



# [12] 发明专利说明书

[21] ZL 专利号 00808870.5

[45] 授权公告日 2004 年 11 月 10 日

[11] 授权公告号 CN 1175709C

[22] 申请日 2000.4.11 [21] 申请号 00808870.5

[30] 优先权

[32] 1999.4.12 [33] US [31] 09/289,065

[86] 国际申请 PCT/EP2000/003223 2000.4.11

[87] 国际公布 WO2000/062579 英 2000.10.19

[85] 进入国家阶段日期 2001.12.12

[71] 专利权人 艾利森电话股份有限公司

地址 瑞典斯德哥尔摩

[72] 发明人 H·古斯塔夫松 I·克莱松

S·诺德霍尔姆

审查员 王晓丽

[74] 专利代理机构 中国专利代理(香港)有限公司

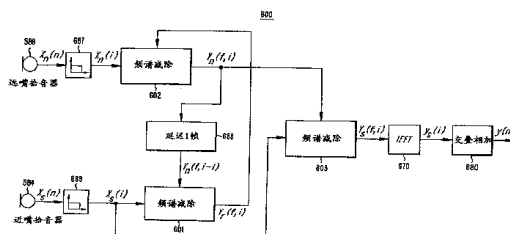
代理人 邹光新 李亚非

权利要求书 3 页 说明书 20 页 附图 6 页

[54] 发明名称 利用频谱减除的双拾音器信号降噪系统和方法

[57] 摘要

在双拾音器降噪系统中，通过采用线性卷积、因果滤波和/或按频谱对频谱减除增益函数进行指数平均的一些频谱减除算法提供了语音增强。按照一些示范性的实施例，在用一个远嘴拾音器与一个近嘴拾音器配合时，只要可以不断地从单个输入样点块估计出噪声谱就能处理非平稳的背景噪声。远嘴拾音器除拾取背景噪声之外，还拾取说话者的语音，但所得到的电平比近嘴拾音器的低。为了加强噪声估计，用一个频谱减除级来抑制远嘴拾音器信号内的语音。为了加强噪声估计，用另一个频谱减除级根据近嘴信号形成一个粗语音估计。最后，用一个第三频谱减除函数用加强的背景噪声估计抑制背景噪声来增强近嘴信号。



1. 一种降噪系统，所述降噪系统包括：

一个第一频谱减除处理器（601），配置成对一个第一信号滤波，提供一个噪声得到减小的输出信号；

5 一个第二频谱减除处理器（602），配置成对一个第二信号滤波，提供一个噪声估计输出信号；以及

一个第三频谱减除处理器（603），配置成按所述噪声估计输出信号对所述第一信号滤波；

一个第一拾音器；以及

10 一个第二拾音器，

其中所述第一信号从所述第一拾音器的输出端得到，而所述第二信号从所述第二拾音器的输出端得到。

2. 权利要求1的系统，其中所述第二频谱减除处理器配置成按所述第一噪声得到减小的输出信号对所述第二信号滤波。

15 3. 权利要求1的系统，所述系统还包括：

一个延迟电路，所述噪声估计输出信号加到所述延迟电路的一个输入端上；以及

其中所述第一频谱减除处理器配置成按所述延迟电路的输出对所述第一信号滤波。

20 4. 权利要求1的系统，其中所述第一拾音器与所述第二拾音器属于不同类型的拾音器。

5. 权利要求1的系统，其中所述第一拾音器比所述第二拾音器接近一个所要求的声波源。

25 6. 权利要求1的系统，其中至少所述第一、第二和第三频谱减除处理器之一的一个增益函数根据一个对一个输入信号的谱密度的估计和根据对所述输入信号中的不希望有的成分的谱密度的估计计算，

其中所述至少所述第一、第二和第三频谱减除处理器之一的输出信号的一个样点块根据所述输入信号的一个相应样点块和根据所述增益函数的一个相应样点块计算，以及

30 其中所述输入信号的相应样点块的阶与所述增益函数的相应样点块的阶之和小于输出信号块的样点数。

7. 权利要求6的系统，其中一个相位加到所述增益函数上使得至

少所述第一、第二和第三频谱减除处理器之一提供因果滤波。

8. 权利要求7的系统，其中所述增益函数具有线性相位。

9. 权利要求7的系统，其中所述增益函数具有最小相位。

10. 一种处理一个有噪声的输入信号和一个噪声信号从而提供一个噪声得到减小的输出信号的方法，所述方法包括下列步骤：

(a) 利用第一和第二拾音器提供输入信号；

(b) 利用第一频谱减除(601)对所述有噪声的输入信号滤波，提供一个第一噪声得到减小的输出信号；

(c) 利用第二频谱减除(602)对所述噪声信号滤波，提供一个噪声估计输出信号；以及

(d) 利用第三频谱减除(603)按所述噪声估计输出信号对所述有噪声的输入信号滤波。

11. 权利要求10的方法，其中所述步骤(b)根据所述第一噪声得到减小的输出信号对所述噪声信号滤波。

12. 权利要求10的方法，所述方法还包括下列步骤：

(d) 延迟所述噪声估计输出信号；以及

其中所述步骤(a)还包括利用所述频谱减除按所述步骤(d)的结果对所述有噪声的输入信号滤波，提供所述第一噪声得到减小的输出信号。

13. 权利要求10的方法，其中至少所述第一、第二和第三频谱减除处理器之一的一个增益函数根据一个对一个输入信号的谱密度的估计和根据一个对所述输入信号中的不希望有的成分的谱密度的估计计算，

其中所述至少所述第一、第二和第三频谱减除处理器之一的输出信号的一个样点块根据所述输入信号的一个相应样点块和根据所述增益函数的一个相应样点块计算，以及

其中所述输入信号的相应样点块的阶与所述增益函数的相应样点块的阶之和小于输出信号块的样点数。

14. 权利要求13的方法，其中一个相位加到所述增益函数上使得至少所述第一、第二和第三频谱减除处理器之一提供因果滤波。

15. 权利要求14的方法，其中所述增益函数具有线性相位。

16. 权利要求14的方法，其中所述增益函数具有最小相位。

17. 一种移动电话机，所述移动电话机包括：  
一个第一拾音器(584)；以及  
一个第二拾音器(586)，  
一个接收从一个第一拾音器(584)得到的第一信号的输入端；  
5 一个接收从一个第二拾音器(586)得到的一个第二信号的输入端；  
一个第一频谱减除处理器(601)，配置成对所述第一信号滤波，提供一个噪声得到减小的输出信号；  
一个第二频谱减除处理器(602)，配置成对所述第二信号滤波，  
10 提供一个噪声估计输出信号；以及  
一个第三频谱减除处理器(603)，配置成按所述噪声估计输出信号对所述第一信号滤波。
18. 权利要求17的移动电话机，其中所述第二频谱减除处理器配置成按所述第一噪声得到减小的输出信号对所述第二信号滤波。
19. 权利要求17的移动电话机，所述移动电话机还包括：  
15 一个延迟电路，所述噪声估计输出信号加到所述延迟电路的一个输入端上；以及  
其中所述第一频谱减除处理器配置成按所述延迟电路的输出对所述第一信号滤波。
20. 权利要求17的移动电话机，其中所述第一拾音器与所述第二拾音器属于不同类型的拾音器。
21. 权利要求17的移动电话机，其中所述第一拾音器比所述第二拾音器接近一个所要求的声波源。

## 利用频谱减除的双拾音器信号降噪系统和方法

## 发明领域

5 本发明与通信系统有关，特别是与减小通信信号中的破坏性背景噪声分量的影响的方法和设备有关。

## 发明背景

现在，技术和用户要求生产体积趋小的移动电话机。随着移动电话机越来越小，拾音器的位置在使用期间越来越离开说话者(近端用户)的嘴。这种距离的增大就更需要语音增强来抑制被拾音器拾取、发送给远端用户的破坏性背景噪声。也就是说，由于使用新的小移动电话机时拾音器与近端用户之间的距离比较大，因此拾音器所拾取的不仅是近端用户的语音，而且还有近端用户处的任何噪声。例如，近端拾音器通常拾取诸如周围的交通、道路和客舱噪声、室内噪声之类的声音。10 所得到的有噪声的近端语音对于远端用户来说可能是很讨厌的，甚至是无法忍受的。因此所希望的的是尽可能减小背景噪声，最好在近端信号处理链中及早处理(例如，在接收到的近端拾音器信号加到近端语音编码器前处理)。

为了抑制背景噪声，一些电话系统包括一个降噪处理器，用来消除加在近端信号处理链输入端上的背景噪声。图1为这样的10 一个系统100的总体方框图。在图1中，降噪处理器110设置在拾音器120的输出端与近端信号处理通路(未示出)的输入端之间。在工作中，降噪处理器110接收来自拾音器120的有噪声的语音信号 $x$ ，加以处理后提供较干净的噪声减小了的语音信号 $s_{NR}$ ，通过近端信号处理链，最终送至远端用户。

