



(12) 发明专利申请

(10) 申请公布号 CN 104508604 A

(43) 申请公布日 2015. 04. 08

(21) 申请号 201380040756. 2

代理人 刘鹏 景军平

(22) 申请日 2013. 07. 30

(51) Int. Cl.

(30) 优先权数据

G06F 3/023(2006. 01)

13/561095 2012. 07. 30 US

(85) PCT国际申请进入国家阶段日

2015. 01. 30

(86) PCT国际申请的申请数据

PCT/US2013/052600 2013. 07. 30

(87) PCT国际申请的公布数据

W02014/022322 EN 2014. 02. 06

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 J. 格里夫斯 G. 克里沃希夫

D. 鲁钦科 P. 森达拉拉彦

T. 佩克 I. 阿尔莫格

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

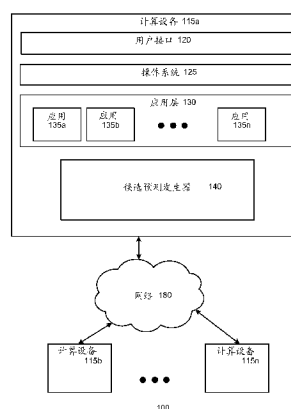
权利要求书1页 说明书9页 附图5页

(54) 发明名称

使用上下文生成字符串预测

(57) 摘要

在移动设备中,为移动设备确定上下文。基于移动设备环境的多种特性来确定上下文,所述特性包括例如正在使用的当前应用、移动设备的用户正在与其交互或者与其对话的任何联系人、当前日期和/或时间、对话的当前话题、对话的当前风格等等。基于与所确定的上下文相关联的一组字符串和用户生成的文本,为用户生成的文本生成一个或者多个字符串预测。字符串预测可以作为用户生成的文本的所建议的完成向用户呈现。



1. 一种方法,包括:

由计算设备接收字符串,其中所接收的字符串与对话相关联;

由计算设备基于对话来为所接收的字符串确定上下文;

由计算设备基于所接收的字符串和所确定的上下文来生成多个字符串预测,其中每个字符串预测包括字符串;

由计算设备提供所述多个字符串预测中的字符串中的一个或者多个;

由计算设备接收所述多个字符串预测中的所提供的一个或者多个字符串中的一个的选择的指示;以及

响应于选择的指示,由计算设备向应用提供所选择的字符串。

2. 如权利要求 1 的方法,其中为所接收的字符串确定上下文包括,基于应用确定上下文。

3. 如权利要求 1 的方法,其中上下文与多个字符串相关联,并且所生成的字符串预测中的至少一个包括,来自与上下文相关联的所述多个字符串中的字符串。

4. 如权利要求 1 的方法,其中所接收的字符串是字符。

5. 如权利要求 1 的方法,其中对话具有一个或者多个参与者,并且为所接收的字符串确定上下文包括,基于所述一个或者多个参与者确定上下文。

6. 如权利要求 1 的方法,其中为所接收的字符串确定上下文包括,确定对话的话题,并且基于对话为所接收的字符串确定上下文进一步包括基于所确定的话题为所接收的字符串确定上下文。

7. 如权利要求 1 的方法,其中为所接收的字符串确定上下文包括,确定对话的风格并且基于所确定的风格为所接收的字符串确定上下文。

8. 一种系统,包括:

至少一个计算设备;以及

候选预测发生器,被适配于:

接收字符串,其中所接收的字符串与一个或者多个参与者相关联;

基于所述一个或者多个参与者为所接收的字符串确定上下文;

基于所接收的字符串和所确定的上下文生成多个字符串预测,其中每个字符串预测包括字符串;以及

提供所述多个字符串预测中的字符串中的一个或者多个。

9. 如权利要求 8 的系统,其中所述候选预测发生器进一步被适配为:

接收所述多个字符串预测中的所提供的一个或者多个字符串中的一个的选择的指示;以及

响应于选择的指示,向应用提供所选择的字符串。

10. 如权利要求 8 的系统,其中所接收的字符串是对话的一部分,并且为所接收的字符串确定上下文进一步包括确定对话的话题,并且为所接收的字符串确定上下文是基于所确定的话题的。

使用上下文生成字符串预测

背景技术

[0001] 键入是移动设备上许多场景的一部分,包括例如写文本消息、写电子邮件、搜索联系人、和键入 URL (统一资源定位符)。移动设备上的键入可能是缓慢和令人沮丧的。这在使用屏幕上的软件键盘时可能尤为真实。

[0002] 对于这个问题的一个解决方案被熟知为软件输入面板。在软件输入面板中,基于用户到目前为止已经输入到文本输入栏中的字符,屏幕的区域可以显示所建议的字预测。随着用户继续输入字符到该栏中,所建议的字改变。当用户看到与用户正在输入的字相匹配的字预测时,用户可以选择所述字预测,并且所选择的字作为对于由用户输入的字符的替换,被放置在文本输入栏中。

[0003] 典型地,通过将输入的字符与字典中的字相匹配并且随后基于每个字是用户打算输入的字的的可能性呈现匹配字的经排序的子集,来生成向用户展示的字预测。所述可能性可能是基于每个匹配字被普通人群使用的总概率的。

[0004] 这样的用于提供字预测的方法的一个缺陷是:其未能考虑到与正输入的字符相关联的上下文。上下文可以包括用户正在使用的应用(即,短消息服务(SMS)或者电子邮件应用)、以及用户正在与其通信的参与者、用户的位置、日期或者一天中的时间、用户较早与其通信的人、或者甚至是与用户相关联的情绪。典型地,用户使用可以依据对话的上下文而变化的一组词汇。例如,与当与朋友进行对话时相比,用户在与商务联系人进行对话时可以使用不同的一组短语,或者与当使用字处理应用时相比,用户在使用 SMS 应用时可以使用不同的一组短语。由于在生成预测时未考虑上下文,该用户可能接收到过于宽泛的、或者对于特定上下文而言是不合适的字预测。

发明内容

[0005] 在移动设备中,为移动设备确定上下文。上下文基于移动设备环境的特性被确定,并可以包括例如正被使用的当前应用、移动设备的用户正在与其交互的或者与其进行对话的任何联系人、当前日期和 / 或时间、对话的当前话题、对话的当前风格等等。基于与所确定的上下文相关联的一组字符串和用户生成的文本,为用户生成的文本生成一个或者多个字符串预测。字符串预测可以作为用户生成文本的所建议的完成来呈现给用户。可替换地或者附加地,可以基于所确定的上下文来生成语音注释或者表情符号。

[0006] 在一个实施方案中,由计算设备接收字符串。所接收的字符串与应用相关联。针对所接收的字符串的上下文由计算设备确定。基于所接收的字符串和所确定的上下文来生成字符串预测。每个字符串预测包括字符串。字符串预测中的字符串中的一个或者多个由计算设备提供。所提供的字符串中的一个的选择的指示由计算设备接收。响应于选择的指示,所选择的字符串由计算设备提供给应用。

