

(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



(10) 国际公布号

WO 2019/232829 A1

(43) 国际公布日

2019 年 12 月 12 日 (12.12.2019)

WIPO | PCT

(51) 国际专利分类号:
G10L 17/00 (2013.01) G10L 25/60 (2013.01)

(21) 国际申请号: PCT/CN2018/092598

(22) 国际申请日: 2018 年 6 月 25 日 (25.06.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810573715.4 2018 年 6 月 6 日 (06.06.2018) 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八卦岭八卦三路平安大厦6楼, Guangdong 518000 (CN)。

(72) 发明人: 涂宏(TU, Hong); 中国广东省深圳市福田区八卦岭八卦三路平安大厦6楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳众鼎专利商标代理事务所(普通合伙)(SHENZHEN ZHONGDING INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市龙岗区龙城街道中心城清林路546号城投商务中心4层/B, Guangdong 518172 (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。

(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,

(54) Title: VOICEPRINT RECOGNITION METHOD AND APPARATUS, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 声纹识别方法、装置、计算机设备及存储介质

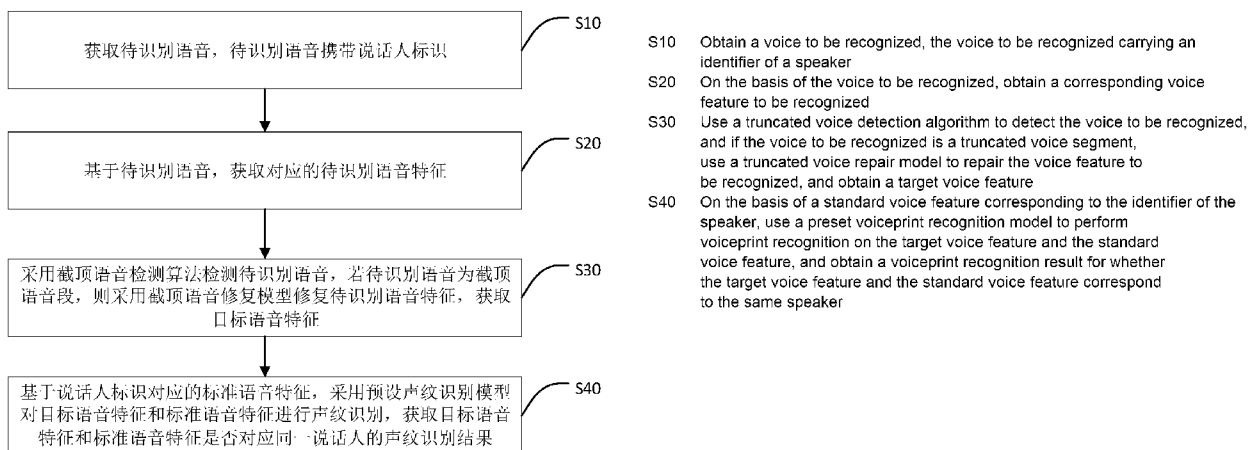


图 2

(57) Abstract: A voiceprint recognition method and apparatus, a computer device and a storage medium, the method comprising: obtaining a voice to be recognized, the voice to be recognized carrying an identifier of a speaker (S10); on the basis of the voice to be recognized, obtaining a corresponding voice feature to be recognized (S20); using a truncated voice detection algorithm to detect the voice to be recognized, and if the voice to be recognized is a truncated voice segment, using a truncated voice repair model to repair the voice feature to be recognized, and obtaining a target voice feature (S30); on the basis of a standard voice feature corresponding to the identifier of the speaker, using a preset voiceprint recognition model to perform voiceprint recognition on the target voice feature and the standard voice feature, and obtaining a voiceprint recognition result (S40). By means of repairing a voice feature to be recognized in a voice to be recognized and then obtaining a target voice feature close to an original voice of a speaker, the described method effectively improves the reliability and accuracy of recognizing the speaker.

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：一种声纹识别方法、装置、计算机设备及存储介质，该方法包括：获取待识别语音，待识别语音携带说话人标识(S10)；基于待识别语音，获取对应的待识别语音特征(S20)；采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征(S30)；基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取声纹识别结果(S40)。该方法通过修复待识别语音中的待识别语音特征后获取接近说话人原始语音的目标语音特征，有效提高了说话人识别的可靠性和准确性。

声纹识别方法、装置、计算机设备及存储介质

本申请以 2018 年 06 月 06 日提交的申请号为 201810573715.4，名称为“声纹识别方法、装置、计算机设备及存储介质”的中国发明申请为基础，并要求其优先权。

技术领域

本申请涉及生物识别技术领域，尤其涉及一种声纹识别方法、装置、计算机设备及存储介质。

背景技术

通讯设备制造商为了使通话音量保持在一个合适的音量范围内，给通讯设备配置有语音增益控制模块，以使语音通话更为友好。自动增益控制模块的工作原理是通过调节语音音量的饱和值也即截顶语音来实现的，具体包括给音量小的语音增添较大增益，给音量大的语音分配较小增益。而这样的设置方式也随之带来问题：通信设备中的语音频繁出现截顶现象，使得基于该通信设备采集到的语音在进行声纹识别时，会削弱声纹识别的准确性。

发明内容

基于此，有必要针对上述技术问题，提供一种可以增强声纹识别准确性的声纹识别方法、装置、计算机设备及存储介质。

一种声纹识别方法，包括：

获取待识别语音，待识别语音携带说话人标识；

基于待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；

基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

一种声纹识别装置，包括：

获取待识别语音模块，用于获取待识别语音，待识别语音携带说话人标识；

获取待识别特征模块，用于基于待识别语音，获取对应的待识别语音特征；

获取目标语音特征模块，用于采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；

获取声纹识别结果模块，用于基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

一种计算机设备，包括存储器、处理器以及存储在存储器中并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现如下步骤：

获取待识别语音，待识别语音携带说话人标识；

基于待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；

基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

一个或多个存储有计算机可读指令的非易失性可读存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行如下步骤：

获取待识别语音，待识别语音携带说话人标识；

基于待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；

基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

本申请的一个或多个实施例的细节在下面的附图和描述中提出，本申请的其他特征和优点将从说明书、附图以及权利要求变得明显。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获取其他的附图。

图 1 是本申请一实施例中声纹识别方法的一应用环境示意图；

图 2 是本申请一实施例中声纹识别方法的一流程图；

图 3 是本申请一实施例中声纹识别方法的另一具体流程图；

图 4 是本申请一实施例中声纹识别方法的另一具体流程图；

图 5 是本申请一实施例中声纹识别方法的另一具体流程图；

图 6 是本申请一实施例中声纹识别方法的另一具体流程图；

图 7 是本申请一实施例中声纹识别装置的一原理框图；

图 8 是本申请一实施例中计算机设备的一示意图。

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作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声纹识别方法，可应用在如图 1 的应用环境中，其中，语音采集终端通过网络与识别服务器进行通信。其中，语音采集终端包括但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备。识别服务器可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

声纹(Voiceprint)信息，是用电声学仪器显示的携带言语信息的声波频谱。人的发声

控制器官包括声带、软颚、舌头、牙齿和唇等，人的发声共鸣器包括咽腔、口腔和鼻腔。这些器官存在大小、形态及功能上的差异，这些差异导致发声气流的改变，造成音质和音色的差别。此外，人发声的习惯有快有慢，用力有大有小，也造成音强和音长的差别。音高、音强、音长和音色在语言学中被称为语音“四要素”，这些因素又可分解成九十余种特征。这些特征表现成不同声音的波长、频率、强度和节奏，通过声学工具可绘制成基于时域的功率谱，也即形成说话人的声纹信息。

声纹识别，生物识别技术的一种，也称为说话人识别，有两类，即说话人辨认和说话人确认。不同的任务和应用会使用不同的声纹识别技术，如缩小刑侦范围时可能需要辨认技术，而银行交易时则需要确认技术。应用于本实施例，基于说话人确认技术进行说明。

在一实施例中，如图 2 所示，提供一种声纹识别方法，以该声纹识别方法应用在图 1 中的识别服务器为例进行说明，包括如下步骤：

S10. 获取待识别语音，待识别语音携带说话人标识。

其中，待识别语音是语音采集终端直接采集到的需要进行识别的语音，该待识别语音携带有说话人标识，用于识别该待识别语音对应的说话人。说话人标识是待识别说话人提供的用于验证身份的说话人标识，包括但不限于：姓名、注册名或身份证号码等可表示说话人唯一身份的标识。

由于通讯设备制造商在制造语音采集终端过程中，在该语音采集终端配置语音增益控制模块，以使采集到的说话人的语音保持在一个合适的音量范围内，使得该语音采集终端直接采集到的待识别语音包括截顶语音段和正常语音段。具体地，语音采集终端录制说话人发出的待识别语音时，若说话人的音量过大或过小，语音采集终端的语音增益控制模块会自适应调整最高音阈值或最低音阈值对应的幅度阈值，然后把待识别语音的音量高于最高音阈值的幅度部分或者低于最低音阈值的幅度部分截去，并记录为幅度阈值，从而形成截顶语音段。相应地，语音采集终端录制的待识别语音中，音量在最低音阈值和最高音阈值之间的语音部分无需通过语音增益控制模块进行增益处理，故为正常语音段。

