

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2013 年 9 月 6 日 (06.09.2013)



(10) 国际公布号
WO 2013/127367 A1

- (51) 国际专利分类号:
G10L 15/26 (2006.01) *H04M 7/00* (2006.01)
H04L 12/58 (2006.01)
- (21) 国际申请号: PCT/CN2013/072086
- (22) 国际申请日: 2013 年 3 月 1 日 (01.03.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210053628.9 2012 年 3 月 2 日 (02.03.2012) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518044 (CN)。
- (72) 发明人: 路依莎 (LU, Yisha); 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518044 (CN)。
- (74) 代理人: 北京三高永信知识产权代理有限公司 (BEIJING SAN GAO YONG XIN INTELLECTU-

AL PROPERTY AGENCY CO., LTD.); 中国北京市海淀区学院路衙门里和景园 A-1-102, Beijing 100088 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: INSTANT COMMUNICATION VOICE RECOGNITION METHOD AND TERMINAL

(54) 发明名称: 一种即时通信的语音识别方法和终端

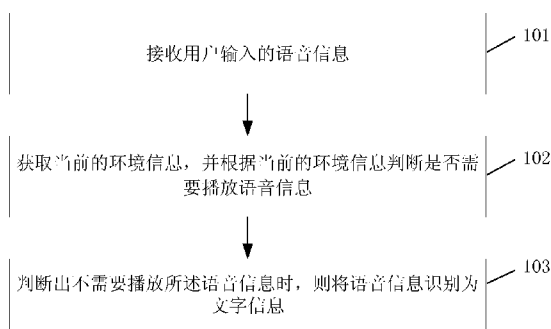


图 1 / Fig.1

- 101 Receiving voice information input by a user
- 102 Obtaining current environment information, and judging whether it is necessary to play the voice information according to the current environment information
- 103 When it is judged that there is no need to play the voice information, recognizing the voice information as word information

(57) Abstract: An instant communication voice recognition method and terminal belong to the field of communications. The method comprises: receiving voice information input by a user; obtaining current environment information, and judging whether it is necessary to play the voice information according to the current environment information; and when it is judged that there is no need to play the voice information, recognizing the voice information as word information. The terminal comprises an acquisition module, a judgment module and a recognition module.

(57) 摘要: 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。



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

说明书

一种即时通信的语音识别方法和终端

本申请要求于 2012 年 3 月 2 日提交中国专利局、申请号为 2012100536289、
5 发明名称为“一种即时通信的语音识别方法和终端”的中国专利申请的优先权，
其全部内容通过引用结合在本申请中。

技术领域

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。

10

背景技术

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。

15

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，
20 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

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：

25

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：

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。

30

其中，所述将所述语音信息识别为文字信息，具体包括：

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；

接收所述语音识别服务器识别得到的文字信息。

其中，所述将语音信息识别为文字信息，具体包括：

5 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。

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：

10 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；

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。

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：

15 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。

20 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：

判断所述当前的噪声信息是否在预设的范围内；

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；

25 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。

其中，所述终端为移动终端或 PC 终端。

本发明实施例还提供了一种终端，所述终端包括：

获取模块，用于接收用户输入的语音信息；

30 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，所述识别模块具体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；

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。

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。

其中，所述判断模块具体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件；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件。

其中，所述判断模块具体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件；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件。

其中，所述判断模块具体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件；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。

其中，所述终端为移动终端或 PC 终端。

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

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。

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；

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；

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；

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；

图 5 是本发明实施例 3 中提供的终端的结构示意图；

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

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。

实施例 1

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：

步骤 101：接收用户输入的语音信息；

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；

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。

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。

实施例 2

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：

步骤 201：接收用户输入的语音信息；

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。

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
5 Devices, 移动互联网设备) 等。

步骤 202: 获取当前的环境信息;

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。

10 其中，终端自身的硬件信息包括但不限于终端当前安装的硬件设备信息，如通过接口连接的音频输入或输出设备，该硬件信息还可以包括音频输入或输出设备的驱动信息等等，如音频编解码器和音频设备驱动等。

用户的操作指令可以是对终端进行的任一操作指令，根据技术人员的开发时的设置，不同的操作指令可以对应不同的功能，如，对语音信息的点击操作
15 可以为用于播放该语音信息的操作指令，而对于语音信息相应的获取文字信息的选项的点击操作可以为用于获取文字信息的操作指令。

当前的噪声信息是指终端所处环境的噪音信息，如该终端当前处于会议室，则该当前的噪声信息是指该会议室内的噪声信息。相应地，获取当前的环境信息即为获取当前的噪声信息，该噪声信息可以由终端的麦克风进行，其具
20 体过程与获取语音的过程类似，在此不作赘述。

