(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 107342088 B
(45)授权公告日 2021.05.18

(21)申请号 201710465049.8

(22)申请日 2017.06.19

(65)同一申请的已公布的文献号
申请公布号 CN 107342088 A

(43)申请公布日 2017.11.10

(73)专利权人 联想(北京)有限公司
地址 100085 北京市海淀区上地信息产业基地创业路6号

(72)发明人 白金才 武亚强 张晓平 李辉 吕琬军

(74)专利代理机构 北京鼎佳达知识产权代理事务所(普通合伙) 11348
代理人 王伟锋 刘铁生

(51)Int.Cl.
G10L 15/26(2006.01)
H04M 1/72454(2021.01)
H04M 1/7243(2021.01)
H04M 1/72439(2021.01)

(56)对比文件
CN 106375548 A,2017.02.01
CN 105677711 A,2016.06.15
CN 105100511 A,2015.11.25

审查员 周家行

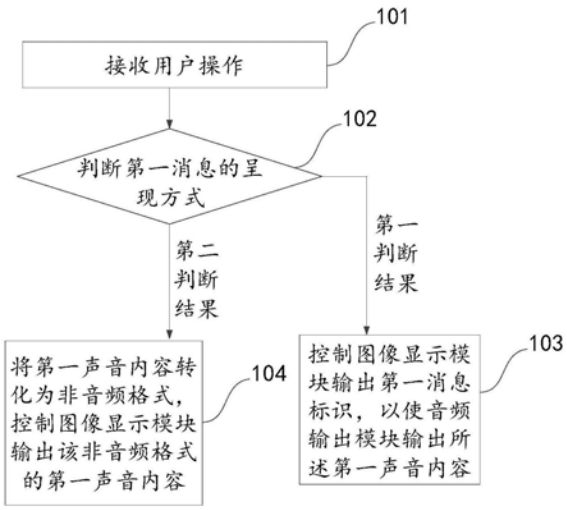
权利要求书2页 说明书13页 附图5页

(54)发明名称

一种声音信息的转换方法、装置及设备

(57)摘要

本发明公开了一种声音信息的转换方法、装置及设备，涉及语音识别转换领域，主要目的在于将用户接收的语音消息有条件地自动转换为文字消息。本发明主要的技术方案为：接收用户操作，所述用户操作用于指示所述第一电子设备呈现第一消息，所述第一消息的内容中包含音频格式的第一声音内容；判断所述第一消息的呈现方式，获得判断结果；若所述判断结果为第一判断结果，则控制图像显示模块输出第一消息标识，以使音频输出模块输出所述第一声音内容；若所述判断结果为第二判断结果，则将所述第一声音内容转化为非音频格式，控制图像显示模块输出所述非音频格式的所述第一声音内容。本发明主要用于对声音信息的转换。



CN 107342088 B

1.一种声音信息的转换方法,应用于第一电子设备,其特征在于,所述方法包括:

接收用户操作,所述用户操作用于指示所述第一电子设备呈现第一消息,所述第一消息的内容中包含音频格式的第一声音内容;

判断所述第一消息的呈现方式,获得判断结果;

若所述判断结果为第一判断结果,则控制图像显示模块输出第一消息标识,以使音频输出模块输出所述第一声音内容;

若所述判断结果为第二判断结果,则将所述第一声音内容转化为非音频格式,控制图像显示模块输出所述非音频格式的所述第一声音内容;

所述判断所述第一消息的呈现方式,包括:

确定语音转换模式以及用于判断所述语音转换模式是否启动对应的转换条件信息;

根据所述第一消息以及所述用户的操作提取对应的转换条件信息;

通过所述转换条件信息与语音转换模式中的启动条件进行匹配,判断语音转换模式是否启动。

2.根据权利要求1所述的方法,其特征在于,判断所述第一消息的呈现方式包括:

判断用户当前是否便于处理音频格式的文件消息。

3.根据权利要求2所述的方法,其特征在于,所述判断用户当前是否便于处理音频格式的文件消息包括:

获取日程信息,根据所述日程信息判断当前时间是否处于特定时间段内。

4.根据权利要求1所述的方法,其特征在于,将所述第一声音内容转化为非音频格式包括:

当所述第一声音内容为语音信息时,识别所述语音信息;

利用语音数据库将所述语音信息转换为预置的文字信息。

5.根据权利要求1所述的方法,其特征在于,将所述第一声音内容转化为非音频格式包括:

当所述第一声音内容为拟声信息时,确定所述拟声信息的发声源;

将所述拟声信息转换为文字呈现的发生源。

6.根据权利要求1所述的方法,其特征在于,将所述第一声音内容转化为非音频格式包括:

当所述第一声音内容为音乐时,确定所述音乐对应的标识信息,所述标识信息包括曲目名称、作者、音乐风格的至少之一;

将所述标识信息转换为文字呈现的标识信息。

7.根据权利要求1-6中任一项所述的方法,其特征在于,所述控制图像显示模块输出所述非音频格式的所述第一声音内容包括:

显示所述第一声音内容对应的文字信息,以及标记在所述文字信息上的转换标签,所述转换标签用于根据用户操作控制音频输出模块输出所述第一声音内容。

8.一种声音信息的转换装置,应用于第一电子设备,其特征在于,所述装置包括:处理器,显示屏以及声音输出模块;

其中,所述处理器,用于接收用户操作,所述用户操作用于指示所述第一电子设备呈现第一消息,所述第一消息的内容中包含音频格式的第一声音内容;

判断所述第一消息的呈现方式，获得判断结果；

当所述处理器的判断结果为第一判断结果时，控制图像显示模块输出第一消息标识，以使所述声音输出模块输出所述第一声音内容；

当所述处理器的判断结果为第二判断结果时，将所述第一声音内容转化为非音频格式，控制所述显示屏输出所述非音频格式的所述第一声音内容；

所述判断所述第一消息的呈现方式，包括：

确定语音转换模式以及用于判断所述语音转换模式是否启动对应的转换条件信息；

根据所述第一消息以及所述用户的操作提取对应的转换条件信息；

通过所述转换条件信息与语音转换模式中的启动条件进行匹配，判断语音转换模式是否启动。

9.根据权利要求8所述的装置，其特征在于，所述处理器还用于，判断用户当前是否便于处理音频格式的文件消息。

10.根据权利要求9所述的装置，其特征在于，所述处理器在判断用户当前是否便于处理音频格式的文件消息时，还包括：

获取日程信息，根据所述日程信息判断当前时间是否处于特定时间段内。

11.根据权利要求8所述的装置，其特征在于，所述处理器将所述第一声音内容转化为非音频格式包括：

当所述第一声音内容为语音信息时，识别所述语音信息；

利用语音数据库将所述语音信息转换为预置的文字信息。

12.根据权利要求8所述的装置，其特征在于，所述处理器将所述第一声音内容转化为非音频格式包括：

当所述第一声音内容为拟声信息时，确定所述拟声信息的发声源；

将所述拟声信息转换为文字呈现的发生源。

13.根据权利要求8所述的装置，其特征在于，所述处理器将所述第一声音内容转化为非音频格式包括：

当所述第一声音内容为音乐时，确定所述音乐对应的标识信息，所述标识信息包括曲目名称、作者、音乐风格的至少之一；

将所述标识信息转换为文字呈现的标识信息。

14.根据权利要求8-13中任一项所述的装置，其特征在于，所述处理器控制显示屏输出所述非音频格式的所述第一声音内容包括：

向所述显示屏输出所述第一声音内容对应的文字信息，以及标记在所述文字信息上的转换标签，所述转换标签用于根据用户操作控制音频输出模块输出所述第一声音内容，以便所述显示屏向用户展示对应的文字信息。

