

(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 10 月 5 日 (05.10.2017)



(10) 国际公布号
WO 2017/166631 A1

- (51) 国际专利分类号:
G10L 15/183 (2013.01) G10L 15/28 (2013.01)
- (21) 国际申请号: PCT/CN2016/096828
- (22) 国际申请日: 2016 年 8 月 26 日 (26.08.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610195611.5 2016 年 3 月 30 日 (30.03.2016) CN
- (71) 申请人: 乐视控股(北京)有限公司 (LE HOLDINGS (BEIJING) CO., LTD.) [CN/CN]; 中国北京市朝阳区姚家园路 105 号 3 号楼 10 层 1102, Beijing 100025 (CN)。乐视致新电子科技(天津)有限公司 (LE SHI ZHI XIN ELECTRONIC TECHNOLOGY (TIANJIN) LIMITED) [CN/CN]; 中国天津市滨海新区生态城动漫中路 126 号动漫大厦 B1 区 2 层 201-427, Tianjin 300467 (CN)。
- (72) 发明人: 王彪 (WANG, Biao); 中国天津市滨海新区生态城动漫中路 126 号动漫大厦 B1 区 2 层 201-427, Tianjin 300467 (CN)。
- (74) 代理人: 北京三聚阳光知识产权代理有限公司 (SUNSHINE INTELLECTUAL PROPERTY INTER-

NATIONAL CO., LTD); 中国北京市海淀区海淀南路甲 21 号中关村知识产权大厦 A 座 5 层 503, Beijing 100080 (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 SIGNAL PROCESSING METHOD, APPARATUS AND ELECTRONIC DEVICE

(54) 发明名称: 语音信号处理方法、装置和电子设备

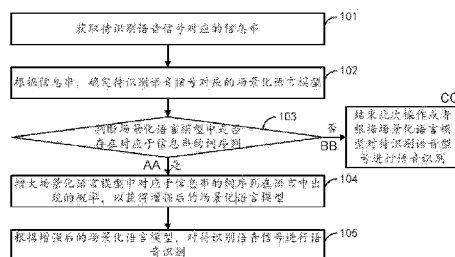


图 1

- 101 Acquiring an information string corresponding to a voice signal to be recognized
- 102 According to the information string, determining a scenarized linguistic model corresponding to the voice signal to be recognized
- 103 Determining whether there is a word sequence corresponding to the information string in the scenarized linguistic model
- 104 Increasing the appearance probability in the language of a word sequence corresponding to the information string in the scenarized linguistic model so as to obtain an enhanced scenarized linguistic model
- 105 According to the enhanced scenarized linguistic model, performing voice recognition on the voice signal to be recognized
- AA Yes
- BB No
- CC Ending this operation or performing voice recognition on the voice signal to be recognized according to a scenarized linguistic model

(57) Abstract: A voice signal processing method, an apparatus and an electronic device. The voice signal processing method comprises: acquiring an information string corresponding to a voice signal to be recognized (101); according to the information string, determining a scenarized linguistic model corresponding to the voice signal to be recognized (102); determining whether there is a word sequence corresponding to the information string in the scenarized linguistic model (103); if a determination result is yes, increasing the appearance probability in the language of a word sequence corresponding to the information string in the scenarized linguistic model so as to obtain an enhanced scenarized linguistic model (104); and according to the enhanced scenarized linguistic model, performing voice recognition on the voice signal to be recognized (105). By adoption of the present embodiment, voice recognition can improve the accuracy rate of voice signal recognition.

(57) 摘要:

[见续页]



WO 2017/166631 A1

根据细则 4.17 的声明:

- 关于申请人有权要求在先申请的优先权(细则 4.17(iii))
- 发明人资格(细则 4.17(iv))

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

一种语音信号处理方法、装置和电子设备。语音信号处理方法包括：获取待识别语音信号对应的信息串（101）；根据信息串，确定待识别语音信号对应的场景化语言模型（102）；判断场景化语言模型中是否存在对应于信息串的词序列（103）；若判断结果为是，增大场景化语言模型中对应于信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型（104）；根据增强后的场景化语言模型，对待识别语音信号进行语音识别（105）。采用本实施例进行语音识别，可以提高语音信号识别的准确率。

语音信号处理方法、装置和电子设备

交叉引用

本申请要求在 2016 年 03 月 30 日提交中国专利局、申请号为 201610195611.5、发明名称为“一种语音信号处理方法及装置”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

本发明实施例涉及语音识别技术领域，尤其涉及一种语音信号处理方法、装置和电子设备。

背景技术

语音识别技术近年来发展迅速，使用户能够通过语音与智能设备进行交互。语音识别技术是通过识别和解析过程把语音信号转变为相应文本或命令的技术。其中，对语音信号的识别和解析过程离不开语言模型(Language Model, LM)，语言模型的目的是建立一个能够描述给定词序列在语言中出现的概率的分布。

