



[12] 发明专利说明书

[21] ZL 专利号 00810493. X

[45] 授权公告日 2004 年 5 月 5 日

[11] 授权公告号 CN 1148721C

[22] 申请日 2000.7.19 [21] 申请号 00810493. X

[30] 优先权

[32] 1999. 7. 19 [33] US [31] 09/356,860

[86] 国际申请 PCT/US2000/019671 2000.7.19

[87] 国际公布 WO01/006491 英 2001.1.25

[85] 进入国家阶段日期 2002.1.17

[71] 专利权人 高通股份有限公司

地址 美国加利福尼亚州

[72] 发明人 S·曼祖那什 A·P·德加科

审查员 杨 叁

[74] 专利代理机构 上海专利商标事务所

代理人 钱慰民

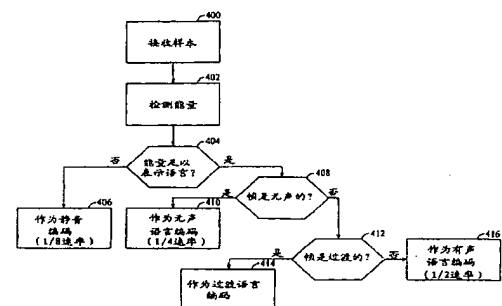
权利要求书 2 页 说明书 10 页 附图 6 页

[54] 发明名称 提供解码器到编码器的反馈以改进
帧删除情况下预测语言编码装置性
能的方法和装置

[57] 摘要

从解码器向编码器提供反馈以改进帧删除情况下预测语音编码装置性能的方法和装置, 它包括如果接收语音编码装置中的解码器(104、110、200、300)没有接收到发射语音编码装置中编码器(100、106)发送的帧, 就通知接收语音编码装置中的编码器(100、106)。响应通知, 将改变的数据包从接收语音编码装置中的编码器(100、106)发送到发射语音编码装置中的解码器(104、110、200、300)。当接收到改变的数据包时, 发射语音编码装置中的解码器(104、110)通知发射语音编码装置中的编码器(100、106)。然后, 发射语音编码装置中的编码器(100、106)处用改变的编码格式将数据包编码。改变的编码格式可以是低记忆编码格式或无记忆编码格式。改变的数据包可以具有被设置为

数字值 1 的删除指示符比特。



1. 一种语音编码系统，其特征在于，它包括：

第一语音编码装置，包括第一编码器和第一解码器；和

第二语音编码装置，包括第二编码器和第二解码器，

其中第一编码器被构造成将语音帧的数据包编码，并通过通信信道将数据包发送到第二解码器，第二解码器被构造成接收并解码数据包，如果第二解码器没有接收到发送帧就发送信号到第二编码器，第二编码器被构造成编码并发送数据包，并响应来自第二解码器的信号改变数据包，第一解码器被构造成接收并解码数据包，并且在接收到来自第二编码器的改变的数据包时发送信号到第一编码器，第一编码器还被构造成响应来自第一解码器的信号，使用改变的编码格式将数据包编码；并且

改变的数据包包括带有被设置为数值 1 的删除指示符比特的数据包。

2. 如权利要求 1 所述的语音编码系统，其特征在于，改变的编码格式包括低记忆编码方案。

3. 如权利要求 1 所述的语音编码系统，其特征在于，改变的编码格式包括无记忆编码方案。

4. 如权利要求 1 所述的语音编码系统，其特征在于，第一语音编码装置存在于用户单元中，第二语音编码装置存在于无线通信系统的基站中。

5. 如权利要求 1 所述的语音编码系统，其特征在于，第二语音编码装置存在于用户单元中，第一语音编码装置存在于无线通信系统的基站中。

6. 从第一语音编码装置中的第一解码器向第二语音编码装置中的第二编码器提供反馈的方法，其中第一语音编码装置包括第一编码器和第一解码器，第二语音编码装置包括第二编码器和第二解码器，其特征在于，该方法包括以下步骤：

如果第一解码器没有接收到第二编码器发送的帧，就通知第一语音编码装置中的第一编码器；

响应通知，将改变的数据包从第一编码器发送到第二解码器；

当第二解码器接收到来自第一编码器的改变的数据包时，通知第二编码器；

和

在第二编码器处用改变的编码格式将数据包编码；

其中，改变的数据包包括带有被设置为数字值 1 的删除指示符比特的数据包。

7. 如权利要求 6 所述的方法, 其特征在于, 改变的编码格式包括低记忆编码方案。

8. 如权利要求 6 所述的方法, 其特征在于, 改变的编码格式包括无记忆编码方案。

9. 如权利要求 6 所述的方法, 其特征在于, 第一语音编码装置存在于用户单元中, 第二语音编码装置存在于无线通信系统的基站中。

10. 如权利要求 6 所述的方法, 其特征在于, 第二语音编码装置存在于用户单元中, 第一语音编码装置存在于无线通信系统的基站中。

11. 语音编码系统中的反馈机构, 语音编码系统包括第一和第二语音编码装置, 第一语音编码装置包括第一编码器和第一解码器, 第二语音编码装置包括第二编码器和第二解码器, 反馈机构包括:

如果第二解码器没有接收到第一编码器发送的帧, 就通知第二编码器的装置;

响应通知将改变的数据包从第二编码器发送到第一解码器的装置;

当第一解码器接收到来自第二编码器的改变的数据包时, 通知第一编码器的装置; 和

在第一编码器处用改变的编码格式将数据包编码的装置;

其中, 改变的数据包包括带有被设置为数字值 1 的删除指示符比特的数据包。

12. 如权利要求 11 所述的反馈机构, 其特征在于, 改变的编码格式包括低记忆编码方案。

13. 如权利要求 11 所述的反馈机构, 其特征在于, 改变的编码格式包括无记忆编码方案。

14. 如权利要求 11 所述的反馈机构, 其特征在于, 第一语音编码装置存在于用户单元中, 第二语音编码装置存在于无线通信系统的基站中。

15. 如权利要求 11 所述的反馈机构, 其特征在于, 第二语音编码装置存在于用户单元中, 第一语音编码装置存在于无线通信系统的基站中。

提供解码器到编码器的反馈以改进帧删除情况下 预测语音编码装置性能的方法和装置

发明背景

I. 发明领域

本发明一般涉及语音处理领域，尤其涉及提供从解码器到协同配置的编码器的反馈以改进帧删除情况下预测语音编码装置性能的方法和装置。

II. 背景技术

用数字技术发送语音已非常普及，尤其在长距离和数字无线电话应用中。接着产生了对确定在信道上发送的最小信息量的兴趣，同时保持重建语音的感觉质量。如果通过简单地采样和数字化发送语音，那么需要大约每秒 64 千比特(kbps)的数据率，以获得传统模拟电话的语音质量。然而，通过使用语音分析，之后通过适当的编码、发送、并在接收机端再合成，可以大大降低数据率。

用于压缩语音的装置发现在电信的众多领域中有用。典型的领域是无线通信。无线通信领域具有很多应用，包括如无绳电话、寻呼、无线局部环路、如蜂窝和 PCS 电话系统的无线电话，移动因特网协议(IP)电话，和卫星通信系统。尤其重要的应用是移动用户的无线电话。

