



(12) 发明专利申请

(10) 申请公布号 CN 113870856 A

(43) 申请公布日 2021. 12. 31

(21) 申请号 202111212386.9

(22) 申请日 2021.10.18

(71) 申请人 深圳追一科技有限公司

地址 518000 广东省深圳市南山区粤海街道科技园社区科苑路8号讯美科技广场3号楼23A、23B

(72) 发明人 王雨辰 刘致远 刘云峰

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 周初冬

(51) Int. Cl.

G10L 15/22 (2006.01)

G10L 15/02 (2006.01)

G10L 15/28 (2013.01)

G06K 9/00 (2006.01)

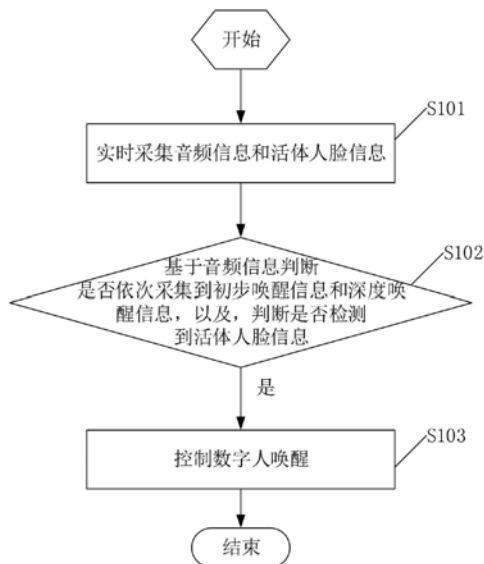
权利要求书2页 说明书11页 附图10页

(54) 发明名称

一种公共场所数字人防误唤醒方法、相关装置和存储介质

(57) 摘要

本发明提供一种公共场所数字人防误唤醒方法、相关装置和存储介质,该方法,包括:实时采集音频信息和活体人脸信息;基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息;若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,控制数字人唤醒;需要分别满足上述多个条件之后,才控制数字人唤醒,另外,该深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;也就是说,在获取到初步唤醒信息之后,还需要与用户进行一个问答的过程,以确认用户唤醒数字人的需求;并不是由用户单方面的音频或人脸图像进行唤醒,降低了误唤醒率,降低数字人的耗能。



1. 一种公共场所数字人防误唤醒方法,其特征在于,包括:  
实时采集音频信息和活体人脸信息;  
基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到所述活体人脸信息;其中,所述深度唤醒信息为在采集到所述初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;  
若依次采集到所述初步唤醒信息和所述深度唤醒信息且检测到所述活体人脸信息,控制所述数字人唤醒。
2. 根据权利要求1所述的公共场所数字人防误唤醒方法,其特征在于,基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,包括:  
基于所述音频信息判断是否采集到预设的唤醒关键字信息;所述唤醒关键字信息为所述初步唤醒信息;  
若是,则启动交互应答模式、输出唤醒询问信息;  
基于所述音频信息判断是否采集到所述唤醒询问信息对应的唤醒确定信息。
3. 根据权利要求2所述的公共场所数字人防误唤醒方法,其特征在于,启动交互应答模式、输出唤醒询问信息,包括:  
启动所述交互应答模式;  
在启动所述交互应答模式的预设时长后,输出唤醒询问信息。
4. 根据权利要求3所述的公共场所数字人防误唤醒方法,其特征在于,在启动所述交互应答模式的预设时长后,输出唤醒询问信息,包括:  
若在启动所述交互应答模式的预设时长内采集到用户的咨询信息,则输出基于所述咨询信息的唤醒询问信息。
5. 根据权利要求3所述的公共场所数字人防误唤醒方法,其特征在于,在启动所述交互应答模式的预设时长后,输出唤醒询问信息,包括:  
若在启动所述交互应答模式的预设时长内未采集到用户的咨询信息,则输出预设的唤醒询问信息。
6. 根据权利要求1-5任一项所述的公共场所数字人防误唤醒方法,其特征在于,控制所述数字人唤醒,包括:  
激活网络通信模式;并开始拾音发送到服务器,进行数字人服务。
7. 根据权利要求1-5任一项所述的公共场所数字人防误唤醒方法,其特征在于,判断是否检测到所述活体人脸信息,包括:  
基于实时采集活体人脸信息时采集到的人脸图像信息,判断所述人脸图像信息是否为用户的正面人脸图像信息;  
若是,则判定为检测到所述活体人脸信息;若否,则判定为未检测到所述活体人脸信息。
8. 一种公共场所数字人防误唤醒装置,其特征在于,包括:  
采集模块,用于实时采集音频信息和活体人脸信息;  
判断模块,用于基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到所述活体人脸信息;所述深度唤醒信息为在采集到所述初步唤醒信息后输出的唤醒询问信息对应的唤醒确定信息;

唤醒模块,用于若依次采集到所述初步唤醒信息和所述深度唤醒信息且检测到所述活体人脸信息,控制所述数字人唤醒。

9.一种电子设备,其特征在于,包括:

一个或多个处理器;

存储装置,其上存储有一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1-7中任一项所述的公共场所数字人防误唤醒方法。

10.一种存储介质,其特征在于,其上存储有计算机程序,其中,所述计算机程序被处理器执行时实现如权利要求1-7中任一项所述的公共场所数字人防误唤醒方法。

## 一种公共场所数字人防误唤醒方法、相关装置和存储介质

### 技术领域

[0001] 本发明属于人工智能技术领域,更具体的说,尤其涉及一种公共场所数字人防误唤醒方法、相关装置和存储介质。

### 背景技术

[0002] 随着科学技术的不断进步和数字人技术的不断发展,数字人已逐渐走入千家万户,市场上也出现了不少数字人,给人们的生活带来便利和乐趣,其中,交互机器人作为智能机器人的一种,能够和人们互动,给人们的生活提供了便利。

[0003] 随着社会生产力的发展和生活水平的提高,人们将越来越多的时间用于休闲和娱乐,从而极大的促进了服务业的发展。近年来,公共场所,比如机场和商城,人口流动剧增。

[0004] 目前,数字人在公共场所只采用红外人体监测误唤醒概率非常高,只采用关键信息唤醒,因为场所内人员交流信息和广播信息到处都是,误唤醒概率也非常高。也即,数字人长时间处于唤醒状态,误唤醒率较大,其耗能较大。

### 发明内容

[0005] 有鉴于此,本发明的目的在于提供一种公共场所数字人防误唤醒方法、相关装置和存储介质,用于降低误唤醒率,降低数字人的耗能。

[0006] 本申请第一方面公开了一种公共场所数字人防误唤醒方法,包括:

[0007] 实时采集音频信息和活体人脸信息;

[0008] 基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到所述活体人脸信息;其中,所述深度唤醒信息为在采集到所述初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;

[0009] 若依次采集到所述初步唤醒信息和所述深度唤醒信息且检测到所述活体人脸信息,控制所述数字人唤醒。

[0010] 可选的,在上述公共场所数字人防误唤醒方法中,基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,包括:

[0011] 基于所述音频信息判断是否采集到预设的唤醒关键字信息;所述唤醒关键字信息为所述初步唤醒信息;

[0012] 若是,则启动交互应答模式、输出唤醒询问信息;

[0013] 基于所述音频信息判断是否采集到所述唤醒询问信息对应的唤醒确定信息。

[0014] 可选的,在上述公共场所数字人防误唤醒方法中,启动交互应答模式、输出唤醒询问信息,包括:

[0015] 启动所述交互应答模式;

[0016] 在启动所述交互应答模式的预设时长后,输出唤醒询问信息。

[0017] 可选的,在上述公共场所数字人防误唤醒方法中,在启动所述交互应答模式的预设时长后,输出唤醒询问信息,包括:

[0018] 若在启动所述交互应答模式的预设时长内采集到用户的咨询信息,则输出基于所述咨询信息的唤醒询问信息。

[0019] 可选的,在上述公共场所数字人防误唤醒方法中,在启动所述交互应答模式的预设时长后,输出唤醒询问信息,包括:

[0020] 若在启动所述交互应答模式的预设时长内未采集到用户的咨询信息,则输出预设的唤醒询问信息。

[0021] 可选的,在上述公共场所数字人防误唤醒方法中,控制所述数字人唤醒,包括:

[0022] 激活网络通信模式;并开始拾音发送到服务器,进行数字人服务。

[0023] 可选的,在上述公共场所数字人防误唤醒方法中,判断是否检测到所述活体人脸信息,包括:

[0024] 基于实时采集活体人脸信息时采集到的人脸图像信息,判断所述人脸图像信息是否为用户的正面人脸图像信息;

[0025] 若是,则判定为检测到所述活体人脸信息;若否,则判定为未检测到所述活体人脸信息。

[0026] 本申请第二方面公开了一种公共场所数字人防误唤醒装置,包括:

[0027] 采集模块,用于实时采集音频信息和活体人脸信息;

[0028] 判断模块,用于基于所述音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到所述活体人脸信息;所述深度唤醒信息为在采集到所述初步唤醒信息后输出的唤醒询问信息对应的唤醒确定信息;

[0029] 唤醒模块,用于若依次采集到所述初步唤醒信息和所述深度唤醒信息且检测到所述活体人脸信息,控制所述数字人唤醒。

[0030] 可选的,在上述公共场所数字人防误唤醒装置中,所述判断模块,包括:

[0031] 第一判断单元,用于基于所述音频信息判断是否采集到预设的唤醒关键字信息;所述唤醒关键字信息为所述初步唤醒信息;

[0032] 交互应答模块,用于若所述判断单元的判断结果为是,则启动交互应答模式、输出唤醒询问信息;

[0033] 所述第一判断单元还用于基于所述音频信息判断是否采集到所述唤醒询问信息对应的唤醒确定信息。

[0034] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块用于启动交互应答模式、输出唤醒询问信息时,具体用于:

[0035] 启动所述交互应答模式;

[0036] 在启动所述交互应答模式的预设时长后,输出唤醒询问信息。

[0037] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块用于在启动所述交互应答模式的预设时长后,输出唤醒询问信息时,具体用于:

[0038] 若在启动所述交互应答模式的预设时长内采集到用户的咨询信息,则输出基于所述咨询信息的唤醒询问信息。

[0039] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块在启动所述交互应答模式的预设时长后,输出唤醒询问信息时,具体用于:

[0040] 若在启动所述交互应答模式的预设时长内未采集到用户的咨询信息,则输出预设

的唤醒询问信息。

[0041] 可选的,在上述公共场所数字人防误唤醒装置中,唤醒模块用于控制所述数字人唤醒时,具体用于:

[0042] 激活网络通信模式;并开始拾音发送到服务器,进行数字人服务。

[0043] 可选的,在上述公共场所数字人防误唤醒装置中,所述判断模块还包括:

[0044] 第二判断单元,用于基于实时采集活体人脸信息时采集到的人脸图像信息,判断所述人脸图像信息是否为用户的正面人脸图像信息;

[0045] 若是,则判定为检测到所述活体人脸信息;若否,则判定为未检测到所述活体人脸信息。

[0046] 本申请第三方面公开了一种电子设备,包括:

[0047] 一个或多个处理器;

[0048] 存储装置,其上存储有一个或多个程序;

[0049] 当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如本申请第一方面任一项所述的公共场所数字人防误唤醒方法。

[0050] 本申请第四方面公开了一种存储介质,其上存储有计算机程序,其中,所述计算机程序被处理器执行时实现如本申请第一方面任一项所述的公共场所数字人防误唤醒方法。

[0051] 从上述技术方案可知,本发明提供的一种公共场所数字人防误唤醒方法,包括:包括:实时采集音频信息和活体人脸信息;基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息;若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,控制数字人唤醒;需要分别满足上述多个条件之后,才控制数字人唤醒,另外,该深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;也就是说,在获取到初步唤醒信息之后,还需要与用户进行一个问答的过程,以确认用户唤醒数字人的需求;并不是由用户单方面的音频或人脸图像进行唤醒,降低了误唤醒率,降低数字人的耗能。

## 附图说明

[0052] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0053] 图1是本发明实施例提供的一种适用于本申请实施例的应用环境示意图;

[0054] 图2是本发明实施例提供的一种公共场所数字人防误唤醒方法的流程图;

[0055] 图3是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0056] 图4是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0057] 图5是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0058] 图6是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0059] 图7是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0060] 图8是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0061] 图9是本发明实施例提供的另一种公共场所数字人防误唤醒方法的流程图;

[0062] 图10是本发明实施例提供的一种公共场所数字人防误唤醒装置的示意图；

[0063] 图11是本发明实施例提供的一种电子设备的示意图。

### 具体实施方式

[0064] 为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0065] 在本申请中，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0066] 请参阅图1，图1示出了一种适用于本申请实施例的应用环境示意图。本申请实施例提供的数据处理方法可以应该用于如图1所示的交互系统100。交互系统100包括终端设备101以及服务器102，服务器102与终端设备101通信连接。其中，服务器102可以是传统服务器，也可以是云端服务器，在此不做具体限定。

[0067] 其中，终端设备101可以是具有显示屏、具有数据处理模块、具有拍摄相机、具有音频输入/输出等功能，且支持数据输入的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机、台式计算机、自助服务终端和可穿戴式电子设备等。具体的，数据输入可以是基于电子设备上具有的语音模块输入语音、字符输入模块输入字符等。具体的，该终端设备101可以是下述实施例中的数字人。

[0068] 其中，终端设备101上可以安装有客户端应用程序，用户可以基于客户端应用程序（例如APP，微信小程序等），其中，本实施例的对话数字人也是配置于终端设备101中的一个客户端应用程序。用户可以基于客户端应用程序在服务器102注册一个用户账号，并基于该用户账号与服务器102进行通信，例如用户在客户端应用程序登录用户账号，并基于该用户账号通过客户端应用程序进行输入，可以输入文字信息或语音信息等，客户端应用程序接收到用户输入的信息后，可以将该信息发送至服务器102，使得服务器102可以接收该信息并进行处理及存储，服务器102还可以接收该信息并根据该信息返回一个对应的输出信息至终端设备101。

[0069] 在一些实施方式中，对待识别数据进行处理装置也可以设置于终端设备101上，使得终端设备101无需依赖于服务器102建立通信即可实现与用户的交互，此时交互系统100可以只包括终端设备101。

[0070] 本申请实施例提供了一种公共场所数字人防误唤醒方法，用于解决现有技术中数字人在公共场所只采用红外人体监测误唤醒概率非常高，只采用关键信息唤醒，因为场所内人员交流信息和广播信息到处都是，误唤醒概率也非常高。也即，数字人长时间处于唤醒状态，误唤醒率较大，其耗能较大的问题。

[0071] 参见图2，该公共场所数字人防误唤醒方法，包括：

[0072] S101、实时采集音频信息和活体人脸信息。

[0073] 需要说明的是,该数字人具备采集器件,如通过音频采集器采集音频信息,通过人脸采集器采集活体人脸信息。

[0074] 活体人脸信息可以包括活体人脸图像。该图像可以是二维图像、三维图像和红外图像等,此处不做具体限定,视实际情况而定即可,均在本申请的保护范围内。

[0075] S102、基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息。

[0076] 具体的,该步骤S102可以划分为两个判断;在两个判断结果均为是的情况下,才能说明步骤S102的判断结果为依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息。

[0077] 需要说明的是,基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息的具体过程可以为:将采集到的音频信息转化为文字信息,再判断该文字信息中是否存在该初步唤醒信息和深度唤醒信息。

[0078] 其中,深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息。

[0079] 也就是说,在采集到初步唤醒信息后,输出唤醒询问信息;在输出唤醒询问信息后,依据当前采集到的音频信息,判断是否采集到与该唤醒询问信息对应的唤醒确定信息。

[0080] 该基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息的具体过程,此处不再一一赘述,视实际情况而定即可,均在本申请的保护范围内。

[0081] 判断是否检测到活体人脸信息的具体过程可以是:判断采集到的人脸图像是否满足预设条件;该预设条件可以包括:面部采集到的面部占比、是否为活体信息等。

[0082] 该判断是否检测到活体人脸信息的具体过程,此处不再一一赘述,视实际情况而定即可,均在本申请的保护范围内。

[0083] 步骤S102中的两个判断步骤是并列,也就是说,其没有先后关系。因此,步骤S102可以是同一时刻两个判断结果为是时,判定为依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息。

[0084] 也可以是设置两个判断之间的时间差;具体的,若任一个判断结果为是,且在预设时长内另一个判断结果为是,则判定为依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息。

