定下游受体音频解码器要使用用于回放环境或情形中的缺省增益简档的缺省动态压缩曲线。在一些实施例中,音频编码器向下游受体音频解码器发送在下游受体音频解码器处定义的一个或更多个动态压缩曲线(或增益简档)中的哪一个要被用于回放环境或情形中的指示。在一些实施例中,对于一个或更多个非缺省增益简档中的每一个,音频编码器发送与该非缺省简档对应的(例如,非缺省等)动态压缩曲线作为由编码音频信号承载的元数据的一部分。这里描述的技术允许通过上游编码器生成与缺省压缩曲线有关的多组差分增益并将其发送到下游解码器。这允许解码器中的DRC压缩器(例如,基于压缩曲线和平滑化动作等计算增益的处

理)设计有大量的自由度,同时,与传送所有增益值相比,使需要的位率保持相对地低。仅出于解释的目的,提到了缺省简档或缺省DRC曲线为与可特别计算非缺省简档或非缺省DRC曲线的差分增益相关的一个。但是,这仅是出于解释的目的;如各种实施例那样,不需要严格区分缺省简档与非缺省简档(例如,在媒体数据流等中),所有其它简档可以是相对于同一特定(例如,"缺省"等)压缩曲线的差分增益。如这里使用的那样,"增益简档"可指的是作为执行DRC动作的压缩器的动作模式的DRC模式。在一些实施例中,DRC模式涉及特定类型的回放装置(AVR对TV对平板)和/或环境(噪声对安静对后夜)。各DRC模式可与增益简档相关。增益简档可由压缩器执行DRC动作所基于的定义数据代表。在一些实施例中,增益简档可以是DRC曲线(可能被参数化)和在DRC动作中使用的时间常数。在一些实施例中,增益简档可以是作为响应音频信号的DRC动作的输出的一组DRC增益。不同DRC模式的简档可与不同的压缩量对应。

[0047] 在一些实施例中,音频编码器基于与缺省增益简档对应的缺省动态范围压缩曲线确定音频内容的一组缺省(例如,全DRC和非DRC、全DRC等)增益,并且,对于一个或更多个非缺省增益简档中的每一个,确定同一音频内容的一组非缺省(例如,全DRC和非DRC、全DRC等)增益。音频编码器可然后确定缺省增益简档的一组缺省(例如,全DRC和非DRC、全DRC等)增益与非缺省增益简档的一组非缺省(例如,全DRC和非DRC、全DRC等)增益之间的增益差,在一组差分增益中包含增益差,等等。作为发送与非缺省回放环境或情形相关的非缺省简档的(例如,非缺省等)动态压缩曲线的替代,作为作为由编码音频信号承载的元数据的一部分的非缺省动态压缩曲线的替代或者除其以外,音频编码器可发送一组差分增益。

[0048] 一组差分增益可能在尺寸上小于一组非缺省(例如,全DRC和非DRC、全DRC等)增益。因此,与直接传送非差分(例如,全DRC和非DRC、全DRC等)增益相比,传送差分增益而不是非差分(例如,全DRC和非DRC、全DRC等)增益可能需要更低的位率。

[0049] 接收这里描述的编码音频信号的音频解码器可能由不同的制造商提供,并且通过不同的部件和设计被实现。音频解码器可在不同的时间释放给最终用户,或者可通过不同版本的硬件、软件、固件被更新。作为结果,音频解码器可具有不同的音频处理能力。在一些实施例中,大量的音频解码器可配有支持诸如由标准、私有要求等规定的缺省增益简档的有限的一组增益简档的能力。大量的音频解码器可配有执行相关的增益生成动作的能力,以基于代表缺省增益简档的缺省动态范围压缩曲线生成缺省增益简档的增益。与传送对音频信号中的缺省增益简档生成/计算的增益相比,传送音频信号中的缺省增益简档的缺省动态范围压缩曲线可能更有效。

[0050] 另一方面,对于非缺省增益简档,音频编码器可参照与特定缺省增益简档对应的特定缺省动态范围压缩曲线重新生成差分增益。响应在由音频编码器生成的音频信号中接收差分增益,音频解码器可基于在音频信号中接收的缺省动态范围压缩曲线生成缺省增益、将接收的差分增益和生成的缺省增益组合成用于非缺省增益简档的非缺省增益、向从音频信号等解码的音频内容施加非缺省增益,同时呈现接收的音频内容。在一些实施例中,非缺省增益简档可被用于补偿缺省动态范围压缩曲线的限制。

[0051] 这里描述的技术可被用于对新的增益简档、特征或增强提供灵活的支持。在一些实施例中,至少一个增益简档不能容易地被缺省或非缺省的动态范围压缩曲线代表。在一些实施例中,至少一个增益简档可能是特定音频内容(例如,特定电影等)特有的。与编码的

音频信号所可承载的相比,非缺省增益简档的代表(例如,参数化DRC曲线、平滑化常数等)还可能在编码音频信号中需要传送更多的参数、平滑化常数等。在一些实施例中,至少一个增益简档可能是特定的音频内容提供方(例如,特定的工作室等)特有的。

[0052] 因此,这里描述的音频编码器能够通过实现用于新增益简档的增益生成动作以及用于与新增益简档相关的缺省增益简档的增益生成动作来带头支持新增益简档。下游受体音频解码器不必执行用于新增益简档的增益生成动作。而是,音频解码器可通过利用由音频编码器生成的非缺省差分增益来支持新增益简档,使得音频解码器不执行新增益简档的增益生成动作。

[0053] 在一些实施例中,在在编码音频信号中编码的简档有关元数据中,根据与一个或更多个(例如,缺省等)动态范围压缩曲线和一组或更多组(例如,非缺省等)差分增益对应的各增益简档,一个或更多个(例如,缺省等)动态范围压缩曲线和一组或更多组(例如,非缺省等)差分增益可被构建、索引等。在一些实施例中,一组非缺省差分增益与缺省动态范围压缩曲线之间的关系可在简档相关元数据中被指示。当在元数据中存在两个或更多个缺省动态范围压缩曲线时,这是特别有用的;或者,不在元数据中而在下游解码器中被限定,基于在简档相关元数据中指示的关系,受体音频解码器可确定哪个缺省动态范围压缩曲线要被用于生成一组缺省增益,这些缺省增益可然后出于生成非缺省增益的目的与接收的一组非缺省差分增益组合,以例如补偿缺省动态范围压缩曲线的限制。

[0054] 当假定在假设的音频解码器处的假设的回放环境、情形等时,这里描述的技术不要求音频解码器锁定(例如,不可逆等)可通过诸如音频编码器等的上游装置执行的音频处理中。例如,为了区分存在于音频内容中的不同响度水平、最小化边界响度水平(例如,最小或最大响度水平等)处或附近的音频感觉质量的损失、保持信道或信道的子集之间的空间平衡等,这里描述的解码器可被配置为基于特定的回放情形定制音频处理动作。

[0055] 接收具有动态范围压缩曲线、基准响度水平、起音时间、释放时间等的编码音频信号的音频解码器可确定在解码器处使用的特定回放环境,并且选择具有与特定回放环境对应的相应基准响度水平的特定压缩曲线。

[0056] 解码器可计算/确定从编码音频信号提取的音频内容的各单个部分(例如,音频数据块、音频数据帧等)中的响度水平,或者如果音频编码器计算和提供了编码音频信号中的响度水平则获得音频内容的各单个部分中的响度水平。基于音频内容的各单个部分中的响度水平、音频内容的前面部分中的响度水平、可用的情况下的音频内容的随后部分中的响度水平、特定压缩曲线、与特定回放环境或情形相关的特定简档等中的一个或更多个,解码器确定音频处理参数,诸如用于动态范围控制的增益(或DRC增益)、起音时间、释放时间等。音频处理参数还可包含用于对准对话响度水平与特定回放环境的特定基准响度水平(可以是用户可调整的)的调整。

[0057] 解码器根据音频处理参数施加包含(例如,多信道、多带等)动态范围控制、对话水平调整等的音频处理动作。由解码器执行的音频处理动作还可包含但不仅限于:基于作为选择的动态范围压缩曲线的一部分或者与其结合提供的起音和释放时间的增益平滑化、用于防止修剪的增益限制等。可通过不同(例如,可调整、阈值依赖、可控等)时间常数执行不同的音频处理动作。例如,通过相对短的时间常数(例如,瞬时、约5.3毫秒等),用于防止修剪的增益限制可被应用于各单个音频数据块、各单个音频数据帧等。

[0058] 在一些实施例中,解码器可被配置为从编码音频信号中的元数据提取ASA参数(例如,听觉事件边界的时间位置、事件确信性测量的时间依赖值等)并且基于提取的ASA参数控制听觉事件中的增益平滑化的速度(例如,使用用于在听觉事件边界处起音的短时间常数、使用长时间常数以在听觉事件内减慢增益平滑化等)。

[0059] 在一些实施例中,解码器还保持用于某个时间间隔或窗口的瞬时响度水平的直方图,并且,例如,通过修改时常常数,使用直方图以控制程序之间、程序与商业等之间的响度水平过渡的增益变化速度。

[0060] 在一些实施例中,解码器支持多于一个的扬声器配置(例如,具有扬声器的便携式模式、具有耳机的便携式模式、立体声模式、多信道模式等)。解码器可被配置为,例如,当回放同一音频内容时,在两个不同扬声器配置之间(例如,在立体声模式与多信道模式等之间)保持相同的响度水平。音频解码器可使用一个或更多个下混式,以对多信道音频内容被编码为音频解码器处的特定扬声器配置的基准扬声器配置下混从编码音频信号接收的多信道音频内容。

[0061] 在一些实施例中,自动增益控制(AGC)可在这里描述的音频解码器中被禁用。

[0062] 在一些实施例中,在这里描述的机制形成媒体处理系统的一部分,包含但不限于:视听装置、平板TV、手持装置、游戏机、电视、家庭剧院系统、平板、移动装置、膝上型计算机、笔记本计算机、蜂窝无线电电话、电子书阅读器、销售终端、台式计算机、计算机工作站、计算机亭、各种其它类型的终端和媒体处理单元等。

[0063] 对于本领域技术人员来说,这里描述的优选实施例和一般原理和特征的各种修改是容易理解的。因此,公开不是要限于所示出的实施例,而要被赋予与这里描述的原理和特征一致的最宽范围。

[0064] 2.动态范围控制

[0065] 在没有定制的动态范围控制的情况下,由于回放装置的特定回放环境可能与已在编码装置处编码编码音频内容的目标回放环境不同,因此,常在回放装置处以不适于回放装置的特定回放环境(例如,包含装置的物理和/或机械回放限制)的响度水平再现输入音频信息(例如,PCM采样、QMF矩阵中的时间频率采样等)。

[0066] 这里描述的技术可被用于支持对各种回放环境中的任一种定制的各种音频内容的动态范围控制,同时保持音频内容的感觉质量。

[0067] 动态范围控制(DRC)指的是将音频内容中的响度水平的输入动态范围修改(例如,压缩、切割、扩展、升高等)到与输入动态范围不同的输出动态范围中的时间依赖音频处理动作。例如,在动态范围控制方案中,柔声可被映射(例如,升高等)到更高的响度水平,并且,响的声音可被映射(例如,切割等)到低的响度值。作为结果,在响度域中,响度水平的输出范围在本例子中变得比响度水平的输入范围小。但是,在一些实施例中,动态范围控制可能是可逆的,使得原始范围被恢复。例如,只要从原始响度水平映射的输出动态范围中的映射响度水平处于修剪水平或者低于它、各唯一原始响度水平被映射到唯一输出响度水平,等等,就可执行扩展动作以恢复原始范围。

[0068] 可以使用这里描述的DRC技术以在某些回放环境或情况下提供更好的收听体验。例如,噪声环境中的柔声会被噪声掩盖,这使得柔声不可听。相反,例如,在一些情况下,为了不影响邻居,不希望大的声音。一般具有小的形状因子的扬声器的许多装置不能以高的

输出水平再现声音。在一些情况下,低的信号水平可能在人听力阈值之下被再现。DRC技术可基于用动态范围压缩曲线观看的DRC增益(例如,缩放音频振幅、升压比、切割比等的缩放因子)执行输入响度水平到输出响度水平的映射。

[0069] 动态范围压缩曲线指的是将从各单个音频数据帧确定的各单个输入响度水平(例如,对话以外的声音等的)映射到各单个增益或用于动态范围控制的增益的函数(例如,查找表、曲线、多段分段线等)。各单个增益中的每一个指示施加到相应的单个输入响度水平的增益量。施加各单个增益之后的输出响度水平代表特定回放环境中的各单个音频数据帧中的音频内容的目标响度水平。

[0070] 除了规定增益与响度水平之间的映射以外,动态范围压缩曲线可在施加增益时包含或者具有特定的释放时间和起音时间。起音指的是连续时间采样之间的信号能量(或响度)的增加,而释放指的是连续时间采样之间的能量(或响度)的减小。起音时间(例如,10毫秒、20毫秒等)指的是当相应信号处于起音模式中时在平滑化DRC增益中使用的时间常数。释放时间(例如,80毫秒、100毫秒等)指的是当相应信号处于释放模式中时在平滑化DRC增益中使用的时间常数。在一些实施例中,另外、任选地或者作为替代方案,在确定DRC增益之前对信号能量(或响度)的平滑化使用时间常数。

[0071] 不同的动态范围压缩曲线可与不同的回放环境对应。例如,用于平板TV的回放环境的动态范围压缩曲线可与用于便携式装置的回放环境的动态范围压缩曲线不同。在一些实施例中,回放装置可具有两个或更多个回放环境。例如,具有扬声器的便携式装置的第一回放环境的第一动态范围压缩曲线可与具有耳机的同一便携式装置的第二回放环境的第二动态范围压缩曲线不同。

[0072] 3.音频解码器

[0073] 图1A示出包括数据提取器104、动态范围控制器106、音频呈现器108等的示例性音频解码器100。

[0074] 在一些实施例中,数据提取器(104)被配置为接收编码输入信号102。这里描述的编码输入信号可以是包含编码(例如,压缩等)输入音频数据帧和元数据的位流。数据提取器(104)被配置为从编码输入信号(102)提取/解码输入音频数据帧和元数据。输入音频数据帧中的每一个包含多个编码音频数据块,每个编码音频数据块代表多个音频采样。各帧代表包含一定数量的音频采样的(例如,恒定的)时间间隔。帧尺寸可随采样率和编码数据率改变。音频采样可以是代表一个、两个或更多个(音频)频带或频率范围中的频谱内容的量化的音频数据要素(例如,输入PCM采样、QMF矩阵中的输入时间频率采样等)。输入音频数据帧中的量化的音频数据要素可代表数字(量化)域中的压力波。量化的音频数据要素可覆盖处于最大可能值处或低于它的有限范围的响度水平(例如,修剪水平、最大响度水平等)。

[0075] 元数据可被各种受体解码器使用以处理输入音频数据帧。元数据可包含与由解码器(100)执行的一个或更多个动作有关的各种动作参数、一个或更多个动态范围压缩曲线、与在输入音频数据帧中代表的对话响度水平有关的归一化参数等。对话响度水平可指的整个节目(例如,电影、TV节目、无线电广播等)、节目的一部分、节目的对话中的对话响度、节目响度、平均对话响度等的(心理声学、感觉等)水平。

[0076] 解码器(104)或模块中的一些或全部(例如,数据提取器104、动态范围控制器106等)的动作和功能可响应从编码输入信号(102)提取的元数据被适应性调整。例如,元数

据—包含但不限于动态范围压缩曲线、对话响度水平等—可被解码器(100)使用以在数字域中生成输出音频数据要素(例如,输出PCM采样、QMF矩阵中的输出时间频率采样等)。输出数据要素然后可被用于驱动音频信道或扬声器以在特定回放环境中的回放期间实现规定响度或基准再现水平。

[0077] 在一些实施例中,动态范围控制器(106)被配置为接收输入音频数据帧和元数据中的音频数据要素中的一些或全部、至少部分地基于从编码音频信号(102)提取的元数据在输入音频数据帧中的音频数据要素上执行音频处理动作(例如,动态范围控制动作、增益平滑化动作、增益限制动作等)。

[0078] 在一些实施例中,动态范围控制器(106)可包含选择器110、响度计算器112、DRC增益单元114等。选择器(110)可被配置为确定与解码器(100)处的特定回放环境有关的扬声器配置(例如,平面模式、具有扬声器的便携式装置、具有耳机的便携式装置、5.1扬声器配置、7.1扬声器配置等)、从由编码输入信号(102)提取的动态范围压缩曲线选择特定的动态范围压缩曲线等。