已经发展了各种空中接口用于无线通信系统，包括如频分多址(FDMA)、时分多址(TDMA)、和码分多址(CDMA)。在与之的连接中，建立了各种国内和国际的标准，包括如高级移动电话业务(AMPS)，全球移动通信系统(GSM)、和临时标准(IS-95)。典型的无线电话通信系统是码分多址(CDMA)系统。电信工业协会(TIA)和其它公知的标准团体颁布了 IS-95 标准及其衍生物 IS-95A、ANSI J-STD-008、IS-95B、提议的第三代标准 IS-95C 和 IS-2000 等等(这里统称为 IS-95)，以规定 CDMA 空中接口在蜂窝或 PCS 电话通信系统中的使用。美国专利号 5,103,459 和 4,901,307 中描述了根据使用 IS-95 标准实质构造的典型无线通信系统，这两个申请都转让给本发明的受让人，并通过引用全面结合于此。

通过使用提取关于人类语音产生模型的参数以压缩语音的技术的装置称为语音编码装置(coder)。语音编码装置将输入语音信号分成时间块或分析帧。语音编

码装置通常包括编码器和解码器。编码器分析输入语音帧，以提取某些相关的参数，然后将参数量化成二进制表示，也就是一组比特或二进制数据包。在通信信道上将数据包从接收机发送到解码器。解码器处理数据包，将它们去量化以产生参数，并使用去量化参数再合成语音帧。

语音编码装置的功能是通过除去语音中的所有固有冗余将数字化的语音信号压缩成低比特率的信号。通过用一组参数表示输入语音帧并使用量化用一组比特表示参数，以实现数字压缩。如果输入语音帧具有比特数 N_i 且语音编码装置产生的数据包具有比特数 N_o ，那么语音编码装置获得的压缩因数为 $C_r = N_i / N_o$ 。问题是在获得目标压缩因数的同时保持解码语音的高语音质量。语音编码装置的性能依赖于(1)语音模型、或上述分析和合成过程的组合执行得多好，(2)以目标比特率每帧 N_o 比特每帧执行的参数量化过程有多好。因此，语音模型的目标是用每帧一小组参数获取语音信号的本质，或目标语音质量。

也许语音编码装置设计中最重要的是搜索一组好的参数(包括向量)，以描述语音信号。一组好的参数要求低的系统带宽，用于感觉精确的语音信号的重建。音调、信号功率、频谱包络(或共振峰)，振幅和相位谱都是语音编码参数的实例。

语音编码装置可以是时域编码装置，它通过使用较高的时间分辨每次处理小片段语音(通常是 5 毫秒(ms)子帧)的编码，以试图获取时域语音波形。对于每个子帧，通过本领域所熟知的各种搜索算法发现了用码本空间的高精度表示。另外，语音编码装置可以是频域编码装置，它试图获取带有一组参数(分析)的输入语音帧的短期频谱，并使用相应的合成过程从频谱参数中重建语音波形。参数量化器根据 A. Gersho & R. M. Gray 所著的《Vector Quantization and Signal Compression》(1992)中描述的量化技术通过用存储的代码向量表示法表示参数，以保存参数。

公知的时域语音编码装置是 L. B. Rabiner & R. W. Schafer 所著的《Digital Processing of Speech Signal》中 396-453 页(1978)中描述的代码受激线性预测编码装置(CELP)，它通过引用全面结合于此。在 CELP 编码装置中，线性预测(LP)分析除去语音信号中的短期相关或冗余，它找出短期共振峰滤波器的系数。将短期预测滤波器施加到输入语音帧产生 LP 剩余信号，用长期预测滤波器参数和随后的随机码本将该信号进一步模型化和量化。因此，CELP 编码将时域语音波形的编码任务分成 LP 短期滤波器系数编码和 LP 剩余编码的分离任务。可以以固定的速率(也就是对每个帧使用相同的比特数目 N_o)或变化的速率(其中不同比特率用于

不同类型的帧内容)执行时域编码。变化速率编码装置试图只使用编码参数编码以达到目标质量水平所需的比特量。美国专利号 5,414,796 中描述了典型的变化速率 CELP 编码装置,该申请转让给了本发明的受让人,并通过引用全面结合于此。

时域编码装置如 CELP 编码装置通常依赖于每帧高的比特数 N_0 ,以保持时域语音波形的精确度。这种编码装置通常传送由相对大的每帧比特数 N_0 (如 8kbps 或更大)提供的较好语音质量。然而,在低比特率(4kbps 或更低)时,由于有限的可获得比特数,时域编码装置就不能保持高质量和稳健的性能。在低比特率时,有限的码本空间箝制了符合传统时域编码装置波形匹配的容量,编码装置在较高速率的工业应用中已成功发展。因此,尽管随着时间推移而进行了改进,但是以低比特率工作的很多 CELP 编码系统遭受感觉上通常表征为噪声的严重的失真。

当前涌现了研究兴趣和较强的商业需要以发展在中到低比特率(也就是在 2.4 到 4kbps 范围内或更小)上工作的高质量的语音编码装置。应用领域包括无线电话、卫星通信、因特网电话、各种多媒体和话音流应用、话音邮件和其它具有存储系统。驱动力是高容量的需要,是在包丢失情况下稳健性能的要求。各种近来的语音编码标准化努力是另一直接驱动力,用于推进低速率语音编码算法的研究和发展。低速率语音编码装置在每个允许的应用带宽中产生更多的信道或用户,而且与适当信道编码附加层耦合的低速率语音编码装置能够适应编码装置规范的总比特预算,并能在信道出错条件下传送稳健的性能。

传统的低到中比特率、预测语音编码装置在引起帧删除的较差信道条件下执行得较差。语音编码装置的预测性质指出在帧删除(也就是在传输中丢失因此没有被接收到的帧)之后接收到的帧将根据较差的先前信息合成,解码器将失去与关联编码器的同步。因此合成的帧退化,语音质量也受影响。因此期望改进较差信道条件下的语音编码装置性能。因此需要使用反馈机构在帧删除情况下改进性能的语音编码装置。

发明内容

本发明是针对使用反馈机构在帧删除情况下改进性能的语音编码装置。因此,在本发明的一个方面,语音编码系统最好包括包含第一编码器和第一解码器的第一语音编码装置;和包含第二编码器和第二解码器的第二语音编码装置,其中第一编码器被构造成将语音帧的数据包编码,并通过通信信道将数据包发送到第二解码器,第二解码器被构造成接收并将数据包解码,如果第二解码器没有接

收到发送帧就发送信号到第二编码器，第二编码器被构造成编码并发送数据包，并响应来自第二解码器的信号改变数据包，第一解码器被构造成接收并解码数据包，并且根据接收到来自第二编码器的改变的数据包发送信号到第一编码器，第一编码器还被构造成响应来自第一解码器的信号，使用改变的编码格式将数据包编码。

