



(12) 发明专利申请

(10) 申请公布号 CN 112802477 A

(43) 申请公布日 2021.05.14

(21) 申请号 202011612549.8

(22) 申请日 2020.12.30

(71) 申请人 科大讯飞云网科技有限公司

地址 230088 安徽省合肥市高新区文曲路  
355号

(72) 发明人 张德昌 丁常坤 夏兵 王江淮  
时代红

(74) 专利代理机构 合肥昊晟德专利代理事务所  
(普通合伙) 34153

代理人 王林

(51) Int. Cl.

G10L 15/26 (2006.01)

H04M 3/51 (2006.01)

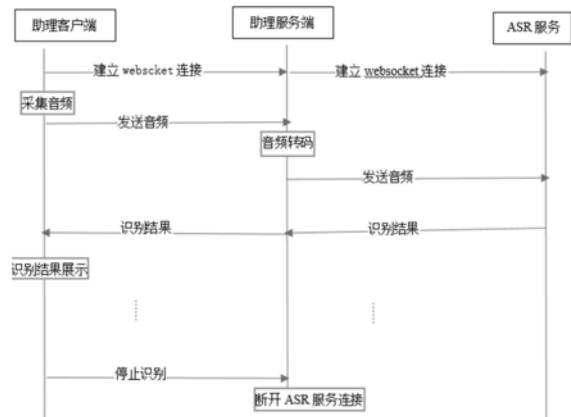
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种基于语音文字转换的客服助理工具服务方法及系统

(57) 摘要

本发明公开了一种基于语音文字转换的客服助理工具服务方法及系统,属于客服服务技术领域,包括以下步骤:S1:调起客户端;S2:语音采集;S3:语音发送;S4:语音转换处理;S5:语音识别;S6:识别结果返回;S7:识别结果展示;S8:停止展示。本发明通过设置助理服务端,通过助理服务端接收语音并返回识别结果,可以在助理服务端录音及保存整通电话的识别结果,便于质检等其他业务;通过接口调用ASR服务,可以更换服务提供商;客服人员可以随时复制聊天文字,或者翻看前面的聊天记录。



1. 一种基于语音文字转换的客服助理工具服务方法,其特征在于,包括以下步骤:

S1: 调起客户端

当客服人员通过前台系统接听到来电,调起助理客户端,开始与客户交流;

S2: 语音采集

助理客户端启动后开始进行音频采集,包括客服人员的声音与客户的声音;

S3: 语音发送

助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

S4: 语音转换处理

利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR服务厂商的websocket接口需要;

S5: 语音识别

助理服务端将处理后的音频数据通过websocket接口发送到ASR服务提供方进行识别,并实时接收返回结果;

S6: 识别结果返回

助理服务端将识别结果通过websocket接口返回给助理客户端;

S7: 识别结果展示

将助理客户端展示界面通过独立窗口进行展示;

S8: 停止展示

通话结束后,助理客户端停止采集音频数据。

2. 根据权利要求1所述的一种基于语音文字转换的客服助理工具服务方法,其特征在于:在所述步骤S1中,客服人员接听来话事件触发时,前台系统通过JavaScript调起助理客户端,并传递当前通话信息。

3. 根据权利要求2所述的一种基于语音文字转换的客服助理工具服务方法,其特征在于:在所述步骤S2中,助理客户端调用windows提供的COM组件,使用Core Audio进行对音频数据进行采集,其中采集客户语音使用loopback模式,采集客服人员语音使用capture模式。

4. 根据权利要求3所述的一种基于语音文字转换的客服助理工具服务方法,其特征在于:在所述步骤S4中,ASR服务websocket接口需要8000hz、单声道、位宽16bit的PCM音频数据,需要对采集到的音频数据进行转码,将音频数据转换成ASR服务websocket接口所需格式。

5. 根据权利要求4所述的一种基于语音文字转换的客服助理工具服务方法,其特征在于:当接听来电事件触发,所述助理客户端启动之后,首先连接所述助理服务端并发送文本信息,所述助理服务端接收到该文本信息之后启动一个websocket连接,连接到ASR服务,并创建相关数据结构,将两个通道关联起来。

6. 根据权利要求5所述的一种基于语音文字转换的客服助理工具服务方法,其特征在于:当通话结束事件触发,所述助理客户端发送结束标志给所述助理服务端,所述助理服务端则断开与ASR服务的websocket连接,并清理相关资源。

