

國立臺灣大學工學院機械工程學系

碩士論文

Department of Mechanical Engineering

College of Engineering

National Taiwan University

Master's Thesis



透過 IMU 和相機融合重建身體運動

Reconstruction Body Motion Via Fusion of IMU and Camera

陳珮甄

Pei-Zhen Chen

指導教授：詹魁元博士

Advisor: Kuei-Yuan Chan, Ph.D.

中華民國 113 年 7 月

July, 2024

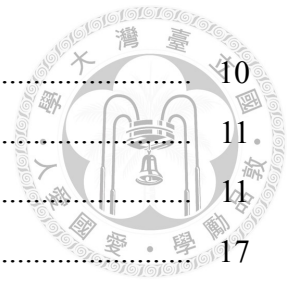




# 目次

目次 .....	i
圖次 .....	iii
表次 .....	iv
第一章 文獻回顧 .....	1
1.1 動作捕捉系統 .....	1
1.1.1 單感測器-光學動作捕捉系統 .....	2
1.1.2 單感測器-慣性動作捕捉系統 .....	3
1.1.3 多感測器動作捕捉系統 .....	3
1.2 個人化骨架 .....	3
1.3 時間對齊 .....	3
1.4 小結 .....	3
第二章 研究方法 .....	5
2.1 系統架設與實驗環境 .....	7
2.1.1 系統架設 .....	7
2.1.2 實驗環境 .....	7
2.2 資料前處理 .....	7
2.2.1 人體立體骨架建立 .....	7
2.2.2 時間軸對齊 .....	9
2.2.3 座標系轉換 .....	9

2.3 探討減少相機數量的可行性 .....	10
2.3.1 實驗方法 .....	11
2.3.2 實驗結果 .....	14
2.3.3 結果與討論 .....	17
2.4 結果可視化 .....	18
2.5 小結 .....	18
參考文獻 .....	19





---

# 圖次

---

2.1	立體骨架建立流程圖 . . . . .	8
2.2	OpenPose 辨識結果 . . . . .	8
2.3	立體骨架建立結果 . . . . .	8
2.4	骨骼座標系定義 . . . . .	9
2.5	IMU local 座標系定義 . . . . .	9
2.6	OpenSim 計算結果 . . . . .	10
2.7	IMU 平移為零 . . . . .	10
2.8	IMU 旋轉為零 . . . . .	10
2.9	相機擺放位置 . . . . .	11
2.10	一台相機到七台相機的組合估計結果 . . . . .	17



---

# 表次

---

2.1	一台相機組合與其估計結果誤差 . . . . .	12
2.2	兩台相機組合與其估計結果誤差 . . . . .	13
2.3	三台相機組合與其估計結果誤差 . . . . .	14
2.4	四台相機組合與其估計結果誤差 . . . . .	15
2.5	五台相機組合與其估計結果誤差 . . . . .	16



---

# 文獻回顧

---

本論文研究主題為評估個人化模型之肌肉參數，在文獻回顧中，首先會介紹關於人體動作量測的技術，包含動作捕捉系統、感測裝置等介紹，接續會討論關於動作模擬與分析的部分，主要切割為三部分，前半部將先引入人體模擬流程與系統介紹，並依照複雜程度進行模型分類，後半部則是關於人體的模擬與分析資訊，最後則會回顧個人化模型的相關文獻，統整現今學者在肌肉參數領域的研究，其針對不同方法來進行介紹，例如以直接量測或是間接估測來取得肌肉參數。第二章節將圍繞在該研究主題的背景與現況來進行探討，最終闡述肌肉參數的重要性。

## 1.1 動作捕捉系統

動作捕捉為現今常用於擷取人體動作、表情、手勢或其他物體動作等資訊的技術，其可應用於運動分析、醫學研究 (cite 復健)、遊戲開發、影片製作等領域，藉由取得的資訊，可進行分析、模擬、辨識等應用。根據感測器的不同，動作捕捉系統可分為單感測器動作捕捉系統及多感測器融合動作捕捉系統，其中單感測器動作捕捉系統又可分為光學動作捕捉系統及慣性動作捕捉系統。而人體姿態 (pose) 則可使用位置資訊 (position) 或朝向資訊 (orientation) 進行定義，位置資訊通常以笛卡爾坐標系描述物體在空間中的位置，因此可使用量測結果為位置資訊的光學動作捕捉系統進行姿態追蹤；而朝向資訊則常以尤拉角 (Euler angles)、旋轉矩陣 (rotation matrix) 或四元數 (quaternions) 來描述物體的朝向狀態，因此可使用量測結果為朝向資訊的慣性動作捕捉系統進行姿態追蹤。以下將針對光學動作捕捉系統、慣性動作捕捉系統、多感測器融合動作捕捉系統，這三大類系統進行介

紹。



### 1.1.1 單感測器-光學動作捕捉系統

光學動作捕捉系統是目前最常見的動作捕捉系統，其原理為透過辨別標記點或特定特徵點的位置來追蹤物體的運動，再進一步由邊際點或特徵點的位置推估物體的姿態。根據受試者身上有無光標記，可將光學動作捕捉系統以有無光標記的分類方法細分為光標記捕捉系統及非光標記捕捉系統。

