

# 三維建物線框模型之無人機影像自動重建

黃郁翎<sup>1</sup> 莊子毅<sup>2\*</sup>

## 摘要

三維建物圖資在智慧城市規劃、管理和能源評估中扮演著重要角色。然而，由於作業繁瑣且自動化不足，針對既有建物構建精確的三維模型依然充滿挑戰。本研究提出基於多視角無人機影像的演算策略，生成具備側向幾何細節的三維線框模型，可做為進行既存建物三維房屋模型建置之基礎，提升作業效率並降低成本。演算程序運用預訓練的角點檢測模型及提出的角點萃取演算，採用「由粗到細」的策略實現角點定位。同時，運用虛擬角點重建策略來降低都市 UAV 影像中無可避免的遮蔽與數據缺失影響。實驗結果顯示，演算策略可適應具曲線形之建築結構，建築角點平均精度約為 30 cm，並可達到 98% 的線框重建完整度。

**關鍵詞：**多視角無人機影像、影像建模、自動化建物線框重建、線框模型、深度學習

## 1. 前言

### 1.1 研究動機與目的

三維空間資訊作為智慧國土的重要元素，可藉由豐富的時空數據協助決策者更全面地理解、規劃與利用環境資源。其中，三維建物模型更可扮演重要的資訊交換與分析平台，可運用於都市規劃、城市管理、能源估算及定位輔助等多個範疇。為實現智慧國土的目標，我國推行的三維國家底圖建置工作，採取「先全面建置，後精進細緻」的策略，基於既有的臺灣通用電子地圖，結合千分之一地形圖及數值地形模型以建構三維圖資。目前能提供符合開放地理空間協會 (Open Geospatial Consortium, OGC) 細緻度標準 (Level of Detail, LoD) 的 LoD1 與 LoD2 三維建物模型，對於重要地標則採用 LoD3 標準進行建置。由於建物模型品質會對後續分析和應用造成直接的影響，並恐於數據整合時產生衝突，因此，如何可靠且高效地產製高精度建物三維模型仍為關鍵研究課題。目前針對 LoD3 規格之三維建物模型製程仍缺乏自動化的作業程序，導致建置成

本高昂且耗時。有鑑於此，本研究採用無人機 (Unmanned Aerial Vehicle, UAV) 對目標建築進行約 45 度角的傾斜拍攝及 90 度角的垂直拍攝，獲取適合三維建築重建的影像數據。同時，本研究以此多視角影像作為唯一的數據類型，提出結合深度學習演算法的自動化處理策略，建置具有側向幾何資訊的建物三維線框模型，可作為既存建物 LoD3 模型建置的基礎模型，在符合精度需求的前提下，獲得降低人工與時間成本的效益。

現行的重建策略多採取結合不同來源與類型的數據，以資訊互補的方式逐步重建房屋模型。然而，不同類型數據的整合往往需要額外處理以妥善和應對數據品質及收集時間的差異性。此外，相較於運用三維點雲為基礎的演算方式，影像資料在資料量、技術發展成熟度及演算複雜度方面相對具有優勢，因此，許多研究著重於運用衛星及航空影像進行大規模的房屋模型重建，但受到影像視角的限制，模型成果無法提供建物的側向幾何資訊，難以滿足高精度和高細緻度的需求。反之，無人機具備高度機動性與作業彈性，並可採用垂直或傾斜拍攝配置提供高解析度影像。儘管如此，實際作業仍不可避

<sup>1</sup> 國立陽明交通大學土木工程學系 碩士

<sup>2</sup> 國立陽明交通大學土木工程學系 副教授

\* 通訊作者, E-mail: jtychuang@nycu.edu.tw

收到日期：民國 113 年 08 月 14 日

修改日期：民國 113 年 09 月 09 日

接受日期：民國 113 年 09 月 11 日

免遭遇目標物被遮蔽之挑戰。如何針對遮蔽區域進行補償以提升重建完整性亦是待改善之課題。

本研究旨在提高 LoD3 建物模型重建任務時的自動化程度，提出基於深度學習的自動化三維線框模型建置演算策略，以此作為輔助 LoD3 建物建模任務的基礎模型，減少人工在幾何判識或檢核操作中的介入。受益於深度學習技術的發展，本研究利用深度學習模型自動偵測建物角點區域，並結合相應的角點萃取演算法，採用「由粗到細」的方式精確定位建物角點位置。同時，透過影像特徵描述元進行特徵點描述並進行多像幅的特徵匹配程序。同時，透過多像幅之共軛點光束交會，估計建物各角點之三維坐標。除此之外，本研究進一步提出了針對描述建物外觀幾何的三維角點的位相關係重建策略，恢復建物外觀的線面關係，從而形成三維線框模型。值得注意的是，為了解決運用無人機影像進行建物重建任務經常遇到的目標遮蔽問題，研究中提出了虛擬角點重建演算程序，重建真實存在於空間中但因為遮蔽或重建程序失敗等因素而缺失之建築物角點。研究假設在建物僅有部分角點缺失的情況下，基於曼哈頓世界假設 (Manhattan World Assumption)，以虛擬角點的形式補全線框模型中的缺失點位，針對結構物主體生成符合 LoD3 精度標準且具側向幾何資訊的三維線框模型。與既有方法相比，本研究具有以下特點：

- (1) 結合深度學習，自動化偵測建物角點特徵，並設計了精確定位角點的演算法。
- (2) 提出角點位相關係重建策略，包括建置虛擬點位，以降低目標物遮蔽及自動化重建過程失敗所導致的缺失。
- (3) 以無人機多視角影像作為唯一數據資料，實現三維建物線框模型的自動化重建演算，生成具側向幾何資訊如陽台、雨遮或曲面特徵的高精度建物線框模型。

## 1.2 相關研究

### 1.2.1 自動化建物模型重建策略

本研究的主要目標是結合深度學習模型，自動

化重建具側面幾何資訊的三維建物線框模型。該線框模型將作為 LoD3 建物模型建置過程中的基礎模型，從而提升其自動化水平。為達成此目標，本研究將重點回顧具側向結構的建物模型重建方法，並深入分析各現有方法的優缺點，以更全面理解在實現自動化重建目標過程中可能面臨的挑戰。

城市地理標記語言 (City Geographic Markup Language, CityGML) 對於各級細緻度的建物模型規範了其精度和簡化程度。在 LoD1 中，平面與高程的絕對精度要求為 5 m，並將平面大於 36 m<sup>2</sup> 和高程超過 3 m 的模型進行簡化，未達標準者可不予考量。LoD2 則需達到平面與高程的絕對精度 2 m，並重建平面面積大於 16 m<sup>2</sup> 以及高程超過 2 m 的特徵。LoD3 的精度要求更為嚴格，平面與高程的絕對精度需小於 0.5 m，並處理平面面積大於 4 m<sup>2</sup> 以及高程超過 1 m 的特徵。LoD4 則需達到平面與高程的絕對精度 0.2 m，並呈現所有結構和開口的細節。這些精度和細緻度的分類與定義提供了關鍵標準，使建物模型能夠更有效地共享、比較和分析，確保模型的精度和一致性。儘管本研究並非直接重建 CityGML 規範下的建物模型，這些標準仍可作為線框模型品質的參考依據。

現行建物模型重建的方法可分為兩大方式：模型導向 (Model-driven) 和資料導向 (Data-driven)。模型導向方法基於資料庫，並根據調整的參數進行建模。此方法對於幾何形狀簡單的建築物 (如平頂屋或山形屋) 非常有效。然而，由於現實世界中的建築外型多樣，資料庫難以涵蓋所有可能的建築形式，造成此方法具有一定的局限性。相較之下，資料導向方法基於建物表面及其位相約制進行重建，更適合處理複雜的細節。其處理流程通常包括三個主要步驟：萃取建物的原始表面特徵、重建建物表面的位相關係，以及構建建物模型 (Ma, 2005)。理論上，資料導向方法能更全面地重建現實世界中各種形狀的建築物。根據處理的數據資料類型，可進一步細分為雷射掃描資料、影像資料和地圖資料。雷射掃描資料主要來自空載雷射掃描儀 (Airborne Laser Scanner, ALS) 和地面雷射掃描儀 (Terrestrial Laser Scanner, TLS)，以點雲形式儲存。相似地，影

像資料根據載具的不同可分為衛星影像、航照影像和無人機影像。經過處理後，影像資料同樣能夠生成點雲資料。不同資料來源各具優缺點，如表 1 所示。為了彌補各自的不足，許多研究採用多種資料來源進行重建 (Flamanc *et al.*, 2003、Gui & Qin, 2021、Wen *et al.*, 2019)。本研究嘗試以無人機影像作為唯一的數據來源，以獨棟建築為目標進行三維線框模型重建。由於無人機具有以高彈性、機動性獲取高解析度影像的優勢，便於掌握和控制所獲得的影像資料，尤其有利於取得側向結構資訊 (Klimkowska *et al.*, 2022) 進行高精度的建物線框模型重建 (Malihi *et al.*, 2016)。

透過整合不同類型資料的優勢，Wen *et al.* (2019) 將點雲萃取的平面特徵與傾斜影像中的邊界約束相結合，實現半自動化的高精度建物模型重建。Chen *et al.* (2014) 則使用 ALS 點雲作為資料來源，採用改良的隨機抽樣一致算法 (RANDOM SAMPLE CONSENSUS, RANSAC) 來分割屋頂平面，並在約束條件下萃取邊緣以構建位相關係，進而優化邊界形狀，自動生成大範圍的建物模型。與之不同，Wang *et al.* (2016b) 聚焦於建物立面結構的重建，通過點雲資料提取立面結構的幾何特徵和約束條件，能快

速且大規模地建置建物立面模型，然而，此方法僅適用於網格狀的基礎立面。此外，Pu & Vosselman (2009a) 同樣以立面結構為重點，透過先驗資訊重建受遮蔽區域，成功克服了資料不足的挑戰。

鑒於上述方法成本較為高昂，為了降低成本並同時獲取高精度資料，近年來無人機影像資料逐漸受到重視 (Murtiyoso *et al.*, 2020)。其中，由於重建細節與獲取坐標資訊的需求，多數研究偏向使用基於影像生成的點雲作為資料來源。McClune *et al.* (2016) 結合正射影像、數值地表模型與影像生成的點雲，通過邊緣線偵測與線擬合，萃取屋頂角點，並利用位相關係完成 LoD2 精度的房屋線框模型自動化重建。然而在處理複雜屋頂幾何時仍有進步空間。Wang *et al.* (2023a) 基於點雲資料與從中萃取的屋頂輪廓，通過演算法將立面抽象化為規則線段，進行二維型態的重建並拉升為三維模型。除此之外，Huang *et al.* (2020) 則是將無人機影像生成之點雲與預設模型資料庫進行整合，運用模型導向的方式完成 LoD3 建物模型重建。近年伴隨著人工智慧技術的發展，三維建物模型重建引入了全新的思維與方法，不僅顯著減少了人工作業的介入，還克服了傳統演算法所面臨的諸多挑戰，大幅提升了資料處

