

行政院國家科學委員會專題研究計畫 成果報告

追蹤人體單純運動之多相機模式比對視覺系統

計畫類別：個別型計畫

計畫編號：NSC92-2213-E-009-081-

執行期間：92 年 08 月 01 日至 93 年 07 月 31 日

執行單位：國立交通大學資訊工程研究所

計畫主持人：陳稔

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 93 年 11 月 3 日

※※※※※※※※※※※※※※※※※※※※※※※※※※
※
※追蹤人體單純運動之多相機模式
※比對視覺系統
※
※※※※※※※※※※※※※※※※※※※※※※※※※※

執行期間： 92 年 8 月 1 日至 93 年 7 月 31

計畫主持人：陳 稔 國立交通大學資工系教授

- ☐赴國外出差或研習心得報告一份
- ☐赴大陸地區出差或研習心得報告一份
- ☐出席國際學術會議心得報告及發表之論文各一份
- ☐國際合作研究計畫國外研究報告書一份

執行單位：交通大學資訊工程學系

中 華 民 國 九 十 三 年 十 月 三 十 日

追蹤人體單純運動之多相機模式比對視覺系統

Model-Based Human Motion Tracking Using a Multi-Camera Vision System

計畫編號：NSC-92-2213-E-009-081

執行期限：92 年 8 月 1 日至 93 年 7 月 31 日

主持人：陳稔 國立交通大學資工系教授

E-mail: zchen@csie.nctu.edu.tw Tel: 03-5731875

1. 摘要

本期計畫利用前期計畫所建立的多相機視覺系統進行人體運動之追蹤，三維運動參數之計算。人體運動分析所面臨之問題包括 (1)人體肢節之偵查 (2)人體肢節的區分標示 (3)肢節基準中心線之計算 (4)人體肢節三維模型重建 (5)人體肢節物理條件之檢查 (6)肢節點及肢節旋轉角度計算。其中所遭遇之挑戰性問題主要有 (1)肢節相互遮蔽 (2)肢節伸縮時之變形 (3)人體輪廓域之完整擷取 (4)肢節之準確三維模型建立。我們採用多相機視覺系統以盡量排除肢節遮蔽問題。藉由多相機的多個影像，我們可以充分地偵查人體肢節各部位，並確立各部位的肢節對應名稱。其次，我們借助事先建立好的三維人體幾何尺寸，試圖做各部位肢節之三維姿態估測。我們先利用二個以上相機所記錄的肢節基準中心線粗略估算肢節的三維姿態參數，再透過逆向運動學方法逐步做修正。

關鍵詞：人體運動、人體肢節模型、動作捕捉、逆向動力學、肢節基準中心線、肢節標示、肢節旋轉

Abstract

This research project utilizes the multi-camera vision system developed in the previous project to track the human motion and to compute the 3D motion parameters. Human motion analysis tackles a series of tasks including (1) human body segmentation, (2) human body labeling, (3) human body spine finding, (4) human skeleton model reconstruction, (5) physical body constraint checking, (6) 3D computation of joints and joint angles. The challenging problems to be faced are related to (1) self occlusion between human bodies, (2) 3D human body deformation, (3) segmentation integrity of the human body silhouette detection, (4) accurate skeleton reconstruction. In this project we shall analyze the human motion without part deformation.

In the current project we propose to use the multi-camera or the multi-view approach to alleviate the self-occlusion problem. Through the availability of multiple images at the same time, we hope to detect all the human body parts and give a correct labeling result. After the part segmentation, we use the 3D human skeleton model gathered beforehand to compute the position and orientation of each joint. First of all, we assume the ribbon spine of the part silhouette can be found in at least two views, so we can roughly estimate the 3D location of the part spine. Then we use the inverse kinematics method to tune the initial skeleton model until its end-effectors agree to the observed positions in the camera views.

Keywords: Human Motion, Human Skeleton Model, Motion Capture, Inverse Kinematics, Body Spine, Body Part Labeling, Joint Rotation.

2. 緣由與目的

近年來人體動作捕捉技術(Motion capture)已被廣泛的使用[1-6]。遊戲業者與或電影業者為了使動畫的效果更加逼真細膩，紛紛使用人體動作捕捉的技術來製作動畫，而在醫學復健動作分析、運動員動作訓練上也都應用到人體動作捕捉技術。因此，人體動作捕捉在現代科技研發上扮演了非常重要的角色。然而，現今的人體動作捕捉技術都需要使用感應器、感應標誌，造成表演者表演時一些不必要的麻煩。

本期計畫利用前一期計畫所建立的多相機視覺系統，進行人體運動的分析。在不需使用感應標誌的前提下，由多相機所拍攝之影像序列推算人體各肢節部位的三維運動參數。所得到的肢節運動參數將可交由三維繪圖軟體如Maya等作後製美工處理。

3. 研究方法

本計畫先使用二台攝影機自不同方向拍攝表

演者動作之影像序列。攝影機架設好後位置與焦距皆固定不變。將得到的兩組影像序列分析後即可得到表演者各肢節部位的三維運動參數。採用步驟如下：

- 攝影機校正：用以建立表演者所在的世界座標系(World coordinate system)與影像座標系(Image coordinate system)之間的對應關係。
- 人體肢節影像擷取：將影像中的表演者與背景分離，得到表演者肢節的二維輪廓影像(2D contour)。
- 肢節基準中心線的計算與標示：由肢節的輪廓影像，訂出各肢節的中心線，並依此推算出人體骨架模型在影像中的投影位置與向量。
- 重建三維骨架模型：將兩台攝影機得到的二維骨架模型，推算出三維肢節方向，再套用正向動力學(Forward kinematics)的方法，訂定出各關節點(Joint)、末端肢節點(End-effector)的區域座標。
- 使用逆向動力學修正骨架模型：前步驟所建立的三維骨架模型，由於在配合表演者肢節的長度之初估後，會導致末端肢節點偏離影像中所看到的實際位置。我們將採用逆向動力學的方式，修正此一誤差，並確保各關節的活動範圍符合物理上的限制。
- 人體運動之重現：將計算所得的骨架模型，以 Stick figure 方式重現出來。

以下分各小節對上述步驟做詳細的描述。

3.1 相機校正

由於本計畫所使用的相機於架設好之後位置與焦距皆不會改變，故可在實際拍攝表演者之前，使用校正物(Calibration object)進行事先校正。

利用 Homogeneous coordinate system，可將世界座標系 WCS 中的 3D 座標點 $[X_w \ Y_w \ Z_w]^T$ 和對應的影像座標系統 ICS 上的 2D 座標點 $[u_i \ v_i]^T$ 之間的關係式表為：

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

其中 w 為一應變數， H 則是所要求的相機投影矩陣 (Projection matrix) 或稱為相機矩陣 (camera matrix)。

在矩陣 H 中共有 12 個未知參數，以傳統的方法，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 。使用的對應點數越多，得到的矩陣 H 精確度越高。

