

公告

申請日期	90. 8. 14
案 號	90119955
類 別	G10L15 ⁰⁰

A4
C4

521262

(以上各欄由本局填註)

發 明 專 利 說 明 書

一、 發明 名稱	中 文	提高口述及命令辨識力之方法
	英 文	A METHOD FOR ENHANCING DICTATION AND COMMAND DISCRIMINATION
二、 發明 創作	姓 名	1. 詹姆士 R. 路易斯 2. 凱利 A. 歐德卡
	國 籍	1. 美國 2. 美國
	住、居所	1. 美國佛羅里達州德瑞海灘莊嚴棕櫚路 4000 號 2. 美國北卡羅萊納州洛利伍曼納大道 1329 號
三、申請人	姓 名 (名稱)	美商·萬國商業機器公司
	國 籍	美國
	住、居所 (事務所)	美國紐約州阿蒙市新果園路
	代 表 人 姓 名	傑拉德羅森瑟爾

裝

訂

線

(由本局填寫)

承辦人代碼：
大類：
IPC分類：

A6

B6

本案已向：

國(地區) 申請專利，申請日期： 案號： 有 無主張優先權

本案已向美國申請專利；申請日：2000年9月20日 案號：09/665,939號

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

經濟部智慧財產局員工消費合作社印製

五、發明說明（ ）

發明領域：

本發明係關於語音辨識領域。更詳而言之，本發明係關於一種提高對各使用者口述、使用者聲音命令與文字嘜以辨識之能力的方法。

發明背景：

語音辨識是指將一話筒所收到的一聲音信號經由一電腦轉變成文字的過程。然後，可以將所認出的文字用來作為種種電腦軟體應用，例如文件準備、資料登記和指令與控制的目的。語音口述系統還提供使用者不需使用手的一電腦系統操作方法。

關於電子文件的準備，目前可使用的語音口述系統提供可使一使用者在一電子文件中選擇一部份文字的使用者聲音命令。一般這樣的使用者聲音命令是使用一語法(syntax)，例如"選擇<文字>"，是代表在使用者聲音命令"選擇"信號命令後方所跟隨的文字應該被選擇或被加亮。在已選擇一部份文字之後，使用者能夠根據所選擇的文字執行任何一連續的後續操作。

因此，如果有一使用者說："選擇你好嗎"，語音口述系統將在電子文件內的一文字正文中查尋文字片語"你好嗎"。置於文字正文內的片語可以被選擇或被加亮。

然後，使用者可以對所選擇的文字執行一操作，例如一刪除操作，一粗體/斜體/底線操作時，或訂正操作。在更進一步的說明中，當"你好嗎"文字被加亮時，使用

五、發明說明（ ）

者所選擇之文字可以被來自一後續使用者表達的不同文字所代替。在這個模式下，使用者能夠執行一電子文件的不需手動之訂正(hands-free correction)。

目前，既有之"選擇"命令的執行，或其他用於選擇文字的類似使用者聲音命令，具有幾個缺點。其中一缺點是使用者想要在一文字正文之內所選擇的片語或字詞可能會出現許多次。例如，在一文字正文之內很有可能會有許多次"這"字詞的出現。因此，如果使用者說："選擇這"語音口述系統可能無法決定哪個"這"被選擇"字詞的出現使用者想要選擇。

為解決此問題，習知的語音口述系統依靠一標準系統來決定哪個使用者所想要字詞或片語是使用者所想要選擇的。舉例來說，一語音口述系統可以從現用視窗的頂部開始搜尋，並選擇該字詞或片語的第一次出現。然而，如果該使用者不想要選擇該字詞或片語的第一次出現，一習知的語音口述系統可以為該使用者提供選擇該字詞另一次出現的能力。尤其一些習知的語音口述系統提供導引聲音命令例如"下一個"或"先前的"。

透過發出"下一個"聲音命令該使用者指示語音口述系統找出和選擇所想要的字詞或片語的下一次出現。

同樣地，"先前的"命令指示語音口述系統會找出並選擇所想要的字詞或片語的先前的出現。雖然如此，習知的系統可使該使用者操縱一特定字詞或片語所想要的出現，但是使用者必須詳盡闡述用以操縱所想要之出現

五、發明說明（ ）

的對策。它可以導致時間的浪費和使用者的挫折感，尤其是當該使用者察覺語音口述系統不十分精確或低效能的時候。

在習知的語音口述系統之內的習知文字選擇方法的另一缺點是，當要查尋使用者所指定的字詞或片語時，如此的語音口述系統一般是搜尋出現在使用者的螢幕上之一文字正文的整個部分。在使用者的螢幕上所出現的每一字詞都在語音口述系統文法之內被活化，而對語音口述系統而言均具有相同之可能性。由於該使用者僅想要一單一字詞或片語，因此允許和搜尋在使用者的螢幕上出現文字正文的整個部分可能是低效能的。此外，此方法可能會增加一錯誤識別將出現的可能性。

習知語音口述系統之內的習知文字選擇方法的另一缺點是，對語音口述系統而言，一使用者是否在語音口述或一聲音命令期間已經發出一字詞並不是立即而明顯，例如使一下拉式(drop-down)選單產生的聲音命令。舉例來說，如果一使用者發出字詞"檔案"，依據其情況，使用者可能是打算使用在選單桿內之檔案選單或把字詞"檔案"插入電子文件中。於是，對普通的語音口述系統而言，一使用者是否使用者的發聲表達是一聲音命令或語音的口述並非總是十分明顯。

因此，雖然目前可利用的語音口述系統提供與一電腦相互作用的方法而可聽見地命令一應用，以在一電子文件中提供語音口述並在電子文件之內選擇文字，而仍然需要

五、發明說明（ ）

用以辨識使用者聲音命令、使用者口述、文字和其結合的一改進的方法。

發明目的及概述：

在此處所揭露的本發明係提供一方法和裝置用以辨識在一電子文件中文字的不同出現，和利用結合一眼睛追蹤系統的一語音口述系統用以辨識一聲音命令的一實例和語音口述的一實例。

本發明的方法和裝置可以有利地包括與一語音口述系統合併發揮的一眼睛追蹤系統(eye-tracking system, ETS)，以決定在一語音口述系統期間一使用者的凝視的焦距點。更詳而言之，與語音口述系統合併發揮的眼睛追蹤系統可以改善"選擇"使用者聲音命令功能性，或在一語音口述系統中的一文字正文之內選擇一部份文字之任何其他使用者聲音命令的精確度。本發明之眼睛追蹤系統的使用也可以透過促進使用者口述和聲音命令之間之辨識來改進系統性能。

依據這些本發明之安排，一種用以在一電子文件中查尋與文字相匹配的方法可包括識別在一使用者界面中的一焦距點和界定焦距點周圍的一周圍區域。更詳而言之，該周圍區域可以包括一文字正文在被安裝用以接收語音命令文字的一使用者界面物體之內。另外，該方法可包括接收一用以在電子文件之內選擇指定的文字並在該周圍區域所包括的文字正文中搜尋符合該指定文字的

五、發明說明（ ）

聲音命令。值得注意的是，該搜尋可以被侷限於在該周圍區域內之文字正文。

如果搜尋步驟中，在文字正文中沒找到一符合指定文字的文字時，用以搜尋在一電子文件中之符合文字一方法可以另包括擴展該周圍區域以包含該使用者界面的一額外範圍。更詳而言之，該擴張所包括的額外範圍能夠包括額外的文字。因此，可以對額外的文字搜尋一符合指定文字的文字。最後，如同先前，該搜尋可以被侷限於文字正文和額外文字。

