

(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 2020/252935 A1

(43) 国际公布日

2020 年 12 月 24 日 (24.12.2020)

(51) 国际专利分类号:

G10L 15/26 (2006.01) G10L 17/00 (2013.01)

(21) 国际申请号:

PCT/CN2019/103843

(22) 国际申请日:

2019 年 8 月 30 日 (30.08.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910522762.0 2019年6月17日 (17.06.2019) 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 518057 (CN)。

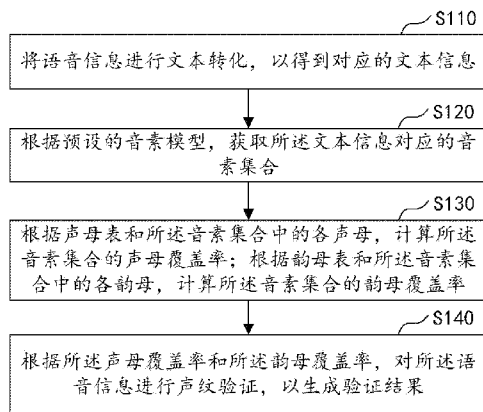
(72) 发明人: 吴冀平(WU, Jiping); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518057 (CN)。 彭俊清(PENG, Junqing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518057 (CN)。 王健宗(WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518057 (CN)。

(74) 代理人: 深圳市力道知识产权代理有限公司(普通合伙)(SHENZHEN LIDAO INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦一路 50 号鹏基商务时空大厦 1717-1719, Guangdong 518000 (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,

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

(54) 发明名称: 声纹验证方法、装置、设备及存储介质



S110 Perform text conversion on speech information to obtain corresponding text information
S120 Acquire, according to a preset phoneme model, a phoneme set corresponding to the text information
S130 Calculate the coverage rate of initials in the phoneme set according to a list of initials and each initial in the phoneme set, and calculate the coverage rate of finals in the phoneme set according to a list of finals and each final in the phoneme set
S140 Perform voiceprint verification on the speech information according to the coverage rate of the initials and the coverage rate of the finals so as to generate a verification result

图 1

(57) Abstract: A voiceprint verification method, apparatus and device, and a storage medium. The method comprises: converting speech information into corresponding text information; acquiring a phoneme set corresponding to the text information; calculating the coverage rate of initials in the phoneme set according to a list of initials and each initial in the phoneme set; calculating the coverage rate of finals in the phoneme set according to a list of finals and each final in the phoneme set; and performing voiceprint verification on the speech information according to the coverage rate of the initials and the coverage rate of the finals so as to generate a verification result.

(57) 摘要: 一种声纹验证方法、装置、设备及存储介质, 该方法包括: 将语音信息转化为对应的文本信息; 获取文本信息对应的音素集合; 根据声母表和音素集合中的各声母, 计算音素集合的声母覆盖率; 根据韵母表和音素集合中的各韵母, 计算音素集合的韵母覆盖率; 根据声母覆盖率和韵母覆盖率, 对语音信息进行声纹验证, 以生成验证结果。

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, 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))。

声纹验证方法、装置、设备及存储介质

本申请要求于2019年6月17日提交中国专利局、申请号为201910522762.0、发明名称为“声纹验证方法、装置、设备及存储介质”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

本申请涉及生物识别领域，尤其涉及一种声纹验证方法、装置、设备及存储介质。

背景技术

在言语无关的说话人识别系统中，声纹注册的语音通常会让用户随意说话，说话时长超过一定的阈值即可，通过这一段语音，提取说话人的发音特征，使用机器学习的方法提取一系列特征向量。一般对于这段语音，要求信噪比在一定阈值以上。然而，信噪比符合要求的语音难以完整地体现出用户的语音特征。比如，用户在说话的这段时间内，一直重复同一个单词，那么这段语音虽然时长和信噪比都可以达标，但是对于所反映的发音特征是非常有限的。

发明内容

本申请提供了一种声纹验证方法、装置、设备及存储介质，为声纹身份验证提供了重要参考。

第一方面，本申请提供了一种声纹验证方法，所述方法包括：

将语音信息进行文本转化，以得到对应的文本信息；

根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；

根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；

根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

第二方面，本申请还提供了一种声纹验证装置，所述装置包括：

文本转化单元，用于将语音信息进行文本转化，以得到对应的文本信息；

音素获取单元，用于根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；

覆盖率计算单元，用于根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；

声纹验证单元，用于根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

第三方面，本申请还提供了一种计算机设备，所述计算机设备包括存储器和处理器；所述存储器用于存储计算机程序；所述处理器，用于执行所述计算机程序并在执行所述计算机程序时实现如上述的声纹验证方法。

第四方面，本申请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使所述处理器实现如上述的声纹验证方法。

本申请公开了一种声纹验证方法、装置、设备及存储介质，通过将语音信息进行文本转化，以得到对应的文本信息；根据预设的音素模型，获取所述文本信息对应的音素集合；根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果，从而能够知悉该语音信息是否具有用户发声音的声纹特征，是否能够涵盖该用户语音的大部分特征，进而找出具有涵盖用户大部分语音特征、用户语音特征完整度高的语音信息，为声纹身份验证提供了重要参考。

附图说明

为了更清楚地说明本申请实施例技术方案，下面将对实施例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的前提下，还可以根据这些附图获得其他的附图。

图1是本申请的实施例提供的一种声纹验证方法的示意图；

图2是图1中的声纹验证方法的子步骤示意图；

图3是本申请一实施例提供的获得音素集合的步骤示意图；

图4是本申请另一实施例提供的获得音素集合的步骤示意图；

图5是本申请一实施例提供的计算声母覆盖率和韵母覆盖率的步骤示意图；

图6是本申请另一实施例提供的计算声母覆盖率和韵母覆盖率的步骤示意图；

图7是本申请的另一实施例提供的声纹验证方法的示意图；

图8是图7中的声纹验证方法的子步骤示意图；

图9是本申请一实施例提供的声纹验证方法的子步骤示意图；

图10是本申请另一实施例提供的声纹验证方法的子步骤示意图；

图11是本申请的再一实施例提供的声纹验证方法的示意图。

图12是本申请的实施例提供的声纹验证装置的示意性框图；

图13是图12中声纹验证装置的子单元的示意性框图；

图14是图13中国音素获取子单元的子模块的示意性框图；

图15是图12中声纹验证装置的子单元的示意性框图；

图16是图12中声纹验证装置的子单元的示意性框图；

图17为本申请一实施例提供的一种计算机设备的结构示意图。

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。

附图中所示的流程图仅是示例说明，不是必须包括所有的内容和操作/步骤，也不是必须按所描述的顺序执行。例如，有的操作/步骤还可以分解、组合或部分合并，因此实际执行的顺序有可能根据实际情况改变。

本申请的实施例提供了一种声纹验证方法、装置、计算机设备及存储介质。该声纹验证方法可用于针对用户声纹注册时，找出用户语音特征完整度高的语音信息，为用户的声纹身份验证提供了重要的参考。

下面结合附图，对本申请的一些实施方式作详细说明。在不冲突的情况下，下述的实施例及实施例中的特征可以相互组合。

请参阅图 1，图 1 是本申请实施例提供的一种声纹验证方法的步骤示意图。

如图 1 所示，该声纹验证方法，具体包括：步骤 S110 至步骤 S140。

S110、将语音信息进行文本转化，以得到对应的文本信息。

在一实施例中，将语音信息进行文本转化的具体过程，即步骤 S110 具体包括：当处于连接外网状态时，将所述语音信息上传至云平台；接收所述云平台根据所述语音信息转化后的文本信息。

具体的，将该语音信息进行压缩打包处理，然后上传到云平台，通过云平台对语音信息进行识别转化为文本信息。其中，云平台是指由多台计算机组成的用于提供语音识别服务的网络平台。

在一实施例中，将语音信息进行文本转化的具体过程，即步骤 S110 具体包括：当处于未连接外网状态时，在本地对所述语音信息进行识别，并转化为文本信息。具体的，在本地安装有对语音进行识别的应用程序，且存储有识别语音的数据库。