以待识别语音形成的典型正弦声波信号为例，若待识别语音的最大振幅为 E_m ，语音采集终端的幅度阈值为 E_q ，发生信号截顶时最大振幅 E_m 超过幅度阈值 E_q ，则会直接导致采样点取值在幅度阈值 E_q ，在波形上显示为大于幅度阈值 E_q 的部分被截断，从而形成本实施例所说的截顶语音段。在实际情况下，当语音采集终端采集大量的待识别语音，语音采集终端可能会自动调整增益大小，可能发生接收到的样本会被随机记录成一个低于幅度阈值 E_q 的值 E_c ，此时， E_c 自适应调整为幅度阈值。

S20. 基于待识别语音，获取对应的待识别语音特征。

其中，待识别语音特征是用以将待识别语音区别于其他人语音的特征，应用于本实施例，可采用梅尔频率倒谱系数 (Mel-Frequency Cepstral Coefficients, 以下简称 MFCC 特征) 作为待识别语音特征。

语音特征在声纹识别技术领域常用 MFCC (Mel-scale Frequency Cepstral Coefficients, 梅尔频率倒谱系数) 特征来表示。根据人耳听觉机理的研究发现，人耳对不同频率的声波有不同的听觉敏感度。从 200Hz 到 5000Hz 的语音信号对语音的清晰度影响对大。两个响度不等的声音作用于人耳时，则响度较高的频率成分的存在会影响到对响度较低的频率成分的感受，使其变得不易察觉，这种现象称为掩蔽效应。由于频率较低的声音在内耳蜗基底膜上行波传递的距离大于频率较高的声音，故一般来说，低音容易掩蔽高音，而高音掩蔽低音较困难。在低频处的声音掩蔽的临界带宽较高频要小。所以，从低频

到高频这一段频带内按临界带宽的大小可以由密到疏安排一组梅尔刻度的带通滤波器，对输入信号进行滤波以使每个频率段对应一个数值。梅尔刻度的滤波器组在低频部分的分辨率高，跟人耳的听觉特性是相符的，这也是梅尔刻度的物理意义所在。

S30. 采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征。

其中，目标语音特征包含正常语音段对应的语音特征，也包含采用截顶语音修复模型修复截顶语音段对应的语音特征所形成的语音特征。即目标语音特征是对待识别语音特征进行语音修复后形成的语音特征。

截顶语音修复模型是可以将输入的待识别语音特征还原，并输出为目标语音特征的模型。应用于本实施例，截顶语音修复模型已预先训练好并存储在识别服务器上，以便识别服务器可实时调用该模型来修复截顶语音段。

本实施例中，截顶语音检测算法是检测语音采集终端采集的待识别语音的种类的算法。待识别语音的种类包括两种：经截顶语音处理后的截顶语音段和未经截顶语音处理的正常语音段。可以理解地，由于正常语音段没有对语音信号进行处理，其保留说话人的声纹特征，而截顶语音段是待识别语音被截除小于最低音阈值或者大于最高音阈值之后获取的语音段，存在语音信号失真现象。若直接基于包括截顶语音段对应的待识别语音进行语音识别，可能存在识别不准确的现象。因此，需要采用截顶语音检测算法首先判定待识别语音的种类，为后续语音识别提供技术基础。

在一实施例中，采用截顶语音修复模型修复待识别语音特征，获取目标语音特征。

其中，截顶语音修复模型是由初始训练模型经训练后形成的、用于修复截顶语音段对应的语音特征的模型。应用于本实施例，可采用 DNN (Deep Neural Networks, 深度神经网络)、DBN (Deep Belief Nets, 深度信念网络) 模型或 CDBN (Convolutional Deep Belief Networks, 卷积的深度信念网络) 模型等作为初始训练模型。下述以采用 DNN 模型为例说明修复截顶语音段的过程：

DNN 模型被广泛应用在很多重要的互联网应用，如语音识别，图像识别，自然语言处理等。DNN 模型由于计算复杂度高，能大幅度提高语音识别的精度，因此被广泛用在很多公司的语音识别产品中。

目前的 DNN 模型的结构包括输入层，数个中间层和输出层。输入层负责接收来自外界的输入信息，并传递给中间层；中间层是内部信息处理层，负责信息变换，根据信息变化能力的需求，中间层可以设计为单中间层或者多中间层结构；中间层传递到输出层的信息经进一步处理后，完成一次学习的正向传播处理过程，由输出层向外界输出信息处理结果。

各层的神经元数一般是几百到几万不等，层以层之间是全连接的网络。DNN 模型的训练计算是一层算完再期待下一层，层与层之间不能并行。一般一次 DNN 训练可以用以下几个阶段表示：前向计算、反向误差计算，最后是根据前向计算和反向误差计算的结果更新每层的权重。前向计算过程是从输入层向后一直算到输出层，计算是串行的。反向计算过程是从输出层向前一直算到第一层，计算也是串行的。

每次输入一小段训练数据称为一个 batch，一个 batch 完成一次训练，也就是说得到一个新的权值后，会用这个权值及新输入的下一个 batch 进行训练，得到更新的一个权值，直到所有的输入计算完毕称为一轮。一般一个完整的训练需要 10~20 轮。

当实际输出与期望输出不符时，进入误差的反向传播阶段。误差通过输出层，按误差梯度下降的方式修正各层权值，向中间层和输入层逐层反传。DNN 训练过程是周而复始的

信息正向传播和误差反向传播过程，是各层权值不断调整的过程，也是神经网络学习训练的过程，此过程一直进行到网络输出的误差减少到可以接受的程度，或者预先设定的学习次数为止。

在一实施例中，在步骤 S30 中，即采用截顶语音修复模型修复待识别语音特征，获取目标语音特征，具体包括如下步骤：

S31. 采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取目标语音特征。

其中，截顶语音修复模型是由 DNN 模型经训练后形成的、用于修复截顶语音段对应的语音特征并输出为目标语音特征的模型。

具体地，本实施例可采用步骤 S30 生成的截顶语音修复模型用以修复待识别语音特征。将待识别语音特征也即待识别 MFCC 特征作为 DNN 模型的输入，经 DNN 模型的训练，可获取 DNN 模型输出的目标语音特征，也即原始 MFCC 特征。

在步骤 S30 中，识别服务器采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取目标语音特征，即将截顶语音段的待识别语音特征（截顶 MFCC 特征）输入截顶语音修复模型，用以经截顶语音修复模型修复后获取修复语音段的目标语音特征（MFCC 特征），以作为语音识别的技术基础。因 MFCC 特征是基于对输入的语音信号不做任何的假设和限制，且利用听觉模型而产生的，具有良好的鲁棒性，更符合人耳的听觉特性，即使当信噪比降低时仍然具有较好的语音识别性能。本步骤中采用基于 DNN 模型生成的截顶语音修复模型来修复截顶语音段，可大幅提高语音修复的准确性。

进一步地，因语音采集终端的增益模块会自适应调整幅度阈值，识别服务器难以通过指定一个固定的幅度阈值来判定截顶语音段，可采用判定高音音量子区间的采样点数的百分比，也即采用本步骤提出的截顶语音检测算法来进行判定，可有效提高判定结果的准确性。

本实施例提出的声纹识别方法，通过截顶语音检测算法可检测出待识别语音是否为截顶语音段，若是则通过截顶语音修复模型可将截顶语音段的待识别语音特征修复为目标语音特征，以对比说话人对于的标准语音特征来识别说话人的真实身份。本实施例通过修复待识别语音中的待识别语音特征后获取接近说话人原始语音的目标语音特征，可有效提高语音识别的可靠性和准确性。

S40. 基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

其中，标准语音是与说话人标识对应的保存在预存标准语音库中、记录未经截顶语音处理的说话人发出的语音。同理地，标准语音特征就是标准语音对应的 MFCC 特征。

预设声纹识别模型是用于对标准语音特征和目标语音特征进行相似度打分的模型，现有多成熟的声纹识别模型，比如 GMM-UBM (Gaussian mixture model-universal background model, 混合模型-通用背景模型) 模型或 i-vector (identity-vector, 身份认证向量) 模型等。应用于本实施例，可采用 i-vector 模型作为预设声纹识别模型。

优选地，为了加快语音识别的处理速度，识别服务器可在数据库中关联存储标准 MFCC 特征对应的标准 i-vector，以便基于 i-vector 进行语音识别时可直接调用该标准 i-vector 作为对比标准。

采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别是通过对比目标语音特征和标准语音特征的余弦距离实现的。余弦距离，也称为余弦相似度，是用向量空

间中两个向量夹角的余弦值作为衡量两个个体间差异的大小的度量。向量，是多维空间中有方向的线段，如果两个向量的方向一致，即夹角接近零，夹角的余弦值也即余弦距离趋近于 1，则这两个向量就相近。应用于本实施例，可根据实际情况设置一距离阈值。当标准 i-vector 和原始 i-vector 的余弦距离大于距离阈值时，可判定目标语音特征和标准语音特征对应同一说话人。

在一实施例中，如图 3 所示，在步骤 S20 中，即基于待识别语音，获取对应的待识别语音特征，具体包括如下步骤：

S21：对待识别语音进行预处理，获取预处理语音数据。

本实施例中，对待识别语音进行预处理，并获取相对应的预处理语音数据。对待识别语音进行预处理能够更好地提取待识别语音的待识别语音特征，使得提取出的待识别语音特征更能代表该待识别语音，以采用该待识别语音特征进行语音区分。

