

第二章 相關研究回顧

Relevant Work

由於迅速發展的計算機硬體技術與計算機軟體系統的不斷改進，使得基於大型數據集合的聲音和圖像的即時運算成爲可能，電腦不再只是幫助設計者畫圖的工具而已，更延伸爲協助設計思考的一種媒材。Tamotsu及Naomasa(1994)最早提出了虛擬輔助設計(Virtual Reality Aided Design, VRAD)的概念。虛擬輔助設計的出現讓設計者可以不經過傳統2D繪圖的過程，直接藉由虛擬建模軟體，在3D的環境中進行設計。在虛擬輔助設計的建模過程中，設計者與建模環境的互動模式是影響設計者最主要的關鍵(Donath 1999)。因此，使用者與介面的關係成爲在研究虛擬輔助建模時一個重要的課題。本章主要在根據研究所說明的1)使用者與介面的關係，2)VRAD建模軟體的發展，以及3)手勢辨識技術(包括3D軌跡追蹤器等)。



2.1 虛擬建模互動要素探討

Bodker (1991)認爲在虛擬建模的環境中，使用者與軟體的互動模式上，主要包含兩個因素。一是如何將使用者所提供的訊息傳送到電腦及如何將電腦的訊息呈現給使用者。另一則是使用者介面透過介面設計的暗示或隱喻，而能整合使用者的互動和訊息表現，進而達到互動之機能。因此，設計者在虛擬輔助設計環境中，一般而言，其重點包括直接操作、介面模型、WIMP 介面等要項，以下將分別闡述之。

2.1.1 直接操作

所謂直接操作就是透過取代訊息或移動訊息等直接動作來互動，並且不需要明確的命令修改訊息，如一般的物件移動等。操作方式如在實體環境下一樣，可立即達到互動的功能。其主要優點在於：1)使用者能立即掌握虛擬環境。2)學習時間相當短。3)能得到立即的回饋。4)發現錯誤並能迅速訂正(Foley 1987)。其主要問題在於如何將使用者日常生活的習慣，反映在軟體的使用上；例如在生活上要拿起茶壺時腦海中會浮現抓取的意像，在虛擬環境中，通常還是會做如此的聯

想，虛擬環境若利用這種思考邏輯去架構，會讓使用者在操作上更易明瞭。不過值得注意的是，在某些操作功能上(如儲存檔案)大致上可得知有些操作功能上，很難找到類似的隱喻或合適模型，來幫助使用者了解詮釋，這時就必須倚靠傳統的視窗選單介面。

2.1.2 介面模型

就介面模型而言，一般泛指所謂概念模型(Conceptual Model)與心智模型(User mental model)(Allen 1990) (圖 2-1)。概念模型則是程式設計師與介面設計師或其他相關工程人員所共同設計的硬體或軟體介面，使用者可以藉由此介面模型控制軟硬體；所謂心智模型而言，則是藉由對軟硬體介面設備之互動之後，使用者對於該系統所形成之心智模型，一般而言是自然地逐步形成，有時甚至會轉化爲一種習慣，並藉由與系統不斷之互動 (Interaction)，使用者勢必會不斷修正其心智模型 (Allen 1990)。那麼當我們進行介面設計時，設計者便必須不斷考量此概念模型與心智模型的演進情形，如此系統設計才將能給予使用者完整易於理解的人機介面環境。

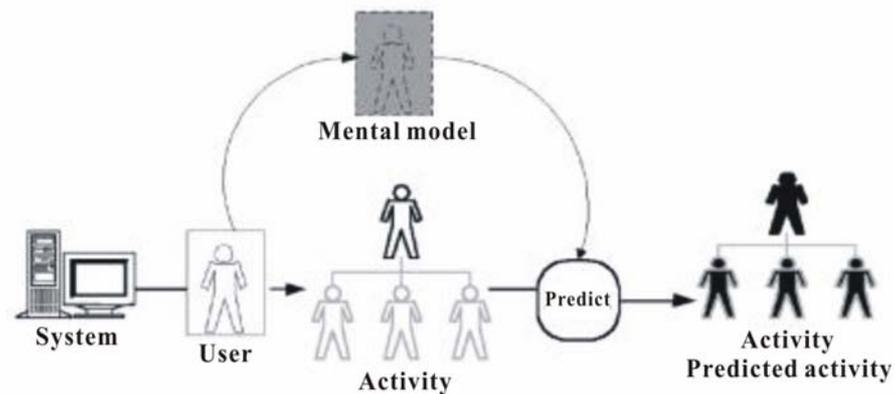


圖 2-1 使用者心智模型(Allen 1990)

