

篇名:基於相片中共同事件之推薦系統

作者:

侯嘉柔; 國立彰化師範大學資訊管理研究所; artuv43@hotmail.com.tw[通訊作者]

江憲坤; 國立彰化師範大學數位內容科技與管理研究所;
hkchiang@cc.ncue.edu.tw

顏韶威; 國立彰化師範大學數位內容科技與管理研究所; swyen@cc.ncue.edu.tw

陳鴻文⁺; ⁺銘傳大學資訊傳播工程學系; hwenzhen@mail.mcu.edu.tw

通訊作者資訊:

姓名: 侯嘉柔

單位: 國立彰化師範大學資訊管理研究所

Email: artuv43@hotmail.com.tw

職稱: 研究生

聯絡地址: 500 彰化市師大路二號 國立彰化師範大學資訊管理研究所

聯絡電話: (04) 7232105 轉 7336

致謝

本文感謝科技部大專研究計畫之經費之贊助, 計畫編號: 104-2815-C-018-022-H。

作者簡介

侯嘉柔現為國立彰化師範大學資訊管理研究所研究生。聯絡地址為 500 彰化市師大路二號國立彰化師範大學資訊管理研究所。電子郵件帳號為 artuv43@hotmail.com.tw。

基於相片中共同事件之推薦系統

A Recommendation System Based on Photographs' Common Events

摘要

聰穎型(smart phone)手機的風行已經到達人手一機的境界，其 CPU 運算速度的精進和內建多個感應器的各式功能，讓各式各樣的行動應用因應而生，其中又以照相功能最為普遍。再則，隨著攝影數位技術的進步和社群網站的普及進步，讓人們樂於用手機隨手拍照並將照片上傳到社群網站，分享給親朋好友們。然而，過量的相片卻造成搜尋和分類上的困難。然而，目前能針對大量照片中資料來找出他們的共同事件，並據此來協助使用者找到相關的照片來進行瀏覽和或進行推薦的研究並不多見。因此，本研究首先訂定相片特徵值，再運用 KNN 分群法來將具有高度照片特徵值的建立共同事件，並具此來使用者分群並，進行類似照片的推薦。

關鍵字：共同事件、合作式過濾、推薦系統

Abstract

Because of the popularity of smartphone, nowadays almost everyone owns one smartphone, with advanced CPU computing power and many built-in sensors, which makes versatile applications possible, especially the photo-taking application. Furthermore, with advances in digital photographic technology and Internet social networking websites, people are willing to take photographs of interest and sharing them families or friends in social networking websites. However, the abundant photographs make the searching or classifying them difficult. Currently, there are few researches in finding common events in photographs and use common events as basis for photography searching or recommendation. Thus, in this study, a photograph classifying and recommending framework is proposed basing on the extracting the useful photograph's information which is then used by KNN algorithm for photograph classification and recommendation.

Keywords: common event, collaborative filtering, recommendation system

壹、緒論

照相，被稱為靜態攝影(Photography)¹。早期，人們使用照相機進行靜態圖片攝影，照片以實體的形式供人們欣賞，只能透過傳閱的方式進行分享。；如今而後，數位相機的出現為人類的生活帶來轉變，照片不再需要透過實體的方式分享而是透過數位化的虛擬方式傳播分享。

以目前行動手機內建照相機影像處理的運作方式設備普及的程度而言，一般人拍攝的照片當中會包含許多資訊，包括如:照片中包含的人物、照片當下發生的事件、照片拍攝的時間、相片拍攝的地點、當中包含的景物。若能基於「人、事、時、地、物」5W 標籤²(Who, Which, When, Where, What)管理的概念(圖 1)，將每張照片註記這些重要資訊，甚至個人化資訊；如此，這些照片就有了獨特的特質，供電腦自動辨識。

