

申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

有主張專利法第二十七條第一項國際優先權：

無主張專利法第二十七條第一項國際優先權：

主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

主張專利法第三十條生物材料：

須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序

註記】

不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

九、發明說明：

【發明所屬之技術領域】

本案係為一種語音辨識方法，尤指一種具有智慧型鑑別調適功能之語音辨識方法。

【先前技術】

使用聲控指令的語音辨識系統的最大問題，就是辨識率不是百分之一百準確。辨識的錯誤會增加許多的不便，甚至有的時候會造成系統運作順暢的風險。

目前，大部分的聲控指令的語音辨識系統，一開始沒有考慮語音辨識錯誤，直接設計對連續的錯誤沒有感覺，也沒有對應系統來改善連續錯誤。使用者面對聲控指令的語音辨識系統，常常因而感覺無奈，無法解決屢次出現的錯誤，以及使用上繁瑣的困擾。最後，只好被動的選擇不用它了。

甚至有些時候，某些聲控指令的辨識錯誤，會造成系統運作相當的風險，對此，習用之聲控指令的語音辨識系統也只是單純的加上針對所有或是部份的辨識指令，進行多一步的確認動作。這樣的設計，增加了使用語音辨識系統上的麻煩。因此，不如更積極的透過智慧型的學習機制，提升部分或是整體指令辨識的鑑別率。

爰是之故，申請人有鑑於習知技術之缺失，發明出本案「具有智慧型鑑別調適功能之語音辨識方

法」，用以改善上述習用手段之缺失。

【發明內容】

本案之主要目的係為提供一種具有智慧型鑑別調適功能之語音辨識方法，其深入的考慮到使用者的感覺，在沒有增加使用者的麻煩下，提升系統的辨識率。此外，加上運用語者辨識的技術，可以擴大上述的學習機制，成為多人使用的操作介面。

根據上述構想，本案係提供一種語音辨識方法，其步驟包含(a)接收一使用者之語音；(b)辨識該語音以產生一辨識結果，該辨識結果具有一分數；(c)對於分數高的辨識結果，不作聲學模型的鑑別調適，但利用該語音之使用率來學習新的語言文法機率模型；(d)對於分數低而後被該使用者確認的辨識結果，進行聲學模型的鑑別調適，調高該語音之出現可能，並利用該語音之使用率來學習新的語言文法機率模型；以及(e)對於分數低而後被該使用者拒絕的辨識結果，進行聲學模型的鑑別調適，調低該語音之出現可能。

如所述之方法，該語音係為一口語指令。

根據上述構想，本案另提供一種語音辨識方法，用以辨識多個使用者之語音，其步驟包含(a)接收一特定使用者之語音；(b)辨識該語音以產生一辨識結果，該辨識結果具有一分數；(c)對於分數高的辨識結

果，自動切換至該特定使用者之系統中；(d)在一般情況下，對於分數低的辨識結果，繼續進行該特定使用者之語音辨識，直到累積出一定信任度，才切換至該特定使用者之系統中；以及(e)在特殊情況下，對於分數低的辨識結果，逕行詢問該特定使用者，並立刻切換至該特定使用者之系統中。

如所述之方法，每一使用者具有各自的系統，紀錄著各自的口語指令之使用成功和錯誤之相關紀錄，亦各自訓練調適出各自的聲學和語言之機率模型。

如所述之方法，該語音係為一口語指令。

如所述之方法，該特殊情況係為當該口語指令發生連續錯誤時。

如所述之方法，該特殊情況係為當進行對該特定使用者之私密資料處理時。

根據上述構想，本案又提供一種語音處理方法，步驟包含(a)接收一使用者之語音；(b)辨識該語音以產生一辨識結果；(c)當該辨識結果連續出現錯誤時，偵測該辨識結果以取得其錯誤型態；以及(d)根據該錯誤型態而進行一調適處理。

