



(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 I583191 B

(45)公告日：中華民國 106 (2017) 年 05 月 11 日

(21)申請案號：105109391

(22)申請日：中華民國 103 (2014) 年 05 月 28 日

(51)Int. Cl. : H04N21/84 (2011.01)

G06F17/30 (2006.01)

(30)優先權：2013/06/03 美國

61/830,227

(71)申請人：雅虎股份有限公司 (美國) YAHOO! INC. (US)
美國

(72)發明人：佩沙文托 蓋瑞 PESAVENTO, GERRY (US) ; 阮 輝 X NGUYEN, HUY X. (US)

(74)代理人：陳傳岳；郭雨嵐

(56)參考文獻：

CN 102763105A

US 2011/0208722A1

WO 2011/104428A1

審查人員：黎世琦

申請專利範圍項數：30 項 圖式數：10 共 53 頁

(54)名稱

相片及視頻搜尋

PHOTO AND VIDEO SEARCH

(57)摘要

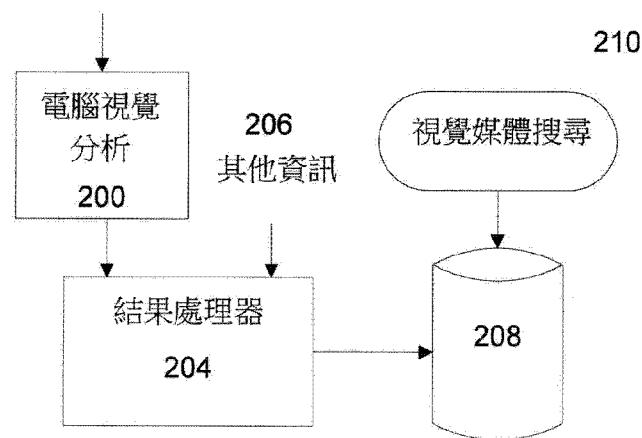
在一個具體實施例內，藉由執行一視覺媒體項目的影像內容的電腦視覺分析，獲得所產生之一組標記，其中該組標記的每一標記都具有一對應的機率。此外，可獲得獨立於該視覺媒體項目的影像內容以外的資訊集合。至少部分根據該資訊集合，可修改至少一部分該組標記的該機率。

In one embodiment, a set of tags that has been generated by performing computer vision analysis of image content of a visual media item may be obtained, where each tag of the set of tags has a corresponding probability. In addition, a set of information that is independent from the image content of the visual media item may be obtained. The probability of at least a portion of the set of tags may be modified based, at least in part, upon the set of information.

指定代表圖：

符號簡單說明：

視覺媒體項目202



- 200 · · · 電腦視覺分析
202 · · · 視覺媒體項目
204 · · · 結果處理器
206 · · · 其他資訊
208 · · · 標記資料庫
210 · · · 視覺媒體搜尋

第二圖

發明摘要

※ 申請案號：105109391 (由 103118543 分割)

※ 申請日：103.5.28

※ IPC 分類：H04N 21/84 (2011.01)
G06F 17/30 (2006.01)

【發明名稱】(中文/英文)

相片及視頻搜尋/ PHOTO AND VIDEO SEARCH

【中文】

在一個具體實施例內，藉由執行一視覺媒體項目的影像內容的電腦視覺分析，獲得所產生之一組標記，其中該組標記的每一標記都具有一對應的機率。此外，可獲得獨立於該視覺媒體項目的影像內容以外的資訊集合。至少部分根據該資訊集合，可修改至少一部分該組標記的該機率。

【英文】

In one embodiment, a set of tags that has been generated by performing computer vision analysis of image content of a visual media item may be obtained, where each tag of the set of tags has a corresponding probability. In addition, a set of information that is independent from the image content of the visual media item may be obtained. The probability of at least a portion of the set of tags may be modified based, at least in part, upon the set of information.

【代表圖】

【本案指定代表圖】：第（二）圖。

【本代表圖之符號簡單說明】：

200 電腦視覺分析

206 其他資訊

202 視覺媒體項目

208 標記資料庫

204 結果處理器

210 視覺媒體搜尋

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

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)

相片及視頻搜尋/ PHOTO AND VIDEO SEARCH

【技術領域】

本揭示具體實施例一般係關於執行電子視覺媒體項目自動標記，幫助電子視覺媒體項目搜尋與組織之電腦實施方法及設備。

【先前技術】

隨著例如智慧型電話以及平板電腦這些行動裝置的流行，大量的收集例如相片與視頻這些視覺媒體。使用者通常將視覺媒體儲存在行動裝置當中，並且透過許多服務與媒體儲存網站上傳儲存其視覺媒體。為了之後可收回該視覺媒體，使用者可手動用標題或類別來標記其視覺媒體。

不幸的是，許多使用者都認為手動標記視覺媒體耗時乏味。更進一步，因為許多使用者都累積大量的視覺媒體，所以該視覺媒體手動標記通常不可行。因此，許多使用者忽略標記其視覺媒體。

【發明內容】

所揭示的具體實施例可由電腦產生與視覺媒體項目相關聯的標記。視覺媒體項目可包含但不受限於相片及視頻。所產生的標記可用來幫助搜尋與組織媒體項目。

在一個具體實施例內，由執行一視覺媒體項目之影像內容的電腦視覺分析得到一組標記，其中該組標記的每一標記都具有一對應的機率。此外，可獲得從該視覺媒體項目的影像內容獨立出來之資訊集合。根據該資訊集合的至少一部分，可改修改少部分該組標記的機率。

在另一個具體實施例內，可執行一視覺媒體項目之影像內容的電腦視覺分析來產生一組標記，其中該組標記的每一標記都具有一對應機率。可得到從該視覺媒體項目的影像內容獨立出來之資訊集合。根據該資訊集合的至少一部分，可修改與至少部分該組標記相關聯的機率。

在另一個具體實施例內，本發明關於一種包含一處理器、記憶體以及一顯示器的裝置。該處理器與記憶體設置成執行上述方法操作之一或多者。在另一個具體實施例內，本發明屬於一種其上儲存電腦程式指令的電腦可讀取儲存媒體，這些指令配置成執行一或多個上述之方法操作。

在以下本發明說明書以及藉由本發明原理所例示的附圖當中，將更詳細呈現本發明的這些與其他特徵和優點。

【圖式簡單說明】

第一圖為例示可實施本發明具體實施例的一範例系統圖。

第二圖為可實施許多具體實施例的一範例系統之方塊圖。

第三圖為例示根據本發明不同的具體實施例標記視覺媒體項目的一範例方法之流程圖。

第四圖為例示根據許多具體實施例透過使用行動裝置感應器所獲得中繼資料之圖式。

第五圖為例示根據本發明不同的具體實施例使用座標來修改標機率的一範例系統之圖示。

第六圖為例示根據許多具體實施例視覺上類似視覺項目的識別之圖式。

第七圖為例示一未監督的特徵學習範例應用之方塊圖。

第八圖為例示可實施許多具體實施例的一網路範例具體實施例之示意圖。

第九圖為例示可實施許多具體實施例的一範例用戶端裝置之示意圖。

第十圖為例示可實施許多具體實施例的一範例電腦系統之示意圖。

【實施方式】

詳細參照本發明的特定具體實施例，這些具體實施例的範例都說明於附圖內。本發明將結合這些具體實施例來說明，但請瞭解這並不將本發明限制在這些具體實施例上。相反地，如申請專利範圍內所定義，將改變、修改以及同等配置包含在本發明的精神與範疇內。在下列說明中，將提出許多特定細節以提供對本發明通盤的了解。不過，在沒有某些或全部這些特定細節的情況下也可實施本發明。在其他實例中，並未詳細說明已知的操作過程，以避免模糊本發明。「實施方式」並非已知概念的延伸或詳細討論，因此在此省略精通相關技術人士已知的細節，或可用總結方式處理。

底下將參考附圖更完整說明本發明，為本說明書的一部分，並且藉由例示顯示特定範例具體實施例。不過，本主張主題可具體實施於許多不同形式，因此所涵蓋或申請主張的主題旨在解釋，並不受限於本說明書所揭示的任何範例具體實施例；範例具體實施例僅為例示。同樣，本發明所申請或涵蓋主張之主題目的在於提供合理寬闊的範疇。除此之外，例如主張之主題可具體實施為方法、裝置、元件或系統。因此，具體實施例可採用例如硬體、軟體、韌體或這些的任意組合(已知並非軟體)之形式。因此，下列詳細說明並不設限。

整個說明書與申請專利範圍中，詞彙具有超出上下文中明確陳述意義

的建議或暗示含意。同樣地，本說明書內使用的詞彙「在一個具體實施例內」並不必要參照相同具體實施例，且本說明書內使用的「在其他具體實施例內」並不必要參照不同的具體實施例。其目的在於例如主張的主題包括全部或部分範例具體實施例的組合。

一般來說，至少部份從上下文的使用當中可了解詞彙。例如：本說明書內使用的「和」、「或」或「和/或」等詞包括許多含義，也至少部分取決於這些用語所使用的範疇。通常「或」若用於關聯一清單，例如A、B或C，就是要表示A、B和C，這用於包括的情況，而A、B或C則用於專屬情況。此外，根據至少部分上下文，本說明書內使用的「一或多個」一詞用來描述單一型態的任何功能、結構或特性，或可用於描述功能、結構或特性的某些組合。同樣地，根據至少部分上下文，例如「一」或「該」這些詞彙同樣表示單次使用或多次使用。此外，「根據」一詞可了解並不必要表達專屬因素集合，而是根據至少部分上下文，允許存在不需明確說明的其他因素。

近年來，網際網路已經成為數百萬使用者的主要資訊來源，這些使用者依賴網際網路搜尋有興趣的資訊。一種讓使用者習以搜尋資訊的方法為透過搜尋服務網頁來搜尋。一般而言，使用者可在搜尋網頁上的輸入框內輸入包含一或多個搜尋詞彙的查詢，然後根據輸入的搜尋詞彙開始搜尋。在回應查詢上，該搜尋服務通常回覆搜尋結果文件的排序清單。

若要取得使用者有興趣的文件，使用者通常透過搜尋服務提出搜尋查詢，並且該搜尋服務回覆搜尋結果文件清單，以回應使用者的搜尋查詢。不過，使用者搜尋視覺媒體項目，例如相片與視頻的能力受限於使用者手

動標記(即是註釋)以及組織其視覺媒體項目的能力。

使用者可藉由使用資料夾(即是專輯)，手動組織其視覺媒體，尤其是，使用者可手動建立具有所要標題的資料夾，例如：視覺媒體項目通常放入依照時間、地點、主題或事件來組織的資料夾內。然後該視覺媒體可移動至相關的資料夾，例如：利用剪下與貼上一視覺媒體項目至所要資料夾內，或拖放該視覺媒體項目至相關資料夾內，可將該視覺媒體項目儲存在特定資料夾內。不過，在視覺媒體項目與一個資料夾以上相關的情況下，該視覺媒體項目通常需要複製到多個資料夾。更進一步，隨著視覺媒體的數量逐漸增加，鮮少有使用者有時間或耐心組織其視覺媒體。

此外，數位視覺媒體可手動標記，尤其是，使用者可使用一或多個標記來標記一已知媒體項目。每一標記都可包含關於該視覺媒體項目的文字，例如：該文字可包含一標題或類型。不過，許多使用者沒有時間標記其媒體，因此許多視覺媒體項目的標記並不完整或全部遺漏。

即使在已經標記視覺媒體項目的情況下，搜尋視覺媒體項目的能力也受限於哪些視覺媒體項目已經標記之精確度。通常搜尋該等視覺媒體項目需要使用者捲動過數千個視覺媒體項目，因此使用者通常對於搜尋與瀏覽視覺媒體項目的能力受限而感到失望。

