

(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
2017 年 1 月 26 日 (26.01.2017)



(10) 国际公布号
WO 2017/012242 A1

- (51) 国际专利分类号:
G10L 15/22 (2006.01) G10L 15/20 (2006.01)
G10L 15/26 (2006.01)
- (21) 国际申请号: PCT/CN2015/096596
- (22) 国际申请日: 2015 年 12 月 7 日 (07.12.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510435887.1 2015 年 7 月 22 日 (22.07.2015) CN
- (71) 申请人: 百度在线网络技术(北京)有限公司
(BAIDU ONLINE NETWORK TECHNOLOGY
(BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
- (72) 发明人: 谢延 (XIE, Yan); 中国北京市海淀区上地
十街 10 号百度大厦三层, Beijing 100085 (CN)。

(74) 代理人: 北京清亦华知识产权代理事务所(普通
合伙) (TSINGYIHUA INTELLECTUAL PROP-
ERTY LLC); 中国北京市海淀区清华园清华大学照
澜院商业楼 301 室, Beijing 100084 (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, IR, IS,
JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ,
BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH,

[见续页]

(54) Title: VOICE RECOGNITION METHOD AND APPARATUS

(54) 发明名称: 语音识别方法和装置

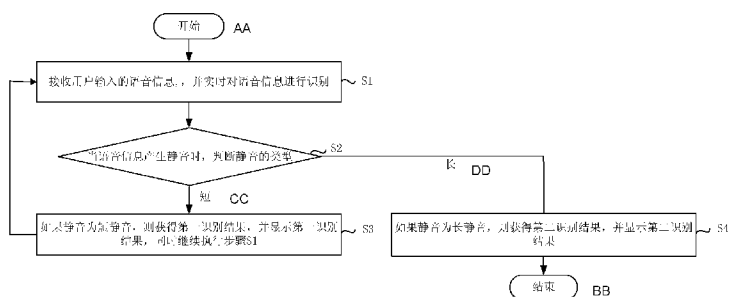


图 1

- S1 RECEIVING VOICE INFORMATION INPUT BY A USER AND RECOGNIZING THE VOICE INFORMATION IN REAL TIME
S2 WHEN THE VOICE INFORMATION PRODUCES A MUTE, DETERMINING THE TYPE OF THE MUTE
S3 IF THE MUTE IS A SHORT MUTE, OBTAINING A FIRST RECOGNITION RESULT, DISPLAYING THE FIRST RECOGNITION RESULT, AND AT THE SAME TIME, CONTINUOUSLY PERFORMING STEP S1
S4 IF THE MUTE IS A LONG MUTE, OBTAINING A SECOND RECOGNITION RESULT AND DISPLAYING THE SECOND RECOGNITION RESULT
AA START
BB END
CC SHORT
DD LONG

(57) Abstract: A voice recognition method and apparatus. The method comprises: receiving voice information input by a user and recognizing the voice information in real time (S1); when the voice information produces mute, determining the type of the mute (S2); if the mute is a short mute, obtaining a first recognition result, displaying the first recognition result, and at the same time, continuously performing step S1 (S3); and if the mute is a long mute, obtaining a second recognition result and displaying the second recognition result (S4). According to the voice recognition method and apparatus, by recognizing voice information input by a user in real time, the waiting time of the user can be shortened, and user usage experience is improved.

(57) 摘要: 一种语音识别方法和装置, 其中, 该方法包括: 接收用户输入的语音信息, 并实时对语音信息进行识别 (S1); 当语音信息产生静音时, 判断静音的类型 (S2); 如果静音为短静音, 则获得第一识别结果, 并显示第一识别结果, 同时继续执行步骤 S1 (S3); 以及如果静音为长静音, 则获得第二识别结果, 并显示第二识别结果 (S4)。该语音识别方法和装置, 通过实时对用户输入的语音信息进行识别, 能够降低用户等待时间, 提升用户使用体验。



WO 2017/012242 A1



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, KM, ML, MR, NE, SN, TD,
TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

语音识别方法和装置

相关申请的交叉引用

5 本申请要求百度在线网络技术（北京）有限公司于 2015 年 7 月 22 日提交的、发明名称为“语音识别方法和装置”的、中国专利申请号“201510435887.1”的优先权。

技术领域

本发明涉及语音识别技术领域，尤其涉及一种语音识别方法和装置。

10

背景技术

随着科技的不断进步，语音识别技术的应用也越来越广泛，例如工业、家电、通信、汽车电子、医疗、家庭服务、消费电子产品等领域，都会应用到语音识别技术。目前，语音识别系统主要通过接收用户输入的语音，对语音进行识别，从而获得语音识别结果。其中，语音搜索类产品不仅可以对用户输入的语音进行识别，还可根据语音识别结果向搜索服务器发送搜索请求，进一步获取搜索结果。

但是，有时候用户输入语音时，内容可能很多，则需要在用户输入语音结束后，等待很长时间才能获取到识别结果。如果是语音搜索类产品，则需要先等待获得识别结果的过程，再等待获取搜索结果的过程，等待时间长，导致用户体验降低。另外，在噪声环境中，由于噪声干扰，有可能出现检测不到语音结束点或者识别结果不准确的情况。

20

发明内容

本发明旨在至少在一定程度上解决相关技术中的技术问题之一。为此，本发明的一个目的在于提出一种语音识别方法，该方法能够降低用户等待时间，提升用户使用体验。

25 本发明的第二个目的在于提出一种语音识别装置。

本发明的第三个目的在于提出一种设备。

本发明的第四个目的在于提出一种非易失性计算机存储介质。

为了实现上述目的，本发明第一方面实施例提出了一种语音识别方法，包括以下步骤：
S1、接收用户输入的语音信息，并实时对所述语音信息进行识别；S2、当所述语音信息产生静音时，判断所述静音的类型；S3、如果所述静音为短静音，则获得第一识别结果，并显示所述第一识别结果，同时继续执行步骤 S1；以及 S4、如果所述静音为长静音，则获得第二识别结果，并显示所述第二识别结果。

30

本发明实施例的语音识别方法，通过接收用户输入的语音信息，并实时对语音信息进行识别，当语音信息产生静音时，判断静音的类型，如果静音为短静音，则获得第一识别结果，并显示第一识别结果，同时继续接收用户输入的语音信息，如果静音为长静音，则获得第二识别结果，并显示第二识别结果，能够有效地降低用户等待时间，提升用户体验。

本发明第二方面实施例提出了一种语音识别装置，包括：接收模块，用于接收用户输入的语音信息，并实时对所述语音信息进行识别；判断模块，用于当所述语音信息产生静音时，判断所述静音的类型；第一识别模块，用于当所述静音为短静音时，获得第一识别结果，并显示所述第一识别结果，同时所述接收模块继续接收搜索用户输入的语音信息；第二识别模块，用于当所述静音为长静音时，获得第二识别结果，并显示所述第二识别结果。