我們所使用的校正物如(圖 1)所示，首先在兩台攝影機皆可拍攝到的範圍內，定出世界座標系統 WCS 的原點 $[0_w \ 0_w \ 0_w]^T$ 當做基準點，接著在不同的平面上貼上特徵點，測量其在世界座標系統中的三維空間座標值，並於攝影機拍攝的影像中找到各特徵點的對應投影點，套用上述方法即可解得兩台攝影機各自的投影矩陣。



(圖 1) 校正物在左右兩台攝影機中的成像

3.2 人體肢節影像擷取

攝影機攝得表演者的影像序列後，必須先將前景的表演者與背景分離，才能進行後續的肢節中心線計算。

在[7]中介紹了數種將前景與背景分離的方法，其中最簡單的就是將含有前景表演者的影像，與事前拍攝只含有背景的影像(見(圖 2))相減，再將差值超過臨界值的像素標示為前景像素。



(圖 2) 左：背景影像、右：含前景的完整影像

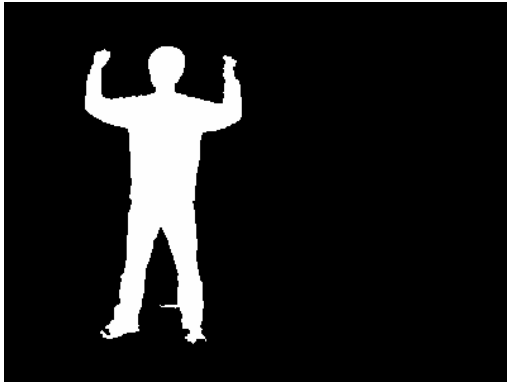
為了處理陰影及光照變化所造成的影響，我們先將彩色影像中各像素的 R、G、B 值轉換為 H、S、I 格式，其中 H 代表色調(Hue)、S 代表飽和度(Saturation)、I 則是亮度(Intensity) [8]。僅取用 H 值來做相減的動作，便可得到不包含陰影的前景像素，見(圖 3)。

經色調比對所得的前景，難免會包含一些雜

訊，由於所要偵測的部份為人體，而雜訊所造成的錯誤前景與實際的人體區塊相較之下都很小，因此只要將(圖三)中的前景找出最大的一個相鄰區塊(Connected component)，該區塊即為我們所要的人體區塊。最後再將人體區塊使用 erosion 與 dilation 將區塊中的小破洞補滿即可得到完整的肢節輪廓影像，如(圖 4)



(圖 3) 色調(Hue)值相減得到的前景像素



(圖 4) 去雜訊並將破洞補滿後的完整肢節輪廓

3.3 肢節基準中心線的計算與標示

在不使用感應標誌的情況下進行動作捕捉，要取得身體各關節的正確位置是件相當困難的事，從影片中擷取下來的影像，除了輪廓線外並沒其它明顯的資訊可供參考，若要以人工的方式進行標定，在如此大量的資料下並不是個可行的方案，因此，如何從大量的影像資料中自動的、有效率的取出末端肢節點(End-effectors)、及各關節點(Joints)的位置則成為一個我們需要克服的難題。

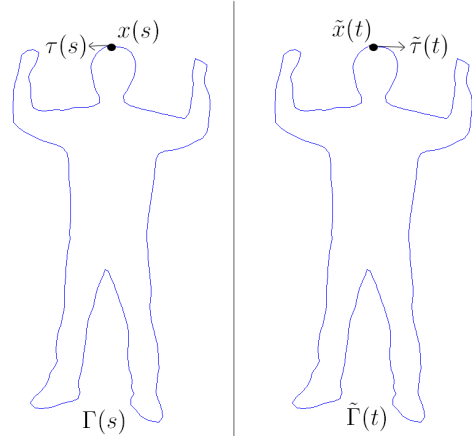
3.3.1 尋找末端肢節點與肢節對稱軸

經由觀察，我們發現人體的各肢節(Body segment)其形狀皆具有對稱性，這是一個相當好的特性，若我們可以找出該肢節的對稱軸線，則該軸線與輪廓線的交點可以視為該肢節的End-effector，再對軸線做Least square 取得最適合的直線來逼近它，則該直線的向量即可當成該

肢節的方向向量，參考[15]。以下是我們的做法：

Step.1: Contour Parameterization

首先，我們對影像做 contour tracing 來取得人體外觀的輪廓線點座標，再對取出來的輪廓線依順時鐘方向及逆時鐘方向各做一次參數化的動作，如此可取得 $\Gamma(s)$ 和 $\tilde{\Gamma}(t)$ ，如(圖 5)。

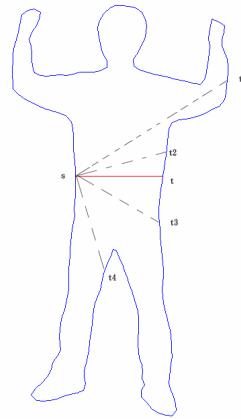


(圖 5) The contour $\Gamma(s)$ and its mirror

version $\tilde{\Gamma}(t)$. $\tau(s)$ and $\tilde{\tau}(t)$ are unit tangents.

Step.2: Matching

接下來沿著參數化曲線，我們試圖對參數 s, t 定義出一種 mapping，使得曲線上的點 $x(s)$ 和 $\tilde{x}(t)$ ，恰好是落在對稱位置。如(圖 6)。



(圖 6) 尋找 s 與 t 的 mapping 使得 $x(s)$ 和 $\tilde{x}(t)$ 對稱

為了找出這樣的一種配對方法，我們定義了一個 energy function，期望讓符合人體對稱性的 (s, t) 配對能有較小的能量，如此一來，尋找 (s, t) 的配對問題則轉換成尋找最小能量的問題。energy function 的型式如下：

$$\int_0^1 F(x, \tau, \tilde{x}, \tilde{\tau}, s', t') d\sigma$$

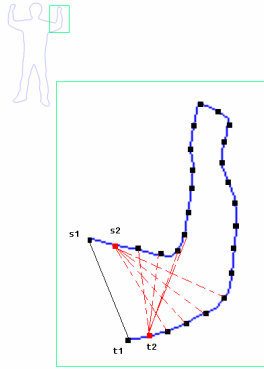
where $x_s, \tilde{x}_t$ are the tangent vectors to

$$\Gamma(s), \tilde{\Gamma}(t) \text{ at } x(s) \text{ and } \tilde{x}(t)$$

$$\tau = x_s / |x_s|, \tilde{\tau} = \tilde{x}_t / |\tilde{x}_t|, s' = ds/d\sigma, t' = dt/d\sigma$$