許多服務以及網站都可用來分享視覺媒體項目，不過由於缺乏標記，所以無法輕鬆依照主題，例如場景、物體或事件將視覺媒體項目分組。因此，使用者無法輕鬆識別要與他人分享的視覺媒體項目。

根據許多具體實施例，視覺媒體經過處理，自動產生或精製（refine）包含一組標記的中繼資料，如此該中繼資料與該視覺媒體相關聯，尤其是

透過自動標記處理，一視覺媒體項目可關聯於包含一組一或多個標記的中繼資料，其中每一標記都具有與之相關聯的機率(即是權重)。如此可產生及/或精製與一視覺媒體項目相關聯的該中繼資料，如底下進一步詳細說明。

產生或精製的中繼資料可用來幫助搜尋視覺媒體項目。此外，該中繼資料可用來幫助組織視覺媒體項目。

一電子視覺媒體項目可包含一或多個數位影像，例如：一電子視覺媒體項目可包含一視頻、照片(即是相片)或電子產生的影像，例如圖片或圖表。相片通常包含視覺影像，而視頻通常包含移動的視覺影像，例如：一視頻可包含一電影或視頻短片。一電子視覺媒體項目可用合適格式的檔案形式儲存。

視覺媒體項目可由行動裝置的相機拍攝，該裝置也可稱為可攜式裝置、穿戴式裝置或手持式裝置，例如：可使用數位相機以數位格式產生相片與視頻。或者，一開始並非用數位格式產生的相片與視頻可轉換成數位格式。

至少部分由一行動裝置及/或相對於行動裝置來說位於遠處的遠端伺服器，可實施許多具體實施例，行動裝置的範例包含但不受限於行動電話、智慧型電話、相機、攝影機、顯示型呼叫器、個人數位助理(PDA，Personal Digital Assistant)、手持式電腦、平板電腦以及膝上型電腦。

範例系統

第一圖為例示實施本發明具體實施例的一範例系統圖。如第一圖內所示，該系統可包含與一網站，例如社群網站，相關聯的一或多個伺服器106。社群網站的範例包含Yahoo、Facebook、Tumblr、LinkedIn、Flickr以及

Instagram。伺服器106可讓網站提供許多服務給其使用者，尤其是，該網站的使用者可維護公開使用者簡介、與網站其他成員互動、上傳檔案(例如照片、視頻)等等。

伺服器106可包含一搜尋伺服器以及一網路伺服器。複數個用戶端102a、102b、102c可透過網路104及/或例如網路伺服器上的網路服務，存取例如該搜尋伺服器上的一搜尋應用(即是搜尋服務)。網路104可採用任何合適的形式，像是廣域網路或網際網路及/或一或多個區域網路(LAN，local area network)。網路104可包含任何合適數量與類型的裝置，例如路由器以及交換器，用於將來自每一用戶端的搜尋或網路物件要求轉發至該搜尋或網路應用，並且將搜尋或網路結果回覆給要求的用戶端。

本發明也可用多種網路環境(由網路104所代表)來實施，包含例如TCP/IP型網路、電信網路、無線網路等等。此外，實施本發明具體實施例的電腦程式指令可儲存在任何一種電腦可讀取媒體內，並且可根據許多計算模型來執行，包含主從式模型、點對點模型、單機計算裝置或根據其中此處所描述許多功能性可生效或運用在不同位置上的分散式計算模型。

一搜尋應用程式一般允許使用者(人或自動化實體)搜尋資訊，該資訊關於一查詢，其包含一或多個搜尋詞彙相關。該等搜尋詞彙可由使用者以任何方式輸入，例如：圖形使用者介面可呈現一輸入功能給該用戶端(如此使用者可輸入包含一或多個搜尋詞彙的查詢)。尤其是，該圖形使用者介面可呈現一個輸入框(即是搜尋查詢區)，讓使用者可輸入包含任何搜尋詞彙數量或其一部分的查詢。在某些具體實施例內，可提供建議的搜尋查詢(即是搜尋建議)。然後使用者透過該圖形使用者介面，選擇建議的搜尋查詢之一者

提交至該搜尋應用程式。

然後可透過相關於一或多個資料來源的搜尋應用程式，來執行該搜尋查詢。該等資料來源可包含本地資料來源(例如用戶端上)及/或透過網路104位於遠端的資料來源。在某些具體實施例內，可執行自然語言處理(NLP，natural language processing)來解析使用者查詢及其意圖。在回應該搜尋查詢方面，該搜尋應用程式可識別並呈現資訊，其可包含與該使用者搜尋查詢最相關的文件或視覺媒體項目。

該搜尋應用程式可透過伺服器106安裝並執行，例如該搜尋伺服器及/或網路伺服器。雖然為了清晰只例示單一伺服器106，不過可在任何數量的伺服器上實施該搜尋應用程式。在其他具體實施例內，該搜尋應用程式可安裝在該用戶端裝置上並執行。

此外，根據許多具體實施例，可自動組織並呈現視覺媒體項目。這種組織可幫助搜尋或由使用者瀏覽。

為了幫助視覺媒體項目的搜尋及/或組織，會用一組標記自動標記該等視覺媒體項目。然後可至少部分根據相關聯標記，其包含該等自動產生的標記，執行該等視覺媒體項目的搜尋及/或組織，當然，與視覺媒體項目相關聯的標記也可包含手動產生的標記。底下進一步詳細說明執行自動標記的許多方法及機制。

本說明書內揭示的具體實施例可透過搜尋伺服器106及/或其他伺服器及/或用戶端102a、102b、102c來實施，例如：透過用戶端102a、102b、102c上的網路瀏覽器及/或應用程式可實施許多功能，所揭示的具體實施例可透過軟體及/或硬體來實施。

搜尋伺服器106 (或多個伺服器)可保留要留存之搜尋資訊的一或多個查詢記錄，例如：該等查詢記錄可保留在一或多個耦合至伺服器106的記憶體108內。每此使用者執行一或多個搜尋詞彙的搜尋時，有關此搜尋的資訊會保留在該等查詢記錄內。例如：該使用者的搜尋要求可內含任何數量的參數，如使用者或瀏覽器識別與該等搜尋詞彙，這可保留在該等查詢記錄內。有關該搜尋的額外資訊，例如一時間戳記，也可搭配該等搜尋要求參數，保留在該查詢記錄內。根據該輸入的搜尋參數將結果呈現給該使用者時，來自此搜尋結果的參數也可保留在該等查詢記錄內。例如：該等特定搜尋結果的識別(例如URL)，例如網站、呈現該等搜尋結果的順序、每一搜尋結果是否有贊助或為算術搜尋結果、每一搜尋結果的擁有者(例如網站)及/或每一搜尋結果是否已經由使用者(若有的話)選取(即是點按)，可保留在該等查詢記錄內。

此外，記憶體108可儲存內含使用者資訊的一或多個使用者記錄(例如使用者資料庫)。此使用者資訊或其一部分可涉及使用者簡介，尤其是，該使用者簡介可包含可用於公開簡介的公開資訊及/或私人資訊。更進一步，該使用者簡介可包含該使用者已經提交的資訊，及/或該系統已經推斷或自動收集的資訊(例如根據使用者的動作)。保留在該等使用者記錄內的該使用者資訊可包含個人資訊，像是人口統計資訊(例如年齡及/或性別)及/或地理資訊(例如全球定位系統(GPS，Global Positioning System)位置、網際網路通訊協定(IP，Internet Protocol)位址、國家區域/州、城市及/或郵遞區號)，其可包含或指出該使用者的居住地址、工作地址及/或目前位置。此外，該使用者資訊可包含使用者的興趣、該使用者購買的產品或服務、該使用者時常

在網際網路上存取的資訊屬性(例如類別)及/或該使用者的職業。在該等使用者記錄當中可藉由一使用者ID(例如使用者帳號ID)、使用者cookie當中的資訊等等，來識別一使用者。該等使用者記錄可保留在一或多個記憶體內。

在此可實施許多種機制來產生一簡介，包含但不受限於收集或探勘瀏覽歷史紀錄、儲存的文件、標記或註釋，僅供參考。簡介建立程式可儲存產生的簡介。一搜尋引擎的使用者簡介，例如，可賦予一搜尋引擎提供者取得註釋、標記、儲存的頁面、瀏覽歷史等等之機制，這對於製作搜尋結果相關文件相當有用，特別是特定使用者。

所揭示的具體實施例可讓視覺媒體項目進行標記、搜尋、瀏覽、組織及/或提供給使用者，以回應一搜尋查詢或該視覺媒體項目內其他興趣指示器。尤其是，伺服器106可透過網站(例如透過網站網頁上的顯示)、透過電子郵件、簡訊服務(SMS，Short Message Service)、透過行動裝置(例如文字訊息)或透過其他媒體，例如連接至網際網路的數位電視，將視覺媒體項目或其識別碼提供給使用者。

第二圖為其中可實施許多具體實施例的一範例系統之方塊圖。電腦視覺分析200可針對視覺媒體項目202，例如相片或視頻，來執行以產生一組標記。此外，執行電腦視覺分析200，可使每一組標記都具有與之相關聯的一對應機率，其中該機率指出該標記精確地描述該視覺媒體項目影像內容的可能性。在擷取視覺媒體項目202時、在上傳視覺媒體項目202時或稍後時，電腦視覺分析200可被執行。

一結果處理器204可取得該組標記和已透過電腦視覺分析200產生與其相關聯的機率，以及獨立於(例如在之外)視覺媒體項目202的影像內容之外

的其他資訊206。結果處理器204可分析該組標記以及其他資訊206，這種分析可包含檢查該組標記與其他資訊206之間的相互關係。結果處理器204可修改與該組標記相關聯的該機率，藉此改善或精製透過電腦視覺分析產生的該組標記，以這種方式，該自動標記處理的精確度能獲得改善。該組標記與相關機率可儲存在標記資料庫208內。在擷取視覺媒體項目202時、在上傳視覺媒體項目202時或稍後時，結果處理器204可精製視覺媒體項目202的該組標記。

雖然結果處理器204在此範例中顯示為單一實體，不過此範例僅為例示。因此，結果處理器204可透過多個處理器或裝置來實施。

根據某些具體實施例，一視覺媒體項目可如本說明書內描述之方式來處理，在該視覺媒體項目不具有與之相關聯的標記時，產生一組標記。因此，若已經手動標記一視覺媒體項目，則該系統不會對該視覺媒體項目產生標記。根據其他具體實施例，一視覺媒體項目可如本說明書內的描述來處理，即使在該視覺媒體項目已經具有與之相關聯的標記時。換言之，即使在先前已經處理過或已經手動標記過一視覺媒體項目，仍可處理該視覺媒體項目。如此，在先前存在的標記內加入自動產生的一組標記，用於一特定視覺媒體項目。

使用者也可隨時手動產生與一視覺媒體項目相關聯的標記，手動產生的標記一開始可具有一特定機率，例如：手動產生的標記可指派給百分之100的機率。

運用與該視覺媒體相關聯的該等標記，視覺媒體搜尋210可透過標記資料庫208過濾並取得視覺媒體。在某些具體實施例內，進一步至少部分根據

與該等標記相關聯的該等機率，可取得該等標記。此外，可使用標記資料庫208自動組織視覺媒體項目。透過一或多個記憶體可實施標記資料庫208，其可在本機上(例如在用戶端上)及/或遠端上(例如在伺服器或雲端上)實施。

第三圖為例示根據本發明具體實施例標記視覺媒體項目的一範例方法之流程圖。一視覺媒體項目可被獲得或識別，例如：可擷取、上傳或從儲存裝置取得該視覺媒體項目，或已經識別。如此在擷取該視覺媒體項目時、在上傳該視覺媒體項目時或稍後時(例如在取得該視覺媒體項目時)，可執行該標記處理。