2.1.3 WIMP 介面

WIMP(Window-Indicator-Menu-Pointer, WIMP)介面是簡而易懂的操做方式，但是由於在使用上並不直覺，所以一般來說，當使用者無法使用直覺性行爲來表達指令動作時，視窗選單才會被選用做爲指令下達方式(Allen 1990)。而常見的視窗選單指令如下(Siegfried 1994)：

- 1) 按鈕：一般作用在於單一動作按鈕，如執行、儲存等。
- 2) 開關：可以在位置內設定系統或是將該系統由一個狀態移至另一狀態。
- 3) 選單：可以選擇的按鈕或是開關的集合，部分選單設計也常看到下拉式選單、捲動選單、階層式選單、移動式選單之設計方式、關聯性控制面板等。
- 4) 指示器：介面設計主動顯示正在發生的某些動作，如計算中、忙碌中等。

- 5) 顯示：控制面板可以顯示部分資訊，如圖形或文字訊息等，甚至可以控制這些顯示及編輯。
- 6) 捲軸：一般設計為設定一定範圍特定的輸入值的輸入介面。譬如音量調節器上之音量控制。
- 7) 拖拉鍵：便於方便顯示主題內容等。
- 8) 其它特殊的控制介面：種類繁多，有結合 Geographical Information System (GIS)方式操作方式或三維之操作介面等。

一般而言，視窗選單介面在被設計的過程中，常常忽略"究竟為誰設計"這個重要的因素，一切只是為了解決現有問題為導向，而忽略了使用者特性(Raskin 1999)。因此，視窗選單雖然簡而易懂，使用者卻必須花費更多時間學習及適應。

2.2 VRAD 使用者介面原則

而從介面的角度來看，在 1987 年蘋果公司發表的三項一般性介面設計原則(Siegfried 1994)：1) 在電腦環境裡，若是能有現實的隱喻(Metaphor)將會帶給使用者熟悉的預期。2) 立即的回饋 (Instantly Feedback) 將會帶給使用者似乎直接操縱現實世界物體的感覺，也提供使用者有價值和事實資訊。3) 必須看到點到的感覺，如此可以獲得使用者的信任而相信自己的記憶。根據這三項原則，Donath(1999)進而提出一個良好的 VRAD 使用介面在操作上應具備以下四點：

- 1) 效率好、品質好：不必思考如何操作，容易預想操作的結果，操作的方法易於理解，沒有錯誤動作。
- 2) 不必學習或容易學習：不必預先學習如何操作，操作方法容易記憶。
- 3) 學習之後不會忘記：使用一次後，其操作方法終身不忘。
- 4) 使用者不會有不愉快，而能達到滿足感：能夠安心的操作，達到滿足感。

此外，Donath(1999)亦針對初學者的使用介面，提出六點應具備的條件：

- 1) 使初學者瞭解操作的目的、系統的概念與模式。
- 2) 採用初學者喜歡且容易理解的操作方式。
- 3) 對於系統的各项操作機能與方法，最初只教導最基本的功能，再慢慢學習其它功能。
- 4) 統一的系統操作方法。
- 5) 舉例示範。
- 6) 為了消除操使用者的不安，應具有隨時能回到原操作畫面之功能。

要滿足這些條件，Donath 認為介面所呈現的訊息，對使用者而言其所認為重要之程度，必須劃分清楚與明瞭。此外，由於在某些情況之下，使用者必須迅速得知參數值的改變情形，介面與使

用者必須藉由直接操作和訊息產生互動。因此，在虛擬建模環境中，自然行為與控制指令的結合是有必要的。

2.3 早期虛擬建模案例

VRAD 之所以不同於 CAD，是因為其強化了 3D 實體建模的功能，並讓使用者可以直接創建 3D 物件(Donath 1999)。而現今的 VRAD 建模軟體已經複雜到足以描述任何"人"可以想像的事物，但並非一開始就如此。本節所要介紹的是早期所發展的虛擬建模軟體，並說明其如何描述及操作虛擬空間。