在语音识别领域中，大多使用通用语言模型，通用语言模型主要包括通用词序列以及通用词序列在语言中出现的概率，用于对通用领域中的语音信号进行识别。但是，随着时代的发展、应用场景的增多以及用户语言习惯的不断变化等，现有通用语言模型显然无法满足这些应用需求，这会降低语音识别的准确率。

发明内容

本发明实施例提供一种语音信号处理方法\装置和电子设备，用以进行

语音识别，提高语音信号识别的准确率。

本发明实施例提供一种语音信号处理方法，包括：

获取待识别语音信号对应的信息串；

根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

本发明实施例提供一种语音信号处理装置，包括：

获取模块，用于获取待识别语音信号对应的信息串；

确定模块，用于根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断模块，用于判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

增强模块，用于若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

识别模块，用于根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

本发明实施例又提供了一种电子设备，包括至少一个处理器；以及，与所述至少一个处理器通信连接的存储器；其中，所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够获取待识别语音信号对应的信息串；根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；判断所述场景化语言模型中是否存在对应于所述信息串的词序列；若判断结果为是，增大所述场景化

语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

本发明实施例还提供了一种非易失性计算机存储介质，所述存储介质存储有计算机可执行指令的所述计算机可执行指令，当由电子设备执行时使得电子设备能够获取待识别语音信号对应的信息串；根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；判断所述场景化语言模型中是否存在对应于所述信息串的词序列；若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

本发明实施例还提供了计算机程序产品，所述计算机程序产品包括存储于非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行上述的语音信号处理方法。

本发明实施例提供的语音信号处理方法、装置和电子设备，根据待识别语音信号对应的信息串，确定待识别语音信号对应的场景化语言模型，并在该场景化语言模型中存在信息串对应的词序列时，增大该词序列在语言中出现的概率，以获得增强后的场景化语言模型，基于增强后的场景化语言模型对待识别语音信号进行语音识别。与现有技术中基于通用语言模型的语音识别方案相比，本发明实施例基于增强后的场景化语言模型，可以提高语音识别的准确率。

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图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 为本发明一实施例提供的语音信号处理方法的流程示意图；

图 2 为本发明另一实施例提供的语音信号处理方法的流程示意图；

图 3 为本发明又一实施例提供的语音信号处理装置的结构示意图；

图 4 为本发明实施例中电子设备的硬件结构示意图。

具体实施方式

下面将结合附图对本发明的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

在本发明的描述中，需要说明的是，术语“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”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”、“第三”仅用于描述目的，而不能理解为指示或暗示相对重要性。

在本发明的描述中，需要说明的是，除非另有明确的规定和限定，术语“安装”、“相连”、“连接”应做广义理解，例如，可以是固定连接，也可以是可拆卸连接，或一体地连接；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，还可以是两个元件内部的连通，可

以是无线连接，也可以是有线连接。对于本领域的普通技术人员而言，可以具体情况理解上述术语在本发明中的具体含义。

此外，下面所描述的本发明不同实施方式中所涉及的技术特征只要彼此之间未构成冲突就可以相互结合。

在语音识别领域中，大多使用通用语言模型，通用语言模型主要包括通用词序列以及通用词序列在语言中出现的概率，用于对通用领域中的语音信号进行识别。但是，随着时代的发展、应用场景的增多以及用户语言习惯的不断变化等，现有通用语言模型显然无法满足这些应用需求，这会降低语音识别的准确率。

针对现有技术存在的问题，本发明提供一种解决方案，主要原理是：确定待识别语音信号对应的场景化语言模型，并增大该场景化语言模型中相应词序列在语言中出现的概率，以获得增强后的场景化语言模型，基于增强后的场景化语言模型对待识别语音信号进行语音识别。与通用语言模型相比，场景化语言模型包含更多与应用场景相关的词序列（也可称为特定词序列），且预先增大了场景化语言模型中与待识别语音信号相关的词序列在语言中出现的概率，所以基于增强后的场景化语言模型对待识别语音信号进行语音识别，能够提高语音识别的准确率。

下面通过具体实施例对本发明技术方案进行详细说明。

图 1 为本发明一实施例提供的语音信号处理方法的流程示意图。如图 1 所示，该方法包括：

101、获取待识别语音信号对应的信息串。

102、根据信息串，确定待识别语音信号对应的场景化语言模型。

103、判断场景化语言模型中是否存在对应于信息串的词序列；若判断结果为是，则执行步骤 104，若判断结果为否，可选的，结束此次操作或者根据场景化语言模型对待识别语音型号进行语音识别。

