I506619

(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 I506619 B**

(45)公告日： 中華民國 104 (2015) 年 11 月 01 日

(21)申請案號：099118245 (22)申請日： 中華民國 99 (2010) 年 06 月 04 日

(51)Int. Cl.： ***G10L15/26 (2006.01)***

(30)優先權：2009/06/05 美國 12/479,477

(71)申請人：蘋果公司 (美國) APPLE INC. (US)
美國

(72)發明人：凡 歐斯 馬塞 VAN OS, MARCEL (NL)；諾維克 葛瑞格里 NOVICK, GREGORY (US)；赫茲 史考特 HERZ, SCOTT (US)

(74)代理人：陳長文

(56)參考文獻：

TW	200638337A	US	2005/0125235A1
US	2008/0201306A1	US	2008/0221880A1
US	2009/0006099A1	US	2009/0112572A1

審查人員：黃衍勳

申請專利範圍項數：27 項 圖式數：9 共 48 頁

(54)名稱

用於情境式語音命令的方法、裝置及非過渡電腦可讀媒體

METHODS, APPARATUSES AND NON-TRANSITORY COMPUTER READABLE MEDIA FOR CONTEXTUAL VOICE COMMANDS

(57)摘要

本發明揭示用於實施情境式語音命令之技術及系統和其他事宜。在一器件上顯示一第一情境中之一資料項。在該器件上接收選擇該第一情境中之該所顯示資料項的一實體輸入。在該器件上接收使該所選資料項與一第二情境中之一操作相關的一語音輸入。在該第二情境中，該操作係對該所選資料項執行。

Among other things, techniques and systems are disclosed for implementing contextual voice commands. On a device, a data item in a first context is displayed. On the device, a physical input selecting the displayed data item in the first context is received. On the device, a voice input that relates the selected data item to an operation in a second context is received. The operation is performed on the selected data item in the second context.

600・・・用於實施情境式語音命令的實例程序

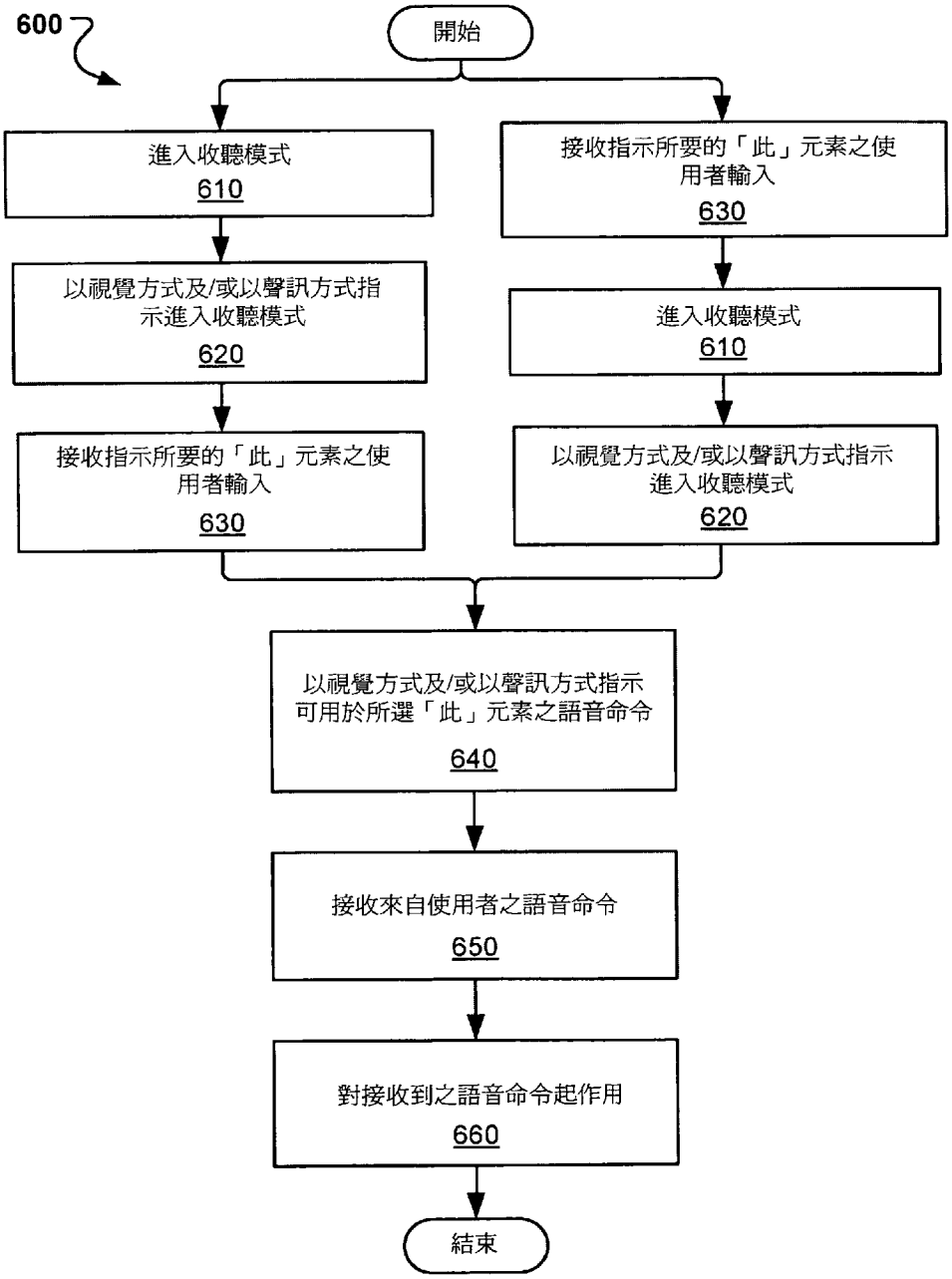


圖6

發明摘要 公告本 104年6月5日修正替換頁

※ 申請案號：099118245

※ 申請日：99.06.04 ※IPC 分類：G10L 15/26 (2006.01)

【發明名稱】

用於情境式語音命令的方法、裝置及非過渡電腦可讀媒體

METHODS, APPARATUSES AND NON-TRANSITORY COMPUTER READABLE MEDIA FOR CONTEXTUAL VOICE COMMANDS

【中文】

本發明揭示用於實施情境式語音命令之技術及系統和其他事宜。在一器件上顯示一第一情境中之一資料項。在該器件上接收選擇該第一情境中之該所顯示資料項的一實體輸入。在該器件上接收使該所選資料項與一第二情境中之一操作相關的一語音輸入。在該第二情境中，該操作係對該所選資料項執行。

【英文】

Among other things, techniques and systems are disclosed for implementing contextual voice commands. On a device, a data item in a first context is displayed. On the device, a physical input selecting the displayed data item in the first context is received. On the device, a voice input that relates the selected data item to an operation in a second context is received. The operation is performed on the selected data item in the second context.

104年6月5日修正替換頁

【代表圖】

【本案指定代表圖】：第（6）圖。

【本代表圖之符號簡單說明】：

600　用於實施情境式語音命令的實例程序

【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

104年6月5日修正替換頁

（本說明書格式、順序，請勿任意更動）

【發明名稱】

用於情境式語音命令的方法、裝置及非過渡電腦可讀媒體

METHODS, APPARATUSES AND NON-TRANSITORY COMPUTER READABLE MEDIA FOR CONTEXTUAL VOICE COMMANDS

對相關申請案之交叉參考

本申請案主張2009年6月5日申請之美國專利申請案第12/479,477號之優先權的權利。

【技術領域】

本申請案係關於話語辨識。

【先前技術】

話語辨識可用來使用一串字元碼之二進位碼將說的話轉換成諸如鍵按壓之機器可讀輸入。典型的話語辨識應用包括語音撥號、來電路由、家用器具控制、簡單資料鍵入及話語至文書處理。在此等應用中，可將話語用作使用者介面之一部分以產生預定義之話語或語音命令。

【發明內容】

可使用實體輸入與語音輸入之組合來實施情境式語音命令，該等語音命令控制一器件(諸如，智慧型電話)上之跨不同情境(例如，應用程式)之器件操作。此外，相比於藉由導航通過一組巢套選單項，使用者可藉由使用情境式語音命令更快地執行所要的操作。又，情境式語音命令可用來教該器件自單一語音命令準確地預測使用者之意圖。另外，情境式語音命令可用以基於所使用的器件之情境來改變使用者提供語音輸入之方式。

104年6月5日修正替換頁

【圖式簡單說明】

圖1a及圖1b展示一實施例中的用於執行情境式語音命令的一器件之正視圖及仰視圖。

圖2a及圖2b說明一實施例中的用於一資料處理器件之情境式語音命令的啟動。

圖3a及圖3b展示一實施例中的一資料處理器件之正視圖，其說明情境式語音命令之使用。

圖4展示一實施例中的實施於另一資料處理器件中之情境式語音命令。

圖5為說明一實施例中的與其他應用程式、輸入單元及輸出單元介接之一情境式語音命令模組的方塊圖。

圖6為一實施例中的用於實施情境式語音命令的一實例程序之程序流程圖。

圖7為展示一實施例中的用於實施確認模型的一實例程序之程序流程圖。

圖8a為展示一實施例中的用於實施最佳假設或學習模型的一實例程序之流程圖。

圖8b展示一實施例中的一資料處理器件，其存取知識庫以識別使用者意圖之最佳假設或最可能的選項。

圖9為一實施例中的一資料處理器件之實例架構之方塊圖。

【實施方式】

各種圖式中之相同參考符號及名稱指示相同元件。

圖1a及圖1b展示一資料處理器件102之正視圖及仰視圖，該器件用於執行情境式語音命令以操作該器件。資料處理器件102可為有能力接收語音及實體輸入且基於接收到之語音及實體輸入而執行一或多個應用程式或情境中之操作的任一器件(諸如，行動器件)。資料處理

器件102包括通信技術(例如，無線技術)以便與其他器件共用資訊。資料處理器件102可包括多種整合式使用者介面單元，或可經由該器件之一或多個通信埠或資料鏈路耦接至該等使用者介面單元。使用者介面單元之一些實例包括(但不限於)語音輸入單元，諸如，麥克風106。使用者介面單元之一些實例包括(但不限於)實體輸入單元，諸如，鍵盤、滑鼠、軌跡球、旋轉撥號盤或輪、觸控板或觸控式螢幕。使用者介面單元之一些實例包括(但不限於)運動感測器，諸如，加速計、磁力計或迴轉儀。可將此等使用者介面單元中之任一者實施為外部單元，其使用有線或無線通信技術與資料處理器件102通信以提供使用者輸入。有線通信技術之實例包括通用串列匯流排(USB)介面、FireWire介面等。無線通信技術之實例包括藍芽(Bluetooth)、Wi-Fi、WiMax、紅外線等。經由此等使用者介面單元，資料處理器件102可接收來自使用者之實體或語音輸入。

資料處理器件102包括輸出單元以向使用者呈現視覺及音訊資訊。輸出單元之一些實例包括(但不限於)顯示單元108及揚聲器104。出於說明性目的，圖1a及圖1b將資料處理器件102展示為行動器件，但資料處理器件102可包括其他類型之器件，諸如，行動電話、個人數位助理、攜帶型電腦及桌上型電腦。

所揭示之語音命令係為情境式的，因為還考慮了除了語音之外的輸入，該等輸入指示針對該等命令的不同等級或類型的情境。所考慮之情境式輸入可包括在特定時間執行之特定活動、所選擇的活動之一特定部分，或此情境中之語音命令。舉例而言，可在使用者正在一資料處理器件上執行(諸如，使用一特定應用程式)的活動之情境中實施情境式語音命令。舉例而言，圖1a展示資料項，諸如，顯示於資料處理器件102之顯示單元108上的圖示。每一圖示表示一可供在資料處理器件102上執行之對應的應用程式。可用之圖示及相關應用程式

104年6月5日修正替換頁

的一些實例包括(但不限於)：電話應用程式圖示110a、電子郵件應用程式圖示110b、網路瀏覽器應用程式圖示110c、音樂播放器應用程式圖示110d、媒體播放器應用程式圖示110e、音樂下載應用程式圖示110f、影像處理應用程式圖示110g、地理定位(例如，GPS、基於Wi-Fi之定位、蜂巢式三角量測)應用程式圖示110h、聯絡人應用程式圖示110i、簡訊服務(SMS)應用程式圖示110j、視訊遊戲應用程式圖示110k及文書處理應用程式圖示110l。又，可使用聲訊資料項(諸如，與可用之應用程式相關聯的所記錄之聲音)向使用者呈現可用之應用程式。所記錄之聲音可為使用者可辨識之話語(諸如，應用程式之名稱)或簡單的聲音(諸如，嗶聲、音調、鐘聲等)。

因此，相比於僅整體上控制器件之習知語音命令，情境式語音命令可更精確。舉例而言，可基於使用者活動或使用中之應用程式來實施不同的情境特定語音命令。此外，情境之範疇可擴展至使用者活動或應用程式之特定部分，諸如，使用者介面之特定區域。以此方式，情境式語音命令可針對使用者活動或應用程式之所選部分選擇性地提供對使用者介面之語音導航。

此等應用程式中之每一者可與一情境式語音命令控制器或模組通信，該情境式語音命令控制器或模組基於接收到之使用者輸入來控制應用程式之操作。情境式語音命令模組可藉由啟動資料處理器件上之情境式語音命令模式來控制應用程式之操作。為了控制應用程式，可將情境式語音命令控制器或模組實施為用於其他應用程式之外掛程式(plug-in)。在其他實施中，資料處理器件可包括一應用程式設計介面(API)，該API允許情境式語音命令模組存取由作業系統或其他導航應用程式提供之程式庫及公用程式。情境式語音命令模組可與輸入單元(例如，語音及/或實體輸入單元)通信以接收用以控制其餘的應用程式之使用者輸入。

102年12月4日修正替換頁

音訊輸入單元可包括接收音訊輸入之使用者介面單元，諸如，麥克風106。如上所述，音訊輸入單元可爲內建式單元或外部單元。使用者可對著麥克風106說話以提供語音輸入，可使用話語辨識技術(例如，基於隱式馬爾可夫模型(Hidden Markov Model)之技術、動態時間扭曲技術等)將該等語音輸入轉譯成情境式語音命令。可使用話語辨識軟體或技術來辨識話語中之語音命令，諸如，市售產品或開放原始碼產品(諸如Open Mind Speech、VoxForge、Julius等)與其他產品。

可使用實體輸入單元(例如，觸控式螢幕、滑鼠、觸控板等)接收實體輸入，諸如，關於顯示於資料處理器件102之輸出單元(例如，輸出單元108)上的一視覺螢幕之位置資訊。可使用一些實體輸入單元(例如，按鈕)接收非位置資訊，諸如，使用者選擇或確認。舉例而言，使用者可使用滑鼠在所顯示之螢幕上執行各種點擊功能。又，使用者可使用觸控式螢幕在所顯示之螢幕上執行相同的功能。除了實體輸入之外，一些使用者介面單元(諸如，鍵盤)還可提供文字輸入。

