

(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 11 月 30 日 (30.11.2017)

(10) 国际公布号
WO 2017/202016 A1

(51) 国际专利分类号:
G10L 15/22 (2006.01)

(72) 发明人: 袁斌(YUAN, Bin); 中国北京市海淀区上地十街10号百度大厦三层, Beijing 100085 (CN)。

(21) 国际申请号: PCT/CN2016/111367

(74) 代理人: 北京清亦华知识产权代理事务所(普通合伙)(TSINGYIHUA INTELLECTUAL PROPERTY 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)。

(22) 国际申请日: 2016 年 12 月 21 日 (21.12.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

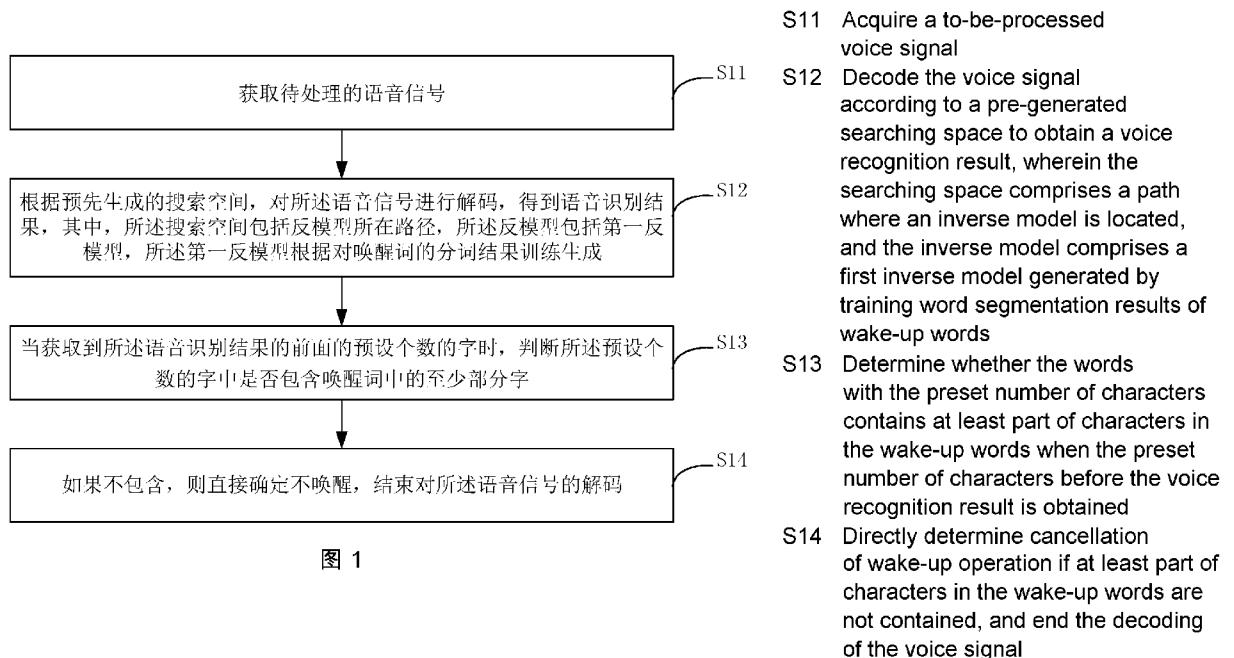
(30) 优先权:
201610357702.4 2016年5月26日 (26.05.2016) CN

(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, DJ, 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, KH, KN, KP, KR, KW, 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,

(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)。

(54) Title: VOICE WAKE-UP METHOD AND DEVICE

(54) 发明名称: 语音唤醒方法和装置



(57) Abstract: The invention discloses a voice wake-up method and a device. The voice wake-up method comprises steps as follows: acquiring a to-be-processed voice signal (S11); decoding the voice signal according to a pre-generated searching space to obtain a voice recognition result, wherein the searching space comprises a path where an inverse model is located, and the inverse model comprises a first inverse model generated by training word segmentation results of wake-up words (S12); determining whether the words with the preset number of characters contains at least part of characters in the wake-up words when the preset number of characters before the voice recognition result is obtained (S13); and directly determining cancellation of wake-up operation if at least part of characters

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, 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))。

in the wake-up words are not contained, and ending the decoding of the voice signal (S14). The method has the advantages that the mistaken wake-up rate can be reduced, and the power consumption can be reduced.

(57) 摘要: 一种语音唤醒方法和装置, 该语音唤醒方法包括: 获取待处理的语音信号 (S11); 根据预先生成的搜索空间, 对所述语音信号进行解码, 得到语音识别结果, 其中, 所述搜索空间包括反模型所在路径, 所述反模型包括第一反模型, 所述第一反模型根据对唤醒词的分词结果训练生成 (S12); 当获取到所述语音识别结果的前面的预设个数的字时, 判断所述预设个数的字中是否包含唤醒词中的至少部分字 (S13); 如果不包含, 则直接确定不唤醒, 结束对所述语音信号的解码 (S14)。该方法能够降低误唤醒率和降低功耗。

说明书

语音唤醒方法和装置

相关申请的交叉引用

本申请要求百度在线网络技术（北京）有限公司于 2016 年 5 月 26 日提交的、发明名称为“语音唤醒方法和装置”的、中国专利申请号“201610357702.4”的优先权。

