



(12)发明专利申请

(10)申请公布号 CN 107871506 A

(43)申请公布日 2018.04.03

(21)申请号 201711132605.6

(22)申请日 2017.11.15

(71)申请人 北京云知声信息技术有限公司  
地址 100191 北京市海淀区花园路2号牡丹科技楼5层A503

(72)发明人 谢书杰 苏牧

(74)专利代理机构 北京尚伦律师事务所 11477  
代理人 张俊国

(51)Int.Cl.

G10L 15/22(2006.01)

G10L 15/197(2013.01)

G10L 15/16(2006.01)

G10L 15/06(2013.01)

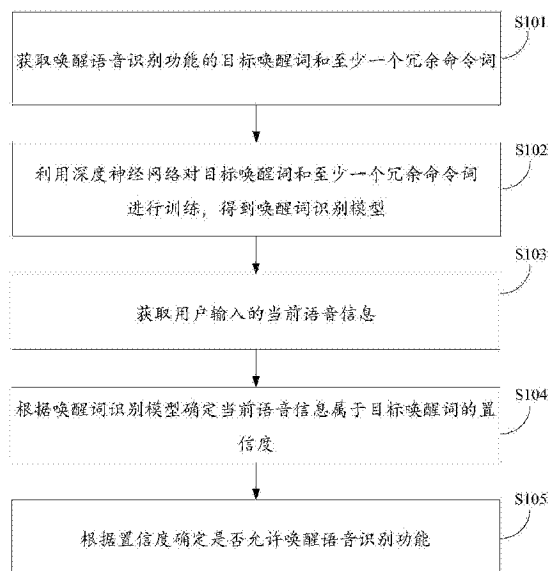
权利要求书2页 说明书8页 附图5页

(54)发明名称

语音识别功能的唤醒方法及装置

(57)摘要

本发明是关于一种语音识别功能的唤醒方法及装置,其中,方法包括:获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;获取用户输入的当前语音信息;根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;根据所述置信度确定是否允许唤醒所述语音识别功能。通过该技术方案,可以保证唤醒词识别的准确率,从而提高语音识别功能唤醒的准确率,降低误唤醒。



1. 一种语音识别功能的唤醒方法,其特征在于,包括:
  - 获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;
  - 利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;
  - 获取用户输入的当前语音信息;
  - 根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;
  - 根据所述置信度确定是否允许唤醒所述语音识别功能。
2. 根据权利要求1所述的方法,其特征在于,所述根据所述置信度确定是否允许唤醒所述语音识别功能,包括:
  - 当所述置信度大于预设置信度时,允许唤醒所述语音识别功能;
  - 当所述置信度小于或者等于所述预设置信度时,禁止唤醒所述语音识别功能。
3. 根据权利要求1所述的方法,其特征在于,所述利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型,包括:
  - 将所述目标唤醒词确定为正例词,将所述至少一个冗余命令词确定为负例词;
  - 利用深度神经网络对所述正例词和所述负例词进行训练,得到所述唤醒词识别模型。
4. 根据权利要求1所述的方法,其特征在于,所述根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度,包括:
  - 获取所述当前语音信息对应的当前音频帧特征序列信息;
  - 将所述当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围;
  - 计算每个音频帧的范围中每个音频帧的声学似然评分,并从所述每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;
  - 根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度。
5. 根据权利要求4所述的方法,其特征在于,所述根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度,包括:
  - 使用所述唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;
  - 从与每个音素对应的多个声学后验得分中选取最大声学后验得分;
  - 根据所有音素的最大声学后验得分计算所述当前语音信息属于所述目标唤醒词的置信度。
6. 一种语音识别功能的唤醒装置,其特征在于,包括:
  - 第一获取模块,用于获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;
  - 训练模块,用于利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;
  - 第二获取模块,用于获取用户输入的当前语音信息;
  - 第一确定模块,用于根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;

第二确定模块,用于根据所述置信度确定是否允许唤醒所述语音识别功能。

7. 根据权利要求6所述的装置,其特征在于,所述第二确定模块包括:

第一处理子模块,用于当所述置信度大于预设置信度时,允许唤醒所述语音识别功能;

第二处理子模块,用于当所述置信度小于或者等于所述预设置信度时,禁止唤醒所述语音识别功能。

8. 根据权利要求6所述的装置,其特征在于,所述训练模块包括:

第一确定子模块,用于将所述目标唤醒词确定为正例词,将所述至少一个冗余命令词确定为负例词;

训练子模块,用于利用深度神经网络对所述正例词和所述负例词进行训练,得到所述唤醒词识别模型。

9. 根据权利要求6所述的装置,其特征在于,所述第一确定模块包括:

获取子模块,用于获取所述当前语音信息对应的当前音频帧特征序列信息;

定位子模块,用于将所述当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围;

选取子模块,用于计算每个音频帧的范围中每个音频帧的声学似然评分,并从所述每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;

第二确定子模块,用于根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度。

10. 根据权利要求9所述的装置,其特征在于,所述第二确定子模块用于:

使用所述唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;

从与每个音素对应的多个声学后验得分中选取最大声学后验得分;

