

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G10L 21/04 (2006.01)



# [12] 发明专利说明书

专利号 ZL 200510112850.1

[45] 授权公告日 2010年1月6日

[11] 授权公告号 CN 100578623C

[22] 申请日 2005.10.14

[21] 申请号 200510112850.1

[30] 优先权

[32] 2005.6.22 [33] JP [31] 2005-181843

[73] 专利权人 富士通株式会社

地址 日本神奈川县

[72] 发明人 远藤香绪里 大田恭士 外川太郎

[56] 参考文献

CN1601912A 2005.3.30

EP0534410A2 1993.3.31

US2004156397A1 2004.8.12

审查员 邓 薇

[74] 专利代理机构 北京三友知识产权代理有限公司

代理人 黄纶伟

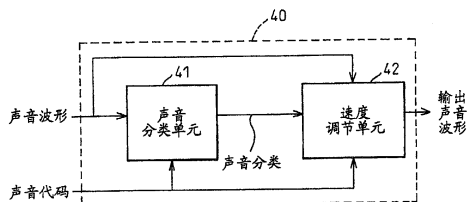
权利要求书 3 页 说明书 14 页 附图 13 页

[54] 发明名称

语音速度转换装置以及语音速度转换方法

[57] 摘要

语音速度转换装置以及语音速度转换方法。本发明涉及语音速度转换，并且提供了一种语音速度转换装置以及一种语音速度转换方法，其用于针对包含有声音的信号，在不使音质劣化、不改变特征的情况下改变声音速度。该语音速度转换装置包括：声音分类单元，向该单元输入声音波形数据以及基于线性预测的声音代码，并且该声音分类单元基于输入信号的特征对所述输入信号进行分类；以及，速度调整单元，该单元基于所述分类，选择利用声音波形的速度转换处理以及利用声音代码的速度转换处理中的一种或者两种处理，并且利用所选择的速度转换方法改变所述输入信号的语音速度。



1、一种语音速度转换装置，其包括：

声音分类单元，向该单元输入声音波形数据和基于线性分析的声音代码，并且所述声音分类单元基于输入信号的特征对所述输入信号进行分类；以及

速度调整单元，其基于所述分类，选择利用所述声音波形的速度转换处理和利用所述声音代码的速度转换处理中的一种或者两种处理，并且通过使用所选择的速度转换方法来改变所述输入信号的语音速度，其中

所述利用声音波形的速度转换处理包括通过以下步骤来转换所述输入信号的语音速度：

计算声音波形的音调周期；以及

按照计算出的音调周期对所述声音波形进行重复或者薄化，并且其中

所述利用声音代码的速度转换处理包括通过以下步骤来转换所述输入信号的语音速度：

通过对所述声音代码的帧的残差信号进行薄化或者插入所述声音代码的新帧的残差信号来修改所述声音代码的残差信号；

通过对所述声音代码的帧的线性预测系数进行薄化或者插入所述声音代码的新帧的线性预测系数来修改所述声音代码的线性预测系数；以及

用修改后的线性预测系数对修改后的残差信号进行滤波。

2、根据权利要求1所述的语音速度转换装置，其中

所述速度转换处理包括：基于所述分类对速度转换等级进行调整。

3、根据权利要求1所述的语音速度转换装置，其中

所述速度调整单元基于所述输入信号的周期性，选择所述利用声音波形的速度转换处理和所述利用声音代码的速度转换处理中的任意一个来改变所述输入信号的语音速度。

4、根据权利要求3所述的语音速度转换装置，其中如果所述输入信号是非周期性的，则所述速度调整单元基于所述输入信号的相似性和稳定性来调整速度转换等级。

5、根据权利要求3所述的语音速度转换装置，其中如果所述输入信号是周期性的，则所述速度调整单元基于所述输入信号的稳定性来调整速度转换等级。

6、根据权利要求1所述的语音速度转换装置，其中所述声音分类单元基于周期性、稳定性以及相似性对所述输入信号进行分类。

7、一种语音速度转换方法，其包括如下步骤：

输入声音波形数据和基于线性预测的声音代码，并且基于输入信号的特征对所述输入信号进行分类；以及

基于所述分类，选择利用所述声音波形数据的速度转换处理和利用所述声音代码的速度转换处理中的一种或者两种处理，并且利用所选择的速度转换方法来改变所述输入信号的语音速度，其中

所述利用声音波形的速度转换处理包括通过以下步骤来转换所述输入信号的速度：

计算声音波形的音调周期；以及

按照计算出的音调周期对所述声音波形进行重复或者薄化，并且其中

所述利用声音代码的速度转换处理包括通过以下步骤来转换所述输入信号的速度：

通过对所述声音代码的帧的残差信号进行薄化或者插入所述声音代码的新帧的残差信号来修改所述声音代码的残差信号；

通过对所述声音代码的帧的线性预测系数进行薄化或者插入所述声音代码的新帧的线性预测系数来修改所述声音代码的线性预测系数；以及

用修改后的线性预测系数对修改后的残差信号进行滤波。

8、根据权利要求7所述的语音速度转换方法，其中

所述速度转换处理包括：基于所述分类对速度转换等级进行调整。

9、根据权利要求7所述的语音速度转换方法，其包括如下步骤：

基于所述输入信号的周期性，选择所述利用声音波形的速度转换处理和所述利用声音代码的速度转换处理中的任意一个来改变所述输入信号的语音速度。

10、根据权利要求7所述的语音速度转换方法，其包括如下步骤：

如果所述输入信号是非周期性的，则基于所述输入信号的相似性和稳定性来调整速度转换等级。

11、根据权利要求7所述的语音速度转换方法，其包括如下步骤：

如果所述输入信号是周期性的，则基于所述输入信号的稳定性来调整速度转换等级。

12、根据权利要求7所述的语音速度转换方法，其中

所述声音分类是基于周期性、稳定性和相似性的对所述输入信号的分类。

## 语音速度转换装置以及语音速度转换方法

### 技术领域

本发明涉及语音速度转换。尤其是本发明涉及一种语音速度转换装置以及一种语音速度转换方法，其用于针对包含有声音的信号，在不降低音质且不改变音色的情况下改变声音速度。

### 背景技术

语音速度转换装置被用于电话系统或者声音再现系统中。通过在再现所接收的声音或者所记录的声音时改变声音的速度，使用者可以对其合适的速度收听所接收或者所记录的内容。例如，当线路另一端的人说话较快，且接电话的人不能容易地理解其声音时，则实时地或者在再现时降低语音速度。利用这一结构，接听者可以很容易地理解语音内容。另一方面，通过在再现时提高声音速度，可以在比实际记录时间更短的时间内听见所记录内容。

图 1 显示了被应用于诸如电话等的声音通信系统的语音速度转换装置的示例。

在图 1 中，电话的接收单元 10 经由数字线等接收声音代码。解码单元 11 将声音代码解码成声音波形信号。包含语音速度转换装置的语音速度转换单元 12 将声音波形信号转换成具有例如更低速度的声音波形信号。诸如受话器的输出单元 13 将接收到的声音输出给外部。当解码单元 11 将声音代码恢复成声音波形时，在本示例中，语音速度转换单元 12 能够直接转换通过接收单元 10 接收到的声音代码的速度，对经速度转换的声音代码进行解码，并且将所解码的声音输入输出单元 13。

作为一种语音速度的转换方法时域谐波换算（time-domain harmonic scaling）是一种公知的方法。根据时域谐波换算，将待改变速度的声音波形以基本频率重复或者将其薄化（thin），因此能够调整速

度。还有通过重复或者薄化波形以转换语音速度的改进方法。一个示例是：把声音分类成几种类型，并且在所分类的声音之间切换速度转换方法。

图 2 显示了利用声音波形的传统语音速度转换装置的结构示例。

在本示例中，声音分类单元 20 将输入声音波形分类为“浊音(voiced sound)”和“非语音(unvoiced sound)”。当所输入声音波形是“浊音”时，音调(pitch)周期计算单元 21 计算“浊音”的音调周期。声音速度转换单元 22 通过基于由声音速度转换单元 22 所计算的音调周期对“浊音”波形输入进行重复或者薄化，来调整声音速度。

根据下述专利文献 1，将声音分类为“元音声(vowel sound)”、“浊辅音(voiced consonant)”、“清辅音(unvoiced consonant)”以及“无音(silence)”。通过按音调周期地对声音波形进行重复或者薄化，来转换“元音声”以及“浊辅音”的速度。根据辅音的特征，不能将“清辅音”扩展或者压缩，或者可通过重复或删除波形以获得预定长度，来进行转换其速度。另一方面，可通过重复或删除波形以获得预定长度，来转换“无音”的速度。

根据下述的专利文献 2，将声音分类为“浊音”、“非语音”以及“无音”。通过按音调周期地重复或者薄化声音波形，来转换“浊音”的速度。不对“非语音”进行处理，并且通过以预定倍率放大或者缩小波形来转换“无音”的速度。

根据下述的专利文献 3，将声音分类为“浊音”、“非语音”以及“无音”。通过按音调周期地重复或者薄化声音波形来转换“浊音”的速度。通过以固定周期(即伪音调)重复或者薄化声音波形，来转换“非语音”的速度。通过以预定的放大及缩小比率重复或薄化波形，来转换“无音”的速度。

图 3 显示了使用声音代码的传统语音速度转换装置的结构的一个示例。

在本示例中，基于对所输入声音的线性预测分析来预先获得所输入声音的残差信号和线性预测系数。音调周期计算单元 30 利用残差信号来

计算所输入信号的音调周期。发声速度转换单元 31 输出基于计算出的音调周期而重复或者薄化的残差信号，从而转换该速度，并且将速度转换信息传送至线性预测系数校正单元 32。

