



(12) 发明专利

(10) 授权公告号 CN 101031963 B

(45) 授权公告日 2010.09.15

(21) 申请号 200580031069. X

(22) 申请日 2005.09.14

(30) 优先权数据

0409819 2004.09.16 FR

(85) PCT申请进入国家阶段日

2007.03.15

(86) PCT申请的申请数据

PCT/FR2005/002284 2005.09.14

(87) PCT申请的公布数据

W02006/032760 FR 2006.03.30

(73) 专利权人 法国电信

地址 法国巴黎

(72) 发明人 克洛德·马罗 西里尔·普拉普斯

帕斯卡尔·斯卡拉尔特

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 罗朋

(51) Int. Cl.

G10L 21/02(2006.01)

(56) 对比文件

CN 1276896 A, 2000.12.13, 说明书第3页第12行至第4页第12行、第6页第12行至第7页18行、第12页第25行至13页第3行及附图1-3.

GB 2398913 A, 2004.09.01, 摘要及附图3.

CN 1488136 A, 2004.04.07, 全文.

CN 1169788 A, 1998.01.07, 全文.

CN 1310840 A, 2001.08.29, 全文.

JP 2003280696 A, 2003.10.02, 说明书第0066-0093段及附图1.

CN 1311891 A, 2001.09.05, 全文.

审查员 黄素霞

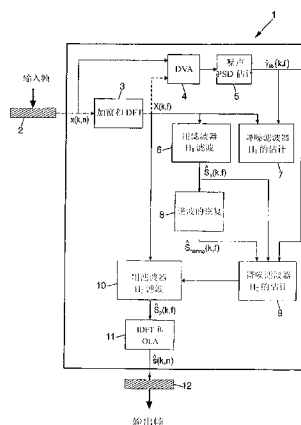
权利要求书 4 页 说明书 13 页 附图 5 页

(54) 发明名称

处理有噪声的声音信号的方法以及实现该方法的装置

(57) 摘要

本发明涉及处理有噪声的声音信号的方法以及实现该方法的装置。该方法包括以下对有噪声的声音信号的每个帧(x(k, n))执行的步骤,包括:实施一个频域变换;估计所述帧的噪声功率谱密度(PSD)(Y_{bb}(k, f));基于所估计的噪声功率谱密度以及与所述帧相应的所需信号的功率谱密度的估计(Y_{ss}(k, f)),计算第一降噪滤波器(H₁(k, f));利用所述计算的降噪滤波器对所述帧(x(k, n); X(k, f))进行滤波,以获得所述帧的第一降噪估计(S₁(k, f));以及,基于所述帧的第一噪音修正估计获得第二信号的一帧(S_{harmono}(k, f)),该第二信号的帧包括基本上与所述有噪声的声音信号的所述帧相对应的所需信号位于相同位置的谐波。



CN 101031963 B

1. 一种处理被组织为连续帧的有噪声的声音信号的方法,其中,该方法包括涉及所述连续帧中至少一个帧的如下步骤:

- 对所述有噪声的声音信号的所述帧实施一个使其转换至频域的变换;
- 估计所述帧的噪声功率谱密度;
- 基于所估计的噪声功率谱密度以及与所述帧相应的有用信号的功率谱密度的估计,计算第一降噪滤波器;以及

- 利用由上述计算得到的所述第一降噪滤波器对所述有噪声的声音信号的所述帧进行滤波,以获得所述帧的第一降噪估计;以及

- 基于所述有噪声的声音信号的所述帧的第一降噪估计获得第二信号的一帧,该第二信号的帧包括与所述有噪声的声音信号的所述帧相对应的有用信号位于相同位置的谐波。

2. 根据权利要求 1 所述的方法,其中,该方法进一步包括以下步骤:

- 基于所估计的噪声功率谱密度,以及所述帧的第一降噪估计的功率和所获得的第二信号的帧的功率的组合,计算第二降噪滤波器;

- 利用由上述计算得到的所述第二降噪滤波器对有噪声的声音信号的所述帧进行滤波,以获得所述帧的第二降噪估计;以及

- 合成所述帧的第二降噪估计。

3. 根据权利要求 2 所述的方法,其中,当利用所述第二降噪滤波器执行的滤波结果位于频域时,在合成所述帧的第二降噪估计之前对所述滤波结果实施一个使其转换至时域的变换。

4. 根据权利要求 1 至 3 中任一项所述的方法,其中,在对所述有噪声的声音信号的所述帧实施使其转换至频域的变换之前,将所述帧乘以一个加权窗。

5. 根据权利要求 4 所述的方法,其中,所述加权窗是汉宁窗,该窗的长度为所述有噪声的声音信号的所述帧的长度 L。

6. 根据权利要求 5 所述的方法,其中,所述加权窗的形式为:

$$w(n) = 0.5 - 0.5 \cdot \cos\left(\frac{\pi n}{L}\right)$$

其中,其中, n 是从 0 到 L-1 的整数, L 等于 256。

7. 根据权利要求 1 至 3 中任一项所述的方法,其中对所述有噪声的声音信号的所述帧实施的所述使其转换至频域的变换是长度为 512 的快速傅立叶变换。

8. 根据权利要求 1 至 3 中任一项所述的方法,其中,在所述有噪声的声音信号的所述帧中检测声音活动,且其中,当在所述帧中没有检测到声音活动时,所述帧的噪声功率谱密度估计与包含于所述帧中的噪声功率谱密度一致;当在所述帧中检测到声音活动时,所述帧的噪声功率谱密度估计与包含在所述有噪声的声音信号的所述帧之前的至少一帧中的噪声功率谱密度一致。

9. 根据权利要求 8 所述的方法,其中,包含于所述有噪声的声音信号的索引为 k_b 且在其中未检测到声音活动的至少一帧中的噪声的功率谱密度受形式为:

$$\hat{\gamma}_{bb}(k_b, f) = \alpha \hat{\gamma}_{bb}(k_b - 1, f) + (1 - \alpha) |X(k_b, f)|^2$$

的指数平滑的影响,其中 $X(k_b, f)$ 表示将所述有噪声的声音信号的索引为 k_b 的帧转换

到频域的变换, α 为平滑量。

10. 根据权利要求 1 至 3 中任一项所述的方法, 其中, 对所述第一降噪滤波器的计算包括实现短期频谱衰减方法的第一轮计算。

11. 根据权利要求 10 所述的方法, 其中, 对所述第一降噪滤波器的计算包括实现具有形如:

$$\hat{H}_{step1}(k, f) = \frac{\eta(k, f)}{1 + \eta(k, f)}$$

的传递函数的开环维纳滤波器的第一轮计算, 其中 k 是所述有噪声的声音信号的所述帧的索引, $\eta(k, f)$ 表示与所述帧对应的有用信号的功率谱密度的估计与噪声的功率谱密度的估计之比。

12. 根据权利要求 11 所述的方法, 其中, 与所述帧相对应的有用信号的功率谱密度 $\hat{\gamma}_{ss}(k, f)$ 的估计包括:

形式为 $\hat{\gamma}_{ss}(k, f) = \beta |\hat{S}(k-1, f)|^2 + (1-\beta) P[|X(k, f)|^2 - \hat{\gamma}_{bb}(k, f)]$ 的直接判决估计, 其中, β 为重心参数, $\hat{S}(k-1, f)$ 是与索引为 $k-1$ 的前一个帧相关的被估计的有用信号的频谱, $X(k, f)$ 表示将所述有噪声的声音信号的索引为 k 的所述帧转换至频域的变换, $\hat{\gamma}_{bb}(k, f)$ 是估计的噪声功率谱密度, P 由以下公式定义: $P[z(k, f)] = \begin{cases} z(k, f) & \text{若 } z(k, f) > 0; \\ 0 & \text{否则} \end{cases}$;

或包括:

指数平滑估计。

13. 根据权利要求 10 所述的方法, 其中, 所述对第一降噪滤波器的计算进一步包括实现短期频谱衰减方法的第二轮计算, 其中与所述帧对应的有用信号的功率谱密度估计考虑了在所所述第一轮计算中执行的计算。

14. 根据权利要求 13 所述的方法, 其中, 考虑了在所所述第一轮计算中执行的计算的、与所述帧对应的有用信号的功率谱密度估计 $\hat{\gamma}_{ss2}(k, f)$ 的形式为:

$$\hat{\gamma}_{ss2}(k, f) = |\hat{H}_{step1}(k, f)X(k, f)|^2$$

其中, $\hat{H}_{step1}(k, f)$ 表示在所所述第一轮计算中被计算的传递函数, $X(k, f)$ 表示将所述有噪声的声音信号的、索引为 k 的所述帧转换至频域的变换,

且其中, 对所述第一降噪滤波器的计算包括实现具有形如:

$$\hat{H}_1(k, f) = \frac{\eta_2(k, f)}{1 + \eta_2(k, f)}$$

的传递函数的开环维纳滤波器的第二轮计算,

其中, $\eta_2(k, f)$ 表示考虑了在所所述第一轮计算中执行的计算的、与所述帧对应的有用信号的功率谱密度的估计与噪声的功率谱密度的估计之比。

15. 根据权利要求 1 至 3 中任一项所述的方法, 其中, 利用所述第一降噪滤波器对所述有噪声的声音信号的所述帧的滤波在将所述帧转换至频域的变换上执行。

16. 根据权利要求 1 至 3 中任一项所述的方法, 其中, 利用所述第一降噪滤波器对所述有噪声的声音信号的所述帧的滤波包括所述第一降噪滤波器的系数选择和所述第一降噪

滤波器的时域冲激响应的加权。

17. 根据权利要求 16 所述的方法,其中,所述第一降噪滤波器的 256 个系数被选择,并且被汉宁窗加权。

18. 根据权利要求 17 所述的方法,其中,所述对所述有噪声的声音信号的所述帧实施的使其转换至频域的变换是长度为 512 的快速傅立叶变换,其中,所述第一降噪滤波器的时域冲激响应被填充 256 个零。

