

(12) 发明专利

(10) 授权公告号 CN 101606196 B

(45) 授权公告日 2012.04.04

(21) 申请号 200880004774.4  
 (22) 申请日 2008.02.01  
 (30) 优先权数据  
 60/901,191 2007.02.14 US  
 12/002,131 2007.12.14 US  
 (85) PCT申请进入国家阶段日  
 2009.08.12  
 (86) PCT申请的申请数据  
 PCT/US2008/001356 2008.02.01  
 (87) PCT申请的公布数据  
 W02008/100385 EN 2008.08.21  
 (73) 专利权人 曼德斯必德技术公司  
 地址 美国加利福尼亚州  
 (72) 发明人 E·施罗默特 Y·高  
 A·拜尼亚斯恩  
 (74) 专利代理机构 北京市中咨律师事务所  
 11247  
 代理人 杨晓光 于静  
 (51) Int. Cl.  
 G10L 19/14 (2006.01)  
 G10L 19/00 (2006.01)  
 (56) 对比文件  
 CN 1650348 A, 2005.08.03, 全文.

US 2005/0004793 A1, 2005.01.06, 说明书第 [0007]-[0013]、[0050] 段和图 1A.

US 2005/0004793 A1, 2005.01.06, 说明书第 [0007]-[0013]、[0050] 段和图 1A.

US 2005/0004793 A1, 2005.01.06, 说明书第 [0007]-[0013]、[0050] 段和图 1A.

Milan Jelinek et al. Advances in Source-controlled variable bit rate wideband speech coding. 《SPECIAL WORKSHOP IN MAUI (SWIM): LECTURES BY MASTERS IN SPEECHPROCESSING》. 2004, 1-8.

Alan McCree et al. An embedded adaptive multi-rate wideband speech coder. 《2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001》. 2001, 第 2 卷 761-764.

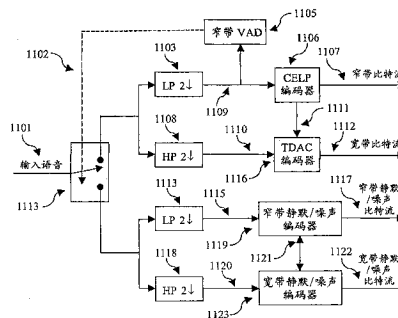
Adil Benyassine et al. ITU-T Recommendation G.729 Annex B: a silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data applications. 《IEEE Communications Magazine》. 1997, 第 35 卷 (第 (续)

审查员 万济萍

权利要求书 3 页 说明书 12 页 附图 12 页

(54) 发明名称  
 嵌入式静默和背景噪声压缩  
 (57) 摘要  
 提供了一种由语音编码器用于编码输入语音信号的方法。所述方法包括：接收所述输入语音信号；判定所述输入语音信号包括活动语音信号还是非活动语音信号；对所述非活动语音信号进行低通滤波以生成窄带非活动语音信号；对所述非活动语音信号进行高通滤波以生成高频带非活动语音信号；使用窄带非活动语音编码器来编码所述窄带非活动语音信号以生成已编码窄带非活动语音；由所述窄带非活动语音编码器根据所述窄带非活动语音信号生成低到高辅助信号；使用宽带非活动语音编码器编码所述高频带非活动语

音信号，以便根据来自所述窄带非活动语音编码器的所述低到高辅助信号生成已编码宽带非活动语音；以及传输所述已编码窄带非活动语音和所述已编码宽带非活动语音。



具有窄带VAD和单独抽取元件的G.729.1的静默/背景噪声编码模式

CN 101606196 B

[ 接上页 ]

(56) 对比文件

9 期), 64-73.

Adil Benyassine et al. ITU-T  
Recommendation G. 729 Annex B: a silence  
compression scheme for use with G. 729

optimized for V. 70 digital simultaneous  
voice and data applications. 《IEEE  
Communications Magazine》. 1997, 第 35 卷 ( 第  
9 期), 64-73.

1. 一种由语音编码器用于编码输入语音信号的方法,所述方法包括:
  - 接收所述输入语音信号;
  - 判定所述输入语音信号包括活动语音信号还是非活动语音信号;
  - 对所述非活动语音信号进行低通滤波以生成窄带非活动语音信号;
  - 对所述非活动语音信号进行高通滤波以生成高频带非活动语音信号;
  - 由宽带非活动语音编码器根据所述高频带非活动语音信号生成高到低辅助信号;
  - 根据来自所述宽带非活动语音编码器的所述高到低辅助信号使用窄带非活动语音编码器来编码所述窄带非活动语音信号以生成已编码窄带非活动语音;
  - 由所述窄带非活动语音编码器根据所述窄带非活动语音信号生成低到高辅助信号;
  - 使用宽带非活动语音编码器编码所述高频带非活动语音信号,以便根据来自所述窄带非活动语音编码器的所述低到高辅助信号生成已编码宽带非活动语音;
  - 传输所述已编码窄带非活动语音和所述已编码宽带非活动语音。
2. 根据权利要求 1 的方法,其中所述传输包括不连续传输 (DTX) 模式。
3. 一种由语音编码器用于编码输入语音信号的方法,所述语音编码器包括宽带非活动语音编码器和窄带非活动语音编码器,所述方法包括:
  - 接收所述输入语音信号;
  - 判定所述输入语音信号包括活动语音信号还是非活动语音信号;
  - 对所述非活动语音信号进行低通滤波以生成窄带非活动语音信号;
  - 对所述非活动语音信号进行高通滤波以生成高频带非活动语音信号;
  - 由所述宽带非活动语音编码器根据所述高频带非活动语音信号生成高到低辅助信号;
  - 使用窄带非活动语音编码器,使用所述高到低辅助信号并根据 ITU-TG. 729 附件 B 建议来编码所述窄带非活动语音信号以生成 G. 729B 已编码窄带非活动语音;
  - 使用所述宽带非活动语音编码器编码所述高频带非活动语音信号以生成已编码宽带非活动语音;
  - 作为 G. 729B 比特流来传输所述 G. 729B 已编码窄带非活动语音;以及
  - 在所述 G. 729B 比特流之后作为宽带基础层比特流来传输所述已编码宽带非活动语音。
4. 根据权利要求 3 的方法,还包括:
  - 编码所述窄带非活动语音信号以生成增强的窄带基础层比特流;
  - 在所述宽带基础层比特流之后传输所述增强的窄带基础层比特流。
5. 根据权利要求 4 的方法,还包括:
  - 编码所述高频带非活动语音信号以生成增强的宽带基础层比特流;
  - 在所述增强的窄带基础层比特流之后传输所述增强的宽带基础层比特流。
6. 根据权利要求 3 的方法,还包括:
  - 编码所述高频带非活动语音信号以生成增强的宽带基础层比特流;
  - 在所述宽带基础层比特流之后传输所述增强的宽带基础层比特流。
7. 根据权利要求 6 的方法,还包括:
  - 编码所述窄带非活动语音信号以生成增强的窄带基础层比特流;

在所述增强的宽带基础层比特流之后传输所述增强的窄带基础层比特流。

8. 一种由语音编码器用于编码输入语音信号的方法,所述方法包括:

接收所述输入语音信号;

对所述输入语音信号进行低通滤波以生成窄带语音信号;

对所述输入语音信号进行高通滤波以生成高频带语音信号;

判定所述窄带语音信号包括活动语音信号还是非活动语音信号;

由宽带非活动语音编码器根据所述高频带语音信号生成高到低辅助信号;

如果判定步骤判定所述窄带语音信号包括所述非活动语音信号,则根据来自所述宽带非活动语音编码器的所述高到低辅助信号使用窄带非活动语音编码器编码所述窄带语音信号以生成已编码窄带非活动语音;

如果判定步骤判定所述窄带语音信号包括所述非活动语音信号,则使用宽带非活动语音编码器编码所述高频带语音信号以生成已编码宽带非活动语音;以及

传输所述已编码窄带非活动语音和所述已编码宽带非活动语音。

9. 根据权利要求 8 的方法,还包括:

由所述窄带非活动语音编码器根据所述窄带语音信号生成低到高辅助信号;

其中所述宽带非活动语音编码器根据来自所述窄带非活动语音编码器的所述低到高辅助信号编码所述高频带语音信号。

10. 根据权利要求 8 的方法,其中所述传输包括不连续传输 (DTX) 模式。

11. 一种适于编码输入语音信号的语音编码器,所述语音编码器包括:

接收器,配置为接收所述输入语音信号;

语音活动检测器,配置为判定所述输入语音信号包括活动语音信号还是非活动语音信号;

低通滤波器,用于对所述非活动语音信号进行低通滤波以生成窄带非活动语音信号;

高通滤波器,用于对所述非活动语音信号进行高通滤波以生成高频带非活动语音信号;

窄带非活动语音编码器,配置为编码所述窄带非活动语音信号以生成已编码窄带非活动语音,并且所述窄带非活动语音编码器还配置为根据所述窄带非活动语音信号生成低到高辅助信号;

宽带非活动语音编码器,配置为编码所述高频带非活动语音信号,以便根据来自所述窄带非活动语音编码器的所述低到高辅助信号生成已编码宽带非活动语音;

传输器,配置为传输所述已编码窄带非活动语音和所述已编码宽带非活动语音;

其中所述宽带非活动语音编码器还配置为根据所述高频带非活动语音信号生成高到低辅助信号,并且其中窄带非活动语音编码器还配置为根据来自所述宽带非活动语音编码器的所述高到低辅助信号编码所述窄带非活动语音信号。

12. 根据权利要求 11 的语音编码器,其中所述传输器配置为根据不连续传输 (DTX) 模式进行传输。

13. 一种适于编码输入语音信号的语音编码器,所述语音编码器包括:

接收器,配置为接收所述输入语音信号;

低通滤波器,用于对所述输入语音信号进行低通滤波以生成窄带语音信号;

高通滤波器,用于对所述输入语音信号进行高通滤波以生成高频带语音信号;

语音活动检测器 VAD,配置为判定所述窄带语音信号包括活动语音信号还是非活动语音信号;

窄带非活动语音编码器,配置为在所述 VAD 判定所述窄带语音信号包括所述非活动语音信号的情况下,编码所述窄带语音信号以生成已编码窄带非活动语音;

宽带非活动语音编码器,配置为在所述 VAD 判定所述窄带语音信号包括所述非活动语音信号的情况下,编码所述高频带语音信号以生成已编码宽带非活动语音;

传输器,配置为传输所述已编码窄带非活动语音和所述已编码宽带非活动语音;

其中所述宽带非活动语音编码器还配置为根据所述高频带语音信号生成高到低辅助信号,并且其中所述窄带非活动语音编码器还配置为根据来自所述宽带非活动语音编码器的所述高到低辅助信号编码所述窄带语音信号。

14. 根据权利要求 13 的语音编码器,其中所述窄带非活动语音编码器还配置为根据所述窄带语音信号生成低到高辅助信号,并且其中所述宽带非活动语音编码器还配置为根据来自所述窄带非活动语音编码器的所述低到高辅助信号编码所述高频带语音信号。

## 嵌入式静默和背景噪声压缩

[0001] 相关申请

[0002] 本申请基于 2007 年 2 月 14 日提交的美国临时申请第 60/901,191 号并要求其优先权,其全部内容在此引入作为参考。

### 技术领域

[0003] 本发明一般地涉及语音编码领域,具体地说,涉及嵌入式静默和背景噪声压缩。

### 背景技术

[0004] 现代电话系统使用数字语音通信技术。在数字语音通信系统中,语音信号被采样并作为数字信号来传输,与普通老式电话系统 (POTS) 中的模拟传输相反。数字语音通信系统的实例包括公共交换电话网络 (PSTN)、已很好地建立的蜂窝网络,以及新兴的基于网际协议的话音传输 (VoIP) 网络。在数字语音通信系统中可以使用各种语音压缩 (或编码) 技术,如 ITU-T 建议 G. 723.1 或 G. 729,以便减小传输语音信号所需的带宽。

[0005] 通过为语音信号的没有实际语音的部分 (如当某人倾听其他谈话者并且未说话时存在的静默时段) 使用较低比特率编码方案,可以实现进一步的带宽减小。语音信号的包括实际语音的部分称为“活动语音”,并且语音信号的不包含实际语音的部分称为“非活动语音”。通常,非活动语音信号包含接听者位置处由麦克风拾取的周围背景噪声。在非常安静的环境中,此周围噪声将非常低并且非活动语音将被感知为静默,而在嘈杂环境中 (例如在机动车中),非活动语音包括环境背景噪声。通常,周围噪声几乎不传达信息,并且因此可以以非常低的比特率来编码和传输。周围噪声的低比特率编码的一种方案仅采用噪声信号的能量 (级别) 和频谱内容。

[0006] 另一种用于带宽减小的常见方案 (其利用背景噪声的静态特性) 是仅发送背景噪声参数的间断更新,而不是发送连续更新。

[0007] 如果传输的比特流具有嵌入式结构,则也可以在网络中实现带宽减小。嵌入式结构意味着比特流包括核心和增强层。可以仅使用核心比特解码和合成语音,而使用增强层比特将提高解码的语音质量。例如,2006 年 5 月的标题为“G. 729-based embedded variable bit-rate coder: An 8-32kbit/sscalable wideband coder bitstream interoperable with G. 729”的 ITU-T 建议 G. 729.1 使用核心窄带层以及若干窄带和宽带增强层,其全部内容在此引入作为参考。