若在contour上共有N個點，則可能的配對方式即有 N^2 種，如何得到正確的N種配對，則是問題的重點。以下是我們的做法

整個想法在於「假設 (s_1, t_1) 為一組正確的配對時，則下一組可能正確的配對為何？」因為matching必須是 monotone and continuous，所以下一組的配對方式我們只需要考慮 $(s_2, t_1), (s_j, t_2)$, where $\{s_j, t_i\}$ are rest points on the contour，如(圖 7)所示。



(圖 7) monotone and continuous matching

我們利用 Dynamic programming 的方法來求出所有 (s, t) 的配對，使得系統的能量最小，在求出所以有的 (s, t) 配對後，軸線的位置即為

$$\frac{x(s) + \tilde{x}(t)}{2}。$$

3.3.2 標示對應的肢節區段

在影像序列中，我們透過上述的方法可以分別找出各張影像中的人體肢節的中心線，但是上述的方法找出來的對稱軸並沒有固定的次序，數量也不盡相同（因為攝影機角度或是肢節間互相遮蔽），因此我們需要一種 mapping 的方法來決定 $T_0, T_1, T_2, \dots, T_n$ 時間找出來的 Body segments or End-effectors 間的正确對映，如此才能完整，正確的記錄每一個 End-effector 的運動軌跡，以下是我們的做法：

Step.1: SA Tree Representation

在這裡，我們採用[15]中的 SA Tree 表示法來表示我們的人體外觀，因為用 contour 來表達人體的形狀雖然仔細，但是無法提供我們足夠的資訊來區分各肢節的位置，及其對應的關係，使用 SA Tree Representation，我們將可以很方便的進行這樣的

處理。

Step.2: SA Tree Mapping

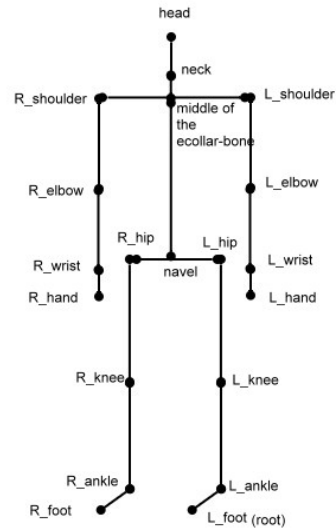
將擷取到的影像透過上述的 SA Tree 的方式編碼後，我們得到一個用來描述人體形狀的 SA Tree，在 SA Tree 中的一個 edge 代表著身體的一個肢節，如此一來，尋找 Segments 的對映則轉換成尋找 SA Tree edge 的對映。

如此只要將不同時間點 $(T_0, T_1, \dots, T_n)$ 的 SA Tree 依[16]之方法進行比對，即可得到在 $T_0 \sim T_n$ 間肢節的運動軌跡。

3.4 重建三維骨架模型

3.4.1 人體骨架模型的定義

理想中的人體模型包含了骨骼、肌肉、皮膚等等，但在動作捕捉系統中，重點在於測得各關節點的位置與方向。人體的骨骼非常複雜，為簡化資料量，本計畫採用(圖 8)所示的簡化版骨架模型 (Skeleton model)。



(圖 8) 本計畫採用的骨架模型

3.4.2 重定肢節點位置

在 3.3 節中我們已求得兩台攝影機影像中，各肢節（雙手、雙腳、頭部）的 End-effectors 之 2D 座標、以及肢節的 2D 向量，但這些資訊通常會存在一些誤差，且對於同一個人，肢節的三維長度是固定不變的，所以必須重定肢節點的位置，使得估算出的肢節長度能符合表演者的骨架模型。

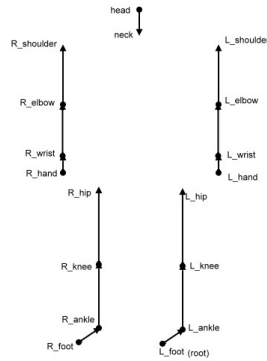
肢節點的位置決定是利用影像上求得的肢節向量，再套上事前量測得到的肢節長度，得出新的肢節點位置。此做法可使肢節的長度符合表演者的骨架模型，並沿用影像中測得的肢節角度。

由於各肢節的 End-effectors 通常沒有被衣

物包覆，在影像上找到的投影點位置會比關節點來得準確。又，假設在左腳著地的情形下，L_foot 必定會座落在三維世界座標系(WCS)中地面的平面上。故我們將L_foot 定義為 root，以此為起點開始重建三維骨架模型。

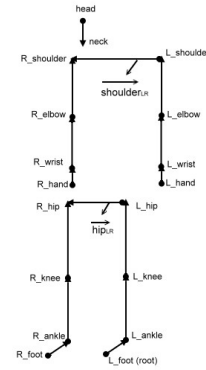
首先，將二張影像座標系(ICS₁和ICS₂)中L_foot (root)的2D位置，利用立體視覺三角定位的方式，反投影可得到其在WCS的3D位置。再將ICS₁與ICS₂中L_foot至L_ankle的2D向量，經由Appendix.A的方法，轉換為WCS中L_foot至L_ankle的3D向量。將此一3D向量套上已知的肢節長度，便可得到L_ankle在WCS中的3D座標。接下來再以L_ankle當端點，求出L_knee的3D位置、以L_knee為端點，求出L_hip的3D位置。

重複同樣步驟，用已知的 End-effectors (L_hand、R_hand、R_foot、head) 為起點，可陸續的求出 wrist、elbow、shoulder、neck 在 WCS 中的3D座標，見(圖9)。



(圖 9) 重定肢節點位置示意圖

然而，這個結果只能讓手、腳、頭各部分的肢節長度固定不變，必需再將各部分整合起來，形成正確的人體模型。在此先整合雙腳和雙手，有了L_hip、R_hip，L_shoulder、R_shoulder 在三維空間中座標，就可以得知它們之間的向量 $\overrightarrow{hip_{LR}}$ 、 $\overrightarrow{shoulder_{LR}}$ ，再套上已知的肢節長度，就可以將雙腳、雙手整合起來，如(圖10)。



(圖 10) 整合雙手與雙腳的肢節

緊接著必須求得身體軀幹向量，才能將雙手和雙腳整合起來，而身體軀幹是由脊椎為骨架，在簡化的人體構造中，脊椎會通過肚臍，而脊椎和鎖骨是垂直的，即是身體軀幹向量會和兩肩所構成的向量垂直。首先求得肚臍的座標，在虛擬的人體骨架模型中，肚臍(navel)的座標為L_hip和R_hip的中心點位置 P_1 ，接下來求通過L_shoulder (P_{LS})、R_shoulder (P_{RS})的直線方程式 L_1 ，利用 L_1 、求得直線方程式 $\overline{P_1P_2}$ ，使得

$$\overline{P_1P_2} \perp L_1, P_2 \in L_1$$

身體軀幹向量和兩肩之間的向量垂直。因為人體是對稱的，所以必需延著 L_2 移動包含L_shoulder的整個區塊和包含R_shoulder的整個區塊，使得

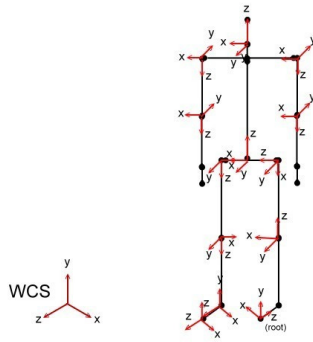
$$\|P'_{LS}P_2\| = \|P_2P'_{RS}\|; P'_{LS}, P'_{RS} \in L_1$$

