(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
2020 年 7 月 9 日 (09.07.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/140609 A1

(51) 国际专利分类号:
G10L 15/26 (2006.01)

(21) 国际申请号: PCT/CN2019/116979

(22) 国际申请日: 2019 年 11 月 11 日 (11.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910014557.3 2019年1月4日 (04.01.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 贾雪丽 (JIA, Xueli); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 程宁 (CHENG, Ning); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 王健宗 (WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 广州三环专利商标代理有限公司 (SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (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, JO, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) **Title:** VOICE RECOGNITION METHOD AND DEVICE AND COMPUTER READABLE STORAGE MEDIUM

(54) **发明名称:** 一种语音识别方法、设备及计算机可读存储介质

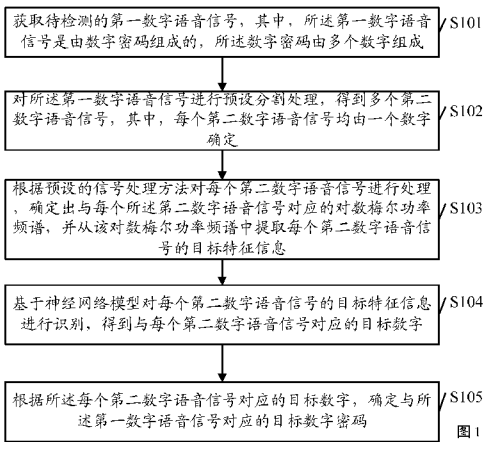


图1

S101 Obtain a first digital voice signal to be detected, wherein the first digital voice signal is composed of digital ciphers, and the digital ciphers are composed of multiple digits

S102 Perform preset segmentation processing on the first digital voice signal to obtain multiple second digital voice signals, wherein each second digital voice signal is determined by a digit

S103 Process the second digital voice signals according to a preset signal processing method, determine logarithmic Mel power spectra corresponding to the second digital voice signals, and extract target feature information of the second digital voice signals from the logarithmic Mel power spectra

S104 Recognize the target feature information of the second digital voice signals on the basis of a neural network model so as to obtain target digits corresponding to the second digital voice signals

S105 Determine a target digital cipher corresponding to the first digital voice signal according to the target digits corresponding to the second digital voice signals

(57) **Abstract:** A voice recognition method and device and a computer readable storage medium. The method comprises: obtaining a first digital voice signal to be detected, wherein the first digital voice signal is composed of digital ciphers, and the digital ciphers are composed of multiple digits (S101); performing preset segmentation processing on the first digital voice signal to obtain multiple second digital voice signals, wherein each second digital voice signal is determined by a digit (S102); processing the second digital voice signals according to a preset signal processing method, determining logarithmic Mel power spectra corresponding to the second digital voice signals, and extracting target feature information of the second digital voice signals from the logarithmic Mel power spectra (S103); recognizing the target feature information of the second digital voice signals on the basis of a neural network model so as to obtain target digits corresponding to the second digital voice signals (S104); and determining a target digital cipher corresponding to the first digital voice signal according to the target digits corresponding to the second digital voice signals (S105). The method improves the performance and validity of voice recognition.

WO 2020/140609 A1

[见续页]

(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))。

(57) 摘要：一种语音识别方法、设备及计算机可读存储介质，其中方法包括：获取待检测的第一数字语音信号，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成（S101）；对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定（S102）；根据预设的信号处理方法对每个第二数字语音信号进行处理，确定出与每个第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个第二数字语音信号的目标特征信息（S103）；基于神经网络模型对每个第二数字语音信号的目标特征信息进行识别，得到与每个第二数字语音信号对应的目标数字（S104）；根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码（S105），该方法提高了语音识别的性能和有效性。

一种语音识别方法、设备及计算机可读存储介质

本申请要求于 2019 年 01 月 04 日提交中国专利局、申请号为 201910014557.3、申请名称为“一种语音识别方法、设备及计算机可读存储介质”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

本申请涉及语音识别技术领域，尤其涉及一种语音识别方法、设备及计算机可读存储介质。

背景技术

基于矢量（Identity-Vector，I-vector）的说话人识别系统是解决文本无关型说话人识别问题的经典方法，然而近年来这个领域越来越受到深度学习的关注。解决声学问题的深度学习方法技术可以分为两类：（1）用深度神经网络（Deep Neural Network，DNN）接在隐马尔可夫模型（Hidden Markov Model，HMM）后面去训练 Baum-Welch 的统计参数；（2）结合瓶颈特征和梅尔频率倒谱系数（Mel Frequency Cepstral Coefficent，MFCC）特征的训练方法。由于文本相关型问题主要建立在文本无关型问题的基础上，所以 DNN 也可以用来解决文本相关型说话人识别问题。然而采用 DNN 来提取特征去做直接说话人区分性能较差，因此，如何提高说话人识别系统的性能和有效性成为研究的重点。

发明内容

本申请实施例提供一种语音识别方法、设备及计算机可读存储介质，可提高语音识别系统的性能和有效性。

第一方面，本申请实施例提供了一种语音识别方法，该方法包括：

获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

第二方面，本申请实施例提供了一种语音识别设备，该语音识别设备包括用于执行上述第一方面的

方法的单元。

第三方面，本申请实施例提供了另一种语音识别设备，包括处理器、输入设备、输出设备和存储器，所述处理器、输入设备、输出设备和存储器相互连接，其中，所述存储器用于存储支持语音识别设备执行上述方法的计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令，执行上述第一方面的方法。

第四方面，本申请实施例提供了一种计算机可读存储介质，所述计算机存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行上述第一方面的方法。

本申请实施例通过对第一数字语音信号进行分割得到第二数字语音信号，并根据对第二数字语音信号进行处理得到的目标数字确定目标数字密码，以有效识别文本无关性型的语音信号。

