

使用二維彩色圖像修補紅外線深度圖像

Repairing IR-depth Image with Use of 2D Color Image

余業緯、王祖鎧

Yeh-Wei Yu, Tzu-Kai Wang

現今流行的深度成像方法包括結構光 (structure light)、飛行時間 (time of light) 和視差立體圖像 (stereoscopic image)。由於任何點雲檢測方法皆可能產生深度資訊遺失, 因此有必要發展深度影像修補方法。由於現今三維彩色圖像的採集皆採用二維彩色圖像與深度圖像相結合, 因此, 我們提出了一種使用二維彩色圖像修補紅外線深度圖像之修補程序, 並提出使用迭代低通擴張法進行深度資訊修補, 具有通用且簡單之特性。我們在實驗中產生較強的紅外雜訊來破壞部分深度圖像, 然後利用新提出之方法成功地對深度圖像進行了修補。本研究亦建立模擬分析流程, 證明了迭代低通擴張法具有極好的準確度和精確度。

The most popular depth imaging methods including structure light, time-of-flight, and stereoscopic image. Since no point cloud detection method can prevent depth information loss, the complement methods is therefore important. Conventional 3D-image technologies produce RGBD images use 2D RGB image combing with depth image. Therefore, we proposed a depth information complement method that repair the IR-depth image with use of 2D RGB image. Besides, we propose a novel iterative low-pass pervasion method for depth image pervasion. In the experiment, we use strong IR noise to destroy parts of the depth image. Then we successfully complement the depth image by the proposed method. This research also builds a simulation analyzing process. It demonstrates that iterative low-pass pervasion method has good accuracy and precision.

一、背景介紹

當前的 3D 圖像技術使用 2D RGB 圖像與深度圖像組合產生 3D 彩色影像 (RGBD 圖像)。其中當前最流行的深度成像方法包括結構光 (structure light), 飛行時間 (time of light, TOF) 和視差立體影像 (stereoscopic image)。上述技術各有其適合之使用場景, 由於結構光和飛行時間技術都使用紅外投影光源, 當目標被太陽光或高強度紅外光照射時, 它們往往會遺失數據。此外, 對於結構光技術和立體圖像, 目標的劇烈深度變化將遮蔽照明光或圖像系統, 從而使數據遺失。表 1 顯示了不同深度信息檢測方法的深度信息遺失問題列表。由於沒

表 1. 不同深度資訊檢測法所面臨之深度資訊遺失議題，其中 ○ 代表優秀，△ 代表中等，× 代表拙劣⁽⁴⁾。

	Structure light	TOF	Stereoscopic image
照明不足	○	○	×
長距離物體	×	△	○
強烈紅外線場景	×	△	○
缺乏特徵點	○	○	×
三角陰影	×	○	×
鏡面物體	×	×	△
可穿透玻璃	×	×	×
回環錯誤	×	×	×

有點雲檢測 (point-cloud detection) 方法可以防止深度信息遺失，因此修補方法有其發展之必要性。

在 2018 年的 CVPR 研討會中，Yinda 成功使用 RGB 圖像與深度學習完成深度圖像 (depth image) 修補⁽¹⁾。在本文中，我們提出了使用 2D RGB 圖像作為精確 2D 框架的圖像補充過程，並且根據精確的 2D 框架修補 IR 深度圖像。與 Yinda 所做的研究不同之處，我們更關注單個物體的精確度，並進一步提出了能夠自行適應不同曲面類型的低通遞迴擴張技術。現今對於三維表面曲面擬合方式有許多先進的解決方案，例如局部加權回歸散點法 (locally weighted scatterplot smoothing, Lowess or Loess)⁽²⁾ 是一個針對散佈在空間中的採樣點做曲面擬合的演算法⁽³⁾，其結合最近鄰居法 (k nearest neighbor, KNN)，針對每一群集抓出趨勢，針對該群集內的點做權重分配，並改寫趨勢，直到其趨勢穩定，其輸出結果相當不錯，但由於此方法須考量所有點的權重並重新計算趨勢，導致運算繁雜，當點數量多且雜訊多時，結果很容易發散，且此方法在儲存時，需保留原始資料、分類標記、分類趨勢與權重，而非儲存方程式。在模型移植上製造了一定的空間複雜度。因而我們提出低通遞迴擴張技術，成為三維曲面擬合中一種通用且簡單的解決辦法。

