

# 以影像辨識實現智慧型導覽系統

朱益進\* 黃惠藩

亞洲大學資訊工程學系（台中市霧峰區柳豐路500號）

\*Jashshadow@yahoo.com.tw

## 摘要

為了使目前的導覽系統能夠更進一步的優化，本研究結合影像識別技術以加強導覽系統對於定位目標的精確度，並以擴增實境的方式呈現導覽資訊。本系統首先建立影像資料庫，對目標場景物件取得不同角度和距離之影像，將其分類並且以SURF方法擷取影像中具特色的特徵點，將特徵點資訊存入資料庫中。使用者所輸入的影像與資料庫的SURF特徵點進行特徵比對，以此推斷影像為何種場景並且同時得到新場景與資料庫影像的尺度、旋轉變化參數，以此參數為基準放入導覽資訊進入新場景影像中，以此提升導覽系統在定位目標的精確度和達到類似擴增實境的效果。本研究中以亞洲大學為例，對五棟主要建築進行測試，準確度達到96%。

**關鍵詞：**特徵點偵測、影像識別、SURF、擴增實境。

## 1. 簡介

近年來行動裝置的普遍與GPS導航系統的快速發展，導覽系統幾乎成了出外旅遊的必備工具（張樹安，2010），同時擴增實境已經在許多領域如建築維修數位化、防災教育、休閒娛樂等應用都展現了驚人的實用性（王國雄，2010）。但目前市面上的導覽系統即使能提供許多資訊，但卻無法很精確的進行更詳細的引導。由於大部分系統只能依靠GPS來確定位置並且提供資訊，但僅僅是使用者的位置是不足的。通常使用者所期望得知的目標並非自身所在的位置，而是由此位置所能觀察到的目標，但是只以GPS定位的導覽系統只能由使用者的位置提供資訊，而不是使用者目標的位置為基準(王湧天等，2006)。由此可能產生某些情況，例如使用者感興趣的並非物件的整體而是某物件的某一部分，如圖1所示，以亞洲大學為例，使用者感興趣的位置可能在於建築右上方的校長室，而非整棟圖書館大樓。使用者站在相同位置所能觀察的範圍非常的廣，如圖1所示，圖1a與圖1b為同一地點不同角度之影像，由圖可知就算是同一地點由使用者觀測的角度不同場景物件也會有很大的變化，如此單純的依靠GPS定位是不足的。



圖 1 情景示意圖

影像辨識在電腦視覺的領域中一直是最具挑戰性的問題之一，物件的尺度、旋轉使描述更加困難。因此影像辨識的一個關鍵問題在於設法取得辨別力高並且獨特的特徵描述，和能因應影像的形變和光線條件變化。早期在描述影像時主要是利用整體影像的特徵，如色彩和紋理 (Swain and Ballard, 1991; Tuceryan and Jain, 1998)。這些方法能由影像直接取出的特徵並且描述整體的影像內容。此種方法雖然在計算的效率上相當優秀，但是整體影像的特徵對於背景變化、遮蔽、視角與光照改變的影響非常敏感。因此，局部的描述影像特徵在近年來越來越受到重視。不同於直接描述整體影像，局部區域描述法的概念則是從視角與尺度不變的局部區域中掘取局部影像特徵。和整體影像的特徵描述相比，局部特徵描述在應對光照改變、影像變化與遮蔽時都能保有較高的辨識度，並且在許多電腦視覺的領域內被應用且有良好的效果，例如影像檢索 (Mikolajczyk and Schmid, 2001) 機器人定位 (Lowe and Little, 2002) 等。許多文獻提出了不同的局部區域描述法，其中最著名的局部區域描述法為SIFT (Lowe, 1999)，被證實為目前最強大且實用的演算法，能夠將影像資料轉換成具有尺度不變性特性的特徵點座標，對旋轉、尺度變化、光度變化具備較高的不變性，以此可降低場景內可能的遮蔽、移動等雜訊。而由此延伸了許多改進的方法如：PCA-SIFT (Ke and Sukthankar, 2004)、GLOH (Mikolajczyk and Schmid, 2005)、HOG (Dalal and Triggs, 2005)、SURF (Herbert Bay, et al., 2008)。其中SURF在保有SIFT在特徵描述上的不變性的基礎上並且將運算速度大為縮減。

