公告本

申請日期	85. 5. 15.
案號	85105749
類別	G06F3/18, 15/40, 15/62

A4
C4

307845

307845

（以上各欄由本局填註）

發明 ~~新型~~ 專利說明書

一、發明 ~~新型~~ 名稱	中文	用以當拼字或說話時改良辨識一大資料庫之字元的可靠度之方法和裝置
	英文	METHODS AND APPARATUS FOR IMPROVING THE RELIABILITY OF RECOGNIZING WORDS IN A LARGE DATABASE WHEN THE WORDS ARE SPELLED OR SPOKEN
二、發明 ~~創作~~ 人	姓名	查里斯　拉魯
	國籍	美國
	住、居所	美國加州拉卡納達市奧利微巷1432號
三、申請人	姓名（名稱）	美商音頻航行系統公司
	國籍	美國
	住、居所（事務所）	美國加州他茲那市維拿頓街3906號
	代表人姓名	

307845

（由本局填寫）

承辦人代碼：	
大　　類：	
ＩＰＣ分類：	

A6
B6

本案已向：

美　國（地區）　申請專利，申請日期：1995.5.9.　案號：08/437,057，□有　□無主張優先權

有關微生物已寄存於：　　，寄存日期：　　，寄存號碼：

（請先閱讀背面之注意事項再填寫本頁各欄）

裝　訂　線

五、發明說明（ 1 ）

相關申請案之總參考

本案係1992年12月31日申請之專利申請案第999062號之一部分，其係1991年3月27日申請之專利申請案第675632號之一部分，現在是美國專利第5274560號，其係1990年12月3日申請之專利申請案621577號之一部分，現在已捨棄。專利申請案第999062號揭露之內容與其中附加的電腦程式，在此併供參考。

發明背景

已大致知道藉由聽覺或說話的溝通功能可強化人機互動。已開發出各種介面，包括可辨識說話的輸入裝置，以及可產生合成語音的輸出裝置。雖然輸出裝置對於良好信號的處理已有大幅進步，但是輸入裝置仍有許多困難的問題。

這種輸入裝置必須將說話即字母，字，或片語轉成電的信號形式，接著必須處理電的信號以辨識說話。例如：以固定間距取樣組成話語的聽覺信號，接著將連續取樣值形成的類型與表示已知話語的儲存類型比較，而儲存類型表示的已知話語若極密切匹配取樣值的類型，即假設其係眞實的話語。理論上輸入裝置有極大的操作可靠度。但是以目前的精製技術而言，在執行程式時其需要極長的處理時間，因而不能在可接受的短時間內完成有效的結果。

一種已上市的語音辨識產品是美國麻州伍本的Lernout與Hauspie語音公司的產品，稱爲：CSR-1000演算法。該公司還銷售一種關鍵字元辨識演算法，產品稱爲：KWS-1000，

五、發明說明（ 2 ）

以及本文轉成語音的演算法，產品稱爲：TTS-1000。這些演算法可在傳統個人電腦上使用，而電腦必須至少具有高性能16位元的固定或浮點DSP處理器，與128KB的RAM記憶體。

CSR-1000提供含有數百個字的基本字彙，而字是以一序列音素儲存。取樣一句話語以產生一序列音素。程式設計師尚未找出處理這種音素序列以辨識說話的正確方式，但是可以相信完成它的方法是，將說話中找出的音素序列與各儲存字元的音素序列比較。此處理程序很費時，這是爲何此演算法僅使用含有數百個字的字彙的原因。

由此可知易於將CSR-1000演算法配成辨識個別說話字母。

而辨識含有30000個不同字或更多字的字彙則需要將話語分成數個音素或其他發音事件，或由使用者將字拼出來，在此情況下藉由其音素內容來辨識字母的發音，接著用字母序列來辨識未知字元。在任何大的字彙系統中，二種方法都需要以確保正確性。使用者會先嘗試以說出字元的方式來辨識該字元，若不成功則使用者接著可以拼字。

但是拼字時會發生一個問題，因爲元素辨識系統不易辨識英文字母。例如幾乎所有的辨識程式都很難辨識B與P，J與K，S與F等。事實上多數字母是由單音節組成，其與其他發音押韻、或者雖然未押韻，仍會與其他發音產生混淆。這種字母的例子是任何成對的字母F，X，S，M，其都具有相同的第一音。類似地，許多發音類似的音素也不能

五、發明說明（ 3 ）

辨識。以下若提及發音類似的字母或音素，即表示包括那些押韻者，以及那些會互相混淆者。很明顯，需要一種語音元素系統以處理因發音類似的字母或音素而產生的錯誤。

專利申請案第999062號(以下稱爲習用專利申請案)揭露一種方法，其使用程式化數位計算系統來辨識複數個發音中的任一個，各發音具有一聽得見的形式，由一序列語音元素表示，而各語音元素在序列中具有一個別位置，該方法包含：在電腦系統中儲存對應各複數個發音之數位表示，並指派一個別辨識指定給各發音；產生一個由複數個資料成的表，各資料與以下二者的唯一組合有關，即一特別語音元素，及發音的聽得見形式之語音元素序列中的特別位置，並且在各資料中儲存複數個發音中任一者之辨識指定，其聽得見的形式由包含特別語音元素的語音元素序列表示，此語音元素在與資料相關的該特別位置中；將人要說出及辨識的發音轉成一序列語音元素，各語音元素在序列中具有一個別位置；讀取至少各表資料，其與語音元素及位置之組合有關並對應以下之組合，即說話序列中的個別位置與說話序列中個別位置的特別語音元素；並決定那一辨識指定最常在資料中出現，而這些資料已在讀取步驟中取了。

爲了實施本方法，要辨識一組字元，即系統的字彙是儲存在第一資料庫中，並指派一辨識指定如辨識碼給各字元。

五、發明說明（ 4 ）

實際上，各說出的字是由一串音素或發音事件組成，其發生在一特定序列中。少數字元與一些字母如e與o可由單音素組成，並可由根據本發明之系統加以辨識。字母通常由一或2個音素組成，並可由根據本發明之系統加以辨識。