技术领域

本申请涉及语音识别技术领域，尤其涉及一种语音唤醒方法和装置。

背景技术

语音唤醒技术是一种具有开关入口属性的功能。用户通过语音唤醒，可以发起人机交互的操作，即机器只有被用户所说的唤醒词唤醒，才会对用户接下来的语音指令进行识别。

相关技术中存在一些语音唤醒技术，但都存在的问题，比如误唤醒率高、抗噪能力差、需要全程联网、功耗高、唤醒词数量仅为一个、唤醒灵敏度低等。

发明内容

本申请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，本申请的一个目的在于提出一种语音唤醒方法。

本申请的另一个目的在于提出一种语音唤醒装置。

为达到上述目的，本申请第一方面实施例提出的语音唤醒方法，包括：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

为达到上述目的，本申请第二方面实施例提出的语音唤醒装置，包括：获取模块，用于获取待处理的语音信号；解码模块，用于根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；判断模块，用于当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；第一处理模块，用于如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本申请实施例至少在一定程度上具有如下技术效果之一：

通过采用根据唤醒词的分词结果训练得到的第一反模型，可以降低误唤醒率。

通过在语音识别结果的前面预设个数的字不包含唤醒词的至少部分内容时，直接确定不唤醒，并结束语音解码，可以降低功耗。

通过对语音信号进行音频处理，可以提高抗噪能力。

通过对语料的聚类结果进行训练生成反模型，可以减小该反模型的规模，从而可以应用在终端本地，以解决需要全程联网的问题。

通过设置多个唤醒词，可以实现通过任一个唤醒词的唤醒。

通过对唤醒词所在路径的加权处理，可以提高唤醒灵敏度。

通过在解码时直接结束异常路径的搜索，可以降低功耗。

通过在搜索空间中包括反模型和唤醒词串联的路径，可以在用户夹杂唤醒词的一句话中依然成功完成唤醒，提高唤醒精度。

本申请实施例还提出了一种终端，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本申请实施例还提出了一种非临时性计算机可读存储介质，当所述存储介质中的指令由终端的处理器执行时，使得终端能够执行一种方法，所述方法包括：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本申请实施例还提出了一种计算机程序产品，当所述计算机程序产品中的指令被处理器执行时，执行一种方法，所述方法包括：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本发明附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

附图说明

本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中

将变得明显和容易理解，其中：

图 1 是本申请一个实施例提出的语音唤醒方法的流程示意图；

图 2 是本申请另一个实施例提出的语音唤醒方法的流程示意图；

图 3 是本申请实施例中搜索空间的示意图；

图 4 是本申请一个实施例提出的语音唤醒装置的结构示意图；

图 5 是本申请一个实施例提出的语音唤醒装置的结构示意图。

具体实施方式

下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的模块或具有相同或类似功能的模块。

下面通过参考附图描述的实施例是示例性的，仅用于解释本申请，而不能理解为对本申请的限制。相反，本申请的实施例包括落入所附加权利要求书的精神和内涵范围内的所有变化、修改和等同物。

如上所示，相关技术中的语音唤醒技术存在的问题。本申请将基于如下思路分别解决上述技术问题。

(1) 构建一个新的反模型，该反模型是通过对唤醒词的分词结果进行训练后生成的。可以避免被唤醒词的部分内容唤醒，以解决误唤醒率高的问题。

(2) 对用户输入的语音信号进行音频处理。音频处理例如降噪、信号增强等处理，以解决抗噪能力差的问题。

(3) 构建一个新的反模型，该反模型是根据对语料的聚类结果进行训练生

成。可以减小该反模型的规模，从而可以应用在终端本地，以解决需要全程联网的问题。

(4) 根据语音识别结果的前几个字直接确定不唤醒。不需要等待全部语音解码完成，可以降低功耗。另外，在解码过程中，如果发现异常路径可以直接结束该路径的搜索，也可以降低功耗。

(5) 不限制唤醒词数量，可以是多个。

(6) 对唤醒词所在路径进行加权处理，可以使得唤醒词更易走到唤醒词所在路径，以提高唤醒灵敏度。

需要说明的是，虽然上述对应每个技术问题对其主要思路进行了说明，但是，为了解决技术问题，具体的技术方案不限于上述的主要思路，还可以与其他特征相互结合，这些不同技术特征之间的结合依然属于本申请的保护范围。

需要说明的是，虽然上述给出了几种要解决的技术问题，但是，本申请并不限于仅能解决上述技术问题，应用本申请给出的技术方案还可以解决的其他技术问题依然属于本申请的保护范围。

需要说明的是，本申请的每个实施例不限于能够全部完美解决所有的技术问题，而在至少在一定程度上解决至少一个技术问题。

需要说明的是，虽然上述给出了本申请的主要思路，以及后续实施例会对一些特别点进行说明。但是，本申请的创新点并不限于上述的主要思路及特别点所涉及的内容，并不排除本申请中一些并未特殊说明的内容依然可能会包含

本申请的创新点。

可以理解的是，虽然上述进行了一些说明，但依然不排除其他可能方案，因此，与后续本申请给出的实施例相同、相似、等同等情况的技术方案依然属于本申请的保护范围。

下面将结合具体实施例对本申请的技术方案进行说明。

本申请的语音唤醒技术可以具体应用于离线场景，即在终端本地应用。当然，可以理解的是，本申请的语音唤醒技术也可以应用在服务端，以实现在线语音唤醒。

本申请中涉及的终端可以是移动终端、车载终端、机载终端、桌面电脑等各种能够应用语音唤醒技术的终端。

图 1 是本申请一个实施例提出的语音唤醒方法的流程示意图。

本实施例可以至少在一定程度上解决误唤醒率高和功耗高的问题。

如图 1 所示，本实施例的流程包括：

S11：获取待处理的语音信号。

初始的语音信号是由用户输入的。

本实施例中，为了提高抗噪性，可以对用户输入的语音信号进行音频处理，以得到待处理的语音信号。具体内容可以参见后续描述。

S12：根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模

型，所述第一反模型根据对唤醒词的分词结果训练生成。

其中，当获取到待处理的语音信号后，可以对该语音信号进行特征提取，获取声学特征，之后采用声学特征在搜索空间中进行最优路径搜索，将最优路径对应的文本确定为语音识别结果。

