



(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 I571796 B

(45)公告日：中華民國 106 (2017) 年 02 月 21 日

(21)申請案號：101138925

(22)申請日：中華民國 101 (2012) 年 10 月 22 日

(51)Int. Cl. : G06F3/16 (2006.01)

G10L15/02 (2006.01)

G10L21/02 (2013.01)

(30)優先權：2011/11/17 美國

13/299,218

(71)申請人：微軟技術授權有限責任公司 (美國) MICROSOFT TECHNOLOGY LICENSING, LLC
(US)

美國

(72)發明人：莫迪拉格胡 MURTHI, RAGHU (US)；吉艾莫三世愛德華 C GIAIMO III, EDWARD
C. (US)

(74)代理人：蔡坤財；李世章

(56)參考文獻：

TW 200919255A

CN 101933051A

US 6070140

US 2005/0041360A1

US 2009/0112605A1

審查人員：陳泰龍

申請專利範圍項數：18 項 圖式數：7 共 52 頁

(54)名稱

用於設備啓動的音頻圖樣比對

AUDIO PATTERN MATCHING FOR DEVICE ACTIVATION

(57)摘要

本發明係關於用於設備啟動的音頻圖樣比對。揭示用於將電設備從待機功率模式啟動到全功率模式的系統及方法。該系統可以包括用於監測電設備附近的音訊訊號的一或多個麥克風，及包括低功率微處理器及非揮發性記憶體之待機功率啟動單元。由該一或多個麥克風擷取的音訊被微處理器數位化並與儲存在非揮發性記憶體中的一或多個預定義的啟動模式相比較。若偵測到數位元音頻圖樣與預定義的啟動模式之間的模式比對，則電設備被啟動。

A system and method are disclosed for activating an electric device from a standby power mode to a full power mode. The system may include one or more microphones for monitoring audio signals in the vicinity of the electric device, and a standby power activation unit including a low-power microprocessor and a non-volatile memory. Audio captured by the one or more microphones is digitized and compared by the microprocessor against predefined activation pattern(s) stored in the non-volatile memory. If a pattern match is detected between the digital audio pattern and a predefined activation pattern, the electric device is activated.

指定代表圖：

符號簡單說明：

101/259 . . . 中央處理單元

462 . . . 麥克風

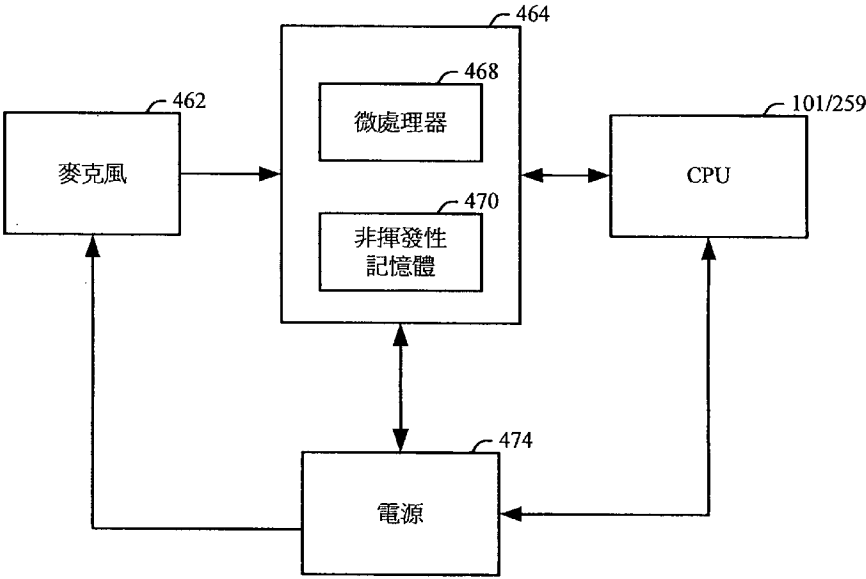
464 . . . 待機啟動單元

468 . . . 處理器

470 . . . 非揮發性記憶體

474 . . . 電源

圖7



發明摘要

※ 申請案號：101138925

※ 申請日：101 年 10 月 22 日

※IPC 分類：

G06F 3/16 (2006.01)

G10L 15/02 (2006.01)

G10L 21/02 (2013.01)

【發明名稱】（中文/英文）

用於設備啟動的音頻圖樣比對 /AUDIO PATTERN
MATCHING FOR DEVICE ACTIVATION

【中文】

本發明係關於用於設備啟動的音頻圖樣比對。揭示用於將電設備從待機功率模式啟動到全功率模式的系統及方法。該系統可以包括用於監測電設備附近的音訊訊號的一或多個麥克風，及包括低功率微處理器及非揮發性記憶體的待機功率啟動單元。由該一或多個麥克風擷取的音訊被微處理器數位化並與儲存在非揮發性記憶體中的一或多個預定義的啟動模式相比較。若偵測到數位元音頻圖樣與預定義的啟動模式之間的模式比對，則電設備被啟動。

【英文】

A system and method are disclosed for activating an electric device from a standby power mode to a full power mode. The system may include one or more microphones for monitoring audio signals in the vicinity of the electric device, and a standby power activation unit including a low-power microprocessor and a non-volatile memory.

Audio captured by the one or more microphones is digitized and compared by the microprocessor against predefined activation pattern(s) stored in the non-volatile memory. If a pattern match is detected between the digital audio pattern and a predefined activation pattern, the electric device is activated.

【代表圖】

【本案指定代表圖】：第（ 7 ）圖。

【本代表圖之符號簡單說明】：

101/259 中央處理單元

462 麥克風

464 待機啟動單元

468 處理器

470 非揮發性記憶體

474 電源

【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)

用於設備啟動的音頻圖樣比對 /AUDIO PATTERN
MATCHING FOR DEVICE ACTIVATION

【技術領域】

【0001】 本發明係關於用於設備啟動的音頻圖樣比對

【先前技術】

【0002】 空轉或關閉待機功率模式下的電設備是共用慣例，其中該設備消耗降低的電量。電設備的典型待機功率例如可以高達大約8瓦。然而，對於幾十億個使用中的電設備消耗而言，即使此小的量仍然合計得到世界總耗電的可觀的一小部分。諸如國際能源機構(IEA)提出的1瓦倡議(One Watt Initiative)之類的倡議正在嘗試到2013年為止將電設備的待機功率減少到0.5瓦。可能有利的是，在待機功率模式下，即使在以0.5瓦操作時仍然允許電設備的簡化的啟動。

【發明內容】

【0003】 本技術大體上係關於電設備中所包括的用於將該設備從待機功率模式啟動到全功率模式的系統。該系統包括一或多個麥克風及待機啟動單元。待機啟動單元包括諸如微處理器之類的處理器及相關聯的非揮發性記憶體。一或多個數位化啟動短語可以儲存在非揮發性記憶體中，該一或

多個數位化啟動短語要麼在使用者首次使用之前被預先記錄，要麼被使用者選擇及輸入。

【0004】 當電設備在待機功率模式下執行時，該一或多個麥克風被供電以感測麥克風附近的聲音。爲了啟動處於待機功率模式時的電設備，使用者可以講出與非揮發性記憶體中儲存的短語比對的預先記錄或使用者選擇的啟動短語。音訊串流被一或多個麥克風偵測，並且數位元化成音頻圖樣。隨後，音頻圖樣由待機啟動單元的處理器與儲存在非揮發性記憶體中的一或多個啟動短語相比較。若所講出的短語比對所儲存的啟動短語，則處理器將識別出比對，並且向功率電路發送訊號以將該電設備啟動到全功率模式。當處理器在所接收的音頻圖樣與一或多個所儲存的啟動短語之間未識別出比對時，電設備保持在待機功率模式。

【0005】 麥克風及待機啟動單元能夠接收音訊串流，將音訊串流數位元化成音頻圖樣並且執行比對操作以將傳入音頻圖樣與一或多個所儲存的啟動短語相比較，該等全部是利用在待機模式下可用的功率進行的。在一個實例中，該功率可以是0.5瓦。

【0006】 在實施例中，本技術係關於一種用於將電設備從待機功率模式啟動的方法，包括：（a）在處於待機功率模式時在該電設備中接收音訊串流；（b）在處於待機功率模式時將該音訊串流數位元化成音頻圖樣；（c）在處於待機功率模式時將在該步驟（b）中數位元化的音頻圖樣與儲存在非揮發性記憶體中的數位化啟動短語相比較；及（d）在該音頻

圖樣在預定義的公差內與該啟動短語比對時啟動該電設備。

【0007】 在另一實例中，本技術係關於一種用於將電設備從待機功率模式啟動的待機啟動系統，該待機啟動系統包括：一或多個麥克風，該麥克風用於偵測該一或多個麥克風附近的音訊串流；及待機啟動單元，包括：非揮發性記憶體，該非揮發性記憶體包括一或多個所儲存的啟動模式；及處理器，該處理器用於使用在待機功率模式下對該電設備可用的功率將從該一或多個麥克風接收的數位元化音頻圖樣與儲存在該非揮發性記憶體中的一或多個啟動模式相比較，該處理器在數位元化音頻圖樣在預定義的公差內與儲存在該非揮發性記憶體中的該一或多個啟動短語中的啟動短語比對時啟動該設備。

【0008】 在另一實例中，本技術係關於一種具有電腦可執行指令的電腦可讀取儲存媒體，該等指令用於將處理器程式設計為執行一種用於將自然使用者介面（NUI）系統的計算設備從待機功率模式啟動的方法，包括：（a）在處於待機功率模式時在與NUI系統相關聯的一或多個麥克風中接收音訊串流；（b）在處於待機功率模式時將該音訊串流數位元化成音頻圖樣；（c）在處於待機功率模式時處理數位元化音訊串流以促進數位元化音訊串流與儲存在計算設備的非揮發性記憶體中的一或多個數位元化啟動短語的比對；（d）在處於待機功率模式時將在該步驟（c）處理的音頻圖樣與儲存在計算設備的非揮發性記憶體中的一或多個數位元化啟動短語相比較；及（e）當音頻圖樣在預定義的公差內與該一或多個啟

動短語中的啟動短語比對時啟動該計算設備。

【0009】 提供本發明內容以便以簡化形式介紹將在以下具體實施方式中進一步描述的一些概念。本發明內容並非意欲識別所主張標的的關鍵特徵或必要特徵，亦不意欲用於幫助決定所主張標的的範圍。此外，所主張標的的不限於解決在本案的任一部分中所提及的任何或所有缺點的實現。

