

南華大學科技學院永續綠色科技碩士學位學程

碩士論文

Master Program of Green Technology for Sustainability

College of Science and Technology


Nanhua University

Master Thesis

應用深度影像技術於實景影像即時三維建模

Applications of Depth Image Technologies on Real-Time 3D

Modeling for Real Images



陳俊宇

Jun-Yu Chen

指導教授：黃冠雄 博士

Advisor: Guan-Shyong Hwang, Ph.D.

中華民國 111 年 7 月

July 2022

南華大學
永續綠色科技碩士學位學程
碩士學位論文

應用深度影像技術於實景影像即時三維建模
Applications of Depth Image Technologies on Real-Time 3D Modeling for Real
Images

研究生： 陳俊宇

經考試合格特此證明

口試委員： 賴信吉
葉文河
黃冠雄

指導教授： 黃冠雄

系主任(所長)： 

口試日期：中華民國 111 年 7 月 / 日

摘要

近年來利用虛擬實境進行虛擬導覽的應用日益普及，但由於多數虛擬導覽裡的場景過於像素化或者平面化，使其影像過於平淡因而使用體驗感受大打折扣。為其增進使用體驗感受，將模型立體化，使其使用沉浸感提升，本研究主要透過 kinetic V2 進行掃描，將場景轉換成 3D 模型。利用深度影像(RGB-D)將場景掃描，輸入至 RTAB-Map 進行運算以輸出 3D 模型，再藉由 Blender 將 3D 模型進行調整，以提升模型精緻程度，並降低模型製作的難度。

關鍵字：深度影像、三維模型

Abstract

In recent years, it becomes more popular that virtual tours have used the technologies of virtual reality for application. However, their experience and perception after the usage of such applications are bland, because most of the scenes in virtual tours are too pixelated or flat. To promote the perceptive experience, construct the three-dimensional models, and enhance the immersive feeling, this study utilizes the device of Kinetic V2 to convert scenes into 3D models. Use RGB-D to scan the scene, input it to RTAB-Map for calculation to output 3D model, and then adjust the 3D model by Blender to improve the model's refinement and reduce the difficulty of model production.

Keywords: RGB-D, 3D Model, RTAB-Map, Blender, Kinect V2

目錄

摘要.....	i
Abstract.....	ii
目錄.....	iii
圖目錄.....	v
表目錄.....	vi
第一章 前言.....	1
1.1研究背景.....	1
1.2研究動機與目的.....	2
1.3文獻回顧.....	3
1.4論文架構.....	9
第二章 軟體與硬體設備介紹.....	10
2.1 Xbox Kinect V2.....	10
2.1.1 Xbox Kinect V2其他相關應用.....	12
2.2 Blender.....	12
第三章 系統架構與研究方法.....	13
3.1 系統架構.....	13
3.2 深度影像(RGB-D).....	17
3.2.1 深度圖(Depth Map)應用.....	19

3.3	RTAB-Map.....	22
第四章	實驗結果.....	24
4.1	實驗結果.....	25
4.2	差異比較.....	28
第五章	結論與未來建議.....	31
5.1	結論.....	31
5.2	未來建議.....	31
參考文獻	32



圖目錄

圖1-1	時間差量測法示意圖.....	4
圖1-2	三角法:單相機量測示意圖.....	6
圖1-3	三角法:雙相機.....	7
圖1-4	立體視覺(Stereo Vision)架構.....	8
圖2-1	Xbox Kinect V2.....	11
圖2-2	Xbox Kinect V2感測元件.....	11
圖2-4	Blender 介面.....	12
圖3-1	三維雷射在古蹟維護應用(a)古蹟實景圖(b)古蹟掃描結果.....	14
圖3-2	實驗流程圖.....	16
圖3-3	深度影像(a)RGB 影像(b)深度圖影像(c)RGB-D 影像.....	18
圖3-4	ZED 2 Stereo Camera.....	19
圖3-5	[12]的實驗結果.....	21
圖3-6	深度圖渲染結果(左)原始深度圖(右)經過處理的深度圖.....	22
圖4-1	實驗室示意圖.....	24
圖4-2	深度影像之三維模型(a)未經過調整(b)經過 Blender 調整.....	27
圖4-3	WebODM 三維模型.....	28
圖4-4	深度影像(a)三角網格(b)三維模型.....	29
圖4-5	WebODM(a)三角網格(b)三維模型.....	30

表目錄

表2-1 Xbox Kinect V2規格[5].....	10
表3-1 SLAM 可輸入及輸出支援表[18].....	22
表4-1 RGB-D 與 WebODM 比較.....	25



第一章 前言

1.1 研究背景

在虛擬實境日益普及的帶動下，虛擬實境中作為場景的三維模型也受到關注，進而開始朝向所謂的虛擬導覽發展。因為虛擬實境中可以跨越場地與地點的限制，讓使用者可以在有限的空間下體驗到有如宇宙般大的場域，畫面上有如身歷其境的感官享受，為了滿足使用者使用虛擬實境進行導覽的期待與想像。

然而三維模型不只可以運用在虛擬實境當中，也可以運用於擴增實境、逆向工程上，可以更靈活的運用三維模型。像是擴增實境可以將製作好的三維模型投射出來，以便於做比對等等運用，又或是當古蹟修復苦於沒有工程圖可以比對時，就可以運用先前掃描好的三維模型，進而比對修復破損的部分，得以最大限度還原出當時的模樣。

近期利用三維模型進行逆向工程進行修復的例子，就是在西元2019年4月15日，法國巴黎聖母院發生大火，當時大家苦於沒有當時建築工程圖難以重建，且巴黎聖母院上方哥德式建築有需多細節，使復原難度相當高。藝術歷史學家「Andrew Tallon」，曾經使用雷射掃描將巴黎聖母院三維建模，誤差僅有5毫米之差，讓巴黎聖母院的修復工程得以順利進行。

1.2 研究動機與目的

將實景製作成三維模型，不僅具有保存典藏的功能之外，還具有龐大的價值，除了可以用在虛擬實境上，還可以運用在擴增實境及逆向工程當中。近年來，利用虛擬實境進行導覽越來越盛行，但虛擬實境中作為場景的三維模型，其模型都是單一較為平面，缺乏真實世界中立體的感覺，為了達到真實景色的效果，將實景製作成三維模型或許是最佳的選擇，光學雷達(LiDAR)進行掃描，所產生的三維模型品質相當優良，在細節也可以最大程度呈現出來，掃描出來的模型產生的誤差也相當低，但光學雷掃描器其設備價錢相當高，掃描出的三維模型僅僅只有其外型，並沒有將三維模型進行上色。