[0079] 响度计算器(112)可被配置为计算由输入音频数据帧中的音频数据要素代表的一种或更多种类型的响度水平。响度水平的类型的例子包括但不仅限于:各单个时间间隔上的各单个信道中的各单个频带上的各单个响度水平、各单个信道中的广(或宽)频率范围上的广带(或宽带)响度水平、从音频数据块或帧确定或在其上面平滑化的响度水平、从多于一个的音频数据块或帧确定或在其上面平滑化的响度水平、在一个或更多个时间间隔上平滑化的响度水平等中的任一个。出于通过解码器(100)的动态范围控制的目的,可以修改这些响度水平中的零个、一个或更多个。

[0080] 为了确定响度水平,响度计算器(112)可确定由输入音频数据帧中的音频数据要素代表的一个或更多个时间依赖物理声波性能,诸如特定音频频率处的空间压力水平等。响度计算器(112)可使用一个或更多个时间改变的物理波性能,以基于模型化人响度感觉的一个或更多个心理声学函数导出一种或更多种类型的响度水平。心理声学函数可以是基于人听觉系统的模型构建的非线性函数,该非线性函数将特定音频频率处的特定空间压力水平转换/映射到特定音频频率的特定响度等。

[0081] 可通过多个(音频)频率或多个频带上的特定响度水平的积分导出多个(音频)频率或多个频带上的(例如广带、宽带等)响度水平。可通过使用实现为解码器(100)中的音频处理动作的一部分的一个或更多个平滑化滤波器,获得一个或更多个时间间隔(例如,比由音频数据块或帧中的音频数据要素代表的时间间隔长)上的时间平均、平滑化等的响度水平。

[0082] 在示例性实施例中,可对一定(例如,256等)采样的每个音频数据块计算不同频带的特定响度水平。可以使用预滤波器以在将特定的响度水平集成到广带(或宽带)响度水平中时向特定响度水平施加频率加权(例如,与IEC B加权等类似)。可以执行两个或更多个信道(例如,左前、右前、中心、左环绕、右环绕等)上的宽响度水平的加算,以提供两个或更多个信道上的总响度水平。

[0083] 在一些实施例中,总响度水平可指的是扬声器配置的单个信道(例如,中心等)中的广带(宽带)响度水平。在一些实施例中,总响度水平可指的是多个信道中的广带(或宽带)响度水平。多个信道可以是扬声器配置中的所有信道。另外、任选地或者作为替代方案,

多个信道可包括扬声器配置中的信道的子集(例如,包含左前、右前和低频效果(LFE)的信道的子集、包含左环绕和右环绕的信道的子集、包含中心的信道的子集、等)。

[0084] (例如,广带、宽带、总体、特有等)响度水平可被用作输入以从选择的动态范围压缩曲线查找相应的(例如,静止、预平滑化、预限制等)DRC增益。用作输入以查找DRC增益的响度水平可首先关于来自从编码音频信号(102)提取的元数据的对话响度水平被调整或归一化。在一些实施例中,在在编码音频信号(102)中的音频内容的一部分中代表的特定空间压力水平被转换或映射成编码音频信号(102)中的音频内容的一部分的特定响度水平之前,与调整对话响度水平有关的调整和归一化可以但不限于在非响度域(例如,SPL域等)中的编码音频信号(102)中的音频内容的一部分上被执行。

[0085] 在一些实施例中,DRC增益单元(114)可配有DRC算法,以生成增益(例如,用于动态范围控制、用于增益限制、用于增益平滑化等)、向由输入音频数据帧中的音频数据要素代表的一种或更多种类型的响度水平中的一个或更多个响度水平施加增益以实现特定回放环境的目标响度水平,等等。这里描述的增益(例如,DRC增益等)的施加可以但不需要在响度域中发生。在一些实施例中,可基于直接对输入信号平滑化和施加的响度计算(可以是Sone或者仅仅是对对话响度水平补偿的SPL值,例如,没有转换)生成增益。在一些实施例中,这里描述的技术可向响度域中的信号施加增益,并然后将信号从响度域转换回(线性)SPL域并且通过评价向响度域中的信号施加增益之前和之后的信号计算要施加到信号的相应增益。比率(或在对数dB代表中代表时的差值)然后确定信号的相应增益。

[0086] 在一些实施例中,DRC算法通过多个DRC参数运算。DRC参数包含已通过上游编码器(例如,150等)计算并且嵌入到编码音频信号(102)中的对话响度水平,并且,可通过解码器(100)从编码音频信号(102)中的元数据获得。来自上游编码器的对话响度水平指示平均对话响度水平(例如,每个程序地,相对于全尺度1kHz正弦波的能量,相对于基准矩形波的能量,等)。在一些实施例中,从编码音频信号(102)提取的对话响度水平可被用于减少程序间响度水平差。在实施例中,在解码器(100)处的同一特定回放环境中的不同程序之间,基准对话响度水平可被设定为同一值。基于来自元数据的对话响度水平,DRC增益单元(114)可向程序中的各音频数据块施加对话响度有关增益,使得在程序的多个音频数据块上平均化的输出对话响度水平上升/下降到程序的(例如,预配置的、系统缺省的、用户可配置的、简档依赖的、等)基准对话响度水平。

[0087] 在一些实施例中,DRC增益可被用于通过根据选择的动态范围压缩曲线升高或切割柔声和/或大声中的信号部分来解决程序内响度水平差。可基于选择的动态范围压缩曲线和从相应的音频数据块、音频数据帧等中的一个或更多个确定的(例如,广带、宽带、总体、特有等)响度水平,通过DRC算法计算/确定这些DRC增益中的一个或更多个。

[0088] 可在短的间隔(例如,约5.3毫秒等)上计算用于通过查找选择的动态范围压缩曲线确定(静止、预平滑化、预增益限制等)DRC增益的响度水平。人听觉系统的集成时间可能长得多(例如,约200毫秒等)。从选择的动态范围压缩曲线获得的DRC增益可通过时间常数被平滑化以考虑人听觉系统的长的集成时间。为了实行响度水平中的快变化速率(增加或减小),可以使用短的时间常数以在与短时间常数对应的短时间间隔中导致响度水平的变化。相反,为了实行响度水平中的慢变化速率(增加或减小),可以使用长的时间常数以在与长时间常数对应的长时间间隔中导致响度水平的变化。

[0089] 人听觉系统能够以不同的集成时间对增加的响度水平和减小的响度水平起反应。在一些实施例中,根据响度水平将增加还是减小,可对平滑化从选择的动态范围压缩曲线查找的静止DRC增益使用不同的时间常数。例如,与人听觉系统的特性对应,用相对短的时间常数(例如,起音时间等)平滑化起音(响度水平增加),而用相对长的时间常数(例如,释放时间等)平滑化释放(响度水平减小)。

[0090] 可通过使用从音频内容的一部分确定的响度水平计算音频内容的一部分(例如,音频数据块、音频数据帧等中的一个或更多个)的DRC增益。用于查找选择的动态范围压缩曲线的响度水平可首先关于(例如,相对于等)从编码音频信号(102)提取的元数据中的对话响度水平(例如,在音频内容是一部分的程序等中)被调整。

[0091] 可对解码器(100)处的特定回放环境规定或建立基准对话响度水平(例如,“Line”模式中的-31$dB_{FS}$、“RF”模式中的-20$dB_{FS}$等)。另外、作为替代方案或者任选地,在一些实施例中,用户可控制解码器(100)处的基准对话响度水平的设定或改变。

[0092] DRC增益单元(114)可被配置为对音频内容确定对话响度有关增益,以导致从对于基准对话响度水平的输入对话响度水平变为输出对话响度水平。

[0093] 在一些实施例中,DRC增益单元(114)可被配置为操作解码器(100)处的特定回放环境中的峰值水平并且调整DRC增益以防止修剪。在一些实施例中,根据第一方法,如果从编码音频信号(102)提取的音频内容包含具有比解码器(100)处的特定扬声器配置的信道多的信道的基准多信道配置的音频数据要素,那么从基准多信道配置到特定扬声器配置的下混合可在确定之前被执行并且出于修剪防止的目的操作峰值水平。另外、任选地或者作为替代方案,在一些实施例中,根据第二方法,如果从编码音频信号(102)提取的音频内容包含具有比解码器(100)处的特定扬声器配置的信道多的信道的基准多信道配置的音频数据要素,那么可以使用下混合式(例如,ITU立体声下混合、矩阵环绕可兼容下混合等)以获得解码器(100)处的特定扬声器配置的峰值水平。峰值水平可被调整以反映从输入对话响度水平变为作为输出对话响度水平的基准对话响度水平。可至少部分地基于峰值水平的逆转(例如,乘以-1等),确定不导致修剪(例如,对于音频数据块、对于音频数据帧等)的最大允许增益。因此,根据这里描述的技术的音频解码器可被配置为精确地确定峰值水平并且专门对解码器侧的回放配置施加修剪防止;音频解码器和音频编码器均不需要进行关于假设解码器处的任何最坏情况情形的假设假定。特别地,上述的第一方法中的解码器可精确地确定峰值水平,并且在不使用会在上述的第二方法中使用的下混合式、下混合信道增益等的情况下在下混合之后施加修剪防止。

[0094] 在一些实施例中,对话响度水平和DRC增益的组合调整在峰值水平上防止修剪,甚至可能在最坏情况下混合(例如,在下混合之后生成最大峰值水平、生成最大下混合信道增益等)中。但是,在一些其它的实施例中,对话响度水平和DRC增益的组合调整仍然可能不足以在峰值水平下防止修剪。在这些实施例中,DRC增益可被确实防止峰值水平中的修剪的最高增益替代(例如,封盖(capped))。

[0095] 在一些实施例中,DRC增益单元(114)被配置为从从编码音频信号(102)提取的元数据得到时间常数(例如,起音时间、释放时间等)。DRC增益、时间常数、最大允许增益等可被DRC增益单元(114)使用以执行DRC、增益平滑化、增益限制等。

[0096] 例如,DRC增益的施加可通过由时间常数控制的滤波器被平滑化。增益限制动作可

由对增益取要施加的增益和最大允许增益中的较小者的min()函数实现,这样,(例如,预限制、DRC等)增益可在相对短时间间隔等上立即被最大允许增益替代,由此防止修剪。

[0097] 在一些实施例中,音频呈现器(108)被配置为在向从编码音频信号(102)提取的输入音频数据施加基于DRC、增益限制、增益平滑化等确定的增益之后生成特定扬声器配置的(例如,多信道等)信道特定音频数据(116)。信道特定音频数据(118)可被用于驱动在扬声器配置中代表的扬声器、耳机等。

[0098] 另外,并且/或者,任选地,在一些实施例中,解码器(100)可被配置为执行与预处理、后处理、呈现等有关的一个或更多个其它动作,这些动作与输入音频数据有关。

[0099] 在这里描述的技术可与与各种不同的环绕声配置(例如,2.0、3.0、4.0、4.1、4.1、5.1、6.1、7.1、7.2、10.2、10～60扬声器配置、60+扬声器配置、物体信号或物体信号的组合等)对应的各种扬声器配置和各种不同呈现环境配置(例如,电影院、停车场、歌剧院、演奏厅、酒吧、家、会堂)一起使用。

[0100] 4.音频编码器

[0101] 图1B示出示例性编码器150。编码器(150)可包含音频内容接口152、对话响度分析器154、DRC基准库156、音频信号编码器158等。编码器150可以是广播系统、基于因特网的内容服务器、空中网络操作器系统、电影制作系统等的一部分。

[0102] 在一些实施例中,音频内容接口(152)被配置为接收音频内容160、音频内容控制输入162等,至少部分地基于音频内容(160)、音频内容控制输入(162)等中的一些或全部生成编码音频信号(例如,102)。例如,音频内容接口(152)可被用于从内容创建器、内容提供方等接收音频内容(160)、音频内容控制输入(162)。

[0103] 音频内容(160)可构成仅包含音频视听等的总媒体数据中的一些或全部。音频内容(160)可包含节目的多个部分、节目、几个节目、一个或更多个商业广告等中的一个或更多个。

[0104] 在一些实施例中,对话响度分析器(154)被配置为确定/建立音频内容(152)的一个或更多个部分(例如,一个或更多个节目、一个或更多个商业广告等)的一个或更多个对话响度水平。在一些实施例中,音频内容由一组或更多组的音频轨道代表。在一些实施例中,音频内容的对话音频内容处于单独的音频轨道中。在一些实施例中,音频内容的对话音频内容的至少一部分处于包含非对话音频内容的音频轨道中。

[0105] 音频内容控制输入(162)可包含用户控制输入、由编码器(510)外部的系统/装置提供的控制输入、来自内容创建者的控制输入、来自内容提供方的控制输入等中的一些或全部。例如,诸如混合工程师等的用户可提供/规定一个或更多个动态范围压缩曲线标识符;标识符可被用于从诸如DRC基准库(156)等的数据库检索最佳地拟合音频内容(160)的一个或更多个动态范围压缩曲线。

[0106] 在一些实施例中,DRC基准库(156)被配置为存储DRC基准参数组等。DRC基准参数组可包含一个或更多个动态范围压缩曲线的定义数据等。在一些实施例中,编码器(150)可(例如,同时等)将多于一个的动态范围压缩曲线编码到编码的音频信号中(102)。动态范围压缩曲线中的零个、一个或更多个可以是基于标准的、专有的、定制的、解码器可修改的,等等。在示例性实施例中,图2A和图2B的两个动态范围压缩曲线均可被(例如,同时等)嵌入到编码的音频信号(102)中。

[0107] 在一些实施例中,音频信号编码器(158)可被配置为从音频内容接口(152)接收音频内容、从对话响度分析器(154)等接收对话响度水平、从DRC基准库(156)检索一个或更多个DRC基准参数组、将音频内容格式化为音频数据块/帧、将对话响度水平、DRC基准参数组等格式化为元数据(例如,元数据容器、元数据栏、元数据结构等)、将音频数据块/帧和元数据编码到编码的音频信号中(102),等等。

[0108] 可通过各种方式中的一种或更多种、诸如以无线的方式、通过有线连接、通过文件、通过因特网下载等,在各种源音频格式中的一个或更多个中接收这里描述的要被编码成编码的音频信号的音频内容。

[0109] 这里描述的编码音频信号可以是总媒体数据位流的一部分(例如,用于音频广播、音频节目、视听节目、视听广播等)。可从服务器、计算机、媒体存储装置、媒体数据库、媒体文件等访问媒体数据位流。可通过一个或更多个无线或有线网络链接广播、传送或接收媒体数据位流。也可通过诸如网络连接、USB连接、广域网络、局域网络、无线连接、光学连接、总线、交叉开关矩阵(crossbar)连接、串行连接等中的一个或更多个的中介传送媒体数据位流。

[0110] 示出的部件中的任一个(例如,图1A、图1B等)可在硬件、软件或硬件和软件的组合中实现为一个或更多个处理和/或一个或更多个IC电路(例如,ASIC、FPGA等)。

[0111] 5.动态范围压缩曲线

[0112] 图2A和图2B示出可被解码器(100)中的DRC增益单元(104)使用以从输入响度水平导出DRC增益的示例性动态范围压缩曲线。如图所示,为了提供适于特定回放环境的总增益,动态范围压缩曲线可以以节目中的基准响度水平为中心。在下表中示出动态范围压缩曲线(例如,包含但不仅限于升压比、切割比、起音时间、释放时间等中的任一个)的示例性定义数据(例如,在编码音频信号102等的元数据中),这里,多个简档(例如,电影标准、电影光、音乐标准、音乐光、语音等)中的各简档代表特定回放环境(例如,在解码器100等处)。

[0113] 表1

[0114]

| 简档 / 参数 | 电影标准 | 电影光 | 音乐标准 | 音乐光 | 语音 |
|---|---|---|---|---|---|
| 时间常数选择 | | | | | |
| 起音阈值(dB) | 15 | 15 | 15 | 15 | 10 |
| 释放阈值(dB) | 20 | 20 | 20 | 20 | 10 |
| 快速起音时间常数(ms) | 10 | 10 | 10 | 10 | 10 |

