



(21)申請案號：099105182

(22)申請日：中華民國 99 (2010) 年 02 月 23 日

(51)Int. Cl. : G10L15/14 (2006.01)

(30)優先權：2009/03/30 美國 12/413,606

(71)申請人：微軟公司(美國) MICROSOFT CORPORATION (US)
美國

(72)發明人：大附克年 OHTSUKI, KATSUTOSHI (JP) ; 梅岡孝史 UMEOKA, TAKASHI (JP)

(74)代理人：蔡坤財；李世章

(56)參考文獻：

TW	283774	TW	200301460A
TW	200400488A	TW	201028997A
TW	201133470A	US	2011/0144991A1
WO	2010/033384A1		

審查人員：涂淑惠

申請專利範圍項數：22 項 圖式數：8 共 39 頁

(54)名稱

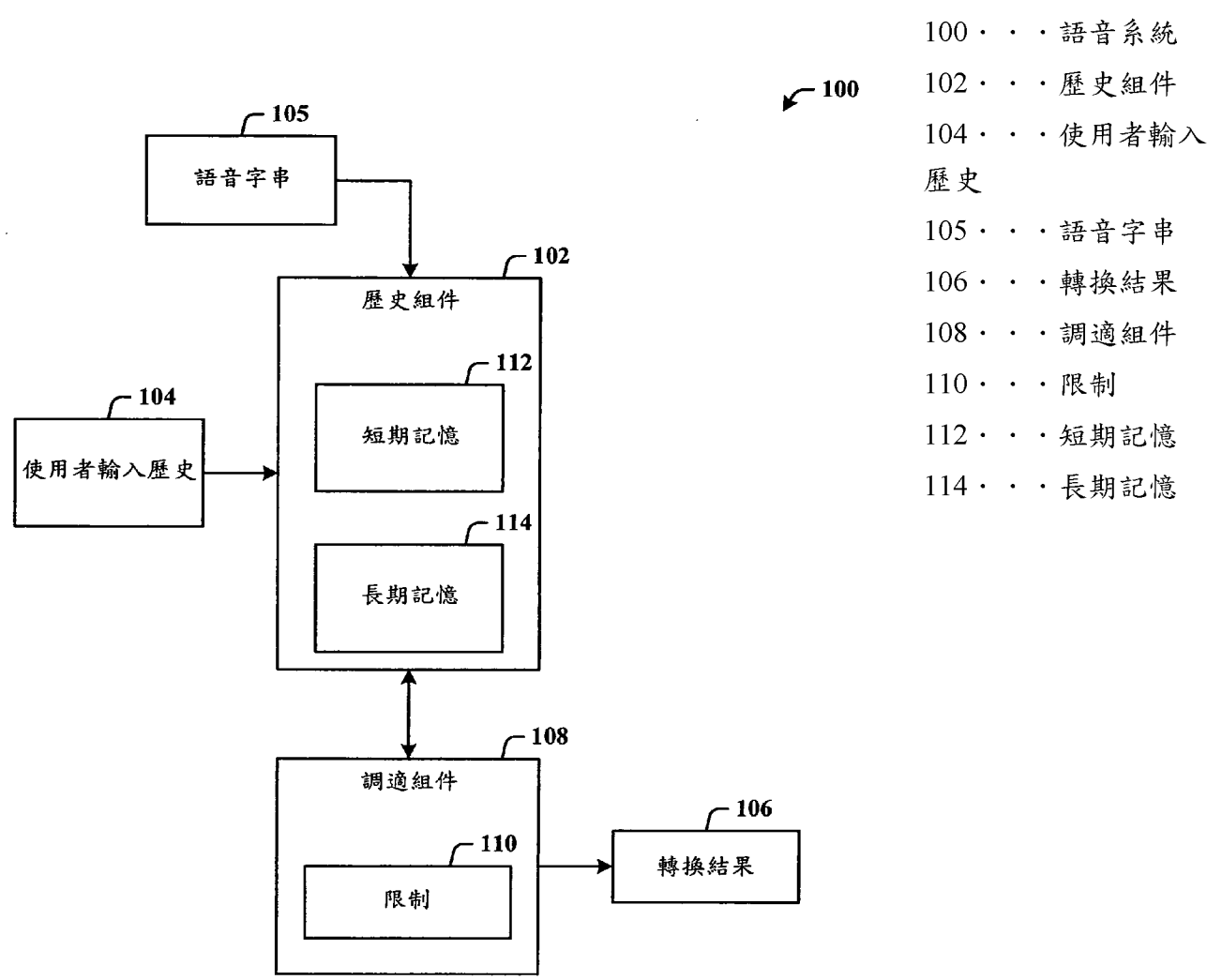
電腦實作的語音系統及方法

COMPUTER-IMPLEMENTED PHONETIC SYSTEM AND METHOD

(57)摘要

本發明揭示一種架構，其藉由對長期及短期記憶施加適當的限制而抑制了字詞會出乎預期的出現。快速調適亦藉由利用該限制來實現。該架構包括一歷史組件，用於藉由輸出轉換結果的一轉換程序，處理使用者輸入歷史來轉換一語音字串，還包括一調適組件，用於基於在該轉換程序期間施加上影響字詞出現的短期記憶之限制，來調適該轉換程序以合於該使用者輸入歷史。該架構基於前後文相關的機率差異(短期記憶)來執行機率增加，且基於字詞前文之頻率(長期記憶)來執行長期記憶與基線語言模型之間的動態線性內插。

Architecture that suppresses the unexpected appearance of words by applying appropriate restrictions to long-term and short-term memory. The quickness of adaptation is also realized by leveraging the restriction. The architecture includes a history component for processing user input history for conversion of a phonetic string by a conversion process that output conversion results, and an adaptation component for adapting the conversion process to the user input history based on restriction(s) applied to short-term memory that impacts word appearances during the conversion process. The architecture performs probability boosting based on context-dependent probability differences (short-term memory), and dynamic linear-interpolation between long-term memory and baseline language model based on frequency of preceding context of word (long-term memory).



- 100 . . . 語音系統
- 102 . . . 歷史組件
- 104 . . . 使用者輸入歷史
- 105 . . . 語音字串
- 106 . . . 轉換結果
- 108 . . . 調適組件
- 110 . . . 限制
- 112 . . . 短期記憶
- 114 . . . 長期記憶

第 1 圖

公告本

發明專利說明書

103年10月22日修正
劃線頁(本)

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫；惟已有申請案號者請填寫)

※ 申請案號：99105182

※ 申請日期：2010 年 2 月 23 日

※IPC 分類：G10L 15/14 (2006.01)

一、發明名稱：(中文/英文)

電腦實作的語音系統及方法

COMPUTER-IMPLEMENTED PHONETIC SYSTEM AND METHOD

二、中文發明摘要：

本發明揭示一種架構，其藉由對長期及短期記憶施加適當的限制而抑制了字詞會出乎預期的出現。快速調適亦藉由利用該限制來實現。該架構包括一歷史組件，用於藉由輸出轉換結果的一轉換程序，處理使用者輸入歷史來轉換一語音字串，還包括一調適組件，用於基於在該轉換程序期間施加到影響字詞出現的短期記憶之限制，來調適該轉換程序以合於該使用者輸入歷史。該架構基於前後文相關的機率差異(短期記憶)來執行機率增加，且基於字詞前文之頻率(長期記憶)來執行長期記憶與基線語言模型之間的動態線性內插。

三、英文發明摘要：

Architecture that suppresses the unexpected appearance of words by applying appropriate restrictions to long-term and short-term memory. The quickness of adaptation is also realized by leveraging the restriction. The architecture includes a history component for processing user input history for conversion of a phonetic string by a conversion

process that output conversion results, and an adaptation component for adapting the conversion process to the user input history based on restriction(s) applied to short-term memory that impacts word appearances during the conversion process. The architecture performs probability boosting based on context-dependent probability differences (short-term memory), and dynamic linear-interpolation between long-term memory and baseline language model based on frequency of preceding context of word (long-term memory).

