

(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
2016年2月4日 (04.02.2016)



(10) 国际公布号
WO 2016/015687 A1

- (51) 国际专利分类号:
H04L 9/32 (2006.01) G10L 17/00 (2013.01)
H04L 29/06 (2006.01)
- (21) 国际申请号: PCT/CN2015/090505
- (22) 国际申请日: 2015年9月24日 (24.09.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410373815.4 2014年7月31日 (31.07.2014) CN
201410389788.X 2014年8月8日 (08.08.2014) CN
201410388869.8 2014年8月8日 (08.08.2014) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。

- (72) 发明人: 孙骁 (SUN, Xiao); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 吴宗俤 (WU, Zongzhuo); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 龙强 (LONG, Qiang); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 李玲 (LI, Ling); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 岑丽芳 (CEN, Lifang); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 张翔 (ZHANG, Xiang); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 黄润佳 (HUANG, Runjia); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。 王晖 (WANG, Hui); 中国广东省深圳市福田区振兴路赛格科技园2栋东403室, Guangdong 518000 (CN)。

[见续页]

(54) Title: VOICEPRINT VERIFICATION METHOD AND DEVICE

(54) 发明名称: 声纹验证方法及装置

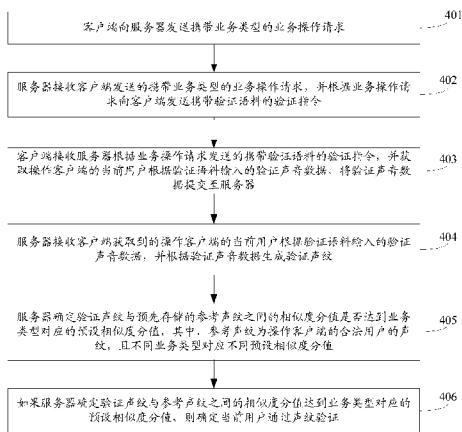
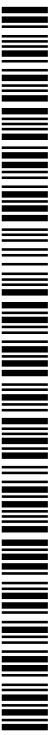


图4 / FIG. 4

401 CLIENT SENDS TO SERVER SERVICE OPERATION REQUEST CONTAINING SERVICE TYPE
 402 SERVER RECEIVES SERVICE OPERATION REQUEST CONTAINING SERVICE TYPE SENT BY CLIENT, AND ON BASIS OF SERVICE OPERATION REQUEST, SERVER SENDS TO CLIENT VERIFICATION COMMAND CONTAINING VERIFICATION PHRASE
 403 CLIENT RECEIVES VERIFICATION COMMAND CONTAINING VERIFICATION PHRASE SENT BY SERVER ON BASIS OF SERVICE OPERATION REQUEST, OBTAINS VERIFICATION VOICE DATA INPUTTED ON BASIS OF VERIFICATION PHRASE BY CURRENT USER OPERATING CLIENT, AND DELIVERS VERIFICATION VOICE DATA TO SERVER
 404 SERVER RECEIVES DATA INPUTTED ON BASIS OF VERIFICATION PHRASE BY CURRENT USER OPERATING CLIENT, AND ON BASIS OF VERIFICATION VOICE DATA SERVER GENERATES VERIFICATION VOICEPRINT
 405 SERVER DETERMINES WHETHER SIMILARITY VALUE BETWEEN VERIFICATION VOICEPRINT AND REFERENCE VOICEPRINT CONFORMS TO PRESET SIMILARITY VALUE CORRESPONDING TO SERVICE TYPE, WHEREIN REFERENCE VOICEPRINT IS VOICEPRINT OF USER AUTHORIZED TO OPERATE CLIENT AND DIFFERENT SERVICE TYPES CORRESPOND TO DIFFERENT PRESET SIMILARITY VALUES
 406 IF SERVER DETERMINES SIMILARITY VALUE BETWEEN VERIFICATION VOICEPRINT AND REFERENCE VOICEPRINT CONFORMS TO PRESET SIMILARITY VALUE CORRESPONDING TO SERVICE TYPE, THEN SERVER DETERMINES CURRENT USER PASSES VOICEPRINT VERIFICATION

(57) Abstract: Embodiments of the present invention provide a voiceprint verification method and device. The voiceprint verification method comprises receiving verification voice data relating to a verification phrase; generating a verification voiceprint on the basis of said verification voice data; determining whether a similarity value between the verification voiceprint and a reference voiceprint conforms to a preset similarity value; and, if the similarity value between the verification voiceprint and a reference voiceprint conforms to a preset similarity value, then determining there is a match between the verification voiceprint and the reference voiceprint.

(57) 摘要: 本发明的一些实施例提供了提供一种声纹验证方法和装置。声纹验证方法包括接收与验证语料相关联的验证声音数据; 根据所述验证声音数据生成验证声纹; 确定所述验证声纹与参考声纹之间的相似度分值是否达到预设相似度分值; 如果所述验证声纹与所述参考声纹之间的相似度分值达到所述预设相似度分值, 则确定所述验证声纹与所述参考声纹匹配。



WO 2016/015687 A1



(74) 代理人: 深圳市深佳知识产权代理事务所(普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市国贸大厦 15 楼西座 1521 室, Guangdong 518014 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布(细则 48.2(h))。
- 包括关于请求恢复一项或多项优先权要求的信息(细则 26 之二.3 和 48.2(b)(vii))。

声纹验证方法及装置

本申请要求 2014 年 7 月 31 日提交中国专利局、申请号为 201410373815.4、发明名称为“身份验证方法、服务器、客户端及系统”以及 2014 年 8 月 8 日提交中国专利局、申请号为 201410389788.X、发明名称为“声纹生成方法、服务器、客户端及系统”以及 2014 年 8 月 8 日提交中国专利局、申请号为 201410388869.8 发明名称为“声纹验证方法、服务器、客户端及系统”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

本公开涉及计算机技术领域，特别涉及一种声纹验证方法及装置。

10 背景技术

随着计算机技术的迅速发展，为了便于用户操作，安装在终端侧涉及用户信息的客户端越来越多，如各种社交应用程序的客户端、银行应用程序的客户端，等等。由于各个客户端中会包括很多与用户信息或用户财产相关的数据，因此，在终端上运行各个客户端时，需要对操作各个客户端的用户进行身份验证。相关技术在进行身份验证时，通常采用指纹识别或人脸识别的方法。其中，如果是指纹识别，则体征信息为指纹信息；如果是人脸识别，则体征信息为人脸图像信息。

采用指纹识别或人脸识别进行身份验证的方式，要求客户端所在的终端配置有能够进行指纹识别或人脸识别的设备，对终端的配置要求比较高。另外，针对于人脸识别来讲，由于人脸识别对环境要求比较高，一旦光线不佳或摄像头效果不佳，则可能无法进行身份验证或通过人脸识别进行身份验证的操作失败，不仅导致身份验证可靠性不高，而且身份验证的效率不高。

发明内容

为了解决现有技术的问题，本发明实施例提供了一种声纹验证方法及装置。

- 2 -

本发明一个实施例提供一种声纹验证方法，包括：接收与验证语料相关联的验证声音数据；根据所述验证声音数据生成验证声纹；确定所述验证声纹与参考声纹之间的相似度分值是否达到预设相似度分值；如果所述验证声纹与所述参考声纹之间的相似度分值达到所述预设相似度分值，则确定所述验证声纹与
5 与所述参考声纹匹配。

本发明一个实施例提供一种声纹验证装置，包括：存储器以及一个或多个处理器，其中所述存储器存储有程序指令，所述程序指令当由所述处理器执行时，配置所述装置执行上述纹验证方法。

附图说明

10 为了更清楚地说明本发明实施例中的技术方案，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，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 是本发明一个实施例提供的一种用于声纹验证的系统的示意图；

15 图 2 是本发明一个实施例提供的一种声纹验证方法的流程图；

图 3 (a) 是本发明一个实施例提供的一种客户端登录界面的流程图；

图 3 (b) 是本发明一个实施例提供的获取当前用户输入声音数据时的客户端界面的示意图；

图 4 是本发明另一示例性实施例提供的一种声纹验证方法的流程图；

20 图 5 是本发明一个实施例的引导用户设置声纹的客户端界面示意图；

图 6 是本发明一个实施例的显示第一提示信息的客户端界面示意图；

图 7 是本发明一个实施例的显示第二提示信息的客户端界面示意图；

图 8 是本发明一个实施例的显示第三提示信息的客户端界面示意图；

图 9 示出了本发明一个实施例的一种生成注册声纹的方法的流程图；

25 图 10 示出了本发明一个实施例的一种存储的注册声音数据及一段时间内每

次声纹验证通过后存储的验证声音数据及每个验证声音数据对应的属性信息的示意图；

图 11 示出了本发明一个实施例的一种筛选进化声音数据的示意图；

图 12 示出了本发明一个实施例的根据拼接的注册声音数据和进化声音数据生成进化声纹的过程示意图；

图 13 是本发明一个实施例的服务器的结构示意图；

图 14 是本发明一个实施例的一种终端的结构示意图。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。

如图 1 所示，其示出了本发明实施例提供的用于通过声纹进行验证（例如进行身份验证）的系统的示意图。如图 1 所示，该系统包括服务器 101 和终端 102。其中，服务器 101 和终端 102 通过无线网络或有线网络连接。

图 2 是根据本发明一个示例性实施例的一种声纹验证方法的流程图。该声纹验证例如可以用于身份验证。参见图 2，本发明该实施例提供的方法包括：

201：服务器向进行身份验证的当前用户对应的客户端发送验证语料。

本发明实施例提供的身份验证方法可以适用于任何在操作之前需要进行身份验证的场景。例如，登录场景、支付场景、进行身份验证的门禁系统，等等。本发明实施例中，客户端例如是位于终端 102 上的应用程序，可以为各种在操作时需要身份验证的客户端，如社交应用客户端、银行账户客户端、支付客户端，门禁验证客户端等等。

本发明实施例提供的身份验证方法通过服务器向进行身份验证的当前用户对应的客户端发送验证语料（例如预设文本信息）来触发对操作客户端的当前用户的身份进行验证。为了与下文提到的第二预设文本信息区分，这里的预设文本信息也称为第一预设文本信息。

- 4 -

例如，如果本发明实施例提供的方法用于对登录客户端的当前用户进行身份验证，则服务器向进行身份验证的当前用户对应的待登录的客户端发送第一预设文本信息是在客户端进入客户端的登录界面后执行的。图 3 (a) 示出了一种客户端的登录界面的示意图。

5 又例如，如果本发明实施例提供的方法用于对在支付客户端进行支付操作的当前用户进行身份验证，则服务器向进行身份验证的当前用户对应的客户端发送第一预设文本信息是在客户端进入客户端的支付界面后执行的。

关于第一预设文本信息的具体内容，本发明实施例不作具体限定。具体实施时，第一预设文本信息包括但不限于为：数字、诗词句、字母、汉字等等。

10 202: 客户端接收服务器发送的验证语料（例如第一预设文本信息），并获取操作客户端的当前用户根据验证语料输入的第一声音数据（即验证声音数据），将第一声音数据提交至服务器。

由于不同用户的声音具有不同的特性，且当根据某一用户的声音数据生成对应的声纹后，后续便可以通过该声纹判断某段声音数据的发声者与该声纹对
15 应的用户是否属于同一用户。因此，本发明实施例提供的方法支持根据操作客户端的当前用户的第一声纹（即验证声纹）与操作客户端的合法用户的第二声纹（即参考声纹）对操作客户端的当前用户进行身份验证。为了使服务器可以获得当前用户的第一声纹，客户端接收服务器发送的第一预设文本信息，并获取当前用户根据第一预设文本信息输入的第一声音数据。另外，为了使服务器
20 能够根据该第一声音数据生成当前用户的第一声纹，从而进一步根据第一声纹对当前用户进行身份验证，客户端将第一声音数据提交至服务器。

如图 3 (b) 所示，其示出了一种客户端获取当前用户根据第一预设文本信息输入的声音数据时的客户端界面示意图。其中，在图 3 (b) 中，第一预设文本信息为“123 453”。

25 203: 服务器接收客户端提交的第一声音数据，并根据第一声音数据生成第

一声纹（即验证声纹）。

关于服务器根据第一声音数据生成第一声纹的方式，可以采用声纹生成算法实现。声纹生成算法例如可以包括 SCHMM(半连续隐马尔科夫模型)、GMM-UBM（高斯混合模型-通用背景模型）等。关于服务器根据声纹生成算法将第一声音数据生成第一声纹的方式，本发明实施例不作具体限定。具体实施时，可以通过提取第一声音数据中的典型特征来实现。例如，由于声音的波长、频率、强度、节奏等特征能体现用户声音的特点，因此，在根据第一声音数据生成声纹时，可以提取第一声音数据中波长、频率、强度、节奏等特征，并确定第一声音数据中波长、频率、强度、节奏等特征的特征值，将第一声音数据中波长、频率、强度、节奏等特征的特征值等作为第一声纹。

204: 服务器确定第一声纹是否与操作客户端的合法用户的第二声纹（即参考声纹）匹配，其中，第二声纹可以是注册声纹也可以是进化声纹。注册声纹可以由合法用户根据注册语料（例如第二预设文本信息）输入的注册声音数据生成。根据本发明的实施例组成第一预设文本信息的字符可以为第二预设文本信息中的至少一个字符。

为了确定操作客户端的当前用户是否为操作客户端的合法用户，服务器需要确定第一声纹是否与操作客户端的合法用户的第二声纹匹配。其中，操作客户端的合法用户的第二声纹也可以包括很多内容，例如，可以包括合法用户的声音的波长、频率、强度、节奏等特征的特征值，等等。

关于服务器确定第一声纹是否与操作客户端的合法用户的第二声纹匹配的方式，包括但不限于：提取第一声纹中与第二声纹的特征一致的特征值，并将该特征值与第二声纹包括的具体内容中各个特征的特征值进行比对来实现。例如，如果第二声纹包括波长、频率，则可以提取第一声纹的波长、频率，并确定第一声纹的波长、频率的特征值，从而通过将第一声纹的波长、频率的特征值与第二声纹中的波长、频率的特征值进行比对来实现。

