



(12) 发明专利

(10) 授权公告号 CN 101138020 B

(45) 授权公告日 2010. 10. 13

(21) 申请号 200680006668. 0

(22) 申请日 2006. 02. 23

(30) 优先权数据
056342/2005 2005. 03. 01 JP

(85) PCT申请进入国家阶段日
2007. 08. 31

(86) PCT申请的申请数据
PCT/JP2006/303290 2006. 02. 23

(87) PCT申请的公布数据
W02006/093019 JA 2006. 09. 08

(73) 专利权人 光荣株式会社
地址 日本兵库县

(72) 发明人 赤木正人 太长根理会子
入江佳洋 柳内久和 田中良种

(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 吴丽丽

(51) Int. Cl.

G10K 11/178(2006. 01)

G10L 21/04(2006. 01)

(56) 对比文件

WO 2004010627 A1, 2004. 01. 29, 全文.

JP 特开 2002-215198 A, 2002. 07. 31, 全文.

JP 特开 2002-251199 A, 2002. 09. 06, 全文.

WO 02054732 A1, 2002. 07. 11, 全文.

JP 特开 2000-3197 A, 2000. 01. 07, 全文.

JP 特开 2002-123298 A, 2002. 04. 26, 全文.

审查员 贾杨

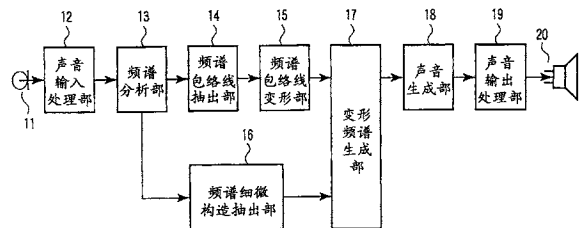
权利要求书 2 页 说明书 8 页 附图 15 页

(54) 发明名称

声音处理方法和装置及存储媒体以及声音系统

(57) 摘要

声音处理装置具有:抽出输入声音信号的频谱包络线的频谱包络线抽出部(14);为了生成变形频谱包络线而对频谱包络线实施变形的频谱包络线变形部(15);抽出输入声音信号的频谱细微构造的频谱细微构造抽出部(16);通过合成变形频谱包络线以及频谱细微构造,生成变形频谱的变形频谱生成部(17);基于变形频谱,生成输出声音信号的声音生成部(18),用输出声音信号发射用于不让第三者听到谈话声音的内容的干扰音。



1. 一种声音处理方法,其特征在于包括:
抽出输入声音信号的频谱包络线;
抽出上述输入声音信号的频谱细微构造;
对上述频谱包络线设定反转轴,通过以该反转轴为中心使上述频谱包络线反转,对上述频谱包络线实施变形,生成变形频谱包络线;
将上述变形频谱包络线以及上述频谱细微构造合成,生成变形频谱;以及
基于上述变形频谱,生成输出声音信号。
2. 一种声音处理方法,其特征在于包括:
抽出输入声音信号的频谱包络线;
抽出上述输入声音信号的频谱细微构造;
对上述频谱包络线实施变形,生成变形频谱包络线;
将上述变形频谱包络线以及上述频谱细微构造合成,生成变形频谱;
抽出上述输入声音信号的频谱的高频成分;
利用所抽出的上述高频成分置换包含在上述变形频谱中的高频成分;
基于上述高频成分被置换后的变形频谱,生成输出声音信号。
3. 一种声音处理装置,其特征在于包括:
频谱包络线抽出部,抽出输入声音信号的频谱包络线;
频谱细微构造抽出部,抽出上述输入声音信号的频谱细微构造;
频谱包络线变形部,为了生成变形频谱包络线,对上述频谱包络线设定反转轴,通过以该反转轴为中心使上述频谱包络线反转,对上述频谱包络线实施变形;
变形频谱生成部,通过合成上述变形频谱包络线以及上述频谱细微构造,生成变形频谱;以及
声音生成部,基于上述变形频谱生成输出声音信号。
4. 一种声音处理装置,其特征在于包括:
频谱包络线抽出部,抽出输入声音信号的频谱包络线;
频谱细微构造抽出部,抽出上述输入声音信号的频谱细微构造;
频谱包络线变形部,为了生成变形频谱包络线,对上述频谱包络线实施变形;
变形频谱生成部,通过合成上述变形频谱包络线以及上述频谱细微构造,生成变形频谱;
高频成分抽出部,抽出上述输入声音信号的频谱的高频成分;
高频成分置换部,用通过上述高频成分抽出部抽出的高频成分置换包含在上述变形频谱中的高频成分;以及
声音生成部,基于上述高频成分被置换后的变形频谱,生成输出声音信号。
5. 如权利要求 3 或者 4 的任意一项所述的声音处理装置,其特征在于:
上述频谱包络线变形部的构成是,在上述频谱包络线上对振幅方向以及频率轴方向的至少一个方向实施上述变形。
6. 如权利要求 3 或者 4 的任意一项所述的声音处理装置,其特征在于:
上述频谱包络线变形部通过改变上述频谱包络线的峰和谷的位置,实施上述变形。
7. 如权利要求 4 所述的声音处理装置,其特征在于:

上述频谱包络线变形部的构成是,对上述频谱包络线设定反转轴,将该反转轴作为中心通过使上述频谱包络线反转来实施上述变形。

8. 如权利要求 3 或者 4 的任意一项所述的声音处理装置,其特征在于:

上述频谱包络线变形部的构成是,通过在频率轴上使上述频谱包络线移位来实施上述变形。

9. 如权利要求 4 所述的声音处理装置,其特征在于:

上述高频成分置换部对用上述高频成分抽出部抽出的高频成分设定置换频带,用上述置换频带内的高频成分置换包含在上述变形频谱中的高频成分。

10. 一种声音系统,其特征在于包括:

麦克风,为了得到上述输入声音信号而收集谈话声音;

权利要求 3 或者 4 的任意 1 项所述的声音处理装置;以及

扬声器,根据上述输出声音信号,发射干扰音。

声音处理方法和装置及存储媒体以及声音系统

技术领域

[0001] 本发明涉及在防止谈话声音的内容被第三者听到的声音系统以及在该声音系统中使用的声音处理方法和装置以及存储媒体。

背景技术

[0002] 如果在公开的场所和隔音室以外的房间中进行谈话,则有时谈话声音向周围泄漏出现问题。例如,在银行内顾客和职员进行谈话,或者在医院内外来患者和接待人员或者医生进行谈话时,如果谈话被第三者听到,则有可能损害机密和个人隐私。

[0003] 因而,提出了利用遮掩效果不让第三者听到谈话的方法(例如,参照佐伯徹郎、藤井健生、山口静马、老末建成(2003)“用于掩盖声音的无意义恒定杂音的选定”,电子信息通信学会论文志, J86-A, 2, 187-191. 以及特开平 5-22391 号公报)。所谓遮掩效果是在能够听到某一声音时如果让其听到一定级别及以上的另一声音,则原来的声音一下被覆盖消失不能听见的现象。作为利用这种遮掩效果不让第三者听到原声音的技术,有把粉红噪声(pink noise)和背景音乐(background music)(BGM)等的声音作为遮掩音,在原来的声音上重叠的方法。如在佐伯徹郎、藤井健生、山口静马、老末建成(2003)“用于掩盖声音的无意义恒定杂音的选定”,电子信息通信学会论文志, J86-A, 2, 187-197 中提出的那样,进行了频带限制后的粉红噪声作为遮掩音被认为最有效。

[0004] 为了把粉红噪声和 BGM 这一恒定发出的声音作为遮掩音使用,需要大于等于原声音级别的级别。因而,这种遮掩音对于听的人来说都感觉是一种噪音,在银行和医院等中使用是困难的。另一方面,如果降低遮掩音的级别则遮掩效果差,特别是在遮掩效果小的频率区域中原来的声音会感知到。进而,即使适宜地调整了遮掩音的级别,因为粉红噪声和 BGM 那样的声音与原本的声音明确地分离能够听到,所以由于在许多音混合存在中能够只听到特定的音的人的听觉特性,所谓的鸡尾酒会效应的作用,有可能原来的声音被听到。

[0005] 本发明的目的在于不会令周围的人感到吵闹,并且不让第三者知道谈话声音的内容。

[0006] 为了解决上述问题,根据本发明的一实施方式,抽出输入声音信号的频谱包络线和频谱细微构造,对频谱包络线实施变形生成变形频谱包络线,通过合成变形频谱包络线以及频谱细微构造生成变形频谱,基于变形频谱生成输出声音信号。

[0007] 根据本发明的另一实施方式,抽出输入声音信号的频谱的高频成分,利用所抽出的高频成分置换包含在变形频谱中的高频成分,基于高频成分被置换的变形频谱,生成输出声音信号。