在一实施例中，所述将语音信息进行文本转化，以得到对应的文本信息，即步骤 S110 之前还包括：接收所述语音信息。

具体的，通过麦克风或话筒等音频输入设备接收用户输入的语音信息。

在一实施例中，用户可以随意说话，也可以朗读预设文本，终端或服务器通过音频输入设备接收用户的语音信息。在接收该语音信息后，获取该语音信息对应的音素集合，直接计算声母覆盖率和韵母覆盖率，从而对该语音信息进行声纹验证。

S120、根据预设的音素模型，获取所述文本信息对应的音素集合。

具体的，所述音素集合包括所述文本信息中每个字所对应的声母和韵母。如图 2 所示，在一实施例中，所述根据预设的音素模型，获取所述文本信息对应的音素集合，具体包括：子步骤 S121、S122 和 S123。

S121、对所述文本信息进行分词处理，以得到多个词串。

具体的，步骤 S121 具体包括：对所述文本信息进行语句切分，以得到切分语句；对各所述切分语句进行分词处理，以得到各所述切分语句对应的词串。

具体的，对所述语音信息进行文本转化后，可对转化后的文本信息进行语句切分，例如可根据标点符号将各个文本切分成一条条完整的语句，从而得到该文本信息对应的若干切分语句。然后，对各个切分语句进行分词处理，从而得到多个词串。

在一实施例中，对各个切分语句进行分词处理的方法可以采用字符串匹配的分词方法，例如正向最大匹配法、反向最大匹配法、最短路径分词法和双向最大匹配法等。其中，正向最大匹配法是指把一个切分的语句中的字符串从左至右来分词。反向最大匹配法是指把一个切分的语句中的字符串从右至左来分词。双向最大匹配法是指正反向（从左到右、从右到左）同时进行分词匹配。

最短路径分词法是指一个切分的语句中的字符串里面要求切出的词数是最少的。

在其他实施例中，对各个切分语句进行分词处理的方法可以为其他任意合适的分词方法，例如通过词义分词法对各个切分后的语句进行分词处理。其中，词义分词法是一种机器语音判断的分词方法，利用句法信息和语义信息来处理歧义现象来分词。

示例性的，获取具有词语集的汉语词典库，通过双向最大匹配法对文本信息与汉语词典库中的词语进行遍历分割匹配，从而实现对所述文本信息进行分词。其中，汉语词典库中的常用词语按首字母排序。例如，汉语词典库可以为《现代汉语词典》。

具体的，假设汉语词典库的最长词组的长度为 m ，文本信息 S 经语句切分后，得到若干切分语句。正反向同时将切分语句中词组长度为 m 的连续字符与汉语词典库中的词语进行匹配。若切分语句与汉语词典库中的各词语匹配不成功，则逐次减小连续字符的长度进行多次扫描匹配，直至该语句与汉语词典库中的某一词语匹配成功，最终将文本信息 S 分解为多个词串，即得到词串 $FS1$ 、 $FS2$ 、...、 FSN 。

S122、对各所述词串进行拼音转换，以得到各所述词串对应的拼音串。

示例性的，文本信息 S 经分词处理后，得到 N 个词串，分别为 $FS1$ 、 $FS2$ 、...、 FSN 。 N 个词串分别经拼音转换后，得到各词串对应的拼音串， $PS1$ 、 $PS2$ 、...、 PSN 。例如，词串“张三”经拼音转换后，得到子拼音串“zhang1san1”，其中数字 1 表示声调为阴平。

S123、将各所述拼音串输入预设的音素模型，以得到音素集合。

在一实施例中，所述将各所述拼音串输入预设的音素模型，以得到音素集合之前，还包括：获取标准发音语音库；根据所述标准发音语音库，对预设的隐马尔科夫模型进行模型训练，以建立音素模型。

在一实施例中，获取标准发音语音库可以具体包括：获取多个原始录音数据以及对应的标注；对各所述原始录音数据和各原始录音数据的标注进行筛选修正处理，以得到标准发音语音库。

具体的，原始录音数据可以来源于互联网，也可以通过录音设备如录音笔录入获取。通过自动或人工方式对原始录音数据和原始录音数据对应的标注进行多轮检查和筛选修正处理，得到标准语音数据。各标准语音数据的集合构造为所述标准发音语音库。

其中，所述标注包括声调标注。对各所述原始录音数据和各原始录音数据的标注进行筛选修正处理，以得到标准发音语音库，具体可以包括：去除各所述原始录音数据中声调发音模糊的数据；根据汉语词典库，修正所述原始录音数据对应的声调标注。

可以理解的，在其他实施例中，获取标准发音语音库可以通过互联网直接获取。

如图 3 所示，在一实施例中，所述音素模型包括声母子模型和韵母子模型。音素集合的获得过程，即步骤 S123，具体包括子步骤 S123a、S123b 和 S123c。

S123a、将所述拼音串输入所述声母子模型，以输出所述词串中各字对应的声母。

具体的，每个音节包括一个韵母，可能还包括一个声母。声母为辅音，韵母由单元音或双元音开头。声母相应于音节的声母部分，韵母相应于音节的韵母部分。汉语词典库中共有 23 个声母。23 个声母中包括汉语拼音中的 21 个声

母、w 和 y。w 和 y 在《汉语拼音方案》中不被作为声母，但根据人们的习惯拼法，会将 w 和 y 使用声母拼韵母的方式拼出，比如将 yan 使用声母拼韵母的方式拼出，即 y-an-yan，故本申请中把 w 和 y 也作为声母。具体的，23 个声母具体如表 1 所示。

表 1 是汉语词典库声母表

b	p	m	f	d	t
n	l	g	k	h	j
q	x	zh	ch	sh	r
z	c	s	w	y	

S123b、将所述拼音串输入所述韵母子模型，以输出所述词串中各字对应的韵母和各所述韵母对应的声调。

其中，汉语词典库中共有 35 个声母，如表 2 所示。

表 2 是汉语词典库韵母表

	i	u	ü
a	ia	ua	
o		uo	
e	ie		üe
ai		uai	
ei		uei	
ao	iao		
ou	iou		
an	ian	uan	üan
en	in	un	ün
ang	iang	uang	
eng	ing	ueng	
ong	iong		

具体的，表 2 中有一部分韵母，在组成音节时会缩写。比如“iou”，“有”字的拼音写成“you”，“有”字的韵母“iou”缩写为“ou”。在一实施例中，在输出韵母时，只考虑表 2 中出现的韵母，缩写的韵母将会被还原成完整的形式。

其中，汉语词典库中的声调包括五种，分别为第一声（亦称阴平或平调）、第二声（亦称阳平或声调）、第三声（亦称上声或折调）、第四声（亦称去声或降调）、轻声。

S123c、根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合。

具体的，将所述词串中各字对应的声母、所述词串中各字对应的韵母和各所述韵母对应的声调构建为音素集合。

如图4所示，在一实施例中，所述音素模型包括音节子模型和韵母子模型。所述根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合之前，还包括步骤S101、将所述拼音串输入所述音节子模型，以输出所述词串中各字对应的音节。

具体的，汉语拼音中共有潜在的3990个音节（声母和韵母的所有可能组合）。但是并非每个声母、韵母和音调的可能组合都能构成合法音节。实际上只有不含声调的大约416个合法音节，和大约1300多个有意义的带调音节。

其中，所述根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合，具体包括：根据所述词串中各字对应的声母、韵母、音节以及各所述韵母对应的声调，构建音素集合。

S130、根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率。

具体的，声母表可以是《汉语拼音方案》中的声母表，韵母表可以是《汉语拼音方案》中的韵母表。

如图5所示，在一实施例中，所述根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率，即步骤S130具体包括子步骤S131、S132和S133。

S131、计算所述音素集合中声母的数量和韵母的数量。

具体的，将音素集合中互不相同的声母进行统计求和，得到音素集合对应的声母的数量。同样的，将音素集合中互不相同的韵母进行统计求和，得到音素集合对应的韵母的数量。

在一实施例中，声母和韵母的数量计算过程，即步骤S131包括：根据所述词串对应的音节和声母，统计所述文本信息中出现的声母的数量；根据所述词串对应的音节和韵母，统计所述文本信息中出现的韵母的数量。

