

(12) 发明专利申请

(10) 申请公布号 CN 101996636 A

(43) 申请公布日 2011.03.30

(21) 申请号 201010536835.0

(22) 申请日 2006.04.05

(30) 优先权数据

11/142,605 2005.05.31 US

(62) 分案原申请数据

200680019541.2 2006.04.05

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 T · 王 K · 科什达

H · A · 海莉尔 X · 孙

W-G · 陈

(74) 专利代理机构 上海专利商标事务所有限公司 31100

代理人 陈斌

(51) Int. Cl.

G10L 19/00 (2006.01)

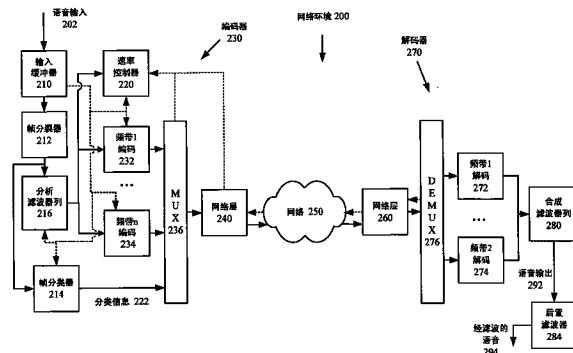
权利要求书 2 页 说明书 20 页 附图 10 页

(54) 发明名称

带多级码本和冗余编码的子带语音编解码器

(57) 摘要

描述了涉及音频信息编码和解码的技术和工具。例如，用于解码当前帧的冗余编码信息包括仅与在前帧一部分相关联的信号历史信息。作为另一个示例，用于解码编码单元的冗余编码信息包括有关一码本级的各参数，该码本级仅当在前编码单元不可用时用于解码该当前编码单元。作为一个示例，每个编码音频单元包括一字段，该字段用以指示编码单元是否包括表示一段音频信号的主编码信息，以及编码单元是否包括用于解码主编码信息的冗余编码信息。



1. 一种方法,包括:

在音频处理工具处,处理有关音频信号的比特流,其中所述比特流包括:

用于当前编码单元的主编码信息,所述主编码信息参考要被用于解码当前编码单元的一段在前编码单元;以及

用于解码所述当前编码单元的冗余编码信息,所述冗余编码信息包括仅在所述在前编码单元不可用的情况下才在解码所述当前编码单元时使用的用于一个或多个额外码本级的一个或多个参数;以及

输出结果。

2. 如权利要求1所述的方法,其特征在于,用于所述当前编码单元的主编码信息包括残留信号参数,所述残留信号参数表示在用于所述当前编码单元的重建和用于所述当前编码单元的预测之间的一个或多个差异。

3. 如权利要求1所述的方法,其特征在于:

所述音频处理工具是音频编码器;以及

处理所述比特流包括生成可任选的冗余编码信息,其中生成所述可任选的冗余编码信息包括在假设没有激励信息用于所述在前编码单元的闭环编码器搜索中确定用于所述一个或多个额外码本级的所述一个或多个参数。

4. 如权利要求1所述的方法,其特征在于:

所述音频处理工具是语音解码器;以及

如果所述在前编码单元对所述解码器不可用,则用于所述码本的所述一个或多个参数就被所述解码器在解码所述当前编码单元时使用;以及

如果所述在前编码单元对所述解码器可用,则用于所述码本的所述一个或多个参数就不被所述解码器在解码所述当前编码单元时使用。

5. 如权利要求1所述的方法,其特征在于,所述码本是跟随在自适应码本级之后的固定码本级中的固定码本,并且其中用于一个或多个额外码本级的所述一个或多个参数包括码本索引和增益。

6. 如权利要求5所述的方法,其特征在于,用于所述自适应码本级中的自适应码本的一个或多个参数表示参考用于所述在前编码单元的激励历史的所述当前编码单元的激励信号,但是用于所述固定码本的一个或多个参数表示不参考所述激励历史的所述激励信号。

7. 如权利要求1所述的方法,其特征在于:

所述音频处理工具是音频解码器;以及

处理比特流包括:

如果所述在前编码单元不可用,就在解码所述当前编码单元时使用所述主编码信息的至少一部分以及用于所述一个或多个额外码本级的所述一个或多个参数;以及

如果在前编码单元可用,就在解码所述当前编码单元时使用所述主编码信息,但不使用用于所述一个或多个额外码本级的所述一个或多个参数。

8. 一种被配置来处理音频信号的比特流并输出结果的音频处理设备,其中所述比特流包括:

用于当前编码单元的主编码信息,所述主编码信息参考要被用于解码当前编码单元的

一段在前编码单元；以及

用于解码所述当前编码单元的冗余编码信息，所述冗余编码信息包括仅在所述在前编码单元不可用的情况下才在解码所述当前编码单元时使用的用于一个或多个额外码本级的一个或多个参数。

9. 如权利要求 8 所述的音频处理设备，其特征在于，用于所述当前编码单元的主编码信息包括残留信号参数，所述残留信号参数表示在用于所述当前编码单元的重建和用于所述当前编码单元的预测之间的一个或多个差异。

10. 如权利要求 8 所述的音频处理设备，其特征在于：

所述音频处理设备是音频编码器；以及

处理所述比特流包括生成可任选的冗余编码信息，其中生成所述可任选的冗余编码信息包括在假设没有激励信息用于所述在前编码单元的闭环编码器搜索中确定用于所述一个或多个额外码本级的所述一个或多个参数。

11. 如权利要求 8 所述的音频处理设备，其特征在于：

所述音频处理设备是语音解码器；以及

如果所述在前编码单元对所述解码器不可用，则用于所述码本的所述一个或多个参数就被所述解码器在解码所述当前编码单元时使用；以及

如果所述在前编码单元对所述解码器可用，则用于所述码本的所述一个或多个参数就不被所述解码器在解码所述当前编码单元时使用。

12. 如权利要求 8 所述的音频处理设备，其特征在于，所述码本是跟随在自适应码本级之后的固定码本级中的固定码本，并且其中用于一个或多个额外码本级的所述一个或多个参数包括码本索引和增益。

13. 如权利要求 12 所述的音频处理设备，其特征在于，用于所述自适应码本级中的自适应码本的一个或多个参数表示参考用于所述在前编码单元的激励历史的所述当前编码单元的激励信号，但是用于所述固定码本的一个或多个参数表示不参考所述激励历史的所述激励信号。

14. 如权利要求 8 所述的音频处理设备，其特征在于：

所述音频处理设备是音频解码器；以及

处理比特流包括：

如果所述在前编码单元不可用，就在解码所述当前编码单元时使用所述主编码信息的至少一部分以及用于所述一个或多个额外码本级的所述一个或多个参数；以及

如果在前编码单元可用，就在解码所述当前编码单元时使用所述主编码信息，但不使用用于所述一个或多个额外码本级的所述一个或多个参数。

带多级码本和冗余编码的子带话音编解码器

[0001] 本发明专利申请是国际申请号为 PCT/US2006/012686, 国际申请日为 2006 年 04 月 05 日, 进入中国国家阶段的申请号为 200680019541.2, 名称为“带多级码本和冗余编码的子带话音编解码器”的发明专利申请的分案申请。

技术领域

[0002] 描述的工具和技术涉及音频编解码器, 尤其涉及子带编码、码本、和 / 或冗余编码。

背景技术

[0003] 随着数字无线电话网络的出现, 经因特网的流媒体音频、以及因特网电话技术、数字传输和语音传送已经变得很平常了。工程师利用多种技术在保证质量的同时来有效处理语音。了解这些技术有助于理解音频信息是如何在计算机中被表示和处理的。

[0004] I. 计算机内的音频信息的表示

[0005] 计算机将音频信息作为一系列表示音频的数字处理。单个数字能够表示一个音频样本, 它是一特定时刻处的幅值。许多因素会影响音频的质量, 包括样本深度和采样速率。

[0006] 样本深度 (或精确度) 显示了用来表示样本的数字的范围。因为能够表示更为细微的幅度变化, 所以通常每个样本的可能值越多则输出质量就会越高。一个 8 位样本具有 256 个可能值, 而 16 位样本则具有 65,536 个可能值。

[0007] 采样速率 (一般被测量作为每秒钟的采样数) 也会影响质量。因为能够表示更高频率的声音, 所以采样速率越高, 质量就越高。一些常见的采样速率是 8,000、11,025、22,050、32,000、44,100、48,000 以及 96,000 采样 / 秒 (Hz)。表 1 示出了带有不同质量等级的多个音频格式, 及其相应的原始比特率成本。

[0008]

样本深度	采样速率	通道模式	原始比特率
(比特 / 样本)	(样本 / 秒)		(比特 / 秒)
8	8,000	单声道	64,000
8	11,025	单声道	88,200
16	44,100	立体声	1,411,200

[0009]

[0010] 表 1 : 不同质量的音频的比特率

[0011] 如表 1 所示, 高质量音频的成本对应高比特率。高质量音频信息消耗大量的计算机存储和传输容量。许多计算机和计算机网络缺少用来处理原始数字音频的资源。压缩 (也成为编码或译码) 通过将信息转换为较低比特率的形式来降低存储及传输音频信息的

成本。压缩可能是无损耗的（其中质量不受损害）或有损耗的（其中质量受到损害，但是从随后的无损压缩得到的比特率降低更为显著）。解压缩（也成为解码）从压缩形式中提取原始信息的重构版本。编解码器是一种编码器 / 解码器系统。

[0012] II. 语音编码器和解码器

[0013] 音频压缩的一个目标是数字化表示音频信号，从而为给定的比特数量提供最佳的信号质量。换句话说，这一目标是在给定的质量等级下用最少的比特来表示音频信号。由编码 / 传送 / 解码引起的诸如传输误差的恢复能力以及对整体延迟的限制的其他目标也可以应用于一些场景。。

[0014] 不同类型的音频信号具有不同的特征。音乐以大范围的频率和振幅为特征，并且通常包含两个或多个信道。另一方面，语音以小范围的频率和振幅为特征，并且一般在一个信道内表示。特定的编解码器以及处理技术适用于音乐和普通音频；其他编解码器及其处理技术则适用于语音。

[0015] 一类常规语音编解码器使用线性预测来实现压缩。该语音编码包含多级。该编码器找出并量化用于线性预测滤波器的系数，该滤波器用于预测各样本值作为在前样本值的线性组合。残留信号（被表示为“激励”信号）表示未被滤波器精确预测的那部分原始信号。在某些级，语音编解码器使用对浊音段（以噪音的声带振动为特征）、清音段、以及无声段使用不同的压缩技术，这是因为不同类型的语音具有不同的特征。浊音段通常呈现出高度重复的发音模式，即使是在残留域中。对于浊音段，该编码器通过将当前残留信号与在前的残留周期进行比较并根据相对于在前周期的延迟或滞后信息对当前残留信号进行编码而实现进一步的压缩。该编码器使用专门设计的码本来处理原始信号与经预测的、被编码表示之间的其他差异。

[0016] 许多语音编解码器通过一些方法在信号中使用暂时冗余度。如上所述，一种常用的方法是根据相对于在前激励周期延迟或滞后，使用基音（pitch）参数的长期预测来预测当前激励信号。使用暂时冗余度能够在质量和比特率方面大幅提高压缩效率，但会把记忆依赖性引入编解码器，即解码器依靠该信号的一个在前解码部分才能正确解码该信号的另一部分。许多有效的语音编解码器都具有明显的记忆依赖。

[0017] 尽管上面所描述的语音编解码器对许多应用而言具有良好的整体性能，但是它们仍然有若干缺陷。更具体地，当语音编解码器被用来与动态网络资源一起使用时就会遇到若干缺陷。在此场景中，编码的语音可能会由于暂时的带宽不足或其他问题而丢失。

[0018] A. 窄带和宽带编解码器

[0019] 许多标准语音编解码器被设计用于带有 8kHz 采样速率的窄带信号。虽然 8kHz 采样速率在许多情况下是足够的，在其他情况下也可以使用更高的采样速率，诸如用来表示更高的频率。

[0020] 带有至少 16kHz 采样速率的语音信号通常被称为宽带语音。虽然这些宽带编解码器正适于表示高频率语音模式，但是它们通常需要比窄带编解码器更高的比特率。这样高的的比特率在一些网络类型中或在一些网络条件下是不可行的。

[0021] B. 动态网络条件中低效的记忆依赖

[0022] 当编码语音由诸如丢失、延迟、破坏或由其他方面造成在传输中不可用而有所缺失，则语音编解码器的性能会由于对丢失信息的记忆依赖而受到损害。有关激励信号的信

息丢失妨碍了随后依赖于那些丢失信号的重建。如果丢失了在前周期，则滞后信息因为其指向了解码器不具有的信息而会变得无用。记忆依赖的另一个例子是滤波器系数插值（用来平滑在各不同合成滤波器之间的转换，尤其是针对浊音信号）。如果丢失了某一帧的滤波器系数，则用于随后帧的滤波器系数可能具有不正确的值。

[0023] 解码器使用各种技术来隐蔽由于分组丢失和其他信息丢失所导致的差错，但是这些隐蔽技术很少能够完全隐蔽这些差错。例如，解码器基于被正确解码的信息重复在前的参数或估计参数。然而，滞后信息可能非常敏感，并且现有技术无法很有效的进行隐蔽。

