

國立臺灣海洋大學

電機工程學系

學士學位論文

指導教授：吳政郎 博士

視覺回授控制多軸機器手臂

Visual Feedback Control of a Multi-axis
Robot Arm

研究生：葉致均 撰

中華民國 112 年 6 月

視覺回授控制多軸機器手臂

Visual Feedback Control of a Multi-axis Robot Arm

研 究 生：葉致均

Student：Zhi-Jun Yeh

指導教授：吳政郎

Advisor：Jenq-Lang Wu

國立臺灣海洋大學

電機工程學系

學士論文

A Thesis

Submitted to the Department of Electrical Engineering College of
Electrical Engineering and Computer Science
National Taiwan Ocean University in partial fulfillment of the
requirements
for the Degree of Bachelor of Science
in
Electrical Engineering
June 2023
Keelung, Taiwan, Republic of China

中華民國 112 年 6 月

致謝

控制機械手臂結合視覺辨識，在一年前的我總覺得，要成功做出專題是遙不可及，而這個目標能夠付諸實踐必須先感謝我的恩師兼指導教授吳政郎老師，恩師給我很高的自主性以培育我獨立思考與解決問題的能力，在我對問題迷茫時，老師總是簡單的一句話就能將我點醒，我總在心底暗暗讚嘆老師的經驗與思維。當過程遭遇困難時，總不厭其煩的與我討論並指引我方向讓我自己去克服，一步一步指導我完成這篇論文，恩師是我做研究與人生中的重要導師，不僅深深影響我對做研究的態度，更使我對於自動控制產生極大的熱忱，在此對於吳政郎老師致上萬分的謝意。此外，也要感謝我的朋友蘇昭融，每天陪著我熬夜寫論文，並耐心帶我看懂每行程式在做什麼，給了我極大的依靠。

最後，將這份論文獻給辛苦培育我的父母葉明峰先生與黃美齡女士，因為有你們的支持與生活上的幫助，讓我在大學階段能無後顧之憂的完成學業。對我來說，現在才剛剛踏入自動控制的領域，在未來的人生裡，我會更努力地鑽研相關知識，不辜負父母對我的期待與教授對我的栽培。

葉致均 謹致於
國立臺灣海洋大學
一百一十二年六月

摘要

本論文主要目的是結合影像視覺與多軸機器手臂，使其突破按照既定路線作業的傳統式機械手臂，對於機械手臂在變化的環境中執行精密任務有很高的實用價值，在工業應用上可提升其適應性及降低機械出錯機率，減少人力資源，做到真正的自動化生產線。

本系統使用 Arduino uno 控制板，具有多組的數位輸出入端及串列埠，可供機器手臂多顆馬達同時控制及資料傳輸使用。本系統的影像處理仰賴 OpenCV，在辨識目標、測得深度並轉換為世界座標後，藉由串列埠將目標物的世界座標傳給 Arduino，使機器手臂能夠自主式的去完成夾取任務。

關鍵字：機械手臂、反向運動學、視覺回授、雙目測距、立體匹配

Abstract

The main objective of this paper is to integrate visual perception with a multi-axis robotic arm, surpassing the traditional mechanical arms that follow predefined paths. This combination has high practical value for the precise execution of tasks by robotic arms in changing environments. It enhances adaptability and reduces the probability of mechanical errors, thereby reducing manpower and achieving a truly automated production line.

This system utilizes an Arduino Uno control board, equipped with multiple sets of digital input/output ports and serial ports. These features allow simultaneous control of multiple motors of the robotic arm and facilitate data transmission. The image processing in this system relies on OpenCV. After recognizing the target and measuring its depth, which is then converted into world coordinates, the world coordinates of the target object are transmitted to the Arduino via the serial port. This enables the robotic arm to autonomously complete the gripping task.

Keywords: robotic arm, inverse kinematics, visual feedback, binocular ranging, Stereo matching.

目錄

摘要.....	IV
Abstract.....	V
目錄.....	VI
圖目錄.....	VIII
表目錄.....	X
第一章、緒論.....	1
1.1 前言.....	1
1.2 研究動機與目的.....	2
1.3 論文架構.....	2
第二章、系統軟硬體架構.....	3
2.1 系統架構簡介.....	3
2.2 機械硬體介紹.....	4
2.2.1 機械臂尺寸與架構.....	4
2.2.1.1 手臂改良.....	5
2.2.2 伺服馬達.....	6
2.2.3 Arduino Uno.....	8
2.2.4 Webcam.....	9
2.3 軟體開發環境與工具.....	10
2.3.1 開發環境 PyCharm.....	10
2.3.2 OpenCV.....	10
2.3.3 Arduino IDE 工具.....	11
第三章、視覺處理.....	13
3.1 影像處理流程.....	13
3.2 色彩空間與色彩模型.....	15
3.2.1 RGB 色彩空間.....	15
3.2.2 HSV 色彩空間.....	16
3.2.3 色彩空間轉換.....	17
3.2.4 二值化.....	18
3.2.5 膨脹與侵蝕.....	19
3.3 影像深度.....	21
3.3.1 雙目測距.....	21
3.3.2 深度轉換相機座標.....	26
第四章、運動學.....	28
4.1 空間中的剛體.....	28
4.2 歐拉旋轉定理.....	29
4.3 反向運動學.....	30

4.4 正向運動學.....	31
4.5 座標系轉換.....	32
4.5.1 相機座標轉圖像座標.....	32
4.5.2 相機座標轉世界座標.....	33
第五章、實驗結果.....	34
5.1 實驗一：立體校正輸出深度圖.....	34
5.1.1 實驗目的.....	34
5.1.2 實驗成果分析.....	34
5.2 實驗二：視覺回授隻機械手臂夾取紅色圓球.....	38
5.2.1 實驗目的.....	38
5.2.2 實驗成果分析.....	38
第六章、結論與未來展望.....	40
6.1 結語.....	40
6.2 未來展望.....	40
參考文獻.....	41

圖目錄

圖 1 Unimate.	1
圖 2 機器手臂系統方塊圖.	3
圖 3 AS-6DOF 尺寸模型.	4
圖 4 接口處加強型的厚型搖臂.	5
圖 5 夾頭改良實際圖.	5
圖 6 MG995 伺服馬達外觀.	6
圖 7 MG995 伺服馬達轉角幅度.	6
圖 8 MG995 伺服馬達控制脈寬區間示意圖.	6
圖 9 RobotBase RB-150MG 伺服馬達.	7
圖 10 Arduino Uno 控制板與腳位簡介(正面)	8
圖 11 Arduino Uno 控制板(背面)	8
圖 12 KINYO PCM-537 攝影機.	9
圖 13 PyCharm 上使用 OpenCV 操作介面.	11
圖 14 Arduino 編譯程式操作介面.	11
圖 15 影像處理流程圖.	14
圖 16 RGB 色彩空間.	15
圖 17 RGB 色彩之加法混色示意圖.	15
圖 18 HSV 色彩空間.	16
圖 19 HSV 圓錐體示意圖.	16
圖 20 二值化前後比較圖.	18
圖 21 膨脹流程示意圖.	19
圖 22 侵蝕流程示意圖.	20
圖 23 雙目測距流程圖.	21
圖 24 使用 MATLAB 對方格紙做雙目標定.	23
圖 25 雙目標定配對誤差直方圖.	23
圖 26 校正前示意圖.	23
圖 27 校正後示意圖.	23
圖 28 拍攝紅色圓球的左右視圖.	25
圖 29 轉為灰度圖後放在同一水平面進行比對.	25
圖 30 深度圖.	26
圖 31 空間中剛體的運動示意圖.	28
圖 32 以 Z 軸為基準遵循右手定則之旋轉示意圖.	29
圖 33 $H \geq 0$ 之二維空間逆向運動學模型.	30
圖 34 相機座標轉換圖像座標.	32
圖 35 相機標定場景.	34
圖 36 兩相機座標旋轉矩陣 R.	35

圖 37	兩相機座標平移矩陣 T	35
圖 38	左相機內參矩陣.....	35
圖 39	右相機內參矩陣.....	35
圖 40	原始 35 組相片配對誤差直方圖.....	35
圖 41	刪除誤差大相片組後留下 12 組較精準配對直方圖.....	35
圖 42	左右相機拍攝原始相片組.....	36
圖 43	立體校正相片組同水平面灰度圖.....	36
圖 44	影像深度圖.....	37
圖 45	紅色圓球的相機座標.....	37
圖 46	伺服馬達連接電路板之電路.....	38

表目錄

表 1 MG995、RB-150MG 伺服馬達規格.....	7
表 2 Arduino Uno 規格.....	8
表 3 Arduino Uno 接腳.....	9
表 4 網路攝影機規格表.....	10
表 5 HSV 色彩閾值.....	17
表 6 夾取實驗結果分析.....	38

第一章、緒論

1.1 前言

隨著科技的不斷進步和人們對機器智能化的需求，機械手臂的應用領域不斷擴展，從傳統的工業領域延伸到服務機器人、醫療保健、農業等多個領域。它的起源可以追溯到上個世紀，當時工業界開始將自動化技術應用於生產線，並發展出了機械手臂作為重要的工業機械設備。機械手臂在自動化生產中展現了出色的效能，可以執行繁重、危險或精細的工作任務，提高了生產效率並降低了勞動強度。所以現今許多產業界皆以此模式來建構生產線。近年驅動手臂的方式也日新月異，從傳統的氣油壓與電動機驅動，到現今採用線趨動控制、串聯式網路控制等的新式技術。因傳統的控制器漸漸無法跟上現代化的硬體設備需求，其缺點包含了擴充性、開放性、系統的相容性不高以及無網路通信等諸多缺點，使得網路控制成了主要發展趨勢。

機械手臂是模仿人類手臂功能與動作的自動化機械，藉由多關節連接，能夠在二維或三維空間下運動或線性移動。透過加入鏡頭模仿人類的眼睛，以取得工作環境的圖像，再透過影像處理分析目標物位置提供手臂夾取資訊，由此一來便不用將零件放置在固定區塊，就能達到抓取的任務。目前機械手臂已廣泛應用在各個工業製造的領域，例如：機械、電子、傳感器、驅動、控制等。

第一台工業用機械手臂為「機械手臂之父」Joseph F.Engelberger，於1956年發明，同時期各國致力發展，三軸機械手臂的技術最為成熟，至今在工業依然非常實用，可用於簡易加工及搬運，為自動化系統的重要指標。總結來說，機械手臂作為一種具有多自由度和精確運動控制能力的機械裝置，在工業、服務機器人、醫療保健和農業等領域得到了廣泛的應用。隨著人工智能和機器學習的不斷發展，機械手臂的智能化和自主性不斷提升。然而，機械手臂仍然面臨著成本、可靠性、感知和智能化等方面的挑戰。未來的研究和發展將致力於解決這些挑戰，進一步提升機械手臂的性能和應用價值，實現更廣泛的應用。



圖 1 Unimate

1.2 研究動機與目的

隨產業逐漸走向高科技，機械手臂也開始出現在精密工業與電子業的生產線上，其對於精度、速度及靈活度相較傳統重工業要求更高，傳統的三軸機械手臂已無法滿足高科技產業，但機械手臂的研發與製造需要龐大資金，因此本研究利用簡易馬達、旋轉基座與鏡頭造出低成本兼具實用性的機械手臂。

製作機械手臂的主要目的為取代人力，去執行高危險性或重複性高而單調的工作。為了能使機械手臂更靈活能自主判別物體位置並夾取，本研究使用視覺處理系統，並將關節數提升至五軸，企圖提高靈活度，使機械手臂工作更有效率。此外，在生產線上用視覺檢測出不良品並挑出，以提升產品良率，降低人為因素對產品品質影響，也可以檢測機械工作有無異常，在發現異常後即時停止並修復，於此便不需要人類時時盯著生產線，達到真正意義上的自動化生產。

1.3 論文架構

本論文共分為五章節：

第一章、緒論

說明研究背景、研究動機與目的、論文架構。

第二章、系統架構

說明本論文的基本硬體與軟體架構，以及其原理與微處理控制板。

第三章、視覺處理

說明影像處理流程與原理，如：RGB 與 HSV 色彩空間、色彩空間轉換、特徵點匹配演算法、二值化、膨脹與侵蝕、雙目測距等。

第四章、運動學

說明機械手臂的運動原理，如剛體運動、歐拉旋轉定理、正向運動學與反向運動學推導等。

第五章、實驗結果

說明本次實驗主要任務，主要是要利用影像處理與機器人運動學的結合，來展現機械手臂的運動狀況與完成指定任務等實驗結果。

第六章、結語與未來展望

說明本實驗的結語與未來展望，問題與解決方法。

第二章、系統軟硬體架構

本論文將機械手臂系統主要分為硬體機構、系統控制與視覺回授等三大系統，此章將依序做簡要的介紹。

2.1 系統架構簡介

本論文主要實現視覺回授控制多軸機器手臂，使機器手臂具有自我組織的能力，在非人為操控下完成指定的工作。依照圖 3 機器手臂系統方塊圖，可以很直觀看到機械手臂中，資訊是如何傳遞。硬體機構部份是由攝影機、伺服馬達、機器手臂所構成；系統控制電路核心是使用 Arduino Uno 控制板，透過 OpenCV 做即時的影像處理，利用影像將目標物的相機座標算出後，再經由一轉換矩陣轉為世界座標，最後將此目標物座標傳給 Arduino 系統進而控制手臂去夾取目標物。