[0085] 若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,则执行步骤S103。

[0086] S103、控制数字人唤醒。

[0087] 需要说明的是,在数字人被唤醒之后,该数字人可以为用户提供相应服务。

[0088] 如用户向数字人提出办理信用卡流程的咨询,则数字人向用户展示办理信用卡的流程。具体的,可以通过数字人自身的显示器显示该流程,也可以同时输出办理信用卡流程的音频。

[0089] 需要说明的是,步骤S101是实时采集的,也就是说,步骤S102和S103的动作并不会影响步骤S101的执行。同理,由于步骤S101实时采集到数据,步骤S102也需实时进行判断,以跟上步骤S101的节奏,对采集到的所有数据均进行判断。

[0090] 在实际应用中,若未依次采集到初步唤醒信息和深度唤醒信息,或检测到活体人



脸信息,则不执行步骤S103。

[0091] 在本实施例中,实时采集音频信息和活体人脸信息;基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息;若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,控制数字人唤醒;需要分别满足上述多个条件之后,才控制数字人唤醒,另外,该深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;也就是说,在获取到初步唤醒信息之后,还需要与用户进行一个问答的过程,以确认用户唤醒数字人的需求;并不是由用户单方面的音频或人脸图像进行唤醒,降低了误唤醒率,降低数字人的耗能。

[0092] 在上述实施例中,步骤S102中所涉及的基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,参见图3,包括:

[0093] S201、基于音频信息判断是否采集到预设的唤醒关键字信息。

[0094] 其中,唤醒关键字信息为初步唤醒信息。

[0095] 该唤醒关键字信息,可以是该数字人的姓名,如“小一”等;当然也可以是其他信息,此处不做具体限定,视实际情况而定即可,均在本申请的保护范围内。

[0096] 该唤醒关键字信息可以是多个,如分别为:“小一”“你好”等;当然也可以是一个;此处不做具体限定,视实际情况而定即可,均在本申请的保护范围内。

[0097] 若基于音频信息判断是否采集到预设的唤醒关键字信息,则执行步骤S202。

[0098] S202、启动交互应答模式、输出唤醒询问信息。

[0099] 需要说明的是,该数字人内部设置有交互应答模块,控制该交互应答模块启动,即可实现启动交互应答模式。

[0100] 该数字人可以是有两级唤醒;如初步唤醒和深度唤醒。上述步骤S103中的唤醒为深度唤醒。

[0101] 具体的,该数字人可以包括采集模块、交互应答模块和网络通信模块。在数字人上电时,该采集模块在持续工作。

[0102] 在交互应答模块启动时,该数字人处于初级唤醒;此时数字人的耗能大于仅采集模块工作时的耗能。

[0103] 在网络通信模块均启动时,数字人处于深度唤醒;此时数字人的耗能大于初级唤醒时的耗能。

[0104] S203、基于音频信息判断是否采集到唤醒询问信息对应的唤醒确定信息。

[0105] 也就是说,该交互应答模式下,该数字人可以与用户进行交互,通过询问以及接收用户的应答来确定是否需要深度唤醒。

[0106] 需要说明的是,若输出唤醒询问信息后的第一时长内,采集到对应的唤醒确定信息,则确定需要进行深度唤醒。若输出唤醒询问信息后的第一时长内未采集到对应的唤醒确定信息或者采集到唤醒取消信息,则确定不需要深度唤醒。

[0107] 如唤醒询问信息为:请问你是想咨询小一问题吗?则唤醒确定信息可以是:是的、对的等带有确定性的词句;唤醒取消信息可以是:不用、不是等带有否定性的词句。

[0108] 当然,各个信息不仅限于上述说明,此处不再一一赘述,视实际情况而定即可,均在本申请的保护范围内。

[0109] 在本实施例中,采用两级唤醒的方式来进行唤醒,其中深度唤醒为数字人获取到

询问信息对应的确定信息之后,再进行唤醒,避免了数字人误唤醒,降低数字人的功耗。

[0110] 在实际应用中,参见图4,步骤S202的具体过程可以是:

[0111] S301、启动交互应答模式。

[0112] S302、在启动交互应答模式的预设时长后,输出唤醒询问信息。

[0113] 也就是说,在启动交互应答模式后并不是直接输出唤醒询问信息,而是等待预设时长后,再输出唤醒询问信息。

[0114] 该预设时长的取值,此处不做具体限定,视实际情况而定即可,均在本申请的保护范围内。

[0115] 当然,本申请也不排除启动交互应答模式后直接输出唤醒询问信息的方案,此处不再一一赘述,视实际情况而定即可,均在本申请的保护范围内。

[0116] 例如,启动交互应答模式,在启动交互应答模式一分钟后,输出“确定唤醒小一吗?”的唤醒询问信息。

[0117] 上述过程仅是一种示例,此处不再一一赘述,视实际情况而定即可,均在本申请的保护范围内。

[0118] 在本实施例中,通过在启动交互应答模式的预设时长后,输出唤醒询问信息,并不是立即输出唤醒询问信息,可以预留预设时长来进行数据处理。

[0119] 在实际应用中,参见图5,步骤S302的具体过程为:

[0120] S401、判断启动交互应答模式的预设时长内是否采集到用户的咨询信息。

[0121] 该预设时长的具体取值,此处不做具体限定,视实际情况而定即可,均在本申请的保护范围内。

[0122] 若在启动交互应答模式的预设时长内采集到用户的咨询信息,则执行步骤S402。

[0123] S402、输出基于咨询信息的唤醒询问信息。

[0124] 具体的,该唤醒询问信息主要为了确定用户的需求,也即如采集到咨询信息,则需进行确认该咨询信息是否为用户需要向数字人进行咨询的信息。

[0125] 例如,用户对数字人说“咨询信用卡办理流程”;也即,用户的咨询信息为:信用卡办理流程。此时,数字人对用户输出“是否确定要咨询信用卡办理流程”;也即,基于咨询信息的唤醒询问信息为:是否确定要咨询信用卡办理流程。

[0126] 在本实施例中,通过判断启动交互应答模式的预设时长内是否采集到用户的咨询信息,并在启动交互应答模式的预设时长内采集到用户的咨询信息时,输出基于咨询信息的唤醒询问信息,进而使数字人输出的唤醒询问信息与用户的咨询信息相关,使数字人的人性化。

[0127] 在实际应用中,参见图6,步骤S302的具体工作过程为:

[0128] S501、判断启动交互应答模式的预设时长内是否采集到用户的咨询信息。

[0129] 若在启动交互应答模式的预设时长内未采集到用户的咨询信息,则执行步骤S502。

[0130] S502、输出预设的唤醒询问信息。

[0131] 具体的,该唤醒询问信息主要为了确定用户唤醒的需求,也即如未采集到咨询信息,则需进行确认用户是否需要唤醒数字人。

[0132] 该预设的唤醒询问信息可以是:您确定要唤醒小一吗;当然,也不仅限于上述示

例,该预设的唤醒询问信息,此处不再赘述,视实际情况而定即可,均在本申请的保护范围内。

[0133] 在本实施例中,通过判断启动交互应答模式的预设时长内是否采集到用户的咨询信息,并在启动交互应答模式的预设时长内未采集到用户的咨询信息时,输出预设的唤醒询问信息;也即,设置数字人的默认唤醒询问信息,进而使数字人的与用户的交互信息明确。

[0134] 在实际应用中,实施例图5和图6可以单独应用,也可以是结合使用。参见图7,将步骤S501和步骤S401是一致的,所以可以省去其中一个,图7以省去步骤S501为例进行展示。

[0135] 具体的,结合后的流程为:

[0136] S401、判断启动交互应答模式的预设时长内是否采集到用户的咨询信息。

[0137] 若在启动交互应答模式的预设时长内采集到用户的咨询信息,则执行步骤S402。

[0138] S402、输出基于咨询信息的唤醒询问信息。

[0139] 若在启动交互应答模式的预设时长内未采集到用户的咨询信息,则执行步骤S502。

[0140] S502、输出预设的唤醒询问信息。

[0141] 在本实施例中,设置了两种情况下的唤醒询问信息,也即,数字人输出的唤醒询问信息与用户的咨询信息相关,或者,为默认唤醒询问信息,使数字人的人性化,多种情况下均有对应的唤醒询问信息。