7. 一种基于语音文字转换的客服助理工具服务系统,利用如权利要求1~6任一所述的服务方法对客户进行服务,包括:

客户端调起模块,用于通过前台系统接听到来电,调起助理客户端,开始与客户交流;

语音采集模块,用于在助理客户端启动后进行音频采集,包括客服人员的声音与客户的声音;

语音发送模块,用于通过助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

语音转换处理模块,用于利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR厂商的websocket接口需要;

语音识别模块,用于通过ASR服务提供方对处理后的音频数据进行识别,并实时接收返回结果;

识别结果返回模块,用于利用助理服务端将识别结果通过websocket接口返回给助理客户端;

结果展示模块,用于将助理客户端展示界面通过独立窗口进行展示;

展示停止模块,用于在通话结束后,停止采集音频数据;

中央处理模块,用于向其他模块发出指令,完成相关动作;

所述客户端调起模块、语音采集模块、语音发送模块、语音转换处理模块、语音识别模块、识别结果返回模块、结果展示模块、展示停止模块均与中央处理模块电连接。

## 一种基于语音文字转换的客服助理工具服务方法及系统

### 技术领域

[0001] 本发明涉及客服服务技术领域,具体涉及一种基于语音文字转换的客服助理工具服务方法及系统。

### 背景技术

[0002] 语音客服人员在接听客户电话的过程中,经常会需要记录或向用户确认一些信息。在信息量较大时,如果没有其他辅助手段或工具而全凭记忆力,一方面会加重客服人员负担,另一方面会导致由于信息不准确而需要重复向客户询问与确认。

[0003] 客服人员在通话中有时需要将客户所说内容录入CRM系统等相关系统中去,如果全部手动打字录入,会占用更多通话时间,或导致分心,工作效率不够高。因此,提出一种基于语音文字转换的客服助理工具服务方法。

### 发明内容

[0004] 本发明所要解决的技术问题在于:如何在客服人员使用的电脑上展示客服人员与客户双方通话的内容,供客服人员参考及复制使用,提供了一种基于语音文字转换的客服助理工具服务方法。

[0005] 本发明是通过以下技术方案解决上述技术问题的,本发明包括以下步骤:

[0006] S1:调起客户端

[0007] 当客服人员通过前台系统接听到来电,调起助理客户端,开始与客户交流;

[0008] S2:语音采集

[0009] 助理客户端启动后开始进行音频采集,包括客服人员的声音与客户的声音;

[0010] S3:语音发送

[0011] 助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

[0012] S4:语音转换处理

[0013] 利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR(自动语音识别技术)厂商的websocket接口需要;

[0014] S5:语音识别

[0015] 助理服务端将处理后的音频数据通过websocket接口发送到ASR服务提供方进行识别,并实时接收返回结果;

[0016] S6:识别结果返回

[0017] 助理服务端将识别结果通过websocket接口返回给助理客户端;

[0018] S7:识别结果展示

[0019] 将助理客户端展示界面通过独立窗口进行展示;

[0020] S8:停止展示

[0021] 通话结束后,助理客户端停止采集音频数据。

[0022] 更进一步地,在所述步骤S1中,客服人员接听来话事件触发时,前台系统通过

JavaScript调起助理客户端,并传递当前通话信息。

[0023] 更进一步地,在所述步骤S2中,助理客户端调用windows提供的COM组件,使用Core Audio进行对音频数据进行采集,其中采集客户语音使用loopback模式,采集客服人员语音使用capture模式。

[0024] 更进一步地,在所述步骤S4中,ASR服务websocket接口需要8000hz、单声道、位宽16bit的PCM音频数据,需要对采集到的音频数据进行转码,将音频数据转换成ASR服务websocket接口所需格式。

[0025] 更进一步地,当接听来电事件触发,所述助理客户端启动之后,首先连接所述助理服务端并发送文本信息,所述助理服务端接收到该文本信息之后启动一个websocket连接,连接到ASR服务,并创建相关数据结构,将两个通道关联起来。

[0026] 更进一步地,当通话结束事件触发,所述助理客户端发送结束标志给所述助理服务端,所述助理服务端则断开与ASR服务的websocket连接,并清理相关资源。

[0027] 本发明还提供了一种基于语音文字转换的客服助理工具服务系统,利用上述的服务方法对客户进行服务,包括:

[0028] 客户端调起模块,用于通过前台系统接听到来电,调起助理客户端,开始与客户交流;

