

(21) 申請案號：106103276

(22) 申請日：中華民國 102 (2013) 年 04 月 08 日

(51) Int. Cl. :

H04N7/15 (2006.01)

G06F3/01 (2006.01)

(30) 優先權：2012/04/09

世界智慧財產權組織

PCT/CN2012/000458

(71) 申請人：英特爾公司 (美國) INTEL CORPORATION (US)

美國

(72) 發明人：童曉芬 TONG, XIAOFENG (CN)；李文龍 LI, WENLONG (CN)；杜楊洲 DU,

YANGZHOU (CN)；胡威 HU, WEI (CN)；張益明 ZHANG, YIMIN (CN)

(74) 代理人：憐軼群；陳文郎

申請實體審查：有 申請專利範圍項數：31 項 圖式數：4 共 49 頁

## (54) 名稱

使用互動化身的通訊技術 (三)

COMMUNICATION USING INTERACTIVE AVATARS

## (57) 摘要

本揭露一般描述一種視訊通訊系統，其以動畫化身替換參與使用者之實際活動影像。一種方法可包括選擇一化身；起始通訊；偵測一使用者輸入；識別該使用者輸入；基於該使用者輸入識別一動畫命令；產生化身參數；及傳輸該動畫命令及該等化身參數中之至少一者。

Generally this disclosure describes a video communication system that replaces actual live images of the participating users with animated avatars. A method may include selecting an avatar; initiating communication; detecting a user input; identifying the user input; identifying an animation command based on the user input; generating avatar parameters; and transmitting at least one of the animation command and the avatar parameters.

指定代表圖：

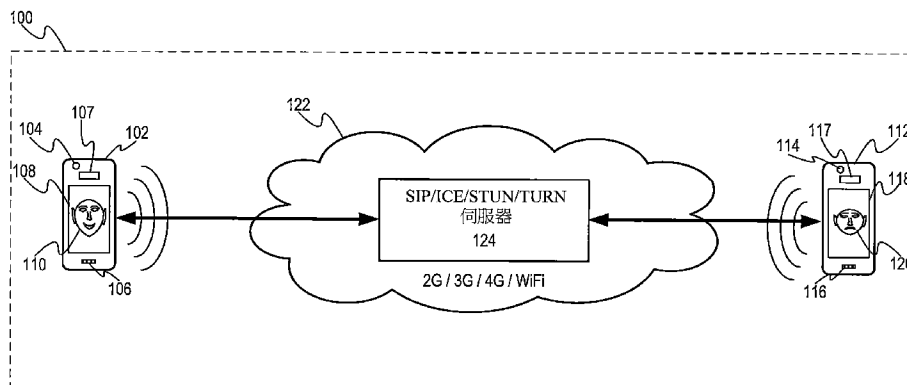


圖 1A

符號簡單說明：

100 . . . 裝置至裝置  
系統/系統102、112 . . . 裝置/  
遠程裝置104、114 . . . 攝影  
機106、116 . . . 麥克  
風107、117 . . . 揚聲  
器108、118 . . . 觸摸  
感應顯示器/顯示器

110、120 . . . 化身

122 . . . 網路

124 . . . 伺服器

# 發明摘要

※ 申請案號：106103276 (由102112332分割)

※ 申請日：102年4月8日

※IPC 分類：  
*H04N 7/15* (2006.01)  
*G06F 3/01* (2006.01)

## 【發明名稱】(中文/英文)

使用互動化身的通訊技術(三)

Communication using Interactive Avatars

## 【中文】

本揭露一般描述一種視訊通訊系統，其以動畫化身替換參與使用者之實際活動影像。一種方法可包括選擇一化身；起始通訊；偵測一使用者輸入；識別該使用者輸入；基於該使用者輸入識別一動畫命令；產生化身參數；及傳輸該動畫命令及該等化身參數中之至少一者。

## 【英文】

Generally this disclosure describes a video communication system that replaces actual live images of the participating users with animated avatars. A method may include selecting an avatar; initiating communication; detecting a user input; identifying the user input; identifying an animation command based on the user input; generating avatar parameters; and transmitting at least one of the animation command and the avatar parameters.

**【代表圖】**

**【本案指定代表圖】**：第（1A）圖。

**【本代表圖之符號簡單說明】**：

- 100...裝置至裝置系統/系統
- 102、112...裝置/遠程裝置
- 104、114...攝影機
- 106、116...麥克風
- 107、117...揚聲器
- 108、118...觸摸感應顯示器/顯示器
- 110、120...化身
- 122...網路
- 124...伺服器

**【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)

使用互動化身的通訊技術(三)

Communication using Interactive Avatars

## 【技術領域】

發明領域

[0001]以下揭示案係有關於視訊通訊，且更特定言之，係有關於使用互動化身的視訊通訊。

## 【先前技術】

發明背景

[0002]行動裝置中可利用的功能性種類之增加使得使用者產生對除單純的呼叫之外經由視訊進行通訊的渴望。例如，使用者可起始「視訊呼叫」、「視訊會議」等等，其中裝置中之攝影機及麥克風捕獲使用者之音訊及視訊，該音訊及視訊實時傳輸至一或多個其他接收者，諸如其他行動裝置、桌上型電腦、視訊會議系統等等。視訊之通訊可涉及實質量之資料的傳輸(例如，取決於攝影機之技術、用來處理經擷取影像資料之特定視訊編碼解碼器，等等)。考慮到現存 2G/3G 無線技術之帶寬限制及新興 4G 無線技術之仍有限的帶寬，許多裝置使用者進行的並行視訊呼叫可超出在現存無線通訊基礎架構中可利用的帶寬，從而可負面地影響視訊呼叫之品質。

## 【發明內容】

[0003] 依據本發明之一實施例，係特地提出一種系統，其包含：一使用者輸入裝置，其組配來擷取一使用者輸入；一通訊模組，其組配來傳輸及接收資訊；以及一或多個儲存媒體，該儲存媒體上單獨地或組合地儲存有指令，當藉由一或多個處理器執行該等指令時導致包含以下之操作：選擇一化身；起始通訊；偵測一使用者輸入；識別該使用者輸入；基於該使用者輸入識別一動畫命令；產生化身參數；以及傳輸該動畫命令及該等化身參數中之至少一者。

### 【圖式簡單說明】

[0004] 所請求標的之各種實施例的特徵及優點將以下隨詳細說明之進行並於參閱圖式之後而變得明顯，圖示中相同數字指定相同部分，且其中：

圖 1A 例示根據本揭示案之各種實施例的示例裝置至裝置系統；

圖 1B 例示根據本揭示案之各種實施例的示例虛擬空間系統；

圖 2 例示根據本揭示案之各種實施例的示例裝置；

圖 3 例示根據本揭示案之至少一實施例的示例系統實行方案；以及

圖 4 為根據本揭示案之至少一實施例的示例操作的流程圖。

[0005] 雖然以下詳細說明係參考說明性實施例來進行，但是熟習此項技術者將明白該等實施例之許多替代例、修改形式及變化形式。

## 【實施方式】

較佳實施例之詳細說明

[0006]通常，本揭示案描述使用互動化身來視訊通訊的系統及方法。與活動影像相對，使用化身大體上減少將要傳輸之資料量，且因此化身通訊需要較少的帶寬。互動化身係組配來藉由基於使用者輸入改變所選化身之顯示而增強使用者體驗。此外，使用者語音可獲擷取及轉換來產生化身語音。化身語音可隨後與使用者語音有關，但可遮掩使用者之身份。音訊轉換可包括例如音調偏移及/或時間延長。

[0007]在一實施例中，啟動耦接至攝影機、麥克風及揚聲器之裝置中的應用程式。該應用程式可組配來允許使用者選擇用於顯示於遠程裝置上、虛擬空間中等等之化身。裝置可隨後組配來起始與至少一其他裝置、虛擬空間等等的通訊。例如，通訊可經由 2G、3G、4G 蜂巢式連接來建立。或者或另外，通訊可經由網際網路，經由 WiFi 連接來建立。通訊建立之後，攝影機可組配來開始擷取影像及/或離物體的距離，且麥克風可組配來開始擷取聲音，例如使用者語音，且將使用者語音轉化成使用者語音信號。

[0008]隨後可判定是否偵測到使用者輸入。使用者輸入可藉由使用者輸入設備擷取。使用者輸入包括藉由觸摸感應顯示器所擷取之觸摸事件及藉由攝影機所擷取的手勢，該攝影機例如組配來擷取離物體之距離的深度攝影機，及/或 web 攝影機。因此，使用者輸入裝置包括觸摸感應顯示

器及/或攝影機。若偵測到使用者輸入，則可識別使用者輸入。對於觸摸事件，使用者輸入識別符可與觸摸類型及一或多個觸摸位置有關。對於手勢(例如張開手)而言，使用者輸入識別符可與手勢識別符有關。動畫命令可隨後基於使用者輸入來識別。動畫命令相應於與使用者輸入相關聯的所要響應，例如響應於所顯示化身之外觀上的單次輕觸而改變所顯示化身外觀的顏色。

[0009]隨後可產生化身參數。化身參數可基於面部偵測、頭部移動及/或動畫命令來產生。化身參數可因此包括基於例如面部偵測及頭部移動的被動組件，及基於動畫命令的互動組件。化身參數可用於使化身於至少一其他裝置上、於虛擬空間內等等成動畫。在一實施例中，化身參數可基於面部偵測、頭部移動及動畫命令來產生。在該實施例中，所得動畫包括基於面部偵測及頭部移動的被動動畫，其藉由基於動畫命令的互動動畫來修改。因此，化身動畫可包括基於例如面部偵測及頭部移動的被動動畫，及基於使用者輸入的互動動畫。

[0010]可隨後傳輸動畫命令及化身參數中之至少一者。在一實施例中，接收遠程動畫命令及遠程化身參數中之至少一者。該遠程動畫命令可使裝置基於遠程動畫命令來判定化身參數以便使所顯示化身成動畫。遠程化身參數可使裝置基於所接收之遠程化身參數來使所顯示化身成動畫。

[0011]音訊通訊可伴隨化身動畫。通訊建立之後，麥克

風可組配來擷取音訊輸入(聲音)，例如使用者語音，且將所擷取的聲音轉化成相應音訊信號(例如使用者語音信號)。在一實施例中，使用者語音信號可轉換成化身語音信號，其可隨後獲編碼及傳輸。所接收之化身語音信號可隨後藉由揚聲器轉化回聲音(例如化身語音)。化身語音可因此基於使用者語音且可保存內容但可改變與所擷取語音相關聯之頻譜資料。例如，轉換包括但不限於音調偏移時間延長及/或轉化回放率。