线性预测系数校正单元 32 对与残差信号（该信号是基于速度转换信息而被重复或者薄化的）相对应的线性预测系数进行校正并且将其输出。组合单元 33 利用来自线性预测系数校正单元 32 的线性预测系数对从发声速度转换单元 31 输入的残差信号进行滤波，然后输出经速度转换的声音波形。

下述专利文献 4 描述了一种进行线性预测分析的方法，以把输入的声音分离成线性预测系数和预测残差信号，并且通过按音调周期重复或者薄化含有强音调的预测残差信号，防止了由于音调提取错误而导致音调分析的劣化。当采用线性预测分析时，为了提高音调分析的精度，通过使用音调比声音波形表现得更强的预测残差，对音调进行提取。以所提取的音调周期重复或薄化预测残差。

下述的专利文献 5 描述了一种通过使用声音代码填充 (fill) “0” 来扩展多路径声源，或者通过削减 (cut) “0” 来缩短声源的速度转换方法。

（专利文献 1） 日本专利公开 No. 2612868

（专利文献 2） 日本专利公开 No. 3327936

（专利文献 3） 日本专利公开 No. 3439307

（专利文献 4） 日本专利申请未审公开 No. 11-311997

（专利文献 5） 日本专利公开 No. 3285472

然而，上述传统技术存在下列问题。

（1） 使用声音波形转换速度时所产生的问题

根据专利文献 1，在“清辅音”中，将除了那些被区分为“流音(liquid sound)”、“爆破音和塞擦音 (plosive and affrictive sound)”以及“猝发音 (burst)”的区间以外的区间的波形重复或者薄化。因此，产生了如下问题：由于对波形的重复或薄化而导致出现了最初不存在的周期性，并且使音质下降。

根据专利文献 2，不对“非语音”进行处理。因此，存在如下问题：当将“非语音”扩展或压缩时，其音长与其他区间的音长之间的平衡被破坏，且音质下降。在此情况下，可扩展或压缩的区间变小，且不能实现大的扩展或压缩。根据专利文献 3，因为按固定周期（即伪音调）对“非语音”进行薄化或重复，则产生出现最初不存在的周期性且使音质下降的问题。

(2) 使用诸如线性预测分析的声音代码来转换速度时出现的问题

根据专利文献 4，存在如下问题：在未特别存在音调周期的浊音区间，以不定音调（即极大或极小的音调值的变化）在极长或极短区间中执行重复或者薄化。结果，在线性预测代码（LPC）系数变化的区间中，在 LPC 系数与预测残差之间出现不匹配，因此降低了音质。

根据专利文献 5，通过利用声音代码填充“0”来扩展多路径声源，或者通过削减“0”来进行缩短。此外，还存在的问题是：在没有音调的非语音区间无法调整速度。因此，其音长与其他被扩展或压缩的区间的音长之间的平衡被破坏，且使音质下降。当填充“0”时，扩展或压缩区间减小。从而无法实现大的扩展或压缩。

## 发明内容

根据上述问题，本发明的目的是提供一种语音速度转换装置以及语音速度转换方法，其用于根据所输入声音的特征，通过在利用声音波形数据和基于线性分析而获得的声音代码的速度调整方法与利用声音波形数据及声音代码中的一个的速度调整方法之间进行适当切换，来调整语音速度而不会降低音质。

根据本发明的一个方面，提供了一种语音速度转换装置，其利用声音波形数据和基于线性预测的声音代码来调整语音速度。

根据本发明的另一方面，提供了一种语音速度转换装置，其包括：声音分类单元，向该单元输入声音波形数据及基于线性分析的声音代码，并且基于输入信号的特征对输入信号进行分类；以及速度调整单元，该单元基于所述分类来选择利用声音波形的速度转换处理和利用声音代码



的速度转换处理中的一种或两种处理，并且利用所选择的速度转换方法来改变输入信号的速度，其中，所述利用声音波形的速度转换处理包括通过以下步骤来转换所述输入信号的语音速度：计算声音波形的音调周期；以及按照计算出的音调周期对所述声音波形进行重复或者薄化，并且其中，所述利用声音代码的速度转换处理包括通过以下步骤来转换所述输入信号的语音速度：通过对所述声音代码的帧的残差信号进行薄化或者插入所述声音代码的新帧的残差信号来修改所述声音代码的残差信号；通过对所述声音代码的帧的线性预测系数进行薄化或者插入所述声音代码的新帧的线性预测系数来修改所述声音代码的线性预测系数；以及用修改后的线性预测系数对修改后的残差信号进行滤波。该速度转换处理包括：基于所述分类对速度转换等级进行调整。

根据本发明的另一方面，提供了一种语音速度转换方法，其用于利用声音波形数据以及基于线性预测的声音代码来调整语音速度。

根据本发明的另一方面，提供了一种语音速度转换方法，其包括如下步骤：输入声音波形数据以及基于线性预测的声音代码，并且基于输入信号的特征对该信号进行分类；基于所述分类来选择利用声音波形的速度转换处理以及利用声音代码的速度转换处理中的一种或两种处理；并且利用所选择的速度转换方法来改变输入信号的速度，其中，所述利用声音波形的速度转换处理包括通过以下步骤来转换所述输入信号的速度：计算声音波形的音调周期；以及按照计算出的音调周期对所述声音波形进行重复或者薄化，并且其中，所述利用声音代码的速度转换处理包括通过以下步骤来转换所述输入信号的速度：通过对所述声音代码的帧的残差信号进行薄化或者插入所述声音代码的新帧的残差信号来修改所述声音代码的残差信号；通过对所述声音代码的帧的线性预测系数进行薄化或者插入所述声音代码的新帧的线性预测系数来修改所述声音代码的线性预测系数；以及用修改后的线性预测系数对修改后的残差信号进行滤波。所述速度转换处理包括：基于所述分类对速度转换等级进行调整。

根据本发明，因为声音波形数据和声音代码二者都被使用，所以能

够基于声音特征来选择性地使用声音波形数据和声音代码中的一个或二者。结果，与仅利用声音波形数据和声音代码中的一个的传统实践所获得的音质相比，显著地提高了转换速度后的音质。

根据本发明，根据输入信号的特征对输入信号进行详细分类。根据所述分类，从利用声音波形数据和声音代码中的一个的方法以及利用声音波形数据和声音代码中的二者的方法中适当选择调整语音速度的方法，因此不产生音质的劣化。结果，与仅利用声音波形数据和声音代码中的一个的传统实践所获得的音质相比，显著提高了转换速度后的音质。如后所述，利用声音波形适当地转换“周期性的”区间的速度。当由于残差的重复或删除导致“非周期性且稳定性的”区间具有不连续区间时，可通过使该区间通过线性预测滤波器来薄化该不连续性。利用声音代码适当转换“非周期性及稳定性的”区间的速度。

根据本发明，当同时使用声音波形数据和声音代码时，并且当将加权的速度调整组合在一起时，可以通过进一步降低音质劣化来调整语音速度。

#### 附图说明

通过如下参照附图阐述的说明将使本发明更清楚地被理解，其中

图 1 是示出将语音速度转换装置应用到声音通信系统的示例的示意图；

图 2 是示出利用声音波形的传统语音速度转换装置的结构的一个示例的示意图；

图 3 是示出利用声音代码的传统语音速度转换装置的结构的一个示例的示意图；

图 4 是示出显示根据本发明的语音速度转换装置的基本结构的示意图；

图 5 是示出图 4 中所示的速度转换单元的结构的一个示例的示意图；

图 6 是示出图 5 所示速度调整单元的结构的一个示意图；

图 7 是示出处理流程的一个示例的流程图；

图 8 是图 5 所示的速度调整单元的结构的一例的示意图；  
图 9 是示出图 8 中所示的处理流程的一例（1）的流程图；  
图 10 是示出图 8 中所示的处理流程的一例（2）的流程图；  
图 11 是根据本发明的一个实施例的处理流程的示意图；  
图 12 是示出图 11 中所示的处理的基本流程的示意图；  
图 13 是示出由声音分类单元执行的输入信号分类处理流程的一例的流程图；  
图 14 是示出图 13 所示的关于周期性的判断的一例的流程图；  
图 15 是示出图 13 所示的关于稳定性的判断的一例的流程图；  
图 16 是示出图 13 所示的关于相似性的判断的一例的流程图；  
图 17 是示出利用代码的速度调整（在压缩时）的一例的流程图；  
而  
图 18 是示出利用代码的速度调整（在扩展时）的一例的流程图。

#### 具体实施方式

图 4 是出根据本发明的语音速度转换装置的基本结构的示意图。

在图 4 中，向速度转换单元 40 输入声音波形和声音代码。速度转换单元 40 根据声音的特征，利用声音波形以及声音代码中的一个或者二者来调整语音速度，并且输出经速度调整的声音。

图 5 是图 4 所示的速度转换单元 40 的结构示例的示意图。

在图 5 中，声音分类单元 41 根据声音的特征对输入声音进行分类。速度调整单元 42 根据声音分类结果，在利用声音波形及声音代码二者的速度调整方法和利用声音波形及声音代码中的一个的速度调整方法中进行适当选择。速度调整单元 42 利用所选择的方法来调整速度，并且输出经调整速度的声音。声音分类单元 41 安装有中央处理器（CPU）和数字信号处理器（DSP），并且由包括只读存储器（ROM）、随机存取存储器（RAM）以及输入/输出（I/O）外围装置的常规 CPU 电路组成。如下面的结构框图所示，速度调整单元 42 也有类似结构。