根据所有音素的最大声学后验得分计算所述当前语音信息属于所述目标唤醒词的置信度。

## 语音识别功能的唤醒方法及装置

### 技术领域

[0001] 本发明涉及语音识别技术领域,尤其涉及一种语音识别功能的唤醒方法及装置。

### 背景技术

[0002] 目前,相关技术中可以通过语音对设备进行控制,设备不是实时录音并识别语音命令的,而是先识别是否收到唤醒词,如果收到唤醒词则激活,然后识别语音命令。而如果误识别了唤醒词,则会导致误唤醒。

### 发明内容

[0003] 本发明实施例提供一种语音识别功能的唤醒方法及装置,用以实现提高语音识别功能唤醒的准确率,降低误唤醒。

[0004] 根据本发明实施例的第一方面,提供一种语音识别功能的唤醒方法,包括:

[0005] 获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;

[0006] 利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;

[0007] 获取用户输入的当前语音信息;

[0008] 根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;

[0009] 根据所述置信度确定是否允许唤醒所述语音识别功能。

[0010] 在该实施例中,利用目标唤醒词和冗余命令词进行训练,得到唤醒词识别模型,进而通过该唤醒词识别模型确定当前语音信息属于目标唤醒词的置信度,这样可以保证唤醒词识别的准确率,从而提高语音识别功能唤醒的准确率,降低误唤醒。

[0011] 在一个实施例中,所述根据所述置信度确定是否允许唤醒所述语音识别功能,包括:

[0012] 当所述置信度大于预设置信度时,允许唤醒所述语音识别功能;

[0013] 当所述置信度小于或者等于所述预设置信度时,禁止唤醒所述语音识别功能。

[0014] 在该实施例中,在当前语音信息属于目标唤醒词的置信度大于预设置信度时,才允许唤醒语音唤醒功能,从而降低语音唤醒功能误唤醒。

[0015] 在一个实施例中,所述利用深度神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型,包括:

[0016] 将所述目标唤醒词确定为正例词,将所述至少一个冗余命令词确定为负例词;

[0017] 利用深度神经网络对所述正例词和所述负例词进行训练,得到所述唤醒词识别模型。

[0018] 在该实施例中,冗余命令词即说这些词时一定不会唤醒语音识别功能。这样,通过将目标唤醒词作为正例,冗余命令词作为负例进行训练,得到的唤醒词识别模型可以提高唤醒词识别的准确率,降低误唤醒。

[0019] 在一个实施例中,所述根据所述唤醒词识别模型确定所述当前语音信息属于所述

目标唤醒词的置信度,包括:

[0020] 获取所述当前语音信息对应的当前音频帧特征序列信息;

[0021] 将所述当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围;

[0022] 计算每个音频帧的范围中每个音频帧的声学似然评分,并从所述每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;

[0023] 根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度。

[0024] 在该实施例中,将当前语音帧特征序列信息与目标唤醒词对应的目标识别模型序列进行对齐强制处理,进而从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围,并从该范围中选取预设数量的目标音频帧,通过唤醒词识别模型和目标音频帧的特征信息确定当前语音信息属于目标唤醒词的置信度,进而确定是否唤醒语音唤醒功能。这样,有效地减少了语音唤醒的计算量,从而能应用于存在低计算资源限制的语音设备中。

[0025] 在一个实施例中,所述根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度,包括:

[0026] 使用所述唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;

[0027] 从与每个音素对应的多个声学后验得分中选取最大声学后验得分;

[0028] 根据所有音素的最大声学后验得分计算所述当前语音信息属于所述目标唤醒词的置信度。

[0029] 在该实施例中,将选出的全部目标语音帧对应的特征信息使用唤醒词识别模型计算对应的声学后验得分,对每个音素的多个声学后验得分选择其中的最大后验得分,并使用全部音素的最大后验得分计算得到当前语音信息属于目标唤醒词的置信度,这样,只需要计算目标音频帧的声学后验得分,无需计算所有的音频帧的得分,在保证识别结果的准确性的基础上,大大减少了计算量,缩短了识别当前语音信息所需的时长,提升了用户的使用体验。

[0030] 根据本发明实施例的第二方面,提供一种语音识别功能的唤醒装置,包括:

[0031] 第一获取模块,用于获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;

[0032] 训练模块,用于利用神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;

[0033] 第二获取模块,用于获取用户输入的当前语音信息;

[0034] 第一确定模块,用于根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;

[0035] 第二确定模块,用于根据所述置信度确定是否允许唤醒所述语音识别功能。

[0036] 在一个实施例中,所述第二确定模块包括:

[0037] 第一处理子模块,用于当所述置信度大于预设置信度时,允许唤醒所述语音识别

功能；

[0038] 第二处理子模块,用于当所述置信度小于或者等于所述预设置信度时,禁止唤醒所述语音识别功能。

[0039] 在一个实施例中,所述训练模块包括:

[0040] 第一确定子模块,用于将所述目标唤醒词确定为正例词,将所述至少一个冗余命令词确定为负例词;

[0041] 训练子模块,用于利用神经网络对所述正例词和所述负例词进行训练,得到所述唤醒词识别模型。

[0042] 在一个实施例中,所述第一确定模块包括:

