



(12) 发明专利

(10) 授权公告号 CN 102652336 B

(45) 授权公告日 2015.02.18

(21) 申请号 201080055064.1

(22) 申请日 2010.10.22

(30) 优先权数据

2009-297147 2009.12.28 JP

(85) PCT国际申请进入国家阶段日

2012.06.05

(86) PCT国际申请的申请数据

PCT/JP2010/006264 2010.10.22

(87) PCT国际申请的公布数据

W02011/080855 JA 2011.07.07

(73) 专利权人 三菱电机株式会社

地址 日本东京

(72) 发明人 古田训 田崎裕久

(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 11038

代理人 许海兰

(51) Int. Cl.

G10L 21/038 (2013.01)

(56) 对比文件

JP 特开 2007-72264 A, 2007.03.22, 全文.

WO 03/019533 A1, 2003.03.06, 全文.

JP 特开平 10-124098 A, 1998.05.15, 全文.

JP 特开平 8-248997 A, 1996.09.27, 全文.

CN 101432804 A, 2009.05.13, 全文.

审查员 康丹丹

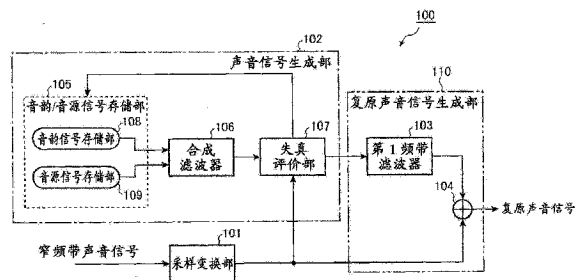
权利要求书2页 说明书11页 附图8页

(54) 发明名称

声音信号复原装置以及声音信号复原方法

(57) 摘要

合成滤波器(106)从声音信号码本(105)中组合宽频带的音韵信号和音源信号来合成多个宽频带声音信号,失真评价部(107)选择从采样变换部(101)输出的与已经上采样的窄频带声音信号的波形失真最小的宽频带声音信号。第1频带滤波器(103)抽出该宽频带声音信号的窄频带以外的频率分量,频带合成部(104)组合到已经上采样的窄频带声音信号中。



1. 一种声音信号复原装置,具备:

合成滤波器,组合音韵信号以及音源信号,生成多个宽频带声音信号;

失真评价部,使用规定的失真尺度,评价具有所述合成滤波器所生成的宽频带声音信号的频带中的至少一部分频带的频率分量的比较对象信号与所述合成滤波器所生成的所述多个宽频带声音信号中的各个宽频带声音信号的波形失真,并根据该评价的结果,选择所述多个宽频带声音信号中的失真最小的宽频带声音信号、或者选择成为所设定的容许范围内的失真的宽频带声音信号的某一个;以及

复原声音信号生成部,使用所述失真评价部所选择的宽频带声音信号,生成复原声音信号。

2. 根据权利要求1所述的声音信号复原装置,其特征在于,

复原声音信号生成部具有频带合成部,该频带合成部组合比较对象信号与失真评价部所选择的宽频带声音信号。

3. 根据权利要求1所述的声音信号复原装置,其特征在于,

失真评价部对比较对象信号与合成滤波器所生成的多个宽频带声音信号中的各个宽频带声音信号的、规定的频带的频率分量的波形失真进行评价。

4. 根据权利要求3所述的声音信号复原装置,其特征在于,

具备采样变换部,该采样变换部对比较对象信号进行采样变换以使其对应于规定的频带,

失真评价部对所述采样变换部进行了采样变换的所述比较对象信号与合成滤波器所生成的多个宽频带声音信号中的各个宽频带声音信号的、所述规定的频带的频率分量的波形失真进行评价。

5. 一种声音信号复原方法,具备:

合成滤波步骤,组合音韵信号以及音源信号,生成多个宽频带声音信号;

失真评价步骤,使用规定的失真尺度,评价具有在所述合成滤波步骤中生成的宽频带声音信号的频带中的至少一部分频带的频率分量的比较对象信号与在所述合成滤波步骤中生成的所述多个宽频带声音信号中的各个宽频带声音信号的波形失真,并根据该评价的结果,选择所述多个宽频带声音信号中的失真最小的宽频带声音信号、或者选择成为所设定的容许范围内的失真的宽频带声音信号的某一个;以及

复原声音信号生成步骤,使用在所述失真评价步骤中所选择的宽频带声音信号,生成复原声音信号。

6. 根据权利要求5所述的声音信号复原方法,其特征在于,

复原声音信号生成步骤具有频带合成步骤,在该频带合成步骤中组合比较对象信号与在失真评价步骤中所选择的宽频带声音信号。

7. 根据权利要求5所述的声音信号复原方法,其特征在于,

在失真评价步骤中,对比较对象信号与在合成滤波步骤中生成的多个宽频带声音信号中的各个宽频带声音信号的、规定的频带的频率分量的波形失真进行评价。

8. 根据权利要求7所述的声音信号复原方法,其特征在于,

具备采样变换步骤,在该采样变换步骤中对比较对象信号进行采样变换以使其对应于规定的频带,

在失真评价步骤中,对在所述采样变换步骤中进行了采样变换的所述比较对象信号与在合成滤波步骤中生成的多个宽频带声音信号中的各个宽频带声音信号的、所述规定的频带的频率分量的波形失真进行评价。

声音信号复原装置以及声音信号复原方法

技术领域

[0001] 本发明涉及从频带被限制为窄频带的声音信号复原宽频带的声音信号、以及对变差或者缺损的频带的声音信号进行复原的声音信号复原装置及其方法。

背景技术

[0002] 在模拟电话中,通过电话线路送来的声音信号的频带被限制为例如 300~3400Hz 这样的窄频带。因此,以往的电话线路的音质不能说很好。另外,在便携电话等数字声音通信中,由于比特率的严格的限制,与模拟线路同样地,频带宽度被限制,所以在该情况下也不能说音质好。

[0003] 另外,近年来,伴随着声音压缩技术(声音编码技术)的发展,能够以低比特率对宽频带(例如 50~7000Hz)的声音信号进行无线传送。但是,发送侧终端以及接收侧终端这双方需要支持对应的宽频带声音编码/解码方法,并且在双方的基站中也需要具备用于宽频带编码的网络,所以仅在一部分的业务通信系统中被实用化,为了在公共电话通信网中实施,不仅在经济上成为大的负担,而且直至普及需要大量的时间。

[0004] 因此,依然未解决以往的模拟电话线路通信以及数字声音通信的音质的问题。

[0005] 因此,针对上述问题,作为在接收侧从窄频带信号虚拟地生成或者复原宽频带信号的方法,例如公开了专利文献 1、2。在专利文献 1 的频带扩展装置中,计算窄频带声音信号的自相关系数而抽出声音的基本周期,并根据该基本周期得到宽频带声音信号。另外,在专利文献 2 的宽频带声音信号复原装置中,通过基于利用合成的分析法的编码方法对窄频带声音信号进行编码,并对作为该编码的最终结果而得到的音源信号或者声音信号,进行零填充处理(oversampling:过采样)而得到宽频带声音信号。

[0006] 专利文献 1:日本专利第 3243174 号(第 3~5 页、图 1)

[0007] 专利文献 2:日本专利第 3230790 号(第 3~4 页、图 1)

发明内容

[0008] 以往的声音信号复原装置由于如上所述构成,所以存在以下叙述的问题。

[0009] 在专利文献 1 公开的频带扩展装置中,需要抽出窄频带声音信号的基本周期。虽然公开了各种抽出声音的基本周期的方案,但难以正确地抽出声音信号的基本周期。在噪声环境下更加困难。

[0010] 在专利文献 2 公开的宽频带声音信号复原装置中,具有无需抽出声音信号的基本周期的优点。然而,所生成的宽频带音源信号虽然是从窄频带信号分析以及生成的信号,但由于是通过零填充处理(过采样)而虚拟地生成的信号,所以混入了重叠失真分量,因此存在不适合宽频带声音信号(尤其是高频信号)、且音质变差这样的问题。

[0011] 本发明是为了解决上述那样的问题而完成的,其目的在于提供一种高质量地复原声音信号的声音信号复原装置以及声音信号复原方法。

[0012] 本发明的声音信号复原装置,具备:合成滤波器,组合音韵信号以及音源信号,生

成多个声音信号；失真评价部，使用规定的失真尺度，评价具有合成滤波器所生成的声音信号的频带中的至少一部分频带的频率分量的比较对象信号与合成滤波器所生成的多个声音信号中的各个声音信号的波形失真，并根据该评价结果，选择多个声音信号中的某一个；以及复原声音信号生成部，使用失真评价部所选择的聲音信号，生成复原声音信号。

[0013] 本发明的声音信号复原方法，具备：合成滤波步骤，组合音韵信号以及音源信号，生成多个声音信号；失真评价步骤，使用规定的失真尺度，评价具有在合成滤波步骤中生成的声音信号的频带中的至少一部分频带的频率分量的比较对象信号与在合成滤波步骤中生成的多个声音信号中的各个声音信号的波形失真，并根据该评价结果，选择多个声音信号中的某一个；以及复原声音信号生成步骤，使用在失真评价步骤中所选择的聲音信号，生成复原声音信号。