15.一种声音信息的转换设备，其特征在于，所述设备中设置有如权利要求8-14中任一项所述的声音信息的转换装置。

一种声音信息的转换方法、装置及设备

技术领域

[0001] 本发明涉及语音识别转换领域，尤其涉及一种声音信息的转换方法、装置及设备。

背景技术

[0002] 随着电子技术以及网络通讯技术的进步与普及，人们沟通、交流的方式也越来越丰富多样，从语音通话、短信到即时通信应用的短消息，再到视频通话、语音消息等，种种新兴的通信方式不断更新迭代。并且，对于不同的通信方式也有着各自适用的应用场景，比如，对于需要及时交流的事情，往往需要通过电话或实时的视频进行通信，而对于普通的问候、聊天等事情，人们则习惯使用即时端消息或语音消息。

[0003] 在语音消息的场景中，对于消息发送者可以省去复杂、繁琐的信息录入步骤，而对于消息的接收者则与文字消息的查看没有明显区别。但是在有些场合下接收者不方便收听发送者的语音信息，从而给接收者查询消息内容带来不便。在不方便收听的场合下，接收者可以选择放弃查询该语音消息，但若想获取语音信息，则需要通过其他操作转变消息的查询方式，比如，通过接入耳机收听消息，或者将该语音消息转换成文字消息后再进行查询。而对于执行将语音消息转换成文字消息的操作，则需要接收者逐一选中待转换的语音消息后，再执行相应的转化操作，其操作繁琐，严重影响了用户获取消息内容的使用体验。

发明内容

[0004] 有鉴于此，本发明提供一种声音信息的转换方法、装置及设备，主要目的在于将用户接收的语音消息有条件地自动转换为文字消息。

[0005] 为达到上述目的，本发明主要提供如下技术方案：

[0006] 一方面，本发明提供了一种声音信息的转换方法，应用于第一电子设备，该方法包括：

[0007] 接收用户操作，所述用户操作用于指示所述第一电子设备呈现第一消息，所述第一消息的内容中包含音频格式的第一声音内容；

[0008] 判断所述第一消息的呈现方式，获得判断结果；

[0009] 若所述判断结果为第一判断结果，则音频输出模块输出所述第一声音内容；

[0010] 若所述判断结果为第二判断结果，则将所述第一声音内容转化为非音频格式，控制图像显示模块输出所述非音频格式的所述第一声音内容。

[0011] 优选的，判断所述第一消息的呈现方式包括：

[0012] 判断用户当前是否便于处理音频格式的文件消息。

[0013] 优选的，所述判断用户当前是否便于处理音频格式的文件消息包括：

[0014] 获取日程信息，根据所述日程信息判断当前时间是否处于特定时间段内。

[0015] 优选的，将所述第一声音内容转化为非音频格式包括：

[0016] 当所述第一声音内容为语音信息时，识别所述语音信息；

[0017] 利用语音数据库将所述语音信息转换为预置的文字信息。

[0018] 优选的，将所述第一声音内容转化为非音频格式包括：

[0019] 当所述第一声音内容为拟声信息时，确定所述拟声信息的发声源；

[0020] 将所述拟声信息转换为文字呈现的发生源。

[0021] 优选的，将所述第一声音内容转化为非音频格式包括：

[0022] 当所述第一声音内容为音乐时，确定所述音乐对应的标识信息，所述标识信息包括曲目名称、作者、音乐风格的至少之一；

[0023] 将所述标识信息转换为文字呈现的标识信息。

[0024] 优选的，所述控制图像显示模块输出所述非音频格式的所述第一声音内容包括：

[0025] 显示所述第一声音内容对应的文字信息，以及标记在所述文字信息上的转换标签，所述转换标签用于根据用户操作控制音频输出模块输出所述第一声音内容。

[0026] 另一方面，本发明还提供了一种声音信息的转换装置，应用于第一电子设备，该装置包括：处理器，显示屏以及声音输出模块；

[0027] 其中，所述处理器，用于接收用户操作，所述用户操作用于指示所述第一电子设备呈现第一消息，所述第一消息的内容中包含音频格式的第一声音内容；

[0028] 判断所述第一消息的呈现方式，获得判断结果；

[0029] 当所述处理器的判断结果为第一判断结果时，控制所述声音输出模块输出所述第一声音内容；

[0030] 当所述处理器的判断结果为第二判断结果时，将所述第一声音内容转化为非音频格式，控制所述显示屏输出所述非音频格式的所述第一声音内容。

[0031] 优选的，所述处理器还用于，判断用户当前是否便于处理音频格式的文件消息。

[0032] 优选的，所述处理器在判断用户当前是否便于处理音频格式的文件消息时，还包括：

[0033] 获取日程信息，根据所述日程信息判断当前时间是否处于特定时间段内。

[0034] 优选的，所述处理器将所述第一声音内容转化为非音频格式包括：

[0035] 当所述第一声音内容为语音信息时，识别所述语音信息；

[0036] 利用语音数据库将所述语音信息转换为预置的文字信息。

[0037] 优选的，所述处理器将所述第一声音内容转化为非音频格式包括：

[0038] 当所述第一声音内容为拟声信息时，确定所述拟声信息的发声源；

[0039] 将所述拟声信息转换为文字呈现的发生源。

[0040] 优选的，所述处理器将所述第一声音内容转化为非音频格式包括：

[0041] 当所述第一声音内容为音乐时，确定所述音乐对应的标识信息，所述标识信息包括曲目名称、作者、音乐风格的至少之一；

[0042] 将所述标识信息转换为文字呈现的标识信息。

[0043] 优选的，所述处理器控制显示屏输出所述非音频格式的所述第一声音内容包括：

[0044] 向所述显示屏输出所述第一声音内容对应的文字信息，以及标记在所述文字信息上的转换标签，所述转换标签用于根据用户操作控制音频输出模块输出所述第一声音内容，以便所述显示屏向用户展示对应的文字信息。

[0045] 另一方面，本发明还提供了一种声音信息的转换设备，该设备中设置有上述任一的声音信息的转换装置。

[0046] 依据上述本发明所提出的一种声音信息的转换方法、装置及设备，能够将用户接收到的含有音频信息的消息根据预置的转换模式进行判断，将能够直接输出声音的消息通过音频输出模块输出声音内容，将不能直接输出声音的消息则通过显示屏以文字内容展示给用户，以便用户在任意场合下都能够方便、快捷的查询到该消息的内容，避免用户由于不方便收听语音信息而错过重要的信息。

附图说明

[0047] 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

[0048] 图1示出了本发明实施例提出的一种声音信息的转换方法的流程图；

[0049] 图2示出了本发明实施例提出的另一种声音信息的转换方法的流程图；

[0050] 图3示出了本发明实施例提出的另一种声音信息的转换方法的流程图；

[0051] 图4示出了本发明实施例提出的另一种声音信息的转换方法的流程图；

[0052] 图5示出了本发明实施例提出的语音转换模式设置界面示意图；

[0053] 图6示出了本发明实施例提出的语音信息转换文字信息的显示界面示意图；

[0054] 图7示出了本发明实施例提出的一种声音信息的转换装置的组成框图。

具体实施方式

[0055] 下面将参照附图更详细地描述本发明的示例性实施例。虽然附图中显示了本发明的示例性实施例，然而应当理解，可以以各种形式实现本发明而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本发明，并且能够将本发明的范围完整的传达给本领域的技术人员。