附图说明

[0008] 图 1 是概略表示本发明的一种实施方式所涉及的声音系统的图。

[0009] 图 2A 是表示在图 1 的声音系统中用微型麦克风收集声音的谈话声音的频谱的一个例子的图。

- [0010] 图 2B 是表示在图 1 的声音系统中从扬声器发射出的干扰音的频谱的图。
- [0011] 图 2C 是表示在图 1 的声音系统中干扰音和谈话声音的融合声音的频谱的一个例子的图。
- [0012] 图 3 是表示本发明的第 1 实施方式所涉及的声音处理装置的结构方框图。
- [0013] 图 4 是表示频谱分析和在频谱分析中附带的处理的一个例子的流程图。
- [0014] 图 5A 是表示输入声音信号的声音频谱的一个例子的图。
- [0015] 图 5B 是表示图 5A 的声音频谱的频谱包络线的一个例子的图。
- [0016] 图 5C 是表示将图 5B 的频谱包络线进行变形后的变形频谱包络线的一个例子的图。
- [0017] 图 5D 是表示图 5A 的声音频谱的频谱细微结构的的一个例子的图。
- [0018] 图 5E 是表示通过合成图 5C 的变形频谱和图 5D 的频谱细微构造而生成的变形频谱的一个例子的图。
- [0019] 图 6 是表示在第 1 实施方式中的声音处理的全体流程的流程图。
- [0020] 图 7A 是声音频谱的频谱包络线的一个例子的图。
- [0021] 图 7B 是表示在第 1 实施方式中对频谱包络线实施振幅方向的频谱变形的方法的第 1 个例子的图。
- [0022] 图 7C 是说明在第 1 实施方式中对频谱包络线实施振幅方向的频谱变形的方法的第 2 个例子的图。
- [0023] 图 7D 是说明在第 1 实施方式中对频谱包络线实施振幅方向的频谱变形的方法的第 3 个例子的图。
- [0024] 图 7E 是说明在第 1 实施方式中对频谱包络线实施振幅方向的频谱变形的方法的第 4 个例子的图。
- [0025] 图 8A 是表示声音频谱的频谱包络线的一个例子的图。
- [0026] 图 8B 是说明在第 1 实施方式中对频谱包络线实施频率轴方向的频谱变形的方法的第 1 个例子的图。
- [0027] 图 8C 是说明在第 1 实施方式中对频谱包络线实施频率轴方向的频谱变形的方法的第 2 个例子的图。
- [0028] 图 9A 是表示摩擦音的频谱的一个例子的图。
- [0029] 图 9B 是表示摩擦音的频谱包络线的一个例子的图。
- [0030] 图 9C 是说明在第 1 实施方式中对摩擦音的频谱包络线实施振幅方向的频谱变形的方法的第 1 个例子的图。
- [0031] 图 9D 是说明在第 1 实施方式中对摩擦音的频谱包络线实施振幅方向的频谱变形的方法的第 2 个例子的图。
- [0032] 图 10 是表示本发明的第 2 实施方式所涉及的声音处理装置的结构方框图。
- [0033] 图 11 是表示在第 2 实施方式中的频谱包络线变形部的处理和低频成分抽出部的处理的一部分的流程图。
- [0034] 图 12A 是表示低频成分强的输入声音信号的声音频谱的一个例子的图。
- [0035] 图 12B 是表示图 12A 的声音频谱的频谱包络线的图。
- [0036] 图 12C 是表示在第 2 实施方式中将图 12A 的声音频谱进行变形后的变形频谱的一

个例子的图。

[0037] 图 12D 是表示在第 2 实施方式中置换在图 12C 的变形频谱中的高频成分所生成的干扰音的频谱的一个例子的图。

[0038] 图 13A 是表示高频成分强的输入声音信号的声音频谱的一个例子的图。

[0039] 图 13B 是表示图 13A 的声音频谱的频谱包络线的图。

[0040] 图 13C 是表示在第 2 实施方式中将图 13A 的声音频谱进行变形后的变形频谱的一个例子的图。

[0041] 图 13D 是表示在第 2 实施方式中置换在图 13C 的变形频谱中的高频成分所生成的干扰音的频谱的一个例子的图。

[0042] 图 14 是表示在第 2 实施方式中的声音处理的整体的流程的流程图。

具体实施方式

[0043] 以下参照附图说明本发明的实施方式。

[0044] 图 1 表示包含本发明的一种实施方式的声音处理装置 10 的声音系统的概念图。声音处理装置 10, 在图中用配置在多名人 1 和人 2 进行谈话的场所的附近的位置 A 上的微型麦克风 11 收集谈话声音, 对所得到的输入声音信号进行处理生成输出声音信号。把从声音处理装置 10 输出的输出声音信号提供给配置在位置 B 上的扬声器 20, 从扬声器 20 发射声音。

[0045] 此时在输出声音信号中, 如果输入声音信号的音源信息得到维持而音韵性被破坏, 则通过把从扬声器 20 发射的声音和谈话声音的声音进行融合, 则在位置 C 上的人 3 不能听取人 1 和人 2 的谈话声音。因为从扬声器 20 发射的声音的目的就是妨碍第三者听取这样的谈话声音, 所以以后称为干扰音。换句话说, 因为从扬声器 20 发射的声音的目的在于防止谈话声音被第三者听见, 所以也可以称为“防听音”。

[0046] 声音处理装置 10 通过对输入声音信号实施处理, 如上所示在维持输入声音信号的音源信息的同时生成破坏音韵性的输出声音信号。根据该输出声音信号, 从扬声器 20 发射破坏了谈话声音的音韵性的干扰音。例如, 如果把用微型麦克风 11 收集声音的谈话声音的频谱作为图 2A, 则经由声音处理装置 10 从扬声器 20 发射的干扰音的频谱例如变成图 2B 所示。在这种情况下, 在图 1 的 C 位置上, 第三者能够听见干扰音和谈话声音的直接音被融合后的具有如图 2C 所示那样的频谱的声音。

[0047] 以下, 详细说明声音处理装置 10 的实施方式。

[0048] (第 1 实施方式)

[0049] 图 3 表示第 1 实施方式的声音处理装置的结构。把微型麦克风 11 例如设置在银行的窗口附近和医院的门诊挂号处等的场所, 收集谈话声音输出声音信号。把来自微型麦克风 11 的声音信号输入到声音输入处理部 12。声音输入处理部 12 例如具有放大器以及 A/D 变换器, 在放大了来自微型麦克风 11 的声音信号(以后, 称为输入声音信号)后, 进行数字化并输出。把来自声音输入处理部 12 的经过数字化的输入声音信号输入到频谱分析部 13。频谱分析部 13 例如用 FFT 对数倒频谱分析、自动语音合成方式的声音分析合成系统的处理进行输入声音信号的分析。

[0050] 用图 4 说明在频谱分析部 13 中使用了对数倒频谱分析时的频谱分析的流程。首

先,对于经过数字化的输入声音信号,例如在加上汉宁(hanning)窗或者汉明(hamming)窗等的时间窗后,采用高速傅立叶变换(FFT)进行短时间频谱分析(步骤S1~S2)。接着,取FFT结果的绝对值(振幅频谱)的对数(步骤S3),进而进行反FFT(IFFT)得到对数倒频谱系数(步骤S4)。以下,对对数倒频谱系数采用对数倒频谱窗进行滤波(liftering),把低倒频(quefreny)部和高倒频(quefreny)部作为对数倒频谱分析结果输出(步骤S5)。

[0051] 在作为频谱分析部13的分析结果得到的对数倒频谱系数中,把低倒频部输入到频谱包络线抽出部14。把对数倒频谱系数中,高倒频部输入到频谱细微构造抽出部16。频谱包络线抽出部14抽出输入声音信号的声音频谱的频谱包络线。频谱包络线表示输入声音信号的音韵信息。例如,如果设输入声音信号的声音频谱为图5A,则频谱包络线表示为图5B。频谱包络线的抽出例如如图4中所示那样通过对对数倒频谱系数的低倒频部实施FFT(步骤S6)进行。

[0052] 对于抽出的频谱包络线使用频谱包络线变形部15实施变形,生成变形频谱包络线。如果设抽出的频谱包络线为图5B,则在频谱包络线变形部15中,如图5C所示通过频谱包络线反转,对频谱包络线实施变形。例如,当在频谱分析部13中使用了FFT对数倒频谱分析的情况下,频谱包络线使用低次的对数倒频谱系数表现。频谱包络线变形部15对于这种低次的对数倒频谱系数进行符号反转。频谱包络线变形部15的更具体的例子以后详细说明。

[0053] 另一方面,频谱细微构造抽出部16抽出输入声音信号的声音频谱的频谱细微构造。频谱细微构造表示输入声音信号的音源信息。例如,如果设输入声音信号的声音频谱为图5A,则频谱细微构造表示为图5D。频谱细微构造的抽出例如如图4中所示通过对对数倒频谱系数的高倒频部实施FFT(步骤S7)而实现。

[0054] 把用频谱包络线变形部15生成的变形频谱包络线、用频谱细微构造抽出部16抽出的频谱细微构造输入到变形频谱生成部17。变形频谱生成部17通过合成变形频谱包络线和频谱细微构造,生成将输入声音信号的声音频谱变形后的频谱即变形频谱。例如,如果设变形频谱包络线为图5C,频谱细微构造为图5D,则通过合成它们生成的变形频谱表示为图5E。

[0055] 把使用变形频谱生成部17生成的变形频谱输入到声音生成部18。声音生成部18生成根据变形频谱进行了数字化的输出声音信号。把经过数字化的输出声音信号输入到声音输出处理部19。声音输出处理部19用D/A变换器把输出声音信号变换为模拟信号,进一步用功率放大器放大并提供给扬声器20。由此,从扬声器20发射干扰音。

[0056] 在图1以及图3中,虽然表示微型麦克风11以及扬声器20是各自1个的情况,但微型麦克风的数量以及扬声器的数量也可以是2个或者2个以上。在这种情况下,声音处理装置对于来自多个微型麦克风的多个通道的输入声音信号各自进行处理,只要从许多扬声器发射干扰音即可。

[0057] 图3所示的声音处理装置10还能够通过数字信号处理装置(DSP)那样的硬件来实现,但使用计算机用程序实现也可以。以下,说明使用图6用计算机实现声音处理装置10的处理时的处理步骤。

[0058] 对在步骤S101中输入的经过数字化的输入声音信号经过频谱分析(步骤S102)抽出频谱包络线(步骤S103),如上所述,进行频谱包络线的变形(步骤S104)以及频谱细

