



(12) 发明专利

(10) 授权公告号 CN 1969315 B

(45) 授权公告日 2012.06.20

(21) 申请号 200580019445.3

G10L 15/20(2006.01)

(22) 申请日 2005.12.20

G10L 15/00(2006.01)

G10L 21/02(2006.01)

(30) 优先权数据

368807/2004 2004.12.21 JP

347641/2005 2005.12.01 JP

(85) PCT申请进入国家阶段日

2006.12.14

(86) PCT申请的申请数据

PCT/JP2005/023336 2005.12.20

(87) PCT申请的公布数据

W02006/068123 JA 2006.06.29

(73) 专利权人 松下电器产业株式会社

地址 日本大阪府

(72) 发明人 野村和也

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 邵亚丽 李晓舒

(51) Int. Cl.

G10L 15/28(2006.01)

(56) 对比文件

EP 0935123 A2, 1999.08.11, 说明书第 87 段, 第 106 段到第 112 段、附图 6.

US 6397188 B1, 2002.05.28, 权利要求 3、说明书第 5 栏第 13 行到第 17 行.

US 20040117187 A1, 2004.06.17, 全文.

CN 1320902 A, 2001.11.07, 全文.

JP 2001-13978 A, 2001.01.19, 说明书第 30 段, 第 47 段、附图 2, 5.

WO 0171480 A2, 2001.09.27, 说明书第 63 页第 24 行到第 64 页第 9 行、附图 42.

CN 1465043 A, 2003.12.31, 全文.

审查员 孙毅

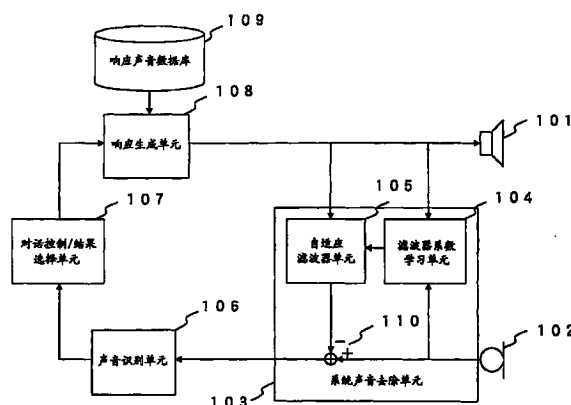
权利要求书 1 页 说明书 10 页 附图 12 页

(54) 发明名称

基于声音的选择装置以及选择方法

(57) 摘要

通过具有:输出单元(101),输出用于引导选择项目的引导声音;声音识别单元(106),识别在由输出单元(101)输出的引导声音的输出中或者其输出之后一定时间内发声的、用于选择选择项目的选择指示;以及对话控制/结果选择单元(107),在由声音识别单元(106)识别出选择指示的情况下,选择被选择指示的选择项目,通过在由输出单元(101)输出引导声音期间,或者其输出结束并经过一定时间前发出用于选择选择项目的声音,可以通过声音识别单元(106)选择该选择项目,即使在引导声音的输出中也可以选择选择项目。



1. 一种基于声音的选择装置,包括:

输出单元,输出用于引导选择项目的引导声音;

声音识别单元,对在由所述输出单元输出的所述引导声音被输出的期间,或者在所述引导声音的输出结束后的一定时间内发声的、用于选择所述选择项目的选择指示进行识别;

对话控制/结果选择单元,在由所述声音识别单元识别出所述选择指示的情况下,选择被选择指示的选择项目;以及

乐曲再现单元,再现对应于所述引导声音的乐曲的一部分或者全部,

所述声音识别单元对在由所述乐曲再现单元再现所述乐曲期间或者再现结束后的一定时间内被输入的选择指示进行声音识别。

2. 如权利要求1所述的基于声音的选择装置,其特征在于,

在由所述输出单元输出的所述引导声音被输出的期间,或者在所述引导声音的输出结束后的一定时间内选择指示未被发声的情况下,所述对话控制/结果选择单元进行控制,以便将用于引导所述选择项目的下一个选择项目的引导声音输出到所述输出单元。

3. 如权利要求1所述的基于声音的选择装置,其特征在于,

所述声音识别单元包括:声音去除单元,从被输入到所述声音识别单元的声音中减去由所述输出单元输出的所述引导声音。

4. 如权利要求1所述的基于声音的选择装置,其特征在于,

其中还包括:生成对应于所述引导声音的图像的图像生成单元,

所述声音识别单元对在由所述图像生成单元生成所述图像期间,或者生成所述图像后的一定时间内被输入的选择指示进行声音识别。

5. 如权利要求1所述的基于声音的选择装置,其特征在于,

其中还包括:输入等待时间设定单元,在由所述输出单元输出的所述引导声音的输出中或者在所述引导声音的输出结束后设定一定时间,

所述声音识别单元对在由所述输入等待时间设定单元设定的所述一定时间内被发声的、用于选择所述选择项目的选择指示进行识别。

6. 一种基于声音的选择方法,包括:

输出步骤,输出用于引导选择项目的引导声音,以及对应于所述选择项目的乐曲的一部分或者全部;

声音识别步骤,对在所述输出步骤中输出的所述引导声音和所述乐曲被输出的期间,或者在所述引导声音和所述乐曲的输出结束后的一定时间内发声的、用于选择所述选择项目的选择指示进行识别;

对话控制/结果选择步骤,在由所述声音识别步骤识别出所述选择指示的情况下,选择被选择指示的选择项目。

基于声音的选择装置以及选择方法

技术领域

[0001] 本发明涉及用于通过声音选择由系统提示的项目的基于声音的选择装置、以及选择方法。

背景技术

[0002] 以往,作为基于声音的选择装置,已知在通过声音指定了控制对象后依次用声音输出并选择控制内容的选择项目的技术(例如,特开平 3-293400 号公报)。

[0003] 按照特开平 3-293400 号公报中记载的技术,可以控制开关而使声音控制系统成为可动作的状态,通过在该状态下将成为控制对象的设备的名称发声来识别该名称,该名称的设备的控制项目依次通过声音合成被发声,在发声了适当的控制项目的阶段通过发声为“是”,执行与该项目相应的控制。

[0004] 而且,还已知以下技术(例如,特开平 6-149534 号公报):利用投影仪将个人计算机的画面设为大画面来显示,并基于该画面,通过发声在该画面上显示的项目来强调显示该项目,如果操作执行按钮,则显示该项目的细节,可以进行该项目的细节的监视、控制。

[0005] 但是,在特开平 3-293400 号公报和特开平 6-149534 号公报中记载的技术中,没有提供具体地受理与系统提示的选择项目重叠的用户的声音的方法,因此,在通常的声音识别方法中,在通过合成声音输出选择项目中难以进行声音识别,同时从系统提示的选择项目的输出方法也被限定于声音,例如,存在不能通过声音直接进行音乐或图像等的选择的问题。

发明内容

[0006] 本发明是处理这样的以往的问题而完成的,提供基于声音的选择装置和选择方法,即使在通过合成声音输出选择项目中,将音乐或图像等作为了选择项目的情况下也可以进行声音识别。

[0007] 本发明的基于声音的选择装置具有以下结构,即包括:输出单元,输出用于引导选择项目的引导声音;声音识别单元,对在由所述输出单元输出的所述引导声音的输出中或者在其输出结束后一定时间内发声的、用于选择所述选择项目的选择指示进行识别;以及对话控制/结果选择单元,在由所述声音识别单元识别出所述选择指示的情况下,选择所述被选择指示的选择项目。

[0008] 通过该结构,可以通过在由输出单元输出引导声音期间、该输出结束并经过一定时间前发出用于选择选择项目的声音,由声音识别单元选择该选择项目,即使在引导声音的输出中也可以选择选择项目。

[0009] 而且,本发明的基于声音的选择装置具有以下结构,即在由所述输出单元输出的所述引导声音的输出中或者在其输出结束后一定时间内选择指示未被发声的情况下,所述对话控制/结果选择单元进行控制,以便将用于引导所述选择项目的下一个选择项目的引导声音输出到所述输出单元。

[0010] 通过该结构,可以在未发出用于选择选择项目的声音的情况下,不断地在一定时间后输出用于引导下一个选择项目的引导声音,在喜爱的选择项目中任意地发出选择指示,并选择该选择项目。

[0011] 而且,本发明的基于声音的选择装置具有以下结构,即所述声音识别单元包括:声音去除单元,从被输入到所述声音识别单元的声音中减去由所述输出单元输出的所述引导声音。

[0012] 通过该结构,由输出单元输出的引导声音被输入声音识别单元,可以尽量排除对声音识别单元进行的选择指示的声音识别产生障碍,即使在引导声音输出中也可以正确地识别选择指示。

[0013] 而且,本发明的基于声音的选择装置具有以下结构,即其中还包括:乐曲再现单元,再现对应于所述引导声音的乐曲的一部分或者全部,所述声音识别单元对在由所述乐曲再现单元再现所述乐曲期间或者再现结束后一定时间内被输入的选择指示进行声音识别。

[0014] 通过该结构,仅通过在用于引导声音的乐曲的再现中,或者在其之后一定时间内发出用于选择指示选择项目的声音,就可以选择并听取该乐曲。

[0015] 而且,本发明的基于声音的选择装置具有以下结构,即其中还包括:生成对应于所述引导声音的图像的图像生成单元,所述图像识别单元对在由所述图像生成单元生成所述图像期间或者生成后一定时间内被输入的选择指示进行声音识别。

[0016] 通过该结构,仅通过在生成并显示对应于引导声音的图像期间,或者在其之后一定时间内发出用于选择指示选择项目的声音,就可以选择该图像,例如,如果为静止图像,则可以原样继续观看静止图像,如果是活动图像,则可以继续该活动画面并观看。

[0017] 而且,本发明的基于声音的选择装置具有以下结构,即其中还包括:输入等待时间设定单元,在由所述输出单元输出的所述引导声音的输出中或者在其输出结束后设定一定时间,所述声音识别单元对在由所述输入等待时间设定单元设定的所述一定时间内被发声的、用于选择所述选择项目的选择指示进行识别。