[0007] 在一个实施方案中,为所接收的字符串确定上下文。确定所接收的字符串是否满足标准。如果所接收的字符串满足标准,则将所接收的字符串添加到与上下文相关联的多个字符串中。

[0008] 提供本发明内容,以便以简化的形式引入概念的选择,将在下文的具体实施方式中进一步描述所述概念。本发明内容不打算标识所要求保护的主题的关键特征或者本质特征,也不打算用来限制所要求保护的主题的范围。

附图说明

[0009] 当连同附图进行阅读时,前述发明内容以及下文的说明性实施例的详细描述被更好地理解。出于图示实施例的目的,在图中示出了实施例的示例构造;然而,实施例不限于所公开的具体方法和手段。在图中:

图 1 是用于使用上下文选择和呈现字符串预测的示例环境的图示;

图 2 是示例候选预测发生器的图示;

图 3 是用于使用上下文提供字符串预测的方法的实施方案的操作流程;

图 4 是用于更新上下文的方法的实施方案的操作流程;以及

图 5 示出其中可以实施示例实施例和方面的示例性计算环境。

具体实施方式

[0010] 图 1 是用于基于上下文选择和呈现字符串预测的示例环境 100 的图示。环境 100 可以包括多个计算设备 115 (例如,计算设备 115a-n)。计算设备 115 可以是包括移动设备的多种计算设备,移动设备例如是蜂窝电话、智能电话、便携式媒体播放器、游戏控制台等等。计算设备 115 可以通过网络 180 彼此通信。网络 180 可以是多种网络类型,包括公共交换电话网络(PSTN)、蜂窝电话网络和分组交换网络(例如,互联网)。适当的计算设备 115 的示例包括关于图 5 描述的计算机系统 500。

[0011] 计算设备 115 中的每个可以包括操作系统 125 和用户接口 120。操作系统 125 可以包括多种众所周知的操作系统。计算设备 115 的用户可以使用用户接口 120 来向操作系统 125 输入一个或者多个字符。用户接口 120 可以包括多种接口,包括例如软件键盘或者硬件键盘。

[0012] 计算设备 115 中的每个还可以执行一个或者多个应用 135(例如,应用 135a-n)。应用 135 可以在应用层 130 中执行。应用 135 可以包括例如电子邮件应用、短消息服务(SMS)或者文本应用、聊天应用、社交网络应用和字处理应用。计算设备 115 可以使用应用 135 来交换基于文本的消息。例如,每个计算设备可以执行 SMS 应用并且可以通过应用发送和接收 SMS 消息。有关的基于文本的消息的序列在本文中被称为对话。

[0013] 计算设备的每次使用可以通过用户接口 120 和操作系统 125 与应用 135 进行交互。在一个实施方案中,用户可以通过用户接口 120 的软输入面板输入字符。所输入的字符可以由操作系统 125 来接收,并且可以被提供给应用 135 中的一个或者多个。

[0014] 对于计算设备 115 的用户而言,逐字符地将文本输入到用户接口 120 中可能是令人沮丧的。相应地,计算设备 115 还可以包括候选预测发生器 140。候选预测发生器 140 可以生成一个或者多个字符串预测。字符串预测可以包括候选预测发生器 140 预测用户打算输入到用户接口 120 中的字符串。字符串预测可以包括字符串的置信度值。可替换地或者附加地,候选预测发生器 140 可以生成语音注释、表情符号、或者其它通信或者消息类型。所预测的字符串可以包括一个或者多个字符的任何序列,包括但不限于字、短语、句、表

情符号、标点记号或者其它符号、或者其组合。

[0015] 如关于图 2 进一步描述的,字符串预测可以部分地基于与所输入的字符或者字符串相关联的一个或者多个上下文。如本文所使用的上下文可以包括与所输入的字符或者字符串相关联的对话或者环境的一个或者多个特性、特征或者属性。上下文的特性或者特征可以例如包括:所接收的文本所针对的特定应用;正在参加对话的参与者、联系人、或者其它用户;当前日期和/或时间;对话的气氛(即,悲伤或者快乐);对话的话题;和对话的风格(即,正式或者非正式);用户的位置;用户较早与其通信的用户;和用户的情绪。

[0016] 每个上下文可以与一组字符串相关联。与上下文相关联的该组字符串可以包括由用户在上下文期间使用或者提供的字符串、或者以其他方式已知为与上下文相关联的字符串。例如,与 SMS 应用相关联的上下文可以与频繁用于文本消息的字符串(诸如“LOL”)相关联。与诸如 1 月 1 日的日期相关联的上下文可以包括与新年假期相关联的字符串,诸如“新年快乐”。与用户的特定联系人相关联的上下文可以包括当用户与该联系人进行交互时输入过的一个或者多个字符串。与上下文相关联的该组字符串中的字符串可以与概率相关联。字符串的概率可以基于由用户或者其他用户针对该上下文提供字符串的次数。

[0017] 在一些实施方案中,当用户输入字符串的一个或者多个字符时,候选预测发生器 140 可以确定所输入的字符串的上下文,并且可以基于所输入的字符串和所确定的上下文来生成一组字符串预测。在一些实施方案中,字符串预测可以包括来自与上下文相关联的该组字符串中的将所输入的字符串作为前缀的一个或者多个字符串,或者以其他方式与所输入的字符串相似的一个或者多个字符串。附加地,字符串预测也可以例如包括基于一个或者多个模型(诸如词典)而生成的字符串预测。

[0018] 通过基于上下文生成字符串预测,而不是唯一地基于词典或者模型生成字符串预测,所生成的字符串预测不限于在词典或者模型中使用的字和短语。而且,因为基于上下文的字符串预测可以比基于词典或者模型的字符串预测更精细地定制,所以字符串可以比基于词典或者模型的字符串预测中的字符串更准确地代表用户打算输入的字符串。

[0019] 图 2 是示例候选预测发生器 140 的图示。如所示出的,候选预测发生器 140 可以包括数个组件,包括但不限于上下文引擎 240 和预测模型 205。候选预测发生器 140 的组件中的一些或者全部可以由计算系统来实施,所述计算系统例如是图 5 中所图示的计算系统 500。

[0020] 候选预测发生器 140 可以从操作系统 125 接收所输入的文本 230。所输入的文本 230 可以包括字符串或者字符,并且可以是由计算设备 115 的用户提供给应用 135 的文本。所输入的文本 230 可以与用户打算或者曾打算提供给应用 135 的字符串有关,或者可以是其一部分。例如,所输入的文本 230 可以是用户打算输入的字符串的第一个字符,或者所输入的文本 230 可以是用户曾打算输入的字符串的拼写错误。因此,如可以领会的,候选预测发生器 140 可以用来向一个或者多个应用 135 提供短语预测和拼写建议两者。

