

(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
2020 年 1 月 9 日 (09.01.2020)



(10) 国际公布号
WO 2020/006879 A1

(51) 国际专利分类号:
G10L 15/06 (2013.01)

(21) 国际申请号: PCT/CN2018/106259

(22) 国际申请日: 2018 年 9 月 18 日 (18.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810735639.2 2018 年 7 月 6 日 (06.07.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 岳鹏昱(YUE, Pengyu); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 闫冬(YAN, Dong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

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

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) Title: RECORDING FILE PROCESSING METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 录音文件处理方法、装置、计算机设备及存储介质

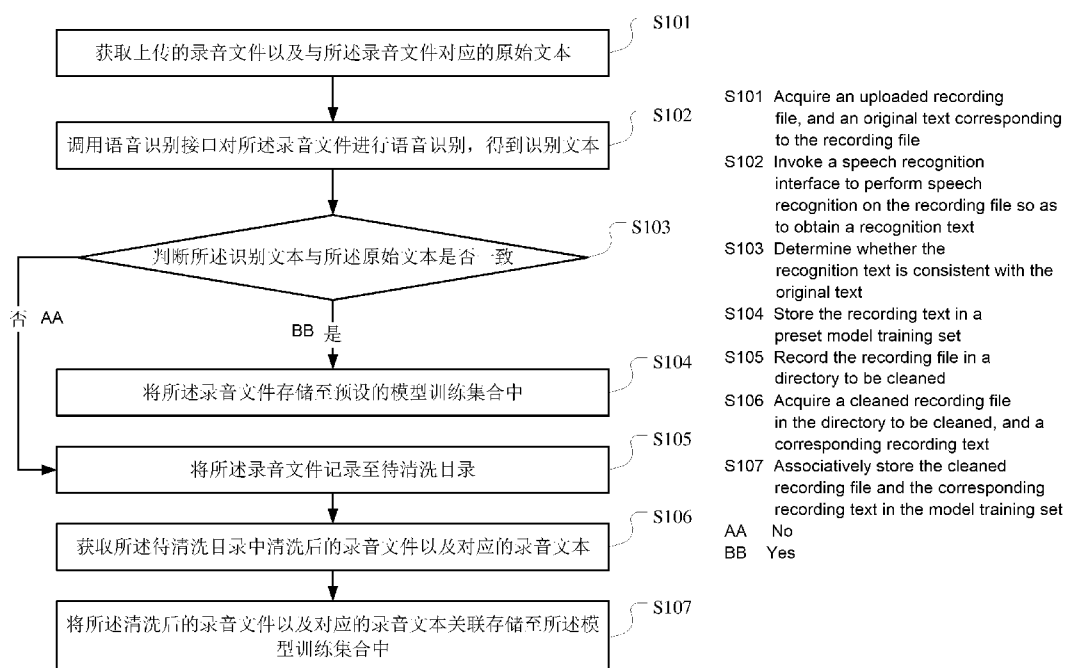


图 2

(57) Abstract: Disclosed are a recording file processing method and apparatus, and a computer device and a storage medium. The method comprises: acquiring an uploaded recording file, and an original text corresponding to the recording file (S101); invoking a speech recognition interface to perform speech recognition on the recording file so as to obtain a recognition text (S102); determining whether the recognition text is consistent with the original text (S103); if so, storing the recording file in a preset model training set (S104); if not, recording the recording file in a directory to be cleaned, wherein a recording file recorded in the directory to be

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

cleaned is listened to by a processing person, and the correct recording text is fed back (S105); acquiring a cleaned recording file in the directory to be cleaned, and a corresponding recording text (S106); and associatively storing the cleaned recording file and the corresponding recording text in the model training set (S107). The method can solve the problems of a low cleaning operation efficiency and nonstandard consolidation of a recording file.

(57) 摘要: 一种录音文件处理方法、装置、计算机设备及存储介质, 该方法包括: 获取上传的录音文件以及与录音文件对应的原始文本 (S101); 调用语音识别接口对录音文件进行语音识别, 得到识别文本 (S102); 判断识别文本与原始文本是否一致 (S103); 若一致, 则将录音文件存储至预设的模型训练集合中 (S104); 若不一致, 则将录音文件记录至待清洗目录待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本 (S105); 获取待清洗目录中清洗后的录音文件以及对应的录音文本 (S106); 将清洗后的录音文件以及对应的录音文本关联存储至模型训练集合中 (S107)。该方法能够解决录音文件的清洗操作效率低下且整理不规范的问题。

录音文件处理方法、装置、计算机设备及存储介质

本申请以 2018 年 07 月 06 日提交的申请号为 201810735639.2，名称为“录音文件处理方法、装置、计算机设备及存储介质”的中国发明专利申请为基础，并要求其优先权。

技术领域

本申请涉及语音识别技术领域，尤其涉及一种录音文件处理方法、装置、计算机设备及存储介质。

背景技术

目前，语音识别技术已经应用极为广泛，很多平台预先训练有语音识别模型，可以对外提供语音识别服务。平台在训练语音识别模型时，需要收集大量的语音文件作为样本提供给语音识别模型来学习和训练，因此，很多平台发布自己的手机 APP 客户端，鼓励广大用户通过该 APP 上传录音作为训练样本。这种方式能够迅速、高效地收集到极具多样性的训练样本，但是也存在录音文件不规范的问题。虽然平台收集的这些录音文件均是通过平台提供原始文本，用户念该原始文件而得到的，可是不同用户情况不同，并且用户不受平台的约束，导致部分用户念出来的内容与原始文本不一致，这种情况下得到的录音文件若提供给语音识别模型训练，不仅无法得到预期训练效果，反而可能减缓语音识别模型的训练进度，严重地甚至损坏语音识别模型。为此，通过上述方式得到的录音文件作为训练样本时，需要在训练之前进行清洗操作。

当前，录音文件的清洗操作均由处理人逐个人工处理，不仅效率低下，而且容易出现录音文件清洗后整理不规范、文件目录不规整等问题。

发明内容

基于此，有必要针对上述技术问题，提供一种可以提高了处理人员清洗录音文本的效率且便于对录音文件的管理和使用的录音文件处理方法、装置、计算机设备及存储介质。

一种录音文件处理方法，包括：

获取上传的录音文件以及与所述录音文件对应的原始文本；

调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

判断所述识别文本与所述原始文本是否一致；

若所述识别文本与所述原始文本一致，则将所述录音文件存储至预设的模型训练集合中；

若所述识别文本与所述原始文本不一致，则将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

一种录音文件处理装置，包括：

录音文件获取模块，用于获取上传的录音文件以及与所述录音文件对应的原始文本；

语音识别模块，用于调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

文本判断模块，用于判断所述识别文本与所述原始文本是否一致；

第一存储模块，用于若所述文本判断模块的判断结果为是，则将所述录音文件存储至预设的模型训练集合中；

文件记录模块，用于若所述文本判断模块的判断结果为否，则将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

