

影像查詢技術應用於展場導覽系統設計之研究

黃國峰 施成憲 林善智 游耿能

國立臺中科技大學多媒體設計系

kfhwang@nutc.edu.tw galetter@livemail.tw rexl1008@gmail.com gny@nutc.edu.tw

摘要

隨著智慧型手機的普及，許多日常生活中與手機互動的方式亦受到改變，影像辨識科技的進步亦影響導覽行為的輔助模式。本研究應用影像查詢技術於展場導覽系統之設計，我們先針對研究目進行相關文獻的探討，而後分兩階段進行研究，第一階段以影像辨識技術建構一個應用於展覽時可使用之行動導覽系統，此導覽系統使用電腦視覺演算法 SIFT 來提取特徵值方式以辨識展覽目標。第二階段則將設計出之導覽系統，實際應用於一個書法展覽中，在現場用隨機方式邀請了 21 位觀眾使用此導覽系統以進行測試，初步結果顯示以影像辨識技術建構之導覽系統，有助於使用者實際的導覽體驗。
關鍵詞：行動導覽、SIFT、影像查詢。

Abstract

With the popularity of smart phones, interaction styles with the phone in our everyday are also subject to change. On the other hand, the advances of image recognition technology will also affect the patterns of mobile guiding behavior. In this study, image retrieval technology is employed in the design of an exhibition guiding system using mobile devices. We first explore and survey relevant literatures for this research. Then the study is conducted with two stages. In the first stage, we constructed a mobile guiding system by applying image recognition technique for guiding people during visiting exhibitions. We adopted the SIFT Computer Vision Algorithm and eigenvalues to extract image features in order to recognize targets in exhibitions. In the second stage, the prototype of the mobile guiding system is used to help viewers of a calligraphy exhibition. In this the scenario of practical application, 21 viewers are randomly chosen and invited to use this system for a preliminary user study, the results show convincing positive response in applying image recognition technology to build the mobile guiding system that helps users navigating experience.

Keywords: Mobile Guiding System, SIFT, Image Search

1. 前言

隨著智慧型手機與行動裝置的普及，行動導覽

也漸漸成為導覽的使用設備。行動導覽不僅能為觀眾提供個人化的導覽解說，同時又能節省實體展覽資源與經費的耗費，行動導覽對展覽來說有利無害。國內常見的行動導覽，通常是語音、影片和互動 app 等導覽方式，而其中導覽查詢應用方面有一維條碼、二維條碼和網路定點呈現資料等。導覽查詢方式雖有多種選擇，但大多以定點給予查詢資料的概念為主，像是在展覽點貼上看板或媒體裝置，沒有更直覺式的查詢方式，若在資訊量多的展覽點上，使用者就必須去閱讀理解訊息，從中找出自己想找的東西，頓時就會耗費許多導覽時間。因此本研究嘗試建構使用這更直覺得查詢導覽系統。

影像處理已發展多年，其效率與實用性已普及在我們生活週遭，如相機中的環景功能、人臉辨識等。尤其是尺度不變特徵點演算法，其能達到抗旋轉、抗光影變化、抵抗遮蔽物等強大功能。然而既然電腦已能辨識影像，那麼將使用者想要找尋的影像交給電腦，再回傳需求的資料，就能輕易的達到使用者直覺選擇查詢項目的問題。為此想嘗試利用影像查詢輔助行動導覽的可行性成為本研究的動機。

2. 文獻探討

2.1. 導覽

本研究技術針對展場導覽，故查詢參考導覽方面研究，並對研究中需理解的參考資料做歸類，分別是導覽功能與行動導覽。

導覽的定義是「引導觀眾，藉由一種有意的安排，來實行引導活動，經由導覽人員來進行，以達到某種教育計畫性的目的」[1]。有導覽的展覽比觀眾自由參觀更易於將作品的意義交給參觀者，這是一種輔助觀眾參觀的方式。導覽在本研究中的功能為：透過導覽系統引導觀眾，使其對展覽標的物有完整的認識，並從中得到資訊、教育、娛樂等作用。

