

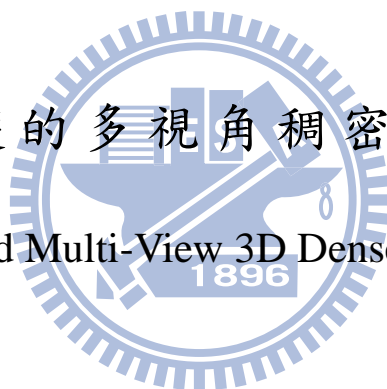
國立交通大學

多媒體工程研究所

碩士論文

以平面為基礎的多視角稠密式三維模型重建

Plane-Based Multi-View 3D Dense Reconstruction



研究生：張嘉峻

指導教授：陳 稔 教授

中華民國 一〇〇 年 七 月

以平面為基礎的多視角稠密式三維模型重建

Plane-Based Multi-View 3D Dense Reconstruction

研究生：張嘉峻

Student : Chia-Chun Chang

指導教授：陳 稔

Advisor : Zen Chen

國立交通大學
多媒體工程研究所
碩士論文



A Thesis
Submitted to Institute of Multimedia Engineering
College of Computer Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master
in
Computer Science

July 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇〇年七月

以平面為基礎的多視角稠密式三維模型重建

研究生：張嘉峻

指導教授：陳 稔

國立交通大學

多媒體工程研究所

摘要

本論文透過多台已校正的相機來進行場景或物體的三維稠密式重建，現今有許多方法是運用固定大小的 window，從多張影像選兩張影像的 window 做 correlation，來看它們之間的 photo consistency。本論文跟這些方法不一樣的是，不再是單純兩兩之間的對應，而是多視角對應，且不使用固定大小的 window。重建的步驟是從擷取影像中特徵點開始，方法是用 local min 或 local max 的 intensity，並使用 epipolar line 做空間幾何的限制，來得到特徵對應點。在三維空間中，整個物體的表面是由許多平面 patch 所逼近的，簡而言之，即物體模型是由一堆有向點或點雲所組成的。初始的平面方程式是由特徵對應點或線所計算出來的，接著使用 iterative optimization 使平面參數的正確性更高，最後完成三維模型。我們運用這重建方法來建合成物來驗證本方法的幾何正確性，並建實拍物來驗證本方法的強韌性。實驗結果比起上面敘述的多視角三維重建方法，看起來更平順，而且更加正確。

Plane-Based Multi-View 3D Dense Reconstruction

Student : Chia-Chun Chang

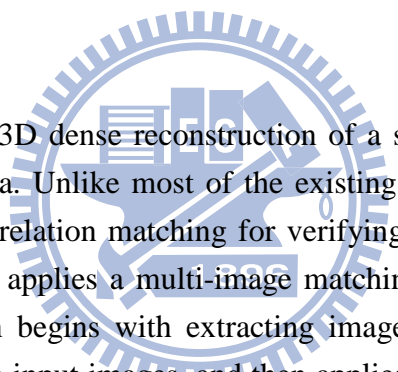
Advisor : Zen Chen

Institute of Multimedia Engineering

College of Computer Science

National Chiao Tung University

Abstract



This paper addresses 3D dense reconstruction of a scene/object from multiple photos taken by a calibrated camera. Unlike most of the existing methods which use a fixed-sized window to do two-view correlation matching for verifying the photo consistency across the multiple views, our method applies a multi-image matching without requiring a fixed-sized window. The reconstruction begins with extracting image features such as local min/max intensity or the like from the input images, and then applies the epipolar constraints to derive the matched feature pairs. The whole object surface is approximated by a set of planar surface patches in 3D space; that is, the object model consists of a set of oriented points or a point cloud in short. Each planar patch is initially represented by a plane equation derived from the associated line/point feature pairs. Next, an iterative optimization process is called for to refine the plane parameter to attain higher accuracy. Finally, a dense model for 3D model is accomplished. We apply our reconstruction method to synthetic images in order to check the geometry correctness and real images for robustness testing. The experimental results show that our reconstructed models generally look smoother and more accurate than the ones reconstructed by a top existing multi-view stereo method.

誌謝

本篇論文的完成，首先要感謝我的指導教授陳稔老師，在我的碩士研究生涯中提供了知識上以及生活上的許多幫助與教誨，老師的研究態度與精神是我學習與敬佩的榜樣。感謝口試委員周宏隆先生、孫樹國學長與賈叢林教授，在論文撰寫方面給予了很多指教，讓論文能順利完成。

接著要感謝博士班的彥旭學長與文昭學長，彥旭學長使我對電腦視覺的領域相當更加瞭解，進而做跟電腦視覺有關的論文，文昭學長常常在我做論文陷入瓶頸時，幫助我度過難關。

感謝實驗室的成員，文軒、岳樺、啟銘學長、佳瑄學姐、秉一、蔡銘、文威學弟，以及研究助理們，一同修課、討論、生活，從你們身上給了我很多勇氣讓我能夠堅持下去。其中感謝岳樺學長指導我會 Matlab，秉一學弟讓我對程式語言變得更加了解，他的鬼點子常能解決我論文的問題，蔡銘學弟讓我了解什麼是做研究的方法及態度，使我漸漸懂得怎麼做研究，而且他還花時間幫我測試程式，文威學弟時常幫我看論文的語句是否通順，讓我減少論文描述的問題。

由於有這些學長姐與學弟的存在，我的研究生活才能夠多彩多姿，僅將這篇論文獻給你們。



張嘉峻 謹誌於
中華民國 100 年 7 月

目錄

| | |
|--|-----|
| 摘要 | i |
| Abstract..... | ii |
| 誌謝 | iii |
| 目錄 | iv |
| 表目錄 | v |
| 圖目錄 | vi |
| 第一章 緒論 | 1 |
| 1.1 研究動機與目標 | 1 |
| 1.2 相關研究 | 1 |
| 1.3 研究方法介紹 | 3 |
| 1.4 論文架構 | 4 |
| 第二章 特徵偵測及對應 | 5 |
| 2.1 特徵點線偵測及對應 | 5 |
| 2.1.1 Textured images 特徵點偵測 | 5 |
| 2.1.2 Structured images 特徵線偵測 | 7 |
| 2.2 運用對應點或線產生世界座標的三維點或線 | 7 |
| 2.3 對 reference 影像畫格子 | 8 |
| 第三章 Iterative optimization 方法描述 | 10 |
| 3.1 Fitness Function | 10 |
| 3.2 求出初始平面參數 | 10 |
| 3.3 Homography 的公式推導 | 13 |
| 3.4 Iterative plane parameter optimization | 15 |
| 第四章 稠密式三維重建 | 19 |
| 4.1 Homography 與 Pixel-wise 標準差 | 19 |
| 4.2 用 MSER 找出標準差小的區域並刪除區域內的特徵點 | 22 |
| 4.3 求出世界座標上的 3D 點 | 25 |
| 4.4 運用空間切割建出三維模型 | 26 |
| 第五章 實驗結果 | 27 |
| 5.1 具豐富紋理平面物體三維重建可行性驗證 | 27 |

| | |
|---|----|
| 5.2 稀疏紋理平面物體三維重建可行性驗證 | 29 |
| 5.3 The 3D reconstruction of Wadham College | 34 |
| 第六章 結論與未來發展 | 41 |
| 參考文獻 | 42 |

表目錄

| | |
|---------------------------------------|----|
| 表 1 iterative optimization 的數據 | 17 |
| 表 2 以 View4 為 reference 影像的投影誤差 | 29 |



圖目錄

| | |
|---|----|
| 圖 1 論文方法流程..... | 4 |
| 圖 2 Input：待建物之影像(concave box)..... | 5 |
| 圖 3 待建物影像以及個別的特徵點..... | 6 |
| 圖 4 三張影像皆有的特徵對應點..... | 6 |
| 圖 5 Canny edge segment map..... | 7 |
| 圖 6 對 reference image 畫格子及每個格子中的特徵點數..... | 9 |
| 圖 7 在影像中格子內取的 3 點以及這 3 點所形成的 bounding window..... | 12 |
| 圖 8 對收斂平面做 homography 的結果..... | 21 |
| 圖 9 對 homography 後的影像做標準差的結果..... | 22 |
| 圖 10 對標準差影像做 MSER 的結果..... | 23 |
| 圖 11 把屬於 MSER 區域內的點標記..... | 24 |
| 圖 12 圖 10(c)做平面切割的結果..... | 26 |
| 圖 13 Furukawa 和我們方法建 Convex cube 之比較..... | 28 |
| 圖 14 重建結果檢驗-View4..... | 29 |
| 圖 15 重建結果檢驗-View2..... | 29 |
| 圖 16 Input image：待建物 mailbox 的影像..... | 30 |
| 圖 17 mailbox 做標準差的結果..... | 31 |
| 圖 18 將標準差用 MSER 抓出來的結果(mailbox)..... | 32 |
| 圖 19 重建三維點雲資料與三維網格模型結果，並與 Furukawa 的重建結果進行比較..... | 33 |
| 圖 20 待建物 wc(Wadham College)的影像..... | 35 |
| 圖 21 透過 Canny edge 演算法計算出 edge 相關資訊..... | 36 |
| 圖 22 標準差以及其擷取平面..... | 38 |
| 圖 23 重建的三維點雲資料與三維網格模型結果，並與 Furukawa 方法重建結果進行比較..... | 39 |
| 圖 24 用 meshlab 所呈現建築物的不同面貌，包括俯瞰圖的兩片牆夾角約為 90 度等..... | 40 |

第一章 緒論

1.1 研究動機與目標

由於這五年來利用 multi-view 影像建立 3D 風行，尤其是 Goesele[1]以 Middlebury dataset[2]進行 benchmark 之後，更是引起許多研究學者的討論風氣，因此現在 multi-view stereo 很熱門，也因熱門的關係，讓我有做這方面的動機。

本論文在於利用多張已校正影像，透過特徵擷取、對應的過程，得到多視圖影像之間的特徵對應關係。透過對一個已校正內部參數的相機所拍攝的影像，進行外部參數的校正，並進行場景或物體的稠密式三維重建。

在進行稠密式三維重建時，需要使用到影像一致性(photo consistency)測試，以得到非前述特徵點的影像區域之空間位置，然而傳統在計算影像一致性時，皆採用兩兩影像之間的比對計算，最後再一併考慮各影像兩兩之間的計算結果，以得到該影像點最後的空間估測位置，本研究提出一個基於多視圖影像資訊的重建方法，而非僅合併許多兩兩影像所重建的部份結果而已，此方法將能增加三維重建的準確性。

另一方面，本研究以小塊平面區塊為出發，並逐漸擴張到原本就有的整個平面，可快速又精準得到屬於共平面的物體表面，而不像 Furukawa 等人[3]以固定的 window size (非如 21 x 21 pixels)的 image patch 當重建單位，一旦該 window 未對應單一平面，則無法找到對應的對象，容易變得支離破碎。屬於平面較多的物體時，此方法將能有效且正確的重建出物體的表面，尤其當該平面是具有稀疏紋理(texture-less)區域時，傳統以固定的 window size 在影像中重建的方式將不易重建出此類區域，這將在實驗中得到印證。本方法以疊代式地對物體做細部的稠密式重建之後，再貼上材質貼圖即能得到物體的高擬真三維模型。

1.2 相關研究

Multi-view stereo 的三維重建方法大致可分為 calibrated 及 uncalibrated 兩種大類型，前者較多針對室內靜態物體進行三維重建，後者則常為針對戶外大型場景多視圖影像進行重建。

為了要判斷重建出來的 3D 幾何是否正確，許多學者提出了不同的演算法，但在絕大部份的 scene-space 方法中皆是利用不同影像之間對應區域的關聯程度(correlation)來判斷重建結果的優劣，而這種方法即是 multi-view 相關技術的主要核心技術，通常被稱為顏色一致性(photo-consistency)的量測[4, 5, 6]。如果各張影像對於場景中物體表面上的某一點觀察到的 intensity 相同，就說這個點是 photo-consistent 的。

使用 photo-consistency 能降低尋找影像之間點對應的難度，Dyer[8]與 Szeliski[9]皆提到使用 photo-consistency 有下列的優點：

- (1) 在某些情況下要取得正確的點對應是非常困難的，尤其是在比對區域內的 intensity 值幾乎一模一樣時。然而 photo-consistency 只使用到正向投影(forward projection)以及比對 pixel 的步驟。因此在給定一個 3D 場景模型時，計算 photo-consistency 時只需判斷在各影像上的投影結果是否一致，因此不需要個別點對應關係。
- (2) 因為要取得「稠密式」的點對應是非常困難的，所以若是使用點對應的演算法來重建場景 3D 模型，就必須捨棄影像中點對應正確性較低的點，而建成較稀疏(sparse)的 3D 資訊，否則點對應錯誤的部份會導致建置出較差品質的 3D 模型。

Hernandez[10]先計算每張影像的 depth map，再將其結果進行合併成一個 cost volume。之後再從初始的 visual hull 開始對 mesh 進行疊代變形(deform)，以此找到最接近物體的表面，因此可以定義一個 optimal surface 來對 global energy ε 做最小化：

$$\varepsilon(S) = \varepsilon_{tex}(S) + \varepsilon_{sil}(S) + \varepsilon_{int}(S)$$

S 是 \mathbb{R}^3 上的一 surface，要找出一 S 可以將 energy $\varepsilon(S)$ 最小化，式中的 $\varepsilon_{tex}(S)$ 是物體本身 texture 的 energy term， $\varepsilon_{sil}(S)$ 是 silhouette 的 energy term， $\varepsilon_{int}(S)$ 是表面模型(surface model)的 regularization term。目前國際上的 multi-view 相關技術都有用到 texture 資訊，目的是利用各個不同視角影像中的 texture 資訊 back project 至空間中來確立該特徵的 3D 位置。然而，雖然有 multi-view images，但大部份都是兩兩利用 NCC 或 SSD 來比較後，再最後計算後做為 multi-view stereo 結果，其實並沒有真正一次同時考慮到所有的 view。而 Hernandez[10]則是少數同時考慮到所有影像的 texture & silhouette 資訊，列成一個 optimization equation，再進行最佳化。

接下來將針對 calibrated 的情況進行文獻探討，Furukawa[3]提出了一種方式來重建

稠密式的 3D 物體模型，此方法不需事先知道物體的 bounding volume，其主要步驟為：

Step 1: Matching: 利用 Harris 或是 Difference-of-Gaussians 來擷取影像上的特徵點，然後在各個角度的影像中計算特徵點對應關係，再以此做為 sparse 的對應區塊。接著再一直重覆下列兩步驟，直至建立起完整的 3D 模型。

Step 2: Expansion: 從前一步驟所計算出來的對應區塊開始由鄰近的 pixel 往外擴張，以得到較 dense 的對應區塊。

Step 3: Filtering: 利用 visibility 的限制來去除不正確的對應區塊。

Habbecke[11, 12]也是利用類似的方式來重建物體的三維模型，首先[11]提出了疊代式的平面法向量最佳化的演算法，透過事先指定的平面區域，再透過多視圖影像進行平面法向量的最佳化。之後 Habbecke 則在[12]中提出以 seed disk 逐漸往外擴張的方式重建，步驟如下：

Step 1: 利用 homography 對應關係，在兩張影像中計算 seed disks 當做初始對應區域，並利用 plane fitting algorithm，綜合其他影像兩兩計算 homography 對應關係的結果來做修正。

Step 2: 利用上一步驟找出的 seed disks 以 greedy growing 的方式來擴展物體表面的區塊。一直擴展直至物體於所有影像中可見的部份都被 disks 覆蓋為止。

在拍攝戶外場景的建築物重建三維模型時，除了特徵點外的區域若要進行三維空間的估測，必須利用前述的 photo consistency 來進行比對。然而，建築物本身經常是缺乏紋理資訊(texture-less)的，在此情況下使用 photo consistency 會導致比對分歧(ambiguity)的發生。

另外，Furukawa[18]則假設建築物都是由三個不同互相垂直方向的平面所構成，也就是所謂的 Manhattan World。首先利用 Furukawa[]重建具紋理物的方法去重建較具方向的點(oriented points)，利用這些點估算出主要座標軸，再由前述已重建出具紋理的點來計算延著各軸的 density peak，最後透過 MRF 及 graph cuts 來指定三個可能的平面給予每一個影像中的點，以完成各視角影像的深度圖。

1.3 研究方法介紹

本論文乃透過已知多視角相機內、外部參數及影像建構出三維模型。下圖是本論文建構三維模型的流程：



圖 1 論文方法流程

1.4 論文架構

本論文第二章首先簡單說明特徵偵測與對應，接著於第三章介紹 Iterative plane parameter optimization 的流程與概念，並於第四章介紹稠密式三維重建，再來於第五章呈現最後的建模結果及進行一些模型的正確性驗證，最後第六、七章式結論以及未來發展。

