(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 102388383 A
(43)申请公布日 2012.03.21

(21)申请号 200780050481.5

(22)申请日 2007.12.10

(30)优先权数据
60/873,812 2006.12.08 US

(85)PCT申请进入国家阶段日
2009.07.27

(86)PCT申请的申请数据
PCT/US2007/087036 2007.12.10

(87)PCT申请的公布数据
WO2008/070877 EN 2008.06.12

(71)申请人 帕特里克·J·霍尔
地址 美国马里兰州

(72)发明人 帕特里克·J·霍尔

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038
代理人 曲瑞

(51)Int.Cl.
G06F 17/28(2006.01)

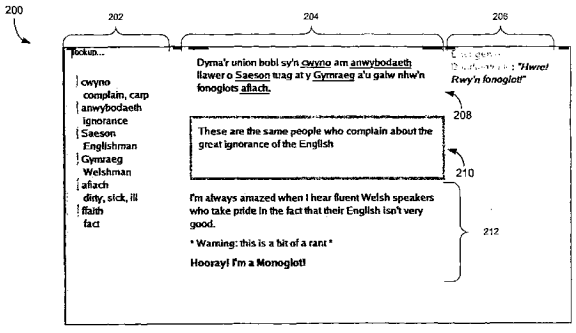
权利要求书 2 页 说明书 7 页 附图 7 页

(54)发明名称

在线计算机辅助翻译

(57)摘要

本发明涉及一种在线计算机辅助翻译方法、程序产品和系统。接收源语言的源文本。将源文本分割成多个片段。从用户接收目标语言的、与所述片段中的第一片段相关联的第一译文输入。将第一译文输入存储在文本数据仓库中。



CN 102388383 A

1. 一种方法，包含：

接收源语言的源文本；

将所述源文本分割成多个片段；

从用户接收目标语言的第一译文输入，所述第一译文输入与所述片段中的第一片段相关联；以及

将所述第一译文输入存储在文本数据仓库中。

2. 根据权利要求 1 所述的方法，其中所述文本数据仓库被存储在一个或更多个服务器中。

3. 根据权利要求 1 所述的方法，进一步包括：

从所述用户接收目标语言的第二译文输入，所述第二译文输入与所述片段中的第二片段相关联；以及

将所述第二译文输入存储在所述文本数据仓库中。

4. 根据权利要求 1 所述的方法，进一步包括：

接收第一语言的查询；

在所述文本数据仓库中检索一个或更多个所述第一语言的、与所述查询匹配的文本串，其中所述文本数据仓库包括与所述第一语言的匹配文本串中的每一个匹配文本串相关联的相应的第二语言的译文；以及

呈现所述第二语言的所述译文。

5. 根据权利要求 1 所述的方法，进一步包括：

将所述源文本存储在所述文本数据仓库中。

6. 根据权利要求 1 所述的方法，进一步包括：

对各片段生成评价线程。

7. 根据权利要求 1 所述的方法，其中所述源文本的片段中的一个或更多个片段与来自用户的所述第二语言的相应的译文输入相关联，所述源文本的所述片段和相应的译文输入被存储在所述文本数据仓库中，所述方法进一步包括：

基于所述译文输入的量和所述源文本的量，确定对所述源文本的翻译完成比率；以及

呈现所述翻译完成比率。

8. 一种计算机程序产品，被编码在有形的程序载体上，可操作以使数据处理装置执行包括以下步骤的操作：

接收源语言的源文本；

将所述源文本分割成多个片段；

从用户接收目标语言的第一译文输入，所述第一译文输入与所述片段中的第一片段相关联；以及

将所述第一译文输入存储在文本数据仓库中。

9. 一种系统，包含：

一个或更多个服务器，可操作以存储文本数据仓库；以及

计算机，可操作以

接收源语言的源文本；

将所述源文本分割成多个片段；

从用户接收目标语言的第一译文输入，所述第一译文输入与所述片段中的第一片段相关联；以及

将所述第一译文输入存储在所述文本数据仓库中。

10. 一种系统，包含：

翻译匹配器，用于将翻译者与对内容的翻译请求进行匹配；

翻译编辑器，用于使得便于进行内容的翻译；和

翻译数据库，用于存储内容的翻译结果。

在线计算机辅助翻译

[0001] 本申请要求 2006 年 12 月 8 日提交的、名称为“在线计算机辅助翻译”的美国临时专利申请第 60/873812 号的基于 35U.S.C. §119 的优先权，通过引用将该美国临时专利申请的全部内容并入于此。

技术领域

[0002] 本公开总体上涉及计算机辅助翻译。

背景技术

[0003] 随着万维网已增长并已变成国际媒体，英语作为用于 Web 上内容的语言选择的主导性也已变小。Web 上的许多内容是以英语之外的语言写的。此现象的例子出现在博客圈 (blogosphere) 中，其中存在许多以英语之外的语言写的博客。非英语博客和非英语 Web 内容的这种增长总体上增大了对在语言之间的差异上架设桥梁的语言翻译的需求。

[0004] 一种翻译选项是机器翻译，其中完全由计算机对内容进行翻译。然而，机器翻译具有其局限性，诸如关于准确性和可由机器翻译处理的语言数量有限的问题。另一种翻译内容的选项是计算机辅助翻译 (CAT)，其中人利用软件程序的协助来翻译内容。然而，当前可获得的 CAT 软件往往很昂贵并且被销售给专业人员。这会抬高 CAT 的成本并且使得这样的服务对许多人或群体是不可获得的。

发明内容

[0005] 一般而言，可在下面的方法中实施本说明书中描述的主题的一个方面，该方法包括如下动作：接收源语言的源文本；将源文本分割成多个片段；从用户接收目标语言的第一译文输入，第一译文输入与所述片段中的第一片段相关联；以及将第一译文输入存储在文本数据仓库中。此方面的其他实施例包括对应的系统、装置、计算机程序产品和计算机可读介质。

