

(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 2017/012243 A1

- (51) 国际专利分类号:
G10L 15/183 (2013.01) G06F 17/30 (2006.01)
G10L 15/26 (2006.01)
- (21) 国际申请号: PCT/CN2015/096622
- (22) 国际申请日: 2015 年 12 月 8 日 (08.12.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510427908.5 2015 年 7 月 20 日 (20.07.2015) 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: 唐立亮 (TANG, Liliang); 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100190
(CN)。贾磊 (JIA, Lei); 中国北京市海淀区上地十
街 10 号百度大厦三层, Beijing 100190 (CN)。彭守
业 (PENG, Shouye); 中国北京市海淀区上地十街 10
号百度大厦三层, Beijing 100190 (CN)。
- (74) 代理人: 北京品源专利代理有限公司 (BEYOND
ATTORNEYS AT LAW); 中国北京市海淀区莲花池
东路 39 号西金大厦 6 层, Beijing 100036 (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,
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)。

[见续页]

(54) Title: VOICE RECOGNITION METHOD AND APPARATUS, TERMINAL DEVICE AND STORAGE MEDIUM

(54) 发明名称: 语音识别方法、装置、终端设备及存储介质

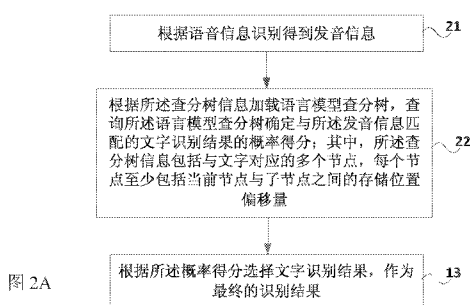


图 2A

- 21 OBTAINING PRONUNCIATION INFORMATION BY RECOGNIZING VOICE INFORMATION
- 22 LOADING A LANGUAGE MODEL SCORE QUERY TREE ACCORDING TO SCORE QUERY TREE INFORMATION, AND QUERYING THE LANGUAGE MODEL SCORE QUERY TREE TO DETERMINE A PROBABILITY SCORE OF A CHARACTER RECOGNITION RESULT MATCHING THE PRONUNCIATION INFORMATION, WHEREIN THE SCORE QUERY TREE INFORMATION COMPRISES A PLURALITY OF NODES CORRESPONDING TO CHARACTERS, AND EACH NODE AT LEAST COMPRISES A STORAGE POSITION OFFSET BETWEEN A CURRENT NODE AND A SUB-NODE
- 13 SELECTING A CHARACTER RECOGNITION RESULT AS A FINAL RECOGNITION RESULT ACCORDING TO THE PROBABILITY SCORE

(57) Abstract: A voice recognition method and apparatus, a terminal device and a storage medium. The method comprises: obtaining pronunciation information according to voice information recognition (21); loading a language model score query tree according to score query tree information, and querying the language model score query tree to determine a probability score of a character recognition result matching the pronunciation information, wherein the score query tree information comprises a plurality of nodes corresponding to characters, and each node at least comprises a storage position offset between a current node and a sub-node (22); and selecting a character recognition result as a final recognition result according to the probability score (13). By means of the method and apparatus, a language model score query tree is directly loaded according to a storage location offset between a current node and a sub node during starting, such that a starting time is greatly shortened.

(57) 摘要:

[见续页]



WO 2017/012243 A1



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种语音识别方法、装置终端设备及存储介质，该方法包括：根据语音信息识别得到发音信息（21）；根据查分树信息加载语言模型查分树，查询语言模型查分树确定与发音信息匹配的文字识别结果的概率得分；其中，查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量（22）；根据概率得分选择文字识别结果，作为最终的识别结果（13）。该方法及装置通过在启动时直接根据当前节点与子节点之间的存储位置偏移量加载语言模型查分树，这样大大的缩短了启动时间。

语音识别方法、装置、终端设备及存储介质

本专利申请要求于 2015 年 07 月 20 日提交的、申请号为 201510427908.5、申请人为百度在线网络技术（北京）有限公司、发明名称为“语音识别方法及装置”的中国专利申请的优先权，该申请的全文以引用的方式并入本申请中。

技术领域

本发明实施例涉及语音识别技术领域，尤其涉及一种语音识别方法、装置、终端设备及存储介质。

背景技术

在嵌入式语音识别领域，语音识别结果由声学模型和语言模型两部分决定。而语言模型有着十分重要的作用，例如，当“北戴河”和“被带河”发音相似，声学模型的得分相差无几，这时就需要使用语言模型来进一步决定哪一个词是语言中会用到的。也就是说，语言模型解决了语音识别中对自然语言顺序的评测问题。

如图 1 所示，为现有技术中提供的语音识别方法，主要包括以下步骤：

S11、从硬盘上读取语言模型资源，资源以节点的方式存储；

其中，每个节点对应一个字，每个节点由节点信息（包括所对应的字或词、孩子信息，例如孩子节点对应的字以及孩子数目），概率列表（ProbList）（存储概率），回退概率列表（BackOff）三部分组成；即如下表一所示：

表一

节点 1(包括 ProbList,	节点 2(包括 ProbList,	
-------------------	-------------------	-------

BackOff, 孩子信息 等)	BackOff, 孩子信息 等)	
---------------------	---------------------	--

S12、根据读取的语言模型资源构建多叉查分树；

构建查分树的过程，具体是：将语言模型资源加载到缓存之后，节点的存储地址发生了变化，因此每个节点只知道自身的孩子节点是哪个字，而不知道其存储地址，因此需要根据每个节点中记录的孩子节点信息，逐一查询其孩子节点的存储地址，并添加至父节点中，从而建立查分树。

S13、加载声学模型和其他语音识别的资源；

S14、接收输入的语音信息，使用维特比算法进行解码；

S15、在解码的过程中，使用声学模型进行语音识别，得到发音信息，并根据发音信息查询语言模型的多叉查分树进行查分；

S16、获得语言模型的识别结果；

S17、输出识别结果，释放资源。

但是，现有的语音识别方法在读取语言模型资源之后，需要对语言模型资源进行动态的加载，构建多叉查分树，这个过程十分浪费时间，导致识别效率较低。

发明内容

本发明实施例提供一种语音识别方法、装置、终端设备及存储介质，能够大大的缩短启动时间。

第一方面，本发明实施例提供了一种语音识别方法，包括：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

第二方面，本发明实施例还提供一种语音识别装置，包括：

发音信息获取模块，用于根据语音信息识别得到发音信息；

概率得分查询模块，用于根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

文字识别模块，用于根据所述概率得分选择文字识别结果，作为最终的识别结果。

第三方面，本发明实施例还提供一种实现语音识别的终端设备，包括：

一个或者多个处理器；

存储器；

一个或者多个模块，所述一个或者多个模块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，进行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置

偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

第四方面，本发明实施例还提供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模块，当所述一个或者多个模块被一个执行语音识别方法的设备执行时，使得所述设备执行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

本发明实施例的技术方案，直接根据当前节点与子节点之间的存储位置偏移量来存储语言模型查分树，无需在启动时动态构建语言模型查分树，这样大大的缩短了启动时间。

附图说明

图 1 为现有技术提供的语音识别方法的流程示意图；

图 2A 为本发明实施例一提供的语音识别方法的流程示意图；

图 2B 为本发明实施例一提供的语音识别方法中的第一种查分子树结构示意图；

图 2C 为本发明实施例一提供的语音识别方法中的第二种查分子树结构示意图；

图 2D 为本发明实施例一提供的语音识别方法中的第三种查分子树结构示意图；

图 2E 为本发明实施例一提供的语音识别方法中的第四种查分子树结构示意图；

图 3 为本发明实施例二提供的语音识别装置的结构示意图；

图 4 为本发明实施例三提供的实现语音识别的终端设备的结构示意图。

具体实施方式

下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释本发明，而非对本发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与本发明相关的部分而非全部结构。

本发明实施例提供的语音识别方法的执行主体，可为本发明实施例提供的语音识别装置，或者集成了所述语音识别装置的终端设备（例如，智能手机、平板电脑等），该语音识别装置可以采用硬件或软件实现。

实施例一

图 2A 为本发明实施例一提供的语音识别方法的流程示意图，如图 2A 所示，具体包括：

S21、根据语音信息识别得到发音信息；

具体的，用户可在本发明实施例提供的语音识别装置中输入语音信息，例如，可在语音识别装置中的输入栏设置一个语音录音按钮，用户通过点击上述

语音录音按钮，即可启动录音功能，对用户说话进行录音，从而获取到所述语音信息。然后通过预先加载的声学模型和语音识别资源对所述语音信息进行识别处理，即可得到需要的发音信息。例如，如果用户想要输入的语音为“北戴河”，则通过上述识别过程可获得的发音信息为“beidaihe”。

S22、根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

其中，所述查分树信息与语言模型资源类似，包括文字对应的多个节点，其中，每个节点至少包括当前节点与子节点之间的存储位置偏移量。除此之外，所述查分树还可以包括每个节点的存储概率（ProbList，即当前节点出现的概率）、当前节点退回的概率（BackOff）以及对应子节点的数目。子节点即孩子节点。父节点与子节点之间是会同时出现的文字组合，例如，“北京”，“京”的节点即为“北”节点的父节点。存储位置偏移量具体是节点与子节点各自存储位置之间的距离。

具体的，预先将语言模型的查分树信息直接写入语言模型资源中，这样初始化上就不需要动态的构建语言模型的查分树资源，而是把动态构建起来的查分树的指针信息当成偏移量写入语言模型资源中，即提前离线构建好语言模型查分树。将当前节点与子节点之间的存储位置偏移量直接写入语言模型资源，所述查分树信息如下述表二所示。在启动加载时，根据查分树信息将离线建立好的语言模型查分树直接进行加载。

表二

节点	节点 1	节点 2	节点 3		节点 n
查分树 信息	(ProbList, BackOff、子 节点数目和 子节点的偏 移量)	(ProbList, BackOff、子 节点数目和 子节点的偏 移量)	(ProbList, BackOff、子 节点数目和 子节点的偏 移量)		(ProbList, BackOff、子 节点数目和 子节点的偏 移量)

当将查分树信息加载到缓存中时，节点之间的存储相对距离不会发生变化，因此，可以基于初始节点的存储位置以及与其他节点的偏移量，确定其他节点的存储位置。

根据上述表二的信息即可将语音模型查分树加载到线上，根据发音信息查询所述查分树。例如，通过上述步骤 S21 得到的发音信息为“beidaihe”，首先在所述查分树的根节点（RootProbList）中查询发音信息“beidaihe”中“he”对应的文字节点，包含多个例如“荷”、“喝”，“河”等，如图 2B 所示，然后在各个“he”对应的文字节点的子节点中查询“dai”对应的文字节点，也包含多个例如“带”、“戴”、“待”等，查询各“dai”对应的文字节点的 ProbList 和 BackOff 中“daihe”对应的概率得分，例如得到如下表三所示的结果：

表三

	ProbList	BackOff
戴荷	1%	98%
带荷	1%	98%
待荷	1%	98%
戴河	99%	1%

带河	80%	5%
待河	60%	30%
戴喝	1%	99%
带喝	1%	98%
待喝	70%	7%

则通过上述表三，可得到“荷”对应的子节点“戴”、“带”和“待”下的文字节点的退回概率 BackOff 均高于 60%，比较高，而 ProbList 均低于 60%，比较低，则“荷”对应的子树被退回。同理，“河”对应的子节点“戴”、“带”和“待”下的文字节点的 ProbList 均高于 60%，而 BackOff 均低于 60%，则“河”对应的子树保留。“喝”对应的子节点“待”下的文字节点的 ProbList 高于 60%，而 BackOff 低于 60%，则“喝”对应的子节点“待”保留，而“喝”对应的子节点“戴”和“带”的 ProbList 均低于 60%，而 BackOff 均高于 60%，则“喝”对应的子节点“戴”和“带”被退回。通过上述选择过程最终可得到的结果如图 2C 和图 2D 所示的两个子树。

在所 2C 和图 2D 所示的子树的基础上，根据发音信息中的“beidaihe”中的“bei”，再次查询“dai”对应的各文字节点的子节点，也包含多个，例如“被”、“北”和“背”等。

查询各“bei”对应的文字节点的 ProbList 和 BackOff 中“beidaihe”对应的概率得分，例如得到如下表四所示的结果：

表四

	ProbList	BackOff
背戴河	30%	70%

被戴河	60%	15%
北戴河	99%	1%
背带河	5%	90%
被带河	1%	95%
北带河	30%	91%
北待河	1%	90%
被待河	1%	98%
背待河	2%	90%
背待喝	1%	99%
北待喝	1%	91%
被待喝	1%	97%

则通过上述表四，可得到“戴”对应的子节点“北戴河”和“被戴河”下的文字节点的 ProbList 均高于 60%，而 BackOff 均低于 60%，则“戴”对应的子节点“北戴河”和“被戴河”保留，而“戴”对应的子节点“背戴河”的文字节点的 BackOff 均高于 60%，而 ProbList 均低于 60%，“戴”对应的子节点“背戴河”被退回。同理，可得出，“待”和“带”对应的子树均被退回。通过上述选择过程最终可得到的结果如图 2E 所示。

