

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：

【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：

美國 US；西元 2005 年 02 月 22 日；11/062,731

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】

本發明係關於使用多型態、基於網路之介面以辨識一使用者。

【先前技術】

一種多型態介面可讓一使用者透過一應用程式或藉由影像或聲音手段(means)之其他系統達成通訊。例如，一使用者可透過圖像使用者介面(GUI)，藉由發出一或更多的說話語調(spoken utterances)，或藉由提供一 GUI 輸入與聲音輸入之組合，以達成通訊。多型態互動可提供使用者更直接的手段，與一系統通訊。

關於可攜式電腦裝置，例如智慧型電話、個人數位助理、及其他具有短小特徵的裝置，多型態介面可供使用者依據環境選擇一互動的手段。在有些情形，利用 GUI 比較有利，而其他時候可能聲音資料輸入可能較有優勢或較方便。舉例而言，在正常的環境下，一使用者可能會選擇聲音互動，因其為較直接且快速的方式。聲音資料輸入更可讓使用者免去使用一微型鍵盤或小型的 GUI，無論是使用二者任一都很累贅或麻煩。

然而，在吵雜的環境下就可能無法實行聲音資料

輸入。環繞著數層的噪音可能會遮蔽掉使用者聲音輸入，至無法精準的辨識使用者的語音的程度。在有些環境之下，例如要辨認一使用者身分時，使用者可能也還是要選擇用鍵盤鍵入或非聲音的互動。典型地，使用者的確認係利用一使用者名稱與一密碼。由於將這種資訊以語音輸入一多型態介面，可使得其他人偷聽到保密的資訊，因此使用者經常利用 GUI 或鍵盤鍵入這類資訊。然而，值得注意的是，在可攜式電腦裝置中使用此方法輸入這類資訊很累贅且麻煩。

使用密碼還有一些其他的缺點。典型地，密碼很難記憶，而且不很安全。使用者要記住其擁有之多組帳號的密碼尤其困難，每一者都需要一個使用者名稱與密碼。為了避免必須要記住每個密碼，使用者傾向於使用一個通用的密碼於多個帳號，將密碼寫在一存取卡的背面，例如提款卡的背面，或者將密碼設定與使用者名稱相同。上述的任一種行為都會使得使用者的帳戶不安全。

儘管如此，像是銀行、航空訂位、產品購買及其他類似的功能都開始設有多型態、基於網路之介面，使用者的辨認變得愈來愈重要。若交易是安全的，而且若是這些系統可確保使用者確為其人，這些系統才會成功。因此，透過一多型態介面提供一辨認使用者

的技術即可獲得利益，克服上述的缺點。

【發明內容】

本發明提供一種解決方案，透過一通訊網絡 (communication network) 供辨認一使用者身份。本發明之一實施例包括一種使用多型態、基於網路之介面以辨認一使用者身分之方法。該方法可包含傳送一多型態介面標記語言文件至一遠端電腦裝置，當遠端電腦裝置呈現時，請求辨認一使用者之身分，並使該使用者之聲音傳送到一多型態、基於網路的應用程式。該使用者之身分辨認及聲音 (recording) 幾乎是同時地從客戶端裝置被接收。將該聲音與一使用者之身分辨認相關聯之聲紋相比較。因此，依據上述之比較所獲得之一結果，可選擇性地允許在遠端電腦裝置之使用者存取系統。

本發明之另一實施例包含一種供辨認一使用者身份之系統。該系統可包含一辨認系統，其具有一聲紋資料庫，係包括複數聲紋，每個聲紋都與一使用者身分辨認相關聯。該系統還可包含一多型態應用程式，其係組態以傳送一多型態介面標記語言文件至一客戶端裝置。執行時，該多型態介面標記語言文件可指定一隨機文字稿並產生一使用者身分辨認與聲音，係包含一語者大聲地朗讀該隨機文字稿之一數位代表，係被傳送回該多型態應用程式。

該多型態應用程式可提供文字稿、聲音、以及該使用者之身分辨認給該辨認系統，藉由比較聲音與選自與使用者之身分辨認相關聯的複數聲紋之一，以辨認該語者之身分。

本發明之另一個實施例係可包括一種機器可讀取之儲存器，其係可被程式化使得一機器執行本發明所述之各種步驟。

【實施方式】

本發明提供一解決方案，透過使用一多型態、基於網路之介面，供辨認一使用者身份。依據本發明，一使用者的身分可透過使用者之聲音辨識技術得以被確認。接收自一多形態、基於網路之應用程式(application)之多型態文件可顯示出一使用者。多形態文件，如一網頁形式的介面，可請求使用者提供一使用者識別符或使用者名稱(以下簡稱「使用者 ID」)以及一聲音樣本。使用者 ID 及聲音樣本兩者皆可再被提供至多形態、基於網路之應用程式。依據聲音樣本並參照使用者 ID，聲音辨識得以執行。依據聲音辨識的結果，即可准許使用者進入一線上系統或拒絕使用者進入。

