

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G10L 9/14

G10L 3/00



[12] 发明专利说明书

[21] ZL 专利号 98800967.6

[45] 授权公告日 2004 年 4 月 14 日

[11] 授权公告号 CN 1145925C

[22] 申请日 1998.6.11 [21] 申请号 98800967.6

[30] 优先权

[32] 1997.7.11 [33] EP [31] 97202166.1

[86] 国际申请 PCT/IB1998/000923 1998.6.11

[87] 国际公布 WO99/03097 英 1999.1.21

[85] 进入国家阶段日期 1999.3.10

[71] 专利权人 皇家飞利浦电子有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 R·陶里 R·J·斯勒伊特

A·J·格尔里茨

审查员 杨 叁

[74] 专利代理机构 中国专利代理(香港)有限公司

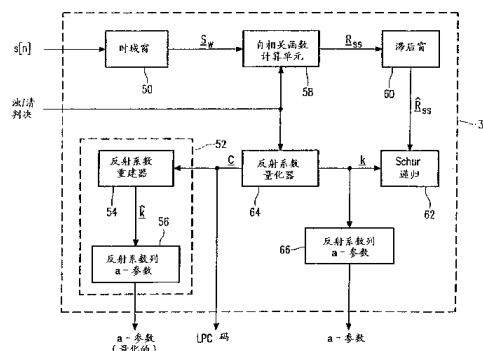
代理人 邹光新 傅 康

权利要求书 2 页 说明书 13 页 附图 9 页

[54] 发明名称 具有改进语音编码器和解码器的发射机

[57] 摘要

在语音编码器(4)中,使用浊音编码器(16)和清音编码器(14)编码语音信号。两个语音编码器(14,16)都使用分析系数表示语音信号。根据本发明,当检测到从浊音向清音(或反之)过渡时,更频繁地确定分析系数。



1. 一种包括具有语音编码器的发射机的传输系统，所述语音编码器包括从语音信号周期性地确定分析系数的分析装置，发射机包括通过传输媒介向接收机发射所述分析系数的发射装置，所述接收机包括具有重建装置的语音解码器，重建装置基于分析系数获得重建的语音信号，其特征

5 在于分析装置在浊音段和清音段之间过渡附近更频繁地确定分析系数，并且重建装置基于更频繁确定的分析系数得到重建的语音信号。

2. 根据权利要求1的传输系统，其特征

10 在于语音编码器包括编码浊音段的浊音编码器，同时语音编码器包括编码清音段的清音编码器。

3. 根据权利要求1或2的传输系统，其特征

在于分析装置对过渡之后的两段更频繁地确定分析系数。

4. 根据权利要求1, 2或3的传输系统，其特征

15 在于分析装置在浊音和清音段或间过渡时，加倍确定分析系数的频率。

5. 根据权利要求4的传输系统，其特征

在于，如果不发生过渡，分析装置每20毫秒确定一次分析系数，同时如果发生过渡，分析装置每10毫秒确定一次分析系数。

6. 一种具有语音编码器的发射机，该语音编码器包括从语音信号周期性地确定分析系数的分析装置，发射机包括发射所述分析系数的发射装置，其特征

20 在于分析装置在浊音段和清音段之间过渡附近更频繁地确定分析系数。

7. 一种接收包括多个周期性的分析系数的编码的语音信号的接收机，所述接收机包括语音解码器，该语音解码器包括基于接收信号中提取的分析系数，获得重建的语音信号的重建装置，其特征

25 在于编码的语音信号携带在浊音段和清音段之间过渡附近更频繁的分析系数，并且重建装置基于更频繁的可用分析系数得到重建的语音信号。

8. 一种包括从语音信号周期性地确定分析系数的分析装置的语音编码方案，其特征

30 在于分析装置在浊音段和清音段之间过渡附近更频繁地确定分析系数。

9. 一种解码包括多个周期性的分析系数的编码的语音信号的语音解码方案，所述语音解码方案包括基于接收信号中提取的分析系数，

获得重建的语音信号的重建装置，其特征在于编码的语音信号携带在浊音段和清音段之间过渡附近更频繁的分析系数，并且重建装置基于更频繁的可用分析系数得到重建的语音信号。

5 10. 一种包括从语音信号周期性地确定分析系数的语音编码方法，其特征在于所述方法包括在浊音段和清音段之间过渡附近更频繁地确定分析系数。

10 11. 一种解码包括多个周期性的分析系数的编码的语音信号的语音解码方法，所述方法包括基于接收信号中提取的分析系数，获得重建的语音信号，其特征在于编码的语音信号携带在浊音段和清音段之间过渡附近更频繁的分析系数，并且基于更频繁的可用分析系数得到重建的语音信号。

12. 一种编码的语音信号，该信号包括在其中周期性地引入的多个分析系数，其特征在于编码的语音信号在浊音段和清音段之间过渡附近更频繁地携带分析系数。

15

具有改进语音编码器和解码器的发射机

技术领域

本发明涉及一个传输系统，该系统包括具有语音编码器的发射机，
5 该语音编码器包括从语音信号中周期性地确定分析系数的分析装置，
该发射机包括通过传输媒介向接收机发射所述分析系数的发射装置，
所述接收机包括具有重建装置的语音解码器，重建装置基于分析系数
获得重建的语音信号。

本发明还涉及发射机，接收机，语音编码器，语音解码器，语音编
10 码方法，语音解码方法，和包括计算机程序在内实现所述方法的有形
媒介。

背景技术

前述的传输系统可从 EP 259 950 得知。

这种传输系统和语音编码器用于语音信号必须通过传输容量有限
15 的传输媒介传输或必须存储在存储容量有限的存储介质中的应用。这
种应用的例子是在 Internet 上传送语音信号，从移动电话向基站以及
反之传送语音信号，和在 CD-ROM 上，在固态存储器中，或在硬盘驱动
器上存储语音信号。

已经尝试了不同的语音编码器工作原理，以便在合适的比特率上获
20 得合理的语音质量。这些工作原理中的一个就是区分浊音信号和清音
信号。这两类语音信号使用不同的语音编码器编码，每一种编码器都
是针对相应类型的语音信号的特性优化的。

另一种操作类型称为 CELP 编码器，语音信号和使用存储在码本中
的多个激励信号导出的激励信号激励合成滤波器得到的合成语音信号
25 比较。为了处理浊音信号这样的周期性信号，使用了所谓的自适应码
本。

在两类语音编码器中，都必须对描述的语音信号确定分析系数。当
降低语音编码器可用的比特率时，可获得的重建语音的语音质量迅速
恶化。

30 发明内容

本发明的目的是提供一个语音信号传输系统，减少降低比特率带来
的语音质量恶化。

因此,根据本发明的传输系统,其特征在于分析装置在浊音段和清音段(或反之)之间过渡附近更频繁地确定分析系数,并且重建装置基于更频繁确定的分析系数得到重建的语音信号。

5 本发明基于这样的认识,即语音信号质量恶化的一个重要原因是从浊音向清音(或反之)过渡期间,分析参数不足以跟踪变化。通过在这样的过渡附近提高分析参数的更新速率,可以充分地提高语音质量。因为过渡并不是非常经常地出现,处理更频繁更新分析参数所要求的额外比特率是适度的。注意到可以在过渡实际发生之前提高确定分析系数的频率,但是也可以在过渡发生之后提高确定分析系数的频率。
10 结合上述提高确定分析系数频率的方法也是可行的。

本发明的一个实施例的特征在于语音编码包括编码浊音段的浊音编码器,同时语音编码器包括编码清音段的清音编码器。

实验显示,在过渡附近提高分析系数的更新速率所获得的改善对使用浊音和清音解码器的语音编码器特别有利。采用这类语音编码器,
15 可能的改进是相当大的。

本发明的另一个实施例的特征在于使分析装置对过渡之后的两段更频繁地确定分析系数。

已经发现对过渡后的两帧更频繁地确定分析系数会显著地提高语音质量。

20 本发明的另一个实施例的特征在于,分析装置在浊音和清音段(或反之)之间过渡时,加倍确定分析系数的频率。

已经证明加倍确定分析系数的频率足以获得显著提高的语音质量。

附图说明

25 现在将参考附图解释本发明。在这里:

图 1, 可以使用本发明的传输系统;

图 2, 根据本发明的语音编码器 4;

图 3, 根据本发明的浊音编码器 16;

图 4, 在根据图 3 的浊音编码器 16 中使用的 LPC 计算装置 30;

30 图 5, 在根据图 3 的语音编码器中使用的基音调谐装置 32;

图 6, 在根据图 2 的语音编码器中用于清音的的语音编码器 14;

图 7, 在根据图 1 的系统中使用的语音解码器 14;

图 8, 在语音解码器 14 中使用的浊音解码器 94;

图 9, 浊音解码器 94 中各点呈现的信号图;

图 10, 在语音解码器 14 中使用的清音解码器 96.

