



(12) 发明专利

(10) 授权公告号 CN 101727910 B

(45) 授权公告日 2012. 07. 04

(21) 申请号 200910205534. 7

US 4630305 A, 1986. 12. 16,

(22) 申请日 2009. 10. 26

Steven F. Boll. Suppression of Acoustic

(30) 优先权数据

2008-274772 2008. 10. 24 JP

Noise in Speech Using Spectral Subtraction.

(73) 专利权人 雅马哈株式会社

《IEEE TRANSACTIONS ON ACOUSTICS, SPEECH,

地址 日本静冈县

AND SIGNAL PROCESSING》. 1979, 第 ASSP-27 卷

(第 2 期),

审查员 刘涛

(72) 发明人 刘恩彩

(74) 专利代理机构 北京天昊联合知识产权代理

有限公司 11112

代理人 何立波 张天舒

(51) Int. Cl.

G10L 21/02 (2006. 01)

(56) 对比文件

CN 1969320 A, 2007. 05. 23,

CN 1354873 A, 2002. 06. 19,

CN 1573930 A, 2005. 02. 02,

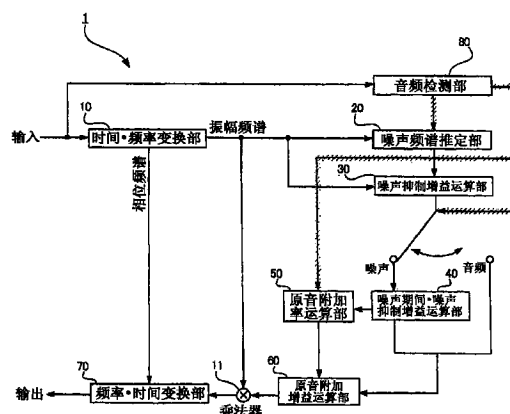
权利要求书 2 页 说明书 15 页 附图 6 页

(54) 发明名称

噪声抑制装置及噪声抑制方法

(57) 摘要

本发明提供一种噪声抑制装置及噪声抑制方法, 该噪声抑制装置不产生音乐噪声, 而且, 即使在音频帧及噪声帧之间切换的情况下, 噪声量的变化也很自然。其具有: 噪声频谱推定部 (20), 其针对 K 个频带 (其中, K 为大于或等于 2 的自然数), 分别推定包含在输入信号中的噪声频谱; 噪声抑制增益运算部 (30), 其基于该推定结果, 计算上述 K 个频带各自的噪声抑制增益; 噪声期间·噪声抑制增益运算部 (40), 其计算上述噪声抑制增益的与上述 K 个频带相关的平均值增益; 以及噪声抑制单元, 其通过对上述输入信号所涉及的上述 K 个频带的全部应用上述平均值增益, 从而抑制包含在该输入信号内的噪声。



1. 一种噪声抑制装置,其特征在于,具有:

噪声频谱推定单元,其基于输入信号,针对 K 个频带分别推定包含在该输入信号内的噪声频谱,其中, K 为大于或等于 2 的自然数;

第 1 增益运算单元,其基于所述噪声频谱推定单元的噪声频谱的推定结果,计算所述 K 个频带各自的噪声抑制增益;

第 2 增益运算单元,其针对所述噪声抑制增益,计算与所述 K 个频带的全部或者一部分相关的平均值增益;

音频检测单元,其通过随时间推移地对所述输入信号中是否包含音频进行检测,从而将该输入信号区分为包含该音频的音频帧、以及不包含该音频的噪声帧;以及

噪声抑制单元,其通过对所述输入信号所涉及的所述 K 个频带的全部或者一部分应用所述平均值增益,从而抑制包含在该输入信号内的噪声,

所述噪声抑制单元向所述输入信号中相当于所述噪声帧的部分,应用所述平均值增益,向所述输入信号中相当于所述音频帧的部分,应用所述噪声抑制增益。

2. 如权利要求 1 所述的噪声抑制装置,其特征在于,

所述第 2 增益运算单元计算将所述平均值增益在时间轴上平滑化后的平滑化增益,

所述噪声抑制单元向所述输入信号中相当于所述噪声帧的部分,应用所述平滑化增益。

3. 如权利要求 1 或 2 所述的噪声抑制装置,其特征在于,

所述噪声频谱推定单元,

在要推定与所述音频帧相关的噪声频谱的情况下,

将针对与该音频帧最接近的噪声帧而已经推定出的噪声频谱,推定为该音频帧的噪声频谱。

4. 如权利要求 1 或 2 所述的噪声抑制装置,其特征在于,

所述噪声频谱推定单元,

在要推定与所述噪声帧相关的噪声频谱的情况下,

在计算出临时频谱推定值后,使用该噪声帧的前一个帧的噪声频谱推定值,计算出将该临时频谱推定值在时间轴上平滑化后的平滑化噪声频谱推定值,将该平滑化噪声频谱推定值推定为该噪声帧的噪声频谱,

在要推定与所述音频帧相关的噪声频谱的情况下,

直接将该音频帧的前一个帧的噪声频谱推定值,推定为该音频帧的噪声频谱。

5. 一种噪声抑制方法,其特征在于,包含:

噪声频谱推定工序,其基于输入信号,针对 K 个频带分别推定包含在该输入信号内的噪声频谱,其中, K 为大于或等于 2 的自然数;

第 1 增益运算工序,其基于所述噪声频谱推定工序中的噪声频谱的推定结果,计算所述 K 个频带各自的噪声抑制增益;

第 2 增益运算工序,其针对所述噪声抑制增益,计算与所述 K 个频带的全部或者一部分相关的平均值增益;

音频检测工序,其通过对所述输入信号中是否包含音频进行检测,从而将该输入信号区分为包含该音频的音频帧、以及不包含该音频的噪声帧;以及

噪声抑制工序,其通过对所述输入信号所涉及的所述 K 个频带的全部或者一部分应用所述平均值增益,从而抑制包含在该输入信号内的噪声,

在所述噪声抑制工序中,向所述输入信号中相当于所述噪声帧的部分应用所述平均值增益,向所述输入信号中相当于所述音频帧的部分,应用所述噪声抑制增益。

噪声抑制装置及噪声抑制方法

技术领域

[0001] 本发明涉及一种噪声抑制装置及噪声抑制方法。

背景技术

[0002] 当前,提出了下述装置等,即:音频播放装置,其与输入信号对应地驱动扬声器等负载;音频通信装置,其在远离地点间传送音频;以及音频识别装置,其通过对音频的种类等进行区分·识别,从而理解其内容等。在上述各装置中,为了对音频准确地进行播放、传递或者识别等,而优选除去包含在其中的噪声的影响。

[0003] 作为这样的噪声抑制的技术,已知例如以下的专利文献 1 及非专利文献 1 至 4 中公开的技术。

[0004] 专利文献 1:日本特开 2007-226264 号公报

[0005] 非专利文献 1:Boll, S.,“Suppression of acoustic noise in speech using spectral subtraction”, IEEE Trans. Vol. ASSP-27, No2, pp. 113-120, 1979.

[0006] 非专利文献 2:M. Berouti, et al,“Enhancement of Speech Corrupted by Acoustic Noise”, Proceedings of ICASSP, pp. 201-211, 1979.

[0007] 非专利文献 3:Lim & Oppenheim,“Enhancement and Bandwidth Compression of Noisy Speech”, Proc. IEEE, Vol67, No12, pp. 1586-1604, 1979

[0008] 非专利文献 4:Y. Ephraim and D. Malah,“Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator”, IEEE Trans. Vol. ASSP-32, No. 6, pp1109-1121, 1984.

发明内容

[0009] 上述各文献中公开的技术,基本上与以下方法相关,即,通过适当的方法减去频带的振幅频谱的电平而抑制噪声的方法,即所谓谱减法 (Spectral Subtraction),通过这些中的任一项技术,都可以得到一定的噪声抑制效果。

[0010] 但是,在这些文献中,仍存在未公开·未解决的课题。

[0011] 例如,上述谱减法基于下述方式进行,即,推定包含在输入信号中的噪声频谱,从振幅频谱中减去该噪声频谱推定值,但在此情况下,存在很可能产生所谓音乐噪声的问题。这是因为,在这里所说的噪声频谱推定值不一定反映实际的噪声频谱。即,存在在某频带中,减去噪声频谱推定值后仍残留噪声的情况,另外,存在在其他的频带中,产生削减过度的情况。因此,如果将减去噪声频谱后的振幅频谱向时域进行再变换,则有可能出现将多个具有随机频率的正弦波合成的信号,通过对其进行播放,可能产生非常刺耳的噪声(即,音乐噪声)。

[0012] 另外,在输入信号中,存在主要由音频占据的部分(音频部分)和几乎没有音频的部分(噪声部分)。

[0013] 在上述状况下,例如,如上述非专利文献 1 所示,在对音频部分应用谱减法、对噪

声部分应用固定增益,由此抑制噪声的情况下,在该固定增益值过小时,在从噪声部分向音频部分切换的情况下,产生背景噪声量增大的现象,而在固定增益值过大时,相反地,可能产生背景噪声量减小的现象。如果对其进行例如播放,则很可能使收听者感到听觉上的不自然感。

[0014] 本发明的课题是,提供一种可以解决上述课题中的至少一部分的噪声抑制装置及噪声抑制方法。

[0015] 为了解决上述课题,本发明所涉及的噪声抑制装置具有:噪声频谱推定单元,其基于输入信号,针对 K 个频带(其中, K 为大于或等于 2 的自然数)分别推定包含在该输入信号内的噪声频谱;第 1 增益运算单元,其基于所述噪声频谱推定单元的噪声频谱的推定结果,计算所述 K 个频带各自的噪声抑制增益;第 2 增益运算单元,其针对所述噪声抑制增益,计算与所述 K 个频带的全部或者一部分相关的平均值增益;以及噪声抑制单元,其通过对所述输入信号所涉及的所述 K 个频带的全部或者一部分应用所述平均值增益,从而抑制包含在该输入信号内的噪声。

[0016] 根据本发明,分别计算 K 个噪声频谱以及 K 个噪声抑制增益。另外,第 2 增益运算单元根据其中的 K 个噪声抑制增益计算平均值增益。即,在本发明中,所谓计算“与 K 个频带的全部…相关的”平均值,是指如果假设噪声抑制增益为 $G(1)$ 、 $G(2)$ 、…、 $G(K)$,则例如由 $G_{ave} = (G(1)+G(2)+\cdots+G(K))/K$ 等计算平均值增益 G_{ave} (此外,在计算该例中所示的平均值增益时,不一定要使用全部 K 个噪声抑制增益。本发明中所称的“与 K 个频带的…‘一部分’相关的平均值增益”,包含如上述所示的情况)。