【圖式簡單說明】

【0010】 圖1A至圖1B圖示使用者正玩遊戲的目標辨識、分析及追蹤系統的示例性實施例。

【0011】 圖2圖示可在目標辨識、分析及追蹤系統中使用的擷取設備的示例性實施例。

【0012】 圖3A圖示可用於在目標辨識、分析及追蹤系統中解釋一或多個姿勢的計算環境的示例性實施例。

【0013】 圖3B圖示可用於在目標辨識、分析及追蹤系統中解釋一或多個姿勢的計算環境的另一示例性實施例。

【0014】 圖4是用於教導及儲存啟動短語的實施例的流程圖。

【0015】 圖5是取樣數位元化啟動模式。

【0016】 圖6是用於在處於待機功率模式時啟動電設備的實施例的流程圖。

【0017】 圖7是用於在處於待機功率模式時啟動電設備的實施例的方塊圖。

【實施方式】

【0018】 現在將參考附圖1A至圖7來描述本技術的實

施例，該等實施例大體上係關於電設備中所包括的一種用於將該設備從待機功率模式啟動到全功率模式的系統。在一個實例中，電設備可以是下文被描述為目標辨識、分析及追蹤系統10的NUI系統。NUI系統可以是一種系統，在該系統中偵測、解釋使用者姿勢並將使用者姿勢用於控制螢幕上人物或軟體應用程式的其他態樣。然而，如下文進一步描述的一般，電設備可以是多種其他計算設備及儀器。在實施例中，本案的系統包括用於監測電設備附近的音訊訊號的一或多個麥克風、及包括低功率微處理器及非揮發性記憶體의待機功率啟動單元。當該設備處於待機功率模式時，麥克風擷取音訊，並且將該音訊轉換成數位元音頻圖樣。隨後由微處理器將該音頻圖樣與儲存在非揮發性記憶體中的一或多個預定義的啟動模式相比較。若偵測到數位元音頻圖樣與預定義的啟動模式之間的模式比對，則電設備被啟動。

【0019】 該一或多個麥克風、微處理器及非揮發性記憶體一起能夠獲得音頻圖樣並且使用少量功率（例如0.5瓦）將該等音頻圖樣與所儲存的啟動模式相比較。因此，本系統可以用於在電設備處於待機功率模式時（亦即要麼為閒置、要麼為關閉時）啟動該電設備。當處於待機功率模式時，下文描述的電設備接收待機功率，該待機功率例如在實例中可以為0.5瓦。能夠理解，待機功率可以在另外的實施例中高於或低於此，並且本系統將能夠以此種更高或更低的待機功率設置來操作。

【0020】 最初參考圖1A至圖2，用於實現本技術的一個

實例的硬體包括目標辨識、分析及追蹤系統10，該系統可用於辨識、分析及/或追蹤諸如使用者18之類的人類目標。目標辨識、分析及追蹤系統10的實施例包括用於執行遊戲或其他應用程式的計算設備12。計算設備12可包括硬體元件及/或軟體元件，使得計算設備12可用於執行諸如遊戲之類的應用程式及非遊戲應用程式。在一個實施例中，計算設備12可以包括可執行儲存在處理器可讀取儲存設備上的指令的處理器，諸如標準化處理器、專用處理器、微處理器等等，該等指令用於在設備12為活動並以全功率執行時執行該設備12的程序。

【0021】 如後文所述，計算設備12亦可以包括可執行儲存在非揮發性記憶體中的指令的第二低功率處理器或微處理器，該等指令用於在設備12執行在待機模式時執行該設備12的啟動程序。

【0022】 系統10亦包括擷取設備20，擷取設備20用於擷取與擷取設備所感測的一或多個使用者及/或物件有關的圖像及音訊資料。在實施例中，擷取設備20可用於擷取與一或多個使用者的身體及手移動及/或姿勢及語音相關的資訊，該資訊由計算環境接收並且用於呈現遊戲或其他應用程式的各態樣、與該等態樣互動及/或控制該等態樣。下文更詳細地解釋計算設備12及擷取設備20的實例。

【0023】 目標辨識、分析及追蹤系統10的各個實施例可連接到具有顯示器14的音訊/視覺（A/V）設備16。設備16可以例如是可向使用者提供遊戲或應用程式視覺及/或音訊的

電視機、監視器、高清電視機（HDTV）等。例如，計算設備12可包括諸如圖形卡之類的視訊配接器及/或諸如音效卡之類的音訊配接器，該等配接器可提供與遊戲或其他應用程式相關聯的音訊/視覺訊號。A/V設備16可從計算設備12接收音訊/視覺訊號，並且隨後可向使用者18輸出與該等音訊/視覺訊號相關聯的遊戲或應用程式視覺及/或音訊。根據一個實施例，音訊/視覺設備16可經由例如S-視訊電纜、同軸電纜、HDMI電纜、DVI電纜、VGA電纜、分量視訊電纜等連接到計算設備12。

【0024】 在實施例中，計算設備12、A/V設備16及擷取設備20可協調以在顯示器14上呈現化身或螢幕上人物19。例如，圖1A圖示使用者18正在玩足球遊戲應用程式。追蹤並使用使用者的移動來使化身19的移動動畫化。在各個實施例中，化身19模仿使用者18在現實世界空間中的移動，使得使用者18可執行控制化身19在顯示器14上的移動及動作的移動及姿勢。在圖1B中，在NUI系統中使用擷取設備20，其中例如，使用者18正在滾動及控制具有呈現在顯示器14上的各種功能表選項的使用者介面21。在圖1B中，計算設備12及擷取設備20可用於辨識及分析使用者身體的移動及姿勢，並且此類移動及姿勢可被解釋為對使用者介面的控制。

【0025】 系統10及系統10之元件的合適的實例在以下共同待審的專利申請案中找到，該等專利申請案全部特此經由引用結合於此：於2009年5月29日提出申請的名稱為「Environment And/Or Target Segmentation（環境及/或目標分割

)」的美國專利申請案第12/475,094號；於2009年7月29日提出申請的名稱為「Auto Generating a Visual Representation (自動產生視覺表示)」的美國專利申請案第12/511,850號；於2009年5月29日提出申請的名稱為「Gesture Tool (姿勢工具)」的美國專利申請案第12/474,655號；於2009年10月21日提出申請的名稱為「Pose Tracking Pipeline (姿態追蹤排序緩衝)」的美國專利申請案第12/603,437號；於2009年5月29日提出申請的名稱為「Device for Identifying and Tracking Multiple Humans Over Time (用於隨時間識別及追蹤多個人類的設備)」的美國專利申請案第12/475,308號；於2009年10月7日提出申請的名稱為「Human Tracking System (人類追蹤系統)」的美國專利申請案第12/575,388號；於2009年4月13日提出申請的名稱為「Gesture Recognizer System Architecture (姿勢辨識器系統架構)」的美國專利申請案第12/422,661號；及於2009年2月23日提出申請的名稱為「Standard Gestures (標準姿勢)」的美國專利申請案第12/391,150號。

【0026】 圖2圖示可在目標辨識、分析及追蹤系統10中使用的擷取設備20的示例性實施例。在示例性實施例中，擷取設備20可被配置成經由任何合適的技術來擷取具有可包括深度值的深度圖像的視訊，該等技術包括例如飛行時間、結構化光、立體圖像等。根據一個實施例，擷取設備20可以將所計算的深度資訊組織成「Z層」，即與從深度相機沿深度相機視線延伸的Z軸相垂直的層。X及Y軸可被界定為與Z軸垂直。Y軸可以是垂直的，而X軸可以是水平的。X、Y及Z軸一起

界定擷取設備20所擷取的3-D真實世界空間。

【0027】 如圖2所示，擷取設備20可包括圖像相機元件22。根據示例性實施例，圖像相機元件22可以是可擷取場景的深度圖像的深度相機。深度圖像可以包括所擷取場景的二維（2-D）像素區域，其中2-D像素區域中的每個像素皆可表示深度值，諸如例如所擷取場景中的物件距相機的以釐米、毫米等為單位的長度或距離。

【0028】 如圖2所示，根據示例性實施例，圖像相機元件22可包括可用於擷取場景的深度圖像的IR光元件24、三維（3-D）相機26，及RGB相機28。例如，在飛行時間分析中，擷取設備20的IR光元件24可將紅外光發射到場景上，並且隨後可使用感測器（未圖示）、用例如3-D相機26及/或RGB相機28來偵測從場景中的一或多個目標及物體的表面反向散射的光。

【0029】 在一些實施例中，可以使用脈衝紅外光，使得可以量測出射光脈衝與相應入射光脈衝之間的時間，並且將該時間用於決定從擷取設備20到場景中的目標或物體上的特定位置的實體距離。附加地，在其他示例性實施例中，可以將出射光波的相位與入射光波的相位相比較來決定相移。該相移隨後可以用於決定從擷取設備20到目標或物體上的特定位置的實體距離。

【0030】 根據另一示例性實施例，可以使用飛行時間分析來藉由經由包括例如快門式光脈衝成像在內的各種技術分析反射光束隨時間的強度來間接地決定從擷取設備20到目

標或物體上的特定位置的實體距離。

【0031】 在另一示例性實施例中，擷取設備20可以使用結構化光來擷取深度資訊。在此種分析中，圖案化光（亦即被顯示成諸如網格圖案或條紋圖案的已知圖案的光）可以經由例如IR光元件24被投射到場景上。在落到場景中的一或多個目標或物體的表面上以後，作為回應，圖案可以變為變形的。圖案的此種變形可以被例如3-D相機26及/或RGB相機28擷取，並且隨後可被分析以決定從擷取設備20到目標或物體上的特定位置的實體距離。

【0032】 根據另一實施例，擷取設備20可包括可以從不同的角度觀察場景的兩個或兩個以上在實體上分開的相機，以獲取可以被解析以產生深度資訊的視覺立體資料。在另一示例性實施例中，擷取設備20可使用點雲資料（point cloud data）及目標數位化技術來偵測使用者的特徵。

【0033】 擷取設備20亦可包括麥克風30。麥克風30可包括可接收聲音並將聲音轉換成電訊號的變換器或感測器。根據一個實施例，麥克風30可用來減少在目標辨識、分析及追蹤系統10中的擷取設備20與計算設備12之間的回饋。附加地，麥克風30可用來接收亦可由使用者提供的音訊訊號，以控制可由計算設備12執行的諸如遊戲應用程式、非遊戲應用程式等應用程式。替代於單個麥克風30，本系統亦可以採用兩個或兩個以上麥克風。多個麥克風允許聲源定位以識別出聲音的來源。

【0034】 在示例性實施例中，擷取設備20亦可包括能

與圖像相機元件22有效通訊的處理器32。處理器32可包括可執行指令的標準處理器、專用處理器、微處理器等，該等指令可包括用於接收深度圖像的指令、用於決定合適的目標是否可包括在深度圖像中的指令、用於將合適的目標轉換成該目標的骨架表示或模型的指令，或任何其他合適的指令。

【0035】 擷取設備20亦可包括記憶體元件34，記憶體元件34可儲存可由處理器32執行的指令、3-D相機或RGB相機擷取到的圖像或圖像的訊框，或任何其他合適的資訊、圖像等。根據示例性實施例，記憶體元件34可包括隨機存取記憶體（RAM）、唯讀記憶體（ROM）、快取記憶體、快閃記憶體、硬碟，或任何其他合適的儲存元件。如圖2所示，在一個實施例中，記憶體元件34可以是與圖像相機元件22及處理器32通訊的分開元件。根據另一個實施例，可將記憶體元件34整合到處理器32及/或圖像相機元件22中。

【0036】 如圖2所示，擷取設備20可以經由通訊鏈路36與計算設備12通訊。通訊鏈路36可以是包括例如USB連接、火線連接、乙太網路電纜連接等有線連接及/或諸如無線802.11b、802.11g、802.11a或802.11n連接等無線連接。根據一個實施例，計算設備12可以經由通訊鏈路36向擷取設備20提供可用於決定例如何時擷取場景的時鐘。

