



(12)发明专利申请

(10)申请公布号 CN 110164437 A

(43)申请公布日 2019.08.23

(21)申请号 201910434480.5

G10L 15/30(2013.01)

(22)申请日 2012.03.02

(62)分案原申请数据

201210053628.9 2012.03.02

(71)申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 路依莎

(74)专利代理机构 北京三高永信知识产权代理
有限责任公司 11138

代理人 张所明

(51)Int.Cl.

G10L 15/22(2006.01)

H04L 12/58(2006.01)

G10L 15/26(2006.01)

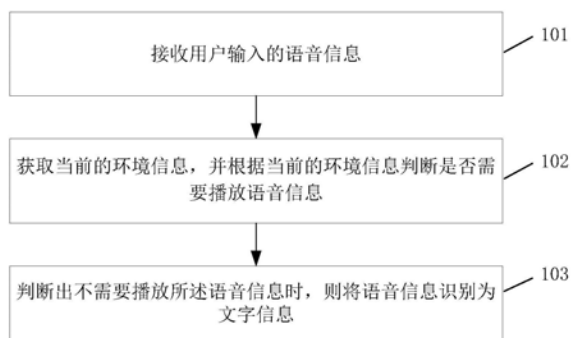
权利要求书2页 说明书7页 附图3页

(54)发明名称

一种即时通信的语音识别方法和终端

(57)摘要

本发明公开了一种即时通信的语音识别方法和终端,属于通信领域。所述方法包括:接收用户输入的语音信息;获取当前的环境信息,并根据所述当前的环境信息判断是否需要播放所述语音信息;判断出不需要播放所述语音信息时,则将所述语音信息识别为文字信息。所述终端包括获取模块、判断模块和识别模块。本发明通过终端接收到即时通信的语音信息时,为语音接收方提供语音识别的功能,以帮助接收方在不方便的情况下正常获知对方要表达的内容。



1. 一种即时通信的语音识别方法,其特征在于,所述方法包括:
 - 接收用户输入的语音信息;
 - 获取当前的环境信息,根据所述当前的环境信息判断是否需要播放所述语音信息;
 - 当根据当前的环境信息判断需要播放所述语音信息时,播放所述语音信息;
 - 当根据当前的环境信息判断不需要播放所述语音信息时,将所述语音信息识别为文字信息,并显示所述文字信息;其中,所述当前的环境信息包括当前的噪声信息,所述根据所述当前的环境信息判断是否需要播放所述语音信息,包括:
 - 判断所述当前噪声信息是否在预设范围内;
 - 当所述当前噪声信息在预设范围内时,判断出需要播放所述语音信息;
 - 当所述当前噪声信息不在预设范围内时,判断出不需要播放所述语音信息。
2. 根据权利要求1所述的方法,其特征在于,所述将所述语音信息识别为文字信息,包括:
 - 发送所述语音信息至语音识别服务器,以便所述语音识别服务器识别得到所述文字信息;
 - 接收所述语音识别服务器识别得到的文字信息。
3. 根据权利要求1所述的方法,其特征在于,将所述语音信息识别为文字信息,包括:
 - 通过内置的语音识别模块对所述语音信息进行语音识别,获取所述文字信息。
4. 根据权利要求1所述的方法,其特征在于,所述当前的环境信息包括用户的操作指令,所述根据所述当前的环境信息判断是否需要播放所述语音信息,包括:
 - 当所述用户的操作指令是获取文字信息的指令时,判断出不需要播放所述语音信息;
 - 当所述用户的操作指令不是获取文字信息的指令时,判断出需要播放所述语音信息。
5. 根据权利要求4所述的方法,其特征在于,所述当所述用户的操作指令是获取文字信息的指令时,判断出不需要播放所述语音信息,包括:
 - 在即时通信的界面中,显示语音信息;
 - 在所述语音信息后方,显示功能键,所述功能键用于指示获取所述文字信息;
 - 当所述功能键被触发时,判断出不需要播放所述语音信息。
6. 根据权利要求5所述的方法,其特征在于,所述语音信息和所述功能键一一对应显示。
7. 根据权利要求1所述的方法,其特征在于,所述当前的环境信息包括终端自身的硬件信息,所述根据所述当前的环境信息判断是否需要播放所述语音信息,包括:
 - 当所述终端自身的硬件信息满足播放所述语音信息的条件时,判断出需要播放所述语音信息;
 - 当所述终端自身的硬件信息不满足播放所述语音信息的条件时,判断出不需要播放所述语音信息。
8. 一种终端,其特征在于,所述终端包括:
 - 获取模块,用于接收用户输入的语音信息;
 - 判断模块,用于获取当前的环境信息,根据所述当前的环境信息判断是否需要播放所述语音信息;