2.3.1 voxDesign

voxDesign 是 igroup 公司於 1985 年所開發的 3D 即時建模軟體，也是最早被開發出來的 VRAD 建模軟體(Kalawsky 1993)。由於技術的限制，這套軟體在空間建構上只提供兩種元件：點與正方體。在這個環境中所創建的物體可以藉由筆式工具(pen-like device)(Polhemus Stylus)點選之後再移動位置，對使用者來說，這個簡單的軟體只要會按鈕和使用選單功能就足夠了(圖 2-2)。

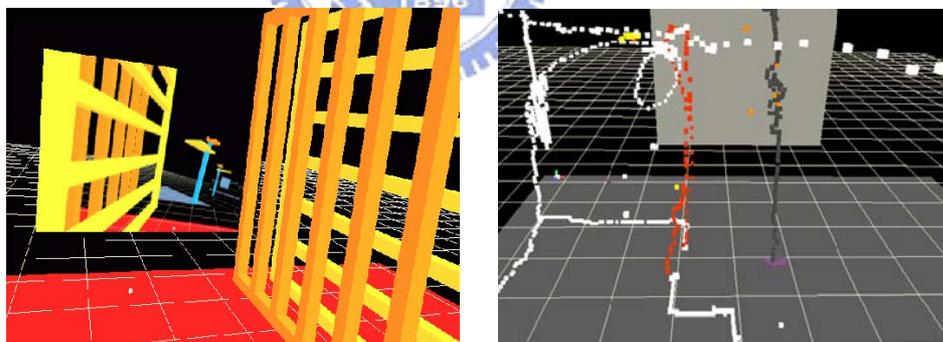


圖 2-2 voxDesign 建模環境(取自 Kalawsky 1993)

2.3.2 planeDesign

planeDesign 是 Loja Virtual 於 1988 開發的虛擬建模軟體(William and Craig 2002)。比起 voxDesign 只多了一個物件：一種可以設定長寬的矩形平面。有了這個物件，在使用上一樣不須經過訓練的過程，但是能夠被清楚創建的型態卻多了上百種。這套軟體是針對建築領域的使用者所開發的，主要目的是提供使用者在設計上更多的幫助(尤其是概念生成階段)，使用者可以將自己每個想法描繪在虛擬空間裡。雖然軟體裡沒有比例或精確位置的概念，使用者還是能夠自由的決定，樓板，

柱子及牆等建築元素的位置(圖 2-3)。

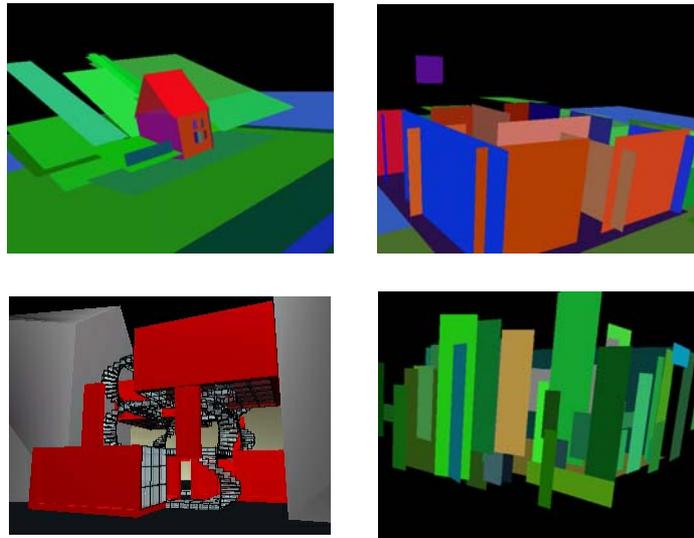


圖 2-3 planeDesign 中由面組成的空間(取自 William and Craig 2002)

2.3.3 VRAM

VRAM(Virtual Reality Aided Modeller)是 Bauhaus University Weimar (BUW)於 1993 針對完全體驗 3D 虛擬環境的目的所開發的(Regenbrecht et al. 1993)。透過頭戴式顯示器(Head-Mounted Display, HMD)以及數據手套(Data Glove)，使用者可以完全擺脫被螢幕及滑鼠所限制住的 2D 環境。在軟體方面包括創建，修改，瀏覽 3D 模型的功能。VRAM 所使用的是 3DUI (Testbed for Threedimensional User Interface)的技術，比起 voxDesign 與 planeDesign 具有更強大的定位及建模功能。正因為如此，其在介面上的設計也比前述兩者來的複雜許多，而無法只倚靠單純的指令動作(圖 2-4)。

