



(12) 发明专利

(10) 授权公告号 CN 101393745 B

(45) 授权公告日 2012. 03. 14

(21) 申请号 200810174735. 0

US 5781696 A, 1998. 07. 14, 全文.

(22) 申请日 2008. 09. 19

CN 1186303 A, 1998. 07. 01, 全文.

(30) 优先权数据

审查员 辛杰

2007-241681 2007. 09. 19 JP

(73) 专利权人 索尼株式会社

地址 日本东京都

(72) 发明人 中村理 安部素嗣

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

代理人 宋鹤 南霆

(51) Int. Cl.

G10L 21/04 (2006. 01)

G11B 20/00 (2006. 01)

(56) 对比文件

WO 02/09090 A2, 2002. 01. 31, 全文.

CN 2162697 Y, 1994. 04. 20, 全文.

US 5717829 A, 1998. 02. 10, 全文.

CN 1189046 A, 1998. 07. 29, 全文.

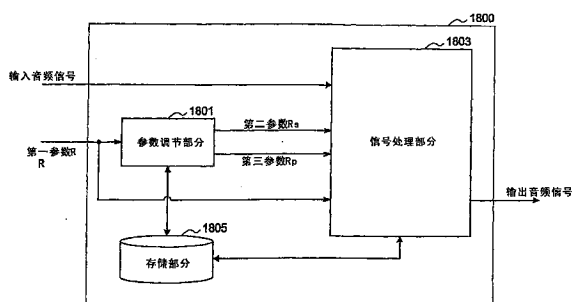
权利要求书 3 页 说明书 36 页 附图 47 页

(54) 发明名称

音频信号处理装置和方法

(57) 摘要

本发明提供了一种信息处理装置、信息处理方法和程序。根据本发明, 提供了参数调节部分, 其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数, 以及信号处理部分, 其基于第二参数和第三参数来调节音频信号的回放速度和声音音高中的至少一个, 其中信号处理部分在所输入的回放速度变量因子小于预定阈值时调节音频信号的回放速度, 并且在所输入的回放速度变量因子大于预定阈值时调节音频信号的回放速度和声音音高。



1. 一种音频信号处理装置,包括:

参数调节部分,其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数;以及

信号处理部分,其基于所述第二参数和所述第三参数,调节音频信号的回放速度和声音音高中的至少一个,其中

所述信号处理部分在所输入的回放速度变量因子小于预定阈值时调节所述音频信号的所述回放速度,并且在所输入的回放速度变量因子在所述预定阈值以上时调节所述音频信号的所述回放速度和所述声音音高,其中

所述信号处理部分包括:

回放速度转换部分,其转换所述音频信号的所述回放速度;以及

音高调节部分,其调节所述音频信号的所述声音音高,其中

所述回放速度转换部分基于所述第二参数来转换所述音频信号的所述回放速度,并且所述音高调节部分基于所述第三参数来调节所述音频信号的所述声音音高。

2. 如权利要求1所述的音频信号处理装置,其中

所述第一参数约等于所述第二参数和所述第三参数的乘积。

3. 如权利要求1所述的音频信号处理装置,其中

所述信号处理部分还包括:

音频信号输出控制部分,其控制将从所述信号处理部分输出的、已被执行了预定信号处理的音频信号的输出,其中

当回放速度和声音音高均经过调节的音频信号被从所述信号处理部分输出时,所述音频信号输出控制部分降低所述回放速度和声音音高均经过调节的音频信号的音频音量。

4. 如权利要求3所述的音频信号处理装置,其中

所述信号处理部分还包括:

拟声声音切换判断部分,其根据所述第一参数,判断是调节所述音频信号的所述回放速度和所述声音音高中的至少一个,还是将所述音频信号切换到指示正在执行高速回放的预定拟声声音,其中

当所述第一参数在所述预定阈值以上时,所述拟声声音切换判断部分判断将所述音频信号切换到所述预定拟声声音;并且

当所述拟声声音切换判断部分判断将所述音频信号切换到所述预定拟声声音时,所述音频信号输出控制部分在将所述音频信号切换到所述预定拟声声音之后输出所述音频信号。

5. 如权利要求1所述的音频信号处理装置,还包括:

内容管理部分,其管理包括所述音频信号的内容,其中

所述参数调节部分根据所输入的所述第一参数来确定第四参数,所述第四参数对将从所述内容管理部分输出到所述信号处理部分的所述音频信号的数据量进行调节。

6. 如权利要求5所述的音频信号处理装置,其中

当所述第一参数在预定阈值以上时,所述参数调节部分减小所述第四参数以减小将从所述内容管理部分输出到所述信号处理部分的所述内容的数据量。

7. 如权利要求5所述的音频信号处理装置,其中

所述第一参数和所述第四参数的乘积约等于所述第二参数和所述第三参数的乘积。

8. 如权利要求 1 所述的音频信号处理装置,还包括:

内容管理部分,其管理包括所述音频信号的内容,其中

所述参数调节部分基于对将从所述内容管理部分输出到所述信号处理部分的音频数据的数据量进行调节的第四参数和所输入的所述第一参数,来确定所述第二参数和所述第三参数。

9. 如权利要求 8 所述的音频信号处理装置,其中

当所述第一参数在预定阈值以上时,所述内容管理部分减小所述第四参数以减小将从所述内容管理部分输出到所述信号处理部分的所述内容的数据量。

10. 如权利要求 8 所述的音频信号处理装置,其中

所述第一参数和所述第四参数的乘积约等于所述第二参数和所述第三参数的乘积。

11. 如权利要求 1 所述的音频信号处理装置,还包括:

存储数据库的存储部分,在所述数据库中,所输入的所述第一参数与所述第二参数和所述第三参数相互相关,其中

所述参数调节部分通过参考存储在所述存储部分中的所述数据库来确定所述第二参数和所述第三参数。

12. 如权利要求 11 所述的音频信号处理装置,其中

当所述第一参数在预定阈值以上时,所述参数调节部分根据所述第一参数与所述预定阈值之差来增大所述第二参数。

13. 如权利要求 11 所述的音频信号处理装置,其中

所述数据库被存储为指示所述第二参数和所述第三参数随所述第一参数的变化的曲线,并且

指示所述第三参数的变化的所述曲线在所述预定阈值之前和之后具有平滑的形状。

14. 如权利要求 5 所述的音频信号处理装置,还包括:

存储数据库的存储部分,在所述数据库中,所输入的所述第一参数与所述第二参数、所述第三参数和所述第四参数相互相关,其中

所述参数调节部分通过参考存储在所述存储部分中的所述数据库来确定所述第二参数、所述第三参数和所述第四参数。

15. 如权利要求 1 所述的音频信号处理装置,其中

当所述第一参数在预定阈值以上时,所述参数调节部分根据所述第一参数与所述预定阈值之差来增大所述第二参数。

16. 一种音频信号处理方法,包括:

参数调节步骤,其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数;以及

信号处理步骤,其基于所述第二参数和所述第三参数,调节音频信号的回放速度和声音高中至少一个,其中

所述信号处理步骤在所输入的回放速度变量因子小于预定阈值时,基于所述第二参数来调节所述音频信号的所述回放速度,并且在所输入的回放速度变量因子在所述预定阈值以上时调节所述音频信号的所述回放速度和所述声音音高,其中

所述信号处理步骤包括：

转换所述音频信号的所述回放速度；以及

调节所述音频信号的所述声音音高，其中

基于所述第二参数来转换所述音频信号的所述回放速度，并且

基于所述第三参数来调节所述音频信号的所述声音音高。

17. 如权利要求 16 所述的音频信号处理方法，其中

在所述参数调节步骤中，确定所述第二参数和所述第三参数以使得所述第一参数可以约等于所述第二参数和所述第三参数的乘积。

18. 如权利要求 16 所述的音频信号处理方法，其中

在所述信号处理步骤中，对所述音频信号的信号波形的振幅进行控制以使得当所述音频信号的所述回放速度和所述声音音高均被调节时，降低所述音频信号的音频音量。

19. 如权利要求 16 所述的音频信号处理方法，其中

在所述信号处理步骤中，当所述第一参数在所述预定阈值以上时，所述音频信号被切换到指示正在执行高速回放的预定拟声声音。

20. 如权利要求 16 所述的音频信号处理方法，其中

在所述参数调节步骤中，还根据所述第一参数确定了第四参数，所述第四参数对在所述信号处理步骤中被处理的所述音频信号的数据量进行调节。

21. 如权利要求 20 所述的音频信号处理方法，其中

在所述参数调节步骤中，确定所述第二参数、所述第三参数和所述第四参数以使得所述第一参数和所述第四参数的乘积可以约等于所述第二参数和所述第三参数的乘积。

22. 如权利要求 20 所述的音频信号处理方法，其中

在所述参数调节步骤中，当所述第一参数在预定阈值以上时，所述第四参数被减小以减小所述音频信号的数据量。

23. 如权利要求 16 所述的音频信号处理方法，其中

在所述参数调节步骤中，根据对将在所述信号处理步骤中被处理的所述音频信号的数据量进行调节的第四参数和所述第一参数来确定所述第二参数和所述第三参数。

24. 如权利要求 23 所述的音频信号处理方法，其中

在所述参数调节步骤中，确定所述第二参数和所述第三参数以使得所述第一参数和所述第四参数的乘积可以约等于所述第二参数和所述第三参数的乘积。

音频信号处理装置和方法

技术领域

[0001] 本发明涉及信息处理装置、信息处理方法和程序。

背景技术

[0002] 近年来,将作为数字数据的由 TV 广播所广播的节目记录到诸如 DVD(数字多功能光盘)或 HDD(硬盘驱动器)之类的具有随机存取性能的记录介质中的视频—记录/回放装置迅速普及。此外,通过因特网对诸如视频和音频之类的内容进行发行变得流行,并且具有内置 HDD 或闪存的回放装置已经广泛普及,通过该回放装置能够在室内或室外欣赏从因特网下载得到的内容。

[0003] 利用数字和随机存取特性,将如上所述的用于数字内容的回放装置实现为具有各种功能。变速回放功能可以作为示例,其在维持声音的恒定音高(pitch)的同时可变地设定回放速度。变速回放功能是减缓或加快视频和音频的回放速度的功能,并且例如,该功能针对语言的初学者等将回放速度减缓大约 20%(慢回放),或者将回放速度加快大约 50%以节省观看时间等(快回放)。变速回放功能是这样的功能,其从数字内容回放装置开始普及起就被一般地实现在该装置中,并且如今,数字内容回放装置已经十分普遍。本发明不仅关注音频内容,而且关注视频内容的音频部分。

[0004] 在数字内容的回放装置中,在维持声音的恒定音高的同时可变地设定回放速度的技术被称作语速转换。在下文中,语速转换意味着在维持声音的恒定音高的同时拉伸或压缩信号的转换。已知几种用于语速转换的方法,例如,用作在对应于数字音频信号的时域上的时间轴拉伸/压缩算法的 PICOLA(指针间隔控制交叠和相加)(参见“Expansion/compression on the audio time-axis using duplication adding method by pointer amount-of-movement control(PICOLA)and its evaluation”, by Morita and Itakura, Acoustic Society of Japan collected papers, October 1986, pp. 149-150)。该算法的优点在于虽然其处理简单轻量,但是可以获得好的声音质量。

发明内容

[0005] 然而,经过语速转换,在维持声音的恒定音高的同时执行了回放速度转换,已经很难在听觉上识别转换后的回放速度。

[0006] 因此,考虑到上述问题而提出本发明,并且期望提供在转换音频信号的回放速度时使得能够在听觉上识别转换后的回放速度的新的并改良的信息处理装置、新的并改良的信息处理方法以及新的并改良的程序。

[0007] 根据本发明的实施例,提供了一种信息处理装置,该信息处理装置包括参数调节部分,其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数;以及信号处理部分,其基于第二参数和第三参数来调节音频信号的回放速度和声音音高中的至少一个,其中信号处理部分在所输入的回放速度变量因子小于预定阈值时调节音频信号的回放速度,并且在所输入的回放速度变量因子在预定阈值以上时调节音频信号的回放速

度和声音音高。

[0008] 通过这样的配置,参数调节部分根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数,并且信号处理部分基于第二参数和第三参数来调节音频信号的回放速度和声音音高的至少一个。在此,信号处理部分在所输入的回放速度变量因子小于预定阈值时调节音频信号的回放速度,并且在所输入的回放速度变量因子在预定阈值以上时调节音频信号的回放速度和声音音高。因此,通过根据本发明的信息处理装置,在音频信号的回放速度被转换的情况下,可以在听觉上识别转换后的回放速度。

[0009] 信号处理部分包括回放速度转换部分,其转换音频信号的回放速度;以及音高调节部分,其调节音频信号的声音音高,并且回放速度转换部分可以基于第二参数来转换音频信号的回放速度,而音高调节部分可以基于第三参数来调节音频信号的声音音高。

[0010] 第一参数可以近似等于第二参数和第三参数的乘积。

[0011] 信号处理部分还包括音频信号输出控制部分,其控制将从信号处理部分中被输出的、被执行了预定信号处理的音频信号的输出,并且当回放速度和声音音高均经过调节的音频信号从信号处理部分中被输出时,音频信号输出控制部分可以降低回放速度和声音音高均经过调节的音频信号的音频音量。

[0012] 信号处理部分还包括拟声声音 (onomatopoeic sound) 切换判断部分,其根据第一参数,判断是调节音频信号的回放速度和声音音高的至少一个,还是将音频信号切换到指示正在执行高速回放的预定拟声声音,并且当第一参数在预定阈值以上时,拟声声音切换判断部分可以判断将音频信号切换到预定拟声声音;并且当拟声声音切换判断部分判断将音频信号切换到预定拟声声音时,音频信号输出控制部分可以在将音频信号切换到预定拟声声音后输出音频信号。

[0013] 信息处理装置还包括内容管理部分,该内容管理部分管理包括音频信号的内容,并且参数调节部分可以根据待输入的第一参数来确定第四参数,该第四参数对将从内容管理部分被输出到信号处理部分的音频信号的数据量进行调节。

[0014] 当第一参数在预定阈值以上时,参数调节部分可以减小第四参数以减小将从内容管理部分被输出到信号处理部分的内容的数据量。

[0015] 第一参数和第四参数的乘积可以近似等于第二参数和第三参数的乘积。

[0016] 信息处理装置还包括内容管理部分,该内容管理部分管理包括音频信号的内容,并且参数调节部分可以基于对将从内容管理部分被输出到信号处理部分的音频数据的数据量进行调节的第四参数和待输入的第一参数,确定第二参数和第三参数。

[0017] 当第一参数在预定阈值以上时,内容管理部分可以减小第四参数以减小将从内容管理部分被输出到信号处理部分的内容的数据量。

[0018] 信息处理装置还包括存储部分,该存储部分存储其中待输入的第一参数与第二参数和第三参数相互相关的数据库,并且参数调节部分可以通过参考在存储部分中所存储的数据库来确定第二参数和第三参数。

[0019] 信息处理装置还包括存储部分,该存储部分存储其中待输入的第一参数与第二参数、第三参数和第四参数相互相关的数据库,并且参数调节部分可以通过参考在存储部分中所存储的数据库来确定第二参数、第三参数和第四参数。

[0020] 当第一参数在预定阈值以上时,参数调节部分可以根据第一参数与预定阈值之间

的差来增大第二参数。

[0021] 数据库被存储为指示第二参数和第三参数随第一参数的变化的曲线,并且指示第三参数的变化的曲线可以在预定阈值前后具有平滑的形状。

[0022] 根据本发明的另一实施例,提供了一种信息处理方法,该信息处理方法包括:参数调节步骤,其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数;以及信号处理步骤,其基于第二参数和第三参数,调节音频信号的回放速度和声音音高的至少一个,其中信号处理步骤在所输入的回放速度变量因子小于预定阈值时,基于第二参数来调节音频信号的回放速度,并且在所输入的回放速度变量因子在预定阈值以上时,基于第二参数和第三参数来调节音频信号的回放速度和声音音高。

[0023] 通过这样的配置,参数调节步骤根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数,并且信号处理步骤基于第二参数和第三参数来调节音频信号的回放速度和声音音高的至少一个。此时,信号处理步骤在所输入的回放速度变量因子小于预定阈值时,基于第二参数来调节音频信号的回放速度,并且在所输入的回放速度变量因子在预定阈值以上时,基于第二参数和第三参数来调节音频信号的回放速度和声音音高。因此,通过根据本发明的信息处理装置,在音频信号的回放速度被转换的情况下,可以在听觉上识别转换后的回放速度。

[0024] 在参数调节步骤中,可以确定第二参数和第三参数以使得第一参数可以近似等于第二参数和第三参数的乘积。

[0025] 在信号处理步骤中,可以对音频信号的信号波形振幅进行控制以使得当音频信号的回放速度和声音音高均经过调节时,音频信号的音频音量可以较小。

[0026] 在信号处理步骤中,当第一参数在预定阈值以上时,音频信号可以被切换到指示正在执行高速回放的预定拟声声音。

[0027] 在参数调节步骤中,还可以根据第一参数来确定第四参数,该第四参数对将在信号处理步骤中被处理的音频信号的数据量进行调节。

[0028] 在参数调节步骤中,当第一参数在预定阈值以上时,第四参数可以被减小以减小音频信号的数据量。

[0029] 在参数调节步骤中,可以根据对将在信号处理步骤中被处理的音频信号的数据量进行调节的第四参数和第一参数来确定第二参数和第三参数。

[0030] 在参数调节步骤中,可以确定第二参数、第三参数和第四参数以使得第一参数和第四参数的乘积可以近似等于第二参数和第三参数的乘积。

[0031] 根据本发明的另一实施例,提供了一种在计算机中实现的程序:参数调节功能,其根据指示所输入的回放速度变量因子的第一参数来设定第二参数和第三参数;以及信号处理功能,其基于第二参数和第三参数来调节音频信号的回放速度和声音音高的至少一个。

[0032] 通过这样的配置,计算机程序被存储在计算机中所包括的存储部分中,并且被计算机中所包括的CPU读取以执行,因此,该程序使计算机用作上述的信息处理装置。此外,还可以提供一种其中记录了计算机程序并且可以被计算机读取的记录介质。记录介质例如是磁盘、光盘、磁光盘和闪存。此外,例如可以经由网络来发布上述的计算机程序而无需使用记录介质。

[0033] 根据上述本发明的实施例,在音频信号的回放速度被转换的情况下,可以在听觉

上识别转换后的回放速度。

附图说明

- [0034] 图 1A 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0035] 图 1B 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0036] 图 1C 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0037] 图 1D 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0038] 图 2A 是示出搜索类似波形长度的示例的说明性示图。
- [0039] 图 2B 是示出搜索类似波形长度的示例的说明性示图。
- [0040] 图 2C 是示出搜索类似波形长度的示例的说明性示图。
- [0041] 图 3A 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0042] 图 3B 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。
- [0043] 图 4A 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0044] 图 4B 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0045] 图 4C 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0046] 图 4D 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0047] 图 5A 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0048] 图 5B 是示出通过 PICOLA 来压缩音频信号的方法的说明性示图。
- [0049] 图 6 是示出通过 PICOLA 来拉伸音频信号的方法的流程图。
- [0050] 图 7 是示出通过 PICOLA 来压缩音频信号的方法的流程图。
- [0051] 图 8 是示出根据 PICOLA 的语速转换装置的配置的框图。
- [0052] 图 9 是示出检测类似波形长度的处理的流程图。
- [0053] 图 10 是示出检测类似波形长度的处理的流程图。
- [0054] 图 11 是示出生成交叉淡入淡出信号的处理示例的流程图。
- [0055] 图 12 是示出降低采样率的方法的说明性示图。
- [0056] 图 13 是示出提高采样率的方法的说明性示图。
- [0057] 图 14A 是示出与回放速度成比例地升高声音音高的处理示例的说明性示图。
- [0058] 图 14B 是示出与回放速度成比例地升高声音音高的处理示例的说明性示图。
- [0059] 图 14C 是示出与回放速度成比例地升高声音音高的处理示例的说明性示图。
- [0060] 图 15A 是示出在相关技术的第一回放装置中回放速度变量因子与语速转换率之间关系的坐标图。
- [0061] 图 15B 是示出在相关技术的第一回放装置中回放速度变量因子与声音音高之间关系的坐标图。
- [0062] 图 16A 是示出在相关技术的第二回放装置中回放速度变量因子与语速转换率之间关系的坐标图。
- [0063] 图 16B 是示出在相关技术的第二回放装置中回放速度变量因子与声音音高之间关系的坐标图。
- [0064] 图 17 是示出包括根据本发明第一实施例的信息处理装置的回放速度转换系统的说明性示图。

- [0065] 图 18 是示出根据本实施例的信息处理装置配置的框图。
- [0066] 图 19A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0067] 图 19B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0068] 图 20 是示出根据本实施例的信息处理装置的处理流程的流程图。
- [0069] 图 21 是示出根据本实施例的信息处理部分的功能的框图。
- [0070] 图 22A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0071] 图 22B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0072] 图 23 是示出根据本实施例的信号处理方法的流程图。
- [0073] 图 24A 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理示例的说明性示图。
- [0074] 图 24B 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理示例的说明性示图。
- [0075] 图 24C 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理示例的说明性示图。
- [0076] 图 24D 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理示例的说明性示图。
- [0077] 图 25A 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理的另一示例的说明性示图。
- [0078] 图 25B 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理的另一示例的说明性示图。
- [0079] 图 25C 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理的另一示例的说明性示图。
- [0080] 图 25D 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理的另一示例的说明性示图。
- [0081] 图 26A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0082] 图 26B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0083] 图 27A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0084] 图 27B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0085] 图 28A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0086] 图 28B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0087] 图 29 是示出根据本实施例的信号处理部分的修改示例的框图。
- [0088] 图 30 是示出根据修改示例的信号处理方法的流程图。
- [0089] 图 31 是示出转换采样率的另一种方法的说明性示图。
- [0090] 图 32 是示意性示出回放速度变量因子随时间的变化的说明性示图。
- [0091] 图 33 是示出根据本发明第二实施例的信息处理装置的功能的框图。
- [0092] 图 34A 是示出第一参数 R 和第四参数 Rt 之间关系的坐标图。
- [0093] 图 34B 是示出第一参数 R 和将被输入到信号处理部分的音频信号数据量之间关系的坐标图。
- [0094] 图 35A 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。

- [0095] 图 35B 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0096] 图 36A 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0097] 图 36B 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0098] 图 37A 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0099] 图 37B 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0100] 图 37C 是示出根据本实施例的调节数据读取速度的方法示例的说明性示图。
- [0101] 图 38A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0102] 图 38B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0103] 图 39 是示出根据本实施例的信息处理装置的处理流程的流程图。
- [0104] 图 40 是示出根据本实施例的信号处理部分的功能的框图。
- [0105] 图 41A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图。
- [0106] 图 41B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。
- [0107] 图 42 是示出根据本实施例的信号处理方法的流程图。
- [0108] 图 43 是示出根据本实施例的信息处理装置的第一修改示例的功能的框图。
- [0109] 图 44 是示出根据本修改示例的信号处理方法的流程图。
- [0110] 图 45 是示出根据本实施例以及本修改示例的信号处理部分的修改示例的框图。
- [0111] 图 46 是示出根据本修改示例的信号处理方法的流程图。
- [0112] 图 47 是示出根据本发明每个实施例的信息处理装置的硬件配置的框图。

具体实施方式

[0113] 在下文中,将参考附图详细描述本发明的优选实施例。应注意,在本说明书和附图中,用相同标号来表示实质上具有相同功能和结构的结构元件,并且省略对这些结构元件的重复说明。

[0114] 顺便提及,以下,由语音构成的信号称作语音信号,由诸如音乐之类的除语音外的声音构成的信号称作声响信号,并且由语音信号和声响信号构成的信号称作音频信号。

[0115] (基础技术描述)

[0116] 首先,在给出对本发明的优选实施例的详细描述之前,将描述实现本发明所基于的技术问题。顺便提及,本发明被配置为能够通过改进如下所述的基础技术来获得显著效果。因此,与改进有关的技术是本实施例的特征。也就是,虽然本实施例遵循如下所述的技术问题的基本概念,但是实施例的本质关注改进,并且应注意,所述配置清楚地区别于基础技术,并且在本实施例的效果和基础技术的效果之间存在清楚的差别。

[0117] (PICOLA 描述)

[0118] 如上所述, PICOLA 是在对应于数字语音信号的时域上的时间轴拉伸 / 压缩算法,并且如下所述,对语音信号执行拉伸和压缩。以下,通过参考图 1A 到图 5B,将描述根据 PICOLA 的信号处理方法。

[0119] 图 1A 到 1D 是示出通过 PICOLA 来拉伸音频信号的方法的说明性视图。顺便提及,在以下描述中,原始波形是信号最初被输入到 PICOLA 时的波形。此外,在图 1A 到 1D 中,纵轴代表信号振幅(即强度),并且横轴代表时间。

[0120] (根据 PICOLA 对波形进行拉伸的处理)

[0121] 根据 PICOLA, 首先, 从原始波形中检测到具有类似波形的周期 A 和周期 B。如图 1A 所示, 周期 A 和周期 B 是连续并且具有相同长度的两个周期, 并且周期 A 的样本数和周期 B 的样本数相同。随后, 生成在图 1B 中示出的波形, 图 1B 的波形在所检测到的周期 A 中保持不变, 然后在所检测到的周期 B 中淡出 (fade out)。类似地, 生成在图 1C 中示出的波形, 图 1C 的波形从周期 A 开始淡入 (fade in) 并且在周期 B 中保持不变。然后, 通过相加在图 1B 和图 1C 中示出的所生成波形, 可以获得在图 1D 中示出的经拉伸的波形。

[0122] 如上所述的相加淡出波形和淡入波形被称作交叉淡入淡出 (cross-fade)。当周期 A 和周期 B 的交叉淡入淡出周期被表示为周期 $A \times B$ 并且上述操作被执行时, 在图 1A 中示出的原始波形的周期 A 和周期 B 变为在图 1D 中示出的经拉伸波形的周期 A、周期 $A \times B$ 和周期 B。

[0123] (类似波形长度检测)

[0124] 在此, 在如上所述的拉伸波长的处理中, 将要从所输入的信号中检测出连续并且具有类似波形的两个周期。在下文中, 通过参考图 2A 到图 2C, 来描述检测具有类似波形的周期 A 和周期 B 的周期长度 W 的方法。图 2A 到图 2C 是示出搜索类似波形长度的示例的说明性示图。顺便提及, 在以下的描述中, 周期 A 和周期 B 的周期长度被称作类似波形长度。

[0125] 首先, 使信号波形中的处理开始位置 P0 作为开始点, 如图 2A 所示, j 个样本的周期 A 和周期 B 被指定。接下来, 如图 2A → 图 2B → 图 2C 所示, j (即, 样本数) 逐渐增大, 并且彼此十分类似的具有周期 A 的 j 和具有周期 B 的 j 被检测到。在此, 作为测量周期 A 与周期 B 之间的相似性的尺度, 例如, 可以使用如以下等式 1 所示的函数 D(j)。

$$[0126] \quad D(j) = \frac{1}{j} \sum_i \{x(i) - y(i)\}^2 \quad (i = 0, 1, 2, \dots, j-1) \dots \text{(等式 1)}$$

[0127] 在类似长度波形的搜索范围的最小值 (WMIN) 到最大值 (WMAX) 的范围内 (即, $WMIN \leq j \leq WMAX$) 计算函数 D(j), 并且导致最小 D(j) 的 j 被获得。该导致最小 D(j) 的参数 j 是周期 A 和周期 B 的周期长度 W。顺便提及, 上述 j、WMIN 和 WMAX 表示周期的样本数。

[0128] 在此, 在上述的等式 1 中, x(i) 代表周期 A 的每个样本值, 并且 y(i) 代表周期 B 的每个样本值。此外, x(i) 代表周期 B 的每个样本值并且 y(i) 代表周期 A 的每个样本值是可以的。顺便提及, 例如, 类似波形长度的搜索频率范围可以近似是 50Hz 到 250Hz。例如, 当采样频率为 8kHz 时, 近似地, WMAX 是 160 并且 WMIN 是 32。在如图 2B 所示的示例中, j 被选作使函数 D(j) 最小的 j。

[0129] 随后, 通过参考图 3A 到图 3B, 将描述利用 PICOLA 来将音频信号拉伸到任意长度的方法。图 3A 和图 3B 是示出通过 PICOLA 来拉伸音频信号的方法的说明性示图。

[0130] 首先, 如参考图 2A 到图 2C 所描述的, 在将处理开始位置 P0 作为开始点的情况下, 使函数 D(j) 最小的 j 被获得, 并且 W 被设定为 j。随后, 周期 301 被复制到周期 303, 并且周期 301 和周期 302 的交叉淡入淡出波形被产生在周期 301 中。然后, 从在 3A 中示出的原始波形的的位置 P0 到位置 P0' 的周期被复制到在图 3B 中示出的经拉伸的波形。通过如上所述的操作, 从在图 3A 中示出的原始波形的的位置 P0 到位置 P0' 的 L 个样本变为在图 3B 中示出的经拉伸波形的 W+L 个样本, 并且样本数变为 r 倍。在此, 利用如下等式 2 来定义表示样本数的拉伸率 (样本数的增长率) 的 r。

[0131]
$$r = \frac{W+L}{L}(1.0 < r) \dots (\text{等式 } 2)$$

[0132] 在此,关于 L 重写上述等式 2 得出如下等式 3。

[0133]
$$L = W \cdot \frac{1}{r-1} \dots (\text{等式 } 3)$$

[0134] 也就是,从等式 3 显然可见,当期望将原始波形的样本数乘以 r 时,可以通过利用如下等式 4 指定位置 P0' 来实现。

[0135]
$$P0' = P0 + L \dots (\text{等式 } 4)$$

[0136] 此外,通过定义如下式 5 所示的参数 Rs,样本数 L 可以被表示为如下等式 6。

[0137]
$$R_s = \frac{1}{r}(R_s < 1.0) \dots (\text{等式 } 5)$$

[0138]
$$L = W \cdot \frac{R_s}{1-R_s} \dots (\text{等式 } 6)$$

[0139] 通过使用如上文所定义的 Rs,使得能够实现诸如原始波形“以 Rs 倍速度被回放”之类的表达。在下文中, Rs 将被称作“语速转换率”。

[0140] 当完成了对原始波形的的位置 P0 到位置 P0' 的处理时,位置 P0' 被切换到将被视为进行处理的新的开始点的位置 P1,并且相同处理被重复。通过重复这样的处理,原始波形可以被拉伸。

[0141] 在如图 3A 和 3B 所示的示例中,样本数 L 近似是 2.5W,因此,根据等式 2 和 5,语速转换率 Rs 近似是 0.7。也就是,如图 3A 和 3B 所示的示例对应于具有近似 0.7 倍速度的慢回放。

[0142] (根据 PICOLA 对波形进行压缩的处理)

[0143] 随后,通过参考图 4A 到图 5B,来描述通过 PICOLA 对波形进行压缩的处理。

[0144] 图 4A 到 4D 是示出通过使用 PICOLA 来压缩音频信号的示例的说明性视图。根据 PICOLA,首先,从如图 4A 所示的原始波形中检测出具有类似波形的周期 A 和周期 B。如图 4A 所示,周期 A 和周期 B 是连续并且具有相同长度的两个周期,并且周期 A 和周期 B 的样本数相同。顺便提及,通过参考图 2A 到图 2C 而描述的方法可以被应用于检测具有类似波形的周期。随后,在周期 A 中淡出的、在图 4B 中示出的波形和从周期 B 开始淡入的、在图 4C 中示出的波形被生成。然后,通过相加在图 4B 和图 4C 中示出的所生成波形,可以获得在图 4D 中示出的压缩波形。通过如上所述的处理,在图 4A 中示出的原始波形的周期 A 和周期 B 变为在图 4D 中示出的压缩波形的周期 A×B。

[0145] 随后,通过参考图 5A 和图 5B,来描述利用 PICOLA 来将音频信号压缩到任意长度的方法。图 5A 和图 5B 是示出通过 PICOLA 来压缩音频信号的方法的说明性视图。

[0146] 首先,如参考图 2A 到图 2C 所描述的,在将处理开始位置 P0 作为开始点的情况下,使函数 D(j) 最小的 j 被获得,并且 W 被设定为 j。随后,周期 501 和周期 502 的交叉淡入淡出波形被产生在周期 502 中。然后,其中从在图 5A 中示出的原始波形的的位置 P0 到位置 P0' 的周期中排除周期 501 的剩余周期被复制到在图 5B 中示出的压缩波形。通过上述操作,从在图 5A 中示出的原始波形的的位置 P0 到位置 P0' 的 W+L 个样本变为在图 5B 中示出的压缩波形的 L 个样本,并且样本数变为 r 倍。在此,利用如下等式 7 来定义表示样本数的压缩率

的 r 。

$$[0147] \quad r = \frac{L}{W+L} (r < 1.0) \dots (\text{等式 } 7)$$

[0148] 在此,关于 L 重写上述等式 7 得出如下等式 8。

$$[0149] \quad L = W \cdot \frac{r}{1-r} \dots (\text{等式 } 8)$$

[0150] 也就是,从等式 8 显然可见,当期望将原始波形的样本数乘以 r 时,可以通过利用如下等式 9 指定位置 $P0'$ 来实现。

$$[0151] \quad P0' = P0 + (W+L) \dots (\text{等式 } 9)$$

[0152] 此外,通过定义如下式 10 所示的参数 R_s ,样本数 L 可以被表示为如下等式 11。

$$[0153] \quad R_s = \frac{1}{r} (1.0 < R_s) \dots (\text{等式 } 10)$$

$$[0154] \quad L = W \cdot \frac{1}{R_s - 1} \dots (\text{等式 } 11)$$

[0155] 通过使用如上文所定义的 R_s ,使得能够实现诸如原始波形“以 R_s 倍速度被回放”之类的表达。当完成了对原始波形的的位置 $P0$ 到位置 $P0'$ 的处理时,位置 $P0'$ 被切换到将被视为进行处理的新的开始点的位置 $P1$,并且 相同处理被重复。通过重复这样的处理,原始波形可以被压缩。

[0156] 在如图 5A 和 5B 所示的示例中,样本数 L 近似是 $1.5W$,因此,根据等式 7 和 10,语速转换率 R_s 近似是 1.7。也就是,如图 5A 和 5B 所示的示例等价于具有近似 1.7 倍速度的快回放。

[0157] (根据 PICOLA 对信号进行拉伸的处理流程)

[0158] 随后,通过参考图 6,来简要描述根据 PICOLA 对信号进行拉伸的处理流程。图 6 是示出利用 PICOLA 来拉伸音频信号的处理流程的流程图。

[0159] 首先,根据 PICOLA,判断在实现了 PICOLA 的信息处理装置等的输入缓冲器中是否存在待处理的音频信号(步骤 S601)。在此,如果判断不存在待处理的音频信号,那么处理终止。然而,如果判断存在待处理的音频信号,那么在使处理开始位置 P 作为开始点的情况下获得使函数 $D(j)$ 最小的 j ,并且 W 被设定为 j (步骤 S602)。随后,根据 PICOLA,根据由使用者所指定的语速转换率 R_s 来获得 L (步骤 S603),并且对应于从处理开始位置 P 起的 W 个样本的周期 A 被输出到实现了 PICOLA 的信息处理装置等的输出缓冲器(步骤 S604)。

[0160] 接下来,根据 PICOLA,具有从处理开始位置 P 起的 W 个样本的周期 A 与具有从周期 A 起连续的下一组 W 个样本的周期 B 之间的交叉淡入淡出被获得,并且被置于周期 A 中(步骤 S605)。随后,输入缓冲器的具有从位置 P 起的 L 个样本的信号被输出到输出缓冲器(步骤 S606)。随后,PICOLA 将处理开始位置 P 移至 $P+L$ (步骤 S607),并且返回到步骤 S601 以重复处理。通过重复这样的处理直到在输入缓冲器中不再有待处理的音频信号,可以执行对音频信号的拉伸处理。

[0161] (根据 PICOLA 对信号进行压缩的处理流程)

[0162] 随后,通过参考图 7,来简要描述根据 PICOLA 对信号进行压缩的处理流程。图 7 是示出利用 PICOLA 来压缩音频信号的处理流程的流程图。

[0163] 首先,根据 PICOLA,判断在实现了 PICOLA 的信息处理装置等的输入缓冲器中是否

存在待处理的音频信号（步骤 S701）。在此，如果判断不存在待处理的音频信号，那么处理终止。然而，如果判断存在待处理的音频信号，那么在使处理开始位置 P 作为开始点的情况下获得使函数 $D(j)$ 最小的 j ，并且 W 被设定为 j （步骤 S702）。随后，根据 PICOLA，根据由使用者所指定的语速转换率 R_s 来获得 L（步骤 S703）。

[0164] 接下来，具有从处理开始位置 P 起的 W 个样本的周期 A 与具有从周期 A 起连续的下一组 W 个样本的周期 B 之间的交叉淡入淡出被获得，并且被置于周期 B 中（步骤 S704）。随后，输入缓冲器的具有从位置 P+W 起的 L 个样本的信号被输出到输出缓冲器（步骤 S705）。随后，PICOLA 将处理开始位置 P 移至 $P+(W+L)$ （步骤 S706），并且返回到步骤 S701 以重复处理。通过重复这样的处理直到在输入缓冲器中不再有待处理的音频信号，可以执行对音频信号的压缩处理。