說出的語言是由數個已定義的不同音素組成，用一特定符號即可辨識各音素，一般認爲英文字包含50-60個不同音素，而在本發明的說明中指派一個別數值給各音素。

不同人對於相同語言中的同一字的發音也不同，所以一已知字元是由互相不同的音素串組成。爲了減少因元素變化而產生的辨識錯誤，因此選擇由不同人說出的話語，決定各說話者說出的字的音素串，以找出各字元部分的平均音素值，說出者由此可產生數個不同的音素。

在上述方法中，於個別字母或音素的翻譯或辨識中計算可能產生的錯誤，方式是給各字元高分，而字元在字母或音素序列的相同位置中，當成相同字母或音素翻譯或辨識，而發音與翻譯的相同者則給予低分。相對分數值的選擇是根據經驗法則或直覺法則，而且於系統已程式化並使用後，仍不改變。因此字元辨識的成功率是個人而異，即依使用者符合已程式化的翻譯器之元素特徵如發音、重音等的程度而定。

發明之概述

本發明之目標是提供先前專利申請案中所述的方法，以及實施此方法的系統之部分改良。

本發明之更特定目標是與先前專利申請案中所述的方法

五、發明說明（ 5 ）

相比能改良字元辨識的可靠度。

本發明之另一目標的允許字元辨識程序自動適應個別使用者。

根據本發明，上述以及其他目標的達成是藉由一種方法與裝置，其使用一程式化的數位資料處理系統來辨識複數個字元中任一者，各字元具有一聽得見的形式，其由一序列說話語音元素表示，而各語音元素在序列中具有一個別位置，數位資料處理系統接至裝置以接收字元的說話語音元素，並翻譯各收到的語音元素。

其中有複數個可能語音元素，各說話語音元素係一語音元素 α，各翻譯語音元素係一語音元素 β，並可將各說話語音元素 α 翻譯成複數個不同語音元素 β 中之任一者，一語音元素 β 與語音元素 α 相同，字元辨識的完成是藉由以下步驟：

將一個別之複數機率 $P_{\alpha\beta}$ 指派給各可能之語音元素，當已說出一語音元素 α 時，即可將語音元素翻譯成一語音元素 β；

儲存表示複數個字元中的所有字元的資料，各字元之資料包括在字元中辨識各語音元素，以及在表示字元之語音元素序列中辨識各語音元素之個別位置；

在接收與翻譯的裝置中，接收一序列由人說出之語音元素並表示一儲存拼字，且翻譯說話中之各語音元素，以及各語音元素在說話語音元素序列中之位置；及

將諸翻譯語音元素與表示複數個字元之各字元之儲存資

五、發明說明（ 6 ）

料比較，並執行一計算，使用與各翻譯語音元素β相關之機率$P_{\alpha\beta}$來辨識複數個字元中之字元，而字元之諸語音元素極密切對應諸翻譯語音元素。

附錄之參考

在此附加的附錄是本說明書的一部分，其中顯示用以實施本發明之程式例子。普通的程式設計師即可寫出此程式以處理發音類似但不押韻的字母。

附錄

押韻拼字檢查程式的演算法

產生資料庫

```
'_____
'產生表2
'_____
將名字表(表1)排序成字母群子集以產生名字ID表(表2)
***第8頁與第9頁***
```

拼字檢查演算法

```
'_____
'產生一個二維押韻表供以下的字母用
'將0放在各押韻群的最後
'_____
'字母A的押韻
RHYME(1,1)=1   'A與A押韻
RHYME(1,2)=8   'H與A押韻
RHYME(1,3)=10  'J與A押韻
```

五、發明說明（ 7 ）

```
RHYME(1,4)=11  'K與A押韻
RHYME(1,5)=0
RHYME(1,6)=0
...
RHYME(1,10)=0

'字母B的押韻
RHYME(2,1)=2  'B與B押韻
RHYME(2,2)=3  'C與B押韻
RHYME(2,3)=4  'D與B押韻
RHYME(2,4)=5  'E與B押韻
RHYME(2,5)=7  'G與B押韻
RHYME(2,6)=16  'P與B押韻
RHYME(2,7)=20  'T與B押韻
RHYME(2,8)=22  'V與B押韻
RHYME(2,9)=26  'Z與B押韻
RHYME(2,10)=0

'字母C的押韻
RHYME(3,1)=3  'C與C押韻
RHYME(3,2)=2  'B與C押韻
RHYME(3,3)=4  'D與C押韻
RHYME(3,4)=5  'E與C押韻
RHYME(3,5)=7  'G與C押韻
```

五、發明說明（ 8 ）

```
RHYME(3,6)=16  'P與C押韻
RHYME(3,7)=20  'T與C押韻
RHYME(3,8)=22  'V與C押韻
RHYME(3,9)=26  'Z與C押韻
RHYME(3,10)=0
...
'_____
'輸入字串
'_____
'從辨識的語音中讀取含有L個字母的字串(12頁第18行)
READ STRING (i) for i=1 to L
'_____
'處理各行
'_____
'在表2(14頁)的各行中
FOR n=1 to L
'將名字表(表2)中的行n讀入記憶體中(13頁)
'READ COL (i) = TABLE 2 (n,m) for m=1 to M
'在行的各邊的字母中(14頁第15行)
FOR j=n-1 to n+1
'從字串中擷取次一字母(14頁第15行)
LET = STRING (j)
'在群m的各押韻字母中(14頁第26行)
FOR r=1 to 10
```

五、發明說明（ 9 ）

```
'從押韻表中找到子集號碼(14頁第26行)
m = RHYME (k,r)
'若未到達押韻群的最後(14頁第26行)
IF m>0 THEN
'在押韻群子集的各名字ID中(14頁第26行)
FOR i=0 to群m中的號碼
'擷取名字號碼ID(15頁第33行)
ID=COL (m,i)
'若剛好匹配則設定WEIGHT=10，或者若押韻則設定
WEIGHT=6 (15頁第33行)
IF m=RHYME (k,l) THEN
   WEIGHT=10
ELSE
   WEIGHT=6
ENDIF
'將HIT計數設爲1(16頁第2行)
HITS=1
'若剛好等於行數則將HIT計數加1(16頁第4行)
IF k=n THEN
   HITS=HITS+1
ENDIF
'若是正確的字母序列則將HIT計數加1(15頁第7行)
IF (n-k)=此名字ID的先前的(n-k) THEN
   HITS=HITS+1
```

