



(12)发明专利申请

(10)申请公布号 CN 109637541 A

(43)申请公布日 2019.04.16

(21)申请号 201811653539.1

(22)申请日 2018.12.29

(71)申请人 联想(北京)有限公司

地址 100085 北京市海淀区上地信息产业
基地创业路6号

(72)发明人 朱冲 何凤龙

(74)专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 杨静

(51)Int.Cl.

G10L 15/26(2006.01)

G06F 3/0487(2013.01)

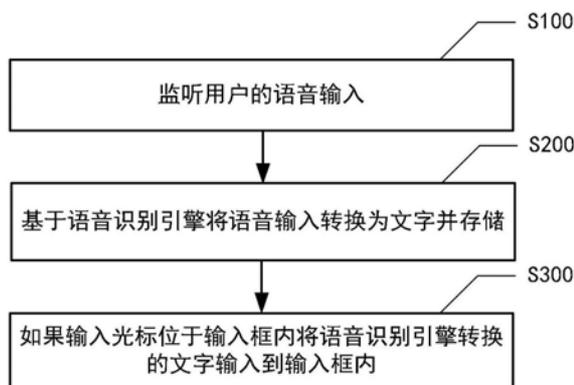
权利要求书2页 说明书10页 附图3页

(54)发明名称

语音转换文字的方法和电子设备

(57)摘要

本公开提供了一种语音转换文字的方法和一种电子设备,其中,语音转换文字的方法包括监听用户的语音输入;基于语音识别引擎将所述语音输入转换为文字并存储;如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内。



1. 一种语音转换文字的方法,包括:

监听用户的语音输入;

基于语音识别引擎将所述语音输入转换为文字并存储;

如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内。

2. 根据权利要求1所述的方法,其中,所述如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内包括以下操作中的至少一种:

如果输入光标位于输入框内,将第一部分文字输入到所述输入框,所述第一部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之后所监听到的语音输入转换的文字;

如果输入光标位于输入框内,将第二部分文字输入到所述输入框,所述第二部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之前以及所述输入光标位于所述输入框之后的一完整单元的语音输入转换的文字。

3. 根据权利要求1所述的方法,其中,所述如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内包括:

如果输入光标位于输入框内,将第三部分文字输入到所述输入框,所述第三部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之前基于所监听到的语音输入转换完成的文字。

4. 根据权利要求3所述的方法,其中,所述如果输入光标位于输入框内,将第三部分文字输入到所述输入框包括:

如果输入光标位于输入框内,输出至少一个选项,以使得用户选择所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成的文字。

5. 根据权利要求4所述的方法,其中,所述输出至少一个选项包括以下形式中的至少一种:

选取与所述输入光标位于所述输入框时间上满足时间范围的语句所对应的文字作为选项;

从所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字中智能选取重要语句所对应的文字作为选项;

从所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字中智能选取与所述输入框属性匹配的语音所对应的文字作为选项;

将所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字的数量标识作为选项。

6. 根据权利要求1所述的方法,还包括:

按照丢弃策略删除所述实时监听用户的语音输入并基于语音识别引擎将所述语音输入转换的文字。

7. 根据权利要求1所述的方法,所述基于语音识别引擎将所述语音输入转换为文字并存储包括:

将所述监听的语音输入传输至服务器;

接收所述服务器传输的文字并存储,其中,所述文字由所述服务器基于语音识别引擎将所述语音输入转换得到。

8. 一种电子设备,包括:
麦克风;
存储单元,用于存储所述麦克风所获得的语音输入通过语音识别引擎转换的文字;
处理器,用于控制所述麦克风监听用户的语音输入,以及在输入光标位于输入框的情况下将所述语音识别引擎转换的文字输入到所述输入框内。
9. 根据权利要求8所述的电子设备,其中,所述处理器还用于基于语音识别引擎将所述监听的语音输入转换为文字并存储至所述处理器。
10. 根据权利要求8所述的电子设备,其中,所述处理器还用于将所述监听的语音输入传输至服务器,并从所述服务器接收文字并存储,其中,所述文字由所述服务器基于语音识别引擎将所述语音输入转换得到。

语音转换文字的方法和电子设备

技术领域

[0001] 本公开涉及一种语音转换文字的方法和电子设备。

背景技术

[0002] 随着科学技术的发展,电子设备的应用覆盖了人们工作和生活的方方面面,与此同时,用户对于电子设备的要求也越来越高。在电子设备的使用过程中,存在许多需要输入文字的场景,例如在文档中输入文字、在搜索栏中输入文字等。在现有技术中,针对上述输入文字的场景,基本上是通过键盘来输入文字,速度较慢,造成用户体验较差。

发明内容

[0003] 有鉴于此,本公开提供了一种语音转换文字的方法,包括:监听用户的语音输入;基于语音识别引擎将所述语音输入转换为文字并存储;如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内。

[0004] 根据本公开的实施例,所述如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内包括以下操作中的至少一种:如果输入光标位于输入框内,将第一部分文字输入到所述输入框,所述第一部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之后所监听到的语音输入转换的文字;如果输入光标位于输入框内,将第二部分文字输入到所述输入框,所述第二部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之前以及所述输入光标位于所述输入框之后的一完整单元的语音输入转换的文字。