[0014] 根据本发明，组合音韵信号以及音源信号来生成多个声音信号，使用规定的失真尺度，分别评价与比较对象信号的波形失真，并根据该评价结果来选择某一个声音信号而生成复原声音信号，所以能够提供将例如由于频带限制或者噪声压制而导致任意的频带的频率分量欠缺了的比较对象信号高质量地进行复原的声音信号复原装置以及声音信号复原方法。

附图说明

[0015] 图 1 是示出本发明的实施方式 1 的声音信号复原装置 100 的结构的框图。

[0016] 图 2 是示意性地示出本发明的实施方式 1 的声音信号复原装置 100 生成的声音信号的曲线图。

[0017] 图 3 是示出本发明的实施方式 2 的声音信号复原装置 100 的结构的框图。

[0018] 图 4 是示出本发明的实施方式 3 的声音信号复原装置 200 的结构的框图。

[0019] 图 5 是示意性地示出本发明的实施方式 3 的声音信号复原装置 200 生成的声音信号的曲线图。

[0020] 图 6 是示意性地示出本发明的实施方式 5 的声音信号复原装置 200 的失真评价部 107 的失真评价处理的曲线图。

[0021] 图 7 是示出图 1 所示的复原声音信号生成部 110 的变形例的框图。

[0022] 图 8 是示意性地示出图 7 所示的复原声音信号生成部 110 生成的声音信号的曲线图。

具体实施方式

[0023] 以下，参照附图，详细说明本发明的实施方式。

[0024] 实施方式 1.

[0025] 在本实施方式 1 中，以用于从由于经由电话线路等传送路径而导致频带被限制为窄频带的声音信号生成宽频带的声音信号的声音信号复原装置为例子进行说明，该声音信号复原装置用于导入了声音通信、声音储存或者声音识别系统的汽车导航、便携电话及对讲机等声音通信系统、免提通话系统、TV 会议系统以及监视系统等的音质改善、声音识别系统的识别率提高。

[0026] 图 1 是示出本实施方式 1 的声音信号复原装置 100 的整体结构的图。

[0027] 在图 1 中,声音信号复原装置 100 包括采样变换部 101、声音信号生成部 102 以及复原声音信号生成部 110。该声音信号生成部 102 包括:具备音韵信号存储部 108 和音源信号存储部 109 的音韵/音源信号存储部 105、合成滤波器 106、以及失真评价部 107。另外,复原声音信号生成部 110 包括第 1 频带滤波器 103 和频带合成部 104。

[0028] 图 2 是示意性地示出通过本实施方式 1 的结构生成的声音信号的图。图 2 的(a)示出输入到采样变换部 101 的窄频带声音信号(比较对象信号)。图 2 的(b)示出采样变换部 101 输出的已经上采样的窄频带声音信号(进行了采样变换的比较对象信号)。图 2 的(c)示出失真评价部 107 从由合成滤波器 106 所生成的多个宽频带声音信号(声音信号)中选择出的失真最小的宽频带声音信号。图 2 的(d)示出第 1 频带滤波器 103 的输出、即从宽频带声音信号抽出了低频分量和高频分量的信号。图 2 的(e)示出声音信号复原装置 100 的输出结果即复原声音信号。另外,图 2 中的各箭头表示处理的顺序,各曲线图的纵轴表示功率,横轴表示频率。

[0029] 以下,根据图 1 以及图 2,说明该声音信号复原装置 100 的动作原理。

[0030] 首先,通过未图示的麦克风等而取入的声音以及音乐等被进行了 A/D(模拟/数字)变换之后,以规定的采样频率(例如,8kHz)被采样并且被分割为帧单位(例如,10ms),进而被频带限制(例如,300~3400Hz)而成为窄频带声音信号,并被输入到本实施方式 1 的声音信号复原装置 100。另外,在本实施方式 1 中,将最终得到的宽频带的复原声音信号的频带设为 50~7000Hz 而进行说明。

[0031] 采样变换部 101 对于所输入的窄频带声音信号例如以 16kHz 进行上采样,并通过低通滤波器去除了重叠失真信号之后,作为已经上采样的窄频带声音信号而输出。

[0032] 在声音信号生成部 102 中,合成滤波器 106 使用音韵信号存储部 108 中保存的音韵信号和音源信号存储部 109 中保存的音源信号来生成多个宽频带声音信号,失真评价部 107 根据规定的失真尺度(distortion scale)来计算与已经上采样的窄频带声音信号的波形失真,选择并输出失真成为最小的宽频带声音信号。另外,该声音信号生成部 102 也可以是与例如 CELP(Code-Excited Linear Prediction:码激励线性预测)编码方式中的解码方法同样的结构,在该情况下,在音韵信号存储部 108 中保存音韵符号,在音源信号存储部 109 中预先保存音源符号。

[0033] 音韵信号存储部 108 采用除了音韵信号以外还一并具有音韵信号的功率或者增益的结构,以能够表现各种宽频带声音信号的音韵形状(频谱图案)的方式,将大量并且多种多样的音韵信号保存到存储器等存储单元中,根据后述的失真评价部 107 的指示将音韵信号输出到合成滤波器 106。能够使用线性预测分析等公知的方案,从宽频带的声音信号(例如,具有 50~7000Hz 的频带)求出这些音韵信号。另外,关于频谱图案,能够以频谱信号自身、或者 LSP(Line Spectrum Pair:线谱对)参数以及倒谱(Cepstrum)等音响参数(acoustic parameter)形式来表现,以能够适用于合成滤波器 106 的滤波系数的方式进行适当变换即可。而且,为了削减存储量,也可以通过标量量化以及矢量量化等公知的方案,对所得到的音韵信号进行压缩。

[0034] 音源信号存储部 109 采用除了音源信号以外还一并具备音源信号的功率或者增益的结构,与音韵信号存储部 108 同样地,以能够表现各种宽频带声音信号的音源信号形状(脉冲串)的方式,将大量并且多种多样的音源信号保存到存储器等存储单元,根据后述

的失真评价部 107 的指示,将音源信号输出到合成滤波器 106。能够使用宽频带的声音信号(例如,具有 50~7000Hz 的频带)和上述音韵信号,通过 CELP 的方案来学习并求出这些音源信号。另外,关于所得到的音源信号,既可以为了削减存储量而通过标量量化以及矢量量化等公知的方案进行压缩,也可以如多脉冲化以及 ACELP (Algebraic CELP:代数码激励线性预测)方式那样通过规定的模型来表现音源信号。另外,还能够如 VSELP (Vector Sum Excited Linear Prediction:矢量和激励线性预测)编码方式那样采用一并具备从过去的音源信号生成的自适应音源码本(adaptive sound source code book)的构造。

[0035] 另外,合成滤波器 106 也可以对音韵信号的功率或者增益、和音源信号的功率或者增益分别进行了调整之后进行合成。在该结构的情况下,从 1 个音韵信号和 1 个音源信号也能够生成多个宽频带声音信号,所以能够削减音韵信号存储部 108 以及音源信号存储部 109 的存储量。

[0036] 失真评价部 107 评价合成滤波器 106 所输出的宽频带声音信号与采样变换部 101 所输出的已经上采样的窄频带声音信号的波形失真。此时,评价失真的频带(规定的频带)仅限于窄频带声音信号的范围,在本例子中限于 300~3400Hz。为了在窄频带声音信号的频带的范围内进行波形失真的评价,例如对于宽频带声音信号和已经上采样的窄频带声音信号这两者,能够使用具有 300~3400Hz 的带通特性的 FIR (Finite Impulse Response:有限冲激响应特性)滤波器进行滤波处理之后,利用下式所示那样的平均波形失真或者利用基于欧几里德距离的评价法。

[0037] 式(1)

$$[0038] \quad E_t = \frac{1}{N} \sum_{n=0}^{N-1} \{s(n) - u(n)\}^2 \quad (1)$$

[0039] 此处,s(n)以及u(n)分别是已经 FIR 滤波处理的宽频带声音信号、已经上采样的窄频带声音信号,N 是声音信号波形的样本数(160 样本、16kHz 采样的情况)。另外,在不进行 300Hz 以下的低频部分的复原的情况下,也可以不使用上述 FIR 滤波器而将宽频带声音信号下采样到窄频带声音信号的频率(8kHz),进行与上采样前的窄频带声音信号的失真评价。另外,失真评价部 107 在以上使用 FIR 滤波器进行了滤波处理,但只要能够适当地进行失真评价,也可以使用例如 IIR (Infinite Impulse Response:无限冲激响应特性)滤波器。

[0040] 另外,失真评价部 107 也可以并非在时间轴上而是在频率轴上进行失真评价,例如,也可以对宽频带声音信号和已经上采样的窄频带声音信号这两者实施了零填充、加窗之后,使用 256 点的 FFT (Fast Fourier Transform:快速傅立叶变换)变换到频谱区域,例如如下式那样将功率谱上的差分的总和评价为失真。在该情况下,与时间轴上的评价不同,无需进行具有带通特性的滤波处理。