五、發明說明（ 10 ）

```
    ENDIF
    '計算每個字母的分數(16頁第14行)
    SCORE=WEIGHT*SCORE/L
    '將hit表加1，表4(16頁第19行)
    TABLE4 (ID) = TABLE4 (ID) + SCORE
    NEXT g
    ENDIF
    NEXT r
    NEXT k
    NEXT n
    '-------------------------------
    '搜尋TABLE4以找出3個最佳分數(16頁第28行)
    '-------------------------------
    FOR i=1 to TABLE4
       IF TABLE4(i)>BestScore(3)  THEN
          BestScore(3)=TABLE4(i, SCORE)
          BestID(3)=TABLE4(i, ID)
          IF BestScore(3)>BestScore(2)  then
          SWAP BestScore(3)，BestScore(2)
          SWAP BestID(3)，BestID(2)
    ENDIF
    IF BestScore(2)>BestScore(1)  then
       SWAP BestScore(2)，BestScore(1)
       SWAP BestID(2)，BestID(1)
```

五、發明說明（ 11 ）

ENDIF

ENDIF

NEXT i

附圖之簡單說明

圖1是實施本發明之系統的方塊圖。

圖2是圖1實施本發明的電腦2具體實例之方塊圖。

較佳具體實例之說明

該圖是繪示一系統之方塊圖，可用該系統實施本發明。系統的心臟是一傳統的通用電腦2如PC，其包含具有至少200KB容量的RAM記憶體4。電腦2通常配有鍵盤、螢幕與連接電腦2至週邊元件之裝置。

與電腦2有關的是儲存裝置6，其可設置在電腦2之中。儲存裝置6可以是硬碟，軟碟，ROM，PROM，EPROM，EEPROM，快閃記憶體，光碟等。可以用光碟或光碟機組成儲存裝置6，以形成另一音響系統的一部分，例如裝在汽車中。

連接電腦2的週邊裝置，包括一語音介面裝置10與語音合成器12。麥克風16可輸入元素介面裝置10，而語音合器12則輸出至喇叭18。若儲存裝置6由音響系統中的光碟機組成，則喇叭18可以由喇叭或該音響系統的喇叭組成。

儲存裝置6的儲存媒體包含辨識說話的作業系統，以及第一資料庫，其包含要辨識的說話表示與各儲存之發音的相關辨識指定，及第二資料庫，以複數個資料組成的表的形

五、發明說明（ 12 ）

式存在。第一資料庫中提供的辨識指示儲存在組成第二資料庫的表的適當資料中，其儲存方式將於以下詳述。

語音合成器12與喇叭18連接以產生並說出組成提示的話語，供系統使用者以及使用者說話的辨識方式用。

儲存媒體可以是光碟類型並能儲存發音，其方式可以直接再生，這可能是經由放大器與/或數位類比轉換器。在這些情況下可用這種元件來替代語音合成器12。

以下是系統的基本操作，操作開始時，必須位於記憶體4中的作業系統部分即從儲存裝置6中載入。載入記憶體4中的作業系統可包括該程式部分用以將說話轉成音素序列，其具有上述CSR-1000演算法的功能。接著麥克風16擷取系統使用者的說話並將其轉成電的類比信號以送至介面10。找出音素序列的方式是依使用的程式而定，介面10可將類比信號置於適當的電壓位準以送入電腦2，或是將類比信號轉成音素序列，並根據本發明以處理此序列以辨識對應說話的儲存發音。接著將一序列與儲存發音有關的音素送入語音合成器12，並由喇叭18以聽得見的形式發出以允許使用者辨識該說話是已正確辨識的字元。電腦可接著在作業系統的控制下再產生聽得見的發音以提示使用者再輸入包含某些資訊的說話，或是輸出從該資訊中導出的資訊，其先前由使用者以說話方式提供。

根據本發明之其他具體實例，說話是拼字的字母。在此情況下藉由將其音素或諸音素與儲存的類型比對，即可確定各字母，以處理組成語音元素序列的最後字母序列，以

五、發明說明（ 13 ）

辨識正確的儲存發音。

此外本發明之具體實例不必以聽得見或任何其他形式將儲存的發音再生。反而，儲存的發音可組成機器指令以對應個別說話，以及使得機器去執行一期望操作。

舉一個廣義的例子，本文所述的技術可用於上述專利申請案第675632號中揭露的那種導航系統，在此情況下會提示使用者以說話方式供應啓點與終點的辨識，接著再提供一串路由方向。若使用者提供的說話資訊是拼出啓點與終點位置的方式，則系統可提示使用者連續輸入各字母。

用一個廣義例子說明本發明，其中第一資料庫包含辨識字元的資料，並指派一辨識碼給各字元。與各字元的說話版本相關的標準或平均音素串是由傳統程序決定。字串的各音素係位於一個別位置n，並指派值m給各個不同的音素。

表1是第一資料庫的結構，其表示包含K個字元的大字彙資料庫，其中提供各字元一個別辨識碼(id#)與表示一字母與/或音素序列的資料，其可用以顯示字元與/或是將字元發音以增與/或找出與特定應用有關的字元的資訊。

表1

id#	字元字母序列	字元音素序列
1	.	. .
2	.	.
.	.	.
101	ALABAMA	a-l-a-b-a-m-a

五、發明說明（ 14 ）

102	ARIZONA	a-r-i-z-o-n-a
103	BRAZIL	b-r-a-z-i-l
104	CHICAGO	c-h-i-c-a-g-o
105	SEATTLE	s-e-a-t-t-l-e
106	ATLANTA	a-t-l-a-n-t-a
107	ARICONE	a-r-i-c-o-n-e
.	.	.
K		

接著在以下表2中作出第二資料庫，其顯示一大字彙資料庫，包含第一資料庫辨識碼的子集{n,m}，以處理在音素或字母字串的位置n，具有音素或字母m的字元。

表2

m ↓ \ n →	1	2	3	4	...	N
1	{1,1}	{2,1}	{3,1}	{4,1}	...	{N,1}
2	{1,2}	{2,2}	{3,2}	{4,2}	...	{N,2}
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.
M-1	{1,M-1}	{2,M-1}	{3,M-1}	{4,M-1}	...	{N,M-1}
M	{1,M}	{2,M}	{3,M}	{4,M}	...	{N,M}