[0165] （根据 PICOLA 的语速转换装置的配置）

[0166] 接下来，通过参考图 8，来描述根据 PICOLA 的语速转换装置的配置。图 8 是示出根据 PICOLA 的语速转换装置的配置的框图。顺便提及，在以下描述中，在图 1A 和图 4A 中的周期 A 和周期 B 的周期长度被称作类似波形长度。

[0167] 如图 8 所示，根据 PICOLA 的信息处理装置 800 例如包括输入缓冲器 801、类似波形长度检测部分 802、连接信号生成部分 803 和输出缓冲器 804。

[0168] 输入缓冲器 801，在缓冲被输入到信息处理装置 800 的音频信号的同时，将所输入的音频信号发送到随后将描述的类似波形长度检测部分 802 和连接信号生成部分 803，并且将根据语速转换率 R_s 而生成的音频信号输出到输出缓冲器 804。顺便提及，将被输入到输入缓冲器 801 的音频信号可以是直接被输入到信息处理装置 800 的数字信号，或者是被信息处理装置 800AD（模拟到数字）转换到数字信号的模拟信号。

[0169] 具体地，基于由随后将描述的类似波形长度检测部分 802 所检测到的类似波形长度 W，输入缓冲器 801 将 $2W$ 个样本的音频信号传送到连接信号生成部分 803。根据语速转换率 R_s ，输入缓冲器 801 将由连接信号生成部分 803 所生成的连接信号存储到输入缓冲器 801 中的适当位置中。此外，根据语速转换率 R_s ，输入缓冲器 801 将输入缓冲器 801 中的音频信号发送到输出缓冲器 804。

[0170] 关于被输入到输入缓冲器 801 的音频信号，类似波形长度检测部分 802 检测使函数 $D(j)$ 最小的参数 j ，并且所检测到的参数 j 被设定为类似波形长度 W ($W = j$)。所检测到的类似波形长度 W 被发送到输入缓冲器 801。顺便提及，所检测到的类似波形长度 W 可以直接被输出到随后将描述的连接信号生成部分 803。此外，所检测到的类似波形长度 W 可以被存储在被配置为 RAM、存储设备等的存储部分（未示出）中。

[0171] 通过使用从输入缓冲器 801 发送的音频信号和类似波形长度 W，连接信号生成部分 803 生成将被用于音频信号的拉伸 / 压缩处理中的连接信号，并且将所生成的连接信号发送到输入缓冲器 801。具体地，连接信号生成部分 803 将所接收的 $2W$ 个样本的音频信号交叉淡入淡出为 W 个样本，并且将经交叉淡入淡出的信号发送到输入缓冲器 801。此外，可以将所生成的连接信号存储在被配置为 RAM、存储设备等的存储部分（未示出）中。

[0172] 输出缓冲器 804 缓冲由输入缓冲器 801 所生成的音频信号，并且在输出缓冲器 804 上执行拉伸 / 压缩处理。被执行了拉伸 / 压缩处理的音频信号在经过 DA（数字到模拟）转换后，经由诸如扬声器之类的输出设备作为输出音频信号被输出。

[0173] (类似波形长度检测流程)

[0174] 随后,通过参考图 9 和图 10,来详细描述检测类似波形长度的处理。图 9 和图 10 是示出检测类似波形长度的处理的流程图。

[0175] 关于检测类似波形长度,首先,作为参数的索引值 j 被设定为初始值 W_{MIN} (步骤 S901)。在此,如上所述, W_{MIN} 是在其中搜索类似波形的搜索范围的最小值。当设定了类似波形长度搜索的初始值时,如图 10 所示的子程序在实现了 PICOLA 的信息处理等中被执行(步骤 S902)。如随后将描述的,该子程序是,用于计算用以判断波形之间的相似性的函数 $D(j)$ 的子程序。在此, $D(j)$ 是通过如下等式 12 被给出的函数。

$$[0176] \quad D(j) = \frac{1}{j} \sum_i \{f(i) - f(j+i)\}^2 (i = 0, 1, 2, \dots, j-1) \dots (\text{等式 12})$$

[0177] 在此,在上述等式 12 中, f 是输入音频信号,并且例如在如图 2A 到图 2C 所示的示例中,在使位置 P_0 作为开始点的情况下,它指示样本。顺便提及,等式 1 和等式 12 表示相同问题。

[0178] 随后,通过子程序而获得的函数 $D(j)$ 的值被指定给变量 $\min$,并且索引 j 被指定给 W (步骤 S903)。然后,索引 j 被增大 1(步骤 S904)。接下来,判断索引 j 是否小于 W_{MAX} (步骤 S905)。如果不小于 W_{MAX} (也就是,如果超过 W_{MAX}),那么处理终止,并且当终止处理时在变量 W 中存储的值是使函数 $D(j)$ 最小的索引 j ,也就是,类似波形长度,并且此时变量 $\min$ 的值是函数 $D(j)$ 的最小值。

[0179] 此外,如果索引 j 小于 W_{MAX} ,那么通过上述的子程序,针对新的索引 j 的函数 $D(j)$ 被获得(步骤 S906)。接下来,判断针对新的索引 j 而获得的函数 $D(j)$ 的值是否小于 $\min$ (步骤 S907)。在此,如果函数 $D(j)$ 的值小于 $\min$,那么函数 $D(j)$ 的值被指定给变量 $\min$,并且索引 j 被指定给 W (步骤 S908),然后处理返回步骤 S904。此外,如果函数 $D(j)$ 的值不小于 $\min$ (也就是,如果超过 $\min$),那么处理返回步骤 S904。通过执行这样的处理,输入音频信号的类似波形部分可以被搜索到,并且类似波形长度可以被检测到。

[0180] (函数 $D(j)$ 的值的计算)

[0181] 随后,通过参考图 10,来详细描述用于计算函数 $D(j)$ 的子程序流程,该函数 $D(j)$ 用于判断波形之间的相似性。

[0182] 当子程序的处理开始时,首先,索引 i 和变量 s 被设定为 0(步骤 S1001)。接下来,判断索引 i 是否小于索引 j (步骤 S1002)。如果索引 i 小于索引 j ,那么执行随后将描述的步骤 S1003,并且如果索引 i 不小于索引 j (也就是,如果索引 i 等于或大于索引 j),那么执行随后将描述的步骤 S1005。在此,索引 j 与如图 9 所示的流程图中的索引 j 相同。

[0183] 在步骤 S1003,输入音频信号的差被平方,然后与变量 s 相加。然后,索引 i 被增大 1(步骤 S1004),并且处理返回步骤 S1002。此外,在步骤 S1005,变量 s 被除以索引 j ,并且商被作为函数 $D(j)$ 的值,然后子程序终止。

[0184] (交叉淡入淡出信号的生成)

[0185] 随后,通过参考图 11,来详细描述在连接信号生成部分 803 中执行的生成交叉淡入淡出信号的方法。图 11 是示出生成交叉淡入淡出信号的处理示例的流程图。

[0186] 关于生成交叉淡入淡出信号,首先,索引 i 被设定为 0(步骤 S1101)。接下来,比较索引 i 和类似波形长度 W (步骤 S1102),并且如果索引 i 不小于 W (也就是,如果索引 i 等于

或大于 W)，那么处理终止。此外，如果索引 i 小于 W ，那么将被用于淡入和淡出的系数 h 被获得（步骤 S1103）。当完成了对系数 h 的计算时，淡入的信号 $x(i)$ 被乘以系数 h ，并且淡出的信号 $y(i)$ 被乘以 $(1-h)$ ，并且这些信号的和被指定给 $z(i)$ （步骤 S1104）。例如，在如图 1A 到 1D 所示的示例中，周期 A 中的信号对应于 $x(i)$ ，并且周期 B 中的信号对应于 $y(i)$ 。此外，在如图 4A 到 4D 所示的示例中，周期 B 中的信号对应于 $x(i)$ ，并且周期 A 中的信号对应于 $y(i)$ 。以这样的方式生成的信号 $z(i)$ 被作为交叉淡入淡出信号。在接下来的处理中，索引 i 被增大 1（步骤 S1105），并且处理返回到步骤 S1102。通过重复这样的处理，交叉淡入淡出信号可以被算出。

[0187] 如在上文中参考图 1A 到图 11 所描述的，通过语速转换算法，PICOLA，使得能够通过任意的语速转换率 R_s ($R_s < 1.0, 1.0 < R_s$) 来拉伸 / 压缩音频信号，并且能够关于语音信号实现尤其良好的声音质量。此外，如果语速转换率 R_s 是 1.0，那么信息处理装置 800 可以将输入音频信号如其原样地用作输出音频信号。

[0188] （关于语速转换处理的考虑）

[0189] 甚至在利用如上所述的语速转换的数字内容回放装置普及之前，针对用于盒式磁带的模拟回放装置等，已经存在可变地设定回放速度的装置。然而，通过这样的模拟回放装置，声音音高与回放速度成比例地改变，并且当回放速度被减缓时，声音音高降低，并且当回放速度被加快时，声音音高升高。

[0190] 例如，当回放内容主要由诸如用于语言学习的内容或新闻节目之类的语音构成时，如果声音音高改变，那么出现语音内容变得很难理解的问题。此外，作为另一个问题，即使声音音高仅略微地改变，识别说话人也变得很难。在其中知道哪个语音由哪个人物发出很重要的内容中，例如戏剧内容等，对于使用者而言，很难通过以不同速度回放的语音来识别说话人是回放装置的缺点。此外，还存在一个问题，即就音乐内容而言，甚至声音音高的轻微改变也会很大程度上改变音乐的情绪。在下文中，将描述如上所述的由于在以不同速度回放时声音音高的改变而引出的问题，该问题被称作第一问题。

[0191] 在维持恒定的声音音高的同时可变地设定回放速度的变速回放，即近年来在许多数字内容回放装置中实现的变速回放功能，解决了第一问题。在回放速度的范围大约是 0.5 到 4.0 倍速度的情况下，可以获得尤其良好的结果。在下文中，其中获得了尤其良好的结果的这个范围被称作第一范围，并且没有落入第一范围内的范围（即低于第一范围的下限的范围和高于第一范围的上限的范围）将被称作第二范围。容易理解，第一范围根据内容而改变。例如，如果内容的说话人的语速较慢，那么即使回放速度被大大加快也可以理解内容。然而，如果内容的说话人的语速较快，那么即使回放速度仅略微被加快也很难理解内容。

[0192] 另一方面，还存在以诸如 10 或 20 倍速度之类的高速来回放声音的需求。例如，虽然由用于盒式磁带的模拟回放装置等所提供的变速回放功能具有第一问题，但是即使当以高速回放时也能够概略地领会内容。对内容的概略领会是诸如“有个人正在说话”、“音乐正在播放”或“没有声音”之类的领会。即使是这个级别的领会，在目标内容中急速搜索期望部分时也是十分有用的。

[0193] 此外，因为回放速度被加快越多，声音音高变得越高，所以能够根据声音音高在听觉上感受到近似的回放速度。存在如下的优点：通过在听觉上识别近似的回放速度，能够本

能地感受到内容中的每个事件之间的时间位置关系（例如，诸如“有个人正在说话”、“音乐正在播放”或“没有声音”之类的事件，等等）。因此，当在目标内容中搜索期望部分时，很容易控制回放速度，例如，“该部分好像不相关所以让我们加快回放速度”或者“该部分好像相关所以让我们减缓回放速度”。因此，当在目标内容中急速搜索期望部分时十分有用。

[0194] （基础技术：转换声音音高的处理）

[0195] 在下文中，将考虑其中声音音高与回放速度成比例地改变的数字内容回放装置，例如，用于盒式磁带的模拟回放装置。作为将被用于与回放速度成比例地改变声音音高的方法的示例，例如，有用于转换采样率的方法。在下文中，通过参考图 12 和图 13，来简要描述用于转换采用率的方法示例。

[0196] （降低采样率的方法）

[0197] 图 12 是示出用于降低采样率的方法的说明性示图（下采样方法）。图 12 的 (a) 是待处理的原始信号，其中 T 是采样周期并且 f_s 是采样频率。

[0198] 在采样率转换过程中，首先，原始信号 (a) 通过低通滤波器 (LPF) 1201。低通滤波器 1201 是将截止频率设定为 $f_s/(2M)$ 的滤波器。原始信号 (a) 被低通滤波器 1201 滤波为信号 (b)。如图 12 的 (b) 所示，原始信号 (a) 的波形通过低通滤波器 1201 变得平滑。随后，下采样器 1202 从信号 (b) 中间隔剔除 (thin out) $M-1$ 个样本并针对每 M 个样本留下一个样本。在如图 12 所示的示例中， M 是 2。因此获得的信号 (c) 具有采样率 f_s/M ，即原始信号 (a) 的采样率的 $1/M$ 倍。此外，信号 (c) 的样本数也是原始信号 (a) 的样本数的 $1/M$ 倍。当低通滤波器 1201 没有被用于上述操作中时，在信号 (c) 中可能生成混迭分量 (aliasing component)。如图 12 所示的包括低通滤波器 1201 和下采样器 1202 的配置被称作抽取器 (decimator)。

[0199] （提高采样率的方法）

[0200] 图 13 是示出用于提高采样率的方法的说明性示图（上采样方法）。图 13 的 (a) 是待处理的原始信号，其中 T 是采样周期并且 f_s 是采样频率。

[0201] 在采样率转换过程中，首先，预定数目的零值被插入原始信号 (a) 中。具体地，上采样器 1301 将 $L-1$ 个零值插入到原始信号 (a) 的每个样本之间。在如图 13 所示的示例中， L 是 2。图中，经上采样的信号是信号 (b)。信号 (b) 具有 f_sL 的采样率，所述 f_sL 是 L 乘以原始信号 (a) 的采样率。此外，信号 (c) 的样本数也是 L 乘以原始信号 (a) 的样本数。随后，随着信号 (b) 通过低通滤波器 1302，信号 (c) 生成。低通滤波器 1302 是将截止频率设定为 $f_s/2$ 的滤波器。此外，在通过低通滤波器 1302 对信号 (b) 进行处理后，经处理的信号的振幅可以被调节。当低通滤波器 1302 没有被用于如上所述的操作中时，在信号 (c) 中生成镜像分量。如图 13 所示的包括上采样器 1301 和低通滤波器 1302 的配置被称作内插器。

[0202] 如图 12 所示的抽取器和如图 13 所示的内插器仅能够转换整数比值的采样率。然而，通过结合这两种配置，使得能够进行有理数采样率的转换。例如，使内插器的参数 L 为 3，并且使抽取器的参数 M 为 2。首先由内插器来处理原始信号以获得经处理的信号 1。随后，由抽取器来进一步处理经处理的信号以获得经处理的信号 2。因此获得的经处理信号 2 以因子 3 被上采样，然后被下采样到 $1/2$ ，因此采样率被转换为原始信号的采样率的 $3/2$ 倍。因此，通过结合抽取器和内插器，使得能够实现 L/M 倍的采样率转换。

[0203] 图 14A 到图 14C 是示出与回放速度成比例地升高声音音高的处理示例的说明性示图。首先,通过根据回放速度利用抽取器和内插器来转换采样率,将在图 14A 中示出的采样率为 $f_s (= 1/T)$ 的原始信号转换为在图 14B 中示出的采样率为 $f_s' (= 1/T')$ 的信号。随后,在图 14B 中示出的采样率为 $f_s' (= 1/T')$ 的信号的采样频率被在图 14A 中示出的原始信号的采样频率 $f_s (= 1/T)$ 取代,并且使其变为在图 14C 中示出的信号。因此获得的在图 14C 中示出的信号的声音音高根据回放速度的改变量,高于在图 14A 中示出的原始信号的声音音高。如图 14A 到图 14C 所示的示例示出其中回放速度是 2 倍的示例。在图 14B 中示出的信号的采样频率是在图 14A 中示出的原始信号的采样频率的 1/2 倍。此外,在图 14C 中示出的信号的声音音高是在图 14A 中示出的原始信号的声音音高的 2 倍,并且在图 14C 中示出的信号的样本数是在图 14A 中示出的原始信号的样本数的 1/2 倍。

[0204] (本实施例的描述)

[0205] 在以下描述中,其中声音音高与回放速度成比例地改变的回放装置将被称作“相关技术的第一回放装置”,并且其中当回放速度改变时恒定的声音音高被维持的回放装置将被称作“相关技术的第二回放装置”。

[0206] (相关技术的第一回放装置)

[0207] 图 15A 是示出在相关技术的第一回放装置中回放速度变量因子和语速转换率之间关系的坐标图,并且图 15B 是示出在相关技术的第一回放装置中回放速度变量因子和声音音高之间关系的坐标图。在此,图 15A 的回放速度变量因子代表回放速度与正常回放速度的比值。例如,当以 2 倍的正常回放速度进行回放时,回放速度变量因子是 2,并且当以一半的正常回放速度进行回放时,回放速度变量因子是 0.5。此外,图 15B 的声音音高代表频率与正常回放时的频率的比值。例如,当以 2 倍的正常回放频率进行回放时,声音音高是 2,并且当以一半的正常回放频率进行回放时,声音音高是 0.5。

[0208] 在相关技术的第一回放装置中,因为没有执行语速转换,所以如图 15A 所示,语速转换率是 1 并且恒定。此外,如图 15B 所示,在相关技术的第一回放装置中,声音音高与回放速度变量因子成比例,并且一般地,声音音高等于回放速度变量因子。

[0209] 顺便提及,图 15A 和图 15B 仅示出以正常速度或以比正常速度更快的速度(换言之,回放速度变量因子为 1 或更大)来进行回放的情况。在下文中,为了避免自变量变得复杂,将讨论比正常速度更快的回放速度。然而,显然,针对以小于正常速度的速度(例如,0.5 倍速度)进行回放的情况,可以做出相同的自变量。

[0210] (相关技术的第二回放装置)

[0211] 图 16A 是示出在相关技术的第二回放装置中回放速度变量因子和语速转换率之间关系的坐标图,并且图 16B 是示出在相关技术的第二回放装置中回放速度变量因子和声音音高之间关系的坐标图。在相关技术的第二回放装置中,因为执行了语速转换,所以如图 16A 所示,语速转换率与回放速度变量因子成比例,并且一般地,语速转换率的值等于回放速度变量因子值。此外,如图 16B 所示,在相关技术的第二回放装置中,声音音高为 1 并且恒定。

[0212] (关于相关技术的语速转换装置的再考虑)

[0213] 在相关技术的第二回放装置中,即使通过语速转换生成了具有超出第一范围的回放速度(换言之,在第二范围内的回放速度)的声音,也很难从听觉上感受到回放速度。例

如,通过诸如如上所述的 PICOLA 之类的语速转换算法,即使指定了例如 10 倍或 20 倍的回放速度,也能够生成相应声音。然而,通过语速转换而获得的声音物理上是 10 倍或 20 倍的速度,而在听觉上感觉在 10 倍速度和 20 倍速度之间实际没有差别。换言之,即使加快了速度,正在收听声音的收听者在转换后也不能够在听觉上感受到加速。因此,存在很难在听觉上感受到第二范围内的回放速度的问题。这样的问题被称作第二问题。

[0214] 如上所述,就相关技术的第一回放装置而言,虽然存在第一问题,但是没有出现第二问题。另一方面,就相关技术的第二回放装置而言,虽然解决了第一问题,但是出现了第二问题。

[0215] 因此,本发明的发明人就上述问题进行了认真的研究,并且实现了一种包括变速回放方法的信息处理装置,该变速回放方法通过在第一范围内的变速回放使能够容易领会语音内容或指定说话人,此外,通过在第二范围内的变速回放使能够在听觉上感受到回放速度(换言之,即能够同时解决第一和第二问题的变速回放)。

[0216] (第一实施例)

[0217] 在下文中,通过参考图 17 到图 32,来详细描述根据本发明第一实施例的信息处理装置。顺便提及,在以下描述中,回放速度变量因子将被称作第一参数,语速转换率将被称作第二参数,并且声音音高将被称作第三参数。

[0218] (回放速度转换系统)

[0219] 图 17 是示出包括根据本实施例的信息处理装置 1701 的回放速度转换系统的说明性示图。如图 17 所示,在回放速度转换系统中,作为用于控制回放速度变量因子的装置的信息处理装置 1701,可以经由诸如因特网和家用网络之类的各种网络 1702 连接到内容服务器 1703 和客户端装置 1704。此外,诸如 AV 设备(例如电视、DVD 录像机和音乐组件、计算机等)之类的各种外部连接装置 1705 可以直接连接到根据本实施例的信息处理装置 1701。

[0220] 在此,内容服务器 1703 是与诸如 URL(统一资源定位符)等之类位置信息、元数据等相关联地来管理包括音频信号的内容的服务器。例如,内容服务器 1703 可以是诸如电视、DVD 录像机和音乐组件、计算机等之类的 AV 设备,或者是符合 DLNA(数字生活网络联盟)规范的 DMS(数字媒体服务器)。此外,客户端装置 1704 是从内容服务器 1703 获得各种内容以进行回放的设备。客户端装置 1704 可以是诸如电视、DVD 录像机和音乐组件、计算机等之类的 AV 设备,或者是符合 DLNA(数字生活网络联盟)规范的 DMP(数字媒体播放器)。

[0221] (根据本实施例的信息处理装置的配置)

[0222] 图 18 是示出根据本实施例的信息处理装置 1800 的配置的框图。如图 18 所示,根据本实施例的信息处理装置 1800 主要包括参数调节部分 1801、信号处理部分 1803 和存储部分 1805。在根据本实施例的信息处理装置 1800 中,音频信号和代表回放速度变量因子的第一参数 R 被输入,并且其回放速度变量因子受第一参数 R 控制的音频信号被作为输出信号输出。

[0223] 顺便提及,在以下描述中,描述其中音频信号从信息处理装置 1800 的外部被输入的情况。然而,本实施例并不局限于这样的情况,并且音频信号可以被存储在信息处理装置 1800 中。

[0224] 参数调节部分 1801 例如被配置为 CPU(中央处理单元)、ROM(只读存储器)、RAM(随机存取存储器)等,并且根据从外部输入的第一参数 R 来调节第二参数 Rs 和第三参数 Rp。随后将详细描述根据第一参数 R 来设定第二参数 Rs 和第三参数 Rp 的方法。参数调节部分 1801 将根据第一参数 R 而确定的第二参数 Rs 和第三参数 Rp 发送到随后将描述的信号处理部分 1803。

[0225] 信号处理部分 1803 例如被配置为 CPU、ROM、RAM 等,并且基于所输入的音频信号和第一参数 R 以及从参数调节部分 1801 发送来的第二参数 Rs 和第三参数 Rp,来调节音频信号的语速和声音音高。此外,信号处理部分 1803 将被调节了语速和声音音高的音频信号作为输出音频信号 输出。信息处理装置 1800 通过未示出的 DA 转换器将这样的输出音频信号转换为模拟信号,并且将该模拟信号从诸如扬声器之类的输出设备输出。

[0226] 存储部分 1805 例如被配置为 RAM、存储设备等,并且存储在根据第一参数 R 来确定第二参数 Rs 和第三参数 Rp 时所使用的各种数据库、将由信息处理装置 1800 执行的各种程序等。此外,除这些数据外,存储部分 1805 可以视需要存储当信息处理装置 1800 执行处理、处理的中间进程等时需要存储的参数。参数调节部分 1801、信号处理部分 1803 等可以在存储部分 1805 中自由地执行读取或写入数据。

[0227] (第一参数与第二参数以及与第三参数的关系)

[0228] 随后,通过参考图 19A 和图 19B,来详细描述根据本实施例的参数调节部分 1801。图 19A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图,并且图 19B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。

[0229] 在如图 19A 和图 19B 所示的示例中,当第一参数 R 是 1 到 4 时,也就是,当以 1 到 4 倍的速度进行回放时,仅执行语速转换(区间 1901 和区间 1903),并且当第一参数 R 大于 4 时,也就是,当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高(区间 1902 和区间 1904)。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0230] 顺便提及,在图 19A 中,以虚线示出区间 1902,因为第二参数 Rs 的值根据改变声音音高的方法而改变。当将如图 12 到 14 所示的方法用作改变声音音高的方法时,样本数随着声音音高的升高而减少,从而导致区间 1902 的虚线。然而,当将其中样本数不减少的方法或者其中减少量较小的方法用作改变声音音高的方法时,将与如图 19 所示的虚线不同地设定区间 1902。

[0231] 在图 19B 的区间 1903 中,当第一参数 R 是 1 到 4 时,第三参数 Rp 是 1 并且恒定。然而,在该区间中的第三参数 Rp 不一定是恒定的。此外,在区间 1904 中第三参数 Rp 的上坡度并不局限于如图所示的示例,并且它可以是任意的,只要第三参数 Rp 具有大于 0 的上坡度。此外,在图 19A 和 图 19B 中,虽然第二参数 Rs 和第三参数 Rp 以连续的方式(模拟地)改变,但是第二参数 Rs 和第三参数 Rp 也可以以离散的方式(数字地)改变。

[0232] (参数调节部分 1801)

[0233] 在根据本实施例的信息处理装置 1800 中,如图 19A 和图 19B 所示的第一参数 R 与第二参数 Rs 以及与第三参数 Rp 的关系的数据库例如被存储在存储部分 1805 中,并且参数调节部分 1801 通过参考这样的数据库,根据第一参数 R 来确定第二参数 Rs 和第三参数

Rp。

[0234] 参数调节部分 1801 在如下所述的四个条件下,通过参考如图 19A 和图 19B 所示的数据库,根据第一参数 R 来确定第二参数 Rs 和第三参数 Rp。

[0235] 条件 1 :当所输入的第一参数 R 存在于区间 1901 中时,将第二参数 Rs 确定为与第一参数 R 成比例 (换言之,确定第二参数 Rs 以使得第二参数 Rs 等于第一参数 R)。

[0236] 条件 2 :当所输入的第一参数 R 存在于区间 1903 中时,将第三参数 Rp 恒定地设定为 1。

[0237] 条件 3 :当所输入的第一参数 R 存在于区间 1904 中时,第三参数 Rp 随第一参数 R 的增大而增大。

[0238] 条件 4 :第一参数 $R = \text{第二参数 } R_s \times \text{样本数的增长率 } R_d$ 。

[0239] 在此,区间 1901 和区间 1903 对应于第一参数 R 的第一范围,并且区间 1902 和区间 1904 对应于第一参数 R 的第二范围。

[0240] 此外,当在改变声音音高的方法中样本数的增长率是 R_d 时,参数调节部分 1801 的第一范围和第二范围二者具有如上述条件 4 所指示的特性。在此,例如,当样本数是 2 倍时,增长率是 2,并且当样本数被减少到一半时,增长率是 $1/2$ 。

[0241] (根据本实施例的控制回放速度变量因子的方法)

[0242] 图 20 是示出根据本实施例的信息处理装置 1800 的处理流程的流程图。首先,信息处理装置 1800 判断是否存在输入音频信号 (步骤 S2001),并且当不存在输入音频信号时,处理终止。此外,当输入音频信号存在时,信息处理装置 1800 的参数调节部分 1801 根据所输入的第一参数 R 来调节第二参数 Rs 和第三参数 Rp (步骤 S2002)。以满足上述条件 1 到 4 的方式来执行调节。随后,信息处理装置 1800 的信号处理部分 1803 根据经调节的第二参数 Rs 和第三参数 Rp 来调节输入音频信号的语速和声音音高 (步骤 S2003)。随后,信息处理装置 1800 输出语速和声音音高经过调节的音频信号 (步骤 S2004)。然后,返回到步骤 S2001,上述处理被重复。

[0243] 通过重复这样的处理,根据本实施例的信息处理装置 1800 被使得能够控制音频信号的回放速度变量因子。

[0244] 如参考图 18 到图 20 所描述的,根据本实施例的控制回放速度变量因子的方法,能够在第一参数 R 的第一范围内仅调节语速,并且在第一参数 R 的第二范围内随同语速一起来调节声音音高。因此,在第一参数 R 的第一范围内解决了第一问题,并且在第一参数 R 的第二范围内解决了第二问题。

[0245] (信号处理部分 1803)

[0246] 随后,通过参考图 21,来详细描述根据本实施例的信号处理部分 1803 的示例。图 21 是示出根据本实施例的信号处理部分 1803 的功能的框图。

[0247] 如图 21 所示,例如,根据本实施例的信号处理部分 1803 主要包括拟声声音 (onomatopoeic sound) 切换判断部分 2101、语速转换部分 2103、音高调节部分 2105 和音频信号输出控制部分 2107。

[0248] 拟声声音切换判断部分 2101 例如被配置为 CPU、ROM、RAM 等,并且基于所发送的第一参数 R,判断是对输入音频信号执行诸如语速和声音音高转换之类的信号处理,还是将输入音频信号切换到拟声声音而不执行信号处理。具体地,拟声声音切换判断部分 2101 将

第一参数 R 的等级与预定阈值进行比较,并且当第一参数 R 大于预定阈值(例如,以大于 20 倍的速度回放)时,判断将音频信号切换到预定拟声声音而不执行语速和声音音高的转换。拟声声音切换判断部分 2101 将判断结果发送到随后将描述的语速转换部分 2103 和音频信号输出控制部分 2107。

[0249] 语速转换部分 2103 例如被配置为 CPU、ROM、RAM 等。输入音频信号和由参数调节部分 1801 确定的第二参数 R_s 被输入到语速转换部分 2103,并且语速转换部分 2103 基于第二参数 R_s 来转换输入音频信号的语速。例如,利用如图 1 到图 7 所示的算法来执行语速转换。语速转换部分 2103 将语速经过调节的音频信号发送到随后将描述的音高调节部分 2105。

[0250] 此外,当被拟声声音切换判断部分 2101 通知“将音频信号切换到拟声声音”的判断结果时,语速转换部分 2103 不必执行转换语速的处理。

[0251] 音高调节部分 2105 例如被配置为 CPU、ROM、RAM 等,并且基于从语速转换部分 2103 发送来的、语速经过调节的音频信号以及从参数调节部分 1801 发送来的第三参数 R_p ,来调节音频信号的声音音高。音高转换的任意方法,例如,如图 12 到图 14C 所示的方法,可以用于音高调节。当完成了声音音高调节时,音高调节部分 2105 将语速和声音音高经过调节的音频信号输出到随后将描述的音频信号输出控制部分 2107。

[0252] 顺便提及,当如图 12 到图 14C 所示的方法被音高调节部分 2105 使用时,该方法中用于改变声音音高的样本数的增长率 R_d 与声音音高成比例,并且样本数的增长率 R_d 变得等于声音音高的上升率。也就是, $R_d = \text{第三参数 } R_p$ 的关系成立。

[0253] 音频信号输出控制部分 2107 例如被配置为 CPU、ROM、RAM 等,并且当输出所输入的音频信号或者从音高调节部分 2105 发送来的音频信号时控制输出。当被拟声声音切换判断部分 2101 通知“将音频信号切换到拟声声音”的判断结果时,音频信号输出控制部分 2107 将所输入的音频信号切换到例如被存储在存储部分 1805 中的预定拟声声音并输出信号。此外,当被拟声声音切换判断部分 2101 通知“不将音频信号切换到拟声声音”的判断结果时,音频信号输出控制部分 2107 输出从音高调节部分 2105 发送的音频信号。

[0254] 此外,音频信号输出控制部分 2107 可以调节待输出的音频信号的音频音量。通过调节指定音频信号的信号波形的绝对值来执行对音频信号的音频音量的调节。当回放速度变量因子超过 1 时,音频信号输出控制部分 2107 可以调小音频信号的音频音量。此外,音频信号输出控制部分 2107 可以独立于回放速度来控制音频音量。

[0255] 图 22A 和图 22B 是示出由包括如图 21 所示的信号处理部分 1803 的信息处理装置 1800 的参数调节部分 1801 所执行的参数调节方法示例的说明性示图。图 22A 是示出第一参数 R 和第二参数 R_s 之间关系的坐标图,并且图 22B 是示出第一参数 R 和第三参数 R_p 之间关系的坐标图。

[0256] 如图 22A 所示,横轴代表第一参数 R 并且纵轴代表第二参数 R_s 的坐标图被配置为具有第二参数 R_s 的不同上升率(换言之,坐标图的坡度(gradient))的至少两个区域。类似地,如图 22B 所示,横轴代表第一参数 R 并且纵轴代表第三参数 R_p 的坐标图被配置为具有第三参数 R_p 的不同上升率的至少两个区域。

[0257] 当信号处理部分 1803 的音高调节部分 2105 通过如图 12 到图 14C 所示的方法来调节音高时,参数调节部分 1801 在下述的四个条件下,通过参考在存储部分 1805 中存储的、

如图 22A 和图 22B 所示的数据库,根据第一参数 R 来确定第二参数 Rs 和第三参数 Rp。

[0258] 条件 1 :当所输入的第一参数 R 存在于区间 2201 中时,将第二参数 Rs 确定为与第一参数 R 成比例 (换言之,确定第二参数 Rs 以使得第二参数 Rs 等于第一参数 R)。

[0259] 条件 2 :当所输入的第一参数 R 存在于区间 2203 中时,将第三参数 Rp 恒定地设定为 1。

[0260] 条件 3 :当所输入的第一参数 R 存在于区间 2204 中时,第三参数 Rp 随第一参数 R 的增大而增大。

[0261] 条件 4' :第一参数 $R = \text{第二参数 } R_s \times \text{第三参数 } R_p$ 在第一范围和第二范围内均成立。

[0262] 在此,区间 2201 和区间 2203 对应于第一参数 R 的第一范围,并且区间 2202 和区间 2204 对应于第一参数 R 的第二范围。

[0263] 在如图 22A 和图 22B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高 声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0264] 至此,已经描述了根据本实施例的信息处理装置 1800 的功能示例。上述结构元件的每个可以被配置为通用组件或电路,或者可以被配置为专用于每个结构元件的功能的硬件。此外,CPU 等可以执行所有的功能。因此,能够根据实现本实施例的各种技术水平来适当改变待使用的配置。

[0265] (根据本实施例的信号处理方法)

[0266] 随后,通过参考图 23,来详细描述根据本实施例的信号处理方法。图 23 是示出根据本实施例的信号处理方法的流程图。

[0267] 首先,信息处理装置 1800 判断是否存在输入音频信号 (步骤 S2301),并且如果没有输入音频信号那么终止处理。此外,当输入音频信号存在时,信号处理部分 1803 的拟声声音切换判断部分 2101 判断所输入的第一参数 R 是否大于预定阈值 (步骤 S2302)。当第一参数 R 小于预定阈值时,参数调节部分 1801 根据所输入的第一参数 R 来调节第二参数 Rs 和第三参数 Rp (步骤 S2303),并且将参数发送到信号处理部分 1803。信号处理部分 1803 的语速转换部分 2103 基于所发送的第二参数 Rs 来调节输入音频信号的语速 (步骤 S2304),并且将语速经过调节的音频信号输出到音高调节部分 2105。音高调节部分 2105 基于所发送的第三参数 Rp 来调节从语速转换部分 2103 发送的音频信号的声音音高 (步骤 S2305)。语速和声音音高经过调节的音频信号被发送到音频信号输出控制部分 2107,并且音频信号输出控制部分 2107 输出语速和声音音高经过调节的音频信号 (步骤 S2306)。然后,返回到步骤 S2301,上述处理被重复。

[0268] 另一方面,当拟声声音切换判断部分 2101 判断第一参数 R 大于预定阈值时,音频信号输出控制部分 2107 输出被存储在存储部分 1805 等中的预定拟声声音,并且将所述拟声声音作为音频信号输出 (步骤 S2307)。然后,返回到步骤 S2301,上述处理被重复。

[0269] 通过重复这样的处理,根据本实施例的信息处理装置 1800 被使得能够以在听觉上可以识别转换后的回放速度的方式来控制音频信号的回放速度变量因子。

[0270] 随后,通过关注在待处理的音频信号中所包括的样本数,来详细描述由根据本实施例的信息处理装置 1800 所执行的信号处理示例。图 24A 到图 24D 是以样本为单位示出由根据本实施例的信息处理装置 1800 所执行的信号处理示例的说明性示图。

[0271] 在如图 24A 到图 24D 所示的示例中,当第一参数 R 是 2.5 时,第二参数 R_s 被调节为 2.0 并且第三参数 R_p 被调节到 1.25。假设在如图 24A 所示的原始信号中,作为在将语速转换的处理开始点 P0 作为开始点的情况下检测到类似波形长度的结果,区间 2401 和区间 2402 被选作交叉淡入淡出区间。区间 2401 的信号和区间 2402 的信号的交叉淡入淡出信号被获得,并且被置于区间 2402 中。随后,区间 2402 的信号被复制到区间 2403 的、如图 24B 所示的信号,并且语速转换的处理开始位置从位置 P0 移至位置 P1。通过将在图 24A 中示出的原始信号转换到在图 24B 中示出的信号,语速变为 2 倍速度(样本数变为 $1/2$ 倍),并且声音音高不变。随后,在图 24B 中示出的信号的采样频率被变为 $4/5$ 倍以获得在图 24C 中示出的信号。当采样频率变为 $4/5$ 倍时,样本数也变为 $4/5$ 倍。通过用在图 24A 中示出的原始信号的采样频率来替换在图 24C 中示出的信号的采样频率,在图 24D 中示出的信号被获得。在图 24D 中示出的信号样本数是 $0.4 = (1/2) \times (4/5)$ 倍的在图 24A 中示出的原始信号的样本数,并且声音音高是 $5/4$ 倍。换言之,回放速度是 $2.5 = 2 \times (5/4)$ 倍的速度并且声音音高是 1.25 倍。