[0006] 一般而言，可在下面的系统中实施本说明书中描述的主题的另一个方面，该系统包括：翻译匹配器，用于将翻译者与对内容的翻译请求匹配；翻译编辑器，用于使得便于进行内容的翻译；和译文数据库，用于存储内容的译文。此方面的其他实施例包括对应的系统、装置、方法、计算机程序产品和计算机可读介质。

[0007] 可实施本说明书中描述的主题的特定的实施例，以实现一个或多个下述优点：翻译项目的源文本和工作产品可以是能够从具有 web 浏览器的计算机访问的，而无需安装专用的软件或附件 (add-on)。委托了翻译项目的客户可以自行检查该项目的进展。翻译者可以与其他翻译者合作以及从其他翻译者寻求协助。

[0008] 在以下附图和描述中阐明本说明书中描述的主题的一个或多个实施例的细节。根据说明书、附图和权利要求书，该主题的其他特征、方面和优点将变得明显。

附图说明

[0009] 图 1 示出了示例的集合器 (aggregator) 界面。

[0010] 图 2 示出了示例的编辑器界面。

[0011] 图 3 示出了在线计算机辅助翻译服务的示例实体关系图。

[0012] 图 4 示出了用于请求翻译的示例处理。

[0013] 图 5 示出了用于接受翻译请求的示例处理。

[0014] 图 6 例示了用于接收和存储译文输入的示例处理。

[0015] 图 7 例示了用于接收和存储源文本和译文的示例系统。

[0016] 各个图中的类似的标号和标记表示类似的组成部分。

具体实施方式

[0017] 计算机辅助翻译 (CAT) 工具可以在线实现。在一些实现方式中，CAT 工具是驻留在网站上的基于 Web 的服务。通过 CAT 网站，翻译者可以选择内容来翻译、输入对于所选择的内容的翻译、以及将翻译的内容公布。

[0018] 在一些实现方式中，在线 CAT 工具包括 :集合器，用于管理源内容和选择内容来翻译 ;编辑器，用于帮助翻译者快速和高效地工作 ;以及输出框，用于将完成的译文组织成输出内容。

[0019] 希望对内容进行翻译的翻译者可以利用 CAT 工具注册账户。CAT 工具可以包括用于账户管理和设置个人偏好的页面。

[0020] 在一些实现方式中，CAT 工具被实现为使用超文本标记语言 (HTML)、JavaScript、可扩展标记语言 (XML)、异步 JavaScript 和 XML(AJAX)、以及其他合适的技术的网页。可在 web 浏览器中呈现网页。

[0021] 图 1 示出了示例的集合器界面。集合器使得翻译者便于选择要翻译的内容 (“源内容”)。集合器呈现出可用于翻译的源内容 (例如，博客、网页)。在一些实现方式中，集合器的布局包括 :区域 102，用于显示诸如博客、网页等的标题之类的可用于翻译的源内容的列表 ;和另一个区域 104，用于显示当前选择的标题下的内容。在一些实现方式中，显示的内容可包括被标记为“请翻译此内容”的链接或者一些其他的用户界面对象。翻译者可以选择该链接或对象以接受要翻译的内容。在一些实现方式中，用户界面被置于编辑器中，翻译者在此进行翻译，以下描述其进一步的细节。在一些其他的实现方式中，内容被添加到翻译者的翻译明细表 (docket) 中。

[0022] 集合器可包括用于将内容添加到集合器的工具。在一些实现方式中，具有用于通过指定博客的统一资源定位符 (URL) 或其内容馈源 (例如，RSS 馈源、Atom 馈源) 而将博客添加到集合器的用户界面。当翻译者提交博客 URL 时，取得博客的内容 (例如，通过访问它的内容馈源)，并且将内容添加到数据库。然后翻译者可浏览添加的内容并选择任何要翻译的内容。在一些实现方式中，可以使用类似的用户界面来添加诸如网页等要翻译的其他内容。在一些其他的实现方式中，也可在没有翻译者的介入的情况下添加内容。例如，集合器可以显示来自其他人的翻译请求，并且翻译者可以浏览这些请求并且选择他们希望接受的请求。作为另一个例子，CAT 工具可以基于诸如涉及的语言和翻译者的技能集合 (skill set) 之类的任何数量的准则，将内容自动指派给翻译者。

[0023] 在一些实现方式中，集合器可以向翻译者呈现可翻译的内容，其中可按照源 (例

如，博客、网站域、翻译的请求方等）对内容进行组织，并且以特定单元（诸如博客帖子、个人网页等）呈现内容。翻译者可以挑选特定单元的内容以将其添加到他们的翻译明细表。例如，翻译者可以将博客添加到集合器。集合器向翻译者呈现来自博客的帖子，并且翻译者可以选择特定的博客帖子以将其添加到他们的一览表。

[0024] 在一些实现方式中，源内容被存储在一个服务器或多个服务器处。例如，源内容可被从博客、网站等提取并被存储在服务器处。作为另一个例子，可将源内容的文件上传到服务器。作为进一步的例子，可（例如，通过键入、拷贝和粘贴等）将源内容文本输入到表单中并将该文本发送到服务器中进行存储。源内容被存储在服务器处的文本数据的仓库（例如，数据库）中。集合器界面可将存储在文本数据仓库中的源内容显示给潜在的翻译者以供选择。在一些实现方式中，源内容文本在服务器被分割成片段。片段可以是句子、段落、表格单元等。

[0025] 图2示出了示例的编辑器界面。翻译者可在编辑器200中进行源内容的翻译。在一种实现方式中，编辑器包括术语表202、编辑区域204、以及用于显示杂项信息的区域206。在一些实现方式中，编辑器将要被翻译的源内容自动划分或分割成较小的单元（“片段”）。在一些实现方式中，这些片段是通过诸如句号、问号等句子结束标点划界而成的单个句子。在一些其他的实现方式中，这些片段是源内容文本中的段落、表格单元等。

