

國立交通大學

資訊工程系

碩士論文

手部平面動作切割



Planar Hand Motion Segmentation

研究生：紀文原

指導教授：陳 稔 博士

中華民國九十四年七月

手部平面動作切割

Planar Hand Motion Segmentation

研究生：紀文原

Student : Wen-Yuan Chi

指導教授：陳 稔 博士

Advisor : Dr. Zen Chen

國立交通大學
資訊工程系
碩士論文



Submitted to Department of Computer Science and Information Engineering
College of Electrical Engineering and Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Computer Science and Information Engineering

July 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年七月

手部平面動作切割

學生：紀文原

指導教授：陳稔 博士

國立交通大學資訊工程學系

摘要

本論文的主要目的在於將手部運動切割成數個有意義的動作片段，以期能將連續的運動表示成動作片段序列，如此一來，即可提供一個的基礎給後續的分析與處理（動作辨識、動作重編輯等等）。

本中利用關節旋轉軸與肢節垂直與否，將肢節在空間中的運動軌跡分為共面與共圓錐兩種基本運動模式，而手部運動即可視為此兩種運動模式搭配不同的旋轉角度的組合情形。

藉由計算肢節運動軌跡在空間中形成的平面，可求出該段運動過程中關節的旋轉軸為何。依照旋轉軸是否發生改變來進行動作切割，即可得到數個動作片段，其運動過程中旋轉軸固定不變。接著，再對片段中的角度變化，以二階的AR模型來做進一步的切割。最後，即可將動作切割成數個“同旋轉軸與同角度變化模式”的動作片段。

在文章中會說明如何計算運動平面及如何利用旋轉軸變化及角度變化來對動作進行切割。最後用實驗驗證文中提出的方法。

Planar Hand Motion Segmentation

Student: Wen-Yuan Chi

Advisor: Dr. Zen Chen

Department of Computer Science and Information Engineering
National Chiao Tung University

Abstract

The main purpose of this thesis is to segment a video clip of planar hand motion into motion subsequences each with some unique semantic meaning. Further analysis or processing on these motion segments such as motion recognition, motion editing can be conducted.

In this thesis, the hand motion segments in 3D space are classified into two types: co-planar and co-axial, depending on whether the rotation axis is perpendicular to the trajectory of each hand joint and whether the hand joints are coplanar. Any hand motion sequence can be treated as a combination of these two types of motion with different motion parameters.

By finding the plane of the hand joint trajectory, the rotation axis of the planar hand motion can be obtained. Then, the projection of the 3D joint position trajectory onto the motion plane is computed. Then, perform the second AR model analysis on the projected trajectory to find any further possible segmentation of planar joint motion. In this way, planar hand motion clip is segmented into sub-segments of the same type of AR model.

The proposed method is mathematically formulated. Experimental simulations applied to various clips of hand motion are shown to obtain the segments of hand motions each described by an AR model.

誌謝

首先，要感謝實驗室提供一個相當不錯的工作環境及設備，讓我能完成本篇論文，接著則是要感謝 陳稔教授兩年來的指導與督促，使我學到如何獨立研究與分析的能力，並且在研究遇到問題時，也都能提供適當的方向與解法來幫助我做的更好。

此外，更要感謝方彥翔學長一路陪我走到最後，尤其是論文結束前時最難熬的那一段日子，還要感謝實驗室的全體成員給予的諸多協助，顧正偉同學、蘇詠迪同學、陳柏聰學弟、吳昶濤學弟，幸好這一路上能有你們從旁一直給我鼓勵與幫助，不然真不曉得這一切該如何結束。

最後，一定要感謝我父母對我的栽培與期望，還有我女朋友鵲如的加油打氣，讓我在一路上都能充滿幹勁與活力。



要感謝的人太多，在此無法一一致謝，故僅以本篇論文獻給他們，表達我的感謝之意。

紀文原 於交大

July, 2005

目錄

摘要	i
Abstract	ii
誌謝	iii
目錄	iv
圖目錄	vi
表目錄	vii
一	緒論.....	1
1.1	研究動機與目標.....	1
1.2	相關研究.....	2
1.3	研究流程概述.....	2
1.4	論文組織.....	4
二	動作捕捉與逆向動力學.....	5
2.1	動作資料擷取.....	5
2.1.1	基礎理論.....	5
2.1.2	相機參數校正.....	6
2.1.3	3D 點位重建.....	7
2.2	應用逆向動力學修正動作資料.....	8
2.2.1	肢節點位重定.....	8
2.2.2	逆向動力學修正.....	9
三	動作切割.....	11
3.1	人體模型定義與運動分析.....	11
3.2	計算運動平面.....	13
3.3	以運動平面進行動作切割.....	14
3.4	計算轉動角度，並利用 AR Model 再進行動作切割.....	18
3.4.1	計算旋轉角度.....	18
3.4.2	Second order AR model 簡介.....	19
3.4.3	AR Model Segmentation Algorithm.....	20
3.5	動作片段的表示.....	21
四	實驗數據與結果.....	25
4.1	檢驗維度退化之容許值參數設定.....	25
4.2	同平面上反覆運動的切割結果.....	30
4.3	不同運動平面切割結果.....	37
4.4	比較有無平面資訊輔助的切割結果.....	51
五	結論與未來發展方向.....	54
5.1	結論.....	54

5.2 未來工作.....	55
參考文獻.....	56



圖目錄

圖 1 - 1 研究方法流程圖	3
圖 2 - 1 座標系統的示意圖	5
圖 2 - 2 實驗校正用影像 (Left-View)	6
圖 2 - 3 實驗校正用影像 (Right-View)	6
圖 2 - 4 求解三度空間點位置示意圖	7
圖 2 - 5 肢節重定示意圖。	9
圖 2 - 6 肢節重定示意圖。	9
圖 3 - 1 人體基本結構	11
圖 3 - 2 運動平面示意圖：	13
圖 4 - 1 實驗一：3D 運動軌跡	25
圖 4 - 2 實驗一：各關節點位	26
圖 4 - 3 實驗二：3D 運動軌跡	28
圖 4 - 4 實驗二：各關節點位	28
圖 4 - 5 實驗三：動作擷圖	33
圖 4 - 6 實驗三：各關節點位座標	34
圖 4 - 7 實驗三：點位到平面誤差	34
圖 4 - 8 實驗三：AR 切割結果	36
圖 4 - 9 實驗四：動作擷圖	39
圖 4 - 10 實驗四：各關節點位座標	40
圖 4 - 11 實驗四：動作片段#1	41
圖 4 - 12 實驗四：動作片段#2	41
圖 4 - 13 實驗四：動作片段#3	42
圖 4 - 14 實驗五：動作擷圖	46
圖 4 - 15 實驗五：各關節點位座標	47
圖 4 - 16 實驗五：動作片段#1	49
圖 4 - 17 實驗五：動作片段#2	49
圖 4 - 18 實驗五：動作片段#2	50
圖 4 - 19 實驗六：關節角度	51
圖 4 - 20 實驗六：AR 切割結果	52

表目錄

表格 1 實驗一：平面切割結果.....	26
表格 2 實驗一：各片段之奇異值.....	26
表格 3 實驗一：片段合併之奇異值.....	27
表格 4 實驗二：平面切割結果.....	28
表格 5 實驗二：各片段之奇異值.....	28
表格 6 實驗二：片段合併之奇異值.....	29
表格 7 實驗三：平面切割結果.....	34
表格 8 實驗三：AR 切割結果.....	36
表格 9 實驗三：AR 參數矩陣.....	35
表格 10 實驗四：平面切割結果.....	41
表格 11 實驗五：平面切割結果.....	47
表格 12 實驗五：平面合併過程之奇異值.....	47
表格 13 實驗五：平面合併結果.....	48
表格 14 實驗六：AR 切割結果.....	52



一 緒論

1.1 研究動機與目標

隨著科技的進步，人們用來記錄資料、傳遞訊息的工具也從文字轉變到圖片、聲音甚至是影片，如何有效在大量資料下尋找出特定的、有用的資料則是共同需要面臨的問題。然而，以影片的方式來描述動作則缺乏像文字一樣的符號集合，所以要對動作進行分析、比較，則變得相當困難。

人體運動缺乏基礎符號集合的主要原因在於，運動是一種隨時間不斷的連續變化，但經由對人體運動的觀察與分析，可以發現某些變化方式不斷的重複出現，或是幾種變化方式依不同的順序發生而產生不同的動作。若是可以透過一些法則將固定的變化切割成獨立的片段，則便可以利用叢集（clustering）的技巧，將相同的動作歸到同一類並給予一個符號來表示，此即為所謂的基礎動作單元（Motion Primitives）。這樣一來，人體的運動過程即可視為一連串基礎動作單元的組合，藉由不同的基礎動作單元來產生出不同的動作。

當這樣的人體基礎動作單元集合被切割出來後，不但可以精簡的將人體運動表示成符號序列；同樣地，更可以進行相當多的分析及應用，如：透過符號序列的比對來辨認特定的動作、人體運動的重建、人體運動的創造、人體動作資料庫的建立等等。

因此，我們期望能將一個完整的人體運動過程切割成數個獨立的動作片段，且每個片段都要能符合感官上的意義，並提供動作片段間相似程度的比較方法，如此一來，即可叢集這些動作片段來找出人體運動的基礎動作單元。

1.2 相關研究

目前對於分析影像序列的研究有以下兩種方面：

1. 訓練過的程序，把影像序列分析成數個動作片段。優點在於一切分析可以自動化，方便處理；但對於人體的動作所產生的錯誤，有很大的空間要改進。例如對於接合點（Joints）動作的測量，增加獲得跟分析的資料數量，諸如此類。
2. 手工對影像序列作切段（Partition），然後根據所估計的接合點轉動角度（Joint Angle）來對已被分段的影像序列作註解[1][2][3]。優點是可以良好的呈現被切段後的動作；但缺點是花費時間相當的久，不適合大量動作序列時的切段使用之。

對於以影像為基礎的機器分析而言，對於不同人的不同動作，和同一人在不同時間的動作，為了做比較，它必須需要一個動作的重新表示。而人體動作的動態系統表示法已經廣泛地被使用在人類使用車子的行為上面 [4][5]。在辨識系統上，動作會被重新表示成一群簡單動作的序列，常見的方式是用線性動態系統（Linear Dynamic System）[6][7][8]；而追蹤跟辨識的方法也已經有使用交換式的動態系統模型，但是需要知道一開始會把動作分為幾種類型（例：[6][8][10]）。

而在使用動態系統做分析的情況下，由於資料量相當的龐大，所以通常以抽取特徵向量（Feature Vector）的方式來將資料作第一步簡化。在對精簡過的資料做運算時，維度簡化（Dimension Reduction）是常用的方法 [11][12][13]。如 PCA（Principle Component Analysis）[12]，Isomap，Spatial-Temporal ([11][13])。

1.3 研究流程概述

在本篇論文中是針對人體手部運動進行分析，先利用[22][23]中的動作捕捉方法與修正取得肩膀、手肘及手腕這三個關節在三維空間中的坐標，再對肢節向量（上臂）在三維空間中的運動軌跡去尋找可能的運動平面，如此即可以將動作切分成若干份動作片段，片段中肢節向量（上臂）的運動軌跡都在同一平面上。

接著，以運動平面的法向量當成該時刻肩膀關節的旋轉軸，計算出肩膀關節的旋轉角度，再利用上臂與前臂的夾角當成手肘關節的旋轉角度。在取得手部運動中的所有關節的角度變量後，將角度變量表示成 AR Model 的狀態向量 (State Vector)，依[14,15,16]中的方法進行再將上述同平面運動的動作片段做進一步的切割，最後，得到一群以轉軸與 AR Model 表示的動作片段集合。

對以轉軸與角度變化表示的動作片段集合，利用 clustering 的技巧，可以將相似的動作分類到同一群，這幾群的動作片段即是所要尋找的基礎動作集合。最後，將該動作的動作片段與找到的基礎動作單元集合進行相似度的比對後，即可將該動作以符號序列的方式表示。

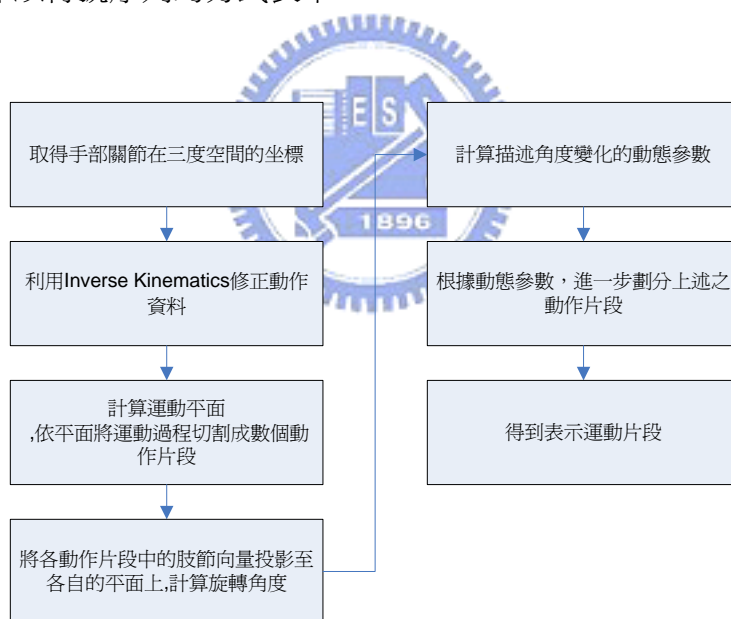


圖 1 - 1 研究方法流程圖

1.4 論文組織

第二章將會說明如何利用多台攝影機取得人體運動資料，再討論捕捉系統誤差的情況，並提出以逆向動力學修正動作捕捉誤差的方法。

第三章則是要探討人體的運動機制及運動模式，並將運動平面的轉換當成初步切割的依據，在取得數個共面運動的動作片段後，採用 AR Model 來做角度變化的分析，再依據 AR Model 中的動態參數變化的劇烈程度，將動作片段做進一步的切割，最後即可利用轉軸與動態參數來表示一個動作片段。

第四章，透過實驗數據來展現切割的結果。

第五章，則是對本文中提出的方法做一個總結，並提出未來可以改進或進行的目標。



二 動作捕捉與逆向動力學

2.1 動作資料擷取

在整個物體重建過程的三度空間座標系統中，使用二台以上的攝影機來拍攝三度空間中運動的人體，藉以計算出人體在 3D 空間中的位置（position）與方向（orientation），並推導三度空間物體和二維影像之間的投影關係，做為計算人體 3D 資訊的依據。

2.1.1 基礎理論

首先定義一個共同的座標系統稱為「世界座標系統」（World Coordinate System, WCS），參考座標軸為 $[X_w \ Y_w \ Z_w]^T$ 。影像座標系統（Image Coordinate System, ICS），參考座標為軸 $[U_1 \ V_1]^T$ 。攝影機座標系統（Camera Coordinate System, CCS），參考座標軸為 $[X_c \ Y_c \ Z_c]^T$ 。三個座標系統的示意圖如圖 2-1 所示：[22]

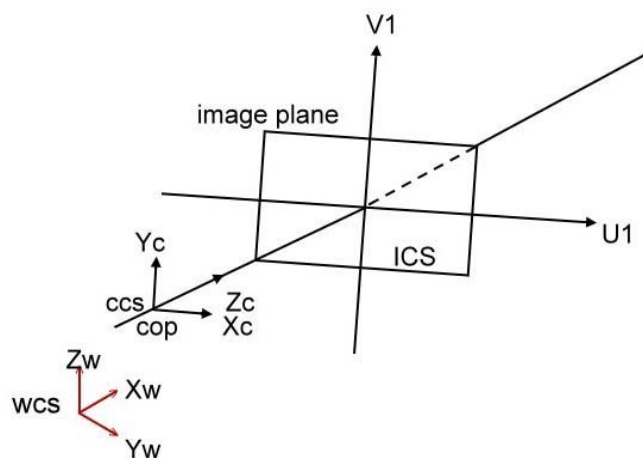


圖 2-1 座標系統的示意圖

2.1.2 相機參數校正

從影像座標系統到世界座標系統存在一個 $H_{3 \times 4}$ 的轉換矩陣：

$$\begin{bmatrix} u_i \cdot w \\ v_i \cdot w \\ w \end{bmatrix} = H \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2.1)$$

在矩陣 H 中共有 12 個未知參數，以傳統的方法[20],[21]，只要用 6 對以上不共面的 3D 點及其對應的 2D 投影點，再利用最小平方方法 (Least square method) 計算出矩陣 H 中的每一個參數，如此即可求得最近似的投影矩陣 H 。



圖 2 - 2 實驗校正用影像 (Left-View)



圖 2 - 3 實驗校正用影像 (Right-View)

2.1.3 3D 點位重建

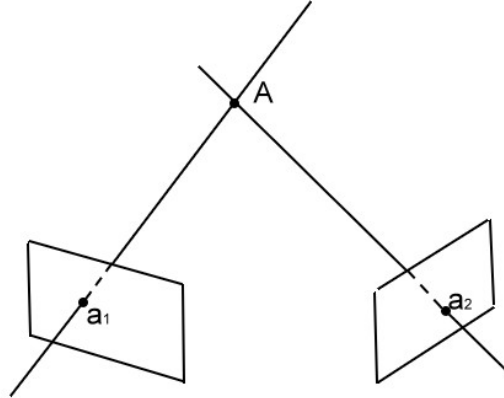


圖 2-4 求解三度空間點位置示意圖

以上圖為例，已知兩張影像上的點 $a_1 = [u_{a_1} \ v_{a_1}]^T$ ， $a_2 = [u_{a_2} \ v_{a_2}]^T$ 為三度空間中的點 $A = [X_A \ Y_A \ Z_A]^T$ 在兩張影像上的投影點，依(2.1)即可列出下列關係式：[22]

$$\begin{bmatrix} u_{a_1} w_1 \\ v_{a_1} w_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} h'_{11} & h'_{12} & h'_{13} & h'_{14} \\ h'_{21} & h'_{22} & h'_{23} & h'_{24} \\ h'_{31} & h'_{32} & h'_{33} & h'_{34} \end{bmatrix} \cdot \begin{bmatrix} X_A \\ Y_A \\ Z_A \\ 1 \end{bmatrix} \quad (2.2)$$

$$\begin{bmatrix} u_{a_2} w_2 \\ v_{a_2} w_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} h''_{11} & h''_{12} & h''_{13} & h''_{14} \\ h''_{21} & h''_{22} & h''_{23} & h''_{24} \\ h''_{31} & h''_{32} & h''_{33} & h''_{34} \end{bmatrix} \cdot \begin{bmatrix} X_A \\ Y_A \\ Z_A \\ 1 \end{bmatrix} \quad (2.3)$$