為了降低製作實景三維模型的難度，並減少製作三維模型的時間，所以本研究目前則透過 Xbox Kinect V2，使用 Ubuntu 作業系統作為處理核心，使用深度影像(RGB-D)相機，建構出三維模型的架構，再透過 RTAB-MAP 與 RGB 影像進行運算，就可以得出一個實景的三維模型。

本論文研究的目的，透過深度影像(RGB-D)相機進行掃描，將實際景色製作成三維模型，讓三維模型更具真實性，藉此達到實景的效果，讓使用者在虛擬實境中的體驗更加真實。

1.3 文獻回顧

在三維模型的相關研究中可以分成主動式、被動式以及3D 電腦圖形軟體(3D computer graphics)三種方式進行三維模型的製作，主動式的方法，主要是以雷射測距儀(Laser rangefinder)為基礎，利用雷射打到物體上往返的時間取得距離，並結合相機拍攝的物體表面影像，例如:三維雷射掃描儀(3D Laser Scanner)、Xbox Kinect V2、光學雷達(Lidar)等等。被動式的方法，則是以拍照為基礎，利用一個或多個相機拍攝物體在多個角度下的照片來重建三維模型，例如:3D Camera Box(3D Cambox)、3D Software Object Modeller(3D SOM)、WebODM 等等。3D 電腦圖形軟體，則是利用3D 電腦圖形軟體直接進行三維模型的建模以及上色，目前常見的軟體如:3ds Max、SketchUp、Blender 等等 [1][2][3]。

三維雷射掃描儀(3D Laser Scanner)能夠快速掃描物體，能夠在不接觸到物體的狀態下快速獲得物體表面高密度及高精度的三維模型，是一種精確的測量工具。三維雷射掃描儀主要構造由雷射測距儀(Laser rangefinder)，加上反射鏡進行垂直方向的掃描，水平方向則透過伺服馬達來達成，可達360度的掃描[4]。根據三維雷射掃描儀測量原理不同，又可以分類為時間差量測(Time-of-flight, TOF)及三角法(Triangulation)兩種，其中三角法又分為單相機(Triangulation principle-single camera solution)以及雙相機(Triangulation principle-double camera

solution)兩種，分別如下[2]:

(一)時間差量測(Time-of-flight, TOF):

由雷射測距儀發射出脈衝雷射光打到待測物體上，再接收經由待測物體反射回來的訊號，計算這段時間的時間長度，由光速 c 、時間 Δt 算出掃描儀與物體之間的距離 ρ ， $\rho = \frac{1}{2}c\Delta t$ 。(如圖1-1所示)。

利用此方法的雷射掃描儀比起使用三角法(Triangulation)可以測量的距離較長，但在較近的距離時，所測量出來的精度相對較差。目前市場上大部分的三維雷射掃描儀多採用此方法，如:Xbox Kinect V2(本研究所使用的機型)、Trimble G200、Riegl LMS-Z420、CyraX2500、Optech ILRIS-3D 等。



圖1-1 時間差量測法示意圖[2]

(二)三角法(Triangulation):

此方法利用三角形幾何關係求出距離。先利用三維雷射掃描儀發射雷射光到待測物體上，藉由入射光與反射光之間的夾角，利用在另一端的相機接收待物體反射的訊號，因為雷射光與感光耦合元件(Charge Coupled Device, CCD)之間的距離長度已經得知了，根據三角形幾何關係推求三維雷射掃描儀與待測物體之間的距離。三角法量測法在近距離的精度和時間差量測法比起來較為精確，但是可以測量的距離範圍也比較短。而另外一種三角法是採用兩台各自架設在兩端的相機，接收待測物體反射的訊號，計算三維雷射掃描儀與待測目標之間的距離。

(1)單相機三角量測法(Triangulation principle-single camera solution):

利用雷射光打到待測物體上時，由另外一側的相機拍攝該光點形成三角關係，而雷射與相機位置為固定基線長度，依此計算待測物體點的座標位置。 $\rho = \frac{L}{\cos \lambda}$ ， ρ 為距離， L 代表反射稜鏡到鏡頭的基線長度， λ 為相機和基線之間的夾角(如圖1-2所示)。

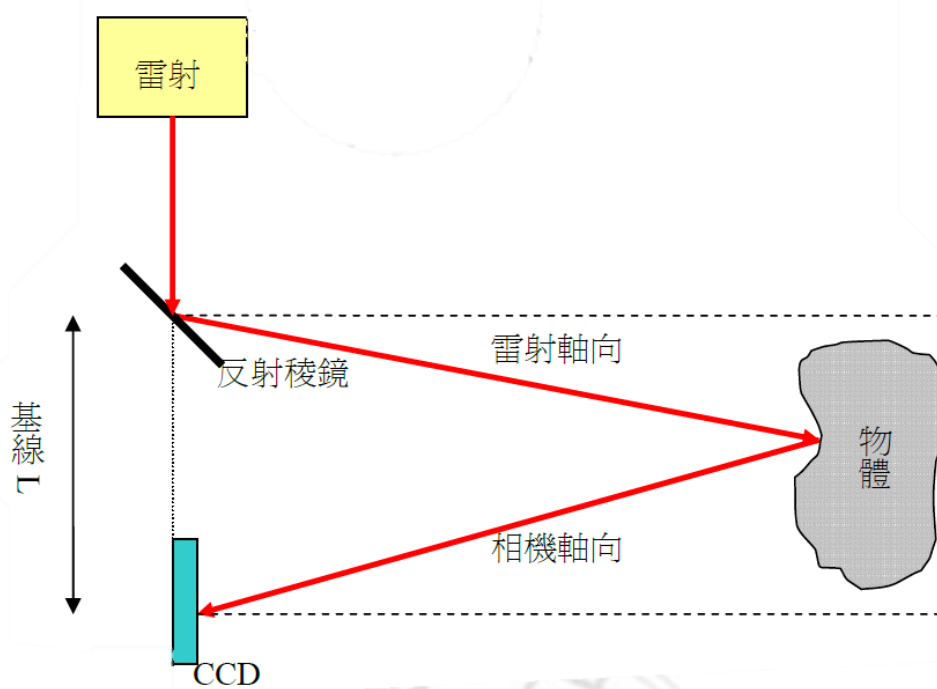


圖1-2 三角法:單相機量測示意圖[2]