因此，本研究以影像識別結合擴增實境的技術以取得使用者所感興趣的導覽資訊，以亞洲大學為例，期望能達到使用者透過行動裝置鏡頭所看到的影像都能迅速的回饋資訊。由於SURF在特徵點描述上具有高度的不變性，本實驗以SURF進行場景物件比對，並且以擴增實境的方式將導覽資訊呈現，使得系統更具直覺性的操作。

## 2. 系統架構

系統架構主要分兩部份，建立影像資料庫和場景物件比對。影像資料庫主要用於儲存場景物件資料以用來和輸入的影像進行比對，因此影像資料庫內只需儲存影像特徵點、物件參數（尺度、角度、中心點等等）、導覽資訊，進行特徵比對時只需要提取資料庫內以儲存的特徵即可進行比對，而不需要再次進行特徵擷取，同時資料庫內的特徵點只保存主體物件的特徵點。場景物件比對時則由資料庫讀取資料和

輸入的影像進行特徵比對，流程如圖2。以下針對系統流程做進一步的詳細說明。

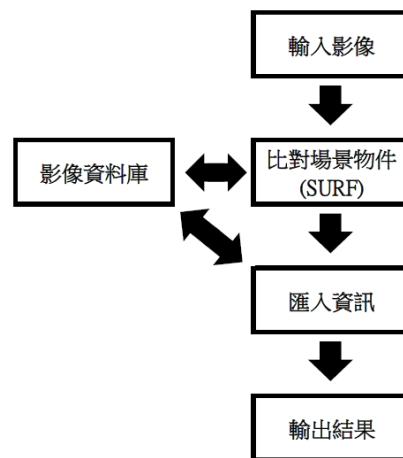


圖 2 場景物件比對流程圖

## 2.1 建立影像資料庫

以亞洲大學為例，首先依照距離和角度等等的不同對單一場景物件進行多次拍攝，並且依照拍攝時的條件進行分類。依據本實驗的結果，拍攝時依每15公尺或15°為條件型一次拍攝，其中由於日夜光線變化差異太大，再次追加中午、夜晚兩項條件。以近距離正面拍攝場景物件75%完整度為基準場景物件，如圖3a，向後推算50公尺為半徑內（距離每15公尺一張、角度每15°一張）蒐集場景物件的影像資訊，如圖3b~d。由實驗所知在15公尺或15°的距離之內對於特徵點的比對保有相當的準確度，依此建立樹狀分類器，以亞洲大學為例，依照日夜、角度、距離建立四層樹狀類。將選定的影像去除背景，只留下物件主體。取得物件主體特徵後，根據反饋的特徵資訊，計算中心位置、角度、距離等參數和特徵點一併存入影像資料庫內。總影像拍攝張數為970張，分別為5棟主體建築，每棟建築包含各種條件共192張。



(a)基準影像



(b) 15公尺15°



(c) 30公尺30°



(d) 15公尺75°

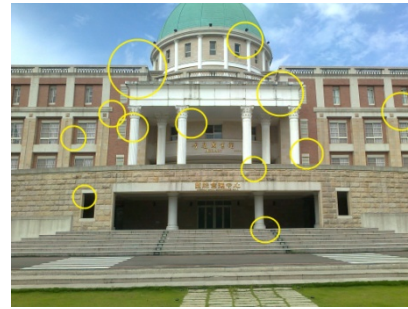
圖 3 資料庫建立示意圖

## 2.2 場景物件特徵比對

本研究採用SURF法對場景物件與影像資料庫取進行特徵點擷取和描述。SURF能夠將影像資料轉換成具有尺度不變性特性的特徵點座標，如圖4，對旋轉、尺度變化、光度變化具備較高的不變性，可降低場景內可能的遮蔽、移動等雜訊。其中SURF可分為三部分：SURF特徵點擷取、SURF特徵點描述、特徵點匹配。



(g) 原始影像



(h) 特徵點顯示

圖 4 影像特徵示意圖

### 2.2.1 SURF特徵點擷取

尋找並取得SURF特徵點是SURF的第一步，目的為取得主要並且明顯的特徵以進行後續的配對。使用Hessian矩陣進行影像中極值偵測，並且以箱式濾波（圖5）近似代替二階高斯濾波（圖6）。

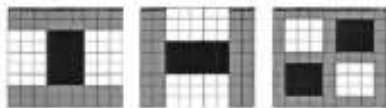


圖 5 箱式濾波

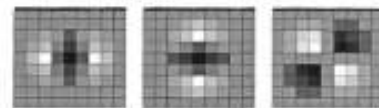


圖 6 高斯濾波

同時以積分影像（如圖 7）來加速捲積以提升運算速度，其積分影像矩形任意區域和為：

$$\Sigma = I_2(x_1 - 1, y_1 - 1) + I_2(x_2, y_2) - I_2(x_1 - 1, y_2 - 1) - I_2(x_2, y_1 - 1) \quad (1)$$