[S]
[S]

四、指定代表圖：

(一)本案指定代表圖為：第(1)圖。

(二)本代表圖之元件符號簡單說明：

100 語音系統

108 調適組件

102 歷史組件

110 限制

104 使用者輸入歷史

112 短期記憶

105 語音字串

114 長期記憶

106 轉換結果

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

六、發明說明：

103年10月22日修正頁(本)

P4-32

【發明所屬之技術領域】

本發明關於統計語言模型的調適性。

【先前技術】

輸入方法可用於將語音字串(閱讀)轉換成東亞語言(例如中文、韓文及日文)的顯示字元，並亦可處理筆劃，如繁體中文的字元。由於同音異義字及多種可能的字詞分段，在轉換時會出現混淆。一種輸入方法嘗試要基於一般(例如基線、預設)語言模型及使用者輸入歷史，來解決這種混淆性。對於該使用者輸入歷史的調適可用數種方式進行，例如短期記憶與長期記憶。短期記憶對應於快速調適，而長期記憶對應於調適的穩定性。轉換結果藉由自該短期與長期記憶加入資訊到該一般語言模型來判定。

短期記憶可藉由基於先前使用者選擇的字詞(使用者輸入歷史)增加字詞分數或改變字詞等級來實作。但是，一些字詞在被用過之後並未很快出現，而一些字詞在被用過之後出乎預期地在無法接受的前後文中出現。長期記憶可藉由累積使用者輸入歷史來實作。但是，雖然有利用長期記憶，一些字詞仍出乎預期地在無法接受的前

後文中出現。

【發明內容】

以下將提供一簡化的發明內容，藉以提供對於此處一些創新具體實施例之基本瞭解。此發明內容並非詳盡的概述，且其並非要辨識關鍵/重要元件或界定其範圍。其唯一的目的是要以一簡化的型式做為一序言來呈現一些觀念，而更為詳細的說明將在稍後呈現。

所揭示的架構藉由施加適當的限制到長期及短期記憶，而抑制了字詞的出乎預期的出現。快速調適亦藉由利用該限制來實現。

該架構包括一歷史組件，用於藉由輸出轉換結果的一轉換程序，處理使用者輸入歷史來轉換一語音字串，及一調適組件，用於基於在該轉換程序期間施加到影響字詞出現的短期記憶之限制，來調適該轉換程序以合於該使用者輸入歷史。該架構基於前後文相關的機率差異(短期記憶)而執行機率增加，且基於字詞前文之頻率(長期記憶)而執行長期記憶與基線語言模型之間的動態線性內插。

為了完成前述及相關的目的，某些例示性態樣將在此處配合以下描述及附加圖面來說明。這些態樣指明多種方式，其中此處所述的該等原理可被實現，且所有態樣

及其同等者皆係在所主張的標的之範圍內。其它好處及創新特徵將可配合該等圖面由以下的實施方式來更加瞭解。

【實施方式】

雖然既有的語音系統之轉換準確性在一般狀況下很高，因為一特定使用者的語言空間不同於該一般性空間，使用者仍會失望。此種現象對於個人名字而言特別明顯，且該等表述喜好自然地根據使用者而改變，因此不能夠由該一般性語言模型處理。

所揭示的架構為一自我調整技術，其中在使用一短期間(例如 2-3 周)的產品之後，該使用者不再需要打開一候選清單。再者，所揭示的自我調整技術改善了一使用者的工作效能。該架構基於前後文相關的機率差異(短期記憶)而執行機率增加，且基於字詞前文之頻率(長期記憶)而執行長期記憶與基線語言模型之間的動態線性內插。

現在將參考圖面，其中類似的參照編號用於參照到所有類似的元件。在以下的說明中，為了解釋起見，提出許多特定細節，藉以提供對其完整瞭解。但是明顯地該等創新具體實施例可以不使用這些特定細節來實施。在其它實例中，熟知的結構及裝置以方塊圖形式顯示，藉此用於進行其說明。本說明書之意圖係要涵蓋所有落在

所主張標的之精神及範圍內的修正、同等者及變化。

第 1 圖為根據所揭示的架構之電腦實作的語音系統 100。系統 100 包括歷史組件 102，用於藉由輸出轉換結果 106 的一轉換程序，處理使用者輸入歷史 104 來轉換語音字串 105，及調適組件 108，用於基於在該轉換程序期間應用到影響字詞出現的短期記憶 112 之限制 110，來調適該轉換程序以合於該使用者輸入歷史 104。

調適組件 108 基於長期記憶 114 執行長期記憶 114 與一基線語言模型之間的動態線性內插。該等限制 110 在當一字詞並非一候選清單的一第一候選者時，增加該字詞的機率。施加於短期記憶 112 的該等限制 110 利用一前後文敏感性的短期記憶雙連字串 (bigram) 機率。施加於短期記憶 112 之該等限制 110 基於一字詞與在一句子中該字詞的前後文增加一機率。該前後文包括相對於在該句子中該字詞之前文與後文。調適組件 108 包括一學習演算法，其基於一候選清單之一第一候選者與該候選清單的一選擇的候選者之間的差異，執行旗標學習 (flag-learning)，且在一下一轉換程序中移動該選擇的候選者到一第一轉換結果位置。

第 2 圖所示為包括第 1 圖之語音系統 100 之額外態樣的系統 200。系統 200 包括歷史組件 102，用於處理使用者輸入歷史 104 來藉由轉換程序 204 轉換語音字串 105，

還包括調適組件 108，用於基於在轉換程序 204 期間施加於影響字詞出現的短期記憶 112 之該等限制 110，調適轉換程序 204 以合於使用者輸入歷史 104。

調適組件 108 基於長期記憶 114 執行長期記憶 114 與基線語言模型 208 之間的動態線性內插。該等限制 110 在當一字詞並非一候選清單的一第一候選者時，增加該字詞的機率。施加於短期記憶 112 的該等限制 110，利用一前後文敏感性的短期記憶雙連字串機率。施加於短期記憶 112 之該等限制 110 基於一字詞與在一句子中該字詞的前後文增加一機率。該前後文包括相對於在該句子中該字詞之前文與後文。調適組件 108 包括一學習演算法，其基於一候選清單之一第一候選者與該候選清單的一選擇的候選者之間的差異，執行旗標學習，且在下一轉換程序中移動該選擇的候選者到一第一轉換結果位置。

系統 200 另包含限制組件 206，用於藉由基於前後文相關的機率差異增加一機率來施加該等限制 110。限制組件 206 亦可藉由基於一前後文相關的機率差異增加一機率，施加一或多個該等限制 110 到長期記憶 114。

換言之，語音系統 200 包括歷史組件 102，用於處理使用者輸入歷史 104，用以在轉換程序 204 期間轉換語音字串 105，還包括限制組件 206，用於在轉換程序 204

期間施加一或多個該等限制 110 到使用者輸入歷史 104。歷史 104 包括短期記憶 112 與長期記憶 114。系統 200 亦包括調適組件 108，用於基於該等限制 110 調適轉換程序 204 以合於使用者輸入歷史 104。

限制組件 206 施加一或多個該等限制 110 到短期記憶 112。所施加的該等限制 110 利用一前後文敏感性短期記憶雙連字串機率，及一或多個限制 110 到長期記憶 114，其基於一前後文相關的機率差異增加一機率。調適組件 108 基於長期記憶 114 執行長期記憶 114 與基線語言模型 208 之間的動態線性內插。該等限制 110 在當一字詞並非一候選清單的一第一候選者時增加該字詞的機率。施加於短期記憶 112 之該等限制 110，基於一字詞與在一句子中該字詞的前後文增加一機率。該前後文包括相對於在該句子中該字詞之前文與後文。

下述為用於統計語言模型的快速與穩定調適之該等運算的實施方式。

一輸入語音字串的輸入方法轉換結果可由以下的機率所決定：

$$P(W) = P(W_1 | \langle S \rangle) \cdot P(W_2 | W_1) \cdot P(W_3 | W_2) \dots P(W_N | W_{N-1}) \cdot P(\langle /S \rangle | W_N)$$

其中 W 為一句子，其中包括一字詞序列 $\{w_1, w_2, w_3, \dots\}$ ，

$w_{N-1}, w_N\}$ ，且 $\langle s \rangle$ 與 $\langle /s \rangle$ 分別為句子開始與句子結束的符號。該等式係用於雙連字串模型，但可用三連字串或更高階的 n 連字串模型來表示。

由於同音異義字及字詞分段化的混淆性，一輸入語音字串可能有許多種可能的字詞序列 W 。

一最有可能的候選句子被選擇做為一轉換結果。

$$\hat{W} = \arg \max_w P(W)$$

每個字詞的機率可被定義成，

$$P(w_n | w_{n-1}) = \alpha \cdot P_{baseline}(w_n | w_{n-1}) + \beta \cdot P_{lrm}(w_n | w_{n-1}) + \delta \cdot P_{stm}(w_n | w_{n-1})$$

其中 α , β , 與 δ 為線性內插係數，其加總成一 ($\alpha + \beta + \delta = 1$)， $P_{baseline}(w_k | w_{k-1})$ 為由該訓練字詞資料庫所估計的一基線雙連字串機率 (當第一次使用該輸入方法時，僅有此機率具有一數值)， $P_{lrm}(w_n | w_{n-1})$ 為該長期記憶的該雙連字串機率，及 $P_{stm}(w_n | w_{n-1})$ 為該短期記憶的雙連字串機率。該長期記憶的雙連字串機率可由該使用者輸入歷史計算，如下式。

$$P_{lrm}(w_n | w_{n-1}) = \frac{C_{user}(w_{n-1}, w_n)}{C_{user}(w_n)}$$

其中 $C_{user}(w_n)$ 為該使用者使用字詞 w_n 的次數，且 $C_{user}(w_{n-1}, w_n)$ 為該使用者使用該字詞序列 w_{n-1}, w_n 的次數。

短期記憶之雙連字串機率 $P_{stm}(w_n | w_{n-1})$ 在當字詞並非該結果的第一候選者時增加該字詞的機率，但使用者從該候選清單中選擇該字詞。

$$P_{stm}(w_n | w_{n-1}) = \frac{C_{user-sel}(w_{n-1}, w_n)}{M}$$

其中 $C_{user-sel}(w_{n-1}, w_n)$ 為使用者從該候選清單中選擇該字詞序列 w_{n-1}, w_n 的次數，而 M 為選擇的最高次數。請注意 $C_{user-sel}(\cdot)$ 不能超過 M 。