五、發明說明（ 15 ）

第二資料庫的各資料是一子集(由{n,m}表示)，其包含第一資料庫中所有字元的辨識碼，位置n的音素或字母則包含音素或字母m。

在表2中，字串中音素或字母位置的總數是極大值N，而不同音素或字母的總數是M。選擇N的值以確保大致上能信任各字元的所有音素或字母。以下爲了簡便將以參考音素爲準。但是該了解的是參考字母也適用於一般規則。

系統又設置有一得分記憶體或表，其包含的位置數等於第一資料庫中的字元數；各位置與一個別字元的辨識碼有關。

分析一句說話字元以找出其特性音素串，該串將具有N個或N個以下的音素，而各音素可以是M個不同值中的任一者。

已辨識出字串(n=1)中第一位置的說話的音素值m，對於子集{1,m}的各元素，將一分數放在各得分記憶體位置，其與子集{1,m}中的辨識碼有關。接著辨識出字串(n=2)中第二位置的說話的音素值m，如上所述，對於子集{2,m}的各元素，將一分數放在各得分記憶體位置，其與子集{2,m}中的辨識碼有關。將此分數加入先前置於任一得分位置中的任何分數，此位置與子集{2,m}的辨識碼有關。

此程序會在說話的整個音素串中，或是在N個音素中持續，N可以不等於說話串中的音素數。已處理音素串之後，即查詢得分記憶體，而字元中，其辨識碼對應得分記憶體位置者，其具有最高分數，即確定爲該說話字元。

五、發明說明（16）

已知此程序會在極高的時間百分比下找出正確的字元。

接著從對應辨識碼的第一資料庫的位置中讀取儲存資料，如上所述可用儲存資料以說話的形式再生字元。

系統接著等使用者的說話回應。

爲了減少辨識錯誤率，根據本發明的系統可從得分表中選擇該發音的發音指示，其接收三個最高分數，並按照分數的高低送出這些指示。與各選擇的辨識指示有關的儲存發音即接著送至語音合成器12，再從喇叭18中以發音形式說出。已說出各合成發音後，系統即等待使用者的回應，如是或否。若各合成發音聽到使用者以否來回應，則可斷定辨識過程失敗並清除得分記憶體，若須要則提示使用者重覆該說話。可預知的是這失敗極少發生。

如上所述以表形式顯示的發音辨識指示儲存法，與習用方法比較大致上是改良了，因爲其大致上減少了必須處理的資料量，以便於辨識說話時到達。特別地，各於說話的語音元素序列中的各位置，不必擷取表的資料，其與該語音元素而以及該說話位置的特別語音元素有關。換言之，爲了正確辨識說話可不必擷取表中的所有資料。

在表2中以m表示的語音元素可以是字母或是音素。比較下，以n表示的語音元素位置可以分別是序列中的音素位置，或是字元中字母的位置。

上述的元素辨識程序，藉由僅讀取各語音元素位置(n)的第二資料庫表的一筆資料即可達成。但是已發現當各話語是字元時，其轉成一串音素，則喇叭可依此方式將該字元

五、發明說明（17）

發音以增減一或2個音素。若此眞的發生，則不會正確辨識該話語。根據本發明之另一特性，增加達成正確辨識機率的方法是，考慮與一特別音素有關的資料，其緊接在正確位置n相關的該筆資料。例如參考表2，若話語中位置n=3的音素與儲存資料比較，而此音素的值是2，則至少在子集{2,2}與{4,2}中的辨識碼會額外的用於包括得分記憶體中的分數。可以相同方式額外使用子集{1,2}與{5,2}。雖然開始時認爲這會減少達成正確辨識的機率，但驚訝的是已發現與此完全相反的事實，即此策略會增加正確辨識的機率。

此外要考慮以下事實，即翻譯各說話音素的裝置或程式會誤解發音類似的音素。爲了減少因這種個別音素的不正確辨識而產生的錯誤，所有的資料其與下列有關，即一已知的音素位置(n)，與音素押韻的音素(m)，其已被系統確定爲已說出，都要讀取，並將分數置於得分表中供已讀取的各辨識碼用。

若以拼字法輸入話語，則語音元素會是字母即字母的字元，而語音元素序列會是字母序列，其中各字母據有一特別位置。以下顯示此實施的簡化例子，其使用具有表lid#101-107的字元。

表3繪示第二資料庫中辨識碼的分配，因此若說話的第一字母是A，則僅需要查詢表3的子集{1,A}，以及剩餘的字母。要注意的是爲了簡便已假設各字元具有極大的7個字元。但是可建立第二資料庫以辨識字元，其具有任何選擇的

五、發明說明（ 18 ）

極大數目字母。

表3

m ↓ \ n→	1	2	3	4	5	6	7
A	101 102 106 107		101 103 105	106	101 104		101 102 106
B	103			101			
C	104			104 107			
D							
E		105					105 107
F							
G						104	
H		104					
I			102 104 107		103		
L		101	106			103 105	
M						101	
N					106	102 107	
O					102 107		104
R		102 103 107					
S	105						
T		106		105	105	106	
Z				102 103			

五、發明說明（19）

在表3的例子中再度假設僅定址話語的各字母的一子集{n,m}。但是已發現藉由應用2種策略即可增加正確辨識字元的機率，該字元是以拼字方式輸入。

首先其不會時常發生，即某人拼字時會省略或增加一個字母。在此情況下若僅讀取包含在第二資料庫中的一筆資料中的子集，其供要辨識的字元的各個字母用，則會減少正確辨識的機率。根據與此實施有關的第一策略，在要查詢的字元中的已知字母位置，也讀取正確資料兩旁的資料，並將分數置於得分記憶體中供各資料用。因此若考慮相信是話語的第四個字母，則對於字母A不僅要讀取{4,A}，而且要讀取子集{3,A}與{5,A}。

根據第二策略，要考慮以下事實，即翻譯各說話音素的裝置或程式會誤解發音類似的音素。例如A很容易與K混淆，而B會與C，D，E等混淆。爲了減少因這種個別字母的不正確辨識而產生的錯誤，所有的資料都與下列有關，即一已知的字母位置(n)，與字母押韻的音素(m)，其已被系統確定爲已說出，都要讀取，並將分數置於得分表中供已讀取的各辨識碼用。