19. 根据权利要求 1 至 3 中任一项所述的方法,其中,获取包括与所述有噪声的声音信号的所述帧相对应的有用信号位于相同位置的谐波的第二信号的一帧的步骤包括:当所述有噪声的声音信号的所述帧的第一降噪估计位于时域时,将一个非线性函数应用于所述有噪声的声音信号的所述帧的第一降噪估计,且当所述有噪声的声音信号的所述帧的第一降噪估计位于频域时,将圆周卷积应用于所述有噪声的声音信号的所述帧的第一降噪估计与将非线性函数转换至频域的变换之间。

20. 根据权利要求 19 所述的方法,其中,所述非线性函数是以下各项中的任一个:单波整流函数、绝对值、所述有噪声的声音信号的所述帧的第一降噪估计和一阈值中的最大值、所述有噪声的声音信号的所述帧的第一降噪估计和一阈值中的最小值。

21. 根据权利要求 2 至 3 中任一项所述的方法,其中,所述第二降噪滤波器的计算实现一个短期频谱衰减方法。

22. 根据权利要求 21 所述的方法,其中,所述第二降噪滤波器的计算实现具有形如:

$$\hat{H}_2(k, f) = \frac{\eta_{harmono}(k, f)}{1 + \eta_{harmono}(k, f)}$$

的传递函数的开环维纳滤波器,其中,

$$\eta_{harmono}(k, f) = \frac{\hat{\gamma}_{ss_harmono}(k, f)}{\hat{\gamma}_{bb}(k, f)}, \hat{\gamma}_{ss_harmono}(k, f) = \rho(k, f) |\hat{S}_1(k, f)|^2 + (1 - \rho(k, f)) |\hat{S}_{harmono}(k, f)|^2$$

其中, $\hat{\gamma}_{bb}(k, f)$ 是估计的噪声功率谱密度, $\hat{S}_1(k, f)$ 表示索引为 k 的所述帧的所述第一降噪估计, $\hat{S}_{harmono}(k, f)$ 表示所述第二信号的帧, $\rho(k, f)$ 为所述第二信号的重注入参数。

23. 根据权利要求 22 所述的方法,其中, $\rho(k, f)$ 为时间和 / 或频率的函数。

24. 根据权利要求 2 至 3 中任一项所述的方法,其中,利用所述第二降噪滤波器对所述有噪声的声音信号的所述帧的滤波在将所述帧转换至频域的变换上执行。

25. 根据权利要求 2 至 3 中任一项所述的方法,其中,利用所述第二降噪滤波器对所述有噪声的声音信号的所述帧的滤波包括所述第二降噪滤波器的系数选择和所述第二降噪滤波器的时域冲激响应的加权。

26. 根据权利要求 25 所述的方法,其中,所述第二降噪滤波器的 256 个系数被选择,并且被汉宁窗加权。

27. 根据权利要求 26 所述的方法,其中,所述对所述有噪声的声音信号的所述帧实施的使其转换至频域的变换是长度为 512 的快速傅立叶变换,其中,所述第二降噪滤波器的时域冲激响应被填充 256 个零。

28. 根据权利要求 2 至 3 中任一项所述的方法,其中,所述帧的第二降噪估计的合成使用块重叠相加或者块重叠保存。

29. 一种用于处理被组织为连续帧的有噪声的声音信号的装置 (1), 其中, 包括涉及所

述连续帧中至少一个帧的如下部分：

- 用于对所述有噪声的声音信号的所述帧实施一个使其转换至频域的变换的分析单元；

- 用于估计所述帧的噪声功率谱密度的功能块；

- 用于基于所估计的噪声功率谱密度以及与所述帧相应的有用信号的功率谱密度的估计，计算第一降噪滤波器的功能块；

- 用于利用由上述计算得到的所述第一降噪滤波器对所述有噪声的声音信号的所述帧进行滤波，以获得所述帧的第一降噪估计的功能块；与

- 用于基于所述有噪声的声音信号的所述帧的第一降噪估计获得第二信号的一帧的功能块，该第二信号的帧包括与所述有噪声的声音信号的所述帧相对应的有用信号位于相同位置的谐波。

处理有噪声的声音信号的方法以及实现该方法的装置

技术领域

[0001] 本发明涉及对有噪声的声音信号的处理,尤其涉及降低上述信号中存在的噪声。

背景技术

[0002] 降低一个声音信号中的噪声,即干扰信号的技术是已知的。该技术的目标在于考虑声音信号所处的声学环境,以提高信号的质量和清晰可辨程度。这些技术包括通过对有噪声的信号进行处理来从声音信号中提取有用的信息。上述技术应用于,例如语音通信,声音信号在多个通话者之间传输的通话应用中,例如远程电话 (telephony)、电话会议、视频会议。此外,它们还用于在嘈杂的环境中进行信号接收的应用,或者声音识别,而当声音在嘈杂的环境中发出时,声音识别的性能会有很大的改变。

[0003] 这些技术通常包括估计一个降噪滤波器 (noise reduction filter) 的传递函数 (transfer function),然后基于频谱域 (spectral domain) 的乘法执行滤波处理。它们所运用的方法被定义为“通过短期频谱衰减 (short-term spectral attenuation) 的降噪 (noise reduction)”。

[0004] 根据这些技术,声音信号 $x(n)$ 包含有用信号成分 $s(n)$ 和噪声成分 $b(n)$, n 代表离散时间的索引。然而,需要注意的是,也可以采用连续时间的信号表示。信号 $x(n)$ 被组织为有固定长度、索引为 k 的连续帧 $x(n,k)$ 。这些帧中的每一个首先被乘以一个加权窗,以使得可以提高随后对计算降噪滤波器所必需的频谱量 (spectral quantities) 的估算的准确度。于是,每个与加权窗相乘的帧在频谱域被分析,例如借助离散傅立叶变换或快速傅立叶变换。这个操作被称为短期傅立叶变换 (STFT, short-term Fourier transformation)。

[0005] 因此,所得到的被处理信号的频域表示 $X(k, f)$ 使得在同一个时间可以估计该降噪滤波器的传递函数 $H(k, f)$,并通过该传递函数与该有噪声的信号的短期频谱的简单相乘将该传递函数应用于频谱域,其中 f 是频率索引。所述滤波的结果可以写成如下形式:

$$[0006] \quad \hat{S}(k, f) = H(k, f)X(k, f)。$$

[0007] 该所得到的信号通过逆频域变换转换回时域。通过块重叠相加 (block overlap and add) 技术 (OLA, 表示“重叠相加 (overlap add)”),或者通过块保留技术 (OLS 表示“重叠保留 (overlap save)”) 来最终合成相应的时间信号。这个在时域中重建信号的操作被称为逆短期傅立叶变换 (ISTFT, inverse short-term Fourier transformation)。

[0008] 短期频域衰减方法的详细描述可以在以下参考文献中找到:“Enhancement and bandwidth compression of noisy speech”, IEEE 学报,1979 年,第 67 卷 1586-1604 页, J. S. Lim、A. V. Oppenheim 著,和“Multirate digital signal processing”, Prentice Hall, 1983 年, R. E. Crochiere、L. R. Rabiner 著。

[0009] 在索引 k 的时间段上借助频率成分 (frequency component) f 应用于被处理的信号 $X(k, f)$ 的短期频谱衰减 $H(k, f)$ 通常是基于本地的信噪比 $\text{SNR}(k, f)$ 的估计来决定的。所有抑制规则 (suppression rules) 的一个共同的特点在于如下所示的渐进性:

$$[0010] \quad H(k, f) \approx 1, \text{SNR}(k, f) \gg 1$$

[0011] $H(k, f) \approx 0, \text{SNR}(k, f) \ll 1$

[0012] 在大多数的方法中,存在如下假设:噪声和有用信号统计不相关,有用信号是间断的(存在无声(silence)期间),且人耳对于信号的相位(通常未通过处理而被修正)不敏感。

[0013] 通常采用的抑制规则可以通过例子来引证:功率谱减法(powerspectral subtraction)、幅度谱减法(amplitude spectral subtraction)和直接采用维纳滤波器(Wiener Filter)。对于这些规则,有用语音信号的频域成分 f 的短期估计可以分别写为:

[0014] 对于功率谱减法(参见前述 J. S. Lim 和 A. V. Oppenheim 所著文献)为:

$$[0015] \quad \hat{S}_{SSP}(k, f) = \sqrt{\frac{\gamma_{ss}(k, f)}{\gamma_{ss}(k, f) + \gamma_{bb}(k, f)}} X(k, f) \quad (1)$$

[0016] 对于幅度谱减法(参见“Suppression of acoustic noise in speech usingspectral subtraction”, IEEE Trans. on Audio, Speech and Signal Processing,, 1979年4月,第27卷第2期,113-120页, S. F. Boll 著)为:

$$[0017] \quad \hat{S}_{SSA}(k, f) = \left[1 - \sqrt{\frac{\gamma_{bb}(k, f)}{\gamma_{ss}(k, f) + \gamma_{bb}(k, f)}} \right] X(k, f) \quad (2)$$

[0018] 对于维纳滤波器(见前述 J. S. Lim 和 A. V. Oppenheim 的文章)为:

$$[0019] \quad \hat{S}_{Wiener}(k, f) = \frac{\gamma_{ss}(k, f)}{\gamma_{ss}(k, f) + \gamma_{bb}(k, f)} X(k, f) \quad (3)$$

[0020] 在这些表达式中, $\gamma_{ss}(k, f)$ 和 $\gamma_{bb}(k, f)$ 分别表示在索引 k 的时间窗上被处理信号 $X(k, f)$ 的频域成分 f 之中的有用信号和噪声的功率谱密度(powerspectral density)。

