

行政院國家科學委員會專題研究計畫 成果報告

子計畫四：動態 VR 運動復健輔助系統之人機溝通介面及使用者情緒瞭解(III)

計畫類別：整合型計畫

計畫編號：NSC92-2213-E-009-112-

執行期間：92年08月01日至93年07月31日

執行單位：國立交通大學電機與控制工程學系

計畫主持人：張志永

報告類型：完整報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 10 月 6 日

行政院國家科學委員會專題研究計畫第三年期末報告

總計畫：泛用型動態虛擬實境操控與運動復健輔助系統研發
—子計畫四：動態 VR 運動復健輔助系統之人機溝通介面
及使用者情緒瞭解(III)

(Subproject IV) A Study of Real-Time Human Emotional Understanding and its Mechanism for Intelligent Human-Computer Interface (III)

計畫編號：NSC-92-2213-E-009-112

執行期限：92/08/01–93/07/31

主持人：張志永

交通大學電機與控制工程學系

臉影像處理，人性化溝通管道，抽象式觀念
表示法，聯想記憶，人臉眼鏡的自動去除

一、中文摘要

本子計畫藉電子攝影機隨時紀錄操控者臉部，藉影像處理技術，偵測出操控者之情緒狀態，並建立操控者與模擬器間的人性化溝通管道，以將操控者的使用感受忠實的告知子計畫一與二進行模擬器的控制變數轉換及後續階層式控制法則之設計，本子計畫已發展出一套即時之人情緒狀態判定系統藉線上拍攝操控者臉部影像序列，由影像處理技術偵測出操控者之臉部變化，進而推敲操控者快樂、生氣、與中性情緒狀態，它的效果經過驗證可達 92 % 辨識率；另為更配合國人常有帶眼鏡之情況，今年度特別處理人臉自動眼睛偵測及眼鏡消除，加入已完成之人機系統中，可適合於有或無帶眼鏡使用者。在人機系統溝通與對映方面，我們設計出一套感覺概念交流所須之對映對映網路的裝置來進行，我們以 Kosko 之模糊關聯網路 Fuzzy Associate Memory 設計之，我們以適應性的局部訓練法則及口袋演算法觀念，所設計的二階聯想記憶(Second-Order Associative Memory)，具有最大吸引範圍。

關鍵詞：虛擬實境，人臉表情情識，序列人

英文摘要

The purpose of this project is to construct a human-machine communication interface between the users and the simulators, which can transfer user's feeling sense of to project 1, to transform the control variables for the project 2 control loop design of the simulator of Steward platform. In this way, the performance of the whole control loop and the method of adjusting parameters can be interacted with the user and the reality of the dynamics can be enhanced. Therefore, the project can be viewed as a higher-level feedback unit of the whole simulator system.

VR system can be more genuine if the truly sense of users can be communicated to the dynamic systems. This can be achieved by on-line monitoring the user's facial emotion and transferring the user's feeling to the machine and updating the appropriate control parameters. The feeling of human being is an abstract

concept and there is no corresponding physical quantity, i.e.,

universe of discourse, for the membership function to define with. Research on how to extract the user facial emotion, represent the abstract concept and devise a method that can communicate with machine simulator constitutes the major concerns to be investigated in this project.

Keywords: Virtual Reality, Facial Emotion Recognition, Facial Image Sequence Processing, Abstract Concept Representation, Associative Memory

二、緣由與目的

動態模擬系統之動態感覺，決定了虛擬實境之逼真程度，但如何表示動態模擬系統使用者之感覺概念，並傳遞給行為轉換與控制模組(子計畫一)及六軸運動平台、力回饋模組(子計畫二、三)做適當的參數調整與控制，是本計畫之重點。為了表示使用者的感覺概念，道，以將操控者的使用感受忠實的告知子計畫一與二進行模擬器的控制變數轉換及後續階層式控制法則之設計，並藉此以改善整個控制迴路及調整參數的方法，而提高了模擬器的真實性，本計畫可視為整個模擬器系統的高階回授單元。本計畫之兩個主要子題：計畫將發展出一套即時之人情緒與注意力狀態判定系統，及能與機器系統做抽象式觀念感覺溝通與對映之方法，下面是我們在這一年的計畫中，對以上兩方面所法提出來實現方法的描述。