圖 1 係依據揭露之發明的安排之一概要圖，闡述使用一基於網路多型態介面執行使用者辨認之一系統

100。如圖所示，系統 100 可包括一客戶端裝置 105、一包含多型態應用程式 120 之伺服器 110，以及一具有一聲紋資料庫(voice print database)130 之辨認系統 125。

在一實施例中，客戶端裝置 105 可為各式各樣的電腦裝置，例如一桌上型電腦、一伺服器、或類似者。在另一實施例中，客戶端裝置 105 可為一可攜式及/或移動式電腦裝置，例如一筆記型電腦、組態供有線或無線網際網路通訊之一個人數位助理、一行動電話、或其他類似者。在任何情形，所有客戶端裝置 105 都可包含適當的軟體用供傳送請求到伺服器 110，並可從伺服器 110 接收訊息。舉例而言，客戶端裝置 105 可包含一適當的瀏覽器，以呈現標記語言文件，其包括有多型態標記語言 (MML) 文件。

MML 文件可指明影像與聲音元件(visual and voice components)，供使用者互動。因此，MML 文件可包含任何依據文件或檔案的標記語言，當成呈現這些 MML 文件時，允許一使用者透過一圖像使用者介面 (GUI) 的視覺手段提供輸入，以及透過說話或提供雙調多頻率處理(dual tone multi-frequency processing)之聲音手段提供輸入。這類標記語言其中的一個例子就是 X+V 標記語言。X+V 是被用來發展多型態網頁的標記語言，其結合可延伸超文字標記語言(XHTML)與語音可延伸標記語言(VoiceXML)的

一子集。

一般說來，XHTML 係 HTML 4.0 的一個版本，其已經被調整成與可延伸標記語言 XML 的規則相符。VoiceXML 為一種以 XML 為基礎的語言，其係透過全球資訊網國際協會 (WWW Consortium, W3C) 發展。VoiceXML 提供一種標準格式，供建立以語音為基礎的應用程式。XHTML 及 VoiceXML (X+V) 的結合，使得網頁開發者可以增加聲音的輸入與輸出至傳統以圖像為基礎的網頁。

另一個 MML 的例子是基於新增聲音應用語言標籤 (SALT) 至一主要標記語言，如 XHTML、超文字標記語言 (HTML)、或可調向量圖像 (SVG)。SALT 係一種將被用來發展多型態網頁的標記語言，其可支援視覺及聲音模式的使用者互動。SALT 係一種以 XML 為基礎的語言，其係透過 SALT Forum 發展。

MML 伺服器 110，例如網路伺服器，可為各種形式的資訊處理系統，其可將請求分區 (fielding)，並可透過一通訊網絡來處理資訊。如圖所示，伺服器 110 與客戶端裝置 105 可透過一通訊網絡 115，如網際網路、區域網路 (LAN)、廣域網路 (WAN)、移動網路或蜂巢網路、及其他各種通訊網絡、或結合上述的各種網

路，而達成通訊。

多型態應用程式 120 可在伺服器 110 內執行。因此，多型態應用程式 120 可接收來自客戶端裝置 105 之請求與資訊，且回饋提供如標記語言文件之資訊，無論其是否為多型態。雖然圖中未顯示，多型態應用程式 120 可包含或存取聲音處理服務，如文字-語音 (text-to-speech; TTS)、聲音辨識、及/或雙音多頻率處理。這些服務可位在伺服器 110 之內，或者可位在可與多型態應用程式 120 通訊相連之不同的電腦系統。

辨認系統 125 可藉由比較使用者記錄的聲音與儲存在聲紋資料庫 130 中的聲紋，辨認使用者身分。聲紋資料庫 130 可包含來自使用者的不同的複數聲紋，這些聲紋被登錄在辨認系統 125 中。每個聲紋可對應到，或相關聯至一個使用者 ID。當接收到一個聲音記錄時，辨認系統 125 可將該聲音記錄與聲紋資料庫 130 中的聲紋作比較，以判斷是否符合。

運作時，客戶端裝置 105 可傳送一請求 135 至多型態應用程式 120 請求一 MML 文件。多型態應用程式 120 可藉由傳送 MML 文件 140，如一多型態網頁，作為回應。依據所呈現的 MML 文件 140，客戶端裝置 105 在 MML 文件的指導之下，可請求一使用者 ID 與

使用者的聲音樣本。在一實施例中，當提供聲音樣本時，該文件可提供一文字稿給使用者朗讀。應注意地，任何的提示都可以是聲音的提示，無論所記錄的是聲音或 TTS，或者可為文字的提示，以顯示在客戶端裝置 105 的一顯示螢幕上。

客戶端裝置 105 可製作使用者聲音的一數位記錄，例如請使用者大聲地朗讀文字稿。一旦客戶端裝置 105 獲得使用者 ID 及一使用者的聲音記錄，客戶端裝置 105 在 MML 文件 140 的導引下，可將這兩者傳送至多型態應用程式 120。記錄 145 與使用者 ID 150 可幾乎在同一時間被傳送。一旦接收使用者 ID 150 與記錄 145，多型態應用程式 120 可提供兩者至辨認系統 125 供處理。辨認系統 125 利用所提供的使用者 ID 150 可找出聲紋資料庫 130 中的一聲紋。然後，辨認系統 125 可將記錄 145 與聲紋相比較，以判斷是否記錄 145 符合聲紋。如果符合，使用者即可獲准存取一線上系統，例如供銀行、訂位、或其他類似之一網站。如果不符合，則拒絕使用者存取。