[0021] 基于上述表达式,有可能研究应用于有噪声的信号的频谱衰减的性能,其作为在给定频域成分 f 上所测量的本地信噪比的函数。图1中画出了上面提到的三个短期抑制规则的曲线。可能要注意的是,当本地信噪比显著时(图1中的右手部分),三个规则给出了基本相同的衰减。在所述处理过程的输出时,最佳功率减法规则(optimal powersubtraction rule)的噪声功率仍然是最大的,这里,最佳是指对高斯模型的最大似然而言(参考“Elimination of the musical noise phenomenon with the Ephraim andMalah noise suppressor”, IEEE Trans. on Speech and Audio Processing, 1994年4月,第2卷第2期, 345-349页, 0. Cappé 著)。对于所述三个抑制规则,我们注意到有关截止值(cutoff value)的本地信噪比的一个小变化就足以从全部衰减($H(k, f) \approx 0$)的情形切换到可忽略的谱修正(spectralmodification) ($H(k, f) \approx 1$)的情形。

[0022] 上述后一个属性构成了“音乐噪声”现象的原因之一。具体地,同时包括确定性和随机成分的环境噪声仅在没有声音活动的时候被特征化。考虑到随机成分的存在,在有声音活动时噪声的频率成分 f 的实际贡献与其在没有声音活动的瞬间基于多个帧所作的平均估计之间存在很大不同。考虑到这个不同,本地信噪比的估计在截止位置可能有波动,因此在所述处理过程的输出时,使得出现然后消失、且平均寿命统计上不超过所考虑的分析窗的数量级的频谱成分上升。这种特征在整个通带(passband)上是普遍的,引入了人耳能够听见并且恼人的残余噪声(residual noise)。

[0023] 已有研究来试图降低这种残余噪声的干扰。所提出的解决方案包括以下几种途径:平均化短期估计(参考前述的 S. F. Boll 所著文献)、噪声功率谱的过估计(参

见“Enhancement of speech corrupted by acoustic noise”, Int. Conf. on Speech, Signal Processing, 1979 年, 第 208-211 页, M. Berouti 等著, 以及“Experiments with a non-linear spectral subtractor, hidden Markov models and the projection for robust speech recognition in cars”, Proc. of EUSIPCO' 91, 1991 年, 第 79-82 页, P. Lockwood、J. Boudy 著) 或者跟踪噪声谱密度的最小值(参见“Spectral subtraction based on minimum statistics”, in Signal Processing VII: Theories and Applications, EUSIPCO' 94, 1994 年 9 月, 第 1182-1185 页, R. Martin 著)。

[0024] 一种相对有效的抑制音乐噪声的解决方案包括称作“直接判决 (directed-decision)”的有用信号 (useful signal) 的功率谱密度估计器(参见“Speech enhancement using a minimum mean square error short-time spectral amplitude estimator”, IEEE Trans. on Audio, Speech and Signal Processing, 1984 年, 第 32 卷第 6 期, 第 1109-1121 页, Y. Ephraim、D. Malah 著, 以及上述 O. Cappé 所著文献)。该估计器实现有用信号的瞬时功率谱密度和长期功率谱密度的折中, 从而能够有效地消除音乐噪声。更进一步, 众所周知, 通过调整这个估计器的固有的延迟可提高这个解决方案的效果(参见专利文献 FR2820227 以及文献“A Two-Step Noise Reduction Technique”, ICASSP, 2004 年 5 月, C. Plapous、C. Marro、L. Mauuary、P. Scalart 著)。

[0025] 已有多个研究基于语音和加性噪声信号 (added noise signal) 的统计模型来建立新的抑制规则。由于这些研究同传统的方法相比具有一个额外的自由度, 使得引入一种称作“软判决”的新算法成为可能(参见“Speech enhancement using a soft-decision noise suppression filter”, IEEE trans. on Audio, Speech and Signal Processing, 1980 年 4 月, 第 28 卷, 第 2 期, 第 137-145 页, R. J. Mac Aulay、M. L. Malpass 著, “Speech enhancement using optimal non-linear spectral amplitude estimation”, Int. Conf. on Speech, Signal Processing, 1983 年, 第 1118-1121 页, 以及上述 Y. Ephraim、D. Malah 所著文献“Speech enhancement using a minimum mean square error short-time spectral amplitude estimator”)。

[0026] 如上所述, 短期频谱衰减的计算依赖于每个频谱成分的信噪比估计。例如, 上面给出的每一个公式都涉及到下述数值:

$$[0027] \quad SNR(k, f) = \frac{\gamma_{ss}(k, f)}{\gamma_{bb}(k, f)}$$

[0028] 因此, 降噪方法的性能, 尤其是在失真和噪声水平的有效降低方面, 受信噪比估计器的相关性的控制。

[0029] 这个缺陷构成了现有语音降噪系统的最大限制。具体地, 目前的降噪系统不能够对很低信噪比的谐波进行降噪。在实践中, 降噪算法使用 SNR 来检测每个频率的语音成分 (speech component) 是否存在。如果所估计的 SNR 很不理想, 则认为没有信号成分并抑制它。因此, 尽管根据先验知识知道谐波肯定存在, 谐波还是有可能被现有的降噪系统破坏。现在, 需要注意的是, 在大多数语言中, 浊音 (voiced sounds) (谐波) 表示所发出声音的非常大的一部分。

发明内容

[0030] 本发明的一个目的是克服现有降噪系统的限制。

[0031] 本发明的另一个目的是提高降噪方法的性能。

[0032] 本发明的另一个方面是提供一种不使信号过分失真的声音信号处理方法。特别是,所执行的信号处理使得保存包含于该信号中的所有或者部分谐波成为可能。

[0033] 本发明的另一个目的是在完成声音信号处理时限制音乐噪声的出现。

[0034] 本发明的另一个目的是获得有用信号的梳状谐波 (harmonic comb) 的一个好的估计。

[0035] 因此,本发明提供了一种处理被组织为连续帧的有噪声的声音信号的方法。该处理包括涉及所述帧中至少一个帧的如下步骤:

[0036] - 对所述有噪声的声音信号的所述帧实施一个使其转换至频域的变换;

[0037] - 估计所述帧的噪声功率谱密度;

[0038] - 基于所估计的噪声功率谱密度以及与所述帧相应的有用信号的功率谱密度的估计,计算第一降噪滤波器;以及

[0039] - 利用由上述计算所得到的所述第一降噪滤波器对所述有噪声的声音信号的所述帧进行滤波,以获得所述帧的第一降噪估计;以及

[0040] - 基于所述有噪声的声音信号的所述帧的第一降噪估计获得第二信号的一帧,该第二信号的帧包括基本上与所述有噪声的声音信号的所述帧相对应的有用信号位于相同位置的谐波。

[0041] 于是,这样的有噪声的语音信号的处理使得应用于所述信号的第一降噪滤波器在输出时获得调和性的再生 (regeneration of harmonicity)。因而所获得的所述第二信号的帧被构建,以消除在第一滤波时可能出现的降噪信号帧的失真。

[0042] 该处理使得获得有用信号的梳状谐波的一个好的估计成为可能,该有用信号可能是随后应用的主体。

[0043] 优选地,在降噪处理中,该方法进一步包括以下步骤:

[0044] - 基于所估计的噪声功率谱密度,以及所述帧的第一降噪估计和所获得的第二信号的帧的功率的结合,计算第二降噪滤波器;

[0045] - 利用由上述计算得到的所述第二降噪滤波器对有噪声的声音信号的所述帧进行滤波,以获得所述帧的第二降噪估计;以及

[0046] - 合成所述帧的第二降噪估计。

[0047] 在本实施例中,由于第二降噪滤波器被来自调和性再生的信号所驱动,该第二降噪滤波器得以得以保留谐波的方式计算。因此,在第二滤波完成时获得的所述帧的第二降噪估计比在传统降噪系统中获得的降噪估计要更有效,因为在传统降噪系统中输入信号的谐波被破坏或者被减弱。

[0048] 第一降噪滤波器的计算与适当时第二降噪滤波器的计算可包括实现短期频谱衰减的第一轮计算 (first pass)。例如,可引用以下方法作为示例:功率谱减法、幅度谱减法和开环维纳滤波器 (open-loop Wiener filter) 等。优选地,还可包括实现短期频谱衰减的第二轮计算 (second pass),例如功率谱减法、幅度谱减法和开环维纳滤波器等,并在考虑第一轮计算中所执行计算的情况下使用与所述帧相应的有用信号的功率谱密度的估计。

[0049] 优选地,获取包括基本上与所述有噪声的声音信号的所述帧相对应的有用信号中基本上位于相同位置的谐波的第二信号的帧的步骤包括:当所述第一估计位于时域时是一

个非线性函数应用,其应用于所述有噪声的声音信号的所述帧的所述第一降噪估计,且当所述第一估计位于频域时是圆周卷积(circular convolution)应用,其应用于所述有噪声信号的所述帧的所述第一降噪估计与将非线性函数转换至频域的变换间。

[0050] 例如,所述非线性函数可以是以下各项中的任一个:单波整流函数(single-wave rectification function)、绝对值、有噪声的声音信号的所述帧的所述第一降噪估计和阈值中的最大者;有噪声的声音信号的所述帧的所述第一降噪估计和阈值中的最小者。

[0051] 本发明进一步提供了一种处理有噪声的声音信号的装置,包括用于实现如上所述方法的装置。

[0052] 本发明也提供了一个位于信息媒介中的计算机程序,其特征在于,当该程序被加载并被计算装置执行时,该程序包括用于实现上述方法的指令。

附图说明

[0053] 通过阅读以下参照附图对非限制性实施例所作的详细描述,本发明的其它特征和优点将会变得更明显。

[0054] - 图 1 为三种已知的短期抑制规则的性能比较图,正如已经说明的;

[0055] - 图 2 为能够实现本发明的一个实施例的各个功能块的示意图;

[0056] - 图 3 为恢复降噪信号中丢失的谐波的例子的示意图;

[0057] - 图 4 为根据本发明的一个实施例信号中的谐波恢复的效果示意图;以及

[0058] - 图 5 为根据本发明的一个实施例第二降噪滤波器估计的例子示意图。

具体实施方式

[0059] 图 2 示出了根据本发明的一个实施例的装置 1。有噪声的声音信号的当前帧 2 进入装置 1 以经历其中的处理过程。根据上文引入的符号,该帧被表示为 $x(k, n)$ 。