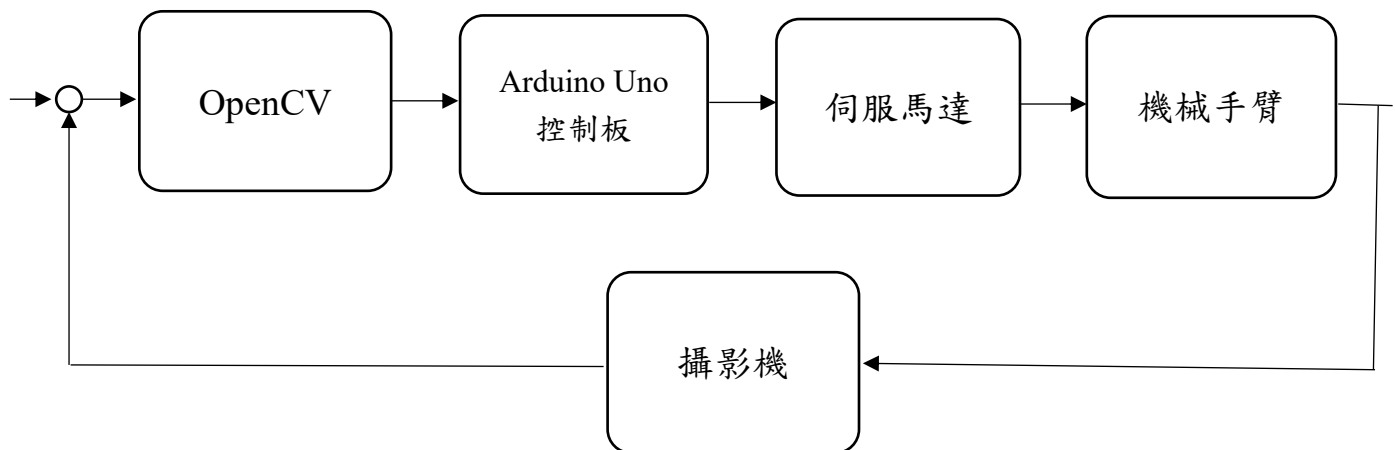


圖 2 機器手臂系統方塊圖

2.2 機器手臂硬體介紹

2.2.1 機械臂尺寸與架構

本論文使用機型為哈爾濱奧松機器人科技公司所產之機械臂，型號為；底部旋轉盤內一馬達控制旋轉角度，使夾頭能與目標物在同一垂直面上；而手臂上三個關節可在該平面中移動使夾頭靠近目標物；前端一馬達控制夾頭角度；最後末端一馬達控制夾頭開合，為一五軸機械手臂：

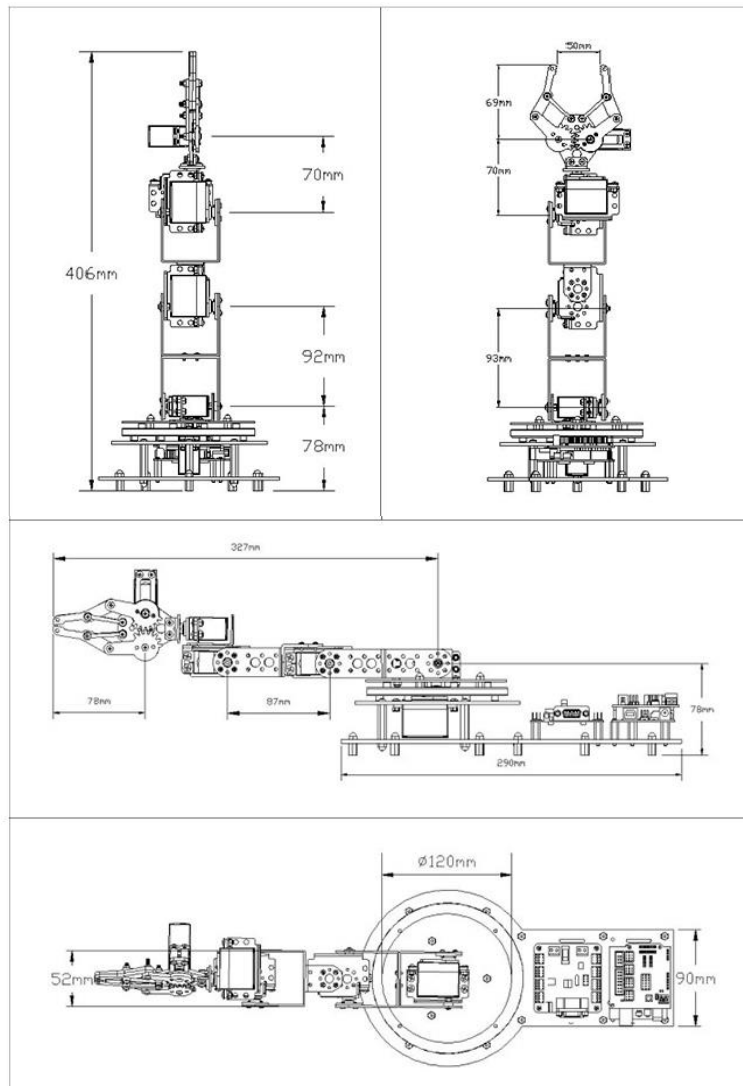


圖 3 AS-6DOF 尺寸模型

2.2.1.1 手臂改良

舵機臂改良:舊型舵機臂易脆化不耐磨，使用久後會影響旋轉精度，因此本論文選用在接口處加強型的厚型搖臂，增強機械臂的精度與耐用度。



圖 4 接口處加強型的厚型搖臂

夾頭改良:由於本論文最後呈現方式為夾取保麗龍球，原夾頭因接觸面積小、摩擦力不夠，導致夾取時易掉落，故在夾頭接觸面貼上泡棉膠，增加接觸面積與摩擦力以解決此問題。



圖 5 夾頭改良實際圖

2.2.2 伺服馬達

MG995



圖 6 MG995 伺服馬達外觀

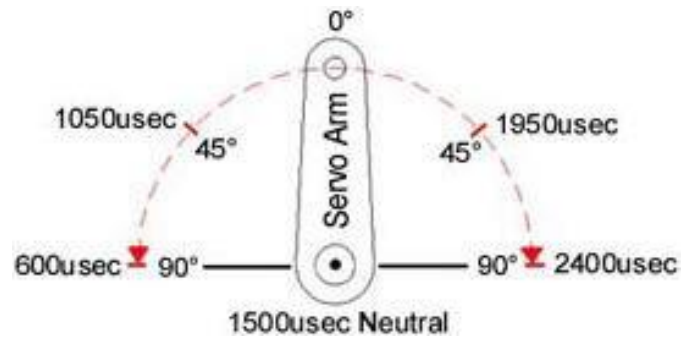


圖 7 MG995 伺服馬達轉角幅度

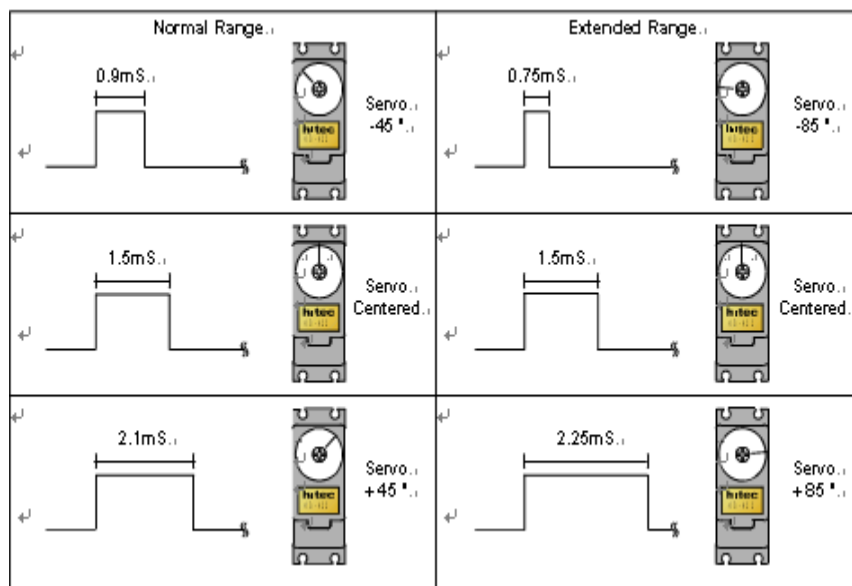


圖 8 MG995 伺服馬達控制脈寬區間示意圖

RB-150MG

RB-150MG 是奧松機器人推出的新一代全金屬齒大扭矩大轉角範圍的機器人專用伺服舵機，其內部採用的電機為直流有刷空心杯電機，內部採用無鐵轉子，具有能量轉換效率高、激活制動響應速度快、運行穩定性可靠、自適應能力強、電磁干擾少等優點，與同等功率的鐵芯電機相比體積小、重量輕；舵機反饋電位器採用導電塑料電位器，其精度和耐磨程度大大優於線繞電位器；電

機控制集成電路(IC)採用數字晶片與功率開關組成H橋電路，電壓控制雙極性驅動方式具有響應速度更快、無響應區範圍小、定位精度高、抗干擾能力強兼容性好等優勢，超越傳統舵機更適合機器人與航模領域廣泛應用。



圖 9 RobotBase RB-150MG 伺服馬達

表一 MG995、RB-150MG 伺服馬達規格

	MG995	RB-150MG
尺寸	40*20*36.5mm	40.8*20.1*38mm
重量	48g	56g
工作電源	3.5V~8.4V	5V~7.2V
轉速	60 度/0.17s(4.8V) 60 度/0.13s(6V)	60 度/0.16s(6V) 60 度/0.14s(7.2V)
扭力	13kg-cm(4.8V) 15kg-cm(6V)	13kg-cm(5V) 15kg-cm(6V) 16kg-cm(7.2V)
工作溫度	-30℃~60℃	0℃~60℃
死區	4us	10us
控制脈波週期	20ms	20ms

2.2.3 Arduino UNO 控制板

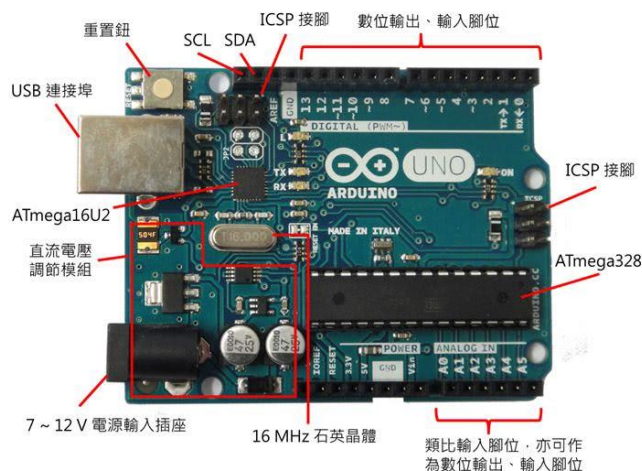


圖 10 Arduino Uno 控制板與腳位簡介(正面)



圖 11 Arduino Uno 控制板(背面)

Arduino Uno 是基於 ATmega328P 的單片機開發板。它有 14 個數字輸入/輸出引腳，其中 6 個可用作脈寬調製輸出，也就是可以輸出可變功率，有助於控制 LED 的亮度。還包含 6 個模擬輸入腳，16MHz 晶振，USB 連接，電源插孔，ICSP 接頭和復位按鈕。作為最入門且最受歡迎的控制板，Arduino Uno 只需一根 USB 線就可以與 IDE 連接，非常適合初學者使用。

表 2 Arduino Uno 規格

名稱	規格
微控制器	ATmega328P
工作電壓	5V
建議輸入電壓	7-12V
限制輸入電壓	6-20V

電源與相關接腳

Arduino Uno 板可通過 USB 連接或外部電源供電。需注意如果供電電壓低於 7V，則 5V 輸出腳位可能電壓不足，電路板可能會變得不穩定。如果使用 12V 以上，電壓調節器可能會過熱並損壞電路板。