如此一來便整合了雙手、雙腳和身體軀幹。最後再以 P_2 為端點，求得脖子到頭的向量，套上已知的肢節長度，人體的三維骨架模型就建構完成了。

以此步驟建構出的三維骨架模型，雖能確保肢節長度與表演者相符，但也使除了 root 之外的 End-effectors 的3D位置產生偏移，因此必須透過3.5節介紹的逆向動力學方式，將End-effectors調整到攝影機所觀測到的位置。

3.4.3 求取各肢節點座標系與關節角度

為方便後續計算，我們將為每個肢節點(包含End-effectors與Joints)訂立各自的區域座標系，見(圖11)。



(圖 11) 定義各肢節點的區域座標系

首先將各區域座標系的 Z 軸定義在相鄰的肢節上。對 1-DOF 的關節點而言(e.g. L_knee)，為方便起見，將相鄰的兩段肢節向量的外積定為 X 軸，亦即該關節的轉軸，Y 軸則可由 X、Z 外積得到。對於 2 或 3-DOF 的關節點，X、Y 軸則可任意設定，只要構成直角座標系即可。

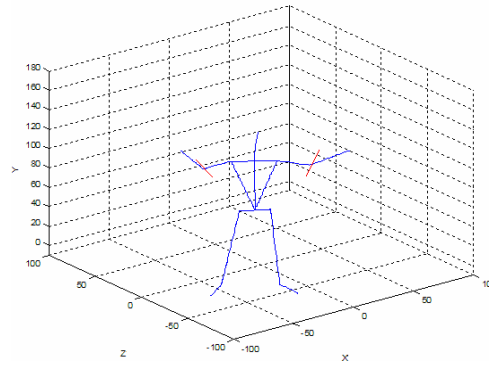
關節的旋轉角度則定義為相鄰兩區域座標系間的旋轉角度。首先將兩座標系轉到同一個座標系統，假設轉到同一個座標系的兩座標系各為 $[OX_1 \ OY_1 \ OZ_1]$ 及 $[OX_2 \ OY_2 \ OZ_2]$ ，可推出：

$$[OX_1 \ OY_1 \ OZ_1] = M[OX_2 \ OY_2 \ OZ_2]$$

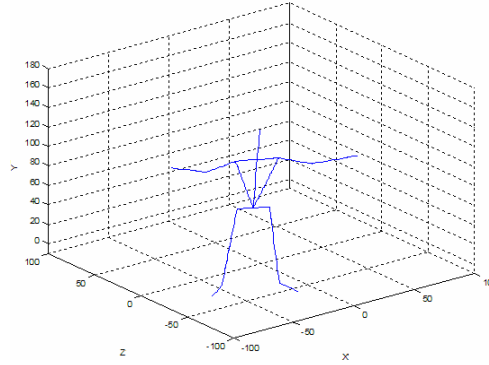
$$M = [OX_1 \ OY_1 \ OZ_1] * [OX_2 \ OY_2 \ OZ_2]^{-1}$$

所求得的 M 稱為旋轉矩陣，只要將 M 分解為 $R_x(\phi_x) * R_y(\phi_y) * R_z(\phi_z)$ 的形式，即可求得關節點對三個轉軸的旋轉角度。

此外，在某些特例情況下，1-DOF 關節點相鄰的二肢節向量趨近平行，便無法以外積定義出關節轉軸，例如手臂伸直時的 L_elbow 或 R_elbow 關節。此時便選用連續影像序列中，時間最相近的一組影像所定義出的關節轉軸作為目前的轉軸。以(圖 12)(圖 13)為例，二圖為時間相鄰的兩組影像重建出的骨架模型，(圖 13)的上臂與下臂近乎平行，故 L_elbow 與 R_elbow 的轉軸需由(圖 12)取得。



(圖 12) 由上臂與下臂向量求得 elbow 之轉軸



(圖 13) 上臂與下臂近乎平行，無法求得 elbow 之轉軸，需由(圖 12)取得

3.5 使用逆向動力學修正骨架模型

前一步驟所建立之三維骨架模型雖確保了肢節長度正確，但也使 End-effectors 的位置產生偏移。因此必須使用逆向動力學(Inverse kinematics)方式，End-effectors 調整到攝影機所觀測到的位置[14]。

End-effector 的位置，在正向動力學(FK)中，是使用者給定關節角度 $\Theta = (\phi_1, \phi_2, \dots, \phi_n)$ 而決定出 End-effector 的位置 X：

$$X = f(\Theta), \quad \Theta = (\phi_1, \phi_2, \dots, \phi_n)$$

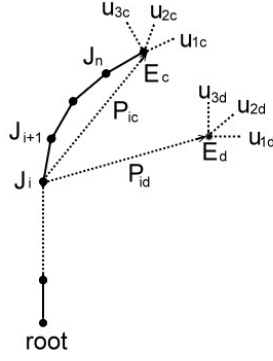
在逆向動力學(IK)中則是反向操作，由使用者給定一個目標位置，藉由這個目標位置和已知的肢節長度來反推決定出關節角度，亦即是求上面式子的反函數：

$$X = f^{-1}(\Theta)$$

在正向動力學(FK)中，只要給各關節角度，則末端肢節點的位置即可確定，但在逆向動力學(IK)並非如此，因為函數 f 並非是線性函數，對於給定的目標位置，所求得的關節角度可能會有多組解。

目前已有不少的逆向動力學(IK)的演算法[9-13]，在這裡我們為了執行速度考量，採用 Cyclic Coordinate Descent (CCD)演算法。簡述如下：

3.5.1 Cyclic Coordinate Descent (CCD)演算法



(圖 14) CCD 演算法示意圖

首先參考(圖 14)定義：

$\overrightarrow{wE_c} = ({}_w x_c, {}_w y_c, {}_w z_c)^T$ 為目前 end_effector 在世界座標系的位置。

$\overrightarrow{wO_c} = [{}_w u_{1c} \quad {}_w u_{2c} \quad {}_w u_{3c}]^T$ 為目前最後一個關節的方向。

$\overrightarrow{wE_d} = ({}_w x_d, {}_w y_d, {}_w z_d)^T$ 為 end_effector 所希望到達的位置。

$\overrightarrow{wO_d} = [{}_w u_{1d} \quad {}_w u_{2d} \quad {}_w u_{3d}]^T$ 為最後一個關節所希望的方向。