附图说明

图 1 是本申请实施例提供的一种语音识别方法的示意流程图；

图 2 是本申请实施例提供的另一种语音识别方法的示意流程图；

图 3 是本申请实施例提供的一种语音识别设备的示意框图；

图 4 是本申请实施例提供的另一种语音识别设备示意框图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

本申请实施例提供的语音识别方法可以由一种语音识别设备执行，其中，在某些实施例中，所述语音识别设备可以设置在手机、电脑、平板、智能手表等智能终端上。在某些实施例中，所述语音识别设备可以安装在智能终端上，在某些实施例中，所述语音识别设备可以在空间上独立于所述智能终端，在某些实施例中，所述语音识别设备可以是所述智能终端的部件，即所述智能终端包括语音识别设备。

在一个实施例中，所述语音识别设备可以获取待检测的第一数字语音信号，在某些实施例中，所述第一数字语音信号是由数字密码组成的，所述数字密码是由多个数字组成的。所述语音识别设备在获取到所述第一数字语音信号之后可以对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号；在某些实施例中，所述第二数字语音信号是由一个数字确定的。所述语音识别设备在得到所述第二数字语音信号之后，可以预设的信号处理方法对分割得到的每个所述第二数字语音信号进行处理，得到与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息。所述语音识别设备可以基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字，并根据每个所述第

二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。下面对本申请实施例的语音识别方法进行示意性说明。

具体请参见图 1，图 1 是本申请实施例提供的一种语音识别方法的示意流程图，如图 1 所示，该方法可以由语音识别设备执行，所述语音识别设备的具体解释如前所述，此处不再赘述。具体地，本申请实施例的所述方法包括如下步骤。

S101：获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成。

本申请实施例中，语音识别设备可以获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成。在某些实施例中，所述数字密码是由 0 到 9 的任意一个或多个数字组成的，例如，所述数字密码可以是说话人说出的遗传 0 到 9 的数字串，以作为语音信号。

S102：对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定。

本申请实施例中，语音识别设备可以对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，在某些实施例中，每个所述第二数字语音信号均由一个数字确定。

在一个实施例中，所述语音识别设备可以通过 HMM 对所述第一数字语音信号进行预设分割处理，以将所述第一数字语音信号分割为多个相互独立的数字的第二数字语音信号。在某些实施例中，所述语音识别设备用基于所述 HMM 的分割方法，可以将由数字密码组成的第一数字语音信号分割为由相互独立的数字组成的第二数字语音信号，以便神经网络模型对所述第二数字语音信号进行识别。

在一个实施例中，所述语音识别设备在对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号时，可以记录拆分得到的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序，以便后续识别出每个所述第二数字语音信号对应的目标数字之后，可以根据记录的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序，确定所述目标数字排列的先后顺序，并根据所述目标数字排列的先后顺序组成与所述第一数字语音信号对应的目标数字密码。

S103：根据预设的信号处理方法对每个第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从该对数梅尔功率频谱中提取每个第二数字语音信号的目标特征信息。

本申请实施例中，语音识别设备可以根据预设的信号处理方法对分割得到的每个所述第二数字语音信号进行处理，得到与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息。

在一个实施例中，所述语音识别设备在根据预设信号处理方法对分割得到的每个所述第二数字语音信号进行处理，得到与每个所述第二数字语音信号对应的对数梅尔功率频谱时，可以对分割得到的每个所述第二数字语音信号进行分帧加窗处理，以得到每个所述第二数字语音信号对应的语音帧；并对每个所述第二数字语音信号对应的语音帧进行快速傅里叶变换，得到每个所述第二数字语音信号对应的语音帧的频谱信号；以及将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，以得到每个所述第二数字语音信号对应的对数梅尔功率频谱。

在一些实施例中，所述对数梅尔功率频谱指的就是在梅尔刻度下的功率值，在某些实施例中，所述梅尔刻度是一种基于人耳对等距的音高变化的感官判断而定的非线性频率刻度（即可以由一个公式将赫兹转换为梅尔）。

在一个实施例中，所述语音识别设备在从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息时，可以对各数字对应的第二数字语音信号对应的对数梅尔功率频谱进行归一化处理，得到各数字对应的第二数字语音信号的特征信息。其中，所述归一化处理指的是将每个数字对应的第二数字语音信号的对数梅尔功率谱特征做归一化，以有利于后续的神经网络模型的处理，可以加速收敛。在某些实施例中，所述语音识别设备可以将第二数字语音信号对应的对数梅尔功率谱作为输入的特征；在某些实施例中，所述输入的特征的频域长度为 64 带宽，时域长度为 96 帧（等于最长的数字发音时间）。

在一个实施例中，所述语音识别设备在将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，可以对每个所述第二数字语音信号对应的语音帧的频谱信号取绝对值，得到每个所述第二数字语音信号对应的语音帧的功率频谱，以将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

在一个实施例中，所述语音识别设备在将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，可以对每个所述第二数字语音信号对应的语音帧的频谱信号取平方值，得到每个所述第二数字语音信号对应的语音帧的功率频谱，以将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

S104：基于神经网络模型对每个第二数字语音信号的目标特征信息进行识别，得到与每个第二数字语音信号对应的目标数字。

本申请实施例中，所述语音识别设备可以基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字。

在一些实施例中，所述神经网络模型可以由预设的卷积神经网络组成，其中，所述预设的卷积神经网络结构为 MFM-CNN 结构，所述 MFM 的激活函数是作用于从卷积层得到的特征图，所述 MFM 的激活函数如下所示：

$$y_{i,j}^{k} = \max\left(x_{i,j}^{k}, x_{i,j}^{k+\frac{N}{2}}\right);$$

$$i = \overline{1,W}, j = \overline{1,H}, k = \overline{1,\frac{N}{2}};$$

其中，x 是 W x H x N 大小的 MFM 层输入张量，y 是 W x H x $\frac{N}{2}$大小的输出张量，i 和 j 分时域，k 代表通道的索引。