[0043] 获取子模块,用于获取所述当前语音信息对应的当前音频帧特征序列信息;

[0044] 定位子模块,用于将所述当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围;

[0045] 选取子模块,用于计算每个音频帧的范围中每个音频帧的声学似然评分,并从所述每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;

[0046] 第二确定子模块,用于根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度。

[0047] 在一个实施例中,所述第二确定子模块用于:

[0048] 使用所述唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;

[0049] 从与每个音素对应的多个声学后验得分中选取最大声学后验得分;

[0050] 根据所有音素的最大声学后验得分计算所述当前语音信息属于所述目标唤醒词的置信度。

[0051] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本发明。

[0052] 本发明的其它特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本发明而了解。本发明的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

[0053] 下面通过附图和实施例,对本发明的技术方案做进一步的详细描述。

## 附图说明

[0054] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本发明的实施例,并与说明书一起用于解释本发明的原理。

[0055] 图1是根据一示例性实施例示出的一种语音识别功能的唤醒方法的流程图。

[0056] 图2是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S105的流程图。

[0057] 图3是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S102的流程图。

[0058] 图4是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S104的流

程图。

[0059] 图5是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S404的流程图。

[0060] 图6是根据一示例性实施例示出的一种语音识别功能的唤醒装置的框图。

[0061] 图7是根据一示例性实施例示出的一种语音识别功能的唤醒装置中第二确定模块的框图。

[0062] 图8是根据一示例性实施例示出的一种语音识别功能的唤醒装置中训练模块的框图。

[0063] 图9是根据一示例性实施例示出的一种语音识别功能的唤醒装置中第一确定模块的框图。

### 具体实施方式

[0064] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0065] 图1是根据一示例性实施例示出的一种语音识别功能的唤醒方法的流程图。该语音唤醒方法应用于终端设备中,该终端设备可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等任一具有语音控制功能的设备。如图1所示,该方法包括步骤S101-S105:

[0066] 在步骤S101中,获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词。

[0067] 例如,空调的唤醒词可以为空调空调,冗余命令词可以为:空调开机,空调关机,升高温度,降低温度,十六度,十七度,十八度,十九度,二十度,二十一度,二十二度,二十三度,二十四度,二十五度,二十六度,二十七度,二十八度,二十九度,三十度,增大风速,减小风速,打开扫风,关闭扫风,制冷模式,加热模式。

[0068] 步骤S102,利用深度神经网络对目标唤醒词和至少一个冗余命令词进行训练,得到唤醒词识别模型;

[0069] 步骤S103,获取用户输入的当前语音信息;

[0070] 步骤S104,根据唤醒词识别模型确定当前语音信息属于目标唤醒词的置信度;

[0071] 步骤S105,根据置信度确定是否允许唤醒语音识别功能。

[0072] 在该实施例中,利用目标唤醒词和冗余命令词进行训练,得到唤醒词识别模型,进而通过该唤醒词识别模型确定当前语音信息属于目标唤醒词的置信度,这样可以保证唤醒词识别的准确率,从而提高语音识别功能唤醒的准确率,降低误唤醒。

[0073] 图2是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S105的流程图。

[0074] 在一个实施例中,上述步骤S105包括步骤S201-S202:

[0075] 步骤S201,当置信度大于预设置信度时,允许唤醒语音识别功能;

[0076] 步骤S202,当置信度小于或者等于预设置信度时,禁止唤醒语音识别功能。

[0077] 在该实施例中,在当前语音信息属于目标唤醒词的置信度大于预设置信度时,才

允许唤醒语音唤醒功能,从而降低语音唤醒功能误唤醒。

[0078] 图3是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S102的流程图。

[0079] 如图3所示,在一个实施例中,上述步骤S102包括步骤S301-S302:

[0080] 步骤S301,将目标唤醒词确定为正例词,将至少一个冗余命令词确定为负例词;

[0081] 步骤S302,利用深度神经网络对正例词和负例词进行训练,得到唤醒词识别模型。

[0082] 在该实施例中,冗余命令词即说这些词时一定不会唤醒语音识别功能。这样,通过将目标唤醒词作为正例,冗余命令词作为负例进行训练,得到的唤醒词识别模型可以提高唤醒词识别的准确率,降低误唤醒。

[0083] 图4是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S104的流程图。

[0084] 如图4所示,在一个实施例中,上述步骤S104包括步骤S401-S404:

[0085] 步骤S401,获取当前语音信息对应的当前音频帧特征序列信息;

[0086] 步骤S402,将当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从当前音频帧序列信息中定位出与目标识别模型序列信息中的每个音素对应的音频帧的范围;

[0087] 步骤S403,计算每个音频帧的范围中每个音频帧的声学似然评分,并从每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;

[0088] 步骤S404,根据唤醒词识别模型和目标音频帧对应的特征信息,确定当前语音信息属于目标唤醒词的置信度。

[0089] 在该实施例中,将当前语音帧特征序列信息与目标唤醒词对应的目标识别模型序列进行对齐强制处理,进而从当前音频帧序列信息中定位出与目标识别模型序列信息中的每个音素对应的音频帧的范围,并从该范围中选取符合预设数量的目标音频帧,通过唤醒词识别模型和目标音频帧的特征信息确定当前语音信息属于目标唤醒词的置信度,进而确定是否唤醒语音唤醒功能。这样,有效地减少了语音唤醒的计算量,从而能应用于存在低计算资源限制的语音设备中。