比如，文本信息“张三喜欢跑步”的拼音为“zhang1san1xi3huan1pao3bu4”。该文本信息中出现了六个声母和六个韵母，六个声母为“zh、s、x、h、d、q”，六个韵母为“ang、an、i、uan、ao、u”。文本信息“张三喜欢跑步”的拼音中没有出现的声母有17个，具体如表3所示。

表3是文本信息“张三喜欢跑步”的拼音中未出现的声母展示表

b	p	m	f	t	n
l	g	k	j	ch	sh
r	z	c	w	y	

其中，文本信息“张三喜欢跑步”的拼音中没有出现的韵母有29个，具体如表4所示。

表 4 是文本信息“张三喜欢跑步”的拼音中未出现的韵母展示表

			ü
a	ia	ua	
o		uo	
e	ie		üe
ai		uai	
ei		uei	
	iao		
ou	iou		
	ian		üan
en	in	un	ün
	iang	uang	
eng	ing	ueng	
ong	iong		

如图 6 所示，在一实施例中，步骤 S131 计算所述音素集合中声母的数量和韵母的数量，之前还可以包括：

S102、对所述音素集合中的声母、韵母和音节进行去重处理，以得到去重音素集合。

具体的，音素集合中某一声母多次重复出现，将该声母重复的部分舍弃，使得该声母在音素集合中只出现一次。同样的，韵母和音节去重的方法可以参照声母的去重方法，在此不再赘述。所述计算所述音素集合中声母的数量和韵母的数量，具体包括：计算所述去重音素集合中声母的数量和韵母的数量。

S132、基于声母覆盖率公式，根据所述声母的数量和所述声母表，计算所述声母覆盖率。

其中，所述声母覆盖率公式为：

$$\alpha = \frac{S}{M}$$

其中， α 为所述声母覆盖率，S 为声母的数量，M 为汉语词典库中声母表的声母的总数量。

比如，文本信息“张三喜欢跑步”的拼音为“zhang1san1xi3huan1pao3bu4”。该文本信息中出现了六个声母，六个声母为“zh、s、x、h、d、q”，声母的数量为 6。汉语词典库中声母表的声母的总数量为 23，声母覆盖率=6/23=26.09%。

S133、基于韵母覆盖率公式，根据所述韵母的数量和所述韵母表，计算所述韵母覆盖率。

具体的，所述韵母覆盖率公式为：

$$\beta = \frac{Y}{Q}$$

其中， β 为所述韵母覆盖率， S 为韵母的数量， M 为汉语词典库中韵母表的韵母的总数量。

比如，文本信息“张三喜欢跑步”的拼音为“zhang1san1xi3huan1pao3bu4”。该文本信息中出现了六个韵母，六个韵母为“ang、an、i、uan、ao、u”，韵母的数量为6。汉语词典库中韵母表的韵母的总数量为35，韵母覆盖率 $=6/35=17.14\%$ 。

S140、根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

具体的，验证结果可以为声纹验证通过或声纹验证未通过两种。声纹验证通过的验证结果可认为用户输入的语音信息具有用户发音的声纹特征，能够涵盖该用户语音的大部分特征，符合声纹注册的深层次需求。声纹验证未通过的验证结果可认为用户输入的语音信息不具有用户发音的声纹特征，不能涵盖该用户语音的大部分特征，因而不符合声纹注册的深层次需求。

如图7所示，在一实施例中，步骤S140根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果之前，还可以包括步骤S103、根据音节表和所述音素集合中的各音节，计算所述音素集合的音节覆盖率。对所述语音信息进行声纹验证的具体过程，即步骤S140具体包括：根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

如图8所示，步骤S103根据音节表和所述音素集合中的各音节，计算所述音素集合的音节覆盖率包括子步骤S103a和S103b。

S103a、计算所述音素集合中音节的数量。

具体的，将音素集合中互不相同的音节进行统计求和，得到音素集合对应的音节的数量。

S103b、根据音节表和所述音节的数量，计算所述音素集合的音节覆盖率。

具体的，音节覆盖率的计算过程，具体包括：基于音节覆盖率公式，根据音节表和所述音素集合中音节的数量，计算所述音素集合的音节覆盖率，从而为判断所输入的语音信息是否能够完整地体现用户的语音特征提供了重要参考。

其中，所述音节覆盖率公式为：

$$\gamma = \frac{P}{U}$$

其中， γ 为所述声母覆盖率， P 为音素集合中音节的数量， U 为音节表中的总数量。

比如，文本信息“张三喜欢跑步”的拼音为“zhang1san1xi3huan1pao3bu4”。该文本信息中出现了六个音节，分别为“zhang1”、“san1”、“xi3”、“huan1”、“pao3”、“bu4”。假设音节表中具有3990个互不相同的音节，则音节覆盖率 $=6/3990=0.1504\%$ 。

如图9所示，在一实施例中，步骤S140所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，具体包括子步骤S141a、S141b和

S141c。

S141a、判断所述声母覆盖率是否大于声母覆盖率阈值，所述韵母覆盖率是否大于韵母覆盖率阈值。

具体的，声母覆盖率阈值和韵母覆盖率阈值可以根据实际应用场景设计为任意适宜的数值，比如声母覆盖率阈值设计为 50%、韵母覆盖率阈值设计为 30%。

S141b、若所述声母覆盖率大于所述声母覆盖率阈值，且所述韵母覆盖率大于所述韵母覆盖率阈值，判定所述语音信息声纹验证通过。

示例性的，假设声母覆盖率阈值为 50%，韵母覆盖率阈值为 30%。经计算，声母覆盖率为 55%，韵母覆盖率为 32%，该声母覆盖率 55% 大于声母覆盖率阈值 50%，且韵母覆盖率 32% 大于韵母覆盖率阈值 30%，此时判定用户输入的语音信息声纹验证通过，生成声纹验证通过的验证结果。

S141c、若所述声母覆盖率不大于所述声母覆盖率阈值；或，所述韵母覆盖率不大于所述韵母覆盖率阈值，判定所述语音信息声纹验证未通过。

具体的，若声母覆盖率不大于声母覆盖率阈值以及所述韵母覆盖率不大于所述韵母覆盖率阈值至少有一个满足条件，判定上述语音信息声纹验证未通过，生成声纹验证未通过的验证结果。

示例性的，假设声母覆盖率阈值为 50%，韵母覆盖率阈值为 30%。经计算，声母覆盖率为 48%，韵母覆盖率为 32%，由于声母覆盖率 48% 小于声母覆盖率阈值 50%，因而判定用户输入的语音信息声纹验证未通过，生成声纹验证未通过的验证结果。

在另一实施例中，步骤 S140 所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，包括：根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

具体的，对所述语音信息进行声纹验证的过程，如图 10 所示，即步骤 S140 包括子步骤 S142a、S142b 和 S142c。

S142a、判断所述声母覆盖率是否大于声母覆盖率阈值，所述韵母覆盖率是否大于韵母覆盖率阈值，判断所述音节覆盖率是否大于音节覆盖率阈值。

具体的，声母覆盖率阈值、韵母覆盖率阈值和音节覆盖率阈值可以根据实际应用场景设计为任意适宜的数值，比如声母覆盖率阈值设计为 50%、韵母覆盖率阈值设计为 30%、音节覆盖率阈值设计为 0.100%。

S142b、若所述声母覆盖率大于所述声母覆盖率阈值，所述韵母覆盖率大于所述韵母覆盖率阈值，且所述音节覆盖率大于所述音节覆盖率阈值，判定所述语音信息声纹验证通过。

示例性的，假设声母覆盖率阈值为 50%，韵母覆盖率阈值为 30%，音节覆盖率阈值为 0.100%。经计算，声母覆盖率为 55%，韵母覆盖率为 32%，音节覆盖率为 0.152%，该声母覆盖率 55% 大于声母覆盖率阈值 50%，韵母覆盖率 32% 大于韵母覆盖率阈值 30%，且音节覆盖率为 0.152% 大于韵母覆盖率阈值 0.100%，此时判定用户输入的语音信息声纹验证通过，生成声纹验证通过的验证结果。

S142c、若所述声母覆盖率不大于所述声母覆盖率阈值；或，所述韵母覆盖率不大于所述韵母覆盖率阈值；或，所述音节覆盖率不大于音节覆盖率阈值，判定所述语音信息声纹验证未通过。