(2)雙相機三角量測法(Triangulation principle-double camera solution):

利用雷射光打到待測物體上時，藉由兩側的相機拍攝該光點形成三角關係，而兩側相機為固定位置，藉此計算待測物體點之座標位置。 $\rho = L_1 \tan \lambda_1 = L_2 \tan \lambda_2$ ， $L = L_1 + L_2$ ， ρ 為距離， L 代表兩個相機鏡頭的基線長度， λ_1 、 λ_2 為相機鏡頭與基線的夾角(如圖1-3所示)。

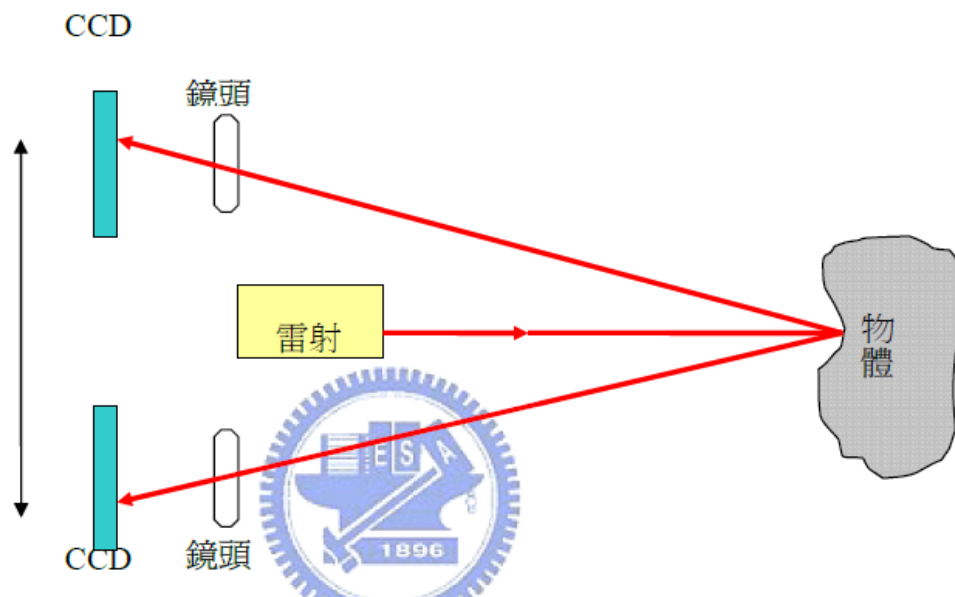


圖1-3 三角法:雙相機[2]

雖然三維雷射掃描儀可區分為時間差量測法和三角法，基本上三維雷射掃描儀仍是距離測量的一種，因此，三維雷射掃描儀的誤差也是隨著掃描儀和待測物體之間的距離而改變。對於距離較近的物體測量，因為距離較短，在時間的精準度上便會影響距離的精度，所以在近距離測量時，三角法的精度跟時間差量測法比起來精度較高，那對於距離較遠的物體測量，因為距離的增加，三角法的精度便會角度的偏差影響精度。

被動式的掃描方法則是以攝影測量法為基礎，利用一個或是多個相機拍攝物體在不同角度的影像以建立三維模型[1][3]。以3DSOM為例，它可以直接將物體進行360度所拍攝的照片來製作三維模型，原理為立體視覺(Stereo Vision)，藉由參考點即可計算出相對的三維座標位置。

立體視覺(Stereo Vision)是由人類雙眼視差的原理，經由兩個以上的相機對

同一個物體進行拍攝，以三角法計算出物體的距離及深度(Depth)資訊[18](如圖1-4所示)。其優勢在於成本較低，但是有效距離較短、延遲性較高。

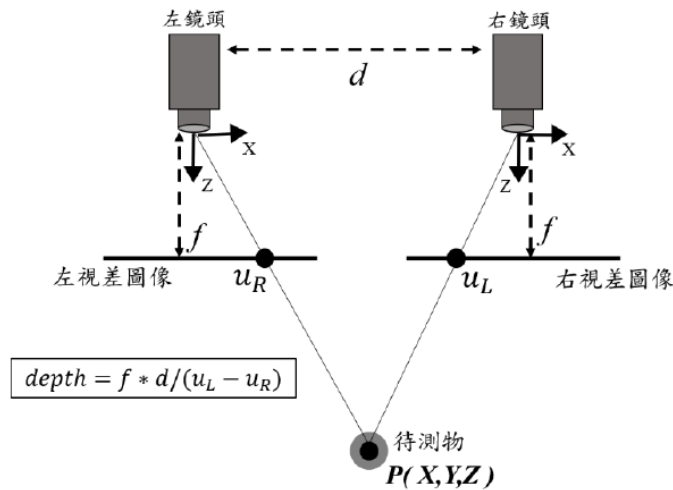


圖1-4 立體視覺(Stereo Vision)架構[18]

3D 電腦圖形軟體，此方法直接透過軟體進行三維模型製作。可以透過這些軟體製作出三維模型，物件三維座標、形狀、顏色乃至於材質都可以藉由3D 電腦圖形軟體製作[4]。但要，要達到擬真的程度，需要花費大量的時間，並且對於軟體操作有相當的技術才能夠完成。

1.4 論文架構

本論文總共有五個章節，各章節的研究內容概述如下：

第一章 前言

簡述本研究之研究背景、動機、目的、文獻回顧以及論文組織架構。

第二章 軟體與硬體設備介紹

講述本研究所使用到的軟硬體。

第三章 系統架構與研究方法

講述本研究系統架構及所採用方法，並說明深度影像(RGB-D)與 RTAB-MAP 演算法。

第四章 實驗結果

探討實驗結果所呈現的成效，並與其他方法作比較所耗費的時間成本。

第五章 結論與建議

依據實驗結果進行分析，已提供未來進行相關研究可供改善的地方，歸納結論與未來改善方向。

第二章 軟體與硬體設備介紹

2.1 Xbox Kinect V2

本研究採用的設備為 Microsoft 生產的 Xbox Kinect V2 如圖 2-1 所示，Kinect V2 主要的感測元件如下：彩色相機(RGB Camera)、深度感應器(Depth Sensor)、紅外發射器(IR Emitter)，四個陣列式麥克風[5](如圖 2-2 所示)。影像規格為 1920*1080，影像擷取速度為 30FPS，其規格如表 2-1 所示。深度感應器由紅外線發射器(IR Emitter)及紅外線相機構成的，使用時間差測量法(Time-Of-Flight, TOF)技術，其影像規格為 512*424，影像擷取速度為 30FPS，其距離為 0.4m-4.5m。