[0060] 装置 1 包括分析单元 3,其用于将有噪声的信号的当前帧 $x(k, n)$ 转换到频域 $X(k, f)$ 。首先,该帧被乘以加权窗 $w(n)$,以提供信号 $x_w(k, n)$:

[0061] $x_w(k, n) = w(n)x(k, n)$ 。

[0062] 然后,利用离散傅立叶变换(DFT)执行到频域的转换,即:

[0063] $X(k, f) = \text{DFT}((x_w(k, n)))$

[0064] 优选地,所述 DFT 可通过快速傅立叶变换(或 FFT)来实现。然而,到频域的其他变换,如小波变换也是可以的。对于逆操作同样如此,即后面将要提到的用于转换回时域的逆离散傅立叶变换(IDFT)。

[0065] 此外,优选地,可在当前帧 2 上执行用于检测声音活动(DVA, detectingvocal activity)的功能 4。DVA 使得能够知晓什么时候更新噪声的功率谱密度(PSD)。因而,对 DVA 检测到的每个“只有噪声”的帧 k_b ,功能块 5 根据下述递归表达式估计噪声功率谱密度 $\hat{\gamma}_{bb}(k_b, f)$:

[0066] $\hat{\gamma}_{bb}(k_b, f) = \alpha(k_b)\hat{\gamma}_{bb}(k_b - 1, f) + (1 - \alpha(k_b))|X(k_b, f)|^2$

[0067] 参数 $\alpha(k_b)$ 控制平滑因子(smoothing factor),其随着时间的变化而变化。

[0068] 如果检测到当前帧 2 不是噪声帧,则噪声的功率谱密度估计被标记,即,保持所获得的前一个噪声帧的 PSD 的最后的值。

[0069] 需要注意的是, $\hat{\gamma}_{bb}(k_b, f)$ 的估计不限于这个指数平滑估计器, 可使用任何其它的功率谱密度估计器。

[0070] 可以使用任何类型的 DVA, 不管后者在时域还是在频域进行操作。甚至也可以省略这样的声音活动检测。

[0071] 然后通过图 2 中的功能块 7 来实现第一降噪滤波器的估计。优选地, 利用两轮计算方法在频域估计该第一降噪滤波器的传递函数 (参见专利文献 FR2820227, 以及文献 “A Two-Step Noise Reduction Technique”, ICASSP, 2004 年 5 月, C. Plapous, C. Marro, L. Mauuary, P. Scalart 著)。在第一轮计算中, 计算具有如下形式:

$$[0072] \quad \hat{H}_{step1}(k, f) = f_{step1}(\hat{\gamma}_{ss}(k, f), \hat{\gamma}_{bb}(k, f))$$

[0073] 的传递函数的滤波器。

[0074] 上述滤波器的传递函数的表达式是一个依赖于有用信号的功率谱密度 $\hat{\gamma}_{ss}(k, f)$ 和噪声的功率谱密度 $\hat{\gamma}_{bb}(k, f)$ 的数学表达式。据此, 可根据短期频谱衰减方法来选择函数 f_{step1} , 以便 $\hat{H}_{step1}(k, f)$ 实现如功率谱减法规则 (公式 (1))、幅度谱减法规则 (公式 (2))、开环维纳滤波器规则 (公式 (3)) 等。可采用其他任何基于频率的噪声抑制规则来估计 $\hat{H}_{step1}(k, f)$ 。

[0075] 上面详细描述了噪声功率谱密度 $\hat{\gamma}_{bb}(k_b, f)$ 的计算。由于在声音活动期间信号和噪声的混合, 故不能直接获得所涉及的频谱量 $\hat{\gamma}_{ss}(k, f)$ 。为了计算该频谱量, 根据以下表达式使用直接判决的估计器 (参见前述文献 “Speechenhancement using a minimum mean square error short-time spectral amplitude estimator”, Y. Ephraim, D. Malah 著):

$$[0076] \quad \hat{\gamma}_{ss}(k, f) = \beta(k) \left| \hat{S}(k-1, f) \right|^2 + (1 - \beta(k)) P \left[\left| X(k, f) \right|^2 - \hat{\gamma}_{bb}(k, f) \right]$$

[0077] 其中, $\beta(k)$ 是能够随着时间变化而变化的重心参数 (barycentric parameter), $\hat{S}(k-1, f)$ 是与索引为 $k-1$ 的前一个帧相关的被估计的有用信号的频谱。确保数值 $\left(\left| X(k, f) \right|^2 - \hat{\gamma}_{bb}(k, f) \right)$ 阈值的函数 P 如下述所示, 该数值在估计错误的情形下有可能为负值:

$$[0078] \quad P[z(k, f)] = \begin{cases} z(k, f) & z(k, f) > 0 \\ 0 & \end{cases}$$

[0079] 需要注意的是, $\hat{\gamma}_{ss}(k, f)$ 的估计不限于该直接判决的估计器。特别地, 也可以使用指数平滑估计器或任何其它功率谱密度估计器。

[0080] 此后, 可重用传递函数 $\hat{H}_{step1}(k, f)$ 来改进有用信号的 PSD $\hat{\gamma}_{ss}(k, f)$ 的估计。然后我们获得由以下公式给出的数值 $\hat{\gamma}_{ss2}(k, f)$:

$$[0081] \quad \hat{\gamma}_{ss2}(k, f) = \left| \hat{H}_{step1}(k, f) X(k, f) \right|^2$$

[0082] 第二轮计算包括基于 $\hat{\gamma}_{ss2}(k, f)$ 计算所述第一降噪滤波器的传递函数的估计值 $\hat{H}_1(k, f)$, 即:

$$[0083] \quad \hat{H}_1(k, f) = f_{step2}(\hat{\gamma}_{ss2}(k, f), \hat{\gamma}_{bb}(k, f))$$

[0084] 与第一降噪滤波器 $\hat{H}_1(k, f)$ 的两个优点相比, 两轮计算的特别之处在于有用信号的 PSD $\hat{\gamma}_{ss}(k, f)$ 的“更快速”更新。一方面, 获得有用信号的非平稳性 (non-stationarities)

的快速跟踪,尤其在其时域包络 (temporal envelope) 快速变换的期间 (例如,在无声 / 语音转换时语音信号的冲击或者消失),另一方面,降噪滤波器被更好地估计,该方法性能的提高证明了这一点。

[0085] $\hat{H}_{step1}(k, f)$ 和 $H_1(k, f)$ 可执行短期频谱衰减的方法,例如功率谱减法规则 (公式 (1))、幅度谱减法规则 (公式 (2)) 或者开环维纳滤波规则 (公式 (3))。也可执行其它任何基于频率的噪声抑制规则来估计 $\hat{H}_{step1}(k, f)$ 和 $H_1(k, f)$ 。当然,也可以仅限于第一轮频谱衰减而无需执行第二轮频谱衰减。

[0086] 装置 1 的功能块 6 根据被计算的第一滤波器 $H_1(k, f)$ 来执行滤波。其可以在频域执行,此时对应于两个频谱相乘。就目前来看,这等效于在时域的圆周卷积操作。因而有必要采取一些预防措施以避免由于时域伪信号 (temporal aliasing) 引起的失真,当随着帧的速度通过点击侦听时,该时域伪信号是明显的。因此,为了满足线性卷积 (linear convolution) 的约束,有必要给每个输入帧加一定数量的零抽样点 (所谓的“填零”方法) 并限制降噪滤波器的冲激响应 (impulse response) 的时域支持 (可在时域或频域执行)。

[0087] 需要注意的是,为了限制降噪滤波器的冲激响应的时域支持 (temporal support), 可以在时域引入约束,这需要:

[0088] i) 第一“逆”频谱变换,使得可以基于滤波器的传递函数 $H_1(k, f)$ 的信息获得冲激响应 $h_1(k, n)$,

[0089] ii) 该冲激响应的点 (point) 数的限制,以获得一个截断的时域滤波器 $h_1'(k, n)$,

[0090] iii) 第二“直接”频谱变换,使得基于被限制的冲激响应 $h_1'(k, n)$ 获得滤波器修正后的传递函数 $H_1'(k, f)$ 。

[0091] 降噪滤波器的 (频域) 传递函数 $\hat{H}_1(k, f)$ 可用的情况下,通过对其进行离散傅立叶变换得到冲激响应 $\hat{h}_1(k, n)$, 即:

$$[0092] \quad \hat{h}_1(k, n) = IDFT(\hat{H}_1(k, f))$$

[0093] 通过用窗 w_{f1111} 对最重要的系数 L_{f1111} 进行选择 and 加权来对上述冲激响应在时域进行限制:

$$[0094] \quad \hat{h}_1'(k, n) = w_{f1111}(n)\hat{h}_1(k, n)$$

[0095] 对噪声滤波器的时域支持的限制呈现出两个优点。一方面,避免了时域伪信号的问题 (与线性卷积相一致)。另一方面,其确保平滑,从而避免了滤波器太过突出的影响。

[0096] 因此,通过离散傅立叶 (DFT) 变换获得降噪滤波器的传递函数 $H_1'(k, f)$:

$$[0097] \quad \hat{H}_1'(k, f) = DFT(\hat{h}_1'(k, n))$$

[0098] 最后,通过频率滤波,所述有噪声信号的频谱与所述降噪滤波器的传递函数相乘获得所述有噪声信号的第一降噪版本,即:

$$[0099] \quad \hat{S}_1(k, f) = \hat{H}_1'(k, f)X(k, f)$$

[0100] 也可以等效的时域滤波方式来执行上述频率滤波步骤,即用 $\hat{h}_1'(k, n)$ 对 $x_w(k, n)$ 直接进行滤波,然后对结果进行离散傅立叶 (DFT) 变换。