[0008] 处理非常大量的语音信道的网络中的业务拥塞取决于每个编解码器使用的平均比特率而不是每个编解码器使用的最大比特率。例如,假设这样的语音编解码器,其最大比特率为 32Kbps,但是平均比特率为 16Kbps。带宽为 1600Kbps 的网络可以处理约 100 个话音信道,因为平均而言所有 100 个信道将仅使用  $100 * 16\text{Kbps} = 1600\text{Kbps}$ 。显然,在小概率中,传输所有信道的总体所需比特率可能超过 1600Kbps,但是如果该编解码器还采用嵌入式结构,则网络可以通过丢弃多个信道的某些嵌入层来容易地解决此问题。当然,如果网络的规划 / 运行基于每个信道的最大比特率,在不考虑平均比特率和嵌入式结构的情况

下,网络将只能够处理 50 个信道。

## 发明内容

[0009] 根据在此概括地描述的本发明的目的,提供了嵌入式语音编码系统中的静默/背景噪声压缩。在本发明的一个示意性方面中,披露了一种能够生成嵌入式活动语音比特流和嵌入式非活动语音比特流两者的语音编码器。所述语音编码器接收输入语音并且使用语音活动检测器(VAD)来判定所述输入语音是活动语音还是非活动语音。如果所述输入语音是活动语音,则所述语音编码器使用活动语音编码模式来生成活动语音嵌入式比特流,后者包含窄带部分和宽带部分。如果所述输入语音是非活动语音,则所述语音编码器使用非活动语音编码模式来生成非活动语音嵌入式比特流,后者可以包含窄带部分和宽带部分。此外,如果所述输入语音是非活动语音,则所述语音编码器调用其中仅发送静默/背景噪声信息的间断更新的不连续传输(DTX)模式。在解码器侧,接收所述活动和非活动比特流并且根据如所述比特流的大小所指示的比特率类型来调用解码器的不同部分。通过确保平滑地改变带宽来为非活动语音保持带宽连续性,即使非活动语音分组信息指示带宽变化也是如此。

[0010] 通过进一步参考以下附图和说明书,本发明的这些和其他方面将变得显而易见。所有此类附加的系统、方法、特性以及优点都旨在包括在此说明书内、在本发明的范围之内以及由所附权利要求来保护。

## 附图说明

[0011] 在详细查看以下具体说明和附图之后,本发明的特性和优点对于本领域技术人员将变得更加显而易见,这些附图是:

[0012] 图 1 示出了根据本发明的一个实施例的 G. 729. 1 比特流的嵌入式结构;

[0013] 图 2 示出了根据本发明的一个实施例的 G. 729. 1 编码器的结构;

[0014] 图 3 示出了根据本发明的一个实施例的具有窄带编码的 G. 729. 1 编码器的备选操作;

[0015] 图 4 示出了根据本发明的一个实施例的 G. 729. 1 的静默/背景噪声编码模式;

[0016] 图 5 示出了根据本发明的一个实施例的具有嵌入式结构的静默/背景噪声编码器;

[0017] 图 6 示出了根据本发明的一个实施例的静默/背景噪声嵌入式比特流;

[0018] 图 7 示出了根据本发明的一个实施例的备选静默/背景噪声嵌入式比特流;

[0019] 图 8 示出了根据本发明的一个实施例的没有可选层的静默/背景噪声嵌入式比特流;

[0020] 图 9 示出了根据本发明的一个实施例的 G. 729. 1 的窄带工作模式的窄带 VAD;

[0021] 图 10 示出了根据本发明的一个实施例的具有窄带 VAD 的 G. 729. 1 的静默/背景噪声编码模式;

[0022] 图 11 示出了根据本发明的一个实施例的具有窄带 VAD 和单独抽取元件的 G. 729. 1 的静默/背景噪声编码模式;

[0023] 图 12 示出了根据本发明的一个实施例的具有 DTX 模块的静默/背景噪声编码器;

- [0024] 图 13 示出了根据本发明的一个实施例的 G. 729. 1 解码器的结构；
- [0025] 图 14 示出了根据本发明的一个实施例的具有静默 / 背景噪声压缩的 G. 729. 1 解码器；
- [0026] 图 15 示出了根据本发明的一个实施例的具有嵌入式静默 / 背景噪声压缩的 G. 729. 1 解码器；
- [0027] 图 16 示出了根据本发明的一个实施例的具有嵌入式静默 / 背景噪声压缩以及共享的上采样 - 滤波元件的 G. 729. 1 解码器；
- [0028] 图 17 示出了根据本发明的一个实施例的基于比特率的解码器控制流程图操作；
- [0029] 图 18 示出了根据本发明的一个实施例的基于带宽历史的解码器控制流程图操作；
- [0030] 图 19 示出了根据本发明的一个实施例的通用化的话音活动检测器；以及
- [0031] 图 20 示出了具有解码器带宽扩展的窄带静默 / 背景噪声传输。

### 具体实施方式

[0032] 在此可以根据功能块组件和各种处理步骤来描述本发明。应理解的是,可以由任何数量的配置为执行所指定功能的硬件组件和 / 或软件组件来实现此类功能块。例如,本发明可以采用各种集成电路组件,如存储器元件、数字信号处理元件、逻辑元件等,它们可以在一个或多个微处理器或其他控制设备的控制下执行各种功能。此外,应指出的是,本发明可以采用任何数量的用于数据传输、信令、信号处理和调节、音调生成和检测等的常规技术。在此未详细描述可能对本领域技术人员公知的此类一般技术。

[0033] 应理解的是,在此示出和描述的特定实施方式只是示意性的并且并非旨在以任何方式限制本发明的范围。实际上,为了简洁,在此可能未详细描述常规的数据传输、信令和信号处理以及通信系统的其他功能和技术方面(以及系统的单个工作组件的组件)。此外,在此包含的各个附图中示出的连接线旨在表示各个元件之间的示意性功能关系和 / 或物理耦合。应指出的是,实际通信系统中可以存在许多备选或附加的功能关系或物理连接。

[0034] 在分组网络(如蜂窝或 VoIP)中,可以在用户终端(例如,蜂窝手机、软电话、SIP 电话或 WiFi/WiMax 终端)处执行语音信号的编码和解码。在此类应用中,网络服务器仅用于传送包含已编码语音信号信息的分组。分组网络中语音的传输消除了对语音频谱带宽的限制,该限制存在于从 POTS 模拟传输技术继承的 PSTN 中。由于语音信息在分组比特流(其提供了原始语音的数字压缩后的表示)中传输,所以此分组比特流既可以表示窄带语音也可以表示宽带语音。麦克风获取语音信号以及耳机或扬声器在最终终端处再现语音信号(作为窄带或宽带表示)仅取决于此类最终终端的能力。例如,在当前的蜂窝电话中,窄带蜂窝电话获得窄带语音的数字表示并使用窄带编解码器(如自适应多速率(AMR)编解码器)经由蜂窝分组网络与其他类似蜂窝电话进行窄带语音通信。类似地,能够使用宽带的蜂窝电话获得语音的宽带表示并使用宽带语音编解码器(如自适应多速率宽带(AMR-WB))经由蜂窝分组网络与其他能够使用宽带的蜂窝电话进行宽带语音通信。显然,由宽带语音编解码器(如 AMR-WB)提供的较宽频谱内容较之窄带语音编解码器(如 AMR)将改进语音的质量、自然性以及可识度。

[0035] 最新采纳的 ITU-T 建议 G. 729. 1 针对分组网络并且采用嵌入式结构以实现窄带和

宽带语音压缩。所述嵌入式结构使用用于语音的基本质量传输的“核心”语音编解码器以及借助每个附加层改进语音质量的附加编码层。G. 729. 1 的核心基于 ITU-T 建议 G. 729, 后者以 8Kbps 编码窄带语音。此核心非常类似于 G. 729, 具有与 G. 729 比特流兼容的比特流。比特率兼容性意味着由 G. 729 编码器生成的比特流可以被 G. 729. 1 解码器解码并且由 G. 729. 1 编码器生成的比特流可以被 G. 729 解码器解码, 两者都没有任何质量下降。

[0036] 在 8Kbps 的核心之上的 G. 729. 1 的第一增强层是速率为 12Kbps 的窄带层。其次的增强层是十 (10) 个从 14Kbps 到 32Kbps 的宽带层。图 1 示出了具有核心以及 11 个附加层的 G. 729. 1 嵌入式比特流的结构, 其中块 101 表示核心 8Kbps 层, 块 102 表示 12Kbps 的第一窄带增强层, 并且块 103-112 表示十 (10) 个宽带增强层, 分别从 14Kbps 到 32Kbps, 步长为 2Kbps。

[0037] G. 729. 1 的编码器生成包括所有 12 个层的比特流。G. 729. 1 的解码器能够解码任何比特流, 从 8Kbps 核心编解码器的比特流开始, 直到 32Kbps 的包括所有层的比特流。显然, 在接收较高的层时, 解码器将产生更佳质量的语音。解码器还允许从一个帧到下一个帧改变比特率, 且几乎没有由于切换人工因素造成的质量下降。G. 729. 1 的此嵌入式结构使得网络能够解决业务拥塞问题而无需操纵或操作比特流的实际内容。通过丢弃比特流的某些嵌入层部分并且仅传送比特流的其余嵌入层部分来实现拥塞控制。

[0038] 图 2 示出了根据本发明的一个实施例的 G. 729. 1 编码器的结构。以 16KHz 采样输入语音 201 并使其通过低通滤波器 (LPF) 202 和高通滤波器 (HPF) 210, 在分别由抽取元件 203 和 211 下采样后生成窄带语音 204 和基带处的高频带语音 212。注意, 以 8KHz 采样速率来采样窄带语音 204 和基带处的高频带语音 212 两者。窄带语音 204 然后被 CELP 编码器 205 编码以生成窄带比特流 206。窄带比特流被 CELP 解码器 207 解码以生成已解码窄带语音 208, 从窄带语音 204 减去已解码窄带语音 208 以生成窄带残差编码信号 209。窄带残差编码信号和基带处的高频带语音 212 被时域混叠消除 (TDAC) 编码器 213 编码以生成宽带比特流 214。(对于编码高频带信号 212 的模块使用了术语“TDAC 编码器”, 尽管对于 14Kbps 层, 所使用的技术通常称为时域带宽扩展 (TD-BWE)。) 窄带比特流 204 包括 8Kbps 层 101 和 12Kbps 层 102, 而宽带比特流 214 包括层 103-112, 分别从 14Kbps 到 32Kbps。为了简化陈述, 图 2 未示出生成 14Kbps 层的 G. 729. 1 的操作的特殊 TD-BWE 模式。还未示出封装元件, 其接收窄带比特流 206 和宽带比特流 214 以产生图 1 中示出的嵌入式比特流结构。例如, 在标题为“RTP Payload Format for the G. 729. 1 Audio Codec”的互联网工程任务组 (IETF) 请求评注 4749 号 (RFC4749) 中描述了此类封装元件, 其全部内容在此引入作为参考。

[0039] 图 3 中示出了 G. 729. 1 编码器的备选操作模式, 其中仅执行窄带编码。将现在以 8KHz 采样的输入语音 301 输入 CELP 编码器 305, 后者生成窄带比特流 306。类似于图 2, 窄带比特流 306 包括如图 1 所示的 8Kbps 层 101 和 12Kbps 层 102。

[0040] 图 4 提供了根据本发明的一个实施例的具有静默 / 背景噪声编码模式的 G. 729. 1 的一个实施例。为了简洁, 将图 2 中的若干元件合并成图 4 中的单个元件。例如, LPF 202 和抽取元件 203 被合并成 LP- 抽取元件 403, 并且 HPF 210 和抽取元件 211 被合并成 HP- 抽取元件 410。类似地, 图 2 中的 CELP 编码器 205、CELP 解码器 207 以及加法器元件被合并成 CELP 编码器 405。窄带语音 404 类似于窄带语音 204, 高频带语音 412 类似于 212, TDAC

编码器 413 与 213 相同,窄带残差编码信号 409 与 209 相同,窄带比特流 406 与 206 相同,并且宽带比特流 414 与 214 相同。图 4 与图 2 的主要不同是增加了静默 / 背景噪声编码器,其由宽带语音活动检测器 (WB-VAD) 模块 416 控制,模块 416 根据本发明的一个实施例接收输入语音 401 并且操作开关 402。使用术语 WB-VAD 是因为输入语音 401 是以 16KHz 采样的宽带语音。如果 WB-VAD 模块 416 检测到实际语音 (“活动语音”),则输入语音 401 通过开关 402 被导向到典型的 G. 729. 1 编码器,后者在此被称为 “活动语音编码器”。如果 WB-VAD 模块 416 未检测到实际语音,其意味着输入语音 401 是静默或背景噪声 (“非活动语音”),则输入语音 401 被导向到静默 / 背景噪声编码器 416,后者生成静默 / 背景噪声比特流 417。图 4 中未示出比特流多路复用和封装模块,后者基本上类似于由诸如 G. 729 的附件 B 或 G. 723. 1 的附件 A 之类的其他静默 / 背景噪声压缩算法使用的多路复用和封装模块并且是本领域技术人员公知的。