接腳列於下表：

表 3 Arduino Uno 接腳

腳位名稱	介紹
VIN	電路板在使用外部電源時的輸入電壓，可以通過此腳位提供電壓。
5V	從電路板上的穩壓器輸出 5V 穩壓電壓。
3V3	由穩壓器產生的 3.3V 電源。最大電流消耗為 50mA。
GND	接地。
IOREF	讓其他設備知道 Arduino 控制板的運作電壓。

類比輸入腳位

A0 至 A5 (A 代表 Analog) 可用來接受類比電壓輸入，但不能輸出類比電壓。每隻腳預設會將 0V 到 5V 轉換為 0 至 1023 的數值。對於輸出電壓為其他範圍的電路模組，可以透過 AREF 腳與函式 `analogReference()`，來提供參考電壓。A0 至 A5 也可作為數位輸出、輸入腳位使用，此時 A0 至 A5 分別可視為 D14 至 D19。

數位輸出、輸入腳位

D0 到 D13 (D 代表 Digital) 可輸出高電位 5V 與低電位 0V 的數位訊號，也可接受數位訊號。

2.2.4 Webcam



圖 12 KINYO PCM-537 攝影機

本次實驗的視覺系統所採用的 Webcam 為 KINYO 的網路攝影機，其型號為 PCM-537，實體如下圖所示：

攝影機架設在機械手臂的末端處，用來偵測目標物位置。此攝影機具 6 顆 LED 照明燈，在光線不足的環境下，可開啟 LED 燈補強光線。其使用 USB 做介面傳輸影像，在本論文中直接與電腦連接，擷取影像後傳到 PC 端進行影像處

理，並分析影像作特徵點匹配，再將影像交由給 OpenCV 處理後計算距離與座標。

PCM-537 透過 USB2.0 進行訊號傳輸，最高動態攝影解析度為 2560x1920，幀率為每秒 30 幀(30FPS)，色彩為 24bit 真彩。

表 4 網路攝影機規格表

廠牌	KINYO PCM-537 網路攝影機
產品尺寸	(W)56×(H)76×(D)53mm
鏡頭	5P 鏡頭
感應器	CMOS Image sensor
畫素	700 萬畫素
傳輸	USB 2.0
FPS	30 FPS/sec

2.3 軟體開發環境與工具

2.3.1 開發環境 PyCharm

PyCharm 是一個跨平台開發環境，主要用於 Python 語言開發，由捷克公司 JetBrains 開發，提供代碼分析、圖形化除錯器，整合測試器、整合版本控制系統(Vcs)，並支援使用 Django 進行網頁開發。擁有 Web 開發、支援網路框架、Python 剖析器、遠端開發、支援資料庫與 SQL 等許多進階功能。

2.3.2 OpenCV

OpenCV 全名為 Open Source Computer Vision，由 Intel 公司開發出來的開源圖形演算法跨平台函式庫，可以製作圖片、視訊、矩陣運算、統計、圖論、資料儲存等相關設計。其應用領域有影像辨識、圖形辨識、電腦圖學、立體視覺、動靜態影像處理、機器學習等。對於影像與視訊 分析處理提供超越 500 種最佳演算法與圖形處理的資料結構，可以整合 C 語言函式庫，做出更加強大的應用。

本論文以 PyCharm 做為開發環境，Python 為主要語法，使用 OpenCV，介面如下：

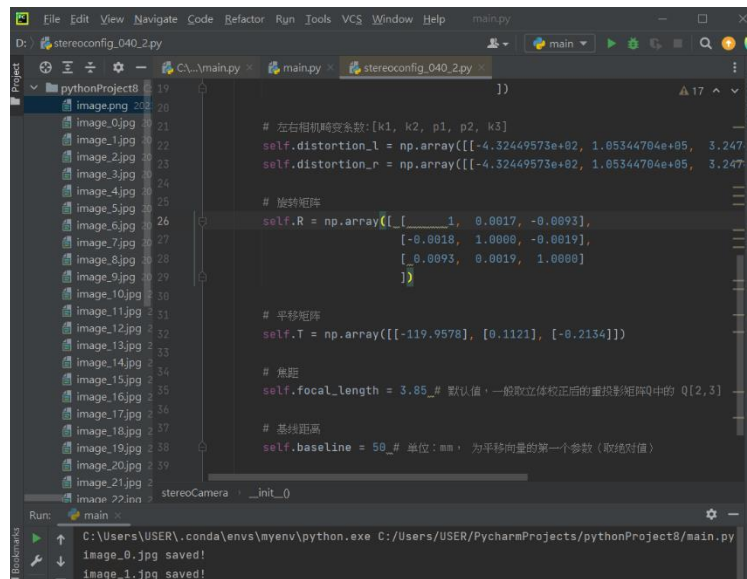


圖 13 PyCharm 上使用 OpenCV 操作介面

2.3.3 Arduino IDE 工具

Arduino IDE 工具可以從 Arduino 官方網站找到最新版本的開發環境，並且免費下載。Arduino 的開發環境是一套用 Java 編寫的跨平台應用軟體，主要使用 C 語言及 C++ 撰寫程式，也提供常見的輸入/輸出函式的 Wiring。在此平台編寫的程式被稱為「sketch」，而其中一個典型的 Arduino C/C++ sketch 程式會包含以下兩個函式，它們會在編譯後合成為 main() 函式：

1. setup()：在程式執行開始時會執行一次，用於初始化設定。
2. loop()：直到 Arduino 硬體關閉前會重複執行函式內的程式碼。

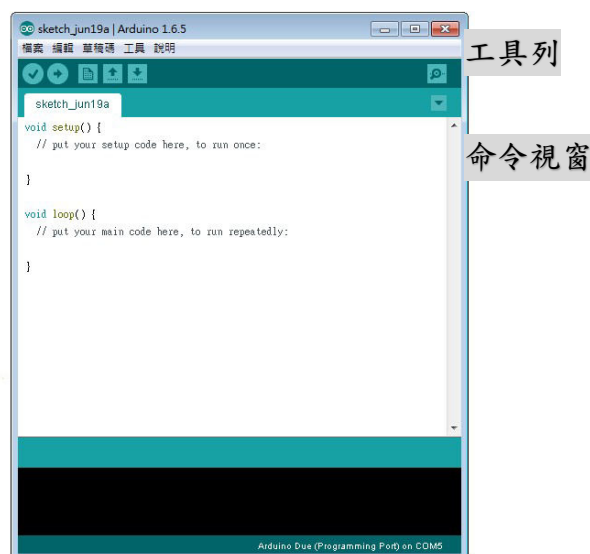


圖 14 Arduino 編譯程式操作介面

以 AVR-GCC 編譯器與 avr-libc、AVRlib 函式庫所構成的開放原始碼之開發環境，可以讓使用者輕易的上手編寫程式，再將程式上傳至 I/O 板上，此工具可以運行於各種作業軟體，如：Windows、Mac、Linux 等。

第三章、視覺處理

在研發機器手臂的過程中，我們期望機器手臂能夠模仿人類，自行完成複雜任務，因而搭配許多種不同的感測器，如：紅外線、超音波、壓力等，其中，能對陌生環境蒐集最多訊息的為視覺回授。若機器擁有視覺功能，就能針對環境的改變作即時的調整，提高機器手臂的穩定性及準確性，縮小後端系統控制的誤差，並大大提升機械手臂對陌生環境的適應力。本章會講述處理影像的流程、並介紹色彩空間與轉換、二值化的原理及目的、以及對影像的膨脹、侵蝕等處理。

3.1 影像處理流程

本論文使用視覺獲取目標物與環境資訊，流程如下：首先打開 Webcam，在畫面中找到紅色圓球時會將其輪廓圈出，按下”p”鍵將影像擷取下來進行處理並計算。辨別紅色是將 RGB 的影像格式資料彙整後，轉換成 HSV 色彩空間描述，再經由閾值的調整來擷取所需要區域的 HSV 數值區間，將此區間的資訊經過二值化轉換以降低資料處理量，提升影像處理的效率。再根據圖片中圓球進行質心運算，找出其的質心位置。

而實際的影像處理過程中，環境與背景都會影響影像運算的精準度，所以除了降低背景顏色、不均勻光線等干擾外，我們也可以在影像運算之前，主動對影像進行降低雜訊。本論文使用的是膨脹與侵蝕的演算法技術，其可以有效地抑制大部分雜訊，使之後的質心運算可以不受到雜訊的影響，降低計算結果的誤差。

影像處理流程圖如圖 15 所示，藉由攝影機擷取影像，分析其 RGB 影像資訊，並轉換成 HSV 色彩模型，並藉由調整閾值來擷取所需要顏色區域後，再將其二值化後，利用膨脹與侵蝕演算法降低雜訊，最後得到目標物在畫面中的質心位置，針對此座標做深度計算。

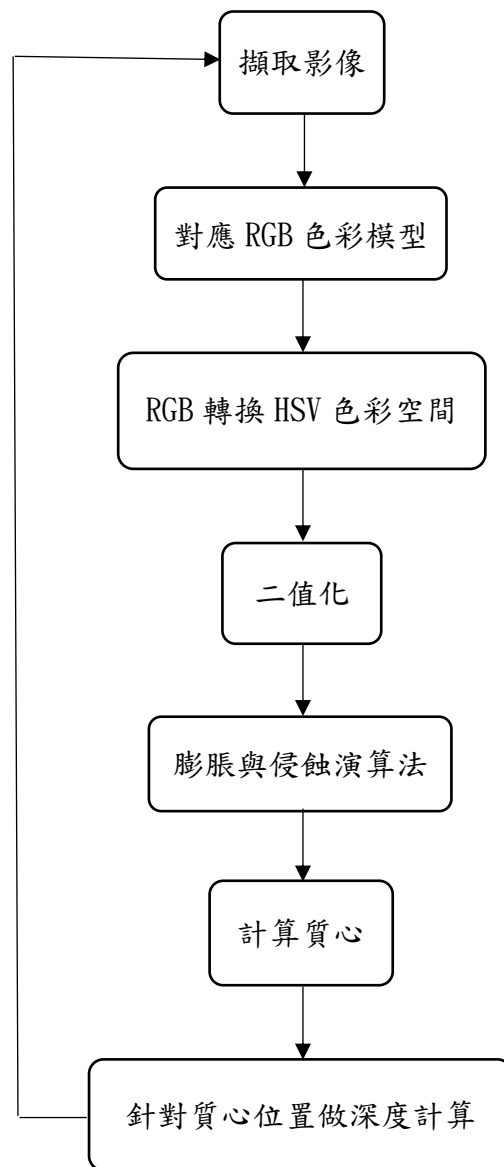


圖 15 影像處理流程圖

3.2 色彩空間與色彩模型

色彩模型(color model)也被稱作色彩空間(color space)或色彩系統(color system)，是利用數值區分不同顏色的數學模型，常見的有 RGB、HSV、CMYK 等的模型。以下介紹本論文使用到的 RGB 模型與 HSV 模型。

3.2.1 RGB 色彩空間

不同人對相同物理量的色彩也會有不同的感受，但機器視覺必須建立一套統一定義量化的標準，因此便建立了色彩模型加以有系統的描述顏色。

RGB 色彩空間是利用顏色相混比例表示色彩空間，其原理與人類視覺系統的原理類似。人類對色彩的判斷是基於眼睛對光的反應，因此色彩屬於一種光的特性。目前最常見的 RGB 色彩空間是 sRGB，在消費級的數位相機、高畫質攝錄影機、電腦顯示器領域尤為明顯。

RGB 顏色空間利用三個顏色分量的線性組合來表示顏色，任何顏色都與這三個分量有關，而且這三個分量是高度相關的，所以連續變換顏色時並不直觀，想對圖像的顏色進行調整需要更改這三個分量才行。而自然環境下獲取的圖像容易受自然光照、遮擋和陰影等情況的影響，即對亮度比較敏感。且人眼對於這三種顏色分量的敏感程度是不一樣的，在單色中，人眼對紅色最不敏感，藍色最敏感，所以 RGB 顏色空間是一種均勻性較差的顏色空間。如果顏色的相似性直接用歐氏距離來度量，其結果與人眼視覺會有較大的偏差。我們很難用精確的三個分量數值來表示某個顏色，所以，RGB 顏色空間適合於顯示系統，卻並不適合於圖像處理。

所謂的 RGB 色彩模型指紅(Red)、綠(Green)與藍(Blue)三種原色所組成的色彩空間，RGB 色彩模型的混色屬於加法混色，混色的方式如圖 27 所示。

