(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101809655 B
(45)授权公告日 2012.07.25

(21)申请号 200880108722.1

(22)申请日 2008.09.09

(30)优先权数据
0718682.8 2007.09.25 GB

(85)PCT申请进入国家阶段日
2010.03.25

(86)PCT申请的申请数据
PCT/US2008/075703 2008.09.09

(87)PCT申请的公布数据
WO2009/042386 EN 2009.04.02

(73)专利权人 摩托罗拉移动公司
地址 美国伊利诺伊州

(72)发明人 乔纳森·阿拉斯泰尔·吉布斯

(74)专利代理机构 中原信达知识产权代理有限责任公司 11219
代理人 穆德骏 陆锦华

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

(56)对比文件
CN 1647155 A,2005.07.27, 说明书第4页第8段,第5页倒数第1段－第6页第5段,第11页第4段,图1,图2.
US 5434948 A,1995.07.18, 全文.
CN 1748247 A,2006.03.15, 全文.
US 2007/0165879 A1,2007.07.19, 全文.
EP 1640971 A1,2006.03.29, 全文.

审查员 张飞弦

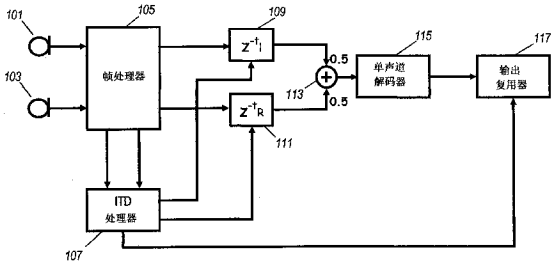
权利要求书 2 页 说明书 9 页 附图 4 页

(54)发明名称

用于编码多信道音频信号的设备和方法

(57)摘要

一种编码设备包括:帧处理器(105),该帧处理器(105)接收至少包括来自第一麦克风(101)的第一音频信号和来自第二麦克风(103)的第二音频信号的多信道音频信号。然后,ITD处理器(107)确定在第一音频信号和第二音频信号之间的间隔时间差,并且一组延迟装置(109、111)响应于间隔时间差信号通过延迟第一和第二音频信号中的至少一个来从多信道音频信号中生成补偿的多信道音频信号。然后,合并器(113)通过合并补偿的多信道音频信号的信道来生成单声道信号,并且单声道信号编码器(115)编码该单声道信号。可以基于确定在第一和第二音频信号之间的互相关通过算法来具体地地确定间隔时间差。



CN 101809655 B

1. 一种用于编码多信道音频信号的设备，所述设备包括：

接收机，所述接收机用于接收所述多信道音频信号，所述多信道音频信号至少包括来自第一麦克风的第一音频信号和来自第二麦克风的第二音频信号；

时间差单元，所述时间差单元用于通过合并在所述第一音频信号和所述第二音频信号之间的互相关的观察的连续值来确定在所述第一音频信号和所述第二音频信号之间的间隔时间差，并且其中，归一化所述互相关来得到使用类维特比算法累加的状态概率以实现具有内置滞后的间隔时间差，并且所述类维特比算法计算每个给定状态的状态概率度量作为来自进入该状态的所有路径的合并的贡献；

延迟单元，所述延迟单元用于响应于所述间隔时间差信号通过延迟所述第一音频信号和所述第二音频信号中的至少一个，从所述多信道音频信号中生成补偿的多信道音频信号；

单声道单元，所述单声道单元用于通过合并所述补偿的多信道音频信号的信道来生成单声道信号；以及，

单声道信号编码器，所述单声道信号编码器用于编码所述单声道信号。

2. 根据权利要求 1 所述的设备，其中，所述时间差单元被配置成确定针对多个时间补偿的在所述第一音频信号和所述第二音频信号之间的互相关，并且响应于所述互相关来确定所述间隔时间差。

3. 根据权利要求 2 所述的设备，其中，所述时间差单元被配置成在所述互相关之前对所述第一音频信号和所述第二音频信号进行低通滤波。

4. 根据权利要求 2 所述的设备，其中，所述时间差单元被配置成在所述互相关之前抽取所述第一音频信号和所述第二音频信号。

5. 根据权利要求 2 所述的设备，其中，所述延迟单元被配置成针对抽取的抽取因子来补偿所述间隔时间差，以便于确定用于所述第一音频信号和所述第二音频信号中的至少一个的延迟。

6. 根据权利要求 2 所述的设备，其中，所述时间差单元被配置成在所述互相关之前将频谱白化应用于所述第一音频信号和所述第二音频信号。

7. 根据权利要求 2 所述的设备，其中，所述时间差单元被配置成在所述互相关之前执行所述第一音频信号和所述第二音频信号的开窗。

8. 根据权利要求 2 所述的设备，其中，所述时间差单元包括：

网格状态机，所述网格状态机具有多个状态，所述多个状态的每个与所述多个时间补偿的时间补偿相对应；

路径单元，所述路径单元用于响应于所述互相关来确定用于所述网格状态机的状态的路径度量；

计算单元，所述计算单元用于响应于与从先前的状态到当前状态的路径相关联的路径度量来确定用于所述每个给定状态的状态概率度量；以及

用于响应于所述状态概率度量来确定所述间隔时间差的单元。

9. 根据权利要求 1 所述的设备，其中，所述延迟单元被配置成，通过响应于第一延迟生成第一补偿的多信道音频信号和响应于第二延迟生成第二补偿的多信道音频信号来从所述第一延迟转变成所述第二延迟，并且合并所述第一补偿的多信道音频信号和所述第二补

偿的多信道音频信号以生成所述补偿的多信道音频信号。

10. 一种用于编码多信道音频信号的方法，所述方法包括：

接收所述多信道音频信号，所述多信道音频信号至少包括来自第一麦克风的第一音频信号和来自第二麦克风的第二音频信号；

通过合并在所述第一音频信号和所述第二音频信号之间的互相关的观察的连续值来确定在所述第一音频信号和所述第二音频信号之间的间隔时间差，并且其中，处理所述互相关来得到使用类维特比算法累加的概率，所述类维特比算法计算每个给定状态的状态概率度量作为来自进入该状态的所有路径的合并的贡献；