[0018] 通过该结构,可以通过在由输出单元输出引导声音期间、或该输出结束并经过被设置的一定时间的输入等待时间前发出用于选择选择项目的声音,由声音识别单元选择该选择项目,即使在引导声音的输出中也可以更确实地选择选择项目。

[0019] 进而,本发明的基于声音的选择方法具有以下结构,即包括:输出步骤,输出用于引导选择项目的引导声音;声音识别步骤,对在所述输出步骤中输出的所述引导声音的输出中或者在其输出结束后一定时间内发声的、用于选择所述选择项目的选择指示进行识别;以及对话控制/结果选择步骤,在由所述声音识别步骤识别出所述选择指示的情况下,选择所述被选择指示的选择项目。

[0020] 通过该结构,通过在输出步骤中引导声音被输出期间、其输出结束并在经过一定期间之前发出用于选择选择项目的声音,可以通过声音识别步骤选择该选择项目,即使在引导声音输出中也可以选择选择项目。

[0021] 如上所述,本发明的基于声音的选择装置包括:输出单元,输出用于引导选择项目的引导声音;声音识别单元,对在由所述输出单元输出的所述引导声音被输出的期间,或者在所述引导声音的输出结束后的一定时间内发声的、用于选择所述选择项目的选择指示进

行识别 ; 对话控制 / 结果选择单元, 在由所述声音识别单元识别出所述选择指示的情况下, 选择被选择指示的选择项目 ; 以及乐曲再现单元, 再现对应于所述引导声音的乐曲的一部分或者全部, 所述声音识别单元对在由所述乐曲再现单元再现所述乐曲期间或者再现结束后的一定时间内被输入的选择指示进行声音识别。

[0022] 通过在由输出单元输出引导声音期间、该输出结束并经过一定时间前发出用于选择选择项目的声音, 由声音识别单元选择该选择项目, 即使在引导声音的输出中也可以选择选择项目。

[0023] 而且, 本发明的基于声音的选择方法具有以下步骤, 即包括 : 输出步骤, 输出用于引导选择项目的引导声音, 以及对应于所述选择项目的乐曲的一部分或者全部 ; 声音识别步骤, 对在所述输出步骤中输出的所述引导声音和所述乐曲被输出的期间, 或者在所述引导声音和所述乐曲的输出结束后的一定时间内发声的、用于选择所述选择项目的选择指示进行识别 ; 对话控制 / 结果选择步骤, 在由所述声音识别步骤识别出所述选择指示的情况下, 选择被选择指示的选择项目。

[0024] 通过在输出步骤中引导声音被输出期间、其输出结束并在经过一定期间之前发出用于选择选择项目的声音, 可以通过声音识别步骤选择该选择项目, 即使在引导声音输出中也可以选择选择项目。

[0025] 本发明的基于声音的选择装置包括 : 输出单元, 输出用于引导选择项目的引导声音 ; 输入等待时间设定单元, 在由所述输出单元输出的所述引导声音的输出中或者在其输出结束后设定一定时间 ; 声音识别单元, 对在由所述输入等待时间设定单元设定的所述一定时间内被发声的、用于选择所述选择项目的选择指示进行识别 ; 以及对话控制 / 结果选择单元, 在由所述声音识别单元识别出所述选择指示的情况下, 选择所述被选择指示的选择项目, 通过在由输出单元输出引导声音期间、该输出结束并经过一定时间前发出用于选择选择项目的声音, 由声音识别单元选择该选择项目, 即使在引导声音的输出中也可以选择选择项目。

附图说明

[0026] 图 1 是表示本发明的实施方式 1 中的基于声音的选择装置的概略结构的方框图。

[0027] 图 2 是表示本发明的实施方式 1 中的基于声音的选择装置的动作的流程图。

[0028] 图 3 是表示本发明的实施方式 1 中的基于声音的选择装置的动作的时序图。

[0029] 图 4 是表示本发明的实施方式 2 中的基于声音的选择装置的概略结构的方框图。

[0030] 图 5 是表示本发明的实施方式 2 中的基于声音的选择装置的动作的流程图。

[0031] 图 6 是表示本发明的实施方式 2 中的基于声音的选择装置的动作的时序图。

[0032] 图 7 是表示本发明的实施方式 3 中的基于声音的选择装置的概略结构的方框图。

[0033] 图 8 是表示本发明的实施方式 3 中的基于声音的选择装置的动作的流程图。

[0034] 图 9 是表示本发明的实施方式 3 中的基于声音的选择装置的动作的时序图。

[0035] 图 10 是表示本发明的实施方式 4 中的基于声音的选择装置的概略结构的方框图。

[0036] 图 11 是表示本发明的实施方式 4 中的基于声音的选择装置的动作的流程图。

[0037] 图 12 是表示本发明的实施方式 4 中的基于声音的选择装置的动作的时序图。

[0038] 标号说明

- [0039] 101 扬声器
- [0040] 102 麦克风
- [0041] 103 系统声音去除单元
- [0042] 104 滤波器系数学习单元
- [0043] 105 自适应滤波器单元
- [0044] 106 声音识别单元
- [0045] 107,1007 对话控制 / 结果选择单元
- [0046] 108,1008 响应生成单元
- [0047] 109 响应声音数据库
- [0048] 110 减法器
- [0049] 411 音乐再现单元
- [0050] 412 音乐数据库
- [0051] 413 混频器单元
- [0052] 700 显示器
- [0053] 711 图像生成单元
- [0054] 712 图像 / 活动画面数据库
- [0055] 1011 输入等待时间设定单元

具体实施方式

[0056] 以下,参照附图对本发明的实施方式进行说明。

[0057] (实施方式 1)

[0058] 图 1 是表示本发明的实施方式 1 中的基于声音的选择装置的概略结构的方框图。

[0059] 如图 1 所示,本发明的实施方式 1 中的基于声音的选择装置包括:作为对用户输出作为引导的声音或声音响应的系统侧的声音的输出单元的扬声器(声音输出单元)101;将用户发出的声音变换为声音信号的麦克风 102;作为将与扬声器 101 输出的引导声相当的输出相当信号从由麦克风 102 输出的声音信号中去除的声音去除单元的系统声音去除单元 103;根据从麦克风 102 输出的、通过系统声音去除单元 103 去除重叠信号部分而得到的声音信号,识别用户声音的发话内容的声音识别单元 106;根据由声音识别单元 106 得到的用户声音的内容选择对应的响应声音而控制与用户的对话,同时选择结果的对话控制/结果选择单元 107;存储了响应声音数据的响应声音数据库 109;根据对话控制/结果选择单元 107 的输出,利用响应声音数据库 109 的数据来生成用于输出到扬声器 101 或系统声音去除单元 103 的声音响应信号的响应生成单元 108。

[0060] 系统声音去除单元 103 包括:滤波器系数学习单元 104,根据从麦克风 102 输出的声音信号和从响应生成单元 108 输出的响应声音信号,例如一边学习利用 LMS(Least Mean Square)/牛顿算法得到的滤波器系数(脉冲响应),一边调整为最佳;自适应滤波器单元 105,根据作为来自滤波器系数学习单元 104 的输出的脉冲响应校正并输出响应声音信号;以及减法器 110,从麦克风 102 输出的声音信号中减去由自适应滤波器单元 105 输出的输出信号。

[0061] 声音识别单元 106 包括:音响处理单元,对从麦克风 102 输出的、由系统声音去除

单元 103 减去了声音响应的相当重叠部分的声音信号进行音响处理；音素识别单元，根据由音响处理单元得到的声音的最小单位选出并识别最相似的音素候选；辞典数据库，存储与声音对话系统的利用目的关联的单词等；语言处理单元，根据在音素识别单元得到的音素和来自辞典数据库的声音数据选定单词的候选，并执行用于利用句子结构、含义、文理等的语言信息得到正确的文章的语言处理。

[0062] 而且，音响处理单元例如构成为：利用 LPC 倒频谱 (Linear PredictorCoefficient Cepstrum: 线性预测系数化倒频谱) 等，将从麦克风 102 输出的声音信号变换为所谓特征量向量的时间系列的向量，并推测声音频谱的大概形状 (频谱包络)。

[0063] 而且，音素识别部例如构成为：利用 HMM (Hidden Markov Model: 隐藏马尔可夫模型) 法等，根据输入的声音，利用由音响处理单元提取的音响参数，进行声音信号的音素记号化，与预先准备的标准音素模型比较，选出与其最相似的音素的候选。

[0064] 另一方面，对话控制 / 结果选择单元 107 根据由声音识别单元 106 识别的声音信号的内容，选择控制响应内容并将其输出到响应生成单元 108，同时选择输出结果。

[0065] 响应生成单元 108 根据由对话控制 / 结果选择单元 107 决定的内容，利用来自响应声音数据库 109 的数据来生成响应声音信号，输出到扬声器 101。

[0066] 接着，利用图 2、图 3 详细地说明本发明的实施方式 1 中的基于声音的选择装置的动作。

[0067] 而且，图 2 是说明本发明的实施方式 1 中的基于声音的选择装置的动作的流程图，图 3 是时序图。

[0068] 首先，在进入选择开始动作时，基于对话控制 / 结果选择单元 107 的控制，选择项目的计数器 N 被设为 1 (步骤 201)。在选择项目的计数器 N 被设为 1 时，接着响应生成单元 108 根据来自对话控制 / 结果选择单元 107 的指令将引导声音从响应声音数据库 109 输出到扬声器 101 (步骤 202)。

[0069] 例如，如图 3 的时序图所示那样，将“请从下面中选择希望的曲名。”那样的引导声音 (301) 输出到扬声器 101。

[0070] 如果引导声音从扬声器 101 输出，则接着进行声音识别的起动，以便可以对来自用户的选择指示进行声音识别 (步骤 203)。由此，声音识别单元 106 如图 3 所示那样起动 (302)。

[0071] 如果声音识别单元 106 起动，则接着根据对话控制 / 结果选择单元 107 的控制，响应生成单元 108 访问响应声音数据库 109，输出对应于第一个选择项目的声音数据 (步骤 204)。

[0072] 即，由此，例如如图 3 所示那样，将“青鳉的学校。”的引导声音 (303) 输出到扬声器 101。于是，“青鳉的学校。”的引导声音 (303) 被输出的期间 308A，以及在该输出结束后的一定期间 (输入等待时间) 308B 成为可以选择第一个选择项目的“青鳉的学校”的时间 308。隐藏，如果在该期间 308 用户发出了指示选择的语言，例如“就是它！”的语言，则可以选择“青鳉的学校”。