【0037】 另外，擷取設備20可提供由例如3-D相機26及/或RGB相機28所擷取的深度資訊及圖像。有了該等設備的說明，可根據本發明的技術開發部分骨架模型，其中經由通訊鏈路36將所得到的資料提供給計算設備12。

【0038】 計算設備12亦可包括用於辨識姿勢的姿勢辨識引擎190。根據本系統，計算設備12亦可以包括模式比對引擎192，並且在實施例中包括語音辨識引擎194，此兩個引擎皆在下文予以說明。

【0039】 圖3A圖示可用於在目標辨識、分析及追蹤系統中解釋一或多個姿勢的計算環境的示例性實施例。上文參考圖1A至圖2所描述的諸如計算設備12等計算環境可以是諸如遊戲控制台等多媒體控制台100。如圖3A所示，多媒體控制台100包括具有一級快取記憶體102、二級快取記憶體104及快閃記憶體ROM 106的中央處理單元（CPU）101。一級快取記憶體102及二級快取記憶體104臨時儲存資料，並且因此減少記憶體存取週期的數量，由此改進處理速度及輸送量。CPU 101可被提供為具有一個以上的核，並且由此具有附加的一級快取記憶體102及二級快取記憶體104。快閃記憶體ROM 106可儲存在多媒體控制台100通電時引導程序的初始階段期間載入的可執行代碼。

【0040】 儘管圖3A中未圖示，但是多媒體控制台100亦可以包括諸如微處理器之類的處理器及非揮發性記憶體以用於將該系統從待機功率模式中啟動。在另外的實施例中，用於將該系統從待機功率模式中啟動的處理器及記憶體分別可以是處理單元101及ROM 106。

【0041】 圖形處理單元（GPU）108及視訊轉碼器/視訊轉碼器（編碼器/解碼器）114形成用於高速及高解析度圖形處理的視訊處理排序緩衝。經由匯流排從GPU 108向視訊轉碼器

/視訊轉碼器114運送資料。視訊處理排序緩衝向A/V（音訊/視訊）埠140輸出資料，用於傳輸至電視機或其他顯示器。記憶體控制器110連接到GPU 108，以便於處理器對各種類型的記憶體112（諸如，但不限於RAM）的存取。

【0042】 多媒體控制台100包括較佳地在模組118上實現的I/O控制器120、系統管理控制器122、音訊處理單元123、網路介面控制器124、第一USB主控制器126、第二USB主控制器128及前面板I/O子部件130。USB控制器126及128用作周邊控制器142(1)-142(2)、無線配接器148及外置記憶體設備146（例如，快閃記憶體、外置CD/DVD ROM驅動器、可移除媒體等）的主機。網路介面124及/或無線配接器148提供對網路（例如，網際網路、家用網路等）的存取，並且可以是包括乙太網路卡、數據機、藍芽模組、電纜數據機等的各種不同的有線或無線配接器元件中的任何一種。

【0043】 提供系統記憶體143來儲存在引導程序期間載入的應用程式資料。提供媒體驅動器144，且媒體驅動器144可包括DVD/CD驅動器、硬碟，或其他可移除媒體驅動器等。媒體驅動器144可以是對多媒體控制台100內置的或外置的。應用程式資料可經由媒體驅動器144存取，以供多媒體控制台100執行、重播等。媒體驅動器144經由諸如串列ATA匯流排或其他高速連接（例如IEEE 1394）等匯流排連接到I/O控制器120。

【0044】 系統管理控制器122提供與確保多媒體控制台100的可用性相關的各種服務功能。音訊處理單元123及音

訊轉碼器132形成具有高保真度及立體處理的相應音訊處理排序緩衝。音訊資料經由通訊鏈路在音訊處理單元123與音訊轉碼器132之間傳輸。音訊處理排序緩衝將資料輸出到A/V埠140，以供外置音訊播放機或具有音訊能力的設備再現。

【0045】 前面板I/O子部件130支援暴露在多媒體控制台100的外表面上的電源按鈕150及彈出按鈕152及任何LED（發光二極體）或其他指示器的功能。系統供電模組136向多媒體控制台100的元件供電。風扇138冷卻多媒體控制台100內的電路系統。

【0046】 多媒體控制台100內的CPU 101、GPU 108、記憶體控制器110及各種其他元件經由一或多條匯流排互連，該一或多條匯流排包括串列及平行匯流排、記憶體匯流排、周邊匯流排及使用各種匯流排體系結構中的任一種的處理器或區域匯流排。作為示例，該等架構可以包括周邊部件互連（PCI）匯流排、PCI-Express匯流排等。

【0047】 當多媒體控制台100通電時，應用程式資料可從系統記憶體143載入到記憶體112及/或快取記憶體102、104中，並且可在CPU 101上執行。應用程式可在導覽到多媒體控制台100上可用的不同媒體類型時呈現提供一致的使用者體驗的圖形化使用者介面。在操作中，媒體驅動器144中所包含的應用程式及/或其他媒體可從媒體驅動器144啟動或播放，以將附加功能提供給多媒體控制台100。

【0048】 多媒體控制台100可藉由簡單地將該系統連接到電視機或其他顯示器而作為獨立系統來操作。在該獨立

模式中，多媒體控制台100允許一或多個使用者與該系統互動、看電影，或聽音樂。然而，在經由網路介面124或無線配接器148可用的寬頻連接整合的情況下，多媒體控制台100亦可作為更大網路細胞服務區中的參與者來操作。

【0049】 當多媒體控制台100通電時，可以保留設定量的硬體資源以供多媒體控制台作業系統作系統使用。該等資源可包括記憶體保留量（諸如，16MB）、CPU及GPU週期的保留量（諸如，5%）、網路連結頻寬的保留量（諸如，8kbs），等等。因為該等資源是在系統引導時間保留的，所保留的資源從應用程式的視角而言是不存在的。

【0050】 具體而言，記憶體保留量較佳地足夠大，以包含啟動核心、併發系統應用程式及驅動程式。CPU保留量較佳地為恆定，使得若所保留的CPU用量不被系統應用程式使用，則閒置執行緒將消耗任何未使用的週期。

【0051】 對於GPU保留量，藉由使用GPU中斷來排程代碼以將快顯視窗渲染為覆蓋圖，從而顯示由系統應用程式產生的輕量訊息（例如，快顯視窗）。覆蓋圖所需的記憶體量取決於覆蓋區域大小，並且覆蓋圖較佳地與螢幕解析度成比例縮放。在完整的使用者介面被併發系統應用程式使用的情況下，較佳使用獨立於應用程式解析度的解析度。定標器可用於設置該解析度，從而無需改變頻率及引起TV重新同步。

【0052】 在多媒體控制台100引導且系統資源被保留之後，執行併發系統應用程式來提供系統功能。系統功能被封裝在上述所保留的系統資源內執行的一組系統應用程式中

。作業系統核心識別出作為系統應用程式執行緒而非遊戲應用程式執行緒的執行緒。系統應用較佳地被排程為在預定時間並以預定時間間隔在CPU 101上執行，以便提供對應用程式而言一致的系統資源視圖。排程是為了使在控制台上執行的遊戲應用程式的快取記憶體中斷最小化。

【0053】 當併發系統應用程式需要音訊時，由於時間敏感性而將音訊處理非同步地排程給遊戲應用程式。多媒體控制台應用程式管理器（如下所描述的）在系統應用程式活動時控制遊戲應用程式的音訊級別（例如，靜音、衰減）。

【0054】 輸入設備（例如，控制器142(1)及142(2)）由遊戲應用程式及系統應用程式共享。輸入設備不是保留資源，而是在系統應用程式及遊戲應用程式之間切換以使輸入設備各自具有設備的焦點。應用管理器較佳控制輸入串流的切換，而無需知道遊戲應用程式的知識，並且驅動器維護關於焦點切換的狀態資訊。相機26、28及擷取設備20可為控制台100界定附加輸入設備。

【0055】 圖3B圖示計算環境220的另一示例性實施例，該計算環境220可以用來解釋目標辨識、分析，及追蹤系統中的一或多個姿勢的圖1A至圖2所圖示的計算設備12。計算系統環境220只是合適的計算環境的一個實例，並且不意欲對當前揭示的標的的使用範圍或功能提出任何限制。亦不應該將計算環境220解釋為對示例性操作環境220中圖示的任一元件或元件組合有任何依賴性或要求。在一些實施例中，各個所示的計算元素可包括被配置成例示本案的特定態樣的電路

系統。例如，本案中使用的術語「電路系統」可包括被配置成藉由韌體或開關來執行一或多個功能的專用硬體元件。在其他示例性實施例中，術語「電路系統」可包括藉由體現可操作以執行一或多個功能的邏輯的軟體指令來配置的通用處理單元、記憶體等。在電路系統包括硬體及軟體的組合的示例性實施例中，實施者可以編寫體現邏輯的原始程式碼，且原始程式碼可以被編譯為可以由通用處理單元處理的機器可讀代碼。因為本領域技藝人士可以明白現有技術已經進化到硬體、軟體或硬體/軟體組合之間幾乎沒有差別的地步，因而選擇硬體或是軟體來實現具體功能是留給實現者的設計選擇。更具體地，本領域技藝人士可以明白軟體程序可被變換成等效的硬體結構，而硬體結構本身可被變換成等效的軟體程序。由此，對於硬體實現亦是軟體實現的選擇是設計選擇之一並留給實現者。

【0056】 在圖3B中，計算環境220包括通常包括各種電腦可讀取媒體的電腦241。電腦可讀取媒體可以是能被電腦241存取的任何可用媒體，而且包含揮發性及非揮發性媒體、可移除及不可移除媒體。系統記憶體222包括諸如ROM 223及RAM 260之類的揮發性及/或非揮發性記憶體形式的電腦儲存媒體。包含諸如在啓動期間有助於在電腦241內的元件之間傳輸資訊的基本常式的基本輸入/輸出系統224（BIOS）通常儲存在ROM 223中。RAM 260通常包含中央處理單元259可立即存取及/或當前正在操作的資料及/或程式模組。作為示例而非限制，圖3B圖示作業系統225、應用程式226、其他程式模組

227，及程式資料228。

【0057】 該計算環境亦可以包括用於將該系統從待機功率模式中啟動的處理器468（該處理器468在實施例中可以是微處理器）及非揮發性記憶體470。記憶體470可以是多種非揮發性記憶體中的任何記憶體，例如包括ROM、PROM、EPROM、EEPROM及快閃記憶體。在實施例中，用於將該系統從待機功率模式中啟動的處理器468及記憶體470可以分別作為一或多個中央處理單元259及ROM 223的一部分被整合。在另外的實施例中，處理器468及記憶體470可以一起整合在所謂的片上系統中。