响应于所述间隔时间差信号通过延迟所述第一音频信号和所述第二音频信号中的至少一个，从所述多信道音频信号中生成补偿的多信道音频信号；

通过合并所述补偿的多信道音频信号的信道来生成单声道信号；以及，

在单声道信号编码器中编码所述单声道信号。

用于编码多信道音频信号的设备和方法

技术领域

[0001] 本发明涉及用于编码多信道音频信号的设备和方法，并且具体地但不排他性地，涉及将立体声语音信号下混合为用于通过诸如码激励线性预测编码器的单声道编码器编码的单声道信号。

背景技术

[0002] 对于越来越多的应用和系统，音频信号的有效编码是必需的。例如，移动通信使用有效语音编码器来减少需要通过空中接口来传送的数据量。

[0003] 例如，国际电信联盟 (ITU) 正在标准化被称为嵌入式可变比特速率编解码器 (EV-VBR) 的语音编码器，其可以以从 8 到 64kbps 范围的数据速率高质量地编码语音信号。该编码器以及许多其他的有效语音编码器使用码激励线性预测 (CELP) 技术来以操作的较低比特速率实现编码处理的高压缩比率。

[0004] 在一些应用中，可能捕获多于一个的音频信号，并且特定地，可以使用两个麦克风来将立体声信号记录在音频系统中。例如，立体声记录通常可以在音频和视频会议以及广播应用中使用。

[0005] 在许多多信道编码系统中，并且特定地，在许多多信道语音编码系统中，低水平的编码基于单信道的编码。在这样的系统中，多信道信号可以被转换为单声道信号，以供由编码器的较低层来进行编码。该单声道信号的生成称为下混合，这样的下混合可以与描述相对于单声道信号的立体声信号的各方面的参数相关联。特定地，下混合可以生成信道间时间差 (ITD) 信息，其表征在左右信道之间的定时差。例如，如果两个麦克风彼此远离，则来自比另一个更接近一个麦克风的扬声器的信号将以相对于第一个的延迟到达后一个麦克风。该 ITD 可以被确定，并且可以在解码器中被使用，以从单声道信号中再现立体声信号。因为已经发现 ITD 是对用于低于大约 1kHz 的频率的立体声位置 (stereo location) 的主要感知影响，所以 ITD 可以显著地改善再现的立体声立体感的质量。因此，还估计 ITD 是必需的。

[0006] 传统地，通过将立体声信号相加在一起来生成单声道信号。然后，单声道信号被编码并且与 ITD 一起被传送到解码器。

[0007] 例如，欧洲电信标准协会在它们的技术规范 ETSI TS126290 “扩展的自适应多速率 - 宽带 (AMR-WB+) 编解码器 ;转码功能”中已经定义了立体声信号下混合，其中，单声道信号如下简单地被确定为左右信道的平均值 ：

[0008] $x_{ML}(n) = 0.5(x_{LL}(n)+x_{RL}(n))$

[0009] 其中，$x_{ML}(n)$ 表示单声道信号的第 n 个样本，$x_{LL}(n)$ 表示左信道信号的第 n 个样本，并且 $x_{RL}(n)$ 表示右信道信号的第 n 个样本。

[0010] 在 2004 年 10 月 5 日至 8 日意大利那不勒斯的第七届国际数字音效会议会议录的第 163-168 页，H.Purnhagen 的“MPEG-4 中的低复杂度参数立体声编码 (Low Complexity Parametric Stereo Coding in MPEG-4)”中提供了下混合的另一个示例。在该文献中，描

述了一种下混合方法，该方法使用关于信道间的强度差（IID）的所获得的信息基于逐个频带的频率来获得作为进入的信道的加权和的输出单声道信号。特定地：

[0011] $M[k,i] = g_{lL}[k,i]+g_{rR}[k,i]$

[0012] 其中，M]k，i]表示单声道信号的第 k 个频率槽（frequency bin）的第 i 个样本，L[k，i]表示左信道信号的第 k 个频率槽的第 i 个样本，并且 R[k，i]表示右信道信号的第 k 个频率槽的第 i 个样本，g_l 是左信道权重，并且 g_r 是右信道权重。

[0013] 这样的方法的特征在于，它们引起具有高混响时间的单声道信号，或者具有高的复杂度和／或延迟的单声道信号。例如，下混合的 AMR-WB+ 方法提供了其混响时间大约是房间的混响时间加两个麦克风之间的行程时间（flight time）的输出。在 Purnhagen 中提供的下混合具有高复杂度，并且施加了由于频率分析和重建而导致的延迟。

[0014] 然而，许多单声道编码器提供了用于具有低混响时间的信号的最佳结果。例如，低比特速率 CELP 语音编码器和采用基于脉冲的激励来表示语音和音频信号的其他编码器，在存在有短混响时间的信号时执行得最佳。因此，编码器的性能和得到的编码的信号的质量往往并不是最好的。

[0015] 因此，改善的系统将是有利的，并且具体地，允许提高的灵活性、便利的实施方式、改善的编码质量、改善的编码效率、降低的延迟和／或改善的性能的系统将是有利的。

发明内容

[0016] 因此，本发明设法单个地或者以任何组合优选地减轻、缓和或者消除上述缺点的一个或多个。

[0017] 根据本发明的一方面，提供了一种用于编码多信道音频信号的设备，所述设备包括：接收机，所述接收机用于接收所述多信道音频信号，所述多信道音频信号至少包括来自第一麦克风的第一音频信号和来自第二麦克风的第二音频信号；时间差单元，所述时间差单元用于确定在所述第一音频信号和所述第二音频信号之间的间隔时间差（intertime difference）；延迟单元，所述延迟单元用于响应于所述间隔时间差通过延迟所述第一音频信号和所述第二音频信号中的至少一个，从所述多信道音频信号中生成补偿的多信道音频信号；单声道单元，所述单声道单元用于通过合并所述补偿的多信道音频信号的信道来生成单声道信号；以及单声道信号编码器，所述单声道信号编码器用于编码所述单声道信号。