需要说明的是，获取的环境信息类型可以由以由技术人员默认设置，还可以由用户在使用过程中调整。对于不同类型的环境信息类型，其具体的判断过程不同，详见步骤 203 的描述。

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
25 步骤 204，若是，则执行步骤 206;

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
30 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。

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：

方式一、当前的环境信息为用户的操作指令；

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，
5 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。

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，
10 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。

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
15 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。

方式二、当前的环境信息为终端自身的硬件信息；

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。
20

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
25 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。

方式三、当前的环境信息为当前的噪声信息；

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
30 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。

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，该实现方式的采用可以由技术人员默认设置，还可以由用户在使用过程中调整。

步骤 204：将语音信息识别为文字信息；

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。

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。

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，然后将该识别结果提供给用户即可。

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。

以下将从移动终端和 PC 终端两种场景下描述本发明方案。

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。

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。

步骤 205: 将该文字信息显示给用户;

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。

需要说明的是, 该文字信息显示时, 可以以独立的信息形式显示在指定区域, 还可以覆盖在语音信息的显示区域上进行显示, 该原地显示提供更直观的视觉效果。

步骤 206: 播放该语音信息。

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。

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。

实施例 3

参见图 5, 本发明实施例提供了一种终端, 所述终端包括:

获取模块 301, 用于接收用户输入的语音信息;

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;

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

其中, 识别模块 303 具体包括发送单元和接收单元;

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;

接收单元, 用于接收语音识别服务器识别得到的文字信息。

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。

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, 获取单元 3021, 用于获取用户的操作指令;

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。

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

获取单元 3021, 用于获取终端自身的硬件信息;

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。

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

获取单元 3021, 用于获取当前的噪声信息;

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。

10 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。

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
15 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。

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。

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
20 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。

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
25 保护范围之内。

权 利 要 求 书

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；

5 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。

2、如权利要求 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，具体包括：

10 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；

接收所述语音识别服务器识别得到的文字信息。

15 3、如权利要求 1 所述的方法，其特征在于，所述将语音信息识别为文字信息，具体包括：

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。

20 4、如权利要求 1-3 任一权利要求所述的方法，其特征在于，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：

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；

25 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。

5、如权利要求 1-3 任一权利要求所述的方法，其特征在于，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：

30 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。

- 5 6、如权利要求 1-3 权利要求所述的方法, 其特征在于, 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:

判断所述当前的噪声信息是否在预设的范围内;

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;

- 10 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。

7、如权利要求 1-3 任一权利要求所述的方法, 其特征在于, 所述终端为移动终端或 PC 终端。

15

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;

- 20 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。

9、如权利要求 8 所述的终端, 其特征在于, 所述识别模块具体包括发送单元和接收单元;

- 25 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;

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。

- 30 10、如权利要求 8 所述的终端, 其特征在于, 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。

11、如权利要求 8-10 任一权利要求所述的终端，其特征在于，所述判断模块具体包括获取单元和判断单元，

5 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。

10 12、如权利要求 8-10 所述的终端，其特征在于，所述判断模块具体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
15 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 8-10 所述的终端，其特征在于，所述判断模块具体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；

