

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G10L 15/06 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200680020928.X

[43] 公开日 2008年6月4日

[11] 公开号 CN 101194305A

[22] 申请日 2006.8.1

[21] 申请号 200680020928.X

[30] 优先权

[32] 2005.8.19 [33] US [31] 11/208,366

[86] 国际申请 PCT/US2006/030011 2006.8.1

[87] 国际公布 WO2007/024427 英 2007.3.1

[85] 进入国家阶段日期 2007.12.12

[71] 申请人 思科技术公司

地址 美国加利福尼亚州

[72] 发明人 凯文·L·切斯特纳特

约瑟夫·B·伯顿

[74] 专利代理机构 北京东方亿思知识产权代理有限  
责任公司

代理人 王怡

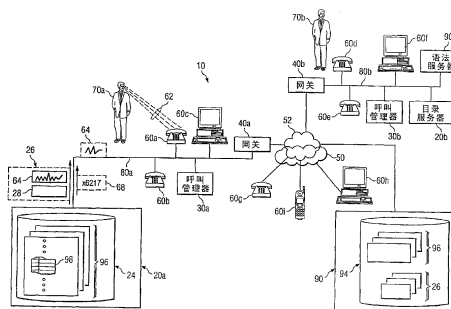
权利要求书6页 说明书22页 附图4页

## [54] 发明名称

用于分发语音识别语法的系统和方法

## [57] 摘要

一种用于分发语音识别语法的方法包括从第一远程元件接收匹配数据。匹配数据包括与远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息。该方法还包括基于匹配数据生成语法条目。语法条目包括第二存储音频数据和与第二存储音频数据相关联的单词标识符。另外，该方法包括将语法条目发送到第二远程元件。



1. 一种用于分发语音识别语法的方法，包括：

从第一远程元件接收匹配数据，所述匹配数据包括与所述远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息；

基于所述匹配数据生成语法条目，其中所述语法条目包括第二存储音频数据和与所述第二存储音频数据相关联的单词标识符；以及

将所述语法条目发送到第二远程元件。

2. 如权利要求 1 所述的方法，其中所述第一远程元件和所述第二远程元件包括不同的远程元件。

3. 如权利要求 1 所述的方法，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的错误发音相关联的音频数据。

4. 如权利要求 1 所述的方法，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的缩略相关联的音频数据。

5. 如权利要求 1 所述的方法，其中所述单词标识符标识一个或多个姓名，并且其中所述第二存储音频数据包含与一个或多个标识的姓名的昵称相关联的音频数据。

6. 如权利要求 1 所述的方法，其中将所述语法条目发送到所述第二远程元件的步骤包括：

从多个远程元件中识别一个远程元件；以及

将所述语法条目发送到识别出的远程元件。

7. 如权利要求 6 所述的方法，其中所述匹配数据指示与所述第一远程元件相关联的语言，并且其中从多个远程元件中识别一个远程元件的步骤包括基于与识别出的所述远程元件相关联的语言来识别所述远程元件。

8. 如权利要求 6 所述的方法，其中所述匹配数据指示与所述第一远程元件相关联的位置，并且其中从多个远程元件中识别一个远程元件的步骤包括基于与识别出的所述远程元件相关联的位置来识别所述远程元件。

9. 如权利要求 1 所述的方法，其中从第一远程元件接收匹配数据的步骤包括：

从多个远程元件接收匹配数据；以及

汇聚从两个或更多个远程元件接收的所述匹配数据，并且其中生成所述语法条目的步骤包括基于汇聚的匹配数据来生成所述语法条目。

10. 如权利要求 9 所述的方法，其中汇聚所述匹配数据的步骤包括：

从所述多个远程元件中选择两个或更多个远程元件；以及

汇聚从所选远程元件接收的所述匹配数据。

11. 如权利要求 1 所述的方法，其中从第一远程元件接收匹配数据的步骤包括：

从所述第一远程元件接收第一组匹配数据；

从所述第一远程元件接收第二组匹配数据；以及

汇聚所述第一组匹配数据和所述第二组匹配数据；并且

其中生成所述语法条目的步骤包括基于汇聚的匹配数据来生成所述语法条目。

12. 如权利要求 1 所述的方法，还包括：

在所述第一远程元件处接收与用户所讲话语相关联的音频信息；

将接收到的所述音频信息与一个或多个语法条目相比较；

基于所述接收到的音频信息和所述第一存储音频数据的比较结果而生成所述匹配数据；以及

将所述匹配数据发送到语法服务器。

13. 一种用于分发语音识别语法的设备，包括：

可用于存储处理器指令的存储器；以及

处理器，其可用于：

从第一远程元件接收匹配数据，所述匹配数据包括与所述远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息；

基于所述匹配数据生成语法条目，其中所述语法条目包括第二存储音频数据和与所述第二存储音频数据相关联的单词标识符；以及

将所述语法条目发送到第二远程元件。

14. 如权利要求 13 所述的设备，其中所述第一远程元件和所述第二远程元件包括不同的远程元件。

15. 如权利要求 13 所述的设备，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的错误发音相关联的音频数据。

16. 如权利要求 13 所述的设备，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的缩略相关联的音频数据。

17. 如权利要求 13 所述的设备，其中所述单词标识符标识一个或多个姓名，并且其中所述第二存储音频数据包含与一个或多个标识的姓名的昵称相关联的音频数据。

18. 如权利要求 13 所述的设备，其中所述处理器可用于通过以下步骤来将所述语法条目发送到所述第二远程元件：

从多个远程元件中识别一个远程元件；以及  
将所述语法条目发送到识别出的远程元件。

19. 如权利要求 18 所述的设备，其中所述匹配数据指示与所述第一远程元件相关联的语言，并且其中所述处理器可用于通过以下步骤来从多个远程元件中识别所述远程元件：基于与识别出的所述远程元件相关联的语言来识别所述远程元件。

20. 如权利要求 18 所述的设备，其中所述匹配数据指示与所述第一远程元件相关联的位置，并且其中所述处理器可用于通过以下步骤来从多个远程元件中识别所述远程元件：基于与识别出的所述远程元件相关联的位置来识别所述远程元件。

21. 如权利要求 13 所述的设备，其中所述处理器可用于通过以下步骤来从所述第一远程元件接收匹配数据：

从多个远程元件接收匹配数据；以及

汇聚从两个或更多个远程元件接收的所述匹配数据，并且其中所述处理器可用于基于汇聚的匹配数据来生成所述语法条目。

22. 如权利要求 21 所述的设备，其中所述处理器可用于通过以下步骤来汇聚所述匹配数据：

从所述多个远程元件中选择两个或更多个远程元件；以及  
汇聚从所述两个或更多个远程元件接收的所述匹配数据。

23. 如权利要求 13 所述的设备，其中所述处理器可用于通过以下步骤来从第一远程元件接收匹配数据：

从所述第一远程元件接收第一组匹配数据；  
从所述第一远程元件接收第二组匹配数据；以及  
汇聚所述第一组匹配数据和所述第二组匹配数据；并且  
其中所述处理器可用于基于汇聚的匹配数据来生成所述语法条目。

24. 如权利要求 13 所述的设备，其中所述处理器还可用于：  
在所述第一远程元件处接收与用户所讲话语相关联的音频信息；  
将接收到的所述音频信息与一个或多个语法条目相比较；  
基于所述接收到的音频信息和所述第一存储音频数据的比较结果而生成所述匹配数据；以及  
将所述匹配数据发送到语法服务器。

25. 一种用于分发语音识别语法的系统，包括：

第一元件，可用于：

接收与话语相关联的音频数据；  
尝试将接收到的所述音频信息与存储的音频数据匹配；  
基于将接收到的所述音频信息与存储的所述音频数据匹配的尝试

来生成匹配信息；以及

将所述匹配信息发送到服务器；

服务器，可用于：

从所述第一元件接收所述匹配数据；  
基于所述匹配数据生成语法条目，其中所述语法条目包括第二存储音频数据和与所述第二存储音频数据相关联的单词标识符；以及  
将所述语法条目发送到远程元件；以及  
所述第二元件，可用于从所述服务器接收所述语法条目。

26. 如权利要求 25 所述的系统，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的错误发音相关联的音频数据。

27. 如权利要求 25 所述的系统，其中所述单词标识符标识一个或多个单词，并且其中所述第二存储音频数据包含与一个或多个标识的单词的缩略相关联的音频数据。

28. 如权利要求 25 所述的系统，其中所述单词标识符标识一个或多个姓名，并且其中所述第二存储音频数据包含与一个或多个标识的姓名的昵称相关联的音频数据。

29. 如权利要求 25 所述的系统，其中所述服务器可用于通过以下步骤来将所述语法条目发送到所述第二远程元件：

从多个远程元件中识别一个远程元件；以及  
将所述语法条目发送到识别出的远程元件。

30. 如权利要求 29 所述的系统，其中所述匹配数据指示与所述第一远程元件相关联的语言，并且其中所述服务器可用于通过以下步骤来从多个远程元件中识别所述远程元件：基于与识别出的所述远程元件相关联的语言来识别所述远程元件。

31. 如权利要求 29 所述的系统，其中所述匹配数据指示与所述第一远程元件相关联的位置，并且其中所述服务器可用于通过以下步骤来从多个远程元件中识别所述远程元件：基于与识别出的所述远程元件相关联的位置来识别所述远程元件。

