

(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
2017 年 5 月 18 日 (18.05.2017)



(10) 国际公布号
WO 2017/080239 A1

- (51) 国际专利分类号:
G10L 17/00 (2013.01) G10L 15/08 (2006.01)
G10L 15/02 (2006.01) G11B 27/10 (2006.01)
G10L 15/06 (2013.01)
- (21) 国际申请号: PCT/CN2016/089033
- (22) 国际申请日: 2016 年 7 月 7 日 (07.07.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510782729.3 2015 年 11 月 15 日 (15.11.2015) CN
- (71) 申请人: 乐视控股(北京)有限公司 (LE HOLDINGS (BEIJING) CO., LTD.) [CN/CN]; 中国北京市朝阳区姚家园路 105 号 3 号楼 10 层 1102, Beijing 100025 (CN)。乐视移动智能信息技术(北京)有限公司 (LEMOBILE INFORMATION TECHNOLOGY (BEIJING) CO., LTD.) [CN/CN]; 中国北京市朝阳区姚家园路 105 号乐视大厦 3 层, Beijing 100025 (CN)。
- (72) 发明人: 蔡竹沁 (CAI, Zhuqin); 中国北京市朝阳区姚家园路 105 号乐视大厦 3 层, Beijing 100025

(CN)。齐峰岩 (QI, Fengyan); 中国北京市朝阳区姚家园路 105 号乐视大厦 3 层, Beijing 100025 (CN)。牛磊 (NIU, Lei); 中国北京市朝阳区姚家园路 105 号乐视大厦 3 层, Beijing 100025 (CN)。关彬 (GUAN, Bin); 中国北京市朝阳区姚家园路 105 号乐视大厦 3 层, Beijing 100025 (CN)。

(74) 代理人: 北京市惠诚律师事务所 (BEIJING HUICHENG LAW FIRM); 中国北京市新中街 68 号聚龙花园 8 号楼 4 层 403, Beijing 100027 (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,

[见续页]

(54) Title: AUDIO RECORDING TAGGING METHOD AND RECORDING DEVICE

(54) 发明名称: 录音标记方法及录音装置

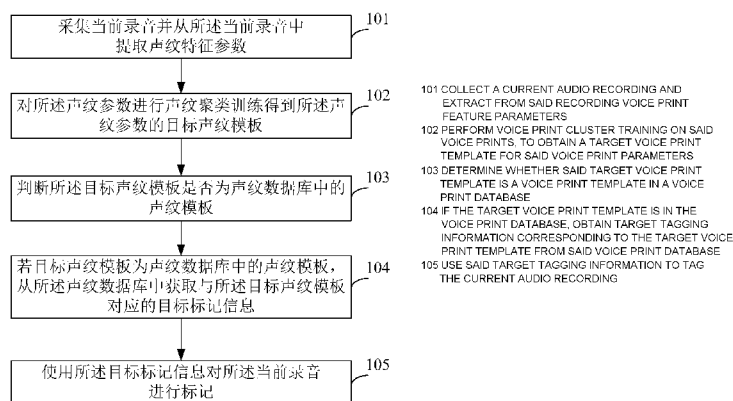


图 1

(57) Abstract: An audio recording tagging method and recording device: collecting a current audio recording and extracting from said recording voice print feature parameters (101); performing voice print cluster training on said voice prints, to obtain a target voice print template for said voice print parameters (102); determining whether said target voice print template is a voice print template in a voice print database (103); if so, obtaining target tagging information corresponding to the target voice print template from said voice print database (104); using said target tagging information to tag the current audio recording (105). Determining a voice print template for a current audio recording by means of a voice print and using a constructed voice print database to obtain corresponding tagging information for the current audio recording, and thereby tagging the current audio recording, makes possible automatic tagging functionality for audio recordings, and saves the user time when adding tags.

(57) 摘要:

[见续页]



WO 2017/080239 A1



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)。

根据细则 4.17 的声明:

— 关于申请人有权要求在先申请的优先权(细则 4.17(iii))

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

录音标记方法及录音装置，通过采集当前录音并从所述当前录音中提取声纹特征参数（101），对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板（102），判断所述目标声纹模板是否为声纹数据库中的声纹模板（103），如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息（104），使用所述目标标记信息对所述当前录音进行标记（105）。通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

录音标记方法及录音装置

本专利申请要求申请日为 2015 年 11 月 15 日、申请号为
5 2015107827293 的中国专利申请的优先权，并将上述专利申请以引用的方式全文引入本文中。

技术领域

本申请实施例涉及电子技术领域，尤其涉及一种录音标记方法及录音装置。

10 背景技术

目前智能手机逐渐融入到了人们日常生活之中，不但成为日常通讯设备，也成为日常易于携带的记录设备。其中，用户通过智能手机的录音应用程序（Application，简称 APP）可以对语音信息进行录制和保存，便于用户快速地保存一段难以直接记忆语音信息，并且还可以
15 以多次使用该录音。

现在很多录音 APP 可以让用户在录音的过程中，手动地对录音进行标记。通过这种手动地打标机的方式，便于整理和使用保存后的录音信息。当用户需要打标记时，只需要在录音时点击标设置好的记按钮，此时就会在录音的位置出现一个标记，在播放录音时该标记也会
20 显示在播放条上。

在录制多人对话的过程中，当讲话者发生变化时，用户就需要相应地在录音中变更与当前讲话人对应的标记，这种通过手动打标记的方法便捷性较差，而且使得用户也无法关注录音的内容，需要花费用户的很多时间。

25 发明内容

本申请实施例提供一种录音标记方法及录音装置，用于解决手动标记时存在便捷性较差以及浪费时间的问题。

为了实现上述目的，本申请实施例提供了一种录音标记方法，包括：

采集当前录音并从所述当前录音中提取声纹特征参数；

对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板；

判断所述目标声纹模板是否为声纹数据库中的声纹模板；

5 如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；

使用所述目标标记信息对所述当前录音进行标记。

为了实现上述目的，本申请实施例提供了一种录音装置，包括：

提取模块，用于采集当前录音并从所述当前录音中提取声纹特征

10 参数；

训练模块，用于对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板；

判断模块，用于判断所述目标声纹模板是否为声纹数据库中的声纹模板；

15 获取模块，用于在所述判断模块的判断结果为是时，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；

标记模块，用于使用所述目标标记信息对所述当前录音进行标记。

另一方面，本申请实施例提供一种录音装置，包括存储器、一个或多个处理器以及一个或多个程序，其中，所述一个或多个程序在由
20 所述一个或多个处理器执行时执行下述操作：采集当前录音并从所述当前录音中提取声纹特征参数；对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板；判断所述目标声纹模板是否为声纹数据库中的声纹模板；如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；使用所述目标标记信息对
25 所述当前录音进行标记。