以上的公式可由指數化來一般化，如下式表示：

$$P_{stm}(w_n | w_{n-1}) = \left(\frac{C_{user-sel}(w_{n-1}, w_n)}{M} \right)^k$$

以下為長期記憶的額外說明。一字詞 w_n 的線性內插加權 α 與 β 根據 $C_{user}(w_{n-1})$ 而改變。此代表該等加權根據先前字詞而有所不同。

當 $C_{user}(w_{n-1})$ 充份大時，即定義並使用該等目標加權

α_{target} 與 β_{target} 。 w_n 實際的加權 α 與 β 可依下式計算，

$$\begin{aligned}\beta &= \omega \cdot \beta_{\text{target}} \\ \alpha &= \alpha_{\text{target}} + \beta_{\text{target}} - \beta = \alpha_{\text{target}} + (1 - \omega) \cdot \beta_{\text{target}} \\ \omega &= \begin{cases} \frac{1}{t} \cdot C_{\text{user}}(w_{n-1}) & \text{if } C_{\text{user}}(w_{n-1}) < t \\ 1 & \text{Otherwise} \end{cases}\end{aligned}$$

第 3 圖為該加權轉換之示意圖 300。示意圖 300 顯示短期記憶 δ 、長期記憶 β 與基線 α ，與標示為 β_{target} 的長期記憶及標示為 α_{target} 的基線之相對垂直範圍分段。示意圖 300 代表當該字詞使用的次數增加時，於時間 t ，該長期記憶的加權到達該 β_{target} 。

當 $C_{\text{user}}(w_{n-1})$ 很小時，該長期雙連字串機率趨向於高，並造成字詞的出乎預期之出現。但是，此加權調整會抑制此類副作用。

以下為短期記憶的額外說明。其使用兩種方法，其可獨立或組合使用：一短期記憶雙連字串機率之前後文敏感性使用，及根據該機率差異來增加機率。

對於第一種方法，一短期記憶雙連字串機率之前後文敏感性使用，當後續的字詞序列之選擇的次數為零時，該機率即被視為零。

$$P_{\text{stm}}(w_n | w_{n-1}) = \begin{cases} \frac{C_{\text{user-sel}}(w_{n-1}, w_n)}{M}, & \text{if } C_{\text{user-sel}}(w_n, w_{n+1}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

類似的結果可使用先前的字詞序列而獲得。

$$P_{\text{stm}}(w_n | w_{n-1}) = \begin{cases} \frac{C_{\text{user-sel}}(w_{n-1}, w_n)}{M}, & \text{if } C_{\text{user-sel}}(w_{n-2}, w_{n-1}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

這些條件可能依該等字詞的語音部份 (Part-of-speech, 「POS」) 而改變。基於這些條件，該機率增加係根據該前後文，且先前選擇的字詞之出乎預期的出現可被抑制。

至於第二種方法，根據機率差異來增加機率，一個一個增加 $C_{\text{user-sel}}(w_{n-1}, w_n)$ 對一些字詞而言可能不足，但對於其它字詞則太多。 $C_{\text{user-sel}}(w_{n-1}, w_n)$ 的適當增加數目係根據該字詞與該字詞所處的前後文。

該使用者由該候選者選擇該字詞，係因為該句子包括該字詞之機率低於另一個句子包括該字詞之機率。因此，為了在下一次取得該字詞，該句子包括該字詞的該機率必須高於另一個句子包括該字詞的機率(在先前轉換中第一個句子)。

第 4 圖所示為一快取加權轉換之示意圖 400。在另一具體實施例中，一快取加權轉換係使用一線性函數來提供，且該快取加權僅用於該雙連字串快取(雙連快取)。該雙連字串快取加權根據該先前字詞的一單元字串快取(單一快取)量。此代表雙連字串快取機率 $P_{\text{bicache}}(w_i | w_{i-1})$

之加權係根據 $C_{unicache}(w_{i-1})$ 。

$$\omega = \begin{cases} \frac{1}{t} \cdot \text{CacheAmount}(w_{i-1}) & \text{if CacheAmount} < t \\ 1 & \text{Otherwise} \end{cases}$$

該旗標加權 $\delta + \varepsilon$ 為固定。該單元字串快取之加權亦為固定，但藉由該稍早的快取加入一偏移值到該單元字串快取總數來降低該等副作用。

$$P_{unicache}(w_i) = P_{unicache-smooth}(w_i) = \frac{C(w_i)}{\sum_k C(w_k) + \text{UniCacheOffset}}$$

$$P_{unicache-raw}(w_i) = \frac{C(w_i)}{\sum_k C(w_k)}$$

旗標學習係視該等機率差異而定。一雙連字串旗標的增加程度，係根據該第一候選者與該選擇的候選者之間所估計的差異量而改變。該選擇的候選者在該周遭前後文為相同時即成為該第一後續轉換結果。

以下的案例可被考慮，且以下的該演算法可涵蓋所有案例。

Case #1: $\{w_a, w_b, w_c\}$ after conversion $\rightarrow \{w_a, w_x, w_c\}$ after editing

Case #2: $\{w_a, w_{b1} \dots w_{bm}, w_c\}$ after conversion $\rightarrow \{w_a, w_x, w_c\}$ after editing

Case #3: $\{w_a, w_b, w_c\}$ after conversion $\rightarrow \{w_a, w_{x1} \dots w_{xn}, w_c\}$ after editing

Case #4: $\{w_a, w_{b1} \dots w_{bm}, w_c\}$ after conversion $\rightarrow \{w_a, w_{x1} \dots w_{xn}, w_c\}$ after editing

提供以下的定義。

$P(w_b | w_a)$ 為學習之前的該字詞雙連字串機率，其包括基線、快取及旗標機率。

$$P(w_b | w_a) = \alpha \cdot P_{\text{baseline}}(w_b | w_a) + \beta \cdot P_{\text{cache}}(w_b | w_a) + \gamma \cdot P_{\text{cache}}(w_b) + \delta \cdot P_{\text{flag}}(w_b | w_a) + \varepsilon \cdot P_{\text{flag}}(w_b)$$

$P_L(w_b | w_a)$ 為學習之後的該字詞雙連字串機率。快取機率的改變在此為了簡化起見而被忽略，且在學習之後僅有該等旗標機率會改變。

$$P_L(w_b | w_a) = \alpha \cdot P_{\text{baseline}}(w_b | w_a) + \beta \cdot P_{\text{cache}}(w_b | w_a) + \gamma \cdot P_{\text{cache}}(w_b) + \delta \cdot P_{\text{flag}}(w_b | w_a)_{-1} + \varepsilon \cdot P_{\text{flag}}(w_b)_{-1}$$

候選字詞之該等旗標數目在學習之後被減一，當一使用者從該候選清單中選擇另一候選者時，該等候選字詞為該等第一候選者。

$$P_L(w_x | w_a) = \alpha \cdot P_{\text{baseline}}(w_x | w_a) + \beta \cdot P_{\text{cache}}(w_x | w_a) + \gamma \cdot P_{\text{cache}}(w_x) + \delta \cdot P_{\text{flag}}(w_x | w_a)_{+d} + \varepsilon \cdot P_{\text{flag}}(w_x)_{+1}$$

該等相對應候選字詞的單元字串旗標數目被加一，該等候選字詞係從該候選清單中所選擇。從該候選清單中所選擇的該等相對應候選字詞的單元字串旗標數目被增加，增加數目有待決定。

該演算法方面，在學習之前，該第一候選者與該選擇的候選者之間的大小關係如下式，

$$P(w_b|w_a) \cdot P(w_c|w_b) > P(w_x|w_a) \cdot P(w_c|w_x)$$

在學習之後的大小成為，

$$P_L(w_b|w_a) \cdot P_L(w_c|w_b) < P_L(w_x|w_a) \cdot P_L(w_c|w_x)$$

該等機率 $P_L(w_b | w_a) \cdot P_L(w_c | w_b)$ 與 $P(w_x | w_a) \cdot P(w_c | w_x)$ 為已知，並需要知道如何計算 $P_L(w_x | w_a) \cdot P_L(w_c | w_x)$ 。

藉由學習之機率的變化可表示成一指數(或乘冪)。

$$P_L(w_x|w_a) \cdot P_L(w_c|w_x) = [P(w_x|w_a) \cdot P(w_c|w_x)]^\varphi$$

因此，

$$P_L(w_b|w_a) \cdot P_L(w_c|w_b) < [P(w_x|w_a) \cdot P(w_c|w_x)]^\varphi$$

則 φ 可如下式計算，

$$\varphi > \frac{\log[P(w_x|w_a) \cdot P(w_c|w_x)]}{\log[P_L(w_b|w_a) \cdot P_L(w_c|w_b)]}$$

現在由 φ 計算 $P_{\text{flag}}(w_x | w_a)_{+d}$ 。

$$P_L(w_x|w_a) \cdot P_L(w_c|w_x) = [P(w_x|w_a) \cdot P(w_c|w_x)]^\varphi = P(w_x|w_a)^\varphi \cdot P(w_c|w_x)^\varphi$$