图 6 是示出图 5 所示的速度调整单元 42 的结构示例的示意图。图 7

是示出处理流程的一个示例的流程图。

在本示例中，利用声音波形数据以及通过线形分析操作所获得的语音代码中的一个来调整语音速度。输入选择单元 43 基于来自声音分类单元 41 的声音分类，选择声音波形数据以及声音代码中的一个，以输入一帧（步骤 S101 和 S102）。

同样，基于声音分类，将后一级互锁开关 44 及 47 转换到声音波形速度调整单元 45 或者声音代码速度调整单元 46（步骤 S103）。速度调整单元 45 或速度调整单元 46（通过输入选择单元 43 将互锁开关 44 及 47 切换到其处）利用相对应的声音波形或声音代码，来执行速度调整处理（步骤 S104 或 S105），并且向输出单元 48 输出经速度调整的声音波形。

因为基于声音分类对用来速度调整的声音波形或者声音代码进行了适当选择，所以与仅使用声音波形或者声音代码来转换速度时相比，显著降低了在转换速度后的音质的劣化。

图 8 是示出图 5 所示的速度调整单元 42 的结构的一示例的示意图。图 9 和 10 是图 8 所示的处理流程的示例的流程图。

在本例中，通过同时使用由线形预测操作所获得的语音波形数据和语音代码二者，来调整语音速度。因此，图 7 所示的输入选择单元 43 不是必需的。将所输入的语音波形以及语音代码直接分别地施加给速度调整单元 45 和速度调整单元 46。将通过速度调整单元 45 对语音波形进行速度转换所获得的语音波形以及通过速度调整单元 46 对语音代码进行速度转换所获得的语音波形输入下一级的输出生成单元 49（步骤 S201-S204）。

输出生成单元 49 基于来自声音分类单元 41 的声音分类，计算两个输入语音波形的权重（步骤 S301 和 S302），将加权的两项语音波形相加，然后输出相加后的结果（步骤 S403）。作为该方法应用的示例，考虑了从使用语音波形的速度调整区间到使用语音代码的速度调整区间的切换。

在此情况下，首先，将权重“1”赋予从使用语音波形的速度调整单元 45 输入的语音波形，将权重“0”赋予从使用语音代码的速度调整单元 46 输出的波形。然后，在预定的区间切换时间内，将来自速度转换单

元 45 的声音波形的权重由“1”逐渐降到“0”。另一方面，将来自速度调整单元 46 的声音波形的权重由“0”逐渐增加到“1”。权重可呈线性或者指数地变化。结果，在本示例中，可以充分地限制由于在声音波形区间和声音代码区间之间进行切换时生成的波形不连续性所造成的噪音。

图 11 是根据本发明一个实施例的处理流程的示意图。利用图 5 所示的声音分类单元 41 和速度调整单元 42 所执行的操作流程来解释该操作。

在本示例中，声音分类单元 41 首先基于帧是否包含有声音将声音分类为“有声音 (voice)”和“非声音 (nonvoice)”(步骤 S401 至 S403)。例如，当所输入信号的短时间能量持续预定时间或者更长时，声音分类单元 41 判定该帧包含有声音。接着，对判定为声音的区间更详细地进行分类。在本示例中，将浊音分类为“周期性的”，而非语音（例如环境噪音）分类为“非周期性的”(步骤 S404)。通过考虑电平变化将“有声音”进一步分类为“周期性且稳定的”和“周期性且不稳定的”(步骤 S405)。

通过考虑电平变化和猝发音，将非语音可以进一步分类为“非周期性、稳定且相似的”和“非周期性、稳定且不相似的”(步骤 S409 和 S410)。此外，通过考虑爆破音等将非语音分类为“非周期性且非稳定的”(步骤 S413)。还可以将类似于上述分类的分类应用于被判定为非语音的区间。

速度调整单元 42 基于上述分类结果，选择适合各个分类的速度调整方法，并且将方法切换到所选择的速度调整方法。在本示例中，利用声音波形，对被判定为“有声音”的区间中的被分类为“周期性且稳定的”区间的速度进行调整。将该速度调整到中间调整等级(步骤 S406)。另一方面，利用声音波形，对被判定为“有声音”的区间中的被分类为“周期性且不稳定的”区间的速度进行调整。将该速度调整到较低调整等级(步骤 S407)。

利用声音代码，对被判定为“声音”的区间中的被分类为“非周期性的”区间的速度进行调整。然而，不对被分类为“非周期性、稳定且相似的”和“非周期性且不稳定的”的区间的速度进行调整。利用声音波形对被判定为“非声音”区间的速度进行调整。将该速度调整到较高

调整等级。

当声音分类单元 41 使用“周期性”、“稳定性”以及“相似性”来对声音进行详细分类时，本示例中的速度调整单元 42 根据该分类，在“周期性”区间中利用声音波形来转换速度（步骤 S404 中的“是”之后）。除了不执行速度转换（步骤 S411 和 S413）的情况以外，声音分类单元 41 在“非周期性”区间利用声音代码来转换速度（步骤 S408 中的“否”之后）。

在周期性区间中，通过根据周期对声音波形进行重复或删除，可以转换速度而不明显劣化音质。然而，在周期性区间中使用声音代码时，对所输入声音的残差信号的重复或删除会影响在线性预测滤波之后的态，且在预测系数与残差信号之间出现不匹配。因此，在周期性区间利用声音波形转换了速度。

另一方面，出于下列原因，在非周期性区中利用声音代码来转换速度。在“非周期性且稳定的”区间（步骤 S409 中的“是”之后），当利用声音波形调整速度时，该波形由于波形的重复或删除而变得不连续。此外，会出现在最初不存在的周期性，且使声音劣化。当在该区间使用声音代码时，即使由于残差的重复或删除而出现了不连续性，该不连续性也会通过最终使该声音通过线形预测滤波而被薄化。“稳定”区间在不包括的滤波器的上升下降区间的频率特性上变化很小。因此，由于残差的重复或删除而导致的对线形预测滤波的状态的影响几乎没有，从而不容易使音质劣化。

出于下列原因，对速度调整单元 42 所执行的速度调整的等级进行确定。

在“非声音”区间中（步骤 S408），速度调整单元 42 搜索在提高速度和降低速度时非声音区间的两端都平滑相连而无间断的声音波形部分。速度调整单元 42 删除夹在这些非声音区间中间的所有区间。在此情况下，速度调整等级变为“高”。

在“周期性且稳定的”区间中（步骤 406），速度调整单元 42 通过在声音信号的周期性且稳定的区间中利用声音波形进行重复或者薄化，

来调整速度而不使音质劣化。在此情况下，当执行重复或薄化的次数变得极端大时，则出现不自然。因此，将速度调整等级设为“中”。“周期性且不稳定的”区间（步骤 S407）具有像声音信号的电平变化的周期性，但能量有所变化。因此，在利用声音波形进行周期性地重复或薄化时，速度调整单元 42 设定速度调整等级为“低”以减少由于能量变化而导致的劣化。

“非周期性、稳定且不相似的”区间（步骤 S112）是具有无关联的信号稳定延续的区间。速度调整单元 42 在该区间中利用声音代码来调整速度。在此情况下，能够通过随机生成固定密码本（codebook），来调整速度（即能够使速度降低）而不生成新的周期性。此外，能够通过压缩（删除）残差信号后利用线形预测滤波来生成输出信号，来限制不连续性。

另一方面，“非周期性、稳定且相似的”区间（步骤 S111）和“非周期性且不稳定的”区间（步骤 S113）是信号变化较大的区间，且声音容易因为速度调整而劣化。因此，速度调整单元 42 不对该区间的速度进行调整。根据本发明，声音分类单元 41 对输入声音进行分类，而速度转换单元 42 选择性地使用速度转换方法。因此，能够增大声音的扩展及压缩区间的比例，而并不使音质劣化。

下面说明上述实施例的详细处理内容。

图 12 是显示图 11 所示的处理的基本流程的流程图。

在图 12 中，图 4 所示速度转换单元 40（即图 5 所示的声音分类单元 41 和速度调整单元 42）首先输入输入信号的一帧（即声音波形和通过执行声音波形的线性预测转换所获得的语音代码）（步骤 S501）。声音分类单元 41 对图 11 所示的输入信号进行分类（步骤 S502），并且速度调整单元 42 基于该分类执行图 11 所示速度转换处理（步骤 S503）。速度转换单元 40 持续上述处理直到输入帧的序列结束（步骤 S504）。

图 13 是通过声音分类单元 41 执行的输入信号分类处理的流程图（图 12 中的步骤 S502）。

在本示例中，基于关于有声音和非声音的判断、以及关于有/无周期

性、有/无稳定性以及有/无相似性的判断，对所输入信号进行分类。首先，将所输入信号大致地分类为“有声音”区间以及“非声音”区间。将被判定为“有声音”的区间进一步分类为“周期性的”区间、“非周期性且稳定的”区间以及“非周期性且不稳定的”区间（见图 11）。

因此，声音分类单元 41 输入声音波形和声音代码的一帧（步骤 S601），并且将所输入信号分类为包含声音的有声音区间和不包含声音的非声音区间（步骤 S602）。接着，声音分类单元 41 在被判定为“有声音”的区间中判断有/无周期性、有/无稳定性以及有/无相似性（步骤 S603 到 S605）。声音分类单元 41 基于判断结果对输入信号进行分类（步骤 S606）。在本示例中，分类项目并不局于周期性、稳定性以及相似性，也可以使用其他分类项目。不需要对未分类项目进行判定。

