



# [12] 发明专利说明书

[21] ZL 专利号 00803641.1

[45] 授权公告日 2004 年 6 月 16 日

[11] 授权公告号 CN 1154086C

[22] 申请日 2000.2.14 [21] 申请号 00803641.1

[30] 优先权

[32] 1999. 2. 12 [33] US [31] 09/249,060

[86] 国际申请 PCT/US2000/003855 2000. 2. 14

[87] 国际公布 WO2000/048170 英 2000. 8. 17

[85] 进入国家阶段日期 2001. 8. 10

[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 A·P·德雅科

审查员 杨 叁

[74] 专利代理机构 上海专利商标事务所

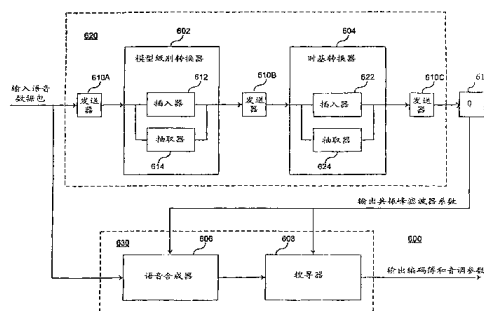
代理人 李家麟

权利要求书 4 页 说明书 10 页 附图 12 页

[54] 发明名称 CELP 转发

[57] 摘要

一种用于基于 CELP 对基于 CELP 声码器数据包变换的方法和装置。装置包括共振峰参数变换器和激励参数变换器。共振峰参数变换器包括模型级别转换器和时基转换器。方法包括下述步骤：将输入数据包的共振峰滤波器系数从输入 CELP 格式变换成输出 CELP 格式，并将输入语音数据包的音调和编码簿参数从输入 CELP 格式变换成输出 CELP 格式。变换共振峰滤波器系数的步骤包括下述步骤：将共振峰滤波器系数模型级别从输入 CELP 格式的模型级别转换成输出 CELP 格式的模型级别，并将得到的系数的时基从输入 CELP 格式时基转换成输出 CELP 格式的时基。



1. 一种将经压缩的语音数据包从一种 CELP 格式转换成另一种 CELP 格式的装置，其特征在于，它包含：

共振峰参数转换器，用来将具有输入 CELP 格式并相应于语音数据包的输入共振峰滤波器系数转换成输出 CELP 格式，而产生输出共振峰滤波器系数；以及

激励参数转换器，用来将具有输入 CELP 格式并相应于所述语音数据包的输入音调和编码簿参数转换成所述输出 CELP 格式，而产生输出音调和编码簿参数，其中，所述激励参数转换器包含：