在本發明的一實施例中，擴展步驟可以包括以一固定的增加量從焦距點向外的擴展周圍區域。或者，擴展步驟可以包括以與文字正文緊接的一可變數量文字使周圍區域擴展。最後，擴展步驟可以包括以一可變的增加量使焦距點的周圍區域向外擴展。

一種用以辨識一聲音命令的一實例和語音口述的一實例兩者的方法可以至少包括在一使用者界面內識別一焦距點；限定一周圍區域圍繞該焦距點；確認使用者界面物件在該周圍區域內；再確認在已確認之使用者界面物件之間那些是被裝配用以接收語音口述文字的使用者界面物件和那些不是用以接收語音口述文字的使用者界面物件；以那些使用者界面物件為根據計算出一已被再次確認是用以接收語音口述文字的使用者界面物件和那些不是用以接收語音口述文字的使用者界面物件的可能性；接收語音輸入；及根據計算所得之可能性判斷下該語音輸入是

五、發明說明（ ）

一聲音命令或是語音口述的一決定。此外，該方法可以包括在使用者界面之外識別一焦距點；及根據一隱含值(default)可能性判斷該語音輸入是一聲音命令或語音的口述的一決定。

圖式簡單說明：

下列之較佳實施例描述與其伴隨之圖示將更詳細的解釋本發明之方法，然而應被瞭解的是本發明不應被這些所顯示之精確安排與工具所侷限。

第1圖是一使用者與在此處揭露的本發明相互作用的一示範說明。

第2圖是說明適用於本發明合適的一電腦系統的方塊圖。

第3圖是在第1圖中之電腦系統的一特有高層級內部結構的方塊圖。

第4圖顯示包括一語音辨識引擎之特有零件的方塊圖。

第5A圖和第5B圖，將一同被參考，組成一流程圖說明在一電子文件中辨識文字的不同出現，和透過使用與一語音口述系統結合的一眼睛追蹤系統辨識一聲音命令的一實例和語音口述的一實例兩者的一方法。

圖號對照說明：

20 電腦系統

22 眼睛追蹤系統

五、發明說明()

- | | |
|---------------|--------------|
| 23 音頻輸出裝置 | 24 作業系統 |
| 26 語音口述系統 | 27 記憶存儲器 |
| 27A 電子隨機存取記憶體 | 27B 大量資料存儲媒體 |
| 28 語音文字處理器 | 29 眼睛追蹤硬體界面 |
| 30 音頻輸入裝置 | |

發明詳細說明：

本發明示利用與一語音口述系統結合的一眼睛追蹤系統(ETS)，以改進一語音口述系統的性能。更具體而言，依據下列的發明安排，一眼睛追蹤系統(ETS)可以幫助一語音口述系統辨識在一文字正文內之文字的多次出現。

另外，一眼睛追蹤系統(ETS)可以在分析語音輸入中幫助語音口述系統辨識聲音命令和語音口述。如此的提升可以藉由檢測在一眼睛追蹤系統(ETS)中一使用者之凝視焦距點的螢幕位置。有助益的是，不論這個螢幕位置是否位於螢幕上或離開螢幕均可傳遞至語音口述系統。基於使用者的凝視焦距點的位置圍繞焦距點的一區域(參考"周圍區域")可以被界定而有助於決定語音輸入是否是聲音命令或語音口述。另外，可以用此周圍區域來確認符合使用者所選擇之文字的一特定出現。

第 1 圖是一使用者與在此處揭露的本發明相互作用的一示範說明。在第 1 圖中，使用者凝視一影像顯示終端機(VDT)32 上的一位置。利用位於影像顯示終端機 32

五、發明說明 ()

之螢幕上的一星號來標示使用者的凝視的焦距點。同時被顯示的是具有一頭嵌式(head-mounted)硬體界面 29 的一眼睛追蹤系統(ETS)。眼睛追蹤系統(ETS)是熟知被用於眼睛追蹤和測量的技術中。眼睛追蹤系統(ETS)，例如由維吉尼亞州 Fairfax 之 LC 技術公司所製造之 EYEGAZE 發展系統，及 EYEMOUSE 和 EYELINK，二兩者由麻州波士頓之 SensoMotoric 設備公司所製造，現在均提供為商業利用。

一眼睛追蹤系統(ETS)的架構可以至少包括一眼睛追蹤硬體界面 29 和一圖像處理系統 34。眼睛追蹤硬體界面 29 可以是可從 LC 技術公司所購得之一桌上型內嵌單元。一示範的桌上型內嵌單元眼睛追蹤單元在第 2 圖中顯示。或者，眼睛追蹤硬體界面 29 可以是可從 SensoMotoric 公司所購得之一頭嵌式單元並顯示在第 1 圖中。不論是一桌上型內嵌單元或一頭嵌式單元，眼睛追蹤硬體界面 29 可以將關於一使用者眼睛的資訊傳遞給圖像處理系統 34。

圖像處理系統可以是一獨立的圖像處理系統，或則也可以存在於一普通電腦內。當圖像處理系統是在一普通電腦內存在時，普通電腦可以利用圖像處理電路系統和圖像處理軟體的一結合以執行一圖像處理系統的功能。那些擅長此項技術之人員應該理解，本發明並不受所選擇之眼睛追蹤系統(ETS)所限制。更確切地說是，任何可以把一使用者的凝視焦距的位置傳遞給電腦的任何

五、發明說明（ ）

合適眼睛追蹤系統(ETS)都可以被使用。

第2圖說明圖像處理系統34是以一普通電腦為基礎的圖像處理系統的環境。更進一步詳細說來，一圖像處理系統34可以包括具有一中央處理器(CPU)，一或更多的記憶體裝置和相關之電路系統的一普通電腦20。普通電腦20可以包括電腦記憶存儲器27，其最好是由一電子隨機存取記憶體27A和大量資料存儲媒體27B所構成，例如一磁碟驅動機。最後，電腦20可以包括一指示裝置21(例如一滑鼠)及至少一個使用者界面顯示單元32，例如一影像顯示終端機(VDT)在操作上連接。

更詳而言之，電腦20可用以完成語音辨識及文字-對-語音(text-to-speech; TTS)轉換。就其本身而言，電腦20可以另包括一音頻輸入裝置30，例如一話筒。此外，電腦20可以包括一音頻輸出裝置23，例如擴音器。音頻輸入裝置30和音頻輸出裝置23二者可以透過合適界面電路系統操作地或"音效卡"(未顯示)連接在電腦20上。透過這個方法，可以利用音頻輸入裝置30接收使用者語音至電腦20裡，而合成語音和其他音頻可以利用音頻輸出裝置23提供給使用者。如上面所描述的普通電腦20所需要的各種硬體通常是任何一各式可達到要求的商用高速多媒體個人電腦可滿足，例如由IBM公司所製造之。

依據這些發明安排，電腦20可以另包括一眼睛追蹤硬體界面29(在此處展示了桌上型種類)，透過電腦20的一通訊端口(未顯示)操作地連接至電腦20，再透過合適的圖

五、發明說明（ ）

像處理電路系統和軟體通訊地連接至電腦20。更特別的是，圖像處理電路系統和軟體可以決定使用者的凝視焦距的位置，也可以把此資訊傳遞給與圖像處理軟體通訊相連接的電腦運用。在本發明中，一語音口述系統可以被通訊地連結至圖像處理軟體，而語音口述系統可利用所收到的資料指示出一使用者的凝視的焦距點的位置。