具体的，若声母覆盖率不大于声母覆盖率阈值、所述韵母覆盖率不大于所述韵母覆盖率阈值和所述音节覆盖率不大于音节覆盖率阈值至少有一个满足条件，判定上述语音信息声纹验证未通过，生成声纹验证未通过的验证结果。

示例性的，假设声母覆盖率阈值为 50%，韵母覆盖率阈值为 30%，音节覆盖率阈值为 0.100%。经计算，声母覆盖率为 48%，韵母覆盖率为 32%，音节覆盖率为 0.152%，由于该声母覆盖率 48% 小于声母覆盖率阈值 50%，因而判定用户输入的语音信息声纹验证未通过，生成声纹验证未通过的验证结果。

如图 11 所示，在一实施例中，步骤 S140 所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果之后，还包括：

S150、若所述验证结果表示所述语音信息声纹验证通过，接收并存储所述语音信息。

在声纹注册应用场景中，即在对所输入的语音信息进行接收与存储之前，先对该语音信息进行声纹验证，若验证结果表示所述语音信息声纹验证通过，再接收并存储该语音信息为注册语音信息，用于供语音识别验证。如此通过声纹验证的语音信息能够更加完整地体现出用户的语音特征，为后续的声纹注册语音识别提供了重要参考，提高了声纹注册的安全性。

当用户输入验证语音信息时，将验证语音信息和注册语音信息输入预先训练好的语音识别模型，以输出语音识别结果。

其中，预先训练好的语音识别模型可以是采用大量的语音-文本样本数据对初始神经网络进行训练获得。初始神经网络可以是各种神经网络，例如，卷积神经网络、循环神经网络、长短期记忆神经网络等。

S160、若所述验证结果表示所述语音信息声纹验证未通过，生成提示信息，以提示用户重新输入语音信息。

示例性的，该提示信息可以为“声纹验证失败，请重新输入语音信息”，用户看到该提示信息后，重新输入语音信息，直至声母覆盖率和韵母覆盖率符合要求为止。

上述声纹验证方法，通过将语音信息进行文本转化，以得到对应的文本信息；根据预设的音素模型，获取所述文本信息对应的音素集合；根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果，从而能够知悉该语音信息是否具有用户发音的声纹特征，是否能够涵盖该用户语音的大部分特征，进而找出具有涵盖用户大部分语音特征、用户语音特征完整度高的语音信息，为声纹身份验证提供了重要参考，确保该语音信息符合声纹注册的深层次需求。

请参阅图 12，图 12 是本申请的实施例还提供一种声纹验证装置的示意性框图，该声纹验证装置用于执行前述任一项声纹验证方法。其中，该声纹验证装置可以配置于服务器或终端中。

其中，服务器可以为独立的服务器，也可以为服务器集群。该终端可以是手机、平板电脑、笔记本电脑、台式电脑、个人数字助理和穿戴式设备等电子设备。

如图 12 所示，声纹验证装置 200 包括：文本转化单元 210、音素获取单元 220、覆盖率计算单元 230 和声纹验证单元 240。

文本转化单元 210，用于将语音信息进行文本转化，以得到对应的文本信息。

音素获取单元 220，用于根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母。

覆盖率计算单元 230，用于根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述

音素集合的韵母覆盖率。

声纹验证单元 240, 用于根据所述声母覆盖率和所述韵母覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果。

如图 13 所示, 在一实施例中, 音素获取单元 220 包括分词处理子单元 221、拼音转化子单元 222 和音素获取子单元 223。

分词处理子单元 221, 用于对所述文本信息进行分词处理, 以得到多个词串。

拼音转化子单元 222, 用于对各所述词串进行拼音转换, 以得到各所述词串对应的拼音串。

音素获取子单元 223, 用于将各所述拼音串输入预设的音素模型, 以得到音素集合。

如图 14 所示, 在一实施例中, 音素获取子单元 223 包括声母输出模块 223a、韵母输出模块 223c 和集合构造模块 223c。

声母输出模块 223a, 用于将所述拼音串输入所述声母子模型, 以输出所述词串中各字对应的声母。

韵母输出模块 223c, 用于将所述拼音串输入所述韵母子模型, 以输出所述词串中各字对应的韵母和各所述韵母对应的声调。

集合构造模块 223c, 用于根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调, 构建音素集合。

如图 15 所示, 覆盖率计算单元 230 包括数量计算子单元 231、声母计算子单元 232 和韵母计算子单元 233。

数量计算子单元 231, 用于计算所述音素集合中声母的数量和韵母的数量。

声母计算子单元 232, 用于基于声母覆盖率公式, 根据所述声母的数量和所述声母表, 计算所述声母覆盖率。

韵母计算子单元 233, 用于基于韵母覆盖率公式, 根据所述韵母的数量和所述韵母表, 计算所述韵母覆盖率。

如图 12 所示, 在一实施例中, 声纹验证装置 200 还包括音节计算单元 201, 用于根据音节表和所述音素集合中的各音节, 计算所述音素集合的音节覆盖率。

在该实施中, 声纹验证单元 240, 用于根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果。

如图 16 所示, 在一实施例中, 声纹验证单元 240 包括覆盖率判断子单元 241、第一判定子单元 242 和第二判定子单元 243。

覆盖率判断子单元 241, 用于判断所述声母覆盖率是否大于声母覆盖率阈值, 所述韵母覆盖率是否大于韵母覆盖率阈值;

第一判定子单元 242, 用于若所述声母覆盖率大于所述声母覆盖率阈值, 且所述韵母覆盖率大于所述韵母覆盖率阈值, 判定所述语音信息声纹验证通过;

第二判定子单元 243, 用于若所述声母覆盖率不大于所述声母覆盖率阈值; 或, 所述韵母覆盖率不大于所述韵母覆盖率阈值, 判定所述语音信息声纹验证未通过。

如图 12 所示, 在一实施例中, 声纹验证装置 200 还包括: 信息存储单元 250 和信息生成单元 260。

信息存储单元 250, 用于若所述验证结果表示所述语音信息声纹验证通过, 接收并存储所述语音信息;

信息生成单元 260, 用于若所述验证结果表示所述语音信息声纹验证未通过, 生成提示信息, 以提示用户重新输入语音信息。

需要说明的是, 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，上述描述的声纹验证装置和各单元的具体工作过程，可以参考前述声纹验证方法实施例中的对应过程，在此不再赘述。

上述的声纹验证装置可以实现为一种计算机程序的形式，该计算机程序可以在如图 17 所示的计算机设备上运行。

请参阅图 17，图 17 是本申请实施例提供的一种计算机设备的示意性框图。该计算机设备可以是服务器或终端。

参阅图 17，该计算机设备包括通过系统总线连接的处理器、存储器和网络接口，其中，存储器可以包括非易失性存储介质和内存存储器。

非易失性存储介质可存储操作系统和计算机程序。该计算机程序包括程序指令，该程序指令被执行时，可使得处理器执行一种声纹验证方法。

处理器用于提供计算和控制能力，支撑整个计算机设备的运行。

内存存储器为非易失性存储介质中的计算机程序的运行提供环境，该计算机程序被处理器执行时，可使得处理器执行一种声纹验证方法。

该网络接口用于进行网络通信，如发送分配的任务等。本领域技术人员可以理解，图 17 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

应当理解的是，处理器可以是中央处理单元（Central Processing Unit, CPU），该处理器还可以是其他通用处理器、数字信号处理器（Digital Signal Processor, DSP）、专用集成电路（Application Specific Integrated Circuit, ASIC）、现场可编程门阵列（Field-Programmable Gate 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。其中，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。

其中，所述处理器用于运行存储在存储器中的计算机程序，以实现如下步骤：

将语音信息进行文本转化，以得到对应的文本信息；根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

本申请的实施例中还提供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，所述计算机程序中包括程序指令，所述处理器执行所述程序指令，实现本申请实施例提供的任一项声纹验证方法。

其中，所述计算机可读存储介质可以是前述实施例所述的计算机设备的内部存储单元，例如所述计算机设备的硬盘或内存。所述计算机可读存储介质也可以是所述计算机设备的外部存储设备，例如所述计算机设备上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。

