



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I490778 B

(45) 公告日：中華民國 104 (2015) 年 07 月 01 日

(21) 申請案號：102113029 (22) 申請日：中華民國 102 (2013) 年 04 月 12 日

(51) Int. Cl. : G06F3/16 (2006.01) G06K9/78 (2006.01)
G10L17/22 (2013.01) G10L25/51 (2013.01)

(30) 優先權：2012/04/27 世界智慧財產權組織 PCT/US12/35538

(71) 申請人：惠普研發公司 (美國) HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.

(US)

美國

(72) 發明人：西斯勒 伊凡 SCHEESSELE, EVAN (US)

(74) 代理人：閻啟泰；林景郁

(56) 參考文獻：

TW	200720991A	TW	201033852A1
TW	201216115A1	CN	101674363A
US	2006/0192775A1	US	2010/0312547A1
US	2011/0310001A1		

審查人員：許人偉

申請專利範圍項數：15 項 圖式數：5 共 26 頁

(54) 名稱

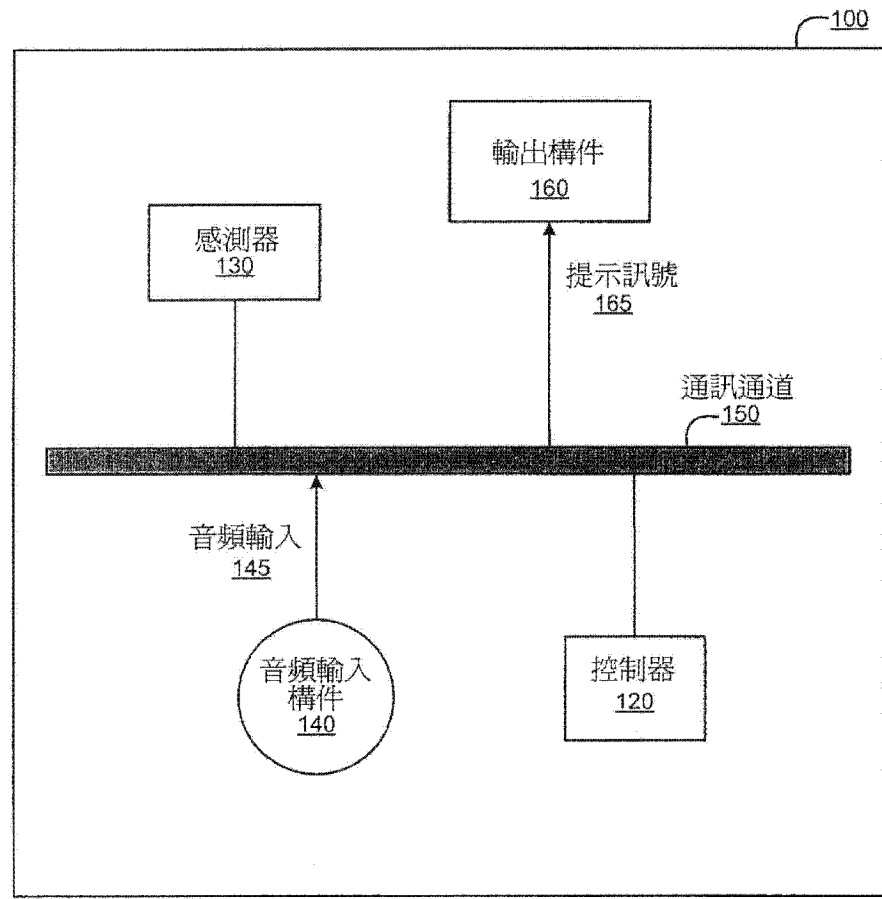
來自使用者的音頻輸入

AUDIO INPUT FROM USER

(57) 摘要

一種計算裝置係偵測一觀看該計算裝置的使用者，並且若該使用者被偵測到為觀看該計算裝置時，輸出一提示訊號。若該使用者持續觀看該計算裝置一段預設的時間量，則該計算裝置係接收來自該使用者的一音頻輸入。

A computing device detects a user viewing the computing device and outputs a cue if the user is detected to view the computing device. The computing device receives an audio input from the user if the user continues to view the computing device for a predetermined amount of time.



- 100 . . . 計算裝置
- 120 . . . 控制器
- 130 . . . 感測器
- 140 . . . 音頻輸入構件
- 145 . . . 音頻輸入
- 150 . . . 通訊通道
- 160 . . . 輸出構件
- 165 . . . 提示訊號

圖1

公告本

發明摘要

※ 申請案號： 102113029

※ 申請日： 102.4.12

※IPC 分類：

G06F 3/16 2006.01

G06K 9/18 2006.01

G10L 17/52 2013.01

G10L 25/51 2013.01

【發明名稱】(中文/英文)

來自使用者的音頻輸入

AUDIO INPUT FROM USER

【中文】

一種計算裝置係偵測一觀看該計算裝置的使用者，並且若該使用者被偵測到為觀看該計算裝置時，輸出一提示訊號。若該使用者持續觀看該計算裝置一段預設的時間量，則該計算裝置係接收來自該使用者的一音頻輸入。

【英文】

A computing device detects a user viewing the computing device and outputs a cue if the user is detected to view the computing device. The computing device receives an audio input from the user if the user continues to view the computing device for a predetermined amount of time.

【代表圖】

【本案指定代表圖】：第（ 1 ）圖。

【本代表圖之符號簡單說明】：

100：計算裝置

120：控制器

130：感測器

140：音頻輸入構件

145：音頻輸入

150：通訊通道

160：輸出構件

165：提示訊號

【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，請勿任意更動)

【發明名稱】(中文/英文)

來自使用者的音頻輸入/ Audio Input from User

【技術領域】

【0001】 本發明係關於用於接收音頻輸入之方法及其計算裝置。

【先前技術】

【0002】 當使用一裝置以輸入一例如是音頻輸入的輸入時，使用者最初可以使用該計算裝置的一例如是滑鼠或鍵盤的輸入構件，以致能該計算裝置的一音頻輸入構件。該使用者可以繼續使用該滑鼠或鍵盤以在該計算裝置上啟動一應用程式，以接收來自該使用者的音頻輸入。當該使用者操縱該輸入構件時，時間可能被浪費掉，而且該使用者可能會犯錯，此導致令人沮喪的使用者經驗。

【發明內容】

【0003】 根據一第一觀點，本申請案提供一種計算裝置，其係包括：一用以偵測一觀看該計算裝置的使用者的感測器；一用以在該使用者被偵測到為觀看該計算裝置時輸出一提示訊號的輸出構件；以及一控制器，其係用以在該感測器偵測到該使用者繼續觀看與被輸出的該提示訊號相關的該計算裝置一段預設的時間量時，利用一音頻輸入構件以接收來自該使用者的一音頻輸入。

【0004】 根據一第二觀點，本申請案提供一種用於接收音頻輸入之方法，其係包括：利用一感測器偵測觀看一計算裝置之一使用者；若該使用者正在觀看該計算裝置，則利用一輸出構件以輸出一提示訊號；判斷該使