[0017] 另外,在本发明中,将上述的平均值增益 G_{ave} 针对 K 个频带的全部或者一部分而应用于输入信号。例如,如果将时域的输入信号变换为频域而得到的振幅频谱设为 $Y(1)$ 、 $Y(2)$ 、…、 $Y(K)$,则频域的输出信号可以由 $G_{ave} \cdot Y(1)$ 、 $G_{ave} \cdot Y(2)$ 、…、 $G_{ave} \cdot Y(K)$ 等而得到。

[0018] 根据上述内容,在本发明中,不但可以得到噪声抑制效果,而且例如,不会发生由上述噪声频谱推定值引起的振幅频谱的过度削减的情况等,因此,可以非常有效地抑制音乐噪声的产生。

[0019] 在本发明所涉及的噪声抑制装置中,也可以构成为,所述第 2 增益运算单元计算将所述平均值增益在时间轴上平滑化后的平滑化增益,所述噪声抑制单元通过对所述输入信号所涉及的所述 K 个频带的全部或者一部分应用所述平滑化增益,从而抑制包含在该输入信号中的噪声。

[0020] 根据该形态,可以计算将所述平均值增益在时间轴上平滑化后的平滑化增益。因此,该平滑化增益是在接受上述平均化处理的基础上,进一步接受平滑化处理而得到的(如果重视这一点,则也可以将该平滑化增益称为平均化·平滑化增益)。此外,所谓“在时间轴上平滑化”是指假定将计算后的平滑化增益沿时间顺序为 $G_{smt-T}(1)$ 、 $G_{smt-T}(2)$ 、…、 $G_{smt-T}(r)$ 、…(r 是适当的整数),则使用例如适当的平滑化系数 δ ,作为 $G_{smt-T}(r) = \delta \cdot G_{smt-T}(r-1) + (1-\delta) \cdot G_{ave}$ 等而进行计算。

[0021] 根据上述内容,由于不会使平均值增益(准确地说是平滑化后的平均值增益。即,本形态所称的“平滑化增益”)随着时间的经过而产生激烈的变化,所以可以维持噪声抑制处理的连续性·一贯性。

[0022] 此外,在本发明中,如后述所示,优选针对随时间推移而划分的每个帧进行处理,但在此情况下,本形态中所称的“时间轴”,更具体地说,可以假设为在将该帧一个一个地按顺序排列的情况下抽象出的轴。关于这一点的更详细的具体例子,在后述的实施方式、特别是式(4)时进行说明。

[0023] 另外,在本发明所涉及的噪声抑制装置中,也可以构成为,还具有音频检测单元,其通过随时间推移地对所述输入信号中是否包含音频进行检测,从而将该输入信号区分为包含该音频的音频帧、以及不包含该音频的噪声帧,所述噪声抑制单元向所述输入信号中相当于所述噪声帧的部分,应用所述平均值增益或者所述平滑化增益。

[0024] 根据该形态,将上述平均值增益或者平滑化增益向噪声帧应用,更优选仅向噪声帧应用。对于噪声帧,由于比较容易产生音乐噪声,而本形态则针对这一问题而进行平均值增益等的应用,所以可以说是用于获得音乐噪声抑制效果的优选方式之一。

[0025] 此外,在本形态中,对于“含有”或者“不含有”音频这样的用语,不能理解为所谓绝对的意思。例如,本形态并不限定于抽象性地假设“全部由音频”占据的帧和“根本不存在音频”的帧这两种形态,针对这两者两极端而进行“音频帧”及“噪声帧”的区分的情况,而且,本形态也不限定于仅以后者为“噪声帧”,除此之外全部为“音频帧”而进行区分的情况。即,在本形态中,即使在认定为“噪声帧”的情况下,也不要求该噪声帧中根本不包含“音频”,本形态中所称的“音频帧”及“噪声帧”的区分,可以以上述2种情况的适当的中间点为基准而进行。

[0026] 在以上所述的意思中,本形态中所称的“包含”或者“不包含”这样的用语或者本形态所涉及的“音频帧”及“噪声帧”的区分,可以说是相对的。

[0027] 在本形态中也可以构成为,所述噪声抑制单元向所述输入信号中相当于所述音频帧的部分,应用所述噪声抑制增益。

[0028] 根据本形态,将所述噪声抑制增益向音频帧应用。该形态可以与之前所述的形态并存,但在此情况下,优选对于音频帧仅使用没有经过平均化处理的所谓通常的噪声抑制增益,而对于噪声帧仅使用经过平均化处理的平均值增益或者平滑化增益。如果鉴于在音频帧中噪声的存在并不那么明显而在噪声帧中却相反这一情况,则本形态所涉及的处理内容,可以得到非常合理的・高效的・有效的噪声抑制效果。

[0029] 在包含所述“音频检测单元”的本发明的形态中,也可以构成为,所述噪声频谱推定单元在要推定与所述音频帧相关的噪声频谱的情况下,将针对与该音频帧最接近的噪声帧而已经推定出的噪声频谱,推定为该音频帧的噪声频谱。

[0030] 根据该形态,音频帧的噪声频谱与其最接近的处理后的噪声帧的噪声频谱相匹配。在这里,由于噪声抑制增益是基于推定的噪声频谱而求出的,所以该音频帧的噪声抑制增益原则上也与该噪声帧中的噪声抑制增益相匹配(但是,在此情况下,两者之间的值当然有可能不同。)

[0031] 根据上述内容,在本形态中,使对最接近的噪声帧执行的噪声抑制处理,在其后的音频帧的噪声抑制处理中沿用,在从噪声帧向音频帧切换的情况下,可以保持噪声抑制处理的一贯性。由此,在该切换的情况下,可以事先防止噪声量急剧变化等现象的发生。

[0032] 另外,在包含所述“音频检测单元”的本发明的形态中,也可以构成为,所述噪声频谱推定单元在要推定与所述噪声帧相关的噪声频谱的情况下,在计算出临时频谱推定值

后,使用该噪声帧的前一个帧的噪声频谱推定值,计算出将该临时频谱推定值在时间轴上平滑化后的平滑化噪声频谱推定值,将该平滑化噪声频谱推定值推定为该噪声帧的噪声频谱,在要推定与所述音频帧相关的噪声频谱的情况下,直接将该音频帧的前一个帧的噪声频谱推定值,推定为该音频帧的噪声频谱。

[0033] 根据该形态,可以适当地推定音频帧及噪声帧各自的噪声频谱。由于针对噪声帧,计算平滑化噪声频谱推定值,所以可以保持噪声抑制处理的一贯性・连续性,针对音频帧,适当地维持如上述所示的“最接近的噪声帧”的噪声频谱(即,根据本方式,如果已针对某个噪声帧推定了噪声频谱,则在此之后,只要持续为音频帧,则继续维持该噪声频谱。)。

[0034] 此外,在本形态中,在仅称为“帧”的情况中,包括“音频帧”的情况,以及“噪声帧”的情况。

[0035] 关于这一点的更详细的具体例子,在后述的实施方式、特别是式(1)时进行说明。

[0036] 另一方面,为了解决上述课题,本发明所涉及的噪声抑制方法包含:噪声频谱推定工序,其基于输入信号,针对K个频带(其中,K为大于或等于2的自然数)分别推定包含在该输入信号内的噪声频谱;第1增益运算工序,其基于所述噪声频谱推定工序中的噪声频谱的推定结果,计算所述K个频带各自的噪声抑制增益;第2增益运算工序,其针对所述噪声抑制增益,计算与所述K个频带的全部或者一部分相关的平均值增益;以及噪声抑制工序,其通过对所述输入信号所涉及的所述K个频带的全部或者一部分应用所述平均值增益,从而抑制包含在该输入信号内的噪声。

[0037] 根据本发明,明确其可以实现与上述针对本发明所涉及的噪声抑制装置记述的作用效果没有本质不同的作用效果。

[0038] 在本发明所涉及的噪声抑制方法中,也可以构成,还包含音频检测工序,其通过对所述输入信号中是否包含音频进行检测,从而将该输入信号区分为包含该音频的音频帧、以及不包含该音频的噪声帧,所述噪声抑制工序向所述输入信号中相当于所述噪声帧的部分,应用所述平均值增益。

[0039] 根据本形态,明确其可以实现与上述针对本发明所涉及的噪声抑制装置的各种形态中向噪声帧应用平均值增益的形态记述的作用效果没有本质不同的作用效果。

[0040] 此外,本形态中所称的“含有”或者“不含有”这样的用语的意义,与上述相同。

[0041] 除上述之外,本发明的更具体化的方式或者由此实现的作用效果,在从下面开始的实施方式的记载中明确记述。

附图说明

[0042] 图1是表示本发明的实施方式1所涉及的噪声抑制装置的结构框图。

[0043] 图2是表示实施方式1所涉及的噪声抑制处理的流程流程图。

[0044] 图3是用于说明实施方式1所涉及的噪声抑制处理的内容的说明图。

[0045] 图4是用于说明现有噪声抑制处理的内容的说明图。

[0046] 图5是表示本发明的实施方式2所涉及的噪声抑制装置的结构框图。

[0047] 图6是表示包含音频信号的情况下的噪声抑制增益 $G(n)$ 的运算例的曲线图。

[0048] 图7是表示不包含音频信号的情况下的噪声抑制增益 $G(n)$ 的运算例的曲线图。

具体实施方式

[0049] <实施方式 1>

[0050] 下面,参照图 1,说明本发明所涉及的实施方式 1。此外,在这里提到的图 1,以及以下所参照的各附图(也包含例如图 6 等的曲线图)中,存在各部分的尺寸比例与实际尺寸相比适当改变的情况。噪声抑制装置 1 如图 1 所示,由时间·频率变换部 10、噪声频谱推定部 20、噪声抑制增益运算部 30、噪声期间·噪声抑制增益运算部 40、原音附加率运算部 50、原音附加增益运算部 60、频率·时间变换部 70、以及音频检测部 80 构成。

[0051] 时间·频率变换部 10 对时域的输入信号进行傅里叶变换,而变换为频域的信号。优选该傅里叶变换是通过将输入信号随时间推移分成规定数量的帧,且向该帧施加适当的窗口函数而进行的。

[0052] 上述频域的信号分为振幅频谱及相位频谱,将其中的相位频谱直接向后述的频率·时间变换部 70 发送。另一方面,将振幅频谱向后述的噪声频谱推定部 20 以后的各部分输送,接受后述的各种处理。