[0101] 如上文所解释,由于某些谐波已经被降噪滤波器 $H_1'(k, f)$ 抑制,所以有用信号的

第一估计 $\hat{S}_1(k, f)$ 有听觉失真。当然,失真的程度直接同 SNR 相关。在一个频率域的噪声功率越大,该频率域的谐波越有可能减少。

[0102] 为了减少上述缺点,图 2 中所示的装置 1 中的功能块 8 创建一个伪信号 $\hat{S}_{harmonic}(k, f)$,该信号具有被破坏的或者被严重衰减掉的谐波。

[0103] 可在时域或者频域获得信号 $\hat{S}_{harmonic}(k, f)$ 。在时域,可通过应用非线性变换来恢复信号的调和性,例如,用以下函数中的任一个:单波整流函数、绝对值、关于阈值的最小/最大值等。如图 3 中的步骤 14 所示,我们得到以下表达式,其中 NL 表示所使用的非线性函数:

$$[0104] \quad \hat{s}_{harmonic}(k, n) = NL(\hat{s}_1(k, n))$$

[0105] 其中,根据图 3 中的步骤 13,通过逆离散傅立叶变换 (IDFT) 得到 $\hat{s}_1(k, n)$:

$$[0106] \quad \hat{s}_1(k, n) = IDFT(\hat{S}_1(k, f))$$

[0107] 如图 3 中的步骤 15 所示,通过离散傅立叶变换 (DFT) 得到 $\hat{S}_{harmonic}(k, f)$:

$$[0108] \quad \hat{S}_{harmonic}(k, f) = DFT(\hat{s}_{harmonic}(k, n))$$

[0109] 需要注意的是,谐波在与有用信号相同的位置被重建。隐含地,使用时域非线性变换来恢复谐波确保了这一点能够实现。

[0110] 进一步需要注意的是,在时域应用非线性变换的情况下,通过将非线性变换应用到从 $x_w(k, n)$ 与 $\hat{h}_1(k, n)$ 的卷积直接获得的信号 $\hat{s}_1(k, n)$,可避免转换到频域。

[0111] 图 4 示出了非线性的效果和优点。第一频谱(曲线 16)对应于没有噪声的有用信号的一个帧,并被当作参考。第二频谱(曲线 17)对应于信号 $\hat{S}_1(k, f)$ 。很显然,在该第二频谱中某些谐波已经被完全破坏,其它的谐波被减弱。最后一个频谱(曲线 18)对应于上述提到的信号 $\hat{S}_{harmonic}(k, f)$ 。显然,应用到信号 $\hat{s}_1(k, n)$ 的非线性变换已经使得可以重建丢失的谐波 $\hat{s}_{harmonic}(k, n)$ 。该信号 $\hat{S}_{harmonic}(k, f)$ 的频谱因而具有在计算新的降噪滤波器时特别有用的谐波信息,这个新的降噪滤波器能够保留被传统的降噪方法破坏的谐波。

[0112] 时域非线性变换具有一个等效的频率,其包括在信号 $\hat{S}_1(k, f)$ 的频谱和与有用信号具有同样谐波结构(在谐波位置的级别)的信号的频谱之间进行圆周卷积。这样的频谱可通过例如与函数 $\hat{s}_1(k, n)$ 相应的时域信号 $d(k, n)$ 得到。因而,可通过以下方式在频域获得信号 $\hat{S}_{harmonic}(k, f)$:

$$[0113] \quad \hat{S}_{harmonic}(k, f) = \hat{S}_1(k, f) \otimes DFT(d(k, n))$$

[0114] 其中,符号 \otimes 对应于圆周卷积运算符。如果所选择的时域非线性变换是,例如,绝对值函数,则信号 $d(k, n)$ 可以被如下表示,其中函数 sgn 代表其所应用的值的符号(sign):

$$[0115] \quad d(k, n) = \text{sgn}(\hat{s}_1(k, n))$$

[0116] 因而,通过时域非线性变换或者频域的圆周卷积获得的信号 $\hat{S}_{harmonic}(k, f)$ 具有所述被破坏或者被减弱的谐波 $\hat{S}_1(k, f)$ 。其本身构成降噪信号的梳状谐波的一个好的估计,其中谐波被保留在与其在有用信号中相同的位置。这个梳状谐波估计非常有益,尤其适用于降噪处理,但不限于此。

[0117] 优选地,更进一步,该信号使得计算能够保留通常被传统的算法破坏的谐波的降噪滤波器成为可能,如下文中所述。

[0118] 需要注意的是,本发明不限于上述获取信号 $\hat{S}_{harmono}(k, f)$ 的例子。实际上可本发明可扩展到任何目的在于重建被降噪滤波器 $H_1'(k, f)$ 抑制的谐波的方法。

[0119] 优选地,通过图2的函数块9估计第二滤波器,图5以如下的方式示出了如何在频域中估计第二降噪滤波器的传递函数 $\hat{H}_2(k, f)$, g 表示一个函数:

$$[0120] \quad \hat{H}_2(k, f) = g(\hat{\gamma}_{ss_harmono}(k, f), \hat{\gamma}_{bb}(k, f))$$

[0121] 这个滤波器的传递函数的表达式是一个依赖于噪声的PSD $\hat{\gamma}_{bb}(k, f)$ 以及利用谐波恢复步骤所估计的有用信号的PSD的数学表达式,即:

$$[0122] \quad \hat{\gamma}_{ss_harmono}(k, f) = \rho(k, f) \left| \hat{S}_1(k, f) \right|^2 + (1 - \rho(k, f)) \left| \hat{S}_{harmono}(k, f) \right|^2$$

[0123] 在作为一个被选择用来创建信号 $\hat{S}_{harmono}(k, f)$ 的非线性函数NL的PSD $\hat{\gamma}_{ss_harmono}(k, f)$ 计算中,参数 $\rho(k, f)$ 被用来控制信号 $\hat{S}_{harmono}(k, f)$ 的重注入(reinjection)的程度。需要注意的是,该参数可以是固定的或者依赖于频率和/或时间。

[0124] 函数 g 可以被选择以使得 $\hat{H}_2(k, f)$ 执行例如功率谱减法规则(公式(1))、幅度谱减法规则(公式(2))或者开环维纳滤波器规则(公式(3))。当然也可执行任何其它基于频率噪声抑制规则来估计 $\hat{H}_2(k, f)$ 。

[0125] 降噪传递函数 $\hat{H}_2(k, f)$ 具有保留信号的谐波的优点,该谐波通常被传统的信号降噪方法破坏。这限制了降噪信号的失真。还要注意的,该滤波器保留了滤波器 $\hat{H}_1(k, f)$ 的好的属性,即对于非平稳性和极小的音乐噪声的良好跟踪性。

[0126] 设备1的功能块10基于所述被估计的第二降噪滤波器执行第二次滤波。这样,对于第一传递函数来说,可通过逆离散傅立叶变换(IDFT)得到与降噪滤波器的(频域)传递函数 $\hat{H}_2(k, f)$ 等效的(时域)冲激响应 $\hat{h}_2(k, n)$,即:

$$[0127] \quad \hat{h}_2(k, n) = IDFT(\hat{H}_2(k, f))$$

[0128] 通过利用窗 w_{f11t2} 对最重要的系数 L_{f11t2} 进行选择 and 加权,可方便地在时域限制该冲激响应,即:

$$[0129] \quad \hat{h}_2(k, n) = w_{f11t2}(n) \hat{h}_2(k, n)$$

[0130] 然后,通过离散傅立叶变换(DFT)得到降噪滤波器的传递函数 $\hat{H}_2'(k, f)$:

$$[0131] \quad \hat{H}_2'(k, f) = DFT(\hat{h}_2(k, n))$$

[0132] 最后,通过频率滤波,即将有噪声信号的频谱与降噪滤波器的传递函数相乘得到降噪频率帧,即:

$$[0133] \quad \hat{S}_2(k, f) = \hat{H}_2'(k, f) X(k, f)$$

[0134] 也可以等效的方式通过时域滤波来执行上述频率滤波步骤,即用 $\hat{h}_2'(k, n)$ 对 $x_w(k, n)$ 直接进行滤波。在这种情形下,我们直接在时域获得信号 $\hat{s}(k, n)$

[0135] 当在频域执行滤波时,回到时域是明智的。这由装置1的功能块11通过应用逆频谱变换(IDFT)来执行,即:

$$[0136] \quad \hat{s}(k, n) = IDFT\left(\hat{S}_2(k, f)\right)$$

[0137] 降噪输出信号由此可以通过块重叠相加技术 (OLA 表示“重叠相加”) 或者块保存技术 (OLS 表示“重叠保存”) 方便地合成。这种在时域重构信号的操作被称作逆短期傅立叶变换 (ISTFT)。

[0138] 在这些步骤完成时, 在设备 1 的输出端得到帧 12, 该输出帧 12 是有限失真的输入帧 2 的一个降噪估计。

[0139] 在本说明书的下面部分, 描述了本发明的实施方式的一些优选的选择。这些选择是一些示范性实施例, 不应该被理解为对本发明的限制。

[0140] 为了便于装置 1 的功能块 3 和 11 实现分析和合成, 优选地, 选择 OLA 技术。此外, 在被分析的帧之间存在重叠。例如, 可以选择 50% 的重叠率。长度为 L 的有噪声的信号在当前帧 $x(k, n) \{n = 0, \dots, L-1\}$ 首先被乘以加权窗 $w(n)$, 以得到信号 $x_w(k, n)$:

$$[0141] \quad x_w(k, n) = w(n)x(k, n)$$

[0142] 在该例中, 使用长度为 L 的汉宁窗 (Hanning window) :

$$[0143] \quad w(n) = 0.5 - 0.5 \cos\left(\frac{\pi n}{L}\right) \quad n = 0, K, L-1$$