识别模块,用于当所述判断模块的判断结果为否时,将所述语音信息识别为文字信息;
其中,所述判断模块,包括:获取单元和判断单元;

所述获取单元,用于获取当前的噪声信息,所述当前的环境信息包括所述当前的噪声信息;

所述判断单元,用于当前的噪声信息在预设的范围内,则判断出需要播放语音信息;当当前的噪声信息不在预设的范围内,则判断出不需要播放语音信息。

9. 根据权利要求8所述的终端,其特征在于,所述识别模块包括发送单元和接收单元,所述发送单元,用于当所述判断模块的判断结果为否时,将所述语音信息发送给语音识别服务器,以便所述语音识别服务器识别得到所述文字信息;

所述接收单元,用于接收所述语音识别服务器识别得到的文字信息。

10. 根据权利要求8所述的终端,其特征在于,

所述识别模块,用于通过内置的语音识别模块对所述语音信息进行语音识别,获取所述文字信息。

11. 根据权利要求8所述的终端,其特征在于,所述判断模块包括获取单元和判断单元,所述获取单元,用于获取用户的操作指令;

所述判断单元,用于当所述用户的操作指令是获取文字信息的指令时,则判断出不需要播放所述语音信息;当所述用户的操作指令不是获取文字信息的指令时,则判断出需要播放所述语音信息。

12. 根据权利要求8所述的终端,其特征在于,所述判断模块包括获取单元和判断单元,所述获取单元,用于获取终端自身的硬件信息;

所述判断单元,用于当所述终端自身的硬件信息满足播放所述语音信息的条件时,则判断出需要播放所述语音信息;当所述终端自身的硬件信息不满足播放所述语音信息的条件时,则判断出不需要播放所述语音信息。

13. 如权利要求8所述的终端,其特征在于,所述终端为移动终端或PC终端。

14. 一种计算机可读存储介质,其特征在于,所述存储介质中存储有程序,所述程序用于指令相关的硬件以完成如权利要求1至7任一所述的即时通信的语音识别方法。

一种即时通信的语音识别方法和终端

[0001] 本申请为2012年3月2日提交的申请号为201210053628.9、发明名称为“一种即时通信的语音识别方法和终端”的中国专利申请的分案申请。

技术领域

[0002] 本发明涉及通信领域,特别涉及一种即时通信的语音识别方法和终端。

背景技术

[0003] 随着网络的发展,越来越多的人逐渐开始依赖移动装置打发时间、与朋友聊天和娱乐消遣,其中,移动终端网民中,近3/4的用户用聊天工具和好友交流。目前主流的聊天工具中,语音输入越来越流行,语音输入摆脱了传统虚拟键盘输入反复点击的操作,用户之间的沟通更快捷、更流畅。

[0004] 目前终端的即时通信工具中大多提供了语音输入的功能,比如PC端QQ、移动终端的微信对讲、手机QQ的语音输入,米聊的对讲等等。但是很多场景下,如在会议室或嘈杂的环境下,或接收方没有耳机或播放器或语音内容比较私密等,基于即时通信进行语音输入时,接收方不太方便直接播放语音,此时,如何让接收方直接获取到语音内容是本领域人员一直关注的问题。