[0005] 根据本公开的实施例,所述如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内包括:如果输入光标位于输入框内,将第三部分文字输入到所述输入框,所述第三部分文字为所述语音识别引擎基于所述输入光标位于所述输入框之前基于所监听到的语音输入转换完成的文字。

[0006] 根据本公开的实施例,所述如果输入光标位于输入框内,将第三部分文字输入到所述输入框包括:如果输入光标位于输入框内,输出至少一个选项,以使得用户选择所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成的文字。

[0007] 根据本公开的实施例,所述输出至少一个选项包括以下形式中的至少一种:选取与所述输入光标位于所述输入框时间上满足时间范围的语句所对应的文字作为选项;从所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字中智能选取重要语句所对应的文字作为选项;从所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字中智能选取与所述输入框属性匹配的语音所对应的文字作为选项;将所述语音识别引擎基于所述输入光标位于所述输入框之前所监听到的语音输入转换完成文字的数量标识作为选项。

[0008] 根据本公开的实施例,语音转换文字的方法还包括:按照丢弃策略删除所述实时监听用户的语音输入并基于语音识别引擎将所述语音输入转换的文字。

[0009] 根据本公开的实施例,所述基于语音识别引擎将所述语音输入转换为文字并存储包括:将所述监听的语音输入传输至服务器;接收所述服务器传输的文字并存储,其中,所述文字由所述服务器基于语音识别引擎将所述语音输入转换得到。

[0010] 本公开还提供了一种电子设备,包括:麦克风;存储单元,用于存储所述麦克风所获得的语音输入通过语音识别引擎转换的文字;

[0011] 处理器,用于控制所述麦克风监听用户的语音输入,以及在输入光标位于输入框的情况下将所述语音识别引擎转换的文字输入到所述输入框内。

[0012] 根据本公开的实施例,所述处理器还用于基于语音识别引擎将所述监听的语音输入转换为文字并存储至所述处理器。

[0013] 根据本公开的实施例,所述处理器还用于将所述监听的语音输入传输至服务器,并从所述服务器接收文字并存储,其中,所述文字由所述服务器基于语音识别引擎将所述语音输入转换得到。

附图说明

[0014] 通过以下参照附图对本公开实施例的描述,本公开的上述以及其他目的、特征和优点将更为清楚,在附图中:

[0015] 图1示意性示出了根据本公开实施例的语音转换文字方法的应用场景;

[0016] 图2示意性示出了根据本公开实施例的语音转换文字方法的流程图;

[0017] 图3示意性示出了根据本公开另一实施例的语音转换文字方法的流程图;

[0018] 图4示意性示出了根据本公开实施例的语音转换文字系统的方框图;

[0019] 图5示意性示出了根据本公开实施例的电子设备的方框图;

[0020] 图6示意性示出了根据本公开另一实施例的电子设备的方框图。

具体实施方式

[0021] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。在下面的详细描述中,为便于解释,阐述了许多具体的细节以提供对本公开实施例的全面理解。然而,明显地,一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0022] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用的术语“包括”、“包含”等表明了所述特征、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、操作或部件。

[0023] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0024] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的装置”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的装置等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来

说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的装置”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的装置等)。

[0025] 本公开的实施例提供了一种语音转换文字的方法,包括:监听用户的语音输入;基于语音识别引擎将所述语音输入转换为文字并存储;如果输入光标位于输入框内将所述语音识别引擎转换的文字输入到所述输入框内。

[0026] 本公开实施例提供的语音转换文字的方法,可以在电子设备运行期间,实时监听用户语音信息并转换为文字输入到界面的输入框中。在界面上任何提示需要输入文字的位置,都可以随时通过语音输入文字,例如网页上的输入框;办公软件能够输入字符的区域(Word、PPT等的正文区域);记录软件的输入记录内容的记录框,聊天软件针对某个通信对象输入内容的输入框。因此,本公开实施例的语音转换文字方法可以完全解放用户双手,代替键盘实现文字输入功能。另外,本公开实施例提供的语音转换文字的方法可以通过硬件实现,例如可以将该语音转换文字的功能固化在电子设备的芯片(例如可以是DSP芯片)上,电子设备被打开后即可启用该语音转换文字的功能,不需像现有的输入软件一样基于用户启动才能实现相应的输入功能。

[0027] 图1示意性示出了根据本公开实施例的语音转换文字方法的应用场景。需要注意的是,图1所示仅为可以应用本公开实施例的语音转换文字方法的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他装置、环境或场景。

[0028] 如图1所示,本公开实施例的语音转换文字方法可以用于电子设备,电子设备例如可以是笔记本电脑1。笔记本电脑上设置有音频接收装置,例如可以是麦克风。语音转换文字方法可以通过硬件实现,例如可以将语音转换文字的功能固化在电子设备的芯片(例如可以是DSP芯片)上,笔记本电脑1在运行期间,可以利用麦克风时刻监听用户的语音信息,并实时将监听到的语音输入转换成文字并进行缓存,如果该笔记本电脑接收到输入操作将该电子设备的光标位于输入框内,将转换的文字显示于输入框中,实现实时语音输入文字的效果。其中,该输入操作为启动了办公软件例如一个word文档,如果该word文档处于运行状态时光标便位于该word文档正文输入框内。