[0053] 还将上述时域的输入信号向音频检测部 80 供给。音频检测部 80 对该输入信号中有无音频信号进行检测。如上述所示,在将输入信号分成帧的情况下,针对每个帧进行音频检测(此外,在实施方式 1 中,将这种处理作为前提。)。在这里,所谓“音频”,特别指会话、口语、音乐、各种信号等对于人来说有意义的声音。即,以下关系成立,在将输入信号通过适当的播放单元进行播放的情况下,如果播放该输入信号中的“音频信号”,则成为该声音。

[0054] 对于该音频信号,是将例如下述条件作为基准而进行检测的,即,输入信号的电平是否超过预先确定的阈值。但是,本发明除此之外还可以采用各种方法。例如,可以采用使用概率·统计的方法推定音频信号的产生概率的方法等,或者,也可以采用不使用上述输入信号,而使用其傅里叶变换后的信号(即,上述所称的频域的信号)作为检测对象的方法等。

[0055] 此外,以下,有时分别将通过该音频检测部 80 判定为存在音频信号的帧称为“音频帧”,将判定为不存在音频信号的帧称为“噪声帧”。此外,在这里,存在·不存在不具有所谓绝对的意义。如上述所示,既然是以规定的阈值为基准,对音频信号的有无进行判断,所以严格来讲,不排除在“噪声帧”中包含被称为音频信号的部分的可能性。

[0056] 噪声频谱推定部 20 基于上述振幅频谱,计算噪声频谱的推定值。在实施方式 1 中,特别地,基于以下的式(1),针对被分割为规定个数的频带分别推定噪声频谱。

[0057] [式 1]

$$[0058] \quad N_t(n) = \begin{cases} \beta N_{t-1}(n) + (1 - \beta)Y(n); & \text{case} \cdot A \\ N_{t-1}(n); & \text{case} \cdot B \end{cases} \quad \dots\dots (1)$$

[0059] 在这里, $N_t(n)$ 是当前处理中某帧的噪声频谱推定值, $N_{t-1}(n)$ 是其前一帧的噪声频谱推定值(因此,“t”是表示当前处理中的帧自身的下标), $Y(n)$ 是输入的振幅频谱,n 是频带(对频带标注的编号。此外,频带被分割为 N 个。此外,该 N 小于或等于本发明中所称的“K 个频带”中的 $K(=N \leq K)$ 。), β 是平滑化系数。另外,在式(1)中,case·A 表示噪声频谱推定部 20 处理噪声帧的情况,case·B 表示其处理音频帧的情况。

[0060] 如上述所示,噪声频谱推定部 20 与当前处理的帧是噪声帧还是音频帧对应地,变更为了计算噪声频谱推定值 $N_t(n)$ 而使用的公式。即,在音频帧处理时(case·B),直接使

用其前一个噪声频谱推定值, 求出噪声频谱推定值 $N_t(t)$, 在噪声帧处理时 (case • A), 通过使输入的振幅频谱在时间轴上平滑化, 计算噪声频谱推定值 $N_t(n)$ 。

[0061] 噪声抑制增益运算部 30 基于上述振幅频谱和通过式 (1) 求出的噪声频谱推定值 $N_t(n)$, 计算噪声抑制增益。在实施方式 1 中, 特别地, 通过以下的式 (2) 计算噪声抑制增益。

[0062] [式 2]

$$[0063] \quad G(n) = \max\left(0, \frac{Y(n) - N_t(n)}{Y(n)}\right) \dots\dots (2)$$

[0064] 在这里, $\max(a, b)$ 表示将 a 和 b 中较大的那一个的值返回的函数 (以下相同。)

[0065] 通过该式 (2), 在输入的振幅频谱 $Y(n)$ 与所对应的噪声频谱推定值 $N_t(n)$ 之间, $Y(n) < N_t(n)$ 成立的情况下, $G(n) = 0$, 在 $Y(n) > N_t(n)$ 成立的情况下, $G(n) = (Y(n) - N_t(n)) / Y(n)$ 。

[0066] 将通过该噪声抑制增益运算部 30 计算出的噪声抑制增益, 分别与通过上述音频检测部 80 区分的音频帧及噪声帧对应地, 经由噪声期间 • 噪声抑制增益运算部 40 或者直接向原音附加增益运算部 60 供给。图 1 所示的噪声抑制装置 1, 具有用于实现上述处理的开关 (参照图中的弧线箭头)。

[0067] 噪声期间 • 噪声抑制增益运算部 40 (以下, 为了简单, 也称为“噪声期间用增益运算部 40”。), 计算应向噪声帧应用的噪声抑制增益。在实施方式 1 中, 为了计算该噪声抑制增益而采用以下方法。

[0068] 首先, 基于通过式 (2) 求出的噪声抑制增益 $G(n)$, 计算由以下的式 (3) 表示的 g 。

[0069] [式 3]

$$[0070] \quad g = \frac{1}{N} \sum_{n=0}^{N-1} G(n) \dots\dots (3)$$

[0071] 从式 (3) 的右边明确可知, 该 g 表示式 (2) 的噪声抑制增益的与频带 n 相关的平均值。

[0072] 接下来, 通过以下的式 (4), 使该式 (3) 的噪声抑制增益平均值 g 平滑化。

[0073] [式 4]

$$[0074] \quad G_t = \mu G_{t-1} + (1 - \mu) g \dots\dots (4)$$

[0075] 在这里, μ 是平滑化系数, G_t 是针对当前处理中的噪声帧的噪声抑制增益, G_{t-1} 是针对其前一个处理的噪声帧的噪声抑制增益。

[0076] 上述式 (1) 中的作为 case • A 而示出的式子也是如此, 但在该式 (4) 中, 由于在计算当前处理中的帧的噪声抑制增益时, 参照其前一个处理的帧的噪声抑制增益, 所以可以说在时间轴上进行了平滑化 (对于后述的式 (7) 也相同。)

[0077] 该式 (4) 中的 G_t , 是应在本噪声期间用增益运算部 40 中求出的、用于向噪声期间应用的噪声抑制增益 (以下, 为了简单, 也称为“噪声期间用增益”。)。

[0078] 噪声期间用增益运算部 40 将按照上述方式求出的噪声期间用增益 G_t , 统一应用于全部频带。以下, 为了表示该情况, 将该统一应用的 G_t , 表示为 $G1(n)$ 。在此情况下, $G1(0)$ 、 $G1(1)$ 、 $\dots$ 、 $G1(N-1)$, 全部都等于 G_t 。

[0079] 原音附加率运算部 50 计算与噪声抑制后的信号相对的原音信号的原音附加率。在实施方式 1 中, 特别地, 基于以下的式 (5) 计算该原音附加率 og 。

[0080] [式 5]

[0081] $og = \max(0, tg - G_t) \cdots \cdots (5)$

[0082] 在这里, tg 是目标噪声抑制增益, 是基于以下的式 (6) 而求出的。

[0083] [式 6]

[0084] $tg = 10^{\frac{-TG}{20}} \cdots \cdots (6)$

[0085] 该式 (6) 中的 TG 是目标噪声抑制量, 以 dB 为单位。该 TG (或者 tg) 通过从装置外部经由未图示的操作部等而人为地提供, 或者, 也可以通过任意适当的方法自动地运算。

[0086] 根据以上的式 (5), 在目标噪声抑制增益 tg 和噪声期间用增益 G_t 之间, $tg < G_t$ 成立的情况下, $og = 0$, 在 $tg \geq G_t$ 成立的情况下, $og = tg - G_t$ 。

[0087] 原音附加增益运算部 60 基于上述原音附加率 og , 计算原音附加后的噪声抑制增益。在实施方式 1 中, 为了计算该噪声抑制增益而采用以下方法。

[0088] 首先, 基于通过式 (5) 求出的原音附加率 og , 计算由以下的式 (7) 表示的 OG_t 。

[0089] [式 7]

[0090] $OG_t = \begin{cases} \lambda OG_{t-1}(n) + (1 - \lambda)og; & \text{case} \cdot A \\ OG_{t-1}; & \text{case} \cdot B \end{cases} \cdots \cdots (7)$

[0091] 在这里, OG_t 是当前处理中的帧的原音附加比例。 OG_{t-1} 是其前一个帧的原音附加比例, λ 是平滑化系数。此外, 式 (7) 中的 $\text{case} \cdot A$ 及 $\text{case} \cdot B$ 的含义, 与上述式 (1) 的情况相同 (在以下的式 (8) 中也相同。)。

[0092] 如上述所示, 原音附加增益运算部 60 与当前处理的帧是噪声帧还是音频帧对应地, 变更为了计算原音附加比例 OG_t 而使用的公式。即, 在音频帧处理时 ($\text{case} \cdot B$), 直接使用其前一个原音附加比例来计算原音附加比例 OG_t , 在噪声帧处理时 ($\text{case} \cdot A$), 通过使上述原音附加率 og 在时间轴上平滑化来计算原音附加比例 OG_t 。

[0093] 接下来, 原音附加增益运算部 60 基于以下的式 (8), 计算原音附加后的噪声抑制增益。

[0094] [式 8]

[0095] $G2(n) = \begin{cases} OG_t(n) + (1 - OG_t) \cdot G1(n); & \text{case} \cdot A \\ OG_t(n) + (1 - OG_t) \cdot G(n); & \text{case} \cdot B \end{cases} \cdots \cdots (8)$

[0096] 在这里, $G1(n)$ 如以上说明所示, 表示在噪声帧中对所有频带统一应用的噪声期间用增益。

[0097] 根据该式 (8), 与上述式 (7) 中的情况分别对应地, 计算原音附加后的噪声抑制增益 $G2(n)$ (以下, 为了简单, 也称为“修正后增益 $G2(n)$ ”)。

[0098] 图 1 所示的乘法器 11, 使按照上述方式求出的修正后增益 $G2(n)$ 与振幅频谱 $Y(n)$ 相乘。即, 进行 $S(n) = G2(n) \cdot Y(n)$ 的运算, 其结果是, 可以得到最终应得到的噪声抑制后的振幅频谱 $S(n)$ 。

[0099] 最后, 频率·时间变换部 70, 基于按照上述方式求出的噪声抑制后的振幅频谱 $S(n)$ 和从时间·频率变换部 10 直接供给的相位频谱, 生成时域的输出信号。在实施方式 1 中, 由于在时间·频率变换部 10 中进行傅里叶变换, 所以频率·时间变换部 70 实施傅里叶逆变换。