[0026] 在图2中，编辑区域204示出用威尔士语(Welsh)写的博客帖子的片段以及已输入的一些片段的译文。在当前句子区域208中示出正被翻译的当前片段（以“Dyma’r union bobl...”开头的文本）。翻译者可在当前片段翻译区域210中键入对当前片段区域208中所示的片段的译文。可在完成翻译区域212中显示已完成的对该片段的译文。在一些实现方式中，按照完成顺序的相反顺序显示完成的译文；当翻译者完成各个单个片段并且（例如，通过按回车键、按“提交”按钮等）提交译文时，该片段被推到完成翻译区域212中的句子列表的顶端。例如，当完成对当前句子的翻译“These are the same people...”时，它将占据其下以“I’m always amazed...”开始的句子之上的位置。这里除了内容的单元是句子而不是单个的帖子之外，类似于博客，文本从顶部向底部流动。在一些实现方式中，当某片段被高亮显示或者被选择进行编辑时，显示该片段的译文。在一些实现方式中，可在并排的视图中显示原始的源内容文本和译文。

[0027] 正被翻译的当前句子中带下划线的词是已在CAT工具的术语表中被找到的那些词，并且可将它们显示在术语表区域202中。

[0028] 在一些实现方式中，与当完成对整个源内容的翻译时进行保存相反，当翻译者提交了片段的译文时，该片段的译文就被保存到服务器。例如，可将译文存储在存储源内容的文本数据仓库中。这样，当翻译者进行对源内容文本的翻译时，可以逐个片段地保存其译文。在文本数据仓库内，可将译文与对应的源内容片段关联起来。

[0029] 具有源内容文本和源内容文本的片段的译文的文本数据仓库是可检索的。例如，编辑器界面200可包括用于从文本数据仓库检索源内容文本的片段的检索框。用户（例如，翻译者）可将文本查询输入到检索框中，并且从文本数据仓库检索包括该文本查询的片段。匹配的片段和它们的译文作为检索结果返回给用户。这样，翻译者可在文本数据仓库中检索文本以看其他翻译者是如何翻译该文本的。

[0030] 在一些实现方式中，可基于源内容的已在文本数据仓库中保存有译文的片段的数

目（或者片段中词/字符的数目）和源内容文本中片段的总数目（或词/字符的总数目），计算对该源内容文本的翻译完成百分比或比率。可将完成百分比和源内容文本一起显示在编辑器界面 200 中。也可向委托翻译的客户显示完成百分比（例如，当客户访问源内容文本和译文以评估翻译的进展时）。

[0031] 在一些实现方式中，文本数据仓库中的源内容和译文无限制地对翻译者和客户开放浏览。然而，可能存在源内容文本和源内容文本的译文包括委托翻译的客户不希望向未被授权方公开的机密信息或其他信息的情况。在一些实现方式中，可将对源内容文本和译文的检索、浏览和编辑限于客户和被授权方（例如，被委托执行翻译的翻译者）。该限制可以针对整篇源内容文本，或者针对各个片段（例如，某些片段对公众开放，而其他片段限于被授权方）。

[0032] 在一些实现方式中，可以生成评价线程 (thread)，并使评价线程与源内容文本的片段相关联。用户（例如，正翻译该片段的翻译者）可以通过使用评价线程请求其他翻译者的协助。这样，评价线程可促进在翻译中进行合作。此外，在一些实现方式中，可以使用对翻译者作出的质量评价（例如，翻译者提供了被其他用户投票为有帮助的协助这样的评价）的数目来确定翻译者的质量或声誉度量。

[0033] 如上所述，可将文本数据仓库存储在一个或多个服务器处。用户（例如，翻译者、客户）可通过基于 Web 的界面（例如，集合器界面 100 和编辑器界面 200）访问文本数据仓库的内容（即，源内容和译文）。

[0034] 图 3 示出了在线计算机辅助翻译服务的示例实体关系图 300。在一些实现方式中，图 300 也表示用于在线计算机辅助翻译的系统。存在具有需翻译材料的消费者，他们希望使材料或内容被翻译。这些消费者可参与翻译市场。在翻译市场中，消费者可提出翻译请求，并且指定或提供要被翻译的材料。消费者还按预先确定的费率支付翻译酬金或者说明他们愿意为翻译支付多少钱（即，招标）。翻译者团体可查看翻译请求并且决定接受哪些请求。也可基于诸如技能集合、涉及的语言和对请求的出价之类的多种准则，将翻译者与请求进行匹配。翻译者也可说明他们的价格费率。

[0035] 当翻译者准备翻译一项内容时，翻译者可以使用翻译编辑器来进行翻译。翻译过的内容被返回给相应的消费者。可将由消费者支付的翻译酬金的至少一部分金额支付给翻译者。翻译市场也可获得由消费者支付的酬金的一部分作为佣金或服务费等。

[0036] 翻译好的内容也被保存在译文数据库中。该数据库可存储原始材料及其多种语言的译文。在一些实现方式中，译文数据库是上述文本数据仓库。翻译者可访问该数据库中的译文以协助他们执行他们的翻译。换句话说，翻译好的内容被保存并可被翻译者在将来用作例子或参考。

[0037] 在一些实现方式中，翻译者也可为其他翻译者的译文评定等级。可将这样的等级评定保存在数据库中。根据这些等级评定，翻译者可以在翻译者团体内和在翻译市场中建立声誉。声誉可以反映在等级评定中，并可被提供给请求翻译的消费者。

[0038] 译文数据库可被视作内容和内容译文的集汇 (corpus)。在一些实现方式中，可以对希望访问该集汇的实体或系统提供应用程序编程接口 (API)。例如，机器翻译系统可以访问该集汇来训练它的翻译算法。可免费提供 API，或者该 API 以付费订阅或许可的方式被提供。

[0039] 图4示出了用于请求翻译的示例处理400。消费者指定要被翻译的内容(402)。在一些实现方式中，消费者指定内容的源语言、内容的目标语言和源格式（例如，内容是否为由URL指示的网页、由馈源URL指示的内容馈源、电子邮件等）。消费者也可以指定其他信息，诸如到期日和对酬金的出价。

[0040] 消费者选择翻译者(404)。在一些实现方式中，消费者可通过姓名请求特定的翻译者。消费者也可通过诸如语言、翻译者等级和专业技能（例如，法律文本的技能、医学文本的技能、关于航空的文本的技能等）多等种准则来搜寻翻译者。