图 14 是图 13 所示的关于周期性的判断（S603）的一个示例的流程图。

在本示例中，将计算自动相关系数的通用方法应用于声音波形。对输入帧进行抽样，并且计算自动相关系数取最大值的频率（步骤 S701 至 S703）。基于该频率与在紧邻的前一帧中使自动相关系数取最大值的频率之间的差异，来判断周期性（步骤 S704）。例如，将预定的阈值与该差异进行比较。当该差异与阈值相等或者比阈值小时，将该区间判定为“周期性的”（步骤 S705）。在其他情况下，将该区间判定为“非周期性的”。

图 15 是图 13 所示的关于稳定性的判断的一个示例的流程图。

在本示例中，使用声音代码来计算能量。首先，输入声音代码的一个帧，然后计算线性预测系数的变化（标准偏差(SD)）（步骤 S801 和 S802）。为此，根据下列公式（1）来计算线性预测系数的 SD。

$$SD = \frac{1}{n} \sum_{i=1}^n (C_i - P_i)^2 \quad (1)$$

其中，n 代表线性预测分析次数， $C_i$  代表当前帧的线性预测系数（第 i 次），而  $P_i$  代表前一帧的线性预测系数（第 i 次）。

接着，根据下述公式（2）计算能量（POW）（步骤 S803）。

$$POW = \frac{1}{m} \sum_{i=1}^m A_i^2 \quad (2)$$



其中,  $m$  代表  $m$  帧的抽样数量, 而  $A_i$  代表当前帧的振幅 (第  $i$  个抽样)。

接着, 根据下述公式 (3) 计算能量的变化 (DP) (步骤 S804)。

$$DP = POW_t - POW_{t-1} \quad (3)$$

其中,  $POW_t$  代表当前帧的能量, 以及  $POW_{t-1}$  代表前一帧的能量。

最后, 基于上述计算结果判断稳定性 (步骤 S805)。在本示例中, 当 SD 与预定阈值相等或者比该值小, 并且当 DP 与预定阈值相等或者比该值小时, 将该区间判定为“稳定的”。在其他情况下, 将该区间判定为“不稳定的”。为判断下一帧, 存储当前帧的能量以及线性预测系数 (步骤 S806)。

图 16 是图 13 所示的关于相似性判断 (步骤 S605) 的一个示例的流程图。

在本示例中, 使用与参照图 14 所说明的相同的自动相关系数来判断相似性。首先, 输入输入信号的声音波形的一帧 (步骤 S901)。其次, 计算自动相关系数, 并且计算该自动相关系数的最大值 (步骤 S902 和 S903)。将自动相关系数的最大值与预定阈值进行比较。当自动相关系数的最大值等于或者大于预定阈值时, 将该区间判定为“相似的”。否则, 将该区间判定为“不相似的”。

下面说明通过速度调整单元 42 执行的速度转换 (图 12 中的步骤 S503) 的详细处理。在图 17 和图 18 所示的示例中说明了使用声音代码执行的处理 (见图 3)。在进行该处理之前, 速度调整单元 42 基于通过声音分类单元 41 执行的分类的结果, 在图 11 所示的流程 (步骤 S406、S407、S408、S411、S412 以及 S413) 中选择一个终端处理。基于时域谐波换算算法等的现有方法, 执行利用声音波形的处理 (见图 2)。

图 17 是示出利用代码的速度调整 (在压缩时) 的一个示例的流程图。

在本示例中, 速度调整单元 42 首先输入声音代码的一帧 (步骤 S1001)。接着, 从前一帧和当前帧, 薄化前一帧的残差信号。结果, 根据这两个帧的残差信号生成一个帧的残差信号 (步骤 S1002)。同时, 从前一帧和当前帧, 薄化紧邻的在先帧的线性预测系数。因此, 根据这两

个帧的线性预测系数来生成一个帧的线性预测系数（步骤 S1003）。将所生成的一个帧的残差信号和所生成的一个帧的线性预测系数输入给线性预测滤波器。因此，通过组合生成了由于压缩而导致速度增大的声音波形。

图 18 是示出利用代码的速度调整（在扩展时）的一个示例的流程图。

在本示例中，速度调整单元 42 首先输入声音代码的一个帧（步骤 S1101）。在此情况下，利用前一帧的残差信号以及当前帧的残差信号来生成一个帧的新的残差信号。因此，将总和为 1 的权重系数乘以前一帧的残差信号以及当前帧的残差信号。将加权残差信号进行相加，以生成新的残差信号。将所生成的残差信号插入在前一帧的残差信号与当前帧的残差信号之间，由此生成三个帧的残差信号（步骤 S1102）。在编码系统具有密码本的情况下，随机地生成密码本的索引，从而生成一个帧的新的残差信号。

接着，对前一帧的线性预测系数和当前帧的线性预测系数进行内插，以生成新的线性预测系数。将所生成的线性预测系数插入在前一帧的线性预测系数与当前帧的线性预测系数之间，因此生成三个帧的线性预测系数（步骤 S1103）。在编码系统具有密码本的情况下，随机地生成密码本的索引，从而生成一个帧的新的残差信号。最后，将所生成的这三个帧的残差信号以及所生成的这三个帧的线性预测系数输入线性预测滤波器。因此，通过组合生成了通过扩展而导致速度降低的声音波形。

如上所述，根据本发明，因为使用了声音波形数据和声音代码二者，所以能够基于声音的特征为选择性地使用信息。与通过仅使用声音波形数据和声音代码中的一个的转换速度所获得的音质相比，能够提高速度转换后的音质。此外，将所输入信号分类为几种声音。基于对声音的分类，能够通过使用声音波形数据和声音代码中的一个或二者的方法来转换输入信号的速度，从而降低了音质的劣化。与通过仅使用声音波形数据和声音代码中的一个的转换速度所获得的音质相比，能够提高速度转换后的音质。