[0100] 下面, 在已参照过的图 1 的基础上, 参照图 2 至图 4, 说明以上所述的实施方式 1 所

涉及的噪声抑制装置 1 的作用、动作以及效果。

[0101] 首先,时间·频率变换部 10 对输入信号实施傅里叶变换,并将其如图 1 所示分解为振幅频谱 $Y(n)$ 及相位频谱(图 2 的步骤 S101)。此时,时间·频率变换部 10 如上述所示,针对每个帧实施处理。

[0102] 另外,与此并行,音频检测部 80 对输入信号中是否包含音频信号进行检测(图 2 的步骤 S102)。该检测处理是可以将输入信号区分为音频帧和噪声帧的处理。音频检测部 80 还进行相关的处理。

[0103] 然后,噪声频谱推定部 20 通过上述振幅频谱 $Y(n)$ 及上述式 (1),针对每个具有规定宽度的频带 n ,计算出噪声频谱推定值 $N_t(n)$ 。在此情况下,如上述所示,与当前处理的帧是噪声帧还是音频帧对应地,进行不同的处理(参照图 2 的步骤 S103)。此外,如图 2 所示,在从该噪声频谱推定值 $N_t(n)$ 的计算处理以后,直至由图 1 所示的乘法器 11 进行的输出信号生成处理(图 2 的步骤 S104)为止的期间,与噪声帧和音频帧之间的区别对应地,展开其内容实质不同的处理。因此,下面,第一、针对噪声帧用的处理,第二、针对音频帧用的处理,分为 [I] 及 [II] 进行说明。

[0104] 此外,上述的分别处理如图 1 所示,是通过与音频检测部 80 的检测结果对应的开关的切换而进行的。

[0105] [I] 首先,在噪声帧用处理中,通过上述式 (1) 中的作为 case • A 示出的公式,计算出噪声频谱推定值 $N_t(n)$ (图 2 的步骤 S201)。如上述所示,这是通过输入的振幅频谱 $Y(n)$ 的平滑化处理进行的。

[0106] 然后,基于上述噪声频谱推定值 $N_t(n)$ 以及上述式 (2),计算出噪声抑制增益 $G(n)$ (图 2 的步骤 S202)。这是通过图 1 中的噪声抑制增益运算部 30 的作用而进行的。如上述所示,在 $Y(n) > N_t(n)$ 成立的情况下,则成为 $G(n) = (Y(n) - N_t(n)) / Y(n)$,在其不成立的情况下,则成为 $G(n) = 0$ 。由此,可以得到例如图 3(C) 所示的噪声抑制增益(此外,分别在图 3(B) 中例示出上述噪声频谱推定值 $N_t(n)$,在图 3(A) 中例示出输入信号的振幅频谱。)

[0107] 然后,通过上述式 (3) 及式 (4),得到该噪声抑制增益 $G(n)$ 的与频带相关的平均值 g ,且通过进行针对该 g 的平滑化处理,计算出噪声期间用增益 G_t (图 2 的步骤 S203)。该经过平均化·平滑化而获得的噪声期间用增益 G_t ,成为所有频带共用的 $G1(n)$ 。这是通过噪声期间用增益运算部 40 的作用而进行的。

[0108] 如上述所示,在实施方式 1 中,其一大特征是,不是将通过式 (2) 求出的噪声抑制增益 $G(n)$ 在原状态下直接使用,而是将针对该 $G(n)$ 进行了式 (3) 的与频带相关的平均化、以及式 (4) 的时间轴上的平均化后的噪声期间用增益 G_t ,作为所有频带用的噪声期间用增益 $G1(n)$ 而使用。

[0109] 此外,在图 3(D) 中,例示出针对噪声抑制增益 $G(n)$ 进行平均化处理的情况的一个例子(也可参照图 3(C) 中所示的虚线)。

[0110] 然后,通过上述噪声期间用增益 G_t 以及上述式 (5),计算出原音附加率 og (图 2 的步骤 S204)。这是通过图 1 中的原音附加率运算部 50 的作用而进行的。在这里,目标噪声抑制增益 tg 乃至目标噪声抑制量 TG 的设定如何,作为一个主要支配因素起作用。即,如果噪声期间用增益 G_t 与目标噪声抑制增益 tg 相比较,则将原音附加率 og 设定为 0,如果不是这样,则设定与噪声期间用增益 G_t 对应的原音附加率 og (即, $og = tg - G_t$)。将这两种情

况分开处理的意义在于,确定根据与目标噪声抑制量 TG 的关系,以何种程度得到通过加上原音而带来的音质改善的效果。即,在后者的情况下,以在通过目标噪声抑制量确定的框内(即, t_g 和 G_t 的差所对应的部分)加上原音,实现音质改善为主要目的,在前者的情况下,由于 $G_t > t_g$ 成立,已经没有改善音质的余地,所以将原音附加率 og 设为 0(在此情况下,由此反而抑制了噪声量增多)。结果,上述式 (5) 及式 (6) 存在实现下述处理的意,即,在以遵守目标噪声抑制量为基准,此外,存在附加原音的余地的情况下,在该框内实现音质的改善。

[0111] 如上述所示,在实施方式 1 中,其一大特征是,原音附加率 og 是通过使用噪声期间用增益 G_t 而求出的。

[0112] 然后,通过上述原音附加率 og 以及上述式 (7) 中的作为 case • A 示出的公式,计算出原音附加比例 OG_t (图 2 的步骤 S205)。该原音附加比例 OG_t 如上述所示,是通过将原音附加率 og 在时间轴上平滑化而求出的。然后,通过按照上述方式求出的原音附加比例 OG_t 以及上述式 (8),计算出原音附加后的噪声抑制增益,即修正后增益 $G2(n)$ 。以上是通过原音附加增益运算部 60 的作用而进行的。

[0113] 在此情况下,该修正后增益 $G2(n)$ 最终具有下述意义,即,是在考虑了上述经过平均化 • 平滑化的噪声期间用增益 $G1(n)$ 和原音附加的程度的基础上确定的增益。

[0114] 此外,优选将装置刚起动后的情况考虑在内,适当地确定作为与上述式 (1) 中的 $N_{t-1}(n)$ 对应的值的初始值(作为上述初始值的 $N_{t-1}(n)$,当然,也可以在后述的音频帧用处理中的噪声频谱推定值 $N_t(n)$ 的计算处理中使用。)。上述式 (4)、式 (7) 中的 $G_{t-1}(n)$ 也与此相同。

[0115] [II] 另一方面,在音频帧用处理中,基本上执行与上述噪声帧处理大致相同的各处理。即,计算出噪声频谱推定值 $N_t(n)$ 及基于其的噪声抑制增益 $G(n)$ (参照图 2 的步骤 S301 • S202),基于原音附加比例 OG_t 计算出修正后增益 $G2(n)$ (图 2 的步骤 S303 • S304),这些与噪声帧处理相同。

[0116] 但是,在该音频帧处理中,与噪声帧处理相比,存在以下所述的不同或者注意点。

[0117] (i) 噪声频谱推定值 $N_t(n)$ 不是通过上述式 (1) 中的作为 case • A 示出的公式求出的,而是通过作为 case • B 示出的公式求出的(图 2 的步骤 S301)。由于该式为 $N_t(n) = N_{t-1}(n)$,所以可以说音频帧处理是所谓维持现状的处理。再稍微详细地说,在该音频帧之前是噪声帧的情况下,将对该噪声帧计算出的噪声频谱推定值 $N_{t-1}(n)$,直接在该音频帧处理中使用,另外,另一方面,在该音频帧之前是音频帧、且在该之前的音频帧之前是噪声帧的情况下,将对该噪声帧计算出的噪声频谱推定值 $N_{t-2}(n)$,直接在该音频帧处理中使用等。

[0118] 总之,在音频帧中,使用对最接近的噪声帧计算出的噪声频谱推定值 $N_{t-p}(n)$ (p 是从该音频帧的前一个帧开始数,直至该最接近的噪声帧为止的帧数(包含两端))。

[0119] (ii) 在使用上述式 (7) 进行的原音附加比例 OG_t 的计算处理中,也可以说是与此相同的。即,由于式 (7) 中的作为 case • B 示出的公式为 $OG_t = OG_{t-1}$,所以在此情况下,在音频帧处理中,还是所谓维持现状(参照图 2 的步骤 S303)。

[0120] 如果使表达与上述情况一致,则在音频帧中,使用对最接近的噪声帧计算出的原音附加比例 $OG_{t-p}(n)$ (p 为从该音频帧的前一个帧开始数,直至该最接近的噪声帧为止的帧数(包含两端))。

[0121] (iii) 噪声抑制增益 $G(n)$ 的运算本身,与区分音频帧及噪声帧无关,使用上述式 (2) 相同地进行。在图 2 的步骤 S202 中,将与 [噪声帧用处理] 和 [音频帧用处理] 对应的边框连接而描述,是为了象征地表示该特征 (但是,式 (2) 中的 $N_t(n)$ 的值,分别与式 (1) 中的 case • A 及 case • B 对应,对于两种帧当然是不同的。)。

[0122] (iv) 在音频帧处理中,不进行与上述式 (3) 及式 (4) 相关的处理、即针对噪声抑制增益 $G(n)$ 的平均化 • 平滑化处理 (参照图 2 的步骤 S203 及该图中的右方)。另外,与此相伴,由于成为所谓不存在正当的噪声期间用增益 G_t 的状态,所以也不进行与上述式 (5) 相关的处理、即计算原音附加率 og 的处理 (参照图 2 的步骤 S204 及该图中的右方)。

[0123] (v) 最终计算出的修正后增益 $G2(n)$,不是通过上述式 (8) 中的作为 case • A 示出的公式求出的,而是通过作为 case • B 示出的公式求出的 (图 2 的步骤 S304)。在此情况下,不同点在于,在噪声帧处理时,使用经过平均化 • 平滑化的噪声期间用增益 $G1(n)$,而在音频帧处理时,直接使用根据式 (2) 求出的噪声抑制增益 $G(n)$ 。

[0124] 经过以上的 [I] 及 [II] 的处理,都可以得到修正后增益 $G2(n)$,如果使该修正后增益 $G2(n)$ 乘以原来的振幅频谱 $Y(n)$,则计算出噪声抑制后的振幅频谱 $S(n)$ (图 2 的步骤 S104)。