【0058】 電腦241亦可以包括其他可移除/不可移除、揮發性/非揮發性電腦儲存媒體。只作為示例，圖3B圖示讀寫不可移除的非揮發性磁性媒體的硬碟238、讀寫可移除的非揮發性磁碟254的磁碟機239及讀寫諸如CD ROM或其他光學媒體之類的可移除的非揮發性光碟253的光碟機240。可在示例性操作環境中使用的其他可移除/不可移除、揮發性/非揮發性電腦儲存媒體包括但不限於，磁帶盒、快閃記憶卡、數位多功能光碟、數位錄影帶、固態RAM、固態ROM等。硬碟238通常經由諸如介面234之類的不可移除記憶體介面連接到系統匯流排221，並且磁碟機239及光碟機240通常經由諸如介面235之類的可移除記憶體介面連接到系統匯流排221。

【0059】 上文所論述的並且在圖3B中所圖示的驅動器及驅動器的相關聯的電腦儲存媒體，為電腦241提供了電腦可讀取的指令、資料結構、程式模組及其他資料的儲存。例如

，在圖3B中，硬碟238被圖示為儲存了作業系統258、應用程式257，其他程式模組256，及程式資料255。注意，該等元件可與作業系統225、應用程式226、其他程式模組227及程式資料228相同，亦可與作業系統225、應用程式226、其他程式模組227及程式資料228不同。在此作業系統258、應用程式257、其他程式模組256及程式資料255被給予了不同的編號，以說明至少作業系統258、應用程式257、其他程式模組256及程式資料255是不同的副本。使用者可經由諸如鍵盤251及定點設備252（通常稱為滑鼠、軌跡球或觸摸墊）之類的輸入設備將命令及資訊輸入電腦241。其他輸入設備（未圖示）可包括麥克風、操縱桿、遊戲手柄、圓盤式衛星天線、掃描器等。該等及其他輸入設備通常經由耦合到系統匯流排的使用者輸入介面236連接到中央處理單元259及微處理器468，但亦可藉由諸如平行埠、遊戲埠或通用序列匯流排（USB）之類的其他介面及匯流排結構來連接。相機26、28及擷取設備20可為控制台100界定附加輸入設備。監視器242或其他類型的顯示設備亦經由諸如視訊介面232之類的介面連接至系統匯流排221。除了監視器以外，電腦亦可包括諸如揚聲器244及印表機243之類的其他周邊輸出設備，該等設備可經由輸出周邊介面233來連接。

【0060】 電腦241可使用到一或多個遠端電腦（諸如，遠端電腦246）的邏輯連接而在網路連結環境中操作。遠端電腦246可以是個人電腦、伺服器、路由器、網路PC、設備或其他共用網路節點，並且通常包括許多或所有以上相對電腦241

所描述的元件，但在圖3B中僅圖示記憶體儲存設備247。圖3B中所圖示的邏輯連接包括區域網路(LAN)245及廣域網(WAN)249，但是亦可以包括其他網路。此類網路連結環境在辦公室、企業範圍的電腦網路、網內網路及網際網路中是常見的。

【0061】 當在LAN網路連結環境中使用時，電腦241經由網路介面或配接器237連接到LAN 245。當在WAN網路連結環境中使用時，電腦241通常包括數據機250或用於經由諸如網際網路等WAN 249建立通訊的其他構件。數據機250可以是內置的或外置的，可經由使用者輸入介面236或其他適當的機制連接到系統匯流排221。在網路連結環境中，相對於電腦241所示的程式模組或程式模組的部分可被儲存在遠端記憶體儲存設備中。作為示例而非限制，圖3B圖示常駐在記憶體設備247上的遠端應用程式248。應當理解，所示的網路連接是示例性的，並且可使用在電腦之間建立通訊鏈路的其他構件。

【0062】 計算設備12結合擷取設備20可以每訊框皆產生使用者身體位置的電腦模型。例如在2010年9月7日提出申請的名稱為「System For Fast, Probabilistic Skeletal Tracking (用於快速概率骨架追蹤的系統)」的美國專利申請案第12/876,418號中揭示產生擷取設備20的視野中的一或多個使用者的骨架模型的此種排序緩衝的一個實例，該申請案的全部內容經由引用結合於此。

【0063】 骨架模型隨後可被提供給計算設備12，使得計算環境可追蹤骨架模型並呈現與該骨架模型相關聯的化身

。計算環境亦可基於例如從骨架模型中辨識出的使用者的姿勢來決定在電腦環境上執行的應用程式中要執行何種控制。例如，如圖所示，在圖2中，計算設備12可包括姿勢辨識引擎190。姿勢辨識引擎190通常可包括姿勢篩檢程式集合，每個姿勢篩檢程式皆包括關於骨架模型（在使用者移動時）可執行的姿勢的資訊。

【0064】 可以將由相機26、28及設備20以骨架模型的形式擷取的資料及與該資料相關聯的移動與姿勢辨識引擎190中的姿勢篩檢程式相比較，以識別（如由骨架模型所表示的）使用者何時執行了一或多個姿勢。彼等姿勢可與應用程式的各種控制相關聯。因此，計算設備12可使用姿勢辨識引擎190來解釋骨架模型的移動並基於該移動來控制應用程式。

【0065】 爲了節省功率並有效利用能量，系統10可以在預定時間段內停留在閒置狀態的情況下下電（power down）爲待機功率模式。替代地，使用者可以在用完系統10時手動地關閉該系統。當被關閉時，該系統亦在待機功率模式下執行。當在待機功率模式下時，電源474（圖6）向計算設備12供應待機功率。在實施例中，該待機功率例如可以是0.3瓦至5.0瓦，並且在另一實施例中可以爲0.5瓦。能夠理解，在另外的實施例中，在待機功率模式下供應給系統10的功率可以高於或低於該範圍。

【0066】 根據本系統，使用者能夠利用可聽啓動短語來啓動系統10的計算設備12。在辨識該啓動短語以後，計算設備12可以從待機功率模式切換到全功率模式，其中在全功

率模式中，所有系統資源皆可用。下文描述使用待機啓動單元464啓動計算設備12。關於A/V設備16及擷取設備20，只要該等設備亦可以處於待機功率模式，則該等設備就可以在計算設備12被啓動的時刻被啓動。A/V設備16及/或擷取設備20可以藉由從計算設備12接收「喚醒」訊號被啓動。替代地，A/V設備16及/或擷取設備20亦可以包括待機啓動單元464，該待機啓動單元464以下文所述與計算設備12相同的方式啓動該等設備。

【0067】 本系統的特徵是，本系統能夠利用僅靠在待機功率模式期間所供應的待機功率操作的元件來辨識啓動短語。該等特徵在下文予以說明，但是大體而言，僅僅使用待機功率，處理器468及記憶體470（圖3B及圖6）就能夠接收數位元化音頻圖樣並且對照儲存在記憶體470中的啓動模式執行受限的模式比對以決定啓動短語是否被講出。

【0068】 啓動短語可以是簡單的兩單詞短語，諸如「啓動系統（activate system）」。當系統10例如是來自微軟公司的Xbox®視訊遊戲控制台時，啓動短語例如可以是「Xbox打開」。舉例而言，該等短語及啓動短語可以是儲存在記憶體470中的任何預定義的短語。在另外的實施例中，啓動短語可以短於兩個單詞或長於兩個單詞。與較長的短語相比，較短的短語可能更頻繁地導致錯誤的肯定比對，但是較長的短語可能需要來自待機功率的更多功率以供處理器468執行比對操作。在實施例中，啓動短語可以被設置成針對可用待機功率最佳化過的長度。亦即，啓動短語可以長得足以最小化

錯誤肯定比對，但是短得足以使得處理器468及記憶體470能夠使用可用待機功率來決定啓動短語何時被接收。

【0069】 在實施例中，啓動短語可以由系統10的使用者講出以便將系統10從待機模式中啓動。然而，能夠理解，啓動短語可以是藉由除語音以外的方法講出或產生的任何有區別的音頻圖樣。此外，在被講出時，啓動短語不需要是已知詞彙表中的單詞。啓動短語可以是由任何有區別的聲音模式構成的音頻圖樣。當啓動短語包括一或多個諸如「啓動系統」或「Xbox打開」之類的單詞時，啓動短語可以以單種語言或以多種語言來儲存。

【0070】 在實施例中，啓動短語可以由系統10供應者預先程式設計在處理器470中（亦即在最終使用者最初使用系統10以前）。替代地，啓動短語可以是使用者界定的。圖4是圖示步驟序列的流程圖，其中一或多個使用者各自可以產生及儲存啓動短語。在步驟400，當系統10爲活動時，使用者可以執行控制姿勢或選擇以進入啓動短語學習模式。一旦處於學習模式，使用者就可以在步驟406講出所期望的啓動短語。使用者可以單次講出所期望的啓動短語，或者系統可以在步驟410提示使用者多次講出該短語。

【0071】 在步驟414，把由使用者提供的啓動短語處理成數位元啓動模式。具體而言，該一或多個所接收的啓動短語的實例可以經過A到D（模數）轉換器以提供短語的數位串流。例如，當短語被多次提供時，反常的頻率及/或發音可以被過濾以得到最佳地逼近由使用者在該多個學習實例中提供

的啟動短語的數位元啟動模式。替代地，來自多個學習實例的啟動短語的每次講出皆可以分開地儲存並且之後用於下述比較。一旦所接收的啟動短語已經被處理成數位元啟動模式，就可以在步驟418將該啟動模式儲存在記憶體470中。

【0072】 圖5中圖示所儲存的數位化啟動模式的示意性表示。圖5中所示的示意性表示是類比音訊串流的數位化表示，類比音訊串流可以被表示成麥克風的振動膜隨時間（t）的位移（D）的繪圖，其中由聲波造成氣壓改變。圖5中所示的數位元啟動模式僅僅是示例性的，並且該模式將根據上文參照圖4的流程圖所述般所講出及處理的啟動短語而變化。

【0073】 單個啟動短語可以經由圖4的步驟來產生及儲存。替代地，可以產生及儲存多個啟動短語。如所述般，單個使用者可以產生經由多次講出相同短語所得到的多個數位化啟動短語。替代地，不同的使用者可以產生不同的啟動短語。例如，某家庭或組的成員各自使用系統10，每個使用者皆可以產生及儲存他/她自己的一或多個啟動短語。由每個使用者產生及儲存的啟動短語可以彼此相同或彼此不同。

【0074】 在此種實施例中，每個使用者皆可以基於該每個使用者的特定啟動短語被辨識。因此，當系統啟動時，NUI系統10可以被個人化為所識別出的使用者，諸如例如提供個人化的問候。此外，在此種實施例中，待機啟動單元464可以基於從所感知的啟動短語中識別出何使用者而採取不同的動作。作為一個實例，計算設備12可以被設置為在家長不想要計算設備12被兒童使用的一天的特定的時間忽略來自該兒

童的啟動命令。

【0075】 在啟動短語已經被使用者預程式設計或儲存為數位元啟動模式以後，使用者可以使用該啟動短語來將計算設備12從待機模式啟動為活動模式。圖6是圖示用於使用啟動短語啟動計算設備12的步驟序列的流程圖，並且圖7是圖示用於使用啟動短語啟動計算設備12的元件的方塊圖。圖7圖示一或多個麥克風462，該一或多個麥克風462耦合到包括微處理器468及非揮發性記憶體470的待機啟動單元464。麥克風462可以在計算設備12中提供，或者麥克風462可以是上文參照圖2描述的一或多個麥克風30。圖7亦圖示電源474。電源474在計算設備12處於活動模式時向計算設備12的所有元件供電，包括該一或多個麥克風462、待機啟動單元464及CPU 101/259。電源474在計算設備12處於待機模式時向一或多個麥克風462及待機啟動單元464供電。