[0041] 可以使用许多方法来使静默 / 背景噪声比特流 417 代表语音的非活动部分。在一种方法中,所述比特流可以在没有任何频带分离和 / 或增强层的情况下代表非活动语音信号。此方法将不允许网络元件操纵静默 / 背景噪声比特流用于拥塞控制,但是这可能不是严重的缺陷,因为传输静默 / 背景噪声比特流所需的带宽非常小。但是,主要的缺点将是使解码器实现带宽控制功能作为静默 / 背景噪声解码器的一部分以在活动语音信号与非活动语音信号之间保持带宽兼容性。图 5 描述了本发明的一个实施例,其包括静默 / 背景噪声 (非活动语音) 编码器,该编码器具有适合于 G. 729. 1 的操作的嵌入式结构,其解决了这些问题。输入非活动语音 501 被馈送到 LP- 抽取元件 503 和 HP- 抽取元件 510 以分别生成窄带非活动语音 504 和基带处的高频带非活动语音 512。窄带静默 / 背景噪声编码器 505 接收窄带非活动语音 504 并且产生窄带静默 / 背景噪声比特流 506。由于静默 / 背景噪声编码器的 G. 729. 1 最小限度操作必须符合 G. 729 的附件 B,所以窄带静默 / 背景噪声比特流 506 必须至少部分地符合 G. 729 的附件 B。窄带静默 / 背景噪声编码器 505 可以与 G. 729 的附件 B 中描述的窄带静默 / 背景噪声编码器相同,但也可以不同,只要其产生与 G. 729 的附件 B (至少部分) 符合的比特流即可。窄带静默 / 背景噪声编码器 505 还可以产生低到高辅助信号 509。低到高辅助信号 509 包含帮助宽带静默 / 背景噪声编码器 513 编码基带中的高频带非活动语音 512 的信息。所述信息可以是窄带重构的静默 / 背景噪声自身或诸如能量 (级别) 或频谱表示之类的参数。宽带静默 / 背景噪声编码器 513 同时接收基带中的高频带非活动语音 512 和辅助信号 509 并产生宽带静默 / 背景噪声比特流 514。宽带静默 / 背景噪声编码器 513 还可以产生高到低辅助信号 508,后者包含帮助窄带静默 / 背景噪声编码器 505 编码窄带语音 504 的信息。类似于图 4,图 5 未示出本领域技术人员公知的比特流多路复用和封装模块。

[0042] 图 6 提供了根据本发明的一个实施例的可以由图 5 的静默 / 背景噪声编码器产生的静默 / 背景噪声嵌入式比特流的描述。静默 / 背景噪声嵌入式比特流 600 包括 :0. 8Kbps 的 G. 729 附件 B (G. 729B) 比特流 601、可选的嵌入式窄带增强比特流 602、宽带基础层比特流 603,以及可选的嵌入式宽带增强比特流 604。相对于图 5,窄带静默 / 背景噪声比特流 506 包括 G. 729B 比特流 601 和可选的窄带嵌入式比特流 602。此外,图 5 中的宽带静默 / 背景噪声比特流 514 包括宽带基础层比特流 603 和可选的宽带嵌入式比特流 604。G. 729B 比特流 601 的结构由 G. 729 附件 B 定义。其包括 10 个用于表示频谱的位和 5 个用于表示能

量（级别）的位。可选的窄带嵌入式比特流 602 包括频谱和能量的改进量化表示（例如，频谱表示的附加码本级或能量量化的改进的时间分辨率）、随机种子信息，或实际量化波形信息。宽带基础层比特流 603 包含高频带静默 / 背景噪声信号的表示的量化信息。该信息可以包括能量信息以及频谱信息，格式为线性预测编码（LPC）格式、子带格式，或其他线性变换系数，如离散傅立叶变换（DFT）、离散余弦变换（DCT）或小波变换。宽带基础层比特流 603 还可以例如包含随机种子信息或实际量化波形信息。可选的宽带嵌入式比特流 604 可以包括宽带基础层比特流 603 中未包括的附加信息，或包括宽带基础层比特流 603 中包括的相同信息的改进的分辨率。

[0043] 图 7 提供了根据本发明的一个实施例的静默 / 背景噪声嵌入式比特流的一个备选实施例。在此备选实施例中，位字段的顺序不同于图 6 示出的实施例，但是各位中的实际信息在两个实施例中是相同的。类似于图 6，静默 / 背景噪声嵌入式比特流 700 的第一部分是 G. 729B 比特流 701，但是第二部分是宽带基础层比特流 703，随后是可选的嵌入式窄带增强比特流 702，然后是可选的嵌入式宽带增强比特流 704。

[0044] 图 6 中的实施例与图 7 中的备选实施例的主要不同是网络的比特流截断的影响。对于图 6 中描述的实施例，网络的比特流截断将除去所有宽带字段，然后才除去任何窄带字段。另一方面，对于图 7 中描述的备选实施例，比特流截断将除去宽带和窄带两者的附加嵌入式增强字段，然后才除去任何基础层字段（窄带或宽带）。

[0045] 如果未将可选的增强层结合到 G. 729. 1 的静默 / 背景噪声嵌入式比特流中，则比特流 600 和 700 变为相同。图 8 示出了此类比特流，其仅包括 G. 729B 比特流 801 和宽带基础层比特流 803。尽管此比特流不包括可选的嵌入式层，但是其仍保持嵌入式结构，其中网络元件可以除去宽带基础层比特流 803，同时保留 G. 729B 比特流 801。在另一选择中，G. 729B 比特流 801 可以是非活动语音的编码器传输的唯一比特流，即使在活动语音编码器传输了包括窄带和宽带信息两者的嵌入式比特流时也是如此。在此情况下，如果解码器接收到活动语音的完整嵌入式比特流，但是仅接收到非活动语音的窄带比特流，则其可以执行合成后的非活动语音的带宽扩展，以便针对合成后的输出信号实现平滑的感知质量。

[0046] 根据图 4 运行静默 / 背景噪声编码模式的主要问题之一在于到 WB-VAD 416 的输入是宽带输入语音 401。因此，如果希望仅使用 G. 279. 1 的窄带操作模式（如图 3 所述），但是具有静默 / 背景噪声编码模式，则应使用可以针对窄带信号进行操作的另一 VAD。

[0047] 一种可能的解决方案是针对 G. 279. 1 的特定窄带操作模式使用特殊的窄带 VAD (NB-VAD)。图 9 描述了根据本发明的一个实施例的此类解决方案，其中窄带输入语音 901 是到控制开关 902 的 NB-VAD 916 的输入。无论 NB-VAD 916 检测到活动语音还是非活动语音，输入语音 901 都分别被路由到 CELP 编码器 905 或窄带静默 / 背景噪声编码器 916。CELP 编码器 905 生成窄带比特流 906 并且窄带静默 / 背景噪声编码器 916 生成窄带静默 / 背景噪声比特流 917。G. 729. 1 的此模式的总体操作非常类似于 G. 729 的附件 B，并且窄带静默 / 背景噪声比特流 917 应部分地或完全地兼容 G. 729 的附件 B。此方案的主要缺点是需要将 WB-VAD 416 和 NB-VAD 916 都结合在 G. 729. 1 静默 / 背景噪声压缩模式的标准和代码中。

[0048] 在频谱的窄带部分（最高到 4KHz）以及在频谱的高频带部分（从 4KHz 到 7KHz）中，活动语音对非活动语音的特征和特性是明显的。此外，多数能量和其他典型语音特性

(如谐波结构)更多地支配窄带部分而不是高频带部分。因此,可以完全使用语音的窄带部分执行语音活动检测。图 10 示出了根据本发明的一个实施例的具有窄带 VAD 的 G. 729. 1 的静默/背景噪声编码模式。LP- 抽取 1002 和 HP- 抽取 1010 元件接收输入语音 1001 以分别生成窄带语音 1003 和基带处的高频带语音 1012。窄带 VAD 1004 使用窄带语音 1003 来生成控制开关 1008 的话音活动检测信号 1005。如果话音活动信号 1005 指示活动语音,则窄带信号 1003 被路由到 CELP 编码器 1006 并且基带中的高频带信号 1012 被路由到 TDAC 编码器 1016。CELP 编码器 1006 生成窄带比特流 1007 和窄带残差编码信号 1009。窄带残差编码信号 1009 用作到生成宽带比特流 1014 的 TDAC 编码器 1016 的第二输入。如果话音活动信号 1005 指示非活动语音,则窄带信号 1003 被路由到窄带静默/背景噪声编码器 1017 并且基带中的高频带信号 1012 被路由到宽带静默/背景噪声编码器 1020。窄带静默/背景噪声编码器 1017 生成窄带静默/背景噪声比特流 1016 并且宽带静默/背景噪声编码器 1020 生成宽带静默/背景噪声比特流 1019。双向辅助信号 1018 代表在窄带静默/背景噪声编码器 1017 与宽带静默/背景噪声编码器 1020 之间交换的辅助信息。

[0049] 图 10 中示出的系统的基础假设是分别由 LP- 抽取 1002 和 HP- 抽取 1010 元件生成的窄带信号 1003 和高频带信号 1012 适合于活动语音编码和非活动语音编码两者。图 11 描述了与图 10 中提供的系统类似的系统,但是使用不同的 LP- 抽取和 HP- 抽取元件用于活动语音编码和非活动语音编码的语音的预处理。例如,如果活动语音编码器的截止频率不同于非活动语音编码器的截止频率,则会出现这种情况。活动语音 LP- 抽取元件 1003 接收输入语音 1101 以产生窄带语音 1109。窄带 VAD 1105 使用窄带语音 1109 来生成控制开关 1113 的话音活动检测信号 1102。如果话音活动信号 1102 指示活动语音,则输入信号 1101 被路由到活动语音 LP- 抽取元件 1103 和活动语音 HP- 抽取元件 1108 以分别生成活动语音窄带信号 1109 和活动语音基带中的高频带信号 1110。如果话音活动信号 1102 指示非活动语音,则输入信号 1101 被路由到非活动语音 LP- 抽取元件 1113 和非活动语音 HP- 抽取元件 1118 以生成非活动语音窄带信号 1115 和非活动语音基带中的高频带信号 1120。应指出的是,将开关 1113 示为作用于输入信号 1101 只是为了图 11 的简洁和简化。实际上,输入语音 1101 可以被连续馈送到所有四个抽取单元 (1103、1108、1113 以及 1118) 并且针对四个输出信号 (1109、1110、1115 以及 1120) 执行实际切换。NB-VAD 1105 可以使用活动语音窄带信号 1109 (如图 11 所示),也可以使用非活动语音窄带信号 1115。类似于图 10,活动语音窄带信号 1109 被路由到 CELP 编码器 1106,后者生成窄带比特流 1107 和窄带残差编码信号 1111。TDAC 编码器 1116 接收活动语音基带中的高频带信号 1110 和窄带残差编码信号 1111 以生成宽带比特流 1112。此外,非活动语音窄带信号 1115 被路由到生成窄带静默/背景噪声比特流 1117 的窄带静默/背景噪声编码器 1119。宽带静默/背景噪声编码器 1123 接收非活动语音高频带信号 1120 并且生成宽带静默/背景噪声编码器 1122。双向辅助信号 1121 代表在窄带静默/背景噪声编码器 1119 与宽带静默/背景噪声编码器 1123 之间交换的信息。

[0050] 由于非活动语音(其包括静默或背景噪声)保存的信息远少于活动语音,所以表示非活动语音所需的位数远小于用于描述活动语音的位数。例如,G. 729 使用 80 个位描述 10 毫秒的活动语音帧,但是仅使用 16 个位描述 10 毫秒的非活动语音帧。此减少的位数有助于减小传输比特流所需的带宽。对于某些非活动语音帧,如果根本不发送信息,则进一步