二、點雲修補技術

3D 影像掃描器之三維空間資訊擷取模組輸出的檔案為 16 位元深度的深度影像，其像素的數值是深度值，基於針孔成像之相機模型，我們可以將深度影像做錐狀投影，投影至三維空間點雲 P_g ，其錐狀投影方式如圖 1，其轉換關係如下：

$$Z_{Global} = Z_{Global}/1000(\text{mm to m}) \quad (1)$$

$$X_{Global} = \frac{X_{Global}}{Z_p} \times Z_{Global} \quad (2)$$

$$Y_{Global} = \frac{Y_{Depth}}{Z_p} \times Z_{Global} \quad (3)$$

利用 2D 影像資訊修補不完整的 3D 點雲演算法之流程如圖 2：首先使用完整邊界偵測

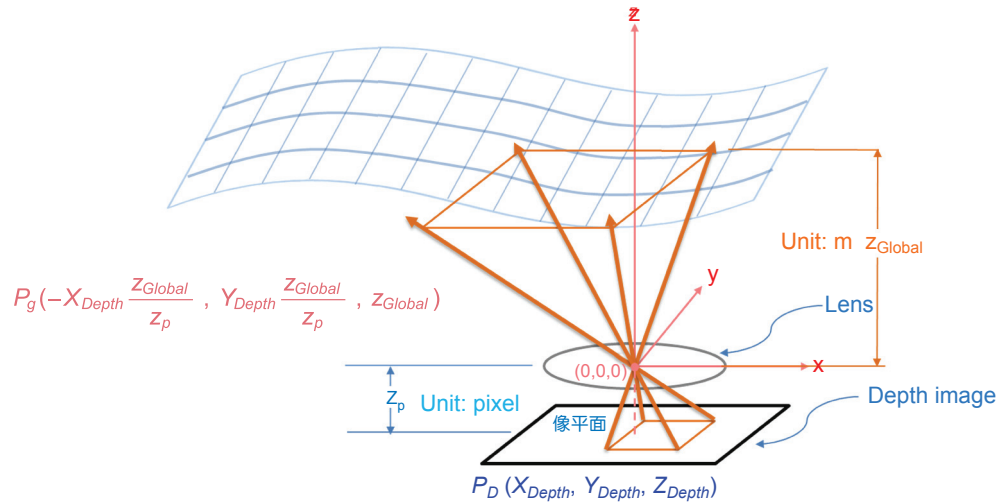


圖 1. 錐狀投影示意圖⁽⁴⁾。

(holistically-nested edge detection, HED)⁽⁵⁾ 對二維彩色影像作邊界偵測，抓出影像的準確邊界，並抓出該邊界 B_i 所屬面的影像座標資訊；接著使用點雲擴張技術進行修補，其修補完深度影像上具深度值 Z_{Depth} 的像素 P_p ，會被錐狀投影至三維空間點雲 P_g 。

遞迴低通擴張空間點雲之方式為：先宣告一運算矩陣 Z_{exp} ， Z_{exp} 大小為原本深度影像大小， $Deepmatrix(X'_{Depth}, Y'_{Depth})$ 由運算區塊給定，按照原深度影像位置填入 Z_{exp} 。低通遮罩為一圓型遮罩，半徑會隨著每次遞迴逐漸大，其目的是將該區塊深度趨勢逐漸掌握、逐漸將其他空間頻率之資料逐步納入與避免雜訊在深度趨勢尚未穩定時影響主訊號，一定程度上達成強化主訊號、降低雜訊影響之目的。此遞迴低通擴張空間點雲之停止條件為當平面趨勢趨近於平緩時，即跳出迴圈，完成空間點雲延伸，其趨勢判斷依據如下：