具体实施例描述

5 在根据图 1 的传输系统中, 语音信号送到发射机 2 的输入端。在发射机 2 中, 语音信号在语音编码器 4 中编码。语音编码器 4 输出端的编码的语音信号被送到发射装置 6。发射装置 6 用于完成对编码的语音信号的信道编码, 交织和调制。

10 发射装置 6 的输出信号送到发射机的输出, 并且通过传输媒介 8 传送到接收机 5。在接收机 5, 信道的输出信号送到接收装置 7。这些接收装置 7 提供 RF 处理, 如调谐和解调, 解-交织 (如果适当) 和信道解码。接收装置 7 的输出信号送到语音解码器 9, 该语音解码器将它的输入信号转换为重建的语音信号。

15 根据图 2, 语音编码器 4 的输入信号 $s.[n]$ 由 DC 陷波滤波器滤波以从输入中消除不希望的 DC 偏置。所述 DC 陷波滤波器的截止频率 (-3dB) 是 15Hz。DC 陷波滤波器 10 的输出信号送到缓冲区 11 的输入。根据本发明, 缓冲区 11 为浊音编码器 16 提供具有 400 个 DC 滤波的语音抽样的块。所述有 400 个抽样的块包括 5 个 10 毫秒的语音帧 (每个 80 个抽样)。它包括要被编码的当前帧, 两个以前和两个后续帧。

20 缓冲区 11 以每帧的间隔将最近接收的有 80 个抽样的帧送到 200Hz 高通滤波器 12 的输入。高通滤波器 12 的输出连接到清音编码器 14 的输入和浊音/清音检测器 28 的输入。高通滤波器 12 为浊音/清音检测器 28 提供有 360 个抽样的块并为清音编码器 14 提供有 160 个抽样的块 (如果语音编码器 4 工作在 5.2 kbit/sec 模式) 或有 240 个抽样的块 (如果

25 语音编码器 4 工作在 3.2 kbit/sec 模式)。上述具有不同抽样的块和缓冲区 11 的输出之间的关系在下表中列出。

组件	5.2 kbit/sec		3.2 kbit/sec	
	抽样数	起始	抽样数	起始
高通滤波器 12	80	320	80	320
浊音/清音检测器 28	360	0...40	360	0...40
浊音编码器 16	400	0	400	0

清音编码器 14	160	120	240	120
将被编码的当前帧	80	160	80	160

浊音/清音检测器 28 确定当前帧是否包含浊音或清音，并且将结果作为浊音/清音标志给出。这个标志送到复用器 22，再送到清音编码器 14 和浊音编码器 16。根据浊音/清音标志的值，激活浊音编码器 16 或清音编码器 14。

在浊音编码器 16 中，输入信号被表示为多个谐波相关的正弦信号。浊音编码器的输出提供一个基音值，一个增益值和一种 16 个预测参数的表示。基音值和增益值被送到复用器 22 相应的输入。

在 5.2 kbit/sec 模式，每 10 毫秒进行一次 LPC 计算。在 3.2 kbit/sec，每 20 毫秒进行一次 LPC 计算，除非在清音到浊音（或反之）之间出现过渡。如果发生这样的过渡，在 3.2 kbit/sec 模式，也是每 10 毫秒进行一次 LPC 计算。

浊音编码器输出的 LPC 系数由霍夫曼编码器 24 编码。在霍夫曼编码器 24 中，比较器比较霍夫曼编码序列的长度和相应的输入序列的长度。如果霍夫曼编码序列的长度大于输入序列的长度，就决定发射未编码的序列。否则决定发射霍夫曼编码序列。所述判决由送到复用器 26 和复用器 22 的“霍夫曼比特”表示。复用器 26 用于根据“霍夫曼比特”的值向复用器 22 传送霍夫曼编码序列或输入序列。在复用器 26 中结合使用“霍夫曼比特”具有确保预测序列的表示的长度不超过一个预定值的好处。不使用“霍夫曼比特”，复用器 26 就可能出现霍夫曼编码序列的长度超过输入序列的长度一部分，这样编码的序列就不再能够放入只为传输 LPC 系数保留了有限数目比特的发射帧。

在清音编码器 14 中确定一个增益值和 6 个预测系数来表示清音信号。这 6 个 LPC 系数由霍夫曼编码器 18 编码，该编码器在它的输出端给出一个霍夫曼编码序列和一个“霍夫曼比特”。霍夫曼编码序列和霍夫曼编码器 18 的输入序列被送到由“霍夫曼比特”控制的复用器 20。霍夫曼编码器 18 和复用器 20 结合的操作与霍夫曼编码器 24 和复用器 20 的操作一样。

复用器 20 的输出信号和“霍夫曼比特”被送到复用器 22 的相应输入端。复用器 22 用于根据浊音-清音检测器 28 的判决选择编码的浊音

信号或编码的清音语音信号。在复用器 22 的输出端得到编码的语音信号。

在根据图 3 的浊音编码器 16 中, 根据本发明的分析装置由 LPC 参数计算单元 30, 精确基音计算单元 32 和基音估计器 38 构成。语音信号 $s[n]$ 送到 LPC 参数计算单元 30 的输入。LPC 参数计算单元 30 确定系数 $a[i]$, 在量化, 编码和解码 $a[i]$ 之后确定量化预测系数 $aq[i]$, 以及确定 LPC 码 $C[i]$, 其中 i 的值从 0-15。

根据发明的概念的基音确定装置包括初始基音确定装置(在这里是基音估计器 38), 和基音调谐装置(在这里是基音范围计算单元 34 和精确基音计算单元 32)。基音估计器 38 确定粗略基音值, 该值由基音范围计算单元 34 用来确定基音值, 这个值再由后面称为精确基音计算单元 32 的基音调谐装置尝试, 确定最终的基音值。基音估计器 38 提供由多个抽样表示的粗略基音周期。精确基音计算单元 32 中使用的基音值是由基音范围计算单元 34 根据下表通过粗略基音周期确定的。

粗略基音周期 p	频率 (Hz)	搜索范围	步长	候选个数
$20 < p < 39$	400...200	$p-3...p+3$	0.25	24
$40 < p < 79$	200...100	$p-2...p+2$	0.25	16
$80 < p < 200$	100...40	P	1	1

在幅度谱计算单元 36 中, 根据下式通过信号 $s[i]$ 确定加窗的语音信号 S_{HAM} :

$$S_{HAM}[i-120] = w_{HAM}[i] \cdot s[i] \quad (1)$$

在 (1) 中 $w_{HAM}[i]$ 等于:

$$w_{HAM} = 0.54 - 0.46 \cos \left\{ \frac{2\pi((i+0.5)-120)}{160} \right\}; \quad 120 \leq i < 280 \quad (2)$$

加窗的语音信号 S_{HAM} 使用 512 点 FFT 变换到频域。所述变换所获得的频谱 S_w 等于:

$$S_w[k] = \sum_{m=0}^{159} S_{HAM}[m] \cdot e^{-j2\pi km/512} \quad (3)$$

其中, 基音计算单元 32 中使用的幅度谱根据下式计算:

$$|S_w[k]| = \sqrt{(\Re\{S_w[k]\})^2 + (\Im\{S_w[k]\})^2} \quad (4)$$

精确基音计算单元 32 通过 LPC 参数计算单元 30 提供的 a-参数和粗略基音值确定精确基音值，该精确基音值使根据 (4) 的幅度谱和包括多个幅度由所述精确基音周期抽样 LPC 谱确定的谐波相关正弦信号的信号的幅度谱之间的误差信号最小。