另一方面，本申请实施例提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可执行指令，所述计算机可执行指令响应于执行使得录音装置执行操作，所述操作包括：采集当前录音并从所述当前录音中提取声纹特征参数；对所述声纹参数进行声纹聚

类训练得到所述声纹参数的目标声纹模板；判断所述目标声纹模板是否为声纹数据库中的声纹模板；如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；使用所述目标标记信息对所述当前录音进行标记。

- 5 本申请实施例的录音标记方法及录音装置，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息，使用所述目标标记
- 10 信息对所述当前录音进行标记。本申请实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

附图说明

- 15 图 1 为本申请实施例一的录音标记方法的流程示意图；
图 2 为本申请实施例一的录音标记方法的应用示例示意图；
图 3 为本申请实施例二的录音标记方法的流程示意图；
图 4 为本申请实施例二的录音标记方法的应用示例示意图之一；
图 5 为本申请实施例一的录音标记方法的应用示例示意图之二；
20 图 6 为本申请实施例一的录音标记方法的应用示例示意图之三；
图 7 为本申请实施例二中的声纹数据库建立方法的流程示意图；
图 8 为本申请实施例三的录音装置的结构示意图；
图 9 为本申请实施例四的录音装置的结构示意图；
图 10 为本申请提供的录音装置又一个实施例的结构示意图；
25 图 11 为本申请提供的用于录音标记的计算机程序产品一个实施例的结构示意图。

具体实施方式

下面结合附图对本申请实施例提供的录音标记方法及录音装置进行详细描述。

实施例一

如图 1 所示,其为本申请实施例一的录音标记方法的流程示意图,

5 该录音标记方法包括:

步骤 101、采集当前录音并从所述当前录音中提取声纹特征参数。

用户可以通过智能手机的用户界面,开启智能手机中下载的录音 APP 的录音功能,录音 APP 开始对当前录音进行采集,在采集的过程中录音 APP 可以对声音进行预处理,例如,采集的数据进行分帧、加窗和滤波等。

进一步地,对采集的当前录音进行特征分析,进而得到当前录音的声波特征参数,其中,声波特征参数包括:声音的能量、共振峰、梅尔倒谱系数(Mel-frequency cepstrum coefficients,简称 MFCC)以及线性预测系数(Linear Prediction Coefficients,简称 LPC)。

15 步骤 102、对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板。

本实施例中,为了识别出录音的模板,设置有声纹聚类训练器,在获取到声纹特征参数后,通过该训练器对声纹特征参数进行声纹聚类训练,就可以得到该当前录音对应的目标声纹模板。

20 步骤 103、判断所述目标声纹模板是否为声纹数据库中的声纹模板。

本实施例中,通过训练器对样本声音进行声纹聚类训练,得到了样本声音对应的样本声纹模板,使用样本声纹模板预先设置了一个声纹数据库存储在录音 APP 中。一般声纹数据库中存储有多个样本声纹模板,以便于用户在录音过程中进行录音标记。在获取到目标声纹模板后,录音 APP 可以在声纹数据库中进行查找,判断该目标声纹模板是否存在于该声纹数据库中。

步骤 104、若目标声纹模板为声纹数据库中的声纹模板,从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息。

在声纹数据库中不仅保存有样本声纹模板，而且还存储有与样本声纹模板对应的标记信息，一般每个样本声纹模板对应有各自的标记信息。当在声纹数据库中获取到与目标声纹模板对应的样本声纹模板时，就可以获取与该目标声纹模板对应的目标标记信息。

5 步骤 105、使用所述目标标记信息对所述当前录音进行标记。

在获取到目标标记信息后，录音 APP 自动使用该目标标记信息对当前录音进行标记。

如图 2 所示，其为本实施例的应用示例示意图，由于说话人的声纹的唯一性可以利用声纹作为区别说话人的唯一特征。对一段录音进行分段标记。首先对当前录音进行声纹分析，对于不同声纹使用不同的标记进行标记。同一个说话人使用相同的标记进行标记。比如一段录音有 5 个说话人，分别使用左斜线、右斜线、横线、竖线以及网格进行标记说话人 A、B、C、D、E。其中，当说话人 A 在这段录音中有两次被其他说话人分开的发言，这两段发言都会使用左斜线标记说话人 A，以表明是同一个说话人的录音段落。为了用户更直观的看到说话人的不同，可以使用不同的颜色标记说话人，例如，分别使用红色、黄色、蓝色、绿色和紫色来标记说话人 A、B、C、D、E。或者当说话人 A 在这段录音中有两次被其他说话人分开的发言，这两段发言都会使用红色标记说话人 A，以表明是同一个说话人的录音段落。

10

15

20

本实施例提供的录音标记方法，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息，使用所述目标标记信息对所述当前录音进行标记。本实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

25

实施例二

如图 3 所示,其为本申请实施例二的录音标记方法的流程示意图,该录音标记方法包括:

步骤 201、采集当前录音并从所述当前录音中提取声纹特征参数。

5 步骤 202、对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板。

步骤 203、判断所述目标声纹模板是否为声纹数据库中的声纹模板

如果判断结果为是,执行步骤 204;否则执行步骤 205。

10 步骤 204、从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息。

步骤 201~步骤 204 的相关内容可参见上述实施例一中步骤 101~步骤 104 中的记载,此处不再赘述。

步骤 205、生成与所述目标声纹模板对应的所述目标标记信息。

15 在识别出目标声纹模板并不存在与声纹数据库中之后,录音 APP 可以为该目标声纹模板设置一个目标标记信息,以通过该目标标记信息对该目标声纹模板进行标记。

步骤 206、使用所述目标标记信息对所述当前录音进行标记。

20 步骤 207、建立所述目标声纹模板与所述目标标记信息之间映射关系并存储在所述声纹数据库中。

步骤 208、接收用户通过终端发送的备注信息。

步骤 209、使用所述备注信息对所述当前录音进行备注。

步骤 210、将所述备注信息更新到所述声纹数据中所述目标标记信息中。

25 接收用户通过终端发送的备注信息,备注信息可以为当前录音的来源名称,在终端获取到备注信息后,指示录音 APP 使用该备注信息对当前录音进行备注。例如,录音 APP 可以为当前录音对应的位置添加一个标签。进一步地,录音 APP 还可以将获取到的备注信息更新到声纹数据中与当前录音对应的目标声纹模板对应的目标标记信息中,