104、增大场景化语言模型中对应于信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型。

105、根据增强后的场景化语言模型，对待识别语音信号进行语音识别。

本实施例提供一种语音信号处理方法，可由语音信号处理装置来执行，用以提高语音信号识别的准确率。

具体的，在对待识别语音信号进行识别之前，语音信号处理装置首先获取待识别语音信号对应的信息串。该信息串是指能够在一定程度上反映待识别语音信号的信息串，例如可以是待识别语音信号对应的拼音串，或者是对待识别语音信号进行初始语音识别获得的初始文本串。然后，语音信号处理装置根据该信息串，确定待识别语音信号对应的场景化语言模型，以便于基于该场景化语言模型对待识别语音信号进行语音识别。

可选的，上述根据该信息串，确定待识别语音信号对应的场景化语言模型的实施方式包括：

对待识别语音信号对应的信息串进行语义解析，确定该信息串中的语法句式 and 实体词；根据该语法句式 and 实体词，确定待识别语音信号表达的用户意图；根据该用户意图，确定待识别语音信号对应的场景化语言模型。例如，待识别语音信号对应的信息串为“我要给小李打电话”，经过语义解析，可以确定该信息串中的语法句式为“我要给...打电话”，而实体词为“小李”，根据该语句句式和实体词，可以确定用户意图是要给某人打电话，根据该用户意图，可以确定待识别语音信号对应的场景化语言模型为电话场景语言模型，而不是搜索场景语言模型。

在确定待识别语音信号对应的场景化语言模型之后，并不是直接基于该场景化语言模型对待识别语音信号进行语音识别，而是对该场景化语言模型中的相应词序列在语言中出现的概率进行增大，以提高语音识别准确率。由

于信息串一定程度上反映着待识别语音信号，所以相比于其它词序列，待识别语音信号被识别为所述对应于信息串的词序列的可能更大，基于此，可以将对应于信息串的词序列作为待识别语音信号对应的场景化语言模型中需要增大概率的相应词序列。当然，在增大对应于信息串的词序列在语言中出现的概率之前，先要判断待识别语音信号对应的场景化语言模型中是否存在对应于信息串的词序列；如果判断结果为是，即该场景化语言模型中存在对应于信息串的词序列，则增大该场景化语言模型中对应于信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型，然后基于增强后的场景化语言模型，对待识别语音信号进行语音识别。

在一可选实施方式中，待识别语音信号对应的场景化语言模型包括语法文件和场景词典。语法文件存储该场景化语言模型对应的应用场景中各种语法句式，即一些固定表达方式，例如“请给...打电话”，“请播放歌曲...”，“请搜索歌曲...的歌词”等。场景词典存储该场景化语言模型对应的应用场景中常用的实体词，例如在电话应用场景下实体词可以是通讯录中联系人的姓名，或者在语音控制播放音乐的应用场景中实体词可以是音乐库中的歌曲名等。

基于上述，判断待识别语音信号对应的场景化语言模型中是否存在对应于信息串的词序列的实施方式包括：

对待识别语音信号对应的信息串进行语义解析，确定信息串中的语法句式和实体词；判断信息串中的固定句式是否包含在该场景化语言模型的语法文件中，并判断信息串中的实体词是否包含在该场景化语言模型的场景词典中；若判断结果均为是，则确定该场景化语言模型中存在对应于信息串的词序列，且由信息串中的固定句式和实体词组合成的词序列即为对应于信息串的词序列。

值得说明的是，在确定待识别信号对应的场景化语言模型的过程与在判

断该场景化语言模型中是否存在对应于信息串的词序列的过程中，均包括对信息串进行语义解析，确定信息串中的语法句式 and 实体词的操作，在具体实现是，该操作可以仅执行一次，也可以在两个过程中分别执行一次。

由于上述可见，本实施例中待识别语音信号对应的场景化语言模型包括了与应用场景相关的词序列，同时该场景化语言模型中可能作为待识别语音信号的识别结果的词序列在语言中出现的概率又被增大了，所以基于增强后的场景化语言模型对待识别语音信号进行识别，能够提高语音识别的准确率。

在一可选实施方式中，可以先采用通用语言模型对待识别语音信号进行语音识别，当采用通用语言模型无法识别待识别语音信号对应的词序列时，再采用本发明实施例提供的方法，对待识别语音信号进行语音识别。该实施方式的流程如图 2 所示，包括以下步骤：