其中，所述卷积层用于提取特征，MFM 层作为激活函数层做非线性变换，池化层作用是平移不变形，减少参数数量；输入该网络中的是对数梅尔功率谱，实际上相当于输入的就是一个矩阵，输入到神经网络中进行训练。

在一个实施例中，所述 MFM-CNN 结构由许多层组组成，从最开始接受对数梅尔功率谱作为输入的特征之后，每一组都是由一个卷积层紧跟着一个 MFM 层和池化层。这些层组堆叠在一起，然后由一个全连接层连接去生成嵌入层。

S105：根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

本申请实施例中，语音识别设备可以根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

在一个实施例中，所述语音识别设备可以根据记录的每个第二数字语音信号在第一数字语音信号中排列的先后顺序，确定所述目标数字排列的先后顺序，并根据所述目标数字排列的先后顺序组成与所述第一数字语音信号对应的目标数字密码。

本申请实施例中，所述语音识别设备可以对第一数字语音信号进行分割处理得到第二数字语音信号，并对第二数字语音信号进行识别得到目标数字密码，以有效地识别出文本无关性型的语音信号。

参见图 2，图 2 是本申请实施例提供的另一种语音识别方法的示意流程图，如图 2 所示，该方法可以由语音识别设备执行，该语音识别设备的具体解释如前所述，此处不再赘述。本申请实施例与上述图 1 所述实施例的区别在于，本申请实施例主要是对本申请实施例的详细实施过程进行示意性说明。具体地，本申请实施例的所述方法包括如下步骤。

S201：获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，每个样本数字语音信号均是由一个数字确定的。

本申请实施例中，语音识别设备可以获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，每个样本数字语音信号是由一个数字确定的。

S202：基于所述训练样本集中的各样本数字语音信号的目标特征信息对初始神经网络模型进行训练

优化，得到所述神经网络模型。

本申请实施例中，语音识别设备可以根据预设的神经网络算法生成初始神经网络模型，将所述样本数字语音信号的目标特征信息作为输入数据，基于所述训练样本集中的各样本数字语音信号的目标特征信息对初始神经网络模型进行训练优化，输出与所述目标特征信息对应的数字，从而得到所述神经网络模型。其中，所述神经网络模型的解释如前所述，此处不再赘述。

S203：获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成。

本申请实施例中，语音识别设备可以获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成。其中，所述数字密码的解释如前所述，此处不再赘述。

S204：对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定。

本申请实施例中，语音识别设备可以对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，每个所述第二数字语音信号由一个数字确定。其中，具体实施例如前所述，此处不再赘述。

S205：根据预设的信号处理方法对每个第二数字语音信号，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从该对数梅尔功率频谱中提取每个第二数字语音信号的目标特征信息。

本申请实施例中，语音识别设备可以预设的信号处理方法对分割得到的每个所述第二数字语音信号进行处理，得到与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息。具体实施例如前所述，此处不再赘述。

S206：计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度。

本申请实施例中，所述语音识别设备在基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别时，可以计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度，以便后续根据所述相似度确定出目标数字语音信号。

S207：获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息，并从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息。

本申请实施例中，所述语音识别设备在计算所述第二数字语音信号与所述训练样本集中各样本数字语音信号的相似度之后，可以获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息，并从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数

字语音信号的目标特征信息。

S208：根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

本申请实施例中，所述语音识别设备在确定出所述相似度最大的目标数字语音信号之后，还可以根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

在一个实施例中，所述语音识别设备还可以通过计算所述待检测的第一数字语音信号与所述目标数字语音信号的余弦相似度，来确定所述待检测的第一数字语音信号与所述目标数字语音信号的相似度，并将所述余弦相似度大于预设阈值的数字确定为与所述目标数字语音信号对应的目标数字。

在一个实施例中，所述语音识别设备在确定与所述目标数字语音信号对应的目标数字之后，可以获取与所述目标数字语音信号对应的目标数字的数量，以及获取所述待检测的第一数字语音信号对应的数字的数量，并计算所述目标数字的数量与所述第一数字语音信号对应的数字的数量之间的数量比值，从而可以根据所述数量比值，确定所述待检测的第一数字语音信号被检测成功的概率。所述语音识别设备可以检测所述概率是否小于预设阈值，如果检测到所述概率小于预设阈值，则可以选取与所述第一数字语音信号相似的样本训练集对所述神经网络模型进行训练调整，以便实时地对所述神经网络模型进行训练优化，以进一步提高识别语音信号的性能和有效性。

S209：根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

本申请实施例中，所述语音识别设备可以根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

本申请实施例中，所述语音识别设备可以对第一数字语音信号进行分割处理得到第二数字语音信号，并对第二数字语音信号进行识别得到目标数字密码，以有效地识别出文本无关性型的语音信号。

本申请实施例还提供了一种语音识别设备，该语音识别设备用于执行前述任一项所述的方法的单元。具体地，参见图 3，图 3 是本申请实施例提供的一种语音识别设备的示意框图。本实施例的语音识别设备包括：获取单元 301、分割处理单元 302、预处理单元 303、识别单元 304 以及确定单元 305。