如果可滿足以下的等式，即可滿足以上的等式。

$$P_L(w_x|w_a) = P(w_x|w_a)^\varphi \text{ and } P_L(w_c|w_x) = P(w_c|w_x)^\varphi$$

考慮 $P_L(w_x | w_a) = P(w_x | w_a)^\varphi$ 。

$$P(w_x|w_a) = \alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) \\ + \delta \cdot P_{\text{flag}}(w_x|w_a) + \varepsilon \cdot P_{\text{flag}}(w_x)$$

$$P_L(w_x|w_a) = \alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) \\ + \delta \cdot P_{\text{flag}}(w_x|w_a)_{+d} + \varepsilon \cdot P_{\text{flag}}(w_x)_{+1}$$

$$\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) + \delta \cdot P_{\text{flag}}(w_x|w_a)_{+d} + \varepsilon \\ \cdot P_{\text{flag}}(w_x)_{+1} \\ = [\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) \\ + \delta \cdot P_{\text{flag}}(w_x|w_a) + \varepsilon \cdot P_{\text{flag}}(w_x)]^\varphi$$

$$\delta \cdot P_{\text{flag}}(w_x|w_a)_{+d} \\ = [\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) \\ + \delta \cdot P_{\text{flag}}(w_x|w_a) + \varepsilon \cdot P_{\text{flag}}(w_x)]^\varphi \\ - [\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) + \varepsilon \\ \cdot P_{\text{flag}}(w_x)_{+1}]$$

$$P_{\text{flag}}(w_x|w_a)_{+d} \\ = \frac{[\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) + \delta \cdot P_{\text{flag}}(w_x|w_a) + \varepsilon \cdot P_{\text{flag}}(w_x)]^\varphi}{\delta} \\ - \frac{[\alpha \cdot P_{\text{baseline}}(w_x|w_a) + \beta \cdot P_{\text{cache}}(w_x|w_a) + \gamma \cdot P_{\text{cache}}(w_x) + \varepsilon \cdot P_{\text{flag}}(w_x)_{+1}]}{\delta}$$

增加量 d 可基於機率的差異來計算。

$$P_{\text{flag}}(w_x|w_a)_{+d} - P_{\text{flag}}(w_x|w_a)$$

該旗標學習機率可藉由對應於該旗標數目來預備。例如該旗標數目的範圍可為 8, 16 或 32。當數目愈高時，此演算法即運作地更精確。

以下包括的為一系列的流程圖，其代表用於執行所揭示的架構之創新態樣的示例性方法。為了簡化說明起見，此處顯示一或多種方法，例如以流程圖的型式，並以一系列的動作來顯示及描述，其將可瞭解到該主題發明並不限於動作的順序，根據本發明，一些動作可與此處所顯示及描述之不同順序下發生，及/或與其它動作同步發生。例如，本技藝專業人士將可瞭解到一方法可另外表示成一系列的交互關連的狀態或事件，例如狀態圖。再者，對於一種創新實施而言，在一方法中例示的全部動作並非都一定是所需的。

第 5 圖所示為一電腦實作的語音方法。在 500 中，該使用者輸入歷史在一轉換程序期間被處理來轉換一語音字串。在 502 中，於該轉換程序期間施加限制到該使用者輸入歷史，該歷史包括短期記憶與長期記憶。在 504 中，該轉換程序基於該等限制被調適以合於該使用者輸

入歷史。

第 6 圖所示為第 5 圖之方法的額外態樣。在 600 中，施加一限制來基於前後文相關的機率差異來增加一機率。在 602 中，基於該長期記憶，在長期記憶與一基線語言模型之間執行動態線性內插。在 604 中，當該字詞並非一候選清單之一第一候選者時，一字詞的機率被增加。

第 7 圖所示為第 5 圖之方法的額外態樣。在 700 中，施加一限制到該短期記憶，其基於一字詞與在一句子中該字詞的前後文增加一機率。在 702 中，基於一候選清單的一第一候選者與該候選清單的一選擇的候選者之間的差異，執行旗標學習。在 704 中，在一下一轉換程序中移動該選擇的候選者到一第一轉換結果位置。

如本申請案中所使用的名詞「組件」及「系統」係要代表一電腦相關的實體，其可為硬體、硬體及軟體的組合、軟體或執行中的軟體。例如，一組件可為(但不限於)在一處理器上運行的一程序、一處理器、一硬碟機、多重儲存驅動器(光學及/或磁性儲存媒體)、一物件、一可執行程式、執行的執行緒、一程式及/或一電腦。藉由例示，在一伺服器上執行的一應用與該伺服器可為一組件。一或多個組件可存在於一程序及/或一執行緒之中，且一組件可位在一電腦上及/或分散在兩個以上的電腦

之間。在此處使用的用語「示例性」係代表做為一示例、實例或例示。在此所述之任何態樣或設計做為「示例性」者並不需要被視為比其它態樣或設計要較佳或較有利。

現在請參照第 8 圖，所示為根據所揭示的架構用於對於一統計語言模型執行快速及穩定調適的運算系統 800 之方塊圖。為了提供其多種態樣的額外內容，第 8 圖及以下的討論係要提供一適當運算系統 800 之一簡短概略性的描述，其中可以實作多種態樣。當以上的說明係由可在一或多部電腦上運轉的電腦可執行指令的一般性內文中做說明時，本技藝專業人士將可瞭解到一創新具體實施例亦可結合其它程式模組及/或硬體及軟體的組合來實施。

用於實作多種態樣之運算系統 800 包括電腦 802，其具有處理單元 804、系統記憶體 806 及系統匯流排 808。處理單元 804 可為多種商用處理器之任一種，例如單一處理器、多處理器、單一核心單元與多核心單元。再者，本技藝專業人士將可瞭解到本發明方法可利用其它電腦系統組態來實施，其中包括迷你級電腦、主機級電腦、以及個人電腦(例如桌上型、膝上型等)、掌上型運算裝置、以微處理器為主或可程式化的消費性電子產品及類似者，其每一個皆可以在運作上結合於一或多個關聯的裝置。

系統記憶體 806 可包括揮發性(Volatile,「VOL」)記憶體 810(例如隨機存取記憶體(Random access memory,「RAM」)及非揮發性記憶體(Non-volatile,「NON-VOL」)812(例如 ROM, EPROM, EEPROM 等)。

一基本輸入/輸出系統(Basic input/output system,「BIOS」)可儲存在非揮發性記憶體 812 中,並包括該等基本例式,其可在電腦 802 內的組件之間進行資料與信號之傳遞,例如在開機期間。揮發性記憶體 810 亦可包括一高速 RAM,例如用於快取資料的靜態 RAM。

系統匯流排 808 提供系統組件的介面,其包括但不限於記憶體子系統 806 到處理單元 804。系統匯流排 808 可為數種匯流排結構中任何一種,其另可互連到一記憶體匯流排(可利用一記憶體控制器,也可不用),及一周邊匯流排(例如 PCI, PCIE, AGP, LPC 等),其可使用多種商用匯流排架構之任何一種。

電腦 802 另包括儲存子系統 814 及儲存介面 816,用於連接儲存子系統 814 到系統匯流排 808 及其它想要的電腦組件。儲存子系統 814 可包括像是一或多個硬碟機(Hard disk drive,「HDD」)、一磁性軟碟機(Floppy disk drive,「FDD」)、及/或光碟儲存驅動器(例如 CD-ROM 驅動器、DVD 驅動器)。儲存介面 816 可包括介面技術,例如像是增強型整合磁碟電路(Enhanced integrated

drive electronics,「EIDE」)、進階技術附加裝置(Advanced technology attachment,「ATA」)、序列進階技術附加裝置(Serial ATA,「SATA」)及 IEEE 1394 標準。

一或多個程式及資料可儲存在記憶體子系統 806 中，可移除式記憶體子系統 818(例如來自快閃碟型式技術)，及/或儲存子系統 814(例如光學、磁性、固態)，其包括作業系統 820、一或多個應用程式 822、其它程式模組 824 及程式資料 826。

一或多個應用程式 822、其它程式模組 824 及程式資料 826 例如可包括第 1 圖的系統 100 與組件，第 2 圖的系統 200 與組件，示意圖 300 與 400 所代表的關係，由第 5-7 圖之流程圖所代表的方法。

概言之，程式包括例式、方法、資料結構及其它軟體組件等，其可執行特殊工作或實作特定的抽象資料型態。作業系統 820、應用程式 822、模組 824 及/或資料 826 之所有或部份亦可快取在記憶體中，例如像是非揮發性記憶體 810。其要瞭解到所揭示的架構可用多種商用作業系統或作業系統之組合(例如做為虛擬機器)來實作。

儲存子系統 814 與記憶體子系統(806 及 818)做為資料、資料結構、電腦可執行指令等之揮發性與非揮發性儲存的電腦可讀取媒體。電腦可讀取媒體可為任何可由

電腦 802 存取的可用媒體，其包括揮發性與非揮發性媒體，可移除與不可移除式媒體。對於電腦 802，該媒體可用一適當的數位格式容納任何資料的儲存。本技藝專業人士應可瞭解到可利用其它種類的電腦可讀取媒體，例如 Zip 驅動器、磁帶、快閃記憶卡、卡匣及類似者，用於儲存電腦可執行指令來執行所揭示架構之創新方法。