圖2a及圖2b說明資料處理器件102上的情境式語音命令模式之啓動。圖2a展示在啓動情境式語音命令模式前的資料處理器件102之正視圖。使用者可藉由與一硬體按鈕或虛擬使用者介面元素(諸如，情境式語音命令圖示110m)互動來啓動情境式語音命令模式且將資料處理器件102置於收聽模式。此外，可實施情境式語音命令模式以向使用者呈現完整的語音控制服務，此服務包括情境式選項及非情境式選項。因此，術語「情境式語音命令模式」應用於本說明書中而包括(但不限於)情境式語音命令。

圖2b展示情境式語音命令模式已被啓動的資料處理器件102之正視圖。在圖2b中展示之實例中，使用者使用(例如)其手指103觸摸且選擇情境式語音命令圖示110m。可使用視覺指示(諸如，黑體邊界、

彩色光輝、反白顯示、不同色彩等)將顯示於顯示單元108上的資料項或元素(例如，情境式語音命令圖示110m)之使用者選擇傳達至使用者。又，除了視覺指示之外，或代替視覺指示，可使用音訊指示。舉例而言，可經由揚聲器播放聲訊「情境式語音命令在作用中」或「我正在收聽」給使用者聽。在一些實施中，向使用者呈現之音訊指示可包括代替可辨識話語或除了可辨識話語之外所播放的提示音調。

一旦情境式語音命令模式在作用中，資料處理器件102便可產生額外的視覺及/或音訊指示以向使用者呈現情境式語音命令的可用選項。舉例而言，顯示器108之一部分可用以向使用者呈現視覺指示(例如，狀態列105)。在圖2b中展示之實例中，狀態列105用以顯示各種文字，其向使用者指示：器件正收聽來自使用者之情境式語音命令，且使用者可選擇一元素(例如，一應用程式)以接收可用於所選元素之情境式語音命令的清單。器件可使用聲訊及視覺指示向使用者提供情境式語音命令及/或指令之可用選項。

圖3a及圖3b展示資料處理器件102之正視圖以顯現使用中的情境式語音命令。根據資料項或元素(例如，所顯示之圖示)之可用選項，使用者可選擇一圖示以啓動一相關聯之應用程式。使用上文在圖2a及圖2b中說明之程序，情境式語音命令模式此時已在作用中。然而，在一些實施中，可在選擇且起始了可用之應用程式中之一者後啓動情境式語音命令模式。在一些實施中，資料處理器件可回應於對自一應用程式切換至另一應用程式之使用者請求而退出情境式語音命令模式。又，使用者可使互動之焦點自一特定應用程式變爲一主螢幕(例如，藉由按壓與主螢幕相關聯之使用者介面按鈕)以退出情境式語音命令模式。在另一實例中，當使用者之焦點在主螢幕上時，情境式語音命令模式可在作用中。在此實例中，當使用者將其焦點自主螢幕變爲一感興趣之應用程式時，資料處理器件可退出或撤銷啓動情境式語音命

102年12月4日修正替換頁

令模式。

此等圖示中之任一者的使用者選擇在資料處理器件102上載入且起始對應的應用程式以供使用者使用。舉例而言，使用者可經由實體輸入單元(諸如，觸控式感測器)實體地觸摸電子郵件圖示110b，以選擇且起始電子郵件應用程式。當情境式語音命令模式在作用中時，使用者亦可使用音訊輸入單元(諸如，麥克風)在有或無位置輸入的情況下進行選擇。

類似於情境式語音命令圖示之使用者選擇，可使用視覺指示(諸如，黑體邊界、彩色光輝、反白顯示、不同色彩等)將可用圖示中之一者的使用者選擇傳達至使用者。又，除了視覺指示之外，或代替視覺指示，可使用音訊指示。舉例而言，可經由揚聲器播放聲訊「電子郵件應用程式」給使用者聽。

對應於使用者所選圖示之應用程式被顯示於顯示單元108上。舉例而言，圖3a展示回應於電子郵件圖示110b之使用者選擇而顯示於顯示單元108上之電子郵件。當情境式語音命令模式在作用中時，可使用另一視覺指示(諸如，狀態列105)顯示可用於該所選應用程式之功能之清單。狀態列105可使用任何數目種視覺技術來顯示可用功能之清單。舉例而言，可將清單顯示爲一串捲動的文字。可將清單分成多個有N個成員的群組，N爲可安全地同時顯示的功能之數目。可以旋轉方式顯示每一功能群組以在每一次旋轉結束時將所有功能顯示一次。又，雖然狀態列105被展示成顯示於顯示單元108之頂部，但狀態列105可位於顯示單元108上任一處。在一些實施中，可在螢幕上之特定觸控點或使用者介面點附近顯示可用之語音命令或選項。舉例而言，當使用者實體地觸摸介面元素或資料項(例如，表示應用程式之圖示)時，可在介面元素及/或使用者之手指附近展示可用於該介面元素的語音命令選項。除了將可用功能顯示於狀態列105上之外，還可

102年12月4日修正替換頁

將可用功能提供爲音訊給使用者聽。在圖3a中展示之實例中，資料處理器件呈現對器件正收聽語音命令之視覺指示，且使用者可選擇顯示於螢幕上之資料項或元素以獲得可用於使用者所選元素之語音命令。又，使用者可使用情境式語音命令來請求且獲得說明。舉例而言，當使用者請求說明進行可用於感興趣之資料項的操作時，器件可向使用者呈現該等可用操作。以下進一步描述說明選項。

在任一應用程式之任一使用者介面上，可啓動情境式語音命令模式以啓用專用於與一使用者所選元素相關聯之特定活動的語音命令。以此方式，語音命令係特定地基於由與螢幕之任一部分或元素相關聯的實體輸入引起的情境。

圖3b展示資料處理器件之正視圖以說明使用中的跨情境之語音命令。顯示於顯示單元108上之電子郵件包括一附加至電子郵件之測試影像107。使用者可(例如，使用其手指103，經由觸控式螢幕)選擇該測試影像以提出與所選測試影像相關聯的情境式語音命令之可用選項。每一情境式語音命令係與一可對所選測試影像107執行之相對應操作相關聯。該等相對應操作可包括針對目前情境(例如，電子郵件應用程式)或針對另一情境(例如，影像處理應用程式)之操作。回應於測試影像107之使用者選擇，資料處理器件102使用狀態列105顯示可用於所選測試影像107的情境式語音命令之可用選項。

圖4展示在另一資料處理器件中使用之情境式語音命令。可在各種資料處理器件中實施情境式語音命令模式。舉例而言，能夠操作情境式語音命令模式之資料處理器件可包括個人電腦400。類似於圖1a至圖3b中展示之行動器件實例，使用者可自顯示單元108選擇一元素(例如，一電子郵件圖示110b)以起始相對應的應用程式。圖4中展示之實例展示使用者使用一滑鼠指標402選擇該電子郵件圖示110b。回應於此實體輸入，載入且起始相對應的電子郵件應用程式。經起始之

102年12月4日修正替換頁

電子郵件應用程式顯示於指示作用中應用程式的一作用中螢幕404上。作用中螢幕404包括一電子郵件406及一附加至該電子郵件之影像107。回應於電子郵件406之使用者選擇，個人電腦400可使用狀態列105顯示與使用者所選元素(例如，電子郵件406)相關聯的可用操作之可用選項。舉例而言，圖4展示一可選狀態列105，其顯示包括「撰寫、回覆、轉發、閱讀、保存、複製、貼上、搜尋、檔案、編輯、檢視、工具、動作、說明等」等之操作之可用選項。然而，在使用者發出情境式語音命令之前，不需要向使用者呈現可用命令之清單的此視覺或聲訊指示。舉例而言，語音命令模組可處理自使用者接收到之口頭輸入且推斷出使用者之意圖，而無需向使用者呈現命令清單。

圖5展示與其他應用程式、輸入單元及輸出單元介接之一情境式語音命令模組的方塊圖。情境式語音命令控制器或模組510m可與資料處理器件上可用之其他應用程式通信且控制該等其他應用程式。與圖1a、圖1b、圖2a、圖2b、圖3a、圖3b及圖4中展示之實例一致，可用之應用程式可包括(但不限於)：電話應用程式510a、電子郵件應用程式510b、網路瀏覽器510c、音樂播放器應用程式510d、媒體播放器應用程式510e、音樂下載應用程式510f、影像處理應用程式510g、地理定位應用程式510h、聯絡人應用程式510i、SMS應用程式510j、視訊遊戲應用程式510k及文書處理應用程式510l。可回應於如圖2a、圖2b、圖3a、圖3b及圖4中所展示的相關聯之圖示之使用者選擇而載入且起始此等應用程式。

情境式語音命令模組510m經由輸入單元520接收使用者輸入。輸入單元可包括音訊輸入單元522(諸如，麥克風)及實體輸入單元524(諸如，觸控式螢幕)。然而，如上文關於圖1a及圖1b所描述，可實施其他輸入單元。回應於接收到之使用者輸入，情境式語音命令模組理解該輸入之情境(例如，所選應用程式或所選應用程式中之一所選元素)

102年12月4日修正替換頁

及相關聯之語音命令。基於接收到之使用者輸入，情境式語音命令模組510m控制該等應用程式中之一或多者及其相關聯之操作或功能。

圖6爲用於實施情境式語音命令的一實例程序600之流程圖。器件(例如，資料處理器件102)進入語音控制或收聽模式(610)。爲了進入收聽模式，使用者可經由一輸入單元提供一預定使用者輸入。舉例而言，使用者可將一實體或虛擬按鈕按住不放，或直接說出一觸發情境式語音命令選項的預定語音命令。舉例而言，可使用諸如「器件收聽」之預定觸發片語使器件進入收聽模式。又，情境式語音命令選項的預設狀態可爲作用中，且使用者可藉由提供一預定觸發輸入來撤銷啓動該選項。舉例而言，器件可基於諸如「退出收聽模式」之語音輸入或諸如觸摸或按鈕按壓之實體輸入而退出收聽模式。又，器件可使用語音輸入與實體輸入之組合進入並退出收聽模式。舉例而言，使用者可觸摸一按鈕以起始進入收聽模式中或自收聽模式退出，及起始一後續的語音命令(諸如，「確認退出或進入」)以結束進入收聽模式中或自收聽模式退出。

在進入收聽模式時，器件可視情況呈現對器件處於收聽模式之視覺及/或聲訊指示(620)。實例聲訊指示可爲向使用者播放之音訊記錄，該音訊記錄指示「我正在收聽且您可說這些話以對元素起作用。」聲訊指示之另一實例可爲簡單的聲音警告，諸如，嗶聲、丁當聲、鐘聲或此等聲音之組合。視覺及/或聲訊指示警告使用者：器件正處於收聽模式，且使用者可提供表示一相對應的情境式語音命令之所要語音輸入。

在接收且處理情境式語音命令時，器件可對情境中之一特定元素進行操作，該元素可爲所指示之元素、當前在作用中之元素或暗示之元素，在本文中被稱作感興趣之元素。在語音命令處理系統之詞彙內，可藉由任一適當術語(諸如，「此」或「項」或「當前項」或「選

104年6月5日修正替換頁

擇」與許多其他術語)來指此感興趣之元素。此外，語音命令系統可能正在接收關鍵字，忽略額外的字或使用額外的字作為指導。因此，由於針對當前感興趣之項使用關鍵字「此」或「所選」，因此可類似地處理「用電子郵件發送此」或「用電子郵件發送所選影像」或「用電子郵件發送此照片」。

器件接收一識別所要的感興趣之元素的使用者輸入，該元素提供情境式語音命令之情境(630)。使用者可藉由使用諸如觸控式螢幕、滑鼠、鍵盤等的實體輸入單元中之任一者來指定且選擇感興趣之元素。又，使用者可藉由使用一聲訊輸入單元說出預定觸發片語(諸如，「第三行之文字」或「底部之影像」，其對螢幕上所顯示之內容提供具體指導)來指定且選擇感興趣之元素。在使用者對感興趣之元素進行其選擇前，器件不需要進入收聽模式。在尚未處於收聽模式時，使用者可在進行對感興趣之元素的選擇後輸入該預定觸發輸入以使器件進入收聽模式中。

可視情況提供額外的視覺及/或聲訊指示以包括可用於所選感興趣之元素的常規語音命令及情境式語音命令之清單(640)。舉例而言，可使用狀態列(例如，圖2b、圖3a、圖3b及圖4中之狀態列105)提供可用之情境式語音命令的視覺指示。可用命令之清單可包括非情境式命令(諸如，「撥打電話給一個人」、「閱讀電子郵件」、「下一個會議」等)及情境式命令(諸如，「將此發送至…」、「加書籤於此」、「在2小時內對我提醒此」、「刪除此」、「在網路上搜尋此」等)。情境式語音命令將被辨識為在「此」之情境中。用於感興趣之元素的術語「此」或「項」或「當前項」或「選擇」或任一類似識別符可與整個所顯示之螢幕、所選作用中螢幕或任一螢幕上之任一元素或任一實體輸入(諸如，硬體按鈕)相關。感興趣之元素可為在使用者觸摸或叫出該元素時被顯示於螢幕上之資料項。感興趣之元素可為應用程式特定

資料項(例如，「清除清單」)。又，感興趣之元素可為呈現於螢幕上之單元(例如，「此郵件訊息」)。

可隱含地(亦即，無需感興趣之元素的顯示使用者指示)識別感興趣之元素。舉例而言，若器件當前正顯示一針對通訊錄中之單一聯絡人之編輯螢幕，則系統可不需要進一步的使用者輸入來處理涉及感興趣之項的「刪除當前項」或「刪除此」或任一類似片語。

器件經由輸入單元接收來自使用者之語音命令(650)。回應於感興趣之元素的使用者選擇及接收到之語音命令，器件對接收到之情境式語音命令起作用(660)。對接收到之情境式語音命令起作用可包括執行一組功能，諸如，識別且呼叫可對使用者所選感興趣之元素起作用之適當應用程式。又，器件可能需要提供額外的視覺及/或聲訊指示，諸如，可用於所選感興趣之元素的功能之清單。情境式語音命令模組(見圖2a、圖2b、圖3a及圖3b)基於經由輸入器件接收到之使用者輸入來理解情境式語音命令中的感興趣之元素的情境，且識別可用於感興趣之元素或資料項的相對應功能。