[0018] 本发明可以提供改善的多信道音频信号的编码。具体地，可以在许多实施例中实现针对给定的数据速率的改善的质量。本发明可以通过减少单声道下混合信号的混响时间来提供来自立体声信号的单声道下混合信号的改善的单声道编码。延迟单元可以根据哪个麦克风最接近（主）音频源来延迟第一音频信号或第二音频信号。间隔时间差可以是在从同一音频源始发的第一音频信号和第二音频信号的对应的音频分量之间的时间差的指示。用于生成单声道信号的单元可以被配置成将与第一音频信号和第二音频信号相对应的合并的多信道音频信号的两个信道相加。在一些实施例中，所述加法可以是加权加法。

[0019] 根据本发明的可选特征，时间差单元被配置成确定针对多个时间补偿的在第一音频信号和第二音频信号之间的互相关性，并且响应于该互相关性来确定间隔时间差。

[0020] 该特征可以允许改善的间隔时间差的确定。该特征可以改善编码的音频信号的质量，和／或可以便利实现和／或降低复杂度。具体地，该特征可以允许根据单声道信号和间

隔时间差渲染的立体声信号的改善的立体声感知。互相关可以指示间隔时间差等于独立的互相关的时间补偿的概率。

[0021] 根据本发明的另一个方面，提供了一种编码多信道音频信号的方法，所述方法包括：接收所述多信道音频信号，所述多信道音频信号至少包括来自第一麦克风的第一音频信号和来自第二麦克风的第二音频信号；确定在所述第一音频信号和所述第二音频信号之间的间隔时间差；响应于所述间隔时间差通过延迟所述第一音频信号和所述第二音频信号的至少一个，从所述多信道音频信号中生成补偿的多信道音频信号；通过合并所述补偿的多信道音频信号的信道来生成单声道信号；以及在单声道信号编码器中编码所述单声道信号。

[0022] 从下述的实施例（多个）中，本发明的这些和其他方面、特征和优点将是显而易见的，并且参考下述的实施例（多个）来阐明本发明的这些和其他方面、特征和优点。

附图说明

[0023] 将参考附图仅通过示例的方式来描述本发明的实施例，其中：

[0024] 图 1 图示了根据本发明的一些实施例的用于编码多信道音频信号的设备的示例；

[0025] 图 2 图示了根据本发明的一些实施例的用于估计间隔时间差的处理单元的示例；

[0026] 图 3 图示了根据本发明的一些实施例的白化处理器的示例；

[0027] 图 4 图示了根据本发明的一些实施例的用于网格状态机的状态更新的示例；以及

[0028] 图 5 图示了根据本发明的一些实施例的用于编码多信道音频信号的方法的示例。

具体实施方式

[0029] 下面的说明集中于本发明的实施例，本发明适用于使用单声道编码器来编码多信道音频信号，并且具体地适用于使用单声道 CELP 编码器来编码立体声语音信号。

[0030] 图 1 图示了根据本发明的一些实施例的用于编码多信道音频信号的设备。在特定的示例中，将立体声语音信号下混合为单声道信号，并且使用单声道编码器来编码立体声语音信号。

[0031] 该设备包括两个麦克风 101、103，该两个麦克风 101、103 从该两个麦克风所位于的音频环境捕获音频信号。在该示例中，两个麦克风用于记录在房间中的语音信号，并且以高达 3 米的核距 (internaldistance) 来设置。在特定应用中，例如，麦克风 101、103 可以记录来自在房间中的多个人的语音信号，并且使用两个麦克风可以提供房间的更好的音频覆盖。

[0032] 将麦克风 101、103 耦合到帧处理器 105，该帧处理器 105 从第一和第二麦克风 101、103 分别接收第一和第二信号。帧处理器将信号分成顺序帧。在特定的示例中，采样频率是 16k 个样本 / 秒，并且帧的持续时间是 20 毫秒，引起每个帧包括 320 个样本。应当注意，因为该帧可能是与用于语音编码的帧相同的帧，或者可能例如在老的语音样本上执行帧处理，所以帧处理无需引起到语音路径的额外的延迟。

[0033] 将帧处理器 105 耦合到 ITD 处理器 107，该 ITD 处理器 107 被配置成确定在第一音频信号和第二音频信号之间的间隔时间差。间隔时间差是在一个信道中的信号相对于在另一个信道中的信号的延迟的指示。在该示例中，根据哪个信道相对于另一个信道被延迟，间

隔时间差可以是正的或者负的。延迟通常由于在主语音源（即当前发声的扬声器）和麦克风 101、103 之间的差而发生。

[0034] 还将 ITD 处理器 107 耦合到两个延迟装置 109、111。第一延迟装置 109 被配置成将延迟引入第一音频信道，并且第二延迟装置 109 被配置成将延迟引入第二音频信道。引入的延迟量取决于估计的间隔时间差。而且，在特定的示例中，在任何给定的时间仅使用延迟装置中的一个。因此，根据估计的间隔时间差的符号，将延迟量引入到第一音频信号或者第二音频信号。将延迟量特定地设置为尽可能地接近估计的间隔时间差。因此，在延迟装置 109、111 的输出处的音频信号被紧密地时间对齐，并且将特定地具有通常接近 0 的间隔时间差。

[0035] 将延迟装置 109、111 耦合到合并器 113，合并器 113 通过合并补偿的多信道音频信号的信道，并且特定地通过合并来自延迟装置 109、111 的两个输出信号来生成单声道信号。在该示例中，合并器 113 是简单的加法单元，其将两个信号相加在一起。而且，在合并之前，通过因子 0.5 来调节信号，以便于保持单声道信号的幅度类似于独立信号的幅度。

[0036] 因此，合并器 113 的输出是单声道信号，该单声道信号是两个捕获的信号的下混合。而且，由于间隔时间差的减少和延迟，所生成的单声道信号已经显著地降低了混响。