[0041] 消费者和翻译者商谈价格(406)。消费者和翻译者都可以出价，直到达成彼此同意的价格。在一些实现方式中，可以按照每个词的费用或者每个字符的费用来表示价格。

[0042] 在一些实现方式中，省略价格商谈。翻译者可提前指定他们的费率，并且消费者可基于价格及其他因素选择翻译者。消费者可以拒绝其费率不合适的翻译者。

[0043] 在选定的翻译者翻译内容后，消费者接收翻译好的内容(408)。

[0044] 图5示出了用于接受翻译请求的示例处理500。翻译者可选择接受翻译请求(502)。可使用各种准则来对请求排序，以使翻译者更高效地找到他们更合意的请求。排序准则的非穷举的例子包括语言、所需的专业技能、价格、要被翻译的内容的量、和到期日。在一些实现方式中，翻译者可以在决定是否接受请求之前查看请求的原始内容。

[0045] 在翻译者接受请求之后，翻译者和提出请求的消费者可以商谈价格(504)。在价格达成一致之后，翻译者开始翻译内容(506)。翻译者可以使用上述编辑器200和相关的工具来进行翻译。在完成翻译之后，将译文传送给消费者。

[0046] 在一些其他的实现方式中，翻译者可以提前指定他们的技能集合和费率，并且消费者可以提出指定所涉及的语言、任何所需技能和价格的请求。可基于所指定的信息自动地将翻译者与请求（或者指派给翻译者的请求）匹配。

[0047] 在进一步的其他实现方式中，翻译者可免费地执行翻译服务。

[0048] 图6示出用于接收和存储翻译输入的示例处理600。接收源语言的源文本(602)。例如，希望委托翻译项目的客户可通过文件或者输入到表格中将源文本提交到系统（例如，系统700，图7）。在一些实现方式中，文件或者表格输入通过前端设备(front-end)704（图7）被接收并被存储在文本数据仓库706（图7）中。

[0049] 在一些实现方式中，由客户提供统一资源定位符(URL)，并且系统700（例如，前端704）可以根据所提供的URL取得源文本。

[0050] 源文本被分割成多个片段(604)。在一些实现方式中，片段是源文本的单个句子，源文本中的每一个句子是一个片段。在一些其他的实现方式中，片段是源文本的段落。可以有其他分割单位。如果源文本足够短以致于在一个片段单元内，那么源文本本身也可以是片段。例如，一个句子就具有一个片段，即该句子本身的源文本（假定句子是分割单位）。在一些实现方式中，源文本以片段的形式被存储在文本数据仓库706中。当在编辑器界面200中向翻译者显示源文本时，源文本被显示为片段。

[0051] 从用户接收对于一个片段的译文输入(606)。用户（例如，翻译者）可以在编辑器界面200中输入对源文本的多个片段的到目标语言的译文。翻译者逐个片段地输入译文。翻译者通过编辑器界面200提交对片段的译文输入，并且系统700接收该输入。

[0052] 译文输入被存储在文本数据仓库中(608)。接收的译文输入被存储在文本数据仓

库中，而不必等待对整个源文本的翻译的完成。

[0053] 翻译者可输入对于其他片段的译文。系统 700 接收译文输入，并且将它们逐片段地存储到文本数据仓库中。

[0054] 图 7 示出用于接收和存储源文本和译文的示例系统。系统 700 包括前端设备 704 和文本数据仓库 706。在一些实现方式中，前端设备 704 是 web 服务器。例如，在一些实现方式中，前端设备 704 充当用以帮助计算机辅助翻译的基于 Web 的界面，所述界面包括集合界面 100 和 / 或编辑界面 200。用户（例如，翻译者、客户）可从能够呈现和显示网页的用户设备 702（例如，具有 web 浏览器应用程序的用户设备）访问基于 Web 的界面。用户设备 702 的例子包括台式计算机、笔记本电脑、智能电话、移动电话、个人数字助理 (PDA) 和平板 (tablet) 计算机等。

[0055] 文本数据仓库 706 存储源文本和对应的译文。在一些实现方式中，源文本在文本数据仓库中被存储为片段，并且存储的译文是对片段的相应的翻译。在一些实现方式中，文本数据仓库 706 充当上面参照图 3 所描述的译文数据库。在一些实现方式中，文本数据仓库 706 可被存储在计算机（例如，服务器）中，或者被存储或分布在多个计算机（例如，服务器）上。

[0056] 可以以数字电子电路，或者以包括本说明书中公开的结构及其等同结构在内的计算机软件、固件或硬件，或者以它们中的一个或更多个的组合，来实现所公开的和其他的实施例以及本说明书中描述的功能操作。可将公开的和其他的实施例实现为一个或更多个计算机程序产品，即，被编码在计算机可读介质上用于由数据处理装置执行或者用于控制数据处理装置的操作的计算机程序指令的一个或更多个模块。计算机可读介质可以是机器可读的存储设备、机器可读的存储基板 (substrate)、存储器设备、生成机器可读的传播信号的合成体、或者它们中的一个或更多个的组合。术语“数据处理装置”包含所有用于处理数据的装置、设备和机器，例如包括可编程处理器、计算机、或者多个处理器或计算机。除了硬件之外，该装置可包括为相关计算机程序创建执行环境的代码，例如，构成处理器固件、协议栈、数据库管理系统、操作系统或它们中一个或更多个的组合的代码。传播信号是人为生成的信号，例如，为了将信息编码以发送到合适的接收器装置而生成的机器生成的电、光或电磁信号。

[0057] 可以采用包括编译语言或解释语言在内的任何形式的编程语言来编写计算机程序（也被称为程序、软件、软件应用程序、脚本或代码），并且可以任何形式（包括作为单独的程序，或者作为适于在计算环境中使用的模块、组件、子例程或其他单元）来配置 (deploy) 该计算机程序。计算机程序不一定对应于文件系统中的文件。可将程序存储在保存其他程序或数据（例如，存储在标记语言文档中的一个或多个脚本）的文件的一部分中、存储在专用于被讨论的程序的单个文件中、或者存储在多个并列文件（例如，存储一个或更多个模块、子程序、或代码部分的多个文件）中。可将计算机程序配置为在位于一个站点处的一个计算机或多个计算机上执行或在分布于多个站点并且通过通信网络互连的多个计算机上执行。