微构造的抽出（步骤 S105）。在此，步骤 S103 以及 S104 和 S105 的处理顺序是任意的。接着，还可以并行进行步骤 S103 及 S104 的处理和步骤 S105 的处理。以下，通过合成经由步骤 S103 以及 S104 生成的变形频谱包络线和通过步骤 S105 生成的频谱细微构造，生成变形频谱（步骤 S106）。最后，从变形频谱中生成声音信号并输出（步骤 S107 ~ S108）。

[0059] 以下，说明频谱包络线的变形方法的具体例子。频谱包络线的变形基本上通过改变频谱包络线的共振峰频率（即，频谱包络线的峰以及谷的位置）来实现。在此的频谱包络线的变形的目的是破坏音韵。因为在音韵的感知中频谱包络线的峰以及谷的位置关系重要，所以使这些峰以及谷的位置和变形前不同。具体地说，可以通过对频谱包络线在振幅方向以及频率轴方向的至少一个方向上实施变形来实现。

[0060] < 频谱包络线的变形方法 1 >

[0061] 图 7A、图 7B、图 7C、图 7D 以及图 7E 表示通过对频谱包络线实施振幅方向的变形来改变峰和谷的位置的方法。频谱包络线变形部 15 因为在振幅方向上改变频谱包络线，所以对图 7A 所示的频谱包络线设定反转轴，以该反转轴作为中心使频谱包络线反转。作为反转轴能够使用各种近似函数。例如，图 7B 是用 \cos 函数设定了反转轴的例子，图 7C 是用直线设定了反转轴的例子，此外图 7D 是用对数设定了反转轴的例子。另一方面，图 7E 是把反转轴设定为频谱包络线的振幅的平均，即设定为与频率轴平行的例子。在图 7B、图 7C、图 7D、图 7E 的任何一个例子中，都可知相对图 7A 的原来的频谱包络线而言峰和谷的位置（频率）变化。

[0062] < 频谱包络线的变形方法 2 >

[0063] 图 8A、图 8B 以及图 8C 表示通过对频谱包络线实施频率轴方向的变形改变峰和谷的位置的方法。为了在频率轴方向改变频谱包络线，使图 8A 所示的频谱包络线如图 8B 所示那样向低频一侧移位，或者如图 8C 所示那样向高频一侧移位。作为频谱包络线的频率轴方向的变形法，除此以外还考虑在频率轴上实施线性伸缩或者非线性伸缩的方法等。此外，为了在频率轴方向改变频谱包络线，还能够组合在频率轴上的移位或者伸缩。进而，无须对频谱包络线的全区域进行频率轴上的变形，可以局部进行。

[0064] < 频谱包络线的变形方法 3 >

[0065] 在上述的频谱包络线的变形方法 1 以及 2 中，因为进行改变输入声音信号的频谱的低频成分的处理，所以如母音那样在第 1 以及第 2 共振峰处于低频的音韵中有效。但是，变形方法 1 以及 2 是第 2 共振峰处于高频的 /e/、/i/，和在高频有特征的摩擦音 /s/、破裂音 /k/ 等中效果不大。因此，希望把改变频谱包络线的对象的频带和反转轴与音韵的频谱形状一致地进行动态控制。

[0066] 例如，当是在摩擦音那样的高频区域上有特征的音韵的情况下，即使改变频谱包络线的峰以及谷的位置，频谱包络线的特征也几乎不改变。图 9A 表示摩擦音的频谱，图 9B 表示摩擦音的频谱包络线。如果把图 9B 的频谱包络线例如和图 7B 一样以 \cos 函数的反转轴为中心反转，则变成如图 9C 所示，频谱包络线的特征变化少。在这种情况下，例如如图 9D 所示，通过和图 7E 一样把设定为频谱包络线的振幅的平均的反转轴为中心使频谱包络线反转，能够使特征变化显著。这是一个例子，只要是频谱包络线的特征显著变化那样的变形即可。

[0067] 如上所述，在第 1 实施方式中，改变输入声音信号的频谱包络线生成变形频谱包

络线,把该变形频谱包络线与输入声音信号的频谱细微构造合成生成变形频谱,根据该变形频谱生成输出声音信号。

[0068] 因而,如图 1 所示那样用配置在位置 A 上的微型麦克风 11 收集谈话声音,对得到的输入声音信号进行上述的处理生成输出声音信号,如果使用输出声音信号从配置在位置 B 上的扬声器 20 发射谈话声音的音韵性受到破坏的干扰音,则在位置 C 上对第三者而言因为干扰音和谈话音的直接音在感觉上融合,所以谈话声音变得不清楚。其结果,谈话声音的内容难以被第三者知道。

[0069] 即,在干扰音中,一边维持由谈话声音产生的输入声音信号的频谱细微构造即音源信息,一边破坏用频谱包络线的形状决定的音韵性。因此,干扰音和谈话声音的直接音很好地融合。因而,如果使用这种干扰音,则不会如使用了粉红噪声和 BGM 这种遮掩音的情况那样让周围感到吵闹,且谈话声音的内容可以不被第三者知道。

[0070] (第 2 实施方式)

[0071] 以下,说明本发明的第 2 实施方式。图 10 表示涉及第 2 实施方式的声音处理装置,对于图 3 所示的涉及第 1 实施方式的声音处理装置追加频谱高频成分抽出部 21 和高分置换部 22。

[0072] 频谱高频成分抽出部 21 经由频谱分析部 13 抽出输入声音信号的频谱的高频成分。频谱的高频成分表示个性信息,例如能够从图 4 中的步骤 S2 的 FFT 结果(输入声音信号的频谱)中抽出,抽出的高频成分被输入到高频成分置换部 22。把高频成分置换部 22 插入到变形频谱生成部 17 的输出和声音生成部 18 的输入之间,进行用通过频谱高频成分抽出部 21 抽出的高频成分置换用变形频谱生成部 17 生成的变形频谱中的高频成分的处理。声音生成部 18 根据高频成分置换后的变形频谱生成输出声音信号。

[0073] 图 11 表示频谱包络线变形部 15 进行图 7B、图 7C 以及图 7D 所示的频谱包络线变形时的处理,和高分置换部 22 的处理的一部分。频谱包络线部 15 检测频谱包络线的斜率(步骤 S201)。接着,频谱包络线变形部 15 根据在步骤 S201 中检测到的频谱包络线的斜率,决定例如 \cos 函数、直线或者对数这样的近似函数(步骤 S202),根据该近似函数进行频谱包络线的反转(步骤 S203)。该频谱包络线变形部 15 的处理和第 1 实施方式一样。

[0074] 另一方面,高分置换部 22 根据从步骤 S201 检测的频谱包络线的斜率决定置换区域,根据由频谱高频成分抽出部 21 抽出的高频成分置换该置换频带内的频谱成分即高频成分。

[0075] 接着,用图 12A~图 12D 以及图 13A~图 13D 说明第 2 实施方式中的具体的处理的例子。例如,当如图 12A 所示那样输入声音信号如母音部那样是低频成分强的频谱的情况下,输入声音信号的频谱包络线如图 12B 所示那样表示负的斜率。在这种情况下,例如通过合成以例如按照上述的 \cos 函数、直线或者对数这种近似函数的反转轴为中心进行了频谱包络线反转的变形频谱包络线,和输入声音信号的频谱构造,生成图 12C 所示的变形频谱。

[0076] 接着,在图 12C 的变形频谱中,对于包含音韵信息的低频成分(例如,2.5~3kHz 以下的频率成分)保持原样,把包含个性信息的高频成分(例如,大于等于 3kHz 的频率成分)用图 12A 的原本的声音频谱的高频成分进行置换,由此生成图 12D 所示那样的频谱的干扰音。在这种情况下,还考虑可以根据频谱包络线的谷位置改变置换频带的下限频率。于是,不管发言人的性别和音质如何都可以决定包含个性信息的频带。

[0077] 另一方面,如图 13A 所示当输入声音信号是如摩擦音和破裂音那样的高频成分强的频谱的情况下,输入声音信号的频谱包络线如图 13B 所示那样表示正的斜率。在这种情况下,通过合成例如如上所述那样以设定为频谱包络线的振幅的平均的反转轴为中心使频谱包络线反转的变形频谱包络线,和输入声音信号的频谱细微构造,生成图 13C 所示的变形频谱。

[0078] 以下,对在图 13C 的变形频谱中包含音韵信息的低频成分保持原样,通过用图 13A 的原本的声音频谱的高频成分置换包含个性信息的高频成分,生成如图 12D 所示那样的频谱的干扰音。但是,在摩擦音等的情况下,因为输入声音信号的频谱的高频成分特别强,所以把置换频带设置在更高频一侧,例如 6kHz 及以上的频带。在这种情况下,还可以根据频谱包络线的峰的位置改变置换频带的下限频率。于是,不管发言人的性别和音质都可以决定包含个性信息的频带。

[0079] 对于图 10 所示的声音处理装置也能够用 DSP 那样的硬件实现,但也可以使用计算机用程序实现。进而,如果采用本发明则能够提供存储有该程序的存储媒体。

[0080] 以下,如果用图 14 说明用计算机实现声音处理装置的处理时的处理步骤,则从步骤 S101 到步骤 S106 的处理和第 1 实施方式相同。在实施方式 2 中,在生成变形频谱的步骤 S106 之后,进行频谱高频成分的抽出(步骤 S109)以及高频成分的置换(步骤 S110)。接着,从高频成分置换后的变形频谱中生成声音信号并输出(步骤 S107 ~ S108)。在此,步骤 S103 ~ S105 以及步骤 S109 的处理顺序是任意的,此外也可以并行进行步骤 S103 以及 S104 的处理和步骤 S105 的处理,或者并行进行步骤 S109 的处理。

