



(12) 发明专利

(10) 授权公告号 CN 1758337 B

(45) 授权公告日 2010.12.08

(21) 申请号 200510109959.X

(22) 申请日 2002.07.10

(30) 优先权数据

0102481-9 2001.07.10 SE

0200796-1 2002.03.15 SE

0202159-0 2002.07.09 SE

(62) 分案原申请数据

02813646.2 2002.07.10

(73) 专利权人 杜比国际公司

地址 荷兰阿姆斯特丹

(72) 发明人 弗雷德里克·海恩

克里斯托弗·克约尔林

拉尔斯·古斯塔夫·里尔耶尔德

乔纳斯·罗丹 乔纳斯·英哥德加德

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 屠长存

(51) Int. Cl.

G10L 19/00(2006.01)

G10L 19/02(2006.01)

H04S 1/00(2006.01)

H04S 3/00(2006.01)

(56) 对比文件

CN 1329810 A, 全文.

US 4523309 A, 1985.06.11, 全文.

EP 0858067 A2, 1998.08.12, 全文.

审查员 孙小蕾

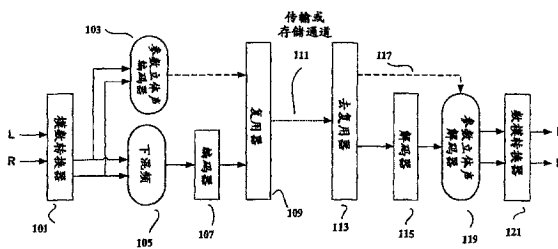
权利要求书 1 页 说明书 9 页 附图 4 页

(54) 发明名称

用于低比特率音频编码应用的高效可标度参数立体声编码

(57) 摘要

本发明提供了对现有技术的音频编解码器的改进,所述音频编解码器通过对所接收到的单声道信号的后处理生成一个立体声幻觉。通过在编码器一侧上提取立体声图像描述参数,将其发送和随后用于在解码器一侧上控制立体声生成器,实现了这些改进。此外,通过使用一种新形式的参数立体声编码,本发明弥补了简单的伪立体声方法和当前的真实立体声编码方法之间的差距。引入了一个立体声平衡参数,使得能够实现更先进的立体声模式,此外,构建了频谱包络立体声编码的新方法的基础,尤其是在使用引导 HFR(高频重建)的系统中使用。作为一种特殊的情况,描述了这种立体声编码方案在可标度的基于 HFR 的编解码器内的应用。



1. 一种在时间连续的两个平衡值之间进行内插的方法,所述平衡值是从具有第一声道和第二声道的立体声信号或者多声道信号得出的,该方法包括以下步骤:

研究从第一声道和第二声道得出的单声道信号,以获得关于声音事件的开始或结束的信息;以及

响应于该信息,计算在所述时间连续的两个平衡值之间的内插值,使得在包含小信号能量的时间片段内执行平衡位置上的改变。

2. 根据权利要求1的方法,其中,所述研究步骤包括得出所述单声道信号的能量包络的步骤。

3. 根据权利要求1或2的方法,其中,所述研究步骤包括检测在特定频带内信号能量的突然增加或降低的步骤。

4. 根据前述权利要求中任一项的方法,包括

通过对能量应用峰值保持而发现声音的开始步骤,其中

使得平衡值作为峰值保持能量的函数而增加,其中小能量值提供大的增量,反之亦然。

5. 根据前述权利要求中任一项的方法,其中,

所述平衡值是第一声道和第二声道的能量的商,

所述时间连续的两个平衡值用对数值来表示,以及

所述计算步骤计算以对数表示的所述内插值。

6. 根据权利要求1的方法,其中,将所述内插值计算为使得所述单声道信号的相应功率的瞬时值控制所述瞬态内插应当采用的陡度。

7. 一种用于在时间连续的两个平衡值之间进行内插的设备,所述平衡值是从具有第一声道和第二声道的立体声信号或者多声道信号得出的,该设备包括:

研究装置,用于研究从第一声道和第二声道得出的单声道信号以获得关于声音事件的开始或结束的信息;以及

计算装置,用于响应于所述信息,计算在所述时间连续的两个平衡值之间的内插值,使得在包含小信号能量的时间片段内执行平衡位置上的改变。

用于低比特率音频编码应用的高效可标度参数立体声编码

[0001] 本申请是申请日为 2002 年 7 月 10 日、申请号为 02813646.2、发明名称为“用于低比特率音频编码应用的高效可标度参数立体声编码”的发明专利申请的分案申请。

技术领域

[0002] 本发明涉及低比特率音频源编码系统。介绍了输入信号的立体声特性的不同参数表示,解释了它们在解码器一侧上的应用,从频谱包络的伪立体声到完全的立体声编码,后者尤其适合于基于 HFR(高频重建)的编解码器。

[0003] 发明背景

[0004] 音频源编码技术可以分成两类:自然音频编码和语音编码。在中到高的比特率上,通常将自然音频编码用于语音和音乐信号,并能进行立体声传输和再现。在仅能获得低比特率的应用中,例如目标为具有慢速电话调制解调器连接的用户在互联网流式音频,或者在新兴的数字调幅广播系统中,不可避免地需要进行音频节目资料的单声道编码。然而,依然希望有立体声印象,尤其当使用耳机收听时,在这种情况下,纯单声道信号给人的感觉像是从“头内”产生的,这可能是一种很不舒服的感觉。

[0005] 解决这个问题的一种方法是在解码器一侧上利用所接收的纯单声道信号合成出一个立体声信号。近年来,已经推荐了几种不同的“伪立体声”生成器。例如在美国专利 US5,883,962 中,描述了通过向未处理的信号添加一个延迟/相移形式的信号来增强单声道信号,从而建立一种立体声幻觉。在这种方法中,以电平相等但是符号相反,对于两个输出中的每个输出将处理后信号添加给原始信号,保证了如果两个声道随后在信号路径中叠加则消除增强信号。在 PCTW098/57436 中,描述了一种类似的系统,尽管没有上述增强信号的单声道的兼容性。现有技术方法的共同点在于将它们作为纯粹的后处理来使用。换句话说,解码器不能获知立体声宽度的信息,更不用说在立体声录音室内的位置。因此,伪立体声信号可能有也可能没有原信号的立体声特性的类似之处。现有技术系统不适用的一种具体的情况是当原信号是一个纯单声道信号时,这通常是在语音录音的情况下。在解码器上将这个单声道信号盲目地转换成一个合成立体声信号,在语音情况下这通常将导致令人不舒服的人为加工感觉,并可能降低清晰度和语音可理解性。