[0144] 其中, $L = 256$, 对应于采样频率 $F_e = 8\text{kHz}$ 、长度为 32ms 的帧。

[0145] 到频域的转换可以方便地使用快速傅立叶变换 (FFT) 或离散傅立叶变换 (DFT) 的数字优化的版本。为了获得更好的频率以及满足线性卷积约束, 使用长度 $L_{\text{FFT}} = 512$ 的 FFT。因此, 有必要先使用 256 个零采样点加长被加窗的块 (windowed block) $x_w(k, n)$ (“填零”) :

$$[0146] \quad x_{\text{FFT}}(k, n) = \begin{cases} x_w(k, n), & n = 0, K, L-1 \\ 0, & n = L, K, L_{\text{FFT}} - 1 \end{cases}$$

[0147] 这样就可以通过以下公式得到输入帧的 FFT $X(k, f)$:

$$[0148] \quad X(k, f) = \text{FFT}(x_{\text{FFT}}(k, n))$$

[0149] 可选地, 通过功能块 4 来执行声音活动的检测。

[0150] 对于所述的噪声功率谱密度的估计 $\hat{\gamma}_{bb}(k_b, f)$ 由功能块 5 来执行, 然后, 可以以下面的方式方便地继续处理。对于由 DVA 检测到的每个“只有噪声”的帧 k_b , 噪声功率谱密度 $\hat{\gamma}_{bb}(k_b, f)$ 的估计由以下递归表达式给出 :

$$[0151] \quad \hat{\gamma}_{bb}(k_b, f) = \alpha \hat{\gamma}_{bb}(k_b - 1, f) + (1 - \alpha) |X(k_b, f)|^2$$

[0152] 如果检测到当前帧不是一个噪声帧, 则噪声的功率谱密度的估计被标记, 平滑量 α 被选择作为一个常量, 并且 $\alpha = 0.8825$, 这对应于 128ms 的时间常量 (已考虑到 50% 的重叠以及采样频率 $F_e = 8\text{kHz}$), 该时间常量被充分地评价以确保在可靠的估计与噪声统计的时域变化的跟踪之间的折中。

[0153] 优选地, 功能块 7 估计的第一降噪滤波器的传递函数 $\hat{H}_{\text{step1}}(k, f)$ 是如下所示的开环维纳滤波器的传递函数 :

$$[0154] \quad \hat{H}_{\text{step1}}(k, f) = \frac{\eta(k, f)}{1 + \eta(k, f)}$$

[0155] 其中, $\eta(k, f)$ 表示先验信噪比 (SNR), 该先验信噪比理论上由有用信号的 PSD 估计 $\hat{\gamma}_{ss}(k, f)$ 与噪声的 PSD 估计 $\hat{\gamma}_{bb}(k, f)$ 的比值来定义, 即 :

$$[0156] \quad \eta(k, f) = \frac{\hat{\gamma}_{ss}(k, f)}{\hat{\gamma}_{bb}(k, f)}$$

[0157] 优选地,通过使用直接判决估计器得到有用信号的 PSD $\hat{\gamma}_{ss}(k, f)$,即:

$$[0158] \quad \hat{\gamma}_{ss}(k, f) = \beta |\hat{S}(k-1, f)|^2 + (1-\beta) P \left[|X(k, f)|^2 - \hat{\gamma}_{bb}(k, f) \right]$$

[0159] 其中选择重心参数 β 为常量,且 $\beta = 0.98$ 。如上所述,函数 P 确保被 $(1-\beta)$ 加权的数值的阈值:

$$[0160] \quad P[z(k, f)] = \begin{cases} z(k, f) & z(k, f) > 0 \\ 0 & \end{cases}$$

[0161] 优选地,传递函数 $\hat{H}_{step1}(k, f)$ 在之后被重用以改进有用信号的 PSD 的估计 $\hat{\gamma}_{ss}(k, f)$, 由此可以通过以下公式获得数值 $\hat{\gamma}_{ss2}(k, f)$:

$$[0162] \quad \hat{\gamma}_{ss2}(k, f) = \left| \hat{H}_{step1}(k, f) X(k, f) \right|^2$$

[0163] 其用来计算如下公式所示的先验信噪比的第二估计 $\eta_2(k, f)$:

$$[0164] \quad \eta_2(k, f) = \frac{\hat{\gamma}_{ss2}(k, f)}{\hat{\gamma}_{bb}(k, f)}$$

[0165] 优选地,第二轮计算包括基于 $\eta_2(k, f)$ 进行第一降噪滤波器的传递函数的估计值 $\hat{H}_1(k, f)$ 的计算,即:

$$[0166] \quad \hat{H}_1(k, f) = \frac{\eta_2(k, f)}{1 + \eta_2(k, f)}$$

[0167] 由功能块 6 执行的滤波操作可以由下文所示的操作来执行。在所述降噪滤波器的 (频率) 传递函数 $\hat{H}_1(k, f)$ 可用的情况下,优选地,可通过使用快速逆傅利叶变换 (IFFT) 得到相应的时域响应 $\hat{h}_1(k, n)$, 快速逆傅利叶变换 (IFFT) 是逆离散傅利叶变换 (IDFT) 的数字优化的版本,即:

$$[0168] \quad \hat{h}_1(k, n) = IFFT(\hat{H}_1(k, f))$$

[0169] 因为信号 $\hat{h}_1(k, n)$ 是实数,首先提供相应的具有因果性的时域滤波器。然后选择与抽样点相应的所述滤波器的系数 $L_{fill1} = 256$, 所述抽样点对所关注的的应用非常重要。然后,用长度为 L_{fill1} 的汉宁窗对所获得的冲激响应进行加权。最后,将加权冲激响应填充 $L_{FFT} - L_{fill1}$ 个零以提供满足避免失真的线性卷积约束的冲激响应 $\hat{h}_1(k, n)$ 。

[0170] 然后,优选地,通过使用快速傅利叶变换 (FFT) 获得降噪滤波器的传递函数 $\hat{H}'_1(k, f)$, 快速傅利叶变换 (FFT) 是离散傅利叶变换 (DFT) 的数字优化的版本,即:

$$[0171] \quad \hat{H}'_1(k, f) = FFT(\hat{h}_1(k, n))$$

[0172] 最后,通过频率滤波,即通过将有噪声信号的频谱与降噪滤波器传递函数相乘获得有噪声的信号的第一降噪版本,即:

$$[0173] \quad \hat{S}_1(k, f) = \hat{H}'_1(k, f) X(k, f)$$

[0174] 在功能块 8 执行的恢复丢失的谐波的步骤中,使用信号 $\hat{S}_1(k, f)$ 来创建一个伪信号 $\hat{S}_{harmo}(k, f)$ 。首先,优选地,通过逆快速傅利叶变换 (IFFT) 获得时域信号 $\hat{s}_1(k, n)$:

[0175] $\hat{S}_1(k, n) = IFFT(\hat{S}_1(k, f))$

[0176] 此后,将单波整流非线性函数应用于该信号,其目的在于恢复信号的调谐性,即:

[0177] $\hat{S}_{\text{harmonic}}(k, n) = \max(\hat{S}_1(k, n), 0)$

[0178] 当然,任何其它除单波整流函数之外的非线性函数(绝对值,阈值的最大/最小值等)都可以用来进行上述操作。最后,优选地,通过使用快速傅利叶变换(FFT)获得信号 $\hat{S}_{\text{harmonic}}(k, f)$,即:

[0179] $\hat{S}_{\text{harmonic}}(k, f) = FFT(\hat{S}_{\text{harmonic}}(k, n))$

[0180] 然后,优选地,可利用该信号来计算第二降噪滤波器的传递函数。

[0181] 在本示例性实施例中,这样的第二滤波器的估计由装置1的功能块9以下述方式执行。第二降噪滤波器的传递函数 $\hat{H}_2(k, f)$ 是如下开环维纳滤波器的传递函数:

[0182]
$$\hat{H}_2(k, f) = \frac{\eta_{\text{harmonic}}(k, f)}{1 + \eta_{\text{harmonic}}(k, f)}$$

[0183] 其中, $\eta_{\text{harmonic}}(k, f)$ 表示先验信噪比(SNR),该先验信噪比理论上由有用信号的PSD估计 $\hat{\gamma}_{\text{ss_harmonic}}(k, f)$ 与噪声的PSD估计 $\hat{\gamma}_{\text{bb}}(k, f)$ 的比值来定义,即:

[0184]
$$\eta_{\text{harmonic}}(k, f) = \frac{\hat{\gamma}_{\text{ss_harmonic}}(k, f)}{\hat{\gamma}_{\text{bb}}(k, f)}$$

[0185] 其中,有用信号的PSD $\hat{\gamma}_{\text{ss_harmonic}}(k, f)$ 由以下公式获得:

[0186]
$$\hat{\gamma}_{\text{ss_harmonic}}(k, f) = \rho(k, f) \left| \hat{S}_1(k, f) \right|^2 + (1 - \rho(k, f)) \left| \hat{S}_{\text{harmonic}}(k, f) \right|^2$$

[0187] 用来控制信号 $\hat{S}_{\text{harmonic}}(k, f)$ 再注入的程度的参数 $\rho(k, f)$ 在应用中被选择为固定不变,并且等于0.5。

[0188] 由于可获得降噪滤波器的(频率)传递函数 $\hat{H}_2(k, f)$,所以优选地,可通过使用快速逆傅利叶变换(IFFT)得到相应的时域响应 $\hat{h}_2(k, n)$,快速逆傅利叶变换(IFFT)是逆离散傅利叶变换(IDFT)的数字优化的版本:

[0189] $\hat{h}_2(k, n) = IFFT(\hat{H}_2(k, f))$

[0190] 因为信号 $\hat{h}_1(k, n)$ 是实数,首先提供相应的具有因果性的时域滤波器。然后选择与抽样点相应的所述滤波器的系数 $L_{\text{filt}2} = 256$,所述抽样点对所述应用非常重要。然后,用长度为 $L_{\text{filt}2}$ 的汉宁窗对所获得的冲激响应进行加权。最后,将加权冲激响应填充 $L_{\text{FFT}} - L_{\text{filt}2}$ 个零以提供满足避免失真的线性卷积约束的冲激响应 $\hat{h}_2(k, n)$ 。

