



(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 I779113 B

(45)公告日：中華民國 111 (2022) 年 10 月 01 日

(21)申請案號：107136063

(22)申請日：中華民國 107 (2018) 年 10 月 12 日

(51)Int. Cl. : G10L25/51 (2013.01)

G10L15/08 (2006.01)

G10L17/22 (2013.01)

(30)優先權：2017/10/12 美國

15/782,287

(71)申請人：美商高通公司(美國) QUALCOMM INCORPORATED (US)

美國

(72)發明人：賈維斯 莫瑞 JARVIS, MURRAY (GB)；塔爾羅 班傑明 TARLOW, BENJAMIN

(GB)；格瑞貝 尼可拉斯 GRAUBE, NICOLAS (FR)；伍爾史特姆 克拉克

WOOLSTENHULME, CLARK (US)；芬奇 賽蒙 FINCH, SIMON (GB)

(74)代理人：林怡芳

(56)參考文獻：

US 9661473B1

US 2017/0154638A1

WO 2017/088026A1

期刊 Mostafa M. A. Shaikh, et al., "Automatic Life-Logging: A novel approach to sense real-world activities by environmental sound cues and common sense", COMPUTER AND INFORMATION TECHNOLOGY, 2008. ICCIT 2008. 11TH INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, (20081224), ISBN 978-1-4244-2135-0, pages 294 - 299, https://www.ubicc.org/files/pdf/548-2011-v3_546.pdf

審查人員：莊榮昌

申請專利範圍項數：28 項 圖式數：11 共 52 頁

(54)名稱

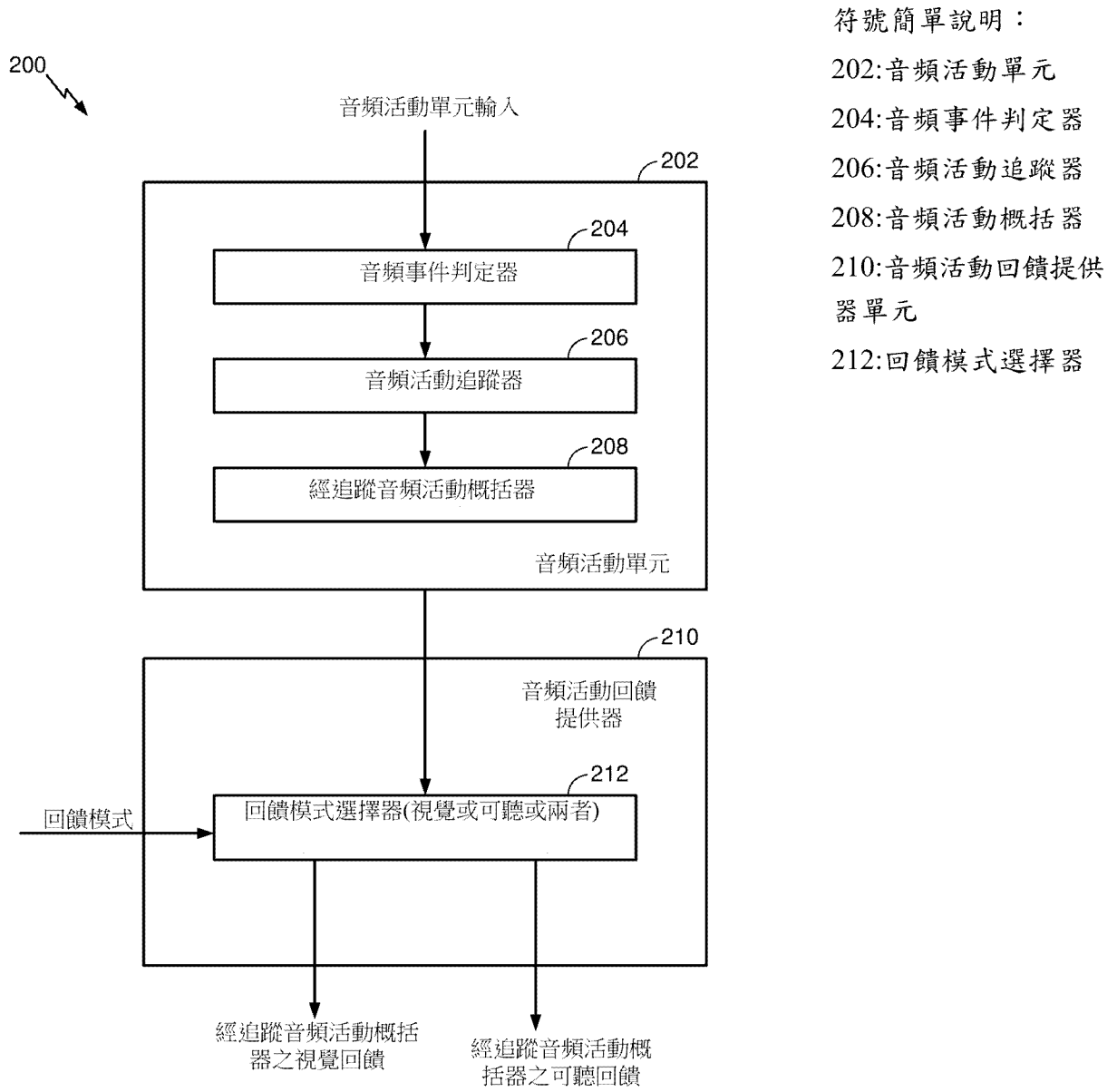
音頻活動追蹤及概括的裝置、方法、設備以及電腦可讀儲存媒體

(57)摘要

各種實施例提供揭示一種裝置之系統及方法，該裝置可用於基於接收一音頻信號判定一音頻事件。基於該經判定音頻事件，可追蹤音頻活動。可基於一概括模式概括該經追蹤音頻活動。該概括之經追蹤音頻活動之回饋可基於一回饋模式。

Various embodiments provide systems and methods which disclose a device which may be used to determine an audio event based on receiving an audio signal. Based on the determined audio event, audio activity may be tracked. The tracked audio activity may be summarized based on a summary mode. Feedback of the summarized tracked audio activity may be based on a feedback mode.

指定代表圖：



【圖2】



I779113

【發明摘要】

公告本

【中文發明名稱】

音頻活動追蹤及概括的裝置、方法、設備以及電腦可讀儲存媒體

【英文發明名稱】

DEVICE, METHOD, APPARATUS AND COMPUTER-READABLE STORAGE MEDIUM FOR AUDIO ACTIVITY TRACKING AND SUMMARIES

【中文】

各種實施例提供揭示一種裝置之系統及方法，該裝置可用於基於接收一音頻信號判定一音頻事件。基於該經判定音頻事件，可追蹤音頻活動。可基於一概括模式概括該經追蹤音頻活動。該概括之經追蹤音頻活動之回饋可基於一回饋模式。

【英文】

Various embodiments provide systems and methods which disclose a device which may be used to determine an audio event based on receiving an audio signal. Based on the determined audio event, audio activity may be tracked. The tracked audio activity may be summarized based on a summary mode. Feedback of the summarized tracked audio activity may be based on a feedback mode.

【指定代表圖】

圖2

【代表圖之符號簡單說明】

202 音頻活動單元

- 204 音頻事件判定器
- 206 音頻活動追蹤器
- 208 音頻活動概括器
- 210 音頻活動回饋提供器單元
- 212 回饋模式選擇器

【發明說明書】

【中文發明名稱】

音頻活動追蹤及概括的裝置、方法、設備以及電腦可讀儲存媒體

【英文發明名稱】

DEVICE, METHOD, APPARATUS AND COMPUTER-READABLE STORAGE MEDIUM FOR AUDIO ACTIVITY TRACKING AND SUMMARIES

【技術領域】

【0001】 本發明大體上係關於音頻處理及使用者介面。

【先前技術】

【0002】 技術進步已導致健身帶或活動追蹤器正變得被更多消費者使用來追蹤其身體活動(例如，睡眠追蹤器(fitbit)裝置或可穿戴式手錶)。此等健身及活動追蹤器通常被穿戴且包括用以量測人的心跳或量測運動的感測器，且耦接至可穿戴式裝置內的收發器。在許多情況下，可穿戴式裝置經由藍牙無線通信將健身或活動(例如，步數、燃燒的卡路里或站立的時間量)無線地傳輸至智慧型手機或其他行動裝置。

【0003】 在過去幾年裏，術語「可聽設備(hearables)」已用於描述可穿戴式裝置及頭戴式耳機之使用。可聽設備亦包括用以播放聲音的揚聲器及用以擷取話音命令或在電話上通話的麥克風。可聽設備可包括追蹤心跳速率、溫度、氧飽和度或步調的感測器。另外，可聽設備可包括用以與智慧型手機或行動裝置通信的藍牙積體電路。一些可聽設備能夠提供允許使用者撥打電話而無需與手機進行無線配對的Wi-Fi及蜂窩連接。語音辨識之進步亦隨機器學習之進步在增加且為可包括於可聽設備之額外技術。

【0004】 隨著可聽設備之技術的持續發展，新的挑戰及機遇可能出現。

【發明內容】

【0005】 各種實施例包括用於包括一或多個處理器之裝置基於接收音頻信號判定音頻事件之方法。該一或多個處理器基於經判定音頻事件追蹤音頻活動且基於概括模式概括經追蹤音頻活動。此外，該一或多個處理器基於回饋模式提供概括之經追蹤音頻活動之回饋。另外，該一或多個處理器可選擇回饋模式，且基於視覺回饋模式顯示概括之經追蹤音頻活動。在相同或替代實施例中，當所選擇回饋模式為可聽回饋模式時，概括之經追蹤音頻活動可為可聽的。

【0006】 存在各種實施例，該等實施例包括一種基於接收音頻信號判定音頻事件的方法。該方法包括基於經判定音頻事件追蹤活動且基於概括模式概括經追蹤音頻活動。此外，該方法包括基於回饋模式提供概括之經追蹤音頻活動之回饋。另外，在一實施例中，該方法包括選擇回饋模式及基於視覺回饋模式顯示概括之經追蹤音頻活動。在相同或替代實施例中，該方法包括概括經追蹤音頻活動，且當所選擇回饋模式為可聽回饋模式時，經追蹤音頻活動可為可聽的。

【0007】 存在包括一種設備的各種實施例，該設備包括用於基於接收音頻信號判定音頻事件之構件及用於基於經判定音頻事件追蹤音頻活動之構件。此外，該設備包括用於基於概括模式概括經追蹤音頻活動之構件及用於基於回饋模式提供概括之經追蹤音頻活動之回饋之構件。

【0008】 存在各種實施例，該等實施例包括其上儲存有指令的非暫時性電腦可讀儲存媒體，該等指令在經執行時使得裝置之一或多個處理器

基於接收音頻信號判定音頻事件且基於接收音頻信號判定音頻事件。該等指令在經執行時亦使得該一或多個處理器基於經判定音頻事件追蹤音頻活動，基於概括模式概括經追蹤音頻活動且基於回饋模式提供概括之經追蹤音頻活動之回饋。

【圖式簡單說明】

【0009】 併入本文中且構成本說明書之一部分的隨附圖式說明申請專利範圍之例示性實施例，且與上文給出的一般描述及下文給出的詳細描述一起用來解釋申請專利範圍之特徵。

【0010】 圖1為說明適合於實施本文中所揭示之各種實施例的裝置之實例之組件方塊圖。

【0011】 圖2說明包括音頻活動單元及音頻活動回饋提供器的方塊圖。

【0012】 圖3說明音頻事件判定器之多個實施例。

【0013】 圖4說明音頻活動追蹤器之例示性實施。

【0014】 圖5說明追蹤活動概括器之例示性實施例。

【0015】 圖6說明可為封包之部分的不同欄位。

【0016】 圖7說明不同類型之個人識別。

【0017】 圖8說明分批處理音頻事件之一實施例。

【0018】 圖9說明顯示器上每日音頻活動之概括之例示性視圖。

【0019】 圖10說明與顯示器上的概括之經追蹤音頻活動相關聯之情緒的鏈接之另一例示性視圖。

【0020】 圖11說明例示性使用案例。

【0021】 將參考隨附圖式詳細地描述各種實施例。在任何可能之

處，將貫穿圖式使用相同附圖標號來指代相同或相似部件。對特定實例及實施之參考為出於說明之目的，且並不意欲限制申請專利範圍之範疇。

【實施方式】

【0022】 本申請案主張2017年10月12日申請之美國專利申請案第15/782,287號之權益，該專利申請案之全部內容以引用之方式併入本文中。

【0023】 下文參考圖式描述本發明之特定實施。在描述中，貫穿圖式藉由共同附圖標號指示共同特徵。如本文中所使用，僅出於描述特定實施之目的而使用各種術語，且該術語並不意欲為限制性的。舉例而言，除非上下文另外明確指示，否則單數形式「一(a/an)」及「該」意欲同樣包括複數形式。可進一步理解，術語「包含(comprise/comprises/comprising)」可與「包括(include/includes/including)」互換地使用。另外，應理解，術語「其中(wherein)」可與「在...的情況下(when)」互換地使用。如本文中所使用，「例示性」可指示實例、實施及/或態樣，且不應被視作限制或視為指示偏好或較佳實施。如本文中所使用，用以修飾元件(諸如，結構、組件、操作等)之序數術語(例如，「第一」、「第二」、「第三」等)本身不指示元件關於另一元件之任何優先權或次序，而是僅將元件與具有相同名稱之另一元件區別開(除非使用序數術語)。如本文中所使用，術語「集合」係指一或多個元件之群組，且術語「複數個」係指多個元件。