清洗后文件获取模块，用于获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

第二存储模块，用于将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现上述录音文件处理方法的步骤。

一个或多个存储有计算机可读指令的非易失性可读存储介质，所述计算机可读存储介质存储有计算机可读指令，使得所述一个或多个处理器执行上述录音文件处理方法的步骤。

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

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述中所需

要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请一实施例中录音文件处理方法的一应用环境示意图；

图 2 是本申请一实施例中录音文件处理方法的一流程图；

图 3 是本申请一实施例中录音文件处理方法步骤 S103 在一个应用场景下的流程示意图；

图 4 是本申请一实施例中录音文件处理方法在一个应用场景下标记清洗时原始文本中重点关注内容的流程示意图；

图 5 是本申请一实施例中录音文件处理方法步骤 S105 在一个应用场景下的流程示意图；

图 6 是本申请一实施例中录音文件处理方法在一个应用场景下筛选优质录音账户的流程示意图；

图 7 是本申请一实施例中录音文件处理装置的结构示意图；

图 8 是本申请一实施例中计算机设备的一示意图。

具体实施方式

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

本申请提供的录音文件处理方法，可应用在如图 1 的应用环境中，其中，客户端通过网络与服务器进行通信。其中，该客户端可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备。服务器可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在一实施例中，如图 2 所示，提供一种录音文件处理方法，以该方法应用在图 1 中的服务器为例进行说明，包括如下步骤：

S101、获取上传的录音文件以及与所述录音文件对应的原始文本；

本实施例中，用户可以使用 APP 客户端录制各个录音文件，APP 客户端通过网络与语音清洗平台的服务器对接，在录制完成后自动将各个录音文件上传至该语音清洗平台

上，从而语音清洗平台可以获取到这些录音文件。

另外，用户录制这些录音文件时，是对照 APP 客户端提供的原始文本念的，可以认为该原始文本即录音文件的标准文本，在语音清洗平台获取录音文件时，应一并获取与所述录音文件对应的原始文本。

S102、调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

在获取到上传的录音文件之后，可以调用该语音清洗平台的语音识别接口对该录音文件进行语音识别，得到识别文本。可以理解的是，本实施例中的语音清洗平台可以自带语音识别功能，也可以通过调用其它平台的语音识别接口来对这些录音文件进行语音识别，完成录音的音转字，得到这些录音文件识别后的识别文本。可知，由于识别文本是采用语音识别技术得到的，因此识别文本可能与原始文本相同，也可能不同。当识别文本与原始文本不相同，造成不相同的原因至少存在三种可能性：1、用户录制录音文件时未按照原始文本念读；2、用户发音不准确；3、语音识别存在误差。这也辅证了，在采用录音文件作为样本训练语音识别模型之前，对录音文件进行清洗是必要的。

S103、判断所述识别文本与所述原始文本是否一致，若是，则执行步骤 S104，若否，则执行步骤 S105；

本实施例中，为了减少处理人员清洗录音文件的工作量，提高清洗效率，在清洗录音文件之前，先判断哪些录音文件的文本内容是与原始文本一致的。因此，可以判断通过语音识别得到的识别文本是否与所述原始文本一致，如果两者一致，则说明了该录音文件录音准确，可以直接用作训练的样本，无需处理人员进行清洗；反之，如果两者不一致，则说明了该录音文件录音不准确，不可以直接用作训练的样本，需要处理人员进行清洗，因此执行步骤 S105 完成清洗操作。

进一步地，本实施例可以在判断录音文件是否需要清洗的同时计算该录音文件对应的识别文本的错词率，以供后续分析使用。具体地，如图 3 所示，所述步骤 S103 包括：

S201、计算所述识别文本相对所述原始文本的错词率；

S202、判断所述错词率是否为 0，若是，则执行步骤 S203，若否，则执行步骤 S204；

S203、确定所述识别文本与所述原始文本一致；

S204、确定所述识别文本与所述原始文本不一致。

对于步骤 S201，可以将所述识别文本与所述原始文本进行比对，计算识别文本相对原始文本的错词率（word error rate）。可知，错词率越高，代表该识别文本与原始文本差异越大；错词率越低，代表该识别文本与原始文本差异越小，当错词率为 0 时，表示该识

别文本与原始文本是一致的。

对于步骤 S202~S204，可知，如果错词率为 0，则表示所述识别文本与所述原始文本一致，然后执行步骤 S104；反之，如果错词率不为 0，则表示所述识别文本与所述原始文本不一致，然后执行步骤 S105。

S104、将所述录音文件存储至预设的模型训练集合中；

对于步骤 104，可以理解的是，在确定所述识别文本与所述原始文本一致之后，可以认为该录音文件是准确的，可以直接用于语音识别模型的训练，因此，可以将所述录音文件存储至预设的模型训练集合中。具体地，可以将该录音文件存储至指定的数据库中，也可以将该录音文件记录在指定的训练文件目录下，在需要样本训练时，查找该训练文件目录即可获取到这些用作训练样本的录音文件。

S105、将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

对于步骤 105，可以理解的是，在确定所述识别文本与所述原始文本不一致之后，代表该录音文件需要进行清洗操作，因此将该录音文件记录至待清洗目录。本实施例中，该待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本，即，当处理人员清洗录音文件时，先查询待清洗目录，得知哪些录音文件需要进行清洗操作，然后获取这些需要清洗的录音文件。接着，处理人员可以直接通过语音清洗平台上内置的播放器播放这些录音文件，处理人员通过该播放器可以对录音文件进行“加速播放”、“减速播放”、“快进”、“后退”、“暂停”等操作，语音清洗平台提供这些功能极大方便了处理人员对录音文件的听取，处理人员听取录音文件的音频后，将听到的文字录入至语音清洗平台上，这些录入的文本即可认为是与该录音文件对应的正确的文本，即所述录音文本。

进一步地，如图 4 所示，在执行下述步骤 S106 之前，本实施例中的录音文件处理方法还可以包括：

S301、获取待清洗的录音文件和对应的原始文本；

S302、将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别，得到各个平台反馈的各个平台识别文本；

S303、将所述原始文本分别与各个所述平台识别文本进行对比，确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容；

S304、标注所述原始文本中除所述部分文本内容以外的文本内容；

S305、将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

需要说明的是，清洗录音文件虽然主要依赖处理人员的人工处理，但是，很多录音文件不仅文本内容繁多，而且通过 APP 客户端收集得到的录音文件数量极其庞大，绝大部分录音文件都是需要进行清洗操作后才可用作训练样本的。这就导致处理人员工作量过大，往往出现人手不足的情况。为了减轻这种状况、提高录音文件的清洗效率，本实施例通过上述步骤 S301~S305 的处理，在处理人员人工处理之前，预先在待清洗的录音文件所对应的原始文本上标注出处理人员需要重点关注的内容，辅助处理人员针对性地听取与核对录音文件、原始文本，有效提高处理人员清洗该录音文件的效率。

对于步骤 S301，首先，可以查询待清洗目录，然后获取待清洗的录音文件，同时，获取该待清洗的录音文件所对应的原始文本。

对于步骤 S302，可以理解的是，由于各个不同平台采用的语音识别模型往往各不相同，因此，不同平台对同一录音文件进行语音识别得到的识别结果也存在差异。本实施例通过利用不同平台对同一待清洗的录音文件进行语音识别，得到各个平台反馈的各个平台识别文本，再依据这些平台识别文本与原始文本的比较来确定出录音文件对应的文本内容中无需重点关注的内容。