[0125] 在图 3(E) 中,为了简单,单纯地示出使图 3(A) 的振幅频谱 $Y(n)$,乘以图 3(C) 的平均化后的噪声抑制增益 (即, g) 而得到的结果。在实施方式 1 中,如上述所示,除此以外,还进行考虑原音附加的程度在内的增益调整 (参照式 (8),特别是 $OG_t(n)$ 的作用),但图 3(E) 充分表示出假设省略对上述原音附加处理的考虑的情况时的处理的本质 (在式 (8) 中仅表现为,如果 $OG_t(n) = 0$,修正后增益 $G2(n)$ 就简单地等于 $G1(n)$ 或 $G(n)$ 。)。

[0126] 根据具有以上所述的结构及作用的噪声抑制装置 1,可以得到以下效果。

[0127] 首先,根据实施方式 1 中的噪声抑制装置 1,可以非常合适地抑制包含在输入信号中的噪声。在这里,在所称的“合适地”中,就实施方式 1 而言,特别地包含以下记述的各项事实。

[0128] (1) 第一,根据实施方式 1,可以非常有效地防止所谓音乐噪声的产生。在这里,所谓音乐噪声,是指在从输入信号的振幅频谱中减去噪声频谱推定值后产生的噪声。

[0129] 例如,基于噪声频谱推定值的噪声抑制增益,可以简单地使用上述式 (2) 中的 $(Y(n)-N(n))/Y(n)$ 而求出,如果假设将其直接在图 1 所示的乘法器 11 中应用的状态,则噪声抑制后的振幅频谱 $S(n)$,作为 $S(n) = \{(Y(n)-N(n))/Y(n)\} \cdot Y(n) = Y(n)-N(n)$ 而被求出。即,在此情况下,通过单纯地从输入信号的振幅频谱中减去噪声频谱推定值,得到噪声抑制后的振幅频谱 $S(n)$ 。

[0130] 但是,由于该情况下的噪声频谱推定值只不过是“推定值”,所以不一定反映实际的噪声频谱。因此,存在在某频带中,减去噪声频谱推定值后还残留噪声的情况,另外,还存在在其他频带中,产生削减过度的情况 (在该削减过度的情况下,由于不存在负的振幅频谱,所以将其设定为 0。)。在图 4 中,将上述情况概念化地进行表示,例如,图 4(C) 中的实线表示减去后还有剩余 (参照标号“KN”),虚线表示削减过度 (参照标号“HS”) 的各情况 (此外,图 4(A) 及 (B) 与图 3(A) 及 (B) 完全相同。另外,图 4(C) 的标号 HS 所指代的部分,是 $Y(n)-N(n) = 0$ 恰巧成立的情况的例示。)。

[0131] 如果将上述振幅频谱 $S(n)$ 向时域进行傅里叶逆变换,则该信号成为由多个具有

随机频率的正弦波合成的信号,如果对其进行播放,则会听到非常刺耳的声音。这就是音乐噪声。

[0132] 如上述所示,产生音乐噪声的主要原因,严格来讲是不可知的实际的噪声频谱与噪声频谱推定值不一致。

[0133] 在实施方式 1 中,非常有效地抑制了上述音乐噪声的产生。其原因是,在噪声帧处理时,使用平均化·平滑化后的噪声期间用增益 G_t 计算出修正后增益 $G2(n)$,并将其应用于振幅频谱 $Y(n)$ (参照图 3(E))。由此,在维持原来的振幅频谱所具有的频率结构的状态下,进行噪声抑制,所以极难产生音乐噪声。

[0134] (1-i) 此外,在计算噪声期间用增益 G_t 时进行的平均化(上述式(3))及平滑化(上述式(4)),分别具有固有的意义。前者的目的,从图 3 中明确可知,主要在于得到上述音乐噪声的抑制效果,后者的目的主要在于从所谓历时的角度来说,维持噪声抑制处理的连续性。根据后者,由于噪声期间用增益 $G_t(n)$ 不会随着时间经过而发生急剧的变化,所以在例如对包含在该噪声帧中的信号进行播放的情况下,不会给收听者造成听觉上的不舒适感(此外,在实施方式 1 中进行的其他的平滑化处理(即,式(1)中的 case·A,式(7)中的 case·A),原则上具有与此实质相同的意义。))。

[0135] (2) 第二,上述(1)是关于预防与噪声帧处理相关的音乐噪声的产生,与此相关联,根据实施方式 1,可以更可靠地实现预防与音频帧处理相关的音乐噪声的产生。这是如上述所示,通过在音频帧处理中,以所谓直接使用不经过平均化·平滑化的噪声抑制增益 $G(n)$ (参照式(2))的形式来计算修正后增益 $G2(n)$ (式(8)中的 case·B,或者上述的 [II] (v))而实现的。

[0136] (3) 而且,第三,根据实施方式 1,在从噪声帧向音频帧切换的情况下,可以保持噪声抑制处理的一贯性。这是如上述所示,通过在音频帧处理时,作为噪声频谱推定值 $N_t(n)$,使用对最接近的噪声帧计算出的 $N_{t-p}(n)$ (参照上述的 [II] (i) 的记载)而实现的。

[0137] 归纳以上的(2)及(3),在实施方式 1 中,对以下方面下了工夫,即,不但在音频帧中进行有效的噪声抑制,此外也尊重噪声帧处理时的噪声抑制处理(特别是其效果),从而使两种帧之间的转换变得更自然。由此,在将实施方式 1 中的噪声抑制装置 1 与某音频播放单元连接的情况下,在从噪声帧向音频帧切换的情况下,不会给收听者造成噪声所涉及的音量感的变化等听觉上的不舒适感。

[0138] 此外,为了抑制音频帧中的音乐噪声,也考虑下述方法,即,取代上述 $S(n) = Y(n) - N(n)$,而采用 $S(n) = Y(n) - \alpha N(n)$,并使该 $\alpha (> 0)$ 的值较大,但其中存在以下缺点,即,使音质急剧恶化的可能性很高。但是,如果减小 α ,则使音乐噪声的抑制不充分。

[0139] 另外,也考虑通过采用以下方法抑制音乐噪声,即,向图 4(C) 中的虚线表示的、噪声抑制后的振幅频谱成为 0 的部分(即,标号 HS 及 HSt 所指代的部分)中,加上固定的值(噪声层(noise floor))。这是出自下述想法,即,想要通过所谓使该部分 HS 及 HSt 增高,掩盖减去后的剩余部分 KN(或者使其不明显)(另外,由于在将该方法和使用上述 α 的方法并用的情况下,可以更小地设定 α ,所以在此情况下,也可以得到防止音质恶化的效果。))。

[0140] 但是,由于附加上述噪声层,就意味着使噪声的绝对量增加,所以不仅从达到抑制噪声这一本来目的的角度来看存在问题,而且存在随着该噪声层量的设定的不同,很可能使噪声抑制效果变得极不充分的问题。

[0141] 从上述角度出发,也可以明确实施方式 1 中的噪声抑制装置 1 非常具有优势。即,在实施方式 1 中,由于没有如使用上述 α 那样,使减去的量机械地增加,所以几乎不可能产生音质恶化,另外,由于不是进行上述噪声层的单纯附加这样的处理,所以不存在牺牲所进行的噪声抑制的效果的情况。另外,如已记述的那样,即使如此,也可以有效地抑制音乐噪声。

[0142] (4) 根据实施方式 1 中的噪声抑制装置 1,由于如参照上述式 (5) ~ 式 (7)、或者图 2 的步骤 S205 及 S303 所说明的那样进行原音附加处理,所以可以更有效地实现噪声抑制效果。根据该原音附加处理,由于可以期待与上述的噪声层的附加处理相同的效果,即,对图 4(C) 中的减去后的残留部分 KN 的掩盖效果,所以可以更有效地抑制音乐噪声、或者防止音质的恶化(但是,上述噪声层始终是“固定”的。这一点是与利用“原音”的情况之间的决定性的不同。)

[0143] 此外,在上述中,为了更加明确地把握通过实施方式 1 中的噪声抑制装置 1 实现的效果,在与使用上述 α 的方法或者使用噪声层的方法进行对比中,存在对该效果进行说明的部分,但本发明不具有积极地排除这些使用 α 或者噪声层来抑制音乐噪声的方法的意图。即,可以将这些方法和本发明及其各种形式并用,根据上述的并用方式,可以汲取该方法的优点,同时,使本发明及其各种方式的效果更加显著。

[0144] 而且,在实施方式 1 中,不是简单地执行原音附加,具有以下几点特征。

[0145] (4-i) 首先,原音附加的比例(即, OG_t)是基于原音附加率 og 而确定的,该原音附加率 og 是与噪声期间用增益 G_t 和目标噪声抑制增益 tg 的大小情况对应地确定的。具体地说,如已记述的那样,由于在原音附加处理中,将作为目标的噪声抑制的程度(即, tg)作为一个主要支配因素,通过与其之间的关系确定原音附加率 og ,所以通过在基于噪音期间用增益 G_t 的处理和原音附加处理之间,平衡地进行使用,可以更有效地得到噪声抑制效果及音乐噪声抑制效果,甚至音质改善效果。

[0146] (4-ii) 另外,对于上述的原音附加处理,在音频帧处理时,作为原音附加比例 OG_t ,使用对最接近的噪声帧计算出的 OG_{t-p} (参照上述的 [II] (ii) 的记载)。这一作法的本质与作为某音频帧的噪声频谱推定值 $N_t(n)$,直接使用其前一个帧的噪声频谱推定值 $N_{t-1}(n)$ 这样的上述想法相同。即,在该原音附加处理中,在噪声帧及音频帧之间的切换的情况下,保持噪声抑制处理的一贯性。

[0147] <实施方式 2>

[0148] 以下,参照图 5 至图 7,说明本发明所涉及的实施方式 2。此外,该实施方式 2,在与上述实施方式 1 的对比中,存在与音频检测处理相关的不同点,而对于其他方面,只要不特别地说明,则与上述实施方式 1 完全相同。因此,下面,主要针对上述不同点进行说明,简化甚至省略对其他方面的说明。另外,对于附图上的标号,除了上述不同点以外沿用标号。

[0149] 该实施方式 2 中的噪声抑制装置 1' 如图 5 所示,具有将音频检测部 801 连接在噪声抑制增益运算部 30 的后段的结构。即,该音频检测部 801 通过使用由上述式 (2) 计算出的噪声抑制增益 $G(n)$,对输入信号中有无音频信号进行检测,或者对音频帧和噪声帧进行区分。

