



(12) 发明专利

(10) 授权公告号 CN 1835077 B

(45) 授权公告日 2011.05.11

(21) 申请号 200510054584.1

US 6311152 B1, 2001.10.30, 全文.

(22) 申请日 2005.03.14

CN 1352450 A, 2002.06.05, 全文.

(73) 专利权人 台达电子工业股份有限公司

CN 1464431 A, 2003.12.31, 全文.

地址 中国台湾桃园县

CN 1136356 A, 1996.11.20, 全文.

(72) 发明人 王瑞璋 蔡锦和 黄良声 沈家麟

审查员 隋欣

(74) 专利代理机构 隆天国际知识产权代理有限公司 72003

代理人 张龙哺 郑特强

(51) Int. Cl.

G10L 15/08(2006.01)

G10L 15/00(2006.01)

(56) 对比文件

JP 2839426 B2, 1998.12.16, 全文.

JP 2002-108389 A, 2002.04.10, 全文.

US 6272464 B1, 2001.08.07, 全文.

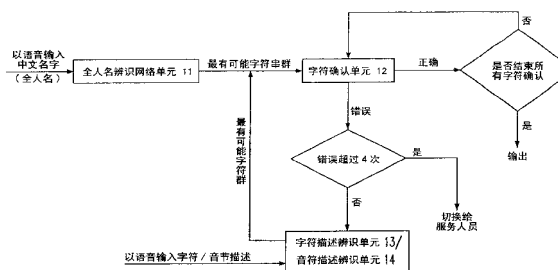
权利要求书 3 页 说明书 6 页 附图 3 页

(54) 发明名称

中文人名自动语音辨识输入方法及系统

(57) 摘要

本发明涉及一种中文人名自动语音辨识输入方法,该方法包含:(a)由一使用者输入一第一语音,以描述所欲辨识的一人名,该人名包含多个字符;(b)利用一全人名辨识网络单元,对该第一语音进行辨识,以得到一人名辨识结果;(c)传送该人名辨识结果至一字符确认单元;(d)利用该字符确认单元对该人名辨识结果的各该字符分别进行确认;(e)若于步骤(d)中确认各该字符辨识正确,则输出确认后的该人名辨识结果;(f)若于步骤(d)中确认各该字符其中之一辨识错误,则该使用者以一型态输入一第二语音描述该辨识错误字符;(g)利用对应该型态的一描述辨识单元对该第二语音进行辨识,且将其辨识结果传送至该字符确认单元等。



1. 一种中文人名自动语音辨识的输入方法,其步骤包含:
  - (a) 由一使用者输入一第一语音,以描述所欲辨识的一人名,该人名包含多个字符;
  - (b) 利用一全人名辨识网络单元,对该第一语音进行辨识,以得到一人名辨识结果;
  - (c) 传送该人名辨识结果至一字符确认单元;
  - (d) 利用该字符确认单元对该人名辨识结果的各该字符分别进行确认;
  - (e) 若于步骤 (d) 中确认各该字符辨识正确,则输出确认后的该人名辨识结果;
  - (f) 若于步骤 (d) 中确认各该字符其中之一辨识错误,则该使用者以一输入型态输入一第二语音描述该辨识错误字符;
  - (g) 利用对应该输入型态的一描述辨识单元对该第二语音进行辨识,且将对第二语音的辨识结果传送至该字符确认单元;以及
  - (h) 重复步骤 (d)-(g)。
2. 根据权利要求 1 所述的方法,其中步骤 (b) 还包含下列步骤:
  - (b1) 获取该第一语音的一特征参数;以及
  - (b2) 依据该特征参数,利用该全人名辨识网络单元对该第一语音进行辨识,其中:
    - 步骤 (b1) 还包含下列步骤:
      - (b11) 对该第一语音进行预处理;以及
      - (b12) 获取该第一语音的该特征参数,其中:
        - 步骤 (b11) 还包含下列步骤:
          - 放大该第一语音信号;
          - 对该第一语音信号正规化;
          - 对该第一语音信号进行预强调;
          - 将该第一语音乘上汉明窗;以及
          - 将该第一语音通过一低通滤波器或一高通滤波器,而其中:
            - 步骤 (b12) 还包含下列步骤:
              - 对该第一语音进行快速傅立叶变换处理;以及
              - 求取该第一语音的梅尔倒频谱参数。
3. 根据权利要求 1 所述的方法,其中步骤 (d) 还包含下列步骤:
  - 逐一输出相对于各该字符的多个字符假设;以及
  - 由该使用者从该多个字符假设中选择正确的各该字符,其中;
  - 该输出以语音播放各该字符假设的一描述词句,或以屏幕显示各该字符假设;及/或该使用者以语音输入选择或以按键输入选择。
4. 根据权利要求 1 所述的方法,其中步骤 (g) 还包含下列步骤:
  - (g1) 获取该第二语音的一特征参数;以及
  - (g2) 依据该特征参数,利用该描述辨识单元对该第二语音进行辨识,其中:
    - 步骤 (g1) 还包含下列步骤:
      - (g11) 对该第二语音进行预处理;以及
      - (g12) 获取该第二语音的该特征参数,其中步骤 (g11) 还包含下列步骤:
        - 放大该第二语音信号;
        - 对该第二语音信号正规化;

