

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

G10L 19/02

G10L 21/02



# [12] 发明专利说明书

专利号 ZL 02812978.4

[45] 授权公告日 2005 年 10 月 19 日

[11] 授权公告号 CN 1223991C

[22] 申请日 2002.5.15 [21] 申请号 02812978.4

[30] 优先权

[32] 2001.5.15 [33] FR [31] 01/06412

[86] 国际申请 PCT/FR2002/001640 2002.5.15

[87] 国际公布 WO2002/093558 法 2002.11.21

[85] 进入国家阶段日期 2003.12.26

[71] 专利权人 维夫康姆公司

地址 法国伊西莱穆利诺

[72] 发明人 弗兰克·比耶特里

胡伯特·卡迪茜尤

审查员 杨永康

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

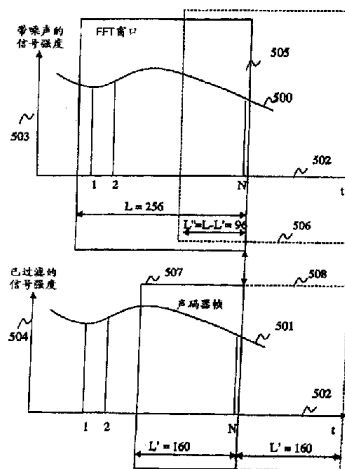
代理人 郭思宇

权利要求书 4 页 说明书 14 页 附图 7 页

[54] 发明名称 声频信号处理设备和方法

[57] 摘要

声频信号处理方法。本发明涉及声频信号处理方法，它包括：源声频信号的第一处理步骤(205)，所述步骤对第一取样序列实施至少一次数学变换，所述第一取样序列是第一分割窗口(505, 506, 700, 701)应用在所述源声频信号上而获得的；及第二声音处理步骤(206)，所述步骤应用在第二取样序列上，所述取样序列是第二分割窗口(507, 508)应用在所述第一步骤时发出的信号上而获得，所述第二分割窗口不同于所述第一分割窗口；两连续第一窗口及/或两连续第二窗口交迭，所述交迭即表示各分割是同步的。



1、 音频信号处理方法，它包括：

——源音频信号的第一处理步骤（205），所述第一处理步骤对第一取样序列实施至少一次数学变换，所述第一取样序列是对所述源音频信号应用第一分割窗口（505，506，700，701）而获得的；及

——第二音频处理步骤（206），所述第二处理步骤应用在第二取样序列上，所述第二取样序列是将第二分割窗口（507，508）应用在所述第一步骤得到的信号上而获得的，所述第二分割窗口不同于所述第一分割窗口；