第3圖說明一語音操作(speech-enabled)電腦系統的特有內部結構，該電腦系統包含一眼睛追蹤系統(ETS)，而電腦系統是用以辨識在一電子文件內之文字的多次出現和辨識一聲音命令的一實例和語音口述的一實例兩者。如第3圖中所示，電腦系統20包括一記憶存儲器27、一作業系統24、一語音口述系統26和一眼睛追蹤系統22。在實施例中顯示，一語音文字處理器28的應用也被提供。然而本發明在這一點上，而且語音口述系統26可以用任何其他允許聲音使用的應用程式。

在第2圖中，語音口述系統26、語音文字處理器28和眼睛追蹤系統22被顯示是分離的應用程式。然而應該注意的是，本發明並沒有被限制在這一點上，而這些不同的應用程式可以執行如一單一而更複雜的應用程式。例如語音口述應用26可以與語音文字處理機應用或是任何與語音口述系統一起使用的其他應用相結合。此外，眼睛追蹤系統22可以是一應用程式儲存於電腦20中，或是位在一獨立眼睛追蹤系統通訊的通訊範圍內而透過一資料連結與電腦20通訊。該系統也可以包括一聲音導引

五、發明說明()

裝置應用(未顯示)，用以協調語音口述系統對其他應用程式的聲音操作的操作，但是在此處所描述之本發明的操作並不需要此應用。

第4圖顯示特有零件的方塊圖，其說明在語音口述系統26中語音-對-文字轉變的一語音信號。一般而言，數位語音信號可以透過在第2圖所顯示的一音頻輸入裝置被接收，並在音頻電路系統中被處理成為類比語音信號。

更具體而言，可以以某種固定的速率抽樣語音信號，以使語音信號可轉換成一組數位化資料。隨後，音頻電路系統可以把數位化語音信號傳遞給語音口述系統26。

表述方塊35可以接收數位化語音信號並可以產生一數位化語音信號的表述，而被使用於在語音辨識過程的後續階段以決定一部份語音信號與一特限定語音事件相符合的可能性。此過程是用以加強從操作系統所接收之語音信號，該語音信號具有不因說話者而異的知覺上重要特徵。

在模式化/分類方塊36中，演算法可以更進一步處理語音信號，以使不因說話者而異的聽覺模式適應那些現今的說話者。最後，在搜尋方塊38中，根據該語音信號，搜尋演算法係用以導引搜尋引擎至與語音信號最符合的字詞。搜尋方塊38中的搜尋過程在聽覺模型40，字詞彙模式42，語言模式44和訓練資料46的幫助下出現。

依據這些發明安排，辨識在一電子文件中文字的不

五、發明說明（ ）

同出現和聲音命令的一實例和語音口述的一實例的一方法和裝置在此處被揭露。本發明的方法和裝置可以包括與一語音口述系統結合的一眼睛追蹤系統(ETS)的合作使用。更詳而言之，此結合可以改進使用者聲音命令功能"選擇"的精準度，或是在一語音口述系統中任何其他用以在一文字正文內選擇一部份文字的使用者聲音命令的精準度。此結合也可以透過協助語音口述系統把語音輸入當作語音口述或聲音命令而改進語音口述系統的性能。

上述對一語音口述系統的增強可以藉由以檢測得到的一使用者的凝視焦距點做為基礎，計算出短暫地近似使用者的凝視焦距點的語音輸入是語音口述之一或是一聲音命令之一可能性。計算所得的可能性可用以使一語音口述系統把語音輸入視為語音口述之一或是一聲音命令。更詳細而言，由於語音口述系統可連續不斷地留存記錄和更新在該周圍區域內的文字和物件的有關資訊，語音口述系統可以界定圍繞該檢測所得到的焦距點的一可調整螢幕區域("周圍區域")。

當收到語音輸入時，語音口述系統可以決定該周圍區域主要是包含有使用者界面物體或是一文字輸入領域。如果此周圍區域主要包含有一文字輸入領域，語音口述系統可以得到如此的結論，應該把語音輸入當作語音口述，以插入文字輸入領域裡。相反地，如果此周圍區域主要包含的是使用者界面物體，語音口述系統可以

五、發明說明（ ）

把語音輸入當作是一聲音命令。最後，當把語音輸入當作是一聲音命令，以在文字輸入領域中的一文字正文中選擇一文字，語音口述系統可以確認所選擇的文字是基於在此周圍區域內的文字，而不是在這個文字輸入領域內的一全部文字正文。以這個模式，語音口述系統的資源可以更有效地致力於一更小的文字區域，而非一電子文件中的一整個文字正文。

第5A圖和第5B圖，一同參考，組成一流程圖說明在一電子文件中辨識文字的不同出現和透過使用與一語音口述系統結合的一眼睛追蹤系統辨識一聲音命令的一實例和語音口述的一實例兩者的一方法。該方法可以與為了一語音口述系統和一眼睛追蹤系統(ETS)的使用二者所設置的一電腦系統相連結而被執行。第5A圖從步驟50開始，其中當該使用者對語音口述系統提供語音輸入時，會自然地凝視在影像顯示終端機32上(在螢幕上)或不在影像顯示終端機32上(不在螢幕上)任一者的各種位置。

在步驟55中，眼睛追蹤系統確認使用者的凝視焦距點的位置。眼睛追蹤系統利用圖像處理電路系統和軟體的幫助來決定使用者的凝視的焦距點的位置是在螢幕上或是不在螢幕上。不論如何，眼睛追蹤系統會把此資訊傳遞給語音口述系統。在步驟60中，語音口述系統已從眼睛追蹤系統得到使用者的焦距點的位置。如果使用者的凝視焦距點的位置在螢幕上，系統便進行步驟70。否

五、發明說明（ ）

則，該系統則繼續步驟65。

如果在步驟60中已決定焦距點的位置是在螢幕上，眼睛追蹤系統也已確認使用者的凝視焦距點的螢幕位置。因此，在步驟70中，可在焦距點周圍界定一周圍區域。在一實施例中，可根據從焦距點向外擴展的一特定半徑的一周邊來界定該周圍區域。或者，也可以藉由在焦距點上覆蓋一預定的幾何區域來界定該周圍區域。

可是，本發明並不僅限於用以計算該周圍區域的方法。更確切的說，任何用於計算該周圍區域的合適方法均可以滿足本發明的目的。此外，任何熟知該項技術之人員都應理解，不管是如何決定該周圍區域或是該周圍區域的最終形狀，在一外部周邊區域之內的預設區域或尺寸均可以是一使用者可調整值。例如，使用者可以指定一預設區域或在其中選擇其一，使用者可以指定該周圍區域應該從焦距點向外擴展的一半徑。

在步驟75中，在界定該周圍區域之後，在該區域之內有關文字和物體的資訊可以被獲得，以用以決定是否應該把語音輸入當作語音口述或是當作一聲音命令，並在一電子文件中確認符合之文字的一特定出現。更特別的是，所獲得之資訊可包括，例如用以顯示使用者界面但不適合用以接收語音口述文字的像素數目，和用以顯示使用者界面並適合用以接收語音口述文字的像素數目。應該被理解的是，經由界定一限制區域可使語音口述系統可提供其資源，而語音口述系統會達到更高的效率。

五、發明說明()