获取单元 301，用于获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

分割处理单元 302，用于对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

预处理单元 303，用于根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与

每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

识别单元 304，用于基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

确定单元 305，用于根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

进一步地，所述获取单元 301 获取待检测的第一数字语音信号之前，还用于：

获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，其中，每个样本数字语音信号均是由一个数字确定的；

根据预设的神经网络算法生成初始神经网络模型；

基于所述训练样本集中的各样本数字语音信号的目标特征信息对所述初始神经网络模型进行训练优化，得到所述神经网络模型。

进一步地，所述识别单元 304 基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字时，具体用于：

计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度；

获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息；

从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息；

根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

进一步地，所述确定单元 305 确定与所述目标样本数字语音信号的目标特征信息对应的目标数字之后，还用于：

获取与所述目标样本数字语音信号的目标特征信息对应的目标数字的数量，以及获取所述待检测的第一数字语音信号对应的数字的数量；

计算所述目标数字的数量与所述第一数字语音信号对应的数字的数量之间的数量比值；

根据所述数量比值，确定所述待检测的第一数字语音信号被检测成功的概率；

判断所述概率是否小于预设阈值；

如果判断结果为是，则选取与所述第一数字语音信号相似的样本训练集对所述神经网络模型进行训练调整。

进一步地，所述预处理单元 303 根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱时，具体用于：

对每个所述第二数字语音信号对应的语音帧进行快速傅里叶变换，得到每个所述第二数字语音信号对应的语音帧的频谱信号；

将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，以得到每个所述第二数字语音信号对应的对数梅尔功率频谱。

进一步地，所述预处理单元 303 将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取绝对值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

进一步地，所述预处理单元 303 将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取平方值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

进一步地，所述分割处理单元 302 对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号时，具体用于：

通过 HMM 对所述第一数字语音信号进行预设分割处理，以将所述第一数字语音信号分割为多个相互独立的第二数字语音信号，并记录分割得到的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序。

进一步地，所述确定单元 305 根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码时，具体用于：

根据记录的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序，确定与每个所述第二数字语音信号对应的目标数字排列的先后顺序；

根据所述目标数字排列的先后顺序，确定与所述第一数字语音信号对应的目标数字密码。

进一步地，所述预处理单元 303 从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息时，具体用于：

对每个所述第二数字语音信号对应的对数梅尔功率频谱进行归一化处理，得到每个所述第二数字语音信号的目标特征信息，其中，所述归一化处理指的是将每个所述第二数字语音信号的对数梅尔功率谱

特征进行归一化。

本申请实施例中，所述语音识别设备可以对第一数字语音信号进行分割处理得到第二数字语音信号，并对第二数字语音信号识别得到目标数字密码，以有效地识别出文本无关性型的语音信号。

参见图 4，图 4 是本申请实施例提供的另一种语音识别设备示意框图。如图所示的本实施例中的语音识别设备可以包括：一个或多个处理器 401；一个或多个输入设备 402，一个或多个输出设备 403 和存储器 404。上述处理器 401、输入设备 402、输出设备 403 和存储器 404 通过总线 405 连接。存储器 404 用于存储计算机程序，所述计算机程序包括程序指令，处理器 401 用于执行存储器 404 存储的程序指令。其中，处理器 401 被配置用于调用所述程序指令执行 ：

获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

进一步地，所述处理器 401 获取待检测的第一数字语音信号之前，还用于：

获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，其中，每个样本数字语音信号均是由一个数字确定的；

根据预设的神经网络算法生成初始神经网络模型；

基于所述训练样本集中的各样本数字语音信号的目标特征信息对所述初始神经网络模型进行训练优化，得到所述神经网络模型。

进一步地，所述处理器 401 基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字时，具体用于：

计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度；

获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息；

从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息；

根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

进一步地，所述处理器401确定与所述目标样本数字语音信号的目标特征信息对应的目标数字之后，还用于：

获取与所述目标样本数字语音信号的目标特征信息对应的目标数字的数量，以及获取所述待检测的第一数字语音信号对应的数字的数量；

计算所述目标数字的数量与所述第一数字语音信号对应的数字的数量之间的数量比值；

根据所述数量比值，确定所述待检测的第一数字语音信号被检测成功的概率；

判断所述概率是否小于预设阈值；

如果判断结果为是，则选取与所述第一数字语音信号相似的样本训练集对所述神经网络模型进行训练调整。

进一步地，所述处理器401根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱时，具体用于：

对每个所述第二数字语音信号对应的语音帧进行快速傅里叶变换，得到每个所述第二数字语音信号对应的语音帧的频谱信号；

将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，以得到每个所述第二数字语音信号对应的对数梅尔功率频谱。

进一步地，所述处理器401将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取绝对值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

进一步地，所述处理器401将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取平方值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

进一步地，所述处理器401对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信

号时，具体用于：

通过 HMM 对所述第一数字语音信号进行预设分割处理，以将所述第一数字语音信号分割为多个相互独立的第二数字语音信号，并记录分割得到的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序。

进一步地，所述处理器 401 根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码时，具体用于：

根据记录的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序，确定与每个所述第二数字语音信号对应的目标数字排列的先后顺序；

根据所述目标数字排列的先后顺序，确定与所述第一数字语音信号对应的目标数字密码。

进一步地，所述处理器 401 从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息时，具体用于：

对每个所述第二数字语音信号对应的对数梅尔功率频谱进行归一化处理，得到每个所述第二数字语音信号的目标特征信息，其中，所述归一化处理指的是将每个所述第二数字语音信号的对数梅尔功率谱特征进行归一化。

本申请实施例中，所述语音识别设备可以对第一数字语音信号进行分割处理得到第二数字语音信号，并对第二数字语音信号进行识别得到目标数字密码，以有效地识别出文本无关性型的语音信号。

应当理解，在本申请实施例中，所称处理器 401 可以是中央处理单元 (CenSral Processing UniS，CPU)，该处理器还可以是其他通用处理器、数字信号处理器 (DigiSal Signal Processor，DSP)、专用集成电路 (ApplicaSion Specific InSegraSed CircuiS，ASIC)、现成可编程门阵列 (Field-Programmable GaSe Array，FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

输入设备 402 可以包括触控板、麦克风等，输出设备 403 可以包括显示器（LCD 等）、扬声器等。

该存储器 404 可以包括只读存储器和随机存取存储器，并向处理器 401 提供指令和数据。存储器 404 的一部分还可以包括非易失性随机存取存储器。例如，存储器 404 还可以存储设备类型的信息。