無線技術的進步，推進了行動導覽開始普遍應用於展場中，最先是由 PDA 開始，1993 年 Visible Interactive 公司開發了第一個互動影音導覽系統 iGo，並將其應用博物館導覽中，利用互動式多媒體，如文字、影、音等大大增加參訪者的個人經驗。後來陸續有許多大型導覽也開始嘗試行動導覽系統，除了行動導覽優越的互動性外，還兼具即時性，比如內容可提供行程表、餐廳資訊等。行動導

覽使得主辦者與參訪者距離大大縮短了，這意味著不僅是導覽，如研討會、都市導覽、觀光導覽等都能利用到行動導覽[2]。

2.2. SIFT 尺度不變特徵

尺度不變特徵之優勢，在於其能夠抵抗圖像的旋轉、縮放、變形、光影變化與抵抗遮蔽等效果，功能強大。而尺度不變特徵能做到這點是因為建構出能因應不同尺度的影像金字塔，並設計豐富訊息且特异性高的特徵點這兩個特性。本研究之影像查詢技術，就是應用尺度不變特徵點原理，找尋影像中特有的特徵以抵抗使用者上傳的各種不同影像變化。尺度不變特徵發展至今已有多種演算法，如 SIFT、PCA-SIFT、GLOH、SURF 等，不同的演算法皆有自身的優勢與特點。而本研究實驗工具應用的演算法是 SIFT 演算法，SIFT 是各類尺度不變特徵演算法的始祖，雖然跟後進演算法相比慢了許多，卻依然是抵抗旋轉與尺度變換效果最好的，也是本研究使用他的理由。

SIFT(Scale-invariant feature transform)是一種電腦視覺演算法，由 David Lowe 於 1999 年所發表，並至 2004 年才給予總結，也立定了尺度不變特徵中圖像金字塔與多訊息特徵點的概念。主要流程大致可分為以下四類：

1. 建立尺度空間。
2. 找尋特徵點。
3. 賦予特徵點方向性。
4. 建立特徵敘述子。

SIFT 演算法內容豐富，因此依以上列各點，再分小節給予詳細的解釋。

2.2.1. 建立尺度空間：

建立尺度空間中又可分為三步驟，第一步，建立圖層金字塔(Pyramid layer)：

將原始圖依等比長寬縮小成各種尺寸，如 1/2、1/4、1/8 等，直至某個最小尺寸為止，影像堆疊起來如同金字塔般，故稱為圖層金字塔，如圖 1。

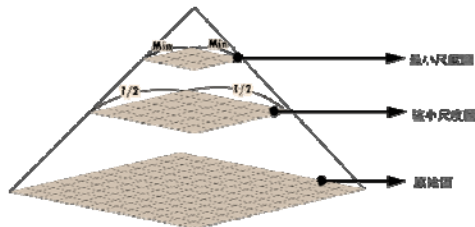


圖 1 圖層金字塔(Pyramid layer)

第二步，進階為高斯金字塔(Gaussian Pyramid)：

若將原始圖做為公式，則以 $I(x,y)$ 已表示，而可變高斯模糊函式則以 $G(x,y,n\sigma)$ 為表示(本研究 σ 表示為尺度大小)，把高斯模糊作為原圖的濾鏡產生的結果則可表示為 $L(x,y,n\sigma)$ ，公式如圖 2。

$$L(x,y,n\sigma) = I(x,y) * G(x,y,n\sigma)$$

圖 2 高斯模糊公式

接著，這邊對每層不同尺度的圖形進行如上述所說的高斯模糊，不同的是，每張圖型會分別以不同變量過濾多次，通常執行至少 4 次，其變量是越來越大，如 σ 、 2σ 、 3σ ...，以結果來看就是變得越來越模糊，高斯金字塔如圖 3。

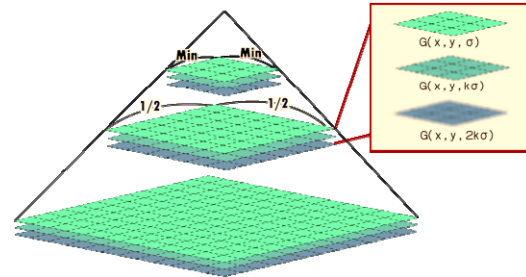


圖 3 高斯金字塔(Gaussian Pyramid)

