

行政院國家科學委員會專題研究計畫 成果報告

人機介面中智慧型虛擬滑鼠技術之研究

計畫類別：個別型計畫

計畫編號：NSC92-2218-E-216-002-

執行期間：92年10月01日至93年07月31日

執行單位：中華大學資訊管理學系

計畫主持人：張勤振

計畫參與人員：邱明毅, 陳伯璋, 潘雅真

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 9 月 6 日

行政院國家科學委員會補助專題研究計畫 成果報告
 期中進度報告

人機介面中智慧型虛擬滑鼠技術之研究
A Study on Techniques for Intelligent Virtual Mouse in Human
Computer Interface

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 92-2218-E-216-002

執行期間：92 年 10 月 01 日至 93 年 07 月 31 日

計畫主持人：張勤振 中華大學資訊管理學系

共同主持人：

計畫參與人員：邱明毅 中華大學資訊管理學系

陳伯璋 中華大學資訊管理學系

潘雅真 中華大學資訊管理學系

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

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

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：中華大學 資訊管理學系

中 華 民 國 93 年 09 月 06 日

行政院國家科學委員會專題研究計畫成果報告

人機介面中智慧型虛擬滑鼠技術之研究

A Study on Techniques for Intelligent Virtual Mouse in Human Computer Interface

計畫編號：NSC 92-2218-E-216-002

執行期限：92年10月01日至93年07月31日

主持人：張勤振 中華大學資訊管理學系

計畫參與人員：邱明毅 中華大學資訊管理學系

陳伯璋 中華大學資訊管理學系

潘雅真 中華大學資訊管理學系

中英文摘要

虛擬滑鼠提供使用者一個新的介面工具，利用電腦視覺之手勢辨識技術，直接用手勢來模擬滑鼠命令與電腦溝通，可以運用在許多新世代的人機互動中。然而，就目前國內外研究之電腦視覺手勢辨識相關技術，尚難同時達到虛擬滑鼠之即時性與高正確性的要求。因此，在本計畫中，我們針對幾個虛擬滑鼠技術困難點，嘗試提出改進方法，開發即時且具有高效能之虛擬滑鼠技術，提供使用者一種方便且自然的溝通工具。

關鍵詞：人機介面、虛擬滑鼠、手勢辨識

Virtual mice provide a new interface for human to interact with computers. Based on computer vision-based gesture recognition techniques, gestures are used to simulate mouse commands to interact with computers and can be used in new generation of HCI. However, previous gesture recognition approaches can not achieve the real time and high accuracy requirements for virtual mice. Therefore, in this project, we try to present improved gesture recognition techniques to solve some of the technology difficulties for intelligent virtual mice and develop real-time gesture recognition techniques with high accuracy for the convenient and natural use of intelligent virtual mouse in HCI.

Keywords: Human Computer Interface; Virtual Mouse; Gesture Recognition

一、前言

近年來，由於電腦的大量使用，人機互動(Human Computer Interaction)變得愈來愈重要，主要是它提供了使用者與電腦之間的互動與操作。目前，大部分的人機互動技術是利用鍵盤與滑鼠的輸入來讓使用者與電腦之間作訊息的溝通工具，然而隨著電腦計算、通訊與顯示技術的進步，在許多的人機互動中，例如多媒體系統、虛擬實境系統等，這些裝置使用起來在速度上與自然性有其限制，將會成為使用者與電腦之間有效互動的瓶頸。

虛擬滑鼠可提供使用者一個新的介面工具，可以運用在許多新世代的人機互動中。現存的虛擬滑鼠技術可分成兩類，其一是利用電子手套(Data Glove)的手勢來輸入滑鼠命令，主要是根據手勢在三度空間的描述做手勢辨識，來模擬滑鼠動作。但由於使用電子手套虛擬滑鼠，要讓使用者戴電子手套在方便性與自然性上有所限制，而且電子手套價格較昂貴。另一是以電腦視覺為基礎(Computer Vision-based)之手勢辨識，根據手勢在二度空間的視覺影像描述做手勢辨識，來模擬滑鼠的命令

動作，使用者只要自然方便的動動手勢，便能輸入資料與電腦溝通。

二、 研究目的

本研究的目的是開發智慧型虛擬滑鼠之手勢辨識技術，來模擬滑鼠的命令動作，其主要研究的核心技術項目包括手勢分割(Gesture Segmentation)技術、手勢特徵擷取(Feature Extraction)技術與手勢辨識(Gesture Recognition)技術三部分。由於在一般環境下，影像的品質、光線變化、遮蔽效應、物體大小等等，會影響手勢辨識之效能。我們將提出改進方法，克服技術困難點，提出即時且具有高正確性之手勢辨識技術，讓智慧型虛擬滑鼠能運用在人機介面中，可使得人機互動變的更有效率與更人性化。