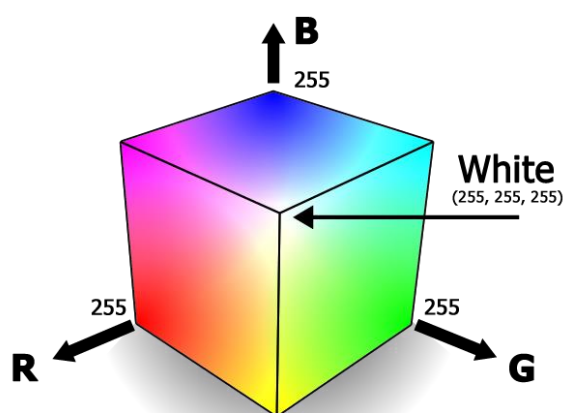


圖 16 RGB 色彩空間

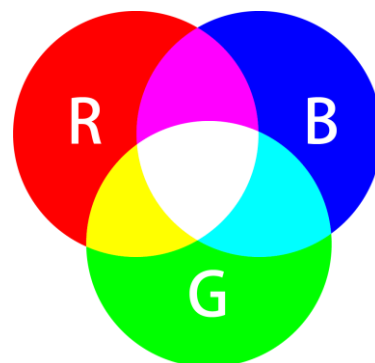


圖 17 RGB 色彩之加法混色示意圖

像素是影像的最小取樣單位，每一個像素點對應的顏色值稱為子像素(sub-pixel)。在 RGB 色彩模型中，子像素為紅、綠、藍三原色所組成的數值，當子像素的數值為(0,0,0)時為黑色，為(255,255,255)則為白色。

3.2.2 HSV 色彩空間

HSV 系統在影像處理上為一種常見的色彩空間，對比 RGB 色彩空間，更接近人們對彩色的感知經驗。非常直觀地表達顏色的色調、鮮豔程度和明暗程度，方便進行顏色的對比，而計算量上保持簡單。HSV 分別代表色調(Hue)、飽和度(Saturation)、亮度(Value)，其優點為三個元素之間的相關性很小，所以我們很容易去控制。

在數學上 HSV 色彩模型的空間可以用圓柱來表示，如圖 29，但概念上可以視為一個圓錐體或六角錐體，如圖 30。

- 色調(Hue)為色彩的基本屬性，在 HSV 空間中為圍繞中心軸的圓盤狀模型，其範圍從 0 度至 360 度，從紅色開始按逆時針方向旋轉，0 度為紅色，120 度為綠色，240 度為藍色，通過改變色調的值可以在範圍內選擇不同顏色。
- 飽和度(Saturation)指色彩的純度，也就是混和白光的多寡。在 HSV 空間中為水平方向表示，越往外飽和度越高，數值越高表示顏色越深，取 0~1 的數值，在色彩模型空間中代表顏色至中心軸的距離，0 表示灰色(無飽和度)，1 則表示完全飽和的純色，通過調整飽和度可以改變顏色的鮮豔程度。
- 亮度(Value)代表顏色明暗度，在色彩模型空間中代表沿中心軸平行方向的高度，範圍為 0 至 255，越往上越亮，0 則為黑色。

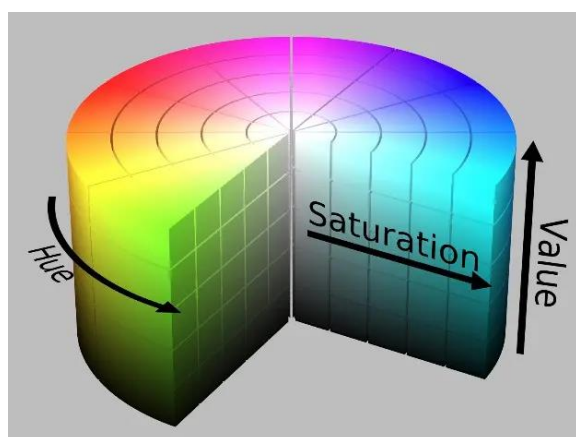


圖 18 HSV 色彩空間

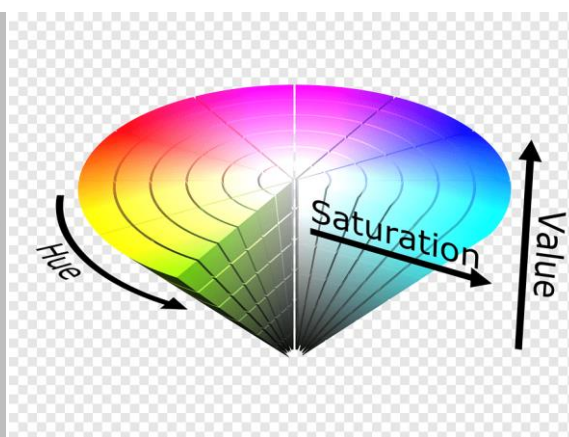


圖 19 HSV 圓錐體示意圖

由於在實際情況下使用 RGB 色彩空間時，模型容易受到干擾，所以在影像分析處理時，必須同時考慮此三種參數才能達到需求，因此本實驗在影像處理前會先將 RGB 色彩空間轉換成 HSV 色彩空間來描述影像的色彩資訊。

表 5 HSV 色彩閾值

	黑	灰	白	红		橙	黄	绿	青	蓝	紫
hmin	0	0	0	0	156	11	26	35	78	100	125
hmax	180	180	180	10	180	25	34;	77	99	124	155
smin	0	0	0	43		43	43	43	43	43	43
smax	255	43	30	255		255	255	255	255	255	255
vmin	0	46	221	46		46	46	46	46	46	46
vmax	46	220	255	255		255	255	255	255	255	255

3.2.3 色彩空間轉換

在數學上，將 RGB 色彩空間轉換成 HSV 色彩空間的座標變換公式如下：
色調(H)為

$$H = \begin{cases} 360^\circ - \theta, G < B \\ \theta, G \geq B \end{cases} \quad (3.1)$$

其中

$$\theta = \cos^{-1} \left(\frac{(R - G) + (R - B)}{2[(R - G)^2 + (R - B)(G - B)]^{\frac{1}{2}}} \right) \quad (3.2)$$

飽和度(S)為

$$S = \begin{cases} 0, & \text{if } \max(R, G, B) = 0 \\ \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}, & \text{otherwise} \end{cases} \quad (3.3)$$

亮度(V)為

$$V = \max(R, G, B) \quad (3.4)$$

HSV 色彩空間將顏色資訊以色調、飽和度和亮度來表示，在本論文的影像處理或顏色辨識上皆較 RGB 空間更準確，但因為計算過程較為繁雜，所以不適合用在即時(real-time)且解析度大的系統中。

3.2.4 二值化

在本篇論文中，當攝影機擷取影像後，會將原始的 RGB 影像轉換成 HSV 色彩空間。此步驟必須透過二值化來將目標物分割出來，把彩色圖像轉為二值圖像，除了可以大大減少計算量，提升計算效率，也能使計算結果更為準確。對於一張平面影像，我們可以將每個像素用一個二維函數 $f(x,y)$ 表示， x 與 y 為像素(Pixel)座標， f 則為該像素的函數值，在本論文中使用 HSV 空間中的 H 分量為主，界定是否為紅色。二值化處理簡言之就是將目標色彩資訊設為門檻值，大於等於該值為設為亮部(255)，小於該值則為暗部(0)。數學式可表達如下：

$$f(x,y) = \begin{cases} 255, & f(x,y) \geq T \\ 0, & f(x,y) < T \end{cases} \quad (3.5)$$

其中 T 為有目標色彩門檻值。

當整張圖片只有乾淨的兩個值(255 及 0)，就很適合我們做輪廓偵測等運算。在本論文中，有助於找出影像中紅色圓球的輪廓，以及找出球的質心等，對於之後計算其深度及座標能大大提升精準度。圖 31 為設定紅色為目標做二值化的前後比較圖。



圖 20 二值化前後比較圖

3.2.5 膨脹與侵蝕

膨脹(Dilation)與侵蝕(Erosion)是形態學的基礎，許多形態學處理的演算法都建立在此兩種原始計算上。在本論文中，圖像做完二值化後，部分區塊會不完整，此時透過膨脹與侵蝕處理後，做到消除雜訊，進而提升面積與質心計算的準確度。

膨脹

膨脹(Dilation)是利用空間中兩個像素集合，A 集合與 B 集合，當 A 集合被 B 集合膨脹時，可以用 $A \oplus B$ 表示，其中 A 為輸入影像，B 為結構元素。

定義如下：

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \phi\} \quad (3.6)$$

A 為輸入影像，也就是被膨脹的目標。

B 為結構元素(structuring element)或稱核心(kernel)，是膨脹目標的基本元素。

$(\hat{B})_z$ 為 B 的反射且平移 Z 單位，B 集合對於其原點的反射並平移 Z 單位，跟 A 不是空集合的部分做 AND 運算。

當輸入像素及其周圍像素相對於結構元素為 1 的像素值有一個以上為 255 時，則將輸入像素的值設為 255。運算結果能對影像的邊界進行擴張，其作用可以使圖像大小往外增加一個像素(3*3)、平滑目標物的邊緣以及減少或填充對象之間的距離。處理效果如下圖所示：

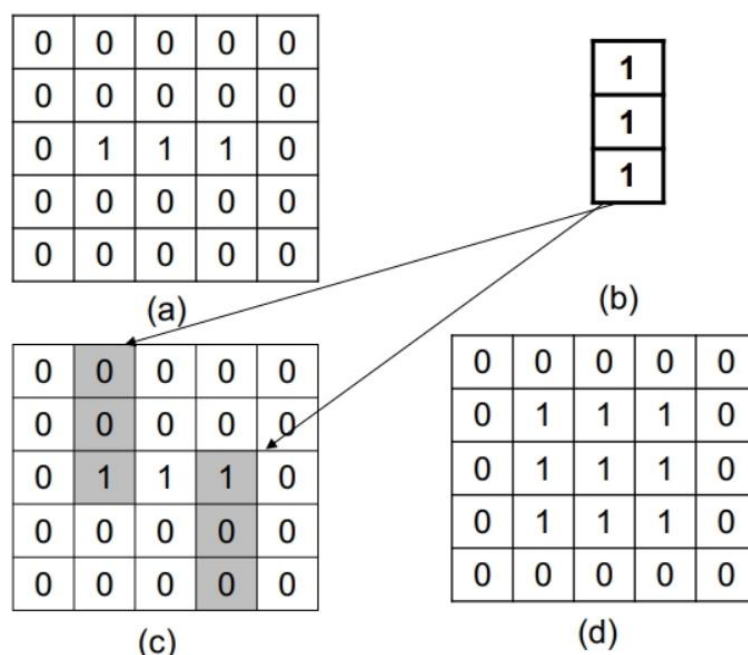


圖 21 膨脹流程示意圖

侵蝕

侵蝕(Erosion)是利用空間中兩像素集合，A 集合與 B 集合，當 A 集合被 B 集合侵蝕，可用 $A \ominus B$ 表示，A 集合為輸入影像，B 集合為結構元素。

定義如下：

$$A \ominus B = \{z \mid (\hat{B})_z \subseteq A\} \quad (3.7)$$

A 為輸入影像，也就是被侵蝕的目標。

B 為結構元素(structuring element)，是侵蝕目標的基本元素。

$(\hat{B})_z$ 為 B 的反射且平移 Z 單位，A 被 B 侵蝕使得 B 位移 Z 後，仍包含在 A 中的所有 Z 點的集合。

侵蝕可以消除目標的邊界點，使邊界向內收縮，在影像處理上用來消除雜訊。卷積核沿著圖像滑動，如果卷積核對應的原圖的所有像素值為 1，那麼中心元素就保持原來的值，否則變為零。侵蝕主要應用在去除白噪聲，也可以斷開連在一起的物體。處理效果如下圖所示：