表 2-1 Xbox Kinect V2 規格[5]

RGB camera	解析度	1920*1080
	FPS	30
Depth camera	解析度	512*424
	FPS	30
Range		0.4m-4.5m



圖2-1 Xbox Kinect V2

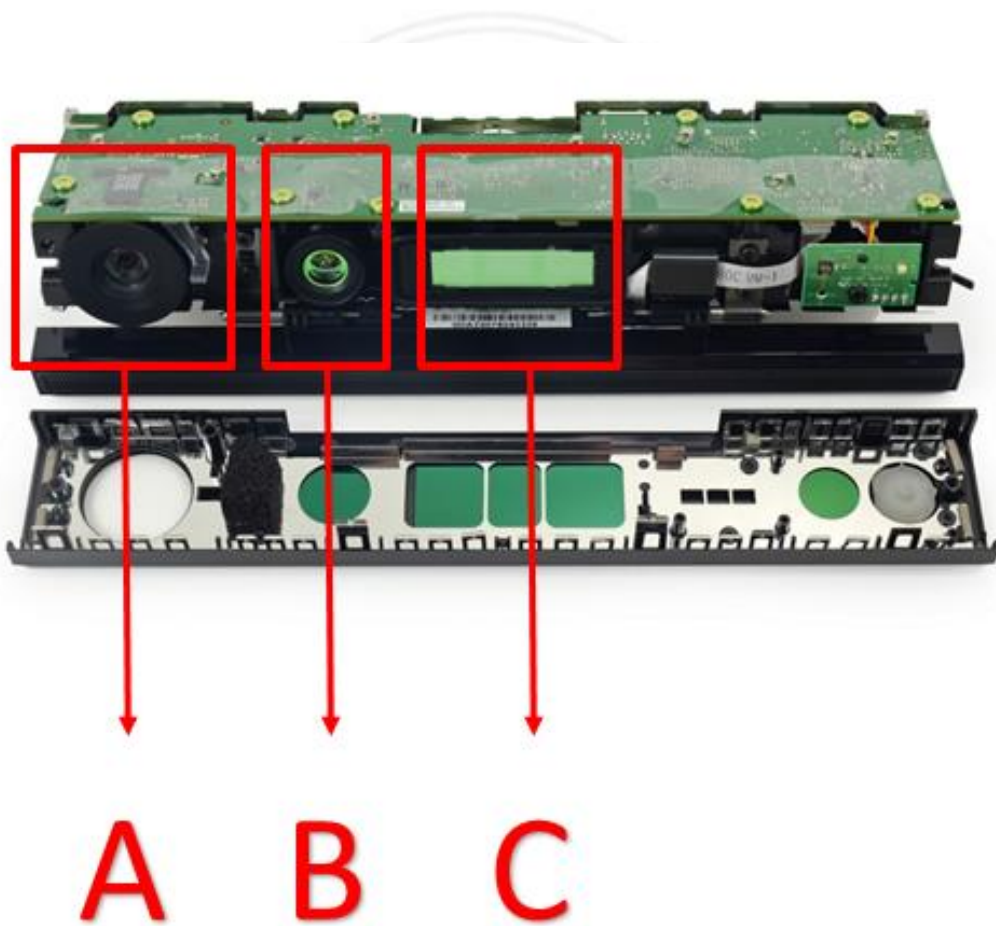


圖2-2 Xbox Kinect V2感測元件

(A)RGB Camera (B)IR Camera (C)IR Emitter

2.1.1 Xbox Kinect V2其他相關應用

Xbox Kinect V2除了作為三維模型的掃描儀器之外，Xbox Kinect V2還有許多應用，我們將相關應用分為以下幾類，包括三維建模、人體姿勢比對、手勢判別等等相關的應用[8]。

2.2 Blender

Blender 為3D 電腦圖形軟體，是一款免費的開源3D 圖形軟體(如圖2-4所示)。其功能相當多元，除了基本的建模之外，還可以製作動畫、模擬、渲染、合成等等，而且 Blender 還具有跨平台的特性[10]。

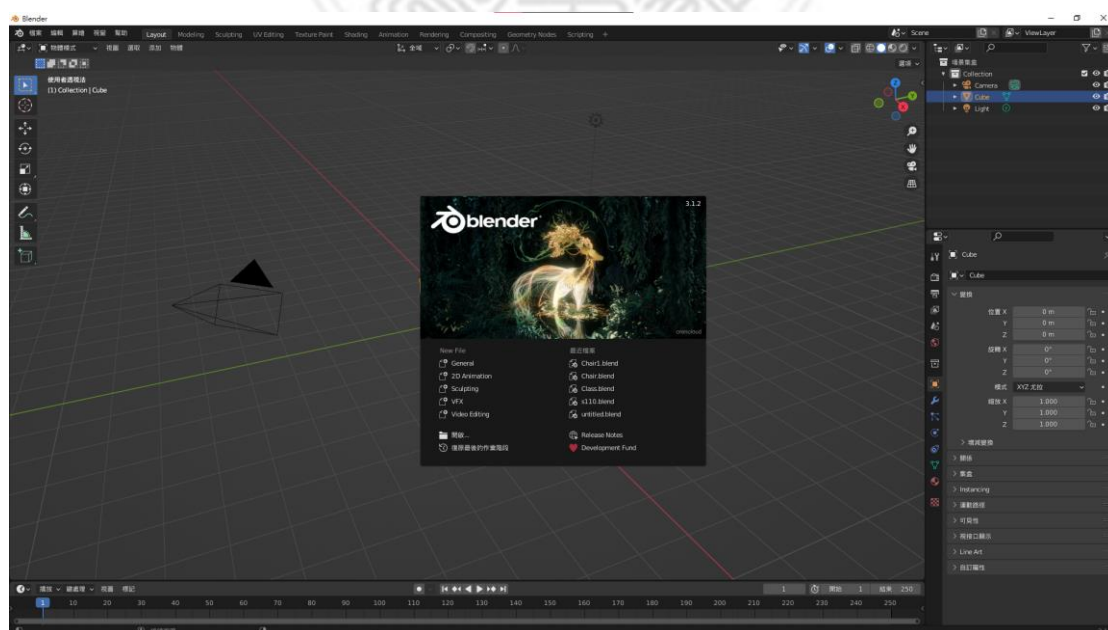


圖2-4 Blender 介面

在本研究中所使用的 Blender 版本為3.1.2，將 Kinect V2所掃描出來的三維模型匯入至 Blender 進行些微的修正，使三維模型更加完整，並根據模型將材質貼在對應的位置上，使三維模型更加貼近真實的效果。