25 一种众所周知的实现图1的降噪处理器110的方法在这个技术领域内称为频谱减除(Spectral Subtraction)。例如可参见S.F.Boll的“利用频谱减除抑制语音中的音频噪声”(“Suppression of Acoustic Noise in Speech using Spectral Subtraction”，IEEE Trans. Acoust. Speech and Sig. Proc., 27:113-120, 1979)，该文在此列作参考全面引用。概括地说，频谱减除利用对噪声谱和有噪声的语音谱的估计形成30 一个基于信噪比(SNR)的增益函数，用它乘以输入频谱来抑制SNR低的频率。虽然频谱减除确实提供了值得注意的降噪，但是它也有

一些众所周知的缺点。例如，频谱减除的输出信号通常含有在这个技术领域称为乐音的不自然现象。此外，经处理的信号块之间的不连续点往往导致从远端用户看来语音质量有所下降。

近年来已经研究了许多对这种基本的频谱减除方法的改进措施。

5 例如参见N. Virage的“基于听觉系统的遮蔽特性的语音增强”（“Speech Enhancement Based on Masking Properties of the Auditory System”，IEEE ICASSP. Proc. 796-799 vol.1, 1995），D. Tsoukalas、M. Paraskevas和J. Mourjopoulos的“应用心理声学准则的语音增强”（“Speech Enhancement using Psychoacoustic Criteria”，IEEE ICASSP. Proc., 359-362 vol.2, 1993），F. Xie和D. Van Compernelle的“利用频谱幅度估计的语音增强，一种统一的方法”（“Speech Enhancement by Spectral Magnitude Estimation - A Unifying Approach”），IEEE Speech Communication, 89-104 vol.19, 1996），R. Martin的“基于频谱减除的最小统计特性”（“Spectral Subtraction Based on

10 Minimum Statistics”，UESIPCO, Proc., 1182-1185 vol ; 2, 1994），以及J. A. Heinen的“改进被有色、非平稳噪声污染的语音的频谱减除方法”（“A Spectral Subtraction Method for Enhancement of Speech Corrupted by Nonwhite, Nonstationary Noise”，IEEE IECON. Proc., 872-877 vol.2, 1995）。

20 法国专利No. 2, 768, 547揭示了一种从语音信号中减去噪声信号以提供噪声得到减小的信号的方法。这种方法用了单一的频谱减法器。在第一减除级中指出频谱中噪声分量增大的估计。这允许产生一个第一噪声得到减小的信号的频谱分量。这个过程还包括通过将一个听觉模型用于第一噪声得到减小的信号的频谱分量计算出一个遮蔽曲线。

25 第二减除级用帧内语音信号的每个频谱分量触发。从每个频谱分量中减去一个取决于一些参数（包括与噪声相应的频谱分量增大的估计与计算得的遮蔽曲线之间的差）的量。结果产生一个噪声得到减小的语音信号。

欧洲专利A0084982也揭示了一种减小语音信号内的噪声的方法。

30 这种方法用了两个拾音器，一个布置成接近于语音源，而第二个布置成离第一拾音器有一段距离。第二拾音器接收的一批信号按照Widrow算法加以修改。第一和第二信号加在一起形成输出信号。以只有0.1秒

的延迟实现噪声相消。

新近，频谱减除已经用正确的卷积和取决于频谱的指数增益函数平均来实现。在1998年5月27日递交的共同待决美国专利申请No.09/084,387“利用线性卷积和因果滤波器的频谱减除减小信号噪声”(“Signal Noise Reduction by Spectral Subtraction using Linear Convolution and Causal Filtering”)和也是1998年5月27日递交的共同待决美国专利申请No.09/084,503“利用取决于频谱的指数增益函数平均的频谱减除减小信号噪声”(“Signal Noise Reduction by Spectral Subtraction using Spectrum Dependent Exponential Gain Function Averaging”)中揭示了这些技术。

频谱减除利用了两个频谱估计，一个是对“受扰”信号的而一个是对“干扰”信号的，形成一个基于信噪比(SNR)的增益函数。受扰频谱乘以这个增益函数就可以提高频谱的SNR。在诸如配合免提电话机使用的单拾音器频谱减除应用中，根据干扰的背景噪声增强语音。噪

声在语音间歇期期间估计，或者用一个在语音期期间的噪声模型估计。这意味着噪声必须是平稳的，在语音期期间具有类似的性质，或者模型适合于活动的背景噪声。不幸的是，对于大多数日常环境中的大多数背景噪声，情况都不是这样。

5 因此，有必要开发一种采用频谱减除技术的降噪系统，可以适合在有大多数日常的可变背景噪声的环境中使用。

#### 发明概要

本发明通过在一个双拾音器系统内提供用频谱减除执行降噪的方法和设备满足了以上所述及其他一些需要。按照一些示范性的实施例，  
10 在用一个远嘴拾音器与一个近嘴拾音器配合时，只要可以不断地从单个输入样点块估计出噪声谱就能处理非平稳的背景噪声。远嘴拾音器除拾取背景噪声之外，还拾取说话者的语音，但所得到的电平比近嘴拾音器的低。为了加强噪声估计，用一个频谱减除级来抑制远嘴拾音器信号内的语音。为了能加强噪声估计，用另一个频谱减除级根据近嘴信号形成一个粗语音估计。最后，用一个第三频谱减除级用加强的  
15 背景噪声估计抑制背景噪声来增强近嘴信号。

以下将结合附图所示的一些例示性例子详细说明本发明的以上所述及其他一些特性和优点。熟悉该技术领域的人员可以理解，在这里所揭示的这些实施例只是为了说明本发明从而能更好地理解本发明，  
20 因此可以想象本发明有许多等效的具体实现方式。

#### 附图简要说明

图1为一个可以实现频谱减除的降噪系统的方框图；

图2示出了一个传统的频谱减除降噪处理器；

图3-4示出了按照本发明的示范性实施例设计的示范性的频谱减除  
25 降噪处理器；

图5示出了在本发明的一个示范性的实施例中近嘴和远嘴拾音器的布置情况；

图6示出了一个示范性的双拾音器频谱减除系统；以及

图7示出了一个用于本发明的一个示范性的实施例的示范性的频谱  
30 减除级。

#### 发明详细说明

为了理解本发明的各个特色和优点，首先来看一下传统的频谱减

除技术是有益的。概括地说，频谱减除建立在通信应用中的噪声信号和语音信号都是随机和不相关的、相加后形成有噪声的语音信号的假设上。例如，如果  $s(n)$ 、 $w(n)$  和  $x(n)$  分别表示语音、噪声和有噪声的语音的短期统计平稳的过程，于是有：

5

$$x(n) = s(n) + w(n) \quad (1)$$

$$R_x(f) = R_s(f) + R_w(f) \quad (2)$$

其中， $R(f)$  表示随机过程的功率谱密度。

10 噪声功率谱密度  $R_w(f)$  可以在语音间歇期间（即此时  $x(n) = w(n)$ ）。为了估计语音的功率谱密度，形成如下估计：

$$\hat{R}_s(f) = \hat{R}_x(f) - \hat{R}_w(f) \quad (3)$$

15 这种估计功率谱密度的传统方式需使用一种周期图（periodogram）。例如，如果  $X_N(f_u)$  为  $x(n)$  的  $N$  点傅里叶变换而  $W_N(f_u)$  为  $w(n)$  的相应傅里叶变换，于是：

$$\hat{R}_x(f_u) = P_{x,N}(f_u) = \frac{1}{N} |X_N(f_u)|^2, \quad f_u = \frac{u}{N}, \quad u=0, \dots, N-1 \quad (4)$$

$$\hat{R}_w(f_u) = P_{w,N}(f_u) = \frac{1}{N} |W_N(f_u)|^2, \quad f_u = \frac{u}{N}, \quad u=0, \dots, N-1 \quad (5)$$

式（3）、（4）和（5）可以合并成：

20

$$|S_N(f_u)|^2 = |X_N(f_u)|^2 - |W_N(f_u)|^2 \quad (6)$$

或者，可以给出更为广义的形式：

$$|S_N(f_u)|^a = |X_N(f_u)|^a - |W_N(f_u)|^a \quad (7)$$

其中，功率谱密度改换成谱密度的广义形式。

5 由于人的耳朵对语音的相位误差并不敏感，干净的语音的相位  $\phi_s(f)$  可以用有噪声的语音的相位  $\phi_x(f_u)$  来近似：

$$\phi_s(f_u) \approx \phi_x(f_u) \quad (8)$$

因此，估计干净的语音的傅里叶变换的广义表达式为

10

$$S_N(f_u) = (|X_N(f_u)|^a - k \cdot |W_N(f_u)|^a)^{\frac{1}{a}} \cdot e^{j\phi_x(f_u)} \quad (9)$$

其中，参数  $k$  用来控制噪声减去量。

为了简化表示，引入矢量形式：

$$X_N = \begin{pmatrix} X_N(f_0) \\ X_N(f_1) \\ \vdots \\ X_N(f_{N-1}) \end{pmatrix} \quad (10)$$