[0081] 如上所述,在第 2 实施方式中使用把通过变形频谱包络线和频谱细微构造的合成而生成的变形频谱的高频成分置换为输入声音信号的高频成分的变形频谱来生成输出声音信号。因而,在用频谱包络线的变形破坏谈话声音的音韵性的同时,能够生成保存了谈话声音的频谱的高频成分即个性信息的干扰音。即,不会因频谱包络线的反转而使干扰音的高频区域的功率增大使音质降低,此外不会因在干扰音中谈话声音的个性的信息也被破坏,使干扰音和谈话声音的融合效应不充分。由此周围不会感到吵闹,能够进一步显著发挥不让第三者听到谈话声音的内容的效果。

[0082] 在第 2 实施方式中,是在生成将变形频谱包络线和频谱细微构造的合成而产生的变形频谱后,进行高频成分的置换,生成高频成分经过置换的变形频谱,但即使只对高频成分以外的频带(低频以及中频)有选择地进行频谱包络线的变形也能够得到同样的效果。

[0083] 如上所述,如果采用本发明的形态,则能够从谈话声音产生的输入声音信号中生成通过频谱包络线的变形而破坏了音韵性的输出声音信号。因而,通过使用该输出声音信号发射干扰音,能够不让第三者听到谈话声音的内容,在保密性和隐私保护方面有效。

[0084] 即,在本发明的形态中因为用在变形频谱包络线中合成了输入声音信号的频谱细微构造的变形频谱生成输出声音信号,所以维持发言人的音源信息,即使具有鸡尾酒会效应这种人的听觉特性,也把原本的谈话声音和干扰音在感觉上融合。由此,对第三者来说谈话声音变得不清楚,难以感知。因而,能够保护谈话的机密和隐私。

[0085] 在这种情况下,因为无须如使用以往的遮掩音的方法那样提高干扰音的级别,所以不会让周围感到吵闹。进而,通过用输入声音信号的频谱的高频成分置换包含在变形频谱中的高频成分,所以能够在干扰音中保存谈话声音的个性的信息,谈话声音和干扰音的