表 1 不同資料類型於建物模型建置之優缺點

資料類型	優點	缺點
雷射掃描	<ol style="list-style-type: none"> <li>1. 高精度的幾何資料，尤以地面雷射掃描表現最佳</li> <li>2. 捕捉建築物細節和結構</li> <li>3. 適用大範圍和複雜環境</li> </ol>	<ol style="list-style-type: none"> <li>1. 成本較高，特別是空載雷射掃描</li> <li>2. 需專業設備和技術</li> <li>3. 處理和分析大量點雲資料可能耗時較長</li> </ol>
影像資料	<p>衛星影像</p> <ol style="list-style-type: none"> <li>1. 覆蓋範圍最廣</li> <li>2. 定期更新</li> <li>3. 長時間序列</li> </ol>	<ol style="list-style-type: none"> <li>1. 相對於其他影像，解析度較低</li> <li>2. 受雲層影像</li> <li>3. 有限的高程感測</li> </ol>
	<p>航照影像</p> <ol style="list-style-type: none"> <li>1. 解析度適中</li> <li>2. 相對於衛星影像，受天氣、雲層影像較小</li> <li>3. 可控制航高</li> <li>4. 覆蓋範圍次之</li> </ol>	<ol style="list-style-type: none"> <li>1. 局部覆蓋</li> <li>2. 成本較高</li> <li>3. 影像更新時間不穩定</li> </ol>
	<p>UAV 影像</p> <ol style="list-style-type: none"> <li>1. 高解析度</li> <li>2. 具高彈性與高機動性</li> <li>3. 即時性</li> <li>4. 較易獲取側向結構與細節</li> </ol>	<ol style="list-style-type: none"> <li>1. 飛行時間有限</li> <li>2. 覆蓋範圍小</li> <li>3. 法規限制</li> </ol>
地圖資料	<ol style="list-style-type: none"> <li>1. 提供地理和地形資訊</li> <li>2. 適合用於城市規劃和區域規劃</li> </ol>	<ol style="list-style-type: none"> <li>1. 缺乏建築物具體細節</li> <li>2. 地圖更新速度較慢</li> <li>3. 有限的結構資訊</li> </ol>

理的自動化程度。Pantoja-Rosero *et al.* (2022) 運用了運動恢復結構 (Structure from Motion, SfM) 技術從影像中生成點雲，並利用這些點雲擬合出建物的平面結構。通過深度學習技術進行語意分割 (Semantic Segmentation)，自動分離出建物的開口處，如門窗，以實現 LoD3 建物模型的重建，成果如圖 1 所示。然而，其成果顯示山形屋頂的滴水線結構未能成功重建。此外，鑒於建物立面結構多數具有高度對稱性，如門窗等矩形結構，Liu *et al.* (2020) 採用了深度學習模型，並結合適用於高度對稱矩形結構的損失函數，成功提高了深度學習模型的預測精度，從近景影像中分割出立面中的元素。



圖 1 結合點雲平面擬合與語意分割技術所建置之 LoD3 建物模型 (Pantoja-Rosero *et al.*, 2022)

此外，運用衛星、航照影像或無人機影像進行建物分割的研究在大範圍的影像處理技術上亦獲得顯著的成果 (Buyukdemircioglu *et al.*, 2022)。衛星與航照影像因涵蓋範圍廣泛，適用於大規模地重建 LoD1 與 LoD2 的建物模型。多數研究往往透過萃取建物屋頂的外框輪廓與多邊形輪廓，結合立體像對解算所得的屋頂高程與數值表面模型 (Digital Surface Model, DSM) 數據，藉以進行建物模型的抬升重建 (Yu *et al.*, 2021、Zhang *et al.*, 2022a)。其中，全卷積網路 (Fully Convolutional Network, FCN) 因其適用於影像分割而廣泛應用。透過 FCN 架構，輸入影像的每個像素皆可產生相應的輸出，應用於建物分割時，能夠精確描繪建物的邊緣。利用專為建物分割設計的 MA-FCN 架構，結合正規化演算法來處理初始多邊形，可大規模且高效地從航照影像中萃取結構化的建物外框 (Wei *et al.*, 2020)，如圖 2 所示。此方法可處理矩形及圓形幾何，展現出較高的適應性。然而，大範圍的高解析度影像往往包含大

量的遮蔽物與干擾物，例如樹木、交通工具或陰影等，這些因素會增加建物分割的難度。為此，Abdollahi *et al.* (2020) 利用生成對抗網路 (Generative Adversarial Network, GAN) 模型，結合 SegNet 的編解碼架構，提升在複雜背景下的分割效能，增強模型抵抗雜訊的能力。此外，Schuegraf & Bittner (2019) 應用 U-Net 架構，訓練出可同時處理多光譜與 DSM 資訊的建物外框萃取模型，即便在低解析度的光譜影像中仍可獲得不錯的性能。然而，由於資料拍攝視角的限制，研究多僅能產製 LoD2 之建物模型。

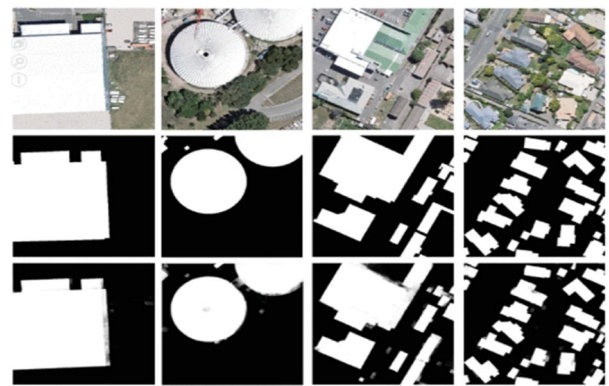


圖 2 利用影像分割結合正規化演算法萃取航空影像中的建物外框 (Schuegraf & Bittner, 2019)。其中，第一列為原始影像；第二列為標記影像 (真值)；第三列為萃成果

## 1.2.2 無人機影像建物模型重建

影像建模首先需要對大量高重疊影像進行影像外方位解算，此過程包括相機姿態估計、影像特徵點萃取、特徵匹配以及方位解算，最終生成密集點雲數據。為獲得良好的資料處理成果，必須仰賴影像中準確可靠的幾何特徵，然而，在存在遮蔽物、光照變化或缺乏紋理的情況下，可能導致解算失敗。此外，獲取精確的影像外方位參數初始值也是影像建模中的一項挑戰 (Aharchi & Ait Kbir, 2020)。近年來，運動恢復結構 (Structure from Motion, SfM) 技術作為一種類似的資料處理方法，能夠從大量無序影像中自動恢復場景的幾何結構。SfM 強調線性解算步驟，並具有與攝影測量方法相似的資料處理步驟，包括特徵點萃取、八點演算法以及結合 RANSAC 的基本矩陣估計等 (Wang, 2013)。但是



SfM 同樣依賴影像中的特徵資訊進行解算，因此會仍面臨相同的資料處理挑戰。

現行的影像房屋模型重建策略除了針對影像生成之點雲進行處理外，直接從影像中萃取角點和邊緣線對也是常見的方法 (Faugeras *et al.*, 1998、Pu & Vosselman, 2009b)。其中，Wang *et al.* (2023b) 利用 Mean Shift 演算法對影像進行內容分割，以獲取建物範圍，並使用尺度不變特徵轉換 (Scale-Invariant Feature Transform, SIFT) 定位角點周圍，進而萃取出建物的近似邊界，最後使用霍夫轉換優化線段成果。在邊緣線偵測方面，Zhu *et al.* (2015) 將傾斜影像作為輸入數據可以獲取屋頂甚至立面的邊線資訊。通過從影像生成網格數據並萃取邊界線，結合平行、正交與對稱等幾何特性，構建線段間的位相關係，如圖 3 所示。將影像中偵測到的二維線段投影至三維模型中，進一步精煉線段結構，以處理遮陽板、滴水線等外凸區域。高重疊影像的優勢在於影像對之間具有較高的紋理相似性，可提高影像匹配的成功率，並提供更多交會觀測，有助於生成高精度數據。此外，Rau & Chen (2003) 利用建物邊緣線重建三維建物模型，提出了「分割—合併—成型」(Split-Merge-Shape, SMS) 演算法。該演算法通過將現有建物線段分割為多個房屋單元，並根據位向關係判斷是否需要進一步合併，最後對共面的部分進行處理，完成建物模型的成型。此方法能有效且快速地重建線段間的位相關係，其特點在於能處理不完整的線段，對於重建受遮蔽的建物具有一定的穩健性。

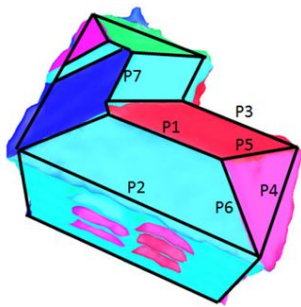


圖 3 建物外框線的位相精煉規則，包括平行、正交、對稱及高度相等條件 (Zhu *et al.*, 2015)

### 1.2.3 建物側向幾何重建

有別於 LoD2 建物模型僅需重建屋頂面結構，在實現更高細緻度與真實性的結構重建中，側向幾何尤其具有挑戰性。目前著重於側向結構的建物模型重建策略，仍然多依賴於多種遙測數據和載具平台的整合，藉以獲取優勢互補的數據資料進行建物模型重建。以點雲作為基礎，可透過整合 ATS 與 TLS 點雲資料補足資料幾何缺失部分。面對側向障礙與遮蔽的挑戰，仍多採用人工精練的濾除程序。相比之下，利用多視角無人機影像生成的點雲可以以較低成本獲取相對完整的幾何細節資訊，通過點雲聚類和特定幾何條件，可以重建特定性質的側向結構，如屋簷、屋頂窗和窗戶等。此外，Wang *et al.* (2016b) 利用模型資料庫作為先驗知識，結合語意分割技術對點雲進行分類，並進行模型導向建物模型重建。這類方法能在屋頂面和側面生成視覺化效果出色的模型，然而，細節結構將受限於資料庫內容。點雲和影像資料各有其優缺點，混合使用可以產生互補效果。例如，將點雲資料與傾斜影像結合使用，可彌補點雲在低紋理結構中表現不佳的問題，並且影像還可提供幾何形狀的約束條件。若進一步利用光譜影像與立面紋理特徵，可更精確地偵測出建物和開口位置 (Gruen *et al.*, 2019)。例如，Zhang *et al.* (2022b) 以既有的含紋理之 LoD2 建物模型作為輸入，針對紋理中的窗戶結構進行重建。該方法使用深度學習模型進行語意分割，對窗戶進行分群與對齊，並以幾何特性進行優化，最終將結果整合回原先的建物模型。然而，該方法目前僅能處理窗戶結構。而 Gadde *et al.* (2018) 開發了應用於影像資料與點雲的立面分割方法。類似研究除了常用的語意分割與聚類技術外，亦引入幾何約束規則於立面重建 (Mesolongitis & Stamos, 2012、Neuhausen & König, 2018)。