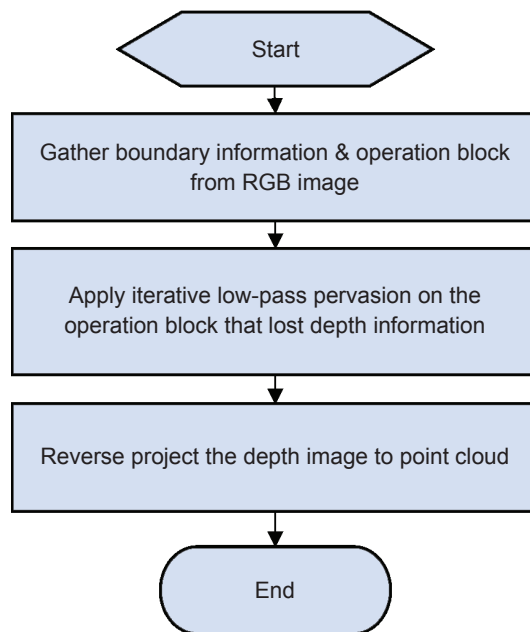


圖 2. 空間點雲修補流程⁽⁴⁾。

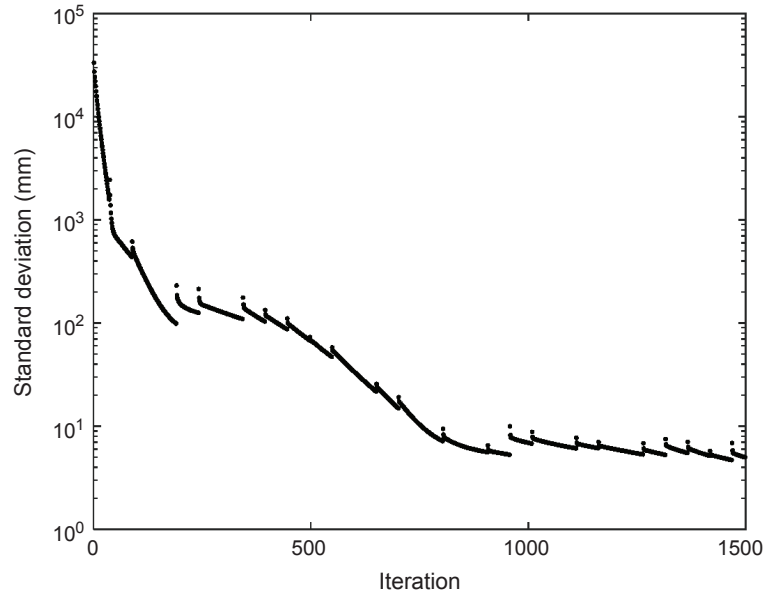


圖 3. 標準差 (對數座標軸) 對迴圈次數作圖。

$$\sigma_i - \sigma_{i-1} < S_{criteria} \parallel i < 1500, S_{criteria} = 0.005 \quad (4)$$

$$\sigma_i = \sqrt{\frac{\sum (P_i - P_{i-1})^2}{N}} \quad (5)$$

其中 σ_i 為與上一迴圈運算區塊內點雲資料之標準差，若與上個迴圈點雲差異不大， $\sigma_i - \sigma_{i-1}$ 小於 $S_{criteria}$ 後，即可跳出迴圈，完成低通擴張空間點雲。

從圖 3 可知，每當低通遮罩孔徑變大時，其標準差數值會震盪，並在後續迴圈收斂，其中 $\sigma_i - \sigma_{i-1}$ 為標準差對迴圈作圖之切線斜率，其數值可以判斷趨勢是否已經收斂，若收斂，則此數值會趨近於 0；受限於運算時間、空間限制，其迴圈數量必須限制與收斂條件必須放寬，其限制條件由多次試驗得出，在 1500 迴圈與斜率小於 0.005 條件下，大多數平面之變化已趨近於平緩。