在另一實施例中，聲音並不需要被記錄在客戶端裝置 105，而是需要來自客戶端裝置 105 的串流 (streamed)，透過資料連接串流到多型態應用程式 120。行動裝置如透過資料通道支援聲音的手機，已知

的如 VoIP，可支援聲音的串流。在有些情況，聲音是來自客戶端裝置 105 的串流，這種聲音可由多型態應用程式 120 記錄。

圖 2 係依據揭露之發明的安排之一流程圖，闡述記錄一使用者之方法供辨認使用者之一方法 200。參照如圖 1 所述的系統或其他適當的系統，可執行方法 200。方法 200 可開始於步驟 205，其中一個線上系統上的一個新使用者，可存取一多型態、基於網路的應用程式。該使用者可透過一客戶端裝置存取多型態、基於網路的應用程式。

在步驟 210 中，多型態、基於網路的應用程式可提供一安全的 MML 文件給客戶端裝置的使用者。應注意，MML 文件可以具體指出一基於網路的多型態界面或網頁。如此，其可具體指出供接收使用者 ID 及其他可能需要的個人資訊之欄位，如電話號碼、地址、電子郵件地址、及其他類似者。MML 文件更可具體指出登錄文字稿。該登錄文字稿可包括文字、數字、符號、字母、及其他類似者。

在步驟 215 中，在呈現 MML 文件之後，客戶端裝置可提示使用者大聲地朗讀登錄文字稿。該提示可為基於文字的提示，如網頁中寫入指令，或亦可為一聲音提示，不

論是 TTS 或預先錄製的聲音。該登錄文字稿可以文字顯示在客戶端裝置的顯示螢幕上。在步驟 220 中，使用者可以大聲地朗讀登錄文字稿，如語音登錄文字稿。在步驟 225 中，客戶端裝置記錄著使用者大聲地朗讀文字稿，並獲得使用者 ID 及其他任何使用者所提供的個人資訊。依據本發明之一實施例，MML 文件中的一記錄指令可指示客戶端裝置製作使用者語音的一數位錄音。例如，在 MML 文件中寫入 X+V，即可使用記錄標籤(record tag)。

步驟 230 中，當使用者結束朗讀登錄文字稿時，客戶端裝置可將其所錄製的聲音轉送到在伺服器執行的多型態應用程式。除了所錄製的聲音之外，任何其他輸入 MML 文件欄位中的資訊，尤其是使用者 ID，都可被傳送到多型態、基於網路的應用程式。應被了解，從 MML 文件欄位所獲得的資料以及所錄製的聲音可以在同一時間被傳送，或者幾乎是在同時地被傳送。於執行時，MML 文件特地讓客戶端裝置轉送這類的資訊。

應注意的是，在另一實施例中，登錄及/或辨認聲音可以用串流，從客戶端裝置串流到多型態、基於網路的應用程式。在本實施例中，多型態、基於網路的應用程式可以錄製聲音。在此情形裡，MML 文件中的一串流指令可指示，使聲音的串流從客戶端裝置串流到多型態、基於網路的應用程式。

在步驟 235 中，多型態、基於網路的應用程式可接收在客戶端裝置輸入到 MML 文件中的聲音或任何其他資訊，例如使用者 ID。在步驟 240 中，多型態、基於網路的應用程式將使用者 ID、聲音、登錄文字稿傳送到辨認系統。在步驟 245 中，辨認系統可從聲音與登錄的文字稿，以建立一個使用者之獨一無二的聲紋。藉此所產生的聲紋可關於使用者 ID，並被儲存在聲紋資料庫之內。使用者成功地登錄之後，在步驟 250 中，多型態、基於網路的應用程式可傳送一個安全的歡迎網頁給使用者，並允許使用者在伺服器或另一個伺服器內安全地存取網頁。

圖 3 係依據揭露之發明的安排之一流程圖，闡述執行辨認使用者之一方法 300。方法 300 可參考圖 1 所述的系統，或其他相容的系統而被執行。在步驟 305 中，一個已註冊的使用者可透過一客戶端裝置存取多型態、基於網路的應用程式，以請求一網頁或請求登入系統。在步驟 310 中，多型態應用程式可傳送一 MML 文件，在這種情形係傳送安全之基於網路的登錄網頁，其可藉由客戶端裝置在步驟 315 中呈現。

該安全的登錄網頁可以有一或多個欄位供接收使用者之輸入，如一使用者 ID。安全的登錄網頁更可指定一文字稿。該文字稿可為一串隨機的字母、數字、及/或文字。

多型態、基於網路的應用程式可隨機地產生文字稿，例如從這類的文句之一集合中隨機地選出文字、字母、及/或數字。在一實施例中，文字稿可包括從登錄文字稿中隨機選出的任何文字、字母、或數字。