205: 如果服务器确定第一声纹与第二声纹匹配, 则确定当前用户通过身份验证。

具体地, 如果服务器确定第一声纹与第二声纹匹配, 则确定当前用户为合法用户, 因此, 确定当前用户通过身份验证。

- 5 可选地, 如果服务器确定第一声纹与第二声纹不匹配, 则服务器可以确定当前用户不是合法用户, 因此, 服务器可以确定当前用户未通过身份验证。进一步地, 如果当前用户身份验证未通过, 则服务器可以禁止当前用户对客户端进行操作, 或者让用户重新进行验证。

根据本发明的一个实施例在采用步骤 201 至步骤 205 对当前用户进行身份
10 验证之前, 还可以确定客户端所在的终端是否为预先认证的安全设备, 如果确定客户端所在的终端为预先认证的安全设备, 则直接通过步骤 201 至步骤 205 对当前用户进行身份验证。如果确定客户端所在的终端为非预先认证的安全设备, 则可以先通过如下方式对当前用户进行初步身份验证:

步骤一: 客户端获取操作客户端的当前用户输入的当前操作验证信息, 并
15 将该当前操作验证信息提交至服务器。

该当前操作验证信息可以为服务器下发的验证码, 也可以为操作该客户端的密码, 等等。例如, 如果正在对登录客户端的当前用户进行身份验证, 则该当前操作验证信息可以为登录该客户端的文本密码。

步骤二: 服务器接收客户端提交的当前操作验证信息, 并确定当前操作验
20 证信息是否正确。

为了初步确定操作客户端的用户是否为客户端的合法用户, 需要确定当前操作验证信息是否正确。关于确定当前操作是否正确的方式, 包括但不限于将当前验证信息与预先存储的客户端的操作验证信息进行比对来实现。例如, 针
25 对于对当前用户的登录操作进行验证的场景, 如果当前验证信息为登录客户端的文本密码, 则可以将登录该客户端的文本密码与预先存储的注册该客户端时

设定的文本密码进行比对来实现。

步骤三：如果服务器确定当前操作验证信息正确，则为了进一步对当前用户的身份进行验证，可以采用上述步骤 201 至步骤 205 继续进行身份验证。

5 可选地，如果服务器确定当前操作验证信息不正确，则可以直接停止当前用户的身份验证。此时，为了保证客户端运行的安全性，服务器可以禁止该终端运行该客户端。

本发明实施例提供的方法，通过服务器将操作客户端的当前用户的第一声纹与操作客户端的合法用户的第二声纹进行匹配来实现身份验证，不仅对客户端所在的终端的配置要求较低，从而适用范围较广，而且不受客户端所处环境的影响，身份验证更具可靠性的同时，能够提高身份验证的效率。

图 4 是根据本发明另一示例性实施例提供的一种声纹验证方法的流程图。参见图 4，本发明实施例提供的方法流程包括：

401：客户端向服务器发送携带业务类型的业务操作请求。

15 如果当前用户需要在客户端操作某一业务类型的业务，则需要通过客户端向服务器发送携带业务类型的业务操作请求。客户端可以在获取到当前用户的操作指令后，触发向服务器发送携带业务类型的业务操作请求。关于客户端获取到当前用户的操作指令的方式，具体实施时，包括但不限于通过检测操作选项或操作按钮是否被选中来实现。其中，检测操作选项或操作按钮是否被选中的方式可以有很多种，例如，可以通过检测操作选项或操作按钮是否被点击、20 是否被双击、操作选项或操作按钮的触摸操作是否达到预设时间，等等。

关于业务操作类型，本发明实施例不作具体限定。具体实施时，可以结合客户端支持的业务类型而定。例如，业务类型可以为登录业务、支付业务、小区门禁业务，等等。

25 例如，如果业务类型为登录业务，为了触发登录操作，当前用户需要通过对应的客户端向服务器发送携带登录业务的登录操作请求。其中，客户端向服

服务器发送携带登录业务的登录操作请求可以是在检测到客户端对应的登录选项被选中后触发的。在检测登录选项是否被选中时，包括但不限于通过检测登录选项是否被点击、是否被双击、登录选项上的触摸操作是否达到预设时间，等等。

5 例如，如果业务类型为小区门禁业务，则当前用户在进入小区之前进行身份验证时，可以通过声纹验证实现。为了触发声纹验证过程，当前用户可以通过客户端进入门禁验证页面，从而触发向服务器发送携带小区门禁业务的业务操作请求。关于当前用户进入门禁验证页面的方式，包括但不限于通过客户端与门禁管理系统建立连接来实现。其中，当前用户通过客户端与门禁管理系统
10 建立连接的方式，包括但不限于通过关注门禁管理系统的公众号实现。

402: 服务器接收客户端发送的携带业务类型的业务操作请求，并根据业务操作请求向客户端发送携带验证语料的验证指令。

为了触发通过声纹验证对当前用户的身份进行验证，服务器根据业务操作请求向客户端发送携带验证语料的验证指令，使当前用户可以根据验证语料输入
15 声音数据。

需要说明的是，服务器根据业务操作请求向客户端发送携带验证语料的验证指令之前，需要获取验证语料。关于获取到的验证语料的具体内容，本发明实施例不作具体限定。具体实施时，该验证语料的具体内容与形成参考声纹时使用的语料有关。后续在介绍参考声纹时，再对验证语料进行具体说明，此处
20 先暂不对验证语料的内容进行说明。

可选地，由于通常的声纹验证过程是由服务器发起的，然而，也可能存在人为伪造发起声纹验证的情况。为了保证声纹验证过程由服务器发起，避免人为伪造发起声纹验证过程，服务器在发起声纹验证过程时，可以生成一个电子票据，从而在与客户端每次进行交互的过程中，均携带该电子票据，以标识该
25 次声纹验证过程的基本信息，保证声纹验证过程的安全性。为了生成电子票据，

服务器在根据业务操作请求向客户端发送携带验证语料的验证指令之前，还需要根据业务操作请求获取当前用户的用户信息（为了与下文的第一用户信息区分，这里的用户信息称作第二用户信息），并根据第二用户信息生成电子票据。

5 关于第二用户信息的具体内容，具体实施时，为了标识声纹验证过程的基本信息，该第二用户信息至少包括但不限于业务类型、验证时刻、当前用户账号信息及当前用户的终端信息。根据第二用户信息生成电子票据的方式，具体实施时，可以包括但不限于通过建立电子票据标识与第二用户信息的映射关系实现。

10 关于获取第二用户信息的方式，包括但不限于通过接收客户端提交的第二用户信息实现。

综上，如果服务器在声纹验证过程中生成了电子票据，则服务器在根据业务操作请求向客户端发送携带验证语料的验证指令时，可以根据业务操作请求向客户端发送携带验证语料及生成的电子票据的验证指令。

15 403: 客户端接收服务器根据业务操作请求发送的携带验证语料的验证指令，并获取操作客户端的当前用户根据验证语料输入的验证声音数据，将验证声音数据提交至服务器。

通常，服务器在生成用户的声纹时，需要根据用户的声音数据生成，因此，客户端获取操作客户端的当前用户根据验证语料输入的验证声音数据，并将验证声音数据提交至服务器。

20 其中，本发明实施例不对客户端获取操作客户端的当前用户根据验证语料输入的验证声音数据的方式进行限定。具体实施时，在当前用户根据验证语料输入验证声音数据时，客户端通过录制当前用户输入的声音数据来实现。具体地，在录制验证声音数据时，可以先获取麦克风或话筒的启动指令，并根据麦克风或话筒的启动指令录制当前用户根据验证语料输入的声音数据。其中，获取
25 取麦克风或话筒的启动指令的方式可以有很多种，具体实施时，包括但不限于

通过检测麦克风或话筒被长按来实现。

可选地，客户端通过接收服务器根据业务操作请求发送的携带验证语料的验证指令后，可以显示验证指令中携带的验证语料，使得当前用户在输入验证声音数据时，可以根据显示的验证语料朗读验证语料，保证用户不会因忘记验证语料的内容而使得输入的验证声音数据不满足验证声音数据的条件，从而给用户带来良好操作体验的同时，能够提高获取当前用户的验证声音数据的效率，从而可以提高声纹验证的效率。

关于将验证声音数据提交至服务器的方式，可以有很多种。例如，客户端在录制当前用户的验证声音数据时，可以在录制验证声音数据的页面提供提交选项，并通过检测该提交选项是否被选中来确定是否将验证声音数据提交至服务器。当客户端检测到提交选项被选中后，触发将验证声音数据提交至服务器。

另外，客户端在向服务器提交验证声音数据时，由于声音数据通常比较大，因此，客户端可以将验证声音数据切分为若干份，并分别提交切分后的声音数据，直至将验证声音数据提交完毕。其中，在切分验证声音数据时，可以将验证声音数据切分为相等大小的声音数据，也可以切分为不等大小的声音数据。本发明实施例不对切分后的声音数据的具体大小进行限定。

例如，可以将验证声音数据切分为相等大小的三等份，得到验证声音数据 1、验证声音数据 2 和验证声音数据 3，因此，客户端在向服务器提交验证声音数据时，将分别提交验证声音数据 1、验证声音数据 2 和验证声音数据 3。

可选地，如果服务器向客户端发送携带验证语料的验证指令时，还发送了电子票据，则为了使服务器能够获得第二用户信息，客户端可以向服务器提交第二用户信息。关于客户端向服务器提交第二用户信息的方式，本发明实施例不作具体限定。另外，当生成电子票据后，服务器与客户端在每次进行交互时，均需要携带该电子票据。因此，客户端在接收服务器发送的携带验证语料的验证指令时，还接收了验证指令中携带的服务器生成的电子票据。此时，客户端

在将验证声音数据提交至服务器时同时还将服务器生成的电子票据提交至服务器。

404: 服务器接收客户端获取到的操作客户端的当前用户根据验证语料输入的验证声音数据, 并根据验证声音数据生成验证声纹。

5 为了获得当前用户的声纹, 服务器接收客户端获取到的操作客户端的当前用户根据验证语料输入的验证声音数据, 并根据验证声音数据生成验证声纹。其中, 本发明实施例不对服务器接收客户端获取到的操作客户端的当前用户根据验证语料输入的验证声音数据的方式进行限定。另外, 关于服务器根据验证声音数据生成验证声纹的方式, 可以采用声纹生成算法实现。关于服务器根据声纹生成算法将验证声音数据生成验证声纹的方式, 本发明实施例不作具体限定。具体实施时, 可以通过提取验证声音数据中的典型特征来实现。具体地, 10 由于声音的波长、频率、强度、节奏等特征能体现用户声音的特点, 因此, 在根据验证声音数据生成验证声纹时, 可以提取验证声音数据中波长、频率、强度、节奏等特征, 并确定验证声音数据中波长、频率、强度、节奏等特征的特15 征值, 将验证声音数据中波长、频率、强度、节奏等特征的特征值等作为验证声纹。

另外, 如果客户端在提交验证声音数据时, 将验证声音数据切分为若干个声音数据, 则服务器在生成验证声纹时, 需要先拼接组成验证声音数据的每个声音数据, 并将拼接后的组成验证声音数据的每个声音数据生成验证声纹。

20 可选地, 如果客户端提交验证声音数据时, 还提交了服务器生成的电子票据, 则为了保证声纹验证过程的安全性, 服务器根据验证声音数据生成验证声纹之前, 还需要判断客户端提交的电子票据与服务器生成的电子票据是否一致; 如果客户端提交的电子票据与服务器生成的电子票据一致, 则执行根据验证声音数据生成验证声纹的步骤。如果服务器确定客户端提交的电子票据与服务器25 生成的电子票据不一致, 则确定当前用户为非法用户, 并终止声纹验证, 从而

禁止当前用户在该客户端操作业务操作请求中携带的业务类型对应的业务。

关于服务器判断客户端提交的电子票据与服务器生成的电子票据是否一致的方式，包括但不限于判断客户端提交的电子票据相对于服务器生成的电子票据是否有改变；如果服务器判断客户端提交的电子票据相对于服务器生成的电子票据未发生改变，则确定客户端提交的电子票据与服务器生成的电子票据一致。

关于终止声纹验证的方法，可以有很多种。具体实施时，包括但不限于通过强制跳出客户端所在的业务类型对应的页面或者发送提示消息以提示当前用户操作失败，等等。例如，如果该次声纹验证用于对当前用户的登录业务进行验证，则可以禁止当前用户登录该客户端。

然而，为了进一步确定当前用户是否为合法用户，还可以继续对当前用户的身份进行验证。因此，服务器也可以采用其它方式对当前用户进行身份验证。其中，服务器在采用其它方式对当前用户进行身份验证时，可以通过弹出其它方式的验证界面，并获取当前用户在该验证界面输入的信息来对当前用户进行身份验证。关于根据操作当前用户在该验证界面输入的信息来对当前用户进行身份验证的具体实现过程，本发明实施例不作具体限定。

可选地，为了保证声纹验证过程在一定的时间范围内有效，服务器还可以设置生成的电子票据的有效时间，从而将在生成的电子票据的有效时间内的完成的声纹验证确定为合法，而如果声纹验证过程超过生成的电子票据的有效时间，则将该声纹验证过程确定为非法。因此，服务器在接收客户端提交的电子票据及验证声音数据之后，还需要确定客户端提交的电子票据是否在生成的电子票据的预设有效时间内；如果确定客户端提交的电子票据在生成的电子票据的预设有效时间内，则执行判断客户端提交的电子票据与生成的电子票据是否一致的步骤。如果确定客户端提交的电子票据不在生成的电子票据的预设有效时间内，则确定当前用户为非法用户，并终止声纹验证。

关于确定预设有效时间的方式，本发明实施例不作具体限定。具体实施时，可以根据经验统计完成声纹验证的基本时间，并将该基本时间作为预设有效时间。当然，还可以采用其它方式确定预设有效时间。其中，本发明实施例不对预设有效时间的具体大小进行限定。

5 关于终止声纹验证的方式，可以有很多种，具体实施时，与上述如果服务器确定客户端提交的电子票据与服务器生成的电子票据不一致时终止声纹验证的原理一致，具体可参见上述内容，在此不再赘述。