[0037] 将合并器 113 耦合到单声道编码器 115，该单声道编码器 115 执行单声道信号的单声道编码以生成编码的数据。在特定的示例中，单声道编码器是根据由国际电信联盟 (ITU) 标准化的嵌入式可变比特速率编解码器 (EV-VBR) 的码激励线性预测 (CELP) 编码器。

[0038] 已知 CELP 编码器提供极为有效的编码，并且特定地，即使对于低数据速率也提供良好的语音质量。然而，CELP 编码器对于具有高混响时间的信号往往也不运行，并且因此不适合于传统生成的单声道下混合的编码。然而，由于延迟补偿和得到的降低的混响，CELP 单声道编码器可以在图 1 的设备中进行使用，以提供语音下混合单声道信号的非常有效的编码。应当认识到，这些优点特别适合于 CELP 单声道编码器，但是不限于此，并且可以适用于许多其他编码器。

[0039] 将单声道编码器 115 耦合到输出复用器 117，还将该输出复用器 117 耦合到 ITD 处理器 107。在该示例中，输出复用器 117 将来自单声道编码器 115 的编码数据和来自 ITD 处理器 117 的表示间隔时间差的数据复用为单个输出比特流。将间隔时间差包括在比特流中可以辅助解码器从解码自编码数据的单声道信号中再现立体声信号。

[0040] 因此，所述系统提供了改善的性能，并且可以具体地针对给定的数据速率来提供改善的音频质量。具体地，诸如 CELP 编码器的单声道编码器的改善的使用可以引起显著改善的质量。而且，所述功能易于实现，并且具有相对低的资源需要。

[0041] 下文中，将参考图 2 来描述由 ITD 处理器 107 执行的间隔时间差估计。

[0042] 由 ITD 处理器 107 使用的算法通过合并在第一和第二音频信号之间的互相关的连续观察来确定针对在信道之间的不同的可能时间补偿的间隔时间差的估计。在抽取的 LPC 剩余域中执行相关 (correlation)，以便于提供更明确的相关性，便利实现并且减少计算需求。在该示例中，处理互相关以得出与在 -12ms 和 +12ms(± ～ 4 米）之间的每个可能的延迟相关联的概率，并且然后，使用修改的类维特比 (Viterbi-like) 算法来累加该概率。结果是具有内置 (in-built) 滞后的间隔时间差的估计。

[0043] ITD 处理器 107 包括抽取处理器 201，该抽取处理器 201 从帧处理器 105 接收两个

信道的样本的帧。抽取处理器 201 首先执行在抽取之前的低通滤波。在特定的示例中，低通滤波器具有大约 2kHz 的带宽，并且使用 4 的抽取因子用于样本频率的 16k 样本 / 秒的信号，得到 4k 样本 / 秒的抽取。滤波和抽取的效果部分地减少了所处理的样本的数目，由此减少了计算需求。然而，另外，该方法允许间隔时间差估计集中于较低频率，其中，间隔时间差的感知重要性最显著。因此，滤波和抽取不仅减少了计算负担，还提供了保证间隔时间差估计与最敏感的频率相关的协同效应。

[0044] 将抽取处理器 201 耦合到白化处理器 203，该白化处理器 203 被配置成在相关之前将频谱白化算法应用于第一和第二音频信号。在有声语音或者音调语音的情况下，频谱白化导致两个信号的时域信号更接近类似一组冲击，由此允许后续的相关引起更明确的互相关值，并且特定地引起更窄的相关峰值（冲击的频率响应与平坦或者白频谱相对应，并且相反地，白频谱的时域表示是冲击）。

[0045] 在特定的示例中，频谱白化包括：计算用于第一和第二音频信号的线性预测系数，并且响应于该线性预测系数来对第一和第二音频信号进行滤波。

[0046] 在图 3 中示出了白化处理器 203 的元件。特定地，将来自抽取处理器 201 的信号馈送到 LPC 处理器 301、303，该 LPC 处理器 301、303 确定用于两个信号的线性预测滤波器的线性预测系数 (LPC)。应当认识到，用于确定 LPC 的不同算法对于本领域的技术人员来说是公知的，并且在不脱离本发明的情况下可以使用任何适当的算法。

[0047] 在该示例中，将两个音频信号馈送到两个滤波器 305、307，将两个滤波器 305、307 耦合到 LPC 处理器 301、303。确定该两个滤波器，使得它们是通过 LPC 处理器 301、303 确定的线性预测滤波器的逆滤波器。特定地，LPC 处理器 301、303 确定用于线性预测滤波器的逆滤波器的系数，并且将该两个滤波器的系数设置为这些值。

[0048] 两个逆滤波器 305、307 的输出在有声语音的情况下类似多组冲击序列，并且由此允许执行比在语音域中可能的明显更为准确的互相关。

[0049] 将白化处理器 203 耦合到相关器 205，该相关器 205 被配置成确定针对多个时间补偿的在两个滤波器 305、307 的输出信号之间的互相关。

[0050] 特定地，相关器可以确定值：

[0051] $$c^t = \sum_N x_n \cdot y_{n-t}$$

[0052] 其中，t 是时间补偿，x 和 y 是两个信号的样本，并且 N 表示在特定帧中的样本。

[0053] 针对一组可能的时间补偿来执行相关。在特定的示例中，针对与 ±12 毫秒的最大时间补偿相对应的总共 97 个时间补偿来执行相关。然而，应当认识到，可以在其他实施例中使用其他组的时间补偿。

[0054] 因此，相关器生成 97 个互相关值，其中，每个互相关与在两个信道之间的特定时间补偿相对应，并且因此与可能的间隔时间差相对应。互相关的值与针对特定的时间补偿的两个信号如何紧密匹配的指示相对应。因此，对于高互相关值，信号紧密地匹配，并且因此存在时间补偿是准确的间隔时间差估计的高概率。相反地，对于低互相关值，信号不紧密匹配，并且因此存在时间补偿是准确的间隔时间差估计的低概率。因此，对于每个帧，相关器 205 生成 97 个互相关值，其中，每个值是对应的时间补偿是正确的间隔时间差的概率的指示。