用者是否正在觀看該輸出構件一段預設的時間量；以及若該使用者正在觀看該輸出構件該段預設的時間量，則啓動該計算裝置的一傾聽模式以接收來自該使用者的一音頻輸入。

【0005】 根據一第三觀點，本申請案提供一種包括指令之非揮發性電腦可讀取媒體，若執行該些指令，其係使得一控制器以：利用一感測器來偵測觀看一計算裝置的一使用者；若該使用者正在觀看該計算裝置，則利用一輸出構件來輸出一提示訊號；偵測該使用者是否持續觀看與被輸出的該提示訊號相關的該輸出構件一段預設的時間量；以及若該使用者正在觀看該輸出構件該段預設的時間量，則啓動該計算裝置的一傾聽模式以接收來自該使用者的一音頻輸入。

【圖式簡單說明】

【0006】 所揭露的實施例的各種特點及優點從以下結合所附的圖式所做的詳細說明將會是明顯的，該詳細說明與圖式係一起描繪例如是所揭露的實施例的特點。

圖 1 係描繪根據一例子的一種具有一感測器、一輸出構件以及一音頻輸入構件之計算裝置。

圖 2 係描繪根據一例子的一輸出提示訊號的輸出構件以及一偵測觀看計算裝置的使用者的感測器。

圖 3 係描繪根據一例子的一偵測來自使用者的音頻輸入之語音應用程式的方塊圖。

圖 4 是描繪根據一例子的一種用於偵測音頻輸入之方法的流程圖。

圖 5 是描繪根據一例子的一種用於偵測音頻輸入之方法的流程圖。

【實施方式】

【0007】 一種計算裝置係包含一例如是影像捕捉構件的感測器，以偵測觀看該計算裝置的使用者。爲了此申請案之目的，若該感測器偵測到使用者的臉或眼睛面對該計算裝置，則該使用者是正在觀看該計算裝置。若該使用者正在觀看該計算裝置，則一例如是視覺指示器、音頻輸出揚聲器及/或觸覺回授的馬達之輸出構件可以提供一提示訊號給該使用者。該提示訊號可以是一視覺、音頻及/或觸覺回授的提示訊號，其係通知該使用者該計算裝置係備妥以接收來自該使用者的音頻輸入。當該提示訊號被輸出時，該計算裝置可以使用該感測器以偵測觀看該計算裝置及/或該輸出構件的使用者一段預設的時間量。

【0008】 若該計算裝置判斷該使用者在該提示訊號被輸出時正繼續觀看該計算裝置該預設的時間量，則該計算裝置可以確認該使用者想要和該計算裝置介接。該計算裝置的一例如是麥克風的音頻輸入構件接著可以接收來自該使用者的一音頻輸入。在一實施例中，當接收音頻輸入時，該計算裝置係啓動該計算裝置的一傾聽模式以致能、取消靜音及/或校準該音頻輸入構件，以接收來自該使用者的音頻輸入。該些音頻輸入可以是一語音命令、一語音詢問及/或任何由該使用者產生並且藉由該計算裝置接收作爲一輸入之額外的聲音。因此，該使用者可以在不實際使用或操縱該計算裝置的輸入構件下，便利地利用音頻輸入和該裝置互動。

【0009】 圖 1 係描繪根據一例子的一種具有一感測器 130、一輸出構件 160 以及一音頻輸入構件 140 之計算裝置 100。在一實施例中，該計算裝置 100 可以是一桌上型電腦、一膝上型電腦、一筆記型電腦、一平板電腦、

一小筆電、一個全一體化系統及/或一伺服器。在另一實施例中，該計算裝置 100 可以是一蜂巢式裝置、一 PDA(個人數位助理)、一電子書、一平板電腦、一攝影機、及/或任何可包含一感測器 130、一輸出構件 160 以及一音頻構件 140 之額外的裝置。

【0010】 該計算裝置 100 係包含一控制器 120、一感測器 130、一輸出構件 160、一音頻輸入構件 140 以及一用於該計算裝置 100 及/或該裝置 100 的一或多個構件彼此通訊的通訊通道 150。在一實施例中，該計算裝置 100 亦包含一儲存在一內含在該計算裝置 100 中或是該計算裝置 100 可存取之非揮發性電腦可讀取的媒體上的語音應用程式。爲了此申請案之目的，該語音應用程式是一可以獨立地被利用及/或結合該控制器 120 被利用的應用程式，以管理該計算裝置 100 的音頻輸入 145。爲了此申請案之目的，音頻輸入 145 係包含一來自使用者而被接收作爲一用於該計算裝置 100 的輸入之語音及/或聲音。

【0011】 當管理音頻輸入 145 時，一感測器 130 最初可被用來偵測一使用者是否正在觀看該計算裝置 100。爲了此申請案之目的，該感測器 130 是一例如爲影像捕捉構件及/或紅外線構件的硬體構件，其可以捕捉在該計算裝置 100 的前面之一視景以偵測一觀看該計算裝置 100 的使用者。偵測觀看該計算裝置 100 的使用者係包含判斷該使用者的臉是否面對該計算裝置 100。在一實施例中，該控制器 320 及/或該語音應用程式可以使用人臉偵測技術及/或視線偵測技術以判斷該使用者的臉或眼睛是否正在觀看該計算裝置 100。

【0012】 若該使用者被偵測到是在觀看該計算裝置 100，則該控制器

120 及/或該語音應用程式係指示一輸出構件 160 來輸出一提示訊號 165 給該使用者。爲了此申請案之目的，一提示訊號 165 是一來自該輸出構件 160 的信號或通知，以通知該使用者該計算裝置 100 係備妥以接收來自該使用者的音頻輸入 145。該輸出構件 160 是該計算裝置 100 的一硬體構件，其可以提供一視覺的提示訊號、一音頻提示訊號及/或一觸覺回授的提示訊號給該使用者。在一實施例中，該輸出構件 160 可以是一視覺指示器、一音頻輸出揚聲器、一觸覺馬達及/或任何可以輸出一提示訊號 165 之額外的構件。

● **【0013】** 該控制器 110 及/或該語音應用程式可以使用該感測器 130 以偵測該使用者是否正在觀看該計算裝置 100 一段預設的時間量。該預設的時間量可以藉由該控制器 120、語音應用程式及/或該使用者來界定。例如，該預設的時間量可被界定爲半秒、2 秒、或是任何另外的時間量。若該感測器 130 偵測到該使用者繼續觀看該計算裝置 100 該預先界定的時間量，則該控制器 110 及/或該語音應用程式係判斷該使用者正在觀看該計算裝置 100 該預設的時間量。

● **【0014】** 在另一實施例中，當偵測該使用者觀看該計算裝置 100 該預設的時間量時，該感測器 130 可進一步偵測該使用者觀看該計算裝置 100 的輸出構件 160。該控制器 120 及/或該語音應用程式可以辨識該使用者的一視線來判斷該使用者是否正在觀看該計算裝置 100 及/或該輸出構件 160。該控制器 120 及/或該語音應用程式可以使用人臉偵測技術及/或視線偵測技術以辨識該使用者的視線。