200、采用通用语言模型对待识别语音信号进行语音识别；

201、判断采用通用语言模型是否识别出待识别语音信号对应的词序列；若判断结果为是，则结束操作；若判断结果为否，则执行步骤 202。

202、获取待识别语音信号对应的信息串。

203、根据该信息串，确定待识别语音信号对应的场景化语言模型；

204、判断场景化语言模型中是否存在对应于信息串的词序列；若判断结果为是，则执行步骤 205，若判断结果为否，可选的，则执行步骤 207。

205、增大场景化语言模型中对应于信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型。

206、根据增强后的场景化语言模型，对待识别语音信号进行语音识别，并结束此次操作。

207、结束此次操作或者根据场景化语言模型对待识别语音型号进行语音识别，并结束此次操作。

其中，通用语言模型又可以称为大语言模型，而场景化语言模型又可以称为小语言模型。

在一可选实施方式中，在上述步骤 105 或步骤 206 中，可以单独基于增强后的场景化语言模型，对待识别语音信号进行语音识别。

在另一可选实施方式中，在上述步骤 105 或步骤 206 中，可以结合通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别。

值得说明的是，本发明实施例中采用通用语言模型或增强后的场景化语言模型，对待识别语音信号进行语音识别的过程，与现有技术中基于通用语言模型对语音信号进行语音识别的过程类似，在此不再详细说明。

另外，上述结合通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别的一种实施方式包括：

可以将增强后的场景化语言模型叠加到通用语言模型中，生成一个复合语言模型（实际上是一个更大的语言模型），然后基于该复合语言模型对待识别语音信号进行语音识别。

上述结合通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别的另一种实施方式包括：

先使用通用语言模型对待识别语音信号进行语音识别，获得待识别语音信号对应的候选词序列以及在通用语言模型中候选词序列在语言中出现的第一概率，从增强后的场景化语言模型中获取候选词序列在语言中出现的第二概率，将候选词序列的第一概率和第二概率进行加权处理，根据加权处理结果从候选词序列中获取待识别语音信号最终对应的词序列。

上述结合通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别的另一种实施方式包括：

使用通用语言模型对待识别语音信号进行语音识别，获取待识别语音信号对应的第一候选词序列及第一候选词序列在语言中出现的概率；使用增强

后的场景化语言模型对待识别语音信号进行语音识别，获取待识别语音信号对应的第二候选词序列及第二候选词序列在语言中出现的概率；根据第一候选词序列在语言中出现的概率和第二候选词序列在语言中出现的概率，从第一候选词序列和第二候选词序列中获取待识别语音信号最终对应的词序列。其中，对于第一候选词序列和第二候选词序列中的相同候选词序列，可以将其对应的两个概率进行加权求和，作为其最终概率。

在上述实施方式中，除了增强场景化语言模型有利于提高语音识别准确率之外，还结合通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别，可以充分利用通用语言模型包含较多通用词序列，而增强后的场景化语言模型中包含较多与应用场景相关的词序列的特点，提高语音识别的准确率。

图 3 为本发明又一实施例提供的语音信号处理装置的结构示意图。如图 3 所示，该装置包括：获取模块 31、确定模块 32、判断模块 33、增强模块 34 和识别模块 35。

获取模块 31，用于获取待识别语音信号对应的信息串。

确定模块 32，用于根据待识别语音信号对应的信息串，确定待识别语音信号对应的场景化语言模型。

判断模块 33，用于判断待识别语音信号对应的场景化语言模型中是否存在对应于信息串的词序列。

增强模块 34，用于若判断结果为是，增大待识别语音信号对应的场景化语言模型中对应于信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型。

识别模块 35，用于根据增强后的场景化语言模型，对待识别语音信号进行语音识别。

在一可选实施方式中，确定模块 32 具体用于：

对待识别语音信号对应的信息串进行语义解析，确定该信息串中的语法句式和实体词；

根据语法句式和实体词，确定待识别语音信号表达的用户意图；

根据用户意图，确定待识别语音信号对应的场景化语言模型。

在一可选实施方式中，待识别语音信号对应的场景化语言模型包括语法文件和场景词典。基于此，判断模块 33 具体用于：

对待识别语音信号对应的信息串进行语义解析，确定该信息串中的语法句式和实体词

判断固定句式是否包含在语法文件中，并判断实体词是否包含在场景词典中；

若判断结果均为是，则确定场景化语言模型中存在对应于信息串的词序列，且由固定句式和实体词组合成的词序列为对应于信息串的词序列。

在一可选实施方式中，获取模块 31 具体用于：

当采用通用语言模型无法识别待识别语音信号对应的词序列时，获取待识别语音信号对应的信息串。

在一可选实施方式中，识别模块 35 具体用于：

根据通用语言模型和增强后的场景化语言模型，对待识别语音信号进行语音识别。

进一步，识别模块 35 具体用于：先使用通用语言模型对待识别语音信号进行语音识别，获得待识别语音信号对应的候选词序列以及在通用语言模型中候选词序列在语言中出现的第一概率，从增强后的场景化语言模型中获取候选词序列在语言中出现的第二概率，将候选词序列的第一概率和第二概率进行加权处理，根据加权处理结果从候选词序列中获取待识别语音信号最终对应的词序列。

