

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G06F 3/00

G06F 17/00 G06F 17/21

[12] 发明专利申请公开说明书

[21] 申请号 01117945.7

[43] 公开日 2001 年 11 月 21 日

[11] 公开号 CN 1322981A

[22] 申请日 2001.5.8 [21] 申请号 01117945.7

[30] 优先权

[32]2000.5.4 [33]US [31]60/202101

[32]2000.10.17 [33]US [31]09/690313

[71] 申请人 国际商业机器公司

地址 美国纽约州

[72] 发明人 J·L·基西

G·J·威尔莫特

[74] 专利代理机构 中国专利代理(香港)有限公司

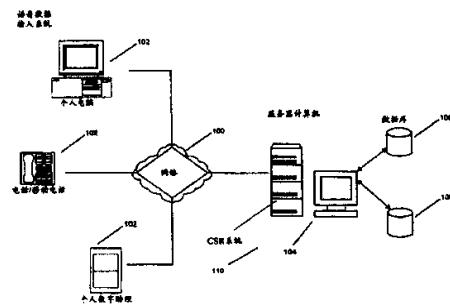
代理人 吴立明 陈景峻

权利要求书 3 页 说明书 6 页 附图页数 3 页

[54] 发明名称 向处理能力有限的设备提供的连续语音识别技术

[57] 摘要

提供一种在设备上输入数据的技术。一开始,在设备上接收语音数据。将语音数据和设备标识符传送到计算机。在计算机上将语音数据翻译成文本。然后确定是否要过滤翻译文本。如果确定要过滤翻译文本,就对翻译文本应用一个过滤器。



ISSN 1008-4274

权 利 要 求 书

1. 一种在设备上进行数据输入的方法，包含：
在设备上接收语音数据；
5 将语音数据和设备标识符传送到计算机；
并在计算机上
将语音数据翻译成文本；
确定是否要过滤翻译文本；和
如果确定翻译文本要被过滤，就对翻译文本应用一个过滤器。
- 10 2. 权利要求 1 的方法，进一步包含在与计算机相连的数据储存库中存储用户个人简档。
3. 权利要求 2 的方法，其中，用户个人简档包含语音印迹。
4. 权利要求 3 的方法，进一步包含用语音印迹将语音数据翻译成文本。
- 15 5. 权利要求 1 的方法，其中的确定包含从翻译文本析取一个或多个关键词。
6. 权利要求 5 的方法，其中，根据一个或多个被析取的关键词选择过滤器。
7. 权利要求 1 的方法，其中，应用过滤器包含格式化翻译文本。
- 20 8. 权利要求 7 的方法，其中，格式化包含为应用程序格式化翻译文本。
9. 权利要求 7 的方法，其中，格式化包含为该设备格式化翻译文本。
10. 权利要求 1 的方法，进一步包含将翻译文本返回到设备。
- 25 11. 权利要求 1 的方法，进一步包含将过滤文本返回到设备。
12. 权利要求 11 的方法，进一步包含通过电子邮件消息返回过滤文本。
13. 权利要求 1 的方法，进一步包含将数据返回到接收语音数据的设备以外的设备。
- 30 14. 一种装置，包含：
一个用于接收和传送数据的设备；
一个与该设备相连的、有与其相连的数据储存库的计算机，其中，