[0029] 语音采集模块,用于在助理客户端启动后进行音频采集,包括客服人员的声音与客户的声音;

[0030] 语音发送模块,用于通过助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

[0031] 语音转换处理模块,用于利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR厂商的websocket接口需要;

[0032] 语音识别模块,用于通过ASR服务提供方对处理后的音频数据进行识别,并实时接收返回结果;

[0033] 识别结果返回模块,用于利用助理服务端将识别结果通过websocket接口返回给助理客户端;

[0034] 结果展示模块,用于将助理客户端展示界面通过独立窗口进行展示;

[0035] 展示停止模块,用于在通话结束后,停止采集音频数据;

[0036] 中央处理模块,用于向其他模块发出指令,完成相关动作;

[0037] 所述客户端调起模块、语音采集模块、语音发送模块、语音转换处理模块、语音识别模块、识别结果返回模块、结果展示模块、展示停止模块均与中央处理模块电连接。

[0038] 本发明相比现有技术具有以下优点:该基于语音文字转换的客服助理工具服务方法,通过设置助理服务端,通过助理服务端接收语音并返回识别结果,可以在助理服务端录音及保存整通电话的识别结果,便于质检等其他业务;通过接口调用ASR服务,可以更换服务提供商;客服人员可以随时复制聊天文字,或者翻看前面的聊天记录,值得被推广使用。

## 附图说明

[0039] 图1是本发明实施例二中助理客户端与助理服务端的交互流程示意图;

[0040] 图2是本发明实施例二中独立窗口界面的示例图。

## 具体实施方式

[0041] 下面对本发明的实施例作详细说明,本实施例在以本发明技术方案为前提下进行实施,给出了详细的实施方式和具体的操作过程,但本发明的保护范围不限于下述的实施例。

[0042] 实施例一

[0043] 本实施例提供一种技术方案:一种基于语音文字转换的客服助理工具服务方法,包括以下步骤:

[0044] S1:调起客户端

[0045] 当客服人员通过前台系统接听到来电,调起助理客户端,开始与客户交流;

[0046] S2:语音采集

[0047] 助理客户端启动后开始进行音频采集,包括客服人员的声音与客户的语音;

[0048] S3:语音发送

[0049] 助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

[0050] S4:语音转换处理

[0051] 利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR(自动语音识别技术)厂商的websocket接口需要;

[0052] S5:语音识别

[0053] 助理服务端将处理后的音频数据通过websocket接口发送到ASR服务提供方进行识别,并实时接收返回结果;

[0054] S6:识别结果返回

[0055] 助理服务端将识别结果通过websocket接口返回给助理客户端;

[0056] S7:识别结果展示

[0057] 将助理客户端展示界面通过独立窗口进行展示;

[0058] S8:停止展示

[0059] 通话结束后,助理客户端停止采集音频数据。

[0060] 在所述步骤S1中,客服人员接听来电事件触发时,前台系统通过JavaScript调起助理客户端,并传递当前通话信息。

[0061] 在所述步骤S2中,助理客户端调用windows提供的COM组件,使用Core Audio进行对音频数据进行采集,其中采集客户语音使用loopback模式,采集客服人员语音使用capture模式。

[0062] 在所述步骤S4中,ASR服务websocket接口需要8000hz、单声道、位宽16bit的PCM音频数据,需要对采集到的音频数据进行转码,将音频数据转换成ASR服务websocket接口所需格式。

[0063] 当接听来电事件触发,所述助理客户端启动之后,首先连接所述助理服务端并发送文本信息,所述助理服务端接收到该文本信息之后启动一个websocket连接,连接到ASR服务,并创建相关数据结构,将两个通道关联起来。

[0064] 当通话结束事件触发,所述助理客户端发送结束标志给所述助理服务端,所述助理服务端则断开与ASR服务的websocket连接,并清理相关资源。

[0065] 本实施例还提供了一种基于语音文字转换的客服助理工具服务系统,利用上述的

服务方法对客户进行服务,包括:

[0066] 客户端调起模块,用于通过前台系统接听到来电,调起助理客户端,开始与客户交流;

[0067] 语音采集模块,用于在助理客户端启动后进行音频采集,包括客服人员的声音与客户的声音;

[0068] 语音发送模块,用于通过助理客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

[0069] 语音转换处理模块,用于利用助理服务端或助理客户端对采集到的语音进行相应转换处理,以适应ASR厂商的websocket接口需要;