模型级别转换器，用来将所述输入共振峰滤波器系数的模型级别从所述输入 CELP 格式的模型级别转换成所述输出 CELP 格式的模型级别；

时基转换器，用来将所述输入共振峰滤波器系数的时基从所述输入 CELP 格式的时基转换成所述输出 CELP 格式的时基；

语音合成器，它采用所述输入音调和编码簿参数以及所述输出共振峰滤波器系数，产生一个目标信号；以及

搜寻器，它用所述目标信号和所述输出共振峰滤波器系数搜寻所述输出编码簿和音调参数。

2. 如权利要求 1 所述的装置，其特征在于，所述共振峰参数转换器包含：

模型级别转换器，用来将所述输入共振峰滤波器系数的模型级别从所述输入 CELP 格式的模型级别转换成所述输出 CELP 格式的模型级别；以及

时基转换器，用来将所述输入共振峰滤波器系数的时基从所述输入 CELP 格式的时基转换成所述输出 CELP 格式的时基。

3. 如权利要求 1 所述的装置，其特征在于，所述搜寻器包含：

另一个语音合成器，它用猜想激励参数和所述输出共振峰滤波器系数产生一个猜想信号；

混合器，它根据所述猜想信号和所述目标信号产生一误差信号；以及

最小化元件，它改变所述猜想激励参数，使所述误差信号为最小。

4. 如权利要求 1 所述的装置，其特征在于，所述模型级别转换器还包含：

共振峰滤波器系数变换器，它在所述语音合成器用来产生第三系数以前，将所述输入共振峰滤波器系数转换成第三 CELP 格式。

5. 如权利要求 4 所述的装置，其特征在于，所述模型级别转换器还包含：

插入器，当所述输入 CELP 格式的模型级别低于所述输出 CELP 格式的所述模型级别的时候，它插入所述第三系数，以产生级别经校正的系数；以及

抽取器，当所述输入 CELP 格式的模型级别高于所述输出 CELP 格式的所述模型级别的时候，它抽取所述第三系数，以产生所述级别经校正的系数。

6. 如权利要求 1 所述的装置，其特征在于，所述语音合成器包含：

编码簿，它用所述输入编码簿参数，产生编码簿矢量；

音调滤波器，它用所述输入音调滤波器参数和所述编码簿矢量产生一音调信号；以及

共振峰滤波器，它用所述输出共振峰滤波器系数和所述音调信号产生所述目标信号。

7. 如权利要求 6 所述的装置，其特征在于，所述猜想激励参数包括猜想音调滤波器参数和猜想编码簿参数，其中，所述另一语音合成器包含：

另一编码簿，它用所述猜想编码簿参数产生另一编码簿矢量；

音调滤波器，它用所述猜想音调滤波器参数和所述另一编码簿矢量产生另一音调信号；以及

共振峰滤波器，它用所述输出共振峰滤波器系数和所述另一音调信号产生所述猜想信号。

8. 如权利要求 2 所述的装置，其特征在于，它还包含：

第一共振峰滤波器系数变换器，它在由所述时基转换器使用之前，将所述输入共振峰滤波器系数变换成第四 CELP 格式。

9. 如权利要求 2 所述的装置，其特征在于，它还包含：

第二共振峰滤波器系数变换器，它将所述时基转换器的输出从所述第四 CELP 格式转换成所述输出 CELP 格式。

10. 如权利要求 4 所述的装置，其特征在于，所述第三 CELP 格式是反射系数 CELP 格式。

11. 如权利要求 8 所述的装置，其特征在于，所述第四 CELP 格式是线谱对 CELP 格式。

12. 一种将经压缩的语音数据包从一种 CELP 格式转换成另一种 CELP 格式的方法，其特征在于，它包含下述步骤：

(a) 将与一语音数据包相应的输入共振峰滤波器系数从输入 CELP 格式变换成输出 CELP 格式，而产生输出共振峰滤波器系数；以及

(b) 将与所述语音数据包相应的输入音调和编码簿参数从所述输入 CELP 格式变换成所述输出 CELP 格式，而产生输出音调和编码簿参数，它包含：

(i) 用所述输入 CELP 格式的所述输入音调和编码簿参数以及所述输出共振峰滤波器系数来合成语音，以产生一个目标信号；以及

(ii) 用所述目标信号和所述输出共振峰滤波器系数搜寻所述输出音调和编码簿参数。

13. 如权利要求 12 所述的方法，其特征在于，所述步骤 (a) 包含下述步骤：

(i) 将所述输入共振峰滤波器系数的模型级别从所述输入 CELP 格式的模型级别转换成所述输出 CELP 格式的模型级别；以及

(ii) 将所述输入共振峰滤波器系数的时基从所述输入 CELP 格式的时基转换成所述输出 CELP 格式的时基。

14. 如权利要求 13 所述的方法，其特征在于，所述步骤 (i) 包含下述步骤：

将所述输入共振峰滤波器系数从所述输入 CELP 格式变换成第三 CELP 格式，以产生第三系数；以及

将所述第三系数的模型级别从所述输入 CELP 格式的模型级别转换成所述输出 CELP 格式的模型级别，以产生级别经校正的系数。

15. 如权利要求 14 所述的方法，其特征在于，所述步骤 (ii) 包含下述步骤：

将所述级别经校正的系数变换成第四格式，以产生第四系数；

将所述第四系数的时基从所述输入 CELP 格式的时基转换成所述输出 CELP 格式的时基，以产生时基经校正的系数；以及

将所述时基经校正的系数从所述第四格式变换成所述输出 CELP 格式，而产生所述输出共振峰滤波器系数。

16. 如权利要求 12 所述的方法，其特征在于，所述搜寻步骤 (ii) 包含下述步骤：

用猜想编码簿和音调参数以及所述输出系数产生一猜想信号；

根据所述猜想信号和所述目标信号产生一误差信号；以及

改变所述猜想编码簿和音调参数，而使所述误差信号为最小。

17. 如权利要求 14 所述的方法，其特征在于，所述步骤 (i) 还包含下述步骤：

当所述输入 CELP 格式的所述模型级别低于所述输出 CELP 格式的所述模型级

别的时候，插入所述第三系数，以产生所述级别经校正的系数；以及

当所述输入 CELP 格式的模型级别高于所述输出 CELP 格式的所述模型级别的时候，抽取所述第三系数，以产生所述级别经校正的系数。

18. 如权利要求 14 所述的方法，其特征在于，所述第三 CELP 格式是反射系数 CELP 格式。

19. 如权利要求 15 所述的方法，其特征在于，所述第四 CELP 格式是线谱对 CELP 格式。

## CELP 转发

### 发明背景

#### 发明领域

本发明涉及编码受激线性预测 (CELP) 语音处理。具体地说, 本发明涉及将数字语音数据包从一种 CELP 格式转换成另一种 CELP 格式。

#### 相关技术领域

采用数字技术进行语音传输已日趋广泛, 特别是在长途和数字无线电话中尤其是这样。这接着引起人们对确定在信道上可以发送的最少信息量的同时保持重新构筑的语音的觉察质量这一方面的兴趣。如果语音是通过简单地进行取样和数字化来传送的, 那么在得到传统的模拟电话语音质量的时候, 需要每秒 64 千比 (kbps) 数量级的数据速率。但是, 通过语音分析、随后进行恰当的编码、传输以及在接收机处进行重新合成, 可以使数据速率显著减小。

通常, 将通过获取与人的发音模型有关的参数而对送气的语音进行压缩的装置称之为声码器。这样的装置是由对输入语音进行分析以获取相关参数的编码器以及采用在信道 (如传输信道) 上进行接收的参数而对语音进行再合成的译码器组成的。将语音分成时间段, 或分析子帧, 在其期间计算各参数。接着, 对每一新的子帧, 修改这些参数。

基于线性预测的时域编码器是迄今为止最普通的语音编码器。这些技术在几个过去的取样中, 从输入的语音取样处获取相关性, 并且仅对信号的不相关部分进行编码。这种技术中所使用的基本线性预测滤波器预测当前取样, 作为过去取样的线性组合。一例这种类型的编码规则见 Thomas E. Tremain 等人的论文: “A 4.8 kbps Code Excited Linear Predictive Coder” (Proceeding of the Mobile Satellite Conference, 1988)。

声码器的作用是通过去掉语音中所有固有的自然冗余位而将数字化的语音信号压缩成低数据位速率的信号。通常, 语音中主要由于嘴唇和舌的筛选而具有较短的冗余位, 并且由于声带的振动而具有较长的冗余位。在 CELP 编码器中, 这些工作是由两个过滤器即短时间共振峰 (short-term formant) 的滤波器和长时间音调 (long-term pitch) 的滤波器来形成模型的。一旦去掉了这些冗余位以后, 所

得到的剩余信号可以形成白高斯噪声，而该白高斯噪声也是经过编码的。

这一技术的基本点是计算两个数字滤波器的参数。一个滤波器称为共振峰滤波器（也称为“LPC（线性预测系数）滤波器”），执行对语音波形的短期预测。另一个滤波器称为音调滤波器，执行对语音波形的长期预测。最后，还必须对这些滤波器进行激励，并且这是在波形对上述两个滤波器进行激励时通过确定编码簿中的几个随机激励波形中哪一个波形最接近原始语音来完成的。因此，所传送的参数涉及三项：（1）LPC 滤波器，（2）音调滤波器以及（3）编码簿激励。

可以将数字语音编码分成两个部分；即，编码和译码，有时也称为分析和合成。图 1 是用于对语音进行数字编码、传输和译码的系统 100 的方框图。该系统包括编码器 102、信道 104 和译码器 106。信道 104 可以是通信系统信道、存储介质等。编码器 102 接收数字化的输入语音、获取描述语音特征的特征，并将这些参数量化成发送到信道 104 的数据位流源。译码器 106 从信道 104 接收数据位流，并用接收到的数据位流中的量化特征来重新构筑输出语音波形。

当前，有许多种格式的 CELP 编码可供使用。为了成功地对 CELP 编码的语音信号进行编码，译码器 106 必须采用相同的 CELP 编码模型（也称为“格式”），来作为产生信号的编码器 102。当采用不同的 CELP 格式的通信系统必须共享语音数据时，要求经常地将语音信号从一种 CELP 编码格式转换成另一种编码格式。

一种传统的转换方法是已知的“汇接编码”。图 2 是用来从输入 CELP 格式转换成输出 CELP 格式的汇接编码系统 200 的方框图。系统包括输入 CELP 格式译码器 206 和输出 CELP 格式编码器 202。输入格式的 CELP 译码器 206 接收已经用一种 CELP 格式（下文中称之为“输入”格式）进行了编码的语音信号（下文中称之为“输入”信号）。译码器 206 对输入信号进行译码，以产生语音信号。输出 CELP 格式编码器 202 接收经译码的语音信号，并用输出 CELP 格式（下文中称为“输出”格式）对其进行编码，以产生输出格式的输出信号。这种方法的主要缺点是在通过多个编码器和译码器时语音信号所经受的觉察劣化。

### 发明概述

本发明是基于 CELP 对基于 CELP 的声码器数据包转换的方法和装置。本发明的装置包括共振峰参数转换器，用来将用于语音数据包的输入共振峰滤波器系数从一种 CELP 格式转换成输出 CELP 格式，以生成输出共振峰滤波器系数；本发明的装置还包括一个激励参数转换器，用来将与语音数据包相应的输入音调和编码簿参数

从一种输入的 CELP 格式转换成输出的 CELP 格式，以产生输出音调和编码簿参数。共振峰参数转换器包括一个模型级别 (order) 转换器，将输入共振峰滤波器的系数的模型级别从输入格式的模型级别转换成输出 CELP 格式的模型级别；本发明的共振峰参数转换器还包括一个时基转换器，用来将输入共振峰滤波器系数的时基从输入 CELP 格式的时基转换成输出 CELP 格式的时基。

本发明的方法包括下述步骤，即，将输入数据包的共振峰滤波器系数从输入 CELP 格式转换成输出 CELP 格式，并将输入语音数据包的音调和编码簿参数从输入 CELP 格式转换成输出 CELP 格式。转换共振峰滤波器系数的步骤包括下述步骤，即，将共振峰滤波器系数从输入 CELP 格式转换成反射系数 CELP 格式、将反射系数的模型级别从输入 CELP 格式的模型级别转换成输出 CELP 格式的模型级别、将合成系数转换成线谱对 (LSP) CELP 格式、将合成系数的时基从输入 CELP 格式时基转换成输出 CELP 格式的时基，以及将合成的系数从 LSP 格式转换成输出 CELP 格式，以生成输出共振峰滤波器系数。转换音调和编码簿参数的步骤包括下述步骤，即，用输入音调和编码簿参数来合成语音，以产生目标信号，并用目标信号和输出共振峰滤波器系数来搜寻输出音调和编码簿参数。