本发明实施例的语音识别装置，通过接收用户输入的语音信息，并实时对语音信息进行识别，当语音信息产生静音时，判断静音的类型，如果静音为短静音，则获得第一识别结果，并显示第一识别结果，同时继续接收用户输入的语音信息，如果静音为长静音，则获得第二识别结果，并显示第二识别结果，能够有效地降低用户等待时间，提升用户体验。

本发明第三方面实施例提供了一种设备，包括：一个或者多个处理器；存储器；一个或者多个程序，所述一个或者多个程序存储在所述存储器中，当被所述一个或者多个处理器执行时，执行本发明第一方面实施例的语音识别方法。

本发明第四方面实施例提供了一种非易失性计算机存储介质，所述计算机存储介质存储有一个或者多个程序，当所述一个或者多个程序被一个设备执行时，使得所述设备执行以本发明第一方面实施例的语音识别方法。

附图说明

图 1 是根据本发明一个实施例的语音识别方法的流程图；

图 2 是根据本发明一个具体实施例的语音识别方法的流程图；

图 3 是根据本发明一个具体实施例的初始化界面效果示意图；

图 4 是根据本发明一个具体实施例的提示界面效果示意图；

图 5 是根据本发明一个具体实施例的接收用户输入的语音信息界面效果示意图；

图 6 是根据本发明一个具体实施例的显示识别结果界面效果示意图一；

图 7 是根据本发明一个具体实施例的显示识别结果界面效果示意图二；

图 8 是根据本发明一个具体实施例的显示识别结果界面效果示意图三；

图 9 是根据本发明一个具体实施例的根据识别结果进行搜索的界面效果示意图；

图 10 是根据本发明一个具体实施例的显示搜索结果的界面效果示意图；

图 11 是根据本发明一个具体实施例的根据识别结果进行搜索的界面效果示意图一；

图 12 是根据本发明一个具体实施例的根据识别结果进行搜索的界面效果示意图二；

5 图 13 是根据本发明一个具体实施例的根据识别结果进行搜索的界面效果示意图三；

图 14 是根据本发明一个具体实施例的根据识别结果进行搜索的界面效果示意图四。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同
10 或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

下面参考附图描述本发明实施例的语音识别方法和装置。

图 1 是根据本发明一个实施例的语音识别方法的流程图。

如图 1 所示，语音识别方法可包括：

15 S1、接收用户输入的语音信息，并实时对语音信息进行识别。

其中，语音信息可以为词组，也可以为短句。

S2、当语音信息产生静音时，判断静音的类型。

在本发明的实施例中，为解决在噪声环境中，静音检测不准确的问题，可根据尾点检测算法检测出静音，并判断静音的类型。其中，静音的类型可包括长静音和短静音。短静音为用户输入语音信息的短暂停顿，而长静音则为用户输入语音信息的结束点（尾点）。
20

具体地，可先在不同环境下采集语音样本，并训练尾点检测模型。然后在对语音信息进行识别时，可通过尾点检测模型判断静音的类型，在噪声环境下能够准确地判断出静音的类型，提高了抗噪性和准确率。相对于本地的尾点检测算法，服务器端的尾点检测算法具有更强大的计算能力，可不断地对尾点检测模型进行优化。在本发明一个实施例中，在对语音信息识别的过程中，可先通过本地的尾点检测算法进行检测，如果无法检测出语音信息的结束点，则再通过服务器端的尾点检测算法进行检测。
25

S3、如果静音为短静音，则获得第一识别结果，并显示第一识别结果，同时继续执行步骤 S1。

具体地，在用户输入语音信息开始时，可实时地对语音信息进行识别，当出现静音时，
30 如果当前出现的静音为短静音，即用户输入语音信息的短暂停顿，则可获得第一识别结果，然后将第一识别结果显示在客户端的屏幕上，反馈给用户。其中，第一识别结果可以为输入语音信息开始至短静音之间的内容，也可以是两个短静音之间的内容。与此同时，用户

还在继续输入语音信息。也就是说，识别过程与接收语音信息过程同步进行，即两个单独且互不干扰的线程并行处理，减少了用户等待的时间。用户在输入语音信息的同时，已经在客户端的屏幕上显示出了一部分的识别结果，由于短静音时间很短，因此在客户端的屏幕上显示的效果相当于用户一边输入语音信息，同时动态地连续不断地显示出识别结果，

5 解决了传统的语音识别中，等待用户输入语音信息结束后，再对语音信息进行整体识别所带来的等待时间过长的的问题，提升了用户使用体验。

此外，在获得第一识别结果之后，还可将第一识别结果作为关键词进行搜索，并获取第一搜索结果。例如：识别系统为语音搜索系统时，可根据实时识别出的识别结果进行搜索。

10 S4、如果静音为长静音，则获得第二识别结果，并显示第二识别结果。

具体地，如果当前出现的静音为长静音，即用户输入语音信息结束，则可获得第二识别结果，然后将第二识别结果显示在客户端的屏幕上，反馈给用户。其中，第二识别结果可以是最后一个短静音与长静音之间的内容，如果用户输入的语音信息没有短静音，则第二识别结果可以为输入语音信息开始与长静音之间的内容。举例来说，实时地对用户输入的语音信息进行识别，当客户端的屏幕显示第一识别结果时，同时还在接收用户输入的语音信息，并实时地对语音信息识别，从而达到减少用户等待时间的目的。

15

另外，还可将第一识别结果与第二识别结果进行对比。若第一识别结果与第二识别结果一致，则可将第一搜索结果作为最终搜索结果。具体地，第一识别结果为语音信息产生短静音时对应的识别结果，第二识别结果为语音信息产生长静音时对应的识别结果。而获得第二识别结果通常需要一个长静音，而在判断当前静音是否为长静音时，已经将其作为短静音并进行了语音识别，获得了第一识别结果，并获取了对应的第一搜索结果。当确定该静音为长静音后，如果第一识别结果和第二识别结果一致，则可直接将第一搜索结果作为最终的搜索结果，而无需将第二识别结果作为关键词再次进行搜索，从而节省了用户等待的时间。

20

25 若第一识别结果与第二识别结果不一致，则可将第一识别结果与第二识别结果进行拼接，生成最终的识别结果，并将识别结果作为关键词进行搜索，以获取最终的搜索结果。

在确定最终的搜索结果后，可在客户端的屏幕显示搜索结果，以反馈给用户。

本发明实施例的语音识别方法，通过接收用户输入的语音信息，并实时对语音信息进行识别，当语音信息产生静音时，判断静音的类型，如果静音为短静音，则获得第一识别结果，并显示第一识别结果，同时继续接收用户输入的语音信息，如果静音为长静音，则获得第二识别结果，并显示第二识别结果，能够有效地降低用户等待时间，提升用户使用体验。