【0024】 如本文中所使用，「耦接」可包括「通信耦接」、「電性耦接」或「實體耦接」，且亦可(或替代地)包括其任何組合。兩個裝置(或組件)可經由一或多個其他裝置、組件、線、匯流排、網路(例如，有線網

路、無線網路或其組合)等直接地或間接地耦接(例如，通信耦接、電性耦接或實體耦接)。作為說明性非限制性實例，電性耦接之兩個裝置(或組件)可包括於相同裝置或不同裝置中，且可經由電子件、一或多個連接器或電感性耦接連接。在一些實施中，諸如在電子通信中通信耦接之兩個裝置(或組件)可諸如經由一或多個導線、匯流排、網路等直接地或間接地發送及接收電信號(數位信號或類比信號)。如本文中所使用，「直接耦接」可包括在無介入組件之情況下耦接(例如，通信耦接、電性耦接或實體耦接)兩個裝置。

【0025】 如本文中所使用，「整合」可包括「與...一起製造或出售」。若使用者購買捆綁或包括作為封裝之部分之裝置的封裝，則該裝置可為整合式的。在一些描述中，兩個裝置可耦接，但未必係整合式的(例如，不同周邊裝置可不整合至一裝置中，但仍可「耦接」)。另一實例可為本文中所描述之可「耦接」至處理器的收發器或天線中之任一者，但未必為包括裝置之封裝之部分。當使用術語「整合」時，可自本文中所揭示之上下文(包括此段)推斷出其他實例。

【0026】 如本文中所使用，「提供」可包括「發送」，例如一或多個處理器可將輸出信號或信號提供至另一組件，可等效地意謂，該一或多個處理器可將該輸出信號或信號發送至另一組件。

【0027】 如本文中所使用，裝置之間的「無線連接」可基於各種無線技術，諸如藍牙、無線保真(Wi-Fi)或Wi-Fi之變體(例如，Wi-Fi直接)。裝置可基於不同蜂窩通信系統「無線連接」，諸如長期演進(LTE)系統、分碼多重存取(CDMA)系統、全球行動通信系統(GSM)系統、無線區域網路(WLAN)系統或某一其他無線系統。CDMA系統可實施寬頻CDMA

(WCDMA)、CDMA 1X、演進資料最佳化(EVDO)、分時同步CDMA (TD-SCDMA)或某一其他版本之CDMA。另外，當兩個裝置在視線內時，「無線連接」亦可基於諸如超音波、紅外線、脈衝射頻電磁能量、結構化光或信號處理中所使用之到達定向技術(例如，音頻信號處理或射頻處理)之其他無線技術。

【0028】 術語「計算裝置」在本文中一般用以指代伺服器、個人電腦、膝上型電腦、平板電腦、行動裝置、蜂巢式電話、智慧筆記型電腦、超級本、掌上型電腦、個人資料助理(PDA)、無線電子郵件接收器、具多媒體網際網路功能之蜂巢式電話、全球定位系統(GPS)接收器、無線遊戲控制器及包括用於無線發送及/或接收資訊的可程式化處理器及電路中之任一者或所有。

【0029】 在一或多個麥克風已擷取或記錄撞擊在一或多個麥克風上的聲波之後，術語「音頻信號」可在不同階段處互換地使用。另外，術語音頻信號亦可在自另一裝置接收封包之不同階段處互換地使用。

【0030】 術語「行動裝置」、「連接裝置」、「無線裝置」及「無線節點」在本文中可互換地用以指代包括用於無線發送及/或接收資訊之電路的任何電子裝置，包括蜂巢式電話、個人或行動多媒體播放機、手錶、腕部顯示器、醫學裝置、耳機、頭戴式耳機、揚聲器、麥克風、計算裝置及其他類似電子件。

【0031】 如本文中所使用，A「及/或」B可意謂「A及B」或「A或B」，或該兩者「A及B」及「A或B」均為可適用或可接受的。

【0032】 圖1為說明裝置100之例示性組件之方塊圖。裝置100可為可聽裝置或可為不必包括人耳內或周圍之揚聲器的另一裝置。舉例而言，

另一裝置可為包括麥克風之家用揚聲器，或智慧型手錶。裝置100亦可包括諸如一或多個感測器、一或多個攝影機及/或亦併入本文中所揭示之實施例中之一些之顯示裝置之例示性組件。

【0033】 在特定實施中，裝置100包括處理器128，該處理器包括：中央處理單元(CPU)；或數位處理器(DSP)；或耦接至記憶體126之圖形處理單元(GPU)。記憶體126包括指令130 (例如，可執行指令)，諸如電腦可讀指令或處理器可讀指令。指令130可包括可由諸如處理器128之電腦執行的一或多個指令。

【0034】 圖1亦說明耦接至處理器128且耦接至裝置100中之一或多個組件的一或多個控制器120。舉例而言，一或多個控制器120可耦接至各種周邊裝置(例如，顯示裝置110、感測器108、攝影機106及麥克風102)。如本文中所描述，顯示裝置110可不與裝置100整合，或在與描述相關聯之實施中，與本文中所揭示之裝置100整合。

【0035】 音頻/語音編碼器/解碼器(CODEC) 114亦可耦接至處理器128及記憶體126。麥克風102可耦接至音頻/語音編解碼器。音頻/語音編解碼器輸出端可為音頻/語音緩衝器136，且可儲存由麥克風102擷取之音頻及/或語音的樣本。樣本可為音頻波形之部分。一般熟習此項技術者可將音頻/語音樣本稱作音頻或語音樣本或兩者。呈語音之時間頻率的音頻樣本可被稱作語音樣本。視訊編碼器/解碼器(CODEC) (未繪示)亦可耦接至處理器128及記憶體126。攝影機106可耦接至視訊編解碼器。視訊編解碼器輸出端可為像素緩衝器131，且可儲存由攝影機106擷取之視訊之像素。由攝影機106擷取之視訊之像素可包括靠近裝置100之環境或場景之影像或影像之一部分。可替代地或另外，由攝影機106擷取之視訊之像素

可包括使用者之影像或影像之部分。一或多個感測器108可耦接至處理器128及記憶體126。感測器108之輸出端可為感測器緩衝器132，且可儲存本文中所描述之由感測器108擷取之位置資料之樣本。

【0036】 裝置100可包括收發器122，且收發器122可耦接至處理器128且耦接至天線112，以使得可將經由天線112及收發器122接收的無線資料提供至處理器128。裝置100亦可包括實施通信協定之網路介面124。在裝置100之一些實施中，網路介面124可直接整合至處理器128中。在一些實施中，處理器128、控制器120、記憶體126及收發器122可包括於系統封裝或系統單晶片裝置中。在一些實施中，各種周邊裝置可整合至裝置100中。

【0037】 此外，在特定實施中，如圖1中所說明，顯示裝置110、感測器108、攝影機106及麥克風102、天線112在系統單晶片裝置外部。在特定實施中，顯示裝置110、感測器108、攝影機106、麥克風102及天線112中之每一者可耦接至諸如一或多個控制器120或記憶體126之系統單晶片裝置的組件。

【0038】 作為說明性非限制性實例，裝置100可包括耳機、智慧型眼鏡或擴增實境裝置、虛擬實境裝置、混合實境裝置、擴展實境裝置、機器人、行動通信裝置、智慧型手機、蜂巢式電話、膝上型電腦、電腦、智慧型眼鏡、平板電腦、個人數位助理、電視、遊戲控制台、音樂播放機、無線電台、數位視訊播放機、作為車輛中之音頻系統之部分、數位視訊光碟(DVD)播放機、獨立攝影機、導航裝置、車輛、車輛之組件或其任何組合。

【0039】 在說明性實施中，記憶體126可包括或對應於儲存指令130

之非暫時性電腦可讀媒體。指令130可包括可由諸如處理器128之電腦執行的一或多個指令。指令130可使處理器128執行本文中所描述之一或多個操作，包括(但不限於)與圖1至圖10中之任一者相關聯之描述的一或多個部分。

【0040】 裝置100可包括音頻活動單元116、音頻活動回饋提供器118、音頻事件緩衝器138、概括經追蹤音頻活動緩衝器144、音頻/語音封包緩衝器142。緩衝器、音頻活動單元116及音頻活動回饋提供器118在本發明內以各種更詳細的方式描述。

【0041】 圖2說明包括音頻活動單元及音頻活動回饋提供器的方塊圖。隨著處理器技術持續按比例縮小至較小節點大小，設想諸如Amazon Echo或Google Home之數位助理的功能亦可併入至可穿戴式裝置中。穿戴裝置之一個優點為能夠跨多個環境(例如，汽車、家、辦公室)追蹤使用者之社交互動。

【0042】 作為實例，穿戴裝置(諸如可聽設備、手錶或智慧型眼鏡)之另一優點為此等裝置之麥克風可不像有時智慧型手機在使用者之一天裏在口袋或錢包中一樣。在一些實施例中，在記錄或擷取音頻信號與穿戴者之代表性體驗之間可存在較好相關性。

【0043】 舉例而言，在一些情況下，可穿戴設備可比在交替置放中更穩定，且可導致與真實穿戴者體驗更局部相關。另外，對於可穿戴設備(例如，可穿戴式感測器)，將感測器置放於身體之不同點處可存在優點，以便獲得對公共實體事件之不同類型的量測。將感測器置放於不同位置處可導致不同挑戰及限制，然而，此類置放之優點在於身體運動(例如，足部感測器、腕部感測器、皮帶周圍之感測器或可聽設備)可增加音頻事件

偵測之信賴度及/或可導致音頻事件偵測。因此，在一些實施例中，其可具有優於手機之優點，例如，其可潛在地漂浮且相對於使用者之運動或使用者之運動的至少部分(例如，手臂、足部等)微弱、流動且未判定的。

【0044】 非可穿戴式裝置亦可併入本文中所示之實施例中之一或多者。亦設想包括音頻活動單元202及/或音頻活動回饋提供器單元210之裝置將具有由無線連接裝置之間的資訊傳送、較大記憶體大小及/或機器學習進步而引起之提高的技術能力。舉例而言，允許本端擷取但可能需要較多分析處理之感測器可獲益於連接及遠端處理。因此，除了社交互動之外，追蹤其他音頻活動之能力旨在擷取整體音頻體驗且將其轉換為活動，且概括除社交互動之外的體驗。另外，可基於不同類型之概括模式提供對音頻活動之概括。在圖5中論述概括模式。

【0045】 除社交互動(亦即，社交互動涉及來回討論的至少兩個人)之外的音頻活動之實例可包括收聽電視或音樂所花費的時間。其他實例可與對情緒(例如，憤怒、高興、中性等)進行分類或基於音頻信號電平(例如，喊叫或私語)對交談(或收聽)所花費之時間進行分類相關。又，其他實例包括收聽不同類型之人(例如，兒童對比成人、或男性對比女性)所花費的時間。基於音頻活動，音頻回饋提供器210可將視覺或可聽回饋提供至使用者。

【0046】 在實施例中，裝置(諸如裝置100)包括音頻活動單元202，該音頻活動單元包括經組態以基於接收音頻信號判定音頻事件之音頻事件判定器204 (參見圖3)。音頻事件基於發出聲音的物體或人而產生。音頻事件之實例已揭示於2015年10月多媒體之IEEE彙刊(IEEE Transactions of Multimedia)，第17卷，第10號，Stowell等人的「Detection and

Classification of Acoustic Scenes and Events」中。音頻事件實例包括敲門、關門、說話、歡笑、清喉、咳嗽、拉抽屜、列印、鍵盤單擊、滑鼠單擊、在桌面上作標記、在桌上置放鑰匙、手機振鈴、短警報(例如，蜂鳴聲)及翻頁。音頻事件不限於僅此等實例。音頻事件之其他非限制性實例包括椅子移動、樂器上弦(例如，吉他)、敲擊(例如，轉鼓)、將瓶子置放於桌上、摻合機聲音、直升機聲音及類似者。

【0047】 在相同或替代實施例中，音頻事件判定器204耦接至經組態以基於經判定音頻事件追蹤音頻活動之音頻活動追蹤器206 (參見圖4)。

【0048】 在相同或替代實施例中，裝置包括耦接至經追蹤音頻活動概括器208 (參見圖5)之音頻活動追蹤器206，該經追蹤音頻活動概括器經組態以基於概括模式概括經追蹤音頻活動。

【0049】 在相同或替代實施例中，裝置包括音頻活動回饋提供器210，該音頻活動回饋提供器包括回饋模式選擇器212。回饋模式選擇器212可經由回饋模式設定進行組態。舉例而言，裝置之使用者可組態回饋模式選擇器212以在概括之音頻上提供可聽或視覺回饋或兩者。在圖9及圖10中說明經追蹤音頻活動之例示性回饋。

【0050】 圖3說明音頻事件判定器204之多個實施例。在實施例中，事件判定器304耦接至音頻路徑選擇器302。音頻路徑選擇器302經組態以基於來自一或多個麥克風102之輸出選擇是否將音頻波形提供至音頻事件判定器304。音頻波形可自編解碼器114輸出。音頻波形可經提供至事件分類器組320，或在相同或替代實施例中，亦可經提供至話音度量判定器單元314。話音度量判定器單元314可計算輔助語音處理及/或分類之一或