对该第二语音信号进行预强调；  
将该第二语音乘上汉明窗；以及  
将该第二语音通过一低通滤波器或一高通滤波器，而其中步骤 (g12) 还包含下列步骤：

对该第二语音进行快速傅立叶变换处理；以及  
求取该第二语音的梅尔倒频谱参数。

5. 根据权利要求 1 所述的方法，其中步骤 (f) 中，该使用者以一字符词组型态描述该辨识错误字符，其中对应该字符词组型态的描述辨识单元为一字符描述辨识单元。

6. 根据权利要求 1 所述的方法，其中步骤 (f) 中，该使用者以一音节词组方式描述该辨识错误字符，其中对应该音节词组型态的描述辨识单元为一音节描述辨识单元。

7. 一种中文人名自动语音辨识的输入系统，包含：

一语音输入装置，用以使一使用者输入一语音，该语音用以描述所欲辨识的一人名，该人名包含多个字符；

一全人名辨识网络单元，用以辨识该语音而得一人名辨识结果；

一字符确认单元，用以分别确认该人名辨识结果的各该字符是否正确；

一字符描述辨识单元，当该使用者以一字符词组型态描述各该字符时，用以辨识各该字符；

一音节描述辨识单元，当该使用者以一音节词组型态描述各该字符时，用以辨识各该字符；以及

一输出装置，用以输出确认后的该人名辨识结果。

8. 根据权利要求 7 所述的系统，其中：

该全人名辨识网络单元还包含一全人名辨识网络引擎以及一人名字符序列语言模型；

该人名字符序列语言模型为根据一基本词汇及一已知人名数据所训练出的语言模型。

9. 根据权利要求 8 所述的系统，其中：

该基本词汇由 408 音节所组成、由 1300 带声调的音节所组成或由 408 音节及 1300 带声调的音节所组成。

10. 根据权利要求 7 所述的系统，其中：

该字符描述辨识单元还包含一字符描述辨识引擎以及一字符描述语言模型；

该字符描述语言模型为根据一基本词汇及一运用字符描述的词组数据所训练出的语言模型。

11. 根据权利要求 10 所述的系统，其中：

该基本词汇由 408 音节所组成、由 1300 带声调的音节所组成或由 408 音节及 1300 带声调的音节所组成。

12. 根据权利要求 7 所述的系统，其中：

该音节描述辨识单元还包含一音节描述辨识引擎、一音节描述语言模型以及一音节对应字符表；

该音节描述语言模型为根据一基本词汇及一运用音节描述的词组数据所训练出的语言模型。

13. 根据权利要求 12 所述的系统,其中:

该基本词汇由 408 音节所组成、由 1300 带声调的音节所组成或由 408 音节及 1300 带声调的音节所组成。

14. 一种全人名辨识网络单元,用于一自动语音辨识输入系统,当一使用者输入用以描述一中文人名的一语音时,该全人名辨识网络单元便可辨识该人名,该全人名辨识网络单元包含:

一人名辨识网络引擎;以及

一人名字符序列语言模型,

其中该人名字符序列语言模型为根据一基本词汇及一已知人名数据所训练出的语言模型,

其中该基本词汇由一音节所组成,

其中该人名辨识网络引擎参考该人名字符序列语言模型以辨识该语音。

15. 一种字符描述辨识单元,用于一自动语音辨识输入系统,当一使用者输入以一字符词组形态描述一字符的一语音时,该字符描述辨识单元便可辨识该字符,包含:

一字符描述辨识引擎;以及

一字符描述语言模型,

其中该字符描述语言模型为根据一基本词汇及一运用字符描述的词组数据所训练出的语言模型,

其中该基本词汇由一音节所组成,

其中该字符描述辨识引擎参考该字符描述语言模型以辨识该语音。

16. 一种音节描述辨识单元,用于一自动语音辨识输入系统,当一使用者输入以一音节词组形态描述一字符的一语音时,该音节描述辨识单元便可辨识该字符,包含:

一音节描述辨识引擎;

一音节描述语言模型;以及

一音节对应字符表,

其中该音节描述语言模型为根据一基本词汇及一运用音节描述的词组数据所训练出的语言模型,

其中该基本词汇由一音节所组成,

其中该音节描述辨识引擎参考该音节描述语言模型以辨识该语音。

## 中文人名自动语音辨识输入方法及系统

### 技术领域

[0001] 本发明涉及一种自动语音辨识输入方法及系统,尤其涉及一种中文人名自动语音辨识输入方法及系统。

### 背景技术

[0002] 随着自动语音辨识技术日渐成熟,提供使用自动语音辨识技术的自动化程序,用以解决惯常繁琐的例行性工作,便可以达到节省巨额人力支出成本的效果。

[0003] 举例来说,目前全国性的查号台(例如:104和105),是使用人工辨识的方式进行人名查询的辨识工作,并无提供自动语音辨识的输入法,而提供104查号台服务的中国电信,则必须使用上千名人力才能应付庞大的查询需求,若能将其服务功能自动化,则不管是节省人力成本方面,甚或是对人力资源做更妥善的运用安排,都能收到显著的成效。

[0004] 现行的中文人名自动语音辨识系统,所采取的方式皆以一个个中文人名作为词组单位训练语言模型,当使用者以语音输入进行人名辨识时,辨识引擎会参考使用中文人名所训练出的语言模型进行比对,待辨识完成后系统即可输出完整的中文人名。然而,此种中文人名自动语音辨识系统却只适用于小量(约几千个人名数据)的中文人名辨识上,一旦遇上大量的人名数据(几万个甚至到几百万个人名数据),辨识成功率就会急遽下降,因此现行的中文人名自动语音辨识系统仅适用于一般公司行号的总机系统,若要应用在全国性的查号系统上,由于使用者的耐心有限,恐怕无法忍受辨识率极低的系统,因此在施行上仍有一定的困难度。

[0005] 因此,申请人鉴于公知技术的缺陷,乃经悉心试验与研究,终于研发出一种中文人名自动语音辨识输入方法及系统。

### 发明内容

[0006] 本发明的主要构想为提供一种中文人名自动语音辨识输入方法,其步骤包含:(a)由一使用者输入一第一语音,以描述所欲辨识的一人名,该人名包含多个字符;(b)利用一全人名辨识网络单元(Name Net Recognizer),对该第一语音进行辨识,以得到一人名辨识结果;(c)传送该人名辨识结果至一字符确认单元(Character Confirmation Unit);(d)利用该字符确认单元对该人名辨识结果的各该字符分别进行确认;(e)若于步骤(d)中确认各该字符辨识正确,则输出确认后的该人名辨识结果;(f)若于步骤(d)中确认各该字符其中之一辨识错误,则该使用者以一型态输入一第二语音描述该辨识错误字符;(g)利用对应该型态的一描述辨识单元对该第二语音进行辨识,且将其辨识结果传送至该字符确认单元;以及(h)重复步骤(d)-(g)。

[0007] 根据上述构想,其中步骤(b)还包含下列步骤:(b1)获取该第一语音的一特征参数;以及(b2)依据该特征参数,利用该全人名辨识网络单元对该第一语音进行辨识。

[0008] 根据上述构想,其中步骤(b1)还包含下列步骤:(b11)对该第一语音进行预处理;以及(b12)获取该第一语音的该特征参数。

[0009] 根据上述构想,其中步骤 (b11) 还包含下列步骤:放大该第一语音信号;对该第一语音信号正规化 (normalization);对该第一语音信号进行预强调 (pre-emphasis);将该第一语音乘上汉明窗 (Hamming Window);以及将该第一语音通过一低通滤波器或一高通滤波器。

[0010] 根据上述构想,其中步骤 (b12) 还包含下列步骤:对该第一语音进行快速傅立叶变换 (Fast Fourier Transform, FFT) 处理;以及求取该第一语音的梅尔倒频谱参数 (Mel-Frequency Cepstrum Coefficients, MFCC)。

[0011] 根据上述构想,其中步骤 (d) 还包含下列步骤:逐一输出相对于各该字符的多个字符假设;以及由该使用者从该多个字符假设中选择正确的各该字符。

[0012] 根据上述构想,其中该输出是以语音播放各该字符假设的一描述词句。

[0013] 根据上述构想,其中该输出是以屏幕显示各该字符假设。

[0014] 根据上述构想,其中该使用者是以语音输入选择。

[0015] 根据上述构想,其中该使用者是以按键输入选择。

[0016] 根据上述构想,其中步骤 (g) 还包含下列步骤:(g1) 获取该第二语音的一特征参数;以及 (g2) 依据该特征参数,利用该描述辨识单元对该第二语音进行辨识。