405: 服务器确定验证声纹与预先存储的参考声纹之间的相似度分值是否达到业务类型对应的预设相似度分值，其中，参考声纹为操作客户端的合法用户
10 的声纹，且不同业务类型对应不同预设相似度分值。

为了确定当前用户能否在客户端操作业务类型对应的业务，服务器需要确定验证声纹与预先存储的操作客户端的合法用户的参考声纹之间的相似度分值是否达到业务类型对应的预设相似度分值。

其中，在确定验证声纹与预先存储的参考声纹之间的相似度分值是否达到
15 业务类型对应的预设相似度分值之前，需要先确定业务操作请求中携带的业务类型对应的预设相似度分值。关于确定业务操作请求中携带的业务类型对应的预设相似度分值的方式，包括但不限于：获取客户端支持的每个业务类型及预设相似度分值之间的映射关系，并根据该映射关系确定业务操作请求中携带的业务类型对应的预设相似度分值。

20 关于获取客户端支持的每个业务类型及预设相似度分值之间的映射关系的方式，本发明实施例不作具体限定。具体实施时，可以从本地获取，也可以从其它节点获取。关于根据映射关系确定业务操作请求中携带的业务类型对应的预设相似度分值的方式，包括但不限于将业务操作请求中携带的业务类型与映射关系中的每个业务类型一一比对，并将映射关系中
25 与业务操作请求中携带的业务类型一致的业务类型对应的预设相似度分值作为业务操作请求中携带的业

务类型对应的预设相似度分值。

如表一所示，其示出了一种获取客户端支持的所有业务类型及预先设定的每个业务类型对应的预设相似度分值的映射关系的表格示意图。

表一

业务类型	预设相似度分值
登录业务	80
...
支付业务	98

5

例如，如表一所示，如果业务操作请求中携带的业务类型为登录业务，则在确定登录业务对应的预设相似度分值时，可以将登录业务与表一包括的每个业务类型一一比对，从而确定登录业务对应的预设相似度分值为 80。

当然，表一中所列出的业务类型及每个业务类型对应的预设相似度分值仅仅用于举例，并不构成对本发明实施例的限制。另外，由于在本发明实施例中，合法用户的参考声纹可能是实时进化的，因此，业务类型对应的预设相似度分值也可能是实时更新的。因此，在获取客户端支持的每个业务类型及预设相似度分值之间的映射关系时，需要获取最新的每个业务类型及预设相似度分值之间的映射关系。

另外，为了确定验证声纹与预先存储的参考声纹之间的相似度分值是否达到业务类型对应的预设相似度分值，还需要确定验证声纹与参考声纹之间的相似度分值。关于确定验证声纹与参考声纹之间的相似度分值的方式，包括但不限于：获取参考声纹，并计算验证声纹与参考声纹之间的相似度分值。

关于获取参考声纹的方式，本发明实施例不作具体限定。需要说明的是，由于参考声纹可以是合法用户在注册声纹时获得的注册声纹，也可能是使用注册声纹进行声纹验证，且声纹验证通过后存储的验证声音数据共同对注册声纹

20

进化后获得的进化声纹。因此，需要确定作为参考声纹的是注册声纹还是进化声纹。其中，关于注册声纹和进化声纹的具体内容，将在后面进行详细描述。

关于计算验证声纹与参考声纹之间的相似度分值的方式，可以有很多种。具体实施时，可以结合声纹的特征所包括的维度、每个维度的特征值以及每个
5 维度所占的权重来实现。具体包括但不限于：确定验证声纹的每个维度的特征值及参考声纹每个维度的特征值，并确定每个维度所占的权重；根据验证声纹每个维度的特征值、参考声纹每个维度的特征值及确定每个维度所占的权重计算验证声纹与参考声纹之间的相似度分值。

例如，如果声纹包括的维度为波长、频率、强度，且在计算验证声纹与参
10 考声纹之间的相似度分值时波长、频率、强度所占的权重分别为 0.3、0.4 和 0.3，验证声纹的波长、频率和强度的特征值分别为 73、75 和 85，参考声纹的波长、频率和强度的特征值分别为 75、80 和 90，则验证声纹与参考声纹之间的相似度分值为： $100 - (0.3 * (75 - 73) + 0.4 * (80 - 75) + 0.3 * (90 - 85)) = 95.9$ 。也就是说，验证声纹与特征声纹的特征值越接近，验证声纹与参考声纹之间的相似度分值
15 越高。

关于服务器确定验证声纹与预先存储的参考声纹之间的相似度分值是否达到业务类型对应的预设相似度分值的方式，包括但不限于将计算得到的验证声纹与参考声纹之间的相似度分值与获取到的业务操作请求中携带的业务类型对应的预设相似度分值进行比较来实现。例如，如果计算得到的验证声纹与参考
20 声纹之间的相似度分值为 80，且获取到的业务操作请求中携带的业务类型对应的预设相似度分值为 75，则确定验证声纹与参考声纹之间的相似度分值达到业务类型对应的预设相似度分值。

其中，由于不同的业务类型与用户的用户数据或用户财产之间的相关性不同，有的业务类型直接关系用户的数据财产安全，有的业务类型对用户数据财产
25 安全的影响较小。因此，在本发明实施例中，为不同业务类型设置不同的声

纹验证等级，并设置不同业务类型对应不同预设相似度分值。该部分内容将在后续进行详细说明，此处暂不赘述。

5 可选地，如上所述，在本发明实施例中，参考声纹可以为注册声纹，也可以为进化声纹。当参考声纹为注册声纹时，获取参考声纹之前，还需要生成注册声纹。

406: 如果服务器确定验证声纹与参考声纹之间的相似度分值达到业务类型对应的预设相似度分值，则确定当前用户通过声纹验证。

10 其中，在本发明实施例中，可以为客户端支持的不同的业务类型设置不同的声纹验证等级，并设置不同的声纹验证等级对应不同的预设相似度分值。例如，可以为客户端支持的不同业务类型设置五个声纹验证等级，如低等、基本、中等、高等和极高等，并设置低等、基本、中等、高等和极高等对应的预设相似度分值为 50-60、60-70、70-80、80-90 及 90-100。此时，当验证声纹与参考声纹之间的相似度分值落在某一等级对应的预设相似度分值范围内时，确定验证声纹与参考声纹之间的相似度达到该等级对应的预设相似度分值，否则，确定
15 验证声纹与参考声纹之间的相似度未达到任一等级对应的预设相似度分值。其中，在该举例中仅以根据不同业务类型将声纹验证划分为五个等级，以及每个等级对应的预设相似度分值进行了举例说明，该举例中的各等级及其对应的预设相似度分值并不构成对本发明实施例的限制。

20 由于不同的业务类型与用户数据及用户财产的相关程度并不相同，因此，在根据不同业务类型设置不同的声纹验证等级之后，还可以根据客户端支持的业务类型为每个业务类型设置相应的声纹验证等级。例如，如果该客户端支持的业务类型包括登录业务、支付业务、小区门禁业务及声纹注册业务，且登录业务又分常用设备登录业务、新设备登录业务。其中，由于支付业务直接与用户的财产安全相关，因此，可以将支付业务的声纹验证等级设置为极高等；由
25 于常用设备登录业务是比较敏感的操作，通常直接与用户数据相关，因此，可

- 17 -

以将常用设备登录业务的声纹验证等级设置为高等；由于在新设备登录客户端时，往往会在声纹验证之前先获取登录客户端的密码，从而通过声纹验证对登录该客户端的当前用户进行二次身份验证，因此，可以将通过密码验证之后的新设备登录业务的声纹验证等级设置为基本等级；另外，在客户端实现对小区
5 门禁业务进行声纹验证时，由于小区门禁业务与用户的财产及数据的相关程度比较低，因此，可以将该业务的等级设置为低等；由于当注册声纹后，初次声纹验证的目的仅用于验证注册的声纹是否有效，因此，可以将初次声纹验证的等级设置为中等。

当然，还可以将声纹验证等级划分为其它等级，具体实施时，可以结合客
10 户端支持的所有业务类型而定。关于客户端支持的所有业务类型的具体内容，本发明实施例不作具体限定。

可选地，如果在进行声纹验证的过程中生成了电子票据，且在生成电子票
据时，服务器获取到第二用户信息，则服务器在确定当前用户通过声纹验证之
后，还可以记录第二用户信息，并获取已经存储的上次声纹验证通过后记录的
15 第三用户信息；确定第二用户信息是否与第三用户信息一致；如果确定第二用
户信息与第三用户信息一致，则将当前用户标记为可信用户。

如，如果获取到的已经存储的上次声纹验证通过后记录的第三用户信息为
用户 44525445 在终端 X, 终端 ID(Identity, 标识):ack5864dsghgd8984562156412,
通过了登录业务的声纹验证。此时，在该次声纹验证时，会检查是否为 44525445
20 这个用户，是否为设备 ack5864dsghgd8984562156412，当前操作业务是否为登
录业务，如果三者都符合，则确定该用户可信，从而将当前用户标记为可信用户。
当将该用户标记为可信用户后，后续如果再次检测到该用户通过该设备进行登
录业务，则可以使用稍微简单的验证方法对该用户进行身份验证。

可选地，确定当前用户通过声纹验证之后，为了对当前用户在该客户端操
25 作业务操作请求中携带的业务类型对应的业务的时间进行限制，还可以确定当

前用户在客户端操作业务类型对应的业务的时间是否达到预设时间，如果当前用户在客户端操作业务类型对应的业务的时间达到预设时间，则终止当前用户在客户端操作业务类型对应的业务。

其中，本发明实施例不对预设时间的具体数值进行限定，具体实施时，可以根据不同的业务类型设置不同的预设时间。例如，针对于支付操作，由于完成支付操作通常会很快，因此，可以将支付业务对应的预设时间设置的稍微小点，如可以设置为 2min (分钟)；又例如，针对于登录业务，通常用户登录客户端的时间间隔比较长，因此，可以将登录业务对应的预设时间设置的比较大，如 30min，等等。

关于确定当前用户在客户端操作业务类型对应的业务的时间是否达到预设时间的方式，本发明实施例不作具体限定。具体实施时，可以声纹验证通过的时刻开始计时，从而可以通过确定当前时刻与声纹验证通过的时刻之间的差值是否达到预设时间。如果当前时刻与声纹验证通过的时刻之间的差值达到预设时间，则确定当前用户在客户端操作业务类型对应的业务的时间达到预设时间；如果当前时刻与声纹验证通过的时刻之间的差值未达到预设时间，则确定当前用户在客户端操作业务类型对应的业务的时间未达到预设时间。

可选地，如果验证声纹与参考声纹未达到业务类型对应的预设相似度分值，则确定当前用户非法，并终止声纹验证。关于终止声纹验证的方法，可以有很多种。具体实施时，包括但不限于通过跳出客户端所在页面或者发送提示消息以提示当前用户操作失败，等等。例如，如果该次声纹验证用于对当前用户的登录业务进行声纹验证，则可以禁止当前用户登录该客户端。又例如，如果该次声纹验证用于对当前用户的支付业务进行声纹验证，则可以禁止当前用户在该客户端进行支付业务，等等。

然而，为了进一步确定当前用户是否为合法用户，还可以继续对当前用户的身份进行验证。此时，服务器也可以采用其它方式对当前用户进行身份验证。

其中，服务器在采用其它方式对当前用户进行身份验证时，可以通过弹出其它方式的验证界面，并获取当前用户在该验证界面输入的对应信息来对当前用户进行身份验证。关于根据操作当前用户在该验证界面输入的对应信息来对当前用户进行身份验证的具体实现过程，本发明实施例不作具体限定。

5 本发明实施例提供的方法，通过服务器为不同业务类型设置不同预设相似度分值，并在进行声纹验证时，客户端向服务器提交当前用户的验证声音数据后，服务器生成验证声纹，并通过确定验证声纹与参考声纹之间的相似度的分值是否达到业务类型对应的预设相似度分值来进行声纹验证，从而可以针对不同的业务类型确定是否声纹验证成功，声纹验证方法更具灵活性。

10 本发明的一个示例实施例公开了一种生成注册声纹的方法，该方法包括但不限于：获取合法用户根据注册语料（例如第二预设文本信息）输入的注册声音数据，并根据获取到的注册声音数据生成合法用户的注册声纹。该方法包括但不限于如下步一至步十三：

第一步：服务器检测客户端所在的终端是否为预先认证的安全设备。

15 在确定客户端所在终端为预先认证的安全设备后，在获取操作客户端的合法用户的注册声音数据时，可以在检测到该客户端在安全设备上成功登录之后，提供引导用户设置声纹的选项，并通过检测该选项被选中的操作来引导用户设置声纹。如图5所示，其示出了一种引导用户设置声纹的客户端界面示意图。

20 第二步：如果客户端所在的终端为预先认证的安全设备，则向客户端发送第二预设文本信息。

其中，本发明实施例不对第二预设文本信息的具体内容进行限定。具体实施时，第二预设文本信息包括但不限于为：数字、诗词句、字母、汉字等等。根据本发明一个实施例，组成第一预设文本信息的字符为第二预设文本信息中的至少一个字符。

25 第三步：客户端接收服务器发送的第二预设文本信息，并获取合法用户根

据第二预设文本信息输入的第二声音数据（即注册声音数据），将第二声音数据提交至服务器。

第四步：服务器接收客户端提交的第二声音数据，并确定第二声音数据是否满足预设条件。

5 预设条件例如可以为第二声音数据的持续时间是否达到预设时长，等等。

第五步：如果服务器确定第二声音数据满足预设条件，则将满足预设条件的第二声音数据作为操作客户端的合法用户的注册声音数据。

可选地，如果服务器确定第二声音数据不满足预设条件，则为了获得满足预设条件的第二声音数据，还可以执行下述步骤：

10 第六步：服务器向客户端发送重新录制声音数据的第一提示信息。

本发明实施例不对服务器向客户端发送重新录制声音数据的第一提示信息的方式进行限定。其中，第一提示信息包括第二预设文本信息，使得合法用户可以根据第一提示信息中的第二预设文本信息重新输入第二声音数据，进而使客户端可以获取到合法用户根据第一提示信息重新输入的第二声音数据。如图 6
15 所示，其示出了一种显示第一提示信息的终端界面示意图。

