

(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4月13日 (13.04.2017)



(10) 国际公布号
WO 2017/059694 A1

- (51) 国际专利分类号:
G10L 21/013 (2013.01) G10L 21/007 (2013.01)
- (21) 国际申请号: PCT/CN2016/086229
- (22) 国际申请日: 2016年6月17日 (17.06.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510644697.0 2015年10月8日 (08.10.2015) CN
- (71) 申请人: 重庆邮电大学 (CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) [CN/CN]; 中国重庆市南岸区崇文路2号, Chongqing 400065 (CN)。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: 刘媛媛 (LIU, Yuanyuan); 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)。刘光军 (LIU, Guangjun); 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)。卢国丽 (LU, Guoli); 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)。付芬 (FU, Fen); 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)。
- (74) 代理人: 广州三环专利代理有限公司 (GUANGZHOU SCIHEAD PATENT AGENT CO., LTD); 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, 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,

[见续页]

(54) Title: SPEECH IMITATION METHOD AND DEVICE

(54) 发明名称: 一种语音模仿方法和装置

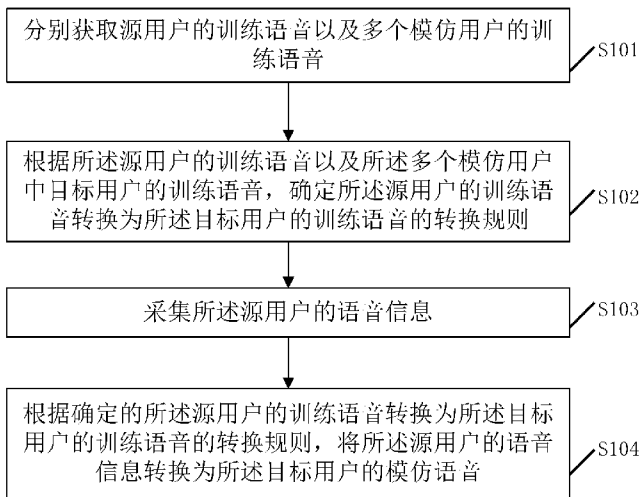


图 1

- S101 SEPARATELY OBTAIN A TRAINING SPEECH OF A SOURCE USER AND TRAINING SPEECHES OF MULTIPLE IMITATION USERS
- S102 DETERMINE, ACCORDING TO THE TRAINING SPEECH OF THE SOURCE USER AND THE TRAINING SPEECH OF A TARGET USER OF THE IMITATION USERS, A CONVERSION RULE FOR CONVERTING THE TRAINING SPEECH OF THE SOURCE USER INTO THE TRAINING SPEECH OF THE TARGET USER
- S103 ACQUIRE SPEECH INFORMATION OF THE SOURCE USER
- S104 CONVERT, ACCORDING TO THE DETERMINED CONVERSION RULE FOR CONVERTING THE TRAINING SPEECH OF THE SOURCE USER INTO THE TRAINING SPEECH OF THE TARGET USER, THE SPEECH INFORMATION OF THE SOURCE USER TO AN IMITATION SPEECH OF THE TARGET USER

(57) Abstract: A speech imitation method and device. The method comprises: separately obtaining a training speech of a source user and training speeches of multiple imitation users (S101); determining, according to the training speech of the source user and the training speech of a target user of the imitation users, a conversion rule for converting the training speech of the source user into the training speech of the target user (S102); acquiring speech information of the source user (S103); and converting, according to the determined conversion rule for converting the training speech of the source user into the training speech of the target user, the speech information of the source user to an imitation speech of the target user (S104). By generating imitation speeches of multiple imitation users from one source user, the present invention provides additional materials for simulating speeches, and improves the efficiency of establishing a speech material database.

(57) 摘要: 一种语音模仿方法和装置。该方法包括: 分别获取源用户的训练语音以及多个模仿用户的训练语音 (S101); 根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则 (S102); 采集所述源用户的语音信息 (S103); 根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音 (S104)。通过一个源用户模仿出多个模仿用户的模仿语音, 从而增加了模仿语音的素材, 提高了建立语音素材库的效率。

建立语音素材库的效率。

WO 2017/059694 A1



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。

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)。

(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,

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种语音模仿方法和装置

本专利申请要求 2015 年 10 月 8 日提交的中国专利申请号为 201510644697.0, 发明名称为“一种语音模仿方法和装置”的优先权, 该申请的全文以引用的方式并入本申请中。

5 技术领域

本申请涉及语音技术领域, 尤其涉及一种语音模仿方法和装置。

背景技术

随着科学技术的发展, 微信、QQ 等即时通信软件可以首先录制用户的语音信息, 然后将录制的语音信息匹配为音色各异的语音, 并将匹配得到的语音发送给其他用户, 从而提高用户之前交流的趣味性。然而, 在用户使用这个变音功能之前, 需要用户从网络服务器下载语音素材包, 为了能模仿更多用户的语音, 网络服务器需要提供大量的语音素材包, 因此, 需要通过语音转换技术得到语音素材包, 在语音转换过程中, 需要提取语音信息中的特征参数, 根据提取的特征参数来进行语音转化, 特征参数可以包括多个共振峰的频率和带宽等等。但是, 常见的即时通信软件中提供的语音素材少, 必然导致趣味性降低, 娱乐效果不能满足更多用户的需求, 并且下载到移动设备中的语音素材包浪费空间存储。进一步地, 常见的即时通信软件中, 将录制的语音信息匹配音色各异的语音时, 录制的语音信息通过一种语音素材匹配只能得到一种语音, 从而使得语音素材的利用率较低。

发明内容

本申请实施例提供一种语音模仿方法和装置。可以通过一个源用户模仿出多个模仿用户的模仿语音, 从而增加了模仿语音的素材, 提高了建立语音素材库的效率。

本申请第一方面提供了一种语音模仿方法, 包括:

分别获取源用户的训练语音以及多个模仿用户的训练语音;

根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语

音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则;
采集所述源用户的语音信息;

根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音。

5

在第一方面的第一种可能的实现方式中, 所述转换规则包括转换矩阵, 所述根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则包括:

10 分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数;

根据所述第一特征参数以及所述第二特征参数, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

15 结合第一方面的第一种可能的实现方式, 在第一方面的第二种可能的实现方式中, 所述根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音包括:

提取所述源用户的语音信息中的第三特征参数;

根据确定的所述转换矩阵以及提取的所述第三特征参数, 确定所述目标用户的语音信息中的第四特征参数;

20 根据所述目标用户的语音信息中的所述第四特征参数, 合成得到所述目标用户的模仿语音。

25 结合第一方面的第二种可能的实现方式, 在第一方面的第三种可能的实现方式中, 所述第四特征参数包括多个共振峰的带宽以及频率, 所述根据所述目标用户的语音信息中的所述第四特征参数, 合成得到所述目标用户的模仿语音包括:

根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率, 构建多个共振峰滤波器;

通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制;

对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

在第一方面的第四种可能的实现方式中,所述根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则,将所述源用户的语音信息转换为所述目标用户的模仿语音之后,还包括:

5 建立语音素材库,所述语音素材库包括转换得到的所述多个模仿用户的模仿语音;

将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

结合第一方面的第四种可能的实现方式,在第一方面的第五种可能的实现方式中,所述将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿之后,还包括:

获取所述其他用户的录制语音以及选择的模仿用户信息;

将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器,以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音;

15 接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

相应地,本申请第二方面提供了一种语音模仿方法,包括:

语音获取模块,用于分别获取源用户的训练语音以及多个模仿用户的训练语音;

20 规则确定模块,用于根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音,确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则;

信息采集模块,用于采集所述源用户的语音信息;

25 语音转换模块,用于根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则,将所述源用户的语音信息转换为所述目标用户的模仿语音。

在第二方面的第一种可能的实现方式中,所述转换规则包括转换矩阵,所述规则确定模块包括:

参数提取单元,用于分别提取所述源用户的训练语音中的第一特征参数以

及所述目标用户的训练语音中的第二特征参数;

矩阵确定单元,用于根据所述第一特征参数以及所述第二特征参数,确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

结合第二方面的第一种可能的实现方式,在第二方面的第二种可能的实现方式中,所述语音转换模块包括:

参数获取单元,用于提取所述源用户的语音信息中的第三特征参数;

参数确定单元,用于根据确定的所述转换矩阵以及提取的所述第三特征参数,确定所述目标用户的语音信息中的第四特征参数;

语音合成单元,用于根据所述目标用户的语音信息中的所述第四特征参数,合成得到所述目标用户的模仿语音。

结合第二方面的第二种可能的实现方式,在第二方面的第三种可能的实现方式中,所述第四特征参数包括多个共振峰的带宽以及频率,所述语音合成单元具体用于:

根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率,构建多个共振峰滤波器;

通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制;

对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

在第二方面的第四种可能的实现方式中,所述装置包括:

语音库建立模块,用于建立语音素材库,所述语音素材库包括转换得到的所述多个模仿用户的模仿语音;

语音库上传模块,用于将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

结合第二方面的第四种可能的实现方式,在第二方面的第五种可能的实现方式中,所述装置还包括语音模仿模块,具体用于:

获取所述其他用户的录制语音以及选择的模仿用户信息;

将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器,以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音;

接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

5 附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对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。

- 10 图 1 是本申请提出的一种语音模仿方法第一实施例的流程图；
图 2 是本申请实施例提供的一种语音合成结构示意图；
图 3 是本申请实施例提供的源用户的语音信息转换为多个模仿用户的模仿语音的结构示意图；
图 4 是本申请提出的一种语音模仿方法第二实施例的流程图；
15 图 5 是本申请实施例提供的一种语音模仿装置的结构示意图；
图 6 是本申请实施例提出的语音模仿装置中规则确定模块的结构示意图；
图 7 是本申请实施例提出的语音模仿装置中语音转换模块的结构示意图；
图 8 为本申请实施例提供的另一种语音模仿装置的结构示意图。

20 具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

- 25 请参考图 1，图 1 是本申请提出的一种语音模仿方法第一实施例的流程图。如图所示，本申请实施例中的方法包括：

S101，分别获取源用户的训练语音以及多个模仿用户的训练语音。

具体实现中，可以录制源用户的训练语音，并从网络服务器中下载多个模仿用户的训练语音，其中，多个模仿用户可以为动画卡通人物、明星人物等任

何用户，训练语音可以为一段短语音。

S102, 根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则。

5 具体实现中, 所述转换规则包括转换矩阵, 可以分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数; 根据所述第一特征参数以及所述第二特征参数, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。其中, 第一特征参数包括所述共振峰带宽、频率等。

10 进一步的, 首先分别对源用户的训练语音以及多个模仿用户的训练语音进行端点检测、滤波等预处理得到的源用户的训练语音数据帧以及多个模仿用户中每个模仿用户的训练语音的数据帧, 然后按照数据帧提取源用户的训练语音的前5个共振峰的带宽以及频率作为第一特征参数, 以及每个模仿用户的训练语音的前5个共振峰的带宽以及频率作为第二特征参数, 最后分别对第一特征
15 参数和第二特征参数进行动态时间归整和矢量量化, 确定第一特征参数到第二特征参数的转换矩阵, 从而可以得到源用户的训练语音转换为每个模仿用户的训练语音的转换矩阵。

S103, 采集所述源用户的语音信息。

具体实现中, 可以录制源用户的语音信息, 为了最终得到的模仿语音能够
20 满足所有用户的语音模仿需求, 录制的语音信息应该包括全部的常见用语以及一些可能不常见用语。

S104, 根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音。

具体实现中, 可以提取所述源用户的语音信息中的第三特征参数; 根据确
25 定的所述转换矩阵以及提取的所述第三特征参数, 确定所述目标用户的语音信息中的第四特征参数; 根据所述目标用户的语音信息中的所述第四特征参数, 合成得到所述目标用户的模仿语音。进一步的, 可以首先对源用户的语音信息进行端点检测、滤波等预处理, 从而提取源用户的语音信息中的第三特征参数, 然后对第三特征参数进行矢量量化, 根据经过矢量量化的第三特征参数以及所

述转换矩阵, 计算得到第四特征参数, 第四特征参数包括多个共振峰的带宽以及频率; 最后根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率, 构建多个共振峰滤波器, 通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制, 对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

5 可选的, 还可以通过 LPC (Linear Predictive Coding, 线性预测编码) 合成技术、LMA (Log Magnitude Approximate) 声道模型技术合成得到所述目标用户的模仿语音。如图 2 所示, 在共振峰合成技术中, 可以通过合成通道经过调制的所述激励信号进行合成得到所述目标用户的模仿语音, 其中, 合成通道可以为并联通道、串联通道或串并联通道。

10 例如: 如图 3 所示, 源用户的训练语音转换为模仿用户 1 的训练语音的转换矩阵 1, 源用户的训练语音转换为模仿用户 2 的训练语音的转换矩阵 2 (未全部标出), 源用户的训练语音转换为模仿用户 3 的训练语音的转换矩阵 3 (未全部标出), 源用户的语音信息为“早上好”, 则分别经过转换矩阵 1、转换矩阵 2、转换矩阵 3 等一系列处理, 最后得到模仿用户 1 的模仿语音“早上好”、模仿用户 2 的模仿语音“早上好”以及模仿用户 3 的模仿语音“早上好”。

20 在本申请实施例中, 首先分别获取源用户的训练语音以及多个模仿用户的训练语音; 然后根据源用户的训练语音以及多个模仿用户中目标用户的训练语音, 其次确定源用户的训练语音转换为目标用户的训练语音的转换规则; 采集源用户的语音信息; 最后根据确定的源用户的训练语音转换为目标用户的训练语音的转换规则, 将源用户的语音信息转换为目标用户的模仿语音, 从而通过一个源用户模仿出多个模仿用户的模仿语音, 增加了模仿语音的素材, 提高了建立语音素材库的效率。

25 请参考图 4, 图 4 是本申请提出的一种语音模仿方法第二实施例的流程图。如图所示, 本申请实施例中的方法包括:

S401, 分别获取源用户的训练语音以及多个模仿用户的训练语音。

具体实现中, 可以录制源用户的训练语音, 并从网络服务器中下载多个模仿用户的训练语音, 其中, 多个模仿用户可以为动画卡通人物、明星人物等任

何用户，训练语音可以为一段短语音。

S402, 根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则。

5 具体实现中, 所述转换规则包括转换矩阵, 可以分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数; 根据所述第一特征参数以及所述第二特征参数, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