[0090] 图5是根据一示例性实施例示出的一种语音识别功能的唤醒方法中步骤S404的流程图。

[0091] 如图5所示,在一个实施例中,上述步骤S404包括步骤S501-S503:

[0092] 步骤S501,使用唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;

[0093] 步骤S502,从与每个音素对应的多个声学后验得分中选取最大声学后验得分;

[0094] 步骤S503,根据所有音素的最大声学后验得分计算当前语音信息属于目标唤醒词的置信度。

[0095] 在该实施例中,将选出的全部目标语音帧对应的特征信息使用唤醒词识别模型计算对应的声学后验得分,对每个音素的多个声学后验得分选择其中的最大后验得分,并使用全部音素的最大的后验得分计算得到当前语音信息属于目标唤醒词的置信度,这样,只需要计算目标音频帧的声学后验得分,无需计算所有的音频帧的得分,在保证识别结果的准确性的基础上,大大减少了计算量,缩短了识别当前语音信息所需的时长,提升了用户的使



用体验。

[0096] 下述为本发明装置实施例,可以用于执行本发明方法实施例。

[0097] 图6是根据一示例性实施例示出的一种语音识别功能的唤醒装置的框图,该装置可以通过软件、硬件或者两者的结合实现成为终端设备的部分或者全部。如图6所示,该语音识别功能的唤醒装置包括:

[0098] 第一获取模块61,用于获取唤醒语音识别功能的目标唤醒词和至少一个冗余命令词;

[0099] 训练模块62,用于利用神经网络对所述目标唤醒词和所述至少一个冗余命令词进行训练,得到唤醒词识别模型;

[0100] 第二获取模块63,用于获取用户输入的当前语音信息;

[0101] 第一确定模块64,用于根据所述唤醒词识别模型确定所述当前语音信息属于所述目标唤醒词的置信度;

[0102] 第二确定模块65,用于根据所述置信度确定是否允许唤醒所述语音识别功能。

[0103] 在该实施例中,利用目标唤醒词和冗余命令词进行训练,得到唤醒词识别模型,进而通过该唤醒词识别模型确定当前语音信息属于目标唤醒词的置信度,这样可以保证唤醒词识别的准确率,从而提高语音识别功能唤醒的准确率,降低误唤醒。

[0104] 图7是根据一示例性实施例示出的一种语音识别功能的唤醒装置中第二确定模块的框图。

[0105] 如图7所示,在一个实施例中,所述第二确定模块65包括:

[0106] 第一处理子模块71,用于当所述置信度大于预设置信度时,允许唤醒所述语音识别功能;

[0107] 第二处理子模块72,用于当所述置信度小于或者等于所述预设置信度时,禁止唤醒所述语音识别功能。

[0108] 在该实施例中,在当前语音信息属于目标唤醒词的置信度大于预设置信度时,才允许唤醒语音唤醒功能,从而降低语音唤醒功能误唤醒。

[0109] 图8是根据一示例性实施例示出的一种语音识别功能的唤醒装置中训练模块的框图。

[0110] 如图8所示,在一个实施例中,所述训练模块62包括:

[0111] 第一确定子模块81,用于将所述目标唤醒词确定为正例词,将所述至少一个冗余命令词确定为负例词;

[0112] 训练子模块82,用于利用神经网络对所述正例词和所述负例词进行训练,得到所述唤醒词识别模型。

[0113] 在该实施例中,冗余命令词即说这些词时一定不会唤醒语音识别功能。这样,通过将目标唤醒词作为正例,冗余命令词作为负例进行训练,得到的唤醒词识别模型可以提高唤醒词识别的准确率,降低误唤醒。

[0114] 图9是根据一示例性实施例示出的一种语音识别功能的唤醒装置中第一确定模块的框图。

[0115] 如图9所示,在一个实施例中,所述第一确定模块64包括:

[0116] 获取子模块91,用于获取所述当前语音信息对应的当前音频帧特征序列信息;

[0117] 定位子模块92,用于将所述当前音频帧序列信息与目标唤醒词对应的目标识别模型序列信息进行强制对齐处理,以从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围;

[0118] 选取子模块93,用于计算每个音频帧的范围中每个音频帧的声学似然评分,并从所述每个音频帧的范围中选取符合预设要求的预设数量的目标音频帧;

[0119] 第二确定子模块94,用于根据所述唤醒词识别模型和目标音频帧对应的特征信息,确定所述当前语音信息属于所述目标唤醒词的置信度。

[0120] 在该实施例中,将当前语音帧特征序列信息与目标唤醒词对应的目标识别模型序列进行对齐强制处理,进而从所述当前音频帧序列信息中定位出与所述目标识别模型序列信息中的每个音素对应的音频帧的范围,并从该范围中选取预设数量的目标音频帧,通过唤醒词识别模型和目标音频帧的特征信息确定当前语音信息属于目标唤醒词的置信度,进而确定是否唤醒语音唤醒功能。这样,有效地减少了语音唤醒的计算量,从而能应用于存在低计算资源限制的语音设备中。