15

这些矢量按元对元进行计算。为了清楚起见，矢量的元对元相乘在这里表示为  $\odot$ 。因此，式(9)可以用增益函数  $G_N$  和矢量表示法写成：

$$S_N = G_N \odot |X_N| \odot e^{j\phi_x} = G_N \odot X_N \quad (11)$$

20 其中，增益函数为：

$$G_N = \left( \frac{|X_N|^a - k \cdot |W_N|^a}{|X_N|^a} \right)^{\frac{1}{a}} = \left( 1 - k \cdot \frac{|W_N|^a}{|X_N|^a} \right)^{\frac{1}{a}} \quad (12)$$

式(12)表示了传统的频谱减除算法,示于图2。在图2中,传统的频谱减除降噪处理器200包括快速傅里叶变换处理器210、幅度平方处理器220、语音活动检测器230、按块平均装置240、按块增益计算处理器250、乘法器260和快速傅里叶逆变换处理器270。

如图所示,有噪声的语音输入信号加到快速傅里叶变换处理器210的输入端上,快速傅里叶变换处理器210的输出端与幅度平方处理器220的输入端和乘法器260的第一输入端连接。幅度平方处理器220的输出端与开关225的第一触点和增益计算处理器250的第一输入端连接。语音活动检测器230的输出端与开关225的投掷输入端连接,而开关225的第二触点与按块平均装置240的输入端连接。按块平均装置240的输出端与增益计算处理器250的第二输入端连接,增益计算处理器250的输出端与乘法器260的第二输入端连接。乘法器260的输出端与快速傅里叶逆变换处理器270的输入端连接,快速傅里叶逆变换处理器270的输出端提供传统的频谱减除系统200的输出。

在工作中,这个传统的频谱减除系统200用如上所述的传统频谱减除算法处理来到的有噪声的语音信号,提供噪声得到减小的较干净的语音信号。在实践中,图2中的各个组成部分可以用任何已知的数字信号处理技术实现,包括通用计算机、一些集成电路和/或专用集成电路(ASIC)。

要注意的是,传统的频谱减除算法有两个参数, $\alpha$ 和 $k$ ,控制噪声减去量和语音质量。将第一参数 $\alpha$ 设置为2就最进行功率谱减除,而将第一参数 $\alpha$ 设置为1就是进行幅度谱减除。此外,将第一参数 $\alpha$ 设置为0.5可以增大降噪,而只使语音有适当的失真。这是由于在从有噪声的语音中减去噪声前频谱受到了压缩。

可以调整第二参数 $k$ 来达到降噪要求。例如,如果选择较大的 $k$ ,就会增大语音失真。在实践中,参数 $k$ 通常根据第一参数的选择设置。减小 $\alpha$ 通常会导致也要减小参数 $k$ ,以便保持语音失真小。在功率谱减

除的情况下，通常是用过减除(即， $k > 1$ )。

传统的频谱减除增益函数(见式(12))是根据一个全块估计得出的，具有零相位。结果，相应的脉冲响应  $g_N(u)$  就是非因果的，长度为  $N$  (等于块长度)。因此，增益函数  $G_N(l)$  与输入信号  $X_N$  的相乘(见式(11))导致与一个非因果滤波器的周期性循环卷积。如上所述，周期性的循环卷积可以导致不合乎需要的在时域的混叠，而滤波器的这种非因果性质可以导致块与块之间的不连续，从而使语音质量较差。有益的是，本发明所给出的方法和设备提供了与一个因果增益滤波器的正确卷积，从而消除了以上所述的时域混叠和块间不连续的问题。

对于时域混叠问题，要注意的是在时域的卷积相当于在频域的相乘。也就是说：

$$x(u) * y(u) \rightarrow X(f) \cdot Y(f), \quad u = -\infty, \dots, \infty \quad (13)$$

在变换用长度为  $N$  的快速傅里叶变换 (FFT) 得到时，相乘的结果不是正确的卷积。确切些说，结果是周期为  $N$  的循环卷积：

$$x_N \textcircled{N} y_N \quad (14)$$

其中，符号  $\textcircled{N}$  表示循环卷积。

为了在用快速傅里叶变换时得到正确的卷积，脉冲响应  $x_N$  和  $y_N$  的阶之和必须小于或等于块长度减1，即小于或等于  $(N-1)$ 。

因此，利用总阶数小于或等于  $N-1$  的一个增益函数  $G_N(l)$  和一个输入信号块  $X_N$  可以解决由于周期性的循环卷积产生的时域混叠问题。

按照传统的频谱减除，输入信号的频谱  $X_N$  具有整个块的长度。然而，按照本发明，用一个长度为  $L$  ( $L < N$ ) 的输入信号块  $x_L$  构造一个阶为  $L$  的频谱。长度  $L$  称为帧长度，因此  $x_L$  为一个帧。由于乘以长度为  $N$  的增益函数的频谱也应该具有长度  $N$ ，因此将帧  $x_L$  补零到整个块长度  $N$ ，得到  $X_{L/N}$ 。

为了构造一个长度为  $N$  的增益函数，按照本发明可以从一个长度为  $M$  ( $M < N$ ) 的增益函数  $G_M(l)$  内插，形成  $G_{M/N}(l)$ 。为了得到本发明的低阶增

益函数  $G_{MN}(l)$ ，可以用任何已知的或还没有公开的频谱估计技术来替换以上所揭示的简单的傅里叶变换周期图。有一些已知的谱估计技术可以使所得到的增益函数的方差较小。例如，参见 J.G. Proakis 和 D.G. Manolakis 的“数字信号处理：原理、算法和应用”（“Digital Signal Processing; Principles, Algorithms, and Applications”，Macmillan, Second Ed., 1992）。

按照例如众所周知的 Bartlett 方法，将长度为  $N$  的块分成  $K$  个长度为  $M$  的子块，然后计算出每个子块的周期图，对结果进行平均后得到一个长为  $M$  的对于整个块的周期图：

10

$$\begin{aligned} P_{x,M}(f_u) &= \frac{1}{K} \sum_{k=0}^{K-1} P_{x,M,k}(f_u), \quad f_u = \frac{u}{M}, \quad u=0, \dots, M-1 \\ &= \frac{1}{K} \sum_{k=0}^{K-1} |\mathcal{F}(x(k \cdot M + u))|^2 \end{aligned} \quad (15)$$

有益的是，在各子块不相关时，与整个块长度的周期图相比，方差减小了一个因子  $K$ 。频率分辨率也降低了相同的因子。

或者，也可以采用 Welch 方法。Welch 方法与 Bartlett 方法类似，只是每个子块加了一个 Hanning 窗，而且允许子块相互交叠，从而得到较多个子块。Welch 方法给出的方差比 Bartlett 方法的还要小。Bartlett 和 Welch 方法只是两种谱估计技术，而其他已知的谱估计技术也可以采用。

无论所实现的是哪种精确的谱估计技术，还可以利用平均技术进一步减小噪声周期图估计的方差。例如，在噪声为长期平稳的假设下，可以对根据上述 Bartlett 和 Welch 方法得到的周期图进行平均。一种技术是指数平均：

$$\bar{P}_{x,M}(l) = \alpha \cdot \bar{P}_{x,M}(l-1) + (1-\alpha) \cdot P_{x,M}(l) \quad (16)$$

在式 (16) 中，函数  $P_{x,M}(l)$  用 Bartlett 或 Welch 方法计算，函数  $\bar{P}_{x,M}(l)$  为对当前块的指数平均，而函数  $\bar{P}_{x,M}(l-1)$  为对上一个块的指数平

均。参数 $\alpha$ 控制指数记忆的长度，通常这长度不应该超过可以将噪声认为是平稳的长度。 $\alpha$ 接近于1导致有较长的指数记忆，可以显著地减小周期图的方差。

5 长度 $M$ 称为子块长度，所得到的低阶增益函数具有长度为 $M$ 的脉冲响应。因此，在组成增益函数中所用的噪声周期图估计 $\bar{P}_{x_L, M}(l)$ 和有噪声的语音周期图估计 $P_{x_L, M}(l)$ 也具有长度 $M$ ：

$$G_M(l) = \left( 1 - k \cdot \frac{\bar{P}_{x_L, M}^a(l)}{P_{x_L, M}^a(l)} \right)^{\frac{1}{\alpha}} \quad (17)$$

10 按照本发明，这是采用例如Bartlett方法从输入帧 $X_L$ 得到较短的周期图估计再加以平均实现的。Bartlett方法(或其他适当的估计方法)减小了所估计的周期图的方差，也减小了频率分辨率。分辨率从 $L$ 个频率区间减小到 $M$ 个频率区间意味着周期图估计 $P_{x_L, M}(l)$ 也具有长度 $M$ 。此外，还可以用如上所述的指数平均进一步减小噪声周期图估计 $\bar{P}_{x_L, M}(l)$ 的方差。