感知融合效应进一步提高。

[0086] 本发明可以在防止谈话声音的内容,或者在手机及其他的电话机的通话人的谈话内容被第三者听到的技术中利用。

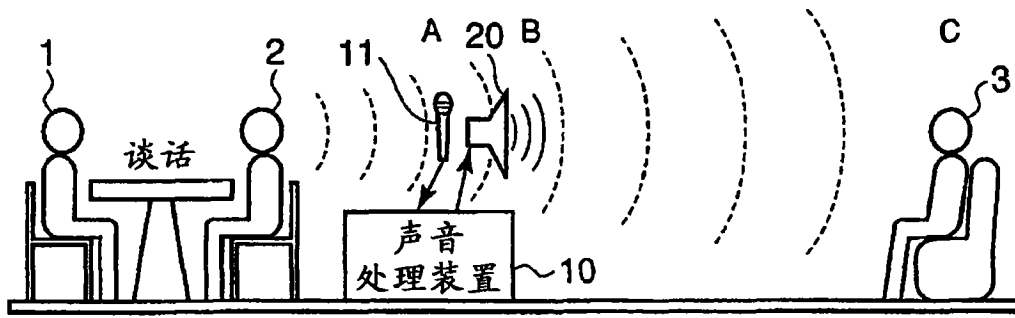


图 1

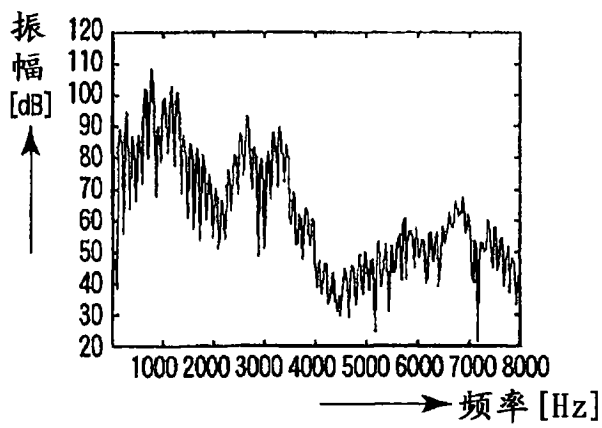


图 2A

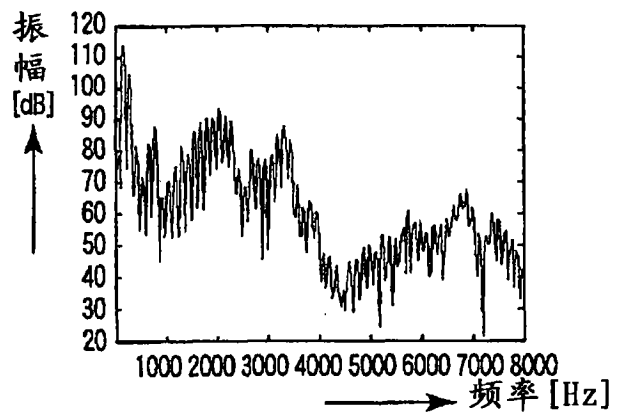


图 2B

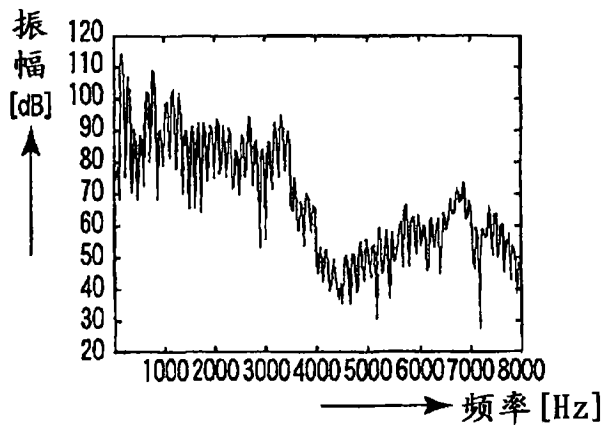


图 2C

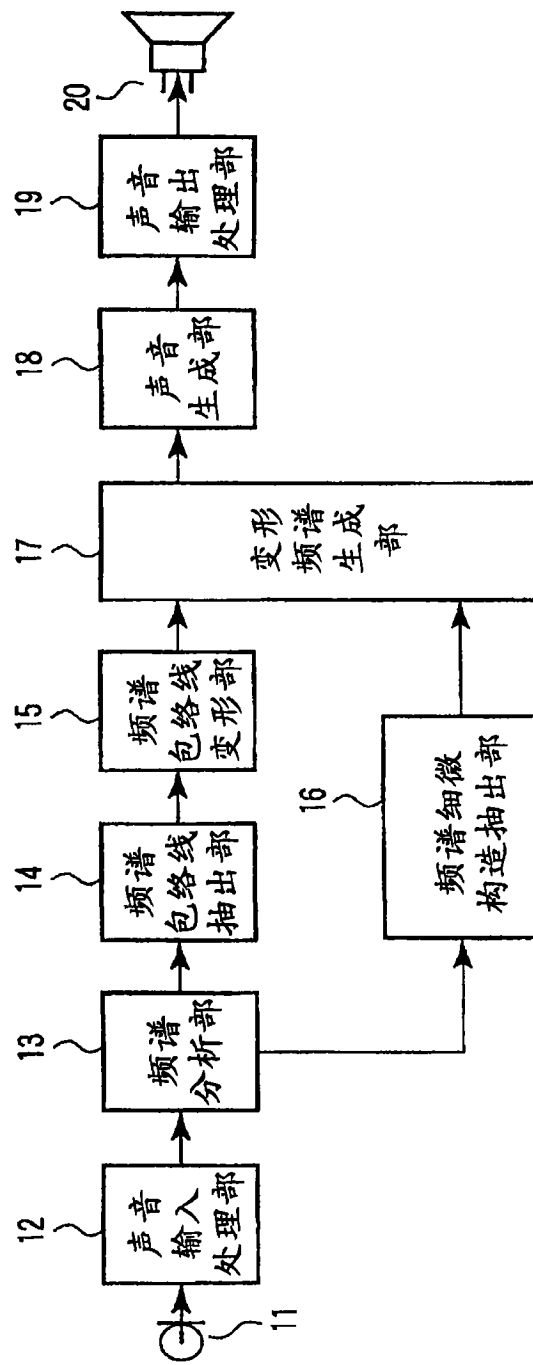


图3

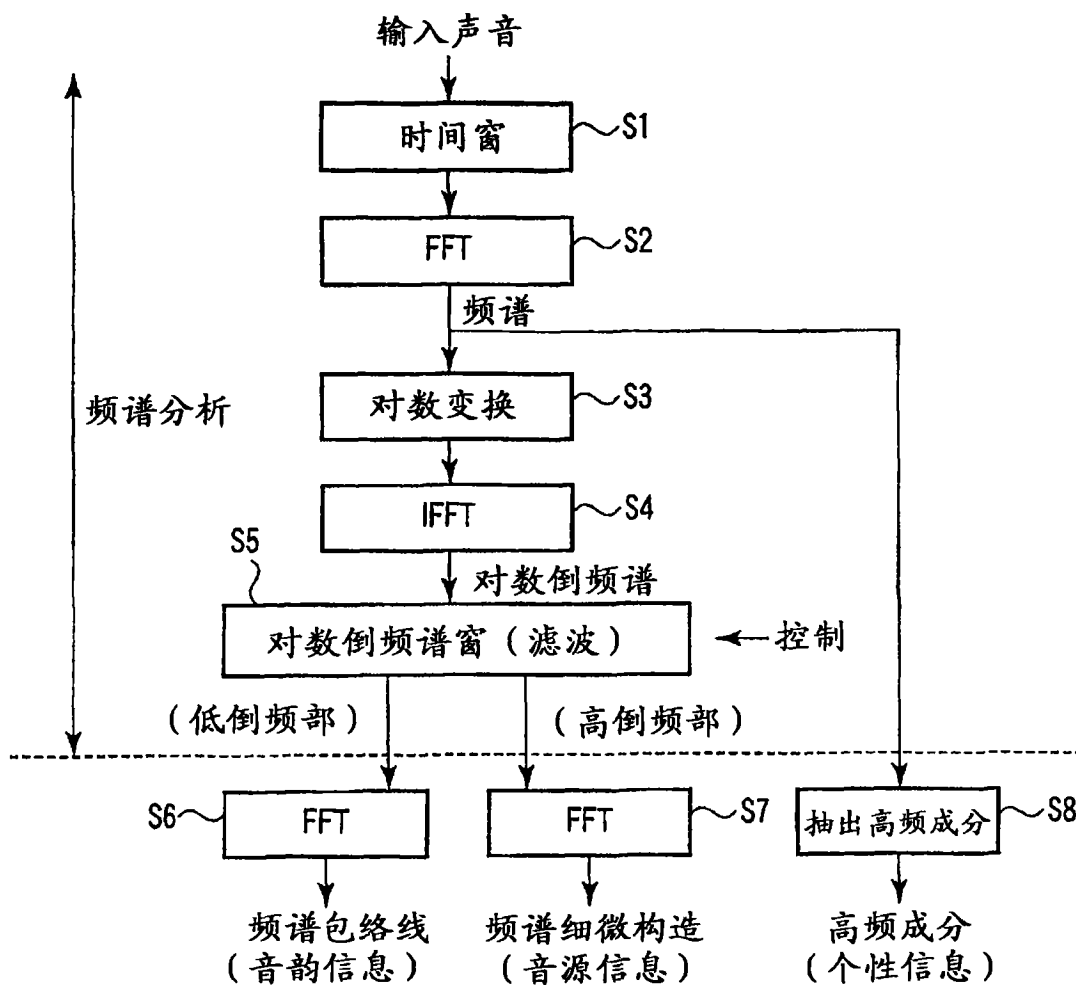


图 4

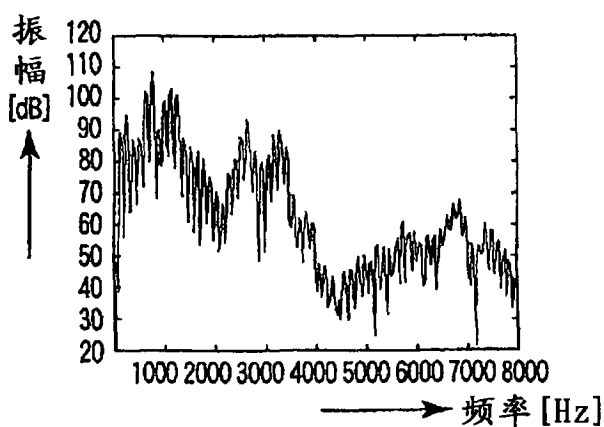
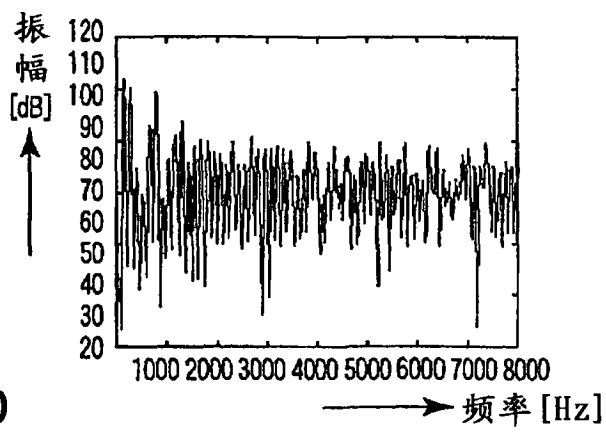
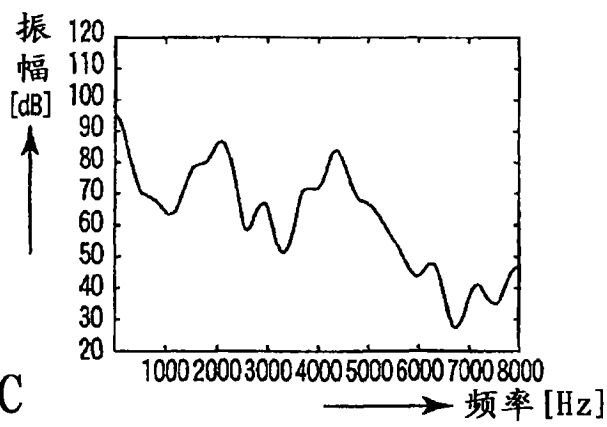
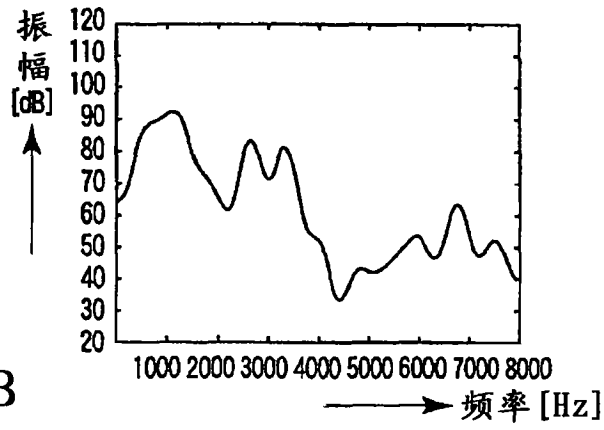