进一步的, 首先分别对源用户的训练语音以及多个模仿用户的训练语音进行
10 端点检测、滤波等预处理得到的源用户的训练语音数据帧以及多个模仿用户中每个模仿用户的训练语音的数据帧, 然后按照数据帧提取源用户的训练语音的前 5 个共振峰的带宽以及频率作为第一特征参数, 以及每个模仿用户的训练语音的前 5 个共振峰的带宽以及频率作为第二特征参数, 最后分别对第一特征参数和
15 第二特征参数进行动态时间归整和矢量量化, 确定第一特征参数到第二特征参数的转换矩阵, 从而可以得到源用户的训练语音转换为每个模仿用户的训练语音的转换矩阵。

S403, 采集所述源用户的语音信息。

具体实现中, 可以录制源用户的语音信息, 为了最终得到的模仿语音能够满足所有用户的语音模仿需求, 录制的语音信息应该包括全部的常见用语以及
20 一些可能不常见用语。

S404, 根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音。

具体实现中, 可以提取所述源用户的语音信息中的第三特征参数; 根据确定的所述转换矩阵以及提取的所述第三特征参数, 确定所述目标用户的语音信息中的
25 第四特征参数; 根据所述目标用户的语音信息中的所述第四特征参数, 合成得到所述目标用户的模仿语音。进一步的, 可以首先对源用户的语音信息进行端点检测、滤波等预处理, 进而提取源用户的语音信息中的第三特征参数, 然后对第三特征参数进行矢量量化, 根据经过矢量量化的第三特征参数以及所述转换矩阵, 计算得到第四特征参数, 第四特征参数包括多个共振峰的带宽以

及频率;最后根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率,构建多个共振峰滤波器,通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制,对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

5 可选的,还可以通过 LPC (Linear Predictive Coding, 线性预测编码) 合成技术、LMA (Log Magnitude Approximate) 声道模型技术合成得到所述目标用户的模仿语音。如图 2 所示,在共振峰合成技术中,可以通过合成通道经过调制的所述激励信号进行合成得到所述目标用户的模仿语音,其中,合成通道可以为并联通道、串联通道或串并联通道。

10 例如:如图 3 所示,源用户的训练语音转换为模仿用户 1 的训练语音的转换矩阵 1,源用户的训练语音转换为模仿用户 2 的训练语音的转换矩阵 2 (未全部标出),源用户的训练语音转换为模仿用户 3 的训练语音的转换矩阵 3 (未全部标出),源用户的语音信息为“早上好”,则分别经过转换矩阵 1、转换矩阵 2、转换矩阵 3 等一系列处理,最后得到模仿用户 1 的模仿语音“早上好”、
15 模仿用户 2 的模仿语音“早上好”以及模仿用户 3 的模仿语音“早上好”。

S405,建立语音素材库,所述语音素材库包括转换得到的所述多个模仿用户的模仿语音。其中,可以根据用户语音模仿的需求增加所述语音素材库中模仿用户的模仿语音。

20 S406,将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

具体实现中,可以首先建立访问云储存服务器的账号和密码,然后通过验证账号和密码登陆到云存储服务器,将语音素材库上传到云存储服务器,从而可以释放装置的存储空间。

25 S407,获取所述其他用户的录制语音以及选择的模仿用户信息。其中,模仿用户信息可以为需要模仿的用户对象名称。

S408,将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器,以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音。

具体实现中,云存储服务器接收到所述其他用户的录制语音以及所述模仿

用户信息之后，首先根据所述其他用户选择的模仿用户信息，确定所述其他用户需要模仿的所述多个模仿用户中的指定用户，然后查找是否存在与所述其他用户的录制语音匹配的所述指定用户的模仿语音，如果存在，则将与所述录制语音匹配的所述指定用户的模仿语音返回给装置。

- 5 例如：其他用户通过即时通信软件获取录制语音“您好”，并选择模仿用户1作为模仿的对象，装置将录制语音“您好”以及模仿用户信息“模仿用户1”发送到云存储服务器，云存储服务器根据模仿用户信息确定该用户需要模仿“模仿用户1”，因此查找语音素材库中“模仿用户1”的模仿语音，最后根据录制语音“您好”，查找“模仿用户1”的模仿语音中的“您好”，并将查找到
- 10 “模仿用户1”的模仿语音中的“您好”返回给装置。

S409,接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

- 具体实现中，在接收到所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音之后，可以首先对与所述录制语音匹配的所述指定用户的模仿语音进行试听，若所述其他用户感觉不满意，可以放弃所述指定用户的模仿语音，进而选择其他模仿用户信息，并重新获取云存储服务器查找的其他模仿用户的模仿语音。若所述其他用户感觉满意，可以通过即时通信软件将该制定用户的模仿语音发送给通话对方，从而提高沟通的趣味性。
- 15

- 在本申请实施例中，首先分别获取源用户的训练语音以及多个模仿用户的训练语音；然后根据源用户的训练语音以及多个模仿用户中目标用户的训练语音，其次确定源用户的训练语音转换为目标用户的训练语音的转换规则；采集源用户的语音信息；最后根据确定的源用户的训练语音转换为目标用户的训练语音的转换规则，将源用户的语音信息转换为目标用户的模仿语音，并且将转换得到的多个模仿用户的模仿语音上传到云储存服务器，通过云储存服务器提供的语音素材包进行语音模仿，从而不仅可以通过一个源用户模仿出多个模仿用户的模仿语音，增加了模仿语音的素材，提高了建立语音素材库的效率，而且减小对移动设备内存的占用。
- 20
- 25

请参考图5,图5是本申请实施例提供的一种语音模仿装置的结构示意图。

如图所示，本申请实施例中的装置包括：

语音获取模块 501，用于分别获取源用户的训练语音以及多个模仿用户的训练语音。

具体实现中，可以录制源用户的训练语音，并从网络服务器中下载多个模仿用户的训练语音，其中，多个模仿用户可以为动画卡通人物、明星人物等任何用户，训练语音可以为一段短语音。

规则确定模块 502，用于根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则。

具体实现中，如图 6 所示，规则确定模块 502 还可以进一步包括：

参数提取单元 601，用于分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数。其中，第一特征参数包括所述共振峰带宽、频率等。具体的，首先分别对源用户的训练语音以及多个模仿用户的训练语音进行端点检测、滤波等预处理得到的源用户的训练语音数据帧以及多个模仿用户中每个模仿用户的训练语音的数据帧，然后按照数据帧提取源用户的训练语音的前 5 个共振峰的带宽以及频率作为第一特征参数，以及每个模仿用户的训练语音的前 5 个共振峰的带宽以及频率作为第二特征参数。

矩阵确定单元 602，用于根据所述第一特征参数以及所述第二特征参数，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

具体的，分别对第一特征参数和第二特征参数进行动态时间归整和矢量量化，确定第一特征参数到第二特征参数的转换矩阵，从而可以得到源用户的训练语音转换为每个模仿用户的训练语音的转换矩阵。