[0012]使用者輸入裝置(例如觸摸感應顯示器及/或攝影機)可組配來擷取使用者輸入，該等使用者輸入係組配來基於至少一其他裝置上之使用者輸入來使化身成動畫。使用者驅動之動畫(基於動畫命令)可另外為基於面部表情及/或頭部移動之動畫。動畫命令可包括但不限於化身顯示之定位改變、面部特徵扭曲、改變特徵來傳達情緒等等。動畫命令可因此修改與基於面部偵測/追蹤之動畫類似的化身動畫及/或除基於面部偵測/追蹤之動畫之外修改化身動畫。動畫命令可產生時間有限之動畫且可基於來自遠程使用者的輸入而使所得動畫例示於本地使用者之所顯示化身上。

[0013]因此，有限帶寬視訊通訊系統可使用化身來實行。音訊可加以轉換且視訊可基於所偵測之使用者輸入及所識別之動畫命令而成動畫，從而利用化身通訊來增強使用者體驗。此外，可使用化身保存匿名性，包括如本文所述之音訊轉換。

[0014]圖 1A 例示與本揭示案之各種實施例一致的裝置至裝置系統 100。系統 100 可通常包括經由網路 122 通訊之裝置 102 及 112。裝置 102 至少包括攝影機 104、麥克風 106、揚聲器 107 及觸摸感應顯示器 108。裝置 112 至少包括攝影機 114、麥克風 116、揚聲器 117 及觸摸感應顯示器 118。網路 122 至少包括伺服器 124。

[0015]裝置 102 及 112 可包括能夠有線通訊及/或無線通訊之各種硬體平台。例如，裝置 102 及 112 可包括但不限於視訊會議系統、桌上型電腦、膝上型電腦、平板電腦、智慧型電話(例如，iPhones®、基於 Android 的電話、Blackberries®、基於 Symbian®的電話、基於 Palm®的電話等等)、蜂巢式手機等等。攝影機 104 及 114 包括用於擷取代表包括一或多個人的環境的數位影像之任何裝置，且可具有足夠解析度以用於如本文所述之外觀分析及/或手勢識別。例如，攝影機 104 及 114 可包括靜物攝影機(例如，組配來擷取靜止照片之攝影機)或視訊攝影機(例如，組配來擷取由多個訊框組成之移動影像的攝影機)。攝影機 104 及 114 可組配來使用可見光譜中的光或利用電磁譜中不限於紅外光譜、紫外光譜等等之其他部分的光來操作。在一實施例中，攝影機 104 及 114 可組配來偵測深度，亦即，攝影機離物體及/或該物體上之點的距離。攝影機 104 及 114 可分別併入裝置 102 及 112 中，或可為組配來與經由有線通訊或無線通訊而與裝置 102 及 112 通訊的獨立裝置。攝影機 104 及 114 之特定實例可包括有線(例如，通用串列匯

流排(USB)、乙太網路、火線等等)或無線(例如，WiFi、藍牙等等)web 攝影機，如其可與電腦、視訊監視器等等相關聯；深度攝影機；行動裝置攝影機(例如，手機或智慧型電話攝影機，其例如整合至先前論述之示例裝置中)；整合式膝上型電腦攝影機；整合式平板電腦攝影機(例如，iPad®、Galaxy Tab®及類似攝影機)等等。

[0016]裝置 102 及 112 可進一步包含麥克風 106 及 116 及揚聲器 107 及 117。麥克風 106 及 116 包括任何裝置，其組配來感測(亦即，擷取)聲音且將感測之聲音轉化成相應音訊信號。麥克風 106 及 116 可分別整合於裝置 102 及 112 內部，或可經由有線通訊或無線通訊與該等裝置交互作用，諸如上文關於攝影機 104 及 114 之實例中所述。揚聲器 107 及 117 包括任何裝置，其組配來將音訊信號轉化成相應聲音。揚聲器 107 及 117 可分別整合於裝置 102 及 112 內部，或可經由有線通訊或無線通訊與該等裝置交互作用，諸如上文關於攝影機 104 及 114 之實例中所述。觸摸感應顯示器 108 及 118 包括任何裝置，該等裝置係組配來顯示文字、靜止影像、移動影像(例如視訊)、使用者介面、圖形等等且係組配來感測諸如輕觸、重擊等等之觸摸事件。觸摸事件可包括觸摸類型及觸摸位置。觸摸感應顯示器 108 及 118 可分別整合於裝置 102 及 112 內部，或可經由有線通訊或無線通訊與該等裝置交互作用，諸如上文關於攝影機 104 及 114 之實例中所述。在一實施例中，顯示器 108 及 118 係組配來分別顯示化身 110 及 120。如本文所

提及，化身係定義為使用者於二維(2D)或三維(3D)中的圖形表示。化身不必類似於使用者之面容，且因此雖然化身可為逼真的表示，但該等化身還可以採取圖畫、卡通、草圖等等的形式。在系統 100 中，裝置 102 可顯示表示裝置 112 之使用者(例如遠程使用者)的化身 110，且同樣地，裝置 112 可顯示表示裝置 102 之使用者的化身 120。以此方式，使用者可看見其他使用者之表示，而不必交換涉及使用活動影像的裝置至裝置通訊的大量資訊。此外，化身可基於使用者輸入而成動畫。以此方式，使用者可與本地及/或遠程化身之顯示器交互作用，進而增強使用者體驗。所得動畫可提供相比於可能僅使用面部偵測及追蹤時更大範圍的動畫。此外，使用者可主動選擇該等動畫。

[0017]如本文所提及，化身音訊(亦即聲音)係定義為經轉換之使用者音訊(聲音)。例如，聲音輸入可包括使用者之**嗓音**，亦即，使用者語音，且相應化身音訊可包括經轉換之使用者語音。化身音訊可與使用者音訊有關。例如，化身語音可相應於音調偏移、時間延長及/或使用語音之其他轉換。化身語音可類似於人類語音或可相應於卡通人物等等。在系統 100 中，裝置 102 可發出表示裝置 112 之遠程使用者的化身音訊，且類似地，裝置 112 可發出表示藉由裝置 102 擷取之音訊的化身音訊(例如，裝置 102 之本地使用者的語音)。以此方式，使用者可聽到可經轉換的其他使用者之嗓音的表示。

[0018]網路 122 可包括各種第二代(2G)、第三代(3G)、

第四代(4G)基於蜂巢式的資料通訊技術、Wi-Fi 無線資料通訊技術等等。網路 122 包括至少一伺服器 124，該伺服器組配來在使用此等技術時建立並保持通訊連接。例如，伺服器 124 可組配來支援網際網路有關的通訊協定，如對話啟動協定(SIP)，其用於建立、修改及終止兩方(單播)及多方(多播)對話；互動連接性建立協定(ICE)，其用於呈現允許協定建立於位元串流連接之頂端的框架；網路存取轉換器(NAT)對話穿越實用機制協定(STUN)，其允許經由 NAT 操作之應用程式，以便發現其他 NAT 之存在；IP 位址及埠，其經分配用於應用程式之使用者資料報協定(UDP)連接以便連接至遠程主機；使用中繼穿越 NAT(TURN)，其允許 NAT 或防火牆背後之元件經由傳輸控制協定(TCP)或 UDP 連接來接收資料，等等。

[0019]圖 1B 例示與本揭示案之各種實施例一致的虛擬空間系統 126。系統 126 可使用裝置 102、裝置 112 及伺服器 124。裝置 102、裝置 112 及伺服器 124 可繼續以與圖 1A 中所例示相似的方式來通訊，但可在虛擬空間 128 中發生使用者交互作用替代以裝置至裝置格式發生使用者交互作用。如本文所提及，虛擬空間可定義為實體位置之數位模擬。例如，虛擬空間 128 可類似於戶外位置，如同城市、道路、人行道、田野、森林、島嶼等等，或戶內位置，如同辦公室、房屋、學校、商場、商店等等。由化身表示之使用者可與現實世界中一樣看起來與虛擬空間 128 交互作用。虛擬空間 128 可存在於與網際網路耦接之一或多個伺

服务器上，且可藉由第三方保持。虛擬空間之實例包括虛擬辦公室、虛擬會議室、如同 Second Life®之虛擬世界、如同 World of Warcraft®之大規模多人線上角色扮演遊戲(MMORPG)、如同 The Sims Online®之大規模多人線上真實生活遊戲(MMORLG)等等。在系統 126，虛擬空間 128 可含有多個對應於不同使用者之化身。替代所顯示化身，顯示器 108 及 118 可顯示包封(例如較小)型式之虛擬空間(VS) 128。例如，顯示器 108 可顯示對應於裝置 102 之使用者的化身在虛擬空間 128 所「看見」內容的透視圖。類似地，顯示器 118 可顯示對應於裝置 112 之使用者的化身在虛擬空間 128 所「看見」內容的透視圖。化身可能在虛擬空間 128 看見的內容之實例包括但不限於虛擬結構(例如建築物)、虛擬車輛、虛擬物體、虛擬動物、其他化身等等。

[0020]圖 2 例示根據本揭示案之各種實施例的示例裝置 102。雖然僅描述裝置 102，但是裝置 112(例如遠程裝置)可包括組配來提供相同或類似功能之資源。如先前所論述，裝置 102 展示為包括攝影機 104、麥克風 106、揚聲器 107 及觸摸感應顯示器 108。攝影機 104、麥克風 106 及觸摸感應顯示器 108 可對攝影機、音訊及觸控螢幕框架模組 200 提供輸入，且攝影機、音訊及觸控螢幕框架模組 200 可提供對揚聲器 107 之輸出(例如音訊信號)。攝影機、音訊及觸控螢幕框架模組 200 可包括慣用的、專屬的、已知的及/或以後開發的音訊及視訊處理碼(或指令集)，該音訊及視訊處理碼通常經明確界定且可操作來至少控制攝影機

104、麥克風 106、揚聲器 107 及觸摸感應顯示器 108。例如，攝影機、音訊及觸控螢幕框架模組 200 可使攝影機 104、麥克風 106、揚聲器 107 及觸摸感應顯示器 108 記錄影像、離物體之距離、聲音及/或觸摸，可處理影像、聲音、音訊信號及/或觸摸，可使影像及/或聲音獲複製，可對揚聲器 107 提供音訊信號，等等。攝影機、音訊及觸控螢幕框架模組 200 可取決於裝置 102，且更尤其取決於裝置 102 中運作之作業系統(OS)而變化。示例作業系統包括 iOS®、Android®、Blackberry® OS、Symbian®、Palm® OS，等等。揚聲器 107 可接收來自攝影機、音訊及觸控螢幕框架模組 200 之音訊資訊，且可組配來複製本地聲音(例如，以便提供使用者嗓音之音訊回饋，該音訊回饋經轉換或未經轉換)及遠程聲音(例如，於虛擬位置中參與電話、視訊呼叫或進行交互作用的另一方或多方之聲音(經轉換或未經轉換))。