[0056] 本发明实施例提供了一种声音信息的转换方法，应用于第一电子设备，如图1所示，该方法是对音频信息转换为文字信息在应用上的具体优化及改进，使得用户能够在需要的场合下实现将音频信息自动转换为文字信息再进行查询，以便及时、快速获取信息内容。本方法实施例的具体步骤包括：

[0057] 101、接收用户操作。

[0058] 本发明实施例中，用户为第一电子设备的使用用户，当第一电子设备接收到包含音频格式的第一声音内容的第一消息时，第一电子设备将根据用户操作确定是否呈现该第一消息，其中第一电子设备呈现第一消息的方式可以是通过显示屏显示该第一消息，也可以是具体显示第一消息的内容，即通过文字形式展现第一消息中的第一声音内容，还可以是通过音频输出模块以声音形式直接输出该第一声音内容。

[0059] 在实际应用中，上述的包含音频格式的第一声音内容的第一消息可以是语音消息，也可以是视频消息，而第一消息中的第一声音内容则不限定为人的语音，或者是自然界的声音，如动物叫声、流水声等，亦或者是音乐等。其中，对于在即时通讯软件的应用场景下，比如微信、Skype等，第一消息是指用户接收或发送的消息，而用户操作是指打开一个会话，进入会话内容的呈现界面，或者是如朋友圈、微博、Twitte等应用中，直接呈现内容流的应用界面中，即在所打开的会话页面中展现第一消息，该第一消息可以是声音消息，也可以

是文字或图像消息。

[0060] 102、判断第一消息的呈现方式，获得判断结果。

[0061] 由于第一电子设备对第一消息具有多种的呈现方式，因此，在接收到用户操作后，第一电子设备将根据预置的判断规则确定一种呈现方式。该判断规则一般是由用户预先设置的，针对在具体应用场景下使用哪一种呈现方式向用户展示第一消息。

[0062] 在本发明实施例中，所得到的判断结果包括两种，当为第一判断结果时，执行步骤103，即控制图像显示模块输出第一消息标识，以使音频输出模块输出第一声音内容；当为第二判断结果时，执行步骤104，即将第一声音内容转化为非音频格式，控制图像显示模块输出该非音频格式的第一声音内容。

[0063] 103、控制图像显示模块输出第一消息标识，以使音频输出模块输出第一声音内容。

[0064] 其中，与本步骤对应的第一判断结果是用户在当前应用场景下能够通过声音的方式读取该第一消息中的声音内容，具体在应用的会话中，是通过图像显示模块在会话界面中显示第一消息标识。针对该第一消息标识，第一电子设备可以通过用户点击操作来播放第一消息中的声音内容，也可以自动对第一消息中的声音内容进行播放，输出声音。此时，第一电子设备将通过音频输出模块将第一消息中的第一声音内容以声音的形式播放给当前用户，对于播放的具体形式则取决于该第一电子设备所具有的硬件设备，比如，当第一电子设备中自带扬声器时，则通过该扬声器进行播放，当第一电子设备中没有扬声器，而带有耳机接口或者是无线耳机连接模块时，则将音频信息通过用户的耳机进行播放；当第一电子设备同时具有上述的音频输出方式时，在输出音频时则可以根据用户自定义的优先级确定一种音频输出方式。

[0065] 104、将第一声音内容转化为非音频格式，控制图像显示模块输出该非音频格式的第一声音内容。

[0066] 与本步骤对应的第二判断结果则是用户在当前的应用场景下需要通过第一电子设备的显示设备展示第一消息，一般是通过屏幕来显示该第一消息，而对于该消息中的第一声音内容则需要通过格式转换将其以非音频格式的内容显示在屏幕上，其中，非音频格式的内容不限于文字、图像等信息。

[0067] 基于上述的实施例，在实际应用场景中较多的是应用于将第一电子设备接收到的语音信息转换为文字信息的情况，即上述实施例中的第一消息为语音消息，对此，通过下面的处理方式来确定是否将语音信息转换为文字信息，或者是向用户播放该语音信息，具体如图2所示，包括：

[0068] 201、在接收语音信息时，判断该语音消息的呈现方式。

[0069] 其中，判断的具体规则可以抽象为不同的语音转换模式，通过语音转换模式的启动判断来确定第一消息的具体呈现方式，该语音转换模式可以是由用户自定义设置的，用户可根据实际需要，在终端设备上同时设置多个不同的语音转换模式，用以在不同的应用场景中实现对所接收的语音信息转换为文字信息。

[0070] 本步骤是在接收到语音信息的同时，查询所设置的语音转换模式是否启动，启动则执行步骤202，否则，执行步骤203。其中，在判断语音转换模式是否启动时，主要是根据在语音转换模式设置时所设置的启动条件，该启动条件与所接收的语音信息中的转换条件信

息相对应，通过获取语音信息中的转换条件信息与启动条件进行匹配，符合启动条件则启动对应的语音转换模式，反之，则不启动对应的语音转换模式。具体在本发明实施例中，所设置的语音转换模式的启动条件一般为其对应的转换条件信息的一个浮动区间或变换范围，或者是特定的多个转换条件信息的集合，通过获取所接收声音信息的转换条件信息，与所对应的语音转换模式的启动条件进行匹配，以此来判断该语音转换模式是否启动。

[0071] 对于该转换条件信息的具体内容由语音转换模式确定，不同的语音转换模式将对应于不同的转换条件信息，而该转换条件信息一般为语音信息的相关属性信息，例如，可以是语音信息的接收时间信息，根据接收的时间点来确定是否将该语音信息转换为文字信息；或者是语音信息的来源信息，也就是该语音信息的发送用户，根据发送用户来确定是否将该语音信息转换为文字信息。可见，该转换条件信息应具有能够区别于其他语音信息的特征。

[0072] 对于判断语音转换模式是否启动的过程，由于所设置的语音转换模式可以为一个，也可以是多个，并且不同的语音转换模式对应于不同的转换条件信息，因此，在设置有多个语音转换模式时，一般会根据语音信息获取到多个转换条件信息，此时，在判断时，将使用所获取的转换条件信息逐一的与其所对应的语音转换模式进行匹配，一旦存在有匹配成功的转换条件信息时，将不再匹配后续未匹配的转换条件信息，直接执行步骤202。而当所有的转换条件信息都未能匹配成功时，则无需对该语音信息进行转换，可直接向用户提示当前存在有待查询的语音信息，并在设备中显示该语音信息以供用户查询。例如，用户设置有两个语音转换模式，一个为时间段模式，可应用于如会议、出差等场景下，另一个是联系人群组模式，主要应用于用户自定义设置部分的重要联系，将其发送的语音主动转化为文字的场景下，针对这两种不同的模式，时间段模式所对应的转换条件信息为接收到语音信息时的时间信息，而联系人群组模式所对应的转换条件信息则为语音信息的发送者信息，为此，当接收到一条语音信息时，将获取该语音信息对应的时间信息和发送者信息，再逐一判断是否符合启动对应语音转换模式，即判断该时间信息是否在时间段模式所设置的时间段内，或者判断发送者信息是否为联系人群组模式中所设置的联系人。