[0041] 式(2)

$$[0042] \quad E_f = \sum_{f=FL}^{FH} \{S(f) - U(f)\} \quad (2)$$

[0043] 此处,S(f)以及U(f)分别是宽频带声音信号的功率谱分量、已经上采样的窄频带声音信号的功率谱分量,FL 以及 FH 是与分别 300Hz、3400Hz 相当的频谱分量编号。

[0044] 失真评价部 107 依次发出从音韵信号存储部 108 以及音源信号存储部 109 输出频谱图案和音源信号的组的指示,使合成滤波器 106 生成宽频带声音信号,并通过上式(1)或者上式(2)计算失真。然后,选择失真最小的宽频带声音信号,输出到第 1 频带滤波器 103。另外,失真评价部 107 还能够在对宽频带声音信号和已经上采样的窄频带声音信号这两者实施了在 CELP 声音编码方式中通常使用的听觉加权处理之后,计算失真。另外,失真评价部 107 无需一定选择失真最小的宽频带声音信号,而也可以选择失真例如第 2 小的宽频带声音信号。或者,也可以设定失真的容许范围来选择成为该范围内的失真的宽频带声音信号,不进行此后的合成滤波器 106 以及失真评价部 107 的处理而削减处理次数。

[0045] 第 1 频带滤波器 103 从宽频带声音信号抽出窄频带声音信号的频带以外的频率分量,并输出到频带合成部 104。即,在本实施方式 1 中,抽出 300Hz 以下的低频分量和 3400Hz 以上的高频分量。在低频分量以及高频分量的抽出中使用 FIR 滤波器、IIR 滤波器等即可。作为声音信号的一般的特性,低频部分的谐波构造在高频部分中也同样地出现的情况较多,相反地,如果在高频部分中能够观察谐波构造,则同样地在低频部分中也出现的情况较多。这样,在低频—高频之间互相关性强,所以通过从以使与窄频带声音信号的失真成为最小的方式生成的宽频带声音信号得到由第 1 频带滤波器 103 抽出的低频分量以及高频分量,从而能够构成最佳的复原声音信号。

[0046] 频带合成部 104 将由第 1 频带滤波器 103 所输出的宽频带声音信号中的低频分量以及高频分量、与由采样变换部 101 所输出的已经上采样的窄频带声音信号进行相加来复原宽频带声音信号,并作为复原声音信号而输出。

[0047] 以上,根据本实施方式 1,提供一种声音信号复原装置 100,从频带被限制为窄频带的窄频带声音信号变换为包含窄频带的宽频带声音信号,该声音信号复原装置 100 构成为具备:采样变换部 101,对窄频带声音信号进行采样变换以使其匹配宽频带;合成滤波器 106,将音韵/音源信号存储部 105 所保存的具有宽频带的频率分量的音韵信号以及音源信号进行组合,生成多个宽频带声音信号;失真评价部 107,使用规定的失真尺度,分别评价采样变换部 101 进行了采样变换的已经上采样的窄频带声音信号与合成滤波器 106 生成的多个宽频带声音信号的波形失真,根据该评价结果来选择失真成为最小的宽频带声音信号;第 1 频带滤波器 103,从由失真评价部 107 所选择的宽频带声音信号抽出窄频带以外的频率分量;以及频带合成部 104,将采样变换部 101 进行了采样变换的已经上采样的窄频带声音信号组合到第 1 频带滤波器 103 抽出的频率分量中。这样,从以使窄频带声音信号的失真成为最小的方式生成的宽频带声音信号得到用于复原声音信号的低频分量以及高频分量,所以能够复原高质量的宽频带的声音信号。

[0048] 另外,根据本实施方式 1,无需抽出声音的基本周期,不会由于基本周期的抽出错误而使质量变差,所以即使在声音的基本周期的分析困难的噪声环境下,也能够复原高质量的宽频带的声音信号。

[0049] 另外,根据本实施方式 1,不对音源信号进行导致变差那样的零填充、全波整流处理等非线性处理,所以能够复原高质量的宽频带的声音信号。

[0050] 另外,根据本实施方式 1,从以使窄频带声音信号的失真成为最小的方式生成的宽频带声音信号得到用于复原声音信号的低频分量以及高频分量,原理上能够使窄频带声音信号与低频分量(或者高频分量与窄频带声音信号)平滑地连接,无需频带合成时的功率校

正等插值处理,能够复原高质量的宽频带的声音信号。

[0051] 另外,上述实施方式 1 的声音信号复原装置 100 在失真评价部 107 中的失真评价结果非常小的情况下,也可以省略第 1 频带滤波器 103 和频带合成部 104 的处理,并将失真评价部 107 所输出的宽频带声音信号直接作为复原声音信号而输出。

[0052] 另外,在上述实施方式 1 中,对于低频以及高频这两方欠缺了的窄频带声音信号,复原这些低频以及高频这两方的频率分量,但不限于此,即使是低频、中频、高频中的至少 1 个频带欠缺了的窄频带声音信号,当然也能够复原。这样,只要是具有合成滤波器 106 所生成的宽频带声音信号的频带中的至少一部分频带的窄频带声音信号,声音信号复原装置 100 就能够复原为与宽频带声音信号相同的频带。

[0053] 实施方式 2.

[0054] 作为上述实施方式 1 的变形例,还能够将窄频带声音信号的分析结果用作用于生成宽频带声音信号的辅助信息。图 3 是示出本实施方式 2 的声音信号复原装置 100 的整体结构的图,是对图 1 所示的声音信号复原装置 100 新追加了声音分析部 111 的结构。关于其他结构要素,对于与图 1 对应的部分附加同一符号,省略详细的说明。

[0055] 声音分析部 111 对于所输入的窄频带声音信号,通过线性预测分析等公知的方案来进行音响特征的分析,抽出窄频带声音信号的音韵信号和音源信号,并分别输出到音韵信号存储部 108 和音源信号存储部 109。此时,作为音韵信号,优选例如插值特性优良的 LSP 参数,但也可以是其他参数。另外,关于音源信号,声音分析部 111 具备在滤波系数中具有例如作为分析结果的音韵信号的逆滤波器,能够将对窄频带声音信号进行滤波处理而得到的残差信号作为音源信号。

[0056] 在音韵 / 音源信号存储部 105 中,将从声音分析部 111 输入的窄频带声音信号的音韵信号和音源信号作为音韵信号存储部 108 和音源信号存储部 109 的辅助信息。在音韵信号存储部 108 中,作为辅助信息的用法,例如能够从宽频带声音信号的音韵信号中去除了 300~3400Hz 的部分,对去除了的部分应用窄频带声音信号的音韵信号。通过应用窄频带声音信号的音韵信号,能够得到与窄频带声音信号更近似的宽频带声音信号的音韵信号。另外,音韵信号存储部 108 能够进行如下那样的预备选择,即,进行窄频带声音信号的音韵信号与宽频带声音信号的例如在频谱上的失真评价,仅将失真少的宽频带声音信号的音韵信号输出到合成滤波器 106。通过进行音韵信号的预备选择,能够削减合成滤波器 106 和失真评价部 107 的处理次数。

[0057] 在音源信号存储部 109 中,作为辅助信息的用法,能够与音韵信号存储部 108 同样地,例如将窄频带声音信号的音源信号添加到宽频带声音信号中或者用作预备选择的信息。通过添加窄频带声音信号的音源信号,能够得到与窄频带声音信号更近似的宽频带声音信号的音源信号。另外,通过进行音源信号的预备选择,能够削减合成滤波器 106 和失真评价部 107 的处理次数。

[0058] 以上,根据本实施方式 2,声音信号复原装置 100 具备声音分析部 111,该声音分析部 111 对于频带被限制为窄频带的窄频带声音信号进行音响分析而生成辅助信息,合成滤波器 106 使用声音分析部 111 所生成的辅助信息,分别组合音韵 / 音源信号存储部 105 所保存的具有宽频带的频率分量的多个音韵信号以及多个音源信号,生成多个宽频带声音信号。因此,通过将窄频带声音信号的分析结果用作辅助信息,能够得到与窄频带声音信号更

近似的宽频带声音信号,能够复原更高质量的宽频带的声音信号。

[0059] 另外,根据本实施方式 2,在生成宽频带声音信号时,能够将窄频带声音信号的分析结果用于辅助信息来预备选择音韵信号以及音源信号,所以能够在确保了高质量的状态下削减处理量。

[0060] 另外,在本实施方式 2 中,在输入到采样变换部 101 之前实施了声音分析部 111 的处理,但即使是采样变换部 101 的处理后也没有关系。在该情况下,进行已经上采样的窄频带声音信号的声音分析。

[0061] 另外,声音分析部 111 也可以对所输入的窄频带声音信号进行例如声音信号和噪声信号的频率分析,生成指定了声音信号频谱功率与噪声信号频谱功率之比(信噪比,以下称为 SN 比)高的频带的辅助信息。在该结构的情况下,采样变换部 101 对窄频带声音信号中的由该辅助信息指定的频带(规定的频带)的频率分量进行采样变换,失真评价部 107 在由该辅助信息指定的频带的频率分量彼此之间进行已经上采样的窄频带声音信号与多个宽频带声音信号的失真评价。而且,第 1 频带滤波器 103 抽出失真评价部 107 选择出的宽频带声音信号中的由该辅助信息指定的频带以外的频率分量,通过频带合成部 104 合成到该频带的已经上采样的窄频带声音信号中。因此,失真评价部 107 不是在窄频带声音信号的整个频带而是仅在由辅助信息指定的频带中进行失真评价,能够削减处理量。

[0062] 实施方式 3.