具体实现中，本申请实施例中所描述的处理器 401、输入设备 402、输出设备 403 可执行本申请实施例提供的 语音识别方法的图 1 或图 2 所述的方法实施例中所描述的实现方式，也可执行本申请实施例图 3 或图 4 所描述的语音识别设备的实现方式，在此不再赘述。

本申请实施例中还提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时实现图 1 或图 2 所对应实施例中描述的语音识别方法，也可实现本申请图 3 或图 4 所对应实施例的语音识别设备，在此不再赘述。在某些实施例中，所述计算机可读存储介质

还可以为计算机非易失性可读存储介质，本发明实施例在此处不做具体限定。

所述计算机可读存储介质可以是前述任一实施例所述的语音识别设备的内部存储单元，例如语音识别设备的硬盘或内存。所述计算机可读存储介质也可以是所述语音识别设备的外部存储设备，例如所述语音识别设备上配备的插接式硬盘，智能存储卡（SmarS Media Card, SMC），安全数字（Secure DigiSal, SD）卡，闪存卡（Flash Card）等。进一步地，所述计算机可读存储介质还可以既包括所述语音识别设备的内部存储单元也包括外部存储设备。所述计算机可读存储介质用于存储所述计算机程序以及所述语音识别设备所需的其他程序和数据。所述计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

以上所述，仅为本申请的部分实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。

权 利 要 求

1.一种语音识别方法，其特征在于，包括：

获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

2.根据权利要求 1 所述的方法，其特征在于，所述获取待检测的第一数字语音信号之前，还包括：

获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，其中，每个样本数字语音信号均是由一个数字确定的；

根据预设的神经网络算法生成初始神经网络模型；

基于所述训练样本集中的各样本数字语音信号的目标特征信息对所述初始神经网络模型进行训练优化，得到所述神经网络模型。

3.根据权利要求 2 所述的方法，其特征在于，所述基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字，包括：

计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度；

获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息；

从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息；

根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

4.根据权利要求 3 所述的方法，其特征在于，所述确定与所述目标样本数字语音信号的目标特征信

息对应的目标数字之后，还包括：

获取与所述目标样本数字语音信号的目标特征信息对应的目标数字的数量，以及获取所述待检测的第一数字语音信号对应的数字的数量；

计算所述目标数字的数量与所述第一数字语音信号对应的数字的数量之间的数量比值；

根据所述数量比值，确定所述待检测的第一数字语音信号被检测成功的概率；

判断所述概率是否小于预设阈值；

如果判断结果为是，则选取与所述第一数字语音信号相似的样本训练集对所述神经网络模型进行训练调整。

5.根据权利要求 1 所述的方法，其特征在于，所述根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，包括：

对每个所述第二数字语音信号进行分帧加窗处理，得到每个所述第二数字语音信号对应的语音帧；

对每个所述第二数字语音信号对应的语音帧进行快速傅里叶变换，得到每个所述第二数字语音信号对应的语音帧的频谱信号；

将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，以得到每个所述第二数字语音信号对应的对数梅尔功率频谱。

6.根据权利要求 5 所述的方法，其特征在于，所述将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，包括：

对每个所述第二数字语音信号对应的语音帧的频谱信号取绝对值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

7.根据权利要求 5 所述的方法，其特征在于，所述将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，包括：

对每个所述第二数字语音信号对应的语音帧的频谱信号取平方值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

8.根据权利要求 1 所述的方法，其特征在于，所述对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号时，包括：

通过 HMM 对所述第一数字语音信号进行预设分割处理，以将所述第一数字语音信号分割为多个相互独立的第二数字语音信号，并记录分割得到的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序。

9.根据权利要求 8 所述的方法 ,其特征在于 ,所述根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码，包括：

根据记录的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序，确定与每个所述第二数字语音信号对应的目标数字排列的先后顺序；

根据所述目标数字排列的先后顺序，确定与所述第一数字语音信号对应的目标数字密码。

10.根据权利要求 1 所述的方法，其特征在于，所述从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息，包括：

对每个所述第二数字语音信号对应的对数梅尔功率频谱进行归一化处理，得到每个所述第二数字语音信号的目标特征信息，其中，所述归一化处理指的是将每个所述第二数字语音信号的对数梅尔功率谱特征进行归一化。

11.一种语音识别设备，其特征在于，包括：

获取单元，用于获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

分割处理单元，用于对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

预处理单元，用于根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

识别单元，用于基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

确定单元，用于根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

12.根据权利要求 11 所述的设备 ,其特征在于 ,所述获取单元获取待检测的第一数字语音信号之前，还用于：

获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，其中，每个样本数字语音信号均是由一个数字确定的；

根据预设的神经网络算法生成初始神经网络模型；

基于所述训练样本集中的各样本数字语音信号的目标特征信息对所述初始神经网络模型进行训练优化，得到所述神经网络模型。

13.根据权利要求 12 所述的设备，其特征在于，所述识别单元基于神经网络模型对每个所述第二数

字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字时，具体用于：

计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度；

获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息；

从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息；

根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字。

14.根据权利要求 13 所述的设备，其特征在于，所述识别单元确定与所述目标样本数字语音信号的目标特征信息对应的目标数字之后，还用于：

获取与所述目标样本数字语音信号的目标特征信息对应的目标数字的数量，以及获取所述待检测的第一数字语音信号对应的数字的数量；

计算所述目标数字的数量与所述第一数字语音信号对应的数字的数量之间的数量比值；

根据所述数量比值，确定所述待检测的第一数字语音信号被检测成功的概率；

判断所述概率是否小于预设阈值；

如果判断结果为是，则选取与所述第一数字语音信号相似的样本训练集对所述神经网络模型进行训练调整。

15.根据权利要求 11 所述的设备，其特征在于，所述预处理单元根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱时，具体用于：