进一步，识别模块 35 具体用于：先使用通用语言模型对待识别语音信

号进行语音识别，获得待识别语音信号对应的候选词序列（通常为多组）以及在通用语言模型中候选词序列在语言中出现的第一概率，从增强后的场景化语言模型中获取候选词序列在语言中出现的第二概率，将候选词序列的第一概率和第二概率进行加权处理，根据加权处理结果从候选词序列中获取待识别语音信号最终对应的词序列。

进一步，识别模块 35 具体用于：使用通用语言模型对待识别语音信号进行语音识别，获取待识别语音信号对应的第一候选词序列及第一候选词序列在语言中出现的概率；使用增强后的场景化语言模型对待识别语音信号进行语音识别，获取待识别语音信号对应的第二候选词序列及第二候选词序列在语言中出现的概率；根据第一候选词序列在语言中出现的概率和第二候选词序列在语言中出现的概率，从第一候选词序列和第二候选词序列中获取待识别语音信号最终对应的词序列。其中，对于第一候选词序列和第二候选词序列中的相同候选词序列，可以将其对应的两个概率进行加权求和，作为其最终概率。

本实施例提供的语音信号处理装置，根据待识别语音信号对应的信息串，确定待识别语音信号对应的场景化语言模型，并在该场景化语言模型中存在信息串对应的词序列时，增大该词序列在语言中出现的概率，以获得增强后的场景化语言模型，基于增强后的场景化语言模型对待识别语音信号进行语音识别，而不是像现有技术那样基于通用语言模型进行语音识别，可以提高语音识别的准确率。

本发明实施例又提供了一种电子设备，包括至少一个处理器810；以及，与所述至少一个处理器810通信连接的存储器800；其中，所述存储器800存储有可被所述至少一个处理器810执行的指令，所述指令被所述至少一个处理器810执行，以使所述至少一个处理器810能够获取待识别语音信号对应的信息串；根据所述信息串，确定所述待识别语音信号对应的场景化

语言模型；判断所述场景化语言模型中是否存在对应于所述信息串的词序列；若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。所述电子设备还包括与所述存储器800和所述处理器电连接的输入装置830和输出装置840，所述电连接优选为通过总线连接。

本发明实施例还提供了一种非易失性计算机存储介质，所述存储介质存储有计算机可执行指令的所述计算机可执行指令，当由电子设备执行时使得电子设备能够获取待识别语音信号对应的信息串；根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；判断所述场景化语言模型中是否存在对应于所述信息串的词序列；若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功

能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

显然，上述实施例仅仅是为清楚地说明所作的举例，而并非对实施方式的限定。对于所属领域的普通技术人员来说，在上述说明的基础上还可以做出其它不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。而由此所引伸出的显而易见的变化或变动仍处于本发明创造的保护范围之内。

权 利 要 求 书

1、一种语音信号处理方法，其特征在于，包括：

获取待识别语音信号对应的信息串；

根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

2、根据权利要求 1 所述的方法，其特征在于，所述根据所述信息串，确定所述待识别语音信号对应的场景化语言模型，包括：

对所述信息串进行语义解析，确定所述信息串中的语法句式 and 实体词；

根据所述语法句式 and 实体词，确定所述待识别语音信号表达的用户意图；

根据所述用户意图，确定所述待识别语音信号对应的场景化语言模型。

3、根据权利要求 1 所述的方法，其特征在于，所述场景化语言模型包括语法文件和场景词典；

所述判断所述场景化语言模型中是否存在对应于所述信息串的词序列，包括：

对所述信息串进行语义解析，确定所述信息串中的语法句式 and 实体词

判断所述固定句式是否包含在所述语法文件中，并判断所述实体词是否包含在所述场景词典中；

若判断结果均为是，则确定所述场景化语言模型中存在对应于所述信息串的词序列，且由所述固定句式和所述实体词组合成的词序列为对应于所述信息串的词序列。

4、根据权利要求 1-3 任一项所述的方法，其特征在于，所述获取待识别

语音信号对应的信息串，包括：

当采用通用语言模型无法识别所述待识别语音信号对应的词序列时，获取所述待识别语音信号对应的信息串。

5、根据权利要求 4 所述的方法，其特征在于，所述根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别，包括：

根据所述通用语言模型和所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

6、一种语音信号处理装置，其特征在于，包括：

获取模块，用于获取待识别语音信号对应的信息串；

确定模块，用于根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断模块，用于判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