对于步骤 S303，在得到各个平台反馈的各个平台识别文本之后，可以将所述原始文本分别与各个所述平台识别文本进行对比，确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容。可以理解的是，如果该原始文本中的部分文本内容均在所述各个平台识别文本中出现，则可以确定的是该录音文本对应该部分文本内容的音频是准确的，从而处理人员无需在关注这部分文本内容所对应的音频。从另一侧面也可以发现，除了该部分文本内容以外，原始文本中其它文本内容所对应的音频部分很可能存在需要清洗、修正的部分，这就需要提醒处理人员重点关注了。

对于步骤 S304 和 S305，如上所述，该原始文本中的所述部分文本内容所对应的音频部分可以认为是无需清洗的，也即处理人员无需重点关注的地方，因此，将该部分内容以外的其它文本内容进行重点标注，增加加粗、下划线、添加底色等突出显示效果，可以使得处理人员在清洗处理该录音文件时，处理人员一边听录音文件一边查看该录音文件所对应的原始文本，并重点关注那些标注出来的文本内容，从而处理人员可以集中注意力在重点的地方，无疑提高了处理人员清洗这些录音文件的效率，更容易发现“错词”所处的位置并修正，使得处理人员更容易一次听取即可完成一个录音文件的清洗工作。

S106、获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

其中，该录音文本是指清洗后经处理人员确定的文本，可以认为该录音文本与录音文件中的音频内容是一致的。处理人员在清洗录音文件之后，可以在语音清洗平台上在原始文本或识别文本的基础上修正为录音文本，提交后语音清洗平台即可获得清洗后该录音文件对应的录音文本。

S107、将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

可以理解的是，清洗后的录音文件已经可以被用于语音识别模型的训练了，因此可以将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

另外，需要说明的是，上述待清洗目录在记录录音文件时，可以对录音文件进行分类，一般来说，由于不同语音识别模型往往按照应用领域来区分，比如财经、新闻、体育、电影对话等，因此，该待清洗目录也可以根据应用领域的不同分类记录待清洗的录音文件。进一步地，如图 5 所示，所述步骤 S105 可以包括：

S401、获取所述录音文件初定的应用领域；

S402、将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上；

在上述步骤 S401 和 S402 的基础上，所述步骤 S106 具体为：获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域，所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域；并且，步骤 S107 具体可以包括下述步骤 S501-S503：

S501、判断所述录音文件的第一应用领域与初定的应用领域是否一致；

S502、若所述录音文件的第一应用领域与初定的应用领域一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上；

S503、若所述录音文件的第一应用领域与初定的应用领域不一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

对于步骤 S401 和 S402，可以理解的是，一个录音文件的应用领域可以根据该录音文件对应的原始文本所确定，由于原始文本是语音清洗平台一方提供的，因此应用领域可以预先确定并获取到。特别需要注意的是，若用户在录制录音文件时未按照原始文本来录音，那么该录音文件必然需要清洗操作，处理人员在对该录音文件清洗时，会听取该录音文件，

然后重新确定该录音文件的应用领域。而关于待清洗目录，该待清洗目录中可以划分多个不同的位置分别记录属于不同应用领域的录音文件的信息，这样更方便对录音文件的管理和应用，在处理人员需要集中清洗某一应用领域的录音文件时，可以从待清洗目录的相应位置快速检索出属于该应用领域的录音文件。

为此，在上述步骤 S401 和 S402 的基础上，所述步骤 S106 具体可以为：获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域，所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域。也即，处理人员在清洗该录音文件时，听取该录音文件后，不仅反馈听取得到的录音文本，而且人为确定出新的应用领域，即该第一应用领域。此时，在步骤 S107 将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中，为了在将清洗后的录音文件归集到模型训练集合中时更加分类清洗，本方法还可以加入上述步骤 S501-S503，先判断该录音文件的第一应用领域与初定的应用领域是否一致，若一致，则表示该录音文件清洗前的初定的应用领域是准确的，因此在归集清洗后的该录音文件时，可以将其和对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上；反之，若不一致，则表示该录音文件清洗前的初定的应用领域不准确，应当以处理人员清洗时一并确定的新的第一应用领域为准，从而将其和对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。可见，通过对清洗后的录音文件进行应用领域上的划分，有助于当使用这些清洗后的录音文件作为样本训练语音识别模型时，可以快速地、有针对性地检索出目标应用领域的录音文件来训练语音识别模型。比如当前准备训练“汽车”领域的语音识别模型，可以在该模型训练集合中“汽车”领域所属的位置上获取各个录音文件作为训练样本。

进一步地，本实施例中，由于服务器通过客户端收集用户录制的录音文件，不同用户录制的录音文件之间会存在显著的差异，有些用户录制的录音文件质量更好，比如错词率很低甚至没有错词，有些用户录制的录音文件质量则较差。显然，考虑到成本问题，录音质量好的用户更受服务器的欢迎。因此，为了从广大用户全体中筛选去录音质量稳定、优质的用户，如图 6 所示，本方法还可以包括：

S501、统计目标账户历史上传的各个录音文件的错词率；

S502、根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率；

S503、判断所述目标账户的平均错词率是否小于预设阈值，若是，则执行步骤 S504，

若否，则执行步骤 S505；

S504、将所述目标账户确定为优质录音账户，优质录音账户上传录音文件时受到预设激励机制的奖励；

S505、按照预设的流程处理。

对于步骤 S501，由上述步骤 S201 可知，服务器在判断录音文件是否需要清洗的同时可以计算录音文件对应的识别文本的错词率，因此当需要时，服务器可以很容易的统计得到各个用户账户历史上传过的各个录音文件以及这些录音文件的错词率。

对于步骤 S502，可以理解的是，该目标账户的平均错词率为该目标账户历史上传过的各个录音文件的错词率的平均值。举例说明，假设目标账户 A 历史上传过 3 个录音文件，这 3 个录音文件的错词率分别为 0.2、0.1 和 0.3，则计算得到该目标账户 A 的平均错词率为 $(0.2+0.1+0.3)/3=0.2$ 。

对于步骤 S503-S505，本实施例中，若某个用户账户的平均错词率低于某个预设阈值，可以认为该用户录制的录音文件是优质的，该用户为服务器所欢迎的用户。其中，该预设阈值可以根据实际使用情况来确定，比如可以确定为 0.1，即 10%，当该目标账户的平均错词率小于 10% 时，可以认为该目标账户为优质用户的录音账户，可以将其确定为优质录音账户，为了鼓励该优质录音账户的用户积极上传录音文件，服务器会在该优质录音账户上传录音文件时或者上传后按照预设的激励机制给予其奖励，比如可以提高该优质录音账户的权限、赠送账户系统的积分、派送小礼品等。反之，当该目标账户的平均错词率大于或等于 10% 时，可以认为该目标账户并非值得关注的优质用户的录音账户，因此可以按照预设的流程处理。这里说的预设的流程具体可以是不将所述目标账户确定为优质录音账户。

另外，本实施例中的服务器还可以查看各个处理人员的清洗工作状况，比如已清洗录音文件的数量、待清洗的录音文件数量，等等

本实施例中，通过对上传的录音文件进行语音识别，可以在处理人员清洗录音文件之前先判断识别出来的识别文本是否与原始文本一致，对于识别文本与原始文本一致的录音文件无需提供给处理人员清洗并直接存储至模型训练集合中，只有识别文本与原始文本不一致的录音文件才会记录到待清洗目录中由处理人员清洗，这样可以节省处理人员一部分的清洗工作，提高了处理人员清洗录音文本的效率；并且，无需清洗的录音文本和清洗后的录音文本最终均存储至模型训练集合中，便于对录音文件的管理和使用，方便了后续语音识别模型使用这些录音文本作为样本进行训练。