[0121] 在一个实施例中,所述第二确定子模块94用于:

[0122] 使用所述唤醒词识别模型计算与每个音素对应的多个目标音频帧中每个目标音频帧的声学后验得分;

[0123] 从与每个音素对应的多个声学后验得分中选取最大声学后验得分;

[0124] 根据所有音素的最大声学后验得分计算所述当前语音信息属于所述目标唤醒词的置信度。

[0125] 在该实施例中,将选出的全部目标语音帧对应的特征信息使用唤醒词识别模型计算对应的声学后验得分,对每个音素的多个声学后验得分选择其中的最大后验得分,并使用全部音素的最大后验得分计算得到当前语音信息属于目标唤醒词的置信度,这样,只需要计算目标音频帧的声学后验得分,无需计算所有的音频帧的得分,在保证识别结果的准确性的基础上,大大减少了计算量,缩短了识别当前语音信息所需的时长,提升了用户的使用体验。

[0126] 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和光学存储器等)上实施的计算机程序产品的形式。

[0127] 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。

[0128] 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。

[0129] 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。

[0130] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

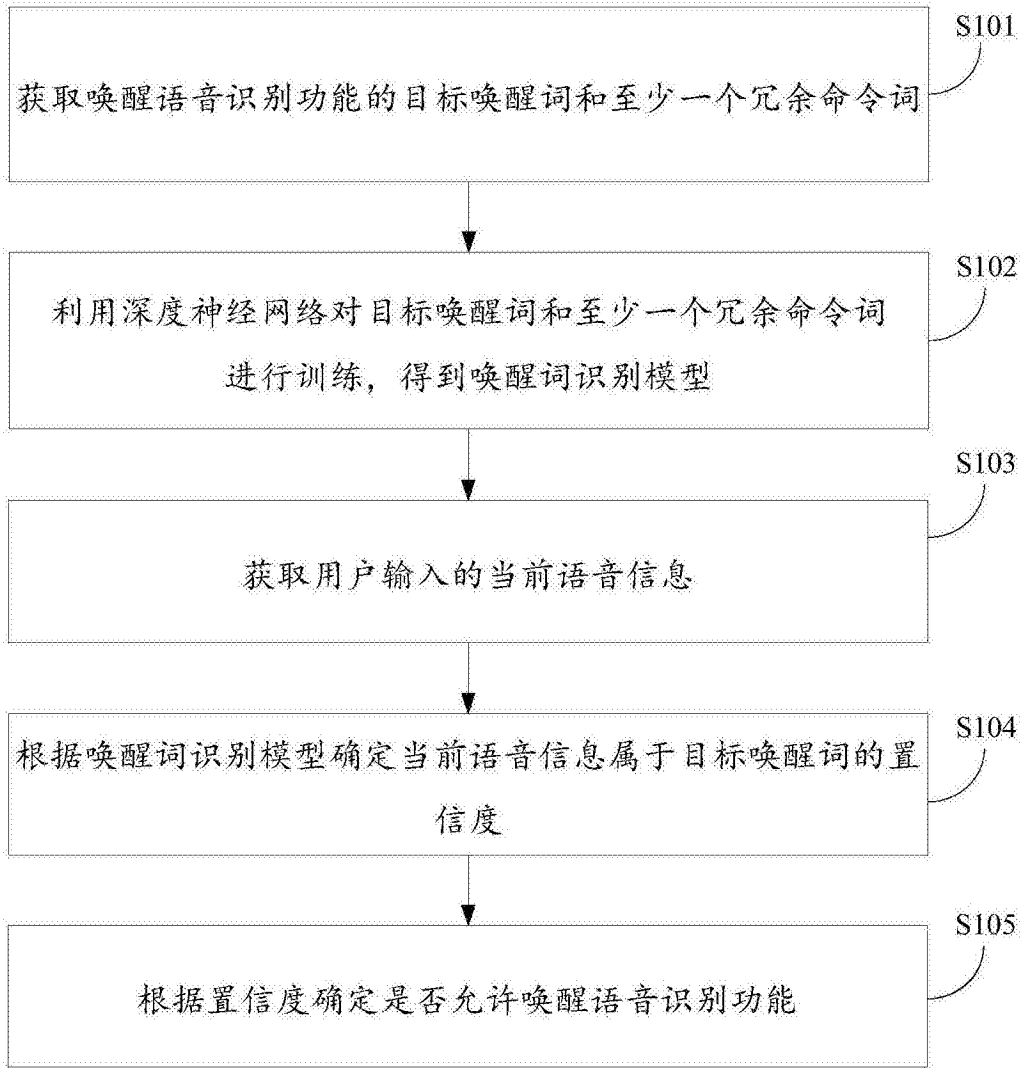


图1

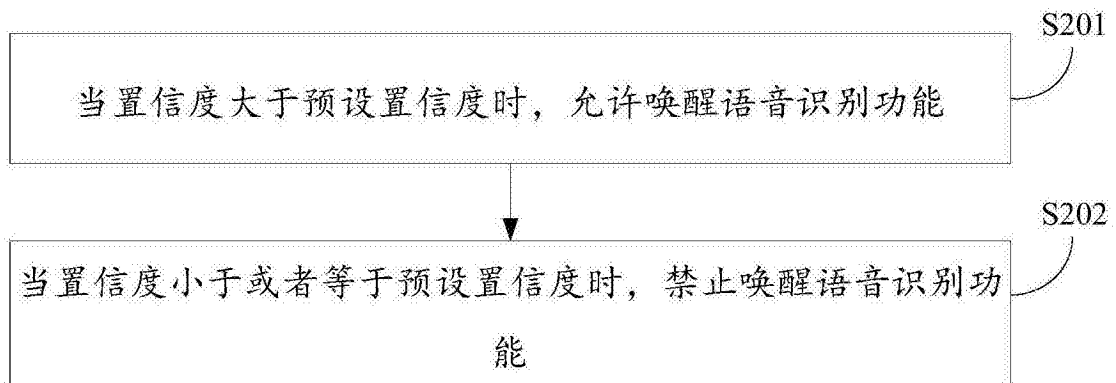


图2

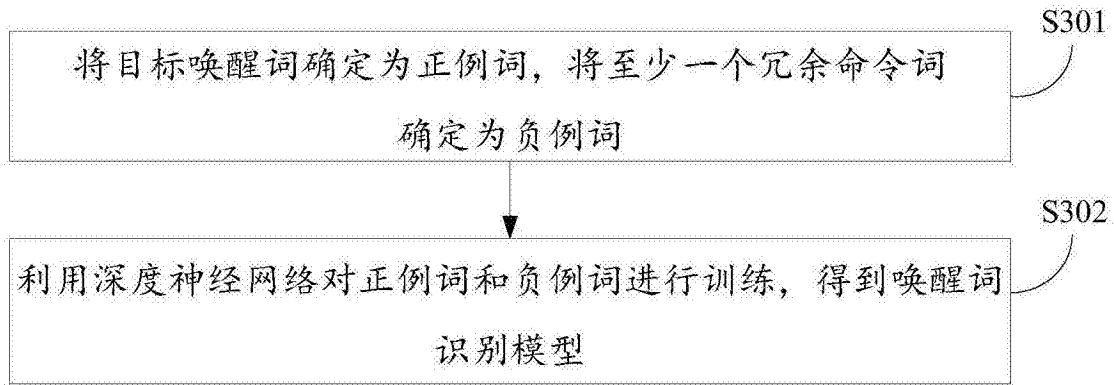