[0029] 例如,在开会过程中,只需将笔记本电脑1打开,使其处于正常运行状态,打开Word或记事本等可以记录文字的应用即可将会议过程中参会人说的话记录下来。由于本实施里是利用电子设备麦克风以及语音识别引擎(即,本地语音识别引擎或者与该电子设备连接的云端识别引擎)实时的进行语音采集和识别。在有输入框存在的情况下就能够将实时采集和识别后的文字输入至该输入框内。从而不需要额外安装和启用任何有关语音输入文字的软件应用。此外,还可以用于网页、PPT、Excel等任何具有输入框的界面上的文字输入。

[0030] 此外,本公开实施例的语音转换文字方法还可以用于记录的电子设备,例如录音笔。录音笔可以与第二电子设备连接,第二电子设备例如可以是笔记本电脑等具有显示屏的电子设备。录音笔与第二电子设备之间可以是有线连接或无线连接,以使两者之间可以相互传输信息。

[0031] 如果录音笔接收到笔记本电脑传输的输入指令的情况下,将相应文字传输至笔记本电脑。其中,该输入指令是笔记本电脑的显示界面上输入光标位于输入框时向录音笔发

送的指令。例如,用户通过上述交互方式使得笔记本电脑的输入光标位于输入框内,笔记本电脑产生并发送输入指令,所述录音笔收到该输入指令且响应所述输入指令将相应部分文字发送至笔记本电脑,以使相应部分文字输入笔记本电脑界面上的输入框中。即,如果所述录音笔收到另外一个电子设备的输入指令,所述输入指令表征输入光标位于输入框。当然,本公开的实施例还可以是用户通过该录音笔在例如笔记本电脑等电子设备上进行输入操作使得所述笔记本电脑的光标位于输入框内;例如,录音笔能够被电子设备的传感器所感应到,从而确定录音笔点击了电子设备的显示屏上所显示的输入框,该电子设备根据该录音笔的输入操作产生输入指令并发送该输入指令给所述录音笔。所述录音笔收到该输入指令且响应所述输入指令将相应部分文字发送至电子设备,以使相应部分文字输入笔记本电脑界面上的输入框中。即,如果所述录音笔收到另外一个电子设备的输入指令,所述输入指令表征输入光标位于输入框。

[0032] 可以理解,图1中的应用场景仅是一种示例,该语音转换文字方法可以用于任何具有音频接收装置处理器和存储器的电子设备,例如手机、平板电脑、智能电视、智能音箱等电子设备。语音转换文字方法也可以用于任何工作和生活的场景中。以智能音箱举例,如果智能音箱有需要进行文字输入的时候,就会通过语音提示需要进行文字输入。如果智能音箱提示需要进行文字输入就表征光标已经位于当前需要进行输入的输入框内。

[0033] 图2示意性示出了根据本公开实施例的语音转换文字方法的流程图。

[0034] 如图2所示,本公开实施例的语音转换文字的方法,包括操作S100~操作S300:

[0035] 在操作S100,监听用户的语音输入。

[0036] 其中,监听可以是指指在电子设备处于运行状态期间,一直利用声音接收装置来实时获取用户语音信息,直至电子设备进入待机状态、休眠状态或关机。

[0037] 在操作S200,基于语音识别引擎将语音输入转换为文字并存储。

[0038] 其中,语音识别引擎可以是指指语音识别算法。利用语音识别算法将监听到的语音输入转换为文字并将文字信息进行缓存。

[0039] 语音识别算法可以存储于本地电子设备中,并由电子设备来执行语音转换文字的过程。

[0040] 此外,语音识别算法也可以存储于云端服务器中,并由服务器来执行语音转换文字的过程。根据本公开的实施例,操作S200可以包括操作S210~操作S220:

[0041] 在操作S210,将监听的语音输入传输至云端服务器;

[0042] 在操作S220,接收云端服务器传输的文字并存储,其中,文字由云端服务器基于语音识别引擎将语音输入转换得到。

[0043] 在得到语音转换的文字后,需要检测界面中输入光标是否位于输入框中。

[0044] 在操作S300,如果输入光标位于输入框内将语音识别引擎转换的文字输入到输入框内。

[0045] 输入框可以是指指任何可以输入文字的区域,例如Word的页面区域、PPT中的文本框、Excel中的单元格、网页中的输入框等。

[0046] 在界面上的输入光标位于输入框的情况下,可以根据预设的规则,将语音转换的文字输入到输入框内。例如,若界面上显示有记事本,且输入光标位于记事本的文字输入区域中,则可以将语音转换的文字输入到输入光标所在的位置。

[0047] 根据本公开的实施例,电子设备只需处于运行状态,就可以实时监听用户语音信息并转换为文字输入到界面的输入框中。不需要额外安装或启用任何语音输入文字的应用程序,在界面上任何提示需要输入文字的位置,都可以随时通过语音输入文字,例如网页、Word、PPT等界面上的输入框。因此,本公开实施例的语音转换文字方法可以完全解放用户双手,代替键盘实现文字输入功能。

[0048] 根据本公开的实施例,操作S300可以包括操作S310和操作S320中的至少一种:

[0049] 在操作S310,如果输入光标位于输入框内,将第一部分文字输入到输入框,第一部分文字为语音识别引擎基于输入光标位于输入框之后所监听到的语音输入转换的文字。

[0050] 例如,在第一时刻(例如可以是12:00:00),输入光标被置于输入框中,在第一时刻之后的第二时刻(例如可以是12:05:00),开始接收到用户的语音输入,在这种情况下,可以将从第二时刻开始到输入光标离开输入框的时刻(例如可以是12:10:00)之间的时间段(例如可以是12:05:00~12:10:00)内接收到的语音输入转换的文字作为第一部分文字,输入到输入光标所在的输入框中。

[0051] 在操作S320,如果输入光标位于输入框内,将第二部分文字输入到输入框,第二部分文字为语音识别引擎基于输入光标位于输入框之前以及输入光标位于输入框之后的一完整单元的语音输入转换的文字。

[0052] 其中,完整单元可以是指具有完整语义的语句,可以是一句话,也可以是具有前后关联性的多句话。

[0053] 第二部分文字为语音识别引擎基于输入光标位于输入框之前以及输入光标位于输入框之后的一完整单元的语音输入转换的文字可以是指:在属于一个完整单元内的语句进行语音输入过程中,输入光标被置入输入框中或者输入光标被从输入框中移出,这种情况下,将该完整单元的语音输入转换的所有文字作为第二部分文字。

[0054] 例如,用户的语音输入例如是“今天开会讨论关于×××的事情”,该句话中的前后文字之间具有关联性,后边的文字可以对前边的文字进行校正,以使整个句子表达出正确且完整的含义,因此,语音识别引擎将该句话识别为一完整单元。若在麦克风接收到“会”字之前输入光标没有位于输入框中,而在麦克风接收到“会”字的时刻,检测到输入光标被置于输入框中,则需要将该完整单元的语音转换成的所有文字作为第二部分文字,即第二部分文字为“今天开会讨论关于×××的事情”,全部输入到输入框中。或者,若在接收到“会”之前,检测到输入光标位于输入框中,而在接收到“会”字的时刻,检测到输入光标离开了输入框,则同样需要将该完整单元的语音转换成的所有文字作为第二部分文字,即第二部分文字为“今天开会讨论关于×××的事情”,全部输入到输入框中。

[0055] 根据本公开的实施例,操作S300还可以包括操作S330:

[0056] 在操作S330,如果输入光标位于输入框内,将第三部分文字输入到输入框,第三部分文字为语音识别引擎基于输入光标位于输入框之前基于所监听到的语音输入转换完成的文字。

[0057] 其中,第三部分文字可以是指输入光标位于输入框之前已经转换完成的文字。例如,在第一时刻(例如可以是10:00:00)和第二时刻(例如可以是10:10:00)之间接收到用户的一段语音输入,该段语音输入在极短的时间内被转换成文字并对文字进行缓存。在第二时刻之后的第三时刻(例如可以是10:12:00),输入光标被置于输入框中,则可以将第一

时刻和第二时刻之间获得的语音输入转换完成的文字作为第三部分文字。第三部分文字可以被全部输入输入框中,也可以有选择性地被输入输入框中。

[0058] 在由录音笔执行语音转换文字的方法的实施例中,可以通过接收到第二电子设备发送的输入指令的时间来确定输入光标位于输入框中的时间,例如可以将接收到第二电子设备发送的输入指令的时间作为输入光标位于输入框中的时间。从而,录音笔可以基于其确定的输入光标位于输入框中的时间来确定将哪部分文字传输至第二电子设备,以使第二电子设备将相应部分的文字输入光标所在的输入框中。

[0059] 根据本公开的实施例,操作S330包括操作S331:

[0060] 在操作S331,如果输入光标位于输入框内,输出至少一个选项,以使得用户选择语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成的文字。

[0061] 在输入光标被置入输入框中时,可以在显示界面上显示至少一个选项,该至少一个选项可以是关于在输入光标位于输入框之前已经转换完成的文字的选项,以使用户可以选择将缓存中的哪些文字输入到输入框中。

[0062] 选项可以是以下拉菜单的形式呈现,也可以是对话框的形式呈现,或者是以快捷菜单的形式呈现。选项显示的内容可以是筛选条件,也可以是筛选条件对应的文字。

[0063] 在用户选择之后,可以基于用户选择的选项,将选项对应的文字作为第三部分文字输入到输入框。

[0064] 根据本公开的实施例,操作S331包括操作S3311~操作S3314中的至少一种:

[0065] 在操作S3311,选取与输入光标位于输入框时间上满足时间范围的语句所对应的文字作为选项。

[0066] 与输入光标位于输入框时间上满足时间范围的语句所对应的文字可以是指输入光标位于输入框之前的最近的一段时间内转换完成的文字。例如,可以是将输入光标位于输入框之前的一分钟或五分钟等时间段内接收的语音转换完成的文字作为选项。

[0067] 在操作S3312,从语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字中智能选取重要语句所对应的文字作为选项。