$\overrightarrow{wJ_j}$ 是目前所要轉動關節在世界座標系的位置。

則可得

$$\begin{aligned} \overrightarrow{wP_{ic}} &= \overrightarrow{wE_c} - \overrightarrow{wJ_i}, \quad \overrightarrow{wP_{id}} = \overrightarrow{wE_d} - \overrightarrow{wJ_i} \\ \overrightarrow{iP_{ic}} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} \overrightarrow{wP_{ic}}, \quad i=1, 2, \dots, n \\ \overrightarrow{iP_{id}} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} \overrightarrow{wP_{id}}, \quad i=1, 2, \dots, n \\ \overrightarrow{iE_c} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} [\overrightarrow{wE_c} - \overrightarrow{wJ_i}]^T, \quad i=1, 2, \dots, n \\ \overrightarrow{iE_d} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} [\overrightarrow{wE_d} - \overrightarrow{wJ_i}]^T, \quad i=1, 2, \dots, n \\ \overrightarrow{iO_c} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} \overrightarrow{wO_c} = [{}_i u_{1c} \quad {}_i u_{2c} \quad {}_i u_{3c}]^T \\ &\quad, \quad i=1, 2, \dots, n \\ \overrightarrow{iO_d} &= \begin{bmatrix} \overrightarrow{w x_i} & \overrightarrow{w y_i} & \overrightarrow{w z_i} \end{bmatrix}^{-1} \overrightarrow{wO_d} = [{}_i u_{1d} \quad {}_i u_{2d} \quad {}_i u_{3d}]^T \\ &\quad, \quad i=1, 2, \dots, n \end{aligned}$$

其次，定義 penalty function：

$$\begin{aligned} {}_i F_p(q) &= \left\| \overrightarrow{iE_d} - \overrightarrow{iE_c} \right\|^2, \quad i=1, 2, \dots, n \\ {}_i F_o(q) &= \sum_{k=1}^3 \left(\left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{i u_{kc}} \right) - 1 \right)^2, \quad i=1, 2, \dots, n \end{aligned}$$

q 為參數(joint variables)。

目的：找出可以使 ${}_i F(q) = {}_i F_p(q) + {}_i F_o(q)$ 為最小值的向量 q。

就求取上述目的解而言，人體關節構造 j_i 可能的自由度分 1 DOF、2 DOF、3 DOF。

CASE1：若 j_i 為 1 DOF

令 j_i 的旋轉軸為 $axis_i$ (以區域座標系而言旋轉軸固定)，轉動角度為 ϕ ，

$$\overrightarrow{iP'_{ic}}(\phi) = R_{axis_i}(\phi) \overrightarrow{iP_{ic}}, \quad i=1, 2, \dots, n$$

$$\overrightarrow{i u'_{kc}}(\phi) = R_{axis_i}(\phi) \overrightarrow{i u_{kc}}, \quad k=1, 2, 3, \quad i=1, 2, \dots, n$$

要求 ${}_i F_p(q) = \left\| \overrightarrow{iE_d} - \overrightarrow{iE_c} \right\|^2$ 的最小值即是求

$$g_1(\phi) = \overrightarrow{iP_{id}} \cdot \overrightarrow{iP'_{ic}}(\phi) \text{ 的最大值} \quad (\text{Eq. 3-1})$$

同樣的要求 ${}_i F_o(q) = \sum_{k=1}^3 \left(\left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{i u'_{kc}} \right) - 1 \right)^2$ 的最小值即是求

$$g_2(\phi) = \sum_{k=1}^3 w_{ok} \left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{i u'_{kc}}(\phi) \right) \text{ 的最大值}$$

$$(\text{Eq. 3-2})$$

結合(Eq. 3-1)、(Eq. 3-2)，且再加入權數(weight factor) w_p 則

$$g(\phi) = w_p g_1(\phi) + g_2(\phi) \quad (\text{Eq. 3-3})$$

解 IK 的問題就變成如下：

找出 ϕ^* ，使得 $g(\phi^*) = \max \{g(\phi) \mid \phi' \leq \phi \leq \phi''\}$

where $\phi' = \phi - \phi^l$ ， $\phi'' = \phi + \phi^u$ ，

ϕ^l 、 ϕ^u 分別為關節的 J_i 轉動角度的下限和上限

ϕ 則是為目前關節的 J_i 轉動角度

又

$$\begin{aligned} R_{axis_i}(\phi) \overrightarrow{iP_{ic}} &= \\ \overrightarrow{axis_i} \left(\overrightarrow{iP_{ic}} \cdot \overrightarrow{axis_i} \right) (1 - \cos_i \phi) + \overrightarrow{iP_{ic}} \cos_i \phi + \left(\overrightarrow{axis_i} \times \overrightarrow{iP_{ic}} \right) \sin_i \phi \end{aligned} \quad (\text{Eq. 3-4})$$

將(Eq. 3-4)代入(Eq. 3-1)

$$\begin{aligned} g_1(\phi) &= \overrightarrow{iP_{id}} \cdot R_{axis_i}(\phi) \overrightarrow{iP_{ic}} \\ g_1(\phi) &= \left(\overrightarrow{iP_{id}} \cdot \overrightarrow{axis_i} \right) \left(\overrightarrow{iP_{ic}} \cdot \overrightarrow{axis_i} \right) (1 - \cos_i \phi) + \\ &\quad \left(\overrightarrow{iP_{id}} \cdot \overrightarrow{iP_{ic}} \right) \cos_i \phi + \overrightarrow{iP_{id}} \cdot \left(\overrightarrow{axis_i} \times \overrightarrow{iP_{ic}} \right) \sin_i \phi \end{aligned} \quad (\text{Eq. 3-5})$$

同樣的(Eq. 3-2)可寫成

$$\begin{aligned} g_2(\phi) &= \sum_{k=1}^3 w_{ok} \left(\overrightarrow{i u_{kd}} \cdot R_{axis_i}(\phi) \overrightarrow{i u_{kc}} \right) \\ g_2(\phi) &= \sum_{k=1}^3 w_{ok} \left\{ \left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{axis_i} \right) \left(\overrightarrow{i u_{kc}} \cdot \overrightarrow{axis_i} \right) (1 - \cos_i \phi) + \right. \\ &\quad \left. \left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{i u_{kc}} \right) \cos_i \phi + \overrightarrow{i u_{kd}} \cdot \left(\overrightarrow{axis_i} \times \overrightarrow{i u_{kc}} \right) \sin_i \phi \right\} \end{aligned} \quad (\text{Eq. 3-6})$$