第七步：客户端接收服务器发送的第一提示信息，并获取合法用户根据第一提示信息重新输入的第二声音数据，将合法用户根据第一提示信息重新输入的第二声音数据提交至服务器。

20 本发明实施例不对客户端接收服务器发送的第一提示信息的方式进行限定。另外，客户端可以显示接收到的第一提示信息，使合法用户可以根据显示的第一提示信息重新根据第二预设文本信息输入第二声音数据。

第八步：服务器接收合法用户根据第一提示信息重新输入的第二声音数据，并确定合法用户根据第一提示信息重新输入的第二声音数据是否满足预设条件，如果不满足预设条件，则返回第六步，直至得到满足预设条件的第二声音
25 数据后，将满足预设条件的第二声音数据作为操作客户端的合法用户的注册声

音数据。

当获取到操作客户端的合法用户的注册声音数据时便可以根据获取到的操作客户端的合法用户的注册声音数据生成操作客户端的合法用户的注册声纹。

为了使生成的操作客户端的合法用户的声纹更加准确，可以多次获取操作客户端的合法用户的注册声音数据。因此，还可以包括如下步骤：

第九步：服务器向客户端发送再次获取声音数据的第二提示信息。

第二提示信息包括第二预设文本信息，使得合法用户可以根据第二提示信息再次输入声音数据，进而使客户端可以再次获取到合法用户根据第二预设文本信息输入的第三声音数据。如图 7 所示，其示出了一种显示第二提示信息的客户端界面示意图。

第十步：客户端接收服务器发送的第二提示信息，并获取合法用户根据第二提示信息再次输入的第三声音数据，将第三声音数据提交至服务器。

第十一步：服务器接收客户端提交的第三声音数据，并确定第三声音数据是否满足预设条件。

第十二步：如果服务器确定第三声音数据满足预设条件，则判断满足预设条件的第二声音数据和满足预设条件的第三声音数据是否一致。

为了确定两次获取到的合法用户的声音数据是否来自于同一用户，或者为了确定两次获取到的合法用户的声音数据是否相关，服务器需要判断满足预设条件的第二声音数据和满足预设条件的第三声音数据是否一致。关于服务器判断满足预设条件的第二声音数据和满足预设条件的第三声音数据是否一致的方式，包括但不限于通过判断满足预设条件的第二声音数据和满足预设条件的第三声音数据的相似度来实现。如果满足预设条件的第二声音数据和满足预设条件的第三声音数据的相似度达到预设阈值，则确定满足预设条件的第二声音数据和满足预设条件的第三声音数据一致。当然，还可以采用其它方式确定满足预设条件的第二声音数据和满足预设条件的第三声音数据是否一致，本发明实

施例对此不作具体限定。

第十三步：如果服务器确定满足预设条件的第二声音数据和满足预设条件的第三声音数据一致，则将满足预设条件的第二声音数据和满足预设条件的第三声音数据作为获取到的合法用户的注册声音数据，并根据满足预设条件的第二声音数据和满足预设条件的第三声音数据生成合法用户的注册声纹。

例如可以将满足预设条件的第二声音数据和满足预设条件的第三声音数据拼接起来，并将拼接得到的声音数据作为合法用户的注册声音数据，从而根据注册声音数据生成注册声纹。

可选地，如果服务器确定满足预设条件的第二声音数据和满足预设条件的第三声音数据不一致，则服务器可以要求客户端从头开始重新获取合法用户的注册声音数据。具体地，服务器可以向客户端发送重新获取第二声音数据的第四提示信息，在客户端获取到合法用户根据第四提示信息重新输入的声音数据，并在客户端向服务器重新发送合法用户根据第二预设文本信息输入的第二声音数据后，服务器返回上述第四步，并从上述第四步开始重新获取合法用户的注册声音数据直至获得满足预设条件的第二声音数据和满足预设条件的第三声音数据，且满足预设条件的第二声音数据与满足预设条件的第三声音数据一致为止。如图 8 所示，其示出了一种显示第四提示信息的客户端界面示意图。

本发明的另一个实施例提供另一种关于生成注册声纹的方式，如图 9 所示。该方法包括但不限于通过如下步骤实现：

步骤 901：服务器获取合法用户的第一用户信息，并获取与第一用户信息相关的注册语料，向客户端发送注册语料。

为了保证在生成注册声纹时，向合法用户展示的注册语料是合法用户感兴趣的内容，在生成注册时，服务器可以获取合法用户的第一用户信息，并获取与第一用户信息相关的注册语料，从而向客户端发送与其用户信息相关的注册语料。

其中，本发明实施例不对第一用户信息的具体内容进行限定。具体实施时，第一用户信息包括但不限于为合法用户的爱好信息、地域信息，等等。例如，如果合法用户爱好数学，则服务器下发的注册语料可以为数字。又例如，如果合法用户为英国人，则服务器下发的注册语料可以为英文字母、英文短语、英文句子，等等。

本发明实施例不对注册语料的具体内容进行限定，具体实施时，注册语料可以为诗词歌赋、随机数字、随机字母、随机词语组合、名词名句，等等。例如，注册语料可以为“0-9”这十个数字，也可以“白日依山尽，黄河入海流”，还可以为“A-H”等。当然，注册语料还可以为诗词歌赋、随机数字、随机字母、随机词语、名词名句中的至少两种的组合，等等。

关于服务器向客户端发送注册语料的方式，本发明实施例不作具体限定。服务器通过根据合法用户的第一用户信息向客户端发送与第一用户信息相关的注册语料，使得不同用户可以使用不同类型的注册语料，不仅选择注册语料的方式灵活多变，而且保证后续生成的注册声纹差异比较大，从而便于后续的声纹验证。

需要说明的是，注册语料的具体内容与后续将注册声纹作为参考声纹进行声纹验证时下发的验证语料的具体内容有关。具体地，验证语料为与注册语料中的全部字符一致的字符串、或者为与注册语料中的部分字符一致的字符串、或者为注册语料中的全部字符组成的倒序字符串、或者为注册语料中的部分字符组成的倒序字符串、或者为注册语料中的全部字符组成的乱序字符串、或者为注册语料中的部分字符组成的乱序字符串、或者为注册语料中的至少一个字符组成的重复字符串中的一种字符串。

例如，如果注册语料为“0-9”这十个数字，则验证语料可以为“0-9”这十个数字；也可以是“0-9”这十个数字中的任意几个数字，如“0-7”、“02468”；还可以是“0-9”这十个数字中的全部或部分数字的倒序字符串，如“9-0”、“7-0”、

“86420”；还可以是“0-9”这十个数字中的全部或部分数字组成的乱序字符串，如“3698521470”、“257369”；还可以是验证语料中的至少一个字符组成的重复字符串，如“000 111”、“115577”等。当然，该举例仅以注册语料为数字时，验证语料的可能形式进行了举例说明。具体实施时，当注册语料为字母、数字、
5 诗词等文本的组合形式时，验证语料的形式与当注册语料为数字时的形式类似，此处不再举例说明。

通过获取合法用户的第一用户信息，并获取与第一用户信息相关的注册语料来触发注册声纹的生成过程。由于注册语料为与合法用户的第一用户信息相关的语料，因此，能成功获取到合法用户根据与其第一用户信息相关的注册语料输入的注册声音数据，不仅能增加声纹生成的成功率，而且使得声纹生成方法更具灵活性。
10

步骤 902: 客户端接收服务器发送的注册语料，并分预设次数获取合法用户根据注册语料输入的注册声音数据，得到预设次数个注册声音数据后，将预设次数个注册声音数据提交至服务器。

15 关于客户端接收服务器发送的注册语料的原理，以及分预设次数获取合法用户根据注册语料输入的注册声音数据的原理，与上述步骤 403 中客户端接收服务器根据业务操作请求发送的携带验证语料的验证指令，并获取操作客户端的当前用户根据验证语料输入的验证声音数据的原理一致，具体可参见上述步骤 403 中的内容，在此不再赘述。

20 需要说明的是，为了使得生成的注册声纹更加准确，以便于后续根据生成的注册声纹进行声纹验证，客户端可以分预设次数获取合法用户根据注册语料输入的注册声音数据，从而得到合法用户的预设次数个注册声音数据。关于预设次数的具体数值，本发明实施例不作具体限定。具体实施时，可以根据需要设定。然而，为了使得生成的注册声纹更加准确，从而使得声纹验证的可靠性
25 更高，该预设次数可以为两次，也可以为三次，等等。

关于将预设次数个注册声音数据提交至服务器的方式，本发明实施不作具体限定。具体实施时，由于声音数据比较大，因此，在提交每次获取到的注册声音数据之前，可以对每次获取到的注册声音数据进行切分。因此，在将预设次数个注册声音数据提交至服务器时，可以分预设次数对合法用户每次根据注册语料输入的注册语音数据进行分片；将合法用户每次根据注册语料输入的注册语音数据分片提交至服务器，使得服务器分预设次数分片获取合法用户每次根据注册语料输入的注册语音数据，并组合合法用户每次根据注册语料输入的注册语音数据，得到预设次数个合法用户根据注册语料输入的注册声音数据。

其中，将合法用户每次根据注册语料输入的注册语音数据分片提交至服务器的原理同上述步骤 403 中客户端向服务器提交验证声音数据的原理一致，具体可参见上述步骤 403 中的内容，在此不再赘述。

可选地，为了确保能够在安静的环境下获取合法用户的注册声音数据，以避免周围噪音或其它声音的干扰，保证获取到的合法用户的语音数据质量比较高，客户端在获取合法用户根据注册语料输入的注册声音数据之前，还可以：
检测周围环境音量是否满足注册声音数据的获取条件；如果周围环境音量满足注册声音数据的获取条件，则获取合法用户根据注册语料输入的注册声音数据；如果周围环境音量不满足注册声音数据的获取条件，则可以向客户端发送提示信息，以提示用户当前不满足注册声音数据的获取条件，需要到安静环境中获取注册声音数据。关于向客户端发送提示信息的方式，本发明实施例不作具体限定。

关于客户端检测周围环境音量是否满足注册声音数据的获取条件的方式，包括但不限于：检测周围环境音量是否达到预设阈值，如果周围环境音量未达到该预设阈值，则确定周围环境音量满足注册声音数据的获取条件。其中，周围环境音量可以是环境噪音，也可以是多个其它用户说话的声音，等等。另外，本发明实施例不对周围环境音量的预设阈值的数值大小进行限定，具体实施时，

可以根据需要设定。

另外，为了保证获取到的注册声音数据确实为合法用户的声音数据，在分预设次数获取合法用户根据注册语料输入的注册声音数据之前，客户端还需要检测客户端所在的设备是否为预先认证的安全设备，且如果客户端所在的设备是否为预先认证的安全设备，则确定可以获取到合法用户的注册声音数据。

关于检测客户端所在的设备是否为预先认证的安全设备的方式，具体实施时，包括但不限于通过获取客户端所在的终端的标识，并将客户端所在的终端的标识与预先认证的安全设备的标识进行比对来实现。其中，获取到的客户端所在的终端的标识包括但不限于客户端所在的终端的IMEI (International Mobile Equipment Identification Number, 国际移动设备识别码) 等。另外，在获取客户端的合法用户的声音数据时，当检测到客户端所在的终端为预先认证的安全设备时，默认操作客户端的当前用户为操作客户端的合法用户。

进一步地，在检测客户端所在的终端是否为预先认证的安全设备之前，服务器需要预先认证安全设备。关于服务器预先认证安全设备的方式，具体实施时，包括但不限于通过确定该客户端在某一终端成功登录的次数是否达到预设次数来实现。如果确定该客户端在某一终端成功登录的次数达到预设次数，则将该终端认证为安全设备。其中，本发明实施例不对预设次数的具体数值进行限定，具体实施时，为了保证认证的安全设备的可靠性更高，该预设次数设置的越大越好。

步骤 903: 服务器分预设次数获取合法用户根据注册语料输入的注册声音数据，得到预设次数个注册声音数据。

关于服务器分预设次数获取合法用户根据注册语料输入的注册声音数据的方式，包括但不限于：如果客户端分预设次数获取合法用户根据注册语料输入的注册声音数据，并分预设次数提交获取到的注册声音数据，则服务器分预设次数接收客户端每次提交的注册声音数据，并将每次接收的注册声音数据作为

获取到的合法用户根据注册语料输入的注册声音数据，从而得到预设次数个注册声音数据。

可选地，为了确定获取到的注册声音数据是否能够作为生成注册声纹的可用声音数据，服务器在每次获取到注册声音数据后，还可以确定每次获取到的注册声音数据是否满足预设条件。其中，本发明实施例不对预设条件的具体内容进行限定，具体实施时，预设条件可以为注册声音数据的持续时间是否达到预设时长，等等。另外，需要说明的是，如果预设条件为注册声音数据的持续时间是否达到预设时长，则在设定该预设时长时，需要结合注册语料的具体内容而定。例如，如果根据统计经验确定普通用户读完注册语料的时长为 30s(秒)，
5 则该预设时长可以设置为 25s。此时，当确定注册声音数据的持续时间达到 25s，
10 则确定注册声音数据满足预设条件。

进一步地，如果获取到的预设次数个注册声音数据不满足预设条件，则可以向客户端下发重新获取合法用户根据注册语料输入的注册声音数据的提示信息，使合法用户根据该提示信息重新输入注册声音数据，直至服务器确定获取
15 到的预设次数个注册声音数据满足预设条件为止。

步骤 904: 服务器确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值。

在获取到预设次数个注册声音数据后，为了确定能否根据获取到的预设次数个注册声音数据生成注册声纹，服务器需要确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值。其中，本发明实施例不对声纹注册相似度分值的具体数值进行限定，具体实施时，可以根据经验设定。然而，
20 为了使得注册声纹可以作为后续声纹验证的可靠参考，该声纹注册相似度分值越高越好。

其中，服务器在确定预设次数个注册声音数据之间的相似度分值是否达到
25 声纹注册相似度分值之前，需要先计算预设次数个注册声音数据之间的相似度

分值。关于计算预设次数个注册声音数据之间的相似度分值的原理与上述步骤 405 中计算验证声纹与参考声纹之间的相似度分值的原理一致，具体可参见上述步骤 405 中的内容，此处不再赘述。