在一实施例中，步骤 S21 中，对待识别语音进行预处理，获取预处理语音数据，包括如下步骤：

S211：对待识别语音作预加重处理，预加重处理的计算公式为 $s'_n = s_n - a * s_{n-1}$ ，其中， s_n 为时域上的信号幅度， s_{n-1} 为与 s_n 相对应的上一时刻的信号幅度， s'_n 为预加重后时域上的信号幅度， a 为预加重系数， a 的取值范围为 $0.9 < a < 1.0$ 。

其中，预加重是一种在发送端对输入信号高频分量进行补偿的信号处理方式。随着信号速率的增加，信号在传输过程中受损很大，为了使接收端能得到比较好的信号波形，就需要对受损的信号进行补偿。预加重技术的思想就是在传输线的发送端增强信号的高频成分，以补偿高频分量在传输过程中的过大衰减，使得接收端能够得到较好的信号波形。预加重对噪声并没有影响，因此能够有效提高输出信噪比。

本实施例中，对待识别语音作预加重处理，该预加重处理的公式为 $s'_n = s_n - a * s_{n-1}$ ，其中， s_n 为时域上的信号幅度，即语音数据在时域上表达的语音的幅值（幅度）， s_{n-1} 为与 s_n 相对的上一时刻的信号幅度， s'_n 为预加重后时域上的信号幅度， a 为预加重系数， a 的取值范围为 $0.9 < a < 1.0$ ，这里取 0.97 预加重的效果比较好。采用该预加重处理能够消除发声过程中声带和嘴唇等造成的干扰，可以有效补偿待识别语音被压抑的高频部分，并且能够突显待识别语音高频的共振峰，加强待识别语音的信号幅度，有助于提取待识别语音特征。

S212：将预加重后的待识别语音进行分帧处理。

本实施例中，在预加重待识别语音后，还应进行分帧处理。分帧是指将整段的语音信号切分成若干段的语音处理技术，每帧的大小在 10-30ms 的范围内，以大概 1/2 帧长作为帧移。帧移是指相邻两帧间的重叠区域，能够避免相邻两帧变化过大的问题。对待识别语音进行分帧处理，能够将待识别语音分成若干段的语音数据，可以细分待识别语音，便于待识别语音特征的提取。

S213：将分帧后的待识别语音进行加窗处理，获取预处理语音数据，加窗的计算公式为 $s'_n = \{0.54 - 0.46 \cos(\frac{2\pi(n-1)}{N-1})\} * s_n$ ，其中， N 为窗长， n 为时间， s_n 为时域上的信号幅度，

s'_n 为加窗后时域上的信号幅度。

本实施例中，在对待识别语音进行分帧处理后，每一帧的起始段和末尾端都会出现不连续的地方，所以分帧越多与待识别语音的误差也就越大。采用加窗能够解决这个问题，可以使分帧后的待识别语音变得连续，并且使得每一帧能够表现出周期函数的特征。加窗处理具体是指采用窗函数对待识别语音进行处理，窗函数可以选择汉明窗，则该加窗的公

式为 $s'_n = \{0.54 - 0.46 \cos(\frac{2\pi(n-1)}{N-1})\} * s_n$ ， N 为汉明窗窗长， n 为时间， s_n 为时域上的信号幅

度， s'_n 为加窗后时域上的信号幅度。对待识别语音进行加窗处理，获取预处理语音数据，能够使得分帧后的待识别语音在时域上的信号变得连续，有助于提取待识别语音的待识别语音特征。

上述步骤 S211-S213 对待识别语音的预处理操作，为提取待识别语音的待识别语音特征提供了基础，能够使得提取的待识别语音特征更能代表该待识别语音，并根据该待识别语音特征进行语音区分。

S22：对预处理语音数据作快速傅里叶变换，获取待识别语音的频谱，并根据频谱获取待识别语音的功率谱。

其中，快速傅里叶变换（Fast Fourier Transformation，简称 FFT），指利用计算机计算离散傅里叶变换的高效、快速计算方法的统称，简称 FFT。采用这种算法能使计算机计算离散傅里叶变换所需要的乘法次数大为减少，特别是被变换的抽样点数越多，FFT 算法计算量的节省就越显著。

本实施例中，对预处理语音数据进行快速傅里叶变换，以将预处理语音数据从时域上的信号幅度转换为在频域上的信号幅度（频谱）。该计算频谱的公式为 $s(k) = \sum_{n=0}^{N-1} s(n)e^{-i2\pi kn/N}$ ，

$1 \leq k \leq N$ ， N 为帧的大小， $s(k)$ 为频域上的信号幅度， $s(n)$ 为时域上的信号幅度， n 为时间， i 为复数单位。在获取预处理语音数据的频谱后，可以根据该频谱直接求得预处理语音数据的功率谱，以下将预处理语音数据的功率谱称为目标待区分语音数据的功率谱。

该计算目标待区分语音数据的功率谱的公式为 $P(k) = \frac{1}{N} |S(k)|^2$ ， $1 \leq k \leq N$ ， N 为帧的大小， s

(k) 为频域上的信号幅度。通过将预处理语音数据从时域上的信号幅度转换为频域上的信号幅度，再根据该频域上的信号幅度获取待识别语音的功率谱，为从待识别语音的功率谱中提取待识别语音特征提供重要的技术基础。

S23：采用梅尔刻度滤波器组处理待识别语音的功率谱，获取待识别语音的梅尔功率谱。

其中，采用梅尔刻度滤波器组处理待识别语音的功率谱是对功率谱进行的梅尔频率分析，梅尔频率分析是基于人类听觉感知的分析。观测发现，人耳就像一个滤波器组一样，只关注某些特定的频率分量（人的听觉对频率是有选择性的），也就是说人耳只让某些频率的信号通过，而直接无视不想感知的某些频率信号。然而这些滤波器在频率坐标轴上却不是统一分布的，在低频区域有很多的滤波器，他们分布比较密集，但在高频区域，滤波器的数目就变得比较少，分布很稀疏。可以理解地，梅尔刻度滤波器组在低频部分的分辨

率高，跟人耳的听觉特性是相符的，这也是梅尔刻度的物理意义所在。

本实施例中，采用梅尔刻度滤波器组处理待识别语音的功率谱，获取待识别语音的梅尔功率谱，通过采用梅尔刻度滤波器组对频域信号进行切分，使得最后每个频率段对应一个数值，若滤波器的个数为 22，则可以得到待识别语音的梅尔功率谱对应的 22 个能量值。通过对待识别语音的功率谱进行梅尔频率分析，使得其分析后获取的梅尔功率谱保留着与人耳特性密切相关的频率部分，该频率部分能够很好地反映出待识别语音的特征。

S24：在梅尔功率谱上进行倒谱分析，获取待识别语音的梅尔频率倒谱系数。

其中，倒谱 (cepstrum) 是指一种信号的傅里叶变换谱经对数运算后再进行的傅里叶反变换，由于一般傅里叶谱是复数谱，因而倒谱又称复倒谱。

本实施例中，对梅尔功率谱进行倒谱分析，根据倒谱的结果，分析并获取待识别语音的梅尔频率倒谱系数。通过该倒谱分析，可以将原本特征维数过高，难以直接使用的待识别语音的梅尔功率谱中包含的特征，通过在梅尔功率谱上进行倒谱分析，转换成易于使用的特征（用来进行训练或识别的梅尔频率倒谱系数特征向量）。该梅尔频率倒谱系数能够作为待识别语音特征对不同语音进行区分的系数，该待识别语音特征可以反映语音之间的区别，可以用来识别和区分待识别语音。

在一实施例中，步骤 S24 中，在梅尔功率谱上进行倒谱分析，获取待识别语音的梅尔频率倒谱系数，包括如下步骤：

S241：取梅尔功率谱的对数值，获取待变换梅尔功率谱。

本实施例中，根据倒谱的定义，对梅尔功率谱取对数值 $\log$ ，获取待变换梅尔功率谱 m 。

S242：对待变换梅尔功率谱作离散余弦变换，获取待识别语音的梅尔频率倒谱系数。

本实施例中，对待变换梅尔功率谱 m 作离散余弦变换 (Discrete Cosine Transform, DCT)，获取相对应的待识别语音的梅尔频率倒谱系数，一般取第 2 个到第 13 个系数作为待识别语音特征，该待识别语音特征能够反映语音数据间的区别。对待变换梅尔功率谱 m

作离散余弦变换的公式为 $c_i = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} m_j \cos\left(\frac{\pi}{N}(j-0.5)i\right)$ ， $i=0, 1, 2, \dots, N-1$ ， N 为帧长， m

为待变换梅尔功率谱， j 为待变换梅尔功率谱的自变量。由于梅尔滤波器之间是有重叠的，所以采用梅尔刻度滤波器获取的能量值之间是具有相关性的，离散余弦变换可以对待变换梅尔功率谱 m 进行降维压缩和抽象，并获得间接的待识别语音特征，相比于傅里叶变换，离散余弦变换的结果没有虚部，在计算方面有明显的优势。