[0068] 其中,重要语句可以由语音识别算法识别出的对用户较为重要的语句。例如,用户重复某个词语或某句话的次数较多,语音识别算法可以将其识别为关键词或关键语句,则与该关键词或关键语句相关的语句可以被认为的重要语句。在输入光标位于输入框中时,可以在显示界面上显示至少一条重要语句作为选项。

[0069] 在操作S3313,从语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字中智能选取与输入框属性匹配的语音所对应的文字作为选项。

[0070] 其中,输入框属性例如可以是指输入框内需要输入特定内容,该特定内容可以作为该输入框的属性。与输入框属性匹配的语音所对应的文字可以是指语音所对应的文字与输入框的特定内容相匹配。特定内容例如可以是姓名、手机号、地址、身份证号等内容。

[0071] 例如,某网页中具有输入框A,输入框A内需要输入身份证号。若在输入光标位于该输入框A之前,监听到的语音输入转换完成的文字中包含身份证号信息,则在输入光标被置入该输入框A时,可以显示转换完成的身份证号作为选项。

[0072] 在操作S3314,将语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字的数量标识作为选项。

[0073] 将输入光标位于输入框之前所监听到的语音输入转换完成文字的数量标识作为选项可以是指：将输入光标位于输入框之前接收的语音转换完成的预定数量的文字作为选项。例如，可以将输入光标位于输入框之前转换完成的100个文字或500个文字作为选项。

[0074] 图3示意性示出了根据本公开另一实施例的语音转换文字方法的流程图。

[0075] 如图3所示，根据本公开的实施例，语音转换文字的方法还包括操作S400：

[0076] 在操作S400，按照丢弃策略删除实时监听用户的语音输入转换的文字。

[0077] 丢弃策略例如可以是删除不重要的信息，保留重要的信息，重要的信息例如可以是上述的重要语句。

[0078] 丢弃策略例如可以是丢弃已添加到输入框中的文字，例如，对于已输入到Word等文档中的文字，这些文字的存储地址相应改变为Word等文档的存储地址，所以可以删除缓存中的相应文字。再例如，对于已输入到微信等社交应用中输入框的文字，点击发送后，这些文字会被保存至服务器中，所以同样可以删除缓存中的相应文字。

[0079] 本公开实施例还提供了一种语音转换文字的系统2。

[0080] 图4示意性示出了根据本公开实施例的语音转换文字系统的方框图。

[0081] 如图4所示，本公开实施例的语音转换文字的系统包括：

[0082] 监听模块100，用于监听用户的语音输入；

[0083] 转换模块200，用于基于语音识别引擎将语音输入转换为文字并存储；

[0084] 输入模块300，用于在输入光标位于输入框内的情况下，将语音识别引擎转换的文字输入到输入框内。

[0085] 具体地，监听模块100例如可以执行上文描述的操作S100，转换模块200例如可以执行上文描述的操作S200，输入模块300例如可以执行上文描述的操作S300，在此不再赘述。

[0086] 根据本公开的实施例，输入模块300可以包括以下子模块中的任意一种：第一输入子模块，用于在输入光标位于输入框的情况下，将第一部分文字输入到输入框，第一部分文字为语音识别引擎基于输入光标位于输入框之后所监听到的语音输入转换的文字；第二输入子模块，用于在输入光标位于输入框的情况下，将第二部分文字输入到输入框，第二部分文字为语音识别引擎基于输入光标位于输入框之前以及输入光标位于输入框之后的一完整单元的语音输入转换的文字。

[0087] 具体地，第一输入子模块例如可以执行上文描述的操作S310，第二输入子模块例如可以执行上文描述的操作S320，在此不再赘述。

[0088] 根据本公开的实施例，输入模块300还可以包括：第三输入子模块，用于在输入光标位于输入框的情况下，将第三部分文字输入到输入框，第三部分文字为语音识别引擎基于输入光标位于输入框之前基于所监听到的语音输入转换完成的文字。

[0089] 具体地，第三输入子模块例如可以执行上文描述的操作S330，在此不再赘述。

[0090] 根据本公开的实施例，第三输入子模块可以包括：选项子模块，用于在输入光标位于输入框的情况下，输出至少一个选项，以使得用户选择语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成的文字；以及基于用户选择的选项，将选项对应的文字作为第三部分文字输入到输入框。

[0091] 具体地，选项子模块例如可以执行上文描述的操作S331，在此不再赘述。

[0092] 根据本公开的实施例,输出至少一个选项包括以下形式中的至少一种:选取与输入光标位于输入框时间上满足时间范围的语句所对应的文字作为选项;从语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字中智能选取重要语句所对应的文字作为选项;从语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字中智能选取与输入框属性匹配的语音所对应的文字作为选项;将语音识别引擎基于输入光标位于输入框之前所监听到的语音输入转换完成文字的数量标识作为选项。

[0093] 根据本公开的实施例,语音转换文字系统还可以包括:删除模块,用于按照丢弃策略删除实时监听用户的语音输入并基于语音识别引擎将语音输入转换的文字。

[0094] 具体地,删除模块例如可以执行上文描述的操作S400,在此不再赘述。

[0095] 根据本公开的实施例,转换模块200可以包括:传输子模块,用于将监听的语音输入传输至服务器;接收子模块,用于接收服务器传输的文字并存储,其中,文字由服务器基于语音识别引擎将语音输入转换得到。