在302上，獲得透過執行視覺媒體項目的影像內容電腦視覺分析所產生之一組標記，其中該標記集合內每一標記都具有一對應機率(即是權重)，指出該標記精確描述或關於該視覺媒體項目影像內容的可能性。尤其是，執行該視覺媒體項目影像內容的電腦視覺分析，以產生該組標記(以及對應的機率)，或另外，可在獲得該組標記(以及對應的機率)之前執行(例如利用另一個實體)。該組標記可儲存在一標記資料庫內，如此每一標記都關聯於該對應機率來儲存。

可執行電腦視覺分析，偵測該視覺媒體項目影像內容內的物體及/或場景。如此，產生的每一標記都可識別該視覺媒體項目內代表的一物體，或描述該視覺媒體項目的一場景，例如：一物體可包含文字、項目、臉孔、人、特定個體或個體群組，一場景可包含但不受限於一主題、設定、位置、區域或事件。

在304上，獲得與該視覺媒體項目的影像內容獨立而出(例如之外)的資

訊集合。以關聯於該視覺媒體項目的方式儲存該資訊集合或其部分。在某些實例中，該資訊集合或其部分可能關聯於該組標記內一特定標記。底下將進一步詳細說明可獲得的許多種資訊之範例。

在306上，至少部分根據該資訊集合來修改與該組標記至少一部分相關聯的該機率，來精製該組標記。例如：可能增加或降低與該組標記內一標記相關聯的機率，如此可更新該標記資料庫內之一或多個標記相關聯的機率。結果，該已修改過的機率可指出，透過該電腦視覺分析產生的該組標記是否精確，藉此改善該機率與標記的精確度。該組標記可儲存在一標記資料庫內，每一標記都關聯於一對應機率來儲存，該對應機率指出該標記精確描述或屬於該視覺媒體項目影像內容的可能性。以這種方式，物體或場景的偵測機率可獲得改善。

此外，也可利用加入具有一對應機率的額外標記，來修改該組標記，例如：對應至一GPS位置的位置標記或一日期可加入具有百分之100機率的一組標記。類似地，利用從該組標記移除一標記，可修改該組標記，例如：在關聯於一已知標記的一已修改機率低於最低臨界量之門檻，這表示該標記與該視覺媒體項目無關，並且可從該組標記當中移除該標記。

在某些具體實施例內，觀看該視覺媒體項目的使用者可看見該組標記。在其他具體實施例內，具有機率高於臨界量(例如百分之80)的標記可呈現給觀看該視覺媒體項目的使用者，而具有機率低於該臨界量的標記則不會呈現給觀看該視覺媒體項目的使用者。在某些具體實施例內，與該等標記相關聯的該等機率可顯示給使用者。

圖形使用者介面可讓使用者提出一搜尋查詢，以取得有關的視覺媒體

項目。透過裝置，例如行動裝置，以及遠端的伺服器，例如網路伺服器，可呈現這種圖形使用者介面。在某些具體實施例內，當使用者輸入一搜尋查詢或一部分時，該圖形使用者介面可呈現建議的查詢。這種建議的查詢可至少部分根據與該視覺媒體及/或一搜尋內容相關聯的該等標記所產生。該搜尋內容包含但不受限於該使用者所存取的用戶端裝置位置及/或與該使用者相關聯的使用者資訊，例如使用者簡介內保有的資訊。

在某些具體實施例內，以關於一特定資料來源、資料夾或專輯之方式來處理該搜尋查詢。在其他具體實施例內，以關於多個資料來源、資料夾或專輯之方式來處理該搜尋查詢。搜尋視覺媒體項目的該(等)資料來源可經過統計設置，或另外可為使用者可選擇或可設置的。

當使用者執行視覺媒體項目的搜尋時，可能搜尋到與該視覺媒體項目相關聯的該等標記。在某些具體實施例內，至少部分根據與該等視覺媒體項目的標記相關聯之該等機率，一搜尋引擎可識別最有可能與該使用者搜尋查詢有關的這些視覺媒體項目。尤其是，該搜尋引擎可考慮這些標記與該使用者搜尋查詢的詞彙相符或與該使用者搜尋查詢相關的該等機率。因此，至少部分根據與要屬於該使用者搜尋查詢的該等標記相關聯之該等機率，可排名該等已識別的視覺媒體項目，並且根據該排名呈現。

在某些具體實施例內，至少部分根據該等對應標記，自動組織(例如索引)該等視覺媒體項目。此外，該組織可進一步至少部分根據與該等標記相關聯的機率。根據許多標記或中繼資料，包含但不受限於時間、事件、地點、人、場景、物體、文字及/或類似體，可執行視覺媒體項目的自動組織。這種組織可幫助在瀏覽期間有效率取得該等視覺媒體項目。此外，該組織

可讓使用者辨識特定資料夾，透過它執行視覺媒體搜尋。

本說明書描述的該處理過程可在行動裝置及/或遠端伺服器上執行。在某些實施當中，至少部分處理過程可由該行動裝置上安裝的軟體應用程式執行。該視覺媒體以及相關聯的中繼資料可儲存在一或多個位置上，其可包含但不受限於一或多個行動裝置、電腦記憶體、儲存裝置、服務供應商及/或網路服務。

視覺媒體項目的影像內容之電腦視覺分析

在一視覺媒體項目的影像內容上，使用許多電腦視覺分析技術以及軟體工具，可執行電腦視覺分析。電腦視覺分析技術以及工具可分析視覺媒體項目的該影像內容，如此在該視覺媒體項目內產生識別物體及/或場景的一組標記。底下將描述這種電腦視覺分析技術與工具的範例。

在該視覺媒體項目上可執行文字辨識，以透過光學字元辨識(OCR，optical character recognition)來識別文字字元，例如：可確定文字位置及/或定義內含該文字的區域之邊框。文字辨識也可執行來辨識文字字元內的字，例如「海灘」、「山脈」或「舊金山」。NLP可執行來幫助文字辨識。

也可執行條碼偵測，來偵測視覺影像的影像內容內之條碼。此外，可執行條碼辨識，來確定條碼編號或者執行條碼轉譯。

臉部分析可執行來偵測臉部(例如一人)或臉部特徵，例如微笑、皺眉、眨眼、紅眼或視覺媒體項目內臉部偵測的方位。此外，可確定視覺媒體項目內臉部(或人)的數量、該視覺媒體項目之內臉部(或人)的位置及/或定義其中該(等)臉部(或人)所在的該視覺媒體項目區域之邊框。

臉部辨識也可執行來辨識特定臉孔當成特定個體(例如John、Sarah等

等)，機率可指出經由臉部分析確定識別該個體的可能性。

物體辨識也可執行來識別該視覺媒體項目內的物體，例如：可確定物體位置及/或定義內含該物體的該視覺媒體項目區域之邊框。物體可包含項目或實體，例如汽車或自行車。此外，物體可包含人，例如個體或個體群組。機率可指出正確識別出該視覺媒體項目內該物體的可能性。

場景辨識也可執行來識別一視覺媒體項目內的場景，例如：一場景可辨識為一海灘或山脈。此外，機率指出該已識別場景確實描繪出該視覺媒體項目的該影像內容之可能性。

顏色分析可執行來偵測該影像內一或多種顏色，尤其是，該顏色分析可偵測該影像內容內的主宰顏色。在某些具體實施例內，該顏色分析可確定該影像內容內許多顏色的位置。再者，該顏色分析可決定該視覺媒體項目內一物體具有特定顏色，例如：該顏色分析可決定該視覺媒體項目內含一輛紅色汽車。

品質分析也可執行來決定該視覺媒體項目的品質，例如：該品質分析可決定相片是否模糊或對比有限。該品質分析可用來評估透過該電腦視覺分析產生的一標記精確描述該對應物體或場景之機率。在此該品質分析決定該視覺媒體項目內一影像的品質不高，已經產生的一或多個標記可指派一低精確度機率，例如：在相片明顯模糊之處，則已知標記準確的機率可設定為20%。

視覺媒體項目的影像內容以外之資訊

可獲得許多種資訊，以精製透過電腦視覺分析產生的一組標記。尤其是，一或多種資訊可分開或結合應用，以適當修改透過電腦視覺分析產生

的一組標記(例如利用修改與該一組標記至少一部分相關聯的機率)。底下將進一步詳細說明可套用來修改標記機率的資訊類型之範例。

中繼資料

當透過相機拍攝例如相片的視覺媒體時，該視覺媒體通常以中繼資料自動標記，例如時間或位置。不過，隨著相片複製、裁切或轉換，此中繼資料通常會遺失。

根據許多具體實施例，可利用由包含相機的行動裝置自動擷取之中繼資料，進一步精製透過電腦視覺分析產生的自動標記。該中繼資料可包含但不受限於擷取該視覺媒體項目的特定地理位置、日期及/或時間之指示，例如：該地理位置可包含一特定城市、州及/或國家。另一個例如，該日期可識別特定日、月及/或年。該中繼資料也指出在該行動裝置擷取該視覺媒體項目時，該行動裝置附近的物體或個體之識別。

該中繼資料也可包含可交換式影像檔案(EXIF，Exchangeable Image File)資料。包含數位相機的行動裝置通常會隨每一影像拍攝時記錄EXIF資料，該EXIF資料可指出確切的相片拍攝時間與日期，此外該EXIF資料也可指示資訊，例如該影像的方位、相機類型以及設定，例如光圈、焦長、影像拍攝期間閃光燈是否開啟以及影像拍攝期間閃光燈是否閃爍。

此外，在一行動裝置擷取該視覺媒體項目時，透過該行動裝置的一或多個感應器可獲取中繼資料。這種感應器可包含例如紅外線(IR，Infra-red)感應器、射頻識別碼(RFID，Radio Frequency Identifier)感應器、加速度計及/或全球定位系統(GPS，Global Positioning System)感應器。如此，該中繼資料可指出該行動裝置的方位及/或透過該行動裝置偵測的聲音。

感應器

第四圖為根據許多具體實施例，使用透過行動裝置感應器所獲得的中繼資料之圖式。如此範例中所示，具有相機的第一行動裝置402可拍攝視覺媒體項目，例如相片或視頻短片。其他物體或裝置可發出信號，可用來獨一識別該等物體或裝置。運用這種信號可確認透過電腦視覺分析產生的標記是否精確，尤其是，該等信號可用來「提高」與許多標記相關聯的機率。

例如：第一行動裝置402可拍攝人物A的第一相片，而該人物具有第二行動裝置404發出第一信號，例如IR或RFID信號。第一行動裝置402可在拍攝該第一相片時感應到該第一信號。在透過電腦視覺分析產生一組標記之處，該等標記之一者可具有62%的機率識別出人物A。根據該第一信號，第一行動裝置402可感應第二行動裝置404為人物A的行動電話。結果，與此標記相關聯的該機率可從62%提高到100%。如此，已經執行的臉部辨識結果為「提高」。因此，可改善在視覺媒體項目內個體以及對應標記產生的偵測精確度。

針對另一個範例，第一行動裝置402可拍攝物體406的第二相片，而物體發出第二信號，例如RFID信號。第一行動裝置402可在拍攝該第二相片時感應到該第二信號。因為該第二信號指出物體406為紅綠燈，該第二信號可用來改善與識別物體406的標記相關聯之機率，例如：在該標記先前以50%機率識別一根電線桿並且該第二信號指示物體406為紅綠燈之處，與該標記相關聯的該機率可下降至20%的機率。如此，在物體透過RFID標記或其他技術發出信號的許多環境下，可改善物體偵測精確度。

位置

全球定位系統(GPS，Global Positioning System)可決定GPS座標，透過一行動裝置的相機拍攝視覺媒體項目時，識別該行動裝置的位置。該等GPS座標可用來精製透過電腦視覺分析產生的一組標記，尤其是，至少部分根據該等GPS座標，可改善與至少該組標記之一者相關聯的該等機率，例如，該等GPS座標可轉譯成一乘法器，可乘上與一標記相關聯的該機率。