S23、根据所述概率得分选择文字识别结果，作为最终的识别结果。

同样以上述步骤 S23 为例，最终得到的识别结果为“北戴河”和“被戴河”，对应的概率得分分别为 99%和 60%，则可按照得分多少，将得分高的文字识别结果放在前边显示，得分低的文字识别结果在显示的原则，同时将它们返回给

用户，即将“北戴河”“被戴河”同时返回给用户，以供用户选择。也可只将得分最高的返回给用户，即将“北戴河”返回给用户。

本实施例，无需在启动时动态构建语言模型查分树，通过当前节点与子节点之间的存储位置偏移量来预先记录语言模型查分树，则需要加载时，可直接根据节点之间存储位置偏移量将查分树加载至缓存中，无需动态建立，这样大大的缩短了启动时间。

示例性的，为了更快的查询，可在根据所述发音信息查询加载的语言模型查分树，确定与所述发音信息匹配的文字识别结果的概率得分之前增加如下步骤：

根据所述发音信息在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；

如果在所述缓存中不存在与所述发音信息匹配的文字识别结果的概率得分，则触发在所述语言模型查分树中进行查询的操作。

其中，所述常用词序列中包含了人们生活中经常用到的一些词汇和热词，例如，包括旅游景点的名称、各省市直辖市自治区的地名、网络名人的名字、歌曲名字等等，将这些常用词汇之间放在缓存中，可以大大提高查询效率。

还可将历史查询的文字识别结果纪录在缓存中，当用户再次输入相同的语音信息时，可直接从缓存中返回给用户，同样节省了查询时间。

由于目前使用的语言模型资源尺寸都比较大，即使是经过裁剪的语言模型，也占用了很大的内存，且经过裁剪还会影响查询效率，为进一步节省内存，本发明实施例将现有的语言模型查分树进行转换，转换为一种更加节约内存的语

言模型。具体的，在根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前进一步增加如下操作，以便形成常用词序列：

将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；

将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；

将所述第一数组和第二数组作为所述常用词序列进行存储。

具体的，将所述语言模型查分树中根节点中包含的部分单个文字或者所有单个文字及其对应的概率得分以数组的形式进行存储。将所述语言模型查分树中父节点以及各子节点对应的文字组合及其概率得分也以数组的形式进行存储。本实施例通过设置设定门限值，可去除所述语言模型查分树中包含的低概率的单个文字和文字组合，以提高查询效率。

举例来说，所述语言模型查分树中根节点中包含的文字包括“北”、“京”、“河”、“荷”、“喝”等，对应的存储概率分别为 P1、P2、P3、P4、P5，具体可采用二维数组来实现，其存储形式如下表五所示：

表五

文字	北	京	河	荷	喝	
出现概率	P1	P2	P3	P4	P5	

同理，对于所述语言模型查分树中父节点以及各子节点对应的文字组合及其概率得分也可以二维数组的形式进行存储，例如如下表六所示，为二元文字的组合：

表六

文字组合	北京	南京	戴河	花荷	待喝	
出现概率	P6	P7	P8	P9	P10	

在后续查询时，可直接从上述数组中进行查询。

示例性的，为了进一步提高查询效率，可通过定位表快速定位要查询的目标区域，具体的可在将所述第一数组和第二数组作为所述常用词序列进行存储之后进一步增加如下步骤：

根据预定规则将所述第二数组中的多个文字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；

将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；

其中，预定规则可根据具体的场景进行设定，对于不同的场景有不同的划分规则，可更好的更快速的匹配到合适的文字识别结果。具体的，可将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ，将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

例如，对于二元文字组合可采用公式一计算得到有序序列数组和无序序列数组，首先计算二元文字组合的特征值 K ：

$$K = M1 \gg 3 + M2 \ll 13$$

其中，第一指定位数取值为 3，第二指定位数取值为 13，“ $\gg$ ”为右移符号，

“<<”为左移符号，M1 为第一文字对应的标识值，M2 为第二文字对应的标识值。

将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

标识值是能够唯一标识文字的数值，例如典型的，以 ASCII 编码标识各文字时，文字的 ASCII 码数值即为标识值。将文字的标识值进行左移和右移之后，计算其特征值 K，相当于根据特征值 K 对各文字组合进行了分类，将特征值 K 相同的文字组合归为一个组。对于组内文字组合数量过少的，则不必设该组。

本领域技术人员可以理解，预设规则也可以是其他公式，不限于左移和右移，以及不限于上述移位的具体位数。

对于三元文字组合，可通过上述公式计算的其中二元文字组合的特征值 K，然后将特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T；将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

例如，可结合上述公式一和如下公式二得到有序序列数组和无序序列数组，首先采用公式一计算得到二元文字组合的特征值 K，然后采用公式二得到三元文字组合的特征值 T

$$T = K \gg 3 + M3 \ll 13$$

其中，第一指定位数取值为 3，第二指定位数取值为 13，K 为 M1 和 M2 组合对应的特征值，M3 为第三文字对应的标识值；

将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

其中，有序序列数组可根据特征值划分为多个子数组，每个子数组存储有相同特征值的文字组合。对于二元文字组合的特征值，采用上述公式一计算得到，对于三元文字组合，可结合公式一和公式二计算得到。

下面通过举例详细介绍定位表的构建过程。

首先根据语音查分树，统计出所有的文字组合，包括二元文字组合、三元文字组合和 n 元文字组合，其中 n 为大于 3 的自然数。较常用的为二元文字组合和三元文字组合。对于二元文字组合，则采用上述公式一计算得到特征值，对于三元文字组合，则采用上述公式一和公式二计算得到特征值。例如，计算得到特征值为 $K1$ 的文字组合包括“北京”、“天津”、“北戴河”、“百度”和“搜狐”，计算得到特征值为 $K2$ 的文字组合包括“杭州好”、“被带河”、“牛奶”和“酸奶”，计算得到特征值为 $K3$ 的文字组合包括“苏州”，计算得到特征值为 $K4$ 的文字组合包括“夏天热”和“被戴河”，计算得到特征值为 $K3$ 的文字组合包括“苏州”则统计特征值为 $K1$ 的文字组合的个数为 5、统计特征值为 $K2$ 的文字组合的个数为 4、统计特征值为 $K3$ 的文字组合的个数为 1，统计特征值为 $K4$ 的文字组合的个数为 2，如果预设数值设为 3，则将相同特征值的文字组合的个数超过 3 的归为有序序列数组，否则，归为无序序列数组，则最终得到的有序序列数组以列表的形式表示，如下述表七所示，其中有序序列数组中还包含特征值和各文字组合出现的概率，此概率可直接从语言模型查分树中获取：