如所述之方法，該語音係為一口語指令。

如所述之方法，該錯誤型態包含一第一類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出相同的錯誤；一第二類型，係該使用者連續重覆

輸入相同的口語指令，卻被連續辨識出各自不同的錯誤；一第三類型，係連續輸入非口語之背景雜音，卻分別被辨識為有意義之語音指令；以及一第四類型，係為其它零散的輸入錯誤連續發生。

如所述之方法，該調適處理包含抑制重覆發生的錯誤選項，暫時性的調整語言文法機率模型。

如所述之方法，該調適處理包含另行建立抑制指令的暫時性資料庫，降低連續被拒絕的選項再出現的機率。

根據上述構想，本案再提供一種語音辨識/處理系統，其包含一語音辨識單元，用以接收並辨識一使用者之語音，以產生一辨識結果；一錯誤偵測單元，連結於該語音辨識單元，用以於該辨識結果連續出現錯誤時，偵測該辨識結果以取得其錯誤型態；以及一錯誤抑制單元，連結於該錯誤偵測單元，用以根據該錯誤型態而進行一調適處理。

如所述之系統，該語音係為一口語指令。

如所述之系統，該錯誤型態包含一第一類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出相同的錯誤；一第二類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出各自不同的錯誤；一第三類型，係連續輸入非口語之背景雜音，卻分別被辨識為有意義之語音指令；以及一第四類型，係為其它零散的輸入錯誤連續發生。

如所述之系統，該調適處理包含抑制重覆發生的錯誤選項，暫時性的調整語言文法機率模型。

如所述之系統，該調適處理包含另行建立抑制指令的暫時性資料庫，降低連續被拒絕的選項再出現的機率。

【實施方式】

本發明所設計的學習機制，是以底下的語音辨識 (Speech Recognition) 系統的架構為前提。語音辨識系統的口語辨識操作步驟包含：口語輸入、語音辨識、自動結果確認、以聲音或影像回應辨識結果、手動結果確認或是其他口語輸入更正。每個口語輸入之後，辨識結果會有一個分數，分數高的口語指令可以不需手動確認就執行；其他分數低的口語指令則需手動確認來輔助執行。此時，系統可能會以聲音或影像等回應的形式，通知使用者進行手動或是口語的再確認步驟，例如，以按鍵確認或拒絕，或是以口語指令回答對或錯。如果使用者已確認，就完成了該口語指令的流程。如果是拒絕動作，則要進行重新輸入的流程，或是進行更正錯誤的流程，直到辨識結果正確為止。

本發明所設計的學習機制，係包含自動的語者辨識 (Speaker Recognition) 技術於系統中。語者辨識系統的使用，包含針對新使用者的學習階段，以及認識使用者之後的正常使用階段。

的鑑別調適；但是利用其口語指令的使用率，學習新的語言文法機率模型。

針對分數低而後被確認的口語指令，本發明將進行聲學模型的鑑別調適，調高其出現可能；且利用其口語指令的使用率，學習新的語言文法機率模型。

針對分數低而後被拒絕的口語指令，本發明將進行聲學模型的鑑別調適，調低其出現可能；卻不拿來用在學習新的語言文法機率模型上。

基本整體的鑑別調適處理，幫助學習使用者的特殊錯誤情形，並建立該使用者特有的聲學和語言模型。

[多語者模型切換模式下之基本整體的鑑別調適處理]

藉由語者辨識技術，上述的基本整體的鑑別調適處理，可以自動的學習出多語者的系統環境，並且使用在多語者操控的口語指令語音辨識系統上。語者辨識和使用者的切換機制，在語者辨識技術的應用下自動的進行。各登錄有名的使用者，有各自的系統環境：紀錄著各自的口語指令的使用成功和錯誤的相關紀錄；也各自訓練調適出各自的聲學和語言的機率模型。請參閱第一圖，其係本案之使用者的切換處理流程圖。切換使用者的機制，可以如下進行。