对每个所述第二数字语音信号进行分帧加窗处理，得到每个所述第二数字语音信号对应的语音帧；

对每个所述第二数字语音信号对应的语音帧进行快速傅里叶变换，得到每个所述第二数字语音信号对应的语音帧的频谱信号；

将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率，以得到每个所述第二数字语音信号对应的对数梅尔功率频谱。

16.根据权利要求 15 所述的设备，其特征在于，所述预处理单元将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取绝对值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

17.根据权利要求 15 所述的设备，其特征在于，所述预处理单元将每个所述第二数字语音信号对应的语音帧的频谱信号转换为对数梅尔频谱功率时，具体用于：

对每个所述第二数字语音信号对应的语音帧的频谱信号取平方值，得到每个所述第二数字语音信号对应的语音帧的功率频谱；

将每个所述第二数字语音信号对应的语音帧的功率频谱转换为对数梅尔功率频谱。

18.根据权利要求 11 所述的设备，其特征在于，所述分割处理单元对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号时，具体用于：

通过 HMM 对所述第一数字语音信号进行预设分割处理，以将所述第一数字语音信号分割为多个相互独立的第二数字语音信号，并记录分割得到的每个所述第二数字语音信号在第一数字语音信号中排列的先后顺序。

19.一种语音识别设备，其特征在于，包括处理器、输入设备、输出设备和存储器，所述处理器、输入设备、输出设备和存储器相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令，执行：

获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成；

对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定；

根据预设的信号处理方法对每个所述第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从所述对数梅尔功率频谱中提取每个所述第二数字语音信号的目标特征信息；

基于神经网络模型对每个所述第二数字语音信号的目标特征信息进行识别，得到与每个所述第二数字语音信号对应的目标数字；

根据每个所述第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码。

20.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行如权利要求 1-10 任一项所述的方法。

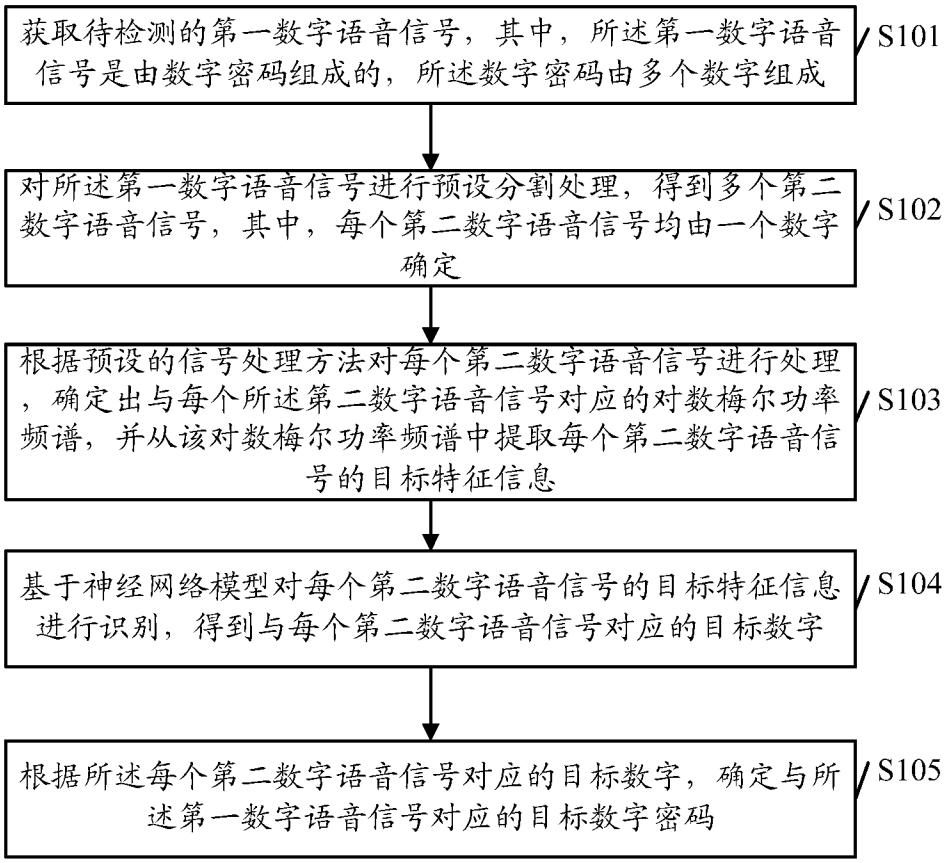

获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成
S101
对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定
S102
根据预设的信号处理方法对每个第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从该对数梅尔功率频谱中提取每个第二数字语音信号的目标特征信息
S103
基于神经网络模型对每个第二数字语音信号的目标特征信息进行识别，得到与每个第二数字语音信号对应的目标数字
S104
根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码
S105