本发明的优点是消除了通常由汇接编码转换所引起的觉察语音质量的劣化。

### 附图简述

读者在阅读了本发明的详细描述以后，将会更清楚地了解本发明的特征、目的和优点。图中，相同的标号所表示的意义相同。

图 1 是对语音进行数字编码、传送和译码的系统的方框图；

图 2 是从输入 CELP 格式转换成输出 CELP 格式的汇接编码系统的方框图；

图 3 是 CELP 译码器的方框图；

图 4 是 CELP 编码器的方框图；

图 5 是描述按照本发明的实施例用于基于 CELP 对基于 CELP 的声码器的数据包变换的方法的流程图；

图 6 描绘的是按照本发明的实施例的基于 CELP 对基于 CELP 的声码器数据包转换器；

图 7、8 和 9 是描绘按照本发明的实施例的共振峰参数转换器工作情况的流程图；

图 10 是描绘按照本发明的实施例的激励参数转换器的工作情况的流程图；



图 11 是描绘搜寻器的工作情况的流程图；以及  
图 12 是激励参数转换器更详细的图。

### 较佳实施例的详细描述

下面详细讨论本发明的较佳实施例。读者应当理解，所讨论的特定步骤、结构和排列仅仅是用于描述的。本领域中的普通技术人员应当理解，在不偏离本发明的精神和范围的情况下，也可以采用其他的步骤、结构和排列。本发明可以用在各种各样的信息和通信系统包括卫星和地面蜂窝电话系统之中。一种较佳的应用是在 CDMA 无线扩展谱通信系统中用于电话服务。

下面分两个步骤来描述本发明。首先描述 CELP 编译码器，包括 CELP 编码器和 CELP 译码器。接着，按照一种较佳实施例来描述数据包转换器。

在描述一种较佳实施例之前，首先描述图 1 所示的典型 CELP 系统的结构。在该结构中，CELP 编码器 102 采用分析—合成的方法来对语音信号进行编码。按照该方法，采用开环的方法计算某些语音参数，而通过试错，以闭环的方式决定其他的语音参数。具体地说，通过求解一组方程式，来决定 LPC 系数。接着，将 LPC 系数施加到共振峰滤波器。随后，再用该共振峰滤波器，采用其余参数（编码簿索引、编码簿增益、音调滞后以及音调增益）的假定值来合成语音信号。然后，将合成的语音信号与实际的语音信号比较，决定这些其余参数中的哪些假定值是合成最精确的语音信号的。

### 受激编码的线性预测（CELP）译码器

语音译码过程包含打开数据包、对接收的参数进行去量化处理，以及由这些参数重新构筑语音信号。语音信号的重新构筑包括采用语音参数对产生的编码簿矢量进行滤波。

图 3 是 CELP 译码器 106 的方框图。CELP 译码器 106 包括编码簿 302、编码簿增益元件 304、音调滤波器 306、共振峰滤波器 308 以及后置滤波器 310。下面综述每一方框的一般用途。

共振峰滤波器 308 也称为 LPC 合成滤波器，可以被看作是模拟声道的舌、齿和唇的，并且其谐振频率接近声道筛选（filtering）所引起的原始语音的谐振频率。共振峰滤波器 308 是一种具有下述形式的数字滤波器：

$$1/A(z) = 1 - a_1z^{-1} - \dots - a_nz^{-n} \quad (1)$$

共振峰滤波器 308 的系数  $a_1 \cdots a_n$  称为共振峰滤波器系数或 LPC 系数。

可以将音调滤波器 306 看作是在浊音发声期间从声带产生的周期脉冲串。浊音是通过声带与肺部气流向外的力之间复杂的非线性相互作用而产生的。浊音的例子如单词“low”中的“O”以及单词“day”中的“A”。在发清音时，音调滤波器基本上是从输入到输出保持不变的。清音是通过强迫使气流通过声道中某一点处收缩而产生的。清音的例子如单词“these”中的“TH”，它是由舌和上齿之间的收缩形成的；以及单词“shuffle”中的“FF”，它是通过下唇与上齿之间的收缩而形成的。音调滤波器 306 是具有下述形式的数字滤波器：

$$1/P(z) = 1(1 + bz^{-L}) = 1 + bz^{+L} + b^2z^{+2L} + \dots$$

式中， $b$  称为滤波器的音调增益，而  $L$  是滤波器的音调滞后。

可以将编码簿 302 看作是清音中的湍流噪声，以及对浊音中声带的刺激。在背景喧闹和寂静期间，编码簿输出由随机噪声所取代。编码簿 302 存储几个称作是编码簿矢量的数据字。编码簿矢量是按照编码簿索引  $I$  来选择的。按照编码簿增益参数  $G$ ，由增益元件 304 来选择编码簿矢量的比例。编码簿 302 可以包括增益元件 304。因此，我们也将编码簿的输出称作是编码簿矢量。增益元件 304 可以用例如乘法器来构成。

采用滤波器 310 来对因参数量化以及编码簿的不完善而加入的量化噪声。该噪声在信号能量很小的频带中可以是显著的，而在信号能量较大的频带中却是觉察不到的。为了利用该性能，后置滤波器 310 尝试在觉察不到的频率范围内加入更多的量化噪声，而在觉察明显的频率范围内加入较少的噪声。这种后置滤波的进一步的讨论见 J-H Chen 和 A. Gersho 的文章：“Real-Time Vector APC Speech Coding at 4800 bps with Adaptive Postfiltering”(Proc. ICASSP(1987))以及 N. S Jayant 和 V. Ramamoorthy 的文章：“Adaptive Postfiltering of Speech”(Proc. ICASSP 829-32) (1986 年 4 月，日本，东京)。

在一种实施例中，每一帧的数字化语音包含一个或多个子帧。对于每一子帧，将一组语音参数施加到 CELP 译码器 106，以产生一个子帧的合成语音  $\cdot(n)$ 。该语音参数包括：编码簿索引  $I$ 、编码簿增益  $G$ 、音调滞后  $L$ 、音调增益  $b$  以及共振峰滤波器系数  $a_1 \cdots a_n$ 。按照索引  $I$  选择编码簿 302 的一个矢量，按照增益  $G$  取比例，并用来激励音调滤波器 306 和共振峰滤波器 308。音调滤波器 306 按照音调增益  $b$  和音调滞后  $L$  对选择的编码簿矢量进行运算。共振峰滤波器 308 按照共振峰滤波器系数  $a_1 \cdots a_n$  对音调滤波器 306 产生的信号进行操作，以产生合成的语音信号  $\cdot(n)$ 。

### 受激编码的线性预测 (CELP) 编码器

CELP 语音编码程序包含确定译码器的输入参数，而这些输入参数使合成的语音信号和输入的数字化语音信号之间的觉察差异为最小。下面描述每一组参数的选择处理过程。编码过程还包括使参数量化，并使之分组而成为用于传输的数据包，这是相关技术领域中的普通技术人员所知道的。