[0063] 在上述实施方式 2 中,说明了用于从频带被限制为窄频带的声音信号生成宽频带的声音信号的声音信号复原装置 100,但在本实施方式 2 中,通过将该声音信号复原装置 100 变形而应用,构成用于将由于噪声压制处理、声音压缩处理等而变差或者缺损了的频带的声音信号进行复原的声音信号复原装置 200。图 4 是示出本实施方式 3 的声音信号复原装置 200 的整体结构的图,是对图 1 所示的声音信号复原装置 100 新追加了噪声压制部 201 以及第 2 频带滤波器 202 的结构。关于其他结构要素,对于与图 1 对应的部分附加同一符号,省略详细的说明。

[0064] 另外,在本实施方式 3 中,为了简化说明,将所输入的噪声混入声音信号的频带设为 $0\sim 4000\text{Hz}$,在所混入的噪声中假设汽车行驶噪音,设为在 $0\sim 500\text{Hz}$ 的频带中混入了噪声。此时,声音信号生成部 102 内部的音韵 / 音源信号存储部 105、合成滤波器 106 以及失真评价部 107、第 1 频带滤波器 103 以及第 2 频带滤波器 202 进行与 $0\sim 4000\text{Hz}$ 的频带对应的动作,或者保持音韵信号以及音源信号。另外,在应用于实际的系统时,当然不限于这些条件。

[0065] 图 5 是示意性地图示通过本实施方式 3 的结构生成的声音信号的图。图 5 的(a)示出噪声压制部 201 所输出的已经压制噪声的声音信号(比较对象信号)。图 5 的(b)示出从由合成滤波器 106 所生成的多个宽频带声音信号(声音信号)中由失真评价部 107 所选择的与已经压制噪声的声音信号的失真成为最小的宽频带声音信号。图 5 的(c)示出第 1 频带滤波器 103 的输出、即从宽频带声音信号抽出了低频分量的信号。图 5 的(d)示出第 2 频带滤波器 202 所输出的已经压制噪声的声音信号的高频分量。图 5 的(e)示出声音信号复原装置 200 的输出结果即复原声音信号。另外,图 5 中的各箭头表示处理的顺序,各曲线图的纵轴表示功率,横轴表示频率。

[0066] 以下,根据图 4 以及图 5,说明该声音信号复原装置 200 的动作原理。

[0067] 噪声压制部 201 输入混入了噪声的噪声混入声音信号,将压制了噪声的声音信号

输出到失真评价部 107 以及第 2 频带滤波器 202。另外,噪声压制部 201 输出用于后级的失真评价部 107 中的失真评价和第 1 频带滤波器 103 使用的、指定了分离为 0~500Hz 的低频和 500~4000Hz 的高频的低频 / 宽频分割频率的频带信息信号。另外,频带信息信号在本实施方式 3 中固定为 500Hz,但是例如所输入的噪声混入声音信号的情况下,例如也可以进行声音信号和噪声信号的频率分析,将噪声信号频谱功率超过声音信号频谱功率的频率(频谱上的 SN 比交叉 0dB 的频率)作为频带信息信号。另外,该频率根据所输入的噪声混入声音信号及其噪声的情况而时刻发生变化,所以例如也可以针对 10ms 的每帧进行变更。

[0068] 此处,作为噪声压制部 201 中的噪声压制处理的方案,例如除了《Steven F. Boll, “Suppression of acoustic noise in speech using spectral subtraction”, IEEE Trans. ASSP, Vol. ASSP — 27, No. 2, Apr. 1979》中公开的基于频谱减法运算的方案、以及《J. S. Lim and A. V. Oppenheim, “Enhancement and Bandwidth Compression of Noisy Speech”, Proc. of the IEEE, vol. 67, pp. 1586 — 1604, Dec. 1979》中公开的根据每个频谱分量的 SN 比而针对每个频谱分量提供衰减量的频谱振幅压制的方案等公知的方法以外,还能够使用组合了频谱减法运算和频谱振幅压制的方案(例如,专利第 3454190 号)等。

[0069] 与上述实施方式 1 同样地,在声音信号生成部 102 中,合成滤波器 106 使用音韵信号存储部 108 中保存的音韵信号和音源信号存储部 109 中保存的音源信号来生成多个宽频带声音信号,失真评价部 107 根据规定的失真尺度来评价与压制了噪声的已经压制噪声的声音信号的波形失真,选择并输出与任意的条件匹配的波形失真的宽频带声音信号。

[0070] 在失真评价部 107 中,作为在评价波形失真时对失真进行评价的频带(规定的频带),限定为比频带信息信号所指定的频率高的范围,在本例子中限定为 500~4000Hz。为了在该范围中进行波形失真的评价,例如能够采用与在上述实施方式 1 中使用的方案同样的方案。失真评价部 107 依次发出从音韵信号存储部 108 以及音源信号存储部 109 输出频谱图案与音源信号的组的指示而使合成滤波器 106 生成多个宽频带声音信号,选择例如波形失真成为最小的宽频带声音信号,并输出到第 1 频带滤波器 103。

[0071] 第 1 频带滤波器 103 从由失真评价部 107 生成的宽频带声音信号,抽出频带信息信号所表示的低频 / 宽频分割频率以下的低频分量,并输出到频带合成部 104。在通过第 1 频带滤波器 103 抽出低频分量时,与实施方式 1 同样地使用 FIR 滤波器、IIR 滤波器等即可。作为声音信号的一般的特性,低频部分的谐波构造在高频部分中也同样地出现的情况较多,相反地,如果在高频部分中能够观察谐波构造,则同样地在低频部分中也出现的情况较多。这样,在低频—高频之间互相关性强,所以通过从以使与已经压制噪声的声音信号的失真成为最小的方式生成的宽频带声音信号得到由第 1 频带滤波器 103 抽出的低频分量,从而能够构成最佳的复原声音信号。

[0072] 第 2 频带滤波器 202 进行与上述第 1 频带滤波器 103 相逆的动作。即,从已经压制噪声的声音信号,抽出频带信息信号所表示的低频 / 宽频分割频率以上的高频分量,并输出到频带合成部 104。在通过第 2 频带滤波器 202 抽出高频分量时,与第 1 频带滤波器 103 同样地使用 FIR 滤波器、IIR 滤波器等即可。

[0073] 频带合成部 104 将第 1 频带滤波器 103 所输出的宽频带声音信号的低频分量、与第 2 频带滤波器 202 所输出的已经压制噪声的声音信号的高频分量进行相加而复原声音信号,并作为复原声音信号而输出。

[0074] 根据本实施方式 3, 提供一种声音信号复原装置 200, 复原由于通过噪声压制部 201 对噪声混入声音信号进行噪声压制处理而变差或者缺损了的已经压制噪声的声音信号, 来生成复原声音信号, 该声音信号复原装置 200 构成为具备: 合成滤波器 106, 将音韵/音源信号存储部 105 所保存的音韵信号以及音源信号进行组合, 来生成多个宽频带声音信号; 失真评价部 107, 使用规定的失真尺度, 分别评价已经压制噪声的声音信号与合成滤波器 106 所生成的多个宽频带声音信号的波形失真, 并根据该评价结果, 选择失真成为最小的宽频带声音信号; 第 1 频带滤波器 103, 从由失真评价部 107 所选择的宽频带声音信号, 抽出变差或者缺损了的频带的频率分量; 第 2 频带滤波器 202, 从已经压制噪声的声音信号, 抽出变差或者缺损了的频带以外的频率分量; 以及频带合成部 104, 组合第 1 频带滤波器 103 抽出的频率分量与第 2 频带滤波器 202 抽出的频率分量。这样, 从以使与压制了噪声的声音信号的失真成为最小的方式生成的声音信号得到用于复原声音信号的低频分量, 所以能够复原高质量的声音信号。

[0075] 另外, 根据本实施方式 3, 无需抽出声音的基本周期, 不会由于基本周期的抽出错误而使质量变差, 所以即使在声音的基本周期的分析困难的噪声环境下, 也能够复原高质量的声音信号。

[0076] 另外, 根据本实施方式 3, 从以使与压制了噪声的声音信号的失真成为最小的方式生成的声音信号得到用于复原声音信号的低频分量, 所以在原理上能够使压制了噪声的声音信号的高频分量与所生成的低频分量平滑地连接, 无需频带合成时的功率校正等插值处理, 能够复原高质量的声音信号。

[0077] 另外, 上述实施方式 3 的声音信号复原装置 200 在失真评价部 107 中的失真评价结果非常小的情况下, 也可以省略第 1 频带滤波器 103、第 2 频带滤波器 202、频带合成部 104 的各处理, 将失真评价部 107 所输出的宽频带声音信号直接作为复原声音信号而输出。

[0078] 另外, 在上述实施方式 3 中, 对于低频变差或者缺损了的已经压制噪声的信号, 复原低频的频率分量, 但不限于此, 也可以对于低频以及高频的一方或者两方变差或者缺损了的已经压制噪声的声音信号, 复原这些频带的频率分量, 还可以根据噪声压制部 201 输出的频带信息信号, 复原例如 800~1000Hz 的中间的频带的频率分量。作为中间的频带变差或者缺损这样的状况, 例如考虑在汽车高速行驶时发生的风噪(Wind noise)等局部频带的噪声混入到声音信号的情况。这样, 在实施方式 3 中也与上述实施方式 1、2 同样地, 只要是具有合成滤波器 106 生成的宽频带声音信号的频带中的至少一部分频带的已经压制噪声的声音信号, 就能够复原该已经压制噪声的声音信号的剩余的频带的频率分量。

[0079] 实施方式 4.