[0006] 目的在于在低比特率上的真实立体声传输的其它现有技术系统通常使用一种求和和差值编码方案。因此,将原始的左(L)和右(R)信号转换成一个求和信号 $S = (L+R)/2$ 和一个差值信号 $D = (L-R)/2$,随后编码和传输。接收机解码 S 和 D 信号,在此基础上通过操作 $L = S+D$ 和 $R = S-D$ 重新建立原始的 L/R 信号。这种方法的优点在于随时可以使用 L 和 R 之间的冗余,因此与 S 相比,需要编码的 D 内的信息更少,需要更少的比特。显然,极端情况是纯单声道信号,即 L 和 R 相等。传统的 L/R 编解码器两次编码这个单声道信号,而一个 S/D 编解码器检测这个冗余, D 信号(理想上)不需要任何比特。另一种极端情况用与“异相”信号对应的 $R = -L$ 的情况来表示。现在, S 信号为零,而 D 信号计算为 L。同样, S/D 方案与标准的 L/R 编码相比具有明显的优点。然而,考虑例如在信号通路中 $R = 0$ 的情况,这在早期的立体声录音中是很常见的。S 和 D 都等于 $L/2$, S/D 方案并不提供任何优点。

相反,L/R 编码方法将这种情况处理得很好 :R 信号并不需要任何比特。因此,现有技术的编解码器根据在给定瞬间哪种方法最有益而在这两种编码方案之间自适应地切换。上面的例子仅是理论性的(除了双单声道的情况之外,双单声道的情况在纯语音的节目中很普遍)。因此,实际的立体声节目资料包含大量的立体声信息,即使执行上述切换,所获得的比特率对于许多应用来说通常依然太高。此外,如从上面的重新合成关系可以看出的,为了进一步降低比特率而非常粗糙地量化 D 信号也是不可行的,因为量化误差将转换成在 L 和 R 信号中不可忽略的电平误差。

[0007] 发明概述

[0008] 本发明在编码和传输之前使用信号立体声特性的检测。在最简单的形式中,检测器测量在输入的立体声信号中出现的立体声感觉的量值。然后,将这个量值作为立体声宽度参数和原始信号的编码单声道之和一起传输。接收机使用受所述参数控制的伪立体声生成器解码这个单声道信号,并应用适当的立体声宽度量值。作为一种特殊的情况,作为一个零立体声宽度来传输一个单声道输入信号,并相应地在解码器中不使用立体声合成。根据本发明,例如,可以根据原始左右声道的差值信号或互相关来确定立体声宽度的有用测量值。可以将如此计算出的量值映射成少量的状态,实时地或者根据需要地在合适的固定速率上传输这些状态。本发明还教导了如何滤波合成的立体声分量,从而降低不能掩蔽通常与低比特率编码信号有关的编码人工噪声的风险。

[0009] 可选择地,在编码器内检测立体声域内的整体立体声平衡或定位。这个信息可选地与上述宽度参数一起作为一个平衡参数与编码的单声道信号一起有效地传输。因而,通过相应地改变两个输出声道的增益,可以在解码器上重建相对于录音室两侧的位置偏移。根据本发明,可以根据左右信号功率的商来获得这个立体声平衡参数。与整体立体声编码相比,这两种参数的传输需要很少的比特,因此总的比特率需求很低。在本发明的一种更好的提供更准确的参数立体声描述的设计中,使用若干平衡和立体声宽度参数,分别代表各个频带。

[0010] 推广到每频带操作的平衡参数和计算为左右信号功率之和的电平参数的相应每频带操作一起实现了立体声信号功率谱密度的一种新的任意详细的表示。除了 S/D 系统也利用的立体声冗余的优点之外,这种表示的一种具体的优点是 :与同样的电平信号相比,可以用更低的精确度来量化平衡信号,因为当转换回立体声频谱包络时量化误差将导致“空间上的误差”,即在立体声全景中可感知的位置,而不是在电平上的误差。类似于传统的切换 L/R 和 S/D 系统,有利于电平 L/ 电平 R 信号,可以自适应地关闭电平 / 平衡方案,当整体信号向任一声道严重偏置时这更加有效。无论何时需要一种功率谱包络的有效编码方法,都可以使用上述的频谱包络编码方案,并可以作为一种工具结合在新的立体声源编解码器内。一个特别感兴趣的应用是在受原始信号高频带包络的相关信息引导的 HFR 系统内。在这样一个系统中,使用任意一个编解码器来编码和解码低频带,在解码器上使用解码的低频带信号和所发送的高频带包络信息重新生成高频带 (PCTWO 98/57436)。此外,通过将包络编码锁定到电平 / 平衡操作,提供了建立一个可标度 (scalable) 的基于 HFR 的立体声编解码器的可能性。在此,将电平值馈送给主比特流,根据实现方式,该主比特流通常解码成一个单声道信号。将平衡值馈送给次比特流,该次比特流是除了主比特流之外靠近发射机的接收机可以获得的,例如一个 IBOC(带内信道上) 数字调幅广播系统。当组合这两个比

特流时,解码器生成一个立体声输出信号。除了电平值之外,主比特流还可以包含立体声参数,例如宽度参数。因此,单独解码这个比特流已经生成了一个立体声输出,当两个比特流都可用时改善了立体声输出。

附图说明

[0011] 现在,参考附图,通过并非限制本发明范围或精神的说明性例子来描述本发明,在附图中:

[0012] 图 1 图示一个音源编码系统,其中包括一个由参数立体声编码器模块增强的编码器和一个由参数立体声解码器模块增强的解码器;

[0013] 图 2a 是一个参数立体声解码器模块的模块示意图;

[0014] 图 2b 是一个带有控制参数输入的伪立体声生成器的模块示意图;

[0015] 图 2c 是一个带有控制参数输入的平衡调整器的模块示意图;

[0016] 图 3 是使用与多频段平衡调整相组合的多频段伪立体声生成的参数立体声解码器模块的模块示意图;