三、實驗驗證

本研究分別使用市售 Kinect v1 3D 影像掃描器進行實驗，其具備二維彩色影像相機模組與三維空間資訊擷取模組，其輸出的檔案分別是 24 位元的彩色影像與 16 位元深度的深度影像，實驗採用的 3D 影像掃描器相關之資料庫 (OpenCV、Point Cloud Library) 內的座標轉換為當下量測之 3D 影像掃描器的數據。

圖 4 顯示了一個實驗場景，其中鹵素燈照亮了桌子並引起較大之紅外噪聲。Kinect V1 用於拍攝場景的 RGB 圖像和深度圖像。在深度圖像中，在桌面上發現了一個大洞。這是因為紅外噪聲比 Kinect V1 的紅外結構光強，並且使得深度信息無法檢測。

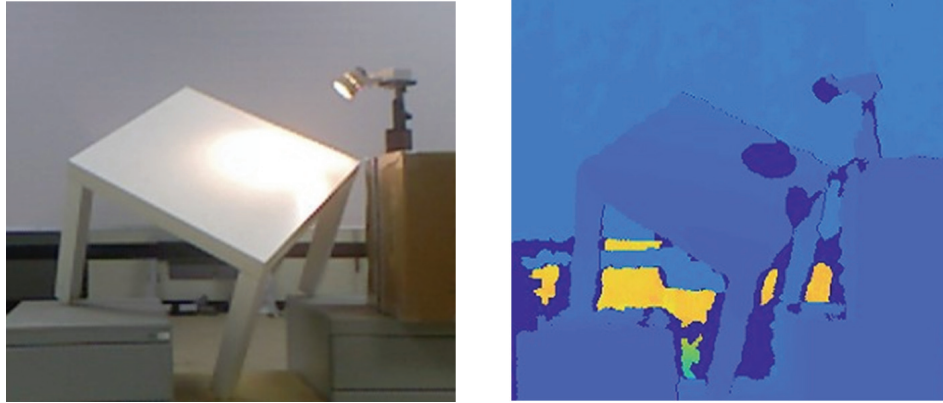


圖 4. Kinect V1 用於拍攝實驗場景的 RGB 圖像和深度圖像。其中我們使用鹵素燈照亮桌子並引起紅外 (雜訊)，造成深度資料遺失⁽⁴⁾。

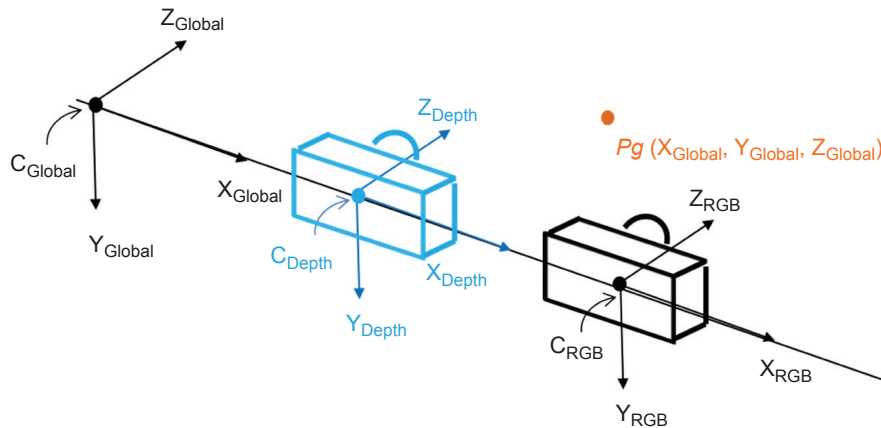


圖 5. 相機模型座標。

由於不同的 Kinect 有組裝上的誤差與深度影像的畸變問題，且官方釋出的校正資料庫對於兩張影像座標校正數據不一致⁽⁶⁾，因此需要對影像座標進行重新校正。我們首先修正二維彩色影像相機模組的畸變像差 (distortion)，畸變像差是三階像差，離軸影像會隨著場高產生較嚴重變形；其校正方式已經法展相當成熟，可使用的方式包含：棋盤方格校正雙眼 3D 視覺相機⁽⁷⁾、校正誤差補償機制⁽⁸⁾、基於已知間距棋盤方格生成空間相機座標之快速單一相機校正方法⁽⁹⁾、針對魚眼與多視角相機校正優化⁽¹⁰⁾，等等。