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

在一实施例中，提供一种录音文件处理装置，该录音文件处理装置与上述实施例中录音文件处理方法一一对应。如图7所示，该录音文件处理装置包括录音文件获取模块601、语音识别模块602、文本判断模块603、第一存储模块604、文件记录模块605、清洗后文件获取模块606和第二存储模块607。各功能模块详细说明如下：

录音文件获取模块601，用于获取上传的录音文件以及与所述录音文件对应的原始文本；

语音识别模块602，用于调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

文本判断模块603，用于判断所述识别文本与所述原始文本是否一致；

第一存储模块604，用于若所述文本判断模块的判断结果为是，则将所述录音文件存储至预设的模型训练集合中；

文件记录模块605，用于若所述文本判断模块的判断结果为否，则将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

清洗后文件获取模块606，用于获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

第二存储模块607，用于将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

进一步地，所述文本判断模块可以包括：

错词率计算单元，用于计算所述识别文本相对所述原始文本的错词率；

错词率判断单元，用于判断所述错词率是否为0；

第一确定单元，用于若所述错词率判断单元的判断结果为是，则确定所述识别文本与所述原始文本一致；

第二确定单元，用于若所述错词率判断单元的判断结果为否，则确定所述识别文本与所述原始文本不一致。

进一步地，所述文件记录模块可以包括：

初定领域获取单元，用于获取所述录音文件初定的应用领域；

第一记录单元,用于将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上;

所述清洗后文件获取模块具体用于:获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域,所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域;

所述第二存储模块具体可以包括:

领域判断单元,用于判断所述录音文件的第一应用领域与初定的应用领域是否一致;

第一存储单元,用于若所述领域判断单元的判断结果为是,则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上;

第二存储单元,用于若所述领域判断单元的判断结果为否,则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

进一步地,所述录音文件处理装置还可以包括:

文件文本获取模块,用于获取待清洗的录音文件和对应的原始文本;

平台识别模块,用于将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别,得到各个平台反馈的各个平台识别文本;

平台识别文本对比模块,用于将所述原始文本分别与各个所述平台识别文本进行对比,确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容;

文本标注模块,用于标注所述原始文本中除所述部分文本内容以外的文本内容;

清洗发送模块,用于将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

进一步地,所述录音文件处理装置还可以包括:

统计模块,用于统计目标账户历史上传的各个录音文件的错词率;

平均错词率计算模块,用于根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率;

平均错词率判断模块,用于判断所述目标账户的平均错词率是否小于预设阈值;

优质账户确定模块,用于若所述平均错词率判断模块的判断结果为是,则将所述目标账户确定为优质录音账户,优质录音账户上传录音文件时受到预设激励机制的奖励。

关于录音文件处理装置的具体限定可以参见上文中对于录音文件处理方法的限定,在

此不再赘述。上述录音文件处理装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 8 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储录音文件处理方法中涉及到的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种录音文件处理方法。

在一个实施例中，提供了一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现上述实施例中录音文件处理方法的步骤，例如图 2 所示的步骤 S101 至步骤 S107。或者，处理器执行计算机可读指令时实现上述实施例中录音文件处理装置的各模块/单元的功能，例如图 7 所示模块 601 至模块 607 的功能。为避免重复，这里不再赘述。

在一个实施例中，提供了一种计算机可读存储介质，该一个或多个存储有计算机可读指令的非易失性可读存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行计算机可读指令时实现上述方法实施例中录音文件处理方法的步骤，或者，该一个或多个存储有计算机可读指令的非易失性可读存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行计算机可读指令时实现上述装置实施例中录音文件处理装置中各模块/单元的功能。为避免重复，这里不再赘述。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程 ROM（PROM）、电可编程 ROM（EPROM）、电可擦除可编程 ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存

存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM（SRAM）、动态 RAM（DRAM）、同步 DRAM（SDRAM）、双数据率 SDRAM（DDRSDRAM）、增强型 SDRAM（ESDRAM）、同步链路（Synchlink） DRAM（SLDRAM）、存储器总线（Rambus）直接 RAM（RDRAM）、直接存储器总线动态 RAM（DRDRAM）、以及存储器总线动态 RAM（RDRAM）等。

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

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

权 利 要 求 书

1. 一种录音文件处理方法，其特征在于，包括：

获取上传的录音文件以及与所述录音文件对应的原始文本；

调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

判断所述识别文本与所述原始文本是否一致；

若所述识别文本与所述原始文本一致，则将所述录音文件存储至预设的模型训练集合中；

若所述识别文本与所述原始文本不一致，则将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

2. 根据权利要求 1 所述的录音文件处理方法，其特征在于，所述判断所述识别文本与所述原始文本是否一致包括：

计算所述识别文本相对所述原始文本的错词率；

判断所述错词率是否为 0；

若所述错词率为 0，则确定所述识别文本与所述原始文本一致；

若所述错词率不为 0，则确定所述识别文本与所述原始文本不一致。

3. 根据权利要求 1 所述的录音文件处理方法，其特征在于，所述将所述录音文件记录至待清洗目录包括：

获取所述录音文件初定的应用领域；

将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上；

所述获取所述待清洗目录中清洗后的录音文件以及对应的录音文本具体为：获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域，所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域；

所述将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中具体包括：

判断所述录音文件的第一应用领域与初定的应用领域是否一致；

若所述录音文件的第一应用领域与初定的应用领域一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上；

若所述录音文件的第一应用领域与初定的应用领域不一致,则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

4. 根据权利要求 1 所述的录音文件处理方法,其特征在于,在获取所述待清洗目录中清洗后的录音文件以及对应的录音文本之前,还包括:

获取待清洗的录音文件和对应的原始文本;

将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别,得到各个平台反馈的各个平台识别文本;

将所述原始文本分别与各个所述平台识别文本进行对比,确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容;

标注所述原始文本中除所述部分文本内容以外的文本内容;

将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

5. 根据权利要求 1 至 4 中任一项所述的录音文件处理方法,其特征在于,所述录音文件处理方法还包括:

统计目标账户历史上传的各个录音文件的错词率;

根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率;

判断所述目标账户的平均错词率是否小于预设阈值;

若所述目标账户的平均错词率小于预设阈值,则将所述目标账户确定为优质录音账户,优质录音账户上传录音文件时受到预设激励机制的奖励。

6. 一种录音文件处理装置,其特征在于,包括:

录音文件获取模块,用于获取上传的录音文件以及与所述录音文件对应的原始文本;

语音识别模块,用于调用语音识别接口对所述录音文件进行语音识别,得到识别文本;

文本判断模块,用于判断所述识别文本与所述原始文本是否一致;

第一存储模块,用于若所述文本判断模块的判断结果为是,则将所述录音文件存储至预设的模型训练集合中;

文件记录模块,用于若所述文本判断模块的判断结果为否,则将所述录音文件记录至待清洗目录,所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本;