[0017] 图 4a 是一个可标度的基于 HFR 的立体声编解码器的编码器一侧的模块示意图,它使用频谱包络的电平/平衡编码;

[0018] 图 4b 是相应的解码器一侧的模块示意图。

[0019] 优选实施例的描述

[0020] 下面所述的实施例仅仅是为了说明本发明的原理。应当理解对于本领域的技术人员来说,对在此所描述的结构和细节进行各种修改和变化将是显而易见的。因此,将仅通过权利要求书的保护范围来限制,而不是由在此对实施例的描述和解释所提出的具体细节来限制。为了清楚起见,下面所有的例子均假设为双声道系统,但是对于本领域的技术人员来说,本发明显然也可以应用于多声道系统,例如 5.1 声道系统。

[0021] 图 1 图示一个可通过根据本发明的参数立体声编码增强的任意声源编码系统,其中包括编码器 107 和解码器 115,编码器和解码器以单声道模式工作。假设 L 和 R 代表馈送给 AD(模数)转换器 101 的左右模拟输入信号。AD 转换器的输出转换成单声道信号 105,然后编码(107)该单声道信号。此外,将立体声信号发送给参数立体声编码器 103,由其计算下面将要描述的一个或多个立体声参数。通过多路复用器 109 组合这些参数与编码的单声道信号,形成一个比特流 111。存储或发送该比特流,然后在解码器一侧使用去复用器 113 进行提取。该单声道信号被解码 115,并由参数立体声解码器 119 使用立体声参数 117 作为控制信号转换成一个立体声信号。最后,将该立体声信号发送给 DA(数模)转换器 121,由其馈送模拟输出 L' 和 R'。根据图 1 的拓扑结构为一组参数立体声编码方法所共有,随后将从较简单的形式开始详细地描述这些参数立体声编码方法。

[0022] 根据本发明的一种立体声特性参数化的方法是在编码器一侧上确定原始信号的立体声宽度。立体声宽度的第一种近似是差值信号 $D = L - R$,因为一般来说 L 和 R 之间的高度类似性将计算出较小的 D 值,反之亦然。特殊的情况是双单声道,其中 $L = R$,因此 $D = 0$ 。因而,即使这种简单的算法也能够检测出通常与新闻广播有关的单声道输入信号的类型,在新闻广播的情况下并不希望伪立体声。然而,在不同电平上馈送给 L 和 R 的单声道信号并不产生为零的 D 信号,即使感觉到的宽度为零。因此,实际上可能需要更精致的检测

器,例如使用互相关的方法。应当确信与总信号电平一起标准化以某种方式描述左右差值或相关性的数值,从而实现与电平无关的检测器。上述检测器的一种问题是在语音到音乐/音乐到语音转换的过程中单声道语音与诸如立体声噪声或背景音乐等非常微弱的立体声信号混合时的情况。在语音暂停时,则检测器将指示一个宽立体声信号。通过标准化立体声宽度值与一个包含先前的总能量电平信息的信号,例如一个总能量的峰值衰落信号,来解决这个问题。此外,为了防止高频噪声或信道不同的高频失真触发立体声宽度检测器,应当通过一个低通滤波器预先滤波检测器信号,该滤波器通常具有一个稍高于语音第二共振峰的截止频率,也可以选择使用一个高通滤波器以避免不平衡的信号偏移或交流声。不考虑检测器的类型,将所计算的立体声宽度映射成有限的一组值,覆盖从单声道到宽立体声的整个范围。

[0023] 图 2a 图示在图 1 中介绍的参数立体声解码器的内部结构的一个例子。随后将描述受参数 B 控制的标有“平衡”的模块 211,现在应当将其视为旁路。标有“宽度”的模块 205 接收一个单声道输入信号,合成地重新建立立体声宽度的印象,其中宽度量值受参数 W 的控制。随后将描述可选的参数 S 和 D。根据本发明,通过结合使用一个包括低通滤波器 (LP) 203 和高通滤波器 (HP) 201 的分频滤波器,从而保持低频范围“固定”和不受影响,通常能够实现主观上更好的音频质量。在此,仅将高通滤波器的输出发送给宽度模块。通过 207 和 209 将宽度模块的立体声输出添加给低通滤波器的单声道输出,形成立体声输出信号。

[0024] 可以将现有技术的任意一种伪立体声生成器用于宽度模块,例如在背景技术部分中所提到的,或者是 Schroeder 类型的早期反射模拟单元(多抽头延迟)或混响器。图 2b 图示馈送一个单声道信号 M 的伪立体声生成器的一个例子。由 215 的增益确定立体声宽度的量值,此增益是立体声宽度参数 W 的函数。增益越高,则立体声印象越宽,零增益对应于纯单声道再现。215 的输出被延迟 (D),221,并使用相反的符号与两个直接信号的例子相加 223 和 225。为了当改变立体声宽度时不明显地改变总的再现电平,可以结合使用直接信号的补偿衰减 213。例如,如果延迟信号的增益是 G,则可以将直接信号的增益选择为 $\sqrt{1-G^2}$ 。根据本发明,可以在延迟信号路径中插入一个高频滚降滤波器 217,这有助于避免伪立体声导致的编码人工噪声无掩蔽。可选择地,可以在比特流中发送分频滤波器、滚降滤波器和延迟的参数,提供模拟原始信号的立体声特性的更高的可能性,如在图 2a 和图 2b 中图示的信号 X、S 和 D。如果使用混响单元来生成立体声信号,有时在一个声音结束之后可能会产生不希望有的混响衰落。然而,仅仅通过改变混响信号的增益就能够轻易地衰减或完全消除这些不希望的混响尾部。可以将为发现声音结尾而设计的检测器用于此目的。如果混响单元在一些特殊的信号例如瞬变信号上产生人工噪声时,用于这些信号的检测器也可以用于衰减人工噪声。