擷取該視覺媒體項目的位置可用來改善場景或物體辨識精確度的機率，例如：在視覺媒體項目為海灘或其他地標的相片之情況下，電腦視覺分析可確定與標記「海灘」相關聯的該機率為百分之56。若該位置靠近海洋，則與該標記「海灘」相關聯的該機率可提高到百分之80。針對另一個範例，在透過電腦視覺分析所產生的標記為「棒球」並且該位置為運動場之情況下，與該標記相關聯的該機率提高。因此，至少部分根據該視覺媒體項目所擷取的位置，可提高場景或物體辨識精確度的機率(例如提高一特定量)。

第五圖為根據本發明不同的具體實施例，使用座標來修改標記機率的一範例系統之方塊圖。如此範例內所示，可使用地圖確定該等座標所識別為該地圖上哪個位置。至少部分根據該(等)標籤是否可能精確給予視覺媒體項目被截取的特定位置，經由電腦視覺分析系統所產生的該組標簽中的至少一者可被修改，例如：該等GPS座標可識別碼頭、海洋、海灘、街道或星巴克咖啡店。若像是相片的視覺媒體項目是在街道上拍攝，則由電腦視覺分析系統標記的該物體為汽車之可能性要高出海洋。

EXIF資料與標記

在特定時間臨界期間之內已經透過行動裝置擷取二或多個視覺媒體項

目之情況下，擷取一視覺媒體項目時的日期及/或時間就特別有用。尤其是，若該EXIF資料或其他中繼資料指出，在一特定預定時間臨界期間之內已經從另一個項目擷取兩個視覺媒體項目，則有可能該等兩個視覺媒體項目之一者的標記與另一個媒體項目相關。

在第一個視覺媒體項目已經由使用者手動標記一或多個標記，並且該使用者並未手動標記該等視覺媒體項目的第二個之情況下，可獲得與該第一視覺媒體項目相關聯的該等標記，進一步精製透過電腦視覺分析產生的該第二視覺媒體項目之該組標記，例如：在已經用事件「Mary的婚禮」來標記第一相片之情況下，則在5分鐘之後拍攝的第二相片也可能有關「Mary的婚禮」。如此，可使用第一視覺媒體項目的該手動產生標記，提高第二視覺媒體項目的電腦視覺分析結果。

類似地，在該等視覺媒體項目之一者具有相當高機率(例如百分之90)的標記之情況下，該標記在精製該等視覺媒體項目之第二者的標記時就有用處，其中該等視覺媒體項目的兩者彼此都已經在一特定時間臨界期間之內擷取。例如：考量兩張相隔5分鐘拍攝的相片，但是第二張相片稍微模糊，其中電腦產生用於該第一相片的第一組標記包含機率百分之90的一標記「人物A」，並且電腦產生用於該第二相片的第二組標記包含機率百分之50的一標記「人物A」，則針對該第二相片，與該標記「人物A」相關聯的該機率可提高到百分之90。

如上面所討論，與在特定時間臨界期間之內彼此擷取的視覺媒體項目相關聯之標記機率可傳播或結合，該等標記可自動或手動產生，例如：該等標記可識別場景、事件、人物或其他合適的標籤。

相似度分析

相似度分析可執行來識別二或多個看起來類似的視覺媒體項目，尤其是，電腦視覺分析可識別在該視覺媒體項目與其他媒體項目之間所分享的視覺特性。如此，可能識別出描繪相同或類似場景或物體的視覺媒體項目。

在某些具體實施例內，相似度權重可指定給其他視覺媒體項目，其中相似度權重指出另一個視覺媒體項目一定程度類似於該視覺媒體項目。視覺上最類似於該視覺媒體項目的視覺媒體項目集合，可由根據其相似度權重藉由排名視覺媒體項目來識別。在某些具體實施例內，若其相似度權重高於預定相似性門檻，則視為視覺媒體項目在視覺上類似於另一個視覺媒體項目。如此，執行關於一視覺媒體項目與其他視覺媒體項目的電腦視覺分析，來識別視覺上類似於該視覺媒體項目的視覺媒體項目集合。

視覺類似媒體項目集合內的每一類似視覺媒體項目相關聯的一組標記可確定，其中與類似媒體視覺項目相關的該標記集合之每一標記都具有一對應機率。在該一組標記內的一標記具有高機率的精確度之情況下，可修改與該視覺媒體項目相關聯的一對應標記(例如一致或類似標記)之機率(例如利用提高與該視覺媒體項目的該標記相關聯之該機率)。例如：在第一相片已經由使用者用標記「棒球」標記，並且在視覺上類似的第二相片已經透過電腦視覺分析用該標記「棒球」標記成具有百分之50的機率之情況下，該機率會提高到百分之80。因此，若該其他視覺媒體項目在視覺上類似於該特定視覺媒體項目，則可使用該其他視覺媒體項目的該等標記(與相關機率)進一步提高特定視覺媒體項目的電腦視覺分析結果。

第六圖為根據許多具體實施例，視覺上類似視覺項目之識別之圖式。

在此範例中，使用者可用文字「Eiger」標記第一相片，這是阿爾卑斯山當中的一座山，該第一相片與第二相片非常類似。如此範例內所示，電腦視覺分析可產生相似度權重0.85，及一標籤「Eiger」其具有該第二相片0.85的對應機率。由於相似於該第一標記相片，所以可提高與該第二相片相關聯的該標記機率(例如至0.95)。

特徵學習與使用者標記

在電腦視覺分析期間，未監督及/或監督下的特徵學習可被執行來訓練電腦實施模型識別特定物體或場景。第七圖為未監督特徵學習的一範例應用之方塊圖。如此範例內所示，使用者可使用標記「滑翔翼」標記一或多張相片。利用執行電腦視覺分析，該系統可決定該(等)手動標記的相片與其他未標記的相片是否共享視覺特徵。在確定手動標記的視覺媒體項目看起來類似於其他未標記視覺媒體項目之情況下，該系統可使用該(等)手動產生的標記，根據其視覺特徵來改善其偵測物體或場景的能力。尤其是，該系統可保存一物體及/或場景清單，其中每一物體或場景都與該對應物體或場景的視覺特徵相關聯。在此範例中，該系統可改善其能力，以精準辨識描述一滑翔翼的視覺媒體項目，並且可改善透過電腦視覺分析所產生的該標記「滑翔翼」之機率權重。如此，該系統可改善其通過使用與其他視覺媒體項目有關聯的手動產生標記，透過電腦視覺分析辨識未標記視覺媒體項目內一物體或場景之能力。

與手動產生標記相關聯之標記密度也用來修改標記機率，例如：在使用者以最小臨界頻率使用一特定標記或標記組合之情況下，至少部分根據該標記或標記組合所使用的頻率，來讓標記機率可被修改。

社交標記與評論

社交標記與評論也可用來精製透過電腦視覺分析產生的標記，例如：若使用者張貼關於一視覺媒體項目的評論，或將另一個標記加入一視覺媒體項目，這可用來精製透過電腦視覺分析所產生的標記。

使用者資料

例如使用者簡介內保留的使用者資料可用來進一步精製電腦視覺分析結果，例如：在該簡介指出該使用者的興趣為釣魚，並且已經將透過電腦視覺分析產生給一視覺媒體項目的標記「釣魚」，以百分之40的機率賦予該等使用者相片之一者的情況下，由於使用者非常有可能會去釣魚，所以機率可能會升高。因此，使用者簡介資料，例如使用者興趣，可用來修改已經透過電腦視覺分析產生的標記之機率。

行事曆

在精製透過電腦視覺分析產生的標記當中，行事曆也很有用。行事曆可指出一或多個事件的日期、時間及/或位置，該行事曆可為使用者、使用者群組的行事曆，或可公開的其他行事曆。自該影像內容獨立出的該資訊之至少一部分可從行事曆取得。在某些具體實施例內，在一或多個行事曆內查找的動作可被執行來識別該日期上的事件及/或擷取該媒體項目的時間。一旦已經識別一事件，則至少部分根據該已識別的事件，可修改與該電腦產生標記集合內標記相關聯的機率，尤其是，若已經在該事件期間擷取該視覺媒體項目，則與一事件相關聯的一標記之機率可被提高。這種事件可為一排程事件、生日或其他節日，例如：在聖誕節期間，可將與該標記「聖誕樹」相關聯的該機率提高一特定量。

例如：假設該電腦視覺分析系統產生以百分之50的機率識別一場棒球賽之一標記給一特定相片，則在該使用者的行事曆上查找該相片拍攝的日期，指出該使用者計畫在當日前往觀看棒球賽。在此實例當中，與該標記「棒球賽」相關聯的該機率可提高到百分之90。

針對另一個範例，假設該電腦視覺分析系統產生具有百分之50機率的一標記「雪」，若使用者的行事曆指出該使用者在當天前往滑雪，則該機率可提高到百分之70。

標記擴充

一旦透過電腦視覺分析產生的標記或手動產生的標記，標記可自動產生來擴充，例如：若一場景已經標記為「狗」，則可自動產生額外的標記「動物」。運用許多資料庫，例如字典、辭典或其他資料庫，可執行自動標記擴充。

使用者回饋

根據許多具體實施例，圖形使用者介面可提供使用者輸入功能，讓使用者提供屬於一對應標記精確度的回饋，例如：該輸入功能可讓該使用者指出在第一狀態下，該標記精確，並且在第二狀態下，該標記不精確。這種回饋可用來修改與該對應標記相關聯的機率。在此範例中，若該使用者指出該標記精確，則該機率可改善為百分之100，或若該使用者指出該標記不精確，則可為百分之0。該回饋也可用來訓練該電腦視覺分析系統，如上述。

此外，可實施群眾外包，進一步提高自動產生標記的精確度。群眾外包的結果可進一步改善自動系統特徵學習。

透過所揭示的具體實施例，可提高執行自動標記的精確度。透過與電腦視覺分析產生的一組標記，利用改善與該組標記相關聯的機率之精確度，至少部分可達成此目的。所以，使用者可提出一搜尋查詢來搜尋視覺媒體項目，即使使用者先前並未手動標記這些媒體項目。由於執行自動標記有較大精確度，所以該使用者的搜尋查詢更可靠。例如：使用者提出例如「找出三年前我跟我的小狗在海灘上拍攝的相片」或「顯示有紅色汽車的所有相片」。

最重要請注意到，上述具體實施例僅為例示，例如：雖然本說明書內描述的許多範例為提高與標記相關聯的機率，不過最重要請注意到，該機率也會降低。更進一步，已經分析的額外資訊可能導致刪除與視覺媒體項目相關聯的標記，以及增加與視覺媒體項目相關聯的標記。

網路

網路可耦合裝置，如此可進行通訊，例如在伺服器與用戶端裝置或其他種裝置之間，包含例如透過無線網路耦合的無線裝置之間。網路也可包含大量儲存裝置，例如網路附加儲存裝置(NAS，network attached storage)、一儲存區域網路(SAN，storage area network)或電腦或機器可讀取媒體的其他形式。網路可包含網際網路、一或多個區域網路(LAN，local area network)、一或多個廣域網路(WAN，wide area network)、有線型連線、無線型連線或這些的任意組合。類似地，在大型網路內可交互運作，例如可採用不同架構或可符合或相容於不同通訊協定之子網路。許多種裝置可例如用來提供交互操作能力給不同的架構或協定。針對一個例示範例，路由器可提供個別與獨立LAN之間的連結。