[0073] 在可以选择“青鳉的学校”的期间 308，如果用户没有发出进行选择指示的语言，例如“就是它！”的语言的情况下，选择项目的计数器 N 中加 1，成为输出对应于下一个选择项

目的引导声音的状态。

[0074] 即,在输出对应于选择项目的声音时(步骤 204),对话控制/结果选择单元 107 执行在该选择项目的提示中或者提示结束后一定时间内用户是否发出了指示选择的语言的声音的判定(步骤 205)。

[0075] 在用户进行了选择指示的情况下(步骤 205 中“是”的情况),此时,例如虽然选择了“青鳉的学校”,但是未进行选择指示的情况下(步骤 205 中“否”的情况下),对选择项目的计数器 N 加 1(步骤 206),成为输出对应于下一个选择项目的声音,即对应于第二个选择项目的引导声音的状态(步骤 204)。

[0076] 由此,响应生成单元 108 访问响应声音数据库 109,将第二个引导声音的例如“扫晴娘。”的引导声音(304)输出到扬声器 101。

[0077] 这时,在输出“扫晴娘。”的引导声音(304)的期间 309A,以及该输出结束之后的一定时间(输入等待时间)309B 也成为可选择第二个选择项目的“扫晴娘”的期间 309,在该期间 309 中如果用户发出了选择指示的语言,例如“就是它!”的语言的声音,则可以选择作为第二个选择项目的“扫晴娘”。

[0078] 在选择项目提示中,或者在提示结束后一定时间内,用户未发出指示选择的语言,例如“就是它!”的声音的情况下,该情况被对话控制/结果选择单元 107 判定(步骤 205),在该判定后,分支到“否”的路径,与以前一样,对选择项目的计数器 N 加 1(步骤 206),输出对应于下面的第三个选择项目的引导声音(步骤 204)。

[0079] 然后,由此响应生成单元 108 访问响应声音数据库 109,作为第三个引导声音,例如将“闪闪的星。”的引导声音(305)输出到扬声器 101。

[0080] 在第三个引导声音“闪闪的星。”(305)被输出期间 310A,以及输出结束后的一定期间(输入等待时间)310B,为可以选择作为第三个选择项目的“闪闪的星”的时间 310,在该期间 310 中,如果用户发出了选择指示的语言,例如“就是它!”的语言的声音,则可以选择作为第三个选择项目的“闪闪的星”。

[0081] 在图 3 中,用户为了对作为第三个选择项目的“闪闪的星”进行选择指示,在第三个引导声音“闪闪的星。”(305)被输出期间,或者在其后的一定期间,发出用于选择它的声音,例如“就是它!”的指示声音(306)。

[0082] 这里,如果在输出引导声音“闪闪的星。”(305)期间,用户发出了“就是它!”的指示声音(306),则引导声音“闪闪的星”(305)和“就是它!”(306)的指示声音重叠地被输入到麦克风 102,但是通过系统声音去除单元 103 将与引导声音相当部分的信号,即“闪闪的星”(305)的声音相当的部分的信号从输入到麦克风 102 的信号中去除,在声音识别单元 106 中可以正确地识别用户发声的指示声音“就是它!”(306)。

[0083] 在图 2 中,在选择项目提示中,或者在提示结束后一定时间内用户发出了选择指示的语言,例如“就是它!”的语言的声音的情况下,它被声音识别单元 106 识别,在对话控制/结果选择单元 107 被判定(步骤 205),分支到“是”的路径。

[0084] 在分支到“是”的路径时,进行声音识别从而结束该声音识别(步骤 207),选择这时的选择项目(步骤 208),以后,根据被选择的结果,例如“闪闪的星”,对话控制/结果选择单元 107 进行对话控制。

[0085] 而且,虽然未图示,但是构成为在最后的選擇項目被提示的阶段,在虽然经过一定

时间,但未从用户发出选择指示的语言的声音的情况下,从扬声器 101 进行超时的警告,结束声音识别而中止选择。

[0086] 如上所述,按照本发明的实施方式 1,在系统基于声音提示选择项目的期间,或者提示结束后被设置的输入等待时间内,用户发出了进行选择指示的语言的声音的情况下,可以选择发出了该选择指示的语言的阶段的选择项目。

[0087] (实施方式 2)

[0088] 图 4 是表示本发明的实施方式 2 中的基于声音的选择装置的概略结构的方框图,图 5 是说明该装置的动作的流程图,图 6 是说明该装置的动作的时序图。

[0089] 而且,在图 4~图 6 中,赋予与图 1~图 3 所示的实施方式 1 相同标号、号码等的部分表示与图 1~图 3 所示的实施方式 1 相同的结构、相同的内容的部分,这里省略其详细的说明。

[0090] 在本实施方式中,除了实施方式 1 的结构,还具有由对话控制/结果选择单元 107 的指令控制的音乐再现单元 411;以及存储多个音乐的音乐数据库 412。

[0091] 音乐再现单元 411 被构成为通过来自对话控制/结果选择单元 107 的指令访问音乐数据库 412,并且再现在对话控制/结果选择单元 107 中指示的音乐。然后,由音乐再现单元 411 再现的音乐经由混频器单元 413 与来自响应生成单元 108 的输出一起输出到扬声器 101。

[0092] 在图 6 中,基于音乐输出的引导音乐 603~605 分别对应于图 3 的引导声音 303~305。

[0093] 因此,按照本实施方式,如图 5、图 6 所示,不仅输出作为选择项目的引导声音,而且还同时输出对应于该选择项目的音乐及其本身,在选择音乐的情况下更方便。

[0094] 即,按照本实施方式,在实施方式 1 中,输出对应于第 N 个选择项目的引导声音的步骤 204 变为输出对应于第 N 个选择项目的音乐的步骤 504,在该步骤 504 中,如果依次输出对应于第 N 个选择项目的引导声音和对应于第 N 个选择项目的音乐两者,则首先曲名被输出,接着音乐被输出,所以在选择音乐的情况下变得更方便。

[0095] 而且,这里,为选择而输出的音乐也可以不是音乐全部(一个乐曲全体),例如,仅序曲或者副曲(サビ)部分就足够。于是,不论为选择而输出的音乐全部或者仅序曲或副曲部分,在为选择而输出的音乐已被选择的情况下,通过音乐再现单元 411,可以原样继续并输出该音乐,也可以暂时返回音乐的最初后输出。

[0096] 这样,按照本实施方式,将音乐作为选择项目提示,并且如果在该期间,或者提示结束后一定时间内用户发出选择指示的语言的声音,则可以容易地选择该用户希望的音乐。

[0097] (实施方式 3)

[0098] 图 7 是表示本发明的实施方式 2 中的基于声音的选择装置的概略结构的方框图,图 8 是说明该装置的动作的流程图,图 9 是说明该装置的动作的时序图。

[0099] 而且,在图 7~图 9 中,赋予与图 1~图 3 所示的实施方式 1 相同标号、号码等的部分表示与图 1~图 3 所示的实施方式 1 相同的结构、相同的内容的部分,这里省略其详细的说明。

[0100] 在本实施方式中,除了实施方式 1 的结构,还包括:由对话控制/结果选择单元

107 的指令控制的图像生成单元 711 ;存储多个静止画面、活动画面等的图像的图像 / 活动画面数据库 712 ;以及显示由图像生成单元 711 生成的图像的显示器 700。

[0101] 图像生成单元 711 被构成为通过来自对话控制 / 结果选择单元 107 的指示访问图像 / 活动画面数据库 712,并输出在对话控制 / 结果选择单元 107 中指示的静止画面、活动画面等图像数据,并且生成图像。然后,图像生成单元 711 生成的图像被显示在显示器 700 中。

[0102] 在图 9 中,基于声音输出的引导声音 901、基于对显示器的显示的引导图像 903 ~ 905 分别对应于图 3 的引导声音 301、303 ~ 305。

[0103] 因此,按照本实施方式,如图 8、图 9 所示,不仅输出成为选择项目的引导声音,而且还同时显示器 700 上显示对应于该选择项目的图像,在选择选择项目时更方便。

[0104] 即,按照本实施方式,在实施方式 1 中,输出对应于第 N 个选择项目的引导声音的步骤 204 变为输出对应于第 N 个选择项目的图像的步骤 804,在该步骤 804 中,输出对应于第 N 个选择项目的引导声音和对应于第 N 个选择项目的图像两者,前者从扬声器 101 作为声音,后者在显示器 700 中作为图像或者活动画面被显示。因此,以它们为基础,可以更容易地选择选择项目。

[0105] 而且,在为选择而输出的图像为活动画面时,也可以不是该活动画面的全部,例如,仅最初的或者子画面的一定时间就足够。于是,无论为选择而输出的图像的全部还是仅一定时间,在为选择而输出的图像已被选择的情况下,都可以通过图像生成单元 711 原样继续并显示该图像,也可以暂时返回活动画面的最初后显示。

[0106] 这样,按照本实施方式,除了选择项目的引导声音,还将与其对应的图像也作为选择项目进行提示,并且如果在该提示的时间或者提示结束后一定时间内用户发出了选择指示的语言的声音,则可以选择该用户希望的选择项目,所以,例如具有绘画或电影等图像本身很好,在音乐的情况下,通过提示护封(jacket)的图像,也可以更容易地选择音乐的效果。

[0107] (实施方式 4)

[0108] 在上述各实施方式中,例如如图 3 所示,没有积极地设置用于选择的时间 308B、309B 等的结构,但是利用图 10 至图 12 对具有设置用于该选择的时间 308B、309B 等的输入等待时间设定单元的基于声音的选择装置进行说明。

[0109] 通过具有该输入等待时间设定单元,可以更确实地进行声音识别。

[0110] 图 10 是表示本发明的实施方式 4 中的基于声音的选择装置的概略结构的方框图,图 11 是说明该装置的动作的流程图,图 12 是说明该装置的动作的时序图。

[0111] 在图 10 中,由于本实施方式的基本的结构与上述实施方式 1 记载的基于声音的选择装置的结构图一样,所以这里仅对两者的不同点进行说明,并且对于其他的基本的结构、动作加以省略。