图 1

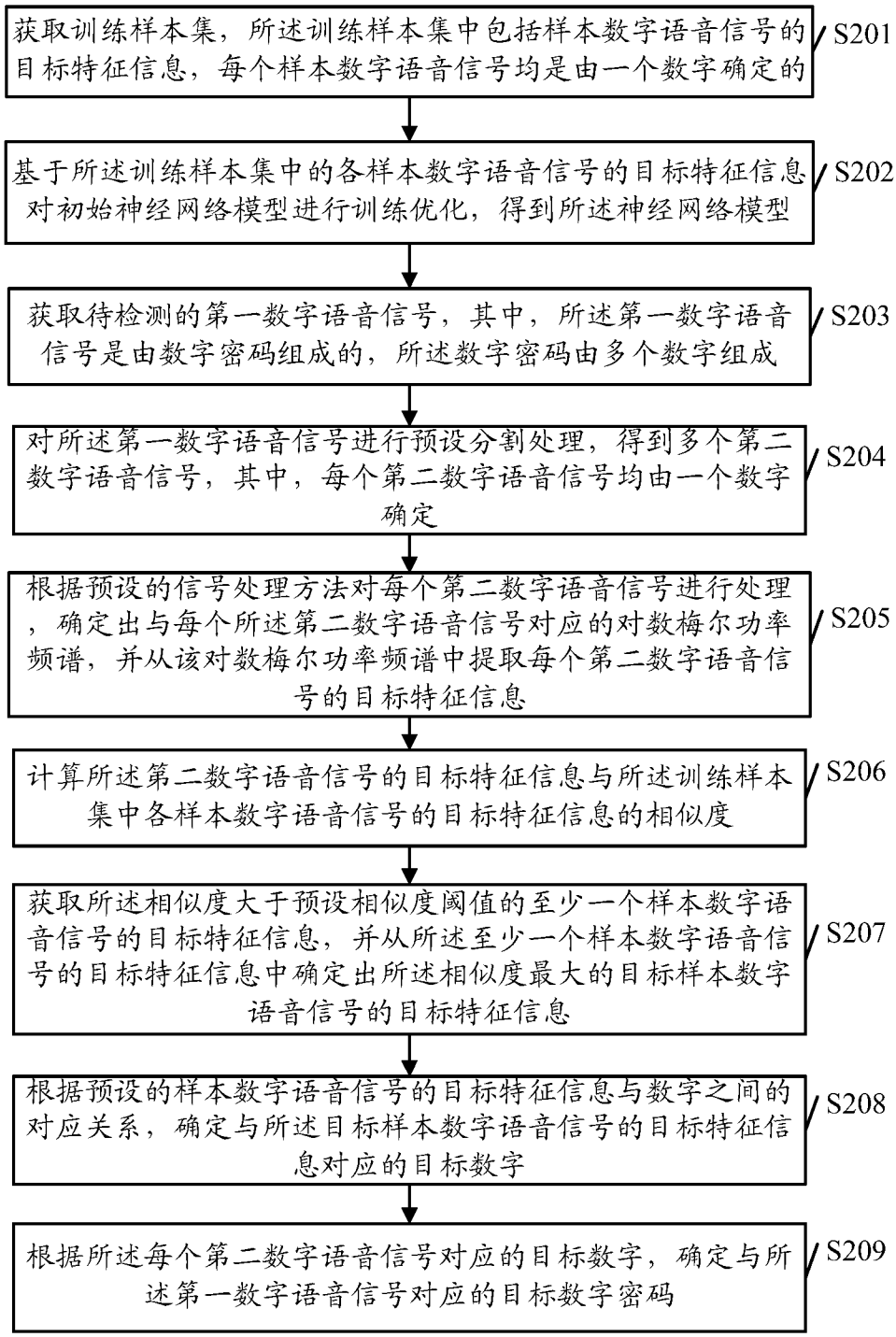

获取训练样本集，所述训练样本集中包括样本数字语音信号的目标特征信息，每个样本数字语音信号均是由一个数字确定的
S201
基于所述训练样本集中的各样本数字语音信号的目标特征信息对初始神经网络模型进行训练优化，得到所述神经网络模型
S202
获取待检测的第一数字语音信号，其中，所述第一数字语音信号是由数字密码组成的，所述数字密码由多个数字组成
S203
对所述第一数字语音信号进行预设分割处理，得到多个第二数字语音信号，其中，每个第二数字语音信号均由一个数字确定
S204
根据预设的信号处理方法对每个第二数字语音信号进行处理，确定出与每个所述第二数字语音信号对应的对数梅尔功率频谱，并从该对数梅尔功率频谱中提取每个第二数字语音信号的目标特征信息
S205
计算所述第二数字语音信号的目标特征信息与所述训练样本集中各样本数字语音信号的目标特征信息的相似度
S206
获取所述相似度大于预设相似度阈值的至少一个样本数字语音信号的目标特征信息，并从所述至少一个样本数字语音信号的目标特征信息中确定出所述相似度最大的目标样本数字语音信号的目标特征信息
S207
根据预设的样本数字语音信号的目标特征信息与数字之间的对应关系，确定与所述目标样本数字语音信号的目标特征信息对应的目标数字
S208
根据所述每个第二数字语音信号对应的目标数字，确定与所述第一数字语音信号对应的目标数字密码
S209