[0115]

| | | | | | |
|---|---|---|---|---|---|
| 慢速起音时间常数（ms） | 100 | 100 | 100 | 100 | 100 |
| 慢速释放时间常数（ms） | 3000 | 3000 | 10000 | 3000 | 1000 |
| 快速释放时间常数（ms） | 1000 | 1000 | 1000 | 1000 | 200 |
| 推迟周期（ms） | 53 | 53 | 53 | 53 | 53 |
| 压缩曲线 | | | | | |
| 最大升压（dB） | 6 | 6 | 12 | 12 | 15 |
| 最大升压范围（<=dB） | -43 | -53 | -55 | -65 | -50 |
| 升压比 | 2:1 | 2:1 | 2:1 | 2:1 | 19:15 |
| 升压范围（dB） | [-43,..., -31] | [-53,..., -41] | [-55,..., -31] | [-65,..., -41] | [-50,..., -31] |
| 空带宽（dB） | 5 | 20 | 5 | 20 | 5 |
| 空带范围（dB） | [-31,..., -26] | [-41,..., -21] | [-31,..., -26] | [-41,..., -21] | [-31,..., -26] |
| 切割比 | 2:1 | 2:1 | 2:1 | 2:1 | 2:1 |
| 切割比范围（dB） | [-26,..., -16] | [-21,..., -11] | [-26,..., -16] | [-21,..., 27] | [-26,..., -16] |
| 二次切割比（dB） | 20:19 | 20:19 | 20:19 | | 20:19 |
| 二次切割比范围（dB） | [-16,..., 4] | [-11,..., 9] | [-16,..., 4] | | [-16,..., 4] |
| 最大切割（dB） | -24 | -24 | -24 | -24 | -24 |
| 最大切割范围（>=dB） | 4 | 9 | 4 | 27 | 4 |

[0116] 一些实施例可接收以$dB_{SPL}$或$dB_{FS}$的响度水平和与$dB_{SPL}$相关的dB的增益描述的一个或更多个压缩曲线，其中，在与$dB_{SPL}$响度水平具有非线性关系的不同响度代表（例如，Sone）中执行DRC增益计算。在DRC增益计算中使用的压缩曲线可然后被转换以关于不同的响度代表（例如，Sone）被描述。

[0117] 6.DRC增益、增益限制和增益平滑化

[0118] 图3示出组合DRC和限制增益的确定/计算的示例性处理逻辑。可通过解码器（100）、编码器（150）等实现该处理逻辑。仅出于解释的目的，解码器（例如，100等）中的DRC增益单元（例如，114）可被用于实现该处理逻辑。

[0119] 可通过使用从音频内容的一部分确定的响度水平计算音频内容的一部分（例如，音频数据块、音频数据帧等中的一个或更多个等）的DRC增益。响度水平可首先关于（例如，相对于等）从编码音频信号（102）提取的元数据中的对话响度水平（例如，在音频内容是一部分的程序等中）被调整。在图3所示的例子中，音频内容的一部分的响度水平与对话响度水平（“dialnorm”）之间的差值可被用作用于从选择的动态范围压缩曲线查找DRC增益的输入。

[0120] 为了防止特定回放环境中的输出音频数据要素中的修剪，DRC增益单元（114）可被配置为操作特定回放情形（例如，编码音频信号102与解码器100处的回放环境的特定组合等特有的）中的峰值水平，该回放情形可以是各种可能的回放情形（例如，多信道情形、下混合情形等）中的一个。

[0121] 在一些实施例中,可作为从编码音频信号(102)提取的元数据的一部分提供特定时间分辨率(例如,音频数据块、几个音频数据块、音频数据帧等)的音频内容的各单个部分的各单个峰值水平。

[0122] 在一些实施例中,DRC增益单元(114)可被配置为确定这些情形中的峰值水平,并且在必要时调整DRC增益。在DRC增益的计算过程中,可通过DRC增益单元(114)使用并行处理以确定音频内容的峰值水平。例如,可对于具有比由解码器(100)使用的特定扬声器配置的信道多的信道的基准多信道配置编码音频内容。基准多信道配置的更多信道的音频内容可被转换成下混合音频数据(例如,ITU立体声下混合、矩阵环绕可兼容下混合等),以对解码器(100)处的特定扬声器配置导出更少的信道。在一些实施例中,根据第一方法,可在出于修剪防止的目的确定和操作峰值水平之前执行从基准多信道配置到特定扬声器配置的下混合。另外、任选地或者作为替代方案,在一些实施例中,根据第二方法,可以使用与下混合音频内容有关的下混合信道增益作为用于调整、导出、计算等特定扬声器配置的峰值水平的输入的一部分。在示例性实施例中,可至少部分地基于用于在解码器(100)处的回放环境中实施从基准多信道配置到特定扬声器配置的下混合动作的一个或更多个下混合式导出下混合信道增益。

[0123] 在一些媒体应用中,可对解码器(100)处的特定回放环境规定或假定基准对话响度水平(例如,“Line”模式中的-31$dB_{FS}$、“RF”模式中的-20$dB_{FS}$等)。在一些实施例中,用户可控制解码器(100)处的基准对话响度水平的设定或改变。

[0124] 可以向音频内容施加对话响度相关增益,以将(例如,输出)对话响度水平调整到基准对话响度水平。因此,峰值水平应被调整以反映这种调整。在例子中,(输入)对话响度水平可以为-23$dB_{FS}$。在具有-31$dB_{FS}$的基准对话响度水平的“Line”模式中,向(输入)对话响度水平的调整为-8dB,以在基准对话响度水平处生成输出对话响度水平。在“Line”模式中,向峰值水平的调整也为-8dB,这与向对话响度水平的调整相同。在具有-20$dB_{FS}$的基准对话响度水平的“RF”模式中,向(输入)对话响度水平的调整为-3dB,以在基准对话响度水平处生成输出对话响度水平。在“RF”模式中,向峰值水平的调整也为3dB,这与向对话响度水平的调整相同。

[0125] 峰值水平与基准对话响度水平(表示为“dialref”)和来自编码音频信号(102)的元数据中的对话响度水平(“dialnorm”)之间的差值的和可被用作计算DRC增益的最大(例如,允许的,等)增益的输入。由于调整的峰值水平以$dB_{FS}$表达(相对于0$dB_{FS}$上的修剪水平),因此,不导致修剪(例如,对于当前音频数据块、对于当前音频数据帧等)的最大允许增益只是调整的峰值水平的逆转(例如,乘以-1等)。

[0126] 在一些实施例中,即使导出DRC增益的动态范围压缩曲线被设计为在一定程度上切割大声,峰值水平也可能超过修剪水平(表示为0$dB_{FS}$)。在一些实施例中,甚至可能在最坏情况混合(例如,生成最大下混合信道增益等)中,对话响度水平和DRC增益的组合调整防止峰值水平的修剪。但是,在一些其它的实施例中,对话响度水平和DRC增益的组合调整仍可能不足以在峰值水平中防止修剪。在这些实施例中,DRC增益可被确实防止峰值水平中的修剪的最高增益替代(例如,封盖等)。

[0127] 在一些实施例中,DRC增益单元(114)被配置为从从编码音频信号(102)提取的元数据得到时间常数(例如,起音时间、释放时间等)。这些时间常数可能或者可能不随对话响

度水平或音频内容的当前响度水平中的一个或更多个改变。从动态范围压缩曲线、时间常数和最大增益查找的DRC增益可被用于执行增益平滑化和限制动作。

[0128] 在一些实施例中,可能增益受限的DRC增益在特定的回放环境中不超过最大峰值响度水平。从响度水平导出的静止DRC增益可通过由时间常数控制的滤波器被平滑化。限制动作可由一个或更多个min()函数实现,这样,(预限制)DRC增益可在短时间间隔等上立即被最大允许增益替代,由此防止修剪。DRC算法可被配置为随着进入的音频内容的峰值水平从高于修剪水平移动到低于修剪水平平稳地从修剪增益释放到下增益。

[0129] 可以使用一个或更多个不同(例如,实时、双行程等)实现来执行图3所示的DRC增益的确定/计算/施加。仅出于解释的目的,向对话响度水平的调整、(例如,静止等)DRC增益、由于平滑化导致的时间依赖增益变动、由于限制导致的增益修剪等已被描述为来自上述的DRC算法的组合增益。但是,在各种实施例中,可以使用为了控制对话响度水平(例如,在不同节目之间等)、为了动态范围控制(例如,对于同一节目的不同部分等)、为了防止修剪、为了增益平滑化等向音频内容施加增益的其它方法。例如,向对话响度水平的调整、(例如,静止等)DRC增益、由于平滑化导致的时间依赖增益变动、由于限制导致的增益修剪中的一些或全部可被部分地/单个地施加、被串行施加、被并行施加、被部分串行部分并行地施加,等等。

[0130] 7.输入平滑化和增益平滑化

[0131] 除了DRC增益平滑化以外,可以在各种实施例中实现根据这里描述的技术的其它平滑化处理。在例子中,可以使用输入平滑化以例如用简单的单极平滑化滤波器平滑化从编码音频信号(102)提取的输入音频数据,以获得具有比没有输入平滑化的特定响度水平的频谱好的时间特性(例如,在时间上更平稳、在时间上不波动等)的特定响度水平的频谱。

[0132] 在一些实施例中,这里描述的不同的平滑化处理可使用不同的时间常数(例如,1秒、4秒等)。在一些实施例中,两个或更多个平滑化处理可使用同一时间常数。在一些实施例中,在这里描述的平滑化处理中使用的时间常数可以是频率依赖的。在一些实施例中,在这里描述的平滑化处理中使用的时间常数可以是频率无关的。

[0133] 一个或更多个平滑化处理可与支持一个或更多个平滑化处理的自动或手动复位的复位处理连接。在一些实施例中,当在复位处理中出现复位时,平滑化处理可通过切换或转移到更小的时间常数加速平滑化动作。在一些实施例中,当在复位处理中出现复位时,平滑化处理的存储器可被复位到一定值。该值可以是对于平滑化处理的最后输入采样。

[0134] 8.多个频带上的DRC

[0135] 在一些实施例中,可以使用特定频带中的特定响度水平以导出特定频带中的相应DRC增益。但是,即使当所有频带上的广带(或宽带)响度水平保持恒定时,由于特定响度水平会在不同的带中也明显改变并由此招致不同的DRC增益,因此,这会导致音色变化。

[0136] 在一些实施例中,不是施加随各单个频带改变的DRC增益,而是施加不随频带改变而随时间改变的DRC增益。跨着所有的频带施加相同的时间改变DRC增益。时间改变DRC增益的时间平均DRC增益可被设定为与基于广带、宽带和/或广带(或宽带)范围或多个频带上的总响度水平从选择的动态范围压缩曲线导出的静止DRC增益相同。作为结果,可以防止在其它的方法中通过在不同的频带中施加不同的DRC增益导致的音色效果的变化。

[0137] 在一些实施例中,用基于广带(或宽带)响度水平确定的广带(或宽带)DRC增益控

制各单个频带中的DRC增益。各单个频带中的DRC增益可在基于广带(或宽带)响度水平在动态范围压缩曲线中查找的广带(或带宽)DRC增益周围动作,使得在一定的时间间隔(例如,比5.3毫秒、20毫秒、50毫秒、80毫秒、100毫秒等长)上时间平均的各单个频带中的DRC增益与在动态范围压缩曲线中指示的广带(或宽带)响度水平相同。在一些实施例中,在信道和/或频带之间,相对于偏离时间平均DRC增益的一定时间间隔的短时间间隔上的响度水平波动是允许的。所述方法确保在动态范围压缩曲线中指示的正确多信道和/或多带时间平均DRC增益的施加,并且防止短时间间隔中的DRC增益太多地偏离在动态范围压缩曲线中指示的这种时间平均DRC增益。

[0138] 9.响度域中的音量调整

[0139] 根据不实现这里描述的技术的其它方法向音频激励信号施加用于音量调整的线性处理可导致低可听信号水平变得不可听(例如,低于人听觉系统的频率依赖听力阈值等)。

[0140] 根据这里描述的技术,可在响度域中(例如,通过Sone代表等)而不是在物理域(例如,通过$dB_{SPL}$代表等)中进行或实现音频内容的音量调整。在一些实施例中,出于保持感觉质量和/或所有音量水平上的所有带之间的响度水平关系的完整性的目的,在响度域中用相同的因子缩放所有带中的响度水平。基于设定和调整这里描述的响度域中的增益的音量调整可被转换回向不同频带中的音频激励信号施加不同缩放因子的物理域(或者在代表物理域的数字域中)中的非线性处理或者通过其被实现。根据这里描述的技术从响度域中的音量调整转换的物理域中的非线性处理用防止音频内容中的低可听水平中的大多数或全部变得不可听的DRC增益来衰减或增强音频内容的响度水平。在一些实施例中,节目内的大声和柔声之间的响度水平差通过这些DRC增益减小但不在感觉上消失,以使低可听信号水平保持高于人听觉系统的听力阈值。在一些实施例中,为了跨着大范围的音量水平保持频谱感觉和感觉音色等的类似性,在低的音量水平上,具有接近听力阈值的激励信号水平的频率或频带很少衰减并因此在感觉上是可听的。

[0141] 这里描述的技术可实现物理域(例如,或在代表物理域的数字域中)中的信号水平、增益等与响度域中的响度水平、增益等之间的转换(例如,往复,等)。这些转换可基于基于人听觉系统的模型构建的一个或更多个非线性函数(例如,映射、曲线、分段线性段、查找表等)的正反版本。

[0142] 10.具有差分增益的增益简档

[0143] 在一些实施例中,这里描述的音频编码器(例如,150等)被构建为向下游的音频解码器提供简档相关元数据。例如,可作为音频相关元数据的一部分连同音频内容一起在编码音频信号中承载简档相关元数据。

[0144] 这里描述的简档相关元数据包含但不限于多个增益简档的定义数据。多个增益简档中的一个或更多个第一增益简档(表示为一个或更多个缺省增益简档)由定义数据包含于简档相关元数据中的一个或更多个相应DRC曲线(表示为一个或更多个缺省DRC曲线)代表。多个增益简档中的一个或更多个第二增益简档(表示为一个或更多个非缺省增益简档)由定义数据包含于简档相关元数据中的关于一个或更多个缺省DRC曲线的一个或更多个相应组差分增益代表。更具体而言,缺省DRC曲线(例如,在简档相关元数据等中)可被用于代表缺省增益简档;关于缺省增益简档的一组差分增益(例如,在简档相关元数据等中)可被

用于代表非缺省增益简档。

[0145] 在一些实施例中,一组差分增益—其代表关于代表缺省增益简档的缺省DRC曲线的非缺省增益简档—包含对非缺省增益简档生成的一组非差分(例如,非缺省等)增益与对缺省增益简档生成的一组非差分(例如,缺省等)增益之间的增益差(或增益调整)。非差分增益的例子包含但不仅仅限于以下方面中的任一个:零增益、DRC增益或衰减、与对话归一化相关的增益或衰减、与增益限制相关的增益或衰减、与增益平滑化相关的增益或衰减,等等。这里描述的增益(例如,非差分增益、差分增益等)可以是时间依赖的,并且,可具有随时间改变的值。

[0146] 为了生成增益简档(例如,缺省增益简档、非缺省增益简档等)的一组非差分增益,这里描述的音频编码器可执行增益简档特有的一组增益生成动作。该组增益生成动作可包括DRC动作、增益限制动作、增益平滑化动作等,这些动作包括但不限于以下动作中的任一个:(1)全局适用于所有增益简档、(2)一个或更多个而不是所有增益简档特有、一个或更多个缺省DRC曲线特有、(3)一个或更多个非缺省DRC曲线特有;(4)相应(例如,缺省、非缺省等)增益简档特有;(5)与超出由媒体编码格式、媒体标准、媒体专有规范等支持的参数化的限制的算法、曲线、函数、动作、参数等中的一个或更多个相关;(6)与还没有在本领域中在音频解码装置中普遍实现的算法、曲线、函数、动作、参数等中的一个或更多个相关;等等。

[0147] 在一些实施例中,音频编码器(150)可被配置为至少部分地基于由缺省DRC曲线(例如,具有编码音频信号的简档相关元数据中的定义数据等)代表的缺省增益简档和与缺省增益简档不同的非缺省增益简档确定用于音频内容(1512)的一组差分增益,并且,作为编码音频信号中的简档相关元数据的一部分包含作为非缺省增益简档的代表(例如,与缺省DRC曲线等相关)的一组差分增益。关于缺省DRC曲线从编码音频信号中的简档相关元数据提取的该组差分增益可被受体音频解码器使用,以在回放环境或情形中对关于缺省DRC曲线由一组差分增益代表的特定增益简档有效且一致地执行增益动作(或衰减动作)。这允许受体音频解码器在不需要受体音频解码器实现可在音频编码器(150)中实现的一组增益生成动作的情况下施加特定增益简档的增益或衰减,以生成增益或衰减。

[0148] 在一些实施例中,一组或更多组差分增益可通过音频编码器(150)包含于简档相关元数据中。可关于一个或更多个缺省增益简档中的相应缺省增益简档从一个或更多个非缺省增益简档中的相应非缺省增益简档导出一组或更多组差分增益中的每一个。例如,可关于第一缺省增益简档从第一非缺省增益简档导出一组或更多组差分增益中的第一组差分增益,而可关于第二缺省增益简档从第二非缺省增益简档导出差分增益组中的第二组差分增益。