以上所述，仅为本申请的具体实施方式，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。因此，本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1. 一种声纹验证方法，包括：

将语音信息进行文本转化，以得到对应的文本信息；

根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；

根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；

根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

2. 根据权利要求1所述的声纹验证方法，其中，所述根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母，包括：

对所述文本信息进行分词处理，以得到多个词串；

对各所述词串进行拼音转换，以得到各所述词串对应的拼音串；

将各所述拼音串输入预设的音素模型，以得到音素集合。

3. 根据权利要求2所述的声纹验证方法，其中，所述音素模型包括声母子模型和韵母子模型；所述将各所述拼音串输入预设的音素模型，以得到音素集合，包括：

将所述拼音串输入所述声母子模型，以输出所述词串中各字对应的声母；

将所述拼音串输入所述韵母子模型，以输出所述词串中各字对应的韵母和各所述韵母对应的声调；

根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合。

4. 根据权利要求1所述的声纹验证方法，其中，所述根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率，包括：

计算所述音素集合中声母的数量和韵母的数量；

基于声母覆盖率公式，根据所述声母的数量和所述声母表，计算所述声母覆盖率；

基于韵母覆盖率公式，根据所述韵母的数量和所述韵母表，计算所述韵母覆盖率。

5. 根据权利要求1所述的声纹验证方法，其中，所述音素集合还包括所述文本信息中每个字所对应的音节；所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果之前，还包括：

根据音节表和所述音素集合中的各音节，计算所述音素集合的音节覆盖率；

所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果，包括：

根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

6. 根据权利要求1所述的声纹验证方法，其中，所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，包括：

判断所述声母覆盖率是否大于声母覆盖率阈值，所述韵母覆盖率是否大于韵母覆盖率阈值；

若所述声母覆盖率大于所述声母覆盖率阈值，且所述韵母覆盖率大于所述韵母覆盖率阈值，判定所述语音信息声纹验证通过；

若所述声母覆盖率不大于所述声母覆盖率阈值；或，所述韵母覆盖率不大于所述韵母覆盖率阈值，判定所述语音信息声纹验证未通过。

7. 根据权利要求1所述的声纹验证方法，其中，所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果之后，还包括：

若所述验证结果表示所述语音信息声纹验证通过，接收并存储所述语音信息；

若所述验证结果表示所述语音信息声纹验证未通过，生成提示信息，以提示用户重新输入语音信息。

8. 一种声纹验证装置，包括：

文本转化单元，用于将语音信息进行文本转化，以得到对应的文本信息；

音素获取单元，用于根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；

覆盖率计算单元，用于根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；

声纹验证单元，用于根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

9. 一种计算机设备，所述计算机设备包括存储器和处理器；

所述存储器用于存储计算机程序；

所述处理器，用于执行所述计算机程序并在执行所述计算机程序时，实现如下步骤：

将语音信息进行文本转化，以得到对应的文本信息；

根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母；

根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率；

根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

10. 根据权利要求9所述的计算机设备，其中，所述处理器在实现所述根据预设的音素模型，获取所述文本信息对应的音素集合，所述音素集合包括所述文本信息中每个字所对应的声母和韵母时，具体实现：

对所述文本信息进行分词处理，以得到多个词串；

对各所述词串进行拼音转换，以得到各所述词串对应的拼音串；

将各所述拼音串输入预设的音素模型，以得到音素集合。

11. 根据权利要求10所述的计算机设备，其中，所述音素模型包括声母子模型和韵母子模型；所述处理器在实现所述将各所述拼音串输入预设的音素模型，以得到音素集合时，具体实现：

将所述拼音串输入所述声母子模型，以输出所述词串中各字对应的声母；

将所述拼音串输入所述韵母子模型，以输出所述词串中各字对应的韵母和各所述韵母对应的声调；

根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合。

12. 根据权利要求9所述的计算机设备, 其中, 所述处理器在实现所述根据声母表和所述音素集合中的各声母, 计算所述音素集合的声母覆盖率; 根据韵母表和所述音素集合中的各韵母, 计算所述音素集合的韵母覆盖率时, 具体实现:

计算所述音素集合中声母的数量和韵母的数量;

基于声母覆盖率公式, 根据所述声母的数量和所述声母表, 计算所述声母覆盖率;

基于韵母覆盖率公式, 根据所述韵母的数量和所述韵母表, 计算所述韵母覆盖率。

13. 根据权利要求9所述的计算机设备, 其中, 所述音素集合还包括所述文本信息中每个字所对应的音节; 所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果之前, 还实现:

根据音节表和所述音素集合中的各音节, 计算所述音素集合的音节覆盖率;

所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果时, 具体实现:

根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果。

14. 根据权利要求9所述的计算机设备, 其中, 所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率, 对所述语音信息进行声纹验证时, 具体实现:

判断所述声母覆盖率是否大于声母覆盖率阈值, 所述韵母覆盖率是否大于韵母覆盖率阈值;

若所述声母覆盖率大于所述声母覆盖率阈值, 且所述韵母覆盖率大于所述韵母覆盖率阈值, 判定所述语音信息声纹验证通过;

若所述声母覆盖率不大于所述声母覆盖率阈值; 或, 所述韵母覆盖率不大于所述韵母覆盖率阈值, 判定所述语音信息声纹验证未通过。

15. 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使所述处理器实现如下步骤:

将语音信息进行文本转化, 以得到对应的文本信息;

根据预设的音素模型, 获取所述文本信息对应的音素集合, 所述音素集合包括所述文本信息中每个字所对应的声母和韵母;

根据声母表和所述音素集合中的各声母, 计算所述音素集合的声母覆盖率; 根据韵母表和所述音素集合中的各韵母, 计算所述音素集合的韵母覆盖率;

根据所述声母覆盖率和所述韵母覆盖率, 对所述语音信息进行声纹验证, 以生成验证结果。

16. 根据权利要求15所述的计算机可读存储介质, 其中, 所述处理器在实现所述根据预设的音素模型, 获取所述文本信息对应的音素集合, 所述音素集合包括所述文本信息中每个字所对应的声母和韵母时, 具体实现:

对所述文本信息进行分词处理, 以得到多个词串;

对各所述词串进行拼音转换, 以得到各所述词串对应的拼音串;

将各所述拼音串输入预设的音素模型, 以得到音素集合。

17. 根据权利要求16所述的计算机可读存储介质, 其中, 所述音素模型包括声母子模型和韵母子模型; 所述处理器在实现所述将各所述拼音串输入预设的音素模型, 以得到音素集合时, 具体实现:

将所述拼音串输入所述声母子模型，以输出所述词串中各字对应的声母；
将所述拼音串输入所述韵母子模型，以输出所述词串中各字对应的韵母和各所述韵母对应的声调；

根据所述词串中各字对应的声母、韵母以及各所述韵母对应的声调，构建音素集合。

18. 根据权利要求 15 所述的计算机可读存储介质，其中，所述处理器在实现所述根据声母表和所述音素集合中的各声母，计算所述音素集合的声母覆盖率；根据韵母表和所述音素集合中的各韵母，计算所述音素集合的韵母覆盖率时，具体实现：

计算所述音素集合中声母的数量和韵母的数量；

基于声母覆盖率公式，根据所述声母的数量和所述声母表，计算所述声母覆盖率；

基于韵母覆盖率公式，根据所述韵母的数量和所述韵母表，计算所述韵母覆盖率。

19. 根据权利要求 15 所述的计算机可读存储介质，其中，所述音素集合还包括所述文本信息中每个字所对应的音节；所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果之前，还实现：