### 1.2.4 小結

綜上所述，本研究旨在以多視角無人機影像作為唯一資料來源，實現單棟建物的三維線框模型自動化重建，重點在於精確重建側向幾何資訊。研究

中採用學習式物件偵測模型來自動偵測建物的角點，並作為建物角點準確定位的初始資訊。此外，考量影像資料在實際收集過程中可能遇到的建物可視性限制及資料處理中的不合格角點，研究提出基於曼哈頓世界假設的虛擬角點重建演算法，以應對資料缺失問題。值得注意的是，此研究旨在基於低作業成本的條件下，僅使用無人機影像進行建物三維線框模型的自動重建，並確保線框模型滿足 CityGML 的 LoD3 精度規範，為後續的加值應用提供基礎模型。

## 2. 研究方法與架構

本研究提出的演算策略包含五個主要處理程序，資料前處理、角點特徵偵測模型、角點萃取演算法、角點三維坐標重建及位相關係重建。資料前處理程序乃通過影像方位解算來獲取影像的內、外方位參數，使用的資料包括垂直及傾斜拍攝的多視角無人機影像，為後續的空間坐標解算奠定基礎。角點特徵偵測模型旨在自動偵測建物角點位置區域，將角點特徵限制於特徵邊界框內，提高後續角點萃取的精度。後續的角點萃取演算法則是基於標界框中的角點特徵，運用提出的角點萃取演算法，自動定位角點的影像坐標。待經過多像幅的角點特徵匹配後使用多光束交會計算建物角點的三維坐標。最後，位相關係重建程序乃根據影像和幾何條件，建立角點之間的點線面關係，即位相關係。同時，為解決角點缺失問題，研究提出虛擬角點重建演算法，自動建置缺失的角點，以確保線框模型的合理性。演算程序最終將產生一組完整的建物三維線框模型，為後續應用提供基礎模型。圖 4 呈現本研究的資料演算流程。

### 2.1 資料前處理

本研究採用無人機影像作為數據來源，分為垂直拍攝與傾斜拍攝兩類資料。無人機的優勢包括低成本以及對氣象變化的高機動性，且普遍配備了直接定位系統，如衛星導航系統 (Global Navigation Satellite System, GNSS) 和慣性測量單元 (Inertial Measurement Unit, IMU)，能提供載具一定程度的三

維坐標和姿態初始值。為了獲得更可靠的方位參數，本研究利用立體視覺技術進行序列影像方位估計，並結合地面控制點，儘可能修正原始偏差，從而優化影像方位解算結果。

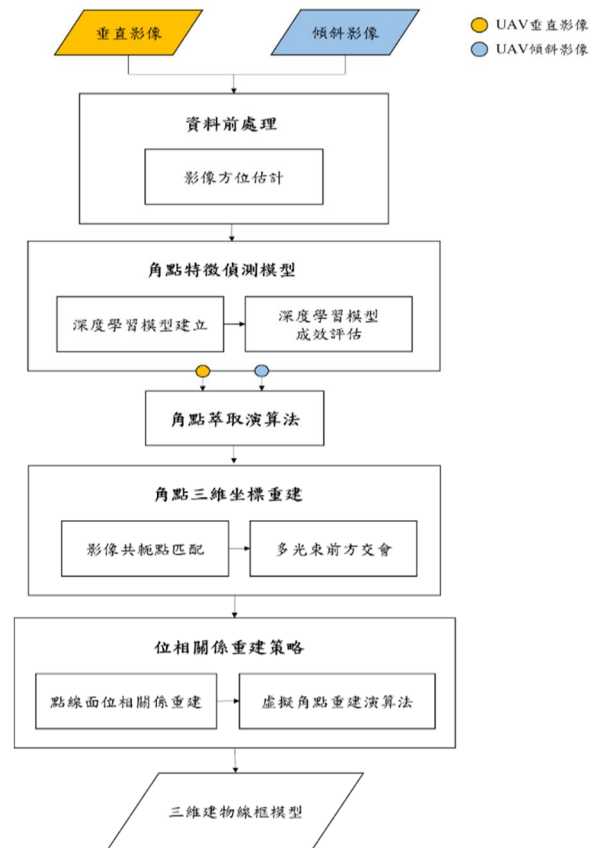


圖 4 演算流程架構圖

### 2.2 角點特徵偵測模型

影像辨識技術在電腦視覺領域取得了顯著突破，並持續有新演算法和架構被提出。著名的深度學習模型包括 Faster R-CNN、YOLO、SSD、Co-DETR 以及 EVA 等。本研究選擇深度學習影像物件辨識技術，以實現建物特徵角點區塊的自動化偵測，並致力於建立一個兼具精度、運算效率與穩健性的模型架構。因此，本研究將採用 YOLOv4 (Bochkovski *et al.*, 2020) 架構進行進一步的訓練和處理。值得注意的是，本研究不涉及對不同演算法的效能比較，所採用的模型可被更先進的演算法所替代。本研究採用微軟建置的 Common Objects in Context (COCO) 數據集進行模型預訓練，作為初始權重，旨在排除基礎常見物件的干擾，並提高模型

對於自定義物件的識別能力。在模型訓練之前，對收集的無人機影像資料進行了標註。本研究的深度學習模型訓練專注於建物角點特徵區塊的偵測，與大多數以整棟建築物為目標的深度學習模型不同。本研究將角點幾何轉折處作為邊界框的中心，如圖 5 所示，以便準確定位角點特徵區塊。

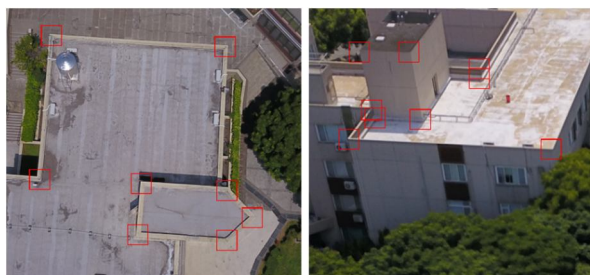


圖 5 建物角點特徵區塊作為標的物示意圖

深度學習模型的訓練過程中，使用了 1511 張 UAV 影像資料，共標註了約 19,082 個建物角點特徵區塊的邊界框，包括不同角度的重複角點特徵。為了確保模型的泛化能力，訓練、測試及驗證資料按 8:1:1 的比例進行分配。為統一影像尺寸並減少硬體負荷，影像被縮放至  $608 \times 608$  像素。此外，考慮到航拍過程已規劃多視角拍攝，無需對影像進行旋轉處理，但為了增加資料的多樣性，採用資料擴增技術對影像的飽和度、曝光度及色調進行隨機調整。經過多次實驗，確定了 YOLOv4 模型的局部最佳超參數設定。在記憶體限制內，每批次 (Batch) 大小設為 8，使用二元交叉熵(Binary Cross-Entropy) 和完全交併度 (Complete IoU, CIoU) 作為損失函數，分別計算分類、可信度和位置的誤差。在誤差反向傳播 (Backward Propagation) 過程中，採用 SGD 優化器進行權重與偏差的調整，動量 (Momentum) 和衰減率 (Decay) 分別設定為 0.949 和 0.0005，初始學習率 (Learning Rate) 設為 0.001。為避免過度擬合 (Overfitting)，在第 80 和 90 個 Epoch 時，學習率將以 0.1 的比率進行衰減。本研究針對深度學習模型成效進行評估，使用三種常用於深度學習物件偵測的指標作為評估標準，包括陽性預測率 (Positive Predictive Value, PPV) 或稱精確率 (Precision)、召回率 (Recall) 及平均精確率 (Average Precision, AP)。模型性能評估基於預先分

類的測試資料作為真值，主要目的是評估角點特徵偵測在建物角點特徵區塊中的位置精度及物件類別分類的品質。此外，本研究將針對不同的 IoU (Intersection over Union) 閾值進行 AP 的評測，以全面評估模型在各種情境下的偵測效能。

## 2.3 角點萃取演算法

儘管影像處理領域已有多種基於影像灰階梯度變化的角點偵測元於。然而，由於缺乏幾何約束，這類演算法通常會偵測出多個點，難以將其準確定位於單一點上，難以滿足本研究的需求。為解決此問題，本研究引入了影像線段特徵萃取技術，並提出一套新的角點萃取演算法，演算法通過幾何約束條件來萃取出最佳的建物外邊角點位置，專注於從角點特徵區塊中確定唯一的建物角點幾何轉折處。演算流程架構如圖 6 所示。

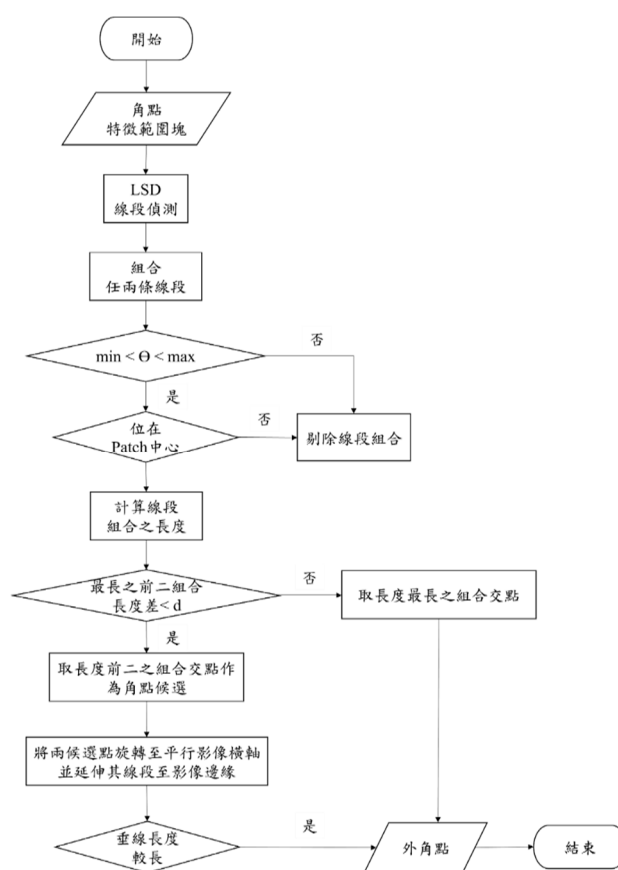


圖 6 角點萃取演算法之演算架構

角點萃取的演算策略中，利用深度學習角點特徵偵測模型所預測的角點區域作為偵測單位，進行



精細化處理。圖 7 與圖 8 分別展示了角點萃取演算法的示意圖與相應的虛擬碼。圖 7(a) 部分顯示了經由 Line Segment Detector (LSD) (Von Gioi *et al.*, 2012) 線段偵測元所萃取出線型特徵；圖 7(b) 部分展示了在角點區域中心(即黃色區域)中，選取夾角在核可範圍內的最長兩組線段組合，並由此交會出角點候選點，其中  $w$  為女兒牆寬度；圖 7(c) 部分則示意將這兩個候選點旋轉至平行於影像橫軸，並進行垂線長度的比較，選擇垂線較長者作為最終角點；圖 7(d) 部分展示了最終角點的示意圖，最終獲得的外側角點為單一像元。