[0149] 在一些实施例中,第一组差分增益包含在基于第一非缺省增益简档生成的第一组非差分非缺省增益与基于第一缺省增益简档生成的第一组非差分缺省增益之间确定的第一增益差(或增益调整),而第二组差分增益包含在基于第二非缺省增益简档生成的第二组非差分非缺省增益与基于第二缺省增益简档生成的第二组非差分缺省增益之间确定的第二增益差。

[0150] 第一缺省增益简档和第二缺省增益简档可相同(例如,由与同一组增益生成动作相同的缺省DRC曲线代表,等等)或者不同(例如,由不同的缺省DRC曲线代表,由具有不同的组的增益生成动作的缺省DRC曲线代表,等等)。在各种实施例中、另外、任选地或者作为替

代方案,第一非缺省增益简档可以或者可以不与第二非缺省增益简档相同。

[0151] 由音频编码器(150)生成的简档相关元数据可承载一个或更多个特定的标记、指示、数据栏等,以指示一个或更多个相应非缺省增益简档的一组或更多组差分增益的存在。简档相关数据还可包含偏好标记、指示、数据栏等,以指示哪个非缺省增益简档优选用于在特定回放环境或情形中呈现音频内容。

[0152] 在一些实施例中,这里描述的音频解码器(例如,100等)被配置为从编码音频信号(102)解码(例如,多信道等)音频内容、从用音频内容传输的响度元数据提取对话响度水平(例如,“dialnorm”等),等等。

[0153] 在一些实施例中,音频解码器(例如,100等)被配置为对诸如第一缺省简档、第二缺省简档等的增益简档执行至少一组增益生成动作。例如,音频解码器(100)可用对话响度水平(例如,“dialnorm”等)解码经编码的音频信号(102);执行一组增益生成动作以获得由可通过音频解码器(100)从编码的音频信号(102)提取定义数据的缺省DRC曲线代表的缺省增益简档的一组非差分缺省增益(或衰减);在解码的过程中施加缺省增益简档的一组非差分缺省增益(例如,基准响度水平与“dialnorm”之间的差值等),以将输出的声音的输出对话响度水平对准/调整到基准响度水平;等等。

[0154] 另外、任选地或者作为替代方案,在一些实施例中,音频解码器(100)被配置为从编码的音频信号(102)提取关于诸如上述的缺省DRC曲线代表非缺省增益简档的至少一组差分增益作为用音频内容传输的元数据的一部分。在一些实施例中,简档相关元数据包含一个或更多个不同组的差分增益—一个或更多个不同组的差分增益中的每一个代表关于代表缺省增益简档的各缺省DRC曲线的非缺省增益简档。可通过在简档相关元数据中承载的一个或更多个标记、指示、数据栏指示在简档相关元数据中存在DRC曲线或一组差分增益。

[0155] 响应于确定存在一组或更多组差分增益,音频解码器(100)可在一个或更多个不同组的差分增益中确定/选择与特定非缺省增益简档对应的一组差分增益。音频解码器(100)可进一步被配置为识别缺省DRC曲线—例如,在简档相关元数据中的一个或更多个不同缺省DRC曲线的定义数据中—关于该缺省DRC曲线,该组差分增益代表特定增益简档。

[0156] 在一些实施例中,音频解码器(100)被配置为执行一组增益生成动作以获得缺省增益简档的一组非差分缺省增益(或衰减)。由音频解码器(100)执行以基于缺省DRC曲线获得一组非差分缺省增益的一组增益生成动作可包含关于一个或更多个标准、专有规范等的一个或更多个动作。在一些实施例中,音频解码器(100)被配置为基于从简档相关元数据提取定义数据的一组差分增益生成特定非缺省增益简档的一组非差分非缺省增益,并且基于缺省DRC曲线生成由一组增益生成动作生成的一组非差分缺省增益;在解码中对非缺省增益简档施加一组非差分非缺省增益(例如,基准响度水平与“dialnorm”之间的差值等),以将输出的声音的输出对话响度水平对准/调整到基准响度水平;等等。

[0157] 在一些实施例中,音频解码器(100)能够执行一个或更多个增益简档的增益相关动作。音频解码器(100)可被配置为基于一个或更多个因素确定和执行特定增益简档的增益相关动作。这些因素可包括但不仅限于以下方面中的一个或更多个:规定特定用户选择增益简档的偏好的用户输入;规定系统选择的增益简档的偏好的用户输入、由音频解码器(100)使用的特定扬声器或音频信道配置的能力;音频解码器(100)的能力;特定增益简档

的简档相关元数据的可用性、增益简档的任何编码器生成偏好标记等。在一些实施例中,音频解码器(100)可实现一个或更多个过程规则、可征求进一步的用户输入等,以在这些因素相互冲突时确定或选择特定增益简档。

[0158] 11.与增益有关的附加动作

[0159] 根据在这里描述的技术,不是在物理域(或代表物理域的数字域)中,而是可在响度(例如,感觉)域中执行诸如动态均衡化、噪声补偿等的其它处理。

[0160] 在一些实施例中,来自诸如DRC、均衡化噪声补偿、修剪防止、增益平滑化等的各种处理中的一些或全部的增益可在响度域中的相同增益中被组合,并且/或者可被并行施加。在一些其它的实施例中,来自诸如DRC、均衡化噪声补偿、修剪防止、增益平滑化等的各种处理中的一些或全部的增益可处于响度域中的单独的增益中,并且/或者可至少部分地被串行施加。在一些其它的实施例中,来自诸如DRC、均衡化噪声补偿、修剪防止、增益平滑化等的各种处理中的一些或全部的增益可被依次施加。

[0161] 12.特定和广带(或宽带)响度水平

[0162] 一个或更多个音频处理要素、单元、部件等,诸如传送滤波器、听觉滤波器组、合成滤波器组、短时傅立叶变换等,可被编码器或解码器使用,以执行在这里描述的音频处理动作。

[0163] 在一些实施例中,可以使用模型化人听觉系统的外和中耳过滤的一个或更多个传送滤波器,以过滤进入的音频信号(例如,编码的音频信号102、来自内容提供方的音频内容等)。在一些实施例中,可以使用听觉滤波器组以模型化人听觉系统的频率选择性和频率展度。来自这些滤波器中的一些或全部的激励信号水平可通过向着更高的频率的更短的频率依赖时间常数被确定/计算和平滑化,以模型化人听觉系统中的能量的积分。随后,可以使用激励信号与特定响度水平之间的非线性函数(例如,关系、曲线等),以获得频率依赖特定响度水平的简档。可通过在频带上积分特定响度获得广带(或宽带)响度水平。

[0164] 特定响度水平的直接(例如,对所有频带具有相同的权重等)加算/积分可对广带信号很好地起作用。但是,这种方法可能低估窄带信号的(例如,感觉等)响度水平。在一些实施例中,不同频率或者不同频带中的特定响度水平被赋予不同的权重。

[0165] 在一些实施例中,上述的听觉滤波器组和/或传送滤波器可被一个或更多个短时傅立叶变换(STFT)替代。可在快速傅立叶变换(FFT)域中施加对传送滤波器和听觉滤波器组的响应。在一些实施例中,例如,当在从物理域(或在代表物理域的数字域中)到响度域的转换中或者之前使用一个或更多个(例如,向前等)传送滤波器时,使用一个或更多个逆传送滤波器。在一些实施例中,例如,当作为听觉滤波器组和/或传送滤波器的替代使用STFT时,不使用逆传送滤波器。在一些实施例中,省略听觉滤波器组;作为替代,使用一个或更多个正交镜像滤波器(QMF)。在这些实施例中,可在不明显影响这里描述的音频处理动作的性能的情况下省略人听觉系统的模型中的基膜的扩展效果。

[0166] 根据这里描述的技术,可在各种实施例中使用不同数量的频带(例如,20个频带、40个感应带等)。另外、任选地或者作为替代方案,也可在各种实施例中使用不同的带宽宽度。

[0167] 13.用于信道的各单个子集的各单个增益

[0168] 在一些实施例中,当特定扬声器配置是多信道配置时,可通过在从物理域(或者在

代表物理域的数字域中)向响度域的转换之前首先加算所有信道的激励信号获得总响度水平。但是,向特定扬声器配置中的所有信道施加相同的增益不能保持特定扬声器配置的不同信道(例如,关于不同信道之间的相对响度水平等)之间的空间平衡。

[0169] 在一些实施例中,为了保持空间平衡使得可最佳地或者正确地保持不同信道之间的相对感觉响度水平,可对每个信道确定或计算各响度水平和基于各响度水平获得的相应增益。在一些实施例中,基于各响度水平获得的相应的增益不等于相同的总增益;例如,相应增益中的一些或全部中的每一个可等于总增益加上(例如,信道特定)小校正。

[0170] 在一些实施例中,为了保持空间平衡,可对信道的每个子集确定或计算各响度水平和基于各响度水平获得的相应增益。在一些实施例中,基于各响度水平获得的相应的增益不等于相同的总增益;例如,相应增益中的一些或全部中的每一个可等于总增益加上(例如,信道特定)小校正。在一些实施例中,信道的子集可包含形成特定扬声器配置中的所有信道的适当子集的两个或更多个信道(例如,包含左前、右前和低频效果(LFE)的信道的子集;包含左环绕和右环绕的信道的子集,等等)。信道的子集的音频内容可构成在编码音频信号(102)中承载的总混合的子混合。可向子混合内的信道施加相同的增益。

[0171] 在一些实施例中,为了从特定扬声器配置生成实际的响度(例如,实际感觉等),可以使用一个或更多个校准参数以关联数字域中的信号水平与由数字域代表的物理域中的相应物理(例如,关于$dB_{SPL}$的空间压力等)水平。一个或更多个校准参数可被赋予特定扬声器配置中的物理声音设备特有的值。

[0172] 14.听觉场景分析

[0173] 在一些实施例中,这里描述的编码器可实现基于计算机的听觉场景分析(ASA)以检测音频内容(例如,编码成编码音频信号102等)中的听觉事件边界、生成一个或更多个ASA参数、将一个或更多个ASA参数格式化为编码音频信号(例如,102等)的要传输到下游装置(例如,解码器100等)的部分。ASA参数可包含但不仅仅限于指示听觉事件边界的位置的参数、听觉事件确信度测量(后面进一步解释)等中的任一个。

[0174] 在一些实施例中,听觉事件边界的(例如,时间上的)位置可在在编码的音频信号(102)内编码的元数据中被指示。另外、任选地或者作为替代方案,听觉事件边界的(例如,时间上的)位置可在检测听觉事件边界的位置的音频数据块和/或帧中被指示(例如,用标记、数据栏等)。

[0175] 这里,听觉事件边界指的是前面的听觉事件结束和/或后续的听觉事件开始的点。各听觉事件在两个连续的听觉事件边界之间出现。

[0176] 在一些实施例中,编码器(150)被配置为通过两个(例如,时间上)连续音频数据帧之间的特定响度频谱中的差值检测听觉事件边界。特定响度频谱中的每一个可包含从连续音频数据帧的相应音频数据帧计算的不平滑化响度的频谱。

[0177] 在一些实施例中,特定响度频谱N[b,t]可被归一化以获得下式所

[0178] $$N_{NORM}[b,t] = \frac{N[b,t]}{\max_b\{N[b,t]\}} \quad (1)$$

[0179] 这里,b表示频带,t表示时间或音频数据帧指数,$\max_b\{N[b,t]\}$是跨着所有频带的最大特定响度水平。

$$A[t]=\begin{cases}0 & D[t]\le D_{min}\\ \frac{D[t]-D_{min}}{D_{max}-D_{min}} & D_{min}<D[t]<D_{max}\\ 1 & D[t]\ge D_{max}\end{cases} \quad (2)$$

[0180] (3)

[0181] 这里,$D_{min}$和$D_{max}$是最小和最大阈值(例如,用户可配置、系统可配置、关于音频内容中的D[t]的过去值分布设定,等等)。

[0182] 在一些实施例中,编码器(150)被配置为检测D[t](例如,在特定的t处等)高于$D_{min}$时的听觉事件边界(例如,特定t等)。

[0183] 在一些实施例中,这里描述的解码器(例如,100等)从编码的音频信号(例如,102等)提取ASA参数,并且使用ASA参数以防止导致听觉事件的感觉畸变的柔声的无意的升压和/或大声的无意切割。

[0184] 解码器(100)可被配置为通过确保在听觉事件内增益更接近恒定并且通过将许多的增益改变约束到听觉事件边界的附近来减少或防止听觉事件的无意畸变。例如,解码器(100)可被配置为响应听觉事件边界处的起音(例如,响度水平增加等)中的增益变化使用相对小的时间常数(例如,与听觉事件的最小持续期相当或者比其短等)。因此,可相对迅速地通过解码器(100)实现起音中的增益变化。另一方面,解码器(100)可被配置为响应听觉事件中的释放(例如,响度水平减小等)的增益变化相对于听觉事件的持续期使用相对长的时间常数。因此,可相对慢地通过解码器(100)实现释放的增益变化,使得应当表现恒定或者逐渐衰减的声音不可听或者在感觉上受到干扰。听觉事件边界处的起音中的迅速响应和听觉事件中的释放的慢速响应允许快速感觉到听觉事件的到达,并且保持听觉事件中的感觉质量和/或完整性—包含通过特定响度水平关系和/或特定时间关系链接的大声和柔声—诸如钢琴弦等。

[0185] 在一些实施例中,解码器(100)使用由ASA参数指示的听觉事件和听觉事件边界以控制解码器(100)处的特定扬声器配置中的信道中的一个、两个、一些或全部的增益变化。

[0186] 15.响度水平过渡

[0187] 例如,在两个节目之间、在节目与大声商业广告之间等,可出现响度水平过渡。在一些实施例中,解码器(100)被配置为基于过去的音频内容(例如,从编码音频信号102接收,过去4秒等)保持瞬时响度水平的直方图。在从响度水平过渡之前到响度水平过渡之后的时间间隔上,具有变高的概率的两个区域可在直方图中被记录。区域中的一个以前面响度水平周围为中心,而区域中的另一个以新的响度水平周围为中心。

[0188] 解码器(100)可将平滑化的响度水平动态地确定为正在处理的音频内容,并且基于平滑化的响度水平确定直方图的相应箱体(例如,包含与平滑化的响度水平相同的值的瞬时响度水平的箱体等)。解码器(100)进一步被配置为比较相应箱体处的概率与阈值(例如,6%、7%、7.5%等),这里,直方图曲线的总区域(例如,所有箱体的和)代表100%的概率。解码器可被配置为通过确定相应箱体处的概率低于阈值检测响度水平过渡的出现。作为响应,解码器(100)被配置为选择相对小的时间常数以相对快速地适于新的响度水平。因此,可减少响度水平过渡内的大声(或柔声)开始的时间持续期。

[0189] 在一些实施例中,解码器(100)使用安静/噪声门,以防止低瞬时响度水平进入到直方图中并且变为直方图中的高概率箱体。另外、任选地或者作为替代方案,解码器(100)可被配置为使用ASA参数以检测包含于直方图中的听觉事件。在一些实施例中,解码器(100)可从ASA参数确定时间平均听觉事件确信度测量$\overline{A}[t]$的时间依赖值。在一些实施例中,解码器(100)从ASA参数确定(例如,瞬时等)听觉事件确信度测量A[t]的时间依赖值,并且从ASA参数等基于(例如,瞬时等)听觉事件确信度测量A[t]的时间依赖值计算时间平均听觉事件确信度测量$\overline{A}[t]$的值。解码器(100)可被配置为如果与响度水平同时的时间平均听觉事件确信度测量$\overline{A}[t]$低于直方图包含阈值(例如,0.1、0.12等)则排除响度水平进入直方图。

[0190] 在一些实施例中,对允许包含于直方图中的(例如,瞬时等)响度水平(例如,相应的$\overline{A}[t]$高于直方图包含阈值等),响度水平被分配与与响度水平同时的(contemporanneous)时间平均听觉事件确信度测量$\overline{A}[t]$的时间依赖值相同、与其成比例等的权重。作为结果,与不接近听觉事件边界的其它响度水平相比,听觉事件边界附近的响度水平对直方图具有更大的影响(例如,$\overline{A}[t]$具有相对大的值等)。