[0058] 可通过一个或更多个可编程处理器执行本说明书中描述的处理和逻辑流程，所述一个或更多个可编程处理器执行一个或更多个计算机程序以通过对输入数据进行操作并生成输出来执行功能。也可通过专用逻辑电路（例如，现场可编程门阵列 (FPGA) 或专用集

成电路 (ASIC)) 执行处理和逻辑流程，也可将装置实现为上述专用逻辑电路。

[0059] 适于执行计算机程序的处理器例如包括 :通用微处理器、专用微处理器、和任何种类的数字计算机的任何一个或更多个处理器。通常，处理器将从只读存储器、随机存取存储器、或它们两者接收指令和数据。计算机的基本元件是用于执行指令的处理器和用于存储指令和数据的一个或更多个存储器设备。通常，计算机还将包括用于存储数据的一个或更多个海量存储设备 (例如，磁盘、磁光盘或者光盘)，或者可操作地耦合以从所述海量存储设备接收数据、或向所述海量存储设备传送数据、或者兼而有之。然而，计算机并不必须具有这样的设备。适于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性存储器、介质和存储器设备，例如包括半导体存储器设备 (例如，EPROM、EEPROM 和闪存设备)、磁盘 (例如，内部硬盘或可移动盘)、磁光盘、和 CD-ROM 及 DVD-ROM 盘。可以用专用逻辑电路来补充处理器和存储器，或者将处理器和存储器并入所述专用逻辑电路。

[0060] 为了提供与用户的交互，可在具有用于向用户显示信息的显示设备 (例如，阴极射线管 (CRT) 或者液晶显示 (LCD) 监视器)、键盘和点击设备 (例如，鼠标或跟踪球) 的计算机上实现所公开的实施例，用户可通过所述显示设备、键盘和点击设备向计算机提供输入。也可使用其他种类的设备以提供与用户的交互，例如，向用户提供的反馈可以是任何形式的感觉反馈 (例如，视觉反馈、听觉反馈或触觉反馈)，并且可以任何形式接收来自用户的输入 (包括声音的、语音或触觉输入)。

[0061] 可在包括后端组件例如数据服务器的计算系统中、包括中间件组件 (例如，应用程序服务器) 的计算系统中、或者包括前端组件 (例如，具有图形用户界面或 Web 浏览器的客户计算机，用户可通过图形用户界面或 Web 浏览器与这里公开的实施方案进行交互) 的计算机系统中、或者在一个或更多个这样的后端组件、中间件组件或前端组件的任何组合中实现所公开的实施例。系统的组件可通过任何数字数据通信的形式或媒介 (例如，通信网络) 互连。通信网络的例子包括局域网 (LAN) 和广域网 (WAN)(例如互联网)。

[0062] 计算系统可包括客户机和服务器。通常客户机和服务器彼此远离并且典型地通过通信网络交互。通过在各自的计算机上运行并彼此具有客户机 - 服务器关系的计算机程序的作用而产生客户机和服务器的关系。

[0063] 虽然本说明书包含许多具体说明，但是这不应被解释为对当前所要求保护的或可能要求保护的范围的限制，而应被解释为对特定实施例的具体特征的描述。也可将在多个分开的实施例的语境下在此说明书中描述的某些特征以组合的方式实现在单个实施例中。反过来，也可将在单个实施例的语境下描述的各种特征分开地或者以任何合适的子组合实现在多个实施例中。此外，虽然以上可能将特征描述为按某种组合来起作用，甚至最初也是如此要求保护的，但是在一些情况下要求保护的组合中的一个或更多个特征可被从该组合中去除，并且所要求保护的组合可以针对子组合或者子组合的变体。

[0064] 类似地，虽然在图中以特定顺序描绘操作，但是这不应被理解成要求以所示的特定顺序或者以顺次的顺序执行这样的操作，或者要求执行所有被示出的操作，以达到期望的结果。在某些情况下，多任务和并行处理可以是有利的。此外，上述实施例中对各种系统组件的划分不应被理解为在所有实施例中都要求这样的划分，而应将它理解为 :一般可将所描述的程序组件和系统一起集成在单个软件产品中，或者打包成多个软件产品。