数据储存库存储数据；和

由该计算机执行的一个或多个计算机程序，用于：

从该设备接收语音数据和设备标识符；

将语音数据翻译成文本；

5 确定是否要过滤翻译文本；和

如果确定翻译文本要被过滤，就对翻译文本应用一个过滤器。

15. 权利要求 14 的装置，进一步包含在与计算机相连的数据储存库中存储用户个人简档。

16. 权利要求 15 的装置，其中，用户个人简档包含语音印迹。

10 17. 权利要求 16 的装置，进一步包含用语音印迹将语音数据翻译成文本。

18. 权利要求 14 的装置，其中的确定包含从翻译文本析取一个或多个关键词。

15 19. 权利要求 18 的装置，其中，根据一个或多个被析取的关键词选择过滤器。

20. 权利要求 14 的装置，其中，应用过滤器包含格式化翻译文本。

21. 权利要求 20 的装置，其中，格式化包含为应用程序格式化翻译文本。

20 22. 权利要求 20 的装置，其中，格式化包含为该设备格式化翻译文本。

23. 权利要求 14 的装置，进一步包含将翻译文本返回到设备。

24. 权利要求 14 的装置，进一步包含将过滤文本返回到设备。

25 25. 权利要求 24 的装置，进一步包含通过电子邮件消息返回过滤文本。

26. 权利要求 14 的装置，进一步包含将数据返回到接收语音数据的设备以外的设备。

27. 一种包含计算机可读的程序存储介质并实施可由计算机执行以处理在设备上输入的数据的一个或多个指令的产品，包含：

在设备上接收语音数据；

30 将语音数据和设备标识符传送到计算机；

在计算机上，

将语音数据翻译成文本；

确定是否要过滤翻译文本；和

如果确定翻译文本要被过滤，就对翻译文本应用一个过滤器。

28. 权利要求 27 的产品，进一步包含在与计算机相连的数据储存库中存储用户个人简档。

5 29. 权利要求 28 的产品，其中，用户个人简档包含语音印迹。

30. 权利要求 29 的产品，进一步包含用语音印迹将语音数据翻译成文本。

31. 权利要求 27 的产品，其中的确定包含从翻译文本析取一个或多个关键词。

10 32. 权利要求 31 的产品，其中，根据一个或多个被析取的关键词选择过滤器。

33. 权利要求 27 的产品，其中，应用过滤器包含格式化翻译文本。

34. 权利要求 33 的产品，其中，格式化包含为应用程序格式化翻译文本。

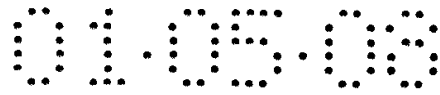
15 35. 权利要求 33 的产品，其中，格式化包含为该设备格式化翻译文本。

36. 权利要求 27 的产品，进一步包含将翻译文本返回到设备。

37. 权利要求 37 的产品，进一步包含将过滤文本返回到设备。

20 38. 权利要求 37 的产品，进一步包含通过电子邮件消息返回过滤文本。

39. 权利要求 27 的产品，进一步包含将数据返回到接收语音数据的设备以外的设备。



说明书

向处理能力有限的设备提供的连续语音识别技术

5 本发明一般地说涉及计算机实现的系统，更具体来说，涉及向诸如个人数字助理（PDA）的处理能力有限的设备提供作为替代输入设备的连续语音识别。

个人数字助理（PDA）是一种将计算与诸如电话和/或网络连接的其它特征功能综合在一起的手持式设备。许多 PDA 被用作个人记事本
10 （organizer），包含日程表、电子邮件系统和字处理器。向 PDA 的输入通常是通过手写笔（stylus）而不是通过键盘或鼠标进行的。手写笔是一种“像钢笔一样的”物体，用来在诸如数字书写板（digital tablet）的屏幕上写数据。手写笔有用来触及数字书写板的电子头，数字书写板含有的电子电路使其能检测手写笔的移动并将这些移动转
15 换成用于计算机的数字信号。

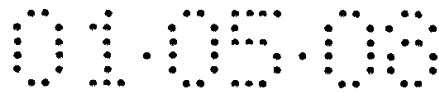
有些 PDA 含有使用户能用手写笔在屏幕上“手写”数据的手写识别功能。然而，一般的手写识别系统有时错误解释手写的数字而要求用户仔细检查和修正手写的数字。

PDA 已经变得非常流行，正越来越多地被各种人广泛使用。遗憾的是，这些小设备只有有限的存储器容量，较小的显示器，和较慢的运行速度。此外，由于用手写笔来输入数据，使有些残疾人不能使用
20 PDA。