第三步，得出高斯差異影像(Difference of Gaussian, DoG)：

DoG 這步驟是利用高斯金字塔產生的結果 $L(x,y,n\sigma)$ ，再兩兩相減產生差異圖形，如圖 4 所示，以利下一步驟找尋差異點。

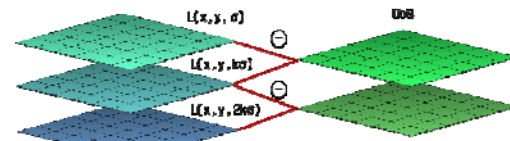


圖 4 高斯差異影像(DoG)

2.2.2. 找尋特徵點

SIFT 特徵點的找尋方式是用極大值檢定法，但他不只檢定單一 DoG 圖層，還包含圖像金字塔同一尺度中前後產生的 DoG 圖做比較，判斷每個像素在周圍鄰近 8 個點與上下兩 DoG 圖層各自的 9 個點，共 26 個點是否是最大或最小值，是則被標記為極值點，如圖 5。

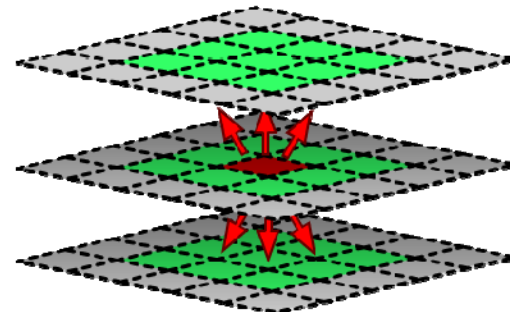


圖 5 極大值檢定法

獲得極值點後，為了得到更為精確的特徵，SIFT 又再次對每個極值點做 Hessian 角點過濾，Hessian 角點過濾特性是可獲得圖像明確的轉角，藉此可刪去在圖像邊緣的極值點，在此殘留下來的極值點可視為特徵點的中心。

2.2.3. 賦予特徵點方向性

找到特徵點後，會利用特徵點為中心，其半徑 $3*1.5\sigma$ 內所有像數進行角度與強度的計算，在將其歸類為8個主方向累計強度，強度最高的方向視為主方向，如圖6。

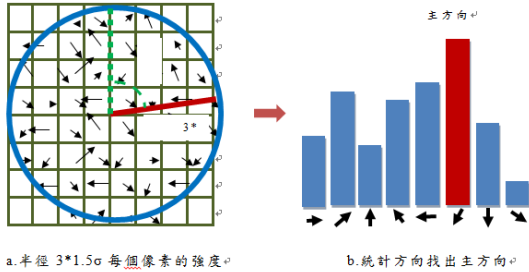


圖 6 SIFT 方向檢測

為了達成影像特徵的不變性，主方向確定後以特徵點為中心，將圖像旋轉 θ 角，再執行下一步的特徵資訊建立。

2.2.4. 建立特徵點敘述子

建立特徵敘述子第一步是以特徵點為中心，將四周分割為4乘4的區塊，再對每區塊計算8方位與其累積強度值，如圖7，最後的結果每個特徵點能獲得 $4*4*8=128$ 個特徵值。

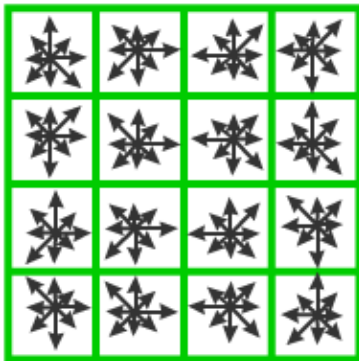


圖 7 SIFT 128 個特徵敘述子的生成

2.3. 尺度不變影像匹配應用

現今尺度不變演算以成熟，已利用於許多商務方面，如相機自動拼接的嘗試應用於衛星影像地圖的拼接研究[3]。利用其圖像辨識能力來解析交通號誌等[4]，以上足以證明尺度不變特徵發展潛力。

2.4. 介面互動設計

使用者介面定義為：整個系統除去影性數值運算、資料儲存、傳輸等所能見到聽到的就是使用者介面[5]，往後推就能明白，使用者介面的設計關聯到設計者、系統與使用者間的關係。Norman 提出一概念模型[6]。敘述這3者的關係，如圖8。