[0112] 本实施方式的对话控制 / 结果选择单元 1007 和响应生成单元 1008 与实施方式 1 的对应部分具有不同的功能,而且,在本实施方式中,具有与对话控制 / 结果选择单元 1007 和响应生成单元 1008 连接的输入等待时间设定单元 1011。

[0113] 与实施方式 1 一样,在对话控制 / 结果选择单元 1007 的控制下声音识别单元 106 起动时,响应生成单元 1008 访问响应声音数据库 109,并且输出对应于选择项目的声音数

据。

[0114] 进而,在对话控制/结果选择单元 1007 中,判定对应于选择项目的声音数据是否完成输出。

[0115] 对话控制/结果选择单元 1007 的判定的结果,在确认了声音数据已被输出时,通过设定用户用于响应的期间的输入等待时间设定单元 1011 设定输入等待时间。

[0116] 在响应生成单元 1008 中,在该输入等待时间的期间,通过对话控制/结果选择单元 1007 禁止其动作。

[0117] 接着,利用图 11 和图 12 对本实施方式的基于声音的选择装置的动作进行说明,但是由于到声音识别的起动为止(步骤 201~步骤 203)与实施方式 1 的动作一样,所以这里省略。

[0118] 在步骤 203 中声音识别单元 106 起动时,接着,根据对话控制/结果选择单元 1007 的控制,响应生成单元 1008 访问响应声音数据库 109,并且输出与第一个选择项目对应的声音数据(步骤 204)。

[0119] 即,由此,例如如图 12 所示,将“青鳉的学校。”的引导声音(303)输出到扬声器 101。

[0120] 接着,对话控制/结果选择单元 1007 进行“青鳉的学校。”的引导声音(303)的输出是否已结束的判定。

[0121] 判定的结果,引导声音(303)的输出结束的情况下,输入等待时间设定单元 1011 在对话控制/结果选择单元 1007 的控制下设定输入等待时间 1208B(步骤 1109)。

[0122] 于是,“青鳉的学校。”的引导声音(303)被输出的期间 308A,以及该输出结束后的一定时间 1208B 成为可选择作为第一个选择项目的“青鳉的学校”的时间 1208。因此,如果在该期间 1208 用户发出了指示选择的语言,例如“就是它!”的语言的声音,则可以选择“青鳉的学校”。

[0123] 对话控制/结果选择单元 1007 在由输入等待时间设定单元 1011 设定的输入时间期间,禁止响应生成单元 1008 的动作,以便不通过响应生成单元 1008 产生下一个引导声音或者由于误动作产生引导声音等。

[0124] 是否禁止响应生成单元 1008 的动作,通过是否经过了输入等待时间设定单元 1011 设定的时间的判定来进行。

[0125] 接着,在可以选择“青鳉的学校”的期间 1208,用户未发出选择指示的语言,例如:“就是它!”的语言的声音的情况下,对选择项目的计数器 N 加 1,成为输出对应于下一个选择项目的引导声音的状态。

[0126] 即,在步骤 204 中输出对应于选择项目的声音时,在该选择项目的提示中,或者提示结束以后,对话控制/结果选择单元 1007 进行在由步骤 S1109 设定的输入等待时间的一定时间内用户是否发出了指示选择的语言的声音的判定(步骤 1105)。

[0127] 接着,在输入等待时间中用户发出了指示选择的语言的声音的情况下(步骤 1105 中的“是”的情况),这里,例如虽然选择了“青鳉的学校”,但是没有发出指示选择的语言的声音的情况下(步骤 1105 中的“否”的情况),对选择项目的计数器 N 加 1(步骤 1106),成为输出与下一个选择项目对应的声音,即与第二个选择项目对应的引导声音的状态(步骤 204)。

[0128] 这时,在图 12 中,对应第二个或者第三个选择项目,引导声音 (304 或者 305) 输出的期间 309A 或者 310A,以及各个输出结束后的各个一定时间 1209B 或者 1210B 成为可以选择第二个或者第三个选择项目的时间 1209 或 1210。

[0129] 这之后的处理与实施方式 1 的图 2 所示的动作一样。

[0130] 如上所述,在本实施方式中,通过具有输入等待时间设定单元 1011,可以等待用户的响应,并且积极地设定等待时间。

[0131] 通过这样积极地设定等待时间,下一个引导声音不会在输入等待时间内被误输出,从而侵占输入等待时间,可以一定确实地设置用户的可响应的期间。

[0132] 本实施方式的具有输入等待时间设定单元 1011 的结构也可以应用于上述实施方式 2 或实施方式 3 的结构,不用说也可以得到与本实施方式一样的动作和效果。

[0133] 而且,本发明也可以适当地组合上述各实施方式,例如,除了选择项目的引导声音,也可以通过将与其对应的图像和乐曲也作为选择项目提示,通过在该提示的时间,或者提示结束后被设置的一定时间的输入等待时间内用户发出选择指示的语言的声音,该用户可以选择希望的选择项目。

[0134] 产业上的可利用性

[0135] 本发明的基于声音的选择装置具有以下结构,即包括:输出单元,输出用于引导选择项目的引导声音;声音识别单元,对在由输出单元输出的引导声音的输出中或者在其输出结束后一定时间内发声的、用于选择选择项目的选择指示进行识别;以及对话控制/结果选择单元,在由声音识别单元识别出选择指示的情况下,选择被选择指示的选择项目,可以广泛地用于汽车音响、汽车空调等车载用电子设备,或者电子黑板、投影仪等电子事务设备、残疾人用的家庭用电子设备等。