20 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。

14、如权利要求 8-10 任一权利要求所述的终端，其特征在于，所述终端为
25 移动终端或 PC 终端。

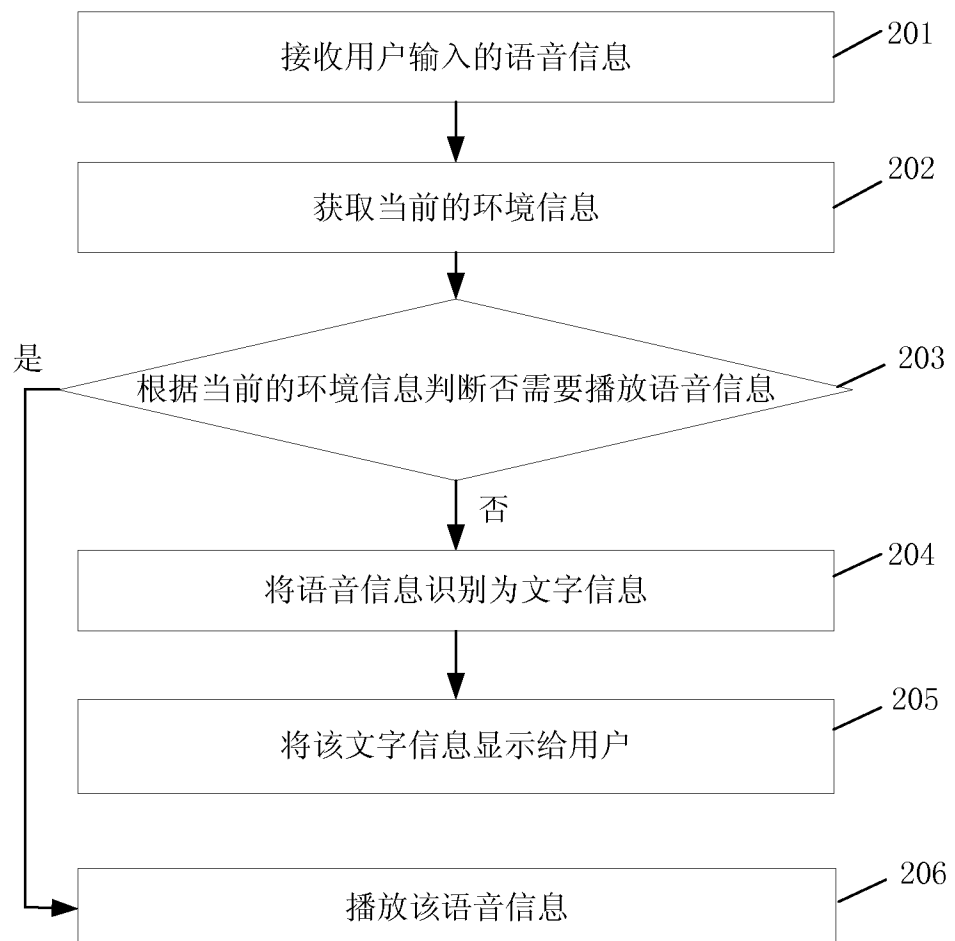
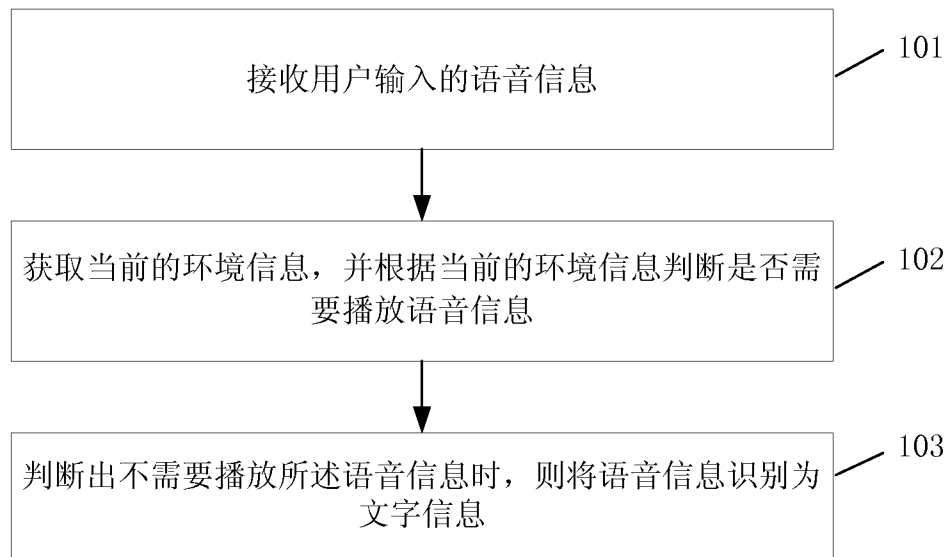