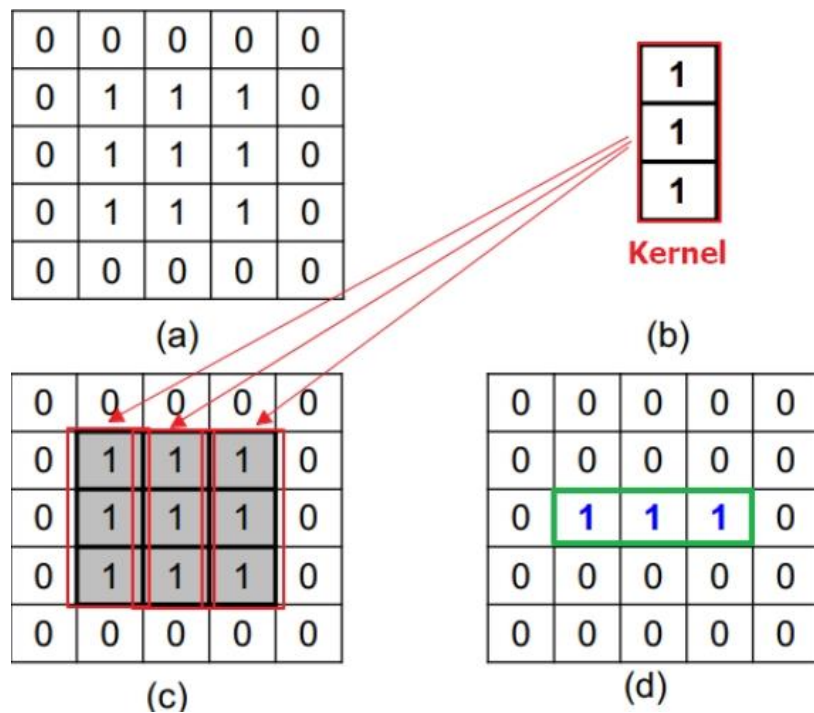


圖 22 侵蝕流程示意圖

藉由設定的 kernel(圖 b)去滾動每個框格，僅有三個單位完全符合 kernel 的，才會留下中間值的數字 1，最後剩下綠色方框。

3.3 影像深度

在求得座標前需要得到目標物的深度，本論文使用雙目測距來獲得，其結果會較單目測距(直接輸入圓球直徑)，來的更靈活且準確。本章節將會說明雙目測距的原理，最後將求得的深度轉換為座標的過程。

3.3.1 雙目測距

雙目測距的基本原理主要是三角測量原理，即通過視差來判定物體的遠近。流程如下：

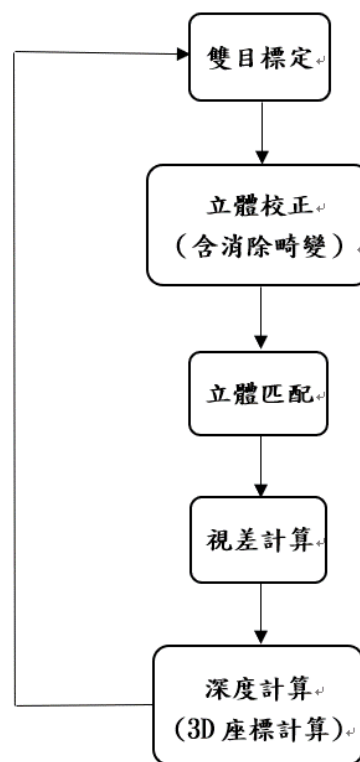


圖 23 雙目測距流程圖

雙目標定的目標是獲得左右兩個相機的內參(在本論文為使用同一相機)、外參和畸變係數，其中內參包括左右相機的 f_x , f_y , c_x , c_y ，外參包括左相機相對於右相機的旋轉矩陣 R 和平移矩陣 T 。畸變係數包括徑向畸變係數 (k_1 , k_2 , k_3) 和切向畸變係數 (p_1 , p_2)。

徑向畸變以數學式表示為：

$$x_{\text{distorted}} = x(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (3.8)$$

$$y_{\text{distorted}} = y(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (3.9)$$

切向畸變以數學式表示為：

$$x_{\text{distorted}} = x + [2p_1xy + p_2(r^2 + 2x^2)] \quad (3.10)$$

$$y_{\text{distorted}} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \quad (3.11)$$

其中，x 和 y 均為平面坐標，按照上述數學式，需要的畸變參數(Distortion coefficients)總共有 5 個：

$$\text{Distortion coefficients} = (k_1 \ k_2 \ k_3 \ p_1 \ p_2) \quad (3.12)$$

雙目相機標定和單目相機標定最主要的區別就是，雙目相機需標定出左右相機座標系間的相對關係，通常會用旋轉矩陣 R 和平移矩陣 T 來描述。計算方式如下：

假設空間中有一點 P，其在全局坐標系下的坐標為 P_w ，其在左右相機座標系下的坐標可以表示為：

$$P_l = R_l P_w + T_l \quad (3.13)$$

$$P_r = R_r P_w + T_r \quad (3.14)$$

其中：

$$P_r = R P_l + T \quad (3.15)$$

綜合上式可推得：

$$R = R_r R_l^T \quad (3.16)$$

$$T = T_r - R T_l \quad (3.17)$$

其中， R_l 、 T_l 為左攝像頭經過單目標定得到的相對標定物的旋轉矩陣和平移向量， R_r 、 T_r 為右攝像頭經過單目標定得到的相對標定物的旋轉矩陣和平移向量。左右相機分別進行單目標定，就可以分別得到 R_l 、 T_l 、 R_r 、 T_r ，帶入上式就可以求出左右相機之間的旋轉矩陣 R 和平移 T。

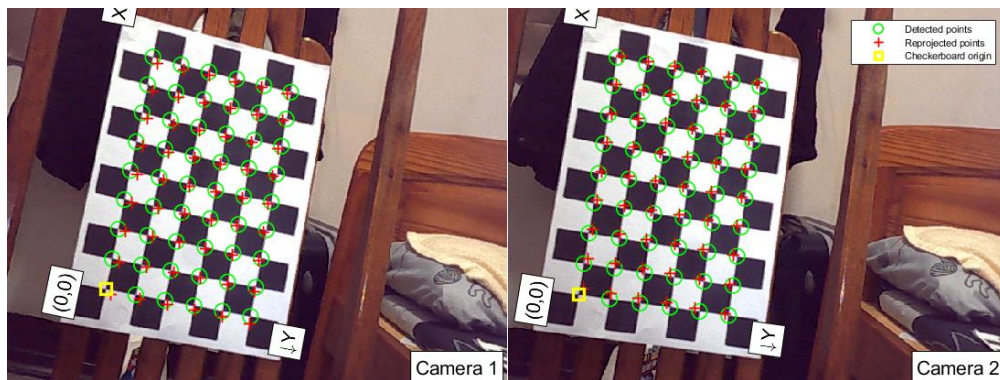


圖 24 使用 MATLAB 對方格紙做雙目標定

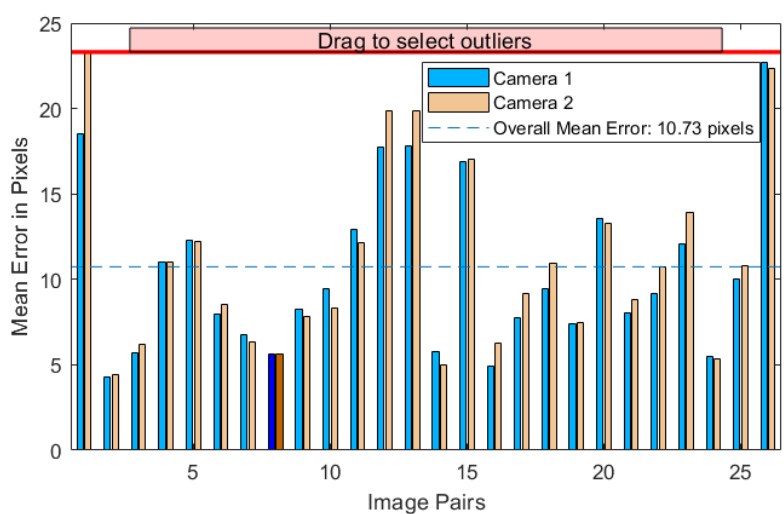


圖 25 雙目標定配對誤差直方圖

立體校正(Recitification)的目的是將拍攝於同一場景的左右兩個視圖進行數學上的投影變換，使得兩個圖像平面共面且平行於基線。經過此校正後，兩幅圖中的極線就會完全水平，使空間中的同一個點在左右兩幅圖中的像素位置位於同一行。在達到共面行對準以後就可以應用三角原理計算距離。

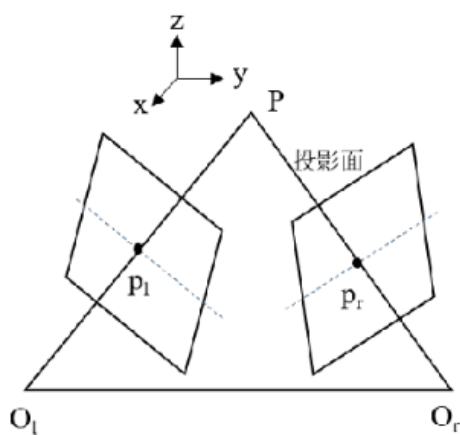


圖 26 校正前示意圖

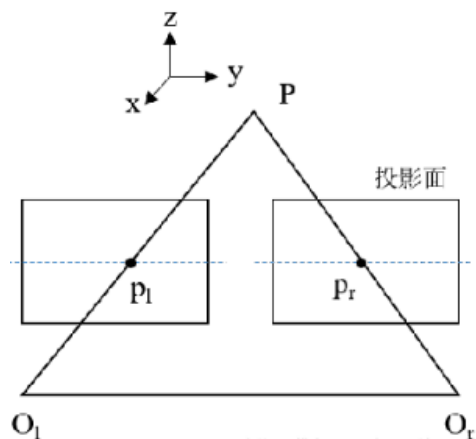


圖 27 校正後示意圖

立體匹配的目的是為左圖中的每一個像素點在右圖中找到其對應點（世界中相同的物理點），這樣就可以計算出視差：

$$\text{disparity} = u_l - u_r \quad (3.18)$$

其中， u_l 和 u_r 分別是兩個對應點在圖像中的列坐標。

立體匹配的計算過程大致可分為：匹配代價計算、代價聚合、視差優化、視差細化。而立體匹配的困難在於：

1. 圖像中可能存在重複紋理和弱紋理，這些區域很難匹配正確。
2. 由於左右相機的拍攝位置不同，圖像中幾乎必然存在遮擋區域，在遮擋區域，左圖中有一些像素點在右圖中並沒有對應的點。
3. 左右相機所接收的光照情況不同。
4. 過度曝光區域難以匹配。
5. 傾斜表面、彎曲表面、非朗伯體表面。
6. 較高的圖像噪聲等。

常用的立體匹配方法基本上可分為局部方法和全局方法，局部方法計算量小，但匹配質量相對較低，全局方法省略了代價聚合而採用了優化能量函數的方法，匹配質量較高，但是計算量也比較大。

本論文使用 SGBM（semi-global block matching）演算法，屬於全局匹配演算法，立體匹配的效果明顯好於局部匹配演算法，但是同時複雜度上也要遠遠大於局部匹配演算法。其匹配原理為通過選取每個像素點的 disparity，組成一個 disparity map，設置一個和 disparity map 相關的全局能量函數，使這個能量函數最小化，以達到求解每個像素最優 disparity 的目的。

SGBM 算法流程如下：

1. 預處理:使用 sobel 算子對源圖像進行處理，並將經 sobel 算子處理後的圖像映射為新圖像，並得到圖像的梯度信息用於後續的代價計算。
2. 代價計算:使用採樣方法對經預處理得到的圖像梯度信息計算梯度代價、使用採樣方法對源圖像計算 SAD 代價。
3. 動態規劃:默認四條路徑，並對路徑規劃的參數 P1，P2 進行設置（包括 P1、P2、cn（圖像通道數量）以及 SADWindowSize（SAD 窗口大小）的設置。
4. 後處理:包括唯一性檢測、亞像素插值、左右一致性檢測、連通區域的檢測。

視差圖是將立體匹配計算得到的視差值映射到相對應的 0~255 像素的灰度空間，在視差圖中，灰度值越高代表像素點越亮，說明物體與相機的距離越

近。相反的，灰度值越低代表像素點越暗，說明物理與攝像機的距離越遠。若視差圖雜訊過多可以對其進行濾波，作用是將稀疏視差轉變為稠密視差，還能降低視差圖噪聲，改善視差圖的視覺效果，但就較依賴初始視差圖的品質。得到了視差圖之後，就可以計算像素深度了，公式如下：

$$depth = \frac{f \times b}{d + (C_{xr} - C_{xl})} \quad (3.19)$$

其中 f 為焦距長度（像素焦距）， b 為基線長度， d 為視差， C_{xl} 與 C_{xr} 為兩個相機主點的列坐標。



圖 28 拍攝紅色圓球的左右視圖



圖 29 轉為灰度圖後放在同一水平面進行比對

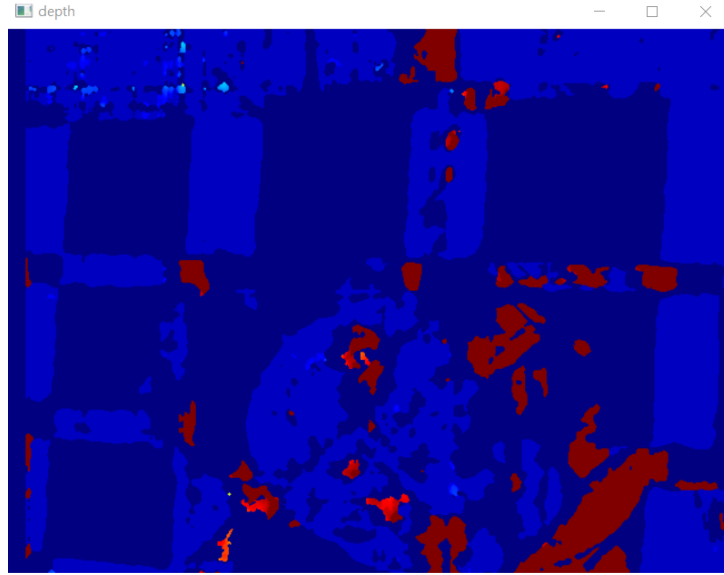


圖 30 深度圖

雙目測距的精度取決於該點處估計的視差 d 的精度。假設視差 d 的誤差恆定，當測量距離越遠，得到的深度精度則越差，因此使用雙目測距不適合測量太遠的目標。若想提高深度的準確度，需要提高相機的基線距離，但是基線距離越大，左右視圖的重疊區域就會變小，就會提高立體匹配的難度。