搜索空间中包括多条路径，具体可以包括：唤醒词所在路径、反模型所在路径，其中，反模型是用于在语音解码时将非唤醒词引导到反模型所在路径。

本实施例中，一个反模型称为第一反模型，该第一反模型是根据对唤醒词的分词结果进行训练生成的。例如，唤醒词是“百度一下”(唤醒词可设置)，则可以先对该唤醒词进行分词，分词原则可以是分为首字、当大于三个字时再两两分词，如分词结果是“百”、“百度”和“一下”，则可以将“百度”和“一下”作为训练数据参与反模型的训练，以得到第一反模型。

通过第一反模型，则“百+非‘度一下’”，或者，“百度+非‘一下’”将走到反模型所在路径，避免误唤醒。

S13：当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字。

其中，预设个数的字例如为3个，则当获取到语音信号对应的前3个字时，可以判断这前3个字中是否包含唤醒词中的至少部分字。例如，唤醒词是“百度一下”，则判断前3个字中是否是“百度一”，或者，判断前3个字中的后两个字是否是“百度”，或者，判断前3个字中的最后一个字是否是“百”。

S14：如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

例如，如果前 3 个字不是“百度一”，且，前 3 个字中的后两个字不是“百度”，且，前 3 个字中的最后一个字不是“百”，则直接确定不唤醒，不执行唤醒操作。

另外，当根据前 3 个字确定不唤醒时，还直接结束对语音信号的解码，即由于语音信号是一段信号，除了前 3 个字还可以有后续的一个或多个字。本实施例中，当识别出前 3 个字，并且根据前 3 个字确定不唤醒时，不需要再继续识别后续的字，而直接结束对这段语音信号的识别，从而可以降低功耗。

本实施例中，通过采用根据唤醒词的分词结果训练得到的第一反模型，可以降低误唤醒率。通过在语音识别结果的前面预设个数的字不包含唤醒词的至少部分内容时，直接确定不唤醒，并结束语音解码，可以降低功耗。

图 2 是本申请另一个实施例提出的语音唤醒方法的流程示意图。

本实施例可以至少在一定程度上解决误唤醒率高、功耗高、全程联网、抗噪能力差、唤醒词数量仅为一个、唤醒灵敏度差等问题。

可以理解的是，本实施例给出了一个相对解决较全面问题的技术方案，但是，本申请不限于本实施例的方案，还可以将解决不同技术问题的技术特征单独组成技术方案，或者，将不同技术特征进行其他方式的任意多个的组合以得到新的技术方案。

如图 2 所示，本实施例的流程包括：

S201：构图生成搜索空间。

其中，参见图 3，搜索空间可以包括多条路径，包括唤醒词所在路径 31 和反模型所在路径 32。

本实施例中，为了解决用户一句话中夹杂唤醒词的问题，搜索空间中还包括唤醒词和反模型串联所在的路径 33。

进一步的，该路径 33 上经过唤醒词后可以直接进入静音（SIL）状态，或者经过反模型再进入 SIL 状态。

本实施例中，为了提高唤醒灵敏度，还可以对唤醒词所在路径进行加权，即对唤醒词所在路径的权重在原始基础上进行增加，以使唤醒词更容易进入到唤醒词所在路径。

唤醒词可以设置为一个或多个。

上述的反模型可以包括第一反模型和第二反模型，具体可以由第一反模型和第二反模型并联而成，或者，加权并联等。

如上所示，第一反模型是对唤醒词的分词结果进行训练后生成的。

本实施例中，第二反模型不是直接对语料进行训练生成，而是对语料的聚类结果进行训练后生成的，以降低第二反模型的规模，更利于应用于终端本地。

具体的，可以利用一些常用的语音语料，对发音的音节进行聚类，如聚类成 26 个，则对应 26 个字，之后可以根据这 26 个字训练生成第二反模型，该第二反模型是一个精简模型。

通过上述流程可以完成搜索空间的生成。

可以理解的是，搜索空间可以是在语音唤醒之前预先生成的。

当需要进行语音唤醒时，还执行如下步骤：

S202：接收用户输入的语音信号。

例如，用户对终端说了一段话。

可以理解的是，在接收语音信号之前还可以进行一些初始化流程。例如，可以设置唤醒词、生成搜索空间，并对音频处理模块进行初始化等。

S203：对所述语音信号进行音频处理。

本实施例中的音频处理可以具体包括：降噪和语音增强处理。

其中，降噪又可以分为对低频噪声的降噪，以及对非低频噪声的降噪。

具体的，空调、车载发动机等噪声均属于低频噪声，可以采用高通滤波技术消除低频噪声。

背景的音乐或者人声等噪声属于非低频噪声，可以采用噪声抑制 (Noise Suppression , NS) 技术消除非低频噪声。

受到硬件麦克风不同增益的影响，有些语音信号的音量可能会处于一个较低的水平，因此，可以采用自动增益控制 (Automatic Gain Control , AGC) 技术进行语音增强，以将音量过低的音频信号的能量增强到可进行识别的水平。