步骤 S21-S24 对待识别语音进行特征提取后获得的待识别语音特征，能够很准确地体现待识别语音的特征，利于基于待识别语音特征进行语音识别。

在一实施例中，如图 4 所示，在步骤 S30 中，即采用截顶语音检测算法对待识别语音进行处理，获取截顶语音段，具体包括如下步骤：

S31. 将待识别语音按时序平均分割出至少两个语音子段。

其中，待识别语音是语音采集终端将待识别说话人的原始语音经截顶语音处理后记录的语音。

具体地，将待识别语音按时序分割成若干不重叠的至少两个语音子段，每一时段可设置为 0.5s，以形成做截顶语音识别检测的最小基本单位。

本步骤通过将待识别语音按时序平均分割出至少两个语音子段进行检测，可以提高检测截顶语音段的精确度。可以理解地，分割出的语音子段越多，检测截顶语音段的精确度

越高。

S32. 将语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量。

其中，高音采样点数量就是在最高音所在的音量子区间获得的语音采样点的数量。

具体地，识别服务器首先获取每一语音子段的幅值（音量）最大值 V_m ，将区间 $[0, V_m]$ 均匀分割为不重叠的至少两个音量子区间。因每一语音子段中的最高音所在的音量子区间可能发生截顶语音处理的现象，若发生截顶语音处理，则最高音所在的音量子区间就是幅度阈值所在的区间。

本步骤通过获取最高音所在的音量子区间的高音采样点数量，可用于判定该语音子段是否为截顶语音段作为技术基础。

S33. 统计所有音量子区间的采样点总数，以获取高音采样点数量相对采样点总数的高音采样百分比。

其中，采样点总数是每个语音字段中对语音音量进行采样的所有采样点数。高音采样百分比是高音采样点数量占采样点总数的百分比。

S34. 若高音采样百分比超过预设阈值，则对应的语音子段为截顶语音段。

其中，预设阈值是根据实际经验设定的高音采样点数量相对采样点总数的百分比。若高音采样百分比超过预设阈值，说明该音量子区间所在的语音子段为截顶语音段。

可以理解地，若最高音所在的音量子区间的高音采样百分比没有都超过预设阈值，说明该音量子区间所在的语音子段的语音采样点的数量属于正常范围，该音量子区间所在的语音子段为正常语音段。

本实施例中，因语音采集终端的增益模块会自适应调整幅度阈值，识别服务器难以通过指定一个固定的幅度阈值来判定截顶语音段，可采用高音采样点数量相对采样点总数的高音采样百分比对比预设阈值进行判定，可有效提高判定方法的灵活性和有利于提高判定结果的准确性。

在一实施例中，如图 5 所示，在步骤 S30 之前，即在采用截顶语音修复模型修复待识别语音特征的步骤之前，该声纹识别方法还包括：

S35. 获取原始训练语音对应的原始训练特征，对该原始训练语音进行截顶语音处理以获取对应的截顶训练语音，再提取该截顶训练语音的截顶训练特征。

其中，原始训练语音是说话人发出的未经截顶语音处理的原始语音。可以理解地，截顶训练语音就是说话人发出的、经截顶语音处理后的语音。

本步骤中，识别服务器通过将原始训练语音进行截顶处理，也即只保留原始训练语音在最高音阈值和最低音阈值之间的语音信号，将超出上述范围的语音信号记录为幅度阈值，以获取对应的截顶训练语音。

S36. 将截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

其中，语音特征在声纹识别技术领域常用 MFCC 特征来表示。因 MFCC 特征是基于对输入的语音信号不做任何的假设和限制，且利用听觉模型而产生的，具有良好的鲁棒性，更符合人耳的听觉特性，即使当信噪比降低时仍然具有较好的语音识别性能。因此，截顶训练特征就是截顶训练语音对应的截顶 MFCC 特征，原始训练特征就是原始训练语音对应的原始 MFCC 特征。

具体地，提取截顶训练语音的截顶 MFCC 特征的实现过程包括：基于傅里叶变换将截顶训练语音从时域信号转换为频域信号；再过滤该频域信号以获取梅尔功率谱；在梅尔功率谱上进行倒谱分析，获取初始语音的梅尔频率倒谱系数也即 MFCC 特征。其中，倒谱 (cepstrum) 是指一种信号的傅里叶变换谱经对数运算后再进行的傅里叶反变换，转换成易于使用的语音特征（用来进行训练或识别的梅尔频率倒谱系数特征向量）。上述过程同样适用于提取原始训练语音对应的原始 MFCC 特征。

DNN 模型的结构包括输入层，数个中间层和输出层。输入层负责接收来自外界的输入信息，并传递给中间层；中间层是内部信息处理层，负责信息变换，根据信息变化能力的需求，中间层可以设计为单中间层或者多中间层结构；最高音量所在的中间层传递到输出层的信息经进一步处理后，完成一次学习的正向传播处理过程，由输出层向外界输出信息处理结果。

具体地，识别服务器将截顶 MFCC 特征和原始 MFCC 特征分别作为 DNN 模型的输入层和输出层，以获取 DNN 模型有关截顶语音修复模型的特征参数。

如果原始 MFCC 特征涉及的特征数量为 n ，则 DNN 模型的输出层包括 n 个神经元，该神经元是针对输入每个神经元的截顶 MFCC 特征输出对应的原始 MFCC 特征值。

若 DNN 网络总共有 n 层，每层有多个神经元，不同层之间全连接；每层设置自己的激励函数 f （激励函数是表示神经网络中的每个神经元接受输入值，并将输入值传递给下一层，神经元之间的输入输出关系的函数）。输入为特征向量 v ，第 i 层到第 $i+1$ 层的转移矩阵为 $w_i(i+1)$ ，第 $i+1$ 层的偏置向量为 $b(i+1)$ ，第 i 层的输出为 out_i ，第 $i+1$ 的输入为 ini_{i+1} ，计算过程为：

$$ini_{i+1} = out_i * w_i(i+1) + b(i+1)$$

$$out_{i+1} = f(ini_{i+1})$$

由此可见 DNN 模型的参数包括层间的转移矩阵 w 和每一层的偏置向量 b 等，训练 DNN 模型的主要任务就是确定上述特征参数，最终生成基于 DNN 模型的截顶语音修复模型。

本实施例中，识别服务器采用基于 DNN 模型生成的截顶语音修复模型来修复截顶语音段，可大幅提高语音修复的精度。优选地，还可采用 DBN (Deep Belief Nets, 深度信念网络) 模型或 CDBN (Convolutional Deep Belief Networks, 卷积的深度信念网络) 模型作为初始训练模型，其中，训练 DBN 模型的网络架构快于 DNN 模型，更适用于训练大型语音数据库的数据；CDBN 模型适用于训练特大型语音数据库的数据。

在一实施例中，在步骤 S30 中，即采用截顶语音修复模型修复截顶语音段，获取修复语音段，具体包括如下步骤：

S37. 采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取修复语音段的目标语音特征。

具体地，识别服务器首先获取截顶语音段的待识别语音特征；然后将截顶语音段的语音特征作为 DNN 模型的输入层，经步骤 S32 获得的基于 DNN 模型的截顶语音修复模型的修复后，可在 DNN 模型的输出层获得对应的目标语音特征。

本实施例中，识别服务器采用基于 DNN 模型的截顶语音修复模型的修复截顶语音段，可有效提高输出层获得的目标语音特征的精确性。

在一实施例中，如图 6 所示，在步骤 S40 中，即基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，具体包括如下步骤：

S41. 采用预设声纹识别模型分别处理目标语音特征和标准语音特征，分别得到原始语音向量和标准语音向量。

其中，预设声纹识别模型是用于对标准语音特征和待识别语音特征进行相似度打分的模型，现有多种成熟的声纹识别模型，比如 GMM-UBM (Gaussian mixture model-universal background model, 混合模型-通用背景模型) 或 i-vector 模型等。

应用于本实施例，可采用 i-vector 模型作为预设声纹识别模型，以获取每个说话人对应的 i-vector 向量。向量，是多维空间中有方向的线段，如果两个向量的方向一致，即夹角接近零，夹角的余弦值也即余弦距离趋近于 1，则这两个向量就相近。本实施例中，原始语音向量就是修复 i-vector，标准语音向量就是标准 i-vector。

具体地，识别服务器采用 i-vector 模型获得原始语音向量的具体实现过程如下：

识别服务器通过目标语音特征来训练表征语音空间的 GMM-UBM；利用训练好的 UBM 计算每帧语音特征的充分统计量，并将该充分统计量映射到总变量空间得到目标语音特征的原始 i-vector。优选地，还可采用 LDA (Linear Discriminant Analysis, 线性鉴别分析) 方法进行信道补偿处理该原始 i-vector，以通过投影矩阵算法最小化同类样本间距离和最大化非同类样本间距离。同理可知获得标准语音向量的实现过程。

S42. 获取原始语音向量和标准语音向量的空间距离。

其中，空间距离应用于本实施例可指两个向量之间的余弦距离。余弦距离，也称为余弦相似度，是用向量空间中两个向量夹角的余弦值作为衡量两个个体间差异的大小的度量。

S43. 根据空间距离与预设的距离阈值，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

其中，距离阈值是根据实际经验设置的一用余弦值表示的数值。

具体地，识别服务器通过对比步骤 S42 得到的标准 i-vector 和原始 i-vector 两个向量的空间距离也即余弦距离。若两者的余弦距离大于根据实际情况预设的距离阈值时，可判定目标语音特征和标准语音特征是源自同一说话人。