图 4 是 CELP 编码器 102 的方框图。CELP 编码器 102 包括编码簿 302、编码簿增益元件 304、音调滤波器 306、共振峰滤波器 308、觉察加权滤波器 410、LPC 发生器 412、加法器 414 以及最小化元件 416。CELP 编码器 102 接收被分隔到几个帧和子帧的数字语音信号  $s(n)$ 。对于每一子帧来说，CELP 编码器 102 产生一组描述该子帧中的语音信号的参数。将这些参数量化，并传送到 CELP 译码器 106。象上文中描述的那样，CELP 译码器 106 采用这些参数来合成语音信号。

参照图 4，以开环方式产生 LPC 系数。采用相关技术领域中所众所周知的方法，从每一子帧的输入语音取样  $s(n)$ ，LPC 发生器 412 计算 LPC 系数。将这些 LPC 系数馈送到共振峰滤波器 308。

然而，通常采用闭环方式（通常也称为是分析—合成的方法）来计算音调参数  $b$  和  $L$  以及编码簿参数  $I$  和  $G$ 。按照该方法，将编码簿和音调参数的假设候选值施加到 CELP 编码器，以合成语音信号  $\hat{s}(n)$ 。在加法器 414 处，将每一猜想的合成语音信号  $\hat{s}(n)$  与输入的语音信号  $s(n)$  比较。将通过比较得到的误差信号  $r(n)$  提供到最小化元件 416。最小化元件 416 选择猜想编码簿和音调参数的不同组合，并决定使误差信号  $r(n)$  最小化的组合。使这些参数和 LPC 发生器 412 产生的共振峰滤波器系数量化，并使之分组用于传输。

在图 4 所示的实施例中，由觉察加权滤波器 410 来对输入语音取样  $s(n)$  加权，从而将经过加权的语音信号提供到加法器 414 的求和输入端。采用觉察加权 (perceptual weighting) 在信号功率较小的频率下对误差进行加权。正是在这些低信号功率频率下，噪声显得更为明显。觉察加权的进一步的讨论见美国专利 5,414,796，其标题是“Variable Rate Vocoder”，并在此引入作为参考。

最小化元件 416 分两个阶段对编码簿和音调参数进行搜寻。首先，最小化元件 416 搜寻音调参数。在音调搜寻期间，是没有来自编码簿的贡献的 ( $G=0$ )。在最小化元件 416 中，音调滞后参数  $L$  和音调增益参数  $b$  的所有可能值都被输入到音调滤波器 306。最小化元件 416 选择使经加权的输入语音和合成的语音之间的误差

$r(n)$ 为最小的那些  $L$  和  $b$  的值。

在找到了音调滤波器的音调滞后  $L$  和音调增益  $b$  以后，以类似的方式进行编码簿搜寻。最小化元件 416 接着产生编码簿索引  $I$  和编码簿增益  $G$  的值。在增益元件 304 中，将按照编码簿索引  $I$  选择的来自编码簿 302 的输出值与编码簿增益  $G$  相乘，得到音调滤波器 306 中使用的值的序列。最小化元件 416 选择使误差  $r(n)$  为最小的编码簿索引  $I$  和编码簿增益  $G$ 。

在一种实施例中，采用觉察加权滤波器 410 对输入语音而采用共振峰滤波器 308 中的加权功能对合成的语音都执行觉察加权。在另一种实施例中，觉察加权滤波器 410 是放在加法器 414 后面的。

#### 基于 CELP 对基于 CELP 的声码器数据包转换

在下文的讨论中，将要转换的语音数据包称作是“输入”数据包，该数据包具有指定“输入”编码簿和音调参数以及“输入”共振峰滤波器系数的“输入”CELP 格式。同样，将变换的结果称为具有指定“输出”编码簿和音调参数以及“输出”共振峰滤波器系数的“输出”CELP 格式的“输出”数据包。这种变换的一种有用的应用是使无线电话系统与互联网接口相连，用来交换语音信号。

图 5 示出的是描述按照较佳实施例的方法的流程图。整个变换分为三个阶段。在第一个阶段，如步骤 502 所示的那样，将输入语音数据包的共振峰滤波器系数从输入 CELP 格式变换成输出 CELP 格式。在第二阶段，如在步骤 504 中所示的那样，将输入语音数据包的音调和编码簿参数从输入 CELP 格式变换成输出 CELP 格式。在第三阶段，用输出 CELP 量化器将输出参数量化。

图 6 描绘的是按照较佳实施例的数据包变换器 600。数据包变换器 600 包括共振峰参数变换器 620 和激励参数变换器 630。共振峰参数变换器 620 将输入共振峰滤波器系数变换成输出 CELP 格式，以产生输出共振峰滤波器系数。共振峰参数变换器 620 包括模型级别转换器 602、时基转换器 604 以及共振峰滤波器系数变换器 610A、B、C。激励参数变换器 630 将输入音调和编码簿参数变换成输出 CELP 格式，以产生输出音调和编码簿参数。激励参数变换器 630 包括语音合成器 606 和搜寻器 608。图 7、8 和 9 是描绘按照较佳实施例的共振峰参数变换器 620 的运行的流程图。

输入语音数据包由变换器 610A 接收。变换器 610A 将每一输入语音数据包的共振峰滤波器系数从输入 CELP 格式变换成适合用于模型级别转换的 CELP 格式。CELP 格式的模型级别描述的是该格式所采用的共振峰滤波器系数的个数。在一种

较佳实施例中，如步骤 702 中所示的那样，输入共振峰滤波器系数被转换成反射系数格式。将反射系数格式的模型级别选择为与输入共振峰滤波器系数格式的模型级别是相同的。执行这样的变换的方法在相关技术领域是众所周知的。当然，如果输入 CELP 格式采用反射系数格式共振峰滤波器系数，那么这样的变换就是不必要的了。

如步骤 704 中所示的那样，模型级别转换器 602 从变换器 610A 接收反射系数，并将反射系数的模型级别从输入 CELP 格式的模型级别转换成输出 CELP 格式的模型级别。模型级别转换器 602 包括插入器 612 和抽取器 614。当输入 CELP 格式的模型级别低于输出 CELP 格式的模型级别的时候，那么如步骤 802 中所示的那样，插入器 612 执行插入操作，以给出附加的系数。在一种实施例中，附加的系数被设置为零。当输入 CELP 格式的模型级别高于输出 CELP 格式的模型级别的时候，如步骤 804 所示的那样，抽取器 614 执行抽取操作，以减少系数的数量。在一种实施例中，简单地就用零来代替不必要的系数的。这种插入和抽取操作在相关技术领域是众所周知的。在系数反射域模型中，级别转换相对来说比较简单，所以似乎是一种合适的选择。当然，如果输入和输出 CELP 格式的模型级别是相同的，那么模型级别转换就是不必要的了。

转换器 610B 从模型级别转换器 602 接收级别经校正的共振峰滤波器系数，并将这些系数从反射系数格式转换成适合于时基转换的 CELP 格式。CELP 格式的时基描述的是对共振峰合成参数进行取样的速率，即，每秒共振峰合成参数的矢量数。在一种较佳实施例中，如步骤 706 所示的那样，将反射系数转换成线谱对 (LSP) 格式。执行这种变换的方法在相关技术领域是众所周知的。