信息采集模块 503，用于采集所述源用户的语音信息。

具体实现中，可以录制源用户的语音信息，为了最终得到的模仿语音能够满足所有用户的语音模仿需求，录制的语音信息应该包括全部的常见用语以及一些可能不常见用语。

语音转换模块 504，用于根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音。

具体实现中，如图 7 所示，语音转换模块 504 还可以进一步包括：

参数获取单元 701，用于提取所述源用户的语音信息中的第三特征参数。
具体的，可以首先对源用户的语音信息进行端点检测、滤波等预处理，进而提取源用户的语音信息中的第三特征参数

5 参数确定单元 702，用于根据确定的所述转换矩阵以及提取的所述第三特征参数，确定所述目标用户的语音信息中的第四特征参数。

具体的，可以对第三特征参数进行矢量量化，根据经过矢量量化的第三特征参数以及所述转换矩阵，计算得到第四特征参数，其中，第四特征参数包括多个共振峰的带宽以及频率。

10 语音合成单元 703，用于根据所述目标用户的语音信息中的所述第四特征参数，合成得到所述目标用户的模仿语音。

具体的，可以首先根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率，构建多个共振峰滤波器，然后通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制，最后对经过调制的所述激励信号进行
15 合成得到所述目标用户的模仿语音。

可选的，还可以通过 LPC（Linear Predictive Coding，线性预测编码）合成技术、LMA（Log Magnitude Approximate）声道模型技术合成得到所述目标用户的模仿语音。如图 2 所示，在共振峰合成技术中，可以通过合成通道经过调制的所述激励信号进行合成得到所述目标用户的模仿语音，其中，合成通道
20 可以为并联通道、串联通道或串并联通道。

例如：如图 3 所示，源用户的训练语音转换为模仿用户 1 的训练语音的转换矩阵 1，源用户的训练语音转换为模仿用户 2 的训练语音的转换矩阵 2（未全部标出），源用户的训练语音转换为模仿用户 3 的训练语音的转换矩阵 3（未全部标出），源用户的语音信息为“早上好”，则分别经过转换矩阵 1、转换矩阵 2、转换矩阵 3 等一系列处理，最后得到模仿用户 1 的模仿语音“早上好”、模仿用户 2 的模仿语音“早上好”以及模仿用户 3 的模仿语音“早上好”。
25

可选的，如图 5 所示，本申请实施例中的装置还可以进一步包括：

语音库建立模块 505，用于建立语音素材库，所述语音素材库包括转换得到的所述多个模仿用户的模仿语音。其中，可以根据用户语音模仿的需求增加

所述语音素材库中模仿用户的模仿语音。

语音库上传模块 506，用于将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

具体实现中，可以首先建立访问云储存服务器的账号和密码，然后通过验证账号和密码登陆到云存储服务器，将语音素材库上传到云存储服务器，从而可以释放装置的存储空间。

可选的，如图 5 所示，本申请实施例中的装置还可以进一步包括：

语音模仿模块 507，用于获取所述其他用户的录制语音以及选择的模仿用户信息；将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器，以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音；接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

具体实现中，所述模仿用户信息可以为需要模仿的用户对象名称。云存储服务器接收到所述其他用户的录制语音以及所述模仿用户信息之后，首先根据所述其他用户选择的模仿用户信息，确定所述其他用户需要模仿的所述多个模仿用户中的指定用户，然后查找是否存在与所述其他用户的录制语音匹配的所述指定用户的模仿语音，如果存在，则将与所述录制语音匹配的所述指定用户的模仿语音返回给装置。装置在接收到所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音之后，可以首先对与所述录制语音匹配的所述指定用户的模仿语音进行试听，若所述其他用户感觉不满意，可以放弃所述指定用户的模仿语音，进而选择其他模仿用户信息，并重新获取云存储服务器查找的其他模仿用户的模仿语音。若所述其他用户感觉满意，可以通过即时通信软件将该制定用户的模仿语音发送给通话对方，从而提高沟通的趣味性。

例如：其他用户通过即时通信软件获取录制语音“您好”，并选择模仿用户 1 作为模仿的对象，装置将录制语音“您好”以及模仿用户信息“模仿用户 1”发送到云存储服务器，云存储服务器根据模仿用户信息确定该用户需要模仿“模仿用户 1”，因此查找语音素材库中“模仿用户 1”的模仿语音，最后根据录制语音“您好”，查找“模仿用户 1”的模仿语音中的“您好”，并将查找

到“模仿用户 1”的模仿语音中的“您好”返回给装置，装置可以通过 QQ、微信等即时通信软对“模仿用户 1”的模仿语音中的“您好”进行播放。

在本申请实施例中，首先分别获取源用户的训练语音以及多个模仿用户的训练语音；然后根据源用户的训练语音以及多个模仿用户中目标用户的训练语音，其次确定源用户的训练语音转换为目标用户的训练语音的转换规则；采集源用户的语音信息；最后根据确定的源用户的训练语音转换为目标用户的训练语音的转换规则，将源用户的语音信息转换为目标用户的模仿语音，并且将转换得到的多个模仿用户的模仿语音上传到云储存服务器，通过云储存服务器提供的语音素材包进行语音模仿，从而不仅可以通过一个源用户模仿出多个模仿用户的模仿语音，增加了模仿语音的素材，提高了建立语音素材库的效率，而且减小对移动设备内存的占用。

请参阅图 8，图 8 为本申请实施例提供的另一种语音模仿装置的结构示意图。如图 8 所示，所述身份验证设备可以包括至少一个处理器 801，例如，CPU，用户接口 803，存储器 804，至少一个通信总线 802。其中，通信总线 802 用于实现这些组件之间的连接通信。其中，用户接口 803 可以包括显示屏 (Display)，可选用户接口 803 还可以包括标准的有线接口、无线接口。存储器 804 可以是高速 RAM 存储器，也可以是非不稳定的存储器 (non-volatile memory)，例如至少一个磁盘存储器。存储器 804 可选的还可以是至少一个位于远离前述处理器 801 的存储装置。其中存储器 804 中存储一组程序代码，且处理器 801 调用存储器 804 中存储的程序代码，用于执行以下操作：

分别获取源用户的训练语音以及多个模仿用户的训练语音；

根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则；

采集所述源用户的语音信息；

根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音。

在可选实施例中，所述转换规则包括转换矩阵，处理器 801 调用存储器 504 中存储的程序代码执行根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音，确定所述源用户的训练语音转换为所述目标用户的

训练语音的转换规则，可以具体包括：

分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数；

5 根据所述第一特征参数以及所述第二特征参数，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

在可选实施例中，处理器 801 调用存储器 504 中存储的程序代码执行根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音，可以具体包括：

提取所述源用户的语音信息中的第三特征参数；

10 根据确定的所述转换矩阵以及提取的所述第三特征参数，确定所述目标用户的语音信息中的第四特征参数；

根据所述目标用户的语音信息中的所述第四特征参数，合成得到所述目标用户的模仿语音。

15 在可选实施例中，所述第四特征参数包括多个共振峰的带宽以及频率，处理器 801 调用存储器 504 中存储的程序代码执行根据所述目标用户的语音信息中的所述第四特征参数，合成得到所述目标用户的模仿语音，可以具体包括：

根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率，构建多个共振峰滤波器；

20 通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制；

对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

25 在可选实施例中，处理器 801 调用存储器 504 中存储的程序代码执行根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音之后，处理器 801 调用存储器 504 中存储的程序代码，还用于执行以下操作：