可用於所選感興趣之元素或資料項的功能或操作可包括與作用中之應用程式或情境相關聯之功能。電子郵件應用程式功能或操作可用於電子郵件元素(諸如，附加至電子郵件的影像107(見圖3a及圖3b))。舉例而言，可將影像107轉發至另一電子郵件地址。一實例命令可為「將此影像用電子郵件發送至email@email.com.」。

此外，用於使用者所選元素的視覺及聲訊清單可包括可用於可對所選元素起作用之任一其他應用程式或情境的功能或操作。此允許情境交越。在選擇一元素後，使用者可在所選元素之情境中提供一情境式語音命令(諸如，「複製此」或「修改選擇」或「刪除項」)以指示情境式語音命令適用於所選元素。

舉例而言，可藉由說出預定觸發片語「修改此影像」而使用影

像處理應用程式修改影像107。回應於情境式語音命令「修改此」，情境式語音命令模組搜尋可修改所選影像之可用之應用程式的清單。當一個以上應用程式可用時，可使用狀態列105顯示可用之應用程式之清單。又，可將可用之應用程式播放為聲訊清單。回應於應用程式之視覺或聲訊清單，使用者可說出所要的應用程式之名稱以啟動該應用程式來修改所選影像。

以此方式，情境式語音命令可用以跨不同情境對使用者所選元素起作用。在上文描述之實例中，在電子郵件情境中顯示之影像107可在影像處理情境中使用。在另一實例中，聯絡人情境中之企業名稱可在網路瀏覽器情境中用來在網路上搜尋該企業名稱。因為使用者可使用簡單且直觀的語音命令(例如，「Google®此企業名稱」)而不是情境相關選單、按鈕等之長清單，所以使用情境式語音命令使實施情境交越更簡單且更快。

舉例而言，使用觸摸及/或語音輸入，使用者可在任一螢幕上進行選擇，諸如，在字處理應用程式情境中選擇一段文字。為針對該段所選文字執行情境交越，使用者可直接說出預定觸發片語(諸如，「將此用電子郵件發送至email@email.com」或「將此移至網路瀏覽器」)以在另一情境中，對該段所選文字起作用。與此情境式語音命令相反，觸控式使用者介面需要使用者複製並保存該段所選文字，開啟電子郵件應用程式，撰寫新的電子郵件，尋找或鍵入電子郵件地址，將該段保存文字貼至電子郵件中，且按壓發送按鈕。

又，使用者可發出諸如「將此用電子郵件發送至John」之命令而無需離開目前情境或應用程式。舉例而言，在傳統使用者介面中，使用者將需要選擇「用電子郵件發送此照片」選項，且接著鍵入將在某一電子郵件使用者介面中接收該電子郵件的人之電子郵件地址，等。情境式語音命令模式允許使用者發出命令且使動作在背景中執行，而

使用者無需離開當前使用之應用程式之情境。又，更多的操作可供使用者自由使用，而不需要螢幕上的空間來配備更多的按鈕或添加長的選單以便找到此等選項。

在一些實施中，使用者可首先提供一輸入以選擇感興趣之資料項。舉例而言，使用者可觸摸所顯示之資料項或與所顯示之資料項互動，或以其他方式使資料處理器件在螢幕上顯示感興趣之資料項(例如，選擇相薄中之一照片以全螢幕顯示所選照片)。接著，使用者可指導資料處理器件進入情境式語音命令模式以與所選資料項互動。

在使用聲訊指示提供使用者互動的過程中，可使用不同的使用者互動模型，諸如，確認模型或最佳假設模型。

圖7為展示用於實施確認模型的一實例程序700之程序流程圖。確認模型回應於自使用者接收到之每一語音輸入而以問題之形式提供聲訊確認。器件接收指示所要的感興趣之元素之使用者輸入及與感興趣之元素之情境相關的一情境式語音命令(710)。回應於接收到之情境式語音命令，情境式語音命令模組識別可用之選項(720)且使用聲訊及/或視覺輸出向使用者請求確認(730)。器件接收回應於確認請求的另一使用者輸入(740)。情境式語音命令模組處理該另一接收到之使用者輸入以判定是否需要額外確認(例如，判定情境式語音命令是否已結束)(750)。當需要額外確認時，情境式語音命令模組識別可用之選項(720)且使用聲訊及/或視覺輸出向使用者請求確認(730)。當不需要額外確認(例如，情境式語音命令已結束)時，器件對接收到之情境式語音命令起作用(760)。

舉例而言，當使用者說出預定觸發片語(諸如，「撥打電話給John」)時，器件提供確認回應「哪個John？」且接著提供名字為John的所有可用之人的聲訊及/或視覺清單。使用者可藉由說出預定觸發片語(諸如，「此人」或「這個John」)來選擇該等名字中之一者，同時

102年12月4日修正頁

使用位置輸入器件來觸摸該清單中所要的John之視覺指示。當完全依靠語音輸入時，使用者可說出所要的John之全名。回應於使用者選擇，器件提供下一個適當的確認(諸如，「好的，哪個號碼？」)且提供可用於所選John的所有電話號碼之聲訊及/或視覺清單。使用者可自該等可用電話號碼進行選擇。

圖8a爲展示用於實施最佳假設或學習模型的一實例程序800之流程圖。最佳假設或學習模型使用使用者之使用或偏好歷史進行最佳假設。學習模型之實例可包括(但不限於)：諸如支援向量機(SVM)之機器學習模型、歸納推論模型、概念學習、決策樹學習、貝氏(Bayesian)學習及其他學習模型。機器學習模型可用以開發一教該資料處理器件基於經由輸入單元接收之或儲存於資料庫中的累積資料來改良其效能之程序。此等機器學習模型可用以基於根據累積資料而設計之規則及型樣自動產生所要的結果。器件接收指示所要的感興趣之元素之使用者輸入及一情境式語音命令(810)。情境式語音命令模組檢查知識庫以識別使用者之意圖之最佳假設或最可能的選項(820)。

圖8b展示一資料處理器件，其存取知識庫以識別使用者之意圖之最佳假設或最可能的選項。在圖8b中展示之實例中，情境式語音命令模式510m在作用中。舉例而言，使用者可使用其手指103經由觸控式螢幕觸摸相對應的情境式語音命令圖示110m。又，使用者選擇應用程式中之一者以載入且起始所選應用程式。在圖8b中展示之實例中，回應於使用者觸摸相對應的電話圖示110a而選擇電話應用程式510a。在實體輸入之情況下，使用者可說出語音輸入802，諸如，「撥打電話給John」。經由各別輸入單元524及522接收實體輸入及語音輸入，且情境式語音命令模組510m存取知識庫806以識別最佳假設。知識庫806可包括儲存於資料處理器件102內之本端知識庫812及位於資料處理器件102外部之外部知識庫814。舉例而言，資料處理器件102

102年12月4日修正替換頁

可經由網路連線804(諸如，網際網路)存取儲存於遠端伺服器上之外部知識庫814。

知識庫806可包括與所請求之情境式語音命令相關的使用歷史816。舉例而言，針對接收到之情境式語音命令「撥打電話給John」，知識庫可包括名爲John的人之清單及撥打電話給每個John的次數。在圖8b中展示之實例使用歷史816中，撥打電話給John Doe的次數最多(例如，31次)。因此，情境式語音命令模組510m將John Doe識別爲使用者之意圖的最佳假設。基於此識別，情境式語音命令模組510m經由輸出單元540(例如，顯示單元108之狀態列105)輸出一訊息，該訊息詢問使用者是否應撥打電話給John Doe。此外，情境式語音命令模組可呈現同一訊息之聲訊指示。回應於此視覺及/或聲訊指示，使用者可確認或校正該假設。

在一些實施中，知識庫可包括其他使用資訊，且可基於其他使用資訊來預測最佳假設。舉例而言，使用資訊可包括最近撥打過電話的John之身分、前些天在同一時間撥撥打過電話的John之身分、標爲最愛的John之身分、住在某一位置的John之身分等。

返回參看圖8a，情境式語音命令模式使用聲訊及/或視覺輸出將經識別之最佳假設傳達給使用者(830)。此允許使用者對經識別之假設進行任何校正。器件等待接收確認或校正經識別之最佳假設的使用者輸入(840)。然而，不需要使用者提示。可預設定或修改等待時間以提供一時間間隔，以供使用者使用言語(諸如「不！John Smith！」)取消或校正該假設命令或使用硬體按鈕之實體按壓取消該假設命令。在該間隔時間內不存在使用者提示之情況下，器件可執行由最佳假設指示之操作。情境式語音命令模組判定使用者輸入是否確認或校正了經識別之最佳假設(850)。當偵測到使用者輸入係確認經識別之最佳假設之輸入時，器件對經識別之最佳假設起作用(860)。

102年12月4日修正替換頁

如上所述，來自使用者之有效確認可包括在該時間間隔期間來自使用者之無回應。

當偵測到使用者輸入係校正經識別之最佳假設之輸入時，情境式語音命令控制器判定是否應進行不同假設或選項(870)。當情境式語音命令控制器偵測到使用者已校正最初識別之最佳假設時，情境式語音命令控制器基於使用者輸入進行不同假設(880)。然而，當情境式語音命令控制器偵測到使用者已僅僅指示經識別之最佳假設不正確時，情境式語音命令返回至知識庫以進行次最佳假設(890)。器件請求對經識別之次最佳假設的確認(830)，且程序800繼續，直至使用者意圖得以正確地識別。

舉例而言，當使用者說出預定觸發片語(諸如，「撥打電話給John」)時，器件檢查知識庫且識別出使用者總是(或時常，等)在其行動電話上撥打電話給John Doe。因此，器件撥打John Doe之行動電話號碼。然而，器件爲使用者提供改變命令之機會。器件可提供視覺或聲訊確認，諸如「您是否想要撥打電話給John Doe的行動電話」。爲此，使用者可藉由在接收到對撥打該號碼之確認後在預定持續時間內說「是」或什麼也不做來作出回應。爲了改變語音命令，使用者可說出預定觸發片語「不，我想撥打給John Smith」。針對此回應，器件檢查使用者歷史，且識別出使用者總是(或時常，等)在家撥打電話給John Smith。因此，器件撥打該號碼。藉由此程序，器件基於使用者歷史來建置知識庫以便減少詢問及確認之數目。基於此知識庫，器件可進行最佳假設。若假設是錯誤的，則讓使用者校正該假設。

爲了建置知識庫，使用者可教器件記住某些命令。舉例而言，情境式語音命令在作用中時，使用者可在影像處理應用程式中使一物件按45度旋轉。最初可使用位置輸入器件來執行使影像旋轉之動作。接著，使用者可藉由說「成45度角之影像」來選擇此旋轉後之物件且

識別此物件以教該器件。下次使用者想要使物件按45度旋轉時，使用者可直接說「按45度旋轉之物件」。

此外，可實施巨集指令來教該資料處理器件。舉例而言，可藉由將多個操作組合成一情境式語音命令來產生巨集指令。以下描述用於產生實例巨集指令之程序：

> 資料處理器件進入情境式語音命令模式。

> 情境式語音命令模組接收指示巨集指令產生之情境式語音命令，諸如，「學習此」。

> 情境式語音命令模組接收感興趣之資料項的使用者選擇及用於執行多個操作的情境式語音命令，諸如，「在備忘錄中使用」、「在2小時對我提醒備忘錄」、「將備忘錄作爲電子郵件發送至聯絡小組」等。

> 情境式語音命令模組接收指示巨集指令產生之結束的情境式語音命令，諸如，「完成學習」。

> 情境式語音命令模組提示使用者命名所產生之巨集指令。

> 情境式語音命令模組接收指示所產生之巨集指令之名稱(諸如，「執行操作」)的使用者輸入。

> 情境式語音命令模組確認接收到之名稱與所產生之巨集指令之關聯性。

> 結束巨集指令產生。

因此，巨集指令可用以使一連串操作與一描述性名稱(例如，「執行操作」)相關聯且教資料處理器件回應於使用者說出該名稱而執行一連串操作。下次使用者想要執行同一連串操作時，使用者可指導資料處理器件對所選感興趣之資料項「執行操作」。此所學習之情境式語音命令可由資料處理器件處理，以對所選感興趣之資料項執行全部一連串操作。

102年12月4日修正替換頁

使用可程式化巨集指令之序列學習可實施爲可在無語音命令之情況下操作的一單獨系統。當實施爲在情境式語音命令模組外部的一單獨系統時，情境式語音命令模組可實施爲控制該單獨系統的層。

爲了加快回應時間，可實施一過濾程式以自動濾除無意義之或在情境外說出之命令。舉例而言，回應於語音命令「撥打電話」，該過濾程式可自視覺及聲訊指示消除電子郵件地址、家庭地址等。在其他實例中，情境式語音命令控制器可經組態以將映射命令限制於街道地址且將加書籤命令限制於URL位址。

除了提供可用之選項的清單之外，視覺及聲訊指示亦可在任何時候用來獲得說明。使用者可獲得之說明包括在特定情形下輔助使用者之有幫助的建議。舉例而言，使用者可能困在一特定應用程式中且不知道如何離開該情形。使用者可藉由說出預定觸發片語(諸如，「幫我」或「幫我，我被困住了」)來請求說明，以獲得必要的指導。情境式語音控制器可輸出視覺及/或聲訊訊息以幫助使用者，諸如「您正試圖用電子郵件發送此影像」。又，使用者可獲得之說明包括用於對所選元素起作用之可能選項。舉例而言，爲了找出使用者可對所選元素做什麼，使用者可直接藉由說出觸發項(諸如，「我可對此做什麼」)進行詢問。

在收聽模式期間，由器件提供之視覺及/或聲訊指示(例如，提示)可回應於自使用者接收到之每一語音輸入而改變。舉例而言，當器件接收到語音輸入「撥打電話給John」且器件偵測到在電話簿中有多個John時，器件播放一要求使用者識別之音訊提示：「哪個John？」。又，視覺指示可變成顯示在使用者之電話簿中列出的名叫John之每個人之聯絡資訊。所顯示之聯絡資訊可包括每個John的相對應電話號碼之名稱(包括行動電話、家庭電話及工作電話)。若使用者說出語音命令以選擇可用之John中之一者(諸如，「John Smith」)且

John Smith與使用者之電話簿中之多個號碼相關聯，則器件可播放一要求使用者識別之音訊提示：「哪個號碼？」，且在螢幕上顯示多個號碼之名稱(例如，行動、工作、家庭)。

在一些實施中，可使用情境式語音命令接收來自器件之關於所選感興趣之元素的資訊或回饋。舉例而言，在媒體專輯應用程式中，使用者可說「尋找關於第四曲目之更多資訊」以獲得關於該曲目之資訊。由於使用者已在媒體專輯情境中，因此情境式語音命令控制器知道該命令之情境。在非情境式語音命令中，使用者將必須(例如)藉由說「尋找關於藝術家X及其歌曲Y之更多資訊」特定地識別藝術家之名字及歌名。可自器件接收其他類型之回饋。舉例而言，使用者可說出預定觸發片語(諸如，「此有多長」)以接收來自器件之答覆。又，使用者可藉由說出預定觸發片語(諸如，「閱讀此」、「寫出此」、「顯示此」等)要求器件提供其他聲訊回饋。