表七

有序序列数组									
特征值	K1					K2			
文字组合	北京	天津	北戴河	百度	搜狐	杭州好	被带河	牛奶	酸奶
概率得分	P6	P11	P8	P12	P13	P14	P15	P16	P17

最终得到的无序序列数组以列表的形式表示，如下述表八所示，其中无序序列数组中同样包含特征值和各文字组合出现的概率，此概率可直接从语言模型查分树中获取：

表八

无序序列数组			
特征值	K3	K4	
文字组合	苏州	夏天热	被戴河
概率得分	P6	P16	P17

由于所述有序序列数组中包含的文字组合个数比较多，在查找时不便于查找，则进一步将有序序列数组中的文字组合进行分割，根据特征值分成多个子数组，例如表七中，可将特征值相同的分为一个子数组，共分为 2 个子数组。如下表九所示：

表九

子数组 1-K1					子数组 2-K2			
北京	天津	北戴河	百度	搜狐	杭州	被带	牛奶	酸奶

		河				河		
P6	P11	P8	P12	P13	P14	P15	P16	P17

而对于无序序列数组，由于其包含的文字组合的个数比较少，则不必将其进行分组。

最后，根据上述表九划分的各个子数组，以及上述表八所示的无序序列数组，构建定位表。即将各子数组对应的特征值及起始存储位置，以及无序序列数组中的文字组合对应的特征值及其概率得分放置在定位表中，例如，得到的定位表如下表十所示

表十

K1	K2	K3	K4
子数组 1 的起止 位置	子数组 2 的起 止位置	苏州- P6	夏天热- P16 被戴河- P17

另外，在构建定位表时，可直接将对应的特征值作为数组的下标，即子数组 1 对应的下标为 K1，子数组 2 对应的下标为 K2，无序序列数组 1 对应的下标为 K3，无序序列数组 2 对应的下标为 K4，则直接将各数组对应的下标存储在定位表中，得到的定位表如下表十一所示：

表十一

K1	K2	K3	K4
		苏州- P6	夏天热- P16 被戴河- P17

相应的，在查询时，根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分具体包

括：

根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；

采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

例如，对于要查询发音信息“beidaihe”对应的文字识别结果，则从所述第一数组存储的单个文字中查询得到发音为“bei”、“dai”、“he”对应的所有文字组合，再根据定位表查询得到各发音为“beidaihe”文字组合的概率得分。例如以表十对应的定位表为例，要查 P（被|带河）的概率，则采用上述公式一和公式二计算得到文字组合“被带河”的特征值为 K2，则根据特征值 K2 查询上述表十的定位表，可知对应的查询范围为子数组 2，则根据定位表中记录的子数组 2 的起止位返回表九去子数组 2 中查询，具体可采用快速查询算法（例如二分法）进行查找，得到 P（被|带河）的概率得分为 P15。例如要查 P（被|戴河）的概率，则采用上述公式一和公式二计算得到文字组合“被戴河”的特征值为 K4，则根据特征值 K4 查询上述表十的定位表，可知对应的查询结果记录在所述定位表中，则直接查询得到 P（被|戴河）的概率得分为 P17。同理，查询得到 P（北|戴河）的概率得分为 P8，比较所有发音为“beidaihe”的文字组合的概率得分，按照概率得分对文字组合进行排序，将排序在前的文字组合返回给用户。

上述各实施例通过根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分，其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节

点之间的存储位置偏移量，根据概率得分即可得到文字识别结果。而无需在启动时动态构建语言模型查分树，本发明实施例通过直接根据当前节点与子节点之间的存储位置偏移量加载语言模型查分树，这样大大的缩短了启动时间。

另外，上述各实施例还通过构建定位表，初步定位要查询的文字组合的大概位置，进一步采用快速查询算法进行精确查找确定与所述发音信息匹配的文字识别结果的概率得分，进一步提高了查询效率。

实施例二

图 3 为本发明实施例二提供的语音识别装置的结构示意图，如图 3 所示，具体包括：发音信息获取模块 31、概率得分查询模块 32 和文字识别模块 33；

所述发音信息获取模块 31 用于根据语音信息识别得到发音信息；

所述概率得分查询模块 32 用于根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

所述文字识别模块 33 用于根据所述概率得分选择文字识别结果，作为最终的识别结果。

本发明实施例所述的语音识别装置用于执行上述各实施例所述的语音识别方法，其技术原理和产生的技术效果类似，这里不再累述。

示例性的，在上述实施例的基础上，所述装置还包括：缓存查询模块 34 和触发模块 35；

所述缓存查询模块 34 用于在所述概率得分查询模块 32 根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分之前，根据所述发音信息在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；

所述触发模块 35 用于如果所述缓存查询模块 34 在缓存中不存在与所述发音信息匹配的文字识别结果的概率得分，则触发在所述语言模型查分树中进行查询的操作。

示例性的，所述装置还包括：第一数组形成模块 36、第二数组形成模块 37 和存储模块 38；

所述第一数组形成模块 36 用于在所述缓存查询模块 34 根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前，将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；

所述第二数组形成模块 37 用于将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；

所述存储模块 38 用于将所述第一数组和第二数组作为所述常用词序列进行存储。

示例性的，所述装置还包括：数组分解模块 39 和定位表构建模块 310；

所述数组分解模块 39 用于在所述存储模块 38 将所述第一数组和第二数组作为所述常用词序列进行存储之后，根据预定规则将所述第二数组中的多个文

字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；

所述定位表模块 310 用于将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；

相应的，所述缓存查询模块 34 具体用于：

根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

示例性的，所述数组分解模块 39 具体用于：

将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ；

将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

示例性的，所述数组分解模块 39 具体用于：

将所述特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T ；

将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

上述各实施例所述的语音识别装置同样用于执行上述各实施例所述的语音

识别方法，其技术原理和产生的技术效果类似，这里不再累述。

实施例三

图 4 为本发明实施例三提供的一种实现语音识别的终端设备的硬件结构示意图，该终端设备包括一个或多个处理器 41、存储器 42，一个或者多个模块，所述一个或者多个模块（例如，附图 3 所示的语音识别装置中的发音信息获取模块 31、概率得分查询模块 32、文字识别模块 33、缓存查询模块 34、触发模块 35 第一数组形成模块 36、第二数组形成模块 37、存储模块 38、数组分解模块 39 和定位表构建模块 310）存储在所述存储器 42 中；图 4 中以一个处理器 41 为例；终端设备中的处理器 41 和存储器 42 可以通过总线或其他方式连接，图 4 中以通过总线连接为例。