[0272] 图 25A 到图 25D 是以样本为单位示出由根据本实施例的信息处理装置所执行的信号处理的另一示例的说明性示图。在如图 25A 到图 25D 所示的示例中,当第一参数 R 是 4.0 时,第二参数 R_s 被调节为 2.0 并且第三参数 R_p 被调节到 2.0。假设在如图 25A 所示的原始信号中,作为在将语速转换的处理开始点 P0 作为开始点的情况下检测到类似波形长度的结果,区间 2501 和区间 2502 被选作交叉淡入淡出区间。区间 2501 的信号和区间 2502 的信号的交叉淡入淡出信号被获得,并且被置于区间 2502 中。随后,区间 2502 的信号被复制到区间 2503 的、如图 25B 所示的信号,并且语速转换的处理开始位置从位置 P0 移至位置 P1。通过将在图 25A 中示出的原始信号转换到在图 25B 中示出的信号,语速变为 2 倍速度(样本数变为 $1/2$ 倍),并且声音音高不变。随后,在图 25B 中示出的信号的采样频率被变为 $1/2$ 倍以获得在图 25C 中示出的信号。当采样频率变为 $1/2$ 倍时,样本数也变为 $1/2$ 倍。通过用在图 25A 中示出的原始信号的采样频率来替换在图 25C 中示出的信号的采样频率,在图 25D 中示出的信号被获得。在图 25D 中示出的信号样本数是 $0.25 = (1/2) \times (1/2)$ 倍的在图 25A 中示出的原始信号的样本数,并且声音音高是 2 倍。换言之,回放速度是 $4.0 = 2 \times 2$ 倍的速度并且声音音高是 2 倍。

[0273] 图 26A 和图 26B 是示出由参数调节部分 1801 所执行的参数调节方法的其他示例的坐标图。图 26A 是示出第一参数 R 和第二参数 R_s 之间关系的坐标图,并且图 26B 是示出第一参数 R 和第三参数 R_p 之间关系的坐标图。

[0274] 如图 26A 所示,横轴代表第一参数 R 并且纵轴代表第二参数 R_s 的坐标图被配置为具有第二参数 R_s 的不同上升率(换言之,坐标图的坡度)的至少两个区域。类似地,如图 26B 所示,横轴代表第一参数 R 并且纵轴代表第三参数 R_p 的坐标图被配置为具有第三参数 R_p 的不同上升率的至少两个区域。

[0275] 在这种情况下,参数调节部分 1801 在下述的五个条件下,通过参考在存储部分 1805 中存储的、如图 26A 和图 26B 所示的数据库,根据第一参数 R 来确定第二参数 R_s 和第

三参数 R_p 。

[0276] 条件 1 :当所输入的第一参数 R 存在于区间 2601 中时,将第二参数 R_s 确定为与第一参数 R 成比例 (换言之,确定第二参数 R_s 以使得第二参数 R_s 等于第一参数 R)。

[0277] 条件 2 :当所输入的第一参数 R 存在于区间 2603 中时,将第三参数 R_p 恒定地设定为 1。

[0278] 条件 3 :当所输入的第一参数 R 存在于区间 2604 中时,第三参数 R_p 随第一参数 R 的增大而增大。

[0279] 条件 4' :第一参数 $R = \text{第二参数 } R_s \times \text{第三参数 } R_p$ 在第一范围和第二 范围内均成立。

[0280] 条件 5 :当所输入的第一参数 R 存在于区间 2602 中时,第二参数 R_s 随第一参数 R 的增大而增大 (换言之,示出第二参数 R_s 变化的曲线的导数大于 0)。

[0281] 在此,区间 2601 和区间 2603 对应于第一参数 R 的第一范围,并且区间 2602 和区间 2604 对应于第一参数 R 的第二范围。

[0282] 在如图 26A 和图 26B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0283] 在如图 26A 和图 26B 所示的示例中,与如图 22A 和图 22B 所示的示例不同的,第二参数 R_s 随着第一参数 R 的增大而增大。换言之,示出第二参数 R_s 变化的曲线的导数大于 0。在图 22A 的区间 2202 中,尽管第一参数 R 增大,但是第二参数 R_s 恒定。换言之,第二参数 R_s 的导数为 0。在这样的情况下,尽管回放速度加快但是语速转换率不变,并且关于所回放的声音可能会感受到不适。另一方面,在图 26A 中的区间 2602 中,因为第二参数 R_s 随第一参数 R 的增大而增大 (因为导数大于 0),所以可以防止尽管语速加快但是语速转换率不变的情况,并且由所回放的声音所导致的不适可以被防止。

[0284] 图 27A 和图 27B 是示出由参数调节部分 1801 所执行的参数调节方法的其他示例的坐标图。图 27A 是示出第一参数 R 和第二参数 R_s 之间关系的坐标图,并且图 27B 是示出第一参数 R 和第三参数 R_p 之间关系的坐标图。

[0285] 如图 27A 所示,横轴代表第一参数 R 并且纵轴代表第二参数 R_s 的坐标图被配置为具有第二参数 R_s 的不同上升率 (换言之,坐标图的坡度) 的至少两个区域。类似地,如图 27B 所示,横轴代表第一参数 R 并且纵轴代表第三参数 R_p 的坐标图被配置为具有第三参数 R_p 的不同上升率的至少 两个区域。

[0286] 在这种情况下,参数调节部分 1801 在下述的五个条件下,通过参考在存储部分 1805 中存储的、如图 27A 和图 27B 所示的数据库,根据第一参数 R 来确定第二参数 R_s 和第三参数 R_p 。

[0287] 条件 1 :当所输入的第一参数 R 存在于区间 2701 中时,将第二参数 R_s 确定为与第一参数 R 成比例 (换言之,确定第二参数 R_s 以使得第二参数 R_s 等于第一参数 R)。

[0288] 条件 2 :当所输入的第一参数 R 存在于区间 2703 中时,将第三参数 R_p 恒定地设定为 1。

[0289] 条件 3:当所输入的第一参数 R 存在于区间 2704 中时,第三参数 R_p 随第一参数 R 的增大而增大。

[0290] 条件 4':第一参数 $R = \text{第二参数 } R_s \times \text{第三参数 } R_p$ 在第一范围和第二范围内均成立。

[0291] 条件 6:区间 2703 和区间 2704 被平滑地连接(换言之,示出第三参数 R_p 变化的曲线在区间 2703 和区间 2704 的连接点处可微分)。

[0292] 在此,区间 2701 和区间 2703 对应于第一参数 R 的第一范围,并且区间 2702 和区间 2704 对应于第一参数 R 的第二范围。

[0293] 在如图 27A 和图 27B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0294] 在如图 27A 和图 27B 所示的示例中,与如图 22A 和图 22B 所示的示例不同的,在第三参数 R_p 中,区间 2703 和区间 2704 被平滑地连接。换言之,示出第三参数 R_p 变化的曲线在区间 2703 和区间 2704 的连接点处可微分。在如图 22A 和图 22B 所示的区间 2203 和区间 2204 的连接点不可微分的情况下,当第一参数 R 逐渐增大时,第三参数 R_p 的单位增大量(微分值)在连接点处激增,并且关于所回放的声音可能会感到不适。另一方面,在如图 27B 中区间 2703 和区间 2704 所示的曲线平滑连接的情况下,当第一参数 R 逐渐增大时,可以防止声音音高在区间 2703 和区间 2704 的连接点处开始激增,并且关于所回放声音的不适可以被防止。

[0295] 图 28A 和图 28B 是示出由参数调节部分 1801 所执行的参数调节方法的其他示例的坐标图。图 28A 是示出第一参数 R 和第二参数 R_s 之间关系的坐标图,并且图 28B 是示出第一参数 R 和第三参数 R_p 之间关系的坐标图。

[0296] 如图 28A 所示,横轴代表第一参数 R 并且纵轴代表第二参数 R_s 的坐标图被配置为具有第二参数 R_s 的不同上升率(换言之,坐标图的坡度)的至少两个区域。类似地,如图 28B 所示,横轴代表第一参数 R 并且纵轴代表第三参数 R_p 的坐标图被配置为具有第三参数 R_p 的不同上升率的至少两个区域。

[0297] 在这种情况下,参数调节部分 1801 在下述的六个条件下,通过参考在存储部分 1805 中存储的、如图 28A 和图 28B 所示的数据库,根据第一参数 R 来确定第二参数 R_s 和第三参数 R_p 。

[0298] 条件 1:当所输入的第一参数 R 存在于区间 2801 中时,将第二参数 R_s 确定为与第一参数 R 成比例(换言之,确定第二参数 R_s 以使得第二参数 R_s 等于第一参数 R)。

[0299] 条件 2:当所输入的第一参数 R 存在于区间 2803 中时,将第三参数 R_p 恒定地设定为 1。

[0300] 条件 3:当所输入的第一参数 R 存在于区间 2804 中时,第三参数 R_p 随第一参数 R 的增大而增大。

[0301] 条件 4':第一参数 $R = \text{第二参数 } R_s \times \text{第三参数 } R_p$ 在第一范围和第二范围内均成立。

[0302] 条件 5 :当所输入的第一参数 R 存在于区间 2802 中时,第二参数 R_s 随第一参数 R 的增大而增大 (换言之,示出第二参数 R_s 变化的曲线的导数大于 0)。

[0303] 条件 6 :区间 2803 和区间 2804 被平滑地连接 (换言之,示出第三参数 R_p 变化的曲线在区间 2803 和区间 2804 的连接点处可微分)。

[0304] 在此,区间 2801 和区间 2803 对应于第一参数 R 的第一范围,并且区间 2802 和区间 2804 对应于第一参数 R 的第二范围。

[0305] 在如图 28A 和图 28B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0306] 在如图 28A 和图 28B 所示的示例中,与如图 27A 和图 27B 所示的示例类似地,在第三参数 R_p 中,区间 2803 和区间 2804 被平滑地连接。换言之,示出第三参数 R_p 变化的曲线在区间 2803 和区间 2804 的连接点处可微分。另一方面,在如图 28A 和图 28B 所示的示例中,与如图 27A 和图 27B 所示的示例不同的,第二参数 R_s 随第一参数 R 的增大而增大。换言之,示出第二参数 R_s 变化的曲线的导数大于 0。在图 27A 的区间 2702 中,尽管第一参数 R 增大,但是存在第二参数 R_s 减小的部分。换言之,存在示出第二参数 R_s 变化的曲线的导数为负的部分。在这样的情况下,尽管回放速度加快但是语速转换率减小,并且关于所回放的声音可能感到不适。另一方面,在图 28A 的区间 2802 中,因为第二参数 R_s 随着第一参数 R 的增大而增大 (因为导数大于 0),所以可以防止尽管语速加快但是语速转换率减小的情况,并且关于所回放的声音的不适可以被防止。

[0307] 如上所述,通过当转换所输入音频信号的回放速度变量因子时,在调节声音音高之前转换语速,可以在语速转换中更精确地执行对输入音频信号的类似波形长度的检测,并且能够将输出音频信号的声音质量维持在其最佳状态。

[0308] (信号处理部分 1803 的修改示例)

[0309] 随后,通过参考图 29,来详细描述根据本实施例的信号处理部分 1803 的修改示例。图 29 是示出根据本实施例的信号处理部分 1803 的修改示例的框图。

[0310] 如图 29 所示,例如,根据该修改示例的信号处理部分 1803 主要包括 拟声声音切换判断部分 2101、音高调节部分 2901、语速转换部分 2903 和音频信号输出控制部分 2107。

[0311] 除了拟声声音切换判断部分 2101 将判断结果输出到音高调节部分 2901 和音频信号输出控制部分 2107 外,拟声声音切换判断部分 2101 具有与根据本发明第一实施例的拟声声音切换判断部分相同的配置和功能,因此,省略其详细描述。

[0312] 音高调节部分 2901 例如被配置为 CPU、ROM、RAM 等,并且基于所发送的输入音频信号以及从参数调节部分 1801 发送来的第三参数 R_p ,来调节音频信号的声音音高。音高转换的任意方法,例如,如图 12 到图 14C 所示的方法,可以用于音高调节。当完成了声音音高调节时,音高调节部分 2901 将声音音高经过调节的音频信号输出到随后将描述的语速转换部分 2903。

[0313] 顺便提及,当如图 12 到图 14C 所示的方法被音高调节部分 2901 使用时,该方法中用于改变声音音高的样本数的增长率 R_d 与声音音高成比例,并且样本数的增长率 R_d 变得

等于声音音高的上升率。也就是, $R_d =$ 第三参数 R_p 的关系成立。

[0314] 此外, 当被拟声声音切换判断部分 2101 通知“将音频信号切换到拟声声音”的判断结果时, 音高调节部分 2901 不必执行转换声音音高的处理。

[0315] 语速转换部分 2903 例如被配置为 CPU、ROM、RAM 等。输入音频信号和由参数调节部分 1801 确定的第二参数 R_s 被输入到语速转换部分 2903, 并且语速转换部分 2903 基于第二参数 R_s 来转换输入音频信号的语速。例如, 利用如图 1 到图 7 所示的算法来执行语速转换。语速转换部分 2903 将语速经过调节的音频信号发送到随后将描述的音频信号输出控制部分 2107。

[0316] 音频信号输出控制部分 2107 例如被配置为 CPU、ROM、RAM 等, 并且当输出所输入的音频信号或者从音高调节部分 2105 发送来的音频信号时控制输出。当被拟声声音切换判断部分 2101 通知“将音频信号切换到拟声声音”的判断结果时, 音频信号输出控制部分 2107 将所输入的音频信号切换到例如被存储在存储部分 1805 中的预定拟声声音并输出信号。此外, 当被拟声声音切换判断部分 2101 通知“不将音频信号切换到拟声声音”的判断结果时, 音频信号输出控制部分 2107 输出从语速转换部分 2903 发送的音频信号。

[0317] 此外, 音频信号输出控制部分 2107 可以调节待输出的音频信号的音频音量。通过调节指定音频信号的信号波形的绝对值来执行对音频信号的音频音量的调节。当回放速度变量因子超过 1 时, 音频信号输出控制部分 2107 可以调小音频信号的音频音量。此外, 音频信号输出控制部分 2107 可以独立于回放速度来控制音频音量。

[0318] 至此, 已经描述了根据修改示例的信号处理部分 1803 的功能示例。上述结构元件的每个可以被配置为通用组件或电路, 或者可以被配置为专用于每个结构元件的功能的硬件。此外, CPU 等可以执行所有的功能。因此, 能够根据实现本实施例的各种技术水平来适当改变待使用的配置。

[0319] (根据修改示例的信号处理方法)

[0320] 随后, 通过参考图 30, 来详细描述根据修改示例的信号处理方法。图 30 是示出根据修改示例的信号处理方法的流程图。

[0321] 首先, 信息处理装置 1800 判断是否存在输入音频信号 (步骤 S3001), 并且当没有输入音频信号时终止处理。此外, 当输入音频信号存在时, 信号处理部分 1803 的拟声声音切换判断部分 2101 判断所输入的第一参数 R 是否大于预定阈值 (步骤 S3002)。当第一参数 R 小于预定阈值时, 参数调节部分 1801 根据所输入的第一参数 R 来调节第二参数 R_s 和第三参数 R_p (步骤 S3003), 并且将参数发送到信号处理部分 1803。信号处理部分 1803 的音高调节部分 2901 基于所发送的第三参数 R_p 来调节所发送的输入音频信号的声音音高 (步骤 S3004), 并且将声音音高经过调节的音频信号输出到语速转换部分 2903。语速转换部分 2903 基于所发送的第二参数 R_s 来调节声音音高经过调节的音频信号的语速 (步骤 S3005)。语速和声音音高经过调节的音频信号被发送到音频信号输出控制部分 2107, 并且音频信号输出控制部分 2107 输出语速和声音音高经过调节的音频信号 (步骤 S3006)。然后, 返回到步骤 S3001, 上述处理被重复。

[0322] 另一方面, 当拟声声音切换判断部分 2101 判断第一参数 R 大于预定阈值时, 音频信号输出控制部分 2107 将被存储在存储部分 1805 等中的预定拟声声音作为音频信号输出 (步骤 S3007)。然后, 返回到步骤 S3001, 上述处理被重复。

[0323] 通过重复这样的处理,根据修改示例的信息处理装置 1800 被使得能够以在听觉上可以识别转换后的回放速度的方式来控制音频信号的回放速度变量因子。

[0324] 如上所述,通过当转换所输入的音频信号的回放速度变量因子时,在转换语速之前调节声音音高,能够减少语速待转换的输入音频信号的样本数,并且能够减少待处理的资源,因此能够实现处理的加速。顺便提及,当转换声音音高经过调节的音频信号的语速时,可以根据音高调节的程度来适当改变执行语速转换的频率范围。

[0325] (转换采样率的其他方法)

[0326] 图 31 是示出通过不同于如图 12 和图 13 所示的转换采样率方法的方法来转换采样率的方法的说明性示图。通常,在如图 12 和图 13 所示的方法中,处理量很大,因此例如在诸如便携式回放装置之类的不期望高处理性能的回放装置中很难实现它们。在这样的情况下,如图 31 所示的转换采样率的方法证明是有用的。图 31 是示出这样的情况的说明确性示图,其中当采样点 n_0 、 n_1 、 n_2 、 n_3 、... 存在于转换前的信号中时,新的采样点 m_0 、 m_1 、 m_2 、... 通过线性插值被获得。例如,关于 m_1 的采样值,线性插值通过计算比值 $p_1:1-p_1$ 来获得在采样点 n_1 和采样点 n_2 之间的采样点 m_1 的位置,并且根据该比值,根据 n_1 的采样值和 n_2 的采样值来获得 m_1 的采样值。

[0327] 如此,在本实施例中,调节声音音高的方法并不限于如图 12 和图 13 所示的那些方法,并且可以使用任意的的方法,例如在图 31 中示出的方法,以及满足根据本实施例的信息处理装置的条件的那些方法。

[0328] (回放速度变量因子的渐变)

[0329] 随后,通过参考图 32,来描述连续改变代表回放速度变量因子的第一参数 R 的情况。图 32 是示意性地示出回放速度变量因子随时间的变化的说明性示图。

[0330] 与代表回放速度变量因子的第一参数 R 被设定为 R_1 并且输出音频信号的信息处理装置 1800 相对地,当在时间点 t_1 输入用于将第一参数 R 变到 R_2 的信号时,例如,如图 32 所示,根据本实施例的信息处理装置 1800 并不立即数字地切换第一参数 R ,而是可以控制第二参数和第三参数从而使得第一参数逐渐从 R_1 切换到 R_2 。

[0331] 在这样的情况下,参数调节部分 1801 使第一参数 R 连续地从 R_1 变到 R_2 ,并且针对渐变中的每个参数 R 来设定第二参数 R_s 和第三参数 R_p 。通过执行这样的处理,即使在改变音频信号的语速和声音音高期间,音频信号的收听者也可以听到音频信号而不会感觉到不适。

[0332] 如上所述,通过根据本实施例的控制回放速度变量因子的方法,当以近似正常的速度进行回放时,回放速度改变然而声音音高不变,并且容易理解说话人的语音内容或者容易识别说话人。此外,在高速回放/低速回放的情况下,当回放速度改变时,因此可以在听觉上感觉到此时的回放速度,并且可操作性可以得到改进。

[0333] (第二实施例)

[0334] 随后,通过参考图 33 到图 46,来详细描述根据本发明第二实施例的信息处理装置 3300。

[0335] 当所谓的内容回放装置回放内容时,装置从内容回放装置的诸如硬盘驱动器、DVD 驱动器和蓝光驱动器之类的记录介质回放装置获得音频信号。但是,这样的记录介质回放装置的数据读取速度存在上限。换言之,每单位时间可以从记录介质中读取的数据量存在

上限。因此,即使能够获得足够以 10 倍速度回放内容的数据量,也不一定能够获得足够以 20 倍速度回放内容的数据量。还存在其他类似情况。例如,近年来,内容数据通常以 MPEG 等进行编码,并且当回放经编码的内容时,首先,它必须被解码。因此,即使诸如硬盘驱动器、DVD 驱动器和蓝光驱动器之类的记录介质回放装置的数据读取速度足够快,如果解码设备的计算能力不够,那么解码处理也不能继续。当连接诸如硬盘驱动器、DVD 驱动器和蓝光驱动器之类的记录介质回放装置,以及 CPU 或存储器的总线带宽不够时,也会出现类似情况。

[0336] 因此,构成内容回放装置的每个结构元件都有其处理能力限制,并且当以变速回放时,整个装置的处理能力限制由处理能力限制最低的结构元件来决定。存在这样的问题,即存在由于该处理能力限制而无法实现期望回放速度的情况。在下文中,这个问题被称作第三问题。

[0337] 因此,本发明的发明人就上述问题进行了认真研究,并且实现了一种变速回放方法,该方法通过在第一范围内的变速回放使能够容易领会语音内容或者指定说话人,此外,通过在第二范围内的变速回放使能够在听觉上感受到回放速度,并且此外,实现回放速度的更高上限。换言之,根据本实施例的变速回放方法是能够同时解决第一、第二和第三问题的变速回放方法。

[0338] (根据本实施例的信息处理装置的配置)

[0339] 首先,通过参考图 33,来详细描述根据本实施例的信息处理装置 3300 的配置。图 33 是示出根据本实施例的信息处理装置 3300 的功能的框图。

[0340] 如图 33 所示,根据本实施例的信息处理装置 3300 例如主要包括参数调节部分 3301、内容管理部分 3303、内容存储部分 3305、信号处理部分 3307 和存储部分 3309。

[0341] 参数调节部分 3301 例如被配置为 CPU、ROM、RAM 等,并且根据从外部输入的第一参数 R 来调节第二参数 Rs、第三参数 Rp 和第四参数 Rt。随后将详细描述根据第一参数 R 来设定第二参数 Rs、第三参数 Rp 和第四参数 Rt 的方法。参数调节部分 3301 将根据第一参数 R 确定的第四参数 Rt 发送到随后将描述的内容管理部分 3303,并且将第二参数 Rs 和第三参数 Rp 发送到随后将描述的信号处理部分 3307。

[0342] 内容管理部分 3303 例如被配置为 CPU、ROM、RAM 等,并且管理包括可以被根据本实施例的信息处理装置 3300 所回放的音频信号的内容。在随后将描述的内容存储部分 3305 中,内容管理部分 3303 例如与内容标题、内容 ID 和属性信息等相关联地来记录包括音频信号的内容。内容管理部分 3303 根据对从信息处理装置 3300 外部输入的内容的回放指令,从内容存储部分 3305 中获得内容,并且将所述内容输出到随后将描述的信号处理部分 3307。当将内容输出到信号处理部分 3307 时,基于从参数调节部分 3301 发送的第四参数 Rt 来确定待发送的数据量。此外,当从内容存储部分 3305 读取的内容数据是经编码的数据时,内容管理部分 3303 通过未示出的解码器对所述内容数据进行解码,并将所述内容数据输出到信号处理部分 3307。

[0343] 此外,内容管理部分 3303 可以经由诸如因特网和家用网络之类的网络 1702 来获得包括待回放的音频信号的内容。内容管理部分 3303 可以将经由网络 1702 而获得的内容记录在内容存储部分 3305 中。

[0344] 内容存储部分 3305 例如被配置为诸如硬盘驱动器、DVD 驱动器和蓝光驱动器之类

的记录介质,并且与内容的标题、ID、属性信息等相关联地存储包括音频信号的内容。此外,包括构成内容存储部分 3305 的各种记录介质的读取速度上限值的控制信息等可以作为数据库被存储在内容存储部分 3305 中。

[0345] 信号处理部分 3307 例如被配置为 CPU、ROM、RAM 等,并且基于从内容管理部分 3303 发送的音频信号、第一参数 R 以及从参数调节部分 3301 发送的第二参数 R_s 和第三参数 R_p ,来调节音频信号的语速和声音音高。此外,信号处理部分 3307 将语速和声音音高经过调节的音频信号作为输出音频信号输出。信息处理装置 3300 通过未示出的 DA 转换器将这样的输出音频信号转换为模拟信号,并且将所述模拟信号从诸如扬声器之类的输出设备输出。

[0346] 存储部分 3309 例如被配置为 RAM、存储设备等,并且存储在根据第一参数来确定第二参数 R_s 、第三参数 R_p 和第四参数 R_t 时所使用的各种数据库,将由信息处理装置 3300 执行的各种程序等。此外,除了这些数据外,存储部分 3309 还可以根据需要存储当信息处理装置 3300 执行处理、处理的中间进程等时需要被存储的各种参数。参数调节部分 3301、内容管理部分 3303、信号处理部分 3307 等可以在存储部分 3309 中自由地执行对数据的读取或写入。

[0347] (第一参数和第四参数之间的关系)

[0348] 随后,通过参考图 34A 和图 34B,来详细描述根据本实施例的通过参数调节部分 3301 来调节第四参数的方法。图 34A 是示出第一参数 R 和第四参数 R_t 之间关系的坐标图,并且图 34B 是示出第一参数 R 和将被输入到信号处理部分 3307 的音频信号数据量之间关系的坐标图。

[0349] 如图 34A 所示,其中横轴代表第一参数 R 并且纵轴代表第四参数 R_t 的坐标图被配置为具有第四参数 R_t 的不同上升率(也就是说,坐标图的坡度)的两个区域。

[0350] 参数调节部分 3301 在下述情况下调节第四参数 R_t 。在此,将在内容管理部分 3303 从内容存储部分 3305 读取内容数据并将所述内容数据发送到信号处理部分 3307 时的数据读取速度的上限简写为 S_m 。顺便提及,在以下描述中,数据读取速度是这样的速度,其包括内容管理部分 3303 从内容存储部分 3305 读取预定内容数据的数据读取速度和当将从内容管理部分 3303 读取的内容数据发送到信号处理部分 3307 时所需要的速度。

[0351] 条件 A:当所输入的第一参数 R 存在于区间 3405 中时,第四参数 R_t 恒定地为 1.0。

[0352] 条件 B:当所输入的第一参数 R 存在于区间 3406 中时,上限速度 $S_m = \text{第一参数 } R \times \text{第四参数 } R_t$ 成立。

[0353] 上限速度 S_m 是根据内容管理部分 3303 和内容存储部分 3305 的处理能力而确定的恒定值,因此,在区间 3406 中,随着第一参数 R 的值变大,第四参数 R_t 变小。

[0354] 图 34B 示出每单位时间输入到信号处理部分 3307 的音频信号量与数据读取速度上限 S_m 的比值。在区间 3407,数据量比值与第一参数 R 成比例。但是,在区间 3408,数据量比值恒定地为 1.0。这是因为数据读取速度根据第四参数 R_t 被调节,所以数据读取速度没有超过其上限 S_m 。因此,可以说第四参数 R_t 是在从内容存储部分 3305 读取内容数据以及将所述内容数据发送到信号处理部分 3307 时数据的间隔剔除率。

[0355] (根据第四参数对数据读取速度的调节)

[0356] 例如,通过如图 35A 到图 37C 所示的方法来执行根据第四参数对数据读取速度的调节。图 35A 到图 37C 是示出根据本实施例的调节数据读取速度的方法示例的说明性示

图。

[0357] 在如图 35A 和图 35B 所示的示例中,诸如区间 3501、区间 3502 和区间 3503 之类的原始信号的片段从在图 35A 中示出的、被记录在记录介质中的原始信号中被选出。在图 35B 中示出的信号代表所读取的信号,并且区间 3504、区间 3505 和区间 3506 分别对应于在图 35A 中示出的原始信号的区间 3501、区间 3502 和区间 3503。从内容存储部分 3305 中读取并被输出到信号处理部分 3307 的信号是由相连接的、在图 35B 中示出的信号的区间 3504、区间 3505 和区间 3506 构成的信号。在此,当连接每个区间时,每个区间的信号可以被淡入或淡出以进行平滑连接。此外,每个区间可以被略微拉长以通过交叉淡入淡出来进行连接。在图 35B 中示出的信号被信号处理部分 3307 处理以在变速回放时形成回放声音。

[0358] 在如图 35A 和图 35B 所示的示例中,关于在图 35A 中示出的原始信号,读取区间长度和跳过区间长度彼此相等(也就是,区间 3501 的长度和介于区间 3501 和区间 3502 之间的水平区段长度彼此相等),因此,第四参数 R_t 等于 $1/2$ 。另一方面,图 36A 和图 36B 示出这样的示例,其中第四参数 R_t 的值不同于如图 35A 和图 35B 所示的示例。在如图 36A 和图 36B 所示的示例中,关于在图 36A 中示出的原始信号,读取区间长度与跳过区间长度的比值是 $3:4$,因此第四参数 R_t 等于 $3/7$ 。

[0359] 图 37A 到图 37C 示出类似于如图 35A 到图 36B 所示的示例,但区别在于,记录在记录介质中的内容数据经过编码。在许多情况下,虽然随编解码器的不同,名称可能有变化,但是通常以一致的单位来管理经编码的数据。例如,就 MPEG 而言,以诸如包或数据包之类的单位 P 来管理经编码的数据。

[0360] 在如图 37A 到图 37C 所示的示例中,诸如区间 3701、区间 3702 和区间 3703 之类的流数据片段从在图 37A 中示出的、记录在记录介质中的流数据(经编码数据)中被读取。在图 37B 中示出的、所读取的流数据的区间 3704、区间 3705 和区间 3706 分别对应于在图 37A 中示出的流数据的区间 3701、区间 3702 和区间 3703。从如图 37B 所示的流数据中读取的区间 3704、区间 3705 和区间 3706 分别被解码器解码,从而变为在图 37C 中示出的音频信号的区间 3707、区间 3708 和区间 3709。在此,当连接每个区间时,每个区间的信号可以被淡入或淡出以进行平滑连接。此外,每个区间可以被略微拉长以通过交叉淡入淡出来进行连接。在图 37C 中示出的音频信号被信号处理部分 3307 处理以在变速回放时形成回放声音。

[0361] 在如图 37A 到图 37C 所示的示例中,关于在图 37A 中示出的流数据,读取区间长度和跳过区间长度彼此相等,因此,第四参数 R_t 等于 $1/2$ 。然而,在经编码信号的情况下,每个管理单位 P 在进行编码前可能在音频数据中具有交叠区间。在这种情况下,可能必须根据交叠区间来读取在图 37A 中示出的流数据中的额外读取区间。此外,根据编解码器,管理信息被添加到每个管理单位,并且管理信息必须被读取以读取下一个管理单位。在这种情况下,即使在跳过区间,也至少必须读取管理信息。因此,当处理流数据时,虽然必须添加取决于编解码器的处理,但是基本处理与在图 35A 到图 36B 中示出的处理是相同的。

[0362] 在以下描述中,对应于其中第四参数 R_t 是 1.0 的区间的第一参数 R 的范围(例如如图 34A 中的区间 3405)被称作第三范围,并且对应于其中第四参数 R_t 受上限速度 S_m 影响的区间的第一参数 R 的范围(例如如图 34A 中的区间 3406)被称作第四范围。

[0363] (第一参数与第二参数以及与第三参数的关系)

[0364] 图 38A 和图 38B 详细描述根据本实施例的参数调节部分 3301 的参数调节方法示例。图 38A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图,并且图 38B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。

[0365] 在根据本实施例的信息处理装置 3300 中,如图 38A 和图 38B 所示的、示出第一参数 R 与第二参数 Rs 以及与第三参数 Rp 关系的数据库,和如图 34A 中所示的、示出第一参数 R 和第四参数 Rt 之间关系的数据库例如被存储在存储部分 3309 中,并且参数调节部分 3301 通过参考这些数据库,根据第一参数 R 来确定第二参数 Rs、第三参数 Rp 和第四参数 Rt。

[0366] 在此,参数调节部分 3301 在下述的四个条件下,通过参考在存储部分 3309 中存储的、如图 38A 和图 38B 所示的数据库,根据第一参数 R 来确定第二参数 Rs 和第三参数 Rp。

[0367] 条件 1 :当所输入的第一参数 R 存在于区间 3801 中时,将第二参数 Rs 确定为与第一参数 R 成比例(换言之,确定第二参数 Rs 以使得第二参数 Rs 等于第一参数 R)。

[0368] 条件 2 :当所输入的第一参数 R 存在于区间 3803 中时,将第三参数 Rp 恒定地设定为 1。

[0369] 条件 3 :当所输入的第一参数 R 存在于区间 3804 中时,第三参数 Rp 随第一参数 R 的增大而增大。

[0370] 条件 4 :第一参数 R $\times$ 第四参数 Rt = 第二参数 Rs $\times$ 样本数的增长率 Rd。

[0371] 在此,在图 38A 的区间 3809 中,因为受上述条件 B 的影响所以第二参数 Rs 减小。顺便提及,从图 38A 和图 38B 中显然可见,Rt 影响第二参数 Rs,但是不影响第三参数 Rp。换言之,当被发送到信号处理部分 3307 的音频信号的数据量减小时,数据量的减小影响语速转换的程度,然而不影响对声音音高的调节。

[0372] 此外,区间 3801 和区间 3803 对应于第一参数 R 的第一范围,并且区间 3802,区间 3809 和区间 3804 对应于第一参数 R 的第二范围。此外,区间 3801 和区间 3802 对应于第一参数 R 的第三范围,并且区间 3809 对应于第一参数 R 的第四范围。

[0373] 在如图 38A 和图 38B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0374] 此外,当第一参数 R 是 1 到 20 时,即当以 1 到 20 倍的速度回放时,信号被连续读取,并且当第一参数 R 大于 20 时,即当以大于 20 倍的速度回放时,信号被断续读取。通过执行这样的处理,可以实现超过 20 倍速度的回放速度,20 倍的速度被认为是连续读取信号情况下的回放上限。

[0375] 顺便提及,在图 38A 中,周期 3802 和周期 3809 以虚线被示出,因为第二参数 Rs 的值随改变声音音高的方法而改变。当将如图 12 到图 14 所示的方法用作改变声音音高的方法时,样本数随声音音高的升高而减少,因此,用虚线示出区间 3802 和区间 3809 的线条。然而,当将样本数不减少的方法或者减少量很小的方法用作改变声音音高的方法时,将不同于如图 38A 所示的虚线来设定区间 3802 和区间 3809。

[0376] 此外,当方法中用于改变声音音高的样本数的增长率是 Rd 时,参数调节部分 3301 具有如上述条件 4 所指示的特性。在此,例如,当样本数是 2 倍时,增长率是 2,并且当样本

数被减少到一半时,增长率是 1/2。

[0377] (根据本实施例的控制回放速度变量因子的方法)

[0378] 图 39 是示出根据本实施例的信息处理装置 3300 的处理流程的流程图。首先,信息处理装置 3300 判断是否存在输入音频信号(步骤 S3901),并且当不存在输入音频信号时,处理终止。此外,当输入音频信号存在时,信息处理装置 3300 的参数调节部分 3301 根据所输入的第一参数 R 来调节第二参数 Rs、第三参数 Rp 和第四参数 Rt(步骤 S3902)。以满足上述条件 1 到 4 以及条件 A 和 B 的方式来执行调节。随后,信息处理装置 3300 的信号处理部分 3307 根据经调节的第二参数 Rs 和第三参数 Rp 来调节从内容管理部分 3303 发送来的音频信号的语速和声音音高(步骤 S3903)。随后,信息处理装置 3300 输出语速和声音音高经过调节的音频信号(步骤 S3904)。然后,返回到步骤 S3901,上述处理被重复。

[0379] 通过重复这样的处理,根据本实施例的信息处理装置 3300 被使得能够控制音频信号的回放速度变量因子。

[0380] 如参考图 33 到图 39 所描述的,根据本实施例的控制回放速度变量因子的方法,能够在第一参数 R 的第一范围内仅调节语速,并且在第一参数 R 的第二范围内随同语速一起来调节声音音高。因此,在第一参数 R 的第一范围内解决了第一问题,并且在第一参数 R 的第二范围内解决了第二问题。此外,在第一参数 R 的第三范围内,信号可以被连续读取,并且在第一参数 R 的第四范围内,信号可以被断续读取。因此,第三问题可以在第四范围内被补救,并且第四范围可以扩展进而回放速度的上限可以提高。

[0381] (信号处理部分 3307)

[0382] 随后,通过参考图 40,来详细描述根据本实施例的信号处理部分 3307 的示例。图 40 是示出根据本实施例的信号处理部分 3307 的功能的框图。

[0383] 如图 40 所示,例如,根据本实施例的信号处理部分 3307 主要包括拟声声音切换判断部分 4001、语速转换部分 4003、音高调节部分 4005 和音频信号输出控制部分 4007。

[0384] 根据本实施例的拟声声音切换判断部分 4001、语速转换部分 4003、音高调节部分 4005 和音频信号输出控制部分 4007 分别与根据本发明第一实施例的拟声声音切换判断部分 2101、语速转换部分 2103、音高调节部分 2105 和音频信号输出控制部分 2107 具有几乎相同的配置,并且实现类似的效果,因此,省略其详细描述。

[0385] 图 41A 和图 41B 是示出由具有如图 40 所示的信号处理部分 3307 的信息处理装置 3300 的参数调节部分 3301 所执行的参数调节方法示例的说明性示图。

[0386] 参数调节部分 3301 包括上述的条件 A 和条件 B 二者。图 41A 是示出第一参数 R 和第二参数 Rs 之间关系的坐标图,并且图 41B 是示出第一参数 R 和第三参数 Rp 之间关系的坐标图。

[0387] 如图 41A 所示,横轴代表第一参数 R 并且纵轴代表第二参数 Rs 的坐标图被配置为具有第二参数 Rs 的不同上升率(换言之,坐标图的坡度)的多于三个的区域。类似地,如图 41B 所示,横轴代表第一参数 R 并且纵轴代表第三参数 Rp 的坐标图被配置为具有第三参数 Rp 的不同上升率的至少两个区域。