30

图 2 是根据本发明一个具体实施例的语音识别方法的流程图，本实施例以搜索 APP 为例进行详细描述。

如图 2 所示，语音识别方法可包括：

S201，开启搜索 APP，并进行初始化。

5 如图 3 所示，在开启终端中的搜索 APP 时，可对运行环境进行初始化。

S202，显示提示界面。

在初始化结束后，可显示如图 4 所示的提示界面。

S203，接收用户输入的语音信息，并实时对语音信息进行识别。

10 当检测到有用户输入语音信息时，如图 5 所示，可在界面中显示如“倾听中”字样，表示正在接收用户输入的语音信息，与此同时正在对输入的语音信息进行识别。

S204，当产生短静音时，获得并显示第一识别结果。

例如，用户输入的语音信息为“百度语音提供技术”，而输入到“百度”时，检测到一个短静音，则可获得并显示对应的识别结果“百度”，如图 6 所示。与此同时，还在接收用户输入的语音信息，且实时地对语音信息进行识别。依此类推，当用户输入到“语音”
15 时，又检测到一个短静音，此时可获得并显示对应的识别结果“语音”，如图 7 所示。

此外，在识别出“百度”的同时，还可以“百度”为关键词，向搜索服务器发送搜索请求，获得“百度”对应的搜索结果。以此类推，在识别出“语音”的同时，还可以“百度语音”为关键词，向搜索服务器发送搜索请求，获得“百度语音”对应的搜索结果。

20 此处检测短静音和长静音使用的方法为尾点检测算法，与上一实施例中的描述一致，故此处不赘述。

S205，当产生长静音时，显示第二识别结果。

例如：当用户输入语音信息“百度语音提供技术”结束时，可检测到产生长静音，则可获得并显示对应的识别结果“提供技术”。由于“百度”、“语音”、“提供技术”是先后显示的，且时间间隔很短，则其效果相当于用户一边输入语音信息，一边连续不断地
25 在客户端的屏幕上显示出识别结果，最终显示出“百度语音提供技术”，如图 8 所示。

S206，将第一识别结果和第二识别结果进行拼接，以生成搜索词，并进行搜索。

在识别结束后，可将每段识别结果进行拼接，生成关键词“百度语音提供技术”，并向搜索服务器发送搜索请求。

S207，获得搜索词对应的搜索结果，并显示搜索结果。

30 具体地，如图 9 所示，在根据关键词“百度语音提供技术”进行搜索时，界面中的状态可显示为“处理中”。然后，在通过搜索服务器获得“百度语音提供技术”对应的搜索结果后，如图 10 所示，显示该搜索结果。

本发明实施例的语音识别方法，通过尾点检测算法对用户输入的语音信息进行分段，能够准确地判断出用户输入的语音信息的暂停点或结束点，提升了语音识别的抗噪性和准确性；通过实时地对语音信息进行识别，可在用户输入语音信息的同时即可显示出已识别的部分，减少了用户等待的时间；通过将识别过程和搜索过程并行处理，降低了整个语音识别搜索系统的响应时间，进而提高了用户使用体验。

为实现上述目的，本发明还提出一种语音识别装置。

图 11 是根据本发明一个实施例的语音识别装置的结构示意图一。

如图 11 所示，该语音识别装置可包括：接收模块 110、判断模块 120、第一识别模块 130 和第二识别模块 140。

其中，接收模块 110 用于接收用户输入的语音信息，并实时对语音信息进行识别。

其中，语音信息可以为词组，也可以为短句。

判断模块 120 用于当语音信息产生静音时，判断静音的类型。

在本发明的实施例中，为解决在噪声环境中，静音检测不准确的问题，判断模块 120 可根据尾点检测算法检测出静音，并判断静音的类型。其中，静音的类型可包括长静音和短静音。短静音为用户输入语音信息的短暂停顿，而长静音则为用户输入语音信息的结束点（尾点）。

具体地，可先在不同环境下采集语音样本，并训练尾点检测模型。然后在对语音信息进行识别时，可通过尾点检测模型判断静音的类型，在噪声环境下能够准确地判断出静音的类型，提高了抗噪性和准确率。相对于本地的尾点检测算法，服务器端的尾点检测算法具有更强大的计算能力，可不断地对尾点检测模型进行优化。在本发明一个实施例中，在对语音信息识别的过程中，可先通过本地的尾点检测算法进行检测，如果无法检测出语音信息的结束点，则再通过服务器端的尾点检测算法进行检测。

第一识别模块 130 用于当静音为短静音时，获得第一识别结果，并显示第一识别结果，同时接收模块继续接收搜索用户输入的语音信息。

具体地，在用户输入语音信息开始时，可实时地对语音信息进行识别，当出现静音时，如果当前出现的静音为短静音，即用户输入语音信息的短暂停顿，则第一识别模块 130 可获得第一识别结果，然后将第一识别结果显示在客户端的屏幕上，反馈给用户。其中，第一识别结果可以为输入语音信息开始至短静音之间的内容，也可以是两个短静音之间的内容。与此同时，用户还在继续输入语音信息。也就是说，识别过程与接收语音信息过程同步进行，即两个单独且互不干扰的线程并行处理，减少了用户等待的时间。用户在输入语音信息的同时，已经在客户端的屏幕上显示出了一部分的识别结果，由于短静音时间很短，因此在客户端的屏幕上显示的效果相当于用户一边输入语音信息，同时动态地连续不断地

显示出识别结果，解决了传统的语音识别中，等待用户输入语音信息结束后，再对语音信息进行整体识别所带来的等待时间过长的的问题，提升了用户使用体验。。

第二识别模块 140 用于当静音为长静音时，获得第二识别结果，并显示第二识别结果。

具体地，如果当前出现的静音为长静音，即用户输入语音信息结束，则第二识别模块 140 可获得第二识别结果，然后将第二识别结果显示在客户端的屏幕上，反馈给用户。其中，第二识别结果可以是最后一个短静音与长静音之间的内容，如果用户输入的语音信息没有短静音，则第二识别结果可以为输入语音信息开始与长静音之间的内容。举例来说，实时地对用户输入的语音信息进行识别，当客户端的屏幕显示第一识别结果时，同时还在接收用户输入的语音信息，并实时地对语音信息识别，从而达到减少用户等待时间的目的。

另外，如图 12 所示，本发明实施例的语音识别装置还可包括搜索模块 150。

搜索模块 150 用于在第一识别模块 130 获得第一识别结果之后，将第一识别结果作为关键词进行搜索，并获取第一搜索结果。例如：识别系统为语音搜索系统时，可根据实时识别出的识别结果进行搜索。