[0021] 上下文引擎 240 可以为所输入的文本 230 确定一个或者多个上下文 210,并且可以从上下文存储装置 260 中获取所确定的一个或者多个上下文 210。在一些实施方案中,每个上下文 210 可以包括多个字符串或者可以与多个字符串相关联。与上下文 210 相关联的多个字符串可以包括由计算设备 115 的用户或者由其它计算设备 115 的其他用户输入的针对上下文 210 的字符串。例如,对于与字处理应用相关联的上下文 210 而言,多个字符串可以

包括由用户频繁输入到字处理应用中的一个或者多个字符串。对于与正式的对话风格相关联的上下文而言,多个字符串可以包括在正式对话期间由用户或者其他用户典型地使用的一个或者多个字符串。在一些实施方案中,与上下文 210 相关联的多个字符串中的字符串可以基于由计算设备的一个或者多个其他用户使用的字符串,并且可以由外部管理员收集并提供给上下文引擎 240。可替换地或者附加地,如下文将进一步描述的,与上下文 210 相关联的字符串可以由上下文引擎 240 确定。

[0022] 如上文所描述的,上下文可以包括用户通过其提供所输入的文本 230 的对话 220 或者环境的一个或者多个特性、特征或者属性。在一些实施方案中,上下文引擎 240 可以通过确定与所输入的文本 230 相关联的特性并且确定与所确定的特性相匹配的上下文存储装置 260 中的一个或者多个上下文 210,来为所输入的文本 230 确定一个或者多个上下文 210。

[0023] 上下文 210 的特性可以包括与所输入的文本 230 相关联的应用 135 的类型。与所输入的文本 230 相关联的应用 135 可以是用户正向其提供所输入的文本 230 的应用 135。应用 135 的类型可以包括例如 SMS 应用、字处理应用、电子邮件应用、和社交网络应用。其它类型的应用可以被支持。上下文引擎 240 可以例如从操作系统 120 确定应用 135 的类型。

[0024] 上下文 210 的特性可以包括与时间和 / 或日期有关的特性。如可以领会的,由用户提供的字或者短语的类型可以基于例如一天中的时间、一周中的一天、或者一个月中的一天而变化。例如,用户可能在下午使用短语“下午好”并且在早上使用短语“早上好”。用户在周末期间使用的字或者短语可能与其在工作日期间使用的字或者短语是不同的。类似地,用户可能更可能在围绕相应假期的一段时间期间使用某些与假期有关的字和短语,诸如“圣帕特里克(St. Patrick)节快乐”或者“母亲节快乐”。上下文引擎 240 可以例如从操作系统 120 确定与时间或者日期有关的特性。

[0025] 上下文 210 的特性还可以根据与所输入的文本 230 相关联的对话 220 来确定。在一些实施方案中,对话 220 可以包括来自用户和作为对话 220 的参与者的一个或者多个其他用户生成的基于文本的消息的一个或者多个字符串。例如,对话 220 可以是在计算设备 115 的用户和其它计算设备 115 的一个或者多个其他用户之间发送的一系列文本、电子邮件、或者社交网络消息。附加地,上下文的特性可以包括在对话期间提到或者提出的实体或者参与者。

[0026] 可以由上下文引擎 240 根据与所输入的文本 230 相关联的对话 220 确定的上下文 210 的特性的示例可以包括,与对话 220 相关联的一个或者多个参与者。例如,用户可以与两个朋友进行电子邮件对话。朋友可以是对话 220 的参与者。用户可以根据他们正在与其通信的参与者(例如,他们的母亲对朋友)而使用不同的字或者短语。上下文引擎 240 可以根据与所输入的文本 230 相关联的应用 135 确定对话 220 的参与者。

[0027] 可以由上下文引擎 240 根据与所输入的文本 230 相关联的对话 220 确定的上下文 210 的特性的另外的示例可以包括,对话 220 的风格或者气氛。对话 220 的风格的示例可以包括正式的、非正式的、或者商务的。气氛的示例可以包括严肃的、悲伤的、快乐的、和激动的。其它语气或者风格可以被支持。典型地,用户使用与对话 220 的风格或者气氛相匹配的对话 220 中的一组字或者短语。

[0028] 上下文引擎 240 可以基于在对话 220 中使用的字或者短语确定对话 220 的风格或

者气氛。在一些实施方案中,上下文引擎 240 可以寻找对话 220 中指示对话的风格或者气氛的字或者短语,并且可以基于所确定的字来确定对话 220 的风格或者气氛。例如,诸如“LOL”那样的非正式字的使用可以指示对话 220 的风格是非正式的,并且感叹词记号的使用可以指示对话的气氛很可能是快乐的或者激动的。在一些实施方案中,上下文引擎 240 可以连续监测对话 220,并且可以随着对话 220 的进行,更新或者改变对话 220 的风格或者气氛。

[0029] 可以由上下文引擎 240 根据与所输入的文本 230 相关联的对话 220 确定的上下文 210 的特性的另外的示例可以包括,对话 220 的话题。与上文描述的风格和气氛类似,上下文引擎 240 可以通过寻找对话 220 中的特定字或者短语来确定对话的话题。例如,对话 220 中的“纽约”和“旅行”的多次出现可以指示对话 220 是关于纽约或者休假的。上下文引擎 240 可以连续监测对话 220,并且可以随着对话 220 的进行,更新或者改变对话 220 的话题。

[0030] 在一些实施方案中,上下文引擎 240 可以从上下文存储装置 260 中获取与所确定的特性中的每个相匹配的上下文 210。因此,如果上下文引擎 240 确定与所输入的文本 230 相关联的上下文 210 具有正式气氛的特性,并且对话的参与者是用户的父母,则上下文引擎 240 可以从上下文存储装置 260 中获取具有正式气氛的特性的上下文 210,并且对话的参与者是用户的父母。如果没有这样的上下文 210 存在于上下文存储装置 260 中,则上下文引擎 240 可以选择具有与所确定的特性最接近或者最相似的特性的上下文 210。可替换地或者附加地,上下文引擎 240 可以从上下文存储装置 260 中为每个特性获取具有该特性的上下文 210,而不是从上下文存储装置 260 中获取具有所确定的所有特性的单个上下文 210。

[0031] 候选预测发生器 140 可以基于所输入的文本 230 和所确定的一个或者多个上下文 210 来生成一个或者多个字符串预测 215。在一些实施方案中,字符串预测 215 可以由候选预测发生器 140 通过从与一个或者多个上下文 210 中的每个相关联的多个字符串中确定将所输入的文本 230 作为前缀的字符串来生成。所确定的字符串的置信度值可以例如通过包括该所确定的字符串的一个或者多个上下文 210 的数量和与所述一个或者多个上下文 210 中的每个字符串相关联的概率来确定。

[0032] 在字符串预测 215 用作拼写建议的实施方案中,候选预测发生器 140 可以从所述多个字符串中确定与所输入的文本 230 相似的字符串,而不是从与一个或者多个上下文 210 中的每个相关联的所述多个字符串中确定将所输入的文本 230 作为前缀的字符串。用于确定字符串的相似性的任何方法可以被使用。可替换地,候选预测发生器 140 可以从与一个或者多个上下文 210 中的每个相关联的多个字符串中确定所输入的文本 230 是否包括任何可能的字符串拼写错误,并且可以使用所确定的字符串来生成字符串预测 215。