在增益计算单元 40 中，与目标谱精确匹配的最佳增益是使用量化的 a-参数再合成的语音信号谱计算的，而不是向精确基音计算单元 32 那样使用非量化的 a-参数。

在浊音编码器 40 的输出端，得到 16 个 LPC 码，精确的基音和增益计算单元 40 计算的增益。LPC 参数计算单元 30 和精确基音计算单元 32 的操作在下面做更详细地描述。

在根据图 4 的 LPC 计算单元 30 中，加窗操作是由加窗处理器 50 在信号 $s[n]$ 上进行的。根据本发明的一个方面，分析长度取决于浊音/清音标志的值。在 5.2 kbit/sec 模式，LPC 计算每 10 毫秒执行一次。在 3.2 kbit/sec 模式，LPC 计算每 20 毫秒执行一次，除非在浊音到清音（或反之）过渡期间。如果出现这样的过渡，LPC 计算每 10 毫秒执行一次。

下表中给出了预测系数判决所涉及的抽样数。

比特率和模式	分析长度 NA 和涉及的抽样	更新间隔
5.2 kbit/sec	160 (120-280)	10 毫秒
3.2 kbit/sec (过渡)	160 (120-280)	10 毫秒
3.2 kbit/sec (非过渡)	240 (120-360)	20 毫秒

对出现过渡的 5.2 kbit/sec 情况和 3.2 kbit/sec 情况，窗可以写为：

$$w_{HAM} = 0.54 - 0.46 \cos\left\{\frac{2\pi((i+0.5)-120)}{160}\right\}; \quad 120 \leq i < 280 \quad (5)$$

加窗的语音信号这样建立：

$$S_{HAM}[i-120] = w_{HAM}[i] \cdot s[i]; \quad 120 \leq i < 280 \quad (6)$$

如果在 3.2 kbit/s 情况下没发生过渡，就在窗口的中间引入 80 个抽样的平顶部分将窗口扩展为跨越从第 120 抽样开始并且以第 360 抽样终止的 240 个抽样。这样，根据下式获得窗 w'_{HAM} ：

$$w'_{HAM} = \begin{cases} w_{HAM}[i] & ; 120 \leq i < 200 \\ 1 & ; 200 \leq i < 280 \\ w_{HAM}[i] & ; 280 \leq i < 360 \end{cases} \quad (7)$$

对加窗的语音信号可以写出下式。

$$S_{HAM}[i-120] = w_{HAM}[i] \cdot s[i]; \quad 120 \leq i < 360 \quad (8)$$

5 自相关函数计算单元 58 确定加窗的语音信号的自相关函数 R_{ss} 。所计算的相关系数的个数等于预测系数的个数+1。如果出现浊音帧，所计算的自相关系数的个数是 17。如果出现清音帧，所计算的自相关系数的个数是 7。出现浊音还是清音帧由浊音/清音标志告知自相关函数计算单元 58。

10 自相关系数被所谓的滞后窗加窗以对自相关系数表示的谱获得一些平滑效果。平滑的自相关系数 $\rho[i]$ 根据下式计算：

$$\rho[i] = R_{ss}[i] \cdot \exp\left(\frac{-\pi f_{\mu} i}{8000}\right); \quad 0 \leq i \leq P \quad (9)$$

在 (9) 中， f_{μ} 是值为 46.4 Hz 的谱平滑常数。加窗的自相关值 $\rho[i]$ 送到 Schur 递归模块 62，用递归的方法计算反射系数 $k[1]$ 到 $k[P]$ 。Schur 递归对本领域的技术人员是众所周知的。

15 在变换器 66 中， P 个反射系数 $\rho[i]$ 被变换为在图 3 中的精确基音计算单元 32 中使用的 a -参数。在量化器 64 中，反射系数被变换为对数域比值，随后这些对数域比值被均匀量化。所得到的 LPC 码 $C[1] \dots C[P]$ 送到 LPC 参数计算单元的输出以便进一步传输。

20 在本地解码器 54 中，LPC 码 $C[1] \dots C[P]$ 被反射系数重建器 54 转换为重建的反射系数 $\hat{k}[i]$ 。随后，重建的反射系数 $\hat{k}[i]$ 被反射系数到 a -参数转换器 56 转换为 (量化的) a -参数。

这个本地解码用于得到在语音编码器 4 和语音解码器 14 中可以得到的相同的 a -参数。

25 在根据图 5 的精确基音计算单元 32 中，基音频率候选选择器 70 通过从基音范围计算单元 34 接收的候选的个数，起始值和步长确定用

于精确基音计算单元 32 的候选基音值。对每个候选者，基音频率候选选择器 70 确定基频 $f_{0,i}$ 。

使用候选频率 $f_{0,i}$ ，谱包络抽样器 72 在谐波位置抽样 LPC 系数所描述的谱包络。第 i 个候选者 $f_{0,i}$ 的 k 次谐波的幅度 $m_{i,k}$ 可以写作：

$$5 \quad m_{i,k} = \left| \frac{1}{A(z)} \right|_{z=2\pi k \cdot f_{0,i}} \quad (10)$$

在 (10) 中， $A(z)$ 等于：

$$A(z) = 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} + \dots + a_p \cdot z^{-p} \quad (11)$$

将 $z = e^{j\theta_{i,k}} = \cos \theta_{i,k} + j \cdot \sin \theta_{i,k}$ 和 $\theta_{i,k} = 2\pi k f_{0,i}$ 代入 11，得到：

$$A(z)|_{\theta=\theta_{i,k}} = 1 + a_1(\cos \theta_{i,k} + j \cdot \sin \theta_{i,k}) + \dots + a_p(\cos \theta_{p,k} + j \cdot \sin \theta_{p,k}) \quad (12)$$

10 将 (12) 分为实部和虚部，可以根据下式获得幅度 $m_{i,k}$ ：

$$m_{i,k} = \frac{1}{\sqrt{R^2(\theta_{i,k}) + I^2(\theta_{i,k})}} \quad (13)$$

其中

$$R(\theta_{i,k}) = 1 + a_1(\cos \theta_{i,k}) + \dots + a_p(\cos \theta_{i,k}) \quad (14)$$

而

$$15 \quad I(\theta_{i,k}) = 1 + a_1(\sin \theta_{i,k}) + \dots + a_p(\sin \theta_{i,k}) \quad (15)$$

根据编码器当前的工作模式，将谱线 $m_{i,k}$ ($1 \leq k \leq L$) 与谱的窗函数 W (根据 (5) 或 (7) 得到的 160 点汉明窗的 8192 点 FFT) 卷积得到候选谱 $|\hat{S}_{w,i}|$ 。可以预先计算 8192 点 FFT 并将结果存储在 ROM 中。在卷积处理中，执行了下抽样操作，因为必须将候选谱与 256 点的参考谱比较，
20 多于 256 点的计算是无用的。因此， $|\hat{S}_{w,i}|$ 可以写作：

$$|\hat{S}_{w,i}[f]| = \sum_{k=1}^L m_{i,k} \cdot W(16 \cdot f - k \cdot f_{0,i}); \quad 0 \leq f < 256 \quad (16)$$

表达式 (16) 只对候选基音 i 给出了幅度谱的大体形状，而不是它的

幅度。因此，谱 $|\hat{S}_{w,i}|$ 必须由增益因子 g_i 修正，该增益因子由 MSE-增益计算器 78 根据 (17) 计算：

$$g_i = \frac{\sum_{j=0}^{256} S_w[j] \cdot \hat{S}_{w,i}[j]}{\sum_{j=0}^{256} (S_w[j])^2} \quad (17)$$

乘法器 82 使用增益因子 g_i 缩放谱 $|\hat{S}_{w,i}|$ 。减法器 84 计算幅度谱计算单元 36 确定的目标谱的系数和乘法器 82 的输出信号之差。随后，求和平方器根据下式计算方差信号 E_i 。

$$E_i = E(f_{0,i}) = \sum_{j=0}^{255} \left(|S_w[j]| - g_i \cdot |\hat{S}_{w,i}[j]| \right)^2 \quad (18)$$