增强模块，用于若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

识别模块，用于根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

7、根据权利要求 6 所述的装置，其特征在于，所述确定模块具体用于：

对所述信息串进行语义解析，确定所述信息串中的语法句式 and 实体词；

根据所述语法句式 and 实体词，确定所述待识别语音信号表达的用户意图；

根据所述用户意图，确定所述待识别语音信号对应的场景化语言模型。

8、根据权利要求 6 所述的装置，其特征在于，所述场景化语言模型包括语法文件和场景词典；

所述判断模块具体用于：

对所述信息串进行语义解析，确定所述信息串中的语法句式 and 实体词

判断所述固定句式是否包含在所述语法文件中，并判断所述实体词是否包含在所述场景词典中；

若判断结果均为是，则确定所述场景化语言模型中存在对应于所述信息串的词序列，且由所述固定句式和所述实体词组合成的词序列为对应于所述信息串的词序列。

9、根据权利要求 6-8 任一项所述的装置，其特征在于，所述获取模块具体用于：

当采用通用语言模型无法识别所述待识别语音信号对应的词序列时，获取所述待识别语音信号对应的信息串。

10、根据权利要求 9 所述的装置，其特征在于，所述识别模块具体用于：

根据所述通用语言模型和所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

11、一种电子设备，其特征在于包括至少一个处理器；以及，与至少一个处理器通信连接的存储器；其中，所述存储器存储有可被至少一个处理器执行的指令，所述指令被至少一个处理器执行，以使至少一个处理器能够

获取待识别语音信号对应的信息串；

根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

12、一种非易失性计算机存储介质，其特征在于：所述存储介质存储有

计算机可执行指令的所述计算机可执行指令，当由电子设备执行时使得电子设备能够：

获取待识别语音信号对应的信息串；

根据所述信息串，确定所述待识别语音信号对应的场景化语言模型；

判断所述场景化语言模型中是否存在对应于所述信息串的词序列；

若判断结果为是，增大所述场景化语言模型中对应于所述信息串的词序列在语言中出现的概率，以获得增强后的场景化语言模型；

根据所述增强后的场景化语言模型，对所述待识别语音信号进行语音识别。

13、一种计算机程序产品，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行权利要求 1-5 所述的方法。