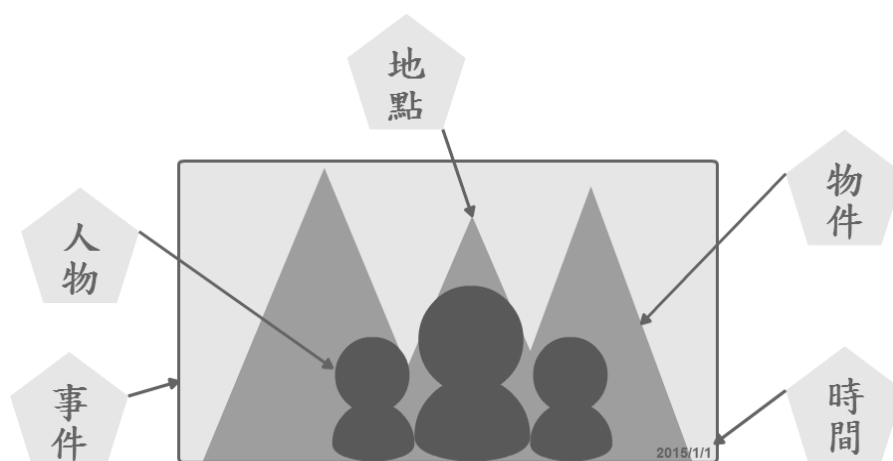


圖 1 5W 標籤示意圖

本研究希望藉由標籤管理的概念，經由系統自動化或個人手動來為每張照片標記 N 個標籤($N \geq 5$)，但此 N 個標籤代表都必須要含有「人事時地物」這五個特徵值。透過 KNN 演算法³，來對照片進行辨別和分類法則，結合共同事件⁴的定義和法則，以決定每張照片的相關性程度(高、中、低)，進而形成共同事件相簿，供系統進行推薦使用。本系統從多個面向來註記發生的事件，進而並統整出多種共同事件相簿，進而提供使用者更多樣化的推薦資訊。

藉由此研究之機制，使用者所儲存的大量個人照片，便不必擔心搜尋不易的問題。再則並且，使用者也更可透過系統統整之共同事件相簿，瞭解好友之高度相關性照片情況，提供使用者一個全新的觀點來觀看其社群網絡，從而進行更進一步之社交活動。最後，系統利用相關性與共同事件資訊，對使用者進行多樣化推薦。

貳、 文獻探討

陳品如⁵認為，隨著行動儲存容量的提升，使用者毋須經常刪除照片。然而，隨手拍攝的數位相片數量持續增加，如何進行數位內容管理便越顯重要。陳耀崧⁶指出智慧型行動裝置普及，改變生活方式，人們習慣以手機拍照其喜歡之物件，幾年下來後，就會累積大量多媒體圖文資訊；但就現有的影像檔案管理應用程式而言，並沒有辦法提供一個有效方便的管理機制。王柏翔⁵也指出現在是個資訊負載的社會，網路上的資訊漸漸大於資訊搜尋者的需求而造成使用者尋找資訊時的負擔。在搜尋方面，王柏翔⁷表示使用者在搜尋問題的時候，不清楚要輸入什麼關鍵字，因此搜尋就變得非常耗時且沒有效率。鄭志宏⁸更認為目前的查詢延伸的技術，大多是基於整個網路使用者的個人資訊得出，卻忽略了社群網路中使用者們的關係。因此，此類技術無法幫助使用者探索社群內其他使用者所擁有的潛藏資訊。因此，本研究之系統將自動分析照片資訊，基於「人、事、時、地、物」5W 標籤管理的概念，註記標籤，事先為每張照片標記 N 個標籤，做為相關性分析研究之基礎。

郭子豪⁴認為現實世界裡，一個事件的發生通常會有許多個參與者。參與者們可能都會在事件發生的過程裡用自己的角度拍攝照片，事後這些照片成為獨立的相簿，而這些相簿的集合就被稱為共同事件相簿(co-event albums)。由於共同事件相簿必為多人參與的事件，共同事件相簿同時也表徵了社交活動。因此，視覺化共同事件相簿與事件參與者之間的關係可以提供使用者，從另一個面

相來看其所屬的社群網路，並且提供使用者一個更容易更新社群活動的介面。然而由於現在人大多數皆擁有大量照片，因此如何建立共同事件相簿的分類，即是一個極具研究的議題。然而，若是分類過少則會造成單個共同事件相簿照片數量過大而難以瀏覽；反之，若是分類過多則會造成使用者相簿查找不易。而本研究利用共同事件的概念，以標籤為基礎，利用購物籃分析與 KNN 演算法做分類。因此，依照共同事件的定義和規範，可提供系統進行快速搜尋高度相關性之共同事件照片。