发明内容

[0005] 为了在即时通信时让接收方直接获取到语音内容,本发明实施例提供了一种即时通信的语音识别方法和终端。所述技术方案如下:

[0006] 本发明实施例提供了一种即时通信的语音识别方法,所述方法包括:

[0007] 接收用户输入的语音信息;

[0008] 获取当前的环境信息,并根据所述当前的环境信息判断是否需要播放所述语音信息;

[0009] 判断出不需要播放所述语音信息时,则将所述语音信息识别为文字信息。

[0010] 其中,所述将所述语音信息识别为文字信息,具体包括:

[0011] 将所述语音信息发送给云端的语音识别服务器,使所述语音识别服务器将所述语音信息识别为文字信息;

[0012] 接收所述语音识别服务器识别得到的文字信息。

[0013] 其中,所述将语音信息识别为文字信息,具体包括:

[0014] 将所述语音信息通过终端内置的语音识别模块进行语音识别,得到文字信息。

[0015] 其中,所述当前的环境信息具体为用户的操作指令,所述根据所述当前的环境信息判断是否需要播放所述语音信息,具体包括:

[0016] 当所述用户的操作指令是获取文字信息的指令时,则判断出不需要播放所述语音;

[0017] 当所述用户的操作指令不是获取文字信息的指令时,则判断出需要播放所述语音

信息。

[0018] 其中,所述当前的环境信息具体包括终端自身的硬件信息,所述根据所述当前的环境信息判断是否需要播放所述语音信息,具体包括:

[0019] 当所述终端自身的硬件信息满足播放所述语音信息的条件时,则判断出需要播放所述语音信息;

[0020] 当所述终端自身的硬件信息不满足播放所述语音信息的条件时,则判断出不需要播放所述语音信息。

[0021] 其中,所述当前的环境信息具体包括当前的噪声信息,所述根据所述当前的环境信息判断是否需要播放所述语音信息,具体包括:

[0022] 判断所述当前的噪声信息是否在预设的范围内;

[0023] 当所述当前的噪声信息在预设的范围内,则判断出需要播放所述语音信息;

[0024] 当所述当前的噪声信息不在预设的范围内,则判断出不需要播放所述语音信息。

[0025] 其中,所述终端为移动终端或PC终端。

[0026] 本发明实施例还提供了一种终端,所述终端包括:

[0027] 获取模块,用于接收用户输入的语音信息;

[0028] 判断模块,用于获取当前的环境信息,根据所述当前的环境信息判断是否需要播放所述语音信息;

[0029] 识别模块,用于当所述判断模块的判断结果为否时,将所述语音信息识别为文字信息。

[0030] 其中,所述识别模块具体包括发送单元和接收单元;

[0031] 所述发送单元,用于当所述判断模块的判断结果为否时,将所述语音信息发送给云端的语音识别服务器,使所述语音识别服务器将所述语音信息识别为文字信息;

[0032] 所述接收单元,用于接收所述语音识别服务器识别得到的文字信息。

[0033] 其中,所述识别模块具体包括:当所述判断模块的判断结果为否时,将所述语音信息通过终端内置的语音识别模块识别为文字信息。

[0034] 其中,所述判断模块具体包括获取单元和判断单元,

[0035] 所述获取单元,用于获取用户的操作指令;

[0036] 所述判断单元,用于当所述用户的操作指令是获取文字信息的指令时,则判断出不需要播放所述语音信息;当所述用户的操作指令不是获取文字信息的指令时,则判断出需要播放所述语音信息。

[0037] 其中,所述判断模块具体包括获取单元和判断单元;

[0038] 所述获取单元,用于获取终端自身的硬件信息;