当被所述一个或者多个处理器 41 执行时，进行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

上述终端设备可执行本发明实施例一和实施例二所提供的方法，具备执行方法相应的功能模块和有益效果。

示例性的，所述处理器 41 在根据所述发音信息查询加载的语言模型查分树，确定与所述发音信息匹配的文字识别结果的概率得分之前，根据所述发音信息

在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；如果在所述缓存中不存在与所述发音信息匹配的文字识别结果的概率得分，则触发在所述语言模型查分树中进行查询的操作。

示例性的，所述处理器 41 在根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前，将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；将所述第一数组和第二数组作为所述常用词序列进行存储。

示例性的，所述处理器 41 将所述第一数组和第二数组作为所述常用词序列进行存储之后，根据预定规则将所述第二数组中的多个文字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

示例性的，所述处理器 41 将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ；将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；将特征值为 K 的二元文字组合的个数小于预设数值的二元文字

组合归为无序序列数组。

示例性的，所述处理器 41 将所述特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T ；将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

实施例四

本发明实施例还提供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模块，当所述一个或者多个模块被一个执行语音识别方法的设备执行时，使得所述设备执行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

上述存储介质中存储的模块被所述设备所执行时，在根据所述发音信息查询加载的语言模型查分树，确定与所述发音信息匹配的文字识别结果的概率得分之前，优选包括：

根据所述发音信息在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；

如果在所述缓存中不存在与所述发音信息匹配的文字识别结果的概率得分，

则触发在所述语言模型查分树中进行查询的操作。

上述存储介质中存储的模块被所述设备所执行时，在根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前，优选包括：

将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；

将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；

将所述第一数组和第二数组作为所述常用词序列进行存储。

上述存储介质中存储的模块被所述设备所执行时，将所述第一数组和第二数组作为所述常用词序列进行存储之后，优选包括：

根据预定规则将所述第二数组中的多个文字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；

将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；

相应的，根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分优选为：

根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；

采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

上述存储介质中存储的模块被所述设备所执行时，根据预定规则将所述第二数组中的二元文字组合分为有序序列数组和无序序列数组优选为：

将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ；

将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；

将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

上述存储介质中存储的模块被所述设备所执行时，根据预定规则将所述第二数组中的三元文字组合分为有序序列数组和无序序列数组优选为：

将所述特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T ；

将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；

将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

注意，上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解，本发明不限于这里所述的特定实施例，对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此，虽然通过以上实施例对本发明进行了较为详细的说明，但是本发明不仅仅限于以上实施例，在不脱离本发明构思的情况下，还可以包括更多其他等效实施例，而本发明的范围由所附的权利要求范围决定。

权 利 要 求 书

1、一种语音识别方法，其特征在于，包括：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

2、根据权利要求1所述的方法，其特征在于，在根据所述发音信息查询加载的语言模型查分树，确定与所述发音信息匹配的文字识别结果的概率得分之前，还包括：

根据所述发音信息在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；

如果在所述缓存中不存在与所述发音信息匹配的文字识别结果的概率得分，则触发在所述语言模型查分树中进行查询的操作。

3、根据权利要求2所述的方法，其特征在于，在根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前，还包括：

将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；

将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；

将所述第一数组和第二数组作为所述常用词序列进行存储。

4、根据权利要求 3 所述的方法，其特征在于，将所述第一数组和第二数组作为所述常用词序列进行存储之后，还包括：

根据预定规则将所述第二数组中的多个文字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；

将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；

相应的，根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分包括：

根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；

采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

5、根据权利要求 4 所述的方法，其特征在于，根据预定规则将所述第二数组中的二元文字组合分为有序序列数组和无序序列数组包括：

将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ；

将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；

将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

6、根据权利要求 5 所述的方法，其特征在于，根据预定规则将所述第二数

组中的三元文字组合分为有序序列数组和无序序列数组包括：

将所述特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T ；

将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；

将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

7、一种语音识别装置，其特征在于，包括：

发音信息获取模块，用于根据语音信息识别得到发音信息；

概率得分查询模块，用于根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

文字识别模块，用于根据所述概率得分选择文字识别结果，作为最终的识别结果。

8、根据权利要求 7 所述的装置，其特征在于，所述装置还包括：

缓存查询模块，用于在所述概率得分查询模块根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分之前，根据所述发音信息在缓存中存储的常用词序列和/或记录的历史查询的文字识别结果中查询与所述发音信息匹配的文字识别结果的概率得分；

触发模块，用于如果所述缓存查询模块在缓存中不存在与所述发音信息匹

配的文字识别结果的概率得分，则触发在所述语言模型查分树中进行查询的操作。

9、根据权利要求 8 所述的装置，其特征在于，所述装置还包括：

第一数组形成模块，用于在所述缓存查询模块根据所述发音信息在缓存中的常用词序列中查询与所述发音信息匹配的文字识别结果的概率得分之前，将所述语言模型查分树中单个文字出现的概率得分高于设定门限值的单个文字及其概率得分，形成第一数组；

第二数组形成模块，用于将所述语言模型查分树中至少两个文字构成的文字组合的概率得分高于设定门限值的文字组合及其概率得分，形成第二数组；

存储模块，用于将所述第一数组和第二数组作为所述常用词序列进行存储。

10、根据权利要求 9 所述的装置，其特征在于，所述装置还包括：

数组分解模块，用于在所述存储模块将所述第一数组和第二数组作为所述常用词序列进行存储之后，根据预定规则将所述第二数组中的多个文字组合分为有序序列数组和无序序列数组，所述有序序列数组中包含至少两个子数组，各子数组中存储有相同特征值的多个文字组合；

定位表构模块，用于将所述无序序列数组中的概率得分，以及起始位置和/或终止位置，以及各子数组的特征值、起始位置和/或终止位置存储在定位表中；

相应的，所述缓存查询模块具体用于：

根据所述发音信息和对应的特征值查询所述定位表，确定与所述发音信息匹配的子数组；采用快速查找算法查询所述匹配的子数组确定与所述发音信息匹配的文字识别结果的概率得分。

11、根据权利要求 10 所述的装置，其特征在于，所述数组分解模块具体用

于：

将所述二元文字组合中第一文字对应的标识值右移第一指定位数与第二文字对应的标识值左移第二指定位数的和作为特征值 K ；

将特征值为 K 的二元文字组合的个数大于或等于预设数值的二元文字组合归为有序序列数组；将特征值为 K 的二元文字组合的个数小于预设数值的二元文字组合归为无序序列数组。

12、根据权利要求 11 所述的装置，其特征在于，所述数组分解模块具体用于：

将所述特征值 K 右移第一指定位数与第三文字对应的标识值左移第二指定位数的和作为特征值 T ；

将特征值为 T 的三元文字组合的个数大于或等于预设数值的三元文字组合归为有序序列数组；将特征值为 T 的三元文字组合的个数小于预设数值的三元文字组合归为无序序列数组。