[0142] 在上述实施例中,参见图8(以在图7的基础之上为例进行展示),步骤S103、控制数字人唤醒的具体过程为:

[0143] S601、激活网络通信模式;并开始拾音发送到服务器,进行数字人服务。

[0144] 需要说明的是,该数字人具备网络通信模块,控制该网络通信模块启动,即可实现激活网络通信模式。

[0145] 激活网络通信模式后,该数字人可以与该网络通信模块进行交互;具体的。

[0146] 具体的,该拾音就是把声音收集的过程。也就是说,数字人将采集到的用户咨询信息发送至服务器,服务器对该咨询信息进行处理,然后将相应的解决方案反馈给数字人,该数据人展示给用户。

[0147] 同时,该服务器还有渲染数字人,以使数字人更加贴近真实的人。如在显示器上展示数字人的虚拟脸,并渲染相应的表情,如嘟嘴、害羞等;也可以是控制数字人的四肢、以使该数字人更生动形象,如通过数字人的肢体指明方向等。

[0148] 在本实施例中,激活网络通信模式;并开始拾音发送到服务器,进行数字人服务,实现智能服务,提高数字人的服务效率。

[0149] 在上述实施例中,参见图9(以在图3的基础之上为例进行展示),步骤S102中所涉及的是否检测到活体人脸信息的具体过程为:

[0150] S701、基于实时采集活体人脸信息时采集到的人脸图像信息,判断人脸图像信息是否为用户的正面人脸图像信息。

[0151] 若是,则判定为检测到活体人脸信息;若否,则判定为未检测到活体人脸信息。

[0152] 判断是否为人脸正面图像信息的过程可以通过判断人脸图像信息是否达到预设占比,或者,两眼连接线与鼻子是否垂直等;该判断是否为人脸正面图像信息的其他实现方

式,此处不再一一赘述,均在本申请的保护范围内。

[0153] 值得说明的是,该判断是否检测到活体人脸信息的其他实现方式,如通过瞳孔检测来判断是否为人脸活体信息等,此处不再一一赘述,均在本申请的保护范围内。

[0154] 在本实施例中,基于实时采集活体人脸信息时采集到的人脸图像信息,判断人脸图像信息是否为用户的正面人脸图像信息,避免将侧脸等视线方向未指向数字人的人脸图像信息作为活体人脸信息,进一步降低误唤醒的概率,降低数字人的功耗。

[0155] 本申请另一实施例提供了一种公共场所数字人防误唤醒装置。

[0156] 参见图10,该公共场所数字人防误唤醒装置,包括:

[0157] 采集模块10,用于实时采集音频信息和活体人脸信息。

[0158] 采集模块10可以包括,音频采集模块和活体人脸采集模块;

[0159] 音频采集模块用于采集音频信息。

[0160] 活体人脸采集模块用于采集活体人脸信息。

[0161] 判断模块20,用于基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息。

[0162] 深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息对应的唤醒确定信息。

[0163] 唤醒模块30,用于若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,控制数字人唤醒。

[0164] 上述各模块的工作过程和原理,详情参见上述实施例,如实施例图2,此处不再一一赘述,均在本申请的保护范围内。

[0165] 在本实施例中,采集模块10实时采集音频信息和活体人脸信息;判断模块20基于音频信息判断是否依次采集到初步唤醒信息和深度唤醒信息,以及,判断是否检测到活体人脸信息;唤醒模块30若依次采集到初步唤醒信息和深度唤醒信息且检测到活体人脸信息,控制数字人唤醒;需要分别满足上述多个条件之后,才控制数字人唤醒,另外,该深度唤醒信息为在采集到初步唤醒信息后输出的唤醒询问信息,其对应的唤醒确定信息;也就是说,在获取到初步唤醒信息之后,还需要与用户进行一个问答的过程,以确认用户唤醒数字人的需求;并不是由用户单方面的音频或人脸图像进行唤醒,降低了误唤醒率,降低数字人的耗能。

[0166] 可选的,在上述公共场所数字人防误唤醒装置中,判断模块20,包括:

[0167] 第一判断单元,用于基于音频信息判断是否采集到预设的唤醒关键字信息;唤醒关键字信息为初步唤醒信息。

[0168] 交互应答模块,用于若判断单元的判断结果为是,则启动交互应答模式、输出唤醒询问信息。

[0169] 第一判断单元还用于基于音频信息判断是否采集到唤醒询问信息对应的唤醒确定信息。

[0170] 上述各模块单元的工作过程和原理,详情参见上述实施例,如实施例图3,此处不再一一赘述,均在本申请的保护范围内。

[0171] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块用于启动交互应答模式、输出唤醒询问信息时,具体用于:

- [0172] 启动交互应答模式。
- [0173] 在启动交互应答模式的预设时长后,输出唤醒询问信息。
- [0174] 交互应答模块用于启动交互应答模式、输出唤醒询问信息的具体过程,详情参见上述实施例,如实施例图4,此处不再一一赘述,均在本申请的保护范围内。
- [0175] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块用于在启动交互应答模式的预设时长后,输出唤醒询问信息时,具体用于:
- [0176] 若在启动交互应答模式的预设时长内采集到用户的咨询信息,则输出基于咨询信息的唤醒询问信息。
- [0177] 交互应答模块用于在启动交互应答模式的预设时长后,输出唤醒询问信息的具体过程,详情参见上述实施例,如实施例图5,此处不再一一赘述,均在本申请的保护范围内。
- [0178] 可选的,在上述公共场所数字人防误唤醒装置中,交互应答模块在启动交互应答模式的预设时长后,输出唤醒询问信息时,具体用于:
- [0179] 若在启动交互应答模式的预设时长内未采集到用户的咨询信息,则输出预设的唤醒询问信息。
- [0180] 交互应答模块用于在启动交互应答模式的预设时长后,输出唤醒询问信息的具体过程,详情参见上述实施例,如实施例图6,此处不再一一赘述,均在本申请的保护范围内。
- [0181] 可选的,在上述公共场所数字人防误唤醒装置中,唤醒模块30用于控制数字人唤醒时,具体用于:
- [0182] 激活网络通信模式;并开始拾音发送到服务器,进行数字人服务。
- [0183] 唤醒模块30用于控制数字人唤醒的具体过程,详情参见上述实施例,如实施例图8,此处不再一一赘述,均在本申请的保护范围内。
- [0184] 可选的,在上述公共场所数字人防误唤醒装置中,判断模块20还包括:
- [0185] 第二判断单元,用于基于实时采集活体人脸信息时采集到的人脸图像信息,判断人脸图像信息是否为用户的正面人脸图像信息。
- [0186] 若是,则判定为检测到活体人脸信息;若否,则判定为未检测到活体人脸信息。
- [0187] 第二判断单元的具体过程和原理,详情参见上述实施例,如实施例图9,此处不再一一赘述,均在本申请的保护范围内。
- [0188] 本申请另一实施例提供了一种电子设备。
- [0189] 参见图11,该电子设备,包括:
- [0190] 一个或多个处理器110。
- [0191] 存储装置120,其上存储有一个或多个程序。
- [0192] 当一个或多个程序被一个或多个处理器110执行时,使得一个或多个处理器110实现如上述实施例提供的公共场所数字人防误唤醒方法。
- [0193] 该公共场所数字人防误唤醒方法的工作过程和原理,详情参见上述实施例,此处不再一一赘述,均在本申请的保护范围内。
- [0194] 本申请另一实施例提供了一种存储介质。
- [0195] 该存储介质上存储有计算机程序,其中,计算机程序被处理器执行时实现如上述实施例提供的公共场所数字人防误唤醒方法。
- [0196] 该公共场所数字人防误唤醒方法的工作过程和原理,详情参见上述实施例,此处

不再一一赘述,均在本申请的保护范围内。

[0197] 本说明书中的各个实施例中记载的特征可以相互替换或者组合,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统或系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统及系统实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0198] 专业人员还可以进一步意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0199] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