以便录制的声音为当前录音对应的音源时可以再次被调用。

如图 4 所示，其为本实施例的应用示例示意图，当录音 APP 对当前录音进行自动标记后，用户可以通过终端向录音 APP 发送备注信息，用于给这段录音中每位说话人添加备注信息。比如，用户可以通过录音 APP 将用左斜线标记的说话人 A 备注为“张老师”。用户可为新说话人添加的备注信息，并直接与该说话人的声纹信息匹配，并作为这段录音的名称。

如图 5 所示，其为本实施例的应用示例示意图，当用户新建一段录音，如果其中包含已保存声音名称的说话人的录音，在声纹分析后，这位说话人的录音段落会直接标记为已保存的标记信息。比如已经保存了之前一段录音的说话人 A 为“张老师”，新建一段包含这个说话人的录音不会再显示说话人 A 的标记，而是显示“张老师”。

如图 6 所示，其为本实施例的应用示例示意图，录音中包含用户保存过的讲话人对应的标记信息，按照所标记的说话人，更快定位需要寻找的录音。比如用户想要寻找张老师的讲课录音，只要寻找“张老师”的标签即可。

在步骤 201 采集当前录音并从所述当前录音中提取声纹特征参数之前，还需要通过样本声音建立一个声纹数据库。

如图 7 所示，其为本申请实施例二中的声纹数据库建立方法的流程图示意图，该声纹数据库建立方法包括：

步骤 301、对样本声音进行分析，提取所述样本声音的所述声纹特征参数。

本实施例中，将录音 APP 在当前录音之前的每次录制的声音作为样本声音。在获取到每次录音后，录音 APP 会对录音的样本声音进行分析，提取出该样本声音的声纹特征参数，其中声纹特征参数包括：声音的能量、共振峰、MFCC 以及 LPC 等。

步骤 302、根据所述样本声音的所述声纹特征参数进行声纹聚类训练生成样本声纹模板。

为了对获取到样本声纹的声纹特征参数进行声纹聚类训练，需要

进一步确定该声纹特征参数是否为同一个音源的声音，具体地，预设时间段内的所述样本声音的所述声纹特征参数，当所述预设时间内的所述样本声音的所述声纹特征参数具有相似性时，对所述样本声音的所述声纹特征参数进行声纹聚类训练生成所述样本声纹模板。如果确定出样本声纹的声纹特征参数的不具有相似性，则需要将声纹特征参数进行缓存，再判断出该声音特征参数具有相似性之后，对声纹特征参数进行声纹聚类训练生成样本声纹模板。

比如，有一段录音中有 5 个说话人，这 5 个说话人就可以做完样本声音，在通过声纹聚类训练后，可以识别出这个 5 个说话人分别为说话人 A、B、C、D、E，并 5 个说话人生成相应的样本声纹模板。

步骤 303、为所述样本声纹模板生成对应的样本标记信息。

在生成样本声纹模板后，为样本声音生成对应的样本标记信息，例如同一个说话人使用相同的标记进行标记。本实施例中，可以使用左斜线、右斜线、横线、竖线以及网格进行标记说话人 A、B、C、D、E。

步骤 304、使用所述样本声纹模板、所述样本标记信息以及所述样本声纹模板与所述样本标记信息之间的映射关系生成所述声纹数据库。

为了提高对录音标记的快捷性，本实施例中，使用样本声纹模板、所述样本标记信息以及所述样本声纹模板与所述样本标记信息之间的映射关系生成所述声纹数据库。每次对录音进行声纹聚类训练后生成的声纹模板都会作为样本声纹模板保存到声纹数据库中，而且会将对该样本声纹模板的标记信息以及两者之间的映射关系也会保存到声纹数据库中，以对声纹数据库进行更新。这样当再次遇到同一说话人的录音时，录音 APP 通过声纹分析，能够很迅速地对该说话人的录音进行标记，提高了录音标记的便捷性。

本实施例通过声纹识别当前录音对应的声纹模板，并利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

实施例三

如图 8 所示，其为本申请实施例三的录音装置的结构示意图。该装置包括：提取模块 11、训练模块 12、判断模块 13、获取模块 14 和标记模块 15。

- 5 其中，提取模块 11，用于采集当前录音并从所述当前录音中提取声纹特征参数。

提取模块 11 对采集的当前录音进行特征分析，进而得到当前录音的声波特征参数，其中，声波特征参数包括：声音的能量、共振峰 MFCC 以及 LPC。

- 10 训练模块 12，用于对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板。

本实施例中，为了识别出录音的模板，设置有训练模块 12，其中训练模块 12 可以为声纹聚类训练器。在获取到声纹特征参数后，通过该训练模块 12 对声纹特征参数进行声纹聚类训练，就可以得到该当前
15 录音对应的目标声纹模板。

判断模块 13，用于判断所述目标声纹模板是否为声纹数据库中的声纹模板。

- 本实施例中，通过训练模块 12 对样本声音进行声纹聚类训练，得到了样本声音对应的样本声纹模板，使用样本声纹模板预先设置了一个声纹数据库。一般声纹数据库中存储有多个样本声纹模板，以便于
20 用户在录音过程中进行录音标记。在获取到目标声纹模板后，判断模块 13 可以在声纹数据库中进行查找，判断该目标声纹模板是否存在于该声纹数据库中。

- 获取模块 14，用于在所述判断模块的判断结果为是时，从所述声
25 纹数据库中获取与所述目标声纹模板对应的目标标记信息。

在声纹数据库中不仅保存有样本声纹模板，而且还存储有与样本声纹模板对应的标记信息，一般每个样本声纹模板对应有各自的标记信息。当在声纹数据库中获取到与目标声纹模板对应的样本声纹模板时，获取模块 14 就可以获取与该目标声纹模板对应的目标标记信息。

标记模块 15，用于使用所述目标标记信息对所述当前录音进行标记。

在获取到目标标记信息后，标记模块 15 自动使用该目标标记信息对当前录音进行标记。

- 5 本实施例提供的录音装置，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息，使用所述目标标记信息对所述当前
- 10 录音进行标记。本实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

实施例四

- 15 图 9 所示，其为本申请实施例四的录像装置的结构示意图。该装置包括上述实施例三中的提取模块 11、训练模块 12、判断模块 13、获取模块 14 和标记模块 15 之外，还包括生成模块 16、建立模块 17 和接收模块 18。