第二章 特徵偵測及對應

2.1 特徵點線偵測及對應

2.1.1 Textured images 特徵點偵測

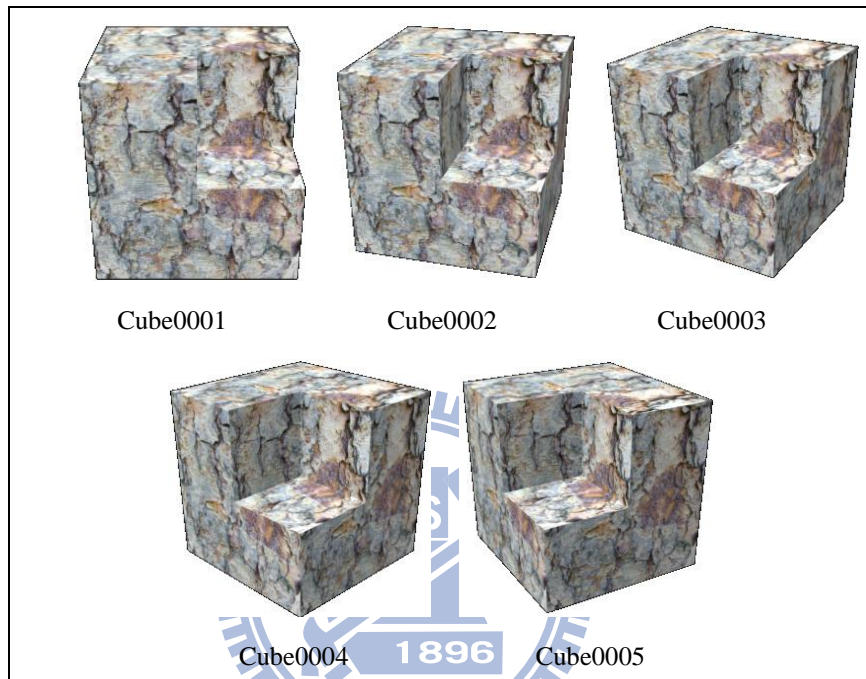
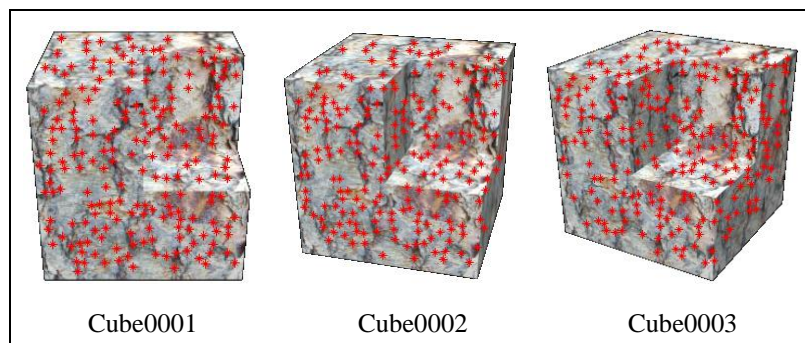


圖 2 Input：待建物之影像(concave box)

首先找出每張影像上的特徵點 (有紋路(texture)或有結構性(structure)的 point) ，本論文使用 local intensity min/max 來偵測特徵點。



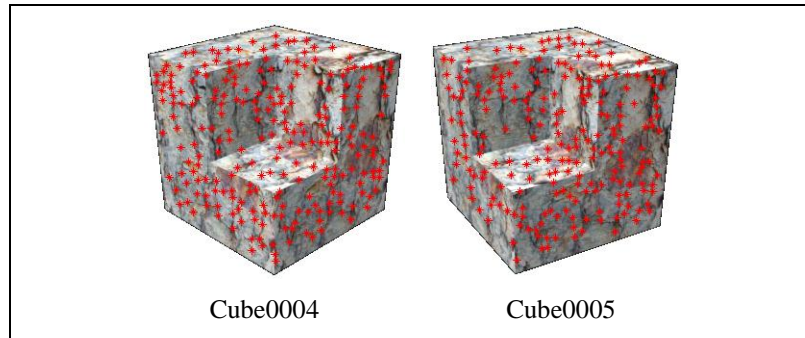


圖 3 待建物影像以及個別的特徵點

圖 3 中特徵點的擷取方式是以一點為中心的 11×11 的範圍中，若在此範圍中其他的點的 intensity 皆比中心點的 intensity 大，則此中心點即為特徵點。

對應的方式是找離 epipolar line 較近的特徵點當特徵對應點，如果只用 2 張影像，限制會不夠，因為 1 張影像上只有另 1 張影像畫過來的 epipolar line，這樣只能單純找到離 epipolar line 近的點，但這些點很多。若使用 3 張影像，1 張影像上會有其他 2 張影像畫過來的 epipolar lines，而這 2 條線通常會交於 1 點，這樣的話，只要找離此交點近的點當特徵對應點就好了。

實際做法為將影像 i 的某特徵點 p_i 在影像 j, k 上各畫一條 epipolar line，分別為 l_{ij} 與 l_{ik} ，並在影像 j, k 中分別找出距離與 epipolar line 小於某門檻值的那些特徵點，並假設在影像 j 上有 n_j 個點離 l_{ij} 很近，影像 k 上有 n_k 個點離 l_{ik} 很近。

將上述在影像 j 上的 n_j 點取一點與在影像 k 上的 n_k 點取一點，各畫一條 epipolar line 到影像 i 上，分別為 l_{ji} 與 l_{ki} ，並算出 l_{ji} 與 l_{ki} 的交點，由於 l_{ji} 有 n_j 種情況， l_{ki} 有 n_k 種情況，所以會得到 $n_j n_k$ 個交點，這些交點與 p_i 的距離小於某個門檻值(因為最小值不一定是正確答案)，就可以找到在影像 j, k 上對應 p_i 的候選特徵點 p_j 及 p_k 。但往往不只一種情況。

為了解決這問題，將候選的 p_j 畫一條 epipolar line 到影像 k 上，候選的 p_k 畫一條 epipolar line 到影像 j 上，分別為 l_{jk} 與 l_{kj} ，此時三張影像中，都離 epipolar line 交點最近的點即為要找的特徵對應點。

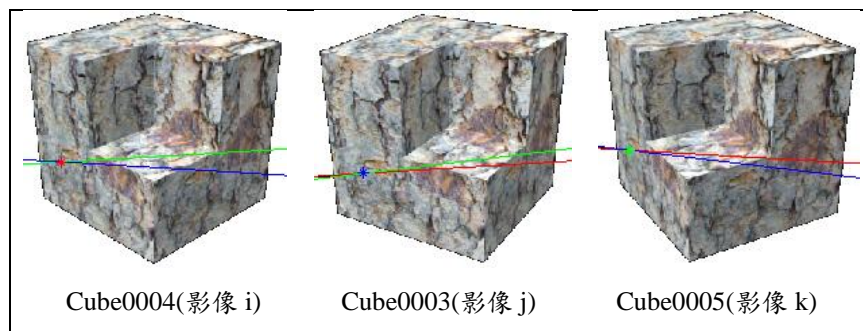


圖 4 三張影像皆有的特徵對應點

圖 4 中紅點為 p_i 、藍點為 p_j 、綠點為 p_k ，紅點在其他影像畫的 epipolar line 以紅線表示，藍點在其他影像畫的 epipolar line 以藍線表示，綠點在其他影像畫的 epipolar line 以綠線表示。若紅點與藍線和綠線的交點、藍點與紅線和綠線的交點與綠點與紅線和藍線的交點皆很近的話，則此 3 點視為特徵對應點。

2.1.2 Structured images 特徵線偵測

至於 Structured images 特徵的偵測則是使用 canny edge detection。

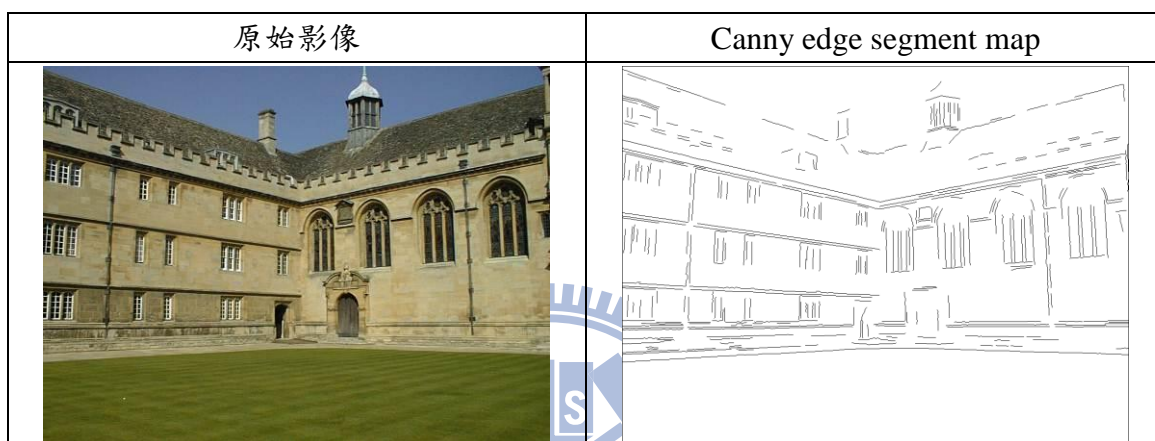


圖 5 Canny edge segment map

而特徵線對應的方式，是參考[19]。

2.2 運用對應點或線產生世界座標的三維點或線

只要找出特徵對應點，就能得到這些對應點的三維座標。

由於相機已校正，所以投影矩陣已知，假設有 m 張影像，每張影像分別對應的 3×4 投影矩陣分別為 M_1, M_2, \dots, M_m ，並設所有特徵對應點可以產生出 n 個 3D 點 P_1, P_2, \dots, P_n 。

設 p_{ij} 為第 j 個 3D 點，投影到在第 i 張影像上的 2D 點，即 $z_{ij}p_{ij} = M_i P_j$ ，所以 $p_{1j}, p_{2j}, \dots, p_{mj}$ 為特徵對應點，其中 $i = 1 \sim m, j = 1 \sim n$ ，可以使用三角定位的公式計算出第 j 組特徵對應點在世界座標下的三維座標。

$$\begin{bmatrix} [p_{1j} \times] M_1 \\ [p_{2j} \times] M_2 \\ \vdots \\ [p_{mj} \times] M_m \end{bmatrix}$$

找到特徵對應線 $l(l = \begin{bmatrix} l_a \\ l_b \\ l_c \end{bmatrix})$ ，代表線方程式 $l_a x + l_b y + l_c = 0$ 後，可以使用計算出這

些對應線在世界座標下線的參數式，而參數式是由線的方向向量及線上的某一點所組成的。

首先，要求出線的方向向量：

$$\begin{bmatrix} l_{1j}^T M_1 \\ l_{2j}^T M_2 \\ \vdots \\ l_{mj}^T M_m \end{bmatrix} = U \Sigma V^T$$

其中 $V = [v_1 \ v_2 \ v_3 \ v_4]$ ，而 $v_i = \begin{bmatrix} v_{1i} \\ v_{2i} \\ v_{3i} \\ v_{4i} \end{bmatrix}$ ， $i = 1 \sim 4$ ， v_1, v_2 表示三維線的兩面式中那兩面的平面參數，也就是指此三維線的兩面式為：

$\begin{cases} v_{11}x + v_{21}y + v_{31}z + v_{41} = 0 \\ v_{12}x + v_{22}y + v_{32}z + v_{42} = 0 \end{cases}$ ，所以這條三維

線的方向向量為這兩個平面法向量的外積，即 $\begin{bmatrix} v_{11} \\ v_{21} \\ v_{31} \end{bmatrix} \times \begin{bmatrix} v_{12} \\ v_{22} \\ v_{32} \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$ 。

接著，要求出線上的某個 3D 點：

$$\begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} & v_{31} & v_{41} \\ v_{12} & v_{22} & v_{32} & v_{42} \end{bmatrix} = U_l \Sigma_l V_l^T$$

假設 V_l 的最後一行為 $\begin{bmatrix} v_{l14} \\ v_{l24} \\ v_{l34} \\ v_{l44} \end{bmatrix}$ ，則此線上的 3D 點為 $\begin{bmatrix} v_{l14} \\ v_{l44} \\ v_{l24} \\ v_{l44} \\ v_{l34} \\ v_{l44} \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ 。

得到線的方向向量 $\begin{bmatrix} u \\ v \\ w \end{bmatrix}$ 與線上的某個 3D 點 $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ ，則此三維線的參數式為 $\begin{cases} x + ut \\ y + vt, t \in R \\ z + wt \end{cases}$ 。

2.3 對 reference 影像畫格子

本論文是以特徵對應點最多的影像，當作 reference 影像。

我們的目標是求出目標平面的三維平面方程式，要算出平面方程式，至少需要 3 個 3D 點。

因此要在目標平面上找出 3 點，但無法得知哪 3 點在目標平面上，因此我們在 reference image 上畫格子，目的是希望一個格子能代表一個平面，但實際上這是不可能的，因為這裡使用的是矩形格子，而欲求的目標平面往往是不規則的多邊形。

此外，格子的大小不好決定，格子不能太大，太大的話，一個格子會包含多個平面；格子的也不能太小，太小的話，一個格子內所含的特徵點會太少。故要找到合適的格子，是很困難的，故本論文在實作中，格子的大小可藉由調參數來改變。

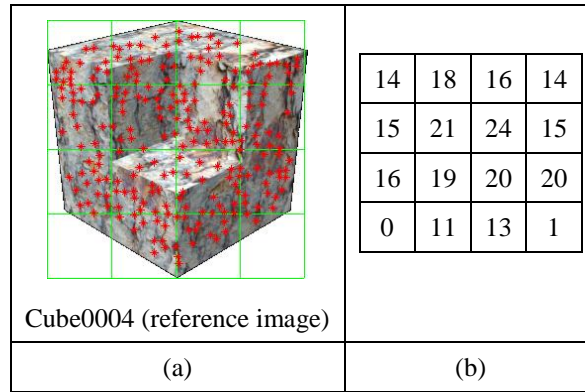
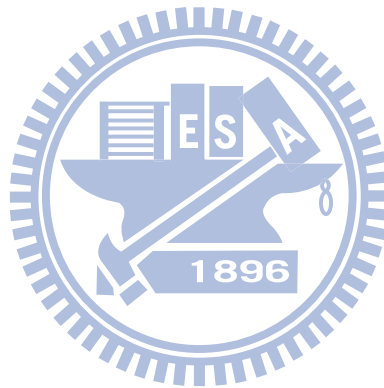


圖 6 對 reference image 畫格子及每個格子中的特徵點數

(a)對 reference image 畫格子。(b)算出每格子的特徵點數，以表格的形式表示，其中紅色的數字代表目前此格子擁有最多特徵點。

畫完格子後，將每個格子個別視為目標平面，先從特徵點數最多的格子開始做，也就是圖 6(b)中紅色的那格，原因是我們的目標只要從格子中取出 3 點，所以從越多點中取 3 點，選到的 3 點共面的機會比較大。



第三章 Iterative optimization 方法描述

本章節是參考 Habbecke[11]運用疊代的方式，來求出最佳化的平面參數之演算法，此演算法為透過指定的平面區域，再透過多視圖影像進行平面參數的最佳化。

3.1 Fitness Function

這裡我們令 fitness function 為

$$E = \sum_{p \in \Omega} (I(p) - J(H(n)p))^2$$

Fitness function 裡每個參數的定義。

1. Ω : 在reference影像上，要找目標平面的區域。但由於怕這個區域中有uniform的部分，所以這裡的 Ω 是指將目標平面的區域做edge detection後有edge的部分。
2. I : reference image。
3. J : comparison image。
4. p : 2D點
5. $I(p)$: 在reference image中， p 這點的intensity。
6. n : 在世界座標下的平面方程式 $ax + by + cz + d = 0$ ，可以用 $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ 表示。
7. $H(n)$: $H(n)$ 是由世界座標下的平面 n 所決定出一個homography matrix。

3.2 求出初始平面參數

本論文的平面參數是指平面方程式： $ax + by + cz + d = 0$ 中的 $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ ，欲求出平面參數，至少需要3個3D點。

假設3個3D點分別為 $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$, $\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$, $\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix}$ ，則 $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} \times \begin{bmatrix} x_3 - x_1 \\ y_3 - y_1 \\ z_3 - z_1 \end{bmatrix}$, $d = -[x_i \ y_i \ z_i] \begin{bmatrix} a \\ b \\ c \end{bmatrix}$,

$i = 1, 2, 3$ 。

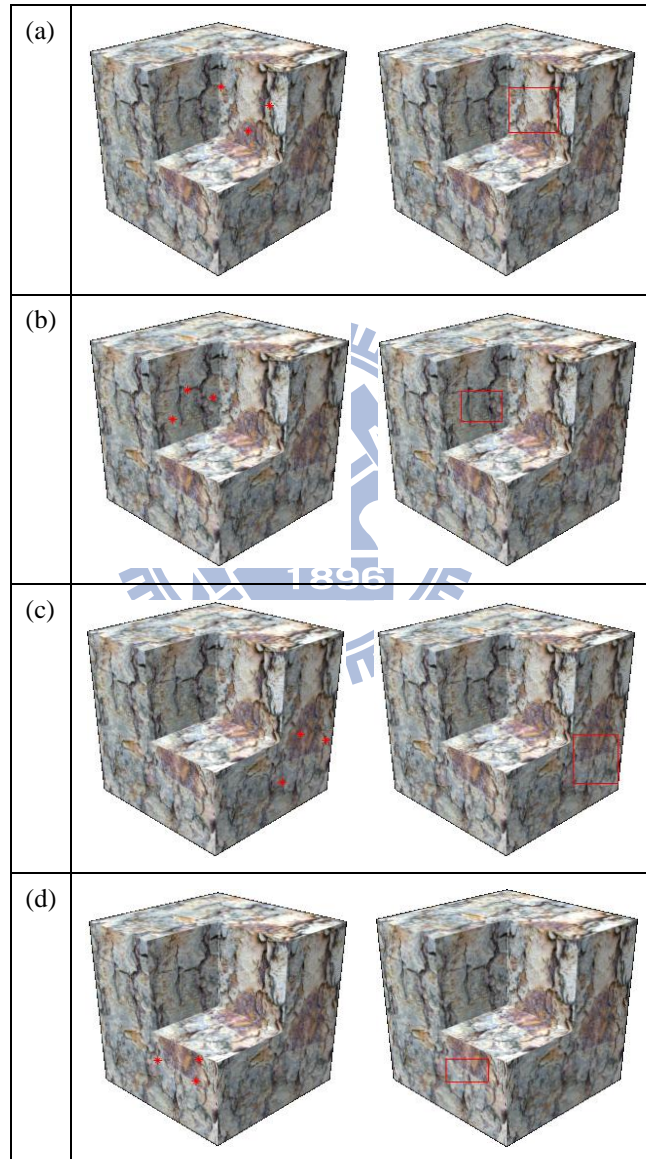
但格子內有那麼多點，要怎麼決定到底哪3點才是最好的？本論文使用 RANSAC 的方式，每次選格子內所產生的3個3D點，求所構成的平面。設此平面為 $a_i x + b_i y +$