[0388] 当信号处理部分 3307 的音高调节部分 4005 通过如图 12 到图 14C 所示的方法来调节音高时,参数调节部分 3301 在下述的四个条件下,通过参考在存储部分 3309 中存储的、如图 41A 和图 41B 所示的数据库,根据第一参数 R 来确定第二参数 Rs 和第三参数 Rp。

[0389] 条件 1:当所输入的第一参数 R 存在于区间 4101 中时,将第二参数 Rs 确定为与第一参数 R 成比例(换言之,确定第二参数 Rs 以使得第二参数 Rs 等于第一参数 R)。

[0390] 条件 2:当所输入的第一参数 R 存在于区间 4103 中时,将第三参数 Rp 恒定地设定为 1。

[0391] 条件 3:当所输入的第一参数 R 存在于区间 4104 中时,第三参数 Rp 随第一参数 R 的增大而增大。

[0392] 条件 4':第一参数 $R \times$ 第四参数 Rt = 第二参数 Rs $\times$ 第三参数 Rp 在第一范围和第二范围(第三范围和第四范围)内成立。

[0393] 在此,在区间 4109 中,因为受上述条件 B 的影响所以第二参数 Rs 减小。顺便提及,如从图 41A 和图 41B 中显然可见的,第四参数 Rt 影响第二参数 Rs,但是不影响第三参数 Rp。换言之,当被发送到信号处理部分 3307 的音频信号的数据量减小时,数据量的减小影响语速转换的程度,但是不影响对声音音高的调节。

[0394] 此外,区间 4101 和区间 4103 对应于第一参数 R 的第一范围,并且区间 4102、区间 4109 和区间 4104 对应于第一参数 R 的第二范围。此外,区间 4101 和区间 4102 对应于第一参数 R 的第三范围,并且区间 4109 对应于第一参数 R 的第四范围。

[0395] 在如图 41A 和图 41B 所示的示例中,当第一参数 R 是 1 到 4 时,即当以 1 到 4 倍的速度回放时,仅执行语速转换,并且当第一参数 R 大于 4 时,即当以大于 4 倍的速度进行回放时,随同对语速进行转换一起来升高声音音高。通过执行这样的处理,当以 1 到 4 倍的速度进行回放时,说话人的语速根据回放速度逐渐加快,并且当以大于 4 倍的速度进行回放时,在说话人语速加快的同时逐渐升高声音音高。

[0396] 此外,当第一参数 R 是 1 到 20 时,即当以 1 到 20 倍的速度回放时,信号被连续读取,并且当第一参数 R 大于 20 时,即当以大于 20 倍的速度回放时,信号被断续读取。通过执行这样的处理,可以实现超过 20 倍速度的回放速度,20 倍的速度被认为是当没有执行间隔(thinned)回放时的回放上限。

[0397] 至此,已经描述了根据本实施例的信息处理装置 3300 的功能示例。上述结构元件的每个可以被配置为通用组件或电路,或者可以被配置为专用于每个结构元件的功能的硬件。此外,CPU 等可以执行所有的功能。因此,能够根据实现本实施例的各种技术水平来适当改变待使用的配置。

[0398] (根据本实施例的信号处理方法)

[0399] 随后,通过参考图 42,来详细描述根据本实施例的信号处理方法。图 42 是示出根据本实施例的信号处理方法的流程图。

[0400] 首先,信息处理装置 3300 的信号处理部分 3307 判断是否存在从内容管理部分 3303 发送的音频信号(步骤 S4201),并且如果没有从内容管理部分 3303 发送的音频信号那么终止处理。此外,当从内容管理部分 3303 发送的音频信号存在时,信号处理部分 3307 的拟声声音切换判断部分 4001 判断所输入的第一参数 R 是否大于预定阈值(步骤 S4202)。当第一参数 R 小于预定阈值时,参数调节部分 3301 根据所输入的第一参数 R 来调节第二参数 Rs、第三参数 Rp 和第四参数 Rt(步骤 S4203),并且将参数发送到信号处理部分 3307。信号处理部分 3307 的语速转换部分 4003 基于所发送的第二参数 Rs 来调节输入音频信号的语速(步骤 S4204),并且将语速经过调节的音频信号输出到音高调节部分 4005。音高调节

部分 4005 基于所发送的第三参数 R_p 来调节从语速转换部分 4003 发送的音频信号的声音音高（步骤 S4205）。语速和声音音高经过调节的音频信号被发送到音频信号输出控制部分 4007，并且音频信号输出控制部分 4007 输出语速和声音音高经过调节的音频信号（步骤 S4206）。然后，返回到步骤 S4201，上述处理被重复。

[0401] 另一方面，当拟声声音切换判断部分 4001 判断第一参数 R 大于预定阈值时，音频信号输出控制部分 4007 将被存储在存储部分 3309 等中的预定拟声声音作为音频信号输出（步骤 S4207）。然后，返回到步骤 S4201，上述处理被重复。

[0402] 通过重复这样的处理，根据本实施例的信息处理装置 3300 被使得能够以在听觉上可以识别转换后的回放速度的方式来控制音频信号的回放速度变量因子。

[0403] （第二实施例的第一修改示例）

[0404] 随后，通过参考图 43，来详细描述根据本发明第二实施例的第一修改示例的信息处理装置 4300 的配置。图 43 是示出根据本修改实施例的信息处理装置 4300 的功能的框图。

[0405] 如图 43 所示的修改示例是其中内容管理部分 4303 设定第四参数 R_t 的示例。例如，当根据本修改示例的信息处理装置 4300 被用作视频记录 / 回放装置时，存在这样的情况，其中对内容的回放和对另一节目的视频记录同时被执行。在这样的情况下，视频记录 / 回放装置必须同时执行回放和记录，并且与仅执行回放的情况相比，可以分配给回放处理的处理量减少。因此，因为回放处理的处理量可能随情况的不同而改变，所以应该根据能够备用于回放处理的处理量来确定间隔率（thinning rate）。根据本修改示例的信息处理装置 4300 通过包括如下所述的内容管理部分 4303 来实现这样的处理。

[0406] 如图 43 所示，根据本修改示例的信息处理装置 4300 例如主要包括参数调节部分 4301、内容管理部分 4303、内容存储部分 4305、信号处理部分 4307 和存储部分 4309。

[0407] 在此，内容存储部分 4305、信号处理部分 4307 和存储部分 4309 分别与根据本发明第二实施例的内容存储部分 3305、信号处理部分 3307 和存储部分 3309 具有几乎相同的配置，并且实现类似的效果，因此省略其详细描述。

[0408] 参数调节部分 4301 例如被配置为 CPU、ROM、RAM 等，并且根据从外部输入的第一参数 R 和从随后将描述的内容管理部分 4303 发送的第四参数 R_t ，来调节第二参数 R_s 和第三参数 R_p 。如在本发明第二实施例中所描述的，通过参考在存储部分 4309 中所存储的、示出第一参数 R 与第二参数 R_s 以及与第三参数 R_p 的关系的数据库，来确定第二参数 R_s 和第三参数 R_p 的设定以满足如在第二实施例中所描述的条件。参数调节部分 4301 将所确定的第二参数 R_s 和第三参数 R_p 发送到信号处理部分 4307。

[0409] 内容管理部分 4303 例如被配置为 CPU、ROM、RAM 等，并且管理包括可以被根据本实施例的信息处理装置 4300 所回放的音频信号的内容。在内容存储部分 4305 中，内容管理部分 4303 例如与内容标题、内容 ID 和属性信息等相关联地来记录包括音频信号的内容。内容管理部分 4303 根据对从信息处理装置 4300 外部输入的内容的回放指令，从内容存储部分 4305 中获得内容，并且将所述内容输出到信号处理部分 4307。当将内容输出到信号处理部分 4307 时，内容管理部分 4303 根据可用于内容输出的资源量来确定对应于数据间隔率的第四参数 R_t ，并且根据第四参数 R_t 来确定待发送的数据量。此外，内容管理部分 4303 将所确定的第四参数 R_t 发送到参数调节部分 4301。顺便提及，当从内容存储部分 4305 读取

的内容数据是经编码的数据时,内容管理部分 4303 通过未示出的解码器对数据进行解码,并将所述数据输出到信号处理部分 4307。

[0410] 此外,内容管理部分 4303 可以经由诸如因特网和家用网络之类的网络 1702 来获得包括待回放的音频信号的内容。内容管理部分 4303 可以将经由网络 1702 而获得的内容记录在内容存储部分 4305 中。

[0411] 内容存储部分 4305 例如被配置为诸如硬盘驱动器、DVD 驱动器和蓝光驱动器之类的记录介质,并且与内容的标题、ID、属性信息等相关联地存储包括音频信号的内容。此外,包括构成内容存储部分 4305 的各种记录介质的读取速度上限值的控制信息等可以作为数据库被存储在内容存储部分 4305 中。

[0412] 至此,已经描述了根据本修改示例的信息处理装置 4300 的功能示例。上述结构元件的每个可以被配置为通用组件或电路,或者可以被配置为专用于每个结构元件的功能的硬件。此外,CPU 等可以执行所有的功能。因此,能够根据实现本修改示例的各种技术水平来适当改变待使用的配置。

[0413] (根据本修改示例的信号处理方法)

[0414] 随后,通过参考图 44,来详细描述根据本修改示例的信号处理方法。图 44 是示出根据本修改示例的信号处理方法的流程图。

[0415] 首先,信息处理装置 4300 的信号处理部分 4307 判断是否存在从内容管理部分 4303 发送的音频信号(步骤 S4401),并且当没有从内容管理部分 4303 发送的音频信号终止处理。此外,当从内容管理部分 4303 发送的音频信号存在时,信号处理部分 4307 的拟声声音切换判断部分判断所输入的第一参数 R 是否大于预定阈值(步骤 S4402)。当第一参数 R 小于预定阈值时,参数调节部分 4301 根据所输入的第一参数 R 和从内容管理部分 4303 发送的第四参数 Rt 来调节第二参数 Rs 和第三参数 Rp(步骤 S4403),并且将参数发送到信号处理部分 4307。信号处理部分 4307 基于所发送的第二参数 Rs 和第三参数 Rp 来调节输入音频信号的语速和声音音高(步骤 S4404)。语速和声音音高经过调节的音频信号被发送到音频信号输出控制部分,并且音频信号输出控制部分输出语速和声音音高经过调节的音频信号(步骤 S4405)。然后,返回到步骤 S4401,上述处理被重复。

[0416] 另一方面,当拟声声音切换判断部分判断第一参数 R 大于预定阈值时,音频信号输出控制部分将被存储在存储部分 4309 等中的预定拟声声音作为音频信号输出(步骤 S4406)。然后,返回到步骤 S4401,上述处理被重复。

[0417] 通过重复这样的处理,根据本修改示例的信息处理装置 4300 被使得能够以在听觉上可以识别转换后的回放速度的方式来控制音频信号的回放速度变量因子。

[0418] (信号处理部分 3307、4307 的修改示例)

[0419] 随后,通过参考图 45,来描述根据本实施例的以及根据本修改示例的信号处理部分 3307、4307 的修改示例。图 45 是示出信号处理部分 3307、4307 的修改示例的框图。

[0420] 如图 45 所示,根据本修改示例的信号处理部分主要包括拟声声音切换判断部分 4001、音高调节部分 4501、语速转换部分 4503 和音频信号输出控制部分 4007。

[0421] 根据本修改示例的拟声声音切换判断部分 4001、音高调节部分 4501、语速转换部分 4503 和音频信号输出控制部分 4007 分别与根据本发明第一实施例的第一修改示例的拟声声音切换判断部分 2101、音高调节部分 2901、语速转换部分 2903 和音频信号输出控制部

分 2107 具有几乎相同的配置,并且实现类似的效果,因此,省略其详细描述。

[0422] (根据本修改示例的信号处理方法)

[0423] 随后,通过参考图 46,来详细描述根据本修改示例的信号处理方法。图 46 是示出根据本修改示例的信号处理方法的流程图。

[0424] 首先,信息处理装置 4300 判断是否存在输入音频信号(步骤 S4601),并且当没有输入音频信号时终止处理。此外,当输入音频信号存在时,信号处理部分 4307 的拟声声音切换判断部分 4001 判断所输入的第一参数 R 是否大于预定阈值(步骤 S4602)。当第一参数 R 小于预定阈值时,参数调节部分 4301 根据所输入的第一参数 R 和从内容管理部分 4303 发送的第四参数 R_t 来调节第二参数 R_s 和第三参数 R_p (步骤 S4603),并且将参数发送到信号处理部分 4307。信号处理部分 4307 的音高调节部分 4501 基于所发送的第三参数 R_p 来调节所发送的输入音频信号的声音音高(步骤 S4604),并且将声音音高经过调节的音频信号输出到语速转换部分 4503。语速转换部分 4503 基于所发送的第二参数 R_s 来调节声音音高经过调节的音频信号的语速(步骤 S4605)。语速和声音音高经过调节的音频信号被发送到音频信号输出控制部分 4007,并且音频信号输出控制部分 4007 输出语速和声音音高经过调节的音频信号(步骤 S4606)。然后,返回到步骤 S4601,上述处理被重复。

[0425] 另一方面,当拟声声音切换判断部分 4001 判断第一参数 R 大于预定阈值时,音频信号输出控制部分 4007 将被存储在存储部分 3309 等中的预定拟声声音作为音频信号输出(步骤 S4607)。然后,返回到步骤 S4601,上述处理被重复。

[0426] 通过重复这样的处理,根据本修改示例的信息处理装置 4300 被使得能够以在听觉上可以识别转换后的回放速度的方式来控制音频信号的回放速度变量因子。

[0427] 如上所述,通过根据第二实施例和根据本发明的每个修改示例的信息处理装置,能够确定音频信号的语速转换率和声音音高转换率,同时通过在发送音频信号时进行间隔剔除来实现构成音频数据的样本数的减少。通过使用这样的装置,当以近似正常速度进行回放时,回放速度改变然而声音音高不变,因而很容易理解说话人的语音内容或者很容易指定说话人。同时,在高速/低速回放的情况下,当转换回放速度时还改变声音音高,因此可以从听觉上感受到此时的回放速度,并且此外,通过调节例如连续读取和断续读取,在高速回放时的回放速度上限可以显著提高。因此,通过根据本实施例的信息处理装置,可操作性能够得到改进。

[0428] (信息处理装置的硬件配置)

[0429] 随后,通过参考图 47,来详细描述根据本发明每个实施例的信息处理装置的硬件配置。图 47 是示出根据本发明每个实施例的信息处理装置的硬件配置的框图。

[0430] 信息处理装置 1800、3300 和 4300 主要包括 CPU4701、ROM4703、RAM4705、主机总线 4707、电桥(bridge)4709、外部总线 4711、接口 4713、输入设备 4715、输出设备 4717、存储设备 4719、驱动器 4721、连接端口 4723 和通信设备 4725。

[0431] CPU4701 用作算术处理设备和控制设备,并且根据在 ROM4703、RAM4705、存储设备 4719 或可移动记录介质 4727 中存储的各种程序,来控制信息处理装置 1800、3300 和 4300 的整个操作或部分操作。RAM4705 临时存储将在 CPU4701 的执行期间所使用的程序、在执行期间根据需要而变化的参数等。这些组件通过主机总线 4707 来彼此连接,通过诸如 CPU 总线之类的内部总线来配置主机总线 4707。

[0432] 主机总线 4707 经由电桥 4709 连接到诸如 PCI (外围组件互连 / 接口) 总线之类的外部总线 4711。

[0433] 输入设备 4715 例如是将由使用者来操作的诸如鼠标、键盘、触摸面板、按钮、开关和控制杆之类的操作装置。此外,输入设备 4715 可以是利用红外线或其他无线电波的遥控装置 (所谓的遥控器), 或者它可以是诸如与信息处理装置 1800、3300 和 4300 的操作相关联的蜂窝式电话、PDA 等之类的外部连接装置 4729。此外,输入设备 4715 基于由使用者例如利用上述操作装置而输入的信息来生成输入信号。信息处理装置 1800、3300 和 4300 的使用者可以将各种数据输入到信息处理装置 1800、3300 和 4300, 或者可以通过在输入设备 4715 上进行操作来指示处理操作。

[0434] 通过能够在视觉上或者在听觉上将所获得的信息通知给使用者的设备来配置输出设备 4717, 例如诸如 CRT 显示器、液晶显示器、等离子显示器、EL 显示器和灯之类的显示设备, 诸如扬声器和头戴式耳机之类的音频输出设备, 打印设备, 蜂窝式电话, 或者传真。4717 例如输出通过由信息处理装置 1800、3300 和 4300 所执行的各种处理而获得的结果。具体地, 显示设备将通过由信息处理装置 1800、3300 和 4300 所执行的各种处理而获得的结果显示为文本或图像。另一方面, 音频输出设备将所回放的音频数据的音频信号、声响数据等转换为模拟信号并且将所述模拟信号输出。

[0435] 存储设备 4719 是作为信息处理装置 1800、3300 和 4300 的存储部分而配置的用于存储数据的设备, 并且例如被配置为诸如 HDD (硬盘驱动器) 之类的磁存储设备、半导体存储设备、光存储设备和磁光存储设备。存储设备 4719 存储将由 CPU4701 执行的程序以及各种数据, 从外部获得的声响信号数据和图像信号数据等。

[0436] 驱动器 4721 是与记录介质结合使用的读取器 / 写入器, 并且被嵌入信息处理装置 1800、3300 和 4300 中或者被配置为外围驱动器。驱动器 4721 读取在装载于其中的诸如磁盘、光盘、磁光盘或半导体存储器之类的可移动记录介质 4727 中所记录的信息, 并且将所述信息输出到 RAM4705。此外, 驱动器 4721 可以将记录写入装载于其中的诸如磁盘、光盘、磁光盘或半导体存储器之类的可移动记录介质 4727 中。可移动记录介质 4727 是 DVD 介质、HD-DVD 介质、蓝光介质、紧凑式闪存 (CF) (注册商标)、存储棒、SD (安全数字) 存储卡等。此外, 可移动记录介质 4727 例如可以是其中嵌入了非接触型 IC 芯片的 IC 卡 (集成电路卡) 或者电子设备。

[0437] 连接端口 4723 是这样的端口, 例如 USB (通用串行总线) 端口、诸如 i.Link 之类的 IEEE1394 端口、SCSI (小型计算机系统接口) 端口、RS-232C 端口、光音频终端和用于直接将设备连接到信息处理装置 1800、3300 和 4300 的 HDMI (高清多媒体接口) 端口。通过将外部连接装置 4729 连接到连接端口 4723, 信息处理装置 1800、3300 和 4300 直接从外部连接装置 4729 获得声响信号数据或图像信号数据, 或者为外部连接装置 4729 提供声响信号数据或图像信号数据。

[0438] 通信设备 4725 例如是被配置为连接到网络 1702 的通信设备等的通信接口。通信设备 4725 例如是用于有线或无线 LAN (局域网) 的通信卡, 蓝牙或 WUSB (无线 USB), 用于光通信的路由器, 用于 ADSL (不对称数字用户线) 的路由器, 或者用于各种通信的调制解调器。通信设备 4725 例如可以向 / 从因特网和其他通信设备发送 / 接收声响信号等。此外, 将被连接到通信设备 4725 的网络 1702 被配置为以有线或无线的方式连接的网络等, 并且

它可以是因特网、家用 LAN、红外通信、无线电波通信、卫星通信等。

[0439] 通过如上所述的配置,信息处理装置 1800、3300 和 4300 可以从各种信息资源获得与声响信号等有关的信息,并且可以将与声响信号等有关的信息发送到与连接端口 4723 或网络 1702 相连接的外部连接装置 4729、内容服务器 1703 和客户端装置 1704,并且此外,信息处理装置 1800、3300 和 4300 可以从外部连接装置 4729、内容服务器 1703 和客户端装置 1704 接收与声响信号有关的信息,并且获得外部连接装置 4729、内容服务器 1703 和客户端装置 1704 中的与声响信号有关的信息。此外,信息处理装置 1800、3300 和 4300 可以通过使用可移动记录介质 4727 来取出与声响信号等有关的信息。

[0440] 至此,已经描述了可以实现根据本发明每个实施例的信息处理装置 1800、3300 和 4300 的功能的硬件配置示例。上述结构元件的每个可以被配置为通用组件,或者可以被配置为专用于每个结构元件的功能的硬件。因此,能够根据实现本实施例的各种技术水平来适当改变待使用的配置。

[0441] 本领域中的技术人员应理解,根据设计需求和其他因素,可以想到各种修改、组合、子组合和变更,只要它们落入随附权利要求书或其等价物的范围内。

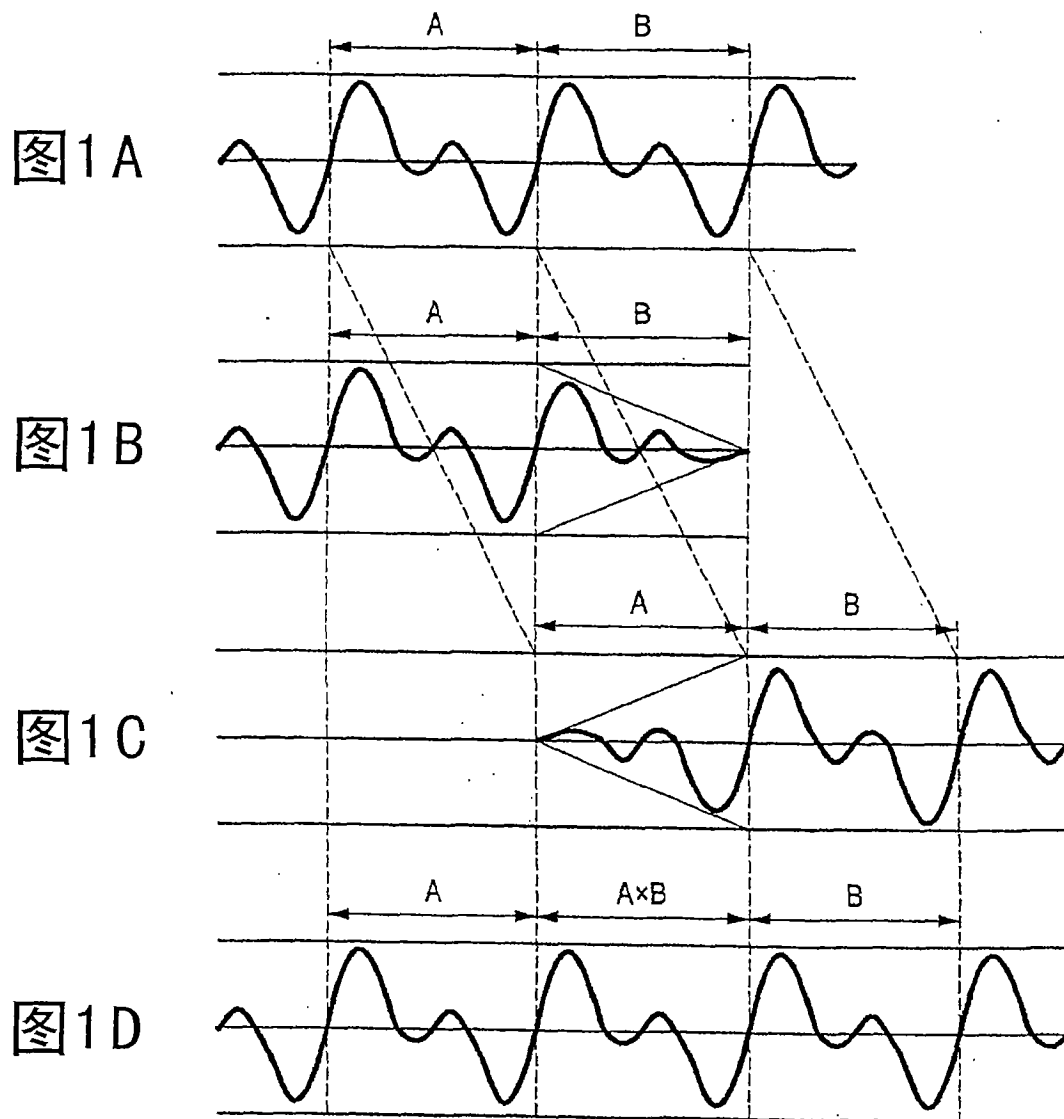
[0442] 例如,在上述的每个实施例中,对这样的情况进行了说明,其中在第一范围内,第一参数 R 是 1 到 4。然而,第一范围并不局限于此,并且第一参数可以具有不同的值。例如,在慢节奏的语音和音乐中,第一参数 R 的第一范围可以是 1 到 6 左右。相反,在快节奏的语音和音乐的情况下,其可以是 1 到 2 左右。

[0443] 此外,在如上所述的第二实施例中,对这样的情况进行了说明,其中 在第三范围内,第一参数 R 是 1 到 20。然而,第三范围并不局限于此,并且它可以具有不同的值。

[0444] 此外,在上述的每个实施例中,PICOLA 被用作语速转换的算法。然而,本发明的语速转换算法并不局限于此,并且可以使用任意算法而无论时间轴和频率轴如何,只要语速转换可以被执行。

[0445] 顺便提及,在上述的每个实施例中,对回放速度比正常速度更快的变速回放示例进行了说明,然而相同的事情可以被视为以小于正常速度的速度进行回放的情况。也就是,例如,0.5 到 1.0 倍的速度对应于第一范围,并且 0.0 到 0.5 倍的速度对应于第二范围。在 0.5 到 1.0 倍速度的范围内仅转换语速,并且在 0.0 到 0.5 倍速度的范围内,在转换语速的同时随着回放速度的减缓来降低声音音高,这是可能的。

[0446] 本发明包含与 2007 年 9 月 19 日递交到日本专利局的日本专利申请 JP2007-241681 相关的主题,该日本专利申请的全部内容通过引用被结合于此。



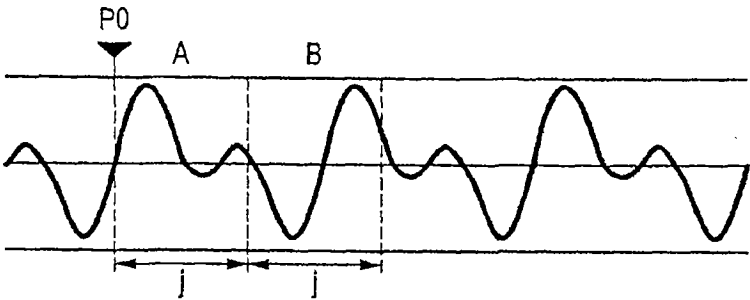


图 2A

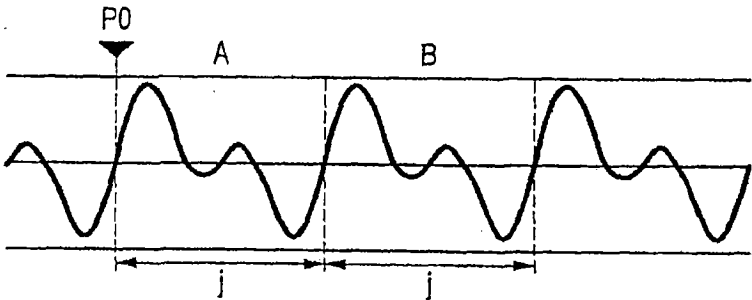


图 2B

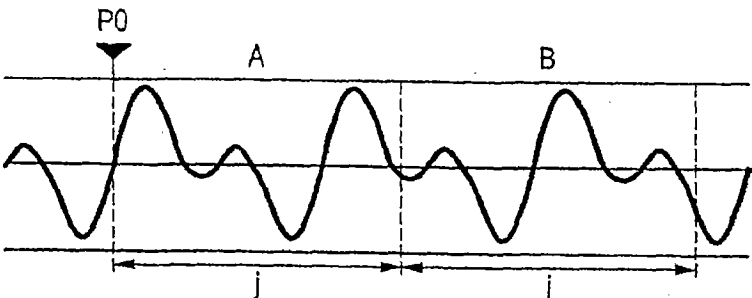


图 2C

图3A

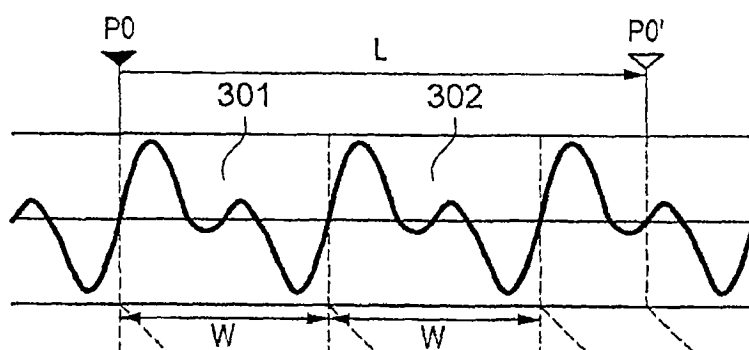
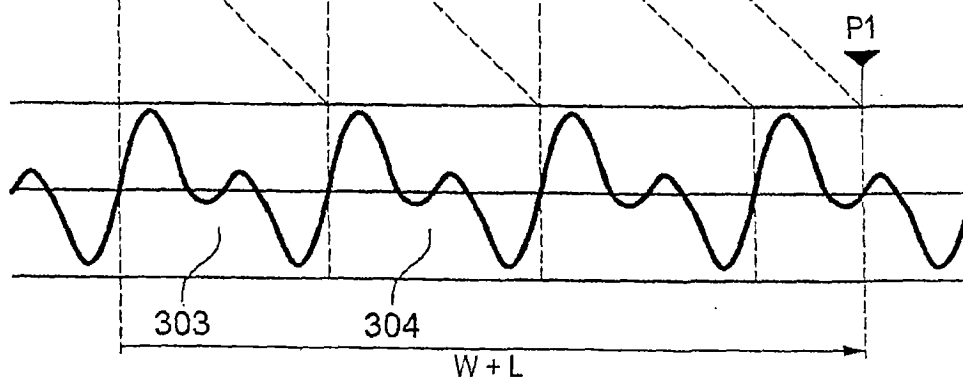


图3B



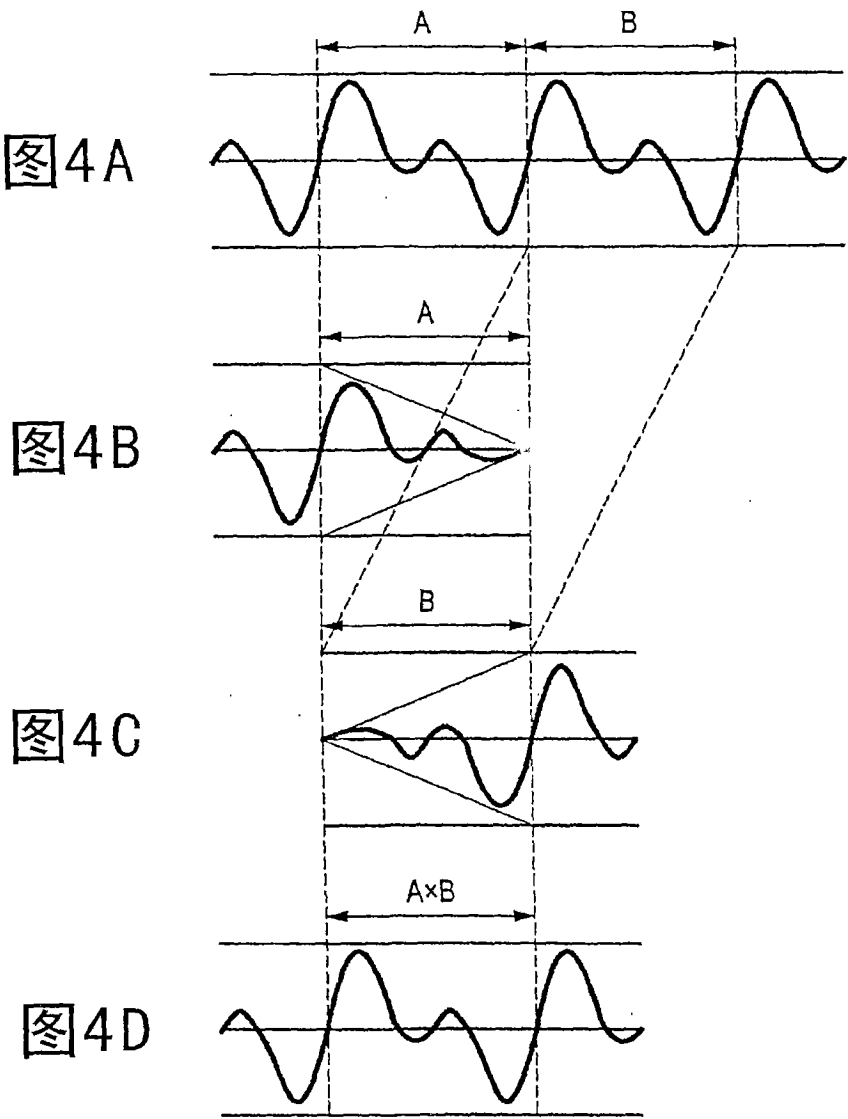


图5A

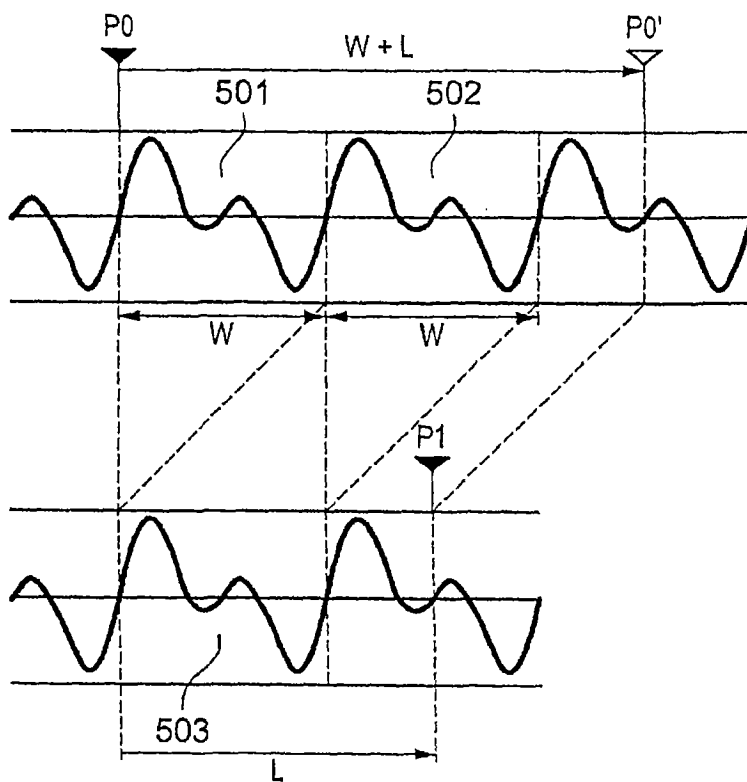
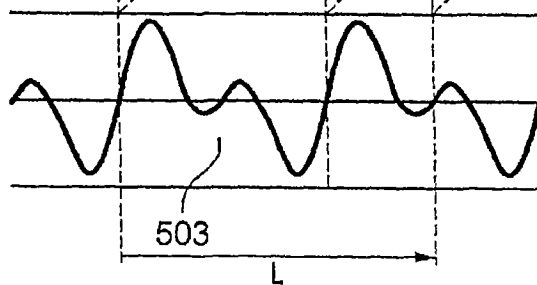