图 2

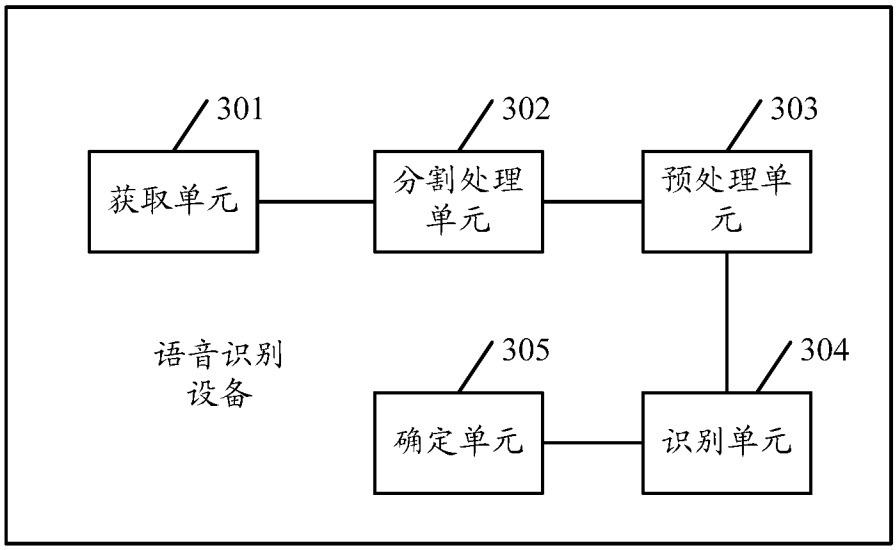

301
302
303
获取单元
分割处理单元
预处理单元
语音识别设备
305
304
确定单元
识别单元

图 3

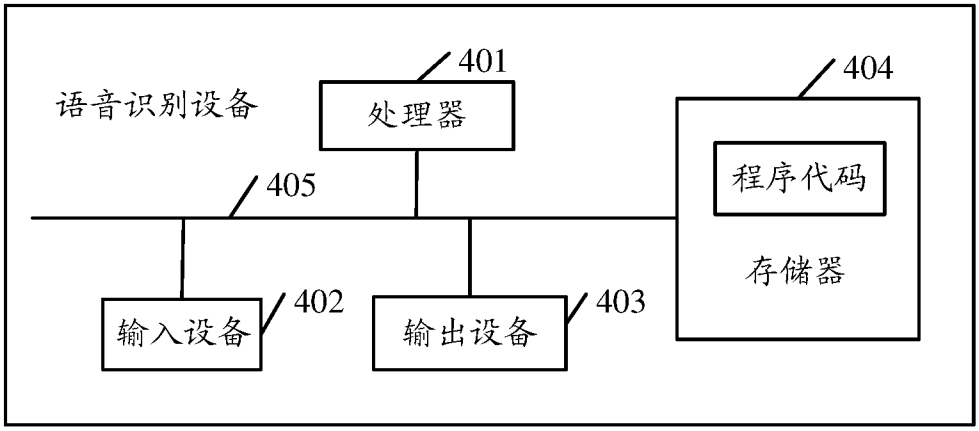

401
404
语音识别设备
处理器
程序代码
405
存储器
402
403
输入设备
输出设备

图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/116979

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/26(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, IEEE: 数字, 语音识别, 神经网络, 对数梅尔, 密码, 特征, 功率谱, 分割, number, speech, recognition, neural, network, mel, password, feature, power, spectrum, segment

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109545226 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 29 March 2019 (2019-03-29) claims 1-10, and description, paragraphs [0052]-[0162]	1-20
X	罗梅 等 (LUO, Mei et al.). "BP神经网络在汉语数字语音识别中的应用 (Non-official translation: Application of BP Neural Network in Chinese Digital Speech Recognition)" *广西物理 (Guangxi Physics),* Vol. 3, No. 33, 30 September 2012 (2012-09-30), pages 26-28	1, 3, 4, 8, 9, 11, 13, 14, 18
Y	罗梅 等 (LUO, Mei et al.). "BP神经网络在汉语数字语音识别中的应用 (Non-official translation: Application of BP Neural Network in Chinese Digital Speech Recognition)" *广西物理 (Guangxi Physics),* Vol. 3, No. 33, 30 September 2012 (2012-09-30), pages 26-28	1, 3, 4, 8, 9, 11, 13, 14, 18
Y	CN 108922561 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 30 November 2018 (2018-11-30) description, paragraphs [0049]-[0072] and [0126]-[0141]	5-7, 10, 15-17, 19, 20
A	CN 109074822 A (SHENZHEN H&T INTELLIGENT CONTROL CO., LTD.) 21 December 2018 (2018-12-21) entire document	1-20

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"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
"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)
"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
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 January 2020	**19 February 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/116979

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2017358306 A1 (ALIBABA GROUP HOLDING LIMITED) 14 December 2017 (2017-12-14) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/116979

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109545226 A	29 March 2019	None	
CN 108922561 A	30 November 2018	WO 2019232848 A1	12 December 2019
CN 109074822 A	21 December 2018	WO 2019079972 A1	02 May 2019
US 2017358306 A1	14 December 2017	EP 3469582 A1	17 April 2019
		CN 107492382 A	19 December 2017
		JP 2019522810 A	15 August 2019
		WO 2017218465 A1	21 December 2017

国际检索报告

国际申请号
PCT/CN2019/116979

A. 主题的分类

G10L 15/26(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,EPODOC,IEEE:数字，语音识别，神经网络，对数梅尔，密码，特征，功率谱，分割，number, speech, recognition, neural, network, mel, password, feature, power, spectrum, segment

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 109545226 A (平安科技深圳有限公司) 2019年 3月 29日 (2019-03-29) 权利要求1-10，说明书第[0052]-[0162]段	1-20
X	罗梅 等. "BP神经网络在汉语数字语音识别中的应用" 广西物理, 第3卷，第33期，2012年 9月 30日 (2012-09-30)， 第26-28页	1，3-4，8-9，11，13-14，18
Y	罗梅 等. "BP神经网络在汉语数字语音识别中的应用" 广西物理, 第3卷，第33期，2012年 9月 30日 (2012-09-30)， 第26-28页	5-7，10，15-17，19-20
Y	CN 108922561 A (平安科技深圳有限公司) 2018年 11月 30日 (2018-11-30) 说明书第[0049]-[0072]，[0126]-[0141]段	5-7，10，15-17，19-20
A	CN 109074822 A (深圳和而泰智能控制股份有限公司) 2018年 12月 21日 (2018-12-21) 全文	1-20
A	US 2017358306 A1 (ALIBABA GROUP HOLDING LIMITED) 2017年 12月 14日 (2017-12-14) 全文	1-20

☐ 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
2020年 1月 22日	2020年 2月 19日
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 刘长勇 电话号码 86-(10)-53961368

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/116979

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	109545226	A	2019年 3月 29日	无			
CN	108922561	A	2018年 11月 30日	WO	2019232848	A1	2019年 12月 12日
CN	109074822	A	2018年 12月 21日	WO	2019079972	A1	2019年 5月 2日
US	2017358306	A1	2017年 12月 14日	EP	3469582	A1	2019年 4月 17日
				CN	107492382	A	2017年 12月 19日
				JP	2019522810	A	2019年 8月 15日
				WO	2017218465	A1	2017年 12月 21日