$c_i z + d_i = 0$ ，要找使 E_i 最小的 $\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix}$ 。

$$E_i = \frac{1}{|\Omega_i|} \sum_{p \in \Omega_i} (I(p) - J(H(n)p))^2$$

其中 $n = \begin{bmatrix} M_r^{-1} & 0 \\ -M_r^{-1}m_r & 1 \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix}$ ， Ω_i 為影像上這 3 點所形成的 bounding window。

在取 3 點的 bounding window Ω_i 時，最害怕的就是 bounding window 會橫跨多個平面，但是大部分的情況，若發生 3 點的 bounding window 跨平面的情形，會造成初始誤差較大，因此在多點中選出較好的 3 點，較難選到這種情形，故不用太擔心這個問題。



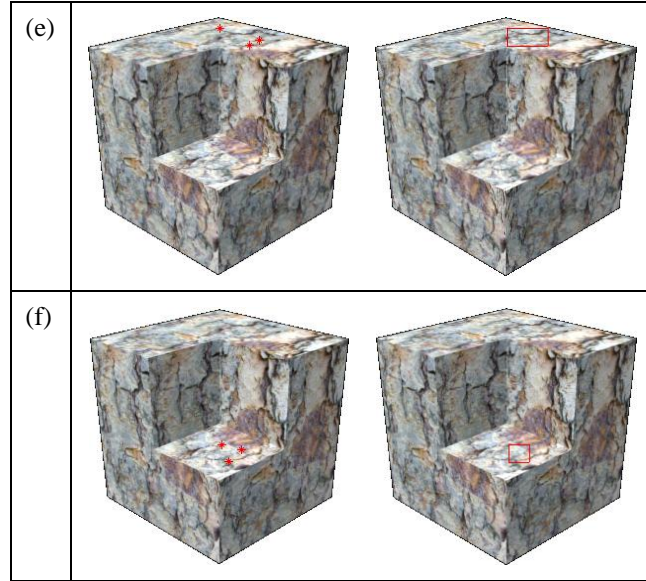


圖 7 在影像中格子內取的 3 點以及這 3 點所形成的 bounding window

- (a) 左：在圖 6(b)中，紅色的那格中選出使 E_i 最小的 3 點(共有 C_3^{24} 種情形)。
右：運用圖 7(a)左的 3 點產生的 bounding window。
- (b) 左：在圖 11(a)右中，紅色的那格中選出使 E_i 最小的 3 點(共有 C_3^{21} 種情形)。
右：運用圖 7(b)左的 3 點產生的 bounding window。
- (c) 左：在圖 11(b)右中，紅色的那 2 格中，各選出使 E_i 最小的 3 點(各 C_3^{20} 種情形)，選出這 2 個誤差最小的 3 點。
右：運用圖 7(c)左的 3 點產生的 bounding window。
- (d) 左：在圖 11(c)右中，紅色的那格中選出使 E_i 最小的 3 點(各 C_3^{19} 種情形)。
右：運用圖 7(d)左的 3 點產生的 bounding window。
- (e) 左：在圖 11(d)右中，紅色的那 2 格中，各選出使 E_i 最小的 3 點(各 C_3^9 種情形)，選出這 2 個誤差最小的 3 點。
右：運用圖 7(e)左的 3 點產生的 bounding window。
- (f) 左：在圖 11(e)右中，紅色的那 2 格中，各選出使 E_i 最小的 3 點(各 C_3^5 種情形)，選出這 2 個誤差最小的 3 點。
右：運用圖 7(f)左的 3 點產生的 bounding window。

最後將使 E_i 最小的 $\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix}$ 當作 Ω_i 的初始平面參數。

線對應的部分：欲求出平面參數，至少需要 2 條 3D 線

假設 2 條 3D 線為 $\begin{cases} x_1 + u_1 t \\ y_1 + v_1 t, t \in R \\ z_1 + w_1 t \end{cases}$ 及 $\begin{cases} x_2 + u_2 t \\ y_2 + v_2 t, t \in R \\ z_2 + w_2 t \end{cases}$ ，照理說 $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} \times \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix}$ ，但

是為了避免 2 條 3D 線發生互相平行的情形，所以會使用 2 條線上各取一點的向量與某

條線的方向向量做外積，即 $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} \times \begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix}$ ， $d = -[x_i \ y_i \ z_i] \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ， $i = 1, 2$ 。

3.3 Homography 的公式推導

設 reference image 的相機參數為 $P_r = [M_r | m_r]$ ，comparison image 的相機參數為 $P_c = [M_c | m_c]$ 。

並設平面方程式為 $n_0x + n_1y + n_2z + d = 0$ ，令 $N = \begin{bmatrix} n_0 \\ n_1 \\ n_2 \\ d \end{bmatrix}$ ， $n = \begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix}$ ，則 $N^T X = 0$ ， $X = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$ 為

平面上的某個 3D 點，其中 n 指的是 plane 的法向量， d 是 plane 與世界座標原點的距離。

$$n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} + d = 0 \Rightarrow d = -n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$z_r p_r = M_r \begin{bmatrix} x \\ y \\ z \end{bmatrix} + m_r \Rightarrow dz_r p_r = dM_r \begin{bmatrix} x \\ y \\ z \end{bmatrix} + m_r d \quad (3.1)$$

將 $d = -n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ 代入(3.1)得：

$$(dz_r) p_r = dM_r \begin{bmatrix} x \\ y \\ z \end{bmatrix} + m_r \left(-n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = (dM_r - m_r n^T) \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

故將 3D 空間中的平面 $n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} + d = 0$ 轉到 reference image 的 homography matrix 為 $(dM_r - m_r n^T)$ 。

同理將 3D 空間中的平面 $n^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} + d = 0$ 轉到 comparison image 的 homography matrix 為 $(dM_c - m_c n^T)$ 。

所以由 reference image 轉到 comparison image 的 homography matrix 為

$$H_c(n) = (dM_c - m_c n^T)(dM_r - m_r n^T)^{-1} \quad (3.2)$$

照理說 homography 的公式到此就推倒結束了，但由於後面我們還需要用此公式來推導 fitness function 微分後的數學式，但(3.2)中的 $(dM_r - m_r n^T)^{-1}$ ，會造成之後的公式不好推導，所以將世界座標會轉換到其他座標系，以簡化(3.2)。

假設 $B = \begin{bmatrix} M_r^{-1} & -M_r^{-1}m_r \\ 0 & 0 & 0 & 1 \end{bmatrix}$ ，則

$$P'_r = P_r B = [M_r | m_r] \begin{bmatrix} M_r^{-1} & -M_r^{-1}m_r \\ 0 & 0 & 0 & 1 \end{bmatrix} = [M_r M_r^{-1} | -M_r M_r^{-1}m_r + m_r] = [\text{Id}_3 | 0]$$

$$P'_c = P_c B = [M_c | m_c] \begin{bmatrix} M_r^{-1} & -M_r^{-1}m_r \\ 0 & 0 & 0 & 1 \end{bmatrix} = [M_c M_r^{-1} | -M_c M_r^{-1}m_r + m_c] = [M'_c | m'_c]$$

$$z_c p_c = P_c X = P_c B B^{-1} X = P_c' X' \Rightarrow X' = B^{-1} X = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix}$$

$$N^T X = N^T B B^{-1} X = (B^T N)^T B^{-1} X = N'^T X' \Rightarrow N' = B^T N = \begin{bmatrix} n'_0 \\ n'_1 \\ n'_2 \\ d' \end{bmatrix}$$

$$N''^T = \frac{N'^T}{d'} = \begin{bmatrix} \frac{n'_0}{d'} \\ \frac{n'_1}{d'} \\ \frac{n'_2}{d'} \\ 1 \end{bmatrix} = \begin{bmatrix} n'_1 \\ 1 \end{bmatrix} \quad (3.3)$$

$$[n'_0 \ n'_1 \ n'_2] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + d' = 0 \Rightarrow d' = -[n'_0 \ n'_1 \ n'_2] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

$$z_r p_r = M'_r \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + m'_r \Rightarrow d' z_r p_r = d' M'_r \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + m'_r = d' \quad (3.4)$$

將 $d' = -[n'_0 \ n'_1 \ n'_2] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$ 代入(3.4)得：

$$(d' z_r) p_r = d' M'_r \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + m'_r \left(-[n'_0 \ n'_1 \ n'_2] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \right) = (d' M'_r - m'_r [n'_0 \ n'_1 \ n'_2]) \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

因為 $M'_r = \text{Id}_3$ ， $m'_r = 0$ ，所以 $(d' z_r) p_r = (d' M'_r - m'_r [n'_0 \ n'_1 \ n'_2]) \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = d' \text{Id}_3 \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$

$$(d' z_c) p_c = d' M'_c \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + m'_c \left(-[n'_0 \ n'_1 \ n'_2] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \right) = (d' M'_c - m'_c [n'_0 \ n'_1 \ n'_2]) \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

故由 reference image 轉到 comparison image 的 homography matrix 為

$$H_c(n') = (d' M'_c - m'_c [n'_0 \ n'_1 \ n'_2]^T) (d' M'_r - m'_r [n'_0 \ n'_1 \ n'_2])^{-1}$$

$$= (d' M'_c - m'_c [n'_0 \ n'_1 \ n'_2]) (d' \text{Id}_3)^{-1} = \frac{1}{d'} (d' M'_c - m'_c [n'_0 \ n'_1 \ n'_2])$$

$$= M'_c - m'_c \frac{[n'_0 \ n'_1 \ n'_2]}{d'} = M'_c - m'_c n' \quad (3.5)$$

(3.2)與(3.5)的差別：

(3.2)是在正常世界座標系下，所得到的 homography 轉換公式。

(3.5)則是將世界座標系透過 $B = \begin{bmatrix} M_r^{-1} & -M_r^{-1}m_r \\ 0 & 0 & 0 & 1 \end{bmatrix}$ ，轉到其他座標系後，所得到的 homography 轉換公式，而(3.5)明顯比(3.2)簡化了許多。

透過 B 轉換後的優點：

1. 原點為 reference 相機鏡心，由於所求平面不會通過鏡心(原點)，即 $d' \neq 0$ ，因此，

新座標的平面參數 $[a' \ b' \ c' \ d']^T$ ，都可簡化為 $\left[\frac{a'}{d'} \ \frac{b'}{d'} \ \frac{c'}{d'} \ 1\right]^T$ ，這樣要估的參數就從 4 個降為 3 個。

2. 由於轉到新座標後， $M'_r = \text{Id}_3$ ， $m'_r = 0$ ，故 homography 轉換公式變得更簡單了。

3.4 Iterative plane parameter optimization

這是使用 fitness function 微分的方式，來找到 fitness function 最小值的方法。

註：這節所提到的 n ，是指公式(3.3)的 n' ，也就是透過 B 轉換得到的法向量。

定義 $n_{k+1} := n_k + \Delta n$ ，則 $E_{k+1} = \sum_{p \in \Omega} (I(p) - J(H(n_{k+1})p))^2 = \sum_{p \in \Omega} (I(p) - J(H(n_k + \Delta n)p))^2$ 。

令

$$J^{(0)} = J$$

$$J^{(k+1)}(p) := J^{(0)}(H(n_k)p)$$

由(3.4)知道 $H(n_k + \Delta n) = M - m(n_k + \Delta n)^T = M - mn_k^T - m\Delta n^T = H(n_k) - m\Delta n^T$

$$\therefore J(H(n_k + \Delta n)p) = J((H(n_k) - m\Delta n^T)p)$$

$$J^{(0)}((H(n_k) - m\Delta n^T)p) = J^{(k+1)}\left((H(n_k))^{-1}(H(n_k) - m\Delta n^T)p\right)$$

$$= J^{(k+1)}\left(\left(\text{Id} - (H(n_k))^{-1}m\Delta n^T\right)p\right)$$

Let $(H(n_k))^{-1}m = \tilde{m}$, then $J^{(k+1)}\left(\left(\text{Id} - (H(n_k))^{-1}m\Delta n^T\right)p\right) = J^{(k+1)}\left(\left(\text{Id} - \right.$

$m\Delta n^T)p$

$$\begin{aligned}
\text{suppose } p &= \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \tilde{m} = \begin{bmatrix} \tilde{m}_0 \\ \tilde{m}_1 \\ \tilde{m}_2 \end{bmatrix}, \therefore J(p) = J \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = J(u, v), \therefore J^{(k+1)}((\text{Id} - \tilde{m}\Delta n^T)p) = J^{(k+1)}(\text{Id}p - \tilde{m}\Delta n^T p) \\
&= J^{(k+1)} \left(\text{Id} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \tilde{m}\Delta n^T p \right) = J^{(k+1)} \left(\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \begin{bmatrix} \tilde{m}_0 \\ \tilde{m}_1 \\ \tilde{m}_2 \end{bmatrix} \Delta n^T p \right) = J^{(k+1)} \left(\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \begin{bmatrix} \tilde{m}_0 \Delta n^T p \\ \tilde{m}_1 \Delta n^T p \\ \tilde{m}_2 \Delta n^T p \end{bmatrix} \right) = J^{(k+1)} \begin{pmatrix} u - \tilde{m}_0 \Delta n^T p \\ v - \tilde{m}_1 \Delta n^T p \\ 1 - \tilde{m}_2 \Delta n^T p \end{pmatrix} \\
&= J^{(k+1)} \begin{pmatrix} \frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} \\ \frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} \\ 1 \end{pmatrix} = J^{(k+1)} \left(\frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p}, \frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} \right)
\end{aligned}$$

二維泰勒展開式

$$\begin{aligned}
f(x, y) &\approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\
&+ \frac{1}{2!} (f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2) + \dots \\
&\approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \quad (3.5)
\end{aligned}$$

令 $x = \frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p}$, $y = \frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p}$, $x_0 = u, y_0 = v$, 代入泰勒展開式(3.5)得：

$$\begin{aligned}
J^{(k+1)}((\text{Id} - \tilde{m}\Delta n^T)p) &= J^{(k+1)} \left(\frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p}, \frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} \right) \\
&\approx J^{(k+1)}(u, v) + \left(\frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - u \right) J_x^{(k+1)}(u, v) + \left(\frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - v \right) J_y^{(k+1)}(u, v) \\
&\approx J^{(k+1)}(p) + \left(\frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - u \right) J_x^{(k+1)}(p) + \left(\frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - v \right) J_y^{(k+1)}(p)
\end{aligned}$$

由於做完泰勒展開式後，有分母 $1 - \tilde{m}_2 \Delta n^T p$ 不好處理，故要簡化之，已知當 iteration 收

斂時， Δn 會快速遞減為 0 向量，故 $\Delta n \rightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ ，所以 $\tilde{m}_2 \Delta n^T p \rightarrow 0$

因此令 $\tilde{m}_2 \Delta n^T p = 0$ ，則得到：

$$\begin{aligned}
&J^{(k+1)}((\text{Id} - \tilde{m}\Delta n^T)p) \\
&\approx J^{(k+1)}(p) + \left(\frac{u - \tilde{m}_0 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - u \right) J_x^{(k+1)}(p) + \left(\frac{v - \tilde{m}_1 \Delta n^T p}{1 - \tilde{m}_2 \Delta n^T p} - v \right) J_y^{(k+1)}(p) \\
&\approx J^{(k+1)}(p) - (\tilde{m}_0 \Delta n^T p) J_x^{(k+1)}(p) - (\tilde{m}_1 \Delta n^T p) J_y^{(k+1)}(p)
\end{aligned}$$

$$\begin{aligned}
E_{k+1} &= \sum_{p \in \Omega} (I(p) - J(H(n_k + \Delta n)p))^2 = \sum_{p \in \Omega} (I(p) - J^{(k+1)}(\text{Id} - \tilde{m} \Delta n^T)p))^2 \\
&\approx \sum_{p \in \Omega} (I(p) - J^{(k+1)}(p) + (\tilde{m}_0 \Delta n^T p) J_x^{(k+1)}(p) + (\tilde{m}_1 \Delta n^T p) J_y^{(k+1)}(p))^2 \\
&= \sum_{p \in \Omega} \left(I(p) - J^{(k+1)}(p) + \begin{pmatrix} \tilde{m}_0 [\Delta n_0 & \Delta n_1 & \Delta n_2] \\ u \\ v \\ 1 \end{pmatrix} J_x^{(k+1)}(p) + \begin{pmatrix} \tilde{m}_1 [\Delta n_0 & \Delta n_1 & \Delta n_2] \\ u \\ v \\ 1 \end{pmatrix} J_y^{(k+1)}(p) \right)^2 \\
&= \sum_{p \in \Omega} (I(p) - J^{(k+1)}(p) + (\Delta n_0 u + \Delta n_1 v + \Delta n_2) (\tilde{m}_0 J_x^{(k+1)}(p) + \tilde{m}_1 J_y^{(k+1)}(p)))^2
\end{aligned}$$