多個話音度量值。例示性話音度量值在圖7中說明為740A1至740E1 (話音活動偵測旗標、話音頻率、雜湊話音頻框、最近語音之持續時間及音調)。

【0051】 在相同或替代實施例中，音頻路徑選擇器302耦接至接收器301。接收器301經組態以自不同裝置接收信號。信號可包括資料封包、音頻封包或語音封包。資料封包可能囊封音頻封包或語音封包。為了易於描述，除非另外明確指示，否則封包可為資料封包、音頻封包或語音封包。音頻封包或語音封包可儲存於音頻/語音封包緩衝器142 (參見圖1) 中。資料封包可儲存於資料封包緩衝器(未圖示)中。音頻路徑選擇器302可自控制器120中之一或多者接收控制器輸入，該等控制器可輔助是否將封包或音頻波形提供至音頻事件判定器304之選擇。

【0052】 在相同或替代實施例中，可將封包提供至音頻事件判定器304。可藉由有效負載資料提取器312接收封包。有效負載資料提取器312可自封包提取識別資料類型(ID) 630及/或個人識別資訊640 (參見圖6)。舉例而言，有效負載資料提取器312可提取個人識別620，亦即指示此人是誰之一或多個位元。若兩個人說話，例如Zoe及Max，則Zoe之裝置可將表示實際上為Zoe之一或多個位元發送至Max之裝置。Max之裝置可包括有效負載資料提取器312。作為Zoe之個人識別可基於語音辨識、面部辨識或定位於Zoe之裝置(Max之裝置的遠端裝置)上之某一其他辨識技術的結果。舉例而言，Max之裝置可為裝置100，且在獨立裝置上擷取Zoe之話音或面部。

【0053】 在相同或替代實施例中，Max之裝置可廣播與Max之裝置相關聯的識別，其可例如作為所描述欄位(630、640或620)中之一者的替

代經包括，或代替欄位中之一者，例如識別資料類型630可識別Max之裝置。在實施例中，Max之裝置的識別可為隨機數字(取決於使用案例而固定或變化)以提供額外的安全及唯一性。識別資料類型630可為由使用者(亦即，Max)設定之組態的結果。

【0054】 群組態或設定可允許將個人識別620或個人識別資訊640與識別資料類型630分離。在實施例中，裝置可在接近(例如，碰撞或使用NFC)內以允許交換包括使用者識別(例如，1234567)之識別資料類型630。隨後，視組態而定，封包中之1234567的廣播使用者識別可由裝置100接收，且封包之欄位(630、640或620)中的資訊可用於裝置100之音頻活動追蹤器206中。

【0055】 其他實施例可存在以交換經由網際網路或私人訊息介導之權限以允許註冊為群且允許特定個體之權限。在相同或替代實施例中，有效負載資料提取器312可提取個人識別資訊640。在圖7中說明個人識別資訊之大量實例(740A1至740E1及740A2至740E2)。一般熟習此項技術者將瞭解，遠端裝置可具有其自身話音度量判定器單元，其可經發送至裝置100。

【0056】 在相同或替代實施例中，亦可自有效負載資料提取器312提取時戳。來自遠端裝置之時戳可與ID資料類型630、個人識別資訊或個人識別在相同時間或相對接近的時間到達，以允許與ID資料類型630、個人識別資訊640或個人識別620相關聯之其他音頻事件的時間相關性(例如，可基於共同時間)。到達的時戳中可用之有效負載資料係在與裝置100不同的實體位置及空間處擷取。時戳及相關聯資訊(亦即，時間指派之有效負載資料欄位(620、630及640))自提供含有有效負載資料及時戳之封包

的一或多個不同遠端裝置傳送至裝置100，該有效負載資料及時戳又允許音頻活動追蹤器206包括於基於時間之分析中。

【0057】 在相同或替代實施例中，封包可直接經提供至音頻/語音解碼器310。在音頻/語音解碼器310減壓音頻/語音封包之後，亦可將經重建音頻及/或語音波形提供至話音度量判定器單元314。另外，可將經重建音頻及/或語音波形提供至事件分類器組320。由於事件分類器組320可基於封包或音頻波形接收輸入，因此在實施例中，事件分類器組320亦可接收指示所提供輸入是否來自本端裝置(例如，裝置100)或遠端裝置之裝置模式作為輸入。視裝置模式而定，可控制事件分類器組320。事件分類器組320可選擇性地使用回應於來自本端裝置(例如，裝置100)之輸入的記憶體或分類器，或替代地，選擇性地使用回應於來自遠端裝置之輸入的記憶體或分類器。

【0058】 在對事件分類器組320之輸入為ID資料類型及/或個人識別資訊及/或經重建音頻及/或語音波形及/或基於經重建音頻及/或語音波形之話音度量的實施例中，裝置模式可具有由一或多個位元表示之一個值且可由裝置100設定。在對事件分類器組320之輸入為音頻波形及/或基於音頻波形之話音度量的實施例中，裝置模式可具有由一或多個位元表示之不同值(與裝置100相關聯)且亦可由裝置100設定。

另外，對事件分類器組320之控制亦視應用而定。因此，事件分類器組320亦可另外輸出遠端裝置ID、遠端裝置之時戳、個人ID、(裝置100之)本端時鐘及本端裝置識別值(亦即，裝置100之ID)。

【0059】 圖4說明音頻活動追蹤器之例示性實施。例示性音頻活動追蹤器402可包括時戳(本端時鐘抑或遠端時戳)、裝置ID (本端裝置ID抑

或遠端裝置ID)、個人ID及/或音頻事件之輸入。雖然在描述圖2時提供音頻事件之非限制性實例，但熟習此項技術者將瞭解，可跨多個情形環境、上下文及應用利用本文中所描述之技術判定音頻事件之數目。

【0060】 在實施例中，音頻活動追蹤器402可包括比較單元404。比較單元404可在當前音頻事件與先前音頻事件之間執行比較，且判定當前音頻事件與先前音頻事件是否係相同音頻事件408或不同音頻事件406。比較可基於減法。然而，在相同或替代實施例中，該比較可基於評估當前音頻事件值(或表示當前音頻事件之值的集合)大於或小於先前音頻事件值(或表示先前音頻事件值之值的集合)。該比較係基於大於或小於不等式可基於應用或設計選擇。另外，比較單元可包括取決於應用或設計選擇而使用對數、絕對值或臨限值。

【0061】 音頻事件資料緩衝器410可儲存表示相同音頻事件408或不同音頻事件406之一或多個值。音頻事件資料緩衝器410亦可儲存當前音頻事件。另外，在相同或替代實施例中，音頻事件資料緩衝器410可儲存時戳、裝置ID或個人ID中之一或多者，如圖4中所說明。在相同或替代實施例中，音頻緩衝器138可單獨儲存音頻事件。

【0062】 文字標籤可基於音頻事件可表示之物生成。舉例而言，音頻事件可表示鍵盤單擊。文字標籤可為「鍵盤按鍵單擊」。熟習此項技術者將瞭解，音頻事件亦可表示若干次鍵盤單擊，且文字標籤可為「鍵盤按鍵單擊」。若噴氣飛行器飛過頭頂，則源分離器可分離出飛過頭頂的噴氣飛行器且文字標籤可為「噴氣飛行器飛過頭頂」。鍵盤按鍵單擊及噴氣飛行器飛過頭頂可視為兩個單獨的音頻事件，或在一些實施例中，可表示「鍵盤按鍵單擊而噴氣飛行器飛過頭頂」之一個音頻事件。

【0063】 一般熟習此項技術者將瞭解，在實施例中，音頻事件之文字標籤亦可儲存於音頻事件資料緩衝器410中。

【0064】 在本發明內，對音頻事件之引用可包括音頻事件之音頻或語音樣本，或對音頻事件之可互換地引用可包括文字標籤。另外，音頻事件可為音頻事件之集合，例如鍵盤按鍵單擊可為一系列鍵盤按鍵單擊之集合。

【0065】 作為文字標籤抑或作為音頻事件之音頻或語音樣本，音頻事件可經分批處理414 (參見圖8)以判定音頻活動。經判定音頻事件之其他背景可判定音頻活動。舉例而言，若音頻事件(或文字)表示呼吸短促及/或快速的腳步聲，則經判定音頻活動可能正在進行。若音頻事件表示快速的腳步聲像在水泥或沙石上，則此可表示經判定音頻活動正在水泥或沙石上進行。時戳、裝置ID及個人ID亦可提供較多背景。舉例而言，經判定音頻活動可為Max (由個人ID指示)在1pm (由時戳指示)在水泥或沙石上跑步且穿戴手錶(由裝置ID指示)。

【0066】 在音頻活動追蹤器402基於一或多個經判定音頻事件及背景判定音頻活動之後，藉由將經判定音頻活動儲存在音頻活動資料緩衝器144中，可保持追蹤經判定音頻活動。經追蹤判定之音頻活動的概括可基於經追蹤音頻活動概括器502 (參見圖5)中之概括模式。

【0067】 一般熟習此項技術者將瞭解，在實施例中，經追蹤判定之音頻活動的文字標籤亦可儲存於音頻活動資料緩衝器144中。

【0068】 在本發明內，對音頻活動或經判定音頻活動之引用可包括音頻活動之音頻或語音樣本，或對音頻活動或經判定音頻活動之互換地引用可包括文字標籤。另外，音頻活動或經判定音頻活動可為音頻活動之集

合，且對音頻活動之集合的引用亦可稱作音頻活動。舉例而言，「Max下午在水泥上跑步」可為Max在1pm與5pm之間跑步至少幾天的一系列音頻活動之集合。音頻活動「Max下午在水泥上跑步」之集合亦可稱作音頻活動。基於一或多個感測器108資料之相關性，該等感測器108中之一或多者可輔助音頻活動之集合且增加音頻事件偵測之信賴度。

【0069】 在相同或替代實施例中，音頻活動「Max在水泥上跑步」之集合可由一或多個感測器108輔助，該一或多個感測器在Max增加步數時基於對水泥之影響偵測呼吸、心跳速率、加速度回應之形狀。舉例而言，偵測器組804中之感測器偵測器816可偵測此等一或多個感測器108之輸出(例如，慣性及光電血管容積圖(PPG)或心跳速率(HR)感測器)(參見圖8)。除了與呼吸、腳步聲及腳步對水泥之影響的所記錄聲音相關聯之聲音度量之外，分批處理800可使用諸如感測度量或視覺度量之其他感測器輸入。因此，分批處理除了具有偵測器組804中的偵測器中之一或多者(參見圖8)之外，包括可基於其他感測器輸入之音頻活動(例如跑步)。

【0070】 在一些實施例中，偵測器組804可包括用以擷取自一或多個攝影機106擷取之本端視訊內容的視覺背景偵測器814，該一或多個攝影機可用於使視訊內容與音頻信號相關聯以增加判定音頻事件之信賴度，或在一些實施例中，輔助音頻事件之判定。類似地，在一些實施例中，偵測器組804可在一些實施例中包括感測器偵測器816，該感測器偵測器可用於使一或多個感測器讀數與音頻信號相關聯以增加判定音頻事件之信賴度，或在一些實施例中，有助於音頻事件之判定。在相同或替代實施例中，感測器偵測器816可由一或多個感測器108(例如，慣性及光電血管容積圖(PPG)或心跳速率(HR)感測器)輔助。

【0071】 在替代實施例中，對在特定日(例如，星期五)與特定人會見或在特定日(例如，星期五)參加不同小組會議之追蹤可用作音頻活動之集合，且對音頻活動(「星期五會議」)之集合的引用亦可被稱作音頻活動。

【0072】 圖5說明追蹤活動概括器之例示性實施例。經追蹤音頻活動概括器502可提供概括之音頻活動作為輸出。音頻活動之概括可基於概括模式。在概括模式選擇器504中說明例示性概括模式。舉例而言，概括模式可為時間模式506、情緒模式508、活動模式510、環境模式512、語音狀態模式514、關聯模式516、重疊模式518及/或建議模式524。

【0073】 在實施例中，時間模式概括模式可基於持續時間提供對音頻活動之概括。舉例而言，持續時間可為歷經持續10分鐘、60分鐘、3小時、一日、一週、一月、一季度、數年等。另外，時間模式可由裝置100之使用者定義。作為實例，裝置100之使用者可定義持續時間，如將下午定義為1pm及5pm。因此，Max過去一年在水泥上跑步多少個下午的概括之經追蹤音頻活動可由經追蹤音頻活動概括器提供。在實施例中，時間模式可概括持續15分鐘或大於15分鐘的會話之統計資料。舉例而言，在人員A(例如，Max)與人員B(例如，Zoe)之間的會話中，經追蹤音頻活動概括器可判定在其二十分鐘的會話期間Zoe講話50%之時間且Max講話在50%的時間之間。

【0074】 在實施例中，情緒模式概括模式可基於由情緒偵測器(參見圖8)偵測之各種情緒提供音頻活動之概括。舉例而言，在實施例中，基於利用情緒偵測器判定笑、哭、尖叫及叫喊之音頻事件，基於高興、興奮、恐懼、憤怒、溫柔或悲傷所導致的音頻活動之概括可由經追蹤音頻活動概

括器502提供。

【0075】 在實施例中，活動模式概括模式可基於各種類型的活動提供音頻活動之概括。舉例而言，運動係一種類型的活動。經追蹤音頻活動概括器502可提供特定運動期間的音頻活動之概括。舉例而言，若裝置100之使用者(例如，Max)打籃球，則音頻活動可指示Max在投籃之前運球多少次或其在比賽中運球多長時間。