[0073] 202、若判断结果为第二判断结果，则将语音信息转换为文字信息，并显示该文字信息。

[0074] 其中，第二判断结果为步骤201中的判断语音转换模式被启动，此时，将自动将所接收的语音信息转换为文字信息，并控制图像显示模块显示该文字信息。其中，在执行转换操作时，优先在本地进行转换操作，一般的，将语音转换为文字需要对语音进行识别，而对于语音则存在许多不同的种类，如不同国家的语言(中文、英语、法语等)，或者不同地区的方言(普通话、广东话、四川话等等)，不同的语言其发音也不同，而要加以识别则需要利用相应的语音数据库来实现，而对于在设备中安装的应用程序而言其受限于程序自身的大小，以及设备的存储、处理资源等因素，只能设置较为常用的语音数据库在设备本地。因此，对于语音的识别有时会出现在本地无法正确地加以识别，甚至是无法识别转换的问题。对此，在本地无法进行识别时，本发明实施例还可以将该语音信息上传至服务器并进行识别转换，将最优的转换结果反馈给该设备。

[0075] 而对于在本地进行转换操作，且本地中还设置有多个语音数据库的情况，本发明实施例优先根据设备操作系统的语言选择语音数据库用以转换操作的语音数据库，例如，

当智能手机的操作系统为中文时，则优先选择中文语音数据库，而当为英文时，则优先选择英语语音数据库执行转换操作。进一步的，对于具有多个方言语音数据库的，则可以对不同的语音数据库设置优先级，根据优先级确定使用的语音数据库，例如，对于中文的操作系统，普通话的语音数据库为高优先级，而其他方言的语音数据库为低优先级，当然，该优先级的设置可以由用户自定义设置，比如该用户为四川人，其语音信息多为四川话，则可以由该用户将四川话的语音数据库设置为高优先级。此外，还可以通过获取设备的位置信息来自动调整语音数据的优先级设置。

[0076] 最终，由图像显示模块将转换后的文字信息显示到设备的显示屏中，以提示用户及时查看，而在显示该文字信息时，本发明实施例则不限定在显示文字信息的同时是否一同显示原始的语音信息。但为了将转换得到的文字信息与其他普通的文字信息加以区分，可以对该文字信息加以区别显示，对此本发明实施例不限定该区别显示的具体方式，既可以是在所显示的文字信息上标注标签，也可以是在一个新的界面下单独显示转化后的文字信息。

[0077] 203、若判断结果为第一判断结果，则显示语音信息的标识。

[0078] 其中，第一判断结果为步骤201中的判断语音转换模式未被启动，此时，不对所接收的语音信息执行转换操作，而是直接显示该语音信息的标识，可以根据用户的进一步查询操作利用设备中的音频输出模块输出该语音信息，也可以在显示该标识的同时，直接通过音频输出模块输出该语音信息的声音内容。

[0079] 结合上述的实现方式可以看出，本发明实施例所采用的声音信息的转换方法，能够将用户接收到的含有音频信息的消息根据预置的转换模式进行判断，将能够直接输出声音的消息通过音频输出模块输出声音内容，将不能直接输出声音的消息则通过显示屏以文字内容展示给用户，以便用户在任意场合下都能够方便、快捷的查询到该消息的内容，避免用户由于不方便收听语音信息而错过重要的信息。特别是在语音信息的应用场景下，通过设置语音信息的语音转换模式来实现在不同的特定场景下自动将语音信息转化为文字信息，并将文字信息进行显示，以方便用户能够及时获取信息的内容。而对于转换的具体过程，通过对语音数据的优化排列也可以更加快速和准确地将语音信息转换为文字信息，从而提高用户获取信息内容的准确性，提升用户的使用体验。

[0080] 为了更加详细地说明本发明提出的一种声音信息的转换方法，特别是针对图1所述实施例的实现方式将通过下面的实施例来具体说明在特殊应用场景下使用该方法所能达到的效果，具体步骤如图3所示，包括：

[0081] 301、接收用户操作。

[0082] 本步骤具体的执行过程与步骤101相同，因此，该内容在此不再赘述，可参考步骤101的描述。

[0083] 302、获取用户设置的日程信息。

[0084] 根据用户的操作，第一电子设备要确定出第一消息的播放方式，即通过屏幕展示或者是通过声音向用户展示。对此，第一电子设备需要判断用户当前是否便于处理音频格式的文件消息，其中，是否便于处理的判断需要根据用户预先设置的具体条件进行判断，在本发明实施例中，用户通过在日程表中设置特定时间段的方式确定用户当前是否便于处理音频格式的文件消息。

[0085] 具体的，用户在第一电子设备的日程表中选择一段时间段，比如，一天中9点至17点，或者是一周的周一至周五等等，通过指定特定的时间段来定义在这些时间段内，用户是否便于处理音频格式的文件消息，一般情况下，这些指定的时间段为用户不方便处理音频格式的文件消息的时间段，因此，在用户设置特定时间段时，同时会给这些时间段标记一个标签，用于确定该时间段内需要将第一消息内容转化非音频格式内容，比如，在定义的时间段9点至11点中加入一个标签为会议，标记在该时间段内用户处于开会中，不便于收听音频消息，在比如在定义的时间段18点至19点中加入一个标签为开车，标记在该时间段内用户处于外出开车中，也不便于收听音频消息。当然，该标签是由用户自定义设置的，用户可随时进行更改或重定义具体内容。

[0086] 303、根据日程信息判断当前时间是否处于特定时间段内。

[0087] 根据步骤302获取的日程信息，第一电子设备将根据接收用户操作的时间判断该时间是否落在日程表中设定的特定时间段内，若没有落在特定的时间段内，则说明此时用户可以接收声音信息，那么此时将通过第一电子设备的扬声器或耳机直接输出第一消息中的第一声音内容。而当前时间落在了特定时间段内时，则说明此时用户不方便接收声音信息，此时需要执行步骤304，将第一声音内容进行转换操作。

[0088] 304、将第一声音内容转化为非音频格式。

[0089] 本步骤中列举了对于不同形式的声音内容将转化为文字信息的几种不同方式，具体为：

[0090] 第一种，当第一声音内容为拟声信息时，即第一消息中的第一声音内容是由自然界的物体发出的声音，比如动物的叫声，雨水、雷生等自然界的拟声声音，此时，设备将针对该声音进行识别，重点匹配出能够发出该声音的物体，即发声源，如具体的动物名称等。若存在多个发声源时，则选择一个最符合的发声源物体，并将该物体的名称以文字的形式进行展示，从而完成将声音信息转化为文字的目的。

[0091] 第二种，当第一声音内容为音乐时，此时，设备将识别该音乐所对应的曲目名称、音乐作者、音乐风格等标识信息，若能够识别出该音乐所对应的标识信息，则将该标识信息以文字信息的形式加以展示，而若无法识别，则告知用户该声音内容为一段无法识别的音乐。

[0092] 第三种，当第一声音内容为语音信息时，此时设备将识别该语音信息，并利用语音数据库将该语音信息转换为预置的文字信息，其中，预置的文字信息可以是根据设备中用户的具体设置输出对应的文字，比如，输出的可以中文或英文，也可以是文字或拼音等等，具体转换的操作可参考上述实施例的步骤202中的内容，而针对语音信息的转换操作也将通过后面的实施例进行具体说明。