We finally set

$$\begin{aligned}
F_0 &:= u F_2 & F_1 &:= u F_2 \\
F_2 &:= \tilde{m}_0 J_x^{(k+1)}(p) + \tilde{m}_1 J_y^{(k+1)}(p) \\
D &:= I(p) - J^{(k+1)}(p)
\end{aligned}$$

$$\begin{aligned}
E_{k+1} &= \sum_{p \in \Omega} (\Delta n_0 F_0 + \Delta n_1 F_1 + \Delta n_2 F_2 + D)^2 \\
\frac{\partial E_{k+1}}{\partial \Delta n_i} &= 2 \sum_{p \in \Omega} F_i (\Delta n_0 F_0 + \Delta n_1 F_1 + \Delta n_2 F_2 + D) = 0 \\
\sum_{p \in \Omega} \begin{bmatrix} F_0^2 & F_0 F_1 & F_0 F_2 \\ F_0 F_1 & F_1^2 & F_1 F_2 \\ F_0 F_2 & F_1 F_2 & F_2^2 \end{bmatrix} \begin{bmatrix} \Delta n_0 \\ \Delta n_1 \\ \Delta n_2 \end{bmatrix} &= - \sum_{p \in \Omega} \begin{bmatrix} F_0 D \\ F_1 D \\ F_2 D \end{bmatrix} \Rightarrow \sum_{p \in \Omega} \begin{bmatrix} F_0^2 & F_0 F_1 & F_0 F_2 & F_0 D \\ F_0 F_1 & F_1^2 & F_1 F_2 & F_1 D \\ F_0 F_2 & F_1 F_2 & F_2^2 & F_2 D \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta n_0 \\ \Delta n_1 \\ \Delta n_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\
&\Rightarrow F \Delta n = \mathbf{0}
\end{aligned}$$

算出 Δn ，可以得到新的法向量 $\mathbf{n}_{k+1} = \mathbf{n}_k + \Delta n$ ，並將新的法向量代入下一次迴圈。

最後當法向量 \mathbf{n}_{k+1} 收斂後，由於這裡的法向量 \mathbf{n} 都是在透過 $B = \begin{bmatrix} M_r^{-1} & -M_r^{-1} m_r \\ 0 & 0 & 0 & 1 \end{bmatrix}$ 轉

換的空間上，故要將其轉回世界座標，所以 $B^{-T} \begin{bmatrix} \mathbf{n}_{k+1} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{conv} \\ b_{conv} \\ c_{conv} \\ d_{conv} \end{bmatrix}$ ，即為所求的收斂平面

參數。

收斂條件：令 $E^{(k)}$ 為第 k 次迴圈的誤差， $E^{(k+1)}$ 為第 $k+1$ 次迴圈的誤差，則當 $\left| \frac{E^{(k+1)} - E^{(k)}}{E^{(k+1)}} \right|$ 小於某一個 threshold 時，則判定 iterative optimization 達到收斂。

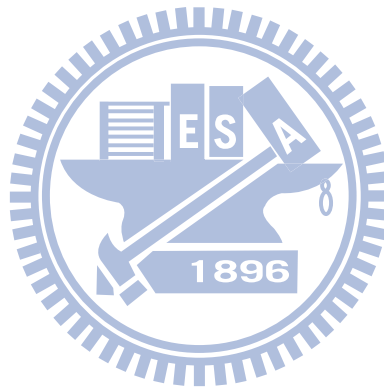
表 1 iterative optimization 的數據

| | 初始平面參數 | 初始誤差 | 迴圈數 | 收斂時誤差 | 收斂時平面參數 |
|-----|--------------------------------|---------|-----|---------|------------------------------------|
| (a) | (0.03, -0.9993, 0.02, -5.064) | 69.3866 | 8 | 60.9051 | (-0.003, -0.9999, -0.0002, 0.6981) |
| (b) | (0.9986, 0.0499, -0.02, 1.332) | 33.136 | 12 | 25.8439 | (0.9999, -0.0048, -0.0061, 0.5892) |

| | | | | | |
|-----|---------------------------------|---------|----|---------|-------------------------------------|
| (c) | (0.9989 0.0244 -0.041, -157.2) | 40.8148 | 21 | 32.5579 | (0.9999, -0.0021, 0.0003, -200.73) |
| (d) | (0.019, -0.999, 0.039, -166.37) | 33.1689 | 15 | 25.7916 | (-0.002, -0.9999, -0.0006, -198.83) |
| (e) | (0.025, 0.04, 0.9988, -452.78) | 74.3815 | 32 | 47.6023 | (-0.0026, -0.0014, 0.9999, -399.41) |
| (f) | (0.029, 0.043, 0.9987, -234.26) | 40.4442 | 11 | 25.0043 | (0.0001, 0.0017, 0.9999, -199.83) |

表 1(a)~(f)分別用圖 6(a)~(f)左的 3 點產生的初始平面參數，並運用 6(a)~(f)右的 bounding window 算出初始誤差，最後做 iterative optimization，得到更好的平面參數與更低的誤差。

註：這裡的誤差為 $\frac{\text{總誤差}}{(\text{bounding window 面積})(\text{影像數}-1)}$ ，其中-1 是減去 reference 影像。

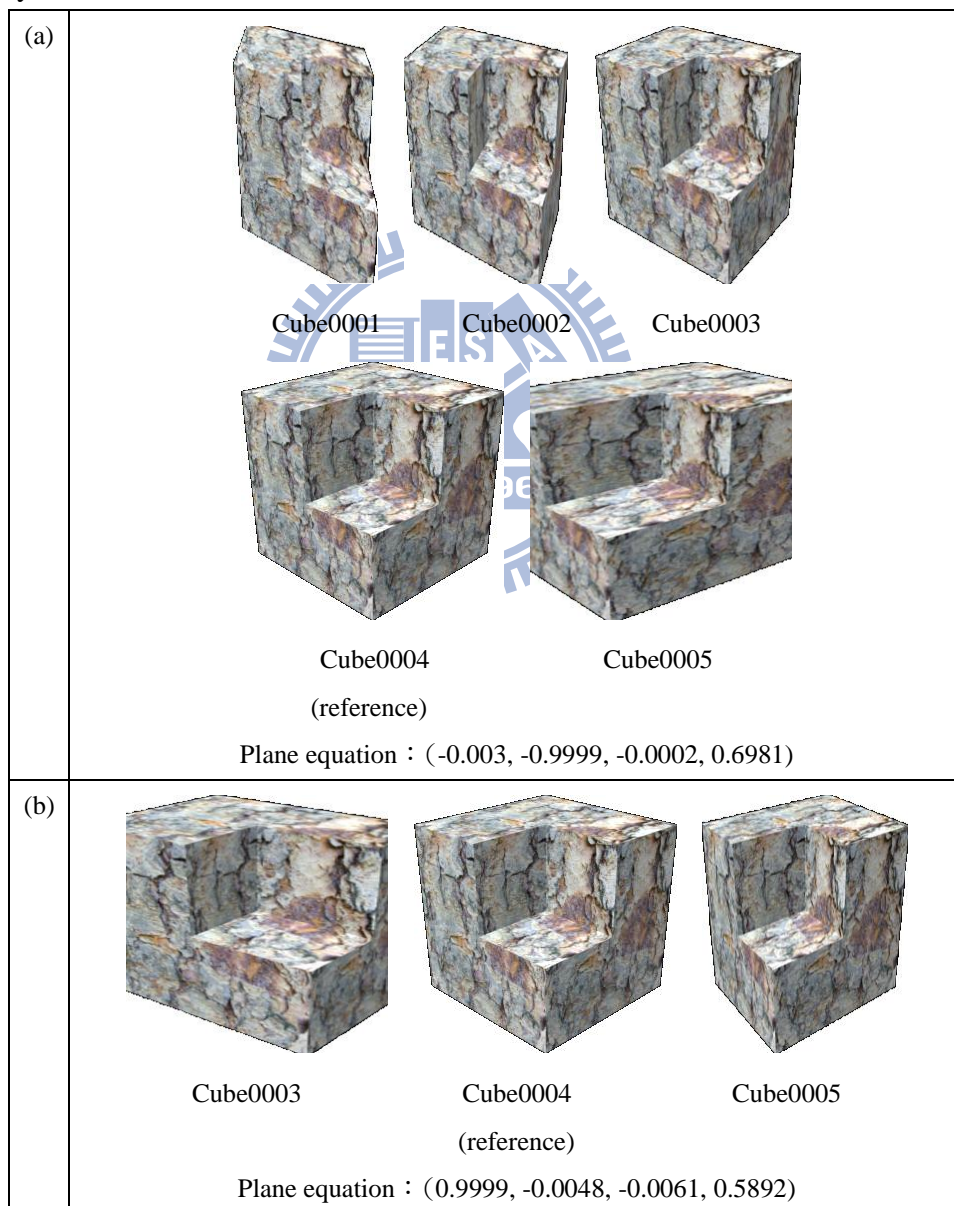


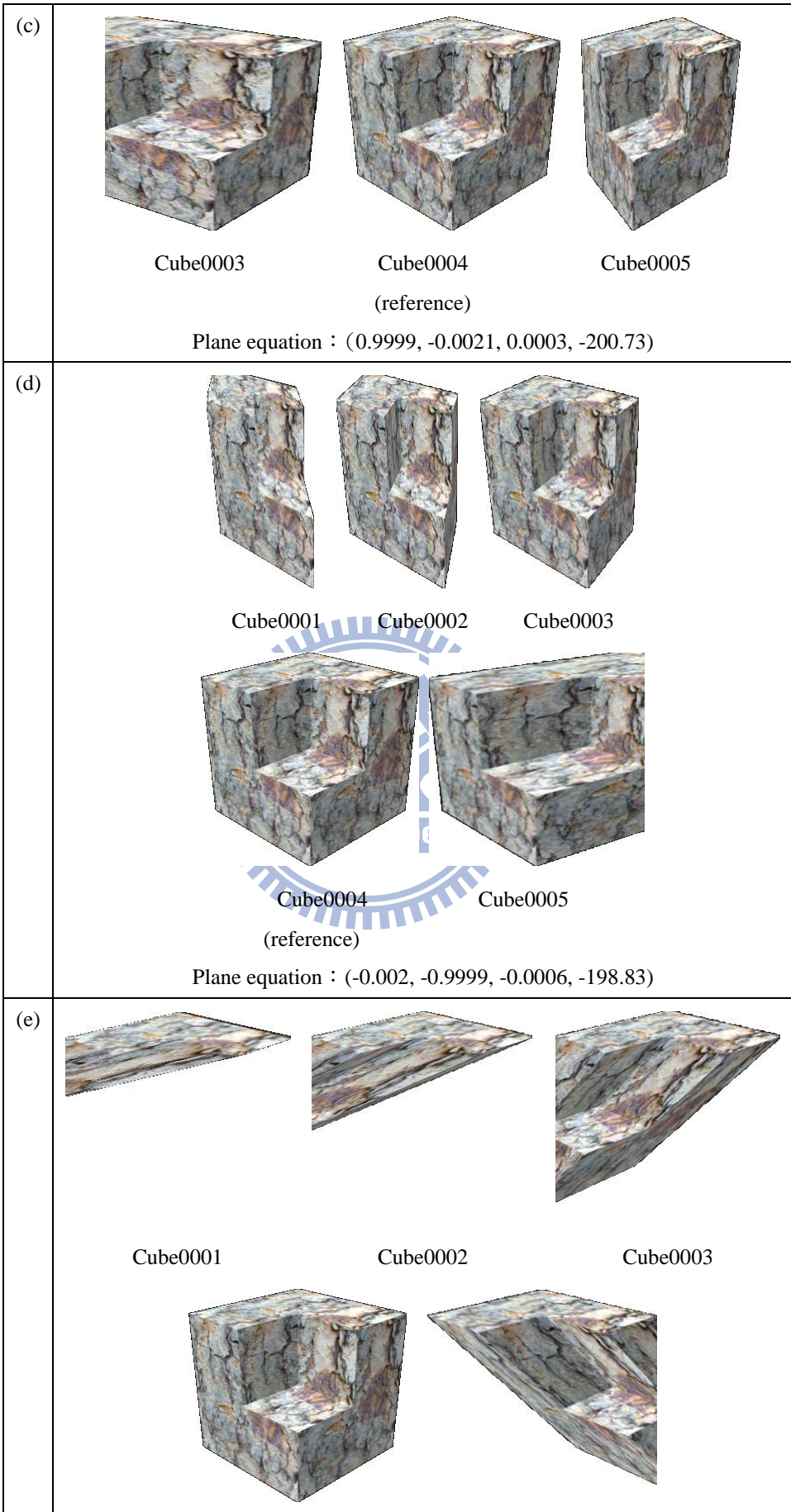
第四章 稠密式三維重建

4.1 Homography 與 Pixel-wise 標準差

做完 iterative optimization 後，算出收斂後平面參數後，會將 reference 影像投影到收斂後平面，再投影到其他的影像上，這裡只要用 3.3 節中的公式(3.2)就能得到經 homography 轉換過的影像。

由於投影到其他影像的座標非整數值，所以會使用 bicubic 內插法來得到 subpixel 的 intensity。





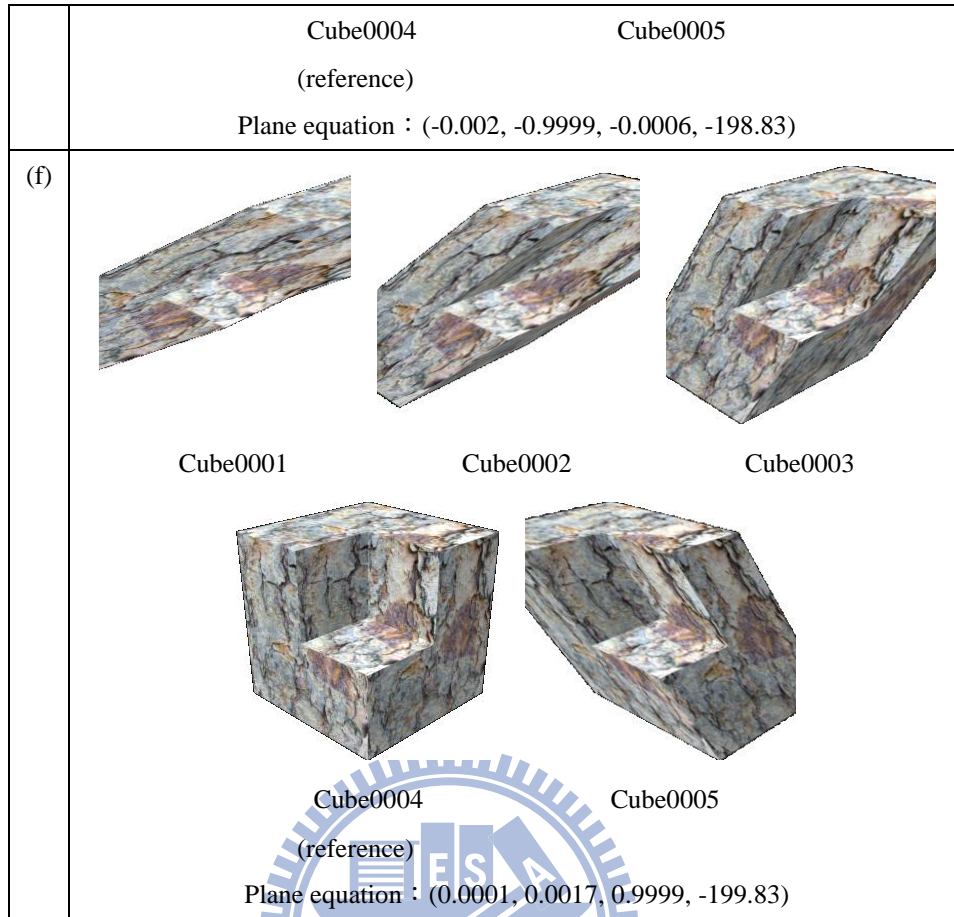


圖 8 對收斂平面做 homography 的結果

圖 8(a)~(f)分別為將 reference image 投影到圖 8(a)~(f)中的平面方程式，在投到其他影像的結果。

我們推測投影後的影像中吻合目標平面的深度及 normal 的部分中的 intensity 變動量小，故標準差小，而不吻合的部分中的 intensity 變動量大，所以標準差大。因此當我們將所有投影後的影像，每個 pixel 的 intensity 做標準差，標準差比較低的視為我們要的部分。

設有 M 個大小皆為 $\text{width} \times \text{height}$ 的影像，令 $p_k(i, j)$ 為第 k 個影像在 (i, j) 上的 intensity, $k=1 \sim M, i=1 \sim \text{width}, j=1 \sim \text{height}$ 。現在欲求 D , D 是 $\text{width} \times \text{height}$ 矩陣，且 $D(i, j) =$

$\frac{1}{M} \sum_{k=1}^M (p_k(i, j) - \overline{p_k(i, j)})^2$ 。其中 $D(i, j)$ 為 $p_1(i, j), p_2(i, j), \dots, p_M(i, j)$ 的標準差，其中

$$\overline{p_k(i, j)} = \frac{1}{M} \sum_{k=1}^M p_k(i, j)。$$

由於這裡的影像都有背景，且背景の intensity 都一樣，這樣的話，會造成背景の標準差值很小，但背景卻不是我們要的部分。這裡的目的本來就要找標準差小的區域，故背景の標準差值小，會造成干擾。因此，本論文の標準差只計算非背景の交集部分，其餘部分，就強迫給大值，這樣，背景の標準差就會很大，不會影響我們擷取低標準差的部分。