【0076】 在實施例中，環境模式概括模式可基於環境之類型提供音頻活動之概括。在某些環境中，聲音可較大(例如，建築場地)、安靜(例如，圖書館)或低水平(例如，在一些辦公室或家裏)。可利用例如人群雜訊之其他類型對環境進行分類，該等類型可有助於識別音頻活動發生的位置。經追蹤音頻活動概括器502可基於環境是否安靜、喧鬧、低水平、存在人群雜訊或環境類型之其他分類提供音頻活動之概括。

【0077】 在實施例中，語音狀態模式概括模式可基於個人之語音狀態提供音頻活動之概括。不同語音狀態之實例可分類為被動收聽(例如，收聽電視或收音機)、收聽(例如，在會話中)或交談(例如，在會話中)。經追蹤音頻活動概括器502可基於環境是否安靜、喧鬧、低水平、存在人群雜訊或語音狀態之其他分類提供音頻活動之概括。

【0078】 在實施例中，關聯模式概括模式可基於關聯模式提供音頻活動之概括。關聯模式可基於個人或一群人之關聯。舉例而言，當人員A及人員B一起執行音頻活動時，概括可期望基於音頻活動。舉例而言，其正在打網球、游泳、喝咖啡聊天、開車去上班等。

【0079】 在替代實施例中，可存在可允許概括之經追蹤音頻活動基於位置資訊的位置模式(未圖示)。位置資訊可能可來自不同源，或在音頻

事件(例如，自記錄之Wi-Fi封包監聽)之後經判定，且可允許基於位置顯示或提供概括之活動。舉例而言，位置可為在家、工作、駕駛、網球俱樂部、咖啡館及類似者。

【0080】 在相同或替代實施例中，可藉由選擇兩個或大於兩個概括模式提供概括。作為實例，在選擇重疊模式518之後，可選擇時間模式506及情緒模式508兩者。另一實例係在選擇重疊模式之後，選擇活動模式510及語音狀態模式514。因此，在重疊模式中，可基於兩個或大於兩個概括模式提供概括。

【0081】 在相同或替代實施例中，可存在將回饋提供至使用者之建議模式524。舉例而言，可存在校正個人說出「emm」、「umm」或過使用某些字語(例如，「如」)之不良說話習慣的回饋。可即時進行回饋及/或事後分析回饋。

在相同或替代實施例中，可經由由包括經追蹤音頻活動概括器502之裝置100處理的說話關鍵字來選擇概括模式。使用者可說出例示性片語「時間模式」、「情緒模式」、「環境模式」、「語音狀態模式」、「關聯模式」、「重疊模式」及/或「建議模式」來選擇一或多個概括模式。在相同或替代實施例中，可在不具有重疊模式本身的情況下，但藉由選擇兩個或大於兩個概括模式執行多個實施例之選擇。

【0082】 在相同或替代實施例中，可經由由包括經追蹤音頻活動概括器502之裝置100處理的說話關鍵字來選擇概括模式。使用者可說出例示性片語「時間模式」、「情緒模式」、「環境模式」、「語音狀態模式」、「關聯模式」、「重疊模式」及/或「建議模式」來選擇一或多個概括模式。在相同或替代實施例中，可在不具有重疊模式本身的情況下，但藉由選擇兩

個或大於兩個概括模式執行多個實施例之選擇。

【0083】 在相同或替代實施例中，經追蹤音頻活動概括器502可整合或耦接至顯示裝置100，且概括模式之選擇可基於觸摸顯示裝置100，其中概括模式為選項按鈕選擇。

【0084】 可在螢幕上觸摸選項按鈕以選擇時間模式506、情緒模式508、活動模式510、環境模式512、語音狀態模式514、關聯模式516、重疊模式518及/或建議模式524。在替代實施例中，在選項按鈕上未展示快速重疊模式518，但藉由觸摸所描述之概括模式中之一或多者來選擇大於一個概括模式之能力可允許經追蹤音頻活動概括器502基於一或多個概括模式提供一或多個音頻活動的概括。

【0085】 在實施例中，為在概括模式選擇器504中選擇概括模式可經由按壓整合或耦接至包括經追蹤音頻活動概括器502之裝置100的一或多個推按按鈕。舉例而言，按壓推按按鈕可選擇時間模式506。第二次按壓推按按鈕可選擇情緒模式508。後續按壓可選擇活動模式510。接著按壓可選擇環境模式512。此後之按壓可選擇語音狀態模式514。按鈕之另一按壓可選擇關聯模式516。

【0086】 在相同或替代實施例中，若存在整合或耦接至包括經追蹤音頻活動概括器502之裝置100之額外按鈕，則可選擇重疊模式518以基於多個概括模式提供概括之音頻活動。舉例而言，在選擇重疊模式518之後，可選擇時間模式506及情緒模式508兩者，且可基於時間及情緒提供概括。

【0087】 對概括之經追蹤音頻活動緩衝器526中之概括的音頻之回饋模式可為可聽或視覺的，或為可聽及視覺的。舉例而言，可經由揚聲器

有聲地描述圖9及圖10中所說明之概括之經追蹤音頻活動。另外，概括之經追蹤音頻活動可顯示於如圖9中所說明之圖形使用者介面(GUI)上，且在圖10以圖表形式或經由說明不同音頻活動之間的部分、百分比或鏈接之另一視覺表示顯示。視覺回饋模式亦可包括文字。舉例而言，音頻活動之清單可經顯示或儲存於記憶體中。概括之經追蹤音頻活動緩衝器526亦被稱作概括之經追蹤音頻活動緩衝器144。

【0088】 圖6說明可為兩個裝置之間通信期間的封包600之部分的不同欄位。當通信基於無線連接時，封包600可被稱作無線封包。封包(不論無線與否) 600可包括前置碼608、標頭610及有效負載資料612。前置碼608可用於使在交換資訊的兩個裝置之間的傳輸及/或接收同步。標頭610可包括關於封包600內攜帶之資料的指令，例如封包600內所含有的欄位或子欄位之長度或數量。

【0089】 在實施例中，封包之有效負載資料612可包括個人識別620，亦即指示此人是誰之一或多個位元。個人識別620可基於語音辨識、面部辨識或某一其他辨識技術之結果。若封包600由裝置100接收，則辨識技術可定位於另一裝置上。在實施例中，本端裝置(亦即，裝置100)可整合辨識技術，且封包600可經傳輸至裝置100外部。

【0090】 在相同或替代實施例中，封包600可包括ID資料類型630。舉例而言，ID資料類型630可指示辨識是否為定位於遠端裝置上且經發送至本端裝置(亦即，裝置100)之語音辨識或面部辨識或某一其他辨識技術之結果。

【0091】 在相同或替代實施例中，有效負載資料612可包括個人識別資訊640。在圖7中說明個人識別資訊之大量實例(740A1至740E1及

740A2至740E2)。

【0092】 圖7說明不同類型之個人識別。可基於一或多個麥克風102擷取之個人識別在圖7中說明為740A1至740E1。話音活動偵測旗標740A1可表示何時存在有聲或無聲語音。有聲頻率740B1可為經量化線性預測係數或人說話之頻率分佈(振幅及共振峰位置)的包封之其他類似表示。雜湊話音頻框740C1可表示對於特定人而言特有之聲紋。最近語音之持續時間740D1可為某人說話的時間長度，例如5秒。音調740E1可表示人說話之音調。兒童及女性傾向於具有較高音調，而男性傾向於具有較低音調。

【0093】 在相同或替代實施例中，可基於一或多個感測器108擷取之額外或替代性個人識別在圖7中說明為740A2至740E2。人頭部之溫度740A2可用於指示其處於過冷抑或過熱之實體環境中。若溫度過熱或過冷，則人可為該人處於壓力下的指示符。使溫度讀數與其他感測器讀數相關聯可提供較多背景以輔助判定所追蹤之音頻活動。人的心跳740B2亦可提供與壓力、平靜及/或鍛煉相關之較多背景以輔助判定所追蹤及概括之音頻活動。人之身體部位運動740C2可指示此人是否正在跑步、行走、站立、抬高手臂、轉動腕部、頭部運動。身體部位之此等運動可提供輔助所追蹤及概括之音頻活動的較多背景。與最接近物體之接近度740D2可為提供有助於追蹤及概括音頻活動之內容的另一指示符。裝置ID之定向740E2可為提供有助於追蹤及概括音頻活動之內容的另一指示符。

【0094】 圖8說明分批處理音頻事件之一實施例。如先前參考圖4所描述，可在音頻活動追蹤器402中分批處理800音頻事件。經判定音頻事件可與先前經判定音頻事件聚集802且經提供至偵測器組804。舉例而言，經判定音頻事件可為鍵盤按鍵單擊。若聚集多個鍵盤單擊，則經判定

音頻事件之聚集可為單擊鍵盤按鍵。另一實例可為經判定音頻事件可為腳踏在沙石上。多個音頻事件(諸如腳步或鍵盤單擊)之聚集可經發送至偵測器組84。

【0095】 偵測器組804可包括環境偵測器806、情緒偵測器808、語音狀態偵測器810及可聽背景偵測器812。在相同或替代實施例中，其他輸入可由偵測器組804處理。舉例而言，ID資料類型、可聽度量(諸如話音度量、視覺度量及/或感測度量)可充當偵測器組804之輸入且供偵測器組804中的偵測器中之一或多者使用。來自感測器之其他背景可提供較多背景。舉例而言，可由某人說出之工作片語或咖啡館片語(亦即，與工作環境相關聯之片語或在咖啡館中)由一或多個麥克風102擷取。基於判定片語之背景的可聽背景偵測器812可輔助自偵測器組804輸出之音頻活動係在辦公室的膝上型電腦上工作或在咖啡館處打字的人。在另一實例中，環境偵測器806可偵測海浪且輔助判定自偵測器組804輸出之音頻活動係在大海附近的沙灘上跑步的人。

【0096】 基於來自一或多個麥克風102之可聽輸入的偵測器組804或事件分類器組320可依賴於機器學習或用於偵測之其他自適應或分類技術。舉例而言，情緒偵測器808及可聽背景偵測器812可基於使用深度神經網路。作為另一實例，諸如對成年人與兒童、或男性與女性之間的差異進行分類之音頻事件可基於分類器或深度神經網路。另一實例在事件分類器組320中可使用格式結構分類器及音調分類器以判定來自正在播放音樂之電視或揚聲器之內容。偵測器組804可基於早先所論述之個人識別資訊640(圖8中未繪示)來判定音頻活動。

【0097】 在一些實施例中，偵測器組804可包括用以擷取自一或多

個攝影機106擷取之本端視訊內容的視覺背景偵測器814，該一或多個攝影機可用於使視訊內容與音頻信號相關聯以增加判定音頻事件之信賴度，或在一些實施例中，輔助音頻事件之判定。類似地，在一些實施例中，偵測器組804可在一些實施例中包括感測器偵測器816，該感測器偵測器可用於使一或多個感測器讀數與音頻信號相關聯以增加判定音頻事件之信賴度，或在一些實施例中，有助於音頻事件之判定。在相同或替代實施例中，感測器偵測器816可由一或多個感測器108（例如，慣性及光電血管容積圖(PPG)或心跳速率(HR)感測器)輔助。

【0098】 一或多個麥克風102或耦接至麥克風102之處理器可進行工作循環以節省能量。此可允許較長時間段持續監視音頻活動。在一些系統中，口頭字語之偵測允許處理器離開閒置模式且接通額外的麥克風。裝置100可具有較接近或遠離人說話之麥克風102。一或多個麥克風102可允許音量監視。舉例而言，可聽設備可具有允許音量監視之外部麥克風。一或多個麥克風102可允許喊叫、交談、私語及類似者之偵測。

【0099】 基於可聽輸入之偵測器亦可包括經過訓練的話音範本。當裝置100可包括天線112及收發器122時，可在兩個裝置之間交換或共享經過訓練的話音範本，此可有助於減少對不具有經過訓練的話音範本之裝置的訓練。不同可信群可共享訓練資料。可信群之實例為一家庭。在家庭成員當中，可在裝置之間交換或共享經過訓練的話音範本。天線112及收發器122允許追蹤及概括待傳輸至另一裝置之音頻活動。傳輸可基於音頻活動追蹤器及概括器裝置(例如，裝置100)之使用者的授權。舉例而言，一人哭了三個小時且未對任何人說話，此可經傳輸至親人。另一實例可為老年親屬在一段時間內尚未進行社交互動，且可對親屬建議打電話或去拜

訪。在又一實例中，可為一個在某個臨限值位準上已暴露於過量背景雜訊中。

【0100】 圖9說明顯示裝置110之GUI 902上每日音頻活動之概括之例示性視圖。在圖9之例示性視圖中，說明人參與之音頻活動904。可在顯示裝置110上向使用者展示906概括模式。不同概括模式可由例如下拉式功能表箭頭908選擇(或可由選項按鈕或藉由在概括模式之名稱中輸入來選擇，作為實例)。舉例而言，若概括模式為時間模式，則指明特定時間段活動之概括的人可為進行以下之人：駕駛汽車(A)、傾聽Joe (B)、聽音樂(C)、與Mary說話(D)、在安靜環境中(E)及收聽電視(F)。餅圖圖表允許顯示比例。

【0101】 在相同或替代實施例中，使用者亦可選擇與概括模式相關聯之臨限值。舉例而言，可不必顯示持續小於10分鐘的音頻活動。因此，可存在人在特定時間段期間參與之但未在顯示裝置110之GUI 902上說明的其他音頻活動。