[0033] 候选预测发生器 140 可以提供所生成的字符串预测 215 中的一个或者多个。其中存在许多字符串预测 215,候选预测发生器 140 可以基于与每个字符串预测 215 相关联的置信度值来选择将提供哪些字符串预测 215。置信度值可以基于与关联于所确定的一个或者多个上下文 210 的字符串中的每个相关联的概率。对于短语预测实施方案而言,所提供的字符串预测 215 中的字符串可以在用户接口 120 的软输入面板中向用户显示。如果用户选择了所显示的字符串,则该字符串可以作为输入提供给应用 135。

[0034] 对于拼写建议实施方案而言,所提供的字符串预测 215 中的字符串可以靠近所输入的文本 230 或者作为菜单选择来显示。如果用户选择了所显示的字符串,则所选择的字

字符串可以在应用 135 中替换所输入的文本 230。

[0035] 除了上下文 210 之外,候选预测发生器 140 也可以使用预测模型 205 生成一个或者多个字符串预测 215。在一个实施方案中,预测模型 205 可以包括词典,并且可以包括多种字符串(例如,字或者短语)以及用于每个字符串的概率。每个字符串的概率可以反映例如一般使用时每个字符串的普及性。概率可以通过对由用户经过一段时间为一个或者多个应用 135 在一个或者多个对话 220 中提供的字符串进行收集和分析来确定,或者可以基于对来自例如报纸、杂志、或者其它语言源的字符串的分析。模型的其它类型可以被使用。

[0036] 在一些实施方案中,候选预测发生器 140 可以将使用一个或者多个上下文 210 生成的字符串预测 215 与由预测模型 205 生成的字符串预测 215 组合。例如,候选预测发生器 140 可以选择使用一个或者多个上下文 210 生成的最前面的字符串预测(即,具有最高的置信度值)以及预测模型 205 的最前面的字符串预测,并且组合所选择的字符串预测。

[0037] 在其它的实施方案中,候选预测发生器 140 可以使用一个或者多个上下文 210 和预测模型 205 生成一个或者多个字符串预测 215。例如,与一个或者多个上下文 210 相关联的多个字符串中的字符串可以被合并到模型 205 中,或者与模型 205 中的字符串相关联的概率可以基于与一个或多个上下文 210 相关联的字符串来调整。由候选预测发生器 140 对预测模型 205 做出的任何改变或者调整可以是暂时的,并且可以在上下文引擎 240 确定一个或者多个不同上下文 210 的情况下被反转。

[0038] 上下文引擎 240 可以基于所输入的文本 230 来保持或者更新所确定的上下文 210。上下文引擎 240 可以从应用 135 接收字符串。该字符串可以是提供给应用的字符串,或者作为与应用 135 相关联的对话 220 的一部分的字符串。例如,计算设备 115 的用户可以将所输入的文本 230 “愿力量与你同在”提供到 SMS 应用中,作为具有特定参与者的对话 220 的一部分。与所输入的文本 230 相关联的(多个)上下文可以由上下文引擎 240 使用 SMS 应用和与对话 220 相关联的参与者的特性来确定。

[0039] 上下文引擎 240 可以确定所输入的文本 230 的字符串是否已经是与一个或者多个上下文 210 相关联的字符串的一部分。可替换地或者附加地,上下文引擎 240 可以确定与所输入的文本 230 的字符串相似的字符串是否已经是与一个或者多个上下文 210 中的每个相关联的字符串的一部分。用于确定相似字符串的任何方法可以被使用。如果所输入的文本 230 的字符串(或者相似的字符串)已经是与一个或者多个上下文 210 相关联的字符串的一部分,则在与一个或者多个上下文 210 相关联的多个字符串中的每个中的字符串(或者相似的字符串)的概率可以增加。例如概率可以增加固定量。

[0040] 如果所输入的文本 230 的字符串(或者相似的字符串)不是与一个或者多个上下文 210 相关联的字符串的一部分,则上下文引擎 240 可以确定是否将所输入的文本 230 的字符串添加到与一个或者多个上下文 210 相关联的多个字符串中。在一些实施方案中,如果所输入的文本 230 的字符串满足与上下文 210 相关联的标准,则其可以被添加到与上下文 210 相关联的多个字符串中。例如,标准可以是字符串作为所输入的文本 230 的一部分多于阈值次数,比如三次。其它阈值次数或者标准可以被使用。每个上下文 210 可以有其自身的标准,或者对于上下文存储装置 260 中的每个上下文 210 而言,标准可以是相同的。因此,继续上文的示例,如果对于上下文 210 中的每个而言,字符串“愿力量与你同在”被接收到超过三次,则字符串“愿力量与你同在”可以被添加到用于 SMS 应用的上下文 210 和用

于与对话 220 相关联的参与者的上下文 210 中。

[0041] 图 3 是用于使用一个或者多个上下文提供字符串预测的方法 300 的实施方案的操作流程。方法 300 可以由例如计算设备 115 的候选预测发生器 140 来实施。

[0042] 在 301 处,接收字符串。字符串可以由候选预测发生器 140 通过用户接口 120 从用户接收,作为所输入的文本 230。例如,用户可能已经使用软件键盘将一个或者多个字符输入到文本栏中。所接收的字符串可以是用户打算作为具有一个或者多个参与者的对话 220 的一部分向应用 135 提供的字符串的前缀。应用 135 可以是例如 SMS 应用、社交网络应用和电子邮件应用中的一个或者多个。

[0043] 在 303 处,为字符串确定上下文。上下文 210 可以通过候选预测发生器 140 的上下文引擎 240 来确定。上下文引擎 240 可以针对所接收的字符串确定单个上下文 210,或者可以确定多个上下文 210。上下文 210 可以根据存储在上下文存储装置 260 中的上下文 210 来确定。每个上下文 210 可以与多个字符串相关联,所述多个字符串是在特定上下文 210 期间由用户或者其他用户所频繁提供的字符串。所述多个字符串中的每个字符串也可以包括一种概率,该概率是对该字符串在该上下文 210 期间被多么频繁地提供的度量。

[0044] 在一些实施方案中,确定上下文可以包括基于与所接收的字符串相关联的应用 135 或者与所接收的字符串相关联的日期或者时间来确定上下文。可替换地或者附加地,例如,上下文可以基于与所接收的字符串相关联的对话 220,并且上下文可以基于对话的参与者、对话的话题、或者对话的风格来确定。对话的话题或者风格可以由上下文引擎 240 通过寻找对话 220 中指示特定风格或者话题的字符串来确定。