在本发明的另一方面，从第一语音编码装置中的第一解码器向第二语音编码装置中的第一编码器提供反馈的方法最好包括以下步骤，如果第一解码器没有接收到第一编码器发送的帧，就通知第一语音编码装置中的第二编码器；响应通知，将改变的数据包从第二编码器发送到第二解码器；当第二解码器接收到来自第二编码器的改变的数据包时，通知第一编码器；并且在第一编码器处用改变的编码格式将数据包编码。

在本发明的另一方面，反馈机构在语音编码系统中，语音编码系统包括第一和第二语音编码装置，第一语音编码装置包括第一编码器和第一解码器，第二语音编码装置包括第二编码器和第二解码器，最好包括如果第二解码器没有接收到第一编码器发送的帧，就通知第二编码器的装置；响应通知将改变的数据包从第二编码器发送到第一解码器的装置；当第一解码器接收到来自第二编码器的改变的数据包时，通知第一编码器的装置；和在第一编码器处用改变的编码格式将数据包编码的装置。

附图说明

图 1 是无线电话系统的框图。

图 2 是语音编码装置在每个终端处终止的通信信道的框图。

图 3 是编码器的框图。

图 4 是解码器的框图。

图 5 是表示语音编码判定过程的流程图。

图 6A 是语音信号振幅相对时间的图像；图 6B 是线性预测 (LP) 剩余振幅相对时间的图像。

图 7 是使用反馈回路的语音编码系统的框图，反馈回路从接收机处解码器到接收机处编码器、从接收机处编码器到发射机处解码器、从发射机处解码器到发射机处编码器。

较佳实施例的详细描述

这里以下所述的典型实施例存在于被构造成使用 CDMA 空中接口的无线电话通信系统中。然而，本领域熟练的技术人员应该理解使用本发明特征的次采样方法和装置可以存在于使用本领域熟练技术人员所熟知的较宽范围技术的任何各种通信系统中。

如图 1 所示，CDMA 无线电话系统一般包括多个移动用户单元 10、多个基站 12、基站控制器(BSC)14，和移动交换中心(MSC)16。MSC 16 被构造成与传统公共交换电话网(PSTN)18 连接。MSC 16 还被构造成与 BSC 14 连接。BSC 14 提供回传线路耦合到基站 12。回传线路可以被构造成支持几个已知接口中的任何一个，如 E1/T1、ATM、IP、PPP、Frame Relay、HDSL、ADSL 或 xDSL。应该理解系统中有多于两个的 BSC 14。每个基站 12 最好包括至少一个扇区(未图示)，每个扇区包括一全向天线或从基站 12 发出的特定径向方向的天线。另一种情况是，每个扇区可以包括两个天线用于分集接收。每个基站 12 最好被设计成支持多个频率分配。扇区的相交和频率分配可以称为 CDMA 信道。基站 12 也熟知为基站收发机子系统(BTS)12。另一种情况是，“基站”可用于工业中以统称为 BSC 14 和一个或多个 BTS 12。BTS 12 也可以表示为“区站”12。另一种情况是，给定 BTS 12 的各个扇区也可称为区站。移动用户单元 10 通常是蜂窝或 PCS 电话 10。系统最好被构造成根据 IS-95 标准使用。

在蜂窝电话系统的典型工作中，基站 12 从移动单元组 10 接收反向链路信号组。移动单元 10 传导电话呼叫或其它通信。给定基站 12 接收到的每个反向链路信号在该基站 12 中被处理。所得的数据被发送到 BSC 14。BSC 14 提供呼叫资源分配和移动管理功能，包括基站 12 之间软越区切换的控制。BSC 14 还将接收到的数据发送到 MSC 16，MSC 16 为与 PSTN 18 的接口提供附加的路由服务。类似地，PSTN 18 和 MSC 16 连接，MSC 16 和 BSC 14 连接，BSC 14 依次控制基站 12 发送前向链路信号组到移动单元组 10。

在图 2 中，第一编码器 100 接收数字化的语音样本 $s(n)$ 并将样本 $s(n)$ 编码，用于在传输介质 102、或通信信道 102 上发送到第一解码器 104。解码器 104 将编码语音样本解码，并合成输出语音信号 $S_{\text{SYNTH}}(n)$ 。为了在反方向上发送，第二编码器 106 将对在通信信道 108 上传输的数字化语音样本 $s(n)$ 编码。第二解码器 110 接收并将编码语音样本解码，产生合成的输出语音信号 $S_{\text{SYNTH}}(n)$ 。

语音样本 $s(n)$ 表示根据本领域中各种熟知的任何方法数字化并量化的语音信号，这些方法如脉冲编码调制(PCM)、压扩 μ 律、或 A 律。如本领域中所熟知的，语音样本 $s(n)$ 组织成输入数据帧，其中每个帧包括预定个数的数字化语音样本 $s(n)$ 。在典型实施例中，使用 8kHz 的采样速率，每个 20ms 的帧包括 160 个样本。在以下描述的实施例中，数据传输速率最好在帧到帧的基础上从 13.2kbps(全速率)变化到 6.2kbps(半速率)到 2.6kbps(四分之一速率)到 1kbps(八分之一速率)。数据传输速率的变化是有利的，因为对于包含相对较少语音信息的帧可以选用较低的比特率。本领域熟练的技术人员能够理解还可以使用其它采样速率和数据传输速率。

第一编码器 100 和第二解码器 110 一同组成第一语音编码装置。语音编码装置可用于发送语音信号的任何通信装置中，包括如以上参考图 1 描述的用户单元、BTS 或 BSC。类似地，第二编码器 106 和第一解码器 104 一同组成第二语音编码装置。本领域熟练的技术人员应该理解语音编码装置可以用数字信号处理器(DSP)、专用集成电路(ASIC)、离散门逻辑、固件、或任何传统可编程软件模块和微处理器实现。软件模块可存在于 RAM 存储器、快闪存储器、寄存器或本领域中所熟知的任何其它形式的可写存储媒介。另一种情况是，任何传统的处理器、控制器或状态机都可以替代微处理器。美国专利号 5,727,123 中中和 1994 年 2 月 16 日提交的美国申请序号 08/197,417 题为“VOCODER ASIC”的申请中都描述了专门为语音编码设计的典型 ASIC，这两个申请都转让给了本发明的受让人，并通过引用全面结合于此。

在图 3 中，可用于语音编码装置的编码器 200 包括模式判定模块 202、音调估计模块 204、LP 分析模块 206、LP 分析滤波器 208、LP 量化模块 210、和剩余量化模块 212。输入语音帧 $s(n)$ 提供给模式判定模块 202、音调估计模块 204、LP 分析模块 206、LP 分析滤波器 208。模式判定模块 202 根据每个输入语音帧 $s(n)$ 的周期、能量信噪比(SNR)或零交叉速率以及其它特征产生模式索引 I_m 和模式 M 。美国专利号 5,911,128 中描述了根据周期将语音帧分类的各种方法，该申请转让给了本发明的受让人，并通过引用全面结合于此。这种方法还结合入电信工业协会工业临时标准 TIA/EIA IS-127 和 TIA/EIA IS-733。1998 年 12 月 21 日提交的美国申请序号 09/217,341 题为“VARAIBLE RATE SPEECH CODING”的申请中还描述了典型的模式判定方案，该申请转让给了本发明的受让人，并通过参考全面结合于此。