此外，如图 13 所示，本发明实施例的语音识别装置还可包括处理模块 160。

处理模块 160 用于将第一识别结果与第二识别结果进行对比，若第一识别结果与第二识别结果一致，则将第一搜索结果作为最终的搜索结果，以及若第一识别结果与第二识别结果不一致，则将第一识别结果与第二识别结果进行拼接，生成最终的识别结果，并将识别结果作为关键词进行搜索，以获取最终的搜索结果。

具体地，第一识别结果为语音信息产生短静音时对应的识别结果，第二识别结果为语音信息产生长静音时对应的识别结果。而获得第二识别结果通常需要一个长静音，而在判断当前静音是否为长静音时，已经将其作为短静音并进行了语音识别，获得了第一识别结果，并获取了对应的第一搜索结果。当确定该静音为长静音后，如果第一识别结果和第二识别结果一致，则可直接将第一搜索结果作为最终的搜索结果，而无需将第二识别结果作为关键词再次进行搜索，从而节省了用户等待的时间。

进一步地，如图 14 所示，本发明实施例的语音识别装置还可包括显示模块 170。

显示模块 170 用于在获取最终的搜索结果之后，显示搜索结果。

本发明实施例的语音识别装置，通过接收用户输入的语音信息，并实时对语音信息进行识别，当语音信息产生静音时，判断静音的类型，如果静音为短静音，则获得第一识别结果，并显示第一识别结果，同时继续接收用户输入的语音信息，如果静音为长静音，则获得第二识别结果，并显示第二识别结果，能够有效地降低用户等待时间，提升用户使用体验。

在本发明的描述中，需要理解的是，术语“中心”、“纵向”、“横向”、“长度”、

“宽度”、“厚度”、“上”、“下”、“前”、“后”、“左”、“右”、“竖直”、“水平”、“顶”、“底”、“内”、“外”、“顺时针”、“逆时针”、“轴向”、“径向”、“周向”等指示的方位或位置关系为基于附图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。

此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中，“多个”的含义是至少两个，例如两个，三个等，除非另有明确具体的限定。

在本发明中，除非另有明确的规定和限定，术语“安装”、“相连”、“连接”、“固定”等术语应做广义理解，例如，可以是固定连接，也可以是可拆卸连接，或成一体；可以是机械连接，也可以是电连接；可以是直接相连，也可以通过中间媒介间接相连，可以是两个元件内部的连通或两个元件的相互作用关系，除非另有明确的限定。对于本领域的普通技术人员而言，可以根据具体情况理解上述术语在本发明中的具体含义。

在本发明中，除非另有明确的规定和限定，第一特征在第二特征“上”或“下”可以是第一和第二特征直接接触，或第一和第二特征通过中间媒介间接接触。而且，第一特征在第二特征“之上”、“上方”和“上面”可是第一特征在第二特征正上方或斜上方，或仅仅表示第一特征水平高度高于第二特征。第一特征在第二特征“之下”、“下方”和“下面”可以是第一特征在第二特征正下方或斜下方，或仅仅表示第一特征水平高度小于第二特征。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

权利要求书

1、一种语音识别方法，其特征在于，包括以下步骤：

S1、接收用户输入的语音信息，并实时对所述语音信息进行识别；

5 S2、当所述语音信息产生静音时，判断所述静音的类型；

S3、如果所述静音为短静音，则获得第一识别结果，并显示所述第一识别结果，同时
继续执行步骤 S1；以及

S4、如果所述静音为长静音，则获得第二识别结果，并显示所述第二识别结果。

10 2、如权利要求 1 所述的方法，其特征在于，在获得所述第一识别结果之后，还包括：
将所述第一识别结果作为关键词进行搜索，并获取第一搜索结果。

3、如权利要求 1 或 2 所述的方法，其特征在于，还包括：

将所述第一识别结果与所述第二识别结果进行对比；

若所述第一识别结果与所述第二识别结果一致，则将所述第一搜索结果作为最终的搜索
搜索结果；

15 若所述第一识别结果与所述第二识别结果不一致，则将所述第一识别结果与所述第二
识别结果进行拼接，生成最终的识别结果，并将所述识别结果作为所述关键词进行搜索，
以获取最终的所述搜索结果。

4、如权利要求 3 所述的方法，其特征在于，在获取最终的所述搜索结果之后，还包括：
显示所述搜索结果。

20 5、如权利要求 1-4 任一项所述的方法，其特征在于，所述判断所述静音的类型，包括：
根据尾点检测算法判断所述静音的类型。

6、一种语音识别装置，其特征在于，包括：

接收模块，用于接收用户输入的语音信息，并实时对所述语音信息进行识别；

判断模块，用于当所述语音信息产生静音时，判断所述静音的类型；

25 第一识别模块，用于当所述静音为短静音时，获得第一识别结果，并显示所述第一识
别结果，同时所述接收模块继续接收搜索用户输入的语音信息；

第二识别模块，用于当所述静音为长静音时，获得第二识别结果，并显示所述第二识
别结果。

7、如权利要求 6 所述的装置，其特征在于，还包括：

30 搜索模块，用于在获得所述第一识别结果之后，将所述第一识别结果作为关键词进行
搜索，并获取第一搜索结果。

8、如权利要求 6 或 7 所述的装置，其特征在于，还包括：

处理模块，用于将所述第一识别结果与所述第二识别结果进行对比，若所述第一识别结果与所述第二识别结果一致，则将所述第一搜索结果作为最终的搜索结果，以及若所述第一识别结果与所述第二识别结果不一致，则将所述第一识别结果与所述第二识别结果进行拼接，生成最终的识别结果，并将所述识别结果作为所述关键词进行搜索，以获取最终的所述搜索结果。

9、如权利要求 8 所述的装置，其特征在于，还包括：

显示模块，用于在获取最终的所述搜索结果之后，显示所述搜索结果。

10、如权利要求 6-9 任一项所述的装置，其特征在于，所述判断判断模块，具体用于：根据尾点检测算法判断所述静音的类型。

11、一种设备，其特征在于，包括：

一个或者多个处理器；

存储器；

一个或者多个程序，所述一个或者多个程序存储在所述存储器中，当被所述一个或者多个处理器执行时，执行如权利要求 1-5 任一项所述的语音识别方法。

12、一种非易失性计算机存储介质，其特征在于，所述计算机存储介质存储有一个或者多个程序，当所述一个或者多个程序被一个设备执行时，使得所述设备执行如权利要求 1-5 任一项所述的语音识别方法。