例如，語音口述系統僅需要在周圍區域之內找到語音口述文法所含之文字的一些符合部分，而不是一整個語音口述文法。

在步驟80中，基於語音口述可被視為是一聲音命令或是語音口述，可計算出一可能性。更具體而言，可以藉由計算該周圍區域的可口述範圍與該周圍區域的總範圍相比較的一比例來計算得到該可能性。例如，如果70%的周圍區域可以接收使用者口述，則可能性會是70%或0.70。可是，本發明並沒有被侷限於計算可能性的特定模式中。而事實上，可能性的其他計算方式可以是以，例如在該周圍區域內之原文的或命令的字詞的數目與在可用於使用者聲音命令的周圍區域內之物體的數目兩者的比率等為基礎。

儘管如此，不管可能性是如何被計算的，應該被瞭解的是，較佳的可能性不會是零也不會是代表後續使用者發聲將一定是使用者口述或使用者的聲音命令的一完全必然結果。由於不容許如此極端可能性值，可容許使用者沒有凝視螢幕卻想要對語音口述系統命令語音的情況。

如果，在決定步驟60中，它決定使用者的凝視焦點是在一離開螢幕的位置時，在步驟65中系統可以將或然率設定成一預設值。

這個預設值被稱為隱含可能性，並可以由使用者預先設定。隱含可能性是代表當使用者的凝視不是在螢幕

五、發明說明()

上時，後續的語音輸入是語音口述之一或一聲音命令的統計可能性。於是，以隱含，可能性為基礎的一統計分析可以表明當使用者不是在看螢幕但該使用者希望語音輸入被當作語音口述的可能性。

預設可能性可有一可調整值範圍從零(0.00)到一(1.00)。更詳而言之，任何熟知該向技術之人員都應理解的是，當分配一高數值給預設可能性時，即代表在語音口述期間該使用者不必直視螢幕的假設。然而，預設可能性不代表在使用者不直視螢幕時所提供的語音輸入應該完全必然被視作是語音口述或一聲音命令會是一較佳的作法。如此的必然可能性可能會在語音口述系統內產生錯誤結果。

在步驟85中，在計算一可能性或依靠一預設可能性以後，語音輸入可以被接收。基於藉由眼睛追蹤系統的協助而取得之可能性，可以分析語音輸入以決定是否應該把該語音輸入視為語音口述或是一聲音命令。然後，本方法可以從跳移圓圈A跳至第5B圖中的決定步驟95以繼續處理語音輸入。

在決定步驟95中，可以決定在步驟85中所收到的語音輸入是否是"選擇"聲音命令，或是用以在一電子文件之內選擇文字的其他的類似聲音命令。如果語音輸入被解釋為不是選擇命令，本方法將繼續步驟97，兩個行動的其中之一。首先，如果語音輸入被認定是另一聲音命令但不是選擇聲音命令時，該聲音命令可以被當作是一

五、發明說明（ ）

習知語音感應應用的情況。接著，如果語音輸入被認定是語音口述時，可以藉由一語音辨識引擎使該語音輸入轉換成為文字。其後，可以把已轉換的文字插入，用以接收轉變的文字的一使用者界面物體。不論哪一種情況，本方法均可以透過跳移圓圈C返回到第5A圖中的步驟50，並且可以重複流程。

再回到決定步驟95中，如果決定在步驟85中所收到的語音輸入是一選擇聲音命令或是用以在一電子文件之內選擇文字的其他的類似聲音命令時，在步驟100中可以決定符合選擇命令之指定的文字是否位在此周圍區域所包含之文字正文中。例如，如果已經把語音輸入視為選擇命令，"選擇老鼠"此選擇命令，可以決定在此周圍區域所包含之文字正文是否包有括字詞"老鼠"。如果在步驟100中為指定文字找到一符合文字，本方法可以繼續進行步驟105。反之，本方法則可以繼續進行步驟110。

依據步驟100，如果為指定文字找到一符合文字，在步驟105中，可以選擇對指定文字而言最適當的符合文字。更具體的說是，如果在此周圍區域之文字正文內只有一符合文字，則該唯一文字符合實例會被選擇，一般是加亮該文字之符合出現。相反地，如果在此周圍區域之文字正文內該符合文字有多次的出現，則可決定在此周圍區域之文字正文中哪一個指定文字的實例是最接近焦距點。如此一來，可以利用使用者的凝視焦距點來決定應該選擇哪個符合文字的實例。但是，本發明並沒有

五、發明說明（ ）

被侷限於這一點，而也可滿足適合用以在符合文字的多次出現中選擇一符合文字的實例。如此的選擇方法可以包括選擇在此周圍區域之文字正文內該符合文字的第一次出現。

一旦已經確認指定文字的適當出現便可選擇該確認文字，一般是將文字加亮。需要瞭解的是，在這種情況下若有一不正確的或不所想要得到的指定文字的出現已被選擇時，一般的聲音命令，例如"先前的"或"下一個"可以被用來在此周圍區域內引導符合文字的其他出現。不論如何，本方法可以返回透過跳移圓圈C返回到第5A圖中的步驟50以再次開始流程。因此，藉由重複流程，本方法可再次計算此周圍區域並決定後續收到的語音輸入是語音口述或一聲音命令的可能性。

現在返回到決定步驟110中，如果在此周圍區域之文字正文內找不到任何符合文字時，可以決定此周圍區域是否包含用以接收語音口述的所有可看見的使用者界面。如果答案是確定，可以假設在螢幕上的文字正文中沒有符合文字，而如在步驟115中，使用者可以被通知。在未顯示於第5B圖的另一實施例中，若沒有符合文字存在於螢幕上時，本系統可以提供使用者額外選擇，以延續和另擴展針對使用者指定文字的搜尋。舉例而言，也可以詢問使用者是否想要搜尋目前所打開之電子文件的剩餘部分。此外，可以把更多挑出的選擇呈現給使用者，例如利用在此周圍區域之後或在其之前的一預定或

五、發明說明（ ）

使用者可調整之數目的字詞或段落來擴展此周圍區域。在任何情況下，本方法隨後仍可以返回透過跳移圓圈C返回到第5A圖中的步驟50以重新開始流程。

相反地，如果在步驟100中決定此周圍區域並沒有包含用以接收語音口述的所有可看見的使用者界面時，便無法假設在螢幕上的文字正文中沒有符合文字。如此一來，繼續步驟120，被此周圍區域所涵蓋的區域可以被擴展，以包含更多的文字。可以利用任何合適的方法來執行此周圍區域的一擴展。例如，可以從使用者焦距點以一預定的或可變地計算值向外均等地朝向所有方向擴展此周圍區域的外面周邊。此外，也可以從使用者焦距點以代表一區域測量的一預定值擴展此周圍區域。

在本發明的一實施例中，一隱含預定值可以被用於決定擴展的範圍。隱含值可以是可調整的以提供一良好調整功能。以這個模式，在一重複搜尋的期間一使用者可指定此周圍區域應該增加至多大。參考先前的實例，如果使用者所指定之文字"老鼠"並沒有在此周圍區域之文字正文內被找到時，此周圍區域的周邊可以從使用者焦距點朝所有方向向外擴展1公分。此外，也可以以5平方公分的一預定範圍或特定數目的像素擴展此周圍區域。

在此周圍區域的擴展之後，在步驟125中，位在此新擴展的周圍區域之內的物件和文字的有關資訊可以被計算、收集和存儲，以提供本發明方法的未來使用。此外，

五、發明說明（ ）

在此新擴展的周圍區域之內的新文字正文可以在語音口述系統文法之內被使用。同樣的，在此新擴展的周圍區域之內的物件的屬性可以被確認。在確認此新擴展的周圍區域之內的文字和物件後之，對文字正文中的符合文字的搜尋可以透過在步驟100中之跳移圓圈B重複開始。以這個模式，本方法可以有系統地和漸增地擴展對一文字正文內之使用者指定文字的搜尋於文字正文在螢幕上的部分並超出其部分。