关于服务器确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值的方式，包括但不限于：在计算得到预设次数个注册声音数据之间的相似度分值之后，将预设次数个注册声音数据之间的相似度分值与预先设定的声纹注册相似度分值进行比对来实现。

需要说明的是，在确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值时，还可以结合预设次数的不同数值而有不同的方法。例如，如果预设次数为两次，则获服务器取到的注册声音数据为两个，此时，在确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值时，需要确定这两次获取到的注册声音数据之间的相似度分值是否达到声纹注册相似度分值来确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值。

又例如，如果预设次数超过两个，如为三个、四个，等等，则在确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值时，可以计算每两个注册声音数据之间的相似度分值，并在确定每两个注册声音数据之间的相似度分值均达到预设声纹注册相似度分值时确定预设次数个注册声音数据之间的相似度分值达到声纹注册相似度分值。

然而，当预设次数超过两次时，在确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值时，还可以根据每两个注册声音数据之间的相似度分值达到声纹注册相似度分值的数量确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值。如果每两个注册声音数据之间的相似度分值达到声纹注册相似度分值的数量超过预设阈值，则确定预设次数个注册声音数据之间的相似度分值是否达到声纹注册相似度分值。例如，如

果预设次数为三个，且确定当每两个注册声音数据之间的相似度分值达到声纹注册相似度分值的数量超过两个时，则确定该三个注册声音数据达到声纹注册相似度分值。

5 步骤 905: 如果服务器确定预设次数个注册声音数据之间的相似度分值达到声纹注册相似度分值，则根据预设次数个注册声音数据生成注册声纹，将注册声纹作为参考声纹。

关于服务器根据预设次数个注册声音数据生成注册声纹的方式，包括但不限于：拼接预设次数个注册声音数据，并根据拼接后的注册声音数据生成注册声纹。

10 其中，在根据拼接后的注册声音数据生成注册声纹时，也可以通过一定的声纹生成算法实现，其具体原理与上述步骤 404 中根据验证声音数据生成验证声纹的原理一致，具体可参见上述步骤 404 中内容，在此不再赘述。在生成注册声纹后，后续便可根据生成的注册声纹为参考进行声纹验证，因此，可以将注册声纹作为参考声纹。

15 可选地，在形成注册声纹之后，如果后续合法用户确定更换其注册声纹，则还可以通过上述步骤一至步骤五重新获取合法用户的注册声音数据，根据重新获取的合法用户的注册声音数据重新生成注册声纹。

需要说明的是，在重新生成注册声纹时，服务器向客户端发送的注册语料可以与之前注册声纹时使用的注册语料不同，从而可以丰富该用户的注册语料，20 使得在后续进行声纹验证时可以通过向客户端发送不同类型的注册语料，且通过给用户下发不同类型的注册语料而使用户在注册声纹的过程中获得良好的操作体验。

另外，在形成注册声纹之后，为了检测是否能进行声纹验证，客户端可以发起初次声纹验证过程。关于初次声纹验证过程的具体实现方式，本发明实施25 例不作具体限定。具体实施时，包括但不限于：在客户端页面提供尝试声纹验

证选项，并在检测到尝试声纹验证选项被选中后，服务器向客户端发送验证语料，从而通过本发明实施例提供的步骤 401 至步骤 406 进行初次声纹验证过程。

5 可选地，服务器根据拼接后的注册声音数据生成注册声纹之后，为了后续可以使用注册声音数据进化注册声纹，并不断扩充注册语料的内容，还可以存储注册语料、预设次数个注册声音数据及注册声纹，并形成注册语料、预设次数个注册声音数据及注册声纹的映射关系。其中，本发明实施例不对存储注册语料、预设次数个注册声音数据及注册声纹的方式进行限定，具体实施时，为了保证存储的注册语料、预设次数个注册声音数据及注册声纹的映射关系的安全性，可以对注册语料、预设次数个注册声音数据及注册声纹进行加密，并存储加密后的注册语料、预设次数个注册声音数据及注册声纹。
10

关于对注册语料、预设次数个注册声音数据及注册声纹进行加密的方式，本发明实施例不作具体限定。关于加密时使用的密码的形式，本发明实施例同样不作具体限定。具体实施时，该密码形式可以为数字密码，也可以为图案密码，等等。

15 如表二所示，其示出了一种存储的注册语料、预设次数个注册声音数据及注册声纹的映射关系的表格示意图。

表二

注册语料	预设次数个注册声音数据	注册声纹
0123456789	注册声音数据 1、注册声音数据 2	频率 85、波长 70

20 当然，上述表二中包括的内容仅用于举例，并不构成对本发明实施例的限制。

根据本发明的一个实施例，确定当前用户通过声纹验证之后，为了使得该次声纹验证过程使用的验证声音数据可以作为后续进化参考声纹时生成进化声

纹的声音数据，服务器还可以存储验证声音数据以及验证声音数据的属性信息，
以将验证声音数据以及验证声音数据的属性信息作为后续进化参考声纹时筛选
进化声音数据的参考。进一步地，随着声纹验证次数的不断增多，在每次声纹
验证通过之后都可以存储声纹验证过程使用的验证声音数据以及每个验证声音
5 数据的属性信息，以将每次声纹验证使用的验证声音数据以及验证声音数据的
属性信息作为后续进化参考声纹时筛选进化声音数据的参考。

其中，本发明实施例不对属性信息的具体内容进行限定。具体实施时，该
属性信息至少包括验证声音数据的时长以及根据该验证声音数据生成的验证声
纹与参考声纹之间的相似度分值。

10 如图 10 所示，其示出了一种存储的注册声音数据及一段时间内每次声纹验
证通过后存储的验证声音数据及每个验证声音数据对应的属性信息的示意图。
在图 10 中，注册声音数据 a 和注册声音数据 b 为生成注册声纹时根据注册语料
获取到的合法用户的声音数据，验证声音数据 1 至验证声音数据 n 为每次声纹
验证通过后存储的验证过程使用的验证声音数据。每个验证声音数据的分值为
15 每次进行声纹验证时验证声纹与参考声纹之间的相似度分值，每个验证声音数
据的长度为每个验证声音数据的时长。例如，验证声音数据 1 对应的“分数”
为通过验证声音数据 1 进行声纹验证时验证声音数据 1 与参考声纹之间的相似
度分值，验证声音数据 1 对应的“长度 3s (秒)”为验证声音数据 1 的时长。

根据本发明的一个实施例，由于本发明实施例提供的方法支持不断进化参
20 考声纹，因此，在存储验证声音数据以及验证声音数据的属性信息之后，还可
以确定是否需要进化参考声纹；如果确定需要进化参考声纹，则可以根据多次
声纹验证通过后存储的验证声音数据的属性信息筛选预设数值个进化声音数
据，并根据注册声音数据和进化声音数据生成进化注册声纹，将进化注册声纹
作为后续声纹验证的参考声纹。

25 其中，确定是否需要进化参考声纹的方式可以有很多种，具体实施时，包

包括但不限于通过如下两种方式来实现:

第一种方式: 由客户端根据需要选择进化参考声纹。具体实施时, 客户端根据需要选择进化参考声纹包括但不限于通过如下步骤来实现:

5 步骤一: 客户端确定是否需要进化参考声纹, 且如果确定需要进化参考声纹, 则向服务器发送进化参考声纹的声纹进化请求。

关于客户端确定是否需要进化参考声纹的方式, 包括但不限于通过确定是否获取合法用户的声纹进化指令来确定是否需要进化参考声纹。如果客户端确定获取到合法用户的声纹进化指令, 则确定需要进化参考声纹。其中, 本发明实施例不对获取合法用户的声纹进化指令的方式进行限定。具体实施时, 包括
10 但不限于在客户端提供声纹进化选项, 并通过检测该声纹进化选项是否被选中来确定是否获取到声纹进化指令。如果检测到声纹进化选项被选中, 则确定获取到声纹进化指令; 如果未检测到声纹进化选项被选中, 则确定未获取到声纹进化指令。关于检测声纹进化选项被选中的方式, 包括但不限于通过检测声纹进化选项是否被点击、长按、双击等实现。

15 关于客户端向服务器发送进化参考声纹的声纹进化请求的方式, 本发明实施例不作具体限定。

步骤二: 服务器确定是否接收到客户端发送的声纹进化请求, 且如果接收到客户端发送的声纹进化请求, 则确定需要进化参考声纹。

当客户端确定进化参考声纹而向服务器发送声纹进化请求触发进化参考声纹后, 服务器通过确定是否接收到客户端发送的声纹进化请求来确定是否进化
20 参考声纹。如果服务器确定接收到客户端发送的声纹进化请求, 则确定需要进化参考声纹。

第二种方式: 由服务器触发进化参考声纹。具体为: 服务器确定是否达到自动进化参考声纹的预设条件; 如果服务器确定达到自动进化参考声纹的预设
25 条件, 则确定需要进化参考声纹。

其中，本发明实施例不对自动进化参考声纹的预设条件的具体内容进行限定。具体实施时，服务器可以设置当满足预设时间间隔后自动进化参考声纹，也可以设置满足预设验证次数后自动进化参考声纹。因此，自动进化参考声纹的预设条件可以为满足自动进化参考声纹的预设时间间隔或者自动进化参考声纹的预设验证次数，等等。关于预设时间间隔及预设验证次数的具体数值，本
5 发明实施例不作具体限定。另外，也可以由客户端根据需要自动触发进化参考声纹，本发明实施例对此不作具体限定。

关于根据多次声纹验证通过后存储的验证声音数据的属性信息筛选预设数值个进化声音数据的方式，本发明实施例不作具体限定。具体实施时，由于验证声音数据的属性信息至少包括验证声音数据的时长及验证声纹与参考声纹的相似
10 度分值，因此，可以根据每次声纹验证通过后已经存储的验证声音数据的时长及验证声纹与参考声纹的相似度分值筛选预设数值个进化声音数据。

其中，在根据验证声音数据的时长筛选进化声音数据时，可以获取已经存储的所有验证声音数据以及每个验证声音数据的时长，计算所有验证声音数据的平均值，获得平均时长。在筛选进化声音数据时，可以过滤掉时长低于平均
15 时长的验证声音数据。在根据验证声纹与参考声纹的相似度分值时，可以设置预设阈值的相似度分值，从而在筛选进化声音数据时，将所有的验证声音数据中低于预设阈值的相似度分值的验证声音数据过滤掉。当通过验证声音数据的时长及验证声纹与参考声纹的相似度分值过滤掉一部分验证声音数据后，可以在
20 在剩余的验证声音数据中筛选同时满足验证声音数据的时长及验证声纹与参考声纹的相似度分值的验证声音数据，并将该类验证声音数据作为进化声音数据。需要说明的是，在该举例中在根据验证声音数据的时长筛选进化声音数据时，仅以按照所有验证声音数据的平均时长为例进行了说明。具体实施时，还可以采用其它时长筛选进化声音数据，本发明实施例不作具体限定。

25 当然，在根据多次声纹验证通过后存储的验证声音数据的属性信息筛选预

设数值个进化声音数据时，还可以仅根据属性信息中的一种筛选进化声音数据。例如，可以仅根据验证声音数据的时长筛选进化声音数据，等等。另外，在进化参考声纹时，服务器也可以不筛选进化声音数据，从而将每次声纹验证通过后存储的验证声音数据均作为进化声音数据。

5 如图 11 所示，其示出了一种筛选进化声音数据的示意图。在图 11 中，每次声纹验证通过后已经存储的验证声音数据包括验证声音数据 a、验证声音数据 b、验证声音数据 c 和验证声音数据 d。如果该次筛选声音数据采用的方法为过滤掉时长低于 2s 的验证声音数据，则通过按照每个验证声音数据的属性信息筛选出验证声音数据 a 和验证声音数据 d 作为进化声音数据。

10 关于根据注册声音数据和进化声音数据生成进化注册声纹的方式，包括但不限于拼接注册声音数据和进化声音数据，并根据拼接的注册声音数据和进化声音数据生成进化注册声纹。其中，拼接注册声音数据和进化声音数据的原理，以及根据拼接的注册声音数据和进化声音数据生成进化声纹的原理与上述生成注册声纹时拼接预设次数个注册声音数据的原理，以及根据拼接的预设次数个注册声音数据生成注册声纹的原理一致，具体可参见上述拼接预设次数个注册声音数据的原理，以及根据拼接的预设次数个注册声音数据生成注册声纹的原理，在此不再赘述。

15 如图 12 所示，其示出了一种根据拼接的注册声音数据和进化声音数据生成进化声纹的过程示意图。在图 12 中，注册声音数据为注册声音数据 a 和注册声音数据 b，筛选出的进化声音数据为验证声音数据 a 和验证声音数据 d。因此，在生成进化声纹时，根据拼接的注册声音数据 a、注册声音数据 b、验证声音数据 a 和验证声音数据 d 生成进化声纹。

需要说明的是，随着操作客户端的次数越来越多，进行声纹验证的次数也越来越多，因此，可以不断存储每次声纹验证通过使使用的验证声音数据，并在后续不断进化参考声纹，从而形成声纹验证、进化声音数据筛选以及声纹进

25

- 35 -

化的循环系统，确保进行声纹验证时使用的参考声纹更具可靠性和准确性，从而使得声纹验证过程具有更高的准确性和更好的识别率。

本发明一个实施例提供了一种计算机可读存储介质，该计算机可读存储介质可以是上述实施例中的存储器中所包含的计算机可读存储介质；也可以是单独存在，未装配入终端中的计算机可读存储介质。该计算机可读存储介质存储有一个或者一个以上程序，该一个或者一个以上程序被一个或者一个以上的处理器用来执行身份验证方法。

图 13 是根据一示例性实施例示出的服务器的结构示意图。参照图 13，服务器 1100 包括处理组件 1122，其进一步包括一个或多个处理器，以及由存储器 1132 所代表的存储器资源，用于存储可由处理组件 1122 的执行的指令，例如应用程序。存储器 1132 中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件 1122 被配置为执行指令，以执行上述任一实施例的方法中服务器所执行的功能。

