

## 利用多重影像產生之點雲的精度評估\*

黃金聰\*\* 陳思翰\*\*\*

論文收件日期：100年11月28日

論文接受日期：102年3月15日

### 摘 要

目前已經有多款利用網路雲端技術進行三維模型建構的軟體發表，並提供免費使用，其中以微軟最早於2008年推出的Photosynth最被常使用。該軟體是運用Scale-Invariant Feature Transform (SIFT) 以及Structure from Motion (SfM) 的技術，將同一場景不同位置拍攝的影像進行特徵萃取及匹配後，重建攝影站位置以及被拍攝物體的三維空間坐標，再結合瀏覽器供3D場景展示。由於操作簡單且運算快速，可以讓使用者使用消費型數位相機輕易建構出虛擬實境，也提供物件嵌入網頁分享成果給其他使用者。本研究透過相關實驗數據探多重影像以Photosynth方法進行物空間坐標重建時定位可達之精度，藉由實驗設計分析相關數據，以全測站經緯儀測算的規標點三維坐標為依據，並將本文提出之方法的成果與近景攝影 (Close-range photogrammetry)、光達等其他空間資訊重建方法的成果做比較，以瞭解運用該法時的三維空間定位精度，並瞭解此方法的限制與可能面臨的問題。

本文所提之方法是利用一般消費型數位相機進行拍攝預先佈設控制規標的實驗區，並利用與已知規標點坐標差異的RMSE為精度評估指標。實驗顯示，在適當的拍攝因素控制下，檢核點精度之三維精度分別為 $\pm 0.027\text{m}$ 、 $\pm 0.065\text{m}$ 以及 $\pm 0.012\text{m}$ 左右。建構出適宜應用精度的三維模型，經濟、便利且可靠的完成空間樣貌紀錄，對於建物保存工作將有所助益，亦提供未來建構三維模型的另一種選擇。

關鍵詞：近景攝影測量、光達、點雲

---

\* 本研究接受行政院國科會NSC 100-2410-H305-065之補助，特此致謝。

\*\* 副教授，國立台北大學不動產與城鄉環境學系，TEL：(02) 86741111#67426，E-mail: jthwang@gm.ntpu.edu.tw

\*\*\* 碩士，國立台北大學不動產與城鄉環境學系

# Accuracy Assessment of Point Cloud Generated by Multiple Images\*

Jin-Tsong Hwang\*\*, Szu-Han Chen\*\*\*

## Abstract

There are many of 3D model reconstruction software based on cloud computation technology released recently. Photosynth is the earliest one which based on Scale-invariant feature transform (SIFT) and Structure from Motion (SfM). SIFT gives function of invariance of scale, rotation, perspective and lighting for feature point extraction. This kind of approach is automatic feature point extraction and matching. SfM can restore camera station and acquire space coordinates by continuous photos. In 2008, Microsoft released a free software named Photosynth which combined with SIFT and SfM technology. It can analyze digital photographs and generate a 3D point cloud of a photographed object.

In this paper, the accuracy was estimated by the index of RMSE of the point cloud generated from Photosynth, Lidar, and close-range photogrammetry by using some empirical data. The results indicate that the 3D positioning accuracy of Photosynth approach under well control is about  $\pm 0.027\text{m}$ ,  $\pm 0.065\text{m}$ , and  $\pm 0.012\text{m}$  respectively. The result is used to compare to that of close-range Photogrammetry and Lidar. We found out the restriction and the problem in this method either.

**Keywords:** Close-range Photogrammetry, Lidar, Point Cloud

---

\* This research has been supported by NSC100-2410-H305-065, which is gratefully acknowledged.

\*\* Associate Professor, Department of Real Estate and Built Environment, National Taipei University, TEL: +886-2-86741111#67426, E-mail: jthwang@gm.ntpu.edu.tw.

\*\*\* Master, Department of Real Estate and Built Environment, National Taipei University.

## 一、前 言

隨著二維平面影像發展已趨於成熟，為了追求更佳的可視化與多元的使用，三維建模技術是目前空間量測與地理資訊系統極力發展的方向之一，就目前技術水準而言，可以透過近景攝影測量、航空攝影測量、光達系統等方式，獲取建物實體資訊來建構三維模型；而也能夠利用數值地形圖或建物平面圖等圖資，配合樓高等相關建物屬性，建構數值化處理的三維模型，如張毅雄（2010）即利用建物測量成果圖來建立地籍建物資訊模型。張智安與陳良健（2006）即利用空載光達點雲與地形圖相互結合，進行房屋重建。然而各種方式就性質上與使用目的上皆有所差異，就使用方便性而言，近景攝影建模是最為快速且資料較容易取得的方法，只要利用一般數位相機拍攝，再經由適當率定及校估，便可作為建模使用，相當具有效率。在2006年微軟推出Photosynth瀏覽器，結合近景攝影、SIFT、SfM等技術（Noah et al., 2007），可進行影像的拼接與展示，是攝影測量上的新發展的技術，此外，可以產出具有三維坐標的點雲資料，可藉此資料建構模型。

在近景攝影建立三維模型上，王正忠（2002）利用航空攝影測量建立較為簡化的三維建物模型，並配合近景攝影測量技術來重建較為細緻的三維建物模型，其以最小二乘法建立模型，以配合影像套疊求解像片外方位元素，並以半自動化方式來重建模型式建物。饒見有等（2005）以近景攝影測量搭配電腦繪圖技術，利用高解析廣角數位影像，快速有效的建構擬真城市模型（Photo-Realistic City Model），並貼附房屋外牆紋理，將模型在地圖上做展示。張桓、蔡富安（2009）以單張影像重建三維建物模型，由具有透視投影的圖畫或是照片中萃取出特徵線，計算滅點與尋找出特徵點。利用滅點進行幾何計算取得特徵點三維坐標，建立三維建物模型，最後敷貼紋理與進行動態展示。方偉凱、黃灝雄（2005）利用地面光達建立建物模型，並對其精度做研究，其結果發現，點雲定位精度約 $\pm 0.05\text{m}$ 。曾義星等（2008）對於掃描儀的定位精度誤差，提出以六參數加上附加參數同時平差計算之方法，可以有效降低儀器系統誤差之影響，並提高觀測精度。Pomaska（2009）敘述由雲端下載Photosynth點雲資料的程序，以及如何利用點雲資料建構三維模型。Dowling et al.（2009）藉由拍攝地表影像，經由Photosynth做匹配處理及點雲產生後，可以快速生成小範圍的數值高程模型（Digital Elevation Model, DEM）資料，辨別地表的起伏情況，並依適當的參考尺度可得知真實DEM的大小。