- 20 其中，所述生成模块 16，用于在所述判断模块的判断结果为否时，生成与所述目标声纹模板对应的所述目标标记信息。

所述建立模块 17，用于在标记模块 15 使用所述目标标记信息对所述当前录音进行标记之后，建立所述目标声纹模板与所述目标标记信息之间映射关系并存储在所述声纹数据库中。

- 25 进一步地，所述提取模块 11，还用于在采集当前录音并从所述当前录音中提取声纹特征参数之前，对样本声音进行分析提取所述样本声音的所述声纹特征参数。

所述训练模块 12，还用于根据所述样本声音的所述声纹特征参数进行声纹聚类训练生成样本声纹模板。

所述生成模块 16，还用于为所述样本声纹模板生成对应的样本标

记信息。

所述建立模块 17，还用于使用所述样本声纹模板、所述样本标记信息以及所述样本声纹模板与所述样本标记信息之间的映射关系生成所述声纹数据库。

- 5 进一步地，所述训练模块 12，具体用于获取预设时间段内的所述样本声音的所述声纹特征参数，在所述预设时间内的所述样本声音的所述声纹特征参数相似时，对所述样本声音的所述声纹特征参数进行声纹聚类训练生成所述样本声纹模板。

- 10 进一步地，接收模块 18，用于在所述标记模块 15 使用所述目标标记信息对所述当前录音进行标记之后，接收用户通过终端发送的备注信息；

所述标记模块 15，还用于使用所述备注信息对所述当前录音进行备注；

- 15 所述建立模块 17，还用于将所述备注信息更新到所述声纹数据库中所述目标标记信息中。

本实施例提供的录音装置的各功能模块可用于执行图 1、3 和 7 所示的录音标记方法的流程，其具体工作原理不再赘述，详见方法实施例的描述。

- 20 本实施例提供的录音装置，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息，使用所述目标标记信息对所述当前录音进行标记。本实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。
- 25

实施例五

图 10 为本申请提供的录音装置又一个实施例的结构示意图。如图 10 所示，本申请实施例的录音装置包括：存储器 61、一个或多个处理器 62 以及一个或多个程序 63。

其中，所述一个或多个程序 63 在由一个或多个处理器 62 执行时
5 执行上述实施例中的任意一种方法。

本申请实施例的录音装置，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与所述目
10 标声纹模板对应的目标标记信息，使用所述目标标记信息对所述当前录音进行标记。本申请实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

15

实施例六

图 11 为本申请提供的用于录音标记的计算机程序产品一个实施例的结构示意图。如图 11 所示，本申请实施例的用于录音标记的计算机程序产品 71，可以包括信号承载介质 72。信号承载介质 72 可以包括一个或更多个指令 73，该指令 73 在由例如处理器执行时，处理器可以提供以上针对图 1-9 描述的功能。例如，指令 73 可以包括：
20 用于采集当前录音并从所述当前录音中提取声纹特征参数的一个或多个指令；用于对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板的一个或多个指令；用于判断所述目标声纹模板是否为声纹数据库中的声纹模板的一个或多个指令；用于如果判断结果为是，
25 从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息的一个或多个指令；以及用于使用所述目标标记信息对所述当前录音进行标记的一个或多个指令。因此，例如，参照图 8，录音装置可以响应于指令 73 来进行图 1 中所示的步骤中的一个或更多个。

在一些实现中，信号承载介质 72 可以包括计算机可读介质 74，诸如但不限于硬盘驱动器、压缩盘（CD）、数字通用盘（DVD）、数字带、存储器等。在一些实现中，信号承载介质 72 可以包括可记录介质 75，诸如但不限于存储器、读/写（R/W）CD、R/W DVD
5 等。在一些实现中，信号承载介质 72 可以包括通信介质 76，诸如但不限于数字和/或模拟通信介质（例如，光纤线缆、波导、有线通信链路、无线通信链路等）。因此，例如，计算机程序产品 71 可以通过 RF 信号承载介质 72 传送给多指滑动手势的识别装置的一个或多个模块，其中，信号承载介质 72 由无线通信介质（例如，符合 IEEE 802.11 标准的无线通信介质）传送。
10

本申请实施例的计算机程序产品，通过采集当前录音并从所述当前录音中提取声纹特征参数，对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板，判断所述目标声纹模板是否为声纹数据库中的声纹模板，如果判断结果为是，从所述声纹数据库中获取与
15 所述目标声纹模板对应的目标标记信息，使用所述目标标记信息对所述当前录音进行标记。本申请实施例通过声纹识别当前录音对应的声纹模板，利用建立的声纹数据库获取与当前录音对应的标记信息，进而对当前录音进行标记，实现了自动标记录音的功能，而且节省了用户添加标记的时间。

20 通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到各实施方式可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件。基于这样的理解，上述技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在计算机可读存储介质中，如 ROM/RAM、
25 磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备）执行各个实施例或者实施例的某些部分所述的方法。

最后应说明的是：以上各实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述各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，或者对其中部分或者全部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范围。