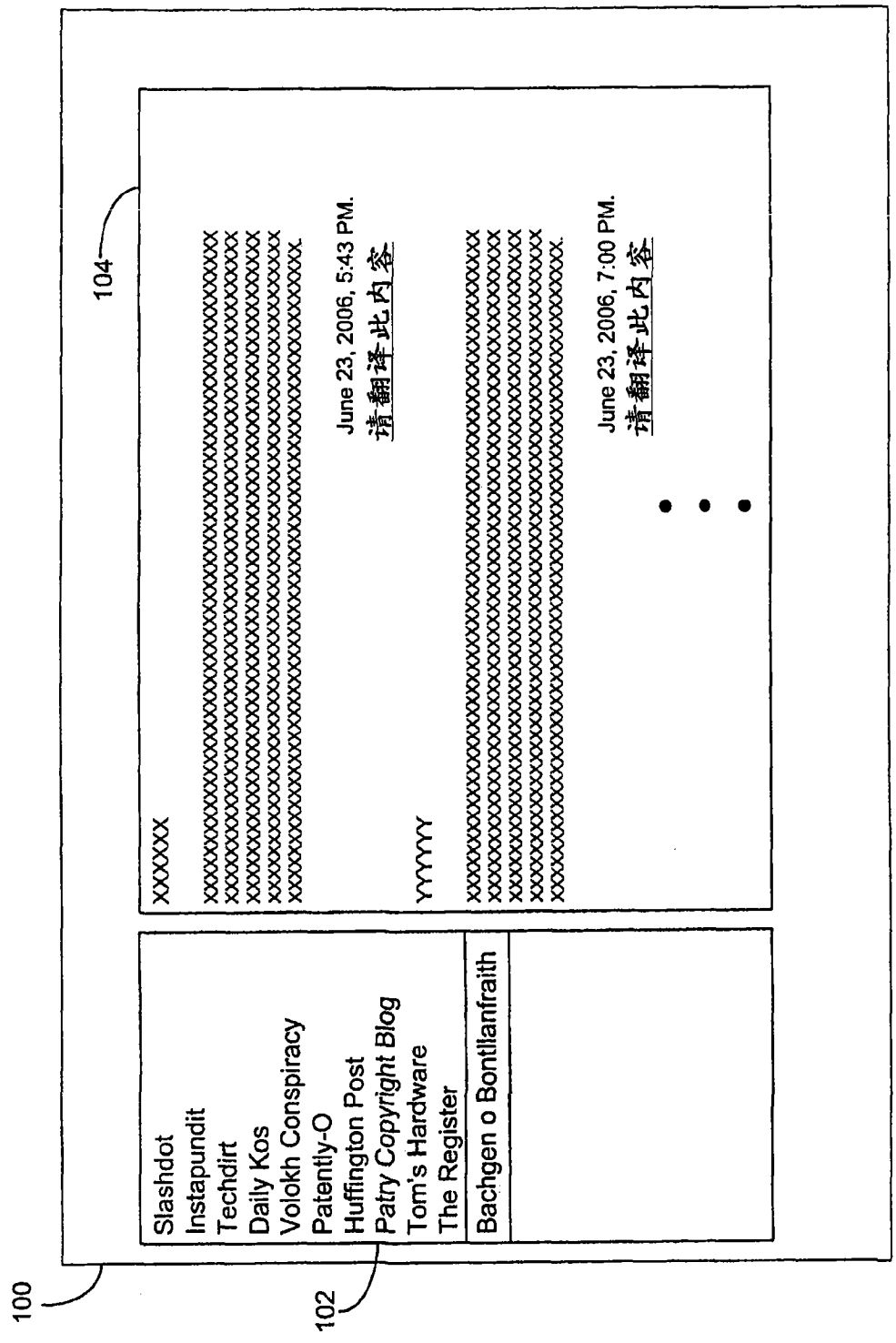

100
102
104
Slashdot
Instapundit
Techdirt
Daily Kos
Volokh Conspiracy
Patently-O
Huffington Post
Patry Copyright Blog
Tom's Hardware
The Register
Bachgen o Bontllanfraith
XXXXXX
June 23, 2006, 5:43 PM.
请翻译此内容
YYYYYY
June 23, 2006, 7:00 PM.
请翻译此内容