[0021] 面部偵測及追蹤模組 202 可組配來識別及追蹤藉由攝影機 104 提供的影像內的頭部、臉部及/或面部區。例如，面部偵測模組 204 可包括慣用的、專屬的、已知的及/或以後開發的臉部偵測碼(或指令集)、硬體及/或軟體，其通常經明確界定且可操作來接收標準格式影像(例如但不限於 RGB 彩色影像)且至少在某種程度上識別影像中的臉部。面部偵測及追蹤模組 202 亦可組配來經由一系列影像(例如處於每秒 24 個訊框下的視訊訊框)追蹤所偵測之臉部且基於所偵測之臉部判定頭部位置。可藉由面部偵測/追蹤模組 202 使用的已知追蹤系統可包括粒子濾波、平均變

動、卡爾曼濾波等等，其中每一者皆可利用邊緣分析、平方和方差分析、特徵點分析、直方圖分析、膚色分析等等。

[0022]特徵提取模組 204 可組配來辨識藉由臉部偵測模組 202 偵測之臉部中的特徵(例如，面部指標(諸如眼睛、眉毛、鼻、嘴等等)之位置及/或形狀)。在一實施例中，化身動畫可直接地基於所感測之面部動作(例如面部特徵之改變)無需進行面部表情識別。化身臉部上之相應特徵點可遵循或模仿真實人的臉部的移動，此稱為「表情仿製」或「表演驅動的面部動畫」。特徵提取模組 204 可包括慣用的、專屬的、已知的及/或以後開發的面部特性辨識碼(或指令集)，其通常經明確界定且可操作來接收來自攝影機 104 之標準格式影像(例如但不限於 RGB 彩色影像)且至少在某程度上提取影像中的一或多個面部特性。此等已知面部特徵系統包括但不限於 Colorado State University 的 CSU 臉部識別評價系統。

[0023]特徵提取模組 204 亦可組配來辨識與所偵測之特徵相關聯的表情(例如，識別先前偵測的臉部是否高興、悲哀、微笑、皺眉頭、驚訝、興奮等等)。因此，特徵提取模組 204 可進一步包括慣用的、專屬的、已知的及/或以後開發的面部表情偵測及/或標識碼(或指令集)，其通常經明確界定且可操作來偵測及/或識別臉部中的表情。例如，特徵提取模組 204 可判定面部特徵(例如眼睛、嘴、頰、牙齒等等)之大小及/或位置，且可將此等面部特徵與面部特徵資料庫比較，該面部特徵資料庫包括具有相應面部特徵類別

(例如、微笑、皺眉頭、興奮、悲哀等等)的多個樣本面部特徵。

[0024]音訊轉換模組 206 係組配來將使用者之嗓音轉換成化身嗓音，亦即，經轉換的使用者之嗓音。轉換包括調整節奏(例如延長時間)、音調(例如音調偏移)及回放率。例如，音訊轉換模組 206 可包括慣用的、專屬的、已知的及/或以後開發的音訊轉換碼(或指令集)，其通常經明確界定且可操作來接收表示使用者之嗓音的嗓音資料，且將該等嗓音資料轉化成經轉換的嗓音資料。嗓音資料可與基於藉由麥克風 106 擷取且藉由攝影機、音訊及觸控螢幕框架模組 200 處理的聲音的音訊信號有關。此類已知嗓音轉換系統包括但不限於聲控開啟式資源音訊處理庫，其係組配來調整音訊串流或音訊檔案之節奏、音調及回放率。

[0025]音訊轉換模組 206 可包括多個預定義嗓音風格，其相應於與轉換使用者之嗓音相關聯的轉換參數。例如，轉換參數可組配來以不同音調及/或節奏保持人聽到的經轉換嗓音輸出。對人類女性或如兒童的嗓音而言，音調可偏移至高頻率；對人類男性的嗓音而言，音調可偏移至較低頻率，可向上或向下調整節奏以便增大或減小語音之速度，等等。在另一實例中，該等轉換參數可組配來產生相應於如動物的嗓音(例如貓)及/或卡通人物類嗓音的經轉換嗓音輸出。此可藉由調整使用者語音之音調、其他頻率分量及/或取樣參數來達成。

[0026]使用者可於起始通訊之前選擇所要音訊轉換輸

出，及/或可在通訊期間選擇所要音訊轉換。音訊轉換模組 206 可組配來提供響應於來自使用者之請求的樣本音訊轉換輸出。在一實施例中，音訊轉換模組 206 可包括允許使用者選擇音訊轉換參數來產生客製音訊轉換輸出的設施。該設施可組配來基於使用者之嗓音輸入提供樣本經轉換音訊輸出。使用者可隨後調整音訊轉換參數(例如，嘗試錯誤法)直至達成適合的轉換輸出。與對使用者之適合輸出相關聯的音訊轉換參數可隨後儲存及/或利用來進行化身通訊，如本文所述。

[0027]觸摸偵測模組 208 係組配來接收來自攝影機、音訊及觸控螢幕框架模組 200 之觸摸資料且基於所接收之觸摸資料識別觸摸事件。觸摸事件識別符可包括觸摸類型及/或觸摸位置。觸摸類型可包括單一輕觸、雙重輕觸、輕觸及保持、輕觸及移動、按壓及延展、重擊等等。觸摸位置可包括觸摸開始位置、觸摸結束位置及/或中間移動觸摸位置等等。觸摸位置可相應於觸摸感應顯示器 108 之坐標。觸摸偵測模組 208 可包括慣用的、專屬的、已知的及/或以後開發的觸摸偵測碼(或指令集)，其通常經明確界定且可操作來接收觸摸資料且識別觸摸事件。

[0028]手勢偵測模組 210 係組配來接收來自攝影機、音訊及觸控螢幕框架模組 200 的深度及/或影像資料，基於所接收之深度及/或影像資料辨識相應手勢，且基於所辨識之手勢判定手勢識別符。深度相應於攝影機至物體之距離。手勢識別符與所辨識之手勢有關。手勢偵測模組 210 可包

括慣用的、專屬的、已知的及/或以後開發的手勢偵測碼(或指令集)，其通常經明確界定且可操作來基於所接收之深度及/或影像資料識別手勢。

[0029]例如，手勢偵測模組 210 可包括預定義手勢之資料庫。該預定義手勢可包括至少一些相對普通、相對簡單的手勢，包括張開手、合緊手(亦即，握拳)、揮手、用手做圓周運動、手自右至左移動、手自左至右移動等等。因此，手勢可包括靜態非移動的手手勢、活動移動的手手勢及/或其組合。在一實施例中，手勢偵測模組 210 可包括訓練設施，其組配來允許使用者改變預定義手勢及/或訓練新手勢。客製手勢及/或新手勢可隨後與手勢識別符相關聯，且該手勢識別符可與動畫命令相關聯，如本文所述。例如，使用者可選擇動畫命令以與來自動畫命令之預定義清單中的手勢相關聯。

[0030]因此，動畫命令與對使用者輸入之所要響應有關。動畫命令可與例如觸摸事件識別符及/或手勢識別符之所識別使用者輸入相關聯。以此方式，使用者可與所顯示化身交互作用及/或可設定手勢以便修改所顯示化身之動畫。

[0031]化身選擇模組 212 係組配來允許裝置 102 之使用者選擇用於在遠程裝置上顯示之化身。化身選擇模組 212 可包括慣用的、專屬的、已知的及/或以後開發的使用者介面構建碼(或指令集)，其通常經明確界定且可操作來向使用者呈現不同化身，以便該使用者可選擇該等化身之一。在

一實施例中，一或多個化身可預定義於裝置 102 中。預定義化身允許所有裝置具有相同化身，且在交互作用期間僅化身之選擇(例如預定義化身之識別)需要與遠程裝置或虛擬空間通訊，從而減少需要交換的資訊之量。化身係於建立通訊之前選擇，但亦可在主動通訊過程中加以改變。因此，可能於通訊期間任何點處發送或接收化身選擇，且接收裝置可能根據所接收之化身選擇來改變所顯示化身。

[0032]化身控制模組 214 係組配來接收基於裝置 102 之使用者輸入的使用者輸入識別符。使用者輸入識別符可包括藉由觸摸偵測模組 208 基於觸摸事件資料所判定之觸摸事件識別符或藉由手勢偵測模組 210 所判定之手勢識別符。觸摸事件資料包括觸摸類型及/或觸摸位置。觸摸位置可相應於與觸摸感應顯示器 108 相關聯的坐標。觸摸位置可對映至所顯示化身上之一或多個點，例如對映至一特徵，例如鼻尖、嘴、唇、耳朵、眼睛等等。所顯示化身上之點可與化身動畫之所要響應(亦即動畫命令)有關。

[0033]化身控制模組 214 係組配來基於使用者輸入識別符(亦即所識別之使用者輸入)判定動畫命令。動畫命令係組配來識別所要化身動畫。例如，所要動畫包括改變所顯示化身之臉部的顏色，改變所顯示化身之一特徵的大小(例如使鼻更大)、使眼色、眨眼、微笑，移除一特徵(例如耳朵)等等。因此，化身控制模組 214 係組配來接收使用者輸入識別符且基於該使用者輸入識別符判定動畫命令。

[0034]化身控制模組 214 係組配來基於動畫命令實行

化身動畫。在一實施例中，對例如裝置 112 之遠程裝置上顯示的互動動畫而言，可傳輸動畫命令且遠程化身控制模組可隨後實行該動畫。在另一實施例中，該等化身參數可經傳輸組配用於化身動畫之立即實行方案。

[0035]基於動畫命令之所實行之互動動畫可具有有限持續時間，在該有限持續時間之後，該化身動畫可回到基於例如如本文所述之面部偵測及追蹤的被動動畫。影響特徵之大小的所實行之互動動畫可組配來逐漸改變大小且逐漸回到初始大小。另外或替代地，影響特徵之大小的動畫可組配來具有一效果梯度。換言之，大小改變的相對量值可取決於相對於例如關鍵頂點之位置。所顯示化身上更接近關鍵頂點之點可經歷比所顯示化身上相對更遠的點更大的改變。

[0036]因此，化身控制模組 214 可接收基於使用者輸入的使用者輸入識別符，可基於該使用者輸入識別符判定動畫命令且可基於該動畫命令實行動畫。基於動畫命令之互動動畫可為時間受限於一時間週期(持續時間)的及/或可包括效果梯度。動畫可在該時間週期之後回到基於面部偵測及追蹤的被動化身動畫。

[0037]化身控制模組 214 係組配來產生用於使化身成動畫之參數。如本文所提及，動畫可定義為改變影像/模型之外觀。動畫包括基於例如面部表情及/或頭部移動的被動動畫及基於使用者輸入的互動動畫。單一動畫(可包括被動動畫及互動動畫)可改變二維靜止影像之外觀，或多個動畫