32. 如权利要求 25 所述的系统，其中所述服务器可用于通过以下步骤来从所述第一远程元件接收匹配数据：

从多个远程元件接收匹配数据；以及

汇聚从两个或更多个远程元件接收的所述匹配数据，并且其中所述服务器可用于基于汇聚的匹配数据来生成所述语法条目。

33. 如权利要求 32 所述的系统，其中所述服务器可用于通过以下步骤来汇聚所述匹配数据：

从所述多个远程元件中选择两个或更多个远程元件；以及

汇聚从所述两个或更多个远程元件接收的所述匹配数据。

34. 如权利要求 25 所述的系统，其中所述服务器可用于通过以下步骤来从第一远程元件接收匹配数据：

从所述第一远程元件接收第一组匹配数据；

从所述第一远程元件接收第二组匹配数据；以及

汇聚所述第一组匹配数据和所述第二组匹配数据；并且

其中所述服务器可用于基于汇聚的匹配数据来生成所述语法条目。

35. 如权利要求 25 所述的系统，其中所述服务器还可用于：

在所述第一远程元件处接收与用户所讲话语相关联的音频信息；

将接收到的所述音频信息与一个或多个语法条目相比较；

基于所述接收到的音频信息和所述第一存储音频数据的比较结果而生成所述匹配数据；以及

将所述匹配数据发送到语法服务器。

36. 一种用于分发语音识别语法的系统，包括：

用于从第一远程元件接收匹配数据的装置，所述匹配数据包括与所述远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息；

用于基于所述匹配数据生成语法条目的装置，其中所述语法条目包括第二存储音频数据和与所述第二存储音频数据相关联的单词标识符；以及

用于将所述语法条目发送到第二远程元件的装置。

## 用于分发语音识别语法的系统和方法

### 技术领域

本发明一般地涉及通信系统，更具体而言，涉及给予通信系统的命令的语音识别（speech-recognition）。

### 背景技术

自动语音识别（ASR）被用于补充或替换计算机键盘、电话键区、鼠标和相关设备以控制计算机应用。例如，ASR 的常见应用允许用户通过经由电话说出命令来控制远程服务器。用户可能能够单独通过语音或通过语音与电话键区的组合来请求来自远程数据库的信息和/或修改远程数据库。

这些系统的有效性由它们可以精确识别的不同单词（word）的数目来左右。由于与特定系统相关联的词汇表不断增大，因此识别难度也在增大。识别精确度还受说话者之间的发音差别所影响。识别出的单词与关于各种预期发音的信息的组合被称为 ASR “语法（grammar）”。

正确的人名尤其难以识别。可能姓名的数目是无限的，并且发音可能取决于姓名的起源、所讲语言和说话者的地区方言或母语（native-language）而明显不同。例如，“Rzegocki”（起源于波兰并且正确的发音为“sha-guts-ki”）对大多数美国英语说话者而言是难以理解的。其需要关于波兰语和波兰变音符号的音译规则的罕见知识。因此，诸如“are-zegockee”之类的错误发音很常见，从而使得 ASR 系统的效率大大降低。这个问题由于全球移民和商业而变得越发复杂。与基于美国英语的 ASR 系统相交互的日语母语或北印度语母语说话者使这种和其他“非母语”姓名的识别进一步复杂化。

不管这种复杂性如何，语音邮件的记录和传送以及通过口语命令的现代通信系统的操作的其他方面都大大简化了使用。结果，对于基于电话和其他适当形式的通信应用而言，非常希望能够识别正确的姓名。例如，寻



址语音邮件消息、传输呼叫、检索联络信息或请求约会都取决于所涉及的人的姓名的准确识别。

### 发明内容

根据本发明，与生成和维护语音识别语法相关联的缺点和问题被大大减少或消除。

根据本发明一个实施例，一种用于分发语音识别语法的方法包括从第一远程元件接收匹配数据。匹配数据包括与远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息。该方法还包括基于匹配数据生成语法条目。语法条目包括第二存储音频数据和与第二存储音频数据相关联的单词标识符。另外，该方法包括将语法条目发送到第二远程元件。

根据本发明另一实施例，一种用于分发语音识别语法的设备包括存储器和处理器。存储器能够存储处理器指令。处理器能够从第一远程元件接收匹配数据。匹配数据包括与远程元件将接收到的音频信息匹配到第一存储音频数据的尝试相关联的信息。处理器可用于基于匹配数据生成语法条目。语法条目包括第二存储音频数据和与第二存储音频数据相关联的单词标识符。处理器还可用于将语法条目发送到第二远程元件。

本发明的某些实施例的技术优点包括提供了一种有效的用于优化语音识别语法的不确定（out-of-box）识别率的技术。其他技术优点包括降低和/或消除对语法的在线调节的需求，并且提供随时间提高识别率的自适应识别系统。

对本领域技术人员而言，本发明的其他技术优点将从以下附图、描述以及权利要求中变得更加明显。此外，虽然以上列出了特定优点，但是各个实施例可能包括所列优点中的所有或某些或不包括这些优点。

### 附图说明

为了更全面地理解本发明及其优点，现在结合附图参考以下描述，在附图中：

图 1 示出根据本发明特定实施例的通信系统；  
图 2 示出可以在通信系统的特定实施例中使用的目录服务器；  
图 3 示出可以在通信系统的特定实施例中使用的语法服务器；以及  
图 4 是示出根据特定实施例的通信系统的示例操作的流程图。

### 具体实施方式

图 1 示出支持用户 70 之间的通信的通信系统 10 的特定实施例。如图所示，通信系统 10 包括通过公共交换电话网（PSTN）50 和/或因特网 52 耦合的多个局域网（LAN）80。每个 LAN 80 额外使得多个通信设备 60、目录服务器 20、呼叫管理器 30 和网关 40 彼此耦合并使它们耦合到 PSTN 50 和/或因特网 52。另外，通信系统 10 包括耦合到 PSTN 50 和/或因特网 52 的语法服务器 90。在特定实施例中，语法服务器 90 收集来自若干设备（例如目录服务器 20）的语音识别结果，并利用这些结果来创建针对特定部署情形调节（tune）的语法和/或为了用在特定设置中而优化现有语法。结果，通信系统 10 可以减小或消除对作为相关语法的部署过程的一部分的手工调节的需求。

通信设备 60 可以分别代表硬件、软件和/或编码逻辑的任意合适的组合以向用户提供通信服务。例如，通信设备 60 可以代表电话、运行电话软件的计算机、视频监视器、照相机或利用通信系统 10 支持语音、视频、文本或其他合适形式的数据的通信的任意其他的通信硬件、软件和/或编码逻辑。在所示实施例中，通信设备 60a、60b、60d 和 60e 代表因特网协议（IP）电话，而通信设备 60c 和 60f 代表具有电话功能的个人计算机（PC）。另外，通信设备 60g 和 60i 代表分别通过 PSTN 50 耦合到通信系统 10 的模拟电话和蜂窝电话。虽然图 1 示出特定数目和配置的通信设备 60，但是通信系统 10 可以设想这样的组件的任意数目或布置来支持媒体通信。

网关 40 可以分别代表执行翻译功能以辅助 LAN 80 与 PSTN 50 和/或因特网 52 之间的无缝通信的硬件和/或软件的任意合适的组合。这些功能可以包括传输格式之间的翻译和通信程序之间的翻译。另外，网关 40 还

可以在音频和视频 CODEC 之间翻译并且在 LAN 侧和 PSTN 侧两者执行呼叫建立和拆除。例如，网关 40a 可以将呼叫管理器 30a 与 PSTN 50 接口。结果，当利用 PSTN 50 通过通信设备 60g 向通信设备 60a 发出呼叫时，网关 40a 将与 PSTN 50 相关联的呼叫格式翻译到与呼叫管理器 30a 相关联的呼叫格式，例如 SIP 或另一合适的基于分组的通信协议。

呼叫管理器 30 包括可用于建立通信系统 10 的主叫组件和被叫组件之间的连接的硬件和/或软件的任意合适的组合，例如在通信设备 60 之间连接呼叫或在特定的通信设备 60 和目录服务器 20 之间连接呼叫的连接。在特定实施例中，呼叫管理器 30 包括语音 IP (VoIP) 服务器，该 VoIP 服务器在分组电话网络内执行信令和会话管理功能。呼叫管理器 30 的信令功能允许跨越网络边界运载呼叫信息。呼叫管理器 30 的会话管理功能提供控制在通信系统 10 中跨越多个网络的端对端呼叫的属性的能力。例如，在特定实施例中，呼叫管理器 30 可以支持 SIP，SIP 是一种用于 IP 系统中的多媒体会议的标准。在这样的实施例中，呼叫管理器 30 可以包括 SIP 代理服务器，该代理服务器处理呼叫处理、SIP 请求和 SIP 响应。虽然以下关于呼叫管理器 30 的描述是针对 SIP 展开的，但是应该理解，呼叫管理器 30 也可以支持其他 IP 电话协议，例如 H.323 协议。在特定实施例中，呼叫管理器 30 还可以为通信设备 60 和位于 LAN 80 中的其他设备提供多种电话服务，包括（但不局限于）呼叫转发、电话会议、网络监视和 CODEC 翻译。