产生最小值的候选基频 $f_{0,i}$ 被选做精确的基频或基音。在根据本例的编码器中，共有 368 个可能的基音周期，需要用 9bit 编码。基音每 10 毫秒更新一次而不管语音编码器的工作模式。在根据图 3 的增益计算器 40 中，发射到解码器的增益是用上面对增益 g_i 描述的不同方法计算的，只是使用量化的 a -参数替代计算增益 g_i 时使用的未量化的 a -参数。发射到解码器的增益因子是 6 比特非线性量化的，对小的 g_i 值使用小量化步长，对较大的 g_i 值使用较大的量化步长。

在根据图 6 的清音编码器 14 中，LPC 参数计算单元 82 的操作类似于根据图 4 的 LPC 参数计算单元 30 的操作。LPC 参数计算单元 82 在高通滤波的语音信号上操作，而不象 LPC 参数计算单元 30 那样在原始语音信号上进行。另外，LPC 计算单元 82 的预测阶数是 6，而不是 LPC 参数基音计算单元 30 使用的 16。

时域加窗处理器 84 根据 (19) 计算通过汉宁窗的语音信号：

$$s_w[n] = s[n] \cdot \left(0.5 - 0.5 \cos \left(\frac{2 \cdot \pi (i + 0.5) - 120}{160} \right) \right); \quad 120 \leq i < 280 \quad (19)$$

在 RMS 值计算单元 86 中，根据 (20) 计算语音帧的幅度的平均值：

$$g_{uv} = \frac{1}{4} \sqrt{\frac{1}{N} \sum_{i=0}^{159} s_w^2[i]} \quad (20)$$

发射到解码器的增益因子 g_{uv} 是 5 比特非线性量化的，对小的 g_{uv}

值使用小量化步长，对较大的 g_{uv} 值使用较大的量化步长。清音编码器 14 不确定激励参数。

在根据图 7 的语音编码器中，为霍夫曼解码器 90 提供霍夫曼编码的 LPC 码和浊音/清音标志。如果浊音/清音标志指示浊音信号，霍夫曼解码器 90 根据霍夫曼编码器 18 使用的霍夫曼表解码霍夫曼编码的 LPC 码。根据霍夫曼比特的值，所接收的 LPC 码由霍夫曼解码器 90 解码或直通到解复用器 92。增益值和接收的精确基音值也送到解复用器 92。

如果浊音/清音标志指示浊音帧，就将精确的基音，增益和 16 个 LPC 码送到谐波语音合成器 94。如果浊音/清音标志指示清音帧，则将增益和 6 个 LPC 码送到清音合成器 96。谐波语音合成器 94 输出的合成浊音信号 $\hat{s}_{v,k}[n]$ 和清音合成器 96 输出的合成清音信号 $\hat{s}_{uv,k}[n]$ 一起送到复用器 98 相应的输入端。

在浊音模式，复用器 98 将谐波语音合成器 94 的输出信号 $\hat{s}_{v,k}[n]$ 送到重叠相加综合模块 100 的输入端。在清音模式，复用器 98 将清音合成器 96 的输出信号 $\hat{s}_{uv,k}[n]$ 送到重叠相加综合模块 100 的输入端。在重叠相加模块 100 中，部分重叠的浊音和清音段被加在一起。重叠相加综合模块 100 的输出信号 $\hat{s}[n]$ 可以写为：

$$\hat{s}[n] = \begin{cases} \hat{s}_{uv,k-1}[n + N_s/2] + \hat{s}_{uv,k}[n] & ; v_{k-1} = 0, v_k = 0 \\ \hat{s}_{uv,k-1}[n + N_s/2] + \hat{s}_{v,k}[n] & ; v_{k-1} = 0, v_k = 1 \\ \hat{s}_{v,k-1}[n + N_s/2] + \hat{s}_{uv,k}[n] & ; v_{k-1} = 1, v_k = 0 \\ \hat{s}_{v,k-1}[n + N_s/2] + \hat{s}_{v,k}[n] & ; v_{k-1} = 1, v_k = 1 \end{cases} \quad (21)$$

在 (21) 中， N_s 是语音帧的长度， v_{k-1} 是前一语音帧的浊音/清音标志，而 v_k 是当前语音帧的浊音/清音标志。

重叠和块的输出信号 $\hat{s}[n]$ 送到后置滤波器 102。后置滤波器通过抑制共振区外的噪声增强感知的语音质量。

在根据图 8 的浊音解码器 94 中，基音解码器 104 解码从解复用器 92 接收的编码的基音并将其转换为基音周期。基音解码器 104 确定的基音周期送到相位合成器 106 的输入端，谐波振荡器组 108 的输入端

和 LPC 谱包络抽样器 110 的第一输入端。

LPC 解码器 112 解码从解复用器 92 接收的 LPC 系数。解码 LPC 系数的方法取决于当前语音帧包含浊音还是清音。因此，浊音/清音标志送到 LPC 解码器 112 的第二输入端。LPC 解码器将量化的 a-参数送到
5 LPC 谱包络抽样器 110 的第二输入端。LPC 谱包络抽样器 112 的操作由 (13), (14) 和 (15) 描述, 因为精确基音计算单元 32 完成相同的操作。

相位合成器 106 用于计算表示语音信号的第 i 阶正弦信号的相位 $\varphi_k[i]$ 。选择的相位 $\varphi_k[i]$ 使第 i 阶正弦信号从一帧到下一帧保持连续。浊
10 音信号是通过合并重叠帧合成的, 每一个重叠帧包括 160 个加窗的抽样。从图 9 中的图 118 和 122 可以看出两个相邻帧之间有 50% 重叠。图 118 和 122 中使用的窗用点划线表示。现在, 相位合成器用于在重叠影响最大的位置提供连续的相位。这里使用的窗口函数的这个位置在抽样 119。当前帧的相位 $\varphi_k[i]$ 现在可以写作:

$$15 \quad \varphi_k[i] = \varphi_{k-1}[i] + i \cdot 2\pi \cdot f_{0,k-1} \frac{3N_s}{4} - i \cdot 2\pi \cdot f_{0,k} \frac{N_s}{4}; \quad 1 \leq i \leq 100 \quad (22)$$

在当前描述的语音编码器中 N_s 的值等于 160。对于最初的浊音帧, $\varphi_k[i]$ 的值初始化为一个预定值。相位 $\varphi_k[i]$ 不断地更新, 即使接收到一个清音帧。在这种情况下,

$f_{0,k}$ 被设置为 50Hz。

20 谐波振荡器组 108 产生多个谐波相关的信号 $\hat{s}_{v,k}[n]$ 来表示语音信号。这个计算是使用谐波幅度 $\hat{m}[i]$, 频率 \hat{f}_0 和合成的相位 $\hat{\varphi}[i]$ 根据 (23) 进行的:

$$\hat{s}'_{v,k}[n] = \sum_{i=1}^L \hat{m}[i] \cos\{(i \cdot 2\pi \cdot \hat{f}_0) \cdot n + \hat{\varphi}[i]\}; \quad 0 \leq n < N_s \quad (23)$$

在时域窗模块 114 中使用汉宁窗对信号 $\hat{s}_{v,k}[n]$ 加窗。这个加窗的信号
25 号如图 9 中的图 120 所示。使用在时间上移位 $N_s/2$ 抽样的汉宁窗对信

号 $\hat{s}_{v,k+1}[n]$ 加窗。这个加窗的信号如图 9 中的图 124 所示。将上述加窗的信号相加得到时域窗模块 144 的输出信号。这个输出信号如图 9 中的图 126 所示。增益解码器 118 从它的输入信号中得到增益值 g_v ，而信号缩放模块 116 使用所述增益因子 g_v 缩放时域窗模块 114 的输出信号，从而获得重建的浊音信号 $\hat{s}_{v,k}$ 。

在清音合成器 96 中，LPC 码和浊音/清音标志送到 LPC 解码器 130。LPC 解码器 130 为 LPC 综合滤波器 134 提供多组 6 个 a-参数。高斯白噪声发生器 132 的输出连接到 LPC 综合滤波器 143 的输入端。LPC 综合滤波器 134 的输出信号被时域窗模块 140 中的汉宁窗加窗。