[0070] 语音识别模块,用于通过ASR服务提供方对处理后的音频数据进行识别,并实时接收返回结果;

[0071] 识别结果返回模块,用于利用助理服务端将识别结果通过websocket接口返回给助理客户端;

[0072] 结果展示模块,用于将助理客户端展示界面通过独立窗口进行展示;

[0073] 展示停止模块,用于在通话结束后,停止采集音频数据;

[0074] 中央处理模块,用于向其他模块发出指令,完成相关动作;

[0075] 所述客户端调起模块、语音采集模块、语音发送模块、语音转换处理模块、语音识别模块、识别结果返回模块、结果展示模块、展示停止模块均与中央处理模块电连接。

[0076] 实施例二

[0077] 本发明的目的是将话务双方的语音转换为文字展示于客服系统中,主要涉及麦克风与声卡语音采集、音频转换、语音识别服务等技术。实际系统应该包括安装于客服人员电脑中的助理客户端和部署于服务器的助理服务端,以及第三方厂商提供的语音识别(ASR)的websocket接口。

[0078] 如图1所示,本实施例的主要流程步骤如下:

[0079] S1:调起客户端

[0080] 当客服人员通过前台系统(如使用浏览器)接听到来电,调起助理客户端,开始与客户交流;

[0081] S2:语音采集

[0082] 助理客户端一旦启动就开始进行语音的捕获,包括从麦克风输入到计算机的客服人员的声音和从计算机声卡输出到耳机的客户的声音;

[0083] S3:语音发送

[0084] 客户端将采集到的两路语音通过websocket接口实时发送到助理服务端;

[0085] S4:语音转换处理

[0086] 助理服务端负责对语音进行相应转换处理,以适应ASR厂商的websocket接口需要;

[0087] 如常见的百度ASR服务接口要求传送频率为8000hz、单声道、位宽16bit的PCM音频数据,而助理客户端传来的音频数据一般不能完全符合这个要求,所以助理服务端需要对此数据转换为完全符合接口要求的格式。音频格式转换完成之后,才能按照一定的频率发送到ASR接口;

[0088] S5:语音识别

[0089] 助理服务端将处理后的音频数据通过websocket接口发送到ASR服务提供方,并实时接收返回结果;

[0090] 常规的语音识别方式为语音文件转写,而客服助理是需要实时语音识别,这里需要ASR服务能提供实时语音识别,实时语音识别接口一般以websocket接口的方式提供,助理服务端会以一定的频率发送音频数据帧,ASR服务端则实时返回临时识别结果和最终识别结果;

[0091] S6:识别结果返回

[0092] 助理服务端将识别结果通过websocket接口返回给助理客户端;

[0093] S7:识别结果展示

[0094] 助理客户端展示界面可以以独立窗口存在,展示形式可以如图2所示;

[0095] 助理客户端是一个需要安装在客服电脑上的程序,该程序单独运行,并且可以提供展示界面。

[0096] S8:停止展示

[0097] 通话结束后,助理客户端将停止采集数据。

[0098] 本实施例的具体实现原理如下:

[0099] 客户端调起:客服人员接听来话事件触发时,助理客户端安装时会自动注册到注册表,浏览器前端可以通过JavaScript调起助理客户端,助理客户端启动后自动连接到助理服务端,并传递当前通话信息。。

[0100] 音频采集:助理客户端调用windows提供的COM组件,使用Core Audio进行捕获。其中捕获客户语音使用loopback模式,捕获麦克风录入语音使用capture模式。

[0101] 音频转码:由于ASR服务websocket接口一般需要8000hz、单声道、位宽16bit的PCM数据,而捕获到的数据编码往往与当前计算机的具体硬件有关,如常见的48000hz、双声道,所以需要转码。转码功能可以在助理客户端实现,也可以在助理服务端实现。音频转码是通过调用FFmpeg实现的,如果出于节省网络流量考虑使用助理客户端转码,则在助理客户端内部集成FFmpeg即可;如果由助理服务端实现转码,则由助理服务端集成FFmpeg,使用FFmpeg是由于这套工具编解码效率高。