3.3.2 深度轉換相機座標

經上述過程得到深度後，要將深度轉為實際物體與相機間距離(Z)，才能將此距離在轉為相機座標。

將深度值轉換為物體到相機的距離(Z)公式如下：

$$Z = \frac{f \times b}{d} \quad (3.20)$$

其中， B 是攝像機的基線長度，表示攝像機兩個鏡頭之間的距離。

由相機與物體間距離(Z)計算物體的相機座標公式如下：

假設物體在相機座標系中的位置為 (X_c, Y_c, Z_c) ，其中 $(0, 0, 0)$ 表示攝像機的光學中心。根據幾何關係，我們可以將物體的座標表示為：

$$X_c = \frac{(x - C_x) \times Z}{f} \quad (3.21)$$

$$Y_c = \frac{(y - C_y) \times Z}{f} \quad (3.22)$$

其中， (x, y) 是物體在畫面上的座標， (C_x, C_y) 是畫面的中心座標。

第四章、運動學

本章將說明機器人的正反運動學，當已知目標物座標，要做的就是將夾頭座標相等於目標物座標，使手臂能成功夾取物體。而此章節將介紹如何利用正反運動學，從手臂末端座標回推每個關節需轉的角度。

4.1 空間中的剛體

剛體指的是在物理學上，具有一定大小與形狀的物體，組成此物體的每個質點之間的相對位置是固定的。因此不管在任何一種情況下，只要知道剛體上任意兩點的運動情形，則整個剛體的運動即可完全掌控。本論文中，機器手臂的連桿可以視為剛體，而在連桿兩端的關節旋轉軸上各有一組座標系，透過座標系定義可以描述剛體在空間中的位置。如圖 36 所示，假設座標系 $O_1-X_1Y_1Z_1$ 於空間中存在著一組剛體，且剛體依附在該座標系的原點，沿著 P 的運動方向平移到的新座標位置，再旋轉成新的參考座標系 $O_2-X_2Y_2Z_2$ ，則此剛體 F 可表示為：

$$F = \begin{bmatrix} {}^1X_2 & {}^1Y_2 & {}^1Z_2 & P \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.1)$$

其中， 1X_2 ， 1Y_2 ， 1Z_2 是剛體相對於座標系的 $O_1-X_1Y_1Z_1$ 的方向， P 為剛體相對於參考座標系的 $O_1-X_1Y_1Z_1$ 的相對位置。

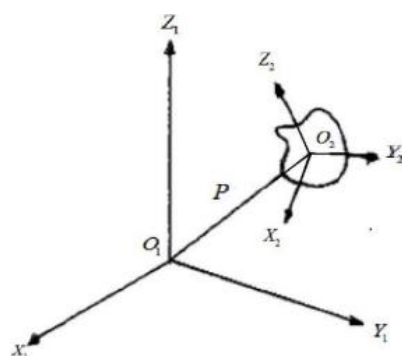


圖 31 空間中剛體的運動示意圖

剛體的運動可分為平移與旋轉，彼此具有獨立性。兩座標系間平移的運動只需移動原始座標系 $O_1-X_1Y_1Z_1$ 的原點位置，即可得到新的座標系 $O_2-X_2Y_2Z_2$ ；而兩座標系間旋轉可利用歐拉旋轉定理(Euler's rotation theorem)來轉換。

4.2 歐拉旋轉定理(Euler's rotation theorem)

歐拉旋轉定理是 Leonhard Paul Euler 基於三維空間剛體動力學，所提出的重要數學定理。其說明一剛體在三維空間中發生位移，剛體內至少有一點固定不動，由該點延伸出一固定軸，則剛體的運動會等價於繞該固定軸的旋轉運動。

簡言之，任何三維剛體的旋轉，都可以分解為 X、Y、Z 三軸的旋轉。這三個軸所旋轉的角度稱為歐拉角，以 (α, β, γ) 表示。

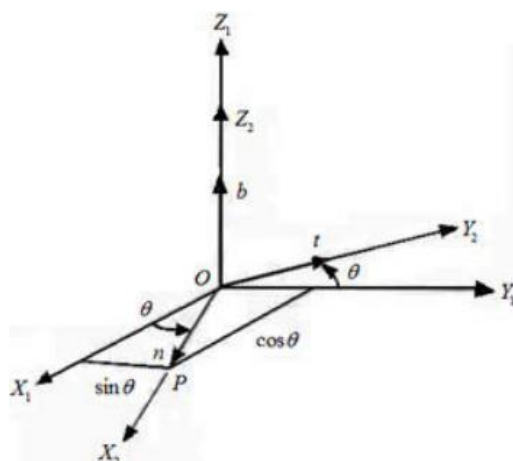


圖 32 以 Z 軸為基準遵循右手定則之旋轉示意圖

參考圖 26，空間中存在著一參考座標系 $O_1-X_1Y_1Z_1$ 與目標點 P，目標點 P 在 $O_1-X_1Y_1Z_1$ 座標系中的座標為 (P_{X1}, P_{Y1}, P_{Z1}) ，此座標系以 Z 軸為基準逆時針旋轉 θ 角，成為新的參考座標系 $O_2-X_2Y_2Z_2$ ，目標點 P 在 $O_2-X_2Y_2Z_2$ 座標系的座標為 (P_{X2}, P_{Y2}, P_{Z2}) ，兩座標轉換關係如下：

$$P_{X1} = l_1 - l_2 = P_{X2} \cos \theta - P_{Y2} \sin \theta \quad (4.2)$$

$$P_{Y1} = l_3 + l_4 = P_{X2} \sin \theta + P_{Y2} \cos \theta \quad (4.3)$$

$$P_{Z1} = P_{Z2} \quad (4.4)$$

將上式以矩陣形式來表示如下：

$$\begin{bmatrix} P_{X1} \\ P_{Y1} \\ P_{Z1} \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_{X2} \\ P_{Y2} \\ P_{Z2} \end{bmatrix} \quad (4.5)$$

令 $R_{X\theta}$ 為參考座標系以 X 軸為基準逆時針旋轉 θ 角度的旋轉矩陣； $R_{Y\theta}$ 為參考座標系以 Y 軸為基準逆時針旋轉 θ 角度的旋轉矩陣； $R_{Z\theta}$ 為參考座標系以 Z 軸為基準逆時針旋轉 θ 角度的旋轉矩陣。則 $R_{X\theta}$ 、 $R_{Y\theta}$ 、 $R_{Z\theta}$ 可分別表示如下：

$$R_{X\theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad (4.6)$$

$$R_{Y\theta} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \quad (4.7)$$

$$R_{Z\theta} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.8)$$

4.3 反向運動學(Inverse kinematics)

反向運動學利用機械手臂末端的世界座標與每軸的長度，回推每個關節需轉的角度。圖 28 虛線為與座標軸垂直之輔助線，實線為機械手臂軸， θ_1 與 θ_2 為角度。 w_1 與 w_2 代表手臂長度， H 為末端點高度， L 為末端點距原點距離，以上皆為已知的值。 θ_x 、 θ_y 、 θ_z 為計算過程中的輔助角度。

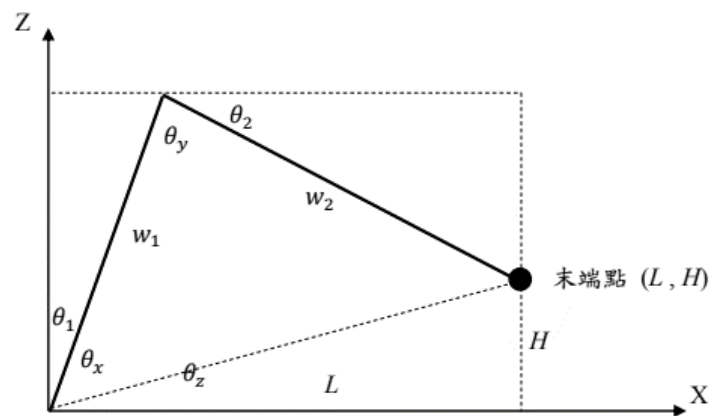


圖 33 $H \geq 0$ 之二維空間逆向運動學模型

圖 28 為假設末端點之高度大於等於零($H \geq 0$)之示意圖。可看出

$$\theta_z = \text{atan}\left(\frac{H}{L}\right) \quad (4.9)$$

$$\theta_x = \arccos\left(\frac{w_1^2 - w_2^2 + (L^2 + H^2)}{2 * w_1 * \sqrt{L^2 + H^2}}\right) \quad (4.10)$$

$$\text{求得}\theta_x、\theta_z\text{後，可得 } \theta_1 = 90 - \theta_x - \theta_y \quad (4.11)$$

$$\theta_2 = 180 - \theta_y - (90 - \theta_1) \quad (4.12)$$

$$\theta_y = \arccos\left(\frac{w_1^2 + w_2^2 - (L^2 + H^2)}{2 * w_1 * w_2}\right) \quad (4.13)$$

若 $H=0$ 則

$$\theta_x = \arccos\left(\frac{w_1^2 + L^2 - w_2^2}{2 * w_1 * L}\right) \quad (4.14)$$

$$\theta_y = \arccos\left(\frac{w_1^2 + w_2^2 - L^2}{2 * w_1 * w_2}\right) \quad (4.15)$$

4.4 正向運動學 (Forward kinematics)

逆向運動學為正向運動學的反推式，可以藉由逆向運動學求得欲求關節角度後代入正向運動學驗證末端點的座標(L,H)。

參考圖 2-1、圖 2-2，將參數 H 與 L 視為欲求目標， θ_1 與 θ_2 為已知參數。圖 2-1 為假設末端點之高度大於等於零($H \geq 0$)，可得

$$L = w_1 * \sin(\theta_1) + w_2 * \cos(\theta_2) \quad (4.16)$$

$$Z = w_1 * \cos(\theta_1) - w_2 * \sin(\theta_2) \quad (4.17)$$

若 $H=0$ 則

$$w_1 * \cos(\theta_1) = w_2 * \sin(\theta_2) \quad (4.18)$$

4.5 座標系轉換

4.5.1 相機座標轉圖像座標

參考圖 34，相機內參作用是將相機座標轉為圖像座標， $P_c = (x_c, y_c, z_c)'$ 在圖像座標中的像素點 m 用齊次座標表示為 $m = (x_p, y_p, 1)'$ 。從圖中三角關係可得：

$$\begin{cases} x_p = \frac{f x_c}{z_c} \\ y_p = \frac{f y_c}{z_c} \end{cases} \quad (4.19)$$

若主點在圖像座標中的座標為 $p = (x_0, y_0, 1)'$ 寫成矩陣形式為：

$$z_c m = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} P_c \quad (4.20)$$

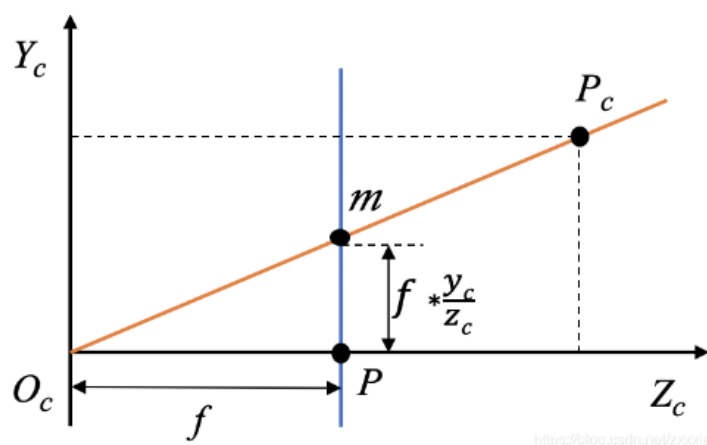


圖 34 相機座標轉換圖像座標

若想將圖像座標用 pixel 表示，首先假設一個像素的長寬為 dx, dy ，設像素座標 $P_{pix} = (u, v, l)'$ ，則用矩陣可表示為：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & 0 \\ 0 & \frac{1}{dy} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} \quad (4.21)$$