接著我們校正影像座標，圖 5 的相機模型定義兩相機鏡心座標與全場域座標。其中 C_{Global} 為空間點座標參考原點， C_{Depth} 與 C_{RGB} 分別為深度相機與二維彩色影像相機的鏡心， X_{Depth} 與 Y_{Depth} 為深度影像上之像素座標， Z_{Depth} 為深度影像上的深度數值， X_{RGB} 與 Y_{RGB} 為二維彩色影像之像素座標，由於二維影像不具有深度值，因此 Z_{RGB} 為一參考單位。

令 C_{Global} 位於 C_{RGB} ，以彩色相機鏡心當作參考座標，接著拍攝具有明顯邊界且有深度落差的場景，根據邊界明顯的色彩與深度變化，取得 C_{Depth} 與 C_{RGB} 之相對關係。藉由已知 C_{Global} 、 C_{Depth} 與 C_{RGB} 之相對關係，再透過下列式子得到正確的影像座標 X'_{Depth} 、 Y'_{Depth} 、 Z'_{Depth} ：



圖 6. 應用完整邊界偵測 (HED) 檢測圖像邊緣並從RGB圖像中收集邊界信息⁽⁴⁾。

$$X'_{Depth} = R \times X_{Global} + (Cx_{RGB} - Cx_{Depth}) \quad (6)$$

$$Y'_{Depth} = R \times Y_{Global} + (Cy_{RGB} - Cy_{Depth}) \quad (7)$$

$$Z'_{Depth} = R \times Z_{Global} + (Cz_{RGB} - Cz_{Depth}) \quad (8)$$

完成深度影像與 RGB 圖像的座標校正之後，我們便可接著進行圖 2 所顯示的影像修補步驟。我們先從 RGB 圖像收集邊界信息並挑選操作塊 (operation block)：應用完整邊界偵測 (HED) 檢測圖像邊緣並從 RGB 圖像中收集邊界信息⁽⁴⁾。然後，分離每個操作塊，並挑選桌面區塊做為需要被修補的操作塊 (圖 6)。接著，我們在桌面的操作塊上應用迭代低通擴張法進行深度圖像擴張。在經過 1500 輪迭代後完成擴張後完成操作塊內的深度資訊修補。最後，我們將深度圖像反轉投影到點雲。圖 7 顯示我們藉由圖像修補程序與迭代低通擴張法成功修補了在桌面操作塊上遺失的數據。