# 交互系统100

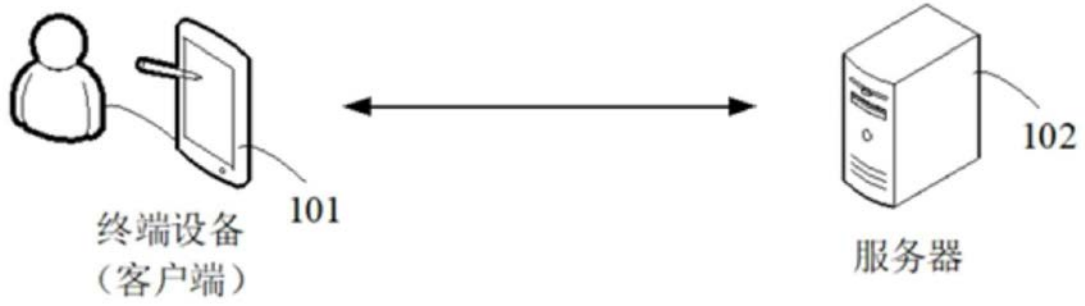


图1

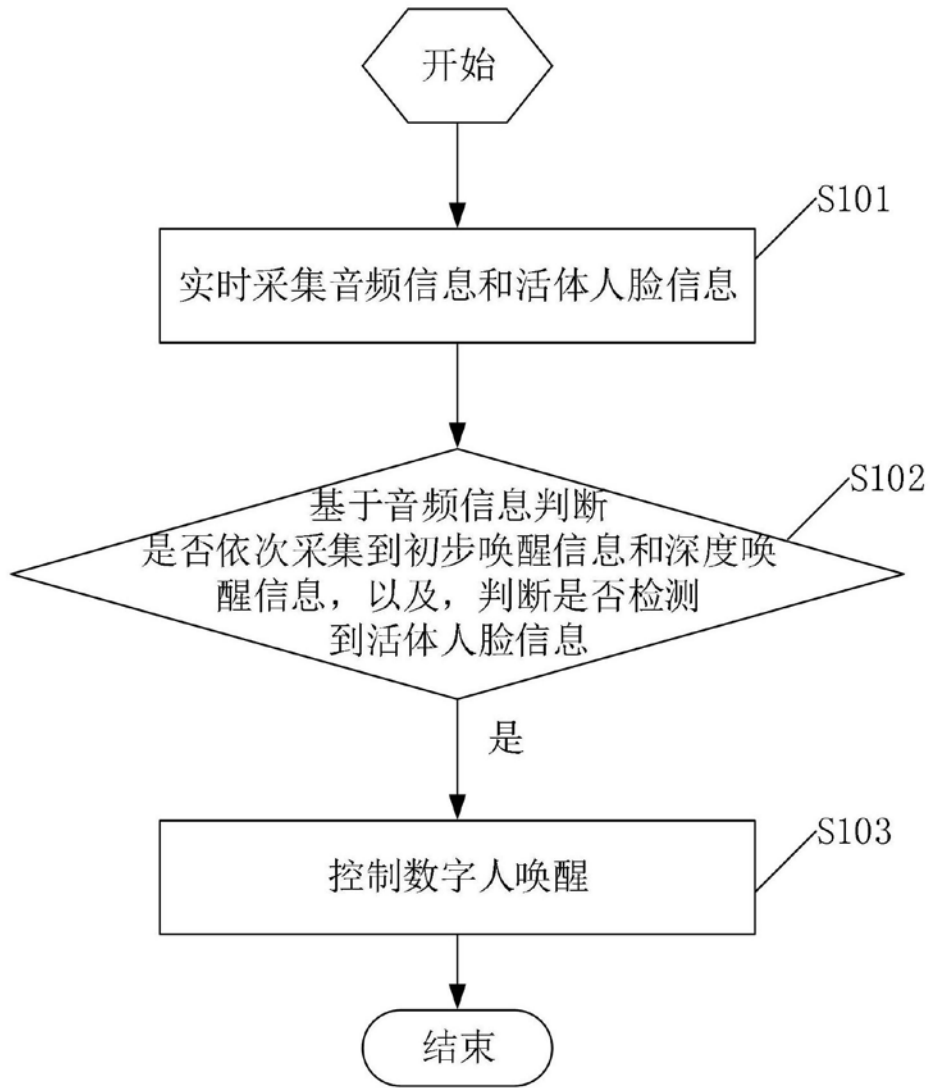


图2



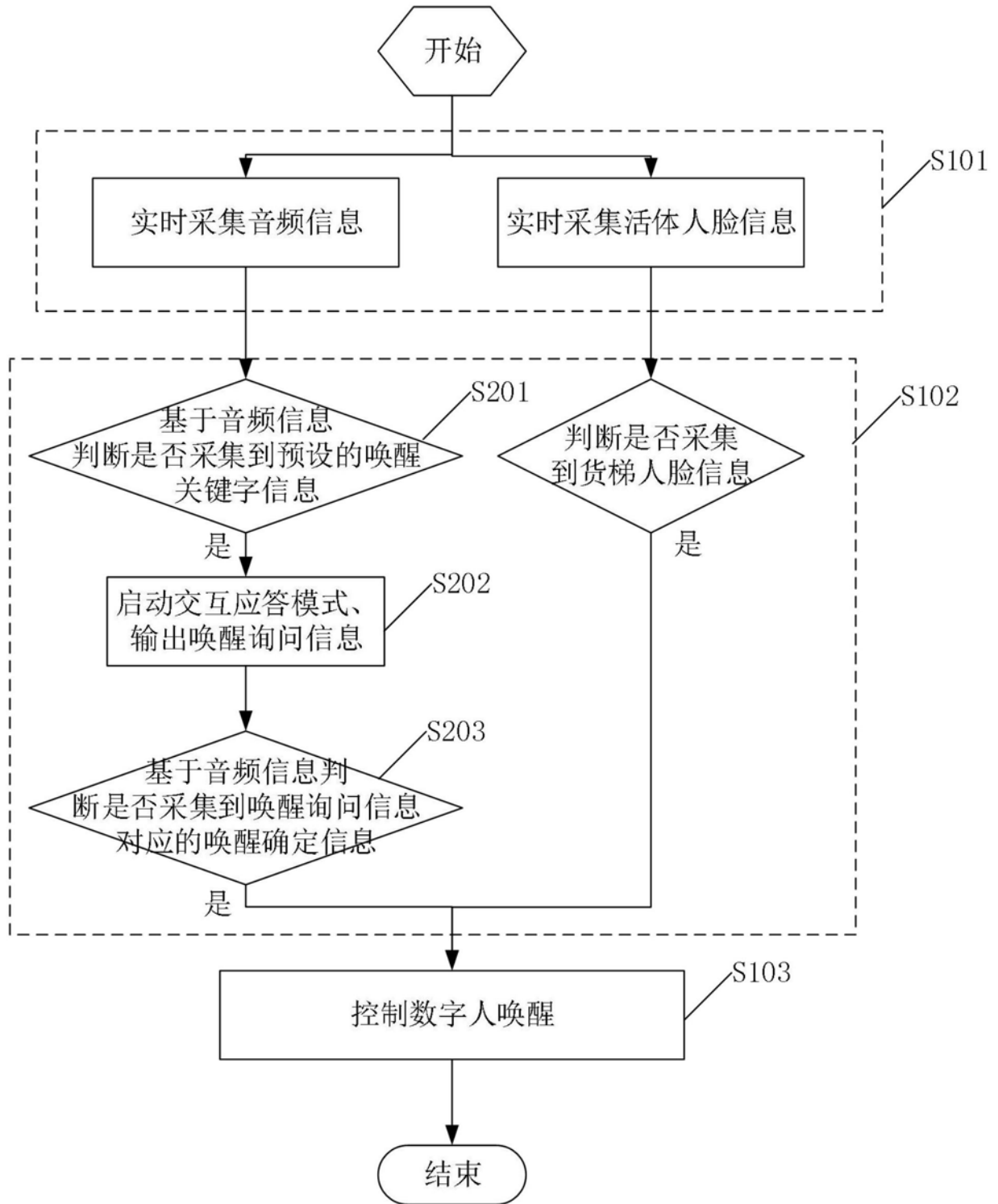