【0015】 若該使用者持續觀看該計算裝置 100 及/或該輸出構件 160 該預設的時間量，則該控制器 120 及/或該語音應用程式可以判斷該使用者

的意圖是要和該計算裝置 100 互動，並且繼續利用一音頻輸入構件 140 以接收來自該使用者的音頻輸入 145。爲了此申請案之目的，該音頻輸入構件 140 是一例如是麥克風的硬體構件，其可以接收來自該使用者的音頻輸入 145。

【0016】 爲了此申請案之目的，一音頻輸入 145 係包含來自該使用者之可聽見的聲音，該控制器 110 及/或該語音應用程式可以接收該聲音作爲一用於該計算裝置 100 的輸入。例如，該音頻輸入 145 可以是一傳送給該計算裝置 100 的語音命令及/或一詢問。在一實施例中，接收音頻輸入 145 係包含該控制器 120 及/或該語音應用程式啓動該計算裝置 100 的一傾聽模式。爲了此申請案之目的，該計算裝置 100 的傾聽模式係對應於該計算裝置 100 的一輸入模式，以接收來自該使用者的音頻輸入 145。若該計算裝置 100 是在該傾聽模式，則該音頻輸入構件 140 係被致能並且取消靜音以接收該音頻輸入 145。

【0017】 圖 2 係描繪根據一例子，一輸出構件 260 係回應於一感測器 230 偵測到一觀看計算裝置 200 的使用者 205 來輸出一提示訊號 265。爲了此申請案之目的，該感測器 230 是該計算裝置 200 的一硬體構件，其係偵測觀看該計算裝置 200 的使用者 205。該感測器 230 可以是一影像捕捉構件、一紅外線構件及/或任何可以偵測觀看該計算裝置 200 的使用者 205 之額外的構件。在一實施例中，該感測器 230 亦可以偵測該使用者 205 在該計算裝置 200 的周圍的身體貼近。該使用者 205 的身體貼近係對應於該使用者到該音頻輸入構件 240 的一位置或距離。該使用者 205 可以是任何可以和該計算裝置 200 互動並且輸入一或多個音頻輸入至該計算裝置 200 的人。

【0018】 該感測器 230 最初可以捕捉在該計算裝置 200 的前面的一視

景，以偵測該使用者 205。若該使用者 205 被偵測到，則該控制器及/或該語音應用程式可以使用人臉偵測技術及/或視線偵測技術來偵測該使用者 205 是否正在觀看該計算裝置 200。該控制器及/或該語音應用程式可以使用該人臉偵測技術及/或視線偵測技術來判斷該計算裝置 200 是否在該使用者 205 的視線內。若該計算裝置 200 是在該使用者的視線內，則該控制器及/或該語音應用程式係判斷出該使用者 205 正在觀看該計算裝置 200。

● 【0019】 若該使用者 205 被偵測到是正在觀看該計算裝置 200，則該輸出構件 260 繼續輸出一提示訊號 265 以通知該使用者 205 該計算裝置 200 是備妥以接收來自該使用者 205 的音頻輸入。在一實施例中，該控制器及/或該語音應用程式可進一步在指示該輸出構件 260 輸出該提示訊號 265 之前，先判斷該使用者是否持續觀看該計算裝置 200 一段預設的時間量。如上所提到的，該預設的時間量可以藉由該控制器、語音應用程式及/或藉由該使用者 205 來加以界定為半秒、2 秒及/或任何的時間量。

● 【0020】 若該使用者 205 持續觀看該計算裝置 200 該預設的時間量，則該輸出構件 260 可被利用以輸出一提示訊號 265。爲了此申請案之目的，該輸出構件 260 是該計算裝置 200 的一輸出該提示訊號 265 的硬體構件，例如一視覺指示器、一顯示器構件、一音頻輸出揚聲器及/或一觸覺馬達。該提示訊號 265 係包含一視覺的信號、一音頻信號、一觸覺回授及/或任何額外的提示訊號 265 中之至少一個，其係通知該使用者 205 該計算裝置 200 是備妥以接收音頻輸入。例如，該視覺的信號可以是一光源照明。該音頻信號可以是一被輸出之可聽見的音調。該觸覺回授可包含一或多個震動。

【0021】 在一實施例中，該輸出構件 260 是一具有有限的視野之方向

性的視覺指示器，並且該提示訊號 265 係根據該使用者 205 的位置而被輸出給該使用者 205。如上所提到的，該感測器 230 可被利用以偵測該使用者 205 的一身體位置或貼近。該方向性的視覺指示器可包含一陣列的光源，其係具有一有限的視野。在該使用者 205 的附近的一或多個光源可被點亮以輸出該提示訊號 265 以讓該使用者 205 看到。其它不在該使用者 205 的附近之光源則不被點亮。因此，該使用者 205 以及該計算裝置 200 的安全性可以增高，此係藉由在該提示訊號 265 的一有限的視野內輸出該提示訊號 265 使得其它個人並不會察覺到該提示訊號 265 並且看見該使用者 205 和該計算裝置 200 互動。

【0022】 當該提示訊號 265 藉由該輸出構件 260 被輸出時，該控制器及/或該語音應用程式係判斷該使用者 205 是否正在觀看和被輸出的提示訊號 265 相關之計算裝置 200 預設的時間量。爲了此申請案之目的，若該使用者 205 在該輸出構件 160 輸出該提示訊號 265 時持續觀看該計算裝置 200，則該使用者 205 可被判斷爲正在觀看和該提示訊號 265 相關的計算裝置 200。例如，若該使用者 205 是和該提示訊號 265 被輸出並行地持續觀看該計算裝置 200 該預設的時間量，則該使用者 205 係被判斷爲正在觀看和被輸出的提示訊號 265 相關之計算裝置 200。

【0023】 在另一實施例中，相對於觀看該計算裝置 200，該控制器及/或該語音應用程式可以偵測觀看和被輸出的提示訊號 265 相關之輸出構件 260 該預設的時間量之使用者 205。在其它實施例中，該感測器 230 亦偵測一在該使用者 205 的臉部表情或是來自該使用者 205 之持續的注視上的變化。當偵測在臉部表情上的變化時，該感測器 230 係偵測一笑容、一皺眉、

一眨眼、一挑眉及/或任何在來自該使用者 205 的臉部表情上之另外的變化。當偵測持續的注視時，該感測器 230 係偵測持續觀看該計算裝置 200 或輸出構件 260 該預設的時間量之使用者 205 的臉或眼睛。