[0096] 根据本公开的实施例的模块、子模块、单元、子单元中的任意多个、或其中任意多个的至少部分功能可以在一个模块中实现。根据本公开实施例的模块、子模块、单元、子单元中的任意一个或多个可以被拆分成多个模块来实现。根据本公开实施例的模块、子模块、单元、子单元中的任意一个或多个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式的硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,根据本公开实施例的模块、子模块、单元、子单元中的一个或多个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0097] 例如,上述的各个模块和子模块中的任意多个可以合并在一个模块中实现,或者其中的任意一个模块可以被拆分成多个模块。或者,这些模块中的一个或多个模块的至少部分功能可以与其他模块的至少部分功能相结合,并在一个模块中实现。根据本公开的实施例,上述的各个模块和子模块中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,上述的各个模块和子模块中的至少一个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0098] 本公开实施例还提供了一种电子设备。

[0099] 图5示意性示出了根据本公开实施例的电子设备的方框图。

[0100] 如图5所示,电子设备500包括:

[0101] 麦克风510;

[0102] 处理器520,用于控制麦克风510监听用户的语音输入;以及在输入光标位于输入框的情况下将语音输入转换的文字输入到输入框内。

[0103] 存储单元530,用于存储通过语音识别引擎将麦克风监听的语音输入转换的文字。

[0104] 其中,电子设备500例如可以是笔记本电脑。

[0105] 在电子设备500运行状态下,麦克风510一直处于工作状态,处理器520实时从麦克

风510接收模拟语音信号,并将模拟语音信号转换为数字语音信号。

[0106] 通过语音识别引擎将数字语音信号转换为文字信息后,利用存储单元530将转换的文字进行存储。

[0107] 处理器520还用于执行:检测显示界面中输入光标是否位于输入框中。在输入光标位于输入框的情况下,将语音输入转换的文字输入到输入框内。

[0108] 根据本公开的实施例,处理器520还用于执行:基于语音识别引擎将监听的语音输入转换为文字并存储至处理器。

[0109] 语音识别引擎可以位于本地,例如可以位于处理器520中,由处理器520来执行数字语音信号转换为文字的过程。转换完成后再由处理器520将文字存储至存储单元530。

[0110] 根据本公开的实施例,处理器520还用于执行:将监听的语音输入传输至服务器,并从服务器接收文字并存储,其中,文字由服务器基于语音识别引擎将语音输入转换得到。

[0111] 语音识别引擎也可以位于云端服务器中,并由云端服务器来执行语音转换文字的过程。处理器20可以将数字语音信号传输至云端服务器。云端服务器基于语音识别引擎将数字语音信号转换为文字信号,并将文字信号传输回处理器20,处理器20将从云端服务器接收到的文字存储至存储单元30中。

[0112] 根据本公开的实施例,处理器20例如可以包括通用微处理器、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC)),等等。处理器20还可以包括用于缓存用途的板载存储器。处理器20可以是用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0113] 图6示意性示出了根据本公开另一实施例的电子设备的方框图。

[0114] 如图6所示,电子设备500还可以包括计算机可读存储介质540、信号发送器550和信号接收器560。

[0115] 计算机可读存储介质540例如可以是非易失性的计算机可读存储介质,具体示例包括但不限于:磁存储装置,如磁带或硬盘(HDD);光存储装置,如光盘(CD-ROM);存储器,如随机存取存储器(RAM)或闪存;等等。

[0116] 计算机可读存储介质540可以包括计算机程序541,该计算机程序541可以包括代码/计算机可执行指令,其在由处理器20执行时使得处理器20执行根据本公开实施例的方法或其任何变形。

[0117] 计算机程序541可被配置为具有例如包括计算机程序模块的计算机程序代码。例如,在示例实施例中,计算机程序中的代码可以包括一个或多个程序模块,例如模块541A、541B等。应当注意,模块的划分方式和个数并不是固定的,本领域技术人员可以根据实际情况使用合适的程序模块或程序模块组合,当这些程序模块组合被处理器20执行时,使得处理器20可以执行根据本公开实施例的方法或其任何变形。

[0118] 处理器520可以与信号发送器550和信号接收器560进行交互,来执行根据本公开实施例的方法或其任何变形。

[0119] 根据本发明的实施例,上述各个模块和子模块中的至少一个可以实现为计算机程序模块,其在被处理器520执行时,可以实现上面描述的相应操作。

[0120] 本公开还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描述的设备/装置/系统中所包含的;也可以是单独存在,而未装配入该设备/装置/

系统中。上述计算机可读存储介质承载有一个或者多个程序，当上述一个或者多个程序被执行时，实现根据本公开实施例的方法。

[0121] 根据本公开的实施例，计算机可读存储介质可以是非易失性的计算机可读存储介质，例如可以包括但不限于：便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (ROM)、可擦式可编程只读存储器 (EPROM或闪存)、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0122] 附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0123] 本领域技术人员可以理解，本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合，即使这样的组合或结合没有明确记载于本公开中。特别地，在不脱离本公开精神和教导的情况下，本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0124] 尽管已经参照本公开的特定示例性实施例示出并描述了本公开，但是本领域技术人员应该理解，在不背离所附权利要求及其等同物限定的本公开的精神和范围的情况下，可以对本公开进行形式和细节上的多种改变。因此，本公开的范围不应该限于上述实施例，而是应该不仅由所附权利要求来进行确定，还由所附权利要求的等同物来进行限定。