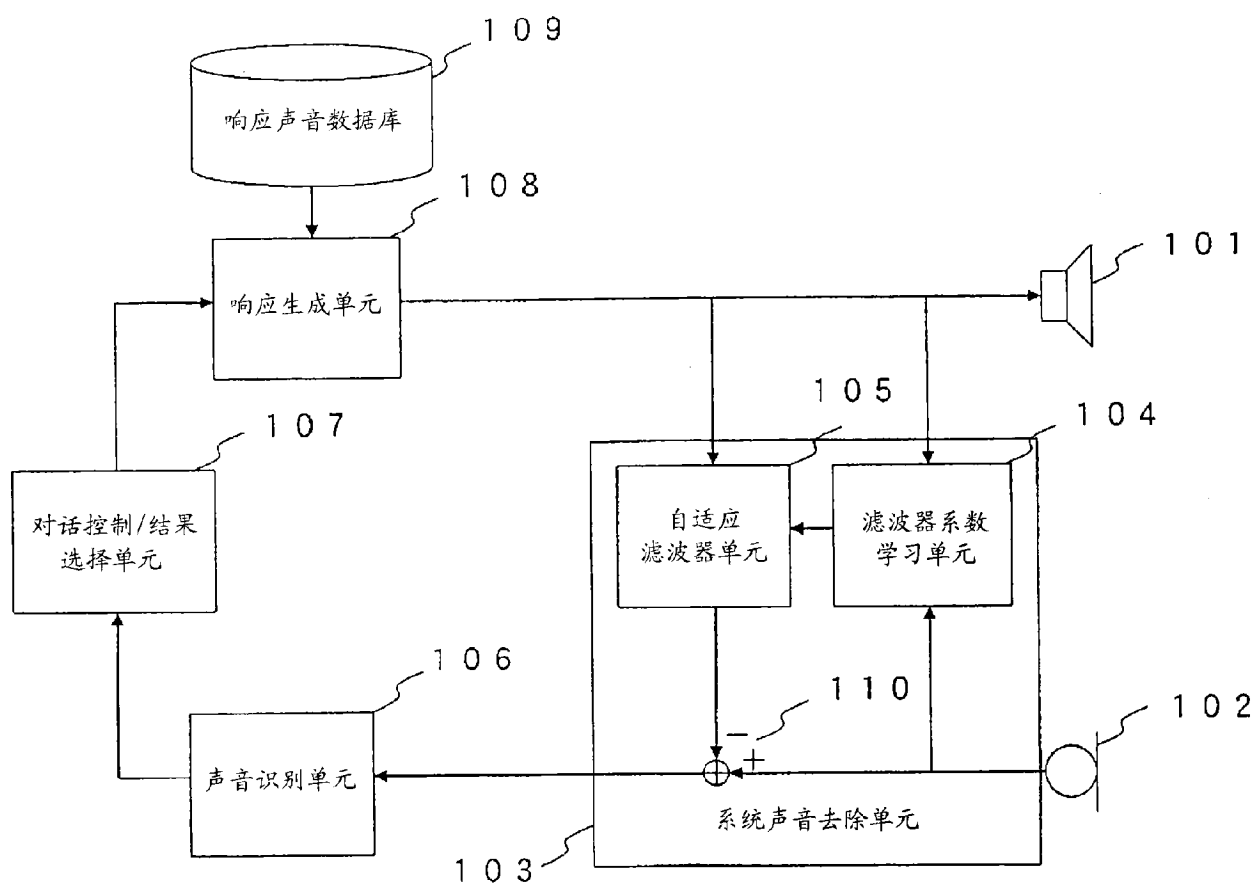


图 1

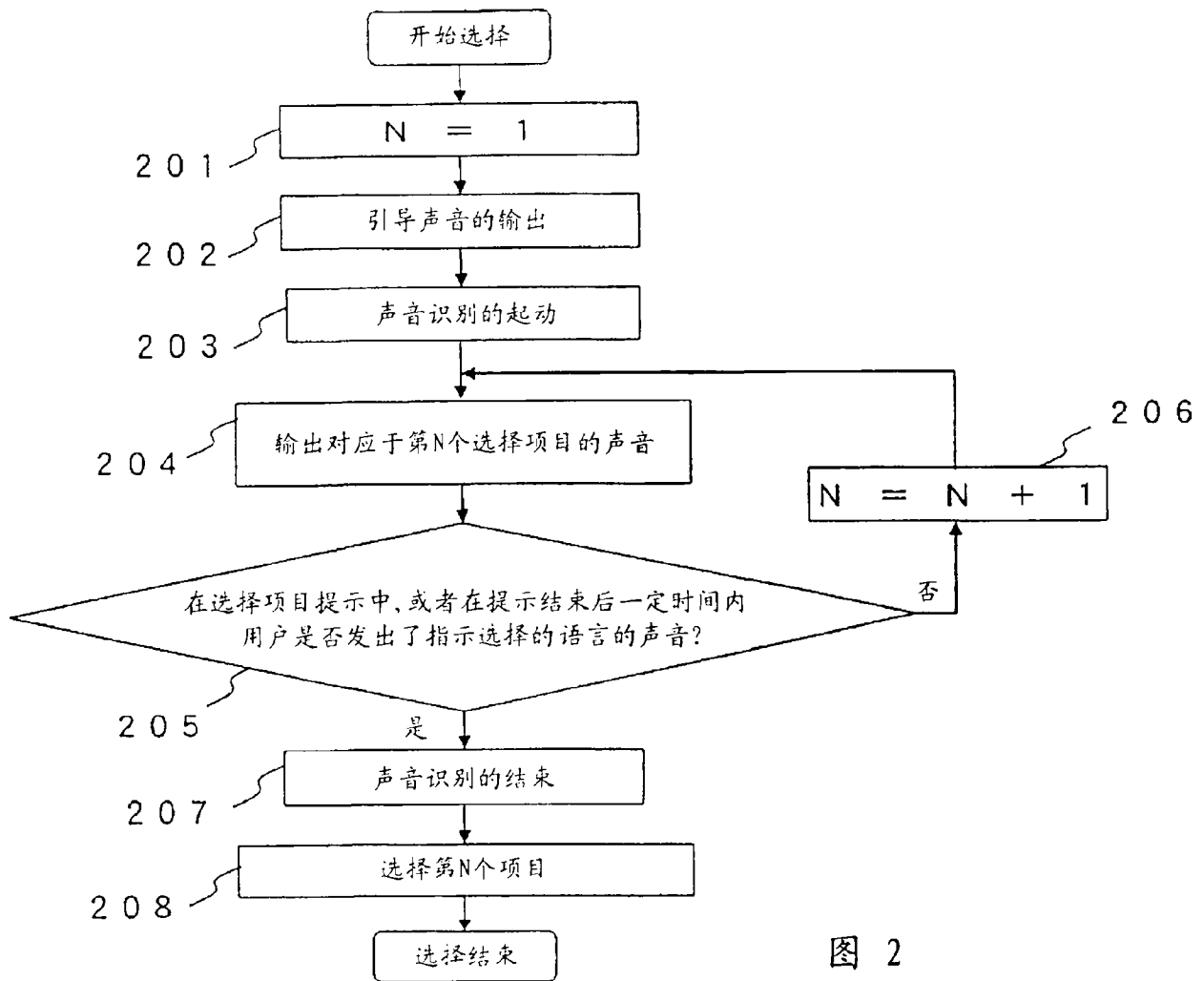


图 2

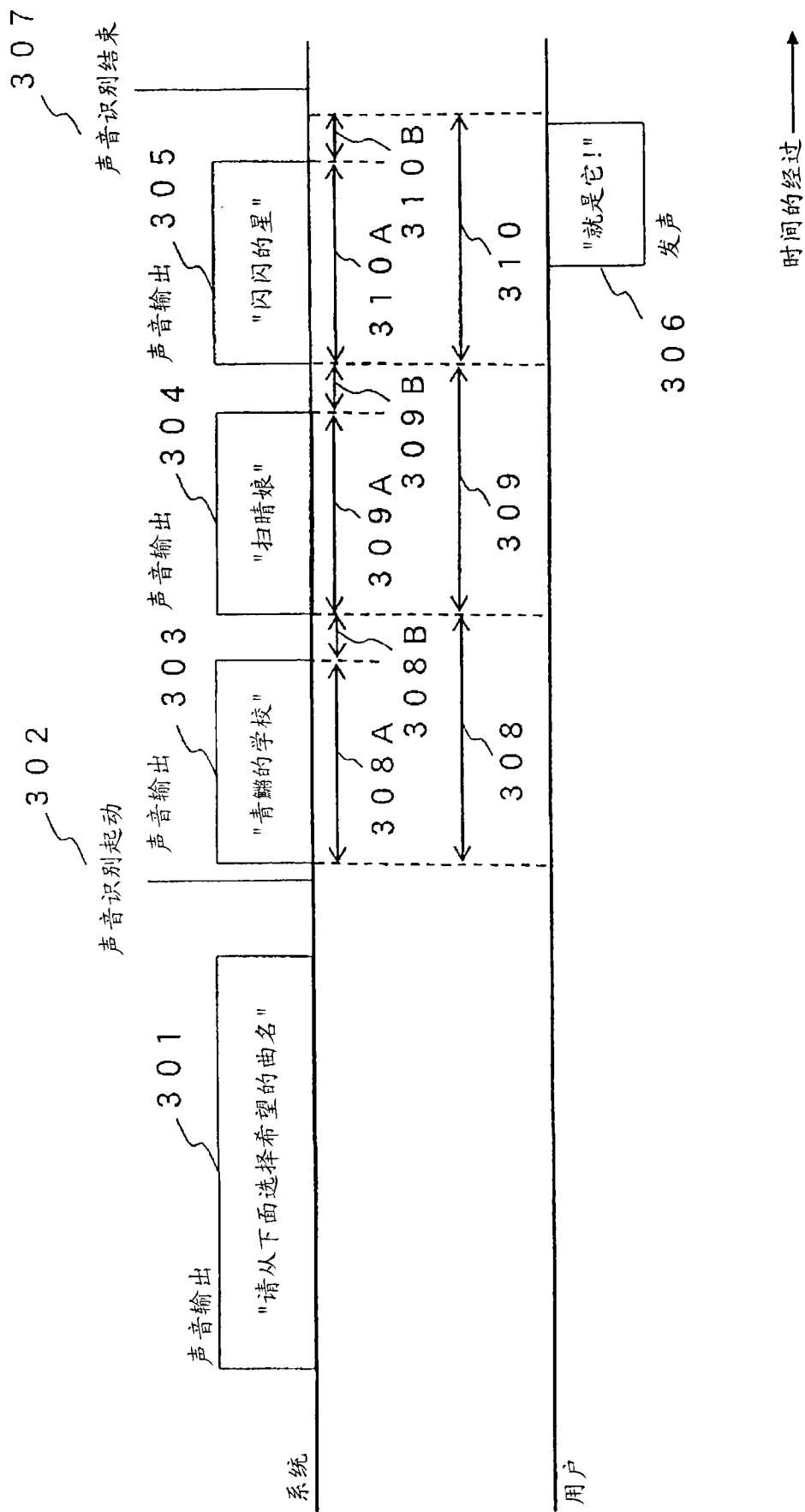
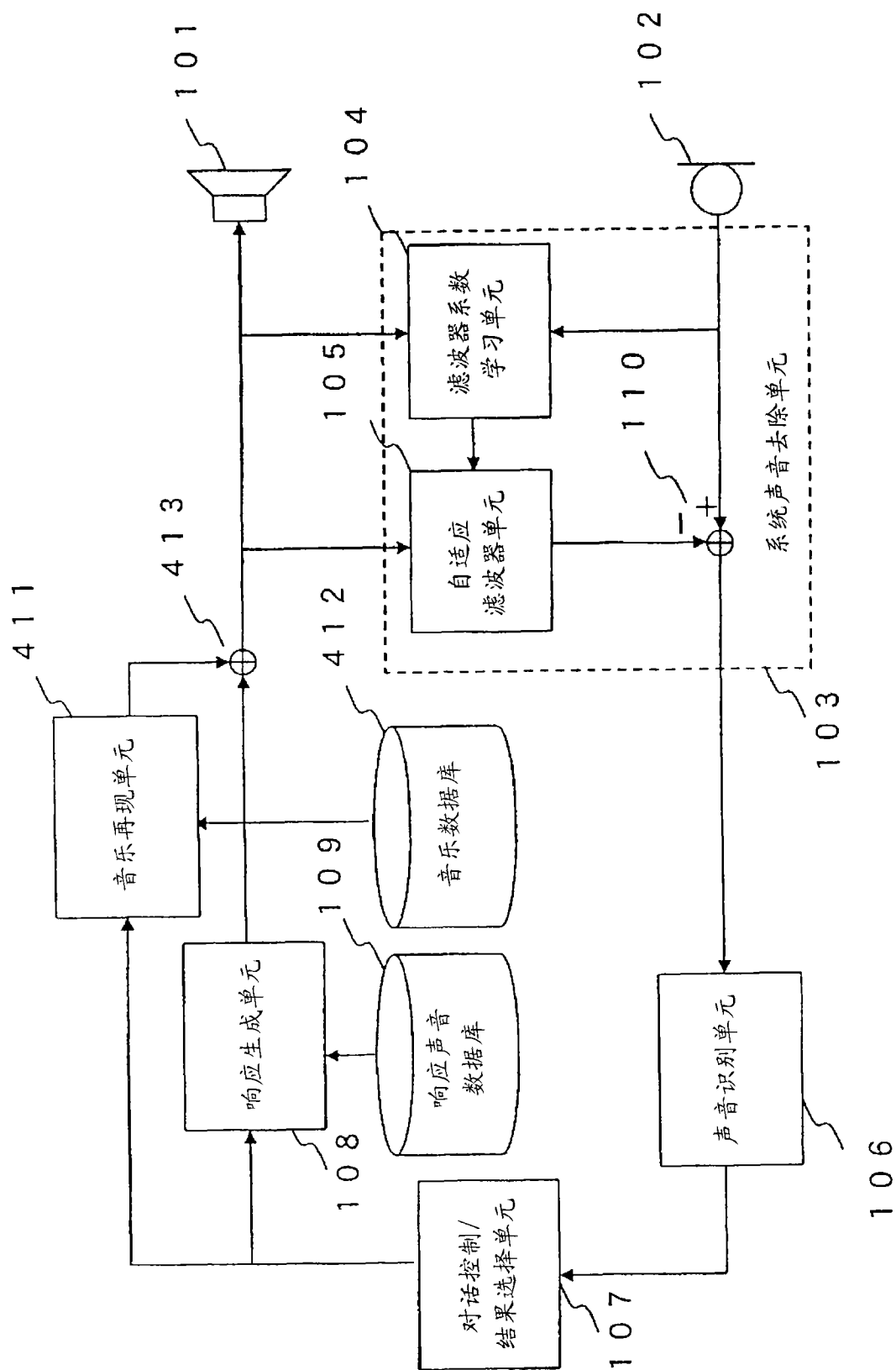