江宗憲⁹學者將最近鄰居之排列方法與其他有利用到最近鄰居的其他知名演算法做比較之實驗結果顯示，最近鄰居法普遍都可以有不錯的表現。而對於一些問題來說，最近鄰居法的結果也略勝於其他利用最近鄰居的知名演算法。他表示若能妥善地利用最近鄰居法，對於解決多標籤分類問題是很有幫助的。**K-nearest neighbors**¹⁰ 中文意思是「離你最接近 K 個鄰居」，意旨假設在一個以你為圓心畫圓的範圍內，離你最近的 10 個人的綜合結果可以當作你的參考結果。**KNN** 演算法的奧義其實就是「見風轉舵」，假設現在剛好在舉辦活動，以顏色作為區分人數多的一方勝利，以你為圓心畫圓的範圍內的那 10 個人穿的衣服有兩種顏色橘色和綠色，橘色 7 人、綠色 3 人，而你發現你外套裡面的衣服剛好是橘色，那麼這時你可以換穿橘色外套以加入勝利的橘色陣營，這就是 **KNN** 演算法簡單的概念。

KNN 屬於機器學習中的監督式學習¹¹(supervised learning)，機器學習就是讓機器接收外界輸入的資料以後，依照某種演算法，訓練出一種模型，這個模型是一種從資料學習出來的東西。有了這個東西，機器看到新的資料的時候，根據所訓練出來的模型，就會有某種程度的經驗和智慧去了解新的資料。非監督式的學習，是指我們給予機器簡單的學習方法，或是一個簡單的價值觀，然後把資料輸入機器，讓機器自行判斷正確答案。**KNN** 一般用來做資料的分類，如果已經有一群分好類別的資料，後來加進去的資料就可以透過 **KNN** 的方式指定新增資料的分類。

K 表示一個常數，簡單而言，**KNN**³ 就是離新增資料最近的 K 個點判斷哪個類別的點最多就把新增的資料也當成那個類別。舉例上述例子而言，就是 10NN。以圖 2 為例，k=10 所以周遭有 7 橘 3 綠，那麼中間的灰色點就會成為橘色。藉由 **KNN** 演算法搭配購物籃分析，照片資料便可以藉由簡單的方式做出極大的效益，以達本研究的目的找出高度關聯性的照片。

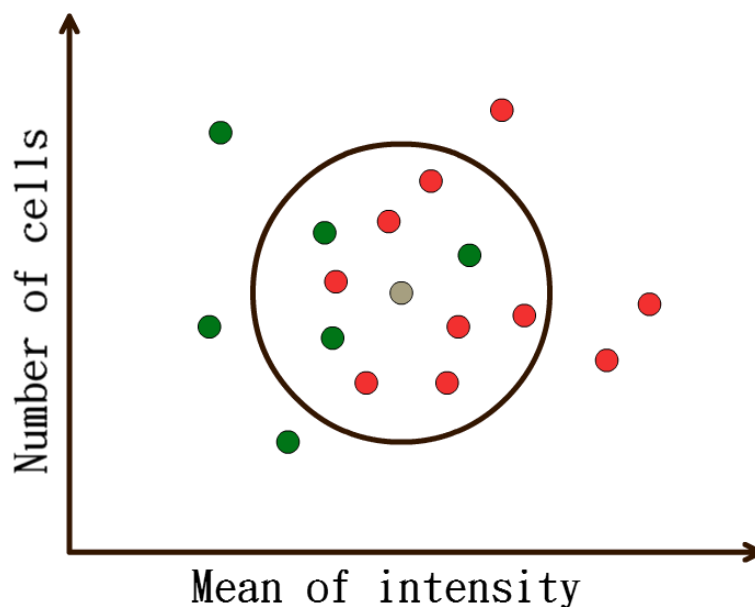


圖 2 K-nearest neighbors algorithm 圖解

Mayer-Schonberger, V. & Cukier, K.¹² 表示分析大量資料時，找出其中的相互關係，再根據這些相互關係做出預測，正是大數據的核心所在。了解資料的相互關係可能不會精確地告訴我們事情發生的原因，也能提供終極答案，但卻可以提供我們各種新的見解和效益，幫助我們更貼近現實。

參、 相片共同事件之推薦

一、 相片共同事件資料之建立

本研究的相片取自現今熱門的名人相片社群網站：Instagram，利用其照片原有的資訊進行研究分析。目前，本研究利用二十位使用者，每人二十張相片進行分析。每張照片所包含的資訊如表 1 和圖 3 所示：