根据音节表和所述音素集合中的各音节，计算所述音素集合的音节覆盖率；

所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证，以生成验证结果时，具体实现：

根据所述声母覆盖率、所述韵母覆盖率和所述音节覆盖率，对所述语音信息进行声纹验证，以生成验证结果。

20. 根据权利要求 15 所述的计算机可读存储介质，其中，所述处理器在实现所述根据所述声母覆盖率和所述韵母覆盖率，对所述语音信息进行声纹验证时，具体实现：

判断所述声母覆盖率是否大于声母覆盖率阈值，所述韵母覆盖率是否大于韵母覆盖率阈值；

若所述声母覆盖率大于所述声母覆盖率阈值，且所述韵母覆盖率大于所述韵母覆盖率阈值，判定所述语音信息声纹验证通过；

若所述声母覆盖率不大于所述声母覆盖率阈值；或，所述韵母覆盖率不大于所述韵母覆盖率阈值，判定所述语音信息声纹验证未通过。

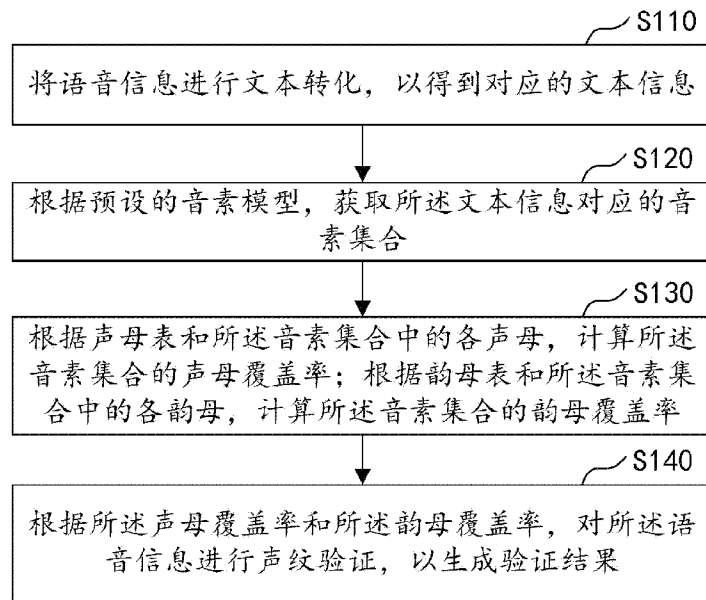


图 1

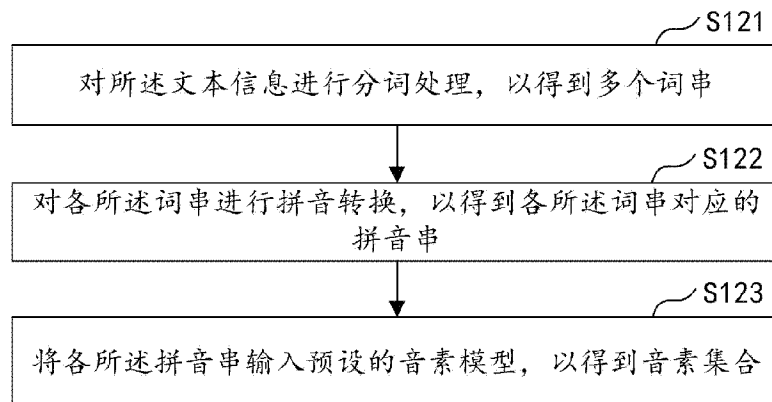


图 2

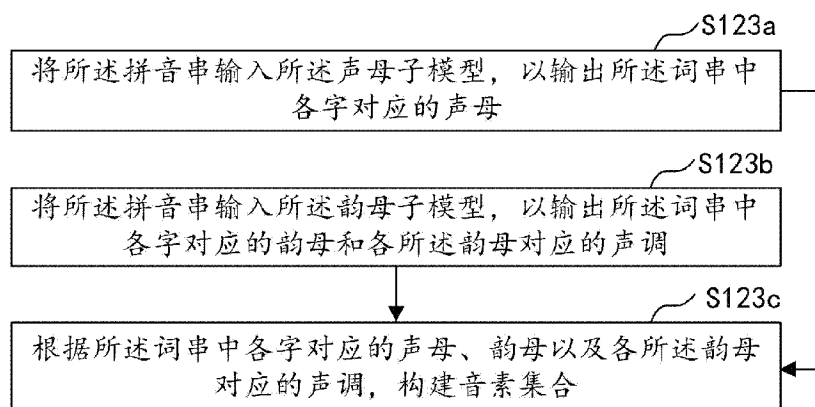


图 3

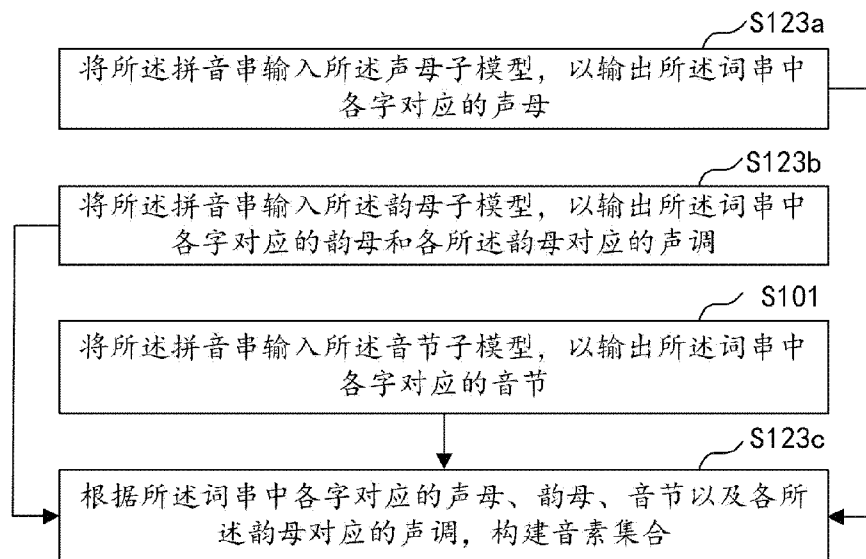


图 4

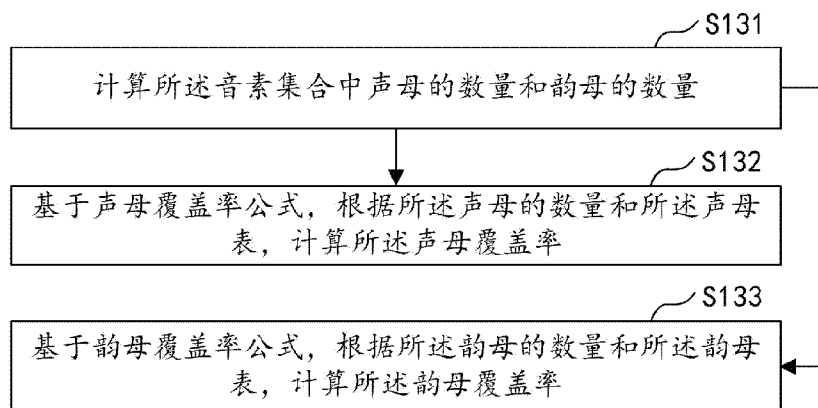


图 5

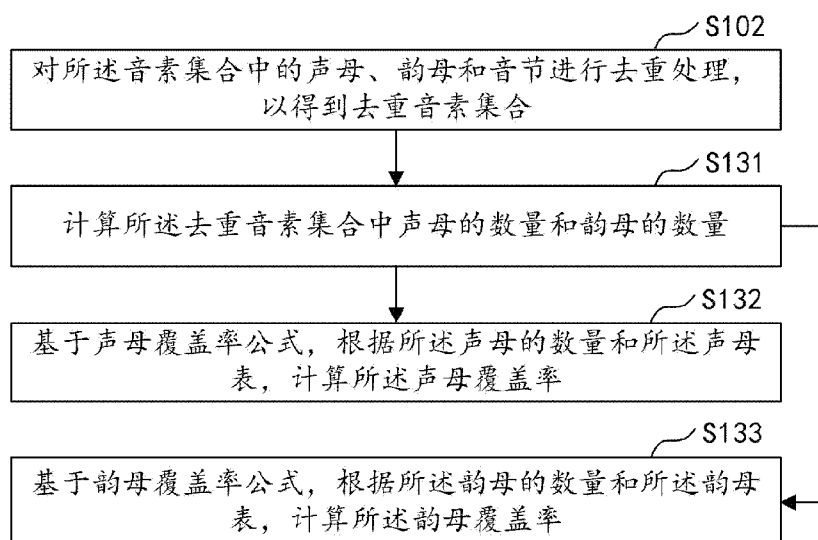


图 6

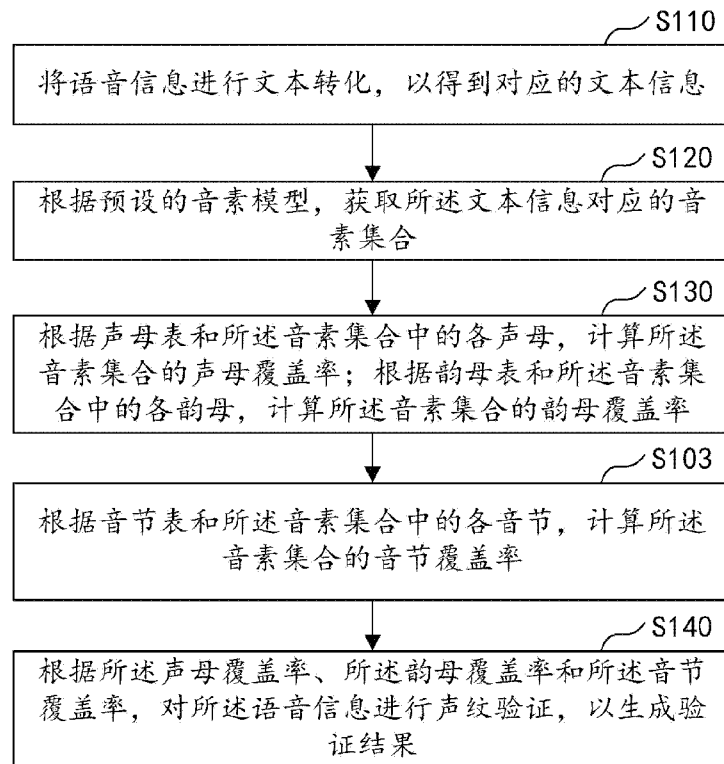


图 7

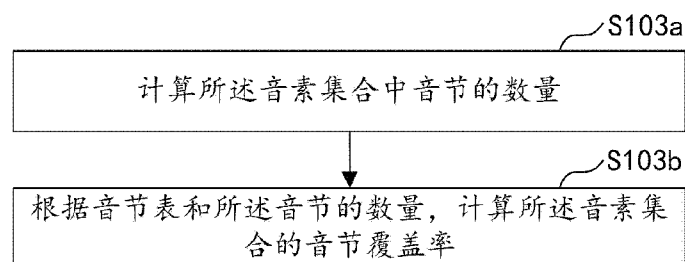


图 8

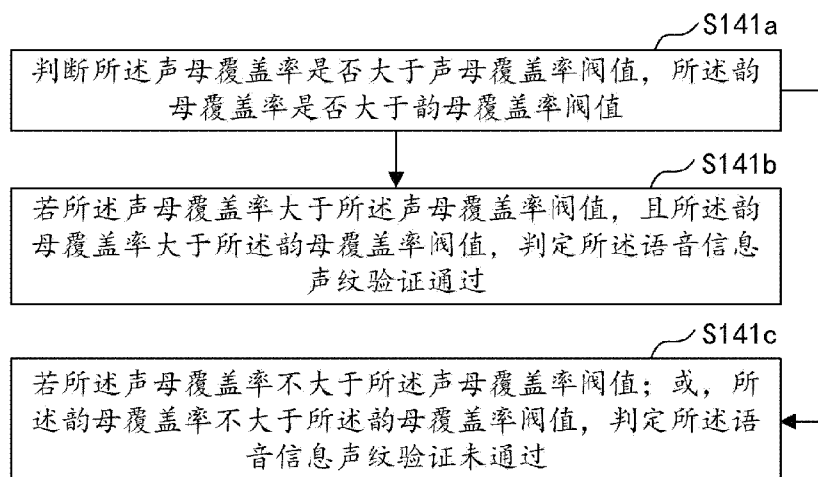


图 9

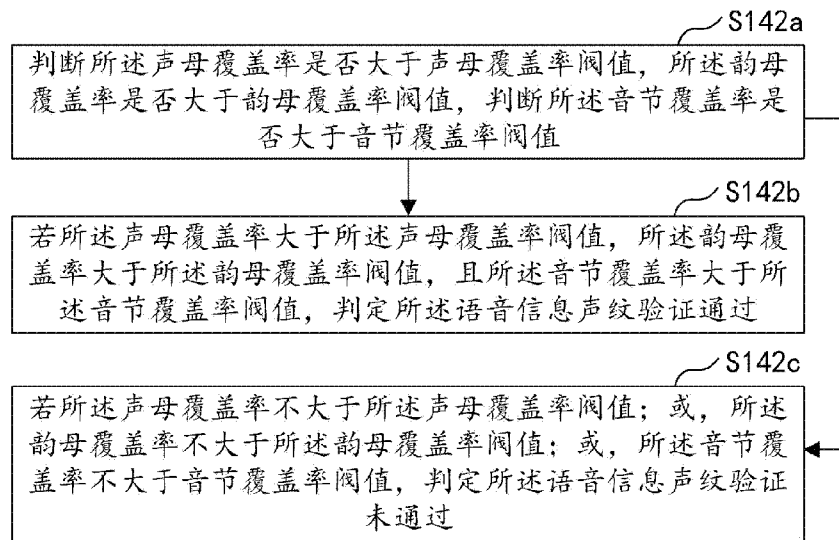


图 10

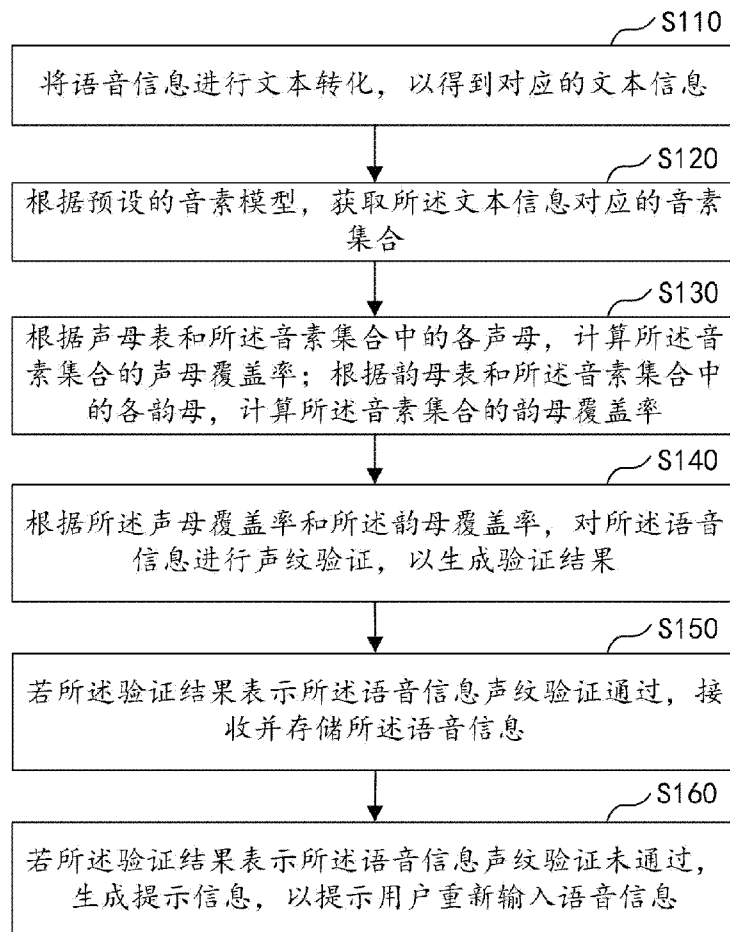


图 11

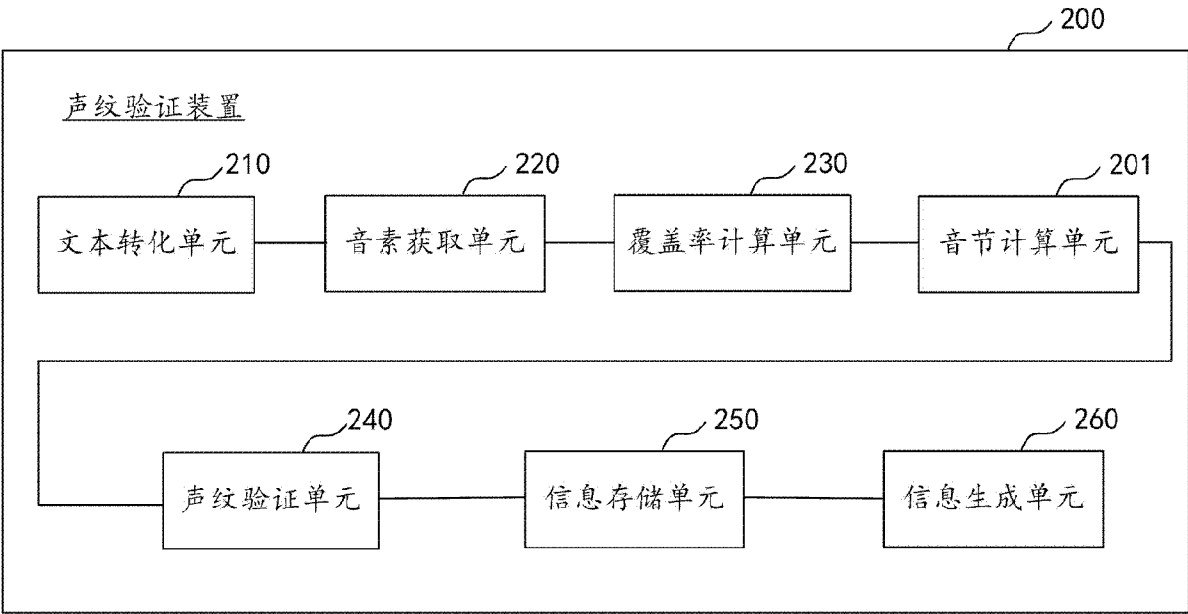


图 12

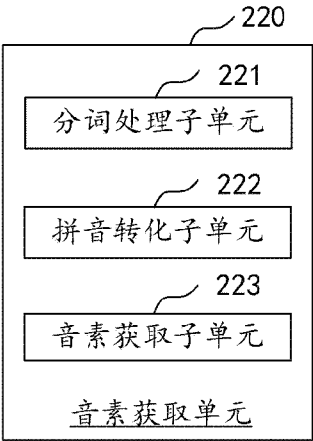


图 13

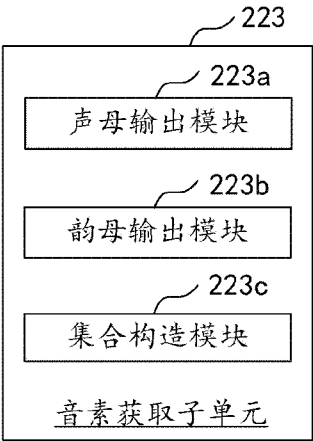


图 14

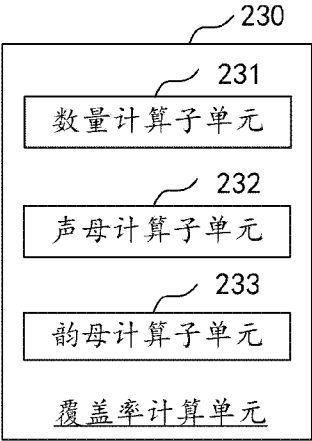


图 15

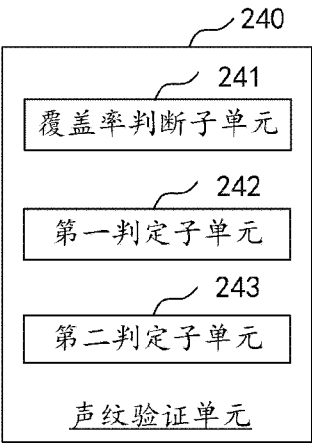


图 16

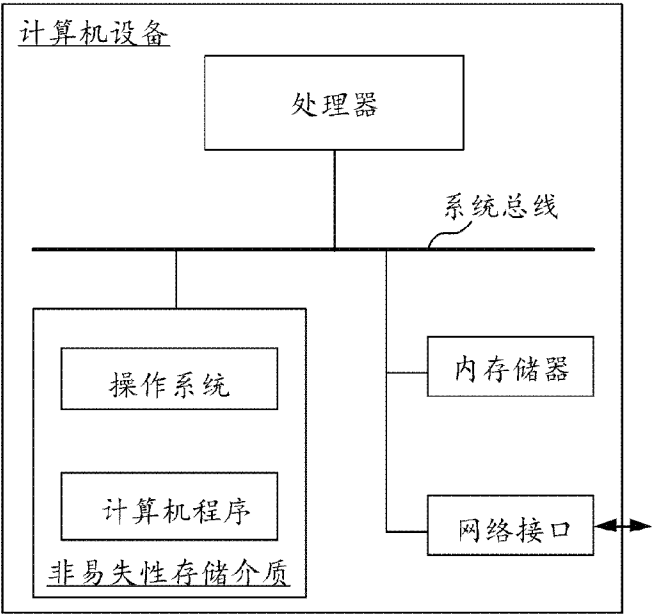


图 17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/103843