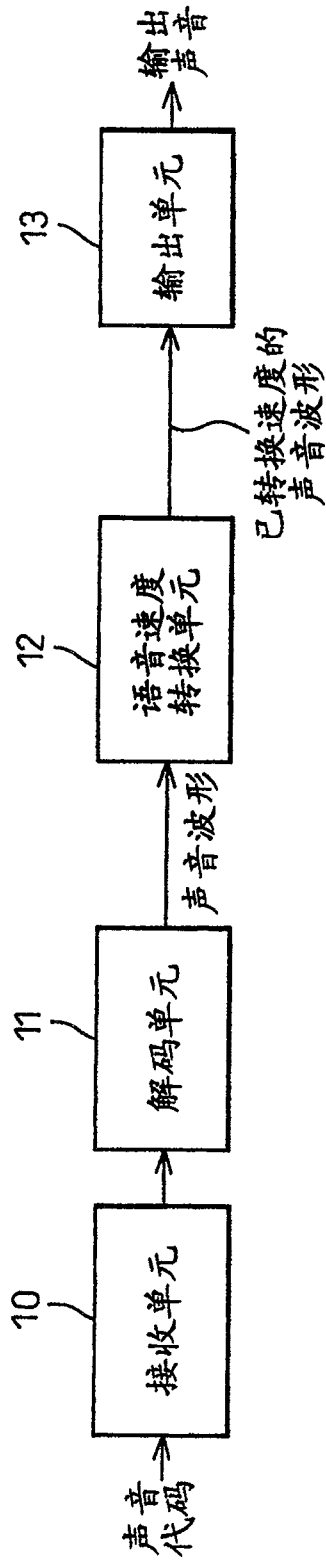


图 1

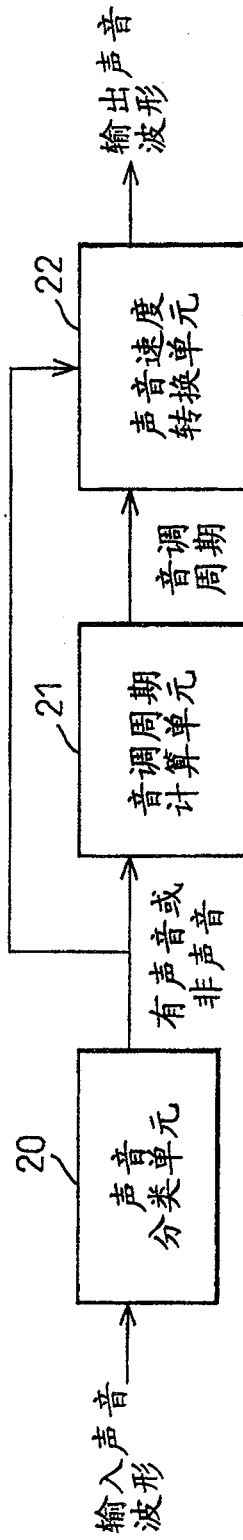


图 2

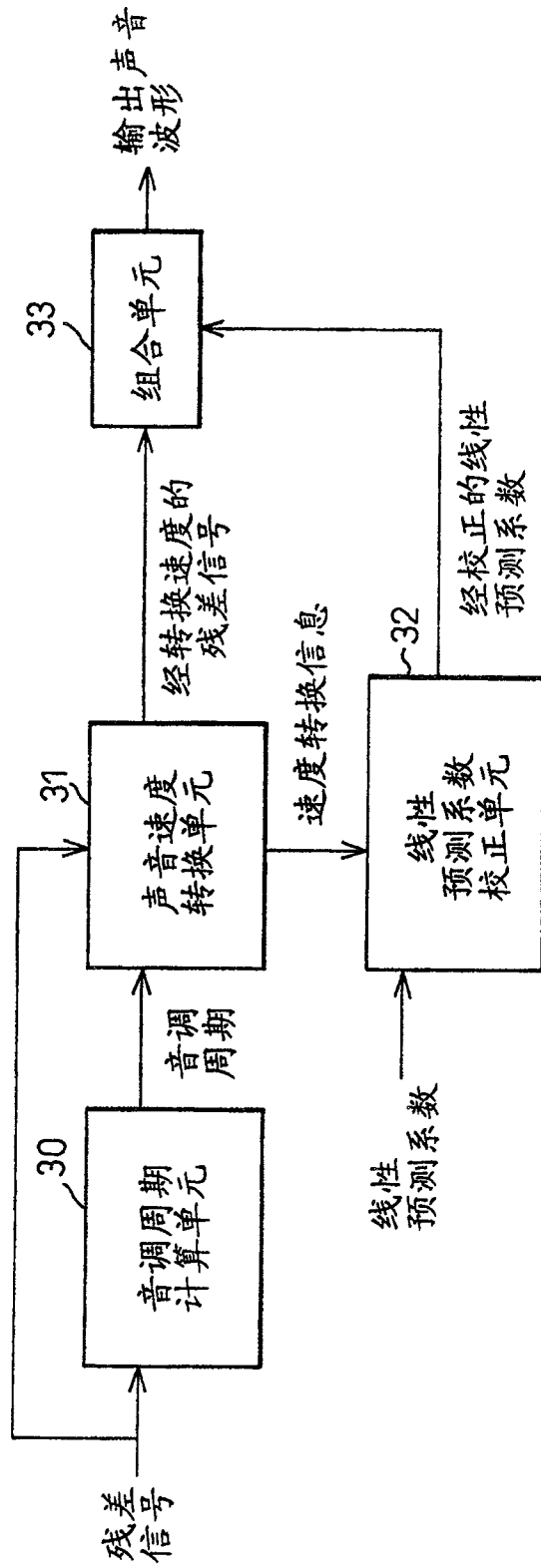


图 3

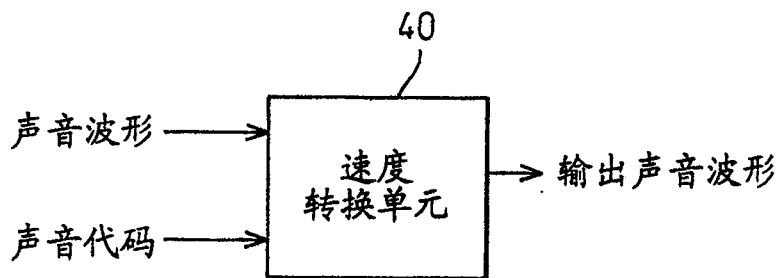


图 4