权 利 要 求 书

1、一种录音标记方法，其特征在于，包括：

采集当前录音并从所述当前录音中提取声纹特征参数；

对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹

5 模板；

判断所述目标声纹模板是否为声纹数据库中的声纹模板；

如果判断结果为是，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；

使用所述目标标记信息对所述当前录音进行标记。

10 2、根据权利要求 1 所述的录音标记方法，其特征在于，所述使用所述目标标记信息对所述当前录音进行标记之前，还包括：

如果判断结果为否，生成与所述目标声纹模板对应的所述目标标记信息。

15 3、根据权利要求 2 所述的录音标记方法，其特征在于，所述使用所述目标标记信息对所述当前录音进行标记之后，还包括：

建立所述目标声纹模板与所述目标标记信息之间映射关系并存储在所述声纹数据库中。

4、根据权利要求 1-3 任一项所述的录音标记方法，其特征在于，所述采集当前录音并从所述当前录音中提取声纹特征参数之前，包括：

20 对样本声音进行分析，提取所述样本声音的所述声纹特征参数；

根据所述样本声音的所述声纹特征参数进行声纹聚类训练生成样本声纹模板；

为所述样本声纹模板生成对应的样本标记信息；

25 使用所述样本声纹模板、所述样本标记信息以及所述样本声纹模板与所述样本标记信息之间的映射关系生成所述声纹数据库。

5、根据权利要求 4 所述的录音标记方法，其特征在于，所述根据所述样本声音的所述声纹特征参数进行声纹聚类训练生成样本声纹模板包括：

获取预设时间段内的所述样本声音的所述声纹特征参数；

在所述预设时间内的所述样本声音的所述声纹特征参数具有相似性时，对所述样本声音的所述声纹特征参数进行声纹聚类训练生成所述样本声纹模板。

- 6、根据权利要求 1-3 任一项所述的录音标记方法，其特征在于，
5 所述使用所述目标标记信息对所述当前录音进行标记之后，还包括：

接收用户通过终端发送的备注信息；

使用所述备注信息对所述当前录音进行备注；

将所述备注信息更新到所述声纹数据中所述目标标记信息中。

- 7、一种录音装置，其特征在于，包括：

- 10 提取模块，用于采集当前录音并从所述当前录音中提取声纹特征参数；

训练模块，用于对所述声纹参数进行声纹聚类训练得到所述声纹参数的目标声纹模板；

- 15 判断模块，用于判断所述目标声纹模板是否为声纹数据库中的声纹模板；

获取模块，用于在所述判断模块的判断结果为是时，从所述声纹数据库中获取与所述目标声纹模板对应的目标标记信息；

标记模块，用于使用所述目标标记信息对所述当前录音进行标记。

- 8、根据权利要求 7 所述的录音装置，其特征在于，还包括：

- 20 生成模块，用于在所述判断模块的判断结果否时，生成与所述目标声纹模板对应的所述目标标记信息。

- 9、根据权利要求 8 所述的录音装置，其特征在于，还包括：

- 25 建立模块，用于在所述标记模块使用所述目标标记信息对所述当前录音进行标记之后，建立所述目标声纹模板与所述目标标记信息之间映射关系并存储在所述声纹数据库中。

10、根据权利要求 7-9 任一项所述的录音装置，其特征在于，所述提取模块，还用于在采集当前录音并从所述当前录音中提取声纹特征参数之前，对样本声音进行分析提取所述样本声音的所述声纹特征参数；

所述训练模块，还用于根据所述样本声音的所述声纹特征参数进行声纹聚类训练生成样本声纹模板；

所述生成模块，还用于为所述样本声纹模板生成对应的样本标记信息；

- 5 所述建立模块，还用于使用所述样本声纹模板、所述样本标记信息以及所述样本声纹模板与所述样本标记信息之间的映射关系生成所述声纹数据库。

- 11、根据权利要求 10 所述的录音装置，其特征在于，所述训练模块，具体用于获取预设时间段内的所述样本声音的所述声纹特征参数，
10 在所述预设时间内的所述样本声音的所述声纹特征参数具有相似时，对所述样本声音的所述声纹特征参数进行声纹聚类训练生成所述样本声纹模板。