目前用來建構三維模型常用的方法是採用近景攝影測量以及光達掃描，但前者需要較多的人為介入，後者則需要較為昂貴的設備方能達成，因此，以低價位非量

測型數位相機作為三維建模的設備，藉由拍攝建物影像來獲取外觀坐標及色彩資訊，可視為經濟、快速的方法。然而此方法所使用之相機未經過率定，影像也未經過校正，在偵測建物特徵與點位匹配上，其有效性及準確性有必要進行檢驗。

本文以全測站經緯儀測算的規標點為參考坐標，藉以比較Photosynth之方法重建成果與近景攝影（以PhotoModeler為處理軟體）、光達等其他空間資訊重建方法的成果，以瞭解運用本文所提方法的三維空間定位精度，並瞭解此方法的限制與可能面臨的問題。

## 二、理論基礎

本研究使用Photosynth的影像重建及瀏覽技術，主要可以分為兩個部分，首先為空間幾何重建，主要是利用尺度不變特徵轉換（Scale-invariant feature transform, SIFT）以及由運動獲取空間結構（Structure from Motion, SfM）方法重建相機位置和視覺場景的幾何關係，透過特徵找尋及匹配，使其能對應不同拍攝角度的影像，再還原拍攝位置與物體間的空間坐標系統，來進行影像式建模；其次是藉由建立照片展示系統，能夠讓使用者自由瀏覽，並透過照片導航點雲的空間位置。

### （一）尺度不變特徵轉換（SIFT）

SIFT是一種電腦視覺的演算法，是Lowe在1999年所提出的，2004年做出演算法的總結（Lowe, 1999, 2004）。此方法主要是用來偵測與描述影像中的局部特徵，所謂的局部特徵，是指如邊緣、點及輪廓等，藉由在特徵空間尺度中尋找影像中的極值點，以找出穩定的特徵，並希望該特徵對於影像旋轉、尺度縮放、視角變換及光照強弱等影像變化亦能夠保持良好的不變性，藉以在不同影像上正確匹配。SIFT主要是解決了尺度上變化的影響，讓不同尺度下的影像可以得到正確的匹配，讓整體成效更快更好。SIFT演算法主要分為特徵點萃取與特徵點匹配兩步驟：

#### 1. 特徵點萃取

主要是利用高斯差分（Difference of Gaussian, DoG）作尺度空間的極值偵測，萃取特徵點位和篩選穩定的特徵點，再描述其特徵向量，使各特徵點具有其代表性，以利於後續在不同影像的匹配。運算程序包含：

##### （1）尺度空間極值偵測

利用尺度空間的極值偵測來尋找不同尺度中具有代表性的特徵，尺度空間的建

立主要由Koenderink (1984) 和Lindeberg (1994) 陸續研究出，其研究指出一張影像經過高斯函數計算後，模糊化的影像能夠代表不同的空間尺度，為了找出尺度空間中穩定的特徵點位，將每組影像的相鄰高斯模糊影像分別計算高斯差分，以得到高斯差分函數DoG，計算如式(1)。

$$D(x, y, \sigma) - L(x, y, k\sigma) - L(x, y, \sigma) \dots\dots\dots (1)$$

其中 $D(x, y, \sigma)$ 為高斯差分後的影像， $k$ 代表高斯模糊的尺度比值，是一常數，一般設為 $\sqrt{2}$ 。DoG的運算方法首先將影像金字塔分為不同尺寸，而每一組都有 $S$ 層，在 $k=2^{1/S}$ 的情況下，每組的影像數必須為 $S+3$ 個，極值偵測才能夠有效涵蓋整組影像。在過程中，原始影像以標準差 $\sigma$ 值作高斯模糊，將 $k$ 值設為 $\sqrt{2}$ 。到另一個階層中，將影像解析力降低兩倍，並將高斯模糊的標準差乘以二倍的 $\sigma$ ，仍然對著個尺度的影像做四次高斯模糊處理。總共要做四組不同取樣倍數。求出四組高斯函數影像後，將每組中的影像，兩兩相減，可得到高斯差分的影像圖。

在取得高斯差分影像後，在每個特徵點上與相鄰8個像元，以及在影像金字塔前、後尺度相鄰各9個像元，一共計26個像元作比較找出極值點。若該點對於其他周遭26個像元相比，皆為極大值或極小值的情況下，則此點被視為是極值點而被保留下來。

### (2) 特徵點定位與篩選

特徵點位置的確認主要是為了提高前一步驟所做極值點的精度到次像元，希望能夠找到最大值的位置，並將低對比或不夠顯著的特徵點位做過濾，其方法主要透過一個三維二次方程式來達成，利用泰勒展開式將此非線性的方程式展開若在方程式的極大或極小值的位置其斜率為0，換言之方程式微分後為0。以偏微分計算極值點的次像元偏移量。

### (3) 特徵點梯度方向指定

為了要給予特徵點描述值，需要確定特徵點的梯度方向，以作為後續的描述向量使用，確認出主要方向之後，便可以達到特徵點旋轉不變的效果。利用10度一個區間的方位直方圖，來代表特徵點在360度的方向，峰值最大則為特徵點主方向，若在直方圖內有其他方向超過最大值80%以上，則該方向也屬於主方向。

### (4) 特徵點描述向量建立

在得到每個特徵點的主方向後，便可以描述特徵點的向量，讓其具有穩定性，不受到尺度、旋轉、視角等變化影響，以作為影像匹配時的依據。由圖1左圖所示，正中央為特徵點位，每小格代表特徵點鄰近周遭尺度空間的像元，共有 $8 \times 8$ 共64格，箭頭方向代表該像元的梯度方向，圓圈代表高斯加權的範圍。一般以 $4 \times 4$ 的