清音增益解码器 136 得到表示当前清音帧的期望能量的增益值 \hat{g}_{uv} 。通过这个增益和加窗的信号的能量，可以确定加窗的语音信号增益的缩放因子 \hat{g}'_{uv} ，以获得具有正确能量的语音信号。这个缩放因子可以写作：

$$\hat{g}'_{uv} = \sqrt{\frac{\hat{g}_{uv}}{\sum_{n=0}^{N_s-1} (\hat{s}'_{uv,k}[n] \cdot w[n])^2}} \quad (24)$$

信号缩放块 142 用缩放因子 \hat{g}'_{uv} 乘时域窗模块 140 的输出信号确定输出信号 $\hat{s}_{uv,k}$ 。

可以改进当前描述的语音编码系统以获得更低的比特率或更高的语音质量。需要更低的比特率的语音编码系统的一个例子是 2 kbit/sec 编码系统。这样的系统可以通过将用于浊音的预测系数的个数从 16 降为 12 并且对预测系数，增益和精确基音使用差分编码获得。差分编码意味着被编码的数据不是独立编码的，而是只发射与后续帧的相应数据之差。在从浊音到清音（或反之）过渡时，第一个新帧的所有系数都独立编码，以为解码器提供初始值。

也可以在 6 kbit/sec 的比特率上获得语音质量更好的语音编码器。这里所做的改进是确定多个谐波相关正弦信号的前 8 个谐波的相位。相位 $\phi[i]$ 是根据 (25) 计算的：

$$\phi[i] = \arctan \frac{I(\theta_i)}{R(\theta_i)} \quad (25)$$

其中 $\theta_i = 2\pi f_0 \cdot i$ 。 $R(\theta_i)$ ， $I(\theta_i)$ 等于：

$$R(\theta_i) = \sum_{n=0}^{N-1} s_w[n] \cdot \cos(\theta_i \cdot n) \quad (26)$$

和

$$I(\theta_i) = -\sum_{n=0}^{N-1} s_w[n] \cdot \sin(\theta_i \cdot n) \quad (27)$$

- 5 这样得到的 8 个相位 $\varphi^{[i]}$ 被均匀量化为 6 比特并且包括在输出比特流中。

- 对 6 kbit/sec 编码器的进一步改进是在清音模式传输附加的增益值。通常是用每 2 毫秒发射一个增益代替每帧一次。在紧接着过渡后的第一帧，发射 10 个增益值，其中 5 个表示当前的清音帧，另 5 个表示清音编码器处理的前一个清音帧。增益是从 4 毫秒的重叠窗中确定的。
- 10