12、根据权利要求 7-9 任一项所述的录音装置，其特征在于，还包括：

- 15 接收模块，用于在所述标记模块使用所述目标标记信息对所述当前录音进行标记之后，接收用户通过终端发送的备注信息；

所述标记模块，还用于使用所述备注信息对所述当前录音进行备注；

- 20 所述建立模块，还用于将所述备注信息更新到所述声纹数据中所述目标标记信息中。

1/8

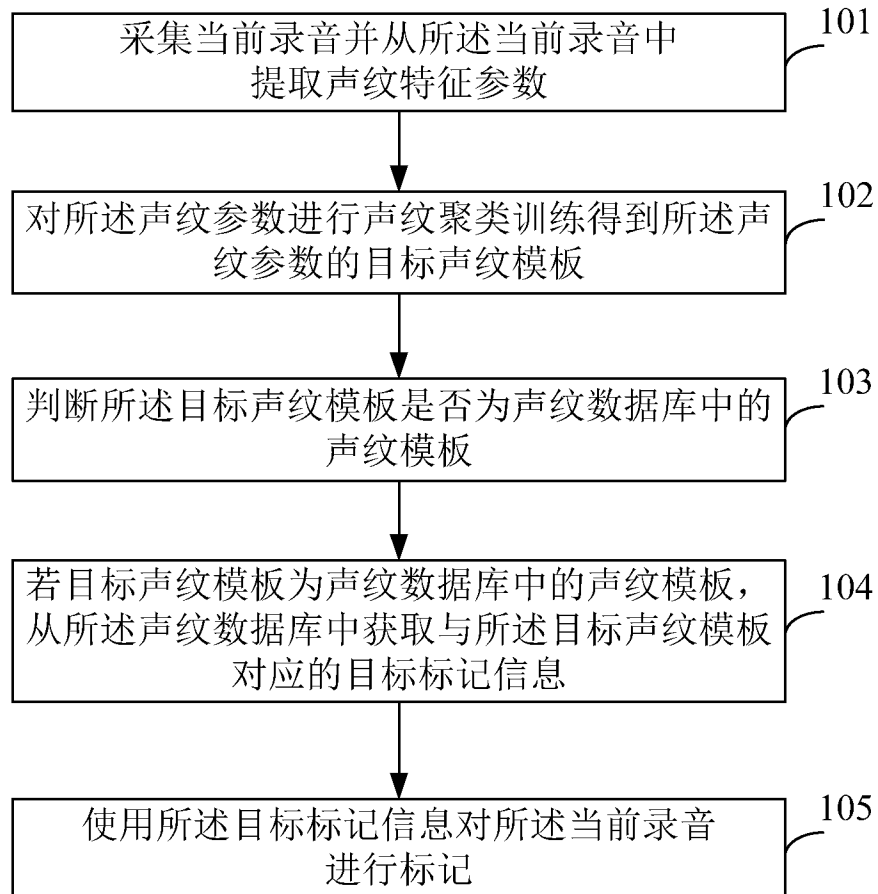


图 1



图 2

3/8

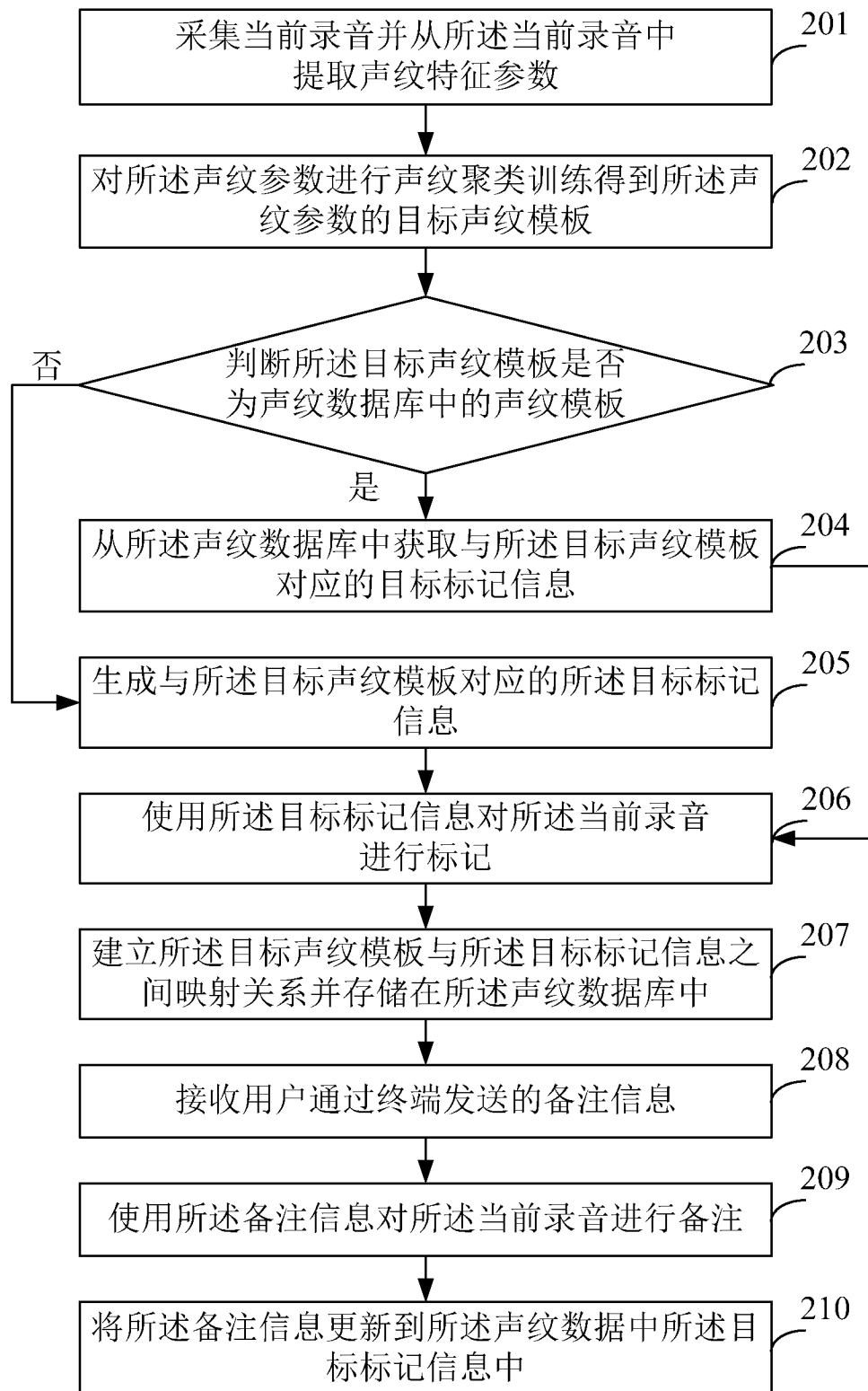


图 3

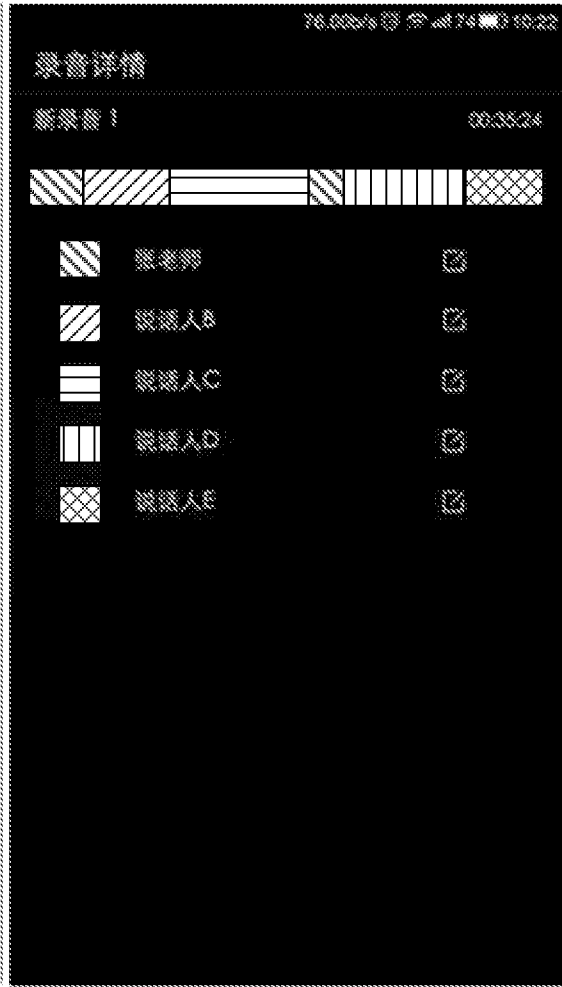


图 4



图 5

6/8

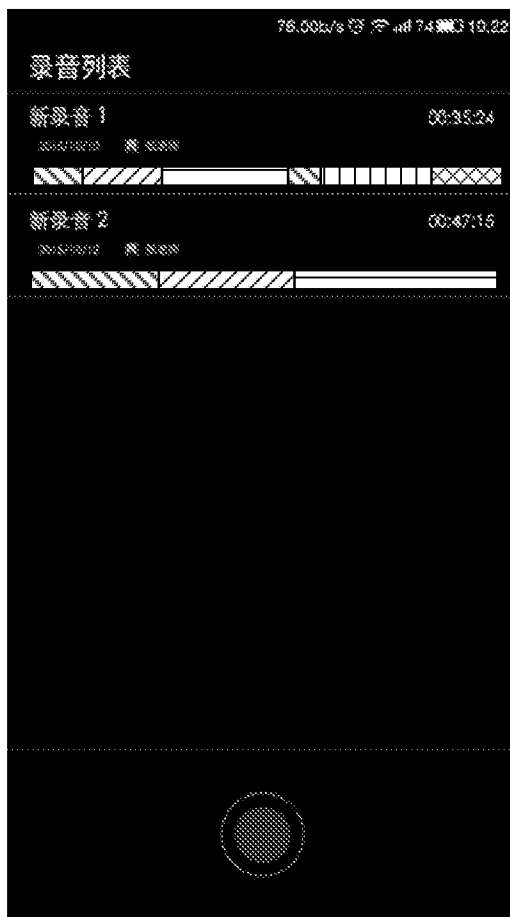


图 6

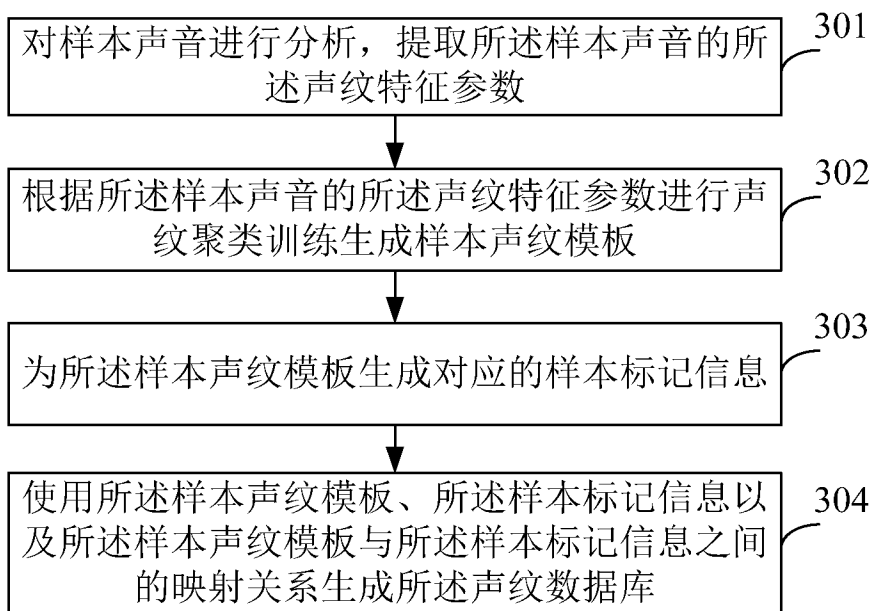


图 7

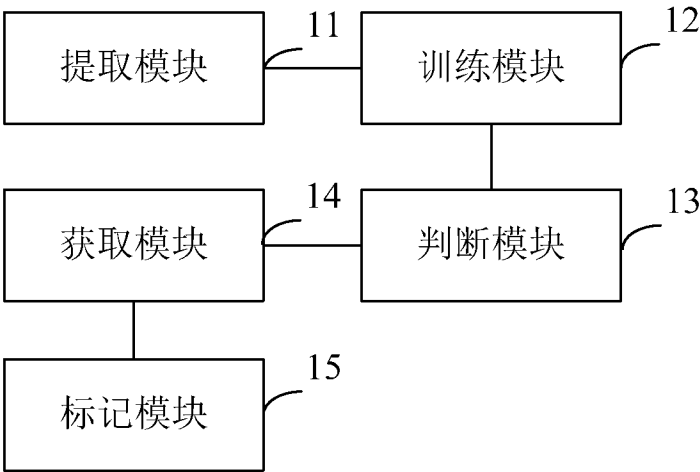


图 8

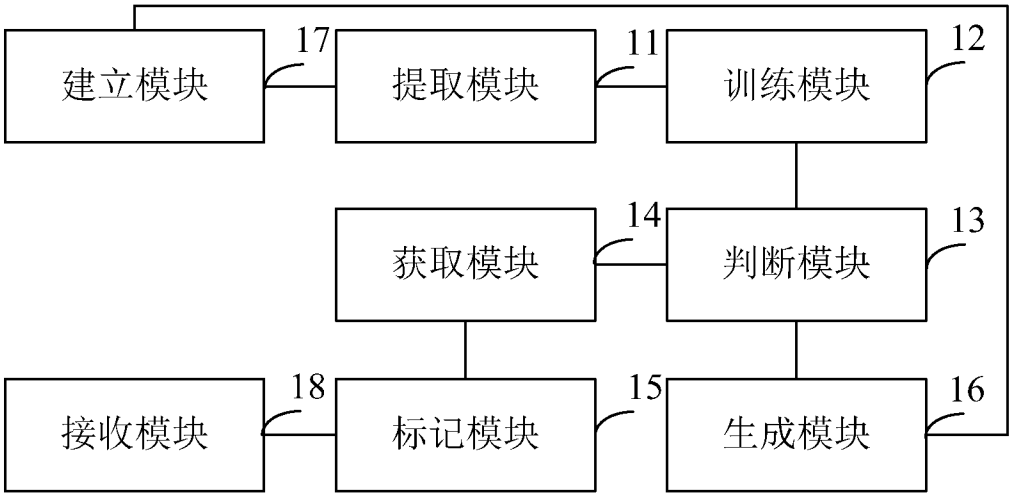


图 9

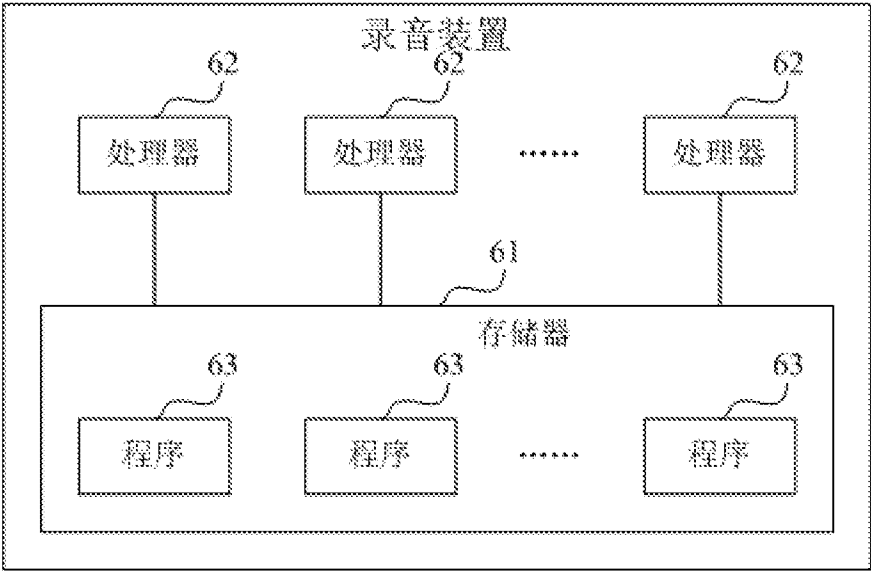


图 10

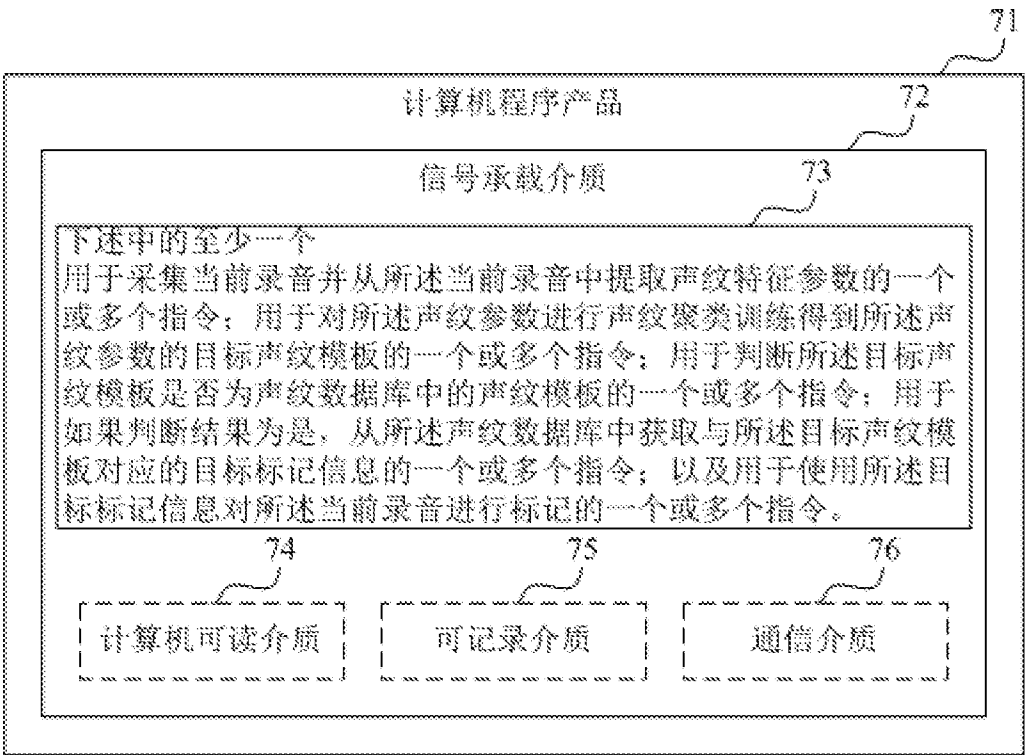


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/089033**A. CLASSIFICATION OF SUBJECT MATTER**

G10L 17/00 (2013.01) i; G10L 15/02 (2006.01) i; G10L 15/06 (2013.01) i; G10L 15/08 (2006.01) i; G11B 27/10 (2006.01) i
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L 17/-; G10L 15/-; G11B 27/-

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)

CNPAT, WPI, EPODOC, CNKI: template, soundprint, model, database, mark+, map, train+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102985965 A (MICROSOFT CORPORATION), 20 March 2013 (20.03.2013), description, paragraphs 27-32, claims 1-3, and figure 3	1-12
A	CN 104978507 A (SINOPEC GROUP et al.), 14 October 2015 (14.10.2015), the whole document	1-12
A	CN 104731927 A (NUBIA TECHNOLOGY CO., LTD.), 24 June 2015 (24.06.2015), the whole document	1-12