其特征在于，相继的两第一窗口及/或相继的两第二窗口交迭，所述交迭使所述分割同步。

2、 根据权利要求1所述的方法，其特征在于，所述第二分割窗口是相继的帧。

3、 根据权利要求1所述的方法，其特征在于，第一序列的最后一个取样在经过第一步骤后也是相应的第二序列的最后一个取样。

4、 根据权利要求2所述的方法，其特征在于，第一序列的最后一个取样在经过第一步骤后也是相应的第二序列的最后一个取样。

5、 根据权利要求1至4中任一项所述的方法，其特征在于，每个所述第一分割窗口（700，701）是完全重建窗口，它是通过卷积以下窗口获得的：

——第一完全重建中间窗口，该窗口具有适合数学变换的频谱特性；及

——第二矩形中间窗口。

6、 根据权利要求1至4中任一项所述的方法，其特征在于，应用于每个第一序列的所述第一处理步骤还包括：

——应用于所述第一序列的预定处理子步骤（303）；

——应用于所述第一序列的已处理取样的逆数学变换子步骤（304）；

——相加步骤(305)，相加来自应用于所述第一序列的所述逆数学变换子步骤的声音取样和来自应用于前一个第一序列的所述逆数学变换子步骤的相应的声音取样。

7、根据权利要求5所述的方法，其特征在于，应用于每个第一序列的所述第一处理步骤还包括：

——应用于所述第一序列的预定处理子步骤(303)；

——应用于所述第一序列的已处理取样的逆数学变换子步骤(304)；

——相加步骤(305)，相加来自应用于所述第一序列的所述逆数学变换子步骤的声音取样和来自应用于前一个第一序列的所述逆数学变换子步骤的相应的声音取样。

8、根据权利要求6所述的方法，其特征在于，所述预定处理子步骤包括降低或消除所述声频信号内的噪声。

9、根据权利要求7所述的方法，其特征在于，所述预定处理子步骤包括降低或消除所述声频信号内的噪声。

10、根据权利要求6所述的方法，其特征在于，所述预定处理子步骤包括至少一种选自下列处理的处理：

——降低或消除声频信号中的回声；

——对声频信号进行声音识别。

11、根据权利要求7所述的方法，其特征在于，所述预定处理子步骤包括至少一种选自下列处理的处理：

——降低或消除声频信号中的回声；

——对声频信号进行声音识别。

12、根据权利要求8所述的方法，其特征在于，所述预定处理子步骤包括至少一种选自下列处理的处理：

——降低或消除声频信号中的回声；

——对声频信号进行声音识别。

13、根据权利要求9所述的方法，其特征在于，所述预定处理子步骤包括至少一种选自下列处理的处理：

——降低或消除声频信号中的回声；

——对声频信号进行声音识别。

14、根据权利要求 1 至 4 中任一项所述的方法，其特征在于，所述一种或多种数学变换选自下列变换：

——快速傅里叶变换（FFT）及其变型；

——快速哈达马特变换（FHT）及其变型；及

——离散余弦变换（DCT）及其变型。

15、根据权利要求 12 所述的方法，其特征在于，所述一种或多种数学变换选自下列变换：

——快速傅里叶变换（FFT）及其变型；

——快速哈达马特变换（FHT）及其变型；及

——离散余弦变换（DCT）及其变型。

16、根据权利要求 13 所述的方法，其特征在于，所述一种或多种数学变换选自下列变换：

——快速傅里叶变换（FFT）及其变型；

——快速哈达马特变换（FHT）及其变型；及

——离散余弦变换（DCT）及其变型。

17、根据权利要求 1 至 4 中任一项所述的方法，其特征在于，所述源声频信号为一声音信号。

18、根据权利要求 15 所述的方法，其特征在于，所述源声频信号为一声音信号。

19、根据权利要求 16 所述的方法，其特征在于，所述源声频信号为一声音信号。

20、声频信号处理装置，它包括：

——源声频信号的第一处理装置，它对第一取样序列实施至少一个数学变换，所述第一取样序列是将第一分割窗口应用在所述源声频信号上而获得的；及

——第二声频处理装置，它应用在第二取样序列上，所述第二取样序列是将第二分割窗口应用在该第一步骤得到的信号上而获得，

所述第二分割窗口不同于第一分割窗口；

其特征在于，相继的两第一窗口及/或相继的两第二窗口相互交迭，所述交迭使所述分割同步。

## 音频信号处理设备和方法

### 技术领域

本发明涉及音频信号处理领域。

更准确地说，本发明尤其涉及降低或消除由数字通讯设备例如数字电话及/或免提型移动无线电话处理的音频信号中夹杂的噪声。

### 背景技术

当声音数字通讯设备应用在噪声环境中（一般乘车时），所述噪声环境会极大干扰音频信号，从而降低通讯的质量。

根据已知技术，利用插入一噪声衰减器或抑制器——所述装置在对音频信号作特殊处理前，作用于麦克风捕捉到的信号——来解决这一问题。

根据第一已知技术，在捕捉音频信号的麦克风和音频信号处理装置之间插入一回声或噪声衰减或降低装置。所述装置提高了有用信号与噪声的比率，或减少回声，以使随后信号可在最佳条件下被处理。但所述现有技术必需一专门的特殊装置，这样既增加了成本，又加大了使用的复杂性。

根据第二已知技术，基于使用快速傅里叶变换（英文 FFT 即“Fast Fourier Transform”）的降低噪声功能集成在数字通讯装置中，所述 FFT 应用于连续的声音取样流。首先，取样流被格式窗分割成 256 个取样的窗口，窗口的一半交迭在一起（一窗口的第一 128 个取样对应于前一窗口的后 128 个取样）。FFT 应用于各窗口，FFT 的结果再通过噪声或回声降低或消除函数进行处理。

所述函数结果再由逆快速傅里叶变换（即 IFFT）进行处理，以复原声音取样流，使其能被声音处理函数处理。

所述现有技术的缺点在于实施相对比较复杂。

## 发明内容

根据其不同特征，本发明的目的尤其在于弥补现有技术的所述缺陷。

更准确地说，本发明的一目的在于，提供一种声音处理方法及装置，它可降低基于数学变换的处理的复杂性，所述数学变换应用在数据块上，同时使应用于声频帧的声音处理效果达到最佳。

本发明的另一目的在于，使基于数学变换的处理及声频处理的集成处理最佳。

本发明的目的还在于最佳化所述处理时间。

本发明的另一目的在于降低所述处理所必需的计算容量。

为此，本发明提出一种声频信号处理方法，它包括：

——源声频信号的第一处理步骤，所述第一处理步骤对第一取样序列实施至少一次数学变换，所述取样序列是第一分割窗口应用在源声频信号上而获得的；及

——第二声音处理步骤，所述第二处理步骤应用在第二取样序列上，所述取样序列是第二分割窗口应用在第一步骤时发出的信号上而获得的，所述第二分割窗口不同于第一分割窗口；

其特征在于，两连续第一窗口及/或两连续第二窗口交迭，所述交迭即表示分割是同步的。

因此，声频处理步骤可连续实施或在一多任务环境中实施。另外，所述实施由于使用了尺寸可预测、准确且经济的存储器而简化了。

根据一特别特征，所述方法的特征在于，第二分割窗口是连续帧。

因此，根据本发明，所述处理方法的处理时间最佳化了。

根据一特别特征，所述方法的特征在于，第一序列的最后一个取样，经过第一步骤后，也是相应的第二序列的最后一个取样。

因此，第二声频处理步骤最好在没有任何无效等待中实施，以优化

总的声音处理时间。

根据一特别特征，所述方法的特征在于，各第一分割窗口为完全重建窗口，它是通过卷积（convolution）以下窗口获得的：

——第一中间完全重建窗口，——所述窗口具有适合数学变换的频谱特性；及

——第二中间矩形窗口。

因此，第一分割窗口的交迭部分为完全重建型，这可使在相对简单的第一处理过程中，能重新组合信号。

另外，由于第一中间窗口适合数学变换（尤其可衰减相对强的窗口的旁瓣，而主瓣仍是平的），提高了相应处理质量。

此外，第二中间窗口为矩形状，相应取样处理简单而有效。

根据一特别特征，所述方法的特征在于，应用于各第一序列的第一处理步骤还包括：

——应用于第一序列的预定处理子步骤；

——应用于第一序列的已处理取样的逆数学变换子步骤；

——相加来自应用于第一序列的相反数学变换子步骤的声音取样和来自应用于前第一步骤的逆数学变换子步骤的相应的声音取样。

根据一特别特征，所述方法的特征在于，预定处理子步骤包括降低或消除声频信号中的噪声。

根据一特别特征，所述方法的特征在于，预定处理子步骤包括至少一种处理，所述处理属于一组，所述组包括：

——降低或消除声频信号内的回声；

——识别声频信号内的声音。

因此，有利地是，所述方法在一装置（如电话、个人计算机或遥控系统）中，把如降低及/或消除噪声及/或回声及/或声音识别结合在一起，这样既降低了复杂性，又优化了所述处理的功效及/或装置的高集成性（因而降低了成本，减少了能量消耗，这些尤其对使用电池作动力的通信设备来说，相对更高）。

根据一特殊特征，所述方法的特征在于，所述数学变换属于这

样一组，所述组包括：

- 快速傅里叶变换（FFT）及其变型；
- 快速哈达马特变换（FHT）及其变型；及
- 离散余弦变换（DCT）及其变型。

因此，有利地是，本发明可使用一种或多种适合第一声频处理的数学变换，所述变换可应用在和第二分割窗口尺寸不同的尺寸的数据块上。

根据一特殊特征，所述方法的特征在于，源声频信号为一声频信号。

因此，本发明还适用于第二声频处理，当它专门针对语音如，语音编码（“声码处理”）及/或语音压缩，以进行存储及/或远距离传输时。

本发明还涉及一声频信号处理装置，它包括：

——源声频信号的第一处理装置，它对第一取样序列实施至少一次数学变换，所述第一取样序列是第一分割窗口应用在源声频信号上而获得的；及

——第二声频处理装置，它应用在第二取样序列上，所述第二取样序列是由第二分割窗口应用在第一步骤时发出的信号上而获得的，所述第二分割窗口不同于第一分割窗口；

其特征在于，两连续第一窗口及/或两连续第二窗口交迭，所述交迭即表示分割是同步的。

另外，本发明还涉及一计算机程序产品，所述程序产品包括程序元件，这些元件记录在由至少一微处理器可读的介质上，其特征在于，程序元件控制一个或多个微处理器，以使其执行：