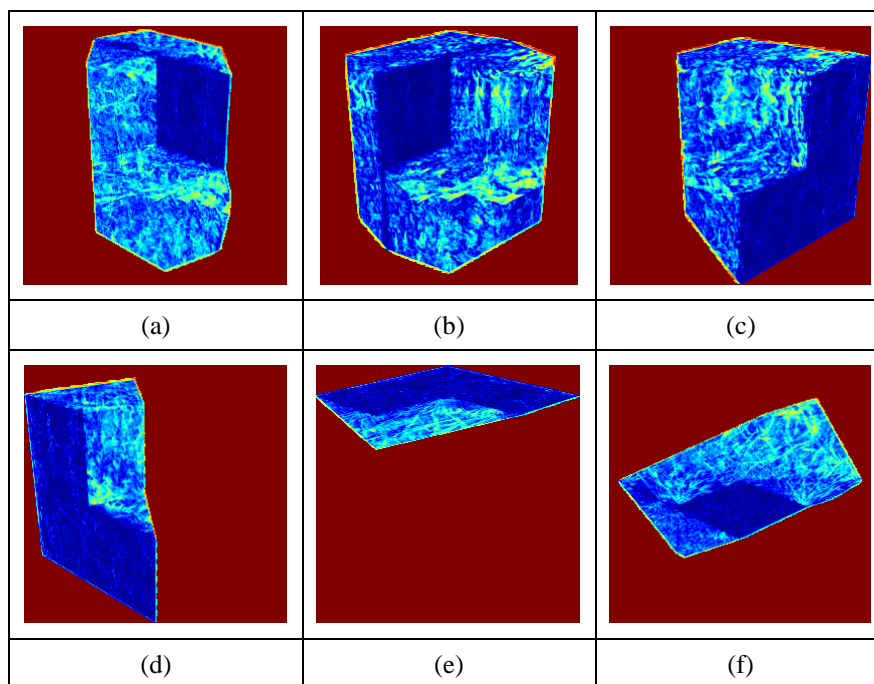


圖 9 對 homography 後的影像做標準差的結果

圖 9(a)~(f)分別為對圖 8(a)~(f)的所有經 homography 轉換後的影像，做完標準差後的結果。

註：這裡的顏色是使用 matlab 的函式 colormap jet 所產生的，越接近深藍色代表 intensity 的值越低，越接近紅色代表 intensity 的值越高。

4.2 用 MSER 找出標準差小的區域並刪除區域內的特徵點

做完標準差後，要取出標準差較低的地方，之前有用 k-means 或是 otsu 門檻法做，雖然能取到標準差較低的地方，但是取出來的東西無法保證連通性，因此我們要能取出標準差值小的地方，而這些地方又需要連通，符合這些條件，最好的方法就是 MSER。

MSER 的全名為 Maximally Stable Extremal Regions，是一種灌水法，將 intensity 從小到大，找出小於 intensity 的連通區域，並從這些區域中，找出較穩定的區域，詳細情形參考[20]。

但是 MSER 通常會抓出很多區域，要怎麼知道哪個區域才是我們要的？在 3.2 節中，有求出 3 個點所形成的 bounding window，這裡我們要用 MSER 找出圖 9(a)~(f)中標準差小的區域，而且分別要包含圖 7(a)~(f)中的 bounding window，這樣就不會煩惱到底要找哪塊區域。

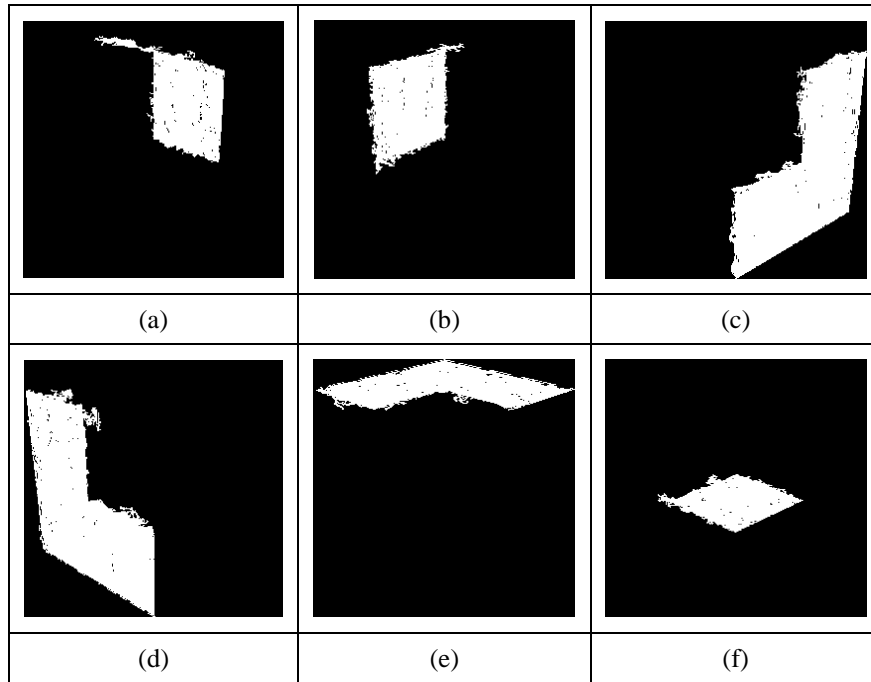


圖 10 對標準差影像做 MSER 的結果

圖 10(a)~(f)為分別對圖 9(a)~(f)的標準差影像，做 MSER 後得到的區域，也就是白色的區域。

值得注意的是，雖然 MSER 會抓到外面去，但很少會刪掉其他平面的特徵點，原因在於 local intensity min/max 是 non-uniform 的，所以標準差值通常較高。

| | | | | | | | | | | | | | | | | | | |
|----------|----------|--|----------|---------|---------|----------|----|----------|----------|----------|----|----|----------|----------|---|----|----------|--------|
| (a) | | <table border="1"> <tbody> <tr> <td>14</td> <td>18</td> <td>9(16-7)</td> <td>13(14-1)</td> </tr> <tr> <td>15</td> <td>21</td> <td>2(24-22)</td> <td>14(15-1)</td> </tr> <tr> <td>16</td> <td>19</td> <td>20</td> <td>20</td> </tr> <tr> <td>0</td> <td>11</td> <td>13</td> <td>1</td> </tr> </tbody> </table> | 14 | 18 | 9(16-7) | 13(14-1) | 15 | 21 | 2(24-22) | 14(15-1) | 16 | 19 | 20 | 20 | 0 | 11 | 13 | 1 |
| 14 | 18 | 9(16-7) | 13(14-1) | | | | | | | | | | | | | | | |
| 15 | 21 | 2(24-22) | 14(15-1) | | | | | | | | | | | | | | | |
| 16 | 19 | 20 | 20 | | | | | | | | | | | | | | | |
| 0 | 11 | 13 | 1 | | | | | | | | | | | | | | | |
| (b) | | <table border="1"> <tbody> <tr> <td>13(14-1)</td> <td>9(18-9)</td> <td>9</td> <td>13</td> </tr> <tr> <td>15</td> <td>2(21-19)</td> <td>1(2-1)</td> <td>14</td> </tr> <tr> <td>16</td> <td>19</td> <td>20</td> <td>20</td> </tr> <tr> <td>0</td> <td>11</td> <td>13</td> <td>1</td> </tr> </tbody> </table> | 13(14-1) | 9(18-9) | 9 | 13 | 15 | 2(21-19) | 1(2-1) | 14 | 16 | 19 | 20 | 20 | 0 | 11 | 13 | 1 |
| 13(14-1) | 9(18-9) | 9 | 13 | | | | | | | | | | | | | | | |
| 15 | 2(21-19) | 1(2-1) | 14 | | | | | | | | | | | | | | | |
| 16 | 19 | 20 | 20 | | | | | | | | | | | | | | | |
| 0 | 11 | 13 | 1 | | | | | | | | | | | | | | | |
| (c) | | <table border="1"> <tbody> <tr> <td>13</td> <td>9</td> <td>9</td> <td>4(13-9)</td> </tr> <tr> <td>15</td> <td>2</td> <td>1</td> <td>1(14-13)</td> </tr> <tr> <td>16</td> <td>19</td> <td>5(20-15)</td> <td>1(20-19)</td> </tr> <tr> <td>0</td> <td>11</td> <td>2(13-11)</td> <td>0(1-1)</td> </tr> </tbody> </table> | 13 | 9 | 9 | 4(13-9) | 15 | 2 | 1 | 1(14-13) | 16 | 19 | 5(20-15) | 1(20-19) | 0 | 11 | 2(13-11) | 0(1-1) |
| 13 | 9 | 9 | 4(13-9) | | | | | | | | | | | | | | | |
| 15 | 2 | 1 | 1(14-13) | | | | | | | | | | | | | | | |
| 16 | 19 | 5(20-15) | 1(20-19) | | | | | | | | | | | | | | | |
| 0 | 11 | 2(13-11) | 0(1-1) | | | | | | | | | | | | | | | |

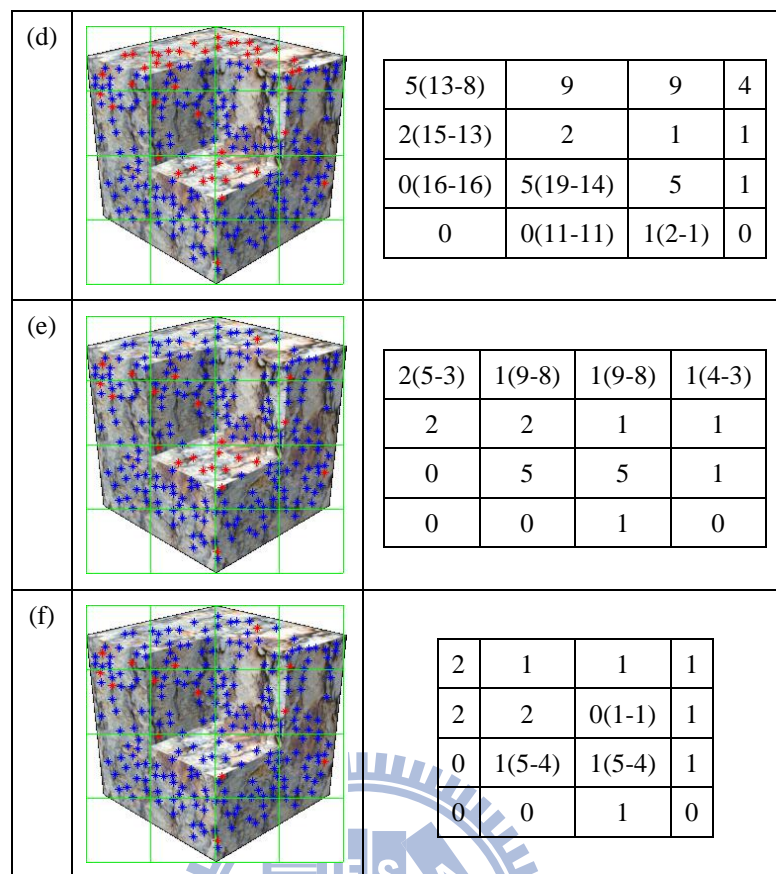


圖 11 把屬於 MSER 區域內的點標記

- (a) 左：將圖 6(a)中的特徵點，屬於 MSER 區域(圖 10(a))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。
右：以圖 6(b)為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應點數量。
- (b) 左：將圖 11(a)左中的特徵點，屬於 MSER 區域(圖 10(b))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。
右：以圖 11(a)右為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應點數量。
- (c) 左：將圖 11(b)左中的特徵點，屬於 MSER 區域(圖 10(c))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。
右：以圖 11(b)右為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應點數量。
- (d) 左：將圖 11(c)左中的特徵點，屬於 MSER 區域(圖 10(d))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。
右：以圖 11(c)右為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應點數量。
- (e) 左：將圖 11(d)左中的特徵點，屬於 MSER 區域(圖 10(e))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。
右：以圖 11(d)右為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應

點數量。

(f) 左：將圖 11(e)左中的特徵點，屬於 MSER 區域(圖 10(f))內的特徵點做標記，以藍色表示，其他未標記的，以紅色表示。

右：以圖 11(e)右為主，將標記的點(藍色的點)都刪除掉，各格子中剩下的特徵對應點數量。

當做到最後每個格子的特徵點數都小於 3 的時候，此演算法結束，如圖 11(f)右。

4.3 求出世界座標上的 3D 點

由 iterative optimization 會計算出的世界座標系中的三維平面： $ax + by + cz + d$ 。我們要將 reference 影像上的 2D 點投影到三維平面來得到此 2D 點的三維座標，做法是由鏡心的世界座標到 2D 點的世界座標這條射線，與三維平面的交點，此交點的三維座標即為所求。因此要先算出鏡心和 2D 點的世界座標。

鏡心的世界座標

令空間中一點 P 在相機座標系統之座標值為 ${}^C P$ ， ${}^C P$ 與其在世界座標系統之座標值 ${}^W P$ 的關係為：

$${}^C P = {}^C R {}^W P + {}^C O_W \quad (4.1)$$

其中 ${}^C R$ 、 ${}^C O_W$ 定義為相機外部參數 R 、 T 。

令參考相機的投影矩陣為 M_{ref} ，內部參數為 K_{ref} ，外部參數為 R_{ref} 、 T_{ref} ，並令 C 為參考相機座標系，則 $R_{ref} = {}^C R$ ， $T_{ref} = {}^C O_W$ 。

欲求在世界座標系下參考相機鏡心的三維座標 ${}^W O_{ref}$ ，由於為已校正影像，故 R_{ref} 、 T_{ref} 為已知，又已知在參考相機座標系下，參考相機鏡心的三維座標為 $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ 。

故由(4.1)式可得參考相機鏡心在世界座標下的座標值 ${}^W O_{ref}$ 計算推導如下：

$${}^C O_{ref} = R_{ref} {}^W O_{ref} + T_{ref}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = R_{ref} {}^W O_{ref} + T_{ref}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - T_{ref} = R_{ref} {}^W O_{ref}$$

$$R_{ref}^{-1} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - T_{ref} \right) = {}^W O_{ref}$$

$${}^W O_{ref} = R_{ref}^{-1} (-T_{ref}) = -R_{ref}^T T_{ref}$$

計算出 ${}^W O_{ref}$ 後再來要計算參考相機 2D 點在深度 1 的成像平面上於世界座標的座標值。

2D 點的世界座標

另 ${}^W p$ 為影像上 2D 點在世界座標系下的三維座標。

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{ref} [R_{ref} \ T_{ref}] \begin{bmatrix} {}^W p \\ 1 \end{bmatrix} \quad (4.2)$$

由於要計算影像上 2D 點的三維座標，所以 $s = 1$ ，將 (4.2) 式寫成 non-homogeneous，並代入 $s = 1$ 。

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= K_{ref} R_{ref} {}^W p + K_{ref} T_{ref} \\ \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - K_{ref} T_{ref} &= K_{ref} R_{ref} {}^W p \\ {}^W p &= (K_{ref} R_{ref})^{-1} \left(\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - K_{ref} T_{ref} \right) \end{aligned} \quad (4.3)$$

有了鏡心與 2D 點的世界座標，可以得到一條射線 ${}^W O_{ref} + t({}^W p - {}^W O_{ref})$ 。

這條射線與目標三維平面的交點，就能得到 2D 點投影到目標三維平面的三維點。

以圖 9(a)~(f) 中 MSER 抓出來的區域的 2D 點為主，將這些 MSER 白色區域的 2D 點，分別投影到表 1(a)~(f) 中收斂的平面，就可以得到 6 個平面的點雲，但有些平面的範圍超出物體外，這時可以透過 3D 空間中平面相交的關係，去除屬於物體之外的錯誤區域。

4.4 運用空間切割建出三維模型

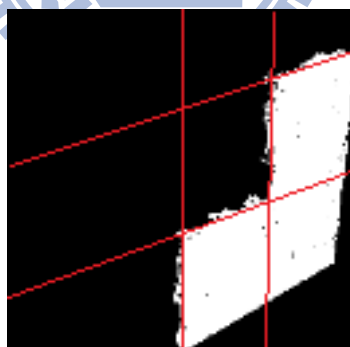


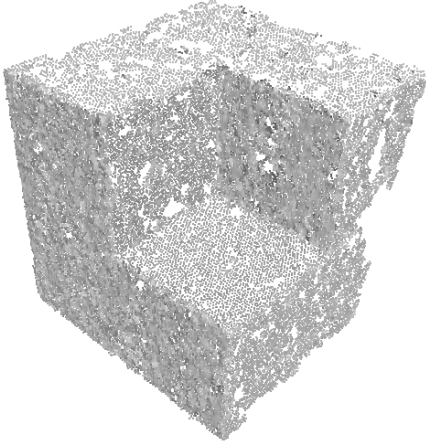
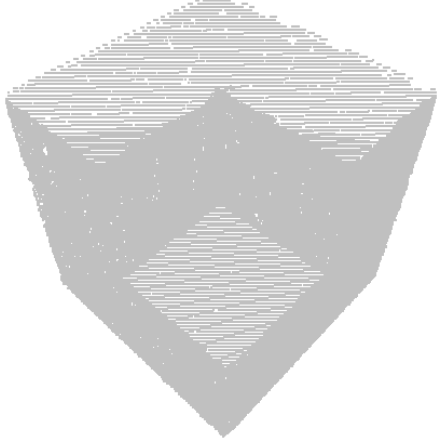
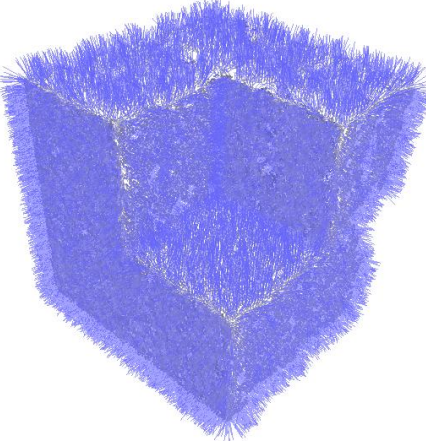
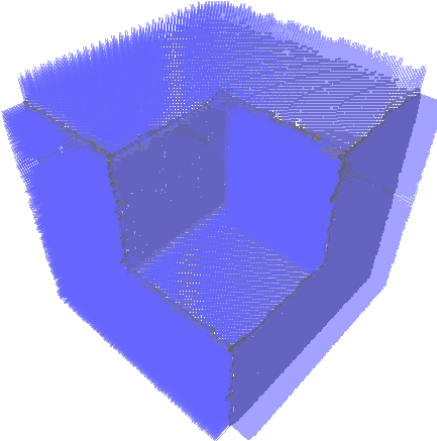
圖 12 圖 10(c) 做平面切割的結果