網格計算方向的累加值，其描述矩陣較為穩定，由於每個矩陣含有8個向量，故每個特徵描述元具有 $4 \times 4 \times 8$ ，共128個方向的特徵向量，即代表特徵點的描述。圖1右圖是以 $2 \times 2$ 的方向矩陣，特徵向量為 $2 \times 2 \times 8$ 共32個為例，藉此展示各網格內的方向性。圖2表示特徵點描述向量在影像上的方向性。

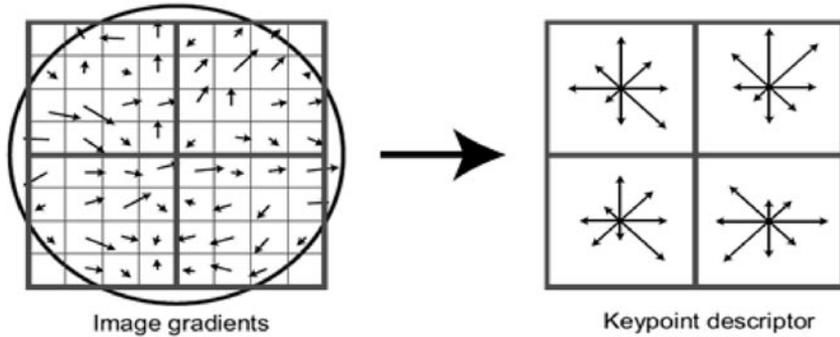


圖1 特徵向量描述示意圖（ $2 \times 2$ 方向矩陣）



圖2 特徵描述向量示意圖

## 2. 特徵點匹配

在特徵點匹配上，主要是考量特徵點匹配的方式，以及如何提高匹配的精度，由於特徵點的特徵向量在保持穩定下，便能在不同影像中達到良好的匹配，Lowe（2004）對於尋找不同影像特徵點匹配方法，主要是利用計算兩張不同影像中特徵點位的歐基里德距離（Euclidean Distance），所謂的歐基里德距離是指在空間上，

兩點之間的真實距離。如果最近的距離與次近的距離的比值小於自定門檻值，則接受匹配，門檻值越低則能夠匹配的特徵點會減少，但匹配結果會更加可靠。此外，利用隨機取樣一致性演算法（RANDOM Sample Consensus, RANSAC），進行匹配點除錯。RANSAC演算法對於分布不均勻的點位具有良好分辨的特性，以多數決的方式迭代，讓少數誤差較大的點位被排除，進而減少特徵點位在匹配上的錯誤。

## （二）由運動獲取空間結構（SfM）

從運動中獲取空間結構是影像式建模很重要的方法，SfM最早在1992年由Tomasi和Kanade（1991）所提出，其方法在於找尋出不同影像中的相同特徵點位，再由SfM恢復拍攝影像時相機的位置與姿態參數，再藉由不同的相機位置推求出拍攝物體的實際三維坐標。Photosynth藉此方法來建立可無縫瀏覽的影像群與具有三維坐標的點雲資料。SfM最主要的目的是由不同拍攝位置的二維影像建立對應的關係，透過分析現場不同影像相對或絕對的位置，來獲取具有坐標系統的三維場景。SfM有三個主要任務：第一，分析移動位置的相機和拍攝目標的相對關係；其次，恢復相機移動的軌跡；最後，對目標建立三維場景（Zhou, 2010）。

SfM在處理影像過程是將影像一張接一張配對，不斷的累加下去，在遇到不匹配的影像則自動排除，進而建構出各個相機相對的位置。在配對過程中，對SIFT演算法所偵測出的特徵點位，對每張影像追蹤，分別觀察場景與相機運動的相互關係。一般SfM處理程序及示意圖3如下：

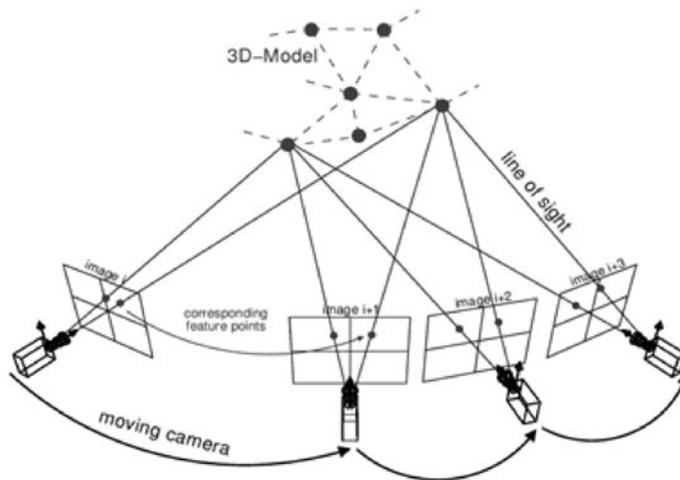


圖3 SfM程序示意圖（資料來源：Zhou, 2010）

SfM處理的程序為：

### 1. 影像特徵萃取

藉由SIFT演算法萃取出影像的特徵點位，並記錄在各影像上，作為爾後連續影像點位追蹤使用。事實上，是利用相機拍攝時像片的EXIF取得各影像基本屬性資料，以作為相機連續移動的估算。

### 2. 估算空間物體模型與相機位置

初步估計建構相機位置，並利用估計的動作來提供資訊，重建標的物。再利用特徵點的核面幾何原理，將相機位置計算出來，並利用特徵點沿運動軌跡方向，建立兩兩圖像中的三角點位對應關係。

### 3. 優化運算成果

光束法平差在SfM方法中，主要是為了優化點位的估計值，利用多張重疊影像互相約制，以共線式將相機內外方位元素與點位進行空間後方交會與空間前方交會，以計算點位坐標，藉由不斷迭代，使其數值收斂。

### 4. 建立點位幾何結構

SfM估算出最後的特徵點位，會根據相對應位置而成自定義的坐標系統，即在Photosynth瀏覽器所呈現的點雲結構，在Photosynth資料中除了包含點位三維坐標外，也包含點位RGB色彩資料。