【0024】 若該使用者 205 正在觀看和被輸出的提示訊號 265 相關之計算裝置 200 及/或輸出構件 260 該預設的時間量，則該控制器及/或該語音應用程式係判斷該使用者想要利用一音頻輸入來和該計算裝置 200 互動。如在圖 2 中所示，一音頻輸入構件 240 可被利用以偵測來自該使用者 205 的音頻輸入。爲了此申請案之目的，該音頻輸入構件 240 是一例如是麥克風的硬體構件，其係偵測語音、聲音及/或雜訊以作爲一用於該計算裝置 200 的音頻輸入。在一實施例中，來自該使用者 205 的音頻輸入可以是一用於該計算裝置 200 執行一指令或命令的語音命令。該指令或命令可以是啓動或終止一應用程式、開啓或關閉一檔案及/或聽寫被輸入到電子郵件、一應用程式及/或一網路瀏覽器中的字母與數字的文字。

【0025】 在另一實施例中，該音頻輸入可以是一針對於來自該計算裝置 200 的資訊之語音詢問。該詢問可以是針對於該計算裝置 200 的一狀態、針對於一本地的氣象預報及/或針對於可從該計算裝置 200 或是另一裝置擷取的資訊。當捕捉來自該使用者 205 的一語音命令及/或一詢問時，該感測器 230 可進一步捕捉該使用者的 205 的嘴部之一視景。該使用者 205 的嘴部的視景可被該控制器及/或該語音應用程式利用來改善辨識來自該使用者 205 的一語音命令及/或一詢問的正確性。在其它實施例中，該使用者 205 可以使用一手、一身體部分及/或一配件來產生音訊以供該音頻輸入構件 240 接收作爲一音頻輸入。

【0026】 圖 3 係描繪根據一例子的一種偵測來自一使用者的音頻輸入之語音應用程式 310 的方塊圖。如上所提到的，該語音應用程式 310 可獨立地及/或結合該控制器 320 而被利用以管理音頻輸入。在一實施例中，該語音應用程式 310 可以是一被內嵌到該計算裝置的一或多個構件之韌體。在另一實施例中，該語音應用程式 310 可以是一可從該計算裝置之一非揮發性電腦可讀取的記憶體利用之應用程式。該電腦可讀取的記憶體是一實體的裝置，其係包含、儲存、通訊或是傳輸該應用程式，以供該裝置使用或是與該裝置連線。在一實施例中，該電腦可讀取的記憶體是一硬碟機、一光碟、一隨身碟、一網路磁碟機或是任何耦接至該計算裝置之其它型式的實體裝置。

【0027】 如在圖 3 中所示，該感測器 330 已經偵測到一觀看該計算裝置的使用者。作為回應的是，該控制器 320 及/或該語音應用程式 310 係指示該輸出構件 360 輸出一視覺、音頻及/或觸覺回授的提示訊號給該使用者。在一實施例中，當該提示訊號被輸出時，該控制器 320 及/或該語音應用程式 310 可以利用臉部辨識技術來驗證該使用者。該控制器 320 及/或該語音應用程式 310 可以指示該感測器 330 來捕捉該使用者的臉的一影像，並且比較該使用者的臉的影像與該計算裝置之驗證過的使用者的影像。該些驗證過的使用者的影像可以是本地儲存在該計算裝置的一儲存構件上、或是在該控制器 320 及/或該語音應用程式 310 可存取的另一裝置上。

【0028】 在一實施例中，若該使用者的影像並不符合一驗證過的使用者的影像，則該使用者係驗證失敗，並且該控制器 320 及/或該語音應用程式 310 係禁能該計算裝置的一傾聽模式。當禁能該傾聽模式時，該控制器

320 及/或該語音應用程式 310 可以禁能該音頻輸入構件 340、靜音該音頻輸入構件 340 及/或拒絕任何來自該使用者的音頻輸入。若該使用者的影像符合一驗證過的使用者的影像，則該使用者將會已經成功地被驗證，並且該控制器 320 及/或該語音應用程式 310 係繼續判斷該使用者是否持續觀看該計算裝置或輸出構件一段預設的時間量。

● 【0029】 若該使用者被判斷出是正在觀看該計算裝置及/或該輸出構件該預設的時間量，則該控制器 320 及/或該輸入應用程式 310 係判斷該使用者想要和該計算裝置互動，並且繼續啟動該傾聽模式。在另一實施例中，該控制器 320 及/或該語音應用程式 310 可進一步在該使用者持續觀看該計算裝置及/或該輸出構件時，偵測來自該使用者的持續的注視或是臉部表情上的變化。藉由偵測該持續的注視或是在臉部表情上的變化，該控制器 320 及/或該語音應用程式 310 可進一步確認該使用者要和該計算裝置互動的意圖。

● 【0030】 該控制器 320 及/或該語音應用程式 310 接著可以啟動該計算裝置的傾聽模式。當啟動該傾聽模式時，該控制器 320 及/或該語音應用程式 310 可以致能該音頻輸入構件 340 並且取消靜音該音頻輸入構件 340，以接收來自該使用者的音頻輸入。在一實施例中，當啟動該傾聽模式時，該控制器 320 及/或該語音應用程式係載入一和該使用者的一身分相關的話音(speech)輪廓。爲了此申請案之目的，該話音輪廓係包含使用者的語音或話音模式的資訊，該控制器 320 及/或該語音應用程式 310 係使用該資訊來改善偵測及辨識來自該使用者的一音頻輸入。該資訊可包含該使用者的話音的一音調、該使用者說話的一速率、該使用者的話音的一發音及/或任何

該使用者的語音或話音模式之額外的資訊。

【0031】 在一實施例中，該控制器 320 及/或該語音應用程式 310 亦指示該感測器 330 偵測該使用者相對於該音頻輸入構件 340 的身體貼近。如上所提到的，該使用者的身體貼近係對應於該使用者到該音頻輸入構件 340 的一位置或距離。根據該使用者的該位置及距離，該控制器 320 及/或該語音應用程式 310 可以校準該音頻輸入構件 340 以改善偵測來自該使用者的音頻輸入的正確性。校準該音頻輸入構件 340 可包含導引該音頻輸入構件 340 以從該使用者的一位置接收音頻輸入。在另一實施例中，校準該音頻輸入構件 340 可包含根據該使用者的該位置及/或距離來施加一雜訊濾波或抵消的演算法。

【0032】 回應於偵測到來自該使用者的音頻輸入，該控制器 320 及/或該語音應用程式 310 係辨識和來自該使用者的音頻輸入相關的該計算裝置的一命令、指令、輸入及/或詢問。在一實施例中，當該音頻輸入被該音頻輸入構件 340 偵測到時，該感測器 330 亦可以捕捉該使用者的嘴部的一視景。藉由捕捉該使用者的嘴部的視景，辨識對應於該音頻輸入的該計算裝置的命令、指令及/或查詢的正確性可被改善。該使用者的嘴部的視景亦可被利用作為一用於該音頻輸入構件 340 開始接收來自該使用者的音頻輸入的提示。

【0033】 圖 4 是描繪根據一例子的一種用於偵測音頻輸入之方法的流程圖。一例如在影像捕捉構件中的感測器最初可以在 400 之處偵測一觀看該計算裝置的使用者。若該使用者正在觀看該計算裝置，則該控制器及/或該語音應用程式可以在 410 之處指示一例如是一視覺指示器、音頻揚聲