[0080] 作为上述实施方式 3 的变形例, 还能够与上述实施方式 2 同样地, 将压制了噪声的声音信号的分析结果用作用于生成宽频带声音信号的辅助信息。具体而言, 在上述实施方式 3 的声音信号复原装置 200 中, 追加图 3 所示那样的声音分析部 111, 该声音分析部 111 对从噪声压制部 201 输入的已经压制噪声的声音信号进行音响特征的分析, 抽出已经压制噪声的声音信号的音韵信号和音源信号, 并分别输出到音韵信号存储部 108 和音源信号存储部 109。

[0081] 根据本实施方式 4, 声音信号复原装置 200 具备声音分析部 111, 该声音分析部 111 对已经压制噪声的声音信号进行音响分析而生成辅助信息, 合成滤波器 106 使用声音分析

部 111 所生成的辅助信息,组合音韵 / 音源信号存储部 105 所保存的音韵信号以及音源信号,来生成宽频带声音信号。因此,通过将已经压制噪声的声音信号的分析结果用作辅助信息,能够得到与已经压制噪声的声音信号更近似的宽频带声音信号,能够复原更高质量的声音信号。

[0082] 另外,根据本实施方式 4,在生成宽频带声音信号时,能够将已经压制噪声的声音信号的分析结果用于辅助信息而预备选择音韵信号以及音源信号,所以能够在确保了高质量的状态下削减处理量。

[0083] 实施方式 5.

[0084] 在上述实施方式 3 中,根据频带信息信号将声音信号 2 分割为低频和高频,在失真评价处理中仅评价了高频部分的失真,但例如还能够对于一部分低频分量也进行加权之后设为失真评价的对象,或者进行与噪声信号的频率特性对应的加权而进行失真评价。另外,本实施方式 5 的声音信号复原装置与图 4 所示的声音信号复原装置 200 在附图上是相同的结构,所以以下使用图 4 来说明。

[0085] 图 6 是用于失真评价部 107 的失真评价的加权系数的一个例子,图 6 的(a)是将一部分低频分量也设为评价对象的情况,图 6 的(b)是将噪声信号的频率特性的逆特性设为权重系数的情况。图 6 中的各曲线图的纵轴表示振幅和失真评价权重值,横轴表示频率。另外,作为失真评价部 107 中的向失真评价的权重系数反映方法,例如考虑对于滤波系数卷积权重系数、或者对功率谱分量乘以权重系数的方法。另外,作为第 1 频带滤波器 103 以及第 2 频带滤波器 202 的特性,既可以与上述实施方式 3 中采用的特性同样地是按照低频和高频进行分离的特性,也可以是表现图 6 的(a)的权重系数的频率特性那样的滤波特性。

[0086] 在如图 6 的(a)那样将低频作为评价对象的原因在于,虽然低频分量的噪声被压制,但声音分量并没有完全消失,通过将该分量加到评价中而生成的宽频带声音信号的质量得到提高。另外,通过如图 6 的(b)那样根据噪声的频率特性的逆特性进行失真评价,能够对 SN 比较高的高频进行加权,所以所生成的宽频带声音信号的质量得到提高。

[0087] 根据本实施方式 5,失真评价部 107 使用进行了频率轴上的加权的失真尺度,来评价波形失真。因此,通过对一部分低频分量进行加权来进行失真评价,从而能够提高所生成的声音信号的质量,复原更高质量的声音信号。

[0088] 另外,根据本实施方式 5,根据噪声的频率特性的逆特性进行加权而进行失真评价,从而能够提高所生成的声音信号的质量,复原更高质量的声音信号。

[0089] 另外,在上述实施方式 5 中,在已经压制噪声的声音信号的复原中实施了失真评价的加权,但也能够同样地应用于上述实施方式 1、2 的声音信号复原装置 100 的从窄频带声音信号向宽频带声音信号的复原。

[0090] 另外,在上述实施方式 1~5 中,作为窄频带声音信号的例子说明了电话声音的情况,但不限于电话声音,也能够应用于通过 MP3 (MPEG Audio Layer - 3) 等音响信号编码技术而截去了高频的信号的高频生成处理。另外,宽频带声音信号的频带也不限于 50~7000Hz,还能够在 50~16000Hz 等更宽的频带中实施。

[0091] 另外,在上述实施方式 1~5 所示的复原声音信号生成部 110 中,通过频带滤波器从声音信号切出特定的频带,并通过频带合成部而与其他的声音信号进行组合来生成复原声音信号,但不限于此,例如也可以对输入到复原声音信号生成部 110 的 2 种声音信号进行加

权相加来生成复原声音信号。图 7 示出将该结构的复原声音信号生成部 110 应用于上述实施方式 1 的声音信号复原装置 100 的情况的一个例子,并且图 8 示意性地图示复原声音信号。另外,图 8 中的各箭头表示处理的顺序,各曲线图的纵轴表示功率,横轴表示频率。

[0092] 如图 7 所示,复原声音信号生成部 110 新具备 2 个权重调整部 301、302。权重调整部 301 将从失真评价部 107 输出的宽频带声音信号的权重(增益)调整为例如 0.2 (图 8 的(a)所示的虚线),权重调整部 302 将从采样变换部 101 输出的已经上采样的声音信号的权重(增益)调整为例如 0.8 (图 8 的(b)所示的虚线),通过频带合成部 104 将两个声音信号进行相加(图 8 的(c)),生成复原声音信号(图 8 的(d))。

[0093] 另外,虽然省略了图示,但也可以将图 7 的结构应用于声音信号复原装置 200。

[0094] 在权重调整部 301、302 中,除了在频率方向上使用一定的权重以外,例如还使用具有随着成为高频而变大那样的频率特性的权重等与所需对应的权重即可。另外,既可以构成为具备权重调整部 301 和第 1 频带滤波器 103 这两者,且第 1 频带滤波器 103 从由权重调整部 301 进行了权重调整的宽频带声音信号抽出与窄频带声音信号相等的频带,相反地,也可以由第 1 频带滤波器 103 从宽频带声音信号抽出与窄频带声音信号相等的频带并通过权重调整部 301 进行权重调整。同样地,也可以构成为具备权重调整部 301 和第 2 频带滤波器 202 这两者。

[0095] 如上所述,本发明的声音信号复原装置根据从由音韵信号以及音源信号合成的多个宽频带声音信号选择出的宽频带声音信号和比较对象信号,生成复原声音信号,所以适用于复原如下比较对象信号的情况,其中,该比较对象信号是由于频带被限制为窄频带而导致一部分频带欠缺、或者由于噪声压制或声音压缩而导致一部分频带变差或缺损了的比较对象信号。另外,在由计算机构成声音信号复原装置 100、200 的情况下,也可以将记述了采样变换部 101、声音信号生成部 102、复原声音信号生成部 110、声音分析部 111、噪声压制部 201 的处理内容的程序保存到计算机的存储器中,并由计算机的 CPU 执行存储器中保存的程序。

[0096] 产业上的可利用性

[0097] 本发明的声音信号复原装置以及声音信号复原方法组合音韵信号以及音源信号来生成多个声音信号,使用规定的失真尺度分别评价与比较对象信号的波形失真,根据该评价结果来选择某一个声音信号而生成复原声音信号,所以适用于从频带被限制为窄频带的声音信号复原宽频带的声音信号、以及复原变差或者缺损了的频带的声音信号的声音信号复原装置及其方法。