S204 :对音频处理后的语音信号进行语音活动检测(voice activity detection , VAD)。

通过 VAD 可以得到待处理的语音信号。

S205 :根据搜索空间 ,对待处理的语音信号进行解码 ,得到语音识别结果。

在解码时 , 可以先对语音信号提取声学特征 , 再将声学特征在搜索空间中进行搜索 , 得到最优路径作为语音识别结果。具体的 , 搜索算法可以是 viterbi 搜索算法。

本实施例中 , 在搜索过程中如果发现某条路径是异常路径 , 则可以直接结束对该路径的搜索 , 这样可以缩小搜索范围、提高搜索效率 , 降低功耗。在异常路径的判断时 , 以声学模型是隐马尔可夫模型(Hidden Markov Model ,HMM) 为例 , 如果在一条路径上搜索时得到相邻状态的声学模型的得分的差值大于预设值 , 则可以将这条路径确定为异常路径。

进一步的 , 如果在一段 VAD 检测的语音当中包含一个以上的唤醒词 , 可以会在检测唤醒词之后 , 立刻对 VAD 进行重置 , 重新开始进行唤醒词检测过程 , 以免在一段 VAD 当中只能命中一个唤醒词的现象发生。

S206 : 当获取到语音识别结果的前 3 个字时 , 判断前 3 个字中是否包含唤醒词中的至少部分字 , 若是 , 执行 S207 , 否则执行 S209。

S207 : 继续语音识别 , 判断整个语音识别结果中是否包含唤醒词 , 若是 , 执行 S208 , 否则执行 S209。

S208 : 执行唤醒操作。

S209 : 不唤醒。

其中 , 当连接 S207 时 , 在确定不唤醒时 , 还直接结束对语音信号的解码。

S210：资源释放。

当不唤醒或唤醒后，可以进行资源释放。

资源释放主要功能是对初始化加载的各类资源所占用的内存进行释放，并完成唤醒模块的重置工作，清除历史缓存数据等内容。

本实施例中，通过采用根据唤醒词的分词结果训练得到的第一反模型，可以降低误唤醒率。通过在语音识别结果的前面预设个数的字不包含唤醒词的至少部分内容时，直接确定不唤醒，并结束语音解码，可以降低功耗。通过对语音信号进行音频处理，可以提高抗噪能力。通过对语料的聚类结果进行训练生成反模型，可以减小该反模型的规模，从而可以应用在终端本地，以解决需要全程联网的问题。通过设置多个唤醒词，可以实现通过任一个唤醒词的唤醒。通过对唤醒词所在路径的加权处理，可以提高唤醒灵敏度。通过在解码时直接结束异常路径的搜索，可以降低功耗。通过在搜索空间中包括反模型和唤醒词串联的路径，可以在用户夹杂唤醒词的一句话中依然成功完成唤醒，提高唤醒精度。

图 4 是本申请一个实施例提出的语音唤醒装置的结构示意图。

如图 4 所示，该装置 40 包括：获取模块 41、解码模块 42、判断模块 43 和第一处理模块 44。

获取模块 41，用于获取待处理的语音信号；

解码模块 42，用于根据预先生成的搜索空间，对所述语音信号进行解码，

得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；

判断模块 43，用于当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

第一处理模块 44，用于如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

一些实施例中，参见图 5，所述获取模块 41 包括：

接收子模块 411，用于接收用户输入的语音信号；

音频处理子模块 412，用于对所述语音信号进行音频处理；

端点检测子模块 413，用于对音频处理后的语音信号进行 VAD，得到待处理的语音信号。

一些实施例中，音频处理子模块 412 具体用于：

对所述语音信号进行高通滤波，以去除低频噪声；

对所述语音信号进行噪声抑制，以去除非低频噪声；

对所述语音信号进行 AGC，以增强语音信号的强度。

一些实施例中，参见图 5，该装置 40 还包括：

第二处理模块 45，用于如果所述预设个数的字中包含唤醒词中的至少部分字，则继续进行语音解码得到所述语音信号对应的整个语音识别结果；如果所述整个语音识别结果中包含唤醒词，则执行唤醒操作。

一些实施例中，所述反模型还包括：第二反模型，所述第二反模型根据语料的聚类结果训练生成。

一些实施例中，所述搜索空间还包括：唤醒词所在路径，所述唤醒词所在路径的权重进行了加权处理。

一些实施例中，所述搜索空间还包括：反模型和唤醒词串联所在路径。

一些实施例中，所述解码模块 42 具体用于：在解码时，如果发现异常路径，则直接结束对所述路径的搜索。

一些实施例中，所述唤醒词为多个。

可以理解的是，本实施例的装置与上述方法实施例对应，具体内容可以参见方法实施例的相关描述，在此不再详细说明。

本实施例中，通过采用根据唤醒词的分词结果训练得到的第一反模型，可以降低误唤醒率。通过在语音识别结果的前面预设个数的字不包含唤醒词的至少部分内容时，直接确定不唤醒，并结束语音解码，可以降低功耗。通过对语音信号进行音频处理，可以提高抗噪能力。通过对语料的聚类结果进行训练生成反模型，可以减小该反模型的规模，从而可以应用在终端本地，以解决需要全程联网的问题。通过设置多个唤醒词，可以实现通过任一个唤醒词的唤醒。通过对唤醒词所在路径的加权处理，可以提高唤醒灵敏度。通过在解码时直接结束异常路径的搜索，可以降低功耗。通过在搜索空间中包括反模型和唤醒词串联的路径，可以在用户夹杂唤醒词的一句话中依然成功完成唤醒，提高唤醒