[0039] 所述判断单元,用于当所述终端自身的硬件信息满足播放所述语音信息的条件时,则判断出需要播放所述语音信息;当所述终端自身的硬件信息不满足播放所述语音信息的条件时,则判断出不需要播放所述语音信息。

[0040] 其中,所述判断模块具体包括获取单元和判断单元;

[0041] 所述获取单元,用于获取当前的噪声信息;

[0042] 所述判断单元,用于当所述当前的噪声信息在预设的范围内,则判断出需要播放所述语音信息;当所述当前的噪声信息不在预设的范围内,则判断出不需要播放所述语

音信息件。

[0043] 其中,所述终端为移动终端或PC终端。

[0044] 本发明实施例提供的技术方案带来的有益效果是:通过终端接收到即时通信的语音信息时,为语音接收方提供语音识别的功能,以帮助接收方在不方便的情况下正常获知对方要表达的内容。例如在公共场合,不方便播放语音时,或在无播放设备的情景下,无法播放语音时,用户能够快捷的获知语音内容。

附图说明

[0045] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅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。

[0046] 图1是本发明实施例1中提供的基于即时通信的语音识别方法流程图;

[0047] 图2是本发明实施例2中提供的基于即时通信的语音识别方法流程图;

[0048] 图3是本发明实施例2中在移动终端下的即时通信的语音识别示意图;

[0049] 图4是本发明实施例2中在PC端下的即时通信的语音识别示意图;

[0050] 图5是本发明实施例3中提供的终端的结构示意图;

[0051] 图6是本发明实施例3中提供的终端中判断模块的结构示意图。

具体实施方式

[0052] 为使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明实施方式作进一步地详细描述。

[0053] 实施例1

[0054] 参见图1,本发明实施例提供了一种基于即时通信的语音识别方法,所述方法包括:

[0055] 步骤101:接收用户输入的语音信息;

[0056] 步骤102:获取当前的环境信息,并根据当前的环境信息判断是否需要播放语音信息;

[0057] 步骤103:判断出不需要播放所述语音信息时,则将语音信息识别为文字信息。

[0058] 本发明实施例提供的方法,通过终端接收到即时通信的语音信息时,为语音接收方提供语音识别的功能,以帮助接收方在不方便的情况下正常获知对方要表达的内容。例如在公共场合,不方便播放语音时,或在无播放设备的情景下,无法播放语音时,用户能够快捷的获知语音内容。

[0059] 实施例2

[0060] 参见图2,本发明实施例提供了一种即时通信的语音识别方法,所述方法包括:

[0061] 步骤201:接收用户输入的语音信息;

[0062] 该步骤中,终端接收到用户输入的语音信息。具体地,本发明实施例中的应用场景,可以为用户在互联网应用的即时场景下输入语音信息,也可以为用户在聊天场景下输入语音信息。

[0063] 其中,本发明实施例中的终端为移动终端或PC终端。该移动终端可以为手机、PDA(Personal Digital Assistant,个人数字助理)和MID(Mobile Internet Devices,移动互联网设备)等。

[0064] 步骤202:获取当前的环境信息;

[0065] 该步骤中,本发明实施例中的当前的环境信息可以为终端自身的硬件信息,也可以为用户的操作指令,也可以为当前的噪声信息等等。本发明实施例中对当前的环境信息的具体内容不做限定。

[0066] 步骤203:根据当前的环境信息判断是否需要播放语音信息,若否,则执行步骤204,若是,则执行步骤206;

[0067] 其中,语音输入确实能够带给用户更快捷、更流畅的沟通过程,但在很多场景下(比如会议室、嘈杂的环境或对方没有耳机或播放器,或者语音内容比较私密等情况下),接收方不太方便播放语音。此时,为了让接收方顺利的获知语音发送者要表达的内容,可以将语音信息识别为文字信息,便于接收方进行选择,接收方可以选择播放语音信息,也可以选择查看识别后的文字信息。