图3

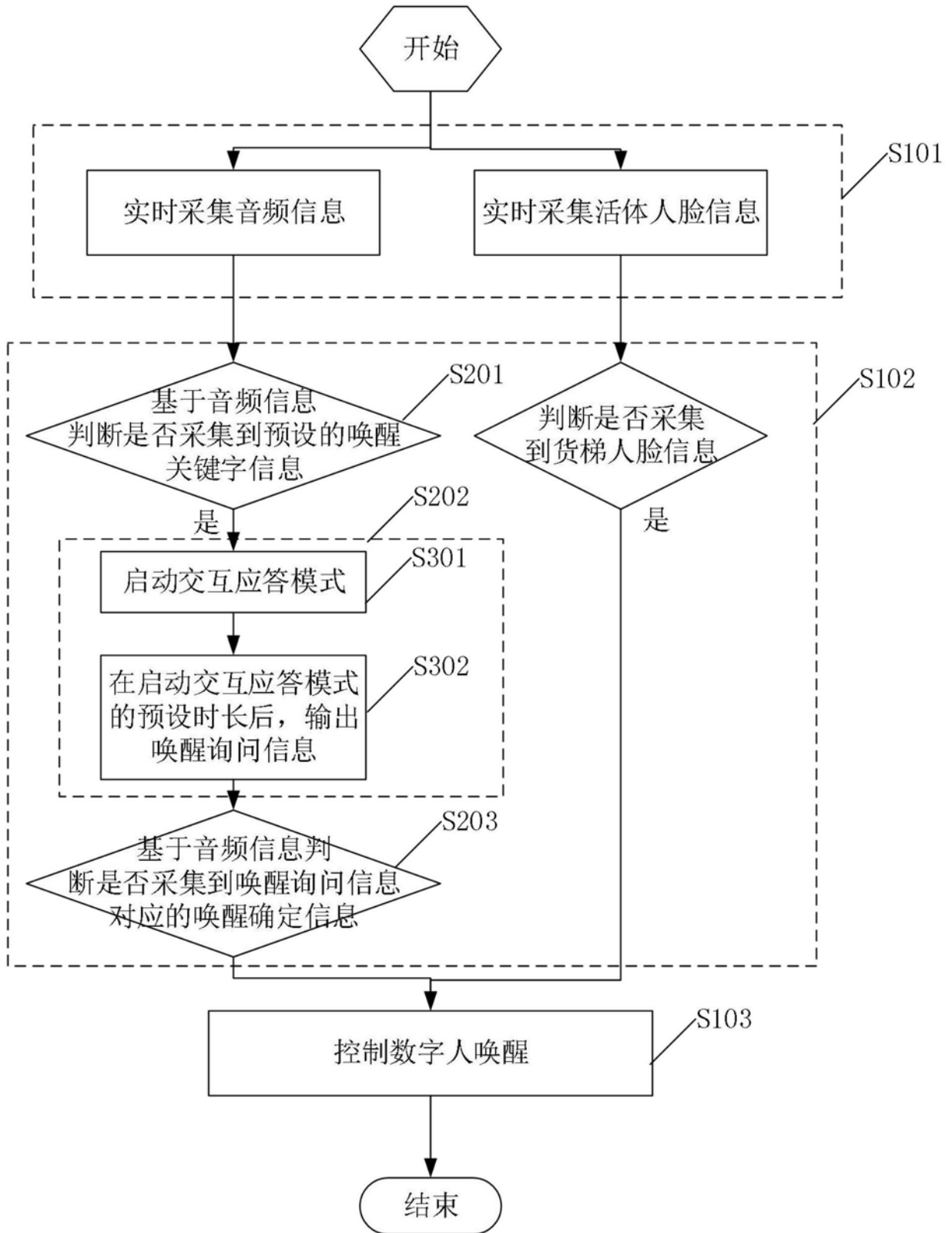


图4

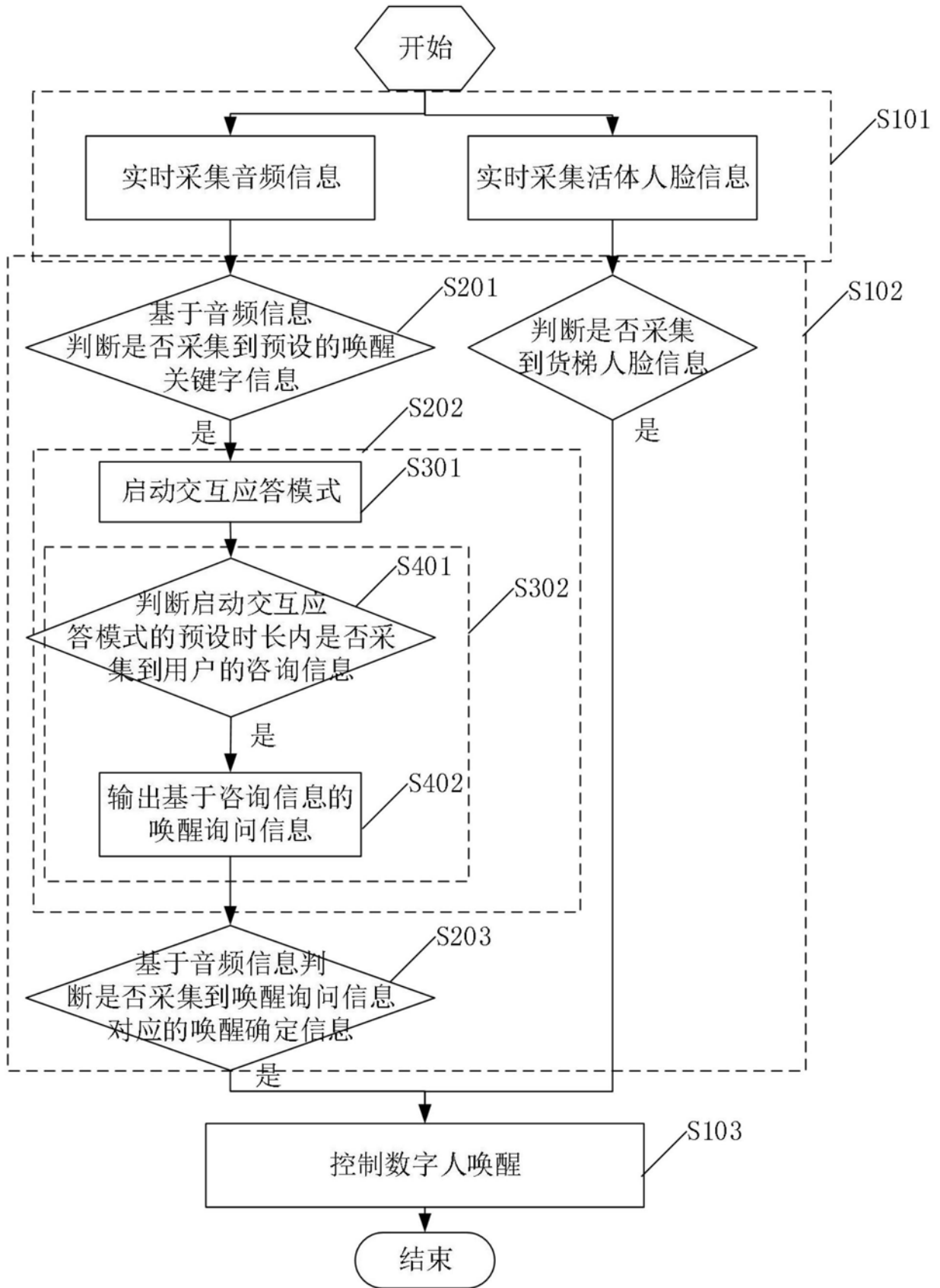


图5