图 5A



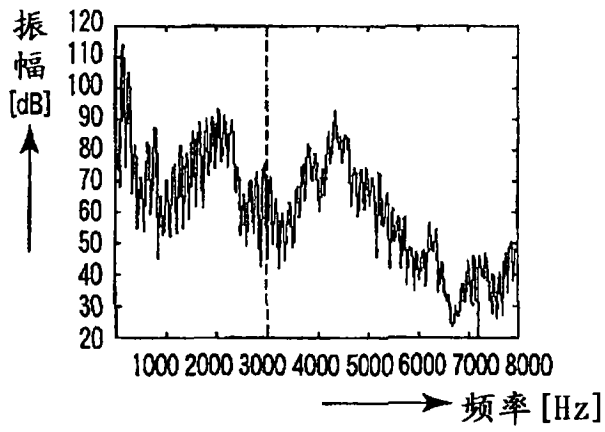


图 5E

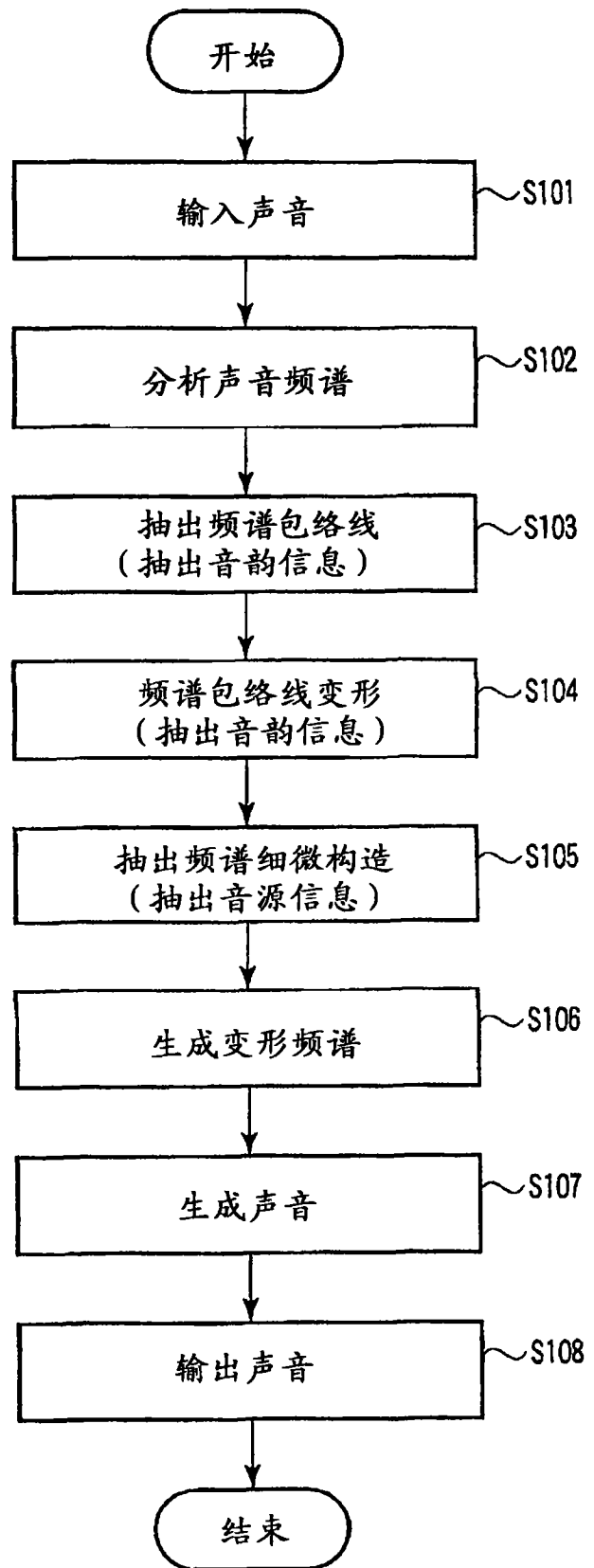


图 6

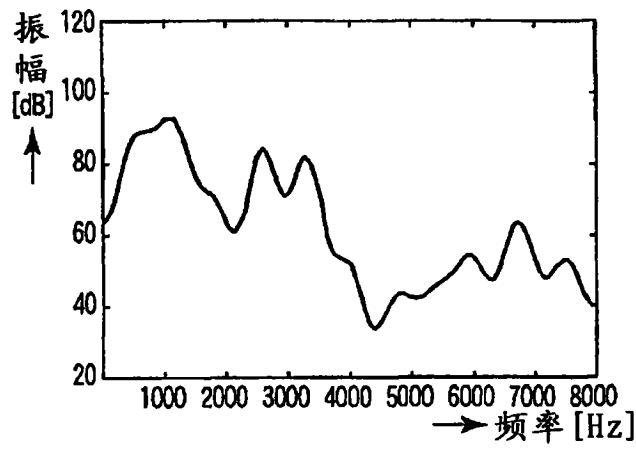


图 7A

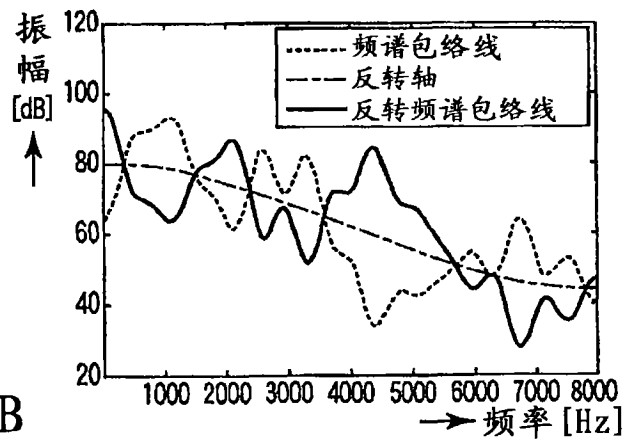


图 7B

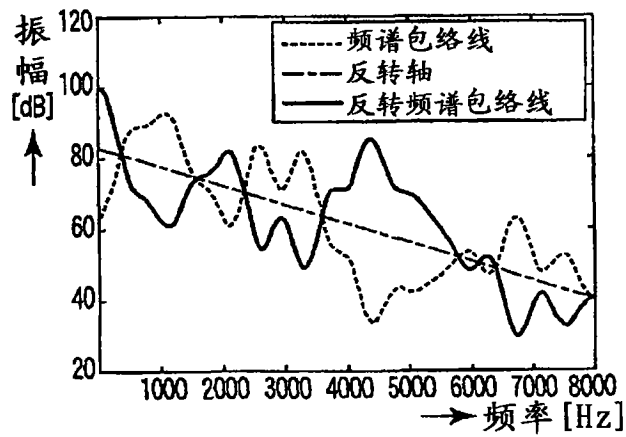
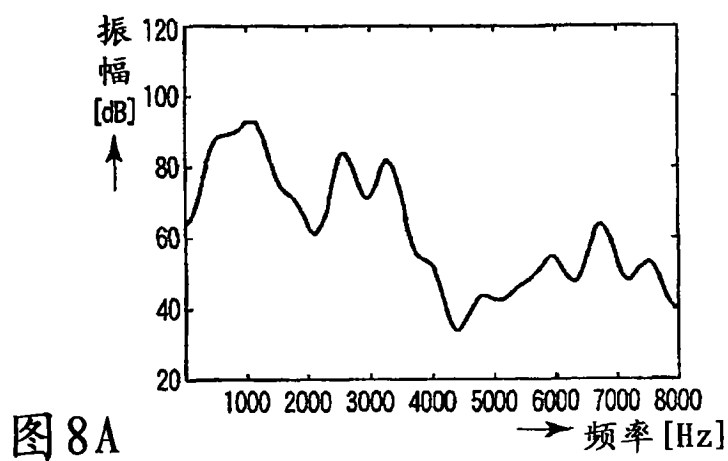
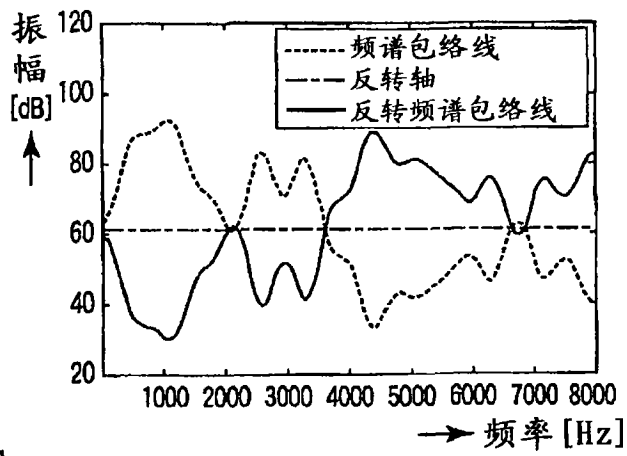
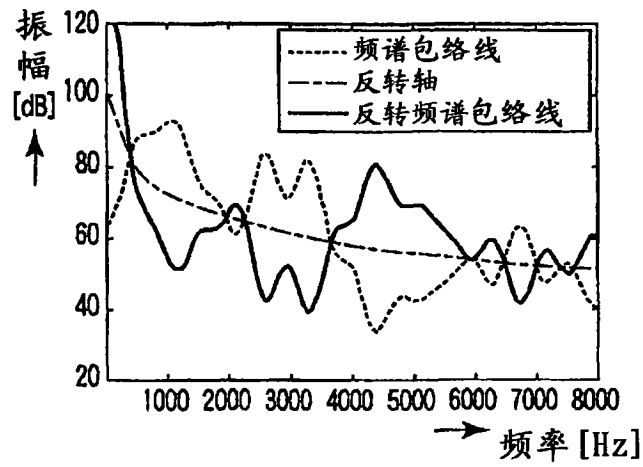
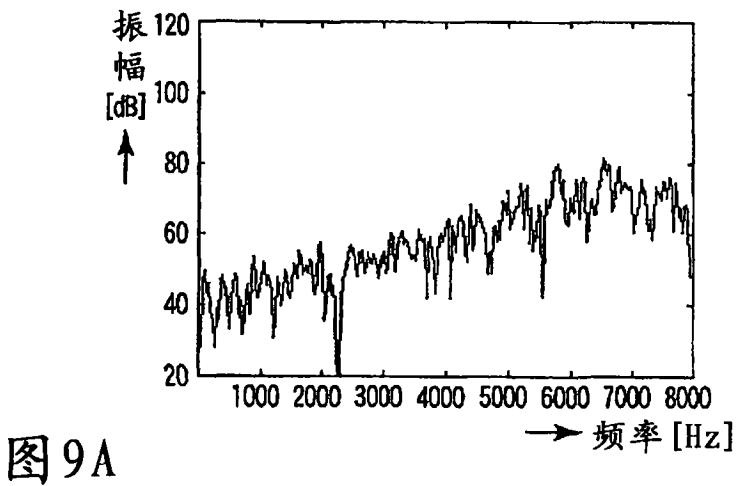
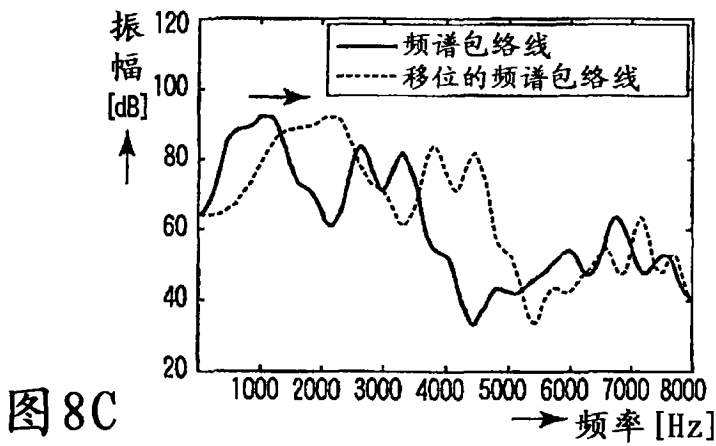
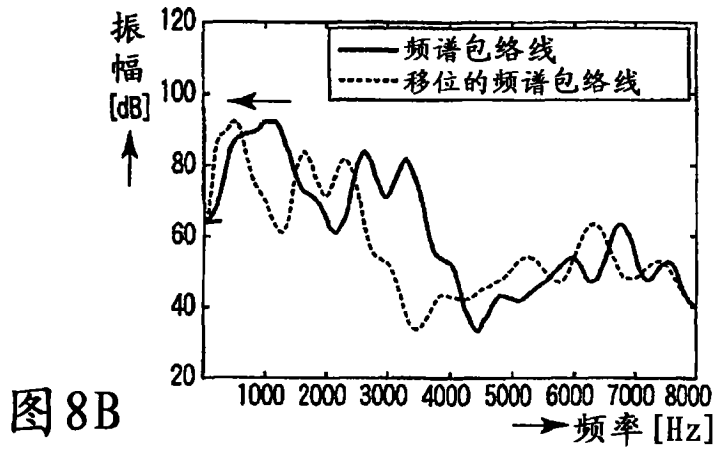
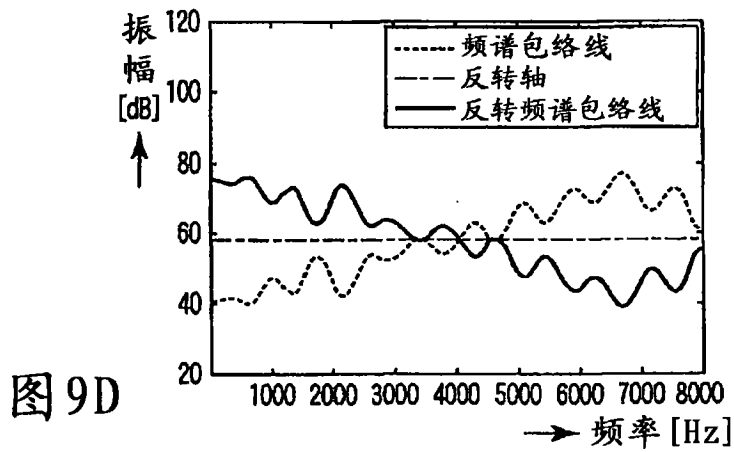
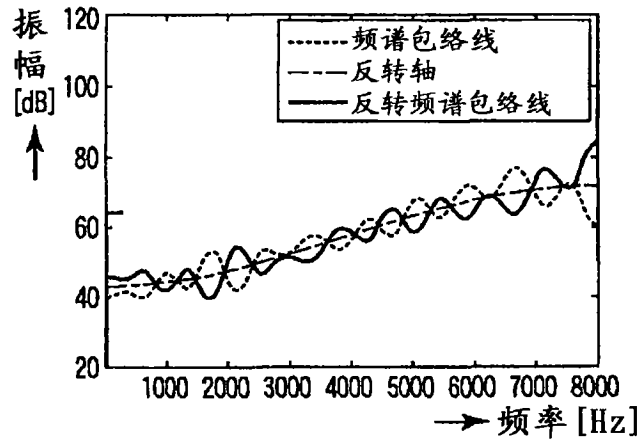
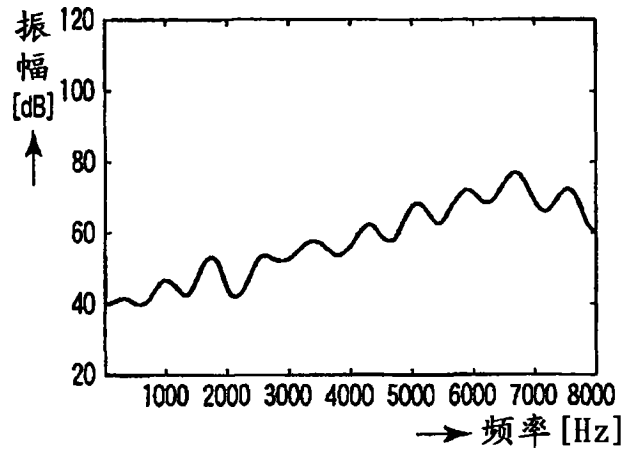