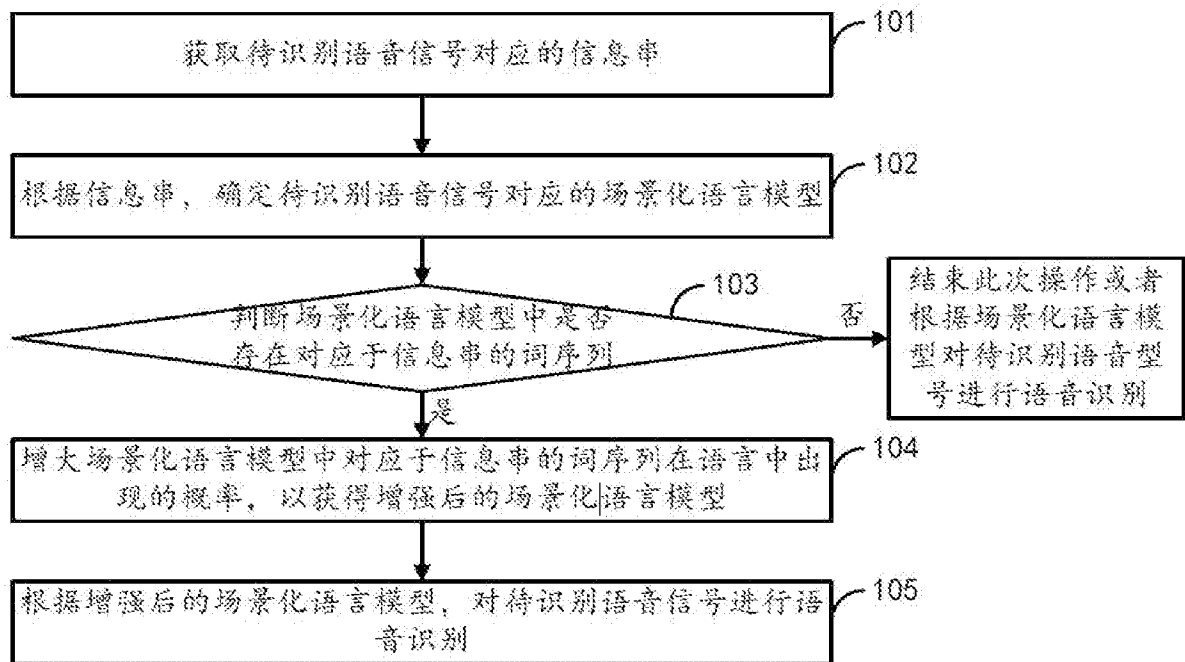


图 1

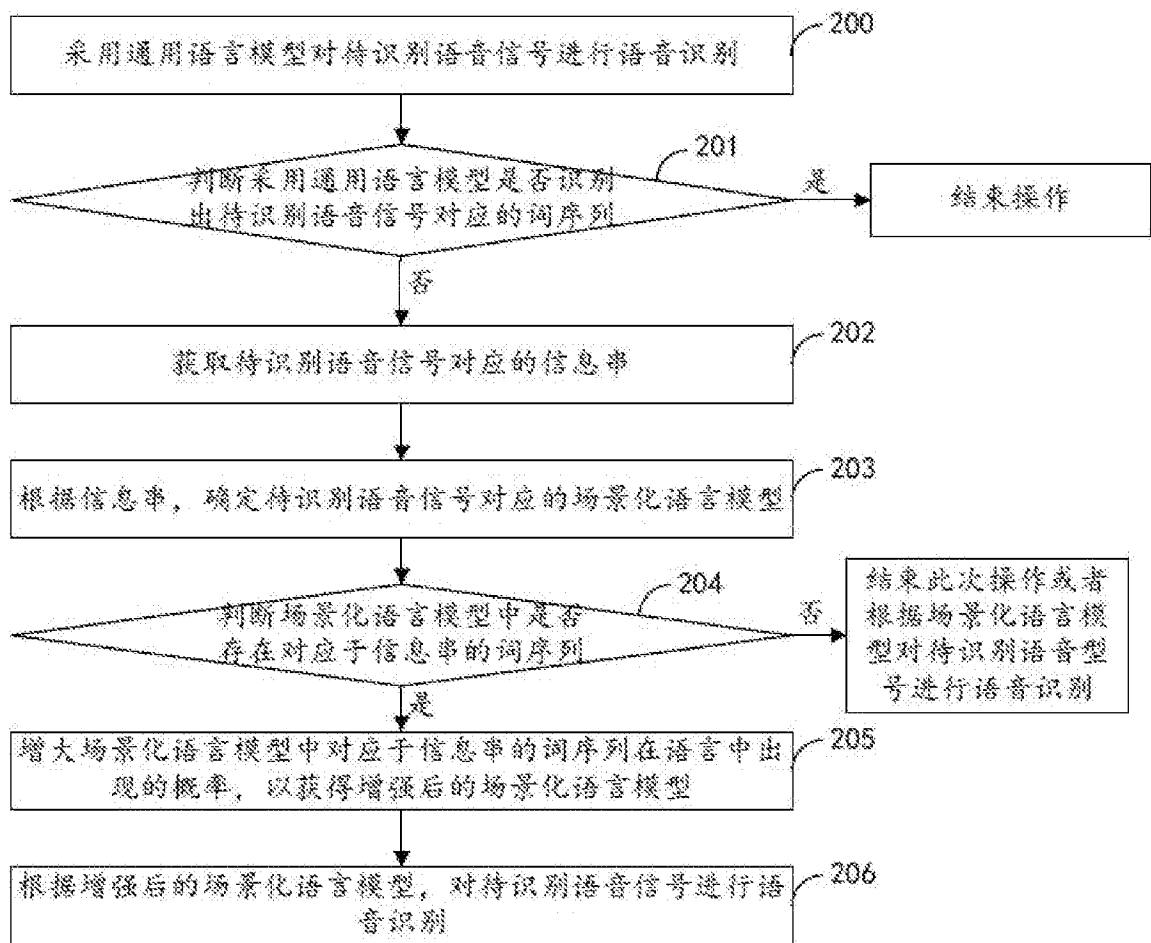


图 2

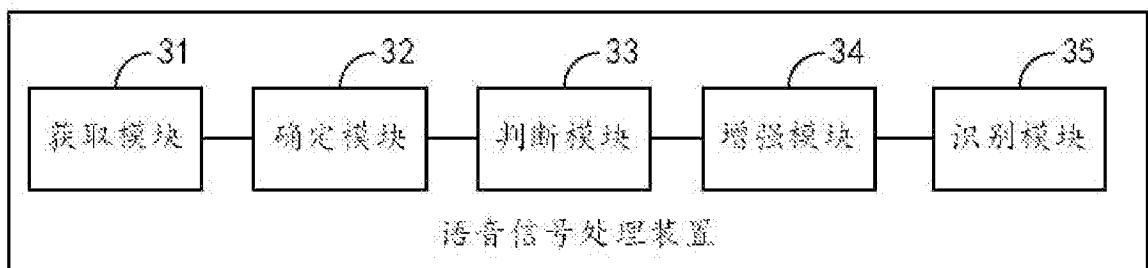


图 3

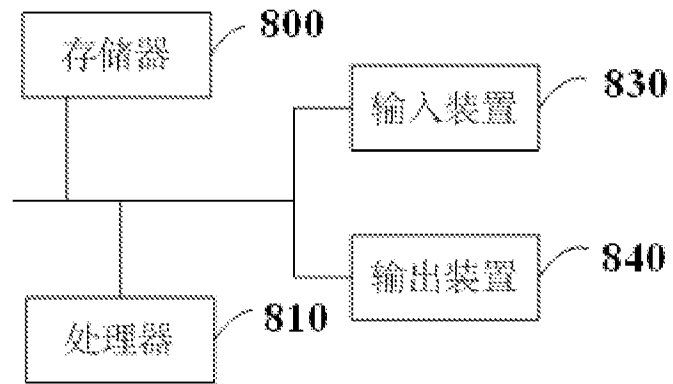


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/096828

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/183 (2013.01) i; G10L 15/28 (2013.01) 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)