建立语音素材库，所述语音素材库包括转换得到的所述多个模仿用户的模仿语音；

将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

在可选实施例中，处理器 801 调用存储器 504 中存储的程序代码执行将所

述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿之后，处理器 801 调用存储器 504 中存储的程序代码，还用于执行以下操作：

获取所述其他用户的录制语音以及选择的模仿用户信息；

5 将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器，以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音；

接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

10 需要说明的是，对于前述的各个方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某一些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本申请所必须的。

15 在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中沒有详细描述的部分，可以参见其他实施例的相关描述。

20 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：闪存盘、只读存储器（英文：Read-Only Memory，简称：ROM）、随机存取器（英文：Random Access Memory，简称：RAM）、磁盘或光盘等。

25 以上对本申请实施例所提供的内容下载方法及相关设备、系统进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权利要求

1、一种语音模仿方法，其特征在于，所述方法包括：

分别获取源用户的训练语音以及多个模仿用户的训练语音；

- 5 根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则；
采集所述源用户的语音信息；

根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音。

10

2、如权利要求 1 所述方法，其特征在于，所述转换规则包括转换矩阵，所述根据所述源用户的训练语音以及所述多个模仿用户中目标用户的训练语音，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则包括：

- 15 分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数；

根据所述第一特征参数以及所述第二特征参数，确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

- 20 3、如权利要求 2 所述方法，其特征在于，所述根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音包括：

提取所述源用户的语音信息中的第三特征参数；

- 25 根据确定的所述转换矩阵以及提取的所述第三特征参数，确定所述目标用户的语音信息中的第四特征参数；

根据所述目标用户的语音信息中的所述第四特征参数，合成得到所述目标用户的模仿语音。

4、如权利要求 3 所述方法，其特征在于，所述第四特征参数包括多个共振峰的带宽以及频率，所述根据所述目标用户的语音信息中的所述第四特征参数，合成得到所述目标用户的模仿语音包括：

5 根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率，构建多个共振峰滤波器；

通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制；

对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

10 5、如权利要求 1 所述方法，其特征在于，所述根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则，将所述源用户的语音信息转换为所述目标用户的模仿语音之后，还包括：

建立语音素材库，所述语音素材库包括转换得到的所述多个模仿用户的模仿语音；

15 将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

6、如权利要求 5 所述方法，其特征在于，所述将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿之后，还包括：

获取所述其他用户的录制语音以及选择的模仿用户信息；

20 将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器，以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音；

接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

25

7、一种语音模仿装置，其特征在于，所述装置包括：

语音获取模块，用于分别获取源用户的训练语音以及多个模仿用户的训练语音；

规则确定模块，用于根据所述源用户的训练语音以及所述多个模仿用户中

目标用户的训练语音, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换规则;

信息采集模块, 用于采集所述源用户的语音信息;

5 语音转换模块, 用于根据确定的所述源用户的训练语音转换为所述目标用户的训练语音的转换规则, 将所述源用户的语音信息转换为所述目标用户的模仿语音。

8、如权利要求 7 所述装置, 其特征在于, 所述转换规则包括转换矩阵, 所述规则确定模块包括:

10 参数提取单元, 用于分别提取所述源用户的训练语音中的第一特征参数以及所述目标用户的训练语音中的第二特征参数;

矩阵确定单元, 用于根据所述第一特征参数以及所述第二特征参数, 确定所述源用户的训练语音转换为所述目标用户的训练语音的转换矩阵。

15 9、如权利要求 8 所述装置, 其特征在于, 所述语音转换模块包括:

参数获取单元, 用于提取所述源用户的语音信息中的第三特征参数;

参数确定单元, 用于根据确定的所述转换矩阵以及提取的所述第三特征参数, 确定所述目标用户的语音信息中的第四特征参数;

20 语音合成单元, 用于根据所述目标用户的语音信息中的所述第四特征参数, 合成得到所述目标用户的模仿语音。

10、如权利要求 9 所述装置, 其特征在于, 所述第四特征参数包括多个共振峰的带宽以及频率, 所述语音合成单元具体用于:

25 根据所述目标用户的语音信息中的所述多个共振峰的带宽以及频率, 构建多个共振峰滤波器;

通过构建的所述多个共振峰滤波器对预设激励源发出的激励信号进行调制;

对经过调制的所述激励信号进行合成得到所述目标用户的模仿语音。

11、如权利要求 7 所述装置，其特征在于，所述装置包括：

语音库建立模块，用于建立语音素材库，所述语音素材库包括转换得到的所述多个模仿用户的模仿语音；

5 语音库上传模块，用于将所述语音素材库上传到云存储服务器以提供给其他用户进行语音模仿。

12、如权利要求 11 所述装置，其特征在于，所述装置还包括语音模仿模块，具体用于：

获取所述其他用户的录制语音以及选择的模仿用户信息；

10 将所述其他用户的录制语音以及所述模仿用户信息发送到所述云存储服务器，以使所述云存储服务器根据所述模仿用户信息确定所述多个模仿用户中的指定用户并获取与所述录制语音匹配的所述指定用户的模仿语音；