表 1 本研究中，相片資訊一覽表。

ID	File Name	Image Description	Artist	Date	XP Keywords	XP Comment
aayan1120	aayan1120_927418_615635498569111_2071796117_n.jpg	I ever had 好吃的流淚啊！Omg	炎亞綸；aayan1120	2015:01:26	#regent;#the;#most;#delicious;#burger	
appleweiting	appleweiting_11899702_1632050030345590_671567757_n.jpg	好日子老闆生日請大家吃蛋糕☺喔耶～老闆生日快樂!!🍰	黃喬歆；appleweiting	2015:08:31	#好日子早午餐;#happybirthday;#lucky	好日子早午餐
byebyechuchu	byebyechuchu_11849183_1113288142033876_341385510_n.jpg	好像比較帥	掰掰啾啾；byebyechuchu	2015:09:20	# shopping	Tokyo Sky Tree Tower (東京スカイツリー)

說明：四百張相片中，以其中三張為例展示

其中 ID 代表帳號、Image Description 為本張發佈照片的標題、Artist 為本使用者名稱和帳號、Date 為相片發布日期、XP Keywords 為 Instagram 之標記、XP Comment 為地點及其他資訊。

Serial Number	ID	Number	File Name	Image Description	Artist	Date	XP Keywords	XP Comment
1	aayan1120	1	aayan1120_927418_615635498569111_2071796117_n.jpg	I ever had 好吃的流淚啊！Omg	炎亞綸aayan1120	2015:01:26	#regent;#the;#most;#delicious;#burger	
2	aayan1120	2	aayan1120_1596678_357210364487418_1783172416_n.jpg	I'm on Facebook mentions now	炎亞綸aayan1120	2015:01:28		
3	aayan1120	3	aayan1120_10724251_981074445242898_605799844_n.jpg	Hey my heart for you	炎亞綸aayan1120	2014:11:03	#heart	food
4	aayan1120	4	aayan1120_10735474_1375885132702565_922100907_n.jpg	Happy #Halloween!! 不可愛 可愛? 炎亞綸aayan1120	炎亞綸aayan1120	2014:11:02	#Happy#Halloween	
5	aayan1120	5	aayan1120_10809531_1579042275659331_1613883154_n.jpg	謝謝貴客的生日驚喜！有火球 #賀! 炎亞綸aayan1120	炎亞綸aayan1120	2014:11:30	#thankyou#賀!#Pot	
6	aayan1120	6	aayan1120_10882005_571531356315358_1997903585_n.jpg	簡單的幾個字就能給一個人一整天 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:03	#fewwords#simple#to#win#to#person#to#make#one#th	
7	aayan1120	7	aayan1120_10899150_913640098675564_1266497250_n.jpg	和好友一起吃飯甜點 很奢侈的享受 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:25	#having#dinner#with#some#friends#feels#great	
8	aayan1120	8	aayan1120_10919468_1529697717299450_957738394_n.jpg	冬天吃冰～#what#is#me? 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:10	#what#is#me?#winter#want#is#me?	
9	aayan1120	9	aayan1120_10932414_357467651127524_71156928_n.jpg	謝謝 很溫暖！雖然讓我改名字🙄 炎亞綸aayan1120	炎亞綸aayan1120	2015:02:02	#thx#have#a#wonderful#day#hope#you#enjoy#your#di	
10	aayan1120	10	aayan1120_10948704_357425047793743_330905380_n.jpg	富士山終於露臉了 所以我不要臉 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:27	#Fuji#is#back#days#富士山	
11	aayan1120	11	aayan1120_10950554_894358497263096_1336707609_n.jpg	#Nikesickstyle 再辛苦都值得的結 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:26	#Nikesickstyle#days#me# crazy	
12	aayan1120	12	aayan1120_10958144_1447169618906610_1342841295_n.jpg	人生就是不斷的抉擇，然而困難 炎亞綸aayan1120	炎亞綸aayan1120	2015:01:20	#Life is#full of#choices but it's not what choices you've made	
13	aayan1120	13	aayan1120_11015547_721202514664822_378194757_n.jpg	一個人的早餐很會吃～#after #gym 炎亞綸aayan1120	炎亞綸aayan1120	2015:03:02	#LastSize#breakfast#one#after#gym	
14	aayan1120	14	aayan1120_11184622_1421899624786603_1448899754_n.jpg	甲奔！#eat 炎亞綸aayan1120	炎亞綸aayan1120	2015:04:27	#eat	