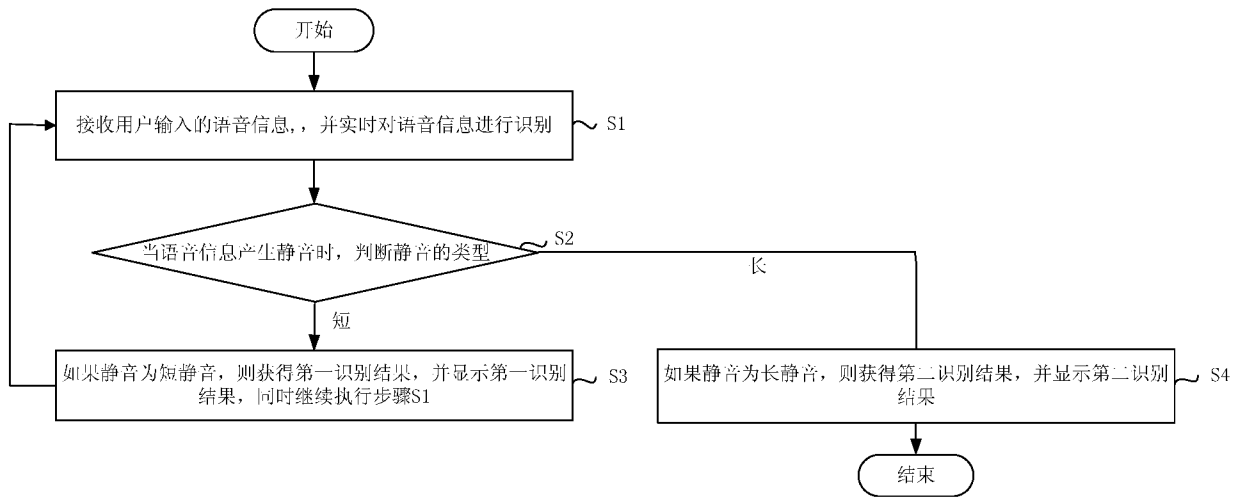


图 1

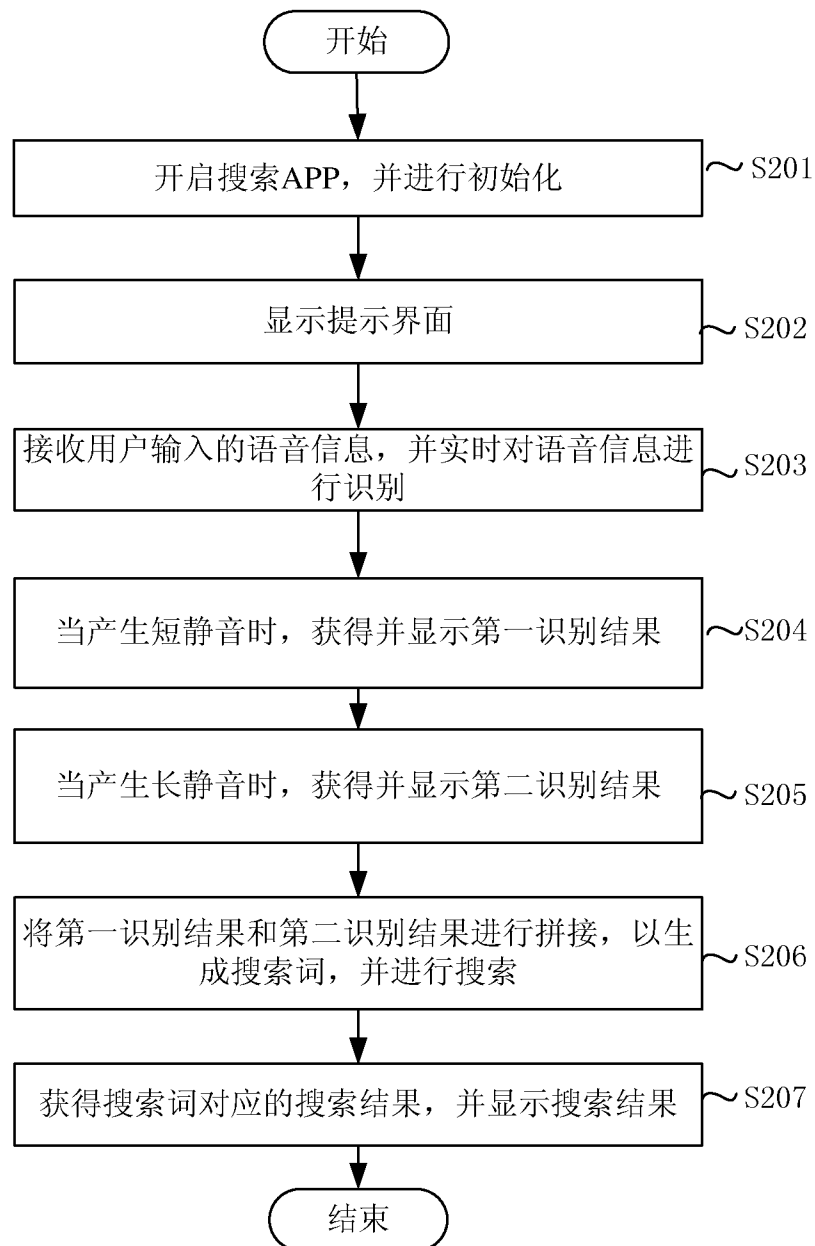


图 2

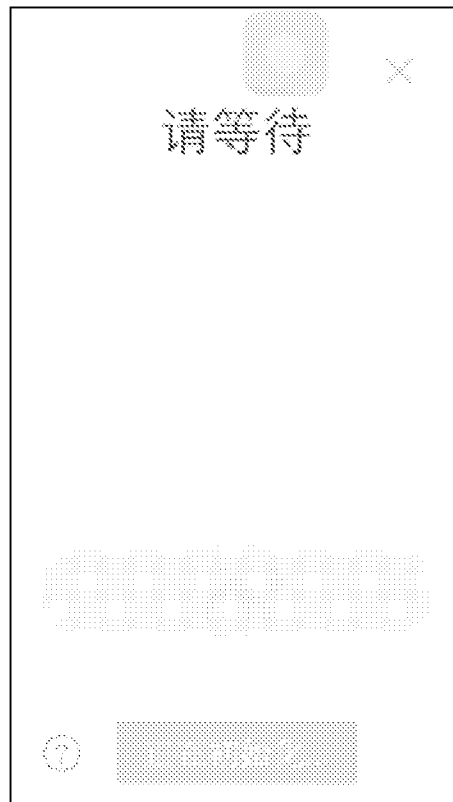


图 3

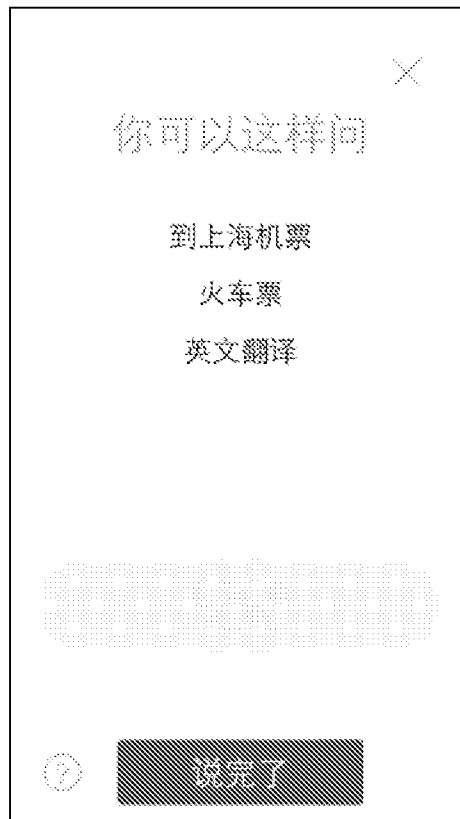


图 4

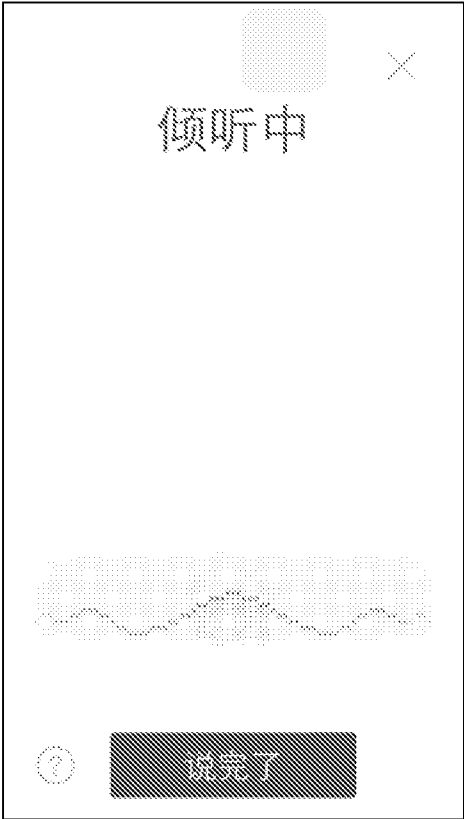


图 5

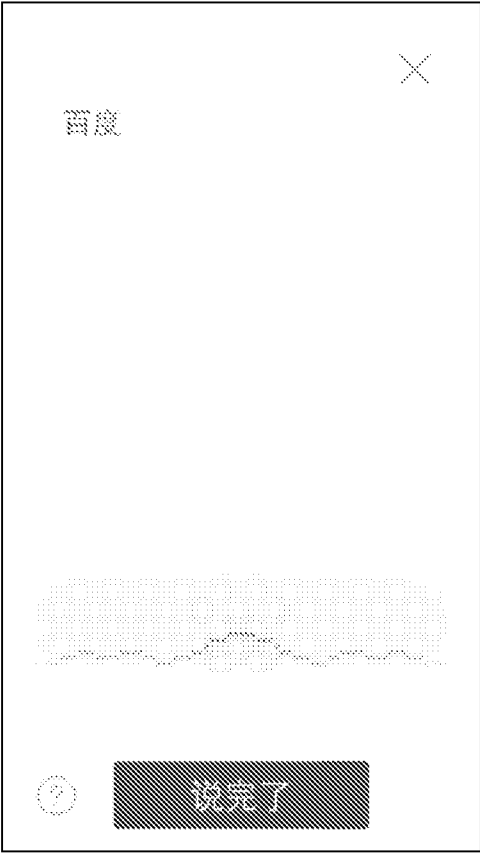


图 6