## (三) 近景攝影與像機率定

所謂的近景攝影測量主要指由攝影站與被拍攝物體間距離較近，不超過300公尺者，主要是應用於建築、醫學、工程及其他不可直接量測的事物上，隨著數位相機的發展快速，近景攝影測量在精度上和影像處理上也達到更好的水準。

在近景攝影作業上，所使用的相機一般分為量測相機（Metric Camera）與非量測相機（Non-Metric Camera），量測相機經過精密的內方位元素率定，具有框標作為量測基準，是專門作為攝影測量之使用；而非量測相機就如同一般大眾所使用的消費型數位相機。近年來隨著數位相機技術發展越來越成熟，解析度有所提升以及價格便宜等優點上，使得目前非量測相機的運用越來越廣泛，但因非量測相機出廠時內方位並未經過率定，透鏡畸變差也較大，且沒有可判讀的框標，故需要經過率定程序才能作為攝影測量使用。以PhotoModeler為例，率定功能主要利用拍攝其軟體附有專用的率定圖檔，此率定圖共100個點，其中四點為控制點，如圖4，其拍攝

原則在於四個方向約45度角拍攝圖紙，不同方向再轉換相機角度，每張照片都需要拍到圖片的四個控制點，拍攝照片數量以8至12張為基準，由軟體自動匹配控制點坐標，並對照片作自率光束法，以求取相機的率定參數，其中包含輻射畸變差與離心畸變差。本文是利用PhotoModeler Pro 5作為近景攝影率定的使用軟體，利用其本身既有的相機率定功能操作。

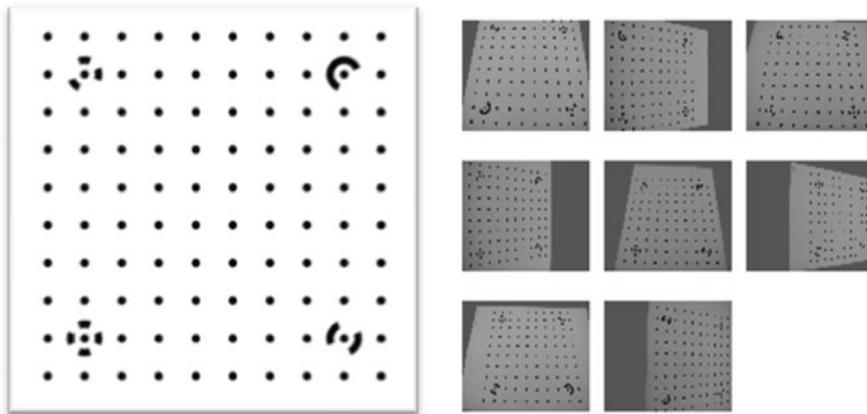


圖4 PhotoModeler的率定圖紙及拍攝影像

Photosynth在拍攝方式上也與PhotoModeler不同，根據其使用手冊所示，在2D平面拍攝上，以平行移動及前後移動作為原則，不需要從不同角度拍攝，即可達到良好的套合成功率與點雲產生量，在本研究中將率定圖紙依上述方法平行拍攝。拍攝完成的影像，需要透過Photosynth軟體上傳至其伺服器作處理，影像群可以由Photosynth瀏覽器作展示。並可使用SynthExport網站 (<http://synthexport.codeplex.com/>) 提供轉檔的程式外，將空間點雲以及Photosynth每張影像相機位置、焦距與透鏡畸變差等資料下載。

由於原PhotoModeler的率定紙為圓標為黑點，在Photosynth產生之點雲不易辨認，因此本研究將圓標附與不同色彩後再進行拍攝再以Meshlab進行點雲編修。在精度分析方面，由於尺度不同造成特徵點除了圖紙給定的點外，在每四個點的中心也會產生特徵點位，如圖中灰點（圖5），其產生原因在於Photosynth特徵點偵測時，灰點正好為於四彩色點中央，對影像而言該位置處於一個極值，故在影像匹配時被保留下來，然而本研究僅保留率定點作為使用，藉由點的位置與顏色來判別，來刪除多餘的點位。

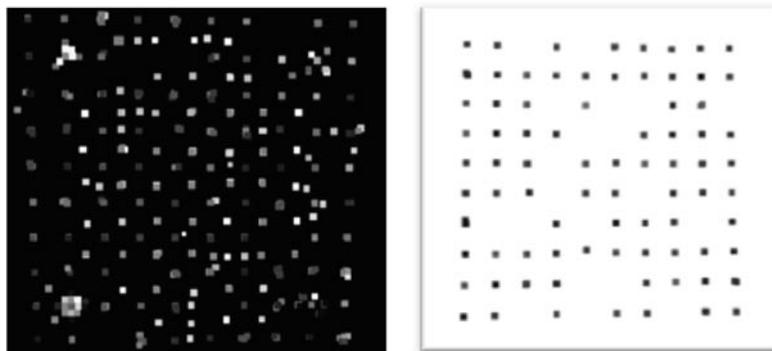


圖5 編修前後之點雲圖

根據PhotoModeler率定紙，四個相鄰控制點的距離相距100cm，其它各點的位置呈等距分布，故可得知實際點位距離，各點X、Y方向皆差距14.2857cm。據此，利用四個控制點來進行六參數轉換，將Photosynth的任意坐標轉換為PhotoModeler率定標的坐標系統及尺度，由量測點位與率定標實際點位計算RMSE。

表1 PhotoModeler與Photosynth像機率定比較表

項目	PhotoModeler	Photosynth
張數	8	7
率定焦距 (mm)	7.7247	7.5985
點數	100	80
RMSE (mm)	±0.0637	全部點位 ±1.9331
		扣除控制點外圍點位 ±1.4152
輻射畸變K1	0.001262	每張照片有單獨率定值
輻射畸變K2	0.00004105	每張照片有單獨率定值
離心畸變差P1	-0.00003033	無
離心畸變差P2	-0.0009758	無

### 三、實驗設計

#### (一) 研究區範圍

本研究區主要以臺北大學三峽校區公共事務學院東北面為實驗範圍，公共事務大樓東北面長約60m、高約35m，圖6為實驗區外觀。在牆面及玻璃上貼附自行設計的規標，作近景攝影、Photosynth及光達之三維定位精度的研究。牆面對攝影站而言有不同景深，規標的分佈儘可能同時顧及平面與景深的均勻分布。