图 7C







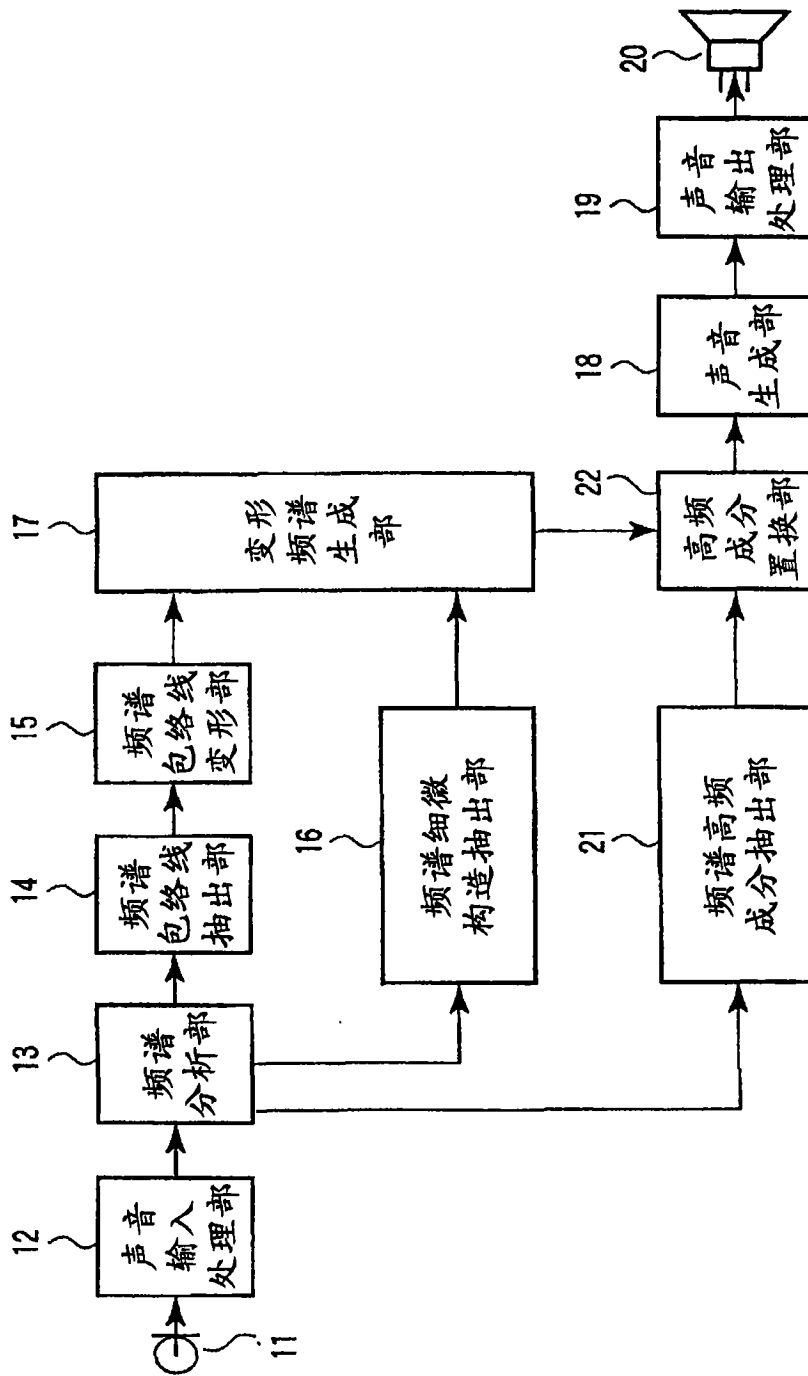


图10

图 10

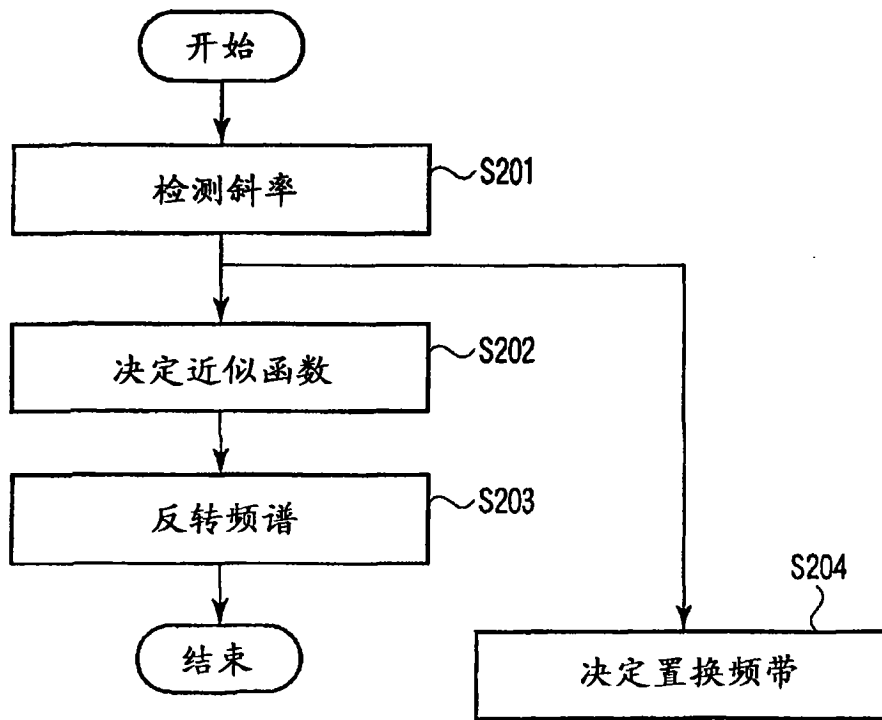


图 11

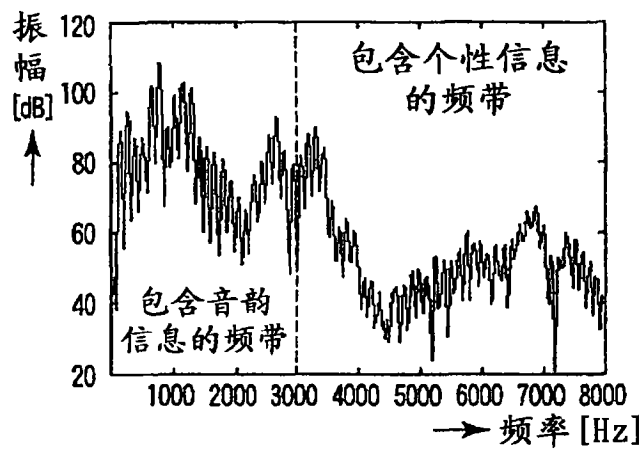
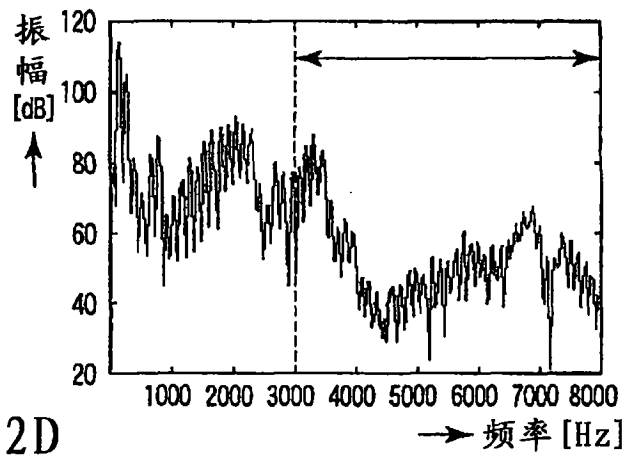
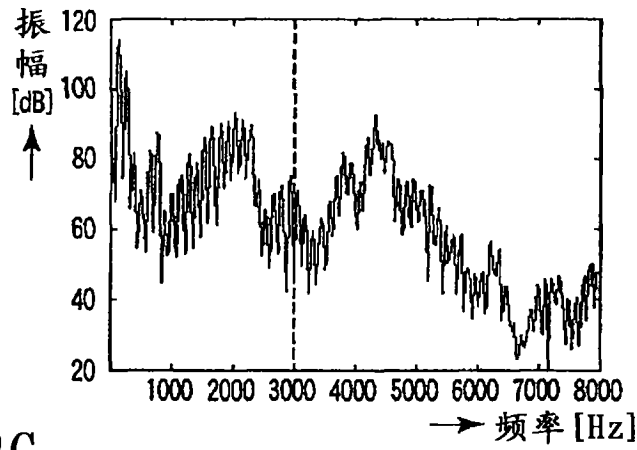
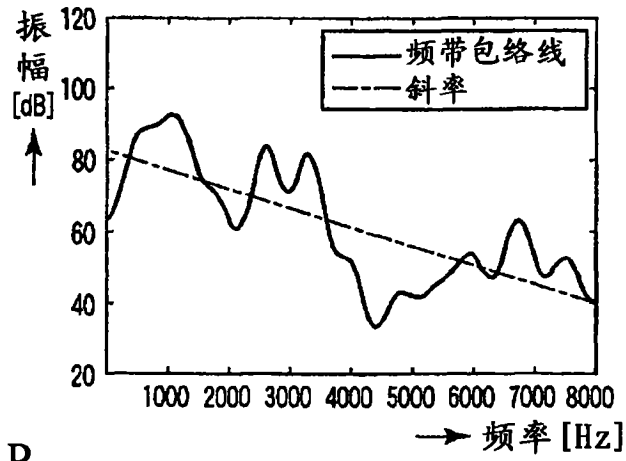