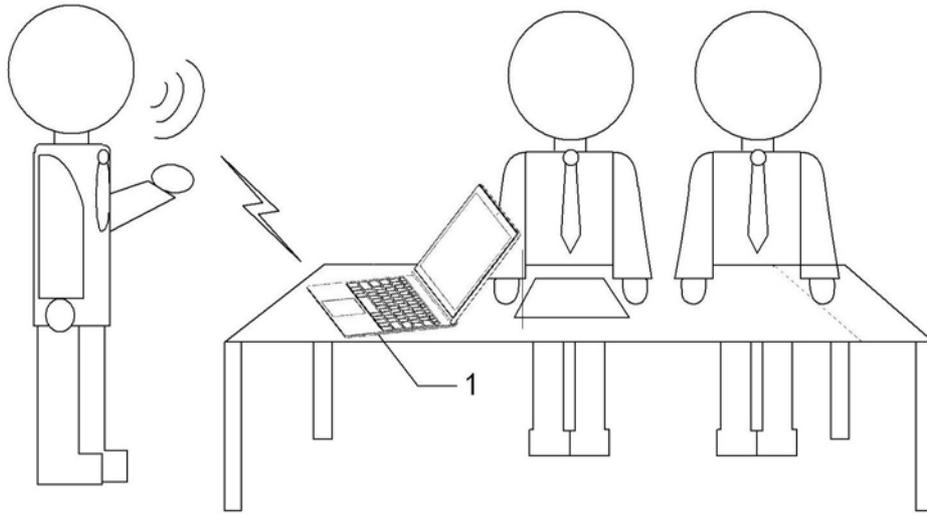


图1

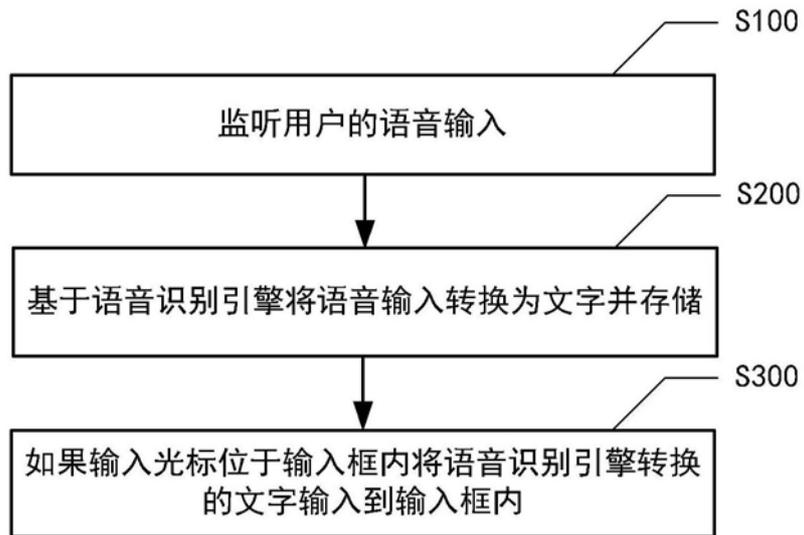


图2

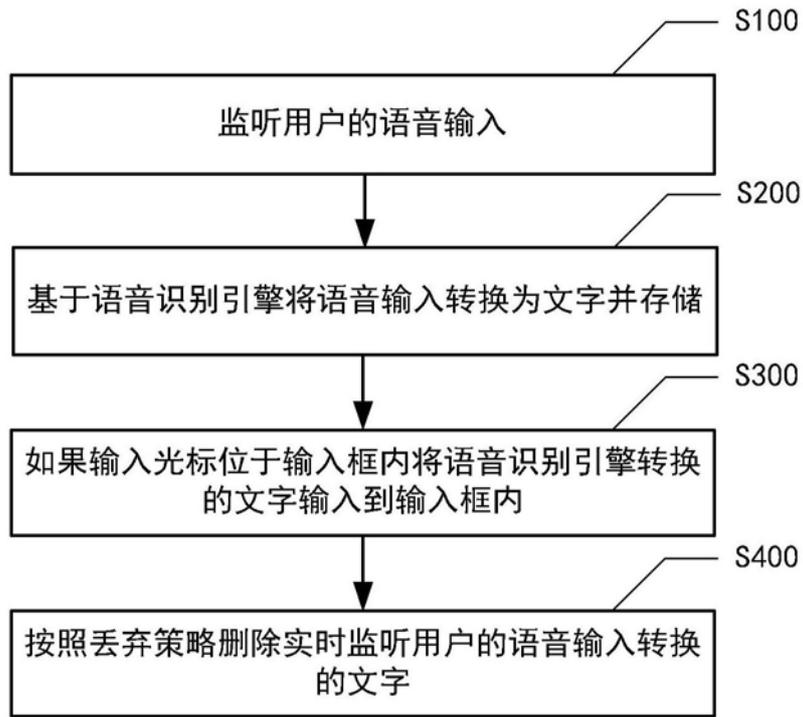


图3