15	aayan1120	15	aayan1120_11203439_419156138254164_1628354441_n.jpg	謝謝可口可樂 炎亞綸aayan1120	炎亞綸aayan1120	2015:05:04	#CocaCola#thank#you#it's#amazing!#可口可樂#很感謝	
16	aayan1120	16	aayan1120_11265820_442383115935259_210864490_n.jpg	夜深了但心還是很澎湃～ 炎亞綸aayan1120	炎亞綸aayan1120	2015:05:04	#夜深#晚安#Goodnight#enjoymusicandlife	
17	aayan1120	17	aayan1120_11272941_470699306438940_1968333179_n.jpg	想念 今天去找了老闆談 炎亞綸aayan1120	炎亞綸aayan1120	2015:06:01	#大南路#芝山岩牛肉麵#小時候的味道#想念#miss#the	
18	aayan1120	18	aayan1120_11373820_16110819232488235_614462537_n.jpg	韓國的早晨好舒服 炎亞綸aayan1120	炎亞綸aayan1120	2015:09:14	#Goodmorning#Korea#sunny#day	
19	aayan1120	19	aayan1120_11410630_111373105867172_1548406315_n.jpg	想知道我在幹嘛快來#松山文創4號 炎亞綸aayan1120	炎亞綸aayan1120	2015:06:22	#松山文創4號倉庫	
20	aayan1120	20	aayan1120_12144163_858449264270683_1049192868_n.jpg	目前為止我最愛的合照 炎亞綸aayan1120	炎亞綸aayan1120	2015:07:27	#目前為止我最愛的#合照	
21	aohueifu	1	aohueifu_10175163_1704817303084245_1400438147_n.jpg	Tiff happy birthday 郭雪芙aohueifu	郭雪芙aohueifu	2015:11:02	#tiff#aymenlo	
22	aohueifu	2	aohueifu_10665453_1129408870419244_1644867563_n.jpg	L 郭雪芙aohueifu	郭雪芙aohueifu	2015:07:20		
23	aohueifu	3	aohueifu_10724972_40899359269781_16612686396_n.jpg	最近拍照眼睛都有點酸🙄 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:27		
24	aohueifu	4	aohueifu_10848192_316402998557741_1712968396_n.jpg	宸希睡眠 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:13	#睡覺#sleep	
25	aohueifu	5	aohueifu_10881981_1394470880850458_135836787_n.jpg	大家辛苦了～ 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:13		
26	aohueifu	6	aohueifu_10882038_1399395383688262_174897677_n.jpg	☺☺☺ 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:05		
27	aohueifu	7	aohueifu_10899359_658409357602390_168520563_n.jpg	吃吃吃～吃到出神 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:13	#吃#eat	
28	aohueifu	8	aohueifu_10899484_736083319803474_2000714292_n.jpg	睡醒啦～ 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:13	#醒#sleep	
29	aohueifu	9	aohueifu_10914389_770931139658348_1010478328_n.jpg	脫窗加大小銀的後果～喜歡嗎?;-) 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:27		
30	aohueifu	10	aohueifu_10932089_1590456801183726_1491317519_n.jpg	晚安 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:27	#睡覺#sleep	
31	aohueifu	11	aohueifu_10932120_597290017081711_950549010_n.jpg	一種秒睡的感覺 郭雪芙aohueifu	郭雪芙aohueifu	2015:01:20	#睡覺#sleep	
32	aohueifu	12	aohueifu_10958685_690360897734879_1506474841_n.jpg	哦～明天見!!! (偷圖了☺) 郭雪芙aohueifu	郭雪芙aohueifu	2015:02:16		
33	aohueifu	13	aohueifu_10994992_859806254058200_792633625_n.jpg	最近在當神隱少女!!! 郭雪芙aohueifu	郭雪芙aohueifu	2015:04:30		
34	aohueifu	14	aohueifu_11022889_914634248586904_1140492841_n.jpg	♬♬♬好吃～海鮮大餐泰選吃 郭雪芙aohueifu	郭雪芙aohueifu	2015:03:02	#吃#eat	
35	aohueifu	15	aohueifu_11055406_876878462353426_448542244_n.jpg	在ㄟ什麼呢? 郭雪芙aohueifu	郭雪芙aohueifu	2015:04:06		
36	aohueifu	16	aohueifu_11056022_1538297211124425_156811496_n.jpg	你好皮 郭雪芙aohueifu	郭雪芙aohueifu	2015:03:15		