一通訊連結或通道可包含例如類比電話線，如雙絞線對、同軸纜線、全部或部分數位線，包含T1、T2、T3或T4型線、整體服務數位網路(ISDN，Integrated Services Digital Network)、數位用戶線(DSL，Digital Subscriber Line)、包含衛星連結的無線連結或其他通訊連結或通道，例如許多精通技術人士所熟知的。更進一步，電腦裝置或其他相關電子裝置可遠端耦合至網路，例如透過電話線或連結。

內容分散網路

一分散式系統可包含一內容分散網路。「內容遞送網路(content delivery network)」或「內容分散網路(content distribution network)」(CDN)一般稱為分散式內容遞送系統，其包含用網路連結的電腦或計算裝置之集合。CDN可運用軟體、系統、通訊協定或技術促進許多服務，例如儲存、快取、內容通訊或串流媒體或應用。服務也可利用配套技術，包含但不受限於「雲端計算」、分散式儲存、DNS要求處置、調配、信號監控與回報、內容定向、個性化或商業智慧。CDN也可讓一實體完整或部分操作或管理另一個站台基礎設施。

點對點網路

相較於運用專屬裝置，例如專屬伺服器的網路，點對點(或P2P)網路可運用網路參與者的計算能力與頻寬；不過，某些網路則同時採用這兩種。P2P網路通常用於透過特別安排或組態來耦合節點，點對點網路可運用某些節點，其可操作當成「用戶端」與「伺服器」。

無線網路

無線網路可使用網路連結用戶端裝置，無線網路可運用單機特別網

路、網狀網路、無線區域網路(WLAN)、行動網路等等。

無線網路另包含終端機、閘道器、路由器或由無線電連結耦合的裝置等等之系統，其可自由、隨機移動或任意組織，如此網路拓撲可隨時迅速改變。無線網路可進一步運用複數個網路存取技術，包含長期演進(LTE，Long Term Evolution)、WLAN、無線路由器(WR，Wireless Router)網、第二、第三或第四代(2G、3G或4G)行動技術等等。無線存取技術可讓廣泛涵蓋裝置，如具有非常高的機動性的用戶端裝置。

● 例如：網路可啟用RF或透過一或多種網路存取技術的無線通訊，例如全球移動通訊系統(GSM，Global System for Mobile communication)、通用移動電信系統(UMTS，Universal Mobile Telecommunications System)、通用封包無線服務(GPRS，General Packet Radio Services)、增強數據GSM環境(EDGE)、3GPP長期演進(LTE，Long Term Evolution)、進階型LTE、寬頻分碼多工存取(WCDMA，Wideband Code Division Multiple Access)、藍牙、802.11b/g/n等等。無線網路可包含任何虛擬無線通訊機構類型，藉其信號可在裝置之間通訊，例如用戶端裝置或計算裝置、網路之間或之內等等。

網際網路通訊協定

透過網路，例如參與數位通訊網路的網路，通訊的信號封包可與一或多個通訊協定相符或相容。所運用的傳訊格式或通訊協定包含例如TCP/IP、UDP、DECnet、NetBEUI、IPX、Appletalk等等。網際網路通訊協定(IP，Internet Protocol)的版本包含IPv4或IPv6。

網際網路為網路的分散式全球網路，網際網路包含LAN、WAN、無線網路或允許信號封包在LAN之間通訊的長途公共網路。信號封包可在網路

節點之間通訊，例如運用本機網路地址通訊至一或多個站台。信號封包可例如透過網際網路，從使用者位置透過連結至網際網路的存取點來通訊。類似地，信號封包可透過網路節點，轉送到透過網路存取節點連結至網路的目標站台。透過網際網路通訊的信號封包可例如透過閘道器、伺服器等等的路徑繞送，其可根據目標位址以及網路路徑可用性，將該信號封包繞送至該目標位址。

網路架構

所揭示的具體實施例可在各種計算環境(computing context)內實施。第八圖為示範例具體實施例的示意圖。在架構或組件種類方面可改變的其他具體實施例也可用來包含在所主張的主題之內。現實設想中，使用者與各種網路環境互動。如所示，第八圖包含多種網路，例如LAN/WAN 705以及無線網路700、多種裝置，例如用戶端裝置701-704，以及多種伺服器，例如內容伺服器707以及搜尋伺服器706。該等伺服器也可包含廣告伺服器(未顯示)。如此範例內所示，用戶端裝置701-704可包含一或多個行動裝置702、703、704。用戶端裝置701-704可例如透過任何電腦類型(例如桌上型、膝上型、平板等等)，實施媒體計算平台(例如有線電視與衛星機上盒)、手持式計算裝置(例如PDA)、行動電話或任何其他種計算或通訊平台。

所揭示的具體實施例可用某些集中方式來實施，這在第八圖內利用伺服器707來表示，這可對應至多重分散式裝置與資料儲存裝置。伺服器707及/或對應的資料儲存裝置可儲存使用者帳號資料、使用者資訊及/或內容。

伺服器

計算裝置可傳送或接收信號，例如透過有線或無線網路，或可處理或

儲存信號，例如在實體記憶狀態的記憶體內，因此可當成伺服器來操作。如此，可當成伺服器的裝置包括例如專用機櫃型伺服器、桌上型電腦、膝上型電腦、機上盒、結合許多功能的整合裝置，例如上述裝置的二或多種功能或等等。

伺服器的設置或能力可非常廣泛，但是伺服器一般可包括一或多個中央處理單元與記憶體。伺服器也可包括一或多個大量儲存裝置、一或多個電源供應器、一或多個有線或無線網路介面、一或多個輸入/輸出介面或一或多個作業系統，例如Windows Server、Mac OS X、Unix、Linux、FreeBSD或等等。

內容伺服器

內容伺服器可包含一裝置，其包含一組態透過網路提供內容給另一個裝置。內容伺服器可例如主控一站台，例如社交網路站台，範例包含但不限於Flicker、Twitter、Facebook、LinkedIn或個體使用者站台(例如部落格、vlog、線上約會網站等等)。內容伺服器也可主控許多其他站台，包含但不限於商業站台、教育站台、字典站台、百科全書站台、維基、財經站台、政府站台等等。

內容伺服器可進一步提供許多服務，包含但不限於網路服務、第三方服務、音頻服務、視頻服務、電子郵件服務、即時傳訊(IM，instant messaging)服務、SMS服務、MMS服務、FTP服務、IP語音(VOIP，voice over IP)服務、行事曆服務、相片服務等等。內容的範例可包含文字、影像、音頻、視頻等等，其可用物理信號方式來處理，如電子信號，或可依照物理狀態儲存在記憶體內。

可操作當成內容伺服器的裝置範例包含桌上型電腦、多處理器系統、微處理器型或可程式消費型電子裝置等等。

用戶端裝置

第九圖為例示其中可實施許多具體實施例的一用戶端裝置範例具體實施例之示意圖。用戶端裝置可包含能夠例如透過有線或無線網路，傳送或接收信號的計算裝置。用戶端裝置可例如包含桌上型電腦或可攜式裝置，例如行動電話、智慧型電話、顯示型呼叫器、射頻(RF，radio frequency)裝置、紅外線(IR)裝置、個體數位助理(PDA，Personal Digital Assistant)、手持式電腦、平板電腦、膝上型電腦、機上盒、穿戴式電腦、結合許多功能，例如前述裝置功能的整合式裝置等等。可攜式裝置也稱為行動裝置或手持式裝置。

如此範例內所示，用戶端裝置800可包含一或多個中央處理單元(CPU，central processing unit) 822，其可透過連線824耦合至電源供應器826和記憶體830。記憶體830可包括隨機存取記憶體(RAM，random access memory) 832以及唯讀記憶體(ROM，read only memory) 834。ROM 834可包含基本輸入/輸出系統(BIOS，basic input/output system) 840。

RAM 832可包含作業系統841。尤其是，用戶端裝置可包含或可執行許多操作系統，包含個人電腦作業系統，例如Windows、iOS或Linux，或行動作業系統，例如iOS、Android或Windows Mobile等等。用戶端裝置800 也可包含或可執行許多可能的應用程式842 (顯示於RAM 832內)，例如用戶端軟體應用程式，如傳訊程式843、啟用與其他裝置的通訊，如通訊一或多種訊息，例如透過電子郵件、簡訊服務(SMS，short message service)或多媒體訊

息服務(MMS，multimedia message service)，包含透過網路，例如社交網路，包含例如Facebook、LinkedIn、Twitter、Flickr或Google，在此僅提供一些範例。用戶端裝置800也可包含或執行應用程式來通訊內容，例如文字內容、多媒體內容等等，其可儲存在資料儲存裝置844內。用戶端裝置也包含或執行一應用程式，例如瀏覽器845，來執行許多可能的工作，例如瀏覽、搜尋、播放許多內容，包含本機儲存或串流的視頻或遊戲(例如夢幻運動聯盟)。

● 用戶端裝置800可透過一或多個介面傳送或接收信號。如此範例內所示，用戶端裝置800可包含一或多個網路介面850。用戶端裝置800可包含一音頻介面852。此外，用戶端裝置800可包含一顯示器854以及一照明器858。用戶端裝置800可進一步包含一輸入/輸出介面860，以及支援觸感反饋技術的觸覺介面862。

● 用戶端裝置800可傳輸與偵測圖案、影像或信號，例如透過該(等)介面的紅外線信號。這種圖案、影像或信號可獨一識別一特定物體(例如個體或項目)，例如：用戶端裝置800可傳輸識別一特定物體的紅外線閃爍圖案，以及偵測一紅外線閃爍圖案。

用戶端裝置800可就性能或功能而不同，主張的主題用於涵蓋多種潛在變化，例如：行動電話可包含鍵盤856，例如數字鍵盤，或功能受限的顯示器，例如單色液晶顯示器(LCD，liquid crystal display)，用來顯示文字。不過，相較之下，針對另一個範例，可上網的用戶端裝置可包含一或多個實體或虛擬鍵盤、大量儲存裝置、一或多個加速度計、一或多個陀螺儀、全球定位系統(GPS) 864或其他位置識別能力，或具有高功能的顯示器，例如觸控彩色2D或3D顯示器。上述用來例示所主張主題要包含廣泛可能的特色

或能力。

根據許多具體實施例，可使用廣泛不同的技術來獲得輸入。例如：從使用者與本機應用程式的互動，例如行動裝置上的行動應用程式、網站或網路式應用程式或服務，透過圖形使用者介面可獲得下載或啟動一應用程式的輸入，並且可達成使用任何許多已知機制來從使用者獲得資訊。不過，應了解這種從使用者獲得輸入的方法僅為範例，可用許多其他方式獲得該輸入。

第十圖例示典型電腦系統，當適當設置或設計時，可當成一系統，透過其可實施許多具體實施例。電腦系統1200包含任意數量的CPU 1202，其耦合至儲存裝置，包含主要儲存裝置1206(一般為RAM)、主要儲存裝置1204(一般為ROM)。CPU 1202可為許多類型，包含微控制器以及微處理器，例如可程式編輯裝置(例如CPLD和FPGA)以及不可程式編輯裝置，例如閘道陣列ASIC或一般用途微處理器。如業界內所熟知，主要儲存裝置1204用來單向傳輸資料與指令至該CPU，並且主要儲存裝置1206通常以雙向方式傳輸資料與指令。這兩種主要儲存裝置都可包含任何合適的電腦可讀取媒體，例如上述的媒體。大量儲存裝置1208也雙向耦合至CPU 1202，並且提供額外資料儲存容量，並且可包含任何上述電腦可讀取媒體。大量儲存裝置1208可用來儲存程式、資料等等，並且通常為次要儲存媒體，例如硬碟。在適當情況下，保存在大量儲存裝置1208之內的資訊可用標準方式與主要儲存裝置1206一部分合併當成虛擬記憶體。特定大量儲存裝置，例如CD-ROM 1214，也可單向將資料傳遞給該CPU。