图3

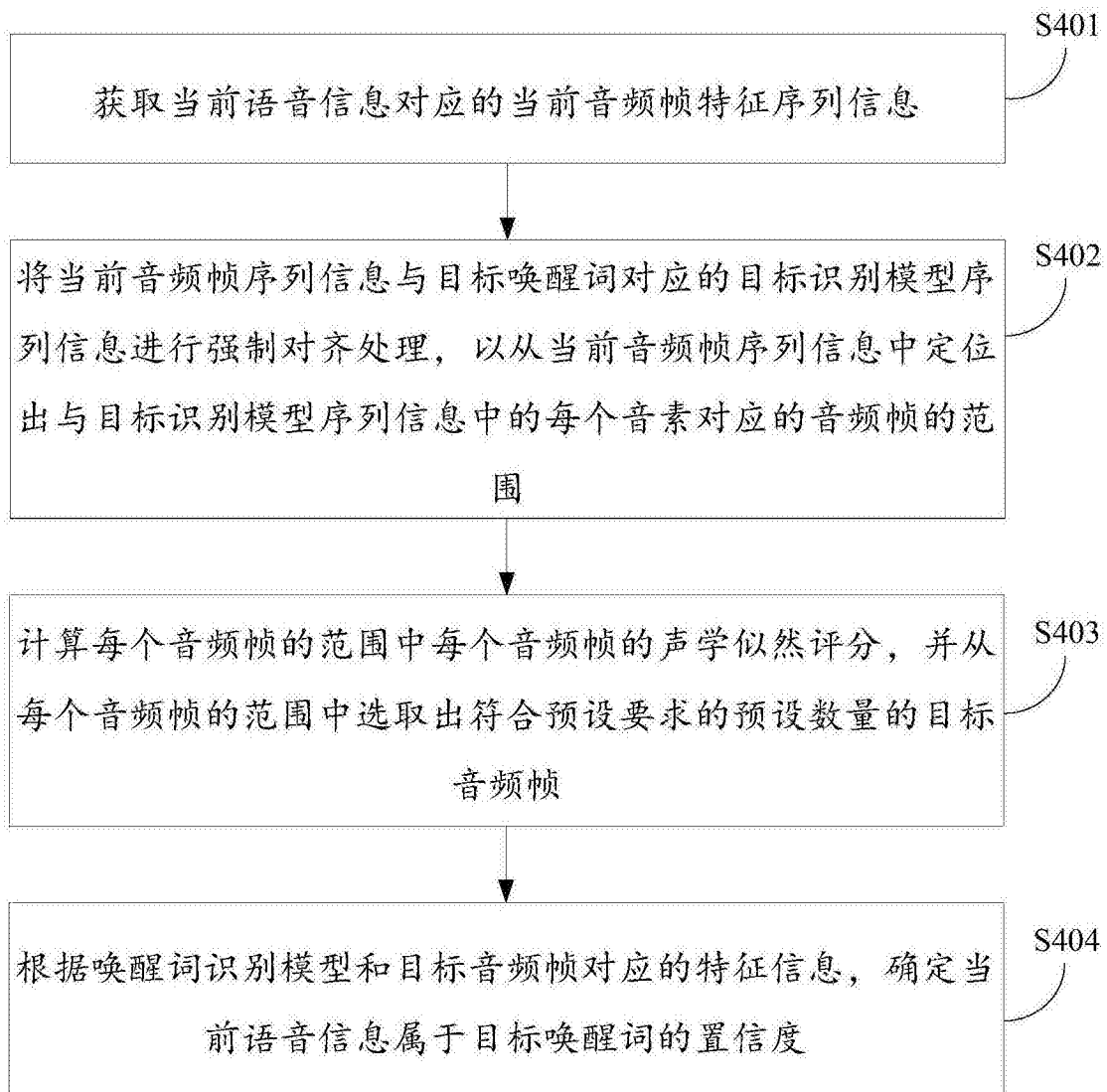


图4

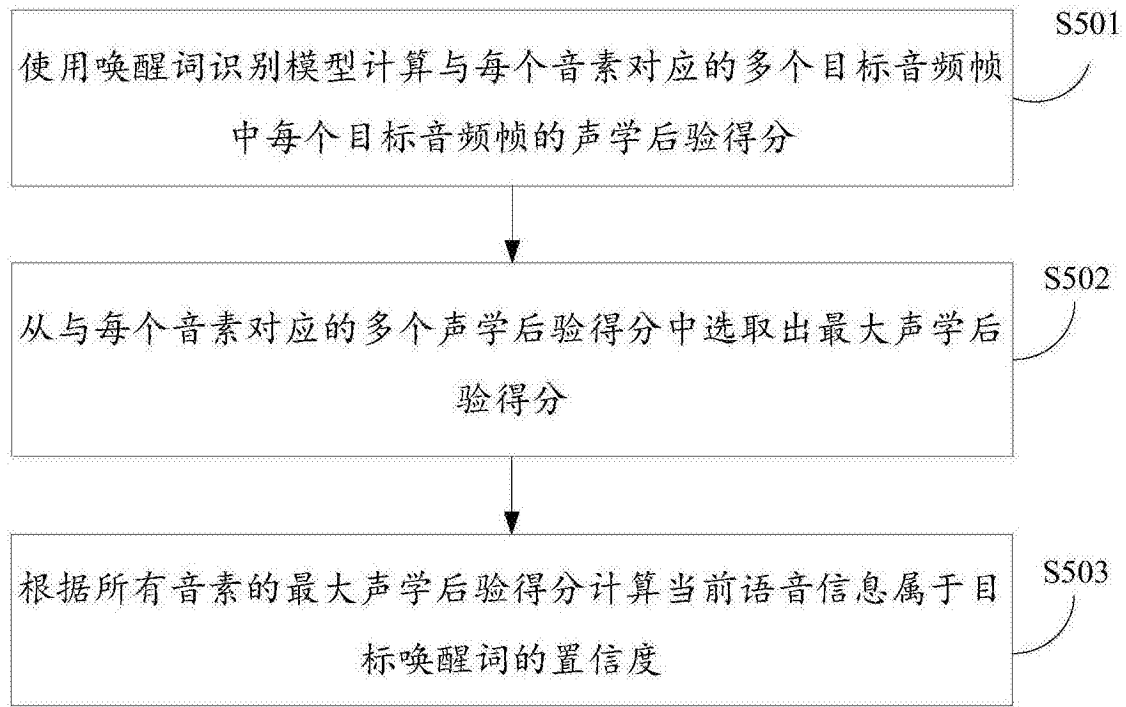


图5

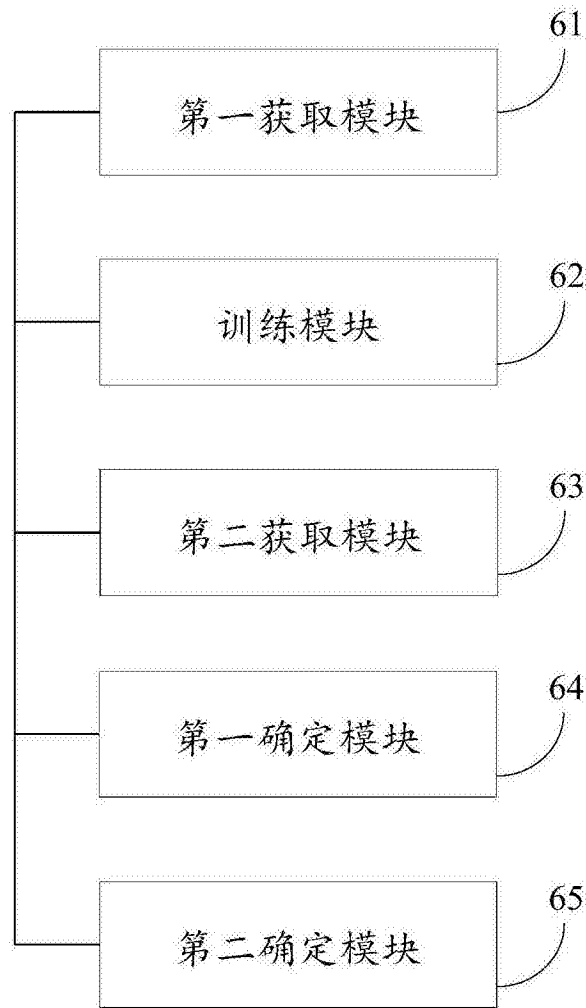


图6

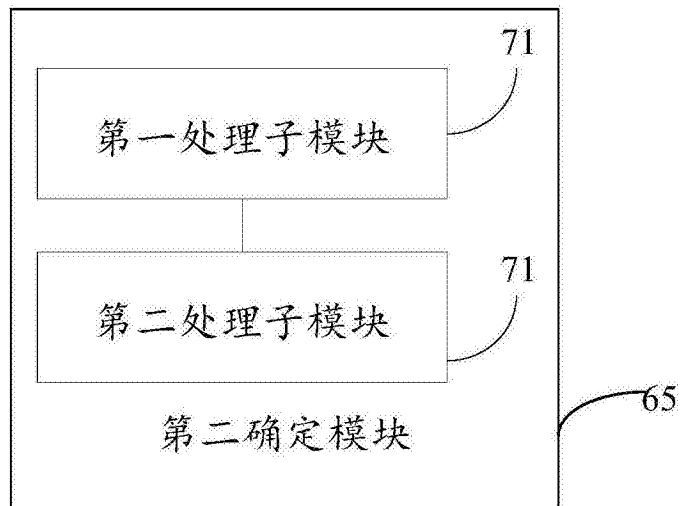


图7

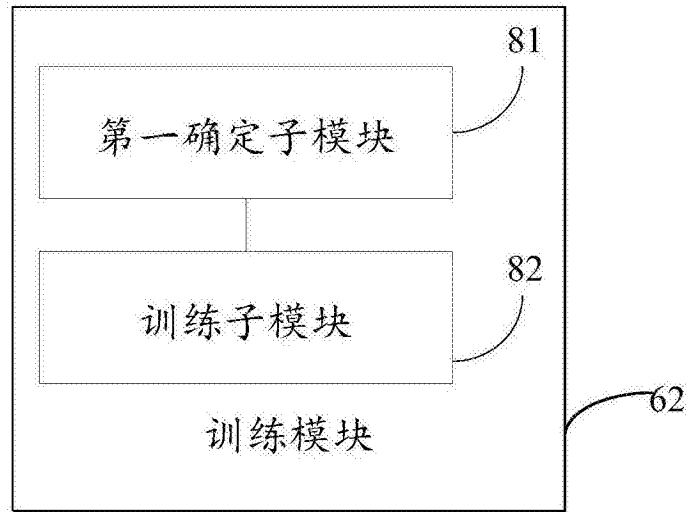


图8

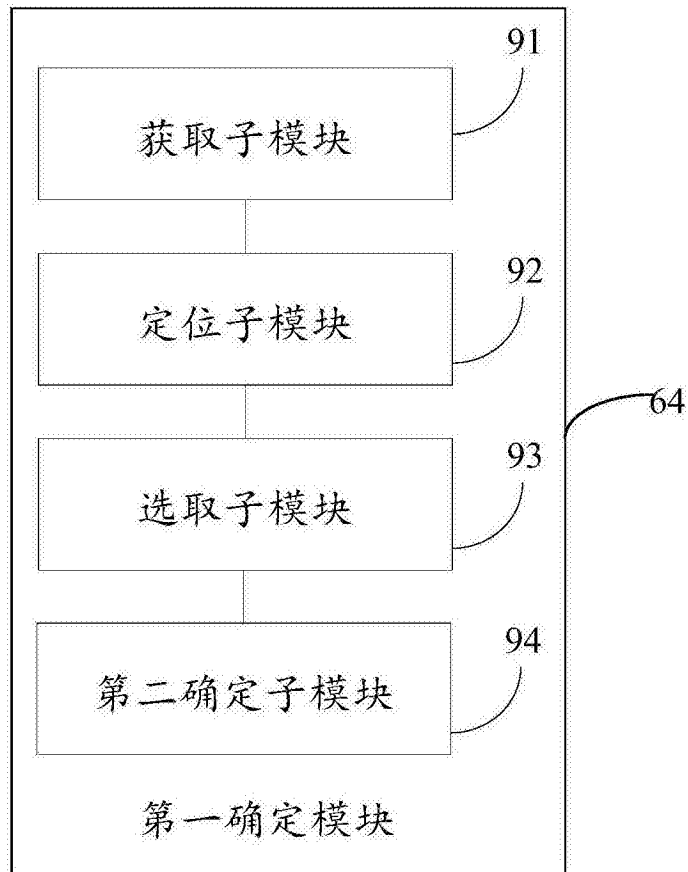


图9