【0102】 圖10說明與顯示器上的概括之經追蹤音頻活動相關聯之情緒的鏈接之另一例示性視圖。在圖10之例示性視圖中，音頻活動(A-F)之百分比可鏈接至與概括之經追蹤音頻活動1050相關聯之情緒且顯示於顯示裝置110之不同GUI 1040上。例示性情緒可包括(但不限於)「無聊」、「中性」、「激動」或「高興」。

【0103】 圖11說明例示性使用案例1100。在圖11中，人在聚會上進行社交活動。在聚會上，可發生多個不同的音頻活動。實例包括：在電視上觀看體育比賽(A1)、唱生日歌(B1)、祝酒(C1)、打開禮品(D1)、吃晚餐(E1)或玩遊戲(F1)。在人聚會之房間或空間內，可存在包括音頻活動單元

202及音頻活動回饋提供器210之裝置1104。裝置1104可包括可以類似於一或多個麥克風102之先前描述之方式運作的麥克風陣列1106。裝置1104可包括具有可顯示音頻活動(例如，A1至F1)之概括的GUI 1108之顯示裝置。

【0104】 在相同或替代實施例中，人中之一些可穿戴包括如本文中所示之音頻活動單元202及音頻活動回饋提供器210之可穿戴式裝置。舉例而言，人具有環繞其腕部之可穿戴式裝置。可穿戴式裝置可為腕帶或手錶(兩者引用為1102A)。另一人穿戴包括如本文中所示之音頻活動追蹤器之可聽設備1102B。另一人具有環繞其頸部之可穿戴式裝置1102C。可穿戴式裝置可為具有麥克風陣列(亦即，一或多個麥克風102)之項鍊，或將來可在社交上可接受穿戴之包括環繞頸部之一或多個麥克風102之裝置。可穿戴物1102A、1102B及1102C可以視覺或可聽形式提供音頻活動之概括。舉例而言，可聽設備1102B及環繞頸部之可穿戴式裝置1102C可將概括之音頻活動傳送至顯示裝置。若可穿戴式裝置1102A為手錶，則存在整合至手錶之顯示裝置。若可穿戴式裝置1102A為不具有顯示裝置之腕帶，則概括之音頻活動可經傳送至顯示裝置。可穿戴物1102A、1102B及1102C亦可包括播放音頻活動之概括的一或多個揚聲器。

【0105】 在特定實施中本文中所揭示之系統及裝置之一或多個組件可整合至解碼系統或設備(例如，電子裝置或其中之處理器)中，整合至編碼系統或設備中，或兩者中。在其他實施中，本文中所揭示之系統及裝置之一或多個組件可整合至無線電話，智慧型眼鏡，或擴增實境裝置、虛擬實境裝置、混合實境裝置、擴展實境裝置之將來版本，機器人，平板電腦，桌上型電腦，膝上型電腦，機上盒，音樂播放機，視訊播放機，娛樂

單元，電視，遊戲控制台，導航裝置，通信裝置，個人數位助理(PDA)，固定位置資料單元，個人媒體播放機或另一種類型的裝置中。

【0106】 結合所描述之技術，一種裝置包括用於基於接收音頻信號判定音頻事件之構件、用於基於經判定音頻事件追蹤音頻活動之構件、用於基於概括模式概括經追蹤音頻活動之構件及用於提供概括之經追蹤音頻事件的回饋之構件。另外，該裝置可包括用於選擇回饋模式之構件。

【0107】 熟習此項技術者將瞭解，在一些實施例中，裝置100不具有周邊裝置(例如，一或多個感測器、一或多個攝影機及/或顯示裝置)或組件(例如，編解碼器、天線、收發器)中之一些，其中裝置100為伺服器，其亦可經組態以基於接收音頻信號判定音頻事件，基於經判定音頻事件追蹤音頻活動，基於概括模式概括經追蹤音頻活動及基於回饋模式提供概括之經追蹤音頻活動之回饋。

【0108】 熟習此項技術者將進一步瞭解，結合本文中所揭示之實施描述之各種說明性邏輯區塊、組態、模組、電路及演算法步驟可實施為電子硬體、由處理器執行之電腦軟體或兩者之組合。上文已大體上就其功能性而言描述各種說明性組件、區塊、組態、模組、電路及步驟。此功能性被實施為硬體抑或處理器可執行指令取決於特定應用及強加於整個系統之設計約束。熟習此項技術者可針對每一特定應用而以變化之方式實施所描述之功能性，但不應將此類實施決策解釋為致使脫離本發明之範疇。

【0109】 結合本文中之揭示內容描述之方法或演算法之步驟可直接以硬體、由處理器執行之軟體模組或兩者之組合來實施。軟體模組可駐存於隨機存取記憶體(RAM)、快閃記憶體、唯讀記憶體(ROM)、可程式化唯讀記憶體(PROM)、可抹除可程式化唯讀記憶體(EPROM)、電可抹除可

程式化唯讀記憶體(EEPROM)、暫存器、硬碟、可移除式磁碟、緊密光碟唯讀記憶體(CD-ROM)或此項技術中已知的任何其他形式之非暫時性儲存媒體中。例示性儲存媒體耦接至處理器，使得處理器可自儲存媒體讀取資訊且將資訊寫入至儲存媒體。以上各者之組合亦應包括於非暫時性電腦可讀媒體之範疇內。另外，方法或演算法之操作可作為程式碼及/或指令中之一者或任何組合或集合而駐存於可併入至電腦程式產品中之非暫時性處理器可讀媒體及/或非暫時性電腦可讀媒體上。在替代例中，儲存媒體可整合至處理器中。處理器及儲存媒體可駐存於特殊應用積體電路(ASIC)中。ASIC可駐存於計算裝置或使用者終端機中。在替代例中，處理器及儲存媒體可作為離散組件駐存於計算裝置或使用者終端機中。

【0110】 先前描述經提供以使熟習此項技術者能夠進行或使用所揭示之實施。熟習此項技術者將容易地顯而易見對此等實施之各種修改，且在不脫離本發明之範疇的情況下，本文中所定義之原理可應用於其他實施。因此，本發明並非意欲限制於本文中所展示之實施，而應符合可能與如以下申請專利範圍所定義之原理及新穎特徵相一致的最廣泛範疇。

【符號說明】

【0111】

100	裝置
102	麥克風
106	攝影機
108	感測器
110	顯示裝置
112	天線

114	音頻/語音編碼器/解碼器
116	音頻活動單元
118	音頻活動回饋提供器
120	控制器
122	收發器
124	網路介面
126	記憶體
128	處理器
130	指令
131	像素緩衝器
132	感測器緩衝器
136	音頻/語音緩衝器
138	音頻事件緩衝器
142	音頻/語音封包緩衝器
144	概括之經追蹤音頻活動緩衝器
202	音頻活動單元
204	音頻事件判定器
206	音頻活動追蹤器
208	音頻活動概括器
210	音頻活動回饋提供器單元
212	回饋模式選擇器
301	接收器
302	音頻路徑選擇器

304	音頻事件判定器
310	音頻/語音解碼器
312	有效負載資料提取器
314	話音度量判定器單元
320	事件分類器組
402	音頻活動追蹤器
404	比較單元
406	不同音頻事件
408	相同音頻事件
410	音頻事件資料緩衝器
414	步驟
502	音頻活動概括器
504	概括模式選擇器
506	時間模式
508	情緒模式
510	活動模式
512	環境模式
514	語音狀態模式
516	縮合模式
518	重疊模式
524	建議模式
526	音頻活動緩衝器
600	封包

608	前置碼
610	標頭
612	有效負載資料
620	個人識別
630	封包提取識別資料類型
640	個人識別資訊
740A1	話音活動偵測旗標
740A1至740E1	話音度量值/個人識別資訊
740A2	人頭部之溫度
740A2至740E2	個人識別資訊
740B1	有聲頻率/語音頻率
740C1	雜湊話音頻框
740D1	最近語音之持續時間
740E1	音調
740B2	人的心跳
740C2	人之身體部位運動
740D2	與最接近物體之接近度
740E2	裝置ID之定向
800	分批處理
802	步驟
804	偵測器組
806	環境偵測器
808	情緒偵測器

810	語音狀態偵測器
812	可聽背景偵測器
814	視覺背景偵測器
816	感測器偵測器
902	GUI
904	音頻活動
906	步驟
908	箭頭
1040	GUI
1050	概括之經追蹤音頻活動
1100	使用案例
1102A	可穿戴式裝置
1102B	可聽設備
1102C	可穿戴式裝置
1104	裝置
1106	麥克風陣列
1108	GUI

【發明申請專利範圍】

【第1項】

一種用於音頻處理之裝置，其包含：

一或多個處理器，其經組態以：

基於接收一音頻信號判定一音頻事件；

基於該經判定音頻事件追蹤音頻活動；

基於由使用者選擇的一概括模式，概括該經追蹤音頻活動，其中該概括模式係以下中之一或多者：一時間模式、一情緒模式、一活動模式、一環境模式、一語音狀態模式、一關聯模式、一重疊模式、或一建議模式；及

基於一回饋模式提供該概括之經追蹤音頻活動之回饋。

【第2項】

如請求項1之裝置，其進一步包含耦接至該一或多個處理器之一音頻解碼器，且該音頻解碼器經組態以自另一裝置接收一封包。

【第3項】

如請求項2之裝置，其中該音頻解碼器耦接至一有效負載資料提取器，且該有效負載資料提取器經組態以自該另一裝置接收之該封包提取一個人識別、識別類型、個人識別資訊、遠端裝置ID或一時戳中之至少一者。

【第4項】

如請求項2之裝置，其進一步包含一音頻路徑選擇器，該音頻路徑選擇器經組態以選擇是否將基於來自一或多個麥克風之輸出的該音頻信號提供至經組態以判定該音頻事件的一音頻事件判定器，或將自該另一裝置接

收之該封包提供至該音頻事件判定器。

【第5項】

如請求項4之裝置，其中該音頻路徑選擇器經組態以接收控制器輸入及提供指示輸入是來自一本端裝置或是一遠端裝置的一裝置模式。

【第6項】

如請求項1之裝置，其進一步包含耦接至一音頻編解碼器之一或多個麥克風，該音頻編解碼器經組態以輸出該音頻信號且經組態以將該音頻信號提供至該一或多個處理器。

【第7項】

如請求項6之裝置，其中該經判定音頻事件基於至少對來自一話音度量判定器單元之一個話音度量進行分類，且該一或多個處理器包括一事件分類器組，該事件分類器組經組態以提供以下各項中至少一者：一本端時鐘、一本端裝置識別值及該音頻事件。

【第8項】

如請求項7之裝置，其中該事件分類器組由指示輸入是來自一本端裝置或是一遠端裝置的一裝置模式控制。

【第9項】

如請求項1之裝置，其進一步包含一記憶體，該記憶體包括一概括之經追蹤音頻活動緩衝器。

【第10項】

如請求項9之裝置，其中該一或多個處理器包括經組態以選擇該回饋模式之一音頻活動回饋提供器。

【第11項】

如請求項10之裝置，其進一步包含一顯示裝置，該顯示裝置經組態以在選擇一視覺回饋模式時顯示概括之經追蹤音頻活動之一視覺回饋。

【第12項】

如請求項10之裝置，其進一步包含一或多個揚聲器，該一或多個揚聲器經組態以在選擇一可聽回饋模式時產生概括之經追蹤音頻活動之一可聽回饋。

【第13項】

如請求項12之裝置，其中該一或多個揚聲器整合至一可聽裝置中。

【第14項】

如請求項1之裝置，其中該一或多個處理器包括一音頻活動追蹤器以基於該概括模式提供該經追蹤音頻活動，且其中該音頻活動追蹤器包含比較一先前音頻事件以判定該經判定音頻事件是否為一相同音頻事件或一不同音頻事件。

【第15項】

如請求項14之裝置，其中該音頻活動追蹤器經組態以基於經聚集的該經判定音頻事件且基於一時戳、裝置識別值或個人識別值中之至少一者分批處理一個以上的該音頻事件或該經判定音頻事件。

【第16項】

如請求項15之裝置，其中該經追蹤音頻活動係基於檢測環境、情緒、語音狀態、可聽背景、視覺背景或一感測器讀數中之至少一者而判定。

【第17項】

如請求項16之裝置，其中該經追蹤音頻活動另外基於一識別資料類

型、個人識別及個人識別資訊中之一者而判定。

【第18項】

如請求項1之裝置，其進一步包含一顯示裝置，該顯示裝置經組態以表示與該概括之經追蹤音頻活動相關聯之一情緒的一鏈接。

【第19項】

一種用於音頻處理之方法，其包含：

基於接收一音頻信號判定一音頻事件；

基於該經判定音頻事件追蹤一音頻活動；

基於由使用者選擇的一概括模式，概括該經追蹤音頻活動，其中該概括模式係以下中之一或多者：一時間模式、一情緒模式、一活動模式、一環境模式、一語音狀態模式、一關聯模式、一重疊模式、或一建議模式；及