图5B



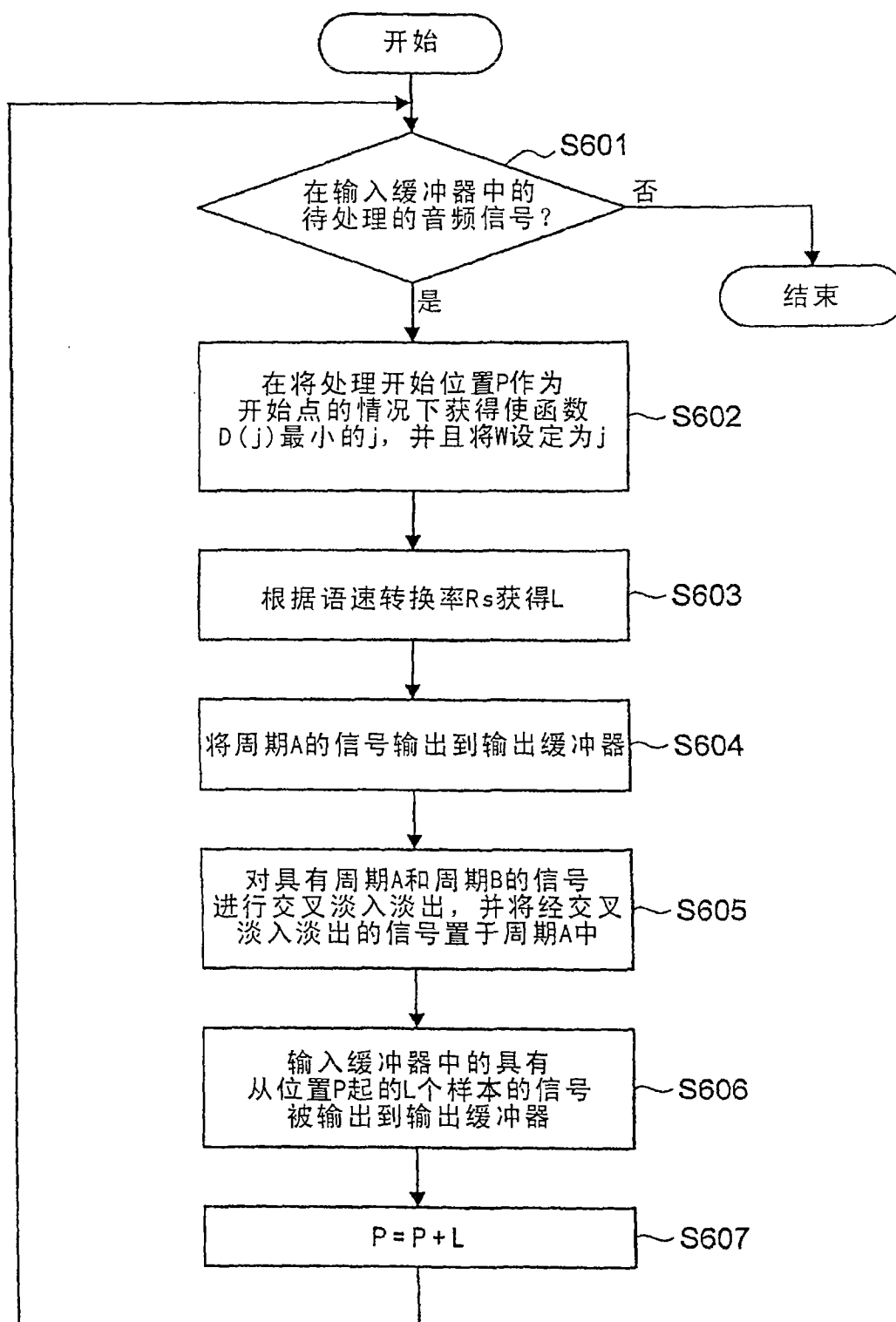


图 6

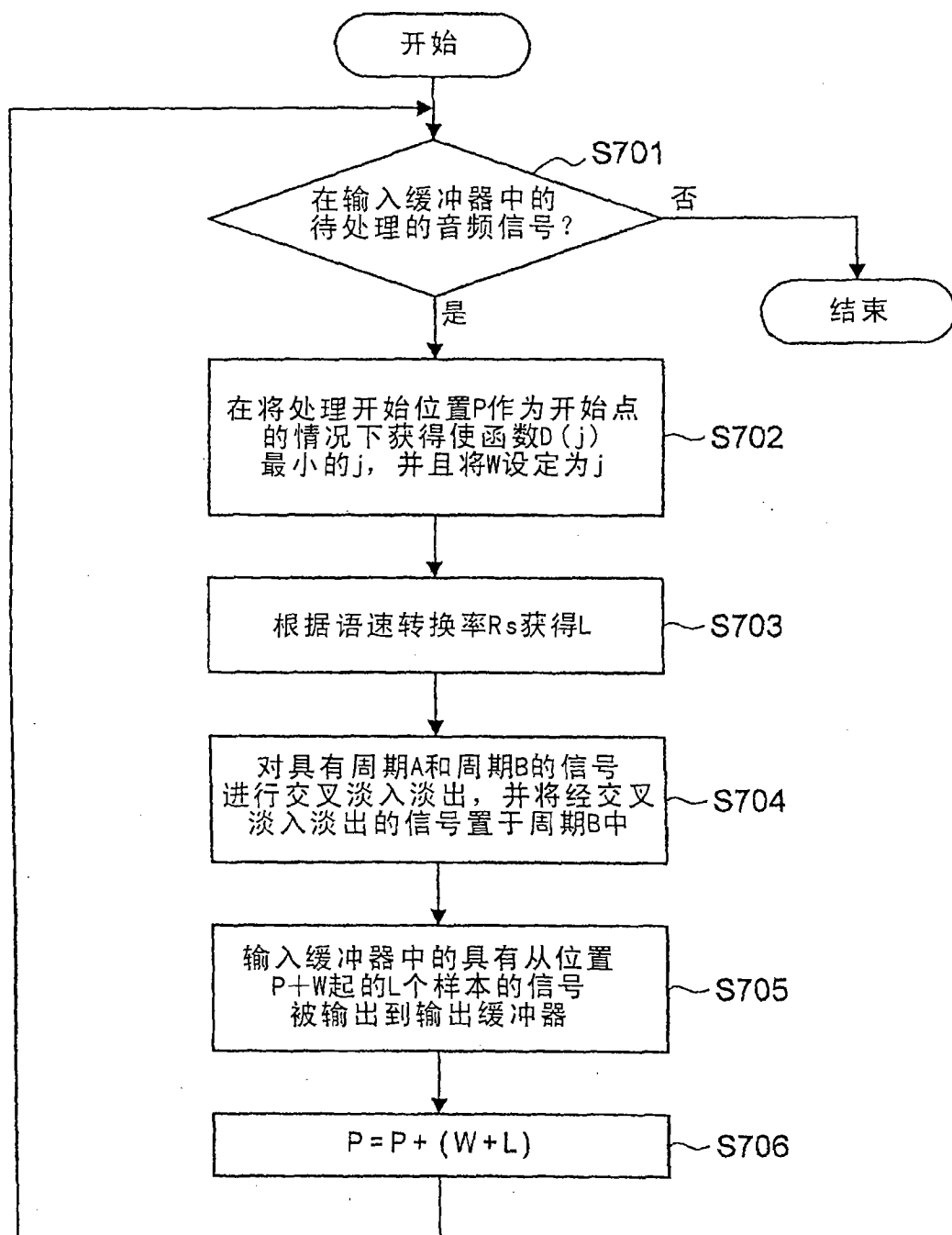


图 7

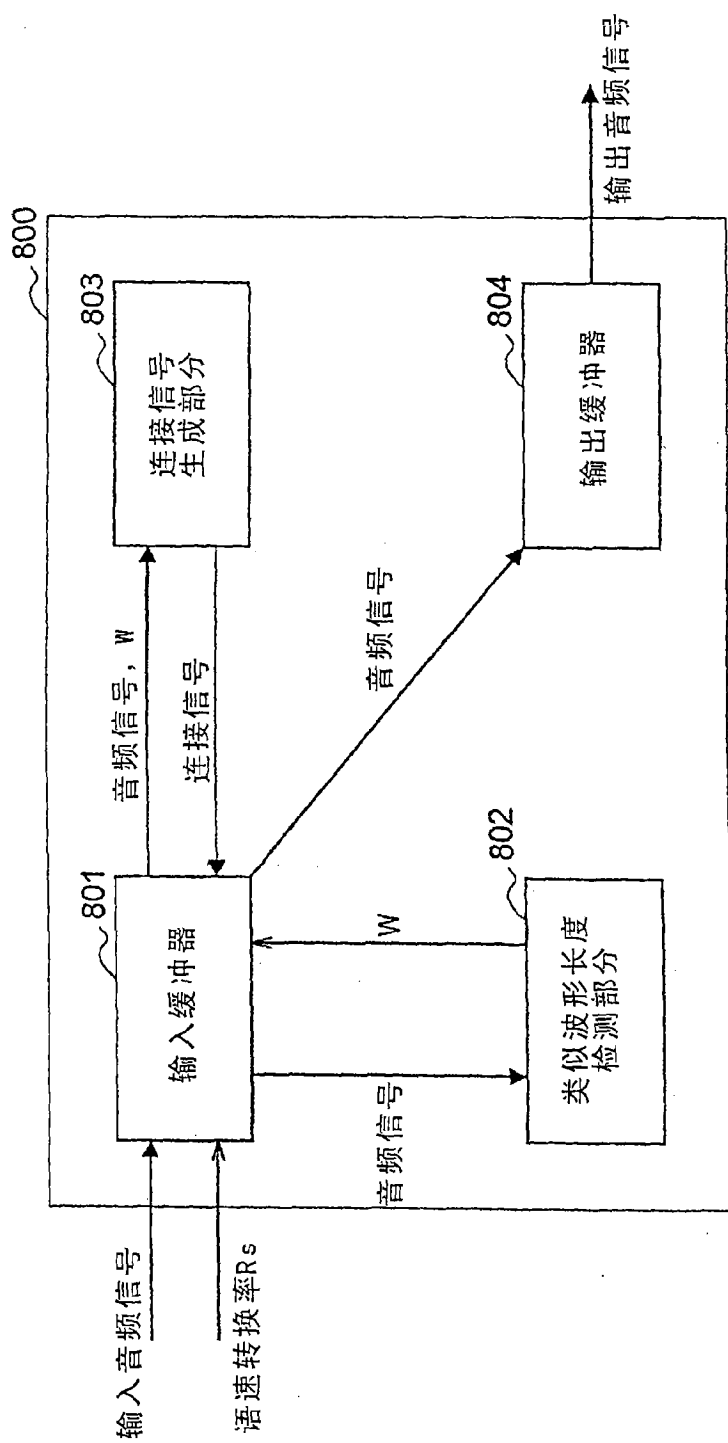


图8

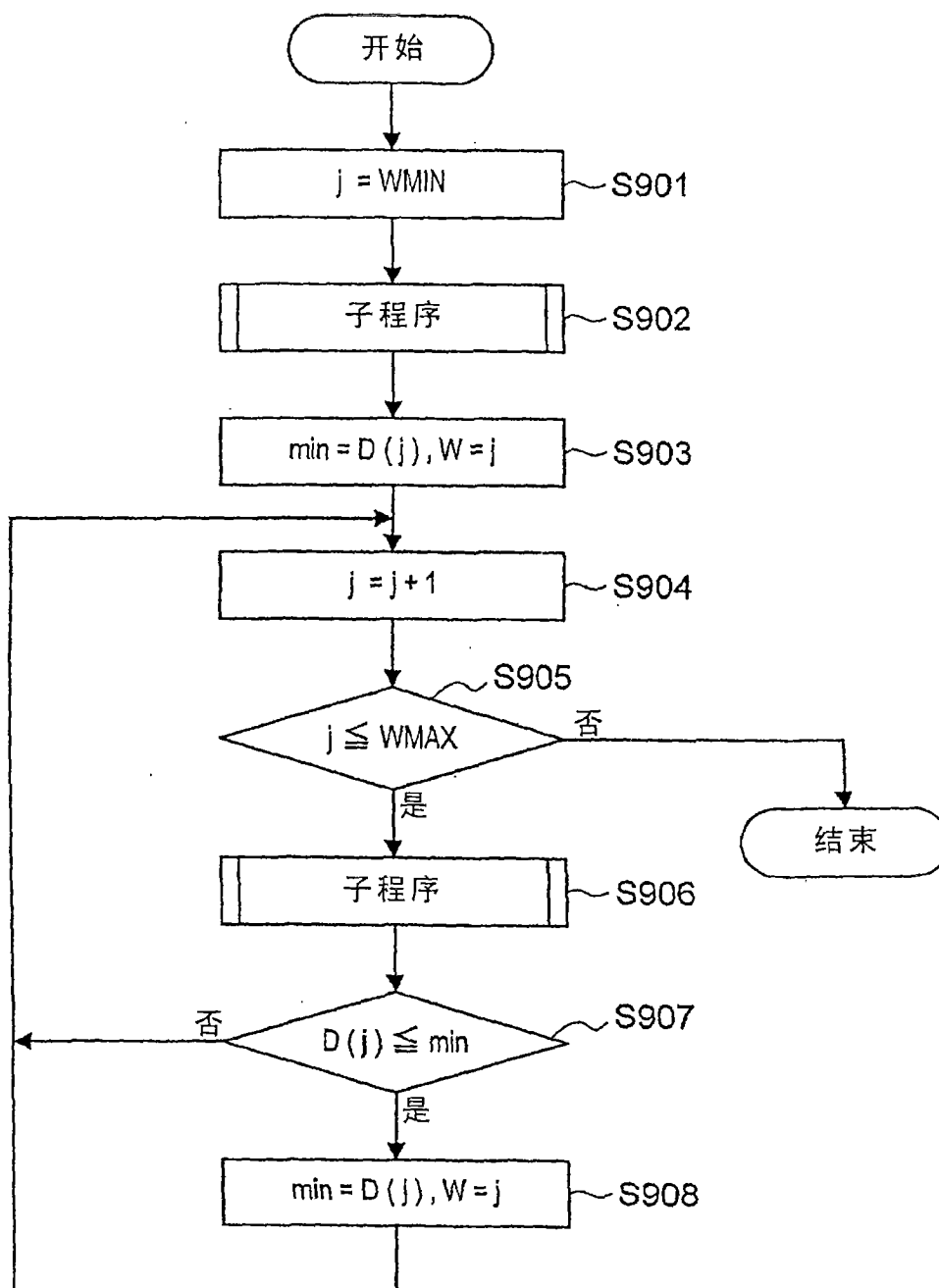


图 9

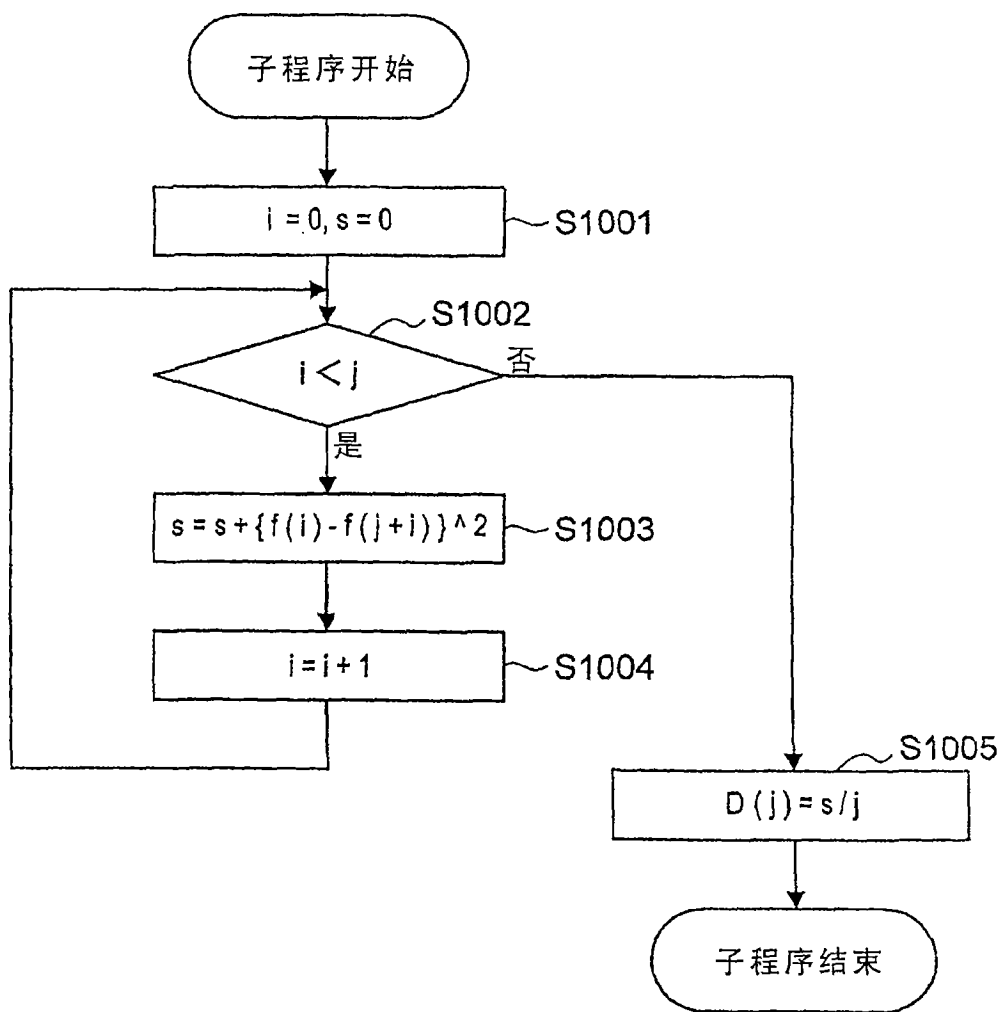


图 10

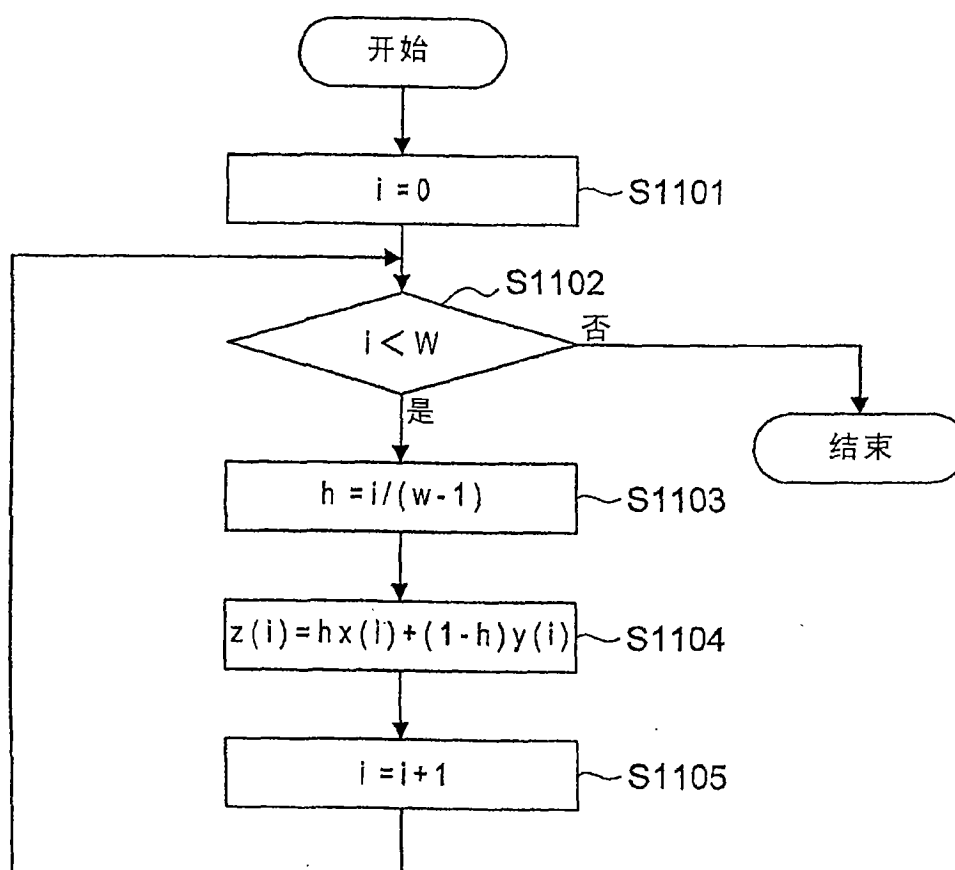
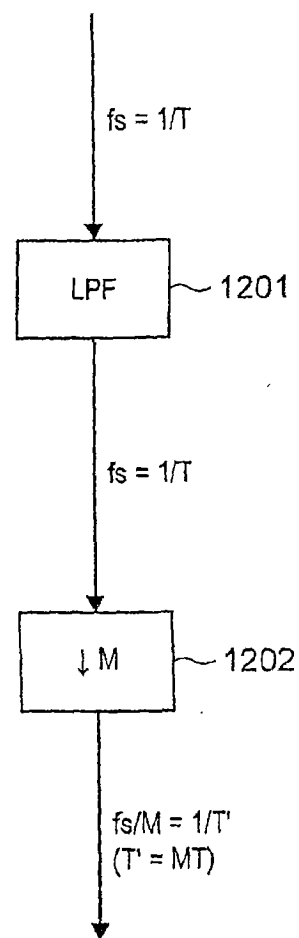
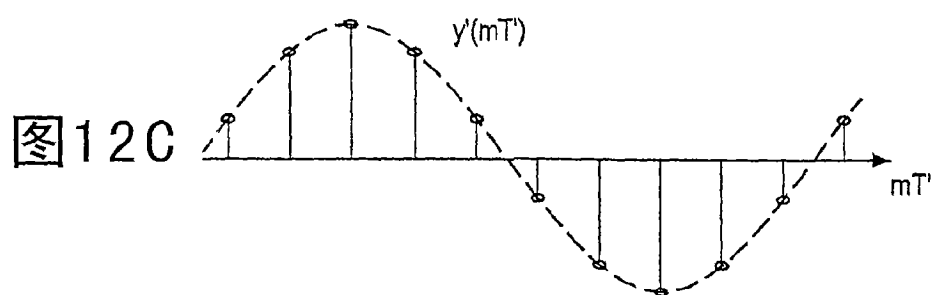
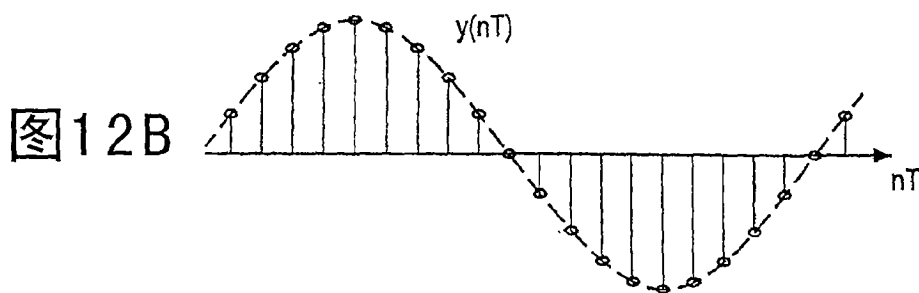
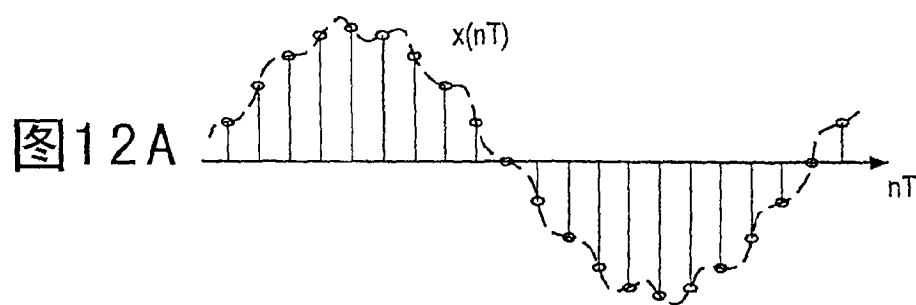


图 11



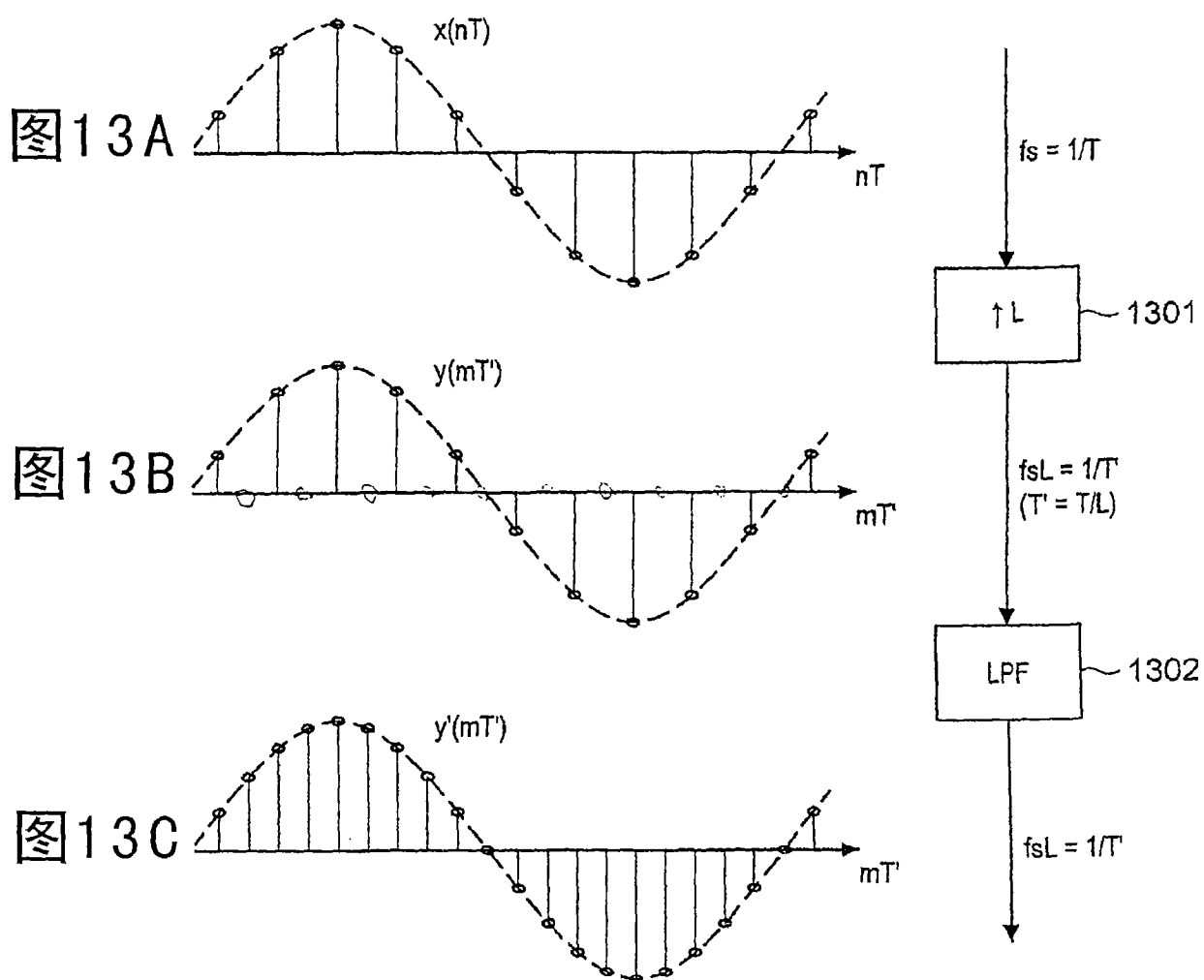


图14A

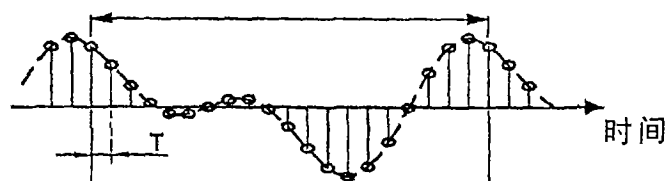


图14B

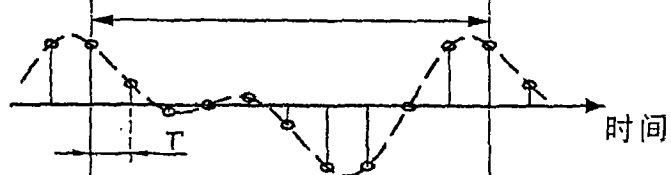


图14C

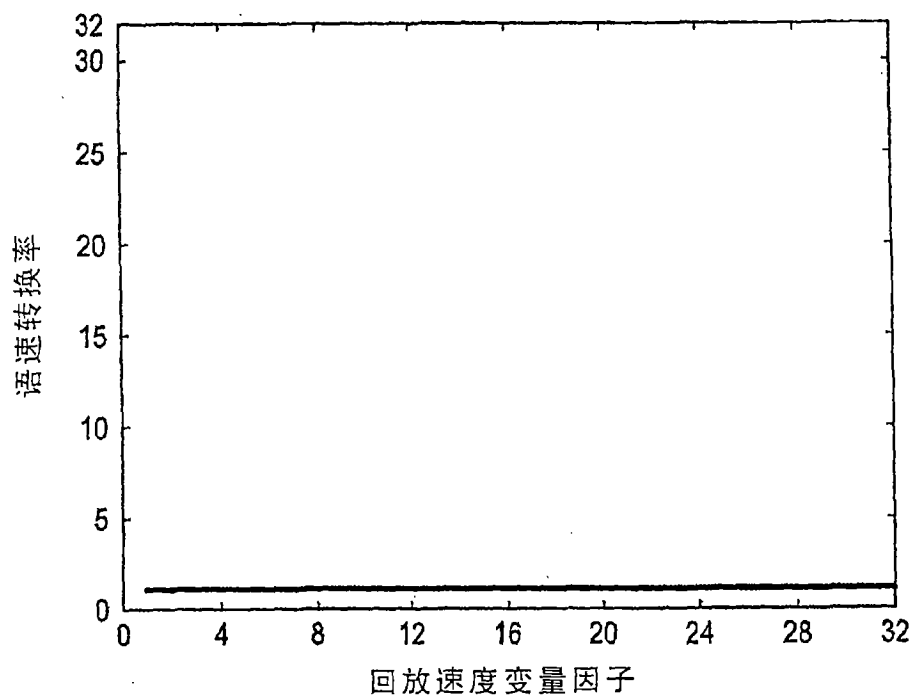
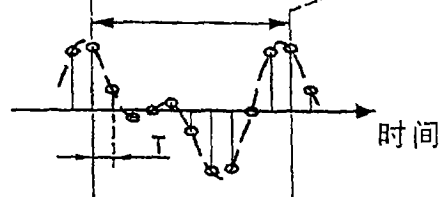