的减小是可能的。此方法被称为不连续传输 (DTX) 并且其中不传输信息的帧被简单地称为非传输 (NT) 帧。如果 NT 帧中的输入语音特征与先前发送的信息 (其可以是过去的若干帧) 相比没有显著变化, 则这是可能的。在此类情况下, 解码器可以根据先前接收的信息生成 NT 帧的输出非活动语音信号。图 12 示出了根据本发明的一个实施例的具有 DTX 模块的静默 / 背景噪声编码器。该静默 / 背景噪声编码器的结构和操作非常类似于作为图 11 的一部分描述的静默 / 背景噪声编码器。输入非活动语音 1201 被路由到非活动语音 LP- 抽取 1203 和非活动语音 HP- 抽取 1216 元件以分别生成窄带非活动语音 1205 和基带中的高频带非活动语音 1218。此外, 窄带非活动语音 1205 被路由到窄带静默 / 背景噪声编码器 1206, 后者生成窄带静默 / 背景噪声比特流 1207。宽带静默 / 背景噪声编码器 1220 接收基带中的高频带非活动语音 1218 并且生成宽带静默 / 背景噪声比特流 1222。双向辅助信号 1214 代表在窄带静默 / 背景噪声编码器 1206 与宽带静默 / 背景噪声编码器 1220 之间交换的信息。主要不同在于引入了生成 DTX 控制信号 1213 的 DTX 元件 1212。窄带静默 / 背景噪声编码器 1206 和宽带静默 / 背景噪声编码器 1220 接收 DTX 控制信号 1213, 后者指示何时发送窄带静默 / 背景噪声比特流 1207 和宽带静默 / 背景噪声比特流 1222。图 12 中未示出的更高级的 DTX 元件可以产生指示何时发送窄带静默 / 背景噪声比特流 1207 的窄带 DTX 控制信号, 以及产生指示何时发送宽带静默 / 背景噪声比特流 1222 的单独的宽带 DTX 控制信号。在此实例实施例中, DTX 元件 1212 可以使用若干输入, 包括输入非活动语音 1201、窄带非活动语音 1205、基带中的高频带非活动语音 1218 以及时钟 1210。DTX 元件 1212 还可以使用由 VAD 模块 (图 11 中示出, 但是图 12 中省略) 计算的语音参数, 以及使用由系统中的任何编码元件 (活动语音编码元件或者非活动语音编码元件) 计算的参数 (为了简单和清晰, 图 12 省略了这些参数路径)。DTX 元件 1212 中实现的 DTX 算法决定何时需要静默 / 背景信息的更新。可以例如基于任何 DTX 输入参数 (例如, 输入非活动语音 1201 的电平) 或基于时钟 1210 测量的时间间隔来做出所述决定。针对静默 / 背景信息的更新发送的比特流被称为静默插入描述 (SID)。

[0051] DTX 方法还可以用于图 4 中示出的非嵌入式静默压缩。类似地, DTX 方法还可以用于图 9 中示出的 G. 729. 1 的窄带操作模式。用于封装比特流并从编码器侧向解码器侧传输比特流以及用于由解码器侧接收和解封装比特流的通信系统是本领域技术人员公知的并且因此不在此进行详细描述。

[0052] 图 13 示出了 G. 729. 1 的典型解码器, 其解码图 2 中呈现的比特流。CELP 解码器 1303 接收窄带比特流 1301 并且 TDAC 解码器 1316 接收宽带比特流 1314。TDAC 解码器 1316 生成基带处的高频带信号 1317 以及生成由 CELP 解码器 1303 接收的重构加权差分信号 1312。CELP 解码器 1303 生成窄带信号 1304。上采样元件 1305 和低通滤波器 1307 处理窄带信号 1304 以生成窄带重构语音 1309。上采样元件 1318 和高通滤波器 1320 处理基带处的高频带信号 1317 以生成高频带重构语音 1322。将窄带重构语音 1309 和低频带重构语音 1322 相加以生成输出重构语音 1324。类似于以上对编码器的讨论, 为解码宽带比特流 1314 的模块使用术语“TDAC 解码器”, 尽管对于 14Kbps 层, 使用的技术通常称为时域带宽扩展 (TD-BWE)。

[0053] 图 14 提供了根据本发明的一个实施例的具有静默 / 背景噪声压缩的 G. 729. 1 解码器的描述, 该解码器适于接收和解码具有如图 4 所示的静默 / 背景噪声压缩的 G. 729. 1

编码器生成的比特流。图 14 的顶部（其描述了活动语音解码器）与图 13 相同且上采样和滤波元件合二为一。CELP 解码器 1403 接收窄带比特流 1401 并且 TDAC 解码器 1416 接收宽带比特流 1414。TDAC 解码器 1416 生成基带处的高频带活动语音 1417 以及生成由 CELP 解码器 1403 接收的重构加权差分信号 1412。CELP 解码器 1403 生成窄带活动语音 1404。上采样 -LP 元件 1405 处理窄带活动语音 1404 以生成窄带重构活动语音 1409。上采样 -HP 元件 1418 处理基带处的高频带活动语音 1417 以生成高频带重构活动语音 1422。将窄带重构活动语音 1409 和 高频带重构活动语音 1422 相加以生成重构活动语音 1424。图 14 的底部提供了静默 / 背景噪声（非活动语音）解码的描述。静默 / 背景噪声比特流 1431 被静默 / 背景噪声解码器 1433 接收, 后者生成宽带重构非活动语音 1434。由于活动语音解码器既可以生成宽带信号也可以生成窄带信号（取决于网络保留的嵌入层的数量），所以确保在最终重构输出语音 1429 中不会听到带宽切换造成的感知人工因素很重要。因此, 宽带重构非活动语音 1434 被馈送到带宽 (BW) 自适应模块 1436, 后者通过将其带宽与重构活动语音 1429 的带宽相匹配来生成重构非活动语音 1438。可以通过比特流解封装模块（未示出）或通过例如在 CELP 解码器 1403 和 TDAC 解码器 1416 的操作内可从活动语音解码器提供的信息, 将活动语音带宽信息提供给 BW 自适应模块 1436。还可以在重构活动语音 1424 上直接测量活动语音带宽信息。在最后的步骤, 根据 VAD 信息 1426, 该信息指示是接收到活动比特流（包括窄带比特流 1401 和宽带比特流 1414）还是静默 / 背景噪声比特流, 开关 1427 分别在重构活动语音 1424 和重构非活动语音 1438 之间进行选择, 以便形成重构输出语音 1429。

[0054] 图 15 提供了根据本发明的一个实施例的具有嵌入式静默 / 背景噪声压缩的 G. 729. 1 解码器的描述, 该解码器适于接收和解码具有例如如图 10 和 11 所示的嵌入式静默 / 背景噪声压缩的 G. 729. 1 编码器生成的比特流。图 15 的顶部（其描述了活动语音解码器）与图 13 和 14 相同且上采样和滤波元件合二为一。活动语音 CELP 解码器 1503 接收窄带比特流 1501 并且活动语音 TDAC 解码器 1516 接收宽带比特流 1514。活动语音 TDAC 解码器 1516 生成基带处的高频带活动语音 1517 以及生成由活动语音 CELP 解码器 1503 接收的活动语音重构加权差分信号 1512。活动语音 CELP 解码器 1503 生成窄带活动语音 1504。活动语音上采样 -LP 元件 1505 处理窄带活动语音 1504 以生成窄带重构活动语音 1509。活动语音上采样 -HP 元件 1518 处理基带处的高频带活动语音 1517 以生成高频带重构活动语音 1522。将窄带重构活动语音 1509 和 高频带重构活动语音 1522 相加以生成重构活动语音 1524。图 15 的底部描述了非活动语音解码器。窄带静默 / 背景噪声解码器 1533 接收窄带静默 / 背景噪声比特流 1531 并且宽带静默 / 背景噪声解码器 1536 接收静默 / 背景噪声宽带比特流 1534。窄带静默 / 背景噪声解码器 1533 生成静默 / 背景噪声窄带信号 1534 并且宽带静默 / 背景噪声解码器 1536 生成静默 / 背景噪声基带处的高频带信号 1537。双向辅助信号 1532 代表在窄带静默 / 背景噪声解码器 1533 与宽带静默 / 背景噪声解码器 1536 之间交换的信息。静默 / 背景噪声上采样 -LP 元件 1535 处理静默 / 背景噪声窄带信号 1534 以生成静默 / 背景噪声窄带重构信号 1539。静默 / 背景噪声上采样 -HP 元件 1538 处理静默 / 背景噪声基带处的高频带信号 1537 以生成静默 / 背景噪声高频带重构信号 1542。将静默 / 背景噪声窄带重构信号 1539 与静默 / 背景噪声高频带重构信号 1542 相加以生成重构非活动语音 1544。根据 VAD 信息 1526, 该信息指示是接收到活动比特流（包括窄带比特

流 1501 和宽带比特流 1514) 还是非活动比特流 (包括窄带静默 / 背景噪声比特流 1531 和静默 / 背景噪声宽带比特流 1534), 开关 1527 分别在重构活动语音 1524 和重构非活动语音 1544 之间进行选择, 以便形成重构输出语音 1529。显然, 切换的顺序和相加的顺序是可互换的, 并且另一实施例可以是一个开关在窄带信号之间进行选择而另一个开关在宽带信号之间进行选择, 同时信号相加元件组合开关的输出。

[0055] 在图 15 中, 假设需要不同的处理 (例如, 不同的截止频率), 则活动语音和非活动语音的上采样 -LP 和上采样 -HP 元件是不同的。如果在活动语音和非活动语音之间, 上采样 -LP 和上采样 -HP 元件中的处理是相同的, 则可以为两种类型的语音使用相同的元件。图 16 描述了具有嵌入式静默 / 背景噪声压缩的 G. 729. 1 解码器, 其中在活动语音和非活动语音之间共享上采样 -LP 和上采样 -HP 元件。活动语音 CELP 解码器 1603 接收窄带比特流 1601 并且活动语音 TDAC 解码器 1616 接收宽带比特流 1614。活动语音 TDAC 解码器 1616 生成基带处的高频带活动语音 1617 以及生成由活动语音 CELP 解码器 1603 接收的活动语音重构加权差分信号 1612。活动语音 CELP 解码器 1603 生成窄带活动语音 1604。窄带静默 / 背景噪声解码器 1633 接收窄带静默 / 背景噪声比特流 1631 并且宽带静默 / 背景噪声解码器 1636 接收静默 / 背景噪声宽带比特流 1635。窄带静默 / 背景噪声解码器 1633 生成静默 / 背景噪声窄带信号 1634 并且宽带静默 / 背景噪声解码器 1636 生成静默 / 背景噪声基带处的高频带信号 1637。双向辅助信号 1632 代表在窄带静默 / 背景噪声解码器 1633 与宽带静默 / 背景噪声解码器 1636 之间交换的信息。根据 VAD 信息 1641, 开关 1619 将窄带活动语音 1604 或静默 / 背景噪声窄带信号 1634 导向到上采样 -LP 元件 1642, 后者产生窄带输出信号 1643。类似地, 根据 VAD 信息 1641, 开关 1640 将基带处的高频带活动语音 1617 或静默 / 背景噪声基带处的高频带信号 1637 导向到上采样 -HP 元件 1644, 后者产生高频带输出信号 1645。将窄带输出信号 1643 和 1645 相加以产生重构输出语音 1646。

[0056] 图 14、15 和 16 中描述的静默 / 背景噪声解码器可以备选地结合根据本发明的备选实施例的 DTX 解码算法, 其中从先前接收的参数外推用于生成重构非活动语音的参数。外推过程对于本领域技术人员是公知的并且不在此进行详细描述。但是, 如果编码器针对窄带非活动语音使用一种 DTX 模式并且编码器针对高频带非活动语音使用另一种 DTX 模式, 则窄带静默 / 背景噪声解码器处的更新和外推将不同于宽带静默 / 背景噪声解码器处的更新和外推。

[0057] 根据所接收的比特流的类型, 具有静默 / 背景噪声压缩的 G. 729. 1 解码器以多种不同的模式工作。所接收的比特流中的比特数 (大小) 确定了接收的嵌入式层的结构 (即, 比特率), 但是所接收的比特流中的比特数还形成解码器处的 VAD 信息。例如, 如果 G. 729. 1 分组 (其代表 20 毫秒的语音) 拥有 640 比特, 则解码器将确定其是 32Kbps 的活动语音分组并且将调用完整的活动语音宽带解码算法。另一方面, 如果该分组拥有 240 比特来代表 20 毫秒的语音, 则解码器将确定其是 12Kbps 的活动语音分组并且将仅调用活动语音窄带解码算法。对于带有静默 / 背景压缩的 G. 729. 1, 如果分组的大小为 32 比特, 则解码器将确定其是仅带有窄带信息的非活动语音分组并且将调用非活动语音窄带解码算法, 但是如果分组的大小为 0 比特 (即, 没有到达的分组), 则其将被视为 NT 帧并且将使用适当的外推算法。比特流大小的变化是由语音编码器 (其根据输入信号使用活动或非活动语音编码) 或

网络元件（其通过截断某些嵌入式层来减小拥塞）造成的。图 17 提供了基于比特率（如接收的分组中的比特流的大小所确定的）的解码器控制操作的流程图。假设活动语音比特流的结构如图 1 所示并且非活动语音比特流的结构如图 8 所示。接收模块 1700 接收比特流。活动 / 非活动语音比较器 1706 首先测试比特流大小，如果比特率大于或等于 8Kbps（160 比特的大小），则比较器 1706 确定其为活动语音比特流，否则，确定其为非活动语音比特流。如果比特流为活动语音比特流，则活动语音窄带 / 宽带比较器 1708 进一步比较其大小，比较器 1708 确定是应由模块 1716 仅调用窄带解码器，还是应由模块 1718 调用完整的宽带解码器。如果比较器 1706 指示非活动语音比特流，则 NT/SID 比较器 1704 检查比特流的大小是等于 0（NT 帧）还是大于 0（SID 帧）。如果比特流是 SID 帧，则非活动语音窄带 / 宽带比较器 1702 进一步测试比特流的大小，以便确定 SID 信息是包括完整的宽带信息还是仅包括窄带信息，以及是应由模块 1712 调用完整的非活动语音宽带解码器还是应由模块 1710 仅调用非活动窄带解码器。如果比特流的大小为 0，即没有接收到信息，则由模块 1714 调用非活动语音外推解码器。应指出的是，比较器的顺序对于算法的操作并不重要并且仅作为示意性实施例提供比较操作的所述顺序。