結合上述兩種座標系變換，可得出相機座標到圖像 pixel 座標的變換矩陣 K 為：

$$K = \begin{bmatrix} \frac{1}{dx} & 0 & 0 \\ 0 & \frac{1}{dy} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.22)$$

其中， K 為內參矩陣。

4.5.2 相機座標轉世界座標

接著找出相機座標系與世界座標系的轉換，就能使世界座標與圖像 pixel 座標直接溝通。若相機座標系中一點 $P_c = (x_c, y_c, z_c)$ ，我們可以藉由旋轉矩陣 R 和平移矩陣 T 來得到該點的世界座標 $P_w = (x_w, y_w, z_w)$ ：

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = L_w \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.23)$$

其中， R 為 3×3 旋轉矩陣、 T 為 3×1 平移矩陣、 0^T 為 $(0,0,0)$ ，簡化後以 L_w 表示 4×4 矩陣。

綜上所述，可以得到世界座標轉換到圖像 pixel 座標的過程為：

$$z_c \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = z_c \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.24)$$

故可得：

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = z_c \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.25)$$

第五章、實驗結果

5.1 實驗一:立體校正輸出深度圖

5.1.1 實驗目的

此實驗目的是利用雙目標定後的鏡頭，拍下的兩張距離基線長的並包含紅球在內的圖片，將兩張照片進行去校正後，進行立體匹配，進而獲得影像深度。在實驗一中，我們除了使用 MATLAB 來標定相機外，也會將得到的參數及轉換矩陣輸入 pycharm 計算，進而匹配出紅球所在位置之深度，我們將會以標出紅色圓球的質心作為此次實驗的目標。

5.1.2 實驗成果分析

此次實驗中，我們令相機分別向右及向左轉 5 度，不會在移動基線長度時使紅色圓球超出鏡頭範圍。標定場景如下所示：



圖 35 相機標定場景

stereoParams.RotationOfCamera2		
1	2	3
0.9907	0.0018	0.1362
-0.0023	1.0000	0.0036
-0.1362	-0.0039	0.9907

圖 36 兩相機座標旋轉矩陣 R

stereoParams.TranslationOfCamera2		
1	2	3
-28.1573	1.2512	-56.7518

圖 37 兩相機座標平移矩陣 T

stereoParams.CameraParameters1.IntrinsicMatrix			
1	2	3	4
1.3432e+03	0	0	
-33.3649	1.3368e+03	0	
271.6028	285.1937	1	

圖 38 左相機內參矩陣

stereoParams.CameraParameters2.IntrinsicMatrix			
1	2	3	4
1.2653e+03	0	0	
6.8175	1.2699e+03	0	
283.5088	269.3237	1	

圖 39 右相機內參矩陣

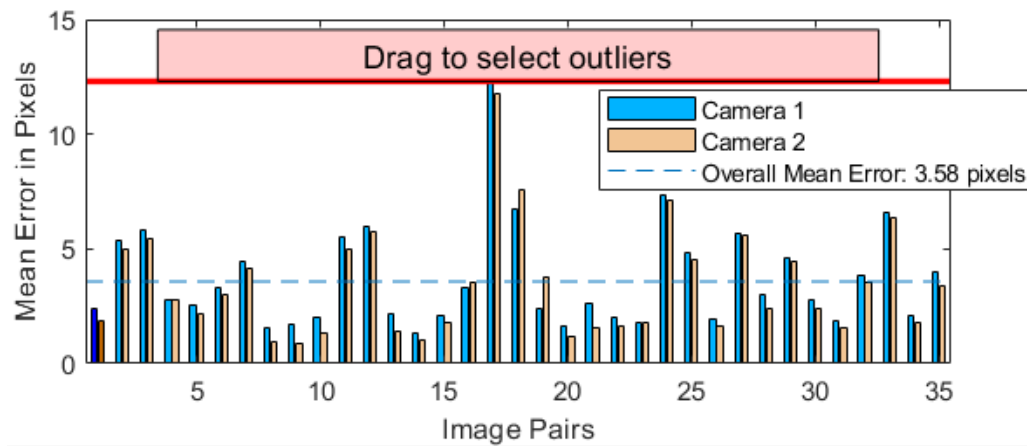


圖 40 原始 35 組相片配對誤差直方圖

上述結果我們會將配對結果較差的相片組刪除，留下 Mean Error Pixels 小於 3 的相片組，以增加相機參數的準確性。刪除後結果如下。

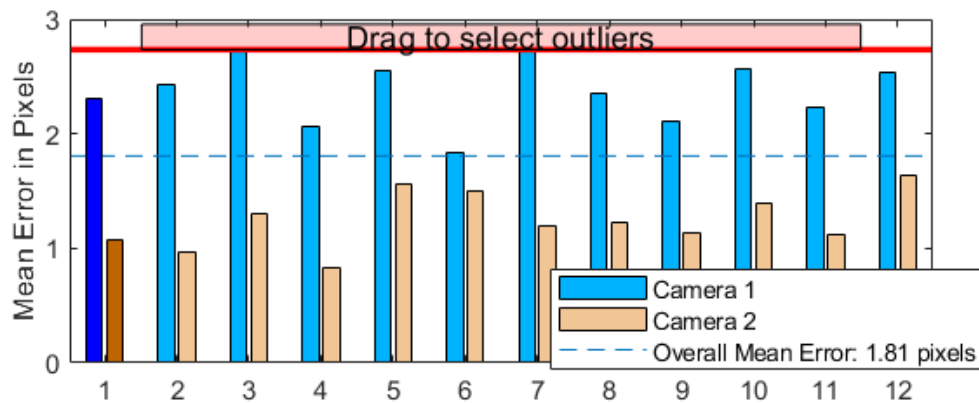


圖 41 刪除誤差大相片組後留下 12 組較精準配對直方圖

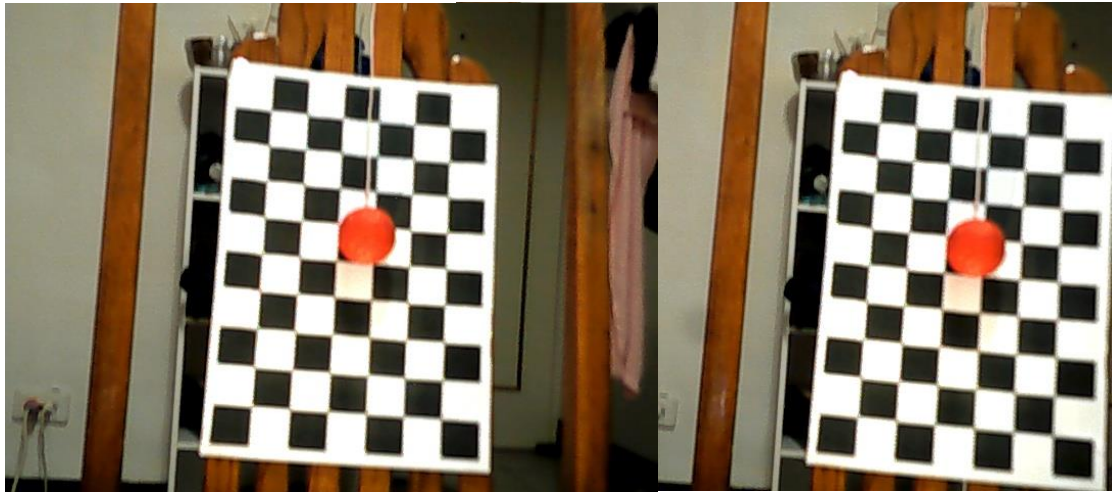


圖 42 左右相機拍攝原始相片組

將原始相片輸入後，我們會將對相片做去畸變等校正，並繪製水平線將兩張相片放置於同一水平方便比對，最後輸出圖片的灰度。輸出結果如下：

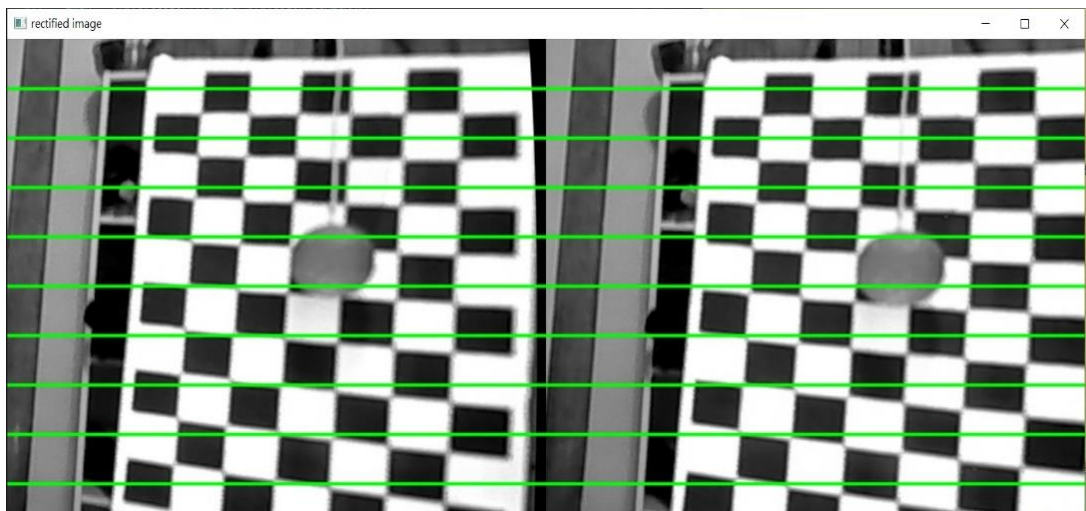


圖 43 立體校正相片組同水平面灰度圖

將兩圖相同像素處重疊配對，在將原灰度圖轉為 RGB 形式，形成深度圖，並計算視差，結果如下：

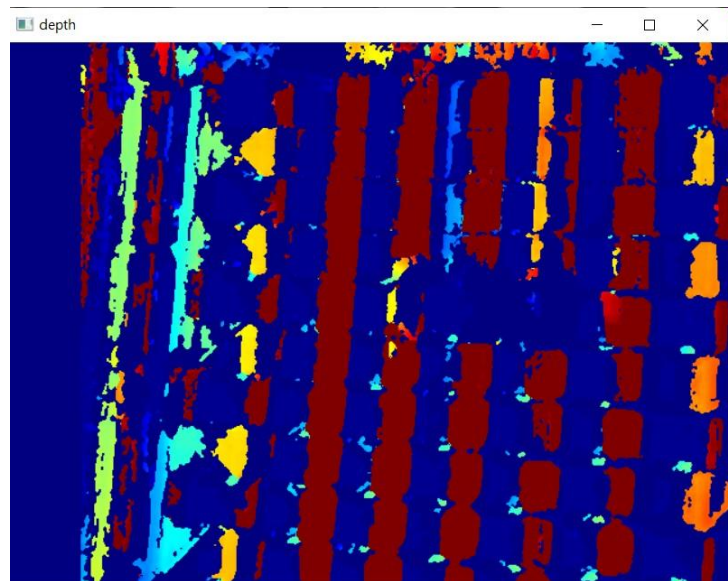


圖 44 影像深度圖

可以看出黑白方格部分有印出(圖中紅色部分)，中間空缺藍色則為紅色圓球位置。

最後輸出紅色圓球的相機座標：

```
i = 479 j = 510 x = -0.47750916973212554 y = 0.5226468524431225 z = 2.97029708913471 (m)
```

圖 45 紅色圓球的相機座標

5.2 實驗二:視覺回授隻機械手臂夾取紅色圓球

5.2.1 實驗目的

實驗一最終獲得了紅色圓球的深度及相機座標，接著就可以利用 4.5 節，相機座標與世界座標的座標系轉換，得出紅色圓球質心的世界座標。此次實驗目的為實現及時性的視覺回授，並控制機械手臂去完成任務。此次實驗設計自主便是紅色圓球座標並夾取，過程利用正反運動學去推導機械手臂各軸轉動角度及世界座標。