LAN 80、PSTN 50 和因特网 52 利用任意适当的通信协议提供耦合到它们的组件之间的连接。在特定实施例中，LAN 80 分别代表连接特定实体（例如公司或大学）的通信设备 60 的以太网的全部或一部分。此外，在特定实施例中，PSTN 50 代表公共交换电话网的全部或一部分。因特网 52 提供 LAN 80 和/或直接耦合到因特网 52 的设备之间的基于分组的数据通信。因特网 52 还可以包括公共交换电话网的全部或一部分。虽然图 1 示出包括耦合通信系统 10 的特定组件的特定类型网络（具体而言，LAN 80、PSTN 50 和因特网 52）的通信系统 10 的特定实施例，但是通信系统 10 可以包括任意适当类型和数目的以任意适当方式配置的网络。

语法服务器 90 生成、维护和/或传播供通信系统 10 的各个元件使用的语法 96。在本说明书以及所附权利要求中所使用的语法 96 可以包括将所讲话语 (utterance) 62 与在操作中由相关组件使用的任意适当形式的信息相关联的任意合适的信息, 包括 (但不局限于) 姓名、命令和/或密码。在特定实施例中, 语法 96 将各种话语 62 与和这些话语 62 相关联的姓名的文本表示相关联。在所示实施例中, 语法服务器 90 在语法服务器 90 的存储器 94 中维护着多个将被一个或多个目录服务器 20 所使用的语法 96。虽然以下描述集中于通信系统 10 的这样一个实施例, 其中语法服务器 90 负责将初始语法 96 提供到通信系统 10 的适当元件, 从这些元件接收信息以更新语法 96, 并生成用于传送到这些元件中的任意多个或全部的修订语法 96, 但是, 在特定实施例中, 语法服务器 90 可能提供这些功能中的任意多个或全部, 和/或与语法 96 的生成、维护和传播相关的任意其他附加功能。语法服务器 90 的特定实施例的内容下面将参考图 3 被更详细地示出。

作为可以使用语法服务器 90 所生成和/或更新的语法 96 之一的特定类型设备的一个示例, 通信系统 10 的所示实施例包括多个目录服务器 20。目录服务器 20 基于用户 70 所讲的语音命令将关于电话号码和/或其他适当形式的通信地址的信息提供给用户 70。在特定实施例中, 目录服务器 20 可以代表生成、存储和播放针对用户 70 的语音邮件消息的语音邮件服务器的全部或一部分。目录服务器 20 的特定实施例的内容参考图 2 被更详细地示出。

每个目录服务器 20 在该目录服务器 20 的存储器 24 中存储与该目录服务器 20 通信的用户 70 所使用的一个或多个语法 96。如上所述, 在通信系统 10 工作期间, 目录服务器 20 使用语法 96 来识别用户所讲的姓名、命令和/或其它类型的单词。目录服务器 20 可以以任意适当的方式获得语法 96。在特定实施例中, 每个目录服务器 20 维护着针对该目录服务器 20 所选择和/或定制的一个或多个语法 96, 所述这—个或多个语法 96 是基于母语、种族划分 (ethnicity)、位置、职业类型和/或与该目录服务器 20 相关联的用户 70 的任意其他合适的特性来选择和/或定制的。另外, 在特

定实施例中，适当的语法 96 的初始版本可以在安装或设置期间被手工加载到目录服务器 20 中。在替换实施例中，语法 96 的初始版本可以由语法服务器 90 传输到目录服务器 20。但是，一般而言，目录服务器 20 可以以任意适当的方式接收它们各自的一个或多个语法 96 的初始版本。目录服务器 20 随后可以与语法服务器 90 交互以获得更新的语法 96，从而提供增强的语音命令特征的操作。

在操作中，目录服务器 20 向用户 70 提供与通信系统 10 的其他用户 70 相关联的电话号码或其他通信地址。作为提供这些服务的一部分，目录服务器 20 可以接受、识别和/或响应于用户 70 所讲的话语 62。例如，用户 70 可以说出话语 62 来指定用户 70 请求目录服务器 20 的操作（例如“查找”）；识别对目录服务器 20 的其他用户（例如消息接收者）；和/或向目录服务器 20 提供密码和其他认证信息。虽然以下描述集中在通信系统 10 的这样一些实施例，其中目录服务器 20 接收其他用户 70 的姓名作为话语 62，但是，用户 70 可以向目录服务器 20 传输任意适当的信息，作为话语 62。另外，由目录服务器 20 提供的地址查找服务被描述为可以受益于通信系统 10 的语法优化技术的语音识别操作的一个示例。然而，通信系统 10 的特定实施例可以在支持任意适当任务的基于语音的控制时利用所述技术，包括（但不局限于）寻址语音邮件消息、会议呼叫期间的主叫识别和/或任意其他适当的语音识别操作。

作为在通信系统 10 的特定实施例中支持的语音命令操作的特定示例，尝试与第二用户 70（这里假设是用户 70b）发起通信的第一用户 70（这里假设是用户 70a）可能通过与第一用户 70 所在的 LAN 80 相关联的呼叫管理器 30 被连接到特定的目录服务器 20。在用户与目录服务器 20a 交互期间，目录服务器 20a 可以提示用户 70a 说出用户 70b 的姓名。用户 70a 随后说出话语 62，该话语 62 被用户 70a 正在用来与目录服务器 20 通信的通信设备 60（这里假设是通信设备 60a）所接收。通信设备 60a 生成包括与用户 70a 所讲话语 62 相关联的音频数据的话语信息 64，并将话语信息 64 发送到目录服务器 20。在特定实施例中，通信设备 60a 基于话语 62 生成话语信息 64，并将话语信息 64 在一个或多个语音 IP（VoIP）分组

中发送到目录服务器 20a。

目录服务器 20a 随后可以尝试通过将接收到的话语信息 64 与存储在适当的语法 96 内的特定语法条目 98 中的音频信息匹配以确定所讲姓名，然后识别针对该姓名的通信地址 68，从而识别出接收者的电话号码、分机（*extension*）和/或另一合适形式的通信地址 68。在特定实施例中，目录服务器 20a 可以通过计算语法 96 中的一个或多个语法条目 98 的置信度量度（*confidence measure*）来识别用户 70a 所讲的姓名。该置信度量度代表话语信息 64 与存储在相关语法条目 98 中的音频数据匹配的概率。一般而言，目录服务器 20 可以使用任意适当的技术来计算该置信度量度。

目录服务器 20a 随后选择具有最高置信度量度的语法条目 98。在特定实施例中，目录服务器 20a 随后可以以基于目录服务器 20a 和/或通信设备 60 的配置和能力的任意适当的方式向用户 70a 呈现与所选语法条目 98 相关联的姓名，从而使得用户 70a 可以确认目录服务器 20a 是否已经将接收到的话语 62 成功地匹配到适当的姓名。例如，在特定实施例中，目录服务器 20a 可以将所选姓名传输到通信设备 60a，从而使得通信设备 60a 在其显示屏上显示所选姓名。用户 70a 随后可以例如通过按下通信设备 60a 上的指定按钮来确认匹配的成功或失败。可替换地，目录服务器 20a 可以尝试将话语信息 64 匹配到一个语法条目 98，然后独立地确定匹配的成功或失败。例如，目录服务器 20a 自身可以基于针对所选语法条目 98 计算出的置信度量度是否高于某个预定最小值来判断尝试的匹配是否成功。该匹配过程以下将参考图 2 来更详细描述。

目录服务器 20a 随后可以识别与匹配的语法条目 98 相关联的通信地址 68，并将识别出的通信地址 68 传输到通信设备 60a 和/或使用识别出的通信地址 68 来执行任意其他适当的动作。作为一个示例，目录服务器 20a 可以向通信设备 60a 发送指定匹配的通信地址 68 的 VoIP 分组，结果，用户 70a 可以被提供以标识出相关的通信地址 68 的音频信息。作为另一示例，目录服务器 20a 可以在用户 70a 指示匹配成功的情况下将通信设备 60a 自动连接到匹配的通信地址 68。一般而言，目录服务器 20a 可以使用匹配的通信地址 68 来完成任意适当的动作。

此外，如果目录服务器 20a 在将话语信息 64 匹配到语法条目 98 的过程中没有成功，目录服务器 20a 则可以执行任意合适的步骤来允许用户 70a 与期望被叫方通信，而不管目录服务器识别适当通信地址 68 的过程已经失败。例如，目录服务器 20a 可以提示用户 70a 使用与通信设备 60a 相关联的键盘输入用户 70b 的姓名，提供多种主叫者可以选择的替换方案，请求主叫者重复姓名，和/或采取任意适当的步骤来确定用户 70a 感兴趣的姓名。

在确定尝试的匹配是否成功之后，目录服务器 20a 可以生成和/或更新匹配数据 26，该匹配数据 26 描述了目录服务器尝试将话语信息 64 匹配到语法条目 98 的任意相关方面。作为一个示例，在特定实施例中，如果目录服务器 20a 无法将话语信息 64 匹配到目录服务器 20a 所维护的任意语法条目 98，目录服务器 20a 则可以生成这样的匹配数据 26，该匹配数据 26 包括目录服务器 20a 无法匹配的话语信息 64，如图 1 所示。这样的统计方式可以识别出相关语法 96 无法成功匹配的特定话语 62。作为另一示例，目录服务器 20a 可以生成这样的匹配数据 26，该匹配数据 26 指示出目录服务器 20a 已经接收到被成功匹配到特定语法条目 98 的话语 62 的次数与目录服务器 20a 已经接收到该目录服务器 20a 无法匹配到同一语法条目 98 的话语 62 的次数（如用户 70 在不成功的匹配之后提供的输入所指示的）。这样的统计方法可以指示没有被相关语法 96 的内容充分代表的某些单词。但是，一般而言，目录服务器 20a 可以生成具有任意合适形式的匹配数据 26，该匹配数据 26 提供关于语音邮件服务器尝试将话语 62 匹配到适当的语法条目 98 的任意适当的信息。