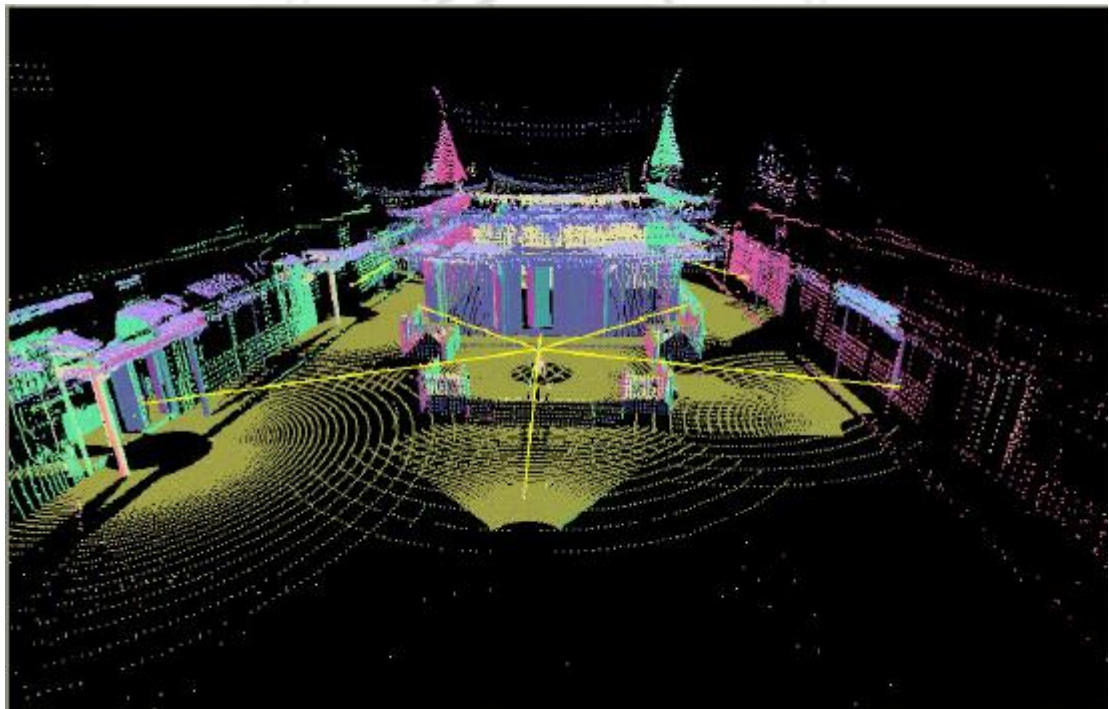
第三章 系統架構與研究方法

3.1 系統架構

本論文研究主要目標為以即時的方式將實景場景製作成三維模型，來提供使用者快速且簡單的製作方法。讓需要實景三維模型的使用者，可以有效率的將實景三維模型投入在虛擬實境導覽中，有效進行文化資產保存。藉由三維雷射掃描可輔助的應用領域包刮了橋樑隧道工程、古蹟維護(如圖3-1所示)、變形監測、機械與逆向工程、外科手術模擬訓練等等[4]。



(a)



(b)

圖3-1 三維雷射在古蹟維護應用(a)古蹟實景圖(b)古蹟掃描結果[4]

本研究實驗步驟如圖3-2所示，將要掃描的空間或物體藉由 Xbox Kinect V2

進行三維掃描，所掃描的資料透過 RTAB-Map(Real-Time Appearance-Based Mapping)運算，輸出後即可得到該空間或物體的三維模型。所得出的三維模型仍然會有些許的破洞與變形的區域，所以在藉由 Blender 電腦圖形軟體進行修整，就可以取的一個完整度相當高的實景三維模型。在下面的章節將會補充 RTAB-Map(Real-Time Appearance-Based Mapping)的相關技術[18]。