如上所述，情境式語音命令之情境可與使用者所選元素所在之特定應用程式相關聯。在一些實施中，語音命令之情境亦與使用該器件之環境相關聯。此環境情境之實例包括汽車情境、導航情境、會議情境等。

在導航情境中，可使用電話之定向或位置來定製情境式語音命令以用於向使用者提供指導。可使用GPS及羅盤提供指導。舉例而言，使用者可說出預定觸發片語(諸如，「我應該走哪條路？」)以自器件接收適當指導。情境式語音命令控制器可使用該器件的基於GPS及羅盤之位置向使用者提供適當的聲訊指導，諸如，「左轉」、「直走」等。此外，使用者可使用情境式語音命令自器件獲得方向資訊。舉例而言，使用者可說出預定觸發術語，諸如，「此器件指向什麼」或「我看到的是什麼」或「我左邊是什麼？」或「我右邊是什麼？」或「我前面是什麼？」或「前面是什麼」或「我後面是什麼」或「哪

102年12月4日修正替換頁

條路到目的地」或「目的地有多遠」等。在此等實例中，器件之定向可爲情境式元素。此等類型之觸發片語可經預定以使器件觸發適當之導航相關應用程式。

針對使用者不能經由位置輸入器件提供輸入之某些情形(諸如，汽車情境)，可定製一聲訊回饋機制以用於語音密集型互動。舉例而言，情境式語音命令控制器可基於情境(諸如，駕駛情境)簡化自器件發送之聲訊回應。舉例而言，可將來自器件之聲訊回應限制於簡單的「是-否」情形，而不是情境式語音命令控制器提供聲訊輸出(諸如「有多個名稱，您想要哪一個？」)以提供選項之清單。舉例而言，器件可詢問「您是否想要此？是或是否。」舉例而言，此簡單的「是-否」回應容易由使用者在駕駛時使用。

可使用各種機制使器件意識到此情境。舉例而言，器件可偵測到使用者正使用藍芽器件提供語音輸入。又，器件可偵測到器件被插入至汽車立體聲之輸入/輸出插口內。又，使用者可按壓、觸摸一實體或虛擬按鈕或將其按住不放以指示此情境。此外，器件可偵測到器件正使用GPS及/或羅盤移動。舉例而言，當位於移動之運輸工具內部時，器件之位置可關於時間改變。可偵測到器件之此移動或位置改變且用來觸發不同回饋模式。

可藉由安裝不同語音來改變由情境式語音命令控制器產生的聲訊輸出之語音。又，可將不同語音用於不同情境。此外，在本說明書中描述之技術、裝置及系統可擴展至任何口頭語言，其包括專用命令語言。舉例而言，可實施語音命令(情境式及非情境式)以允許器件辨識英語、法語、德語、韓語、漢語(普通話及廣東話)、日語、意大利語、西班牙語、波斯語等。又，此等口頭語言中之任一者可用以產生專用命令語言，諸如，經設計以自器件得出高效回應之一命令集。

情境式語音命令不限於當前功能性，而是可經調適以控制未來

102年12月4日修正替換頁

的功能性。舉例而言，器件之任何額外功能性可受情境式語音命令模組控制。藉由將情境式語音命令模組實施爲新功能性之外掛程式，使此成爲可能。又，第三方開發者可具備向系統指示額外的非情境式命令及情境式命令可用的一選項(例如，使用軟體開發套件(SDK))。

圖9爲資料處理器件之實例架構900之方塊圖。資料處理器件102及400可包括：一記憶體介面902、一或多個資料處理器、影像處理器及/或中央處理單元904，及一周邊裝置介面906。該記憶體介面902、該一或多個資料處理器904及/或該周邊裝置介面906可爲單獨的組件或可整合於一或多個積體電路中。資料處理器件102及400中之各種組件可藉由一或多個通信匯流排或信號線耦接在一起。

感測器、器件及子系統可耦接至周邊裝置介面906以促進多個功能性。舉例而言，一運動感測器910、一光感測器912及一接近度感測器914可耦接至周邊裝置介面906以促進定向、照明及接近度功能。一位置處理器915(例如，GPS接收器)可連接至周邊裝置介面906以提供地理定位。一磁性羅盤積體電路916亦可連接至周邊裝置介面906以提供定向(例如，以判定正北方向)。

一相機子系統920及一光學感測器922(例如，電荷耦合器件(CCD)或互補金屬氧化物半導體(CMOS)光學感測器)可用以促進相機功能，諸如，記錄照片及視訊片段。

可經由一或多個無線通信子系統924促進通信功能，該一或多個無線通信子系統可包括射頻接收器與發射器及/或光學(例如，紅外線)接收器與發射器。通信子系統924之具體設計及實施可取決於通信網路，資料處理器件102及400意欲在該(等)通信網路上操作。舉例而言，資料處理器件102及400可包括經設計以在GSM網路、GPRS網路、EDGE網路、Wi-Fi或WiMax網路及Bluetooth™網路上操作之通信子系統924。詳言之，無線通信子系統924可包括代管協定，使得資料

處理器件102及400可被組態爲用於其他無線器件之基地台。

一音訊子系統926可耦接至一揚聲器928及一麥克風930以促進語音啓用功能，諸如，語音辨識、語音複製、數位記錄及電話功能。

I/O子系統940可包括一觸控式螢幕控制器942及/或其他輸入控制器944(如關於圖1a及圖1b描述)。舉例而言，I/O子系統940可包括一麥克風(內部及/或外部)、一揚聲器及一語音命令辨識引擎。I/O子系統940可經由全雙工通信來接收語音命令並呈現音訊輸出。舉例而言，可實施不同於正規蜂巢式語音通信的傳輸技術，諸如，網路電話。

又，可使用兩遍式程序來處理語音命令。器件上之語音命令模組可處理接收到之語音命令以執行一般化辨識。可將接收到之語音命令的音訊資料發送至伺服器以提供更詳細且準確的處理。相比於行動器件，伺服器可經更好地配備(例如，使用更快且更強大的處理器)以執行語音命令辨識。爲了減少頻寬要求及延時問題，可不將音訊資料全部發送至伺服器。舉例而言，器件上之語音命令模組可處理語音命令以識別成串的數字，但不能識別精確的語音命令。因此，器件上之語音命令模組可判定語音命令或言語含有「一些數字」。可將音訊資料之較大周圍區段發送至伺服器，且伺服器可不同步地傳回對該等語音命令中實際所說內容的好得多之理解。藉由以此方式使用伺服器，可獲得伺服器處理之益處，同時減少或最小化伺服器處理所涉及之成本。

觸控式螢幕控制器942可耦接至觸控式螢幕946。觸控式螢幕946及觸控式螢幕控制器942可(例如)使用複數種觸摸敏感性技術(包括(但不限於)電容性、電阻性、紅外線及表面聲波技術，以及其他接近度感測器陣列或用於判定與觸控式螢幕946之一或多個接觸點的其他元件)中之任一者來偵測接觸及移動或其中斷。

該(等)其他輸入控制器944可耦接至其他輸入/控制器件948，諸

如，一或多個按鈕、搖臂開關、拇指旋輪、紅外線埠、USB埠及/或指標器件(諸如，電筆)。一或多個按鈕(未繪示)可包括用於揚聲器928及/或麥克風930之音量控制的一上/下按鈕。

在一實施中，歷時第一持續時間的對該按鈕之按壓可解開觸控式螢幕946之鎖定；且歷時比第一持續時間長之第二持續時間的對該按鈕之按壓可接通或切斷至資料處理器件102及400之電力。使用者可能能夠定製該等按鈕中之一或多者的功能性。舉例而言，觸控式螢幕946亦可用以實施虛擬或軟按鈕及/或鍵盤。

在一些實施中，資料處理器件102及400可呈現所記錄之音訊及/或視訊檔案，諸如，MP3、AAC及MPEG檔案。在一些實施中，資料處理器件102及400可包括MP3播放器(諸如，iPod Touch™)之功能性。

記憶體介面902可耦接至記憶體950。記憶體950可包括高速隨機存取記憶體及/或非揮發性記憶體，諸如，一或多個磁碟儲存器件、一或多個光學儲存器件及/或快閃記憶體(例如，NAND、NOR)。記憶體950可儲存一作業系統952，諸如，Darwin、RTXC、LINUX、UNIX、OS X、WINDOWS或嵌式作業系統(諸如VxWorks)。作業系統952可包括用於處置基本系統服務及用於執行硬體相依任務之指令。在一些實施中，作業系統952可爲核心(例如，UNIX核心)。

記憶體950亦可儲存通信指令954，其用以促進與一或多個額外器件、一或多個電腦及/或一或多個伺服器之通信。記憶體950可包括：圖形使用者介面指令956，其用以促進圖形使用者介面處理；感測器處理指令958，其用以促進與感測器相關之處理及功能；電話指令960，其用以促進電話相關之程序及功能；電子傳訊息指令962，其用以促進與電子傳訊息相關之程序及功能；網路瀏覽指令964，其用以促進與網路瀏覽相關之程序及功能；媒體處理指令966，其用以促

進與媒體處理相關之程序及功能；GPS/導航指令968，其用以促進與GPS及導航相關之程序及功能；相機指令970，其用以促進與相機相關之程序及功能；及語音命令指令972，其用以促進使用情境式語音命令對資料處理器件102進行的操作(如參看圖1a、圖1b、圖2a、圖2b、圖3a及圖3b所描述)。在一些實施中，GUI指令956及/或媒體處理指令966實施參看圖1至圖8描述之特徵及操作。

記憶體950亦可儲存其他軟體指令(未繪示)，諸如：網路視訊指令，其用以促進與網路視訊相關之程序及功能；及/或網路購物指令，其用以促進與網路購物相關之程序及功能。在一些實施中，媒體處理指令966被分成音訊處理指令及視訊處理指令，其用以分別促進與音訊處理相關之程序及功能及與視訊處理相關之程序及功能。啓動記錄及國際行動設備識別碼(IMEI)974或類似的硬體識別符亦可儲存於記憶體950中。

以上識別之指令及應用程式中之每一者可對應於用於執行上文描述之一或多個功能的一指令集。不需要將此等指令實施爲單獨的軟體程式、程序或模組。記憶體950可包括額外指令或較少指令。此外，資料處理器件102及400之各種功能可實施於硬體中及/或軟體中，包括實施於一或多個信號處理及/或特殊應用積體電路中。

可使用一或多個電腦程式(包含儲存於有形電腦可讀媒體上且在資料處理器件102上執行之可執行程式碼)來實施如圖1至圖9中所描述的用於實施情境式語音命令之技術。電腦可讀媒體可包括硬碟機、快閃記憶體器件、諸如DRAM及SDRAM之隨機存取記憶體器件、諸如CD-ROM及DVD-ROM之抽取式儲存媒體、磁帶、軟碟、緊密快閃記憶卡、安全數位(SD)記憶卡或某一其他儲存器件。在一些實施中，電腦可執行程式碼可包括多個部分或模組，每一部分經設計以執行關於圖6至圖8所描述之一特定功能。在一些實施中，可使用硬體(諸如微

102年12月4日修正替換

處理器、微控制器、具有內部記憶體嵌式微控制器或可抹除可程式化唯讀記憶體(EPROM))來實施該等技術，該硬體編碼用於執行關於圖6至圖8所描述之技術的電腦可執行指令。在其他實施中，可使用軟體與硬體之組合來實施該等技術。

適合於執行電腦程式之處理器包括(舉例而言)通用微處理器與專用微處理器，及任何種類之數位電腦之任何一或多個處理器(包括諸如GPU之圖形處理器)。通常，處理器將自唯讀記憶體或隨機存取記憶體或兩者接收指令及資料。電腦之元件爲：用於執行指令之一處理器及用於儲存指令及資料之一或多個記憶體器件。通常，電腦將亦包括用於儲存資料之一或多個大容量儲存器件(例如，磁碟、磁光碟或光碟)，或經操作性地耦接以自該一或多個大容量儲存器件接收資料或傳送資料至該一或多個大容量儲存器件或者接收且傳送資料。適合於體現電腦程式指令及資料之資訊載體包括所有形式之非揮發性記憶體，包括(舉例而言)：半導體記憶體器件，例如，EPROM、EEPROM及快閃記憶體器件；磁碟，例如，內部硬碟或抽取式碟片；磁光碟；及CD ROM及DVD-ROM碟片。處理器及記憶體可由專用邏輯電路進行補充或併入於專用邏輯電路中。

爲了提供與使用者之互動，可將此處描述之系統裝置及技術實施於一資料處理器件上，該資料處理器件具有：一顯示器件(例如，CRT(陰極射線管)或LCD(液晶顯示器)監視器)，其用於向使用者顯示資訊；及一位置輸入器件，諸如鍵盤及指標器件(例如，滑鼠或軌跡球)，使用者可藉由該位置輸入器件將輸入提供至電腦。亦可使用其他種類之器件以提供與使用者之互動；舉例而言，提供給使用者之回饋可爲任何形式之感官回饋(例如，視覺回饋、聽覺回饋或觸覺反饋)；且可以任何形式接收來自使用者之輸入，包括聲學、話語或觸覺輸入。

102年12月4日修正替換頁

雖然本說明書含有許多細節，但此等細節不應解釋爲對任何發明或所主張內容之範疇的限制，而是應解釋爲對特定發明之特定實施例所特有之特徵的描述。本說明書中在單獨之實施例之情況下描述之某些特徵亦可在單一實施例中以組合實施。反過來，在單一實施例之情況下描述的各種特徵亦可單獨地實施於多個實施例中或以任何合適的子組合實施。此外，儘管上文可將特徵描述爲以某些組合起作用且甚至最初按此主張，但來自所主張之組合之一或多個特徵在一些情況下可自該組合被刪除，且所主張之組合可針對子組合或子組合之變化。

類似地，雖然在圖式中按特定次序描繪操作，但不應將此理解爲：需要按展示之特定次序或按順序執行此等操作或需要執行所有說明之操作來達成所要結果。在某些情況下，多任務及並行處理可爲有利的。此外，不應將上述實施例中各種系統組件之分離理解爲所有實施例中需要此分離，且應理解，所描述之程式組件及系統可通常在單一軟體產品中整合在一起或封裝至多個軟體產品中。