[0024] 在大部分情况下，解码器最终会从由于丢失信息所导致的差错中恢复。随着分组的接收和解码，参数被逐渐调整到它们正确的值。但是质量很可能劣化直到解码器能够恢复正确的内部状态。在许多最为高效的语音编解码器中，回放质量会在一延长时段（例如，长至一秒）内劣化，引起高度失真并通常将语音描绘地难以理解。恢复时间在发生例如无声帧的明显变化时更快，因为这为许多参数提供了一自然重置点。一些编解码器比较不容易出现分组丢失，因为它们移除了帧间依赖。然而，这种编解码器需要明显更高的比特率来完成与带有帧间依赖的传统 CELP 编解码器相同的话音质量

[0025] 在此给出了在计算机系统内用于表示语音信号的压缩和解压缩的重要性，因而语音的压缩和解压引起了对其的研究和标准化的行为就不足为奇了。无论现有技术和工具有怎样的优点，但它们都不具有这里描述的技术和工具的优点。

发明内容

[0026] 总的说来，详细描述涉及用于音频编解码器的各种技术和工具，尤其涉及有关子带编码、音频编解码器码本、和 / 或冗余编码的工具和技术。描述的实施例执行了一个或多个所描述的技术和工具，包括但不限于以下各项：

[0027] 一个方面，音频信号的比特流包括用于当前帧并参考了一段在前帧用以解码当前帧的主要编码信息，以及用于解码该当前帧的冗余编码信息。该冗余编码信息包括与在前帧被参考段相关联的信号历史信息。

[0028] 另一方面，音频信号的比特流包括用于当前编码单元并参考了一段在前编码单元用以解码当前编码单元的主要编码信息，以及用于解码该当前编码单元的冗余编码信息。该冗余编码信息包括用于仅当在前编码单元不可用时被用来解码当前编码单元的有关一个或多个额外码本级的一个或多个参数。

[0029] 另一个方面，比特流包括多个编码音频单元，且每个编码单元包括字段。该字段指示了编码单元是否包括表示一段音频信号的主要编码信息，以及编码单元是否包括用于解码该主要编码信息的冗余编码信息。

[0030] 在另一个方面，音频信号被解压成多个频率子带。每个子带都根据激励编码（code-excited）线性预测模型而被编码。该比特流可能包含各自表示一段音频信号的多个编码单元，其中上述多个编码单元包括表示第一多个频率子带的第一编码单元以及表示第二多个频率子带的第二编码单元，该第二多个子带与第一多个子带会因为与第一编码单元或第二编码单元有关的子带信息的下降特性而有所不同。第一子带可以根据第一编码模式来编码，且第二子带可以根据不同的第二编码模式来编码。该第一和第二编码模式可以使用不同数量的码本级。每个子带可以被分别编码。此外，实时语音编码器可以处理比特流，

包括将音频信号解压为多个频率子带并编码上述多个频率子带。处理比特流可能包括解码多个频率子带并合成多个频率子带。

[0031] 另一方面,用于音频信号的比特流包括与用于表示音频信号第一段的第一组码本级有关的参数,第一组码本级包括多个固定码本级的第一集合。多个固定码本级的第一集合可以包括多个随机固定的码本级。固定码本级可以包括脉冲码本级和随机码本级。第一组码本级可以进一步包括自适应码本级。比特流可以进一步包括与用于表示音频信号第二段的第二组码本级有关的参数,该第二组具有和第一组不同数量的码本级。第一组码本级中的码本级的数量可以基于包括音频信号第一段的一个或多个特征内的一个或多个因素中选择。第一组码本级中的码本级的数量可以基于包括编码器和解码器之间的网络传输条件的一个或多个因素中选择。该比特流可以包括用于多个固定码本级中每一个的分离的码本索引和分离的增益。利用该分离的增益有助于信号匹配,并且利用该分离的码本索引则能够简化码本搜索。

[0032] 另一方面,比特流对于使用自适应码本描述的多个可参数化的单元中的每个单元而言,包括指示自适应码本参数是否用于该单元的字段。该单元可以是多个音频信号帧的子帧。诸如实时语音编码器的音频处理工具可以处理比特流,包括确定是否在每个单元使用自适应码本参数。确定是否使用自适应码本参数可以包括确定自适应码本增益是否在阈值之上。同样,确定是否使用自适应码本参数可以包括估计该帧的一个或多个特征。此外,确定是否使用自适应码本参数可以包括估计编码器和解码器之间的一个或多个网络传输特征。该字段可以是每个浊音单元一个的一比特标记。该字段可以是音频信号的浊音帧的每个子帧一个的一比特标记,并且其他类型的各帧可能不需要包括该字段。

[0033] 各种技术和工具可以被组合或独立地使用。

[0034] 其他特征和优点从下面参考附图的不同实施例的详细描述中将变得显而易见。

附图说明

- [0035] 图 1 是一种在其中可以实现一个或多个描述的实施例的适合计算机环境的框图。
- [0036] 图 2 是结合其可以实现一个或多个描述的实施例的网络环境的框图。
- [0037] 图 3 是描绘了可以用于子带编码的有关一子带结构的一组频率响应的子带图示。
- [0038] 图 4 是一种结合其可以实现一个或多个描述的实施例的实时语音频带编码器的框图。
- [0039] 图 5 是描绘一实现中的码本参数确定的流程图。
- [0040] 图 6 是一种结合其可以实现一个或多个描述的实施例的实时语音频带解码器的框图。
- [0041] 图 7 是包括当前帧和在前帧的重新编码部分的激励信号历史的图示。
- [0042] 图 8 是描绘一实现中有关一额外随机码本级的码本参数确定的流程图。
- [0043] 图 9 是使用额外的随机码本级的实时语音频带解码器的框图。
- [0044] 图 10 是有关各帧的比特率格式的框图,其中上述帧包括有关可与某些实施例一并使用的不同冗余编码技术的信息。
- [0045] 图 11 是有关各分组的比特率格式的框图,其中上述分组包括具有可与某些实施例一并使用的冗余编码信息的各帧。

具体实施方式

[0046] 描述的实施例涉及用于在编码和解码中处理音频信息的技术和工具。使用这些技术就能够改善从诸如实时语音编解码器之类的语音编解码器中所得到的语音质量。这种提高可以是分别或组合地利用各种技术和工具的结果。

[0047] 这些技术和工具可以包括使用诸如 CELP 之类的线性预测技术而对子带的编码和 / 或解码。

[0048] 该技术还可以包括具有包括脉冲和 / 或随机固定码本在内的固定码本的多级。码本级的数量可以改变从而为给定比特率提供最佳质量。另外，取决于诸如期望的比特率以及当前帧或子帧的特征之类的因素，可以打开或关闭自适应码本。

[0049] 此外，帧可以包括有关当前帧依赖在前帧的部分或全部的冗余编码信息。这种信息可由解码器在在前帧丢失的情况下用来解码当前帧，而不无需请求多次发送整个在前帧。这些信息可以与当前或在前帧一样以相同的比特率被编码，或以更低的比特率被编码。此外，该信息可以包括近似激励信号的期望部分的随机码本信息，而非该激励信号的期望部分的整个重新编码。

[0050] 尽管出于表示的目的以具体的顺序描述了各种技术的操作方法，但应该理解，除非要求一具体次序，否则这种描述的方法涵盖了在操作顺序的可选重新排列。例如，随后描述的操作在一些情况下可以被重新排列或并发执行。此外，出于简化的目的，流程图没有示出在其中特定技术可与其他技术结合使用的各种方法。

I. 计算环境

[0052] 图 1 示出了在其中可以实现一个或多个描述的实施例的适合计算环境 (100) 的概括示例。该计算环境 (100) 并不旨在暗示对本发明的使用或功能范围的任何限制，因为本发明可以在完全不同的通用或专用计算环境中实现。

[0053] 参考图 1，计算环境 (100) 包括至少一个处理单元 (110) 和存储器 (120)。图 1 中，在虚线范围内包括最基本配置 (130)。处理单元 (110) 执行计算机可执行指令并且可以是真实的或虚拟的处理器。在多处理系统中，多个处理单元执行计算机可执行指令来增加处理功率。存储器 (120) 可以是易失性存储器（例如，寄存器、高速缓存、RAM）、非易失性存储器（例如，ROM、EEPROM、闪存等）、或者这两者的组合。存储器 (120) 存储用于为语音编码器或解码器执行子带编码、多级码本、和 / 或冗余编码技术的软件 (180)。

[0054] 计算环境 (100) 可以具有额外的特征件。图 1 中，计算环境 (100) 包括存储 (140)，一个或多个输入设备 (150)，一个或多个输出设备 (160)，以及一个或多个通信连接 (170)。诸如总线、控制器、或网络的互连机构（未示出）互连计算环境 (100) 的各部件。通常操作系统软件（未示出）为在计算环境 (100) 中执行的其他软件提供操作环境，并且协调计算环境 (100) 部件的活动。

[0055] 存储 (140) 可以是可移动或不可移动的，并且可以包括磁盘，磁带或盒式录音带、CD-ROM、CD-RW、DVD、或其他任何能够存储信息并且能够在计算环境 (100) 中被访问的介质。存储 (140) 存储软件 (180) 的指令。

[0056] 输入设备 (150) 可以是触摸式输入设备，例如键盘、鼠标、笔、或跟踪球、话音输入设备、扫描设备、网络适配器、或提供输入到计算环境 (100) 的另一设备。对于音频，输入设

备(150)可以是声卡、麦克风或其他以模拟或数字格式接受音频输入的设备、或提供音频样本到计算环境(100)的CD/DVD读卡器。输出设备(160)可以是显示器、打印机、扬声器、CD/DVD写入器、网络适配器、或者或提供来自计算环境(100)的输出的另一设备。

[0057] 通信连接(170)能够经通信介质和另一个计算实体通信。通信介质传送信息,诸如计算机可执行指令、压缩语音信息、或其他已调制数据信号。已调制数据信号指的是这样一种信号,其一个或多个特征以在信号中编码信息的方式被设定或更改。作为示例,而非限制,通信介质包括用电、光学、RF、红外线、声学的或其他载体实现的有线或无线技术。

[0058] 本发明可以在计算机可读介质的通用上下文中描述。计算机可读介质是任何能够在计算环境中被访问的可用介质。作为示例而非限制,结合计算环境(100),计算机可读介质包括存储器(120)、存储(140)、通信介质、以及上述任何组合。

[0059] 本发明可以在诸如那些包括在程序模块中的计算机可执行指令的通用上下文中描述,以在目标真实或虚拟处理器上计算环境中执行的。程序模块通常包括例程程序、库、对象、类、部件和数据结构等,用以执行特定的任务或实现特定的抽象数据类型。可以按需在不同实施例中的编程模式之间组合或分割各程序模块。用于程序模块的计算机可执行指令可以是在本地或分布的计算环境中执行。

[0060] 出于表示的目的,详细描述使用像“确定”、“生成”、“调整”以及“应用”之类的术语来描述在计算环境中的计算机操作。这些术语是由计算机执行操作的高级抽象,并且不应该与人执行的动作相混淆。对应于这些术语的实际计算机操作则根据执行而变化。

[0061] II. 概括的网络环境及实时语音编解码器

[0062] 图2是结合一个或多个所述实施例被执行的概括的网络环境(200)的框图。网络(250)区分各种编码器侧的部件和各种解码器侧的部件。

[0063] 编码器侧和解码器侧的主要功能分别是语音编码和解码。在编码器侧,输入缓冲器(210)接受并存储语音输入(202)。语音编码器(230)从输入缓冲器(210)中获取语音输入(202)并对其进行编码。

[0064] 更具体地,帧分裂器(212)将语音输入(202)的样本分成各帧。在一个实现中,帧统一为20ms长——8kHz输入下的160个样本以及16kHz输入下的320个样本。在其他实现中,帧具有不同的持续时间,且不均匀或重叠,和/或输入(202)的采样速率不同。帧可以在超帧/帧、帧/子帧、或其他用以编码和解码各不同级的配置中被组织。

[0065] 帧分类器(214)根据一个或多个准则进行帧分类,这些准则诸如可以是信号能量、零交叉率、长期预测增益、增益微分和/或其他用于子帧或整个帧的准则。基于该准则,帧分类器(214)将不同帧分为诸如无声的、清音的、浊音的、和过渡的(例如,从清音到浊音)的各类。另外,可以根据冗余编码的类型(用于帧,如果有的话)对帧进行分类。帧分类会影响将用于计算编码帧的参数。另外,帧分类会影响用其编码的参数的解析和丢失恢复能力,以便为更重要的帧分类和参数提供更多的解和丢失恢复能力。例如,无声帧通常以极慢的速率编码,若丢失则可由隐蔽简单地恢复,并且无需丢失防护。浊音帧通常以略高的速率编码,若丢失则可由隐蔽合理简单地恢复,并且无需显著的丢失防护。清音和过渡帧通常取决于帧的复杂性以及过渡的呈现而用更多的比特进行编码。清音和过渡帧若有丢失则难以恢复,因而需要更为显著的丢失防护。可选地,帧分类器(214)利用其他的和/或额外的帧分类。

[0066] 在向有关帧的子带信息应用诸如 CELP 编码模型的编码模型之前,可以将输入语音信号分为子带信号。可以利用一系列的一个或多个分析滤波器列(例如 QMF 分析滤波器)(216)来实现。例如,如果使用 3 频带结构,则使用通过让信号穿过低通滤波器来分出低频带。同样地,使用通过让信号穿过高通滤波器来分出高频带。使用通过让信号穿过顺序包括一个低通滤波器和一个高通滤波器的带通滤波器来分出中间频带。可选的,使用可以使用用于子带分解和 / 或滤波定时(例如,在帧分配之前)的其他滤波器排列类型。如果近对部分信号解码一个频带,则该部分绕过该分析滤波器列(216)。CELP 编码在语音信号时通常比 ADPCM 和 MLT 具有更高的编码效率。