[0017] 根据上述构想,其中步骤 (g1) 还包含下列步骤:(g11) 对该第二语音进行预处理;以及 (g12) 获取该第二语音的该特征参数。

[0018] 根据上述构想,其中步骤 (g11) 还包含下列步骤:放大该第二语音信号;对该第二语音信号正规化 (normalization);对该第二语音信号进行预强调 (pre-emphasis);将该第二语音乘上汉明窗 (Hamming Window);以及将该第二语音通过一低通滤波器或一高通滤波器。

[0019] 根据上述构想,其中步骤 (g12) 还包含下列步骤:对该第二语音进行快速傅立叶变换 (Fast Fourier Transform, FFT) 处理;以及求取该第二语音的梅尔倒频谱参数 (Mel-Frequency Cepstrum Coefficients, MFCC)。

[0020] 根据上述构想,其中步骤 (f) 中,该使用者是以一字符词组型态描述该辨识错误字符。

[0021] 根据上述构想,其中对应该字符词组型态的描述辨识单元为一字符描述辨识单元 (Character Description Recognizer, CDR)。

[0022] 根据上述构想,其中步骤 (f) 中,该使用者是以一音节词组方式描述该辨识错误字符。

[0023] 根据上述构想,其中对应该音节词组型态的描述辨识单元为一音节描述辨识单元 (Syllable Spelling Recognizer, SSR)。

[0024] 本发明的另一构想在于提供一种中文人名自动语音辨识输入系统,其包含:一语音输入装置,其用以使一使用者输入一语音,该语音用以描述所欲辨识的一人名,该人名包含多个字符;一全人名辨识网络单元 (Name Net Recognizer),其用以辨识该语音而得一人名辨识结果;一字符确认单元 (Character Confirmation Unit),用以分别确认该人名辨识结果的各该字符是否正确;一字符描述辨识单元 (Character Description Recognizer, CDR),其当该使用者以一字符词组型态描述各该字符时,用以辨识各该字符;一音节描述辨识单元 (Syllable Spelling Recognizer, SSR),其当该使用者以一音节词组型态描述各该

字符时,用以辨识各该字符;以及一输出装置,用以输出确认后的该人名辨识结果。

[0025] 根据上述构想,其中该全人名辨识网络单元还包含一全人名辨识网络引擎以及一人名字符序列语言模型。

[0026] 根据上述构想,其中该人名字符序列语言模型为根据一基本词汇及一已知人名数据所训练出的语言模型。

[0027] 根据上述构想,其中该基本词汇由 408 音节所组成。

[0028] 根据上述构想,其中该基本词汇由 1300 带声调的音节所组成。

[0029] 根据上述构想,其中该基本词汇由 408 音节及 1300 带声调的音节所组成。

[0030] 根据上述构想,其中该字符描述辨识单元还包含一字符描述辨识引擎以及一字符描述语言模型。

[0031] 根据上述构想,其中该字符描述语言模型为根据一基本词汇及一运用字符描述的词组数据所训练出的语言模型。

[0032] 根据上述构想,其中该基本词汇由 408 音节所组成。

[0033] 根据上述构想,其中该基本词汇由 1300 带声调的音节所组成。

[0034] 根据上述构想,其中该基本词汇由 408 音节及 1300 带声调的音节所组成。

[0035] 根据上述构想,其中该音节描述辨识单元还包含一音节描述辨识引擎、一音节描述语言模型以及一音节对应字符表。

[0036] 根据上述构想,其中该音节描述语言模型为根据一基本词汇及一运用音节描述的词组数据所训练出的语言模型。

[0037] 根据上述构想,其中该基本词汇由 408 音节所组成。

[0038] 根据上述构想,其中该基本词汇由 1300 带声调的音节所组成。

[0039] 根据上述构想,其中该基本词汇由 408 音节及 1300 带声调的音节所组成。

[0040] 本发明的又一构想在于提供一种全人名辨识网络单元 (Name NetRecognizer), 其用于一自动语音辨识输入系统, 当一使用者输入一语音, 其用以描述一中文人名, 该全人名辨识网络单元便可辨识该人名, 其包含: 一人名辨识网络引擎以及一人名字符序列语言模型, 其中该人名字符序列语言模型为根据一基本词汇及一已知人名数据所训练出的语言模型, 其中该基本词汇由 408 音节或 1300 带声调的音节或 408 音节及 1300 带声调的音节所组成, 其中该人名辨识网络引擎参考该人名字符序列语言模型以辨识该语音。