僅描述了幾個實施及實例，且可基於本申請案中所描述及說明之內容進行其他實施、增強及變化。

【符號說明】

102	資料處理器件
103	手指
104	揚聲器
105	狀態列
106	麥克風
107	影像
108	顯示單元
110a	電話應用程式圖示

102年12月4日修正替換頁

110b	電子郵件應用程式圖示
110c	網路瀏覽器應用程式圖示
110d	音樂播放器應用程式圖示
110e	媒體播放器應用程式圖示
110f	音樂下載應用程式圖示
110g	影像處理應用程式圖示
110h	地理定位應用程式圖示
110i	聯絡人應用程式圖示
110j	簡訊服務(SMS)應用程式圖示
110k	視訊遊戲應用程式圖示
110l	文書處理應用程式圖示
110m	情境式語音命令圖示
400	個人電腦/資料處理器件
404	作用中螢幕
406	電子郵件
510a	電話應用程式
510b	電子郵件應用程式
510c	網路瀏覽器
510d	音樂播放器應用程式
510e	媒體播放器應用程式
510f	音樂下載應用程式
510g	影像處理應用程式
510h	地理定位應用程式
510i	聯絡人應用程式
510j	簡訊服務(SMS)應用程式
510k	視訊遊戲應用程式

102年12月4日修正替換頁

510l	文書處理應用程式
510m	情境式語音命令模組
520	輸入單元
522	音訊輸入單元
524	實體輸入單元
540	輸出單元
600	用於實施情境式語音命令的實例程序
700	用於實施確認模型的實例程序
800	用於實施最佳假設或學習模型的實例程序
802	語音輸入
806	知識庫
812	本端知識庫
814	外部知識庫
816	使用歷史
820	網路
900	資料處理器件之實例架構
902	記憶體介面
904	資料處理器、影像處理器及/或中央處理單元
906	周邊裝置介面
910	運動感測器
912	光感測器
914	接近度感測器
915	位置處理器
916	磁性羅盤積體電路
920	相機子系統
922	光學感測器

924	無線通信子系統
926	音訊子系統
928	揚聲器
930	麥克風
940	輸入/輸出(I/O)子系統
942	觸控式螢幕控制器
944	其他輸入控制器
946	觸控式螢幕
948	其他輸入/控制器件
950	記憶體
952	作業系統
954	通信指令
956	圖形使用者介面指令
958	感測器處理指令
960	電話指令
962	電子傳訊息指令
964	網路瀏覽指令
966	媒體處理指令
968	全球定位系統(GPS)/導航指令
970	相機指令
972	語音命令指令
974	啓動記錄及國際行動設備識別碼(IMEI)

第099118245號專利申請案
中文申請專利範圍替換本(104年6月)

申請專利範圍

104年6月5日修正本

1. 一種用於情境式語音命令之方法，其包含：

在一器件上顯示：

一圖像，其代表在一第一顯示區域中之一資料項，及

遠離該第一顯示區域之一第二顯示區域；

在該器件上接收該圖像之一選擇；

回應於該選擇，在該第二顯示區域中顯示可對該資料項執行之複數個操作，其中該複數個操作之至少一操作可回應於一語音輸入以對該資料項執行，且其中該第二顯示區域係與在該第一顯示區域中之該圖像同時顯示；

在該器件上接收對應於可對該資料項執行之該至少一操作之一語音輸入；及

回應於接收該語音輸入以在該器件上對該資料項執行該至少一操作。

2. 如請求項1之方法，其進一步包含：

呈現該複數個操作之一或多者之一聲訊指示。

3. 如請求項2之方法，其中呈現該複數個操作之一或多者之該聲訊指示包含呈現在一或多個聲訊問題中之該一或多個操作至該使用者；及

接收該語音輸入包含接收對一聲訊問題之一口頭答覆以指示該至少一操作的一使用者選擇。

4. 如請求項1之方法，其中該圖像係顯示於與一第一應用程式環境相關聯之一介面內。

5. 如請求項1之方法，其中該第二顯示區域不阻擋該圖像。

6. 如請求項5之方法，其中該第一顯示區域及該第二顯示區域係由

一背景區域之一部分而視覺地分離。

7. 如請求項1之方法，其中該經顯示之複數個操作係可經由語音輸入而選擇且不可經由觸覺輸入而選擇。

8. 如請求項1之方法，其進一步包含：

顯示一第二圖像，其代表在不同於該第一顯示區域之一第三顯示區域中之一第二資料項；

在該器件上接收該第二圖像之一選擇；

回應於該選擇，在該第二顯示區域顯示可對該第二資料項執行之第二複數個操作。

9. 一種用於情境式語音命令之方法，其包含：

在一資料處理器件處接收對應於用於一個以上操作的一語音輸入；

藉由該資料處理器件搜尋一知識庫，該知識庫包括一使用歷史，其回應於與該經接收之語音輸入相似的一先前語音輸入以儲存先前已被執行之複數個操作；

藉由該器件識別該複數個操作中之一者；

在該資料處理器件之一顯示器上之一使用者介面中呈現對應於該複數個操作之該經識別者的一物件以回應於接收該語音輸入以指示將被執行之一操作；及

當一預定時間間隔失效而無接收一取消請求時，或當一確認係在該預定時間間隔之期間被接收時，執行該複數個操作之該經識別者。

10. 如請求項9之方法，其包含：

在該器件處呈現一聲訊通知，其提示一使用者確認該複數個操作之該經識別者是否被執行。

11. 如請求項9之方法，其包含：回應於該經接收之語音輸入，藉由

該器件濾除儲存於該使用歷史中之該複數個操作之一或多者。

12. 如請求項9之方法，其中因該語音輸入匹配於在與該器件之一使用者相關聯之一通訊錄中的一個以上之聯絡人，該語音輸入對應於一個以上之操作。

13. 如請求項9之方法，其中因該語音輸入匹配於在與一個以上之電話號碼或電子郵件地址相關聯之一通訊錄中的一聯絡人，該語音輸入對應於一個以上之操作。

14. 一種用於情境式語音命令之裝置，其包含：

一輸出單元，其：

呈現一圖像，該圖像代表在一第一顯示區域中之一資料項；及

回應於該圖像之一選擇以在遠離該第一顯示區域之一第二顯示區域中呈現可對該資料項執行之複數個操作，其中該複數個操作之至少一操作可回應於一語音輸入以對該資料項執行，且其中該第二顯示區域係與在該第一顯示區域中之該圖像同時顯示；

一輸入單元，該輸入單元接收使用者輸入，其包含對應於可對該該資料項執行之該至少一操作之一語音輸入；及

一控制器，其與該輸入單元及該輸出單元通信以回應於接收該語音輸入以對該資料項執行該至少一操作。

15. 如請求項14之裝置，其中該控制器經組態以造成該輸出單元呈現該複數個操作之一或多者之一聲訊指示。

16. 如請求項15之裝置，其中該控制器經組態以造成該輸出單元在一或多個聲訊問題中呈現該一或多個操作。

17. 如請求項15之裝置，其中該控制器經組態以：

判定該經接收之語音輸入包含對應於該複數個操作之一者以

上的一第一語音輸入；及

回應於該判定：

搜尋一知識庫，該知識庫包括一使用歷史，其回應於與該經接收之第一語音輸入相似的一先前語音輸入以儲存先前已被執行之複數個操作，及

在該使用歷史中選擇該複數個操作中之一者作為將對該資料項執行之該至少一操作。

18. 如請求項17之裝置，其中該控制器經組態以指導該輸出單元向使用者呈現一提示，以確認該所選操作是否將被執行；且

該輸入單元經組態以接收一第二語音輸入，其確認或取消該所選操作。

19. 如請求項15之裝置，其中：

該接收到之語音輸入包含對應於該複數個操作之一者以上的一第一語音輸入；

該輸出單元經組態以藉由向該使用者提出一聲訊問題來呈現該一個以上相對應操作之一選項；且

該輸入單元經組態以接收一作為對該問題之一口頭答覆的使該資料項與該等相對應操作中之一者相關的第二語音輸入。

20. 如請求項15之裝置，其中該控制器經組態以在該輸出單元呈現該複數個操作之前濾除一操作。

21. 如請求項15之裝置，其中該輸入單元包含一外部麥克風。

22. 如請求項15之裝置，其中該輸入單元經組態以接收一實體輸入，其選擇該圖像。

23. 如請求項17之裝置，該輸出單元進一步經組態以呈現一物件，該物件對應於該複數個操作之該經識別者以回應於接收該語音輸入以指示將被執行之一操作；及

該控制器進一步經組態以當一預定時間間隔失效而無接收一取消請求時，或當一確認係在該預定時間間隔之期間被接收時，執行該複數個操作之該經識別者。

24. 一種非過渡電腦可讀媒體，其體現指令，該等指令在由一處理器執行時使該處理器執行包含以下各項之操作：

指導一輸出單元以：

呈現一圖像，該圖像代表在一第一顯示區域中之一資料項；及

回應於該圖像之一選擇以在遠離該第一顯示區域之一第二顯示區域中呈現可對該資料項執行之複數個操作，其中該複數個操作之至少一操作可回應於一語音輸入以對該資料項執行，且其中該第二顯示區域係與在該第一顯示區域中之該圖像同時顯示；

處理經由一語音輸入單元接收到之一語音輸入，該語音輸入對應於可對該資料項執行之該至少一操作；及回應於接收該語音輸入以對該資料項執行該至少一操作。

25. 如請求項24之電腦可讀媒體，其中指導該輸出單元以呈現該複數個操作包含指導該輸出單元呈現該複數個操作中之一或多者之一聲訊指示。

26. 一種用於情境式語音命令之方法，其包含：

在一資料處理器件處顯示一情境中之一資料項；

在該資料處理器件處接收指示該資料項之一選擇的一輸入；

在該資料處理器件處接收使該所選資料項與包含一關鍵字之一識別符相關聯的一語音輸入；

在該資料處理器件處接收另一語音輸入，其包含該關鍵字及包含將被執行之一操作；及

回應於接收該另一語音輸入以對該資料項執行包含於該另一語音輸入之該操作，包含於該另一語音輸入之該關鍵字係與該資料項相關聯。

27. 如請求項26之方法，其包含：

在該資料處理器件處接收修改該資料項之輸入；及

在該資料處理器件處接收使該經修改之資料項與包含反映該修改之一第二關鍵字之另一識別符相關聯的另一語音輸入。

圖式

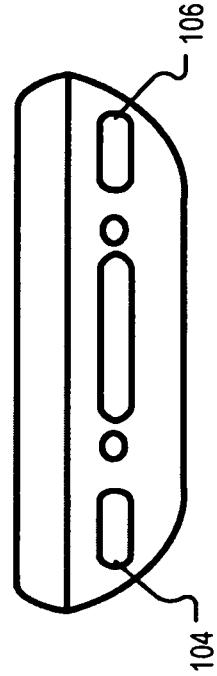
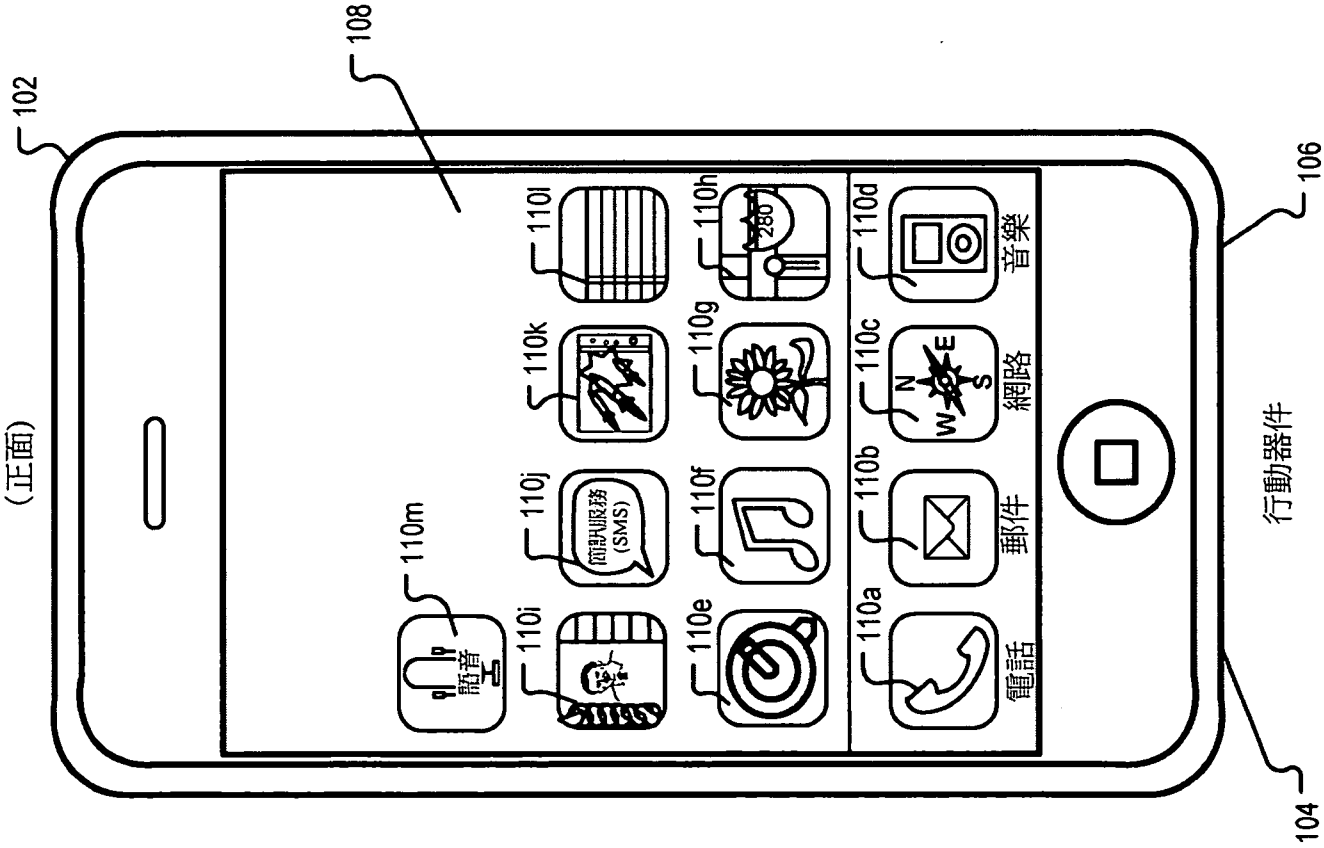
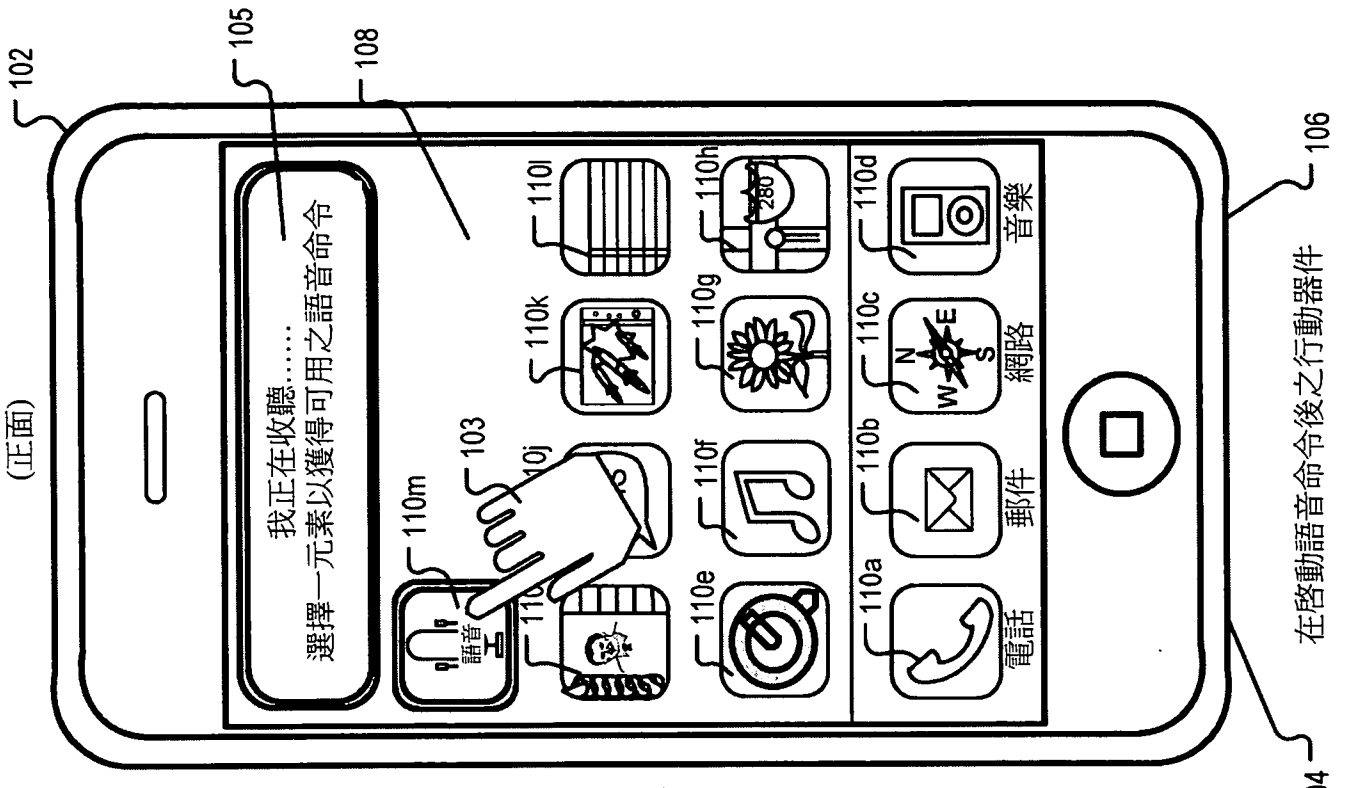