#### 光標記動作捕捉系統-Vicon

光標記動作捕捉系統，如 Vicon [1] 是目前最常見且最精準的動作捕捉系統，其量測誤差約落在 1 (cm) 的那篇文獻，因此常被當作黃金標準，用於確認其他動作捕捉系統的準確性。光標記動作捕捉系統所需量測設備為多個紅外光攝影機及多個反光標記點，紅外光攝影機可發出紅外光，也可接收反光標記點反射的紅外光，透過計算標記點的位置來推估受試者的姿態。由於光標記動作捕捉系統主要傳遞訊息的媒介為紅外光，需要嚴格控管環境光（例如陽光）的影響，減少環境光造成的雜訊，因此實驗環境多為使用人造光為主要照明的室內，無法在室外進行量測，實驗環境及實驗設置如圖插入 Vicon 實驗環境的圖片所示，且由於需使用多台攝影機，因此器材架設不易，價格昂貴。

#### 非光標記動作捕捉系統

因為光標記動作捕捉系統的限制，非光標記動作捕捉系統逐漸受到重視，其中 OpenPose [2] [3] [4] [5]、MediaPipe [6] 等系統是目前常見的非光標記動作捕捉系統。透過單一或多個攝影機拍攝受試者的影像，利用深度學習技術辨識出受試者在二維平面上的關節點位置，辨識關節點位置的方法分為上到下方法 (top-down method) 及下到上方法 (bottom-up method) 兩種 [7]，上到下方法如圖 Single-Stage Multi-Person Pose Machines 的圖上半，先辨識出受試者的周圍方塊 (detected bounding boxes)，再由周圍方塊內的範圍辨識出關節點位置，例如 MediaPipe 即為上到下方法，其缺點為若在辨識過程中找不到人，無法匡出周圍方塊，則無法辨識出關節點位置，且計算量會隨著出現在畫面中的人數增加而增加；下到上方法如圖 Single-Stage Multi-Person Pose Machines 的圖下半，為先偵測到關節，再將關節組成群組，最後形成特定姿勢的方法，例如 OpenPose 即為下到上方法，其缺點為若受試者沒有完整的出現在畫面中，則有可能辨識點無法組



成一個人的群組而被判斷為一個人，且也有可能會把非人類，但形狀與人類相似的關節點辨識出來。由於最容易被大眾應用的相機為 RGB 相機，因此缺乏受試者的深度資訊，需再透過棋盤格方法相機校正及三角測量計算，取得受試者的三維姿態。

非光標記動作捕捉較不受環境光限制，因此可以解決實驗場地必須設置於室內的限制，可拓展至室外進行量測，也增加了可量測的範圍及動作多樣性，且器材架設相對容易，使用手機也可進行影像錄製，價格也相對較低，但其缺點為易受到遮擋而影響辨識結果，若受試者被遮擋，則無法辨識出被遮擋的關節。

### 1.1.2 單感測器-慣性動作捕捉系統

- IMU - 只能量測到方向，無法量測到位置 - 量測時常拉長後會產生 drift 的問題

### 1.1.3 多感測器動作捕捉系統

- 相機與 IMU 融合 - 3DPW、totalcapture、real-time full-body...、wearable fusing ...

## 1.2 個人化骨架

個人化骨架，還沒頭緒要寫什麼 - 可能可以提及前人的做法，例如使用 Vicon 作為骨架，但是這樣的骨架不夠個人化，所以我要做這些事情

## 1.3 時間對齊

還沒頭緒要寫什麼 - 可能可以提及前人的做法，別人怎麼對齊的等等

## 1.4 小結

描述目前文獻不足的地方，所以我要做這些事情我的部分可能就是描述目前使用的文獻都是使用 Vicon 作為骨架，我想要改善這一點，還有資料對其的地方，要再思考這不分要如何在前面 2.3 的章節先提到，而且可以承接章節 2.2

## - 骨架 - 時間對齊

本章節回顧了與研究主題相關的論文，首先從最基礎的量測技術開始介紹，像是動作捕捉系統、力感測裝置與 EMG 等常見的量測方法，這些儀器能夠量化人體的動作表現，透過數據處理針對特定主題來深入探討，亦可作為模擬的輸入或是輸出的驗證比較，其中量測的精準度將對應用有大幅影響，因此在校正量測結果或是提高估計準確度上皆是重要的議題；接續介紹關於人體動作模擬的工作流程、系統、模型，以及相關的研究文獻，研究探討不再僅限於臨床實驗當中，另外像是肌肉力量、關節扭矩等這些難以量測的資訊，皆可在電腦中快速地取得，獲取資訊的種類增加意味著能探討的問題更加多元，大幅增廣在生物力學上的觀點；最後針對個人化模型議題進行更深入的調查，模型準確度代表著與受試者是否有密切關聯，若僅透過通用模型進行研究，模擬結果將與實際情況有所落差，而個人化模型能對所感興趣的資訊更真實的呈現 [8]，因此根據不同受試者來建立對應的個人化模型是必要的。