如步骤 708 所示，时基转换器 604 从变换器 610B 接收 LSP 系数，并将 LSP 系数的时基从输入 CELP 格式的时基转换成输出 CELP 格式的时基。时基转换器 604 包括插入器 622 和抽取器 624。当输入 CELP 格式的时基低于输出 CELP 格式的时基的时候（即，每秒采用更少的取样数），如步骤 902 所示的那样，插入器 622 执行插入操作，以增加取样数。当输入 CELP 格式的时基高于输出 CELP 格式的模型级别的时候（即，每秒采用更多的取样数），那么如步骤 904 所示的那样，抽取器 624 执行抽取操作，以减少取样数。这样的插入和抽取操作在本领域中是众所周知的。当然，如果输入 CELP 格式的时基与输出 CELP 格式的时基是相同的，那么就不必有时基转换了。

如步骤 710 所示的那样，变换器 610C 从时基转换器 604 接收经时基校正的共

共振滤波器系数，并将这些系数从 LSP 格式转换成输出 CELP 格式，以产生输出共振峰滤波器系数。当然，如果输出 CELP 格式采用 LSP 格式共振峰滤波器系数，那么该变换就是不必要的了。如步骤 712 所示的那样，量化器 611 从变换器 610C 接收输出共振峰滤波器系数，并量化输出该共振峰滤波器系数。

在变换的第二阶段，如步骤 504 所示的那样，将输入语音数据包的音调和编码簿参数（也称为“激励”参数）从输入 CELP 格式转换成输出 CELP 格式。图 10 是描述按照本发明的较佳实施例的激励参数变换器 630 的操作的流程图。

参照图 6，语音合成器 606 接收每一输入语音数据包的音调和编码簿参数。语音合成器 606 采用输出共振峰滤波器系数产生称为“目标信号”的语音信号，而输出共振峰滤波器系数是由共振峰参数变换器 620 产生的，并且还产生输入编码簿和音调激励参数，如步骤 1002 所示。接着，如上所述，在步骤 1004，搜寻器 608 用与上述 CELP 译码器 106 所使用的类似的搜寻程序得到输出编码簿和音调参数。搜寻器 608 随后使输出参数量化。

图 11 是描绘按照本发明的较佳实施例的搜寻器 608 运行的流程图。搜寻中，如步骤 1104 所示的那样，搜寻器 608 采用共振峰参数变换器 620 所产生的输出共振峰筛选系数和语音合成器 606 的产生的目标信号以及候选编码簿和音调参数来产生候选信号。如步骤 1006 所示的那样，搜寻器 608 将目标信号和候选信号比较，以产生误差信号。搜寻器 608 接着改变候选编码簿和音调参数，如步骤 1008 中所示的那样，使误差信号为最小。选择使误差信号为最小的音调和编码簿参数的组合，作为输出激励参数。这些过程将在下文中作更详细的描述。

图 12 更详细地描绘了激励参数变换器 630。如上所述，激励参数变换器 630 包括语音合成器 606 和搜寻器 608。参照图 12，语音合成器 606 包括编码簿 302A、增益元件 304A、音调滤波器 306A 和共振峰滤波器 308A。如同上文中针对译码器 106 所描述的那样，语音合成器 606 根据激励参数和共振峰滤波器系数，产生一个语音信号。具体地说，语音合成器 606 用输入激励参数和输出共振峰滤波器系数产生一个目标信号  $s_T(n)$ 。将输入编码簿索引  $I_1$  施加到编码簿 302A，以产生一个编码簿矢量。由增益元件 304A 采用输入编码簿增益参数  $G_1$  对编码簿矢量取比例。音调滤波器 306A 用已取比例的编码簿矢量和输入音调增益和音调滞后参数  $b_1$  和  $L_1$  产生音调信号。共振峰滤波器 308A 用由共振峰参数变换器 620 产生的音调信号和输出共振峰滤波器系数  $a_{01} \cdots a_{0n}$ ，产生目标信号  $s_T(n)$ 。本领域中的普通技术人员将会理解，输入和输出激励参数的时基可以是不同的，但所产生的激励信号具有相同的时

基（按照一种实施例，为每秒 8000 个激励取样）。所以，激励参数的时基插入在这处理过程中是固有的（inherent）。

搜寻器 608 包括第二语音合成器、加法器 1202 以及最小化元件 1216。第二语音合成器包括编码簿 302B、增益元件 304B、音调滤波器 306B 以及共振峰滤波器 308B。如上文中针对译码器 106 所描述的那样，第二语音合成器根据激励参数和共振峰滤波器系数，产生语音信号。

具体地说，语音合成器 606 用共振峰参数变换器 620 生成的候选激励参数和输出共振峰滤波器系数产生候选信号  $s_c(n)$ 。将猜想的编码簿索引  $I_c$  施加到编码簿 302B 以产生编码簿矢量。由增益元件 304B 采用输入编码簿增益参数  $G_c$  对编码簿矢量取比例。音调滤波器用已取了比例的编码簿矢量以及输入音调增益和音调滞后参数  $b_c$  和  $L_c$  产生音调信号。共振峰滤波器 308B 用该音调信号和输出共振峰滤波器系数  $a_{01} \cdots a_{0n}$ ，产生猜想的信号  $s_c(n)$ 。

搜寻器 608 将候选信号与目标信号进行比较，产生误差信号  $r(n)$ 。在一种较佳实施例中，将目标信号  $s_T(n)$  施加到加法器 1202 的和输入端，而将猜想信号  $s_c(n)$  施加到加法器 1202 的差输入端。加法器 1202 的输出是误差信号  $r(n)$ 。

将误差信号  $r(n)$  提供到最小化元件 1216。最小化元件 1216 选择编码簿和音调参数的不同组合，并采用与上述针对 CELP 编码器 102 的最小化元件 416 类似的方法，确定使误差信号  $r(n)$  为最小的组合。使通过搜寻得到的编码簿和音调参数量化，并采用由数据包变换器 600 的共振峰参数变换器产生和量化的共振峰滤波器系数来产生在输出 CELP 格式中的语音数据包。

## 结论

上文中较佳实施例的描述使得本领域中的普通技术人员能够制造和使用本发明。很明显，本领域中的普通技术人员还能对这些实施例作各种修改，并且在没有发明人的帮助的情况下，将本文中所揭示的原理应用于其他的实施例。所以，本发明并非仅限于这里所描述的实施例，读者应当从最宽的范围来理解本文中所揭示的原理和新特征。