[0045] 在 305 处,基于所接收的字符串和所确定的上下文,生成多个字符串预测。多个字符串预测可以包括字符串预测 215 并且可以由候选预测发生器 140 使用所接收的字符串和与所确定的上下文 210 相关联的多个字符串来生成。在一些实施方案中,字符串预测 215 中的每个可以包括来自与所确定的上下文 210 相关联的多个字符串中的字符串,并且可以通过从所述多个字符串中确定包括所接收的字符串作为前缀的字符串来生成。可替换地,字符串预测 215 可以通过从所述多个字符串中确定与所接收的字符串相似的字符串来生成。字符串预测 215 中的每个还可以包括基于多个字符串中的字符串的概率的置信度值。

[0046] 在一些实施方案中,多个字符串预测 215 也可以通过候选预测发生器 140 使用预测模型 205 和所确定的上下文 210 来生成。预测模型 205 可以包括或者基于例如词典。

[0047] 在 307 处,提供所述多个字符串预测中的一个或者多个字符串。可以由候选预测发生器 140 按照字符串预测 215 的置信度值来提供所述一个或者多个字符串。在一些实施方案中,可以提供字符串预测 215 中的具有最高置信度值的字符串。可以通过例如在用户接口 120 的软输入面板中显示字符串来提供字符串。当字符串作为针对所接收的字符串的拼写建议而提供时,字符串可以通过操作系统 125 临近所接收的字符串来显示。

[0048] 在 309 处,接收所提供的一个或者多个字符串中的一个的选择的指示。选择的指示可以由操作系统 125 从用户接口 125 接收,并可被提供给候选预测发生器 140。例如,用户可以在用户接口 120 的软输入面板中选择或者触摸字符串中的一个。

[0049] 在 311 处,所选择的字符串作为输入提供给应用。所选择的字符串可以由候选预测发生器 140 作为对话 220 的一部分提供给应用 135,或者作为针对所接收的字符串的拼写替换提供给应用 135。

[0050] 在一些实施方案中,所选择的字符串也可以用来通过候选预测发生器 140 的上下文引擎 240 来更新所确定的上下文 210。例如,与关联于所确定的上下文 210 的多个字符串中的所选择字符串相关联的概率可以通过候选预测发生器 140 的上下文引擎 240 来增加。

[0051] 图 4 是用于更新上下文的方法 400 的实施方案的操作流程。方法 400 可以通过例如候选预测发生器 140 来实施。

[0052] 在 401 处,接收字符串。字符串可以由候选预测发生器 140 从操作系统 125 接收。字符串可能已经由应用 135 的用户输入或者选择。字符串可以包括一个或者多个字或者短语。

[0053] 在 403 处,为所接收的字符串确定上下文。上下文 210 可以通过候选预测发生器 140 的上下文引擎 240 来确定。上下文 210 可以与多个字符串相关联,该多个字符串是由用户或者其他用户在特定上下文 210 期间频繁提供的字符串。该多个字符串中的每个字符串也可以包括概率,该概率是对在该上下文 210 期间该字符串被多么频繁地提供的度量。

[0054] 在 405 处,确定与所确定的上下文相关联的多个字符串是否包括所接收的字符串。可以由候选预测发生器 140 的上下文引擎 240 做出确定。如果所述多个字符串包括该字符串,则方法 400 可以在 407 处继续。否则,方法 400 可在 409 处继续。

[0055] 在 407 处,所述多个字符串中的该字符串的概率增加。该概率可以由候选预测发生器 140 的上下文引擎 240 来增加。例如,上下文引擎 240 可以通过将固定量添加到所述多个字符串中的所接收的字符串的概率来增加该概率。

[0056] 在 409 处,确定该字符串是否满足标准。可以通过候选预测发生器 140 的上下文引擎 240 做出确定。在一些实施方案中,标准可以是字符串被接收了多于阈值次数。例如,在字符串已经由用户向应用 135 输入或者提供了多于三次(或者另一个预先确定的次数,这取决于实施方案)之后,该字符串可以被添加到与所确定的上下文 210 相关联的多个字符串中。其它的标准可以被使用。如果标准得以满足,那么方法 400 可以在 411 处继续。否则,方法 400 可以在 413 处退出。

[0057] 在 411 处,字符串被添加到与上下文相关联的多个字符串中。字符串可以通过候选预测发生器 140 的上下文引擎 240 来添加。在一些实施方案中,与所添加的字符串相关联的概率可以被设定成例如针对新添加的字符串的默认值。

[0058] 图 5 示出其中可以实施示例实施例和方面的示例性计算环境。计算系统环境仅是适当的计算环境的一个示例,并且不打算建议关于用途或者功能性的范围的任何限制。

[0059] 可以使用若干其它通用或者专用计算系统环境或者配置。可以适用于使用的众所周知的计算系统、环境和 / 或配置的示例包括但不限于个人计算机、服务器计算机、手持式或者膝上型设备、多处理器系统、基于微处理器的系统、网络 PC、小型计算机、大型计算机、嵌入式系统、包括上文的系统或者设备中的任何一个的分布式计算环境等等。

[0060] 可以使用由计算机执行的计算机可执行指令,诸如程序模块。一般地,程序模块包括执行特定任务或者实施特定抽象数据类型的例程、程序、对象、组件、数据结构等等。可以使用分布式计算环境,其中任务由通过通信网络或者其它数据传输介质链接的远程处理设备来执行。在分布式计算环境中,程序模块和其它数据可以位于本地和远程计算机存储介质(包括存储器存储设备)两者中。

[0061] 参考图 5,用于实施本文描述的方面的示例性系统包括计算设备,诸如计算系统

500。在其最基本的配置中,计算系统 500 典型地包括至少一个处理单元 502 和存储器 504。取决于计算设备的确切配置和类型,存储器 504 可以是易失性(诸如随机存取存储器(RAM))、非易失性(诸如只读存储器(ROM)、闪速存储器等等)或者这两者的某组合。这个最基本的配置在图 5 中由虚线 506 图示出。

[0062] 计算系统 500 可以具有附加的特征 / 功能性。例如,计算系统 500 可以包括附加存储装置(可移除和 / 或不可移除的),包括但不限于,磁性或者光学盘或者带。这样的附加存储装置在图 5 中通过可移除存储装置 508 和不可移除存储装置 510 图示出。

[0063] 计算系统 500 典型地包括多种计算机可读介质。计算机可读介质可以是可以由计算系统 500 访问的任何可用介质,并且包括易失性和非易失性介质、可移除和不可移除介质两者。

[0064] 计算机存储介质包括以用于信息(诸如计算机可读指令、数据结构、程序模块或者其它数据)的存储的任何方法或者技术实施的易失性和非易失性、以及可移除和不可移除介质。存储器 504、可移除存储装置 508、和不可移除存储装置 510 都是计算机存储介质的示例。计算机存储介质包括但不限于 RAM、ROM、电可擦除程序只读存储器(EEPROM)、闪速存储器或者其它存储器技术、CD-ROM、数字万能盘(DVD) 或者其它光学存储装置、磁盒、磁带、磁盘存储装置或者其它磁存储设备、或者可以用来存储合期望的信息并且可以由计算系统 500 访问的任何其它介质。任何这样的计算机存储介质可以是计算系统 500 的一部分。