5.2.2 實驗成果分析

電路連接如下，再依序將各 pin 腳接上伺服馬達。

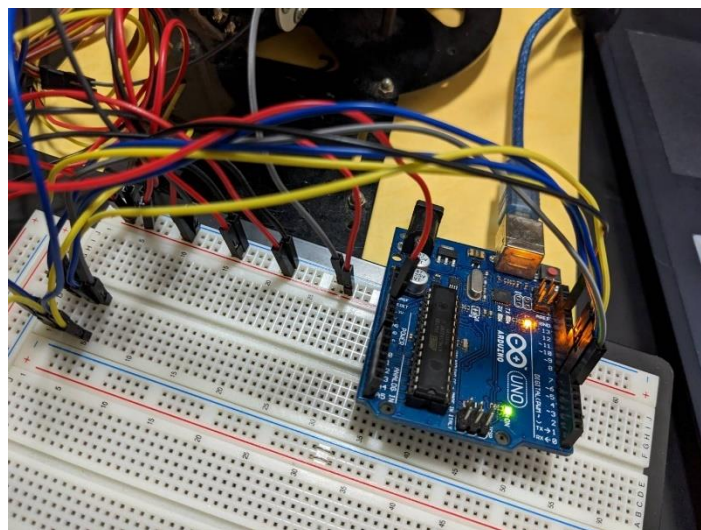


圖 46 伺服馬達連接電路板之電路

若想提高此次實驗成功率，則需要提高實驗一立體匹配的精準度。夾取實驗紀錄如下：

表 6 夾取實驗結果分析

夾取實驗分析	第一次	第二次	第三次	第四次	第五次	第六次	第七次	第八次
深度誤差是否 $\leq 10\%$	否	是	是	是	否	是	否	否
是否碰到紅球	否	是	是	是	是	是	否	是
是否成功夾取	否	否	否	是	否	否	否	否

由表可看出，實驗八次，深度誤差 $\leq 10\%$ 的機率為 62.5%，而在此條件下，夾頭可以碰到紅球的機率為 100%，而成功夾取率為 12.5%。

此次實驗的結果成功率非常低，其誤差產生主要原因為：

1. 使用機械手臂拍照時的轉動角度過小，由於距離較短，所以僅是轉 10 度就會讓紅色圓球跑出鏡頭外。但若旋轉角度過小會導致空間中三角形的角度太小，使得計算出的 \cos 值誤差被放大，進而造成深度不夠準確。
2. 實驗用的機械手臂因尺寸緣故，只能伸展 20cm，但鏡頭在 20cm 下只能拍攝到 5*4 的黑白方格紙，導致沒有足夠的點可以匹配，進而造成深度不準確，但若將距離拉遠則可改善深度的精確度，卻會導致機械手臂到目標物的距離過遠無法正常夾取。
3. 因此次實驗同時控制 5 顆伺服馬達，所以只用電腦 USB 輸出的電流過低，使得馬達有時無力，有時會不正常抖動，進而影響實驗的準確性。

改善方法如下：

1. 使用更大尺寸的機械手臂，使距離能拉得更遠，兩相機的角度也能增大。
2. 將配對用的黑白方格紙邊長縮小，以增加同面積可配對的角點。
3. 額外供給電流使馬達運作更為穩定且有力。

第六章、結論與未來展望

6.1 結論

本論文主要做影像辨識，立體匹配，及機械手臂的關節運動等，完成一個能基於視覺回授作自主辨識及夾取的機械手臂系統。在此次實驗中結合了許多不同領域的專業，在設計實驗的過程常會遇到互相矛盾的情況，如：在計算雙目測距的三角形時，我們希望雙目的距離大一些，可以避免小角度史 \cos 值發生誤差，但若分得太開卻會導致兩相機沒有可以立體匹配的重複點，使匹配的準確度下降；抑或者是使用特徵點匹配時若背景過於雜亂，會影響匹配準確度，但若背景太過單一或具有重複紋路，也會導致配對錯誤進而使準確度下降，所以每種方法都有其優缺點要去取捨。

若未來若研習相關知識我會繼續深入此次沒用到的包括 D-H 轉換矩陣，可以同時分析兩軸，雖原因較艱澀，寫出矩陣卻大大減少了推正反運動學的時間，非常方便。再者我想將 Arduino 與 Pycharm 平台用通訊協定做對接，讓我實驗時不會因要在多平台跳轉而手忙腳亂。

6.2 未來展望

這兩年迎來隨著 AI 的大躍進，ChatGPT 的出現勢必是對現今世界發展的奇異點，未來 AI 侵入我們生活的腳步只會越來越快。目前業界已有許多先進機械手臂具有視覺回授的功能，目的就是能更及時的處理非預期狀況。未來若加入 AI 的學習技能，機械手臂將不在只是能處理重複性高的任務，而是任何危險性高、高精密度、抑或者是需要設計的產業，都能交由 AI 的機械手臂完成。

我認為此次實驗最可貴的並非結果的數據，而是在實驗過程中不斷去探索、嘗試每個不同方法的原理與結果，有時會迷茫的研究許久才發現不適用此次實驗，但這並不代表這些時間是毫無價值的，或許未來到研究所甚至是職場都會在重新接觸到這些演算法與控制方法。雖最終兩個實驗僅有實驗一的結果較令人滿意，但也使我反省了實驗設計的流程，教授更是在驗收結果時提點了我許多改進部分，使我在此次專題實驗受益良多。未來研究所我將會持續關注與研究相關議題，希望能比這次更進步，實驗結果更精確。

參考文獻

- [1] J. J. Craig, "Introduction to Robotics: Mechanics & Control," Addison-Wesley, 1989.
- [2] L. W. Tsai, "Robot Analysis - The Mechanics of Serial and Parallel Manipulator," Wiley Interscience, 2005.
- [3] 陳偉海、陳泉柱、張建斌、張穎, "線驅動擬人臂機器人逆向運動學分析," Chinese Journal of Mechanical Engineering, vol.4, pp. 14-21, 2007
- [4] J. Denavit and R. S. Hartengbart, "A Kinematic Notation for Lower Pair Mechanisms Based on Matrices," Transactions of American Society Mechanical Engineers, pp. 215-221, 1955.
- [5] S. B. Niku, "Introduction to Robot-Analysis, Systems, Application," Prentice Hall, 2004.
- [6] 葉忠憲, "移動式機械臂之抓物控制設計," 國立交通大學電機工程學系碩士論文, 2006。
- [7] 程海強, "基於視覺回授並搭載機械手臂之四足移動機器人" 國立臺灣海洋大學電機工程學系碩士論文, 2018。
- [8] 徐柏剛, "六足機器人之步態分析與控制," 國立台灣海洋大學電機工程系 碩士論文, 2013。
- [9] Download:
http://cc.ee.ntu.edu.tw/~rbwu/rapid_content/course/IoT_Intro/Ch1_Arduino.pdf
- [10] Download:
<http://portal.lib.ntnu.edu.tw:8080/server/api/core/bitstreams/10755ea9-0834-4816-90c7-9195d924e175/content>
- [11] Download: <https://www.epcio.com.tw/paper/六軸機械臂之控制理論分析與應用.pdf>
- [12] Download: <https://www.epcio.com.tw/paper/運動學反運動學控制硬體 IP 設計開發.pdf>
- [13] 鍾哲瑋, "多軸機器手臂平台之機構設計與電腦視覺回授控制," 國立台灣海洋大學電機工程系 碩士論文, 2013。
- [14] 張國卿, "光流演算法於影像伺服機械臂應用," 國立台灣海洋大學電機工程系 碩士論文, 2012。
- [15] 葉忠憲, "移動式機械臂之抓物控制設計," 國立交通大學電機工程學系 碩士論文, 2006。
- [16] 洪卉萱, "多軸機器手臂之視覺回授控制" 國立臺灣海洋大學電機工程學系 碩士論文, 2016。

- [17] 劉一忠，“整合雙目視覺及適應性夾爪之六自由度自動分揀機械手臂設計與控制”龍華科技大學電機工程學系 碩士論文，2019。
- [19] 蔡碩倫，“應用深度學習及雙目立體視覺於機械手臂夾取物體作業”國立臺北科技大學工業工程與管理系 碩士論文，2019。
- [19] 蕭志宏，“機械手臂與視覺辨識之研究”國立雲林科技大學電子工程系 晶片與系統組 碩士論文，2019。
- [20] 徐逢廷，“視覺回授機械手臂繪圖系統之實現”國立臺灣海洋大學電機工程學系碩士論文，2017。
- [21] 黃脩瑞，“立體視覺於六軸機械手臂之應用”龍華科技大學機械工程學系
- [22] <https://shop.mirotek.com.tw/arduino/arduino-start-1/> Arduino
- [23] https://zhuanlan.zhihu.com/p/149795622?fbclid=IwAR3KDHxWVd9_EnSImy6zONkSX7jH7LZfg2UmfrCs5-FXE5cKgSKlxV9r1zM D-H 轉換矩陣。
- [24] https://en.wikipedia.org/wiki/Denavit-Hartenberg_parameters?fbclid=IwAR3MfDeTBqaOFMWE46HWrw7L1QflXIA2wnqcIsw6wuv_zHVB-9NA5UREaC4 D-H 轉換矩陣。
- [25] https://zhuanlan.zhihu.com/p/149795622?fbclid=IwAR3mF5CVqluY_cmgO_0-AZuHp7apHeKqvPde19RpvJjtRuxbTAwf6PB_8HQ D-H 轉換矩陣。
- [26] https://blog.csdn.net/maple_2014/article/details/105612912?utm_medium=distribute_pc_relevant.none-task-blog-baidujs_title-2&spm=1001.2101.3001.4242&fbclid=IwAR31Me6XLBf7e_oFY8kR6w9SvFbZiKOjMBDzcxpLripaPRqY-ELAhosVBJS D-H 參數。
- [27] <https://blog.csdn.net/gwplovekimi/article/details/90172544> camera matrix。
- [28] https://www.bilibili.com/video/BV1Q34y1n7ot/?p=5&vd_source=c5175e354f9afc5b1b225fbaf7bec795 camera calibration
- [29] <https://zhuanlan.zhihu.com/p/379269769> 座標系轉換。
- [30] <https://blog.csdn.net/dulingwen/article/details/98071584> 雙目測距。
- [31] <https://www.cnblogs.com/lijin-note/p/16801644.html> 相機內參。
- [32] <https://micro.rohm.com/tw/deviceplus/how-tos/arduino-guide/smart-robotic-arm-part-1-mechanics-and-wiring/> 機械手臂接線。
- [33] [https://medium.com/@ckwang19/slam 學習之路-4-三維空間剛體運動-8a71dcaa9f16](https://medium.com/@ckwang19/slam-學習之路-4-三維空間剛體運動-8a71dcaa9f16) n0 三維空間剛體運動。
- [34] <https://blog.csdn.net/wanggsx918/article/details/23272669> HSV 色彩空間。
- [35] <https://blog.csdn.net/hjxu2016/article/details/77834599> cv2.inRange 函數。
- [36] <https://hackmd.io/@cws0701/B1AxRjijq> 形態學-輪廓處理。
- [37] https://blog.csdn.net/weixin_42216109/article/details/89840323 輪廓處理與繪

製。

[38] <https://chtseng.wordpress.com/2016/12/05/opencv-contour> 輪廓/

[39] <https://sites.google.com/a/ms.ttu.edu.tw/cse2012dance-robot/yan-jiu-cheng-guo/opencv-ruan-ti-she-ji/qin-shi-yu-peng-zhang> 侵蝕與膨脹。

[40] <https://zhuanlan.zhihu.com/p/532797417> 相機標定。

[41] <https://www.guyuehome.com/36265> 相機標定。

[42] <https://developer.aliyun.com/article/666716> 立體匹配。

[43] <https://kknews.cc/zh-hk/tech/e5m9qyq.html> SGBM 立體匹配演算法。

[44] https://blog.csdn.net/zhubaohua_bupt/article/details/51866567 SGBM 立體匹配演算法。

[45] <https://blog.csdn.net/candice5566/article/details/115639723> 視差圖深度轉換。

[46]

https://www.youtube.com/watch?v=h4oVoMvn3EI&list=PLI6pJZaOCtF1YE9wnxSOxFJzoBqub6M-K&index=4&ab_channel=DeltaMOOCx 機器人學。

[47] <https://zhuanlan.zhihu.com/p/32199990> 雙目立體視覺。

[48] <https://blog.csdn.net/dulingwen/article/details/100115157> MATLAB 雙目標定。

[49] <https://blog.csdn.net/LuohenYJ/article/details/104697062> OpenCV 相機標定。

[50] <https://xiaobaidiy.github.io/2020/05/11/robot-arm-4/> 正反運動學。

[51] https://www.youtube.com/watch?v=ODpMM1pn_sc&list=PLI6pJZaOCtF1YE9wnxSOxFJzoBqub6M-K&index=1&ab_channel=DeltaMOOCx D-H 轉換矩陣。

[52] <https://openhome.cc/Gossip/ComputerGraphics/HomogeneousCoordinate.htm> 齊次座標。