圖6 實驗區範圍

#### (二) 研究流程

本研究在三維建模上，主要採近景攝影測量、Photosynth與地面光達測量三種方法。近景攝影測量與Photosynth都是利用數位相機拍攝影像來建模，而地面光達則利用儀器掃描方式產生點雲資料。

定位精度的評估，主要以公共事務學院東北面的規標點為依據，先在本建物牆上及玻璃上佈設60個規標，作為量測的控制點與檢核點。坐標系統的建立是利用實驗區附近的兩個地面控制點以GPS靜態測量，解算時與鄰近的台北大學衛星接收站和樹林高鐵衛星接收站一起平差，求得兩個地面控制點的X, Y, Z坐標。在其中一個地面控制點上架設全測站經緯儀，以另一個地面控制點定向後。觀測並解算各規標的三維坐標，以作為三種不同方法在量測精度上的評估依據。

在近景攝影與Photosynth部分，是利用數位相機拍攝含規標的牆面，再進行後續的工作。拍攝位置距離牆面大約介於10公尺至30公尺之間。PhotoModeler需要將

每張影像目視選取特徵點位置，並且人為指定不同像片中共軛特徵點的對應關係後，以直接線性轉換方式進行多張像片的平差處理獲得物空間點位坐標；而Photosynth則需將拍攝的像片上傳至雲端的伺服器完成處理後，再下載點雲進行編修及提取規標點的點位。在光達部分，本研究使用單站掃描的方式，以距離100公尺點雲間距2公分的模式掃描，由於融入了光達儀器反射專用規標圖形的設計，掃描後可由程式自動辨識規標點的位置。規標點位置萃取後，皆採用同一組已具地面控制點坐標系統的規標點做為坐標系統轉換的控制點。藉由這三種方法獲得的規標點點位三維坐標再與全測站經緯儀的觀測成果進行比較，以作為定位精度評估的依據。

### (三) 使用設備

坐標系統的建構是使用GPS接收儀、全測站經緯儀進行控制點坐標測算，其中GPS是使用Topcon GRS-1衛星接收儀結合G3-A1天線，可有效消除多路徑影響，以提高靜態測量的精度。此外使用Leica TPS 700全測站，來量測各規標點位之三維坐標，以作為後續精度分析的參考。光達資料則以Leica Scanstation C10雷射掃描儀進行掃描，掃描後的點雲以選取的規標點做為坐標轉換的依據。

### (四) 規標點位佈設

本研究規標主要對Leica地面光達儀器規標作改良，讓其能夠在地面光達、Photosynth中都能被辨認，圖7中心圓形的黑白部分，是Leica地面光達儀器規標，藉此能夠被Leica軟體辨識並直接萃取點位；在PhotoModeler在選點上，由於其黑白部分刻畫出十字形狀，對影像作人工選點時，也容易辨識。而在Photosynth選點上，由於中心圓形部分對於其點雲產生並不是一個很好的特徵點，依據實驗，可知其對顏色的特徵有較好的辨認效果，所以可以透過四周具有顏色（藍、紅）的圓形標來確定規標點的中心位置（如圖8），如果無法產生兩色四點的規標點雲時該點將被放棄使用。

規標點位佈設上，在公共事務學院東北面共佈設60個規標，其中20個為控制點，作為三維坐標轉換依據，其餘40個則為檢核點使用，如圖9所顯示的位置。

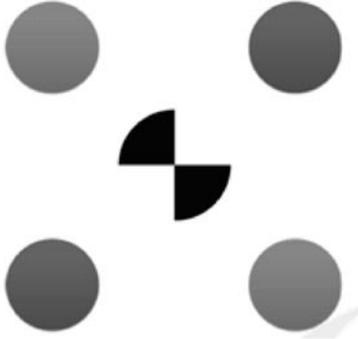


圖7 規標示意圖



圖8 Photosynth規標點鄰近點雲



圖9 規標點分布位置示意圖

#### (五) 近景攝影測量處理程序

近景攝影測量是以PhotoModeler軟體為主，三維精度分析處理程序如下：

1. 像機率定，進行相機焦距、像主點以及透鏡畸變差的率定。
2. 攝影，匯入研究區挑選良好重疊與不同拍攝位置的影像共12張（圖10）。
3. 標記影像中的各規標點位，共60個點（圖11）。
4. 將不同影像之共軛點位進行關聯，讓各影像間的規標點能互相對應。
5. 自動化點空間坐標解算，計算出各點位的局部參考坐標。
6. 匯出各點位坐標成XYZ檔案格式。將局部參考坐標以20個控制點利用七參數轉換為與全測站相同之大地坐標系統。