圖1b



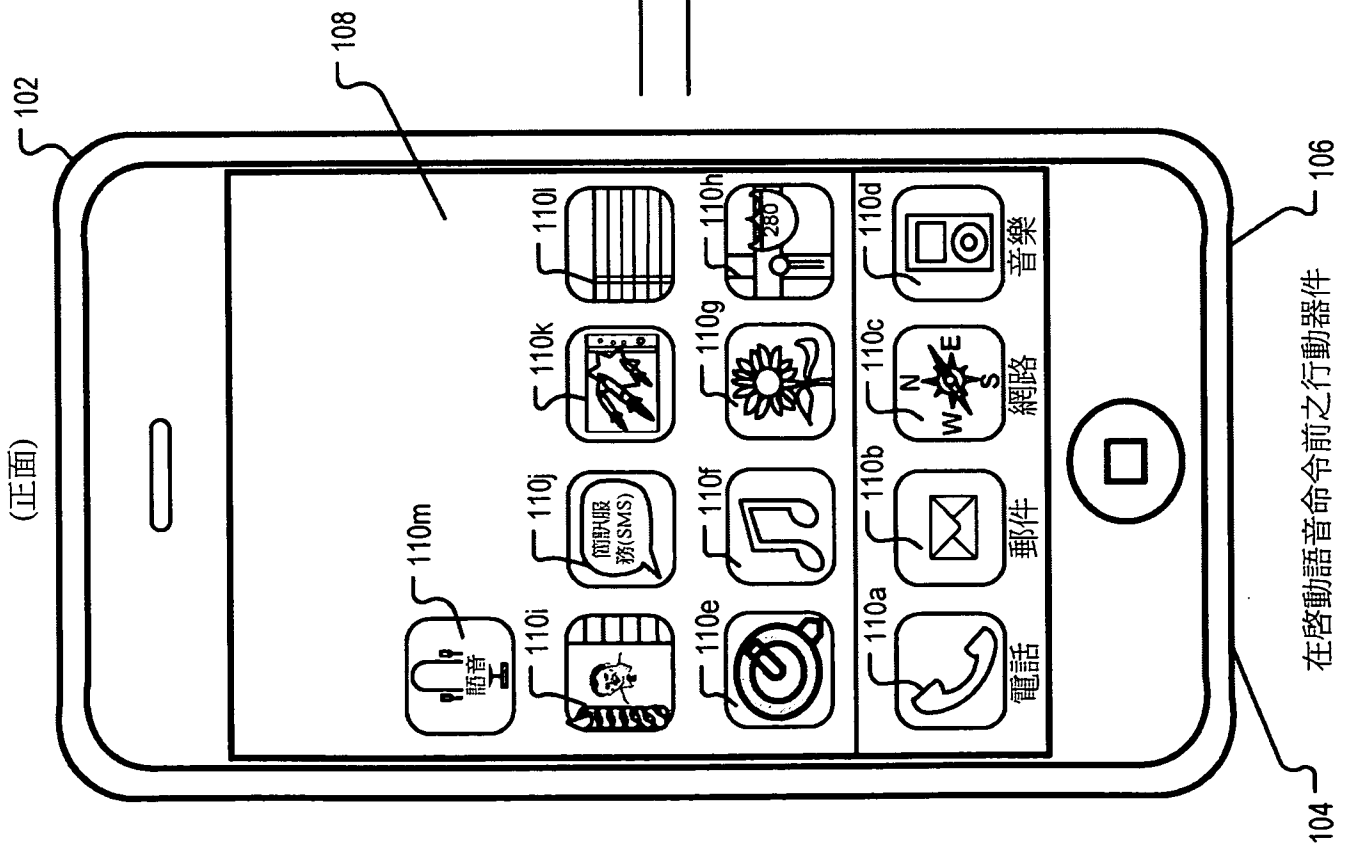
行動器件

圖1a



在啓動語音命令後之行動器件

圖2b



在啓動語音命令前之行動器件

圖2a

C148305DBX201312C.pdf

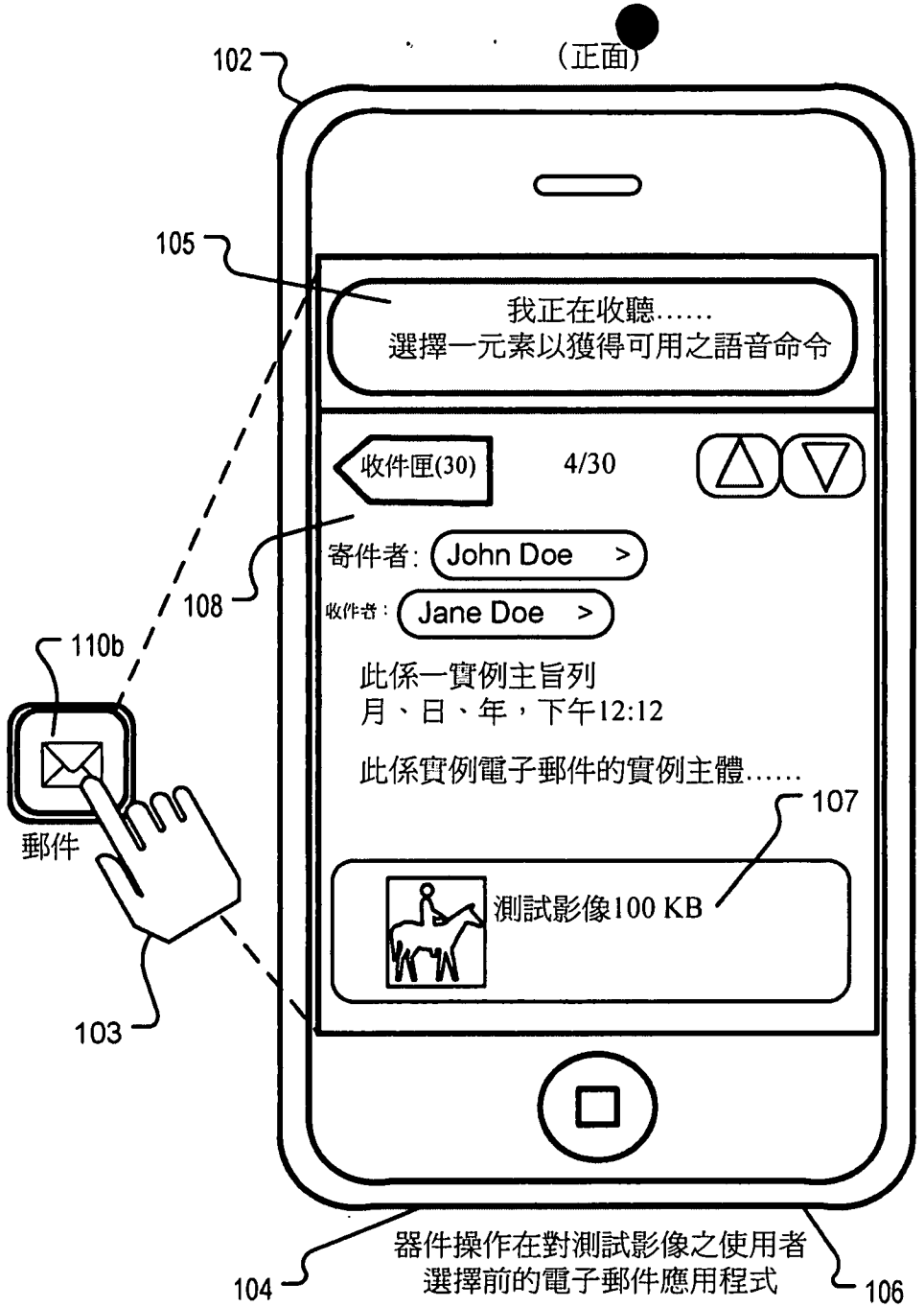


圖3a

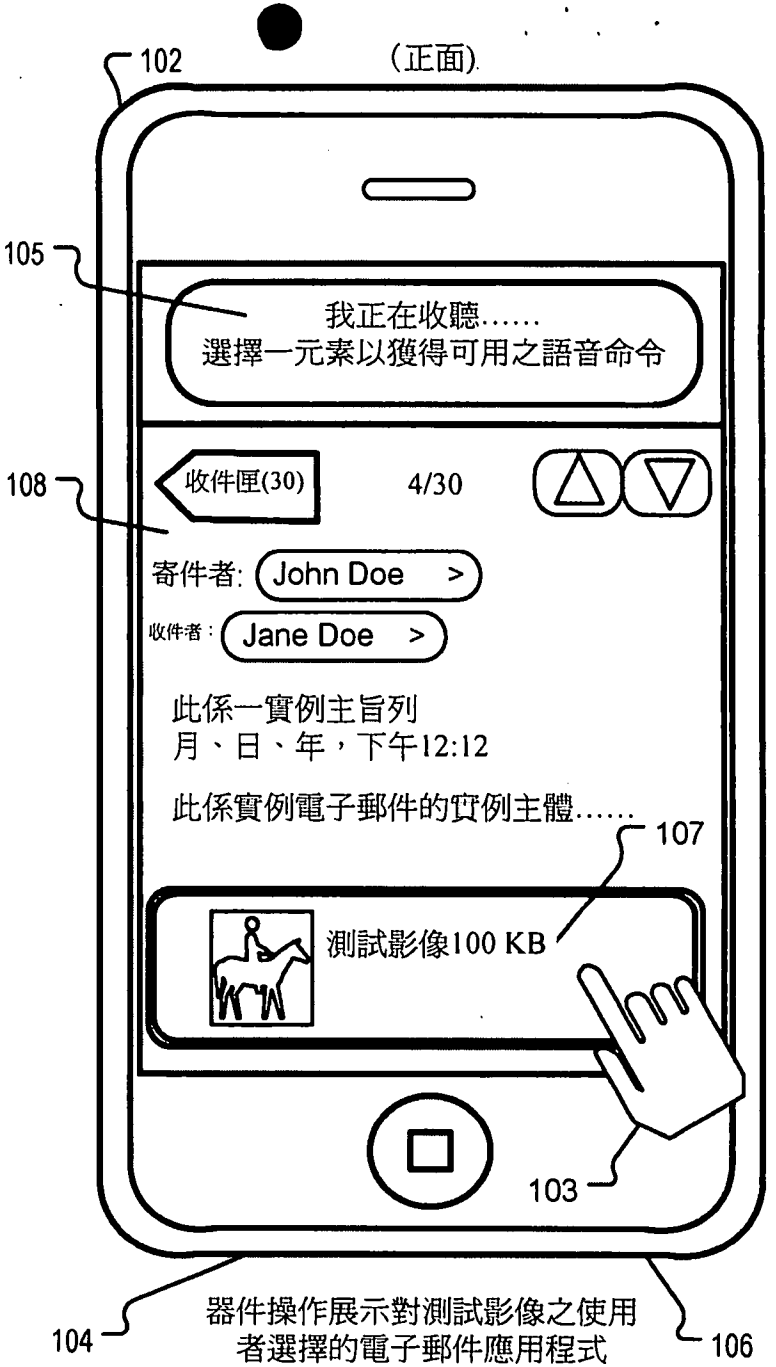


圖3b

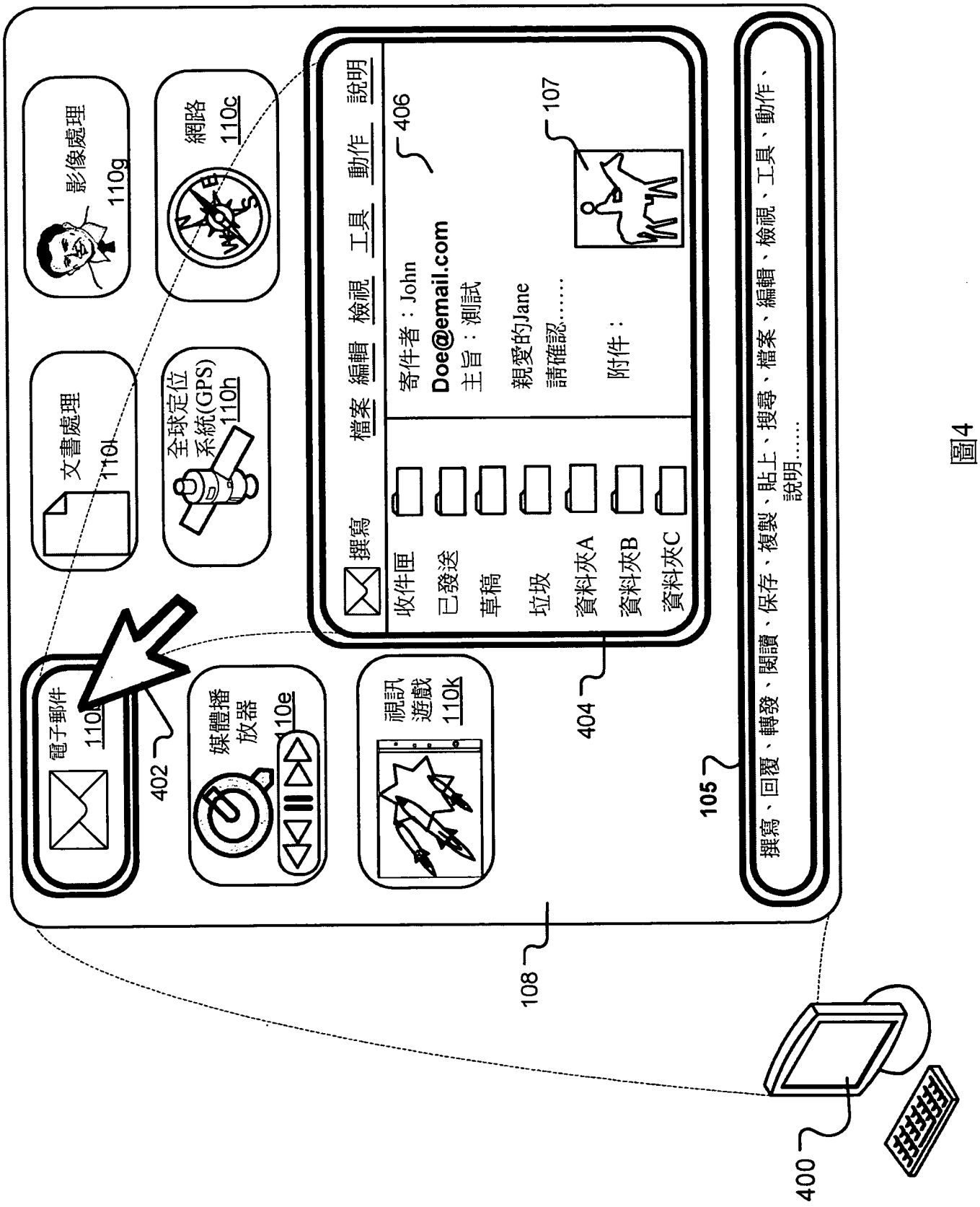

影像處理
110g
網路
110c
文書處理
110l
全球定位系統(GPS)
110h
電子郵件
110i
媒體播放器
110e
視訊遊戲
110k
撰寫
檔案
編輯
檢視
工具
動作
說明
收件匣
已發送
草稿
垃圾
資料夾A
資料夾B
資料夾C
寄件者：John
Doe@email.com
主旨：測試
親愛的Jane
請確認……
附件：
406
107
404
402
105
108
400
撰寫、回覆、轉發、閱讀、保存、複製、貼上、搜尋、檔案、編輯、檢視、工具、動作、說明……

圖4

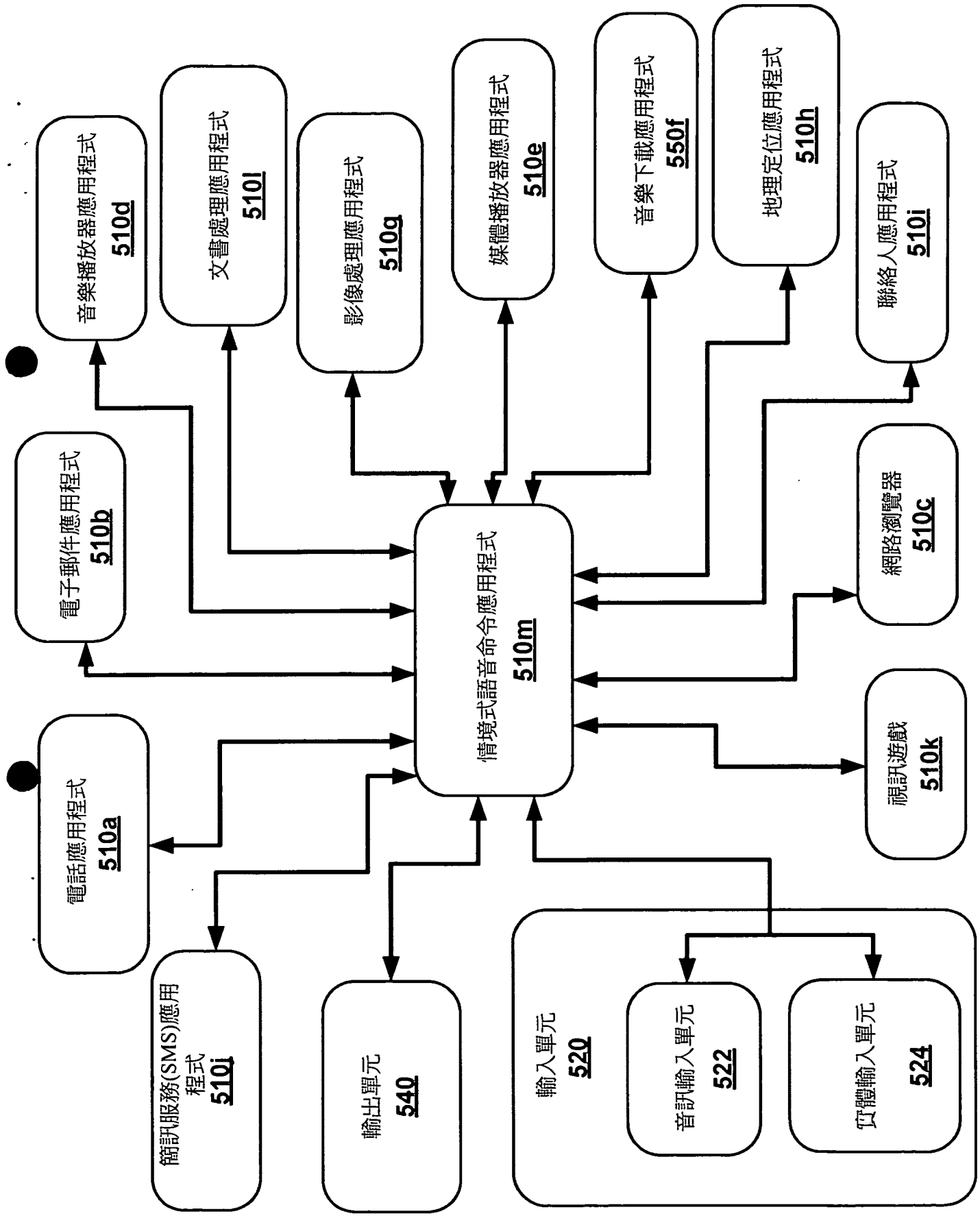

音樂播放器應用程式
510d
文書處理應用程式
510l
影像處理應用程式
510g
媒體播放器應用程式