[0065] 计算系统 500 可以含有允许设备与其它设备和 / 或接口进行通信的(多个)通信连接 512。计算系统 500 也可以具有(多个)输入设备 514,诸如键盘(软件或者硬件)、鼠标、笔、语音输入接口、触摸接口等等。也可以包括(多个)输出设备 516,诸如显示器、扬声器、打印机等等。所有这些设备都在本领域是众所周知的,并且不需要在此处详尽地讨论。

[0066] 应该理解的是,本文描述的各种技术可以连同硬件或者软件或者在适当情况下两者的组合来实施。因此,本公开主题的方法和装置或者其某些方面或者部分可以采用在有形介质中具体实现的程序代码(即,指令)的形式,所述有形介质诸如是软盘、CD-ROM、硬驱动器、或者任何其它机器可读存储介质,其中当程序代码被载入机器(诸如计算机)并且由该机器执行时,该机器变成用于实践本公开主题的装置。

[0067] 虽然示例性实施方案可能涉及在一个或者多个单独计算机系统的上下文中利用本公开主题的方面,但是主题不如此受限,相反可以连同任何计算环境(诸如网络或者分布式计算环境)实施。更进一步地,本公开主题的方面可以在多个处理芯片或者设备中或者跨多个处理芯片或者设备实施,并且相似地,存储装置可以跨多个设备实现。这样的设备可以包括例如个人计算机、网络服务器、和手持式设备。

[0068] 虽然主题已经以具体于结构特征和 / 或方法论动作的语言进行了描述,但是将理解的是,在所附权利要求中定义的主题不必受限于上文描述的具体特征或者动作。相反地,上文描述的具体特征和动作作为实施权利要求的示例形式而公开。