因此，在该领域中需要有向具有有限资源的设备输入数据的改进技术。

25 为了克服以上所述的现有技术中的缺陷，克服在阅读和理解本说明书时将逐渐明了的其它缺陷，本发明公开一种用于向诸如个人数字助理（PDA）的处理能力有限的设备提供作为替代输入设备的连续语音识别的技术的方法、装置和产品。

按照本发明一个实施例，提供一项在设备上输入数据的技术。一开始，在设备上接收语音数据。将语音数据和设备标识符传送到计算机。在计算机上，将语音数据翻译成文本。然后确定是否要过滤翻译
30 文本。如果确定翻译文本要过滤，就对翻译文本应用过滤器（filter）。



现在参看以下附图 - 其中相同的标注号自始至终代表相应的部件:

图 1 是表示本发明一个实施例的硬件环境的示意图;

图 2 是表示本发明一个实施例中的 CSC 系统 212 及其环境的示意图;

图 3 是表示由本发明一个实施例中的 CSC 系统 212 执行的过程的流程图。

在以下对本发明实施例的说明中, 参照构成其一部分的附图, 说明书是以能实现本发明的特定实施例的方式说明的。应当理解, 可以采用其它实施例, 因为在不偏离本发明范围的情况下可以进行结构上的修改。

图 1 是表示本发明一个实施例的硬件环境的示意图, 更具体来说, 该图表示一个典型的分布式计算机系统, 它用网络 100 把语音数据输入设备 102 (客户机) 与执行计算机程序的服务器计算机 104 连接, 并把服务器系统 104 与数据源 106 连接。数据源 106 可以存储例如包括语音印迹 (voice print) 记录的用户个人简档 (profile)。典型的资源组合可包括语音数据输入设备 102 - 例如是个人电脑或工作站、电话或移动电话、或个人数字助理 (PDA)。服务器计算机 104 例如可以是个人电脑、工作站、小型计算机或大型机。这些系统是由 LAN、WAN、SNA 网及因特网在内的各种网络连接的。有些语音数据输入设备 102 (例如个人电脑或个人数字助理) 和服务器计算机 104 另外还包含一个操作系统和一个或多个计算机程序。

服务器软件包括一个连续语音识别 (CSR) 系统 110, 它包含一个或多个计算机程序用于将语音翻译成文本、过滤文本以及将文本转换成适当格式。服务器计算机 104 也用数据源接口和可能的其它计算机程序来连接到数据源 106。语音数据输入设备 102 通过线路或无线系统与服务器计算机 104 双向连接。服务器计算机 104 进而与数据源 106 双向连接。

操作系统和计算机程序是由指令组成的, 这些指令在被语音数据输入设备 102 和服务器计算机 104 读取并执行时, 使设备和计算机执行实现和/或使用本发明所需的步骤。一般地说, 操作系统和计算机程序被安装在和/或可读自诸如存储器的设备、载体或介质, 其它数据存

储设备和/或数据通信设备。在操作系统的控制下，可以将计算机程序从存储器、其它数据存储设备和/或数据通信设备装入计算机的存储器，供在实际运行期间使用。

5 所以，可以用生产软件、固件、硬件及其任意组合的标准程序设计和/或工程技术，把本发明实现为方法、装置或产品。本文中所使用的术语“产品”（或者“计算机程序产品”）的意思包括能从任何计算机可读设备、载体或介质获得的计算机程序。当然，本领域的技术人员知道，在不偏离本发明范围的条件下可以对这个结构作出许多修改。

10 本领域的技术人员将认识到，图 1 中所示的示意性环境并非是要限定本发明。实际上，本领域的技术人员将认识到，在不偏离本发明范围的条件下，可以采用其它替代性的硬件环境。

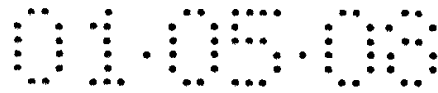
连续语音识别系统

15 在一个实施例中，本发明提供一个连续语音识别（CSR）系统。该 CSR 系统使具有有限处理能力的设备能进行连续语音识别。就是说，大多数手持式设备（例如 PDA 或移动电话）不具有进行连续语音识别的处理能力。这一点加上它们体积小，用户只好用手写笔在输入区上点触，这使得这些设备极其难以为残疾人所用。这也妨碍使用者快速记录笔记、更新日程表或发送电子邮件。