器及/或一觸覺馬達的輸出構件來輸出一視覺、音頻及/或觸覺回授提示訊號。該提示訊號可被利用來通知該使用者該計算裝置係備妥以啓動一傾聽模式，以接收來自該使用者的音頻輸入。當該提示訊號被輸出時，該控制器及/或該語音應用程式可以在 420 之處判斷該使用者是否在該提示訊號被輸出時繼續觀看該輸出構件一段預設的時間量。若該使用者持續觀看該輸出構件該預設的時間量，則該控制器及/或該語音應用程式係判斷該使用者想要輸入一語音命令及/或詢問，並且該計算裝置的一傾聽模式係在 430 之處針對一音頻輸入構件來加以啓動，以接收來自該使用者的一音頻輸入。該方法接著完成。在其它實施例中，圖 4 的方法係包含除了那些在圖 4 中描繪的步驟之外及/或替代的額外步驟。

【0034】 圖 5 是描繪根據一例子的一種用於偵測音頻輸入之方法的流程圖。一例如是影像捕捉構件的感測器可被利用來捕捉一使用者的視景，並且該控制器及/或該語音應用程式可以在 500 之處使用人臉偵測技術及/或視線偵測技術來判斷該使用者是否正在觀看該計算裝置。該控制器及/或該語音應用程式係在 510 之處藉由判斷該使用者的臉是否面對該計算裝置及/或該使用者的視線是否觀看該計算裝置，以偵測該使用者是否正在觀看該計算裝置。若該使用者並未觀看該計算裝置，則該控制器及/或該語音應用程式繼續在 500 之處偵測觀看該計算裝置的使用者。

【0035】 若該使用者被判斷出是正在觀看該計算裝置，一輸出構件可在 520 之處被利用以輸出一視覺、音頻及/或觸覺回授的提示訊號來通知該使用者該計算裝置係備妥以啓動一傾聽模式。當該提示訊號被輸出時，該控制器及/或該語音應用程式可以在 530 之處判斷該使用者是否持續觀看該

輸出構件一段預設的時間量。若該使用者並不觀看該輸出構件該預設的量，則該控制器及/或該語音應用程式繼續在 500 之處偵測觀看該計算裝置的使用者。

【0036】 在一實施例中，若該使用者被判斷出是正在觀看該輸出構件該預設的時間量，則該控制器及/或該語音應用程式可以在 540 之處判斷在臉部表情上的一變化，例如一挑眉、一眨眼及/或一笑容是否已經在該使用者上偵測到。若在臉部表情沒有偵測到變化，則該控制器及/或該語音應用程式在 530 之處繼續判斷該使用者是否正在觀看該輸出構件該預設的時間量。若在臉部表情偵測到一變化，則該使用者接著可以在 550 之處利用臉部辨識技術來加以驗證。若該使用者驗證失敗，則一傾聽模式並不予以啟動，並且該控制器及/或該語音應用程式在 500 之處繼續偵測是否另一使用者正在觀看該計算裝置。

【0037】 若該使用者成功地加以驗證，則該控制器及/或該語音應用程式係在 560 之處載入一和該使用者的身分相關的話音輪廓。該話音輪廓係被用來改善偵測及辨識來自該使用者的一音頻輸入。該控制器及/或該語音應用程式接著可以在 570 之處藉由致能一例如是麥克風的音頻輸入構件來啟動該計算裝置的傾聽模式，以接收來自該使用者的一音頻輸入。在一實施例中，該感測器及/或該音頻輸入構件亦可以在 580 之處辨識該使用者的身體貼近，並且根據該使用者的身體貼近來繼續校準該音頻輸入構件。在另一實施例中，當該傾聽模式被啟動時，該控制器及/或該語音應用程式亦可以禁能在該計算裝置上之其它形式的輸入，例如觸控輸入、鍵盤輸入、滑鼠輸入及/或其它形式的用於該計算裝置的輸入。該方法係接著完成。在

其它實施例中，圖 5 的方法係包含除了那些在圖 5 中描繪的步驟之外及/或替代之額外的步驟。

【符號說明】

【0038】

100：計算裝置

120：控制器

130：感測器

140：音頻輸入構件

145：音頻輸入

150：通訊通道

160：輸出構件

165：提示訊號

200：計算裝置

205：使用者

230：感測器

240：音頻輸入構件

260：輸出構件

265：提示訊號

310：語音應用程式

320：控制器

330：感測器

340：音頻輸入構件

360 : 輸出構件

400-430 : 步驟

500-580 : 步驟

申請專利範圍

1.一種計算裝置，其係包括：

一用以偵測一觀看該計算裝置的使用者的感測器；

一用以在該使用者被偵測到為觀看該計算裝置時輸出一提示訊號的輸出構件；以及

一控制器，其係用以在該感測器偵測到該使用者繼續觀看與被輸出的該提示訊號相關的該計算裝置一段預設的時間量時，利用一音頻輸入構件以接收來自該使用者的一音頻輸入，

其中，該感測器可進一步捕捉該使用者的嘴部之一視景，且其中該使用者的嘴部的視景可被該控制器利用來改善辨識來自該使用者的該音頻輸入的正確性。

2.如申請專利範圍第 1 項之計算裝置，其中該提示訊號係包含一視覺信號、一音頻信號、一觸覺回授中的其中至少一個，其係對該使用者指出該計算裝置係備妥以啟動一傾聽模式來接收音頻輸入。

3.如申請專利範圍第 1 項之計算裝置，其中該音頻輸入是一來自該使用者的語音命令或詢問。

4.如申請專利範圍第 1 項之計算裝置，其中該感測器係包含一影像捕捉構件以及一紅外線構件中的其中至少一個。

5.如申請專利範圍第 1 項之計算裝置，其中該輸出構件係包含一視覺指示器、一音頻揚聲器以及一觸覺馬達中的其中至少一個，以輸出該提示訊號。

6.如申請專利範圍第 1 項之計算裝置，其中該控制器係在偵測該使用者

觀看該計算裝置時，偵測來自該使用者之一持續的注視或是在該使用者的一臉部表情上的一變化。

7.一種用於接收音頻輸入之方法，其係包括：

利用一感測器偵測觀看一計算裝置之一使用者；

若該使用者正在觀看該計算裝置，則利用一輸出構件以輸出一提示訊號；

判斷該使用者是否正在觀看該輸出構件一段預設的時間量；以及

若該使用者正在觀看該輸出構件該段預設的時間量，則啟動該計算裝置的一傾聽模式以接收來自該使用者的一音頻輸入，

其中，該感測器可進一步捕捉該使用者的嘴部之一視景，且其中該使用者的嘴部的視景可被該計算裝置利用來改善辨識來自該使用者的該音頻輸入的正確性。

8.如申請專利範圍第 7 項之用於接收音頻輸入之方法，其進一步包括若該使用者正在觀看該輸出構件，則利用臉部辨識技術來驗證該使用者。

9.如申請專利範圍第 7 項之用於接收音頻輸入之方法，其進一步包括載入一和該使用者的一身分相關的語音輪廓。

10.如申請專利範圍第 7 項之用於接收音頻輸入之方法，其進一步包括利用該感測器來辨識該使用者的一身體貼近，並且根據該使用者的該身體貼近來校準該計算裝置的一音頻輸入構件。