[0191] 然后,优选地,通过使用快速傅利叶变换(FFT)获得降噪滤波器的传递函数 $\hat{H}_2'(k, f)$,快速傅利叶变换(FFT)是离散傅利叶变换(DFT)的数字优化的版本,即:

[0192] $\hat{H}_2'(k, f) = FFT(\hat{h}_2(k, n))$

[0193] 可选地,可以对所述降噪滤波器 $\hat{H}_2'(k, f)$ 的模(modulus)的值进行阈值限制,以避免不平滑的降噪。

[0194] 最后,通过功能块10执行的频率滤波获得降噪频率帧,即通过将有噪声信号的频谱与降噪滤波器传递函数相乘:

[0195] $\hat{S}_2(k, f) = \hat{H}'_2(k, f)X(k, f)$

[0196] 然后在功能块 11 级别通过逆频谱变换执行将信号转换回时域的变换, 优选地, 使用逆快速傅立叶变换 (IFFT), 即

[0197] $\hat{s}(k, n) = IDFT(\hat{S}_2(k, f))$

[0198] 然后通过块重叠相加方法 OLA 来合成降噪输出信号。

[0199] 此外, 需要注意的是, 图 2 中所示的装置 1 作为所应用的功能, 可以被放置在一种特殊的设备中, 诸如声音采集设备、通信设备或者其它的语音识别设备。

[0200] 此外, 需要注意的是, 装置所执行的处理可以本地方式执行, 例如在通信终端中, 或者在网络中以集中化方式执行, 例如在语音识别服务器中。

[0201] 另外, 在实际中, 根据本发明, 用于处理有噪声的声音信号的装置 1 包括适用于执行根据本发明的处理信号方法的 (电子) 硬件装置和 / 或软件装置。

[0202] 根据本发明的一个优选实现, 处理根据本发明的有噪声的声音信号的方法的步骤由根据本发明的设备中使用的计算机程序的指令来确定。

[0203] 当上述程序被加载到包含在所述设备中的计算装置时, 根据本发明的方法被执行, 该方法的执行由运行所述程序来控制。

[0204] 这里, “计算机程序” 被理解为一个或多个计算机程序形成的 (软件) 集合, 当该 “计算机程序” 被适当的计算装置运行时, 其目的在于实现本发明的方法。

[0205] 因此, 本发明的主题也包括这样的计算机程序, 尤其是以软件的形式存储在信息媒介中的计算机程序。所述的信息媒介可包括任何能够存储根据本发明的程序的实体或者装置。

[0206] 例如, 所述的媒介可包括硬件存储装置, 如 ROM, 例如 CD ROM、微电子电路 ROM, 或者是磁记录装置, 如硬盘。作为一个变化例, 信息媒介还可以是包含所述程序的集成电路, 该电路适合于执行所述方法, 或者适合于被用于所述方法的执行中。

[0207] 此外, 该信息媒介还可以是可传输的非物质介质, 如能够通过无线电或其他方法, 用电缆或者光缆传送的电信号或者光信号。特别的, 可从因特网类型的网络中下载基于本发明的程序。

[0208] 从设计的观点来看, 根据本发明的计算机程序可使用任何语言, 可以是源代码、目标码或者是在源代码和目标码之间的中间码 (如部分编译的形式), 或者是其它任何可实现根据本发明的方法的形式。