7. 以40個檢核點與全測站測得的坐標比較計算RMSE。

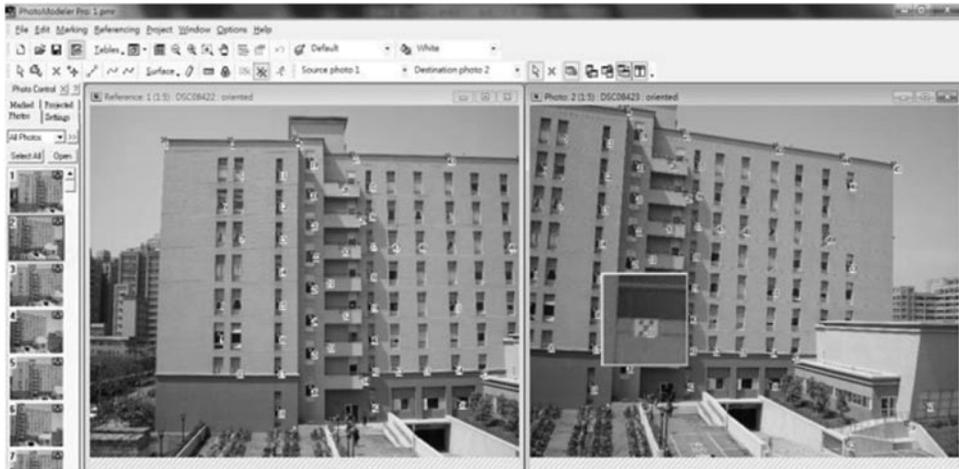


圖10 PhotoModeler中拍攝影像與點位標定及對應

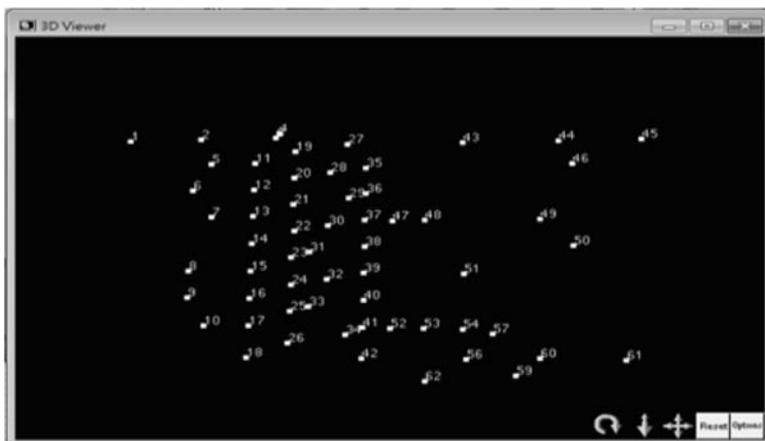


圖11 PhotoModeler中控制點三維點位展示

#### (六) Photosynth 處理程序

拍攝影像時以不同距離、不同高度拍攝，站與站之間距離大約1~2公尺間，但礙於實際空間的限制，僅有少數幾個攝影站有較高的約制。利用Photosynth產生點雲時，同一目標物可能由不同像片中重建，因此會產生多個點，在辨認規標點的點雲時以顏色為過濾依據（本文為藍色及紅色組成的規標），將具規標色彩的點雲選出後計算這群點雲的重心做為規標點位。Photosynth三維精度分析處理程序如下：

1. 利用Photosynth產生點雲資料，以SynthExport下載點雲資料。利用meshlab編修點雲，清除非規標點的雜點，保留規標範圍點雲（如圖12）。
2. 將規標範圍內點雲計算重心，求出規標點位。
3. 匯出各點位坐標成XYZ檔案格式。將局部參考坐標以20個控制點利用七參數轉換為與全測站相同之大地坐標系統。
4. 以40個檢核點與全測站測得的坐標比較計算RMSE。



圖12 三維研究區規標點雲編修後示意圖

### （七）光達處理程序

Leica Scanstation C10測站位置大約距離實驗區牆面約30公尺，與最高之規標點仰角約30~40度間。低仰角與高仰角的光達點雲密度光達三維精度分析處理程序如下：

1. 利用Leica Scanstation C10對研究區掃描。（如圖13）
2. 確定規標點數量，給予編號。
3. 利用Scan Control軟體辨識規標，自動擷取規標中心。（如圖14）
4. 匯出各點位坐標成XYZ檔案格式。將局部參考坐標以20個控制點利用七參數轉換為與全測站相同之大地坐標系統。
5. 以40個檢核點與全測站測得的坐標比較計算RMSE。



圖13 光達掃瞄主站與規標點雲圖

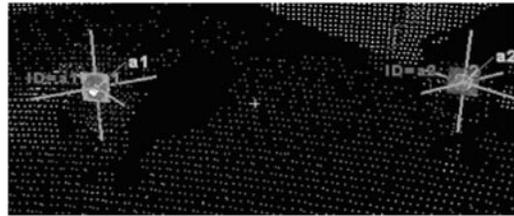


圖14 自動擷取規標中心

#### 四、實驗成果分析

在定位精度評估的設計上，以佈設的60個規標的20個點作為控制點，其餘40個為檢核點，以光達、近景攝影以及Photosynth的方法分別萃取出這60個點位後，分別與全測站測得的點位坐標比較，計算 RMSE (Root Mean Square Error) 做為定位精度的指標。

精度分析主要以圖9中點號為1, 6, 9, 13, 15, 18, 19, 22, 25, 27, 34, 36, 39, 42, 45, 48, 50, 56, 57, 58共20點作為控制點，作為坐標轉換的控制，挑選原則是該20點皆出現於三種方法的點雲資料中，且均勻分佈在研究區範圍。但由於光達與Photosynth在施測與計算過程中，各有少部分點位受到遮蔽，或是無法產生規標位置點雲等情況，為避免轉換時依據不同，故挑選此20點，其餘點位則為檢核點。接著分別對三種方法進行七參數坐標轉換至全測站之大地坐標系統，再以全測站所測點位為實際坐標，計算各方法RMSE，作為精度的衡量依據。

考量不同方法的定位精度時，Lidar、PhotoModeler與Photosynth各組控制點與檢核點精度分析匯整於表2、3。由表3的檢核點精度可以看出，Lidar三維定位精度不管在X, Y, Z任一軸，RMSE都低於其他兩種方法，達到公釐等級的精度水準，整體RMSE為 $\pm 0.007\text{m}$ 。另外兩種方法，在精度上受限拍攝距離、交會角度與建物高度，部分樓層須以高仰角方式才能完整拍攝，故精度都僅在公分等級；在PhotoModeler部分，整體RMSE仍普遍小於Photosynth，約在 $\pm 0.069\text{m}$ 左右。

在Photosynth部分，若考量使用的拍攝像片數對精度的影響，改變相片數後進行點雲產稱與定位分析，成果如圖15、16所示。對本研究區拍攝影像數量介於300至500張時，檢核點精度較為穩定，整體RMSE約維持在 $\pm 0.119\text{m}$ 至 $\pm 0.187\text{m}$ 。過多的影像則易造成太多點雲且易含匹配錯誤，過少的像片則可能無法產生足夠的特徵點，由圖中可知，大於500張或小於300張時，整體RMSE最大達 $\pm 0.604\text{m}$ 左右，差距相當大。另考慮拍攝仰角的因素與定位精度的關係時，當拍攝高度及仰角越低