[0067] 频带的数量 n 可由采样速率确定。例如,在一实现中,单个频带结构使用被用于 8kHz 采样速率。对于 16kHz 和 22.05kHz 采样速率而言,则可使用如图 3 所示的 3 频带结构。在图 3 的 3 频带结构中,低频频带(310)延伸至全部带宽 F 的一半(从 0 到 0.5F)。带宽的另一半在中间频带(320)和高频频带(330)之间等分。频带的交叉点附近,响应于一频带的频率会逐渐地从通过级减至向停止级,其特征随着交叉点的接近而衰减信号的两侧。也可以使用频率带宽的其他划分。例如,对于 32kHz 采样速率而言,可以使用等分的 4 频带结构。

[0068] 低频频带通常是语音信号最重要的频带,因为信号能量一般朝着更高频率范围而衰减。因此,低频频带通常使用比其他频带更多的比特来编码。相比较单个频带编码结构而言,子带结构更灵活,并且允许更好地控制横跨各频段的比特分布 / 量化噪声。因此,可以相信通过使用子带结构能有效提高所感知的话音质量。

[0069] 图 2 中,每个子带如编码部件(232,234)所示的那样被分别编码。部件虽然分别示出了频带编码部件(232,234),但是所有的频带编码可以由一个编码器来完成,或者它们可以通过分开的编码器来编码。这样的频带编码将在以下参考图 4 来更详细地描述。可选地,编解码器可以作为一单独的编解码器操作。

[0070] 通过多路复用器(“MUX”)(236)将编码语音的结果提供给用于一个或多个网络层(240)的软件。网络层(240)处理编码的语音以供经网络(250)的传输。例如,该网络层软件将编码的语音信息打包成遵循 RTP 协议的分组,这些分组经使用 UDP、IP、和各种物理层协议的网络来中继。可选地,使用还可以使用其他和 / 或附加的软件层或网络协议。该网络(250)是分组交换的广域网,例如因特网。可选地,网络(250)也可以是局域网或其他种类的网络。

[0071] 在解码器侧,用于一个或多个网络层(260)的软件接收并处理被传输的数据。解码器 - 侧网络层(260)内的网络、传输及更高层协议和软件一般与编码器侧网络层(240)中的那些部件相对应。网络层通过多路分配器(“DEMUX”)(276)向语音解码器(270)提供编码的语音信息。解码器(270)如解码模块(272,274)中描述的那样分别解码每个子带。所有的子带可以由单个解码器解码,或者可由分开的频带解码器解码。

[0072] 该解码子带接着在输出解码语音(292)的一系列的一个或多个合成滤波器列(例如, QMF 合成滤波器)(280)中合成。可选地,使用可以使用其他类型的滤波器排列以供子带合成。如果仅仅出现单个频带,则该解码频带可以绕过滤波器列(280)。

[0073] 该解码语音输出(292)也可以被传递通过一个或多个后置滤波器(284)来提高所得的经滤波语音输出(294)的质量。同样,每个频带可以在进入滤波器列(280)之前分别

通过一个或多个后置滤波器。

[0074] 下面参考图 6 描述一种概括的实时语音频带解码器，但是也可以代替地使用其他的语音解码器。另外，描述的部分或全部工具和技术都可以结合例如音乐编码器和解码器，或通用型音频编码器和解码器的其他类型的音频编码器及解码器来使用。

[0075] 除了这些主要编码和解码功能，部件也可以共享信息（图 2 中用虚线表示的）以控制编码语音的速率、质量、和 / 或丢失恢复能力。该速率控制器（220）考虑多种因素，诸如输入缓冲器（210）中当前输入的复杂性、编码器（230）或其他设备中输出缓冲器的缓冲充满度、期望的输出速率、当前网络带宽、网络拥堵 / 噪音条件和 / 或解码器丢失速率。解码器（270）向速率控制器（220）反馈解码器丢失速率信息。网络层（240, 260）收集或估计有关当前网络带宽以及拥堵 / 噪音条件的信息，则这些信息则被反馈到速率控制器（220）。可选地，速率控制器（220）考虑其他和 / 或附加的因素。

[0076] 速率控制器（220）引导语音编码器（230）改变被编码语音的速率、质量、和 / 或丢失恢复能力。编码器（230）可以通过调整有关参数的量化因素或改变表示参数的熵码的解来改变速率和质量。另外，编码器通过调整冗余编码的速率或类型还可以改变丢失恢复能力。因此，编码器（230）可以取决于网络条件来改变主要编码功能和丢失恢复能力功能之间的比特分配。

[0077] 速率控制器（220）可以基于若干因素为每个帧的每个子带确定编码模式。这些要素可以包括每个子带的信号特征、比特流缓冲历史、以及目标比特速率。例如，如上所述的，例如清音和无声帧之类的较简单的帧通常需要的比特较少，而像过渡帧的更复杂的帧需要的比特则较多。另外，例如高频频带之类的某些频带需要的比特较少。此外，如果比特流历史缓冲器中的平均比特率小于目标平均比特率，则可为当前帧使用更高的比特率。如果平均比特率大于目标平均比特率，则可为当前帧选择较低的比特率以降低平均比特率。另外，可以从一个或多个帧中省略一个或多个频带。例如，从清音帧中省略中间帧和高频帧，或者它们从所有帧中被忽略一段时间，从而降低在那段时间内的比特率。

[0078] 图 4 是结合一个或多个所描述的实施例来实现的概括的语音频带编码器（400）的框图。频带编码器（400）通常与图 2 的频带编码器（232, 234）中的任一个相对应。

[0079] 频带编码器（400）在信号（例如，当前帧）被分成多个频带情况下接受来自滤波器列（或其他滤波器）的频带输入。如果当前帧没有被分为多个频带，则频带输入（402）包括表示整个带宽的采样。该频带编码器产生编码的频带输出（492）。

[0080] 如果信号被分为多个频带，则下采样部件（420）可以对每个频带上执行下采样。作为例子，如果采样速率被设置为 16kHz 且每个帧的持续时间是 20ms，则每个帧包括 320 个样本。如果没有执行下采样且帧被分为图 3 所示的 3 频带结构，则会对该帧进行样本个数 3 倍的（即，每频带 320 个样本，或总共 960 个样本）编码和解码。然而，每个频带可以被下采样。例如，低频频带（310）可以从 320 个样本下采样至 160 个样本，且每个中间频带（320）和高频频带（330）可以从 320 个样本下采样至 80 个样本，这里频带（310, 320, 330）分别疏至频率范围的一半，四分之一和四分之一。（在该实现中下采样（420）程度的变化与频带（310, 320, 330）的频率范围有关。然而，其他实现也是可能的。在随后的各级中，频带越高使用的比特通常较少，因为信号能量通常朝着更高频率范围衰减。）因此，这就为该帧提供了总共 320 个样本来进行编码和解码。

[0081] 可以相信就算使用了每个频带的下采样,该子带编解码器仍可比单个频带编解码器生成更高的话音质量输出,因为该子带编解码器更灵活。例如,它可以更灵活地以每个频带为基础控制量化噪声,若不是对整个频谱使用相同的手段。多个频带的每一个都能够被编码带有不同的属性(例如,以下将会讨论的不同数量和/或类型的码本级)。这些属性可以按包括每个子带的信号特征、比特流缓冲历史以及目标比特率在内的上述若干因素的基础而由速率控制来确定。如上所述,诸如清音帧和无声帧的“简单”帧通常需要较少的比特,而诸如过渡帧的“复杂”帧则需要更多的比特。如果比特流历史缓冲器中的平均比特率小于目标平均比特率,可以为当前帧使用较高的比特速率。否则就选择较低的比特速率以降低平均比特率。在子带编解码器中,每个频带可以以这种方式为特征并被相应地编码,而非以相同的方式特征化整个频谱。另外,速率控制通过为一个或多个帧忽略一个或多个较高频率的频带就能够降低比特率。

[0082] LP 分析部件 (430) 计算线性预测系数 (432)。在一实现中,LP 滤波器对 8kHz 输入使用 10 个系数而对 16kHz 输入使用 16 个系数,并且 LP 部件分析部件 (430) 为每个频带计算每帧的一组线性预测系数。可选地,LP 分析部件 (430) 为每个频带计算每帧的两组系数,每组各用于以不同位置为中心上的两个窗口中的一个,或者 LP 分析部件 (430) 计算每频带和/或每帧的不同数量的系数。

[0083] LPC 处理部件 (435) 接收并处理线性预测系数 (432)。通常 LPC 处理部件 (435) 为了更有效地量化并编码部件而把 LPC 值转换为不同表示。例如,LPC 处理部件 (435) 将 LPC 值转换为线光谱对 [“LSP”] 表示,并且该 LSP 值被量化(例如通过矢量量化)并编码。LSP 值可被内编码或从其他 LSP 值预测。各种表示、量化技术和编码技术都有可能用于 LPC 值。以一些格式来提供的 LPC 值作为编码频带输出 (492) 的一部分用于分组和传输(连同任何量化参数以及重建所需的其他信息一起)。对于编码器 (400) 中随后的使用,LPC 处理部件 (435) 重建 LPC 值。该 LPC 处理部件 (435) 可以为 LPC 值(诸如,等效的 LSP 表示或另一表示)执行插值,以平滑不同的 LPC 系数组之间的转换或者用于帧的不同子帧使用的 LPC 系数之间的转换。

[0084] 合成(或“短期预测”)滤波器 (440) 接收重建的 LPC 值 (438) 并将它们合并到滤波器中。合成滤波器 (440) 接受一激励信号并生成原始信号的近似值。对于指定帧,合成滤波器 (440) 会缓冲来自预测开始之前的在前帧中的多个重建样本(例如,1 每 10 接头滤波器 10 个)。

[0085] 感知加权部件 (450,455) 将感知加权应用于原始数据以及合成滤波器 (440) 的模型化输出,以便选择性地削减语音信号共振峰结构的重要性,从而使听觉系统对量化误差不那么敏感。感知加权部件 (450,455) 使用例如掩蔽的心理声学现象。在一实现中,感知加权部件 (450,455) 基于从 LP 分析部件 (430) 中导出的原始 LPC 值 (422) 应用权重。可选地,感知加权部件 (450,455) 应用其他和/或额外权重。

[0086] 感知加权部件 (450,455) 之后,编码器 (400) 计算被感知加权的原始信号和被感知加权的合成滤波器输出之间的区别,以产生差信号 (434)。可选地,编码器 (400) 利用不同的技术来计算语音参数。

[0087] 在最小化被感知加权的原始值与合成信号(根据被加权的均方误差或其他准则)之间差异方面,激励参数化部件 (460) 搜索并找出自适应码本索引、固定的码本索引以及

增益码本索引的最佳组合。可以为每一子帧计算许多参数,但是更普遍的是为每一个超帧、帧、或子帧计算参数。如上所述,用于帧或子帧的不同频带的参数可以不同。表 2 示出了在一实现中用于不同帧分类的可用参数类型。

[0088]

帧分类	参数
无声	分类信息; LSP; 增益 (每帧, 对于生成的噪声)
清音	分类信息; LSP; 脉冲, 随机以及增益码本参数
浊音	分类信息; LSP; 自适应、脉冲、随机以及增益码本参数 (每子帧)
过渡	

[0089] 表 2 :用于不同帧分类的参数

[0090] 图 4 中,激励参数化部件 (460) 将帧分为子帧并为每个子帧适当地计算码本索引和增益。例如,要使用的码本级的数量和类型使用、以及码本索引的解都可由一编码模式来最初确定,其中该模式可由上述速率控制部件来规定。一具体模式也可以规定除了码本级的数量和类型之外的编码和解码参数,例如,码本索引的解析。每个码本级的参数通过优化参数来确定,以最小化目标信号与码本级对合成信号的贡献 (contribution) 之间的误差。(使用在此使用的术语“优化”表示相对于执行参数空间上的完全搜索,而在诸如失真降低、参数搜索时间、参数搜索复杂性、参数比特率等的应用限制下找到合适的解决方案。类似的,术语“最小化”可以在可用限制下找到合适的解决方案这一方面进行理解)。例如,可以使用修改的均方误差技术来实现优化。每一级的目标信号是残留信号与各在前码本级(如果有的话)对合成信号的贡献总和之间的差异。可选地,可以使用其他优化技术。

[0091] 图 5 根据一实现示出了一种确定码本参数的技术。激励参数化部件 (460) 潜在地结合例如速率控制器的其他部件来执行该技术。可选地,编码器中的其他部件执行该技术。

[0092] 参照图 5,对于浊音或过渡帧内的每个子帧,激励参数化部件 (460) 确定 (510) 自适应码本 (ACB) 是否可被用于当前子帧。(例如,速率控制可以规定没有自适应码本被用于一特定帧。) 如果自适应码本不被使用,随后自适应码本转换将指示没有自适应码本被使用 (535)。例如,这可以通过在帧层处设置指示了没有自适应码本被用于该帧的一比特标记来实现,或者通过为每个子帧设置指示了没有自适应码本被用于该子帧的一比特标记来实现。

[0093] 例如,速率控制部件可以排除用于帧的自适应码本,从而移除帧之间最为明显的记忆依赖。尤其对于浊音帧而言,一种典型的激励信号以周期模式为特征。该自适应码本包括表示滞后的索引,该滞后指示了历史缓冲器内一段激励的位置。该段在前激励被调节为自适应码本对该激励信号的贡献。在解码器上,自适应码本信息通常对重建激励信号相当重要。如果在前帧丢失且自适应码本索引指向在前帧的一段,则该自适应码本索引一般没有用,因为它指向不存在的历史信息。即使执行隐蔽技术来恢复该丢失信息,将来的重建也要基于这一未完善恢复的信号。这将会导致随后个帧内的误差,因为滞后信息通常是敏感的。