清洗后文件获取模块,用于获取所述待清洗目录中清洗后的录音文件以及对应的录音文本;

第二存储模块,用于将所述清洗后的录音文件以及对应的录音文本关联存储至所述模

型训练集合中。

7. 根据权利要求 6 所述的录音文件处理装置, 其特征在于, 所述文本判断模块包括: 错词率计算单元, 用于计算所述识别文本相对所述原始文本的错词率;

错词率判断单元, 用于判断所述错词率是否为 0;

第一确定单元, 用于若所述错词率判断单元的判断结果为是, 则确定所述识别文本与所述原始文本一致;

第二确定单元, 用于若所述错词率判断单元的判断结果为否, 则确定所述识别文本与所述原始文本不一致。

8. 根据权利要求 6 所述的录音文件处理装置, 其特征在于, 所述文件记录模块包括: 初定领域获取单元, 用于获取所述录音文件初定的应用领域;

第一记录单元, 用于将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上;

所述清洗后文件获取模块具体用于: 获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域, 所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域;

所述第二存储模块具体包括:

领域判断单元, 用于判断所述录音文件的第一应用领域与初定的应用领域是否一致;

第一存储单元, 用于若所述领域判断单元的判断结果为是, 则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上;

第二存储单元, 用于若所述领域判断单元的判断结果为否, 则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

9. 根据权利要求 6 所述的录音文件处理装置, 其特征在于, 所述录音文件处理装置还包括:

文件文本获取模块, 用于获取待清洗的录音文件和对应的原始文本;

平台识别模块, 用于将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别, 得到各个平台反馈的各个平台识别文本;

平台识别文本对比模块, 用于将所述原始文本分别与各个所述平台识别文本进行对比, 确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容;

文本标注模块, 用于标注所述原始文本中除所述部分文本内容以外的文本内容;

清洗发送模块,用于将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

10. 根据权利要求 6 至 9 中任一项所述的录音文件处理装置,其特征在于,所述录音文件处理装置还包括:

统计模块,用于统计目标账户历史上传的各个录音文件的错词率;

平均错词率计算模块,用于根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率;

平均错词率判断模块,用于判断所述目标账户的平均错词率是否小于预设阈值;

优质账户确定模块,用于若所述平均错词率判断模块的判断结果为是,则将所述目标账户确定为优质录音账户,优质录音账户上传录音文件时受到预设激励机制的奖励。

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

获取上传的录音文件以及与所述录音文件对应的原始文本;

调用语音识别接口对所述录音文件进行语音识别,得到识别文本;

判断所述识别文本与所述原始文本是否一致;

若所述识别文本与所述原始文本一致,则将所述录音文件存储至预设的模型训练集合中;

若所述识别文本与所述原始文本不一致,则将所述录音文件记录至待清洗目录,所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本;

获取所述待清洗目录中清洗后的录音文件以及对应的录音文本;

将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

12. 根据权利要求 11 所述的计算机设备,其特征在于,所述判断所述识别文本与所述原始文本是否一致包括:

计算所述识别文本相对所述原始文本的错词率;

判断所述错词率是否为 0;

若所述错词率为 0,则确定所述识别文本与所述原始文本一致;

若所述错词率不为 0,则确定所述识别文本与所述原始文本不一致。

13. 根据权利要求 11 所述的计算机设备,其特征在于,所述将所述录音文件记录至待清洗目录包括:

获取所述录音文件初定的应用领域；

将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上；

所述获取所述待清洗目录中清洗后的录音文件以及对应的录音文本具体为：获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域，所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域；

所述将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中具体包括：

判断所述录音文件的第一应用领域与初定的应用领域是否一致；

若所述录音文件的第一应用领域与初定的应用领域一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上；

若所述录音文件的第一应用领域与初定的应用领域不一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

14. 根据权利要求 11 所述的计算机设备，其特征在于，在获取所述待清洗目录中清洗后的录音文件以及对应的录音文本之前，所述处理器执行所述计算机可读指令时还实现如下步骤：

获取待清洗的录音文件和对应的原始文本；

将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别，得到各个平台反馈的各个平台识别文本；

将所述原始文本分别与各个所述平台识别文本进行对比，确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容；

标注所述原始文本中除所述部分文本内容以外的文本内容；

将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

15. 根据权利要求 11 至 14 中任一项所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还实现如下步骤：

统计目标账户历史上传的各个录音文件的错词率；

根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率；

判断所述目标账户的平均错词率是否小于预设阈值；

若所述目标账户的平均错词率小于预设阈值，则将所述目标账户确定为优质录音账户，优质录音账户上传录音文件时受到预设激励机制的奖励。

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

获取上传的录音文件以及与所述录音文件对应的原始文本；

调用语音识别接口对所述录音文件进行语音识别，得到识别文本；

判断所述识别文本与所述原始文本是否一致；

若所述识别文本与所述原始文本一致，则将所述录音文件存储至预设的模型训练集合中；

若所述识别文本与所述原始文本不一致，则将所述录音文件记录至待清洗目录，所述待清洗目录记录的录音文件由处理人员听取并反馈正确的录音文本；

获取所述待清洗目录中清洗后的录音文件以及对应的录音文本；

将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中。

17. 根据权利要求 16 所述的非易失性可读存储介质，其特征在于，所述判断所述识别文本与所述原始文本是否一致包括：

计算所述识别文本相对所述原始文本的错词率；

判断所述错词率是否为 0；

若所述错词率为 0，则确定所述识别文本与所述原始文本一致；

若所述错词率不为 0，则确定所述识别文本与所述原始文本不一致。

18. 根据权利要求 16 所述的非易失性可读存储介质，其特征在于，所述将所述录音文件记录至待清洗目录包括：

获取所述录音文件初定的应用领域；

将所述录音文件记录至待清洗目录中所述初定的应用领域所属的位置上；

所述获取所述待清洗目录中清洗后的录音文件以及对应的录音文本具体为：获取所述待清洗目录中清洗后的录音文件、所述清洗后的录音文件对应的录音文本、以及第一应用领域，所述第一应用领域由处理人员在清洗所述录音文件时确定的应用领域；

所述将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中具体包括：

判断所述录音文件的第一应用领域与初定的应用领域是否一致；

若所述录音文件的第一应用领域与初定的应用领域一致，则将所述清洗后的录音文件以及对应的录音文本关联存储至所述模型训练集合中所述初定的应用领域所属的位置上；

若所述录音文件的第一应用领域与初定的应用领域不一致，则将所述清洗后的录音文

件以及对应的录音文本关联存储至所述模型训练集合中所述第一应用领域所属的位置上。

19. 根据权利要求 16 所述的非易失性可读存储介质，其特征在于，在获取所述待清洗目录中清洗后的录音文件以及对应的录音文本之前，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

获取待清洗的录音文件和对应的原始文本；

将所述待清洗的录音文件和对应的原始文本发送至各个不同平台的语音识别服务接口进行语音识别，得到各个平台反馈的各个平台识别文本；

将所述原始文本分别与各个所述平台识别文本进行对比，确定所述原始文本中与各个所述平台识别文本均一致的部分文本内容；

标注所述原始文本中除所述部分文本内容以外的文本内容；

将标注后的所述原始文本发送至指定终端以供处理人员进行清洗处理。

20. 根据权利要求 16 至 19 中任一项所述的非易失性可读存储介质，其特征在于，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