隨機產生的文字稿可防止冒充者企圖騙過辨認系統。如果一冒充者想要使用預先錄製的一已註冊者之語音，預先錄製的語音與使用於聲紋中之預期的使用者語音之比對將不符合。預先錄製的使用者語音將包含有不同於隨機產生的指定文字稿。因此，冒充者的使用者辨認將會失敗。所以依據本發明之一實施例，聲音辨認之外可附加上語音辨識，以防止冒充者利用預先錄製使用者的語音來欺騙辨認系統。

在步驟 320 中，使用者可將其使用者 ID 輸入至顯示於客戶端裝置之 MML 文件中適當的欄位。使用者可利用語音、鍵盤輸入、輸入筆或其他類似者輸入其使用者 ID。無論使用者使用哪種特殊的方式輸入其使用者 ID 都並非用來限制本發明。在任何的情況之下，使用者 ID 都可暫時地被儲存在客戶端裝置。

在步驟 325 中，可導引使用者大聲地朗讀顯示在客戶端裝置的文字稿。可藉由一 TTS 提示、一聲音提示、或由所呈現的 MML 文件顯示的文字提示，提示使用者。在步

驟 330 中，客戶端裝置可將使用者大聲朗讀文字稿的內容製作成一數位錄音。在一實施例中，該數位錄音可被儲存成 PCM 資料。然而應了解的是，特定的數位語音檔案格式並非係用以限制本發明，任何各式各樣的數位語音格式，如 ADPCM、MP3、AAC 或被壓縮的 DSR 等都可被使用，其係以聲音串流，透過一資料連結至多型態、基於網路的應用程式。

在步驟 335 中，當使用者結束說話，客戶端裝置可傳送文字稿及使用者 ID 至多型態、基於網路的應用程式。無論是一輸入錄音或聲音的串流，客戶端裝置幾乎是同時地提供使用者 ID 與聲音，即辨認聲音，且多型態、基於網路的應用程式幾乎是同時地接收使用者 ID 與聲音。在步驟 340 中，多型態、基於網路的應用程式可接收聲音及使用者 ID。在步驟 345 中，多型態、基於網路的應用程式可轉送聲音、文字稿及使用者 ID 至辨認系統。

在步驟 350 中，辨認系統依據使用者 ID，從聲紋資料庫中找出或擷取出使用者的聲紋。在步驟 355 中，辨認系統可將聲音與符合使用者 ID 的聲紋相比對，以判斷該聲音是否符合聲紋。如果符合，則來到步驟 360，多型態、基於網路的應用程式提供一個安全的歡迎網頁給在客戶端裝置的使用者，該安全的歡迎網頁可為傳統的視覺網頁或是多型態網頁。經確認過的使用者即被允許存取該網站

或其他安全的網頁。如果所錄製的聲音與聲紋不相符，在步驟 365 中，由多型態、基於網路的應用程式拒絕該使用者進行存取。

本發明提供一解決方案，供使用基於多型態、基於網路之介面以辨認一使用者身份。依據本發明，語者辨認系統依據實際講話者的聲音特性，以辨認該講話者。透過多型態的介面，使用者可為已經註冊者，亦或可為利用辨認系統事後確認者。顯示登錄文字稿予使用者作為視覺文字，因此使用者無需記住一長串的通關文字，將促進登錄之便利性。

語者辨認的部份需要使用者唸出由多型態應用程式所提供之隨機文字稿。隨機文字稿可防止冒充者利用預先錄製之另一使用者的語音來欺騙辨認系統。預先錄製的語音，將無法符合依據隨機產生的文字稿指定之預期之使用者語音。用來辨認一使用者身份的文字稿，類似於登錄文字稿，亦可視覺地被呈現。利用多型態界面的辨認，相較於單一型態界面係更安全的，因其同時需要密碼及語音辨認兩者。

本發明可為硬體、軟體或硬體與軟體之結合。本發明還可為一集中式的方式存在於一電腦系統中，或一分散式的方式，其中不同元件可分散於數個彼此互連的電腦系

統。只要是可以實現本發明於此描述之方法，何形式的電腦系統或其他設備都是合適的。典型的硬體與軟體之結合可為一般用途的一電腦系統，係被安裝並執行一電腦程式，控制該電腦系統以實現本發明於此所述之方法。

本發明亦可內建於一電腦程式產品，其包括實施本發明於此所述之方法的所有特徵，且當其被安裝於電腦系統時，可實現本發明之方法。本文中，電腦程式、軟體應用程式、及/或其他各種名詞，以任何的語言、程式碼、或標記(notation)，表一組指示之任何表示(expression)，以產生一系統，係具有一處理資訊能力，以直接或間接地執行下列特定的功能之一或兩者：a)轉換成另一種語言、程式碼、或標記；或 b)再製成為不同材料的形式。