20 用 CSR 系统，向设备输入数据变得像说话一样简单。可以设想 CSR 系统能消除对点触输入设备的需要。CSR 系统也便于将体积小得不能有输入板或屏幕的设备（如手腕佩戴设备）用作输入设备。

25 图 2 是表示本发明一个实施例中的 CSC 系统 212 及其环境的示意图。CSC 系统 212 位于语音识别服务器 210。CSC 系统 212 在一个或多个客户设备（有限处理能力的设备）与一个或多个语音识别服务器之间建立一种共生的关系。为方便解释，图中描绘了一个客户设备 200 和一个语音识别服务器 210。客户设备 200 能记录和/或中继语音。CSC 系统 212 包含语音-文本翻译软件 214 和文本过滤和转换软件 216。

30 一般地说，客户设备 200 捕获语音并将其发送到语音识别服务器 210 供翻译和转换。语音识别服务器 210 将转换过的信息发送回客户设备 200，后者然后将其结合到目标应用程序（例如日程表、电子邮件或笔记本）中。



5 在使用 CSC 系统 212 之前，用户向语音识别服务器 210 提交信息。该信息包含一个在数据储存库中存储的用户个人简档 218。用户个人简档包含一个与用户说话方式相关联的“语音印迹”，关于一个或多个要接收数据的目标应用程序的信息，一个或多个标识用户所用的特定设备的客户设备（单元）标识符（ID），和包括电子邮件地址在内的用户联系信息。

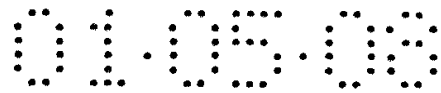
10 一开始，用户录制在语音识别服务器上作为语音印迹存储的语音。例如，可以要求每个用户说出特定的文字，诸如一本书中的某段落。所说的文字就是语音印迹。每个用户说话的方式都有些不同，在停顿和重音上都略有差异。所以，可以用语音印迹来标识用户。此外，语音印迹也被 CSR 系统 212 用来更好地将语音转换成文本。

15 用户个人简档被存储在语音识别服务器 210 后，用户就能通过向位于客户设备 200 的语音录音/中继机说话而向客户设备 200 输入语音数据。用户说关键词和其它语句。关键词向 CSR 系统 212 指出紧接着的信息的特定类型。关键词的例子包括 - 但不限于 - 下列：日程表项 (CALENDAR ENTRY)、日期 (DATE)、时间 (TIME)、发送记事 (SEND NOTE)、地址项 (ADDRESS ENTRY)、记事本项 (NOTEPAD ENTRY)。要在日程表应用程序中安排会议时间，用户可以向客户设备 200 说出下列信息：日程表项日期 (CALENDAR ENTRY DATE) 2000 年 12 月 1 日时间 (TIME) 上午 20 10 点主题 (SUBJECT) 关于项目的会议。

25 客户设备 200 用该语音数据生成一个语音包，语音包的组成是语音数据（例如短语）、适合目标应用程序（例如日程表应用程序）的数据和单元 ID（客户设备标识符）。客户设备 200 通过任何能用的通信系统 - 如移动调制解调器连接和/或因特网连接，把语音包发送到语音识别服务器 210。

语音识别服务器 210 接收语音包，从中析取单元 ID，用它从数据储存库检索用户的语音印迹。语音-文本翻译软件 214 利用语音印迹把语音包中的语音数据翻译成文本。结果生成“翻译文本”。

30 下一步，文本过滤和转换软件 216 试图从翻译文本中析取一个或多个关键词。在一个实施例中，期望这一个或多个关键词位于翻译文本的开始。如果找不到关键词，CSR 系统 212 就例如通过电子邮件将翻译文本返回到客户设备 200。另一方面，如果析取出一个或多个关