A	CN 104427292 A (ZTE CORP.), 18 March 2015 (18.03.2015), the whole document	1-12
A	JP 2010060850 A (NEC CORP.), 18 March 2010 (18.03.2010), the whole document	1-12

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

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 14 September 2016 (14.09.2016)	Date of mailing of the international search report 26 September 2016 (26.09.2016)
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 CHEN, Cheng Telephone No.: (86-10) 62413594

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/089033

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102985965 A	20 March 2013	CN 102985965 B	01 April 2015
		US 2014100849 A1	10 April 2014
		US 8606579 B2	10 December 2013
		WO 2011149647 A2	01 December 2011
		TW 201142823 A	01 December 2011
		US 2011288866 A1	24 November 2011
		EP 2577655 A2	10 April 2013
		TW I536365 B	01 June 2016
CN 104978507 A	14 October 2015	None	
CN 104731927 A	24 June 2015	None	
CN 104427292 A	18 March 2015	WO 2014173370 A1	30 October 2014
		WO 2015024413 A1	26 February 2015
JP 2010060850 A	18 March 2010	None	

国际检索报告

国际申请号

PCT/CN2016/089033

A. 主题的分类		
G10L 17/00(2013.01)i; G10L 15/02(2006.01)i; G10L 15/06(2013.01)i; G10L 15/08(2006.01)i; G11B 27/10(2006.01)i		
按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
G10L17/-;G10L15/-;G11B27/-		
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))		
CNPAT, WPI, EPODOC, CNKI:声纹, 模板, 数据库, 标记, 标识, 映射, 训练, soundprint, model, database, mark+, map, train+.		