图 3



4

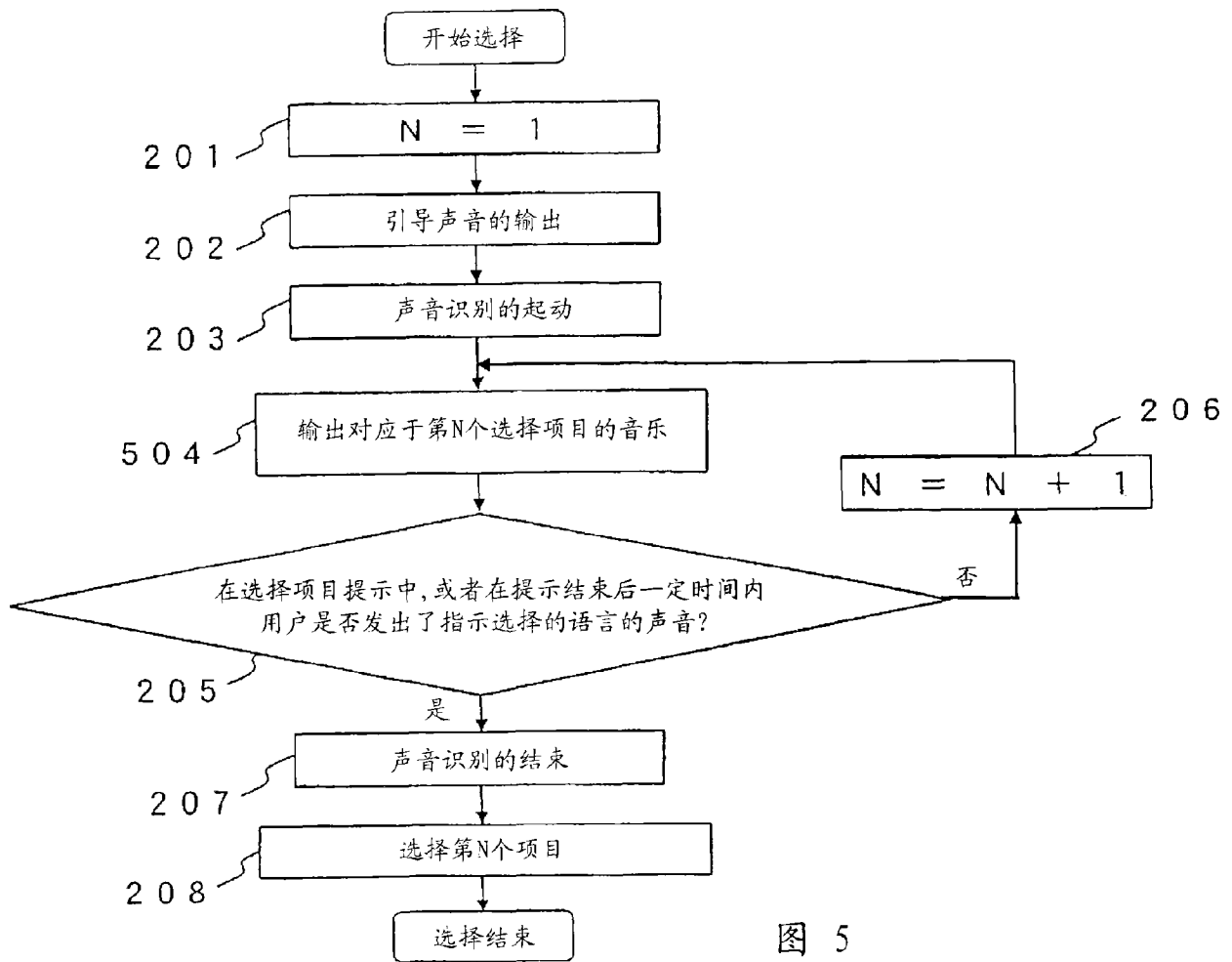


图 5

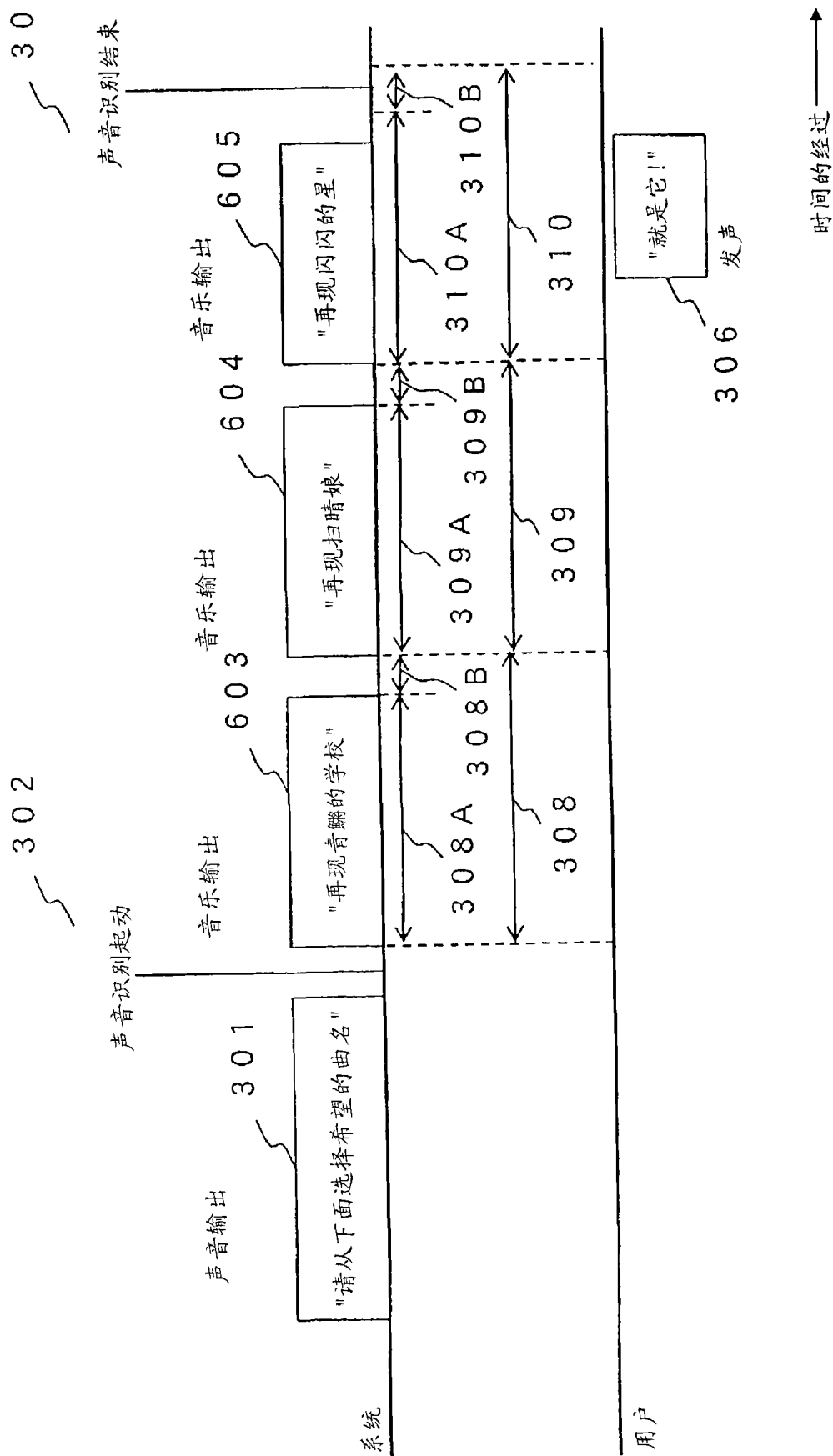


图 6