個人化模型的建立將會複雜許多，以肌肉參數為例，除了參數取得不易外，還有可能會發生肌肉代償問題 [9]，而在同一條肌肉的參數間，其亦具有參數不可識別性 [10]，在運動軌跡預測作為驗證的文獻中 [11]，於比較軌跡的圖形裡 (文獻 Fig. 6)，這個量級的誤差是否為可接受的？在肌肉代償與參數間會互相抗衡的情況下，微小的誤差就會造成參數的估計錯誤，若誤差大到一定程度，其評估效果也將不及通用模型，故在估計與驗證參數的過程是十分重要的 [12]。

本論文著重在肌肉參數的估計中，對於上述的問題進行深入探討與驗證，提出一套適用的估計參數研究方法，並以已建立好的上肢模型作為驗證，除了進行參數估計外，也會呈現參數間抗衡的範例，來顯現該問題的重要性。



---

## 研究方法

---

本研究目的是透過最佳化方法估計肌肉參數，省去醫療器材量測之高成本，藉以生成個人化模型，應用於復健規劃、運動訓練與輔具設計等領域。本研究假設為已知一肌肉骨骼模型，該模型除了欲評估之肌肉參數外，其餘幾何特徵、肌肉參數等資訊皆為已知，而其可透過 OpenSim 軟體來執行任何模擬。研究方法主要分為三大主軸，介紹如下：

- **主軸一：敏感度分析**

透過擾動肌肉參數，來計算出任務與參數間之敏感度，其指標是藉由預測任務之誤差來表示。敏感度高意味著該肌肉和該任務具有高相關性，其結果提供最佳化與模型驗證之任務挑選的參考依據。

- **主軸二：多運動軌跡預測最佳化**

從敏感度分析挑選數個適當任務作為輸入，並將多運動軌跡預測任務之平均預測誤差視作目標函數，藉由最佳化方法最小化平均預測誤差，以此估計出最佳模型之肌肉參數。

- **主軸三：模型驗證**

從敏感度分析挑選適當任務作為輸入，指定最佳模型完成並檢視其預測誤差，確認該模型是否於其它任務仍具有低誤差表現，以此驗證模型正確性。

本研究之流程圖如圖所示，首先輸入標準模型和分析參數來執行敏感度分析，藉由全因子實驗設計法來評估出所有任務之敏感度指標，接下來挑選適合的組合集作為最佳化之評估任務，以生成最佳模型，最終透過高敏感度任務進行模型驗證，來評估最佳模型之正確性。若於模型驗證階段成功，代表評估正確，並結束

程式執行；若於模型驗證階段失敗，則返回至最佳化步驟重新評估；若模型驗證失敗次數過多，則返回至敏感度分析流程，重新挑選合適的評估任務。程式碼將發布於 <https://github.com/solab-ntu/MuscleParamEstimation> 該網址中，執行細節可檢視程式碼中的註解說明。





## 2.1 系統架設與實驗環境

### 2.1.1 系統架設

123123

#### 硬體-相機擺放

123123

#### 軟體-xsens

123123

#### 軟體-OpenPose

123123

#### 感測器融合方法

IMU image pytorch

### 2.1.2 實驗環境

123123

## 2.2 資料前處理

利用章節 3.1 介紹的系統蒐集完實驗資料後，接下來將進行資料前處理，以利後續進行感測器融合。資料前處理分為人體立體骨架建立、時間軸對齊、座標系轉換三個部分，將於本章節進行討論。

### 2.2.1 人體立體骨架建立

在文獻 [13] 中，作者使用 Vicon 系統所提供的立體骨架資料，並以其為基礎，進行 IMU 計算；但是在非實驗室的環境中無法使用 Vicon 進行量測，進而取得立體骨架資料。因此本研究透過 Pose2Sim [14] [15] [16] 提出之方法，使用影像辨識技術，辨識每一視角中人體的關節點在影像中的位置，並利用相機校正技術建立出每一關節在空間中的位置，以此方式建立人體立體骨架。章節 ??將會利用

TotalCapture Dataset [17] 提供的影像資料，進行影像辨識與三角測量計算，並將結果與 TotalCapture Dataset 提供之 Vicon 立體骨架資料比對，以驗證此方法的可行性。



## 建立方法

本方法流程如圖 2.1 所示。以 T pose 的姿勢拍攝影片後，進入後續軟體端處理流程。首先，使用 OpenPose [2] [3] [4] [5] 影像辨識技術辨識每一視角人體關節點在影像中的位置 (使用 BODY-25B model)，辨識結果如圖 2.2 所示，其可辨識出左右肩膀、左右手肘、左右手腕、左右骨盆、左右膝蓋、左右腳踝等關節位置，還在考慮要不要加上 OpenPose 會辨識出的關節點圖；接著，利用棋盤格方法之相機校正技術計算出各個相機的內部參數及外部參數；再來，利用方才相機校正之結果，經由三角測量計算建立出關節在空間中的位置，並使用距離公式計算每一關節之間的距離；最後，基於 TotalCapture Dataset 提供之 vicon 立體骨架，利用計算出的距離進行立體骨架伸縮，得到與受試者四肢、身高相似之立體骨架，如圖 2.3。