將(Eq. 3-4)、(Eq. 3-5)整合到(Eq. 3-3)

$$g(\phi) = k_1 (1 - \cos \phi) + k_2 \cos \phi + k_3 \sin \phi$$

where k_1 、 k_2 、 k_3 皆為係數

$$k_1 = w_p \left(\overrightarrow{iP_{id}} \cdot \overrightarrow{axis_i} \right) \left(\overrightarrow{iP_{ic}} \cdot \overrightarrow{axis_i} \right) + \left(\sum_{k=1}^3 w_{ok} \left(\overrightarrow{i u_{kd}} \cdot \overrightarrow{axis_i} \right) \left(\overrightarrow{i u_{kc}} \cdot \overrightarrow{axis_i} \right) \right)$$

$$k_2 = w_p (\overrightarrow{P_{id}} \cdot \overrightarrow{P_{ic}}) + \sum_{k=1}^3 w_{ok} (\overrightarrow{u_{kd}} \cdot \overrightarrow{u_{kc}})$$

$$k_3 = \text{axis}_i \cdot \left[w_p (\overrightarrow{P_{id}} \times \overrightarrow{P_{ic}}) + \sum_{k=1}^3 w_{ok} (\overrightarrow{u_{kd}} \times \overrightarrow{u_{kc}}) \right]$$

若關節角度沒有限制的話，則 $g(\phi)$ 的最大值成立於

$$\frac{dg(\phi)}{d\phi} = (k_1 - k_2) \sin_i \phi + k_3 \cos_i \phi = 0$$

且

$$\frac{d^2 g(\phi)}{d^2 \phi} = (k_1 - k_2) \cos_i \phi - k_3 \sin_i \phi < 0$$

但實際上關節角度都是有限制的，

令上兩式所求的角度為 $\bar{\phi}$ ，

If $\phi^l \leq \bar{\phi} \leq \phi^u$

then $\phi^* = \bar{\phi}$ ，

elseif $g(\phi^u) > g(\phi^l)$

then $\phi^* = \phi^u$

else

$\phi^* = \phi^l$

CASE2：若 j_i 為 2、3 DOF

可將 n DOF 轉換成 n 個 1 DOF，但在角度限制要另外檢查。以 3 DOF 為例，令第 j_i 的關節為 3 DOF，將 3 DOF 的關節 j_i 轉換成 3 個 1 DOF，其關節限制的上下限各為 $\phi_x^u, \phi_x^l, \phi_y^u, \phi_y^l, \phi_z^u, \phi_z^l$ 。CCD 方法未執行時， $\overrightarrow{w_{j_i j_{i+1}}}$ 在第 j_i 個關節區域座標系的向量為 $\overrightarrow{V}$ ，假設 CCD 方法執行到第 k 個疊代結束時，

$$\overrightarrow{V}^* = {}_i M_{k_i} M_{k-1} \dots {}_i M_1 \overrightarrow{V} = {}_i M^k \overrightarrow{V}，$$

$$\text{其中 } {}_i M^k = {}_i M_{k_i} M_{k-1} \dots {}_i M_1$$

$$\text{又 } {}_i M^k = R_x(\phi_x^*) R_y(\phi_y^*) R_z(\phi_z^*)，$$

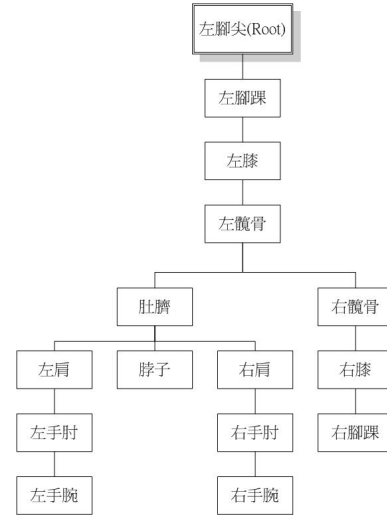
如此一來就可以得到各個旋轉軸的旋轉角度，接下來只要檢查並修正角度是在合理的範圍裡面，若有不合理的角度出現，就重新設定，使角度在合理範圍內而得到新的旋轉角度 $\phi_x^*, \phi_y^*, \phi_z^*$ ，然後再將角度轉換成旋轉矩

$${}_i M^K = R_x(\phi_x^*) R_y(\phi_y^*) R_z(\phi_z^*)，$$

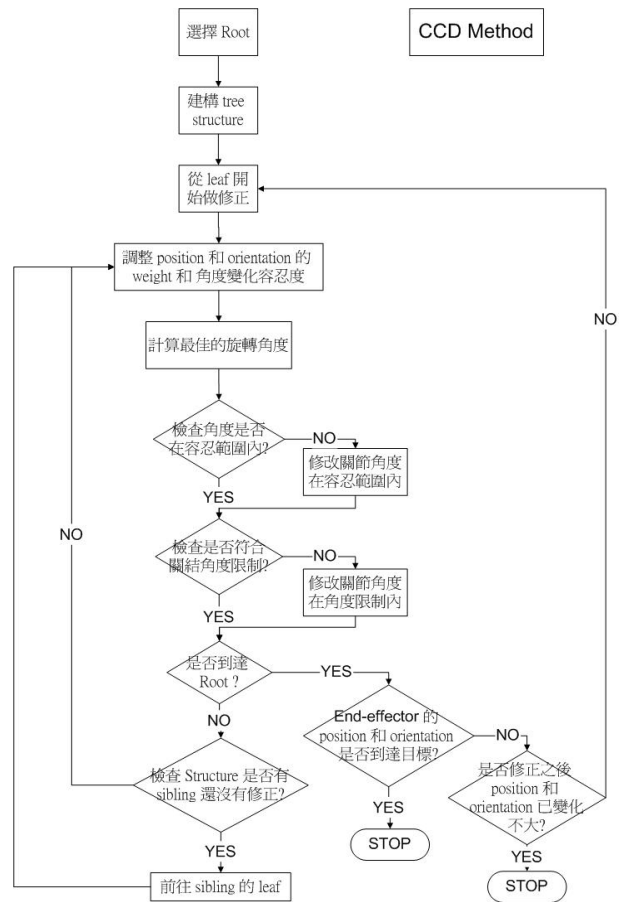
繼續一下次的疊代。每次疊代的次序可以建構成樹狀結構如(圖 15)。

每次疊代修正的關節順序由樹狀圖的 leaf 開始往 parent 修正，當所有的 children 都修正完之後能輪到本身做修正。假設要修正樹狀圖中的肚臍的

話，必需等到樹狀圖所肚臍所有的 children(左肩、脖子、右肩、左手肘、左手腕、右手肘、右手腕)修正完後才能換肚臍做修正。(圖 16)為整個 CCD 演算法的流程圖。



(圖 15) CCD 疊代順序樹狀圖



(圖 16) CCD 演算法流程圖

3.5.2 加強 CCD 演算法

我們對原始的 CCD 演算法作以下三點改進，

以改善收斂的效果並使其符合人體關節先天的限制。分述如下：

(1) 人體關節角度限制：

因為求解 IK 單純只是數學上的運算求出肢節的旋轉角度，並沒有考量到人體的骨架結構，導致雖然肢節端點可以到達目標位置，卻不是人體所能擺出來的姿勢，所以必須考量人體運動時肢節所能轉動的角度範圍，確保不會有不合理的人體肢節角度出現。當計算出來的角度限制超過資料中人體肢節所能轉動的角度範圍，就將計算出來的角度重新設定在合理的肢節角度範圍內，使其動作能夠符合人體骨架的限制。