[0093] 以上三种情况基本涵盖了所有音频信息的形式，而对于如何区分以上三种声音，则可以由用户设置的顺序逐一对第一声音内容进行匹配，一般的，对声音的识别是通过对声音的音域、音频等指标进行判断，人类的语音、自然声以及音乐都具有不同的声音特征，据此，可以先通过这些特征对第一声音内容进行分类，即判断其所属的上述类别，在确定后再进一步根据上述的方式进行转换，并最终以文字的形式加以呈现。

[0094] 305、控制图像显示模块输出非音频格式的第一声音内容。

[0095] 根据步骤304所转换得到的文字信息，即非音频格式的第一声音内容。设备将该文

字信息通过显示模块输出展示。

[0096] 在输出展示该文字信息的同时，该设备会在该文字信息上标记一个转换标签，用于说明该文字信息是通过转换操作而得到的，并且，该转换标签还具有跳转连接功能，也就是用户通过触发给转换标签可以将原始的第一声音内容通过设备的音频输出模块以声音的形式输出，该功能的主要目的是当用户在看到转换后的文字信息后，如果没有看明白或者存在疑问时，可以通过该转换标识直接调取原始的音频信息进行收听以确认第一消息的内容。

[0097] 基于以上实施例的内容，本发明实施例还具体提供了一种声音信息的转换方法，用于对上述图2所示的方式进行具体说明，即在所接收的第一消息为语音信息时，第一电子设备将根据用户设置的语音转换模式中所确定的判断条件确定该语音信息的具体呈现方式，具体如图4所示，其具体步骤为：

[0098] 401、确定语音转换模式以及用于判断该语音转换模式是否启动所对应的转换条件信息。

[0099] 其中，转换条件信息是与所接收到的语音信息相关联的信息，即对于不同的语音信息，所获取的转换条件信息是不同的，而该转换条件信息则用于判断是否触发将该语音信息转换为文字信息的操作，即判断是否启动所对应的语音转换模式。也就是说，在设置有语音转换模式后，对于所接收到的语音信息将逐一地判断该语音信息所对应的转换条件信息是否能够启动对应的语音转换模式。

[0100] 在本发明实施例中，语音转换模式的设置是由用户自定义设置的。为此，用户可通过应用中提供的设置界面来定义语音转换模式，如图3所示，在该设置界面中，为用户提供有多种的场景设置，例如，会议模式、联系人分组模式等，通过用户的选择，在确定了应用的场景后，该设置界面将向用户提供基于所选模式对应的具体启动条件，例如，当用户选择会议模式后，用户将可以进一步的确定会议的具体时间段，以便在该时间段内将语音信息转换为文字信息，而该时间段不限定时间的间隔，也就是说，用户可以设置某一日中的一个或多个时间段，也可以设置多日中的一个或多个或多时间段，具体的，可以通过调用系统中的日程表，由用户在日程表中指定具体的会议时间，即设定将语音信息转换为文字信息的转换时间或转换时间段；而当用户选择联系人分组模式时，用户将可以进一步的选择具体的联系人，将选中联系人所发送的语音信息转换为文字信息，具体的，可以通过调用该应用中的通讯录，由于用户在该通讯录中选择至少一个联系人生成转换联系人群组。此外，还可以将两者加以结合，组成混合模式，在不同的时间段确定不同的转换联系人群组。

[0101] 而对于图5中所示出的界面，可以是在多个界面中实现用户的操作，即先让用户确定模式，在根据模式确定进一步的启动条件。但也可以一同设置在同一界面中，通过操作权限的设置来确定用户可选择设置的部分内容。具体方式在本发明实施例中不做具体限定。需要说明的是，该设置需要预先设置，在生效后执行相应的转换操作。并且，该设置也可以实时修改调整。通过对语音转换模式的设置，同时确定了不同语音转换模式所对应的转换条件信息，如此，在后续执行接收到语音信息时，系统就可以根据所设置的语音转换模式来获取该语音信息所关联的转换条件信息。

[0102] 402、根据所接收的语音信息以及用户的操作提取对应的转换条件信息。

[0103] 根据步骤401中的具体设置，当接收到语音信息，同时，接收到用户在第一电子设

备上选择查看该语音信息的操作时，将根据所设置的具体内容获取该语音信息相关联转换条件信息，例如，对应上述步骤中的例子，当设置的为会议模式时，将获取接收到语音信息的时间信息作为转换条件信息；而当设置的为联系人分组模式时，将获取语音信息的发送来源，即发送联系人，将该联系人信息作为转换条件信息；而当设置的为混合模式时，则需要获取该语音信息对应的接收时间信息和发送联系人信息，将两者结合生成对应的转换条件信息。

[0104] 403、利用转换条件信息判断语音转换模式是否启动。

[0105] 本步骤是将步骤402中获取的转换条件信息与步骤401中用户设置的语音转换模式中的启动条件进行匹配，判断转换条件信息是否符合语音转换模式中的启动条件，若符合，则执行步骤404，若不符合，则执行步骤405。对应本发明实施例中上述步骤中的例子，当设置的为会议模式时，在所提取的语音接收时间信息中提取时间点，再判断该时间点是否落在设置的会议时间段中，若在，则确定该语音信息需要进行转换，即会议模式已启动，反之则不对该语音信息进行转换处理；当设置的为联系人分组模式时，则提取该语音信息的联系人信息的联系人，再判断该联系人是否在所设置的联系人群组中，若在，则确定该语音信息需要进行转换，反之则不对该语音信息进行转换处理；而当设置的为混合模式时，则需要提取该语音信息中的接收时间点和发送联系人，先判断接火速时间点是否落在了设置的时间段中，若在则进一步判断发送联系人是否为该时间段中所设置的联系人群组中的联系人，若是，则确定该语音信息需要进行转换，否则将不对该语音信息进行转换处理。

[0106] 对于设置有多种模式的情况，则逐一判断各个转换条件信息，若存在符合对应的语音转换模式时，就将该语音信息转换为文字信息。

[0107] 404、若语音转换模式启动，则将语音信息转换为文字信息，并显示该文字信息。

[0108] 根据步骤403的判断，当转换条件信息与语音转换模式相匹配时，将执行对该声音信息的转换操作，本发明实施例中对于将语音信息转换为文字信息的具体步骤不做具体限定，具体可参考上述实施例中在步骤203中的实施方式。

[0109] 而对于显示转换得到的文字信息时，本发明实施例中提供两种不同的显示方式：

[0110] 一种是同时显示语音信息与转换后对应的文字信息，即在完成对声音信息的转换后，同时显示语音信息与转换后的文字信息，并在这两个信息中标记相互转换的标识信息，或者是将两个信息绑定为一组信息，如图6中所示的显示方式，其中，A中是在界面中对语音信息与文字信息标注标识信息，B中是将语音信息与文字信息绑定为一组信息在单独界面中显示。从而在用户查看消息时，能够清楚的区分出哪些文字信息是经过转换操作得到的，并且还可以快速识别出转换后的文字信息所对应的语音信息是哪一条，以便于在转换的文字信息转换不准确时，用户能够快速找到对应的原始语音信息进行查询。