[0058] 还可能的是网络元件将截断活动语音分组的宽带嵌入式层，同时保持非活动语音分组的宽带嵌入式层不变。这是因为除去活动语音分组的宽带嵌入式层中的大量比特可以显著有助于拥塞减小，而截断非活动语音分组的宽带嵌入式层将只是少量地有助于拥塞减小。因此，非活动语音解码器的操作还依赖于活动语音解码器的操作历史。具体地说，如果当前接收的分组中的宽带信息不同于先前接收的分组，则需要特别注意。图 18 提供了示出在非活动语音解码中使用先前和当前带宽信息的算法的步骤的流程图。决策模块 1800 测试先前的比特流信息是否为宽带。如果先前比特流为宽带，则决策模块 1804 测试当前非活动语音比特流。如果当前非活动语音比特流为宽带，则调用非活动语音宽带解码器。如果当前非活动语音比特流为窄带，则执行宽带扩展以便避免输出静默 / 背景噪声信号上的突然带宽变化。此外，如果对于预定数量的分组，接收的带宽保持窄带，则可以执行适度的宽带减小。如果决策模块 1800 确定先前比特流为窄带，则决策模块 1802 测试当前非活动语音比特流。如果非活动语音比特流为窄带，则调用非活动语音窄带非活动语音解码器。如果当前非活动语音比特流为宽带，则截断非活动语音比特流的宽带部分并且调用窄带非活动语音解码器，避免输出静默 / 背景噪声信号上的突然带宽变化。此外，如果对于预定数量的分组，接收的带宽保持宽带，则可以执行适度的宽带增加。应指出的是，非活动语音外推解码器（尽管未在图 18 中隐含地指定）被视为非活动语音解码器的一部分并且始终遵循先前接收的带宽。

[0059] 图 4、9、10 和 11 中示出的 VAD 模块区分活动语音和非活动语音，后者被定义为静默或周围背景噪声。许多当前通信应用除语音信号外还使用音乐信号，如在中断时的音乐中或在个性化回铃音中。音乐信号既不是活动语音也不是非活动语音，但是如果为音乐信号的片段调用非活动语音编码器，则会严重降低音乐信号的质量。因此，设计为处理音乐信号的通信系统中的 VAD 检测音乐信号并且提供音乐检测指示很重要。音乐信号的检测和处理在使用宽带语音的语音通信系统中甚至更加重要，因为针对音乐信号的活动语音编解码器的内在质量相对较高，并且因此由针对音乐信号使用非活动语音编解码器导致的质量下降可能具有较强的感知影响。图 19 示出了接收输入语音 1902 的通用化的话音活动检测器

1901。输入语音 1902 被馈送到活动 / 非活动语音检测器 1905 (其类似于图 4、9、10 和 11 中提供的 VAD 模块) 和馈送到音乐检测器 1906。活动 / 非活动语音检测器 1905 生成活动 / 非活动语音指示 1908 并且音乐检测器 1906 生成音乐指示 1909。可以以多种方式使用音乐指示。其主要目的是避免使用非活动语音编码器并且对于该任务,可以通过忽略不正确的非活动语音决策来将其与活动 / 非活动语音指示符相结合。其还可以控制专用或标准的噪声抑制算法 (未示出),后者在输入语音到达编码器之前预处理输入语音。音乐指示还可以控制活动语音编码器的操作,例如其音高轮廓线平滑算法或其他模块。

[0060] 网络截断非活动语音的宽带增强层可能需要解码器扩展带宽以维持活动语音片段与非活动语音片段之间的带宽连续性。类似地,还可以使编码器仅发送窄带信息以及使解码器在活动语音是宽带语音的情况下执行宽带扩展。图 20 示出了非活动语音编码器 2000,其接收输入非活动语音 2002 并将静默 / 背景噪声比特流 2006 传输到非活动语音解码器 2001,后者生成重构非活动语音 2024。注意,输入非活动语音 2002 和重构非活动语音 2024 都是以 16KHz 采样的宽带信号。LP- 抽取元件 2003 接收输入非活动语音 2002 并且生成非活动语音窄带信号 2004,后者被窄带静默 / 背景噪声编码器 2005 接收以生成窄带静默 / 背景噪声比特流 2006。窄带静默 / 背景噪声比特流 2006 被窄带静默 / 背景噪声解码器 2007 接收以生成窄带非活动语音 2009 和辅助信号 2014。辅助信号 2014 可以包括能量和频谱参数以及窄带非活动语音 2009 本身。宽带扩展模块 2016 使用辅助信号 2014 来生成基带中的高频带非活动语音 2018。所述生成可以使用频谱扩展,其应用于具有能量轮廓线匹配和平滑的宽带随机激励。上采样 -LP 2010 接收窄带非活动语音 2009 并且生成低频带输出非活动语音 2012。上采样 -HP 2020 接收基带中的高频带非活动语音 2018 并且生成高频带输出非活动语音 2022。将低频带输出非活动语音 2012 和 2022 相加以产生重构非活动语音 2024。

[0061] 以上提供的方法和系统可以存在于软件、硬件或器件上的固件中,后者可以在微处理器、数字信号处理器、专用集成电路、现场可编程门阵列 (FPGA) 或它们的任意组合中实现而不脱离本发明的精神。此外,本发明可以体现在其他特定形式中而不脱离其精神或本质特性。所述实施例在所有方面都将被视为只是示例性的而非限制性的。