利用最小平方法（least square method）求解，得到 A 點在三維空間的坐標

$$A = (M^T M)^{-1} M^T b$$

2.2 應用逆向動力學修正動作資料

在動作捕捉的過程中，因為肩膀與手肘這兩個關節在運動過程中常會發生遮蔽、重疊的情況，會使得關節位置的標定變得相當困難，也因此會造成重建出來的 3D 點位不夠精準，但相較於肩膀與手肘，在運動過程中手腕較為明顯、易標定的關節，因此手腕重建出來的點位相對於肩膀與手肘重建的點位是較可以信任的資料。因此，需要採用逆向動力學中的方法來修正肩膀與手肘的點位。[22][22]

2.2.1 肢節點位重定

由於影像處理所求得的肢節點位置不可能絕對正確，計算出來的肢節長度也會有所誤差。然而對於同一個人，肢節的長度是固定不變的，所以必須重定肢節點的位置，使得重建出來的肢節長度能符合表演者的肢節長度。

肢節點的初始位置決定是利用影像處理所求得的肢節向量，再套上已知的肢節長度得出新的肢節點位置。此做法可使肢節的長度符合已知的人體模型。[22]

首先，將手腕在二張影像座標系統上 (ICS_1 和 ICS_2) 的座標位置，依照 2.1.3 節中提到的方法重建出手部關節在三維空間中的座標位置後，計算出上臂與前臂的肢節 low_arm 和 $upper_arm$ 在三維空間的向量。接著，從固定的肩膀位置依上臂向量加上事前測量的上臂長度可得修正後手肘在三維空間中的座標位置；同樣地，也可得到手腕在三維空間中的座標位置。

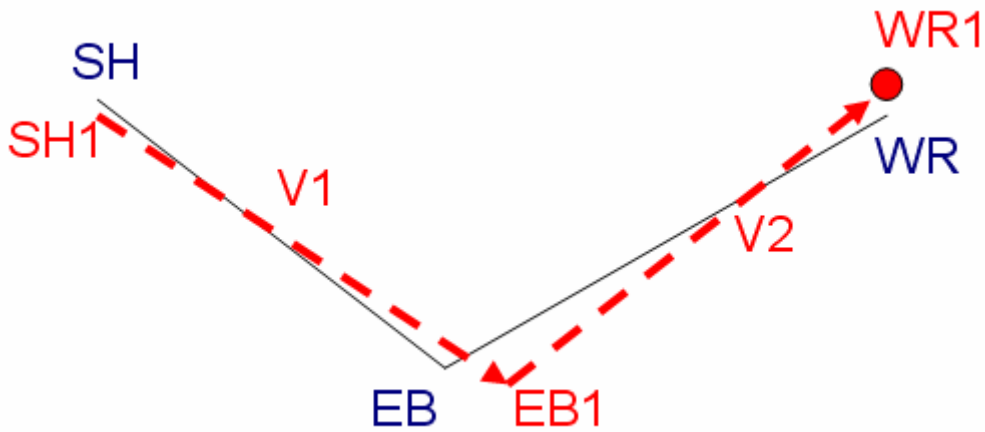


圖 2-5 肢節重定示意圖。

SH,EB,WR 為真實點位，SH1,EB1,WR1 為重建點位

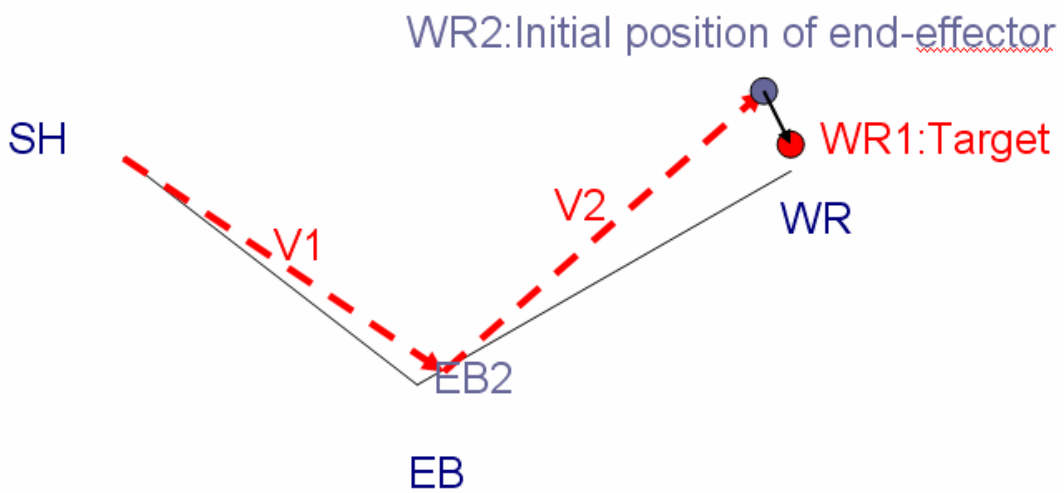


圖 2-6 肢節重定示意圖。

EB2,WR2 重定後點位

2.2.2 逆向動力學修正

在 2.2.1 中，修正了肢節長度不等長及肩膀不對齊的問題，但也影響了手腕的位置，而在動作捕捉重建的資料中，手腕的位置重建出來的點位是較清楚且誤差較小的點位，因此，需要重新調整關節的角度，使得手腕座標能符合實際拍攝重建的座標。

把從影像中重建的手腕坐標 **WR1** 定為逆向動力學修正時的目標 (**target**)，將肢節重定後的手腕坐標 **WR2** (圖 2-6) 當成初始解，代入[23]的方法進行修正，即可求解出肩膀與手肘的關節角度，使得 **WR2** 到 **WR1** 的距離最小。透過這樣的修正，我們即可得到可靠的手部運動資料。



三 動作切割

3.1 人體模型定義與運動分析

人體是由骨骼所組成的骨架堆成，而骨骼又受肌肉牽引而轉動，此為人體運動的機制。然而，人體的骨架非常複雜，為了使 Motion Capture 的資料簡化，本文制定了既方便計算又不失人體基本結構的骨架架構[22]，如圖 3-1：

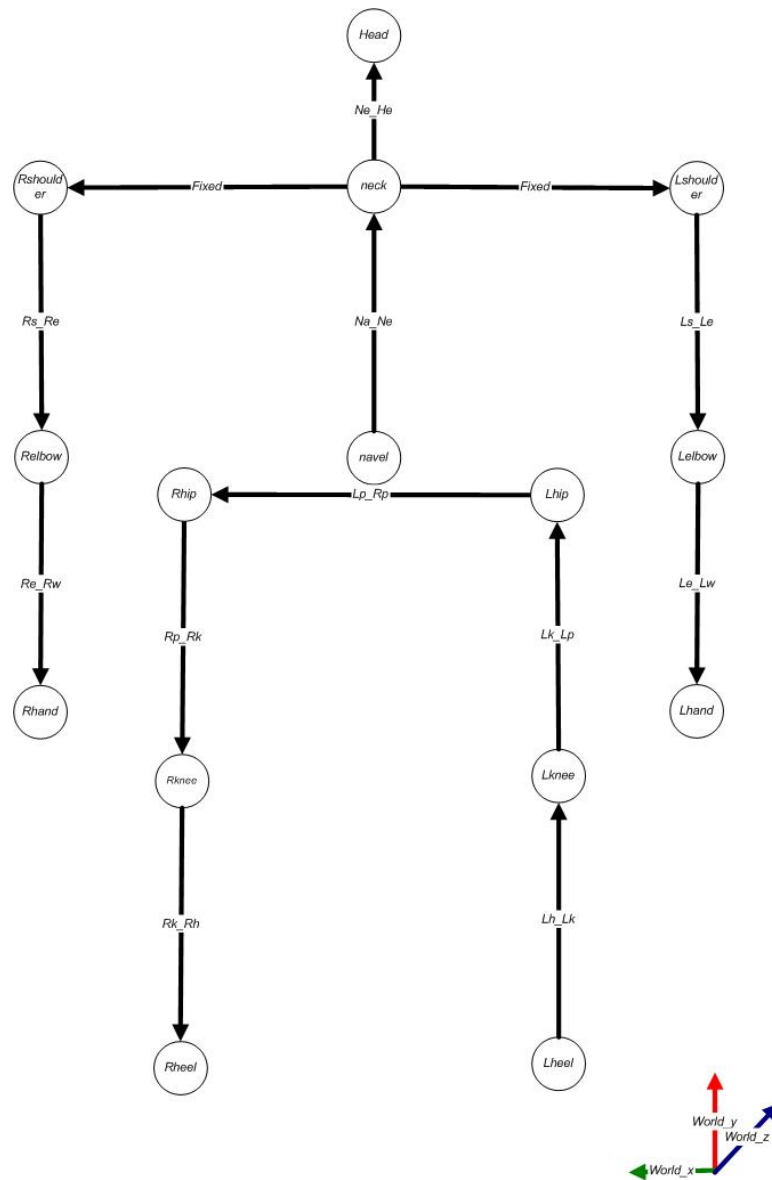


圖 3 - 1 人體基本結構

運動的發生是由全身各關節點的轉動，再加上腳部關節轉動所造成的平移而組成的。因此，可以透過觀察關節角度的變化來判斷目前的運動情況。

在本文中是以手部運動為例，因此只需要考慮手部關節（肩膀、手肘）的轉動，其它的關節則是固定不動。在上圖基本結構的定義中，我們定義肩膀為 3DOF 的關節，手肘為 1DOF 的關節。因為手肘是 1DOF 的關節，其所連接的肢節只能在固定的平面上運動，所以要計算手肘關節轉動角度只需要看兩個肢節（前臂與上臂）的夾角，即可表示手肘的旋轉角度；而肩膀是 3DOF 的關節，可以在三維空間中自由地轉動，但若找出旋轉時的旋轉軸，即可將 3DOF 降為 1DOF，再將前後時刻上臂向量依 quaternion 計算法，即可求出肩膀的旋轉角度。

上臂在三維空間中的運動軌跡根據肢節與轉軸垂直與否主要可分為共面（肢節與轉軸垂直）、共圓錐（肢節與轉軸不垂直）等兩種模式。藉由這兩種模式的組合而產生出各種不同的運動，而共圓錐模式在短時間內亦可視為共面。所以第一個工作即是要找出動作中的共面運動。

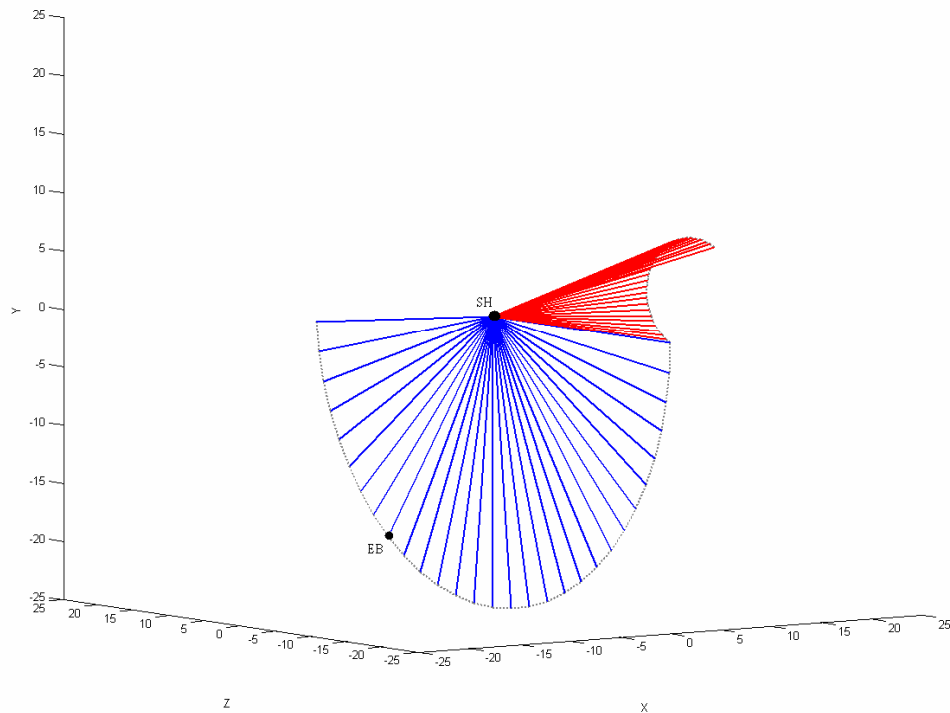


圖 3 - 2 運動平面示意圖：

由共面運動（藍色線段）與共圓錐運動（紅色線段）組成，
 灰色虛線為手肘運動軌跡
 SH:肩膀坐標、EB:手肘坐標。

3.2 計算運動平面

在運動過程中，利用肩膀到手肘的肢節向量來尋找動作的旋轉平面，該平面的法向量即是動作的旋轉軸；而要求空間中兩個向量的旋轉軸最快的方法莫過於對這兩個向量做外積。但是，實際情況並非如此簡單，除了轉軸未必與肢節垂直外，從捕捉系統得到的 3D 點位並不是非常準確，些微的誤差都會讓做外積得到的向量發生偏移，而誤差的來源有

- 相機參數校正不夠準確
- 關節點位標定不精確
- 示範人員的晃動，

而使得一個平順的動作，其關節點位沒有落在同一平面，而是在該平面附近

跳動，而這些誤差則會帶來相當大的影響。

因為誤差的存在，只能改求最適合平面，使得所有點位到平面的距離最短，即求解一個平面 $aX + bY + cZ + d = 0$ ，使得 $D = \sum_{t=1}^N |a * x(t) + b * y(t) + c * z(t) + d|$ 最小，套用 SVD 的方法即是求解

$$A \cdot \bar{x} = \bar{0}, \text{ 其中 } A = \begin{bmatrix} x(1) & y(1) & z(1) & 1 \\ x(2) & y(2) & z(2) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x(N) & y(N) & z(N) & 1 \end{bmatrix}_{N-by-4}, \bar{x} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix},$$

假設 $A\bar{x} = \varepsilon$ ($\varepsilon = 0$ or $\varepsilon \neq 0$), $A = UDV^T$ (U, V 為 orthonormal matrix)

$$\varepsilon^2 = (A\bar{x})^T A\bar{x} = \bar{x}^T A^T A\bar{x} = \bar{x}^T V D^T U^T U D V^T \bar{x} = \bar{x}^T V D^T D V^T \bar{x}$$

$$= [a \ b \ c \ d] V \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}^T \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix} V^T \bar{x}$$

將 V^T 表示為 $V^T = \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \bar{r}_3 \\ \bar{r}_4 \end{bmatrix}$, \bar{x} 表示為 V^T 行向量線性組合 $\bar{x} = a_1 \bar{r}_1^T + a_2 \bar{r}_2^T + a_3 \bar{r}_3^T + a_4 \bar{r}_4^T$

$$\text{如此 } \bar{x}^T V D^T D V^T \bar{x} = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + a_3^2 \sigma_3^2 + a_4^2 \sigma_4^2$$

If $\sigma_4 \neq 0$, no exact solution

If $\sigma_4 = 0$, then nontrivial solution of \bar{x} is \bar{r}_4^T

取矩陣 $V_{4 \times 4}$ 的第 4 列即是尋找的平面方程式

3.3 以運動平面進行動作切割

在上一節中，介紹了如何替共面的一群點位找出其在空間中的最適合平面。然而，事先並無從得知何時該將動作分段，自然也就無從將同一片段的關節點位代入上一節的方法中去尋找平面。盲目的將點位代入計算若代入的點位分屬於兩

個或多個平面，則以這群點位擬合（Fitting）出來的平面，將不具任何意義。所以在做平面擬合之前，需要一個方法來判斷一群點位是否同屬一個平面。若是，才將這群點位做平面擬合（Plane Fitting）來估算肢節的運動平面。

判斷法則是從“同一平面上的點位，到平面的距離為零”來出發，但是因為資料來源並不是理想的情況，所求得的平面也只是最適合平面而已，只能改用“點位到其所屬的平面距離必然遠小於到其它平面的距離”，且設定一個容許值，來限定距離多少以內者視為同一平面，而在距離大於容許值時，將動作切開。

另外，還需要考慮到兩個動作平面間的少數影格。因為動作才剛發生變化，其點位與到上一平面的距離可能會小於我們設定的容許值，如此一來會被視為動作捕捉上的誤差而將加入上一個運動平面中。要等到動作發生一段時間後，點位與上一個運動平面的距離大於容許值時才會進行切割。如此一來，得到的分段點會與真實的分段點會有時間的延遲，且將不屬於平面的點位加入，會導致擬合出來的平面發生偏移，此一偏移會連帶影響到下一個平面的準確性，而使得切割的結果愈來愈不精確。為了避免這樣的誤差一直傳遞下去，將這少數影格捨棄不列入平面擬合的計算中。

我們提出了下列的方法：

Plane Segmentation Algorithm

Input: a set of points $S = [p_1, p_2, p_3, \dots]$, and distance threshold d_{th}

Output: a segmented sets of points $S_1 = [p_1, p_2, \dots, p_{t_1}]$, $S_2 = [p_{t_1+1}, p_{t_1+2}, \dots]$, ...

Method:

1. **Form the initial planes** $E = [E_1, E_2, \dots, E_n]$

(1) For the input set S, assume that the first m points to form a initial plane

E_1' .

And initiate $S_1' = \{p_1, p_2, \dots, p_m\}$.

(2) Check if a point $p_i \in \{p_{m+1}, p_{m+2}, \dots\}$ belongs to E_1' ,

i.e., if $d(p_i, E_1') < d_{th}$, then add p_i to S_1' , and refit E_1' with S_1' .

Until $d(p_i, E_1') > d_{th}$. Let this point be p_{t_1} .

(3) Then remove the last k points,

$$S_1 = S_1' - \{p_{t_1-k}, \dots, p_{t_1-1}, p_{t_1}\} = \{p_1, p_2, \dots, p_{t_1-k-1}\},$$

and refit $E_1 = E_1'$ with S_1 .