可依次存在以模擬影像之運動(例如轉頭、點頭、眨眼、交談、皺眉頭、微笑、大笑、使眼色、眨眼等等)。用於三維模型之動畫的實例包括使三維線框模型變形、應用紋理對映及重新計算用於正常顯現之模型頂點。所偵測之臉部及/或所提取之面部特徵之位置的改變可轉化成使化身之特徵類似於使用者臉部之特徵的參數。在一實施例中，所偵測之臉部的一般表情可轉化成使化身顯示相同表情的一或多個參數。化身之表情亦可予以誇示以強調該表情。當化身參數可通常應用於所有預定義化身時，對所選擇之化身的認識可並非必需。然而，在一實施例中，化身參數可對所選擇之化身為特定的，且因此可在選擇另一化身之情況下加以改變。例如，人類化身可需要與動物化身、卡通化身等等不同的參數設置(例如，可改變不同化身特徵)來演示如高興、悲哀、生氣、驚訝等等之情緒。

[0038]化身控制模組 214 可包括慣用的、專屬的、已知的及/或以後開發的圖形處理碼(或指令集)，其通常經明確界定且可操作來使藉由化身選擇模組 212 基於面部偵測及追蹤模組 202 所偵測的臉部/頭部位置、特徵提取模組 204 所偵測的面部特徵及/或觸摸偵測模組 208 及/或手勢偵測模組 210 所判定的使用者輸入識別符而選擇的化身成動畫。對基於面部特徵之動畫方法而言，二維化身動畫可例如用影像翹曲或影像漸變來完成，而三維化身動畫可用自由形式變形(FFD)或藉由利用頭部三維模型中定義之動畫結構來完成。Oddcast 為可用於二維化身動畫之軟體資源的

一實例，而 FaceGen 為可用於三維化身動畫之軟體資源的一實例。

[0039] 例如，對包括拉長三維化身顯示之鼻部的互動動畫而言，可定義(例如選擇)與鼻尖有關的關鍵頂點  $v_k$ 。相關聯三維運動向量  $d_k$  ( $dx$ 、 $dy$ 、 $dz$ )及效果半徑  $R$  可定義用於關鍵頂點  $v_k$ 。效果半徑  $R$  內之其他頂點可在互動動畫中改變(亦即移動)，而效果半徑  $R$  外部之頂點可保持不因互動動畫而改變。互動動畫可具有相關聯之持續時間，即動畫時間  $T$ ，其可延伸達多個訊框。暫時效果參數  $\eta_t$  可基於時間  $t$  及動畫時間  $T$  來定義，如：

$$\eta_t = \begin{cases} t/T, & 0 \leq t < T \\ 2 - t/T, & T \leq t < 2T \\ 0, & \text{否則} \end{cases}$$

[0040] 效果半徑  $R$  內相對更接近  $v_k$  之頂點可比相對更遠離關鍵頂點  $v_k$  之頂點相對更大地改變。一頂點  $v_i$  之空間效果參數  $\eta_i$  可定義為：

$$\eta_i = \exp\left(-\frac{\|v_k - v_i\|^2}{R^2}\right)$$

[0041] 而頂點  $v_i$  在時間  $t$  的運動向量  $d_i^t$  則可定義為  $d_i^t = \eta_t \cdot \eta_i \cdot d_k$ 。互動動畫化身之新坐標則為  $v_i^t = v_i^0 + d_i^t$ ，其中  $v_i^0$  相應於頂點  $v_i$  之坐標，其基於面部偵測及追蹤，亦即，被動動畫。

[0042] 因此，可對包括修改互動動畫之被動動畫的所顯示化身實行動畫。互動動畫可受限於總體持續時間，且該動畫之效果的量值可在該持續時間內變化。互動動畫可組

配來僅影響化身的一部分，且該等效果可對較接近關鍵頂點之點而言較大。互動動畫完成之後，動畫可基於如本文所述的面部偵測及追蹤來繼續。

[0043]此外，在系統 100 中，化身控制模組 214 可接收遠程化身選擇及可用於顯示相應於遠程裝置處之使用者的化身並使其成動畫的遠程化身參數。動畫可包括被動動畫以及互動動畫。化身控制模組可使顯示模組 216 於顯示器 108 上顯示化身 110。顯示模組 216 可包括慣用的、專屬的、已知的及/或以後開發的圖形處理碼(或指令集)，其通常經明確界定且可操作來根據示例裝置至裝置實施例在顯示器 108 上顯示化身且使其呈動畫。例如，化身控制模組 214 可接收遠程化身選擇且可使該遠程化身選擇解譯以相應於預定化身。顯示模組 216 可隨後在顯示器 108 上顯示化身 110。此外，化身控制模組 214 中所接收的遠程化身參數可獲解譯，且可將命令提供至顯示器模組 216 以使化身 110 成動畫。在一實施例中，兩個以上的使用者可參與視訊呼叫。當兩個以上的使用者在視訊呼叫中交互作用時，顯示器 108 可分割或分段以允許一個以上的相應於遠程使用者之化身將同時顯示 或者，在系統 126 中，化身控制模組 214 可接收資訊，該資訊使顯示器模組 216 顯示相應於裝置 102 之使用者的化身在虛擬空間 128 中(例如，自該化身之虛擬透視角度)「看見」的內容。例如，顯示器 108 可顯示虛擬空間 128 中表示的建築物、物體、動物，其他化身，等等。

[0044]在一實施例中，化身控制模組 214 可組配來使顯

示器模組 216 顯示「回饋」化身 218。回饋化身 218 表示所選擇化身如何呈現在遠程裝置上、在虛擬位置中等等。詳言之，回饋化身 218 呈現為藉由使用者選擇之化身且可使用藉由化身控制模組 214 產生的相同參數來成動畫。以此方式，使用者可確認遠程使用者在其交互作用期間看見之內容。回饋化身 218 亦可用來顯示藉由裝置 112 之遠程使用者輸入所引起的互動動畫。因此，本地使用者可與其回饋化身(例如化身 218 及裝置 102 之使用者)交互作用，以使其相關聯化身之互動動畫顯示至裝置 112 上的遠程使用者。本地使用者可類似地與遠程使用者之所顯示化身(例如化身 110)交互作用，從而使遠程使用者之回饋化身的互動動畫顯示於裝置 112 上。

[0045] 通訊模組 220 係組配來傳輸及接收資訊以用於選擇化身、顯示化身、使化身成動畫、顯示虛擬位置透視圖等等。通訊模組 220 可包括慣用的、專屬的、已知的及/或以後開發的通訊處理碼(或指令集)，其通常經明確界定且可操作來傳輸化身選擇、化身參數、動畫命令、互動化身參數及接收遠程化身選擇、遠程化身參數、遠程動畫命令及遠程互動化身參數。通訊模組 220 亦可傳輸及接收相應於基於化身之交互作用的音訊資訊。通訊模組 220 可經由網路 122 傳輸及接收以上資訊，如先前所述。

[0046] 處理器 222 係組配來執行與裝置 102 及其中所包括模組的一或多者相關聯之操作。

[0047] 圖 3 例示根據至少一實施例之示例系統實行方

案。裝置 102'係組配來經由 WiFi 連接 300 來無線地通訊(例如在工作時)，伺服器 124'係組配來經由網際網路 302 協商裝置 102'與 112'之間的連接，且裝置 112'係組配來經由另一 WiFi 連接 304 來無線地通訊（例如在家時）。在一實施例中，基於裝置至裝置化身之視訊呼叫應用程式在裝置 102'中啟動。在化身選擇之後，應用程式可允許選擇至少一遠程裝置(例如裝置 112')。應用程式可隨後使裝置 102'起始與裝置 112'之通訊。通訊可以裝置 102'經由企業存取點(AP) 306 傳輸連接建立請求至裝置 112'來起始。企業 AP 306 可為可用於商業設置之 AP，且因此可支援比家 AP 314 高的資料通量及更多的並行無線客戶端。企業 AP 306 可接收來自裝置 102'之無線信號，且可經由各種商用網路，經由閘道 308 進行對連接建立請求的傳輸。連接建立請求可隨後通過防火牆 310，該防火牆可組配來控制流入及流出 WiFi 網路 300 之資訊。

[0048]裝置 102'之連接建立請求可隨後藉由伺服器 124'處理。伺服器 124'可組配來登記 IP 位址、鑑別目的地位址及 NAT 穿越，以便連接建立請求可導向網際網路 302 上的正確目的地。例如，伺服器 124'可自接收自裝置 102 的連接建立請求中的資訊來解析所欲之目的地(例如遠程裝置 112')，且可將信號安排路由傳遞穿過正確 NAT、埠及因此到達目的地 IP 位址。此等操作可僅必須在連接建立期間執行，此取決於網路組態。在一些情況下，可在視訊呼叫期間重複操作以便向 NAT 提供通知來保持連接有效。媒

體及信號路徑 312 可在已建立連接之後將視訊(例如化身選擇及/或化身參數)及音訊資訊指導攜帶至家 AP 314。裝置 112'可隨後接收連接建立請求且可組配來判定是否接受該請求。判定是否接受該請求可包括例如向查詢關於是否接收來自裝置 102'之連接請求的裝置 112'之使用者呈現視覺敘事。裝置 112'之使用者接收該連接(例如，接收該視訊呼叫)，即可建立該連接。攝影機 104'及 114'可組配來隨後開始分別擷取裝置 102'及 112'之各自使用者的影像，以用於是藉由各使用者選擇之化身成動畫。麥克風 106'及 116'可組配來隨後開始擷取來自各使用者之音訊。當在裝置 102'及 112'之間開始資訊交換時，顯示器 108'及 118'可顯示相應於裝置 102'及 112'之使用者的化身且使該等化身成動畫。

[0049]圖 4 例示與本揭示案之一實施例一致的示範性操作的流程圖 400。該等操作可例如藉由裝置 102 及/或 112 執行。詳言之，流程圖 400 描繪組配來實行化身動畫(包括被動動畫及/或互動動畫)及/或音訊轉換以用於裝置之間經由網路的通訊的操作。假定面部偵測及追蹤、特徵提取及被動化身動畫如本文所述加以實行及操作。

[0050]化身模型可在操作 402 選擇。化身模型可包括視訊化身選擇及音訊轉換選擇。可顯示多個視訊化身模型，使用者可自該等視訊化身模型選擇一所要化身。在一實施例中，選擇視訊化身模型可包括相關聯音訊轉換。例如，如貓的化身可與如貓的音訊轉換相關聯。在另一實施例中，音訊轉換可獨立於該視訊化身選擇來選擇。

[0051]包括音訊轉換之化身模型可在啟動通訊之前選擇，但亦可在活動通訊的過程中加以改變。因此，可能於通訊期間任何點處發送或接收化身選擇及/或改變音訊轉換選擇，且接收裝置可能根據所接收之化身選擇來改變所顯示化身。