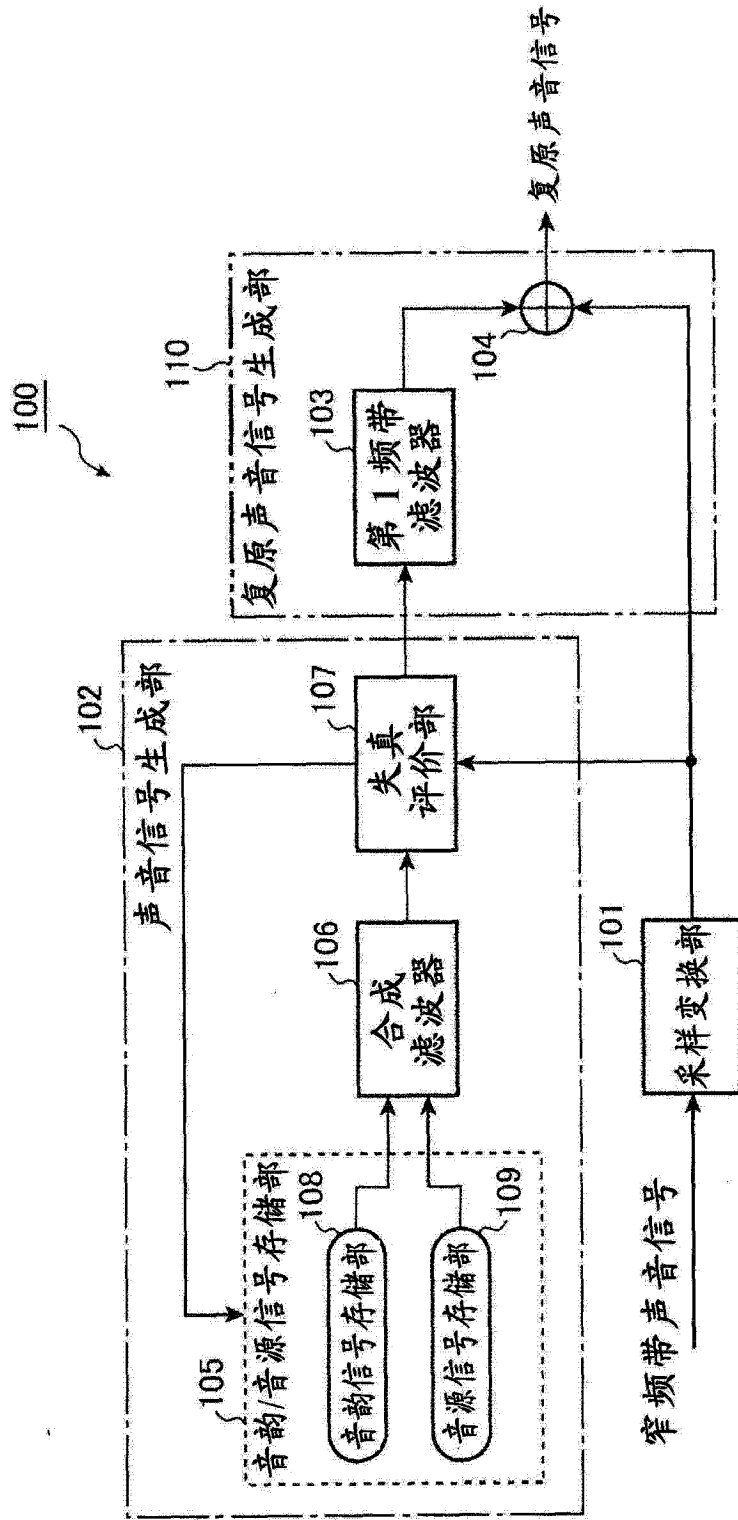


图 1

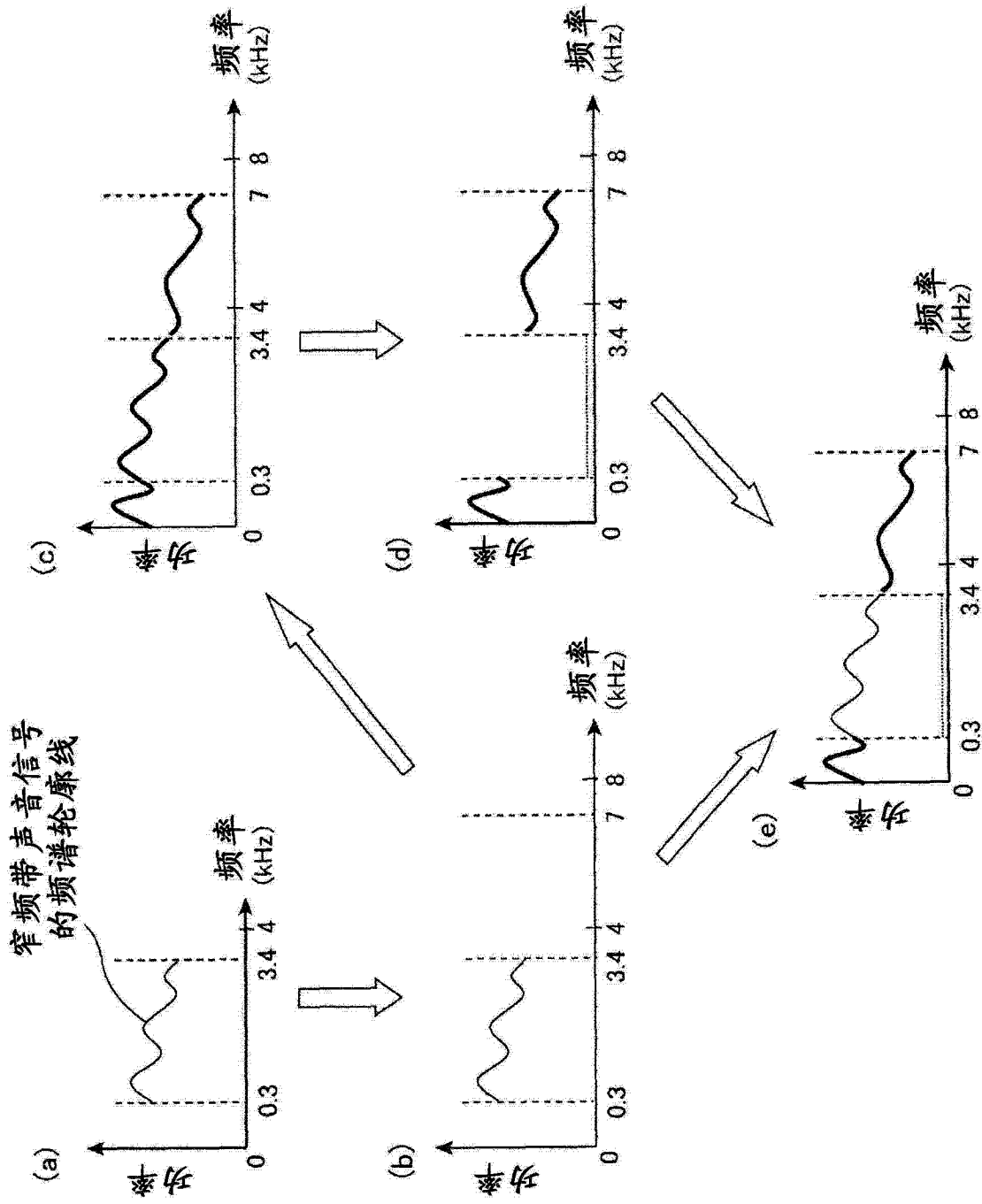


图 2

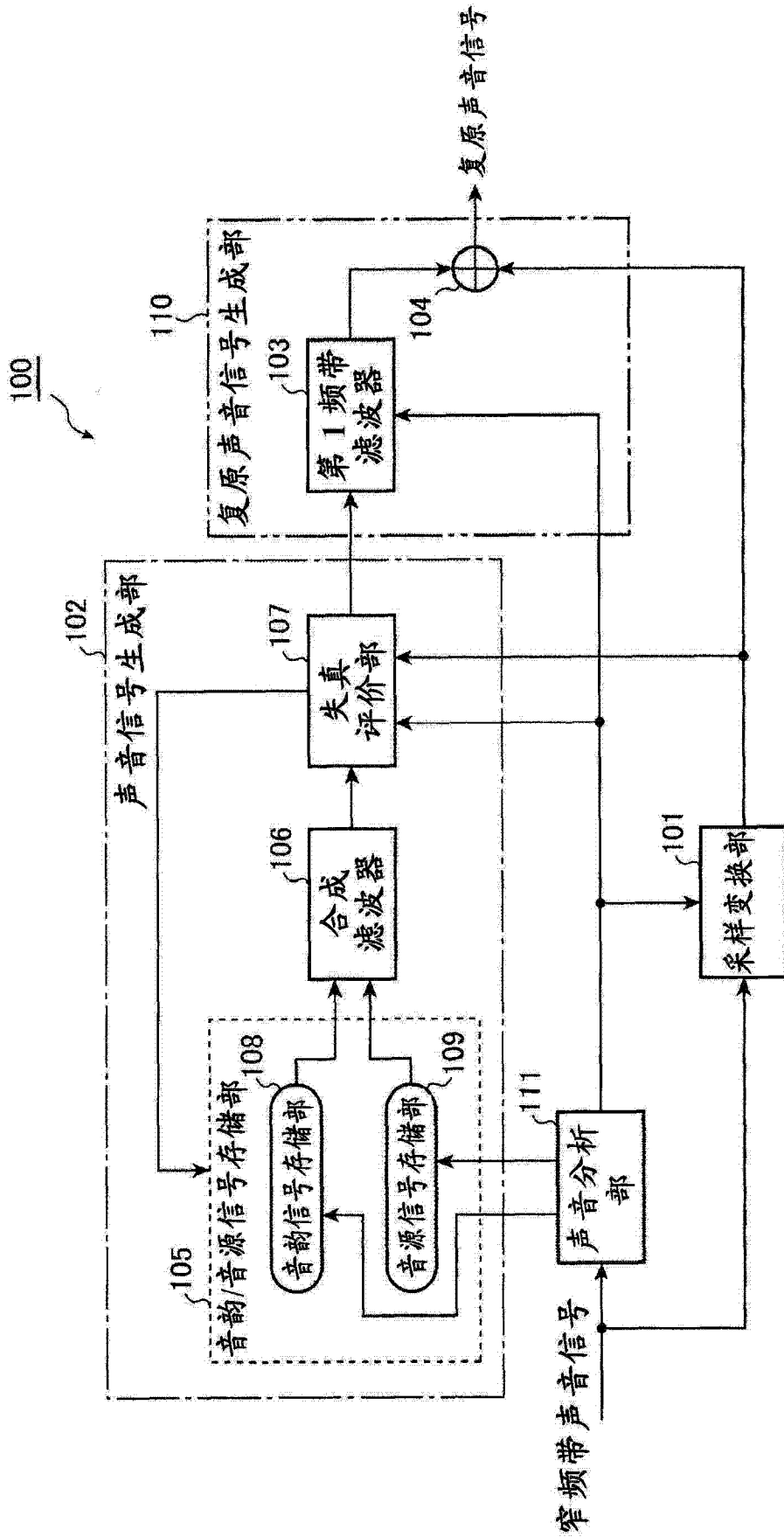