另外，匹配数据 26 可以包括标识生成匹配数据 26 的目录服务器 20 或以其他方式描述该目录服务器 20 的特性的信息。例如，匹配数据 26 可以包括数据上下文信息 28，该信息 28 描述了与匹配数据 26 相关联的目录服务器 20 和/或用户 70，如图 1 所示。数据上下文信息 28 可以指示与相关目录服务器 20 和/或使用该目录服务器 20 的用户 70 相关联的位置、母语、主要种族划分和/或任意其他合适的特性。

在生成匹配数据 26 之后，目录服务器 20a 将匹配数据 26 传输到语法

服务器 90。在特定实施例中，目录服务器 20a 可以在生成匹配数据 26 时将匹配数据 26 实时地传输到语法服务器 90。在替换实施例中，目录服务器 20a 可以在通信系统 10 的操作期间随时间汇聚匹配数据 26 并在适当的时刻将匹配数据 26 传输到语法服务器 90。此外，在特定实施例中，目录服务器 20a 将匹配数据 26 以超文本传输协议（HTTP）响应的形式传输到语法服务器 90。但是，一般而言，目录服务器 20a 可以收集匹配数据 26 并将匹配数据 26 以任意适当的方式传输到语法服务器 90。

语法服务器 90 从一个或多个目录服务器 20 接收匹配数据 26。语法服务器 90 随后可以以任意适当的方式比较、汇聚和/或以其他方式处理接收到的匹配数据 26，以生成指示对通信系统 10 中的目录服务器 20 所使用的一个或多个语法 96 的潜在改进的信息。在特定实施例中，这样的处理可以包括人类干预，例如识别潜在趋势或基于识别出的趋势确定将作出的适当修改。在替换实施例中，该处理可以是完全自动的。

作为来自多个目录服务器 20 的匹配数据 26 的汇聚、匹配数据 26 随时间的汇聚、来自处理类似特性的目录服务器 20 的匹配数据 26 的关联和/或对匹配数据 26 的任意其他适当的处理的结果，语法服务器 90 和/或语法服务器 90 的操作者可能能够识别出如下经常发生的情形：目录服务器 20 未能将话语 62 匹配到语法 96 中的适当的语法条目 98。基于该信息，语法服务器 90 随后可以生成附加语法条目 98，并将这些新的语法条目 98 添加到现有的语法 96 或生成包括这些新语法条目 98 的新语法 96。这些新语法条目 98 可以包括想要将经常接收的话语 62 以期望提高目录服务器 20 将话语 62 成功匹配语法条目 98 的能力的方式匹配到适当的单词的音频信息。语法服务器 90 随后可以将新的语法条目 98 添加到由语法服务器 90 维护的语法 96。

作为结果，语法服务器 90 可以适应语法 96 以改善对难以发音的姓名的识别。例如，语法服务器 90 可以基于匹配数据 26 确定用户 70 频繁错误地将姓名“Rzegocki”发音为“ar-ze-gaw-key”。因此，语法服务器 90 可以生成一个或多个语法条目 98，它们将来自话语“ar-ze-gaw-key”的音频数据与姓名“Rzegocki”相关联。作为结果，相关的语法 96 随后可以



包括将正确发音映射到“Rzegocki”的语法条目 98 和将错误发音也映射到“Rzegocki”的语法条目 98 两者。

作为另一示例，在特定实施例中，可以基于匹配数据 26 识别昵称（nickname）或目录服务器 20 频繁无法匹配到正确姓名的特定姓名的缩略（abbreviation）版本。作为结果，匹配数据 26 可能能够突出语法 96 中可能漏掉的公知昵称或缩略。例如，语法服务器 90 可能能够基于匹配数据 26 确定一个或多个语法 96 不包括将针对话语“Bill”的音频信息与姓名“William”相关联的任意语法条目 98。响应于这一点的确定，语法服务器 90 可以生成一个或多个语法条目 98，这些语法条目 98 将针对话语“Bill”的音频数据与包括“William”的姓名相关联。此外，在通信系统 10 的特定实施例中，除了突出公知的没有出现在语法 96 中的缩略和昵称之外，匹配数据 26 还可以教导操作者在对该操作者而言是外语的语言中利用的缩略和昵称或者在特定于区域的方言中使用的缩略和昵称。例如，通过汇聚和回顾来自服务于大量拉丁美洲种族用户的目录服务器 20 的匹配数据 26，操作者或目录服务器 20 本身可能能够确定“Nacho”是“Ignacio”的昵称。作为结果，目录服务器 20 可以生成将针对“Nacho”的发音的音频数据与包括“Ignacio”的姓名相关联的语法条目 98。

作为又一示例，在通信系统 10 的特定实施例中，目录服务器 20 可以维护针对一个或多个母语/所讲语言对的语法。例如，语法服务器 90 可以维护针对法国公司的日本部门定制的语法。语法服务器 90 可以利用这样的语法 96 识别出与说话者相关联的匹配数据 26，并且可以汇聚所有这样的匹配数据 26。基于该匹配数据 26，语法服务器 90 可以识别出在日语为母语的说话者之中常见的法语姓名（或其它类型的法语单词）的错误发音。语法服务器 90 随后可以生成将该错误发音与被错误发音的姓名相关联的语法条目 98。

如上所述，语法服务器 90 可以将新生成的语法条目 98 添加到已经存储在存储器 94 中的语法 96 和/或可以生成包括新语法条目 98 的新语法 96。在特定实施例中，语法服务器 90 维护多个语法 96，其中每个语法 96 想要用在通信系统 10 内的一个或多个目录服务器 20 中。语法服务器 90

可以基于与将要使用这些语法 96 的目录服务器 20 相关联的用户 70 的特性，将新的语法条目 98 添加到一个或多个特定语法 96。同样如上所述，在特定实施例中，匹配数据 26 可以包括数据上下文信息 28，该数据上下文信息 28 描述与匹配数据 26 相关联的目录服务器 20 和/或用户 70。数据上下文信息 28 可以指示与相关目录服务器 20 和/或使用该目录服务器 20 的用户 70 相关联的位置、母语、主要种族划分和/或任意其他合适的特性。作为结果，语法服务器 90 可以基于包括在匹配数据 26 中的数据上下文信息 28 来确定语法服务器 90 会将新创建的语法条目 98 添加到的一个或多个适当的语法 96。

例如，返回上述示例，定制的语法 96 可以是针对服务于法国公司的日本部门的目录服务器 20 的使用而设计的。作为与这些各个目录服务器 20 相关联的用户 70 的说话方式和所讲单词之间的预期相似性的结果，基于从由类似公司操作的目录服务器 20 接收到的匹配数据 26 而生成的新的语法条目 98 可以被添加到该定制的语法 96。因此，语法服务器 90 可以基于针对工作在类似环境中的目录服务器 20 接收的匹配数据 26 来优化该定制的语法 96，以用于尝试发音法语姓名的日语母语说话者和/或尝试发音日本姓名的法语母语说话者。

语法服务器 90 随后可以将新的或修改后的语法 96 发送到一个或多个目录服务器 20。在特定实施例中，语法服务器 90 可以基于任意适当的标准、考虑和/或因素来选择一个或多个目录服务器 20 来发送新的或修改后的语法 96。在特定实施例中，在从第一目录服务器 20 接收到匹配数据 26 之后，语法服务器 90 可以基于第一目录服务器 20 和所选目录服务器 20 之间的相似性来选择一个或多个目录服务器 20 来发送修改后的语法 96。此外，在特定实施例中，语法服务器 90 可以基于包括在相关匹配数据 26 中的数据上下文信息 28 来选择接收这些语法 96 的目录服务器 20。

作为一个示例，语法服务器 90 可以从位于特定地理区域（例如美国南部的州）内的第一目录服务器 20 接收匹配数据 26。作为响应，语法服务器 90 可以识别一般与位于该同一地理区域中的目录服务器 20 相关联的语法 96。在基于接收到的匹配数据 26 生成修改后的语法 96 之后，语法服

务器 90 可以将修改后的语法 96 发送到位于该同一地理区域中的其他目录服务器 20。

作为另一示例，语法服务器 90 可以从与一种或多种语言相关联的第一目录服务器 20 接收匹配数据 26，并且作为响应，可以识别出与这些相同语言相关联的语法 96。语法服务器 90 随后可以生成修改后的语法 96 并将修改后的语法 96 发送到与正讨论的相同的一种或多种语言相关联的一个或多个目录服务器 20。再次返回更早的示例，语法服务器 90 可以从服务于法国公司的日本部门的第一目录服务器 20 接收匹配数据 26。在基于接收到的匹配数据 26 生成修改后的语法 96 之后，语法服务器 90 可以将修改后的语法 96 发送到服务于讲法语的用户 70、讲日语的用户 70、希望发音日语姓名或单词的讲法语的用户 70 和/或希望发音法语姓名或单词的日语用户的其他目录服务器 20。