雖然再一次，開始時認爲這些策略會減少正確辨識已用拼字法輸入的話語的機率，但驚訝的是已發現與此完全相反的事實。得分記憶體中顯示的諸錯誤翻譯的字母分數，會一直小於正確字元的分數。

將得分記憶體概念化爲具有如以下表4的結構，其中各字元分數的累加是根據該字元的id#於表2或3讀取的次數，此

五、發明說明（ 20 ）

讀取是根據上述的一種程序。

表4

id#	分數
1	字元1的累加分數
2	字元2的累加分數
.	.
.	.
.	.
K	字元K的累加分數

爲了得分目的，每次要在第二資料庫的資料中要找出一特別的id#，此組成該id#的成功得分。依使用的策略而定，即考慮押韻音素或字母，或在要辨識的話語中加減一個字母或一個或2個音素，可讀取數筆資料供要辨識的字元的各語音元素位置用。在此情況下可根據各種因素加權各次成功。

例如當擷取第二資料庫或讀取時即可實施以下的得分策略，以占要辨識的說話的各語音元素位置。

對於列n資料中的各id#其剛好匹配說話的相關字母或音素，即指派權重10；對於一列資料中的各id#，該列與一字母或音素有關，其與列m的資料押韻，則指派權重6。此適用於行n中的id#，其與特別的說話語音元素位置有關，而且在第二資料庫(n±1,2)的各其他行中當說話時即考慮讀取要增減的字母或音素。指派成功值1給每一個這種id#。

五、發明說明（ 21 ）

接著，對m列內每一正好對屬於說話語音元素位置的成功值而言，成功值 1 被加至前一個成功值中；

若列m中的id#剛好匹配說話的相關字母或音素(以及列中的同一id#剛好匹配說話的相關字母或音素)，這是剛好在說話的語音元素位置的前面，則將另一成功值1加入前一成功值。

接著將總成功值乘上權重再除以說話中的語音元素的數目，即字母或音素。此除法可確保較長的字元不會因爲其長度而得到較高的分數。

將最後的分數置於各id#的得分記憶體。

最後對於各id#，其儲存的字元中其音素串的字母的長度剛好等於話語的長度，則可將等於10的額外分數加入得分記憶體。

將說話的所有語音元素位置與儲存資料比較後，即可按照分數的高低將包含分數的得分記憶體位置排序，並輸出與三個最高分數有關的id#。

由此可知上述的資料處理序列較簡單。用已知方法處理過說話以找出一字母或音素串序列後，即將字串的各位置中的字母或音素當作參考以決定讀取第二資料庫中的那些資料，在得分記憶體中累加分數，當說話串中的所有位置都已查詢後，即將得分記憶體排序。

根據本發明各儲存的字元的分數是以各辨識的或翻譯的語音元素如字母或音素爲準，其方式大致與上述的類似，但是相對分數是根據翻譯機率的決定。因此在字母或音素

五、發明說明（ 22 ）

序列的各位置，各儲存字元的排序是根據翻譯的字母或音素的唯一性，以及該儲存字元中字母或音素眞實說出的機率。

再根據本發明，於使用字元增系統時更新機率值，以改正翻譯器與使用者的元素特性中原有的缺點。

在本發明的較佳具體實例中指派機率$P_{\alpha\beta}$給說話語音元素α與辨識或翻譯的語音元素β之各種可能組合。這是當說出一語音元素α時一語音元素已翻譯成一語音元素β的機率。對於語音元素序列中的各位置，將說出的語音元素翻譯成元素β時，即給說出的語音元素一分數以對應特定的機率$P_{\alpha\beta}$，並指派給翻譯元素β的組合，而序列中相同位置的眞實語音元素α則表示該字元。接著對於各儲存的字元，將所有位置的分數累加，而具有最高總分的字元即辨識爲說出字元。

根據另一種方法，開始時可以將前面數個序列位置中的所有儲存字元的分數累加，接著僅對於一選擇的儲存字元數持續給予分數，該字元於前面數個序列位置中的初始累加中具有最高的分數。

藉由以下例子可得到用以實施本發明的機率$P_{\alpha\beta}$：

數個不同喇叭將各字母發音，各喇叭將所有的字母發音一或數次；

接收器接收各發音的字母並將其翻譯；

每當發音字母α被翻譯成字母B時即得到一次計數值，而各α與β可以是字母A至Z中的任一者，而α可與β相同或

五、發明說明（23）

是不同；及

對於α與β的各種組合的機率$P_{\alpha\beta}$，是以次數N除以總次數N而得到，$N_{\alpha\beta}$是元素α翻譯成元素β的次數，而ΣN_{β}是說話元素翻譯成元素β的次數。

下表顯示理論上上述程序執行時收集到的部分資料：

表

$\alpha \backslash \beta \rightarrow$ ↓	A		C	D		F		U
A	89	---	0	0	---	1	---	0
B	1		5	3		0		0
C	1		51	0		0		0
D	0		11	5		1		0
E	1		4	4		0		0
L	0		0	0		1		0
P	1		2	2		0		0
R	0		0	0		1		0
S	1		0	0		36		1
U	0		0	0		0		111
ΣN_{β}	113		131	18		163		123

五、發明說明（ 24 ）

在表中ΣN_β的值包括出現在格子中但未顯示的計數值，從上表中計算得到的典型機率$P_{\alpha\beta}$的值如以下例子所示：

$$P_{AA}=\frac{89}{113}=0.788；$$

$$P_{BD}=\frac{3}{18}=0.167；$$

$$P_{PA}=\frac{1}{113}=0.009。$$

這些機率的意義如下：

P_{AA}-說話字母A翻譯成字母A的次數爲89次，而其他說話字母翻譯成A的次數爲113次。當一說話字母翻譯成A時該說話字母是眞的A的機率是89/113=0.788。

P_{BD}-說話字母B翻譯成字母D的次數爲3次，而其他說話字母翻譯成D的次數爲18次，因此當一說話字母翻譯成D時該說話字母是眞的B的機率是3/18=0.167。