统计目标账户历史上传的各个录音文件的错词率；

根据统计得到的所述历史上传的各个录音文件的错词率计算所述目标账户录制录音文件的平均错词率；

判断所述目标账户的平均错词率是否小于预设阈值；

若所述目标账户的平均错词率小于预设阈值，则将所述目标账户确定为优质录音账户，优质录音账户上传录音文件时受到预设激励机制的奖励。

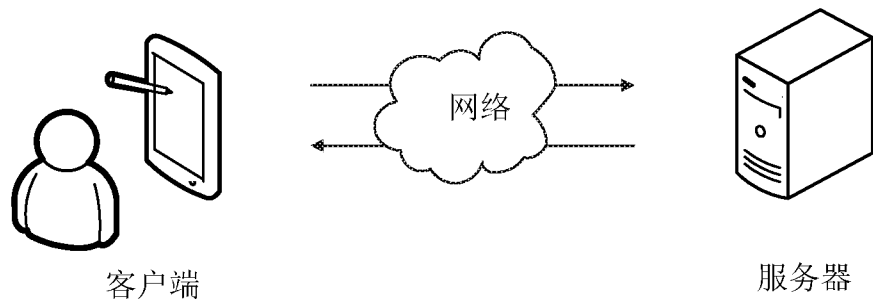


图 1

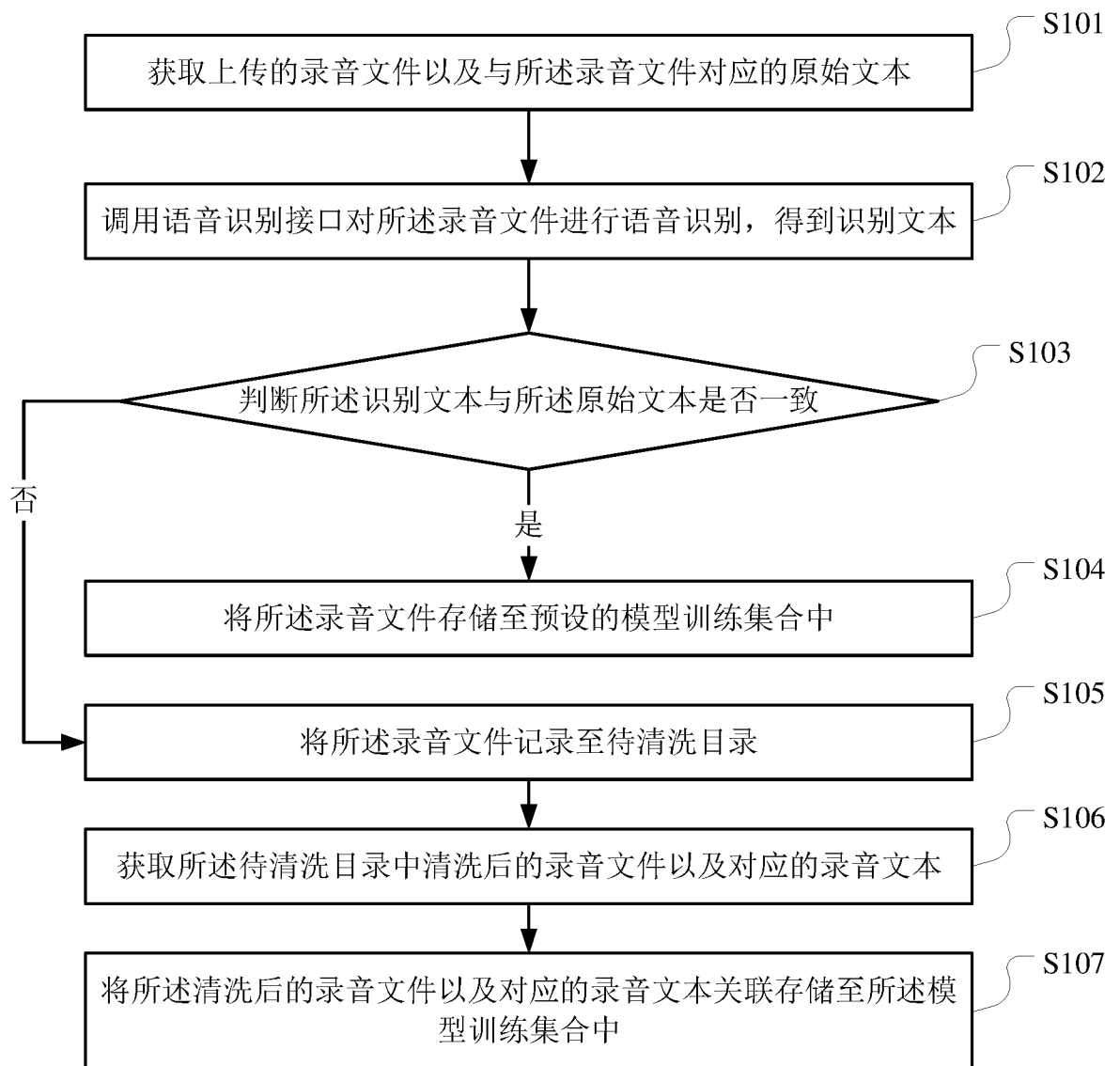


图 2

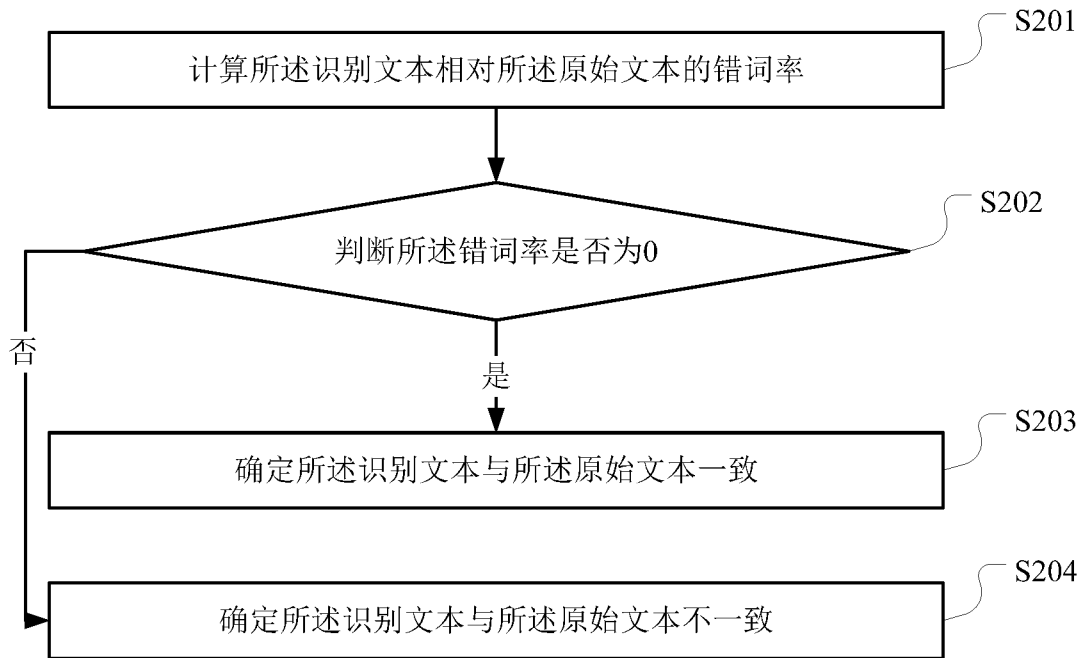


图 3

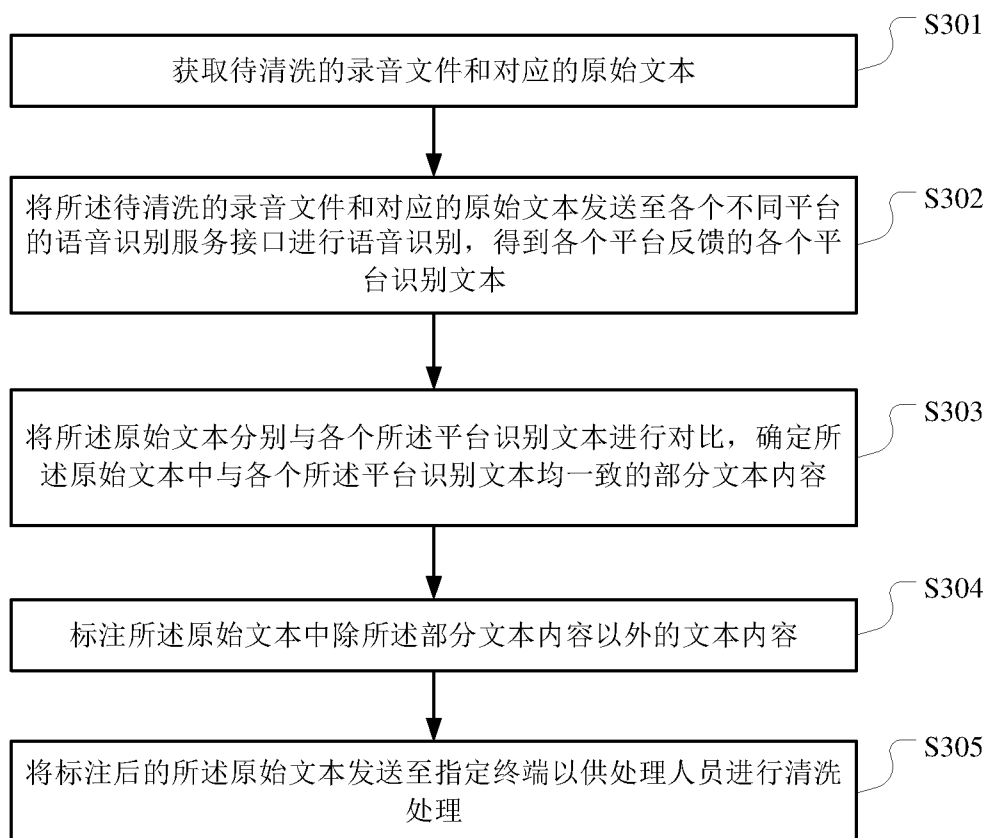


图 4

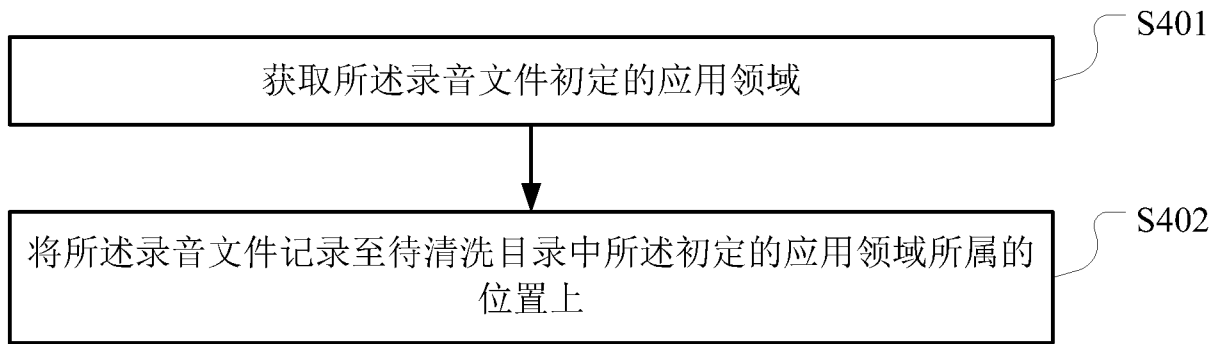


图 5

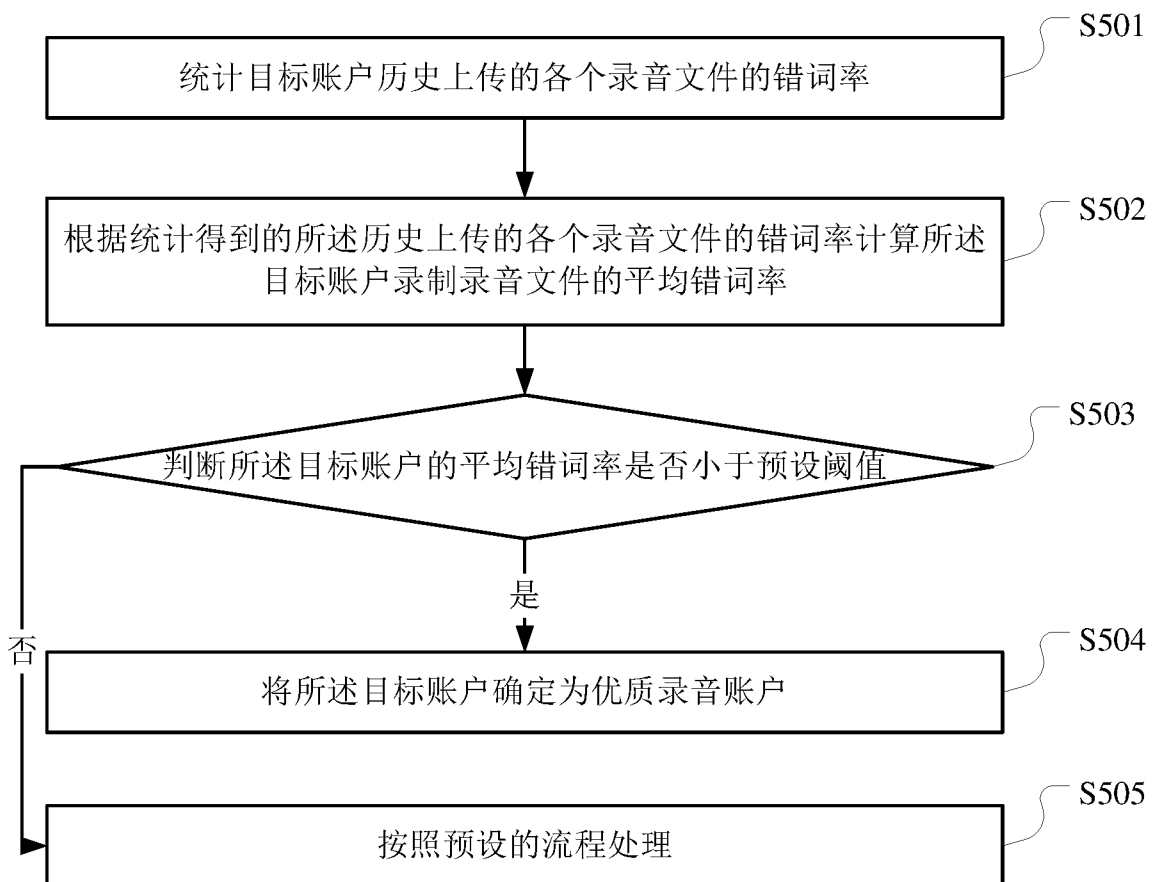


图 6

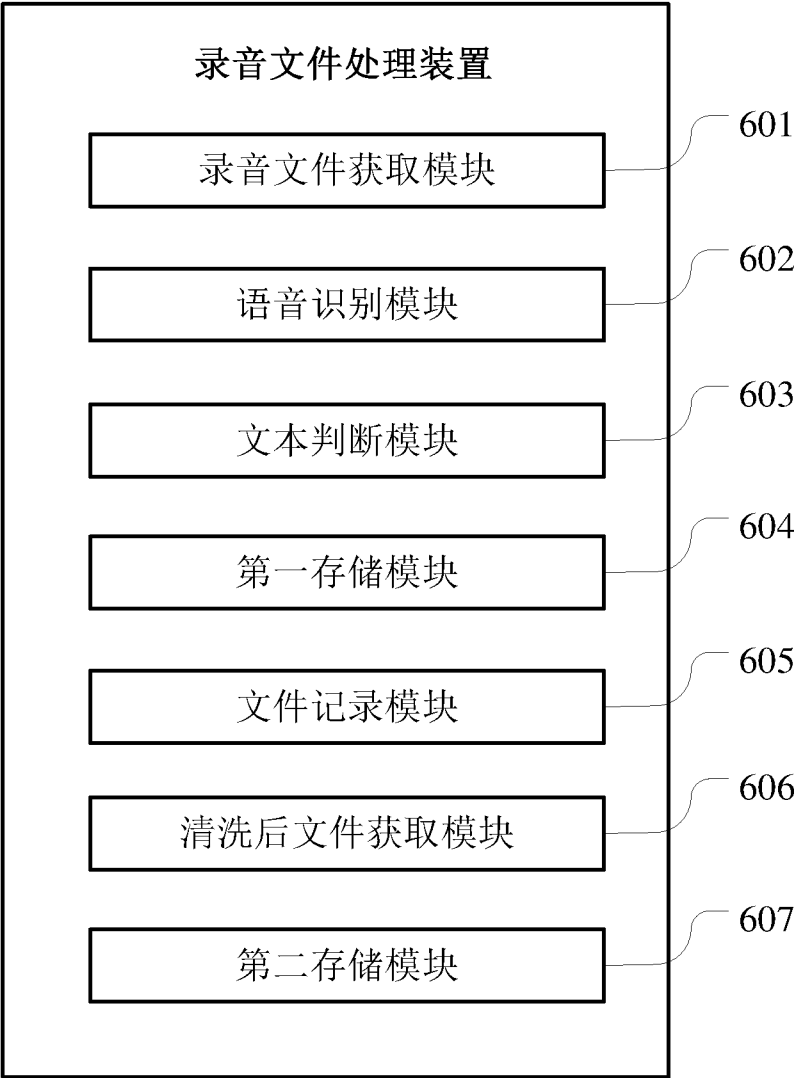


图 7

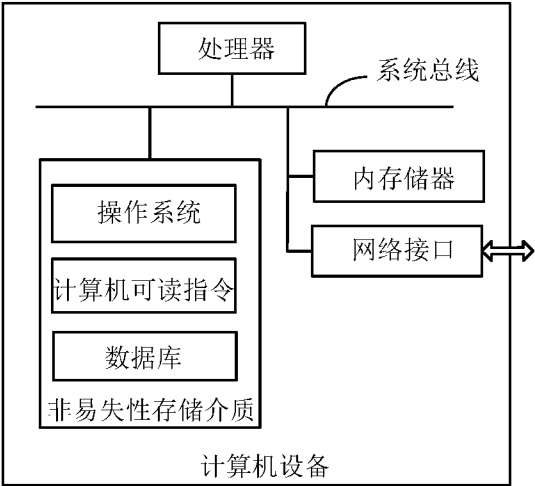


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/106259

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/06(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 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)

CNABS, CNTXT, DWPI, SIPOABS: 语音, 语料, 音频, 标注, 纠错, 修正, 校对, 校正, voice file, voice metadata, correct, label, mark, transcript

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106228980 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 14 December 2016 (2016-12-14) description, paragraphs [0020]-[0048], and figures 1-2	1-20