[0094] 因此,受到随后自适应码本依赖的分组的丢失会导致扩大的劣化,这种劣化需

要等到许多分组已经被解码之后或者在遇到没有自适应码本的帧时才会逐渐消失。这一问题可以通过有规律地在分组流之间插入在各帧之间没有记忆依赖的所谓“内部帧”来减轻。这样，误差将只会传播直到下一个内部帧。因此，在较好的话音质量和较好的分组丢失性能之间存在一折衷，因为自适应码本的编码效率通常要高于固定的码本的编码效率。速率控制部件可以确定什么时候阻止用于特定帧的自适应码本是有利的。该自适应码本转换被用于防止用于特定帧的自适应码本的使用，从而消除通常最为显著的对在前各帧的依赖（LPC 插值以及合成滤波器记忆在某种程度上也依赖于在前各帧）。因此，该自适应码本转换可由速率控制部件动态地基于诸如分组丢失速率之类的因素来创建准内部帧（quasi-intra-frame）（即，当分组丢失速率高时，可以插入更多的内部帧以允许更快的记忆重置）。

[0095] 仍然参照图 5，如果使用自适应码本，则部件（460）确定自适应码本参数。那些参数包括指示了激励信号历史的期望段的索引、或基音值、以及要应用于该期望段的增益。在图 4 和 5 中，部件（460）执行一个闭环基音搜索（520）。该搜索由图 4 中可选开环基音搜索部件（425）所确定的基音开始。开环基音搜索部件（425）分析由加权部件（450）生成的加权信号以估计其基音。由此估计的基音开始，闭环基音搜索（520）优化该基音值以减少目标信号与从激励信号历史的指示段中生成的加权合成信号之间的误差。自适应码本增益值（525）也被优化。该自适应码本增益值指示应用于基音预测值（该值来自于激励信号历史的指示段）的乘法器，以调整上述各值比例。乘以基音预测值的该增益是自适应码本对用于当前帧或子帧的激励信号的贡献。增益优化（525）产生增益值以及索引值，该索引值最小化目标信号与由自适应码本贡献的加权合成信号之间的误差。

[0096] 当确定基音和增益值之后，就确定（530）自适应码本贡献是否显著到足以使其值得由各自适应码本参数使用的比特数。如果自适应码本增益小于阈值，则关闭自适应码本来为如下讨论的固定码本保存比特。在一个实施例中，使用阈值 0.3，尽管其他可选值也可以作为阈值。作为例子，如果当前编码模式使用自适应码本加上具有 5 个脉冲的脉冲码本，则可以在关闭自适应码本时使用 7 脉冲码本，并且比特总数将仍会相同或更少。如上所述，可以用于每个子帧的一比特标记可用来指示有关该子帧的自适应码本转换。因此，如果没有使用自适应码本，设置该转换以指示在子帧中没有使用自适应码本（535）。同样地，如果使用自适应码本，则设置该转换以指示在子帧中使用了自适应码本并且在比特流中发信号通知这些自适应码本参数（540）。尽管图 5 示出了在确定之后发信号通知，但是也可以直到技术完成一帧或超帧才成批处理信号。

[0097] 该激励参数化部件（460）同样确定（550）是否使用脉冲码本（pulse CB）。在一个实施例中，使用或不使用脉冲码本被作为用于当前帧的整个编码模式的一部分来指示，或者也可在其他方面被指示或确定。脉冲码本是一类指定要对该激励信号做出贡献的一个或多个脉冲的固定码本，来组成激励信号。该脉冲码本参数包括索引和符号（增益可能是正的或负的）对。每一对指示一个被包含在激励信号内的脉冲，其中索引指示脉冲位置而符号则指示脉冲极性。包含在脉冲码本内并用于贡献激励信号的脉冲数量会依赖于编码模式而变化。此外，脉冲数量还依赖于是否使用自适应码本。

[0098] 如果使用脉冲码本，则优化脉冲码本参数（555）来最小化指示脉冲的贡献和目标信号之间的误差。如果没有使用自适应码本，目标信号就是加权的原始信号。如果使用自

适应码本，则目标信号是加权原始信号和自适应码本对加权合成信号的贡献之间的差。在一些点上（未示出），脉冲码本参数接着在比特流中发信号通知。

[0099] 激励参数化部件 (460) 也能确定 (565) 是否使用任何随机固定码本。随机码本级的数量（如果有的话）被指示作为用于当前帧的全部编码模式的一部分，尽管可以在其他方面被指示或确定。随机码本是一类为其编码的值使用预定义信号模型的固定码本。该码本参数可以包括用于信号模型指示段的起始点以及可能或正或负的符号。该指示段的长度或范围通常是固定的，因此一般不用信号通知，但是另外也可发信号通知指示段的长度或范围。增益乘以指示段中的值以生成随机码本对激励信号的贡献。

[0100] 如果至少使用一个随机码本 (random CB) 级，则优化适用于该码本级的码本级参数 (570) 从而最小化随机码本级的贡献与目标信号之间的误差。目标信号是加权的原始信号与自适应码本（如果有的话）、脉冲码本（如果有的话）、以及在前确定的随机码本级（如果有的话）对加权合成信号的贡献总和之间的差。在一些点上（未示出），接着在比特流中发信号通知该随机码本参数。

[0101] 部件 (460) 接着确定 (580) 是否要使用任何更多的随机码本级。如果是，接着优化 (570) 下一个随机码本级的各参数并如上所述那样发信号通知。这将持续直到用于随机码本的所有参数都被确定。所有随机码本级可以使用相同的信号模型，虽然它们可是指示与该模型不同的段并具有不同的增益值。可选地，可以为不同的随机码本级使用不同的信号模型。

[0102] 每个激励增益可以被独立地量化，或两个或多个增益可以被同时量化，正如由速率控制器和 / 或其他部件所确定的那样。

[0103] 虽然已在此阐述了用于优化各不同码本参数的特定次序，也可以使用其他次序和优化技术。因此，虽然图 5 示出了不同码本参数的顺序计算，但另外也可以共同优化两个或更多不同的码本参数（例如，根据一些非线性优化技术共同改变参数并估计结果）。此外，可以使用码本的其他配置或激励信号参数。

[0104] 在该实现中的激励信号是自适应码本、脉冲码本、以及一个或多个随机码本级的任何贡献之和。可选地，部件 (460) 可以为激励信号计算其他的和 / 或附加的参数。

[0105] 参照图 4，用于激励信号的码本参数被发信号通知或通过其他方式被提供给本地解码器 (465)（在图 4 中用虚线圈出）以及频带输出 (492)。因此，对于每个频带而言，编码器输出 (492) 包括来自上述 LPC 处理部件 (435) 的输出，以及来自激励参数化部件 (460) 的输出。

[0106] 输出 (492) 的比特率部分依赖于码本所用的参数，并且编码器 (400) 可以通过在不同的码本索引集合之间转换，使用嵌入式编码，或使用其他技术来控制比特率和 / 或质量。码本类型和级的不同组合可以产生用于不同帧、频带、和 / 或子帧的不同编码模型。例如，一种清音帧可以仅使用一个随机码本级。自适应码本以及脉冲码本可以用于低速率浊音帧。高速率帧则可以使用自适应码本、脉冲码本、以及一个或多个随机码本级来编码。在一帧中，对于所有子带的所有这些编码模式的组合被统称为模式集合。存在用于每个采样速率的若干预定义的模式集合，这些模式集合带有与不同编码比特率相对应的不同模式。速率控制模块能够确定或影响用于每个帧的模式集合。

[0107] 可能的比特率范围对于所描述的实现而言可能非常大，并且会对所得质量产生显

著的改进。在标准编码器中,用于脉冲码本的比特的数量也可被改变,但是太多比特可以仅仅产生过度密集的脉冲。类似地,当仅使用单个码本时,添加更多的比特就能使用更大的信号模型。但这会显著增加用于该模型优化段搜索的复杂度。相反,可以添加码本的附加类型以及附加的随机码本级而不会显著增加各自码本搜索的复杂度(与搜索单个的组合码本相比)。此外,多个随机码本级和多类固定码本允许多个增益因素来提供更灵活的波形匹配。

[0108] 仍然参照图4,激励参数化部件(460)的输出被码本重建部件(470,472,474,476)以及与参数化部件(460)所用的各码本相对应的增益应用部件(480,482,484,486)接收。码本级(470,472,474,476)以及相应的增益应用部件(480,482,484,486)重建码本的贡献。总计这些贡献以产生激励信号(490),该信号由合成滤波器(440)接收,在其中该信号连同后继的线性预测由此产生“预测”样本一起使用。激励信号的延迟部分还由自适应码本重建部件(470)用作激励历史信号来重建后继的自适应码本参数(例如,基音贡献),并由参数化部件(460)来计算后继的自适应码本参数(例如,基音索引以及基音增益值)。

[0109] 参照回图2,由MUX(236)接收用于每个频带的频带输出,以及其他参数。这些其他参数包括来自帧分类器(214)的帧分类信息(222)和帧编码模式的信息。MUX(236)构造应用层分组来传递给其他软件,或者MUX(236)遵循例如RTP的协议而将数据放入到分组的净荷。该MUX缓冲参数以允许选择性的重复参数,以供在随后各分组中的前向纠错。在一实现中,MUX(236)将有关一帧的主要编码语音信息,连同有关全部或部分的一个或多个在前帧的向前误差纠正信息一起打包成一个独立的分组。

[0110] MUX(236)出于速率控制的目的提供诸如当前缓冲充满度之类的反馈。更一般地,编码器(230)的各个部件(包括帧分类器(214)和MUX(236))可以向诸如图2所示那样的速率控制器(220)提供信息。

[0111] 图2中的比特流DEMUX(276)接收编码语音信息作为输入并解析它来识别和处理参数。这些参数可以包括帧分类、LPC值的一些表示、以及码本参数。帧分类可以指示对于给定帧存在那些其他参数。更具体地,DEMUX(276)使用编码器(230)所用的协议并从编码器(230)打包成的分组中提取参数。对于经动态分组转换网络接收分组,DEMUX(276)包括抖动缓冲器,用来平滑给定时段内分组速率内的短期波动。在一些情况下,解码器(270)调节缓冲器延迟并管理何时从缓冲器中读取分组以便把延迟、质量控制、遗漏帧的隐蔽等归并到一起来解码。在其他情况下,应用程序层部件管理抖动缓冲器,而该抖动缓冲器则以可变速率填充并通过解码器(270)以一恒定或相对恒定速率耗尽。

[0112] DEMUX(276)可以接收用于给定段的各参数的多个版本,包括主要编码版本以及一个或多个次要纠错版本。当纠错失败,解码器(270)则基于被正确接收的信息而使用诸如参数重复或估计的隐蔽技术。

[0113] 图6是一种结合其可以实现一个或多个描述的实施例的实时语音频带解码器的框图。频带解码器(600)一般对应于图2中频带解码部件(272,274)的任何一个。

[0114] 频带解码器(600)接收用于频带(可以是完整频带,或多个子带之一)的编码语音信息作为输入并在解码之后生成重建的输出(602)。该解码器部件(600)具有编码器(400)内对应的部件,但是整体上解码器(600)更为简单,因为它没有用于感知加权、激励处理循环以及速率控制的部件。

[0115] LPC 处理部件 (635) 接收带有由频带编码器 (400) 提供的格式的表示 LPC 值的信息 (以及重建所需的任何量化参数和其他信息)。LPC 处理部件 (635) 使用先前应用与 LPC 值的变换、量化、编码等的逆来重建 LPC 值 (638)。LPC 处理部件 (635) 还可以为 LPC 值执行插值 (在 LPC 表示或例如 LSP 的另一表示中) 来平滑 LPC 系数不同集合之间的过渡。

[0116] 码本级 (670, 672, 674, 676) 以及增益应用部件 (680, 682, 684, 686) 解码用于激励信号的任何相应码本级的参数并计算所使用的每个码本级的贡献。更具体地, 码本级 (670, 672, 674, 676) 和增益部件 (680, 682, 684, 686) 的配置和操作对应于编码器 (400) 中的码本级 (470, 472, 474, 476) 和增益部件 (480, 482, 484, 486) 的配置和操作。总计所用的码本级贡献, 并将所得激励信号 (690) 送入合成滤波器 (640)。激励信号 (690) 的延迟值也在计算用于激励信号后继部分的自适应码本的贡献中由该自适应码本 (670) 用作激励历史。

[0117] 合成滤波器 (640) 接收重建的 LPC 值 (638) 并将它们并入滤波器。该合成滤波器 (640) 存储在前重建的样本用于处理。激励信号 (690) 被传递通过合成滤波器以形成原始语音信号的近似值。参见回到图 2, 如上所述, 如果存在有多个子带, 就在滤波器列 (280) 中合成用于每个子带的子带输出以形成语音输出 (292)。

[0118] 如图 2-6 所示的关系指示了一般信息流;为了简化没有示出其他关系。取决于实现及期望的压缩类型, 各部件可以被添加, 省略, 分割成多个部件, 与其他部件组合, 和 / 或由类似部件替换。例如, 在图 2 所示的环境 (200) 中, 速率控制器 (220) 可以与语音编码器 (230) 相组合。可能添加的部件包括多媒体编码器 (或回放) 应用, 它管理语音编码器 (或解码器) 以及其他编码器 (或解码器) 并收集网络和解码器条件信息, 并且执行自适应纠错功能。在可选的实施例中, 各部件的不同组合及配置使用这里描述的技术来处理语音信息。