使用 LSD 線段偵測元提取角點區域周圍的線型特徵。針對篩選出的線型特徵，進行兩兩交會並計算其之間的夾角。這些夾角必須符合預設的閾值才能保留該組線段組合。角點的閾值設置在垂直影像和傾斜影像中有所不同。在垂直影像中，大部分建物的角點夾角接近九十度，而傾斜影像則需要使用較寬的閾值進行篩選。此閾值為可調變參數，可根據待重建的建物型態進行設定。由於深度學習模型預測的角點特徵範圍應以角點轉折處為中心，

本研究將角點位置限縮於影像中心範圍內。如果線段組合所形成的角度符合規範，但其交會點不在中心範圍內，則會剔除該組合。此外，考慮到某些建物屋頂具有女兒牆結構，本演算策略針對內外牆部分進行判斷。研究主要關注整體建物的線框模型，並基於外側角點進行建構。在演算中，將保留外側角點並捨棄內部角點。在方法上，將兩組最長的線段組合的各別交會點視為候選點，並將這兩個候選點旋轉至平行於影像橫軸。隨後，從影像邊緣向候選點垂直投射，並依據外角的垂線比內角長的原則來判斷出外側角點。然而，角點萃取演算法在處理內凹結構時仍存在侷限性，難以準確獲得外側角點。針對內凹的女兒牆外牆結構，本研究提出了一種特殊的角點標籤方法以克服此問題，如圖 9 所示。具體而言，此特殊標籤方法並不以建物轉折處為中心，而是在標籤過程中，將女兒牆內側角點置於角點萃取演算法中的位置核可範圍外。透過這種方法，限縮了角點特徵區域的範圍，確保角點萃取演算法在處理過程中，不會將內側角點歸納為候選點。

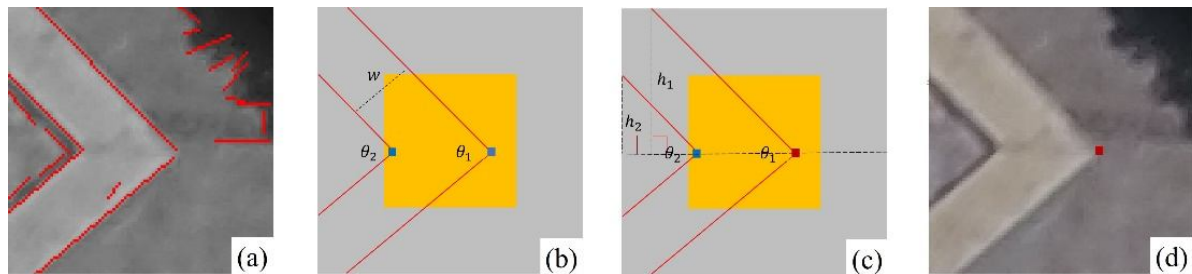


圖 7 角點萃取演算法之示意圖

#### Algorithm Corner Extraction Algorithm

```

1: Function extractCorners(image)
2:   cornerFeatures = deepLearningCornerDetection(image)
3:   lineSegments = LSD(image)
4:   for each segment1 in lineSegments
5:     for each segment2 in lineSegments
6:       intersectionPoint = calculateIntersection(segment1, segment2)
7:       if isWithinCenterRegion(intersectionPoint, image) And
8:         meetsAngleThreshold(estimateAngle(segment1, segment2), image) And
9:         isOuterCorner(segment1, segment2)
10:        addCornerToResult(intersectionPoint)
11:      end if
12:    end for
13:  end for
14: end Function

```

圖 8 角點萃取演算法之虛擬碼





圖 9 內凹女兒牆結構之角點標籤方式

為評估角點萃取演算法的性能，研究中選擇了均方根誤差 (Root Mean Square Error, RMSE) 作為量化精度的指標。由於無人機影像具有高重疊率，因此處理的角點特徵範圍影像會產生大量數據。為了有效評估，本研究採用隨機取樣方法，對實驗區域中十分之一的角點影像的點位坐標進行人工判釋，並將這些數據視為真值。隨後，將角點萃取演算法獲得的點位坐標視為預測值，據此計算 RMSE，以評估演算法的精度。

## 2.4 角點三維坐標重建

經過前述演算程序可成功獲得建物角點的二維影像坐標。為了將這些影像角點轉換為三維絕對坐標系中的表示，本研究採用基於多影像的多光束前方交會法進行角點的三維坐標計算。針對在多視角高重疊序列影像中偵測到的建築角點，結合電腦視覺技術，逐一構建以角點為中心的特徵描述元。此程序的優勢在於所需建立的特徵描述元數量較少，並且在匹配和後續解算過程中能顯著減少計算量。隨後，通過對角點局部特徵進行影像特徵匹配 (Distance Ratio Matching)，建立影像間角點特徵的共軛關係，從而進一步執行空間前方交會，實現角點三維絕對空間坐標的精確估計。

多光束前方交會乃基於共線條件構成數學模型，如圖 10 所示。在假設光線直線傳播的情況下，影像的透視中心、像點與物點之間可建立共線條件方程式 (式 1)。利用多個共軛點及其外方位參數，可形成多組共軛光束的非線性方程式，從而解算角點的物空間坐標，其解算方程如式(2) 至式(5) 所示。理論上，相較於傳統的兩光束方法，多光束交會方法更具優勢，因為多組共軛光束能提供更多的觀測約束，有助於減少誤差並提高交會點位的精度，亦

能補償短基線造成的高程精度下降。在先前步驟中，已獲取所需的參數，包括內外方位參數及共軛點的影像坐標。由於共軛點坐標位於影像坐標系中，因此在解算之前，需將其轉換至像片坐標系，以確保在相同坐標框架下進行計算。

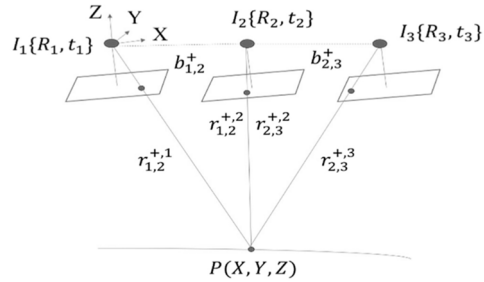


圖 10 多光束前方交會示意圖(Chuang & Sung, 2020)

$$\begin{aligned} x - x_0 &= -f \frac{m_{11}(X-X_L) + m_{12}(Y-Y_L) + m_{13}(Z-Z_L)}{m_{31}(X-X_L) + m_{32}(Y-Y_L) + m_{33}(Z-Z_L)} \\ y - y_0 &= -f \frac{m_{21}(X-X_L) + m_{22}(Y-Y_L) + m_{23}(Z-Z_L)}{m_{31}(X-X_L) + m_{32}(Y-Y_L) + m_{33}(Z-Z_L)} \end{aligned} \quad (1)$$

其中， $f$  為焦距； $x$ 、 $y$  為像片坐標； $X_L$ 、 $Y_L$ 、 $Z_L$  為透視中心坐標； $X$ 、 $Y$ 、 $Z$  為點位之物空間坐標； $m_{11} \sim m_{33}$  為由拍攝當下之姿態角構成的旋轉矩陣  $M_\kappa M_\phi M_\omega$  計算而得之參數。點位之物空間坐標可利用最小二乘法中的附有參數之條件平差進行解算：

$$A\xi + (y + e) = d, e \sim (0, \Sigma = \sigma_0^2 P^{-1}) \dots \dots \dots (2)$$

其中  $y$ 、 $e$ 、 $d$ 、 $\xi$  及  $P$  代表觀測量向量、誤差向量、常數向量、未知參數及權矩陣。 $A$  為針對未知數偏微分之係數矩陣。式(2)可再改寫為式(3)：

$$A\xi + e = w \dots \dots \dots (3)$$

其中，未知參數  $\xi$  可通過式(4)求得，而  $w = d - y$  代表差異向量；式(5)則可用於計算後驗單位權方差作為平差模式之正確性檢核。 $r$  代表多於觀測數目。

$$\xi = (A^T P A)^{-1} A^T P w \dots \dots \dots (4)$$

$$\hat{\sigma}_0^2 = e^T P e / r \dots \dots \dots (5)$$

研究中使用密匹配技術取得高密度的點雲數據，並以人工方式進行點雲平面擬合。透過將擬合的平面進行交會獲得交會角，以此作為本研究之檢核點，視其為真值角點坐標。在精度評估分為平面

精度、垂直精度、點位精度等面項，其中點位精度是表示結合了平面與垂直精度的整體精度，三項精度皆採用 RMSE 作為精度指標，並以擬合平面所交出之點為作為檢核點。

## 2.5 位相關係重建

本研究採用幾何規則與拓撲約束規則的方法，通過三維角點的幾何資訊與二維影像中的線段特徵，構建建物的幾何結構，其中包括對弧形結構的處理。通過引入幾何規則，可確保所構建的建物線框模型在空間中的相對位置及形狀具備合理性。此外，針對可能發生的遮蔽現象與角點重建失敗問題，本研究設計虛擬角點重建演算法，該方法可在必要時補充缺失的角點，從而確保整體線框模型的完整性。以下分別說明重建線面關係、虛擬角點重建演算法與完整性評估等三個步驟。

### 2.5.1 點線面位相關係重建

如圖 11 所示，點與點之間的位相關係重建方法乃針對垂直與傾斜之二維原始影像進行 Canny 邊緣偵測與霍夫轉換，以萃取房屋影像邊緣線。接著，將已有的三維角點坐標反投影至對應的二維影像，以此重建三維角點之間的線段關係。在線段偵測過程中，由於房屋邊緣線可能呈現不完整的線段，本研究將對偵測到的每一條線段進行延伸。此外，由於反投影過程中的誤差及三維角點本身的偏差，反投影後的角點可能未落在影像中的建物轉折處。為此，本研究在判斷線段關係時設置了偏差容許範圍，並對延伸線進行兩側加寬。此偏差容許範圍根據影像品質和三維角點的精度進行設定。如此，可通過判斷延伸線段是否穿過任意兩個角點來確定角點之間的連線關係。

然而，在某些情況下，延伸線可能穿過超過兩個角點，如圖 12 所示。圖 12(a) 為示意圖，在本研究策略中，P1 與 P2 將被視為一條線段，而 P3 與 P4 將被視為另一條線段。圖 12(b) 則展示了一個實例案例。針對此情況，本研究僅為最接近原始線段的兩個角點賦予連線關係。此外，在反投影過程中，由於建物高度和遮蔽的影響，部分點位可能無法在

特定拍攝視角的影像中準確呈現。為了解決此問題，本研究設置了基於高程的評判條件，即針對可連成水平線段的點位，角點間的高程差異不可超過一定的門檻值。該門檻值除了依據三維角點的品質設定外，亦取決於欲重建之建物類型。例如，在山型屋頂的建物中，此門檻值應設置為較高。