[0191] 16.复位

[0192] 在一些实施例中,这里描述的编码器(例如,150等)被配置为检测复位事件且在由编码器(150)生成的编码音频信号(例如,102等)中包含复位事件的指示。在第一例子中,编码器(150)响应于确定出现连续(例如,250毫秒,可通过系统和/或用户配置等)的相对安静周期检测复位事件。在第二例子中,编码器(150)响应于确定跨着所有频带在激励水平中出现大的瞬时下降来检测复位事件。在第三例子中,编码器被提供出现需要复位的内容(例如,节目开始/结束、场景改变等)中的过渡的输入(例如,元数据、用户输入、系统控制等)。

[0193] 在一些实施例中,这里描述的解码器(例如,100等)实现可用于瞬时加速增益平滑化的复位机构。复位机构是有用的,并且可在出现信道或视听输入之间的切换时被调用。

[0194] 在一些实施例中,解码器(100)可被配置为通过确定是否出现连续(例如,250毫秒,可通过系统和/或用户配置等)的相对安静周期、是否出现跨着所有频带的激励水平的大的瞬时下降等确定是否出现复位事件。

[0195] 在一些实施例中,解码器(100)可被配置为响应接收通过上游编码器(例如,150等)在编码音频信号(102)中提供的指示(例如,复位事件的指示等)确定出现复位事件。

[0196] 可使复位机构在解码器(100)确定出现复位事件时发出复位。在一些实施例中,复位机构被配置为利用DRC压缩曲线的略微更激进的切割行为以防止(例如,大声的节目/信道/视听源等的)硬开始。另外、任选地或者作为替代方案,解码器(100)可被配置为实现保护措施以在解码器(100)检测到复位被错误地触发时温和地恢复。

[0197] 17.编码器提供的增益

[0198] 在一些实施例中,音频编码器可被配置为计算编码成编码音频信号的音频内容的各单个部分(例如,音频数据块、音频数据帧等)的一组或更多组增益(DRC增益等)。由音频编码器生成的多组增益可包含以下方面中的一个或更多个:包含所有信道(左前、右前、低频效果或LFE、中心、左环绕、右环绕等)的单个广带(或宽带)增益的第一组增益;包含信道的各单个子集的各单个广带(或宽带)增益的第二组增益;包含信道的各单个子集和第一数

量(例如,两个等)单个频带(例如,各信道中的两个频带等)中的每一个的各单个广带(或宽带)增益的第三组增益;包含信道的各单个子集和第二数量(例如,四个等)单个频带(例如,各信道中的四个频带等)中的每一个的各单个广带(或宽带)增益的第四组增益;等等。这里描述的信道的子集可以是包含左前、右前和LFE信道的子集、包含中心信道的子集、包含左环绕和右环绕信道的子集等中的一个或更多个。

[0199] 在一些实施例中,音频编码器被配置为以时间同步方式传送音频内容的一个或更多个部分(例如,音频数据块、音频数据帧等)和对音频内容的一个或更多个部分计算的一个或更多个组的增益。接收音频内容的一个或更多个部分的音频解码器可以以很少的延迟或者在没有延迟的情况下选择和施加一组或更多个组的增益中的一组增益。在一些实施例中,音频编码器可实现在图4所示的一个或更多个子帧中承载(例如,通过差分编码等)一组或更多组增益的子帧技术。在例子中,可在计算增益的音频数据块或音频数据帧内编码子帧。在另一例子中,可在计算增益的音频数据块或音频数据帧前面的音频数据块或音频数据帧内编码子帧。在另一非限制性例子中,可在从计算增益的音频数据块或音频数据帧在一定的时间内的在音频数据块或音频数据帧内编码子帧。在一些实施例中,可以使用Huffman和差分编码,以占据和/或压缩承载多组增益的子帧。

[0200] 18.示例性系统和处理流程

[0201] 图5示出非限制性示例性实施例中的示例性编解码器系统。可以是诸如150等的音频编码器中的处理单元的内容创建器被配置为向编码器单元(“NGC Encoder”)提供音频内容(“Audio”)。编码器单元将音频内容格式化成音频数据块和/或帧,并且将音频数据块和/或帧编码成编码的音频信号。内容创建器还被配置为在音频内容和一个或更多个动态范围压缩曲线标识符(“Compression curve IDs”)中建立/生成一个或更多个节目、商业广告等的一个或更多个对话响度水平(“dialnorm”)。内容创建器可从音频内容中的一个或更多个对话音频轨道确定对话响度水平。可至少部分地基于用户输入、系统配置参数等选择动态范围压缩曲线标识符。内容创建器可以是利用工具以生成音频内容和dialnorm的人(艺术家、音频工程师等)。

[0202] 基于动态范围压缩曲线标识符,编码器(150)生成包含但不限于由一个或更多个动态范围压缩曲线支持的多个回放环境的相应基准对话响度水平(“Reference levels”)的一个或更多个DRC参数组。在编码的音频信号的元数据中,这些DRC参数组在编码音频信号的元数据中可通过音频内容被带内编码、通过音频内容被带外编码等。作为生成可传输到诸如100等的音频解码器的编码音频信号的一部分,可以执行诸如压缩、格式化多路复用(“MUX”)等的动作。可用支持音频数据要素、DRC参数组、基准响度水平、动态范围压缩曲线、函数、查找表、在压缩中使用的Huffman代码、子帧等的输送的句法编码编码的音频信号。在一些实施例中,在一些实施例中,句法允许上游装置(例如,编码器、解码器、变码器等)将增益传送到下游装置(例如,解码器、变码器等)。在一些实施例中,用于将数据编码成编码的音频信号和/或从其解码数据的句法被配置为支持向后的兼容,使得依赖于由上游装置计算的增益的装置可任选地继续这样做。

[0203] 在一些实施例中,编码器(150)计算音频内容的一组、两组或更多组增益(例如,DRC增益、增益平滑化,通过适当的基准对话响度水平等)。在以音频内容编码成编码音频信号的元数据中,多组增益可具有一个或更多个动态范围压缩曲线。第一组增益可与(例如,

缺省等)扬声器配置或简档中的所有信道的广带(或宽带)增益对应。第二组增益可与扬声器配置或简档中的所有信道中的每一个的广带(或宽带)增益对应。第三组增益可与扬声器配置或简档中的所有信道中的每一个的两个频带中的每一个的广带(或宽带)增益对应。第四组增益可与扬声器配置或简档中的所有信道中的每一个中的四个频带中的每一个的广带(或宽带)增益对应。在一些实施例中,可用元数据中的扬声器配置的(例如,参数化等)动态范围压缩曲线传送对扬声器配置计算的多组增益。在一些实施例中,对扬声器配置计算的多组增益可替代元数据中的扬声器配置的(例如,参数化等)动态范围压缩曲线。可根据这里描述的技术支持附加的扬声器配置或简档。

[0204] 解码器(100)被配置为例如通过诸如解压缩、去格式化、去多路复用("DEMUX")等的动作从编码的音频信号提取音频数据块和/或帧和元数据。提取的音频数据块和/或帧可通过解码器单元("NGC Decoder")解码成音频数据要素或采样。解码器(100)进一步被配置为确定要呈现音频内容的解码器(100)处的特定回放环境的简档,并且从从编码音频信号提取的元数据选择动态范围压缩曲线。数字音频处理单元("DAP")被配置为出于生成在特定回放环境中驱动音频信道的音频信号的目的在音频数据要素或采样上施加DRC和其它动作。解码器(100)可基于从音频数据块或帧确定的响度水平和选择的动态范围压缩曲线计算和施加DRC增益。解码器(100)也可基于与选择的动态范围压缩曲线相关的基准对话响度水平和从编码音频信号提取的元数据中的对话响度水平调整输出对话响度水平。解码器(100)可随后施加与音频内容和特定回放环境相关的回放情形特有的增益限制器。因此,解码器(100)可呈现/播放针对回放情形定制的音频内容。

[0205] 图5A示出另一示例性解码器(可以与图5的解码器100相同)。如图5A所示,图5A的解码器被配置为例如通过诸如解压缩、去格式化、去多路复用("DEMUX")等的动作从编码的音频信号提取音频数据块和/或帧和元数据。提取的音频数据块和/或帧可通过解码器单元("Decoder")解码成音频数据要素或采样。图5A的解码器进一步被配置为对一组缺省增益基于缺省压缩曲线、与缺省压缩曲线有关的平滑化常数等执行DRC增益计算。图5A的解码器进一步被配置为从元数据中的简档相关元数据提取非缺省增益简档的一组差分增益、确定要呈现音频内容的图5A的解码器处的非缺省增益简档的一组非差分增益、出于生成在特定回放环境中驱动音频信道的DRC增强音频输出的目的在音频数据要素或采样上施加一组非差分增益和其它动作。因此,即使图5A的解码器自身可能或者可能没有实现用于执行一组增益生成动作以直接对非缺省增益简档获得一组非差分增益的支持,图5A的解码器也可根据非缺省增益简档呈现/播放音频内容。

[0206] 图6A～6D示出示例性处理流程。在一些实施例中,媒体处理系统中的一个或更多个计算装置或单元可执行该处理流程。

[0207] 图6A示出可通过这里描述的音频解码器实现的示例性处理流程。在图6A的块602中,第一装置(例如,图1A的音频解码器100等)接收包含音频内容和一个或更多个动态范围压缩曲线的定义数据的音频信号。

[0208] 在块604中,第一装置确定特定回放环境。

[0209] 在块606中,第一装置基于从音频信号提取的一个或更多个动态范围压缩曲线的定义数据建立特定回放环境的特定动态范围压缩曲线。

[0210] 在块608中,第一装置在从音频信号提取的音频内容的一个或更多个部分上执行

一个或更多个动态范围控制(DRC)动作。一个或更多个DRC动作至少部分地基于从特定动态范围压缩曲线获得的一个或更多个DRC增益。

[0211] 在实施例中,一个或更多个动态范围压缩曲线的定义数据包含起音时间、释放时间或与一个或更多个动态范围压缩曲线中的至少一个相关的基准响度水平中的一个或更多个。

[0212] 在实施例中,第一装置进一步被配置为执行以下过程:计算音频内容的一个或更多个部分的一个或更多个响度水平;基于特定动态范围压缩曲线和音频内容的一个或更多个部分的一个或更多个响度水平确定一个或更多个DRC增益;等等。

[0213] 在实施例中,对音频内容的一个或更多个部分计算的响度水平中的至少一个是与一个或更多个频带相关的特定响度水平、跨着广带范围的广带响度水平、跨着宽带范围的宽带响度水平、跨着多个频带的广带响度水平、跨着多个频带的宽带响度水平等中的一个或更多个。

[0214] 在实施例中,对音频内容的一个或更多个部分计算的响度水平中的至少一个是瞬时响度水平或在一个或更多个时间间隔上平滑化的响度水平中的一个或更多个。

[0215] 在实施例中,一个或更多个动作包括与调整对话响度水平、增益平滑化、增益限制、动态均衡化、噪声补偿等中的一个或更多个相关的一个或更多个动作。

[0216] 在实施例中,第一装置进一步被配置为执行以下过程:从编码的音频信号提取一个或更多个对话响度水平;将一个或更多个对话响度水平调整到一个或更多个基准对话响度水平,等等。

[0217] 在实施例中,第一装置进一步被配置为执行以下过程:从编码音频信号提取一个或更多个听觉场景分析(ASA)参数;改变在平滑化施加到音频信号的增益时使用的一个或更多个时间常数,增益与一个或更多个DRC增益、增益平滑化或增益限制等中的一个或更多个有关。

[0218] 在实施例中,第一装置进一步被配置为执行以下过程:基于复位事件的指示确定复位事件在音频内容的一个或更多个部分中出现,该复位的指示是从编码的音频信号提取的;和响应于确定复位事件在音频内容的一个或更多个部分中出现,在在确定复位事件在音频内容的一个或更多个部分中出现时执行的一个或更多个增益平滑化动作上采取一个或更多个行动。

[0219] 在实施例中,第一装置进一步被配置为执行以下过程:保持瞬时响度水平的直方图,该直方图被从音频内容中的时间间隔计算的瞬时响度水平占据;确定特定响度水平是否在直方图的高概率区域中高于阈值,特定响度水平是从音频内容的一部分计算的;和响应于确定特定响度水平在直方图的高概率区域中高于阈值,执行以下过程:确定出现了响度过渡;缩短在增益平滑化中使用的时间常数以加速响度过渡。

[0220] 图6B示出可通过这里描述的音频编码器实现的示例性处理流程。在图6B的块652中,第二装置(例如,图1B的音频编码器150等)接收源音频格式的音频内容。

[0221] 在块654中,第二装置检索一个或更多个动态范围压缩曲线的定义数据。

[0222] 在块656中,第二装置生成包含音频内容和一个或更多个动态范围压缩曲线的定义数据的音频信号。

[0223] 在实施例中,第二装置进一步被配置为执行以下过程:确定一个或更多个动态范

围压缩曲线的一个或更多个标识符;和基于一个或更多个标识符从基准数据库检索一个或更多个动态范围压缩曲线的定义数据,等等。

[0224] 在实施例中,第二装置进一步被配置为执行以下过程:对音频内容的一个或更多个部分计算一个或更多个对话响度水平;用音频内容的一个或更多个部分将一个或更多个对话响度水平编码到编码的音频信号中,等等。

[0225] 在实施例中,第二装置进一步被配置为执行以下过程:在音频内容的一个或更多个部分上执行听觉场景分析(ASA);基于音频内容的一个或更多个部分上的ASA的结果生成一个或更多个ASA参数;用音频内容的一个或更多个部分将一个或更多个ASA参数编码到编码的音频信号中,等等。

[0226] 在实施例中,第二装置进一步被配置为执行以下过程:确定一个或更多个复位事件在音频内容的一个或更多个部分中出现;和用音频内容的一个或更多个部分将一个或更多个复位事件的一个或更多个指示编码到编码的音频信号中,等等。

[0227] 在实施例中,第二装置进一步被配置为将音频内容的一个或更多个部分编码到音频数据帧或音频数据块中的一个或更多个中。

[0228] 在实施例中,一个或更多个DRC增益中的第一DRC增益向与特定回放环境对应的特定扬声器配置中的一组所有信道中的第一适当子集中的各信道施加;而一个或更多个DRC增益中的第二不同DRC增益向与特定回放环境对应的特定扬声器配置中的一组所有信道中的第二适当子集中的各信道施加。

[0229] 在实施例中,一个或更多个DRC增益中的第一DRC增益向第一频带施加,而一个或更多个DRC增益中的第二不同DRC增益向第二不同的频带施加。

[0230] 在实施例中,音频内容的一个或更多个部分包含音频数据帧或音频数据块中的一个或更多个。在实施例中,编码音频信号是视听信号的一部分。

[0231] 在实施例中,一个或更多个DRC增益在响度域中被限定。

[0232] 图6C示出可通过这里描述的音频解码器实现的示例性处理流程。在图6C的块662中,第三装置(例如,图1A的音频解码器100、图5的音频解码器、图5A的音频解码器等)接收包含音频内容以及一个或更多个动态范围压缩曲线(DRC)曲线和一组或更多组差分增益的定义数据的音频信号。

[0233] 在块664中,第三装置对特定回放环境中的增益简档识别一组或更多组差分增益之中的特定组的差分增益;第三装置还识别一个或更多个DRC曲线中的与特定组的差分增益相关的缺省DRC曲线。

[0234] 在块666中,第三装置至少部分地基于缺省DRC曲线生成一组缺省增益。

[0235] 在块668中,至少部分地基于一组缺省增益与特定组的差分增益的组合,第三装置在从音频信号提取的音频内容的一个或更多个部分上执行一个或更多个动作。

[0236] 在实施例中,一组缺省增益包含通过至少部分地基于缺省DRC曲线执行一组增益生成动作生成的非差分增益。

[0237] 在实施例中,缺省DRC曲线代表缺省增益简档。在实施例中,关于缺省DRC曲线的特定组的差分增益代表非缺省增益简档。在实施例中,音频信号不包含与非缺省增益简档对应的非缺省DRC曲线的定义数据。

[0238] 在实施例中,特定组的差分增益包含对非缺省增益简档生成的一组非差分非缺省

增益和对由缺省DRC曲线代表的缺省增益简档生成的一组非差分缺省增益之间的增益差。一组非差分非缺省增益与一组非差分缺省增益由编码音频信号的上游音频解码器生成。

[0239] 在实施例中,一组非差分非缺省增益或一组非差分缺省增益中的至少一组不作为音频信号的一部分被提供。

[0240] 图6D示出可通过这里描述的音频解码器实现的示例性处理流程。在图6D的块672中,第四装置(例如,图1A的音频编码器150、图5的音频编码器等)接收源音频格式的音频内容;