[0025] 下面描述根据本发明的检测立体声特性的另一种方法。再次假设 L 和 R 代表左右输入信号。然后用 $P_L \sim L^2$ 和 $P_R \sim R^2$ 代表相应的信号功率。现在,可以将立体声平衡的测量值计算为两个信号功率的商,或者更具体地说计算为 $B = (P_L + e) / (P_R + e)$,其中 e 是任意的非常小的数值,它避免被零除。可以通过关系 $B_{dB} = 10 \log_{10}(B)$ 用 dB 表示平衡参数 B。例如,三种情况 $P_L = 10P_R$, $P_L = P_R$ 和 $P_L = 0.1P_R$ 分别对应于 +10dB、0dB 和 -10dB 的平衡值。很显然,这些值映射成位置“左”、“中心”和“右”。实验已经表明平衡参数的范围例如可以限制为 +/-40dB,因为这些极限数值已经可以被视为声音完全从两个扬声器或耳机驱动器

之一发出。这种限制降低了在传输中要覆盖的信号空间,因而降低了比特率。此外,可以使用渐进的量化方案,由此在零附近使用较小的量化步长,而在上限上使用较大的量化步长,这进一步降低了比特率。通常在延长路径的时间上平衡是恒定的。因此,可以采用的显著地降低所需要的平均比特的数量的最后一个步骤是:在传输一个初始化平衡量值之后,仅传输相邻平衡量值之间的差值,从而使用熵编码。非常普遍地,这个差值为零,因而可以用可能的最短的码字来表示以传输。显然地,在可能存在比特误差的应用中,必需在合适的时间间隔上重新设置这个 Δ 编码,从而消除不受控制的误差传播。

[0026] 通过将单声道信号馈送给两个输出,并使用控制信号 B 相应地调整增益,如图 2c 中模块 227 和 229 所图示的,平衡参数的最基本的解码器使用方法简单地将单声道信号向两个再现声道之一上偏移。这类似于调整混合台上的“全景”调节器,合成地在两个立体声扬声器之间“移动”单声道信号。

[0027] 除了上述的宽度参数之外,还可以发送平衡参数,提供以受控方式在录音室内定位和传播声音图像的可能性,并提供当模仿原始的立体声压缩时的灵活性。组合上述的伪立体声生成和参数控制平衡的一个问题是在远离中心位置的平衡位置上伪立体声生成器的不希望有的信号影响。这通过在立体声宽度值上应用一个有利于单声道的函数来解决,所述函数导致在最靠边的位置上的平衡位置上立体声宽度值的较大衰减,并在靠近中央位置的平衡位置上的较小或无衰减。

[0028] 上述方法用于很低的比特率应用。在能够获得较高的比特率的应用中,可以使用上述宽度和平衡方法的更精致的形式。可以在多个频带上执行立体声宽度检测,导致各个立体声宽度值分别用于每个频带。类似地,平衡计算可以通过多频带的方式进行,这相当于将不同的滤波器曲线应用于馈送一个单声道信号的两个声道。图 3 图示一个参数立体声解码器的例子,它使用根据图 2b,用模块 307、317 和 327 代表的一组 N 个伪立体声生成器,并组合如图 2c 所示用模块 309、319 和 329 代表的多频带平衡调整。通过将单声道输入信号 M 馈送给一组带通滤波器 (BP) 305、315 和 325 来获得各个通频带。平衡调整器输出的带通立体声输出被相加,311、321、313 和 323,形成立体声输出信号 L 和 R。现在,用阵列 $W(k)$ 和 $B(k)$ 来替代原先的标量宽度和平衡参数。在图 3 中,每个伪立体声生成器和平衡调整器具有特有的立体声参数。然而,为了降低将要发送或存储的数据总量,可以在编码器上分组地平均若干频带的参数,数量减少的这些参数可以在解码器上映射到相应的宽度和平衡模块组。显然地,可以将不同的分组方案和长度用于阵列 $W(k)$ 和 $B(k)$ 。 $S(k)$ 代表在宽度模块内延迟信号路径的增益,和 $D(k)$ 代表延迟参数。同样地,在比特流中 $S(k)$ 和 $D(k)$ 是可选的。

[0029] 参数平衡编码方法可特别适用于较低的频带,假设由于频率分辨率较低,或者由于在同一时间上但是在不同的平衡位置上在一个频带内出现过多的声音事件,导致稍微不稳定的性能。这些平衡错误的特征通常为在非常短的时间周期内一个不正常的平衡值,通常是更新速率计算出的一个或多个连续的数值。为了避免扰乱平衡错误,可以在平衡数据上应用一个稳定化处理。这个处理可以在当前的时间位置之前和之后使用多个平衡值来计算这些数据的中值。该中值随后可以用作当前平衡值的限制值,即当前平衡值应当不允许低于该中值。然后,将当前值限制在最后一个数值和中值之间的范围内。可选择地,可允许当前的平衡值以某一过量因子超过该限制值。此外,过量因子以及用于计算中值的平

衡值数量应当被视为频率相关特性的,因此各自用于每个频带。

[0030] 在较低的平衡信息的更新比例上,时间分辨率缺乏可能导致立体声图像和实际的声音事件的运动之间同步上的错误。为了改善同步方面的性能,可以使用以识别声音事件为基础的内插方案。在此,内插是指在时间连续的两个平衡值之间的内插。通过在接收机一侧上研究单声道信号,能够获得不同声音事件的开始和结束的相关信息。一种方法是检测在特定频带内信号能量的突然地增加或降低。内插应当在能量包络在时间上引导之后,以确保最好应当在包含小信号能量的时间片段内执行平衡位置上的改变。因为人耳对声音的开始部分比对声音的结束部分更为敏感,例如通过对能量应用峰值保持,然后使平衡值作为峰值保持能量的函数而增加,其中较小的能量值提供较大的增加,反之亦然,内插方案的优点在于发现一个声音的开始部分。对于包含在时间上不均匀分布能量的时间片段来说,即对于一些固定信号来说,这种内插方法相当于两个平衡值之间的线性内插。如果平衡值是左右能量的商,因为左右对称的原因,优选对数平衡值。在对数域内使用整个内插算法的另一个优点是使电平与对数标度相关的趋势。

[0031] 而且,对于立体声宽度增益值的较低的更新比例,也可以应用内插。一种简单的方法是在时间上连续的两个立体声宽度值之间线性地内插。通过在包含多个立体声宽度参数的一个较长的时间片段上平滑立体声宽度增益值,能够实现立体声宽度更稳定的特性。通过利用通过不同的上升和释放时间常数的平滑,实现了一种尤其适合于包含混合或交织的语音和音乐的节目资料的系统。因此在立体声中对音乐开始部分的立即响应,使用一个短的上升时间常数来获得一个短的上升时间,并使用一个长的释放时间来获得一个长的下降时间,实现这种平滑滤波器的一种合理设计。为了快速地从宽立体声模式切换到单声道模式,这可能是突然的语音开始部分所希望的,存在通过通知这个事件来旁路或重置该平滑滤波器的可能性。此外,上升时间常数、释放时间常数和其它的平滑滤波器特性也可以由编码器来通知。