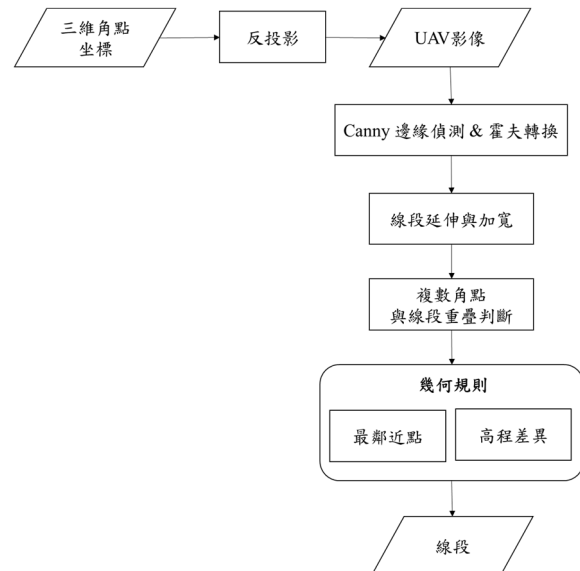


圖 11 點與點的位相關係重建方法

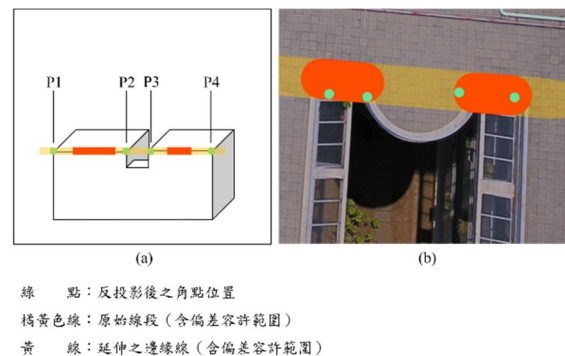


圖 12 延伸線上具有兩個以上角點之示意圖與案例

為了豐富可處理的幾何形態，本研究除了針對基礎的直線結構進行探討外，亦對弧形結構進行偵測，以重建弧形結構之房屋形態，如圖 13 所示。圖中，A 與 B 為位於霍夫圓偵測所得之圓形內的角點。通過距離轉換方法，可從像素分布中判別弧形的朝向。綠色點表示具有連貫性的像素，而藍色點則表示不具連貫性，故以綠色點較多的方向判定為弧形的朝向。P 點為 AB 線段以圓心 O 為中心，沿弧形朝向繪製鉛錘線後所形成的交點；a 與 b 則分別表示該圖形的半長軸與半短軸。

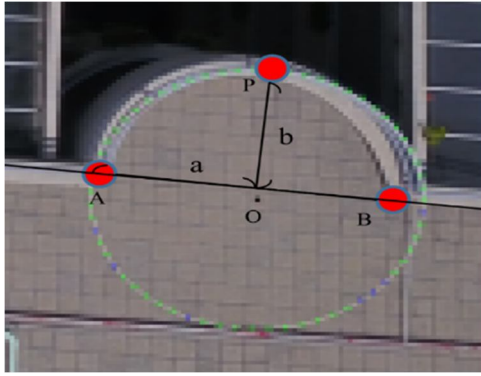


圖 13 弧形結構偵測策略

在此過程中，運用 Canny 邊緣偵測與霍夫圓偵測技術，旨在於二維影像中識別出圓形分布。霍夫圓偵測演算法允許設置篩選圓的大小，參考 CityGML 中對 LoD3 平面細緻度的規範，本研究將可被重建的弧形結構預設為長軸大於 2 m。同樣地，將三維角點反投影至二維影像中，針對已識別出的弧形進行進一步判斷，以確認是否存在兩個角點位於該弧形周圍。此周圍的範圍可設置偏差容許值。若弧形周圍的角點數量超過兩個，則篩選出最接近弧形直徑的兩個角點。接著，將這兩個角點的連線之中點視為圓心，並以圓心向弧形的朝向劃設鉛錘線。鉛錘線與圓形的交點即為該弧形的弧頂，圓心與該交點的距離及兩角點距離的一半將分別作為弧形的半短軸與半長軸。其中，弧形朝向的判定是利用距離轉換 (Distance Transform) 方法，通過影像中的像素差異來確定該弧形結構在圓形中的位置。最終，將保存兩個角點、圓心、弧頂的坐標，以及弧形的半長軸與半短軸資訊。

在初步建立線段資訊後，下一步驟是重建線與線之間的關係。所形成的平面將用於判斷體閉合狀態，並作為後續虛擬角點建置的依據。在此過程中，首先整理各線段相接的線段編號，並通過深度優先搜尋演算法 (Depth-First-Search, DFS) (Tarjan, 1972)，以樹狀結構的形式遍歷所有線段節點的連接關係。此階段通過編號進行搜尋，識別由線段節點所形成的閉合面。本研究期望獲得的面為平面，因此，對於具有多組法向量的面將不予紀錄。在執行策略中，

透過線段編號回推相關節點，即角點坐標，並運用三維角點坐標資訊，針對每一個閉合面計算其平面向程式，以判斷其是否為單一法向量平面。另一方面，若線段連結的終點位於一既有平面上，同樣視為閉合，此情況多見於具女兒牆結構的建物類型。如圖 14 所示，其中，黃色和綠色各為一單位閉合面，儘管在進行 DFS 時，兩者會被歸納為一組閉合面，但由於其發生轉折等因素，使得法向量不唯一，則會被拆分為不同面。



圖 14 單位閉合面之示意圖

## 2.5.2 虛擬角點重建演算法

在完成初步的線與面重建後，將進入虛擬角點重建演算法的環節，其演算流程如圖 15 所示。在本研究中，虛擬角點定義為在現實空間中實際存在，但因遮蔽或在先前過程中重建失敗而未被記錄的點位。虛擬角點重建演算法基於對稱性和點線面連結性原則，通過推論創建虛擬角點，以補充實際存在但重建失敗或被遮蔽的點位。該演算法的主要設計原則是用於補足建物底部以及陽台後側的不足，其實際應用情境如圖 16 所示。其中，圖 16(a) 之建物底部角點受樹木之遮蔽，導致無法呈現於照片中；圖 16(b) 則是受到視角與其他人為造物影響，導致部分陽台後側點位無法被自動識別。

在虛擬角點重建演算法中，多數程序遵循曼哈頓世界假設，即假定三維空間中的所有物體、牆壁、地板等皆由垂直與水平的直線構成，因此空間中的物體均符合直角幾何。虛擬角點重建演算法將持續迭代，直至實現體閉合才會結束。



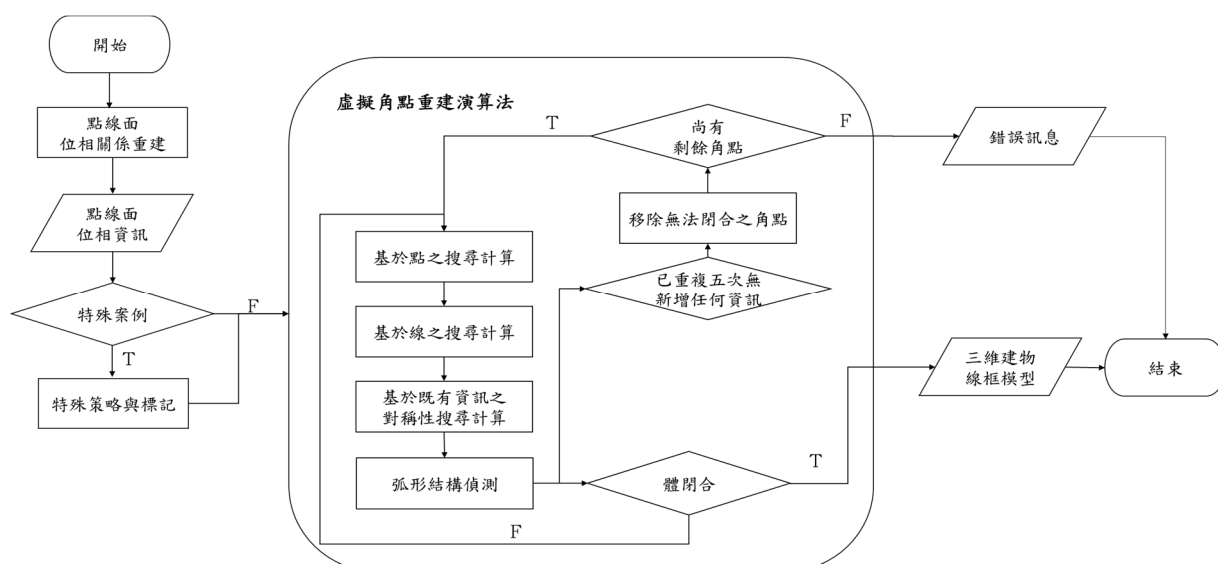


圖 15 虛擬角點重建演算法流程圖

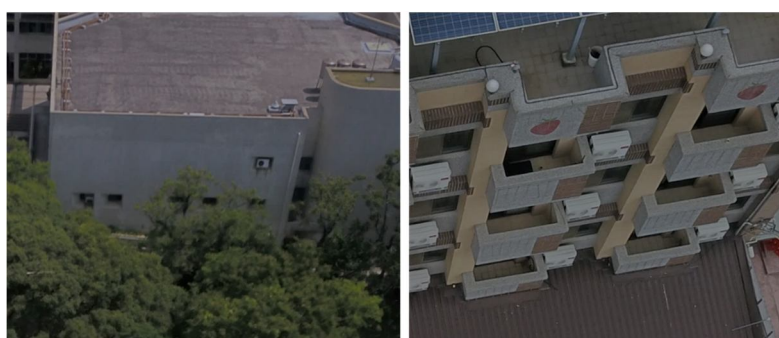


圖 16 建物幾何角點受遮蔽之影像示意圖

演算流程依序進行基於點、基於線面、及基於既有資訊對稱性的搜尋。首先利用既有點重建合理的線段，並通過新生成的線段與既有線面及對稱性資訊進行後續重建。在基於線面搜尋的階段，會先運用圖 18 中的判斷式，隨後再執行圖 19 中的判斷。如果任一點無法達成閉合，則該點將自動被刪除，以完成體閉合並建構出閉合的線框模型(Watertight Wireframe Model)。具體規則與策略如下所述：

#### (1) 基於點之搜尋

- A. 基於曼哈頓世界假設
- B. 以任一點至少可以連結成兩條線為原則
- C. 目的在於重建線段資料
- D. 在線資料內搜索僅使用到一次或沒有跟任何點連成線段之點，並尋找與之最鄰近之面。針對僅使用到一次的點資料，將計算其線段與最鄰近之面邊緣線是否具平行關係，若在誤差容許範圍內，則對其進行線面重建，如圖 17 之

Case A 所示，在此過程中不會產生新的虛擬點，而是新增線段與面資料。若為單一點的情況，則計算該點與最鄰近面之最鄰近點是否可形成一條與最鄰近面垂直之線段，若可，則在此新增一條新的線段資料，如圖 17 之 Case B 所示。