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/26(2006.01)i; G10L 17/00(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)

WPI, EPODOC, CNPAT, CNKI, IEEE: 声纹, 验证, 注册, 语音, 文本, 音素, 声母, 韵母, 覆盖, 特征, 识别, voiceprint, authenticat+, enrollment, speech, text, phoneme, initial, final, coverage, feature?, recognition

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	王琳琳 (WANG, Linlin). "说话人识别中的时变鲁棒性问题研究 (Research on time-varying robustness in speaker recognition)" 中国博士学位论文全文数据库 信息科技辑 (China Doctoral Dissertations Full-text Database, Information Science & Technology), No. 7, 15 July 2014 (2014-07-15), ISSN: 1674-022X, , pages 27-30	1-20
A	CN 108989341 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 11 December 2018 (2018-12-11) entire document	1-20
A	CN 109473108 A (SHENZHEN ONECONNECT INTELLIGENT TECHNOLOGY CO., LTD.) 15 March 2019 (2019-03-15) entire document	1-20
A	CN 107016994 A (ALIBABA GROUP HOLDING LIMITED) 04 August 2017 (2017-08-04) entire document	1-20
A	US 2015073796 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 12 March 2015 (2015-03-12) 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

21 February 2020

Date of mailing of the international search report

18 March 2020

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/103843

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108989341	A	11 December 2018	None			
CN	109473108	A	15 March 2019	None			