三、 文獻探討

目前世界國內外許多研究單位都在積極研究開發虛擬滑鼠(Virtual Mouse)相關技術，IBM開發利用電子手套在虛擬環境系統中模擬虛擬滑鼠，根據手勢在三度空間的描述做手勢辨識，來模擬滑鼠動作。美國伊利諾大學黃煦濤教授進行許多手勢分析、追蹤與辨識方面的研究，可應用在人機介面中，如利用轉換學習法做彩色手勢追蹤，如發展可適性自我組織彩色分割演算法用於手勢定位中。MIT 媒體實驗室A. Pentland等教授利用隱藏式馬可夫模型 (Hidden Markov Model) 做美國手語辨識技術，能解決連續影像之時間模型上的問題，達到不錯之辨識效能。

在國內則有清華大學黃仲陵教授利用隱藏式馬可夫模型進行手勢辨識方面的研究。中山大學蔣依吾教授開發虛擬滑鼠技術，利用以電腦視覺為基礎之方法進行手勢辨識之研究，並實作出虛擬滑鼠手勢輸入介面系統應用於電腦操作上，其主要的特色是在於手不需要貼任何顏色標記或不需要帶有顏色之手套或感應裝置。然而，就目前國內外研究之手勢辨識相關技術，尚難同時達到虛擬滑鼠之即時且具有高正確性之要求。

四、 研究方法

智慧型虛擬滑鼠核心技術之手勢辨識工作內容可分為三個部分：手勢分割技術、手勢特徵擷取技術和手勢辨識技術。研究方法敘述如下：

a. 手勢分割技術：

決定適當手勢模型，能有效表示之手勢之特性，將有利於手勢特徵之擷取。我們將採用形狀邊界(Shape Boundary)模型，利用物體實際外形描述手勢，可以在即時的時間擷取手勢之特徵。

由於手指的靈活度高，配合手勢的角度變化多，在一般環境下，如何快速而有效的找出目標物是主要的困難點。傳統上，很多相關的技術提出，其中有一種稱為背景分割法 (Background Subtraction) 的技術可快速找出目標物。背景分割法主要分為二部分，第一部份為背景模組建立階段，第二部分則為手勢分割階段。在背景模組建立階段：去背法主要的概念就是利用統計的方式來區分出影像中前景與背景的物件。首先需拍攝一段背景影像，然後計算所有像素的統計數據，如：平均值、變異數等當作背景的參數。在手勢抽取階段：有了背景參數之後，我們可以將新的影像 (含有前景物體)，輸入系統，計算出每個像素的相關統計數據，並與背景參數加以比較，將變異量大者歸類為前景，變異量小者則歸為背景，快速而有效的找出目標物。然而，在一般環境下，很難找出完整之目標物，容易受雜訊的影響，會影響擷取特徵的正確性。

所以，在此情況下，我們提出讓使用者穿帶有顏色手套的方式解決此問題。穿帶一般的手套並不會對使用者造成很大的影響，所以適合運用在虛擬滑鼠上。因此，我們可利用簡單有效的目標物分割方法，正確的找出目標物。目標物找出後，我們再利用簡單的影像處理技術找出手勢形狀

輪廓，然後存成二值影像(Binary Image)，以便於手勢特徵擷取。

b. 手勢特徵擷取技術：

此部分為分析手勢形狀以便找出具有代表性之手勢特徵參數。在先前的相關技術中，如力距描述子 (Moments Descriptors) 和傅立葉描述子 (Fourier Descriptors)，能支援形狀整體性(Global)之描述，但對形狀局部性(Local)無法有效的描述。因此，在本計畫我們利用曲率格化空間(Curvature Scale Space)形狀描述子描述手勢形狀，因為根據 MPEG-7 範式，曲率格化空間形狀描述子為當形狀經縮放、位移和旋轉後仍保有很穩定的特徵描述。

但是由於手勢形狀是高度形變的，手勢曲率格化空間特徵變化隨著手勢形狀變化而差異極大，先前的特徵校準 (Alignment) 技術會受手勢形變的干擾，影響手勢辨識結果。我們針對手勢形狀的特性，提出可適性(Adaptive)多重手勢曲率格化空間特徵集解決特徵不穩定的問題。由於特徵校準基準點具有不確定性，我們試著利用模糊推理(Fuzzy Reasoning)的技巧找出可能的特徵校準基準的點，模糊推理的概念敘述如下。