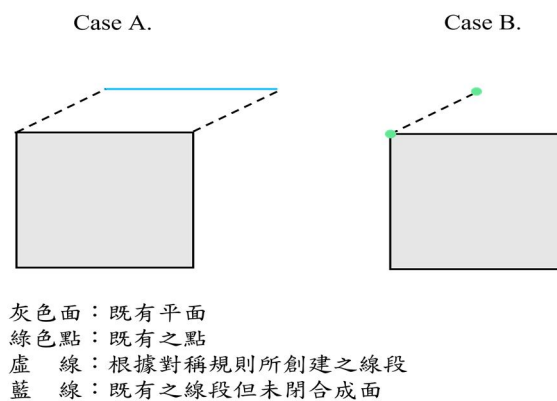


圖 17 基於點搜尋之線段重建示意圖



(2) 基於線面之搜尋

- A. 以所有線段皆具有鄰邊為假設；基於曼哈頓世界假設。
- B. 以任一線段至少可組成兩個面為原則。
- C. 目的在於重建虛擬點資料。
- D. 在面資料內搜索僅使用到一次的線段，暫稱此單一為 F1，於目標線段之轉折處，規劃一垂直於 F1 之平面 F2。若 F2 於其對稱方向已存在既有之面，則根據其對稱面之高程坐標以及目標線段之平面坐標賦予缺失點坐標，如圖 18 之 Case A 所示。若無，則將 F2 視為垂直於地平面之面，以參數最低起算高程之值作為其缺失點高程坐標，其餘則以對稱目標線為原則建置，如圖 18 之 Case B 所示。
- E. 使用上述 DFS 演算法，找出已連結部分卻未閉合之線段組合。首先，判定未閉合之兩端點連結是否可形成一對稱面，若是，則可徑直補足最後一條線段，添增一個面，如圖 19 之 Case A 所示。若否，則將既有之線段產生相互之對稱線段，並組成一面，如圖 19 之 Case B 所示。

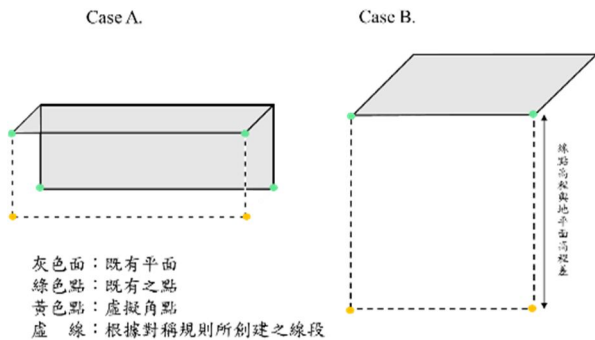


圖 18 基於線段僅被單一面使用之角點重建示意圖

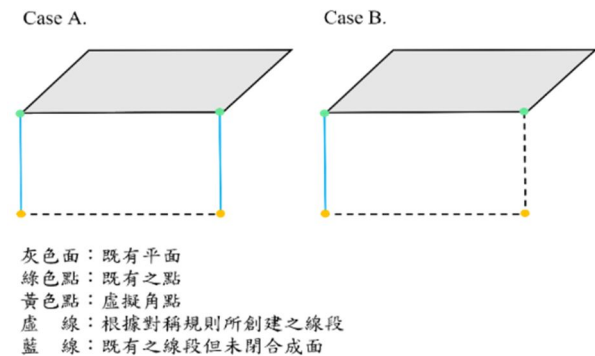


圖 19 基於未閉合之線段組合之角點重建示意圖

(3) 基於既有資訊之對稱性搜尋

- A. 不限於矩形面之假設，但僅限於水平方向面之線段組，且其實際之三條邊線關西需為兩條互相平行，並且各與第三條垂直。
- B. 既有之未閉合線段組僅有一組解可達到對稱，其狀況如圖 20 所示。
- C. 目的在於重建特殊對稱形狀之幾何。
- D. 在方法上將鎖定垂直之兩條邊線，分別取其中點劃設延伸之鉛錘線，並計算鉛錘線兩側之既有點到鉛錘線之距離與點數，若扣除缺少之點數後，兩側點到鉛錘線之距離皆相等，則以該鉛錘線為對稱性，重建出虛擬點。

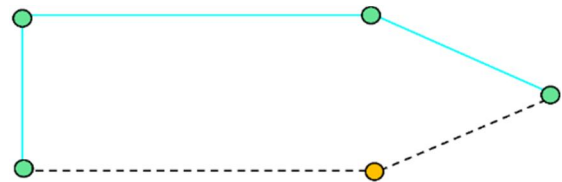


圖 20 基於既有資訊之對稱性重建角點之示意圖

在特定情況下，以上規則及策略將不適用，包括具女兒牆的建物頂樓內側與山型屋屋頂與底面之連接，如圖 21 所示。由於這些結構的特殊性，於進入虛擬角點建置環節前，將針對這兩種情況先行進行自動化判斷。首先，是否為具女兒牆之建物頂樓內側，將透過內外側面積比較及角點位置的相近程度進行判斷。針對該區域，本研究設定允許其以單一面形式存在，不需要透過其他線或面進行搜尋與整合，且內外牆皆不受限於單一法向量之面約束。山型屋的判斷則基於該區域最高高程的面結構之斜率。山型屋頂並不採用將屋頂角點直接向下連結以生成線面的方案。若真實的底面角點已被重建，則將其直接向上連接至屋頂平面，並在相交處建立一個屬於屋頂平面與底面連接之虛擬角點。若資料中未包含底面角點資訊，則依據屋頂角點的位置，向內縮一定距離，並結合最低起算高程參數，創建虛擬角點。隨後，將底部角點向上連接至屋頂平面，並賦予一虛擬角點以完成線段連結。

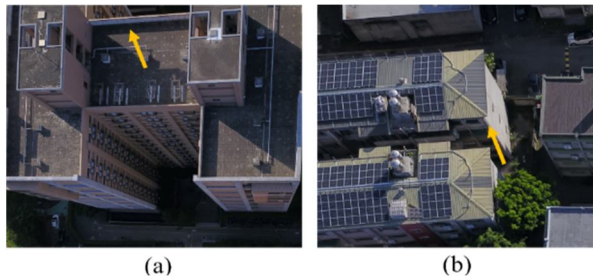


圖 21 虛擬角點重建演算法中之特殊案例。其中，(a)箭頭處示意具女兒牆之建物頂樓內側、(b)箭頭處示意山型屋之滴水線結構

### 3. 測試驗證

#### 3.1 獨棟建物三維線框模型重建

為了驗證本研究提出方法之成效，實驗中針對四棟建築物進行重建並分析其成果表現。包括國立陽明交通大學工程二館、苗栗縣大湖鄉台三線之白色民宿、苗栗縣大湖鄉南湖國中的健行館及育賢館，如圖 22 所示，以下將分別稱之為建物一至建物四。

此測試使用 DJI Mavic Pro 無人機進行資料蒐集，建物一的資料蒐集配置為採用 4.73 mm 之焦距分別以 50 m 及 45 m 之航高進行垂直及傾斜拍攝，對應之地面取樣距離為 0.017 m 及 0.022 m。其中，垂直拍攝之前後與左右重疊率為 80% 及 70%；建物二之資料蒐集配置為採用 4.76 mm 之焦距分別以 40 m 及 30 m 之航高進行垂直及傾斜拍攝，對應之地面取樣距離為 0.013 m 及 0.013 m。其中，垂直拍攝之前後與左右重疊率為 90% 及 80%；建物三與建物四則採用 6.09 mm 之焦距分別以 45 m 及 25 m 之航高進行垂直及傾斜拍攝，對應之地面取樣距離為

0.012 m 及 0.014 m。其中，垂直拍攝之前後與左右重疊率為 80% 及 80%。

針對與角點獲取相關的前三項程序，包括角點特徵偵測模型預測、角點萃取演算法與角點三維坐標重建進行成果分析，其對應成果列於表 2 至表 4。角點特徵偵測模型預測的成果如圖 23 所示。

可以發現使用整棟建物的所有影像進行預測時，成果表現更加出色。儘管整棟建物的複雜度較高，但標準且清晰的建物角點占整體建物的比例較高，這有助於提升整體的精確率與召回率。同理，表 3 在角點萃取的結果中，若角點特徵區域清晰且無干擾，則其成果相對會有所提升；而建物三、四，其特色在於鐵皮屋頂、條狀紋理及周圍環繞的水管，這些特徵匯集了許多干擾因素，導致角點萃取演算法較易發生錯誤偵測，故其精度相對較低。從表 4 中可以發現，建物三與建物四的平均點位精度相對較差，其中建物三的最差精度甚至達到 40 cm。

經由位相關係重建程序後，建物線框模型的成果如圖 24 與所示。其中，圖 24(a) 為線框模型上視圖；圖 24(b) 為線框模型側視圖；圖 24(c) 與 (d) 分別為線框模型上視圖與側視圖之對照影像；圖 24(e) 為線框模型檢核上視圖；圖 24(f) 為線框模型檢核側視圖。而圖 24 中之藍色點為一般點；紅色點為虛擬點；綠色點為女兒牆內部平台之點。類似地，圖 25(a) 為線框模型上視圖；圖 25(b) 為線框模型側視圖；圖 25(c) 與 (d) 分別為線框模型上視圖與側視圖之對照影像；圖 25(e) 為線框模型檢核上視圖；圖 25(f) 為線框模型檢核側視圖，藍色點為一般點；紅色點為虛擬點。



交大光復校區工程二館  
(建物一)

苗栗縣大湖鄉台三線之民宿  
(建物二)

苗栗縣大湖鄉南湖國中  
(建物三與建物四)

圖 22 實現建物影像

表 2 角點特徵偵測模型用於預測整棟建物之表現

	角點特徵偵測模型					
	Precision	Recall	AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>90</sub>
建物一	0.99	0.95	0.56	0.96	0.79	0
建物二	1.00	0.96	0.48	0.92	0.60	0
建物三	0.98	0.88	0.51	0.94	0.72	0
建物四	0.99	0.91	0.53	0.94	0.76	0.02

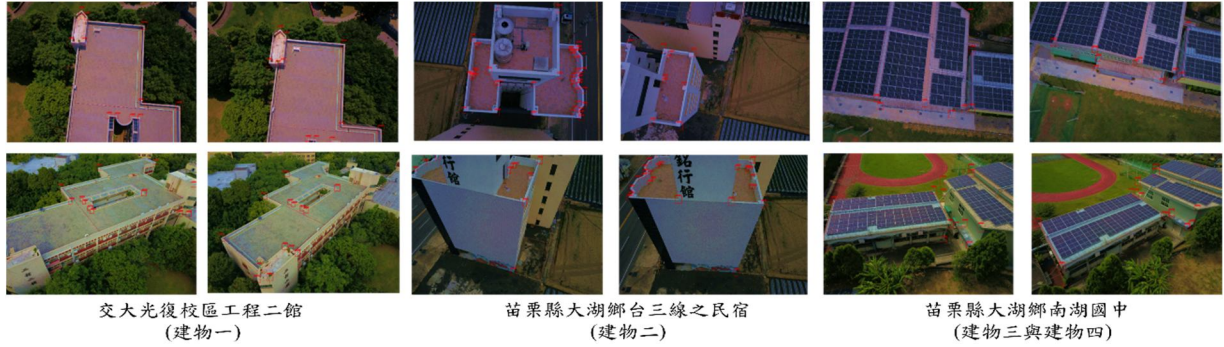


圖 23 角點特徵偵測模型預測結果

表 3 角點萃取演算法成果精度

	RMSE (px)
建物一	6.3
建物二	5.7
建物三	11.1
建物四	8.0

表 4 角點三維坐標重建成果精度

單位：m	建物一	建物二	建物三	建物四
平面精度	0.17	0.15	0.19	0.17
高程精度	0.18	0.15	0.21	0.21
點位精度	0.25	0.21	0.28	0.27
精度最高之平面精度	0.07	0.06	0.04	0.05
精度最高之高程精度	0.04	0.07	0.08	0.06
精度最高之點位精度	0.08	0.09	0.09	0.08
精度最低之平面精度	0.24	0.18	0.31	0.24
精度最低之高程精度	0.24	0.24	0.28	0.23
精度最低之點位精度	0.34	0.30	0.42	0.33