接收所述云存储服务器发送的与所述录制语音匹配的所述指定用户的模仿语音并进行播放。

15

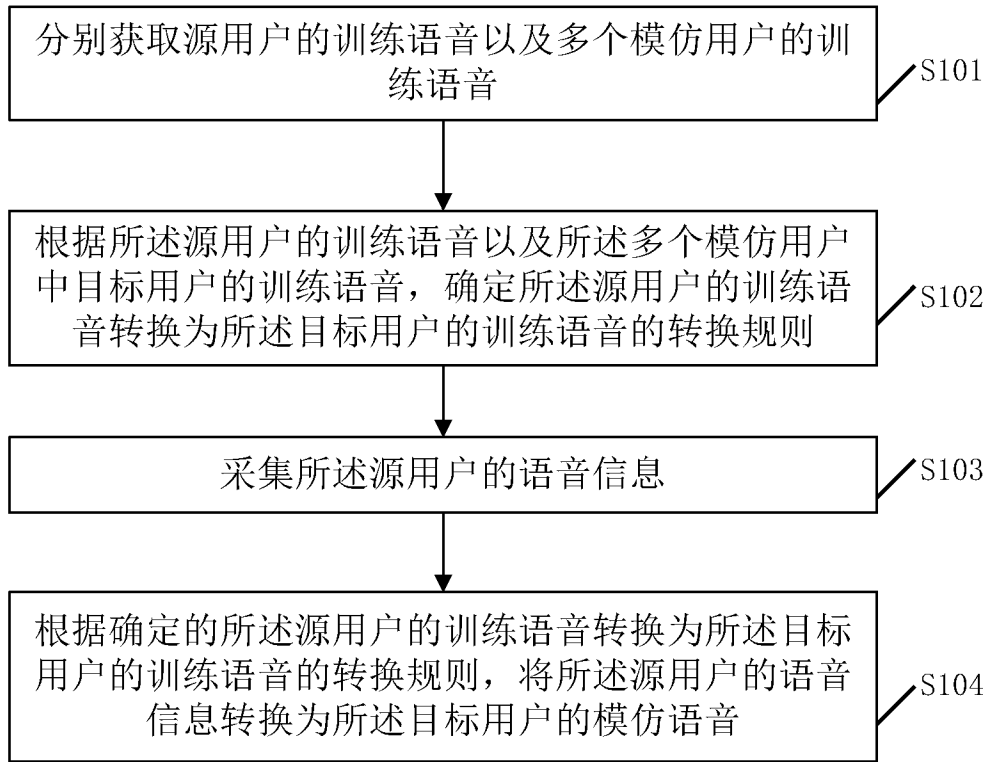


图 1

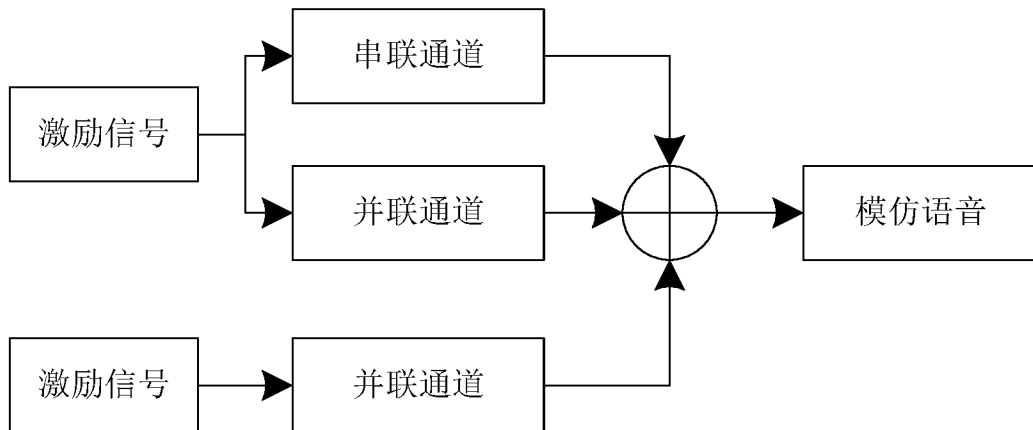


图 2

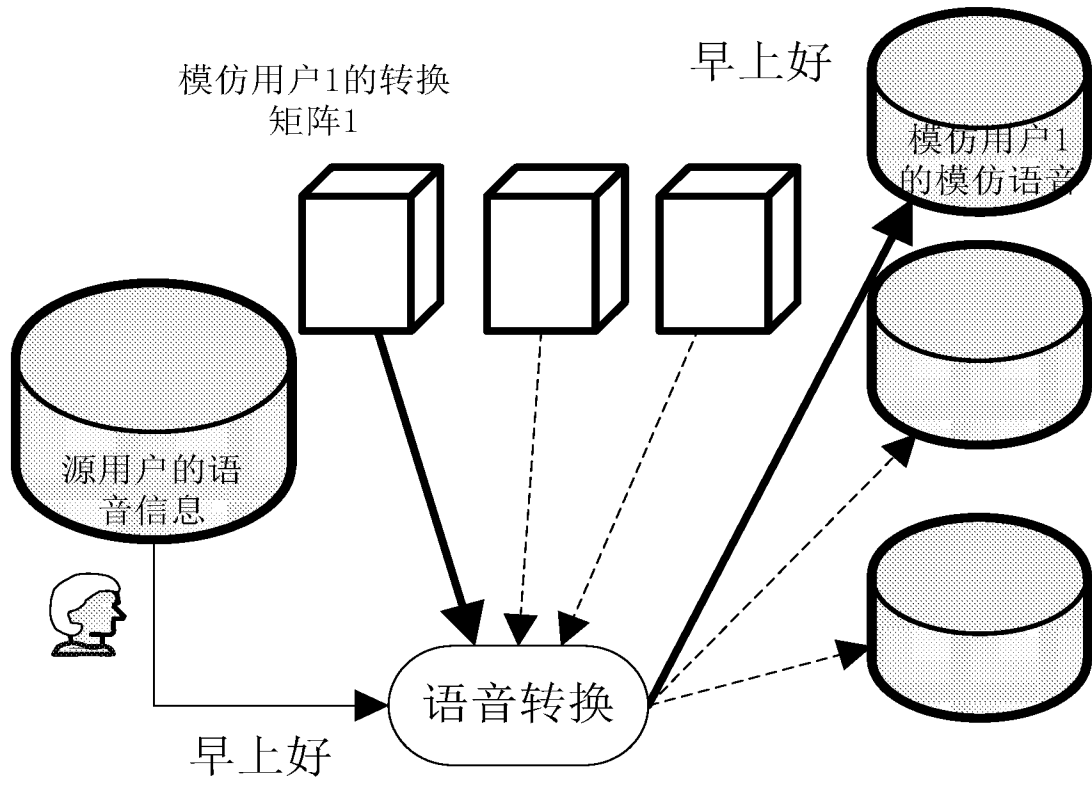


图 3

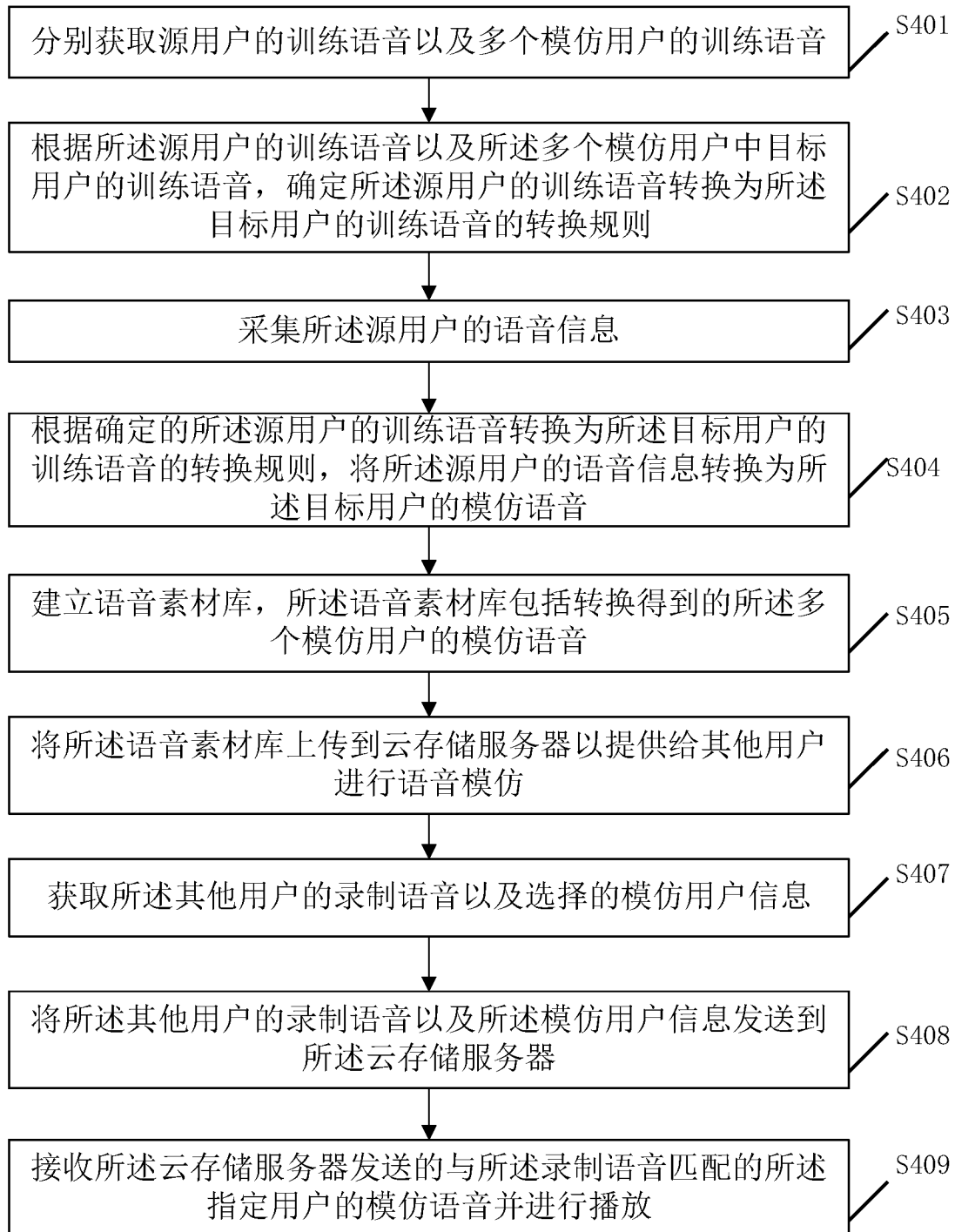


图 4

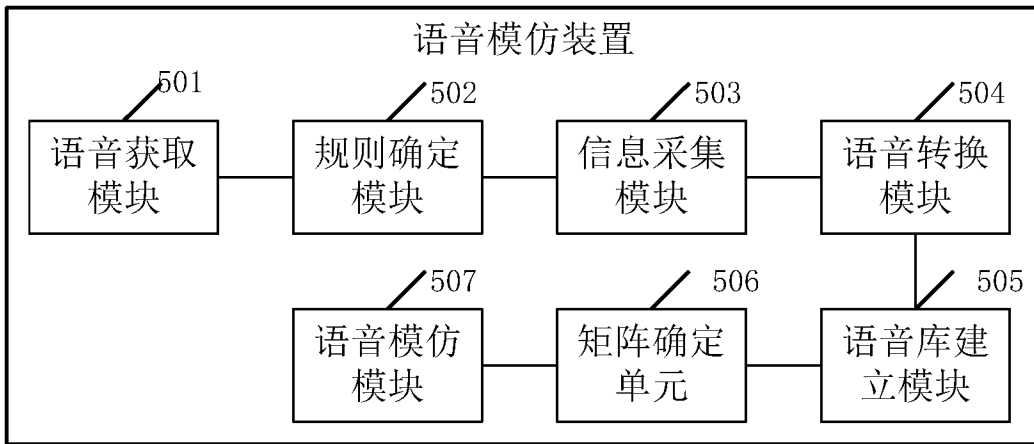


图 5

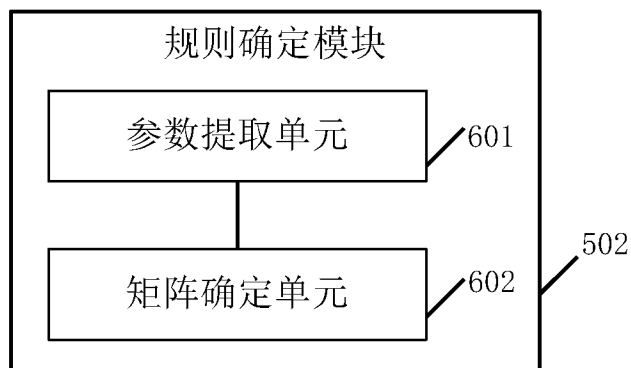


图 6

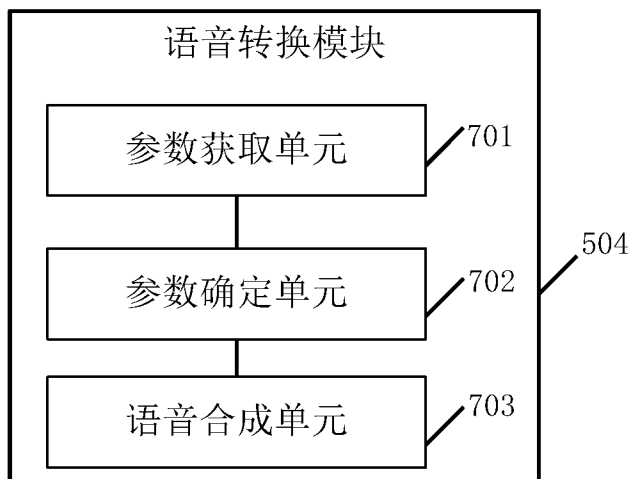


图 7

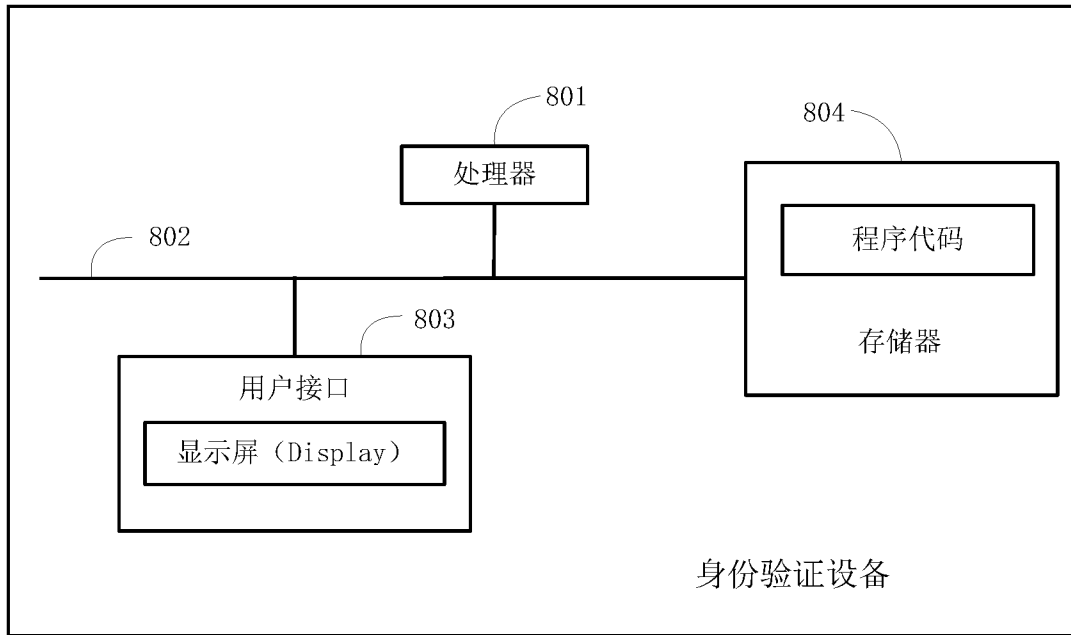


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/086229

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/013 (2013.01) i; G10L 21/007 (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 21/-; G10L 15/-;

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: audio?, speech+, voice?, sound?, convert+, conversion, simulat+, formant?, source?, target?, object?, server?, save, memory, storage