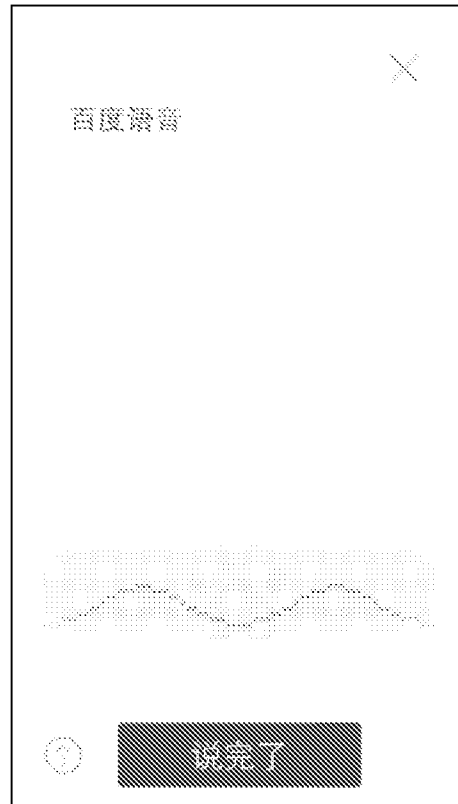


图 7

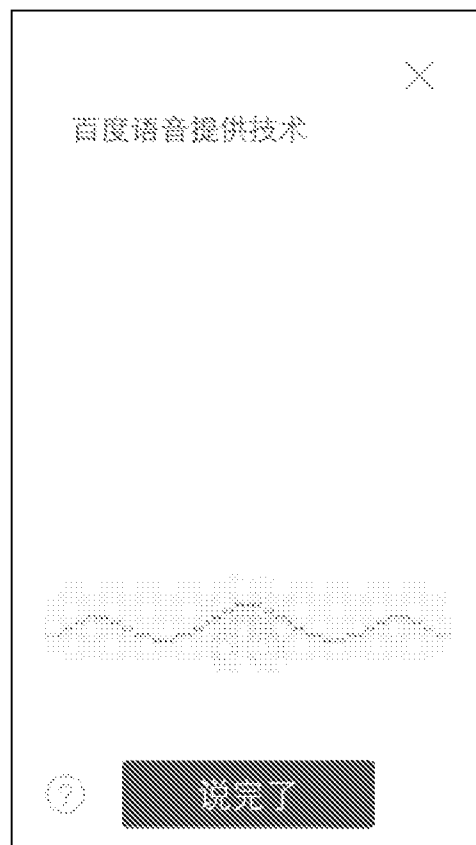


图 8

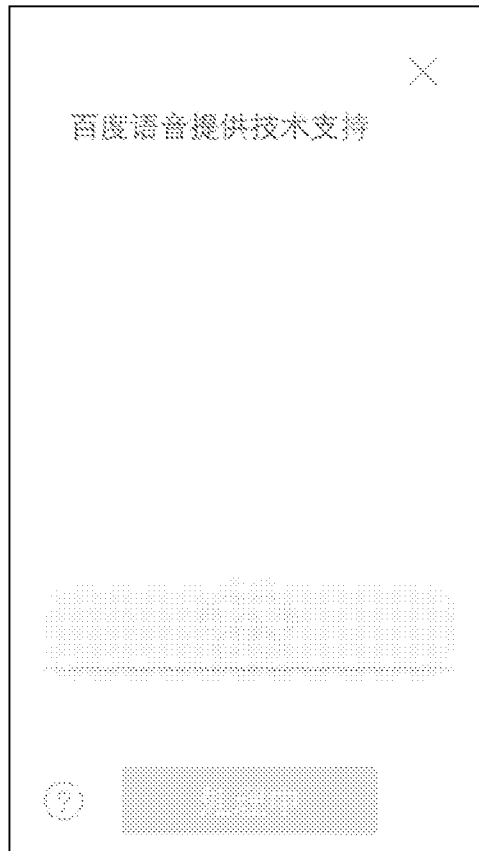


图 9



图 10

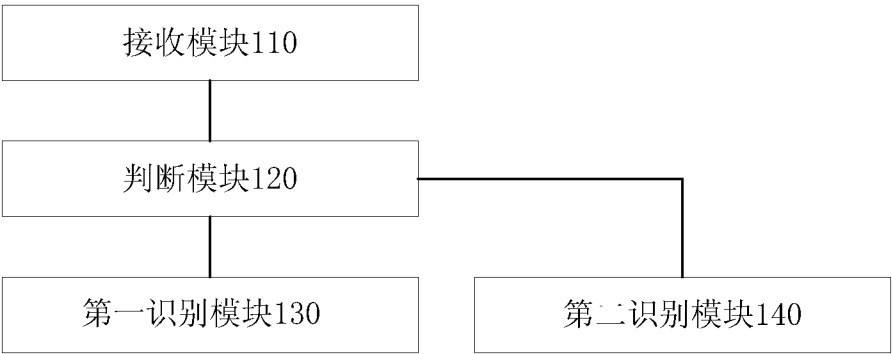


图 11

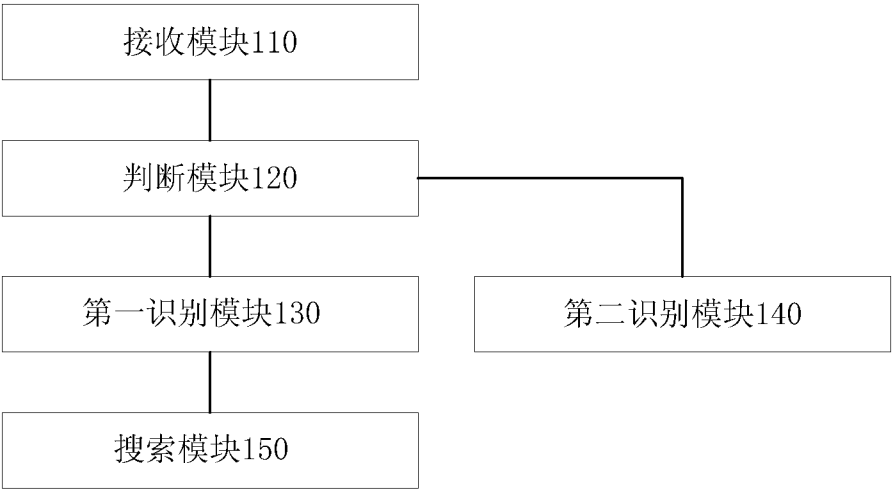


图 12

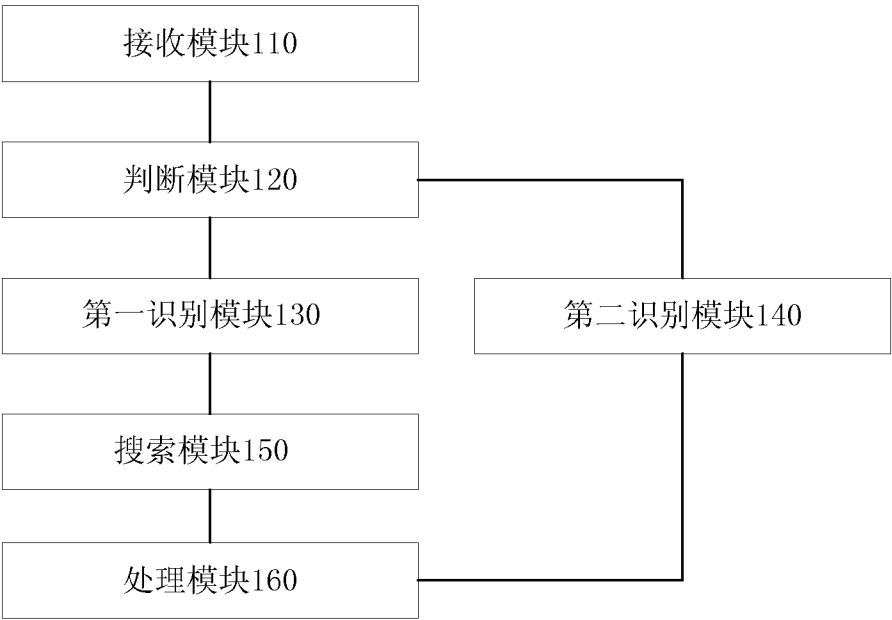


图 13