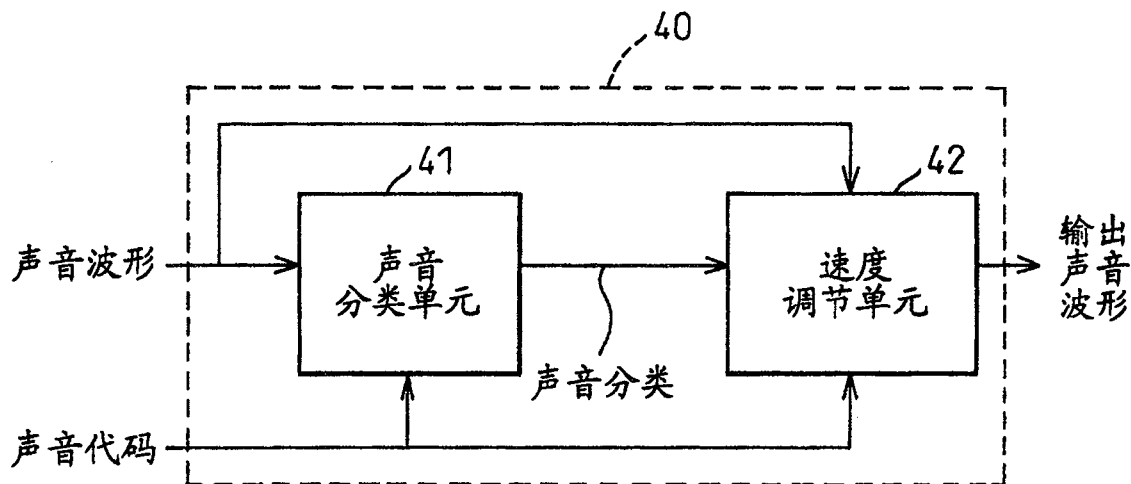


图 5

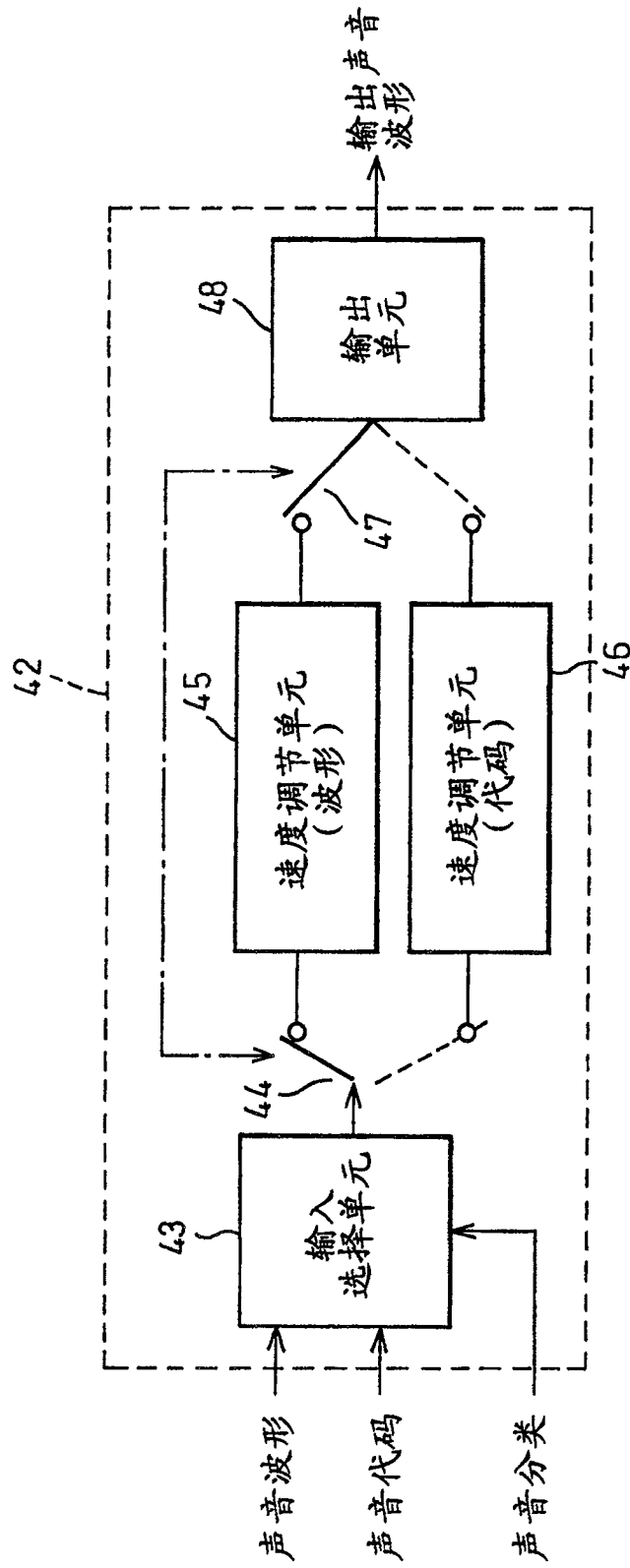


图 6

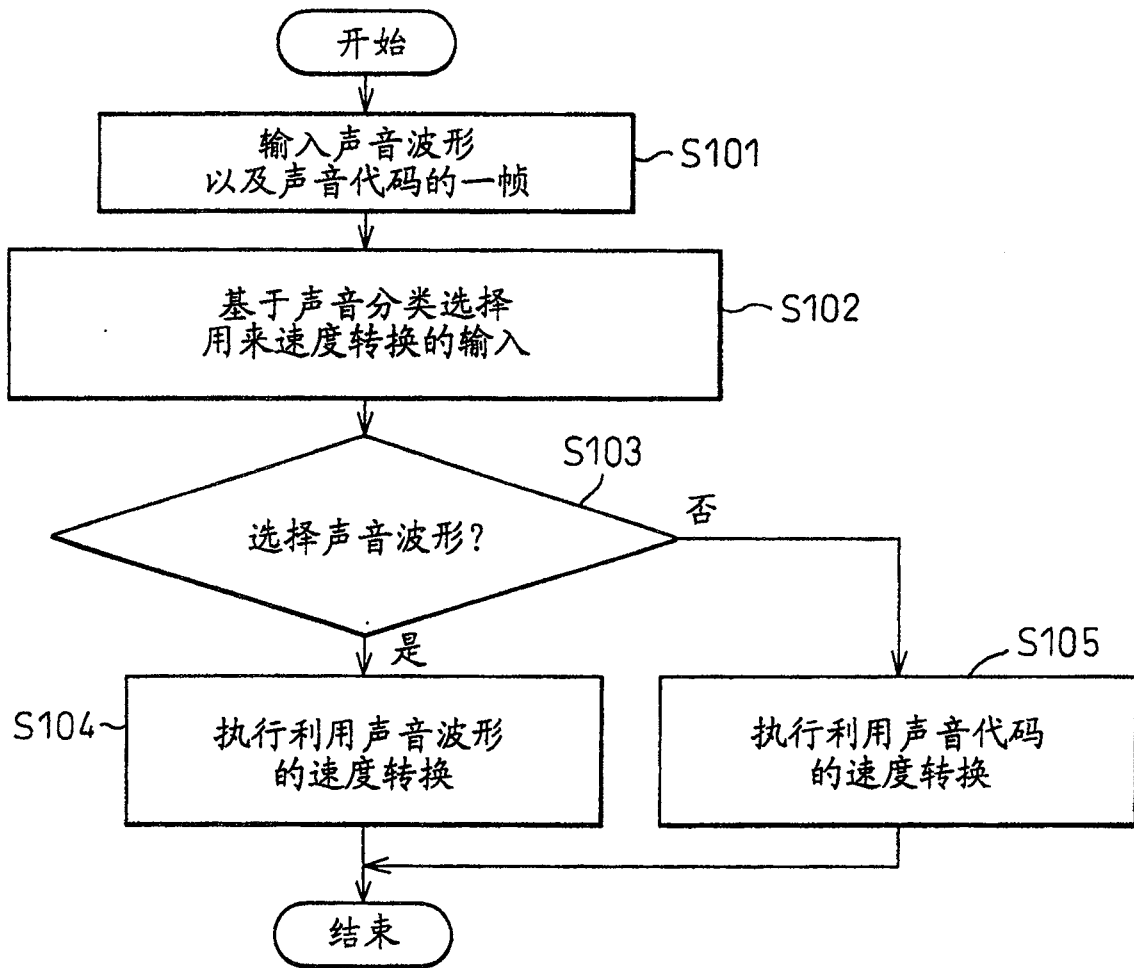


图 7



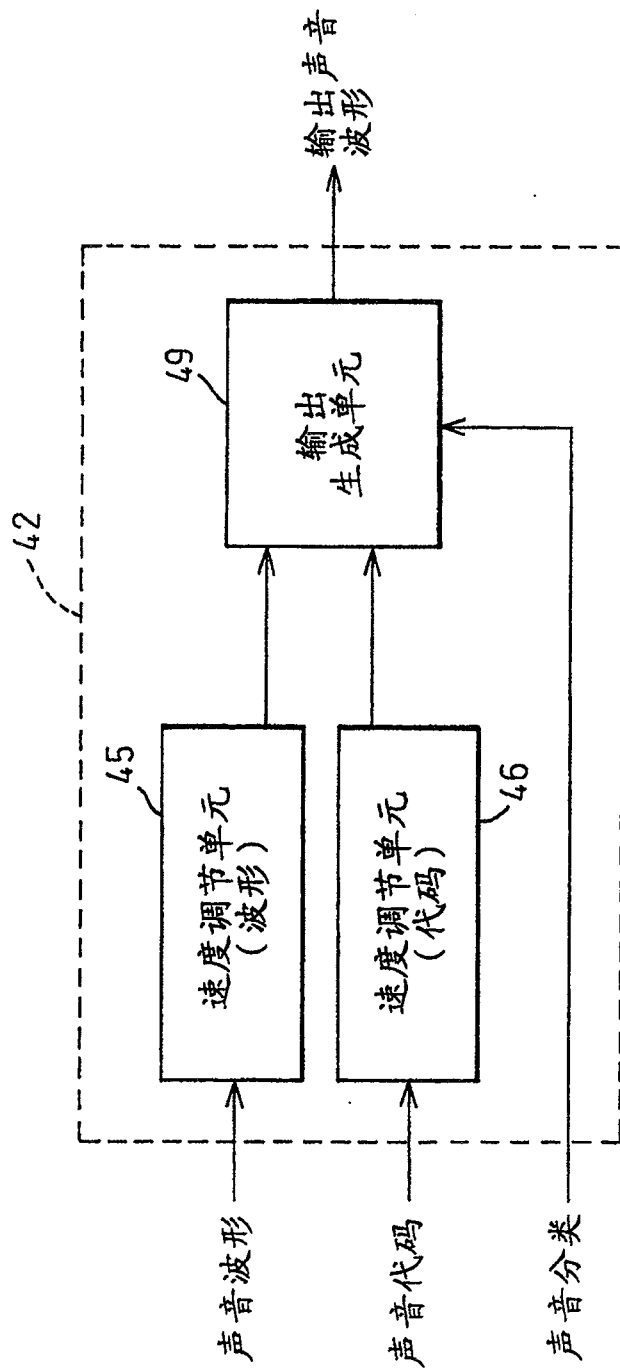


图 8