服务器 1100 还可以包括一个电源组件 11211 被配置为执行服务器 1100 的电源管理，一个有线或无线网络接口 1150 被配置为将服务器 1100 连接到网络，和一个输入输出 (I/O) 接口 1158。服务器 1100 可以操作基于存储在存储器 1132 的操作系统，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 或类似。其中，一个或者一个以上程序存储于存储器中，且经配置以由一个或者一个以上处理器执行，所述一个或者一个以上程序包含用于进行本申请前面的实施例的方法。

请参考图 14，其示出了一种终端的结构示意图，该终端包括本发明实施例提供的客户端。终端 1300 可以包括 RF (Radio Frequency, 射频) 电路 110、包括有一个或一个以上计算机可读存储介质的存储器 120、输入单元 130、显示单元 140、传感器 150、音频电路 160、WiFi (Wireless Fidelity, 无线保真) 模块 170、包括有一个或者一个以上处理核心的处理器 180、以及电源 190 等部件。

- 36 -

RF 电路 110 可用于收发信息或通话过程中，信号的接收和发送。通常，RF 电路 110 包括但不限于天线、至少一个放大器、调谐器、一个或多个振荡器、用户身份模块（SIM）卡、收发信机、耦合器、LNA（Low Noise Amplifier，低噪声放大器）、双工器等。此外，RF 电路 110 还可以通过无线通信与网络和其他设备通信。

存储器 120 可用于存储软件程序以及模块，处理器 180 通过运行存储在存储器 120 的软件程序以及模块，从而执行各种功能应用以及数据处理。存储器 120 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据终端 1300 的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 120 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

输入单元 130 可用于接收输入的数字或字符信息，以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。具体地，输入单元 130 可包括触敏表面 131 以及其他输入设备 132。除了触敏表面 131，输入单元 130 还可以包括其他输入设备 132。

显示单元 140 可用于显示由用户输入的信息或提供给用户的信息以及终端 1300 的各种图形用户接口，这些图形用户接口可以由图形、文本、图标、视频和其任意组合来构成。显示单元 140 可包括显示面板 141 以及覆盖显示面板 141 的触敏表面 131。终端 1300 还可包括至少一种传感器 150，比如光传感器、运动传感器以及其他传感器。光传感器可包括环境光传感器及接近传感器。

音频电路 160、扬声器 161，传声器 162 可提供用户与终端 1300 之间的音频接口。音频电路 160 可将接收到的音频数据转换后的电信号，传输到扬声器 161，由扬声器 161 转换为声音信号输出；另一方面，传声器 162 将收集的声音信号转换为电信号，由音频电路 160 接收后转换为音频数据，再将音频数据输

出处理器 180 处理后, 经 RF 电路 110 以发送给比如另一终端, 或者将音频数据输出至存储器 120 以便进一步处理。

WiFi 属于短距离无线传输技术, 终端 1300 通过 WiFi 模块 170 可以帮助用户收发电子邮件、浏览网页和访问流式媒体等, 它为用户提供了无线的宽带互
5 联网访问。虽然图 14 示出了 WiFi 模块 170, 但是可以理解的是, 其并不属于终端 1300 的必须构成, 完全可以根据需要在不改变发明的本质的范围内而省略。

处理器 180 是终端 1300 的控制中心, 利用各种接口和线路连接整个手机的各个部分, 通过运行或执行存储在存储器 120 内的软件程序和/或模块, 以及调用存储在存储器 120 内的数据, 执行终端 1300 的各种功能和处理数据, 从而对
10 手机进行整体监控。可选的, 处理器 180 可包括一个或多个处理核心; 优选的, 处理器 180 可集成应用处理器和调制解调处理器, 其中, 应用处理器主要处理操作系统、用户界面和应用程序等, 调制解调处理器主要处理无线通信。可以理解的是, 上述调制解调处理器也可以不集成到处理器 180 中。

终端 1300 还包括给各个部件供电的电源 190 (比如电池), 优选的, 电源可
15 以通过电源管理系统与处理器 180 逻辑相连, 从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成, 也可以通过程序来指令相关的硬件完成, 所述的程序可以存储于一种计算机可读存储介质中, 存储介质可以是只读存储器, 磁盘或光盘等。

20 以上所述仅为本发明的较佳实施例, 并不用以限制本发明, 凡在本发明的精神和原则之内, 所作的任何修改、等同替换、改进等, 均应包含在本发明的保护范围之内。

权 利 要 求

- 1、一种声纹验证方法，包括：
由处理器接收与验证语料相关联的验证声音数据；
5 由处理器根据所述验证声音数据生成验证声纹；
由处理器确定所述验证声纹与预先存储的参考声纹之间的相似度分值是否达到预设相似度分值；
如果所述验证声纹与所述参考声纹之间的相似度分值达到所述预设相似度分值，则处理器确定所述验证声纹与所述参考声纹匹配。
- 10 2、根据权利要求1所述的方法，
其中，接收与验证语料相关联的验证声音数据包括：
接收客户端获取到的用户根据所述验证语料输入的验证声音数据，
- 3、根据权利要求2所述的方法，其中如果所述验证声纹与所述参考声纹匹配，则确定所述用户通过身份验证。
- 15 4、根据权利要求1所述的方法，还包括接收携带业务类型的业务操作请求，
其中，不同业务类型对应不同预设相似度分值，
其中，确定所述验证声纹与参考声纹之间的相似度分值是否达到预设相似度分值包括：
确定所述验证声纹与参考声纹之间的相似度分值是否达到所述业务类型对
20 应的预设相似度分值。
- 5、根据权利要求1所述的方法，还包括：
获取与注册语料相关联的注册声音数据；
根据所述注册声音数据生成所述参考声纹。
- 6、根据权利要求5所述的方法，还包括：
25 检测环境音量是否满足所述注册声音数据的获取条件。
- 7、根据权利要求5所述的方法，

其中所述验证语料包括与所述注册语料中的至少部分字符一致的字符串、或者为所述注册语料中的至少部分字符组成的倒序字符串、或者为所述注册语料中的至少部分字符组成的乱序字符串、或者为所述注册语料中的至少一个字符组成的重复字符串中的一种。

5 8、根据权利要求 5 所述的方法，还包括：

如果确定所述验证声纹与所述参考声纹匹配，则存储所述验证声音数据以及所述验证声音数据的属性信息；

根据所述属性信息选择所述存储的验证声音数据；

根据注册声音数据和选择的验证声音数据更新所述参考声纹。

10 9、根据权利要求 1 所述的方法，还包括：

获取用户信息；

获取与所述用户信息相关的注册语料；

分预设次数获取与所述注册语料相关联的注册声音数据；

确定所述注册声音数据之间的相似度分值是否达到声纹注册相似度分值；

15 如果所述注册声音数据之间的相似度分值达到所述声纹注册相似度分值，则根据所述注册声音数据生成所述参考声纹。

10、根据权利要求 1 所述的方法，还包括：

接收客户端发送的携带业务类型的业务操作请求，

获取用户信息，并根据所述用户信息生成电子票据；和

20 根据所述业务操作请求向所述客户端发送携带验证语料及生成的电子票据的验证指令，

其中，接收与验证语料相关联的验证声音数据，包括：

接收所述客户端提交的电子票据及所述客户端获取到的操作所述客户端的当前用户根据所述验证语料输入的验证声音数据，

25 其中，根据所述验证声音数据生成验证声纹包括：

- 40 -

判断所述客户端提交的电子票据与所述生成的电子票据是否一致；和
如果所述客户端提交的电子票据与所述生成的电子票据一致，则根据所述
验证声音数据生成验证声纹。

11、一种声纹验证装置，包括：

5 存储器；以及

一个或多个处理器，

其中所述存储器存储有程序指令，所述程序指令当由所述处理器执行时，
配置所述装置执行以下操作：

接收与验证语料相关联的验证声音数据；

10 根据所述验证声音数据生成验证声纹；

确定所述验证声纹与预先存储的参考声纹之间的相似度分值是否达到预设
相似度分值；

如果所述验证声纹与所述参考声纹之间的相似度分值达到所述预设相似度
分值，则确定所述验证声纹与所述参考声纹匹配。

15 12、根据权利要求 11 所述的装置，

其中，接收与验证语料相关联的验证声音数据包括：

接收客户端获取到的用户根据所述验证语料输入的验证声音数据，

13、根据权利要求 12 所述的装置，其中如果所述验证声纹与所述参考声纹
匹配，则确定所述用户通过身份验证。

20 14、根据权利要求 11 所述的装置，所述操作还包括接收携带业务类型的业
务操作请求，

其中，不同业务类型对应不同预设相似度分值，

其中，确定所述验证声纹与参考声纹之间的相似度分值是否达到预设相似度
分值包括：

25 确定所述验证声纹与参考声纹之间的相似度分值是否达到所述业务类型对

应的预设相似度分值。

15、根据权利要求 11 所述的装置，所述操作还包括：
获取与注册语料相关联的注册声音数据；
根据所述注册声音数据生成所述参考声纹。

5 16、根据权利要求 15 所述的装置，所述操作还包括：
检测环境音量是否满足所述注册声音数据的获取条件。

17、根据权利要求 15 所述的装置，
其中所述验证语料包括与所述注册语料中的至少部分字符一致的字符串、
或者为所述注册语料中的至少部分字符组成的倒序字符串、或者为所述注册语
10 料中的至少部分字符组成的乱序字符串、或者为所述注册语料中的至少一个字
符组成的重复字符串中的一种。