13、一种实现语音识别的终端设备，其特征在于，包括：

一个或者多个处理器；

存储器；

一个或者多个模块，所述一个或者多个模块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，进行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置

偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

14、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模块，其特征在于，当所述一个或者多个模块被一个执行语音识别方法的设备执行时，使得所述设备执行如下操作：

根据语音信息识别得到发音信息；

根据所述查分树信息加载语言模型查分树，查询所述语言模型查分树确定与所述发音信息匹配的文字识别结果的概率得分；其中，所述查分树信息包括与文字对应的多个节点，每个节点至少包括当前节点与子节点之间的存储位置偏移量；

根据所述概率得分选择文字识别结果，作为最终的识别结果。

说明书附图

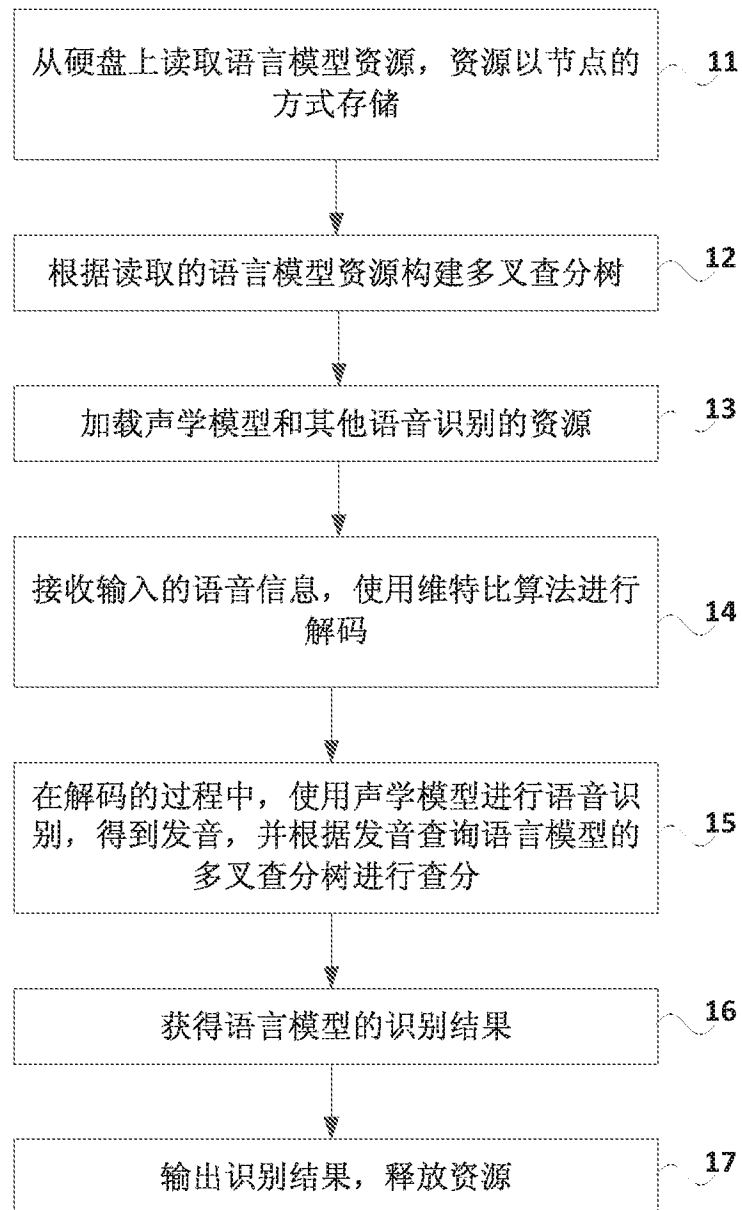


图 1