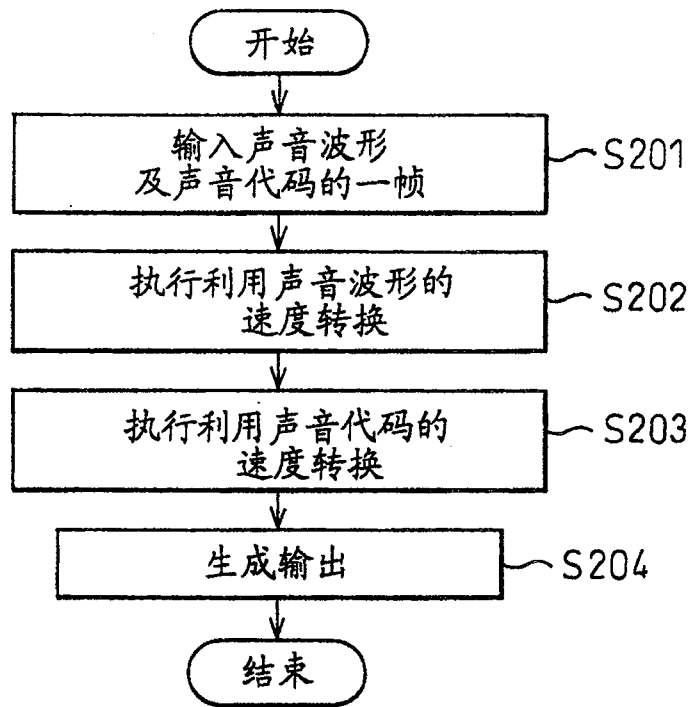


图 9

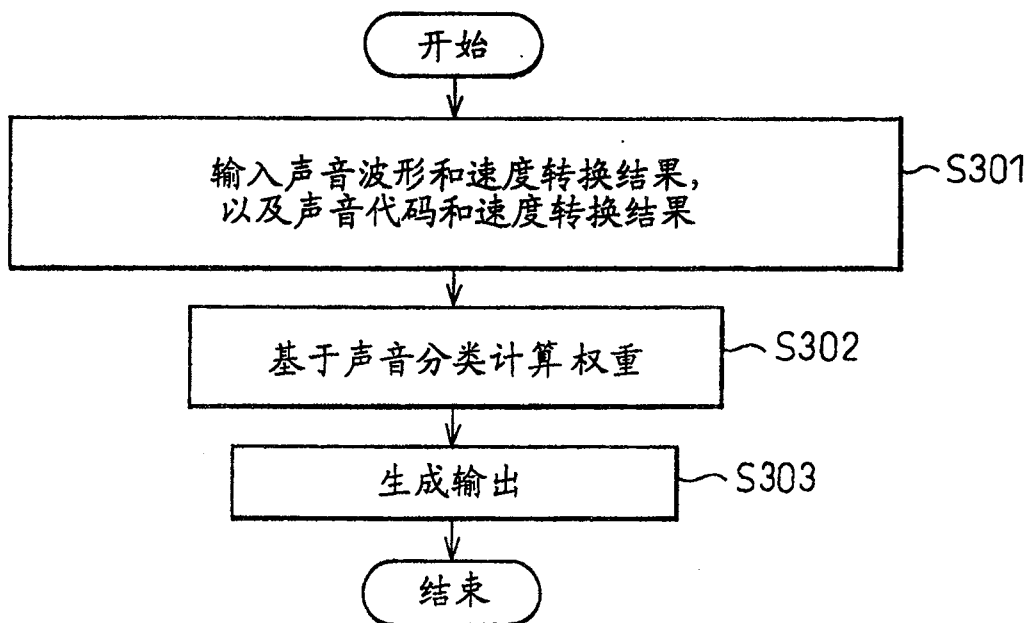


图 10

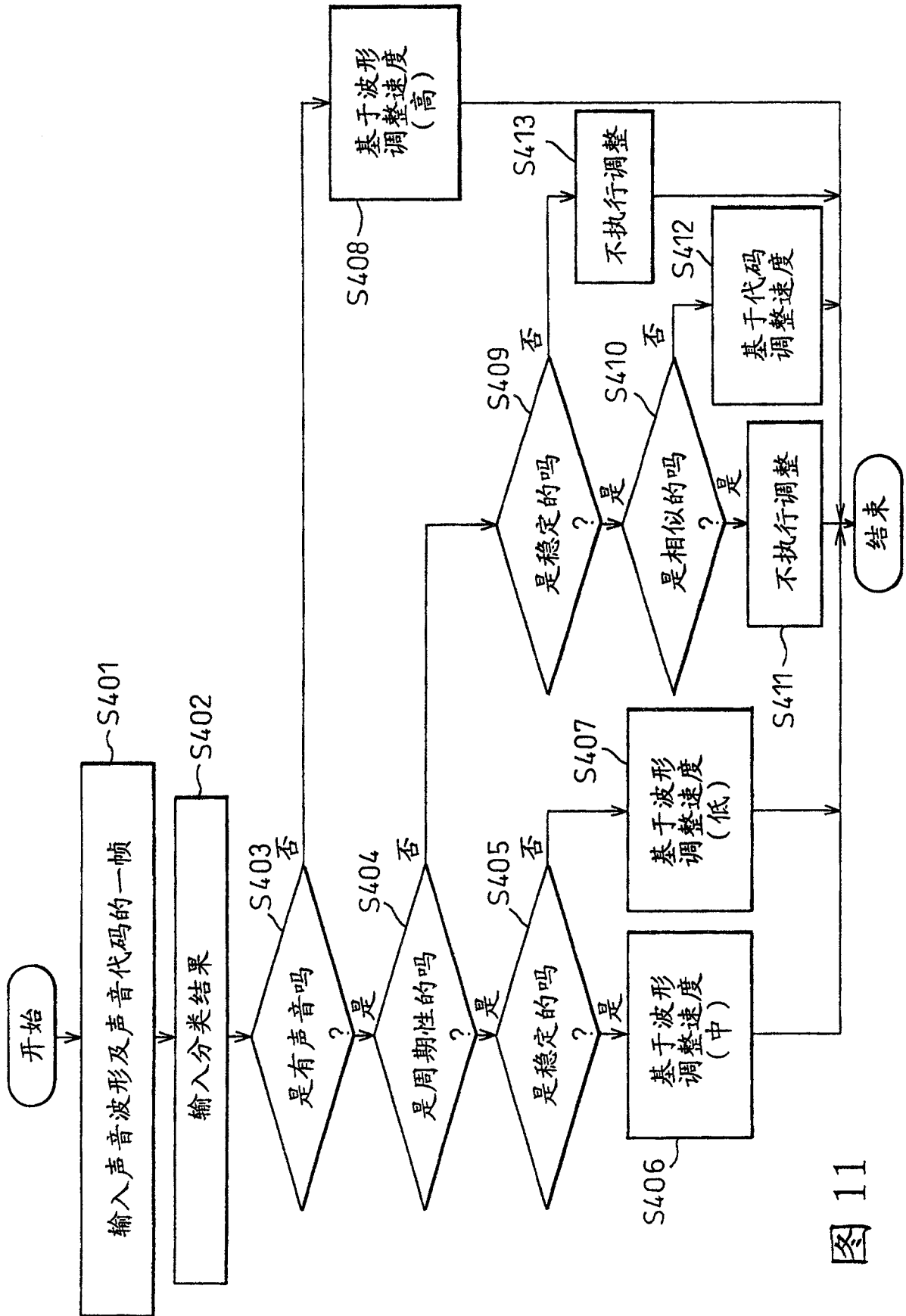
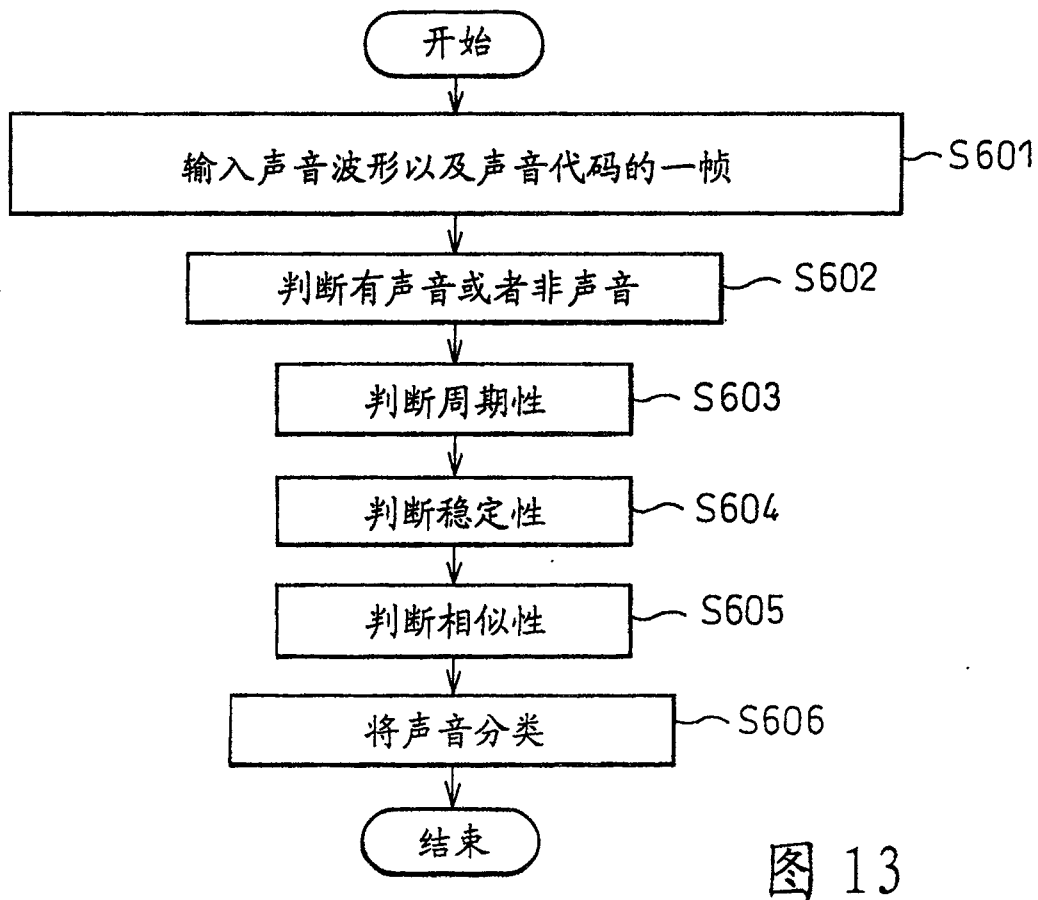
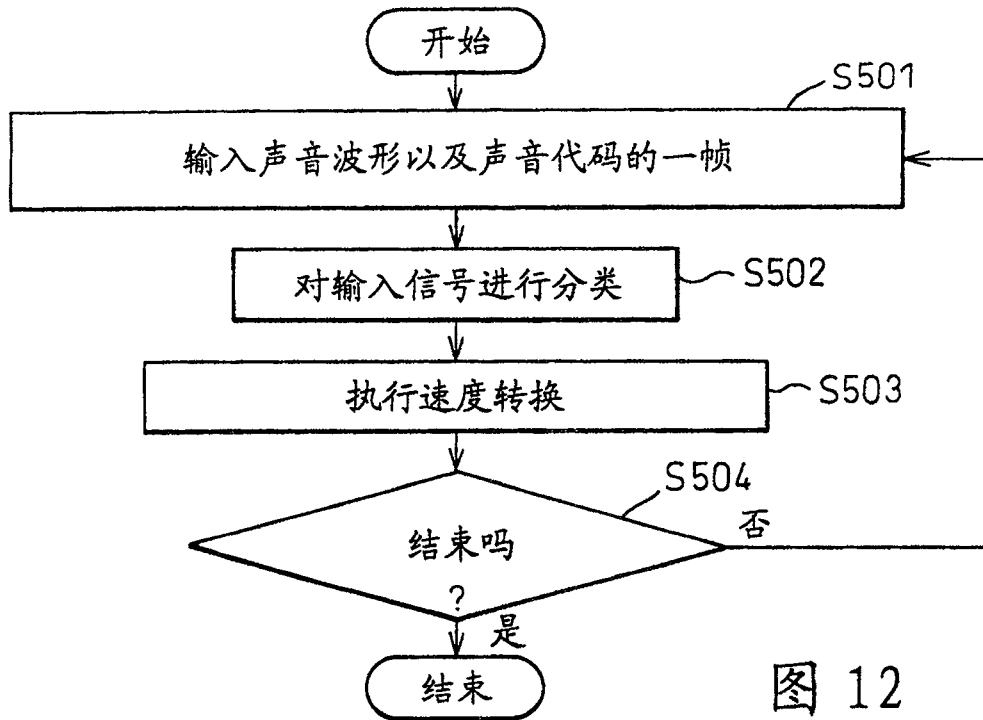