[0041] 本发明的又一构想在于提供一种字符描述辨识单元 (CharacterDescription Recognizer, CDR), 其用于一自动语音辨识输入系统, 当一使用者输入一语音, 其以一字符词组形态描述一字符时, 该字符描述辨识单元便可辨识该字符, 其包含: 一字符描述辨识引擎以及一字符描述语言模型, 其中该字符描述语言模型为根据一基本词汇及一运用字符描述的词组数据所训练出的语言模型, 其中该基本词汇由 408 音节或 1300 带声调的音节或 408 音节及 1300 带声调的音节所组成, 其中该字符描述辨识引擎参考该字符描述语言模型以辨识该语音。

[0042] 本发明的又一构想在于提供一种音节描述辨识单元 (Syllable Spelling Recognizer, SSR), 其用于一自动语音辨识输入系统, 当一使用者输入一语音, 其以一音节词组形态描述一字符时, 该音节描述辨识单元便可辨识该字符, 其包含: 一音节描述辨识引擎, 一音节描述语言模型以及一音节对应字符表, 其中该音节描述语言模型为根

据一基本词汇及一运用音节描述的词组数据所训练出的语言模型,其中该基本词汇由 408 音节或 1300 带声调的音节或 408 音节及 1300 带声调的音节所组成,其中该音节描述辨识引擎参考该音节描述语言模型以辨识该语音。

[0043] 本发明的又一构想在于提供一种包含上述中文人名自动语音辨识输入系统的人名查询系统。

[0044] 根据上述构想,其中该人名查询系统用于查号台。

[0045] 根据上述构想,其中该人名查询系统用于自动总机对话系统。

[0046] 根据上述构想,其中该人名查询系统用于语音入口网站。

[0047] 本发明得借助下列附图与实施例的说明,得一更深入的了解。

### 附图说明

[0048] 图 1 所示为本发明的中文人名自动语音辨识输入方法的一流程图。

[0049] 图 2 所示为本发明的全人名辨识网络单元应用在中文人名自动语音辨识输入系统的示意图。

[0050] 图 3 所示为本发明的字符描述辨识单元应用在中文人名自动语音辨识输入系统的示意图。

[0051] 图 4 所示为本发明的音节描述辨识单元应用在中文人名自动语音辨识输入系统的示意图。

[0052] 其中,附图标记说明如下:

[0053] 1:中文人名自动语音辨识输入系统

[0054] 11:全人名辨识网络单元 (Name Net Recognizer)

[0055] 12:字符确认单元 (Character Confirmation Unit)

[0056] 13:字符描述辨识单元 (Character Description Recognizer, CDR)

[0057] 14:音节描述辨识单元 (Syllable Spelling Recognizer, SSR)

[0058] 具体实施方式

[0059] 本发明将可由以下的实施例说明而得到充分了解,使得本领域技术人员可以据以完成,然而本发明的实施并非可由下列实施例而被限制其实施形态。

[0060] 请参阅图 1,其为本发明的中文人名自动语音辨识输入方法的一流程图。该方法借助中文人名自动语音辨识输入系统 1 所执行,该中文人名自动语音辨识输入系统 1 包含语音输入装置(图上未示)、全人名辨识网络单元 (NameNet Recognizer) 11、字符确认单元 (Character Confirmation Unit) 12、字符描述辨识单元 (Character Description Recognizer, CDR) 13、音节描述辨识单元 (Syllable Spelling Recognizer, SSR) 14 以及输出装置(图上未示)。

[0061] 首先,使用者经由语音输入装置,以语音输入所欲辨识的中文人名。其后,将所输入的语音传送至全人名辨识网络单元 11,以进行人名辨识,全人名辨识网络单元 11 的辨识结果为中文人名的每个字符的可能的字符串群,经过切割后分成多个单一字符群,并分别传送给字符确认单元 12。字符确认单元 12 在接收到单一字符群之后,便针对每一字符群,由输出装置逐一输出可能的字符,让使用者进行分段确认并选择正确的字符,若每个字符都辨识正确无误,则由输出装置输出正确的结果,若其中有字符辨识失败,则系统 1 会引导