(2) 角度變化容忍度：

經由 3.3、3.4 節方法所求得關節角度初始值雖會有誤差，但基本上與實際的角度相去不遠，在做修正的時候關節角度不可能會有大幅度的改變，若求解出來的角度有大幅度改變的話反而會影響到下一次的疊代。因此可以加入加角度變化容忍度，使得關節角度的修正量受到限制。

基於上面的想法，修正角度過大，之後的疊代勢必會將角度再修回來，如此一來有可能增加疊代的次數，也有可能導致做修正時陷入錯誤的區域最佳值中。所以加入了加角度變化容忍度除了可以使整個系統速度提升之外，也可以避免求得錯誤的關節角度。

(3) 隨時更新位置與方向的 weight：

由於 CCD 方法在做修正時，位置誤差(position error)和方向誤差(orientation error)隨時都在變動，因此要隨時更新位置和方向的 weight。當位置誤差修正變小時，有可能導致方向誤差變大，因此要將方向的 weight 加大，使得下次要做修正時在方向的考量方面加大。位置誤差 weight 調整，是以目前的 End-effector 和目標的距離來做調整，當距離越遠，weight 就會加大，距離越近，weight 就越小。至於方向誤差的 weight 調整，則採三個方向向量各別考慮，當方向的夾角越大，weight 就會加重，反之則變小。

4. 實驗結果

● 拍攝影像

我們架設兩部高解析數位攝影機，由表演者的左前方與右前方同步拍攝。攝影機架設好後，位置與焦距皆固定不變。攝影機的解析度為 1024x768 像素，拍攝速度為每秒 15 張畫面。

將攝得的影片轉換為連續的靜態影像序列，再

將各張影像使用 3.2 節的方法轉換為人體肢節輪廓，如(圖 1)~(圖 4)所示。

● 肢節基準中心線計算與標示

利用 3.2 中描述的方法，我們找出人體的肢節輪廓影像，見(圖 17, column1)，再利用 3.3 節中提到的方法，我們實際計算出人體肢節的對稱軸線，由圖(圖 17, column2)我們可以看到找出來的結果是相當令人滿意的，接著對軸線進行 least square 取出肢體的方向向量(圖 17, column3)也符合人眼辨識的結果，且在不同時間點的肢節對映也相當的成功，見(圖 17, column2)以特定顏色表示特定肢節。

● 重訂所得肢節點與 IK 修正過後的肢節點比較

為了檢視 3.4 節重建得到的骨架模型與真實人體姿勢的差距，我們將模型中各重要肢節點，投影至原始的攝影機影像中，以白色方塊標示，見(圖 18)中的(a1)~(d1)子圖。由圖中可以看出，root 的肢節點與原始影像完全吻合，距離 root 越遠的肢節點則偏離的越嚴重，尤其是手掌的肢節點以完全跑出人體輪廓之外。

以同樣的方法，將 3.5 節逆向動力學修正得到的骨架模型肢節點，投影至原始影像中，以白色方塊標示，見(圖 17)中的(a2)~(d2)子圖。可看出圖中各肢節點幾乎都能和影像中的人體吻合，原本偏離很嚴重的手掌肢節點也能修正的回來，證實我們提出的 IK 修正演算法有相當程度的效果。

● 動作捕捉所得的三維肢節運動

(圖 19)顯示的是本計畫最終的目的，動作捕捉得到人體動作序列。最右側直行中的藍色 Stick figure 就左側二張原始影像經過 3.1~3.5 節各步驟演算法，運算得到的三維骨架模型以 MATLAB render 得到的結果。