(4) For the remaining points of the input set $S = S - S_1'$, repeat part (1) ~ (3) to obtain E_2, \dots, E_n .

2. **Refine** $E = [E_1, E_2, \dots, E_n]$ **to remove outlier points.**

(1) Choose l outlier points whose distances d_i are above the threshold.

$i = 1 \sim n$.

(2) By removing them, refine E_i with the remaining points.

透過上述的方法，即可將輸入的點位集合切分成數個共面的點位集合。如此一來，就可以利用上一節中的方法來計算運動平面。

然而，這樣的方法雖然能將共面的點位集合切割出來，但是，第二種共圓錐模式的運動卻會被切割成多個平面。因此，我們還需要對切出來的平面進行檢查，將能形成共圓錐模式的平面合併（Merge），找出其真正的運動平面。

一組散佈在三維空間中的點位集合，其維度自然為三。但如果可以在空間中找到一個平面，該平面包含這組點位集合，則可推論這組點位的維度從三維退化

(degenerate) 到二維。而利用這樣的觀念，我們就可以檢查一群點位集合中的所有點是不是同屬一個平面。

隨意將不同平面的點位合併起來，自然無法在空間中找到適合的平面，所以該群點位的維度還是三維。但原屬一個平面卻被誤切成數個平面的共圓錐模式，雖然在上述的方法中過度切割成數個片段，但是將這些平面合併起來，因為該平面確實存在，所以我們檢驗其維度還是會退化到二維。

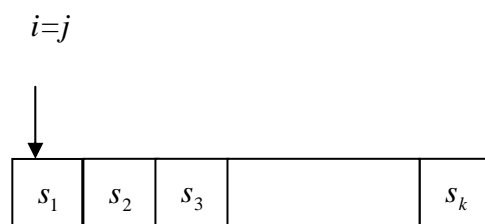
同樣的，由於誤差的的存在，其退化的特徵值 (Eigenvalues) 並不會為零，而是一個相對小的量。因此，我們必須設定一個容許值來判斷是否有維度退化的情形。以下是對此類型平面作合併的演算法：

Plane Merge Algorithm

Input: A series of segmented sets of points $s_1, s_2, \dots, s_k \in 2D$, each one with sorted eigenvalues $\lambda_1, \lambda_2, \lambda_3$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3$. And a threshold of eigenvalue, e_{th} .

Output: A merged sets of points $S_1, S_2, \dots, S_m, m < k$

Method:



1. Initial i points to $s_1 (i = 1)$.
- 2.

For the set s_j and s_{j+1} , merge them, and recalculate its dimension.

i.e.,

<i> Find out the eigenvalues $\lambda_1, \lambda_2, \lambda_3$,

<ii> Normalize $\lambda_1, \lambda_2, \lambda_3$, such that $\lambda_1 = 1, \lambda_2 = \frac{\lambda_2}{\lambda_1}, \lambda_3 = \frac{\lambda_3}{\lambda_1}$

<iii> Compare if $\lambda_3 < e_{th}$.

(1) If $\lambda_3 < e_{th}$, then the dimension of the merged set is 2.

j set to $j+1$, and repeat part 2..

(2) Else, the dimension of the merged set is 3.

then split s_j and s_{j+1} , merge the sets $s_i \sim s_j$ into the larger one.

i.e., $S_m = \{s_i, s_2, \dots, s_j\}$.

i set to $j+1$, j set to i , m set to $m+1$, and repeat part 2..

Until $s_i = s_k$.

3.4 計算轉動角度，並利用 AR Model 再進行動作切割

根據[14][15][16]，我們可以用 Second order AR Model 來描述將關節角度的變化，並藉由觀察參數矩陣的變化來進行動作切割[16]。

3.4.1 計算旋轉角度

在套用 AR Model 切割之前，先介紹如何計算旋轉角度。在 3.3 求解出運動平面後，把 t 時刻的運動平面法向量 \vec{n} ，當成 t 時刻的旋轉軸，代入 quaternion 的運算法則中，即可求得從 $t-1$ 時刻肢節 \vec{v}_{t-1} 旋轉到 t 時刻肢節 \vec{v}_t 的旋轉角度 θ_t 。

$$\vec{v}_t = \left(\hat{A} + (I - \hat{A})\cos\theta_t + (A^*)\sin\theta_t \right) \vec{v}_{t-1} \quad (3.1)$$

其中，

$$\hat{A} = \vec{n}\vec{n}^T = \begin{bmatrix} n_x \cdot n_x & n_x \cdot n_y & n_x \cdot n_z \\ n_y \cdot n_x & n_y \cdot n_y & n_y \cdot n_z \\ n_z \cdot n_x & n_z \cdot n_y & n_z \cdot n_z \end{bmatrix}, (A^*)\vec{v}_{t-1} = (\vec{n} \times \vec{v}_{t-1}) = \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_{t-1_x} \\ \vec{v}_{t-1_y} \\ \vec{v}_{t-1_z} \end{bmatrix}$$

求解(3.1)的線性方程式，即可算出 θ_t

3.4.2 Second order AR model 簡介

AR (auto-regression) Model 是用來描述一組資料中，其 t 時刻狀態 $x(t)$ 與先前時刻狀態 $x(t-1), x(t-2), \dots$ 有相關的一種動態模型 (Dynamic Model)，一階 AR 模型 (First-order AR Model) 表示為

$$x(t) = c_1 x(t-1) + e(t),$$

t 時刻的資料 $x(t)$ 是由 $t-1$ 時刻的資料 $x(t-1)$ 再加上一個隨機亂數 $e(t)$ 所組成，係數 c_1 用來表示與先前資料的關係。

然後在[17]中提到，對於人體運動的關節角度變化無法用一階 AR 模型來完整描述，需要二階 (Second-order) 或更高次的 AR 模型。但隨著 order 的增加，計算量也跟著大幅增加。而[17]中也說明二階 AR 模型即足夠，超過二階在描述上並無顯著的改進，但卻要付出大量的計算。因此，採用二階 AR 模型來進行手部運動分析，即表示為

$$x(t) = c_1 x(t-1) + c_2 x(t-2) + e(t)$$

為了接下來分析上的便利，再令 $X(t) = \begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix}$ ，則可將二階 AR 模型以矩陣的形式表示成

$$X(t) = A \cdot X(t-1) + B \cdot e(t)$$

其中， $A = \begin{bmatrix} 0 & I \\ A_2 & A_1 \end{bmatrix}$ 用來表示模型中的決定性參數， $B = \begin{bmatrix} 0 \\ B_0 \end{bmatrix}$ ，用來表示模型中的隨機性參數， e_t 為 zero mean 的高斯雜訊。

以手部運動為例，手部運動需要考慮肩膀及手肘這兩個關節的角度變化，此時表示在 AR Model 中的狀態向量就是肩膀與手肘這兩個關節的旋轉量，即

$$x(t) = \begin{bmatrix} \theta_{shoulder}(t) \\ \theta_{elbow}(t) \end{bmatrix}。$$

3.4.3 AR Model Segmentation Algorithm

在一個穩定的運動中，關節角度應該呈現出穩定增加或減少的情況。故在此一穩定運動過程中，其 AR 模型中的參數矩陣 A 應該保持不變。換句話說，當動作發生變化時，參數矩陣 A 也必需修正以符合新的變化模式。如此一來，動作切割的問題就轉變成偵測參數矩陣 A 變化的問題，並在參數矩陣 A 發生劇烈變化時予以切割。對於這樣的問題，可以採用 recursive system estimation 的方法來求解[18]。

利用 $W(t)$ 來表示參數矩陣變化的部分，參數矩陣 A 變化在 State Space Model 中表示如下：[16]

$$\begin{aligned} A(t+1) &= A(t) + W(t) \\ X(t) &= A(t) \cdot X(t-1) + E(t) \end{aligned} \quad (3.2)$$

透過最小化觀察值 $X(k), k=1 \sim t$ 與預測值 $\hat{X}(k) = A \cdot X(k-1), k=1 \sim t$ 的誤差，可以計算出參數矩陣 $A(t)$ [16]

$$\begin{aligned} V_t(A(t)) &= \sum_{k=1}^t \beta(t, k) \|X(k) - A(t) \cdot X(k-1)\|^2 \\ \beta(t, k) &= \prod_{j=k}^t \lambda, \lambda \text{ is a constant} \end{aligned} \quad (3.3)$$

對(3.2)式的 A 為偏微分，可以得到估算的參數矩陣 $\hat{A}^T(t)$ ，

$$\hat{A}^T(t) = \left(\sum_{k=1}^t \beta(t, k) X(k-1) \cdot X^T(k-1) \right)^{-1} \sum_{k=1}^t \beta(t, k) X(t-1) \cdot X^T(t) \quad (3.4)$$

再將計算相鄰時刻參數矩陣 $A(t)$ 的差異 $d(t) = \|\hat{A}(t) - \hat{A}(t-1)\|$ ， $d(t)$ 的大小反映出相鄰矩陣的差異程度，同時也表現出系統的變化程度。如此一來，可以取 $d(t)$ 中的區域最大值（Local Maximum）當作動作的切割點。

根據[14]提出的動作切割演算法總結如下：

AR Model Segmentation Algorithm

- Step 1. 將輸入的角度資料 $X(t), t = 1 \sim n$ 利用 low pass filter 去除雜訊，並減去平均值，得 $Y(t), t = 1 \sim n$
- Step 2. 將角度資料 $Y(t)$ 代入公式 (3.2) 中估算參數矩陣 A ，並計算相臨時刻的差異矩陣 $D(t) = \hat{A}(t) - \hat{A}(t-1)$
- Step 3. 計算差異矩陣的 norm $d(t) = \|D(t)\| = \|\hat{A}(t) - \hat{A}(t-1)\|$ ，在這裡利用 MSE (mean square error) 來計算 matrix norm， $A = (a_{ij})_{m \times n}$ ， $\|A\| = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2$
- Step 4. 將 $d(t)$ 代入 1D RLS
- Step 5. 從 Step4 中的結果，取出序列中的最大值，該處即是切割點
- Step 6. 將切割出來的片段，進行動態模型擬合 (Dynamic Model Fitting)，若是模型計算出的資料與觀察到的資料誤差超過容許值時，回到 Step5，對此一片段重新計算切割點

經過上述演算法的處理後，我們得到一序列的動作片段且每個片段都能有一個 AR Model 描述。



3.5 動作片段的表示

只要給定動態模型的參數矩陣 A 及初始狀態向量 $X(0)$ ，就可以根據該動態模型的定義來計算 t 時刻的狀態向量 $X(t)$ 。

$$\begin{aligned} X(t) &= AX(t-1) \\ &= A \cdot A\bar{X}(t-2) = A^2 X(t-2) \\ &= \dots \\ &= A^t X(0) \end{aligned} \tag{3.5}$$

根據參數矩陣 A 的特徵值 (Eigenvalues)，可以分成以下兩種情況：

Case 1: 特徵值為非複數

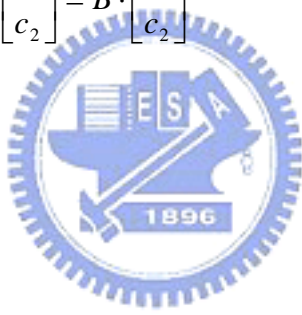
令參數矩陣 A 的特徵值為 λ_1, λ_2 ，特徵向量為 \bar{v}_1, \bar{v}_2 ，則初始向量 $X(0)$ 即可表示成：

$$X(0) = c_1 \bar{v}_1 + c_2 \bar{v}_2 \quad (3.6)$$

將(3.6)代入(3.5)，得：

$$\begin{aligned} X(t) &= A^t X(0) \\ &= [\bar{v}_1 \quad \bar{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^t [\bar{v}_1 \quad \bar{v}_2]^{-1} X(0) \\ &= [\bar{v}_1 \quad \bar{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^t \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= c_1 \lambda_1^t \bar{v}_1 + c_2 \lambda_2^t \bar{v}_2 \end{aligned}$$

利用 Least square 從已知的觀察值 $X(k), k = 1 \sim t$ 可推算 $X(0)$ 的係數 $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ ，即

$$\begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(n) \end{bmatrix} = \underbrace{\begin{bmatrix} [\bar{v}_1 \quad \bar{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^1 \\ \vdots \\ [\bar{v}_1 \quad \bar{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n \end{bmatrix}}_{2n-by-2} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = B \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$


$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = (B^T B)^{-1} B^T \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(n) \end{bmatrix}$$

因此，即可利用初始向量 $X(0)$ 及參數矩陣 A 來表示此組資料。

Case 2: 特徵值為複數

令參數矩陣 A 的特徵值為 $\lambda_{1,2} = p \pm iq = r(\cos \phi + i \sin \phi)$ ，特徵向量為 $\bar{v} \pm i\bar{w}$ ，

$$S = [\bar{w} \quad \bar{v}]$$

Then

$$\begin{aligned}
X(t) &= r^t S \begin{bmatrix} \cos(\phi t) & -\sin(\phi t) \\ \sin(\phi t) & \cos(\phi t) \end{bmatrix} S^{-1} X(0) \\
&= r^t \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} \cos(\phi t) & -\sin(\phi t) \\ \sin(\phi t) & \cos(\phi t) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
&= r^t \begin{bmatrix} s_{11} \cos(\phi t) + s_{12} \sin(\phi t) & -s_{11} \sin(\phi t) + s_{12} \cos(\phi t) \\ s_{21} \cos(\phi t) + s_{22} \sin(\phi t) & -s_{21} \sin(\phi t) + s_{22} \cos(\phi t) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
&= r^t \begin{bmatrix} c_1 s_{11} \cos(\phi t) + c_1 s_{12} \sin(\phi t) - c_2 s_{11} \sin(\phi t) + c_2 s_{12} \cos(\phi t) \\ c_1 s_{21} \cos(\phi t) + c_1 s_{22} \sin(\phi t) - c_2 s_{21} \sin(\phi t) + c_2 s_{22} \cos(\phi t) \end{bmatrix} \\
&= r^t \begin{bmatrix} (c_1 s_{11} + c_2 s_{12}) \cos(\phi t) + (c_1 s_{12} - c_2 s_{11}) \sin(\phi t) \\ (c_1 s_{21} + c_2 s_{22}) \cos(\phi t) + (c_1 s_{22} - c_2 s_{21}) \sin(\phi t) \end{bmatrix} \\
&= r^t \begin{bmatrix} a \cos(\phi t) + b \sin(\phi t) \\ c \cos(\phi t) + d \sin(\phi t) \end{bmatrix}, \begin{cases} a = c_1 s_{11} + c_2 s_{12} \\ b = c_1 s_{12} - c_2 s_{11} \\ c = c_1 s_{21} + c_2 s_{22} \\ d = c_1 s_{22} - c_2 s_{21} \end{cases} \\
&= r^t \begin{bmatrix} \sqrt{a^2 + b^2} \cos(\phi t + h_1) \\ \sqrt{c^2 + d^2} \cos(\phi t + h_2) \end{bmatrix} \\
&= r^t \begin{bmatrix} m_1 \cos(\phi t + h_1) \\ m_2 \cos(\phi t + h_2) \end{bmatrix}
\end{aligned}$$



則未知數： m_1, m_2, h_1, h_2

已知 $X(t) = \begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix}$ ，所以只需要取 $x(t) = r^t m_2 \cos(\phi t + h_2)$ 代表，即可推算未知

數，利用 Least square 推算 m_2 與 h_2 ，即：

$$\begin{aligned}
\begin{bmatrix} x(1) \\ \vdots \\ x(n) \end{bmatrix} &= \begin{bmatrix} r \cdot m_2 \cos(\phi + h_2) \\ \vdots \\ r^n \cdot m_2 \cos(\phi n + h_2) \end{bmatrix} \\
&= \begin{bmatrix} r \cdot m_2 (\cos(\phi) \cos(h_2) - \sin(\phi) \sin(h_2)) \\ \vdots \\ r^n \cdot m_2 (\cos(\phi n) \cos(h_2) - \sin(\phi n) \sin(h_2)) \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} r \cos \phi & r \sin \phi \\ \vdots & \vdots \\ r^n \cos(\phi n) & r^n \sin(\phi n) \end{bmatrix}}_{n-by-2} \underbrace{\begin{bmatrix} m_2 \cos(h_2) \\ m_2 (-\sin(h_2)) \end{bmatrix}}_{2-by-1} = B \cdot \begin{bmatrix} m_2 \cos(h_2) \\ m_2 (-\sin(h_2)) \end{bmatrix}
\end{aligned}$$

$$\begin{bmatrix} m_2 \cos(h_2) \\ m_2(-\sin(h_2)) \end{bmatrix} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{bmatrix}$$

得到

$$m_2 = \sqrt{(m_2 \cos(h_2))^2 + (m_2(-\sin(h_2)))^2} \quad (3.7)$$

$$h_2 = \cos^{-1}\left(\frac{m_2 \cos(h_2)}{m_2}\right) \quad (3.8)$$



四 實驗數據與結果

4.1 檢驗維度退化之容許值參數設定

在 3.3 節中，提到了點位集合的維度退化情形，下面以模擬數據（理想狀態與加入亂數）來觀察特徵值變化的情形，且藉此來設定演算法中的容許值大小。

實驗一：

首先，設定肩膀坐標為 $[0 \ 0 \ 0]^T$ 手肘坐標為 $[17.6777 \ 0 \ 17.6777]^T$ ，依旋轉軸 $[1 \ 3 \ 2]^T$ 以等角速度旋轉 360 度後，得到一組點位集合其運動平面是共圓錐模式。運動情形如（圖 4-1），共 120 個影格。