[0032] 对于包含心理声学编解码器的掩蔽失真的信号来说,引入基于编码单声道信号的立体声信息的一个共同的问题是失真的未掩蔽效应。这种通常称作“立体声未掩蔽”的现象是并不符合掩蔽标准的非居中声音的结果。通过在解码器一侧上引入用于这种情况的检测器可以解决或部分地解决立体声未掩蔽的问题。可以使用测量信号与掩蔽之比的公知技术来检测潜在的立体声未掩蔽。一旦检测到,可以明确地通知,或者可以仅简单地降低立体声参数。

[0033] 在编码器一侧上,如本发明所教导的,一种选择是将一个希耳伯特变换器用于输入信号,即引入在两个声道之间的 90 度相移。当随后通过相加两个信号形成单声道信号时,实现了中央摆动单声道信号和“真实的”立体声信号之间更佳的平衡,因为希耳伯特变化为中央信息引入了 3dB 的衰减。实际上,这改善了诸如当前流行音乐的单声道编码,例如通常使用单声道音源来录音引导声音和低音电吉他。

[0034] 多频带平衡参数方法并不限制于图 1 所描述的应用类型。只要目标是有效地编码一个立体声信号的功率谱包络,就可以有效地使用该方法。因此,可以在立体声编解码器中将其用作工具,其中除了立体声频谱包络之外,还编码一个相应的立体声残余信号。假设总能量 P ,用 $P = P_L + P_R$ 来定义,其中 P_L 和 P_R 是如上所述的信号功率。注意到这种定义并不考虑从左至右的相位关系。(例如,相等的左和右信号但是符号相反,并不产生一个零的总能

量)。类似于 B, 可以用 dB 将 P 表示为 $P_{dB} = 10 \log_{10}(P/P_{ref})$, 其中 P_{ref} 是一个任意的参考功率, Δ 的值被熵编码。与平衡的情况相反, 不将渐进的量化用于 P。为了表示一个立体声信号的频谱包络, 为一组频带计算 P 和 B, 一般地但并不必需地, 使用与人耳的临界频带有关的带宽。例如, 通过在一个常数带宽滤波器组内对声道分组可以形成这些频带, 由此将 P_L 和 P_R 计算为对应于相应频带和时间上相应周期的子频带平方的时间和频率的平均值。这些组 $P_0, P_1, P_2, \dots, P_{N-1}$ 和 $B_0, B_1, B_2, \dots, B_{N-1}$, 其中下标代表 N 个频带表示中的频带, 被 Δ 和霍夫曼编码, 发送或存储, 并最终解码成在编码器中计算的量化值。最后一个步骤是将 P 和 B 转换回 P_L 和 P_R 。如根据 P 和 B 的定义很容易看出的, 反向的关系为 (当忽略 B 定义中的 ϵ) $P_L = BP/(B+1)$, 和 $P_R = P/(B+1)$ 。

[0035] 上述包络编码方法的一种特别有用的应用是为基于 HFR 的编解码器编码高频带的频谱包络。在这种情况下, 不发送高频带的残余信号。而是根据低频带获得这个残余信号。因而, 残余和包络表示之间不存在严格的关系, 包络量化更重要。为了研究量化的效果, 假设 P_q 和 B_q 分别代表 P 和 B 的量化值。则将 P_q 和 B_q 插入在上述关系中, 总和为:

$$[0036] \quad P_L q + P_R q = B_q P_q / (B_q + 1) + P_q / (B_q + 1) = P_q (B_q + 1) / (B_q + 1) = P_q.$$

[0037] 在这里感兴趣的特征是消除了 B_q , 总功率上的误差仅仅由 P 的量化误差来确定。这意味着即使 B 被严重量化, 所感觉到的电平也是正确的, 假设在 P 的量化中使用了足够高的精确度。换句话说, B 中的失真映射成了空间上的失真, 而不是电平上的失真。只要声源随着时间在空间内是稳定的, 则立体声感觉上的这个失真也是稳定的, 并很难注意到。如已经描述的, 立体声平衡的量化在上限附近也可以比较粗糙, 因为当到中心线的角度很大时, 由于人类听觉的特性, 用 dB 表示的给定误差对应于在所观察角度上的较小的误差。

[0038] 当量化与频率相关的数据例如多频带立体声宽度增益值或多频带平衡值时, 能够有利地选择量化方法的分辨率和范围, 以匹配听觉标度特性。如果这种标度取决于频率, 可以为不同的频带选择不同的量化方法或者所谓的量化种类。因此, 在一些情况下, 即使数值相同, 代表不同频带的编码参数值也应当用不同的方式来解释, 即解码成不同的数值。

[0039] 类似于切换 L/R 到 S/D 的编码方案, 可以自适应地用 P_L 和 P_R 信号来替代 P 和 B 信号, 从而更好地应付极端信号。如 PCT/SE00/00158 所教导的, 根据在特定时刻上在比特数量方面哪个方向最有效, 可以将包络抽样的 Δ 编码从时间上的 Δ 切换成频率上的 Δ 。平衡参数也可以采用这种方案: 例如考虑一个随着时间进入立体声域的音源。显然地, 这对应于平衡值随着时间的连续改变, 这取决于音源的速度与参数更新速率之比, 可能对应于较大的时间上的 Δ 值, 当使用熵编码时对应于较大的码字。然而, 假设音源在频率上具有均匀的声音辐射, 平衡参数在频率上的 Δ 值在时间上的每个点上为零, 则再次对应于一个较小的码字。因而, 在这种情况下, 当使用频率 Δ 编码方向时实现了较低的比特率。另一个例子是在房间内固定但是具有非均匀辐射的音源。现在, 频率上的 Δ 值较大, 而时间上的 Δ 值是优选的。