[0119] III. 冗余编码技术

[0120] 语音编解码器的一种可能的应用是针对 IP 网络电话 (voice over IP network) 或其他分组转换网络。这些网络具有优于现有电路转换基础设施的一些优点。然而, 在 IP 网络电话中, 分组经常会由于网络拥堵而被延迟或衰落。

[0121] 许多标准语音编解码器具有较高的内部帧依赖。于是对于这些编解码器而言, 一帧的丢失会导致殃及随后许多帧的严重话音质量劣化。

[0122] 在其他编解码器中可以独立地解码每一帧。这样的帧可以应付分组丢失。然而就质量和比特速率而言, 编码效率则由于不允许内部帧依赖而显著下降。因此, 这些编解码器通常需要更高的比特率来实现与传统的 CELP 编码器相似的话音质量。

[0123] 在一些实施例中, 下面将讨论的冗余编码技术有助于在不显著增加比特率的情况下实现良好的分组丢失恢复性能。该技术可以一并的用于编解码器中, 也可以分开使用。

[0124] 在如上参照图 2 和 4 所述的编码器实现中, 自适应码本信息通常是对其他帧的主要依赖源。如上所述, 该自适应码本索引指示了历史缓冲器中一段激励信号的位置。在前激励信号的这一段被 (根据增益值) 调节成当前帧 (或子帧) 激励信号的自适应码本贡献。如果包含用于重建在前的编码激励信号的信息的在前分组丢失, 则该当前帧 (或子帧) 滞后信息因为它指向非现存的历史信息而不可用。因为滞后信息是敏感的, 所以这通常会导致所得语音输出的扩大的劣化, 这种劣化需要等到许多分组已经被解码之后才会逐渐消失。

[0125] 下面的技术被设计用来至少在某种程度上去除当前激励信号对来自因被延迟或丢失的不可用在前帧的重建信息的依赖。

[0126] 诸如上述参考图 2 所示的编码器 (230) 之类的编码器能够基于逐帧或其他而在下列的各编码技术之间转换。诸如上述参考图 2 所示的解码器 (270) 之类的解码器则能够基于逐帧或其他而转换相应的分析 / 解码技术。可选地，另一编码器、解码器、或音频处理工具也可以执行以下技术的一个或多个。

[0127] A. 主要自适应码本历史重新编码 / 解码

[0128] 在主要自适应码本历史重新编码 / 解码中，激励历史缓冲器不用于解码当前帧的激励信号，即使激励历史缓冲器在解码器处可用（在前帧的分组接收，在前帧解码等）。代替地，在编码器上，为当前帧分析基音信息以确定需要多少激励历史。激励历史的必要部分被重新编码并连同有关当前帧的编码信息（例如，滤波器参数、码本索引以及增益）被一起发送。当前帧的自适应码本贡献参考了连同当前帧一起发送的重新编码激励信号。这样就为每一帧确保了冗余激励历史对解码器可用。这种冗余编码在诸如清音帧的当前帧没有使用自适应码本的情况下不是必要的。

[0129] 激励历史的被参考部分的重新编码可以连同当前帧的编码一起完成，并且可以通过与如上所述对有关当前帧的激励信号的编码相同的方式来完成。

[0130] 在一些实现中，激励信号的编码基于子帧完成，并且该段重新编码激励信号从包括当前子帧的当前帧的开始部分延伸回到超出对当前帧的最远自适应码本依赖的子帧边界。重新编码的激励信号因此可用于参考与该帧内的多个子帧有关的基音信息。可选地，激励信号的编码可以基于诸如逐帧的其他方式来实现。

[0131] 图 7 中示出了一个描述了激励历史 (710) 的例子。帧边界 (720) 和子帧边界 (730) 分别由较大和较小的虚线描述。使用自适应码本来编码当前帧 (740) 的子帧。线 (750) 描述了用于当前帧的子帧的任何自适应滞后索引的最远依赖点。因此，重新编码历史 (760) 从当前帧的开始处延伸越过最远点 (750) 的下一子帧边界。该最依赖远点可以使用上述的开环基音搜索 (425) 的结果来估计。因为该搜索不精确，然而有可能该自适应码本依赖超出了估计的最远点的激励信号的某些部分，除非随后的基音搜索被限定。因此，重新编码历史可以包括超出估计的最远依赖点的附加样本，从而为寻找匹配基音信息提供额外的空间。在一实现中，至少有十个超出估计的最远依赖点的附加样本被包括在重新编码历史中。当然，也可以包括十个以上样本，从而增加重新编码历史延伸到足以包括匹配当前子帧内各基音周期的基音周期的概率。

[0132] 可选地，仅有当前帧的子帧内被实际参考的先前激励信号的各段被重新编码。例如，具有适当持续时间的一段先前激励信号被重新编码以供解码该持续时间内的单个当前段使用。

[0133] 主要自适应码本历史重新编码 / 解码消除了对先前帧的激励历史的依赖。同时，它允许使用自适应码本，并且不需要重新编码整个在前帧（或者甚至是在前帧的整个激励历史）。然而，相比下面描述的技术而言，重新编码自适应码本记忆需要非常高的比特率，尤其是在该重新编码历史被用来以与带有内部帧依赖的编码 / 解码相同的质量级别而进行主要编码 / 解码时。

[0134] 作为主要自适应码本历史重新编码 / 解码的副产品，重新编码激励信号可用于恢

复用于在前丢失帧的激励信号的至少部分。例如,在当前帧的各子帧解码期间重建重新编码激励信号,并且把重新编码激励信号输入使用实际或估计的滤波器系数重建的LPC合成滤波器中。

[0135] 所得的重建输出信号可用作部分在前帧输出。该技术还有助于估计用于当前帧的合成滤波器记忆的原始状态。使用重新编码历史以及估计的合成滤波器记忆,就能够以与常规编码相同的方式生成当前帧的输出。

[0136] B. 次要自适应码本历史重新编码 / 解码

[0137] 在次要自适应码本历史重新编码 / 解码技术中,当前帧的主要自适应码本编码不变。类似地,当前帧的主要解码不变;它在接收到在前帧的情况下使用在前帧激励历史。

[0138] 使用时,如果先前激励历史没被重建,则使用与前述主要自适应码本历史重新编码 / 解码技术相同的方式来顺序地重新编码激励历史缓冲器。然而相较于主要编码 / 解码,只有很少的比特是用于重新编码,这是因为在没有分组丢失的情况下话音质量不受重新编码信号的影响。用来重新编码激励历史的比特数量能够通过改变各种参数而降低,诸如使用更少的固定码本级,或者在脉冲码本中使用更少的脉冲。

[0139] 当在前帧丢失时,重新编码的激励历史被用于在解码器中生成用于当前帧的自适应码本激励信号。正如在主要自适应码本历史重新编码 / 解码技术中那样,重新编码的激励历史还可用与恢复与在前丢失帧有关的至少部分激励信号。

[0140] 同样地,所得的重建输出信号可用作在前帧输出的部分。这一技术还有助于估计有关该当前帧的合成滤波器记忆的原始状态。使用重新编码的激励历史和估计的合成滤波器记忆,就能够使用与常规编码相同的方式来生成当前帧的输出。

[0141] C. 额外码本级

[0142] 正如次要自适应码本历史重新编码 / 解码技术中那样,在额外码本级技术中,主激励信号编码与参照图 2-5 所述的常规编码相同。然而,也可以确定用于额外码本级的参数。

[0143] 在如图 8 所示的这种编码技术中,假设 (810) 在当前帧开始处的在前激励历史缓冲器全部为零,因此不存在来自在前激励历史缓冲器的贡献。除了用于当前帧的主编码信息之外,一个或多个额外码本级也可用于使用自适应码本的每个子帧或其他段。例如,额外码本级使用了随机的固定码本,诸如参照图 4 描述的那些码本。

[0144] 在此技术中,通常编码当前帧以产生当在前帧可用的情况下供解码器使用的主编码信息(可以包括用于主码本级的主码本参数)。在编码器侧,假设没有来自在前帧的激励信息,则在闭环内确定用于一个或多个额外码本级的冗余参数。在第一顺序中,这一确定可以在不使用任何主码本参数的情况下做出。可选地,在第二实现中,确定使用用于当前帧的至少部分主码本参数。那些主码本参数可以连同额外码本级参数一起用来在如下所述在前帧丢失的情况下解码当前帧。一般说来,该第二实现可以使用额外码本级所需的更少比特来实现与第一实现类似的质量。

[0145] 根据图 8,额外码本级的增益以及最后存在的脉冲或随机码本的增益被共同地在编码器闭环搜索中优化,从而最小化编码误差。在常规编码中形成的大部分参数被保存并在优化中使用。在优化中,确定 (820) 是否有任何随机或脉冲码本级在普通编码中使用。如果是,则优化 (830) 最后存在的随机或脉冲码本级(诸如,图 4 中的随机码本级 n) 的校正

增益,从而最小化该码本级的贡献与目标信号之间的误差。用于该优化的目标信号是残留信号与任何前述随机码本级(即,所有前述码本级,但是来自在前帧各段的自适应码本贡献被设置为零)的贡献总和之间的差。

[0146] 额外随机码本级的索引和增益参数被类似地优化(840)以最小化该码本贡献与目标信号之间的误差。用于该额外随机码本级的目标信号是残留信号与自适应码本、脉冲码本(如果有的话)以及任何常规随机码本(连同具有修正增益的最后存在的常规随机或脉冲码本)的贡献总和之间的区别。最后存在的常规随机或脉冲码本的校正增益与额外随机码本级的增益可以被分别或共同优化。

[0147] 当处于常规解码模式时,解码器不使用额外随机码本级,且根据以上描述(例如,如图6所示)来解码信号。

[0148] 图9A示出了一种可以在自适应码本索引指向已经丢失的一段在前帧的情况下使用额外码本级的子带解码器。该框架通常与在图6中描述并示出的解码框架相同,且图9子带解码器(900)中的许多部件和信号的功能与图6中相应的部件和信号相同。例如,接收编码子带信息(992),LPC处理部件(935)使用该信息重建线性预测系数(938),并将这些系数据供给合成滤波器(940)。然而当在前帧缺失时,重置部件(996)发信号通知零历史部件(994),用来将用于缺失帧的激励历史设置为零,并将该历史提供给自适应码本(970)。增益(980)被应用于自适应码本的贡献。自适应码本(970)于是在其索引指向与该缺失帧的历史缓冲器时就具有零贡献,但是在前索引指向当前帧内部一段时则可能具有一些非零贡献。固定码本级(972,974,976)应用它们用子带信息(992)接收的常规索引。类似地,除了最近常规码本部件(986)之外的固定码本增益部件(982,984)也应用它们的常规索引来生成对激励信号(990)的各自贡献。

[0149] 如果额外随机码本级(998)可用并且在前帧缺失,那么重置部件(996)发信号通知转换(998)传递带有残留增益(987)的最后常规码本级(976)的贡献来与其他码本贡献进行总计,而非优于传递带有常规增益(986)的最后常规码本级(976)贡献来用于总计。校正增益在有关在前帧的激励历史被设置为零的情况下被优化。此外,额外码本级(978)应用其索引而在相应的码本中指示该随机码本模型信号的一段,并且随机码本增益部件(988)对那一段应用有关该额外随机码本级的增益。转换(998)传递要与在前码本级(970,972,974,976)进行总计的额外码本级贡献以产生激励信号(990)。因此,用于额外随机码本级的冗余信息(例如额外级索引和增益)以及最后主随机码本级的校正增益(代替有关最后主随机码本级的常规增益使用)被用来将当前帧快速重置到一已知状态。可选地,该常规增益可用于最后主随机码本级和/或一些其他参数可用于发信号通知额外级随机码本。

[0150] 额外码本级技术所需的比特如此之少以至于对其使用的比特率损失通常是无关紧要的。另一方面,它能够显著减少当存在内部帧依赖时由帧丢失所导致的质量劣化。

[0151] 图9B示出了与图9A类似但是没有常规随机码本级的子带解码器。于是在该实现中,校正增益(987)在有关在前丢失帧的残留历史被设置为零时为脉冲码本(972)而被优化。因此,当帧缺失时,各自适应码本(970)(连同有关在前缺失帧的残留历史被设置为零)、脉冲码本(972)(连同校正增益)、以及额外随机码本级(978)的贡献被总计以产生激励信号(990)。

[0152] 在有关缺失帧的残留历史被设置为零的情况下被优化的额外码本级可以结合码本的实现和组合和 / 或残留信号的其他表示一起使用。

[0153] D. 在各冗余编码技术之间的折衷

[0154] 相比其他的而言,上述三个冗余编码技术中的每一个都具有优点和缺点。表 3 示出了被认为是在这三种冗余编码技术之间折衷的概括性结论。比特率损失指的是利用该技术所需的比特总量。例如,假设与在常规编码 / 解码中使用的比特率相同,则在标准解码期间,较高的比特速率损失通常对应于较低的质量,这是因为会有更多的比特被用于冗余编码,于是则是更少的比特被用于常规编码信息。降低记忆依赖的效率指的是当有一个或多个在前帧丢失时用于改进所得语音输出质量的技术的效率。用于恢复在前帧的有效性指的是当在前帧丢失时使用冗余编码信息来恢复一个或多个在前帧的能力。表中的结论是概括性的,并且无需在特定实现中应用。

[0155]