——源声频信号的第一处理步骤，对第一取样序列实施至少一次数学变换，所述第一取样序列是第一分割窗口应用在声频源信号上而获得的；及

——第二声频处理步骤，它应用在第二取样序列上，所述第二取样序列是第二分割窗口应用在第一步骤时发出的信号上而获得，所

述第二分割窗口不同于第一分割窗口；

两连续第一窗口及/或连续第二窗口交迭，所述交迭即表示分段是同步的。

另外，本发明还涉及一计算机程序产品，其特征在于，所述程序包括指令系列，当所述程序在计算机上执行时，所述系列指令可实施声频处理方法。

声频信号处理装置、计算机程序产品的优点与声频信号处理方法的优点相同，这里就不再详述了。

### 附图说明

本发明的其它特征和优点将在后文参照附图、以非限制性方式描述一最佳实施例中体现出来。附图中：

——图 1 示出了根据一特别实施例的符合本发明的一无线电话机的总框架图；

——图 2 示出了图 1 中的无线电话机对一声频信号实施的连续处理；

——图 3 示出了根据图 2 的降低或消除噪声算法；

——图 4 示出了根据图 2 的声音处理应用于一帧；

——图 5 描述了如图 3、4 的处理所实施的取样流窗口化；

——图 6 示出了一已知成型窗口；

——图 7 示出了一成型窗口，所述窗口最佳化，并运用在根据本发明一最佳实施例的图 3 的窗口化操作中；及

——图 8 详细描述了图 3 中所示的一降低噪声型处理。

### 具体实施方式

本发明的基本原理即以下处理的同步：

——基于 FFT 的处理，尤其消除或降低噪声处理；及

——语音编码型声音处理。

事实上，FFT（快速傅里叶变换算法）与 IFFT 处理窗口包括有

2 的整数次方个取样（一般为 128 或 256）。

相反地，语音编码考虑到了大小不同的窗口（通常，GSM 领域内的声音处理考虑 160 取样窗口）。

例如，若为符合 ETSI（“European Telecommunication Standard Institute”）公布的 GSM 标准的无线电话，声音信号在以每帧 20ms 的压缩形式传输给用户前，在 8 千赫兹频率被采样。

可看出，根据 GSM 标准，语音编码由一声码器在 160 取样帧上实施。所述编码是所需速率的函数，下列文件尤其专门描述过它：

——《Full Rate (FR) speech transcoding》(GSM06.10)（《全速率语音代码转换》）

——《Half Rate (HR) speech transcoding》(GSM06.20)（《半速率语音代码转换》）

——《Enhanced Full Rate (EFR) speech transcoding》(GSM06.60)（《增强的全速率语音代码转换》）；及

——《Adaptive Multi-Rate (AMR) speech transcoding》(GSM 06.90)（《自适应多速率语音代码转换》）

根据目前技术现状，通过考虑一已作声音处理的 160 取样窗口，噪声及/或回声降低或消除装置处理一 256 长度的窗口，所述窗口可分段成长度 160 的三个窗口。其中，现有技术固有的异步性使得这些处理变得复杂，还必需大尺寸存储器、计算容量及/或 DSP（用于计算的 *Processeur de Traitement de Signal*，数字信号处理器）时钟。

根据本发明，可系统地使一噪声及/或回声消除或降低窗口的末端与声音处理帧、最好与声音处理帧的末端重合，以使两类处理同步。因此，如果噪声降低或消除窗口的尺寸等于 256 取样，如果声音处理帧的尺寸等于 160 取样，则回声降低或消除窗口会包含一完整声音处理帧及前一窗口的 96 取样（即 256 减去 160）。

因此，可保持噪声降低或消除窗口与声音处理帧之间的同步性，最佳化总处理时间。

根据本发明，成型窗口（适合于有关 160 点的取样帧及 256 点的 FFT）最好：

——为完全重建型，即两窗口重合的振幅和等于 1（在重叠部分上）；

——长度 256 的窗口各侧有 96 的长度重合。

例如，这种窗口可通过把宽度 97 的 Hanning 窗口（记作 Hanning(97)）与长度 160 的矩形窗口（记作 Rect(160)）卷积（convolution）而获得。

因此，一 256 点 FFT 可应用在与 160 取样帧上同步的各 256 取样窗口。FFT 的实施已为本领域技术人员熟知，尤其在 Press W.H.、Teukolsky S.A.、Vetterling W.T.及 Flannery B.P.撰写的书籍《Numerical Recipes in C, 2<sup>nd</sup> edition》（《C 语言数字方法，第二版》）（1992 年由剑桥大学出版社出版）中详细讲述过。

在实施逆变换（记作 IFFT）前，可把本领域技术人员已知的各种降低噪声算法运用在所述 256 取样块上。

因此，256 取样块连续被处理。IFFT 后，当前窗口已处理的前 96 个取样再添加在前一窗口已处理的后 96 个取样中。添加后，当前窗口的前 160 个取样被传输给声码器，以根据本领域人员已知的语音编码方法、必要时符合已实施标准，进行处理。

图 1 示出了实施本发明的一无线电话。

图 1 简略示出了根据一最佳实施例的符合本发明的无线电话的总框架图。

无线电话 100 包括以下各元件，所述各元件之间通过一地址及数据总线 103 连接起来：

——麦克风 107；

——模拟/数字转换器 108；

——扬声器 109；

——数字/模拟转换器 110；

——信号处理器（DSP）104；

- 非易失存储器 105;
- 随机存储器 106;
- 无线接口 111;
- 数据帧交换和协议的管理及控制单元 112; 及
- 人/机接口 (一般为一键盘和一屏幕) 113.