[0068] 该步骤中,根据当前的环境信息不同,判断是否需要播放语音信息的方式也不同,本发明实施例提供下面几种实现方式:

[0069] 方式一、当前的环境信息为用户的操作指令;

[0070] 根据当前的环境信息判断是否需要播放语音信息,具体包括:判断该操作指令是否为获取文字信息的指令,当该操作指令为获取文字信息的指令时,则判断出不需要播放语音信息,执行步骤204的操作;当用户的操作指令不是获取文字信息的指令时,则判断出需要播放语音信息,执行步骤206的操作。

[0071] 具体地,用户通过执行操作指令读取终端上接收的语音信息,该操作指令可以为点击操作或触摸操作,本发明在此不做限定。终端根据用户的操作指令判断出用户选择哪种获取信息的方式,其中,终端接收到用户的操作指令后,根据该操作指令判断出终端是否需要播放语音信息,若用户不需要播放语音信息时,则需要将该语音信息识别为文字信息并提供给用户;若用户需要播放语音信息时,则终端将该语音信息直接播放给用户即可,不需要做其他的特殊处理操作。

[0072] 例如,终端上设置至少两个按钮A和B,按钮A上显示“语音”、“喇叭形状”等字符,按钮B上显示“文字”、“ABC”等字符,其中,用户若点击或触摸按钮A,则表明获取到语音信息,播放出该语音信息;用户若点击或触摸按钮B,则表明获取到文字信息,将该文字信息显示出来。

[0073] 方式二、当前的环境信息为终端自身的硬件信息;

[0074] 其中,终端满足播放语音信息的硬件信息至少包括音频编解码器和音频输出设备。

[0075] 根据当前的环境信息判断是否需要播放语音信息,具体包括:判断终端自身的硬件条件是否具备音频解码器和音频输出设备,当该终端自身的硬件信息具备音频编解码器和音频输出设备时,则该终端自身的硬件信息满足播放语音信息的条件,此时判断出该终端需要播放语音信息,执行步骤206的操作;当该终端自身的硬件信息不具备音频编解码器,或不具备音频输出设备,或不具备音频编解码和音频输出设备时,则该终端自身的硬件

信息不满足播放语音信息的条件,判断出不需要播放语音信息,执行步骤204的操作。

[0076] 方式三、当前的环境信息为当前的噪声信息;

[0077] 本发明预设一定的噪声范围,如预设的范围为30分贝至90分贝,根据当前的环境信息判断是否需要播放语音信息,具体包括:判断当前的噪声信息是否在预设的范围内;当当前的噪声信息在预设的范围内,则判断出需要播放语音信息,执行步骤206的操作;当当前的噪声信息不在预设的范围内,则表明当前环境为无声环境或特别嘈杂的环境,判断出不需要播放语音信息,执行步骤204的操作。

[0078] 其中,该方式一,方式二,方式三为或的关系,实际应用中,可以采用方式一的实现方式,也可以采用方式二的实现方式,也可以采用方式三的实现方式,也可以设置三种实现方式的优先级,根据优先级的高低进行实现。

[0079] 步骤204:将语音信息识别为文字信息;

[0080] 具体地,本发明实施例提供的将获取的语音信息识别为文字信息的方式,可以在云端进行识别,也可在本地内置识别模块,本发明对此不做限制。

[0081] 其中,在云端将语音信息识别为文字信息时,终端将该语音信息发送给云端的语音识别服务器,由语音识别服务器进行语音识别,云端即指网络上,云端的语音识别服务器即指网络上的语音识别服务器,然后语音识别服务器将识别结果传输给终端,终端接收到该语音识别服务器识别得到的文字信息。其中,终端将该语音信息传输到网络上指定的语音识别服务器,该语音识别服务器可以对一种或多种语音信息进行识别。