应该注意到 LPC 系数的个数是 12 并且可能使用差分编码。

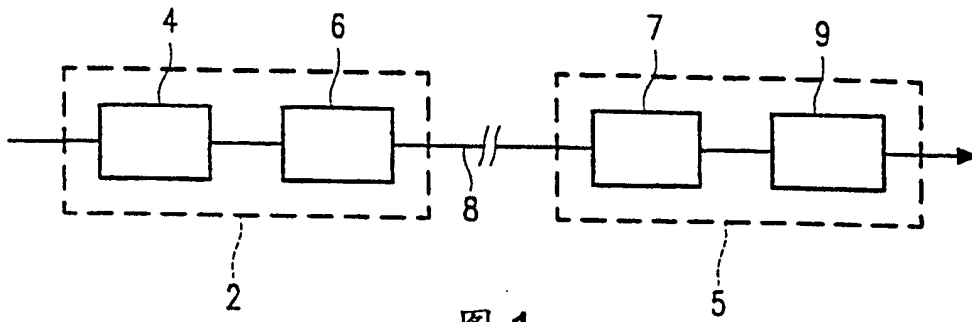


图 1

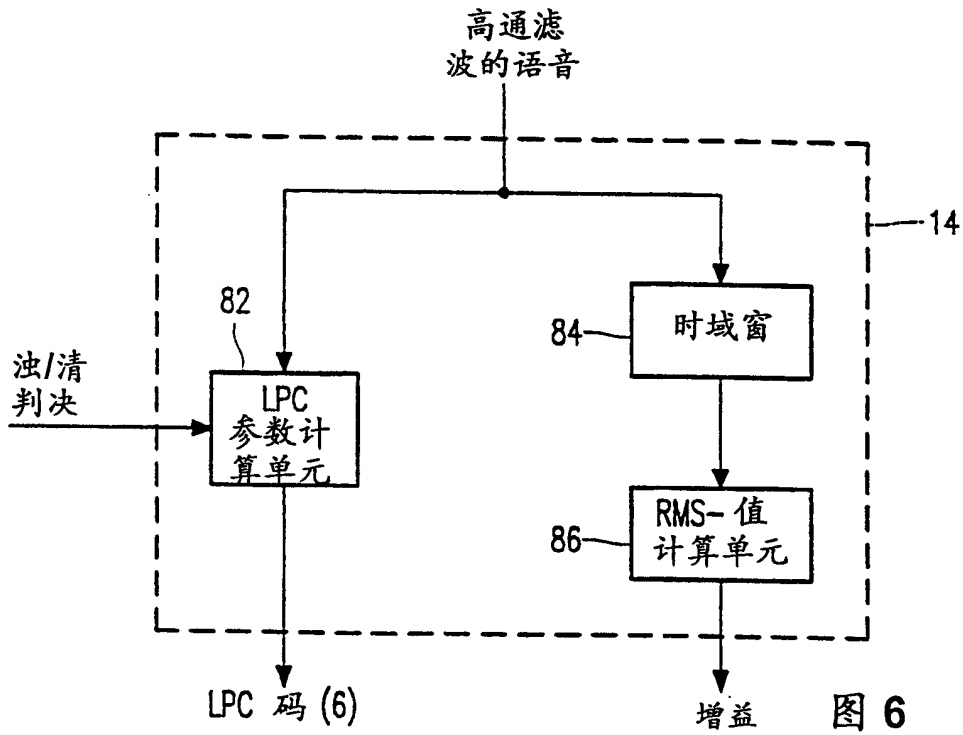


图 6

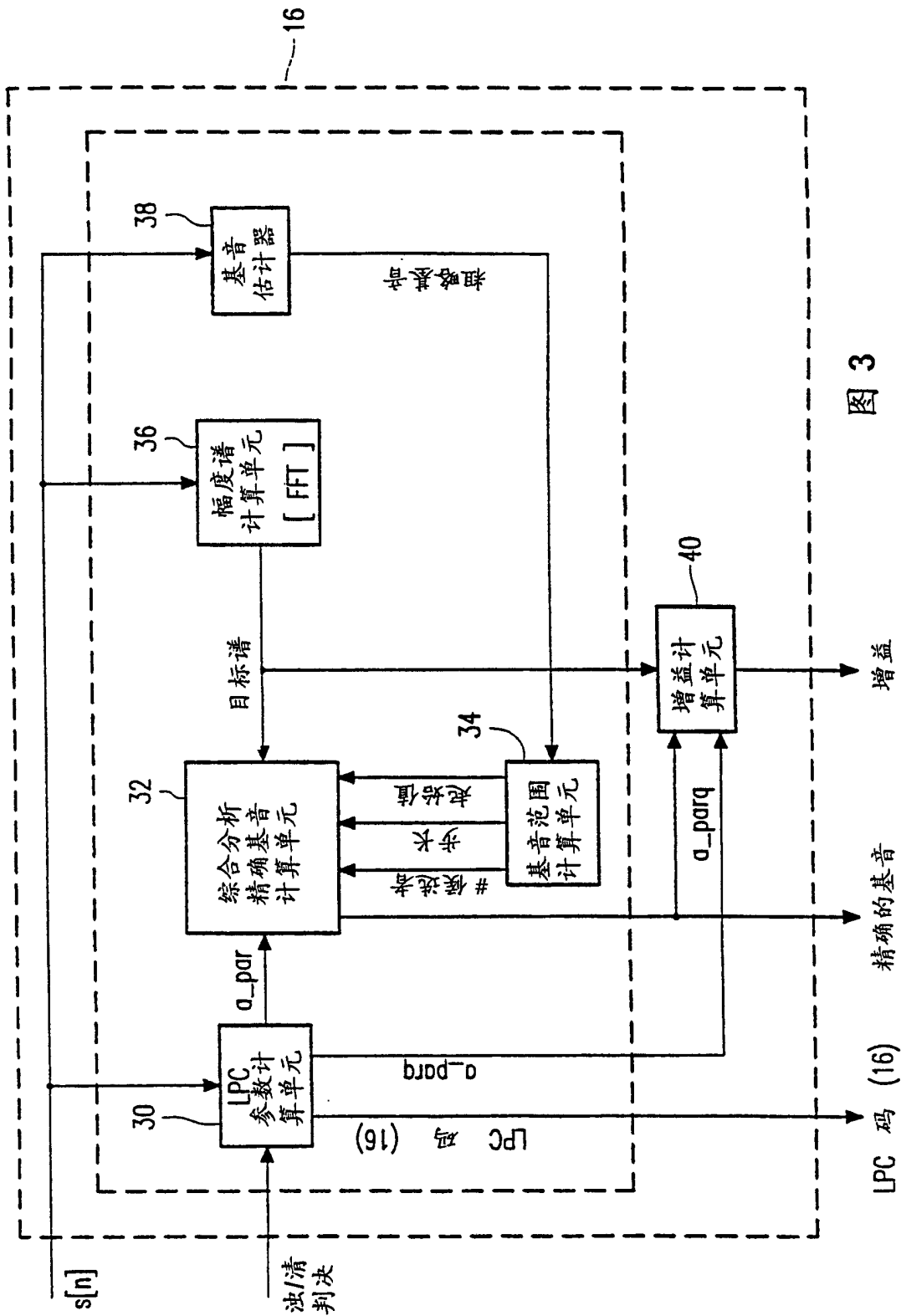
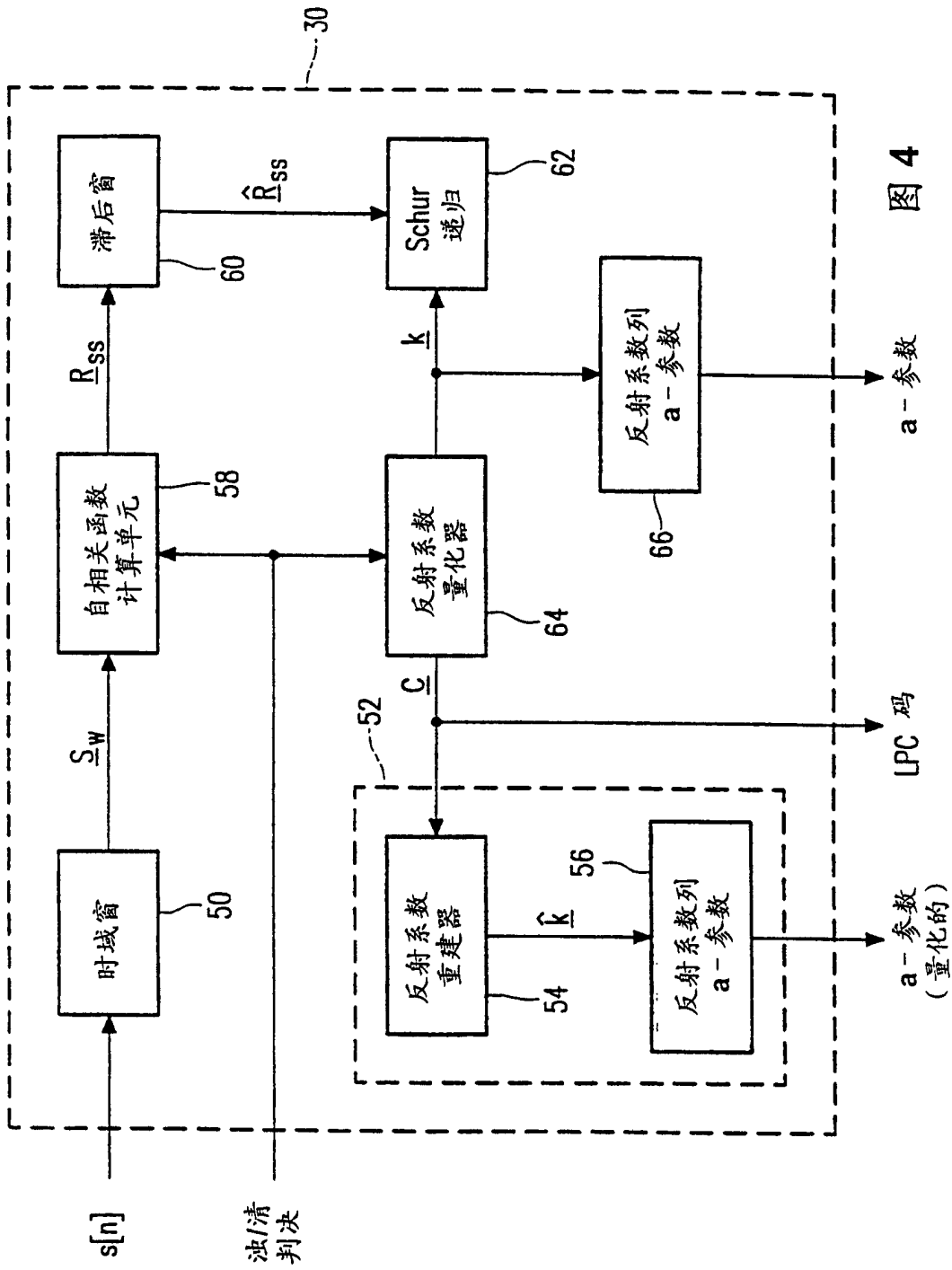