在一实施例中，在步骤 S42 中，即获取原始语音向量和标准语音向量的空间距离，具体包括如下步骤：

S424. 采用余弦相似度算法获取原始语音向量和标准语音向量的空间距离。

具体地，获取原始语音向量和标准语音向量的空间距离可由以下公式进行判定：

$$\text{similarity} = \cos \theta = \frac{AB}{|A||B|} = \frac{\sum_{i=1}^n A_i * B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} * \sqrt{\sum_{i=1}^n (B_i)^2}}$$

其中， A_i 和 B_i 分别代表原始语音向量和标准语音向量的各个分量。由上式可知，相似度范围从 -1 到 1，其中 -1 表示两个向量方向相反，1 表示两个向量指向相同；0 表示两个向量是独立的。在 -1 和 1 之间表示两个向量之间的相似性或相异性，可以理解地，相似度越接近 1 表示两个向量越接近。应用于本实施例，可根据实际经验预先设定 $\cos \theta$ 的距离阈值。若原始语音向量和标准语音向量的相似度大于距离阈值，则认为原始语音向量和标准语音向量相似，也即可判定目标语音特征和标准语音特征对应同一说话人的声纹识别结果。

本实施例中，通过余弦相似度算法即可判定原始语音向量和标准语音向量的相似度，简单快捷，利于快速确认识别结果。

综上，本实施例提出一种声纹识别方法，通过截顶语音检测算法可检测出待识别语音

是否为截顶语音段，若是则通过截顶语音修复模型可将截顶语音段的待识别语音特征修复为目标语音特征，以对比说话人对于的标准语音特征来识别说话人的真实身份。本实施例通过修复待识别语音中的待识别语音特征后获取接近说话人原始语音的目标语音特征，可有效提高语音识别的可靠性和准确性。

进一步地，识别服务器可对待识别语音进行特征提取后获得的目标语音特征，能够很准确地体现待识别语音的特征，利于基于待识别语音特征进行语音识别；采用高音采样点数量相对采样点总数的高音采样百分比对比预设阈值进行判定，可有效提高判定方法的灵活性和有利于提高判定结果的准确性；采用基于 DNN 模型生成的截顶语音修复模型来修复截顶语音段，可大幅提高语音修复的精度；通过余弦相似度算法即可判定原始语音向量和标准语音向量的相似度，简单快捷，利于快速确认识别结果。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

图 7 示出与上述实施例中声纹识别方法一一对应的声纹识别装置的原理框图。如图 7 所示，该声纹识别装置包括获取待识别语音模块 10、获取待识别特征模块 20、获取目标语音特征模块 30 和获取声纹识别结果模块 40。其中，获取待识别语音模块 10、获取语音段模块 20、修复语音段模块 30 和确定说话人模块 40 的实现功能与上述实施例中声纹识别方法对应的步骤一一对应，为避免赘述，本实施例不一一详述。

获取待识别语音模块 10，用于获取待识别语音，待识别语音携带说话人标识。

获取待识别特征模块 20，用于基于待识别语音，获取对应的待识别语音特征。

获取目标语音特征模块 30，用于采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征。

获取声纹识别结果模块 40，用于基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

优选地，获取待识别特征模块 20 包括获取语音数据单元 21、获取功率谱单元 22、获取梅尔功率谱单元 23 和获取梅尔系数单元 24。

获取语音数据单元 21，用于对待识别语音进行预处理，获取预处理语音数据。

获取功率谱单元 22，用于对预处理语音数据作快速傅里叶变换，获取待识别语音的频谱，并根据频谱获取待识别语音的功率谱。

获取梅尔功率谱单元 23，用于采用梅尔刻度滤波器组处理待识别语音的功率谱，获取待识别语音的梅尔功率谱。

获取梅尔系数单元 24，用于在梅尔功率谱上进行倒谱分析，获取待识别语音的梅尔频率倒谱系数。

优选地，获取目标语音特征模块 30 包括分割语音子段单元 31、获取采样点数量单元 32、获取采样百分比单元 33 和超过预设阈值单元 34。

分割语音子段单元 31，用于将待识别语音按时序平均分割出至少两个语音子段。

获取采样点数量单元 32，用于将语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量。

获取采样百分比单元 33，用于统计所有音量子区间的采样点总数，以获取高音采样点数量相对采样点总数的高音采样百分比。

超过预设阈值单元 34，用于若高音采样百分比超过预设阈值，则对应的语音子段为截

顶语音段。

优选地，该声纹识别装置还包括获取原始语音单元 35 和生成修复模型单元 36。

获取原始语音单元 35，用于获取原始训练语音对应的原始训练特征，对该原始训练语音进行截顶语音处理以获取对应的截顶训练语音，再提取该截顶训练语音的截顶训练特征。

生成修复模型单元 36，用于将截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

优选地，该声纹识别装置还包括获取原始特征单元 37。

获取原始特征单元 37，用于采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取修复语音段的目标语音特征。

优选地，获取声纹识别结果模块 40 包括采用识别模型单元 41、获取空间距离单元 42、获取识别结果单元 43。

采用识别模型单元 41，用于采用预设声纹识别模型分别处理目标语音特征和标准语音特征，分别得到原始语音向量和标准语音向量。

获取空间距离单元 42，用于获取原始语音向量和标准语音向量的空间距离。

获取识别结果单元 43，用于根据空间距离与预设的距离阈值，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

关于声纹识别装置的具体限定可以参见上文中对于声纹识别方法的限定，在此不再赘述。上述声纹识别装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一实施例中，提供一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 8 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储与声纹识别方法相关的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种声纹识别方法。

在一实施例中，提供一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现以下步骤：获取待识别语音，待识别语音携带说话人标识；基于待识别语音，获取对应的待识别语音特征；采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

在一实施例中，提取训练语音数据对应的训练语音特征，处理器执行计算机可读指令时实现如下步骤：对训练语音数据进行预处理，获取预处理语音数据；对预处理语音数据作快速傅里叶变换，获取训练语音数据的频谱，并根据频谱获取训练语音数据的功率谱；采用梅尔刻度滤波器组处理训练语音数据的功率谱，获取训练语音数据的梅尔功率谱；在梅尔功率谱上进行倒谱分析，获取训练语音数据的 MFCC 特征。

在一实施例中，采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，处理器执行计算机可读指令时实现如下步骤：将待识别语音按时序平均分割出至少两个语音子段；将语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量；统计所有音量子区间的采样点总数，以获取高音采样点数量相对采样点总数的高音采样百分比；若高音采样百分比超过预设阈值，则对应的语音子段为截顶语音段。

在一实施例中，在采用截顶语音修复模型修复待识别语音特征的步骤之前，处理器执行计算机可读指令时还实现如下步骤：获取原始训练语音，并截顶处理原始训练语音，以获取对应的截顶训练语音；将截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

在一实施例中，采用截顶语音修复模型修复待识别语音特征，获取目标语音特征，处理器执行计算机可读指令时实现如下步骤：采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取修复语音段的目标语音特征。

在一实施例中，基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，处理器执行计算机可读指令时实现如下步骤：采用预设声纹识别模型分别处理目标语音特征和标准语音特征，分别得到原始语音向量和标准语音向量；获取原始语音向量和标准语音向量的空间距离；根据空间距离与预设的距离阈值，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

在一实施例中，获取原始语音向量和标准语音向量的空间距离，处理器执行计算机可读指令时实现如下步骤：采用余弦相似度算法获取原始语音向量和标准语音向量的空间距离。

在一实施例中，提供一种计算机可读存储介质，其上存储有计算机可读指令，计算机可读指令被处理器执行时实现以下步骤：获取待识别语音，待识别语音携带说话人标识；基于待识别语音，获取对应的待识别语音特征；采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，则采用截顶语音修复模型修复待识别语音特征，获取目标语音特征；基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

在一实施例中，提取训练语音数据对应的训练语音特征，计算机可读指令被处理器执行时实现以下步骤：对训练语音数据进行预处理，获取预处理语音数据；对预处理语音数据作快速傅里叶变换，获取训练语音数据的频谱，并根据频谱获取训练语音数据的功率谱；采用梅尔刻度滤波器组处理训练语音数据的功率谱，获取训练语音数据的梅尔功率谱；在梅尔功率谱上进行倒谱分析，获取训练语音数据的 MFCC 特征。

在一实施例中，采用截顶语音检测算法检测待识别语音，若待识别语音为截顶语音段，计算机可读指令被处理器执行时实现以下步骤：将待识别语音按时序平均分割出至少两个语音子段；将语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量；统计所有音量子区间的采样点总数，以获取高音采样点数量相对采样点总数的高音采样百分比；若高音采样百分比超过预设阈值，则对应的语音子段为截顶语音段。

在一实施例中，在采用截顶语音修复模型修复待识别语音特征的步骤之前，计算机可

读指令被处理器执行时实现以下步骤：获取原始训练语音，并截顶处理原始训练语音，以获取对应的截顶训练语音；将截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