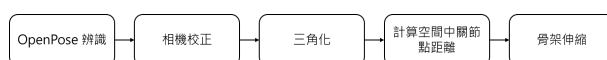


圖 2.1: 立體骨架建立流程圖

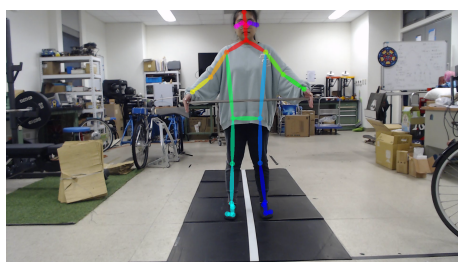


圖 2.2: OpenPose 辨識結果

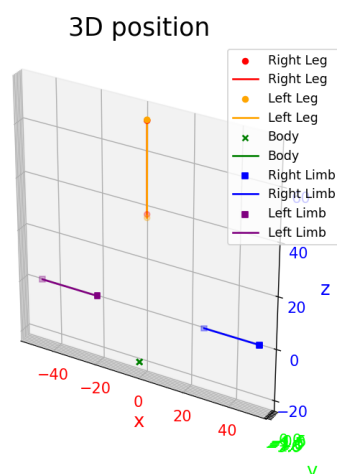


圖 2.3: 立體骨架建立結果



## 2.2.2 時間軸對齊

拍手及加速度判斷

## 2.2.3 座標系轉換

在人體量測領域中，每一量測方法都有其固有且常用之座標系，因此在進行感測器融合時，必須將各感測器之座標系轉換至同一座標系，以方便後續感測器融合計算。本研究共會涉及四種座標系，分別為骨骼座標系，IMU local 座標系，IMU global 座標系，全域座標系。其中，骨骼座標系以 TotalCapture Dataset 提供之 Vicon 立體骨架所在之座標系為基礎，如圖 2.4所示，其以人體骨盆中心為原點，+x、+y、+z 方向分別為向右 (red)、向後 (green)、向下 (blue)；IMU local 座標系即為感測器座標系，以感測器本身為原點，於 Xsens 系統中，定義其 x 軸為長邊延伸方向，y 軸為短邊延伸方向，z 軸則沿最大平面法向，如圖 2.5 所示；IMU global 座標系則以所在地為原點，y 軸沿子午線指向正北為正，z 軸指向天頂為正，x 軸與 yz 平面垂直，指向正東為正；

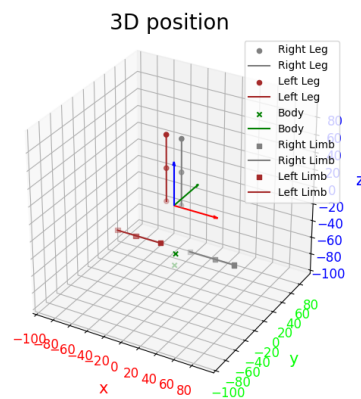


圖 2.4: 骨骼座標系定義

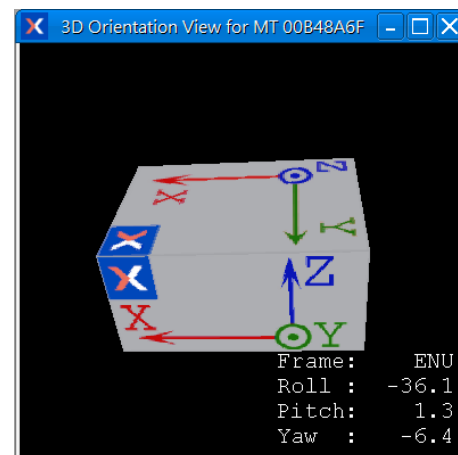


圖 2.5: IMU local 座標系定義

## IMU - 全域座標旋轉矩陣計算

$$R_{ig}$$

## IMU - 骨骼座標旋轉矩陣計算

使用 IMU 進行肢體朝向量測時，由於 IMU 黏貼於肢體上的方向及位置可能會和骨骼定義之方向及位置有所偏差，因此需使用 IMU - 骨骼座標旋轉矩



陣，將位於 IMU local 座標系的量測數值轉換至骨骼座標系，此矩陣 ( $R_{ib}$ ) 可經由 OpenSim [18] 處理並進一步計算而得，其取得及計算方式如下。首先，量測結束之 IMU 資料將透過 OpenSim 軟體進行處理，並取得 IMU 相對每一骨骼之相對位置及角度，處理結果可經由 OpenSim 軟體可視化，如圖 2.6 所示。此項數據紀錄於 OpenSim 輸出之 .osim 檔案的 <BodySet> 元素中，<BodySet> 以骨骼為單位，其中包含之 <components> 則記錄一附著於該骨骼上之 IMU 的相對位置及角度，<translation> 代表該 IMU 的原點相對於其附著之骨骼原點的平移，<orientation> 則代表該 IMU 的坐標軸相對於其附著之骨骼坐標軸的旋轉。因此若將 <translation> 數值更改為零，則可看到 IMU 原點與骨骼原點重合，如圖 2.7，而若將 <orientation> 數值更改為零，則可看到 IMU 坐標軸與骨骼坐標軸平行，如圖 2.8。因此，從 .osim 檔案取得每一 IMU 相對於骨骼之相對位置及角度後，即可計算出 IMU - 骨骼座標四元數。