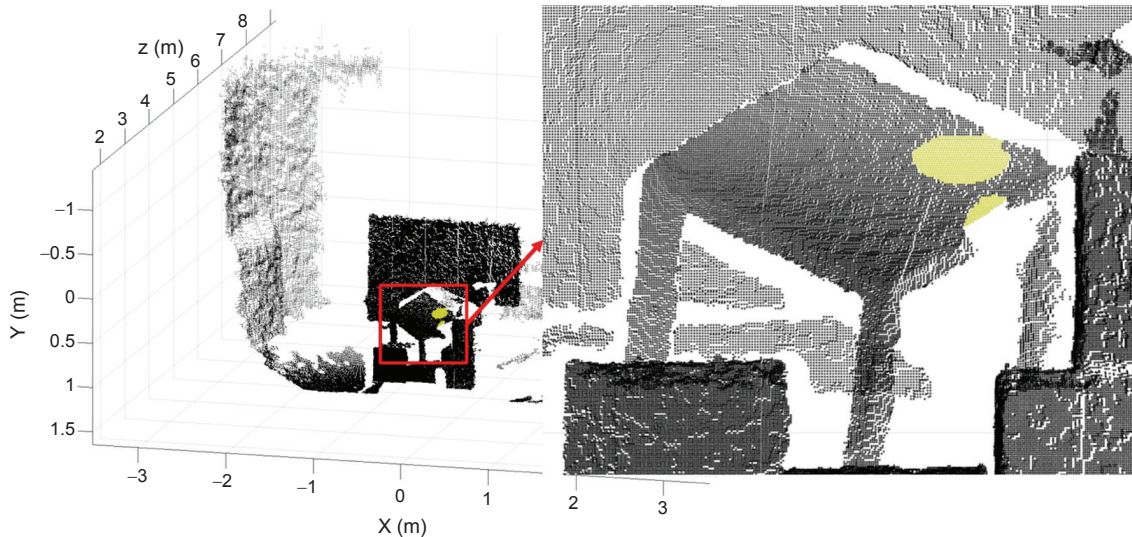


圖 7. 我們藉由圖像修補程序與迭代低通擴張法成功修補了在桌面操作塊上遺失的數據⁽⁴⁾。

四、迭代低通擴張法之模擬分析

為了驗證迭代低通擴張法之在曲面修補上之效能，我們模擬使用圓柱面做為被修補之曲面，評估迭代低通擴張法在不同資料缺失率下之準確性與精確性，並使用二次曲面擬合法作為對照組。二次曲面擬合法雖然無法通用於形狀多變之未知空間，對於本實驗中之圓柱面之樣本而言，卻為理論上最適合之擬合曲面，以其作為對照組，可以有效比對出迭代低通擴張法之效能。我們分別使用準確度與精確度進行重建影像品質之評價：

1. 準確度 (accuracy) 分析之定義為⁽¹⁰⁾：

$$\text{accuracy} = \frac{1}{N} \sum P_{fit_i} - \frac{1}{N} P_{original} \quad (9)$$

2. 精確度 (precision) 分析之定義為⁽⁵⁵⁾：

$$\text{precision} = \sqrt{\frac{1}{N} \left[\sum (P_{fit_i} - P_{original_i}) - \text{accuracy} \right]^2} \quad (10)$$

其中 P_{fit_i} 為單位像素延伸深度值， $P_{original_i}$ 為單位像素對應之原始資料深度值， N 為延伸深度總個數；準確度是分析與原始資料整體偏離量，精確度是分析相對於原始資料之資料離散程度，因此精確度數值越小表示與原始資料 $P_{original_i}$ 整體偏離較小，精確度越小表示與原始資料 $P_{original_i}$ 離散度相似。