[0111] 另一种是仅显示转换后的文字信息，同样，为了与其他普通的文字信息加以区别，还可以在转换后的文字信息上标注其由语音信息转换而来的标识信息。进一步的，由于不显示原始的语音信息，若用户在查看转换后的文字信息后还需要再次查询原始的语音信息时，可以在所显示的文字信息上设置显示该文字信息对应语音信息的标签，该标签具有触发操作的功能，可以通过触发该标签实现显示该文字信息对应的语音信息，也可以奖该标签与文字信息自身进行合并，使得用户通过触发该文字信息就可以进一步地显示该文字信息所对应的原始语音信息，也就是说，在该显示方式下，文字信息或者标记在转换后的文字

信息上的标签被设置为显示原始语音信息的触发条件，使得用户在看到转换后的文字信息时无需再查找对应的语音信息，仅需要一步操作便可实现对原始语音信息的调取查询，这样不仅提高了用户使用的便捷性，更提升了应用于用户操作的互动性，大幅提高的应用的操作体验。

[0112] 405、若语音转换模式未启动，则显示语音信息。

[0113] 本步骤在语音转换模式未启动，而直接显示了接收的语音信息之后，设备还能够根据用户针对该语音信息的操作，将该语音信息转换为文字信息，而用户的操作就是将语音信息转换为文字信息的触发指令。而对于该触发指令也不限定是在该语音信息的具体状态，即用户是否已经查询过该语音信息，或者是未查询该语音信息。

[0114] 通过本发明实施例中的实现方式，用户可以通过设置界面针对不同的应用场景设置多种不同的语音转换模式，并且根据所接收到的语音信息中对应的转换条件信息判断是否存在启动的语音转换模式，使得用户在不方便收听语音信息的场景下能够直接获取该语音信息转换后的文字信息，并且在显示转换后的文字信息时，也提供了多种的显示方式供用户选择，以使得用户能够根据自身的喜好或阅读方式选择适合的显示方式，提升用户的使用体验。对于不在语音转换模式范围内的语音信息，同样向用户提供了将其转换为文字信息的实现方式，这对于不习惯收听或无法收听语音信息的用户则提供了以文字信息的查询方式，提供了更多的信息查询途径。

[0115] 进一步的，作为对上述方法的实现，本发明实施例提供了一种声音信息的转换装置，该装置实施例与前述方法实施例对应，为便于阅读，本装置实施例不再对前述方法实施例中的细节内容进行逐一赘述，但应当明确，本实施例中的装置能够对应实现前述方法实施例中的全部内容。该装置用于智能设备中，具体如图7所示，该装置包括：处理器51，显示屏52以及声音输出模块53，所述声音输出模块包括扬声器531以及耳机接口532；

[0116] 其中，所述处理器51，用于接收用户操作，所述用户操作用于指示所述第一电子设备呈现第一消息，所述第一消息的内容中包含音频格式的第一声音内容；

[0117] 判断所述第一消息的呈现方式，获得判断结果；

[0118] 当所述处理器51的判断结果为第一判断结果时，控制所述声音输出模块53输出所述第一声音内容；

[0119] 当所述处理器51的判断结果为第二判断结果时，将所述第一声音内容转化为非音频格式，控制所述显示屏52输出所述非音频格式的所述第一声音内容。

[0120] 进一步的，所述处理器51还用于，判断用户当前是否便于处理音频格式的文件消息。

[0121] 进一步的，所述处理器51在判断用户当前是否便于处理音频格式的文件消息时，还包括：

[0122] 获取日程信息，根据所述日程信息判断当前时间是否处于特定时间段内。

[0123] 进一步的，所述处理器51将所述第一声音内容转化为非音频格式包括：

[0124] 当所述第一声音内容为语音信息时，识别所述语音信息；

[0125] 利用语音数据库将所述语音信息转换为预置的文字信息。

[0126] 进一步的，所述处理器51将所述第一声音内容转化为非音频格式包括：

[0127] 当所述第一声音内容为拟声信息时，确定所述拟声信息的发声源；

[0128] 将所述拟声信息转换为文字呈现的发生源。

[0129] 进一步的，所述处理器51将所述第一声音内容转化为非音频格式包括：

[0130] 当所述第一声音内容为音乐时，确定所述音乐对应的标识信息，所述标识信息包括曲目名称、作者、音乐风格；

[0131] 将所述标识信息转换为文字呈现的标识信息。

[0132] 进一步的，所述处理器51控制显示屏输出所述非音频格式的所述第一声音内容包括：

[0133] 向所述显示屏52输出所述第一声音内容对应的文字信息，以及标记在所述文字信息上的转换标签，所述转换标签用于根据用户操作控制音频输出模块输出所述第一声音内容，以便所述显示屏向用户展示对应的文字信息。

[0134] 进一步的，本发明实施例提供了一种声音信息的转换设备，该设备中设置有上述的声音信息的转换装置，其包括但不限于智能手机、pad等。

[0135] 综上所述，本发明实施例所采用的声音信息的转换方法、装置及设备，是在用户根据应用场景的需要设置的多种不同的语音转换模式下，获取所接收的语音信息中与语音转换模式相对应的转换条件信息，当通过该转换条件信息确定语音转换模式启动时自动执行声音信息的转换操作，得到对应的文字信息，使得用户在不方便收听语音信息时可以及时、快速地查看该语音信息中的内容，提升了用户使用语音信息转换功能的便捷性。并且对于转换后的文字信息在设备上的具体显示，本发明实施例也通过多种具体的显示方式实现了用户在查看消息时可以方便、清楚的查询到转换后的文字信息以及对应的原始语音信息，进一步的简化用户查询语音信息的操作步骤，提升用户操作的使用体验。此外，对语音信息进行转化时，通过采用本地转换或服务器转换的方式还进一步的提高了语音信息转化为文字信息的准确率，而对于不在语音转换模式范围内的语音信息，也向用户提供了转换为文字信息的实现方式，使得用户可以通过更多途径查询语音信息。

[0136] 所述声音信息的转换装置包括处理器和存储器，由处理器执行存储在存储器中的上述程序单元来实现相应的功能。

[0137] 处理器中包含内核，由内核去存储器中调取相应的程序单元。内核可以设置一个或以上，通过调整内核参数来实现将用户接收的语音消息在符合预置的条件下自动的转换为文字消息。

[0138] 存储器可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)，存储器包括至少一个存储芯片。

[0139] 本申请还提供了一种计算机程序产品，当在数据处理设备上执行时，适于执行初始化有如下方法步骤的程序代码：接收用户操作，所述用户操作用于指示所述第一电子设备呈现第一消息，所述第一消息的内容中包含音频格式的第一声音内容；判断所述第一消息的呈现方式，获得判断结果；若所述判断结果为第一判断结果，则控制图像显示模块输出第一消息标识，以使音频输出模块输出所述第一声音内容；若所述判断结果为第二判断结果，则将所述第一声音内容转化为非音频格式，控制图像显示模块输出所述非音频格式的所述第一声音内容。

[0140] 本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序

产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0141] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0142] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0143] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0144] 在一个典型的配置中，计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0145] 存储器可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。存储器是计算机可读介质的示例。

[0146] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体(transitory media)，如调制的数据信号和载波。

[0147] 还需要说明的是，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、商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、商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，并不排除在包括要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0148] 本领域技术人员应明白，本申请的实施例可提供为方法、系统或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存

储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0149] 以上仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