P_{PA}-藉由類似分析，當一說話字母翻譯成A時，該說話字母是眞的P的機率是1/113=0.009。

用上表的資料來說明根據本發明之字元辨識法。

假設儲存在資料庫中的資料的字元包括：

ABACUS

APPLES

ABSURD

使用者拼出A-B-A-C-U-S而翻譯器聽到或是將該拼字翻譯成A-D-A-C-U-F。根據本發明要執行的工作是辨識眞實拼出的字元。因此將翻譯的元素序列與各字元的元素序列比較，其中藉由擷取機率而將資料儲存在資料庫中，當翻

五、發明說明（25）

譯的元素序列的各位置的元素翻譯成元素β時，說話元素α是儲存的字元素序列的相同位置中的儲存字元的元素。接著將各儲存字元的所有相關機率相加$\Sigma P_{\alpha\beta}$，並判定產生最大總和的儲存字元是眞實說出的元素的字元。

例如各數字元的各元素或字母的機率是：

A	B	A	C	U	S	=	$\Sigma P_{\alpha\beta}$
.788	.167	.788	.389	.902	.221		3.245
A	P	P	L	E	S	=	1.129
.788	.111	.009	.000	.000	.221		
A	B	S	U	R	D	=	.972
.788	.167	.011	.000	.000	.006		

因此系統指示說出的字元是ABACUS，因爲該字元的元素的$\Sigma P_{\alpha\beta}$是最大的。

根據本發明之另一特性，用以計算機率$P_{\alpha\beta}$的資料於每次字元辨識過程後要根據過程的結果而更新。特別地，假設辨識的字元其元素是由使用者說出，則翻譯的字元的各元素β與該辨識的字元的相同位置中的元素α有關，資料表中與α，β有關的各計數值存在於翻譯的字元的元素之間，若將翻譯的字元加1，則計算各行的新值ΣN_{β}，其中該行與翻譯的元素β有關。

五、發明說明（ 26 ）

因此在上述例子中從翻譯字ADACUF中辨識出是ABACUS，計數值N_{AA}會加2，而N_{BD}，N_{CC}，N_{UU}與N_{SF}則各加1。同理，ΣN_A會加2，而ΣN_C，ΣN_D，ΣN_P與ΣN_U會各加1。

此更新允許根據本發明之系統能以較大可靠度來辨識說話，方法是補償翻譯器的缺點以及/或是使用者部分的口吃。例如因此這種翻譯器的缺點以及/或是使用者的口吃，例如翻譯器常會將說話字母B翻譯成字母V。若眞的發生則更新程序會改變資料以逐漸增加P_{BV}的值。這會增加辨識包含字母B的說話字元的正確性。換言之，每當正確辨識字母V時，ΣN_{VV}即加1。在二種情況下ΣN_V都會加1。

這種程序對於單一使用者的系統特別有用。若系統具有多名使用者，各使用者計算機率的資料可以分開儲存，並提供一選擇器給系統，以便由使用者操作來辨識該使用者。根據選擇器的操作可用專屬於該使用者的資料來計算機率，並於每次使用後更新。

上述對於發音或說話字母的說明也適用於一種系統，其中說出字元的本身並將元素信號分成數個音素。在這二種情況下可使用傳統的翻譯器，因爲初始語音元素的判定並非本發明之新式功能。

根據本發明之辨識字元系統，藉由適當程式化的數位電腦即可實施該系統，如圖2所示。此系統包括：一CPU22，以控制系統操作，一第一資料儲存單元24，以儲存表示複數個字元的資料，一第二資料儲存單元26，以儲存用以計

五、發明說明（27）

算機率$P_{\alpha\beta}$的資料，元素介面10，以接收說話語音以及將說話語音元素轉成數位信號，及翻譯器30，接至介面10以接收數位信號並翻譯該信號以產生各說話語音元素之辨識，及一比較與計算單元32，以使用單元26中資料產生的機率$P_{\alpha\beta}$，將翻譯的語音元素與儲存在單元24中的資料比較，以辨識由該儲存資料表示的字元，該資料對應說話語音元素。在CPU22的控制下將單元24，26與翻譯器30連接以供應資料與信號至單元32。又連接單元32以與辨識字元的語音元素有關的資料，翻譯器30經由CPU22而連接，以供應與翻譯的語音元素有關的資料給單元26以允許資料在單元26中更新。單元32的一輸出接至輸出裝置如語音合成器12，其以看得見或聽得見的方式提供辨識的字元。最後將輸入面板36與CPU22連接以允許使用者執行適當的人工控制動作。

雖然上述說明係針對本發明之特別具體實例，將可了解的是在不偏離本發明之精神下，可作許多其他修正。附屬的申請專利範圍其目的在使這些修正落實在本發明之眞實範圍與精神中。

因此可以視本文揭露之具體實例爲一般說明性質而非限制，本發明之範圍如附屬的申請專利範圍所示，而非上述說明，因此符合申請專利範圍之等效者的意義與範圍的所有變化皆包括在其中。

A5
B5

四、中文發明摘要（發明之名稱：用以當拼字或說話時改良辨識一大資料庫之字元的可靠度之方法和裝置）

一種辨識複數個字元中之任一字元之方法與裝置，各字元具有一聽得見之形式，其中一序列說話語音元素表示，而各語音元素在序列中有一個別位置，其包括：接收一字元之諸說話語音元素並翻譯各收到之語音元素，其中各說話語音元素係一語音元素 α，各翻譯之語音元素係一語音元素 β，並可將各說話語音元素 α 翻譯成複數個不同語音元素 β 中之任一者，一語音元素 β 與語音元素 α 相同；將一個別之複數機率 $P_{\alpha\beta}$ 指派給各可能之語音元素，當已說出一

英文發明摘要（發明之名稱：METHODS AND APPARATUS FOR IMPROVING THE RELIABILITY OF RECOGNIZING WORDS IN A LARGE DATABASE WHEN THE WORDS ARE SPELLED OR SPOKEN）

A method and apparatus for identifying any one of a plurality of words, each word having an audible form represented by a sequence of spoken speech elements, with each speech element having a respective position in the sequence, which involves: receiving spoken speech elements of a word and interpreting each received speech element, wherein each spoken speech element is a speech element α, each interpreted speech elements is a speech element β, and each spoken speech element α may be interpreted as any one of a plurality of different speech elements β, one of the speech elements β being the same as speech element α; assigning to each of the possible speech elements a respective plurality of probabilities, $P_{\alpha\beta}$, that the speech element will be interpreted as a speech element β when a