一般而言，语法服务器 90 可以基于任意适当的标准、考虑和/或因素来选择适当的目录服务器 20 来发送修改后的或新的语法 96。随后，语法服务器 90 可以识别出将对各个目录服务器 20 所利用的语法 96 作出的修改和/或添加。作为结果，语法服务器 90 可以基于相关目录服务器 20 服务的用户 70 和/或 LAN 80 的特性来调整对由特定目录服务器 20 存储的语法 96 的添加或修改。作为结果，语法服务器 90 可以使用匹配数据 26 来基于每个目录服务器 20 期望接收的话语 62 来优化该目录服务器 20 所维护的语法 96。

因此，概括地讲，在通信系统 10 的特定实施例中，语法服务器 90 可以利用匹配数据 26 来识别用户 70 所讲的目录服务器 20 可能难以匹配到适当的语法条目 98 的话语 62。语法服务器 90 随后可以修改一个或多个语法 96 来提高使用修改后的语法 96 的目录服务器 20 将正确地匹配相关话语 62 的可能性。作为结果，通信系统 10 的特定实施例可以优化目录服务器 20 所使用的语法 96 和/或提供其他操作优点。然而，通信系统 10 的特定实施例可以提供所述优点中的某些或全部或不提供任何所述优点。

图 2 示出目录服务器 20 的特定实施例的内容。如图 2 所示，目录服务器 20 包括处理器 210、存储器 220 和网络接口 230。如上所述，目录服

务器 20 从通信系统 10 的通信设备 60 接收话语信息 64，并尝试将接收到的话语信息 64 匹配到由目录服务器 20 存储的语法 96 中的语法条目 98。目录服务器 20 随后向相关的通信设备 60 传输与在目录 260 中的匹配的语法条目 98 相关联的通信地址 68。另外，目录服务器 20 可以与语法服务器 90 交互以优化目录服务器 20 所存储的语法 96。虽然图 2 示出可以在通信系统 10 的特定实施例中与语法服务器 90 结合使用的语音识别设备的一个特定示例，但是由语法服务器 90 支持的语法优化技术可被用来优化任意适当类型的语音识别行为的语法。

处理器 210 可用于执行与目录服务器 20 提供的服务相关联的指令。处理器 210 可以代表能够处理和/或传输电子信息的任意合适的设备。处理器 210 的示例包括（但不局限于）专用集成电路（ASIC）、现场可编程门阵列（FPGA）、数字信号处理器（DSP）和任意其他合适的特定或通用处理器。

存储器 220 存储处理器指令、语法 96、目录 260、匹配数据 26 和/或目录服务器 20 在操作中使用的任意其他适当的信息。存储器 220 可以代表适合于存储数据的易失性或非易失性、本地或远程设备的任意集合和布置。存储器 240 可以代表或包括随机访问存储器（RAM）设备、只读存储器（ROM）设备、磁存储设备、光存储设备或任意其他合适的数据存储组件。

网络接口 230 包括可用于辅助目录服务器 20 和语法服务器 90、呼叫管理器 30、通信设备 60 和/或通信系统 10 的其他组件之间的通信的接口。网络接口 230 可以包括适合于辅助目录服务器 20 和通信系统 10 的其他组件之间的交互的硬件和/或软件的任意适当的组合。在特定实施例中，网络接口 230 可以包括网络接口卡（NIC）和任意适当的控制逻辑，这种控制逻辑适合于利用以太网或令牌环协议来支持目录服务器 20 和相关联的 LAN 80 之间的通信。

语法 96 包括多个语法条目 98。每个语法条目将存储的一组特定的音频数据 112 与一个单词标识符 114 相关联。单词标识符 114 代表目录服务器 20 对其支持语音识别的姓名、命令、参数和/或任意其他适当的值。单

词标识符 114 可以包括目录服务器 20 可用的任意适当形式的信息。在特定实施例中，单词标识符 114 包括通信系统 10 的各个用户 70 的姓名的文本表示。存储的音频数据 112 可以代表任意适当形式的音频数据，其描述了与相关的单词标识符 114 相关联的话语 62。取决于通信系统 10 的配置，包括在每个语法条目 98 中的存储的音频数据 112 可以代表与相关的单词标识符 114 所标识的单词的标准发音、相关单词的错误发音、该单词的缩略的发音、相关单词（例如昵称）的发音和/或目录服务器 20 被配置为映射到该相关单词标识符 114 的任意其他适当的话语相关联的音频信息。

目录 260 包括多个目录条目 262。每个目录条目 262 将特定的单词标识符 114 与通信地址 68 相关联。通信地址 68 可以代表全电话号码、分机、电子邮件地址和/或任意其他适当形式的通信地址。虽然图 2 出于简化目的示出了目录 260 的一个实施例，其中每个目录条目 262 包括单个代表特定用户 70 的姓的单词标识符 114，但是目录条目 262 可以分别包括多个单词标识符 114，这些单词标识符 114 被以任意适当的方式级联或组合以形成相关用户的全名的文本表示。

在操作中，目录服务器 20 向操作通信设备 60 的用户 70 提供目录服务。作为提供该目录服务的一部分，目录服务器 20 接收与用户 70 所讲话语 62 相关联的音频数据。更具体而言，目录服务器 20 从通信设备 60 接收话语信息 64。在特定实施例中，目录服务器 20 利用因特网协议（IP）来与通信设备 60 通信，并且话语信息 64 可以包括一个或多个语音 IP（VoIP）分组的多个部分。但是，一般而言，目录服务器 20 和通信设备 60 可以利用任意适当的协议通信，并且话语信息 64 可以代表以任意合适的方式构建的音频数据。

在接收到话语信息 64 之后，目录服务器 20 尝试将话语信息 64 与存储的音频数据 112 匹配，音频数据 112 包括在存储在存储器 220 中的语法 96 内的一个或多个语法条目 98 中。出于该说明书和所附权利要求的目的，目录服务器 20 可以通过执行任意适当的动作、操作和/或计算来识别与接收到的话语信息 64 相对应的语法条目，从而尝试将话语信息 64 匹配

到存储的音频数据 112。在特定实施例中，话语信息 64 和存储的音频数据 112 包括指定针对相关音频信号中的各个频率分量的幅度的数据。目录服务器 20 可以将话语信息 64 中的一个或多个频率分量的幅度与一个或多个语法条目 98 的存储的音频数据 112 中的相同频率分量的幅度相比较。目录服务器 20 随后可以基于该比较结果计算每个语法条目 98 的置信度量度 264。如上所述，置信度量度 264 代表话语信息 64 匹配相关语法条目 98 的可能性。

目录服务器 20 随后可以至少部分基于针对特定语法条目 98 计算出的置信度量度 264 来选择该语法条目 98。在特定实施例中，目录服务器 20 仅仅基于置信度量度 264 来选择语法条目 98，因此，选择具有最高置信度量度 264 的语法条目 98。在所示示例中，目录服务器 20 基于语法条目 98c 与最高置信度量度 264 相关联这一事实来选择语法条目 98c。

但是，目录服务器 20 可以考虑任意其他适当的因素。例如，在特定实施例中，目录服务器 20 可以维护关于特定用户利用其来呼叫其他用户的频率的数据。在这样的实施例中，目录服务器 20 可以在选择语法条目 98 时额外考虑该频率使用（frequency-of-use）数据，从而在选择中赋予频率使用数据任意适当的权重。作为结果，在这样的实施例中，在某些情况下，基于第一语法条目 98 与用户 70 频繁呼叫的另一方的姓名相关联这一事实，可以选择第一语法条目 98，而不是具有更高置信度量度的第二语法条目 98。

另外，在特定实施例中，目录服务器 20 可以维护多个语法 96。作为结果，目录服务器 20 可以在尝试匹配之前选择适当的语法 96 来匹配话语信息 64。例如，目录服务器 20 可以基于用户 70 的简介（profile）来确定用户 70 是母语为西班牙语的说话者，并且可以基于这一事实来选择特定语法。一般而言，目录服务器 20 可以基于任意适当的标准、考虑和/或因素来选择适当的语法 96。

在特定实施例中，用户 70 随后可以向目录服务器 20 指示目录服务器 20 是否成功地将话语信息 64 匹配到适当的语法条目。例如，目录服务器 20 可以将来自所选语法条目 98 的单词标识符 114 发送到相关的通信设备

60。通信设备 60 随后可以将所选单词标识符 114 显示在通信设备 60 的显示屏上。用户 70 随后可以例如通过按下通信设备 60 的适当按钮来指示目录服务器 20 是否已经成功地将话语信息 64 匹配到适当的语法条目 98。

如果用户 70 指示目录服务器 20 成功地将话语信息 64 匹配到适当的语法条目 98，目录服务器 20 则前进以查找与包括在所选语法条目 98 中的单词标识符 114 相关联的通信地址 68。在所示实施例中，目录服务器 20 通过将相关单词标识符 114 匹配到目录 260 内的特定目录条目 262 中的单词标识符 114 来执行该查找。因此，在所示示例中，所选单词标识符 114 匹配目录条目 262c。