在不脫離本發明之精神和必要特徵內，本發明可以其他形式被實施。因此，當參照後附之申請專利範圍界定本發明之範圍，而非前述之說明書部份。

【圖式簡單說明】

本發明之圖式所顯示較佳實施例，然而應被了解的是，本發明並非以所示的確切圖式安排及結構所限制。

圖 1 係依據揭露之發明的安排之一概要圖，闡述使用一基於網路多型態介面執行使用者辨認之一系統。

圖 2 係依據揭露之發明的安排之一流程圖，闡述記錄一使用者之方法供辨認使用者之一方法。

圖 3 係依據揭露之發明的安排之一流程圖，闡述執行辨認使用者之一方法。

【主要元件符號說明】

系統 100

伺服器 110

多型態應用程式 120

聲紋資料 130

MML 文件 140

使用者 ID 150

客戶端裝置 105

辨認系統 125

請求 135

記錄 145

方法 200、300

五、中文發明摘要：

一種使用多型態、基於網路之介面以辨認一使用者身分之方法，其可包含傳送一多型態介面標記語言文件(markup language document)至一遠端電腦裝置，當遠端電腦裝置呈現(rendered)時，請求辨認一使用者之身分(identifier)並使該使用者之聲音傳送到一多型態、基於網路的應用程式(multimodal Web-based application)。該使用者之身分辨認(identifier)及聲音幾乎是同時地從客戶端裝置被接收。將該聲音與一聲紋相比較，且該聲紋係與使用者之身分辨認相關聯。依據上述之比較所獲得之一結果，可選擇性地允許在遠端電腦裝置之使用者存取系統。

六、英文發明摘要：

A method of verifying a user identify using a Web-based multimodal interface can include sending, to a remote computing device, a multimodal markup language document that, when rendered by the remote computing device, queries a user for a user identifier and causes audio of the user's voice to be sent to a multimodal, Web-based application. The user identifier and the audio can be received at about a same time from the client device. The audio can be compared with a voice print associated with the user identifier. The user at the remote computing device can be selectively granted access to the system according to a result obtained from the comprising step.