基於一回饋模式提供該概括之經追蹤音頻活動之回饋。

【第20項】

如請求項19之方法，於基於該回饋模式提供該概括之經追蹤音頻活動之回饋之前，該方法進一步包含選擇該回饋模式的種類。

【第21項】

如請求項20之方法，其中當該回饋模式為一視覺回饋模式時，顯示該概括之經追蹤音頻活動之該回饋。

【第22項】

如請求項20之方法，其中當該回饋模式為一可聽回饋模式時，該概括之經追蹤音頻活動之該回饋為可聽的。

【第23項】

如請求項19之方法，其中該概括之經追蹤音頻活動基於比較一先前音頻事件以判定該經判定音頻事件是否為一相同音頻事件或一不同音頻事件。

【第24項】

如請求項19之方法，其中該概括之經追蹤音頻活動基於聚集該經判定音頻事件分批處理大於一個音頻事件。

【第25項】

如請求項19之方法，其中該概括之經追蹤音頻活動經傳輸至另一裝置。

【第26項】

一種用於音頻處理之設備，其包含：

用於基於接收一音頻信號判定一音頻事件之構件；

用於基於該經判定音頻事件追蹤一音頻活動之構件；

用於基於由使用者選擇的一概括模式，概括該經追蹤音頻活動之構件，其中該概括模式係以下中之一或多者：一時間模式、一情緒模式、一活動模式、一環境模式、一語音狀態模式、一關聯模式、一重疊模式、或一建議模式；及

用於基於一回饋模式提供該概括之經追蹤音頻活動之回饋的構件。

【第27項】

如請求項26之設備，其進一步包含用於選擇該回饋模式之構件。

【第28項】

一種非暫時性電腦可讀儲存媒體，其上儲存有指令，該等指令在執行時使得一裝置之一或多個處理器：

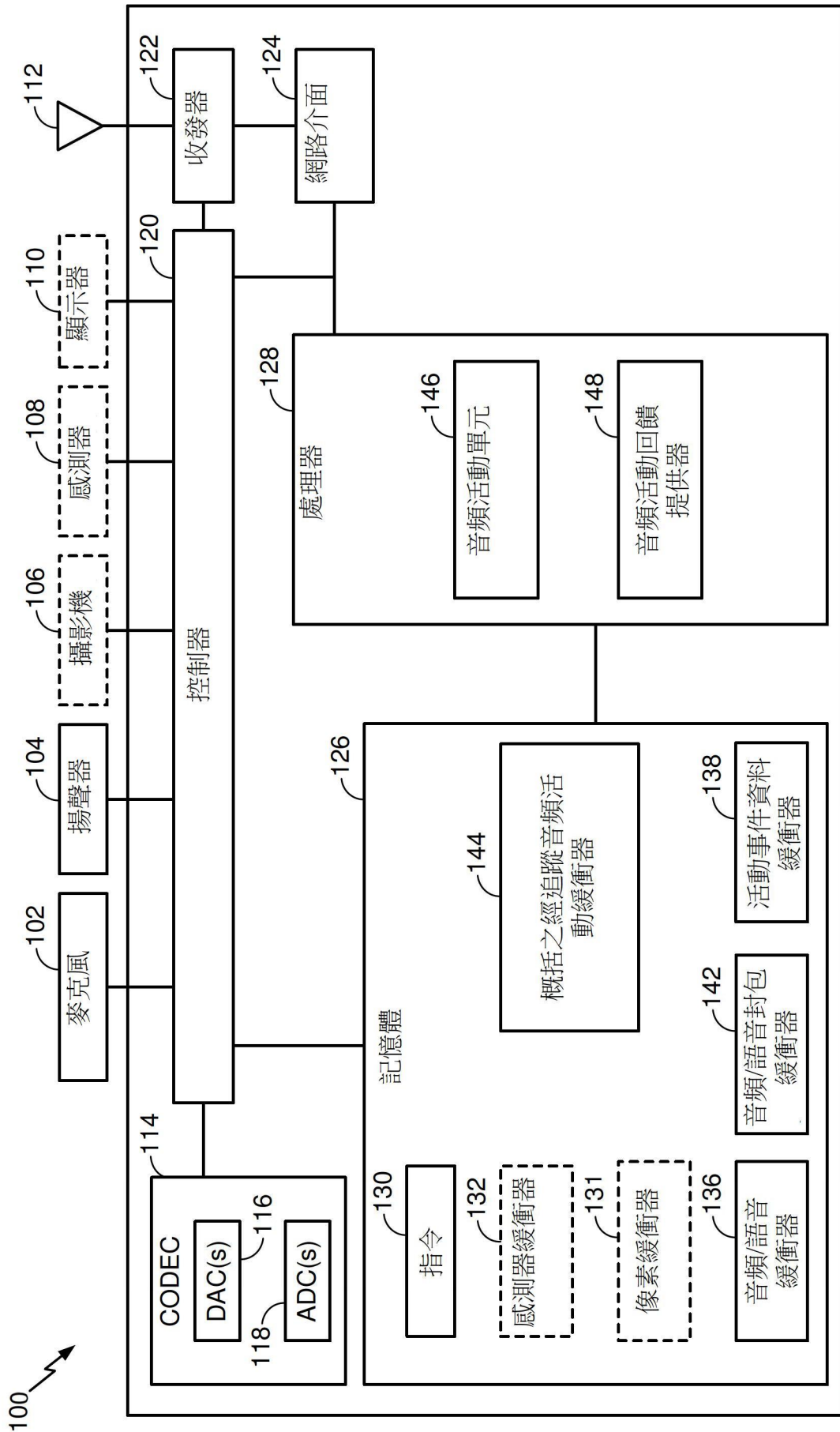
基於接收一音頻信號判定一音頻事件；

基於該經判定音頻事件追蹤一音頻活動；

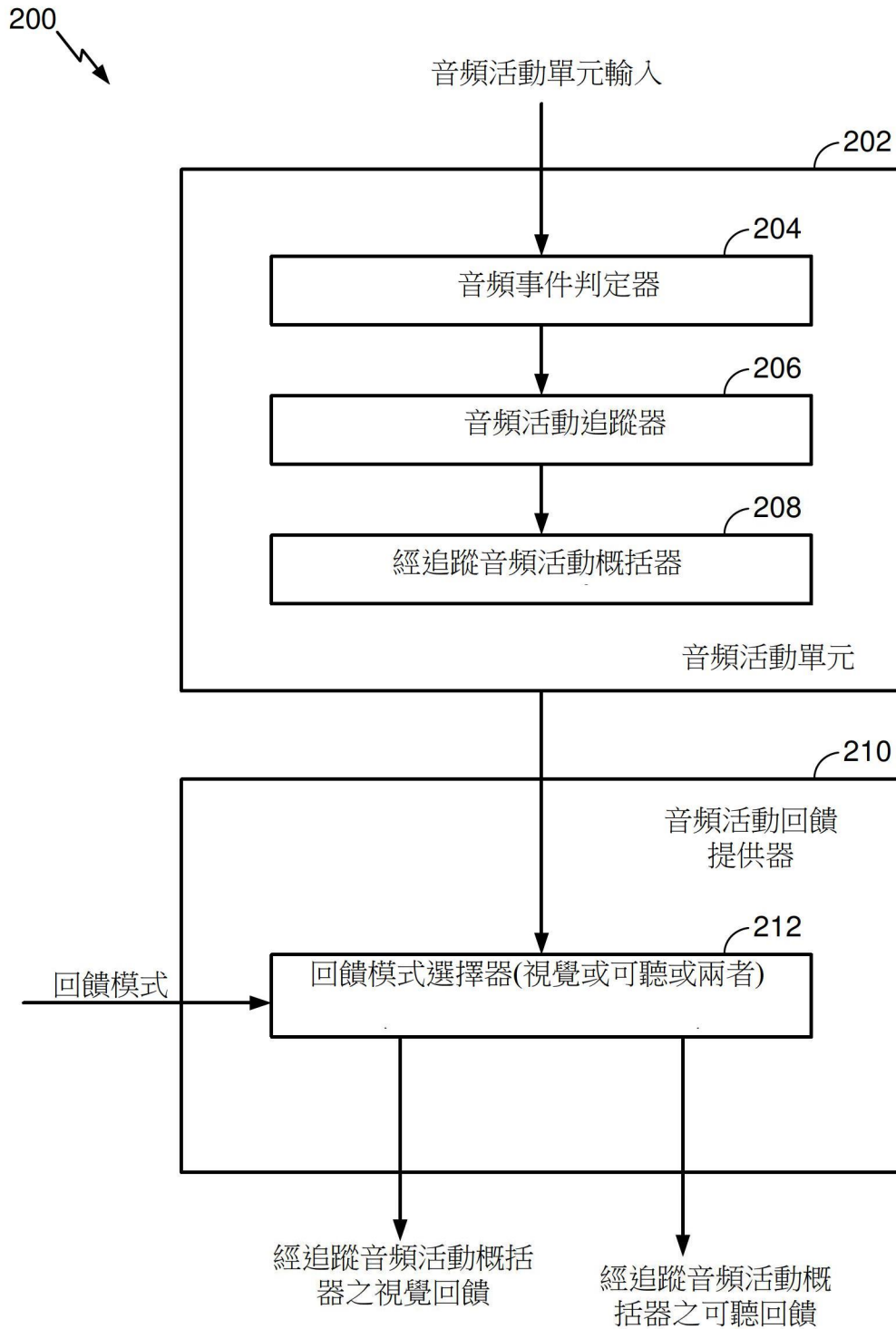
基於由使用者選擇的一概括模式，選擇概括該經追蹤音頻活動，其中該概括模式係以下中之一或多者：一時間模式、一情緒模式、一活動模式、一環境模式、一語音狀態模式、一關聯模式、一重疊模式、或一建議模式；及