在识别出匹配的目录条目 262 之后，目录服务器 20 可以向相关的通信设备 60 传输包括在匹配的目录条目 262 中的通信地址 68，如图 2 所示。通信设备 60 随后可以将该通信地址 68 显示给用户 70，或者可以基于匹配的通信地址 68 发起与另一通信设备 60 的通信。可替换地，目录服务器 20 自身可以发起主叫通信设备 60 和另一与匹配的通信地址 68 相关联的通信设备 60 之间的通信，和/或基于匹配的通信地址 68 采取任意其他适当的动作。

如果用户 70 指示目录服务器 20 没有成功地将话语信息 64 匹配到适当的语法条目 98，目录服务器 20 则可以采取任意合适的步骤来识别适当的通信地址 68 以提供给用户 70。例如，目录服务器 20 可以向相关的通信设备 60 传输与具有高置信度量度 264 的语法条目 98 相关联的单词标识符 114。通信设备 60 可以将该单词标识符 114 显示给用户 70，并提示用户 70 指示该第二单词标识符 114 是否是针对用户 70 所讲话语 62 的适当匹配。目录服务器 20 可以重复该过程，直到用户 70 指示目录服务器 20 已经选择了适当的语法条目 98 为止。目录服务器 20 随后可以尝试将适当的单词标识符 114 匹配到目录条目 262，如上所述。可替换地，目录服务器 20 可以响应于用户 70 指示目录服务器 20 未能将话语信息 64 匹配到适当的语法条目 98 而中止匹配过程，而不为用户 70 识别任何通信地址 68。

此外，在将话语信息 64 匹配或尝试匹配到适当的语法条目 98 之后，目录服务器 20 可以生成匹配数据 26，其以任意适当的方式描述了目录服

务器将话语信息 64 匹配到适当的语法条目 98 的尝试。作为一个示例，目录服务器 20 可以针对每次不成功的尝试（将话语信息 64 匹配到适当的语法条目 98 的尝试）生成匹配数据 26，并且可以包括话语信息 64 和基于用户 70 的随后输入的被确定为针对该话语信息 64 的适当匹配的单词标识符 114。作为另一示例，目录服务器 20 可以针对所有尝试（成功的和不成功的）生成匹配数据 26，并且在匹配数据 26 中包括所选语法条目 98 的单词标识符和关于单词标识符 114 是否被正确选择的指示。但是，一般而言，匹配数据 26 可以包括描述目录服务器尝试将接收到的话语信息 64 匹配到适当的语法条目 98 的任意适当的信息，包括（但不局限于）话语信息 64；在尝试匹配话语信息 64 时正确和/或不正确地选择的单词标识符 114；数据上下文信息 28，其描述与相关的用户 70、目录服务器 20 或 LAN 80 相关联的母语、种族划分、位置和/或其他合适的特性，和/或可被用于优化语法 96 的内容的任意其他适当的信息。

目录服务器 20 随后可以通过网络接口 230 将匹配数据 26 发送到语法服务器 90。目录服务器 20 可以利用任意合适的通信技术以任意适当的形式将匹配数据 26 发送到语法服务器 90。例如，目录服务器 20 可以将匹配数据 26 作为 HTTP 响应（例如响应于语法服务器 90 发送的 HTTP 请求）或简单邮件传输协议（SMTP）消息发送到语法服务器 90。语法服务器对发送的匹配数据 26 的使用随后将参考图 3 来描述。在特定实施例中，目录服务器 20 可以将匹配数据 26 存储在存储器 220 中，以用于随后的使用、检索和/或传输。

图 3 示出语法服务器 90 的特定实施例。如图所示，语法服务器 90 包括处理器 310、存储器 320 和网络接口 330。如上所述，语法服务器 90 为目录服务器 20 或提供语音识别能力的通信系统 10 的其他组件提供和/或更新语法 96。在特定实施例中，语法服务器 90 可以包括用户接口组件（图 3 中未示出）以允许操作者手工控制匹配数据 26 的处理部分和/或语法条目 98 的创建。作为结果，在语法服务器 90 的操作中的任意或所有所述步骤可以由人类操作者或通信系统 10 的另一组件来发起、控制和/或终止。



处理器 310 可用于执行与目录服务器 20 提供的服务相关联的指令。处理器 310 可以代表能够处理和/或传输电子信息的任意合适的设备。处理器 310 的示例包括（但不局限于）专用集成电路（ASIC）、现场可编程门阵列（FPGA）、数字信号处理器（DSP）和任意其他合适的特定或通用处理器。

存储器 320 存储处理器指令、语法 96、匹配数据 26 和/或目录服务器 20 在操作中使用的任意其他适当的信息。存储器 320 可以代表适合于存储数据的易失性或非易失性、本地或远程设备的任意集合和布置。存储器 340 可以代表或包括随机访问存储器（RAM）设备、只读存储器（ROM）设备、磁存储设备、光存储设备或任意其他合适的数据存储组件。

网络接口 330 包括可用于辅助语法服务器 90 和目录服务器 20、呼叫管理器 30、通信设备 60 和/或通信系统 10 的其他组件之间的通信的接口。网络接口 330 可以包括适合于辅助语法服务器 90 和通信系统 10 的其他元件之间的交互的硬件和/或软件的任意适当的组合。在特定实施例中，网络接口 330 可以包括网络接口卡（NIC）和任意适当的控制逻辑，这种控制逻辑适合于利用以太网或令牌环协议来支持语法服务器 90 和相关联的 LAN 80 之间的通信。

在操作中，语法服务器 90 向目录服务器 20 和/或提供语音识别服务的通信系统 10 的其他组件提供语法 96。在特定实施例中，语法服务器 90 可以向每个目录服务器 20 发送初始语法 96。在替换实施例中，初始语法 96 可以例如通过在目录服务器 20 的初始安装期间从致密盘加载一个或多个语法 96 而本地提供，并且语法服务器 90 可以将更新的、修改的和/或新的语法 96 提供到目录服务器 20。

在接收到初始语法 96 之后，目录服务器 20 开始操作并从通信设备 60 接收话语信息 64，如上所述。目录服务器 20 尝试将接收到的话语信息 64 匹配到它们各自的初始语法 96 中的语法条目 98。另外，目录服务器 20 以某种适当的方式基于这种将话语信息 64 匹配到语法条目 98 的尝试来生成匹配数据 26。目录服务器 20 随后将匹配数据 26 发送到语法服务器 90。

语法服务器 90 通过网络接口 330 从一个或多个目录服务器 20 接收一组或多组匹配数据 26。语法服务器 90 可以以任意合适的方式汇聚、比较和/或以其他方式处理匹配数据 26，以根据语法服务器 90、目录服务器 20 和通信系统 10 的整体的配置和能力来确定对现有语法 96 作出的适当改变和/或确定将由语法服务器 90 生成的新语法 96 的内容。

语法服务器 90 随后可以基于处理后的匹配数据 26 来生成一个或多个新的语法条目 98（在图 3 中示为“新语法条目” 398a、398b 和 398c）。在各个实施例中，语法服务器 90 可以以任意适当的方式来利用匹配数据 26 判断是否生成新的语法条目 98 以及这些新的语法条目 98 应该具有什么内容。此外，语法服务器 90 可以在生成附加语法条目 98 之后选择一个或多个将添加附加语法条目 98 的现有语法 96 和/或选择将包括在将包含这些附加语法条目 98 的新语法 96 中的其他语法条目 98。在特定实施例中，语法服务器 90 可以与匹配数据 26 一起接收数据上下文信息 28，并且可以基于与匹配数据 26 一起接收的数据上下文信息 28 来执行任一判断。

作为一个示例，在通信系统 10 的特定实施例中，语法服务器 90 可以从一个或多个目录服务器 20 接收多组匹配数据 26，并且可以汇聚该匹配数据 26，如图 3 中的“汇聚的匹配数据 326a”所示。基于汇聚的匹配数据 326a，语法服务器 90 可以识别出对所有用户 70 和/或目录服务器 20 或其特定子集而言常见的错误发音。语法服务器 90 随后可以生成新的语法条目 98，该新的语法条目 98 将存储的针对该常见错误发音的音频数据 112 与希望得到的姓名的单词标识符 114 相关联。结果，语法服务器 90 可以适应语法 96 以改善难以发音的姓名的识别。例如，语法服务器 90 可以基于汇聚的匹配数据 326a 确定用户 70 频繁将姓名“Rzegocki”错误发音成“ar-ze-gaw-key”。因此，语法服务器 90 可以生成被示为“新语法条目 398a”的语法条目 98，其将存储的针对话语“ar-ze-gaw-key”的音频数据 112 与针对“Rzegocki”的单词标识符 114 相关联。语法服务器 90 随后可以识别出一个或多个适当的语法 96，并将新语法条目 398a 添加到识别出的一个或多个语法 96。例如，语法服务器 90 可以判决将新语法条目 398a 添加到语法服务器 90 存储的所有语法 96，添加到当前具有包含

“Rzegocki”的正确发音的语法条目 98 的所有语法 96，或添加到由语法服务器 90 维护的语法 96 的任意其他适当的子集。

作为另一示例，在特定实施例中，语法服务器 90 可以基于匹配数据 26 识别出与特定姓名相关联的昵称或目录服务器 20 频繁无法匹配到正确姓名的特定姓名的缩略版本。结果，匹配数据 26 可能能够突出语法 96 中可能漏掉的公知昵称或缩略。例如，语法服务器 90 可能能够基于匹配数据 26 确定一个或多个语法 96（例如语法 96b）不包括将存储的针对“Bill”的音频数据 112 与针对“William”的单词标识符 114 相关联的任意语法条目 98。响应于这一点的确定，语法服务器 90 可以生成一个或多个新的语法条目 98（图 3 中被示为新语法条目 398b），它们将存储的针对话语“Bill”的音频数据 112 与针对“William”的单词标识符 114 相关联。语法服务器 90 随后可以识别出适当的语法 96 并将新语法条目 398b 添加到识别出的语法 96（在此情况下是语法 96b）。