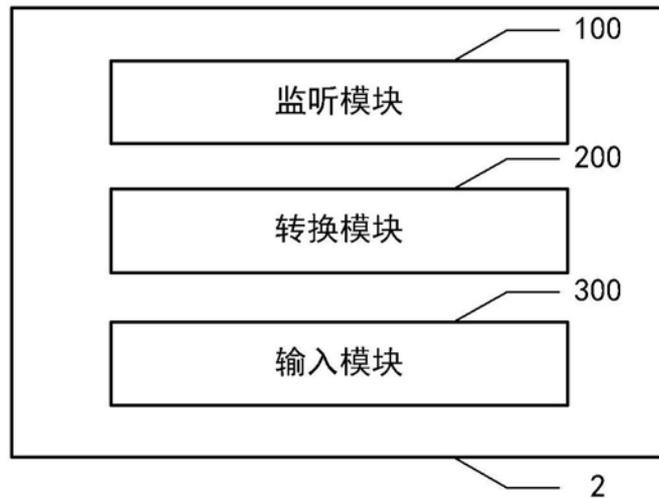


图4

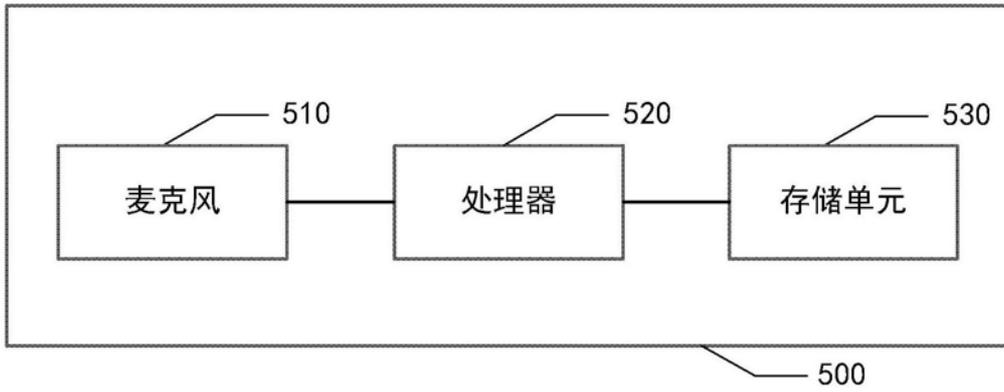


图5

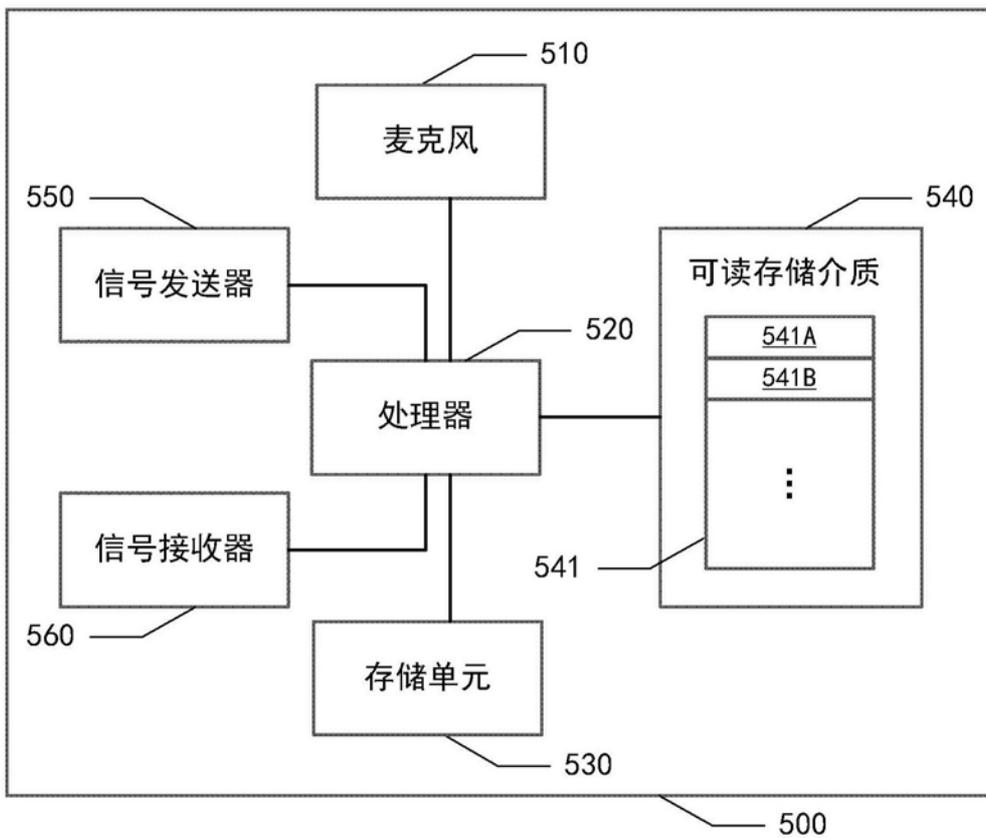


图6