三、研究方法與成果

(1) 人類表情與注意力的辨識

(1a) 人類表情與情緒的辨識

人類表情與情緒的辨識此子系統第一部份為臉部特徵的選取。臉部特徵的選取對於辨識人臉和臉部表情是很重要的。人類臉部有些明顯的特徵如眼睛、嘴巴、眉毛、臉部輪廓等等[3-6]。在本計畫中，我們擷取眉毛、眼睛及嘴巴作為主要的特徵。我們定義 30 個臉部特徵點；來描述這三個臉部特徵的位置和形狀。臉部快樂、生氣、與中性情緒狀態表情我們藉由一放射狀函數基礎之神經網路辨識，它的效果經過驗證可達 92% 辨識率，此部份已在第一年完成與報告。

(1b) 人類注意力的辨識

在人機系統溝通與對映方面，我們正在進行以觀察臉部影像序列，偵測使用者之感覺與精神狀態[11-13]，精神不濟或昏昏欲睡時，機器會提出暫停機台操作之警告，以策安全。我們以眼睛縮小閉合時間比例(PERCLOS)及眨眼頻率(Blink Duration)兩種特徵參數，偵測精神不濟或昏昏欲睡之程度，我們再利用模糊積分(fuzzy integral)的概念[10]，發展出上述兩種證據跡象資訊整合的技術，以增加偵測的可靠度。經測試偵測精神不濟或昏昏欲睡之正確率可達 90%。

(2) 人機系統溝通與對映法

人機系統溝通與對映之法 Takagi 等人係利用關聯網路 (Associate Memory; 簡稱 AM) 設計之赫比式(Hebbian)學習方法來建構之對映傳遞之關係[1, 2]，但 AM 之兩對映是由(0, 1)之串列對映至(0, 1)之串列，與上述之感覺對應，為 0~1 間的任意實數對映，我們以某一層的觀念節點的節點活性度值組表示一個感覺概念，以機電整合受控體之物理量為輸入層，感覺概念的溝通與學習可由函數大約化之對映方法來執行感覺概念之交流，以使電

腦有感覺概念意識與溝通的能力，而對於觀念節點溝通所需之函數大約化，我們設計出一套對映網路的裝置來進行，亦即在受控機器中建立並學習某一感覺概念交流所須之對映網路。第一年我們以 Kosko 之模糊關聯網路 Fuzzy Associate Memory 設計，以二階非對稱式雙向聯想記憶(Second-Order Asymmetric Bidirectional Associative Memory)實現；第二年我們以傳統二階對稱式雙向聯想記憶(Second-Order Bidirectional Associative Memory)實現，並訓練儲存圖樣能有最大的吸引範圍，已經證明出我們所提出的局部訓練法則效果優於其他主要的 AM 的設計 [7-9]。

(3) 戴眼鏡使用者自動眼睛偵測及眼鏡去除方法

眼睛偵測[16-18]在各種應用上扮演一個很重要的腳色，例如：人類臉部辨識[14,15]、眼睛注視的方向、昏睡偵測和其他人類與電腦相互關係的應用。雖然已經有很多方法提出來解決這些偵測的問題，但似乎遺漏人們戴眼鏡的問題。再加上東方人，尤其是華人，有很多戴眼鏡者。因此處理戴眼鏡的問題[19-22]是相當迫切且需要的。

在這些應用上，我們今年特別針對眼鏡對精神不濟或昏睡偵測系統與臉部辨識系統所造成的干擾問題來做處理。在昏睡偵測系統中，我們利用影像處理的技術來測量眼睛的閉合並且計算 PERCLOS 與眨眼頻率。然而當一個駕駛員戴著眼鏡時，因為眼鏡鏡框與眼睛重疊，造成系統無法正確的偵測到眼睛的位置。而且戴眼鏡是最常阻礙到一個經由臉部辨識系統來做個人識別的問題。我們所提出的自動眼鏡去除方法有助於正確地偵測到眼睛的位置並且消