（請先閱讀背面之注意事項再填寫本頁各欄）

裝 訂 線

A5
B5

四、中文發明摘要（發明之名稱：　　）

語音元素α時，即可將語音元素翻譯成一語音元素β；儲存表示各字元之資料，各字元之資料包括字元中各語音元素之辨識，以及在表示字元之語音元素序列中各語音元素個別位置之辨識；接收一序列由一個人說出之語音元素並表示一儲存字元，且翻譯說話中之各語音元素，以及各語音元素在說話語音元素序列中之位置；並將諸翻譯語音元素與表示複數個字元之各字元之儲存資料比較，並執行一計算，使用與各翻譯語音元素β相關之機率$P_{\alpha\beta}$來辨識複數個字元中之字元，而字元之諸語音元素極密切對應諸翻譯語音元素。

英文發明摘要（發明之名稱：　　）

speech element α has been spoken; storing data representing each word, the data for each word including identification of each speech element in the word and identification of the respective position of each speech element in the sequence of speech elements representing the word; receiving a sequence of speech elements spoken by a person and representing one of the stored words, and interpreting each speech element of the spoken word and the position of each speech element in the sequence of spoken speech elements; and comparing the interpreted speech elements with stored data representing each word of the plurality of words and performing a computation, using the probability, $P_{\alpha\beta}$, associated with each interpreted speech element β to identify the word of the plurality of words whose speech elements correspond most closely to interpreted speech elements.

（請先閱讀背面之注意事項再填寫本頁各欄）

裝　訂　線

六、申請專利範圍

1. 一種使用一程式化數位資料處理系統以辨識複數個字元中之任一者之方法與裝置，各字元具有一聽得見之形式，其中一序列說話語音元素表示，而各語音元素在該序列中具有一個別位置，該數位資料處理系統接至裝置以接收一字元之諸說話語音元素，並翻譯各收到之語音元素，

其中有複數個可能語音元素，各說話語音元素係一語音元素α，各翻譯語音元素係一語音元素β，並可將各說話語音元素α翻譯成複數個不同語音元素β中之任一者，一語音元素β與語音元素α相同，該方法包含：

將一個別之複數機率$P_{\alpha\beta}$指派給各可能之語音元素，當已說出一語音元素α時，即可將該語音元素翻譯成一語音元素β；

儲存表示複數個字元中之各字元之資料，各字元之資料包括字元中各語音元素之辨識，以及在表示該字元之語音元素序列中各語音元素個別位置之辨識；

在接收與翻譯之裝置中，接收一序列由一個人說出之語音元素並表示一儲存拼字，且翻譯該說話中之各語音元素，以及各語音元素在該說話語音元素序列中之位置；及

將該等翻譯語音元素與表示該複數個字元之各字元之儲存資料比較，並執行一計算，使用與各翻譯語音元素β相關之機率$P_{\alpha\beta}$來辨識該複數個字元中之字元，而字元之諸語音元素極密切對應諸翻譯語音元素。

六、申請專利範圍

2. 根據申請專利範圍第1項定義之方法，其中該執行一計算之步驟，包含：相加該機率$P_{\alpha\beta}$，其相關於該收到語音元素序列之翻譯語音元素β與相同位置中之語音元素α，作爲至少數個該複數字元之翻譯語音元素，及決定該字元數之字元，其與最大總和相關。

3. 根據申請專利範圍第2項定義之方法，包含以下初始步驟，使各可能語音元素說出一已知次數N_{α}，在裝置中翻譯各說話語音元素以接收並翻譯，決定各說話語音元素α被翻譯成一語音元素β之次數$N_{\alpha\beta}$，以及對於一個別說話語音元素α與一個別翻譯語音元素β之各種組合，計算一機率$P_{\alpha\beta}$，其等於α翻譯成β之次數$N_{\alpha\beta}$除以所有$N_{\alpha\beta}$之總和，$N_{\alpha\beta}$係指所有說話語音元素α個別翻譯成語音元素β之次數。

4. 根據申請專利範圍第1項定義之方法，又包含該步驟，於該等比較與執行一計算步驟後藉由增加各$N_{\alpha\beta}$一單位，再計算該機率$P_{\alpha\beta}$，$N_{\alpha\beta}$與各翻譯語音元素β相關，而該語音元素α之位置與該翻譯語音元素在辨識字元中之位置相同。

5. 根據申請專利範圍第1項定義之方法，其中當拼出一字元時各語音元素係一說話字母。

6. 根據申請專利範圍第1項定義之方法，其中當說出一字元時各語音元素係一發音音素。

7. 一種程式化之數位資料處理系統，用以辨識複數個字元中之任一者，各字元具有一聽得見之形式，其由一序列

六、申請專利範圍

說話語音元素表示，而各語音元素在該序列中具有一個別位置，其中有複數個可能語音元素，各說話語音元素係一語音元素α，各翻譯語音元素係一語音元素β，並可將各說話語音元素α翻譯成複數個不同語音元素β中之任一者，一語音元素β與語音元素α相同，該裝置包含：

第一資料儲存裝置，以儲存各可能語音元素之個別複數機率$P_{\alpha\beta}$，當已說出一語音元素α時，會將該語音元素翻譯成一語音元素β；

第二資料儲存裝置，以儲存表示複數個字元中之各字元之資料，各字元之資料包括字元中各語音元素之辨識，以及在表示該字元之語音元素序列中各語音元素個別位置之辨識；

裝置，用以接收一序列由一個人說出之語音元素並表示一儲存拼字，且翻譯該說話中之各語音元素，以及各語音元素在該說話語音元素序列中之位置；及

裝置，連接以便將該等翻譯語音元素與表示複數個字元之各字元之儲存資料比較，並執行一計算，使用與各翻譯語音元素β相關之機率$P_{\alpha\beta}$來辨識該複數個字元中之字元，而字元之諸語音元素極密切對應諸翻譯語音元素。

8. 根據申請專利範圍第7項定義之系統，其中該用以比較與執行一計算之裝置，包含：裝置，用以相加該機率$P_{\alpha\beta}$，其相關於該收到語音元素序列之翻譯語音元素β與