一使用者可使用外部使用者輸入裝置 828(例如鍵盤與滑鼠)來與電腦 802、程式與資料進行互動。其它外部使用者輸入裝置 828 可包括一麥克風、一紅外線(Infrared, 「IR」)遙控器、搖桿、遊戲手把、相機辨識系統、光筆、觸控螢幕、姿勢系統(例如眼睛移動、頭部移動等)，及/或類似者。該使用者可使用電路板上使用者輸入裝置 830(例如觸控板、麥克風、鍵盤等)與電腦 802、程式及資料進行互動，其中電腦 802 例如可為一攜帶式電腦。這些及其它輸入裝置透過系統匯流排 808 經由輸入/輸出(I/O)裝置介面 832 連接至處理單元 804，但可由其它介面連接，例如並列埠、IEEE 1394 標準序列埠、遊戲埠、USB 埠、IR 介面等。I/O 裝置介面 832 亦可使用輸出周邊 834，例如印表機、音訊裝置、相機裝置等等，例如音效卡及/或電路板上音訊處理能力。

一或多個繪圖介面 836(亦通常稱之為繪圖處理單元

(Graphics processing unit,「GPU」)提供電腦 802 與外部顯示器 838(例如 LCD, 電漿)及/或電路板上顯示器 840(例如用於攜帶式電腦)之間的繪圖與視訊信號。繪圖介面 836 亦可製造成電腦系統電路板上的一部份。

電腦 802 經由一有線/無線通訊子系統 842 使用邏輯連接到一或多個網路及/或其它電腦在一網路化環境(例如 IP)中操作。其它的電腦可包括工作站、伺服器、路由器、個人電腦、微處理器式娛樂裝置、一端點裝置或其它共用網路節點,且基本上包括相對於電腦 802 所述之許多或所有的組件。該等邏輯連接可包括有線/無線連接到一區域網路(Local area network,「LAN」)、一廣域網路(Wide area network,「WAN」)、熱點(hotspot)等等。LAN 及 WAN 網路化環境為辦公室與公司內常見到,其可實施整個企業內的電腦網路,例如企業內網路,其所有可以連接至一全球通訊網路,例如網際網路。

當用於網路化環境中時,電腦 802 經由一有線/無線通訊子系統 842 連接至該網路(例如一網路介面轉接器、電路上收發器子系統等),以通訊於有線/無線網路、有線/無線印表機、有線/無線輸入裝置 844 等。電腦 802 可以包括一數據機,或具有其它手段來在該網路上建立通訊。在一網路化環境中,相對於電腦 802 的程式及資料可儲存在遠端記憶體/儲存裝置中,其係關聯於一分散式

系統。其將可瞭解到所示的網路連接為示例性，其可應用在電腦之間建立通訊的其它手段。

電腦 802 可使用像是 IEEE 802.xx 家族的標準之無線電技術用於與有線/無線裝置或個體進行通訊，像是可運作地設置成無線通訊之無線裝置(例如 IEEE 802.11, 空中傳輸調變技術)，藉由例如印表機、掃描機、桌上型及/或攜帶式電腦、個人數位助理(Personal digital assistant, 「PDA」)、通訊衛星、任何關連於一無線偵測標籤(如服務亭、新聞架、盥洗室)及電話之設備或位置的任何片段。此包括至少熱點用的無線保真度(Wireless Fidelity, 「Wi-Fi」), WiMax 及藍芽無線技術。因此，該等通訊可為如同與一習用網路之一預先定義的結構，或僅為至少兩個裝置之間的一即插即用通訊。Wi-Fi 網路使用稱為 IEEE 802.1x(a, b, g 等)之無線電技術來提供安全可靠及快速的無線連接。一 Wi-Fi 網路可用於將電腦彼此連接，連接到網際網路，及連接到有線網路(其使用 IEEE 802.3 相關的媒體及功能)。

以上所描述的包括所揭示架構的範例。當然，其不可能描述每一項可以想到的組件或方法之組合，但本技藝專業人士應可瞭解另有可能有許多其它的組合及排列。因此，該創新架構係要包含所有這些改變、修正及位在附屬申請專利範圍的精神及範疇內的變化。再者，在實

施方式或申請專利範圍中的用語「包括」，涵蓋之方式係要類似用語「包含」在一申請專利範圍內做為一轉折語時所詮釋的意義。

【圖式簡單說明】

- 第 1 圖為根據所揭示的架構之電腦實作的語音系統。
- 第 2 圖為包括第 1 圖之語音系統之額外態樣的系統。
- 第 3 圖為該加權轉換之示意圖。
- 第 4 圖為一快取加權轉換之示意圖。
- 第 5 圖為一電腦實作的語音方法。
- 第 6 圖為第 5 圖之方法的額外態樣。
- 第 7 圖為第 5 圖之方法的額外態樣。
- 第 8 圖為根據所揭示的架構用於對於一統計語言模型執行快速及穩定調適的一運算系統之方塊圖。

【主要元件符號說明】

100 語音系統	110 限制
102 歷史組件	112 短期記憶
104 使用者輸入歷史	114 長期記憶
105 語音字串	200 系統
106 轉換結果	204 轉換程序
108 調適組件	206 限制組件

- | | |
|--------------------|---------------------------|
| 208 基線語言模型 | 822 應用程式 |
| 300 示意圖 | 824 程式模組 |
| 400 示意圖 | 826 程式資料 |
| 800 運算系統 | 828 外部使用者輸入裝置 |
| 802 電腦 | 830 電路板上使用者輸入
裝置 |
| 804 處理單元 | 832 輸入/輸出裝置介面 |
| 806 記憶體子系統 | 834 輸出周邊 |
| 808 系統匯流排 | 836 繪圖介面 |
| 810 揮發性記憶體 | 838 外部顯示器 |
| 812 非揮發性記憶體 | 840 電路板上顯示器 |
| 814 儲存子系統 | 842 有線/無線通訊子系統 |
| 816 儲存介面 | 844 網路、電腦、無線周
邊、無線輸入裝置 |
| 818 可移除式記憶體子系
統 | |
| 820 作業系統 | |

七、申請專利範圍：

1. 一種電腦實作的語音系統，其包含下列：

一歷史組件，該歷史組件經配置以處理用於由一轉換程序轉換一語音字串的使用者輸入歷史；

一調適組件，該調適組件經配置以基於所施加的一或更多個限制來調適該轉換程序以合於該使用者輸入歷史，該使用者輸入歷史包含短期記憶及長期記憶，該短期記憶係前後文相關之機率差異，該長期記憶係一字詞之前文的頻率，其中該調適組件調適該轉換程序以合於該使用者輸入歷史係為了在該轉換程序期間抑制非預期之字詞的出現；及

一微處理器，該微處理器經配置以執行電腦可執行指令，該等指令關聯於該歷史組件及該調適組件。

2. 如申請專利範圍第1項所述之系統，另包含一限制組件，用於藉由基於前後文相關的機率差異增加一機率之方式來施加該等限制。

3. 如申請專利範圍第1項所述之系統，另包含一限制組件，用於藉由基於前後文相關的機率差異增加一機率之方式來施加一限制到該長期記憶。

4. 如申請專利範圍第1項所述之系統，其中該調適組件執行該長期記憶與一基線語言模型之間的動態線性內插。

5. 如申請專利範圍第 1 項所述之系統，其中當一字詞並非一候選清單之第一候選者時，該限制增加該字詞的機率。
6. 如申請專利範圍第 1 項所述之系統，其中施加於該短期記憶的限制利用一前後文敏感性短期記憶雙連字串 (bigram) 機率。
7. 如申請專利範圍第 1 項所述之系統，其中施加於該短期記憶之限制係基於一字詞與在一句子中該字詞的前後文來增加一機率。
8. 如申請專利範圍第 7 項所述之系統，其中該前後文包括在該句子中相對於該字詞的前文與後文。
9. 如申請專利範圍第 1 項所述之系統，其中該調適組件包括一學習演算法，該學習演算法基於一候選清單之一第一候選者與該候選清單的一所選候選者之間的差異來執行旗標學習 (flag-learning)，且該學習演算法在一下一轉換程序中移動該所選候選者到一第一轉換結果位置。
10. 一種電腦實作的語音系統，其包含下列：
 - 一歷史組件，該歷史組件經配置以處理用於在一轉換程序期間轉換一語音字串的使用者輸入歷史；
 - 一限制組件，該限制組件經配置以在該轉換程序期間施加限制到該使用者輸入歷史，該使用者輸入歷