图 3



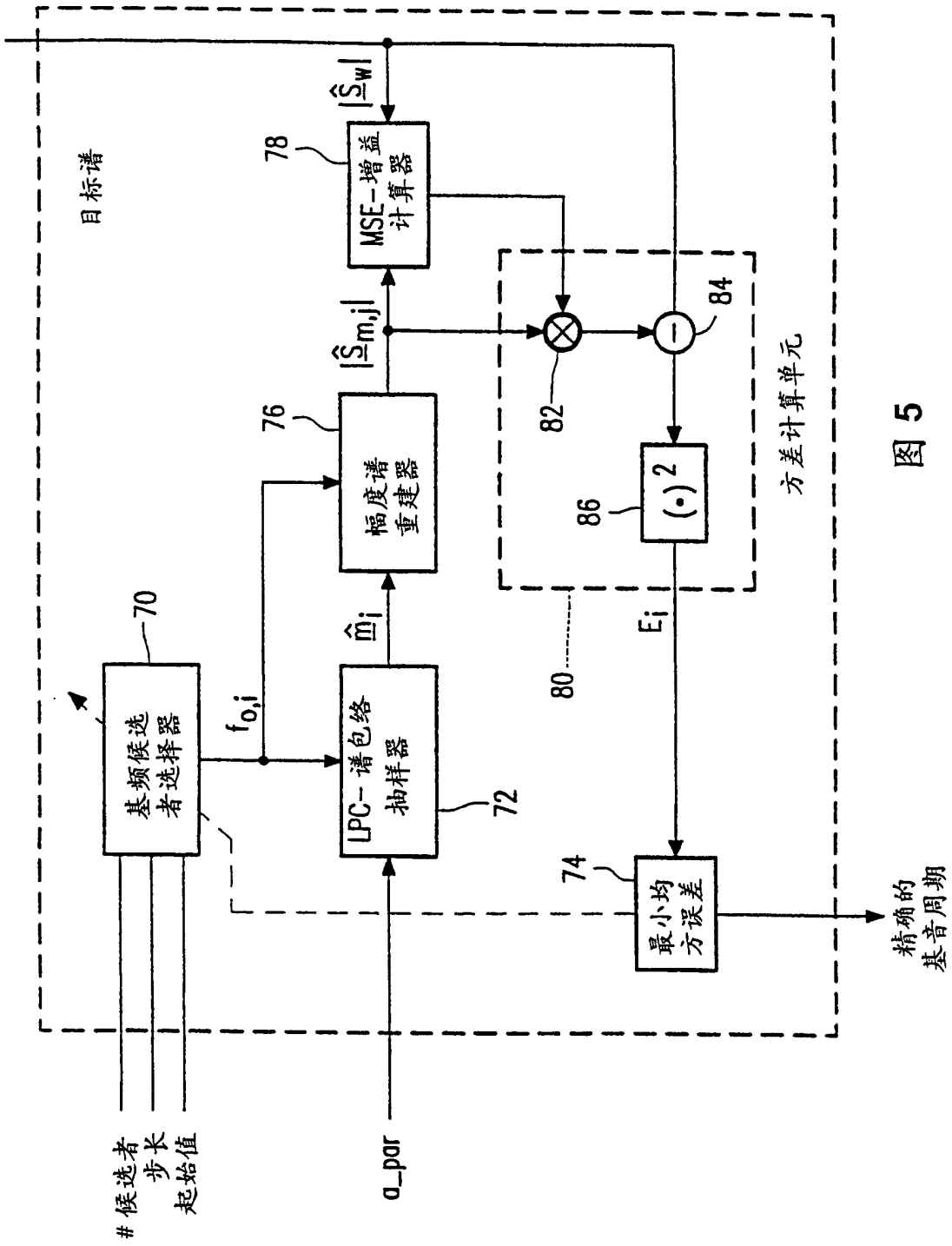


图 5

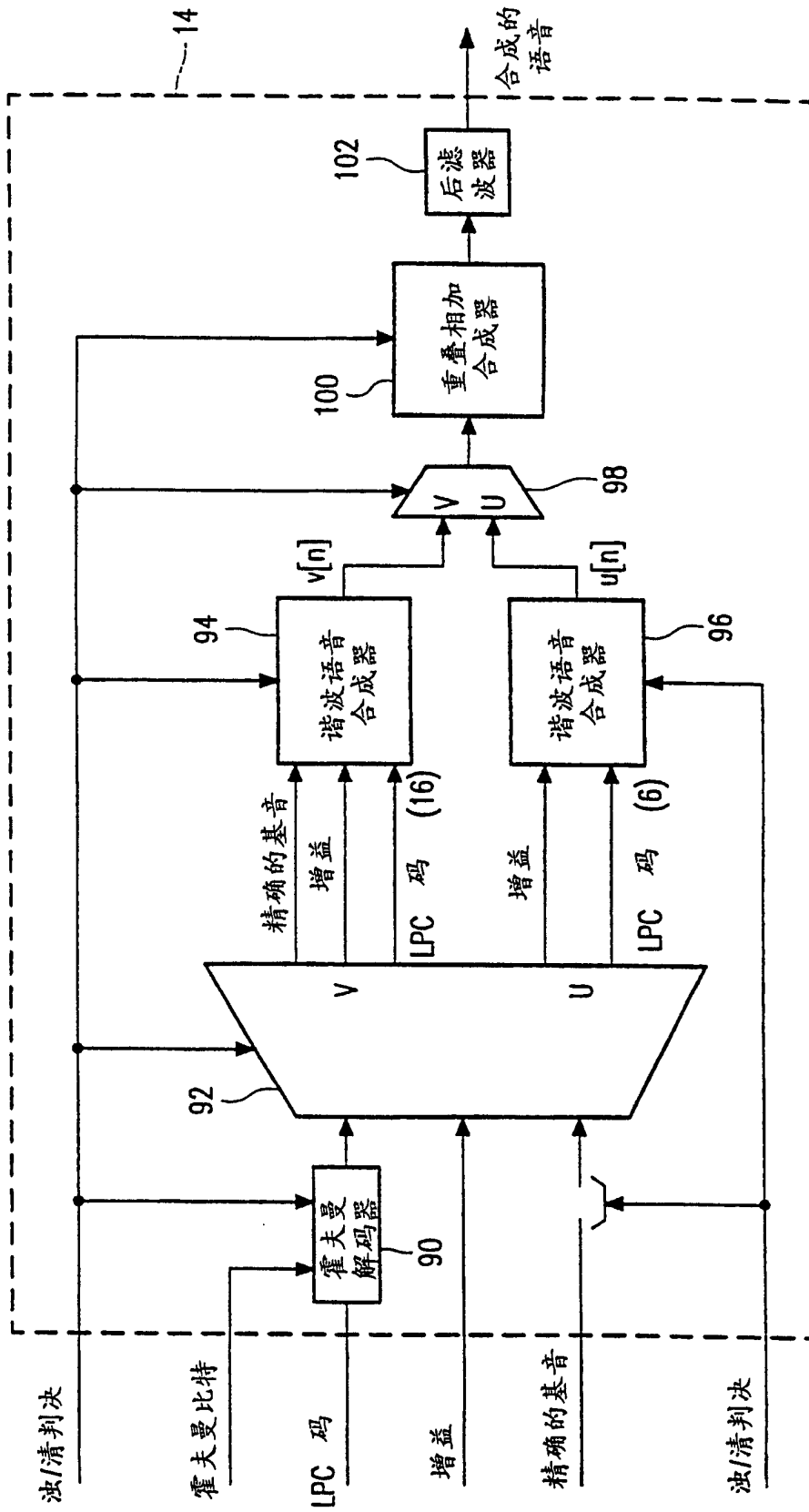


图 7