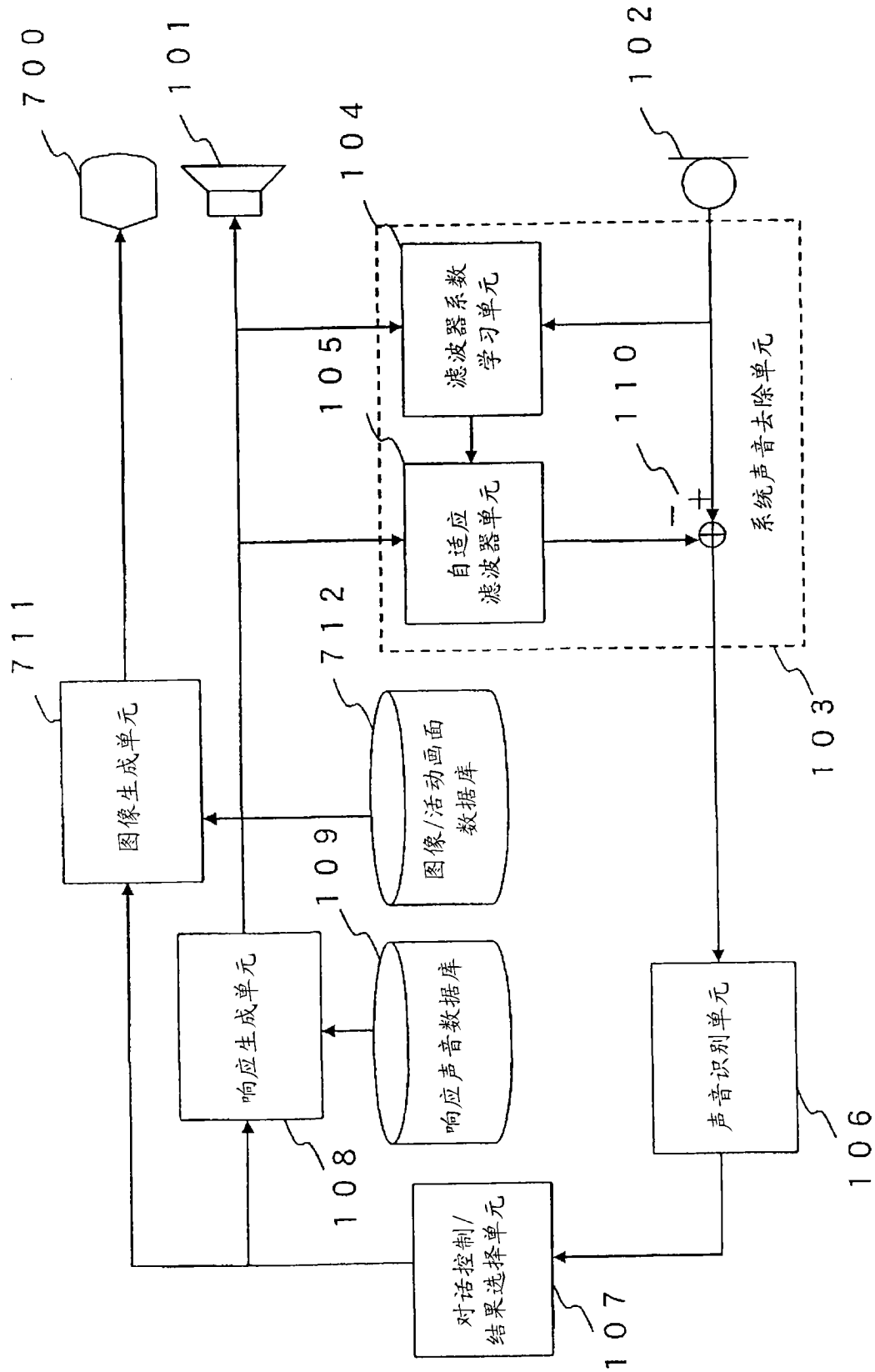


图 7

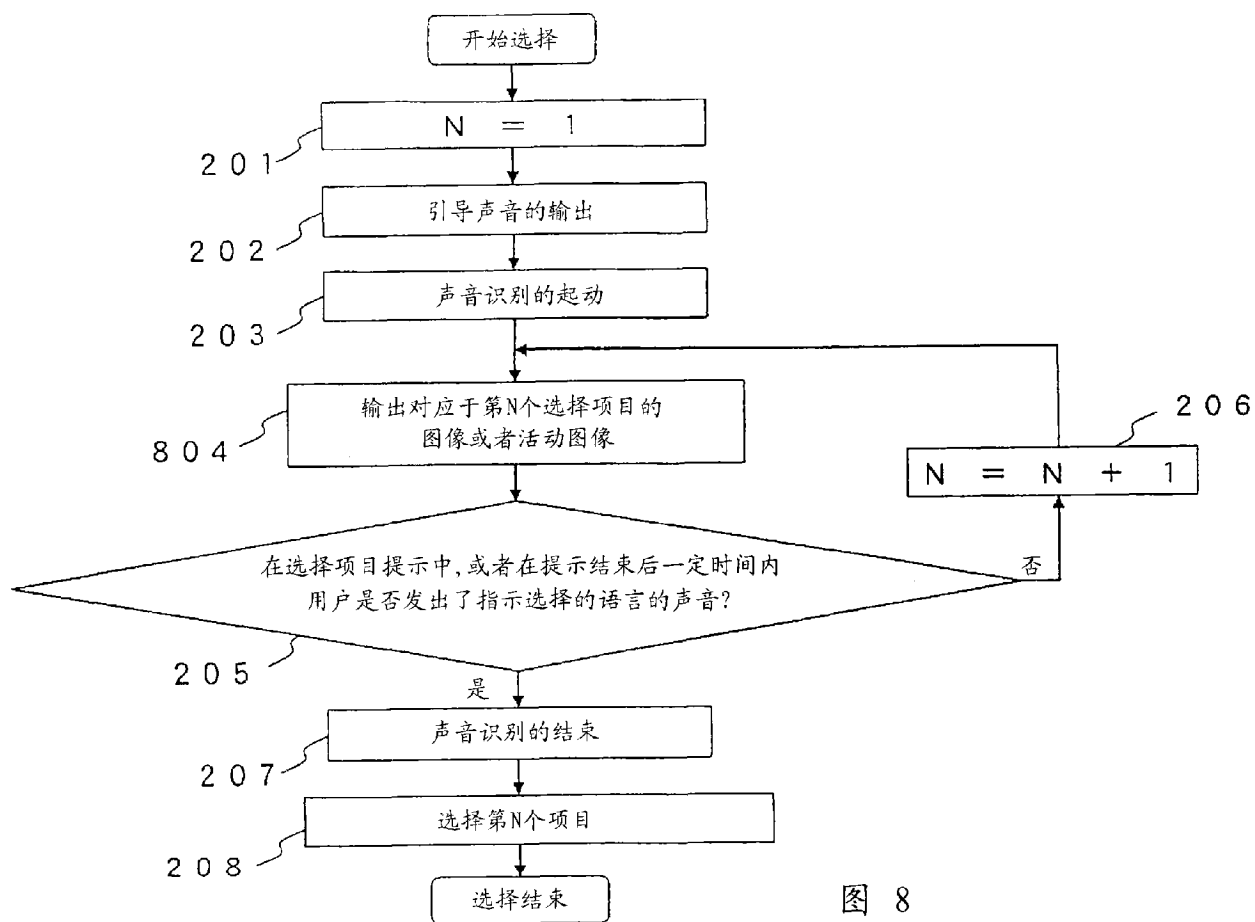
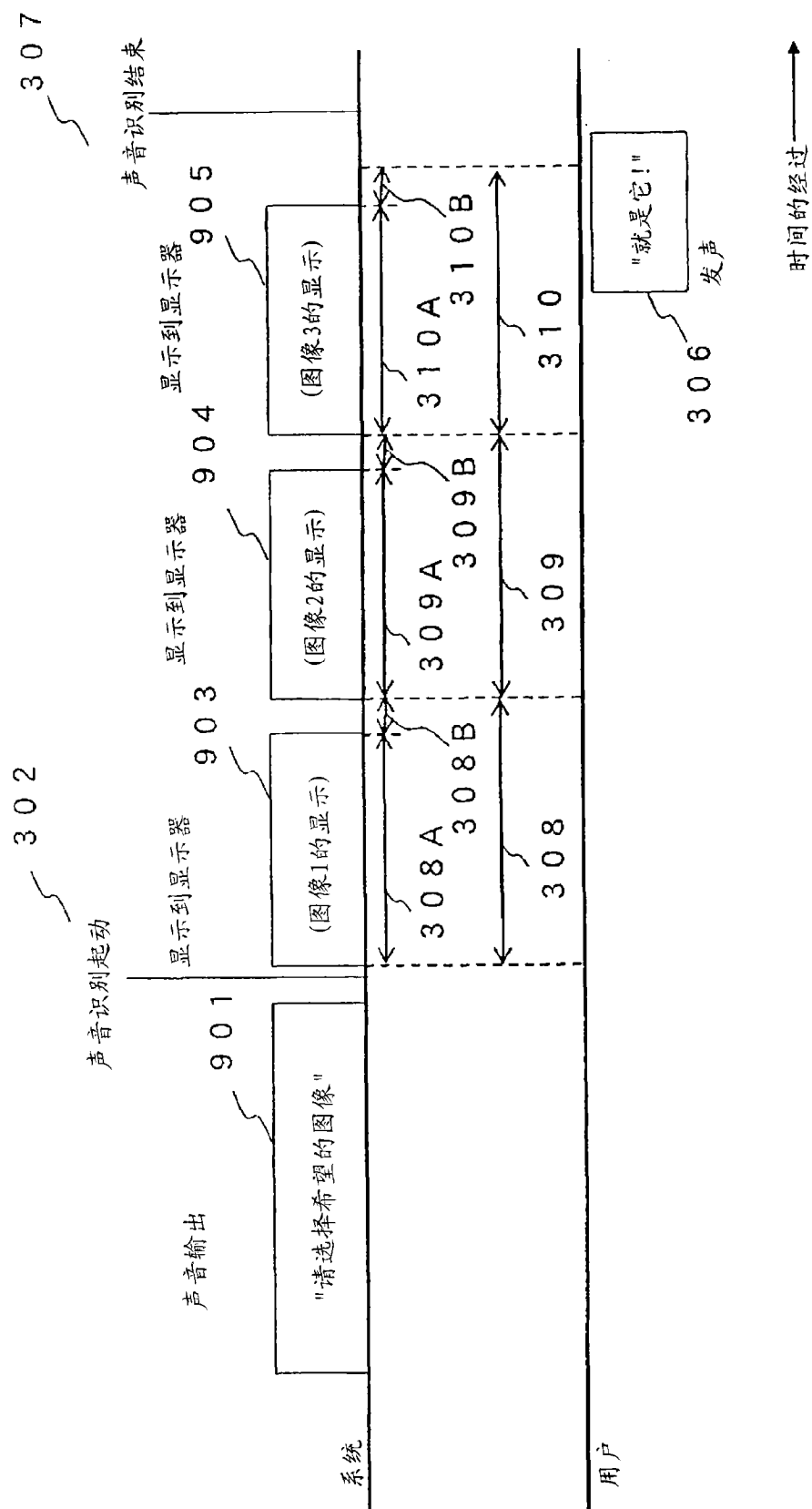


图 8



9
[X]