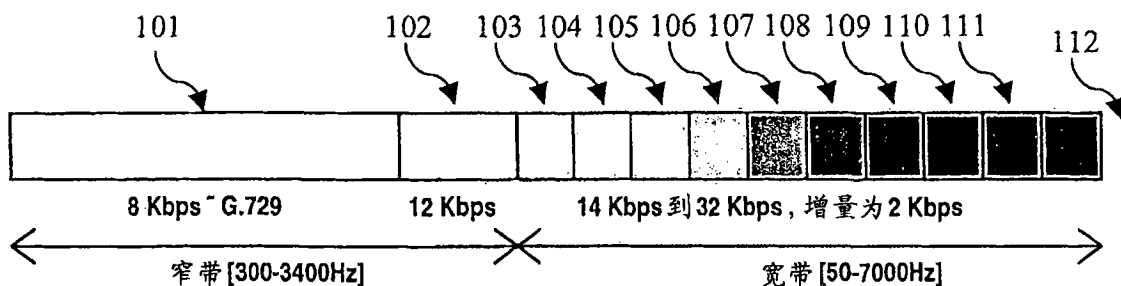


图 1G.729.1 比特流的嵌入式结构

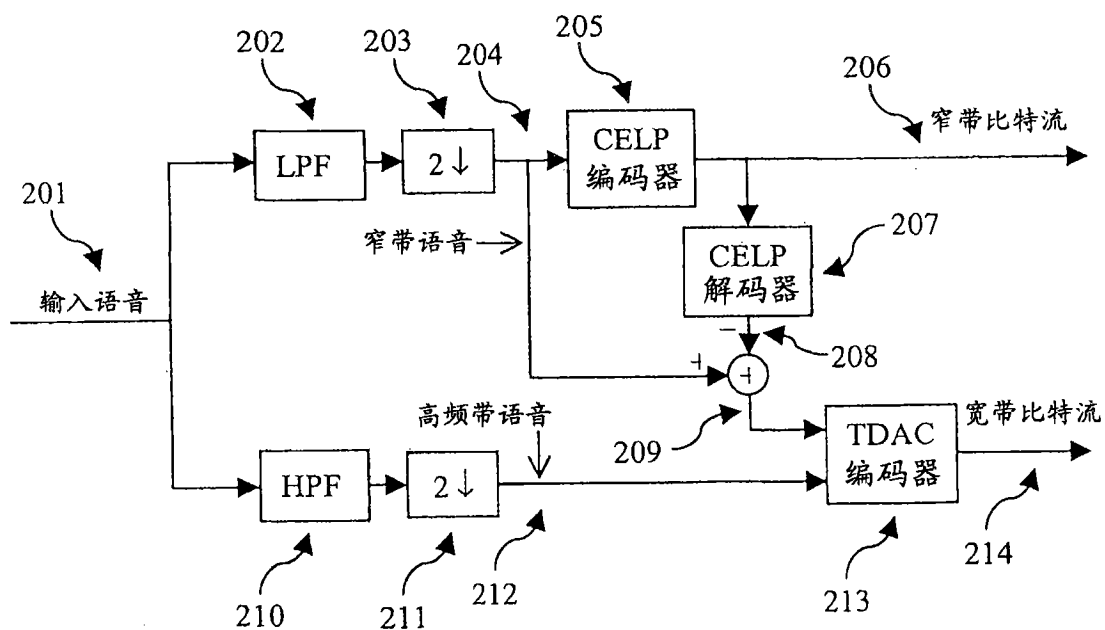


图 2G.729.1 编码器的结构

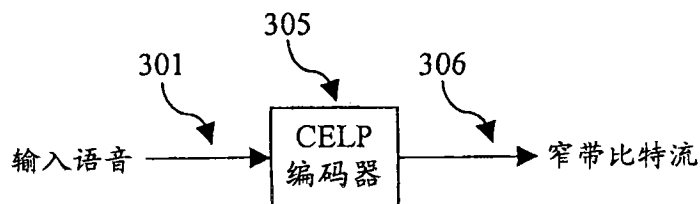


图 3G.729.1 编码器的备选操作 - 窄带编码

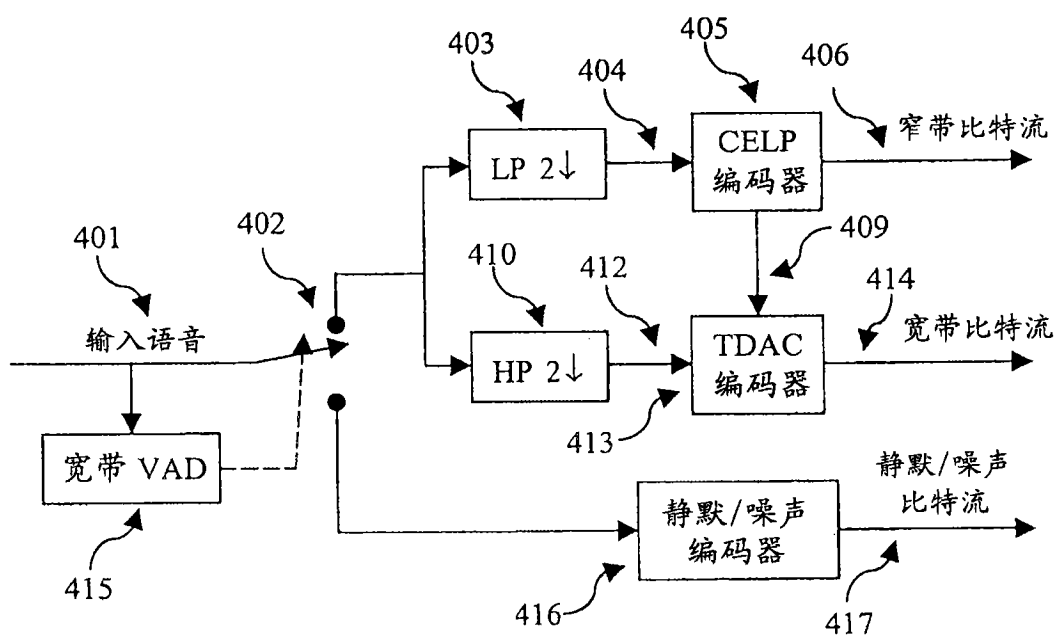


图 4G.729.1 的静默 / 背景噪声编码模式

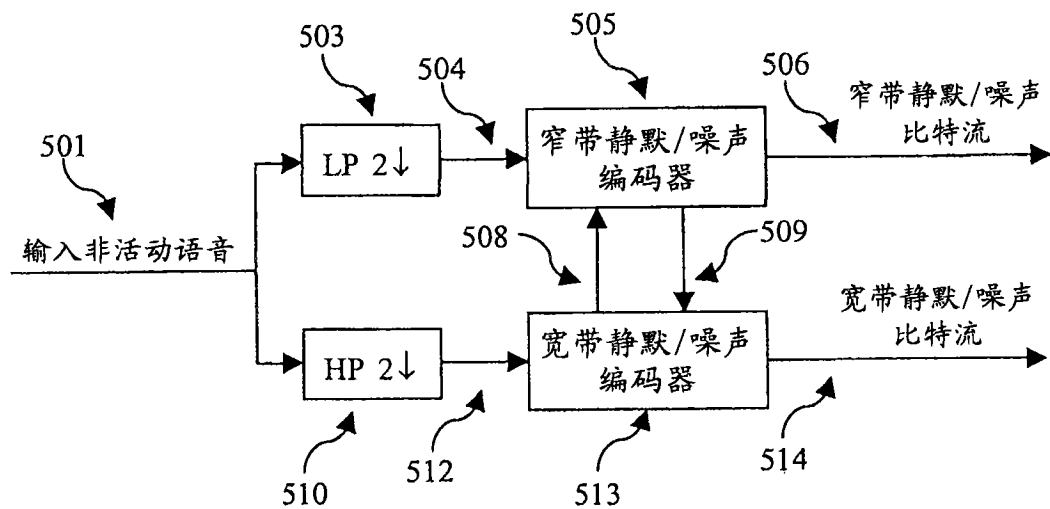


图 5 具有嵌入式结构的静默 / 背景噪声编码器

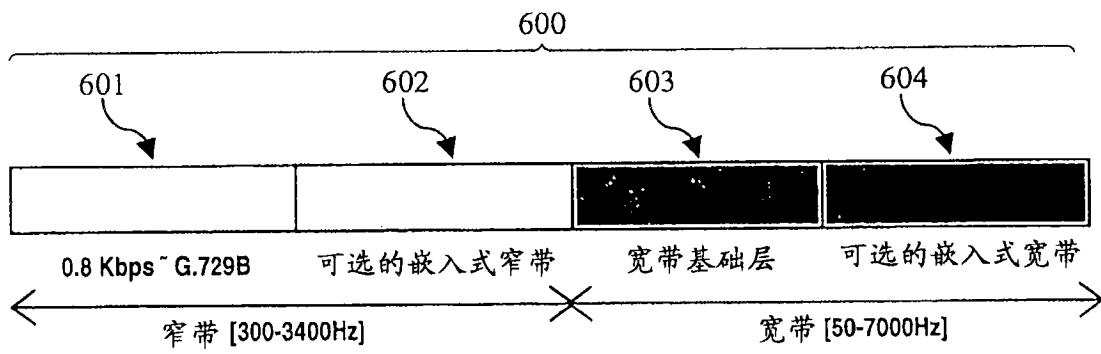


图 6 静默 / 背景噪声嵌入式比特流

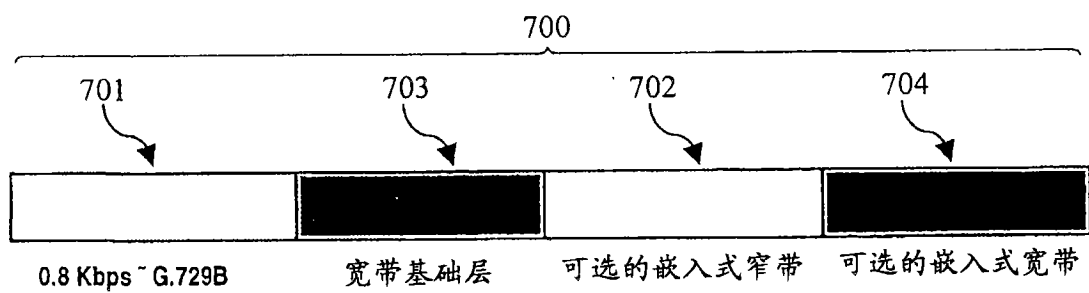


图 7 备选静默 / 背景噪声嵌入式比特流