除這些的阻礙。以下為自動眼睛偵測及眼鏡去除方法之說明。

(3a) 人類表情與情緒的辨識

在一張影像中搜尋人的位置是辨識所有人類相關動作的第一個步驟。這裡我們是採取 Chai *et al.*[14]所提出的方法。此方法是有以下四個步驟。

步驟A：顏色分割。此步驟的目的是為了將輸入影像的像素值分類成皮膚與非皮膚兩個區域。在 YCrCb 彩色空間中，我們定義且測試出 $R_{C_r} = [133, 173]$ 與 $R_{C_b} = [77, 127]$ 分別為 Cr 和 Cb 所對應皮膚顏色的範圍。當(x,y)的 Cr 與 Cb 的值屬於 $R_{C_r} = [133, 173]$ 而且也屬於 $R_{C_b} = [77, 127]$ 的範圍，則分割結果 O_A 等於 1，其他則等於 0，其中 $x = 0, \dots, M/2-1, y = 0, \dots, N/2-1$ 而 M 與 N 分別影像的高度都寬度。**步驟B：強度規則化。**此步驟的目的是為了考量在步驟 A 中所產生的雜訊干擾，這些雜訊可能因為一些未偵測到的臉部特徵，像是眼睛、嘴巴，而造成在臉部區域出現一些小的空洞，或是背景中有著與皮膚顏色相同的物體所產生的雜訊。因此我們利用形態學中的膨脹來填補臉部的空洞，利用侵蝕去除背景中的物體所造成的雜訊。**步驟C：幾何修正。**此步驟的目的是利用垂直與水平掃描來消除 $O_B(x, y)$ 中零星分佈的結構。**步驟D：輪廓抽取。**最後一個步驟的目的是為了從步驟 C 中所得到的輸出圖，利用步驟 A 的邊緣資訊，將所偵測到臉部範圍還原出來。經過上述的四個步驟處理後，即可正確且有效的將臉部的輪廓擷取出來。圖一為測試某人臉部影像經過上述的四個步驟所得到的結果。

(3b) 眼睛偵測與眼鏡去除

眼睛的偵測

在參考論文[19]中，其方法是找到一些介於兩眼間的“between-eyes”而不是直接去偵測眼睛的位置。在“between-eyes”左右兩側包含了眼睛與眉毛，故此處是屬於比較黑暗的部分，相反的在其上下兩側包含了額頭跟鼻樑，故此處是屬於比較明亮的部分。此篇論文是利用一個 circle-frequency filter 來找到“between-eyes”的位置。假設以 (x, y) 為圓心來作一個圓，其圓上的素 $\|F_n\|^2$ 稱為 circle-frequency-n filter 的輸出。根據傅立葉轉換的特性，可知當 $n=2$ 時，會產生一個高值，稱為 circle-frequency filter 或是 CF-filter。接著再利用以下的兩個條件，來搜尋“between-eyes”的位置。

眼鏡的去除

眼鏡的鼻樑架是最常見的眼鏡特徵，藉由“between-eyes”與眼睛位置的資訊來找到“between-eyes”周圍的區域與兩眼中心的位置，可以從邊界存在的範圍裡找到眼鏡的鼻樑架。再利用梯度邊緣偵測的方法，我們可以在每一點 (x, y) 求出梯度的大小與相位，若有一個垂直方向的像素個數的數目超出一個固定的臨界值，則可判定眼鏡存在。

如果眼鏡存在，我們必須去除眼鏡的干擾進而找到眼睛的位置並量測眼睛的閉合狀況，再來應用在昏睡偵測系統上。首先我們必須計算出眼睛的範圍，這樣才能去除眼鏡所造成的干擾。基於這樣的論點，我們將 HSI 與 RGB 彩色空間所作的梯度邊緣偵測資訊以及在 YCrCb 彩色空間所偵測到非皮膚顏色的區域作結合。我們再對戴眼鏡的邊緣圖作侵蝕的動