图 15A

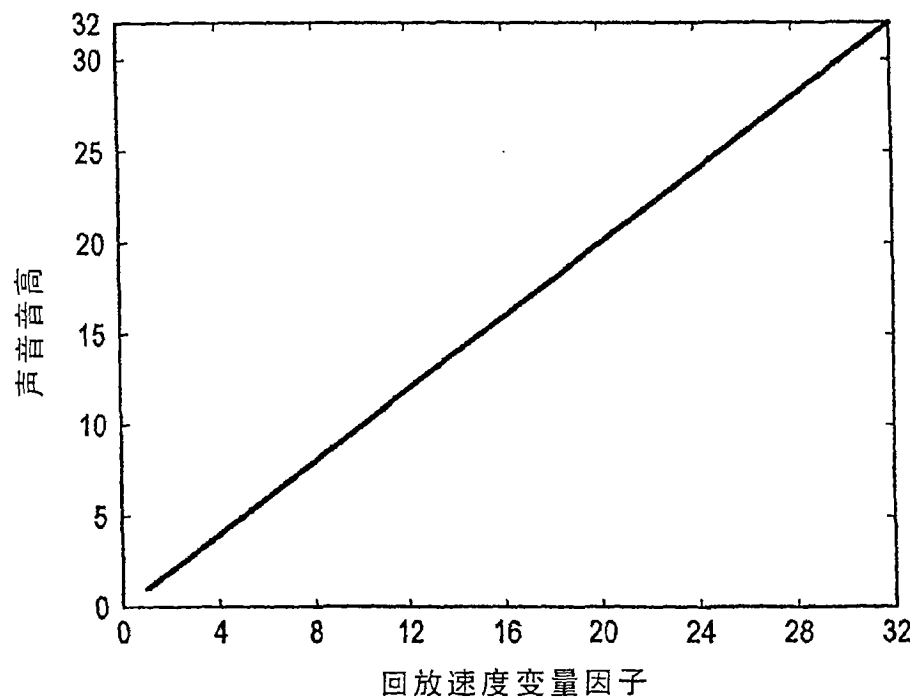


图 15B

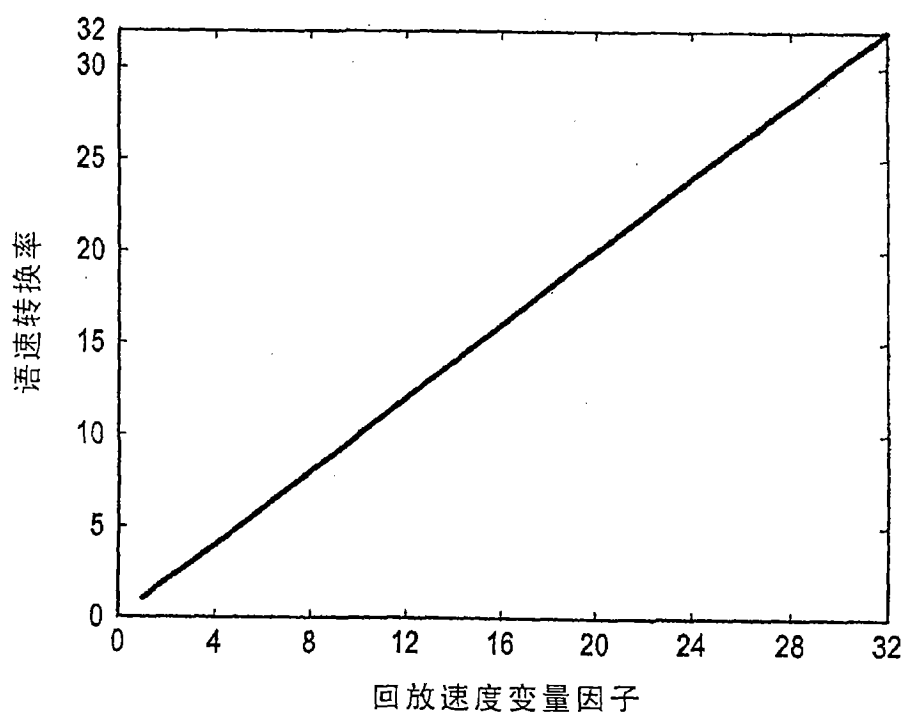


图 16A

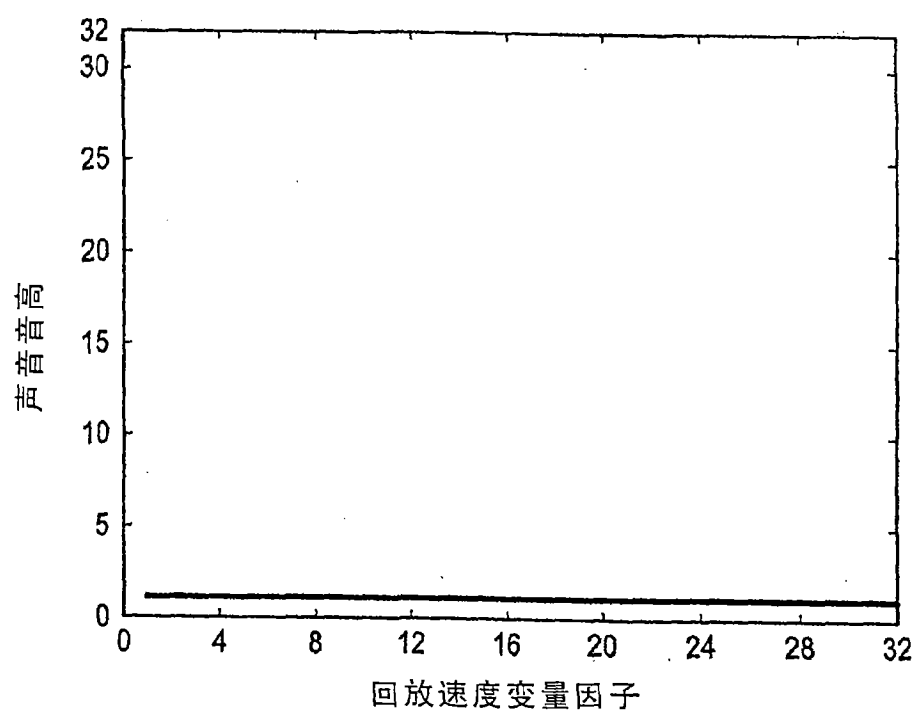


图 16B

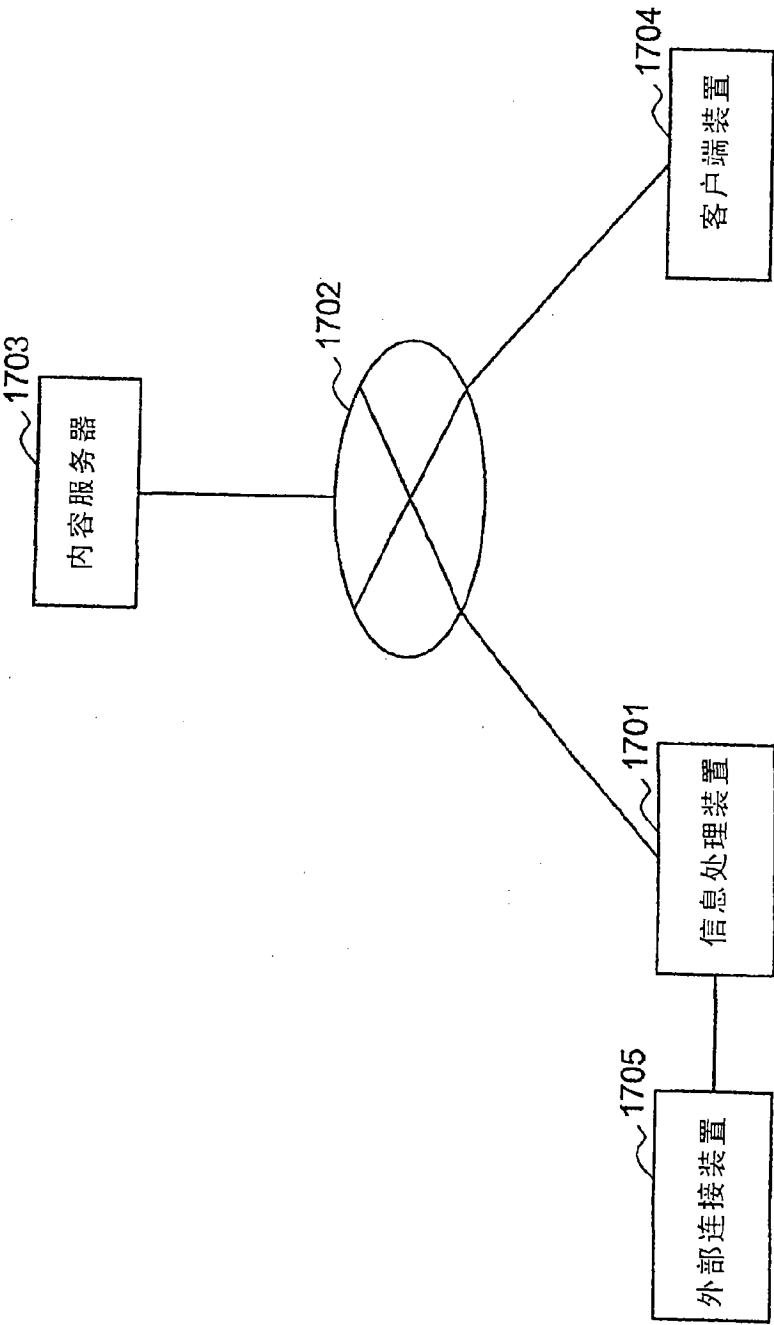


图17

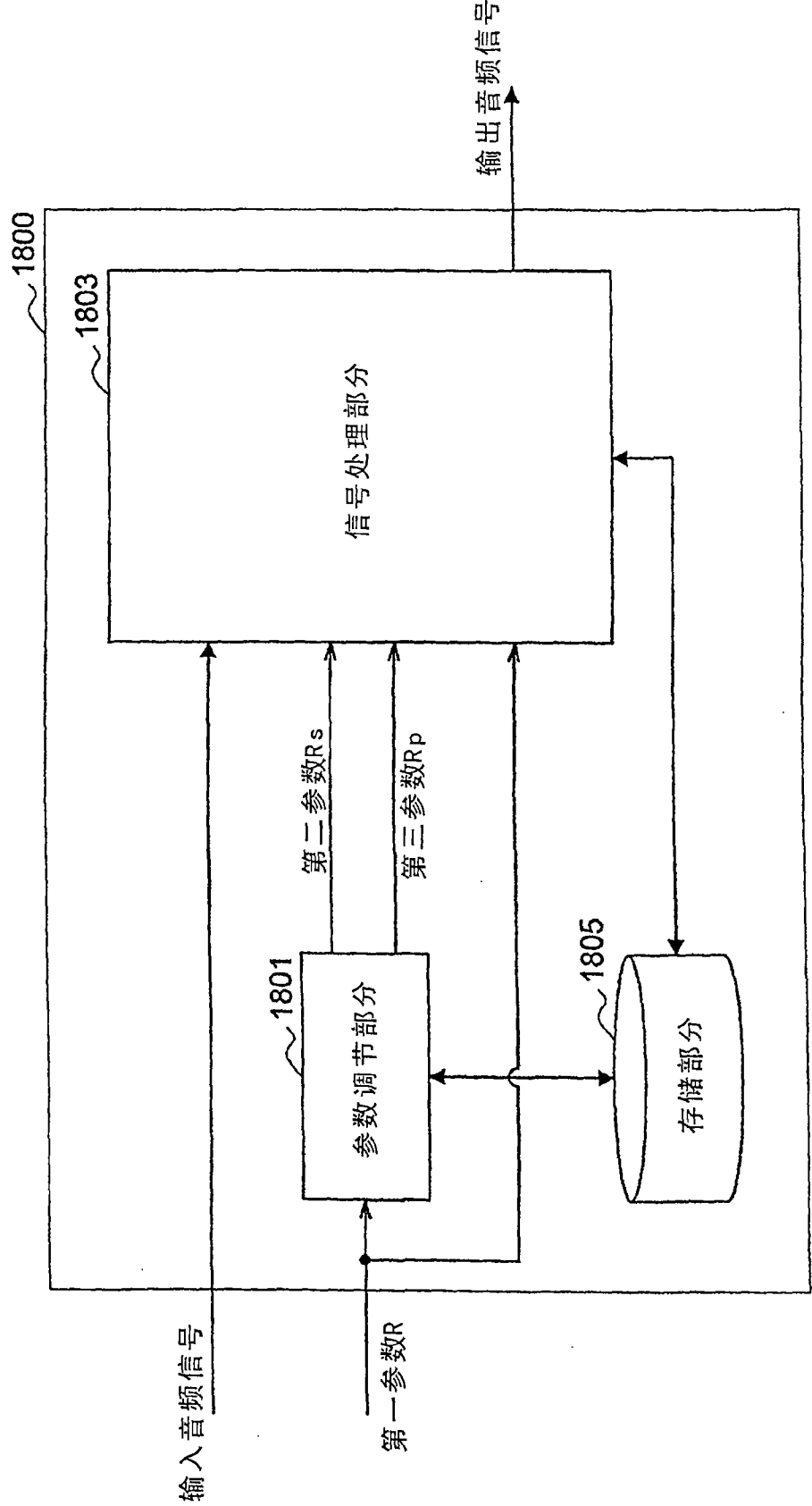


图18

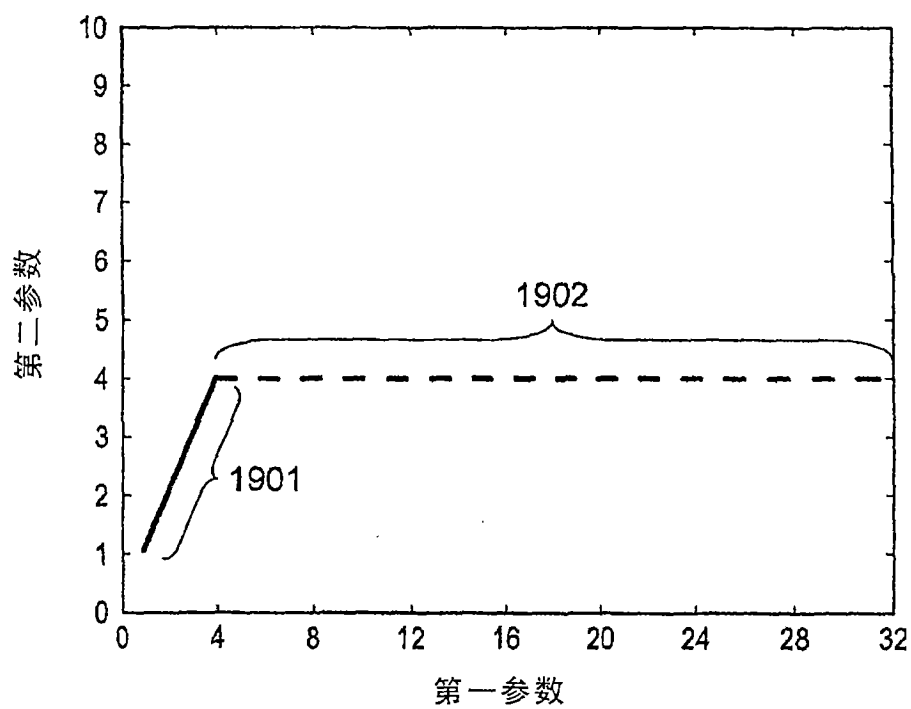


图 19A

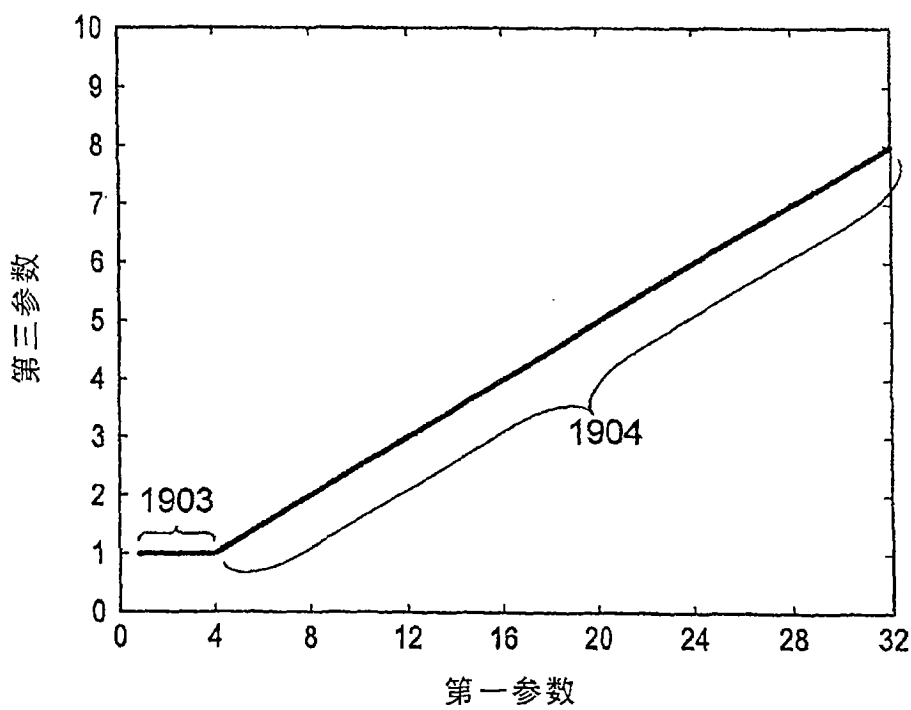


图 19B

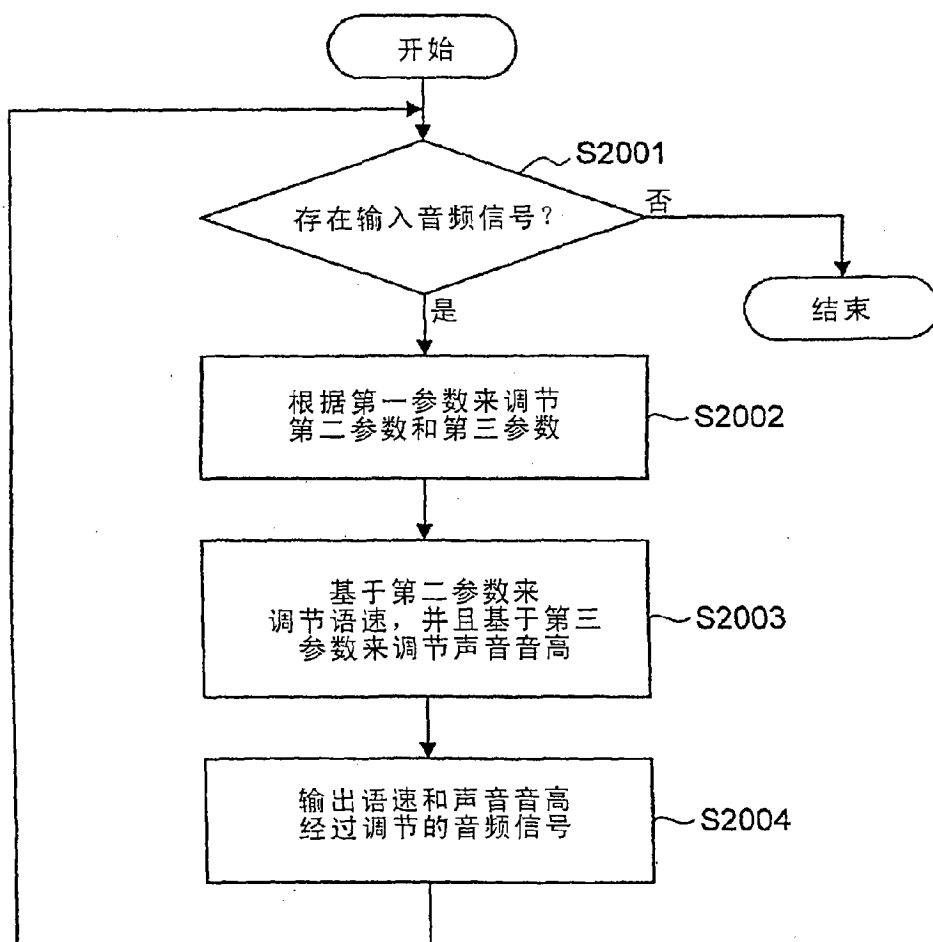


图 20

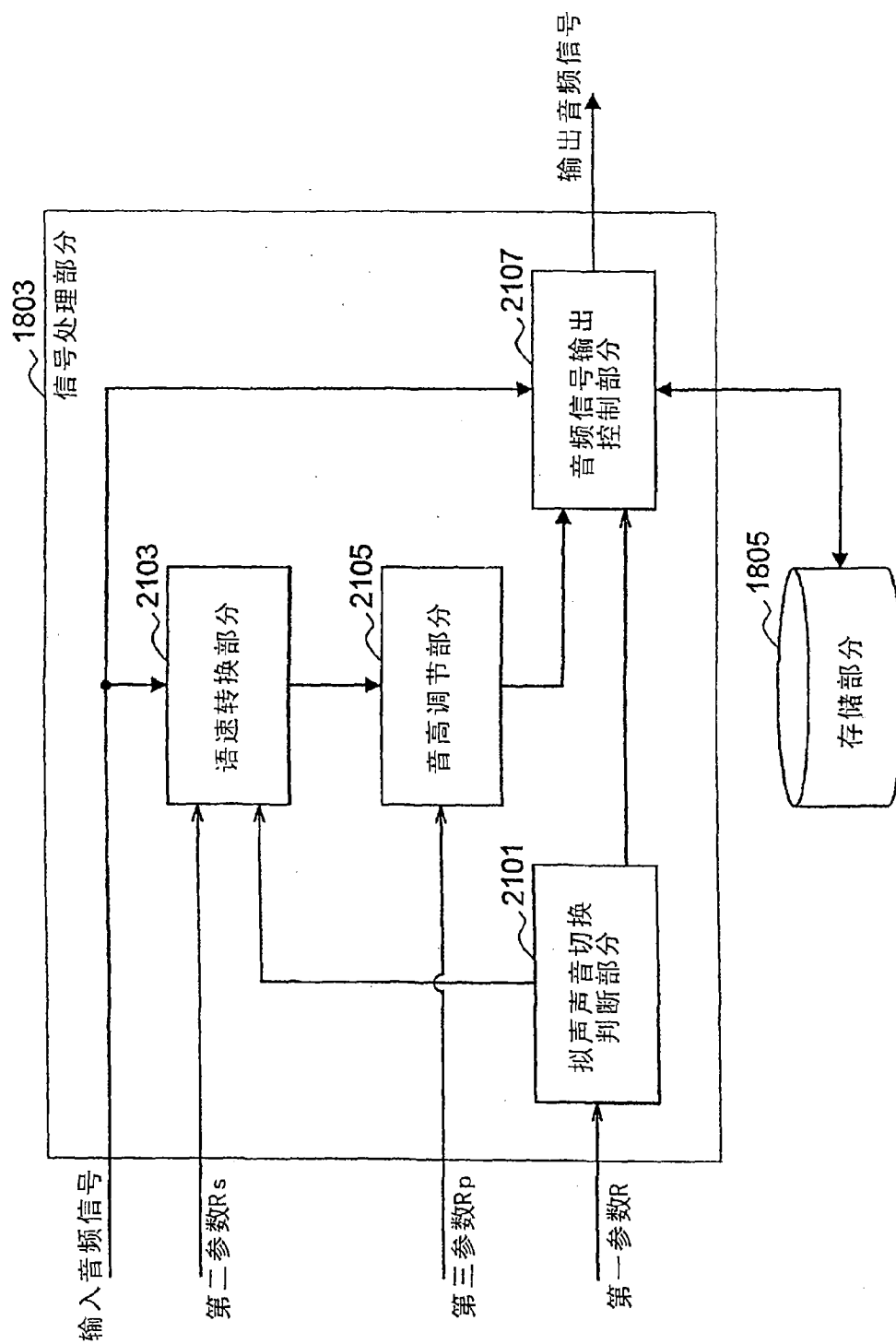


图21

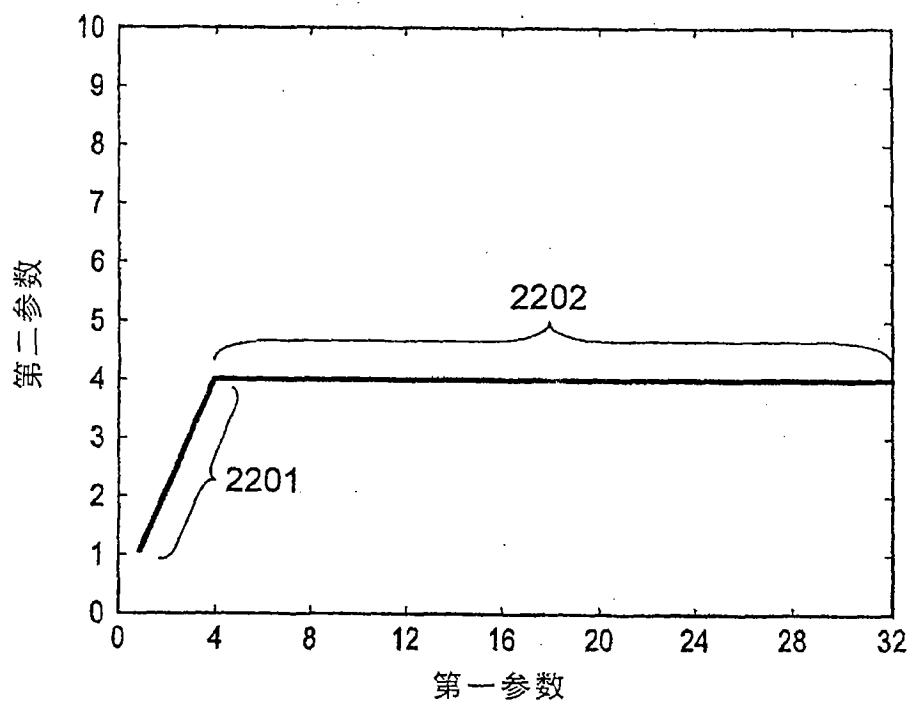


图 22A

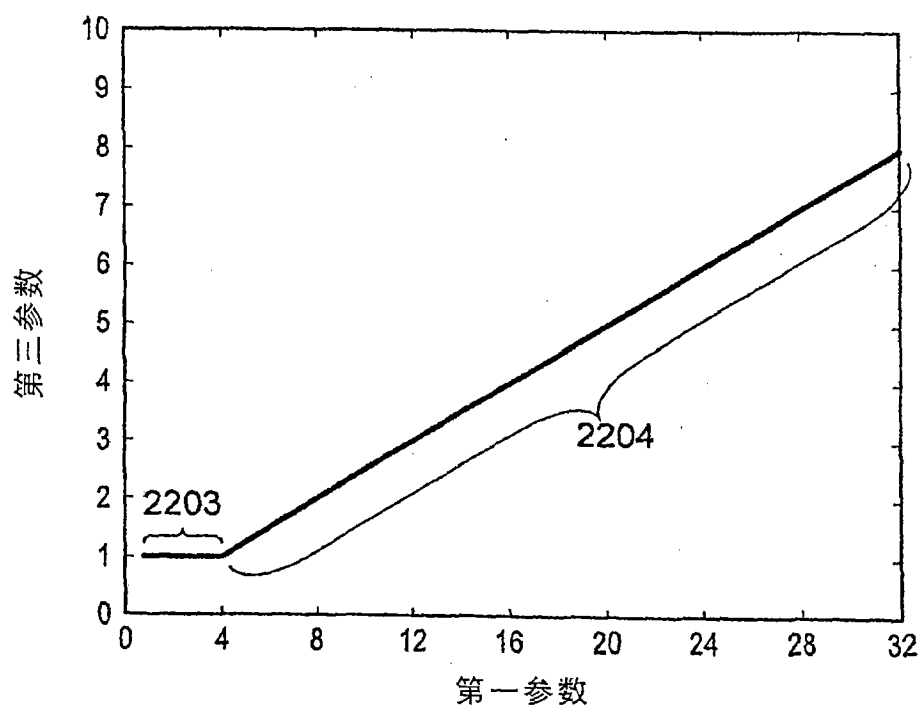


图 22B

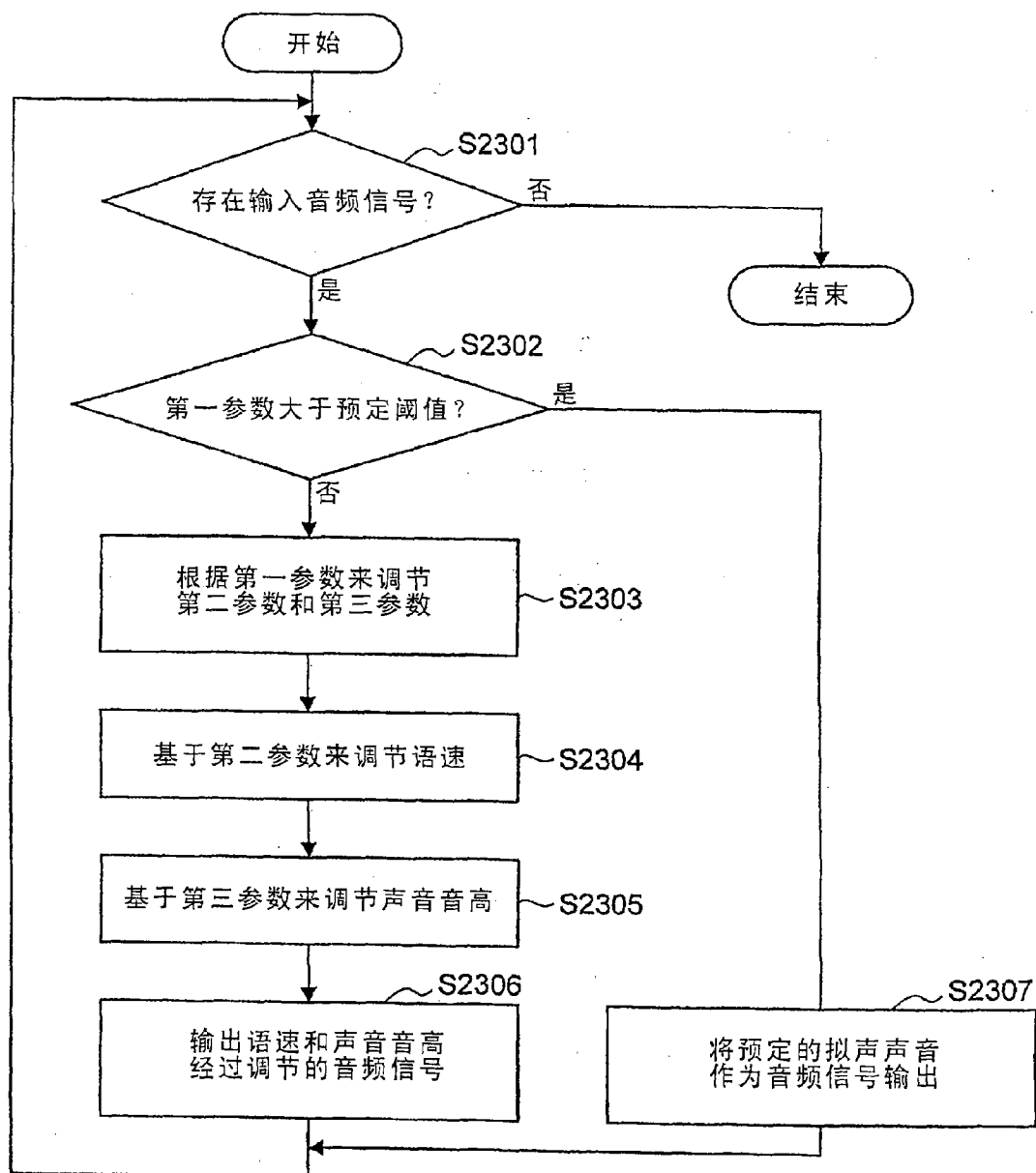
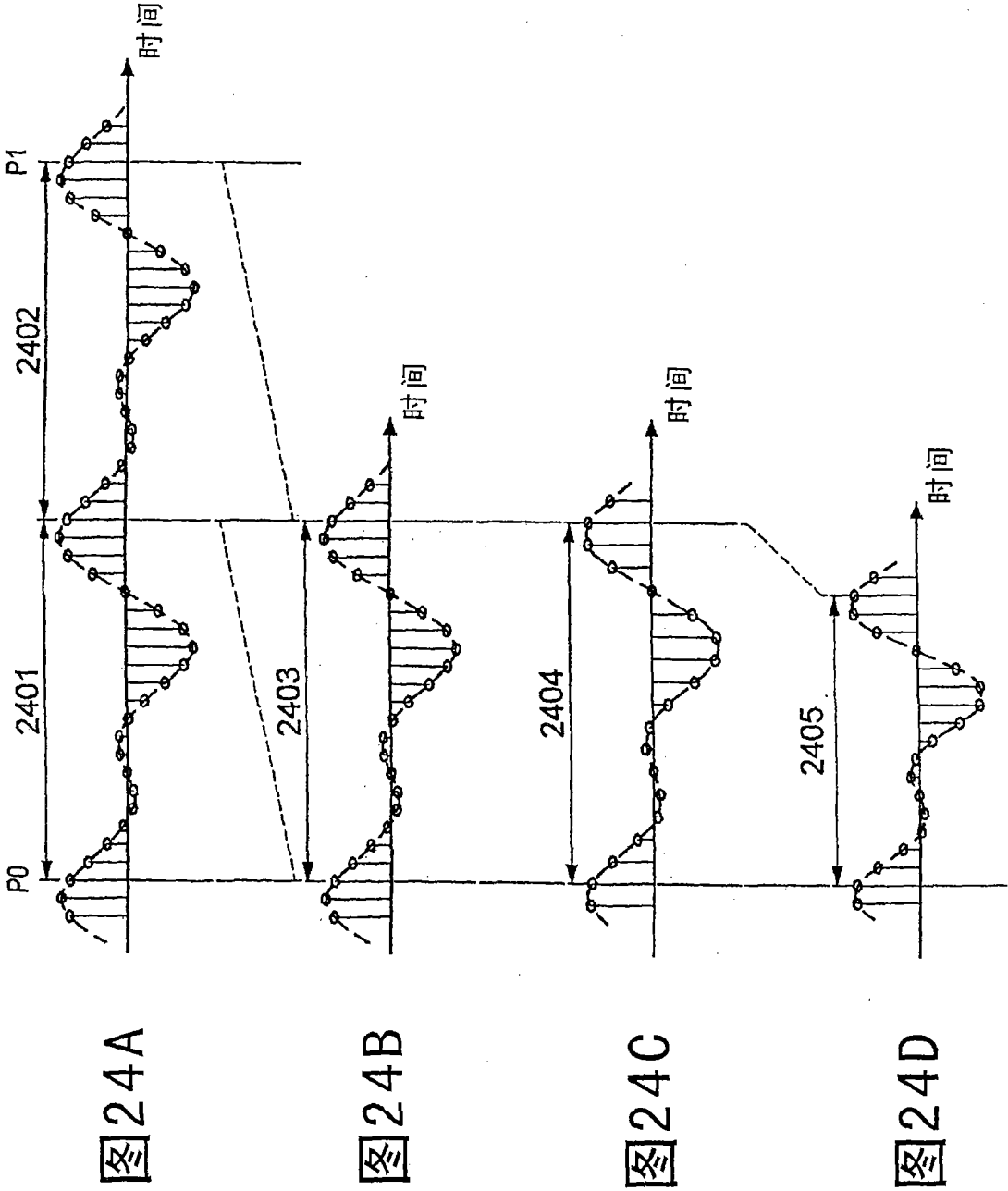