100

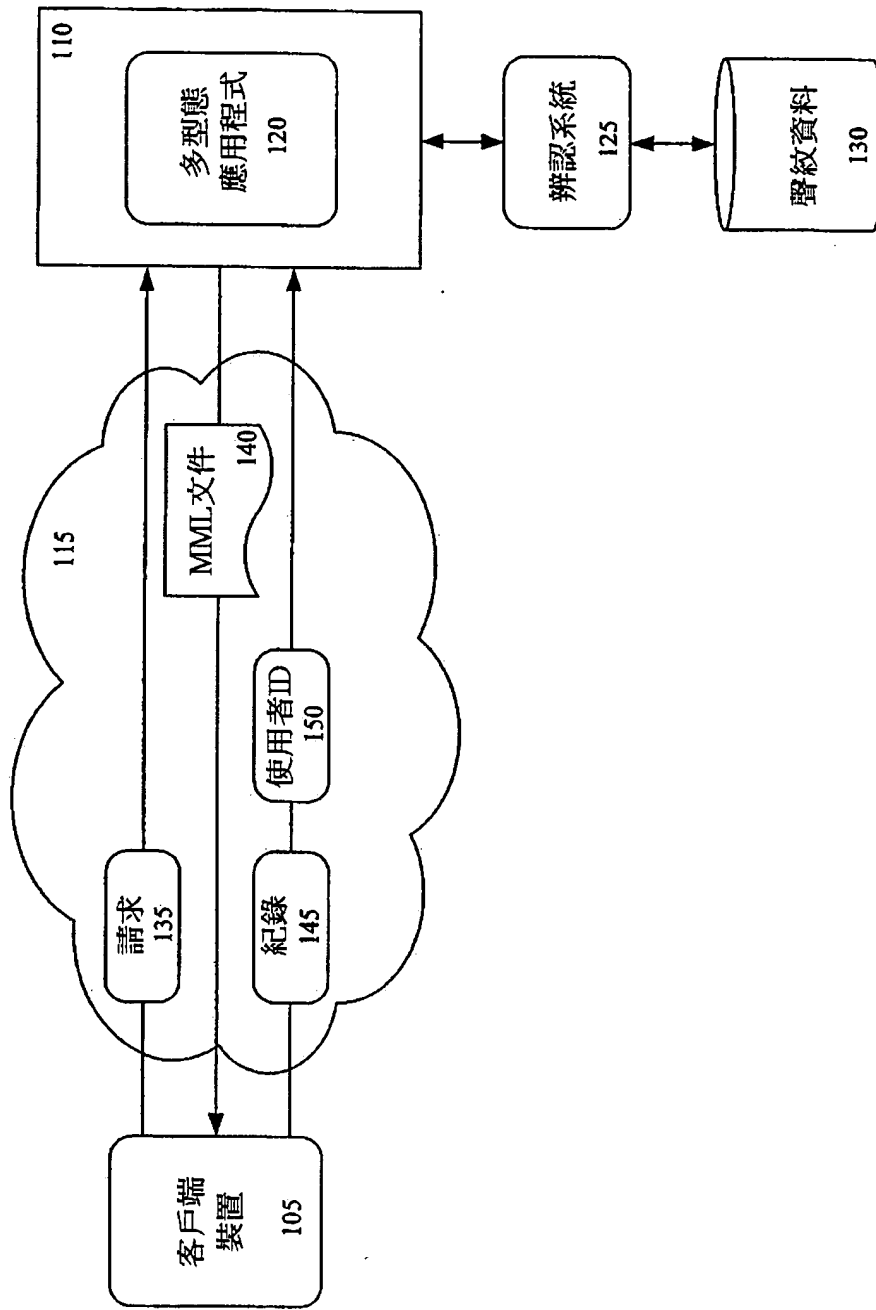


圖1

200

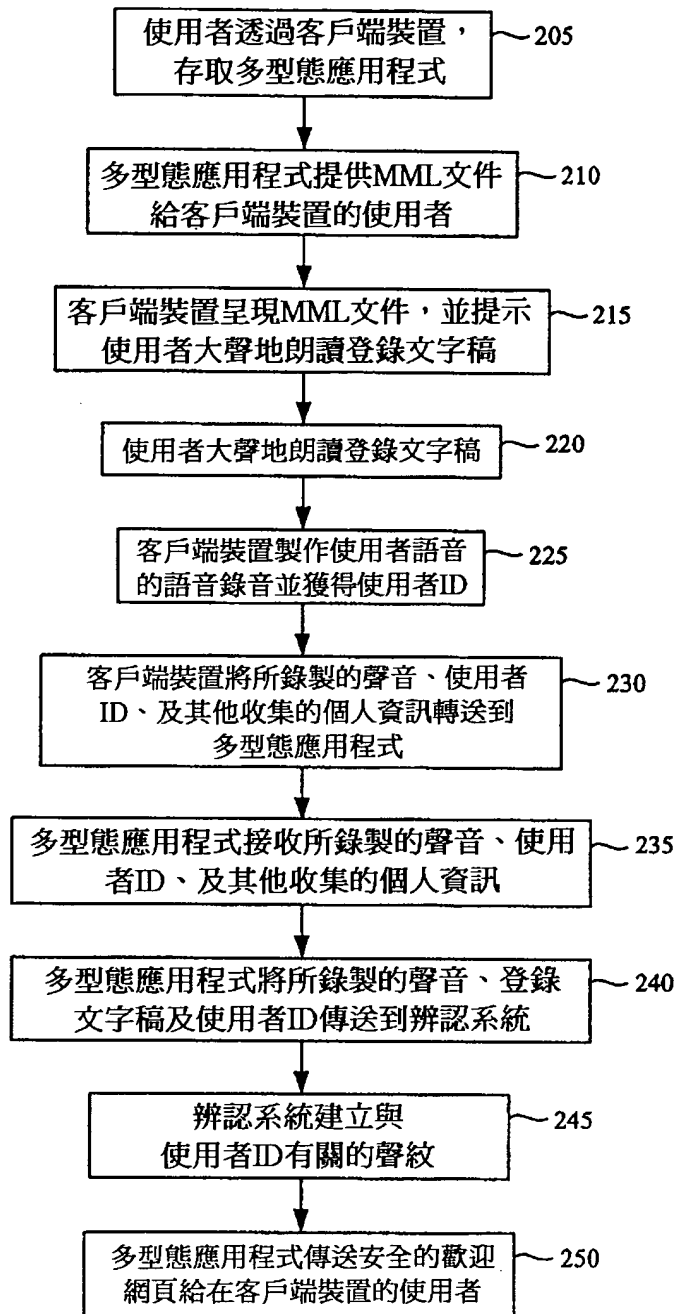


圖2

300

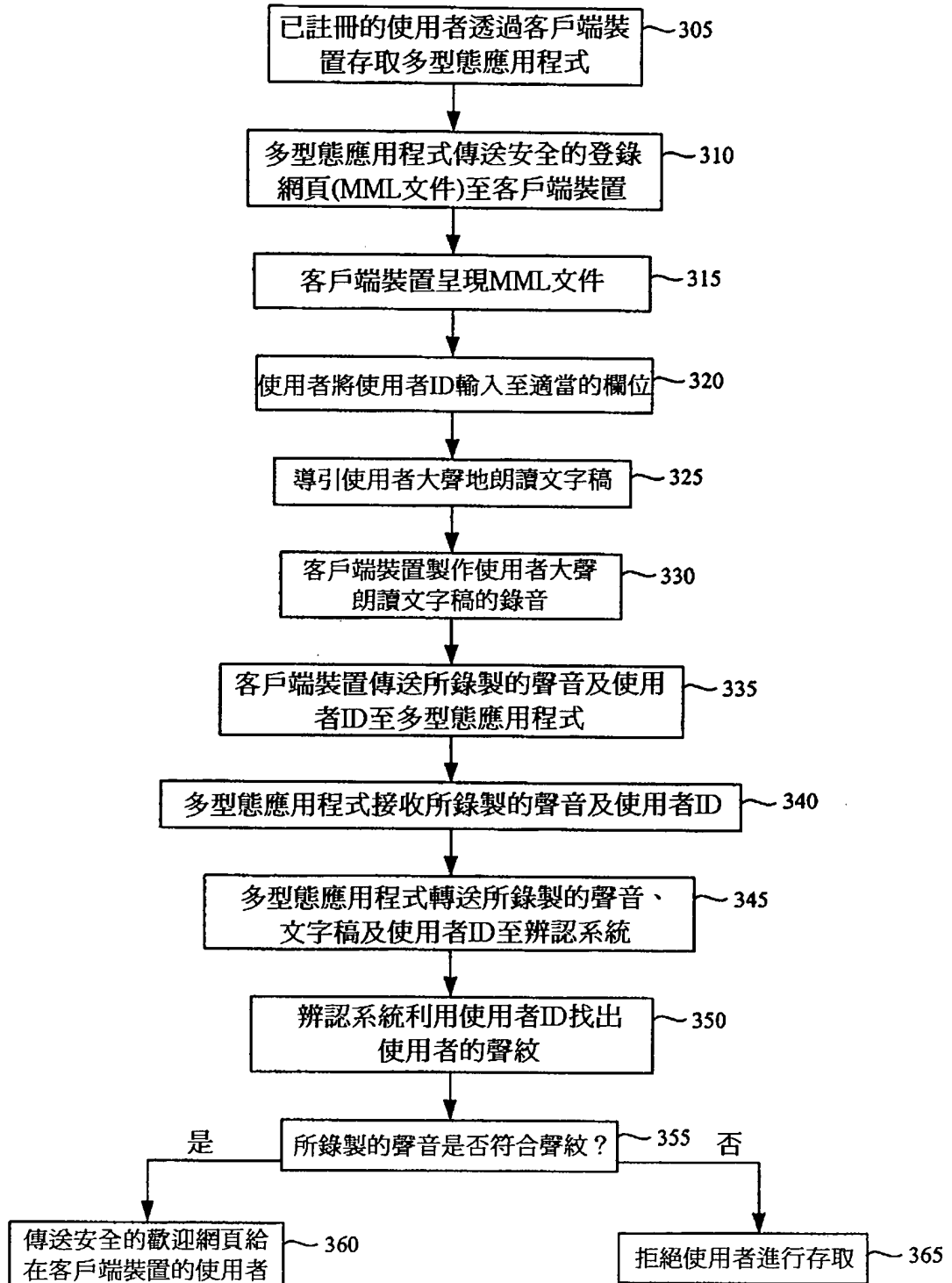


圖3

七、指定代表圖：

(一)本案指定代表圖為：圖 1。

(二)本代表圖之元件符號簡單說明：

- 100. 系統
- 115. 客戶端裝置
- 110. 伺服器 110
- 120. 多型態應用程式
- 125. 辨認系統
- 130. 聲紋資料
- 135. 請求
- 140. MML 文件
- 145. 記錄
- 150. 使用者 ID

八、本案若有化學式時，請揭示最能顯示發明特徵的
化學式：無

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：95104025

※ 申請日期：95年02月07日

※IPC 分類：H04L9/32 (sub.op)

一、發明名稱：(中文/英文)

使用語者辨認及基於網路之多型態介面以辨認使用者

VERIFYING A USER USING SPEAKER VERIFICATION AND
A MULTIMODAL WEB-BASED INTERFACE

二、申請人：(共 1 人)

姓名或名稱：(中文/英文)(簽章)

紐安斯通訊公司

NUANCE COMMUNICATIONS, INC.

代表人：(中文/英文)(簽章)

琳恩 J. 費茲傑拉德 / FITZGERALD, LEANNE J.

住居所或營業所地址：(中文/英文)