图 3

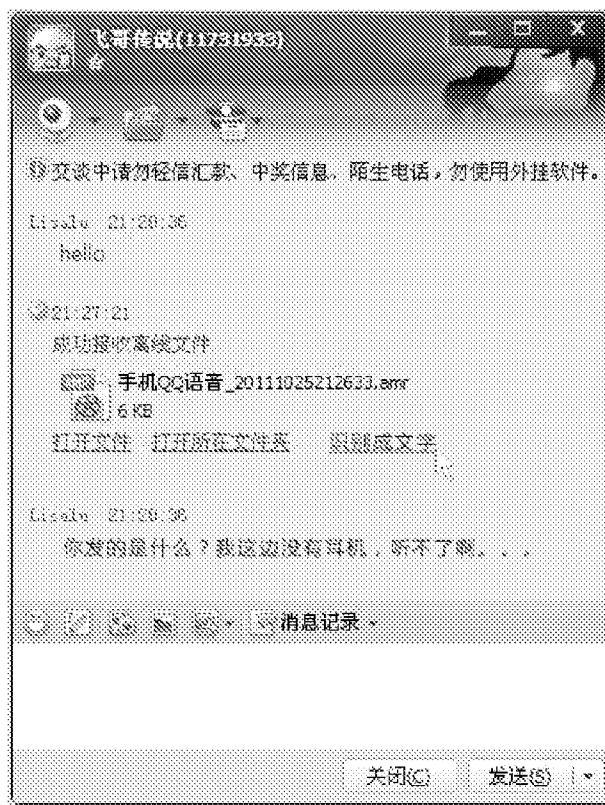


图 4

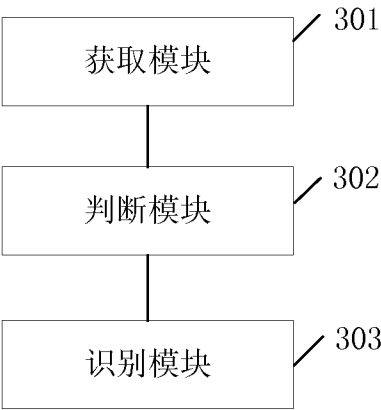


图 5

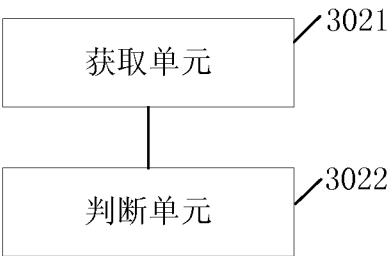


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/072086

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L 12/-, H04M 3/-, H04M 7/-, G10L 15/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: instant communication voice media word chose client

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101079836 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 28 November 2007 (28. 11. 2007) description, pages 1 to 3 and figure 1	1-5, 7-12, 14
A		6, 13
A	CN 101098309 A (FOSHAN SHUNDE DISTRICT SHUNDA COMPUTER FACTORY CO., LTD.) 02 January 2008 (02. 01. 2008) the whole document	1-14
A	CN 101453611 A (SIM INFORMATION TECHNOLOGY SHANGHAI GROUP LTD.) 10 June 2009 (10. 06. 2009) the whole document	1-14
A	US 2011137660 A1 (SKYPE LTD.) 09 June 2011 (09. 06. 2011) the whole document	1-14

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 23 April 2013 (23. 04. 2013)	Date of mailing of the international search report 06 June 2013 (06. 06. 2013)
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer MA, Shanshan Telephone No. (86-10) 82245571

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2013/072086

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101079836 A	28. 11. 2007	None	
CN 101098309 A	02. 01. 2008	None	
CN 101453611 A	10. 06. 2009	None	
US 2011137660 A1	09. 06. 2011	WO 2011070033 A1	16. 06. 2011
		GB 2476041 A	15. 06. 2011

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/072086

CLASSIFICATION OF SUBJECT MATTER

G10L 15/26 (2006.01) i
H04L 12/58 (2006.01) i
H04M 7/00 (2006.01) i

国际检索报告

国际申请号

PCT/CN2013/072086

A. 主题的分类

参见附加页

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L12/-,H04M3/-,H04M7/-,G10L 15/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT,CNKI,WPI,EPODOC: 即时 通信 通讯 语音 文字 识别 信息 instant communication voice media word chose client

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN101079836A (腾讯科技(深圳)有限公司) 28.11 月 2007 (28.11.2007) 说明书第 1-3 页, 附图 1	1-5, 7-12, 14
A		6, 13
A	CN101098309A (佛山市顺德区顺达电脑厂有限公司, 神达电脑股份有限公司) 02.1 月 2008 (02.01.2008) 全文	1-14
A	CN101453611A (希姆通信息技术(上海)有限公司) 10.6 月 2009 (10.06.2009) 全文	1-14
A	US2011137660A1 (SKYPE 公司) 09.6 月 2011 (09.06.2011) 全文	1-14

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期
23.4 月 2013 (23.04.2013)

国际检索报告邮寄日期
06.6 月 2013 (06.06.2013)

ISA/CN 的名称和邮寄地址:
中华人民共和国国家知识产权局
中国北京市海淀区蓟门桥西土城路 6 号 100088
传真号: (86-10)62019451

授权官员

马珊珊
电话号码: (86-10) **82245571**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/072086

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN101079836A	28.11.2007	无	
CN101098309A	02.01.2008	无	
CN101453611A	10.06.2009	无	
US2011137660A1	09.06.2011	WO2011070033A1	16.06.2011
		GB2476041A	15.06.2011

主题的分类

G10L 15/26(2006.01) i

H04L 12/58(2006.01)i

H04M 7/00(2006.01)i