18、根据权利要求 15 所述的装置，所述操作还包括：

如果确定所述验证声纹与所述参考声纹匹配，则存储所述验证声音数据以
及所述验证声音数据的属性信息；

15 根据所述属性信息选择所述存储的验证声音数据；
根据注册声音数据和选择的验证声音数据更新所述参考声纹。

19、根据权利要求 11 所述的装置，所述操作还包括：

获取用户信息；

获取与所述用户信息相关的注册语料；

20 分预设次数获取与所述注册语料相关联的注册声音数据；

确定所述注册声音数据之间的相似度分值是否达到声纹注册相似度分值；

如果所述注册声音数据之间的相似度分值达到所述声纹注册相似度分值，
则根据所述注册声音数据生成所述参考声纹。

20、根据权利要求 11 所述的装置，所述操作还包括：

25 接收客户端发送的携带业务类型的业务操作请求，

- 42 -

获取用户信息，并根据所述用户信息生成电子票据；和
根据所述业务操作请求向所述客户端发送携带验证语料及生成的电子票据
的验证指令，

其中，接收与验证语料相关联的验证声音数据，包括：

5 接收所述客户端提交的电子票据及所述客户端获取到的操作所述客户端的
当前用户根据所述验证语料输入的验证声音数据，

其中，根据所述验证声音数据生成验证声纹包括：

判断所述客户端提交的电子票据与所述生成的电子票据是否一致；和

10 如果所述客户端提交的电子票据与所述生成的电子票据一致，则根据所述
验证声音数据生成验证声纹。

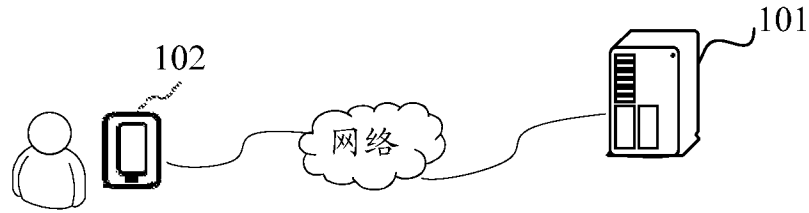


图 1

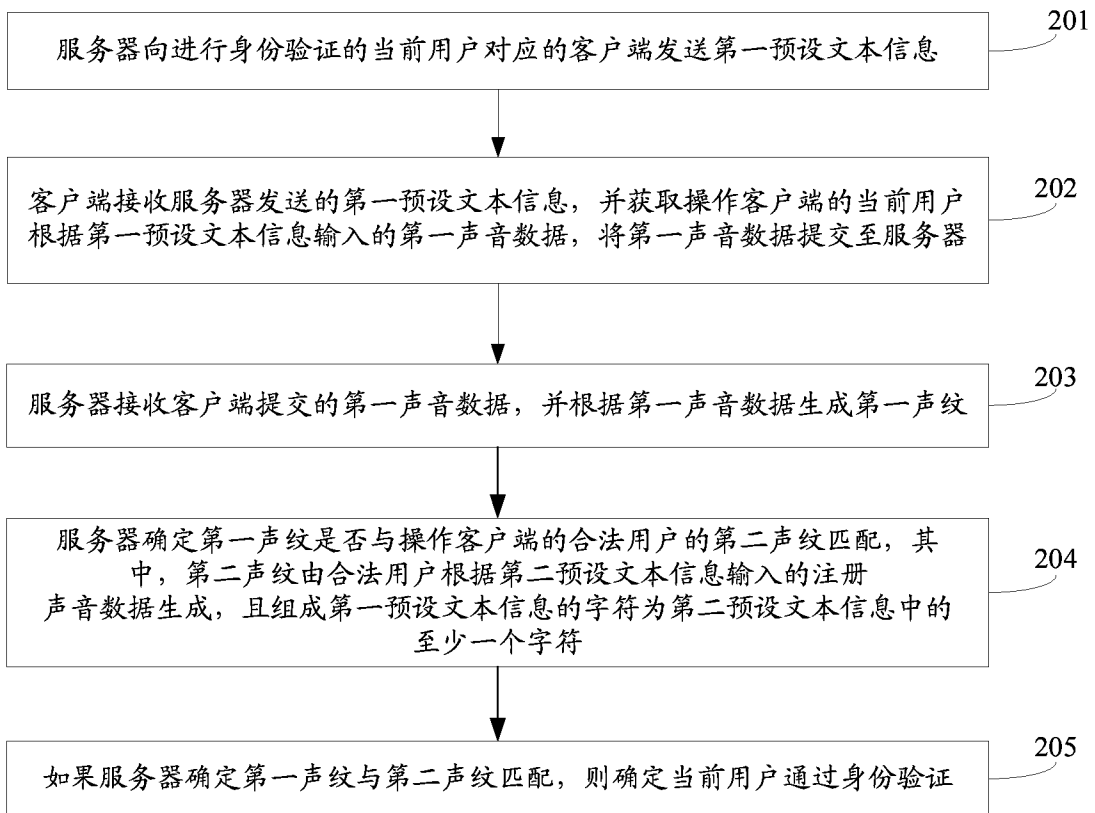


图 2

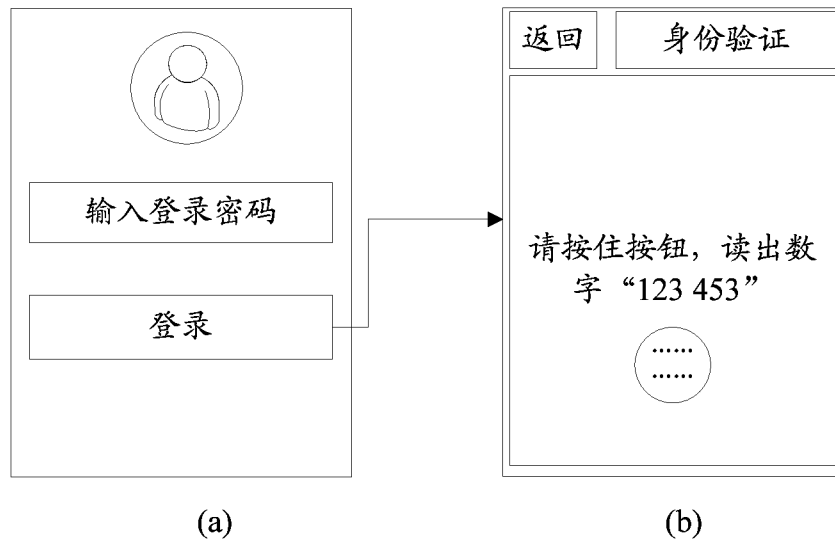


图 3

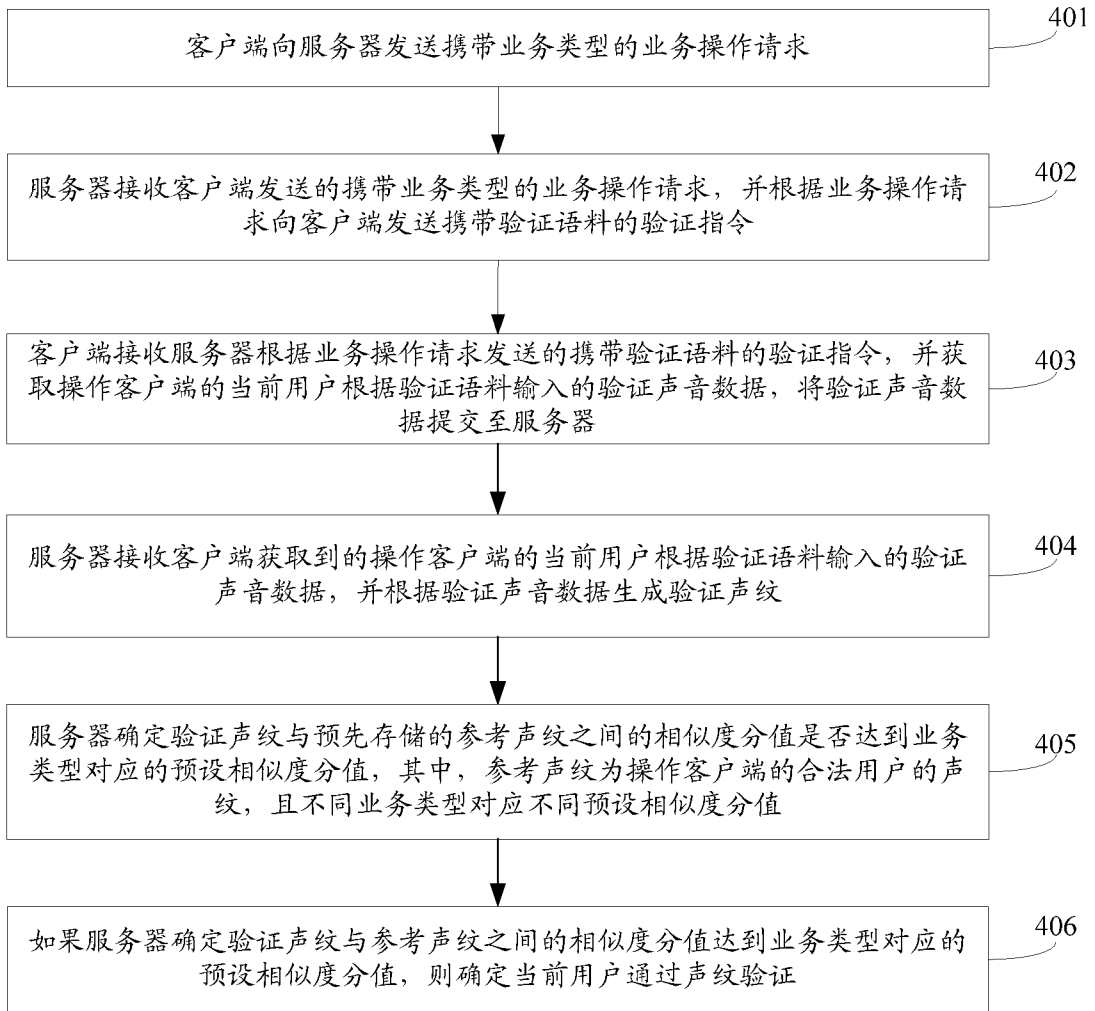


图 4

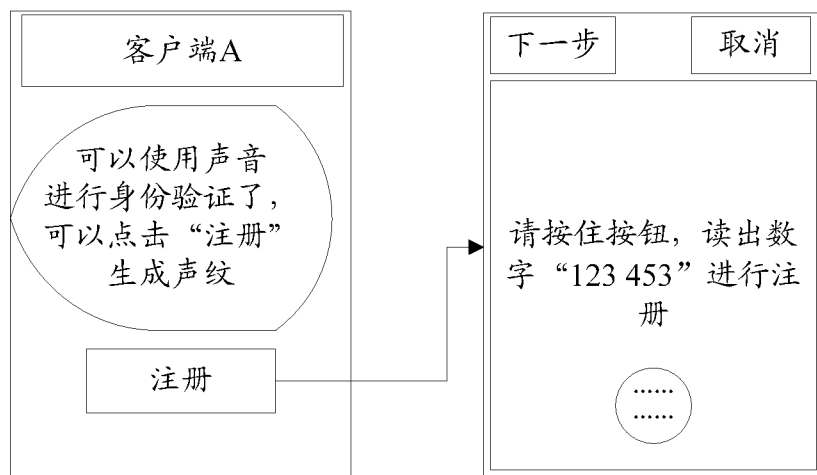


图 5

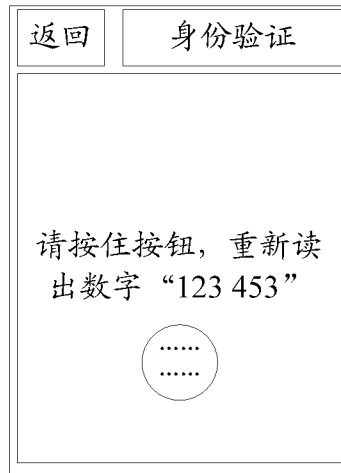


图 6

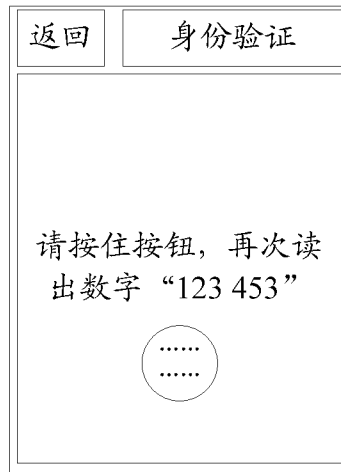


图 7

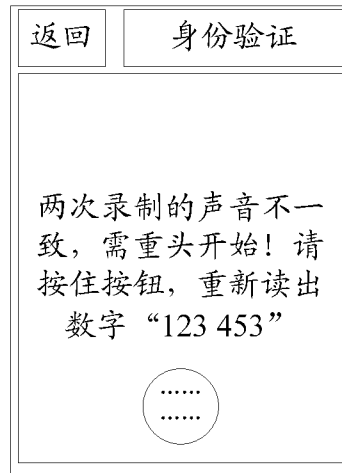


图 8

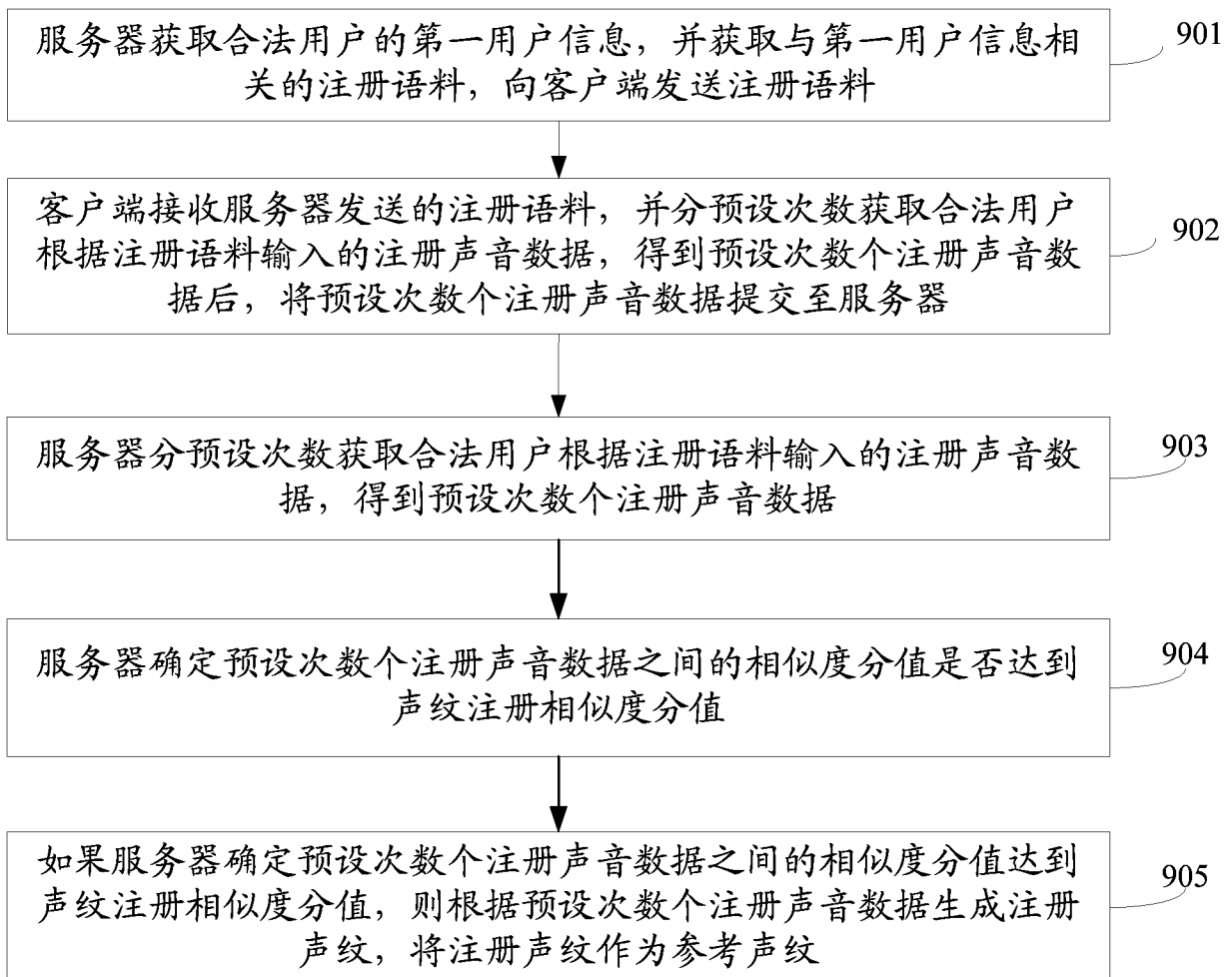


图 9

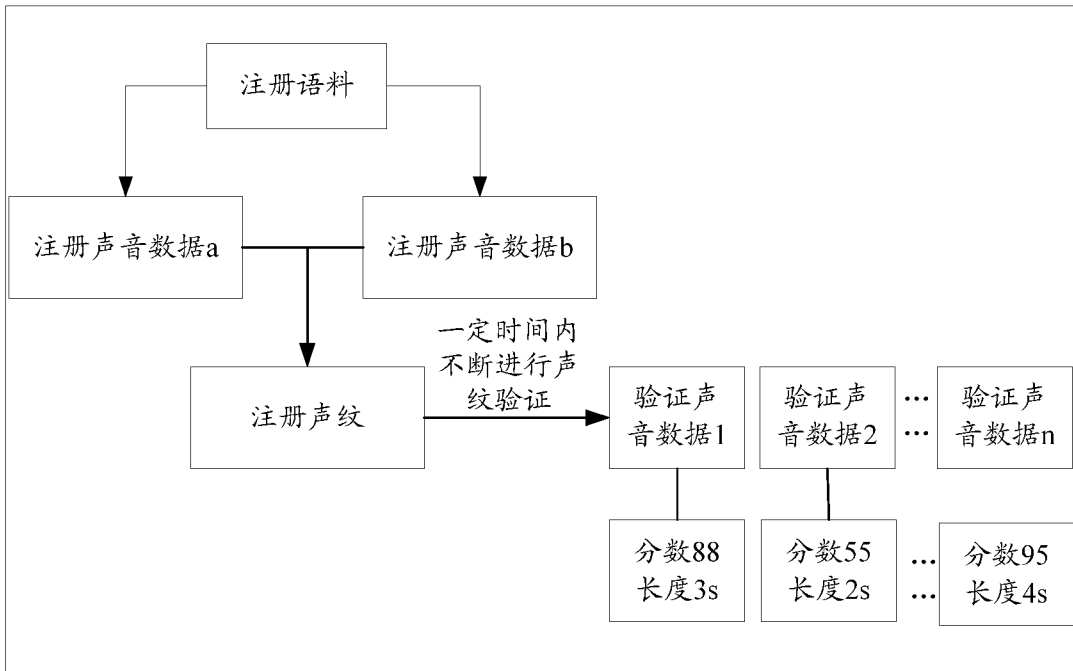


图 10

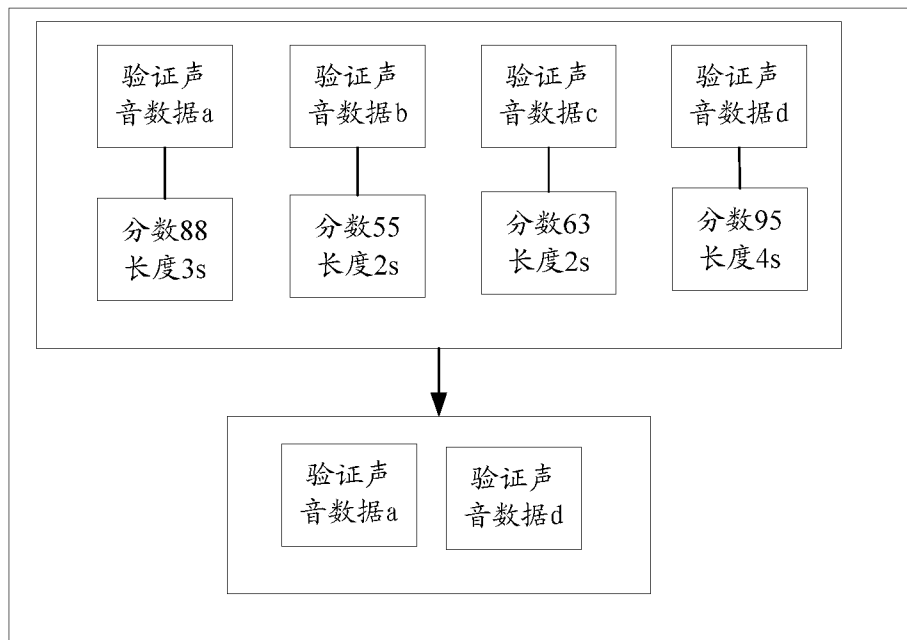


图 11

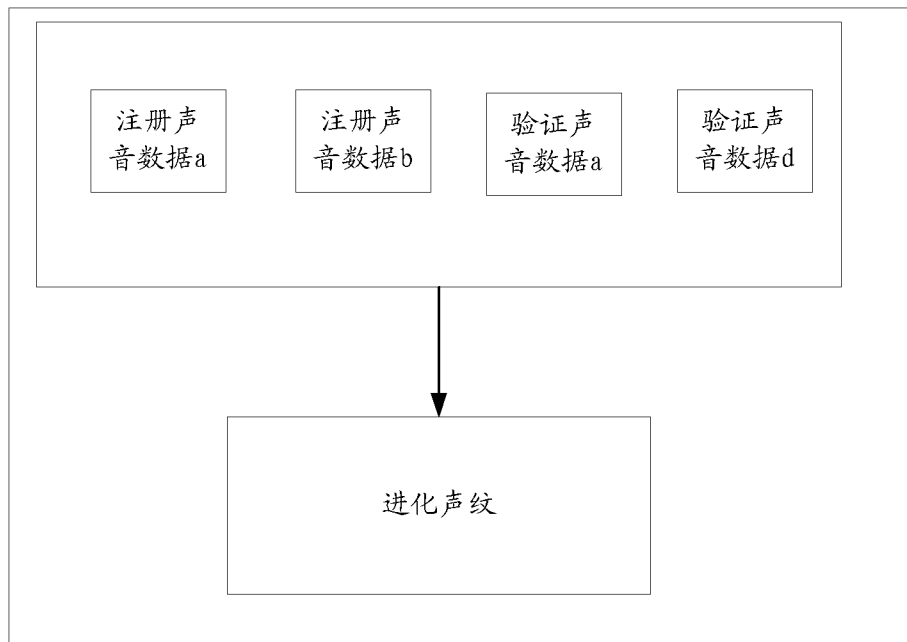


图 12

- 8/9 -

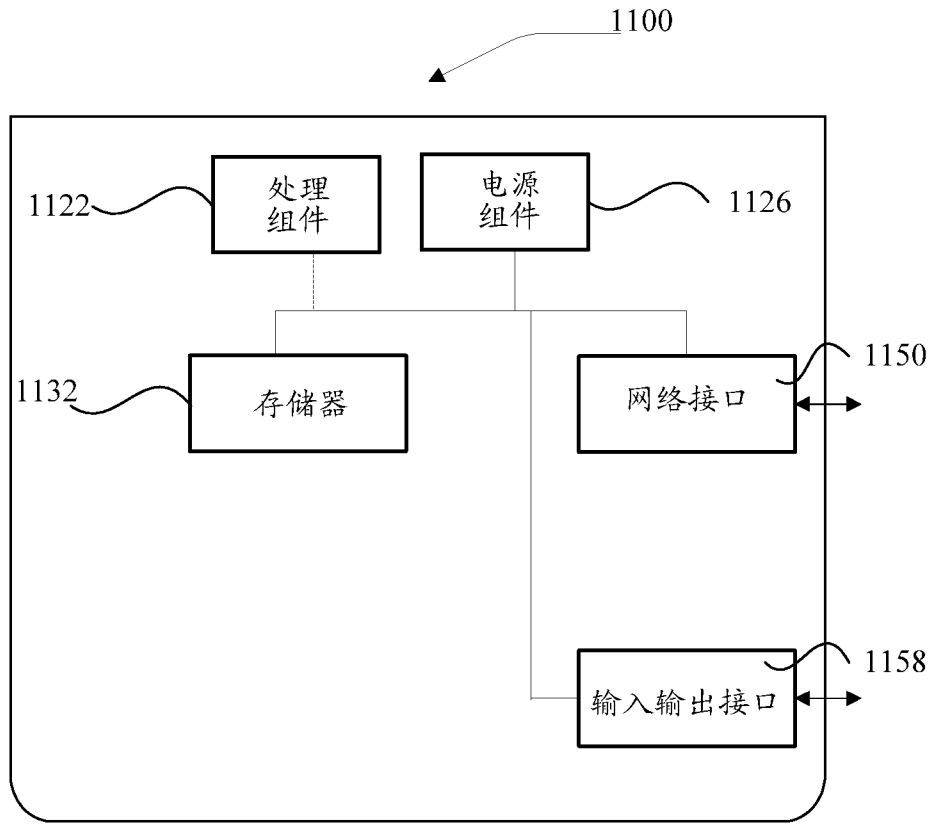


图 13

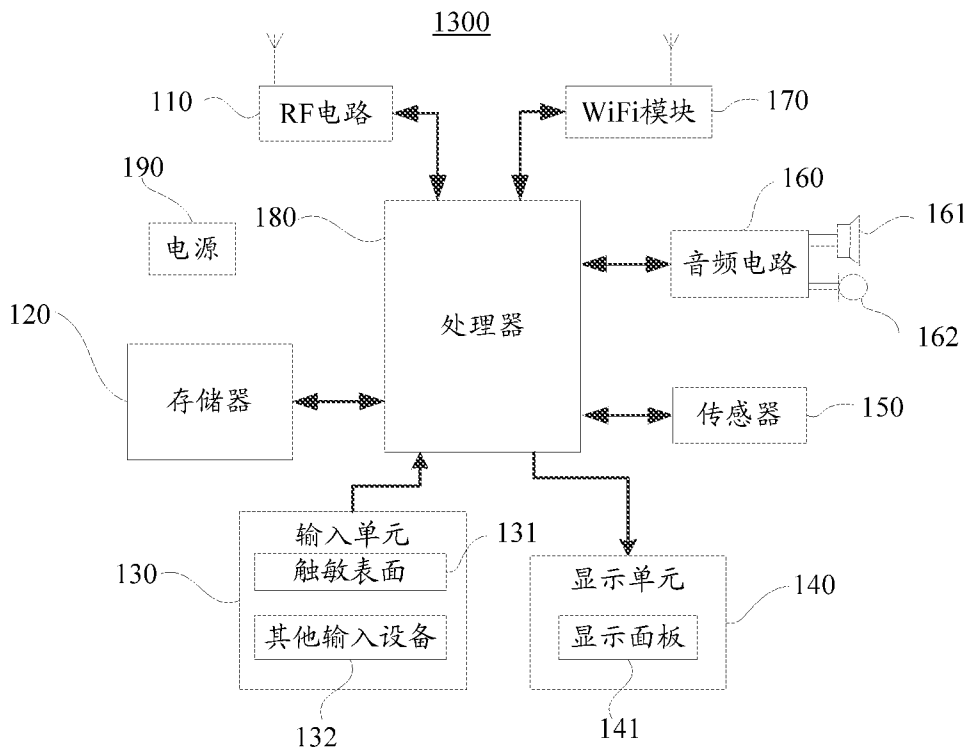


图 14

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/090505

A. CLASSIFICATION OF SUBJECT MATTER

H04L 9/32 (2006.01) i; H04L 29/06 (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)

H04L, 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)

CPRSABS; CNTXT; CNKI; VEN: voice w print, voice, identificat+, verificat+, user, client, score, mark, similar+, match+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102404287 A (SHENGYUE INFORMATION TECHNOLOGY (SHANGHAI) CO., LTD.), 04 April 2012 (04.04.2012), description, paragraphs 23-32, and figures 1-4	1-20
X	CN 103730114 A (SHANGHAI JIAO TONG UNIVERSITY WUXI RESEARCH INSTITUTE), 16 April 2014 (16.04.2014), description, paragraphs 44-89, and figure 1	1-20
X	US 2010228656 A1 (NICE SYSTEMS LTD.), 09 September 2010 (09.09.2010), description, paragraphs 24-110, and figure 5	1-20
PX	CN 104168270 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 26 November 2014 (26.11.2014), the whole document	1-20
PX	CN 104184587 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 03 December 2014 (03.12.2014), the whole document	1-20
PX	CN 104219050 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 17 December 2014 (17.12.2014), the whole document	1-20

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
24 November 2015 (24.11.2015)

Date of mailing of the international search report
29 December 2015 (29.12.2015)

Name and mailing address of the ISA/CN:
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No.: (86-10) 62019451

Authorized officer
PANG, Yan
Telephone No.: (86-10) **62089462**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2015/090505

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102404287 A	04 April 2012	None	
CN 103730114 A	16 April 2014	None	
US 2010228656 A1	09 September 2010	US 8145562 B2	27 March 2012
CN 104168270 A	26 November 2014	None	