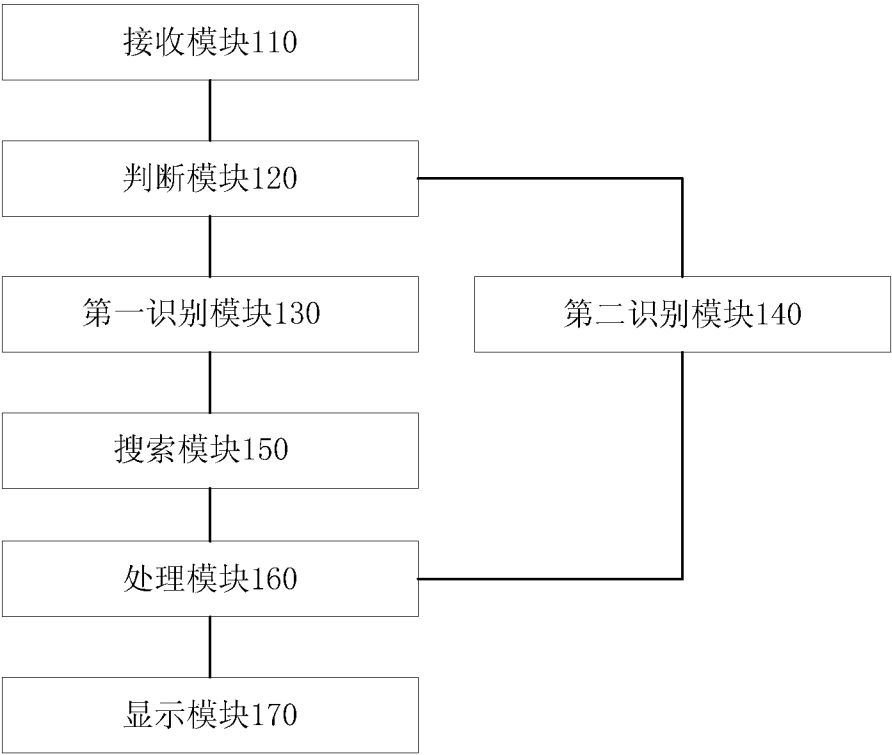


图 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/096596

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22 (2006.01) i, G10L 15/26 (2006.01) i, G10L 15/20 (2006.01) i
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: G10L 15; G10L; G06F; H04