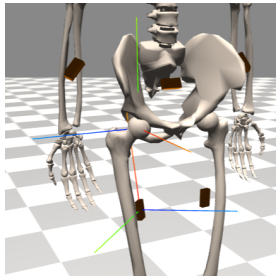


圖 2.6: OpenSim 計算結果

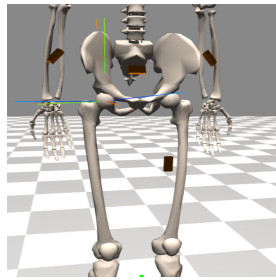


圖 2.7: IMU 平移為零

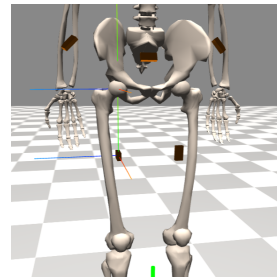


圖 2.8: IMU 旋轉為零

## 2.3 探討減少相機數量的可行性

可以從增加實驗架設方便性方面作為開頭著手，TotalCapture Dataset [17] 提供 8 台相機與 13 個 IMU 的量測資料，而在文獻 [13] 中，作者使用到 4 台相機及 8 個 IMU 的資料；另外在 TotalCapture Dataset 發表的文獻 [19] 中則有提及嘗試減少相機的硬體數量，準確度隨著相機數量的減少而下降，因此本章節嘗試以減少相機數量並選擇相機擺放位置進行感測器融合計算的方式，探討減少相機使用數量對於動作捕捉的影響，並嘗試探討相機擺放位置的選擇。



### 2.3.1 實驗方法

TotalCapture Dataset 實驗環境為一個 4x6 (m) 的方形空間，每一面牆面上方架設兩台相機，四面牆共計八台相機，每台相機距離地面高度皆為 2.5 (m)，擺放位置以上視的視角呈現，如圖 2.9 所示。本次實驗將分成七類情況進行，第一類情況為從八台相機中任選一台相機進行感測器融合計算，並將得到的姿勢估計結果與 TotalCapture Dataset 提供的 Vicon 位置資料進行比對，計算 MPJPE，第二類為任選兩台相機進行姿勢估計，以此方式遞增至第七類，任選七台相機進行姿勢估計，每一類情況皆會產生一組 MPJPE。

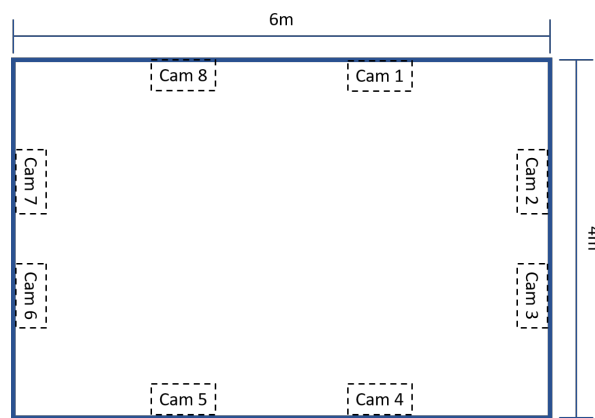


圖 2.9: 相機擺放位置

### 2.3.2 實驗結果

經過上述七類情況之實驗後，得到的 MPJPE 結果如以下所示。

#### 一台相機

將圖 2.9 中的相機任選一台進行估計，共計有 8 種估計結果，每一結果皆可計算出 MPJPE，MPJPE 計算結果如表 2.1 所示，可以發現無論選擇哪一個位置的相機，其估計誤差皆超過 500 (mm)。標準差為 87.6016 (mm)，平均值為 599.8855 (mm)。



表 2.1: 一台相機組合與其估計結果誤差

相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE
1	514.76	2	782.01	3	611.84	4	520.97
5	653.96	6	573.41	7	599.78	8	542.36

## 兩台相機

將圖 2.9 中的相機任選兩台進行排列組合，共計有 28 種組合方式之估計結果，每一結果皆可計算出 MPJPE，MPJPE 計算結果如表 2.2 所示，其中最佳組合方式為相機 18，其 MPJPE 為 46.9798 (mm)；而最差組合方式為相機 25，其 MPJPE 為 300.1637 (mm)。標準差為 61.7928 (mm)，平均值為 109.8694 (mm)。經由平均值可以發現兩台相機的姿勢估計表現相對一台相機的姿勢估計表現有大幅度的進步，進步幅度約為 500 (mm)；但是由標準差可知兩台相機的姿勢估計結果仍不穩定。

表 2.2: 兩台相機組合與其估計結果誤差

相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE
12	194.4957	13	81.9440	14	47.4927	15	76.7795
16	66.8653	17	58.7619	18	46.9798		
23	227.5722	24	182.6378	25	300.1637	26	183.0773
27	181.6520	28	145.2458				
34	99.3426	35	109.4178	36	92.7550	37	97.7435
38	81.3640						
45	87.4750	46	66.8800	47	67.4949	48	52.3318
56	126.3387	57	95.1469	58	64.6374		
67	100.7632	68	63.7904				
78	77.1948						