15 为了满足总阶数小于或等于 $N-1$ 的要求，应使帧长度 $L$ 加上子块长度 $M$ 小于 $N$ 。结果，可以形成所要求的输出块：

$$S_N = G_{M1M}(l) \odot X_{L1N} \quad (18)$$

20 有益的是，按照本发明设计的这种低阶滤波器还提供了解决由于传统的频谱减除算法中增益滤波器的非因果特性引起的问题(即块间不连续和语音质量降低)的机会。具体地说，按照本发明，可以为增益函数加上一个相位来提供一个因果滤波器。按照一些示范性的实施例，这个相位可以根据幅度函数设计，按需要可以是线性相位也可以最小相位。

25 为了按照本发明构造一个线性相位滤波器，首先注意到的是，如果FFT的块长度为长度 $M$ ，那么在时域的循环移位就是在频域乘以一个相位函数：

$$g(n-l)_M \approx G_M(f_u) \cdot e^{-j2\pi ul/M}, f_u = \frac{u}{M}, u = 0, \dots, M-1 \quad (19)$$

在当前情况下， $l$ 等于 $M/2+1$ ，因为脉冲响应的第一个位置应该具有零延迟(即，一个因果滤波器)。因此：

5

$$g(n-(M/2+1))_M \approx G_M(f_u) \cdot e^{-j\pi u(1+\frac{2}{M})} \quad (20)$$

从而得到线性相位滤波器 $\bar{G}_M(f_u)$ 为

$$\bar{G}_M(f_u) = G_M(f_u) \cdot e^{-j\pi u(1+\frac{2}{M})} \quad (21)$$

10 按照本发明，增益函数也内插到长度 $N$ ，这例如可以用一种平滑内插来执行。添加给增益函数的相位因此改变，得到：

$$\bar{G}_{M|N}(f_u) = G_{M|N}(f_u) \cdot e^{-j\pi u(1+\frac{2}{M})\frac{M}{N}} \quad (22)$$

15 有益的是，构造线性相位滤波器也可以在时域执行。在这种情况下，用IFFT将增益函数 $G_M(f_u)$ 变换到时域，执行循环移位。经移位的脉冲响应补零到长度 $N$ 后再用 $N$ 点FFT变换回来。这样就得到一个所需的经内插的因果线性相位滤波器 $\bar{G}_{MN}(f_u)$ 。

20 按照本发明设计的因果最小相位滤波器可以用Hilbert变换关系根据增益函数构造。例如，可参见A. V. Oppenheim和R. W. Schaffer的“离散时间信号处理”(“Discrete-Time Signal Processing”, Prentice-Hall, Inter. Ed., 1989)。Hubert变换关系隐含着一个复值函数的实部与虚部之间的独特关系。有益的是，对于幅度与相位之间的关系也可以使用这个关系，如果采用复信号的对数：

$$\begin{aligned} \ln\left(|G_M(f_u)| \cdot e^{j \arg(G_M(f_u))}\right) &= \ln(|G_M(f_u)|) + \ln\left(e^{j \arg(G_M(f_u))}\right) \\ &= \ln(|G_M(f_u)|) + j \arg(G_M(f_u)) \end{aligned} \quad (23)$$

在当前情况下，相位为零，从而得到一个实函数。函数  $\ln(|G_M(f_u)|)$  用长度为  $M$  的 IFFT 变换到时域后，形成  $g_M(n)$ 。将这个时域函数重新整理后，有：

$$\bar{g}_M(n) = \begin{cases} 2 \cdot g_M(n), & n=1, 2, \dots, M/2-1 \\ g_M(n), & n=0, M/2 \\ 0, & n=M/2+1, \dots, M-1 \end{cases} \quad (24)$$

函数  $\bar{g}_M(n)$  用  $M$  点 FFT 变换回频域，产生  $\ln(|\bar{G}_M(f_u)| \cdot e^{j \arg(\bar{G}_M(f_u))})$ 。按此形成函数  $\bar{G}_M(f_u)$ 。然后将因果最小相位滤波器  $\bar{G}_M(f_u)$  内插到长度  $N$ 。这个内插以与在如上所述的线性相位情况中相同的方式进行。这样得到的经内插的滤波器  $G_{M/N}(f_u)$  是因果的，具有大致最小相位。

图3示出了按照发明实现的上述频谱减除设计。在图3中，所示的提供线性卷积和因果滤波器的频谱减除降噪处理器300包括Bartlett处理器305、幅度平方处理器320、语音活动检测器330、块平均处理器340、低阶增益计算处理器350、增益相位处理器355、内插处理器356、乘法器360、快速傅里叶逆变换处理器370和交叠增添处理器380。

如图所示，有噪声的语音输入信号加到Bartlett处理器305的输入端上和加到快速傅里叶变换处理器310的输入端上。Bartlett处理器305的输出端与幅度平方处理器320的输入端连接，快速傅里叶变换处理器310的输出端与乘法器360的第一输入端连接。幅度平方处理器320的输出端与开关325的第一触点和低阶增益计算处理器350的第一输入端连接。语音活动检测器330的控制输出端与开关325的投掷输入端连接，开关325的第二触点与按块平均装置340的输入端连接。

按块平均装置340的输出端与低阶增益计算处理器350的第二输入端连接，低阶增益计算处理器350的输出端与增益相位处理器355的输入端连接。增益相位处理器355的输出端与内插处理器356的输入端连

接，内插处理器356的输出端与乘法器360的第二输入端连接。乘法器360的输出端与快速傅里叶逆变换处理器370的输入端连接，快速傅里叶逆变换处理器370的输出端与交叠相加处理器380连接。交叠相加处理器380的输出端给出这个示范性的降噪处理器300的噪声得到减小的、干净的语音输出。

在工作中，频谱减除降噪处理器300用如上所述的线性卷积因果滤波算法处理来到的有噪声的语音信号，提供噪声得到减小的、干净的语音信号。在实践中，图3中的各个组成部分可以用任何已知的数字信号处理技术实现，包括通用计算机、一些集成电路和/或专用集成电路(ASIC)。

有益的是，本发明的增益函数 $G_M(l)$ 的方差还可以按照本发明通过受控制的指数增益函数平均进一步减小。按照一些示范性的实施例，这种平均按当前的块频谱 $P_{x,M}(l)$ 与平均噪声谱 $\bar{P}_{x,M}(l)$ 之间的差异进行。例如，在差异小时，可以对增益函数 $G_M(l)$ 进行长时间平均，这相当于是平稳背景噪声情况。相反，在差异大时，可以对增益函数 $G_M(l)$ 进行短时间平均，这相当于有语音的情况或者背景噪声改变很快的情况。

为了处理从语音期到背景噪声期的暂态转换，对增益函数的平均不是与差异的减小成正比增大，因为这样的话会引入一个听得到的寄生语音(因为适合于语音频谱的增益函数要保持一段长的时间)。代之以允许渐渐增大平均，可以提供使增益函数适应平稳的输入的时间。

按照一些示范性实施例，频谱之间的差异的测度定义为

$$\beta(l) = \frac{\sum_u |P_{x,M,u}(l) - \bar{P}_{x,M,u}(l)|}{\sum_u \bar{P}_{x,M,u}(l)} \quad (25)$$

其中， $\beta(l)$ 限制为

$$\beta(l) = \begin{cases} 1, & \beta(l) > 1 \\ \beta(l), & \beta_{min} \leq \beta(l) \leq 1, 0 \leq \beta_{min} < 1 \\ \beta_{min}, & \beta(l) < \beta_{min} \end{cases} \quad (26)$$

其中,  $\beta(l)=1$  导致不对增益函数进行指数平均, 而  $\beta(l)=\beta_{\min}$  提供最大限度的指数平均。

参数  $\bar{\beta}(l)$  为频谱之间差异的指数平均, 示为

$$\bar{\beta}(l) = \gamma \cdot \bar{\beta}(l-1) + (1-\gamma) \cdot \beta(l) \quad (27)$$

5

式(27)中的参数  $\gamma$  用来保证在出现从频谱之间差异大的时期过渡到差异小的时期时增益函数适应新的电平。如上所述, 这是为了防止寄生语音。按照这些示范性的实施例, 这种自适应是在由于  $\beta(l)$  减小而开始增大对增益函数的指数平均前完成的。因此:

10

$$\gamma = \begin{cases} 0, & \bar{\beta}(l-1) < \beta(l) \\ \gamma_c, & \bar{\beta}(l-1) \geq \beta(l), \quad 0 < \gamma_c < 1 \end{cases} \quad (28)$$

在差异  $\beta(l)$  增大时, 参数  $\beta(l)$  随着增大, 但是在差异减小时, 对  $\beta(l)$  用了指数平均, 形成平均的参数  $\bar{\beta}(l)$ 。增益函数的指数平均示为:

$$\bar{G}_M(l) = (1 - \bar{\beta}(l)) \cdot \bar{G}_M(l-1) + \bar{\beta}(l) \cdot G_M(l) \quad (29)$$