CPU 1202也耦合至介面1210，其連接至一或多個輸入/輸出裝置，例如

視訊顯示器、軌跡球、滑鼠、鍵盤、麥克風、觸控顯示器、傳感器讀卡機、磁帶或紙帶讀取器、平板、觸控筆、語音或手寫辨識器或其他已知的輸入裝置，例如其他電腦。最後，CPU 1202可使用外部連線，一般顯示為1212，選擇性耦合至外部裝置，例如資料庫、電腦或通訊網路。運用這種連線，可預期該CPU可在執行本說明書所描述的方法步驟之下，從網路接收資訊，或可輸出資訊至網路。

不管系統的組態為何，可運用一多個記憶體或記憶體模組，設置來儲存資料、一般處理操作用的程式指令及/或本說明書內描述的發明技術。例如，該等程式指令可控制作業系統及/或一或多個應用程式的操作。記憶體也可設置成儲存指令來執行所揭示方法、要關聯於所揭示方法顯示的圖形使用者介面等等。

因為這種資訊與程式指令可用來實施本說明書內描述的系統/方法，因此所揭示的具體實施例係關於包含程式指令、狀態資訊等等的機器可讀取媒體，用來執行本說明書內描述的許多操作。機器可讀取媒體的範例包含但不限於像是硬碟、軟碟和磁帶的磁性媒體、像是CD ROM光碟的光學媒體、像是光碟的磁性光學媒體以及特別設置成儲存並執行程式指令的硬體裝置，像是ROM和RAM。程式指令的範例包含像是由編譯器產生的機器碼，以及內含可由電腦使用解譯器執行的高階程式碼之檔案。

實施許多具體實施例的電腦程式指令可儲存在任何一種電腦可讀取媒體內，並且可根據許多計算模型來執行，包含主從式模型、點對點模型、單機計算裝置或根據其中此處所描述許多功能性可生效或運用在不同位置上的分散式計算模型。

所揭示的技術可用任何合適的軟體及/或硬體系統組合來實施，例如網路伺服器或桌上型電腦系統。再者，實施許多具體實施例的系統可為可攜式裝置，例如膝上型或行動電話。一設備及/或網路瀏覽器可特別建造用於所要目的，或可為一般用途電腦，其可利用該電腦內儲存的電腦程式及/或資料結構，選擇性啟動或重新設置。此處呈現的處理本質上與任何特定電腦或其他設備無關。尤其是，許多一般用途機器可運用依照此處技術撰寫的電腦程式，或可更加便利來建構更特殊的設備，來執行所揭示的方法步驟。

雖然為了清楚瞭解，本發明已經用某些細節來描述，但在申請專利範圍內可實施特定變更與修改。因此，本具體實施例僅用於說明，並不設限，並且本發明並不受限於此處說明的細節，但是可在附加之申請專利範圍的領域下進行修改。

【符號說明】

【0100】

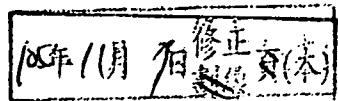
102a 用戶端	204 結果處理器
102b 用戶端	206 其他資訊
102c 用戶端	208 標記資料庫
104 網際網路	210 視覺媒體搜尋
106 伺服器	302-306 流程
108 記憶體	402 第一行動裝置
200 電腦視覺分析	404 第二行動裝置
202 視覺媒體項目	406 物體

105709391

700	無線網路	845	瀏覽器
701-704	用戶端裝置	850	網路介面
705	LAN/WAN(網路)	852	音頻介面
706	搜尋伺服器	854	顯示器
707	內容伺服器	856	鍵盤
800	用戶端裝置	858	照明器
822	中央處理單元	860	輸入/輸出介面
824	連線	862	觸覺介面
826	電源供應器	864	全球定位系統
830	記憶體	1200	電腦系統
832	隨機存取記憶體	1202	處理器
834	唯讀記憶體	1204	主要儲存裝置
840	基本輸入/輸出系統	1206	主要儲存裝置
841	作業系統	1208	大量儲存裝置
842	應用程式	1210	介面
843	傳訊程式	1212	連線
844	資料儲存裝置	1214	CD-ROM

申請專利範圍

1. 一種用以相片及視頻搜尋之方法，包含：



獲得一視覺媒體項目的一組標記，該組標記的每一標記具有一對應的機率，其中該機率指出該標記精確地描述該視覺媒體項目的影像內容的可能性；

獲得獨立於該視覺媒體項目的影像內容之外的一資訊集合；以及至少部分根據該資訊集合，修改至少一部分該組標記的該機率。

2. 如申請專利範圍第 1 項之方法，其中該資訊包含：

在一行動裝置擷取一視覺媒體項目時透過該行動裝置獲取中繼資料。

3. 如申請專利範圍第 1 項之方法，其中該資訊集合包含一特定地理位置的指示、一日期、一時間及/或在一行動裝置擷取該視覺媒體項目時，在該行動裝置附近的物體或個體之識別的指示。

4. 如申請專利範圍第 1 項之方法，進一步包含：

執行有關該視覺媒體項目與其他視覺媒體項目的相似度分析，來識別與該視覺媒體項目類似的一視覺媒體項目集合；

其中所述獲得獨立於該視覺媒體項目的影像內容之外的一資訊集合，包含確定與該視覺媒體項目集合內每一類似視覺媒體項目相關聯之一組標記，與每一類似視覺媒體項目相關聯的該組標記之每一標記具有一對應機率。

5. 如申請專利範圍第 1 項之方法，進一步包含：

確定擷取該視覺媒體項目的一日期或時間；

其中該資訊集合包含一組標記，其關聯於在距離該視覺媒體項目擷取日期及/或時間一段臨界時期之內所擷取的至少另一視覺媒體項目。

6. 如申請專利範圍第 1 項之方法，其中該資訊集合包含已被張貼的關於該視覺媒體項目的一評論。
7. 如申請專利範圍第 1 項之方法，其中所述獲得該資訊集合包含：

從一行事曆取得該資訊的至少一部分，該資訊的至少一部分識別一事件。
8. 如申請專利範圍第 1 項之方法，其中該資訊包含關於至少該組標記之一者的精確度之使用者回饋。
9. 如申請專利範圍第 1 項之方法，進一步包含：

接收一搜尋查詢；
至少部分基於與一複數個視覺媒體項目的標記相關聯之機率，識別與該搜尋查詢相關的該複數個視覺媒體項目；以及
提供該複數個視覺媒體項目以透過一行動裝置顯示；
其中，該複數個視覺媒體項目包含該視覺媒體項目。
10. 如申請專利範圍第 1 項之方法，進一步包含：

識別儲存在一行動裝置上的一複數個視覺媒體項目；
至少部分基於與該複數個視覺媒體項目及對應的機率相關聯之標記，利用該行動裝置的一處理器，組織在一複數個資料夾中的該複數個視覺媒體項目；
其中，該複數個視覺媒體項目包含該視覺媒體項目。
11. 如申請專利範圍第 10 項之方法，進一步包含：

透過該行動裝置，接收該複數個資料夾之一的一選擇。

12. 如申請專利範圍第 1 項之方法，進一步包含：

至少部分根據一日期及/或透過一行動裝置擷取該視覺媒體項目的時間，在一行事曆內執行一查找以獲得該資訊集合的至少一部分。

13. 如申請專利範圍第 12 項之方法，其中該資訊集合的至少一部分識別在該行事曆內的一事件，其中該事件發生在該日期及/或透過該行動裝置擷取該視覺媒體項目的時間；

其中修改該組標記的至少一部分的該機率，包含至少部分根據該事件是否對應至一標記，增加或減少在該組標記中該標記的該機率。

14. 如申請專利範圍第 1 項之方法，其中該資訊集合包含在一行動裝置擷取該視覺媒體項目時，在該行動裝置附近的物體或個體之識別的指示；

其中獲得該資訊集合包含透過該行動裝置擷取的該視覺媒體項目，偵測該物體或個體之該識別。

15. 如申請專利範圍第 1 項之方法，其中該視覺媒體項目為一相片。

16. 如申請專利範圍第 1 項之方法，其中該視覺媒體項目的該組標記由執行該視覺媒體項目的影像內容的該電腦視覺分析所產生。

17. 一種用以相片及視頻搜尋之設備，包含：

一處理器；以及

一記憶體，該處理器或該記憶體之至少一者適用於：

獲得一視覺媒體項目的一組標記，該組標記的每一標記具有一對應的機率，其中該機率指出該標記精確地描述該視覺媒體項目的影像內容的可能性；

獲得獨立於該視覺媒體項目的影像內容之外的一資訊集合；以及
至少部分根據該資訊集合，修改至少一部分該組標記的該機率。

18. 如申請專利範圍第 17 項之設備，其中該資訊集合包含：

在一行動裝置擷取一視覺媒體項目時透過該行動裝置獲取中繼資
料。

19. 如申請專利範圍第 17 項之設備，其中該資訊集合包含一特定地理位置
的指示、一日期、一時間及/或在一行動裝置擷取該視覺媒體項目時，
在該行動裝置附近的物體或個體之識別的指示。

20. 如申請專利範圍第 17 項之設備，該處理器或該記憶體之至少一者設置
成執行操作，進一步包含：

執行有關該視覺媒體項目與其他視覺媒體項目的相似度分析，來識
別與該視覺媒體項目類似的一視覺媒體項目集合；

其中所述獲得獨立於該視覺媒體項目的影像內容之外的一資訊集
合，包含確定與該視覺媒體項目集合內每一類似視覺媒體項目相關聯
之一組標記，與每一類似視覺媒體項目相關聯的該組標記之每一標記
具有一對應機率。

21. 如申請專利範圍第 17 項之設備，其中該資訊集合包含已被張貼的關於
該視覺媒體項目的一評論。

22. 如申請專利範圍第 17 項之設備，其中所述獲得該資訊集合包含：
從一行事曆取得該資訊的至少一部分，該資訊的至少一部分識別一
事件。

23. 如申請專利範圍第 17 項之設備，其中該視覺媒體項目的該組標記由執

行該視覺媒體項目的影像內容的該電腦視覺分析所產生。

24. 一種其上儲存電腦可讀指令來執行操作的非暫態電腦可讀取儲存媒體，包含：

獲得一視覺媒體項目的一組標記，該組標記的每一標記具有一對應的機率，其中該機率指出該標記精確地描述該視覺媒體項目的影像內容的可能性；

獲得獨立於該視覺媒體項目的影像內容之外的一資訊集合；以及至少部分根據該資訊集合，修改至少一部分該組標記的該機率。

25. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，其中該資訊集合包含：

在一行動裝置擷取一視覺媒體項目時透過該行動裝置獲取中繼資料。

26. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，其中該資訊集合包含一特定地理位置的指示、一日期、一時間及/或在一行動裝置擷取該視覺媒體項目時，在該行動裝置附近的物體或個體之識別的指示。

27. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，另包含：

執行有關該視覺媒體項目與其他視覺媒體項目的相似度分析，來識別與該視覺媒體項目類似的一視覺媒體項目集合；

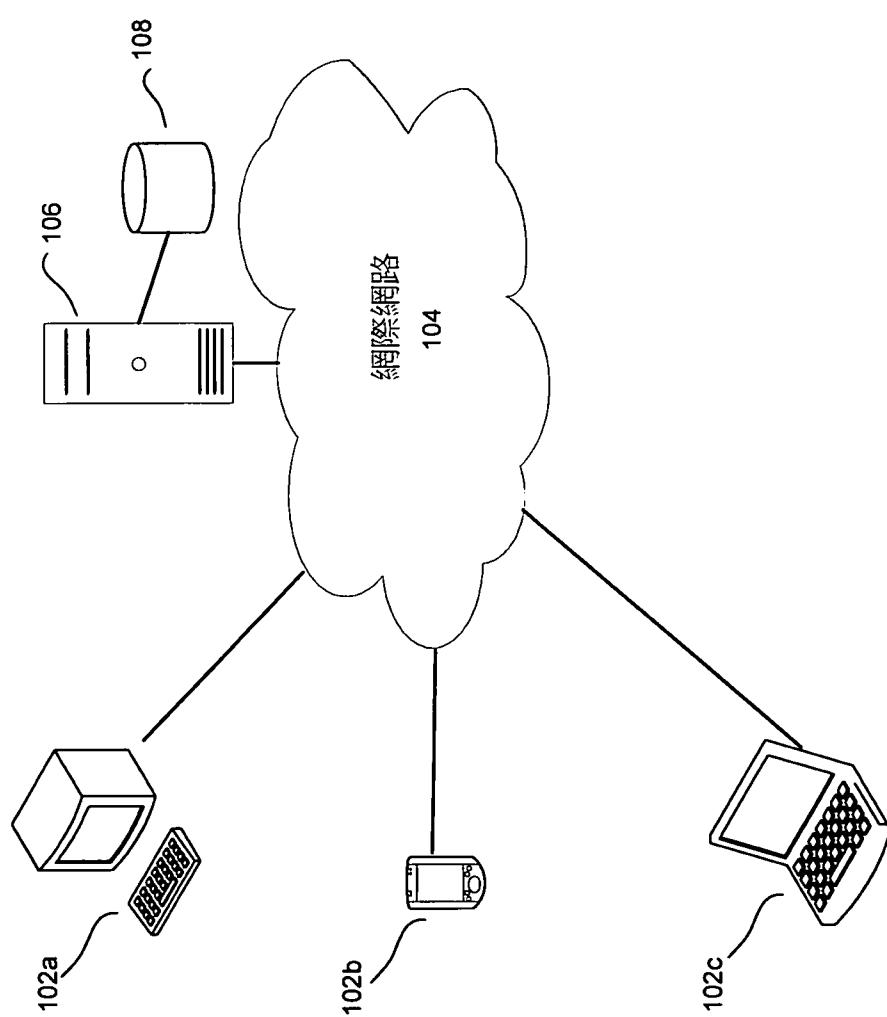
其中獲得獨立於該視覺媒體項目的影像內容之外的一資訊集合，包含確定與該視覺媒體項目集合內每一類似視覺媒體項目相關聯之一組標記，與每一類似視覺媒體項目相關聯的該組標記之每一標記具有一對應機率。

28. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，其中該資訊集合包含已被張貼的關於該視覺媒體項目的一評論。
29. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，其中所述獲得該資訊集合包含：

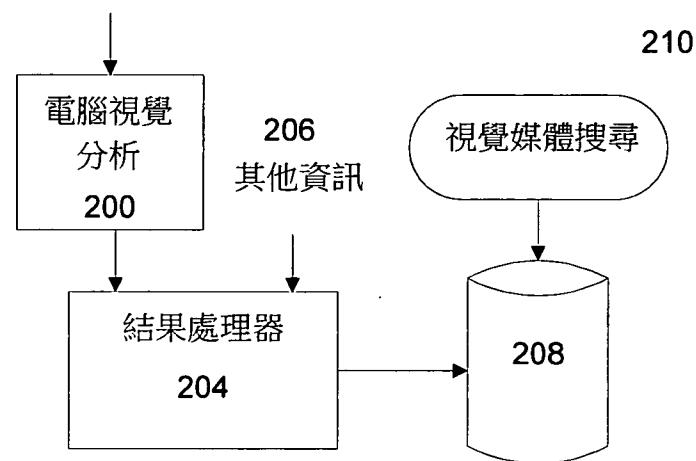
從一行事曆取得該資訊集合的至少一部分，該資訊集合識別一事件。
30. 如申請專利範圍第 24 項之非暫態電腦可讀取儲存媒體，其中該視覺媒體項目的該組標記由執行該視覺媒體項目的影像內容的該電腦視覺分析所產生。