### 三台相機

將圖 2.9 中的相機任選三台進行排列組合，共計有 56 種組合方式之估計結果，每一結果皆可計算出 MPJPE，MPJPE 計算結果如表 2.3 所示，其中最佳組合方式為相機 137，其 MPJPE 為 27.1579 (mm)；而最差組合方式為相機 257，其 MPJPE 為 82.8158 (mm)。標準差為 13.8423 (mm)，平均值為 42.9001 (mm)。經由平均值可以發現：相較兩台相機的姿勢估計結果，三台相機的姿勢估計結果準確性仍有明顯提升，進步幅度約為 60 (mm)；由標準差可以知道三台相機的姿勢估計結果漸趨穩定。

表 2.3: 三台相機組合與其估計結果誤差

相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE
123	41.8538	124	43.2544	125	70.4747	126	41.7470	127	54.5721
128	46.7714								
134	30.9862	135	29.0275	136	33.6116	137	27.1579	138	29.9961
145	34.4860	146	32.7227	147	29.9785	148	31.5573		
156	38.3802	157	33.2107	158	33.4297				
167	32.4098	168	33.4044	178	32.2636				
234	52.0193	235	76.4435	236	52.6657	237	65.5297	238	44.5936
245	76.6047	246	42.4841	247	54.8719	248	45.2496		
256	74.8608	257	82.8158	258	62.4336				
267	58.7578	268	39.2619	278	62.7102				
345	35.9928	346	42.7418	347	35.2381	348	33.5695		
356	42.9933	357	35.1894	358	29.6995				
367	40.0411	368	35.8360	378	34.9221				
456	40.0658	457	37.2030	458	32.7289				
467	35.6906	468	33.8860	478	35.3151				
567	41.5287	568	34.8289	578	37.5083				
678	34.8273								

## 四台相機

將圖 2.9 中的相機任選四台進行排列組合，共計有 70 種組合方式之估計結果，每一結果皆可計算出 MPJPE，MPJPE 計算結果如表 2.4 所示，其中最佳組合方式為相機 1357，其 MPJPE 為 24.5789 (mm)；而最差組合方式為相機 2567，其 MPJPE 為 38.9142 (mm)。標準差為 3.6860 (mm)，平均值為 31.0114 (mm)。經由平均值可以發現，四台相機的姿勢估計結果相對三台相機的姿勢估計結果有進一步提升，進步幅度約為 10 (mm)，相較兩台相機與三台相機的進步幅度有逐漸平穩的趨勢；而由標準差可以知道四台相機組合的表現都相當穩定，無論如何選擇都不會有太大的影響。

表 2.4: 四台相機組合與其估計結果誤差

相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE
1234	30.9409	1235	28.8549	1236	30.3748	1237	28.8213	1238	31.1560
1245	33.1706	1246	31.5377	1247	32.1031	1248	33.6613		
1256	34.5089	1257	34.6787	1258	35.2625				
1267	31.7009	1268	33.0636	1278	34.6861				
1345	26.9753	1346	28.2925	1347	25.4715	1348	26.7584		
1356	27.7651	1357	24.5789	1358	25.8578				
1367	26.3845	1368	26.9445	1378	26.1466				
1456	29.9311	1457	26.6848	1458	27.9429				
1467	26.8121	1468	27.5472	1478	27.1979				
1567	27.5824	1568	29.1892	1578	27.4552	1678	27.8378		
2345	34.0972	2346	36.5508	2347	36.7314	2348	33.7047		
2356	35.6585	2357	36.3698	2358	30.8585				
2367	36.8662	2368	32.6160	2378	34.8430				
2456	35.6440	2457	37.1359	2458	33.4067				
2467	35.8266	2468	33.2654	2478	36.6081				
2567	38.9142	2568	33.7187	2578	38.7211	2678	34.3506		
3456	32.4429	3457	27.7338	3458	27.6554				
3467	31.0735	3468	29.9925	3478	27.9168				
3567	29.2876	3568	28.3170	3578	26.8460	3678	29.4004		
4567	29.8717	4568	29.7963	4578	28.2055	4678	28.9104	5678	29.5824

## 五台相機

將圖 2.9 中的相機任選三台進行排列組合，共計有 56 種組合方式之估計結果，每一結果皆可計算出 MPJPE，MPJPE 計算結果如表 2.5 所示，其中最佳組合方式為相機 137，其 MPJPE 為 27.1579 (mm)；而最差組合方式為相機 257，其 MPJPE 為 82.8158 (mm)。標準差為 13.8423 (mm)，平均值為 42.9001 (mm)。經由平均值可以發現：相較四台相機的姿勢估計結果，五台相機的姿勢估計結果並無明顯提升，進步幅度約為 4 (mm)；由標準差可以知道五台相機的姿勢估計結果十分集中。