[0055] 在该示例中，相关器 205 被配置成在互相关之前在第一和第二音频信号上执行开窗。特定地，以 20ms 的窗口来开窗该两个信号的每个帧样本块，该 20ms 的窗口包括 14ms 的矩形中心部分和在每端处的 3ms 的两个汉恩 (Hann) 部分。该开窗可以改善准确度，并且减少在相关窗口边缘处的边界效应的影响。

[0056] 而且，在该示例中，归一化互相关。归一化特定地保证可以实现的最大互相关值（即当所述两个信号相同时）具有单一值。归一化支持相对独立于输入信号的信号电平和所测试的相关时间补偿的互相关值，由此提供更准确的概率指示。具体地，允许用于一系列帧的改善的比较和处理。

[0057] 在简单实施例中，可以直接地评估相关器 205 的输出，并且可以将用于当前帧的间隔时间差设置为具有如由互相关值指示的最高概率的值。然而，这样的方法往往在语音信号从有声到无声到静音进行波动时提供不太可靠的输出，并且在所述示例中，相关器向状态处理器 207 进行馈送，状态处理器 207 处理用于多个状态的相关值，以提供更准确的间隔时间差估计。

[0058] 在该示例中，相关值被用作对在状态处理器 207 中实现的维特比算法度量累加器的更新步骤。

[0059] 因此，状态处理器 207 特定地实现度量累加器，其具有与时间补偿相对应的多个状态。每个状态因此表示时间补偿，并且具有相关联的累加度量值。

[0060] 因此，以度量累加器形式的基于维特比的网格状态机存储用于针对其已经计算了相关值的时间补偿的每一个的度量值（即在特定示例中的 97 个状态 / 时间补偿）。每个状态 / 时间补偿特定地与概率度量相关联，该概率度量指示间隔时间差与该状态的时间补偿相对应的概率。

[0061] 在每个帧中重新计算针对所有时间补偿的概率度量，以考虑已经针对当前帧确定的相关值。特定地，根据互相关来针对状态 / 时间补偿计算路径度量。在特定的示例中，通过应用公式 $\log(0.5+p_i)$ 来将互相关转换到对数域中，其中，p_i 是第 i 个相关值（其由于归一化处理而在 0 和 1 之间，并且对应于与关联时间补偿的间隔时间差相对应的概率）。

[0062] 在该示例中，从该时间补偿的前一概率度量和针对当前帧计算的补偿的相关值来确定对于给定的概率度量的贡献。另外，对应于间隔时间差随值改变的情况（即，使得最可能的状态从作为一个时间补偿的最可能的状态改变为作为另一个时间补偿的最可能的状态），从与相邻时间补偿相关联的相关值做出贡献。

[0063] 基本上低于用于根据同一状态的路径的路径度量来加权用于根据与邻近的间隔时间差值相对应的相邻状态的路径的路径度量。特定地，实验已经示出了，对于相邻相关值被加权了比用于同一状态的互相关高了至少多 5 倍，已经发现了特别有利的性能。在特定示例中，通过因子 0.009 来加权邻近状态路径度量，并且通过因子 0.982 来加权同一状态的路径度量。

[0064] 图 4 图示了用于网格状态机的帧 t 的度量更新的示例。在特定的示例中，根据来自包括时间 t-1 时的状态 s_n 和时间 t-1 时的邻近状态 s_{n-1} 和 s_{n+1} 的先前的状态的子集的路径的路径度量来计算在时间 t 时状态 s_n 的状态概率度量。特定地，通过下式来给出针对状态 s_n 的状态概率度量：

[0065] $$S_n^t = S_n^{t-1} + P_n^t + P_{n-1}^t + P_{n+1}^t$$

[0066] 其中，P_x^t 是在帧 t 中从状态 x 到状态 n 的计算的加权路径度量。

[0067] 在该示例中，通过从所有的状态概率度量中减去最低的状态概率度量来在每个帧中修改概率度量。这减轻了来自持续增加的状态概率度量的溢出问题。

[0068] 在该示例中，对于包括补偿本身和邻近补偿的补偿的子集，仅包括对给定的时间补偿度量的贡献。然而，应当认识到，在其他实施例中，可以考虑时间补偿的其他子集。

[0069] 在该示例中，在每个帧中更新用于网格状态机的状态度量。然而，与传统的维特比算法相反，状态处理器 207 不选择每个状态的优选路径，而是计算给定状态的状态概率度量作为来自进入该状态的所有路径的合并的贡献。而且，状态处理器 207 不通过网格执行追溯来确定仍然存在的路径。相反地，在该示例中，可以简单地选择当前的间隔时间差估计作为与当前具有最高的状态概率度量的状态相对应的时间补偿。因此，在状态机中不发生延迟。而且，由于概率状态度量取决于先前的值（和其他状态），因此固有地实现滞后。

[0070] 特定地，将状态处理器 207 耦合到 ITD 处理器 209，该 ITD 处理器 209 从与具有最高状态概率度量的状态相关联的时间补偿中来确定间隔时间差。特定地，可以将间隔时间差直接地设置为等于具有最高状态概率度量的状态的时间补偿。

[0071] 将 ITD 处理器 209 耦合到延迟处理器 211，该延迟处理器 211 确定要应用于延迟装置 109、111 的延迟。首先，延迟处理器 211 通过在抽取处理器 201 中应用的抽取因子来补偿间隔时间差。在简单的实施例中，可以将估计的间隔时间差给出作为抽取的样本数（例如以与 250μs 分辨率相对应的 4kHz），并且这可以通过使其乘以抽取因子来被转换为非抽取的样本数（例如通过将其乘以因子 4 来转换为 16kHz 的样本）。

[0072] 在该示例中，延迟处理器 211 设置用于两个延迟装置 109、111 的值。特定地，根据间隔时间差的符号，将延迟的一个设置为 0，并且将另一个延迟装置设置为计算的非抽取的样本数。