作，則此邊緣圖將會被分成很小的片段。之後藉由計算每一個部分之間的標準誤差，選取最小的標準誤差且是最大相連的部分，由此部分即可輕易的將眼睛從眼鏡的輪廓中區分出來。接下來，我們將眼睛部分的像素值以皮膚顏色的像素值來取代。為了使還原的更自然，我們利用 modified fuzzy rule-base (MFRB) filter [23, 24] 將眼鏡的像素值作平滑的動作並且使其更接近沒戴眼鏡的影像。圖二為測試某人臉部眼鏡去除，經過與沒經過 MFRB filter 步驟所得到的影像比較結果；圖三為測試某人眼鏡去除，影像經過上述的各個步驟所得到的結果，由圖二、三的結果可看出所提出的眼鏡去除方法良好，具體可行 [25]。

四、結論與討論

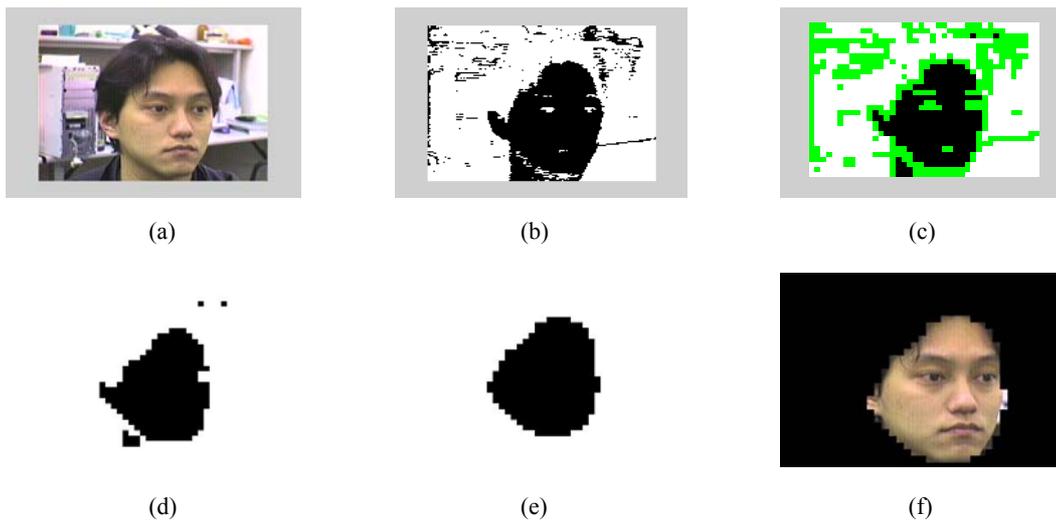
一理想的虛擬實境系統，應能隨時掌握操控者的情況與感受，並能讓操控者與系統模擬器做直覺性的溝通，而讓使用者有完全沈浸其中的感受，因此研究發展任何虛擬實境技術時，使用者與機器系統之間的互動與溝通為相當重要的一環。本子計畫藉電子攝影機隨時紀錄操控者臉部，藉影像處理技術偵測出操控者之情緒狀態，並建立操控者與模擬器間的人性化溝通管道，以將操控者的使用感受忠實的告知子計畫一與二進行模擬器的控制變數轉換及後續階層式控制法則之設計，並藉此以改善整個控制迴路及調整參數的方法，而提高了模擬器的真實性，本計畫可視為整個模擬器系統的高階回授單元。本子計畫已發展出一套即時之人情緒狀態判定系統藉線上拍攝操控者臉部影像序列，由影像處理技術偵測出操控者之臉部變化，進而推敲操控者快樂、生氣、與中性情緒狀態，它的效果經過驗證可達 92% 辨識率。另外，我們以眼睛縮小閉合時間及眨眼頻率兩種特徵參數，偵測精神不濟或昏昏欲睡之程度，本偵測系統提出非接觸式之使用者是否