图 3

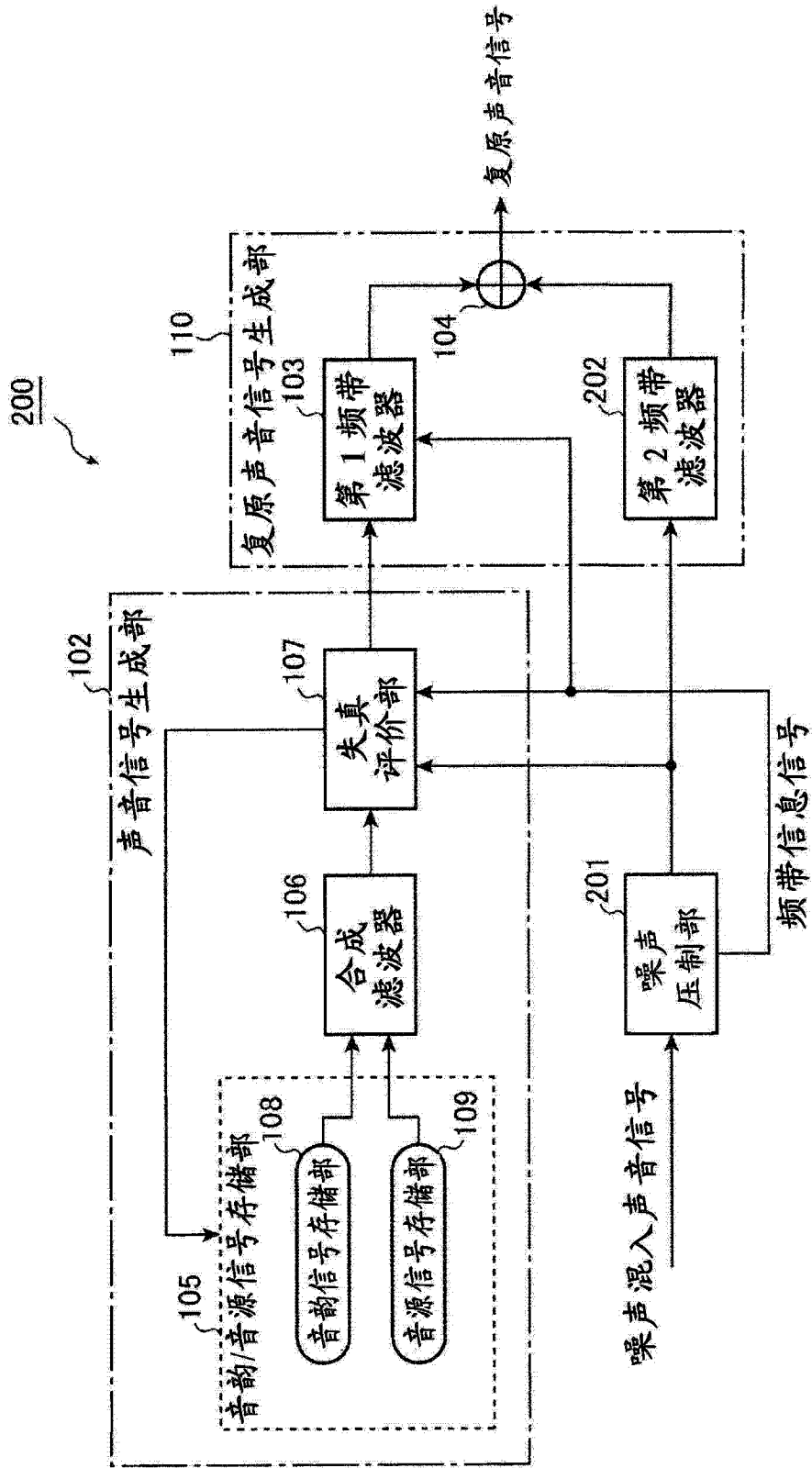


图 4

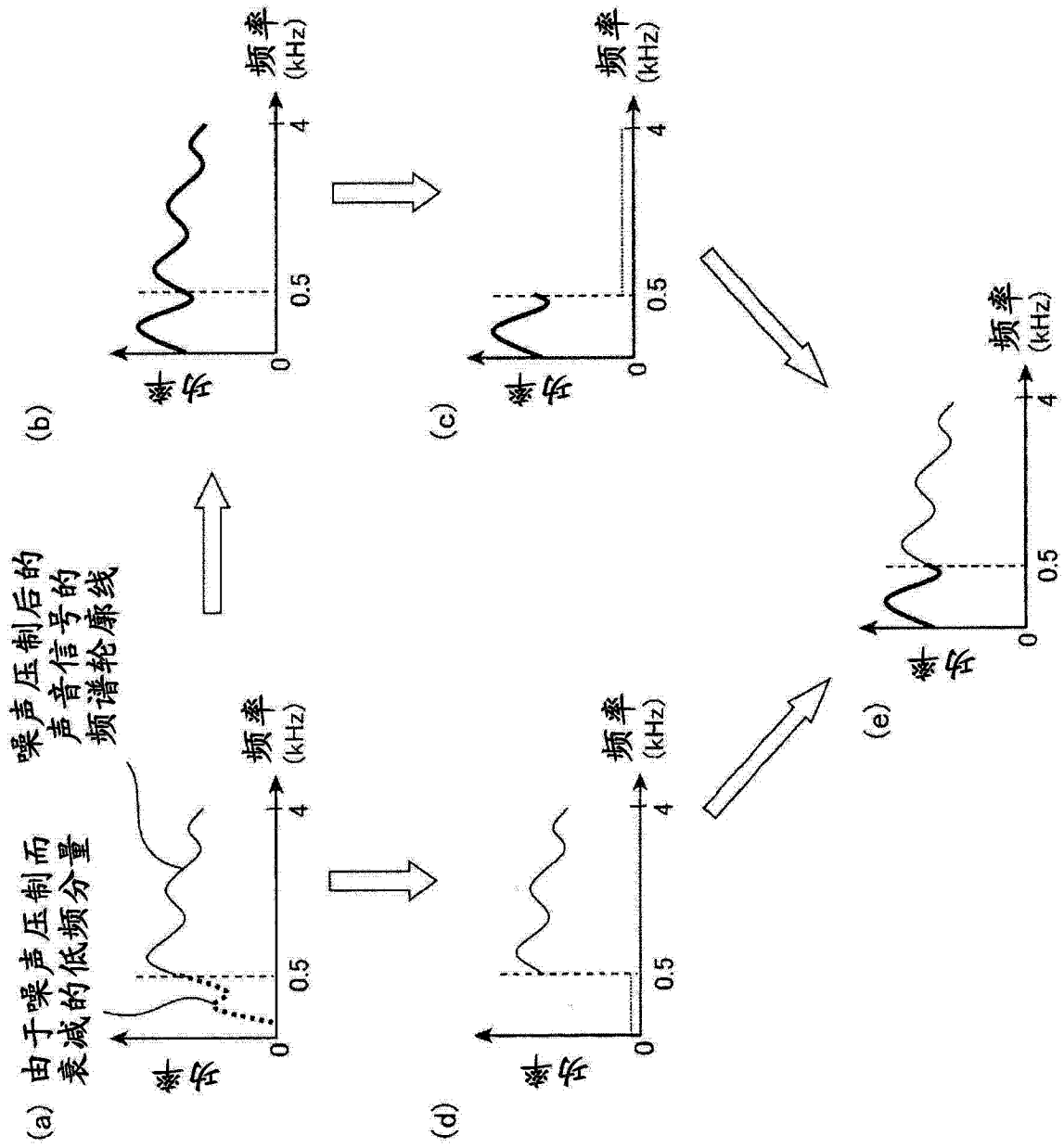


图 5