A	CN 107578769 A (IFLYTEK CO., LTD.) 12 January 2018 (2018-01-12) entire document	1-20
A	US 2006265221 A1 (DICTAPHONE CORP.) 23 November 2006 (2006-11-23) entire document	1-20

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

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

11 December 2018

Date of mailing of the international search report

20 December 2018

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

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/106259

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106228980	A	14 December 2016	None			
CN	107578769	A	12 January 2018	None			
US	2006265221	A1	23 November 2006	WO	2006127342	A2	30 November 2006
				WO	2006127342	A3	23 April 2009
				US	8380510	B2	19 February 2013
				US	2013179171	A1	11 July 2013
				EP	1882249	A2	30 January 2008
				EP	1882249	A4	04 November 2009
				CA	2608566	A1	30 November 2006
				US	8655665	B2	18 February 2014
				EP	1882249	B1	19 October 2011
				AT	529853	T	15 November 2011

国际检索报告

国际申请号

PCT/CN2018/106259

A. 主题的分类

G10L 15/06(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L 15

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, DWPI, SIPOABS: 语音, 语料, 音频, 标注, 纠错, 修正, 校对, 校正, voice file, voice metadata, correct, label, mark, transcript

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 106228980 A (百度在线网络技术北京有限公司) 2016年 12月 14日 (2016 - 12 - 14) 说明书第[0020]-[0048]段, 附图1-2	1-20
A	CN 107578769 A (科大讯飞股份有限公司) 2018年 1月 12日 (2018 - 01 - 12) 全文	1-20
A	US 2006265221 A1 (DICTAPHONE CORP.) 2006年 11月 23日 (2006 - 11 - 23) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2018年 12月 11日

国际检索报告邮寄日期

2018年 12月 20日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

游晓梅

传真号 (86-10)62019451

电话号码 (86-10)62089539

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/106259

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	106228980	A	2016年 12月 14日	无			
CN	107578769	A	2018年 1月 12日	无			
US	2006265221	A1	2006年 11月 23日	WO	2006127342	A2	2006年 11月 30日
				WO	2006127342	A3	2009年 4月 23日
				US	8380510	B2	2013年 2月 19日
				US	2013179171	A1	2013年 7月 11日
				EP	1882249	A2	2008年 1月 30日
				EP	1882249	A4	2009年 11月 4日
				CA	2608566	A1	2006年 11月 30日
				US	8655665	B2	2014年 2月 18日
				EP	1882249	B1	2011年 10月 19日
				AT	529853	T	2011年 11月 15日

表 PCT/ISA/210 (同族专利附件) (2015年1月)