時，精度會有所提升，原因在於從地面高仰角拍攝時，僅少數幾張從上方往下拍攝的角度約制，故交會角度不佳，在拍攝低仰角時，影響會較小。研究結果顯示出在四層樓以下精度可以提升至 $\pm 0.072\text{m}$ 左右，可見高樓層部分點位誤差較大，圖17所示為拍攝相片456張時的攝影站位置與牆面點雲間的關係圖。

表2 控制點精度 (單位：公尺)

	點數	RMSE X	RMSE Y	RMSE Z	Total RMSE
Lidar	20	$\pm 0.004$	$\pm 0.002$	$\pm 0.003$	$\pm 0.005$
PhotoModeler (12張)	20	$\pm 0.044$	$\pm 0.059$	$\pm 0.025$	$\pm 0.065$
Synth456張	20	$\pm 0.072$	$\pm 0.083$	$\pm 0.067$	$\pm 0.129$
1至頂樓 (456張)	18	$\pm 0.044$	$\pm 0.085$	$\pm 0.062$	$\pm 0.114$
1至8樓 (456張)	15	$\pm 0.030$	$\pm 0.071$	$\pm 0.026$	$\pm 0.080$
1至4樓 (456張)	8	$\pm 0.027$	$\pm 0.039$	$\pm 0.015$	$\pm 0.049$

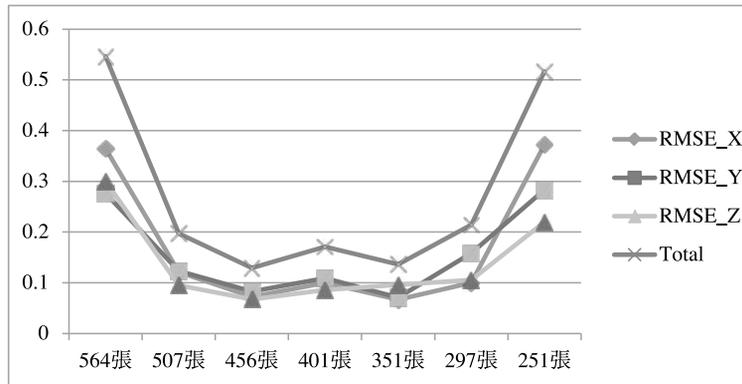


圖15 拍攝影像數與精度關係 (比較20個控制點，單位：公尺)

表3 檢核點精度 (單位：公尺)

	點數	RMSE X	RMSE Y	RMSE Z	Total RMSE
Lidar	37	$\pm 0.005$	$\pm 0.003$	$\pm 0.004$	$\pm 0.007$
PhotoModeler (12張)	40	$\pm 0.038$	$\pm 0.052$	$\pm 0.024$	$\pm 0.069$
Synth456張	39	$\pm 0.059$	$\pm 0.082$	$\pm 0.089$	$\pm 0.135$
1至頂樓 (456張)	38	$\pm 0.058$	$\pm 0.078$	$\pm 0.091$	$\pm 0.133$
1至8樓 (456張)	33	$\pm 0.034$	$\pm 0.072$	$\pm 0.028$	$\pm 0.084$
1至4樓 (456張)	15	$\pm 0.028$	$\pm 0.065$	$\pm 0.012$	$\pm 0.072$

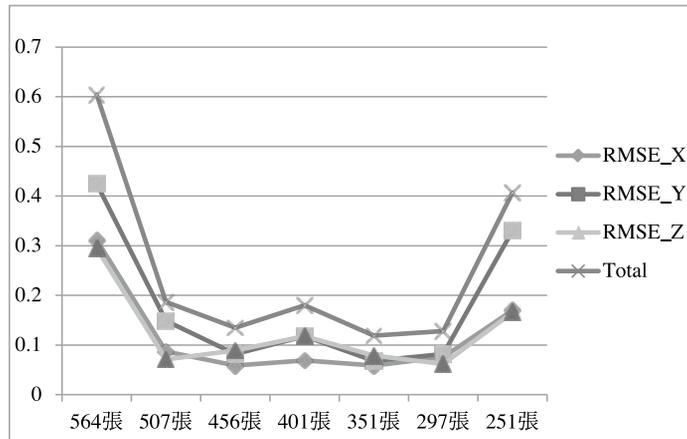


圖16 拍攝影像數與精度關係（比較38~40個檢核點，單位：公尺）

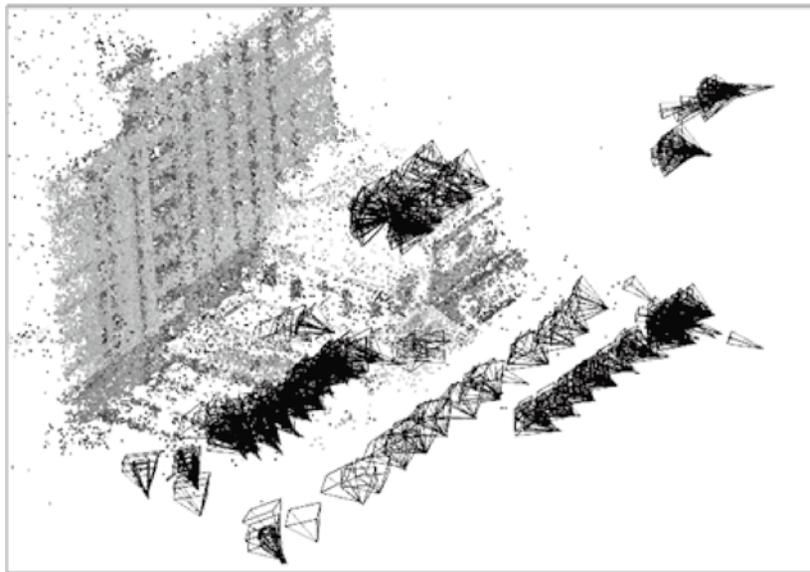


圖17 拍攝相片456張時的攝影站位置與牆面點雲間的關係圖

## 五、結論與建議

本文藉由近景攝影測量、光達及Photosynth的理論研究及實證分析，求得各方法對於定位上所能達到不同的水準。Photosynth是近年來才開始發展的影像處理技術，透過拍攝多張影像，自動化的建構出物體三維空間，並能夠直接在網頁上瀏覽，產生點雲資料，而其所使用的SIFT演算法與SfM持續不斷的改良，藉以克服影