15

对于不同的输入信号情况, 以上这些等式可以说明如下。在噪声期期间, 方差减小。只要噪声谱对于每个频率都具有稳定的平均值, 就可以对它进行平均, 以减小方差。噪声电平改变将导致平均的噪声谱  $\bar{P}_{x,M}(l)$  与当前块的频谱  $P_{x,M}(l)$  之间的差异。因此, 这种受控的指数平均方法减小对增益函数的平均, 直到噪声电平稳定到一个新的电平上。这样就能处理噪声电平的改变, 既可以在平稳的噪声期期间减小方差, 也可以迅速响应噪声的改变。高能量的语音通常具有一些时变的频谱峰。在这些频谱峰以不同的块平均时, 它们的频谱估计含有这些尖峰的平均, 因此看来像一个较宽的频谱, 这将导致语音质量下降。因此, 指数平均在高能量的语音期期间保持为最小。由于平均的噪声谱  $\bar{P}_{x,M}(l)$  与当前高能量语音谱  $P_{x,M}(l)$  之间的差异大, 因此不对增益函数进行指数

20

25

平均。在较低能量的语音期期间，采用取决于当前低能量语音谱与平均噪声谱之间的差异的短记忆的指数平均。所以，对于低能量的语音，方差减小比在背景噪声期期间的小，而比高能量语音期的大。

图4示出了按照本发明设计的上述频谱减除方案。在图4中，提供  
5 线性卷积、因果滤波和受控指数平均的频谱减除降噪处理器400示为包括图3所示系统300的Bartlett处理器305、幅度平方处理器320、语音活动检测器330、按块平均装置340、低阶增益计算处理器350、增益相位处理器355、内插处理器356、乘法器360、快速傅里叶逆变换处理器370和交叠相加处理器380，以及平均控制处理器445、指数平均处理器  
10 446和可选的固定FIR后置滤波器465。

如图所示，有噪声的语音输入信号加到Bartlett处理器305的输入端上和加到快速傅里叶变换处理器310的输入端上。Bartlett处理器305的输出端与幅度平方处理器320的输入端连接，快速傅里叶变换处理器310的输出端与乘法器360的第一输入端连接。幅度平方处理器320的输  
15 出端与开关325的第一触点、低阶增益计算处理器350的第一输入端和平均控制处理器445的第一输入端连接。

语音活动检测器330的控制输出端与开关325的投掷输入端连接，开关325的第二触点与按块平均装置340的输入端连接。按块平均装置340的输出端与低阶增益计算处理器350的第二输入端和平均控制器445  
20 的第二输入端连接。低阶增益计算处理器350的输出端与指数平均处理器446的信号输入端连接，平均控制器445的输出端与指数平均处理器446的控制输入端连接。

指数平均处理器446的输出端与增益相位处理器355的输入端连接，增益相位处理器355的输出端与内插处理器356的输入端连接。内  
25 插处理器356的输出端与乘法器360的第二输入端连接，可选的固定FIR后置滤波器465的输出端与乘法器360的第三输入端连接。乘法器360的输出端与快速傅里叶逆变换处理器370的输入端连接，快速傅里叶逆变换处理器370的输出端与交叠相加处理器380的输入端连接。交叠相加处理器380的输出端给出这个示范性的系统400的干净的语音信号。

30 在工作中，按照本发明实现的频谱减除降噪处理器400用如上所述的线性卷积、因果滤波和受控指数平均算法处理来到的有噪声的语音信号，提供噪声得到减小的改善语音信号。就象图3的实施例那样，图

4中的各个组成部分可以用任何已知的数字信号处理技术实现，包括通用计算机、一些集成电路和/或专用集成电路(ASIC)。

5 注意，由于帧长度L和子块长度M之和按照示范性的实施例选择成不到N-1，因此可以如图4所示增添一个长度 $J \leq N-1-L-M$ 的附加固定FIR滤波器465。后置滤波器465将内插的滤波器脉冲响应与信号谱相乘，如图所示。这个内插到长度N是通过将滤波器补零再应用N点FFT来实现的。后置滤波器465用来滤除电话带宽之外或恒定的音调成分。或者，也可以将后置滤波器465的功能直接包括在增益函数内。

10 在实践中，上述算法的参数根据实现算法的具体应用设定。以下将以GSM移动电话机为例说明参数选择情况。

首先，根据GSM规范，将帧长度L设置为160个样点，提供为20 ms的帧。在其他系统中可以选用其他的L。然而，应指出的是帧长度L增大相当于延迟增大。使子块长度M(例如，Bartlett处理器的周期图长度)小可以提供大的方差缩减M。由于用FFT来计算周期图，因此将长度  
15 M设置为2的幂更为方便。于是，频率分辨率为：

$$B = \frac{F_s}{M} \quad (30)$$

GSM系统的采样率为8000 Hz。因此，长度M为16、32和64给出的频率分辨率分别为500Hz、250Hz和125Hz。

20 为了在一个诸如移动电话机之类的噪声可变的系统中采用上述的频谱减除技术，本发明应用了一种2拾音器系统。图5例示了这种2拾音器系统，其中，582为移动电话机，584为近嘴拾音器，而586为远嘴拾音器。在用一个远嘴拾音器与一个近嘴拾音器配合时，只要可以不断地从单个输入样点块估计出噪声谱就能处理非平稳的背景噪声。

25 远嘴拾音器586除拾取背景噪声之外，还拾取说话者的语音，但所得到的电平比近嘴拾音器584的低一些。为了加强噪声估计，用一个频谱减除级来抑制远嘴拾音器586的信号内的语音。为了能加强噪声估计，用另一个频谱减除级根据近嘴信号形成一个粗语音估计。最后，用一个第三频谱减除级通过滤除增强的背景噪声来增强近嘴信号。

30 由于语音和噪声估计只能根据一个短的数据样点块形成，因此采

用上述技术的一个潜在的问题是需要作出小方差的滤波器（即增益函数）估计。为了减小增益函数的变动，采用上面所讨论的单拾音器频谱减除算法。这样，这种方法利用减小方差的Bartlett谱估计方法来减小增益函数的变动。这种方法也减小了频率分辨率，而这个属性用来实现因果的实线性卷积。在本发明的一个示范性的实施例中，用噪声与有噪声的语音的谱估计之间的差异大小所控制的自适应平均进一步减小了增益函数的变动。

在本发明的这种2拾音器系统中，如图6所示，有两个信号：来自近嘴拾音器584的语音占优势的连续信号 $x_s(n)$ ；以及来自远嘴拾音器586的噪声更占优势的连续信号 $x_n(n)$ 。来自近嘴拾音器584的信号加到缓存器689的输入端上，在缓存器689内分解为一些块 $x_s(i)$ 。在本发明的所有示范性实施例中，缓存器689也是一个语音编码器。来自远嘴拾音器586的信号加到缓存器687的输入端上，在缓存器687内分解为一些块 $x_n(i)$ 。缓存器687和689也都可以包括进一步增强本发明的性能的诸如回声消除器之类的其他信号处理。模拟数字（A/D）变换器（未示出）将来自拾音器584、586的模拟信号变换成数字信号，从而可由本发明的这些频谱减除级处理。A/D变换器可以设在缓存器687、689前或后。

第一频谱减除级601将一个近嘴信号块 $x_s(i)$ 和上一帧的噪声估计 $Y_n(f, i-1)$ 作为它的输入。上一帧的噪声估计是通过将第二频谱减除级602的输出耦合到延迟电路688的输入端来产生的。延迟电路688的输出端与第一频谱减除级601连接。这个第一频谱减除级用来对语音作出一个粗估计 $Y_r(f, i)$ 。第一频谱减除级601的输出提供给第二频谱减除级602，第二频谱减除级602利用这个估计 $(Y_r(f, i))$ 和一个远嘴信号块 $x_n(i)$ 估计当前帧的噪声谱 $Y_n(f, i)$ 。最后，第二频谱减除级602的输出提供给第三频谱减除级603，第三频谱减除级603利用这个当前噪声谱估计 $Y_n(f, i)$ 和一个近嘴信号块 $x_s(i)$ 估计出噪声减小了的语音 $Y_s(f, i)$ 。第三频谱减除级603的输出端与快速傅里叶逆变换处理器670的输入端连接，快速傅里叶逆变换处理器670的输出端与交叠相加处理器680的输入端连接。交叠相加处理器680的输出端提供一个干净的语音信号作为这个示范性的系统600的输出。

在本发明的一个示范性的实施例中，频谱减除级601-603各有一个控制减除量的参数。这个参数最好为根据拾音器的输入SNR和所用的降

噪方法设置。此外，在本发明的另一个示范性的实施例中，用一个控制器动态地按可变的噪声环境更为精确地设置各频谱减除级601-603的这个参数。此外，由于远嘴拾音器信号用来估计要从近嘴有噪声的语音谱中减去的噪声谱，因此本发明的性能在背景噪声谱在两个拾音器内具有相同的特性时就较好。例如，在用有一个有方向性的近嘴拾音器时，它所得到的背景特性就与一个全方向的远嘴拾音器所得到的不同。为了补偿在这种情况下差别，两个拾音器信号或者其中一个拾音器信号应该予以滤波，以减小的谱的差别。