【0076】 如上所述，待機啟動單元464可以在單個片上系統上實現。替代地，麥克風468及記憶體470可以是彼此通訊的單獨元件。在另外的實施例中，可以省略單獨的微處理器468，相反，下文所述之待機啟動程序由系統10的CPU（圖3A的多媒體控制台100的CPU 101、圖3B的計算環境220的CPU 259）來執行。在此種實施例中，CPU 101/259在計算設備12處於待機功率模式時接收待機功率，例如0.5瓦。當處於待機功率模式時，CPU 101/259將具有有限的處理能力。隨後，當計算設備12被啟動時，CPU 101/259將以更高功率執行以提供完全的系統功能。

【0077】 在另外的實施例中，可以省略單獨的記憶體470，相反，啟動短語的儲存可以由計算設備12中的其他非揮發性記憶體來完成，諸如多媒體控制台100的ROM 106或者計算環境220的ROM 223。在此種實施例中，ROM 106/223可以被分割，使得在待機模式下操作時，非揮發性記憶體的僅僅一小部分對處理器可用。

【0078】 現在參考圖6的流程圖，在步驟420，該一或多個麥克風462監測周圍環境以監聽音訊串流。若在步驟424偵測到聲音，則在步驟430過濾音訊。例如，可以使用多路回波對消技術來從所接收的音訊串流中過濾背景雜訊。可以在步驟430使用其他雜訊過濾技術來過濾所接收的音訊串流。

【0079】 在步驟432，經由A到D轉換器將所接收的音訊串流數位化成數位元音頻圖樣。隨後將該模式與儲存在記憶體470中的一或多個啟動模式相比較。該比較是由在處理器468中執行的模式比對引擎192執行的。若在步驟440，在某個預定義公差內找到資料串流數位元化音頻圖樣的任何部分與所儲存的啟動模式之間的模式比對，則模式比對引擎192向電源474發送向計算設備12提供全功率的訊號，以便啟動該設備。

【0080】 另一方面，若在步驟440未找到模式比對，則模式比對引擎192可以在步驟444嘗試進一步處理所接收的數位元化音頻圖樣。若未找到模式比對（步驟440）並且模式比對引擎192不能進一步處理音頻圖樣（步驟444），則計算設備12保持在待機功率模式，並且流程返回到步驟420以監聽另

外的音訊。若模式比對引擎192在步驟444能夠進一步處理所接收的數位元化音頻圖樣，則在步驟446處理音頻圖樣，並且在步驟440再次將音頻圖樣與一或多個所儲存的啟動模式相比較。

【0081】 作為可以在步驟446對所接收的數位元化音頻圖樣執行的處理的一個實例，模式比對引擎192可以在時間上拉伸及/或收縮所接收的音頻圖樣以查看該音頻圖樣彼時是否比對一或多個所儲存的啟動模式。可以將各種過濾及雜訊對消技術應用於音頻圖樣以查看該音頻圖樣彼時是否比對一或多個所儲存的啟動模式。構思了對音頻圖樣的其他處理。

【0082】 上文參照步驟446所述之對音頻圖樣的處理可以替代於或附加於上述流程在步驟436將數位元音頻圖樣與一或多個所儲存的數位元啟動模式相比較以前進行，（其中步驟446的處理是僅在步驟440的比較失敗以後才執行的）。亦構思了可以省略步驟444及446。在該實施例中，未執行對音頻圖樣的處理。若在步驟440未找到所接收的數位元音頻圖樣與一或多個所儲存的啟動短語之間的比對，則不啟動設備12。

【0083】 若在步驟440找到模式比對，則在步驟450啟動計算設備12。在此時，待機啟動單元464用訊號通知電源向計算設備12供應全功率，並且所有設備資源於是變得可用。

【0084】 能夠辨識語音（例如啟動短語）的富語音辨識系統是已知的。然而，語音辨識系統需要比在待機功率模式下對計算設備12可用的處理功率更大量的處理功率。另一

方面，由待機啓動單元464將所接收的數位元化音頻圖樣與一或多個所儲存的啓動模式進行的有限的比較需要相對少量的處理功率，並且可以在例如0.5瓦功率可用時在待機模式下執行。

【0085】 在步驟436將所接收的音頻圖樣與若干所儲存的啓動模式進行的比較可能由處理器468消耗大量功率。類似地，在步驟446對所接收的音頻圖樣的顯著處理可能由處理器468消耗大量功率。當本系統的特定實施方式以低待機功率水平操作時，具有若干所儲存的啓動模式及對所接收的音頻圖樣的顯著處理兩者也許是不可能的。在該實例中，不同的實施例可以不同地分配資源。例如，一實施例可以具有大量所儲存的啓動模式及相對少的對所接收數位元音頻圖樣的處理。相反，另一實施例可以具有僅僅一個或相對少的所儲存的啓動模式，但是具有對所接收數位元化音頻圖樣的更穩健的處理。所儲存的啓動模式的數目及長度及可用於所接收的音頻圖樣的處理的量可以基於在計算設備12處於待機功率模式時可用於該等程序的功率量來設置。

【0086】 如上所述，所接收的音頻圖樣可以在兩個數位元模式在某個公差內彼此比對時被認為是與所儲存的啓動短語比對。在此，公差是指所接收的數位元化音頻圖樣需要與所儲存的啓動模式比對的程度。公差可以在不同實施例中以幾種不同方式來量測。公差可以在數位元音頻圖樣中的每個資料點與所儲存的啓動模式之間量測，要求每個點，或者所有點合在一起量測，以具有不小於某個預定幅度的差異。

替代地，該比較可以在數位元化音頻圖樣的資料點組與所儲存的啓動模式之間進行。

【0087】 在一個實例中，所構思的是，數位元化音頻圖樣與所儲存的啓動模式之間的比對需要是完美的或者近似完美的。然而，由於相同的人可能在不同的時間不同地講出相同的短語，從而要求完美比對可能導致使用者難以嘗試啓動計算設備12。因此，在實施例中，公差可以相對低。此可能導致錯誤的肯定啓動。亦即，待機啓動單元464在未提供真實啓動短語時啓動計算設備12。此將擷取到使用者打算啓動計算設備12的實例，但是亦可能在沒有使用者或使用者不打算啓動計算設備12時啓動該系統。

【0088】 當公差爲低使得發生錯誤肯定啓動時，本系統的實施例亦可以包括啓動確認程序，啓動確認程序的一個實例在圖6的步驟454及456中予以圖示。如上所述，富語音辨識引擎可能不可以在待機模式下可用的稀少功率操作。然而，一旦計算系統12如上所述被待機啓動單元464啓動，語音辨識引擎194（圖2）於是可以在步驟454確認該使用者是否實際上講出了正確的啓動短語。若是如此，則計算設備可以保持爲啓動的。若否，則可以在步驟456向電源474發送訊號以回復到待機模式。隨後，流程返回到步驟420以再次監聽可能的啓動短語。

【0089】 在步驟454，語音辨識引擎194可以處理在步驟420接收的類比音訊串流以決定是否講出了合適的啓動短語。替代地，所啓動的計算設備12可以提示使用者講出附加

的短語以供語音辨識引擎194分析。語音辨識引擎可以使用比由待機啓動單元464所執行的模式比對更複雜的演算法來以高得多的確定性確認啓動。

【0090】 能夠理解，在另外的實施例中，啓動確認程序可以替代於或附加於語音辨識引擎194使用元件及/或軟體引擎。例如，一旦被啓動，可以提示使用者執行預定義的啓動確認姿勢，該姿勢被NUI系統10辨識以確認使用者啓動計算系統12並與系統10互動的期望。構思了其他啓動確認程序。

【0091】 此外，能夠理解，步驟454及456的確認程序可以一起被省略。當確認程序被省略時，計算設備12可能在偵測到認為接收到啓動短語以後啓動。隨後，若在預定的時間段內未偵測到與系統10的進一步互動，則計算設備12可以再次進入待機功率模式。

【0092】 上述的取樣實施例描述了NUI系統10的計算設備12中所包括的待機啓動單元464。然而，如上所述，待機啓動單元464可以作為其他電設備的一部分來提供，包括在圖3A及圖3B分別描述的通用遊戲控制台或計算環境。此種系統可以是PC、膝上型電腦、智慧型電話、控制器及/或其他手持計算設備。在另外的實施例中，待機啓動單元464可以包括在電器內，諸如但不限於洗衣機/乾燥器、咖啡機、電視及立體音響系統及電車庫門。具有待機啓動單元的電設備亦可以包括但不限於汽車、防盜警報系統及照明系統。能夠理解，在另外的實施例中，使用模式比對的待機啓動單元464可以用在多種其他電設備中。

【0093】 上述實施例揭示偵測計算設備12附近的音訊的一或多個麥克風。然而，可能發生的是，包括待機啓動單元464的設備亦能夠在待機功率模式下操作時發送及/或接收資料串流的傳輸。在此種實施例中所構思的是，音訊源可以處於計算設備12的遠端。

【0094】 例如，音訊源可以在第一位置被偵測到，並且隨後例如經由諸如無線網路、區域網路、寬頻網路及/或網際網路之類的網路被傳輸給第二位置處的設備。音訊源可以由第一或第二位置處的設備來數位化及/或處理。隨後，第二位置處的設備中的待機啓動單元464可以執行上述比對比較以決定從第一位置處的設備接收的音訊是否是合適的啓動短語。若是如此，則可以啓動第二位置處的設備。替代地，第二位置處的設備可以向第一位置處的設備發送回傳輸或者向不同於第一或第二位置的第三位置處的設備發送傳輸以啓動第一或第三位置處的設備。

【0095】 本發明系統的前述詳細描述是出於說明及描述的目的而提供的。此並不意欲窮舉本發明系統或將本發明系統限於所揭示的精確形式。鑒於上述教導，許多修改及變型皆是可能的。選擇所述實施例以最好地解釋本發明系統的原理及本發明系統的實踐應用，從而允許本領域技藝人士能夠在各種實施例中並採用各種適於所構想的特定用途的修改來最好地利用本發明系統。本發明系統的範圍意欲由所附申請專利範圍來界定。

【符號說明】

【0001】

10 目標辨識、分析及追蹤系統

12 計算設備

14 顯示器

16 A/V設備

18 使用者

19 化身或螢幕上人物

20 擷取設備

21 使用者介面

22 圖像相機元件

24 IR光元件

26 3-D相機

28 RGB相機

30 麥克風

32 處理器

34 記憶體元件

36 通訊鏈路

100 多媒體控制台

101 中央處理單元

101/259 中央處理單元

102 一級快取記憶體

104 二級快取記憶體

106 快閃記憶體ROM

108 圖形處理單元

- 110 記憶體控制器
- 112 記憶體
- 114 視訊轉碼器/視訊轉碼器
- 118 模組
- 120 I/O控制器
- 122 系統管理控制器
- 123 音訊處理單元
- 124 網路介面控制器
- 126 第一USB主控制器
- 128 第二USB主控制器
- 130 前面板I/O子部件
- 132 音訊轉碼器
- 136 系統供電模組
- 138 風扇
- 140 A/V（音訊/視訊）埠
- 142(1) 周邊控制器
- 142(2) 周邊控制器
- 143 系統記憶體
- 144 媒體驅動器
- 146 外置記憶體設備
- 148 無線配接器
- 150 電源按鈕
- 152 彈出按鈕
- 190 姿勢辨識引擎