圖 3 四百份照片詳細資料部分內容畫面截取圖。

二、 資料分群

在相片資訊的擷取及匯整方面，本研究利用 ExifTool¹³ 將每張照片背後詳細資訊匯出 txt 檔案(圖 4)後，再利用 c++程式進行特定部分資訊(表 1)擷取與統整成二十份 csv 檔案(圖 5)，以利後續分析使用。


名稱	類型	大小
 aayan1120_927418_615635498569111_2071796117_n	文字文件	4 KB
 aayan1120_1596678_357210364487418_1783172416_n	文字文件	5 KB
 aayan1120_10724251_981074445242898_605799844_n	文字文件	4 KB
 aayan1120_10735474_1375885132702565_922100907_n	文字文件	4 KB
 aayan1120_10809531_1579042275659331_1613883154_n	文字文件	4 KB
 aayan1120_10882005_571531356315358_1997903585_n	文字文件	4 KB
 aayan1120_10899150_913649098675564_1296497250_n	文字文件	4 KB
 aayan1120_10919468_1529697717299450_957738394_n	文字文件	4 KB
 aayan1120_10932414_357467651127524_71156928_n	文字文件	5 KB
 aayan1120_10948704_357425047793743_330905380_n	文字文件	4 KB
 aayan1120_10950554_894358497263096_1336707609_n	文字文件	4 KB
 aayan1120_10958144_1447169618906610_1342841295_n	文字文件	6 KB
 aayan1120_11015547_721202514664827_378194757_n	文字文件	4 KB
 aayan1120_11184622_1421899824786603_1448899754_n	文字文件	4 KB
 aayan1120_11203439_419156138254164_1628354441_n	文字文件	4 KB
 aayan1120_11265820_442383115935259_210864490_n	文字文件	5 KB
 aayan1120_11272941_470699306438940_1968333179_n	文字文件	4 KB
 aayan1120_11373820_1611081932488235_614462537_n	文字文件	5 KB
 aayan1120_11410650_111373105867172_1548496315_n	文字文件	4 KB
 aayan1120_12144163_858449264270683_1049192868_n	文字文件	5 KB

圖 4 ExifTool 匯出之一位使用者照片詳細資訊 txt 檔案。





















名稱	類型	大小
 aayan1120_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 aohsuehfu_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 appleweiting_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 bingbing_fan_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 blackielovelife_exif	Microsoft Excel 逗點分隔值檔案	5 KB
 breeze_wayne_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 butterfly092288_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 byebyechuchu_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 chenbolin_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 cherngyang_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 hsiang_5208_exif	Microsoft Excel 逗點分隔值檔案	6 KB
 janetagram_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 kikihsieh_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 queenshoplife_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 sandra0315306_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 sqwhat_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 t2025242322_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 waawei_exif	Microsoft Excel 逗點分隔值檔案	3 KB
 wuchengan_exif	Microsoft Excel 逗點分隔值檔案	4 KB
 yuyanpeng_exif	Microsoft Excel 逗點分隔值檔案	3 KB

圖 5 二十位使用者照片詳細資訊 csv 檔案。

KNN classification (KNN 分類法)

本研究根據所拍攝的照片特徵值，運用 KNN 分群法將相片分類。當系統形成共同事件相片群後，當有新相片被增加時，系統將可為此照片分入關聯性最近的一群，而相關的推薦訊息便可提供給此相片的擁有者。圖 6 是本研究根據 Weka 資料挖掘程式所要求的 arff 格式，建立出的照片資料。圖 7 是 Weka 範例程式，用 knn 演算法來對第 1 筆資料和第 2 筆資料進行分類的程式，圖 8 是程式執行結果。

```

1 @relation aayan1120-instagram-photos
2
3 @attribute PersonsNumber numeric
4 @attribute Object numeric
5 @attribute Time numeric
6 @attribute Mood numeric
7 @attribute type numeric
8
9 @data
10
11 0,1,1,1,0
12 1,6,0,1,1
13 0,1,3,1,0
14 0,1,2,1,0
15 0,3,3,1,2
16 0,1,0,1,0
17 0,1,2,1,0
18 1,1,1,1,0
19 0,1,0,1,0
20 1,4,0,1,3
21 0,5,3,1,4
22 1,6,2,4,5
23 0,1,0,1,0
24 0,1,3,1,0
25 0,1,2,1,0
26 1,6,4,1,5
27 2,0,3,1,6
28 0,1,0,1,0
29 1,3,2,1,2
30 2,6,3,1,6