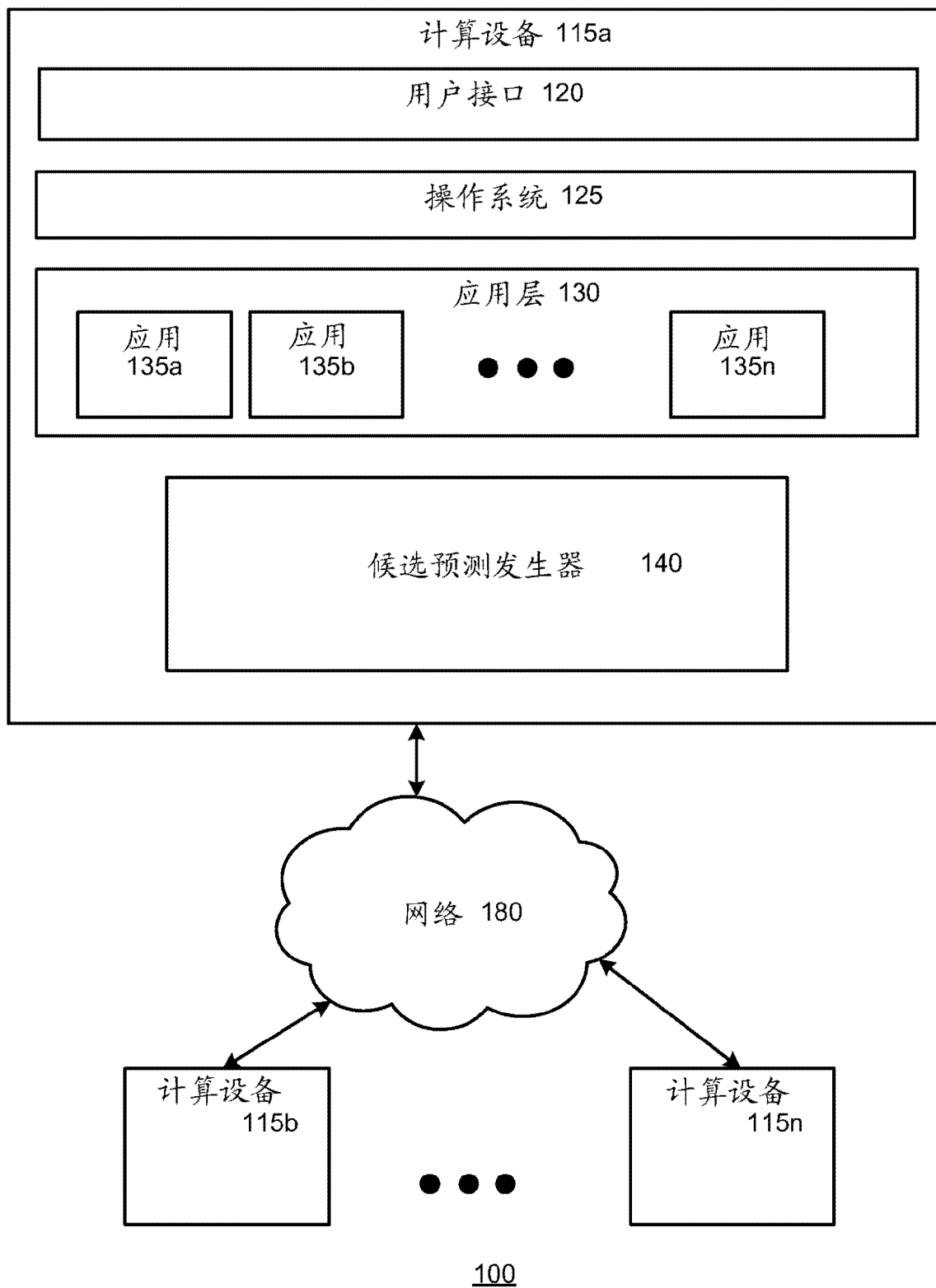


图 1

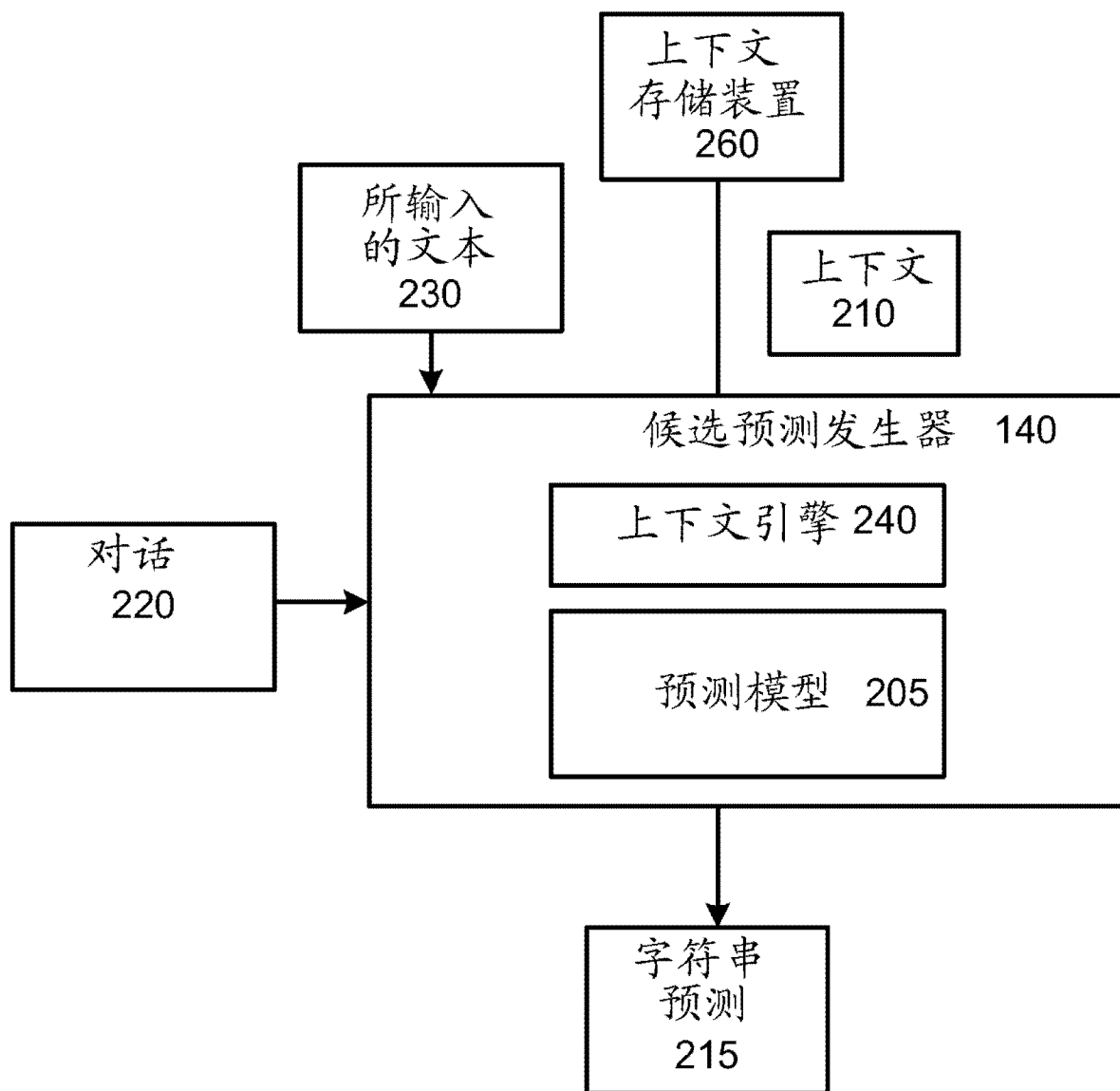
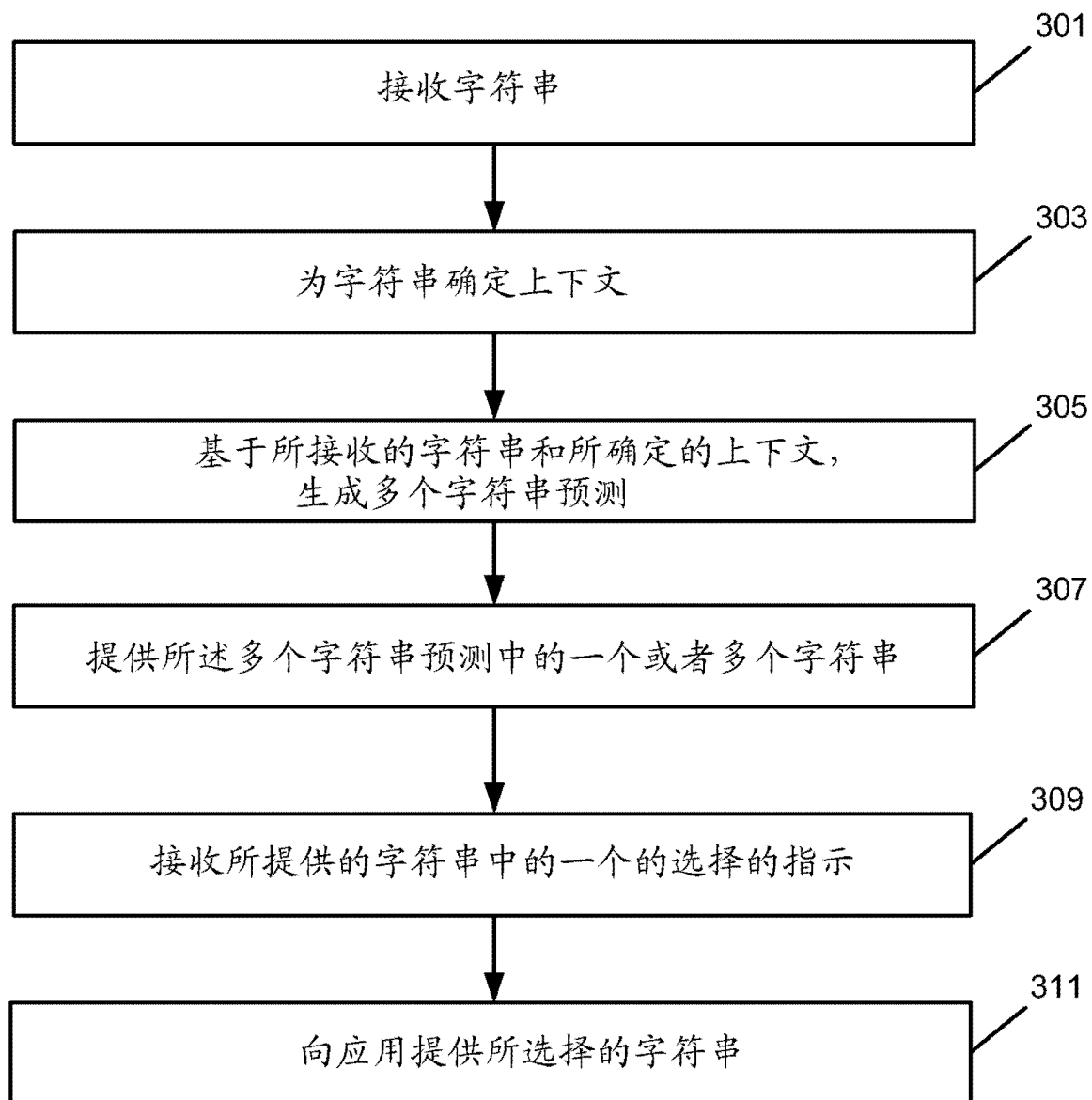
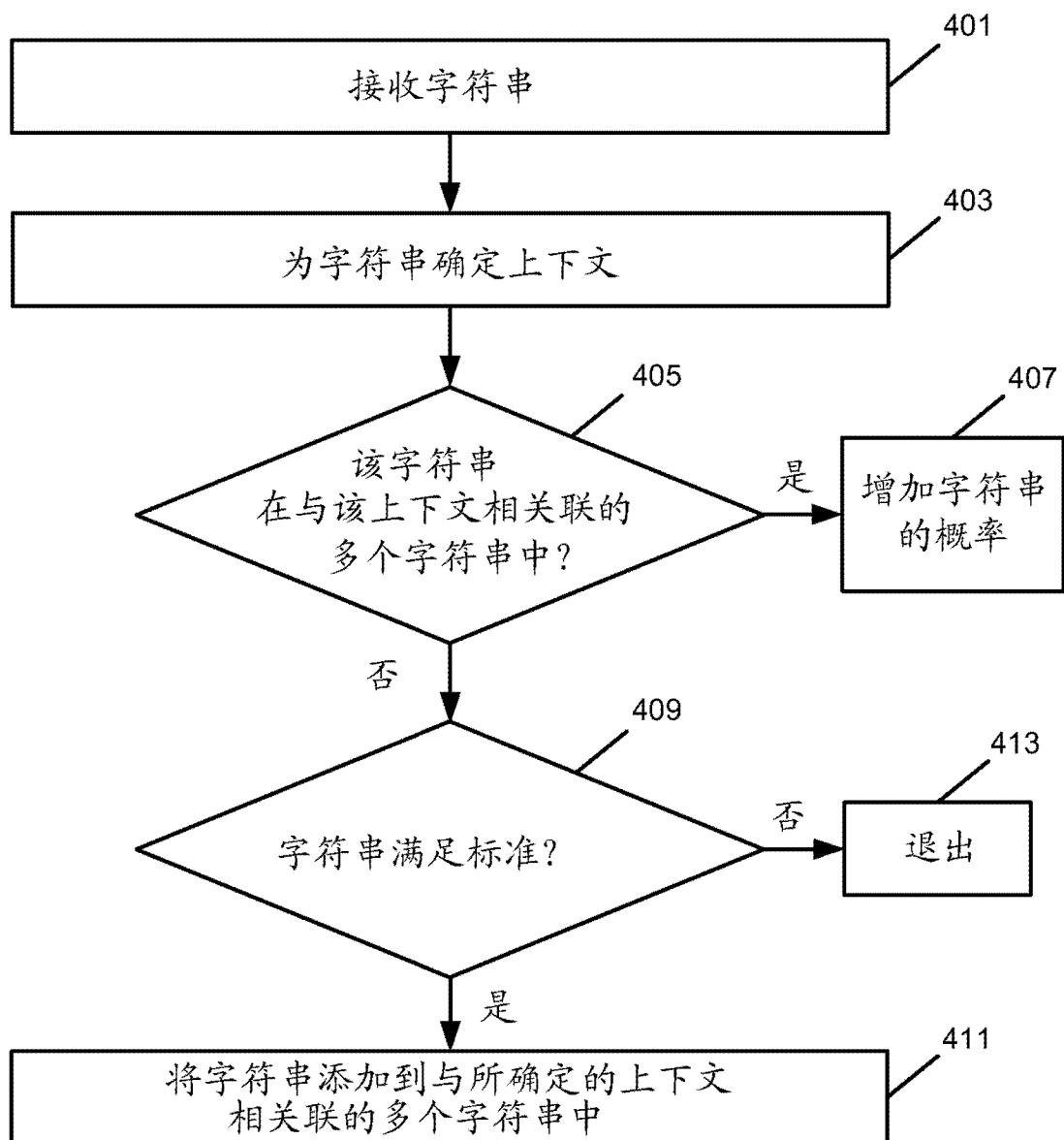