作为又一示例，在通信系统 10 的特定实施例中，目录服务器 20 可以维护针对一个或多个母语/所讲语言对的语法，例如上述针对尝试发音法语姓名的日语母语说话者的示例语法，在图 3 中被示为语法 96c。语法服务器 90 可以基于语法服务器 90 接收的各组匹配数据 26 中的数据上下文信息 28 来识别与这样的说话者相关联的匹配数据 26，并且可以汇聚所有这样的匹配数据 26，如图 3 中的汇聚的匹配数据 326c 所示。基于汇聚的匹配数据 326c，语法服务器 90 可以识别出在日语母语说话者中常见的法语姓名（或其它类型法语单词）的错误发音。语法服务器 90 随后可以生成语法条目 98（在图 3 中由新语法条目 398c 示出），其将该错误发音与针对被错误发音的姓名的单词标识符 114 相关联。另外，语法服务器 90 随后可以将新语法条目 398c 添加到语法 96c。

在将任意新语法条目 398 添加到适当的语法 96 和/或生成包括新的语法条目 98 的任意新语法 96 之后，语法服务器 90 可以识别一个或多个目录服务器 20 来接收新的或修改后的一个或多个语法 96。语法服务器 90 可以基于这些目录服务器 20 当前存储的语法 96、语言位置和/或与这些目录服务器 20 相关联的说话者和/或相关的目录服务器 20 的任意其他适当的

特性来识别适当的目录服务器 20 来接收新的或修改后的语法 96。例如，在特定实施例中，语法服务器 90 维护服务器表 440，其中，语法服务器 90 维护关于语法服务器 90 用来识别发送特定语法 96 的适当目录服务器 20 的目录服务器 20 的信息。此外，在这样的实施例中，语法服务器 90 还可以使用服务器表 440 作为与匹配数据 26 一起接收的数据上下文信息 28 的补充或替代。

在识别出适当的目录服务器 20 之后，语法服务器 90 将新的或修改后的语法 96 发送到识别出的目录服务器 20。语法服务器 90 可以利用任意合适的技术以任意适当的格式将相关语法 96 发送到识别出的目录服务器 20。例如，在特定实施例中，语法服务器 90 可以将相关语法 96 作为 HTTP 请求的一部分或作为 SMTP 邮件消息发送到目录服务器 20。可替换地，语法服务器 90 可以不发送完整的新的或修改后的语法 96，而仅仅发送新的语法条目 98 和/或新的或修改后的语法 96 的任意适当部分。如上所述，识别出的目录服务器 20 随后可以存储接收的语法 96 和语法条目 98，并利用它们开始匹配这些目录服务器 20 接收到的话语信息 64。

作为结果，在特定实施例中，语法服务器 90 可能能够收集来自若干目录服务器 20 或通信系统 10 中支持语音识别服务的其他设备的匹配数据 26，并优化这些设备所利用的语法 96。具体而言，语法服务器 90 可能能够利用特定语法 96 来识别目录服务器 20 中经常发生的错误发音。另外，语法服务器 90 可能能够识别与在特定语法 96 中列出的特定姓名相关联的流行昵称或缩略，并将存储的与这些昵称或缩略相关联的音频数据 112 添加到适当的语法 96。一般而言，语法服务器 90 还可能能够基于用户 70、位置、使用和/或相关目录服务器 20 的其他特性来针对目录服务器 20 开发专门的语法 96。因此，通信系统 10 可以提供若干优点。通信系统 10 的各个实施例可以展现这些优点中的某些或所有或不展现这些优点。

图 4 是示出根据特定实施例的语法服务器 90 的示例操作的流程图。具体而言，图 4 示出在基于接收自多个目录服务器 20 的匹配数据 26 修改语法 96 时语法服务器 90 的操作。虽然图 4 示出以特定顺序执行特定步骤的语法服务器 90 的实施例的操作，但是替换实施例可以省略这些步骤中

的任意步骤或包括任意适当的附加步骤，并且可以以任意适当的顺序来执行所示步骤。

操作开始于步骤 400，其中语法服务器 90 从多个目录服务器 20 接收多组匹配数据 26。在特定实施例中，每组匹配数据 26 包括相关目录服务器 20 没有成功匹配的话语信息 64、正在说话的用户 70 指示用户 70 希望查找的姓名的单词标识符 114，以及描述了发送该组匹配数据 26 的目录服务器 20 的数据上下文信息 28。在步骤 410 处，语法服务器 90 基于与汇聚的多组匹配数据 26 相关联的数据上下文信息 28 来汇聚两组或更多组匹配数据 26。

在步骤 420 处，语法服务器 90 基于接收到的匹配数据 26 生成新的语法条目 98。语法服务器 90 随后在步骤 430 处基于数据上下文信息 28 识别出将添加新的语法条目的一个或多个语法 96。在步骤 440 处，语法服务器 90 将新的语法条目 98 添加到识别出的语法 96。

语法服务器 90 随后在步骤 450 处识别修改后的语法 96 将发送到的一个或多个目录服务器 20。在步骤 460 处，将修改后的语法 96 发送到识别出的目录服务器 20。在特定实施例中，语法服务器 90 的操作随后可以结束。

虽然已经参考若干实施例描述了本发明，但是可以向本领域技术人员建议多种改变、变化、替代、变形和修改，并且希望本发明包含落在所附权利要求的范围内的这样的改变、变化、替代、变形和修改。