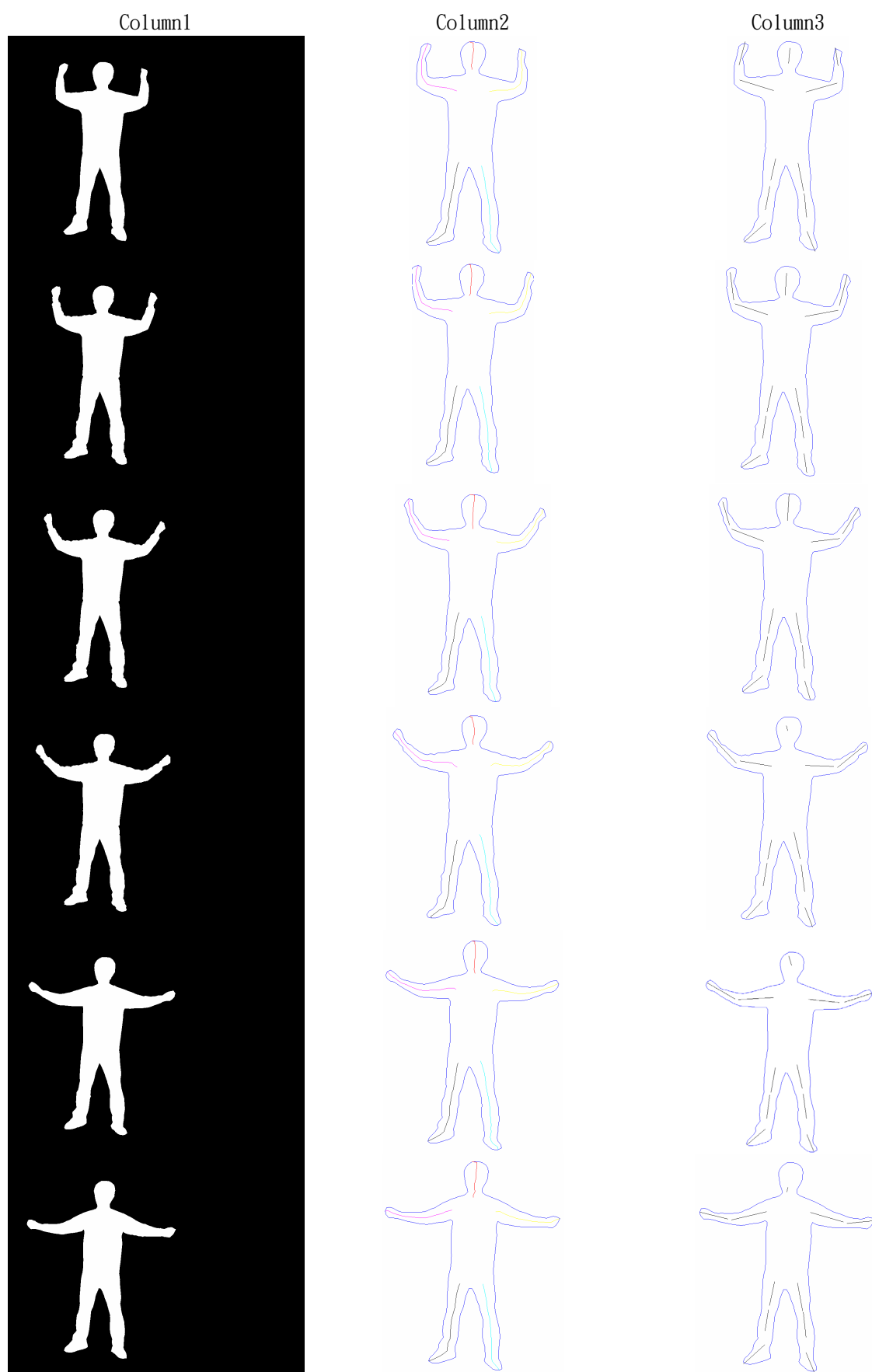
5. 結論

本計畫提出一套完整的動作捕捉系統，可由二台數位攝影機拍攝得到的影像，在不需額外 markers 的前提下，全自動將表演者的動作，轉換為三維的動作參數，並用我們定義的人體骨架模型來表示。這些動作參數將可作為未來更進一步動作分析(Motion analysis)研究的基礎。

6. 參考文獻

[1] T. Moeslund "Summaries of 107 computer

- vision-based human motion capture papers", Technical Report LIA 99-01, University of Aalborg, March 1999.
- [2] T. Moeslund "Computer vision-based human motion capture-a survey", Technical Report LIA 99-02, University of Aalborg, March 1999.
- [3] D.M. Gavriila and L.S. Davis, "3-D model-based tracking of human in action: A multi-view approach," in Proc. IEEE Int. Conf. Computer Vision pattern Recognition, CVPR' 96, pp.73-80, June 1996.
- [4] Gray R.Bradschi and James Davis, "Motion Segmentation and Pose Recognition with Motion History Gradients" Proc IEEE Workshop on Applications of Computer Vision, pp. 238-244, Dec. 2000.
- [5] Jia-Ching Cheng and Jose M. F. Moura, "Capture and Representation of Human Walking in Live Video Sequences", IEEE Trans On Multimedia, Vol.1, No.2, pp. 144-156, June 1999.
- [6] M. K. Leung and Y.H. Yang, "First Sight: A Human Body Outline Labeling System" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.17, no.4, pp. 359-377, April 1995.
- [7] E. Durucan and T. Ebrahimi, "Change Detection and Background Extraction by Linear Algebra," Proceedings of the IEEE, Vol.89, No.10, pp.1368-1381, Oct. 2001.
- [8] Rafael C. Gonzalez & Richard E. Woods "Digital Image Processing"
- [9] Li-Chun Tommy Wang and Chih Cheng Chen, "A Combined Optimization Methods for Solving the Inverse Kinematics Problem of Mechanical Manipulators", IEEE Transactions on Robotics and Automation, Vol. 7, NO. 4, August 1991
- [10] A. Sezgin Abal, "Animation of human option with inverse kinematics using nonlinear programming", b_ilkent university, master thesis, September, 2001
- [11] Chris Welman, "Inverse kinematics and geometric constraints for articulated figure manipulation", Simon Fraser University, master thesis, September 1993
- [12] Jianmin Zhao and Norman I. Badler, "Inverse Kinematics Positioning Using Nonlinear Programming for Highly Articulated Figures", University of Pennsylvania, ACM Transactions on Graphics, Vol. 13, No.4, October1994, Pages313 336
- [13] Naganand Madhavapeddy and Stuart Ferguson, "Specialised Constraints for an Inverse Kinematics Animation System Applied to Articulated Figures", The Queen's University of Belfast
- [14] 張家銘, "三維動畫中動作重置之研究", 國立清華大學, 碩士論文, 民國 90 年
- [15] D. Geiger, Tyng-Luh Liu, and R.V. Kohn, "Representation and Self-Similarity of Shapes", IEEE Transactions on Pattern Analysis & Machine Intelligence, Vol.25, No.1, January 2003.
- [16] R. Basri, L. Costa, D. Geiger, and D. Jacobs, "Determining the similarity of deformable shapes", Journal of Vision Research 38, pp.2365-2385, 1998.



(圖 17) Camara2 擷取到的影像，由上而下依照時間先後排序。
 Column 1: 肢節輪廓影像；Column 2: 肢節軸線；Column 3: 肢節向量



(a1) Time=10 / Camera1 重訂所得肢節點



(a2) Time=10 / Camera1 經 IK 修正後的肢節點



(b1) Time=10 / Camera2 重訂所得肢節點



(b2) Time=10 / Camera2 經 IK 修正後的肢節點



(c1) Time=18 / Camera1 重訂所得肢節點



(c2) Time=18 / Camera1 經 IK 修正後的肢節點



(d1) Time=18 / Camera2 重訂所得之節點



(d2) Time=18 / Camera2 經 IK 修正後的肢節點

(圖 18) 選用 Time=10, Time=18 兩個時間點重建得到的骨架模型與 IK 修正後的骨架模型，將其肢節點投影至 Camera1 與 Camera2 的原始影像中，標示為白色方塊。可以看出 IK 演算法修正的效果。

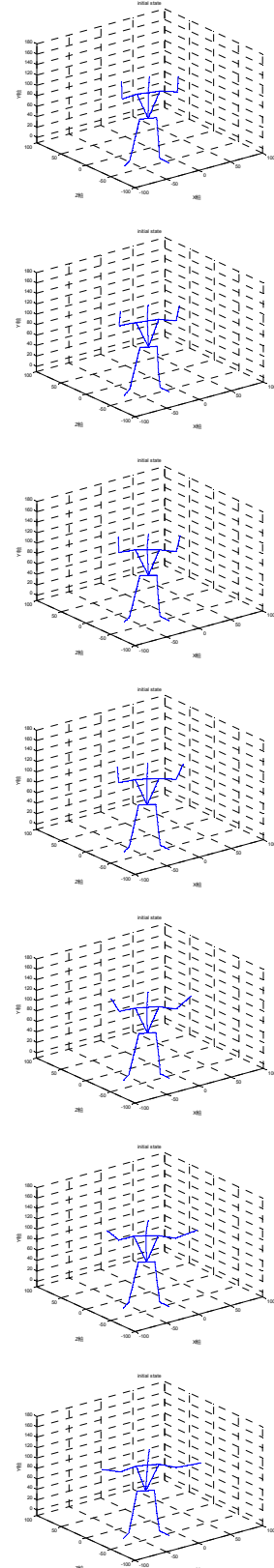
Camera1 原始影像



Camera2 原始影像



Motion Capture 得到的三維骨架



(圖 19) Motion Capture 所得的三維骨架模型，由上而下依照時間先後排序。