[0150] 在实施方式 2 中,为了对有无音频信号进行检测,而采用以下方法。

[0151] 首先,基于通过式 (2) 求出的噪声抑制增益 $G(n)$,计算出由以下的式 (9) 表示的

Var。

[0152] [式 9]

$$[0153] \quad Var = \frac{1}{N} \sum_{n=0}^{N-1} (G(n) - g)^2 \dots\dots (9)$$

[0154] 在这里, g 是通过在上述实施方式 1 中使用的式 (3) 表示的 g , 总而言之, 是针对 $G(n)$ 的与频带 n 相关的平均值 (在实施方式 2 中, 不仅通过噪声期间用增益运算部 40, 而且也通过音频检测部 801 进行该 g 的运算。当然, 也可以将由两者中的一方得出的运算结果, 在两者间共用。)

[0155] 从公式明确可知, 该式 (9) 的 Var 表示 $G(n)$ 的方差。

[0156] 然后, 对该 Var 是否超过规定值进行判断。该判断的意义如下。

[0157] 通常, 通过式 (2) 计算出的噪声抑制增益 $G(n)$, 在包含音频信号的情况下和不包含音频信号的情况下, 表现出大不同的形态。图 6 及图 7 表示其一个例子, 前者是包含音频信号的情况下的噪声抑制增益 $G(n)$ 的运算例, 后者是不包含音频信号的情况下的噪声抑制增益 $G(n)$ 的运算例。将这些图进行对比明确可知, 如果计算两种情况下的各自的 $G(n)$ 的方差, 则容易地推测出两者间产生较大的差距这一结论。即, 可以以相当程度的可靠性进行以下判断, 即, 如果针对某帧的 $G(n)$ 的方差值大到一定程度, 则其包含音频信号, 如果不是这样, 则不包含音频信号。

[0158] 针对上述 Var 的大小判断的意义就在于此。换言之, 假设存在某个规定值 VB , 如果 $Var > VB$, 则在该帧中存在音频信号, 因此, 将其区分为“音频帧”, 如果 $Var \leq VB$, 则该帧中没有音频信号, 因此, 将其区分为“噪声帧”。

[0159] 此外, 在图 5 的结构中, 与图 1 的结构不同, 噪声频谱推定部 20 无法使用音频信号的检测结果。即, 噪声频谱推定部 20 不以区分音频帧及噪声帧为前提, 而计算噪声频谱推定值 $N_t(n)$ 。

[0160] 上述情况下的噪声频谱推定值 $N_t(n)$, 可以通过例如以下的式 (10) 及式 (11) 求出。

[0161] [式 10]

$$[0162] \quad PA_t(n) = \alpha PA_{t-1}(n) + (1 - \alpha) Y(n) \dots\dots (10)$$

[0163] [式 11]

$$[0164] \quad N_t(n) = \begin{cases} \gamma N_{t-1}(n) + \frac{1-\gamma}{1-\beta} (PA_t(n) - \beta PA_{t-1}(n)); & \text{case} \cdot C \\ PA_t(n); & \text{case} \cdot D \end{cases} \dots\dots (11)$$

[0165] 在这里, $PA_t(n)$ 是当前处理中的帧中的输入信号中的振幅频谱, 是平滑化后的振幅频谱, $PA_{t-1}(n)$ 是其前一个帧中的该振幅频谱, 是平滑化后的振幅频谱, α 是平滑化系数, $\gamma \cdot \beta$ 是控制参数。另外, 在式 (11) 中, case · C 表示 $PA_t(n) > N_{t-1}(n)$ 成立的情况, case · D 表示除此以外的情况。

[0166] 在此情况下, 式 (11) 中的作为 case · D 表示的公式与式 (10) 的组合, 实质上与上述式 (1) 中的作为 case · A 表示的公式大致同义。

[0167] 另一方面, 式 (11) 中的作为 case · C 表示的公式, 不是上述式 (1) 中所对应的那个。但是, 由于该式在如上述所示, 即, $PA_t(n) > N_{t-1}(n)$ 成立的情况下, 也就是说, 在当前处理中的帧的振幅频谱超过其前一个帧的噪声频谱推定值的情况下发挥作用, 所以并不是

不可以认为该 case •C 能够暗示出该当前处理中的帧是音频帧的可能性（如果对于大量的 $n (= 0, 1, 2, 3, \dots)$ 满足上述条件,则可以说其可能性更高。但是,毕竟只是“暗示”。)。

[0168] 可以说这些式 (10) 及式 (11) 在上述的意义范围内与上述式 (1) 具有共通性。

[0169] 总之,适当地计算出噪声频谱推定值这一点是不变的。

[0170] 如上述所示,根据实施方式 2 可以得到以下的效果。

[0171] 首先,可以明确通过本实施方式 2,也可以得到与通过上述实施方式 1 得到的作用效果没有本质区别的作用效果。即,在本实施方式 2 中,可以得到与针对上述实施方式 1 记述的从 (1) 至 (4) 的效果大致相同的效果。

[0172] 而且,将图 1 与图 5 进行对比明确可知,根据本实施方式 2 可以得到处理效率的提高、电路构成的简化等效果。这是因为,实施方式 1 中的音频检测是所谓独立地进行的,取代之,实施方式 2 中的音频检测是通过利用噪声抑制增益 $G(n)$ 而进行的,进行了其从属化。

[0173] 在本发明中,由于噪声抑制增益 $G(n)$ 的运算是必须进行的处理,所以利用其运算结果进行音频检测处理,自然可以获得处理的高效化・合理化。而且,其检测性能相当高(参照对比图 6 及图 7)。

[0174] 以上,对本发明所涉及的实施方式进行了说明,但本发明所涉及的噪声抑制装置,并不限于上述方式,可以进行各种变形。

[0175] (1) 在上述实施方式 1 及 2 中,使噪声期间用增益 G_t 在频率轴上平均化,在时间轴上平滑化,但本发明并不限于上述方式。在上面已经叙述,对于平均化处理和平滑化处理,由于其主要目的不同,所以特别是对于平滑化处理,可以根据情况而省略。如图 3(E) 所示,即使仅实施平均化处理,也可以一定程度地得到音乐噪声的抑制效果。

[0176] (2) 另外,在上述实施方式 1 及 2 中,噪声期间用增益 G_t 是经过上述式 (3) 的平均化处理及上述式 (4) 的平滑化处理而求出的,但本发明并不拘泥于上述式 (3) 及式 (4) 的形式。

[0177] 首先,在本发明中,噪声抑制增益平均值 g 并不限于通过式 (3) 进行计算的形式。

[0178] 即,虽然在式 (3) 中, g 是使用全部 N 个频带(全部为 N 个,即,第 $0, 1, 2, \dots, N-1$ 个频带)进行计算的,但该 g 也可以仅使用例如其中的一部分频带而进行计算。在此情况下,可以考虑使用除了极低频带(接近 DC 成分的频带)和极高频带(接近乃奎斯特频率的频带)这两者或者一方以外的频带。

[0179] 另外,在计算噪声抑制增益平均值 g 时,可以对各个频带进行不同的加权。例如,仅使特定的频带乘以某个特定的加权系数,或者,使所有的频带乘以连续地、阶段性地增加或者减少的加权系数,等等。

[0180] 另外,在本发明中,噪声期间用增益 G_t 并不限于通过上述式 (4) 进行计算的形式。

[0181] 即,虽然在式 (4) 中, G_t 是通过将噪声抑制增益平均值 g 在时间轴上平滑化而求出的,但该 G_t 也可以作为例如相邻的帧的 g 的平均值而计算出。

[0182] (3) 而且,在上述实施方式 1 及 2 中,虽然将平均化・平滑化的噪声期间用增益 G_t 及 $G1(n)$ 应用于所有频带(参照式 (8) 中的 case •A 或者图 3(E)),但本发明并不限于上述方式。

[0183] 例如,可以仅向上述除了极低频带和极高频带这两者或者一方以外的频带,应用该噪声期间用增益 G_t 及 $G_1(n)$ 。在此情况下,对于排除在该应用范围以外的频带,可以应用固定值的增益。

[0184] (4) 在上述实施方式 1 及 2 中,噪声抑制增益 $G(n)$ 是通过式 (2) 计算出的,但本发明并不限于上述方式。例如,除此之外,可以使用维纳滤波器法、MMSE (Minimum Mean-Square Error) 法等 (关于这些方法,参照上述的非专利文献 3 及 4)。也可以推定 SNR (音频 (信号)/噪声比例),基于该 SNR 计算噪声抑制增益 $G(n)$ 。

[0185] (5) 在上述实施方式 2 中,为了在音频帧及噪声帧之间进行区分,而通过上述式 (9) 得到噪声抑制增益 $G(n)$ 的在频率轴上的方差,但本发明并不限于上述方式。

[0186] 例如,当然可以取代方差而使用标准偏差,也可以使用时间轴上的方差或者标准偏差。另外,也可以基于针对每个频带的噪声抑制增益 $G(n)$ 中,有多少落在由规定的 2 个基准值划分成的空间内等,进行音频帧及噪声帧之间的区分 (例如,如果其数量较大,则可以判断为噪声抑制增益 $G(n)$ 集中存在于一定的位置上,所以可以说其分散的程度较小,因此,该帧被认定为噪声帧,等等。)。另外,上述的各种判断方法,可以根据情况而并用。由此,例如,在参照上述方差和落在上述空间内的噪声抑制增益 $G(n)$ 的数量这两者的基础上,对分散的程度进行判断等。