图 1

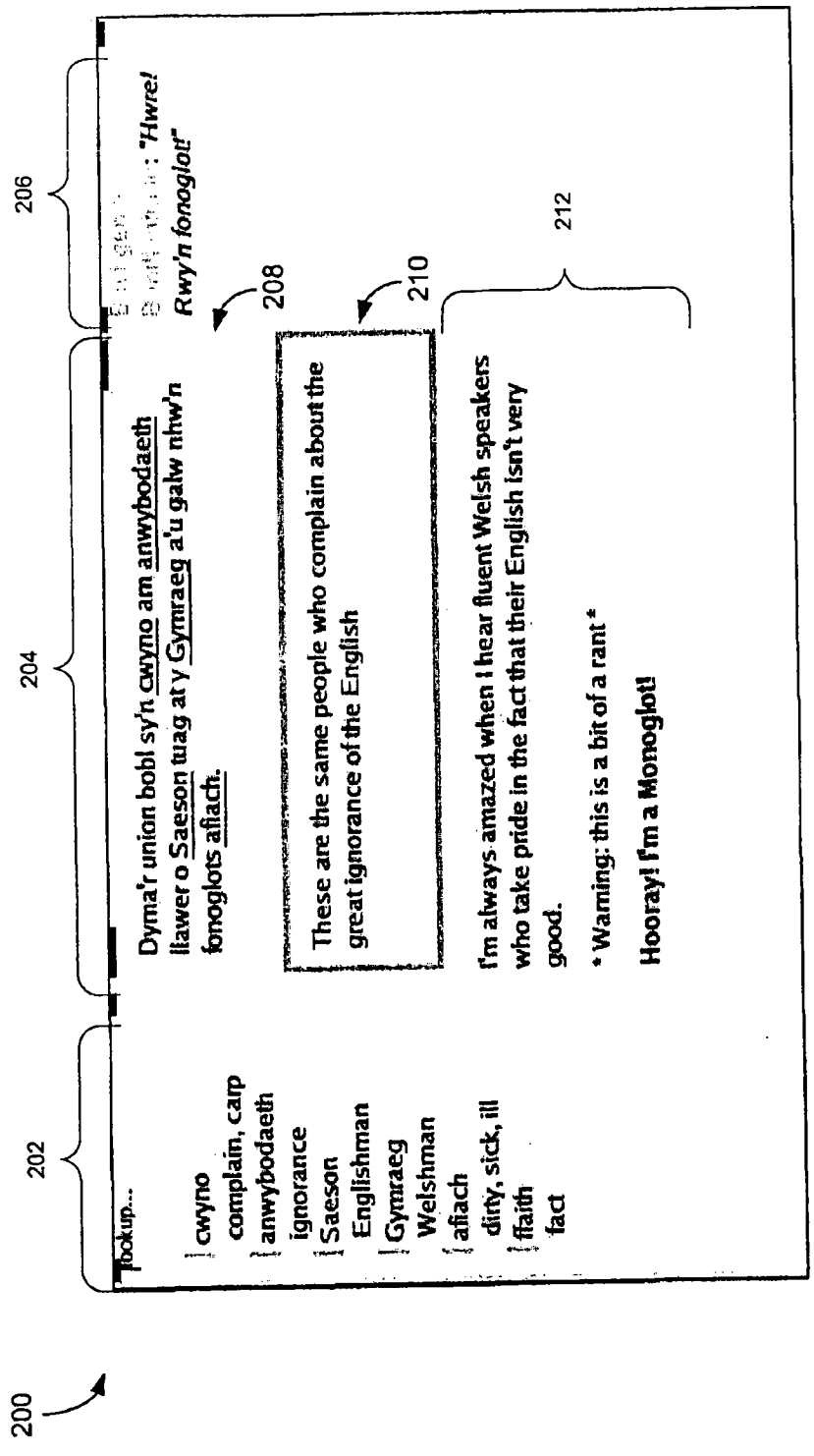

200
202
204
206
208
210
212
lookup...
cwyno
complain, carp
anwybodaeth
ignorance
Saeson
Englishman
Gymraeg
Welshman
afiach
dirty, sick, ill
ffaith
fact
Dyma'r union bobl sy'n cwyno am anwybodaeth llawer o Saeson tuag at y Gymraeg a'u galw nhw'n fonoglots afiach.
These are the same people who complain about the great ignorance of the English
I'm always amazed when I hear fluent Welsh speakers who take pride in the fact that their English isn't very good.
* Warning: this is a bit of a rant *
Hooray! I'm a Monoglot!
Rwy'n fonoglot"

图 2

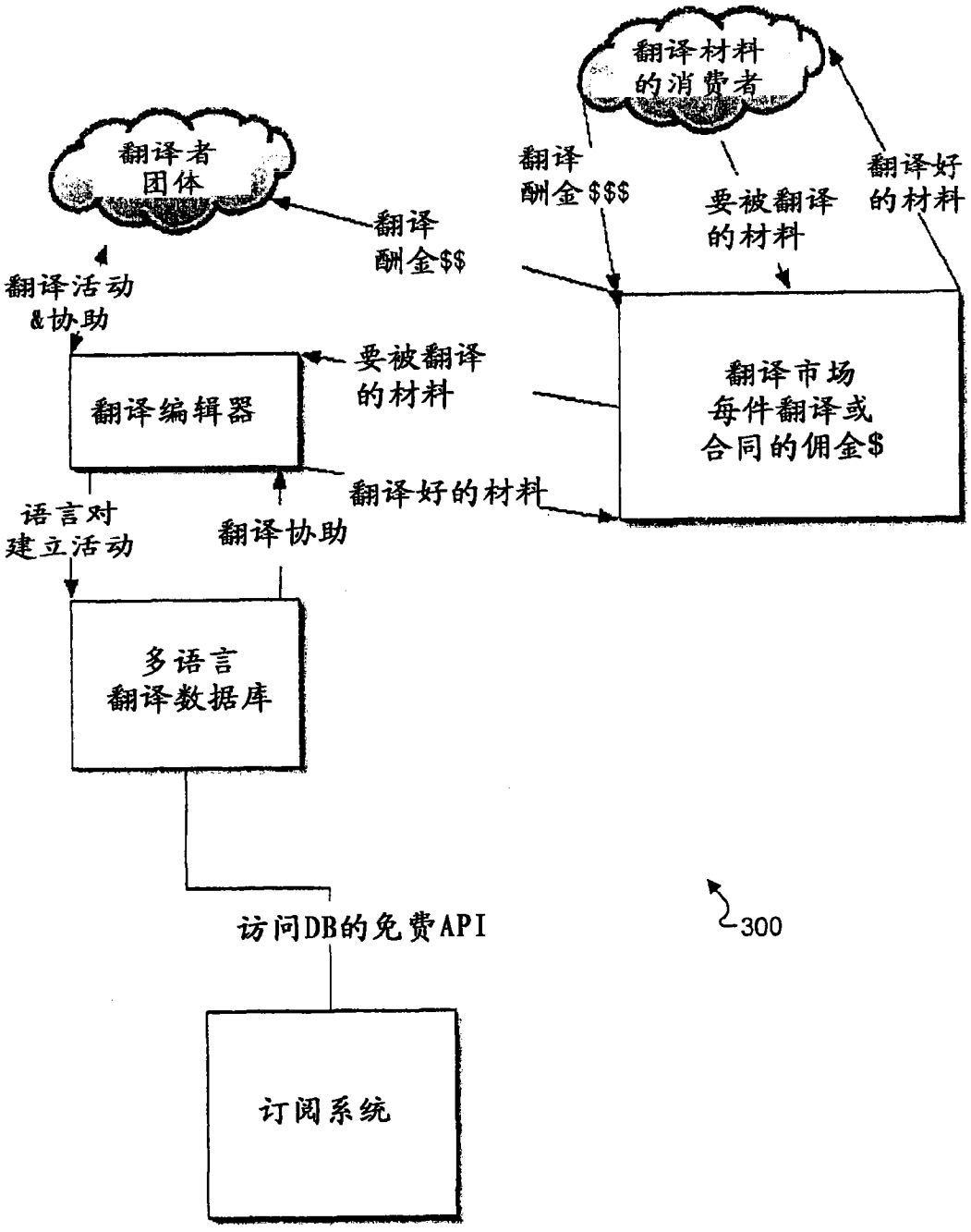

翻译材料
的消费者
翻译者
团体
翻译
酬金$$
翻译
酬金$$$
要被翻译
的材料
翻译好
的材料
翻译活动
&协助
翻译编辑器
要被翻译
的材料
翻译市场
每件翻译或
合同的佣金$
翻译好的材料
语言对
建立活动
翻译协助
多语言
翻译数据库
访问DB的免费API
300
订阅系统