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104464725 A (FUJIAN STAR-NET E-VIDEO INFORMATION SYSTEM CO., LTD.), 25 March 2015 (25.03.2015), description, paragraphs [0066]-[0136], and figures 1-2	1-3, 7-9
Y	CN 104464725 A (FUJIAN STAR-NET E-VIDEO INFORMATION SYSTEM CO., LTD.), 25 March 2015 (25.03.2015), description, paragraphs [0066]-[0136], and figures 1-2	4-6, 10-12
Y	ZHAO, Li; "Voice Signal Processing", 31 March 2003 (31.03.2003), page 187, and chapter 8.2, paragraph 2	4, 10
Y	CN 102693729 A (BEIJING AOXINTONG SCIENCE AND TECHNOLOGY DEVELOPMENT CO., LTD.), 26 September 2012 (26.09.2012), description, paragraphs [0043]-[0050] and [0060], and figure 1	5-6, 11-12
X	KR 20020049061 A (TRUST & OBEY CO., LTD.), 26 June 2002 (26.06.2002), description, paragraphs [0004]-[0099], and figures 1-3	1-3, 7-9
X	CN 101751922 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES), 23 June 2010 (23.06.2010), description, paragraphs [0036]-[0041], and figure 1	1-3, 7-9

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
22 August 2016 (22.08.2016)

Date of mailing of the international search report
14 September 2016 (14.09.2016)

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
KANG, Dandan
Telephone No.: (86-10) **62413593**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/086229

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101004911 A (INTERNATIONAL BUSINESS MACHINES CORPORATION), 25 July 2007 (25.07.2007), description, page 6, paragraph 1 to page 7, paragraph 4 and page 12, paragraphs 3-4	1-3, 7-9
X	US 2008082320 A1 (NOKIA CORP.), 03 April 2008 (03.04.2008), description, paragraphs [0007]-[0009], and figure 9	1-3, 7-9
A	CN 104766603 A (ANHUI USTC IFLYTEK CO., LTD.), 08 July 2015 (08.07.2015), the whole document	1-12

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/086229

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104464725 A	25 March 2015	None	
CN 102693729 A	26 September 2012	CN 102693729 B	03 September 2014
KR 20020049061 A	26 June 2002	KR 351590 B	05 September 2002
CN 101751922 A	23 June 2010	CN 101751922 B	07 December 2011
CN 101004911 A	25 July 2007	US 2007185715 A1	09 August 2007
		US 8401861 B2	19 March 2013
		CN 101004911 B	27 June 2012
US 2008082320 A1	03 April 2008	None	
CN 104766603 A	08 July 2015	None	

国际检索报告

国际申请号

PCT/CN2016/086229

<p>A. 主题的分类</p> <p>G10L 21/013(2013.01)i; G10L 21/007(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>G10L 21/-; G10L 15/-;</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>CNPAT, WPI, EPODOC, CNKI: 音频, 声音, 语音, 语音, 转换, 转化, 模仿, 模拟, 共振峰, 源, 目标, 服务器, 存储, 储存, audio?, speech+, voice?, sound?, convert+, conversion, simulat+, formant?, source?, target?, object?, server?, save, memory, storage.</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 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2</td> <td>1-3, 7-9</td> </tr> <tr> <td>Y</td> <td>CN 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2</td> <td>4-6, 10-12</td> </tr> <tr> <td>Y</td> <td>赵力. 《语音信号处理》, 2003年 3月 31日 (2003 - 03 - 31), 第187页第8.2节第2段</td> <td>4, 10</td> </tr> <tr> <td>Y</td> <td>CN 102693729 A (北京奥信通科技发展有限公司) 2012年 9月 26日 (2012 - 09 - 26) 说明书第[0043]- [0050], [0060]段, 图1</td> <td>5-6, 11-12</td> </tr> <tr> <td>X</td> <td>KR 20020049061 A (TRUST & OBEY CO., LTD.) 2002年 6月 26日 (2002 - 06 - 26) 说明书第[0004]-[0099]段, 图1-3</td> <td>1-3, 7-9</td> </tr> <tr> <td>X</td> <td>CN 101751922 A (中国科学院自动化研究所) 2010年 6月 23日 (2010 - 06 - 23) 说明书第[0036]-[0041]段, 图1</td> <td>1-3, 7-9</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 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2	1-3, 7-9	Y	CN 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2	4-6, 10-12	Y	赵力. 《语音信号处理》, 2003年 3月 31日 (2003 - 03 - 31), 第187页第8.2节第2段	4, 10	Y	CN 102693729 A (北京奥信通科技发展有限公司) 2012年 9月 26日 (2012 - 09 - 26) 说明书第[0043]- [0050], [0060]段, 图1	5-6, 11-12	X	KR 20020049061 A (TRUST & OBEY CO., LTD.) 2002年 6月 26日 (2002 - 06 - 26) 说明书第[0004]-[0099]段, 图1-3	1-3, 7-9	X	CN 101751922 A (中国科学院自动化研究所) 2010年 6月 23日 (2010 - 06 - 23) 说明书第[0036]-[0041]段, 图1	1-3, 7-9
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	CN 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2	1-3, 7-9																					
Y	CN 104464725 A (福建星网视易信息系统有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0066]-[0136]段, 图1-2	4-6, 10-12																					
Y	赵力. 《语音信号处理》, 2003年 3月 31日 (2003 - 03 - 31), 第187页第8.2节第2段	4, 10																					
Y	CN 102693729 A (北京奥信通科技发展有限公司) 2012年 9月 26日 (2012 - 09 - 26) 说明书第[0043]- [0050], [0060]段, 图1	5-6, 11-12																					
X	KR 20020049061 A (TRUST & OBEY CO., LTD.) 2002年 6月 26日 (2002 - 06 - 26) 说明书第[0004]-[0099]段, 图1-3	1-3, 7-9																					
X	CN 101751922 A (中国科学院自动化研究所) 2010年 6月 23日 (2010 - 06 - 23) 说明书第[0036]-[0041]段, 图1	1-3, 7-9																					
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <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>2016年 8月 22日</p>	<p>国际检索报告邮寄日期</p> <p>2016年 9月 14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>康丹丹</p> <p>电话号码 (86-10)62413593</p>																						

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 101004911 A (国际商业机器公司) 2007年 7月 25日 (2007 - 07 - 25) 说明书第6页第1段-第7页第4段, 第12页第3-4段	1-3, 7-9
X	US 2008082320 A1 (NOKIA CORP.) 2008年 4月 3日 (2008 - 04 - 03) 说明书第[0007]-[0009]段, 图9	1-3, 7-9
A	CN 104766603 A (安徽科大讯飞信息科技股份有限公司) 2015年 7月 8日 (2015 - 07 - 08) 全文	1-12

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/086229

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	104464725	A	2015年 3月 25日	无			
CN	102693729	A	2012年 9月 26日	CN	102693729	B	2014年 9月 3日
KR	20020049061	A	2002年 6月 26日	KR	351590	B	2002年 9月 5日
CN	101751922	A	2010年 6月 23日	CN	101751922	B	2011年 12月 7日
CN	101004911	A	2007年 7月 25日	US	2007185715	A1	2007年 8月 9日
				US	8401861	B2	2013年 3月 19日
				CN	101004911	B	2012年 6月 27日
US	2008082320	A1	2008年 4月 3日	无			
CN	104766603	A	2015年 7月 8日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)