音调估计模块 204 根据每个输入语音帧 $s(n)$ 产生音调索引 I_p 和滞后值 P_0 。LP 分析模块 206 执行每个输入语音帧 $s(n)$ 的线性预测分析, 以产生 LP 参数 a 。LP 参数 a 提供给 LP 量化模块 210。LP 量化模块 210 还接收模式 M , 以此执行模式依赖方式的量化过程。LP 量化模块 210 产生 LP 索引 I_{LP} 和量化 LP 参数 $\hat{a}$ 。除了输入语音帧 $s(n)$ 外 LP 分析滤波器 208 还接收量化 LP 参数 $\hat{a}$ 。LP 分析滤波器 208 产生 LP 剩余信号 $R[n]$, 它表示输入语音帧 $s(n)$ 和根据量化线性预测参数 $\hat{a}$ 重建的语音之间的误差。LP 剩余 $R[n]$ 、模式 M 和量化 LP 参数 $\hat{a}$ 被提供给剩余量化模块 212。根据这些值, 剩余量化模块 212 产生剩余索引 I_R 和量化剩余信号 $\hat{R}[n]$ 。

在图 4 中, 可用于语音编码装置的解码器 300 包括 LP 参数解码模块 302、剩余解码模块 304、模式解码模块 306 和 LP 合成滤波器 308。模式解码模块 306 接收并将模式索引 I_M 解码, 从中产生模式 M 。LP 参数解码模块 302 接收模式 M 和 LP 索引 I_{LP} 。LP 参数解码模块 302 将接收到的值解码, 产生量化 LP 参数 $\hat{a}$ 。剩余解码模块 304 接收剩余索引 I_R 、音调索引 I_p 、和模式索引 I_M 。剩余解码模块 304 将接收到的值解码, 产生量化剩余信号 $\hat{R}[n]$ 。量化剩余信号 $\hat{R}[n]$ 和量化 LP 参数 $\hat{a}$ 被提供给 LP 合成滤波器 308, 它从中合成解码输出语音信号 $\hat{s}[n]$ 。

本领域中熟知图 3 中编码器 200 和图 4 中解码器 300 中各种模块的操作和实施, 并且在上述美国专利号 5,414,196 和 L.B.Rabiner & R.W.Schafer 所著的《Digital Processing of Speech Signals》中 396-453 页(1978)中对其进行了描述。

如图 5 中流程图所示, 根据一个实施例的语音编码装置遵循处理语音样本用于发送的一组步骤。在步骤 400 中, 语音编码装置接收连续帧中的语音信号的数字样本。一旦接收给定的帧, 语音编码装置就进入步骤 402。在步骤 402 中, 语音编码装置检测帧的能量。能量是帧中语音活动性的衡量。通过将数字化语音样本振幅的平方求和并将所得的能量和阈值比较, 执行语音检测。在一个实施例中, 阈值根据背景噪声电平的变化而自适应改变。上述美国专利号 5,414,796 中描述了典型的可变阈值语音活动性检测器。一些无声的语音声很可能是低能量样本, 它可能被当做背景噪声而错误编码。为了防止这种情况的发生, 使用低能量样本的频谱倾斜, 以区别无声语音和背景噪声, 如上述美国专利号 5,414,796 中所描述的。

在检测帧的能量之后, 语音编码装置进入步骤 404。在步骤 404 中, 语音

编码装置确定检测到的帧能量是否足以将帧分类为包含语音信息。如果检测到的帧能量低于预定的阈值电平,那么语音编码装置执行步骤 406。在步骤 406 中,语音编码装置将帧作为背景噪声(也就是无语音或静音)编码。在一个实施例中,以 1/8 速率或 1kbps 将背景噪声帧编码。如果在步骤 404 中检测到的帧能量符合或超过预定的阈值电平,那么帧被分类为语音,语音编码装置执行步骤 408。

在步骤 408 中,语音编码装置确定帧是否为无声语音,也就是语音编码装置测试帧的周期。各种已知的周期确定方法包括如使用零交叉和使用归一化自相关函数(NACF)。尤其,上述美国专利号 5,911,128 中和美国申请序号 09/217,341 中描述了使用零交叉和 NACF 检测周期。此外,以上用于区别有声语音和无声语音的方法被结合入电信工业协会临时标准 TIA/EIA IS-127 和 TIA/EIA IS-733。如果在步骤 408 中确定帧是无声语音,语音编码装置执行步骤 410。在步骤 410 中,语音编码装置将帧作为无声语音编码。在一个实施例中,以 1/4 速率或 2.6kbps 将无声语音帧编码。如果在步骤 408 中,确定帧不是无声语音,语音编码装置执行步骤 412。

在步骤 412 中,语音编码装置使用本领域中熟知的周期检测方法确定帧是否是过渡语音,如上述美国专利号 5,911,128 中所描述的方法。如果确定帧是过渡语音,语音编码装置执行步骤 414。在步骤 414 中,将帧作为过渡语音(也就是从无声语音过渡到有声语音)编码。在一个实施例中,根据多脉冲内插编码法将过渡语音帧编码,1999 年 5 月 7 日提交的美国申请序号 09/307,294 题为“MULTIPULSE INTERPOLATIVE CODING OF TRANSITION SPEECH FRAMES”的申请中描述了该方法,该申请转让给了本发明的受让人,并通过引用全面结合于此。在另一实施例中,以全速率或 13.2kbps 将过渡语音帧编码。

如果在步骤 412 中语音编码装置确定帧不是过渡语音,那么语音编码装置执行步骤 416。在步骤 416 中,语音编码装置将帧作为有声语音编码。在一个实施例中,可以以半速率或 6.2kbps 将有声语音帧编码。也可以以全速率或 13.2kbps(或者在 8k CELP 编码装置中以全速率 8kbps)将有声语音帧编码。然而,本领域熟练的技术人员将理解以半速率将有声帧编码,通过开发有声帧的稳态本质,允许编码装置节省有用的带宽。此外,不考虑有声语音编码所用的速率,最好使用来自过去帧的信息将有声语音编码,因此称为预测编

码。

本领域熟练的技术人员应该理解遵循图 5 的步骤可以将语音信号或相应的 LP 剩余编码。从图 6A 的曲线图中可以看出作为时间函数的噪声、无声、过渡和有声语音的波形特征。从图 6B 的曲线图中可以看出作为时间函数的噪声、无声、过渡和有声 LP 剩余的波形特征。

如图 7 所示，在一个实施例中，语音编码系统 500 被构造成提供从接收机处解码器到接收机处编码器、从接收机处编码器到发射机处解码器、从发射机处解码器到发射机处编码器的反馈回路。如下所述，从接收机解码器到发射机编码器的反馈回路最好使语音编码系统 500 能够通过避免坏帧记忆的传播，在帧删除的情况下改进性能。