G10L15; G06F 17

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: SCENE, SITUATION, MODEL, PROBABILITY, WORD, word sequence, substantives

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105845133 A (LETV HOLDING BEIJING CO., LTD. et al.) 10 August 2016 (10.08.2016) claims 1-10, description, paragraph [0085]	1-13
Y	US 2014025380 A1 (KOCH FERNANDO LUIZ et al.) 23 January 2014 (23.01.2014) description, paragraphs [0071]-[0079], and figures 6 and 7	1-13
Y	CN 102074231 A (WANYINDA CO., LTD.) 25 May 2011 (25.05.2011) description, paragraphs [0035]-[0039], and figure 1	1-13
Y	CN 101593518 A (CHINESE ACAD SCIENCE AUTOMATION INSTITUTE) 02 December 2009 (02.12.2009) description, page 4, line 11 to page 8, line 3, and figures 1 and 2	1-13
A	US 2006015326 A1 (IBM CORP.) 19 January 2006 (19.01.2006) the whole document	1-13
A	JP 2013142870 A (NIPPON TELEGRAPH & TELEPHONE) 22 July 2013 (22.07.2013) the whole document	1-13

☐ 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 17 November 2016	Date of mailing of the international search report 05 January 2017
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 WANG, Xiaoping Telephone No. (86-10) 62085739

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/096828

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105845133 A	10 August 2016	None	
US 2014025380 A1	23 January 2014	US 9043208 B2	26 May 2015
		US 2014025377 A1	23 January 2014
		US 9053708 B2	09 June 2015
CN 102074231 A	25 May 2011	None	
CN 101593518 A	02 December 2009	None	
US 2006015326 A1	19 January 2006	US 7917350 B2	29 March 2011
		JP 4652737 B2	16 March 2011
		US 2008228463 A1	18 September 2008
		JP 2006031295 A	02 February 2006
JP 2013142870 A	22 July 2013	JP 5800718 B2	28 October 2015

国际检索报告

国际申请号

PCT/CN2016/096828

A. 主题的分类

G10L 15/183(2013.01)i; G10L 15/28(2013.01)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)

G10L15; G06F17

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:场景, 情景, 模型, 概率, 词序列, 实体词, SCENE, SITUATION, MODEL, PROBABILITY, WORD

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 105845133 A (乐视控股北京有限公司等) 2016年 8月 10日 (2016 - 08 - 10) 权利要求1-10, 说明书第【0085】段	1-13
Y	US 2014025380 A1 (KOCH FERNANDO LUIZ等) 2014年 1月 23日 (2014 - 01 - 23) 说明书第【0071】—【0079】段, 图6-7	1-13
Y	CN 102074231 A (万音达有限公司) 2011年 5月 25日 (2011 - 05 - 25) 说明书第【0035】—【0039】段, 图1	1-13
Y	CN 101593518 A (中国科学院自动化研究所) 2009年 12月 2日 (2009 - 12 - 02) 说明书第4页第11行至第8页第3行, 图1-2	1-13
A	US 2006015326 A1 (IBM CORP.) 2006年 1月 19日 (2006 - 01 - 19) 全文	1-13
A	JP 2013142870 A (NIPPON TELEGRAPH & TELEPHONE) 2013年 7月 22日 (2013 - 07 - 22) 全文	1-13

☐ 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 11月 17日

国际检索报告邮寄日期

2017年 1月 5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

受权官员

王晓萍

电话号码 (86-10)62085739

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/096828

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	105845133	A	2016年 8月 10日	无			
US	2014025380	A1	2014年 1月 23日	US	9043208	B2	2015年 5月 26日
				US	2014025377	A1	2014年 1月 23日
				US	9053708	B2	2015年 6月 9日
CN	102074231	A	2011年 5月 25日	无			
CN	101593518	A	2009年 12月 2日	无			
US	2006015326	A1	2006年 1月 19日	US	7917350	B2	2011年 3月 29日
				JP	4652737	B2	2011年 3月 16日
				US	2008228463	A1	2008年 9月 18日
				JP	2006031295	A	2006年 2月 2日
JP	2013142870	A	2013年 7月 22日	JP	5800718	B2	2015年 10月 28日

表 PCT/ISA/210 (同族专利附件) (2009年7月)