100

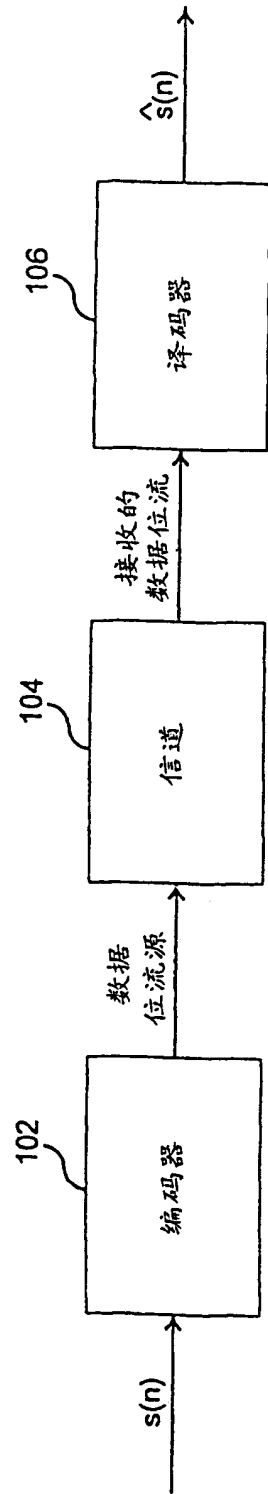


图 1



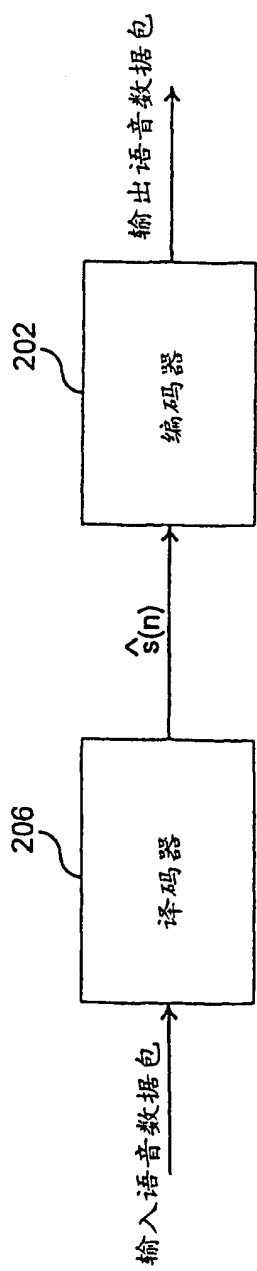


图 2

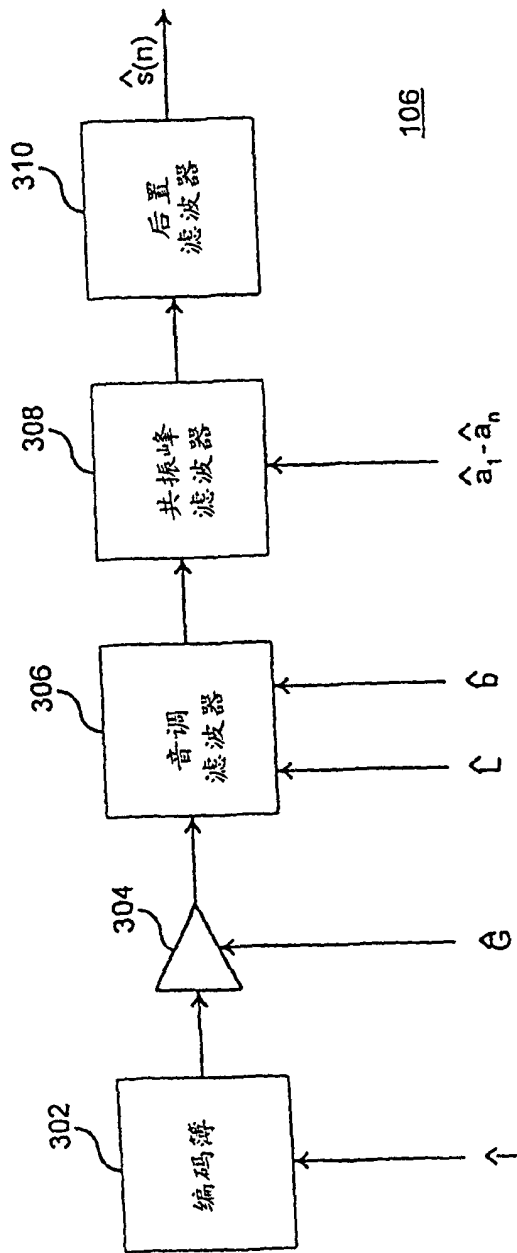


图 3

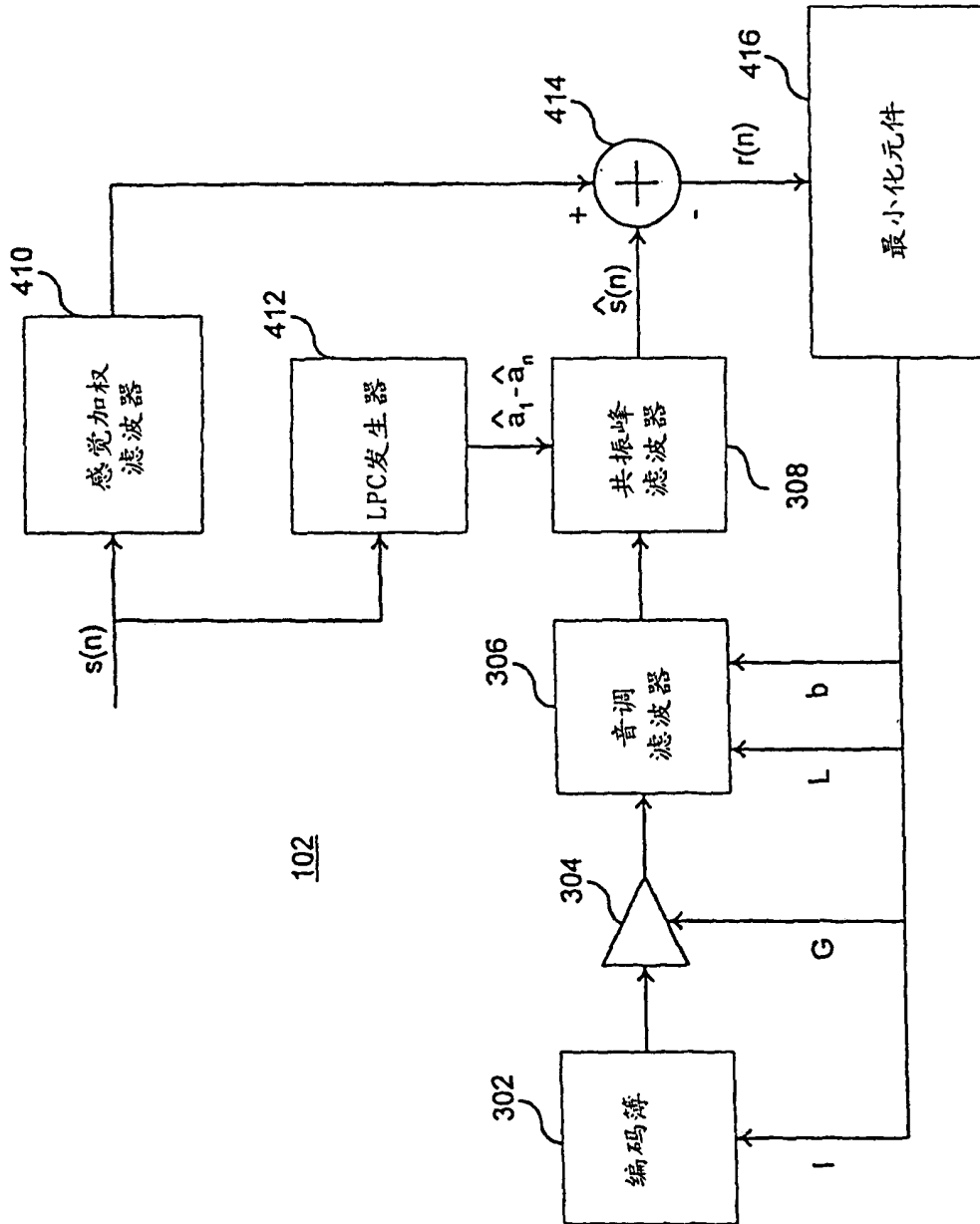