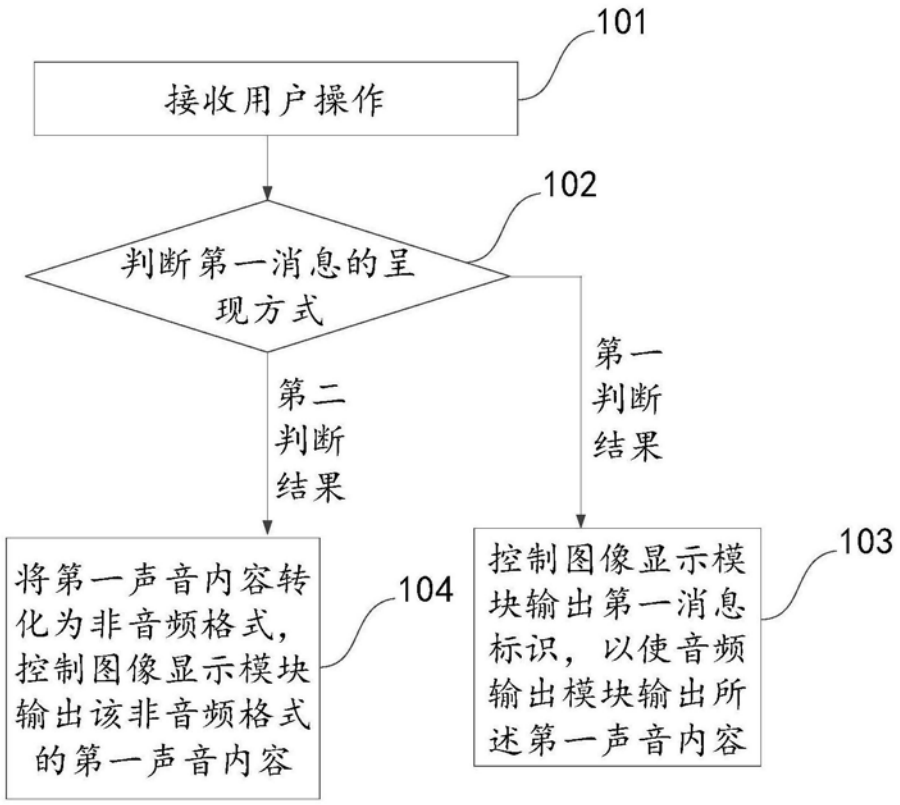

101
接收用户操作
102
判断第一消息的呈现方式
第一判断结果
第二判断结果
103
控制图像显示模块输出第一消息标识，以使音频输出模块输出所述第一声音内容
104
将第一声音内容转化为非音频格式，控制图像显示模块输出该非音频格式的第一声音内容

图1

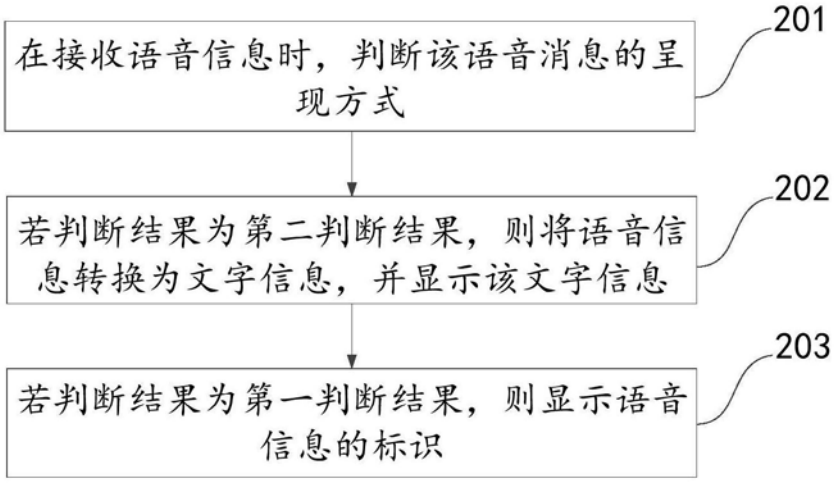

201
在接收语音信息时，判断该语音消息的呈现方式
202
若判断结果为第二判断结果，则将语音信息转换为文字信息，并显示该文字信息
203
若判断结果为第一判断结果，则显示语音信息的标识

图2

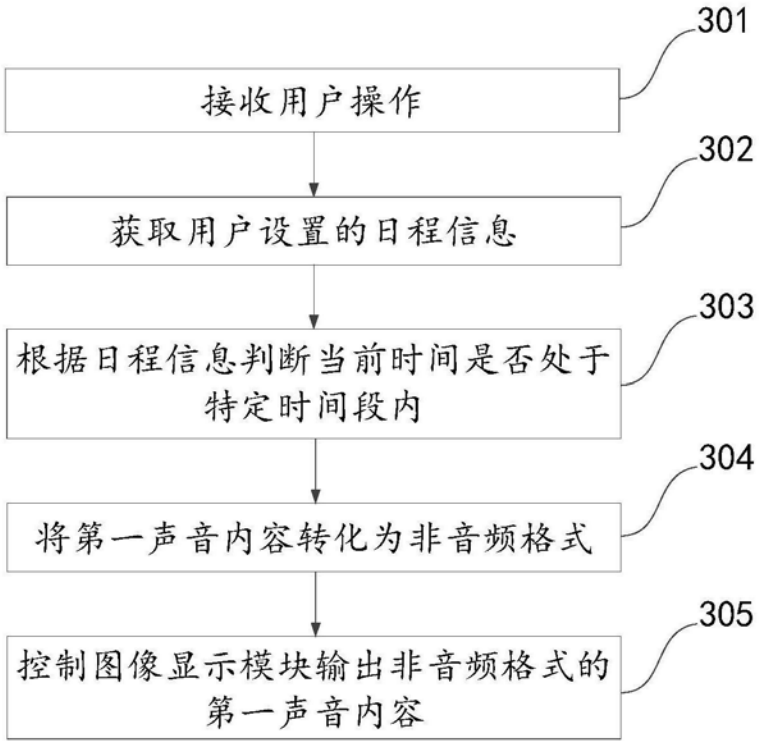

301
接收用户操作
302
获取用户设置的日程信息
303
根据日程信息判断当前时间是否处于特定时间段内
304
将第一声音内容转化为非音频格式
305
控制图像显示模块输出非音频格式的第一声音内容