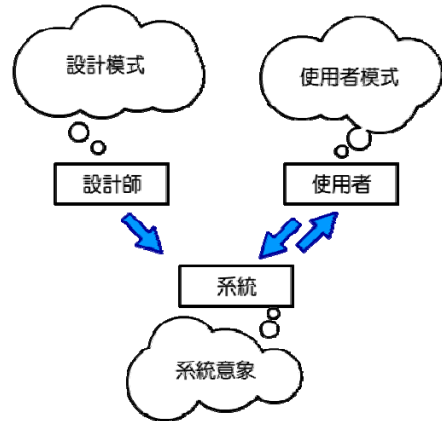


圖 8 設計者、系統與使用者關係圖

圖中表現出，設計師對於系統與使用者對於系統的認知是不同的，然而介面設計是設計師與使用者的對話，需要顧及互動性與操作過程，方能同理使用者的想法，減少使用者經驗上的挫折，如此做出有善於使用者的人機介面。

3. 系統設計

3.1. 系統設計方向

為了方便展方管理與編輯資料，為系統建立了給展方編輯用的後端系統，而後端系統又分兩個版本，分別是桌機與行動裝置版，後端行動裝置版系統是為了讓編輯者能快速取的與編修展場影像資料，而後端桌機版系統則是為了對於指展覽的詳細編修而設計，雙版本的設計除了各有優勢，還能同時一起使用，加速後端資料的編輯速度。後端系統完成後也為參訪者建立前端系統，前端系統相較於後端系統單純，僅有行動裝置板，而較特別的設計部分，是為了方便驗證與收集使用者反應，在前端系統尾端多設計了一套問卷系統，方便參訪者測試後直接填寫並傳回伺服器已統計，同時方便本研究做統計。

除了視覺介面上的前端系統與後端系統，視覺介面外的伺服器系統設計也是一大工程。伺服器系統本身負責與資料庫的溝通與回覆視覺介面的命令需求，同時解析影像特徵、影像匹配也是在這裡完成。查詢導覽系統簡易架構圖如圖9。

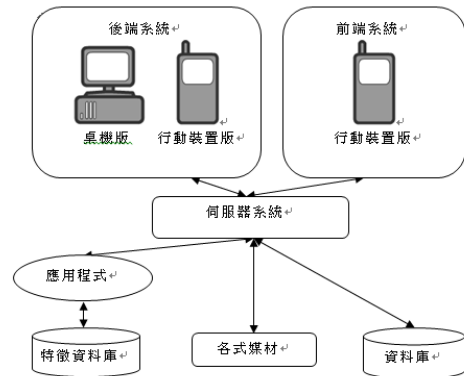


圖 9 查詢導覽系統簡易架構圖

3.2. 後端系統設計

後端編輯介面又分桌機板與行動裝置版，行動裝置版設計目的，是想利用行動裝置攜帶方便的特性，迅速把展場上的影像都收集歸類至資料庫內。為此目的，讓編輯者可在展覽會場直接拍照並做簡易的文字編輯，再上傳於伺服器中。相對於行動裝置版，桌機板的設計是為了詳細編修行動裝置版中建構的資料，也可快速的新增展覽項目、超連結與額外的多媒體檔案。行動裝置版與桌機板的設計，意味著資料建構能同時多人進行。

行動裝置版是讓編輯者能快速、簡易新增資料與配對測試的工具。因此依功能分類到編輯模式與使用者模式中，編輯模式主要讓編輯者拍照新增資料、瀏覽項目、更新與刪除等管理動作。使用者模式是在編輯者完成編輯後，可用在展場中使用此模式測試影像查詢結果。若無法查詢到照片，可能是照片拍照時搖晃造成模糊，抑或是亮度不足，造成解析度不足等問題，此時只要在編輯模式，重新拍照更新資料即可。編輯者能用以上方式，立即排除資料庫中有問題的圖像資料。介面架構圖如圖 10。