语音编码系统 500 包括第一和第二语音编码装置 502、504。仅仅为了解释的目的第一语音编码装置 502 表示为发射机语音编码装置，第二语音编码装置 504 表示为接收机语音编码装置。第一语音编码装置 502 包括编码器 506 和解码器 508。第二语音编码装置 504 包括编码器 510 和解码器 512。任一语音编码装置 502、504 最好作为 DSP 的一部分实现，并存在于如 PCS 或蜂窝电话系统中的用户单元或基站中、或在卫星系统的用户单元或网关中。

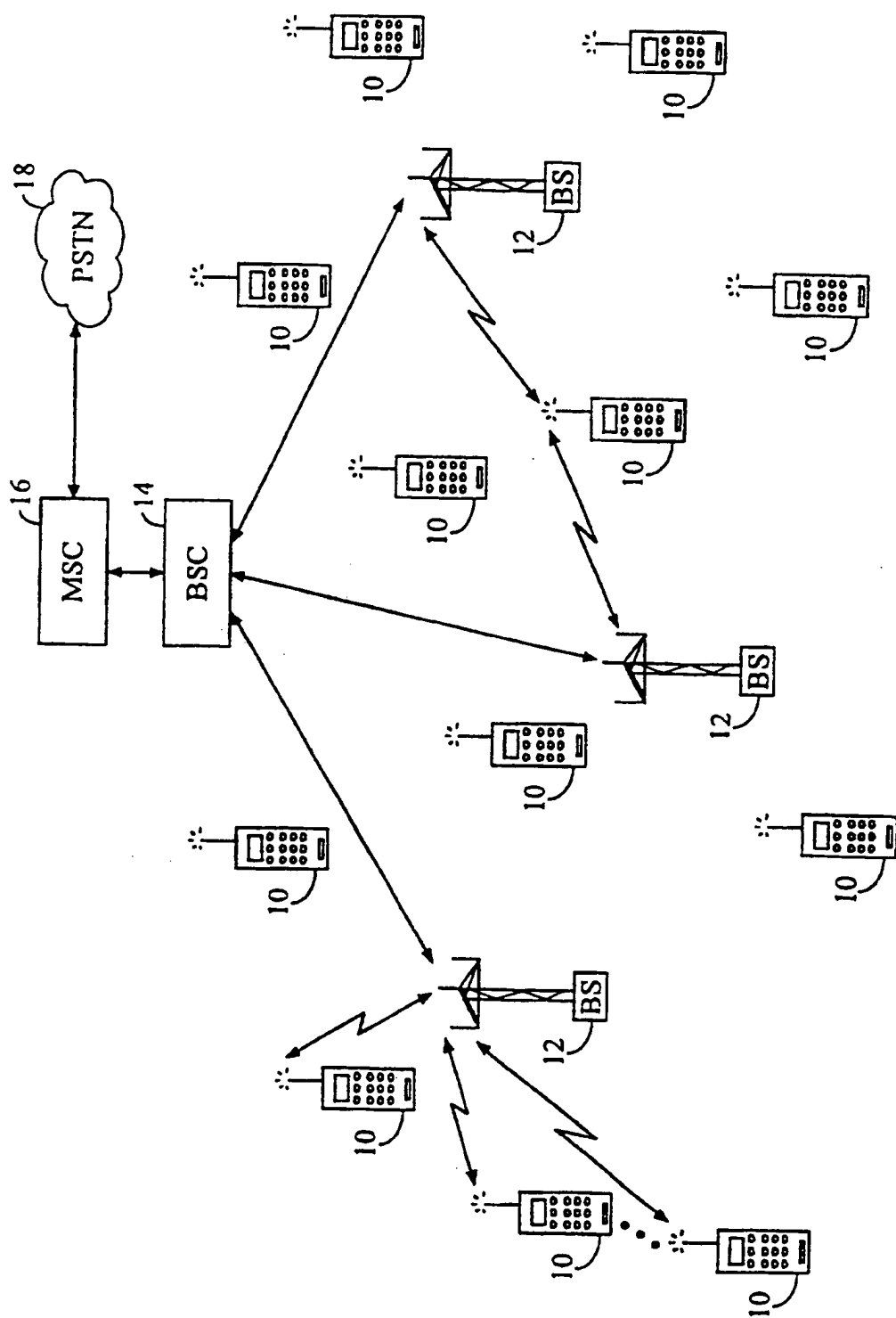
编码器 506 通过通信信道发送数据包。解码器 512 接收数据包。如果帧在传输期间丢失(也就是由于较差或有噪声的信道条件)，解码器 512 发送信号到编码器 510，表示接收到一帧删除。然后编码器 510 在将要发送的下一数据包上将表示为删除符指示比特(EIB)的特定比特值设置为 1。然后编码器 510 发送数据包。解码器 508 接收数据包。解码器 508 发送信号到编码器 506，表示接收到了 EIB 设置为 1 的数据包。根据接收到的来自解码器 508 的信号，编码器 506 发送低记忆编码数据包作为下一数据包。在特殊的实施例中，编码器 506 发送无记忆编码数据包作为下一数据包。

由于以下原因语音编码系统 500 是有益的。通常，预测语音编码装置使用相当大存储量。因此，当编码时每个帧(在特定实施例中，每个帧长 20ms)使用来自过去编码帧的信息。这影响了帧删除情况下的语音编码装置性能。例如，如果一个帧(或多个帧)删除了，在基于预测的语音编码装置(使用来自过去帧的信息，预测当前帧)中该删除之后的帧的质量将受到影响。在低比特率语音编码装置中这尤其显著，其中有大量的预测。然而根据上述实施例，当接收机侧语音解码器 512 接收到删除帧时，解码器 512 发送反馈到发射机

侧的语音编码器 506 表示解码器 512 见到一删除, 因此请求低记忆(最小预测)编码或无记忆(非预测)编码, 使得接收机侧语音解码器 512 的输出和记忆和发射机侧语音编码器 506 的输出和记忆再次同步。因此, 如上所述, 接收机侧的语音解码器 512 通知接收机侧的语音编码器 510 发送 EIB 和下一数据包。然后发射机侧的语音解码器 508 通知发射机侧的语音编码器 506 接收到 EIB。从而发射机侧的语音编码器 506 执行低记忆(最小预测)编码或无记忆(非预测)编码, 并发送相应的数据包到接收机侧的语音解码器 512。然后接收机侧的语音解码器 512 将低记忆或无记忆数据包解码, 并使用解码后的数据包重置或再次同步它的记忆和发射机侧语音编码器 506 的记忆。接收机侧语音解码器 512 在接收到低记忆或无记忆编码数据包之前必须等待的最大时间是一个帧周期(因为接收机侧的编码器 510 可能已经开始产生数据包)加上另一帧周期(因为当接收到 EIB 时发射机侧的编码器 506 可能已经开始产生数据包)再加上单向发送延迟时间。