六、申請專利範圍

相同位置中之語音元素α，作爲至少數個該複數字元之翻譯語音元素，及裝置，用以決定該字元數之字元，其與最大總和相關。

9. 根據申請專利範圍第8項定義之系統，又包含裝置以執行一初始步驟，使各可能語音元素說出一已知次數N_{α}，在裝置中翻譯各說話語音元素以接收並翻譯，決定各說話語音元素α被翻譯成一語音元素β之次數$N_{\alpha\beta}$，以及對於一個別說話語音元素α與一個別翻譯語音元素β之各種組合，計算一機率$P_{\alpha\beta}$，其等於α翻譯成β之次數$N_{\alpha\beta}$除以所有$N_{\alpha\beta}$之總和，$N_{\alpha\beta}$係指所有說話語音元素α個別翻譯成語音元素β之次數。

10. 根據申請專利範圍第7項定義之系統，又包含裝置藉由增加各$N_{\alpha\beta}$一單位再計算該機率$P_{\alpha\beta}$，$N_{\alpha\beta}$與各翻譯語音元素β相關，而該語音元素α之位置與該翻譯語音元素在辨識字元中之位置相同。

11. 一種使用一程式化數位計算系統以辨識複數個字元中之任一者之方法，各字元具有一聽得見形式，由一語音元素序列表示，各元素在序列中具有一個別位置，其中各語音元素具有：至少一可辨識之聽覺特徵，及複數個語音元素，其大致相等於至少一可辨識之聽覺特徵，該方法包含：

在該數位計算系統中儲存對應各該複數個字元之數位表示；

接收一序列由一個人說出之語音元素，並表示一複數

六、申請專利範圍

個字元之聽覺形式，與儲存諸收到語音元素之表示及其在該說話序列中之個別位置；

在該說話序列中之各位置決定各語音元素，而非儲存之語音元素表示，其大致相等於該儲存之語音元素表示，其相對於至少一可辨識之聽覺特徵；

將儲存之諸語音元素表示與一字元之決定之諸語音元素之諸組合，與諸儲存字元比較；及

辨識該比較之儲存字元，以產生與一語音元素組合之最佳匹配。

12. 根據申請專利範圍第11項定義之方法，又包含再生該儲存字元，其於該辨識步驟中已辨識出。

第 85105749 號中文申請案
中文圖式修正本(86 年 3 月)

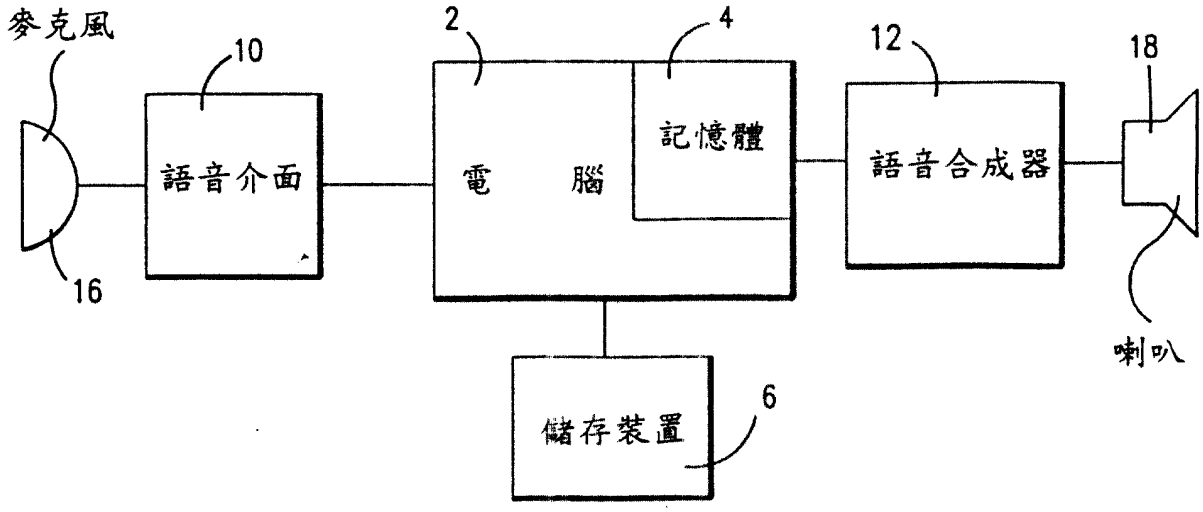

圖 1

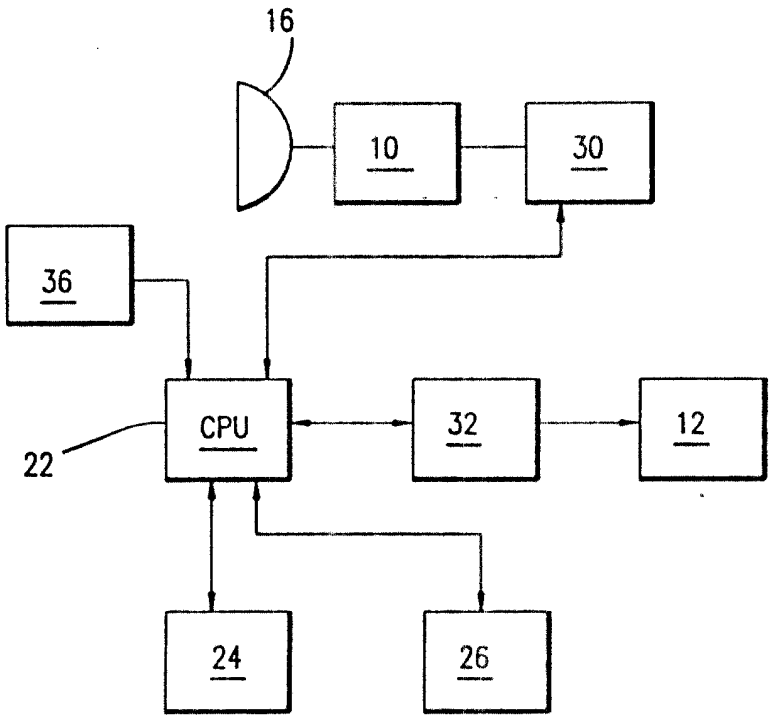

圖 2