CN 104184587 A	03 December 2014	None	
CN 104219050 A	17 December 2014	None	

国际检索报告

国际申请号

PCT/CN2015/090505

<p>A. 主题的分类</p> <p>H04L 9/32(2006.01)i; H04L 29/06(2006.01)i; G10L 17/00(2013.01)i</p> <p>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</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04L, G10L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CPRSABS;CNTXT;CNKI;VEN:声纹, 声音, 验证, 用户, 客户, 分值, 评分, 相似, 匹配, voice w print, voice, identificat+, verificat+, user, client, score, mark, similar+, match+</p>																							
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 102404287 A (盛乐信息技术上海有限公司) 2012年 4月 4日 (2012 - 04 - 04) 说明书第23-32段, 图1-4</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 103730114 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 说明书第44-89段, 图1</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>US 2010228656 A1 (NICE SYSTEMS LTD.) 2010年 9月 9日 (2010 - 09 - 09) 说明书第24-110段, 图5</td> <td>1-20</td> </tr> <tr> <td>PX</td> <td>CN 104168270 A (腾讯科技深圳有限公司) 2014年 11月 26日 (2014 - 11 - 26) 全文</td> <td>1-20</td> </tr> <tr> <td>PX</td> <td>CN 104184587 A (腾讯科技深圳有限公司) 2014年 12月 3日 (2014 - 12 - 03) 全文</td> <td>1-20</td> </tr> <tr> <td>PX</td> <td>CN 104219050 A (腾讯科技深圳有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 102404287 A (盛乐信息技术上海有限公司) 2012年 4月 4日 (2012 - 04 - 04) 说明书第23-32段, 图1-4	1-20	X	CN 103730114 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 说明书第44-89段, 图1	1-20	X	US 2010228656 A1 (NICE SYSTEMS LTD.) 2010年 9月 9日 (2010 - 09 - 09) 说明书第24-110段, 图5	1-20	PX	CN 104168270 A (腾讯科技深圳有限公司) 2014年 11月 26日 (2014 - 11 - 26) 全文	1-20	PX	CN 104184587 A (腾讯科技深圳有限公司) 2014年 12月 3日 (2014 - 12 - 03) 全文	1-20	PX	CN 104219050 A (腾讯科技深圳有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文	1-20
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	CN 102404287 A (盛乐信息技术上海有限公司) 2012年 4月 4日 (2012 - 04 - 04) 说明书第23-32段, 图1-4	1-20																					
X	CN 103730114 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 说明书第44-89段, 图1	1-20																					
X	US 2010228656 A1 (NICE SYSTEMS LTD.) 2010年 9月 9日 (2010 - 09 - 09) 说明书第24-110段, 图5	1-20																					
PX	CN 104168270 A (腾讯科技深圳有限公司) 2014年 11月 26日 (2014 - 11 - 26) 全文	1-20																					
PX	CN 104184587 A (腾讯科技深圳有限公司) 2014年 12月 3日 (2014 - 12 - 03) 全文	1-20																					
PX	CN 104219050 A (腾讯科技深圳有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文	1-20																					
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																							
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																							
<p>国际检索实际完成的日期</p> <p>2015年 11月 24日</p>	<p>国际检索报告邮寄日期</p> <p>2015年 12月 29日</p>																						
<p>ISA/CN的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>	<p>受权官员</p> <p>庞艳</p> <p>电话号码 (86-10)62089462</p>																						

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/090505

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	102404287	A	2012年 4月 4日	无			
CN	103730114	A	2014年 4月 16日	无			
US	2010228656	A1	2010年 9月 9日	US	8145562	B2	2012年 3月 27日
CN	104168270	A	2014年 11月 26日	无			
CN	104184587	A	2014年 12月 3日	无			
CN	104219050	A	2014年 12月 17日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)