因此, 已经描述了从解码器向编码器提供反馈以改进帧删除情况下预测语音编码装置性能的新颖方法和装置。本领域熟练的技术人员将理解结合这里所揭示实施例描述的各种说明性逻辑模块和算法步骤可以通过以下装置实现或执行: 数字信号处理器(DSP)、专用集成电路(ASIC)、离散门或晶体管逻辑、离散硬件如寄存器和 FIFO, 执行一组固件指令的处理器, 或任何传统可编程软件模块和处理器。处理器最好是微处理器, 但是另外处理器也可以是任何传统的处理器、控制器、微控制器、或状态机。软件模块可存在于 RAM 存储器、快闪存储器、寄存器或本领域中所熟知的任何其它形式的可写存储媒介。本领域熟练的技术人员还应该理解以上描述中参考的数据、指令、命令、信息、信号、比特、码元和码片最好用电压、电流、电磁波、磁场或粒子、光场或粒子或它们的任何组合表示。

显示并描述了本发明的较佳实施例。然而对本领域熟练的技术人员显而易见的是不脱离本发明的精神和范围可以对这里揭示的实施例进行各种变化。因此, 本发明只受以下权利要求书的限制。



1

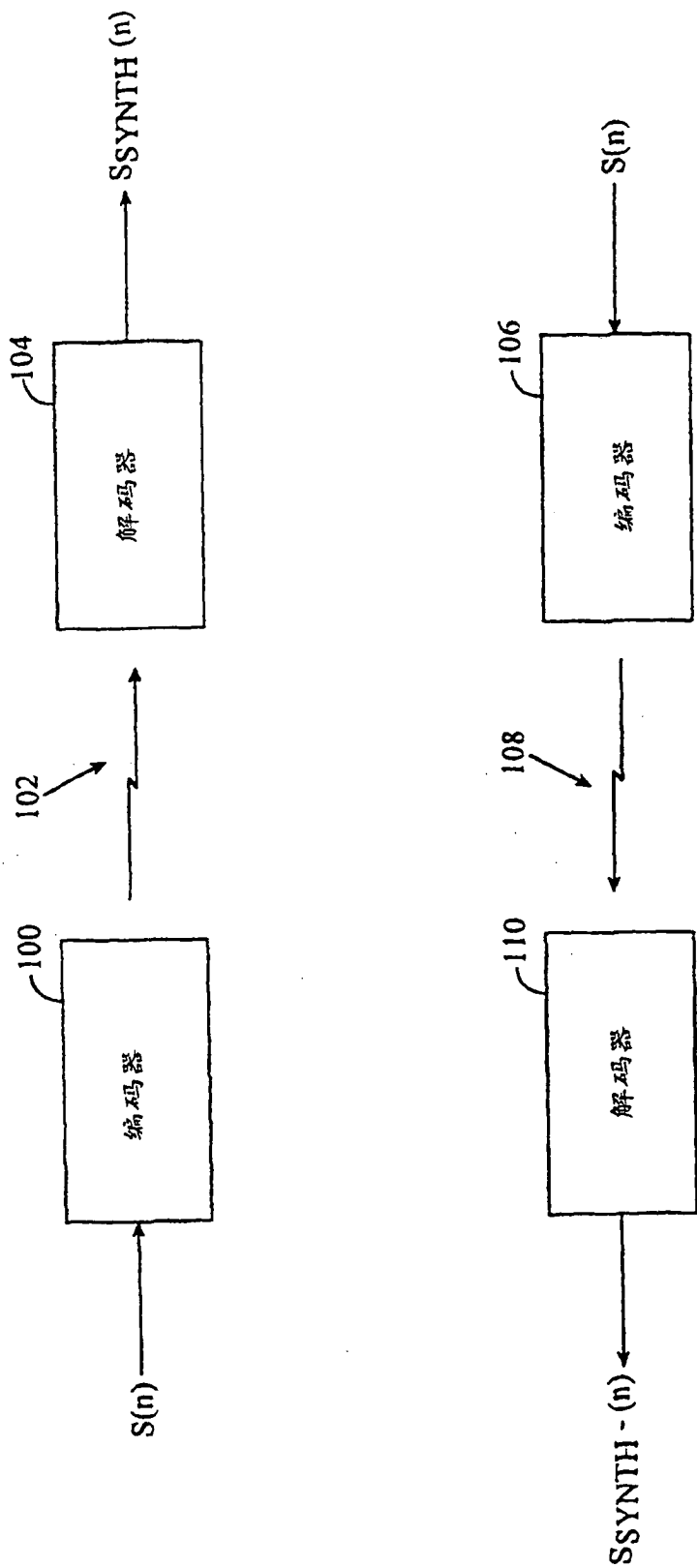


图 2

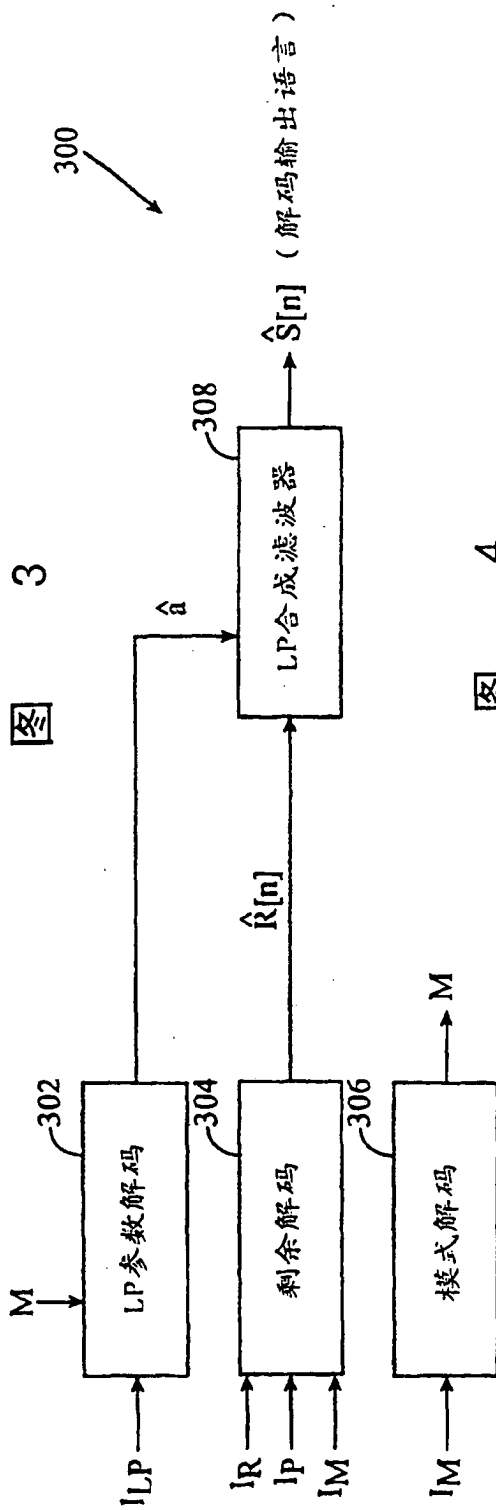
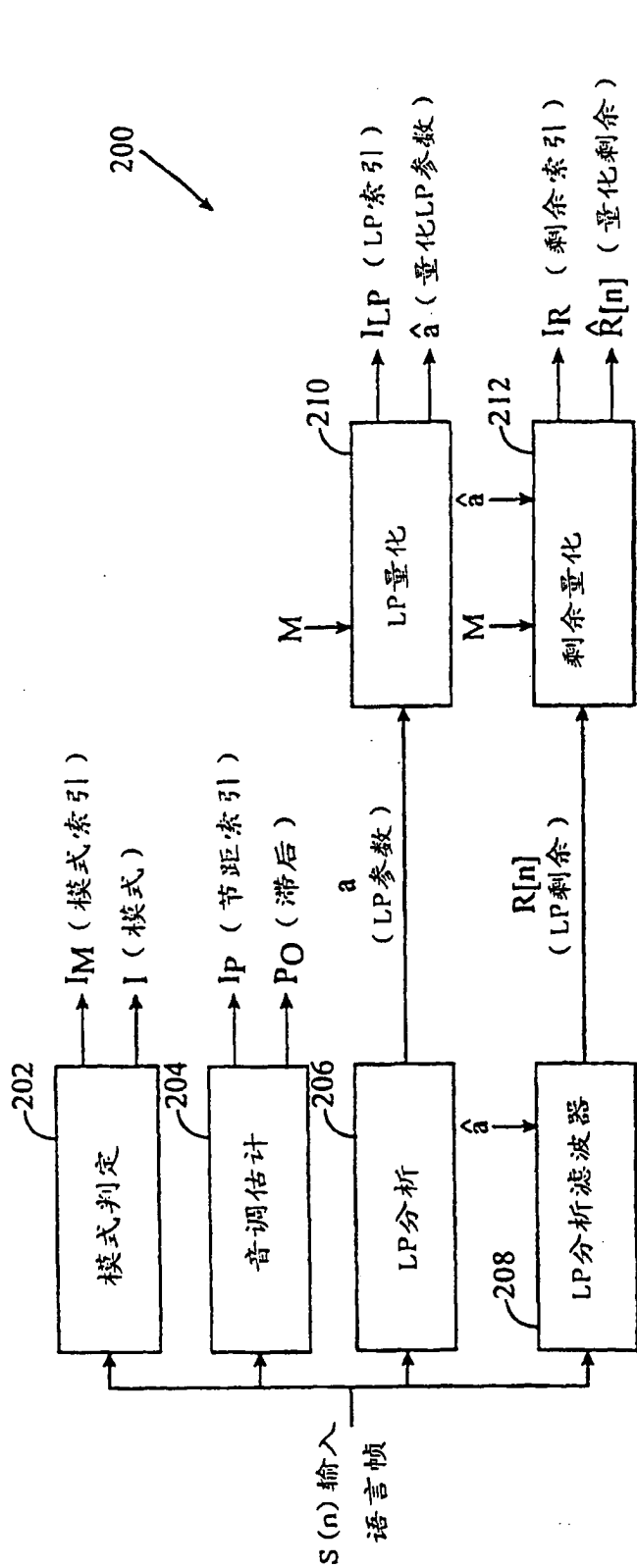