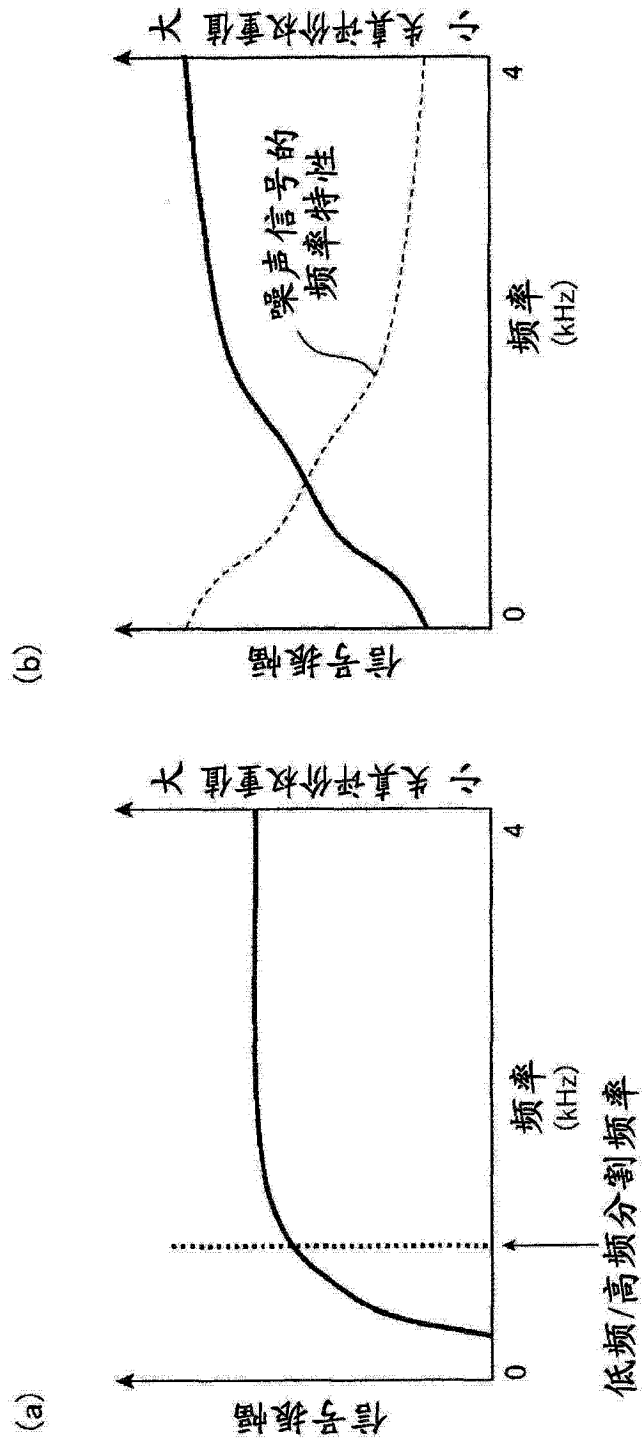


图 6

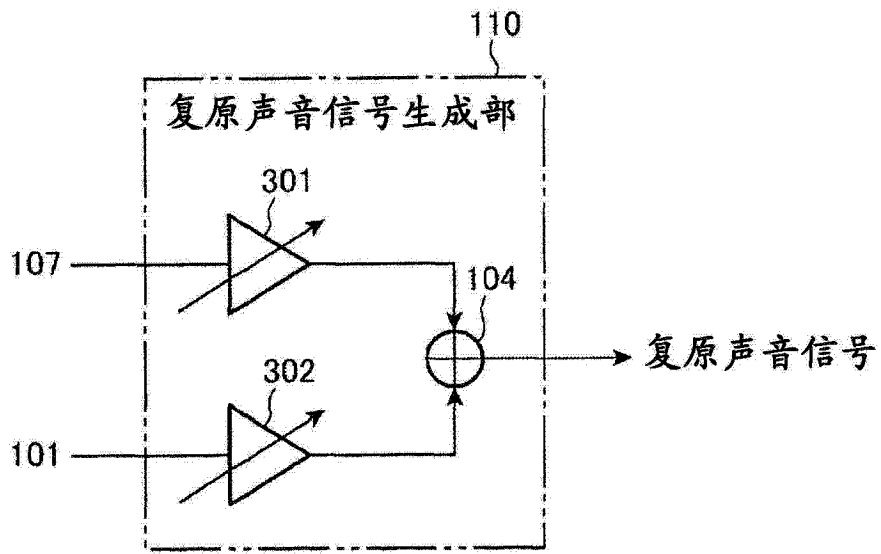


图 7

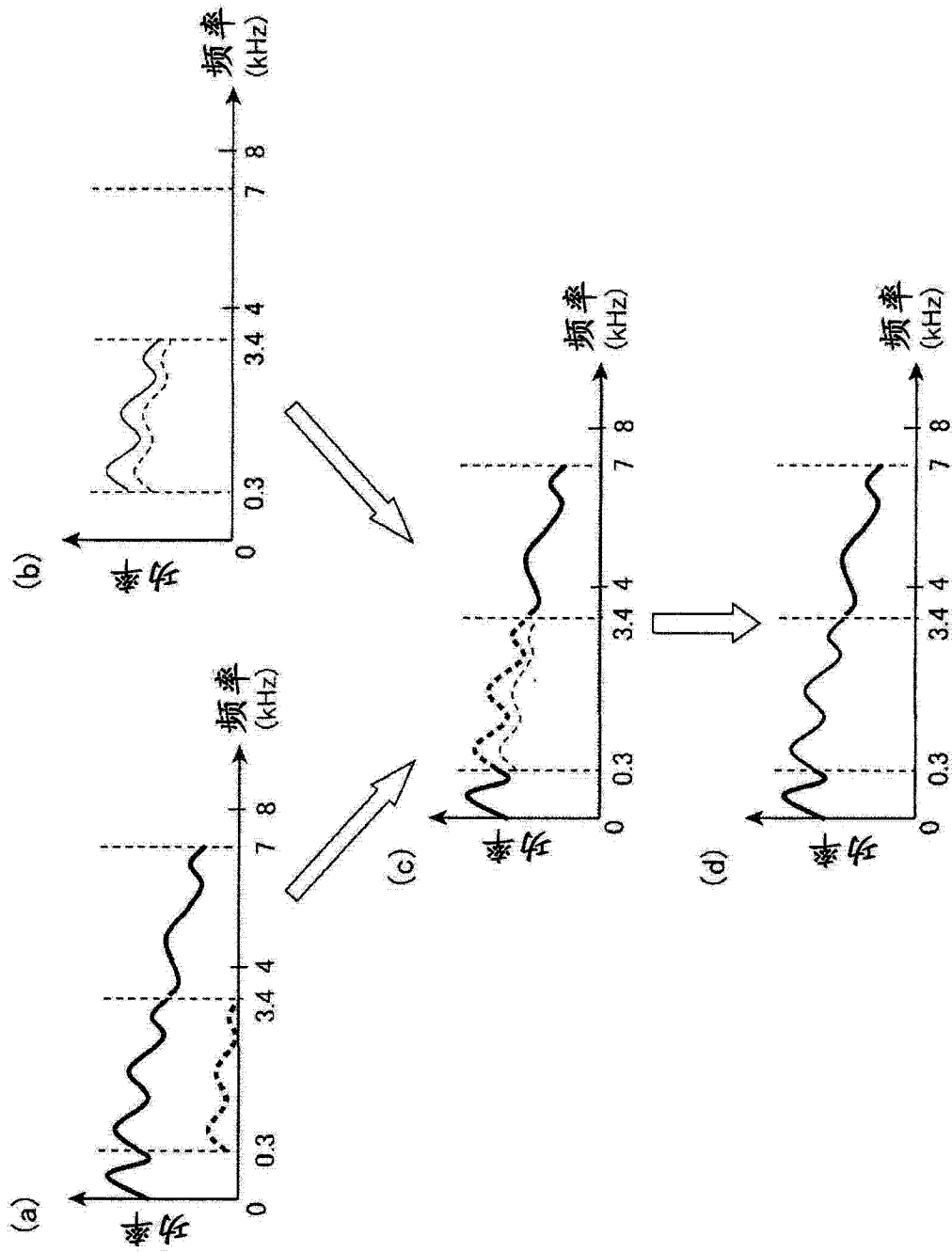


图 8