- 192 模式比對引擎
- 194 語音辨識引擎
- 220 計算環境
- 221 系統匯流排
- 222 系統記憶體
- 223 ROM
- 224 基本輸入/輸出系統
- 225 作業系統
- 226 應用程式
- 227 其他程式模組
- 228 程式資料
- 232 視訊介面
- 233 輸出周邊介面
- 234 介面
- 235 介面
- 236 使用者輸入介面
- 237 配接器
- 238 硬碟
- 239 磁碟機
- 240 光碟機
- 241 電腦
- 242 監視器
- 243 印表機
- 244 揚聲器

- 245 區域網路
- 246 遠端電腦
- 247 記憶體儲存設備
- 248 遠端應用程式
- 249 廣域網
- 250 數據機
- 251 鍵盤
- 252 定點設備
- 253 光碟
- 254 磁碟
- 255 程式資料
- 256 其他程式模組
- 257 應用程式
- 258 作業系統
- 259 中央處理單元
- 260 RAM
- 400 步驟
- 406 步驟
- 410 步驟
- 414 步驟
- 418 步驟
- 420 步驟
- 424 步驟
- 430 步驟

432 步驟

436 步驟

440 步驟

444 步驟

446 步驟

450 步驟

454 步驟

456 步驟

462 麥克風

464 待機啓動單元

468 處理器

470 非揮發性記憶體

474 電源

【生物材料寄存】

國內寄存資訊【請依寄存機構、日期、號碼順序註記】

無

國外寄存資訊【請依寄存國家、機構、日期、號碼順序註記】

無

【序列表】(請換頁單獨記載)

無

申請專利範圍

1. 一種用於將一電設備從一待機功率模式啟動的方法，該方法包括以下步驟：
 - (a) 在處於待機功率模式時在該電設備中接收一音訊串流；
 - (b) 在處於待機功率模式時將該音訊串流數位化成一音頻圖樣；
 - (c) 在處於待機功率模式時將在該步驟 (b) 中數位元化的該音頻圖樣與儲存在一非揮發性記憶體中的一數位元化啟動短語相比較；
 - (d) 在該音頻圖樣在一預定義的公差內與該啟動短語比對(match)時啟動該電設備；及
 - (e) 在該電設備於該步驟(d)中被啟動以後，在該電設備啟動時利用該電設備可用的資源來確認該音頻圖樣與該啟動短語比對。
2. 如請求項1述及之方法，其中一麥克風連續地監測一周圍環境以獲得該步驟 (a) 的該音訊串流。
3. 如請求項1述及之方法，亦包括在將該數位元化音頻圖樣與一數位元化啟動短語相比較以前從該所接收的音訊串流中過濾雜訊的步驟。
4. 如請求項3述及之方法，其中該過濾步驟包括多路雜訊回

波對消。

5. 如請求項1述及之方法，進一步包括處理該數位化音頻圖樣的步驟。
6. 如請求項5述及之方法，其中該數位化音頻圖樣之該處理在比較該數位化音頻圖樣與該數位化啟動模式之前進行。
7. 如請求項5述及之方法，其中該數位化音頻圖樣之該處理僅在該步驟(c)的一比較未能識別該預定義的公差內的一模式比對以後進行。
8. 如請求項5述及之方法，其中該數位化音頻圖樣之該處理包括在時間上拉伸及收縮該數位化音頻圖樣。
9. 如請求項1述及之方法，其中在一使用者最初使用該電設備以前將該數位元化啟動模式儲存在該電設備的該非揮發性記憶體中。
10. 如請求項1述及之方法，其中在該電設備提示以後由一使用者選擇該數位元化啟動模式。
11. 一種用於將一電設備從一待機功率模式啟動的待機啟動

系統，待機啟動系統包括：

一或多個麥克風，該一或多個麥克風用於偵測該一或多個麥克風附近的一音訊串流；及

一待機啟動單元，包括：

一非揮發性記憶體，該非揮發性記憶體包括一或多個所儲存的啟動模式；及

一處理器，該處理器用於使用在待機功率模式下對該電設備可用的功率將從該一或多個麥克風接收的一數位元化音頻圖樣與儲存在該非揮發性記憶體中的該一或多個所儲存的啟動模式相比較，該處理器在該數位元化音頻圖樣在一預定義的公差內與儲存在該非揮發性記憶體中的一或多個啟動短語中的一個啟動短語比對(match)時啟動該設備，且若該電設備可用的資源在啟動時未確認該音訊串流為一講出的啟動命令，該處理器將該電設備於啟動時回復到待機功率模式。

12. 如請求項11述及之待機啟動系統，其中用於在待機功率模式下將該數位元化音頻圖樣與該一或多個所儲存的啟動模式相比較的該處理器是一第一處理器，該電設備包括與該第一處理器分開的一第二處理器，該第二處理器用於在該設備在一被啟動的全功率模式下操作時控制該電設備的功能。

13. 如請求項11述及之待機啟動系統，其中包括一或多個所

儲存的啟動模式的該非揮發性記憶體是一第一非揮發性記憶體，該電設備包括與該第一非揮發性記憶體分開的一第二非揮發性記憶體，該第二非揮發性記憶體用於在該設備在一被啟動的全功率模式下操作時儲存該電設備的資料。

14. 如請求項11述及之待機啟動系統，其中該非揮發性記憶體及處理器被整合為一片上系統而至該電設備內。

15. 如請求項11述及之待機啟動系統，其中在待機功率模式中對該電設備可用的該功率係在5.0瓦與0.3瓦之間。

16. 如請求項11述及之待機啟動系統，其中在待機功率模式下對該電設備可用的該功率為0.5瓦。

17. 一種用於將一NUI系統的一計算設備從一待機功率模式啟動的方法，該方法包括以下步驟：

(a) 在處於待機功率模式時在與該NUI系統相關聯的一或多個麥克風中接收一音訊串流；

(b) 在處於待機功率模式時將該音訊串流數位化成一音頻圖樣；

(c) 在處於待機功率模式時處理該數位化音訊串流以促進該數位化音訊串流與儲存在該計算設備的一非揮發性記憶體中的一或多個數位化啟動短語的比對(matching)；

(d) 在處於待機功率模式時將在該步驟處理的該音頻圖樣與儲存在該計算設備的該非揮發性記憶體中的一或多個數位化啟動短語相比較；

(e) 當該音頻圖樣在一預定義的公差內與該一或多個啟動短語中的一個啟動短語比對時啟動該計算設備；及

(f) 在該音頻圖樣與該啟動短語之間比對(match)時，在該計算設備於該步驟(e)中被啟動以後，確認該啟動是否為所意欲的，該確認的步驟於該計算設備完全啟動時，使用一語音辨識引擎，該語音辨識引擎是該計算設備可用的。

18. 如請求項17述及之方法，進一步包括以下步驟：一或多個使用者產生儲存在該非揮發性記憶體中的該啟動短語的該步驟，在該計算設備執行在一啟動全功率模式中時進行的一或多個使用者產生該啟動短語的該步驟。