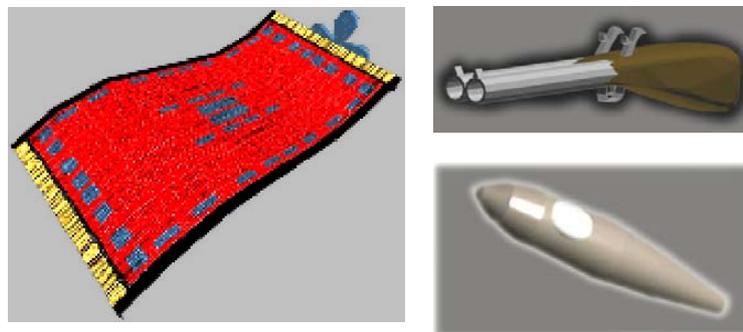


圖 2-4 VRAM 已可製作複雜的模型(Regenbrecht et al. 1993)

2.3.4 結論

透過這些案例的回顧，我們可以知道，隨著硬體及軟體技術的快速發展，虛擬建模軟體為了滿足

使用者的需求而變的越來越複雜，以致於無法使用單純的指令行為達到創建或修改 3D 實體模型的目的。為了解決這個問題，在虛擬建模的環境下，最後還是必須倚靠 WIMP 的功能。

2.4 3D 軌跡追蹤器

要如何將人在實體環境中的動作轉換為電腦可讀取的資訊，3D 軌跡追蹤器扮演了很重要的角色。在過去，設計者只關心追蹤器是不是能提供足夠的更新率，然而現在卻要考慮其它因素，包括靜態準確度(static accuracy)，動態準確度(dynamic accuracy)以及即時相延遲(phase lag)等(Ouyang 1996)。特別是時相延遲，由於現在的追蹤器都有處理器進行運算，如果物體持續移動，輸出結果將永遠不等於目前位置。如果我們看到的動作和實際的步調不一致，那會產生很多問題，如物件怎樣都抓不到，因為他們看到的手和他們手的實際動作並不能配合。

為了改善顯示的及追蹤到的動作不一致問題，現在也使用了預測技術(Melianthe 2000)，預先推測下一畫面的位置。根據 Ouyang(1996)的分析，追蹤器的原理與方式有下列五種，根據不同的使用而有不同的定位方式：

- 1) 機械式：利用機械臂及彎曲測量儀器中各關節的轉動角度，就可以計算受測物的位置。具有高精準度、低時間延遲（對頭部位置追蹤很重要），不用收發裝置，所以不易受外界環境影響。但是追蹤範圍很小，會磨損，對四肢運動產生機械的遲滯(inertia)，而且價格昂貴。
- 2) 電磁式：由三個正交的發信器發出磁場，經由三組接收器接收，可由收發的信號變化計算出受測物的位置。體積小巧(適合裝在頭部，手上及筆尖)，可將收發器組合，分時多工，以得到更大的追蹤區域。缺點是易受無線電等電磁干擾，如果附近有像鐵的磁性金屬類物品，也會產生干擾。系統時間延遲約 100ms，也算比較長的。目前已經致力於減低延遲。
- 3) 光學式：將光源，如紅外線，放在受測物上，再由安放於四周的接收器（攝影機）接收，據以計算位置。好處是價位低、追蹤範圍廣大(取決於光強度及同相位程度)，但是必須要光能照到，也就是光的路徑中不能有障礙物。
- 4) 超音波式：由幾個發聲器發出聲音（超音波），讓幾個麥克風接收。經由測定時間差來計算受測物距離。需要幾組收發器才能得到完整的 3D 位置資料。另外一種較少見的技術是經由計算收發器信號的相位差，測出受測物的位移。轉換器容易取得，可以以低成本製作。但是音速受外界溫度，氣壓及溼度影響而有不同，音響式系統必須先考慮這些，不然會有