(1)語者辨識技術，在語音辨識功能進行之後接

著進行(S21)。當辨識為同一語者時(S22)，則系統不動作。

(2)當辨識出不同語者之後，針對分數高的語者辨識結果，進行自動切換的動作，並顯示在被操控的機器上的角落。

(3)一般情況下，針對分數低的語者辨識結果，留著隨後進行的口語指令再一次進行語者辨識的確認，直到累積出一定信任，才進行使用者切換的動作(S23)。

(4)特殊情況下，針對分數低的語者辨識結果，逕行詢問使用者，並立刻進行切換的動作(S24)。例如，當口語指令發生連續錯誤時，可以進行切換語者，立即改善辨識品質。或是例如，當進行對使用者之私密資料處理時，可以逕行詢問使用者，以正確的進行使用者之私密資料處理。

〔針對連續錯誤的情形進行暫時性的鑑別調適處理〕

本發明的設計將針對連續錯誤的發生，設計抑制錯誤再發生的法則，進行暫時性的鑑別調適處理，以有效抑制錯誤的連續發生，維持口語操作介面的方便自然度。連續錯誤的定義，是指被操控的機器在相同的狀態下，口語指令的語音辨識結果連續出現錯誤而未能進行指令的動作。所謂的「機器在相同的狀態」包含TV的電台沒有更動，音量的大小沒有變化，以

及亮度的強暗等等口語指令操控的範圍，沒有更動。如果符合「機器在相同的狀態」下，可以進一步假設口語指令的連續錯誤發生，是輸入了相同的口語指令，因而可藉此偵測抑制相同的錯誤再發生。

請參閱第二圖，其係本案之語音辨識/處理系統之方塊圖。該系統包含一語音辨識單元 21、一錯誤偵測單元 22、及一錯誤抑制單元 23。本案之暫時性的鑑別調適處理係經由該錯誤偵測單元 22 來偵測出錯誤型態，再經由該錯誤抑制單元 23 來處理不同型態的錯誤抑制動作。而該錯誤偵測單元 22 所偵測的連續錯誤分為以下 A ~ D 類型：

A 型:使用者連續重複輸入相同的語音指令，卻遭到連續辨識出相同的錯誤。

B 型:使用者連續重複輸入相同的語音指令，卻遭到連續辨識出各自不同的錯誤。

C 型:係連續輸入了非口語的背景雜音，卻分別被辨識為有意義的語音指令。

D 型:其它零散的輸入錯誤連續發生。

請參閱第三圖，其係本案之連續辨識錯誤的鑑別處理流程圖。如第三圖所示，當連續錯誤發生 N 次時(S31)，系統會偵測語音能量是否大於等於一預設值 E(S32)，如否，則判定該語音屬於 C 型。當語音能量大於該預設值 E 時，系統會偵測該語音之錯誤相似度(全段)是否大於等於一預設值 P1%，如是，則

判定該語音屬於 A 型。當該語音之全段錯誤相似度小於該預設值 P1%時，系統會偵測該語音之中段(去除指定比例句頭及指定比例句尾的語句中段)錯誤相似度是否大於等於一預設值 P2%，如是，則判定該語音屬於 B 型。剩下其他情形，則判定該語音屬於 D 型。

本案之錯誤抑制單元 23 係根據偵測的錯誤型態，分別給予調適處理。處理主要是抑制重複發生的錯誤選項，暫時性的調整語言文法機率模型，或是另行建立抑制指令的暫時性資料庫，降低連續被拒絕的選項再出現的機率。在機器狀態改變之後，就視為新的條件狀態，應將暫時性的調解除除，回到原來正常的的使用狀態，並重新計算錯誤發生的連續次數。

暫時性的語言文法機率調整，可以是某一比例的機率降低，甚至降低至零。可以直接調整在進行中的語言文法機率模型，但是要另外儲存常態模型，才可以在解除暫時性的調整之後，回覆至常態模型；也可以另外儲存一個語言文法的抑制機率模型，在計算進行中的語言文法機率時，採取常態模型和抑制模型相減的結果。