	主要 ACB 历史 编码	次要 ACB 历史 编码	额外码本进程
比特率损失	高	中	低
减轻记忆依赖的效率	最好	好	很好
恢复丢失在前帧的有效性	好	好	无

[0156] 表 3 :各冗余编码技术之间的折衷

[0157] 编码器可以在编码期间为任何空中 (on the fly) 帧选择任一种冗余编码方案。冗余编码对一些帧分类可能完全无用 (例如,用于浊音帧,不用于无声或清音帧),并且如果它被使用,自需要以诸如每十帧的周期为基础或以某些其他基础来用于每个帧。这可以由诸如速率控制部件的部件在考虑各种因素的情况下控制,各因素诸如上述的折衷,可用信道带宽,以及有关分组丢失状态的解码器反馈。

[0158] E. 冗余编码比特流格式

[0159] 该冗余编码信息可以以各种不同的格式在比特流中发送。以下是用于发送上述冗余编码信息以及向解码器用信号通知其表示的一种格式的实现。在该实现中,比特流内的每一帧都以被称为帧类型的两比特字段开始。帧类型用于识别有关如下各比特的冗余编码模式,并且也可以用于编码和解码中其他目的。表 4 给出了表示帧类型字段的冗余编码模式。

[0160]

帧类型比特	冗余编码模式
00	无 (常规帧)
01	额外码本级
10	主要 ACB 历史编码

11	次要 ACB 历史编码
----	-------------

[0161] 表 4 :帧类型比特的描述

[0162] 图 10 示出了在比特流帧格式中这些代码的四种不同的组合,其中这些代码用信号通知常规帧和 / 或各冗余编码类型的存在。对于包括有关该帧的主编码信息而没有任何冗余编码位的常规帧 (1010) 而言,跟随在帧开始处的字节边界 (1015) 之后的是帧类型代码 00。帧类型代码之后则跟有有关常规帧的主编码信息。

[0163] 对于带有主要自适应码本历史的冗余编码信息的帧 (1020) 而言,跟随在帧开始处字节边界 (1025) 之后的是帧类型代码 10,该代码用信号通知有关该帧的主要自适应码本历史信息的存在。帧类型代码之后则跟有与带有主编码信息和自适应码本历史信息的帧有关的编码单元。

[0164] 当次要历史冗余编码信息被包括在帧 (1030) 内时,跟随在帧开始处字节边界 (1035) 之后的是包括帧类型代码 00(用于常规帧的代码) 的编码单元,而代码 00 之后则跟随有关常规帧的主编码信息。然而,跟随在主编码信息结尾处的字节边界 (1045) 之后,另一编码单元包括帧类型 11,该代码 11 用来指示将有可任选的次要历史信息 (1040) (而不是有关帧的主编码信息) 跟随。因为次要历史信息 (1040) 仅当在前帧丢失时才使用,所以可以给予分组器或其他部件选择省略该信息的可选项。这样做可能出于不同的原因,诸如当整体比特率需要被减少时,分组丢失率较低时,或者在前帧被包含在带有当前帧的分组内时。或者,可以给予多路分配器或其他部件当常规帧 (1030) 被成功接收时选择跳过该次要历史信息的可选项。

[0165] 类似地,当额外码本级冗余编码信息被包括在帧 (1050) 内时,跟随在编码单元开始处的字节边界 (1055) 之后的是帧类型代码 00(用于常规帧的代码),而代码 00 之后则跟随有关常规帧的主编码信息。然而,跟随在主编码信息结尾处的字节边界 (1065) 之后,另一编码单元包括帧类型 01,该代码 01 用来指示将有可任选的额外码本级信息 (1060) 跟随。如同次要历史信息 (1040) 一样,额外码本级信息 (1060) 仅当在前帧丢失时才使用。因此仍如同次要历史信息一样,可以给予分组器或其他部件选择省略该额外码本级信息的可选项,或者可以给予多路分配器或其他部件选择跳过该额外码本级信息的可选项。

[0166] 应用程序 (例如,执行传输层分组的应用程序) 可以决定将多个帧进行组合以形成更大的分组来降低分组报头所需的额外比特。在该分组内部,应用程序通过扫描比特流可以确定帧边界。

[0167] 图 11 示出了具有四个帧 (1110,1120,1130,1140) 的多个分组 (1100) 的可能比特流。可以假设该单个分组内的所有帧在它们中任何一个被接收的情况下都将被接收 (即,没有部分数据讹误),并且自适应码本滞后,或基音通常小于帧长度。在这个例子中,一般不为帧 2(1120)、帧 3(1130)、以及帧 4(1140) 使用任何可任选的冗余编码信息,因为如果当前帧存在则在前帧通常也存在。因此,可以移除分组 (1110) 内用于除了第一帧之外所有帧的可任选冗余编码信息。这样就得到了压缩分组 (1150),其中帧 1(1160) 包括可任选的额外码本级信息,但是所有可任选的冗余编码信息都已经从残留帧 (1170,1180,1190) 中移除。

[0168] 如果编码器使用主要历史冗余编码技术,应用程序会在把各帧一起打包为单个分组时丢掉任何这些比特,,因为无论在前帧是否丢失都要使用该主要历史冗余编码信息。然

而,该应用程序如果知道这一帧将在多帧分组内并且不会是这一分组中的第一帧,则会迫使编码器像编码常规一样地对这一帧进行编码。

[0169] 虽然图 10 和 11 及其相关描述示出了在各帧与信息类型之间的字节对准边界,但是可选地,这些边界也可以不是字节对准的。此外,图 10 和 11 及其相关模式示出了示例性的帧类型代码和帧类型的组合。可选地,编码器和解码器使用其他和 / 或附加的帧类型或帧类型的组合。

[0170] 已经参考描述的实施例描述并示出了本发明的原理,将会认识到所描述的实施例可以在排列和细节上进行修改且不背离这些原理。应该理解除非另外指出,否则在此描述的程序、进程或方法并不关联于或限制于任何特定类型的计算环境。各种类型的通用或专用计算环境都可与根据在此描述教示的操作一并使用或执行。软件中示出的所述实施例的元素也可以由硬件实现,并且反之亦然。