图3

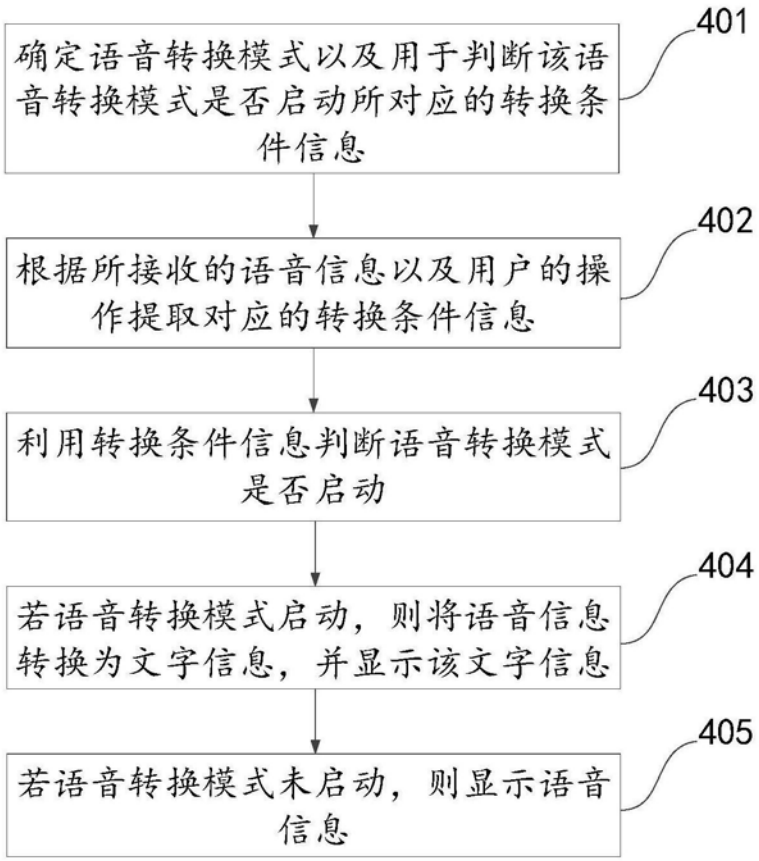

401
确定语音转换模式以及用于判断该语音转换模式是否启动所对应的转换条件信息
402
根据所接收的语音信息以及用户的操作提取对应的转换条件信息
403
利用转换条件信息判断语音转换模式是否启动
404
若语音转换模式启动，则将语音信息转换为文字信息，并显示该文字信息
405
若语音转换模式未启动，则显示语音信息

图4

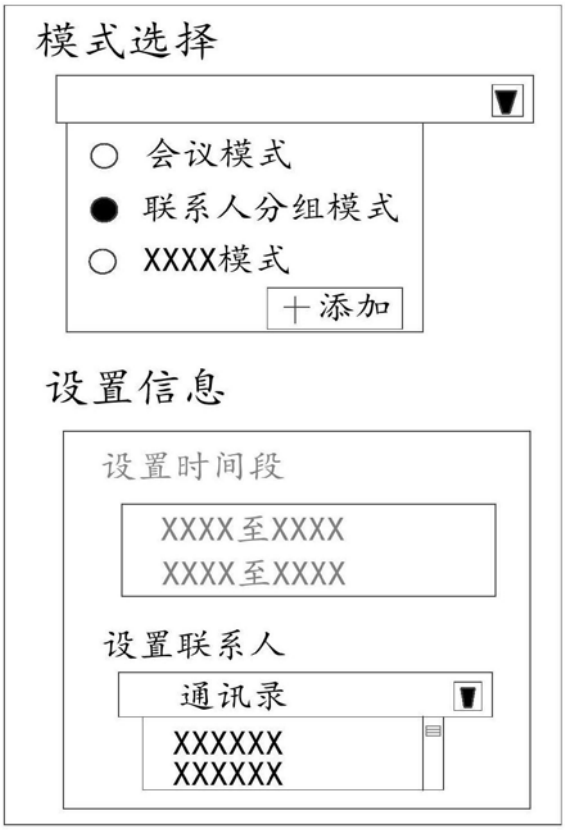

模式选择
会议模式
联系人分组模式
XXXX模式
+添加
设置信息
设置时间段
XXXX至XXXX
XXXX至XXXX
设置联系人
通讯录
XXXXXX
XXXXXX

图5

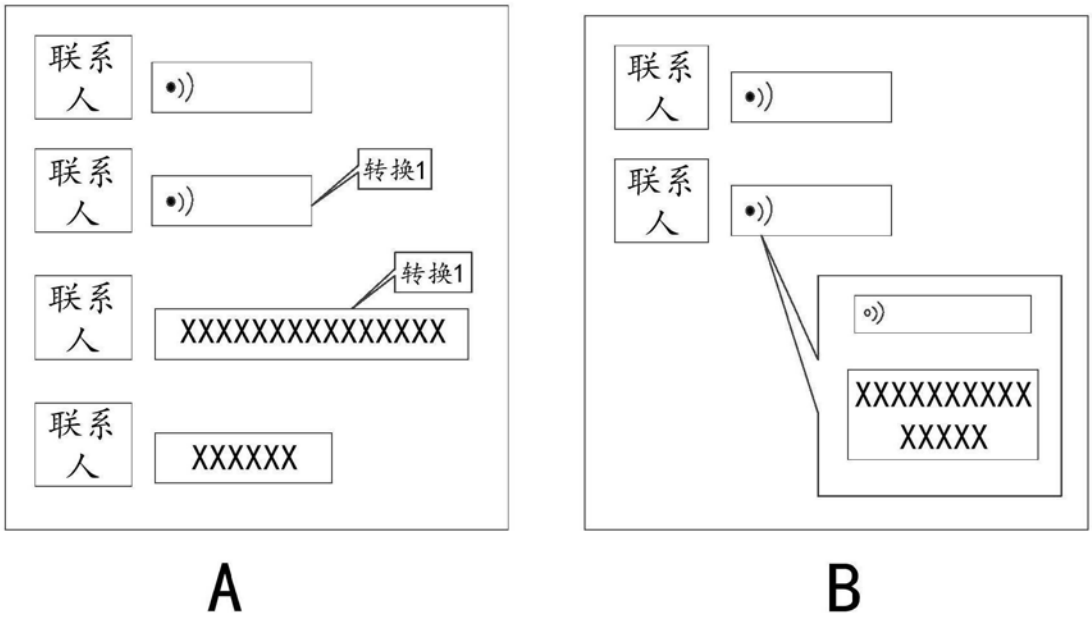

联系人
联系人
转换1
联系人
转换1
XXXXXXXXXXXXXXX
联系人
XXXXXX
A
联系人
联系人
XXXXXXXXXX
XXXXX
B

图6

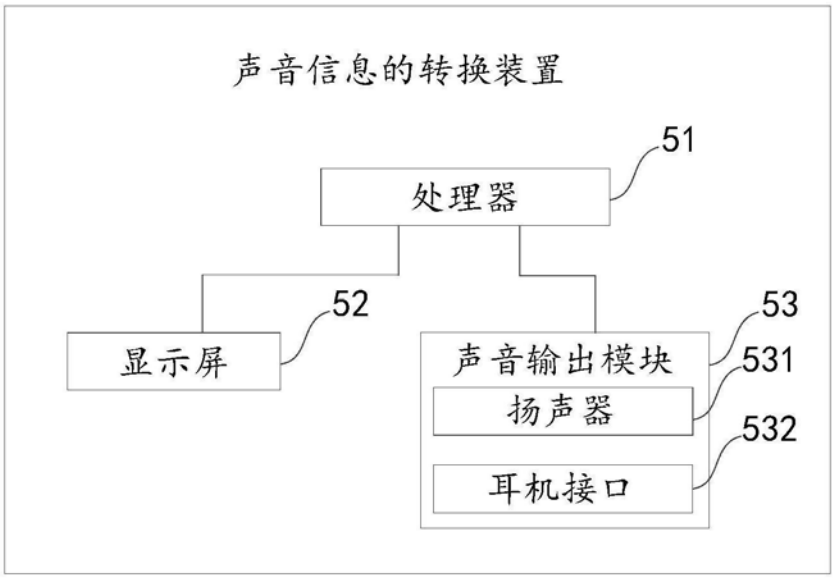


图7