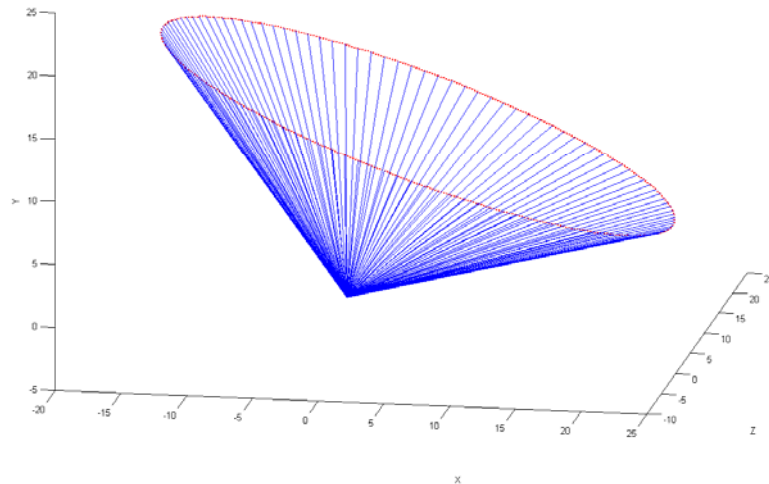


圖 4 - 1 實驗一：3D 運動軌跡

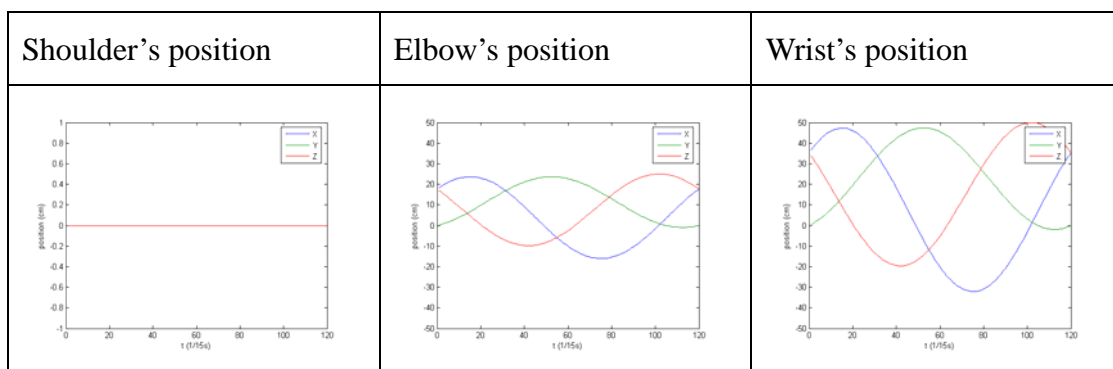


圖 4-2 實驗一：各關節點位

代入 3.3 中的平面切割演算法可得到 5 個動作片段如（表格 1）：

表格 1 實驗一：平面切割結果

Segments	Start Frame	End Frame
#1	1	21
#2	27	47
#3	53	73
#4	79	99
#5	105	120

將各片段中的手肘點位分別排列成一個 n -by-3 的矩陣後，利用 SVD 分解可得到三個奇異值（Singular Eigenvalues） λ_1 、 λ_2 、 λ_3 。其結果如（表格 2）：

表格 2 實驗一：各片段之奇異值

Segment #1	λ_1	λ_2	λ_3
eigenvalue	29.0339	4.1386	0.0000
比例	1.0000	0.1425	0.0000
Segment #2	λ_1	λ_2	λ_3
eigenvalue	29.0339	4.1386	0.0000
比例	1.0000	0.1425	0.0000
Segment #3	λ_1	λ_2	λ_3
eigenvalue	29.0339	4.1386	0.0000
比例	1.0000	0.1425	0.0000
Segment #4	λ_1	λ_2	λ_3

eigenvalue	29.0339	4.1386	0.0000
比例	1.0000	0.1425	0.0000
Segment #5	λ_1	λ_2	λ_3
eigenvalue	19.5398	2.1074	0.0000
比例	1.0000	0.1079	0.0000

再將片段合併起來計算奇異值，其結果如（表格 3）：

表格 3 實驗一：片段合併之奇異值

Merge Segment #1~2	λ_1	λ_2	λ_3
eigenvalue	86.0528	26.2373	0.0000
比例	1.0000	0.3049	0.0000
Merge Segment #1~3	λ_1	λ_2	λ_3
eigenvalue	127.8490	70.6904	0.0000
比例	1.0000	0.5529	0.0000
Merge Segment #1~4	λ_1	λ_2	λ_3
eigenvalue	143.3715	119.2874	0.0000
比例	1.0000	0.8320	0.0000
Merge Segment #1~5	λ_1	λ_2	λ_3
eigenvalue	148.1827	142.8590	0.0000
比例	1.0000	0.9641	0.0000

最後發現將五個片段合併起來其維度仍然是二，因此，可以推斷空間中存在一個平面能包含所有的點位，而該平面才是此動作真正的運動平面。而這樣的結果也與之前的設計相同

實驗二：

接著，爲了貼近真實情況，將上述的理想點位集合加入一個常態分佈的隨機亂數（平均值爲 0，標準差爲 0.7），使得產生出來的數據能與 Motion Capture 得到的資料相似。運動情形如（圖 4-3），共 120 個影格。

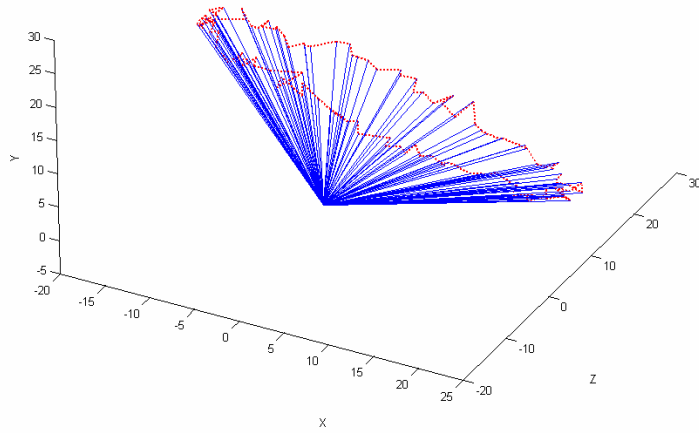


圖 4 - 3 實驗二：3D 運動軌跡

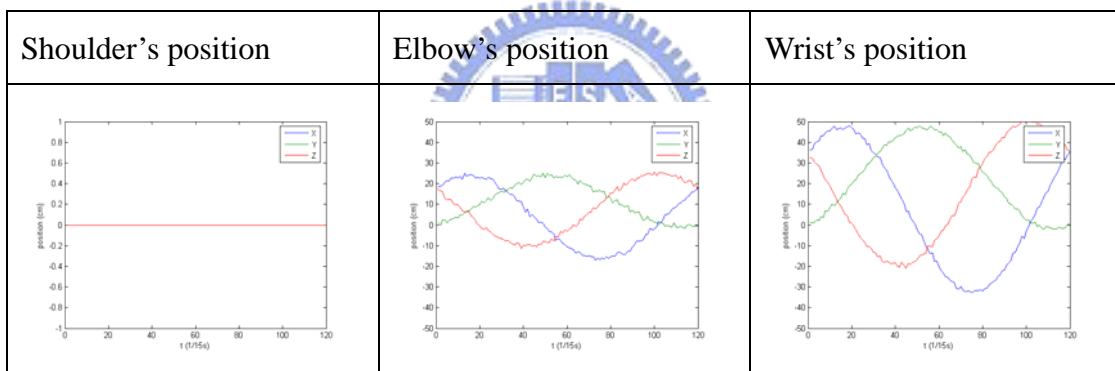


圖 4 - 4 實驗二：各關節點位

同樣的，依上述的流程可以得到表 4、表 5、表 6 的結果：

表格 4 實驗二：平面切割結果

Segments	Start Frame	End Frame
#1	1	21
#2	27	47
#3	53	73
#4	79	99
#5	105	120

表格 5 實驗二：各片段之奇異值

Segment #1	λ_1	λ_2	λ_3

eigenvalue	30.1621	4.1783	3.6137
比例	1.0000	0.1385	0.1198
Segment #2	λ_1	λ_2	λ_3
eigenvalue	21.3196	5.0758	2.2967
比例	1.0000	0.2381	0.1077
Segment #3	λ_1	λ_2	λ_3
eigenvalue	19.3592	4.2322	2.9449
比例	1.0000	0.2186	0.1521
Segment #4	λ_1	λ_2	λ_3
eigenvalue	26.4485	5.4504	2.4696
比例	1.0000	0.2061	0.0934
Segment #5	λ_1	λ_2	λ_3
eigenvalue	38.8900	7.4889	3.5531
比例	1.0000	0.1926	0.0914

表格 6 實驗二：片段合併之奇異值

Merge Segment #1~2	λ_1	λ_2	λ_3
eigenvalue	76.8263	21.3422	4.7751
比例	1.0000	0.2778	0.0622
Merge Segment #1~3	λ_1	λ_2	λ_3
eigenvalue	113.7831	54.0074	5.6508
比例	1.0000	0.4747	0.0497
Merge Segment #1~4	λ_1	λ_2	λ_3
eigenvalue	139.3933	99.1061	6.2470
比例	1.0000	0.7110	0.0448
Merge Segment #1~5	λ_1	λ_2	λ_3
eigenvalue	147.8146	142.0463	7.4151
比例	1.0000	0.9610	0.0502

實驗中對每一個動作片段做 SVD 分解並觀察奇異值的變化，可以觀察到維

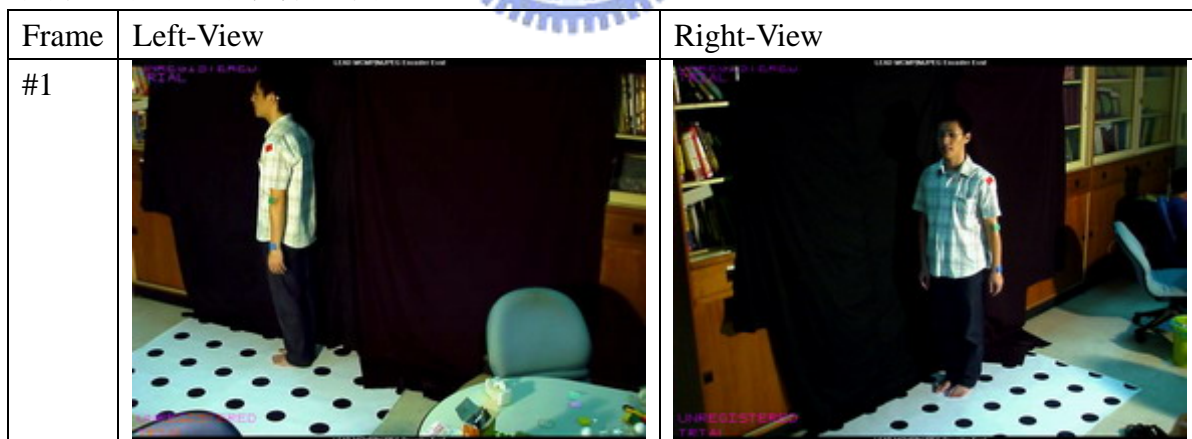
度退化的情況，即在一個完整的平面運動中 λ_1 、 λ_2 值的大小相似， λ_3 則衰減到很小，另外，在過度切割的時候，因為片段中的點位太少，甚至有退化到一維的情況（ λ_2 也很小）。依表格 6 中的數據及實驗的實際數據，本文中將容許值設為 0.1。

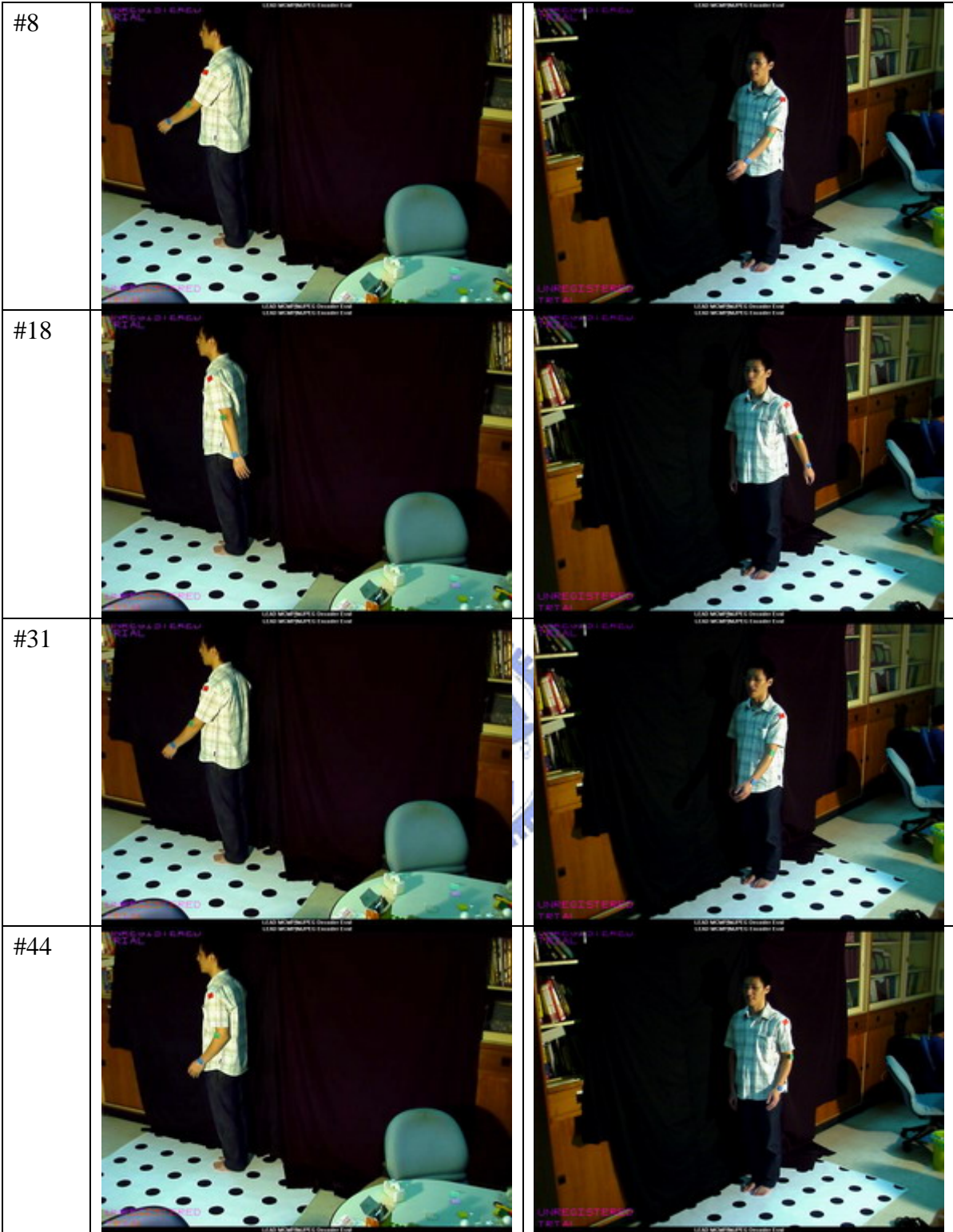
4.2 同平面上反覆運動的切割結果

在平面切割的結果中，會遭遇到一種情況即是運動在同一平面上連續重覆發生或是在同一平面上的兩種動作連續發生。當這樣的運動發生時，依 3.3 中的方法，會將所有的連續重覆動作劃分在同一片段，其原因在於運動過程中，手肘點位皆在同一面上運動，而這樣的切割顯得過於粗糙，因此還需要利用運動過程中角度的變化情形再做一次切割。以”擺手”與”出拳”這兩種動作所組合的運動來測試。