精度。

可以理解的是，上述各实施例中相同或相似部分可以相互参考，在一些实施例中未详细说明的内容可以参见其他实施例中相同或相似的内容。

本申请实施例还提出了一种终端，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本申请实施例还提出了一种非临时性计算机可读存储介质，当所述存储介质中的指令由终端的处理器执行时，使得终端能够执行一种方法，所述方法包括：获取待处理的语音信号；根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

本申请实施例还提出了一种计算机程序产品，当所述计算机程序产品中的

指令被处理器执行时,执行一种方法,所述方法包括:获取待处理的语音信号;根据预先生成的搜索空间,对所述语音信号进行解码,得到语音识别结果,其中,所述搜索空间包括反模型所在路径,所述反模型包括第一反模型,所述第一反模型根据对唤醒词的分词结果训练生成;当获取到所述语音识别结果的前面的预设个数的字时,判断所述预设个数的字中是否包含唤醒词中的至少部分字;如果不包含,则直接确定不唤醒,结束对所述语音信号的解码。

需要说明的是,在本申请的描述中,术语“第一”、“第二”等仅用于描述目的,而不能理解为指示或暗示相对重要性。此外,在本申请的描述中,除非另有说明,“多个”的含义是指至少两个。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本申请的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本申请的实施例所属技术领域的技术人员所理解。

应当理解,本申请的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如,如果用硬件来实现,和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有

用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列(PGA)，现场可编程门阵列(FPGA)等。

本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通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，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本申请各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读取存储介质中。

上述提到的存储介质可以是只读存储器，磁盘或光盘等。

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申请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一定指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何的一个或多个实施例或示例中以合适的方式结合。

尽管上面已经示出和描述了本申请的实施例，可以理解的是，上述实施例

是示例性的，不能理解为对本申请的限制，本领域的普通技术人员在本申请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、一种语音唤醒方法，其特征在于，包括：

获取待处理的语音信号；

根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；

当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

2、根据权利要求1所述的方法，其特征在于，所述获取待处理的语音信号，包括：

接收用户输入的语音信号；

对所述语音信号进行音频处理；

对音频处理后的语音信号进行VAD，得到待处理的语音信号。

3、根据权利要求2所述的方法，其特征在于，所述对所述语音信号进行音频处理，包括：

对所述语音信号进行高通滤波，以去除低频噪声；

对所述语音信号进行噪声抑制，以去除非低频噪声；

对所述语音信号进行 AGC，以增强语音信号的强度。

4、根据权利要求 1-3 任一项所述的方法，其特征在于，还包括：

如果所述预设个数的字中包含唤醒词中的至少部分字，则继续进行语音解码得到所述语音信号对应的整个语音识别结果；

如果所述整个语音识别结果中包含唤醒词，则执行唤醒操作。

5、根据权利要求 1-4 任一项所述的方法，其特征在于，所述反模型还包括：第二反模型，所述第二反模型根据语料的聚类结果训练生成。

6、根据权利要求 1-5 任一项所述的方法，其特征在于，所述搜索空间还包括：唤醒词所在路径，所述唤醒词所在路径的权重进行了加权处理。

7、根据权利要求 1-6 任一项所述的方法，其特征在于，所述搜索空间还包括：反模型和唤醒词串联所在路径。

8、根据权利要求 1-7 任一项所述的方法，其特征在于，在解码时，如果发现异常路径，则直接结束对所述路径的搜索。

9、根据权利要求 1-8 任一项所述的方法，其特征在于，所述唤醒词为多个。

10、一种语音唤醒装置，其特征在于，包括：

获取模块，用于获取待处理的语音信号；

解码模块，用于根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；

判断模块，用于当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

第一处理模块，用于如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

11、根据权利要求 10 所述的装置，其特征在于，所述获取模块包括：

接收子模块，用于接收用户输入的语音信号；

音频处理子模块，用于对所述语音信号进行音频处理；

端点检测子模块，用于对音频处理后的语音信号进行 VAD，得到待处理的语音信号。

12、根据权利要求 10-11 任一项所述的装置，其特征在于，还包括：

第二处理模块，用于如果所述预设个数的字中包含唤醒词中的至少部分字，则继续进行语音解码得到所述语音信号对应的整个语音识别结果；如果所述整个语音识别结果中包含唤醒词，则执行唤醒操作。