11.如申請專利範圍第 8 項之用於接收音頻輸入之方法，其中若該使用者的驗證失敗，則該計算裝置的該傾聽模式並不被致能。

12.一種包括指令之非揮發性電腦可讀取媒體，若執行該些指令，其係

使得一控制器以：

利用一感測器來偵測觀看一計算裝置的一使用者；

若該使用者正在觀看該計算裝置，則利用一輸出構件來輸出一提示訊號；

偵測該使用者是否持續觀看與被輸出的該提示訊號相關的該輸出構件一段預設的時間量；以及

若該使用者正在觀看該輸出構件該段預設的時間量，則啟動該計算裝置的一傾聽模式以接收來自該使用者的一音頻輸入，

其中，該感測器可進一步捕捉該使用者的嘴部之一視景，且其中該使用者的嘴部的視景可被該計算裝置利用來改善辨識來自該使用者的該音頻輸入的正確性。

13.如申請專利範圍第 12 項之非揮發性電腦可讀取媒體，其中在該音頻輸入構件偵測來自該使用者的該音頻輸入時，該感測器係另外捕捉該使用者的嘴部之一視景。

14.如申請專利範圍第 12 項之非揮發性電腦可讀取媒體，其中該控制器係辨識與音頻輸入以及該使用者的嘴部的該視景相關之該計算裝置的一輸入命令。

15.如申請專利範圍第 12 項之非揮發性電腦可讀取媒體，其中該輸出構件是具有一有限的視野之方向性的視覺指示器，並且該提示訊號係根據該使用者的一位置而被輸出給該使用者。