一個平面的與許多平面的許多交線，會將平面分成許多區塊，如圖 12，共有 9 個區域，如果每個區塊中的點數和 MSER 取的點數(圖 12 中的白色的點數)的比率過低，則將那些點刪掉，最後剩下的是沒有雜點的平面。

第五章 實驗結果

5.1 具豐富紋理平面物體三維重建可行性驗證

重建的三維點雲資料與三維網格模型結果如下，並與 Furukawa 方法的重建結果進行比較：

| | Furukawa's result | Our result |
|--|---|--|
| Point set |  |  |
| Oriented point set (Point cloud with surface normal) |  |  |

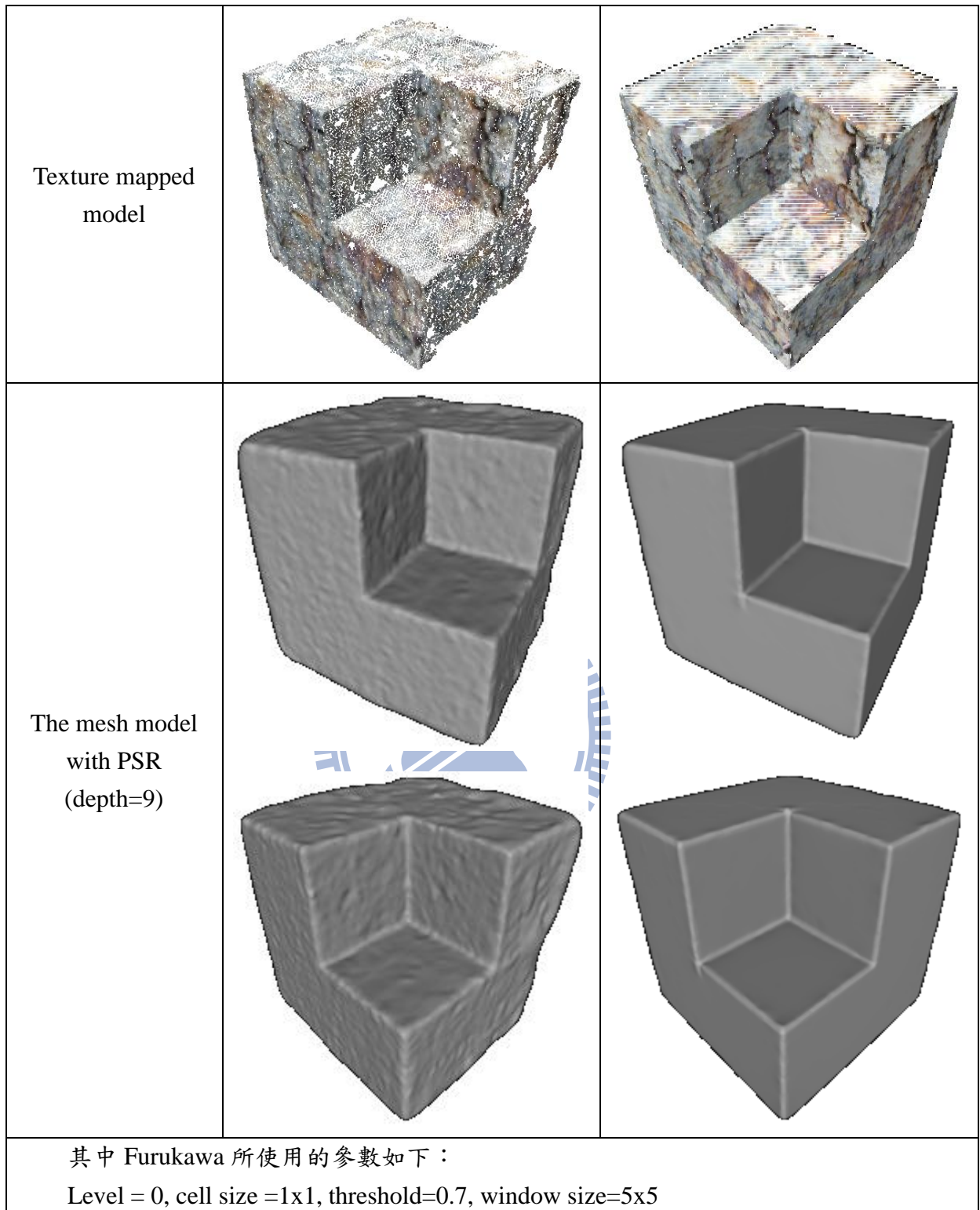


圖 13 Furukawa 和我們方法建 Convex cube 之比較

從圖 13 可以看出 Furukawa 的重建結果並不如我們方法的平順，主要原因為 Furukawa 用許多小 patch 來長出平面，因此同一平面的每個小 patch 的法向量皆有差異，造成平面不平順的結果。而本論文的方法是一個平面就是一塊 patch，且法向量相同，故較平順。

以 reference 影像的特徵點投到空間平面，再投回其他影像的投影誤差距離，如表 2。

表 2 以 View4 為 reference 影像的投影誤差

| 特徵點 | View1 | View2 | View3 | View4 (reference) | View5 |
|-----|--------|--------|--------|----------------------|--------|
| 1 | 0.5138 | 0.5073 | 0.3847 | 0 | 0.2946 |
| 2 | X | 0.5806 | 0.073 | 0 | 0.2161 |
| 3 | X | 0.8999 | 0.5642 | 0 | 0.8407 |
| 4 | 1.1575 | 0.6935 | 0.125 | 0 | 0.4402 |
| 5 | 1.7308 | 1.1002 | 0.3851 | 0 | 0.6317 |
| 6 | 0.4636 | 0.224 | 0.5902 | 0 | 1.6874 |

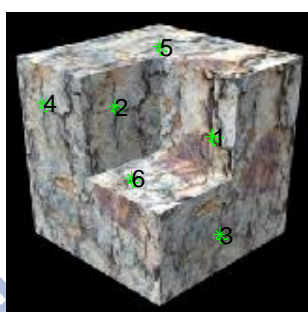


圖 14 重建結果檢驗-View4

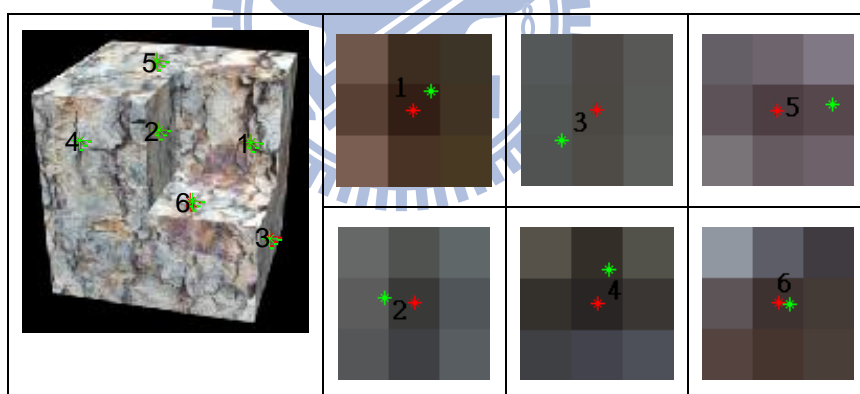


圖 15 重建結果檢驗-View2

圖 15 中紅點為 View2 的特徵點，綠點為 View4 的特徵點到 View2 的反投影點。

5.2 稀疏紋理平面物體三維重建可行性驗證

實驗環境：

使用 1 台相機，(故只有 1 個內部參數 K)，其型號為 Canon EOS 500D，對一靜止物體做多視角拍攝(相機動，物體不動)。

拍攝光源使用發光連續的棚燈，拍攝時光圈、快門、ISO 固定，並使用相同的白平衡演算法(鎢絲燈，因棚燈是鎢絲燈)，目的是使每張影像顏色不會差太多。

先利用一些平面上的特徵對應點(landmark)來得到 reference 影像與其他影像的 homography 矩陣，每個 view 的影像都透過各自的 homography matrix 轉至 reference 影像。

以 homography 轉換後與 reference 影像差異性小的區域往外長，差異性小的區域包含 texture & uniform 區域，直至有明確的 boundary 為止，這點與傳統技術中由特徵點往外 expand 的方式不一樣。

- 寧可讓算出的範圍超出正確邊緣而形成過大的區域，之後再用 3D 空間關係來刪除 ambiguous 的區域。(凸多邊形)
- 不願讓算出的範圍小於正確邊緣的區域，因為這樣還是沒辦法找出正確的邊緣。

本實驗首先以 mailbox (共有 5 張影像，如下圖)的上蓋為要重建的平面範圍，並以其他平面的 edge 資訊做為平面擴張的邊緣。

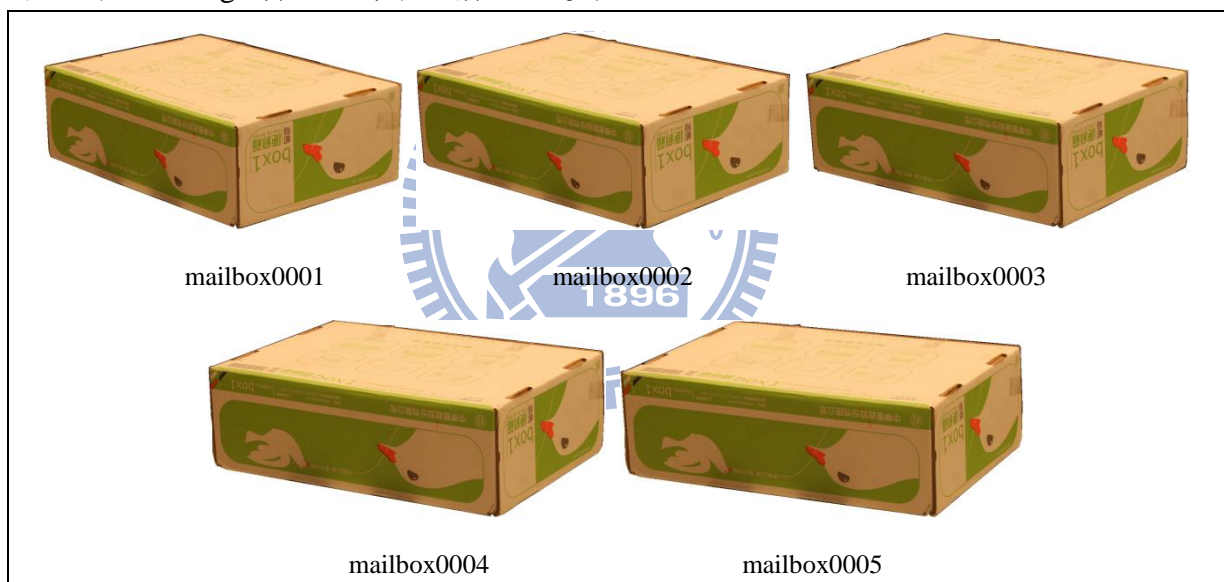
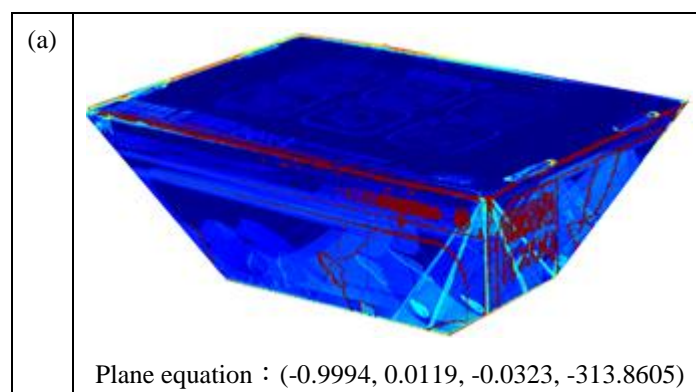


圖 16 Input image：待建物 mailbox 的影像



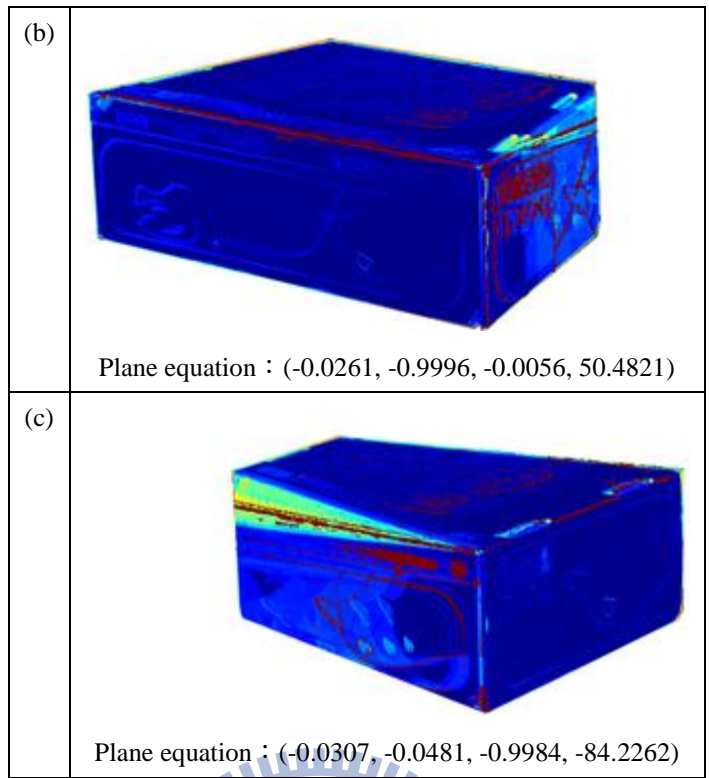
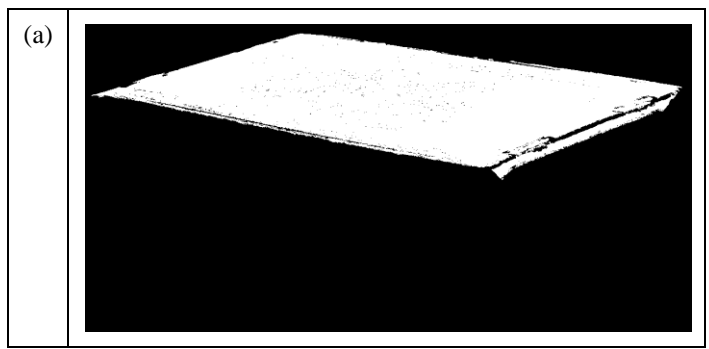


圖 17 mailbox 做標準差的結果

這些標準差皆是以 mailbox0003 當作 reference 所得到的。

- (a) 以 mailbox 的上面為主，所產生的標準差圖以及其收斂後平面參數。
- (b) 以 mailbox 的左面為主，所產生的標準差圖以及其收斂後平面參數。
- (c) 以 mailbox 的右面為主，所產生的標準差圖以及其收斂後平面參數。



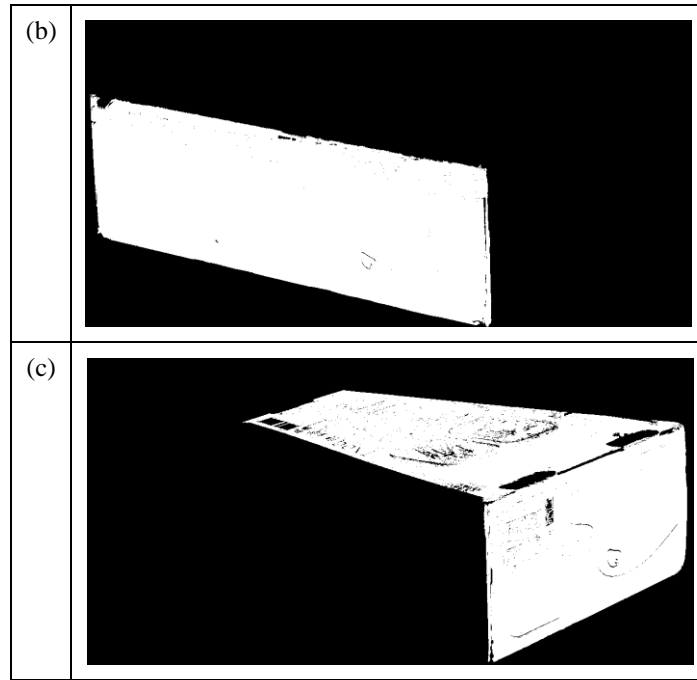
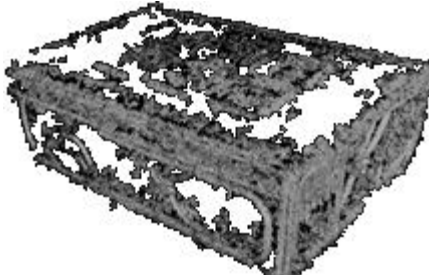
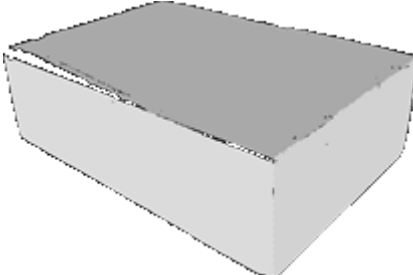


圖 18 將標準差用 MSER 抓出來的結果(mailbox)