5 关键词，CSR 系统 212 就确定并提取一个转换过滤器 220，它将被用来将翻译文本格式化成为特定一个格式（例如特定于某特定应用程序和/或特定设备）。例如，如果一个或多个关键词表明语音数据是与日程表应用程序相关的，代表一个日程表项（CALENDAR ENTRY），文本过滤和转换软件 216 就确定要被使用的转换过滤器，并从转换过滤器 220 中提取一个日程表过滤器，以便把要向客户设备 200 发送的数据格式化成日程表项。这种格式化不仅为特定一个应用程序（例如日程表应用程序）格式化翻译文本，也为特定一个客户设备 200（例如特定品牌的 PDA）格式化翻译文本。然后，CSR 系统 212 用适当的通信通道（例如通过在移动调制解调器上和/或因特网上发送电子邮件）把过滤过的文本返回到客户设备 200。客户设备 200 接收经过翻译和转换的语音包，将其传送到目标应用程序（例如日程表应用程序）去处理。

10 如果客户设备 200 是移动电话，用户能通过移动电话输入语音。该语音和单元 ID 被发送到语音识别服务器 210。位于语音识别服务器 210 的 CSR 系统 212 将该语音数据转换成翻译文本，应用适当的过滤器来生成过滤文本，并按用户个人简档中的规定，通过电子邮件向用户的设备返回翻译文本或过滤文本。

15 所以，有了 CSR 系统 212，要在日程表应用程序中安排会议时间，用户可以向客户设备 200 说出下列信息：日程表项日期(CALENDAR ENTRY DATE) 2000 年 12 月 1 日时间 (TIME) 上午 10 点主题 (SUBJECT) 关于项目的会议。然后，CSR 系统 212 将该语音数据格式化位日程表项，准备好插入日程表中。而在一般的系统中，用户则要打开日程表应用程序，确定日期和时间，再键入或写入主题信息。在 PDA 上，这通常要求使用手写笔，这对许多人来说是难以使用的，对残疾人来说尤其如此。此外，用一般的系统也不可能仅以移动电话就能生成一个日程表项。

20 图 3 是表示由本发明一个实施例中的 CSC 系统 212 执行的过程的流程图。应当理解，在一个实施例中，CSR 系统 212 既包含语音-文本翻译软件 214 也包含文本过滤和转换软件 216。

25 在方框 300 中，CSR 系统 212 接收一个包含语音印迹和单元 ID 在内的用户个人简档 218，并在语音识别服务器 210 存储该用户个人简档 218。在方框 302 中，客户设备 200 接收语音数据并将该语音数据和单元 ID 转发到语音识别服务器 210。在方框 304 中，位于语音识别

服务器 210 的 CSR 系统 212 根据单元 ID 提取用户的语音印迹。在方框 306 中，CSR 系统 212 利用语音印迹将语音数据转换成文本，生成翻译文本。在方框 308 中，CSR 系统 212 确定是否要应用一个过滤器。如果是，CSR 系统 212 就继续到方框 312，否则，CSR 系统 212 就继续到方框 310。在方框 310 中，CSR 系统 212 将翻译文本返回到客户设备 200。在方框 312 中，CSR 系统 212 选择并提取一个转换过滤器 220。在方框 314 中，CSR 系统 212 将该转换过滤器应用到翻译文本，生成过滤文本。在方框 316 中，CSR 系统 212 将过滤文本返回到客户设备 200。在一个实施例中，CSR 系统 212 将过滤文本返回到客户设备 200 上的一个应用程序。

至此说明了本发明的实施例。以下说明用于实现本发明的一些其它形式的实施例。例如，任何类型的计算机（如大型机、小型计算机或个人电脑）或计算机结构（如分时大型机、局域网或独立个人电脑）都可以用于本发明。