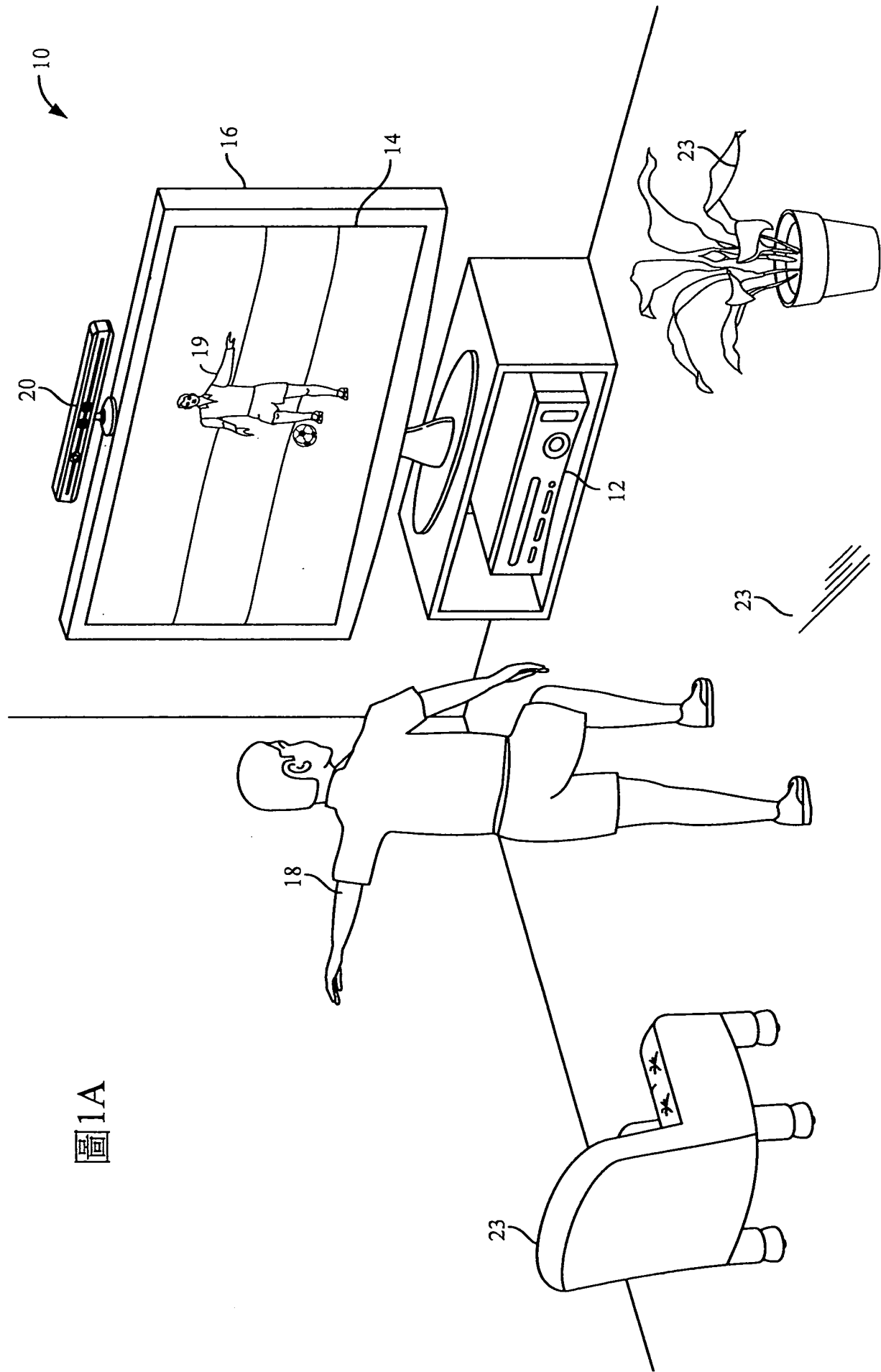


圖 1A

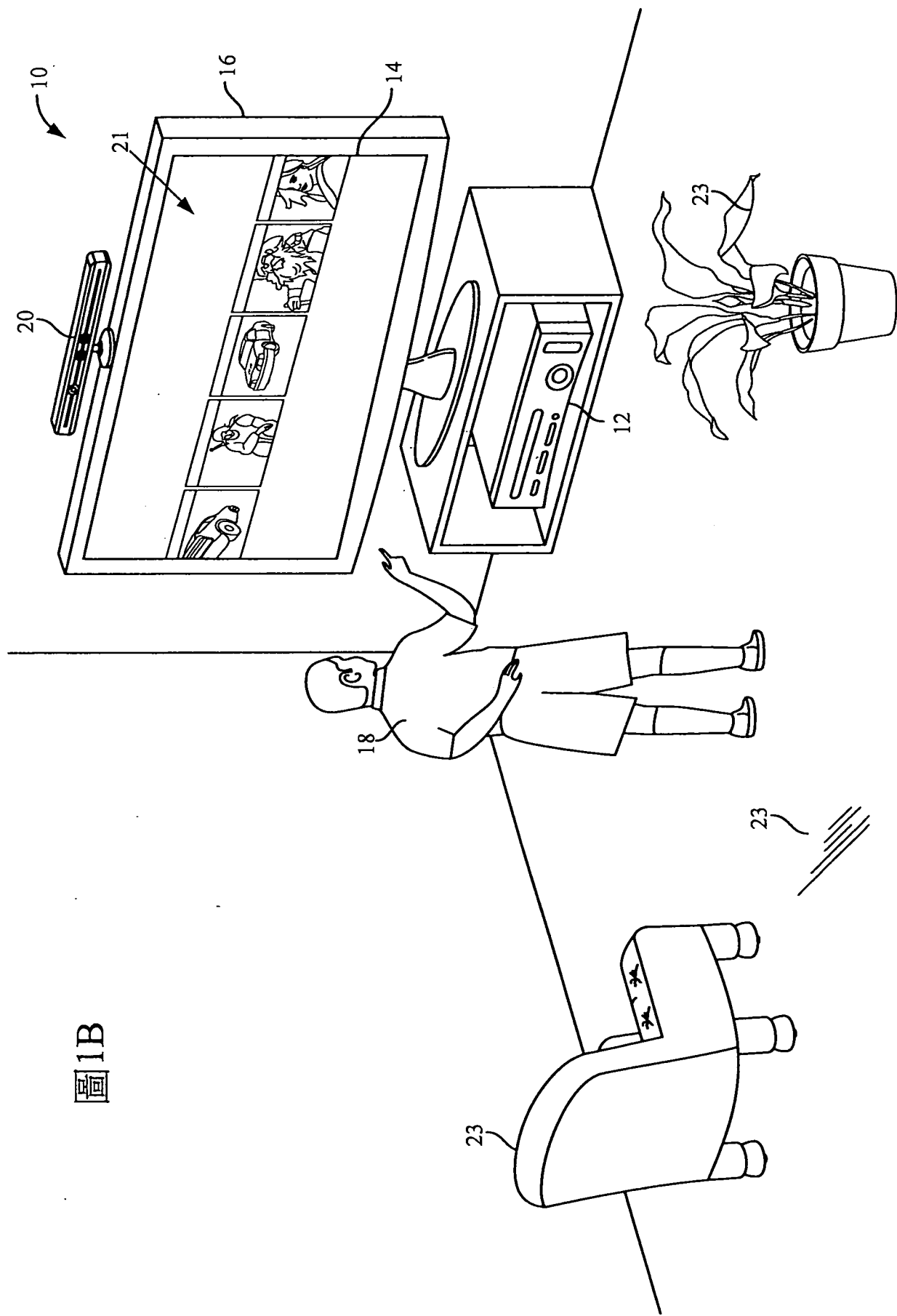
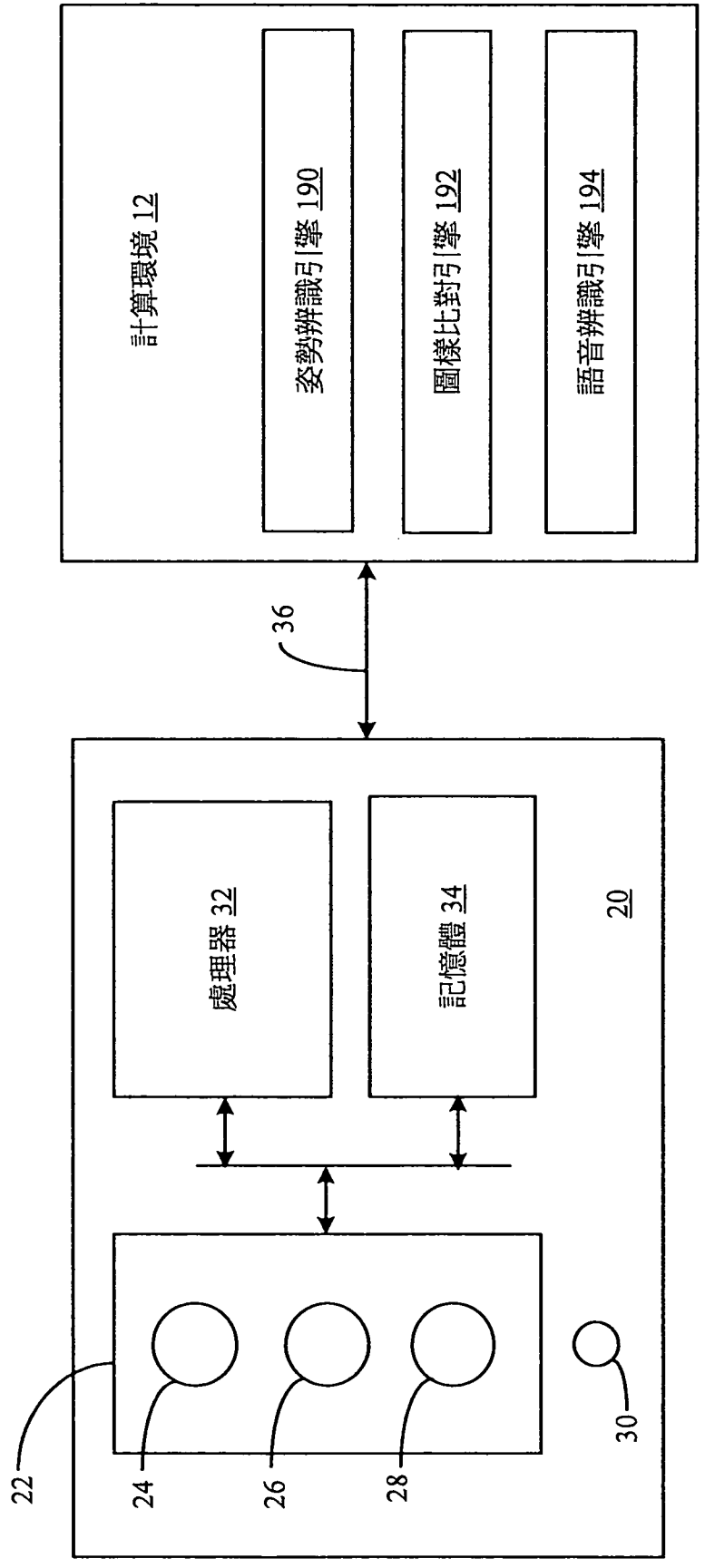