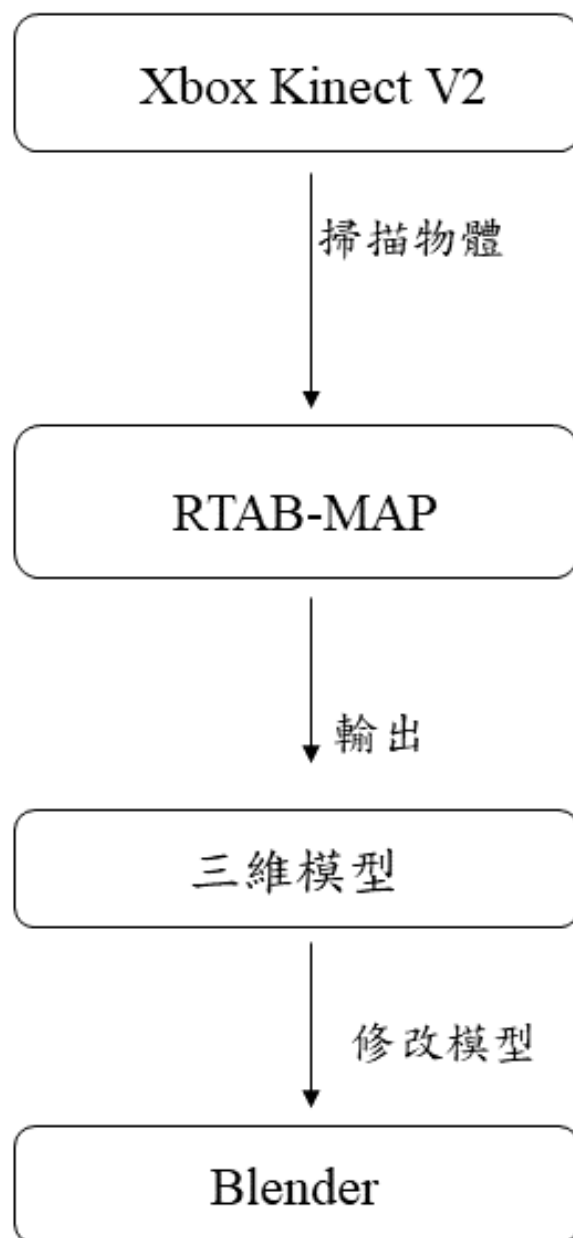


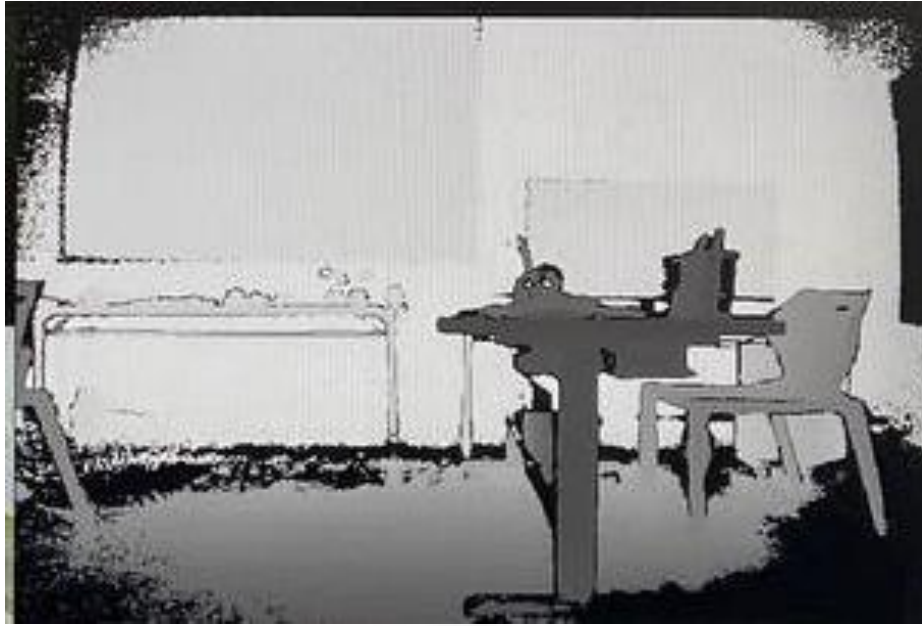
圖3-2 實驗流程圖

3.2 深度影像(RGB-D)

深度影像(RGB-D)為 RGB 和深度圖(Depth Map)，深度圖(Depth Map)為物件的距離的數據。本研究則是透過 Xbox Kinect V2作為深度影像(RGB-D)的輸入裝置。深度圖(Depth Map)也可以透過如立體視覺(Stereo Vision)技術獲取。如圖3-3所示，分別為 RGB 影像、深度圖(Depth Map)影像以及深度影像(RGB-D)。深度圖(Depth Map)根據亮度可以判斷距離，當越亮的部分代表著所計算出來的距離較遠，黑色區域代表著沒有取得深度值。



(a)



(b)



(c)

圖3-3 深度影像(a)RGB 影像(b)深度圖影像(c)RGB-D 影像

想要獲取深度影像(RGB-D)的數據，所以除了可以透過 Kinect V2之外，還可以利用立體相機(Stereo Camera)如:ZED 2 Stereo Camera(如圖3-4所示)等，透過立體視覺(Stereo Vision)技術取得深度影像[6]。



圖3-4 ZED 2 Stereo Camera

3.2.1 深度圖(Depth Map)應用

使用深度圖(Depth Map)為基礎的硬越來越多，除了作為三維模型建模重要的數據之外，還有基於深度圖的渲染(Depth Image Base Rendering, DIBR)[5]。舉例來說，在文獻[12]探討了深度圖的品質如何影響基於深度圖的渲染。在這篇論文中，作者提出新的方法來處理深度圖邊緣的問題。其結果如圖3-5所示，圖3-5(a)為彩色影像，圖3-5(b)為未處理的深度圖，圖3-5(c)將影像放大未處理的深度圖及輸出的結果，(d)為經過處理的深度圖及輸出的結果。可以看到當深度圖進行處理過後在邊緣區域較為完整。



(a)彩色影像



(b)深度圖



(c)未處理的深度圖及輸出的結果



(d)處理過的深度圖及輸出的結果

圖3-5 [12]的實驗結果

在文獻[13]這篇論文中，則是將深度圖用於渲染3D場景。在圖3-6可以看出來深度圖對於3D場景的重要性。可以看的出來圖3-6(左)所示，出來的結果有較多的破洞。圖3-5(右)相比之下，所呈現出來的3D場景明顯優於左圖。因此，深度圖的處理對於後續如三維建模、3D場景渲染等，都是影響結果的關鍵。



圖3-6 深度圖渲染結果(左)原始深度圖(右)經過處理的深度圖[13]

3.3 RTAB-Map

RTAB-Map(Real-Time Appearance-Based Mapping)演算法為即時定位地圖建構(Simultaneous Localization And Mapping, SLAM)的其中一種，包括了其他演算法如:RTAB-Map(本研究所使用的演算法)、GMapping、DVO-SLAM、ORB-SLAM2、MCPTAM、RGBDSLAMV2等等[18]。完整規格如表3-1，SLAM可以運用範圍也相當廣泛，比如自動駕駛、無人機、掃地機器人等等。

表3-1 SLAM 可輸入及輸出支援表[18]

	Inputs							Online Outputs			
	Camera				Lidar		Odom	Pose	Occupancy		Point Cloud
	Stereo	RGB-D	Multi	IMU	2D	3D			2D	3D	
GMapping					✓		✓	✓	✓		
TinySLAM					✓		✓	✓	✓		
Hector SLAM					✓			✓	✓		
ETHZASL-ICP					✓	✓	✓	✓	✓		Dense
Karto SLAM					✓		✓	✓	✓		
Lago SLAM					✓		✓	✓	✓		
Cartographer					✓	✓	✓	✓	✓		Dense
BLAM						✓		✓			Dense
SegMatch						✓					Dense
VINS-Mono				✓				✓			
ORB-SLAM2	✓	✓						✓			
S-PTAM	✓							✓			Sparse
DVO-SLAM		✓						✓			
RGBiD-SLAM		✓									
MCPTAM	✓		✓					✓			Sparse
RGBDSLAMv2		✓					✓	✓		✓	Dense
RTAB-Map	✓	✓	✓		✓	✓	✓	✓	✓	✓	Dense