对本发明的实施例的以上说明是为方便解释和说明而陈述的。上述说明不是穷举的，也不是要把本发明完全限定于所陈述的形式。按照以上原理，有许多可能的修改和变体。本发明的范围不受该详细说明的限定，而是受后附的权利要求的限定。

说明书附图

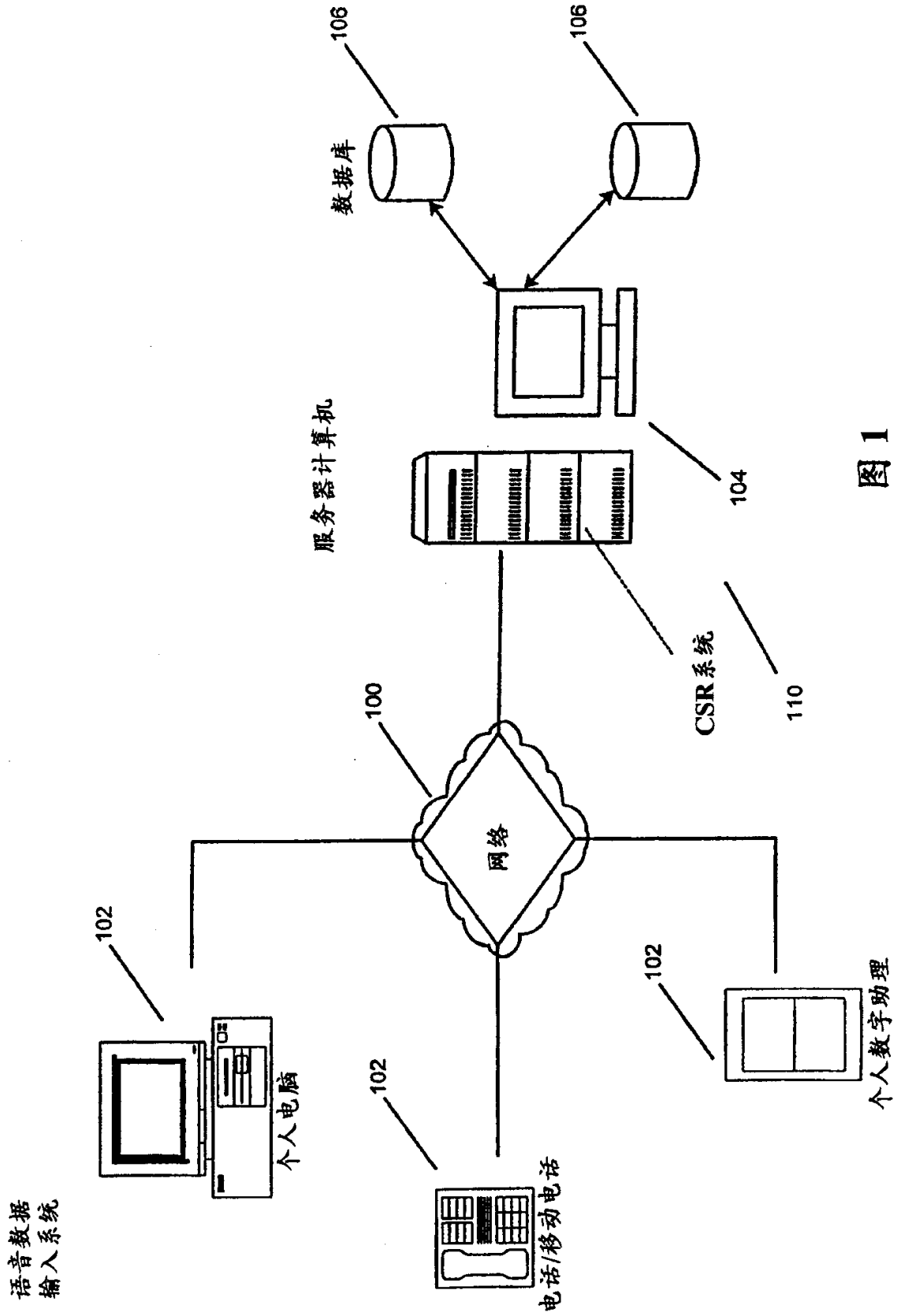


图 1