[0052]化身通訊可在操作 404 啟動。例如，使用者可運行組配來使用如本文所述化身傳達音訊及視訊之應用程式。操作 404 可包括組配通訊及建立連接。通訊組態包括識別參與視訊呼叫之至少一遠程裝置或虛擬空間。例如，使用者可自儲存於應用程式內、儲存於與另一系統相關聯的裝置內(例如智慧型電話、手機等等中的聯絡人清單)、遠程儲存於諸如網際網路(例如，如 Facebook、LinkedIn、Yahoo、Google+、MSN 等等的社交媒體網站)上的的遠程使用者/裝置之清單中進行選擇。或者，使用者可選擇在如 Second Life 的虛擬空間中進行線上操作。

[0053]在操作 406，裝置中之攝影機可隨後開始擷取影像及/或深度，且裝置中之麥克風可開始擷取聲音。影像可為靜止影像或活動影像(例如，依次擷取的多個影像)。深度可與影像一起擷取或可獨立地擷取。深度相應於攝影機之視場中攝影機至物體(及物體上之點)的距離。可在操作 408 判定是否偵測到使用者輸入。使用者輸入包括藉由影像及/或深度攝影機擷取的手勢及在觸摸感應顯示器上偵測到之觸摸輸入。若偵測到使用者輸入，則可在操作 410 識別使用者輸入。使用者輸入識別符包括觸摸識別符或手勢識別

符。觸摸識別符可基於對觸摸感應顯示器的觸摸來判定且可包括觸摸類型及觸摸位置。手勢識別符可基於所擷取影像及/或深度資料來判定且可包括辨識手勢。

[0054]可在操作 412 識別動畫命令。動畫命令可組配來使顯示於遠程裝置上的使用者之所選擇化身成動畫，或使亦顯示於遠程使用者之裝置上的遠程使用者之回饋化身成動畫。動畫命令相應於與使用者輸入相關聯的所要響應。例如，觸摸所顯示化身的臉部(使用者輸入)可產生所顯示化身的臉部之顏色改變(藉由動畫命令識別的所要響應)。動畫命令可基於所識別之使用者輸入來識別。例如，各使用者輸入可與具有使用者輸入識別符及動畫命令之資料庫中的動畫命令有關(例如與之相關聯)。

[0055]操作 414 包括產生化身參數。化身參數包括被動組件且可包括互動組件。若未偵測到使用者輸入，則化身參數可包括被動組件。若偵測到使用者輸入，則化身參數是否可包括互動組件取決於動畫命令並因此取決於使用者輸入。對於相應於組配來使使用者之所選擇化身成動畫的動畫命令之使用者輸入而言，動畫命令可與僅包括被動組件之化身參數一起傳輸或可在傳輸之前應用於化身參數，以便所傳輸之化身參數包括被動組件及互動組件。對於相應於組配來使顯示於遠程使用者之裝置上的遠程使用者之回饋化身成動畫的動畫命令之輸入而言，可僅傳輸動畫命令。

[0056]操作 416 包括轉換及編碼所擷取音訊。所擷取音

訊可轉化成音訊信號(例如使用者語音信號)。使用者語音信號可根據操作 402 之化身選擇的音訊轉換部分來轉換。經轉換之使用者語音信號相應於化身語音信號。化身語音信號可使用已知用於經由網路傳輸至遠程裝置及/或虛擬空間的技術來編碼。可在操作 418 處傳輸經轉換及編碼之音訊。操作 418 可進一步包括傳輸動畫命令及化身參數中之至少一者。傳輸動畫命令係組配來允許遠程裝置藉由根據動畫命令修改化身參數而使本地所顯示化身成動畫。已在傳輸之前根據動畫命令修改的經傳輸化身參數可直接用來使顯示於遠程裝置上的化身成動畫。換言之，由動畫命令表示的對化身參數之修改可在本地執行或遠程執行。

[0057]操作 420 包括接收可為經轉換音訊之遠程編碼音訊。操作 420 進一步包括接收遠程動畫命令及遠程化身參數中之至少一者。遠程動畫命令可用來修改相應於遠程使用者之所顯示化身或本地使用者之所顯示回饋化身的化身參數。動畫命令及化身參數係組配來產生基於使用者輸入加以修改的化身動畫。在操作 422 處，所接收之音訊可獲解碼及播放，且在操作 424 處，化身可獲顯示及成動畫。

[0058]所顯示化身之動畫可基於所偵測及識別之使用者輸入，如本文所述。在裝置至裝置通訊(例如系統 100)之示例中，遠程化身選擇或遠程化身參數中至少一者可接收自遠程裝置。相應於遠程使用者之化身可隨後基於所接收之遠程化身選擇來顯示，且可基於所接收之遠程化身參數而成動畫。在虛擬位置交互作用(例如系統 126)之示例

中，可接收允許裝置顯示相應於裝置使用者之化身所看見的內容的資訊。

[0059]可在操作 426 處判定通訊是否完成。若通訊完成，即可在操作 428 處結束程式流。若通訊未完成，程式流即可繼續進行至操作 406，擷取影像、深度及/或音訊。

[0060]雖然圖 4 例示根據一實施例之各種操作，但是要理解的是，並非圖 4 中描繪的所有操作皆為其他實施例所必需。事實上，本文完全涵蓋的是，本揭示案之其他實施例、圖 4 中描繪之操作及/或本文描述之其他操作均可以一方式組合，該組合方式並未明確展示於隨附圖式之任何圖式中，但仍完全與本揭示案一致。因此，針對並未確切展示於一圖式中的特徵及/或操作的請求項被視為屬於本揭示案之範疇及內容。

[0061]如本文中任何實施例所使用，「應用程式(app)」一詞可以代碼或指令體現，該等代碼或指令可在諸如主機處理器的可規劃電路或其他可規劃電路上執行。

[0062]如本文中任何實施例所使用，「模組」一詞可代表 app、軟體、韌體及/或電路，其組配來執行上述操作中之任何操作。軟體可體現為套裝軟體、記錄於至少一非暫時性電腦可讀儲存媒體上之代碼、指令、指令集及/或資料。韌體可體現為硬編碼(例如非依電性)於記憶體裝置中的代碼、指令或指令集及/或資料。

[0063]如本文中任何實施例所使用，「電路」可包含例如單獨的或呈任何組合的硬連線電路；可規劃電路，諸如

包含一或多個單獨指令處理核心之電腦處理器；狀態機電路及/或儲存藉由可規劃電路執行之指令的韌體。模組可共同地或單獨地體現為形成大型系統之部分的電路，例如積體電路(IC)、系統單晶片(SoC)、桌上型電腦、膝上型電腦、平板電腦、伺服器、智慧型電話等等。

[0064]如此所描述之任何操作可實行於包括一或多個儲存媒體之系統中，該等儲存媒體上儲存有單獨的或呈組合的指令，在藉由一或多個處理器執行該等指令時，該等指令執行該等方法。在此，處理器可包括例如伺服器 CPU、行動裝置 CPU 及/或其他可規劃電路。此外，本文描述之操作意欲可跨越多個實體裝置來分散，該等實體裝置諸如處在一個以上不同實體位置處的處理結構。儲存媒體可包括任何類型的有形媒體，例如，任何類型之碟片，包括硬碟、軟碟片、光碟、光碟片-唯讀記憶體(CD-ROM)、可重寫光碟片(CD-RW)及磁光碟；半導體裝置，諸如唯讀記憶體(ROM)、隨機存取記憶體(RAM)(諸如動態及靜態 RAM)、可抹除可規劃唯讀記憶體(EPROM)、電氣可抹除可規劃唯讀記憶體(EEPROM)、快閃記憶體、固態碟片(SSD)、磁性或光學卡；或者適合於儲存電子指令的任何類型之媒體。其他實施例可實行為藉由可規劃控制裝置執行之軟體模組。儲存媒體可為非暫時性的。

[0065]因此，本揭示案提供一種用於使化身交互地成動畫以替代活動影像來進行視訊通訊的方法及系統。與活動影像之發送相比，化身之使用減少要交換的資訊之量。該

系統及方法進一步組配來藉由例如音調偏移及/或使所擷取音訊信號時間延長而將使用者語音轉換成化身語音。化身之互動動畫可基於所偵測之使用者輸入，包括觸摸及手勢。互動動畫係組配來修改基於面部偵測及追蹤判定之動畫。

[0066]根據一態樣，提供一種系統。該系統可包括：使用者輸入裝置，其組配來擷取使用者輸入；通訊模組，其組配來傳輸及接收資訊；以及一或多個儲存媒體。此外，該一或多個儲存媒體上儲存有單獨的或呈組合的指令，在藉由一或多個處理器執行該等指令時產生以下操作，包含：選擇化身；起始通訊；偵測使用者輸入；識別使用者輸入；基於使用者輸入識別動畫命令；產生化身參數；以及傳輸動畫命令及化身參數中之至少一者。

[0067]另一示例系統包括前述組件且進一步包括：麥克風，其組配來擷取聲音且將所擷取之聲音轉化成相應音訊信號；以及指令，當藉由一或多個處理器執行該等指令時產生以下額外操作：擷取使用者語音且將使用者語音轉化成相應使用者語音信號；將使用者語音信號轉換成化身語音信號；以及傳輸化身語音信號。

[0068]另一示例系統包括前述組件且進一步包括：攝影機，其組配來擷取影像；以及指令，當藉由一或多個處理器執行該等指令時產生以下額外操作：擷取影像；偵測影像中的臉部；自臉部提取特徵；以及將特徵轉化成化身參數。

[0069]另一示例系統包括前述組件且進一步包括：顯示器；以及指令，當藉由一或多個處理器執行該等指令時產生以下額外操作：顯示至少一化身；接收遠程動畫命令及遠程化身參數中之至少一者；以及基於遠程動畫命令及遠程化身參數中之至少一者使一所顯示化身成動畫。

[0070]另一示例系統包括前述組件且進一步包括：揚聲器，其組配來將音訊信號轉換成聲音；以及指令，當藉由一或多個處理器執行該等指令時產生以下額外操作：接收遠程化身語音信號；以及將遠程化身語音信號轉化成化身語音。

[0071]另一示例系統包括前述組件，且該使用者輸入裝置為組配來擷取距離之攝影機且該使用者輸入為手勢。

[0072]另一示例系統包括前述組件，且該使用者輸入裝置為觸摸感應顯示器且該使用者輸入為觸摸事件。

[0073]另一示例系統包括前述組件，且該轉換包含音調偏移及時間延長中之至少一者。

[0074]根據另一態樣，提供一種方法。該方法可包括選擇化身；起始通訊；偵測使用者輸入；識別使用者輸入；基於使用者輸入識別動畫命令；基於動畫命令產生化身參數；及傳輸動畫命令及化身參數中之至少一者。

[0075]另一示例方法包括前述操作且進一步包括：擷取使用者語音且將使用者語音轉化成相應使用者語音信號；將使用者語音信號轉換成化身語音信號；以及傳輸化身語音信號。