- (a) 利用 MSER 來尋找圖 17(a)中標準差較低的區域，其 MSER 參數設定為： $\delta = 1$, $\text{variation} = 0.1$ 。
- (b) 利用 MSER 來尋找圖 17(b)中標準差較低的區域，其 MSER 參數設定為： $\delta = 1$, $\text{variation} = 0.1$ 。
- (c) 利用 MSER 來尋找圖 17(c)中標準差較低的區域，其 MSER 參數設定為： $\delta = 1$, $\text{variation} = 0.1$ 。此結果較為不理想，逸出不少區域至上蓋，不過這可以透過 3D 空間中平面相交的關係，去除屬於物體之外的錯誤區域。

| | Furukawa's result | Our result |
|--------|---|--|
| Vertex |  |  |

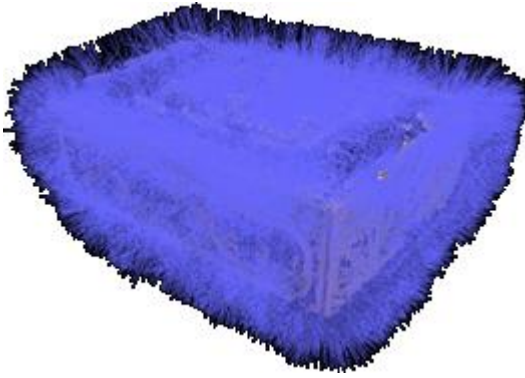
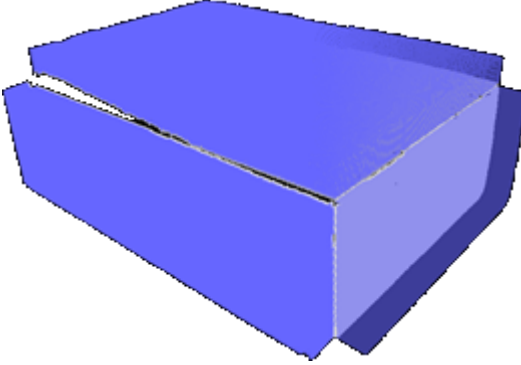


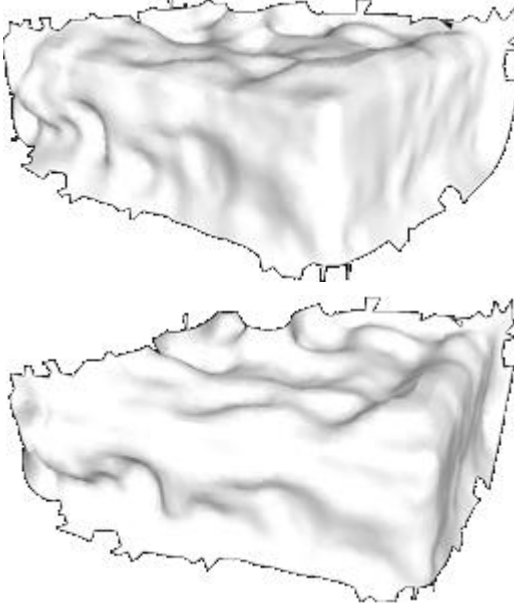
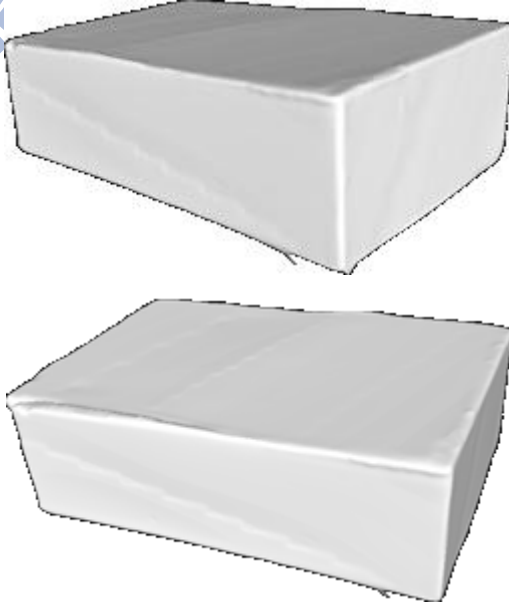
| | | |
|--|---|--|
| Vertex with normal |  |  |
| Texture mapped vertex |  |  |
| The mesh model after PSR (depth=9) |  |  |
| <p>其中 Furukawa 所使用的參數如下：Level = 1, cell size = 1x1, threshold = 0.7, window size = 5x5。並去除經 PSR 處理後的 mesh model 中邊長過大的 face。</p> | | |

圖 19 重建三維點雲資料與三維網格模型結果，並與 Furukawa 的重建結果進行比較

從圖 19 可以看出 Furukawa 的重建結果並不好，其原因來自於 texture-less 的區域。Furukawa 方法中是以特徵點為基礎，且是由穩定、具紋理資訊的特徵點往外做拓展 (expansion) 的動作，較無共平面的概念在內。然而拓展至 texture-less 的區域時，因僅能以附近的特徵點法向量及 3D 位置做為其拓展 patch 的初始值，一旦 texture-less 的區域過大，就無法透過此方法有效地產生鄰近的新 patch。且 Furukawa 所採取去除誤差過大的 3D 座標點的 filtering 步驟，亦無法有效發揮效用，造成最後的 point cloud 其實是非常 noisy 的。透過 PSR 重建結果可以明顯看出其凹凸不平的表面，主要原因即是沒有考慮到整體平面的關係，拓展過程中若遇到無紋理資訊的區域，即無法處理，且後續空間中的 smoothness constraint 也無法發揮作用。

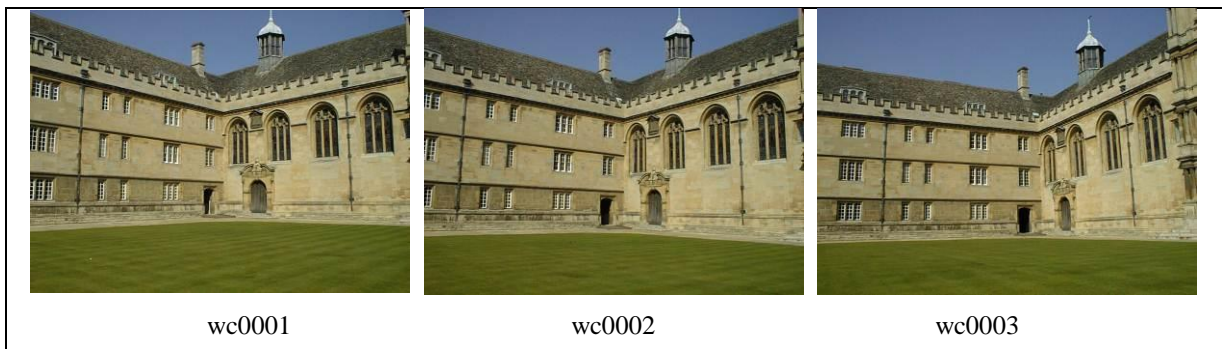
而我們方法則是透過 homography 轉換(轉換公式為(3.2))後，同時一次計算多視圖影像之間的標準差，再擷取出 MSER 的區域，如此針對較不具紋理的平面物體，將具有明顯改善的功效。

5.3 The 3D reconstruction of Wadham College

此組實驗乃針對網路上所抓取的影像資料 (<http://www.robots.ox.ac.uk/~vgg/data2.html>) 進行建築物的三維重建，其五張原始影像 views 1~5 如下：

實驗環境：

使用一台數位相機，其型號為 Olympus C-820L，以多視角拍攝一建築物(相機動，物體不動)。



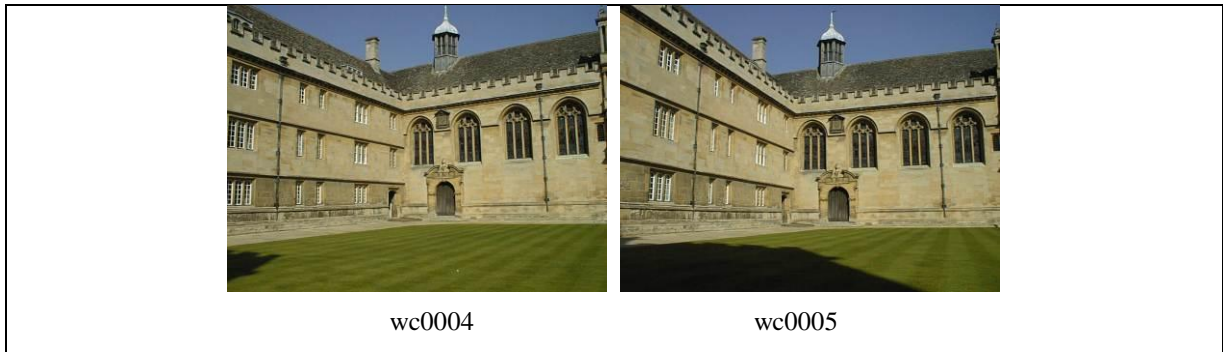
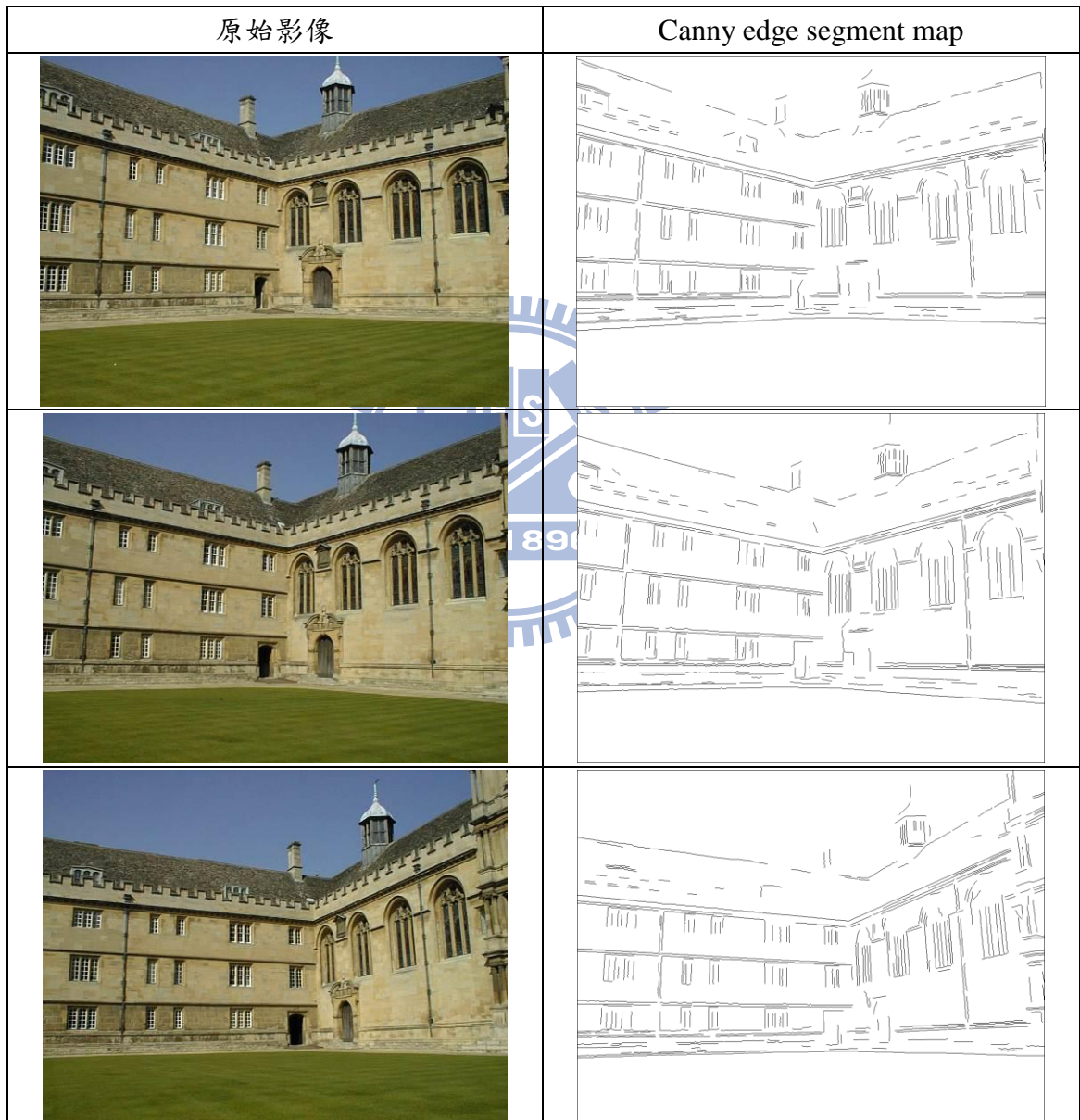


圖 20 待建物 wc(Wadham College)的影像



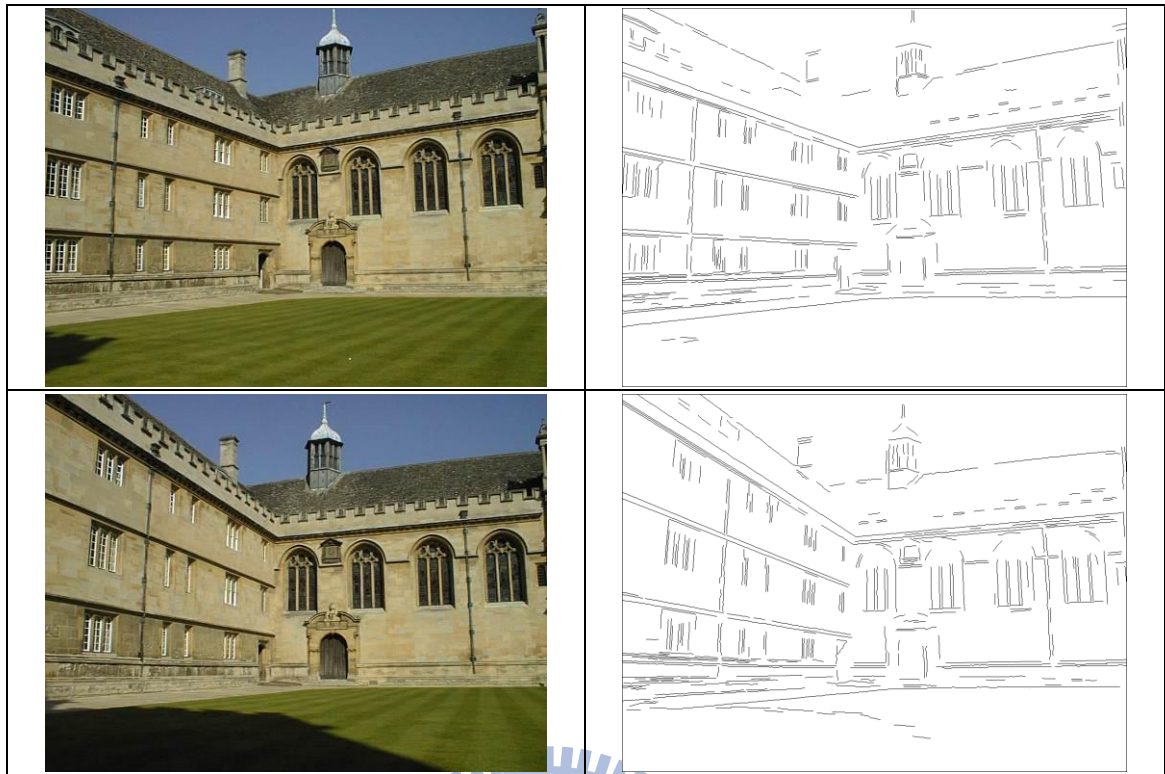
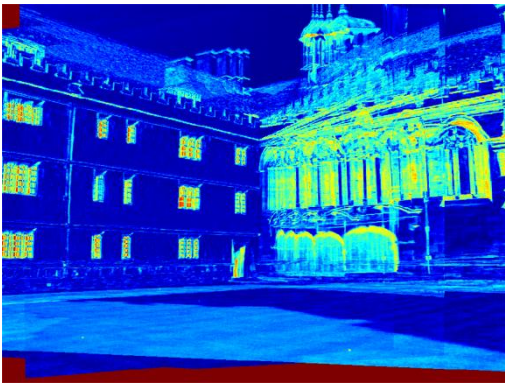
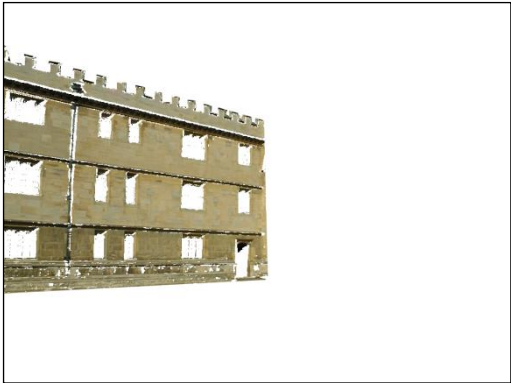
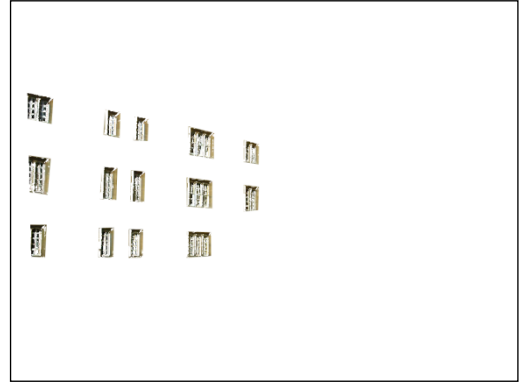
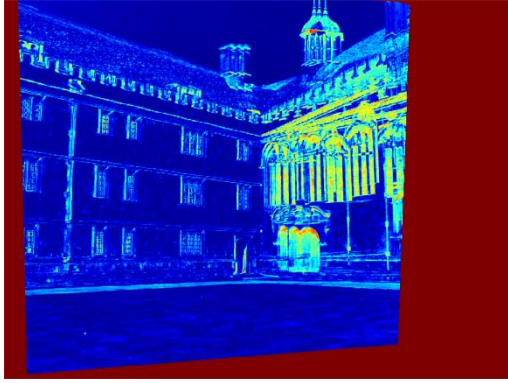