CN	107016994	A	04 August 2017	None			
US	2015073796	A1	12 March 2015	KR	20150030337	A	20 March 2015

国际检索报告

国际申请号

PCT/CN2019/103843

A. 主题的分类 G10L 15/26(2006.01)i; G10L 17/00(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)) WPI, EPDOC, CNPAT, CNKI, IEEE: 声纹, 验证, 注册, 语音, 文本, 音素, 声母, 韵母, 覆盖, 特征, 识别, voiceprint, authenticat+, enrollment, speech, text, phoneme, initial, final, coverage, feature?, recognition		
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
X	王琳琳. “说话人识别中的时变鲁棒性问题研究” 中国博士学位论文全文数据库 信息科技辑, 第7期, 2014年 7月 15日 (2014 - 07 - 15), ISSN: 1674-022X, , 第27-30页	1-20
A	CN 108989341 A (平安科技深圳有限公司) 2018年 12月 11日 (2018 - 12 - 11) 全文	1-20
A	CN 109473108 A (深圳壹账通智能科技有限公司) 2019年 3月 15日 (2019 - 03 - 15) 全文	1-20
A	CN 107016994 A (阿里巴巴集团控股有限公司) 2017年 8月 4日 (2017 - 08 - 04) 全文	1-20
A	US 2015073796 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 2015年 3月 12日 (2015 - 03 - 12) 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 2020年 2月 21日		国际检索报告邮寄日期 2020年 3月 18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 马鑫 电话号码 86-(10)-53961354

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/103843

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	108989341	A	2018年 12月 11日	无	
CN	109473108	A	2019年 3月 15日	无	
CN	107016994	A	2017年 8月 4日	无	
US	2015073796	A1	2015年 3月 12日	KR 20150030337	A 2015年 3月 20日