很大誤差。此外，聲音碰到"硬表面"會產生回音，造成"鬼影"現象。收發器間也不可以有東西擋住。這使它的準確度及追蹤範圍都受限。

- 5) 慣性式及重力式：慣性式用陀螺儀測量三軸的轉動量。過去用於飛機或飛彈的導引裝置。現在也應用於一些防震(no shake)手提攝影機及無線遙控直昇機。重力式則用地心引力加上單擺測量相對於地的旋轉角度。優點是無收發裝置。但是陀螺儀難以和其它裝置溝通，且會隨時間偏移。(每分鐘幾度)，也需要做溫度校正。不過新一代的"固態(solid-state)陀螺儀"(實為一種壓電設備)可減低這些效應兩三倍。另外，重力式與陀螺儀皆無法準確得知平移的位置，但對旋轉很敏感，只能提供旋轉資訊。

機械式與光學式為目前最常被使用的 3D 軌跡追蹤方式，主要因為機械式的軌跡追蹤高精準度與低時相延遲，且可配合軟體製造力回饋(Force Feedback)的效果；而光學式的軌跡追蹤則輕便便宜，且價格便宜，在辨識度上也有不錯的表現。

2.5 手勢辨識



手勢動作包含手的結構和手的運動，因此手勢辨識是包括了時間上和空間上手形變化的分析。手勢辨識的研究在近十年來已經激發了許研究者的興趣，早期的研究大部分是與機器人的操作有關。Simeon 及 Peter(1998)則提出最近在手勢的應用研究上相關的介面技術共分為三大類：以數據手套為基礎，以手勢影像辨識為基礎和以 3D 滑鼠為基礎。

2.5.1 數據手套

在需要利用手指的軌跡和手運動來操作的工作上，我們常使用特殊的手套裝置(如圖 2-5)。以手套為基礎的裝置是最適合用在以操作為目的手勢。假如使用者想要在虛擬的環境下操作，應該較適合使用類似手套的裝置。自從 80 年代以後，這些以手套為基礎的裝置已經使用某些附加在手套上的感測器(機械或光學)，再利用轉換器能將手指的彈動及伸展轉換成電的信號(Burdea and Richardand 1996)。原始發展的目的是針對機器人的應用，以做標記的方式用做於手功能和損傷之臨床分析(Lee et al. 1998)。資料手套裝置是手勢辨識中，利用手套為基礎的研究當中使用最廣的。

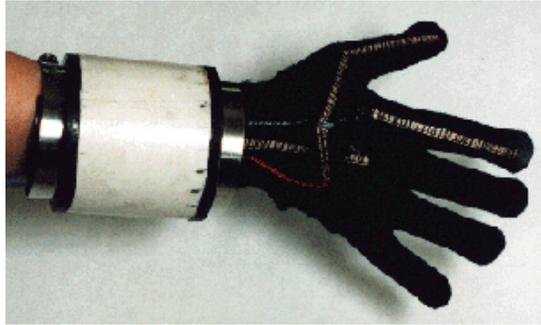


圖 2-5 數據手套(取自 Burdea and Richardand 1996)

2.5.2 手勢影像辨識

手勢影像辨識是以電腦視覺(computer vision)影像辨識為基礎的手勢分析(圖 2-6)。其技術可由分析 2D 的手勢外形之時變來分析關節的運動和其它的運動軌跡。以電腦視覺為基礎的手勢辨識，原則上是一種建構人類與電腦之手勢介面較為自然的方法。它與人類利用視覺獲得視訊資料的系統相類似。但由於目前電腦視覺技術的發展仍有其限制，結果很難令人滿意。最主要的困難是當 2D 的手勢外形從任何角度觀察都很像時(如握拳時)，電腦無法判斷當下的手勢動作到底為何。而其他困難如手指和手臂的關節運動，2D 運動軌跡觀察點的敏銳度，手勢運動和背景運動之間的干擾，關節物體(手)被部份遮蔽的問題都不易解決(Huang and Pavlovic 1995)。為了設計一個較可行的手勢辨識系統，手勢動作常常被局限在一個有限的空間和時間的定義域中。此外，對於一個較優良的視覺基礎演算法而言，最主要的假設是畫面的取樣率要足夠的快，尤其手的形狀和手指的運動變化量，在兩個連續的框間是小量的(Burdea and Richardand 1996)。以電腦視覺為基礎的辨識方法，主要是利用分析不同的手指和不同的移動軌跡的變異來解釋手勢。由於從相機所拍攝到的影像去偵測手的結構是高難度的問題，因此，在克服這個問題上，可用一些特殊手勢辨識的技術，如使用紅外線 LED 系統或反射標記系統。反射標記通常放在指尖上。一般來說，使用立體鏡頭相機和黑色手套並在指尖、手腕、手肘或肩膀塗上顏色，都可以減少手勢在辨識上的困難度。即便是如此，在辨識精準度上還是不如數據手套(Ouyang 1996)。