像上不同的變化，以提供Photosynth更好的匹配成果。

本文在實證研究方面，主要是要比較近景攝影測量、光達及Photosynth之定位精度，並了解三種方法在模型建構的程序與成果。基於上述目標，本研究提出以下結論：

1. 在三維定位精度分析方面，以臺北大學公共事務大樓東北側作為本研究範圍，實驗結果顯示光達的控制點精度略高於檢核點精度。Y方向由於為雷射測距方向，故景深精度其較X、Z定位精度好，X與Z方向的檢核點RMSE約為 $\pm 0.005\text{m}$ 與 $\pm 0.004\text{m}$ ，Y方向景深RMSE約為 $\pm 0.003\text{m}$ ，而控制點整體RMSE約在 $\pm 0.005\text{m}$ 左右檢核點的整體RMSE為 $\pm 0.007\text{m}$ ，具有公釐水準等級的量測精度。在近景攝影測量方面，受制於拍攝仰角與高度，PhotoModeler的X與Z方向定位精度大約在 $\pm 0.038\text{m}$ 與 $\pm 0.024\text{m}$ ，而Y方向景深精度受至於拍攝交會角度、距離，故精度較差，約為 $\pm 0.052\text{m}$ ，整體控制點精度在 $\pm 0.065\text{m}$ 左右，檢核點精度在 $\pm 0.069\text{m}$ 左右，皆較光達的成果精度低許多。
2. 在Photosynth方面，依影像數不同分析三維定位之精度可以發現，拍攝實驗中約 $60\text{m} \times 35\text{m}$ 的牆面，拍攝位置距離牆面約 $10\text{m}$ 至 $30\text{m}$ 距離時，拍攝像片數在300至500張中可獲得較穩定的定位精度，以實驗中較佳的整體定位精度為拍攝351張者，其X、Y與Z方向精度最好可以分別為 $\pm 0.059\text{m}$ 、 $\pm 0.068\text{m}$ 與 $\pm 0.079\text{m}$ ，而Y方向為景深方向，若拍攝交會角度太銳或太鈍時將影響此方向的精度，因此大部分的情況下Y方向精度皆低於X與Z方向，整體精度部分為 $\pm 0.119\text{m}$ 左右。而低於300張或高於500張，的情形，若以564張為例，其整體RMSE增加到 $\pm 0.604\text{cm}$ ，顯示出影像數過多與過少，精度都會較差的情況。而依拍攝仰角不同的精度分析可以發現，仰角越低，其精度越高，僅計算1至4樓的定位精度，在控制點整體RMSE上可以提升為 $\pm 0.049\text{m}$ ，而檢核點可達到 $\pm 0.072\text{m}$ ，由此可知若拍攝仰角較低時，Photosynth的定位精度會較佳。
3. 由於近景攝影測量之PhotoModeler軟體中的所有空間重建的點位，皆須藉由人力量取所有像片中的同名點，當像片數增加時所費人力將驟增，相較於本文所提之Photosynth的方法，增加像片數只增加計算機的運算時間，重建空間的點雲數量在一定條件下也會因拍攝像片數量增加而更豐富，且不但可使用不同焦距像機拍攝影像，對於拍攝旋角、亮度變化以及尺度變化也較不敏感，因此在自動化以及可塑性皆具優勢。

## 參考文獻

- 王正忠，2002，以近景攝影測量進行模型式建物重建，國立成功大學航空測量研究所碩士論文。
- 方偉凱、黃灝雄，2005，地面光達資料建立立體模型的精度研究，第 24 屆測量學術及應用研討會論文集。
- 張桓、蔡富安，2009，以單張影像重建三維建物模型，第 28 屆測量及空間資訊研討會論文集。
- 張智安、陳良健，2006，利用光達資料模塑建物之研究，航測及遙測學刊，第 11 卷，第 2 期，頁 175-186。
- 張毅雄，2010，地籍建物資訊模型建立之研究，國立台北大學不動產與城鄉環境學系碩士論文。
- 曾義星、林見福、蔡漢龍、陳鶴欽、曾耀賢，2008，地面光達系統誤差分析及校正，地籍測量：中華民國地籍測量學會會刊，第 27 卷，第 1 期，頁 39-50。
- 黃漢哲，2009，SIFT 演算法應用於航測影像拼接之研究，國立中山大學海洋環境及工程學系碩士論文。
- 趙輝，2006，SIFT 特徵匹配技術講義，山東大學信息學院。
- 饒見有、張智安、陳良健、蔡富安、蕭國鑫、徐偉城，2005，建構像真城市模型之研究，台灣地理資訊學會年會暨學術研討會。
- Dowling, T., I., Read, A., M. and Gallant, J., C., 2009, Very high resolution DEM acquisition at low cost using a digital camera and free software, The 18th World IMACS Congress and MODSIM09 International Congress on Modeling and Simulation, Cairns, Australia.
- Fischler, M., A. and Bolles, R., C., 1981, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM, 24(6), pp. 381-395.
- Koenderink, J., J., 1984, The structure of images, Biological Cybernetics, 50, pp.363-396.
- Lindeberg, T., 1994, Scale-space theory: A basic tool for analysing structures at different scales, Journal of Applied Statistics, 21(2), pp. 224-270.
- Lowe, D. G., 1999, Object recognition from local scale-invariant features, International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157.

- Lowe, D. G., 2004, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2), pp. 91-110.
- Pomaska, G., 2009, Utilization of Photosynth Point Clouds for 3D Object Reconstruction, 22nd CIPA Symposium, Kyoto, Japan.
- Snavely, N., Steven M. Seitz, Richard Szeliski, 2007, Modeling the World from Internet Photo Collections, *International Journal of Computer Vision*, DOI 10.1007/s11263-007-0107-3.
- Tomasi, C. and Kanade, T., 1991, Detection and Tracking of Point Features, Carnegie Mellon University Technical Report CMU-CS-91-132.
- Zhou, K., 2010, Structure & Motion, *Structure in Pattern Recognition*, Vienna University of Technology, Faculty of Informatics, Institute of Computer Graphics and Algorithms, Pattern Recognition and Image Processing Group.