在本发明的一个示范性的实施例中，可取的是使电话通信中的延迟尽可能小，从而防止回声干扰和不自然的间歇。在信号块长度与移动电话系统的语音编码器的块长度匹配时，本发明使用与语音编码器相同的样点块。这样，缓存信号块不会引入额外的延迟。因此，引入的延迟只是本发明的降噪计算时间加上在最后的频谱减除级内进行增益函数滤波的群延迟。如在第三级所示，可以使振幅增益函数加上一个最小相位，从而在因果滤波器的制约下给出一个短的延迟。

由于本发明使用两个拾音器，因此不再需要使用如对图3和4中使用单个拾音器的频谱减除所述的VAD 330、开关325和平均块340。也就是说，远嘴拾音器可用来提供一个在语音期和非语音期期间都恒定的噪声信号。此外，IFFT 370和交叠相加电路380移到了最终输出级，示为图6中的670和680。

上述在双拾音器实现中所用的各个频谱减除级都可以如图7所示那样实现。在图7中，提供线性卷积、因果滤波和受控指数平均的频谱减除级700示为包括Bartlett处理器705、频率抽取器722、低阶增益计算处理器750、增益相位处理器和内插处理器755/756和乘法器760。

如图所示，有噪声的语音输入信号 $X_{(i)}$ 加到Bartlett处理器705的输入端和快速傅里叶变换处理器710的输入端。符号 $X_{(i)}$ 用来表示加到频谱减除级601-603的输入端上的 $X_n(i)$ 或 $X_s(i)$ ，如图6中所示。长度为N的干扰信号的振幅谱 $Y_{(c,m)}(f,i)$ 、 $Y_{(i)}(f,i)$ 加到频率抽取器722的输入端上。符号 $Y_{(i)}(f,i)$ 用来表示 $Y_n(f,i-1)$ 、 $Y_s(f,i)$ 或 $Y_n(f,i)$ 。频率抽取器722的输出为具有长度M的 $Y_{(c,m)}(f,i)$ 振幅谱，其中 $M < N$ 。此外，频率抽取器722使输出振幅谱的方差比输入振幅谱的小。Bartlett处理器705的振幅谱输出和频率抽取器722的振幅谱分别加到低阶增益计算处理器750的输入端

上。快速傅里叶变换处理器710的输出端与乘法器760的第一输入端连接。

低阶增益计算处理器750的输出端与可选的指数平均处理器746的信号输入端连接。指数平均处理器746的输出端与增益相位和内插处理器755/756的输入端连接。处理器755/756的输出端与乘法器760的第二输入端连接。经滤波的谱 $Y_r(f,i)$ 因此就是乘法器760的输出，其中符号 $Y_r(f,i)$ 用来表示 $Y_r(f,i)$ 、 $Y_n(f,i)$ 或 $Y_s(f,i)$ 。图7中所用的增益函数为：

$$G_M(f,i) = \left( 1 - k_{(c)} \cdot \frac{|Y_{(c),M}(f,i)|^\alpha}{|X_{(c),M}(f,i)|^\alpha} \right)^{\frac{1}{\alpha}} \quad (31)$$

其中 $|X_{(c),M}(f,i)|$ 为Bartlett处理器705的输出， $|Y_{(c),M}(f,i)|$ 为频率抽取器722的输出， $\alpha$ 为谱指数， $k_{(c)}$ 为控制用于具体频谱减除级的抑制量的减除因子。可以将增益函数选择成自适应平均的。这种增益函数相应于一个非因果时变滤波器。得到一个因果滤波器的一种方法是强加一个最小相位。得到一个因果滤波器的另一种方法是强加一个线性相位。为了得到一个有着与输入块 $X_{(c),N}(f,i)$ 相同的FFT区间数的增益函数 $G_M(f,i)$ ，将增益函数内插为 $G_{MN}(f,i)$ 。增益函数 $G_{MN}(f,i)$ 现在就相应于一个具有长度M的因果线性滤波器。经传统的FFT滤波后，可以得到一个没有周期性效应的输出信号。

在工作中，按照本发明设计的频谱减除级700用如上所述的线性卷积、因果滤波和受控指数平均算法处理来到的有噪声的语音信号，提供噪声得到减小的改善语音信号。就象图3和4的实施例那样，图7-8中的各个组成部分可以用任何已知的数字信号处理技术实现，包括通用计算机、一些集成电路和/或专用集成电路(ASIC)。

总之，本发明提供了改进的利用线性卷积、因果滤波和/或对增益函数的受控指数平均的双拾音器频谱减除的方法和设备。熟悉该技术领域的人员很容易认识到本发明可以提高诸如音乐之类的任何音频信号的质量，而不只是局限于语音音频信号。这些示范性的方法可以处理非平稳的背景噪声，因为本发明所根据的不只是对只有噪声的那些时段进行的噪声测量。此外，在短期平稳的背景噪声期间，由于可以

在只有噪声和也有语音期间估计背景噪声，因此语音质量也得到了改善。此外，本发明可以使用有或没有方向性的拾音器，而且两个拾音器的类型也可以不同。此外，降噪量可以调节到适当的程度，以适应具体要求的语音质量。

- 5       熟悉该技术领域的人员可以理解，本发明并不局限于在这里为了说明而例举的这些具体的示范性实施例，有许多可能的实施方式也是可行的。例如，虽然本发明以移动通信应用为例进行说明，但是熟悉该技术领域的人员可以理解，本发明的原理同样可用于任何需要除去特定信号分量的信号处理应用。因此，本发明的专利保护范围由所附
- 10 权利要求书限定而不是上述说明，所附权利要求书涵盖了所有含意与之一致的等效实现方式。