基於一回饋模式提供該概括之經追蹤音頻活動之回饋。

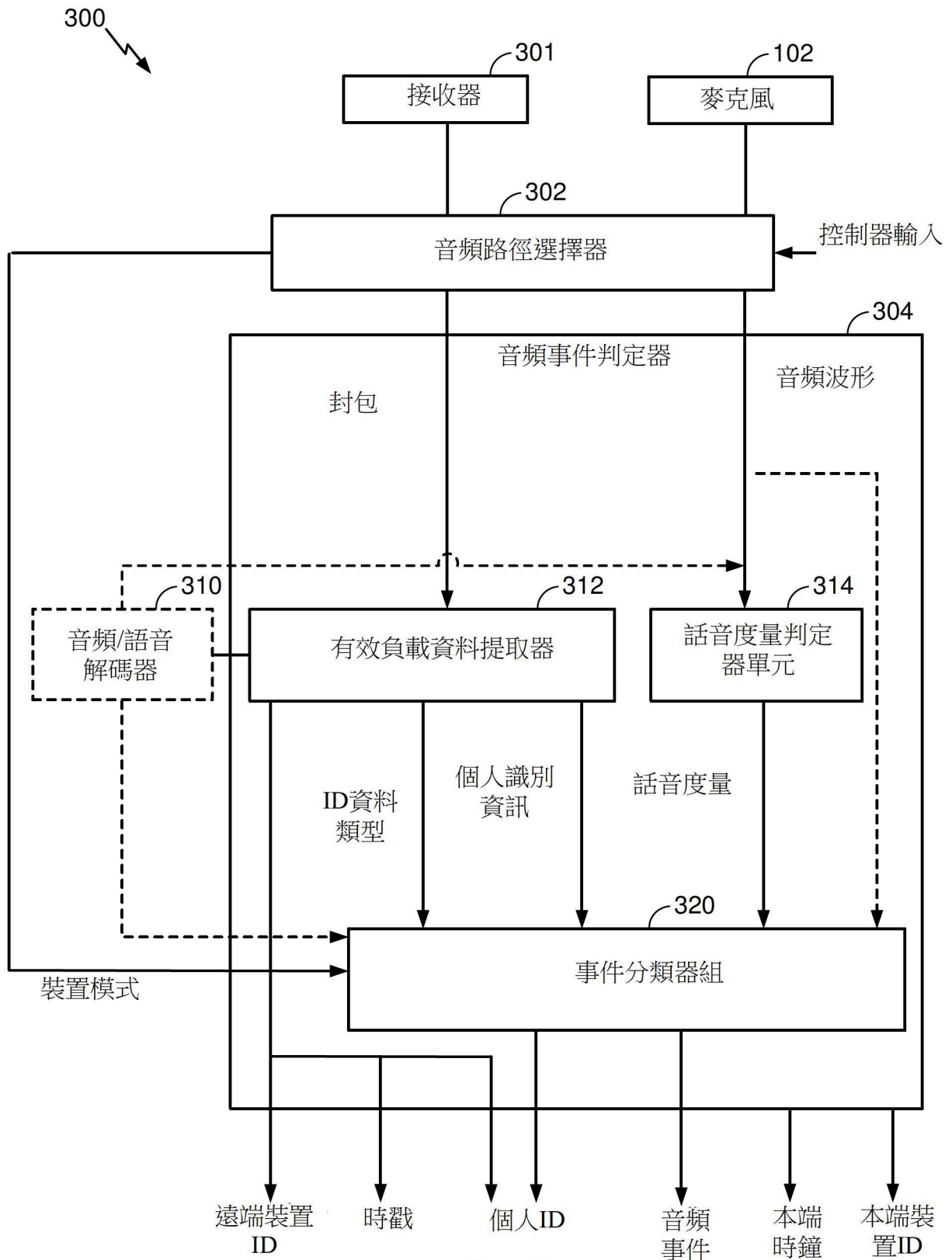
【發明圖式】



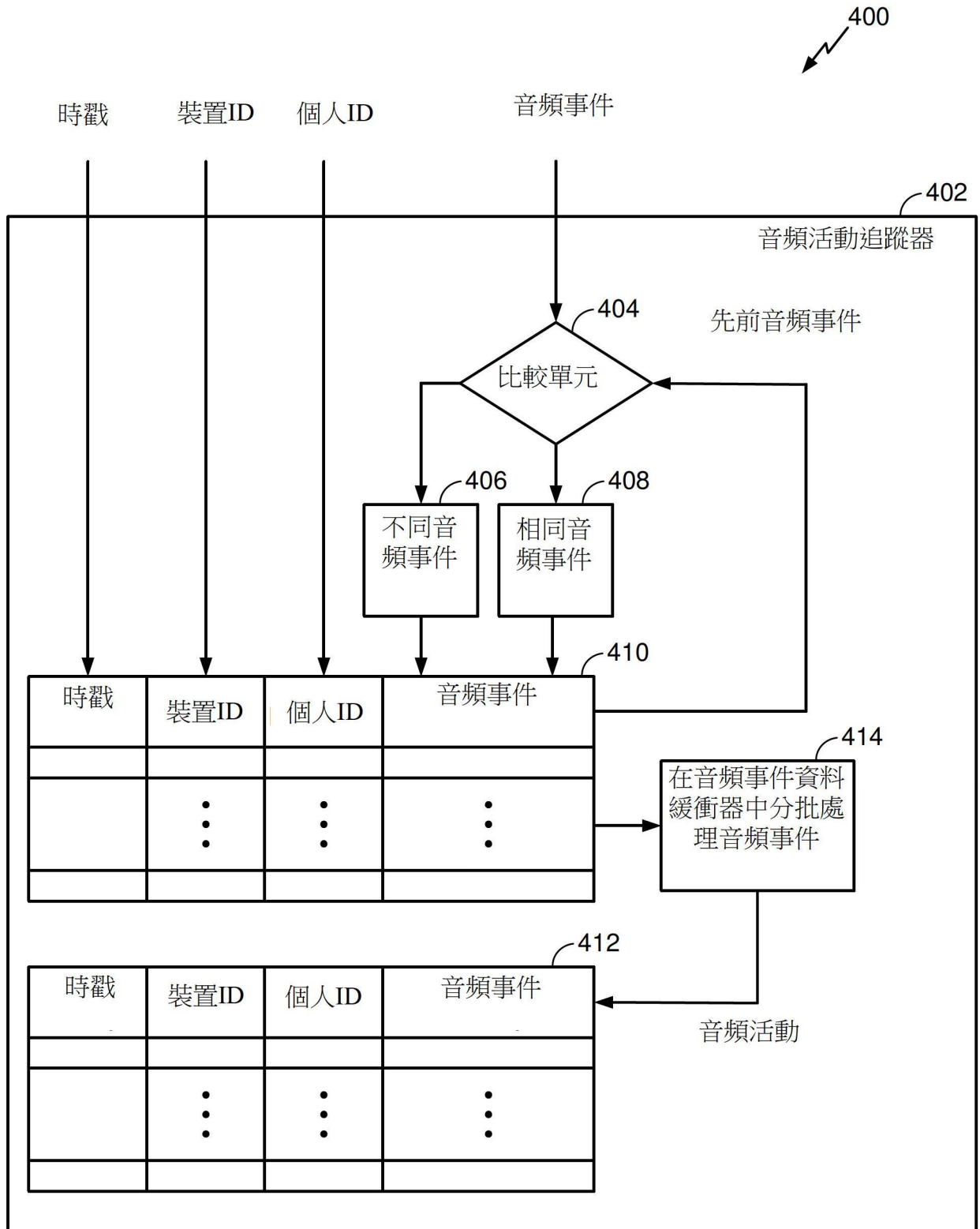
【圖1】



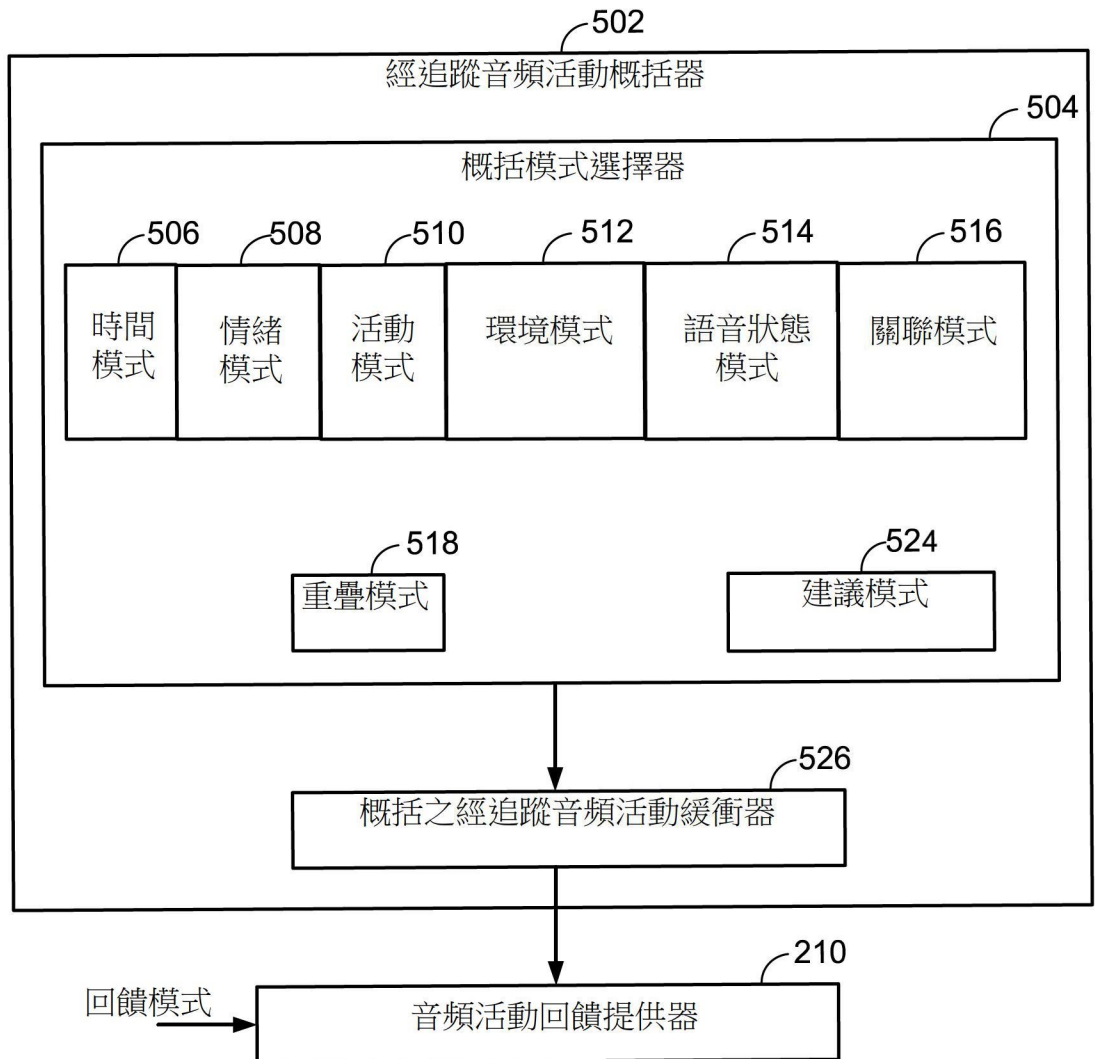
【圖2】



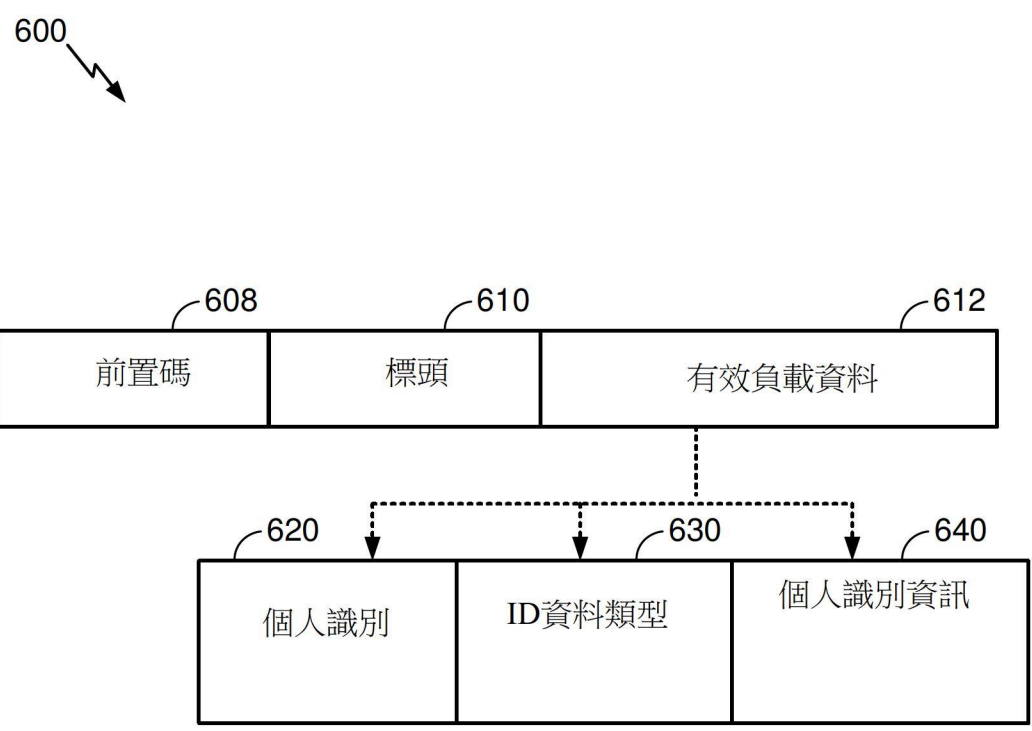
【圖3】



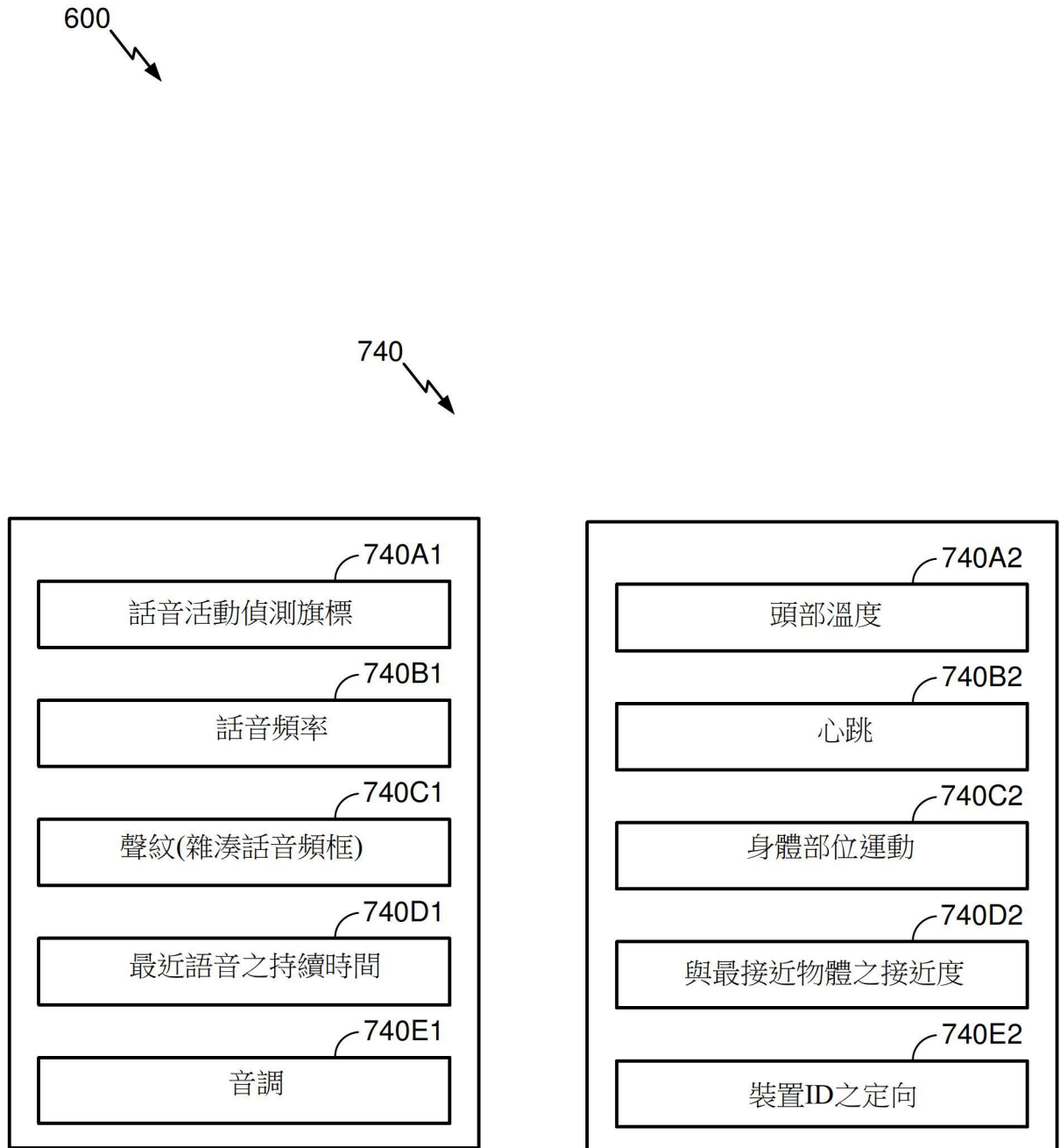