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)

CNABS, CNKI, CNTXT, VEN: voice, speech, sound, audio, vocal+, acoustic+, recogni+, silen+, paus+, halt+, intermission, intermit+, soundless, break+, mute, short, long, length, time, duration, period, interval, research, +, retriev+, realtime, internet, reduce

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101042866 A (FUJITSU LTD.) 26 September 2007 (26.09.2007), description, page 15, the third paragraph to page 17, the third paragraph, and figures 5 and 6	1, 5, 6, 10, 11, 12
Y	CN 101042866 A (FUJITSU LTD.) 26 September 2007 (26.09.2007), description, page 15, the third paragraph to page 17, the third paragraph, and figures 5 and 6	2-4, 7-9
Y	CN 102280106 A (SAMSUNG ELECTRONICS CO., LTD. et al.) 14 December 2011 (14.12.2011), the abstract	2-4, 7-9
X	EP 0770986 A2 (DRAGON SYSTEMS INC.) 02 May 1997 (02.05.1997), description, page 3, line 15 to page 5, line 27, and figures 1-6	1, 5, 6, 10, 11, 12
X	JP 2003255978 A (SUZUKI MASAKAZU et al.) 10 September 2003 (10.09.2003), description, paragraphs [0010]-[0027], and figure 1	1, 5, 6, 10, 11, 12

☒ 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	
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“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
“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	“&” document member of the same patent family

Date of the actual completion of the international search
21 April 2016

Date of mailing of the international search report
06 May 2016

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
YANG, Shilin
Telephone No. (86-10) 62085717

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/096596

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 102231278 A (ANHUI USTC IFLYTEK CO., LTD.) 02 November 2011 (02.11.2011), the whole document	1-12
A	CN 103871401 A (LENOVO BEIJING CO., LTD.) 18 June 2014 (18.06.2014), the whole document	1-12
A	JP H07230293 A (SONY CORP.) 29 August 1995 (29.08.1995), the whole document	1-12

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/096596

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101042866 A	26 September 2007	JP 2007256482 A	04 October 2007
		JP 4906379 B2	28 March 2012
		US 7805304 B2	28 September 2010
		US 2007225982 A1	27 September 2007
		CN 101042866 B	04 July 2012
CN 102280106 A	14 December 2011	None	
EP 0770986 A2	02 May 1997	EP 0770986 A3	17 June 1998
JP 2003255978 A	10 September 2003	None	
CN 102231278 A	02 November 2011	CN 102231278 B	21 August 2013
CN 103871401 A	18 June 2014	US 2014163984 A1	12 June 2014
JPH 07230293 A	29 August 1995	None	

国际检索报告

国际申请号

PCT/CN2015/096596

A. 主题的分类

G10L 15/22(2006.01)i; G10L 15/26(2006.01)i; G10L 15/20(2006.01)i

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)

G10L 15; G10L; G06F; H04

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))

CNABS, CNKI, CNTXT, VEN, 互联网: 语音, 声音, 话音, 文语, 识别, 辨识, 静音, 静默, 静谧, 停顿, 暂停, 间歇, 间隙, 间断, 无声, 长, 短, 时间, 搜索, 检索, 查找, 查询, 实时, 即时, 缩短, 降低, 减少, 节省, 用时, voice, speech, sound, audio, vocal+, acoustic+, recogni+, silen+, paus+, halt+, intermission, intermit+, soundless, break+, mute, short, long, length, time, duration, period, interval, research+, retriev+, realtime

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	CN 101042866 A (富士通株式会社) 2007年 9月 26日 (2007 - 09 - 26) 说明书第15页第3段-第17页第3段, 附图5和6	1, 5, 6, 10, 11, 12
Y	CN 101042866 A (富士通株式会社) 2007年 9月 26日 (2007 - 09 - 26) 说明书第15页第3段-第17页第3段, 附图5和6	2-4, 7-9
Y	CN 102280106 A (三星电子株式会社 等) 2011年 12月 14日 (2011 - 12 - 14) 摘要	2-4, 7-9
X	EP 0770986 A2 (DRAGON SYSTEMS INC.) 1997年 5月 2日 (1997 - 05 - 02) 说明书第3页第15行-第5页第27行, 附图1-6	1, 5, 6, 10, 11, 12
X	JP 2003255978 A (SUZUKI MASAKAZU 等) 2003年 9月 10日 (2003 - 09 - 10) 说明书第[0010]段-第[0027]段, 附图1	1, 5, 6, 10, 11, 12
A	CN 102231278 A (安徽科大讯飞信息科技股份有限公司) 2011年 11月 2日 (2011 - 11 - 02) 全文	1-12

☒ 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

2016年 4月 21日

国际检索报告邮寄日期

2016年 5月 6日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
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

杨士林

电话号码 (86-10)62085717

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
A	CN 103871401 A (联想北京有限公司) 2014年 6月 18日 (2014 - 06 - 18) 全文	1-12
A	JP H07230293 A (SONY CORP.) 1995年 8月 29日 (1995 - 08 - 29) 全文	1-12

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/096596

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	101042866	A	2007年 9月 26日	JP	2007256482	A	2007年 10月 4日
				JP	4906379	B2	2012年 3月 28日
				US	7805304	B2	2010年 9月 28日
				US	2007225982	A1	2007年 9月 27日
				CN	101042866	B	2012年 7月 4日
CN	102280106	A	2011年 12月 14日	无			
EP	0770986	A2	1997年 5月 2日	EP	0770986	A3	1998年 6月 17日
JP	2003255978	A	2003年 9月 10日	无			
CN	102231278	A	2011年 11月 2日	CN	102231278	B	2013年 8月 21日
CN	103871401	A	2014年 6月 18日	US	2014163984	A1	2014年 6月 12日
JP	H07230293	A	1995年 8月 29日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)