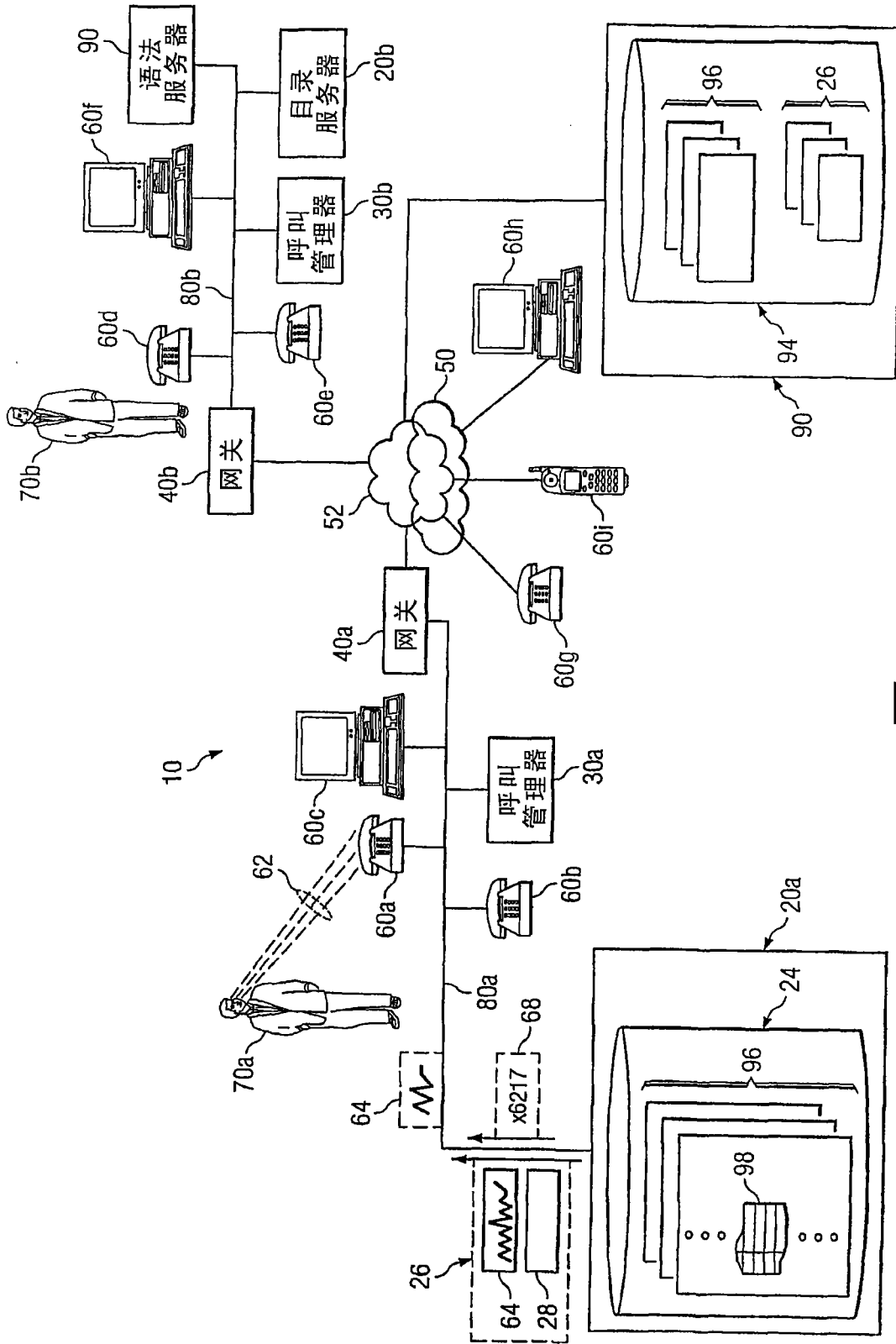


图1

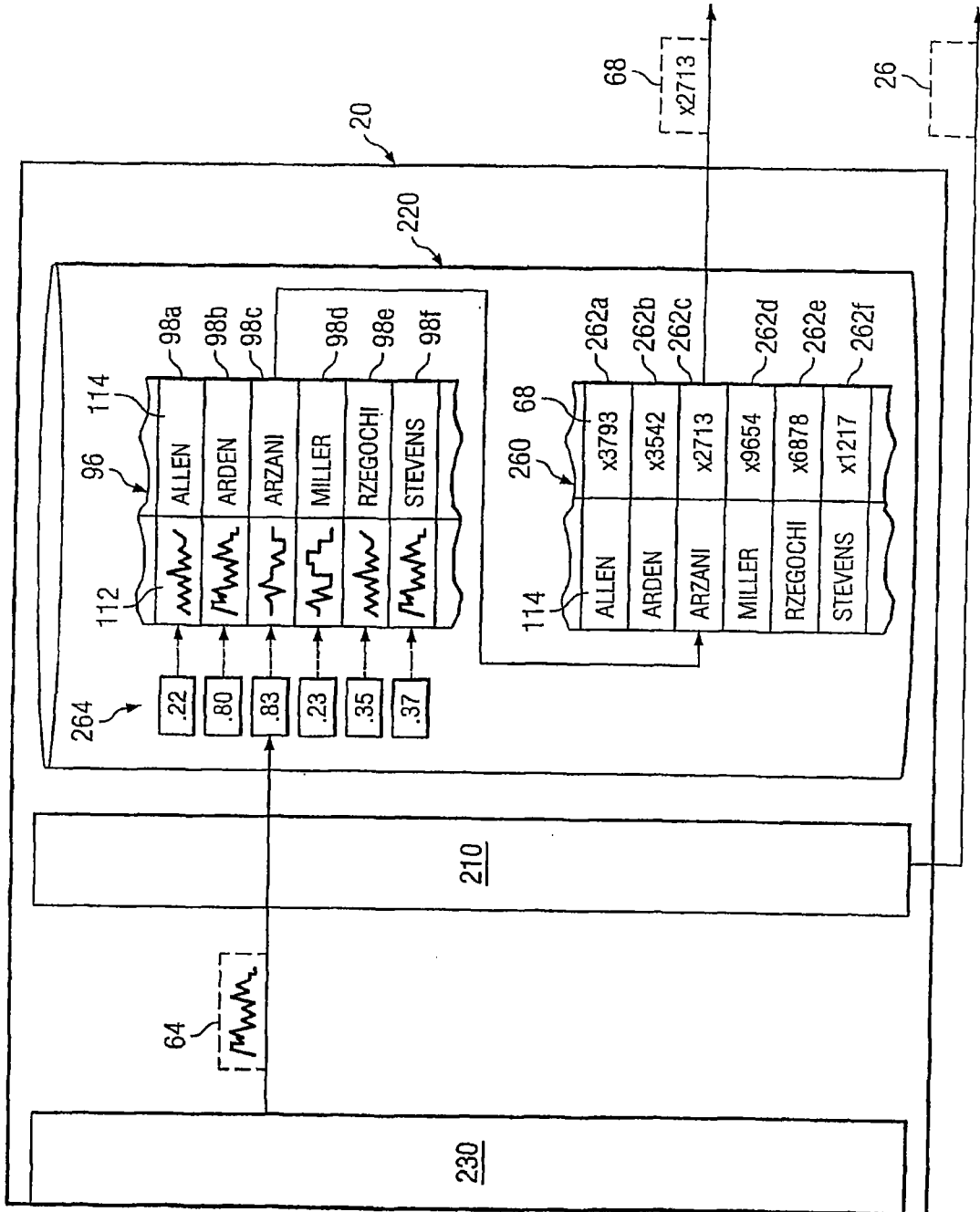


图2

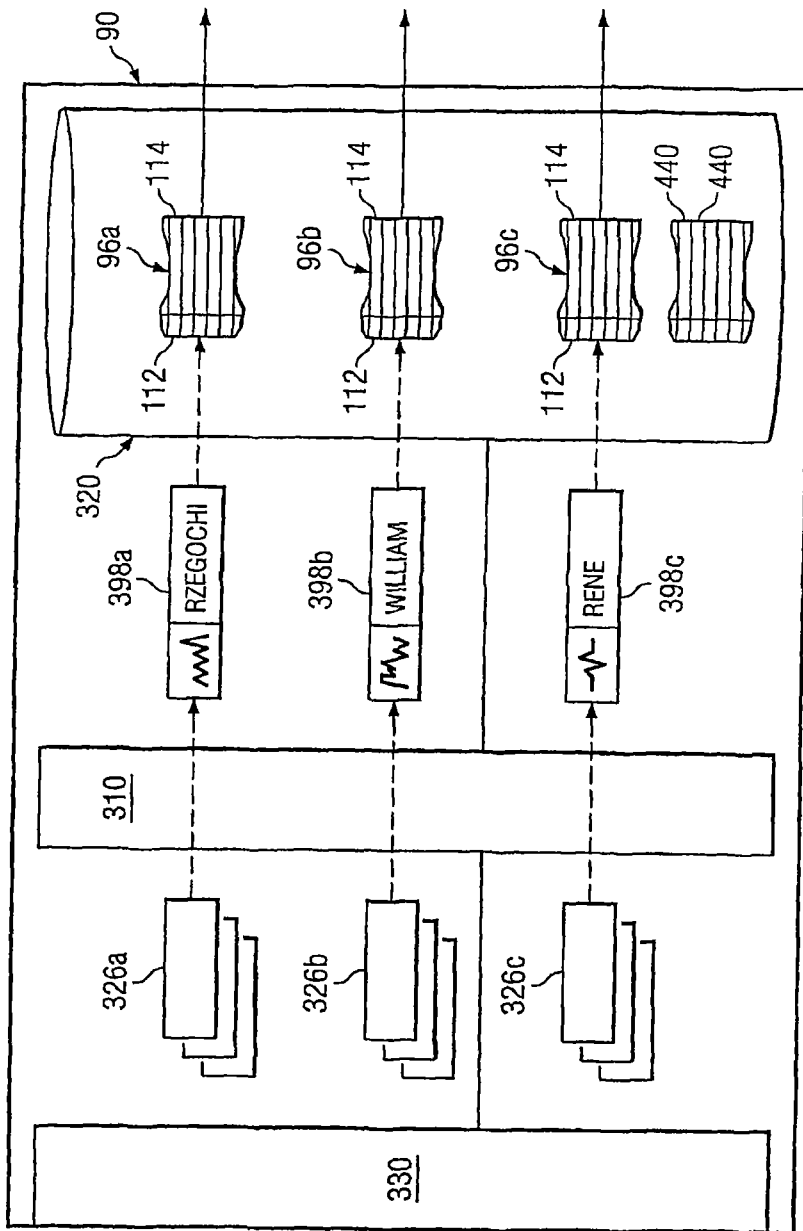


图3



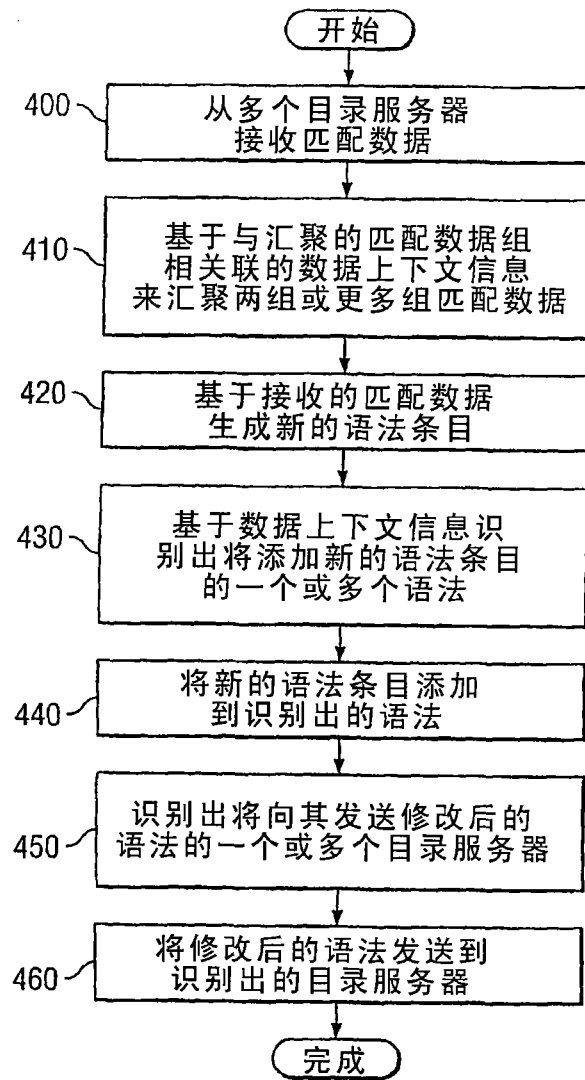


图4