图 4

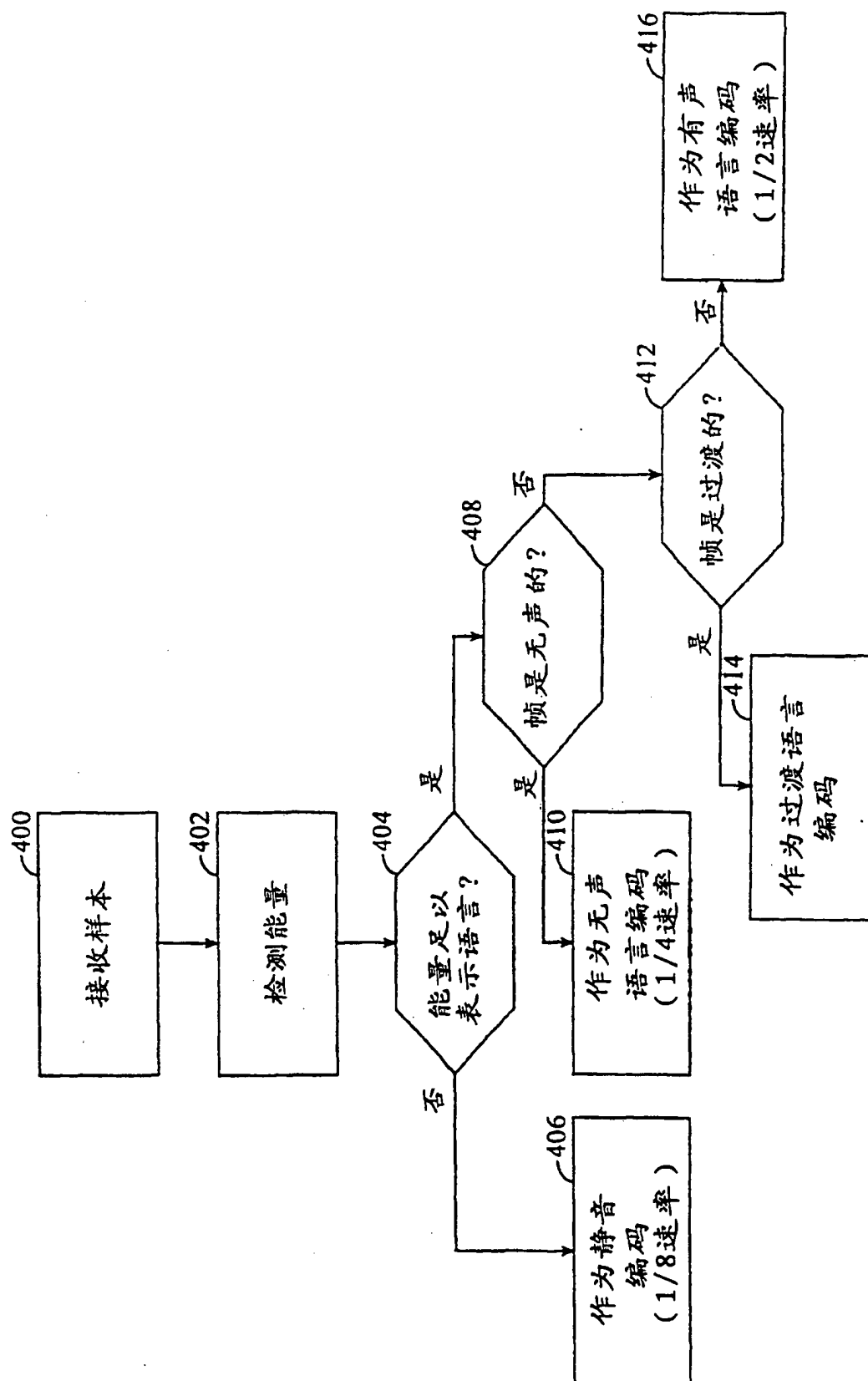


图 5

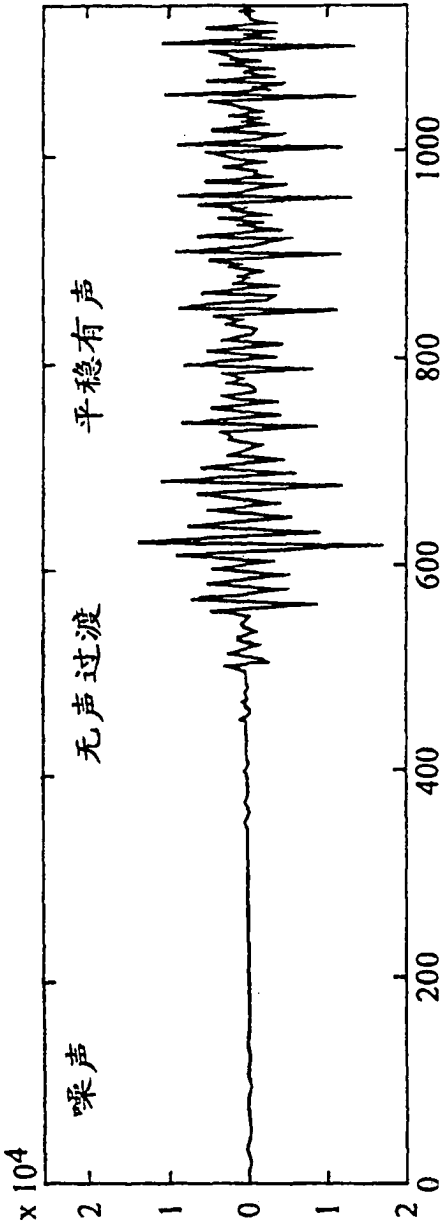