圖式

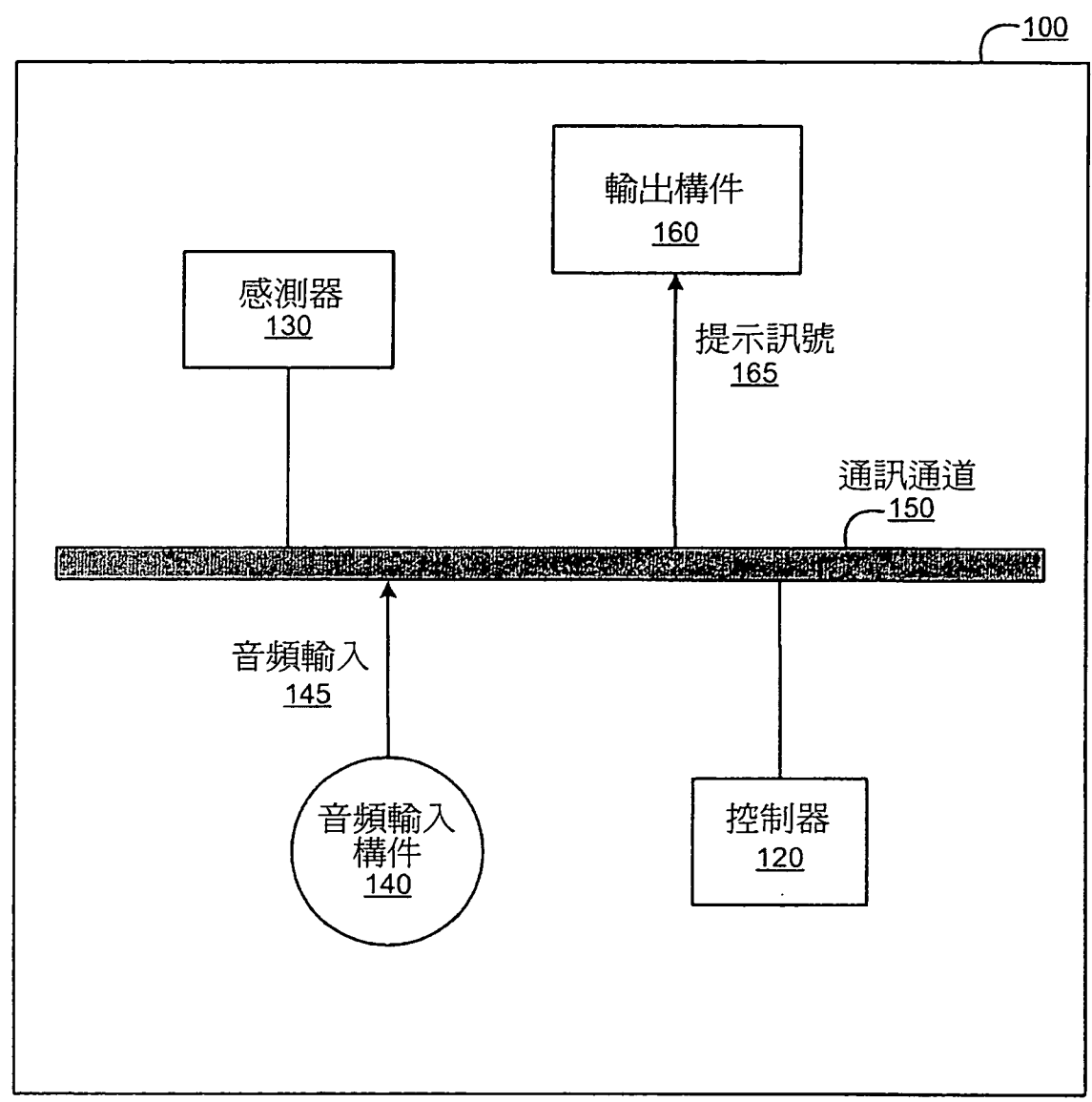


圖1

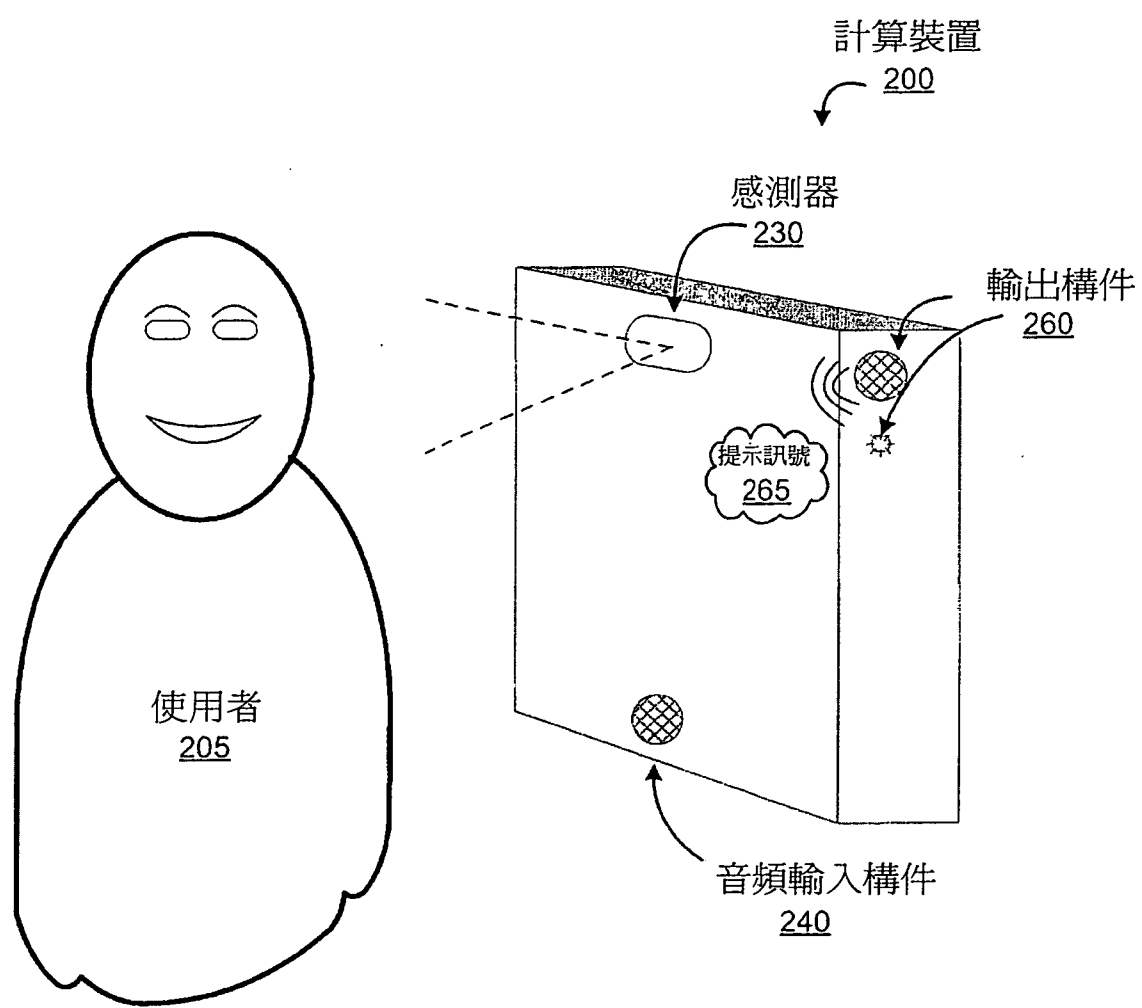


圖2