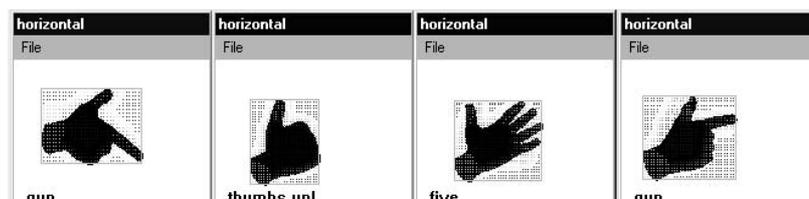


圖 2-6 影像辨識技術(Gross and Kemp 2001)

2.5.3 3D 滑鼠

另一種在手勢辨識中所使用的方法是利用 3D 滑鼠做為工具(圖 2-7),再以輸入的參數來計算符合手掌方向和關節角度。其主要是利用軌跡球,來分析手關節的運動。最近發展即時手勢辨識系統,以配合使用 3D 柱狀模型,且設計成具有 27 個自由度,指尖和每一段指骨被選擇當成匹配模型的特徵,而它們是由單一或立體影像中所萃取出(Burdea and Coiffet 1994)。開始的方法是適應手腕位置的部份,可以移動手的模型從任意的位置和方向到手影像中附近的一個。最後的方法是適應手指位置的部份。在以上的辨識法中,最消耗時間的步驟是手指適應的過程。



圖 2-7 3D 滑鼠 (取自 Burdea and Richardand 1996)

2.5.4 三種辨識法之比較

在瞭解了三種辨識法之後,本研究將其比較資料列表如下:

	以手套為基礎	以影像辨識為基礎	以 3D 滑鼠為基礎
輸入工具	數據手套	攝影機	3D 滑鼠
辨識精細度	極高	低	低
操控性	極高	低	中
輕巧度	低	極高	中
手勢轉換的情況	十分良好	良好	非直覺性轉換
價格	高	低	中
市場佔有率	高	極低	高

表 2-1 三種辨識法比較資料

三種辨識法各有其缺點,哪一種比較好事實上也沒有一定的標準,端看使用者最後的目的為何,一般來說,以手套為基礎的模式大多應用在虛擬實境的瀏覽與遠距離機械操作上,因為這些動作通常需要一定的精準程度(Bimber 1999)。以影像辨識為基礎的模式雖然有很多相關的研究,但是由於這種方式辨識精細度過低,不過價錢便宜,所以容易普及,目前真正有實際應用這種技術的

以娛樂器材為主。以 3D 滑鼠為基礎的模式功能與以手套為基礎的模式類似，不過為非直覺性的轉換模式在空間瀏覽上常常出現控制不良的情況，因此雖然這種模式是前述三種中最早商品化的，但是在使用上還是無法普及。

2.6 手勢與虛擬建模環境整合之案例

虛擬建模概念初期，手勢輸入並非一種被其廣泛應用的輸入方式，直到 1894 年時，Nemeth(1984) 首次將手勢輸入應用在 VRAD，研究者才發現這是一種值得探討的概念，不過手勢與 VRAD 環境在整合上一直都存在著某些問題。以下將簡單介紹幾個著名的案例。

2.6.1 the Maestro project

手勢輸入一開始的目的並不是為了空間設計而做的，在 1982 年時，Thonmas(Lee et al. 1998)設計出第一套手勢輸入工具，其目的在於利用手勢動作控制機器手臂，直至 1984 年時 Nemeth(1984) 在 the Maestro project 中提出了利用手勢做為輸入工具時，手勢才算真正的應用到虛擬環境中。由於當時硬體與軟體技術的限制，the Maestro project 並不算是一個完整的虛擬建模環境，基本上只是一些內建幾何多邊體組合。但是 Nemeth 認為，一個能夠滿足設計者需求的虛擬建模環境，在指令上一定相當的繁瑣，所以便研究了設計者在設計過程中所會利用到的行為，包括手繪圖，製作實體模型時的切割，沾黏等動作，編輯了一套適用於設計者的虛擬建模手勢。為了追求完整的描述設計者的行為，這套手勢系統在內容上相當的龐大，甚至龐大到難以學習的地步。因為 Nemeth 當時認為根據一般設計者的自然行為一定可以找到一個可被所有人了解的自然手勢準則，可惜的是事實卻不是如此，就算排除每個人對自然手勢定義不同的問題，實體空間的動作也不能完全類比到虛擬空間的指令上，因此，開始有質疑手勢與虛擬環境的整合是否可行的問題(圖 2-8)。