實驗三：

擺手與出拳混合運動，共 101 張影格，#1~40 為兩次擺手運動，#41~105 為兩次出拳運動，運動影像擷取如（圖 4-5）：





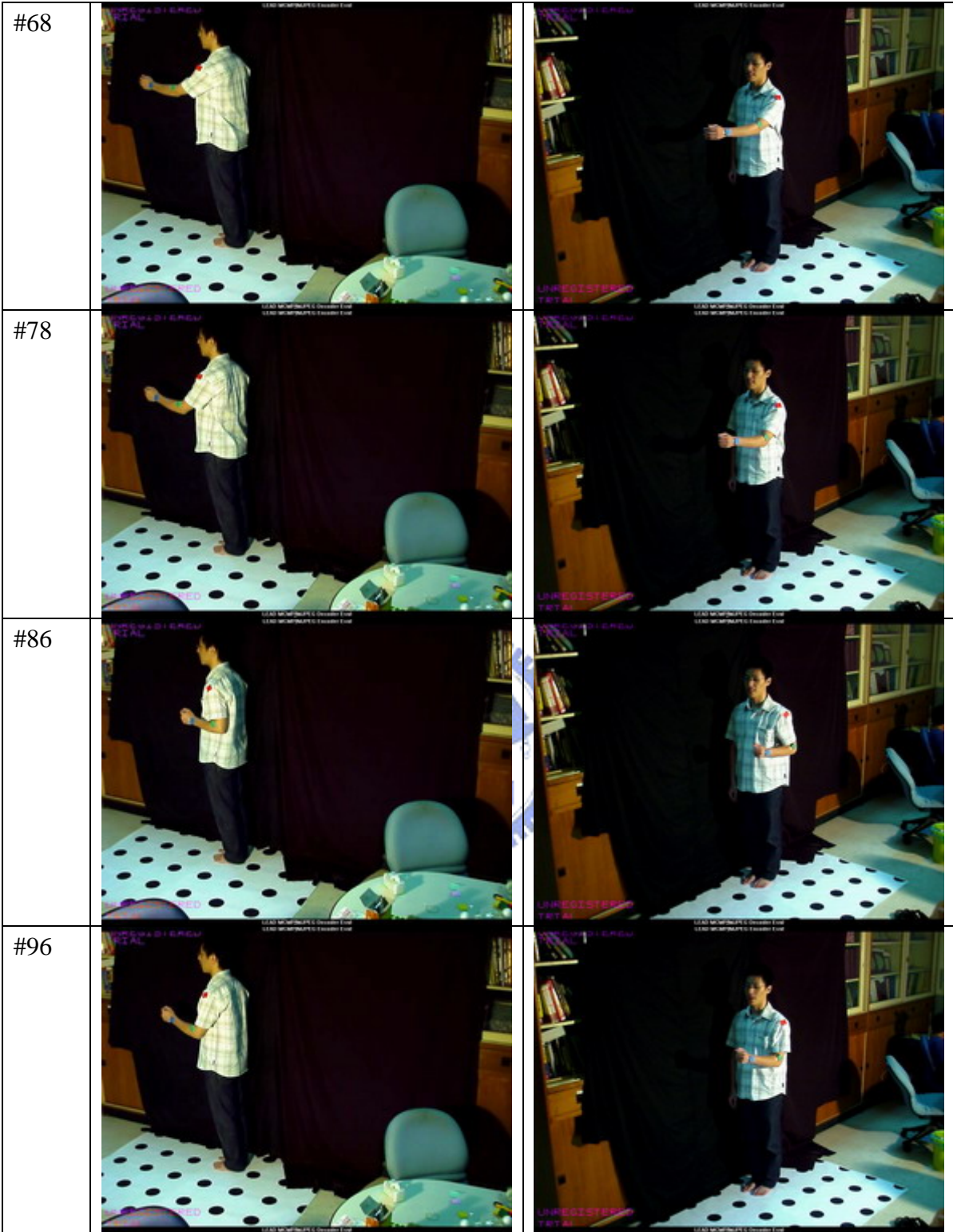
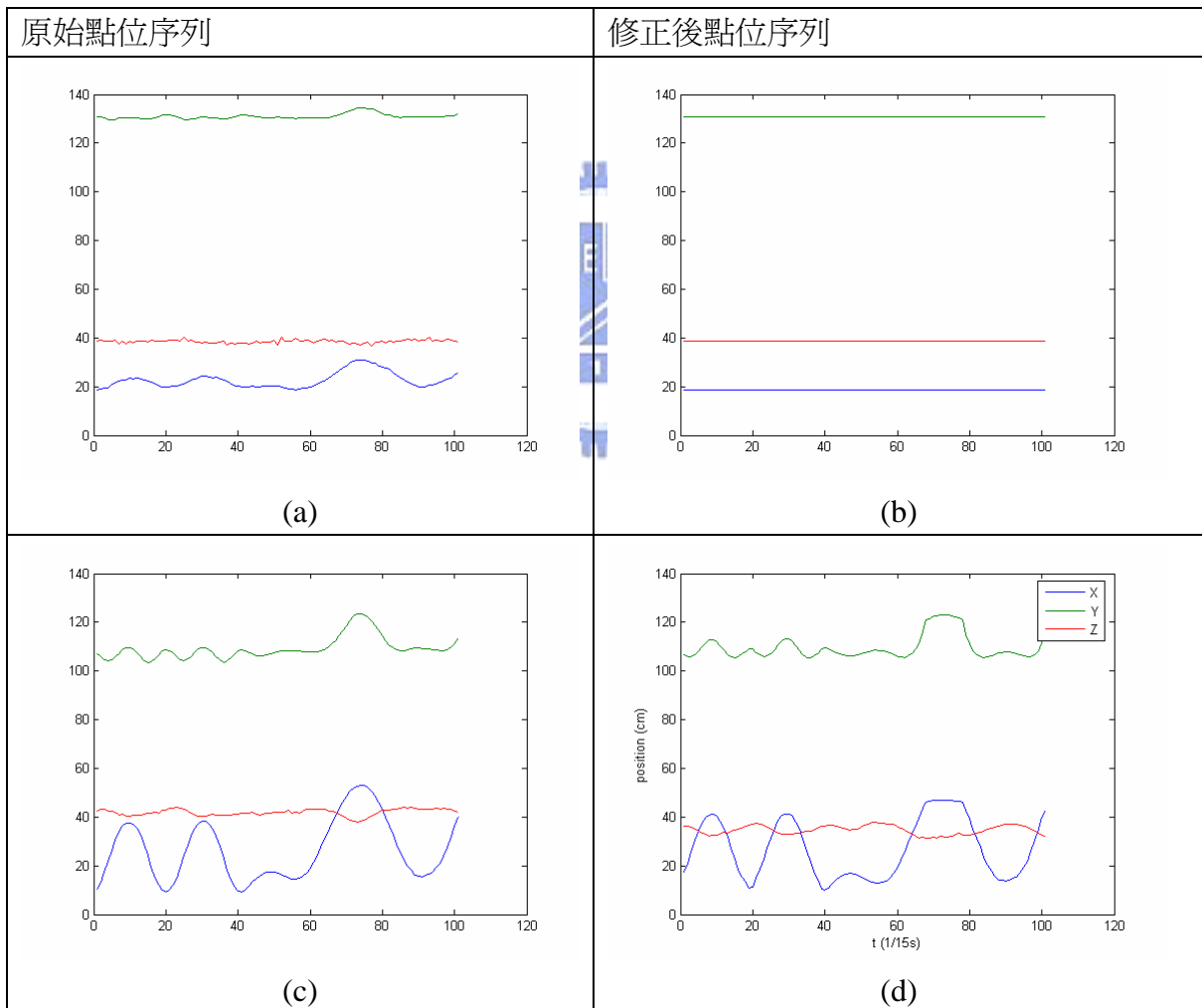




圖 4-5 實驗三：動作擷圖

首先，以 2.2, 2.3 中的方法修正動作捕捉系統取得關節點位坐標的誤差。(圖 4-6)



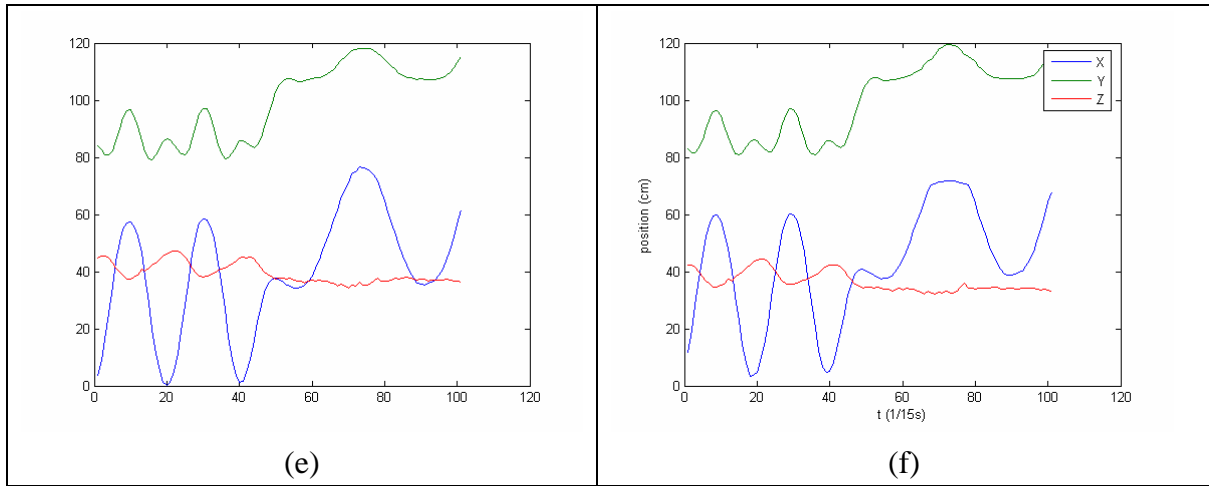


圖 4-6 實驗三：各關節點位座標

(a)(b)為原始及修正的肩膀點位序列，(c)(d)為原始及修正的手肘點位序列，
(e)(f)為原始及修正的手腕點位序列

將修正後的點位序列代入 3.3 的方法，得到的平面切割結果（表格 7）顯示此組運動為同一平面。該平面的法向量為 $[-0.0701, -0.2305, -0.9705]$ ，所有點位到平面的平均距離為 0.8682 (cm)，標準差為 1.07 (cm)

表格 7 實驗三：平面切割結果

Segments	Start Frame	End Frame
#1	1	101

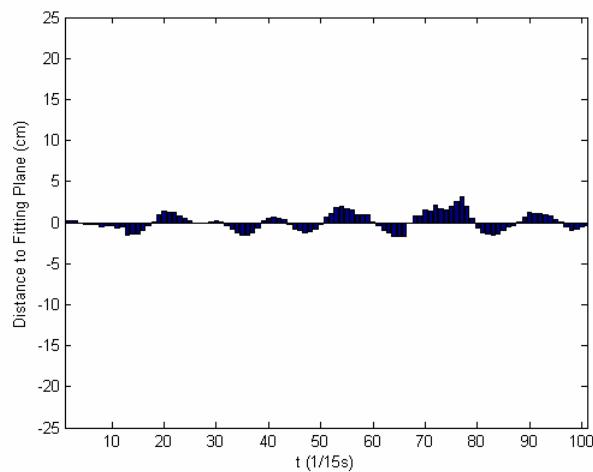


圖 4-7 實驗三：點位到平面誤差

計算出運動平面後，將平面的法向量當成旋轉軸，依 3.4.1 的方法計算肩膀的旋轉角度序列 $\theta_{\text{shoulder}}(t)$ ，依上臂與前臂的夾角計算出手肘關節的旋轉角度序列

$\theta_{\text{elbow}}(t)$ 。再將角度序列表示成 AR Model 的狀態向量序列後，代入計算(表格 8)。

得到的切割點即是 [1, 8, 18, 31, 44, 68, 78, 86, 96, 101] (表格 9) (圖 4-8)

表格 8 實驗三：AR 參數矩陣

Segment #1 Matrix A				Segment #6 Matrix A			
0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1
-1.1483	-1.4154	2.0500	0.8743	-1.0665	-0.2095	2.0941	0.2550
0.1663	-1.4100	-0.1457	0.5145	0.4353	-0.3968	-0.4595	1.3488
Segment #2 Matrix A				Segment #7 Matrix A			
0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1
-0.9332	-0.5685	1.8369	0.2148	-1.1216	-0.1992	2.0974	0.2195
-0.0005	-0.9969	0.0175	1.2910	0.6080	-0.3144	-0.5569	1.2884
Segment #3 Matrix A				Segment #8 Matrix A			
0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1
-1.0738	-0.2806	1.9894	0.5648	-0.8050	-0.0027	1.7671	-0.0005
0.0322	-0.4699	-0.0634	0.6938	0.4110	-0.4015	-0.3519	1.4140
Segment #4 Matrix A				Segment #9 Matrix A			
0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1
-0.9711	-0.3219	1.9221	0.5235	-1.0998	-0.1211	2.1146	0.1637
-0.0663	-0.5445	-0.0424	0.7370	0.6187	-0.3977	-0.6318	1.3664
Segment #5 Matrix A							
0	0	1	0				
0	0	0	1				
-0.9633	0.1236	1.8955	-0.1153				
-0.0584	-0.8407	-0.0093	1.8040				

表格 9 實驗三：AR 切割結果

Segments	Start Frame	End Frame
#1	1	8
#2	8	18
#3	18	31
#4	31	44
#5	44	68
#6	68	78
#7	78	86
#8	86	96
#9	96	101

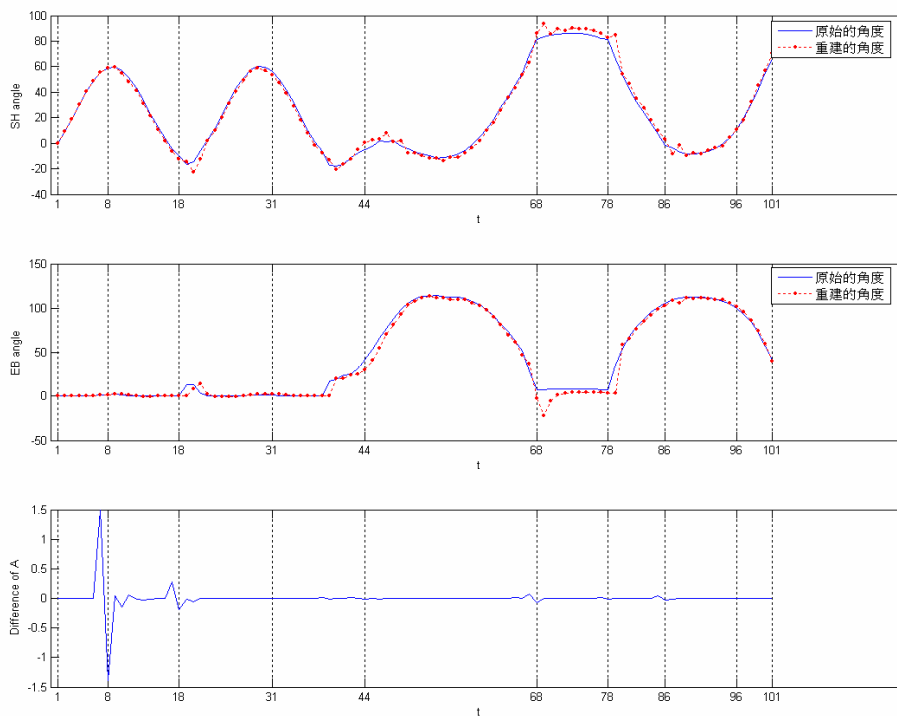


圖 4 - 8 實驗三：AR 切割結果

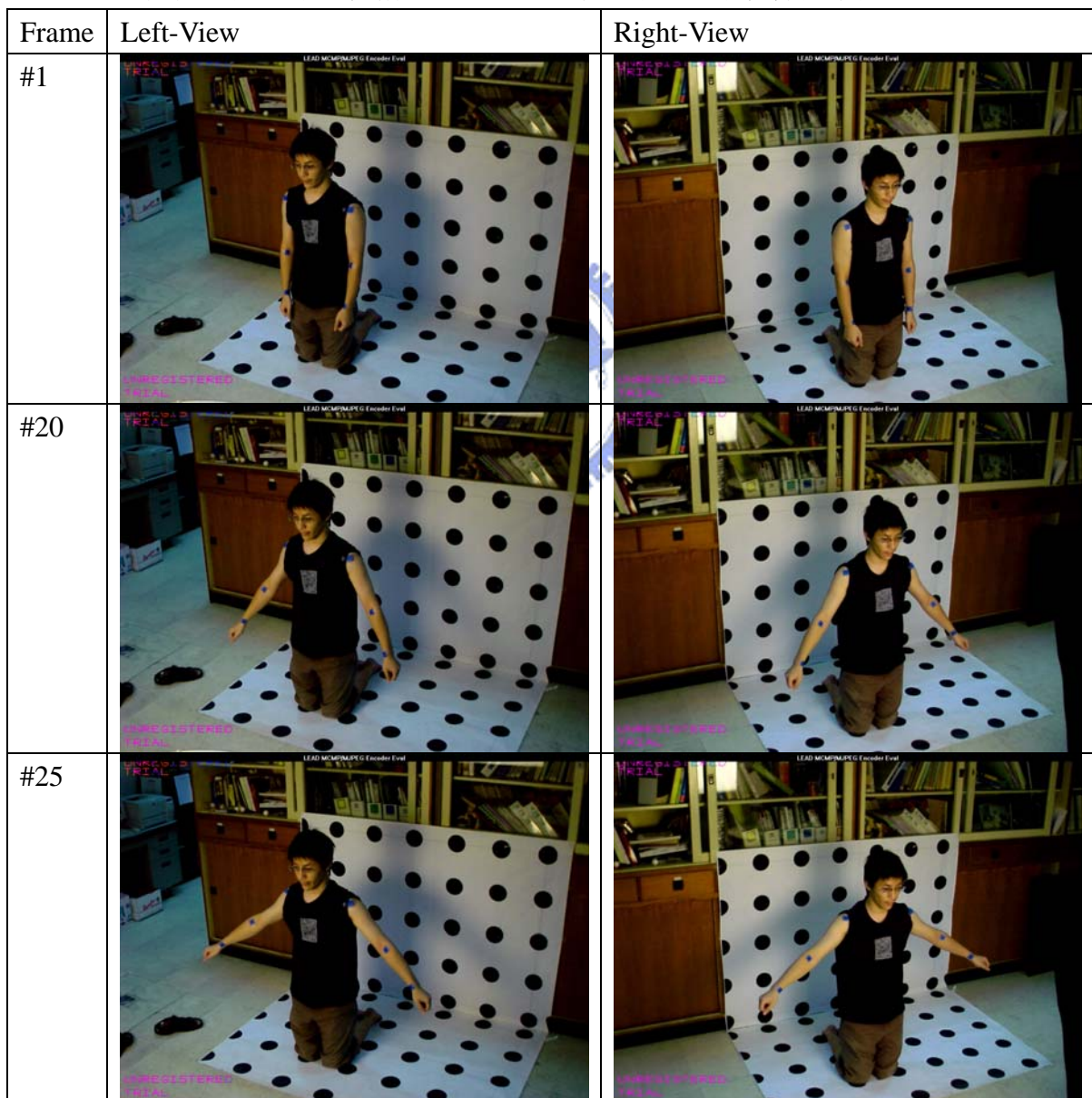
從 (圖 4-8) 中，可發現 Segment #1 為上擺手、Segment #2 下擺手、Segment #3 為上擺手、Segment #4 下擺手、Segment #5 出拳、Segment #6 為靜止、Segment #7 收拳、Segment #8 為出拳、Segment #9 為出拳，切割結果與影像相符。