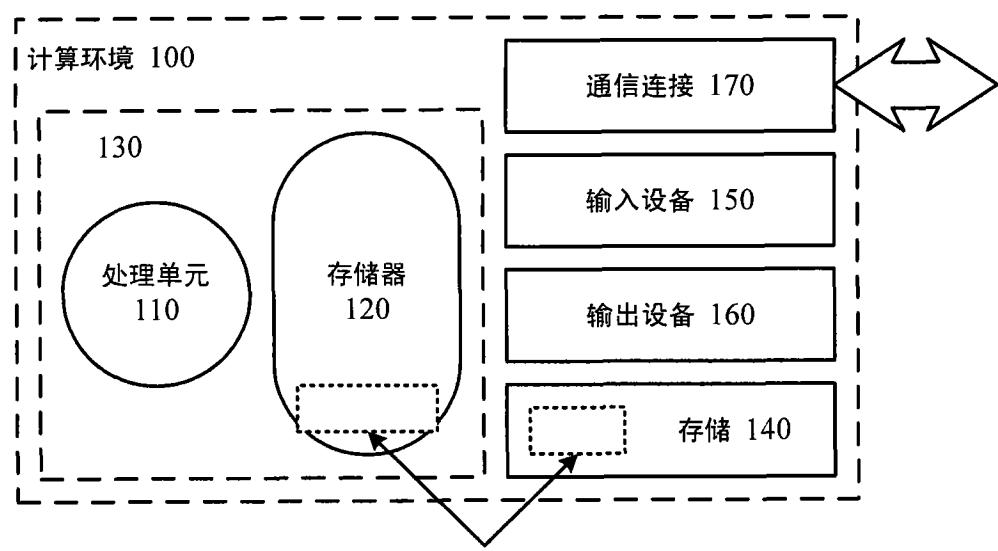


图 1

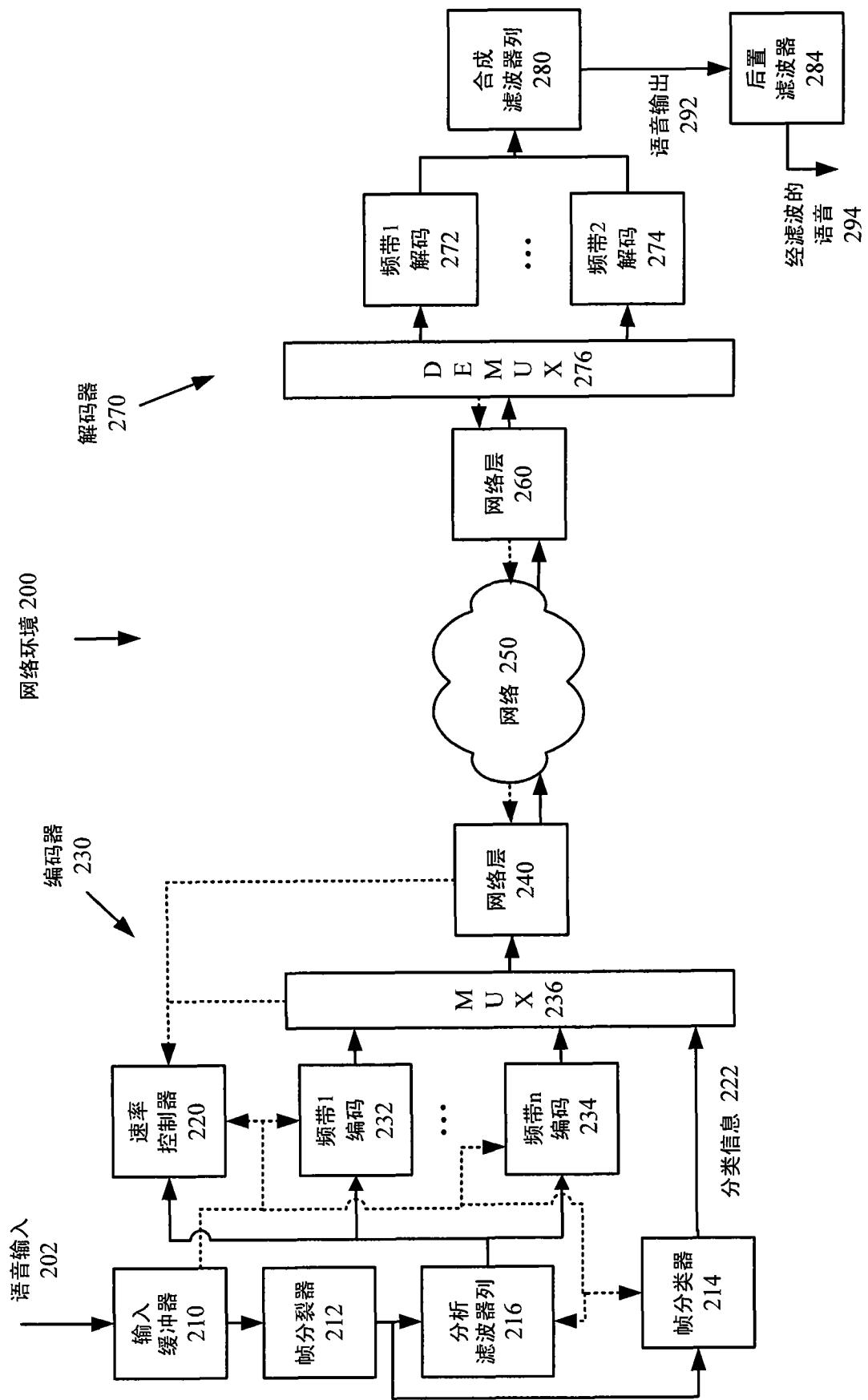


图 2

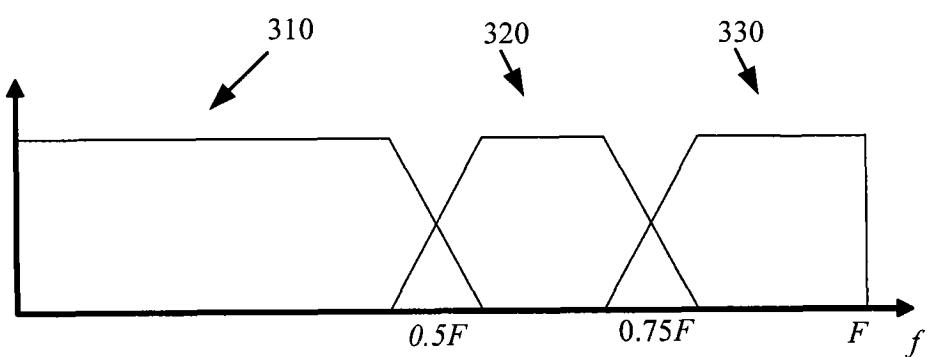


图 3

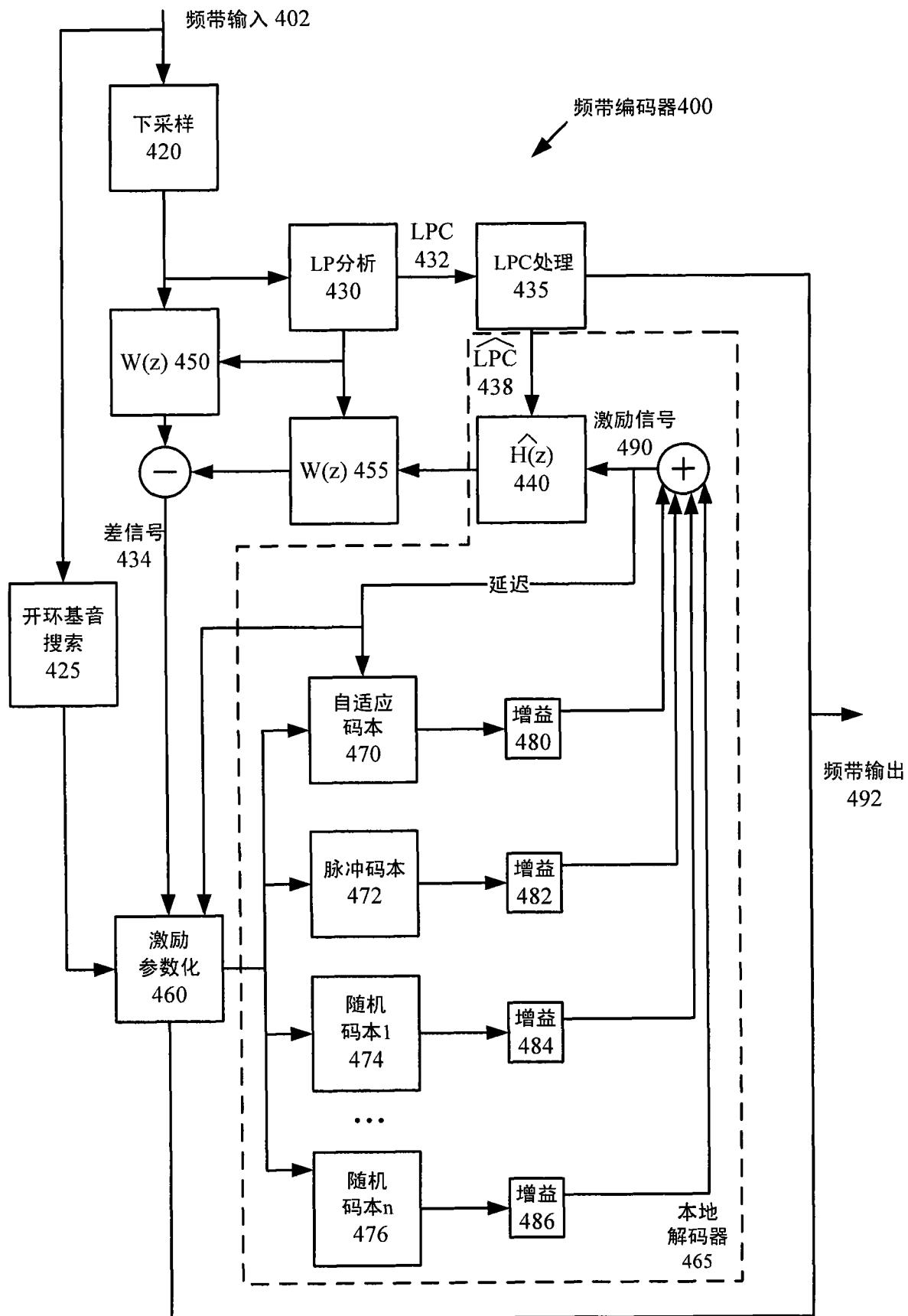


图 4

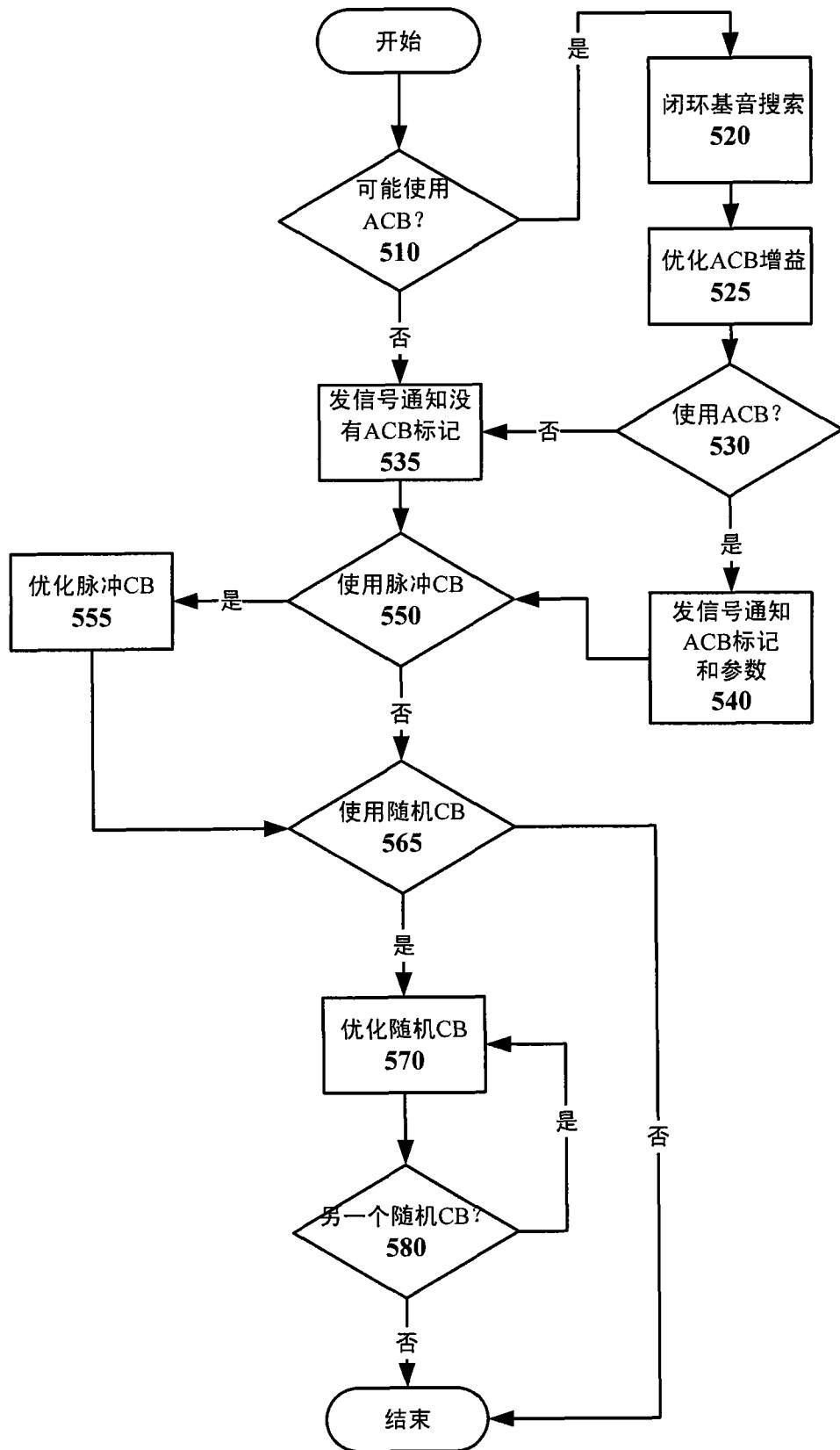


图 5

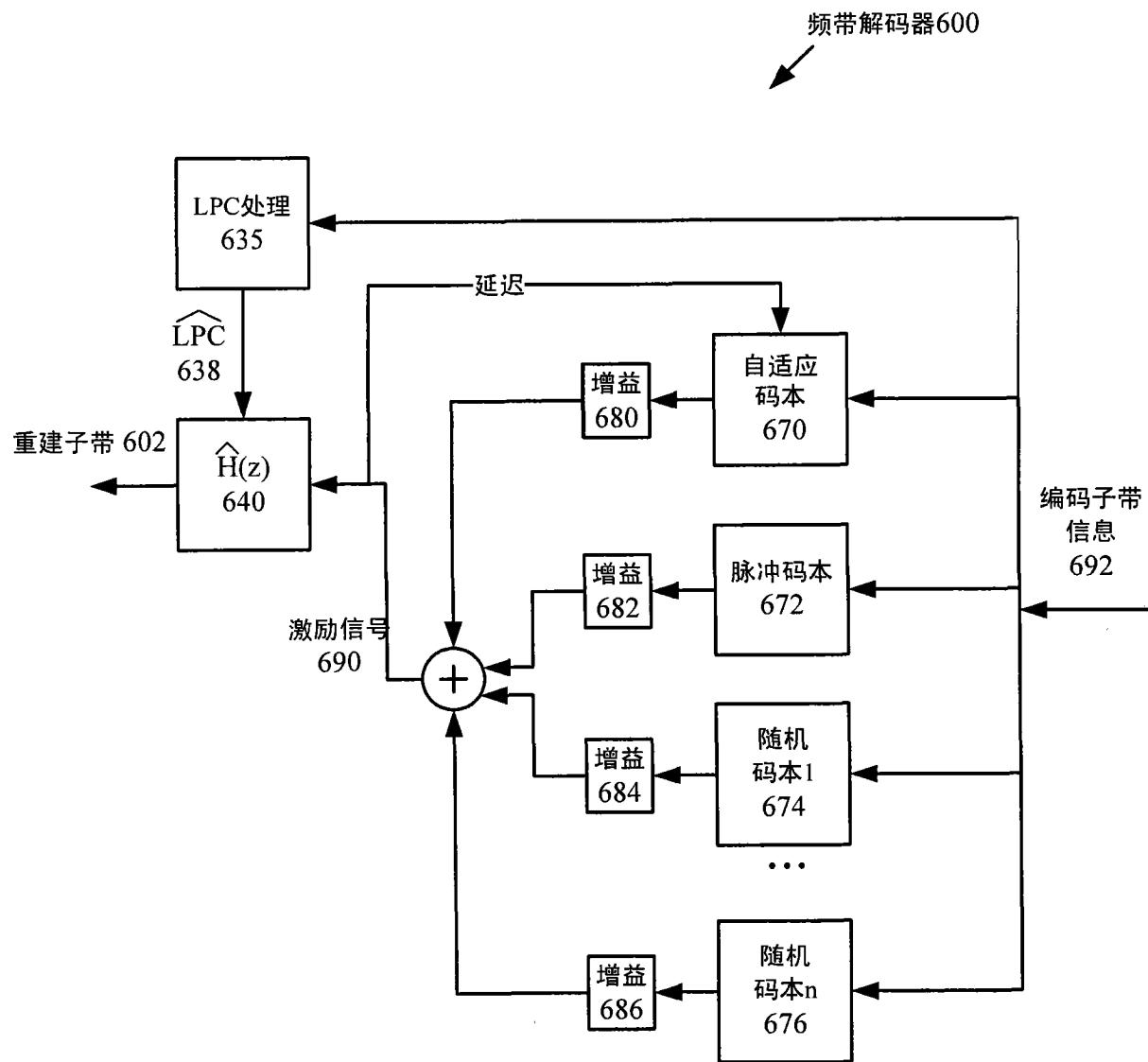


图 6