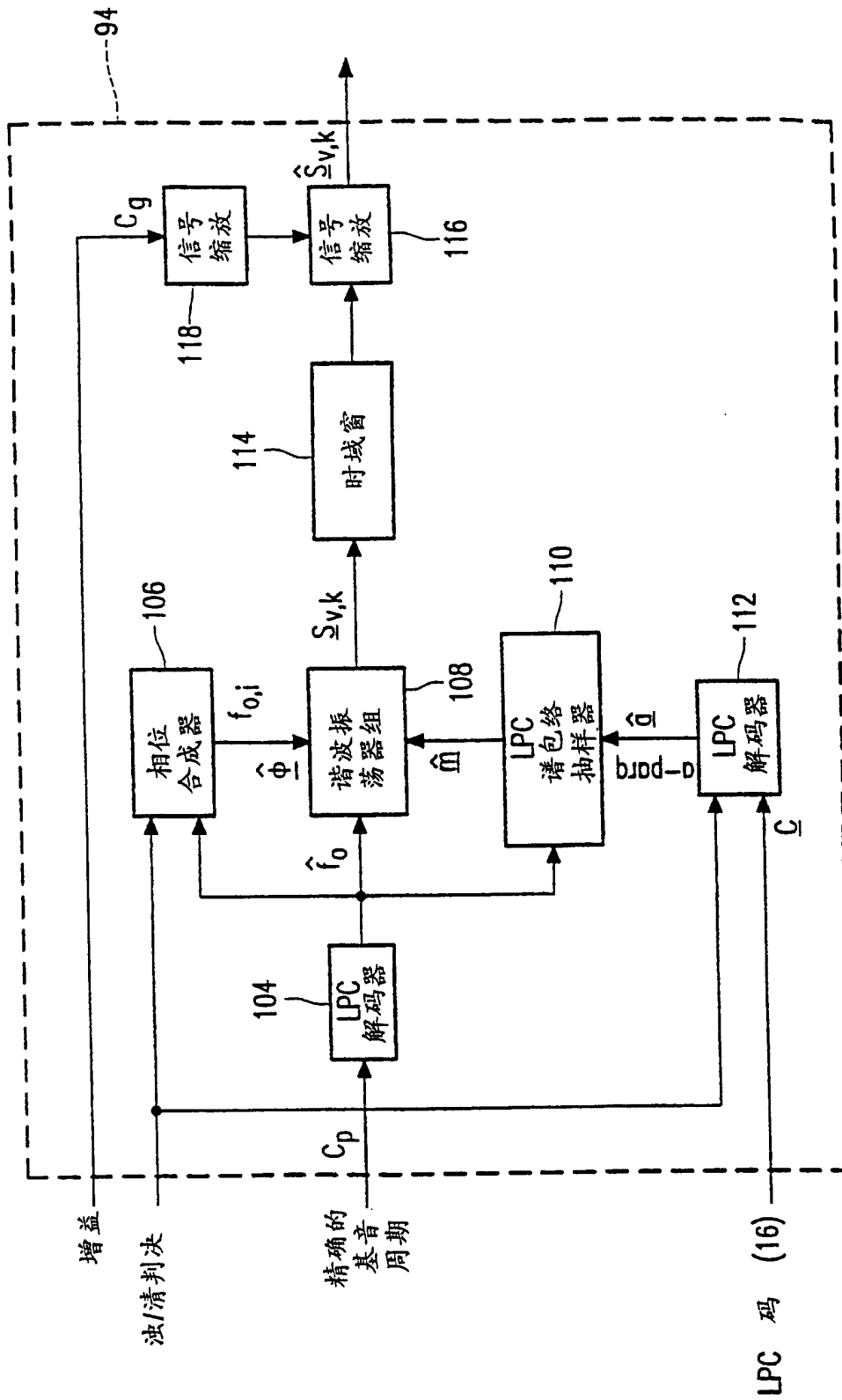


图 8

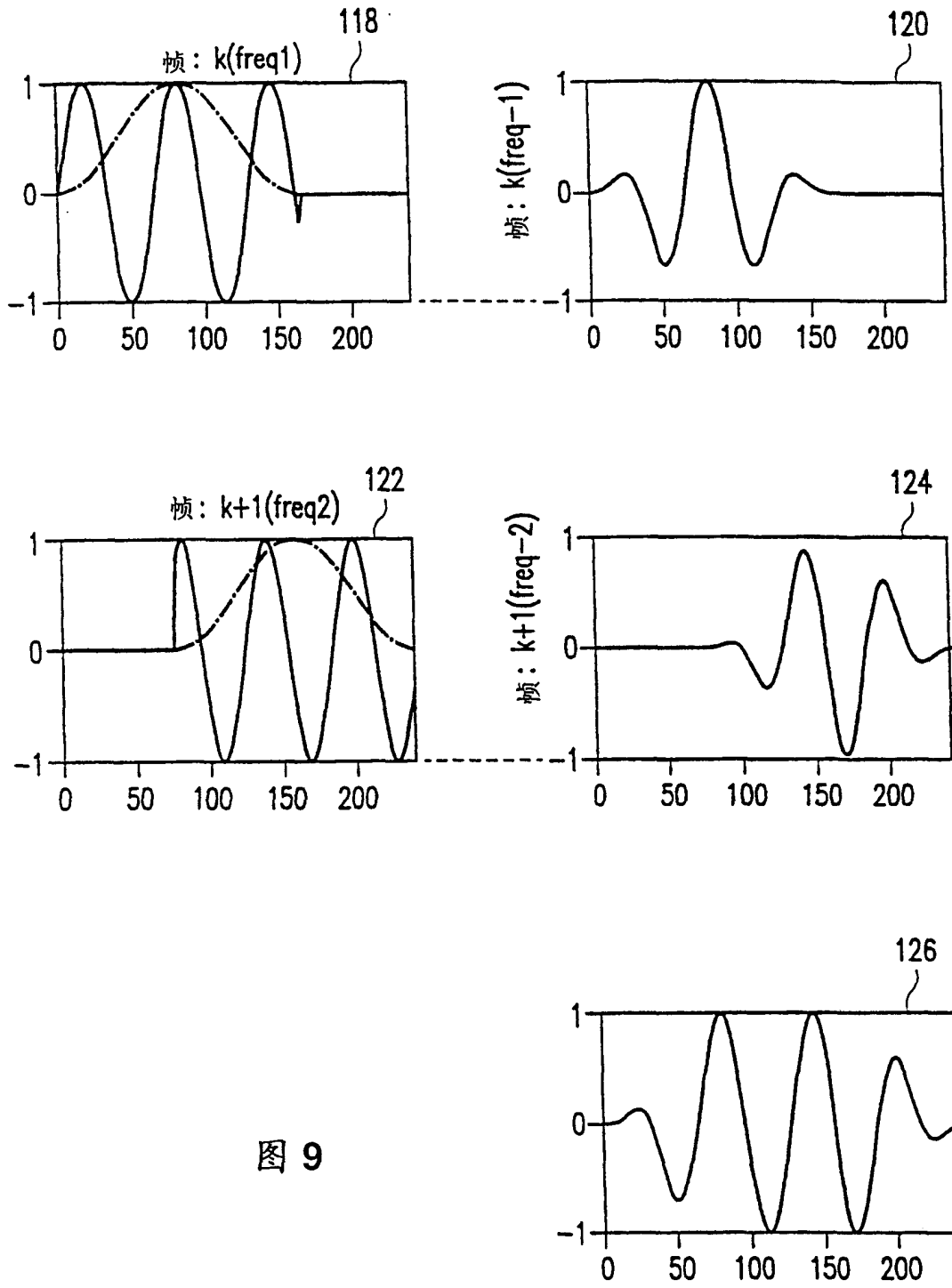


图 9

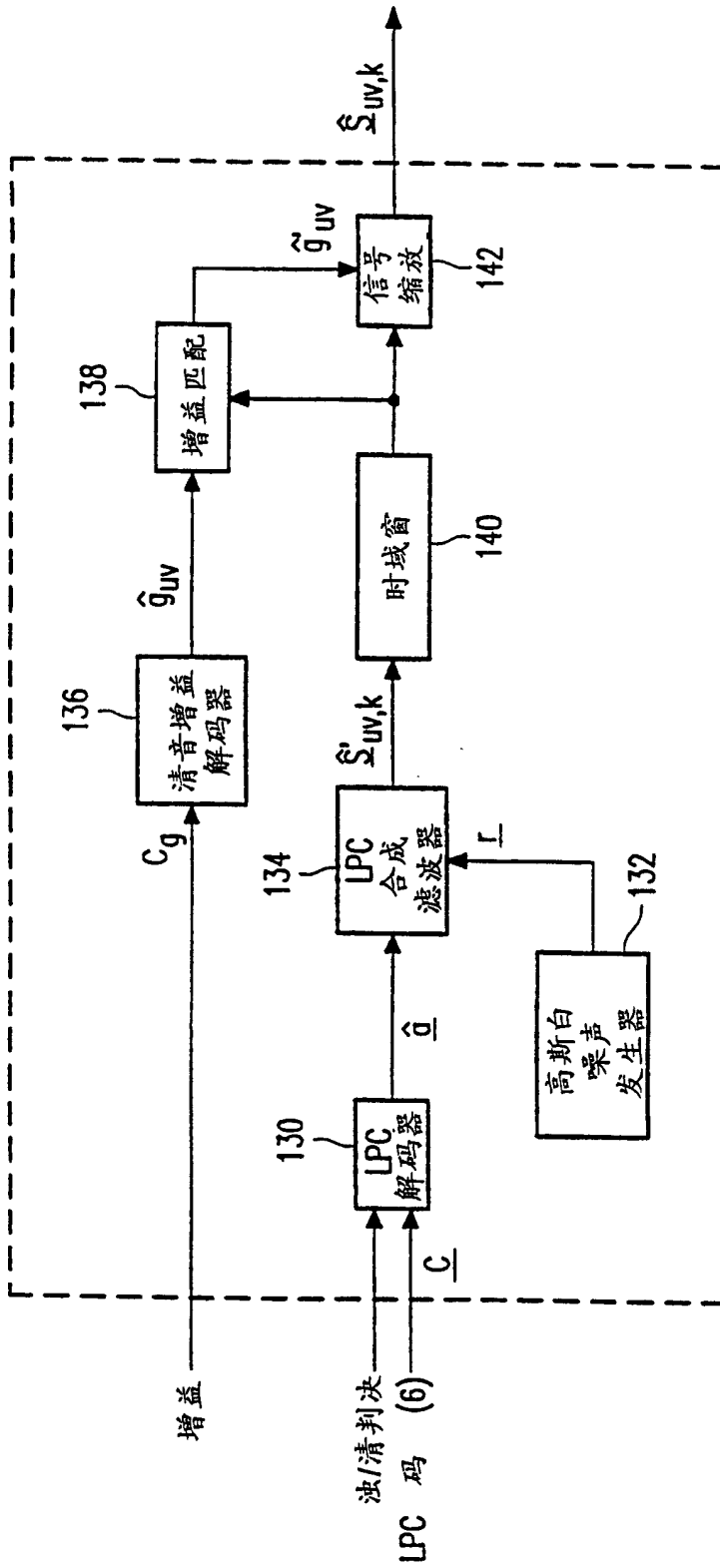


图 10