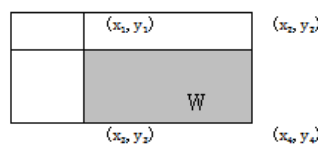


圖 7 積分影像示意圖

積分影像計算捲積的計算量不會因為W區域大小改變而有不同，計算量縮小後相對的計算速度大幅提升。以不同尺度的箱式濾波對固定的積分影像製作影像空間，在不同尺度的影像空間中用Hessian矩陣求極值，然後在3x3x3的立體鄰近區域內，將本身以及前後兩尺度周圍26個鄰近域的極大極小值都設為特徵點。

### 2.2.2 SURF特徵點描述

以特徵點為中心，半徑為6s（s為特徵點所在的尺度值）的點在x、y方向的Harr小波（波長為4s），並且以特徵點為中心由高斯函數進行加權，將360°分為6等份每60°的範圍內的值相加以形成新的向量，並且以最長的向量為主要向量。於主要向量上，以特徵點為中心建立邊長為20s的方形矩陣，將其分

為4×4的小矩陣。對每一小矩陣內分別計算（單一計算範圍為5×5）相對於主向量的Harr小波並以高斯函數加權。將每個小區域內的值相加可得到一個四維向量：

$$V_{sub-region} = [\Sigma d_x, \Sigma d_y, \Sigma |d_x|, \Sigma |d_y|] \quad (2)$$

因此，每一個特徵點會擁有4×（4×4）=16維的描述向量。將向量正規畫以消除光照影響。

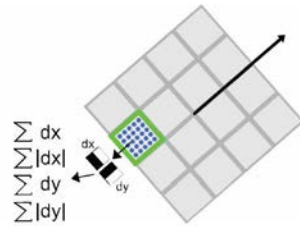


圖 8 SURF描述示意圖

### 2.2.3 特徵點匹配

基於運算速度和效能的考量，採用各歐氏離函數作為相似性度量來計算特徵向量的距離，從而判斷兩個不同的SURF特徵點是否能以互相配對。首先利用KD Tree，找到與待配對點距離最小和次最小的點，然後根據最小距離與次最小距離的比率來確定待配對點與距離最小的點是否為正確的配對點。由於場景中出現特徵相同但是位置不同的特徵點的機率極高，這使得在特徵比對時容易出現錯誤，並非每一個匹配成功的特徵點都是正確的。因此，由Hough法（Richard and Peter,2008）進行對特徵點影像投影運算，可有效降低位置不同的錯誤。

### 2.3 匯入導覽資訊

計算資料庫內影像的物件參數（中心點、角度、大小）和輸入影像參數的差。以此調整附加的影像資訊，將其縮放、旋轉並調整座標，最終和輸入的影像進行合成。

### 2.4 輸出結果

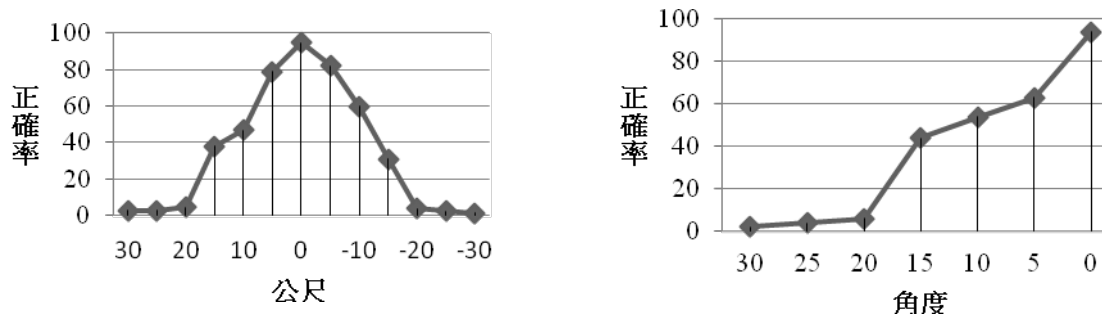
根據計算資料庫內影像的物件參數調整附加導覽資訊所顯示的位置，導覽資訊內包含文字影像、相關連結等等，以圖9a、圖9b為例，包含圖書館位置、館藏查詢、續借申請等資訊，以擴增實境的方式呈現，使用者點選後可直接連結各種服務。請參見實驗結果。