圖1B

圖2

10



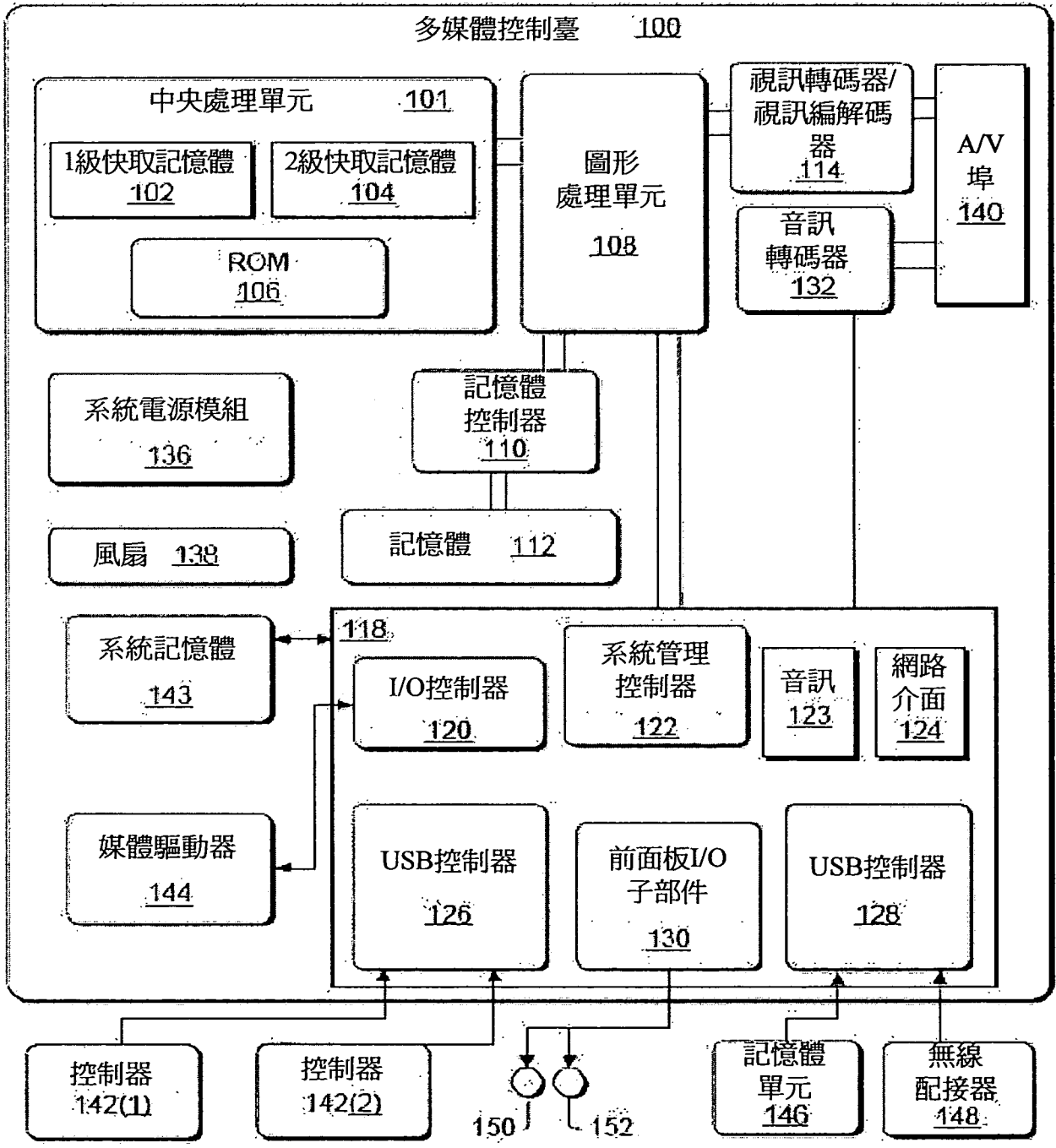


圖3A

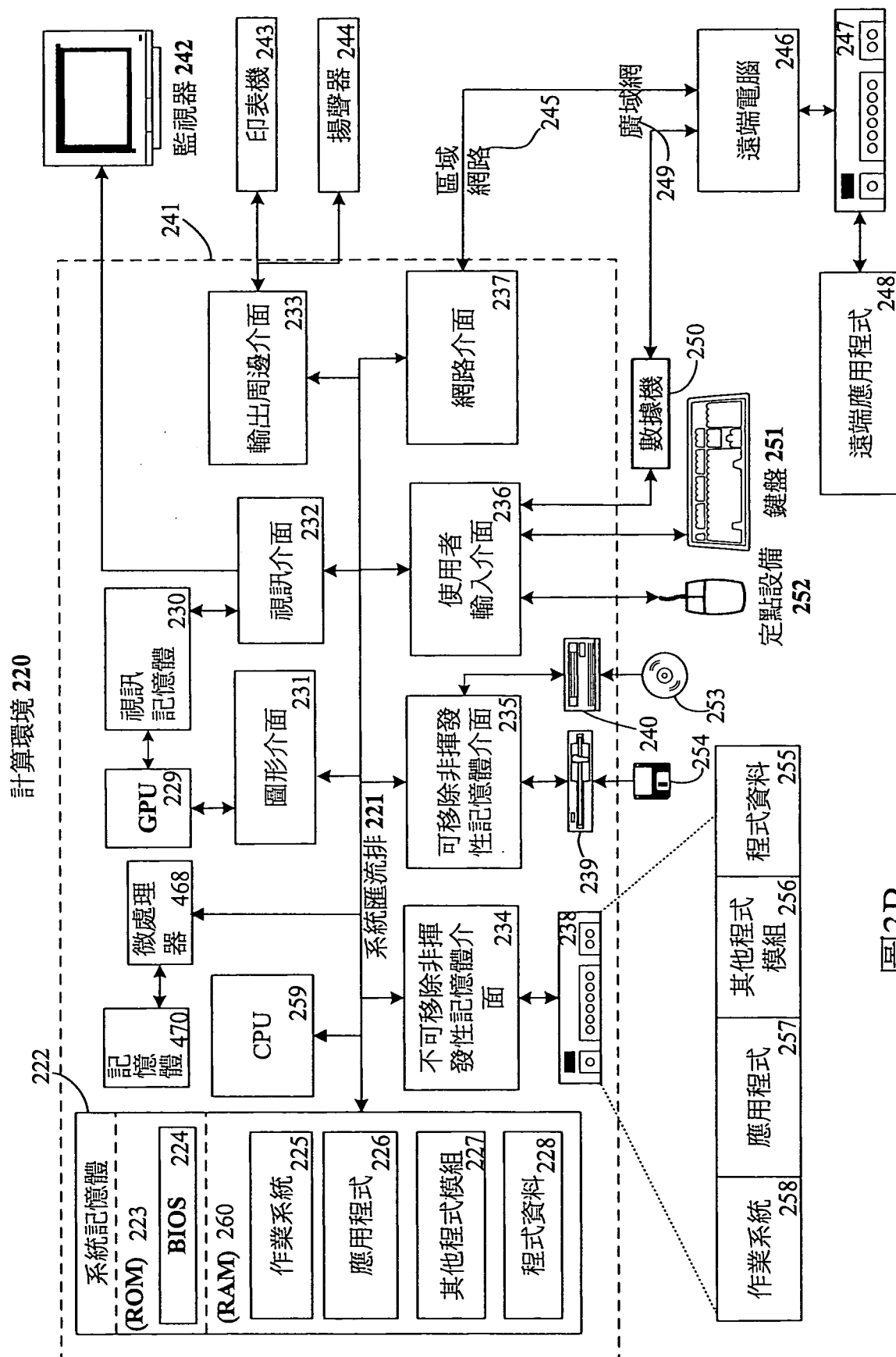


圖4

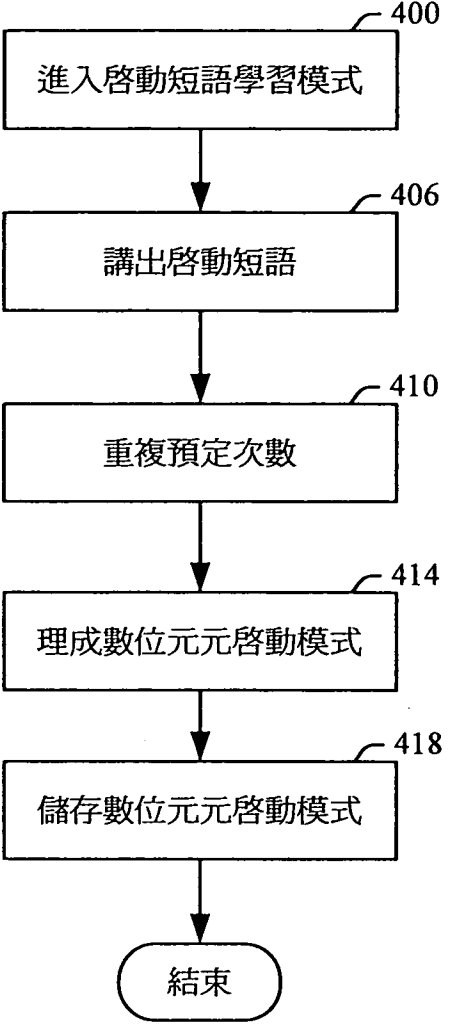


圖5

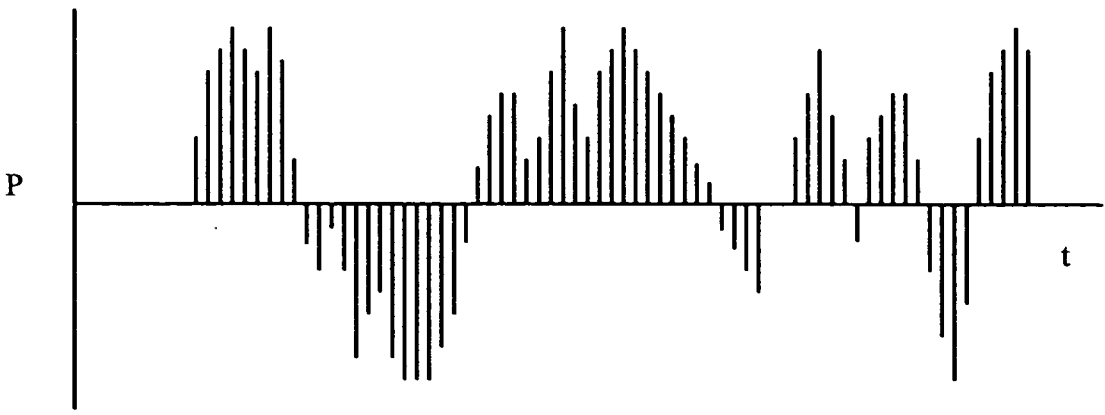


圖6

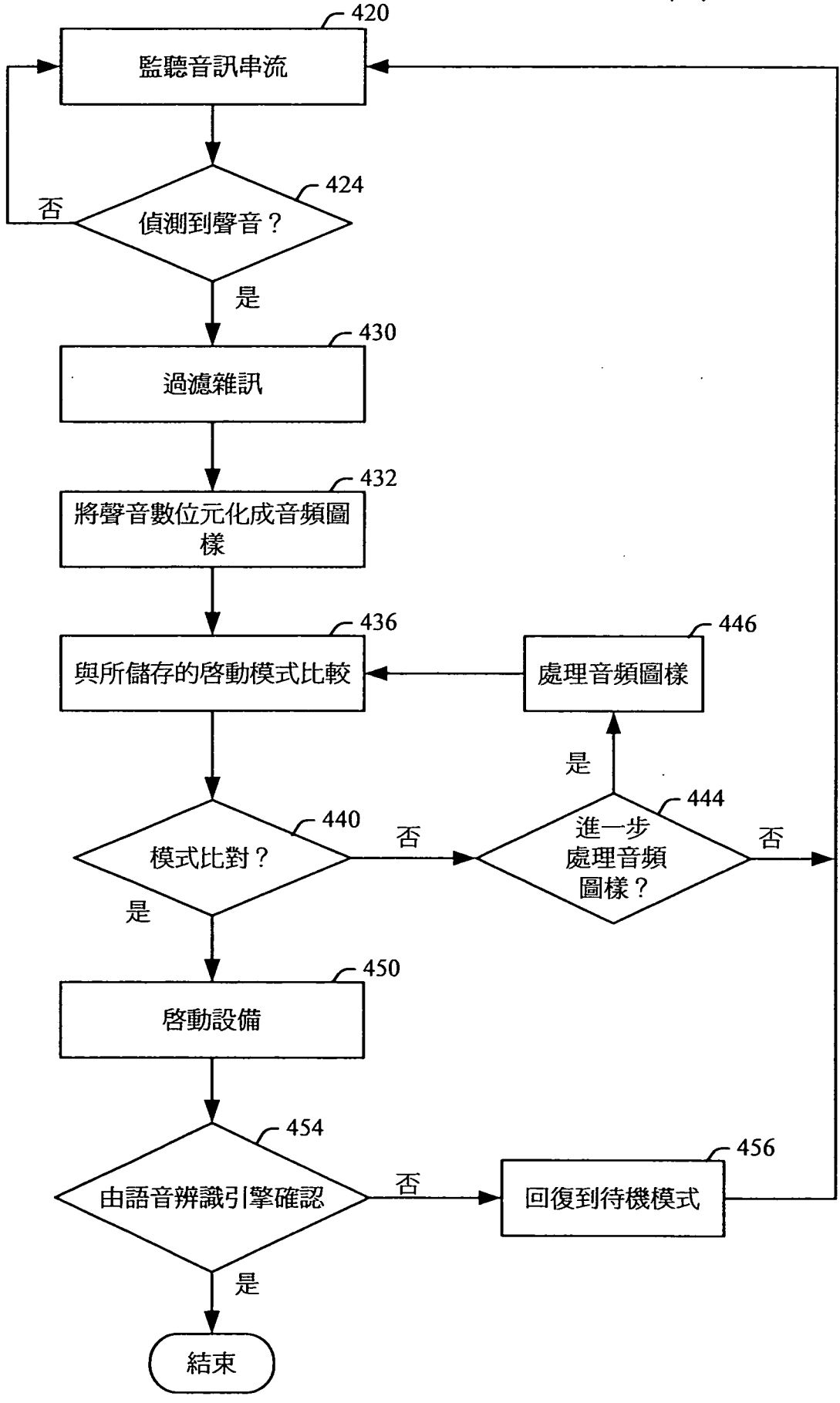


圖7