綜上所述，本案係提供一種具有智慧型鑑別調適功能之語音辨識方法，其深入的考慮到使用者的感覺，在沒有增加使用者的麻煩下，提升系統的辨識率。此外，加上運用語者辨識的技術，可以擴大上述

的學習機制，成為多人使用的操作介面。因此，本案能有效改善習知技術之缺失，是故具有產業價值，進而達成發展本案之目的。

本案得由熟悉本技藝之人士任施匠思而為諸般修飾，然皆不脫如附申請專利範圍所欲保護者。

【圖式簡單說明】

第一圖：其係本案之使用者的切換處理流程圖。

第二圖：其係本案之語音辨識/處理系統之方塊圖。

第三圖：其係本案之連續辨識錯誤的鑑別處理流程圖。

【主要元件符號說明】

21:語音辨識單元

22:錯誤偵測單元

23:錯誤抑制單元

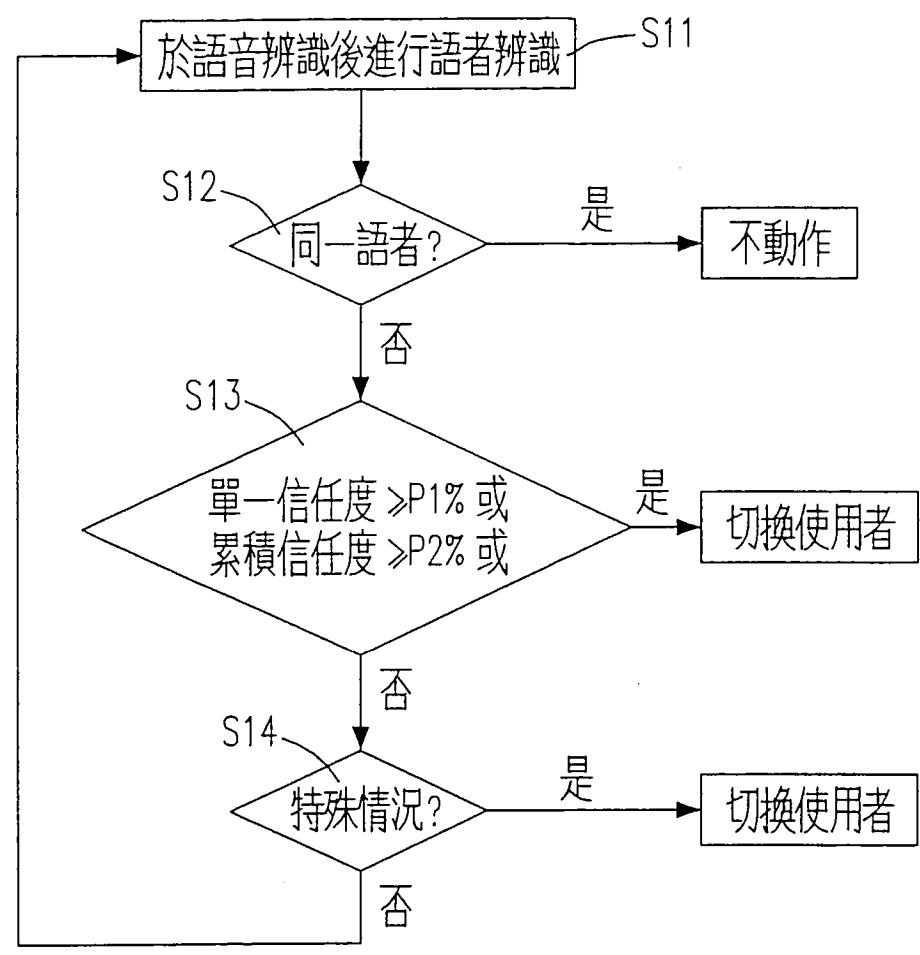
五、中文發明摘要：

本案係為一種語音辨識方法，其步驟包含(a)接收一使用者之語音；(b)辨識該語音以產生一辨識結果，該辨識結果具有一分數；(c)對於分數高的辨識結果，不作聲學模型的鑑別調適，但利用該語音之使用率來學習新的語言文法機率模型；(d)對於分數低而後被該使用者確認的辨識結果，進行聲學模型的鑑別調適，調高該語音之出現可能，並利用該語音之使用率來學習新的語言文法機率模型；以及(e)對於分數低而後被該使用者拒絕的辨識結果，進行聲學模型的鑑別調適，調低該語音之出現可能。

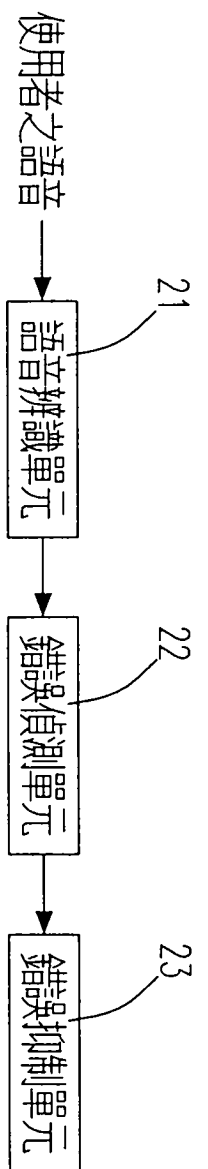
六、英文發明摘要：

A speech recognition method is provided. The speech recognition method includes the steps of (a) receiving a speech from a user; (b) recognizing the speech to generate a recognition result, the recognition result having a score; (c) not performing the adjustment for the acoustic model but using the utility rate of the speech to learn a new language and grammar probability model when the recognition result has a high score; (d) performing the adjustment for the acoustic model to increase the occurrence probability of the speech and using the utility rate of the speech to learn the new language and grammar probability when the recognition result has a low score and is subsequently recognized by the user; and (e) performing the adjustment for the acoustic model to decrease the occurrence probability of the speech when the recognition result has a low score and is subsequently rejected by the user.