表 2.5: 五台相機組合與其估計結果誤差

相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE	相機配置	MPJPE
12345	27.2222	12346	28.7270	12347	27.2410	12348	28.1424	12356	27.5630
12357	25.4345	12358	26.9508	12367	27.4309	12368	27.9631	12378	27.5871
12456	29.2157	12457	27.4585	12458	29.1856	12467	28.0690	12468	28.8999
12478	28.9294	12567	28.2794	12568	29.9585	12578	29.0917	12678	29.1141
13456	26.5308	13457	24.2506	13458	24.7973	13467	25.4456	13468	25.5140
13478	24.5152	13567	24.7163	13568	25.2622	13578	23.9568	13678	25.0847
14567	25.6222	14568	26.3839	14578	25.2386	14678	25.3628	15678	26.0199
23456	31.3515	23457	28.8724	23458	28.4882	23467	32.1120	23468	30.4884
23478	29.8341	23567	29.8954	23568	28.7786	23578	27.9967	23678	30.1241
24567	30.4360	24568	30.1827	24578	29.4012	24678	30.1792	25678	
34567		34568		34578		34678		35678	
45678									

## 六台相機、七台相機

由於六台相機與七台相機之估計結果皆與四台相機及五台相機的估計結果相近，因此不再完整將結果列出。六台相機姿勢估計誤差之標準差為 13.8423 (mm)，平均值為 42.9001 (mm)，七台相機姿勢估計誤差之標準差為 13.8423 (mm)，平均值為 42.9001 (mm)。將兩者的標準差及平均值與四台相機及五台相機的標準差及平均值進行比較，可以發現六台相機及七台相機的表現皆與四台相機及五台相機的表現結果相近。

### 2.3.3 結果與討論

將一台相機到七台相機的全部 MPJPE 繪製成圖 2.10，可以發現相機數量從一台增加到四台時，MPJPE 隨著相機數量的增加有明顯下降的趨勢，從四台相機開始，MPJPE 的下降幅度逐漸減緩，因此可以推斷，當相機數量增加到四台時，MPJPE 的表現已經相當穩定，且相機數量增加到五台以上時，MPJPE 的表現並不會有太大的改善。另外，若觀察每一相機數量的最佳表現 (即 MPJPE 最小值)，可以發現兩台相機及三台相機的最佳表現皆有到達誤差 50 (mm) 以下，所以，若希望盡可能減少相機數量，且容許誤差 50 (mm) 以下，則可以選擇兩台相機或三台相機進行姿勢估計，因此，綜上所述可推斷，為增加實驗架設方便性，減少相機數量是可考慮的選擇。

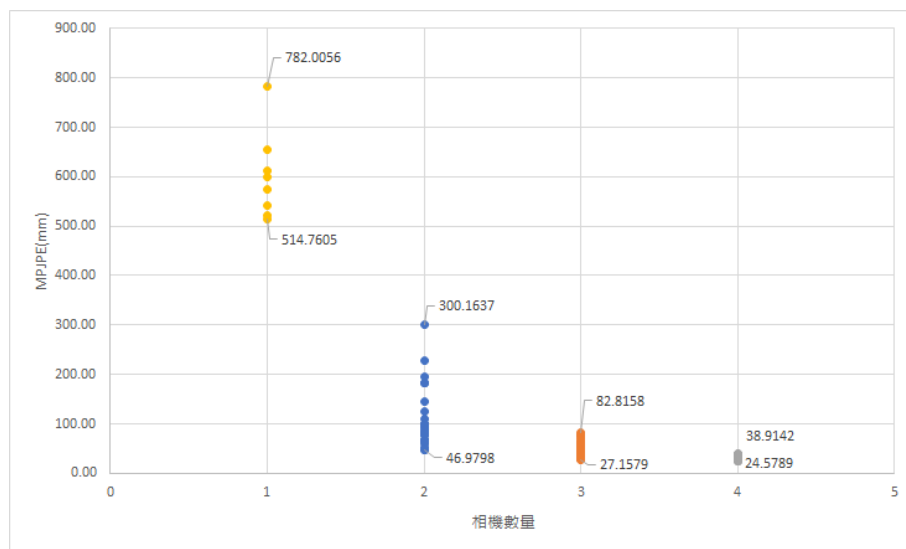


圖 2.10: 一台相機到七台相機的組合估計結果



## 2.4 結果可視化

還沒想到要寫甚麼

## 2.5 小結

本章節首先介紹將會使用到的背景知識，像是透過希爾式肌肉模型來模擬肌肉，藉此理解肌肉的機械與生理特性，而藉由 OpenSim 的協助，不但能快速建立肌肉骨骼模型，還可以執行正向動力學與肌肉計算控制等模擬，且得到一個可信的結果，接下來介紹本研究的核心模擬——運動軌跡預測任務，後續的敏感度分析、最佳化過程與模型驗證，皆是以預測任務為基礎來延伸，透過預測任務來檢視模型的表現，亦即預測誤差。本研究透過敏感度分析來得知肌肉與任務之間的關係，藉此挑選合適的任務集作為參數評估輸入，搭配最佳化演算法來進行參數估計，其中多組預測任務具有讓目標函數更加明確的功用，避免參數不具識別性的原因，掉入至局部最小值結果當中，最終透過間接方法來驗證模型的正確性。