图 4

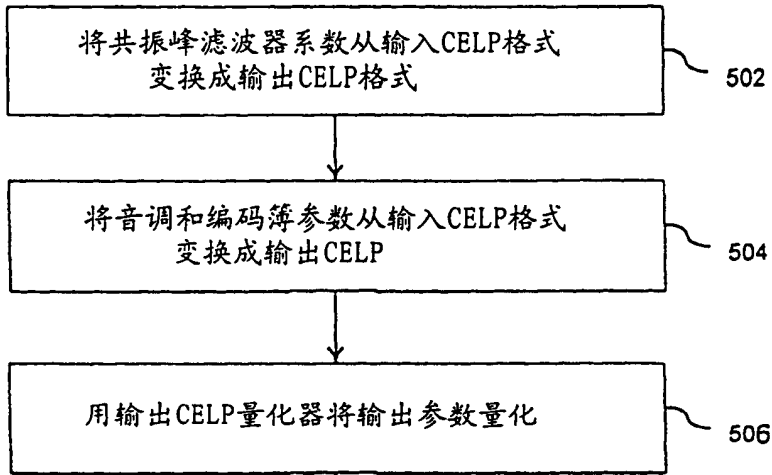


图 5

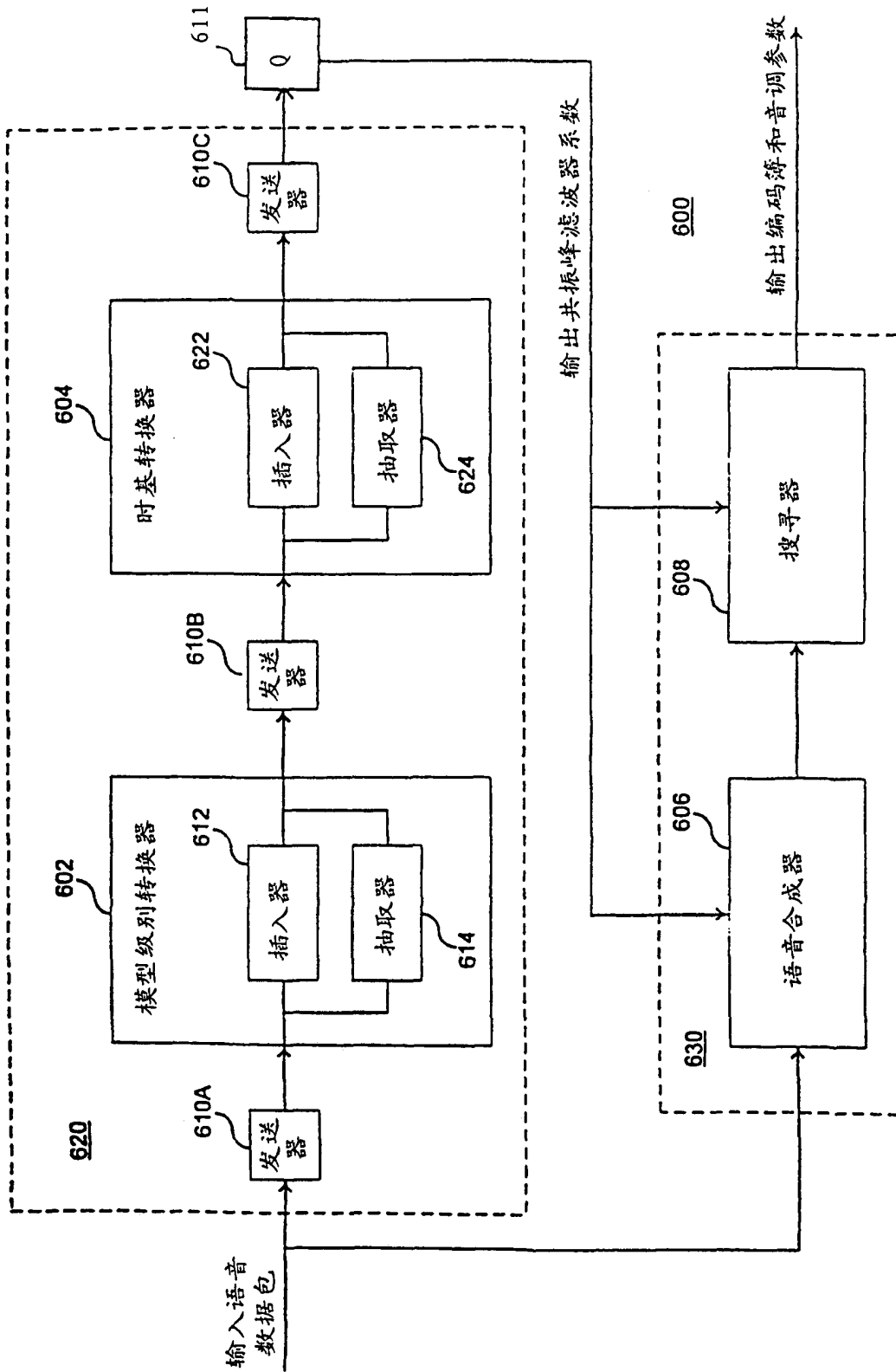


图 6

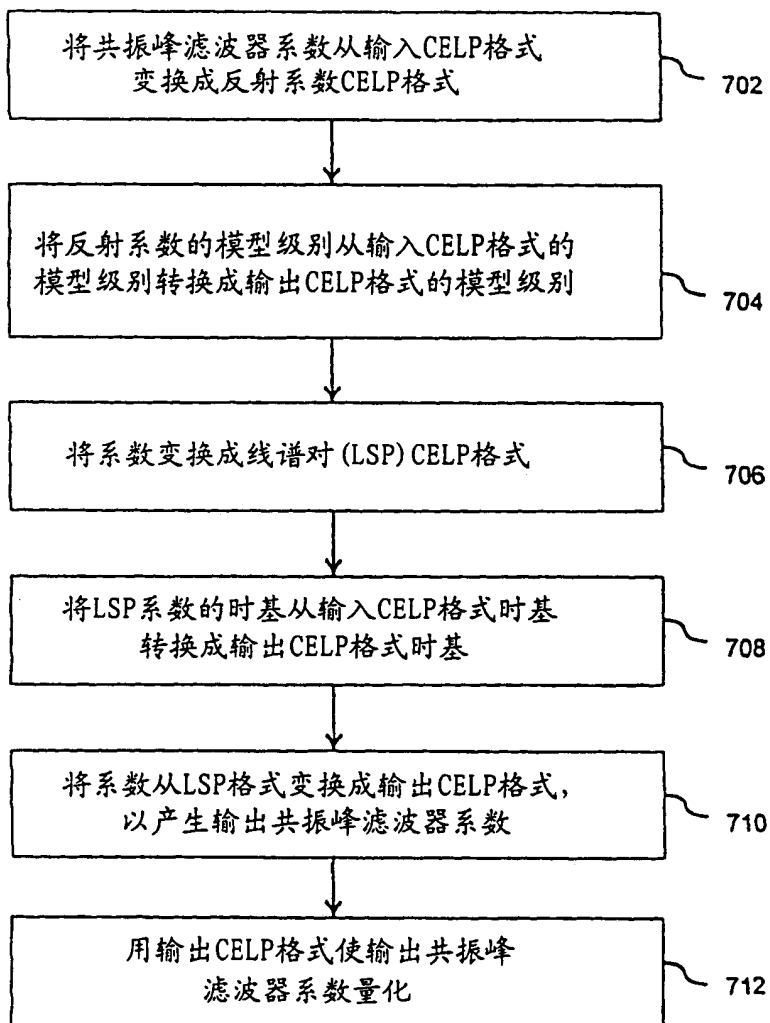


图 7

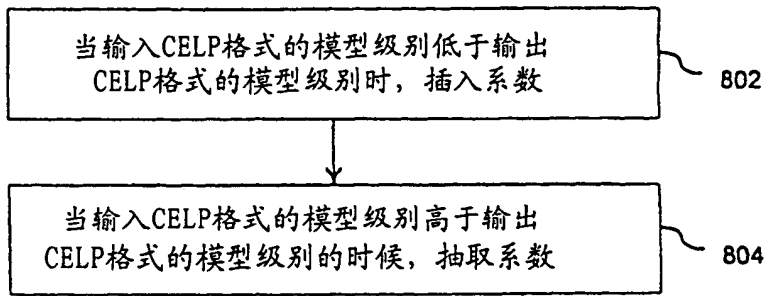


图 8