在一实施例中，采用截顶语音修复模型修复待识别语音特征，获取目标语音特征，计算机可读指令被处理器执行时实现以下步骤：采用基于 DNN 模型的截顶语音修复模型修复待识别语音特征，获取修复语音段的目标语音特征。

在一实施例中，基于说话人标识对应的标准语音特征，采用预设声纹识别模型对目标语音特征和标准语音特征进行声纹识别，计算机可读指令被处理器执行时实现以下步骤：采用预设声纹识别模型分别处理目标语音特征和标准语音特征，分别得到原始语音向量和标准语音向量；获取原始语音向量和标准语音向量的空间距离；根据空间距离与预设的距离阈值，获取目标语音特征和标准语音特征是否对应同一说话人的声纹识别结果。

在一实施例中，获取原始语音向量和标准语音向量的空间距离，计算机可读指令被处理器执行时实现以下步骤：采用余弦相似度算法获取原始语音向量和标准语音向量的空间距离。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程 ROM (PROM)、电可编程 ROM (EPROM)、电可擦除可编程 ROM (EEPROM) 或闪存。易失性存储器可包括随机存取存储器 (RAM) 或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权 利 要 求 书

1. 一种声纹识别方法，其特征在于，包括：

获取待识别语音，所述待识别语音携带说话人标识；

基于所述待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，则采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征；

基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

2. 如权利要求 1 所述声纹识别方法，其特征在于，所述基于所述待识别语音，获取对应的待识别语音特征，包括：

对所述待识别语音进行预处理，获取预处理语音数据；

对所述预处理语音数据作快速傅里叶变换，获取所述待识别语音的频谱，并根据所述频谱获取所述待识别语音的功率谱；

采用梅尔刻度滤波器组处理所述待识别语音的功率谱，获取所述待识别语音的梅尔功率谱；

在所述梅尔功率谱上进行倒谱分析，获取所述待识别语音的梅尔频率倒谱系数。

3. 如权利要求 1 所述声纹识别方法，其特征在于，所述采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，包括：

将所述待识别语音按时序平均分割出至少两个语音子段；

将所述语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量；

统计所有所述音量子区间的采样点总数，以获取所述高音采样点数量相对所述采样点总数的高音采样百分比；

若所述高音采样百分比超过预设阈值，则对应的所述语音子段为截顶语音段。

4. 如权利要求 1 所述声纹识别方法，其特征在于，在所述采用截顶语音修复模型修复所述待识别语音特征的步骤之前，所述声纹识别方法还包括：

获取原始训练语音对应的原始训练特征，对该原始训练语音进行截顶语音处理以获取对应的截顶训练语音，再提取该截顶训练语音的截顶训练特征；

将所述截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将所述原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准所述 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

5. 如权利要求 4 所述声纹识别方法，其特征在于，所述采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征，包括：

采用所述基于 DNN 模型的截顶语音修复模型修复所述待识别语音特征，获取目标语音特征。

6. 如权利要求 1 所述声纹识别方法，其特征在于，所述基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，包括：

采用预设声纹识别模型分别处理所述目标语音特征和所述标准语音特征，分别得到原始语音向量和标准语音向量；

获取所述原始语音向量和所述标准语音向量的空间距离；

根据所述空间距离与预设的距离阈值，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

7. 如权利要求 6 所述声纹识别方法，其特征在于，所述获取所述原始语音向量和所述标准语音向量的空间距离，包括：

采用余弦相似度算法获取所述原始语音向量和所述标准语音向量的空间距离。

8. 一种声纹识别装置，其特征在于，包括：

获取待识别语音模块，用于获取待识别语音，所述待识别语音携带说话人标识；

获取待识别特征模块，用于基于所述待识别语音，获取对应的待识别语音特征；

获取目标语音特征模块，用于采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，则采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征；

获取声纹识别结果模块，用于基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

9. 一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，其特征在于，所述处理器执行所述计算机可读指令时实现如下步骤：

获取待识别语音，所述待识别语音携带说话人标识；

基于所述待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，则采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征；

基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

10. 如权利要求 9 所述的计算机设备，其特征在于，所述基于所述待识别语音，获取对应的待识别语音特征，包括：

对所述待识别语音进行预处理，获取预处理语音数据；

对所述预处理语音数据作快速傅里叶变换，获取所述待识别语音的频谱，并根据所述频谱获取所述待识别语音的功率谱；

采用梅尔刻度滤波器组处理所述待识别语音的功率谱，获取所述待识别语音的梅尔功率谱；

在所述梅尔功率谱上进行倒谱分析，获取所述待识别语音的梅尔频率倒谱系数。

11. 如权利要求 9 所述的计算机设备，其特征在于，所述采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，包括：

将所述待识别语音按时序平均分割出至少两个语音子段；

将所述语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量；

统计所有所述音量子区间的采样点总数，以获取所述高音采样点数量相对所述采样点

总数的高音采样百分比；

若所述高音采样百分比超过预设阈值，则对应的所述语音子段为截顶语音段。

12. 如权利要求 9 所述的计算机设备，其特征在于，在所述采用截顶语音修复模型修复所述待识别语音特征的步骤之前，所述处理器执行所述计算机可读指令时还实现如下步骤：

获取原始训练语音对应的原始训练特征，对该原始训练语音进行截顶语音处理以获取对应的截顶训练语音，再提取该截顶训练语音的截顶训练特征；

将所述截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将所述原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准所述 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

13. 如权利要求 12 所述的计算机设备，其特征在于，所述采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征，包括：

采用所述基于 DNN 模型的截顶语音修复模型修复所述待识别语音特征，获取目标语音特征。

14. 如权利要求 9 所述的计算机设备，其特征在于，所述基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，包括：

采用预设声纹识别模型分别处理所述目标语音特征和所述标准语音特征，分别得到原始语音向量和标准语音向量；

获取所述原始语音向量和所述标准语音向量的空间距离；

根据所述空间距离与预设的距离阈值，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

15. 一个或多个存储有计算机可读指令的非易失性可读存储介质，其特征在于，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

获取待识别语音，所述待识别语音携带说话人标识；

基于所述待识别语音，获取对应的待识别语音特征；

采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，则采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征；

基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

16. 如权利要求 15 所述的非易失性可读存储介质，其特征在于，所述基于所述待识别语音，获取对应的待识别语音特征，包括：

对所述待识别语音进行预处理，获取预处理语音数据；

对所述预处理语音数据作快速傅里叶变换，获取所述待识别语音的频谱，并根据所述频谱获取所述待识别语音的功率谱；

采用梅尔刻度滤波器组处理所述待识别语音的功率谱，获取所述待识别语音的梅尔功率谱；

在所述梅尔功率谱上进行倒谱分析，获取所述待识别语音的梅尔频率倒谱系数。

17. 如权利要求 15 所述的非易失性可读存储介质，其特征在于，所述采用截顶语音检测算法检测所述待识别语音，若所述待识别语音为截顶语音段，包括：

将所述待识别语音按时序平均分割出至少两个语音子段；

将所述语音子段按音量变化均匀分割成至少两个音量子区间，获取最高音所在的音量子区间的高音采样点数量；

统计所有所述音量子区间的采样点总数，以获取所述高音采样点数量相对所述采样点总数的高音采样百分比；

若所述高音采样百分比超过预设阈值，则对应的所述语音子段为截顶语音段。

18. 如权利要求 15 所述的非易失性可读存储介质，其特征在于，在所述采用截顶语音修复模型修复所述待识别语音特征的步骤之前，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

获取原始训练语音对应的原始训练特征，对该原始训练语音进行截顶语音处理以获取对应的截顶训练语音，再提取该截顶训练语音的截顶训练特征；

将所述截顶训练语音对应的截顶训练特征作为 DNN 模型的输入层，将所述原始训练语音对应的原始训练特征作为 DNN 模型的输出层，校准所述 DNN 模型的特征参数，以生成基于 DNN 模型的截顶语音修复模型。

19. 如权利要求 18 所述的非易失性可读存储介质，其特征在于，所述采用截顶语音修复模型修复所述待识别语音特征，获取目标语音特征，包括：

采用所述基于 DNN 模型的截顶语音修复模型修复所述待识别语音特征，获取目标语音特征。

20. 如权利要求 15 所述的非易失性可读存储介质，其特征在于，所述基于所述说话人标识对应的标准语音特征，采用预设声纹识别模型对所述目标语音特征和所述标准语音特征进行声纹识别，包括：