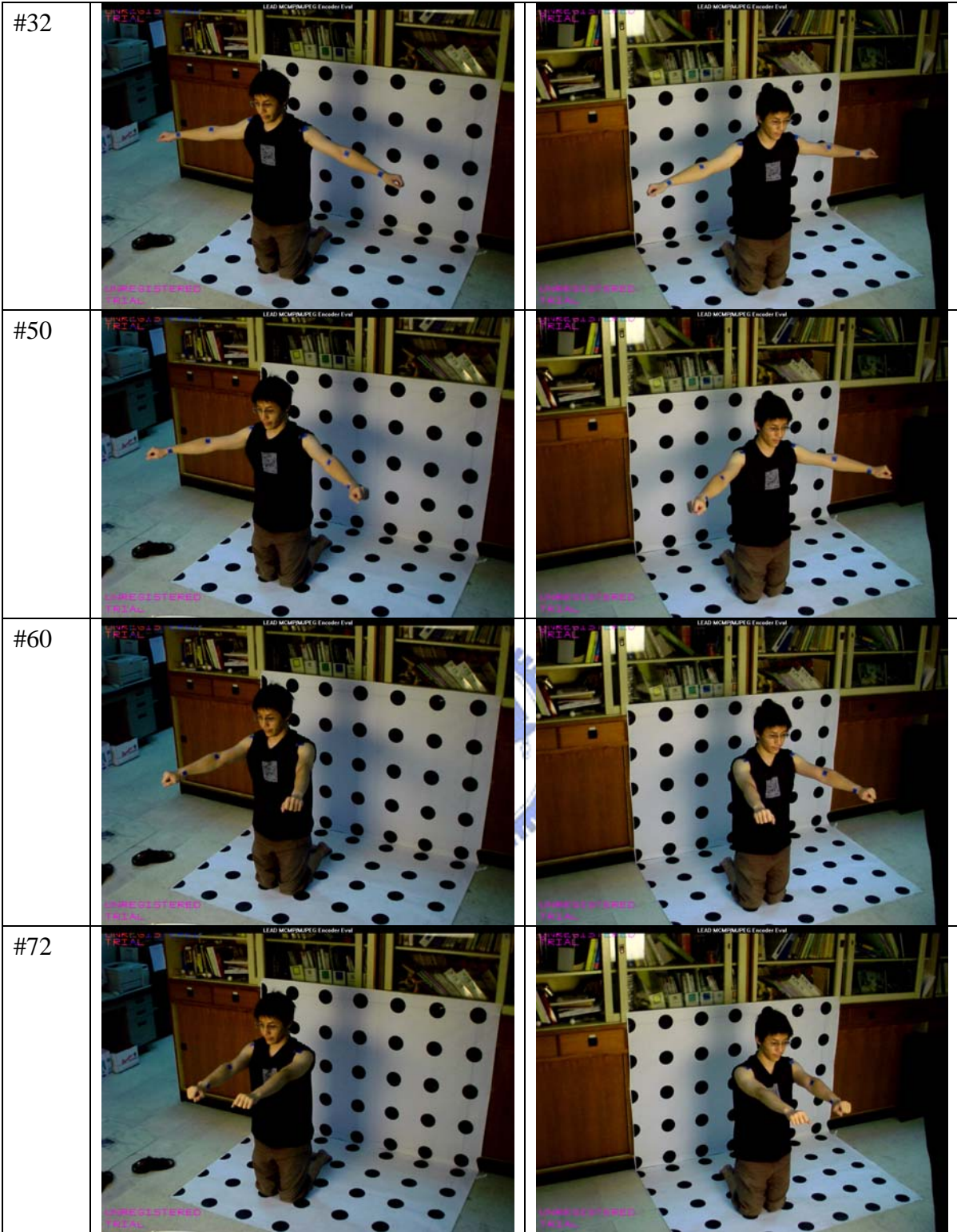
4.3 不同運動平面切割結果

透過實際運動資料，來檢驗本文中提出平面切割方法。

實驗四：

側抬、向前收攏、放下混合運動，共 110 個影格，影格#1~32 為側抬運動，影格 #33~72 為向前收攏運動，影格#73~110 為放下運動，運動影像擷取如（圖 4-9）：





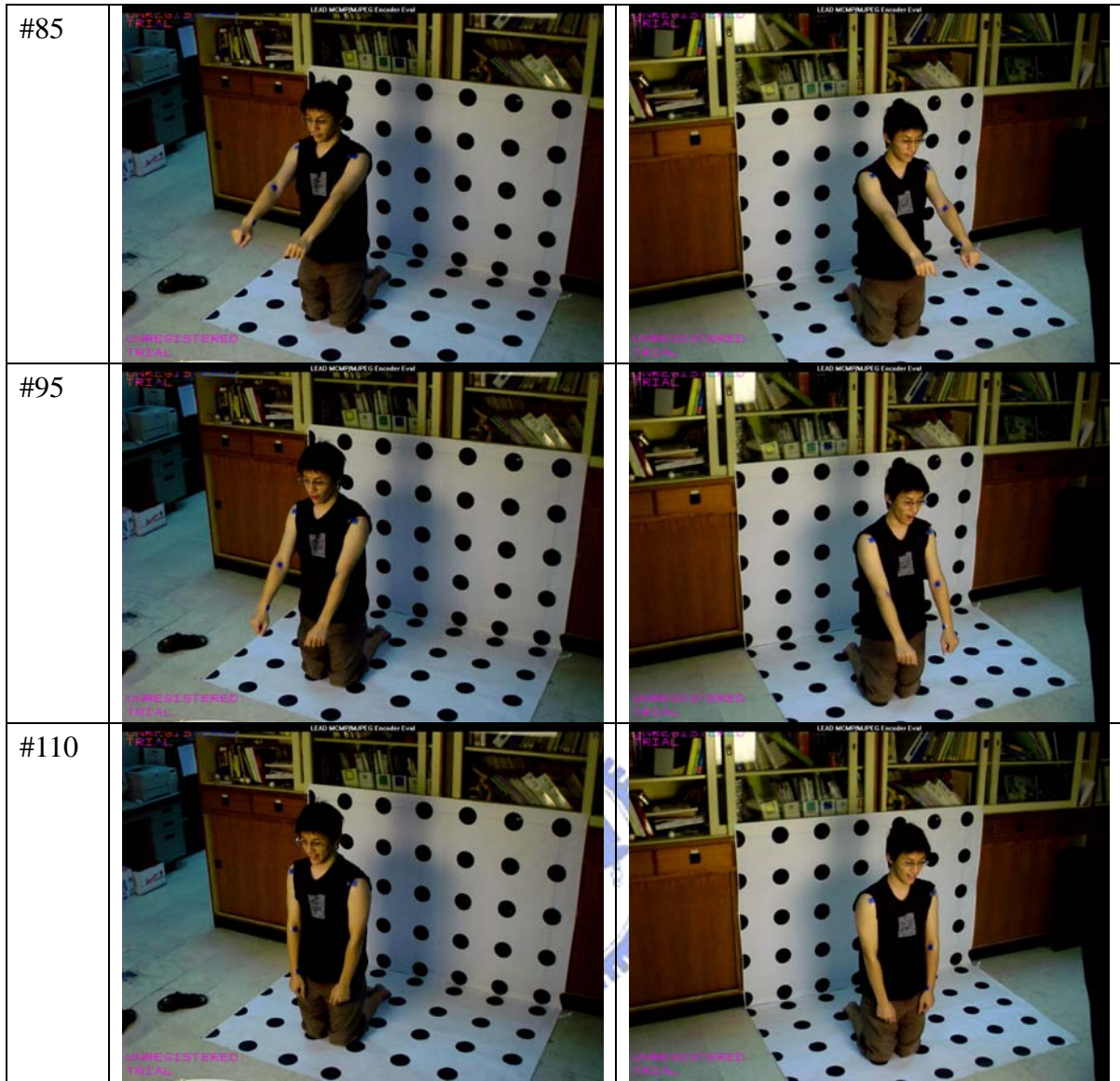
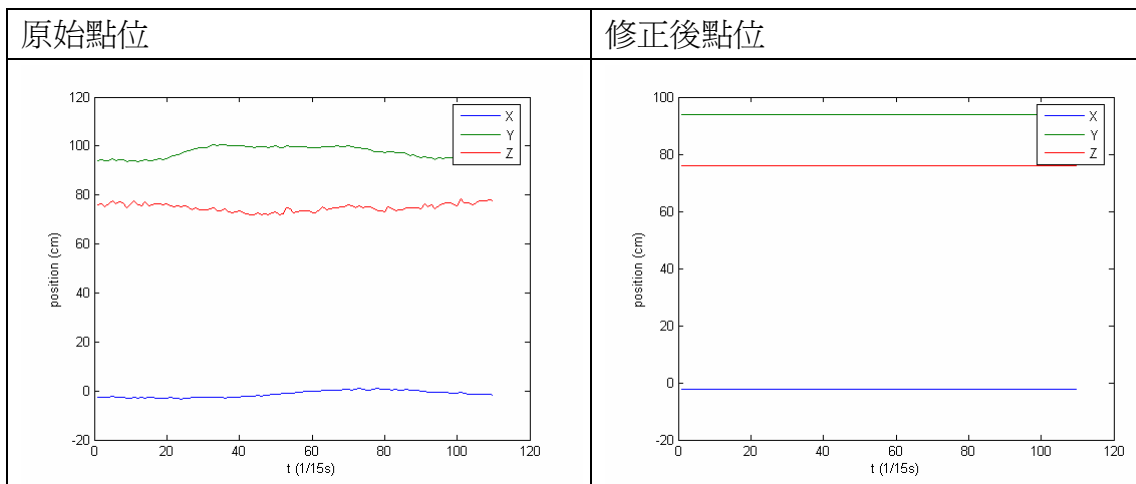


圖 4-9 實驗四：動作擷圖

同樣，利用 Motion Capture 取得表示運動的點位集合，再依 2.2，2.3 中的方法修正。(圖 4-10)



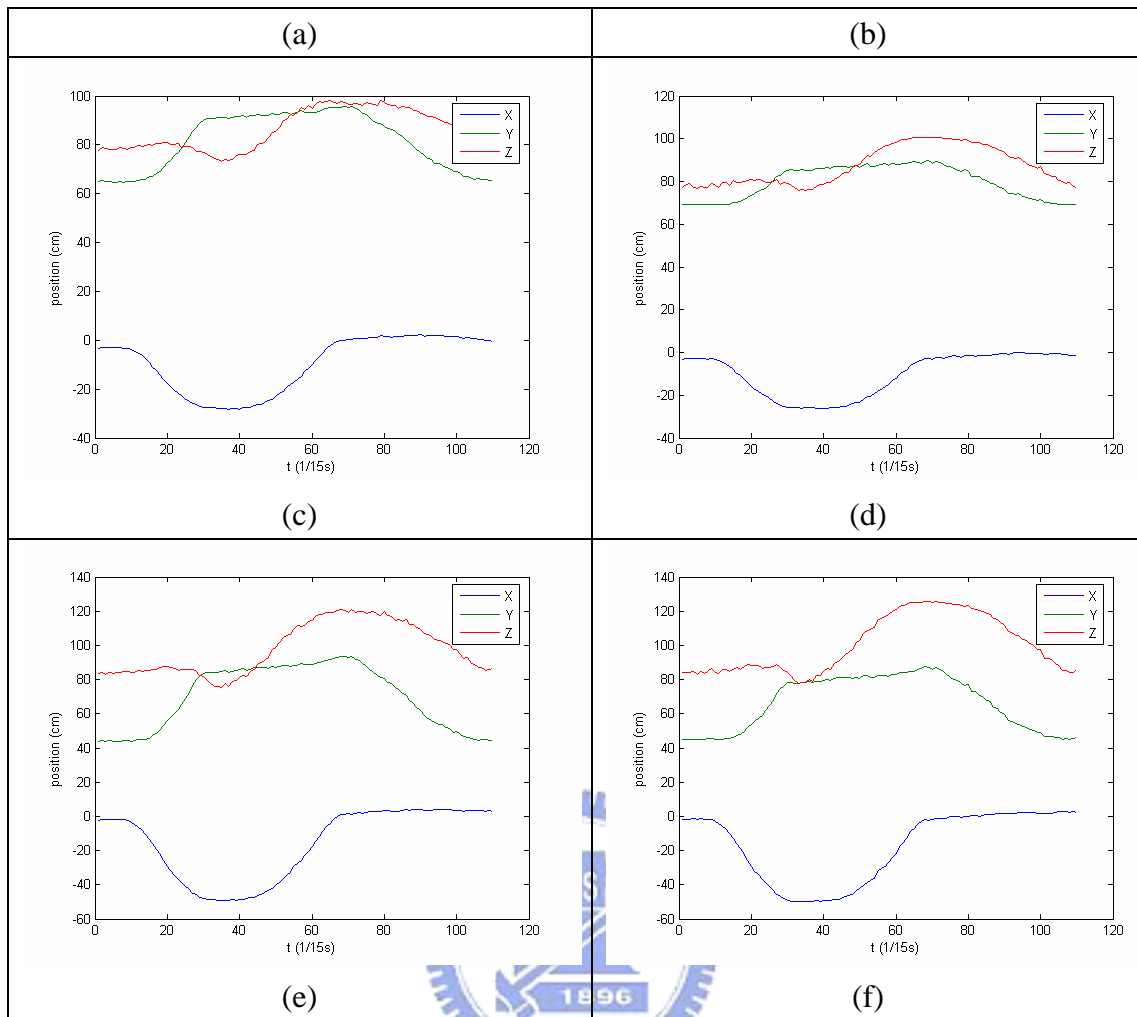


圖 4 - 10 實驗四：各關節點位座標

(a)(b)為原始及修正的肩膀點位序列，(c)(d)為原始及修正的手肘點位序列，
(e)(f)為原始及修正的手腕點位序列

再將修正後的點位利用 3.3 中的方法進行切割，得到的平面切割結果（表格 10）顯示此組運動為三個平面運動的組合。動作片段一為側抬運動其運動平面法向量為 $[-0.1251, -0.09, -0.9881]$ ，所有點位到該平面的平均距離為 0.7802，標準差 0.9013(圖 4-11)；動作片段二為向前收攏運動其運動平面法向量為 $[-0.2800, 0.9496, 0.1410]$ ，所有點位到該平面的平均距離為 0.8622，標準差 1.0229（圖 4-12）；動作片段三為放下運動其運動平面法向量為 $[-0.9930, -0.1137, -0.0334]$ ，所有點位到該平面的平均距離為 0.5506，標準差 0.7124（圖 4-13）。

切割出來的三個片段，手肘點位到運動平面的距離都相當的小，顯示手肘真的在該平面上運動。

表格 10 實驗四：平面切割結果

Segments	Start Frame	End Frame
#1	1	32
#2	33	72
#3	73	110

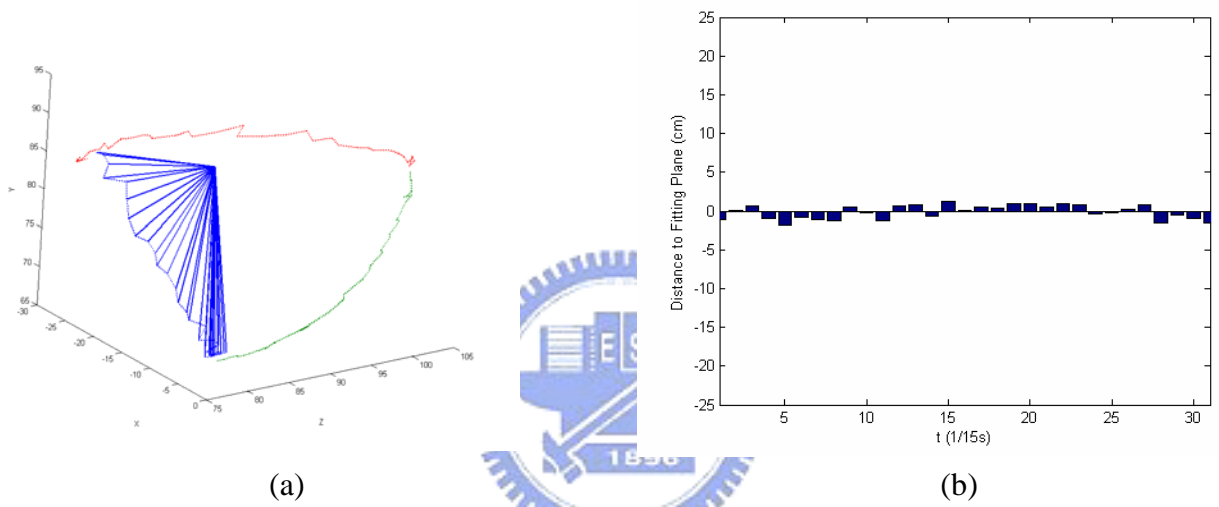


圖 4 - 11 實驗四：動作片段#1
(a)3D 運動軌跡 (b)點位到平面誤差

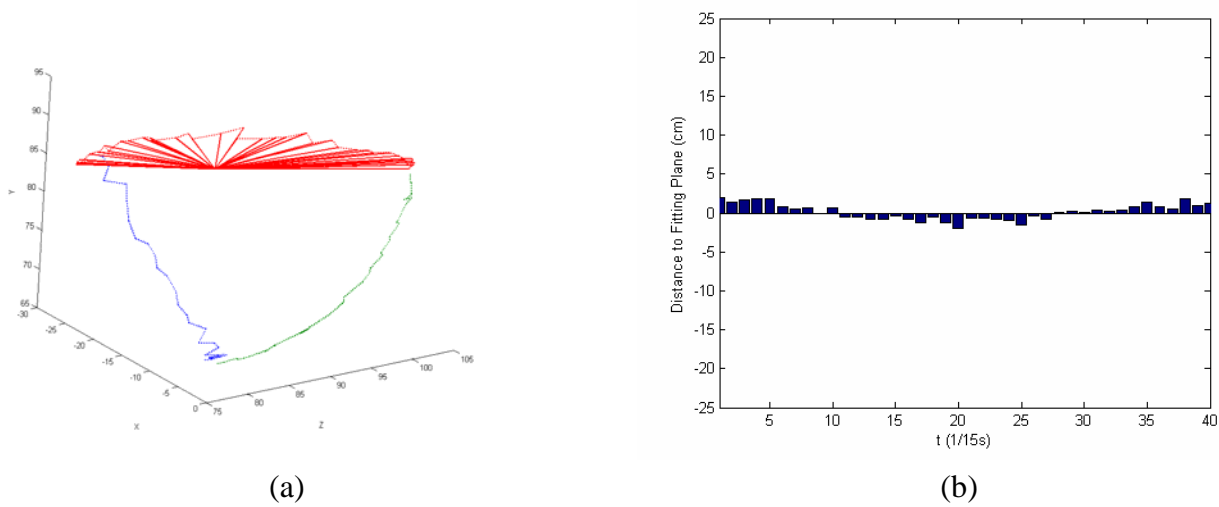
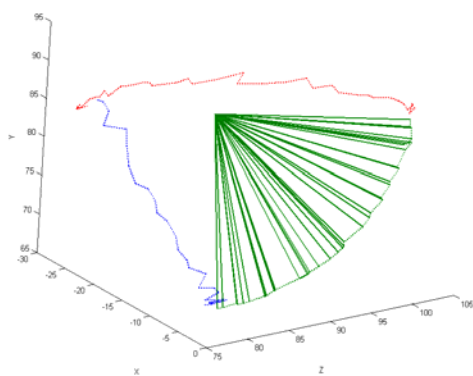
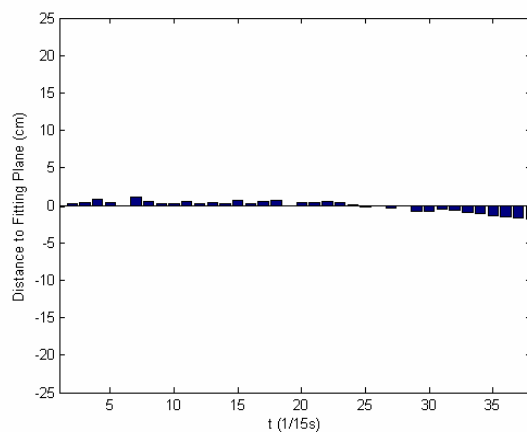