即時定位地圖建構(SLAM)主要透過各式傳感器如:光學雷達、RGB-D相機、IMU 等等產生即時定位地圖的演算法。目前 SLAM 主要分為兩大類，基於光學雷達的 SLAM(Lidar SLAM)和基於視覺的 V-SLAM(Visual SLAM)。其中光

學雷達 SLAM，主要使用光學雷達進行掃描測距，優點是所測出來的數據相當穩定、地圖資訊準確，其缺點則有光學雷達價格較高且體積也比較大。V-SLAM 雖然在準確度及穩定性相對於 Lidar SLAM 較差，但 V-SLAM 所使用的感測器大幅低於光學雷達。目前主流的演算法有 ORB-SLAM、LSD-SLAM、RTAB-Map 等等。

ORB-SLAM:可支援的輸入儀器包括了立體相機和 RGB-D 相機，可以即時計算出相機軌跡和在各種環境中進行三維掃描，該演算法其主要核心為 ORB(Oriented FAST and Rotable BRIEF)特徵點。

LSD-SLAM(Large-Scale Direct SLAM):該演算法主要特色在於它不是藉由計算特徵點。該演算法的設計概念就是將利用影像中的資訊不斷計算相機的所在位置，在過程中不斷計算出深度圖(Depth Map)進而形成點雲(Point Cloud)。

RTAB-Map(Real-Time Appearance-Based Mapping):該演算法則是站在 ORB-SLAM 演算法的基礎上，在處理速度上也優於前面兩者，且所計算出的結果可用於3D 點雲(Point Cloud)以及2D 平面圖等。

第四章 實驗結果

本章節將會和 WebODM 所建立出的實景三維模型的進行差異比較，本研究將以系上實驗室如圖4-1所示，將作為實景三維建模的參考標準。下面將會分別展示深度影像(RGB-D)與 WebODM 所建立的三維模型，並對其所建立的三維模型進行差異比較分析。



圖4-1 實驗室示意圖

4.1 實驗結果

本節將分別展示深度影像(RGB-D)與 WebODM 兩種三維建模的成果，並比較兩者方法所建立出的實景三維模型的差異。在同一實景空間中的掃描下，深度影像(RGB-D)將實景空間建立為實景三維模型所需要的時間大約為5分鐘，所建立出的實景三維模型的三角網格數量為367,937，產生約1,103,805個頂點，檔案大小為25,099KB(.obj)。WebODM 建模需要環繞場域拍攝一圈後，將照片傳輸至電腦藉由 WebODM 進行計算，大約需費時24小時才可以得到實景三維模型，所建立出來的實景三維模型三角網格數量為308,903，產生約個926,709頂點，檔案34,196KB(.obj)大小為。完整模型數據如表4-1。

表4-1 RGB-D 與 WebODM 比較

	RGB-D	WebODM
點數	367,937	308,903
三角網格	1,103,805	926,709
檔案大小(.obj)	25,099KB	34,196KB
掃描時間	5 分鐘	24 小時

(一)深度影像(RGB-D)建模實驗結果

透過實驗結果可以發現藉由深度影像(RGB-D)技術所產生的實景三維模

型，所建立出的三維模型相當接近真實場景的狀態，如圖4-2所示，為本研究實驗結果。圖4-2(a)為未經過修正的三維模型，可以看到所輸出的三維模型仍有些許的不平整及破圖，為了有效提升三維模型的精緻度，所掃描出來的三維模型再經由 Blender 進行修正，其結果如圖4-2(b)所示，將邊緣的部份的模型進行修正，便可以取的完整度極高的三維模型。



(a)未經過調整



(b) 經過 Blender 調整

圖4-2 深度影像之三維模型(a)未經過調整(b)經過 Blender 調整

(二)WebODM 三維建模實驗結果

WebODM 為被動式的方法之一，製作方法主要為兩步，根據實景場景進行多個角度的拍攝，將拍攝到的照片輸入至電腦內，藉由 WebODM 進行運算得出三維模型。如圖4-3所示，可以明顯發現三維模型在許多地方都有白點的情形，在邊緣處更可以看到明顯的變形，雖然白點的點都不大，但三維模型破圖的狀況仍然明顯且數量不少，所呈現的三維模型雖然能大致看出是怎樣的物體，但與實際物品。



圖4-3 WebODM 三維模型

4.2 差異比較

經由4.1節，可以發現深度影像(RGB-D)與 WebODM 在三維掃描技術上，將實景建立成三維模型的時間上有明顯且顯著的差異。利用深度影像(RGB-D)技術僅需要五分鐘就可以將一間實驗室完成實景三維模型，而透過 WebODM 進行三維建模，則需要大約24小時，所耗費的時間與深度影像(RGB-D)有明顯的差異。

從兩者三維模型的三角網格可以更明顯看出兩種方法所建立的三維模型之間的差異。深度影像(RGB-D)所建立出來的三維模型，雖然在掃描時，會有產生多餘不必要的三角網格，但仍可以發現所建立出的三維模型在完整性是相當接近實景的狀態。如圖4-4所示，可以看到圖中建立的三維模型會參雜著場景中其他的物件，進而導致建立出的實景三維模型容易產生缺陷以及破損。