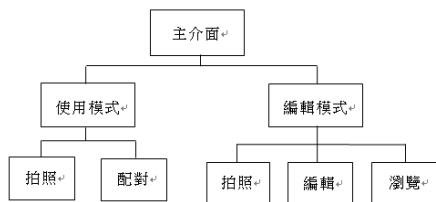


圖 10 行動裝置版介面架構圖

介面設計部分以簡單的漸層灰色調為主，如下圖 11 行動裝置版介面所示。

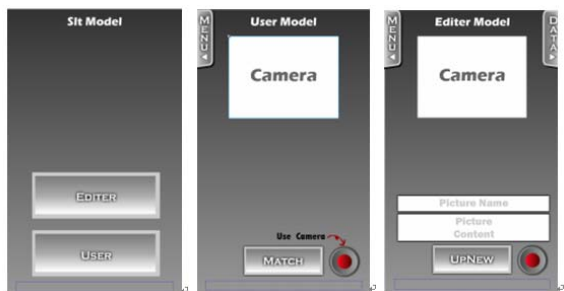


圖 11 行動裝置版主頁與各模式首頁

如上小節所說行動裝置版擁有使用者與編輯者兩模式，因此行動裝置版主頁，如圖 11a 簡單設計兩個按鈕 EDITER 與 USER 能連結進入這兩個模式。在各模式下，只要點擊左上角 MENU 按鈕，即可回到首頁，藉此切換模式。

使用者模式，如圖 11b，僅有此頁面，若要擷取照片可按下圖中紅色按鈕進入 Android 系統內建拍照模式進行拍照，之後照片會取代圖中 Camera 的部分。接著點擊 Match 便上傳照片至伺服器與資料庫中所有的照片配對，並將結果傳回圖中空白處。以上拍照與配對動作，可重複執行。

編輯者模式，如圖 11c，一進去是資料新增頁面，資料新增動作前得先完成拍攝，步驟如同使用

者模式，點擊畫面中紅色按鈕即可進入 Android 內建拍攝模組，取得照片後便取代 Camera 圖像。接著在 PictureName 與 PicContent 兩文字輸入欄位輸入圖像名稱與簡易的內容，最後點擊 UpNew 按鈕即可將資料新增於資料庫中。圖像名稱欄位無資料，則以日期時間替代名稱，紀錄於資料庫裡。欲檢查資料庫中現有的資料內容，可點擊右上 Data 按鈕切換到資料頁面，如圖 12。



圖 12 編輯者介面-資料頁面

資料頁面列表中，是現有資料的 id 資訊、影像縮圖、名稱與刪除按鈕。欲刪除資料點擊刪除按鈕即可。點擊 id 資訊、影像縮圖、名稱皆會進入資料更新頁面。頁面中 All Clear 的按鈕，是刪除資料庫所有資料功能，是為了方便測試而設立的功能。資料列表底下的 PAGE 物件，是分頁顯示物件，能提醒使用者目前的頁數。而右上角 New 按鈕，功能是回資料新增頁面。

資料管理於行動裝置版中，會發現不提供影片、聲音資料與網路連結的管理，因為行動裝置版定義是快速新增影像資料，上傳影音媒體的步驟與輸入連結網址等相對於桌機處理麻煩許多，尤其是將媒體檔案上傳伺服器，會對網路成為笨重的負擔。況且需新增影音資料通常已準備好在桌機裡，沒必要由行動裝置中執行，因此詳細管理資料部分交給桌機版處理即可。

在資料更新頁面中如圖 13，會發現並無拍攝按鈕，此頁面資料更新僅提供更新圖像名稱與內容，同樣於 PictureName 與 PicContent 兩文字欄位輸入圖像名稱與內容，再點擊 UpData 按鈕即可更新資料庫資訊。位於此頁右上方的兩個按鈕 New 與 Data 分別連接回到資料新增頁面與資料頁面。



圖 13 編輯者介面資料更新頁面

桌機板的設計期望是快速與詳細的編輯資料，因此介面偏重功能導向，沒有分頁，僅分出項目列表與資料編輯欄位兩個區塊，能讓編輯者一目了然方便操作。起始頁面如下圖 14，此時右邊資料編輯欄位成灰色表示未選擇資料。