[0241] 在块674中,第四装置至少部分地基于代表缺省增益简档的缺省动态范围压缩(DRC)曲线生成一组缺省增益。

[0242] 在块676中,第四装置生成非缺省增益简档的一组非缺省增益。

[0243] 在块678中,至少部分地基于一组缺省增益和一组非缺省增益,第四装置生成一组差分增益,该组差分增益代表关于缺省DRC曲线的非缺省增益简档。

[0244] 在块680中,第四装置生成包含音频内容以及一个或更多个DRC曲线和一组或更多组差分增益的定义数据的音频信号,一个或更多个DRC曲线包含缺省DRC曲线,一组或更多组差分增益包含一组差分增益。

[0245] 在一些实施例中,非缺省增益简档由DRC曲线代表。在实施例中,音频信号不包含代表非缺省增益简档的DRC曲线的定义数据。在一些实施例中,非缺省增益简档不由DRC曲线代表。

[0246] 在实施例中,提供包含处理器且被配置为执行这里描述的方法中的任一种的装置。

[0247] 在实施例中,提供包含软件指令的非暂时性计算机可读存储介质,这些软件指令当通过一个或更多个处理器被执行时导致执行这里描述的方法中的任一种。注意,虽然这里讨论了各单独的实施例,但这里讨论的实施例和/或部分实施例的任意组合可被组合以形成其它的实施例。

[0248] 19.实现机构—硬件概要

[0249] 根据一个实施例,通过一个或更多个专用计算装置实现这里描述的技术。专用计算装置可被硬接线以执行所述技术,或者可包含被永久编程以执行技术的诸如一个或更多个应用特定集成电路(ASIC)或场可编程门阵列(FPGA)的数字电子器件,或者可包含被编程以根据固件、内存、其它存储器或组合中的程序指令执行所述技术的一个或更多个通用硬件处理器。这种专用计算装置还可组合定制硬接线逻辑、ASIC或FPGA与定制编程以实现所述技术。专用计算装置可以是台式计算机系统、便携式计算机系统、手持设备、联网设备或加入硬接线和/或程序逻辑以实现所述技术的任何其它装置。

[0250] 例如,图7是示出可实现实施例的计算机系统700的框图。计算机系统700包含用于传送信息的总线702或其它通信机构和与总线702耦合的用于处理信息的硬件处理器704。硬件处理器704可以例如为通用微处理器。

[0251] 计算机系统700还包括与总线702耦合的用于存储信息和要通过处理器704执行的指令的主内存706,诸如随机存取内存(RAM)或其它动态存储装置。主内存706也可用于在执行要通过处理器704执行的指令时存储时间变量或其它中间信息。这种指令,当存储于处理器704可访问的非暂时性存储介质中时,将计算机系统700转变成被定制以执行在指令中规

定的动作的专用机器。

[0252] 计算机系统700还包括只读内存(ROM)708或与总线702耦合的用于存储静态信息和处理器704的指令的其它静止存储装置。用于存储信息和指令的存储装置710,诸如磁盘或光盘,被设置并且与总线702耦合。

[0253] 计算机系统700可通过总线702与用于向计算机用户显示信息的诸如阴极射线管(CRT)的显示器712耦合。用于向处理器704传送信息和命令选择的包含数字字母和其它键的输入装置714与总线702耦合。另一类型的用户输入装置是用于向处理器704传送方向信息和命令选择并且用于控制显示器712上的光标移动的光标控制716,诸如鼠标、跟踪球或光标方向键。该输入装置一般沿两个轴即第一轴(例如,x)和第二轴(例如,y)具有两个自由度,这两个自由度允许装置规定面中的位置。

[0254] 计算机系统700可通过使用与计算机系统组合以导致计算机系统700或者将其编程为专用机器的定制硬接线逻辑、一个或更多个ASIC或FPGA、固件和/或程序逻辑实现这里描述的技术。根据一个实施例,响应执行包含于主内存706中的一个或更多个指令的一个或更多个序列的处理器704,通过计算机系统700执行这里的技术。这种指令可从诸如存储装置710的另一存储介质被读入到主内存706中。执行包含于主内存706中的指令序列导致处理器704执行这里描述的处理步骤。在替代性实施例中,作为软件指令的替代,或者,与其组合,可以使用硬接线电路。

[0255] 这里使用的术语“存储介质”指的是存储导致机器以特定的方式动作的数据和/或指令的任何非暂时性介质。这种存储介质可包含非易失性介质和/或易失性介质。非易失性介质包含例如光盘或磁盘,诸如存储装置710。易失性介质包含动态内存,诸如主内存706。存储介质的常见形式包含例如软盘、柔性盘、硬盘、固态驱动、磁带或任何其它的磁数据存储介质、CD-ROM、任何其它光学数据存储介质、具有孔图案的任何物理介质、RAM、PROM和EPROM、FLASH-EPROM、NVRAM、任何其它的内存芯片或盒子。

[0256] 存储介质与传送介质明显不同但可与其组合使用。传送介质参与在存储介质之间传送信息。例如,传送介质包含共轴电缆、铜导线和光纤,包括包含总线702的导线。传送介质也可采取声波或光波的形式,诸如在无线电波和红外-红数据通信中生成的那些。

[0257] 各种形式的介质可参与向处理器704携带用于执行的一个或更多个指令的一个或更多个序列。例如,可首先在远程计算机的磁盘或固态驱动上承载指令。远程计算机可将指令加载到其动态内存中并且通过使用调制解调器在电话线上发送指令。计算机系统700本地的调制解调器可接收电话线上的数据并且使用红外-红传送器以将数据转换成红外-红信号。红外-红检测器可接收在红外-红信号中承载的数据并且,适当的电路可在总线702上放置数据。总线702将数据携带到主内存706,从该主内存706,处理器704检索并且执行指令。通过主内存706接收的指令可任选地在被处理器704执行之前或之后存储于存储装置710上。

[0258] 计算机系统700可包括与总线702耦合的通信接口718。通信接口718提供与与局域网络722连接的网络链接720的双向数据通信耦合。例如,通信接口718可以是集成服务数字网络(ISDN)卡、电缆调制解调器、卫星调制解调器或向相应类型的电话线提供数据通信连接的调制解调器。作为另一例子,通信接口718可以是向兼容LAN提供数据通信连接的局域网络(LAN)卡。也可实现无线链接。在任何这种实现中,通信接口718发送和接收承载代表各

种类型的信息的数字数据流的电气、电磁或光学信号。

[0259] 网络链接720一般通过一个或更多个网络向其它数据装置提供数据通信。例如,网络链接720可通过局域网络722向主机计算机724或向通过因特网服务提供商(ISP)726操作的数据设备提供连接。ISP 726又通过现在常称为“因特网”728的世界范围数据包数据通信网络提供数据通信服务。局域网络722和因特网728均使用承载数字数据流的电气、电磁或光学信号。关于计算机系统700携带数字数据的通过各种网络的信号和网络链接620上以及通过通信接口718的信号是传送介质的示例性形式。

[0260] 计算机系统700可通过网络、网络链接720和通信接口718发送消息并且接收数据,包含程序代码。在因特网的例子中,服务器730会通过因特网728、ISP 726、局域网络722和通信接口718传送对应用程序的请求代码。

[0261] 接收的代码可在其被接收时由处理器704执行,并且/或者,存储于存储装置710或其它的非易失性存储器中,以供以后的执行。

[0262] 20.等同物、扩展、替代和杂项

[0263] 在以上的说明书中,参照可在实现之间改变的大量的特定细节描述了本发明的实施例。因此,本发明的申请人意图的什么是本发明的专有和专用的指示是包含任何随后的校正的一组权利要求,这些权利要求以这些权利要求发出的特定的形式从本申请发出。这里对于包含于这些权利要求中的术语明确阐述的任何定义应掌控在权利要求中使用的这些术语的意思。由此,没有在权利要求中明确详述的限制、要素、性能、特征、优点或属性不应以任何的方式限制这些权利要求的范围。因此,说明书和附图应视为解释性而不是限制性的。