圖 4 - 12 實驗四：動作片段#2
(a)3D 運動軌跡 (b)點位到平面誤差



(a)











(b)

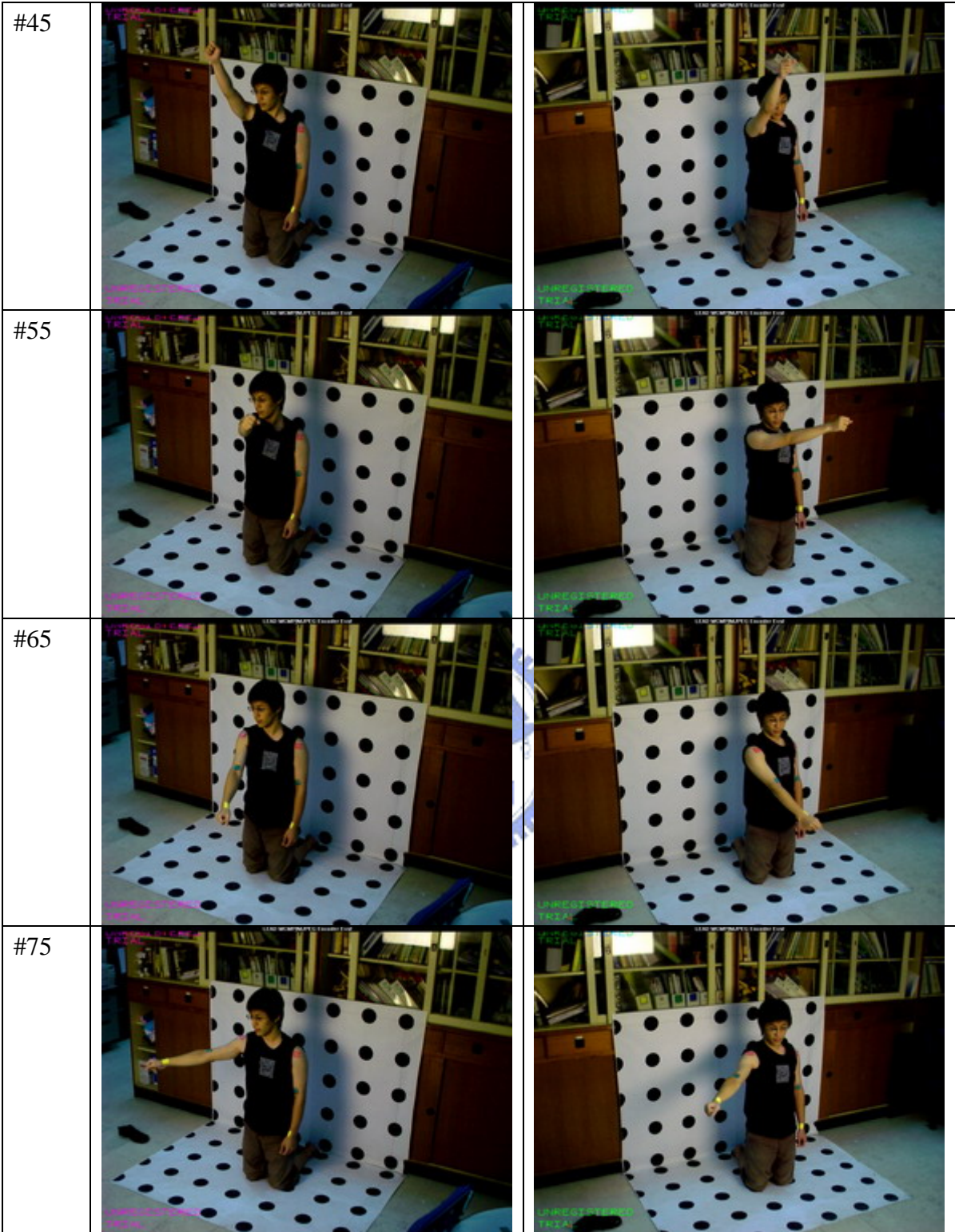
圖 4 - 13 實驗四：動作片段#3
 (a)3D 運動軌跡 (b)點位到平面誤差

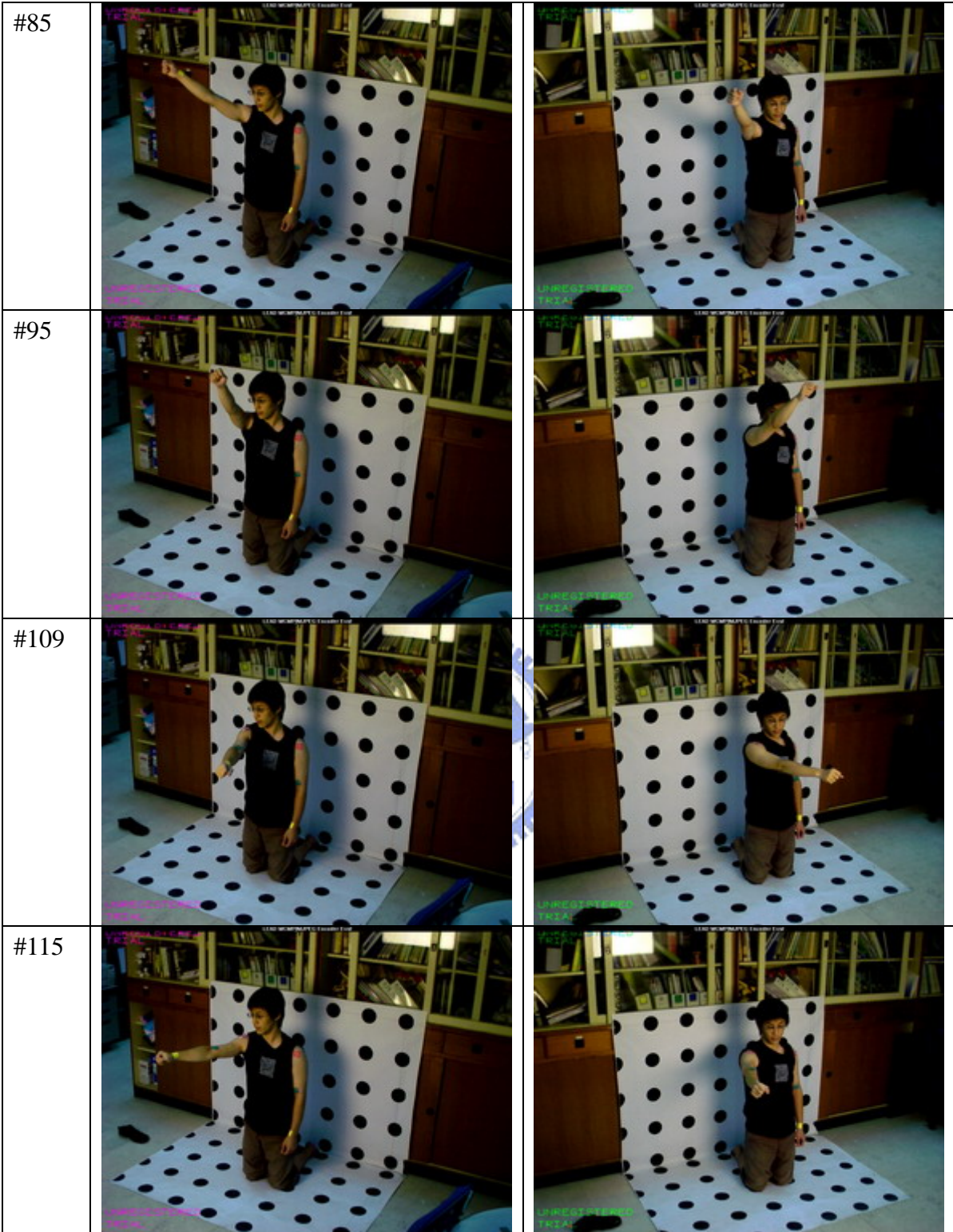


實驗五：

平揮、畫圓、平揮混合運動，共 125 個影格，影格#1~25 為平揮運動，影格#26~108 為畫圓運動，影格#109~125 為平揮運動，運動影像擷取如（圖 4-14）：

Frame	Left-View	Right-View
#1		
#15		
#25		
#35		





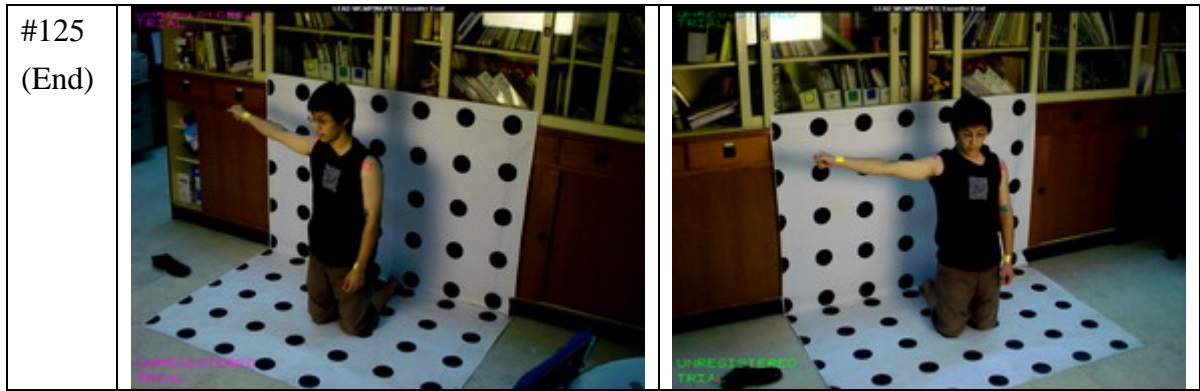
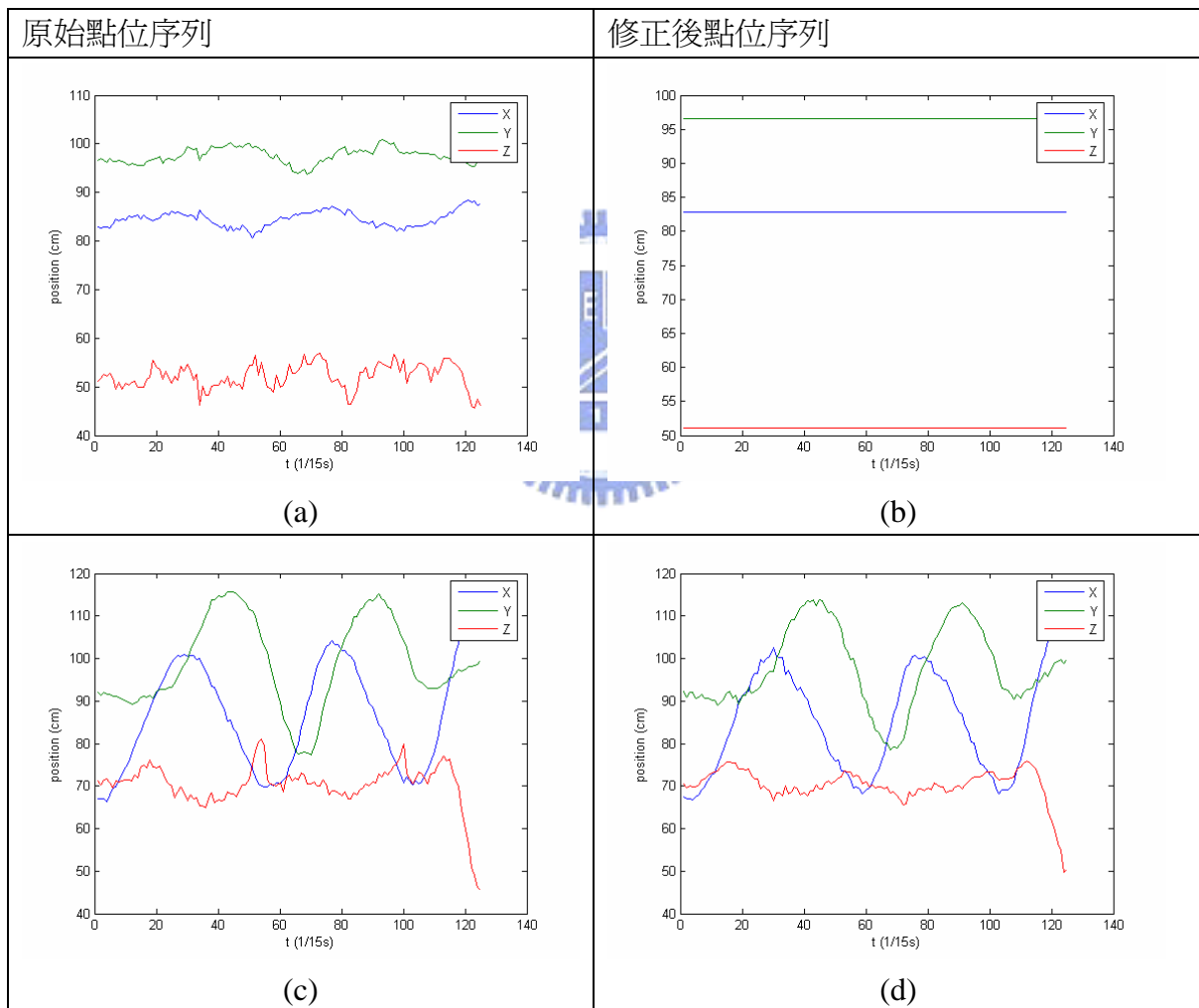


圖 4-14 實驗五：動作擷圖

同樣，利用 Motion Capture 取得表示運動的點位集合，再依 2.2，2.3 中的方法修正。(圖 4-15)



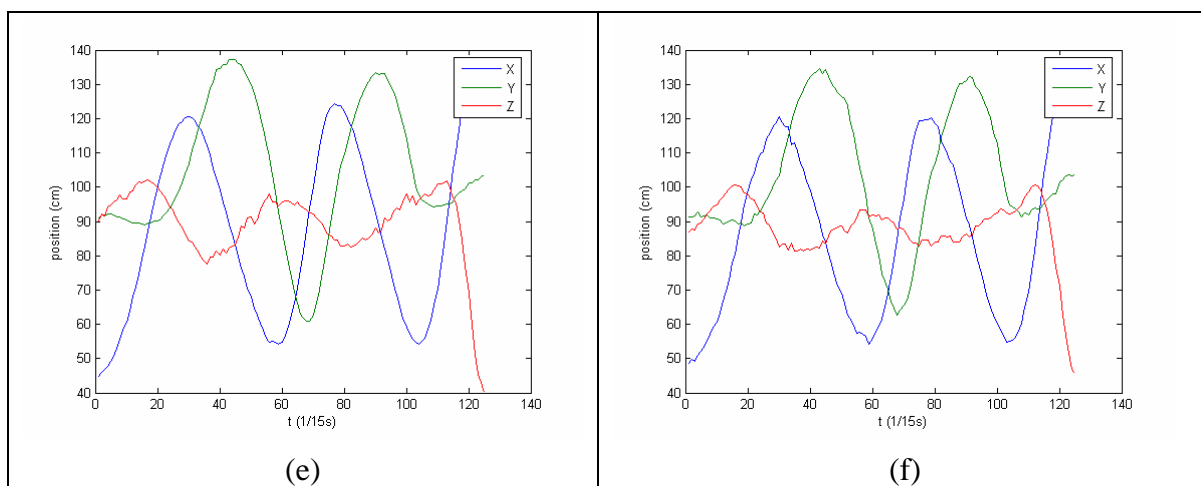


圖 4-15 實驗五：各關節點位座標

(a)(b)為原始及修正的肩膀點位序列，(c)(d)為原始及修正的手肘點位序列，

(e)(f)為原始及修正的手腕點位序列

再將修正後的點位利用 3.3 中的方法進行切割，得到的平面切割結果（表格 11）顯示此組運動為 8 個平面運動的組合。但是發現 Segment #2~7 的影格數都偏少，此時依 3.3 中的平面合併演算法對此 8 個平面進行測試（表格 12）。

表格 11 實驗五：平面切割結果

Segments	Start Frame	End Frame
#1	1	25
#2	31	41
#3	47	56
#4	62	67
#5	73	83
#6	89	93
#7	99	103
#8	109	125

表格 12 實驗五：平面合併過程之奇異值

Merge Segment #1~2	λ_1	λ_2	λ_3
eigenvalue	77.954	36.9639	9.4406
比例	1.0000	0.4742	0.1211
Merge Segment #2~3	λ_1	λ_2	λ_3
eigenvalue	50.2002	20.581	2.9786

比例	1.0000	0.41	0.0593
Merge Segment #2~4	λ_1	λ_2	λ_3
eigenvalue	67.2451	43.6419	4.9604
比例	1.0000	0.649	0.0738
Merge Segment #2~5	λ_1	λ_2	λ_3
eigenvalue	72.7189	61.8164	6.4631
比例	1.0000	0.8501	0.0889
Merge Segment #2~6	λ_1	λ_2	λ_3
eigenvalue	73.5075	66.5219	6.5885
比例	1.0000	0.905	0.0896
Merge Segment #2~7	λ_1	λ_2	λ_3
eigenvalue	80.7864	67.9598	7.18
比例	1.0000	0.8412	0.0889
Merge Segment #2~8	λ_1	λ_2	λ_3
eigenvalue	102.5122	73.0391	29.8636
比例	1.0000	0.7125	0.2913

發現 Segment#2~7 在合併的過程中 λ_2 漸漸與 λ_1 與相同，且而 λ_3 始終維持在一個很小的量，因此可以將 Segment#2~7 進行合併，所以最後得到三個動作片段（表格 13）。

表格 13 實驗五：平面合併結果

Segments	Start Frame	End Frame
#1	1	25
#2	31	103
#3	109	125

最後切分出來的動作片段一為側抬運動其運動平面法向量為[0.0602, -0.9139, -0.4015]，所有點位到該平面的平均距離為 1.0093，標準差 0.8907（圖 4-16）；動作片段二為向前收攏運動其運動平面法向量為[-0.1365, 0.0444, -0.9896]，所有點位到該平面的平均距離為 0.9995，標準差 1.0566（圖 4-17）；