```

圖 6 使用者 ID:aayan1120 二十張相片屬性資料示意圖

```

8 import weka.core.Instances;
9
10 public class KNN1 {
11     public static BufferedReader readDataFile(String filename) {
12         BufferedReader inputReader = null;
13
14         try {
15             inputReader = new BufferedReader(new FileReader(filename));
16         } catch (FileNotFoundException ex) {
17             System.err.println("File not found: " + filename);
18         }
19
20         return inputReader;
21     }
22
23     public static void main(String[] args) throws Exception {
24         BufferedReader datafile = readDataFile("aayan1120.arff");
25
26         Instances data = new Instances(datafile);
27         data.setClassIndex(data.numAttributes() - 1);
28
29         //do not use first and second
30         Instance first = data.instance(0);
31         Instance second = data.instance(1);
32         data.delete(0);
33         data.delete(1);
34
35         Classifier ibk = new IBk(7);
36         ibk.buildClassifier(data);
37
38         double class1 = ibk.classifyInstance(first);
39         double class2 = ibk.classifyInstance(second);
40
41         System.out.println("first: " + class1 + "\nsecond: " + class2);
42     }
43 }

```

圖 7 KNN 分群法程式碼

本系統運用 Weka¹⁴ Library 來執行 KNN clustering

```

C:\Windows\system32\cmd.exe
Microsoft Windows [版本 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Users\JiaRou>d:

D:\>cd weka_dm

D:\weka_dm>d:\myjava.bat

D:\weka_dm>set path=d:\jdk18\bin;C:\ProgramData\Oracle\Java\javapath;C:\Windows\
system32;C:\Windows;C:\Windows\System32\Wbem;C:\Windows\System32\WindowsPowerShe
ll\w1.0\;;C:\Program Files (x86)\Microsoft VS Code\bin;.

D:\weka_dm>set classpath=d:\jdk18\lib;D:\progra~1\Weka-3-6\weka.jar;.

D:\weka_dm>javac KNN1.java

D:\weka_dm>java KNN1
first: 0.0
second: 1.0

D:\weka_dm>

```

圖 8 KNN 分群程式執行後結果

三、 基於相片中共同事件之推薦系統

本研究將根據系統離型法，設計和實作一個能夠根據者用者照片集之特徵值、相關性、共同事件，來彙整、分類和建立每位使用者之「相片推薦集」，以提供相片共同事件之推薦系統。