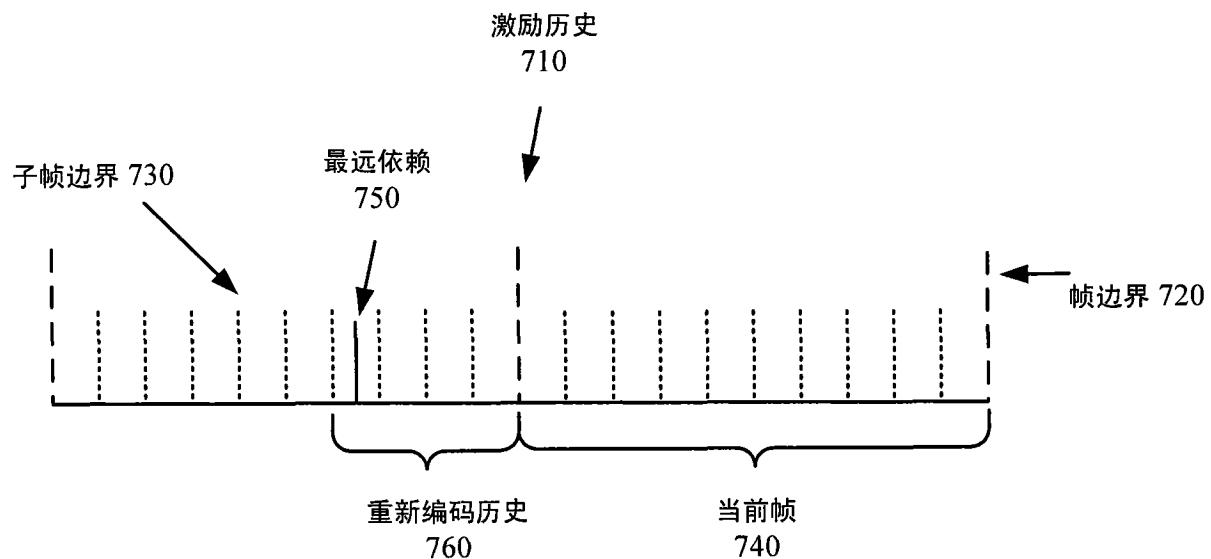


图 7

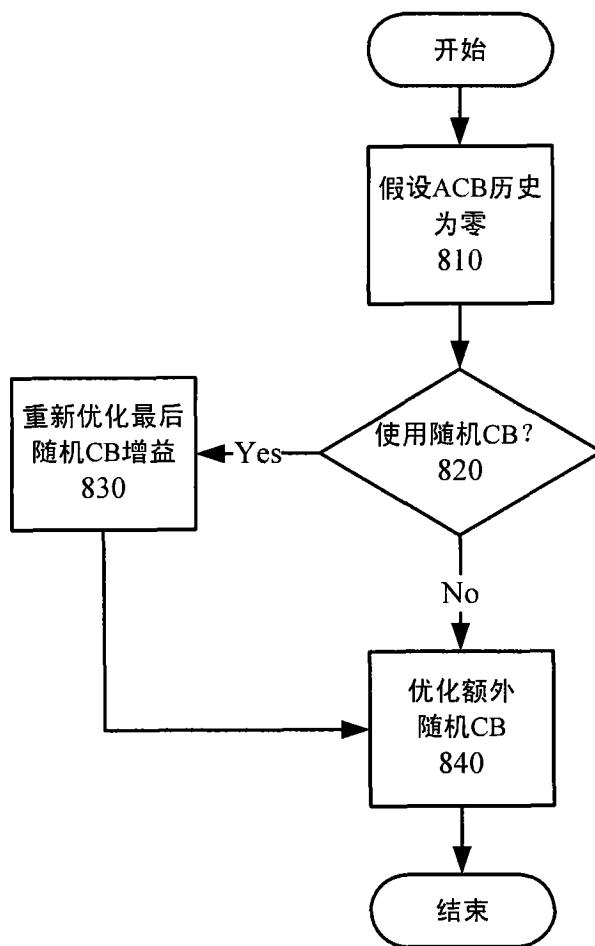


图 8

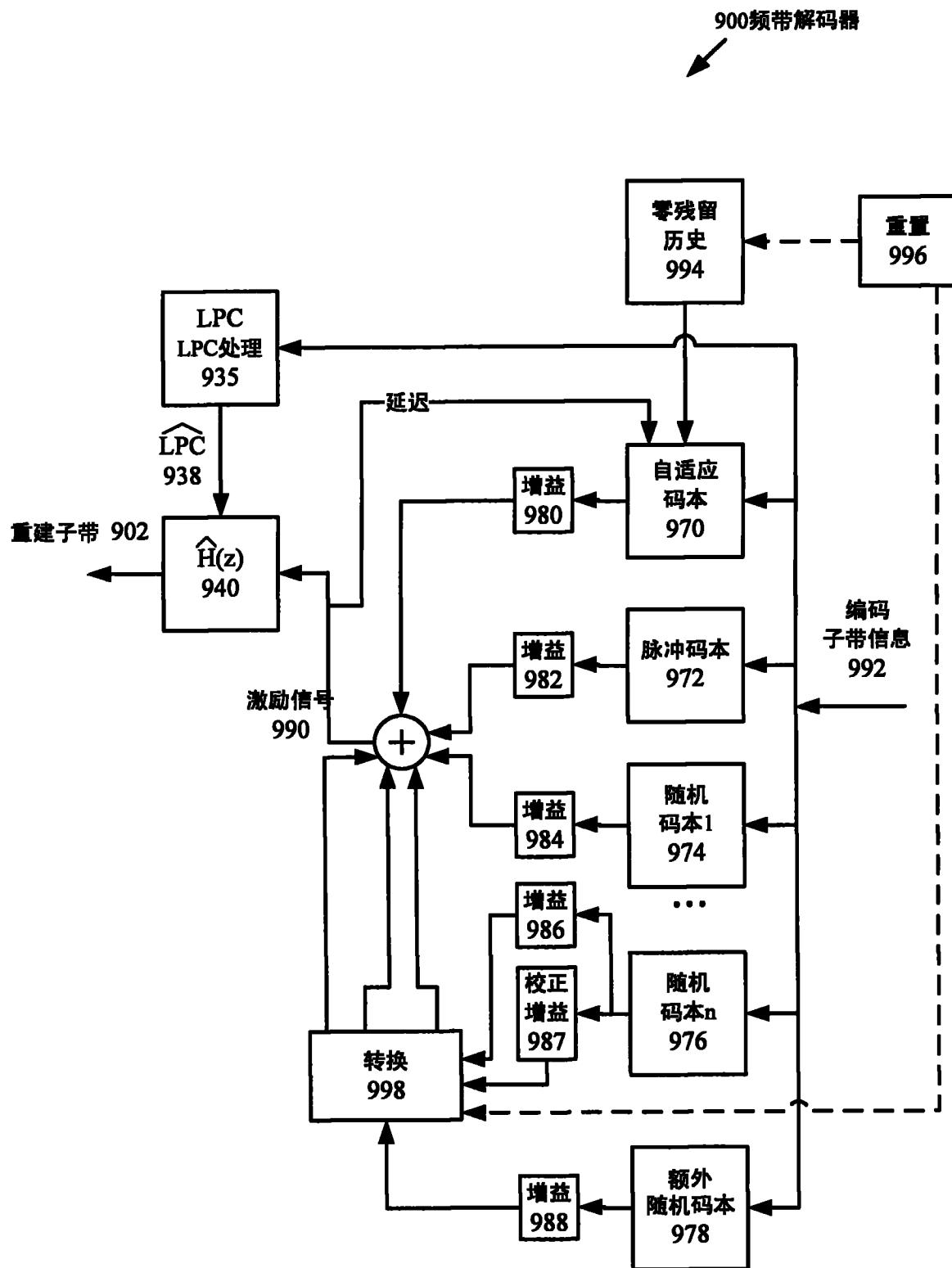


图 9A

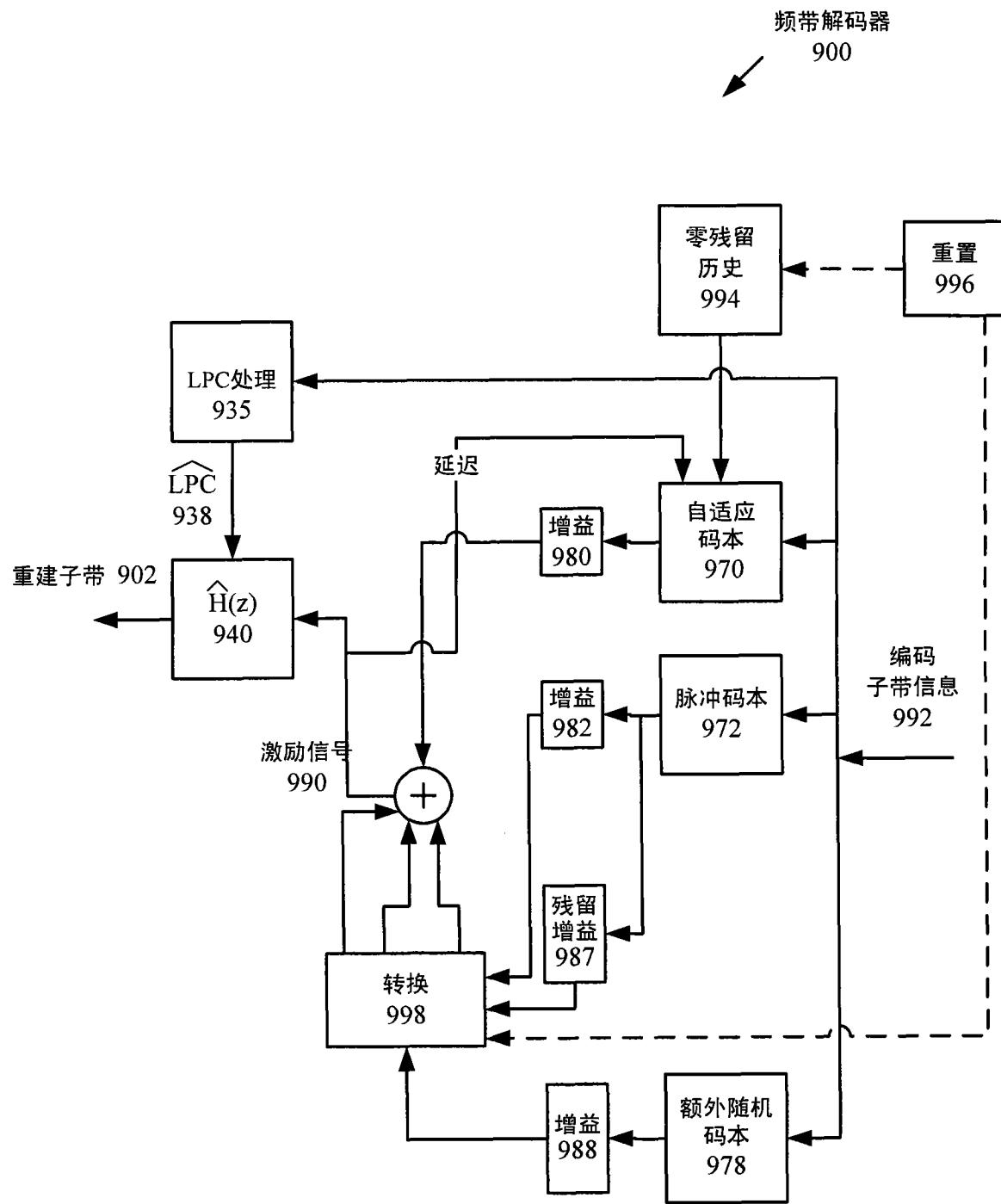


图 9B

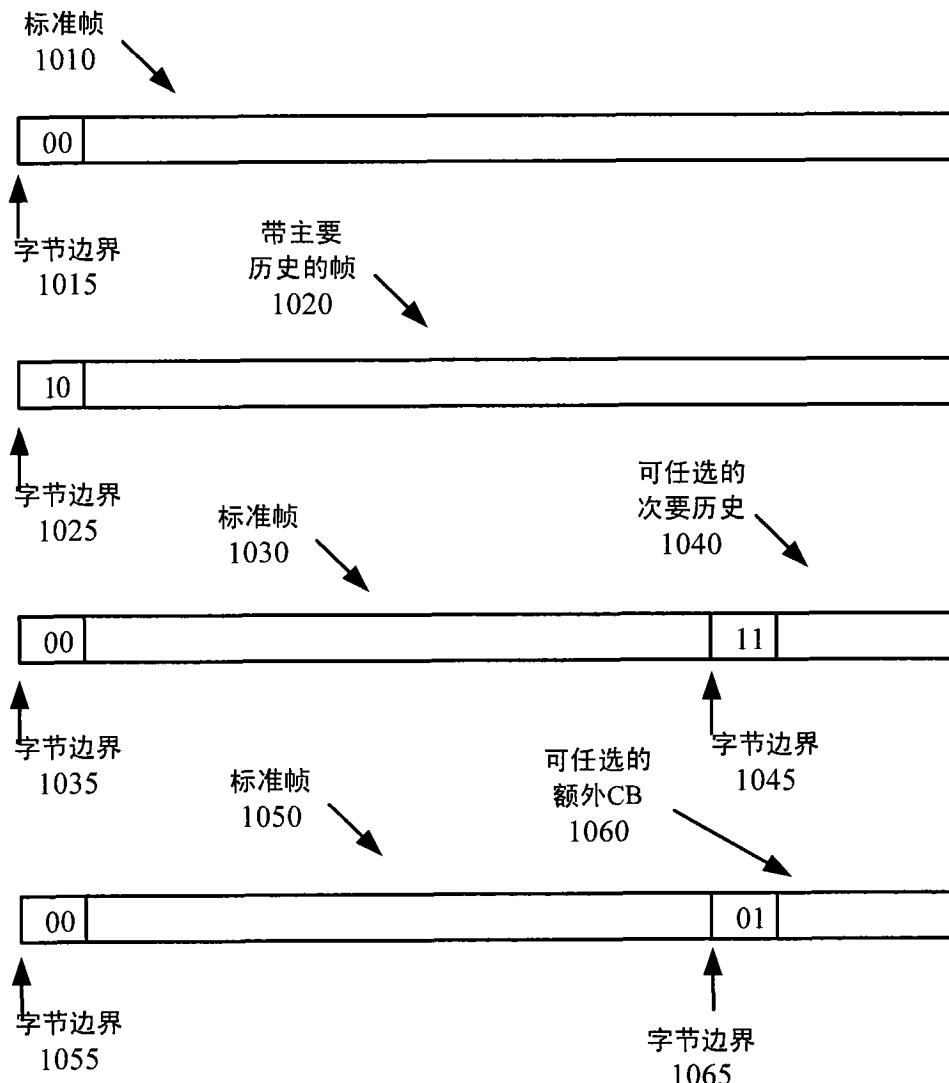


图 10

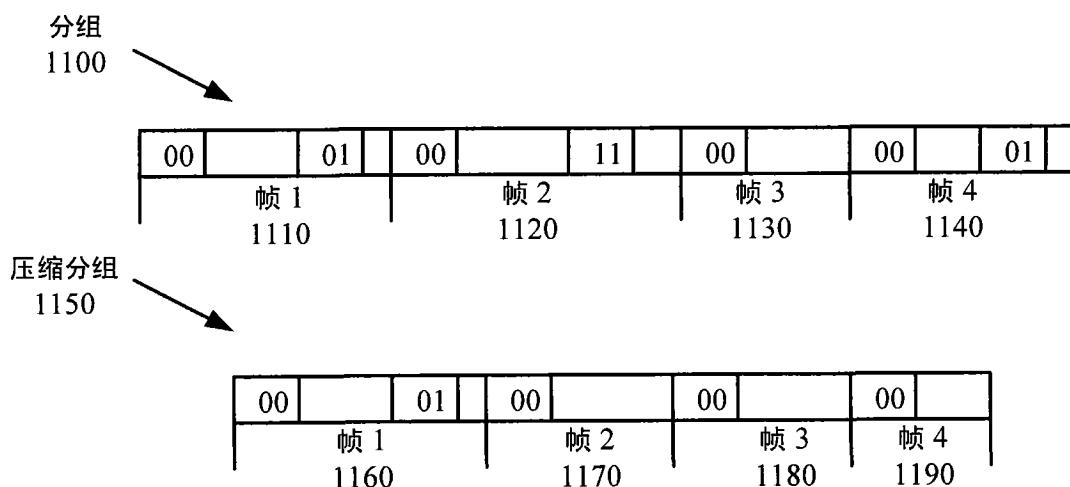


图 11