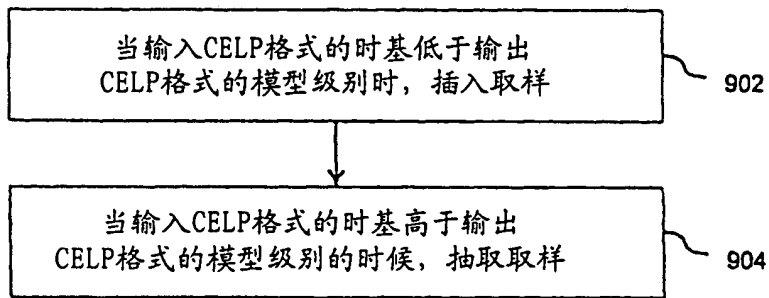


图 9



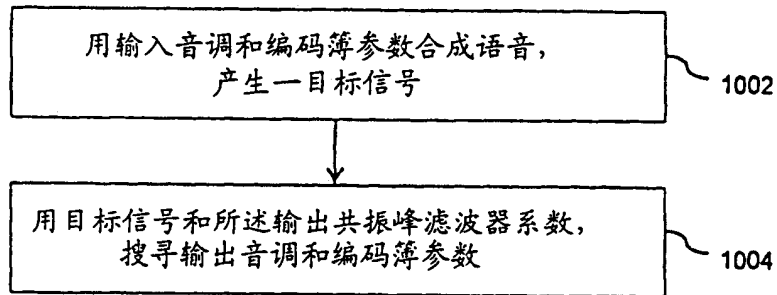


图 10

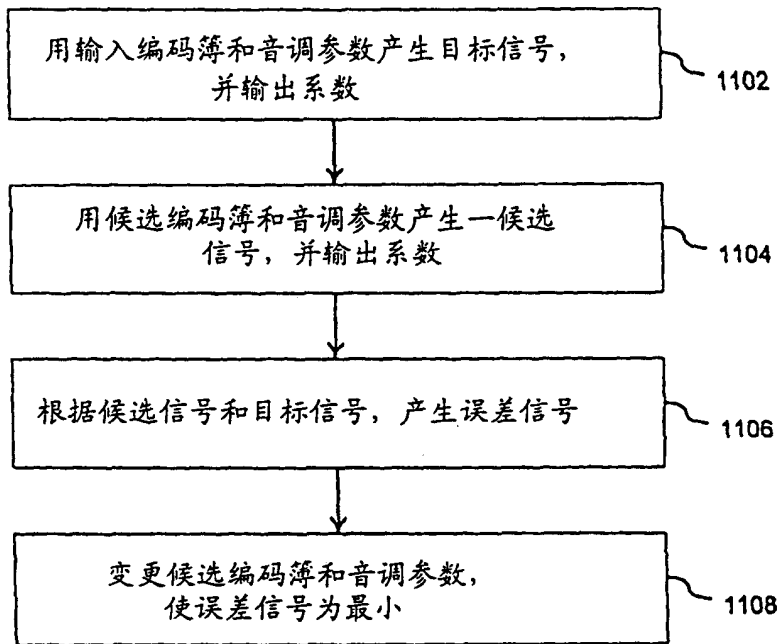


图 11

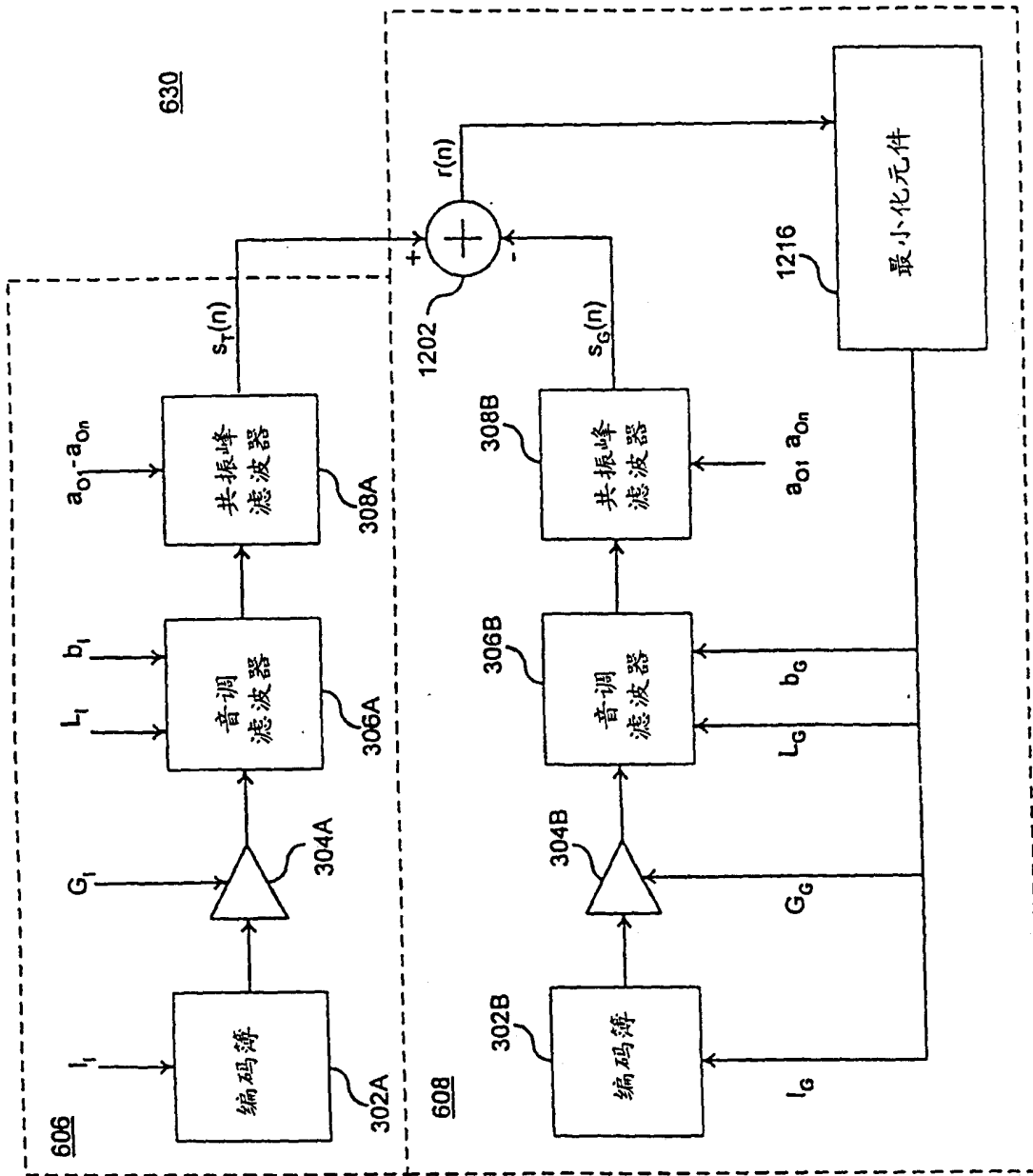


图 12