图 6A

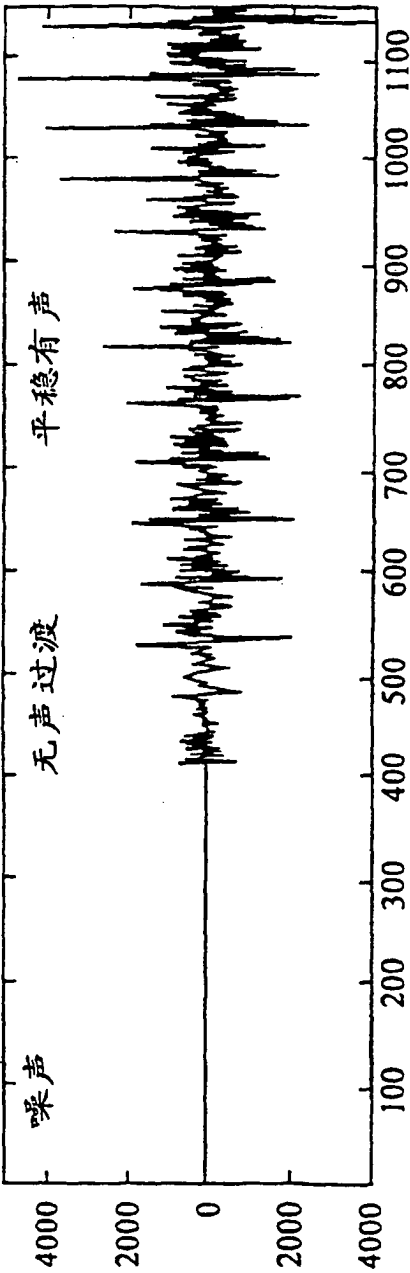


图 6B

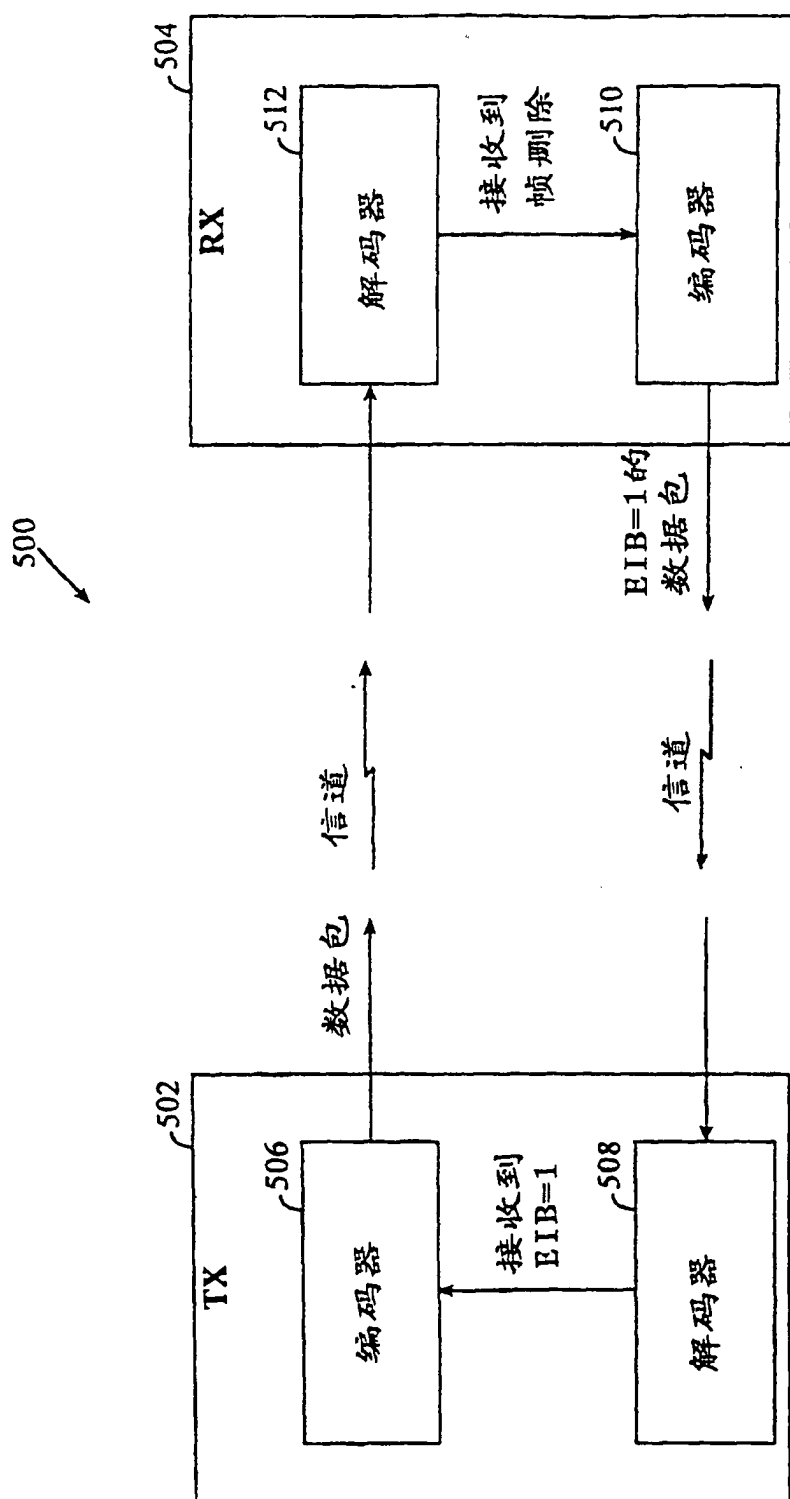


图 7