圖式

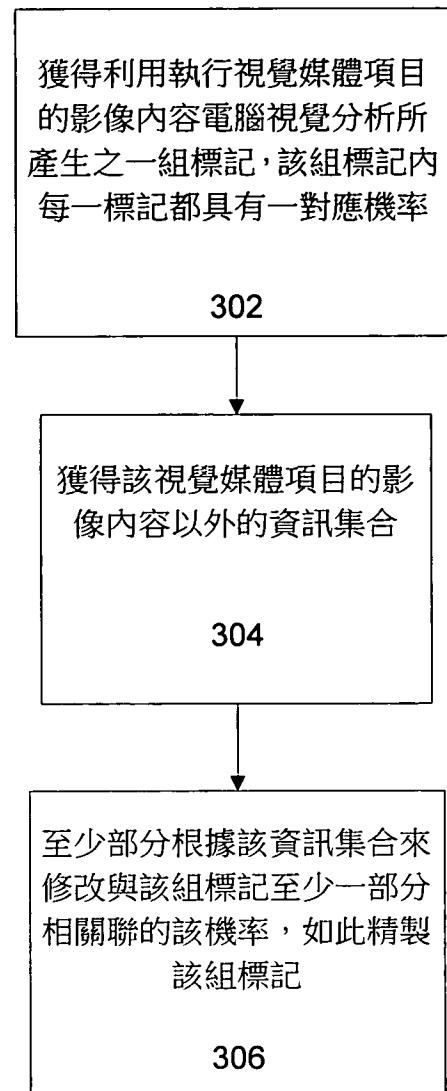
第一圖



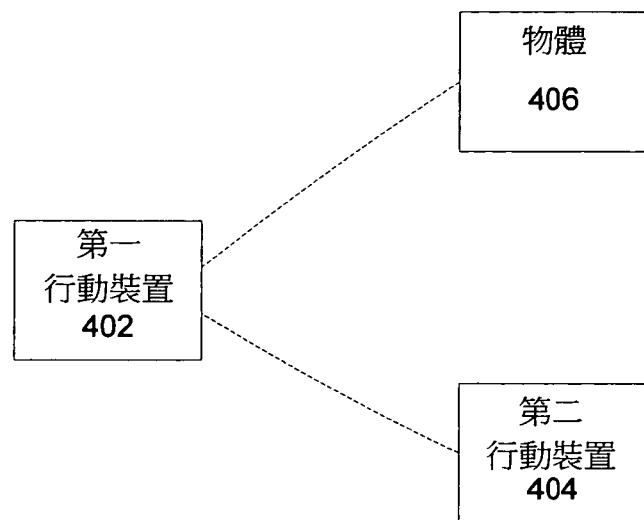
視覺媒體項目202



第二圖



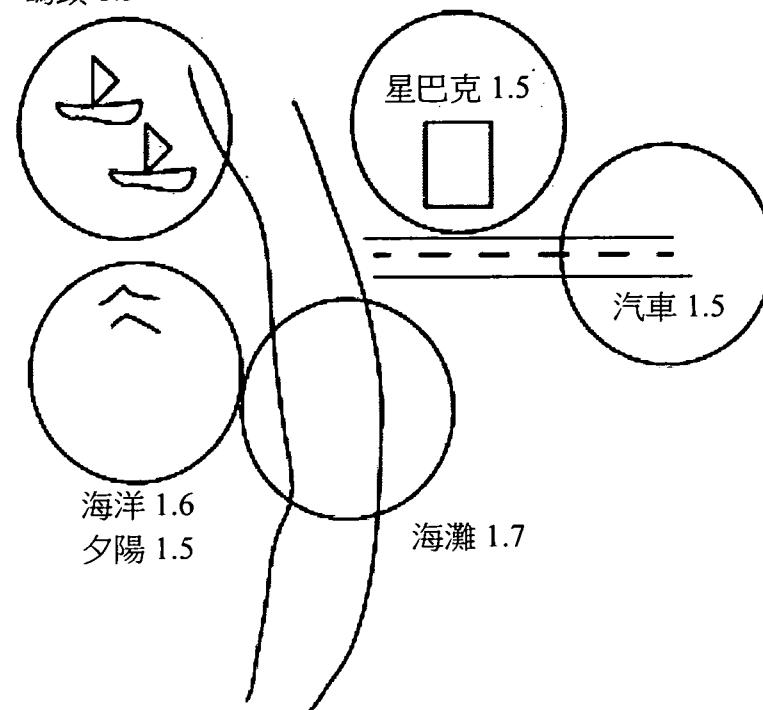
第三圖



第四圖

地理處理器

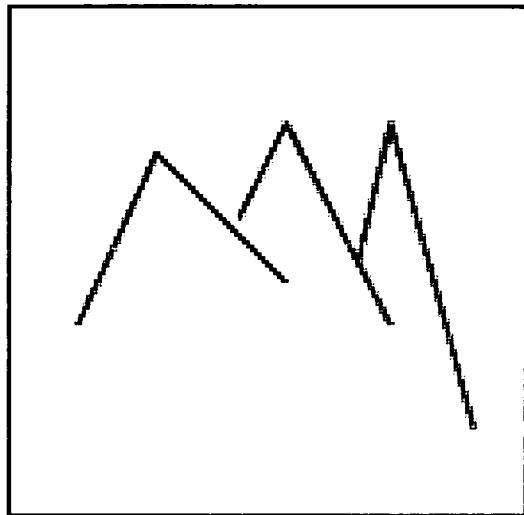
船、小船、碼頭 1.5



第五圖

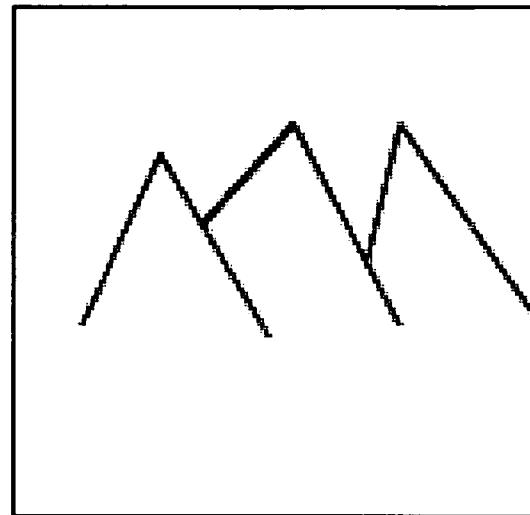
相似度標記處理器

相片



使用者標記：eiger

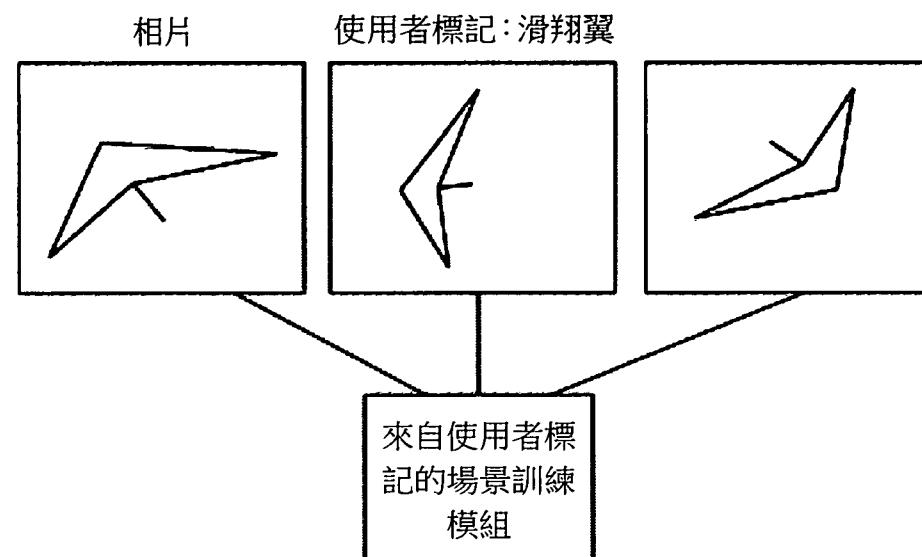
相似的相片



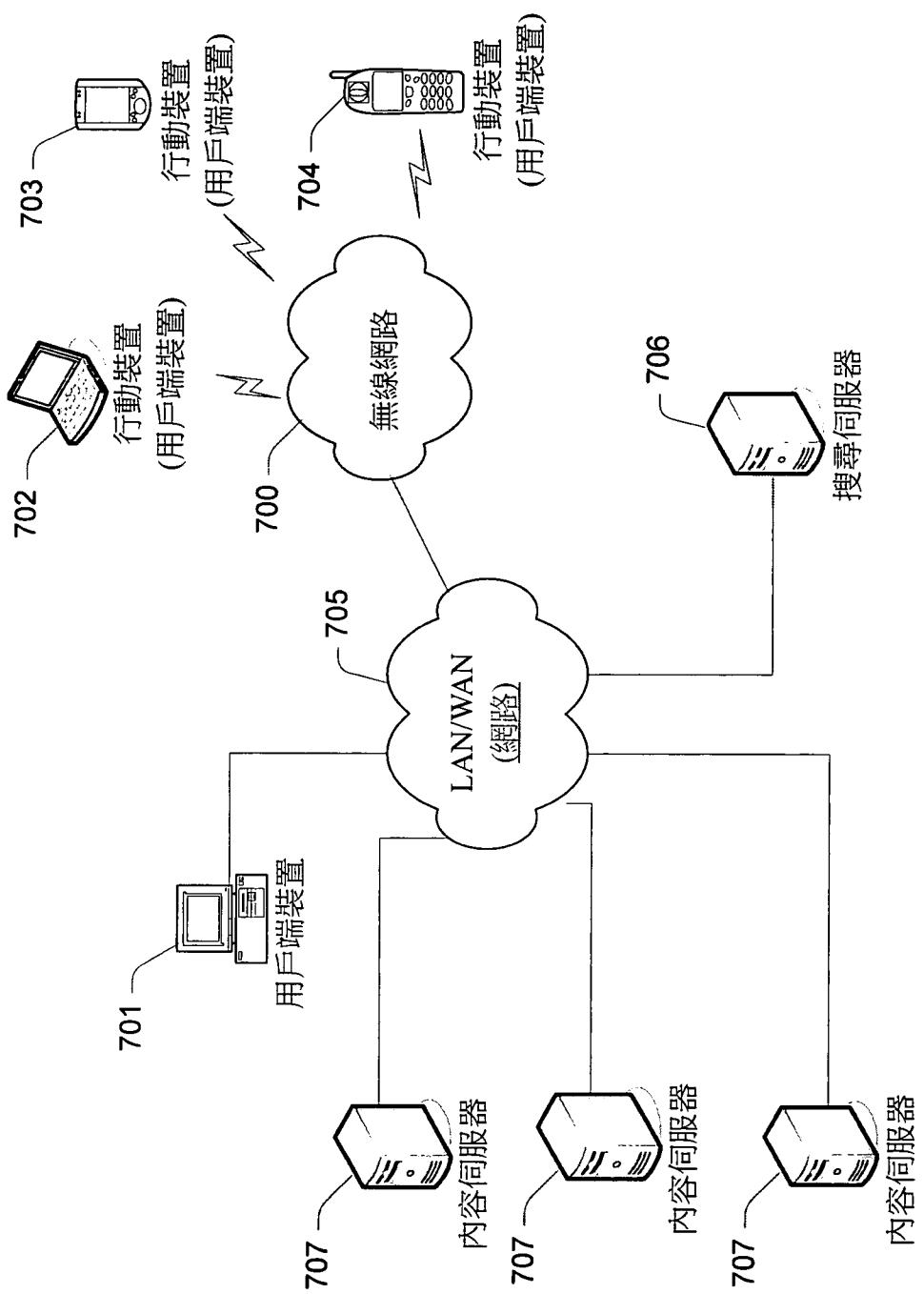
相似度權重：0.85
標記：eiger 0.85

第六圖

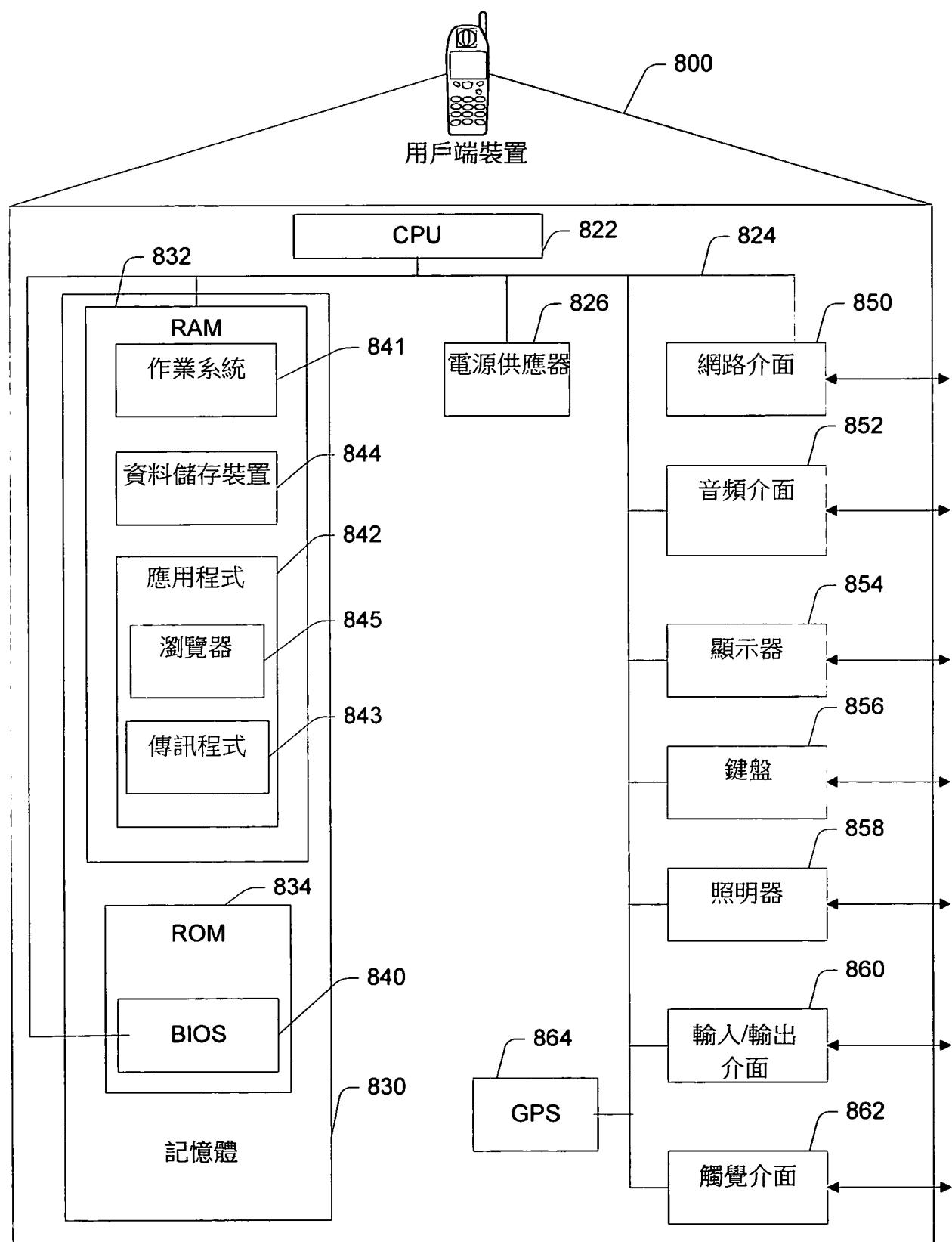
自我場景學習



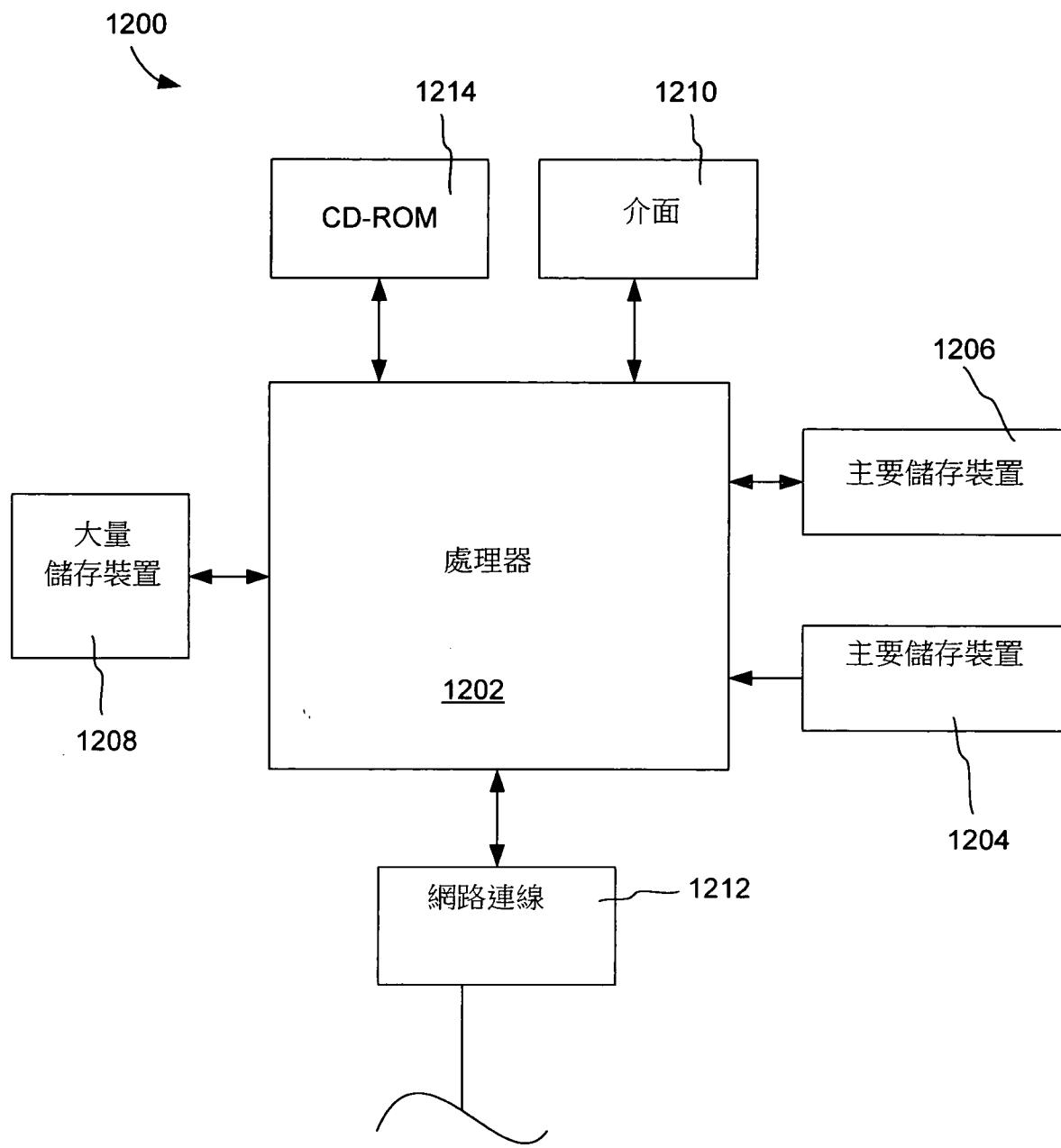
第七圖



第八圖



第九圖



第十圖