图 3

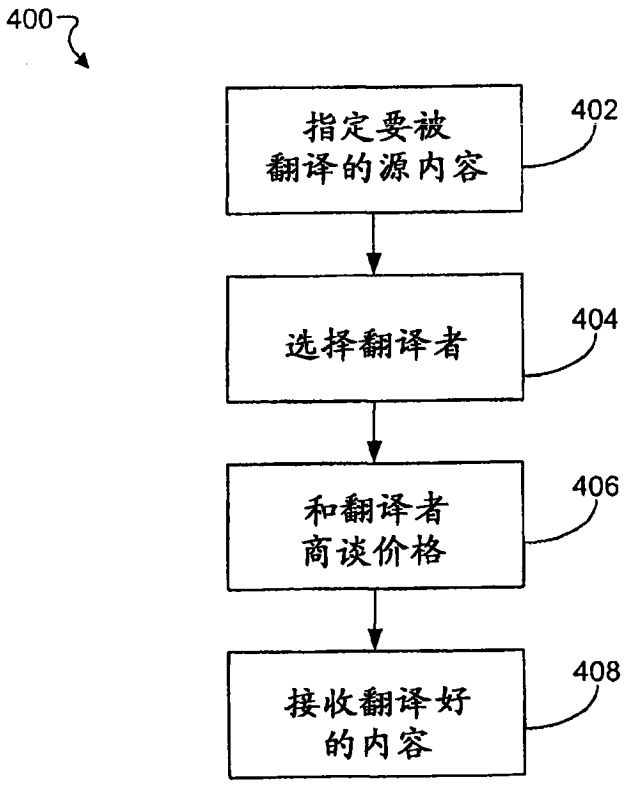


图 4

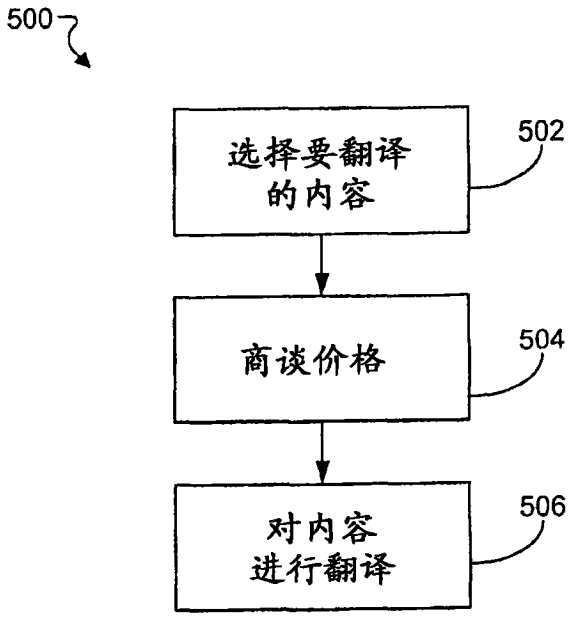


图 5

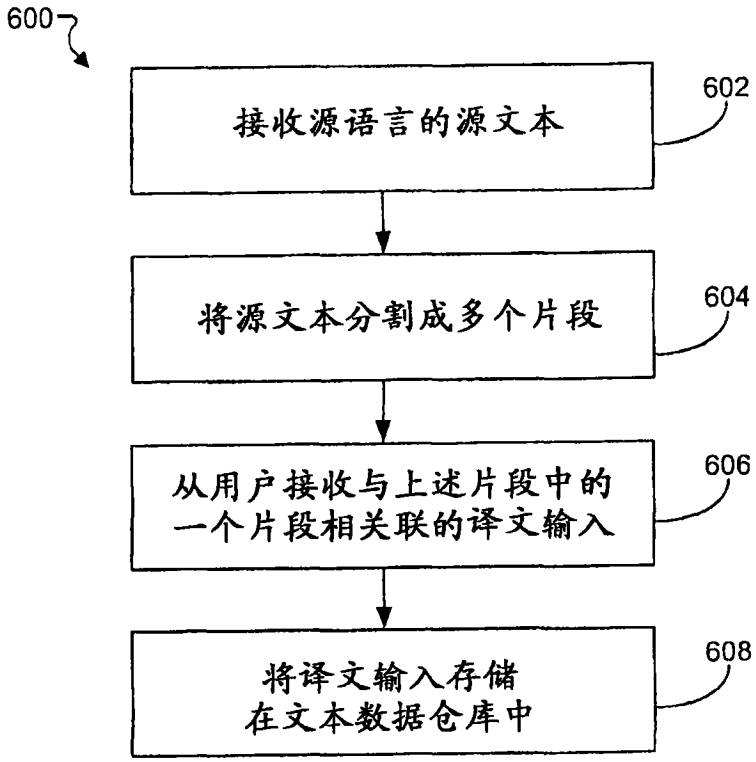

600
接收源语言的源文本
602
将源文本分割成多个片段
604
从用户接收与上述片段中的一个片段相关联的译文输入
606
将译文输入存储在文本数据仓库中
608

图 6

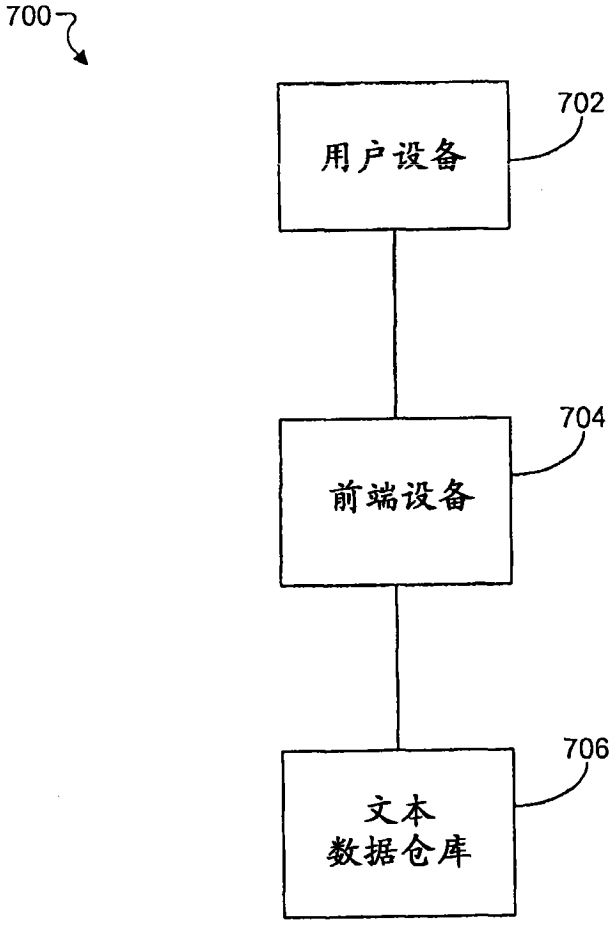


图 7