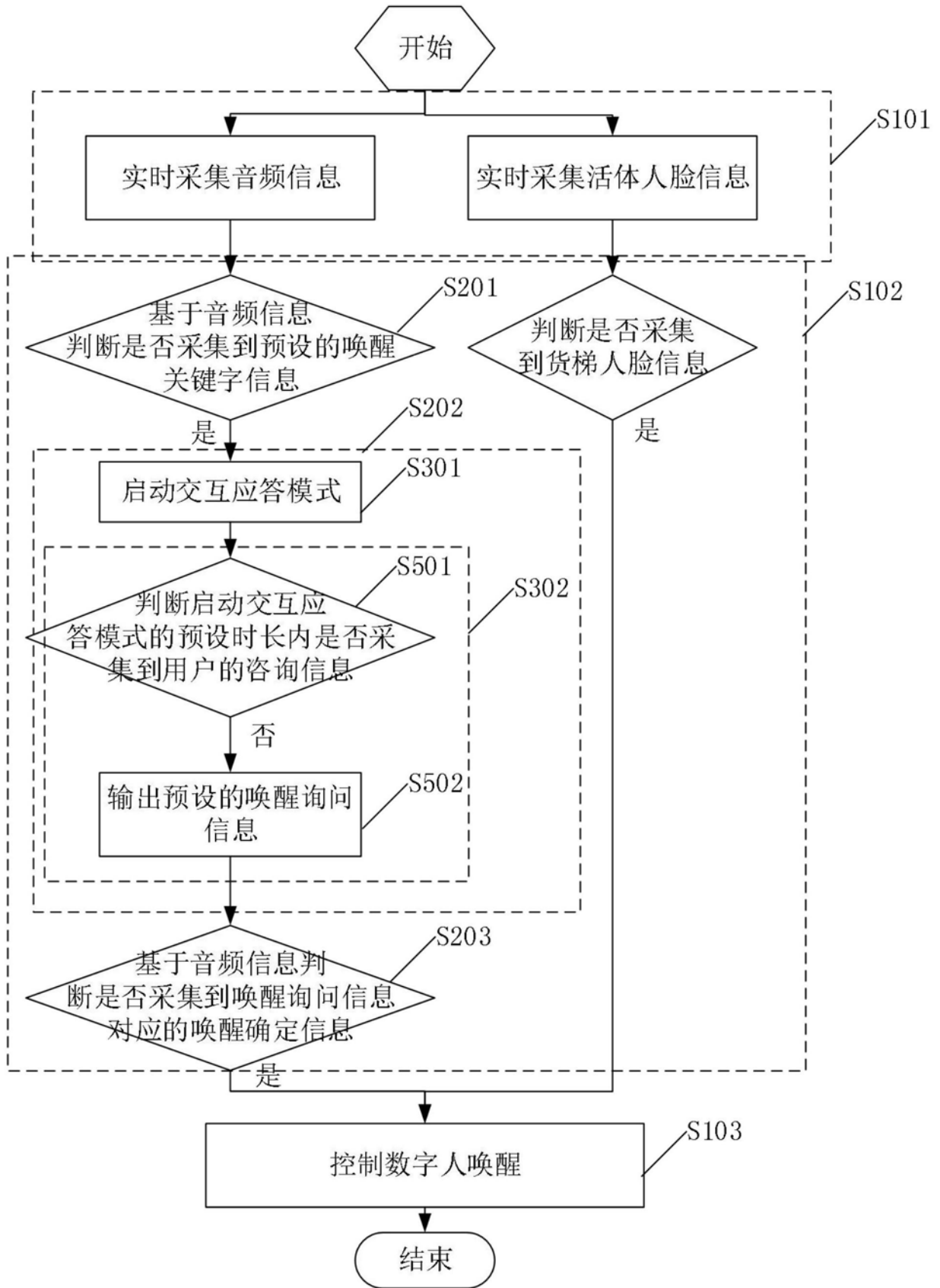


图6

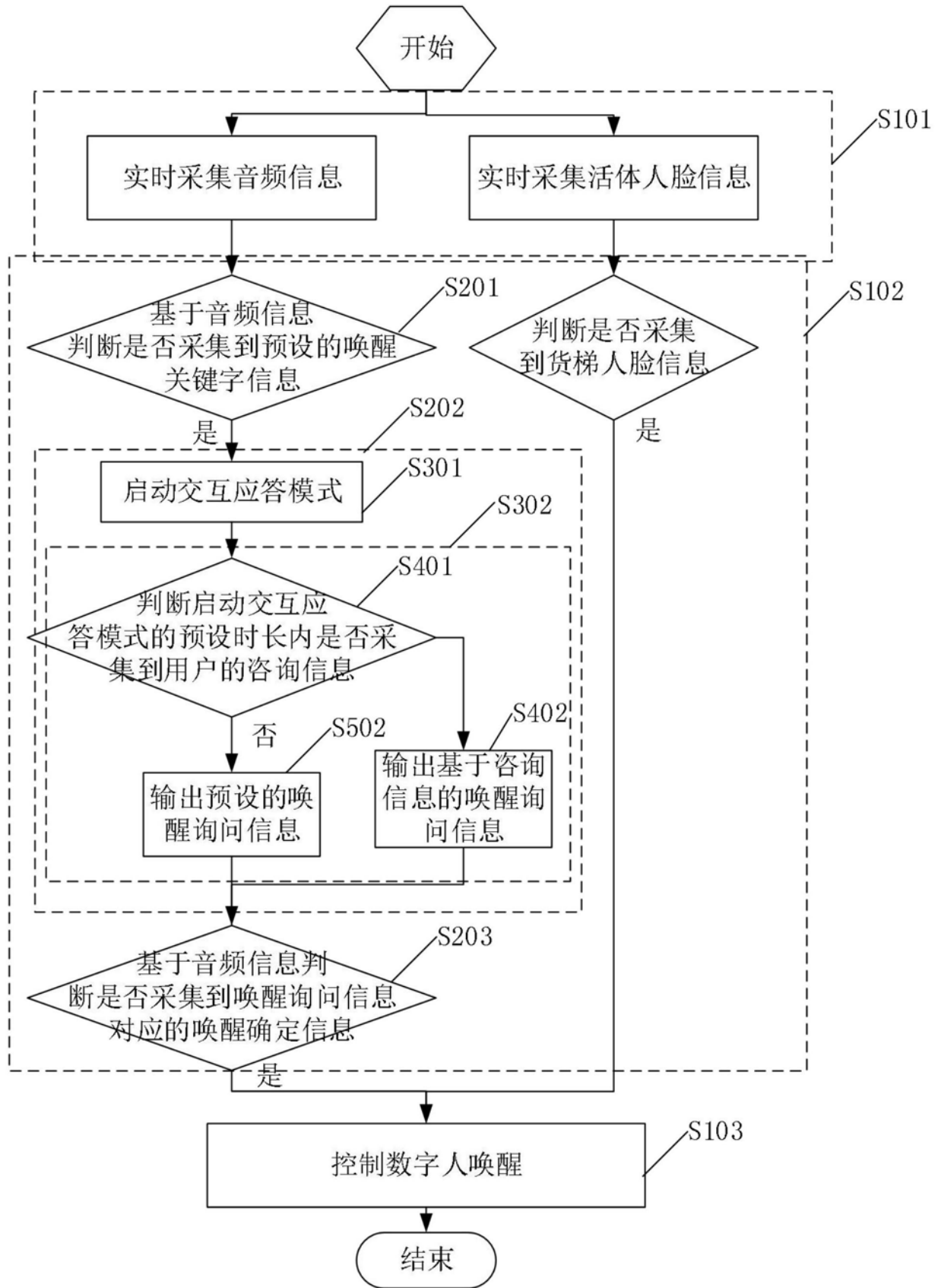


图7

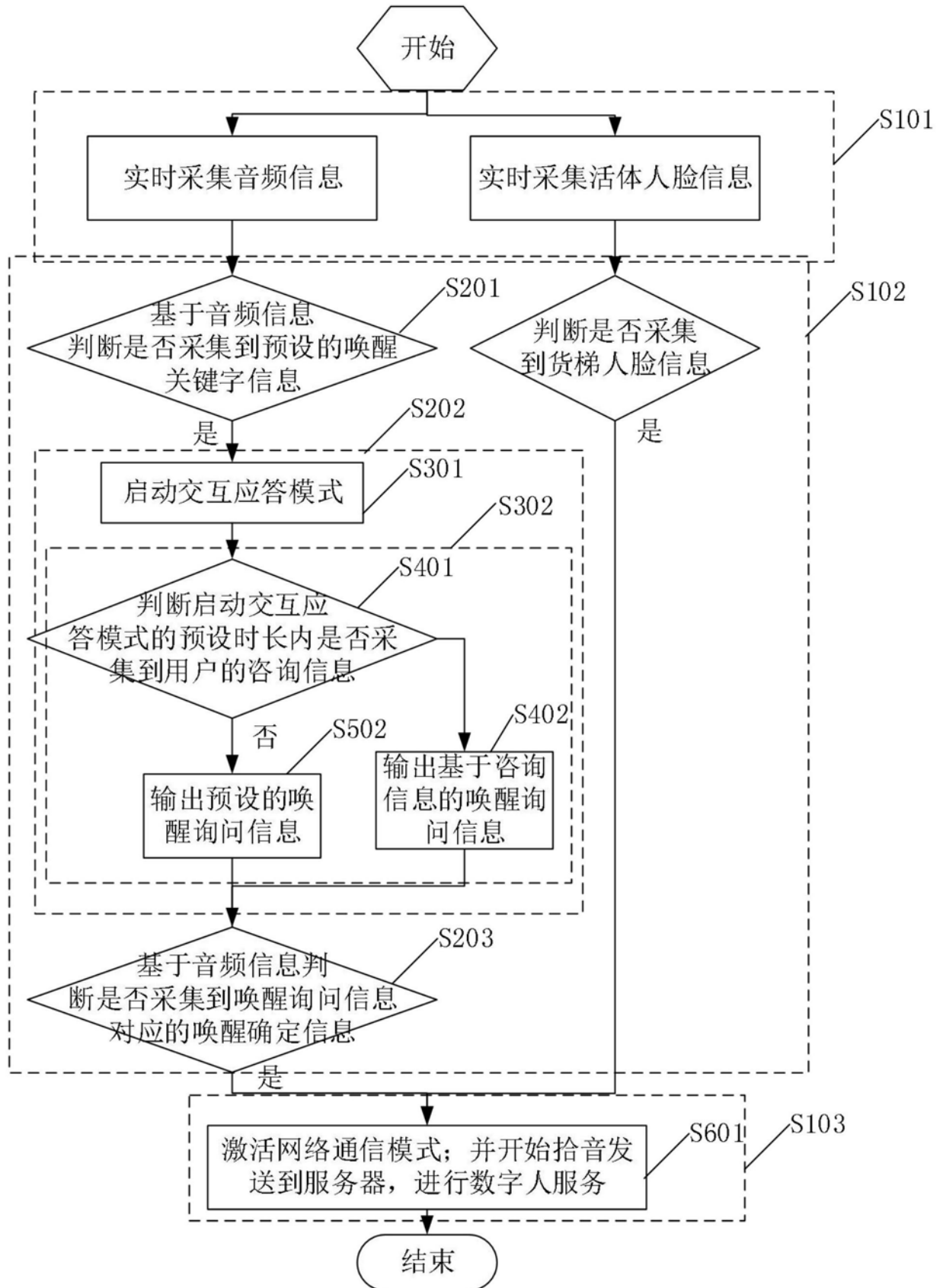


图8