## 3. 實驗結果

科技技術的快速發展，許多數位元影像處理工具不斷的發展，其中OpenCV（Open Source Computer Vision Library）是一種跨平臺的影像演算法的函式庫（Gary and Adrian,2008），由Intel公司所開發，可用於製作影像處理、電腦視覺、圖形識別等等相關的C語言程式設計，並且以授權在商業和研究領域上免費使用，使得OpenCV普遍的受到歡迎。因此，本研究採用OpenCV作為開發環境。本實驗尚未將系統架構於行動裝置，為了驗證實驗所使用的環境為個人電腦Aspire 5750G作業系統為Windows 7以方便及時調整實驗流程。

由於SURF對旋轉、尺度變化、光度變化具備較高的不變性，本研究使用SURF進行特徵擷取和比對。其中SURF特徵點擷取由於影像資料庫內已具備場景物件的特徵點，使得系統運算時間縮短了70%。而進行特徵比對時，由Hough法進行對特徵點影像投影運算，將準確度提升23%。

圖9顯示所有影像SURF特徵點比對平均正確率（比對正確之特徵點數除以總物件特徵點數）與距離變化與角度變化的關係。由圖9a所知，在與資料庫內的場景物件相距15公尺以上比對特徵點的正確率值大幅下滑，而15公尺內則還保有30%以上的正確率，而特徵點正確率高於30%以上則依然具有高辨識準確度。對角度而言（圖9b）在15°的角度偏差內可達到30%以上的特徵點正確率。



(a) 距離

(b) 角度

圖 9 SURF 特徵點比對正確百分比

由此設計樹狀分類器，每15公尺和15° 設定一筆資料並且與輸入的場景影像比對，由此將每個場景物件的拍攝張數訂為192張。對於亞洲大學的主要建築（圖書館、資訊大樓、健康大樓、管理大樓、體育館），準確度達到96%，系統運算時間為4秒。由圖10a和圖10b可看出對於系統對於拍攝角度的不同達到了良好的處理結果，圖10b中的位置標示隨著影像的不同而改變，同時美術中心也應為拍攝角度的不同而沒有顯示在輸出結果上，圖10c則識別了建築物正反面的不同並加入了其他導覽資訊。其中導覽資訊為黃色底色的資訊代表可以直接連結各種服務，以圖10a~c為例，可直接點選以進入亞洲大學圖書館的館藏系統內，或是進行續借圖書的申請和線上預約會議室的使用。圖10d和圖10e則可以看出在不同的距離系統依然能準確的識別場景物件的位置並且以擴增實境的方式給予相關導覽資訊。



(a) 圖書館\_正面



(b) 圖書館\_正面\_右側



(c) 圖書館\_背面



(d) 資訊大樓\_正面\_近



(e) 資訊大樓\_正面\_遠



(f) 管理大樓\_正面

圖 10 實驗結果

## 結論

由於科技的快速發展，導覽系統幾乎成了出外旅遊的必備工具，但是只依靠GPS依然不足，通常使用者所期望得知的目標並非自身所在的位置，而是由此位置所能觀察到的目標，由實驗可知就算是同一地點由使用者觀測的角度不同場景物件也會有很大的變化，使用者所需的是更精確的資料和更直覺性的判斷。本研究利用影像辨識技術結合擴增實境技術將導覽資訊呈現，以強化導覽系統的精確度和操作上的直覺性。

本研究建立影像資料庫，影像資料庫內需儲存場景物件的特徵點、物件參數、導覽資訊，並且依據實驗結果建立樹狀分類器。同時利用SURF對使用者輸入的影像和資料庫內的特徵點進行比對，並且在各個步驟設立了不同的閾值，降低了實際應用時可能產生的誤判和相對速度的提升。最後將所得附加導覽資訊以擴增實境的方式顯示，導覽資訊內包含文字影像、相關連結等等。實驗結果中，以亞洲大學為例，成功的比對資料庫中的場景，並且根據輸入影像的不同，輸出資訊也隨之改變，同時也測試了不同的影像輸入皆有良好的成效。以亞洲大學為例，對於亞洲大學的主要建築，準確度達到96%，。