更詳而言之，本發明可以利用硬體、軟體或硬體和軟體的一結合實行。本發明的方法可以在一電腦系統中以一集中方式或是以一分發方式實行，而不同元件被分散於數個相互連接的電腦系統之間。任何種類的電腦系統或其他的設備可用以實行此處所描述之方法均應適用。其中硬體和軟體的一常見結合可以是包含一電腦程式的一一般用途電腦系統，當本發明方法被載入與執行時可控制該電腦系統。

本發明也可以內建於一電腦程式產品中，其包括使本發明方法執行的所有特點，而且當電腦程式產品被載入於一電腦系統中便可以實行這些方法。電腦程式機制或在本語文機制中使用各種形式、任何語言、代碼或符號的一組電腦指令預期將使具有資訊處理能力的一系統可直接或在下列兩者或其中之一過程之後執行一特定功能：a)轉變至另一種語言、代碼或符號；b)以一不同材料形式再生產。

五、發明說明()

本發明已經按照一些實施例被描述。然而，並非將本發明侷限於說明和實施例的描述。其他不脫離本發明基本架構與精神的形式，皆應為本專利所主張之權利範圍，更確切的說是，應由下列的專利範圍界定本發明的範圍。

(請先閱讀背面之注意事項再填寫本頁)

訂

線

四、中文發明摘要(發明之名稱：)

提高口述及命令辨識力之方法

本發明係關於一種用以辨識一聲音命令的一實例和語音口述的一實例兩者的方法，該方法可以包括：確認一焦距點在一使用者界面上；界定一周圍區域在該焦距點四周；在該周圍區域內確認使用者界面物件；再確認在該已確認使用者界面物件之間那些使用者界面物件是用以接收語音口述正文，而那些使用者界面物件不是用以接收語音口述正文；以已再次確認是用以接收語音口述正文的那些使用者界面物件與已再次確認不是用以接收語音口述正文的那些使用者界面物件為基礎，計算一可能性；接收語音輸入；以所計算出的可能性為基礎，決定該語音輸入

英文發明摘要(發明之名稱：)

METHOD FOR ENHANCING DICTATION AND COMMAND DISCRIMINATION

A method for discriminating between an instance of a voice command and an instance of speech dictation can include identifying a focus point in a user interface; defining a surrounding region about the focus point; identifying user interface objects in the surrounding region; further identifying among the identified user interface objects those user interface objects which are configured to accept speech dictated text and those user interface objects which are not configured to accept speech dictated text; computing a probability based upon those user interface objects which have been further identified as being configured to accept speech dictated text and those user interface objects which have been further identified as not being configured to accept speech dictated text; receiving speech input; and, biasing a determination of whether the speech input is a voice command or speech dictation based upon the computed

四、中文發明摘要(發明之名稱：)

是一聲音命令或是語音口述。此外，此方法可另包括確認一焦距點在該使用者界面之外；及以一隱含可能性為基礎，決定該語音輸入是一聲音命令或是語音口述。

英文發明摘要(發明之名稱：)

probability. Additionally, the method can include identifying a focus point outside of the user interface; and, biasing a determination of whether the speech input is a voice command or speech dictation based upon a default probability.

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

六、申請專利範圍

1. 一種用以在一電子文件中搜索符合文字的方法，該方法至少包括下列步驟：

確認一焦距點在一使用者界面上；

界定一周圍區域在該焦距點四周，該周圍區域包含一文字正文在一使用者界面物件範圍之內，該使用者界面物件係用以接收語音口述文字；

接收一聲音命令，以挑選在該電子文件中的指定文字；及

在該周圍區域所包含之文字正文中搜尋該指定文字的一符合文字，該搜尋係侷限於該周圍區域內之文字正文中。

2. 如申請專利範圍第1項所述之方法，其中更包括下列步驟：

如果在該搜尋步驟中沒找到該指定文字的一符合文字，擴展該周圍區域以包含該使用者界面的一額外範圍，該額外區域包含有額外文字；及

在該額外文字中搜尋該指定文字的一符合文字，該搜尋係侷限於該文字正文和該額外文字中。

3. 如申請專利範圍第2項所述之方法，其中上述之擴展步驟更包括下列步驟：

從該焦距點以一不變增加量向外的擴展該周圍區域。

六、申請專利範圍

4. 如申請專利範圍第2項所述之方法，其中上述之擴展步驟更包括下列步驟：

以緊接該文字正文的一不變數量的文字擴展該周圍區域。

5. 如申請專利範圍第2項所述之方法，其中上述之擴展步驟更包括下列步驟：

從該焦距點以一可變增加量向外的擴展該周圍區域。

6. 一種用以辨識一聲音命令的一實例和語音口述的一實例的方法，該方法至少包括下列步驟：

確認一焦距點在一使用者界面上；

界定一周圍區域在該焦距點四周；

在該周圍區域內確認使用者界面物件；

再確認在該已確認使用者界面物件之間那些使用者界面物件是用以接收語音口述正文，而那些使用者界面物件不是用以接收語音口述正文；

以已再次確認是用以接收語音口述正文的那些使用者界面物件與已再次確認不是用以接收語音口述正文的那些使用者界面物件為基礎，計算一可能性；

接收語音輸入；及

以所計算出的可能性為基礎，判定該語音輸入是一聲音命令或是語音口述的一決定。

六、申請專利範圍

7. 如申請專利範圍第6項所述之方法，其中更包括下列步驟：

確認一焦距點在該使用者界面之外；及

以一隱含可能性為基礎，判定該語音輸入是一聲音命令或是語音口述的一決定。

8. 一種機器可讀取記錄媒體，載有一軟體程式，該軟體程式用以在一電子文件中搜索符合文字；其中該軟體程式至少包括複數個程式碼節，可由一機器來執行，以使該機器執行至少下列步驟：

確認一焦距點在一使用者界面上；

界定一周圍區域在該焦距點四周，該周圍區域包含一文字正文在一使用者界面物件範圍之內，該使用者界面物件係用以接收語音口述文字；

接收一聲音命令，以挑選在該電子文件中的指定文字；及

在該周圍區域所包含之文字正文中搜尋該指定文字的一符合文字，該搜尋係侷限於該周圍區域內之文字正文中。

9. 如申請專利範圍第8項所述之機器可讀取記錄媒體，其中更包括下列步驟：

如果在該搜尋步驟中沒找到該指定文字的一符合文字，擴展該周圍區域以包含該使用者界面的一額

六、申請專利範圍

外範圍，該額外區域包含有額外文字；及

在該額外文字中搜尋該指定文字的一符合文字，
該搜尋係侷限於該文字正文和該額外文字中。

10. 如申請專利範圍第9項所述之機器可讀取記錄媒體，其中上述之擴展步驟更包括下列步驟：

從該焦距點以一不變增加量向外的擴展該周圍區域。

11. 如申請專利範圍第9項所述之機器可讀取記錄媒體，其中上述之擴展步驟更包括下列步驟：

以緊接該文字正文的一不變數量的文字擴展該周圍區域。

12. 如申請專利範圍第9項所述之機器可讀取記錄媒體，其中上述之擴展步驟更包括下列步驟：

從該焦距點以一可變增加量向外的擴展該周圍區域。

13. 一種機器可讀取記錄媒體，載有一軟體程式，該軟體程式用以辨識一聲音命令的一實例和語音口述的一實例兩者；其中該軟體程式至少包括複數個程式碼節，可由一機器來執行，以使該機器執行至少下列步驟：