[0076]另一示例方法包括前述操作且進一步包括：擷取影像；偵測影像中的臉部；自臉部提取特徵；以及將特徵轉化成化身參數。

[0077]另一示例方法包括前述操作且進一步包括：顯示至少一化身；接收遠程動畫命令及遠程化身參數中之至少一者；以及基於遠程動畫命令及遠程化身參數中之至少一者使一所顯示化身成動畫。

[0078]另一示例方法包括前述操作且進一步包括：接收遠程化身語音信號；以及將遠程化身語音信號轉化成化身語音。

[0079]另一示例方法包括前述操作且該使用者輸入為手勢。

[0080]另一示例方法包括前述操作且該使用者輸入為觸摸事件。

[0081]另一示例方法包括前述操作且該轉換包含音調偏移及時間延長中之至少一者。

根據另一態樣，提供一種系統。該系統可包括一或多個儲存媒體，該一或多個儲存媒體上儲存有單獨的或呈組合的指令，在藉由一或多個處理器執行該等指令時產生以下操作，包括選擇化身；起始通訊；偵測使用者輸入；識別使用者輸入；基於使用者輸入識別動畫命令；產生化身參數；以及傳輸動畫命令及化身參數中之至少一者。

[0082]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且亦包括：擷取使用者語音且將使用者語音轉化成相應使用者語音信號；將使用

者語音信號轉換成化身語音信號；以及傳輸化身語音信號。

[0083]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且亦包括：擷取影像；偵測影像中的臉部；自臉部提取特徵；以及將特徵轉化成化身參數。

[0084]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且亦包括：顯示至少一化身；接收遠程動畫命令及遠程化身參數中之至少一者；以及基於遠程動畫命令及遠程化身參數中之至少一者使一所顯示化身成動畫。

[0085]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且亦包括：接收遠程化身語音信號；以及將遠程化身語音信號轉化成化身語音。

[0086]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且該使用者輸入為手勢。

[0087]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且該使用者輸入為觸摸事件。

[0088]另一示例系統包括指令，當藉由一或多個處理器執行該等指令時產生前述操作，並且該轉換包含音調偏移及時間延長中之至少一者。

[0089]本文已使用之用詞及表述係用作描述之用詞且並非限制，且在使用此等用詞及表述時，不欲排除所展示及所描述的特徵之任何等效物(或其部分)，且應認識到，在

申請專利範圍之範疇內，可能存在各種修改。因此，申請專利範圍意欲涵蓋所有此類等效物。

### 【符號說明】

100...裝置至裝置系統/系統

102、112、102'...裝置/遠程裝置

104、114、104'/114'...攝影機

106、116、106'、116'...麥克風

107、117...揚聲器

108、118...觸摸感應顯示器/顯示器

108'、118'...顯示器

110、120...化身

112'...裝置/遠程裝置

122...網路

124、124'...伺服器

126...系統

128...虛擬空間

200...攝影機、音訊及觸控螢幕框架模組

202...面部偵測及追蹤模組/面部偵測/追蹤模組/臉部偵測模組

204...特徵提取模組

206...音訊轉換模組

208...觸摸偵測模組

210...手勢偵測模組

212...化身選擇模組

214...化身控制模組

- 216...系統
- 218...回饋化身
- 220...通訊模組
- 222...處理器
- 300、304...WiFi 連接
- 302...網際網路
- 306...企業 AP
- 308...閘道
- 310...防火牆
- 312...媒體及信號路徑
- 314...家 AP
- 400...流程圖
- 402~428...操作

## 申請專利範圍

1. 一種具有指令儲存在上的一或多個非暫時性電腦可讀儲存裝置的集合，該等指令在受一第一計算裝置的至少一處理器執行時，致使進行包含下列者的操作：
  - 致能一第一化身的選擇；
  - 識別出該第一計算裝置之一使用者的一或多個面部特徵；
  - 產生會被傳輸至一第二計算裝置的資訊，以使得被選擇的該第一化身在該第二計算裝置的一顯示器上出現動畫；其中，該資訊係以被識別出的該第一計算裝置的該使用者的該一或多個面部特徵為基礎；以及
  - 將該第一計算裝置之該使用者的聲學資訊轉換為目標聲學資訊，該目標聲學資訊會被傳輸至該第二計算裝置；其中，該轉換為使用一或多個嗓音效果以扭曲該第一計算裝置之該使用者的該聲學資訊。
2. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該一或多個面部特徵是從該第一計算裝置的該使用者之一或多個視訊影像所識別出。
3. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該等指令在受到該第一計算裝置的該至少一處理器執行時，致使進行包含下列者的額外操作：
  - 處理會被傳輸到該第二計算裝置的該第一計算裝置之使用者的音訊資訊。

4. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該等指令在受到該第一計算裝置的該至少一處理器執行時，致使進行包含下列者的額外操作：

致能一第二化身的選擇；

產生會被傳輸至該第二計算裝置的第二資訊，以使得被選擇的該第二化身在該第二計算裝置的該顯示器上出現動畫；其中，該第二資訊係以被識別出的該第一計算裝置之該使用者的該一或多個面部特徵為基礎；以及

致使被選擇的該第二化身顯示於該第一計算裝置上而使得該第一計算裝置的該使用者能夠觀察在該第二計算裝置上之被選擇的該第二化身的一外觀。

5. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該等指令在受到該第一計算裝置的該至少一處理器執行時，致使進行包含下列者的額外操作：

基於一使用者輸入命令而致能被選擇的該第一化身的動畫；其中，該使用者輸入命令是不同於該一或多個面部特徵，且該使用者輸入命令是由一使用者輸入裝置在受到該第一計算裝置的該使用者控制時所產生。

6. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該等指令在受到該第一計算裝置的該至少一處理器執行時，致使進行包含下列者的額外操作：

致使被選擇的該第一化身顯示於該第一計算裝置上而使得該第一計算裝置的該使用者能夠觀察在該第

二計算裝置上之被選擇的該第一化身的一外觀。

7. 如請求項 1 之一或多個非暫時性電腦可讀儲存裝置的集合，其中，該一或多個嗓音效果包含一音調平移嗓音效果。

8. 一第一計算裝置，包含：

可儲存指令與資料的記憶體電路；

可顯示一化身的一顯示裝置；

可處理一或多個指令以進行包含下列操作的處理器電路：

致能一第一化身的選擇；

識別出該第一計算裝置之一使用者的一或多個面部特徵；

產生會被傳輸至一第二計算裝置的資訊，以使得被選擇的該第一化身在該第一計算裝置的一顯示器上出現動畫；其中，該資訊係以被識別出的該第一計算裝置的該使用者的該一或多個面部特徵為基礎；並且

轉換該第一計算裝置之該使用者的聲學資訊為目標聲學資訊，該目標聲學資訊會被傳輸至該第二計算裝置；其中，該轉換為使用一或多個嗓音效果以扭曲該第一計算裝置之該使用者的該聲學資訊。

9. 如請求項 8 的第一計算裝置，進一步包含：

可擷取一或多個視訊影像的一視訊攝錄裝置，其中，該一或多個面部特徵是從該第一計算裝置之該使用者的一或多個視訊影像所識別出。

10. 如請求項 8 的第一計算裝置，進一步包含一音訊擷取裝置，該音訊擷取裝置可擷取會被傳輸至該第二計算裝置的該第一計算裝置之使用者的音訊資訊。
11. 如請求項 8 的第一計算裝置，其中，可處理一或多個指令的處理器可進行包含下列者的額外操作：
  - 致能一第二化身的選擇；
  - 產生會被傳輸至該第二計算裝置的第二資訊，以使得被選擇的該第二化身在該第一計算裝置的該顯示器上出現動畫；其中，該第二資訊係以被識別出的該第一計算裝置的該使用者的該一或多個面部特徵為基礎；並且
  - 致使被選擇的該第二化身顯示於該第一計算裝置上，而使得該第一計算裝置的該使用者能夠觀察在該第一計算裝置上之被選擇的該第二化身的一外觀。
12. 如請求項 8 的第一計算裝置，其中，可處理一或多個指令的處理器可進行包含下列者的額外操作：
  - 基於一使用者輸入命令而致能被選擇的該第一化身的動畫；其中，該使用者輸入命令是不同於一或多個面部特徵，且該使用者命令是由一使用者輸入裝置在受到該第一計算裝置的該使用者控制時所產生。
13. 如請求項 8 的第一計算裝置，其中，可處理一或多個指令的處理器可進行包含下列者的額外操作：
  - 致使被選擇的該第一化身顯示於該第一計算裝置上而使得該第一計算裝置的該使用者能夠觀察在該第一計

算裝置上之被選擇的該第一化身的一外觀。

14. 如請求項 8 的第一計算裝置，其中，該一或多個嗓音效果包含一音調平移嗓音效果

15. 一種方法，該方法包含下列步驟：

藉由一第一計算裝置而致能一第一化身的選擇；

藉由該第一計算裝置而識別出該第一計算裝置之一使用者之一或多個面部特徵；

藉由該第一計算裝置而產生會被傳輸至一第二計算裝置的資訊，以使得被選擇的該第一化身在該第二計算裝置的一顯示器上出現動畫；其中，該資訊係以被識別出的該第一計算裝置之該使用者的該一或多個面部特徵為基礎；並且

藉由該第一計算裝置而將該第一計算裝置之該使用者的聲學資訊轉換為目標聲學資訊，該目標聲學資訊會被傳輸至該第二計算裝置；其中，該轉換為使用一或多個嗓音效果以扭曲該第一計算裝置之該使用者的該聲學資訊。

16. 如請求項第 15 項的方法，其中，該一或多個面部特徵是從該第一計算裝置的該使用者之一或多個視訊影像所識別出。

17. 如請求項第 15 項的方法，進一步包含下列步驟：

藉由該第一計算裝置而處理會被傳輸到該第二計算裝置的該第一計算裝置之使用者的音訊資訊。

18. 如請求項第 15 項的方法，進一步包含下列步驟：

藉由該第一計算裝置而致能一第二化身的選擇；

藉由該第一計算裝置而產生會被傳輸至該第二計算裝置的第二資訊，以使得被選擇的該第二化身在該第二計算裝置的該顯示器上出現動畫；其中，該第二資訊係以被識別出的該第一計算裝置之該使用者的該一或多個面部特徵為基礎；並且

藉由該第一計算裝置而致使被選擇的該第二化身顯示於該第一計算裝置上，而使得該第一計算裝置的該使用者能夠觀察在該第二計算裝置上之被選擇的該第二化身的一外觀。

19. 如請求項第 15 項的方法，進一步包含下列步驟：

藉由該第一計算裝置基於一使用者輸入命令而致能被選擇的該第一化身的動畫；其中，該使用者輸入命令是不同於該一或多個面部特徵，且該使用者輸入命令是由一使用者輸入裝置在受到該第一計算裝置的該使用者控制時所來產生。