图 1 所示各元件为本领域技术人员已知。所以此处就不再描述这些共同元件了。

还可看出, 整个说明书所使用的词《寄存器》在所有所述存储器中, 既指一低容量 (几个二进制数据) 存储区, 也指一高容量存储区 (可储存整个程序或所有事务处理数据序列)。

非易失存储器 105 (即 ROM) 在寄存器中, 出于方便考虑, 仍使用和它们保存的数据相同的名, 它们保存:

- 在《prog》寄存器 308 中, DSP 104 的运行程序;
  - 值 L (一般为 256), 它表示第一分割窗口尺寸, 所述窗口尺寸与寄存器 115 中 FFT 所考虑的点数相对应;
  - 值 L' (一般为 160), 它表示第二窗口尺寸, 所述窗口尺寸和寄存器 115 中声码器处理的帧的尺寸相对应; 及
  - 值  $\alpha$ 、 $\beta$ 、 $\gamma$ 、 $\kappa$  与  $\beta_r$ , 它们表示信号中噪声的降低。
- 随机存储器 106 保存数据、变量及中间处理结果, 尤其包括:
- 寄存器 117, 其中保存着已接收信号的噪声取样值;
  - 寄存器 118, 其中保存着已处理取样值; 及
  - 用于声码器的已处理取样序列。

DSP 尤其适合傅里叶变换与语音编码型处理。例如, 可使用由《DSP GROUP》(注册商标)公司生产的 DSP 芯, 以《OAK》(注册商标)标识。

图 2 示出了图 1 的无线电话对声音信号实施的连续处理。

可看出, 进入麦克风 107 的信号为下列之和 203:

- 一声音信号, 所述声音信号可能受到回声干扰 (以发生信号 200 与发生延迟信号之和标记); 及

### ——一噪声 202

麦克风 107 捕捉到的带噪声的信号被传送给模拟/数字转换器 204, 此处, 步骤 204 中, 它将其变换成数字取样序列。根据 GSM 标准, 可看出, 采样一般在 8 千赫兹的频率上。

步骤 205 中, 处理数字取样序列。

步骤 206 中, 已处理取样帧  $L'$  (160) 由声码器根据一已知方法 (一般在 GSM 标准中规定) 进行编码。

步骤 207 中, 已“声码处理的”帧通过单元 112 格式化 (成型), 再由无线模块 111 根据已知技术 (例如根据 GSM 标准) 传输。

图 3 示出了图 2 中处理步骤 205 中实施的噪声消除或降低算法。

启动步骤 300 中, DSP 104 在 RAM 106 中, 初始化第一 96 取样块为 0, 它对应于所接收的最后取样及处理 205 良好运行所必需的变量。

步骤 301 中, DSP 104 在前面接收的取样后, 把来自转换器 108 的 160 个输入取样序列存储在 RAM 106 中。

步骤 302 中, DSP 104 把长度 256 的分段窗口应用于已接收的最后 256 取样形成的序列。(可看到, 所述窗口将在后文参照图 7 进行描述)

于是, 256 点的 FFT 型数学变换, 可运用在通过应用分段窗口而获得的序列中。

步骤 303 中, 噪声降低型处理 (后文将参照图 8 进行描述) 应用于数学变换产生的序列。

步骤 304 中, 步骤 302 中变换的逆 IFFT 型变换应用在已处理序列中。

步骤 305 中, 必要时 (即第一次重复后), DSP 104 把前面已处理的序列的后 96 个取样添加在当前序列的已处理的前 96 个取样中。

步骤 306 中，当前已处理前 160 个取样形成的序列或帧传输给声码器。

步骤 307 中，把和步骤 305 中被传输 160 个取样相对应的已接收 160 个取样从存储器 106 中抹去。

重复步骤 301。

图 4 示出了图 2 的步骤 206 所实施的语音编码。

初始化步骤 400 中，DSP 104 在 RAM 106 中初始化编码 206 良好运行所必需的所有变量。

步骤 401 中，DSP 104 把步骤 307 中已传输的 106 取样帧存储在 RAM 106 中。

步骤 402 中，DSP 104 根据一已知技术，把语音编码处理应用在 160 个取样帧中。

步骤 403 中，已编码帧被格式化，传输给单元 102，以发送给收信方。

步骤 404 中，160 取样帧从存储器 RAM 106 中抹去。

重复步骤 401。

图 5 描述了如图 3、4 的处理所实施的取样序列的窗口化。

第一图描画出了作为时间  $t$  502 的函数的，由转换器 108 直接接收到的信号强度 503 的曲线 500。

第二图描画出了作为时间  $t$  502 的函数的，步骤 205 中已处理的信号强度 504 的曲线 500。

从第一图中可看出，时间被分割成长度  $L$  为 256 的两连续窗口 505、506，所述两窗口在长度为 96 的  $L''$  上交迭且在步骤 302 中获得。

从第二图中还可看出，时间被分割成长度为 160 的  $L'$  的两连续帧 507、508，它们没有交迭，且在传输步骤 306 中获得。

信号分段是这样的：窗口 505（或 506）与 507（或 508）完全同步。

因此，根据最佳实施方式，窗口 505（或 506）与 507（或

502) 在处理前后 (根据步骤 303、304 及 305) 在同一取样上结束。

采用这种方式, 在长度  $L''$  上是交迭的。

图 6 示出了一已知成型窗口。

图中给出作为取样序列 601 的函数的长度为 256 且重叠 128 的 Hanning 窗口 603、604。

可看出, 根据所述已知分段, 窗口无论如何都不可能 与 160 个取样帧分段同步。

图 7 示出了根据本发明已优化的已成型窗口 700、701 (分别对应图 5 的窗口 505、506, 但更准确)。

和上述的一样, 图中给出作为取样 601 函数的窗口振幅 602。

可看出, 窗口 700、701 为 Hanning 窗口, 所述窗口是通过把长度 97 的中间 Hanning 窗口与长度 160 的矩形窗口卷积而形成。因此, 通过 160 个取样的窗口的连续偏移, 可获得完全重建窗口。

图 8 详细描述了图 3 中所示的噪声降低型处理步骤 303。

所述噪声降低处理尤其在下列文件中描述过:

——《Spectral substration based on minimum statistics》(《基于最小统计的频谱扣减》), 作者 R. Martin, 发表在《Signal Processing VII: Theories and applications, 1994, EURASIP》, 1182 至 1185 页;

——《Computationally efficient speech enhancement by spectral minima tracking in subbands》(《通过研究分波段里的频谱最小值, 提高用于计算的有效声音》), 作者 G. DOBLINGER, 发表在大会《ESCA. EUROPSPEECH'95, 4<sup>th</sup> European Conference on speech communication and technology》的报告 (第 1513 至 1516 页); 及

——《A combination of noise reduction and improved echo cancellation》(《改善的噪声降低及回声消除法》), 由 Darmstadt 技术大学以德语发表在《Fachgebiet Theorie der Signale》上。