图 12A



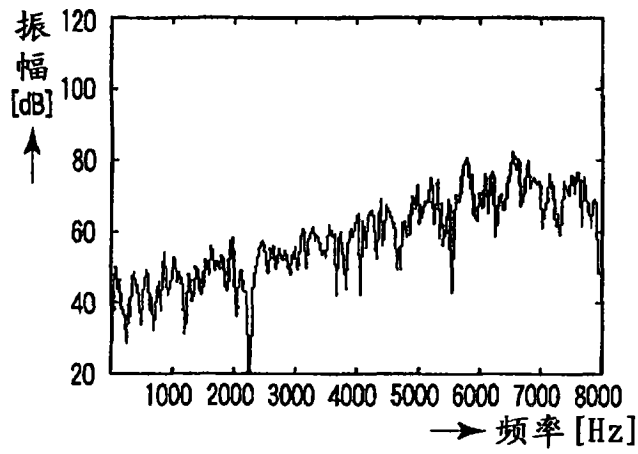


图13A

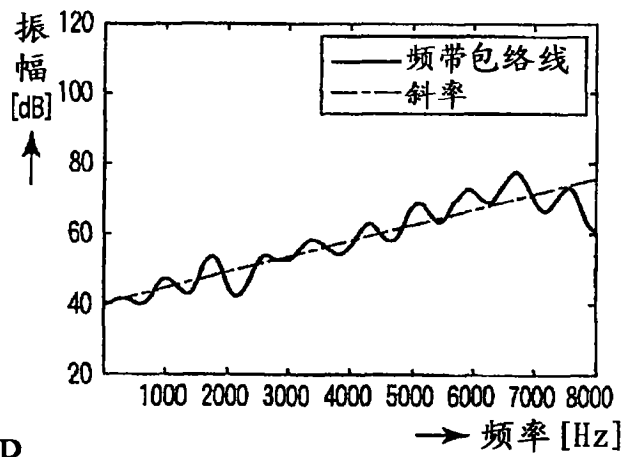


图13B

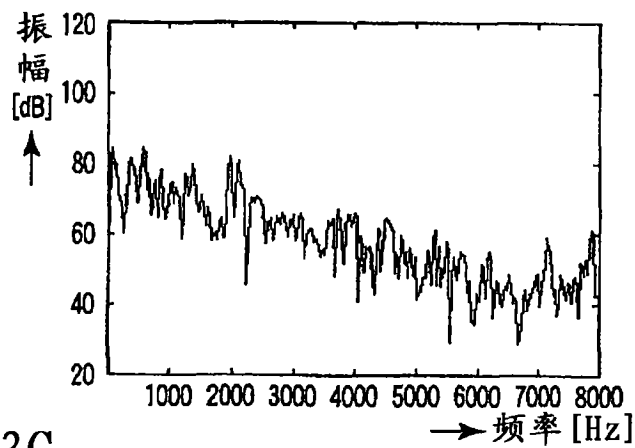


图13C

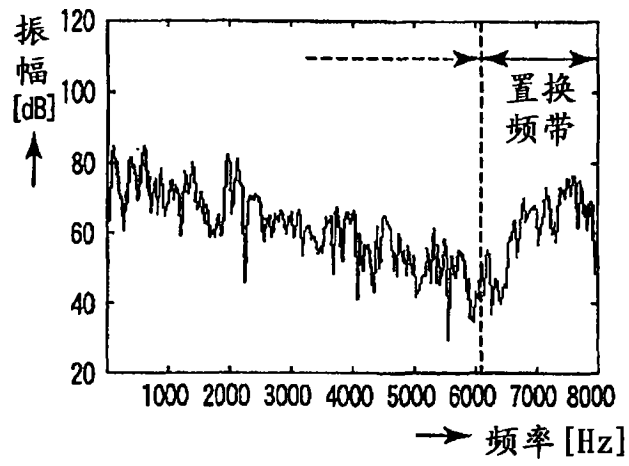


图 13D

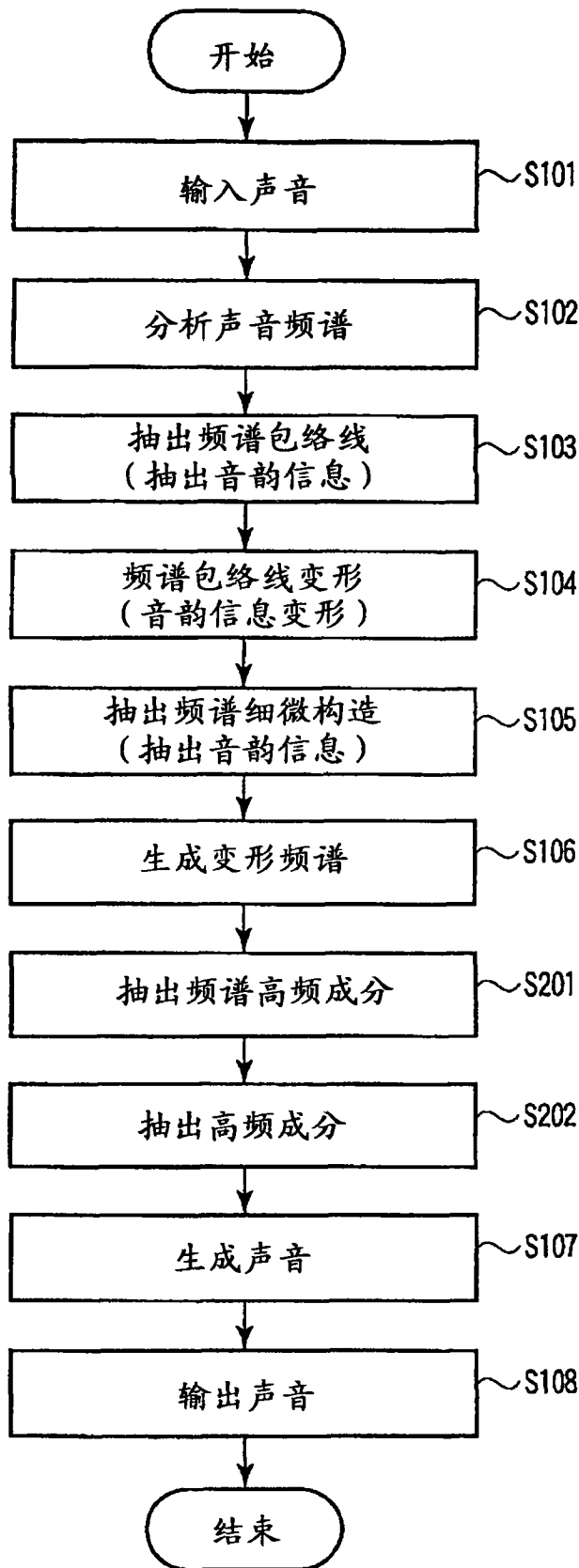


图 14