六、申請專利範圍

確認一焦距點在一使用者界面上；

界定一周圍區域在該焦距點四周；

在該周圍區域內確認使用者界面物件；

再確認在該已確認使用者界面物件之間那些使用者界面物件是用以接收語音口述正文，而那些使用者界面物件不是用以接收語音口述正文；

以已再次確認是用以接收語音口述正文的那些使用者界面物件與已再次確認不是用以接收語音口述正文的那些使用者界面物件為基礎，計算一可能性；

接收語音輸入；

以所計算出的可能性為基礎，判定該語音輸入是一聲音命令或是語音口述的一決定。

14. 如申請專利範圍第13項所述之機器可讀取記錄媒體，其中更包括下列步驟：

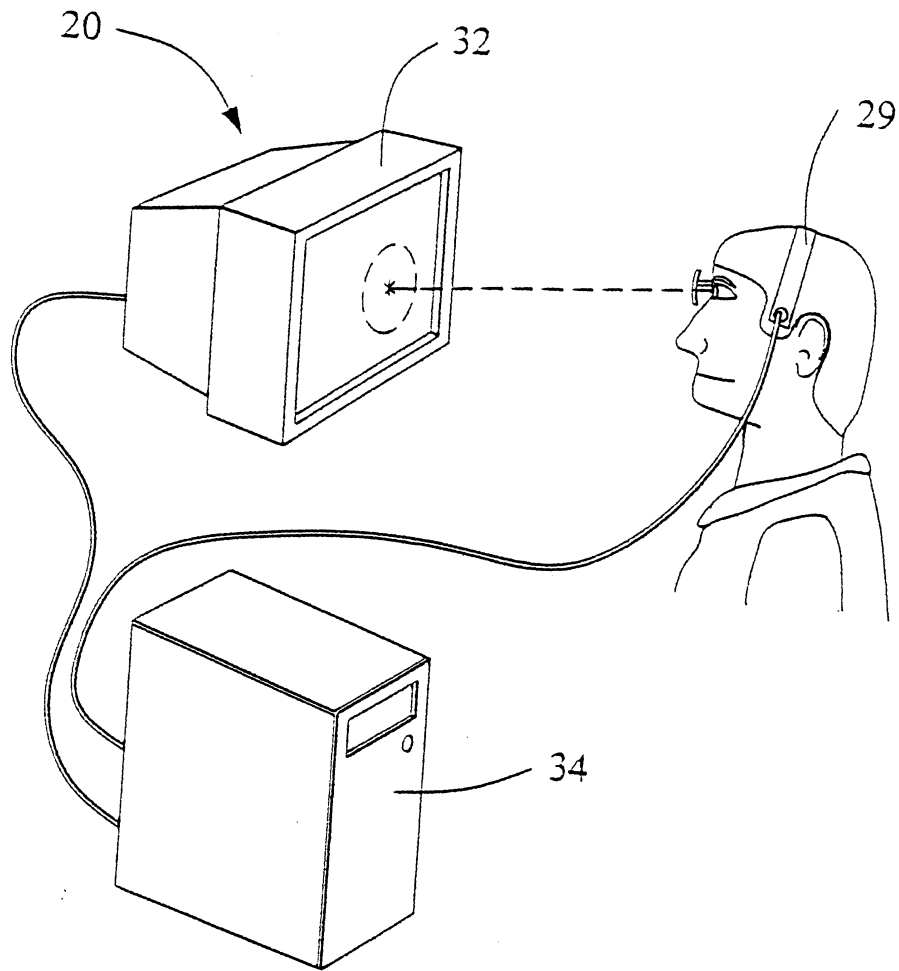
確認一焦距點在該使用者界面之外；及

以一隱含可能性為基礎，判定該語音輸入是一聲音命令或是語音口述的一決定。

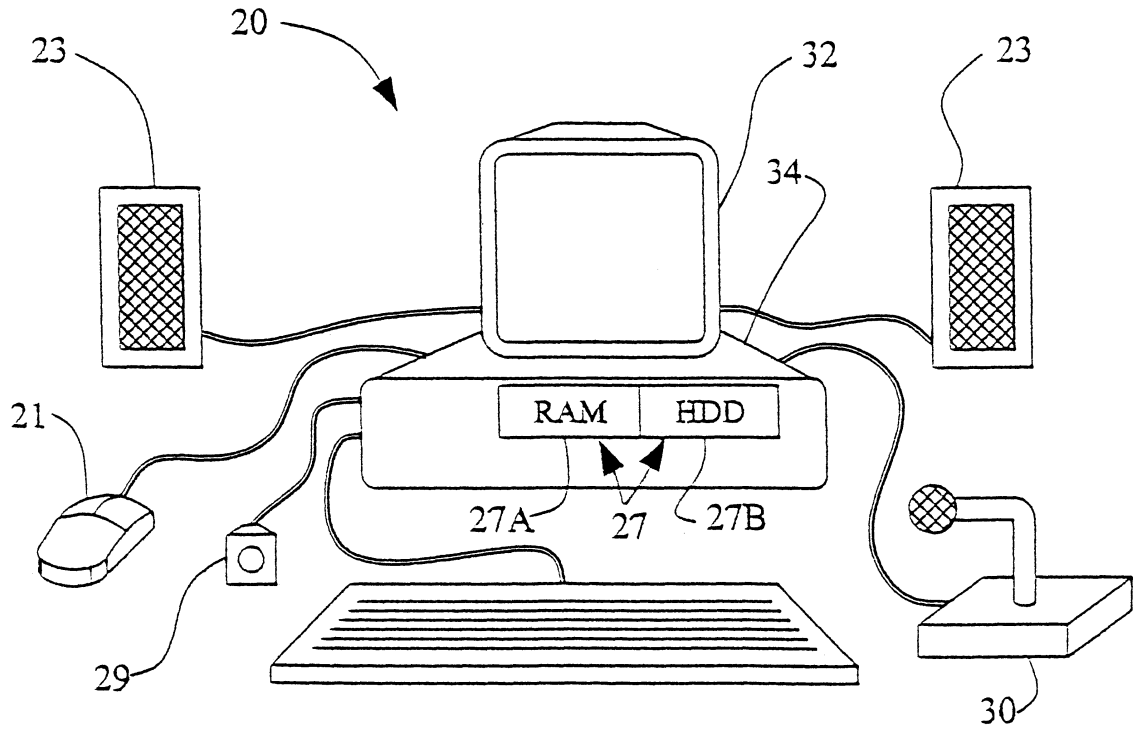
(請先閱讀背面之注意事項再填寫本頁)