使用者做更进一步的辨识动作。

[0062] 假使有字符辨识失败,使用者必须针对辨识错误的字符,以另一种型态输入语音描述辨识错误的字符,然后传给对应该型态的描述辨识单元进行辨识。若使用者以字符词组型态重新描述辨识错误的字符时,便利用字符描述辨识单元 13 对该字符进行辨识,若使用者以音节词组型态重新描述辨识错误的字符时,则是利用音节描述辨识单元 14 对该字符进行辨识,不论采取何种型态描述该字符,所辨识的结果均是该字符可能的字符群,字符描述辨识单元 13 及音节描述辨识单元 14 随后会将该字符群传送至字符确认单元 12,由字符确认单元 12 辅助使用者作更进一步的确认。若有字符再度辨识失败,则继续重复描述该字符的步骤。然而,为了防止一再辨识错误会造成使用者感到不耐烦,若字符确认单元 12 在确认辨识失败超过一定次数(例如:4 次)后,则系统 1 便会将使用者切换给服务人员,由服务人员提供使用者所需的查询服务。

[0063] 在实施细节方面,该全人名辨识网络单元 11、字符描述辨识单元 13 以及音节描述辨识单元 14 皆必须先对所输入的语音进行特征参数获取的处理。该语音信号先经过适当步骤的预处理(例如:信号放大、正规化(normalization)、预强调(pre-emphasis)、乘上汉明窗(Hamming Window)、通过低通滤波器或高通滤波器等),接着便进入特征参数获取处理的步骤。特征参数获取处理以数据帧为单位,针对每一数据帧先进行例如:快速傅立叶变换(Fast Fourier Transform,FFT)处理将语音信号转为频谱,接着进一步对该频谱求得梅尔倒频谱参数(Mel-Frequency Cepstrum Coefficients,MFCC)。各辨识单元(11、13、14)求得语音信号的特征参数后,便会与各辨识单元(11、13、14)所具备的语言模型比对,以找出每一字符最有可能的字符群,并且依出现频率排序,再将结果传送至字符确认单元 12。

[0064] 请参阅图 2,其为本发明的全人名辨识网络单元应用在中文人名自动语音辨识输入系统的示意图。全人名辨识网络单元包含全人名辨识网络引擎以及人名字符序列语言模型,其中该人名字符序列语言模型为根据一基本词汇及一已知人名数据所训练出的语言模型,以供全人名辨识网络引擎参考比对用,其中该基本词汇可由 408 音节、或 1300 带声调的音节、或 408 音节及 1300 带声调的音节所组成。当语音输入后,经由上述特征参数获取处理,全人名辨识网络引擎便会参考人名字符序列语言模型,而得出各个字符最有可能的字符群。

[0065] 请参阅图 3,其为本发明的字符描述辨识单元应用在中文人名自动语音辨识输入系统的示意图。字符描述辨识单元包含一字符描述辨识引擎以及一字符描述语言模型,其中该字符描述语言模型为根据一基本词汇及一运用字符描述的词组数据所训练出的语言模型,以供字符描述辨识引擎参考比对用,其中该基本词汇可由 408 音节、或 1300 带声调的音节、或 408 音节及 1300 带声调的音节所组成。当以一字符描述型态的语音输入后,经由上述特征参数获取处理,字符描述辨识引擎便会参考字符描述语言模型,而得出该字符最有可能的字符群。

[0066] 请参阅图 4,其为本发明的音节描述辨识单元应用在中文人名自动语音辨识输入系统的示意图。音节描述辨识单元包含一音节描述辨识引擎以及一音节描述语言模型,其中该音节描述语言模型为根据一基本词汇及一运用音节描述的词组数据所训练出的语言模型,以供音节描述辨识引擎参考比对用,其中该基本词汇可由 408 音节、或 1300 带声调的音节、或 408 音节及 1300 带声调的音节所组成。当以一音节描述型态的语音输入后,经由

上述特征参数获取处理,音节描述辨识引擎便会参考音节描述语言模型,而辨识出该音节描述的可能结果,其后再参照音节对应字符表,而找出对应该音节的可能字符群。

[0067] 举例来说,若使用者今欲辨识“王小明”,则当使用者以语音输入“王小明”(Wang Xiao Ming)时,全人名辨识网络单元便会将人名切割成“王”、“小”、“明”分别辨识,随后将发音为“王”、“小”、“明”的最有可能字符传给字符确认单元,由字符确认单元辅助使用者进行确认工作,假设输出装置为屏幕,则屏幕会分别显示发音为“王”的最有可能字符,例如:1. 王 2. 亡,当使用者选择1之后,再显示发音为“小”的最有可能字符,例如:1 小 2. 晓 3. 筱,由使用者再进行确认,以此类推。最有可能字符的显示是以出现频率由高往低排序,以方便使用者能更快速的确认,而亦可每隔一段时间便依据查询的机率动态更新出现频率。若输出装置为语音的话(例如电话查询),则最有可能字符可用常用的词组型态描述,例如:“2. 春晓的晓”,使用者亦可通过电话按键进行确认。