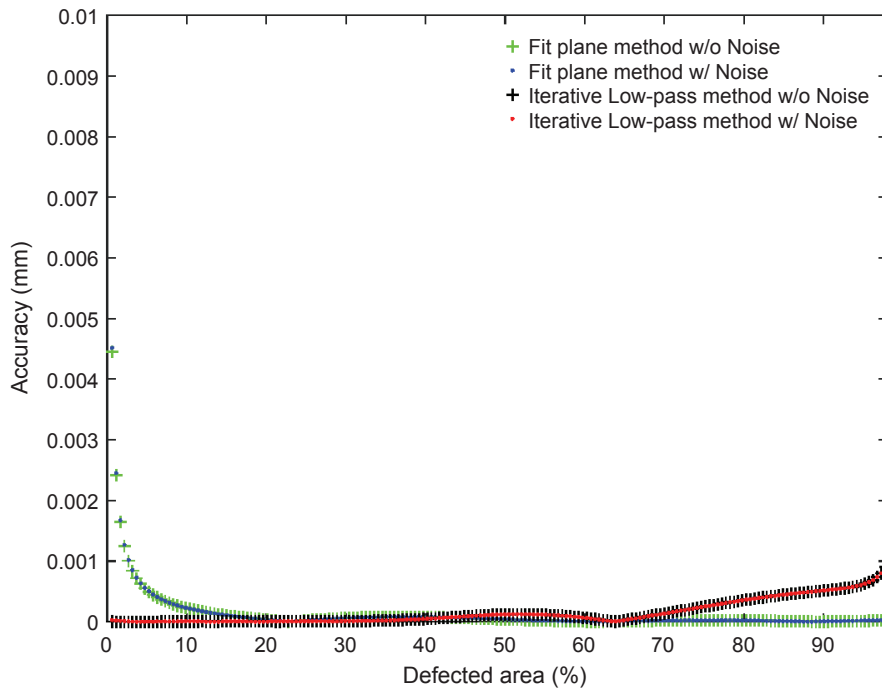


圖 8. 柱面深度影像刪除面積比率與準確度分析。

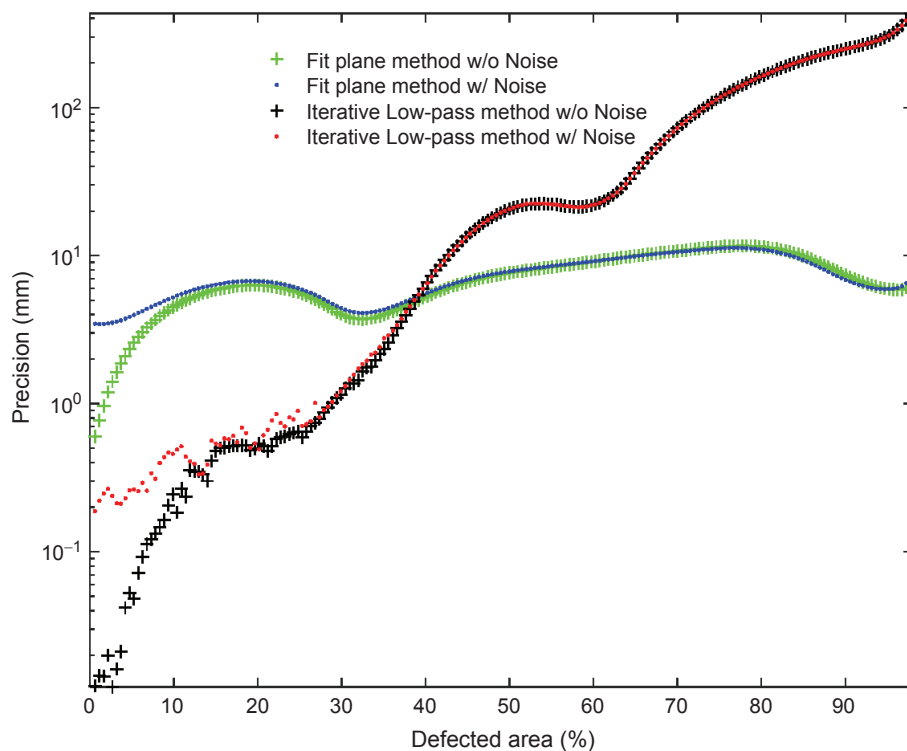


圖 9. 柱面深度影像刪除面積比率與精確度分析。

模擬分析中，我們首先生成一無雜訊之柱面深度影像與球面深度影像，大小為 540×960 pixels。皆著我們刪除部分深度影像，給定破損面積，每次刪除 $533 \times 5n$ pixels， n 為界於 1–190 自然數；我們也比較畫面存在雜訊時之修補效能，雜訊給定方式乃是基於 Kinect 之點雲雜訊量測結果，設定成分布範圍 ± 0.031 m、標準差 0.0065 m 之常態分佈雜訊，分別使用迭代低通擴張法與二次曲面擬合法進行曲面修補。

當圖面資訊遺失比例由 0% 趨近 98% 時，圖 8 為迭代低通擴張法 (iterative low-pass) 與二次曲面擬合法 (fit plane method) 之準確度曲線。可以看出其準確度皆小於 0.01 mm，而在 63.85% 破損率時，準確度有最小值，而大於此破損率之準確度有上升趨勢；由準確度分析可知，使用此兩種技術修補完成之點雲與原始點雲位置相當靠近。