510e
音樂下載應用程式
550f
地理定位應用程式
510h
聯絡人應用程式
510i
電子郵件應用程式
510b
網路瀏覽器
510c
情境式語音命令應用程式
510m
電話應用程式
510a
視訊遊戲
510k
簡訊服務(SMS)應用程式
510j
輸出單元
540
輸入單元
520
音訊輸入單元
522
實體輸入單元
524

圖5

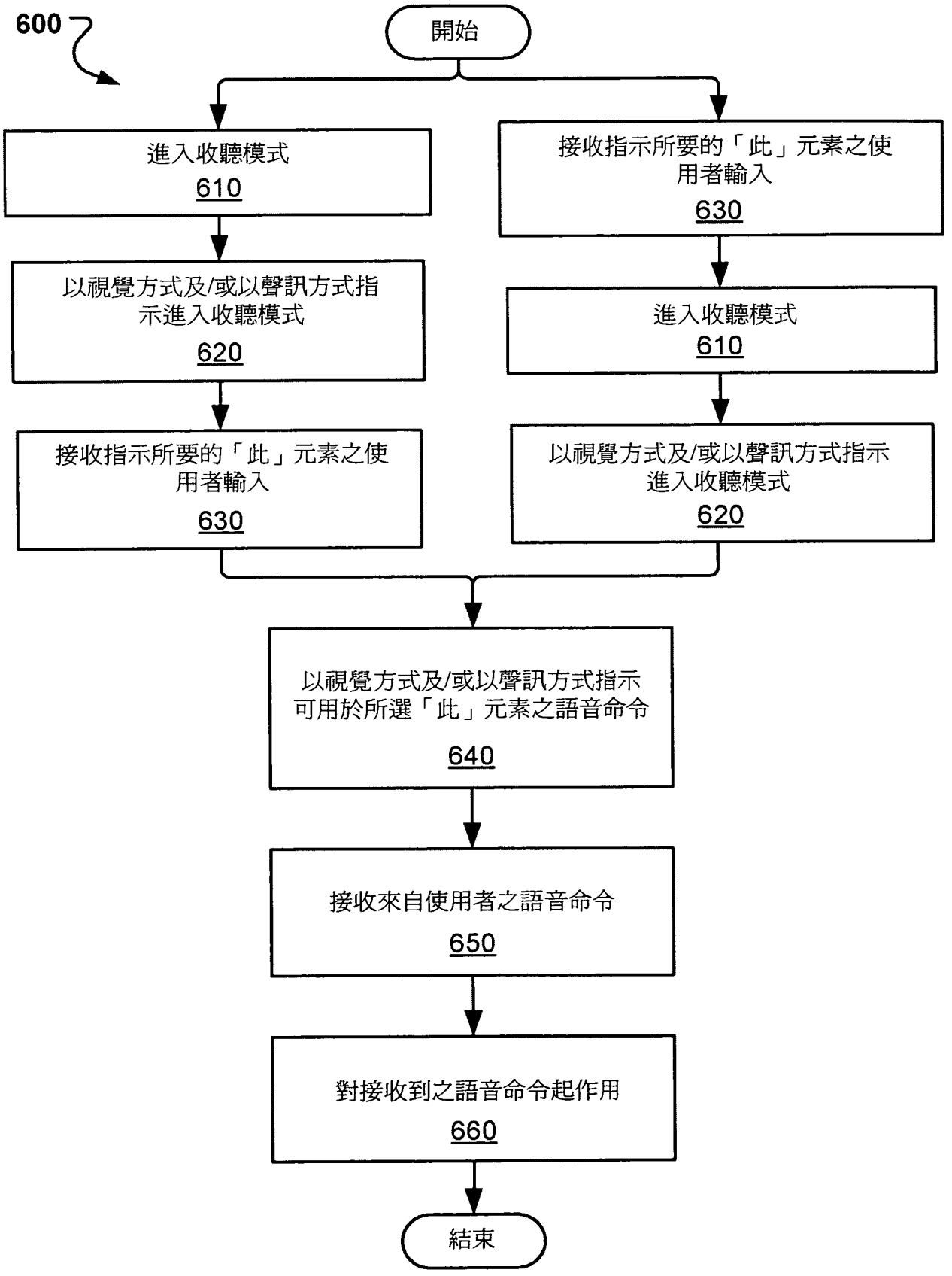

600
開始
進入收聽模式
610
以視覺方式及/或以聲訊方式指示進入收聽模式
620
接收指示所要的「此」元素之使用者輸入
630
接收指示所要的「此」元素之使用者輸入
630
進入收聽模式
610
以視覺方式及/或以聲訊方式指示進入收聽模式
620
以視覺方式及/或以聲訊方式指示可用於所選「此」元素之語音命令
640
接收來自使用者之語音命令
650
對接收到之語音命令起作用
660
結束

圖6

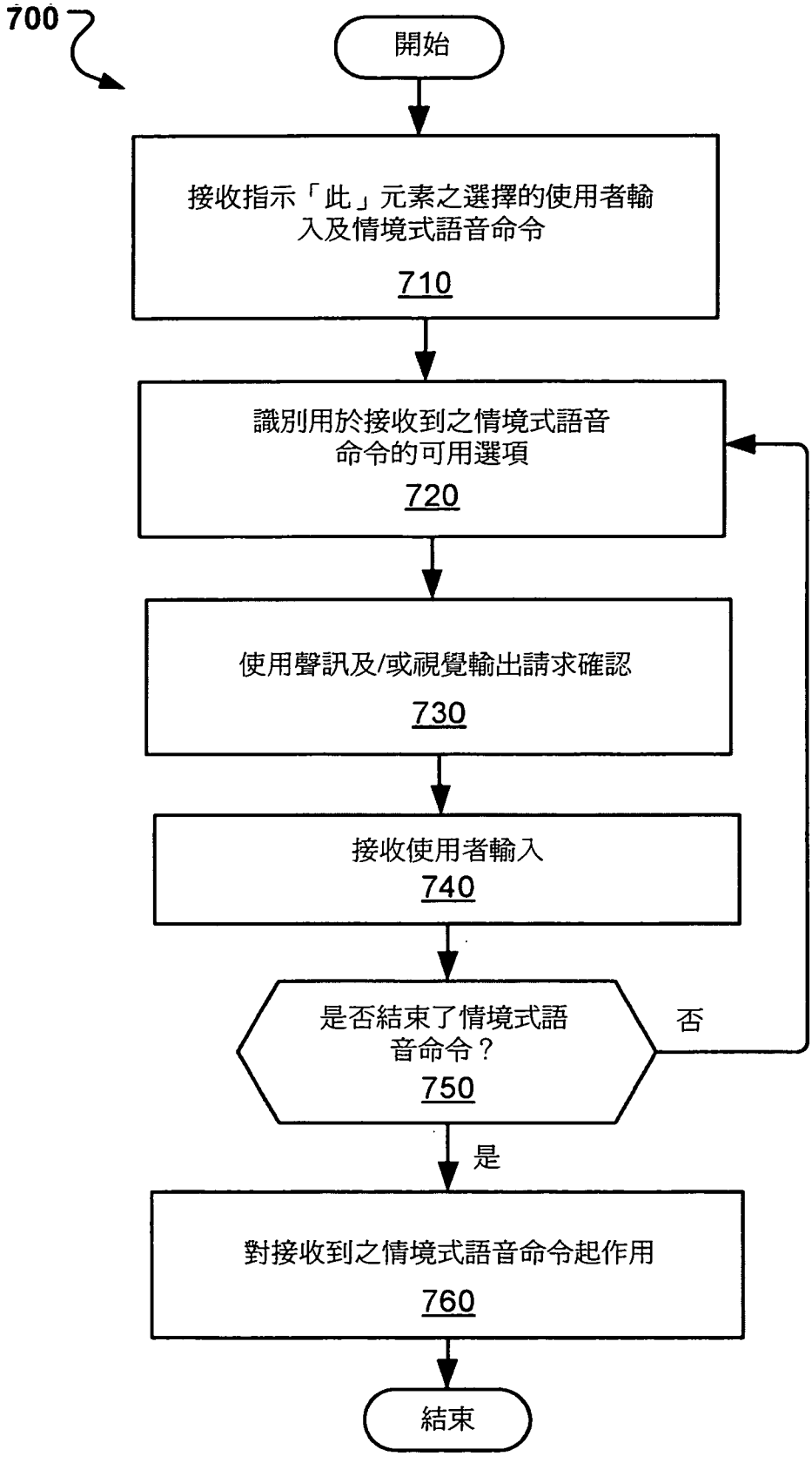

700
開始
接收指示「此」元素之選擇的使用者輸入及情境式語音命令
710
識別用於接收到之情境式語音命令的可用選項
720
使用聲訊及/或視覺輸出請求確認
730
接收使用者輸入
740
是否結束了情境式語音命令？
750
否
是
對接收到之情境式語音命令起作用
760
結束