13、根据权利要求 10-12 任一项所述的装置，其特征在于，所述解码模块具体用于：在解码时，如果发现异常路径，则直接结束对所述路径的搜索。

14、一种终端，其特征在于，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

获取待处理的语音信号；

根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，

其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；

当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

15、一种非临时性计算机可读存储介质，其特征在于，当所述存储介质中的指令由终端的处理器执行时，使得终端能够执行一种方法，所述方法包括：

获取待处理的语音信号；

根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述第一反模型根据对唤醒词的分词结果训练生成；

当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

16、一种计算机程序产品，其特征在于，当所述计算机程序产品中的指令被处理器执行时，执行一种方法，所述方法包括：

获取待处理的语音信号；

根据预先生成的搜索空间，对所述语音信号进行解码，得到语音识别结果，其中，所述搜索空间包括反模型所在路径，所述反模型包括第一反模型，所述

第一反模型根据对唤醒词的分词结果训练生成；

当获取到所述语音识别结果的前面的预设个数的字时，判断所述预设个数的字中是否包含唤醒词中的至少部分字；

如果不包含，则直接确定不唤醒，结束对所述语音信号的解码。

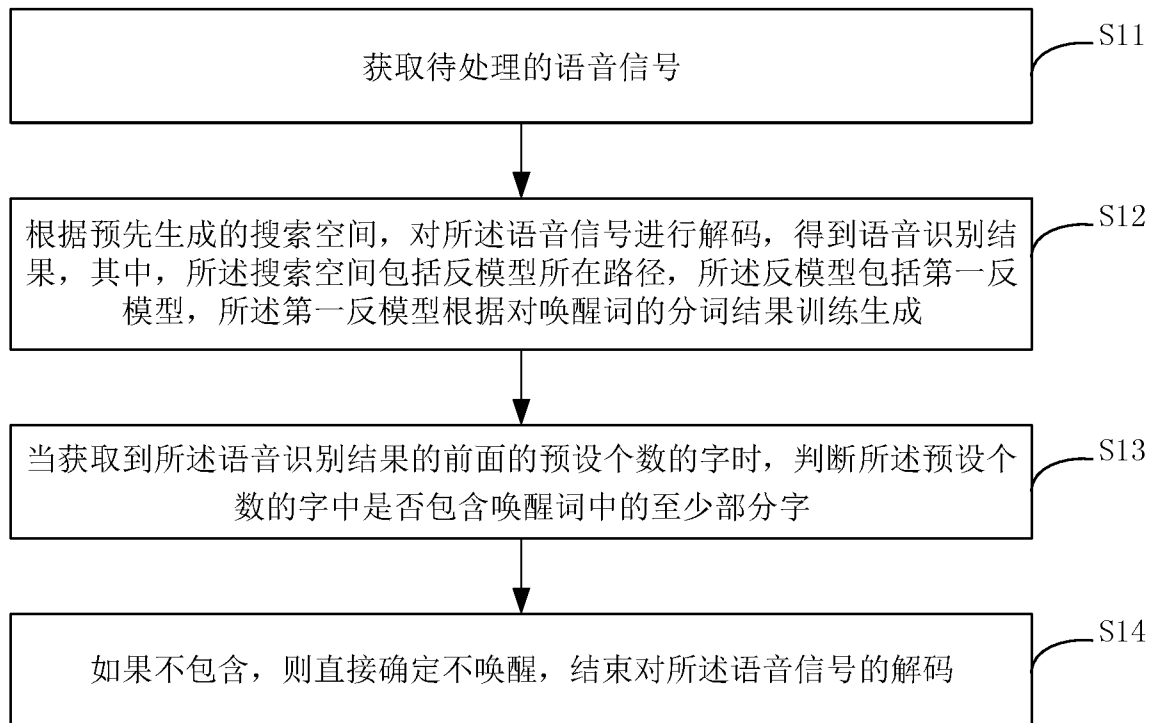


图 1

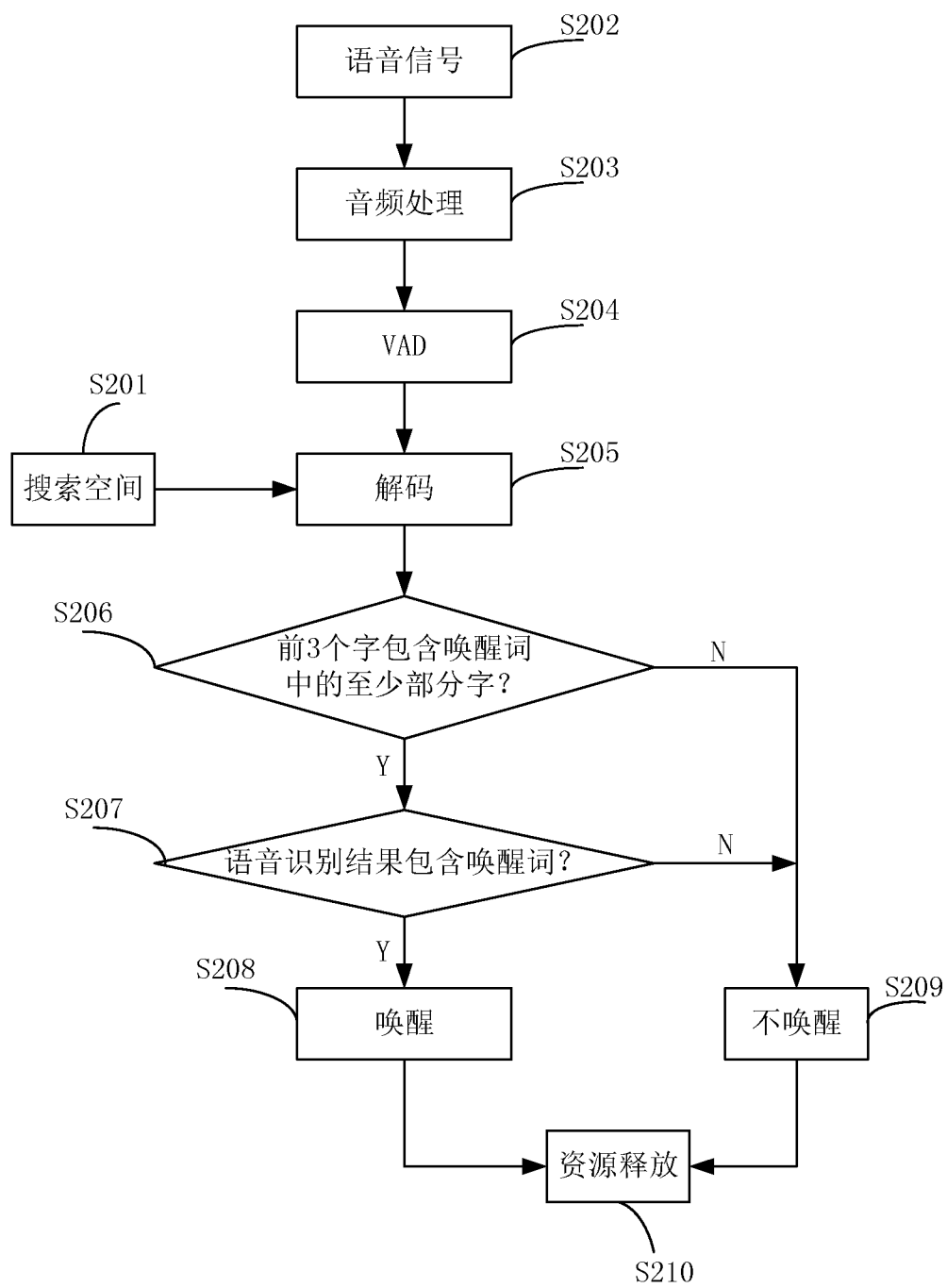


图 2

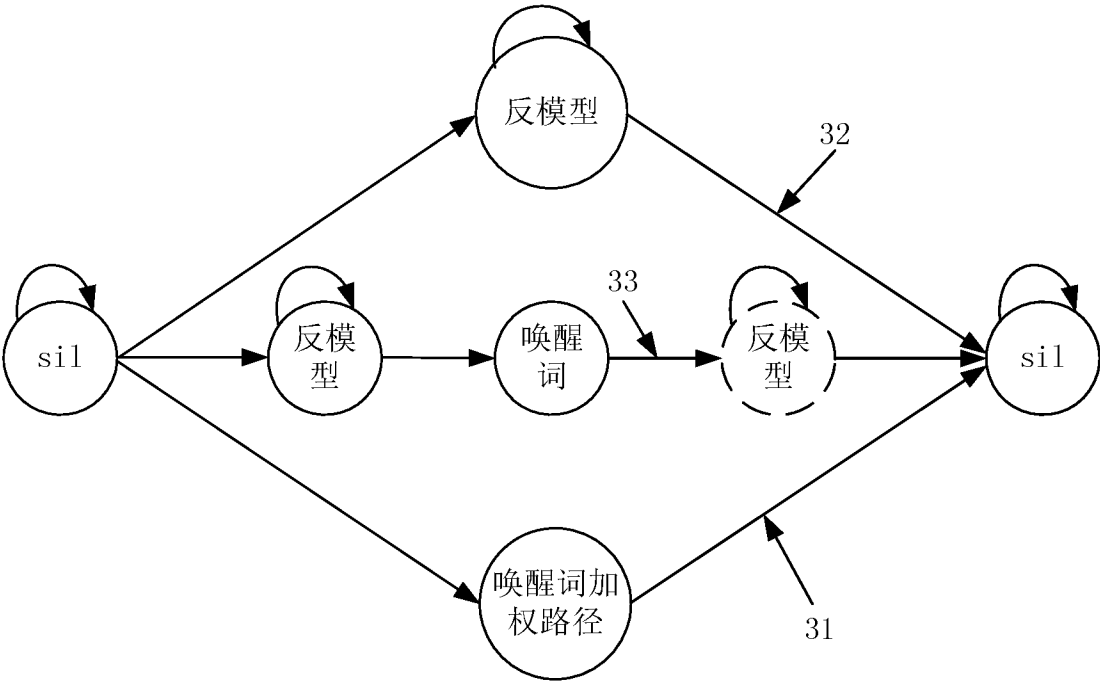


图 3

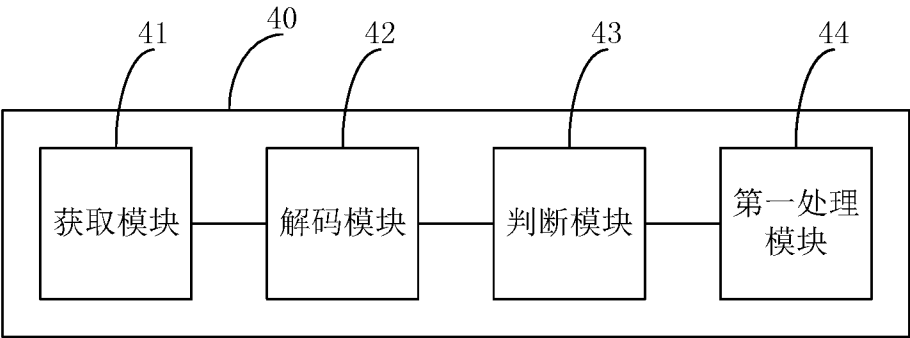


图 4

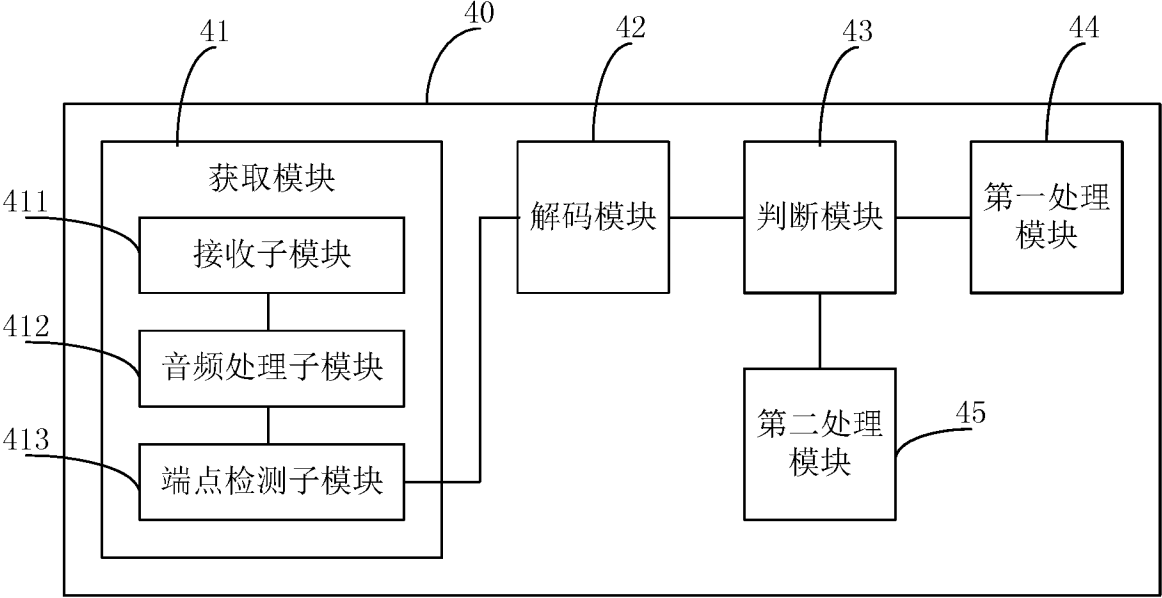


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/111367

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22 (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)

G10L

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)

CNKI, CNPAT, WPI, EPODOC: voice wakeup, voice starting, voice activation, voice command, voice instruction, partial word, word graph network, word segmentation, clustering, corpus, voice, wake-up, activ+, deactiv+, command?, cancel+, recognition, pregenerated, searching space, path, inverse model, utterance, word?, segmentation, keyword, path?, search+, first, front+, portion, segments, part