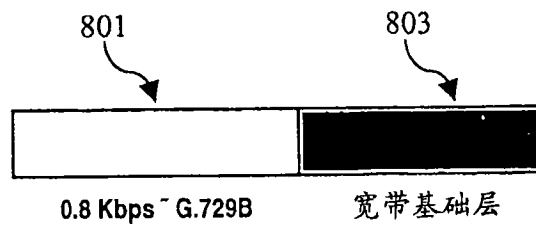


图 8 没有可选层的静默 / 背景噪声嵌入式比特流

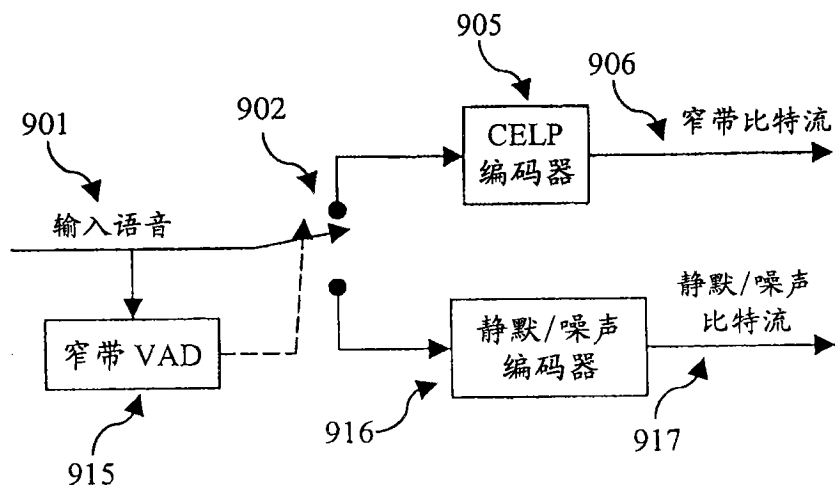


图 9G. 729. 1 的窄带工作模式的窄带 VAD

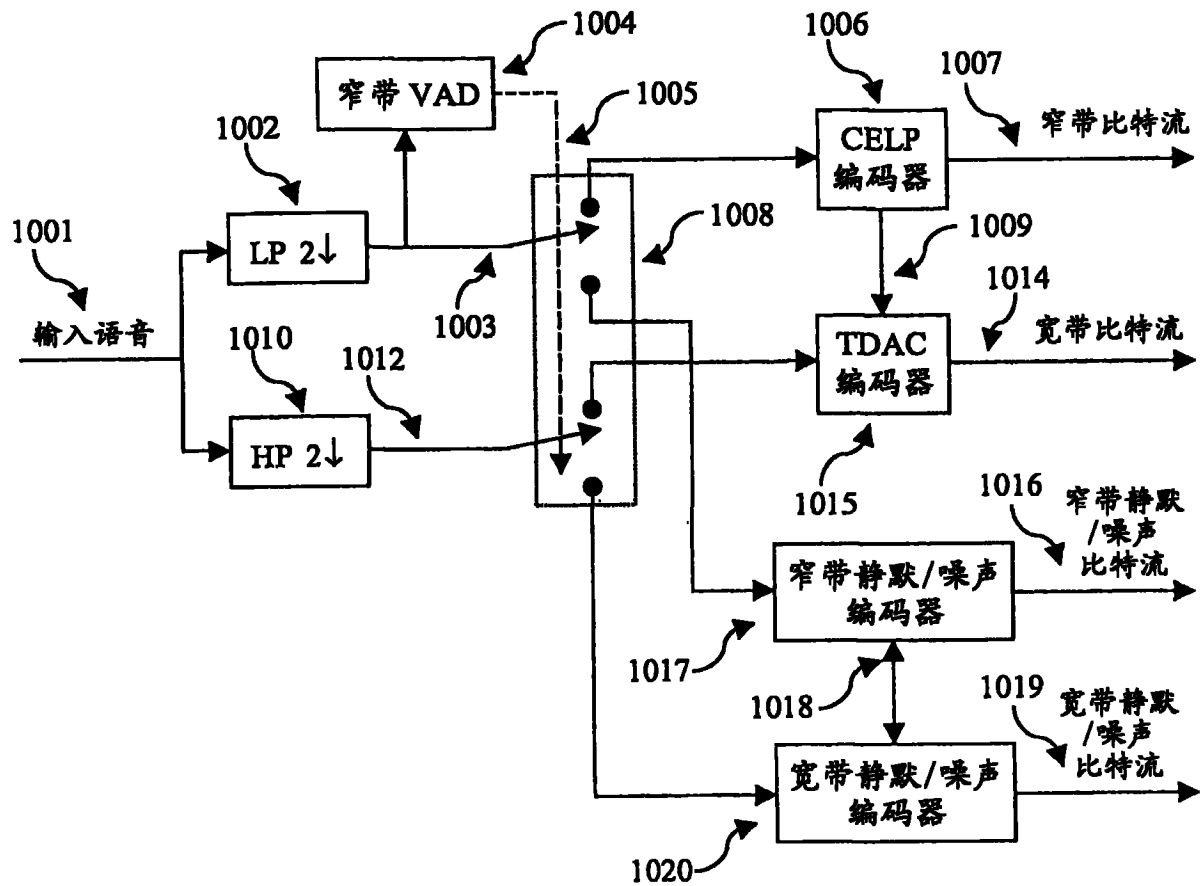


图 10 具有窄带 VAD 的 G. 729. 1 的静默 / 背景噪声编码模式

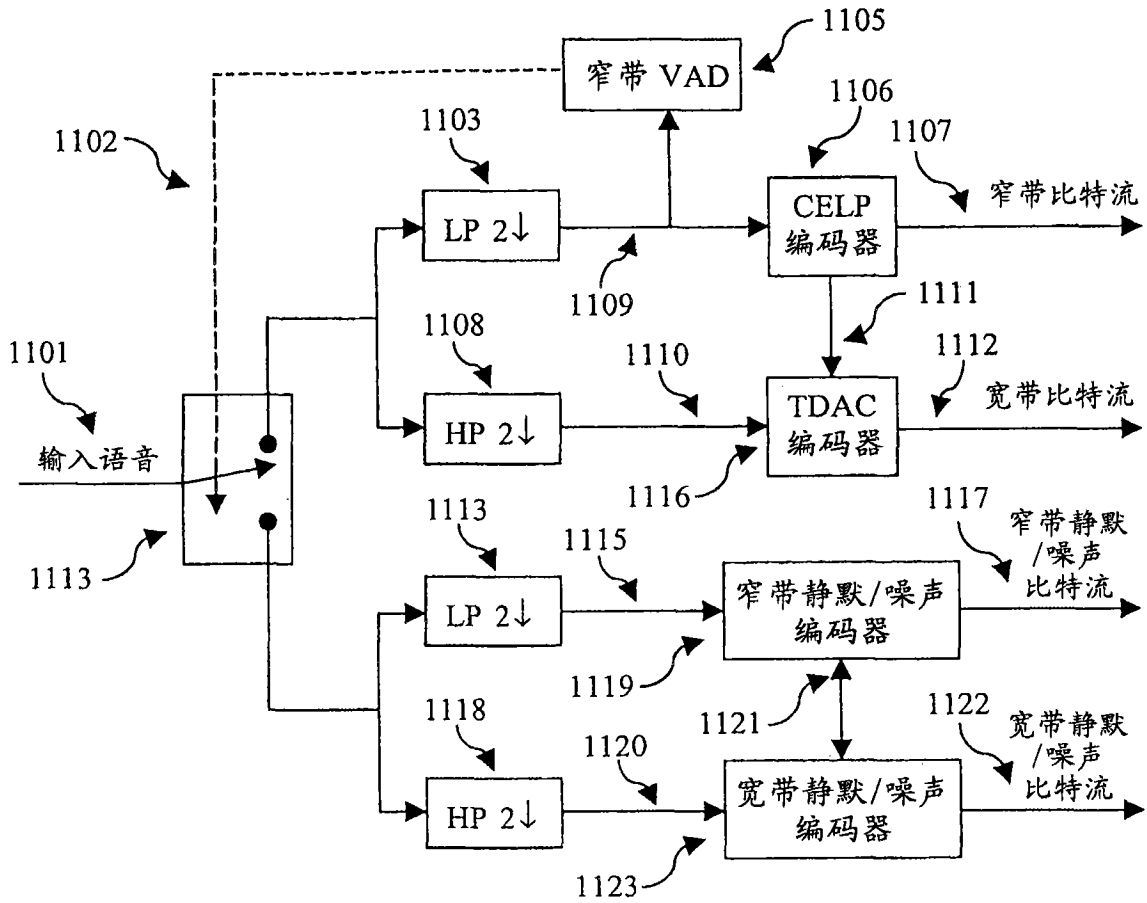


图 11 具有窄带 VAD 和单独抽取元件的 G.729.1 的静默 / 背景噪声编码模式

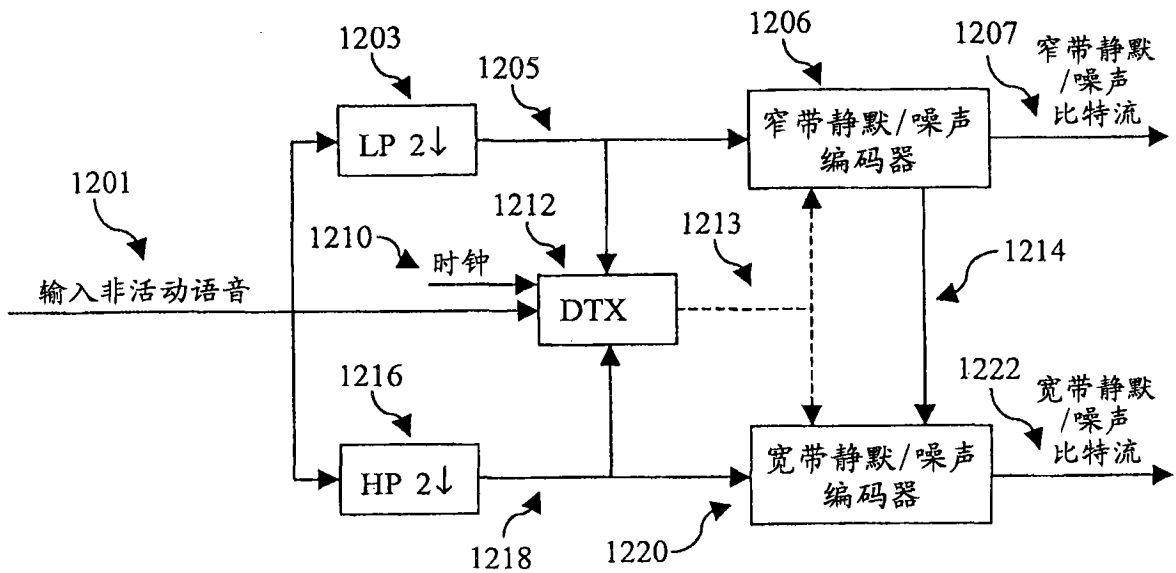


图 12 具有 DTX 模块的静默 / 背景噪声编码器

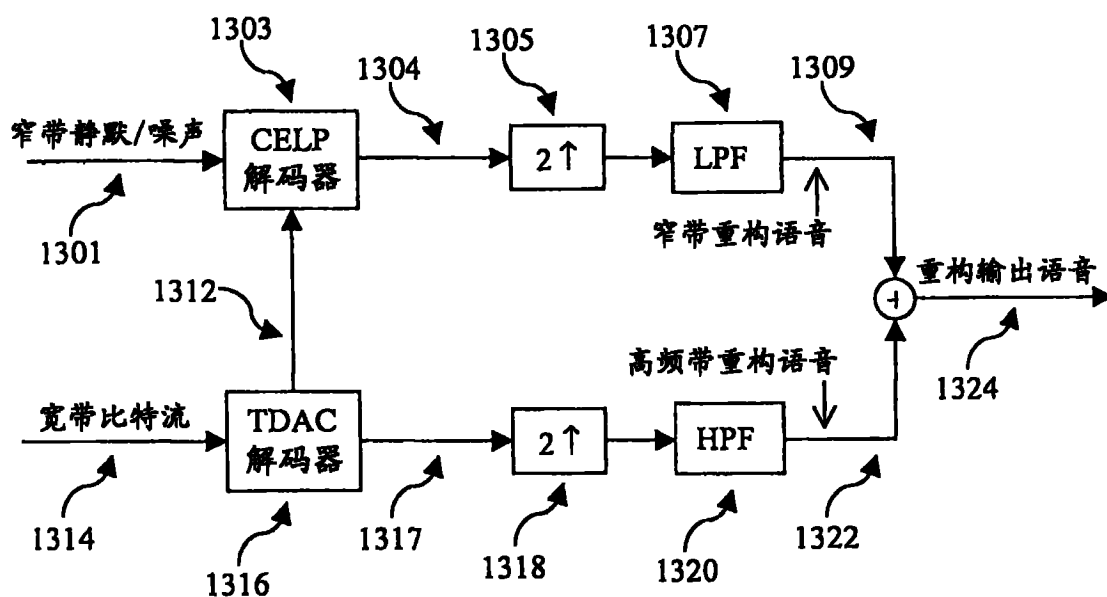


图 13G. 729. 1 解码器的结构

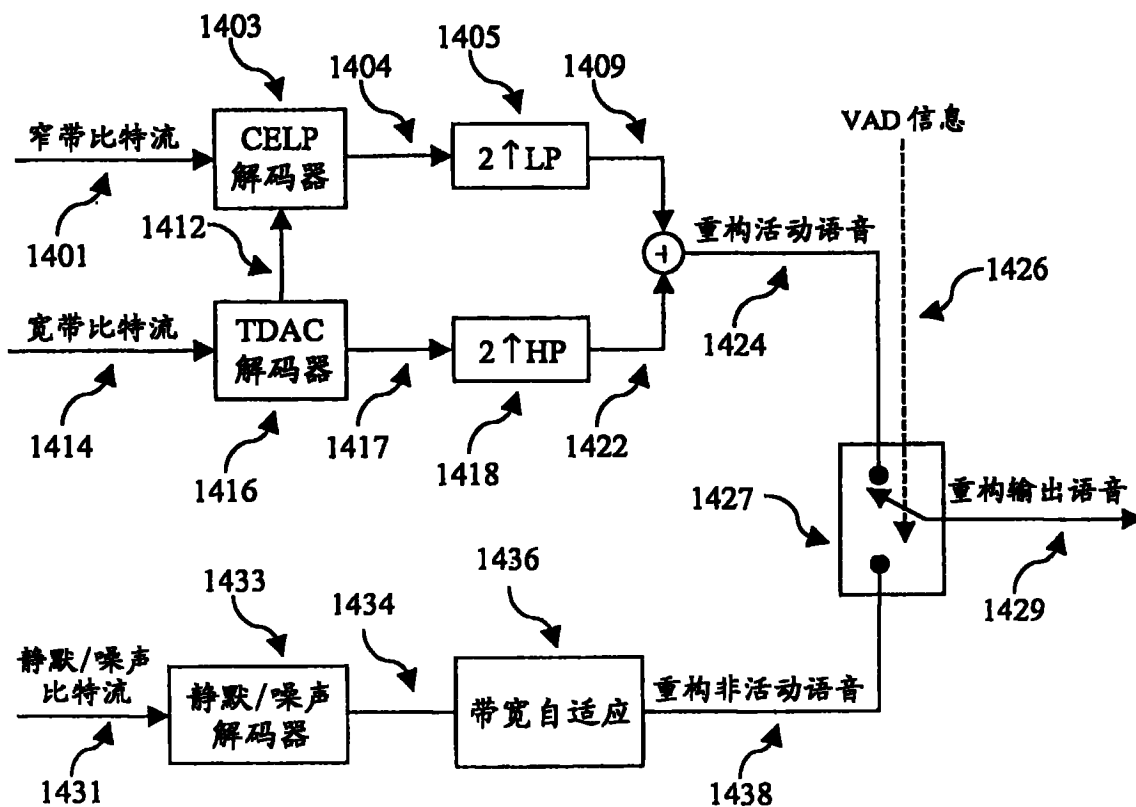