[0073] 用于计算间隔时间差的所述方法提供了改善的编码的信号的质量，并且具体地，在编码之前提供了降低的单声道信号的混响，由此改善了 CELP 单声道编码器 115 的操作和性能。

[0074] 已经执行了特定的测试，其中，以不同的配置在具有一对麦克风的会议房间中记录了三个立体声测试信号。在第一配置中，相距 1m 地放置麦克风，并且两个男性交谈者在两个麦克风的每一个之外的轴上落座，并且记录测试会话。在第二配置中，相距 3m 地放置两个麦克风，并且该男性交谈者再一次在两个麦克风的每一个之外的轴上落座。在最后一个配置中，相距 2m 地放置麦克风，并且两个交谈者在麦克风的轴的宽侧，但是在面对两个麦克风的每一个的轴的相对侧上。在所有这些情况下，算法良好地跟踪延迟，并且当通过用于 ITU-T EV-VBR编解码器的基线算法来编码得到的单声道信号时，在每种情况中都观察到在 SEGSNR 和 WSEGSNR 中的大约 0.3dB 的增益。

[0075] 在一些实施例中，通过改变样本数来简单地实现从一个延迟到另一个的转变，适当的信号由延迟装置 109、111 延迟。然而，在一些实施例中，可以包括用于执行从一个延迟到另一个延迟的平滑转变。

[0076] 特定地，设备可以被配置成通过生成第一信号和第二信号来从第一延迟转变成第二延迟，该第一信号在转变之前由延迟装置来进行延迟，该第二信号在转变之后由延迟装置来进行延迟。然后，合并第一和第二信号以生成合并的信号，该合并的信号包括来自在转

变之前的信号和在转变之后的信号二者的贡献。逐渐地改变来自两个信号的贡献，使得贡献初始地主要或者排他地来自第一信号，而在转变结束时，贡献主要或者排他地来自第二信号。

[0077] 因此，设备可以在延迟转变期间合成与初始和最后延迟相对应的两个信号。可以通过加权加法来合并该两个信号，该加权加法诸如：

[0078] $s = a \cdot s_1 + b \cdot s_2$

[0079] 其中，s_1 和 s_2 表示第一和第二信号，并且 a 和 b 是在转变间隔（其可以特定地等于单个帧）期间修改的权重。特定地，值可以初始地被设置为 $a = 1$ 并且 $b = 0$，并且最后的值可以被设置为 $a = 0$ 和 $b = 1$。在这些值之间的转变可以根据任何适当的函数来执行，并且可以特定地在转变期间保持关系 $a+b = 1$。

[0080] 因此，在这样的实施例中，通过合并用于两个延迟的信号并且在时域中从一个逐渐地转变为另一个来实现在不同的延迟之间的平滑转变。

[0081] 在特定的示例中，应用 20ms 的半汉恩重叠相加窗口以保证尽可能地感觉不到从一个延迟到下一个延迟的转变。

[0082] 图 5 图示了根据本发明的一些实施例的编码多信道音频信号的方法。

[0083] 该方法在步骤 501 中发起，其中，接收多信道音频信号，该多信道音频信号至少包括来自第一麦克风的第一音频信号和来自第二麦克风的第二音频信号。

[0084] 步骤 501 后是步骤 503，其中，确定在第一音频信号和第二音频信号之间的间隔时间差。

[0085] 步骤 503 后是步骤 505，其中，响应于间隔时间差通过延迟第一和第二立体声信号中的至少一个，从多信道音频信号中生成补偿的多信道音频信号。

[0086] 步骤 505 后是步骤 507，其中，通过合并补偿的多信道音频信号的信道来生成单声道信号。

[0087] 步骤 507 后是步骤 509，其中，通过单声道信号编码器来编码单声道信号。

[0088] 应当认识到，为了清晰，以上的描述已经参考不同的功能单元和处理器描述了本发明的实施例。然而，显然，在不偏离本发明的情况下，可以使用在不同的功能单元或者处理器之间的功能的任何适当分布。例如，图示为由分立的处理器或者控制器执行的功能可以由同一处理器或者控制器来执行。因此，对于特定功能单元的参考仅仅被看作对于用于提供所述功能的适当组件的参考，而不是指示严格的逻辑或者物理结构或者组织。

[0089] 可以以任何适当的形式来实现本发明，任何适当形式包括硬件、软件、固件或者其任何组合。本发明可以可选地被至少部分地实现为在一个或多个数据处理器和 / 或数字信号处理器上运行的计算机软件。可以以任何适当的方式来物理地、功能上地和逻辑地实现本发明的实施例的元件和部件。事实上，可以在单个单元中、在多个单元中或者作为其他功能单元的一部分来实现功能。如此一来，本发明可以在单个单元中被实现，或者可以物理地或功能上地分布在不同的单元和处理器之间。

[0090] 虽然已经结合一些实施例描述了本发明，但是不希望限于在此阐述的特定形式。相反地，本发明的范围仅由所附权利要求来限定。另外，虽然结合具体实施例描述了特征，但是本领域内的技术人员应当认识到，可以根据本发明来合并所述实施例的各种特征。在权利要求中，术语包括不排除存在其他元件或者步骤。

[0091] 而且，虽然独立地列出，但是多个单元、组件、元件或者方法步骤可以由例如单个单元或者处理器来实现。另外，虽然可以在不同的权利要求中包括独立的特征，但是这些能够被有利地合并，并且包括在不同的权利要求中不意指特征的组合不是可行的和 / 或有利的。而且，将特征包括在一类权利要求中不意指限于该类，而是指示该特征适当地等同适用于其他权利要求类别。而且，在权利要求中的特征的顺序不意指该特征必须以其工作的任何特定顺序，并且特定地，方法权利要求中的独立步骤的顺序不意指必须以该顺序来执行步骤。相反地，可以以任何适当的顺序来执行步骤。