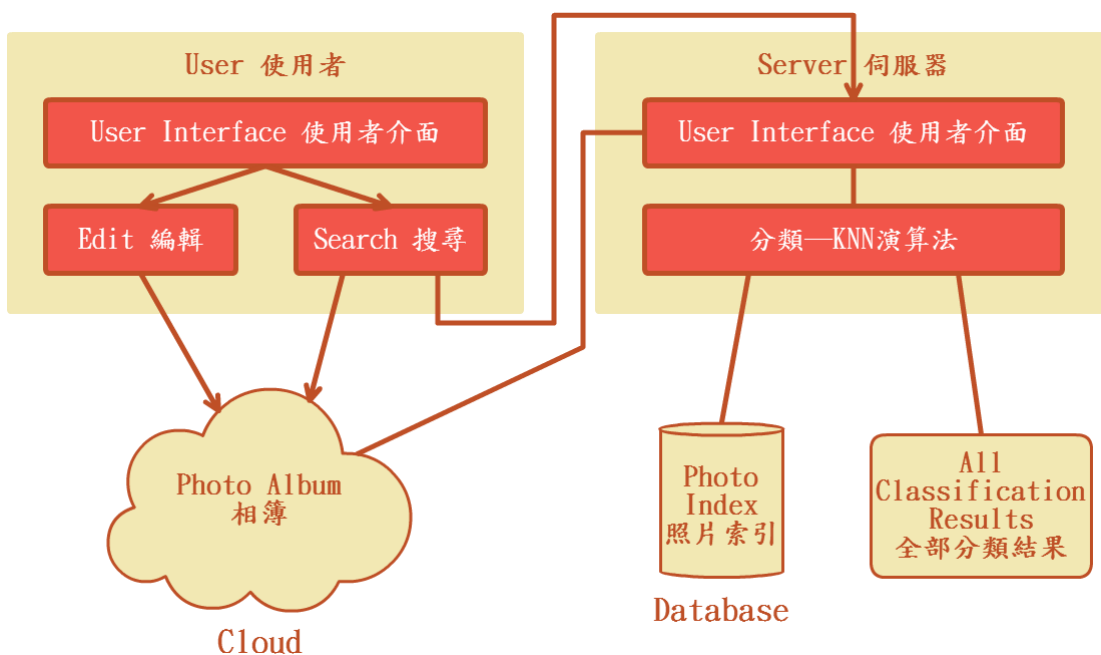


圖 9 系統架構圖

肆、 討論與結論

現代人擁有聰穎型手機後，不論到任何地方做任何事情，用餐、旅遊、工作、玩遊戲等，皆容易拍攝大量相片。而大量相片可能被分享於各個不同的社群網站、上傳備份於雲端或是儲存於個人手機或其他設備中。然而，相片背後的大量資訊卻多沉寂無用。大數據在今天已是廣為人知的大潮流，席捲而來。在不久的將來，當數據化更為普遍，而每個人的數據資料量也會快速累積。因此，以照片而言，如何快速找到在社群網絡中，跟自己有興趣的照片或網友，是一個具有實務性的議題。

本研究提出一個基於具有共同事件照片的推薦系統的基本架構，並基於「人、事、時、地、物」5W 標籤管理的概念，將每張照片以自動化或半自動化的方式製訂特徵值屬性，然後再運用 KNN 演算法將訓練樣本(training set)建立起分類模型。如此，應用本系統，即可對未來的某張照片進行分類。目前本系統已經提出基本概念和架構，並正在進行訓練樣本和測試樣本的檢測，以便能夠找出真正有用的照片特徵值，建立具共同事件照片的分類模型。期待本研究未來實做出的分類和推薦系統，可以讓使用者快速且準確的找到跟她的相片相關性高的相片或網友。

參考文獻

1. 吳鋼(2006)。攝影史話，中國攝影出版社。
2. David Olson& Yong Shi.(2007). Introduction to business data mining. The McGraw-Hill Companies, Inc.
3. 白 昌 永 (2014). [Machine Learning] KNN 分 類 演 算 法 ， 取 自 <http://enginebai.logdown.com/posts/241676/knn>
4. 郭子豪(2011)。基於社群網路之共同事件相簿呈現，碩士論文。台北市：台灣大學。
5. 陳品如(2013)。數位相片註記結合行事曆資訊之研究，碩士論文。桃園市：國防大學。
6. 陳耀崧(2014)。用於社群用戶有效率蒐集、標註、分享和管理照片的 Web 應用程式開發平台，碩士論文。南投縣：國立暨南國際大學。
7. 王柏翔(2012)。虛擬社群之搜尋經驗設計，碩士論文。新竹市：國立清華大學。
8. 鄭志宏(2012)。基於社群網路分析中字詞共現關係的用戶化查詢優化研究，碩士論文。桃園市：國立中央大學。
9. 江宗憲(2011)。基於最近鄰居之排列方法於多標籤分類問題，碩士論文。台北市：台灣大學。
10. Misgod (2014). 輕 鬆 聊 之 KNN 演 算 法 ， 取 自 <http://35around.blogspot.tw/2014/07/knn.html>

11. Thursday (2007). MMDays - KNN 演算法，取自
<http://mmdays.com/2007/05/16/knn/>
12. Mayer-Schonberger, V. & Cukier, K. (2014). Big data: A revolution that will transform how we live,work, and think, Master NO.506。
13. Kathryn Gronsbell (2013). ExifTool Introduction. From
<http://u88.n24.queensu.ca/~bogdan/>
14. Jason Brownlee (2014). How to Run Your First Classifier in Weka. From
<http://machinelearningmastery.com/how-to-run-your-first-classifier-in-weka/>