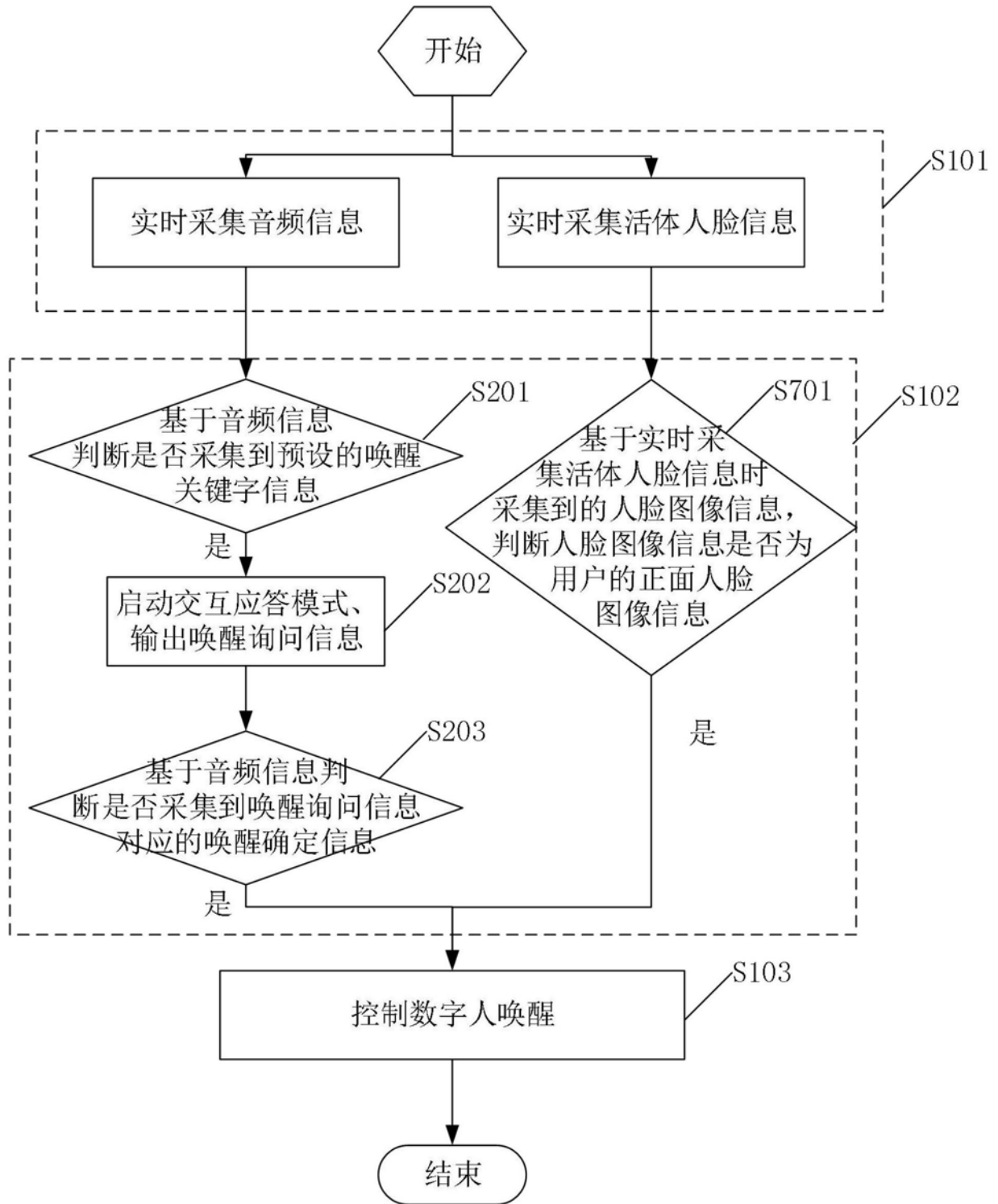


图9

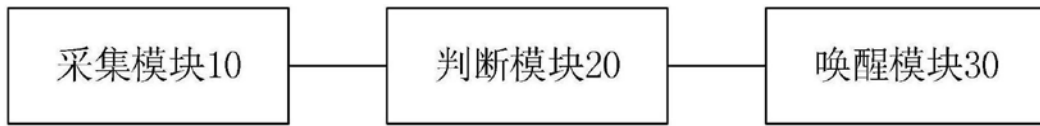


图10

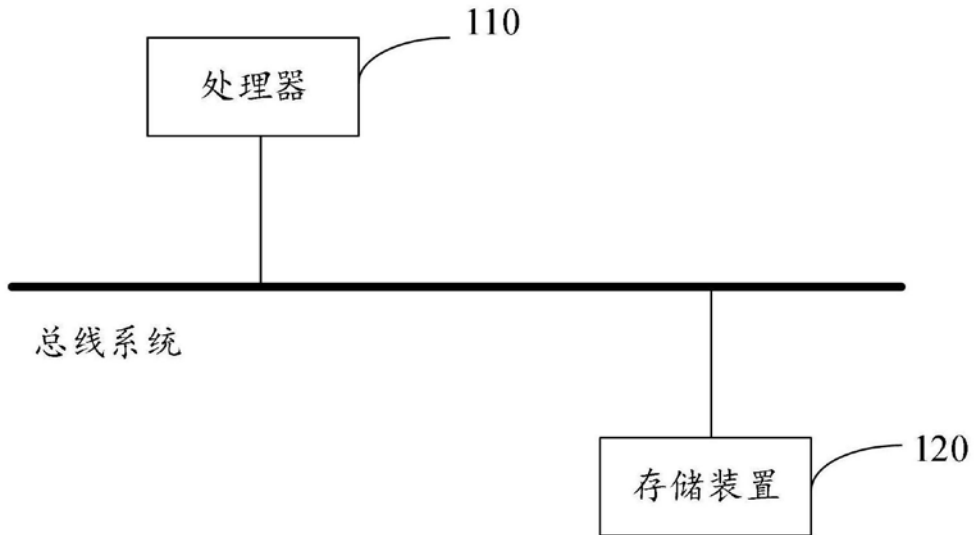


图11