該研究方法之應用可分為兩種討論，第一是模擬研究，即為本研究使用之方法，透過純模擬研究來檢視該方法的可行性，優點是其具有明確答案可供參考，且無需考慮量測造成的不確定性，在進入到實際應用前，也必須先確認模擬研究是可執行的；第二種則是實際應用，其可透過 EMG 與動作捕捉系統來達成，藉由 EMG 量測肌肉訊號作為輸入，動作捕捉系統量測運動軌跡作為輸出之結果比較，兩者結合並搭配本研究之最佳化方法，即可達成肌肉之參數估計，但存在龐大的量測不確定性情況下，由於微小的軌跡偏離，即會造成肌肉參數的變動與抗衡，縱使估計出肌肉參數，其結果之正確性仍有待商榷，因此於現今之科技發展，要達成實際應用仍有一段距離。

本章節主要介紹論文之研究方法，下章節會以實際模型與動作進行討論，藉由上方所提及之方法與流程，針對模型進行肌肉參數評估的前置作業與套用說明。



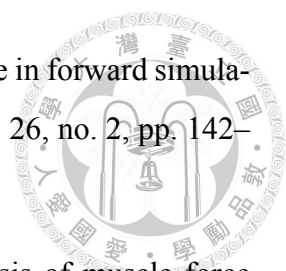


---

## 參考文獻

---

- [1] Vicon Motion Systems Ltd., “Vicon.” WebPage. Online available at: <https://www.vicon.com>; accessed at 19/5/2023.
- [2] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, “Openpose: Real-time multi-person 2d pose estimation using part affinity fields,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [3] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, “Convolutional pose machines,” in *CVPR*, 2016.
- [4] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, “Hand keypoint detection in single images using multiview bootstrapping,” in *CVPR*, 2017.
- [5] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *CVPR*, 2017.
- [6] Google Inc., “Mediapipe.” WebPage. Online available at: <https://developers.google.com/mediapipe>; accessed at 19/5/2023.
- [7] X. Nie, J. Feng, J. Zhang, and S. Yan, “Single-stage multi-person pose machines,” in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 6951–6960, 2019.
- [8] R. Akhundov, D. J. Saxby, L. E. Diamond, S. Edwards, P. Clausen, K. Dooley, S. Blyton, and S. J. Snodgrass, “Is subject-specific musculoskeletal modelling worth the extra effort or is generic modelling worth the shortcut?,” *PLoS One*, vol. 17, no. 1, p. e0262936, 2022.

- 
- [9] M. Xiao and J. Higginson, “Sensitivity of estimated muscle force in forward simulation of normal walking,” *Journal of Applied Biomechanics*, vol. 26, no. 2, pp. 142–149, 2010.
- [10] P. Bujalski, J. Martins, and L. Stirling, “A monte carlo analysis of muscle force estimation sensitivity to muscle-tendon properties using a hill-based muscle model,” *Journal of Biomechanics*, vol. 79, pp. 67–77, 2018.
- [11] R. Hinson Jr, K. Saul, D. Kamper, and H. Huang, “Sensitivity analysis guided improvement of an electromyogram-driven lumped parameter musculoskeletal hand model,” *Journal of Biomechanics*, vol. 141, p. 111200, 2022.
- [12] J. L. Hicks, T. K. Uchida, A. Seth, A. Rajagopal, and S. L. Delp, “Is my model good enough? best practices for verification and validation of musculoskeletal models and simulations of movement,” *Journal of Biomechanical Engineering*, vol. 137, no. 2, 2015.
- [13] Z. Zhang, C. Wang, W. Qin, and W. Zeng, “Fusing wearable imus with multi-view images for human pose estimation: A geometric approach,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2200–2209, 2020.
- [14] D. Pagnon, M. Domalain, and L. Reveret, “Pose2sim: An end-to-end workflow for 3d markerless sports kinematics—part 1: Robustness,” *Sensors*, 2021.
- [15] D. Pagnon, M. Domalain, and L. Reveret, “Pose2sim: An end-to-end workflow for 3d markerless sports kinematics—part 2: Accuracy,” *Sensors*, 2022.
- [16] D. Pagnon, M. Domalain, and L. Reveret, “Pose2sim: An open-source python package for multiview markerless kinematics,” *Journal of Open Source Software*, 2022.
- [17] M. Trumble, A. Gilbert, C. Malleson, A. Hilton, and J. Collomosse, “Total capture: 3d human pose estimation fusing video and inertial sensors,” in *2017 British Machine Vision Conference (BMVC)*, 2017.
- [18] S. L. Delp, F. C. Anderson, A. S. Arnold, P. Loan, A. Habib, C. T. John, E. Guendelman, and D. G. Thelen, “Opensim: open-source software to create and analyze

dynamic simulations of movement,” *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.

- [19] M. Trumble, A. Gilbert, C. Malleson, A. Hilton, and J. Collomosse, “Total capture: 3d human pose estimation fusing video and inertial sensors,” in *Proceedings of 28th British Machine Vision Conference*, pp. 1–13, 2017.