動作片段三為放下運動其運動平面法向量為 $[0.1279, -0.9772, -0.1696]$ ，所有點位到該平面的平均距離為 0.5309，標準差 0.6949（圖 4-18）。切割出來的三個片段，手肘點位到運動平面的距離都相當的小，顯示手肘真的在該平面上運動。

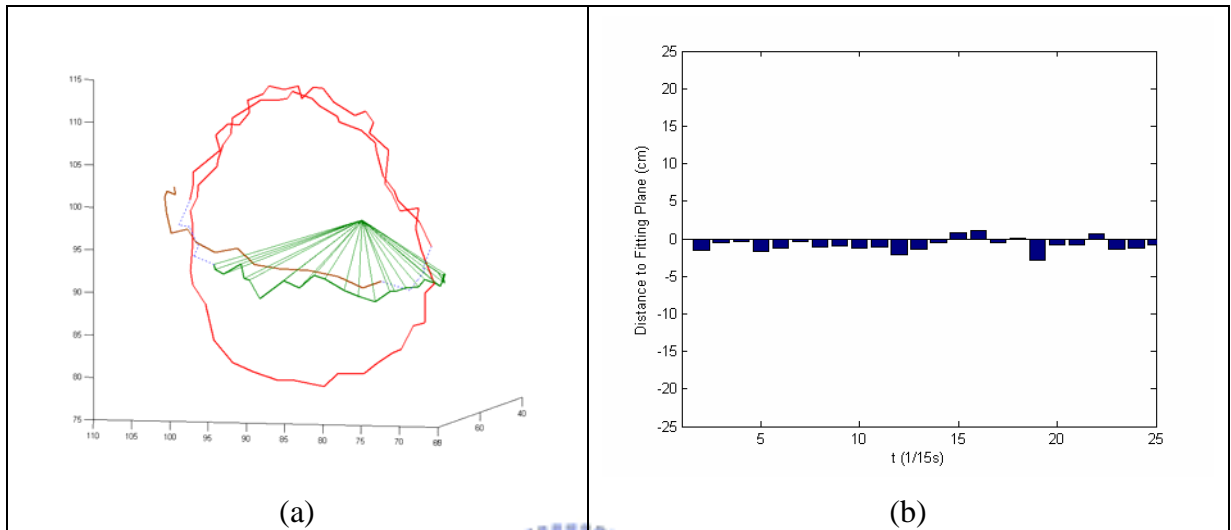


圖 4 - 16 實驗五：動作片段#1
(a)3D 運動軌跡 (b)點位到平面誤差

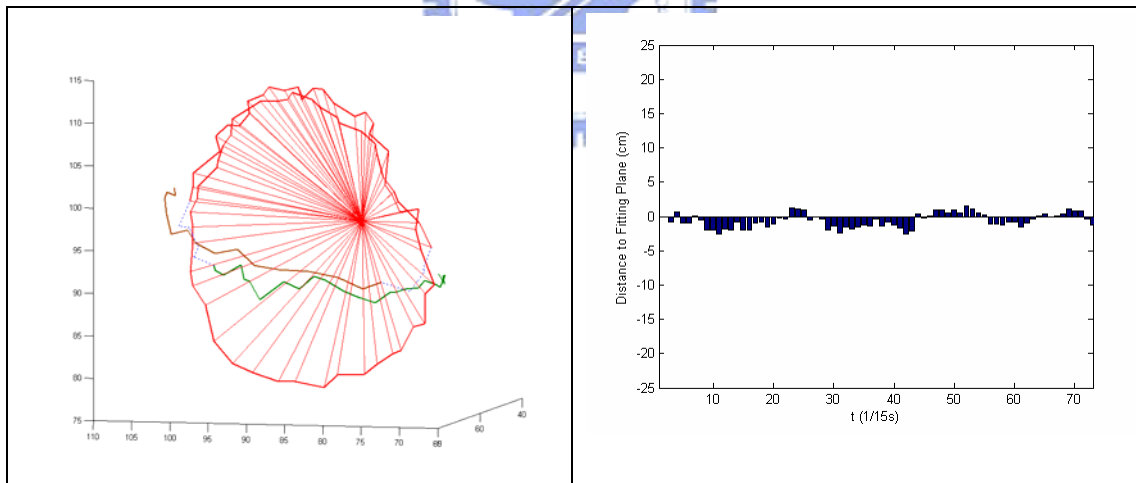


圖 4 - 17 實驗五：動作片段#2
(a)3D 運動軌跡 (b)點位到平面誤差

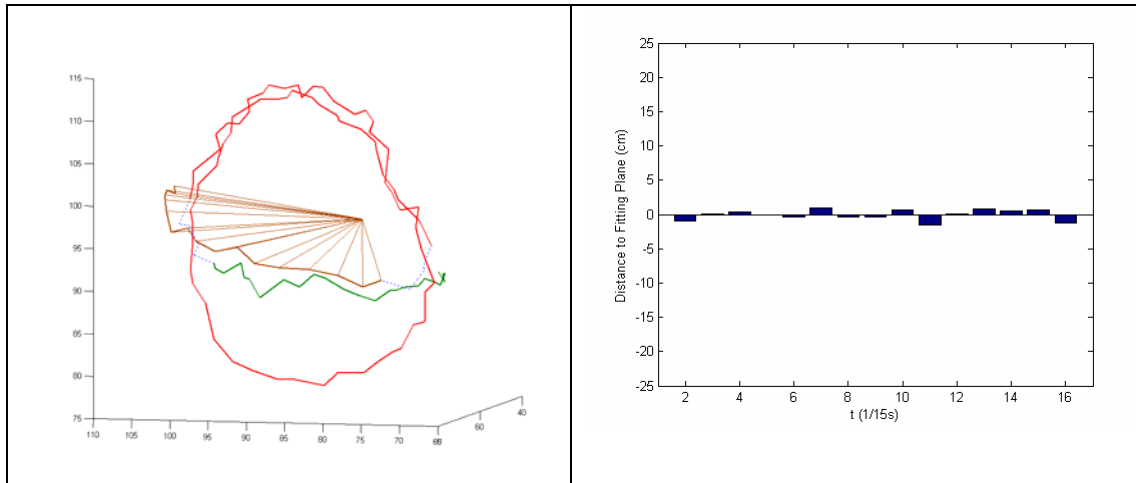


圖 4 - 18 實驗五：動作片段#2
(a)3D 運動軌跡 (b)點位到平面誤差

實驗四與實驗五的結果與人工切割的結果相同。證明本中提出的平面切割演算法能切割出符合人體感官的動作片段。



4.4 比較有無平面資訊輔助的切割結果

最後，要比較本文中提出的方法是否在切割結果上有所幫助。

實驗六：

實驗動作數據同實驗四。

在不用平面資訊的情況下，手部運動過程中，需要考慮 3DOF 肩膀關節在三個軸向上的角度變化量及 1DOF 手肘關節在一個軸向上的角度變化量

(圖 4-19)，表示成 AR Model 的狀態向量即為

$$\begin{bmatrix} \theta_{\text{shoulder_x-dof}} \\ \theta_{\text{shoulder_y-dof}} \\ \theta_{\text{shoulder_z-dof}} \\ \theta_{\text{elbow}} \end{bmatrix}。$$

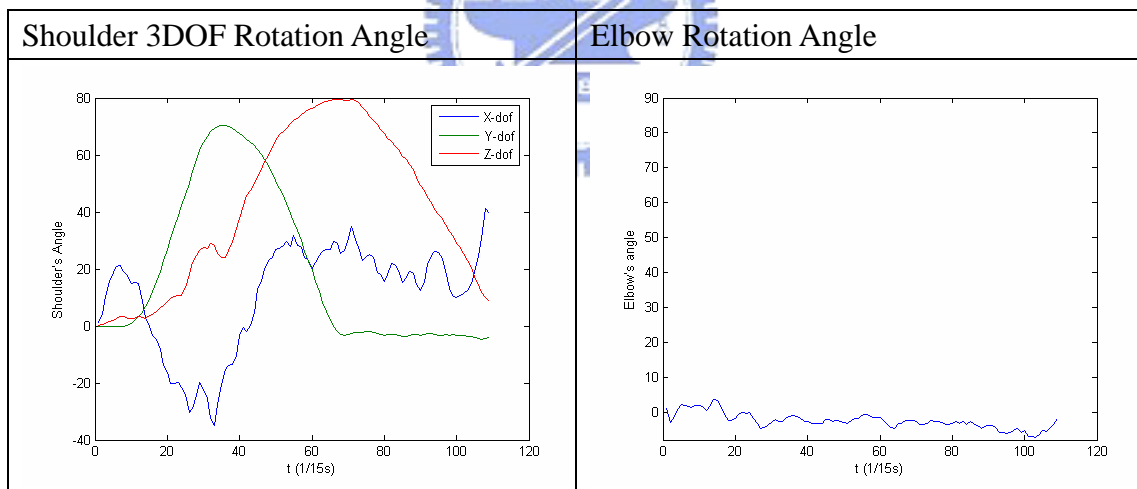


圖 4 - 19 實驗六：關節角度

再依照 3.4 中的 AR Model Segmentation 方法直接切割成數個動作片段，其結果如（表格 14）：

表格 14 實驗六：AR 切割結果

Segments	Start Frame	End Frame
#1	1	12
#2	12	17
#3	17	27
#4	27	34
#5	34	44
#6	52	60
#7	68	87
#8	92	104
#9	104	109

檢視切割結果與實際的運動影像，發現 AR Model Segmentation 會有過度切割 (Over-Segmentation) 的情況，切割出來的動作片段並不具感官上的意義，只是該片段中的角度變化符合某一組動態模型，而 AR 模型在處理有雜訊的訊號常常會因為雜訊的干擾而導致有過度切割的情況。

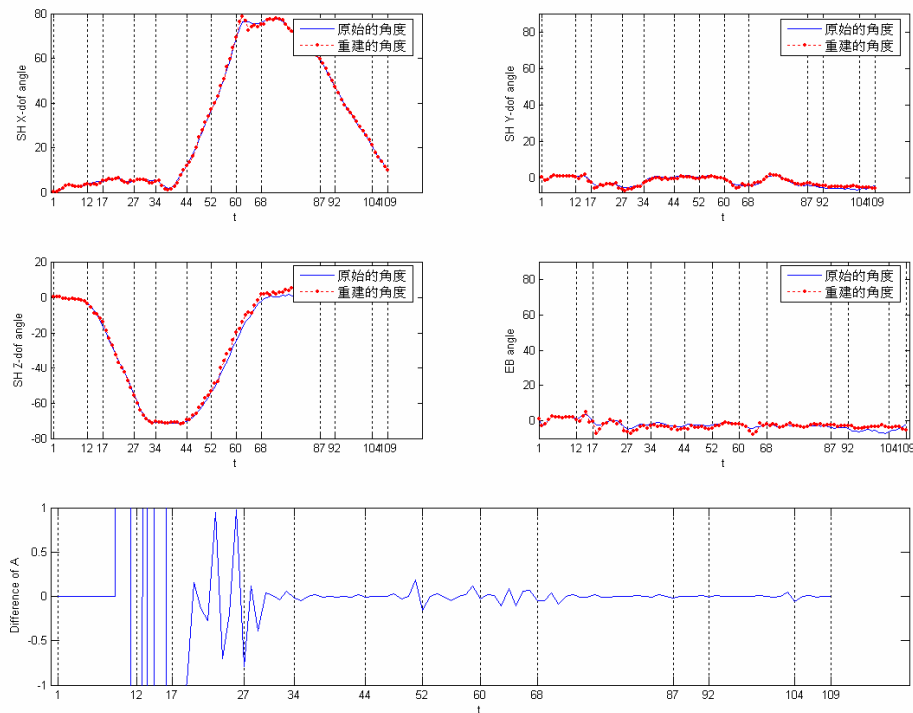


圖 4-20 實驗六：AR 切割結果

比較實驗四的切割結果，該實驗最後將動作切割成三個片段，而在本次實驗中卻過度切割成 9 個片段，因此，透過平面的資訊可以有效的降低誤差對 AR Model 的影響，並且得到的動作片段比起 AR Model 的片段更具語義。



五 結論與未來發展方向

5.1 結論

本研究依照關節旋轉軸與肢節垂直與否，將肢節在空間中的運動軌跡分為共面與共圓錐兩種基本運動模式，而手部運動即可視為此兩種運動模式搭配不同的旋轉角度的組合情形。藉由將肢節運動做 **plane fitting** 出來的平面法向量為旋轉軸，觀察旋轉軸是否發生改變，將手部運動切割成數個獨立的動作片段。接著，再對動作片段中的角度變化，以二階的 **AR** 模型在參數變化劇烈的地方來做進一步的切割。最後，即可將動作切割成數個“同旋轉軸與同角度變化模式”的動作片段。



就實驗結果來看，運動軌跡上點位維度退化成二維這樣的性質，可以從 4.1 實驗中得到驗證。而 **AR Model** 來描述角度變化及切割的能力，可從 4.2 實驗中得到目標中的切割結果。將手部運動過程視為由數個不同的共面運動的組合的假設，可以在 4.3 實驗中證明此假設是成立的，同時 4.3 實驗中而我們也成功的將動作切割成數個平面運動，而切割出來的結果與人工切割的結果相符。在 4.4 實驗中，可以明顯觀察到 **AR Model** 有過度切割的問題，而以平面先進行切割則可降低過度切割的問題。

此外，由於捕捉系統的誤差。不精確的點位數據不止造成平面擬合上的困難，也讓旋轉角度資料的錯誤進而導致 **AR Model** 有過度切割的情況。而誤差的大小也會影響演算法中容許值的大小，過大的容許值往往會造成切割時間的延遲。因此，切割的結果與資料的精確度有密切的關聯性。

5.2 未來工作

受限於時間及資料取得不便，本文中實驗雖只對手部運動來進行分析與實驗，但手部與腳部的構造相同，腳部的運動也是由腕部、膝蓋、腳踝的轉動所組成，因此本文中提出的方法同樣可以用於腳部運動切割。

未來，我們將要對全身的運動進行分析，將全身分成五個部位（身軀、雙手、雙腳），分別對各部位的運動依照本文中的方法，去找出動作片段，再利用 **clustering** 的計巧將相似的動作片段叢集成基礎動作單元集合，然後，即可依五個部位的動作組合定義出更高層次的基礎單元。



參考文獻

- [1] G. Ahanger and T. Little, "A Survey of Technologies for Parsing and Indexing Digital Video," *J. Visual Comm. and Image Representation*, vol. 7, no. 1, pp. 28-43, 1996.
- [2] H. Wactlar, T. Kanade, M. Smith, and S. Stevens, "Intelligent Access to Digital Video: The Imformedia Project," *IEEE Computer*, vol. 29, no. 5, pp. 46-52, 1996.
- [3] T. Yen and R. Radwin, "A Video-Based System for Acquiring Biomechanical Data Synchronized with Arbitrary Events and Activities," *IEEE Trans. Biomedical Eng.*, vol. 42, no. 9, pp. 944-948, 1995.
- [4] *HumanMotor Behavior*. J.A.S. Kelso ed., Lawrence Erlbaum Associates, 1982.
- [5] F.A. Mussa-Ivaldi and E. Bizzi, "Learning Newtonian Mechanics," *Self-organization, Computational Maps, and Motor Control*, Morasso and Sanguineti, eds., pp. 491-501, Elsevier, 1997.
- [6] C. Bregler, "Learning and Recognizing Human Dynamics in Video Sequences," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 1997.
- [7] A.J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement Imitation with Nonlinear Dynamical Systems in Humanoid Robots," *Proc. IEEE Int'l Conf. Robotics and Automation*, 2002.
- [8] V. Pavlovic and J. Rehg, "Impact of Dynamic Model Learning on Classification of Human Motion," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 788-795, 2000.
- [9] D. Del Vecchio, R. Murray, and P. Perona, "Primitives for Human Motion: A Dynamical Approach," *15th IFAC World Congress on Automatic Control*, 2002.
- [10] N. Howe, M. Leventon, and W. Freeman, "Bayesian Recognition of 3D Human Motion from Single-Camera Video," *Proc. Neural Information Processing Systems Conf. (NIPS '99)*, 1999.
- [11] Odest Chadwicke Jenkins and Maja J Matatic, "Automated Derivation of Behavior Vocabularies for Autonomous Humanoid Motion", *Proc. Autonomous Agents and Multi Agent Systems Conf.*, 2003
- [12] Le Lu, Gregory D. Hager, and Laurent Younes, "A Three Tiered Approach for Articulated Object Action Modeling and Recognition", *Neural Information Processing Systems Conf.*, 2004
- [13] Han-Ming Wu and Henry Horng-Shing Lu, "Supervised Motion Segmentation by Spatial-Frequential Analysis and Dynamic Sliced Inverse Regression", *National Chiao-Tung University, Taiwan*.

- [14] C. M. Lu and N. J. Ferrier, "Automated Analysis of Repetitive Joint Motion", IEEE Trans. Information Technology in Biomedical System, 2003.
- [15] C. M. Lu and N. J. Ferrier, "Repetitive Motion Analysis : Segmentation and Event Classification", IEEE Trans. Pattern Anal. Machine Intell., Feb. 2004.
- [16] C. M. Lu, H. Z. Liu, and N. J. Ferrier, "Multidimensional motion segmentation and identification," Proc. IEEE Conf. Computer Vision and Pattern Recog., pp. 629–636, 2000.
- [17] A. Blake and M. Isard, *Active Contours*, New York: Springer-Verlag, 1998.
- [18] L. Ljung and T. Soderstrom. *Theory and Practice of Recursive Identification*, The MIT Press, 1983.
- [19] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1991.
- [20] 陳郁中, 藉由平面投射轉換執行影像嵌合及物體重建, 國立交通大學, 碩士論文, 民國 89 年
- [21] 黃聿文, 三維空間多攝影機視覺監控系統, 國立交通大學, 碩士論文, 民國 91 年
- [22] 吳政峰, 人體肢節三維動作捕捉之電腦視覺技術, 國立交通大學, 碩士論文, 民國 92 年
- [23] 蘇詠迪, 利用逆向動力學修正立體視覺人體動作原始資料, 國立交通大學, 碩士論文, 民國 93 年