未來希望在影像資料庫中加入更多的場景物件，但是這勢必會增加計算量和分類的複雜度，整體系統的運算速度對於一般使用者的習慣仍嫌不足，必須不斷的優化演算法，以使得能夠有更廣泛的應用。同時，某些特殊的光線角度和過度曝光等拍攝時的誤差對於實驗的精確性也影響很大，降低建立資料庫時所需拍攝的數量也是一個很有挑戰性的課題。本研究以個人電腦為實驗環境，但是導覽系統必須用於行動裝置上才能取得最佳效果，未來將系統架構轉換為以雲端方式連結行動裝置的模式是一個必要的研究課題。

## 參考文獻

- 王國雄 (2010)。輔助古蹟寺廟行動學習之多標記擴增實境遊戲的互動模式研究。大同大學工業設計研究所碩士論文。
- 王湧天, 林偉, 劉越, 鄭偉 (2006)。亦真亦幻的戶外增強現實系統-圓明園的數字重建。中國科學基金 第 20 卷, 第2期: 76-80。
- 林武, 洪景新, 張昊, 李林 (2009)。快速有效的視頻圖像序列拼接方法。計算機工程與應用: 1002-8331-24-0173-03。
- 張樹安 (2010)。3D擴增實境應用於行動導覽之研究。政治大學數位內容研究所碩士論文。
- 郭其綱, 鄭泰昇 (2008)。擴增實境定位技術應用於建築與城市戶外導覽之研究—以「古蹟導覽系統」與「隱形招牌」應用為例。中華民國建築學會「建築學報」第66 期, 145~166 頁。
- 彭輝, 文友先, 翟瑞芳, 羅俊, 劉善梅 (2010)。結合SURF運算元和極線約束的柑橘立體圖像對匹配。計算機工程與應用: 1002-8331-08-0157-04。
- G. Bradski and A. Kaehler. (2008). *Learning OpenCV*. O'Reilly Media, Inc.
- H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, Vol.110, No.3, pp.346—359.
- K. Mikolajczyk and C. Schmid (2001). Indexing based on scale invariant interest points. *8th IEEE International Conference on Computer Vision*, vol. 1, pp. 525–531.
- K. Mikolajczyk and C. Schmid (2005). A performance evaluation of local descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615-1630.
- Lowe, David G. (1999). *Object recognition from local scale-invariant features*. Proceedings of the International Conference on Computer Vision.
- M. Swain and D. Ballard. (1991). Color indexing.. *International Journal of Computer Vision*, vol. 7, no. 1, pp.11-32.
- M. Tuceryan and A. Jain (1998). Texture analysis. *Handbook of Pattern Recognition and Computer Vision*, 2<sup>nd</sup> Edition, World Scientific Publishing Co., pp. 207-248,
- N. Dalal and B. Triggs (2005). Histograms of oriented gradients for human detection. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 886-893.
- R. Duda and P. Hart (2008) . Use of the hough trasformtion to detect lines and curves in pictures. *Comm.ACM*, Vol 15, No.1, pp.11-15.
- S. Se, D. Lowe and J. Little (2002). Global localization using distinctive visual features.in *Proc. IEEE/RSJ International Conference on Intelligent Robots and System*, vol. 1, pp. 226–231.
- Y. Ke and R. Sukthankar (2004). PCA-SIFT: a more distinctive representation for local image descriptors. *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 506-513.

## DEVELOPMENT OF A SMART NAVIGATION SYSTEM USING IMAGE MATCHING



*Yi-Chin Chu \* Hui-Fuang Ng*

Department of Computer Science and Information Engineering, Asia University(500,  
Lioufeng Rd., Wufeng, Taichung 41354, Taiwan)

\*Jashshadow@yahoo.com.tw

#### **ABSTRACT**

In this study, image matching combined with augmented reality techniques were used to enhance the functionality and accuracy of a navigation system. First, an image database for the scene was built. Images from different angles and distances for each of the objects of interest in the scene were taken and image features were extracted using SURF. Only the SURF features were stored in the database, along with the navigation information associated with each of the objects. During navigation, SURF features of the image provided by a user will be extracted and matched with the features in the database. The best match object in the database is identified and the navigation information of the matched object is retrieved. The scaling and rotation results from the feature matching process are used to perform necessary geometric transformation to the navigation information before showing them on the input image, thus achieving accurate augmented reality effect. A navigation system for Asia University campus was built to test the performance of the proposed method, with an accuracy of 96% for identifying the five main buildings in the campus.

**Keywords:** Feature Detection, Image Matching, SURF, Augmented Reality.