采用预设声纹识别模型分别处理所述目标语音特征和所述标准语音特征，分别得到原始语音向量和标准语音向量；

获取所述原始语音向量和所述标准语音向量的空间距离；

根据所述空间距离与预设的距离阈值，获取所述目标语音特征和所述标准语音特征是否对应同一说话人的声纹识别结果。

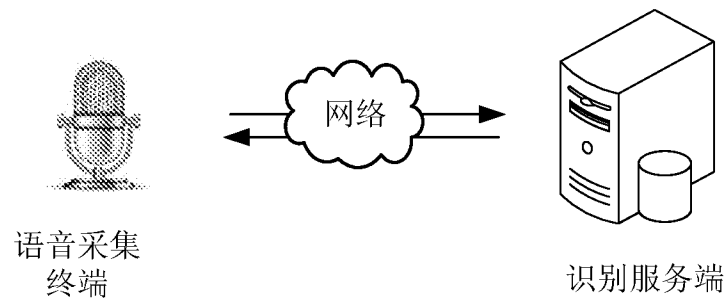


图 1

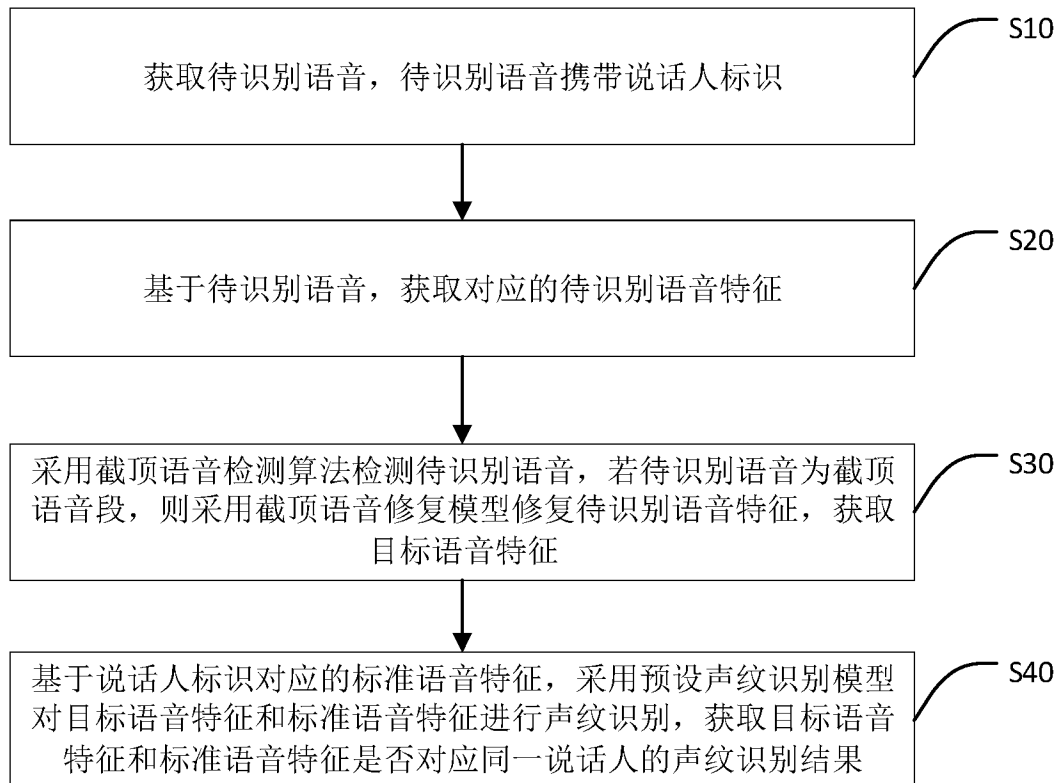


图 2

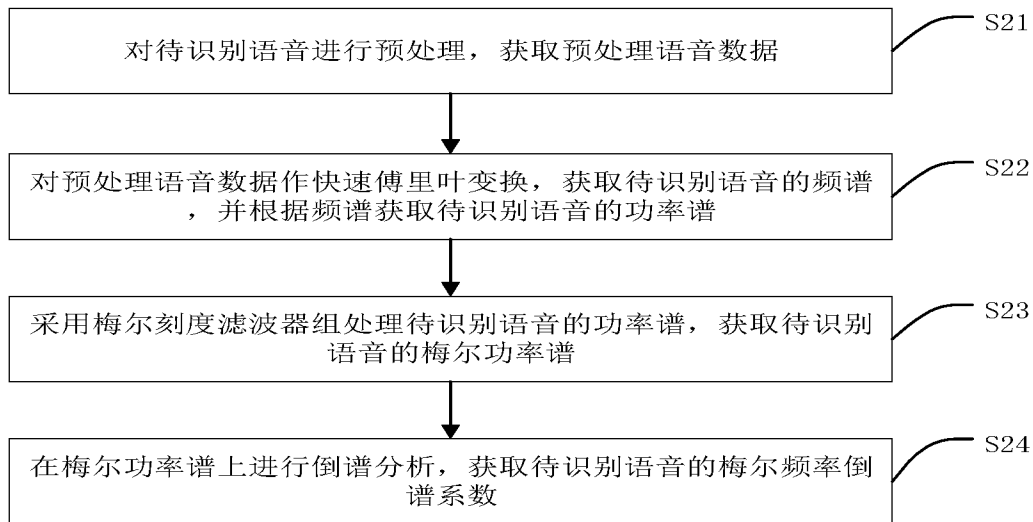


图 3

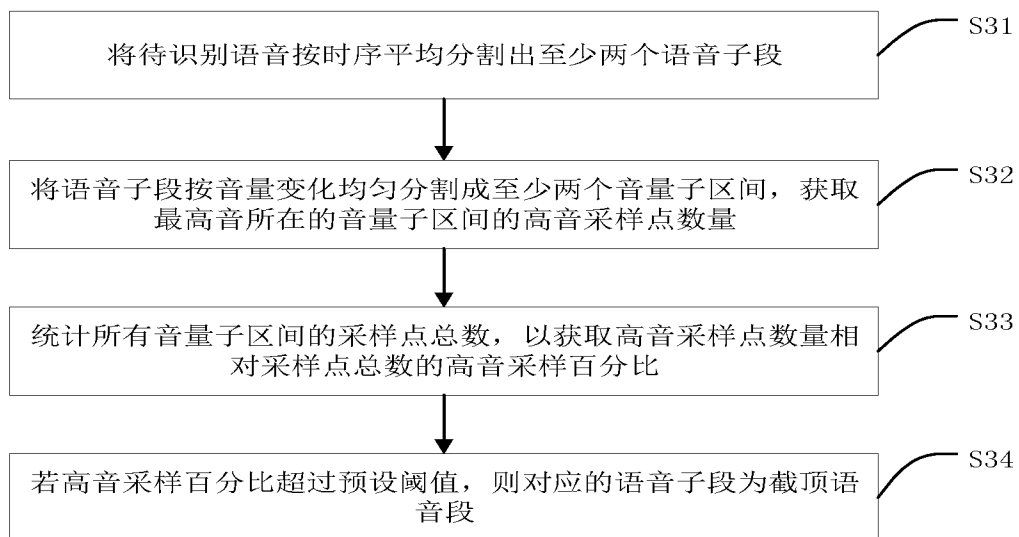


图 4

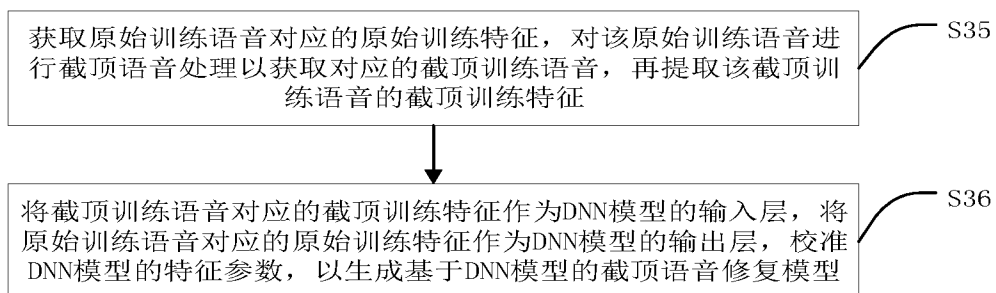


图 5

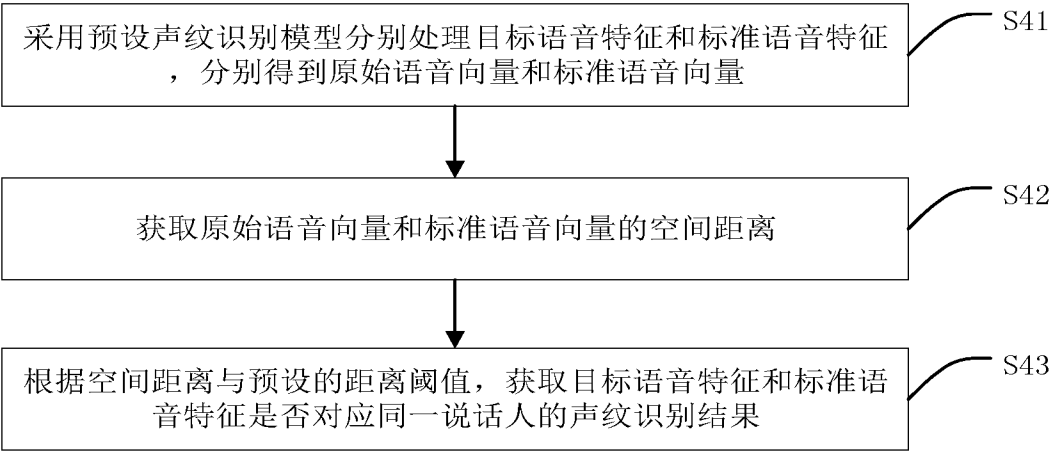


图 6

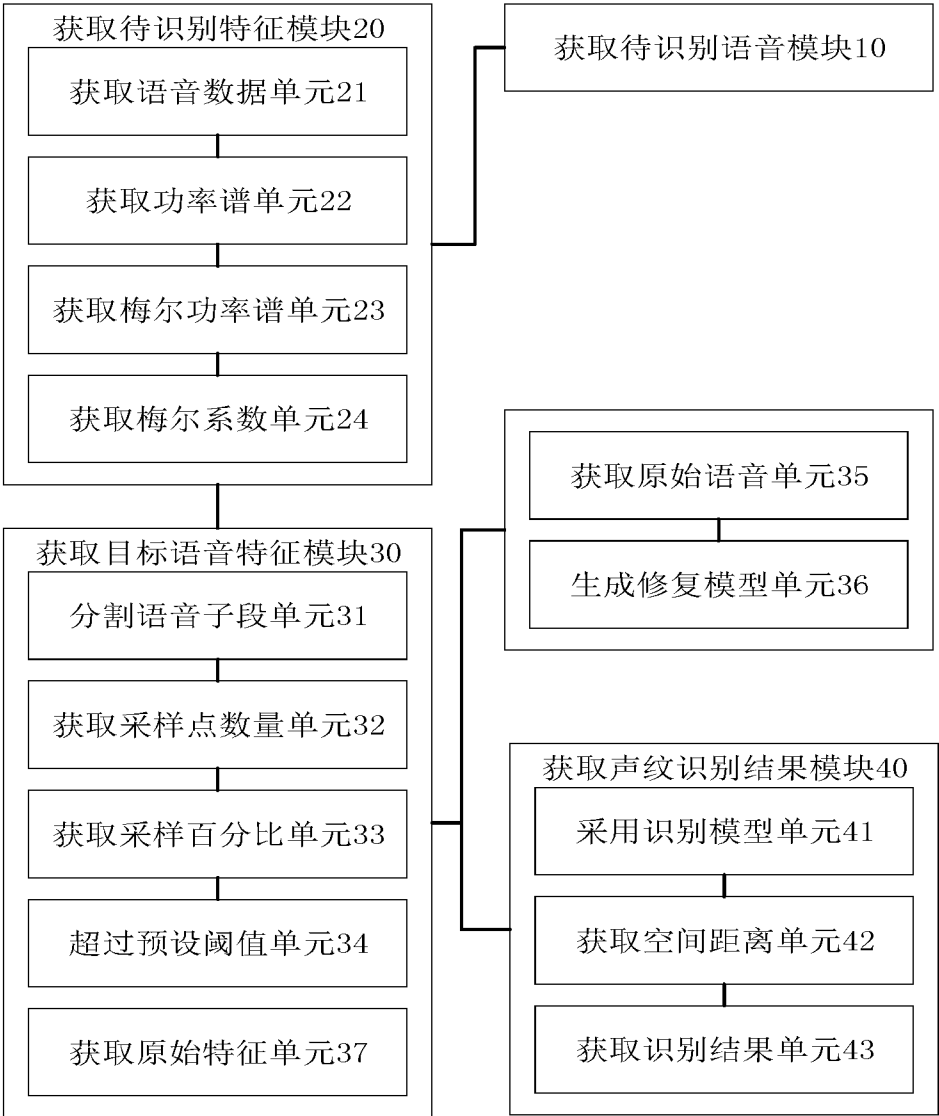


图 7

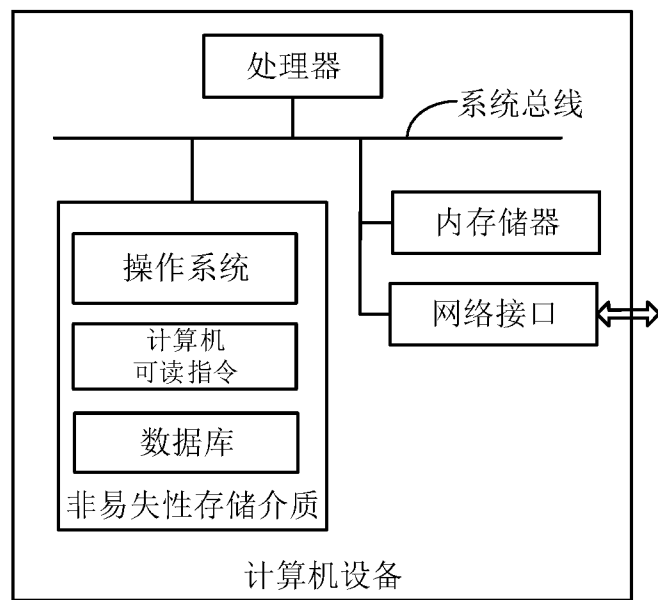


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/092598

A. CLASSIFICATION OF SUBJECT MATTER

G10L 17/00(2013.01)i; G10L 25/60(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)

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)

CNABS, DWPI, CNKI, IEEE, 百度学术, BAIDU XUESHU: 截顶, 截幅, 削顶, 修复, 重建, clipped signal, clipped, speech, detect, DNN

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 108899032 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 27 November 2018 (2018-11-27) claims 1-10	1-20
X	Fanhu Bie et al. "Detection and Reconstruction of Clipped Speech for Speaker Recognition" <i>Speech Communication</i> , 02 July 2015 (2015-07-02), 218-231	1-20
X	LI Chun-zhi et al. "Restoration of Clipped Vibration Signal Based on BP Neural Network" <i>2010 International Conference on Measuring Technology and Mechatronics Automation</i> , 14 March 2010 (2010-03-14), 251-253	1-20
A	CN 108091352 A (TENCENT MUSIC ENTERTAINMENT TECHNOLOGY (SHENZHEN) CO., LTD.) 29 May 2018 (2018-05-29) entire document	1-20
A	CN 106847292 A (PING AN TECHNOLOGY SHENZHEN CO., LTD.) 13 June 2017 (2017-06-13) 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

02 March 2019

Date of mailing of the international search report

14 March 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/092598

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101605111 A (HUAWEI TECHNOLOGIES CO., LTD.) 16 December 2009 (2009-12-16) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/092598

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108899032	A	27 November 2018	None			
CN	108091352	A	29 May 2018	None			
CN	106847292	A	13 June 2017	KR	20180104595	A	21 September 2018
				CN	106847292	B	19 June 2018
				SG	11201803895R	A	27 September 2018
				AU	2017341161	A1	30 August 2018
				WO	2018149077	A1	23 August 2018