昏睡及精神不濟的判定技術，可增加單一特徵參數正確率 12%，整合正確率達 95%。在人機系統溝通與對映方面，我們以模糊 SOAM 設計，實現感覺概念交流所須之對映網路，模擬證明我們所提出的局部訓練法則效果優於其他主要的 AM 設計。另為更配合國人常有帶眼鏡之情況，今年度特別處理人臉自動眼睛偵測及眼鏡消除，加入已完成之人機系統中，可適合於有或無帶眼鏡使用者。

五、參考文獻

- [1] H. Ushida, and T. Yamaguchi, and T. Takagi, "Fuzzy associate memory based knowledge construction with a application to a human machine interface," *IEEE Trans. Ind. Electron.*, vol. 46, pp. 851-860, 1999.
- [2] T. Takagi, A. Imura, H. Ushida, and T. Yamaguchi, "Conceptual fuzzy sets as a meaning representation and their inductive construction," *International Journal of Intelligent Systems*, vol. 10, pp. 929-945, 1995.
- [3] P. Ekman, and W. V. Friesen, *The Facial Action Coding System*, Consulting Psychologist Press, San Francisco, CA, 1978.
- [4] H. Kobayashi, and F. Hara, "Analysis of the neural network recognition characteristics of six basic facial expressions," in *Proceedings of 3rd IEEE Int. Workshop on Robot and Human Communication*, Nagoya, Japan, pp. 222-227, 1994.
- [5] M. A. Turk A. P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [6] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, Vol. 8, No. 2, pp. 99-111, 1992.
- [7] T. Wang, X. Zhuang, and X. Xing, "Designing bidirectional associative memories with optimal stability," *IEEE Trans. on Syst. Man, Cybern.*, vol. 24, pp. 778-790, May. 1994.
- [8] C.-S. Leung, L.-W. Chan, and E. Lai, "Stability and statistical properties of second order bidirectional associative memory," *IEEE Trans. on Neural Networks*, vol. 8, pp. 267-277, Mar. 1997.
- [9] H. Shi, Y. Zhao, and X. Zhuang, "A general model for bidirectional associative memories," *IEEE Trans. on Syst. Man, Cybern. B*, vol. 28, pp. 511-519, Aug. 1998.
- [10] S. B. Cho and J. H. Kim, "Combining multiple neural networks by fuzzy integral for robust classification," *IEEE Trans. Syst., Man, Cybern. B*, vol. 25, no. 1, pp 113-120, 1995.
- [11] K. Ogawa and M. Shimotani, "Drowsiness detection system," *Technical Report of Mitshbish Electric*, Mar. 1997.
- [12] H. Sako and A. V. W. Smith, "Real-time facial expression recognition based on features' positions and dimensions," in *Proc. IEEE Int. Pattern Recognition Conf.*, 1996.
- [13] C. D. Wylie *et al.*, "Commercial motor vehicle driver fatigue and alertness study," *Project Report* (Report No.