经步骤 302 处理后, 包括 256 频谱分量的帧 801——对应一带噪

声音信号——可根据后面所述的处理 303 进行处理。

用  $X_k(m)$  表示第  $m$  个带噪声声音信号帧的第  $k$  个分量。

步骤 802 中，DSP 104 把直角坐标帧 801 的分量转换为极坐标的，以把相位与频谱振幅分离开。

在不同处理过程中，只有频谱振幅会改变，而相位保持不变。

在步骤 803，首先根据下列关系式估计短期的信号功率  $P_{xk}(m)$ ：

$P_{xk}(1) = (1-\alpha)|X_k(1)|^2$ （其中，可能加入一校正值，以提高估计的收敛速度）；

$$P_{xk}(m) = \alpha P_{xk}(m-1) + (1-\alpha)|X_k(m)|^2 \quad \text{其中 } m > 1$$

而“遗忘”系数  $\alpha$  的值介于 0.7 至 0.9 之间，这可确保寻找到合适的短期固定语音的频谱。

这些关系式尤其体现出两大好处：

- 其计算简单；及
- 不必引入任何测量延时。

根据一实施变型，可使用一经改善的噪声降低算法。但在所述算法中引入了一附加延时，这就要求存储器的尺寸更大，以存储复数值的频谱分量。

再根据以下非线性估计器（某种程度上它在寻找  $P_{xk}(m)$  的临时最小值），估计噪声的频谱功率  $P_{nk}(m)$ ：

$$P_{nk}(1) = P_{xk}(1)$$

其中  $m$  严格大于 1 ( $m > 1$ )：

如果  $P_{nk}(m-1) < P_{xk}(m)$

$$\text{则 } P_{nk}(m) = \gamma P_{nk}(m-1) + \{(1-\gamma)/(1-\beta)\} (P_{xk}(m) - \beta P_{xk}(m-1)) ;$$

否则， $P_{nk}(m) = P_{xk}(m)$

随后步骤 806 中，DSP 104 根据下列关系式计算实数值的增益因数  $g_k(m)$ ：

$$g_k(m) = 1 - \sqrt{\kappa P_{nk}(m) / P_{xk}(m)} \quad \text{如果 } g_k(m) > \beta_f$$

否则， $g_k(m) = \beta_f$

系数  $\kappa$  为一引入的噪声的高估因数，以使降低噪声算法性能更

佳。

$\beta_f$  为一最低频谱值。 $\beta_f$  把噪声降低过滤器的衰减限制在一正值，以使信号内的噪声最小。

步骤 807 中，DSP 104 把振幅  $|X_k(m)|$  乘以相应的增益因数  $g_k(m)$ ，以根据下列关系式获得改善后的信号振幅：

$$|Y_k(m)| = g_k(m) \cdot |X_k(m)|, \quad k \text{ 在 } 1 \text{ 至 } 256 \text{ 之间。}$$

在极坐标转换为直角坐标的步骤 808 中，DSP 104 根据步骤 807 中确定的振幅  $|Y_k(m)|$  与步骤 802 中提取的信号相位，构建噪声已降低的信号 809。

于是，信号 809 根据逆傅里叶变换步骤 304 被处理。

当然，本发明并不局限于上述实施例中。

尤其地，本领域技术人员可在本发明的应用中实施各种变型，并不只局限在移动电话（尤其是 GSM、UMTS、IS95 型……）中，还可扩展到任何类型的装置，所述装置包括在对输入声频信号进行数学变换前、后的一声频编码。

另外，本发明不仅适用于处理源声频信号，而且适用于任何类型的声频处理。

根据本发明，所实施的数学变换尤其为可应用在特殊长度的取样块上，所述特殊长度不等于根据声频处理所处理的帧的尺寸，或并不是接收所述帧尺寸的倍数或除数。因此，本发明可适合这种情况：声频帧的尺寸等于 160，或更广泛地讲，不是 2 的几次方，及数学变换适用于长度 256、128、512 或更广泛地  $2^n$ （此处， $n$  为一整数）的块的尺寸上，尤其是 FFT，FHT（Fast 哈达马特 Transform）或 DCT（离散余弦变换：Discrete Cosine Transform）或这些变换的变型（例如，通过把这些变换中的一种或几种与其它一种或几种变换结合起来）……

另外，本发明适用于任何与数学变换相关的处理，所述处理可在语音编码前、后实施，尤其当识别声音或消除及/或减小回声时。

需注意，本发明并不局限于单纯的设备安装，它还可以为计算

机程序的序列指令形式或综合部分硬件、部分软件的任何混合形式。  
当本发明部分或全部以软件形式安装时，相应的指令序列可存储在一可拆卸或不可拆卸的存储装置（如软盘、**CD-ROM** 或 **DVD-ROM**）中，所述存储装置可由计算机或微处理器部分地或完全地读取。