图 23



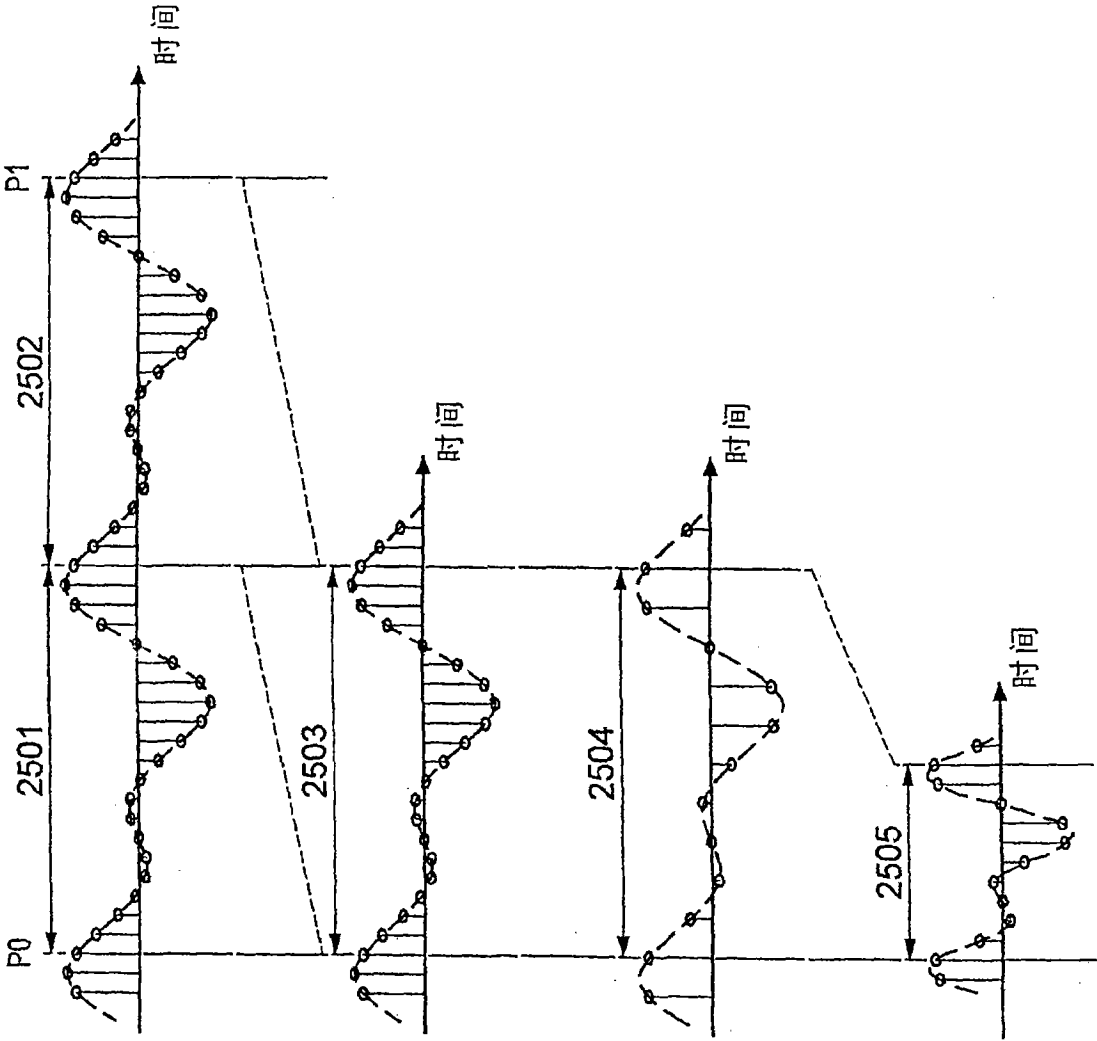


图25A

图25B

图25C

图25D

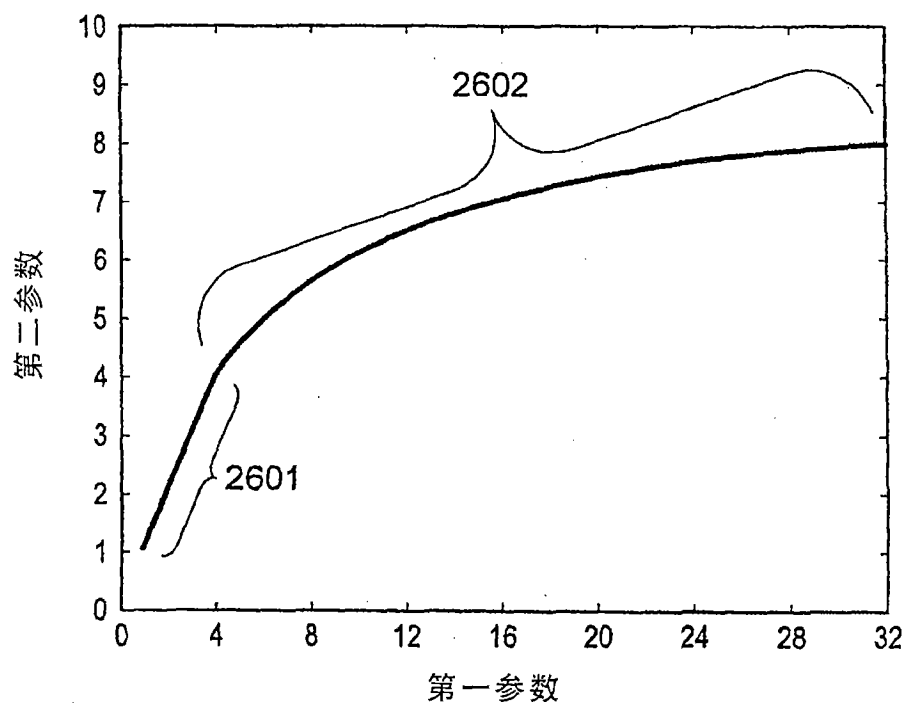


图 26A

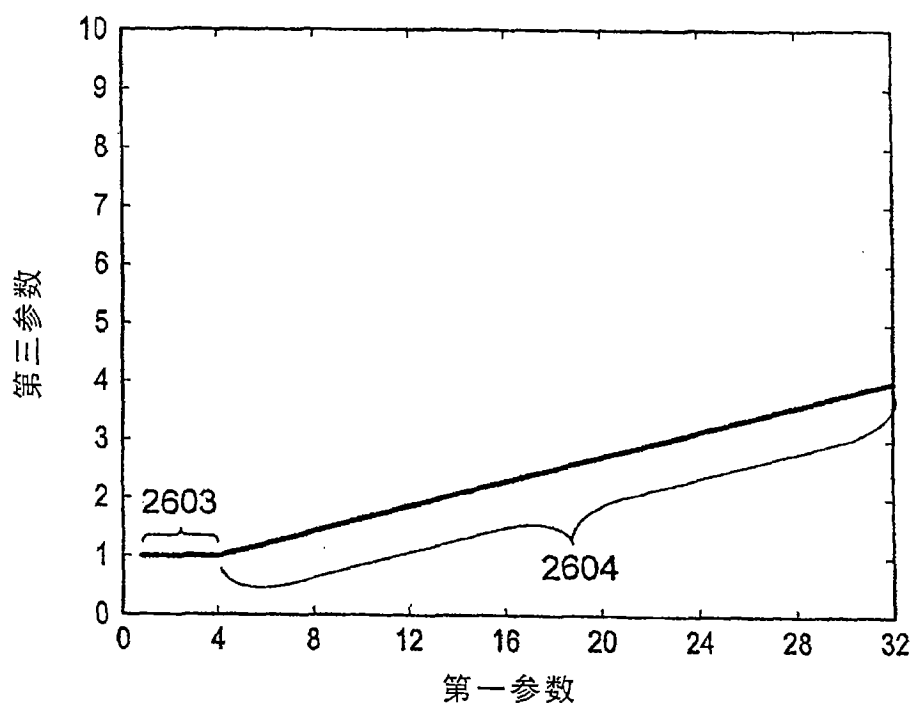


图 26B

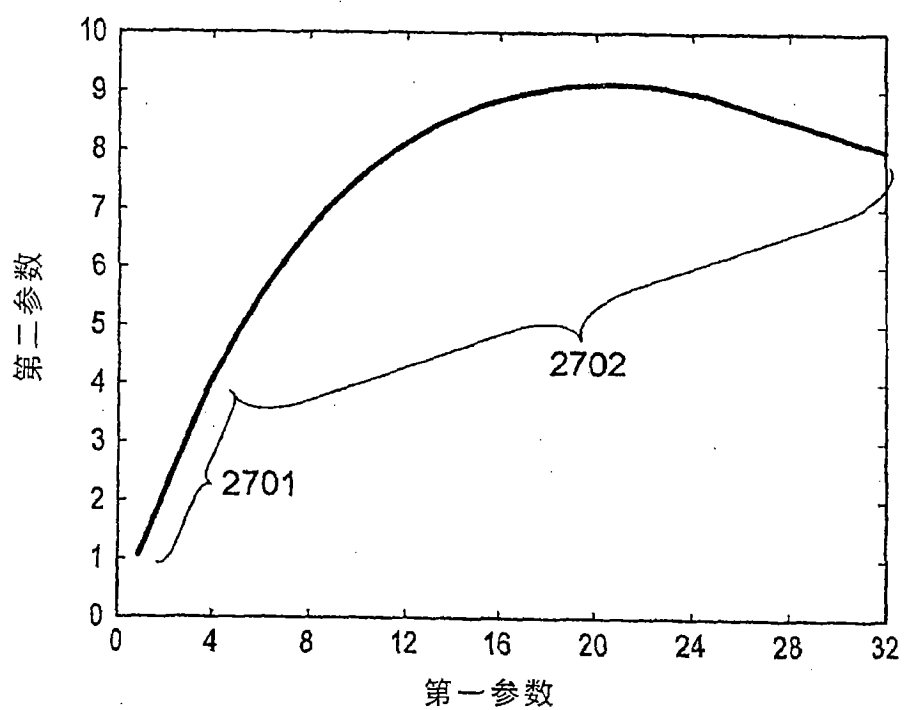


图 27A

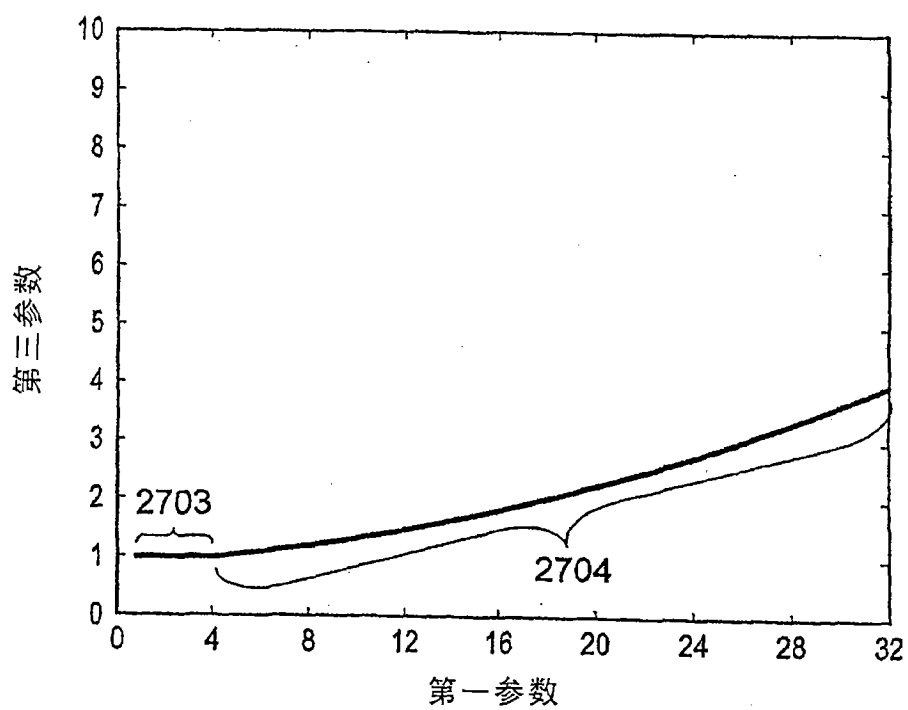


图 27B

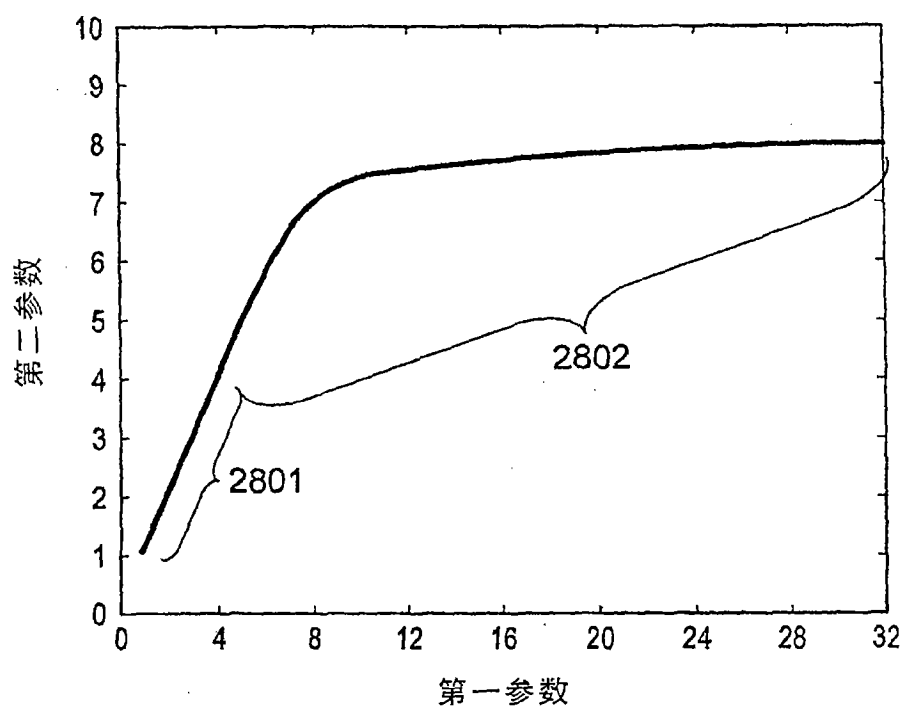


图 28A

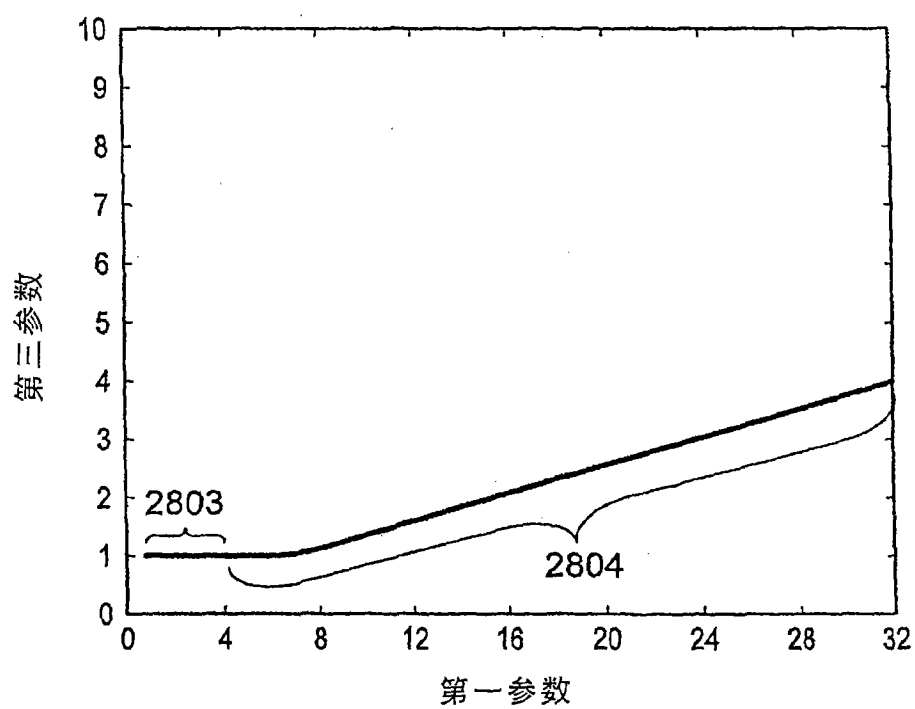


图 28B

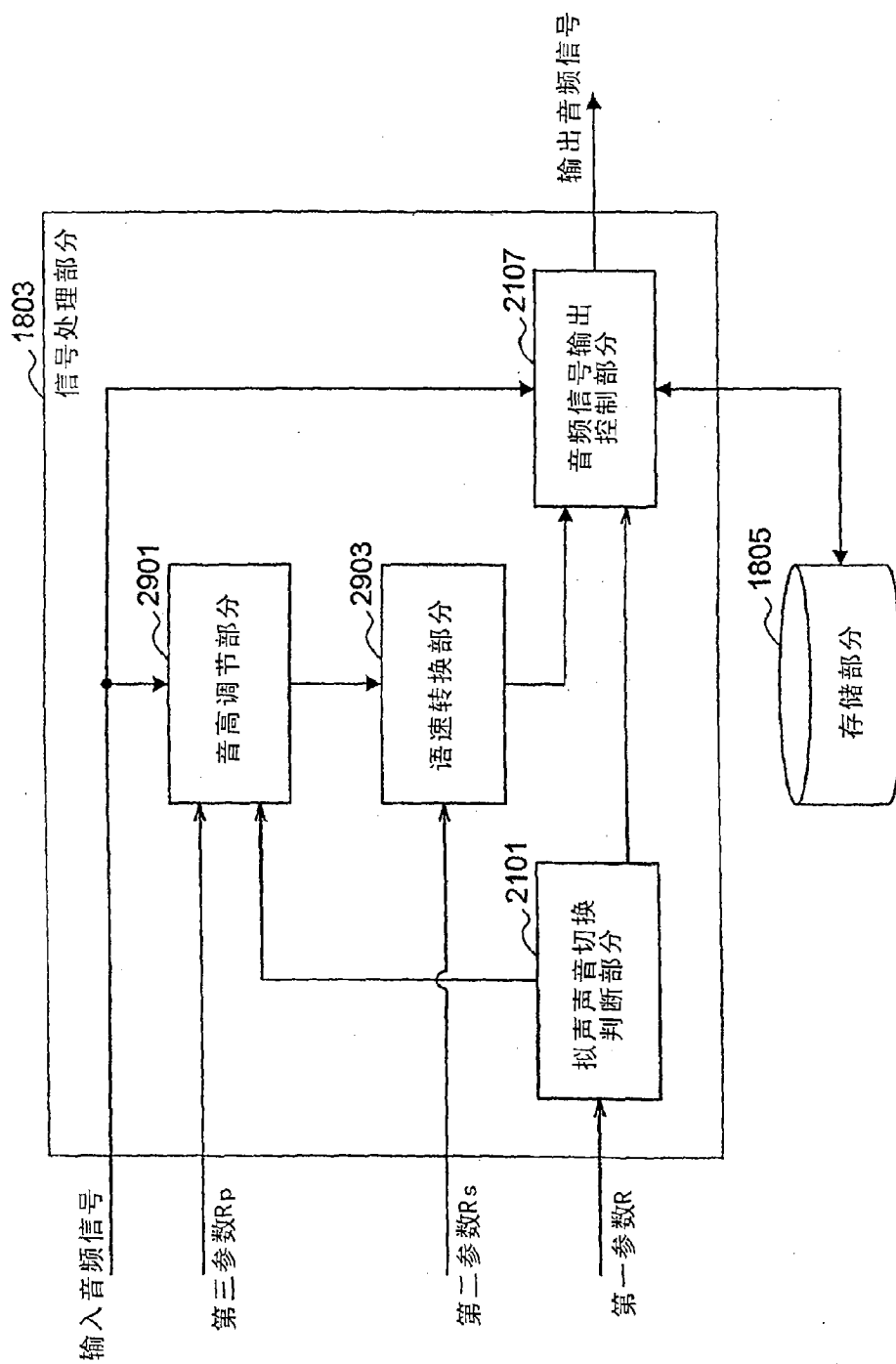


图29

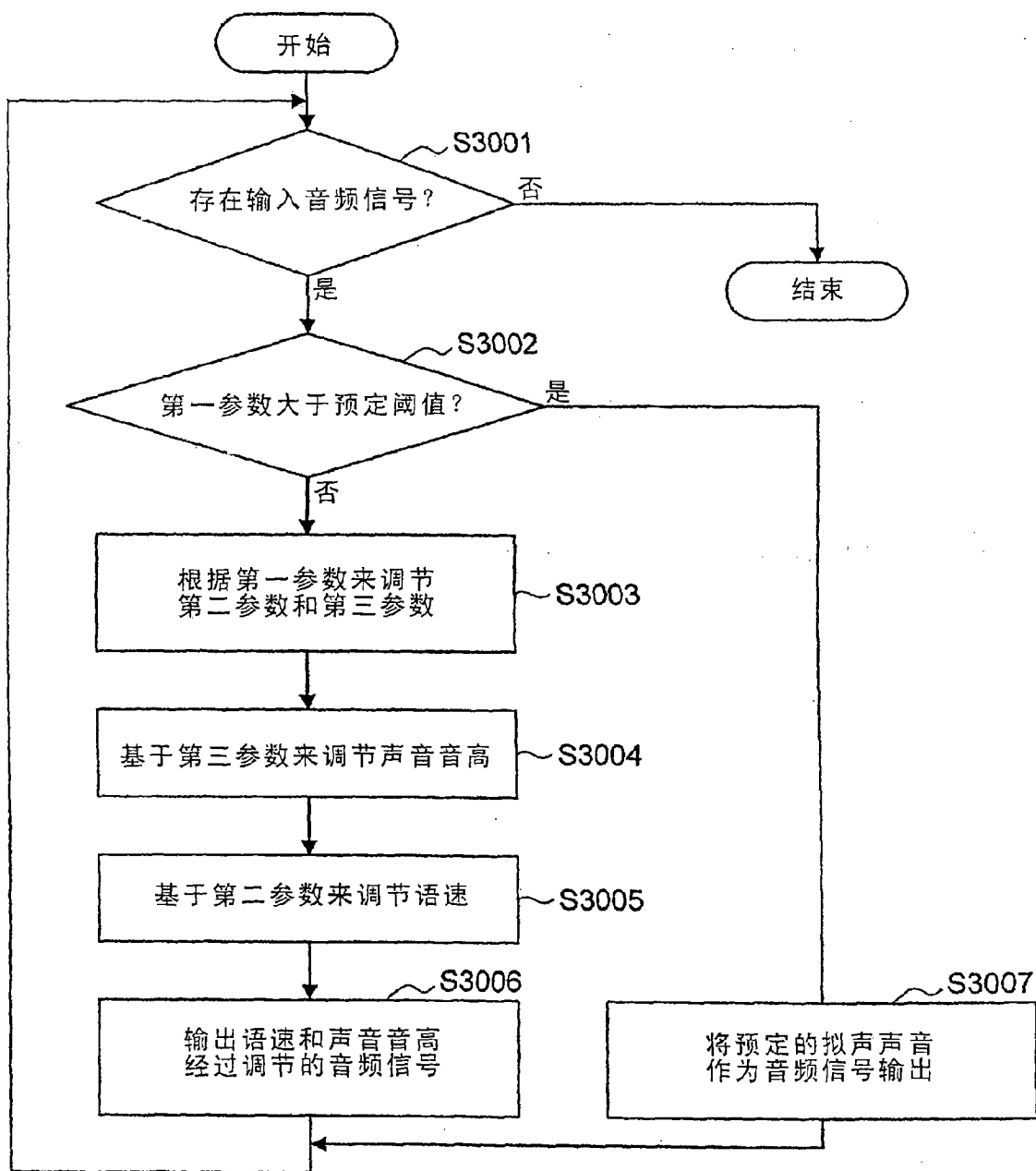


图 30

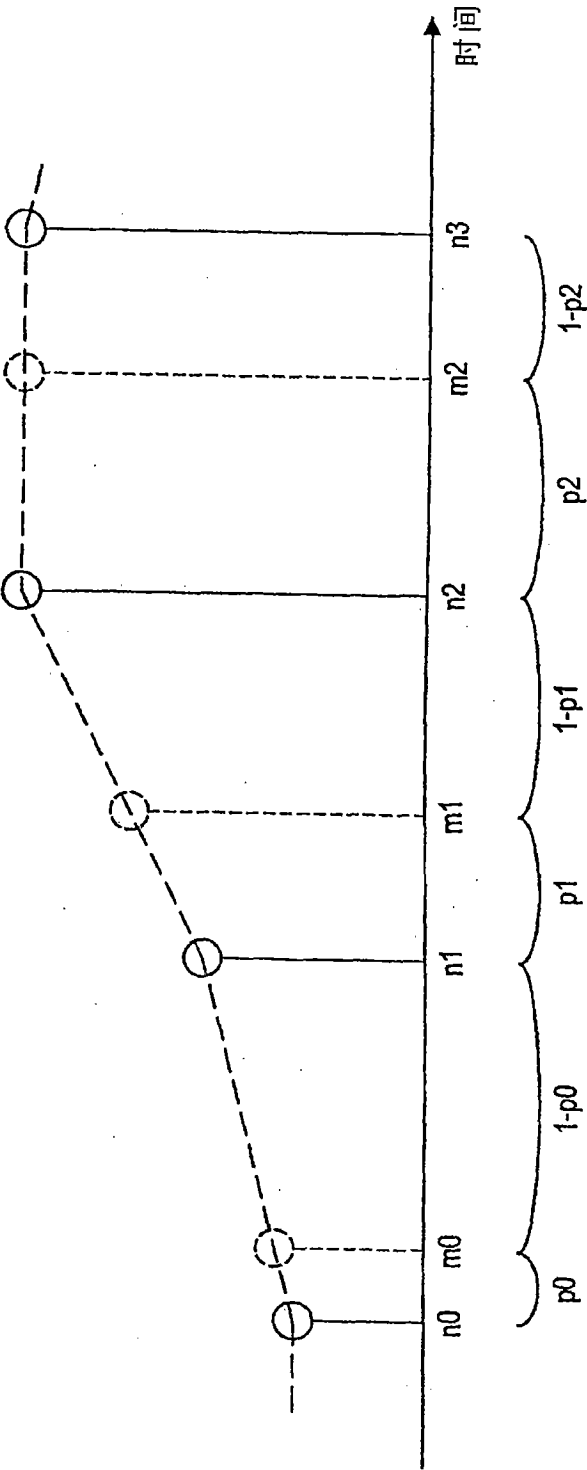


图31

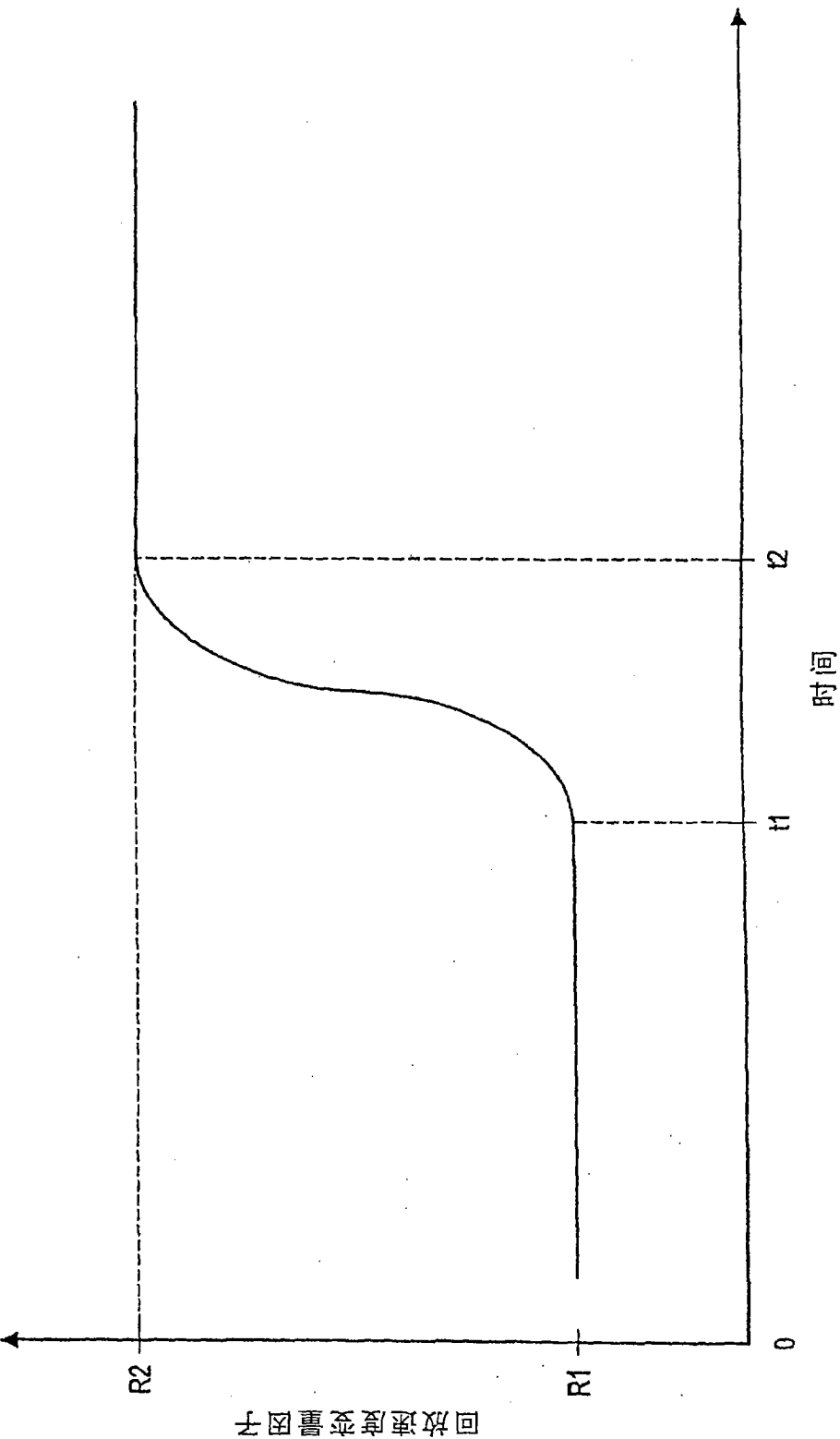


图32

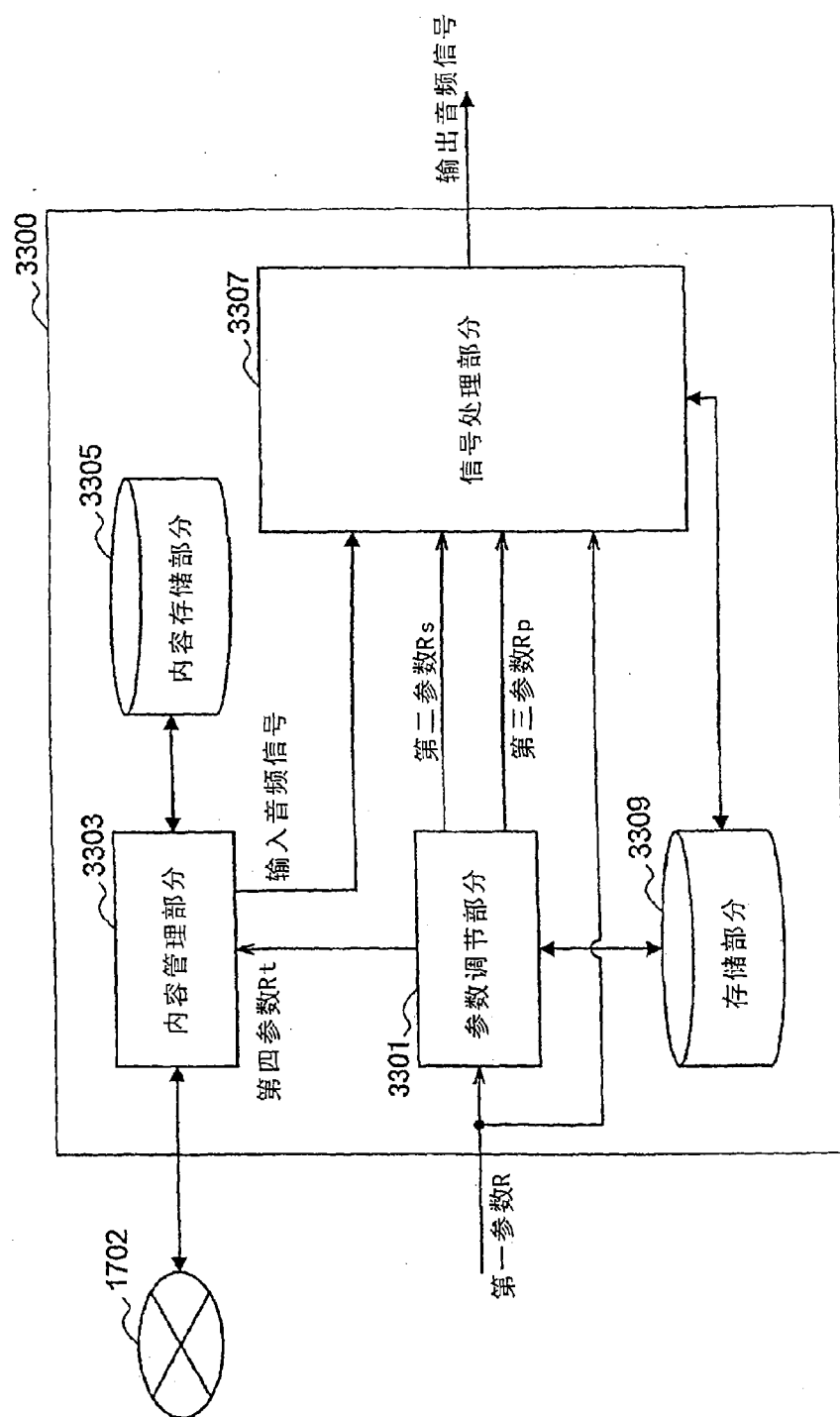


图33

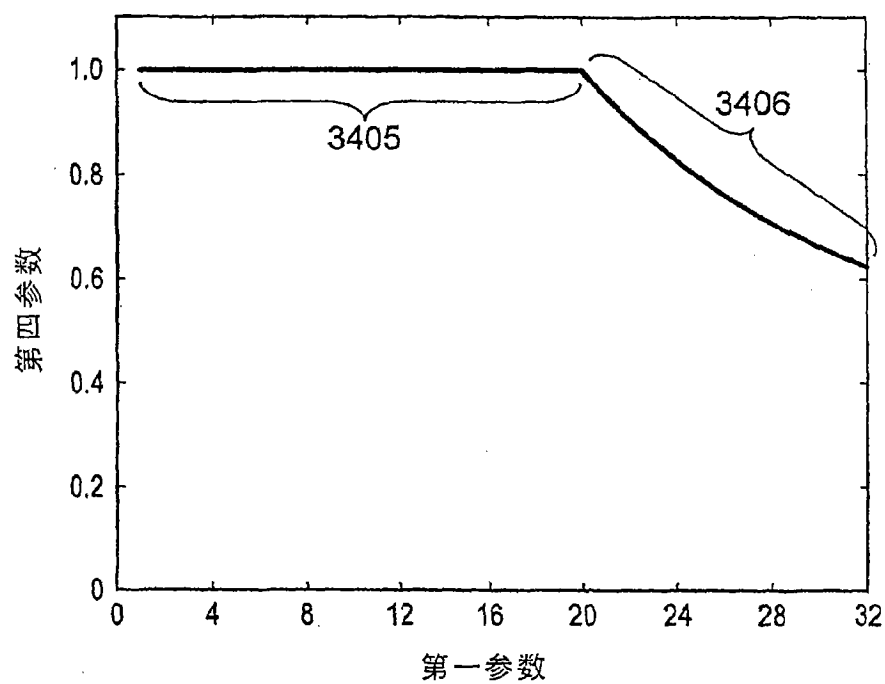


图 34A

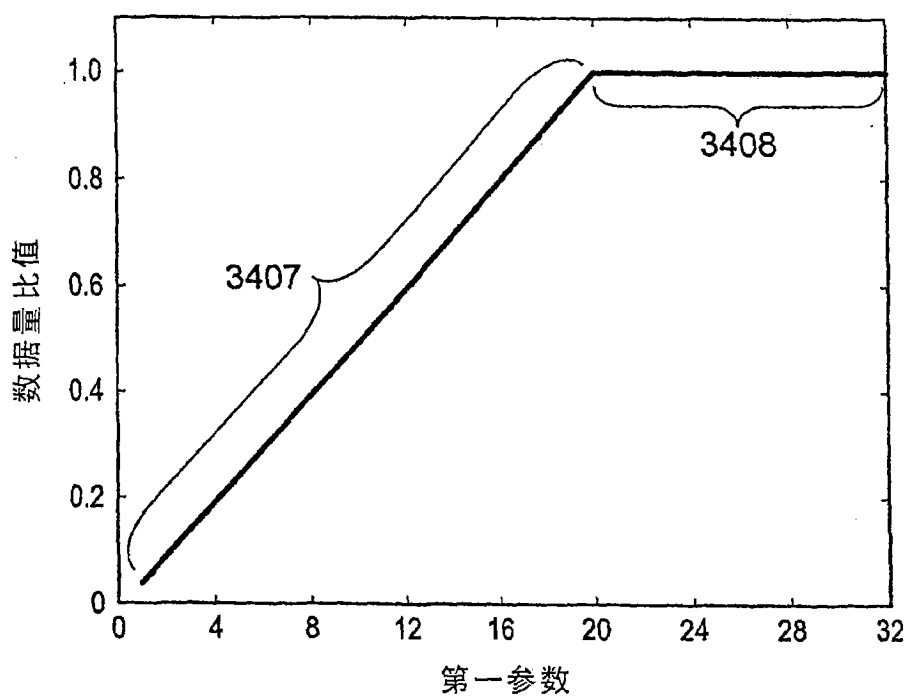
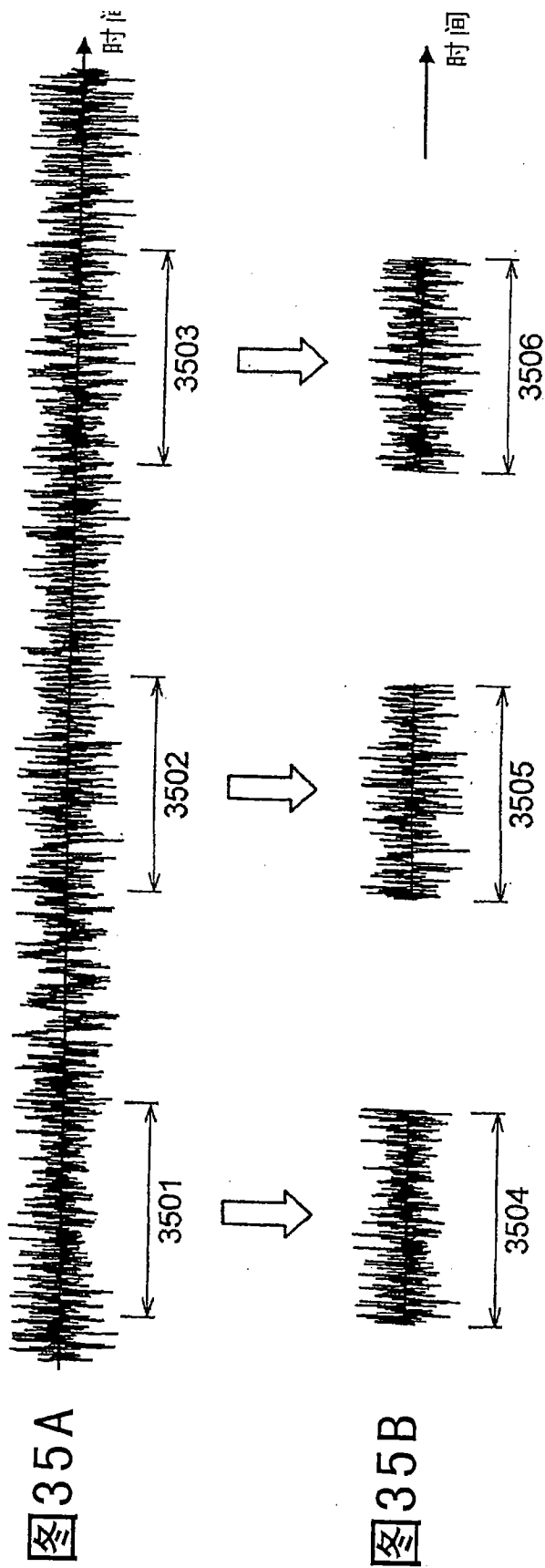
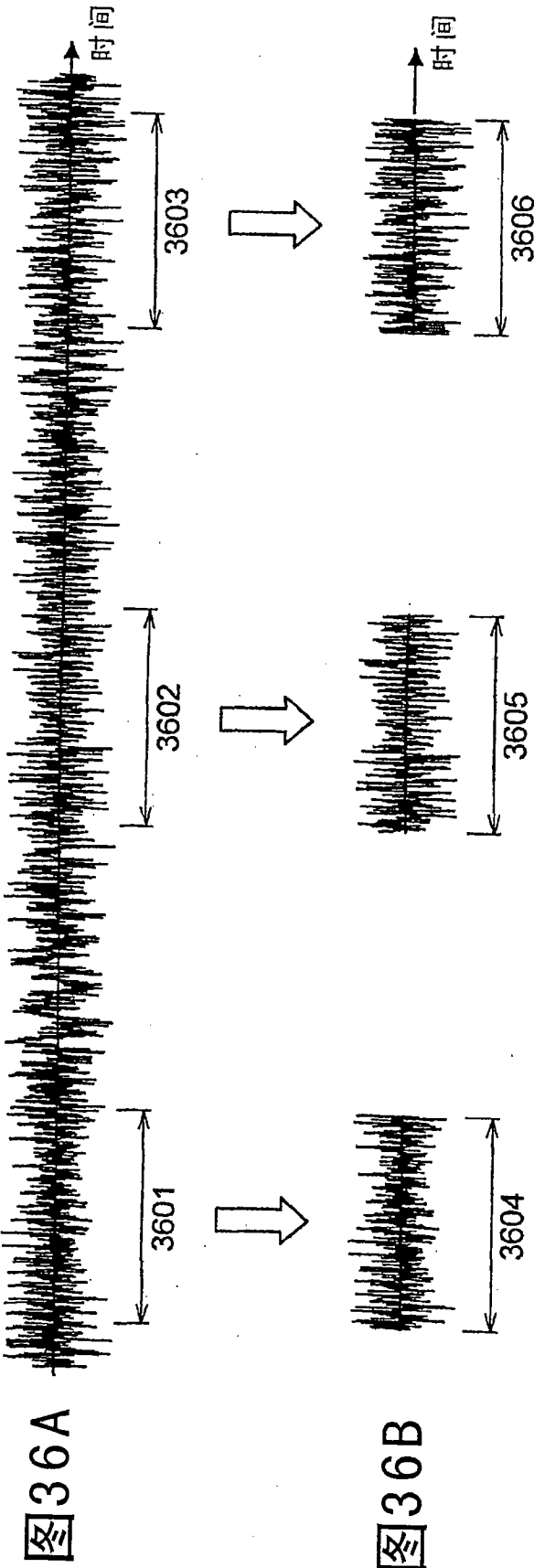
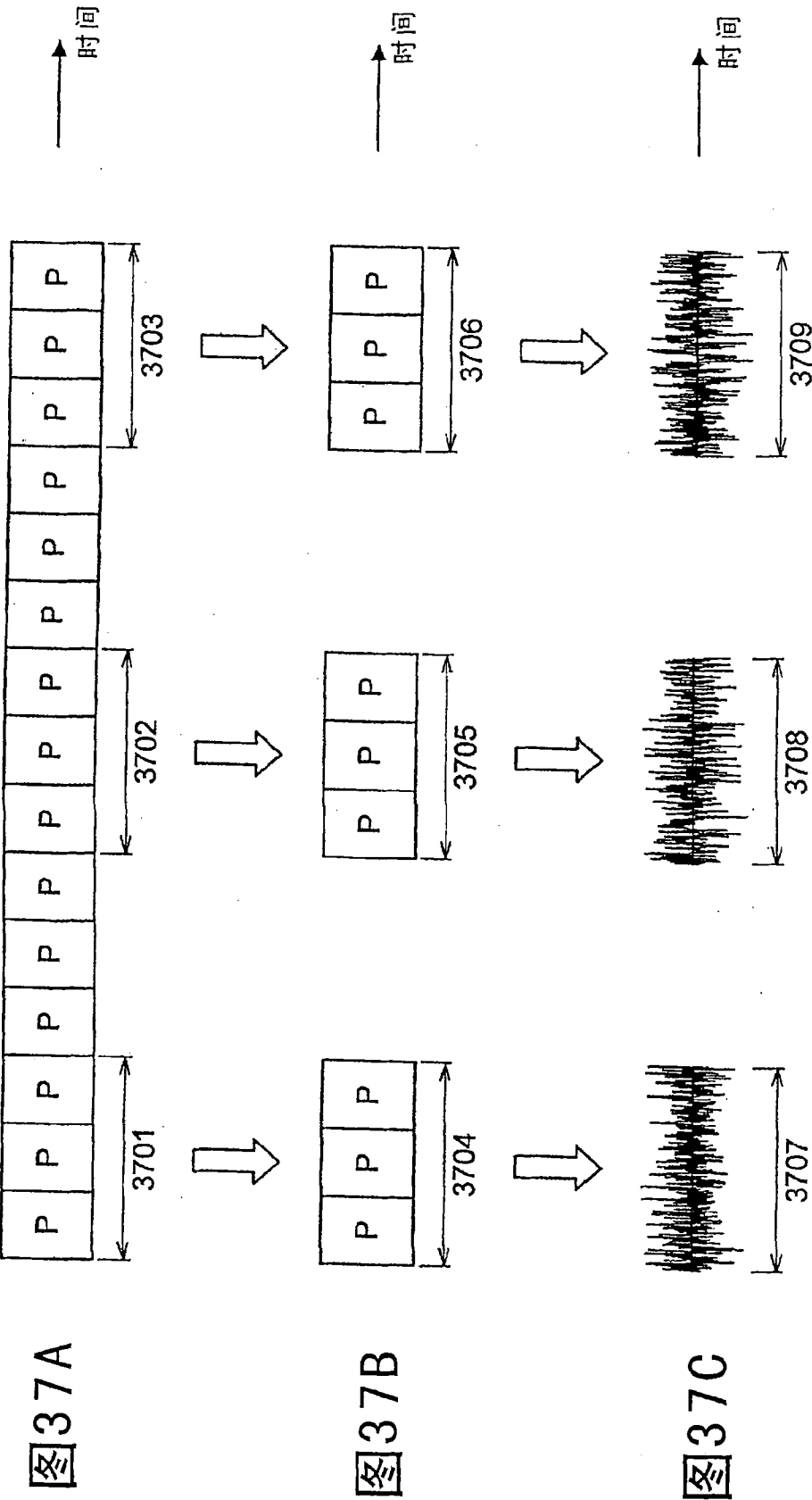