[0082] 其中,在本地内置识别模块将语音信息识别为文字信息时,终端将该语音信息通过终端内置的语音识别模块进行语音识别,得到文字信息。其中,终端在安装即时通信工具时,就已经内置了语音识别模块,终端欲识别该语音信息时,直接使用该模块即可对待识别的语音信息进行语音识别,然后直接将该识别结果提供给用户即可。

[0083] 本发明实施例中,无论是云端识别还是本地内置识别模块,都需要用到语音识别技术,比如:google、百度、讯飞等,本发明实施例对具体采用哪种语音识别技术不做限定,只要能将语音信息识别为文字信息即可。

[0084] 以下将从移动终端和PC终端两种场景下描述本发明方案。

[0085] 移动终端场景:在移动终端的即时通信的聊天场景下,如图3所示,在语音信息后面,提供“文字”的功能键,用户点击此键时,表示用户的操作指令为获取文字信息的指令,此时移动终端将语音信息识别成文字信息。

[0086] PC终端场景:在PC终端的即时通信的聊天场景下,如图4所示,在语音信息后面,提供“识别成文字”的功能键,用户点击此键,表示用户的操作指令为获取文字信息的指令,此时PC终端将该语音信息识别成文字信息。

[0087] 步骤205:将该文字信息显示给用户;

[0088] 其中,将语音信息识别为文字信息后,将该文字信息显示在终端屏幕上。

[0089] 步骤206:播放该语音信息。

[0090] 具体地,终端判断出需要播放语音信息时,则直接播放该语音信息。

[0091] 本发明实施例提供的方法,通过终端接收到即时通信的语音信息时,为语音接收方提供语音识别的功能,以帮助接收方在不方便的情况下正常获知对方要表达的内容。例如在公共场合,不方便播放语音时,或在无播放设备的情景下,无法播放语音时,用户能够

快捷的获知语音内容。

[0092] 实施例3

[0093] 参见图5,本发明实施例提供了一种终端,所述终端包括:

[0094] 获取模块301,用于接收用户输入的语音信息;

[0095] 判断模块302,用于获取当前的环境信息,根据所述当前的环境信息判断是否需要播放所述语音信息;

[0096] 识别模块303,用于当判断模块302的判断结果为否时,将语音信息识别文字信息。

[0097] 其中,识别模块303具体包括发送单元和接收单元;

[0098] 发送单元,用于当判断模块302的判断结果为否时,将语音信息发送给语音识别服务器,使语音识别服务器将所述语音信息识别为文字信息;

[0099] 接收单元,用于接收语音识别服务器识别得到的文字信息。

[0100] 其中,识别模块303具体包括:当判断模块302的判断结果为否时,将语音信息通过终端内置的语音识别模块识别为文字信息。

[0101] 其中,参见图6,判断模块302具体包括获取单元3021和判断单元3022,

[0102] 获取单元3021,用于获取用户的操作指令;

[0103] 判断单元3022,用于当用户的操作指令是获取文字信息的指令时,则判断出不需要播放语音信息;当用户的操作指令不是获取文字信息的指令时,则判断出需要播放语音信息。

[0104] 其中,判断模块302具体包括获取单元3021和判断单元3022;

[0105] 获取单元3021,用于获取终端自身的硬件信息;

[0106] 判断单元3022,用于当终端自身的硬件信息满足播放语音信息的条件时,则判断出需要播放语音信息;当终端自身的硬件信息不满足播放语音信息的条件时,则判断出不需要播放语音信息。

[0107] 其中,判断模块302具体包括获取单元3021和判断单元3022;

[0108] 获取单元3021,用于获取当前的噪声信息;

[0109] 判断单元3022,用于当当前的噪声信息在预设的范围内,则判断出需要播放语音信息;当当前的噪声信息不在预设的范围内,则判断出不需要播放语音信息。

[0110] 其中,该终端具体为移动终端或PC终端。该移动终端可以为手机、PDA(Personal Digital Assistant,个人数字助理)和MID(Mobile Internet Devices,移动互联网设备)等。