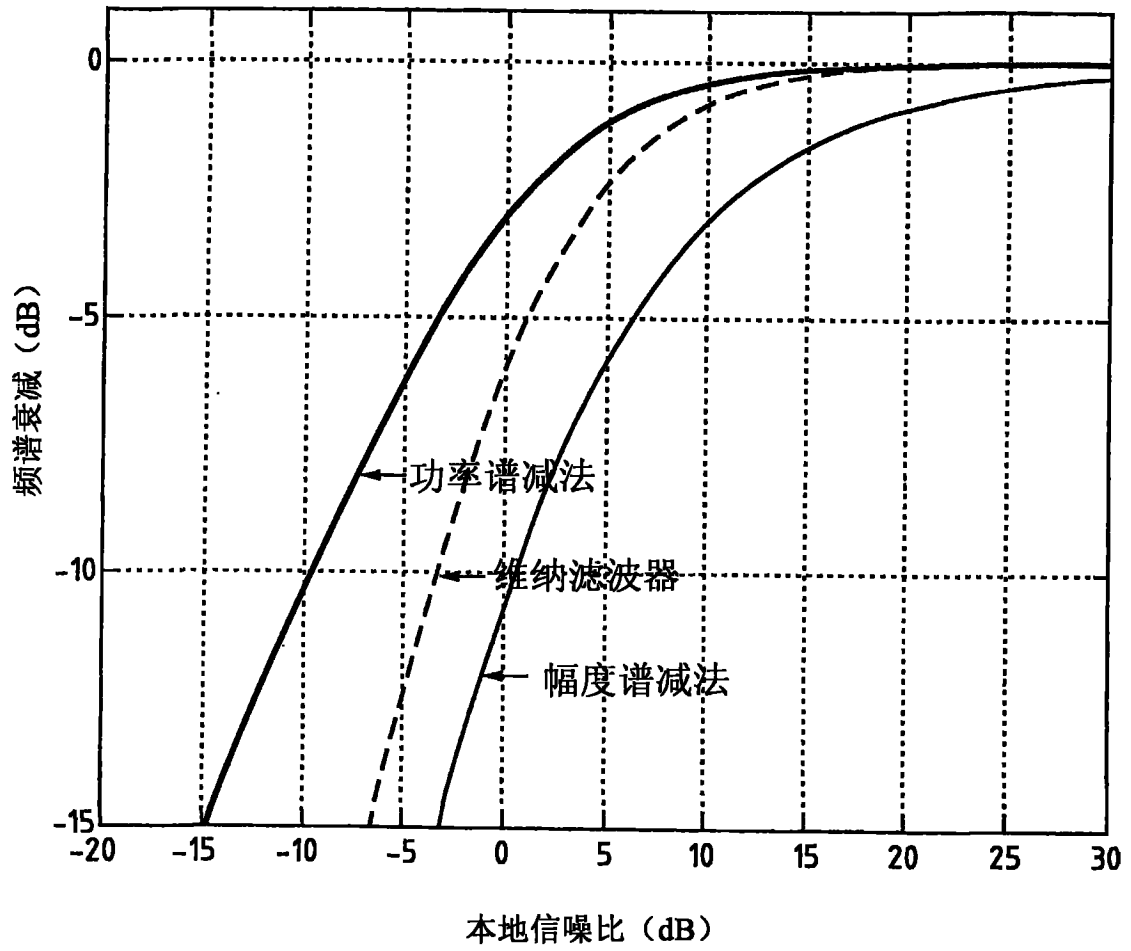


图 1

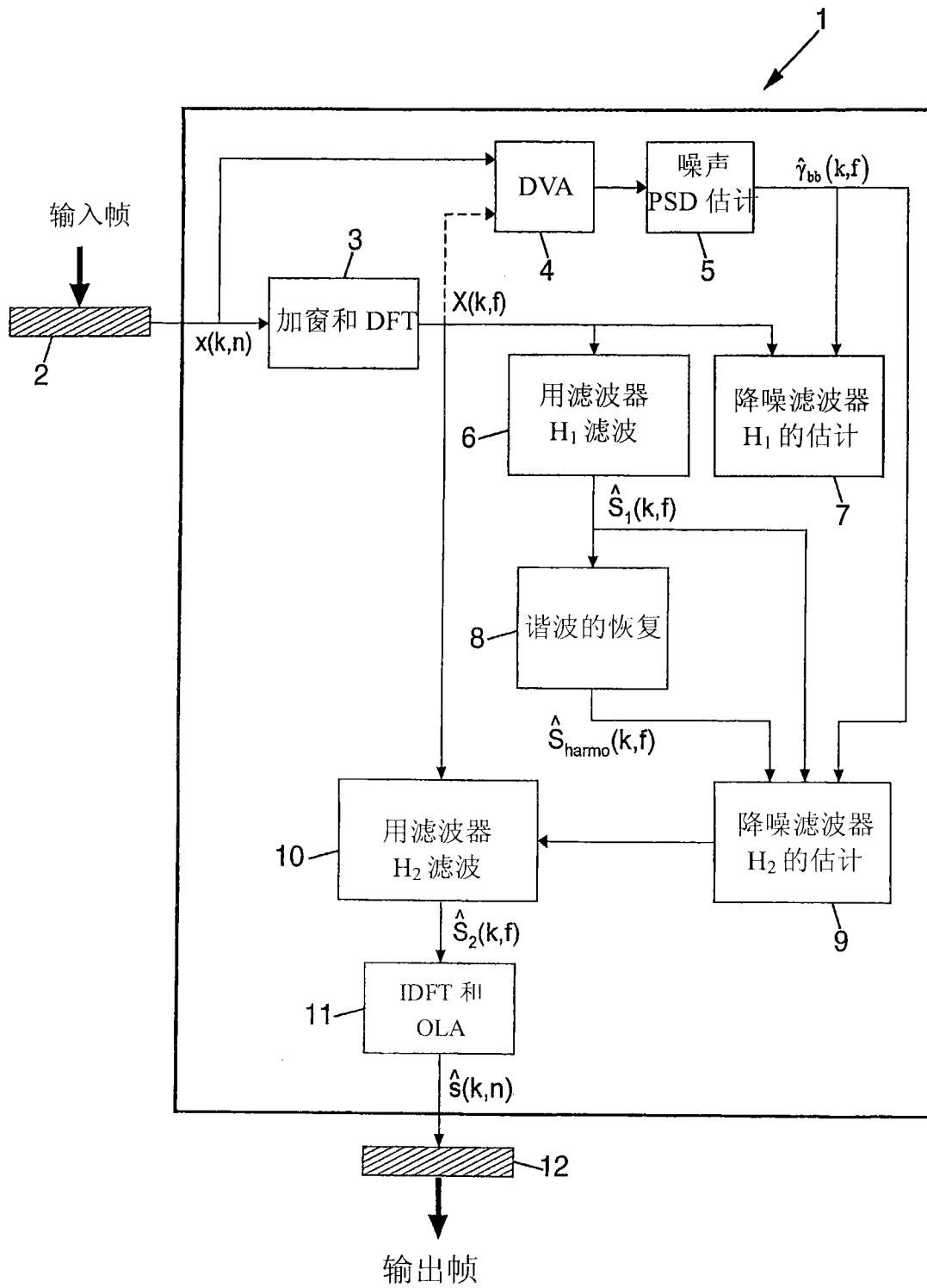


图 2

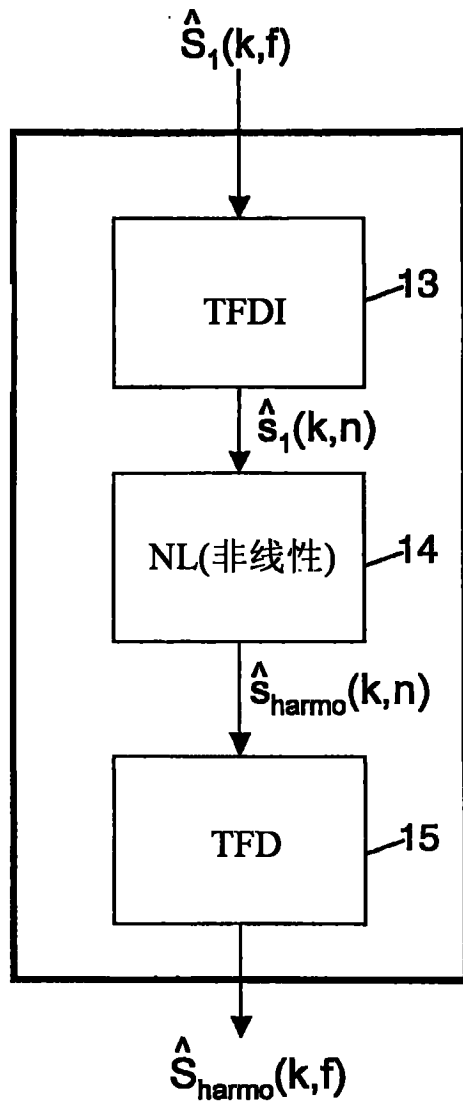


图 3

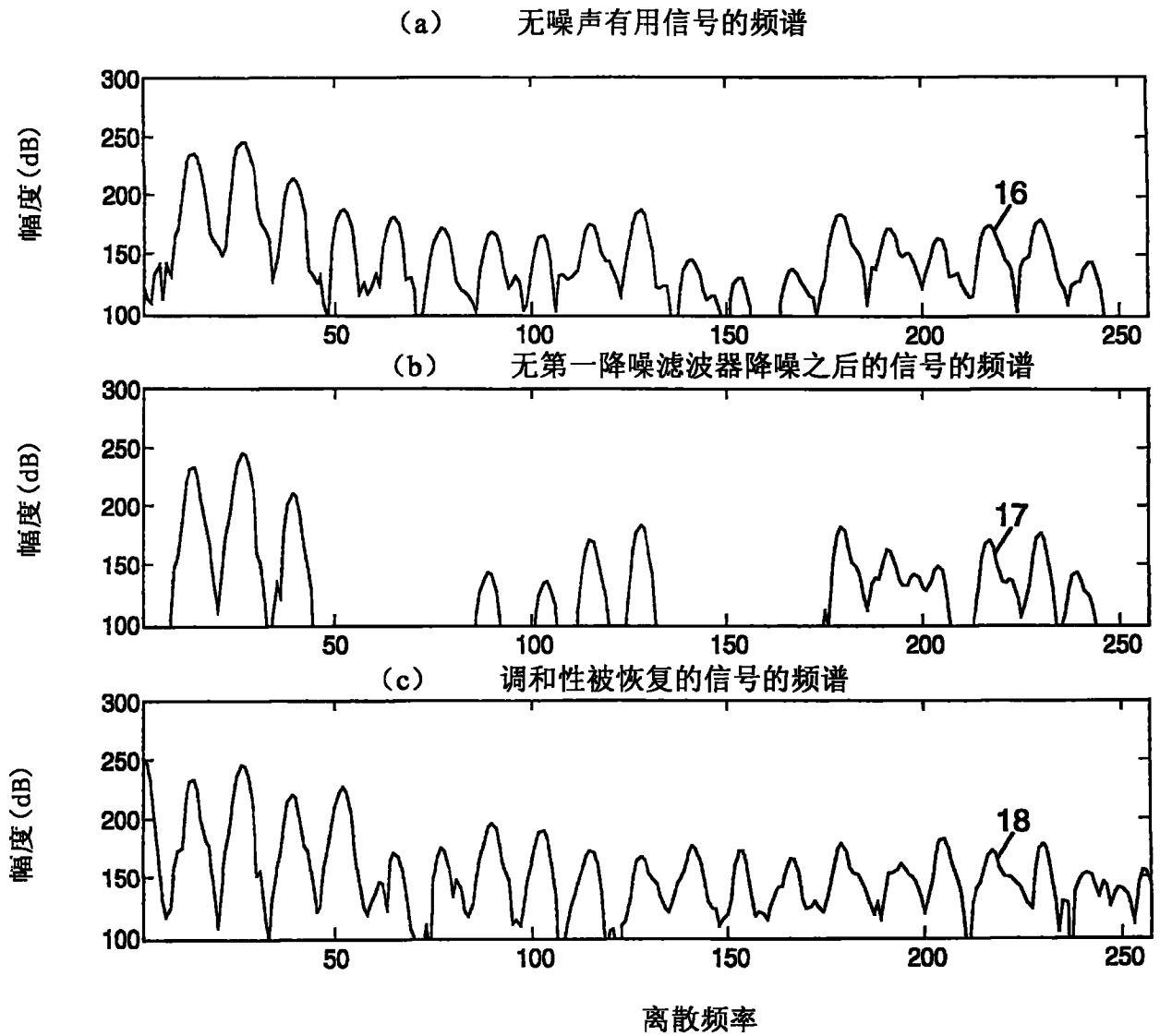


图 4

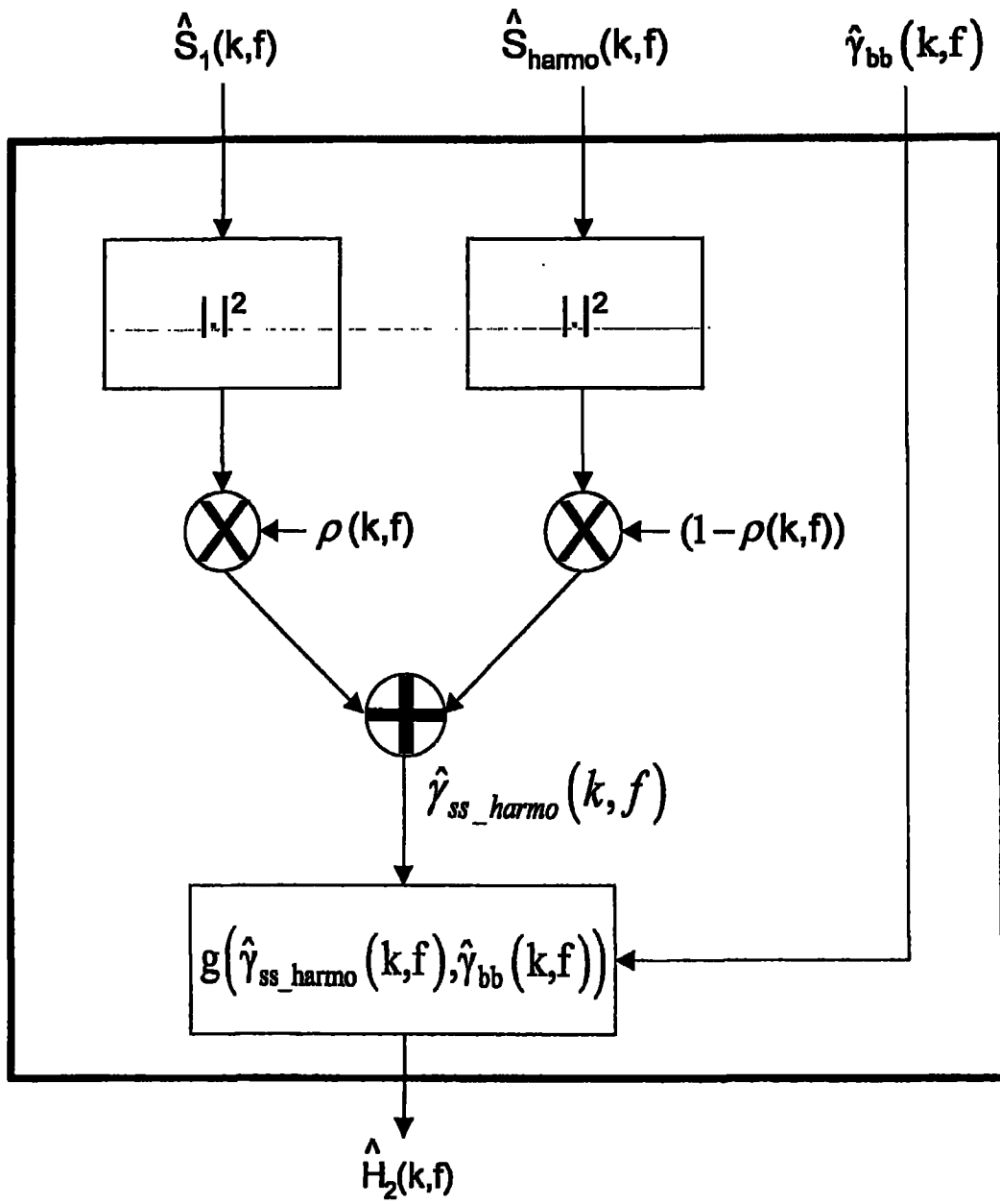


图 5