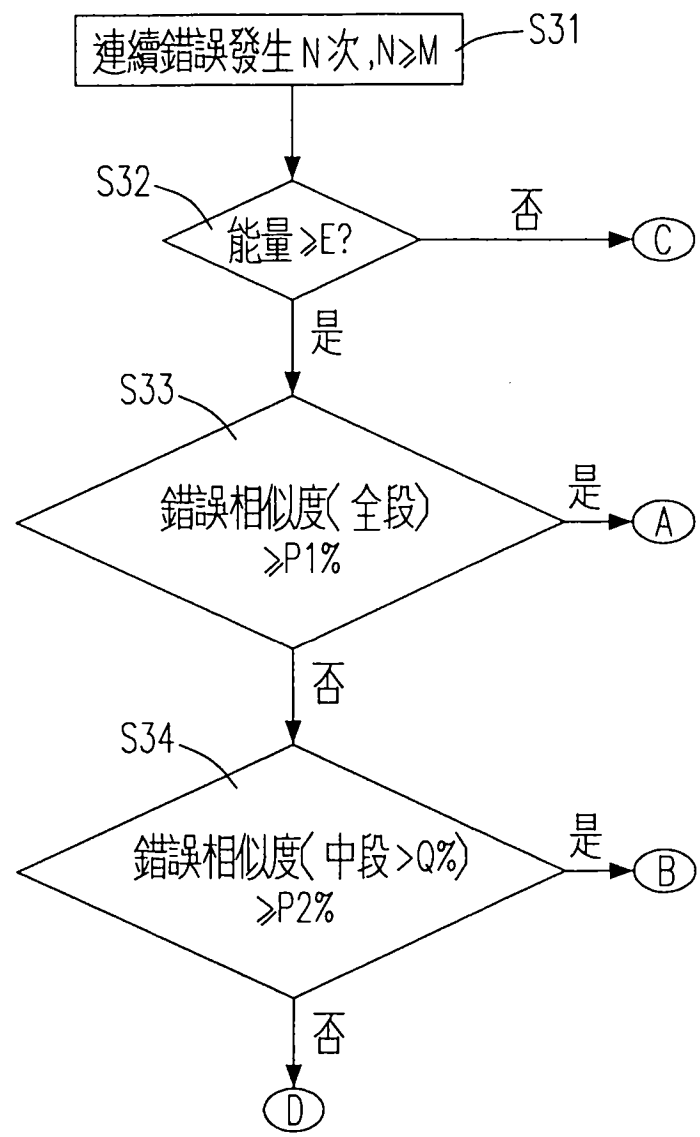
十一、圖式：



第一圖



第二圖



第三圖

七、指定代表圖：

(一)本案指定代表圖為：第(一)圖。

(二)本代表圖之元件符號簡單說明：

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：95146777

※ 申請日期：95.12.13

※IPC 分類：G10L15/06 (2006.01)

一、發明名稱：

具有智慧型鑑別調適功能之語音辨識方法及系統 /
SPEECH RECOGNITION METHOD AND SYSTEM WITH INTELLIGENT
CLASSIFICATION AND ADJUSTMENT

二、申請人：(共 1 人)

名稱：台達電子工業股份有限公司 / Delta Electronics, Inc.

代表人：鄭崇華 / Chong-Hua Zheng

住居所地址：桃園縣龜山鄉山頂村興邦路 31 之 1 號 / 31-1 Shien
Pan Road, Kuei San Industrial Zone, Taoyuan Hsien 333,
Taiwan, R.O.C.

國 籍：中華民國 / TW

三、發明人：(共 1 人)

(一)姓 名：王瑞璋 / Wang, Jui-Chang

國 籍：中華民國 / TW

四、聲明事項：

主張專利法第二十二條第二項 第一款或 第二款規定之事

實，其事實發生日期為： 年 月 日。

申請前已向下列國家(地區)申請專利：

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：95146777

※ 申請日期：95.12.13

※IPC 分類：G10L15/06 (2006.01)

一、發明名稱：

具有智慧型鑑別調適功能之語音辨識方法及系統 /
SPEECH RECOGNITION METHOD AND SYSTEM WITH INTELLIGENT
CLASSIFICATION AND ADJUSTMENT

二、申請人：(共 1 人)

名稱：台達電子工業股份有限公司 / Delta Electronics, Inc.

代表人：鄭崇華 / Chong-Hua Zheng

住居所地址：桃園縣龜山鄉山頂村興邦路 31 之 1 號 / 31-1 Shien
Pan Road, Kuei San Industrial Zone, Taoyuan Hsien 333,
Taiwan, R.O.C.

國 籍：中華民國 / TW

三、發明人：(共 1 人)

(一)姓 名：王瑞璋 / Wang, Jui-Chang

國 籍：中華民國 / TW

四、聲明事項：

主張專利法第二十二條第二項 第一款或 第二款規定之事

實，其事實發生日期為： 年 月 日。

申請前已向下列國家(地區)申請專利：

在針對新使用者的學習階段，語者的聲紋模型需要建立；此時，必須透過 G U I 圖形使用介面的輔助，讓需要專用環境 user profile 的使用者，在尚未有足夠的聲紋資料前，先行以圖形化介面或是鍵盤輸入做為選擇語者的操作介面。然後在自然的使用口語語音辨識時，紀錄下該語者的聲紋資料。待該語者的聲紋比對資料充足之後，使用者就不需要每次透過 G U I 圖形介面或是鍵盤輸入的方式，選擇自己的名字或代號，才能開始使用該系統。