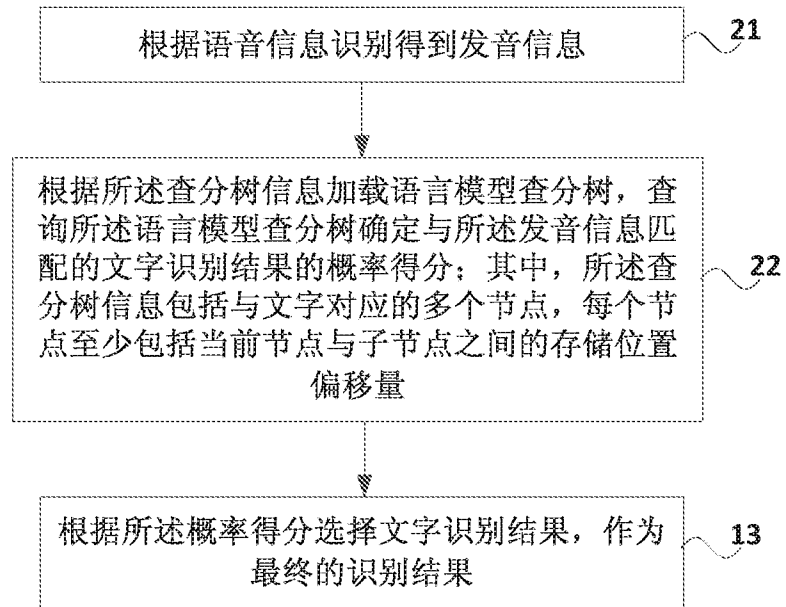


图 2A

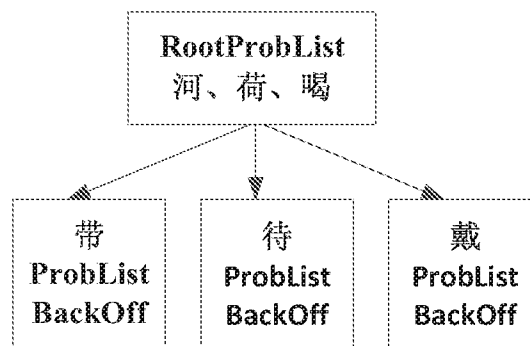


图 2B

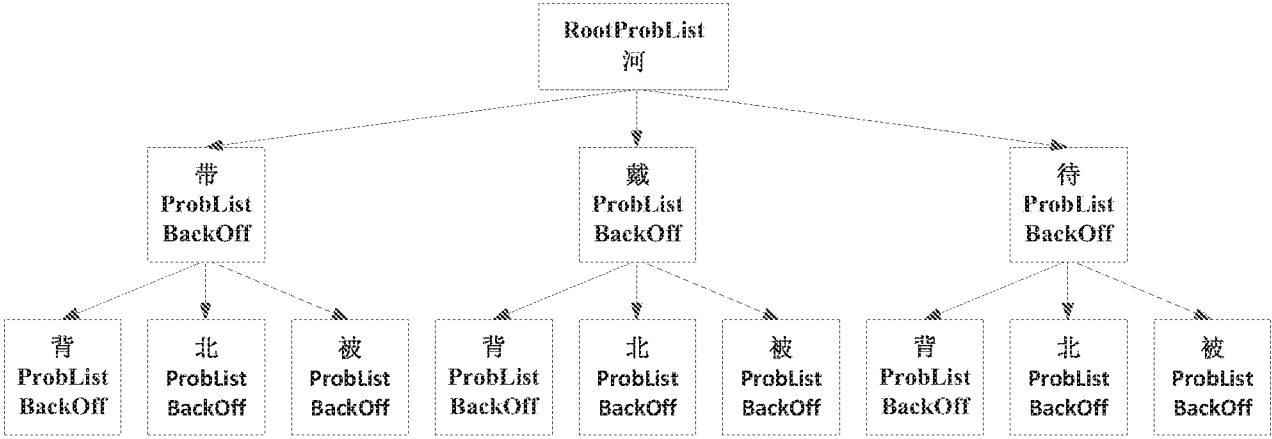


图 2C

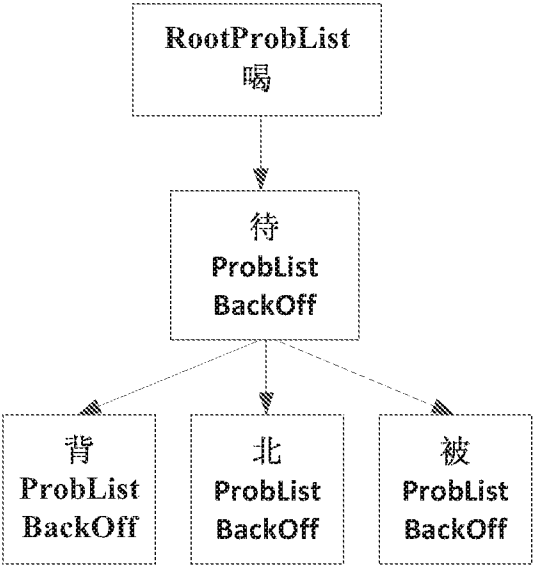


图 2D

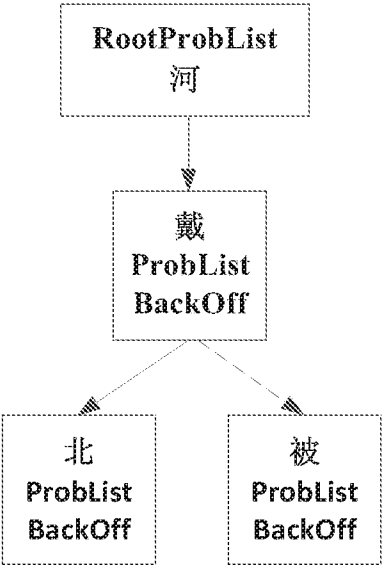


图 2E

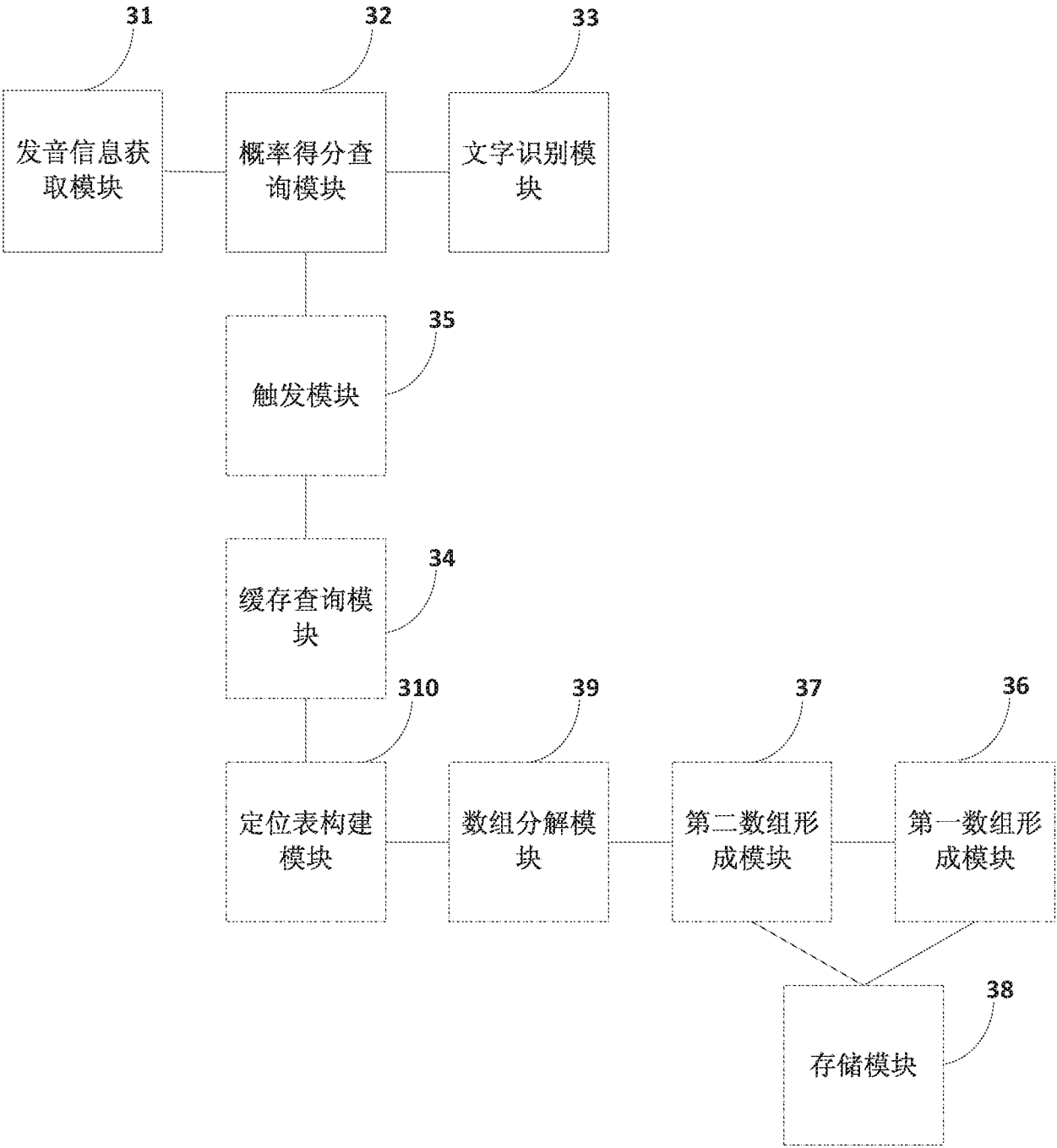


图 3

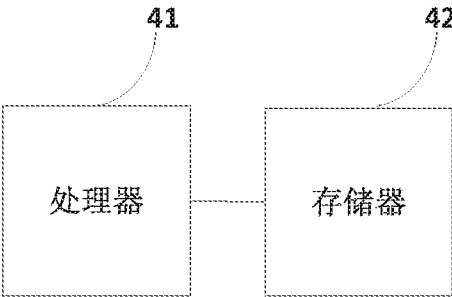


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/096622

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/183 (2013.01) i; G10L 15/26 (2006.01) i; G06F 17/30 (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; G06F