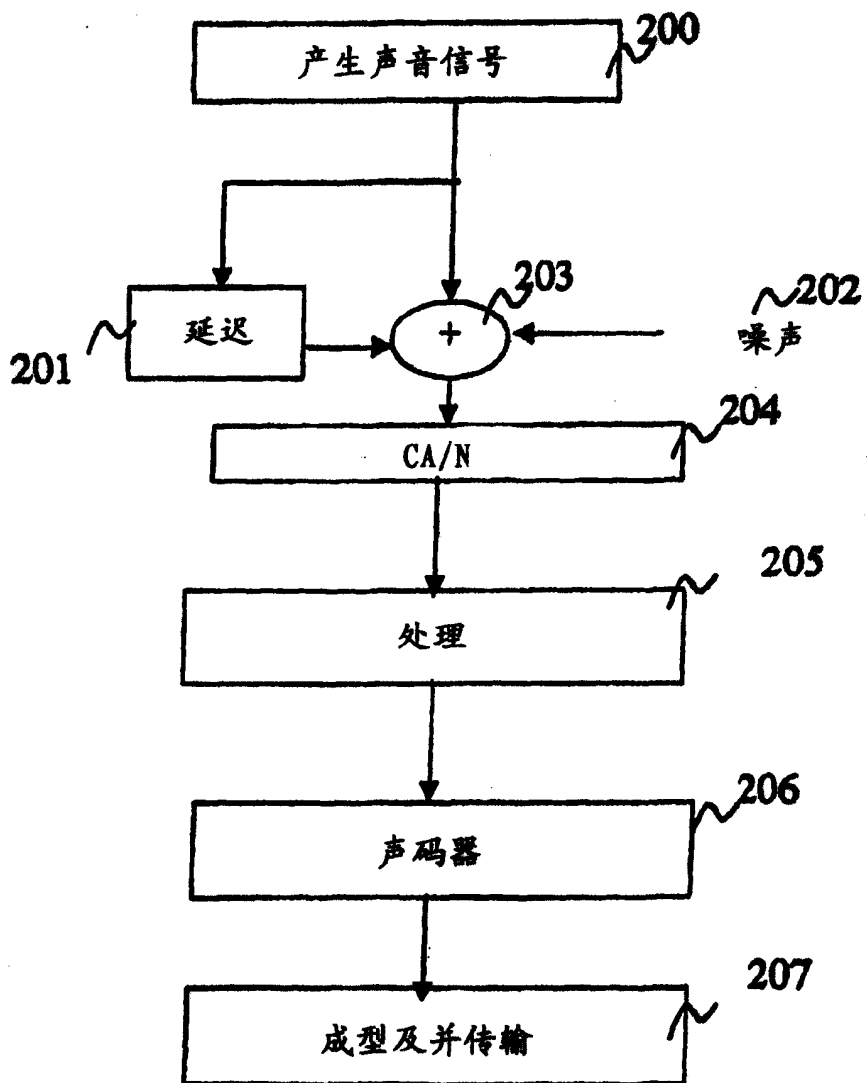


图2

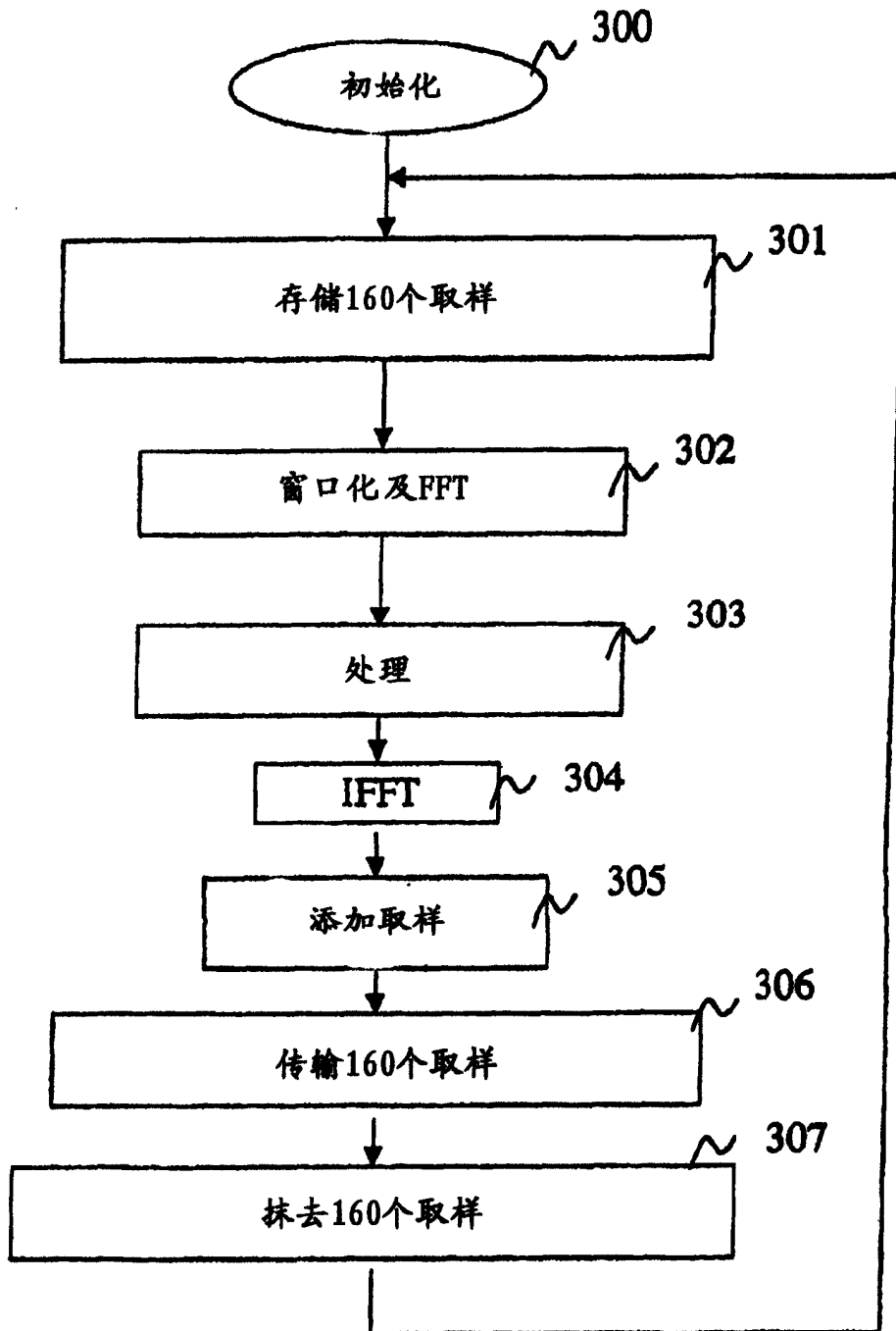


图3