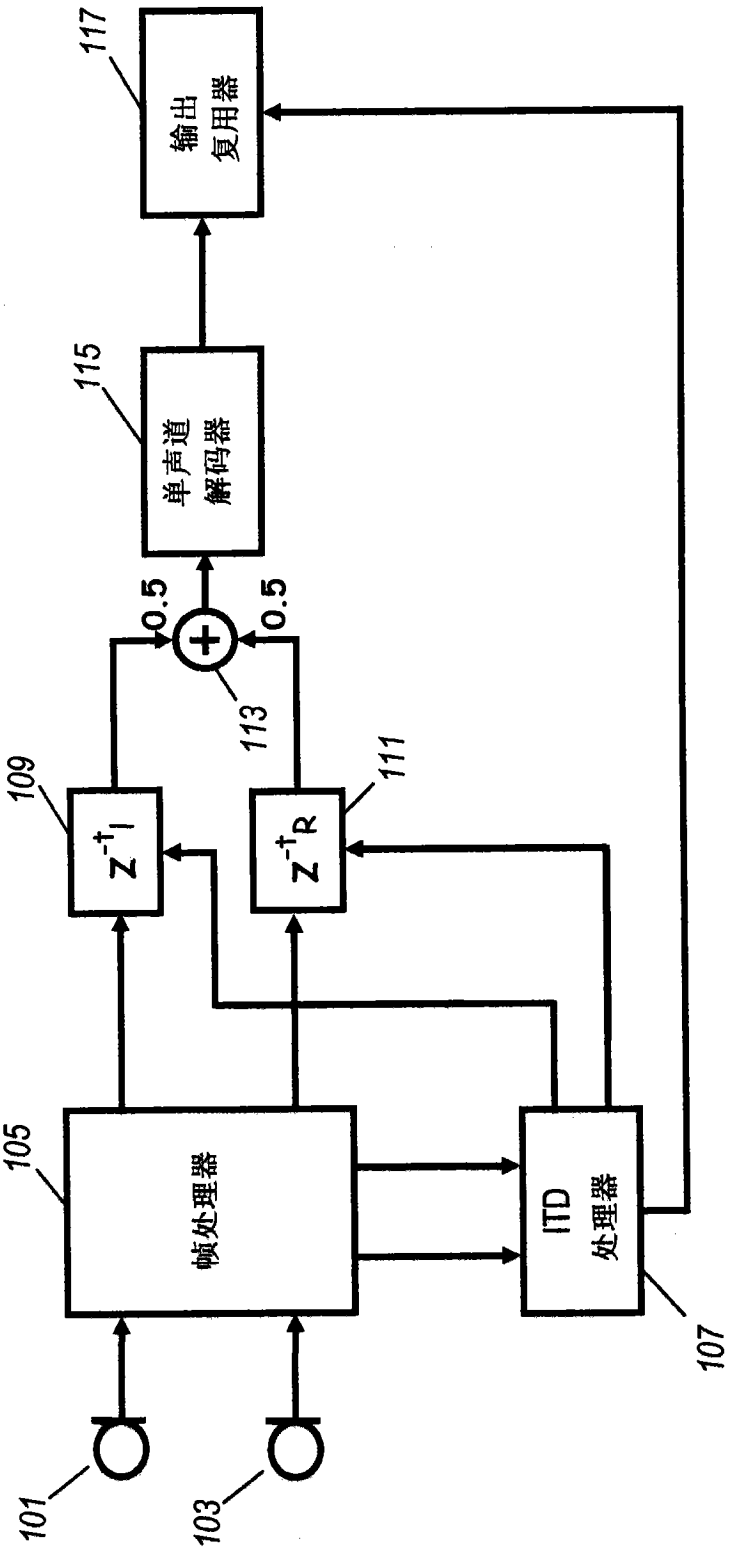

117
输出
复用器
115
单声道
解码器
0.5
0.5
113
109
111
z^{-t_I}
z^{-t_R}
105
帧处理器
ITD
处理器
107
101
103