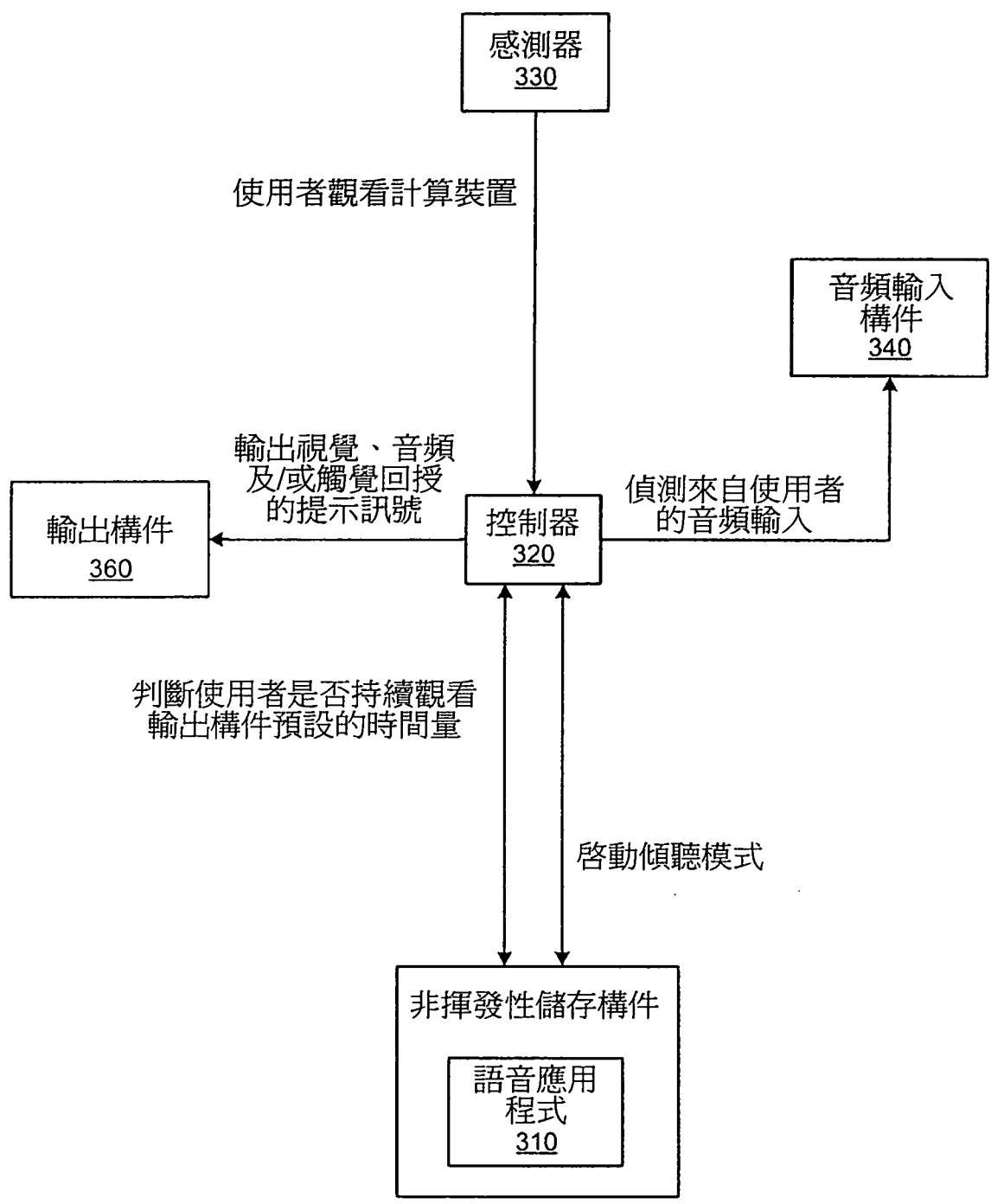


圖3

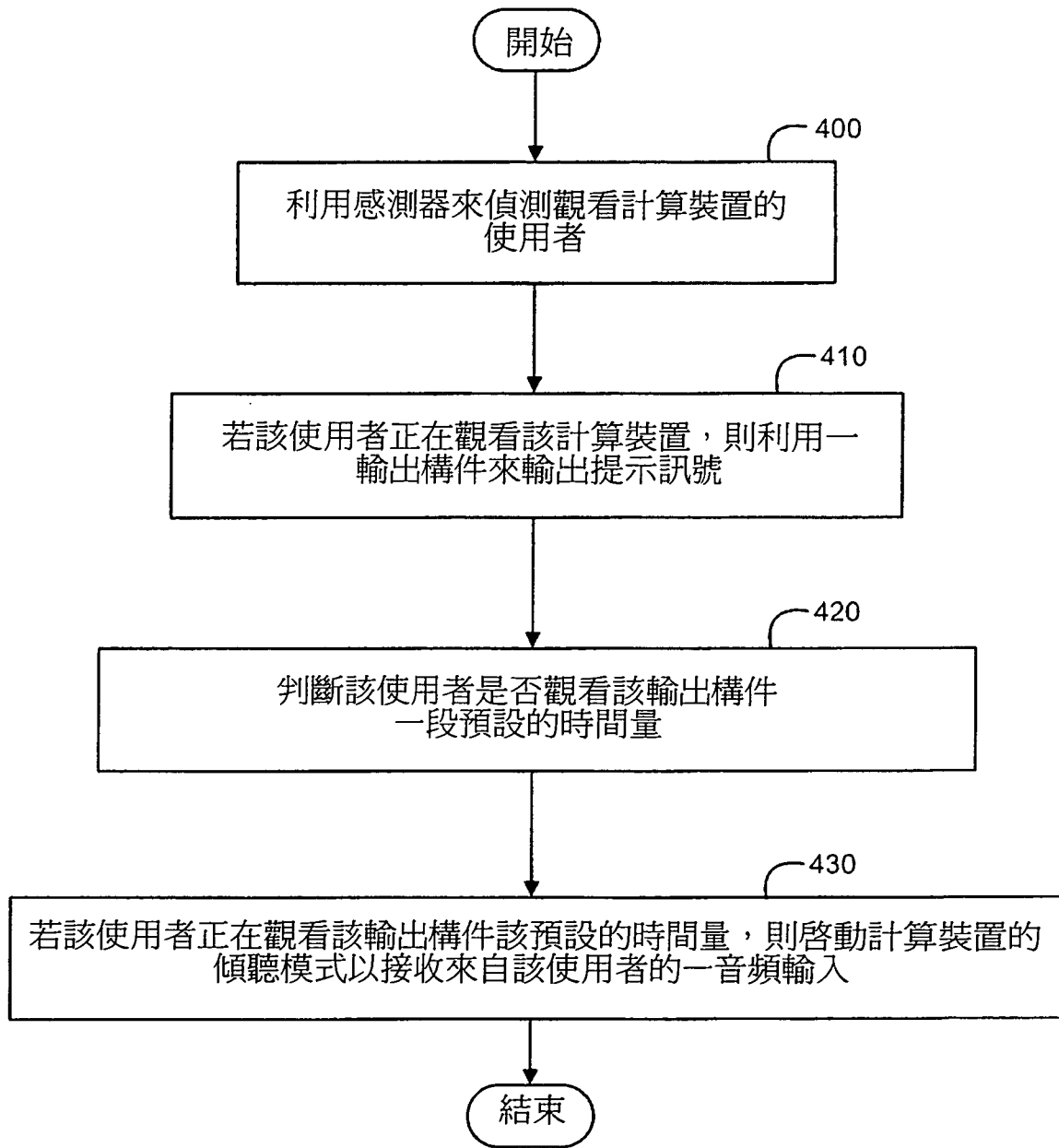


圖4

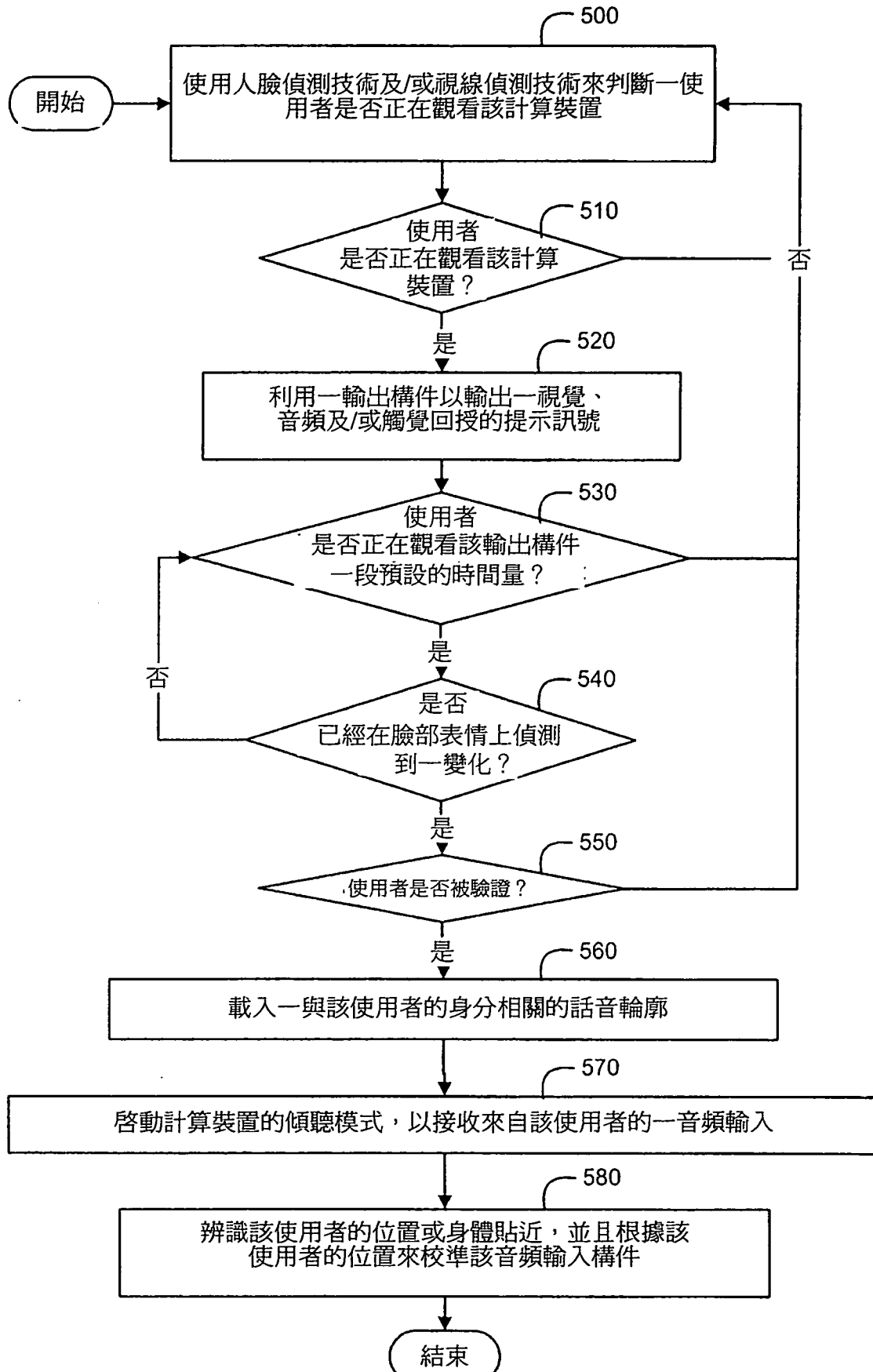


圖5