[0040] P/B 编码方案提供了建立一个可标度的基于 HFR 的编解码器的可能性, 参见图 4。可标度的编解码器的特点在于将比特流分割成两个或更多的部分, 其中可以选择高阶部分的接收和解码。该例子假设两个比特流部分, 在下文中称作主部分 419 和次部分 417, 但是扩展成更多的部分显然也是可以的。图 4a 所示编码器一侧包括: 任意立体声低频带编码器 403, 它在输入信号 IN 上操作 (在该图中未图示详细的 AD 和相应的 DA 转换步骤); 参数立

体声编码器,它消除高频带频谱包络,和可选的附加立体声参数 401,它也在立体声输入信号上操作;和两个复用器(MUX)415和413,分别用于主和次比特流。在这个应用中,将高频带包络编码锁定到P/B操作,通过415将P信号407发送给主比特流,而通过413将B信号405发送给次比特流。

[0041] 对于低频带编解码器来说,存在不同的可能性:它可能始终工作在S/D模式中,并将S和D信号分别发送给主和次比特流。在这种情况下,主比特流的解码产生一个完整频带的单声道信号。当然,可以通过根据本发明的参数立体声方法来增强这个单声道信号,在这种情况下立体声参数也必须位于主比特流内。另一种可能性是将一个立体声编码低频带信号馈送给主比特流,可选择地与高频带和平衡参数一起。现在,主比特流的解码产生低频带的真实的立体声,高频带的非常逼真的伪立体声,因为低频带的立体声特性被反映在高频的重新构建上。描述另外一种方式:即使可用的高频带包络表示或频谱粗略的结构是在单声道内,合成后的高频带残余或频谱精细结构也不是在单声道内。在这种实施方式中,次比特流可能包含更低频带的信息,当将其与主比特流组合时,产生更高质量的低频带再现。图4的拓扑结构表示两种情况,因为分别连接到415和417的主和次低频带编码器输出信号411和409可能包含上述任一种信号类型。

[0042] 发送或存储比特流,仅将419或同时将419和417馈送给解码器,图4b。由423将主比特流去复用成低频带核心解码器主信号429和P信号431。类似地,由421将次比特流去复用成低频带核心解码器次信号427和B信号425。将这个(些)低频带信号发送给低频带解码器433,它生成一个输出435,在仅解码主比特流的情况下,该输入也可以是上述任意类型的(单声道或立体声)。将信号435馈送给HFR单元437,其中生成一个合成高频带,并根据P来调整,所述P也连接到HFR单元。在HFR单元内组合解码后的低频带与高频带,在最终馈送给系统输出之前,可选地通过伪立体声生成器(也位于HFR单元内)来增强低频带和/或高频带,形成输出信号OUT。当存在次比特流417时,HFR单元也获得B信号作为一个输入信号425,435是立体声的,因此该系统生成一个完全立体声的输出信号,并旁路伪立体声生成器,如果有的话。

[0043] 换句话说,一种输入信号的立体声特性的编码方法包括:在编码器上,计算一个代表所述输入信号的立体声宽度的宽度参数;和在解码器上,生成一个立体声输出信号,使用所述宽度参数控制所述输出信号的立体声宽度。该方法还可以包括在所述编码器上根据所述输入信号形成一个单声道信号,其中,在所述解码器上所述生成是指在所述单声道信号上操作的伪立体声方法。其中,所述伪立体声方法可以指将所述单声道信号分成两个信号,并在所述宽度参数控制的电平上将所述单声道信号的延迟形式添加给所述两个信号。其中,可以在添加给所述两个信号之前,将所述延迟形式高通滤波,并在较高的频率上渐进地衰减。其中,所述宽度参数可以是一个矢量,所述矢量的各单元对应于各个频带。其中,如果所述输入信号是双单声道类型的,则所述输出信号也可以是双单声道类型的。

[0044] 另一种输入信号的立体声特性的编码方法包括:在编码器上,计算代表所述输入信号的立体声平衡的平衡参数;和在解码器上,生成一个立体声输出信号,使用所述平衡参数控制所述输出信号的立体声平衡。

[0045] 在这种方法中,可以在所述编码器上根据所述输入信号形成一个单声道信号;在所述解码器上,所述生成是指将所述单声道信号分成两个信号,所述控制是指调整所述两

个信号的电平。该方法还可以包括：计算所述输入信号的每个声道的功率，并根据所述功率之间的商计算所述平衡参数。其中，所述功率和所述平衡参数可以是其中每个单元对应于一个特定频带的矢量。该方法还可以包括：在所述解码器上，在所述平衡参数的时间上连续的两个值之间内插，以便所述单声道信号的相应功率的瞬时值控制所述瞬时内插应当采用的陡度。其中，可以在表示为对数值的平衡值上执行所述内插方法。其中，可以将所述平衡参数的数值限制在前一平衡值和由一个中值滤波或其它滤波处理从其它平衡值提取出的一个平衡值之间的范围内，所述范围可以通过用某一因数移动所述范围的边界来进一步地扩展。其中，所述提取用于平衡值的限制边界的方法对于一个多频段系统来说可以是取决于频率的。其中，可以将一个附加的电平参数计算为所述功率的矢量之和并发送给所述解码器，从而向所述解码器提供所述输入信号的频谱包络的表示。其中，所述电平参数和所述平衡参数可以自适应地用所述功率来替换。其中，所述频谱包络可以用于控制一个解码器内的 HFR 处理。其中，可以将所述电平参数馈送给一个可标度的基于 HFR 的立体声编解码器的主比特流，并将所述平衡参数馈送给所述编解码器的次比特流。其中，可以将所述单声道信号和所述宽度参数馈送给所述主比特流。其中，可以通过一个函数来处理所述的宽度参数，所述函数为与更远离中心位置的平衡位置相对应的平衡值给出较小的数值。其中，所述平衡参数的量化可以在中心位置附近使用较小的量化步长，在较偏外的位置上使用较大的步长。其中，可以使用一个量化方法来量化所述宽度参数和所述平衡参数，所述量化方法对于一个多频带系统来说在分辨率和范围方面是取决于频率的。其中，可以在时间或者在频率上自适应地 Δ 编码所述平衡参数。其中，在形成所述单声道信号之前，可以使所述输入信号通过一个希耳伯特变换器。