图 1

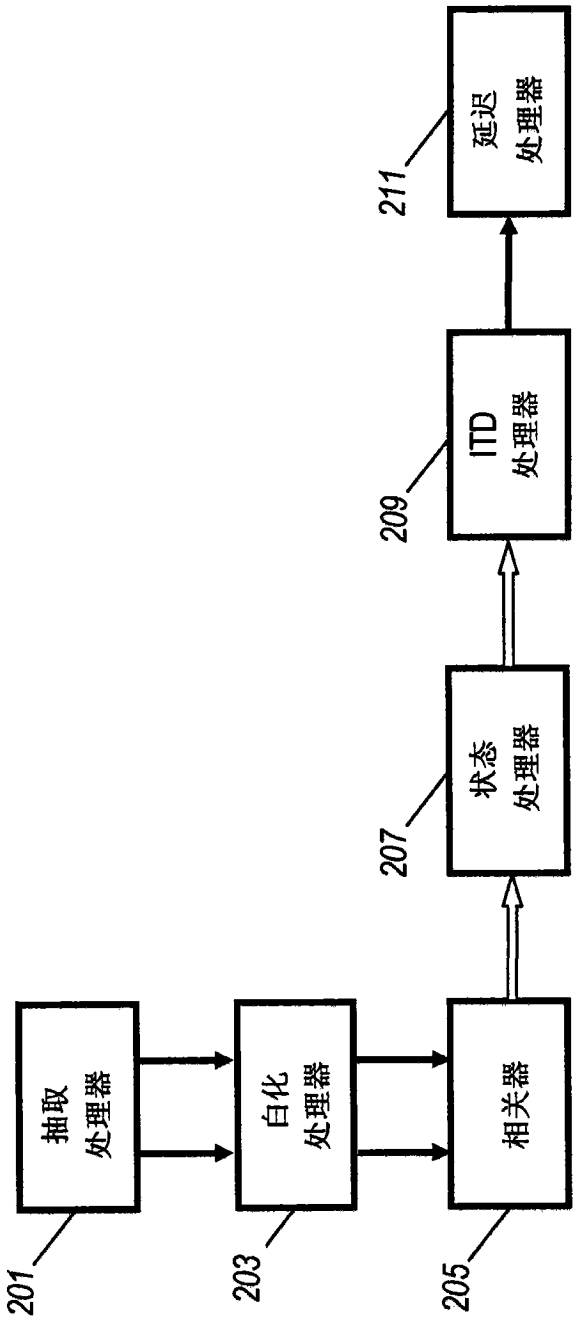


图 2

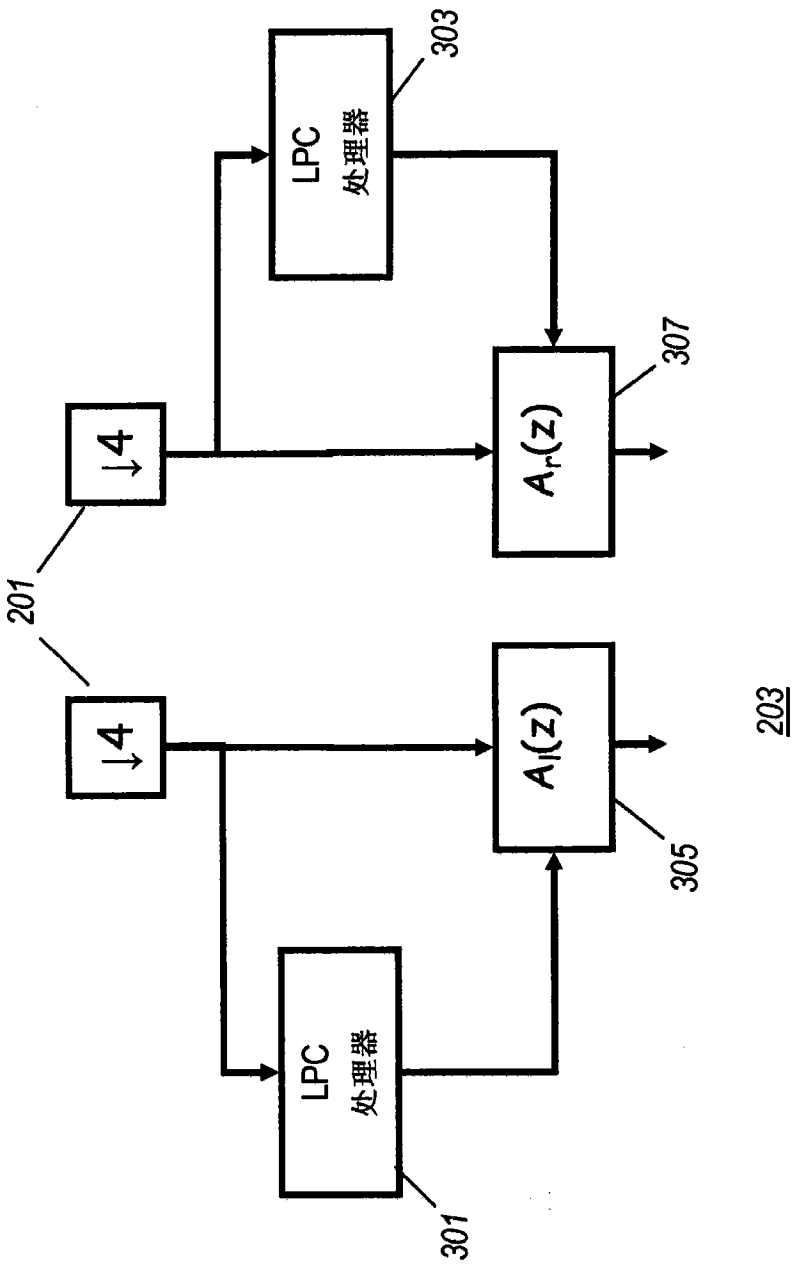

LPC
处理器
303
↓4
$A_r(z)$
307
201
↓4
$A_l(z)$
305
203
LPC
处理器
301

图 3

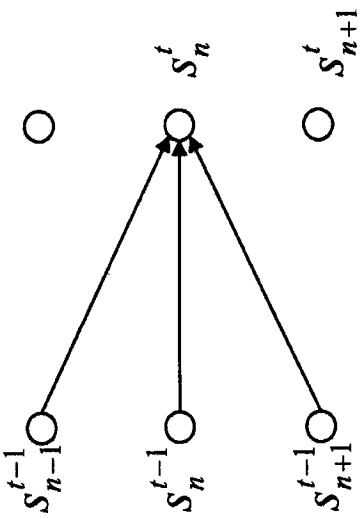

S_n^t
S_{n+1}^t
S_{n-1}^{t-1}
S_n^{t-1}
S_{n+1}^{t-1}

图 4

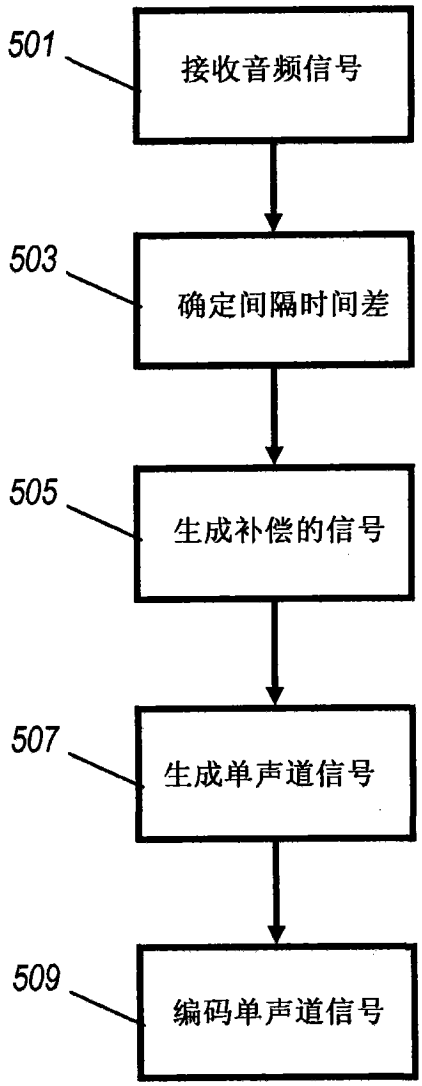


图 5