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, CNTXT, CNKI, VEN: speech model, language model, lexical tree, lexicon, storage position, relative, deviate, speech recognition, language, tree, thesaurus, dictionary, load+, creat+, build+, construct+, establish+, child node, leaf node, sub node, address, position, location, offset, excursion, departure, save, shorten, time, probability, score

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105096944 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 25 November 2015 (25.11.2015), claims 1-12	1-14
Y	US 2007055525 A1 (KENNEWICK, R.A. et al.), 08 March 2007 (08.03.2007), description, paragraphs [0022]-[0042], and figures 1-3	1-3, 7-9, 13-14
Y	US 7810024 B1 (ADOBE SYSTEMS INC.), 05 October 2010 (05.10.2010), abstract	1-3, 7-9, 13-14
A	JP 2009156941 A (ATR ADVANCED TELECOMM. RES. INST.), 16 July 2009 (16.07.2009), the whole document	1-14
A	CN 101604522 A (BEIJING SIMBOK INTELLIGENCE TECHNOLOGY CO., LTD.), 16 December 2009 (16.12.2009), 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 20 April 2016 (20.04.2016)	Date of mailing of the international search report 28 April 2016 (28.04.2016)
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 YU, Xin Telephone No.: (86-10) 62085168

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/096622

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 103577548 A (UC MOBILE LTD.), 12 February 2014 (12.02.2014), the whole document	1-14
A	CN 104238991 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 24 December 2014 (24.12.2014), the whole document	1-14
A	CN 1346112 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES), 24 April 2002 (24.04.2002), the whole document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/096622

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105096944 A	25 November 2015	None	
US 2007055525 A1	08 March 2007	US 2010049501 A1	25 February 2010
		US 2011231188 A1	22 September 2011
		US 2010049514 A1	25 February 2010
		US 8150694 B2	03 April 2012
		WO 2007027989 A3	15 November 2007
		WO 2007027989 A2	08 March 2007
		EP 1934971 A4	27 October 2010
		US 7983917 B2	19 July 2011
		EP 1934971 A2	25 June 2008
		US 8069046 B2	29 November 2011
		US 7634409 B2	15 December 2009
US 7810024 B1	05 October 2010	None	
JP 2009156941 A	16 July 2009	JP 5120749 B2	16 January 2013
CN 101604522 A	16 December 2009	CN 101604522 B	28 September 2011
CN 103577548 A	12 February 2014	None	
CN 104238991 A	24 December 2014	WO 2014201834 A1	24 December 2014
		US 2014379335 A1	25 December 2014
CN 1346112 A	24 April 2002	CN 1161703 C	11 August 2004

A. 主题的分类 G10L 15/183(2013.01)i; G10L 15/26(2006.01)i; G06F 17/30(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; G06F 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, CNTXT, CNKI, VEN: 语音识别, 语言模型, 树, 词树, 词库, 词典, 加载, 构建, 创建, 建立, 构造, 子节点, 叶节点, 存储位置, 地址, 偏移, 相对, 偏离, 偏差, 节省, 节约, 缩短, 时间, 概率, 得分, 分值, 打分, speech recognition, language, tree, thesaurus, dictionary, load+, creat+, build+, construct+, establish+, child node, leaf node, sub node, address, position, location, offset, excursion, departure, save, shorten, time, probability, score		
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 105096944 A (百度在线网络技术北京有限公司) 2015年 11月 25日 (2015 - 11 - 25) 权利要求1-12	1-14
Y	US 2007055525 A1 (KENNEWICK ROBERT A. 等) 2007年 3月 8日 (2007 - 03 - 08) 说明书第[0022]-[0042]段及图1-3	1-3, 7-9, 13-14
Y	US 7810024 B1 (ADOBE SYSTEMS INC.) 2010年 10月 5日 (2010 - 10 - 05) 摘要	1-3, 7-9, 13-14
A	JP 2009156941 A (ATR ADVANCED TELECOMM. RES. INST.) 2009年 7月 16日 (2009 - 07 - 16) 全文	1-14
A	CN 101604522 A (北京森博克智能科技有限公司) 2009年 12月 16日 (2009 - 12 - 16) 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 2016年 4月 20日		国际检索报告邮寄日期 2016年 4月 28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 喻新 电话号码 (86-10)62085168

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 103577548 A (优视科技有限公司) 2014年 2月 12日 (2014 - 02 - 12) 全文	1-14
A	CN 104238991 A (腾讯科技深圳有限公司) 2014年 12月 24日 (2014 - 12 - 24) 全文	1-14
A	CN 1346112 A (中国科学院自动化研究所) 2002年 4月 24日 (2002 - 04 - 24) 全文	1-14

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/096622

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	105096944	A	2015年 11月 25日	无	
US	2007055525	A1	2007年 3月 8日	US	2010049501 A1 2010年 2月 25日
				US	2011231188 A1 2011年 9月 22日
				US	2010049514 A1 2010年 2月 25日
				US	8150694 B2 2012年 4月 3日
				WO	2007027989 A3 2007年 11月 15日
				WO	2007027989 A2 2007年 3月 8日
				EP	1934971 A4 2010年 10月 27日
				US	7983917 B2 2011年 7月 19日
				EP	1934971 A2 2008年 6月 25日
				US	8069046 B2 2011年 11月 29日
				US	7634409 B2 2009年 12月 15日
US	7810024	B1	2010年 10月 5日	无	
JP	2009156941	A	2009年 7月 16日	JP	5120749 B2 2013年 1月 16日
CN	101604522	A	2009年 12月 16日	CN	101604522 B 2011年 9月 28日
CN	103577548	A	2014年 2月 12日	无	
CN	104238991	A	2014年 12月 24日	WO	2014201834 A1 2014年 12月 24日
				US	2014379335 A1 2014年 12月 25日
CN	1346112	A	2002年 4月 24日	CN	1161703 C 2004年 8月 11日

表 PCT/ISA/210 (同族专利附件) (2009年7月)