在認識使用者之後的正常使用階段，語者辨識系統能夠自動的辨識出語者，顧及使用操作的便利。此時，系統將能夠透過語者辨識系統，不僅能夠自動的辨識出語者，並且能夠自動的切換使用者環境，提供更為便利的資訊服務。

底下根據上述的系統和操作資訊，說明智慧的學習機制如何進行。大致而言，針對語音聲學模型和語言文法機率模型，分別進行底下兩種處理工作：基本整體的鑑別調適處理，及針對連續性錯誤的情形進行暫時性的鑑別調適處理。

[基本整體的鑑別調適處理]

口語指令分為高分數自動過關、低分數確認通過、以及低分數拒絕通過三種類別。

針對分數高的口語指令，本發明並不作聲學模型

補充修正日期：民國 100 年 3 月 3 日。

十、申請專利範圍：

1. 一種語音辨識方法，其步驟包含：

(a) 接收一使用者之語音；

(b) 辨識該語音以產生一辨識結果，該辨識結果具有一分數；

(c) 對於分數高的辨識結果，不作聲學模型的鑑別調適，但利用該語音之使用率來學習新的語言文法機率模型；

(d) 對於分數低而後被該使用者確認的辨識結果，進行聲學模型的鑑別調適，調高該語音之出現可能，並利用該語音之使用率來學習新的語言文法機率模型；以及

(e) 對於分數低而後被該使用者拒絕的辨識結果，進行聲學模型的鑑別調適，調低該語音之出現可能。

2. 如申請專利範圍第 1 項所述之方法，其中該語音係為一口語指令。

3. 一種語音處理方法，其步驟包含：

(a) 接收一使用者之語音；

(b) 辨識該語音以產生一辨識結果；

(c) 當該辨識結果連續出現錯誤時，偵測該辨識結果以取得其錯誤型態；以及

(d) 根據該錯誤型態而進行一調適處理。

4. 如申請專利範圍第 3 項所述之方法，其中該語音係

補充修正日期：民國 100 年 3 月 3 日。

為一口語指令。

5. 如申請專利範圍第 4 項所述之方法，其中該錯誤型態包含：

一第一類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出相同的錯誤；

一第二類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出各自不同的錯誤；

一第三類型，係連續輸入非口語之背景雜音，卻分別被辨識為有意義之語音指令；以及

一第四類型，係為其它零散的輸入錯誤連續發生。

6. 如申請專利範圍第 4 項所述之方法，其中該調適處理包含抑制重覆發生的錯誤選項，暫時性的調整語言文法機率模型。

7. 如申請專利範圍第 4 項所述之方法，其中該調適處理包含另行建立抑制指令的暫時性資料庫，降低連續被拒絕的選項再出現的機率。

8. 一種語音辨識/處理系統，其包含：

一語音辨識單元，用以接收並辨識一使用者之語音，以產生一辨識結果；

一錯誤偵測單元，連結於該語音辨識單元，用以於該辨識結果連續出現錯誤時，偵測該辨識結果以取得其錯誤型態；以及

一錯誤抑制單元，連結於該錯誤偵測單元，用以

補充修正日期：民國 100 年 3 月 3 日。

根據該錯誤型態而進行一調適處理。

9. 如申請專利範圍第 8 項所述之系統，其中該語音係為一口語指令。

10. 如申請專利範圍第 9 項所述之系統，其中該錯誤型態包含：

一第一類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出相同的錯誤；

一第二類型，係該使用者連續重覆輸入相同的口語指令，卻被連續辨識出各自不同的錯誤；

一第三類型，係連續輸入非口語之背景雜音，卻分別被辨識為有意義之語音指令；以及

一第四類型，係為其它零散的輸入錯誤連續發生。

11. 如申請專利範圍第 8 項所述之系統，其中該調適處理包含抑制重覆發生的錯誤選項，暫時性的調整語言文法機率模型。

12. 如申請專利範圍第 8 項所述之系統，其中該調適處理包含另行建立抑制指令的暫時性資料庫，降低連續被拒絕的選項再出現的機率。