图 14 具有静默 / 背景噪声压缩的 G. 729. 1 解码器

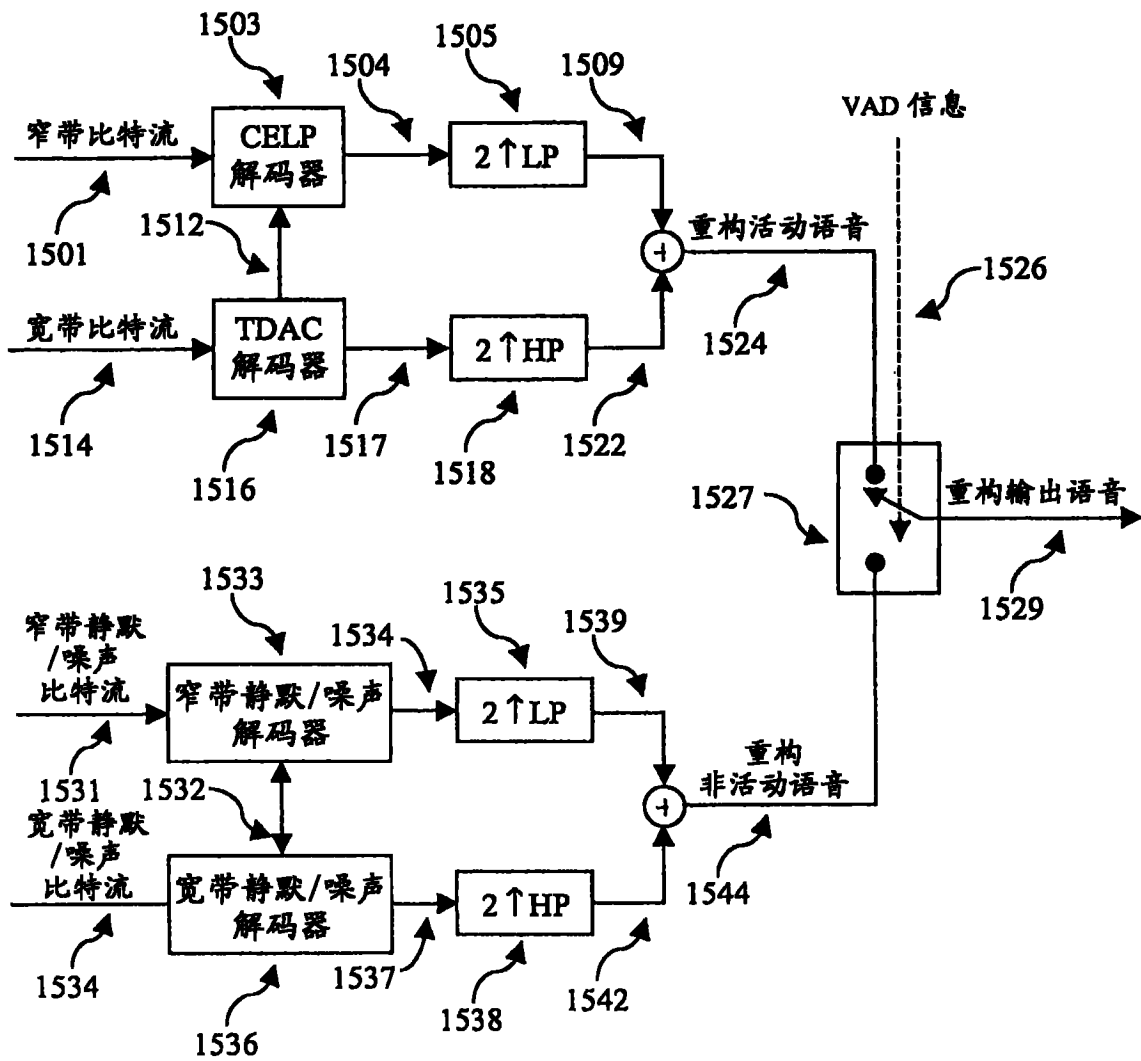


图 15 具有嵌入式静默 / 背景噪声压缩的 G.729.1 解码器

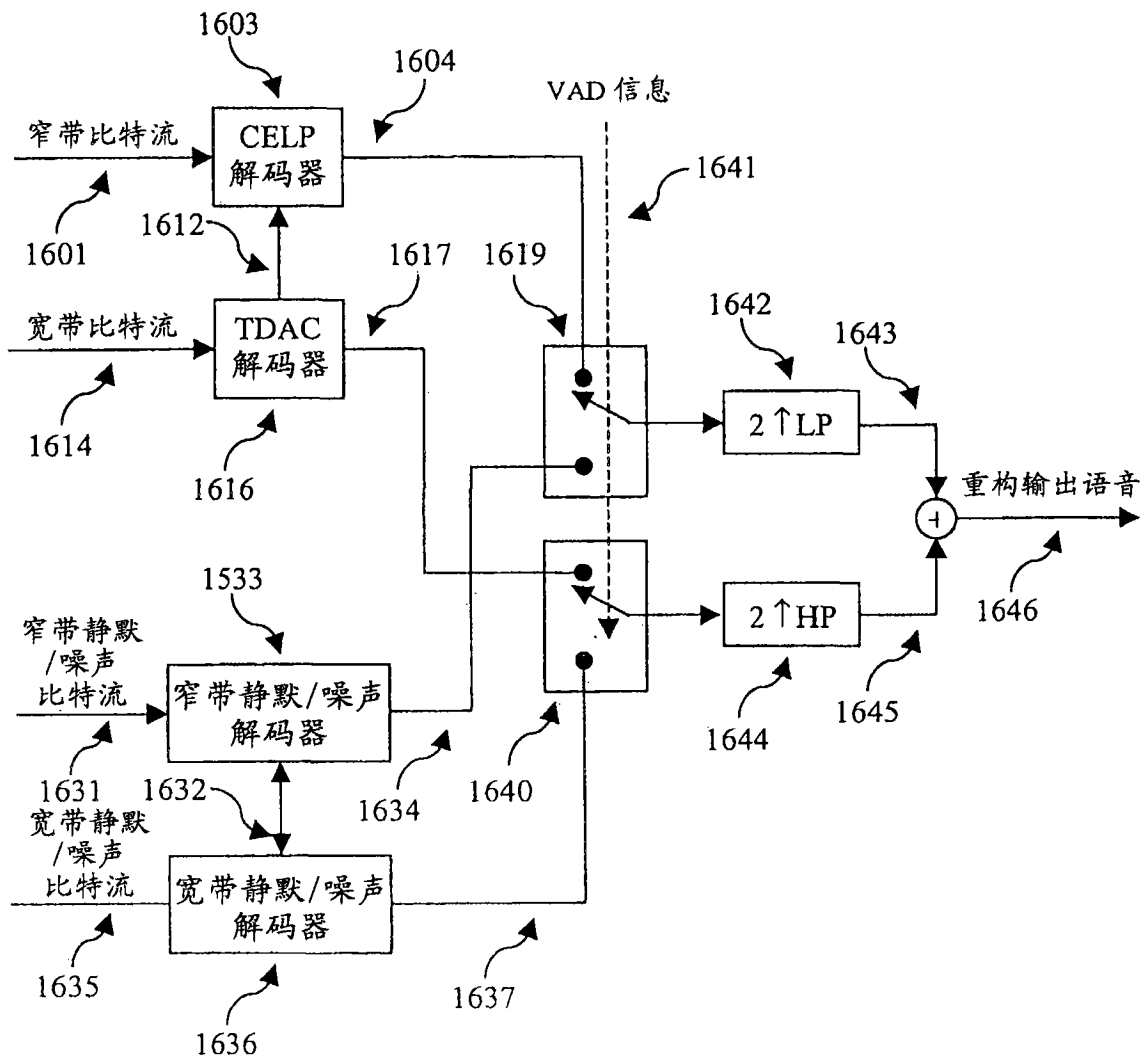


图 16 具有嵌入式静默 / 背景噪声压缩以及共享的上采样 - 滤波元件的 G.729.1 解码器

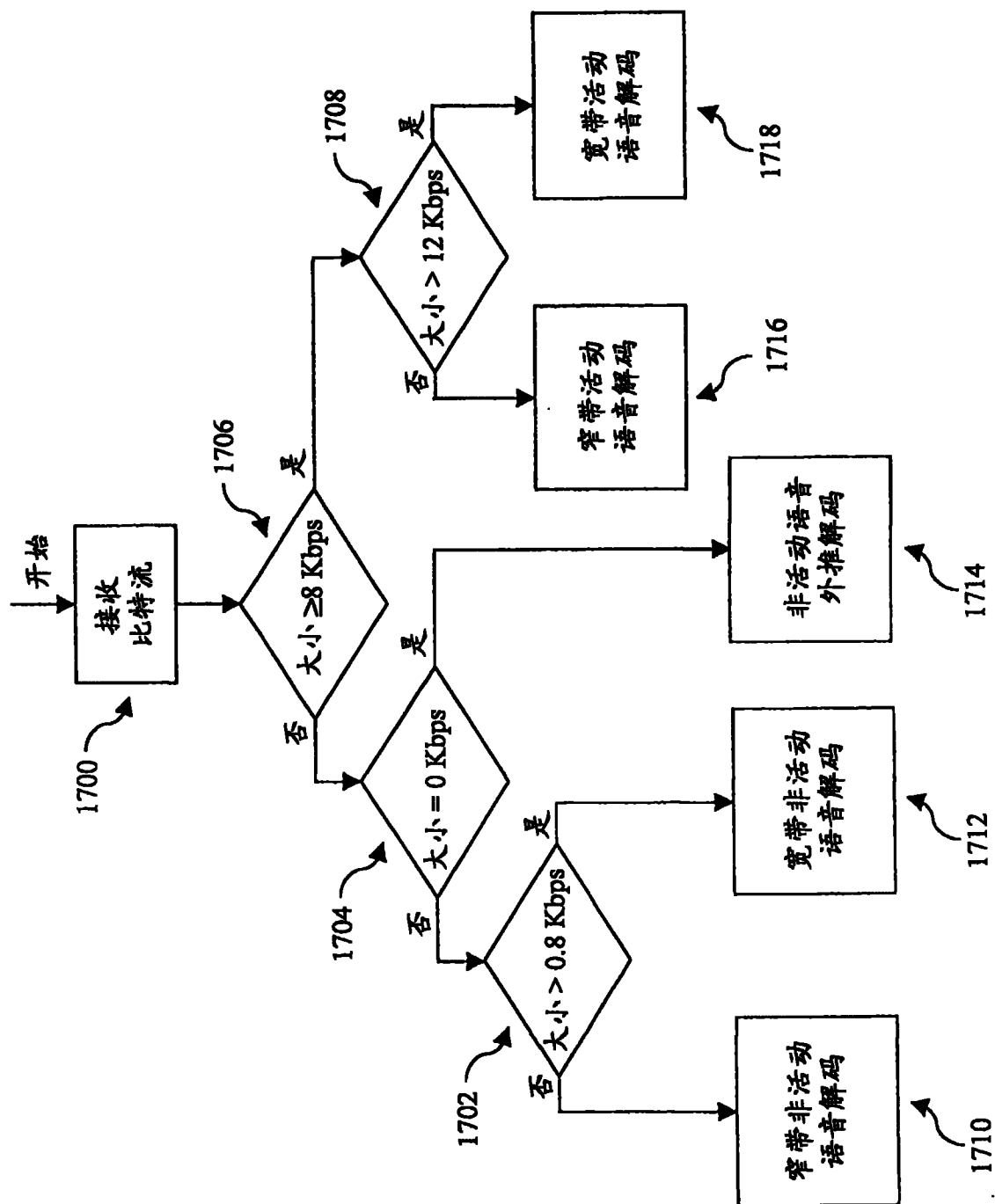


图 17 基于比特率的解码器控制操作

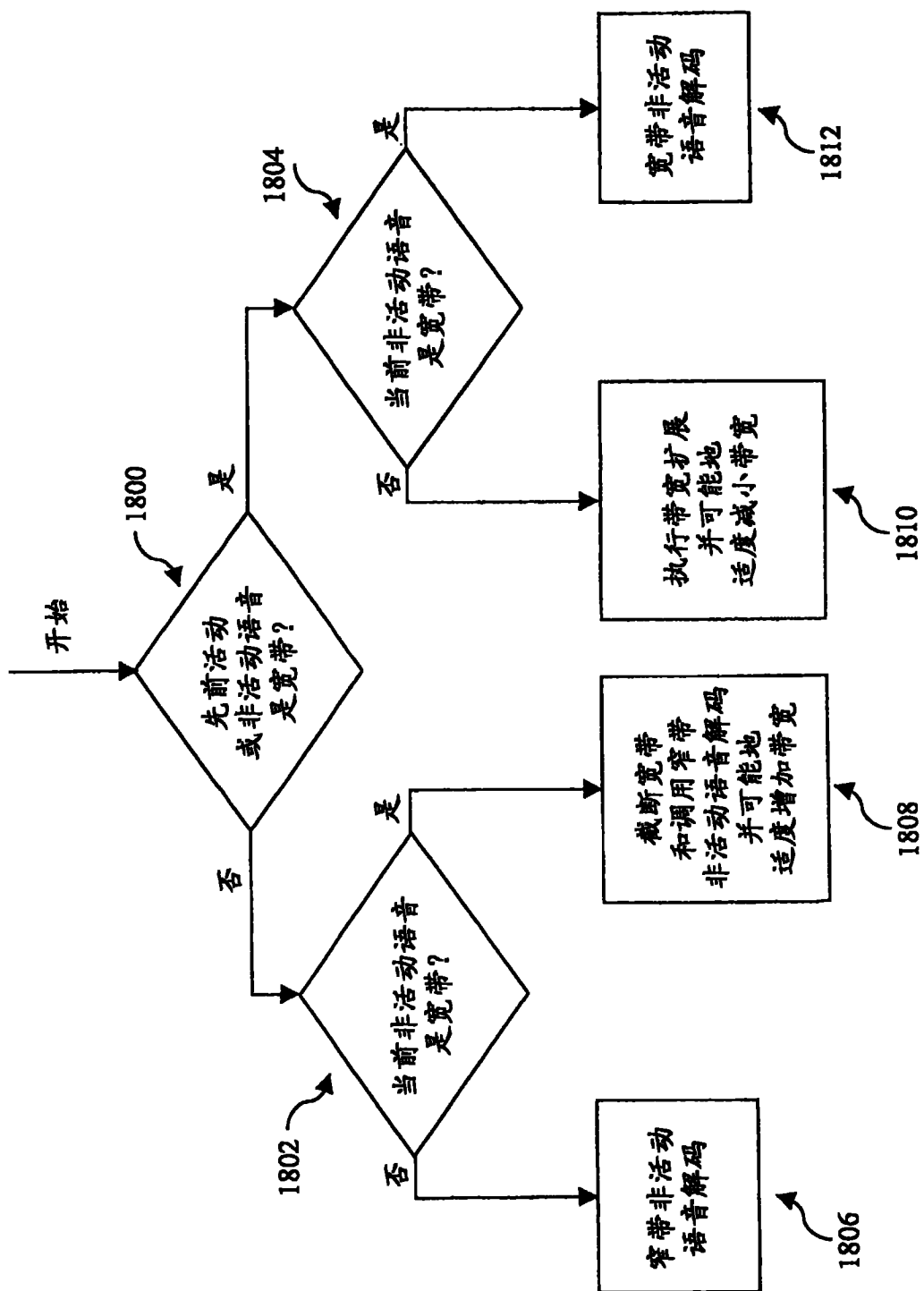


图 18 基于带宽历史的解码器控制操作

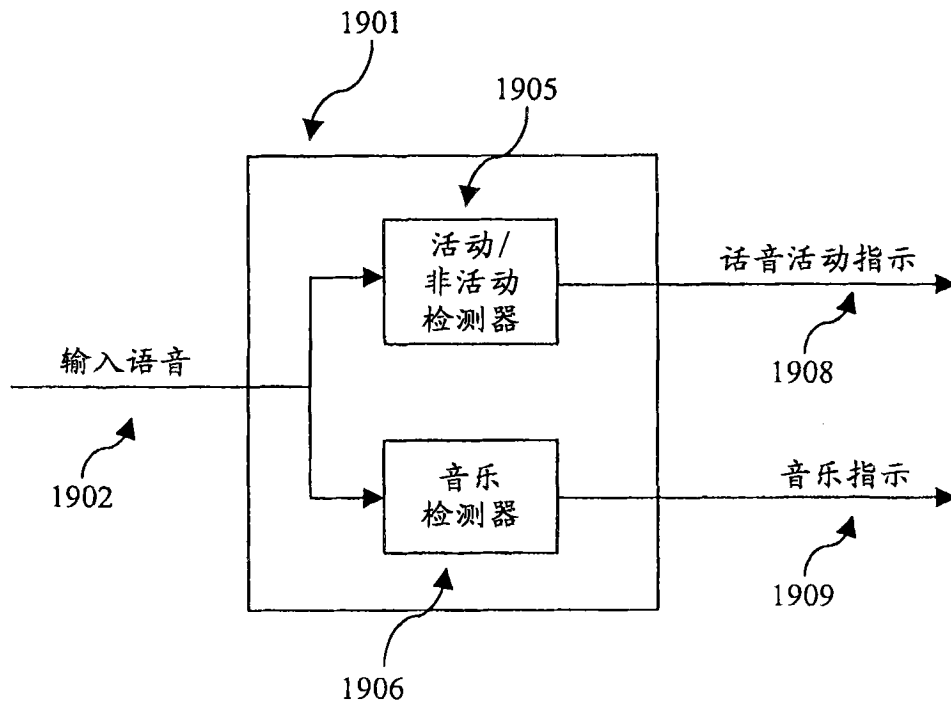


图 19 具有音乐检测的通用化的语音活动检测

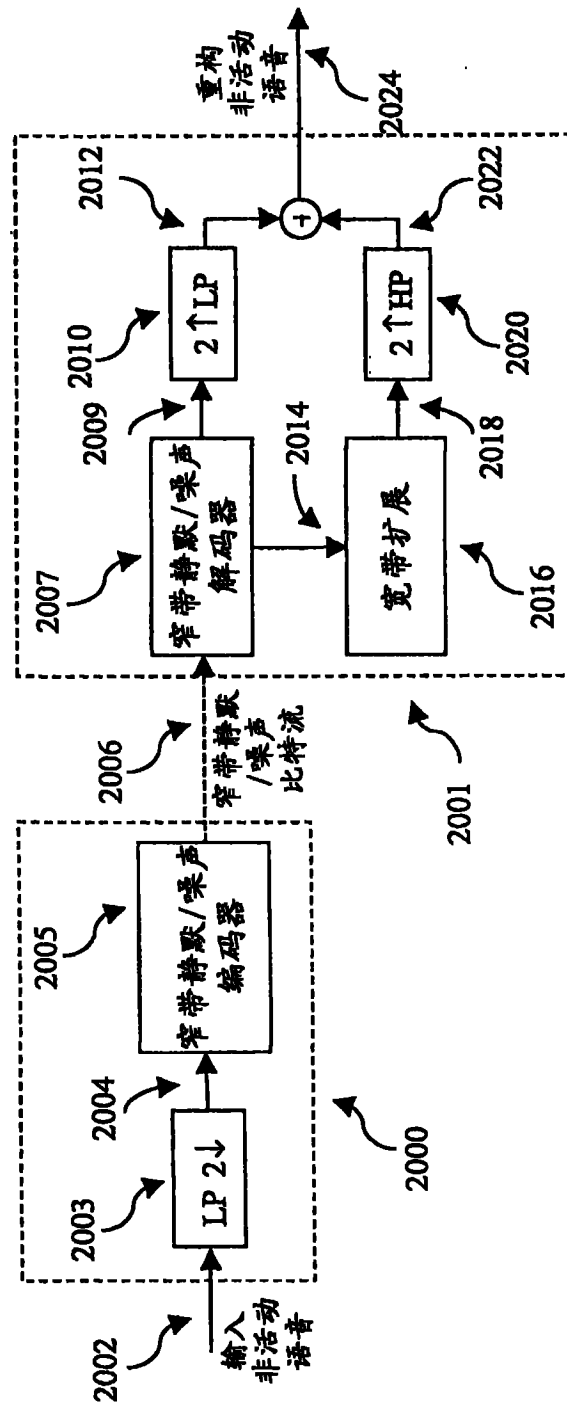


图 20 具有解码器带宽扩展的窄带静默 / 背景噪声传输