史包括短期記憶與長期記憶，該短期記憶係前後文相關之機率差異，該長期記憶係一字詞之前文的頻率；

一調適組件，該調適組件經配置以基於該等限制，來調適該轉換程序以合於該使用者輸入歷史；及

一微處理器，該微處理器經配置以執行電腦可執行指令，該等指令關聯於該歷史組件、該限制組件及該調適組件。

11. 如申請專利範圍第 10 項所述之系統，其中該限制組件施加一限制到該短期記憶及該長期記憶，該短期記憶利用一前後文敏感性短期記憶雙連字串 (bigram) 機率，該長期記憶基於一前後文相關的機率差異增加一機率。
12. 如申請專利範圍第 10 項所述之系統，其中該調適組件執行該長期記憶與一基線語言模型之間的動態線性內插。
13. 如申請專利範圍第 10 項所述之系統，其中當一字詞並非一候選清單之第一候選者時，該限制增加該字詞的機率。
14. 如申請專利範圍第 10 項所述之系統，其中施加於該短期記憶的該限制，基於一字詞與在一句子中該字詞的前後文來增加一機率，該前後文包括在該句子中相對於該字詞的前文與後文。

15. 如申請專利範圍第 10 項所述之系統，其中該調適組件包括一學習演算法，該學習演算法根據在一候選清單中之一第一候選者及該候選清單中之一所選候選者之間的一差異來執行旗標學習 (flag-learning)，且該學習演算法在一下一轉換程序中將該所選候選者移動至一第一轉換結果位置。
16. 一種電腦實作的語音方法，其包含以下步驟：
- 處理用於在一轉換程序期間轉換一語音字串的使用者輸入歷史；
- 在該轉換程序期間施加限制到該使用者輸入歷史，該使用者輸入歷史包括短期記憶與長期記憶，該短期記憶係前後文相關之機率差異，該長期記憶係一字詞之前文的頻率；及
- 基於該等限制調適該轉換程序以合於該使用者輸入歷史。
17. 如申請專利範圍第 16 項所述之方法，另包含下列步驟：施加一限制，其基於前後文相關的機率差異增加一機率。
18. 如申請專利範圍第 16 項所述之方法，另包含下列步驟：執行在該長期記憶與一基線語言模型之間的動態線性內插。
19. 如申請專利範圍第 16 項所述之方法，另包含下列步