圖 2-8 the Maestro project(Nemeth 1984)

2.6.2 VRAM

到了 1993 年，由 Regenbrecht(1993)所主導的案子 VRAM(Virtual Reality Aided Modeller)，打破了這中質疑。Regenbrecht 認為一個好的虛擬建模介面，最重要的就是讓使用者覺得“有趣”，而且必須簡單易學，因此，此案所建立的虛擬環境中，手勢動作所替代的指令只有選去動作而已，其他複雜難學的指令，全都交給虛擬環境中的 WIMP 介面解決。透過這種方式，Regenbrecht 讓手勢真正能利用在虛擬建模上，可惜的是，這套系統一般人是用不到的，因為“VRAM”的發展目的並非真正為了“建模”，正確的說，因該是為了“體驗”，所以必須要有工作站級的電腦以及及精密的手勢辨識裝置做為支援。雖然如此，Regenbrecht 還是證明了手勢與虛擬建模健模整合的可能性(圖 2-9)。



圖 2-9 VRAM(Regenbrecht 1993)

2.6.3 Gesture Modelling

近十年來，由於影像辨識技術的快速發展，讓手勢辨識可以在不借助於任何穿戴式機械設備的情況下，依然可以達到某種程度的精準，由 Christopher 所帶領的團隊便是致力於這方面的研究。

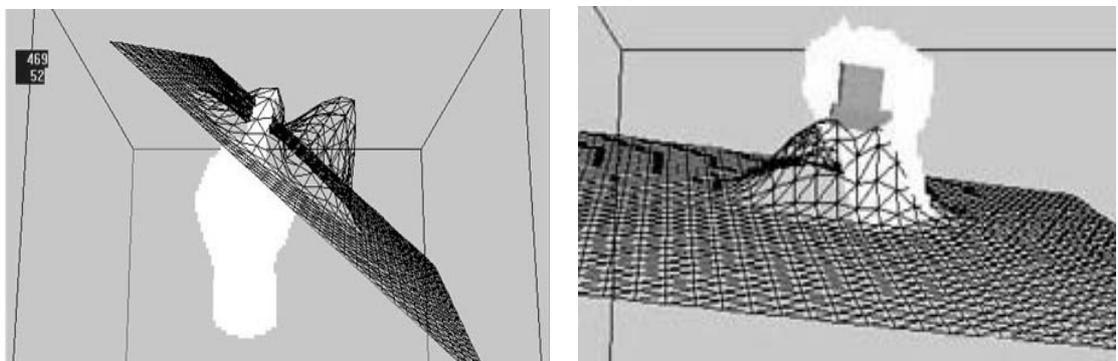


圖 2-10 Gesture Modelling(Mark 2000)

Gross 及 Kemp(2001)嘗試著將其應用到虛擬建模上，並命名為 Gesture Modelling。這種不需要

手套的辨識技術有幾個好處：輕便，手勢動作不易受到影響，可辨識區域擴及全身，價格便宜。唯一的缺點就是無法辨識連續動作，及相近的手勢。因此，在 *Gesture Modelling* 中，能夠被辨識的手勢只有兩種，其建模指令也受限於此，無法完成高複雜度的模型(圖 2-10)。

2.6.4 結論

從各種條件考量的結果，近幾年來，在虛擬建模研究中手勢辨識的方式，還是以彎曲感應器或光纖感應器搭配光學式 3D 軌跡追蹤器為主(Ouyang 1996)。而目前大部分這類研究討論的，只是單一手勢行為在在配套的建模環境下使用，或是整套設備太過於昂貴，一般的設計者根本無法使用。在這種情況下，要談手勢虛擬建模的真實應用還有一段距離。因此，本研究的目的並非完全在探討手勢建模的技術，而是如何利用手勢輸入的方式來幫助目前市面上主流之虛擬建模軟體，在使用介面上能夠更人性化。