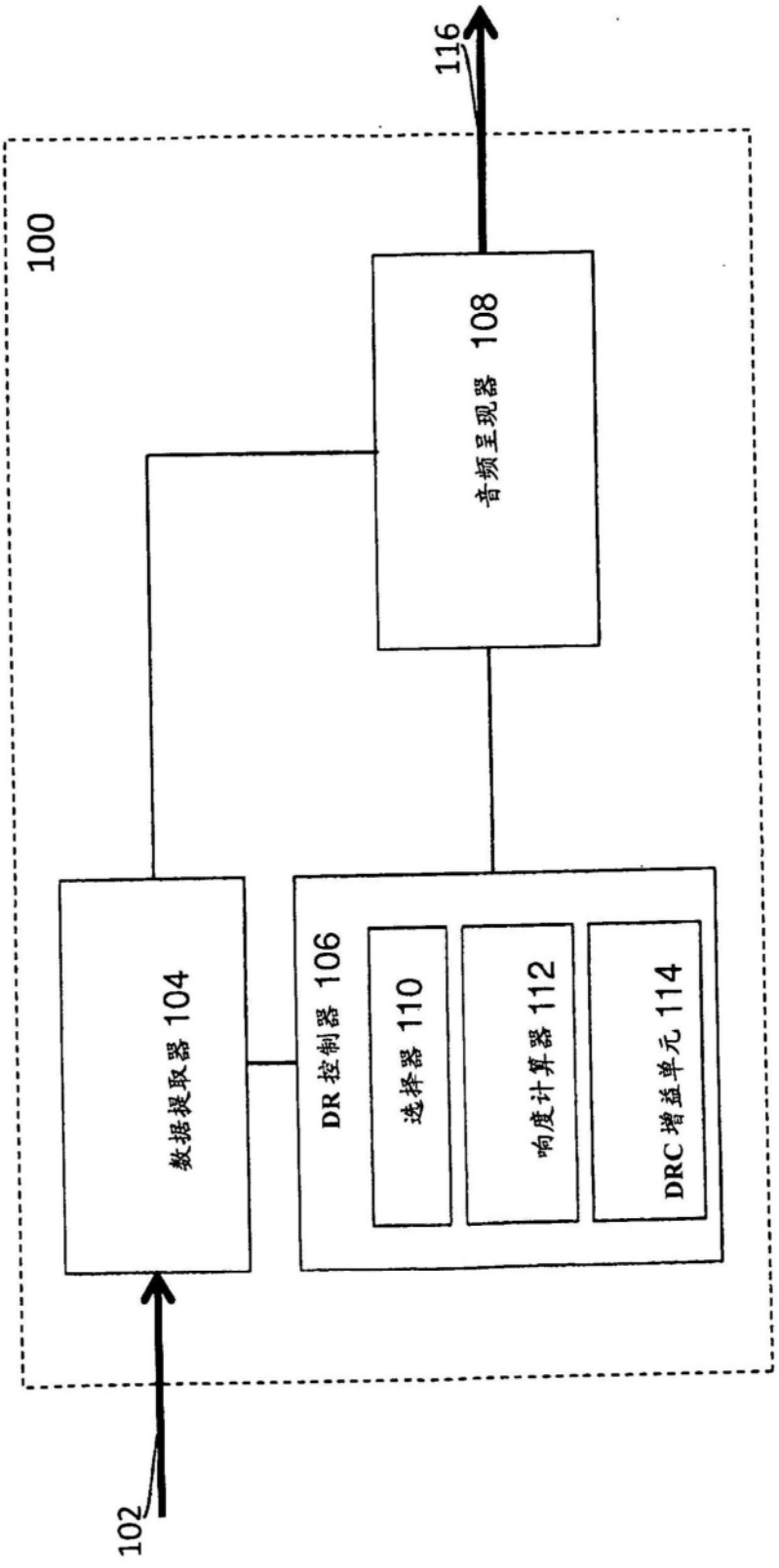

116
100
音频呈现器 108
数据提取器 104
DR 控制器 106
选择器 110
响度计算器 112
DRC 增益单元 114
102

图1A

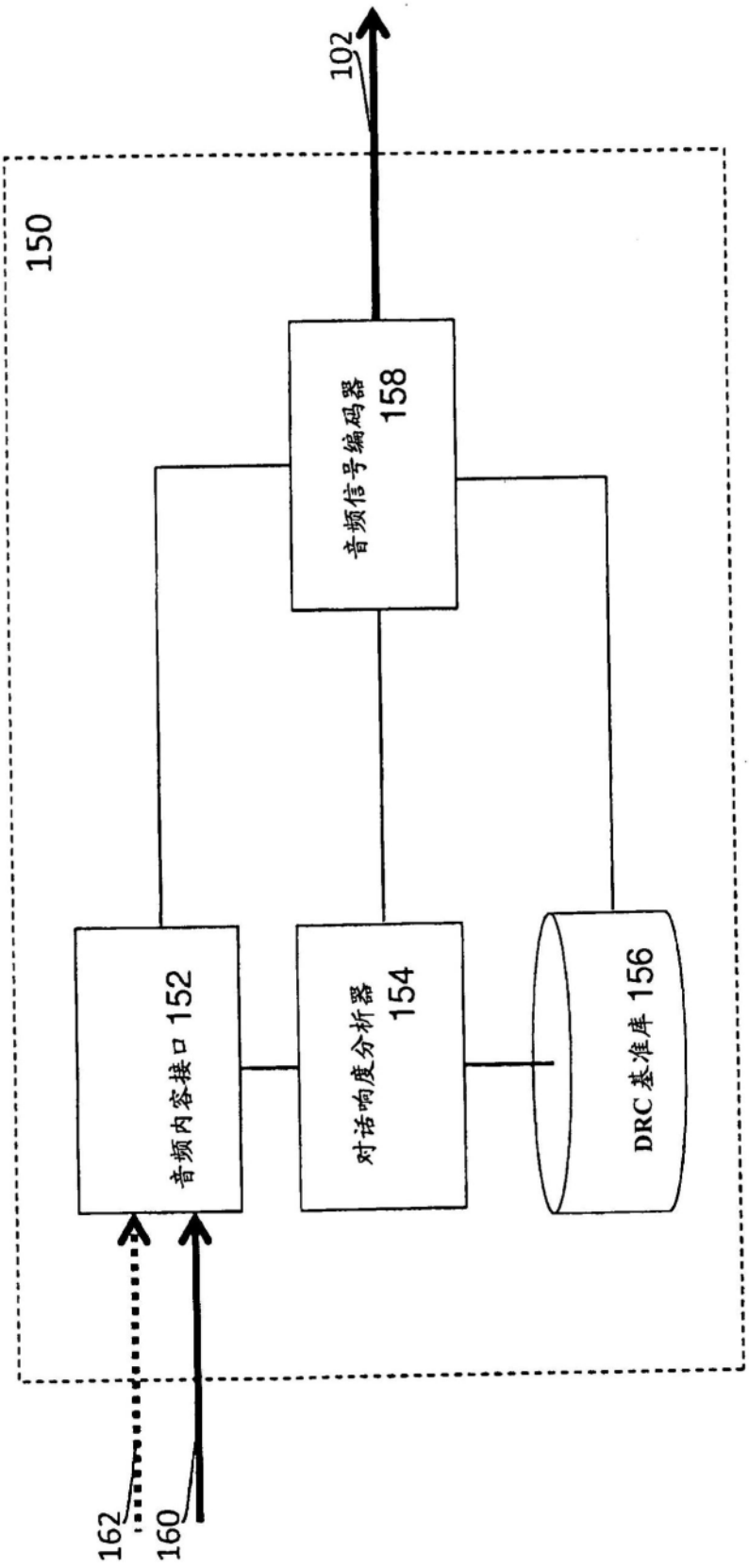

102
150
音频信号编码器 158
音频内容接口 152
对话响度分析器 154
DRC 基准库 156
162
160

图1B

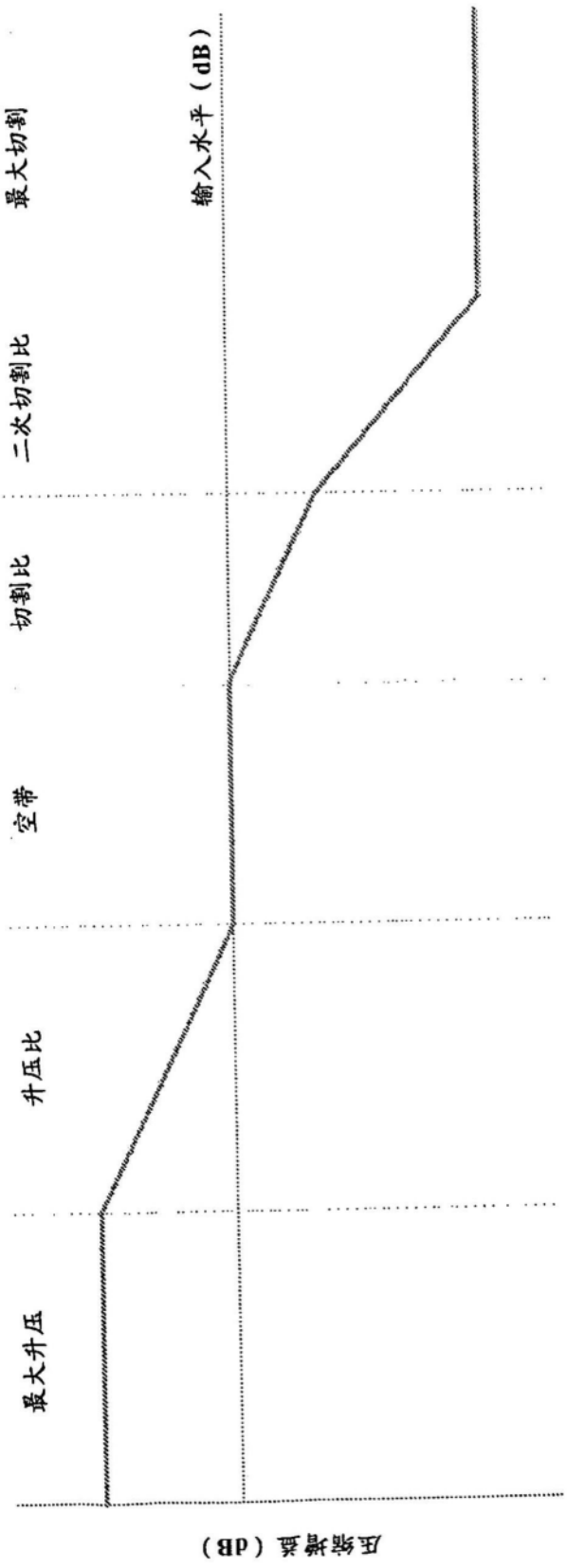



图2A

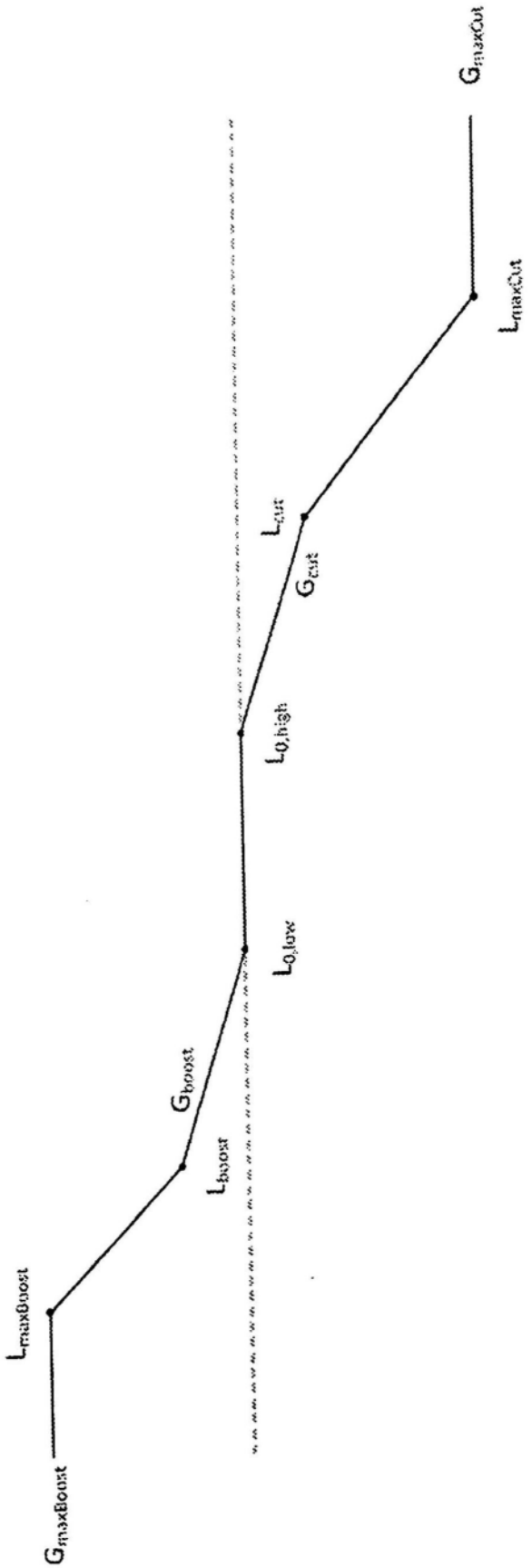

$G_{maxCut}$
$L_{maxCut}$
$L_{cut}$
$G_{cut}$
$L_{0,high}$
$L_{0,low}$
$G_{boost}$
$L_{boost}$
$L_{maxBoost}$
$G_{maxBoost}$

图2B

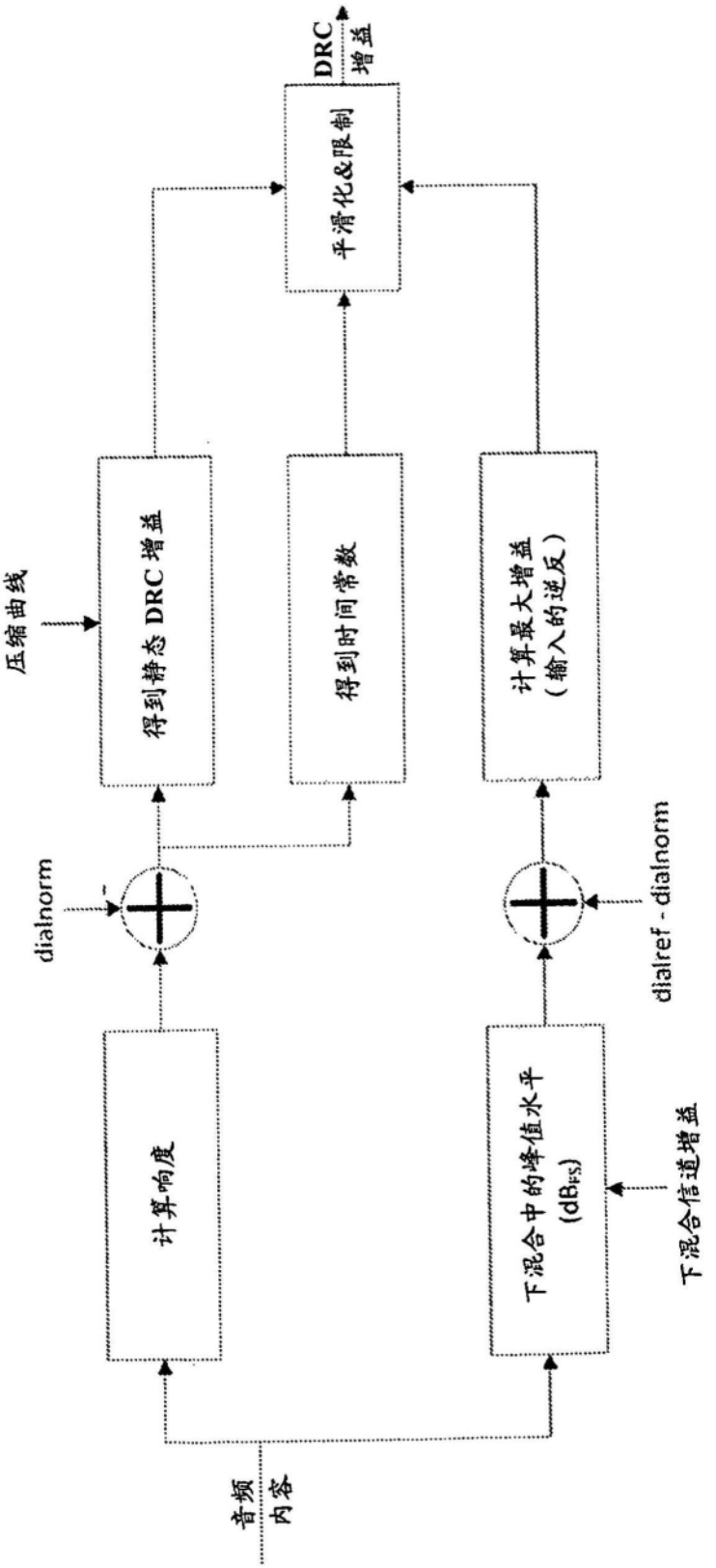

DRC
增益
平滑化&限制
压缩曲线
得到静态 DRC 增益
得到时间常数
计算最大增益
(输入的逆反)
dialnorm
dialref - dialnorm
计算响度
下混合中的峰值水平
(dBFS)
下混合信道增益
音频
内容

图3

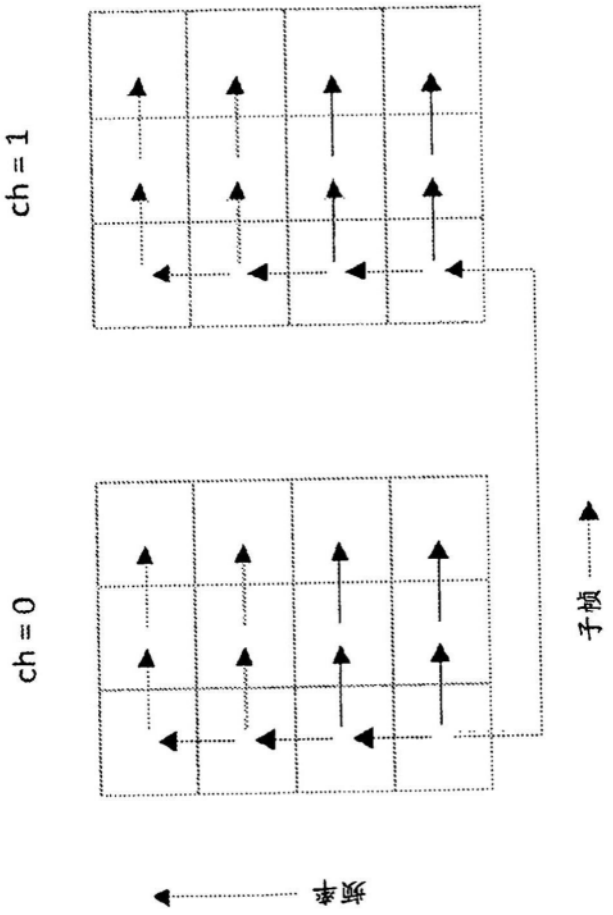



图4

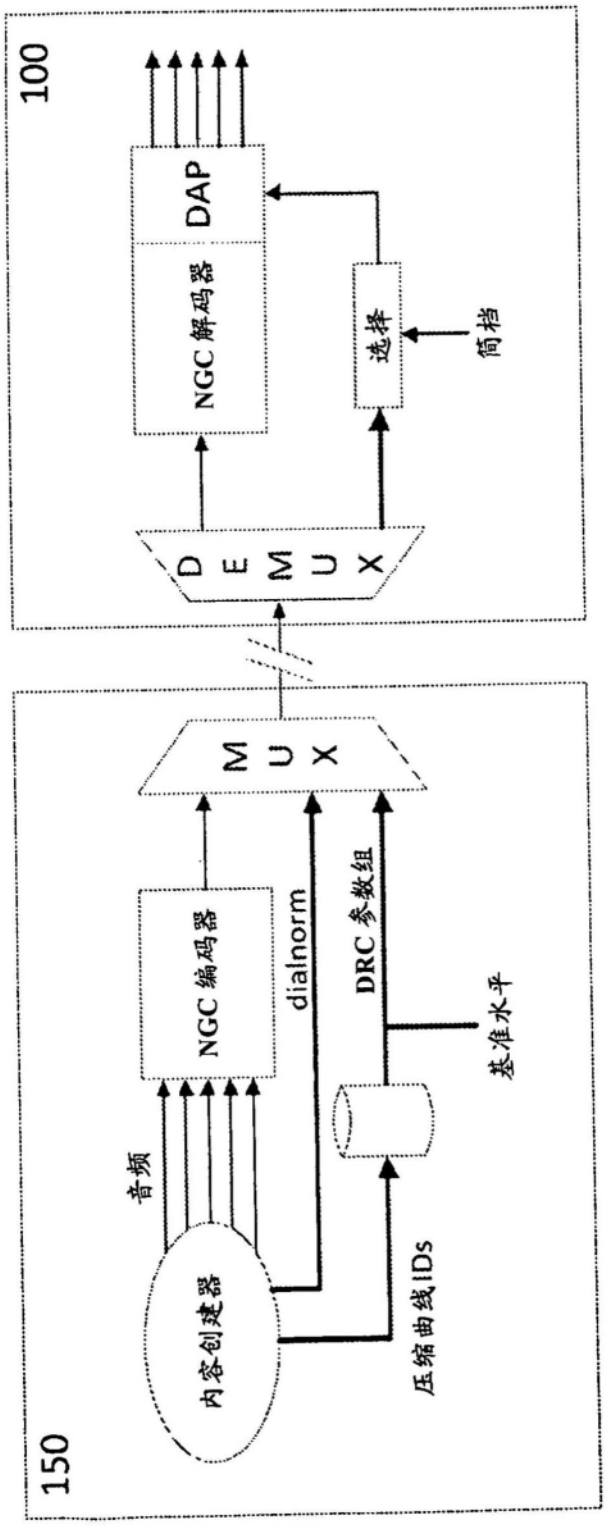

100
DAP
NGC 解码器
选择
简档
D E M U X
M U X
NGC 编码器
dialnorm
DRC 参数组
基准水平
音频
内容创建器
压缩曲线IDs
150

图5

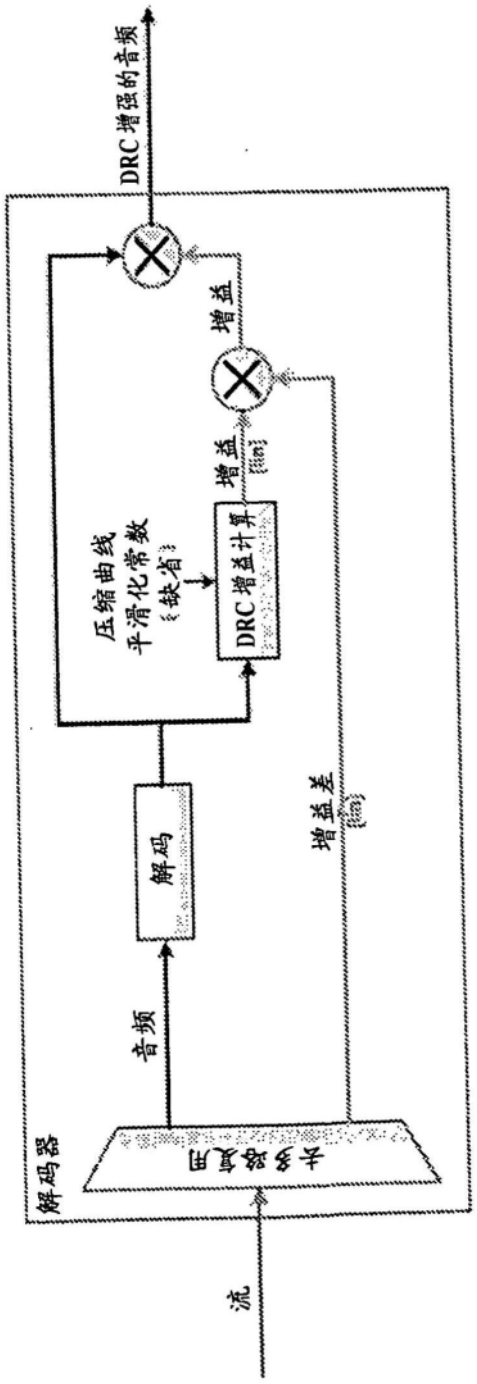



图5A

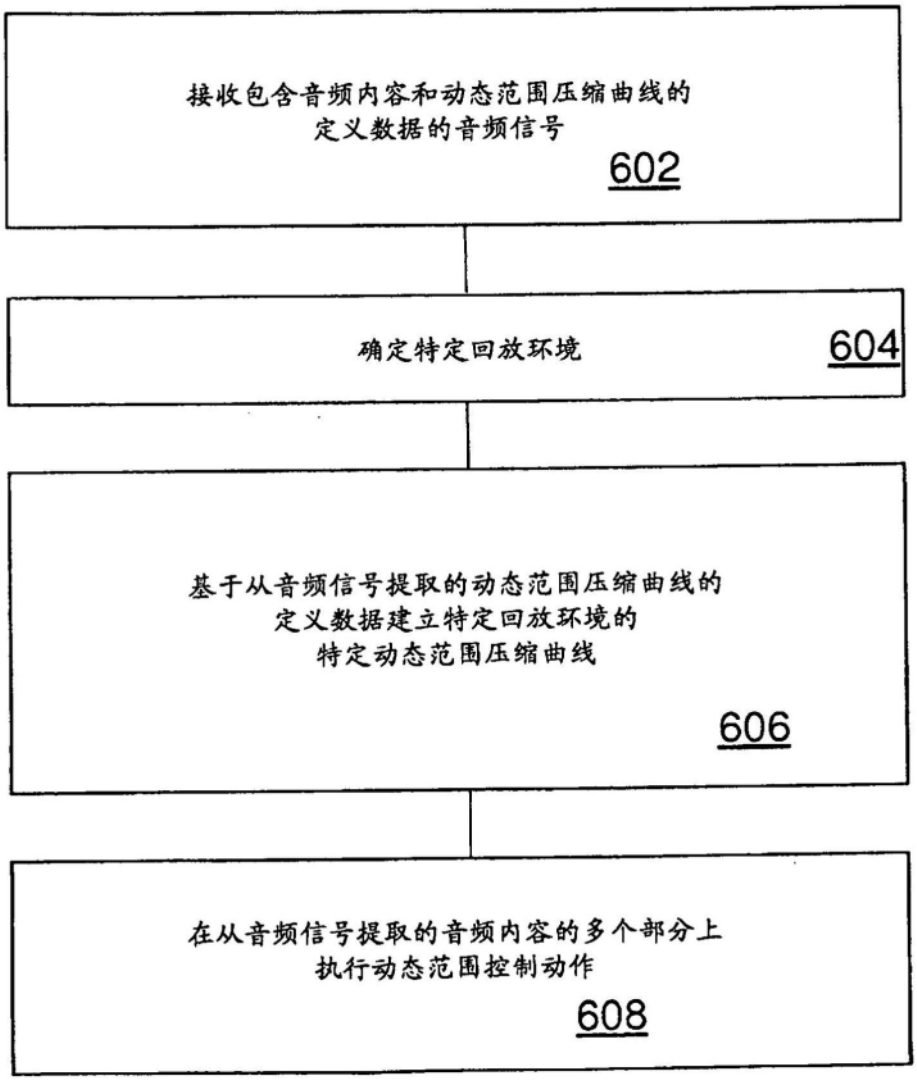

接收包含音频内容和动态范围压缩曲线的
定义数据的音频信号
602
确定特定回放环境
604
基于从音频信号提取的动态范围压缩曲线的
定义数据建立特定回放环境的
特定动态范围压缩曲线
606
在从音频信号提取的音频内容的多个部分上
执行动态范围控制动作
608