驟：當該字詞並非一候選清單之第一候選者時，增加一字詞的機率。

20. 如申請專利範圍第 16 項所述之方法，另包含下列步驟：施加一限制到該短期記憶，其基於一字詞與在一句子中該字詞的前後文增加一機率。

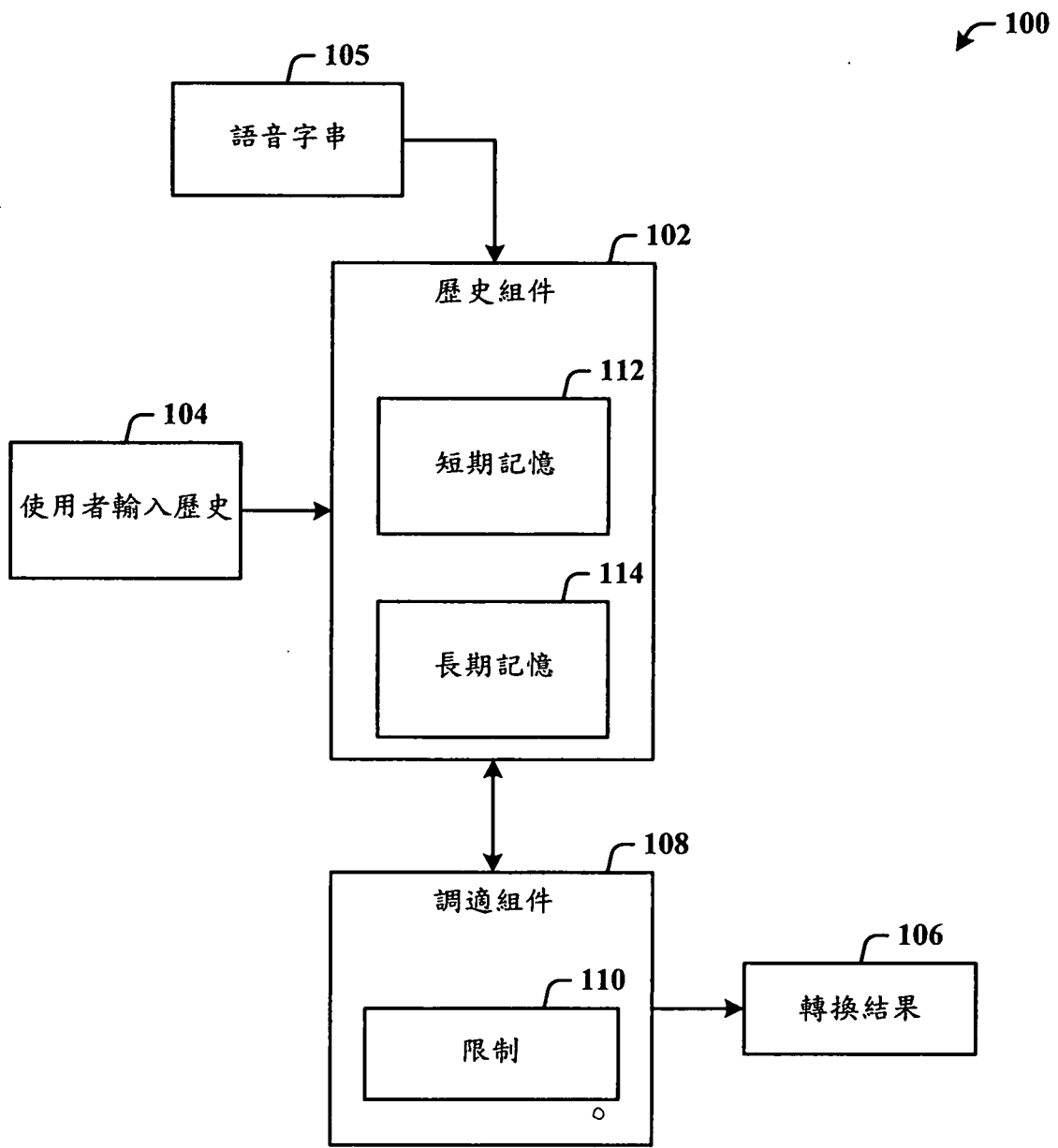
21. 如申請專利範圍第 16 項所述之方法，另包含下列步驟：

基於一候選清單的一第一候選者與該候選清單的一所選候選者之間的差異，來執行旗標學習 (flag-learning)；及

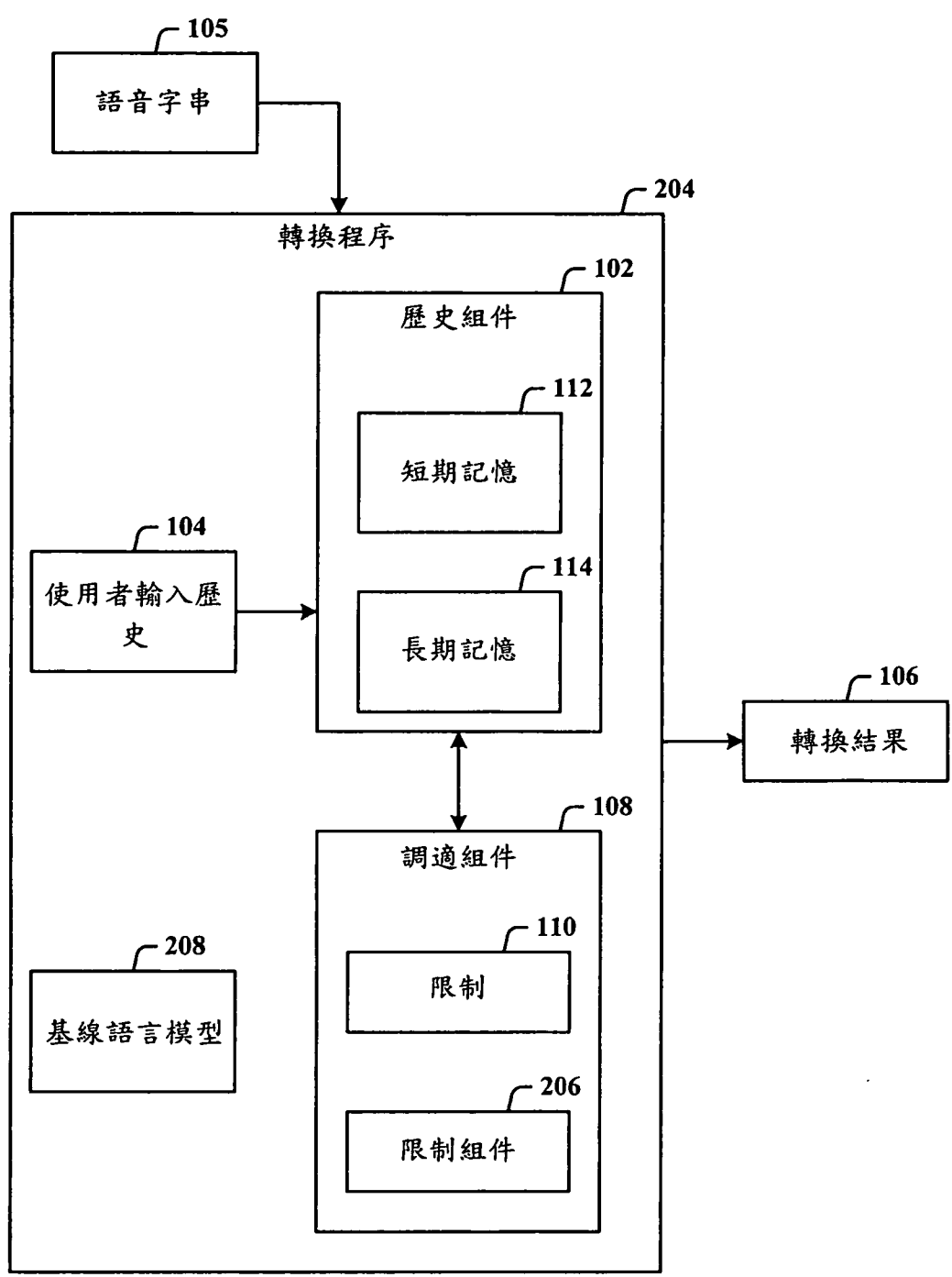
在一下一轉換程序中，移動該所選候選者到一第一轉換結果位置。

22. 如申請專利範圍第 20 項所述之方法，其中該前後文包括在該句子中相對於該字詞的前文及後文。