【圖4】



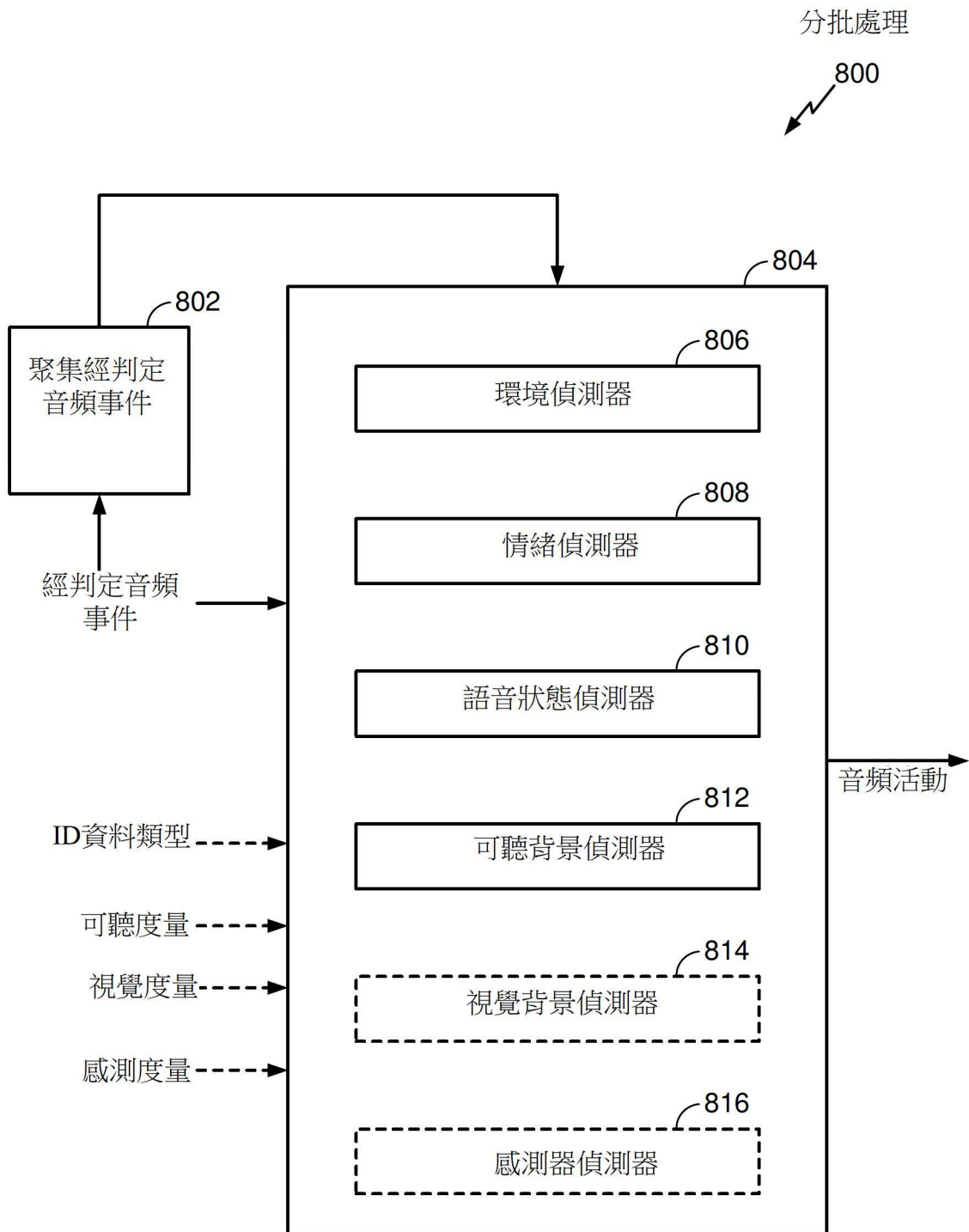
【圖5】



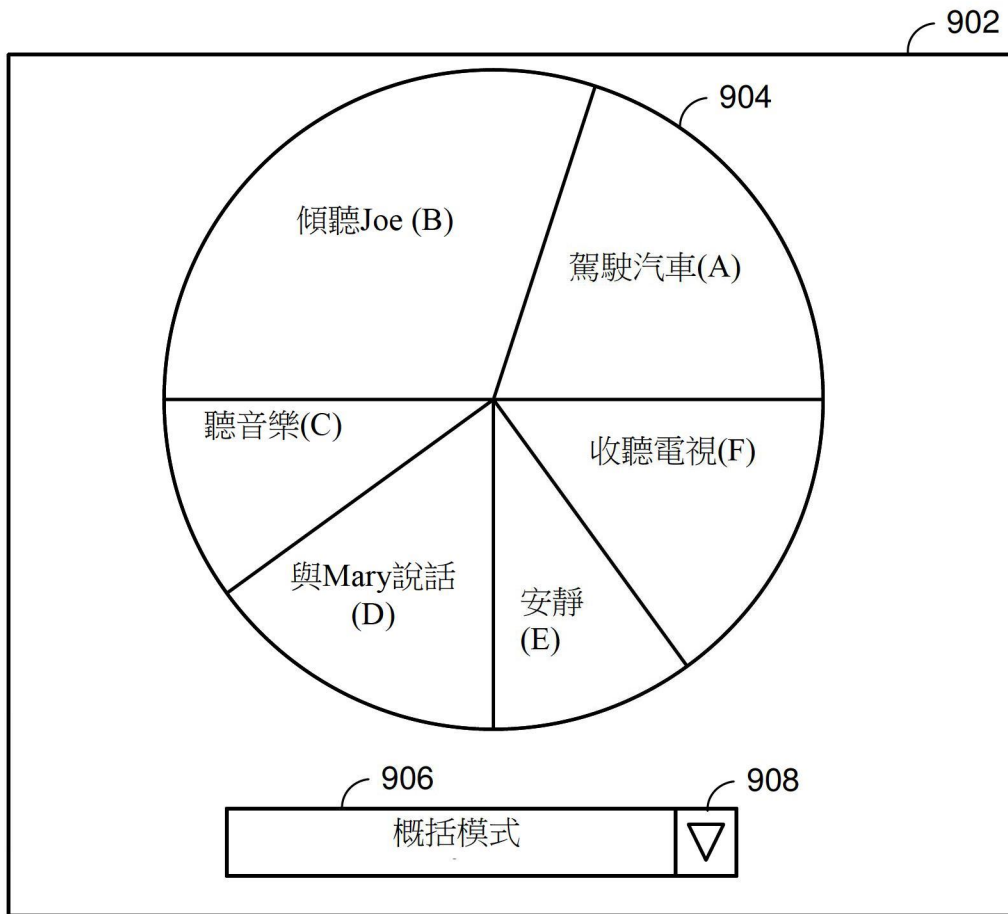
【圖6】



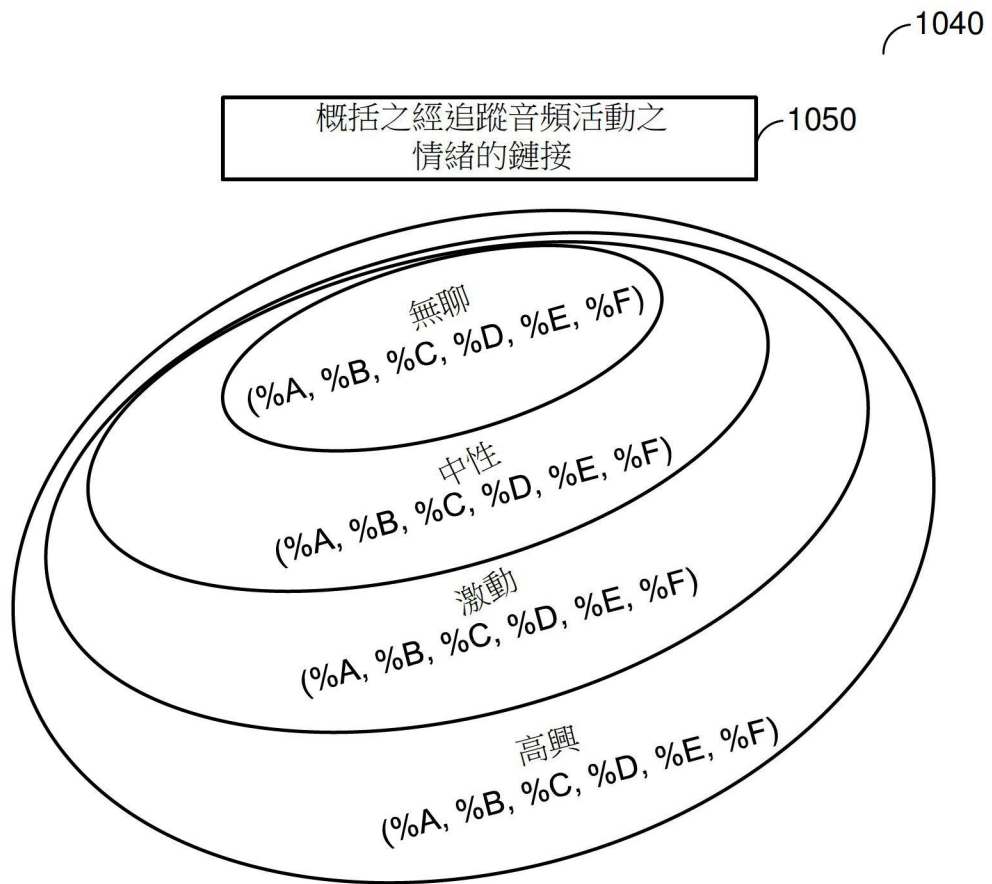
【圖7】



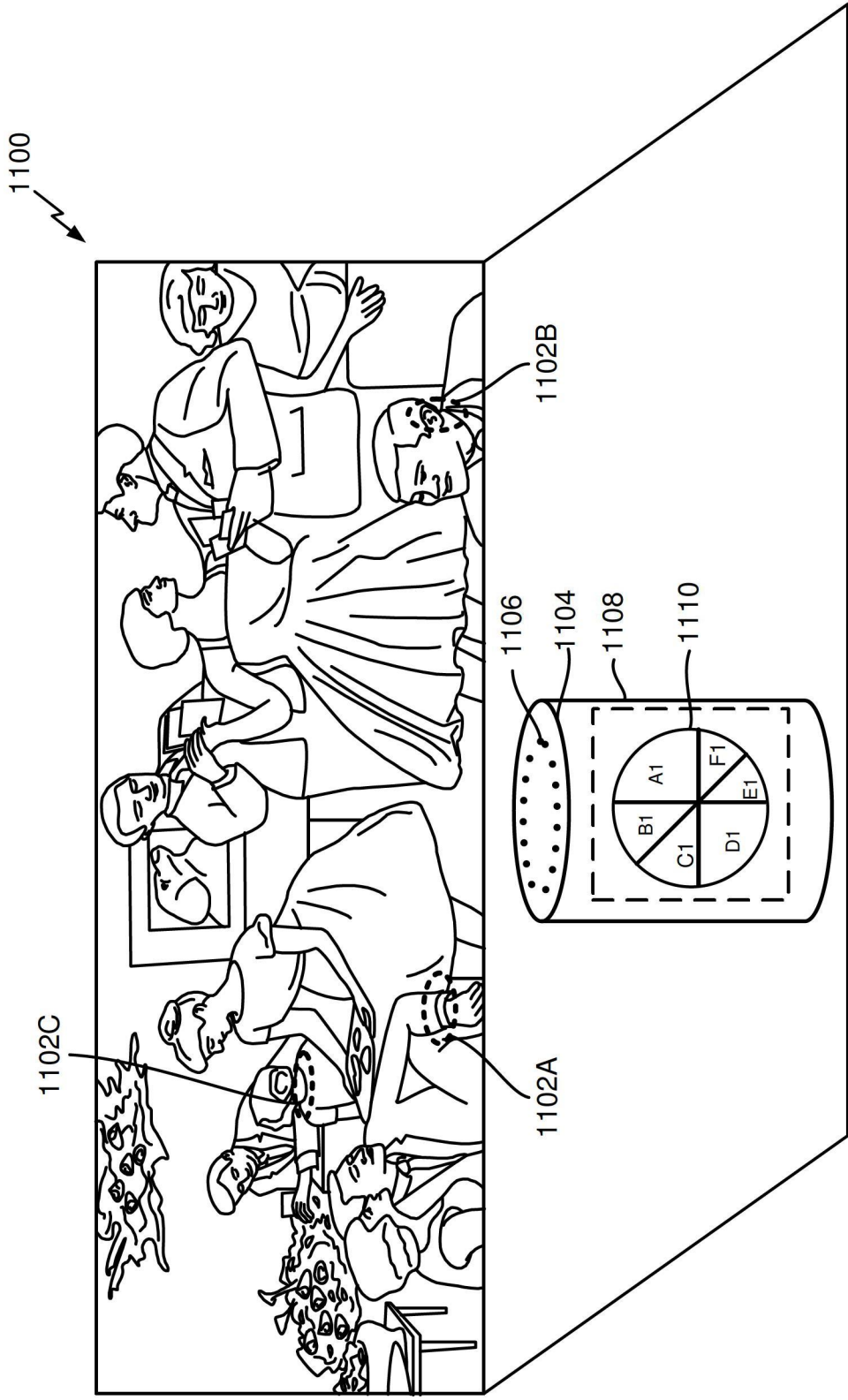
【圖8】



【圖9】



【圖10】



【圖11】