图 11



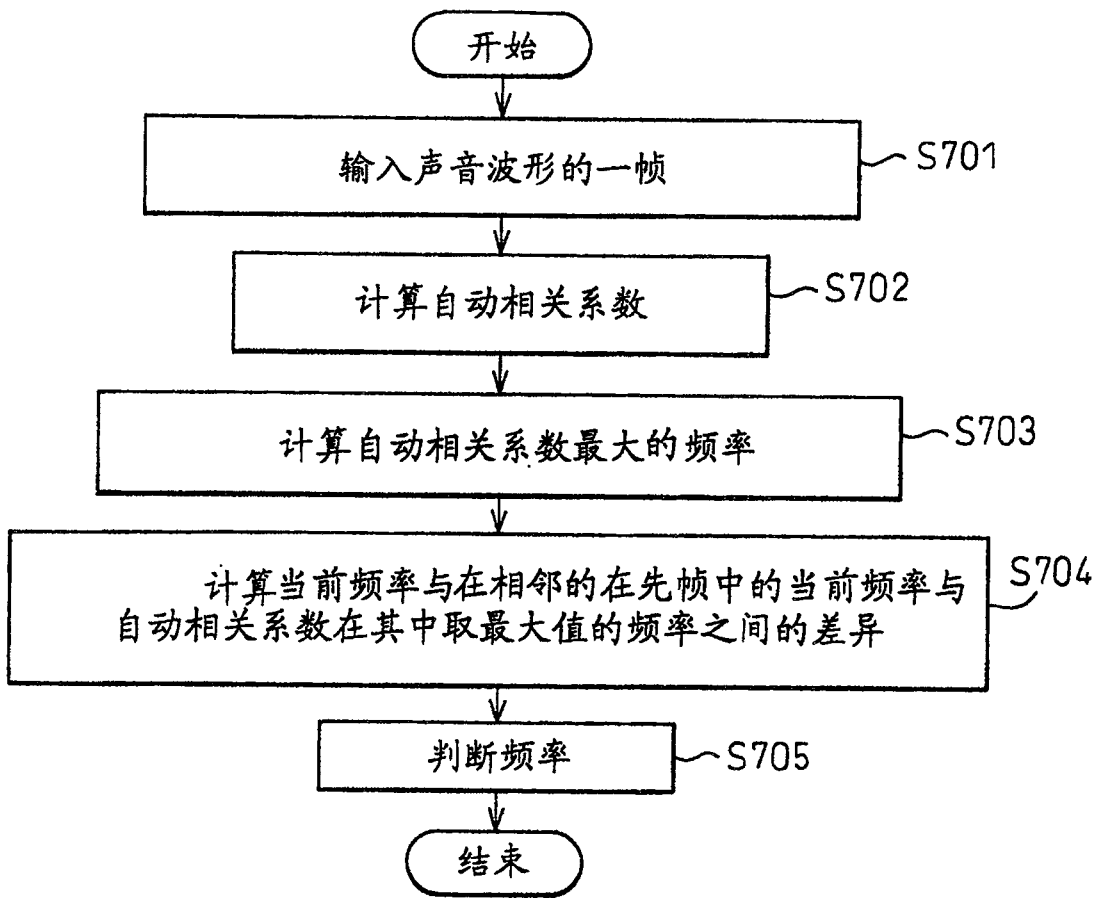


图 14

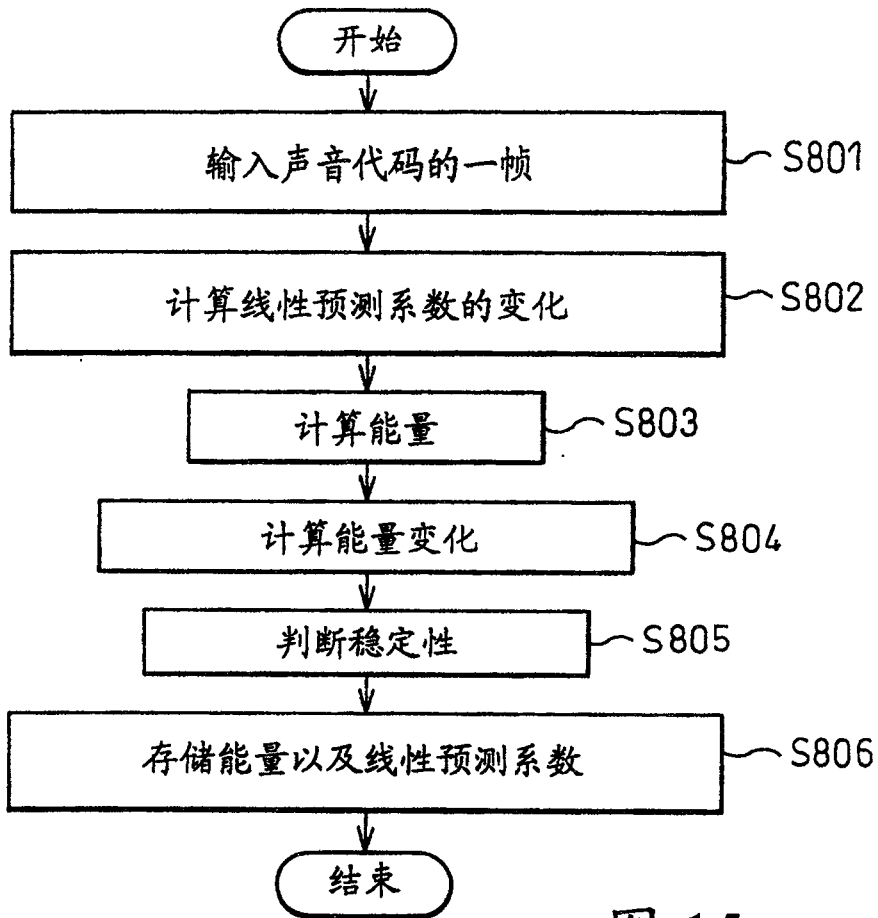


图 15

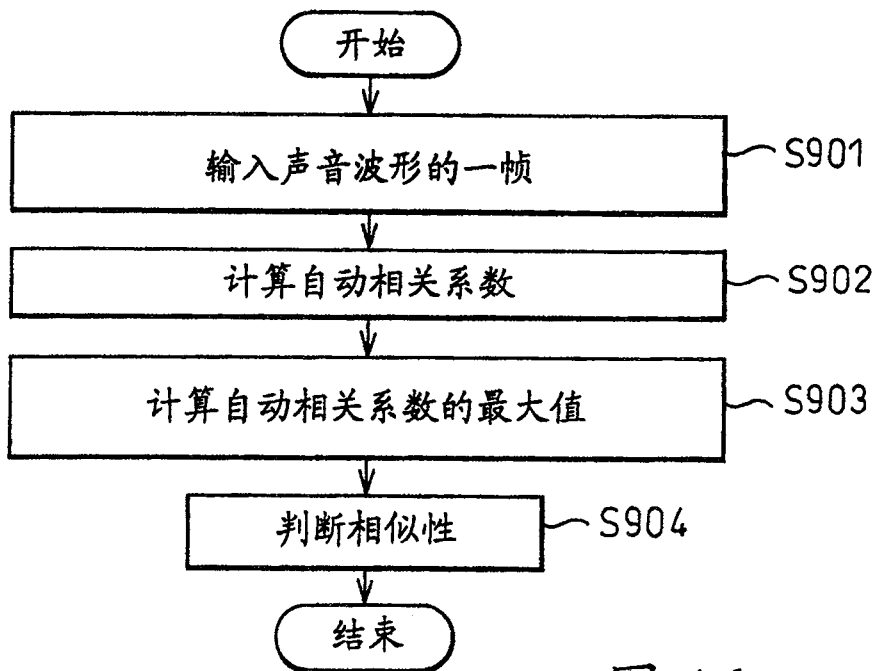


图 16

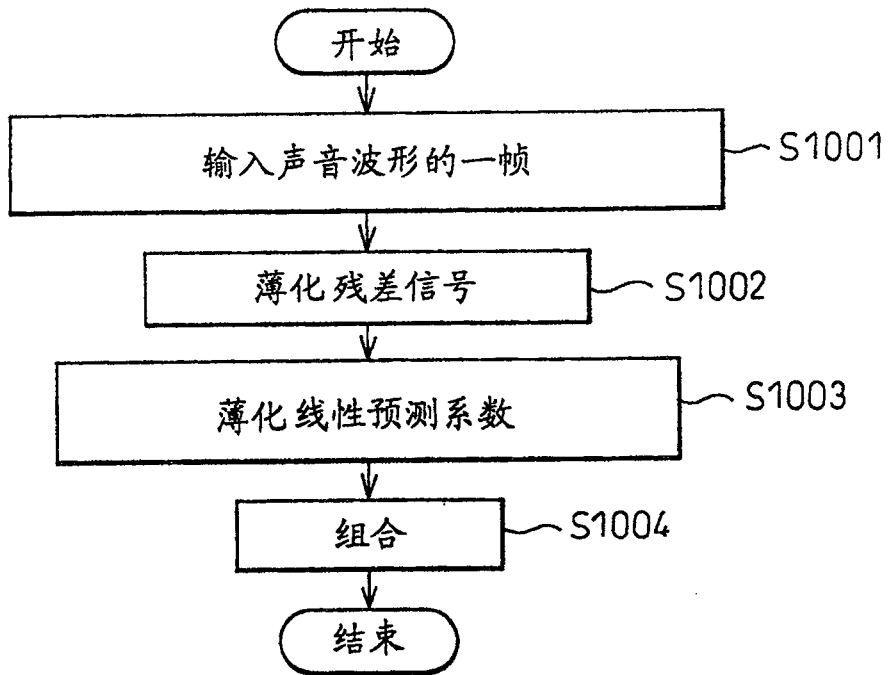


图 17

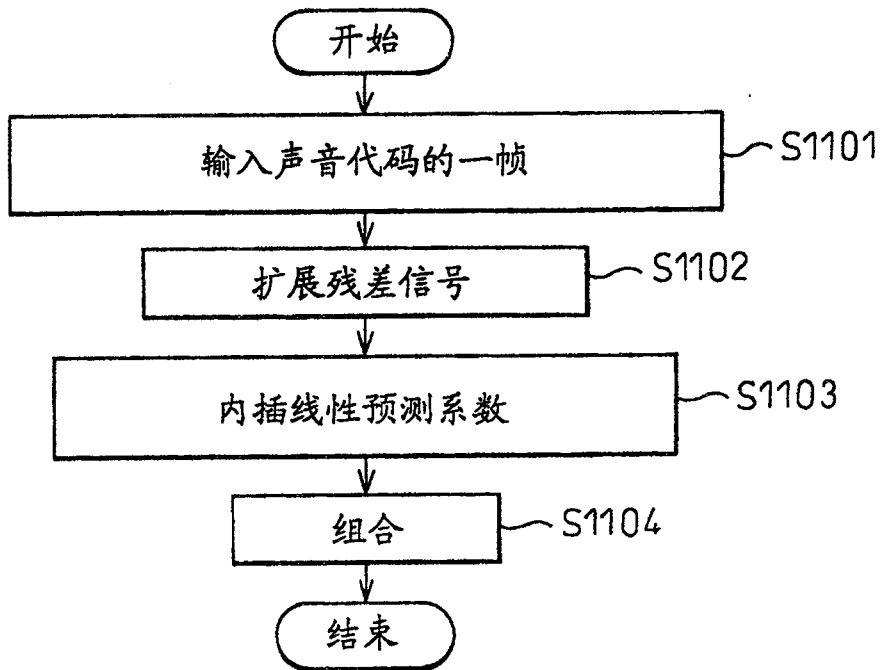


图 18