[0046] 一种用于参数立体声编码的设备包括：在编码器上，用于计算代表一个输入信号的立体声宽度的宽度参数的装置和用于根据所述输入信号形成一个单声道信号的装置；以及在解码器上，用于根据所述单声道信号生成一个立体声输出信号并使用所述宽度参数来控制所述输出信号的立体声宽度的装置。

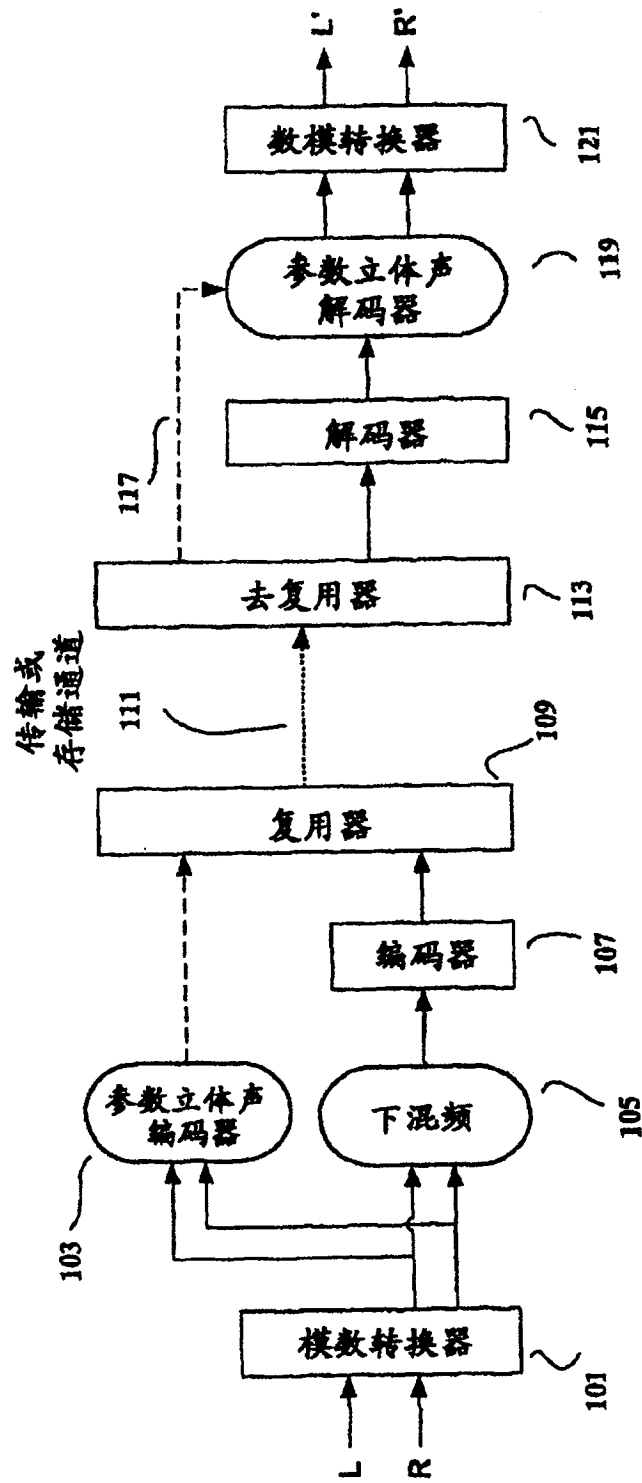


图 1

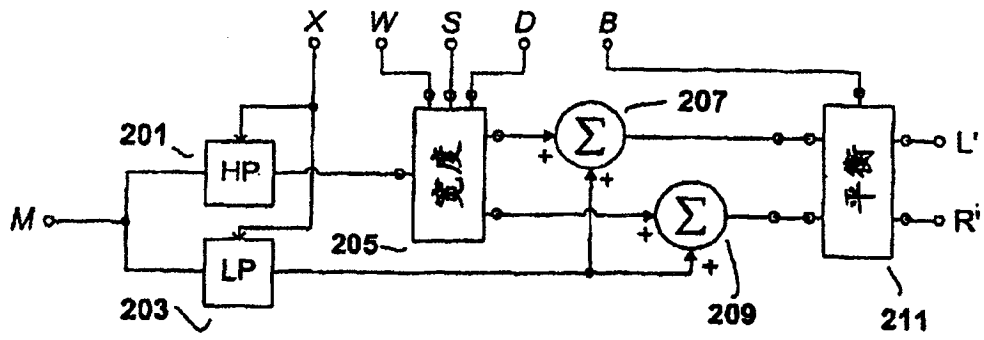