美國麻省伯靈頓市道邊路 1 號

One Wayside Road, Burlington, MA 01803, USA

國籍：(中文/英文)

美國 / US

三、發明人：(共 2 人)

姓名：(中文/英文)

1. 大衛 傑洛米歐 / JARAMILLO, DAVID

2. 麥克柏 傑洛德 馬修 / MCCOBB, GERALD MATTHEW

國籍：(中文/英文)

1-2 皆為美國 / US

十、申請專利範圍：

1. 一種使用多型態 (multimodal)、基於網路之介面以辨認一使用者身分之方法，包含：

從一遠端電腦裝置接收用於使用者證明的一多型態介面標記語言文件 (markup language document)之一請求；

傳送該多型態介面標記語言文件 至該遠端電腦裝置，當該遠端電腦裝置呈現 (rendered) 時，請求一使用者身分辨認 (identifier) 以及該使用者之聲音，其中該多型態介面標記語言文件指定一文字稿以供該使用者大聲地朗讀；

接收幾乎是同時來自該遠端電腦裝置之該使用者身分辨認及聲音於一多型態、基於網路的應用程式 (multimodal Web-based application)；

比較該聲音與一和該使用者身分辨認相關聯之聲紋 (voice print)；以及

依據上述比較步驟中所獲得之一結果，選擇性地允許該使用者在該遠端電腦裝置存取一線上系統。

2. 如請求項1所述之方法，其中該多型態介面標記語言文件包括一指令使該聲音產生串流，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令，從該遠端電腦裝置使該聲音的串流，流