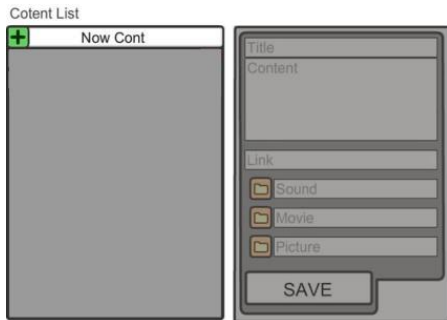


圖 14 桌機版起始介面(未有資料)

當按下 Now Cont 按鈕會增加一筆新的展覽項目，資料編輯欄位也跟著亮起，此時可編輯展覽項目，如標題、內容、連結與各項媒體檔案，如圖 15 所示。

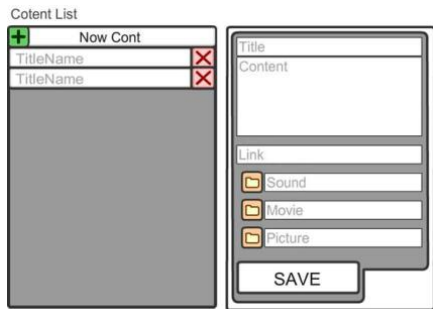


圖 15 桌機版起始介面(新增資料)

其中資料夾符號表示為瀏覽檔案，會開啟瀏覽檔案視窗，方便使用者選擇媒體，編輯完成再按 SAVE 按鈕方可儲存入資料庫，此時會順便將存入的圖片做特徵值處理。而項目中 X 的按鈕表示刪除項目，此功能會清除所有內容、儲存在伺服器端的媒體檔案與特徵檔。

3.3. 前端系統介面

前端系統簡易架構如下圖 16，主要有 7 個頁面，系統實際操作時的主體是拍攝頁、資料展示頁與多選頁。

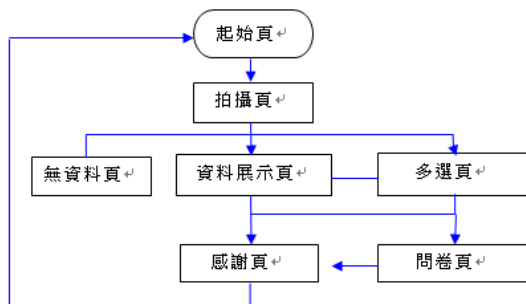


圖 16 前端介面架構圖

3.4. 伺服器設計

編輯者與使用者各自擁有分開的介面，各介面最後會將命令傳到到伺服器上傳或取得各式的資料與媒體，伺服器整體架構如圖 17 所示。來到伺服器的命令後續動作可分 A、B、C 三類，A 為透過網路動態語言與資料庫做溝通，此時取得資料與媒材檔案名稱或將資訊建於資料庫中，接著回傳結果於使用介面。B 為透過網路動態語言啟動應用程式，接著應用程式會執行特徵值資料庫相關的處理動作，可能是分析製作資料或比對特徵。C 部分是上傳媒材檔案或取得媒材的過程。

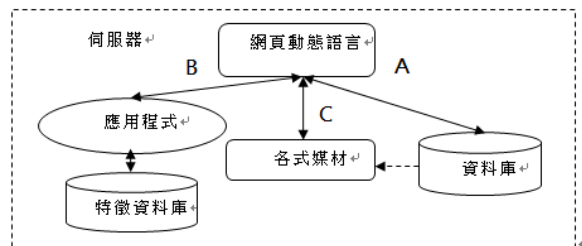


圖 17 伺服器簡易架構

這部分包含資料庫結構與命令流程設計，而命令流程是指使用者向伺服器回傳的命令項目，可分成四類，分別是：

- a. 上傳資料與特徵值處理
- b. 修改資料
- c. 刪除資料
- d. 匹配與查詢資料

4. 實驗結果與分析

系統完成後選擇 2013 中區書法聯展為實驗對象，展場狀況如圖 18，目的考驗系統對於平面展覽作品的抗旋轉、抗光影與抗角度的成效。在實驗中所使用的行動裝置載體為 Asus Transformer TF 101 平板電腦。



圖 18 2013 中區書法聯展，台中科技大學展次

2013 中區書法聯展是國立台中科技大學、國立中興大學與逢甲大學三校書法社聯展，作品繁多在三校中輪流展出，本次實驗是台中科技大學場次，辦於台中科技大學，中商大樓一樓川堂，展場示意圖如圖 19 所示。