20. 如請求項第 15 項的方法，進一步包含下列步驟：

藉由該第一計算裝置而致使被選擇的該第一化身顯示於該第一計算裝置上，而使得該第一計算裝置的該使用者能夠觀察在該第二計算裝置上之被選擇的該第一化身的一外觀。

21. 如請求項第 15 項的方法，其中，該一或多個嗓音效果包含一音調平移嗓音效果。

22. 一第一計算裝置，包含：

一化身選擇模組，其可致能一第一化身的選擇；

一特徵提取模組，其可識別出該第一計算裝置之一使用者之一或多個面部特徵；

一化身控制模組，其可進行下列動作：

產生會被傳輸至一第二計算裝置的資訊，以使得被選擇的該第一化身在該第一計算裝置的一顯示器上出現動畫；其中，該資訊係以被識別出的該第一計算裝置之該使用者的該一或多個面部特徵為基礎；以及

一音訊轉換模組，其可將該第一計算裝置之該使用者的聲學資訊轉換為目標聲學資訊，該目標聲學資訊會被傳輸至該第二計算裝置；其中，該轉換為使用一或多個嗓音效果以扭曲該第一計算裝置之該使用者的該聲學資訊。

23. 如請求項 22 的第一計算裝置，進一步包含一面部偵測與追蹤模組，該面部偵測與追蹤模組可偵測與追蹤該第一計算裝置之該使用者的一面部。
24. 如請求項 22 的第一計算裝置，進一步包含一音訊擷取模組，該音訊擷取模組可擷取會被傳輸至該第二計算裝置的該第一計算裝置之使用者的音訊資訊。
25. 如請求項 22 的第一計算裝置，其中，該化身選擇模組可進一步致能一第二化身的選擇。
26. 如請求項 25 的第一計算裝置，其中，該化身選擇模組可進一步產生會被傳輸至該第二計算裝置的第二資訊，以

使得被選擇的該第二化身在該第一計算裝置的該顯示器上出現動畫；其中，該第二資訊係以被識別出的該第一計算裝置之該使用者的該一或多個面部特徵為基礎。

27. 如請求項 26 的第一計算裝置，進一步包含一顯示模組，該顯示模組可使被選擇的該第二化身顯示於該第一計算裝置上，而使得該第一計算裝置的該使用者能夠觀察在該第一計算裝置上之被選擇的該第二化身的一外觀。
28. 如請求項 22 的第一計算裝置，其中，該化身控制模組進一步可基於一使用者輸入命令而致能被選擇的該第一化身的動畫；其中，該使用者輸入命令為獨立於該一或多個面部特徵，且該使用者輸入命令是由一使用者輸入裝置在受到該第一計算裝置的該使用者控制時所產生。
29. 如請求項 22 的第一計算裝置，進一步包含一顯示模組，該顯示模組可使被選擇的該第一化身顯示於該第一計算裝置上，而使得該第一計算裝置的該使用者能夠觀察在該第一計算裝置上之被選擇的該第一化身的一外觀。
30. 如請求項 22 的第一計算裝置，其中，該一或多個嗓音效果包含一音調平移嗓音效果。
31. 如請求項 22 的第一計算裝置，進一步包含一視訊擷取裝置，該視訊擷取裝置可擷取該第一計算裝置的該使用者的一或多個視訊影像；其中，該一或多個面部特徵是從該計算裝置的該使用者之一或多個視訊影像所識別出。