[0111] 本发明实施例提供的终端,通过终端接收到即时通信的语音信息时,为语音接收方提供语音识别的功能,以帮助接收方在不方便的情况下正常获知对方要表达的内容。例如在公共场合,不方便播放语音时,或在无播放设备的情景下,无法播放语音时,用户能够快捷的获知语音内容。

[0112] 本实施例提供的终端,具体可以与方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0113] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0114] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

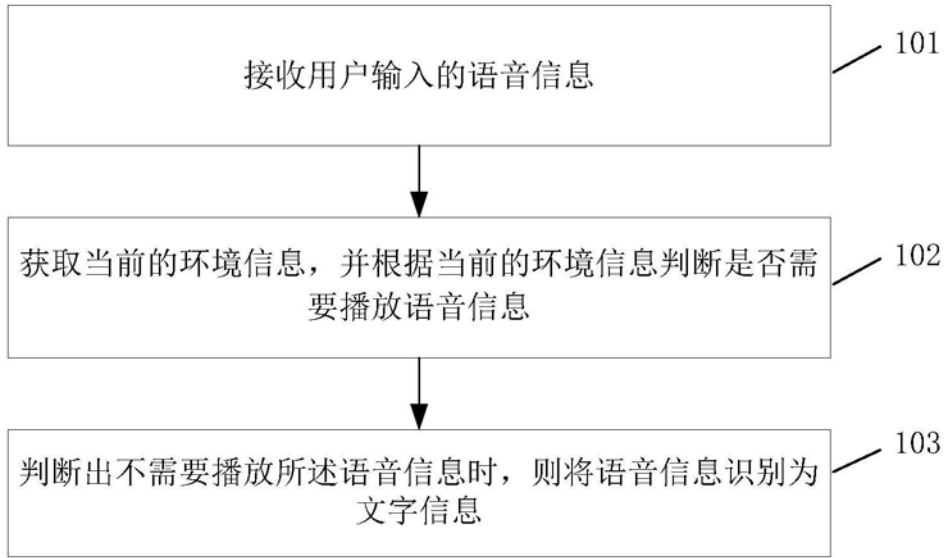


图1

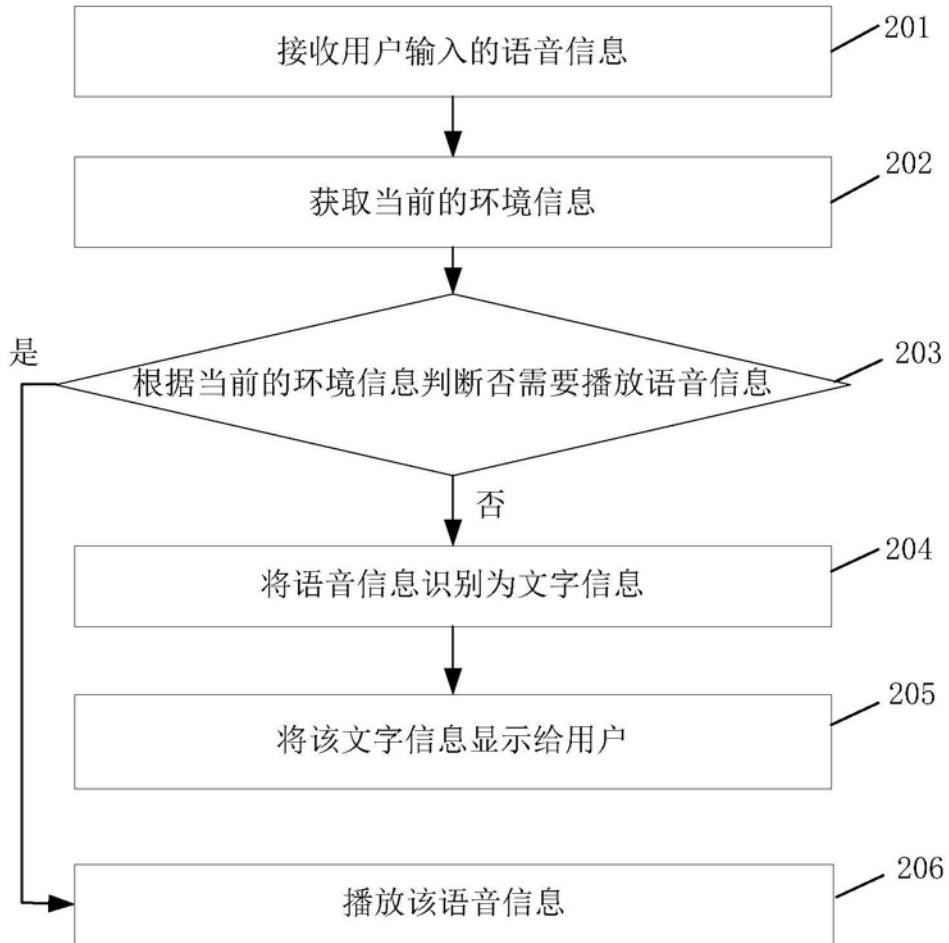


图2



图3

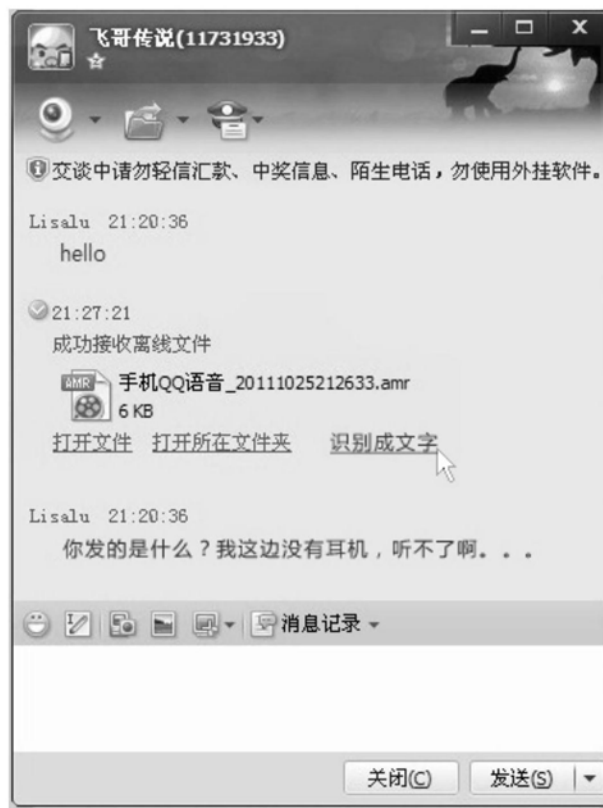


图4

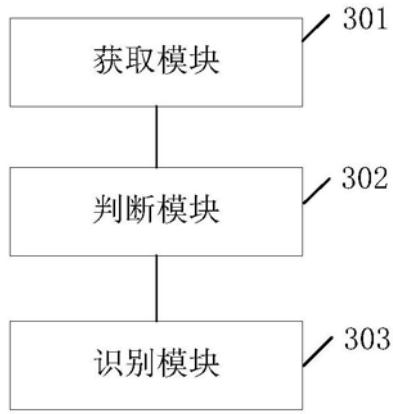


图5

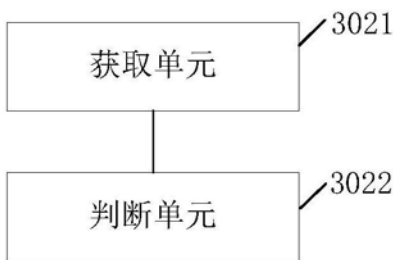


图6