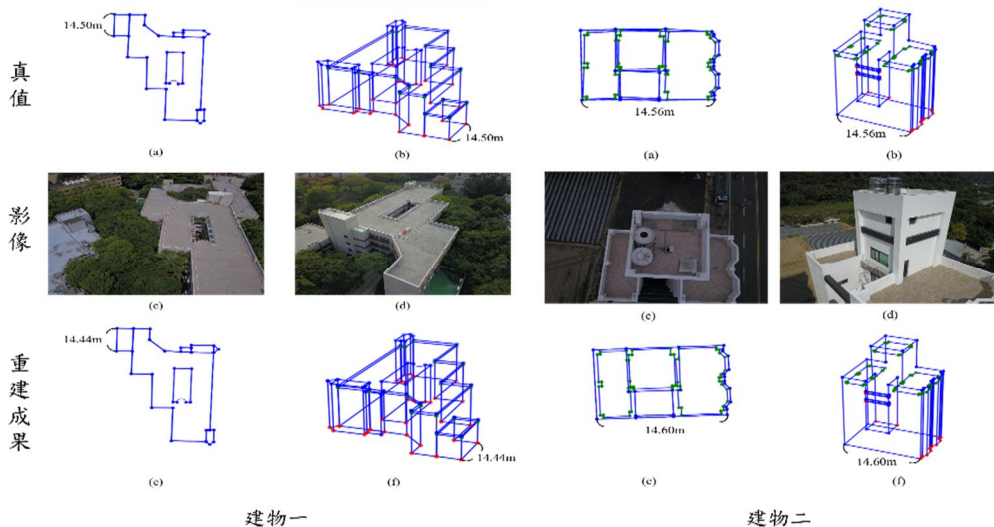


圖 24 建物一與二之線框角點偵測成果對照圖，(a)與(c)為上視圖；(b)與(d)為側視圖



建物模型的等角圖則如圖 26 所示，而位相關係重建的完整性記錄於表 5 中。本次實驗中，對於近水平面進行了高程平差，使得結果更加合理。從表中的數據可以看出，整體線段重建效果極佳。在一般點與虛擬點部分，虛擬點有效地彌補了一般點的缺失，使得重建出的線框模型具有更高的完整性。其中，建物一的線段匹配率 (MR) 為 0.01。未能重建出的線段均位於五邊形幾何形狀內，由於受最大閉合面運算影響而被移除。

點特徵範圍預測和角點萃取方面表現優異。這主要歸因於整體建物中標準且乾淨的角點數量足夠，使得角點特徵偵測模型可獲得良好的精確率和召回率。在角點萃取結果以建物一與建物二的表現較佳，可歸因於這兩個建物的特徵區域較為清晰且無雜訊。相對而言，建物三和建物四的資料具有挑戰性，其條紋狀鐵皮屋頂和周圍的水管等特徵容易引起角點萃取演算法的誤偵測，導致較低的精度。除此之外，位相關係重建的結果顯示一般點和虛擬點的重建獲得了實際效用，成功地填補了缺失點位，使最終的線框模型更加完整。

### 3.2 成果討論

依據實驗成果可以觀察到整體建物的重建在角

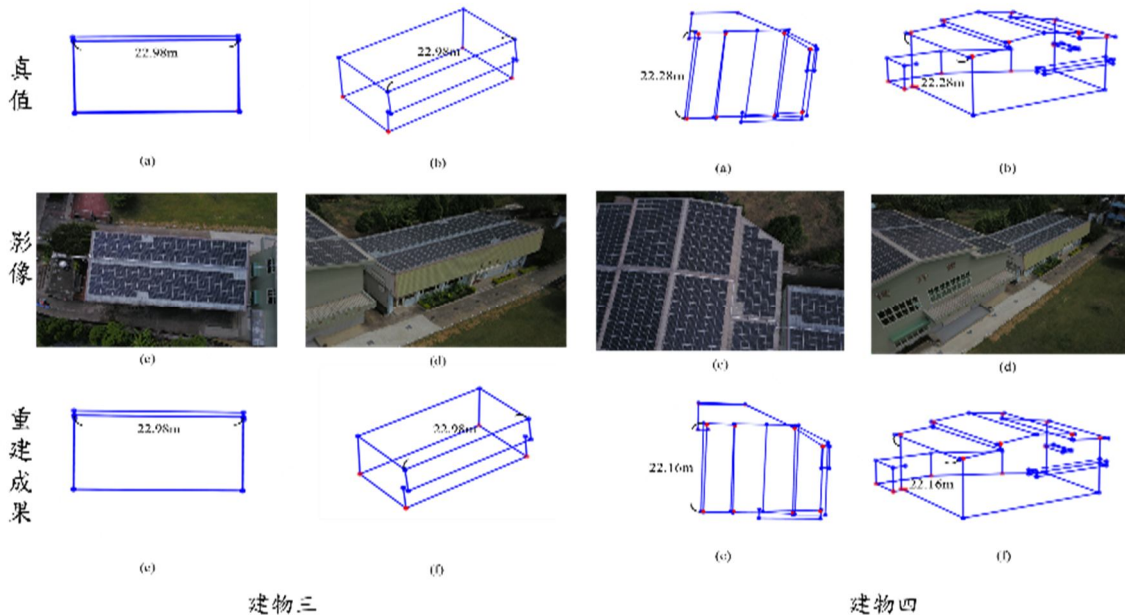


圖 25 建物三與四之線框角點偵測成果對照圖，(a)與(c)為上視圖；(b)與(d)為側視圖

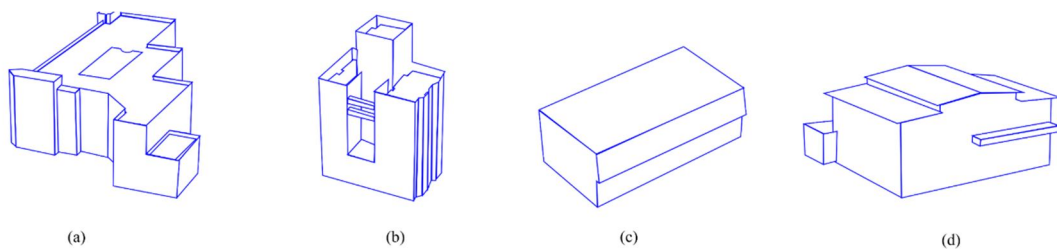


圖 26 建物三維線框模型重建成果，(a)至(d)為建物一至建物四

表 5 建物線框模型完整性評估

	一般點			虛擬點			整體點			線段		
	總檢驗數	F <sub>1</sub>	MR	總檢驗數	F <sub>1</sub>	MR	總檢驗數	F <sub>1</sub>	MR	總檢驗數	F <sub>1</sub>	MR
建物一	63	0.99	0.02	38	1.00	0	100	1.00	0	137	0.99	0.01
建物二	113	0.98	0.04	20	1.00	0	128	1.00	0	157	1.00	0
建物三	8	1.00	0	4	1.00	0	12	1.00	0	17	1.00	0
建物四	44	0.99	0.02	15	1.00	0	58	1.00	0	62	1.00	0



總體而言，所重建的模型在精度上符合 LoD3 的要求，平均精度約為 30 cm，最佳點位的精度可達到 10 cm 內。與其它研究相比，本研究的優勢主要體現在以下幾個方面：

- (1) 具側向結構物：大多數自動化重建研究中，立面結構的重建主要集中於如門窗等開口的偵測，而本研究提出的框架則針對整體建物的結構進行重建，這為後續細緻的 LoD3 建物模型搭建提供了不同的貢獻。
- (2) 對複雜建物的適應性：本研究處理的建物型態包括山型屋頂、具女兒牆的建物、弧形結構及側向結構（如陽台、水泥雨遮等）。在紋理方面，研究涵蓋了條狀鐵皮、條紋狀牆面與格狀地板。儘管每一項都有進步空間，但研究確實能成功自動化重建多種複雜建物，且平均精度達到 30 cm。
- (3) 虛擬角點重建演算法：為達到自動化建模的穩健性，本研究提出了一套虛擬角點重建演算法，以增強所建構之建物線框模型的合理性與穩健性。
- (4) 保留原始結構型態：本研究旨在盡可能保留建物原有型態，減少水平平面的簡化，僅對高程進行平差，最大限度地保留弧形結構、五邊型屋頂、複雜的女兒牆結構等非矩形特徵。

## 4. 結論

本研究提出基於多視角無人機影像的建物線框模型自動化處理框架。有別於多數現行研究，聚焦於使用單一資料源，基於無人機影像數據進行建物重建，強調資料獲取迅速、成本較低、精度適中、演算高度自動化等優勢。研究提出的演算策略結合深度學習模型自動偵測建物角點特徵，並提出精確定位角點之演算法，實現具側向結構建物的自動化重建。整體平均精度達到 30 cm 內，符合 LoD3 的精度要求，且整體點線重建完整性超過 98%。研究成果有助於協助 LoD3 或更精細建物模型的建置任務，提升作業效率。

儘管本研究實現了具側向幾何資訊的建物模型自動化重建，演算方法上仍存在限制和挑戰，需在後續研究進行探討和解決。本研究的建物線框模

型建置框架在很大程度上依賴於影像中的特徵資訊。與大多數研究相似，採用影像資料將不可避免地面臨陰影、反光及覆蓋物的影響，這些因素會影響角點的精度，甚至影響模型的完整性。為此，研究過程中只能盡量避免遮蔽嚴重的區域。若僅為部分遮蔽，尚可依賴虛擬角點重建演算法進行處理；然而，若遮蔽超過一定比例或關鍵點缺失，本方法將難以有效適用。另一方面，本研究提出的虛擬角點重建演算法多基於曼哈頓世界假設，因此虛擬重建出的幾何形態多為矩形。針對影像中未被拍攝或無法拍攝的區域，除非進行現地考證，否則無法驗證其是否真為矩形。因此，本研究所產製的線框模型僅可作為建物的基礎模型，對於掌握度較低的建物，仍需依賴人工進行後續檢核與修正。同時，鑑於三維模型細緻度的需求，在實務操作中，要在城市尺度下實現自動化的三維建物模型重建仍然面臨諸多挑戰。現行上多針對具特殊意義之建築進行 LoD3 等級的建模。如何在資料收集、資料運算策略及演算效能方面，實現大規模且高精度、高細緻度的自動化模型建置，依然是有待解決的課題。

## 參考文獻

- Abdollahi, A., Pradhan, B., Gite, S., and Alamri, A., 2020. Building footprint extraction from high resolution aerial images using Generative Adversarial Network (GAN) architecture, *IEEE Access*, 8: 209517–209527, DOI: 10.1109/ACCESS.2020.3038225.
- Aharchi, M., and Ait Kbir, M., 2020. A review on 3D reconstruction techniques from 2D images, Presented at the Innovations in Smart Cities Applications Edition 3: The Proceedings of the 4th International Conference on Smart City Applications 4, Springer, pp. 510–522, DOI: 10.1007/978-3-030-37629-1\_37.
- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection, arXiv preprint arXiv:2004.10934, DOI: 10.48550/arXiv.2004.10934.