當圖面資訊遺失比例由 0% 趨近 98% 時，圖 9 為迭代低通擴張法與二次曲面擬合法之精確度曲線，迭代低通擴張技術在破損面積大於 50%，精確度開始大於 25 mm，而二次曲面擬合則在 0–38.14% 破損面積之間，精確度皆大於迭代低通擴張技術，表示在低破損面積情況下，迭代低通擴張技術更貼近原始資料。而在高破損面積情況下，二次曲面擬合理所當然擁有較佳表現。

五、結論

當前的 3D 圖像技術產生 RGBD 圖像使用 2D RGB 圖像與深度圖像組合產生 3D 彩色影像技術，包括結構光，飛行時間和立體圖像等深度成像方法皆無法防止深度訊息遺失，因此深度訊息補充技術具有其存在價值。在此研究當中，我們提出了深度訊息修補技術。在第一

步中，我們應用完整邊界偵測法檢測圖像邊緣並從 RGB 圖像中收集邊界信息，進而分開操作塊並定義待修補操作塊。在第二步中，我們對遺失深度信息的操作塊應用迭代低通擴張方法，迭代低通擴張法相較於使用函數曲面擬合法而言具有極大的使用彈性，在使用前不需事先知道待修補曲面之函數特性。最後，我們將深度圖像反轉投影到點雲，完成修補。我們在驗證實驗中使用鹵素燈照亮桌子並引起嚴重的紅外噪聲。接著我們應用所提出的圖像修補程序與迭代低通擴張法成功修補了遺失的數據。為了測試迭代低通擴張法的修補效能，我們進行圓柱曲面樣本的模擬分析，模擬結果顯示，即使 50% 的圓柱表面深度信息被刪除，在使用迭代低通擴張方法後，我們仍能將修補的曲準確度降至低於 0.01 mm，精確度低於 25 mm。

誌謝

作者要感謝科技部計畫的資助 (106-3114-8-008 -001-; 104-2221-E-008 -078 -MY3; 105-2218-E-035 -009 -MY3)。

參考文獻

1. Y. Zhang, and T. Funkhouser, "Deep Depth Completion of a Single RGB-D Image." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 175 (2018).
2. W. S. Cleveland, *Journal of the American statistical association*, **74**, 829 (1979).
3. L. Wilkinson, "Statistics," in *The Grammar of Graphics* (Springer), 165 (1999).
4. Yeh-Wei Yu, Tzu-Kai Wang, Chi-Chung Lau, Jia-Ching Wang, Tsung-Hsun Yang, Jann-Long Chern, and Ching-Cherng Sun, "Repairing IR depth image with 2D RGB image," *Proc. SPIE 10745, Current Developments in Lens Design and Optical Engineering XIX*, 1074505, November 2 (2018).
5. S. Xie, and Z. Tu, "Holistically-nested edge detection," *Proceedings of the IEEE international conference on computer vision*, 1395 (2015).
6. D. Herrera, J. Kannala, and J. Heikkilä, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 2058 (2012).
7. Z. Zhang, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22** (11), 1330 (2000).
8. J. Heikkilä, and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Computer Vision and Pattern Recognition, Proceedings of IEEE Computer Society Conference on IEEE*, 1106 (1997).
9. D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on IEEE*, 5695 (2006).
10. S. Urban, J. Leitloff, and S. Hinz, *ISPRS Journal of Photogrammetry and Remote Sensing*, **108**, 72 (2015).
11. S. W. Smith, [The scientist and engineer's guide to digital signal processing], California Technical Pub., (1997).

作者簡介

余業緯先生為國立中央大學光電科學與工程博士，現為逢甲大學光電科學與工程學系助理教授。
Yeh-Wei Yu received his Ph.D. in the Department of Electrical Engineering from National Central University. He is currently an assistant professor in the Department of Photonics at Feng Chia University.

王祖鎧先生為國立中央大學光電科學與工程研究所碩士。
Tzu-Kai Wang received his M.S. in the Department of Optics and Photonics from National Central University.