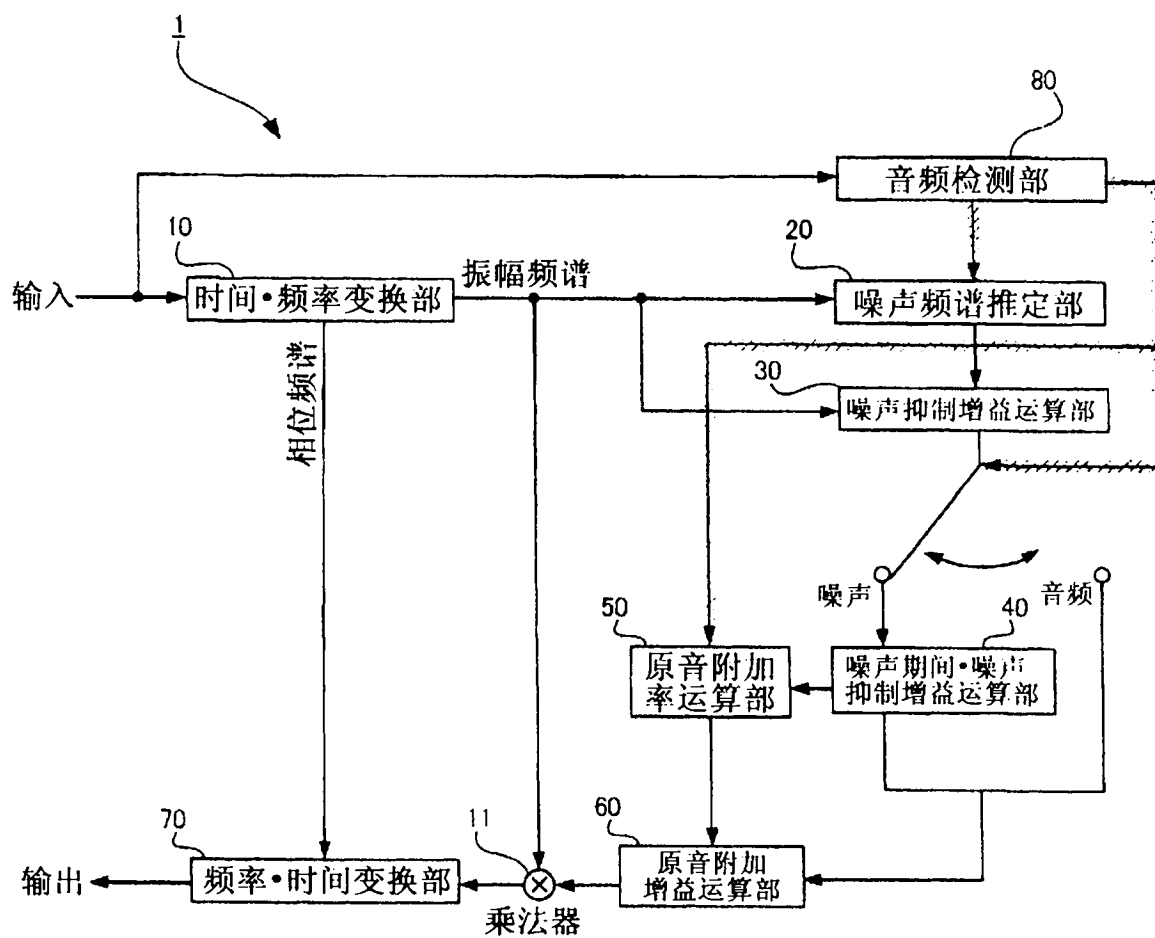


图 1

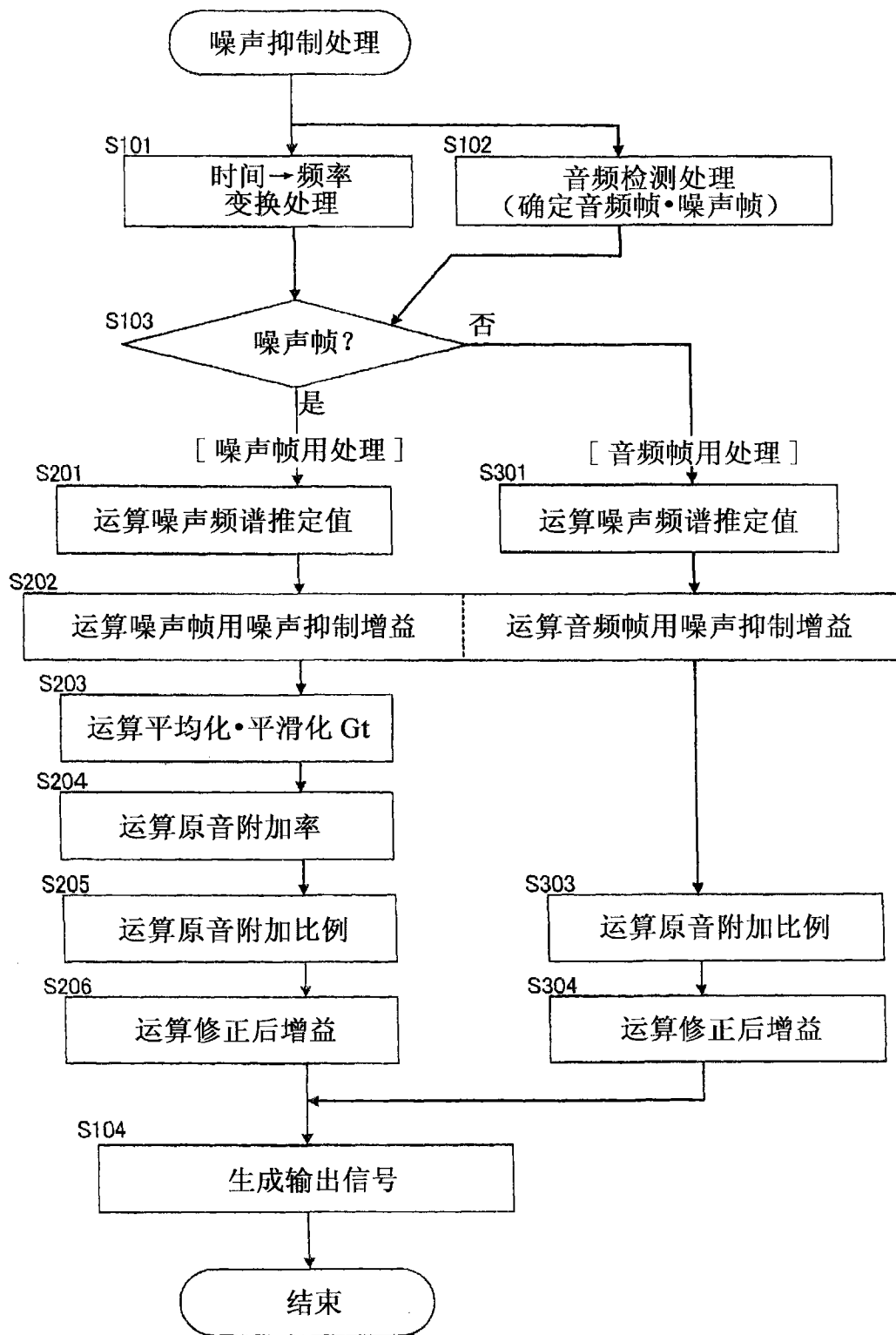
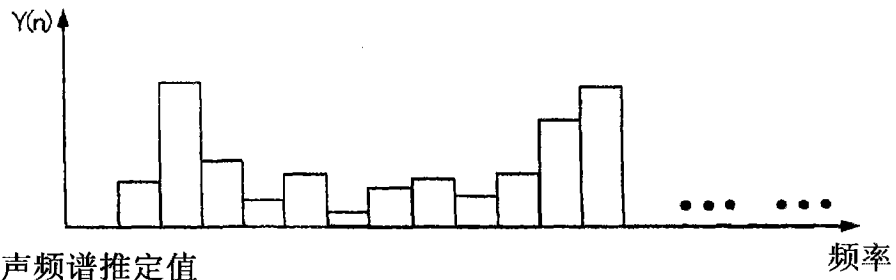
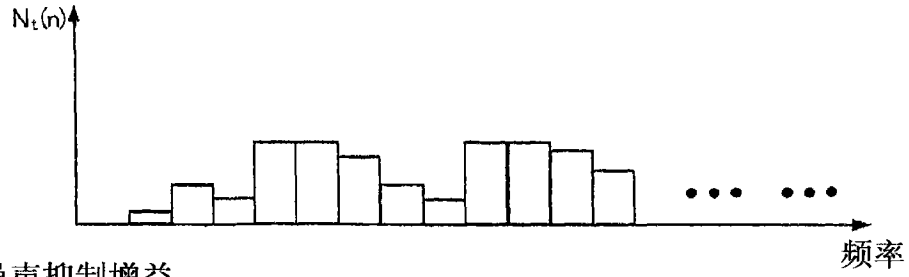


图 2

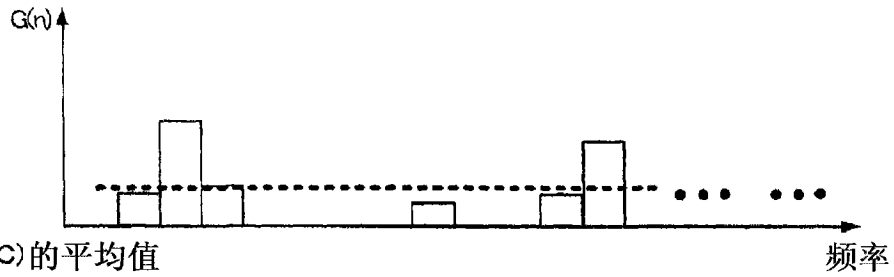
(A) 输入信号的振幅频谱



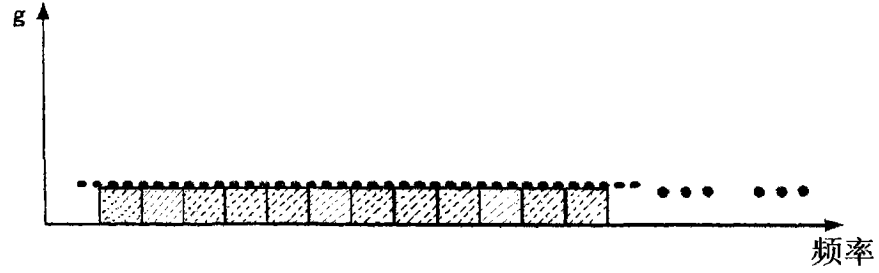
(B) 噪声频谱推定值



(C) 噪声抑制增益



(D) (C) 的平均值



(E) 噪声抑制后的振幅频谱 (=阴影线部分) (= (A) × (C))