图 34B







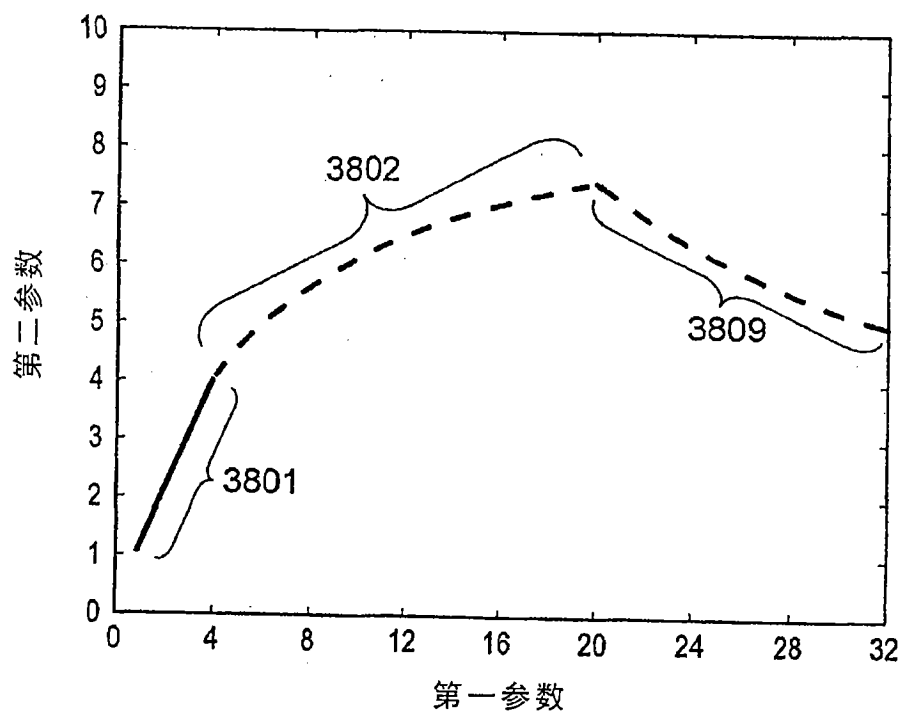


图 38A

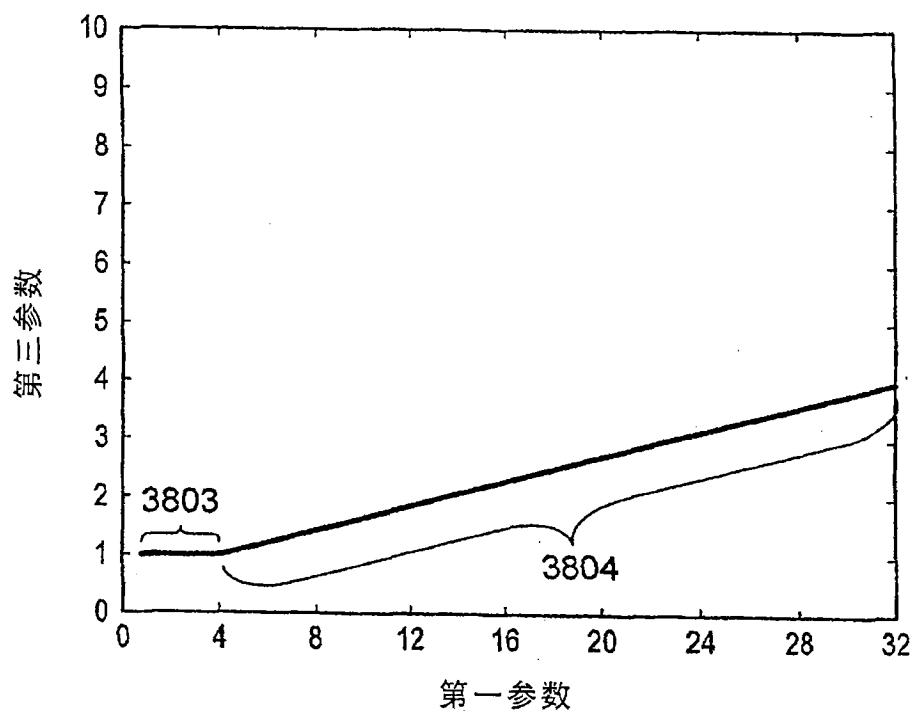


图 38B

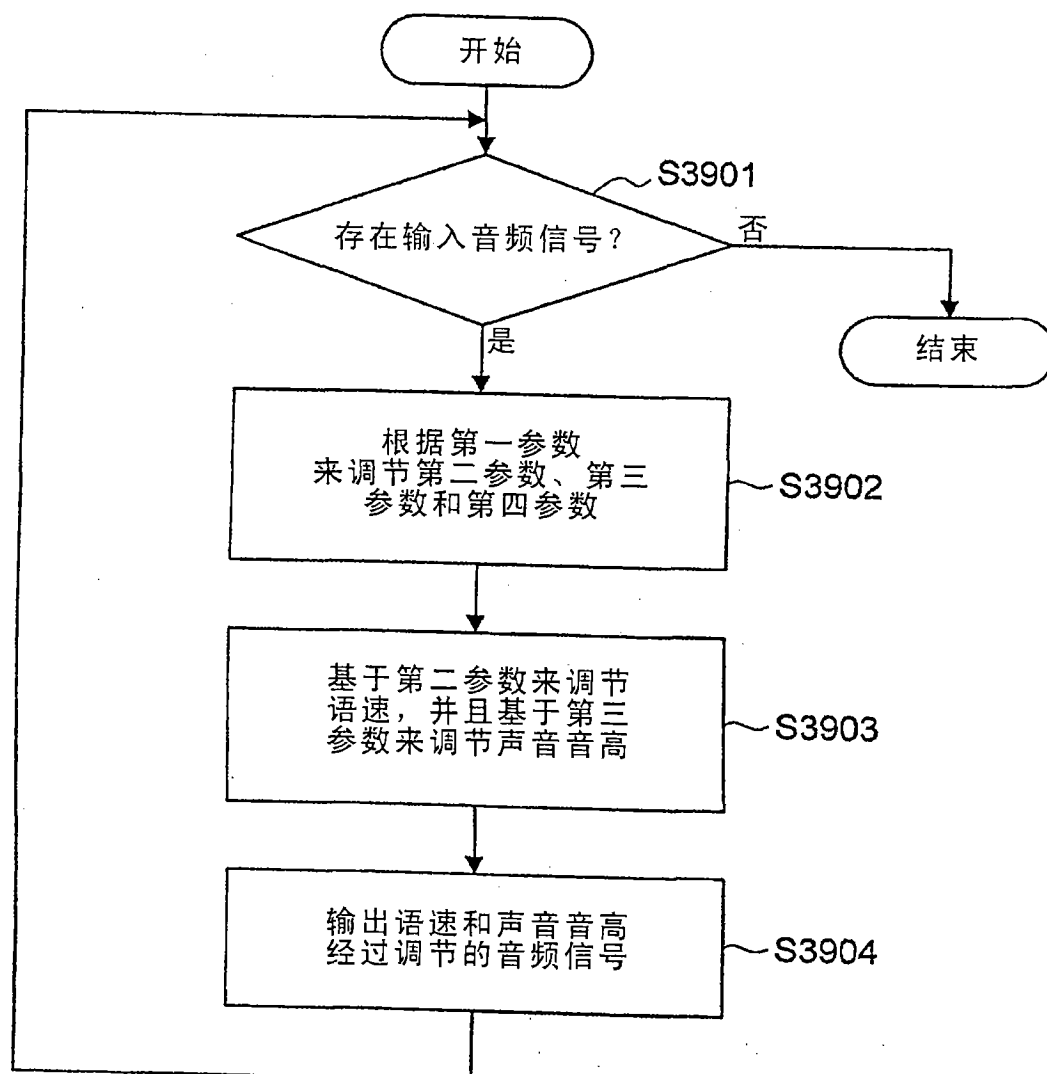


图 39

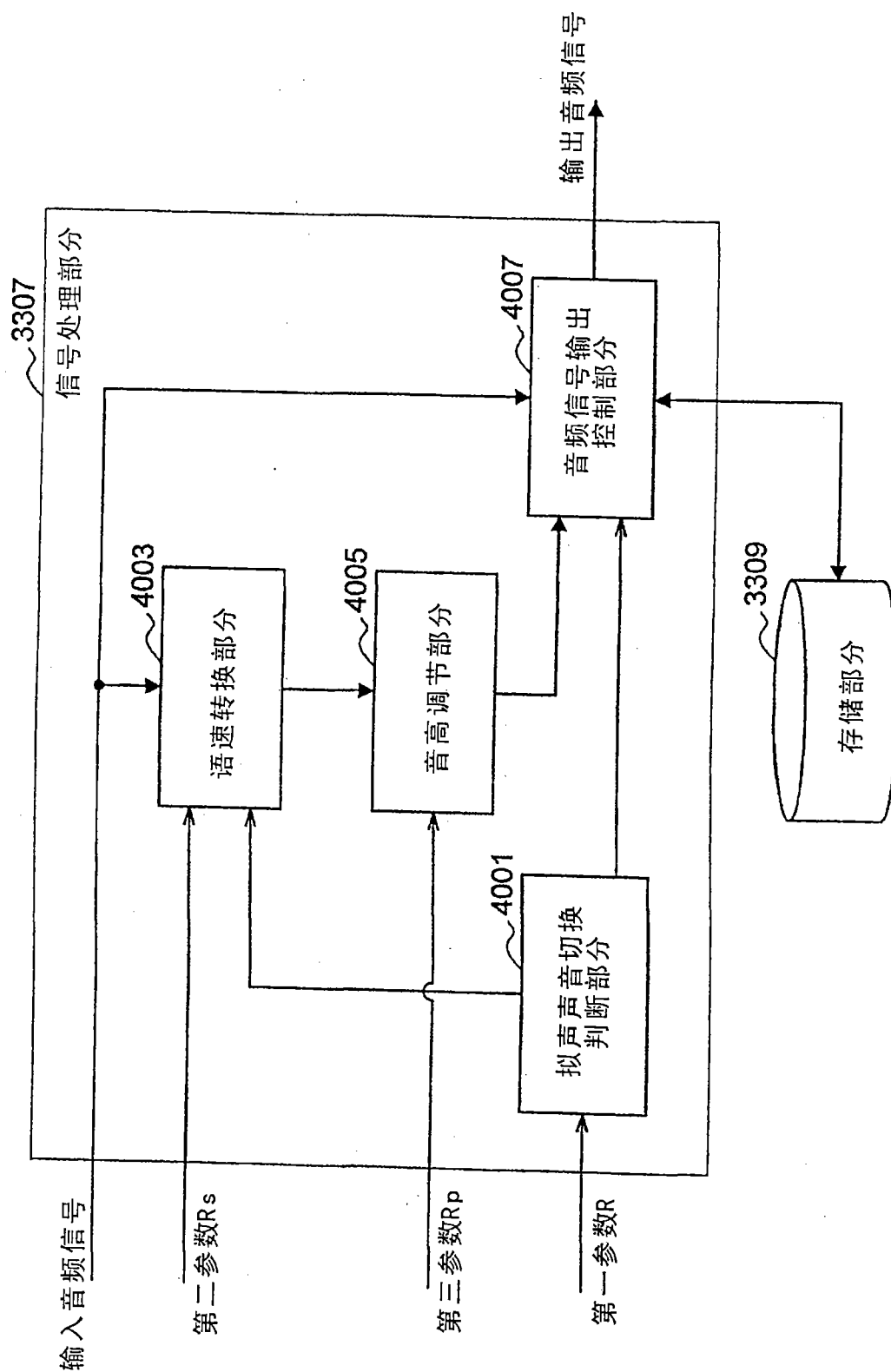


图40

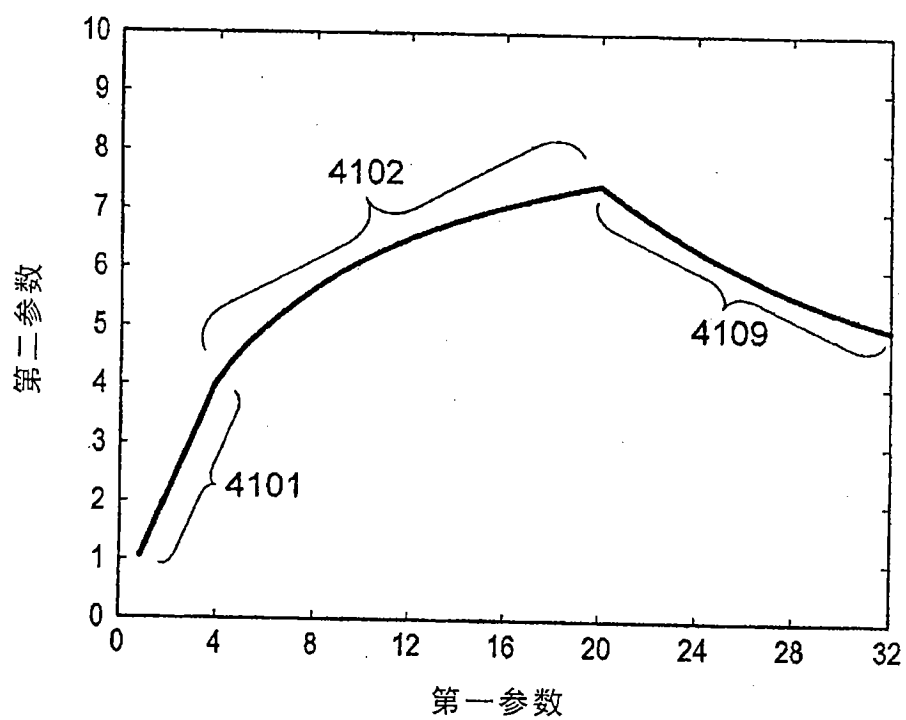


图 41A

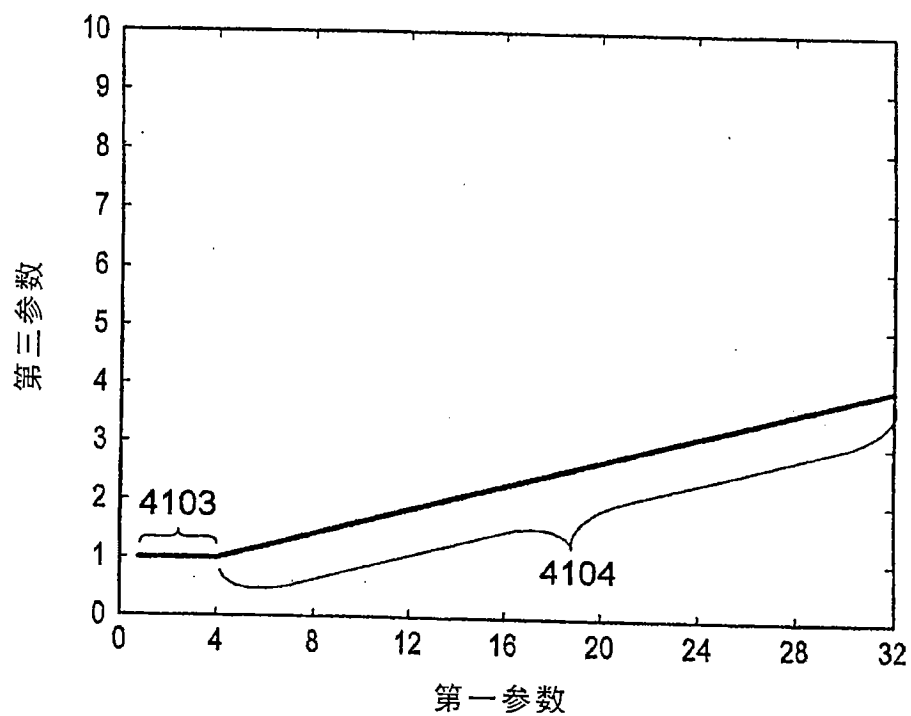


图 41B

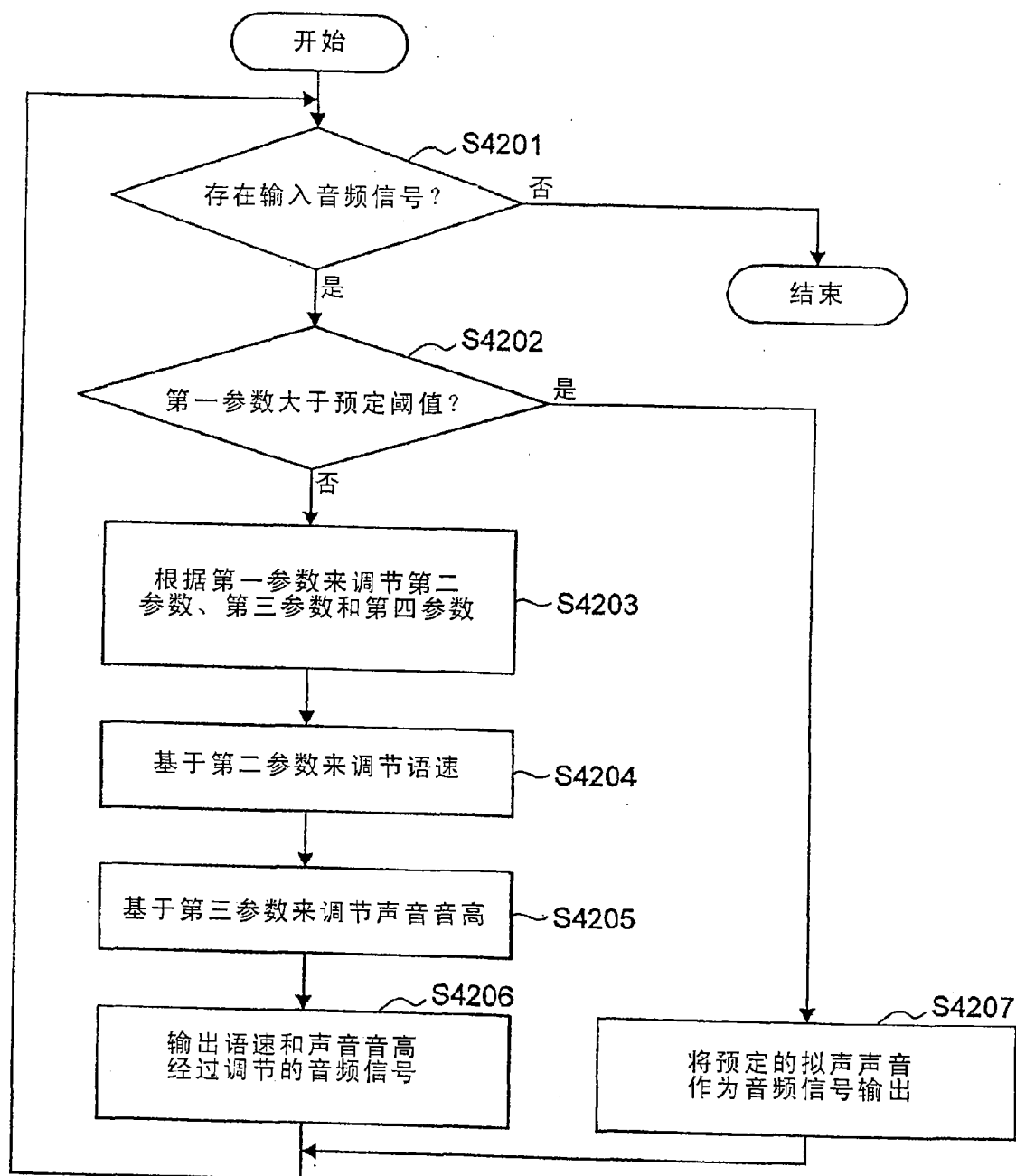


图 42

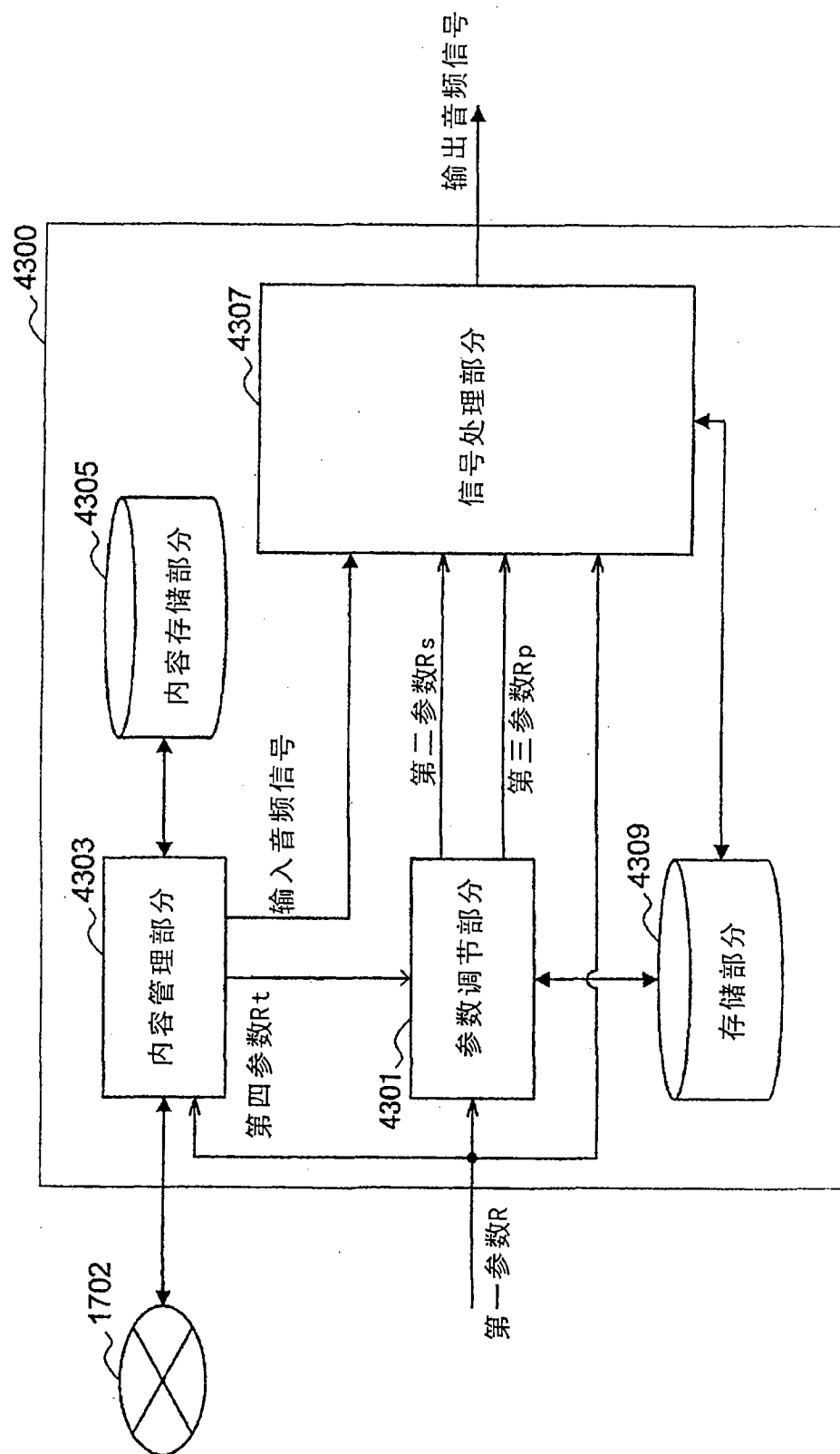


图43

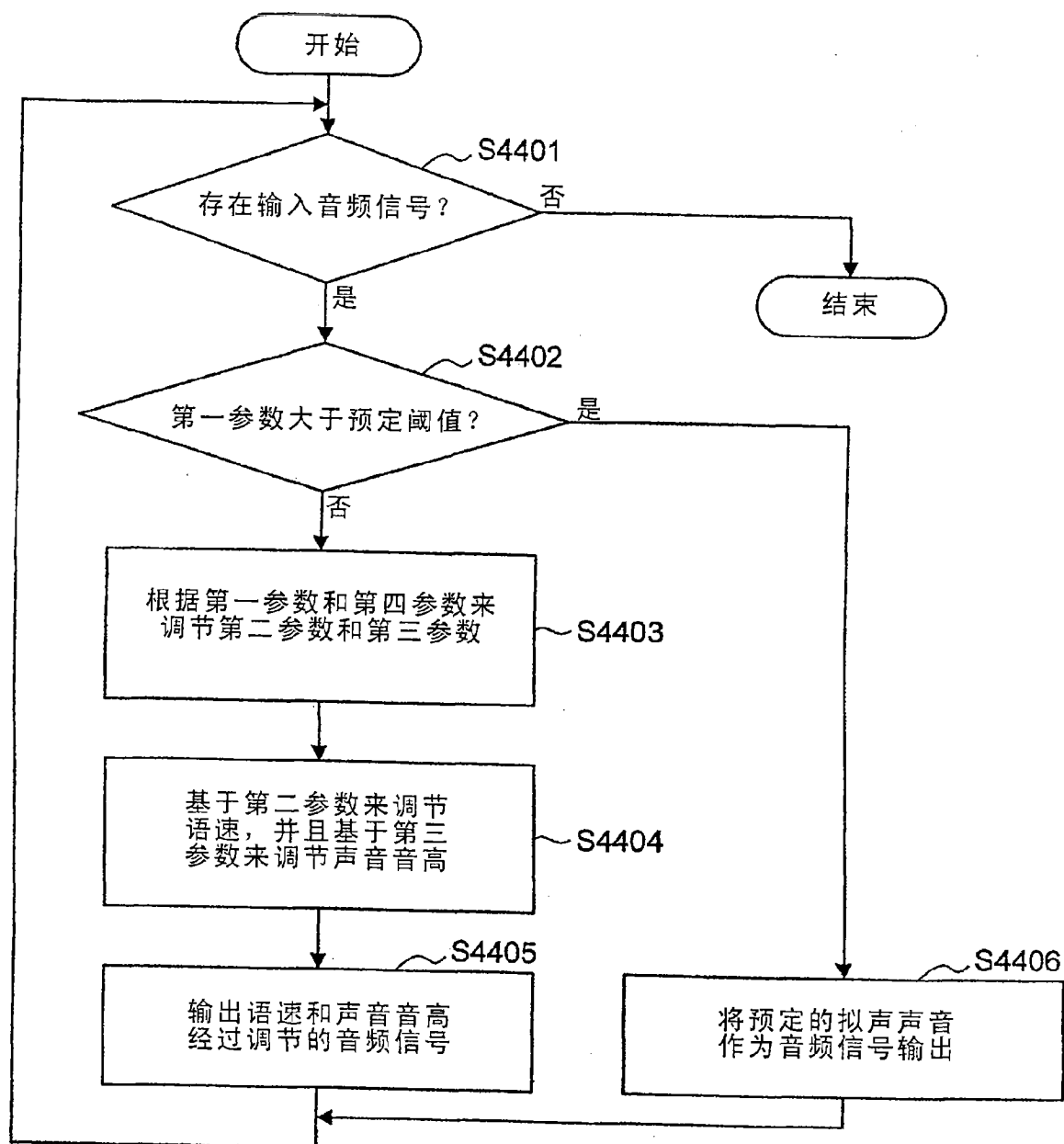


图 44

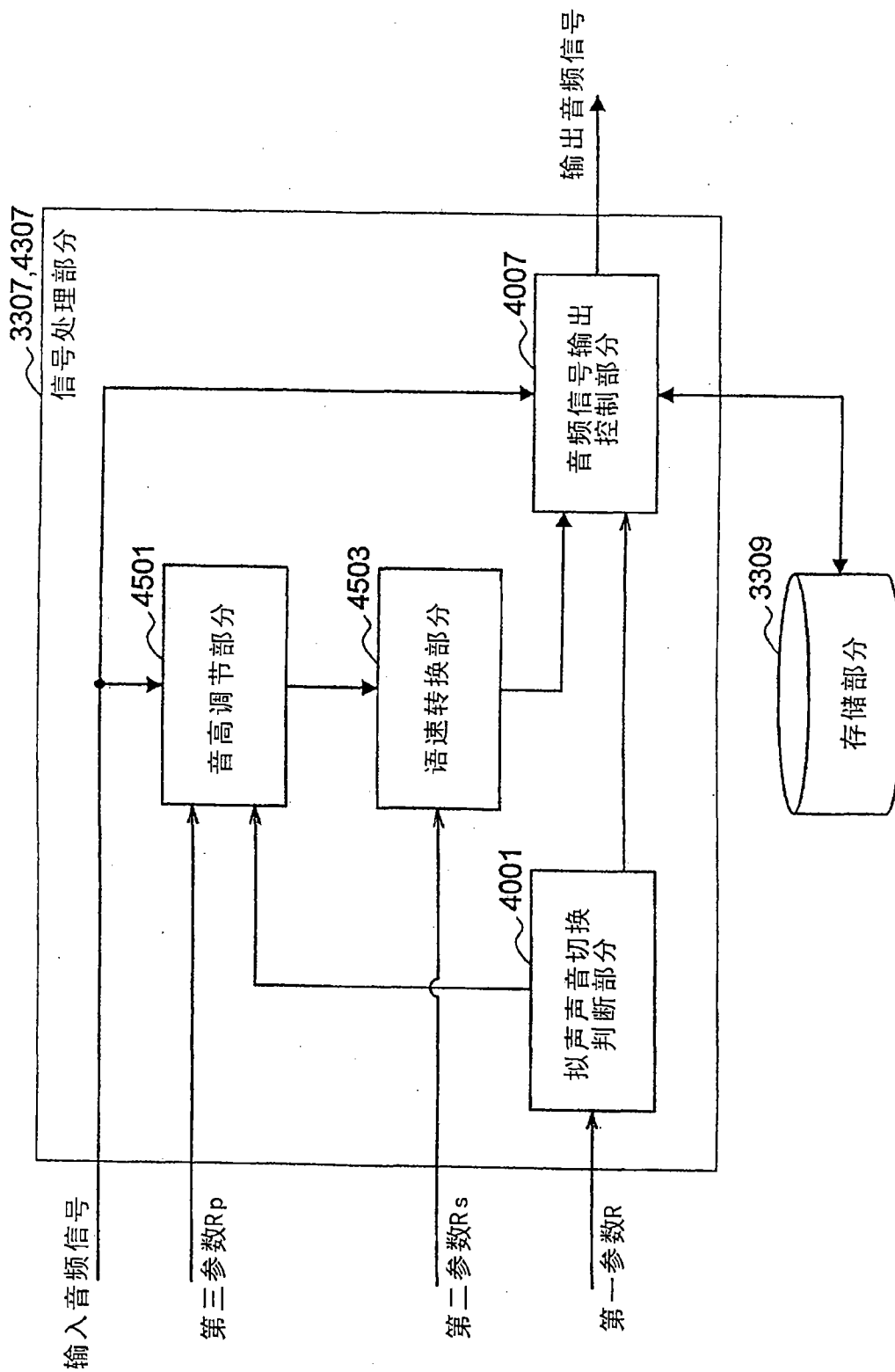


图45

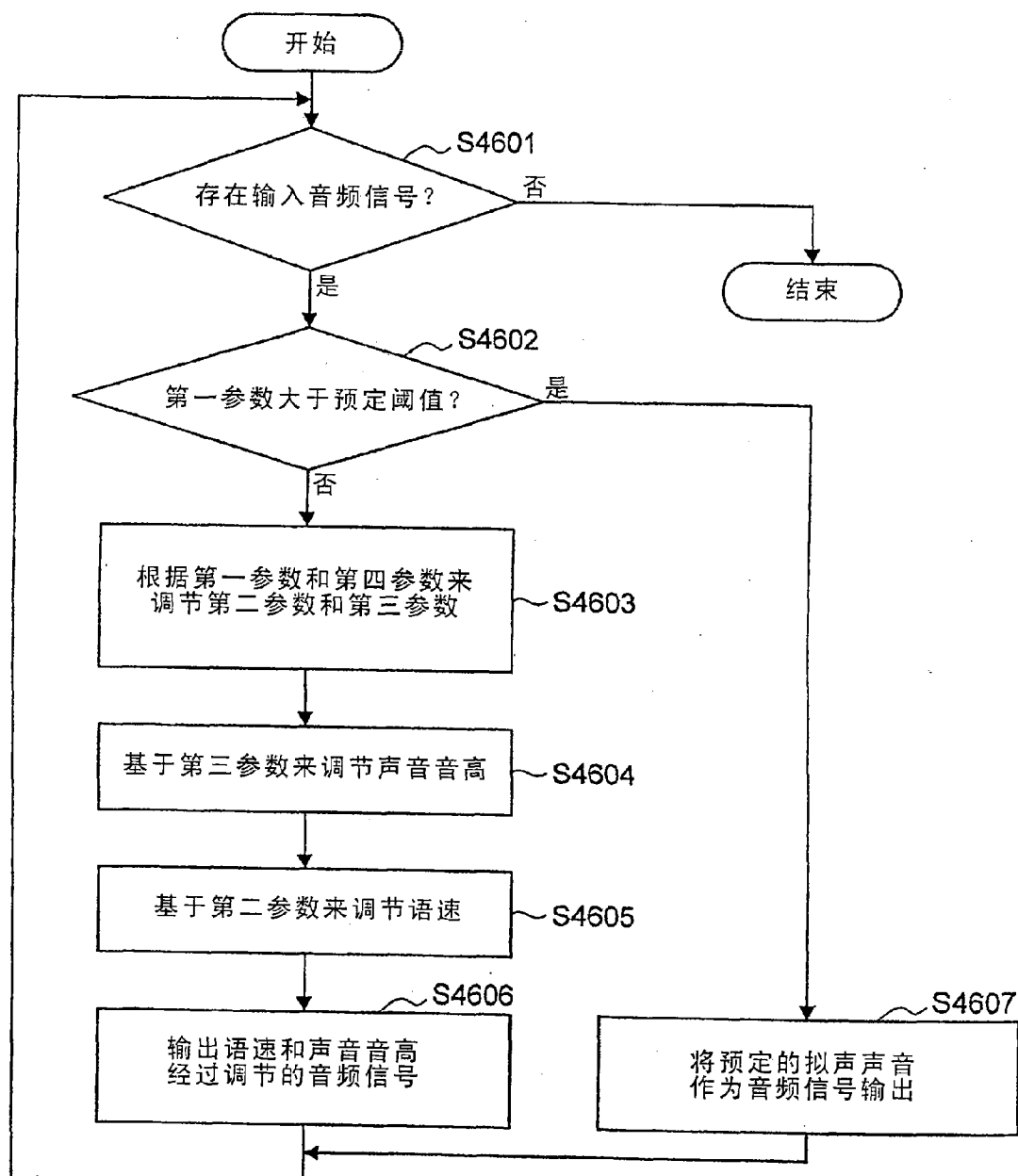


图 46

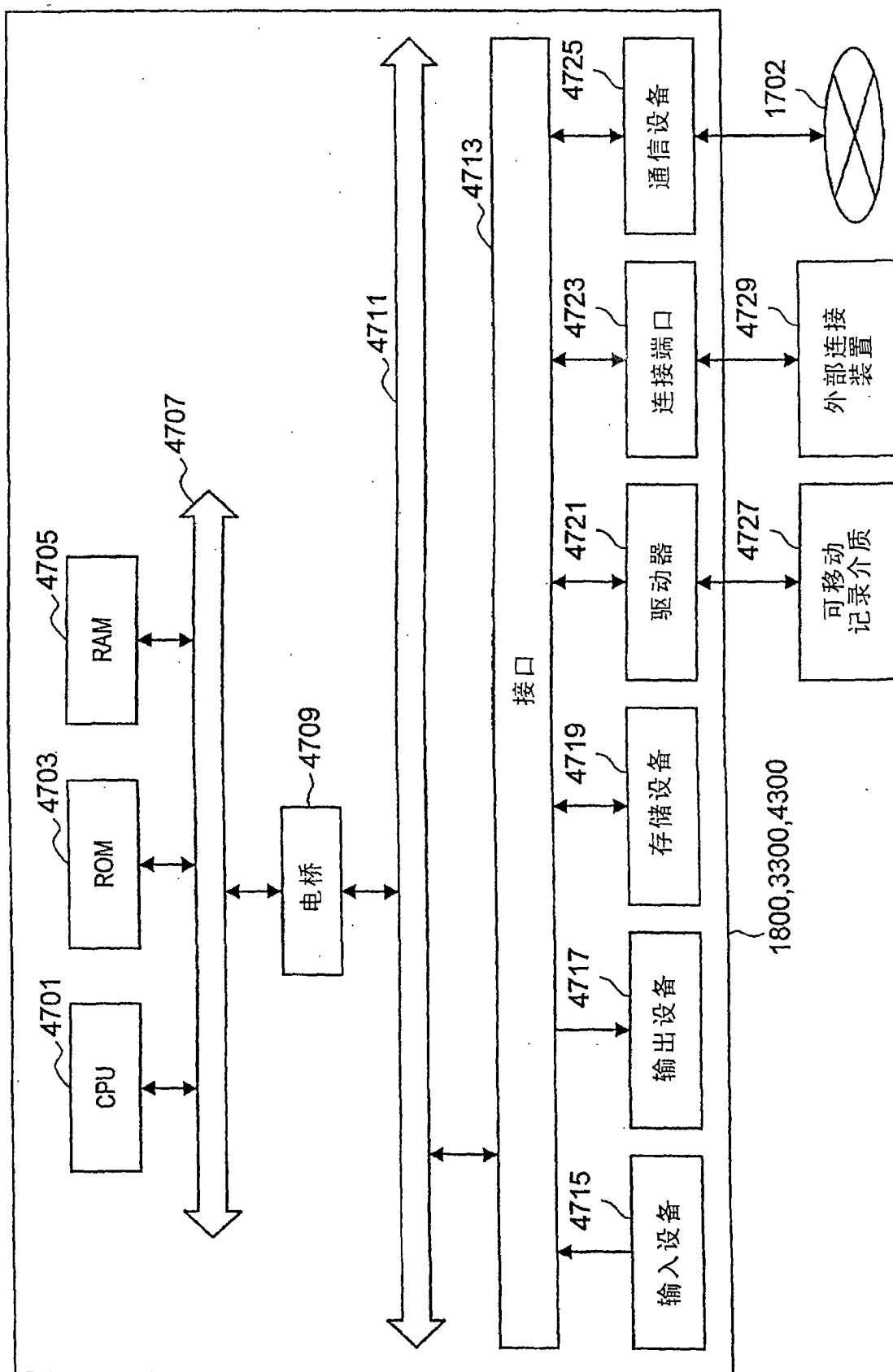


图47