模糊集中，元素可部分(Partially)屬於一個集合。令全域(Universal Set)為 X ，一個模糊集合 A 可定義為

$$u_A(x): X \rightarrow [0,1]$$

其中 $u_A(x)$ 為模糊集合 A 的歸屬函數 (Membership Function)。歸屬函數表示元素 x 在全域 X 中屬於模糊集合 A 的程度。

模糊推論系統整合模糊集、模糊若則規則以及模糊推理，對於不確定性 (Uncertainty) 的事件提供一個非常好的推

論系統。模糊推論系統包括三部分：

- (1) 規則庫(Rule Base): 包括一些模糊若則規則的集合。
- (2) 資料庫(Database): 定義用在模糊若則規則中模糊集合之歸屬函數。
- (3) 推理機制(Reasoning Mechanism): 根據給定的輸入和模糊若則規則，用模糊推理推理出結果。

在模糊推論系統中，輸入(Input)的部分可以為模糊輸入(Fuzzy Input)，也可以為明確輸入(Crisp Input)，輸出(Output)的部分通為模糊輸出(Fuzzy Output)。當需要明確輸出(Crisp Output)時，可利用反模糊化(Defuzzifier)把模糊輸出轉為明確輸出。

可能的特徵校準基準的點利用模糊推理找到後，接著利用這些校準基準的點分別校準曲率格化空間特徵而得到多重手勢曲率格化空間特徵集。由於手勢形狀的特性，可推導出最差狀況之校準基準點的數目，因此，特徵集並不會複雜，仍能達到即時性的要求。

c. 手勢辨識技術：

手勢辨識技術主要分為二部分，第一部份為待辨識手勢的儲存部分，第二部分則為欲辨識手勢的辨識部分。在手勢的儲存部分方面，系統還未執行辨識之前，必須讓電腦知道我們所要辨識的手勢，因此須將各種手勢影像收集，然後擷取其手勢特徵，然後儲存在資料庫中，以待辨識。

手勢辨識部分則利用最近鄰居法 (Nearest Neighbor Method) 將手勢特徵資料加以辨認，得到辨識結果進行虛擬滑鼠命令。在此部分中，手勢經過 CCD 拍攝進來是連續的影像，我們擷取其中一張手勢影像，經過手勢形狀分割以及手勢特徵擷取，然後進行可適性多重手勢曲率格化空間特徵集特徵比對 (Feature Matching) 方法，最後再利用最近鄰居法找出正確的結

果，以執行虛擬滑鼠命令。

五、結果與討論

在本研究中，我們完成下列成果。

(一)、提出適用於虛擬滑鼠之手勢分割方法，能快速且正確的找出手勢形狀邊界範圍。(二)、提出可適性多重手勢曲率格化空間特徵集解決特徵不穩定的問題，能擷取穩定且正確的手勢形狀特徵。(三)、提出可適性多重手勢曲率格化空間特徵集之特徵比對方法，能達到即時且具高正確性之辨識手勢。

由於整個計畫執行時間短，我們完成初步之核心技術，但尚未有足夠的時間開發智慧型虛擬滑鼠介面離形系統，希望未來後續研究能完成。不過，我們所開發之核心技術具有相當學術價值，我們將規劃投稿到相關的會議論文與國際期刊。

參考文獻

- [1] J.Y. Chen, "Virtual Mouse : Vision-Based Gesture Recognition," (Supervised by Dr. J. Y. Chiang), *M.S. Thesis*, Department of Computer Science and Engineering, National Sun YetSen University, Kaoshing, Taiwan, R.O.C., 2003.
- [2] E. Gose, R. Johnsonbaugh and S. Jost, *Pattern Recognition and Image Analysis*, Prentice Hall, 1996.
- [3] T.S. Huang and V.I. Pavlovic, "Hand Gesture Modeling, Analysis and Synthesis," in *International Workshop on Automatic Face and Gesture Recognition*, pp. 73-79, 1995.
- [4] F. Mokhtarian and A.K. Mackworth, "Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 1, pp. 34-43, 1986.
- [5] F. Mokhtarian and A.K. Mackworth, "A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 8, pp. 789-805, 1992.
- [6] V.I. Pavlovic, R. Sharma and T.S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 496-513, 1988.
- [7] C.H. Teh and R.T. Chin, "On Image Analysis by the Methods of Moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, pp. 677-695, 1997.
- [8] Y. Wu and T.S. Huang, "View-independent Recognition of Hand Postures," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2000)*, Vol. II, pp.88-94, 2000.
- [9] C.T. Zhan and R.Z. Roskies, "Fourier descriptor for plane closed curves," *IEEE Transactions on Computer*, Vol. 21, No. 3, pp. 269-281, 1972.
- [10] Overview of the MPEG-7 Standard (version 4.0). ISO/IEC JTC1/SC29/WG11 N3753, Oct. 2000.