[0102] 助理客户端与助理服务端的交互:助理客户端与助理服务端之间的交互也使用websocket。当接听来电事件触发,助理客户端启动之后,首先连接助理服务端并发送文本信息,其中包含当前电脑系统采集音频的采样率、采样位数、声道数等信息。助理服务端接收到该文本信息之后启动一个websocket连接,连接到ASR服务,并创建相关数据结构,将两个通道关联起来。此数据结构包含音频数据缓冲区、与助理客户端的websocket连接通道、与ASR服务的websocket连接通道。其中音频数据缓冲区是用于存储助理客户端发送过来并经过转码之后的音频数据,助理服务端向ASR服务接口发送音频数据也是从这个缓冲区取数据,使用这个缓冲区是由于助理客户端发送数据的频率与助理服务端向ASR服务接口发送数据的频率不匹配。此后,助理客户端将以一定时间间隔不断发送音频数据,而助理服务端则将ASR服务返回的识别结果以文本发送到助理客户端。当通话结束事件触发,助理客户端发送结束标志给助理服务端,助理服务端则断开与ASR服务的websocket连接,并清理相关资源。

[0103] 综上所述,上述实施例的基于语音文字转换的客服助理工具服务方法,通过设置助理服务端,通过助理服务端接收语音并返回识别结果,可以在助理服务端录音及保存整通电话的识别结果,便于质检等其他业务;通过接口调用ASR服务,可以更换服务提供商;客服人员可以随时复制聊天文字,或者翻看前面的聊天记录,值得被推广使用。

[0104] 尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

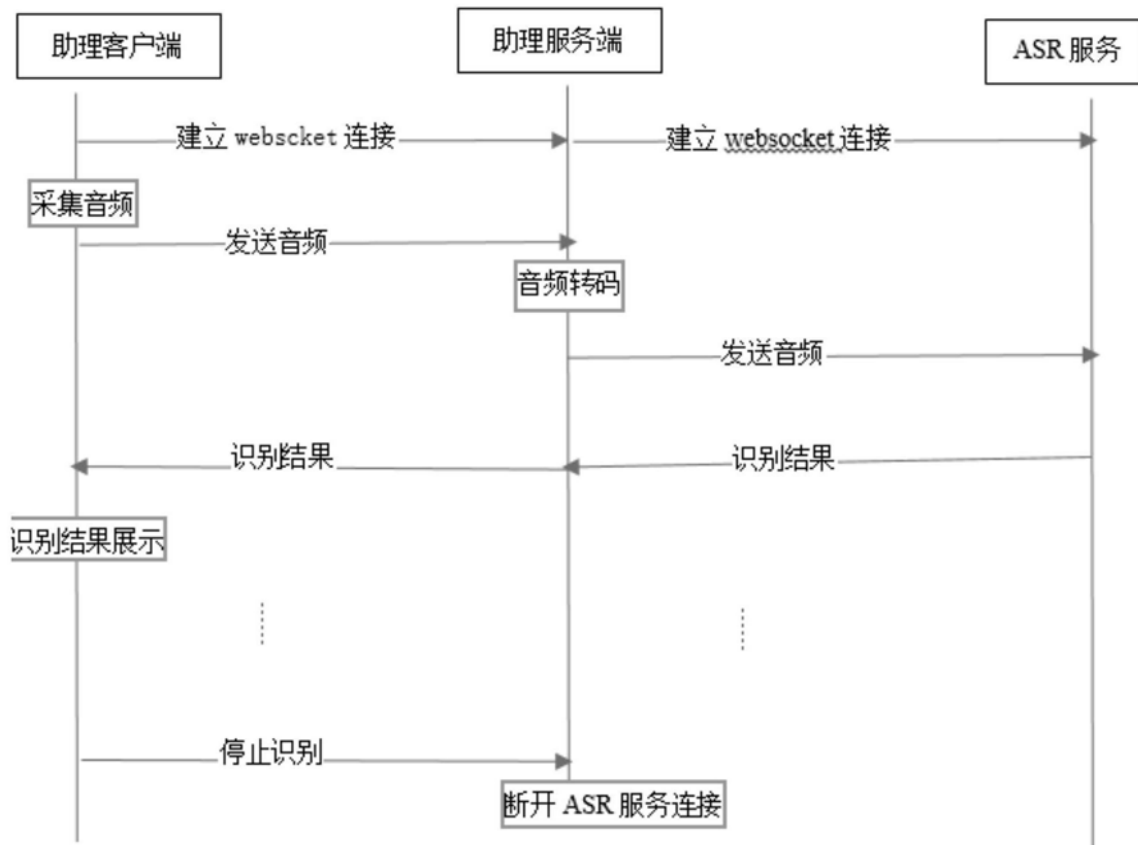


图1



图2