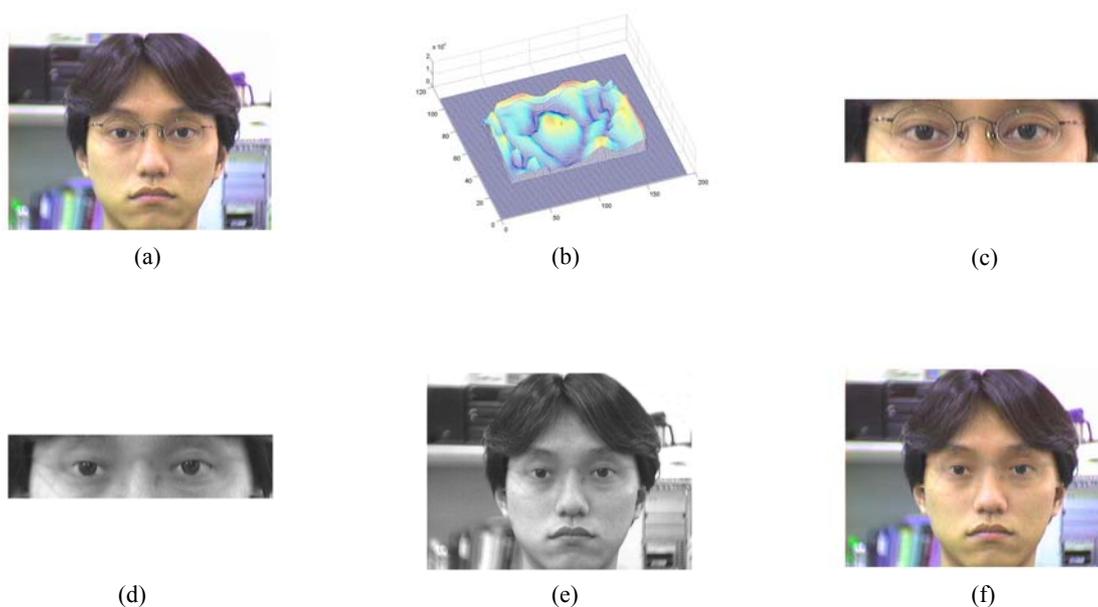
- FHWAMC-97-002), Washington, D. C: Federal Highway Administration Office of Motor Carriers, Oct. 1996.
- [14] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 551–564, 1999.
- [15] R. L. Hsu, M. A. Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 696–706, 2002.
- [16] S. A. Sirohey and A. Rosenfeld, "Eye detection in a face image using linear and nonlinear filters," *Pattern Recognition*, vol. 34, pp. 1367–1391, 2001.
- [17] G. C. Feng and P. C. Yuen, "Multi-cues eye detection on gray intensity image," *Pattern Recognition*, vol. 34, pp. 1033–1046, 2001.
- [18] R. Thilak Kumar, S. Kumar Raja, and A. G. Ramakrishnan, "Eye detection using color cues and projection functions," in *Proc. IEEE Int. Conf. Image Processing*, 2002, vol. 3, pp. 24–28.
- [19] S. Kawato and J. Ohya, "Two-step approach for real-time eye tracking with a new filtering technique," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2000, vol. 2, pp. 1366–1371.
- [20] X. Jiang, M. Binkert, B. Achermann, and H. Bunke, "Towards detection of glasses in facial images," in *Proc. 14th Int. Conf. Pattern Recognition*, 1998, vol. 2, pp. 1071–1073.
- [21] Z. Jing and R. Mariani, "Glasses detection and extraction by deformable contour," in *Proc. 15th Int. Conf. Pattern Recognition*, 2000, vol. 2, pp. 933–936.
- [22] C. Wu, C. Liu, H. Y. Shum, Y. Q. Xu, and Z. Zhang, "Automatic eyeglasses removal from face images," in *Proc. 5th Asia Conf. Computer Vision*, 2002, pp. 23–25.
- [23] J. Y. Chang and S. M. Lu, "Image blocking artifact suppression by the modified fuzzy rule-based filter," in *Proc. IEEE Int. Conf. Syst., Man., Cybern.*, 2003, vol. 1, pp. 486–491.
- [24] K. Arakawa, "Fuzzy rule-based signal processing and its application to image restoration," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 1495–1502, Dec. 1994.
- [25] S. W. Wang, "Automatic Eye Detection and Glasses Removal," Master Thesis, Department of Electrical and Control Engineering, National Chiao Tung University, Taiwan, 2004.



圖一. Example of face segmentation algorithm. (a) Input image. (b) Result of Stage A. (c) Density map in Stage B. (d) Result of Stage B. (e) Result of Stage C. (f) Result of Stage D.



圖二. Example which has removed glasses with and without MFRB filtering. (a) Without MFRB filtering. (b) With MFRB filtering.



圖三. Example of eye detection and glasses removal system. (a) Original image. (b) CF-filtered image of face region. (c) Detected eye region. (d) Glasses removal with MFRB filtered in gray level image. (e) Attaching the eye region to (a) in gray level. (f) The result of chromatic level.