图 2



300

图 3



400

图 4

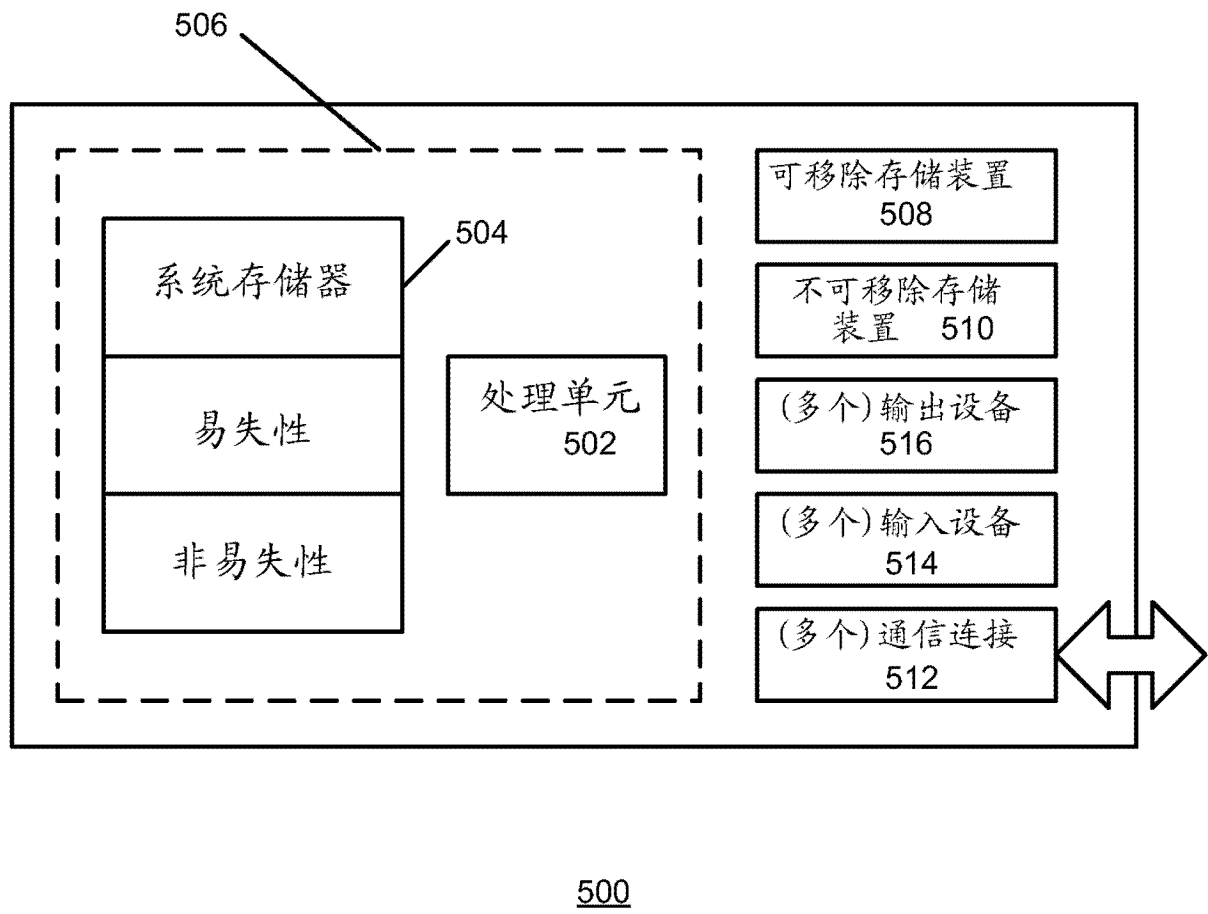


图 5