图 2a

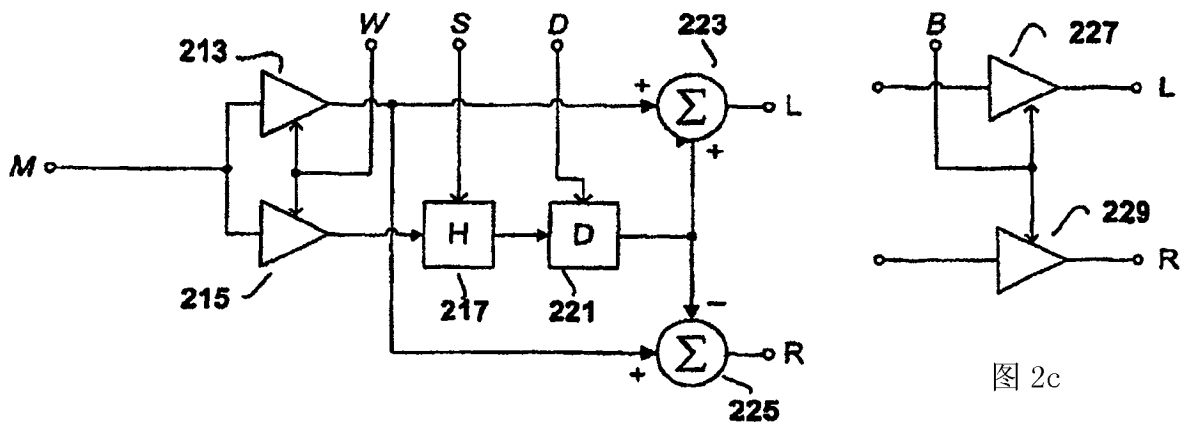


图 2c

图 2b

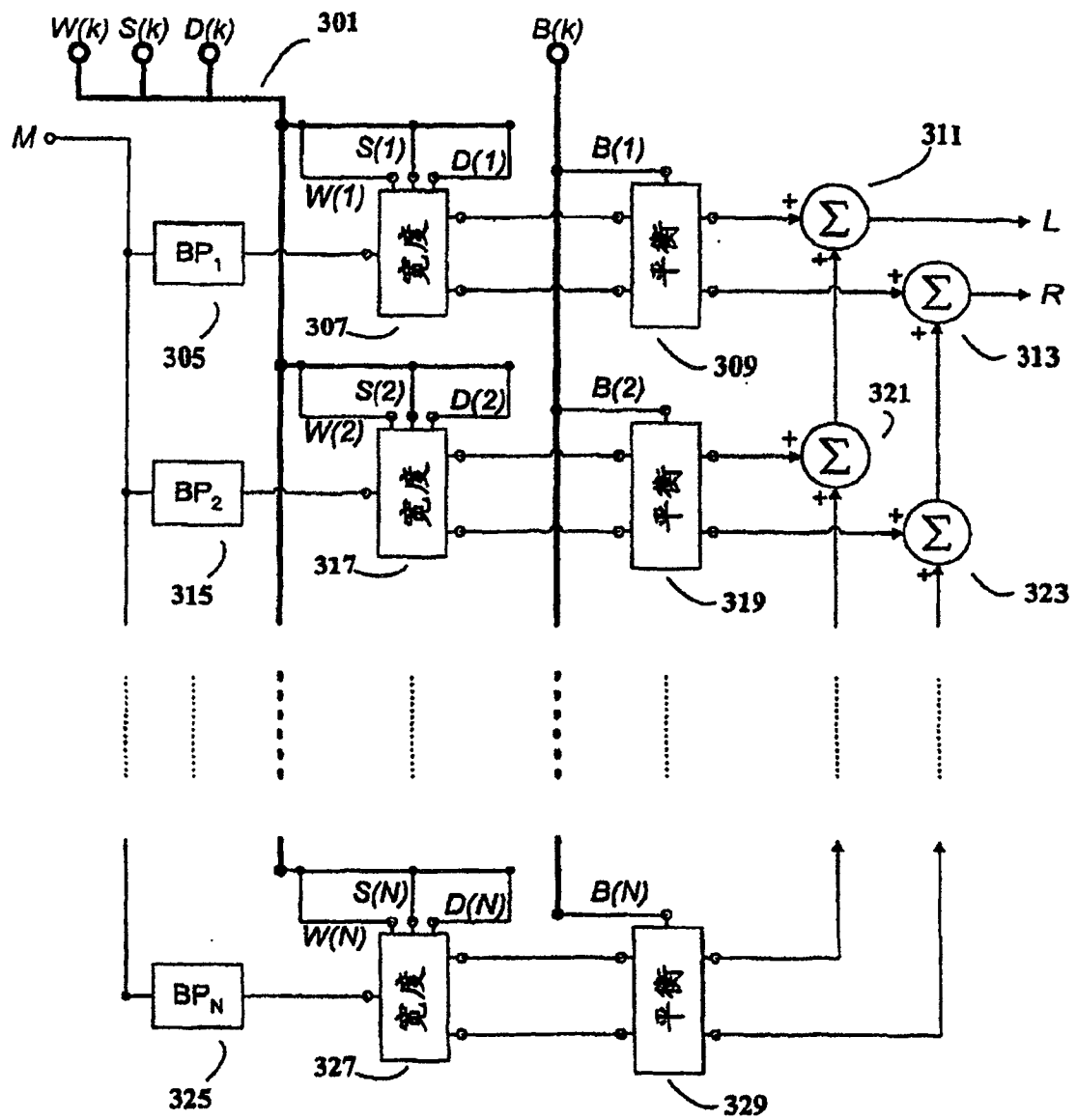


图 3

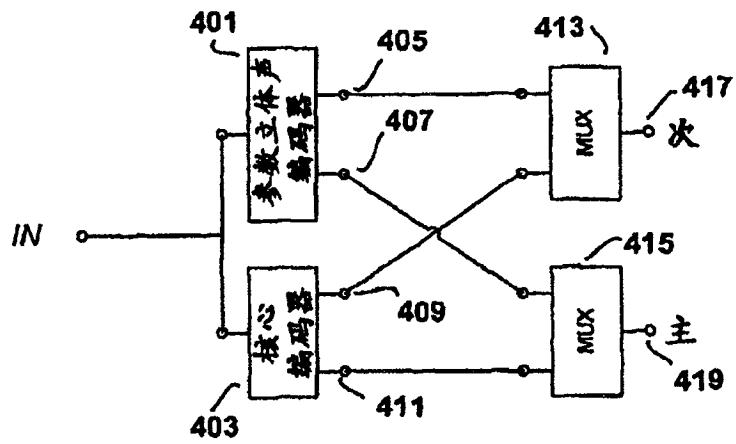


图 4a

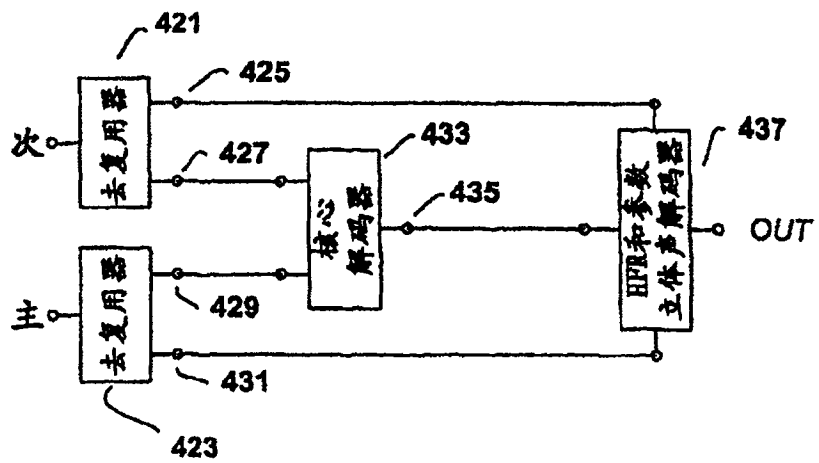


图 4b