[0068] 若全人名辨识网络单元辨识无误,使用者得以据此选择正确字符的话,则输出使用者确认的人名,以继续进行使用者所要求的服务(例如连结至数据库找出“王小明”的电话。若其中有字符辨识错误的话,则系统会引导使用者以其它方式重新描述该字符。例如:全人名辨识网络单元将“明”辨识成“民”,使用者无正确字符可选择,因此使用者可采用字符词组型态重新描述“明”,例如:“明天的明”,此时便由字符描述辨识单元进行辨识,并将最有可能字符群传送给字符确认单元,让使用者做进一步确认,使用者亦可采用音节词组型态重新描述“明”,例如“ㄇ一ㄥ明,二声明”(M-i-n-g),此时便由音节描述辨识单元进行辨识,并将最有可能字符群传送给字符确认单元,让使用者做进一步确认。倘若字符确认的动作超过四次,则服务人员便会接手,以人力提供使用者所需的服务。

[0069] 综上所述,本发明确实提供一种中文人名自动语音辨识输入方法及系统,通过本发明的中文人名自动语音辨识输入方法将中文人名“分割并征服”(divide and conquer)的方式,中文人名自动语音辨识输入系统将不再受限于一定数量的中文人名,而可以在维持辨识正确率的情形下,辨认无限多的中文人名,不但可以节省巨额的人力成本,亦可对人力资源做更妥善的安排。