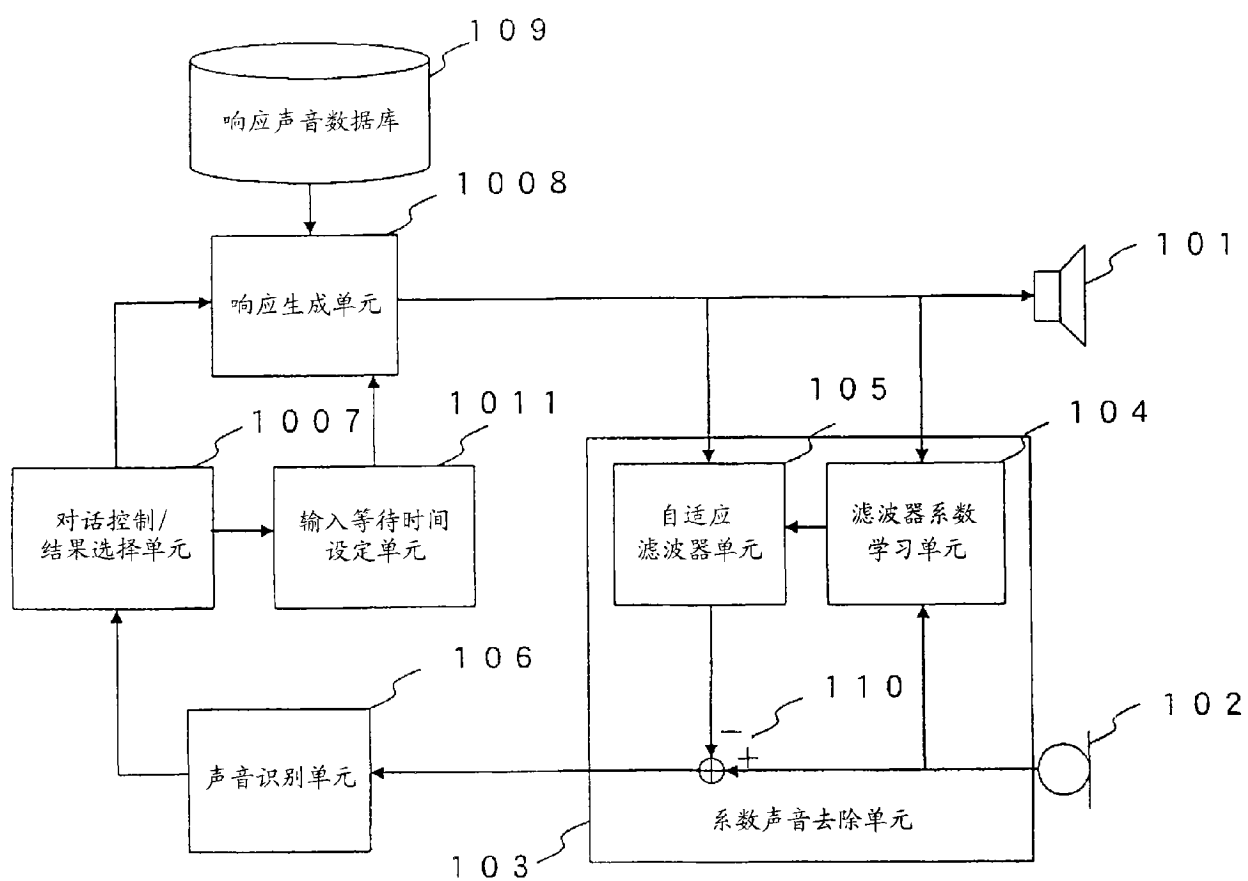


图 10

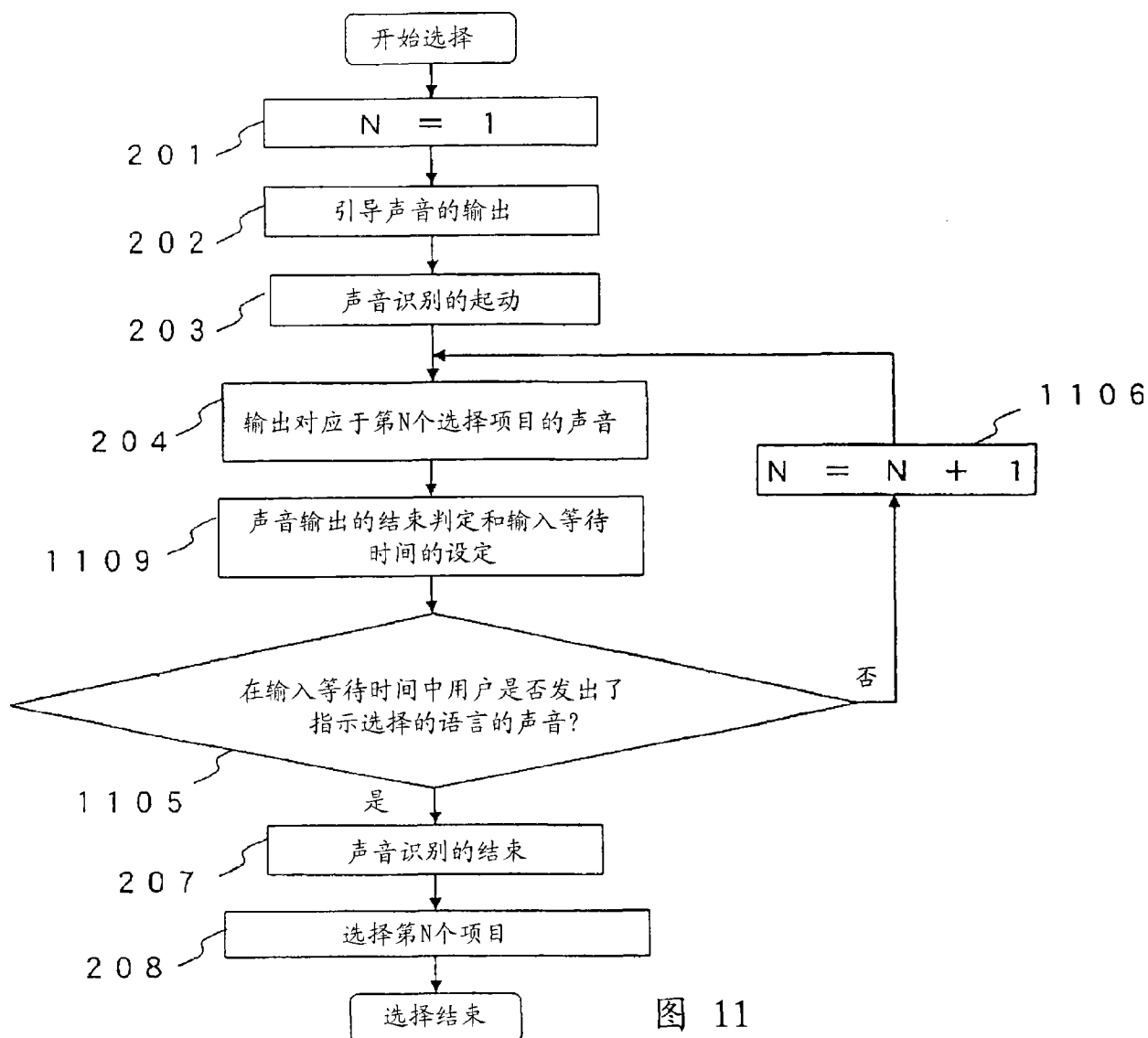


图 11

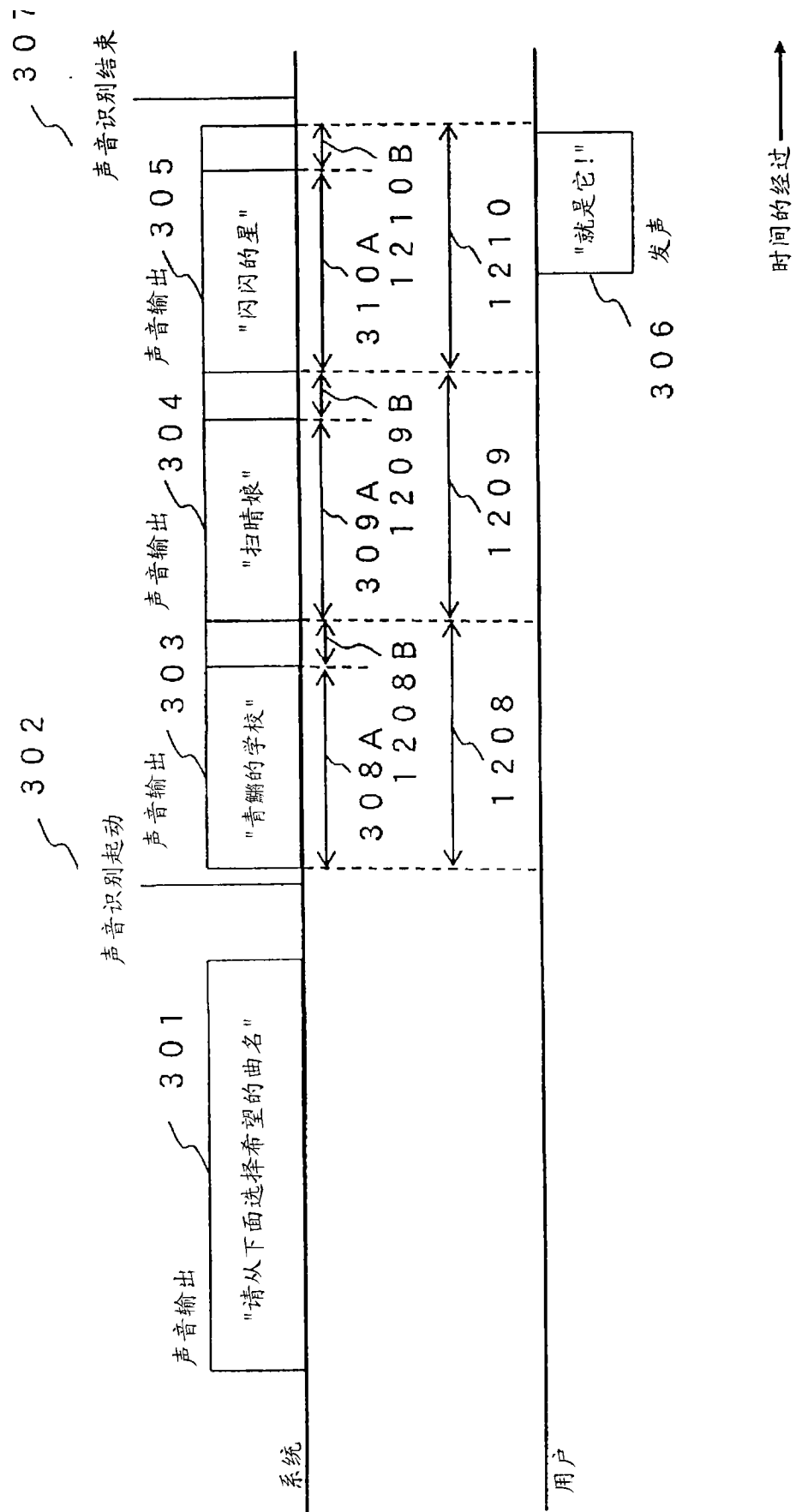


图 12