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	CN 102985965 A (微软公司) 2013年 3月 20日 (2013 - 03 - 20) 说明书第27-32段, 权利要求1-3, 附图3	1-12
A	CN 104978507 A (中国石油化工集团公司等) 2015年 10月 14日 (2015 - 10 - 14) 全文	1-12
A	CN 104731927 A (努比亚技术有限公司) 2015年 6月 24日 (2015 - 06 - 24) 全文	1-12
A	CN 104427292 A (中兴通讯股份有限公司) 2015年 3月 18日 (2015 - 03 - 18) 全文	1-12
A	JP 2010060850 A (NEC CORP.) 2010年 3月 18日 (2010 - 03 - 18) 全文	1-12
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件,或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布,与申请不相抵触,但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件,单独考虑该文件,认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件,当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时,要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2016年 9月 14日		2016年 9月 26日
ISA/CN的名称和邮寄地址		受权官员
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		陈 斌
传真号 (86-10)62019451		电话号码 (86-10)62413594

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/089033

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	102985965	A	2013年 3月 20日	CN	102985965	B	2015年 4月 1日
				US	2014100849	A1	2014年 4月 10日
				US	8606579	B2	2013年 12月 10日
				WO	2011149647	A2	2011年 12月 1日
				TW	201142823	A	2011年 12月 1日
				US	2011288866	A1	2011年 11月 24日
				EP	2577655	A2	2013年 4月 10日
				TW	1536365	B	2016年 6月 1日
CN	104978507	A	2015年 10月 14日	无			
CN	104731927	A	2015年 6月 24日	无			
CN	104427292	A	2015年 3月 18日	WO	2014173370	A1	2014年 10月 30日
				WO	2015024413	A1	2015年 2月 26日
JP	2010060850	A	2010年 3月 18日	无			