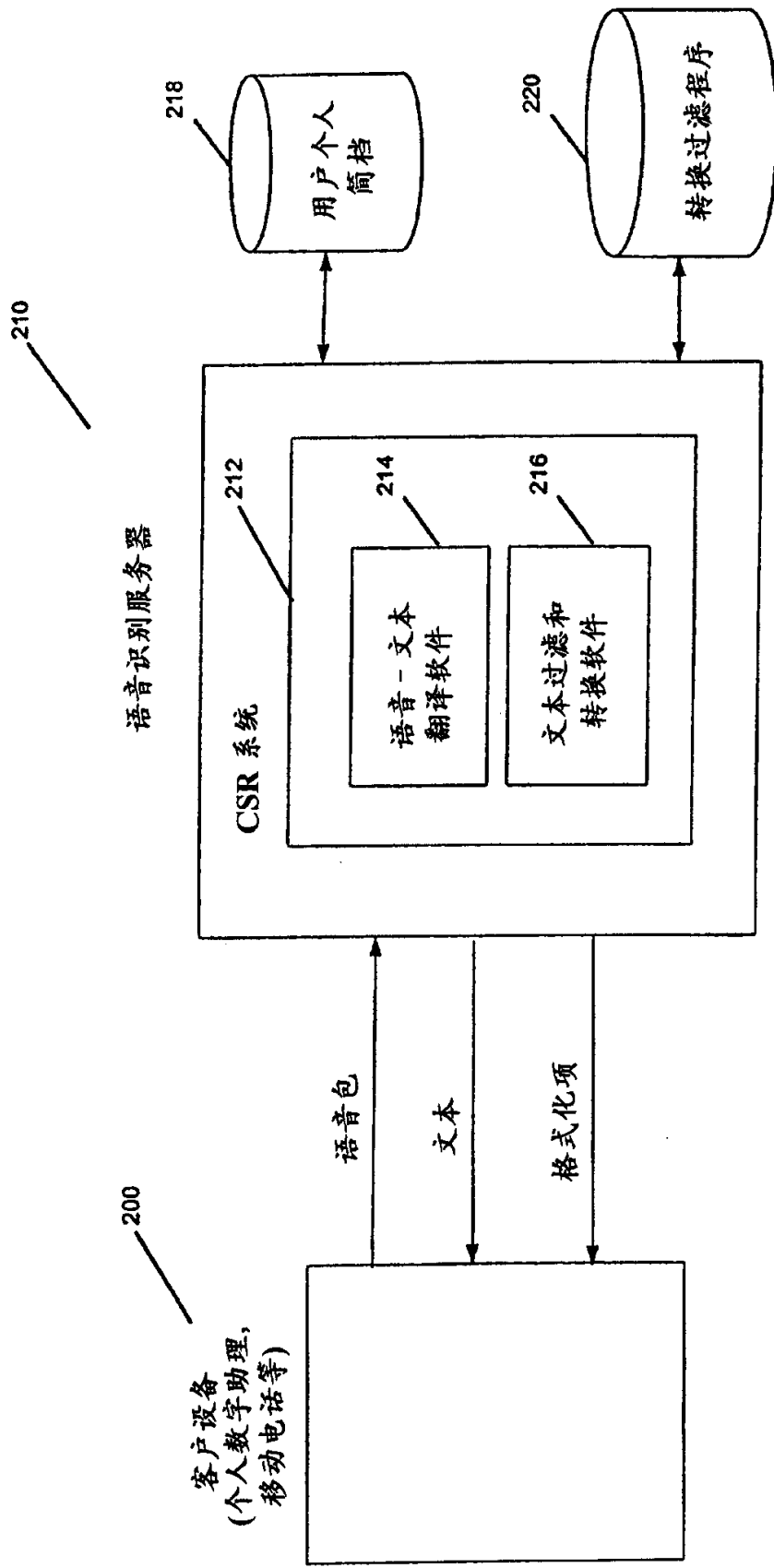


图 2

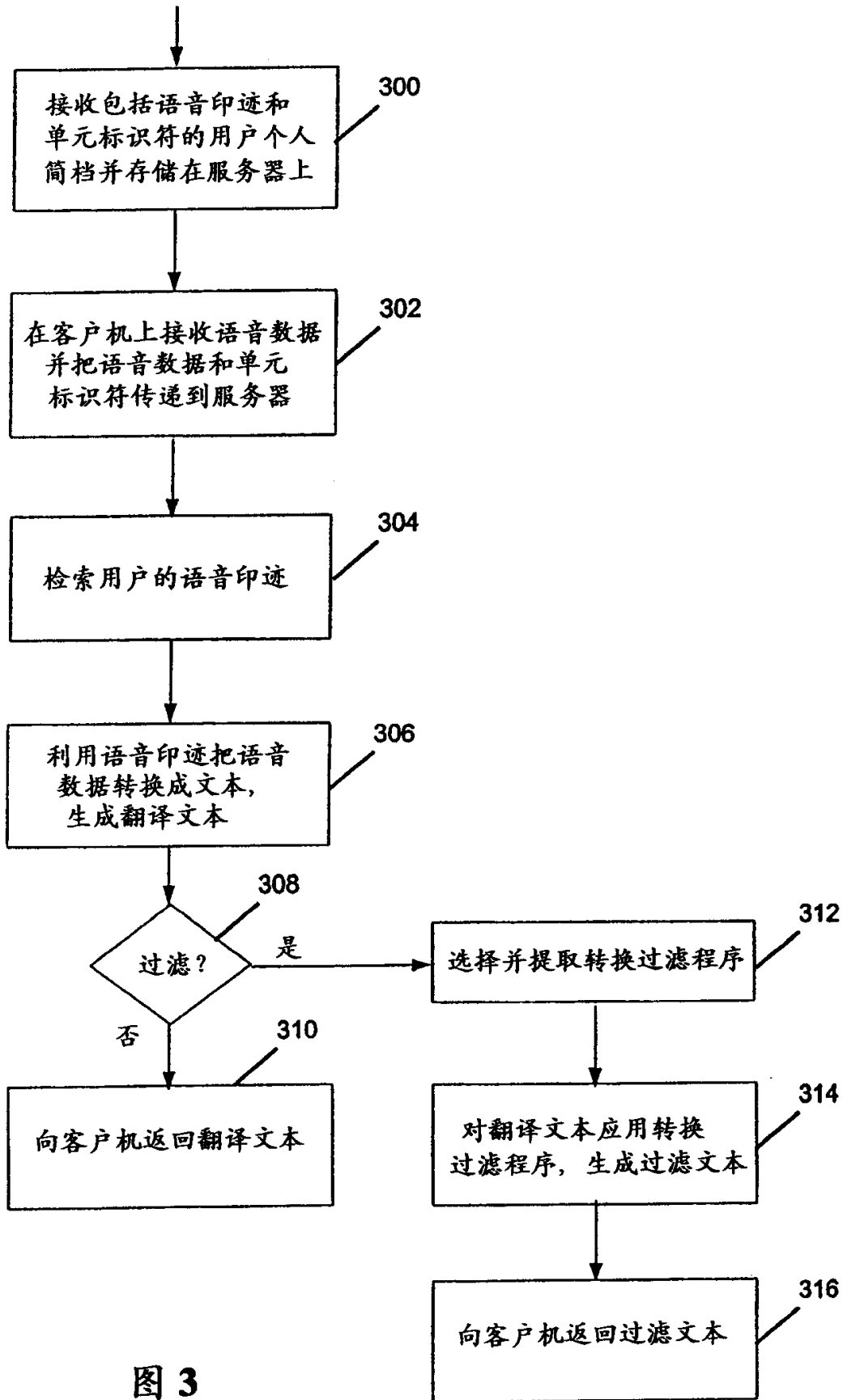


图 3