图6A

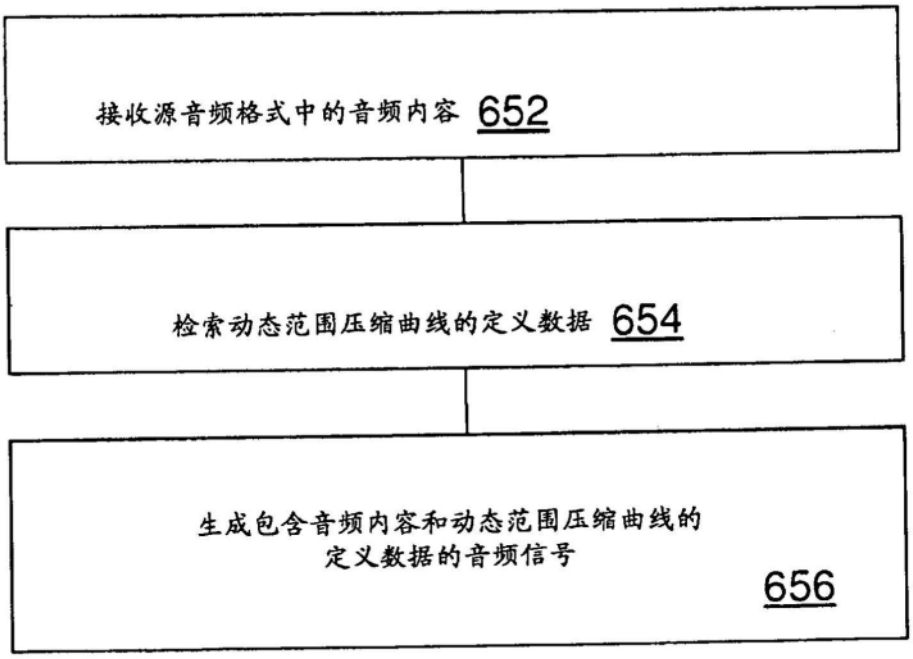

接收源音频格式中的音频内容 652
检索动态范围压缩曲线的定义数据 654
生成包含音频内容和动态范围压缩曲线的
定义数据的音频信号
656

图6B

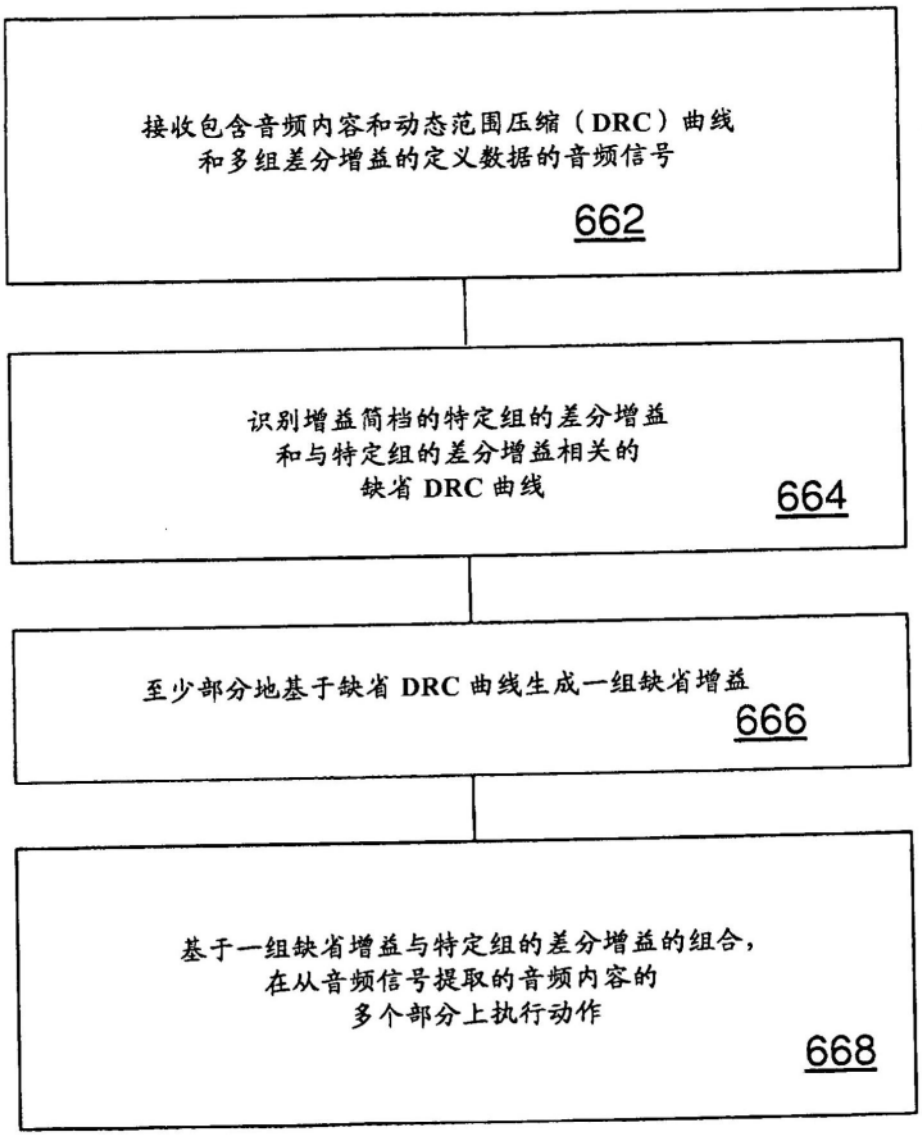

接收包含音频内容和动态范围压缩（DRC）曲线
和多组差分增益的定义数据的音频信号
662
识别增益简档的特定组的差分增益
和与特定组的差分增益相关的
缺省 DRC 曲线
664
至少部分地基于缺省 DRC 曲线生成一组缺省增益
666
基于一组缺省增益与特定组的差分增益的组合，
在从音频信号提取的音频内容的
多个部分上执行动作
668

图6C

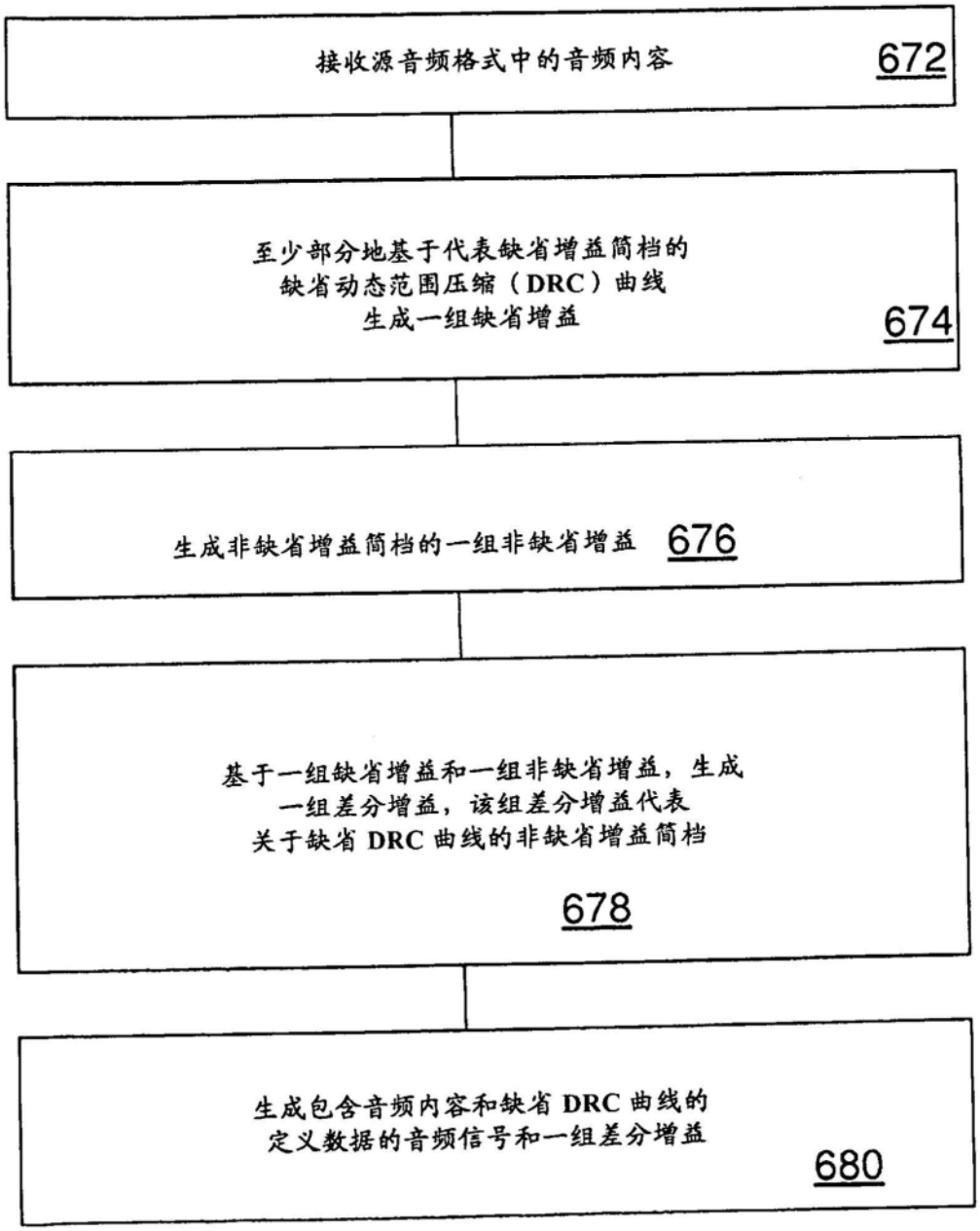

接收源音频格式中的音频内容
672
至少部分地基于代表缺省增益简档的
缺省动态范围压缩（DRC）曲线
生成一组缺省增益
674
生成非缺省增益简档的一组非缺省增益
676
基于一组缺省增益和一组非缺省增益，生成
一组差分增益，该组差分增益代表
关于缺省 DRC 曲线的非缺省增益简档
678
生成包含音频内容和缺省 DRC 曲线的
定义数据的音频信号和一组差分增益
680

图6D

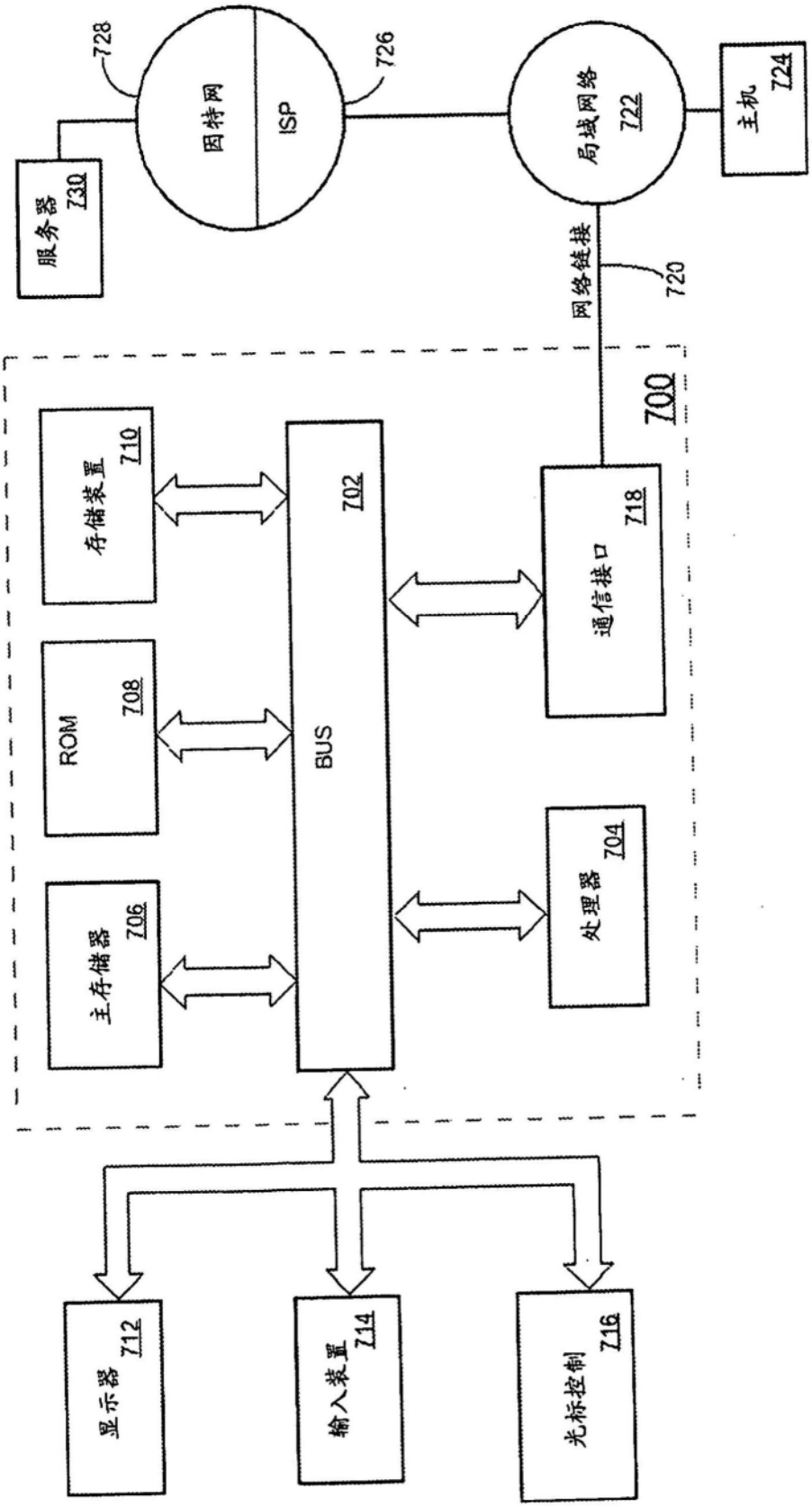

728
因特网
ISP
726
局域网络
722
主机
724
服务器
730
网络链接
720
700
存储装置
710
ROM
708
主存储器
706
BUS
702
通信接口
718
处理器
704
显示器
712
输入装置
714
光标控制
716

图7