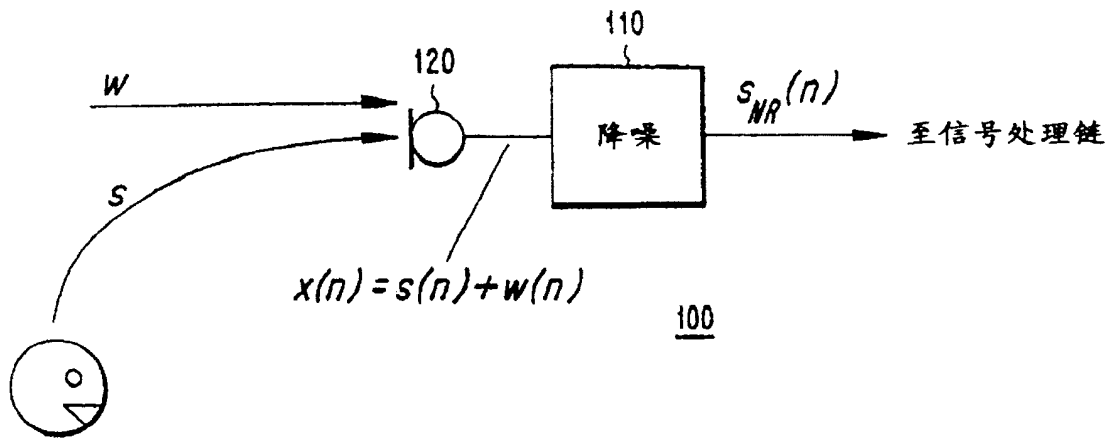


图 1

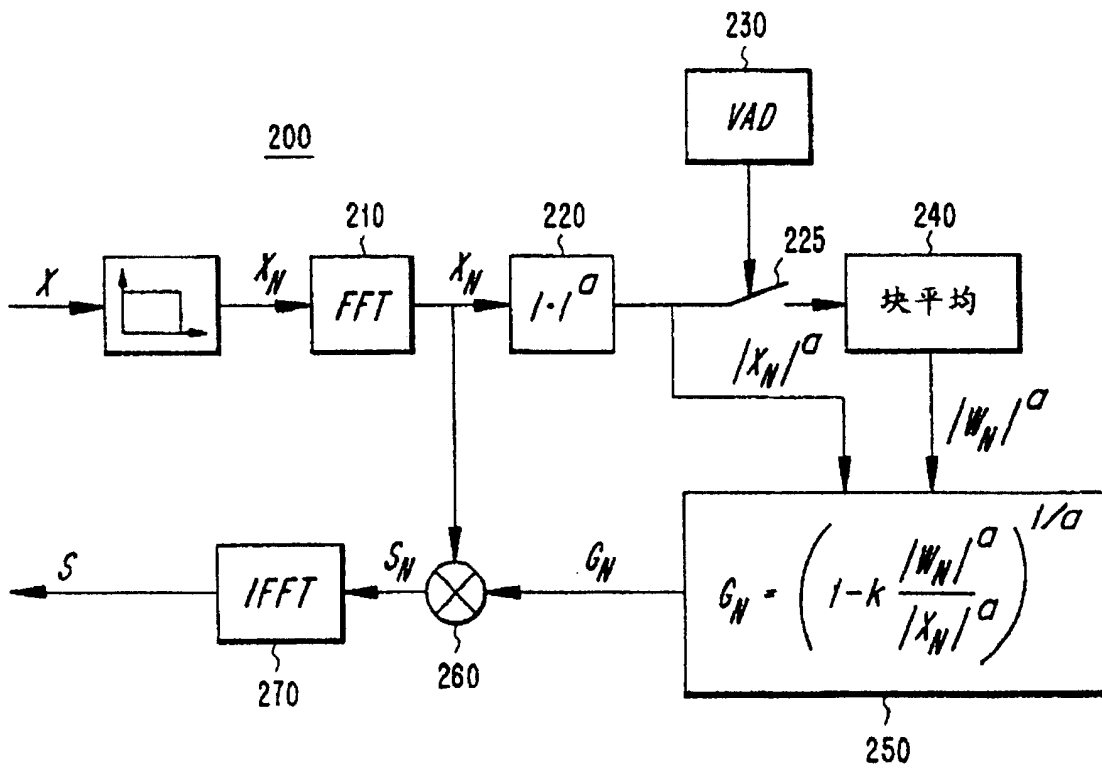


图 2

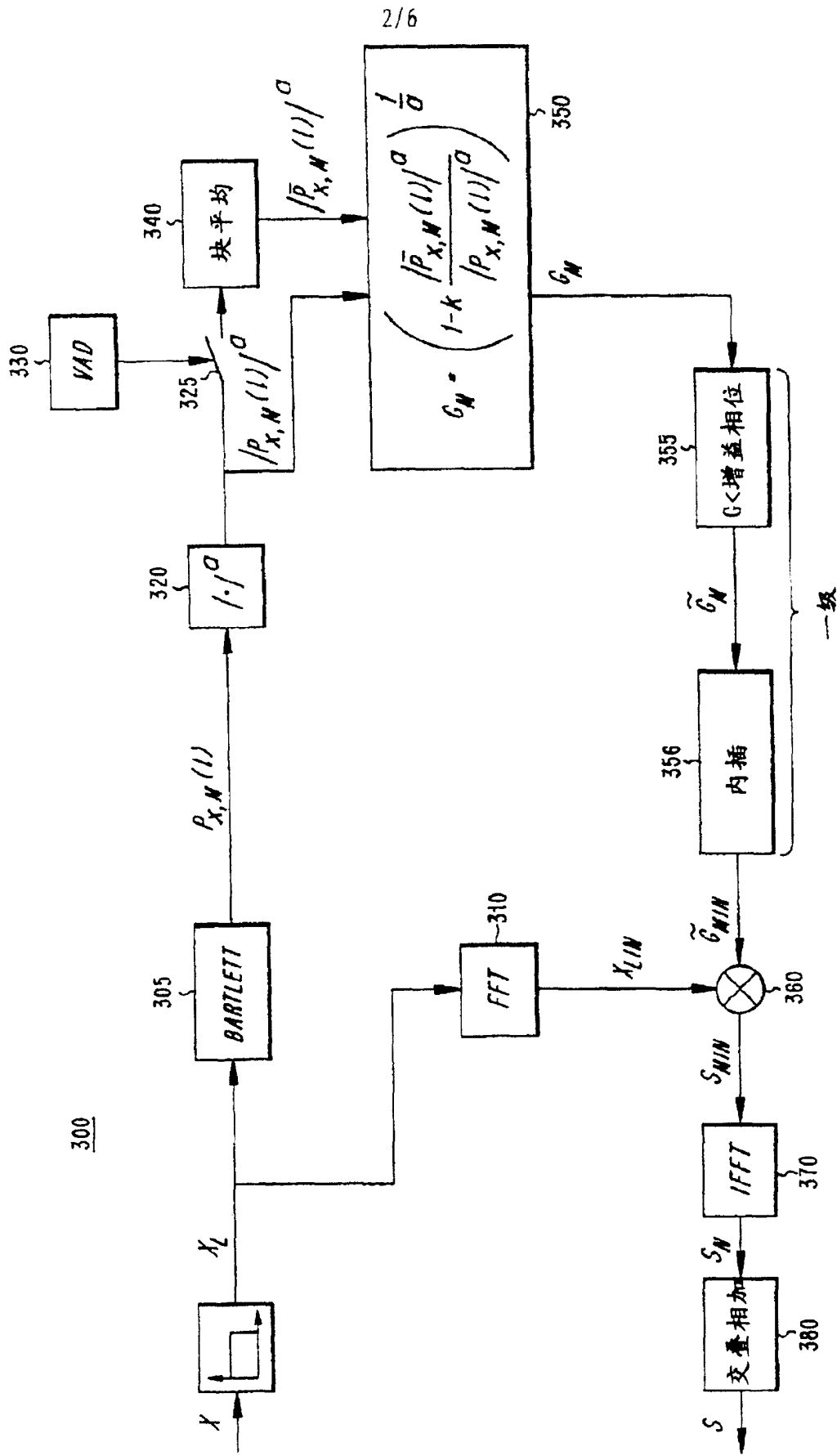


图 3

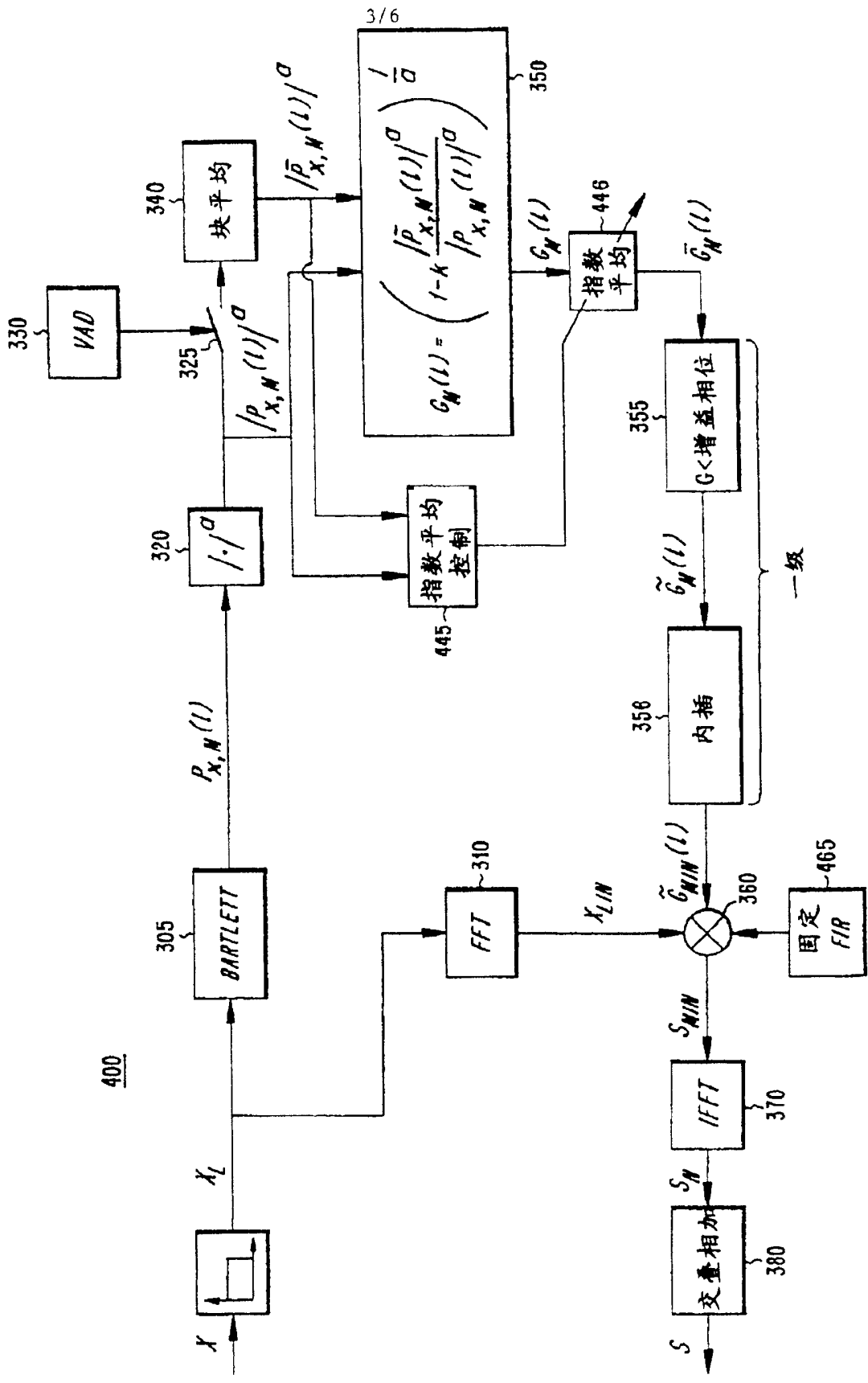