- Buyukdemircioglu, M., Kocaman, S., and Kada, M., 2022. Deep learning for 3D building reconstruction: A review, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2022: 359–366, DOI: 10.5194/isprs-archives-XLIII-B2-2022-359-2022.
- Chen, D., Zhang, L., Mathiopoulos, P.T. and Huang, X., 2014. A methodology for automated segmentation and reconstruction of urban 3-D buildings from ALS point clouds, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7: 4199–4217, DOI: 10.1109/JSTARS.2014.2349003.
- Chuang, T.Y., and Sung, C.C., 2020. Learning and SLAM based decision support platform for sewer inspection, *Remote Sensing*, 12(6): 968, DOI: 10.3390/rs12060968.
- Faugeras, O., Robert, L., Laveau, S., Csurka, G., Zeller, C., Gauclin, C., and Zoghiani, I., 1998. 3-D reconstruction of urban scenes from image sequences, *Computer Vision and Image Understanding*, 69(3): 292–309, DOI: 10.1006/cviu.1998.0665.
- Flamanc, D., Maillet, G., and Jibrini, H., 2003. 3D city models: An operational approach using aerial images and cadastral maps, *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34: 53–58.
- Gadde, R., Jampani, V., Marlet, R., and Gehler, P.V., 2018. Efficient 2D and 3D facade segmentation using auto-context, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40: 1273–1280, DOI: 10.1109/TPAMI.2017.2696526.
- Gruen, A., Schubiger, S., Qin, R., Schrotter, G., Xiong, B., Li, J., Ling, X., Xiao, C., Yao, S., and Nuesch, F., 2019. Semantically enriched high resolution LoD 3 building model generation, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42: 11–18, DOI: 10.5194/isprs-archives-XLII-4-W15-11-2019.
- Gui, S., and Qin, R., 2021. Automated LoD-2 model reconstruction from very-high-resolution satellite-derived digital surface model and orthophoto, *ISPRS Journal of Photogrammetry and Remote Sensing*, 181: 1–19, DOI: 10.1016/j.isprsjprs.2021.08.025.
- Huang, H., Michelini, M., Schmitz, M., Roth, L., and Mayer, H., 2020. LOD3 building reconstruction from multi-source images, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43: 427–434, DOI: 10.5194/isprs-archives-XLIII-B2-2020-427-2020.
- Klimkowska, A., Cavazzi, S., Leach, R., and Grebby, S., 2022. Detailed three-dimensional building Façade reconstruction: A review on applications, data and technologies, *Remote Sensing*, 14(11): 2579, DOI: 10.3390/rs14112579.
- Liu, H., Xu, Y., Zhang, J., Zhu, J., Li, Y., and Hoi, S.C., 2020. DeepFacade: A deep learning approach to facade parsing with symmetric loss, *IEEE Transactions on Multimedia*, 22(12): 3153–3165, DOI: 10.1109/TMM.2020.2971431.
- Ma, R., 2005. Building model reconstruction from LIDAR data and aerial photographs (Ph.D.), ProQuest Dissertations and Theses, The Ohio State University, United States.
- Malihi, S., Valadan Zoej, M., Hahn, M., Mokhtarzade, M., and Arefi, H., 2016. 3D building reconstruction using dense photogrammetric point cloud, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41: 71–74, DOI: 10.5194/isprs-archives-XLI-B3-71-2016.
- McClune, A.P., Mills, J.P., Miller, P.E., and Holland, D.A., 2016. Automatic 3D building reconstruction

- from a dense image matching dataset, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41: 641–648, DOI: 10.5194/isprs-archives-XLI-B3-641-2016.
- Mesolongitis, A., and Stamos, I., 2012. Detection of windows in point clouds of urban scenes, in *proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Providence, USA, pp. 17–24, DOI: 10.1109/CVPRW.2012.6238910.
- Murtiyoso, A., Veriandi, M., Suwardhi, D., Soeksmantono, B., and Harto, A.B., 2020. Automatic Workflow for roof extraction and generation of 3D citygml models from low-cost uav image-derived point clouds, *ISPRS International Journal of Geo-Information*, 9(12): 743, DOI: 10.3390/ijgi9120743.
- Neuhausen, M., and König, M., 2018. Automatic window detection in facade images, *Automation in Construction*, 96: 527–539, DOI: 10.1016/j.autcon.2018.10.007.
- Pantoja-Rosero, B.G., Achanta, R., Kozinski, M., Fua, P., Perez-Cruz, F., and Beyer, K., 2022. Generating LOD3 building models from structure-from-motion and semantic segmentation, *Automation in Construction*, 141: 104430, DOI: 10.1016/j.autcon.2022.104430.
- Pu, S., and Vosselman, G., 2009a. Knowledge based reconstruction of building models from terrestrial laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 64: 575–584, DOI: 10.1016/j.isprsjprs.2009.04.001.
- Pu, S., and Vosselman, G., 2009b. Building facade reconstruction by fusing terrestrial laser points and images, *Sensors*, 9: 4525–4542, DOI: 10.3390/s90604525.
- Rau, J.-Y., and Chen, L.-C., 2003. Robust reconstruction of building models from three-dimensional line segments, *Photogrammetric Engineering & Remote Sensing*, 69: 181–188, DOI: 10.14358/PERS.69.2.181.
- Schuegraf, P., and Bittner, K., 2019. Automatic building footprint extraction from multi-resolution remote sensing images using a hybrid FCN, *ISPRS International Journal of Geo-Information*, 8(4): 191, DOI: 10.3390/ijgi8040191.
- Tarjan, R., 1972. Depth-first search and linear graph algorithms, *SIAM Journal on Computing*, 1(2): 146–160, DOI: 10.1137/0201010.
- Von Gioi, R.G., Jakubowicz, J., Morel, J.-M., and Randall, G., 2012. LSD: A line segment detector, *Image Processing On Line*, 2: 35–55, DOI: 10.5201/ipol.2012.gjmr-lsd.
- Wang, F., Zhou, G., Hu, H., Wang, Y., Fu, B., Li, S., and Xie, J., 2023a. Reconstruction of LoD-2 building models guided by Façade structures from oblique photogrammetric point cloud, *Remote Sensing*, 15(2): 400, DOI: 10.3390/rs15020400.
- Wang, J., Xu, Y., Remil, O., Xie, X., Ye, N., Yi, C., and Wei, M., 2016a. Automatic modeling of urban facades from raw LiDAR point data, *Computer Graphics Forum*, 35(7): 269–278, DOI: 10.1111/cgf.13024.
- Wang, M., Yuan, S., and Pan, J., 2023b. Building detection in high resolution satellite urban image using segmentation, corner detection combined with adaptive windowed hough transform, in *proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, Australia, pp. 508–511, DOI: 10.1109/IGARSS.2013.6721204.
- Wang, Q., Yan, L., Zhang, L., Ai, H., and Lin, X., 2016b. A semantic modelling framework-based method for building reconstruction from point clouds, *Remote Sensing*, 8(9): 737, DOI: 10.3390/rs8090737.

- Wang, R., 2013. 3D building modeling using images and LiDAR: A review, *International Journal of Image and Data Fusion*, 4: 273–292, DOI: 10.1080/19479832.2013.811124.
- Wei, S., Ji, S., and Lu, M., 2020. Toward automatic building footprint delineation from aerial images using CNN and regularization, *IEEE Transactions on Geoscience and Remote Sensing*, 58: 2178–2189, DOI: 10.1109/TGRS.2019.2954461.
- Wen, X., Xie, H., Liu, H., and Yan, L., 2019. Accurate reconstruction of the LoD3 building model by integrating multi-source point clouds and oblique remote sensing imagery, *ISPRS International Journal of Geo-Information*, 8(3): 135, DOI: 10.3390/ijgi8030135.
- Yu, D., Ji, S., Liu, J., and Wei, S., 2021. Automatic 3D building reconstruction from multi-view aerial images with deep learning, *ISPRS Journal of Photogrammetry and Remote Sensing*, 171: 155–170, DOI: 10.1016/j.isprsjprs.2020.11.011.
- Zhang, C., Cui, Y., Zhu, Z., Jiang, S., and Jiang, W., 2022a. Building height extraction from GF-7 satellite images based on roof contour constrained stereo matching, *Remote Sensing*, 14(7): 1566, DOI: 10.3390/rs14071566.
- Zhang, X., Chen, K., Johan, H., and Erdt, M., 2022b. A semantics-aware method for adding 3D window details to textured LoD2 CityGML models, in *proceedings of the 2022 International Conference on Cyberworlds (CW)*, Japan, pp. 63–70, DOI: 10.1109/CW55638.2022.00018.
- Zhu, Q., Jiang, W., and Zhang, J., 2015. Feature line based building detection and reconstruction from oblique airborne imagery, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40: 199–204, DOI: 10.5194/isprsarchives-XL-4-W5-199-2015.



# Automatic UAV Image Reconstruction for 3D Building Wireframe Models

Yu-Ling Huang<sup>1</sup>      Tzu-Yi Chuang<sup>2\*</sup>

## Abstract

3D building data is vital in smart city planning, management, and energy assessment. However, constructing accurate 3D models for existing buildings remains challenging due to the labor-intensive processes and insufficient automation. This study proposes an algorithmic strategy based on multi-view UAV imagery to generate 3D wireframe models with detailed lateral geometric features, serving as a foundation for constructing 3D building models of existing structures. This approach aims to improve operational efficiency and reduce costs. The algorithm employs a pre-trained corner detection model and a novel corner extraction algorithm, utilizing a "coarse-to-fine" strategy to achieve precise corner localization. Additionally, a virtual corner reconstruction strategy is employed to mitigate the inevitable occlusion and data loss in urban UAV imagery. Experimental results demonstrate that this algorithmic strategy adapts well to buildings with curved architectural structures, achieving an average corner localization accuracy of approximately 30 cm and up to 98% completeness in wireframe reconstruction.

**Keywords: Multi-View UAV Imagery, Image Modeling, Automated Building Wireframe Reconstruction, Wireframe Models, Deep Learning**

---

<sup>1</sup> Master, Department of Civil Engineering, National Yang Ming Chiao Tung University

<sup>2</sup> Associate Professor, Department of Civil Engineering, National Yang Ming Chiao Tung University

\* Corresponding Author, E-mail: jyuchuang@nycu.edu.tw

Received Date: Aug. 14, 2024

Revised Date: Sep. 09, 2024

Accepted Date: Sep. 11, 2024