到該多型態、基於網路的應用程式。

3. 如請求項1所述之方法，其中該多型態介面標記語言文件包括一指令，係用以錄製該聲音，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令，使該遠端電腦裝置製作該使用者聲音之一數位錄音。
4. 如請求項1所述之方法，其中該文字稿包括至少一部份隨機產生的文字。
5. 如請求項1所述之方法，其中該聲音係為該使用者大聲地朗讀該文字稿之一數位代表 (representation)。
6. 如請求項5所述之方法，該比較步驟進一步包含：該多型態、基於網路的應用程式首先傳送該聲音、該文字稿及該使用者身分辨認至一辨認系統。
7. 如請求項1所述之方法，進一步包含：

首先傳送一多型態介面標記語言文件至該遠端電腦裝置，係用以指定一登錄(enrollment)文字稿；

接收該使用者之身分辨認及聲音，係包含該使

用者朗讀該登錄文字稿之一數位代表；以及

藉由判斷一聲紋與該使用者朗讀該登錄文字稿之聲音，及相關聯該聲紋係與該使用者身分辨認，以登錄該使用者至一線上系統。

8. 一種辨認一使用者身分 (identity) 之系統，包含：

一辨認系統，其具有一聲紋資料庫，係包括複數聲紋，每一聲紋都與一使用者身分辨認相關聯；

一多型態應用程式，被組態以接收來自一客戶端裝置用於使用者證明的一多型態介面標記語言文件的一請求以及傳送一多型態介面標記語言文件至該客戶端裝置，該多型態介面標記語言文件指定一文字稿以供該使用者大聲地朗讀以及當該客戶端裝置呈現 (rendered) 時，使該使用者身分辨認及聲音被傳送回至該多型態應用程式，其中該聲音以及該使用者身份辨認幾乎是同時從該客戶端裝置提供至該多型態應用程式並且係基於該多型態介面標記語言文件的控制；以及

其中該多型態應用程式能夠提供該聲音、以及該使用者身分辨認給該辨認系統，比較該聲音與一聲紋以辨認該說話者之身分，而該聲紋係與該使用者身分辨認相關聯。

9. 如請求項8所述之系統，其中該聲音與該使用者身

分辨認幾乎同時地從該客戶端裝置被提供給該多型態應用程式。

10. 如請求項8所述之系統，其中該多型態介面標記語言文件包括一指令使該聲音產生串流，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令，從該遠端電腦裝置使該聲音的串流，流到該多型態、基於網路的應用程式。
11. 如請求項8所述之系統，其中該多型態介面標記語言文件包括一指令，係用以錄製該聲音，藉以達成：回應呈現該多型態介面標記語言文件及轉譯該指令，使該遠端電腦裝置錄製該使用者大聲地朗讀該文字稿。
12. 如請求項8所述之系統，其中該文字稿包括至少一部份隨機產生的文字。
13. 如請求項8所述之系統，其中該多型態介面標記語言文件包括一指令，係用以串流該聲音，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令以串流該聲音，使該遠端電腦裝置串流該聲音到該多型態應用程式。

14. 一種機器可讀取之儲存器，其儲存一具有由一機器執行之複數程式碼段 (code section) 之電腦程式，供該機器執行以下之步驟：

傳送一多型態介面標記語言文件至一遠端電腦裝置，當該遠端電腦裝置呈現時，請求一使用者身分辨認指定一文字稿以供該使用者大聲地朗讀，並使該使用者之聲音傳送到一多型態、基於網路的應用程式；

接收幾乎是同時來自該客戶端裝置之該使用者身分辨認及聲音；

比較該聲音與一和該使用者身分辨認相關聯之聲紋；以及

依據上述比較步驟中所獲得之一結果，選擇性地允許該使用者在該遠端電腦裝置存取一線上系統。

15. 如請求項14所述之機器可讀取之儲存器，其中該多型態介面標記語言文件包括一指令使該聲音產生串流，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令，從該遠端電腦裝置使該聲音的串流，流到該多型態、基於網路的應用程式。

16. 如請求項14所述之機器可讀取之儲存器，其中該多

型態介面標記語言文件包括一指令，係用以錄製該聲音，該方法進一步包含：回應呈現該多型態介面標記語言文件及轉譯該指令，使該遠端電腦裝置製作該使用者聲音之一數位錄音。

17. 如請求項14所述之機器可讀取之儲存器，其中多型態介面標記語言文件包括指定一文字稿以供該使用者大聲地朗讀。
18. 如請求項15所述之機器可讀取之儲存器，其中該文字稿包括至少一部份隨機產生的文字。
19. 如請求項17所述之機器可讀取之儲存器，其中該聲音包含該使用者大聲地朗讀該文字稿之一數位代表。
20. 如請求項19所述之機器可讀取之儲存器，該比較步驟進一步包含：該多型態、基於網路的應用程式首先傳送該聲音、該文字稿及該使用者身分辨認至一辨認系統。
21. 如請求項14所述之機器可讀取之儲存器，進一步包含：
首先傳送一多型態介面標記語言文件至該遠

端電腦裝置，係用以指定一登錄文字稿；

接收該使用者之身分辨認及聲音，係包含該使用者朗讀該登錄文字稿之一數位代表；以及

藉由判斷一聲紋與該使用者朗讀該登錄文字稿之聲音，及相關聯該聲紋係與該使用者身分辨認，以登錄該使用者至一線上系統。