圖 21 透過 Canny edge 演算法計算出 edge 相關資訊



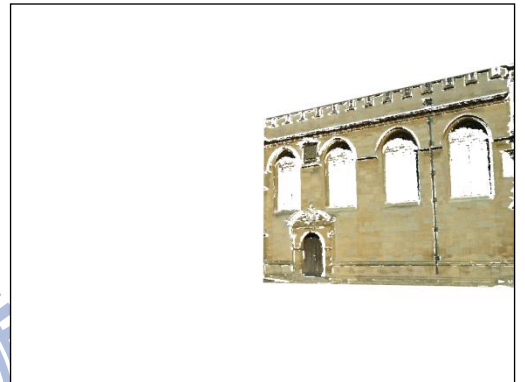
| | 多視圖影像之亮度標準差 | Planar patch extracted |
|--|---|--|
| 左側牆 |  |  |
| Plane Equation: (-0.122207, 2.73108, 0.00356, 0.44526) | | |

左側窗框



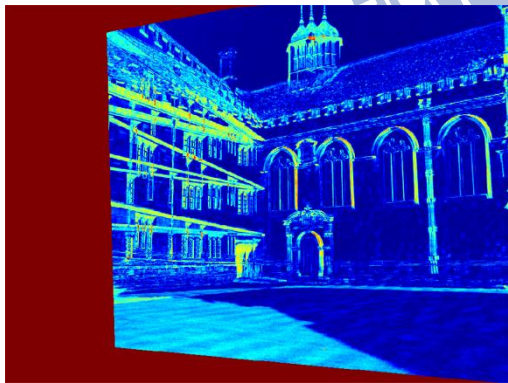
Plane Equation: $(-0.17636, 8.56395, 0.075354, -0.40249)$

右側牆



Plane Equation: $(6.35967, 10.22689, -0.042374, 48.66512)$

右側窗框



Plane Equation: $(-0.124784, -0.00636, 7.17101, -50.93003)$

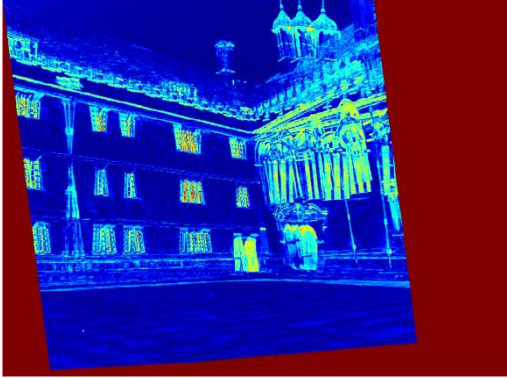

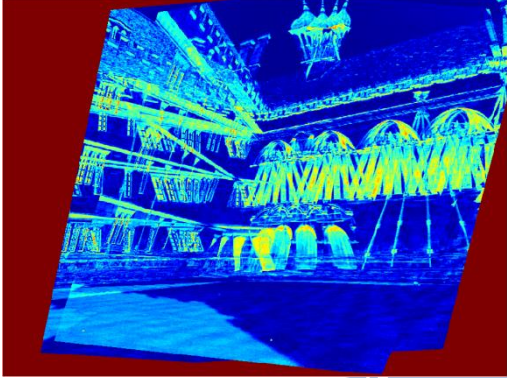
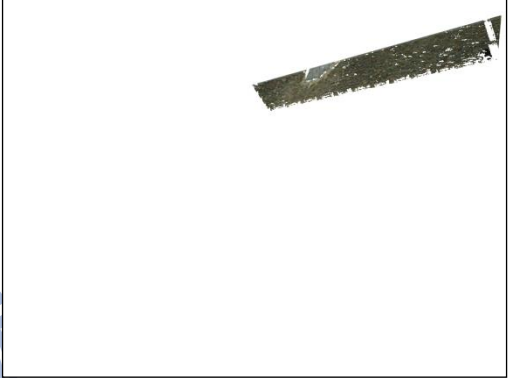
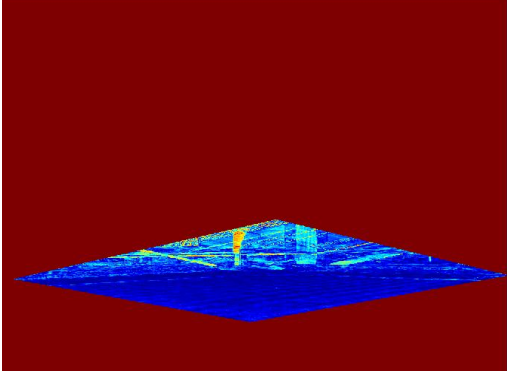
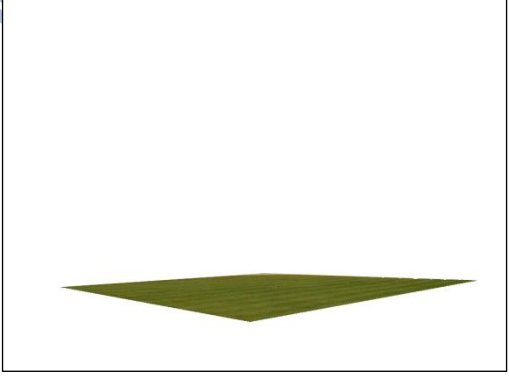
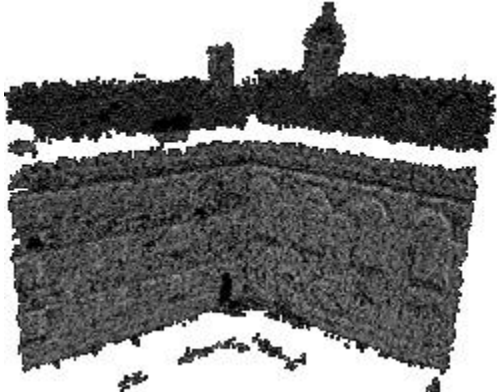

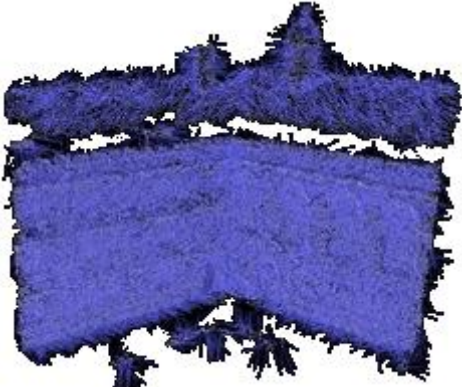
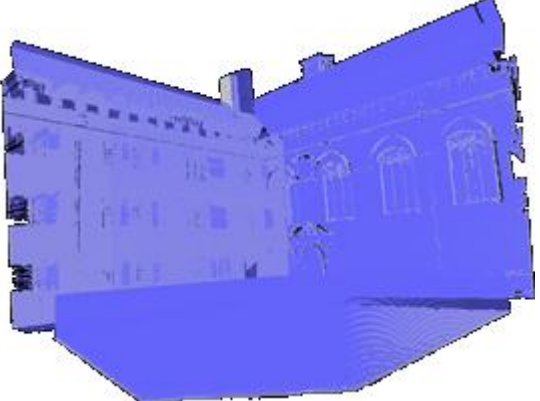

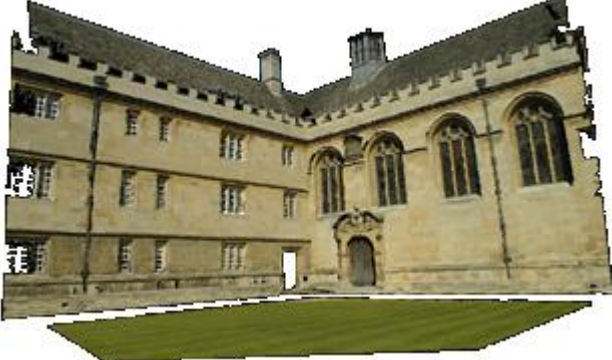
| | | |
|--|---|--|
| 左側屋頂 |  |  |
| Plane Equation: (6.35967, 10.22689, -0.04237, 48.66512) | | |
| 右側屋頂 |  |  |
| Plane Equation: (-25.628434, 0.92826, -31.87700, 24.96075) | | |
| 草皮 |  |  |
| Plane Equation: (-5.20439, -0.04653, -0.00406, 4.07834) | | |

圖 22 標準差以及其擷取平面

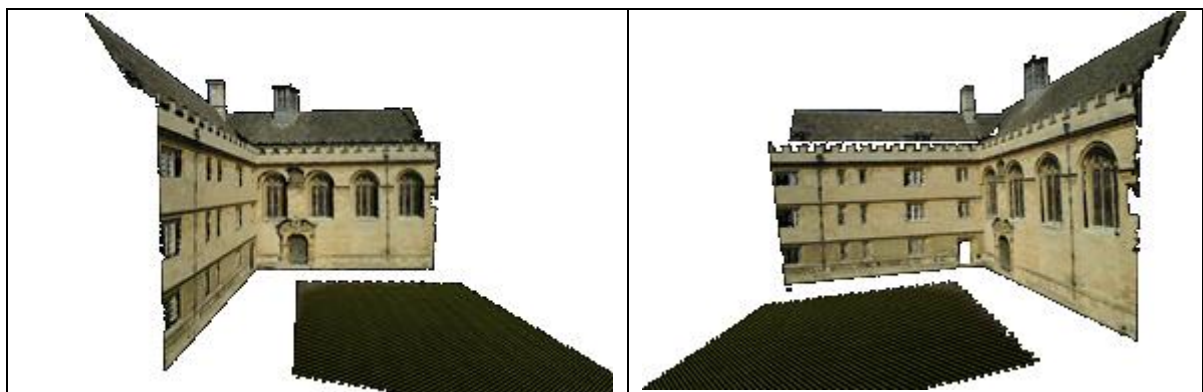
另外，我們也對屋頂煙窗、觀景台做重建，其面積較小，不予列出。

| | | |
|--|-------------------|------------|
| | Furukawa's method | Our method |
|--|-------------------|------------|

| | | |
|----------------------|---|--|
| Point set |  |  |
| Oriented point set |  |  |
| Texture mapped model |  |  |

其中 Furukawa 所使用的參數如下：Level = 1, cell size = 1x1, threshold = 0.75, window size = 5x5

圖 23 重建的三維點雲資料與三維網格模型結果，並與 Furukawa 方法重建結果進行比較



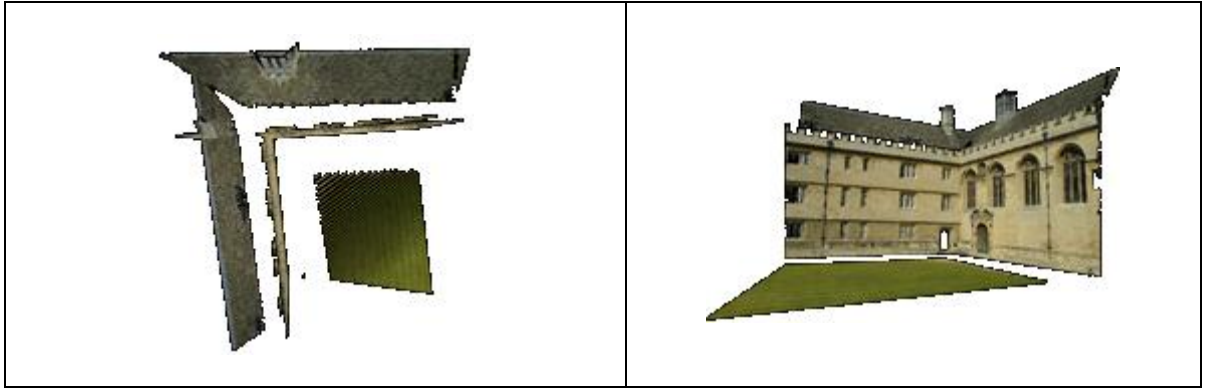
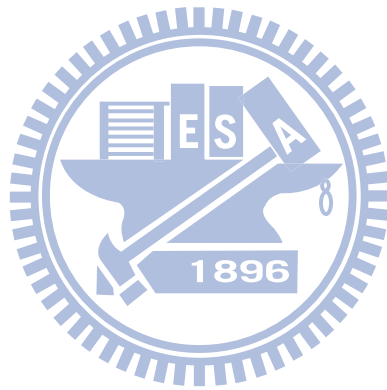


圖 24 用 meshlab 所呈現建築物的不同面貌，包括俯瞰圖的兩片牆夾角約為 90 度等



第六章 結論與未來發展

本論文利用所發展的多視圖三維重建法實施在二組電腦合成影像，用以檢視其準確度，同時也應用於二組實拍影像，以檢視其實用性。我們這些重建例子與用網站上所提供的 Furukawa 方法所重建結果做比較。我們方法與 Furukawa 的方法差別主要有下列幾點：

- (1) 我們不像 Furukawa 方法使用固定的 window size 來做 feature matching。因為所要重建的物體表面有大大小小的平面，固定的 window size 往往包括只單一的 planar patch，多視圖之間的比對，就無法成功。再者，大塊的 planar patch 若使用固定的 window size 比對，會分成好多次執行，而往往造成原本大塊平面數個小平面而變成不平。我們的方法會隨物體平面實際大小而造出平面來。
- (2) Furukawa 方法因使用固定的 window size，無法框住一個完整重覆性圖樣 (repetitive pattern)，造成比對不良，故該方法處理具重覆性紋理物體的效果不好。
- (3) Furukawa 方法中的特徵點群 (f, f') 只考慮兩兩間的極線幾何限制 (pairwise epipolar constraint)，所以初始所重建的 patches 會較疏離並可能有許多的錯誤。為了移除這些可能的錯誤對應，他們假設看得到該 patch 的影像，其相機光軸與 patch 的法向量之間的夾角必須小於 60 度。
- (4) Furukawa 的方法在決定 patch 的初始法向量時，是以 patch 的中心與投影至參考影像的連線做為其初始法向量，之後再利用 conjugate gradient 最佳化來疊代修正。但如果初始法向量與物體上的實際方向相差甚遠時，在最佳化過程中，可能就不會收斂。因此初始重建出的 patch 可能數量不多，將造成後續在做拓展 (expansion) 至鄰近物體表面時的困難。

由重建實驗成果來看，我們的大體上較平滑而準確。同時，我們實驗中處理了下列問題：

- (1) 解決 wide-baseline 所拍攝的視圖，其幾何變形較大，造成特徵點比對困難
- (2) 戶外景物 repeated image patterns 的比對混淆問題
- (3) 由於相機移動距離大，會拍到不同範圍的場景，如何界定同一座標系統，以免影響重建不一致性。

参考文献

- [1]. M. Goesele, B. Curless, and S. Seitz, "Multi-view stereo revisited," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2006, vol. 2, pp.2402-2409.
- [2]. S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. (2007). The Middlebury Computer Vision Pages – The Multi-view Stereo Page. Available: <http://vision.middlebury.edu/mview/>
- [3]. Y. Furukawa, and J. Ponce, "Accurate, Dense, and Robust Multi-view Stereopsis," in *Proc. IEEE Conf. Pattern Analysis and Machine Intelligence*, 2010.
- [4]. S. M. Seitz and C. R. Dyer, "Photorealistic scene reconstruction by voxel coloring," *Int. J. Computer Vision*, vol. 35, no.2, 1999, pp.151-173.
- [5]. K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," *Int. J. Computer Vision*, vol. 38, no.3, 2000, pp.199-218.
- [6]. K. Kutulakos, "Approximate N-view stereo," in *Proc. European Conf. on Computer Vision*, 2000, vol.1, pp.67-83.
- [7]. O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, "Real time correlation-based stereo: Algorithm, implementations and applications," Technical Report 2013, INRIA, 1993.
- [8]. C. R. Dyer, "Volumetric scene reconstruction from multiple views," in *Foundations of Image Understanding*, pp. 469-489. Kluwer, Boston, 2001.
- [9]. R. Szeliski, "Prediction error as a quality metric for motion and stereo," in *Proc. Int. Conf. on Computer Vision*, 1999, pp.781-788.
- [10]. C. Hernandez and F. Schmitt, "Silhouette and stereo fusion for 3D object modeling," *Computer Vision and Image Understanding*, vol.96, no.3, 2004, pp.367-392.
- [11]. M. Habbecke and L. Kobbelt, "Iterative Multi-View Plane Fitting," VMV, 2006.
- [12]. M. Habbecke and L. Kobbelt, "A surface-growing approach to multi-view stereo reconstruction," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.
- [13]. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Computer Vision*, vol. 60, no. 2, 2004, pp. 91-110.
- [14]. I. Gordon and D. G. Lowe, "Scene modeling, recognition and tracking with invariant image features," in *Int. Symp. Mixed and Augmented Reality*, 2004, pp. 110-119.
- [15]. D. Martinec and T. Pajdla, "3D reconstruction by gluing pair-wise Euclidean reconstructions, or how to achieve a good reconstruction from bad images," in *Int. Symp.*

- 3D Data Processing, Visualization and Transmission*, 2006, pp.25-32.
- [16].X. Zabulis and K. Daniilidis, “Multi-camera reconstruction based on surface normal estimation and best viewpoint selection,” in *Int. Symp. 3D Data Processing, Visualization and Transmission*, 2004, pp.733-740.
- [17].D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang, and M. Pollefeys, “Real-time plane sweeping stereo with multiple sweeping directions,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.
- [18].Yasutaka Furukawa, Brian Curless, Steven M. Seitz and Richard Szeliski, “Manhattan-world Stereo,” in CVPR 2009.
- [19].Cordelia Schmid and Andrew Zisserman, “Automatic Line Matching across Views,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 666-671, 1997.
- [20].J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” *Image and Vision Computing* , vol. 22, pp.761–767, 2004.