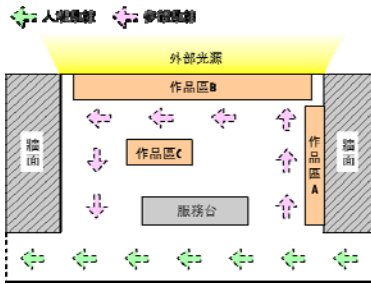


圖 19 展場示意圖

透過展場示意圖中，可了解展覽分 A、B、C 三個展區，A 展區靠於牆面上，B 展區則是靠於在窗戶上，C 展區則是平躺在桌面上的作品，展場示意如圖 19。

從以上可對展覽看出幾以下幾點：

- a. B 區光線對外開放，各展區光源皆會受到天候影響，尤其 B 展區拍攝上逆光嚴重。
- b. C 區展覽作品平躺於桌上，參觀者無法平面拍攝。

以第一點來說，雖然光線影響嚴重，但卻可以用來測試本系統查詢能力是否受光源影響。而對於第二點來說，原先資料庫中的對應影像是平面的，但到了現場，平躺角度關係讓作品拍攝後變形過份嚴重，如圖 20 所示。

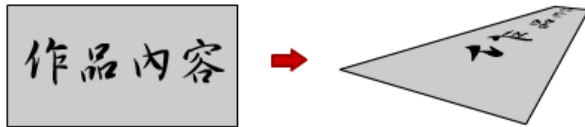


圖 20 拍攝變形示意圖

拍攝到的角度，無法每次都順利辨識，因此把其當成立體作品看待，對此作品在展覽中拍攝 4 個面，如圖 21 分別將不同的照片與特徵值和相同的作品敘述，分別紀錄於資料庫中，目的就是使用者的角度建構以減少搜尋誤差。

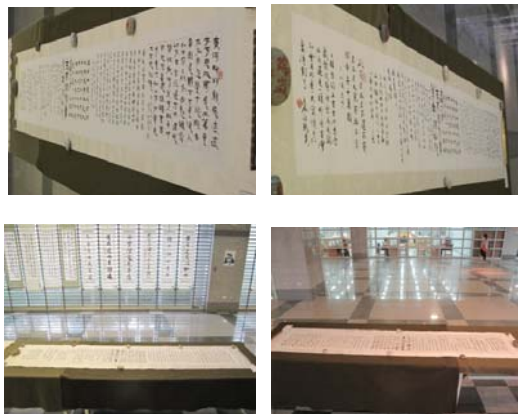


圖 21 作品區 C 重新拍攝的 4 個面

經過使用者操作後，訪問結果表示，有少數使用者認為本系統幫助不大，原因是輔助的媒體不夠多。而且其中有女性表示，長時間使用平板電腦對於女性使用者來說太過沉重，但使用者普遍認為此輔助導覽裝置是便利的。

5. 結論與未來展望

本研究以影像辨識技術建構一個應用於展覽的行動導覽系統，此導覽系統運用 SIFT 演算法來辨識使用者欲導覽之目標。目前系統所使用 SIFT 演算法對於比對資料仍有極限，資料庫中超過 50 筆資料，運作上會出現緩慢的情形，未來可嘗試 SURF 或 CCH SIFT 等其他方式來改善資料過多所產生的緩慢情形。

參考文獻

- [1]余少卿，梁朝雲，莊育振 (2003)。行動數位導覽之博物館應用探討。圖書資訊學刊，卷期:1，頁 1-24。
- [2]蕭顯勝，黃向偉，洪琬諦 (2007)。行動導覽系統於博物館學習之研究。高雄師大學。
- [3]李慧，蘭啟忠，劉慶杰(2012)。基於 FAST 和 SURF 的遙感圖片自動配準方法。國土資源遙感，卷期:93:2，頁 28-33。
- [4]胡錦城，李實英，李仁發(2012)。基於高穩定 SURF 特徵的交通標誌識別。計算機應用研究，卷期:29:8，頁 3179-3181。
- [5]Lauesen, S. (2005), User interface design: A software engineering perspective. Addison Wesley
- [6]Norman Donald A. (2007), 設計&日常生活:如何選擇安全好用的日常生活用品，台北：遠流。