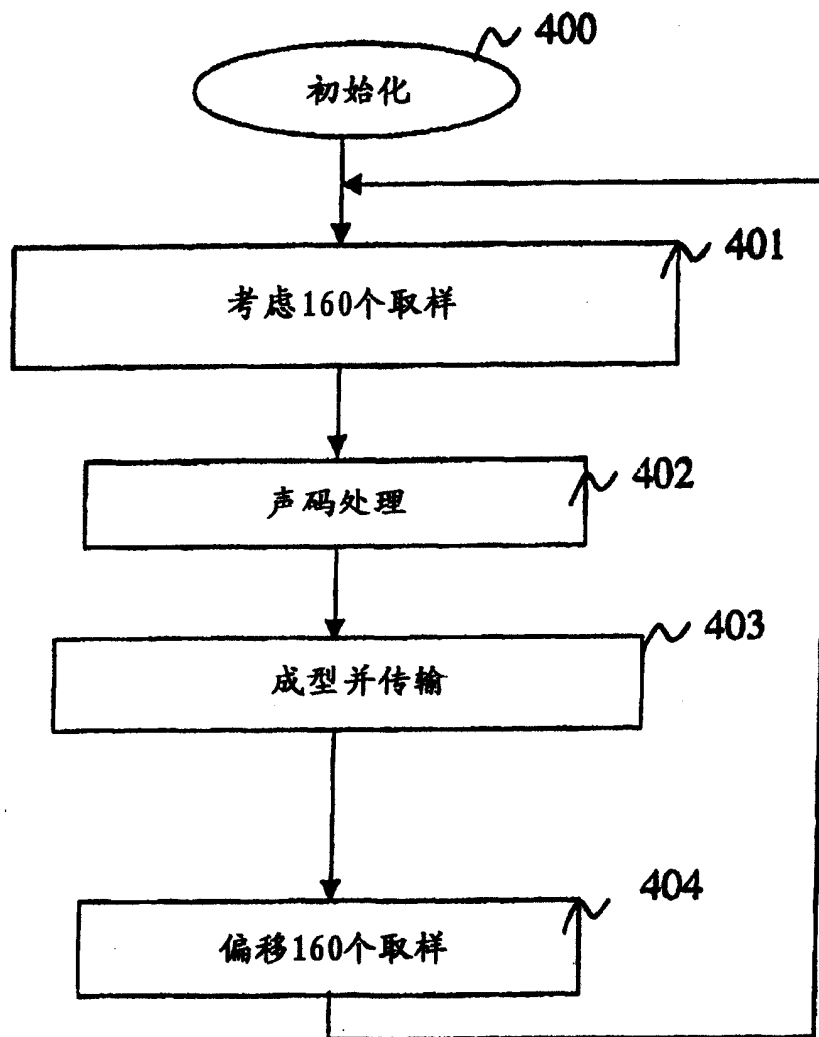


图4

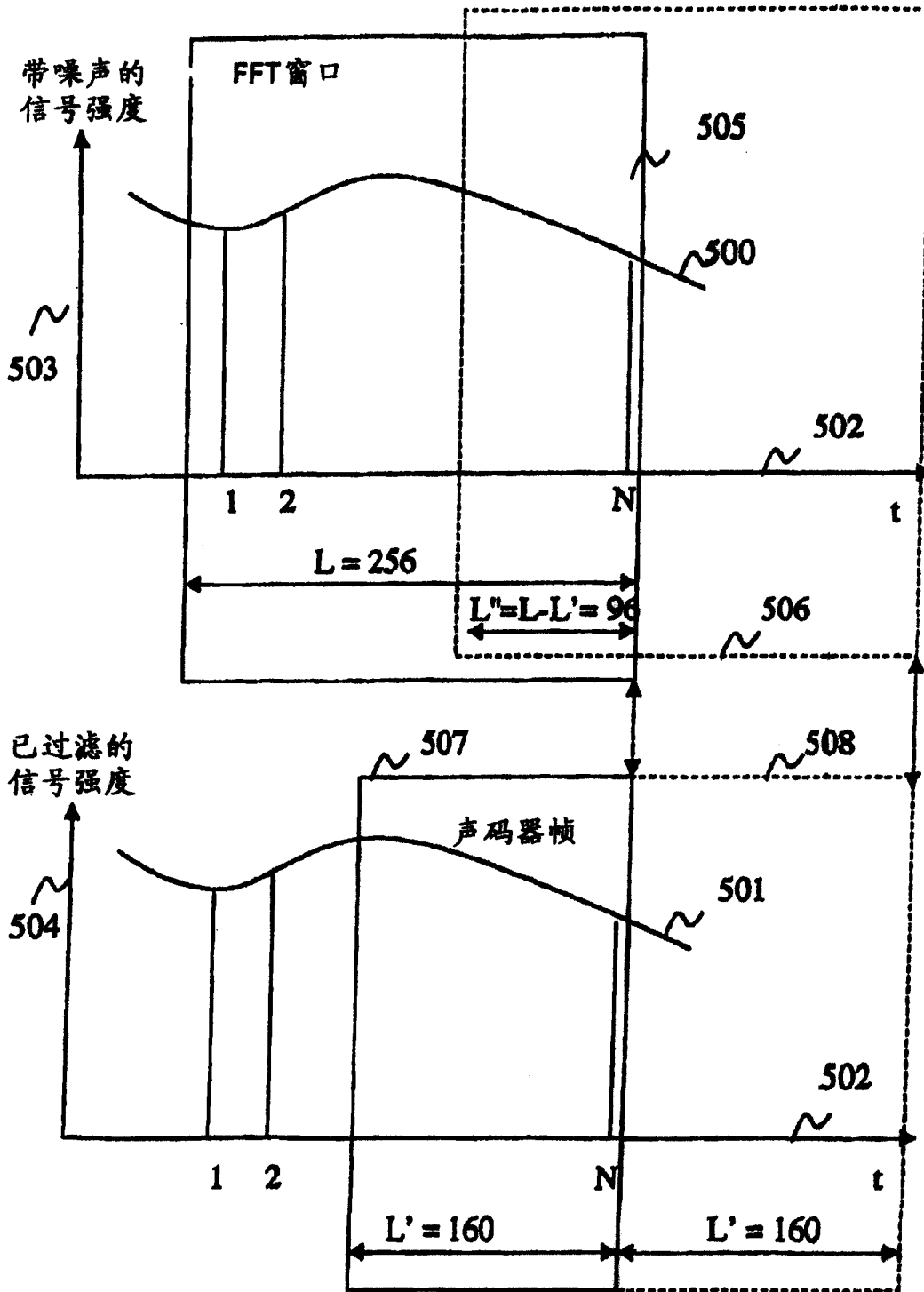


图5

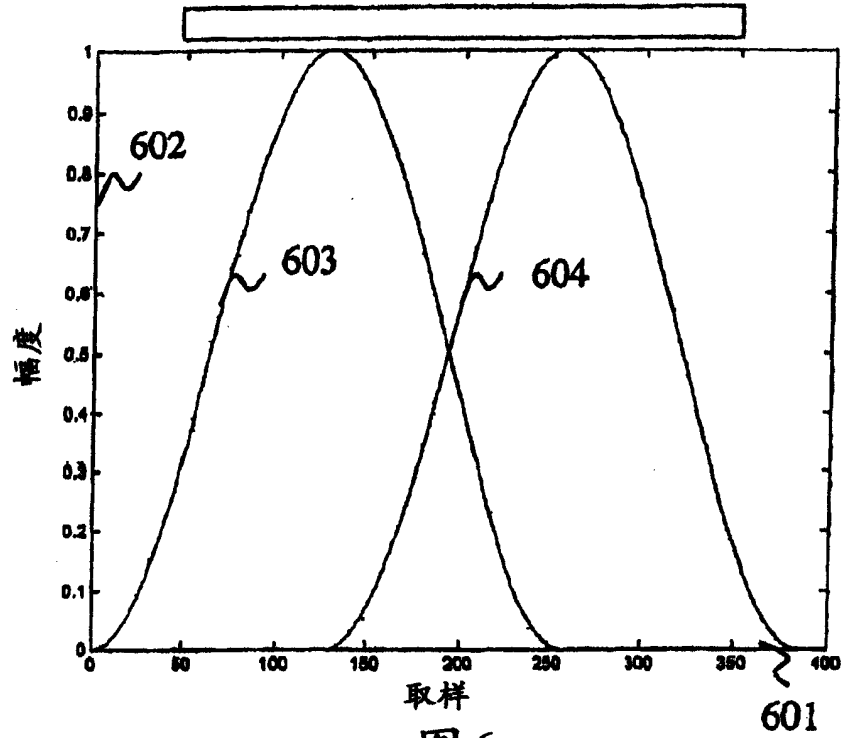