圖式

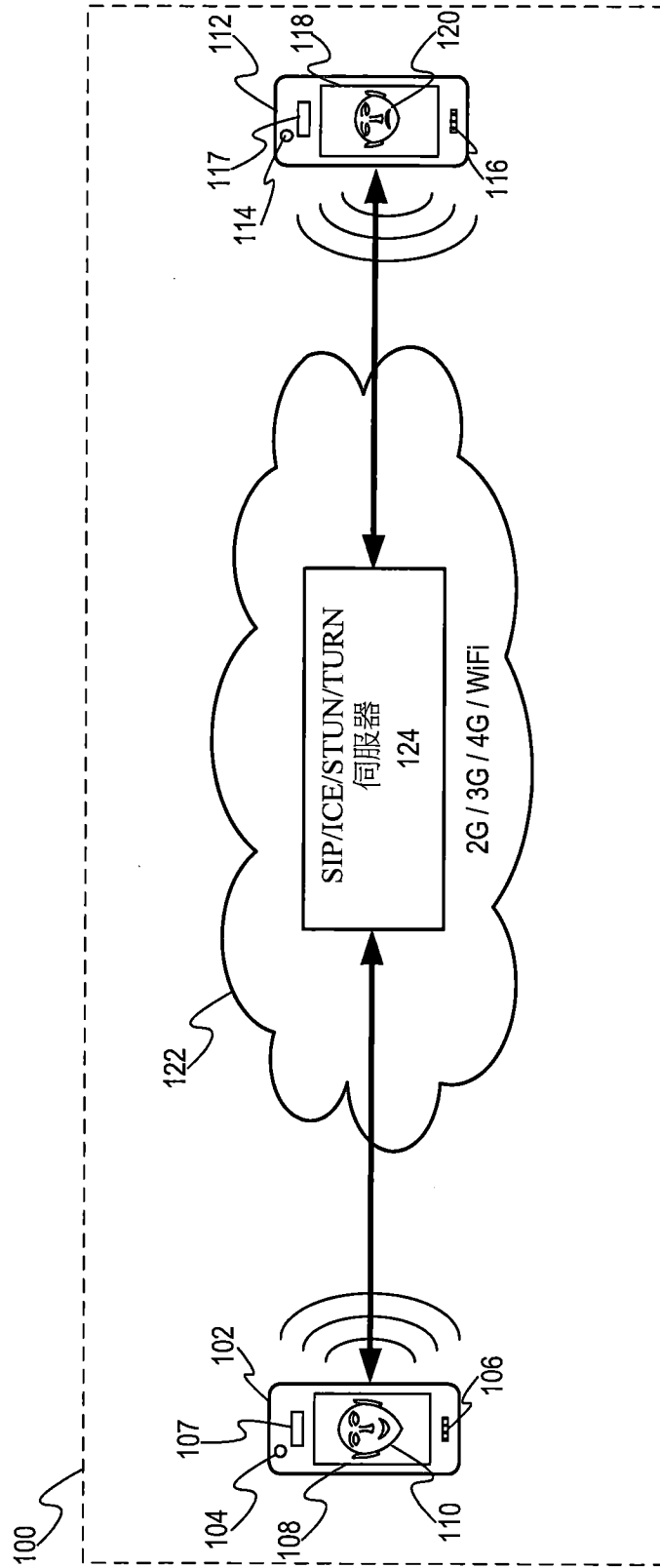


圖 1A

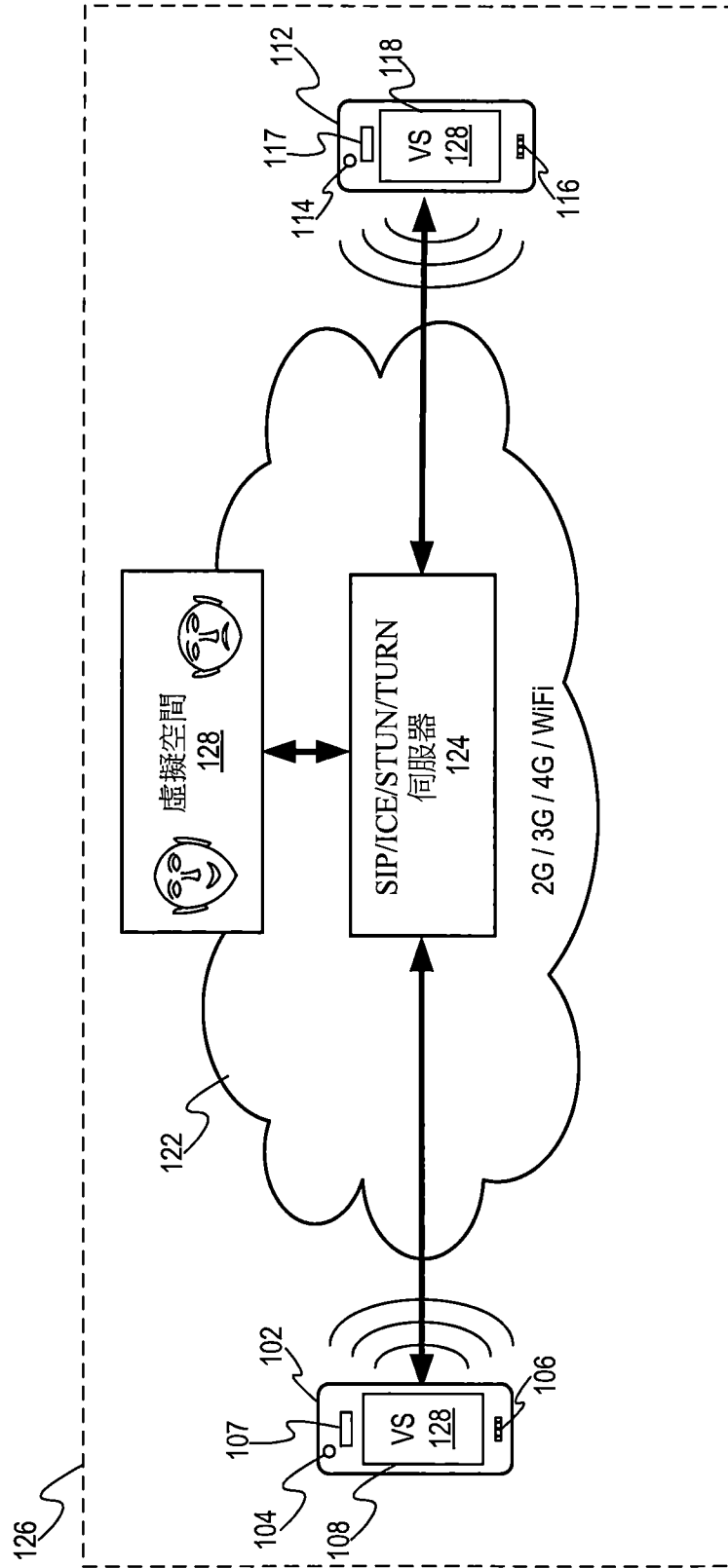


圖 1B

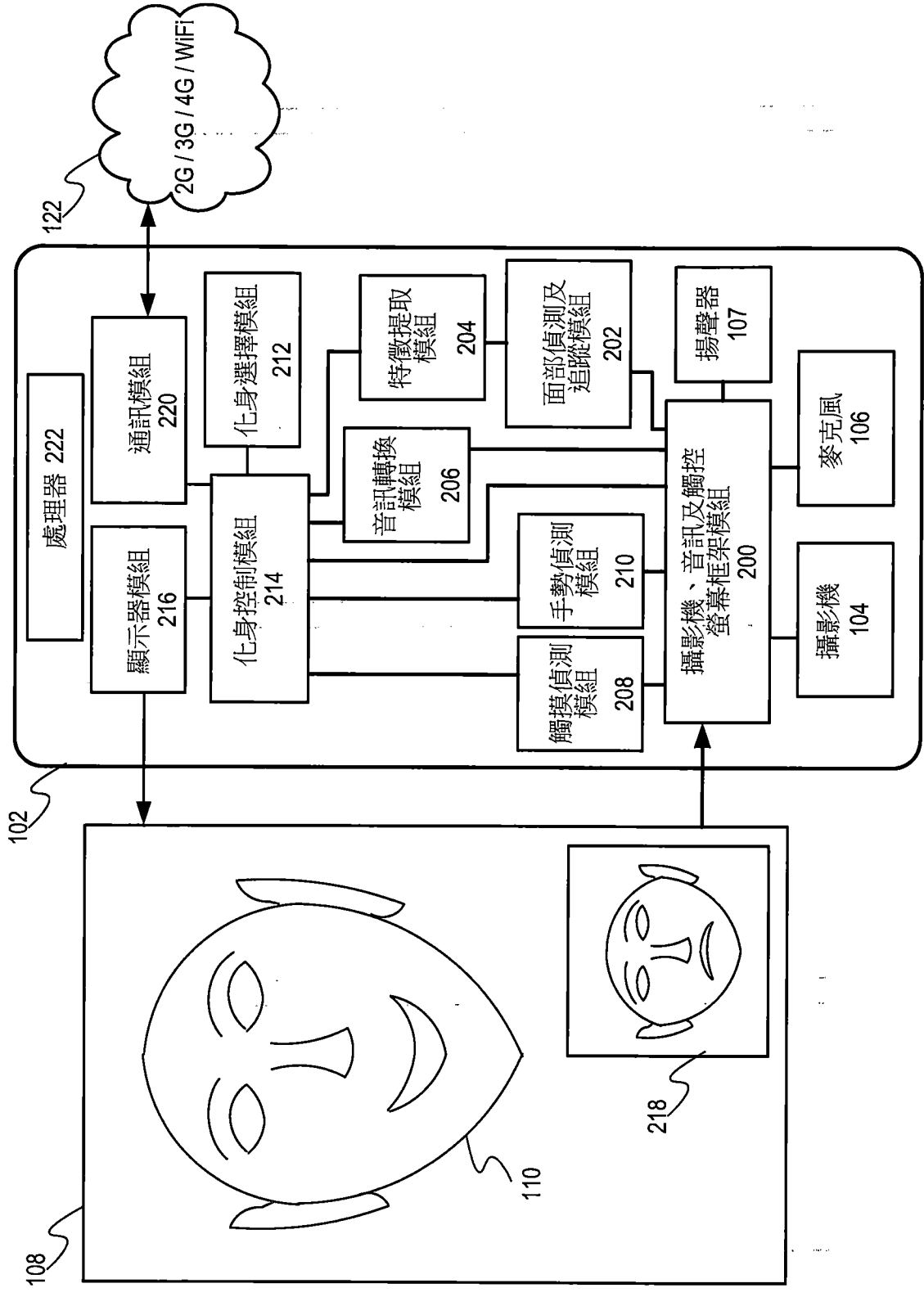


圖 2

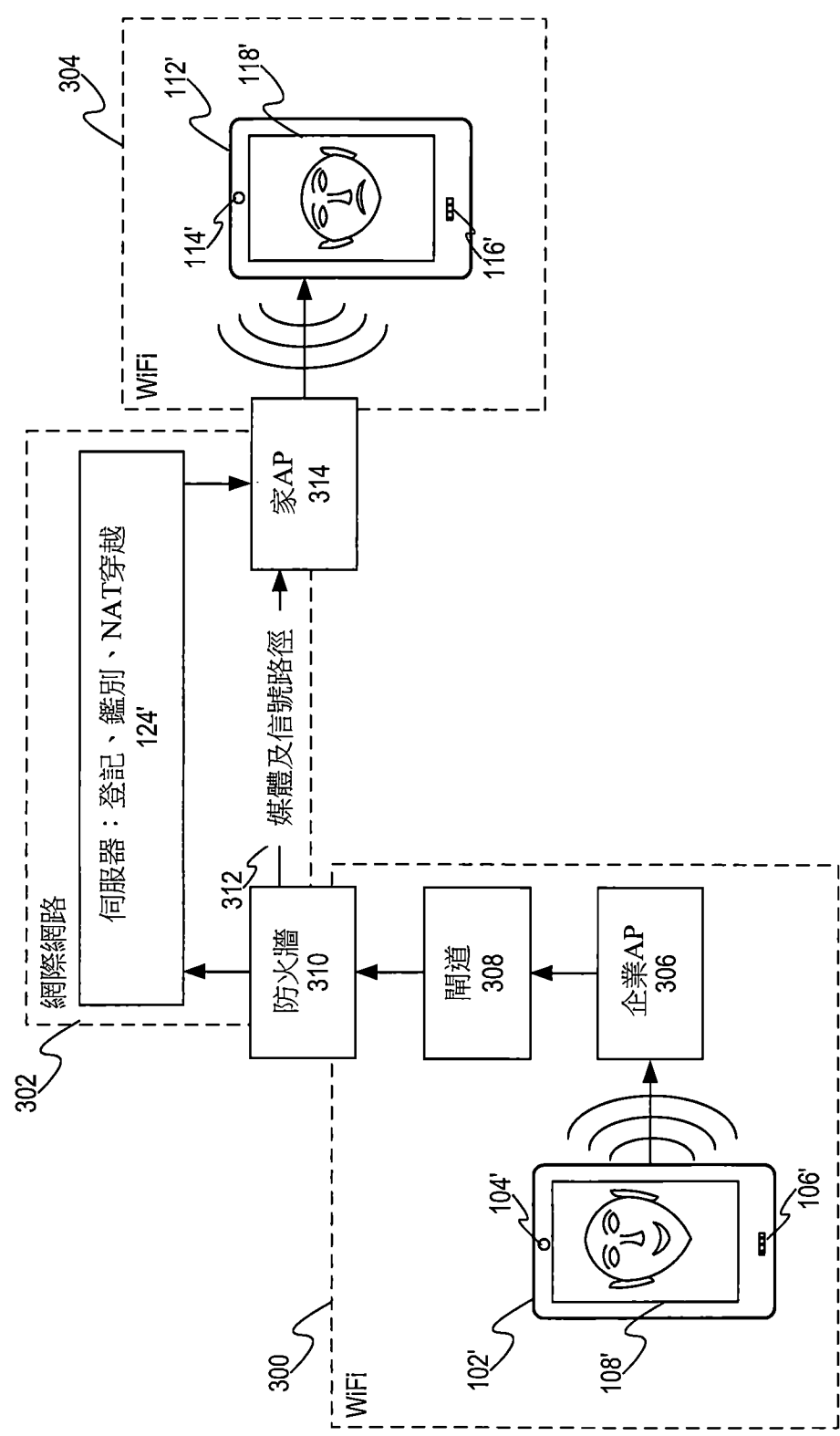


圖 3

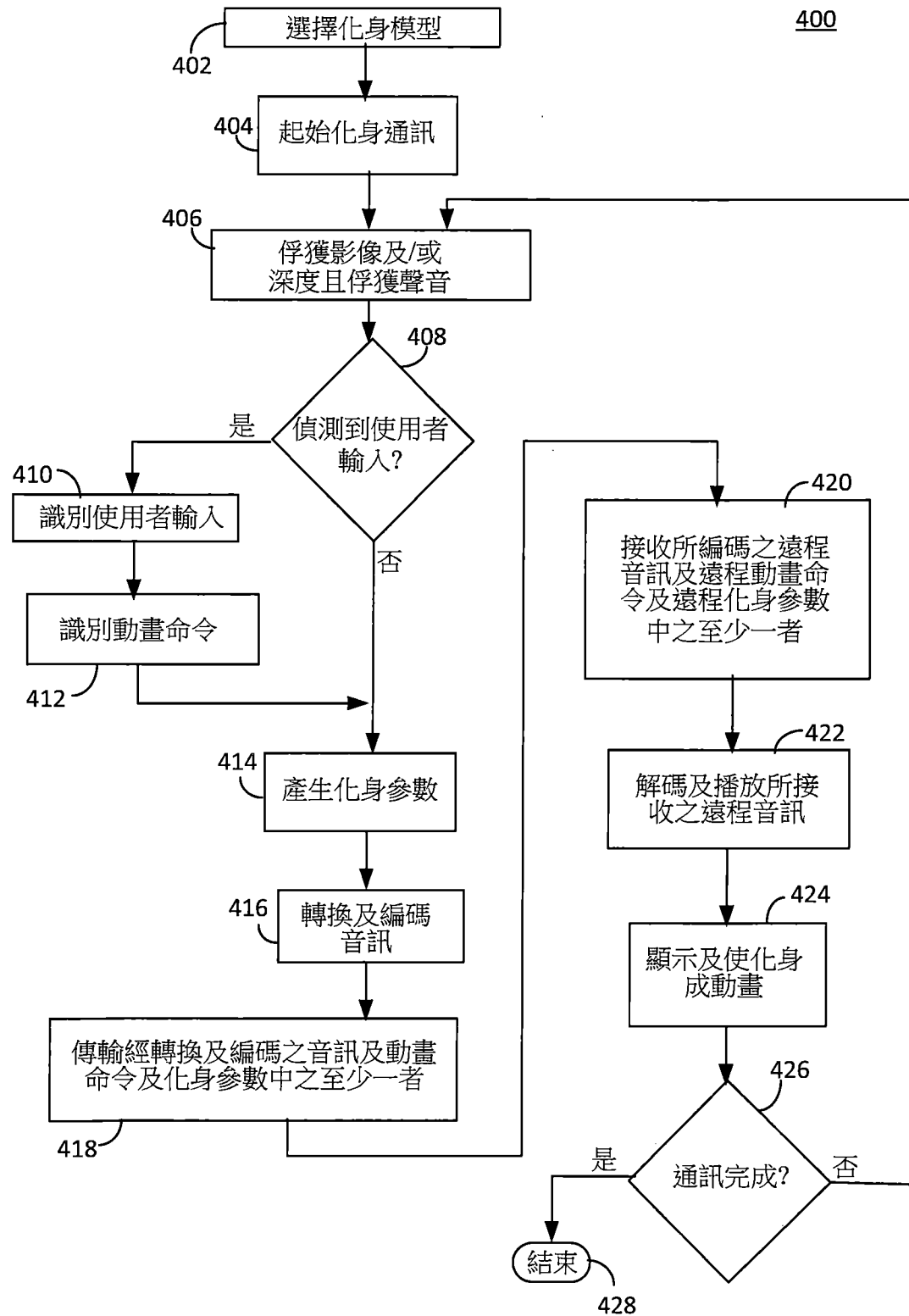


圖 4