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105869637 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 17 August 2016 (17.08.2016), claims 1-13	1-16
Y	CN 104464723 A (IFLYTEK CO., LTD.), 25 March 2015 (2015-013-25), description, paragraphs [0037]-[0072] and [0074]-[0077]	1-16
Y	CN 105210146 A (QUALCOMM INC.), 30 December 2015 (30.12.2015), description, paragraphs [0028], [0033] and [0064]-[0069]	1-16
Y	CN 105489222 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 13 April 2016 (13.04.2016), description, paragraphs [0026]-[0028] and [0058]-[0072], and figure 4	2-9, 11-13
A	CN 103021409 A (ANHUI USTC IFLYTEK CO., LTD.), 03 April 2013 (03.04.2013), the whole document	1-16
A	WO 2016064556 A1 (QUALCOMM INC.), 28 April 2016 (28.04.2016), the whole document	1-16

☐ 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 16 January 2017 (16.01.2017)	Date of mailing of the international search report 26 January 2017 (26.01.2017)
Name and mailing address of the ISA/CN: 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 WANG, Yue Telephone No.: (86-10) 62413592

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/111367

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105869637 A	17 August 2016	None	
CN 104464723 A	25 March 2015	None	
CN 105210146 A	30 December 2015	WO 2014182458 A1	13 November 2014
		KR 20160005045 A	13 January 2016
		JP 2016526205 A	01 September 2016
		EP 2994909 A1	16 March 2016
		US 2014334645 A1	13 November 2014
CN 105489222 A	13 April 2016	None	
CN 103021409 A	03 April 2013	CN 103021409 B	24 February 2016
WO 2016064556 A1	28 April 2016	US 2016118039 A1	28 April 2016