图6

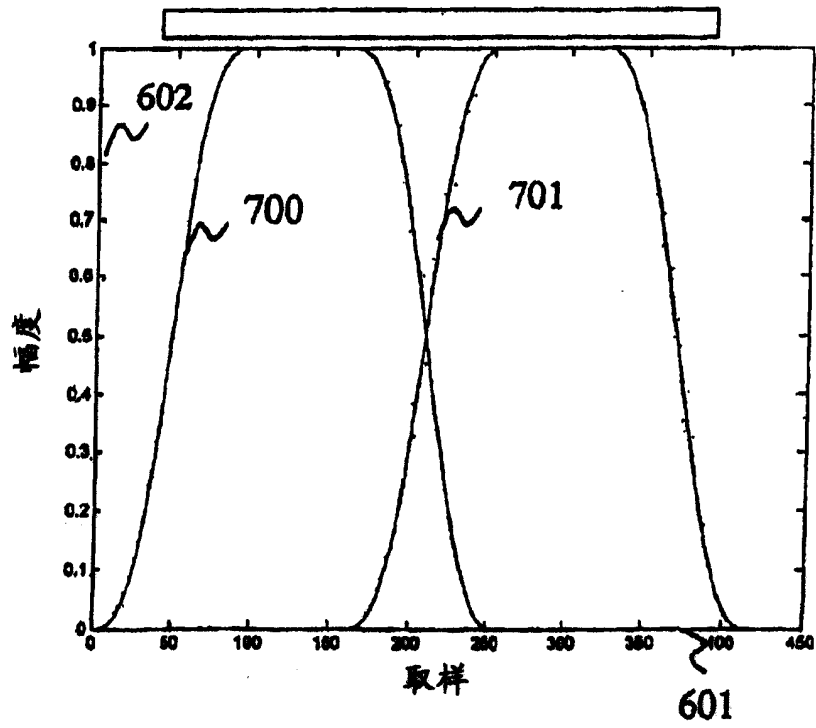


图7

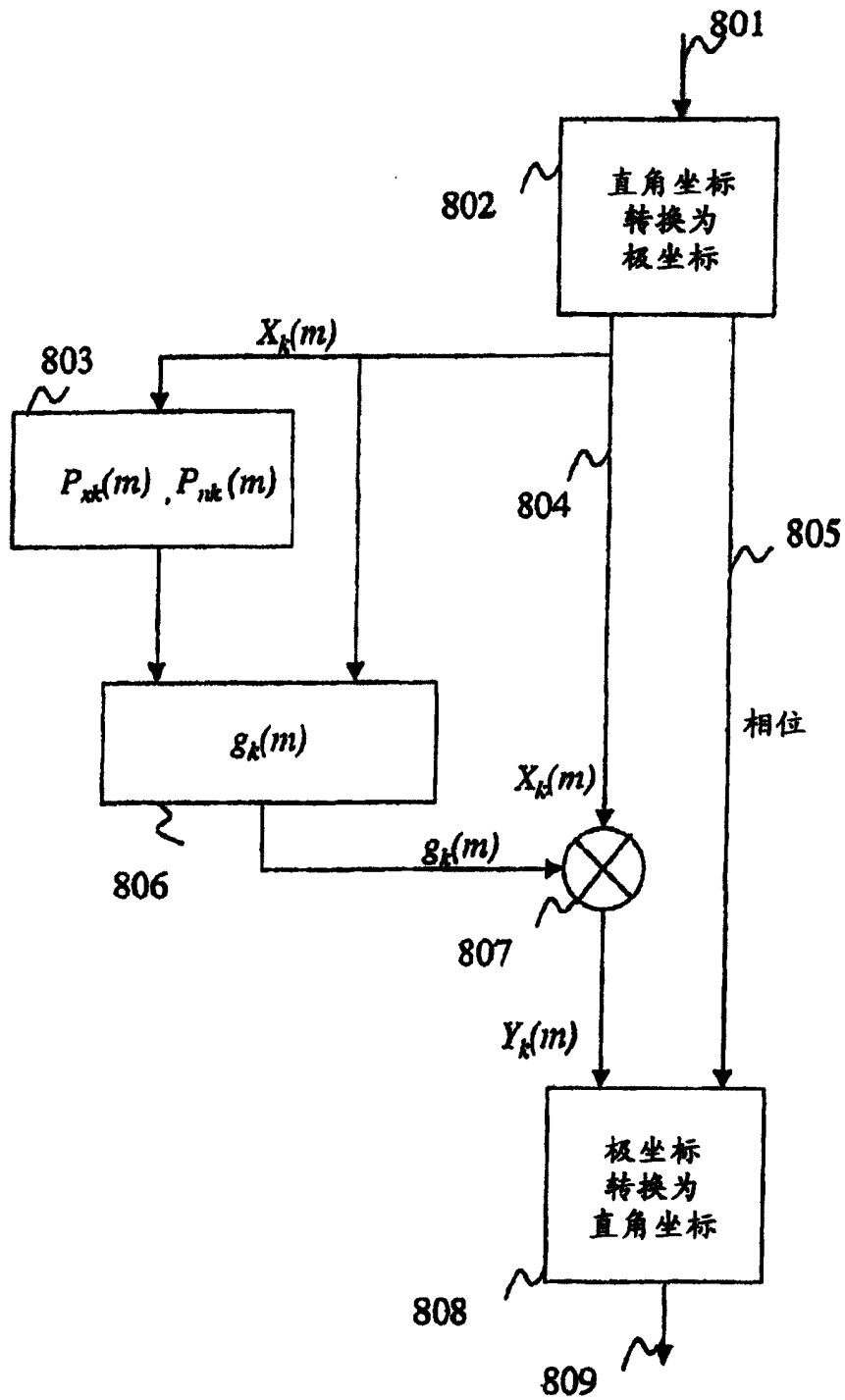


图 8