[0070] 纵使本发明已由上述的实施例详细叙述而可由本领域技术人员任施匠思而作一些修饰,然而皆不脱离本发明所欲保护的范畴。

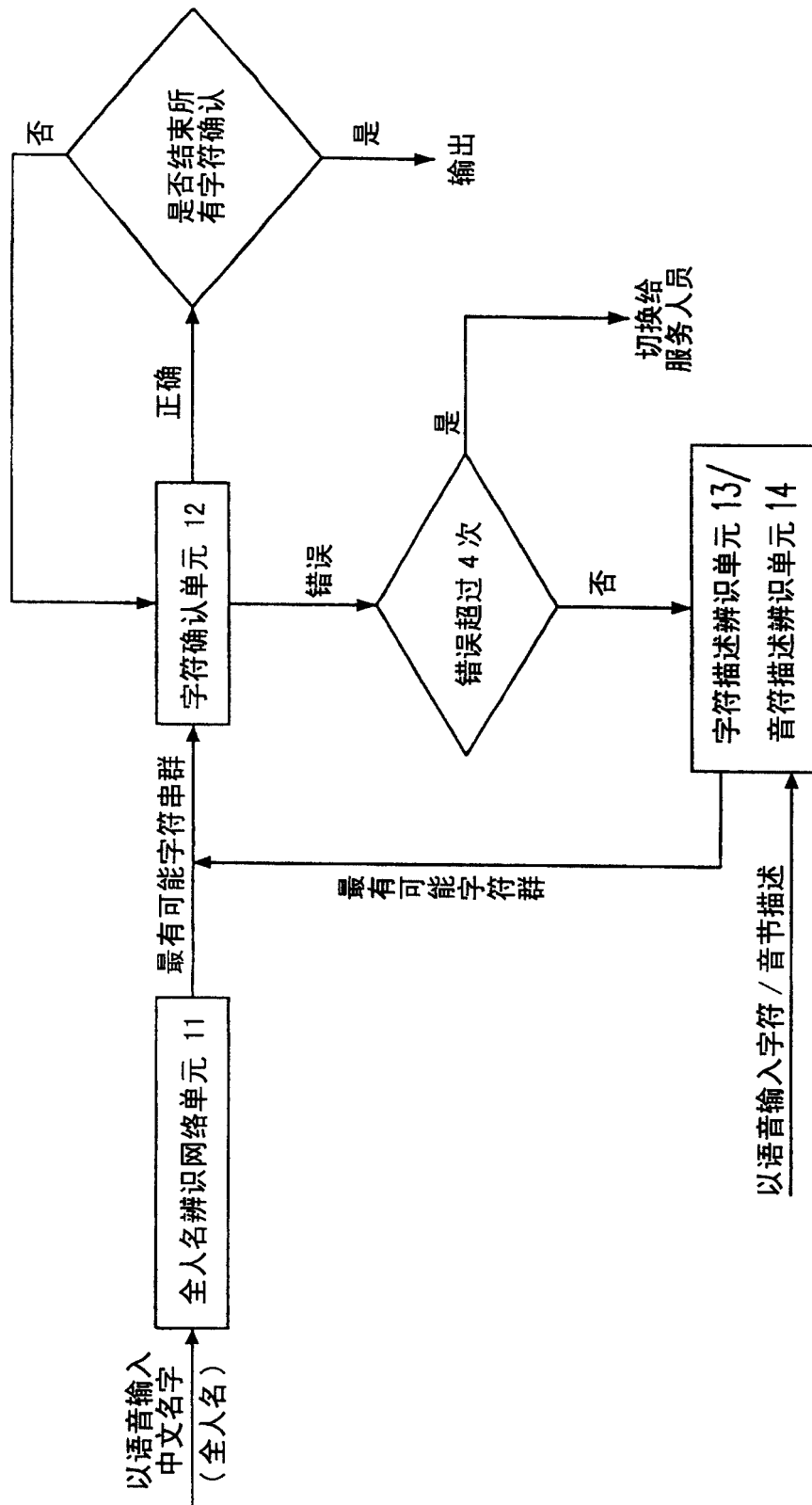


图 1

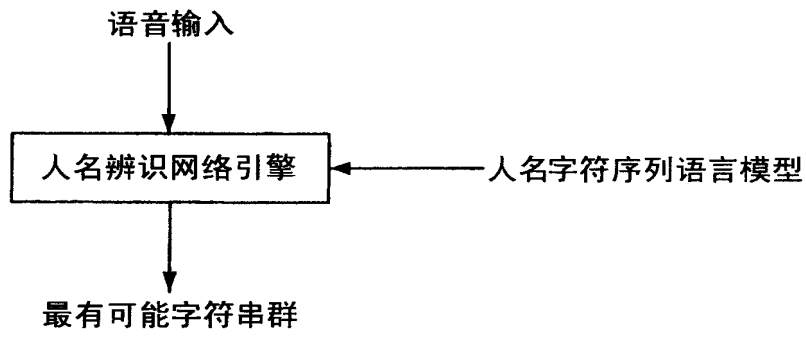


图 2

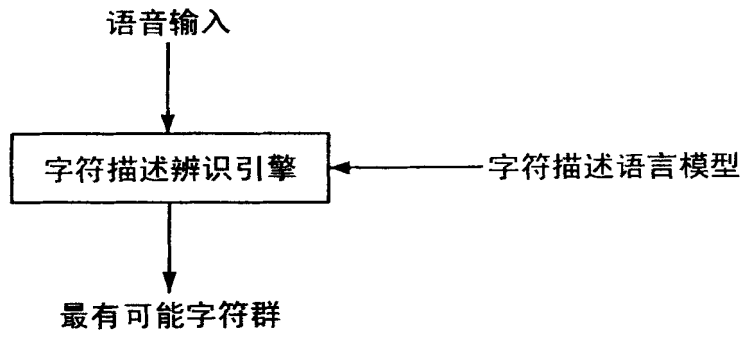


图 3

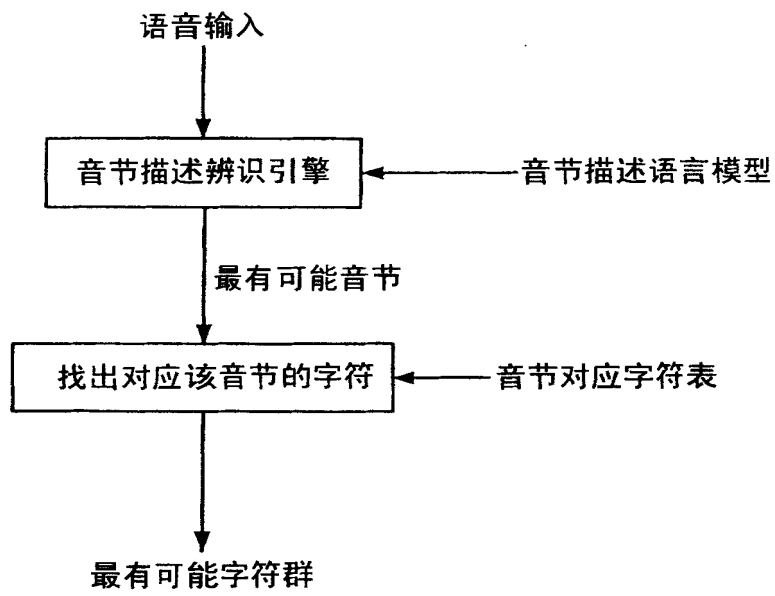


图 4