A. 主题的分类 G10L 15/22 (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 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)) CNKI, CNPAT, WPI, EPDOC: 语音唤醒, 语音启动, 语音激活, 语音命令, 语音指令, 搜索, 路径, 部分字, 部分词, 词图网络, 分词, 聚类, 语料, voice, wake-up, activ+, deactiv+, command?, cancel+, recognition, pre-generated, searching space, path, inverse model, utterance, word?, segmentation, keyword, path?, search+, first, front+, portion, segments, part		
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
PX	CN 105869637 A (百度在线网络技术北京有限公司) 2016年 8月 17日 (2016 - 08 - 17) 权利要求1-13	1-16
Y	CN 104464723 A (科大讯飞股份有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第37-72, 74-77段	1-16
Y	CN 105210146 A (高通股份有限公司) 2015年 12月 30日 (2015 - 12 - 30) 说明书第28, 33, 64-69段	1-16
Y	CN 105489222 A (百度在线网络技术北京有限公司) 2016年 4月 13日 (2016 - 04 - 13) 说明书26-28, 58-72段, 图4	2-9, 11-13
A	CN 103021409 A (安徽科大讯飞信息科技股份有限公司) 2013年 4月 3日 (2013 - 04 - 03) 全文	1-16
A	WO 2016064556 A1 (QUALCOMM INC.) 2016年 4月 28日 (2016 - 04 - 28) 全文	1-16
☐ 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		国际检索报告邮寄日期
2017年 1月 16日		2017年 1月 26日
ISA/CN的名称和邮寄地址		授权官员
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		王玥 电话号码 (86-10) 62413592

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/111367

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	105869637	A	2016年 8月 17日	无			
CN	104464723	A	2015年 3月 25日	无			
CN	105210146	A	2015年 12月 30日	WO	2014182458	A1	2014年 11月 13日
				KR	20160005045	A	2016年 1月 13日
				JP	2016526205	A	2016年 9月 1日
				EP	2994909	A1	2016年 3月 16日
				US	2014334645	A1	2014年 11月 13日
CN	105489222	A	2016年 4月 13日	无			
CN	103021409	A	2013年 4月 3日	CN	103021409	B	2016年 2月 24日
WO	2016064556	A1	2016年 4月 28日	US	2016118039	A1	2016年 4月 28日