CN	101605111	A	16 December 2009	EP	2448204	A4	02 May 2012
				US	2012095580	A1	19 April 2012
				US	8862257	B2	14 October 2014
				EP	2448204	A1	02 May 2012
				KR	101388540	B1	23 April 2014
				JP	2012531142	A	06 December 2012
				CN	101605111	B	04 July 2012
				EP	2448204	B1	16 August 2017
				WO	2010149053	A1	29 December 2010
				JP	5512809	B2	04 June 2014
				KR	20120013431	A	14 February 2012

国际检索报告

国际申请号

PCT/CN2018/092598

A. 主题的分类

G10L 17/00(2013.01)i; G10L 25/60(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)

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))

CNABS, DWPI, CNKI, IEEE, 百度学术: 截顶, 截幅, 削顶, 修复, 重建, clipped signal, clipped, speech, detect, DNN

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 108899032 A (平安科技深圳有限公司) 2018年 11月 27日 (2018 - 11 - 27) 权利要求1-10	1-20
X	Fanhu Bie 等. "Detection and reconstruction of clipped speech for speaker recognition" Speech Communication, 2015年 7月 2日 (2015 - 07 - 02), 218 - 231	1-20
X	LI Chun-zhi 等. "Restoration of Clipped vibration signal Based on BP Neural Network" 2010 International Conference on Measuring Technology and Mechatronics Automation, 2010年 3月 14日 (2010 - 03 - 14), 251-253	1-20
A	CN 108091352 A (腾讯音乐娱乐科技深圳有限公司) 2018年 5月 29日 (2018 - 05 - 29) 全文	1-20
A	CN 106847292 A (平安科技深圳有限公司) 2017年 6月 13日 (2017 - 06 - 13) 全文	1-20
A	CN 101605111 A (华为技术有限公司) 2009年 12月 16日 (2009 - 12 - 16) 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

2019年 3月 2日

国际检索报告邮寄日期

2019年 3月 14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

受权官员

游晓梅

传真号 (86-10)62019451

电话号码 (86-10)62089539

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/092598

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	108899032	A	2018年 11月 27日	无			
CN	108091352	A	2018年 5月 29日	无			
CN	106847292	A	2017年 6月 13日	KR	20180104595	A	2018年 9月 21日
				CN	106847292	B	2018年 6月 19日
				SG	11201803895R	A	2018年 9月 27日
				AU	2017341161	A1	2018年 8月 30日
				WO	2018149077	A1	2018年 8月 23日
CN	101605111	A	2009年 12月 16日	EP	2448204	A4	2012年 5月 2日
				US	2012095580	A1	2012年 4月 19日
				US	8862257	B2	2014年 10月 14日
				EP	2448204	A1	2012年 5月 2日
				KR	101388540	B1	2014年 4月 23日
				JP	2012531142	A	2012年 12月 6日
				CN	101605111	B	2012年 7月 4日
				EP	2448204	B1	2017年 8月 16日
				WO	2010149053	A1	2010年 12月 29日
				JP	5512809	B2	2014年 6月 4日
				KR	20120013431	A	2012年 2月 14日

表 PCT/ISA/210 (同族专利附件) (2015年1月)