訂

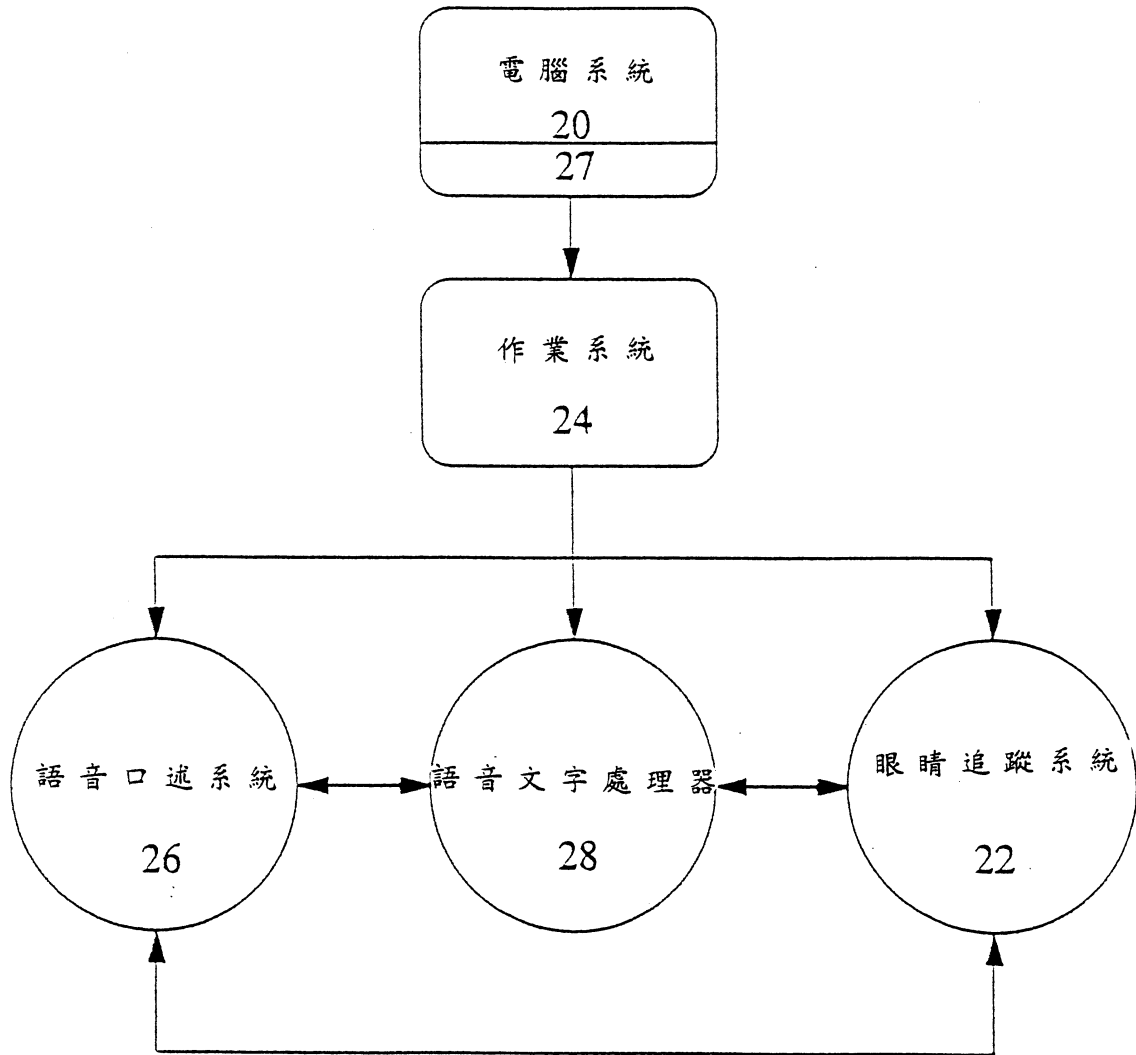
線



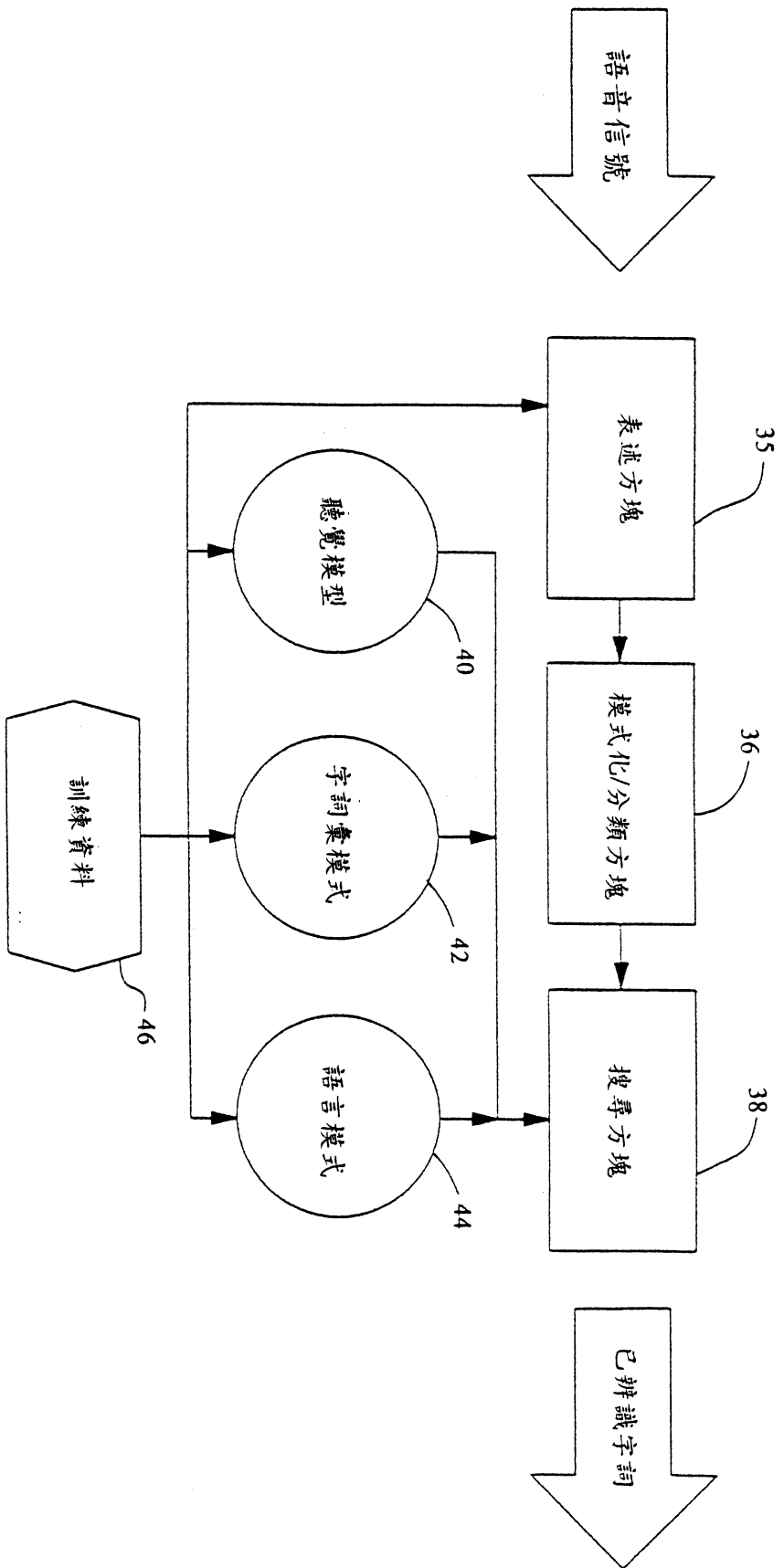
第 1 圖