圖7

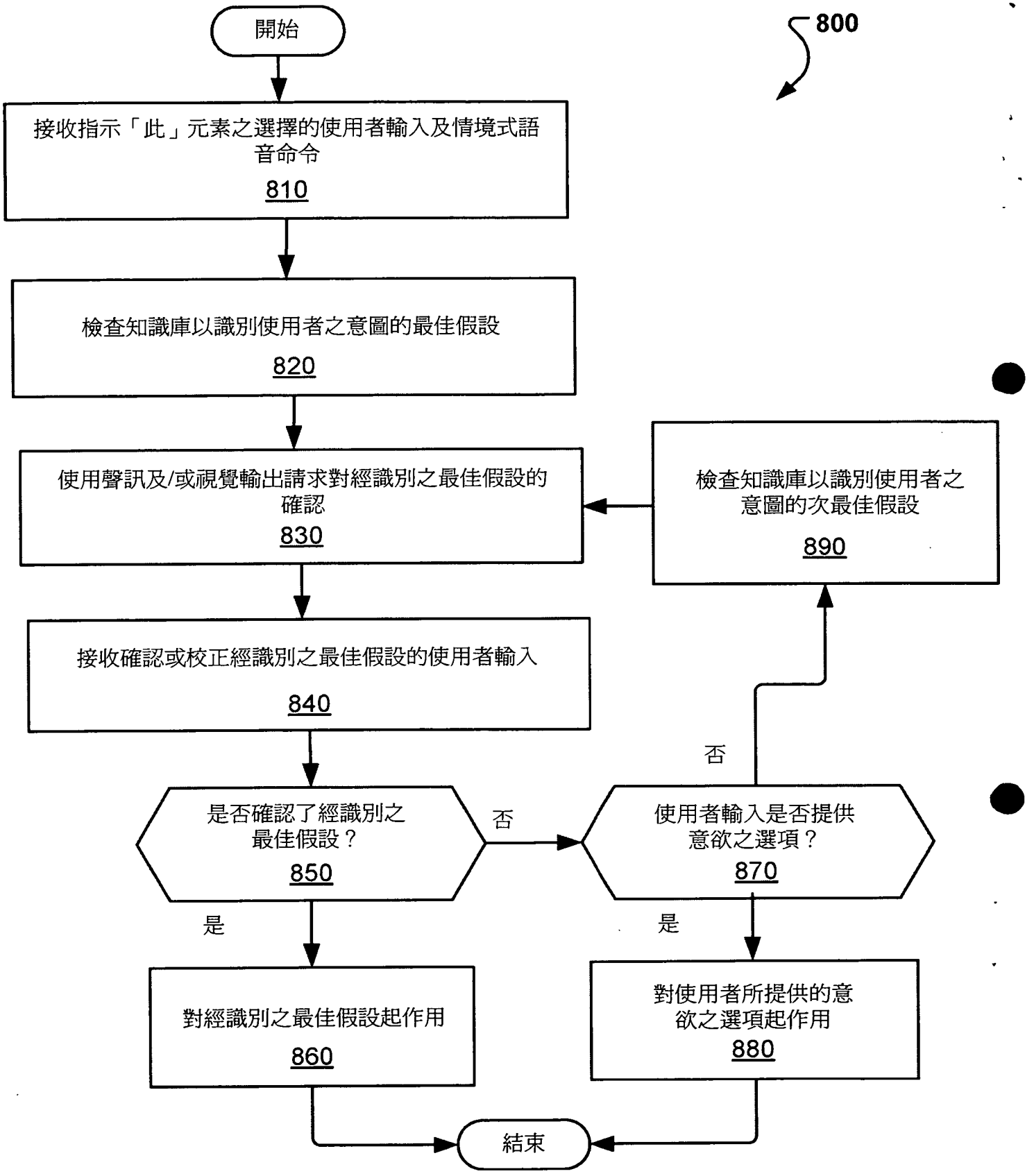

800
開始
接收指示「此」元素之選擇的使用者輸入及情境式語音命令
810
檢查知識庫以識別使用者之意圖的最佳假設
820
使用聲訊及/或視覺輸出請求對經識別之最佳假設的確認
830
檢查知識庫以識別使用者之意圖的次最佳假設
890
接收確認或校正經識別之最佳假設的使用者輸入
840
否
是否確認了經識別之最佳假設？
850
否
使用者輸入是否提供意欲之選項？
870
是
是
對經識別之最佳假設起作用
860
對使用者所提供的意欲之選項起作用
880
結束

圖8a

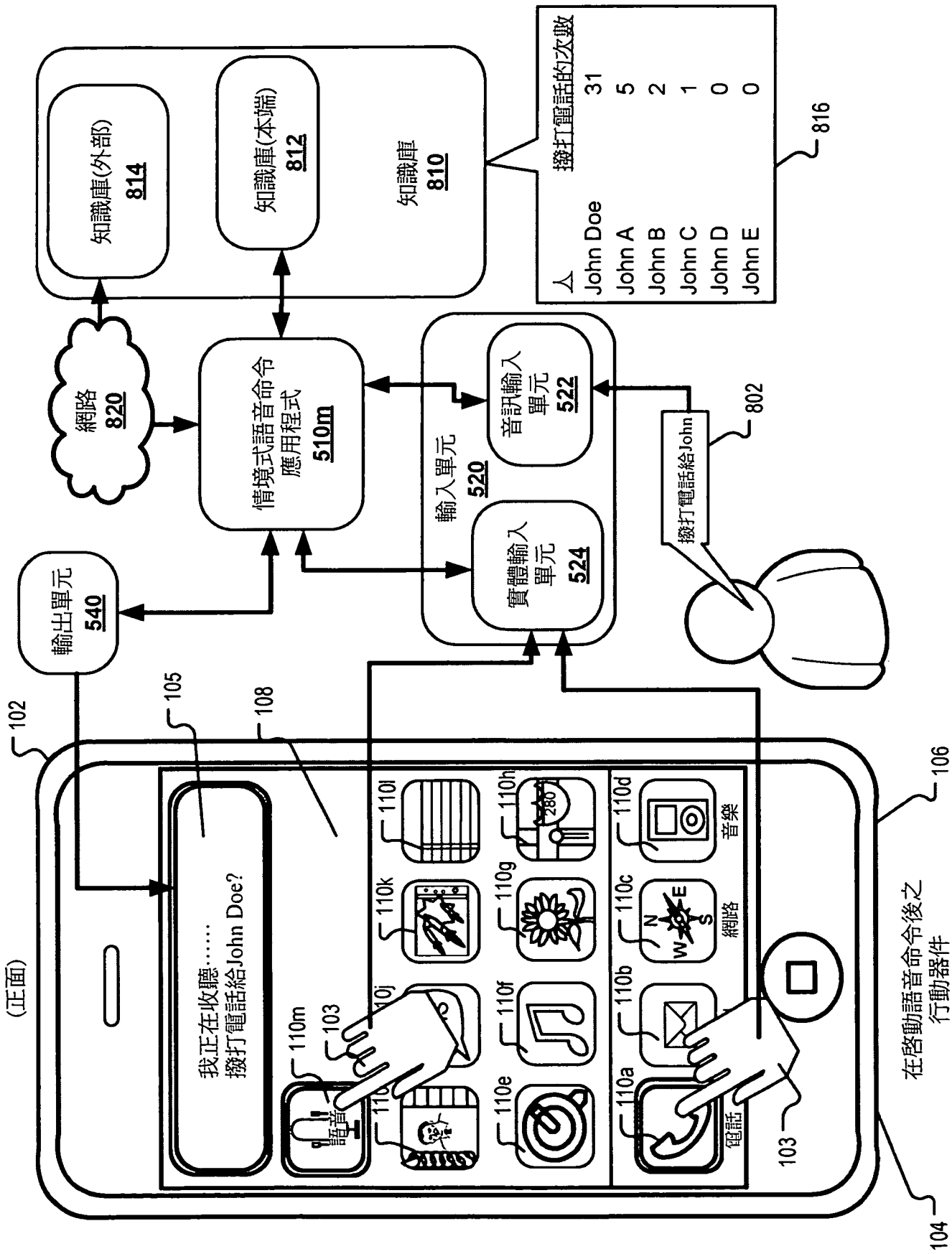

知識庫(外部)
814
知識庫(本端)
812
知識庫
810
撥打電話的次數
人
John Doe 31
John A 5
John B 2
John C 1
John D 0
John E 0
816
網路
820
情境式語音命令
應用程式
510m
輸入單元
520
音訊輸入
單元
522
實體輸入
單元
524
802
撥打電話給John
輸出單元
540
102
105
108
(正面)
我正在收聽……
撥打電話給John Doe?
110m
103
110l
110k
110j
110h
110g
110f
110e
110d
110c
110b
110a
音樂
網路
電話
106
104
在啓動語音命令後之
行動器件

圖8b

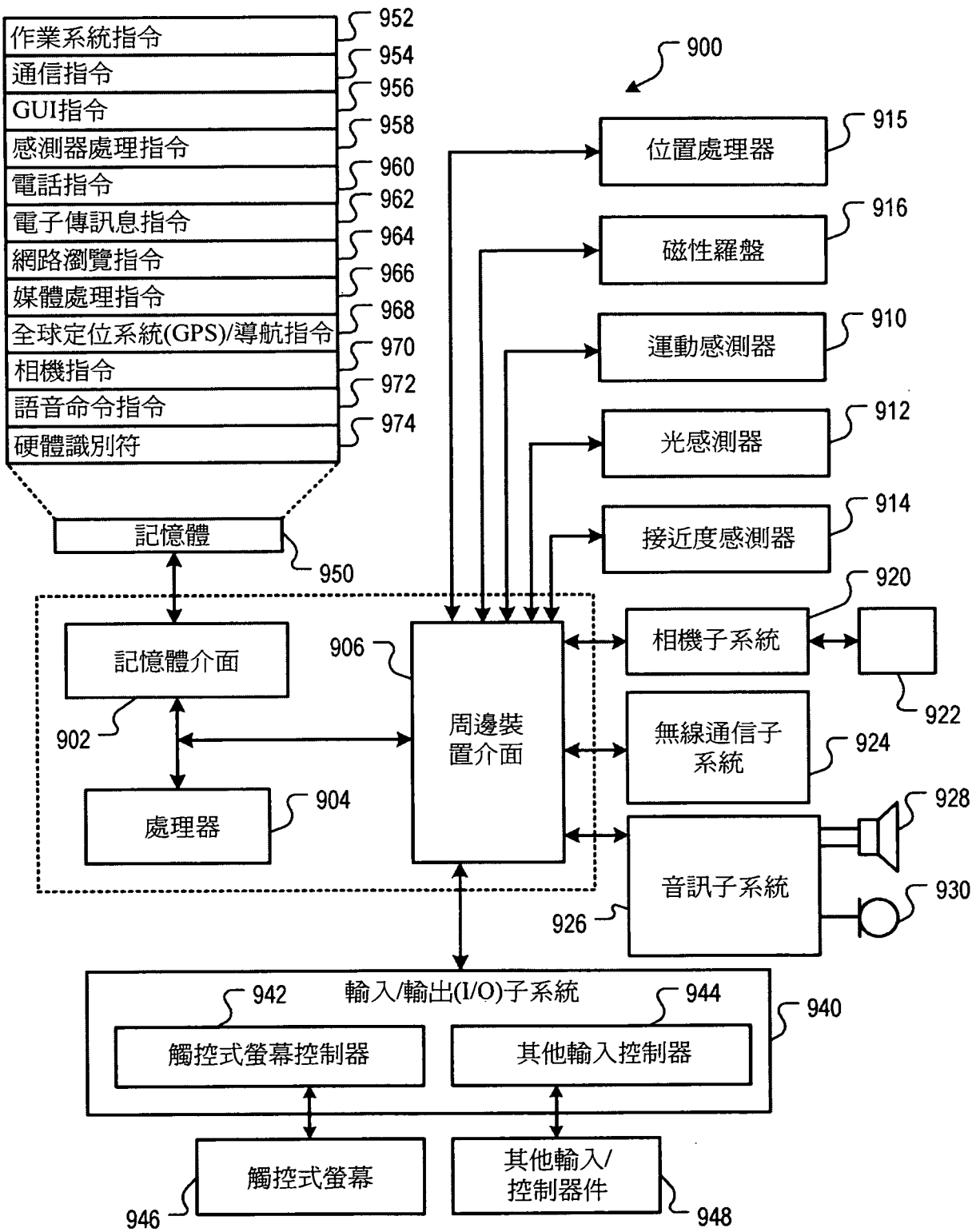

作業系統指令
通信指令
GUI指令
感測器處理指令
電話指令
電子傳訊息指令
網路瀏覽指令
媒體處理指令
全球定位系統(GPS)/導航指令
相機指令
語音命令指令
硬體識別符
952
954
956
958
960
962
964
966
968
970
972
974
900
位置處理器
915
磁性羅盤
916
運動感測器
910
光感測器
912
接近度感測器
914
記憶體
950
記憶體介面
902
906
周邊裝置介面
相機子系統
920
922
無線通信子系統
924
處理器
904
928
音訊子系統
930
926
輸入/輸出(I/O)子系統
942
944
940
觸控式螢幕控制器
其他輸入控制器
觸控式螢幕
其他輸入/控制器件
946
948

圖9