(a)三角網格

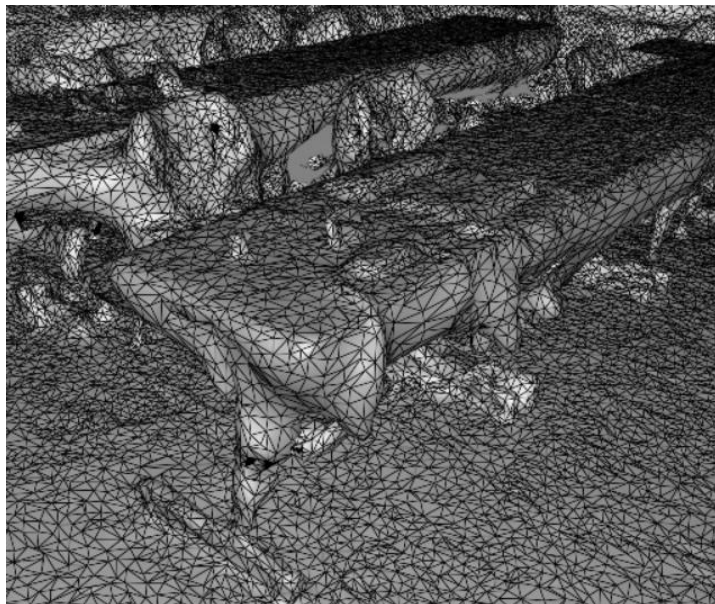


(b)三維模型

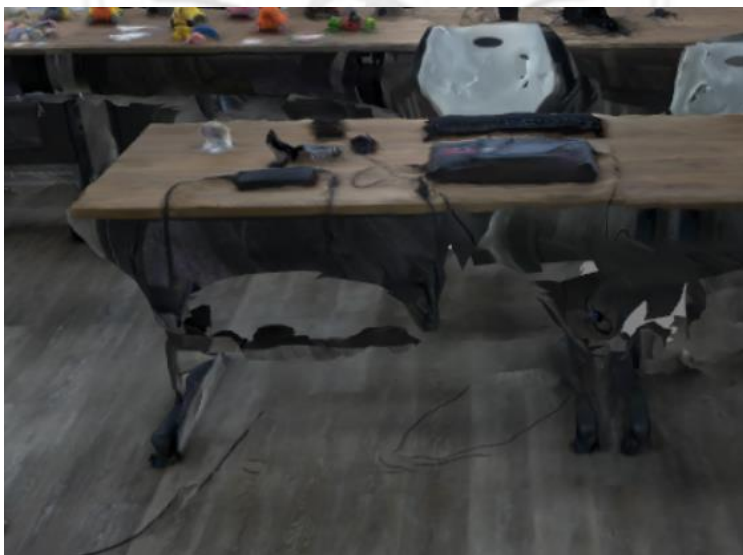
圖4-4 深度影像(a)三角網格(b)三維模型

WebODM所產生的三角網格，可以明顯看到三維模型的缺陷，因為此方法是藉由多個角度的照片來計算完成，來去構成三維模型建立的基礎，容易因為光線的差異及拍攝角度的完整性導致三維模型的成像是否精確。如圖4-5所示，

圖4-5(a)可以觀察出 WebODM 所建立的三維模型有明顯的變形問題，圖4-5(b)所呈現的三維模型僅能大致分辨出是怎樣的物體。



(a)三角網格



(b)三維模型

圖4-5 WebODM(a)三角網格(b)三維模型

第五章 結論與未來建議

5.1 結論

本研究透過以深度影像(RGB-D)技術做為三維掃描的核心技術，藉由 Xbox Kinect V2 收集深度影像(RGB-D)的數據，將數據導入至 RTAB-Map 輸出為三維模型，再經由 Blender 進行貼圖。其三維模型效果明顯提升可作為虛擬實境導覽使用。所以本研究所採用的方法，在時間上比起完全透過3D 電腦圖形軟體建模較為有效率外，也比利用照片快速生成的模型精度較高。

5.2 未來建議

目前本研究所掃描出來的三維模型，仍需要使用者透過3D 電腦圖形軟體將三維模型進行上修正的動作，但未來可朝向 AI 自動化處理，利用大數據進行比對自動將三維模型修正並將照片貼在正確的位置上，讓使用者不需要自行手動處理，便可以更快速地獲得實景三維模型，不用在自行手動修改三維模型。

參考文獻

- [1] 錢泓嶧，三維影像建模技術之研究與應用，2011
- [2] 曾信翰，三維雷射掃描儀反射鏡定為精度提升之研究-以 Trimble Mensi GS200為例，2008
- [3] 楊政儒，以雙繞射光學元件設計廣角結構光圖案之設計與實現，2021
- [4] 吳瑞一，三維雷射掃描技術應用之研究-3D 物件建模與變形模擬為例，2004
- [5] 林科瑜，對於 Kinect v2 RGB-D 影片之深度圖品質優化，2017
- [6] 蕭元昇，基於深度學習之 RGB-D 視覺辨識系統，2016
- [7] 吳建廷，地面三維雷射掃描儀影響精度因子之研究，2013
- [8] 廖承恩，以 RGB-D 攝影機為基礎的人體動作分析與理解，2016
- [9] 郭哲維，Blender 標記追蹤之探討於臉部表情動畫製作的應用，2017
- [10] 董擘欣，雲端即時三維建模檢視系統開發，2018
- [11] 鐘國豪，使用 RGB-D 攝影機實現移動式機器人即時同步定位與建構地圖，2013
- [12] Xu, Xuyuan, Lai-Man Po, Ka-Ho Ng, Litong Feng, Kwok-Wai Cheung, Chun-Ho Cheung and Chi-Wang Ting. "Depth Map Misalignment Correction and Dilation for Dibr View Synthesis." Signal Processing: Image Communication 28, no. 9 (2013): 1023-1045.

- [13] Cho, Ji-Ho, Kwan H Lee and Kiyoharu Aizawa. "Enhancement of Depth Maps with Alpha Channel Estimation for 3-D Video." IEEE Journal of Selected Topics in Signal Processing 6, no. 5 (2012): 483-494.
- [14] Si-Pin Weng, " Depth Map Enhancement based on Its Associated High-Resolution RGB Video," MS thesis, National Chiao Tung University, July 2015
- [15] Gandhi, Vineet, Jan Čech and Radu Horaud. "High-Resolution Depth Maps Based on Tof-Stereo Fusion." In Robotics and Automation (ICRA), 2012 IEEE International Conference on, 4742-4749: IEEE, 2012.
- [16] Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. Rgb-d mapping: Using depth cameras for dense 3d modeling of indoor environments. In the 12th International Symposium on Experimental Robotics (ISER), volume 20, pages 22—26, 2010.
- [17] Qing Zhua , Leveraging photogrammetric mesh models for aerial-ground feature point matching toward integrated 3D reconstruction , 2020
- [18] Mathieu Labbé , Francois Michaud , RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation , 2018
- [19] Burak Teke, Minna Lanz, Joni-Kristian Kamarainen, Antti Hietanen "Real-time and Robust Collaborative Robot Motion Control with Microsoft Kinect v2" , 2018