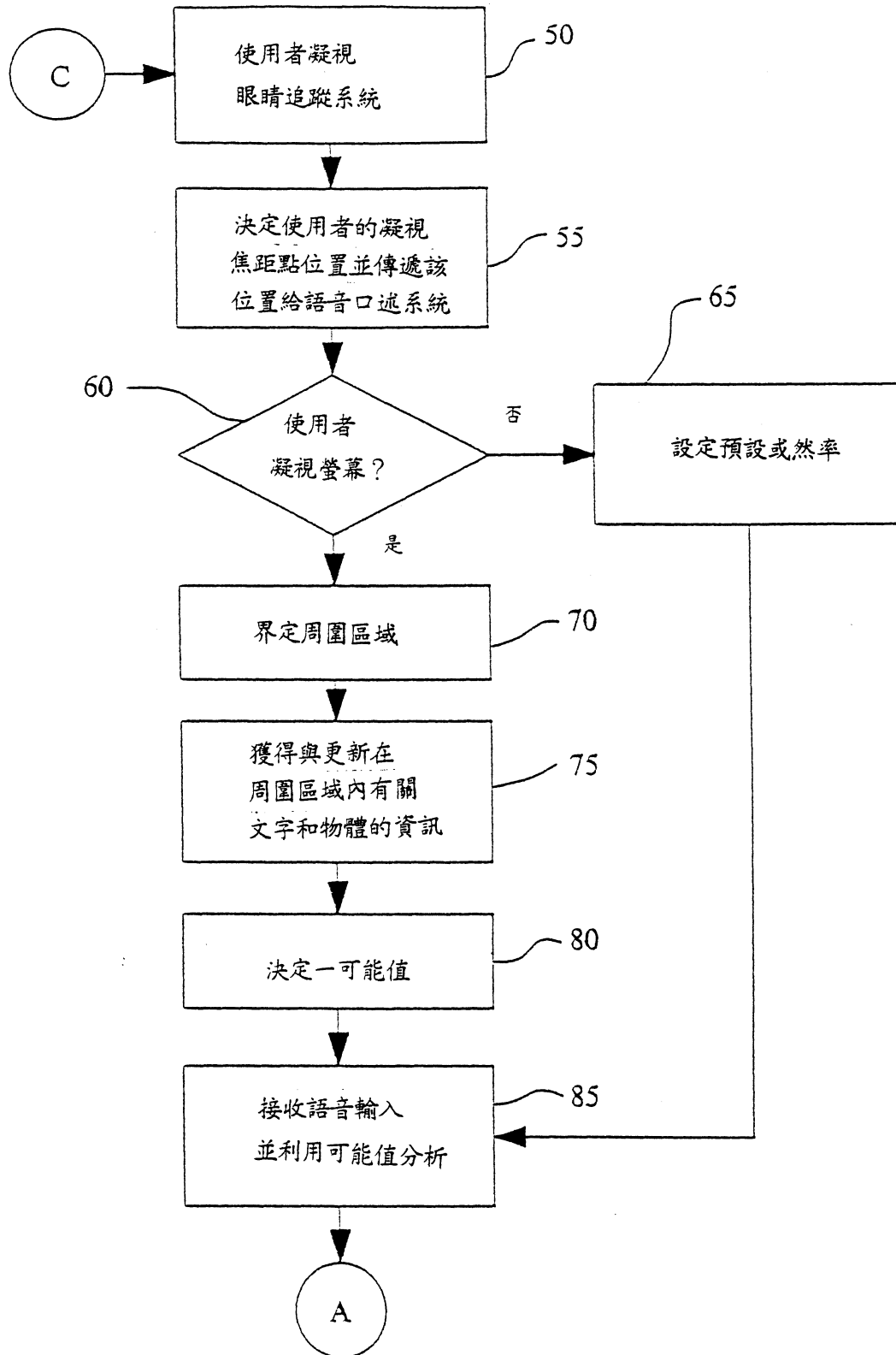
第 2 圖



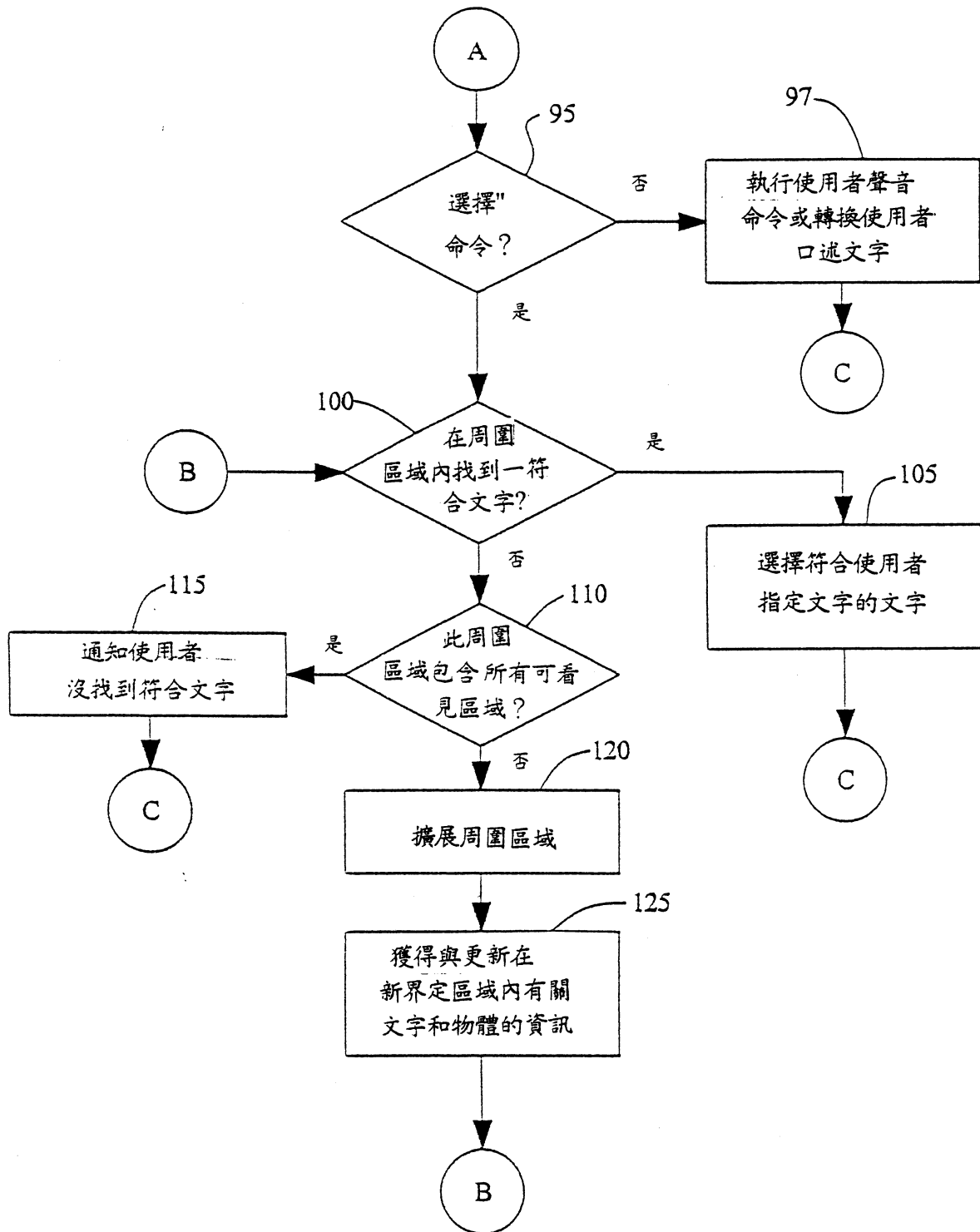
第 3 圖



第 4 圖



第 5A 圖



第 5B 圖