图 4

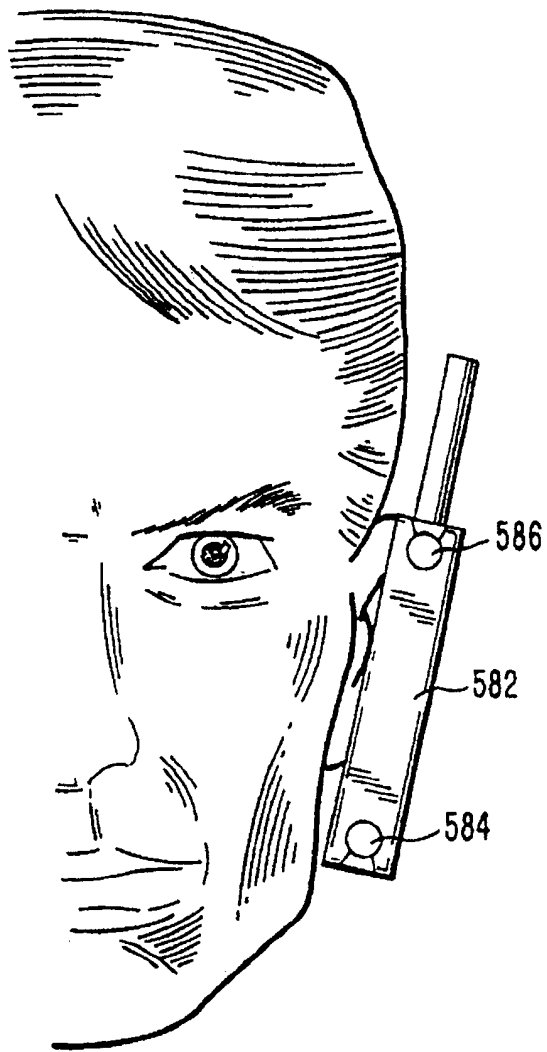


图 5

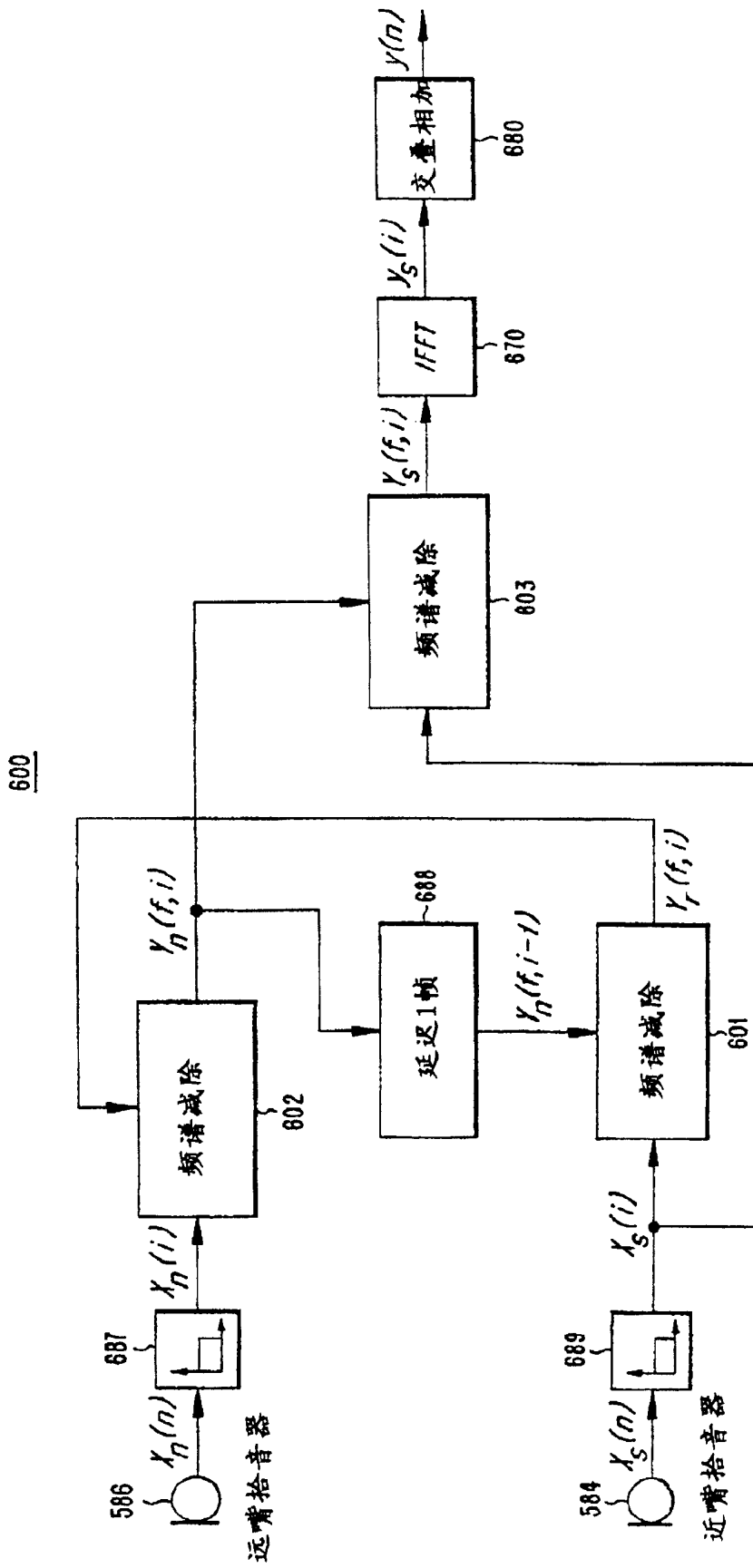


图 6

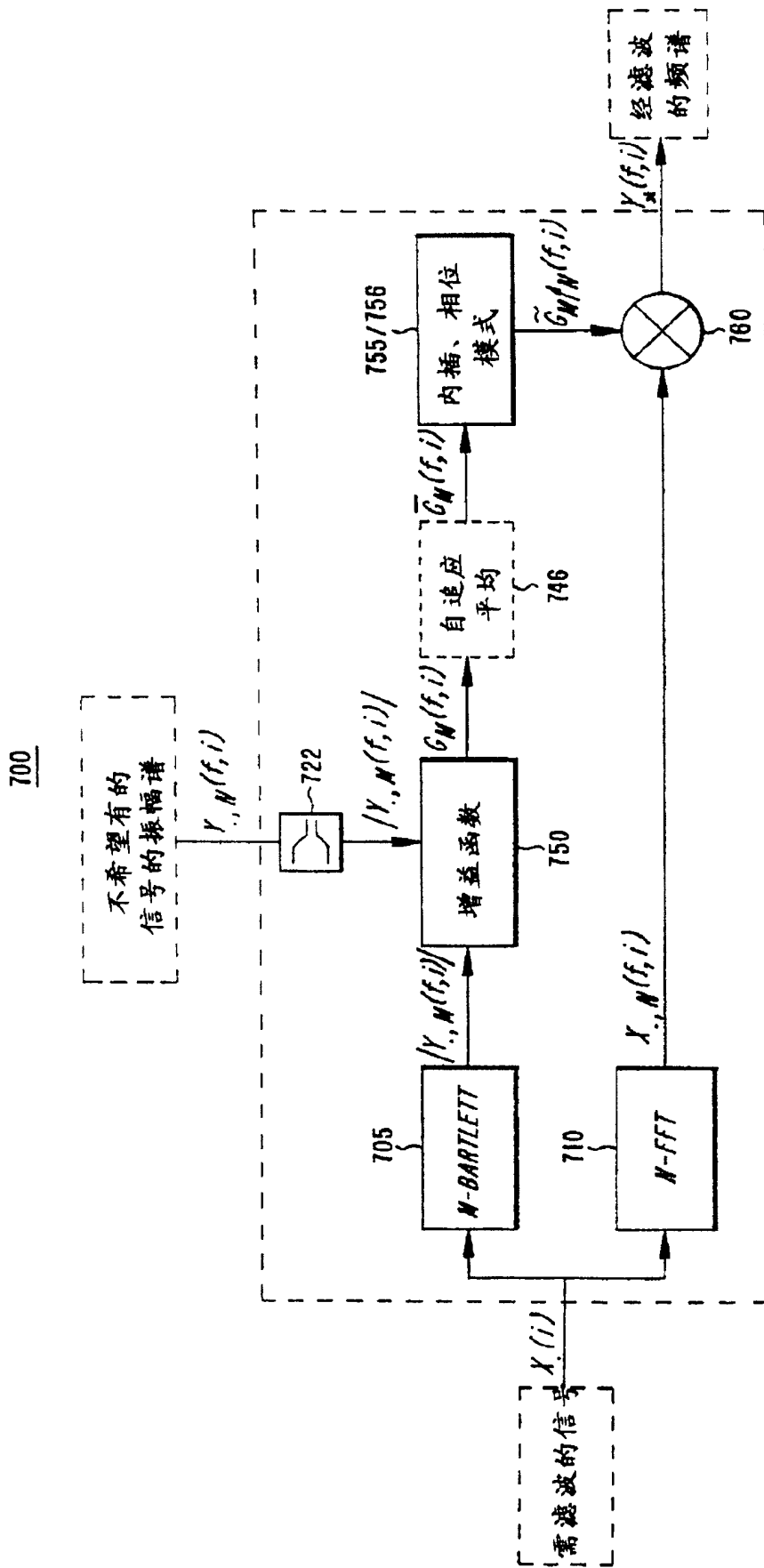


图 7