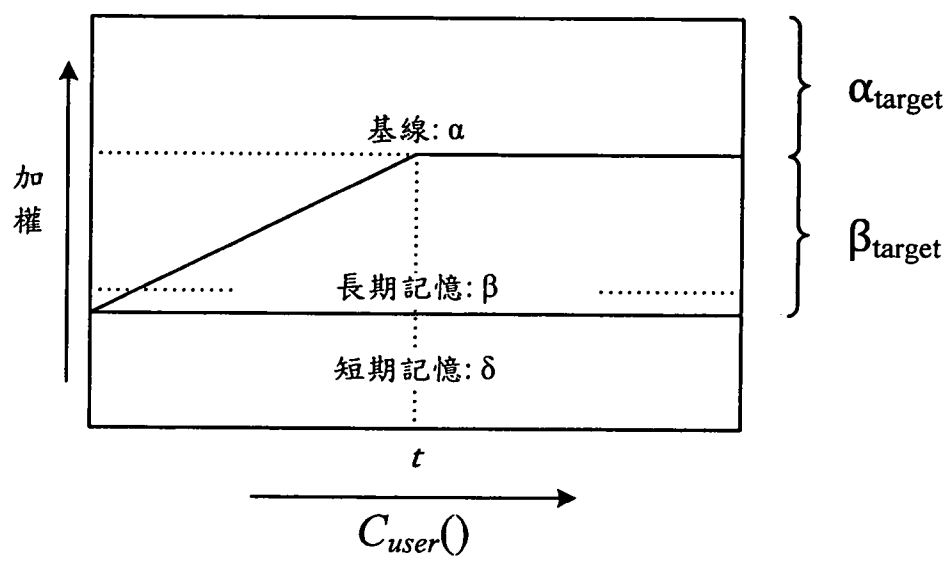
八、圖式：



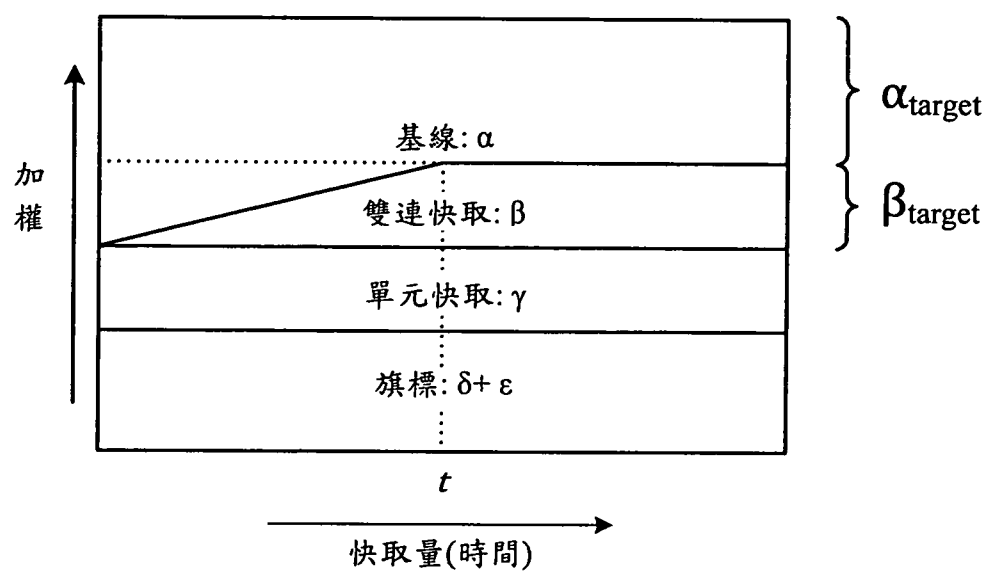
第 1 圖



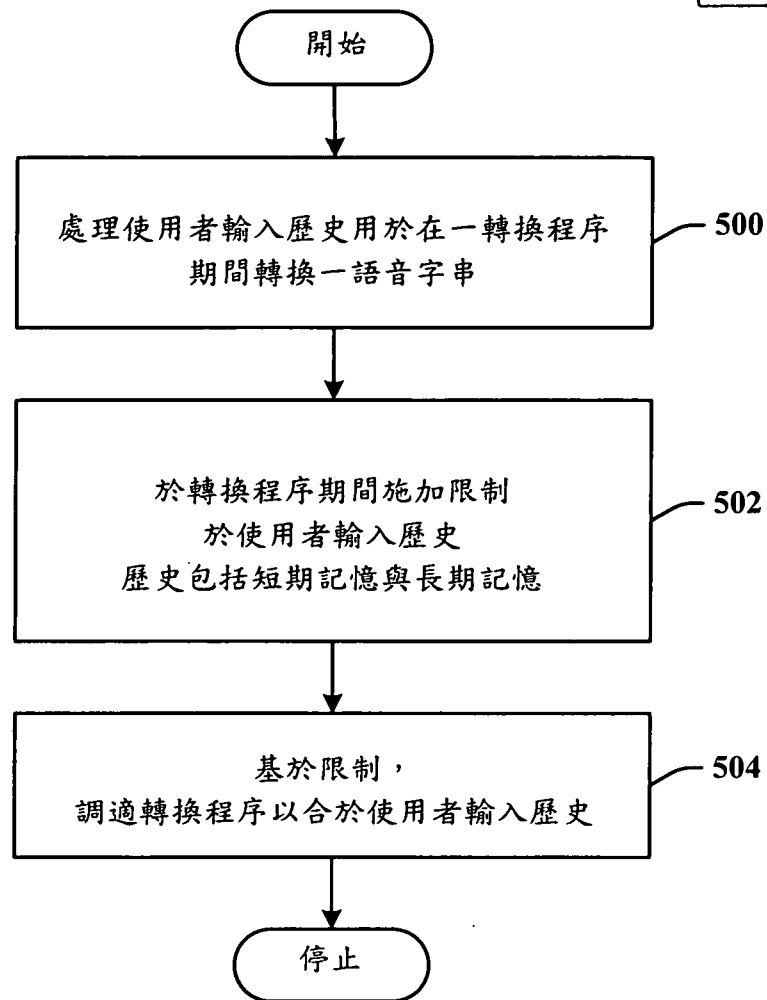
第 2 圖



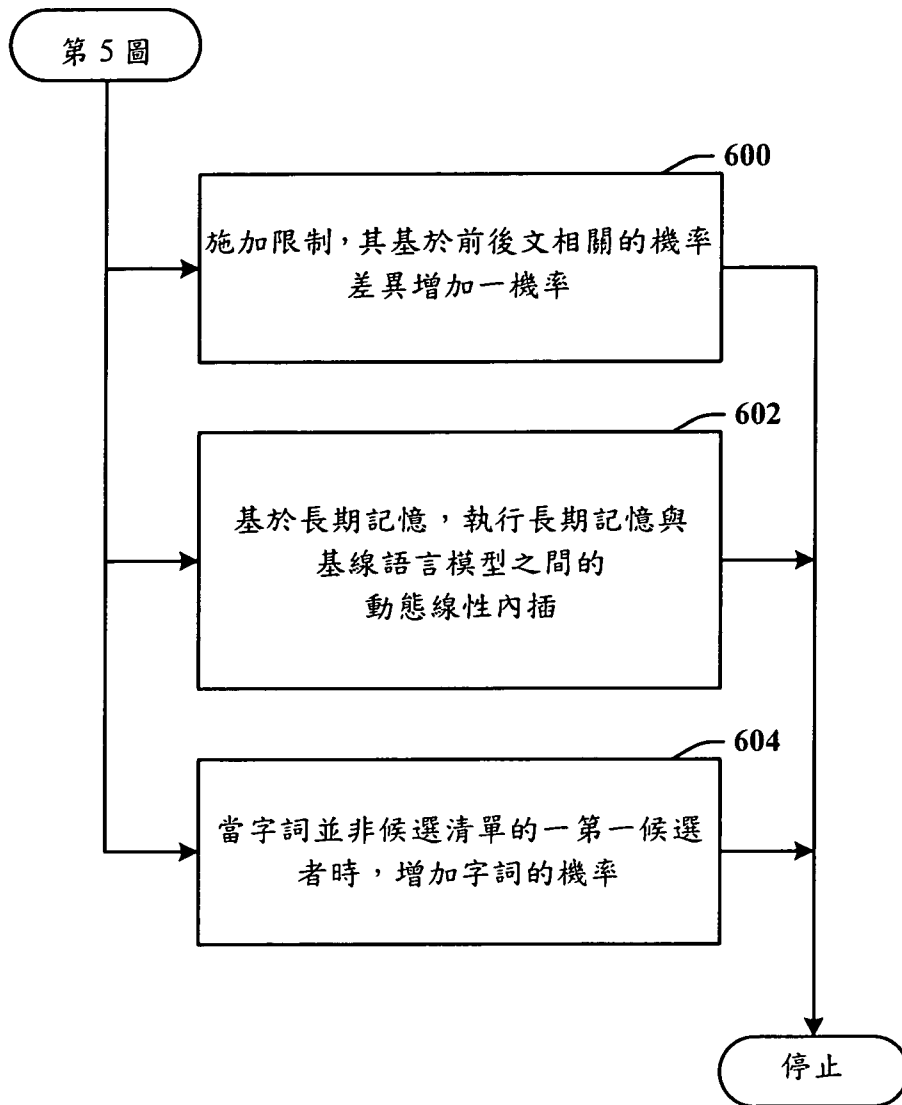
第 3 圖



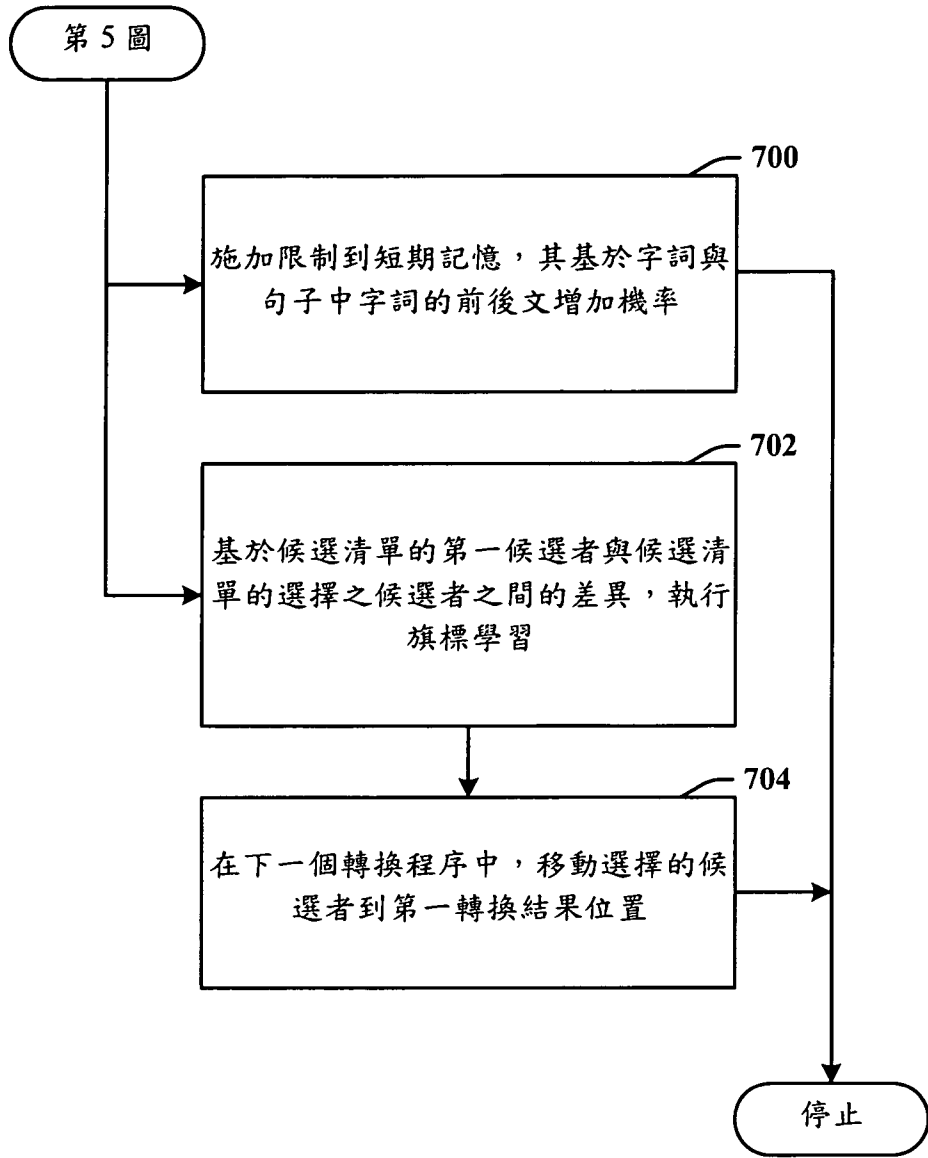
第 4 圖

103年10月22日修正
劃線頁(本)

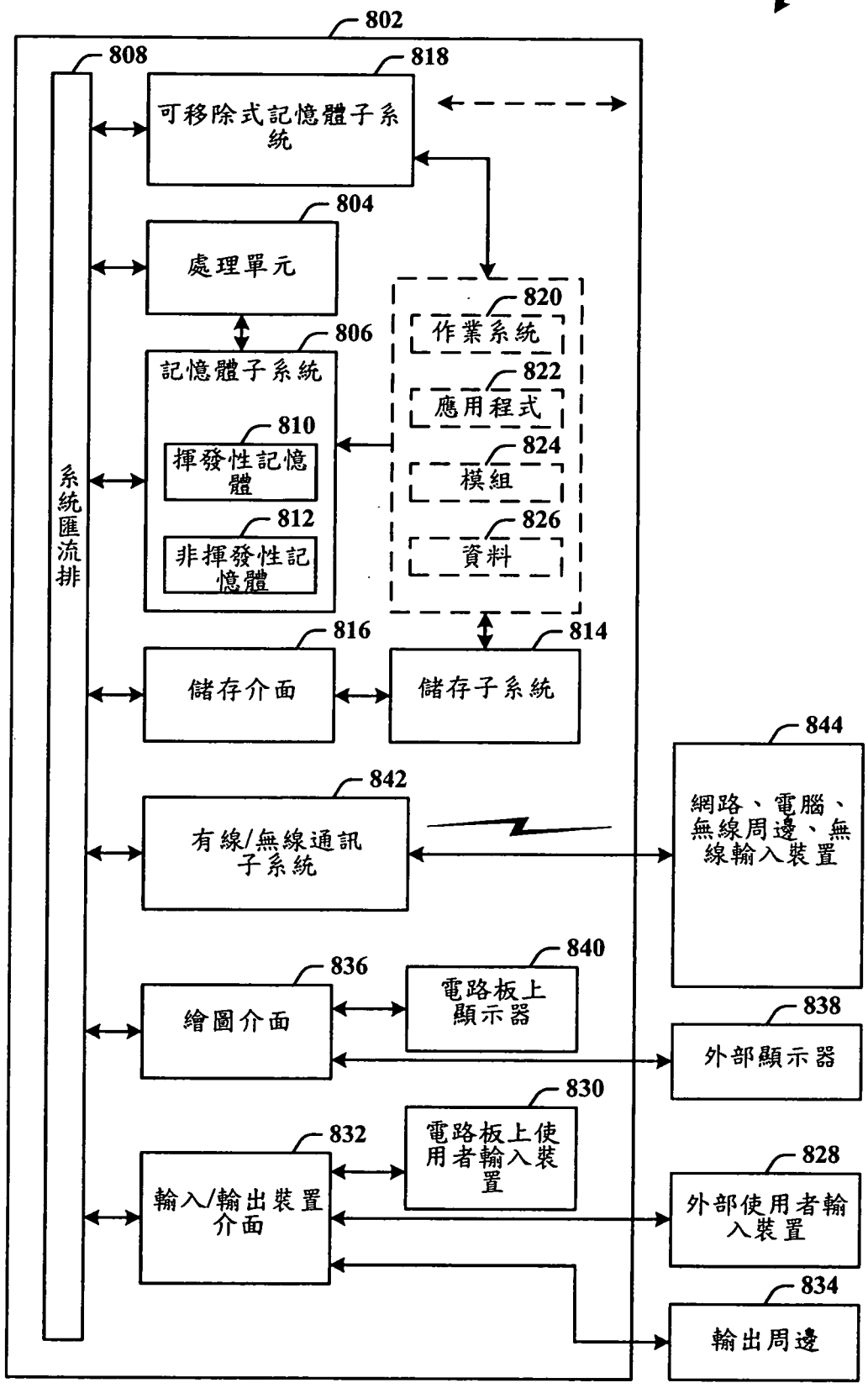
第 5 圖



第 6 圖



第 7 圖



第 8 圖