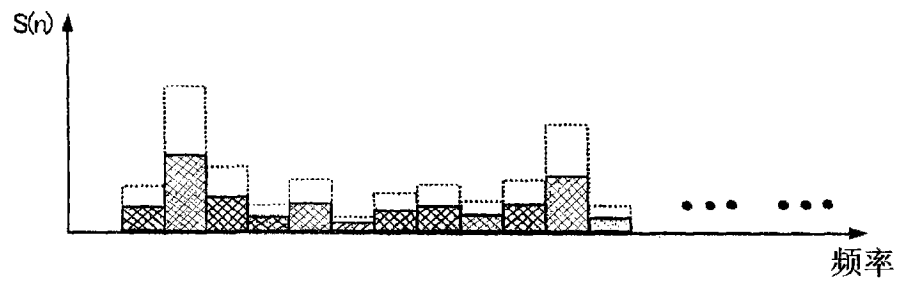


图 3

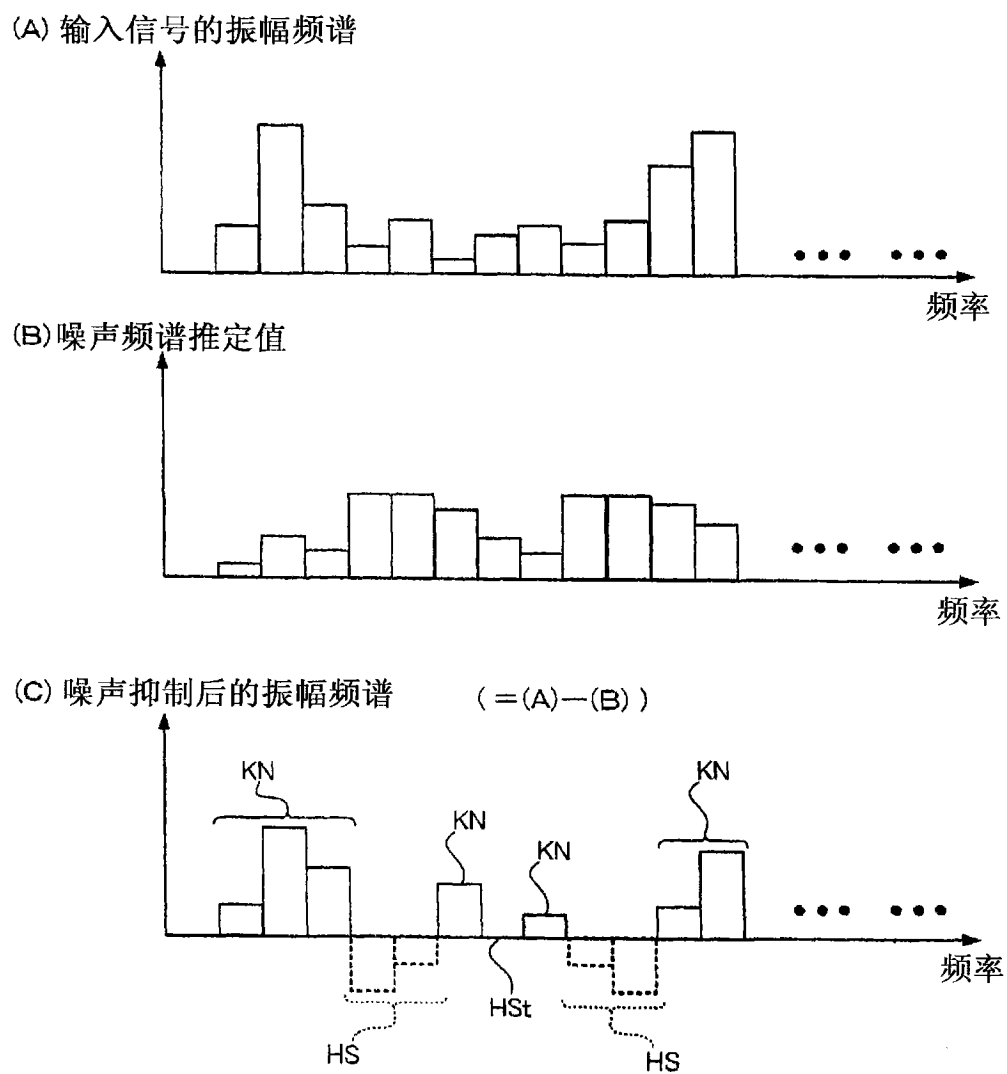


图 4

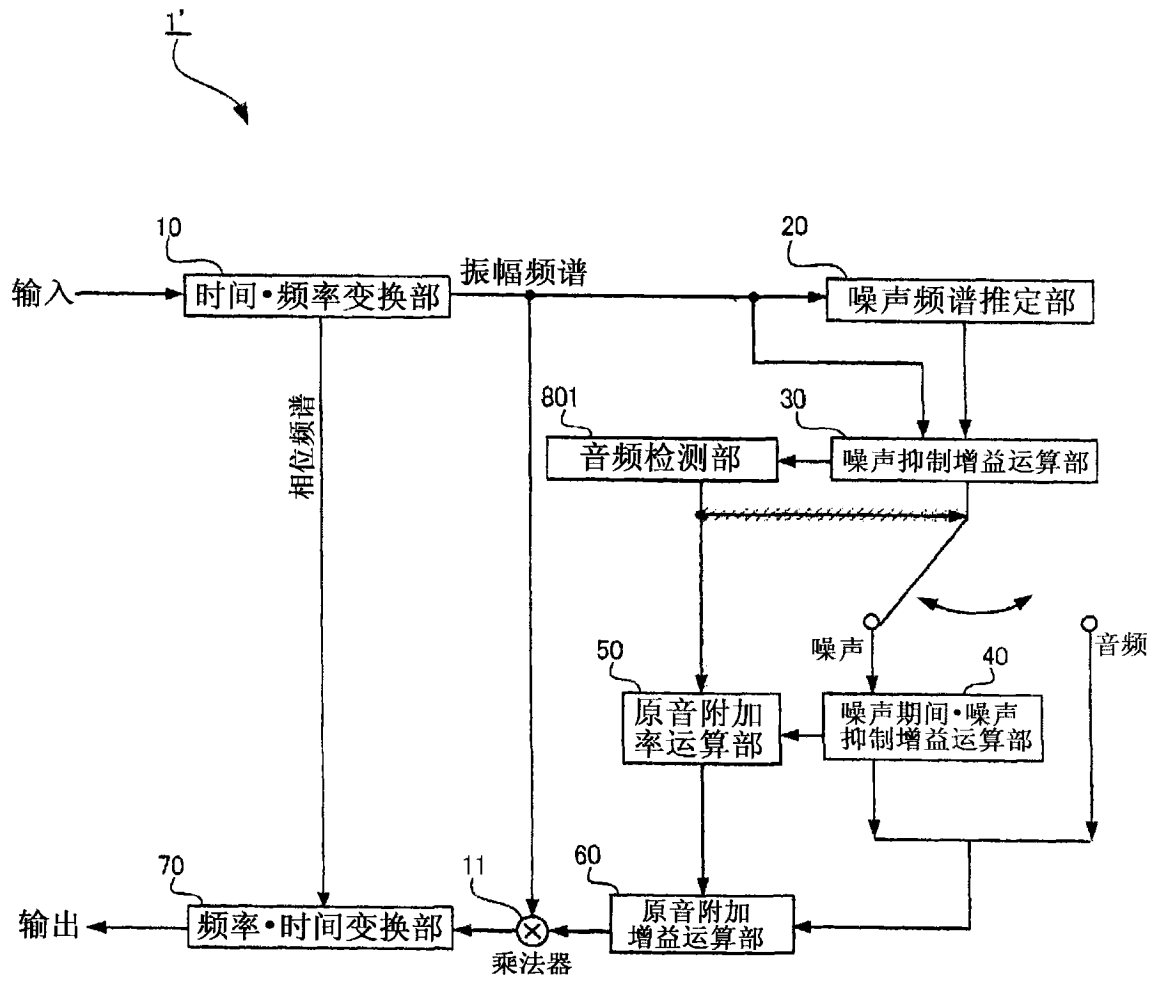


图 5

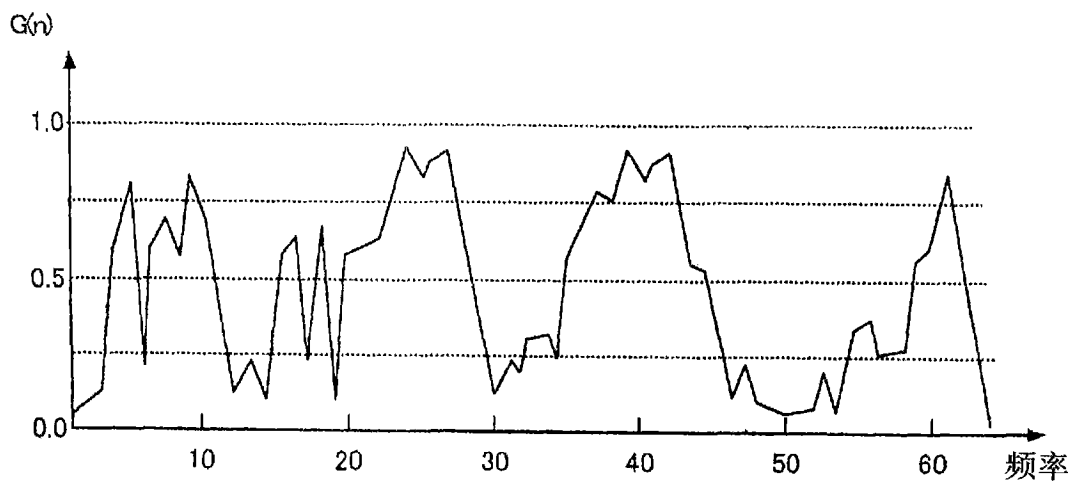


图 6

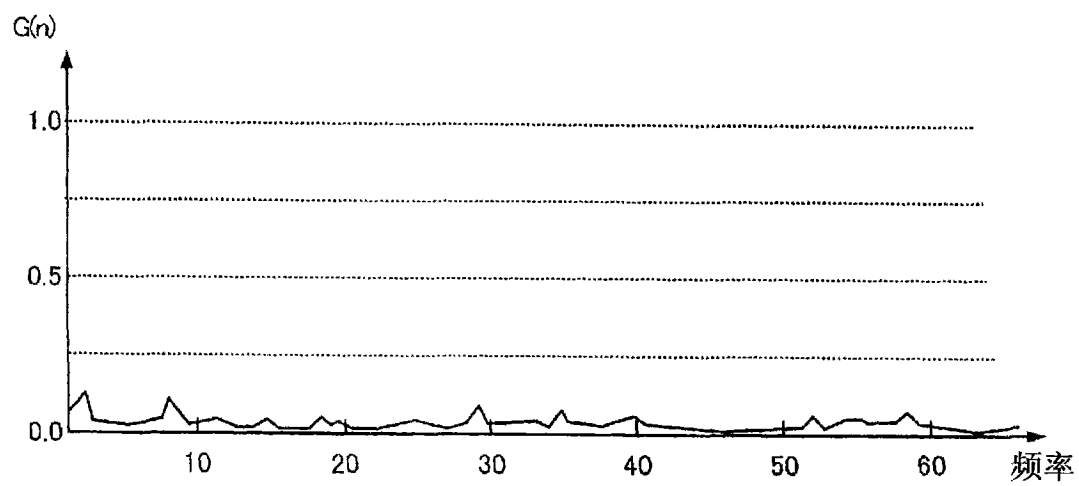


图 7