

人臉追蹤與疲勞偵測系統

Face Tracking and Tiredness Detection System

白能勝

Neng-Sheng Pai
勤益科技大學
電機工程系
Pai@ncut.edu.tw

張世平

Shih-Ping Chang
勤益科技大學
電機工程系
s49812103
@student.ncut.edu.tw

郭逸平

Yi-Pin Kuo
遠東科技大學
電子工程系
ypkuo@cc.feu.edu.tw

黃聖富

Sheng-Fu Huang
勤益科技大學
電機工程系
s39312020@
student.ncut.edu.tw

摘要

研究內容由兩個主要部份所組成，分別為人臉偵測(face detection)、特徵擷取(feature extraction)，在人臉偵測部份，主要利用膚色偵測等影像處理技術將臉部區塊，從複雜背景中取出。在特徵擷取部份，主要利用唇色偵測、邊緣偵測、眼睛偵測等影像處理技術從臉部區塊中取出嘴巴、眼睛等特徵。在本研究中，建立了一套「疲勞偵測警示裝置」，此系統是透過攝影裝置擷取影像，再利用影像處理技術與疲勞判別技術來監視使用者的精神狀況，偵測其是否屬於清醒狀態。

關鍵字：人臉偵測、頭部追蹤、影像處理

1. 前言

影像處理是近來熱門的話題，由於電腦技術的進步，以往只能以筆紙推算的理論，現在都可以在電腦上逐一實現了，加上人類使用上的需求，影像處理的應用也就越來越廣泛，在網路發達的時代，聲音已不能符合人類的需求，再加上影像才会有真實感。臉部表情是我們人類用來表達喜怒哀樂複雜的心情，而最基本的臉部也可以作為判斷的依據，如點頭搖頭等，我們可以利用這樣的資訊來做一些簡單的定義，做為人機介面與電腦的溝通，這樣的設計可以將電腦的控制更加人性化與便利性。鑒於目前辨識技術日新月異，辨識技術不停有新的突破，除了臉紋判別還有眼紋判別、指紋判別、聲音判別等等，但是每一種判別都容易受各種不同的因素所干擾而影響辨識的準確率。而在那麼多種辨識技術下。本研究利用眼部來判斷使用者的疲乏程度，將外在環境的干擾除去，以眼部辨別為基礎，再加上頭部的點頭動作，作出準確的判斷。

近十幾年來，眾多研究人員在人臉偵測及人臉辨識方面投入了許多心力，經由不斷的技術突破跟經驗的累積，人臉偵測相關的研究已逐漸成熟；然而在人臉辨識方面，由於特徵選取以及資料庫處理上的限制，目前大多只應用在特定場所針對少數人進行辨識。在人臉偵測方面，Pentland et al.[10]提

出主要分量分析(principal component analysis, PCA)，使用較少維度的特徵去描述與比對人臉。Sung 與 Poggio[11]是評估輸入影像跟“人臉”與“非人臉”兩個群集之間的距離來確認是否為人臉。Rowley et al.[12]使用一個搜尋視窗在多重解析度影像中移動，然後將視窗影像輸入類神經網路確認。Jeng et al.[13]使用人臉的幾何樣板找出複雜影像中人臉的位置，以上研究都是採用灰階影像。除了灰階影像以外，Sobottka 與 Pitas[14]，Chen et al.[15]都嘗試從彩色影像中分離膚色區域來縮小人臉的搜尋範圍。而近年來還有許多著名的方法，如PCA(principle component analysis) [16][17]以及SVM(support vector machine) [18][19][20]。但這些方法需要大量的訓練樣本，包含著每個人各種角度的訓練樣本，因此用來辨識大量的人物時，需要的樣本數將相當地驚人。

往常要驗證或證明一個人的身分，通常會使用類似身份證之類的證件來提供辨識，於是便造成了每個人身上的各種證件琳琅滿目。它們一個共同的目標就是要證明一個人的身分，而現今各種證件竊取及偽造事件頻繁，證件本身的安全性及可靠性就愈來愈受質疑，因此利用每個人身上容易取得的生物特徵。隨著電子科技的快速發展，生物特徵相關的研究逐漸被重視，人臉、指紋與虹膜都是很好的生物特徵，而“人臉特徵”在生物認證中占了重要的一角，因此人臉偵測與辨識的相關研究也開始蓬勃起來。近年來電腦運算能力與取像設備快速地進步，且彩色影像所能提供用來做特徵抽取的資訊較灰階影像多，故相關的研究多已從灰階影像轉而使用彩色影像來進行偵測及辨識。儘管如此，人臉偵測與辨識還是存在許多問題需要解決，如：人臉的角度、人臉的大小、背景及光源的影響．．．等。

2. 色彩辨識理論

我們所看到的色彩(color)基本上是眼睛察覺到光所產生的反應。因此色彩是光的一種特性，故可以用物理方法來表示，同時色彩一是人類視覺的一

種感受，由於不同人的視覺感受不同，所以對於相同物理量，不同的人會有不同的色彩感受。

2.1 色彩空間

要找出膚色區域必須先定義什麼是皮膚顏色，首先要選擇一種合適的色彩空間，這樣才能用色彩參數將皮膚顏色具體化描述。在過去的許多研究中，曾經用過的色彩空間有RGB、HSV、YCbCr、YIQ、HSL等，由於環境光線變化對膚色的影響很大，通常會選擇能夠從色彩中分離出亮度因素的色彩空間，其中最常見的是HSV與YCbCr 色彩空間。

2.2 RGB 色彩空間

在RGB色彩空間中，每種色彩是以其紅、綠、藍的主要頻譜成分來顯現。此空間是建立在直角座標系統的基礎上。其中RGB值是在三個頂點上；青色、紫紅和黃色是在另外三個頂點上；黑色在原點上；而白色是在離頂點最遠的頂點上。在此模型中，灰階(有相同RGB值的點)是從黑色到白色沿著連接這兩個點之間的線。此模型中的不同色彩是位於立方體上或其內部，且用從原點出發的向量來延伸定義，為了方便我們假定所有色彩值已經被正規化，也就是說，所有RGB的值都假定在[0,1]的範圍內。

以RGB色彩空間表示的影像是由三個分量影像所組成，每個原色對應一個分量影像。當送入RGB顯示器時，這三幅影像在螢幕上組合起來產生一幅複合的彩色影像。用來表示在RGB空間每個像素所用的位元數稱為像素深度(pixel depth)。考慮一個RGB影像，其中的紅色、綠色和藍色影像都是一個8位元的影像，在這種情況下，每個RGB彩色像素(亦即三個為一組的(R,G,B)值)有24位元的深度(每個平面位元數乘以3個影像平面)，全彩(full-color)影像這個術語通常用來表示24位元的RGB彩色影像。一個24位元RGB影像的總色彩數是 $(2^8)^3=16777216$ ，圖1為RGB色彩模型結構。

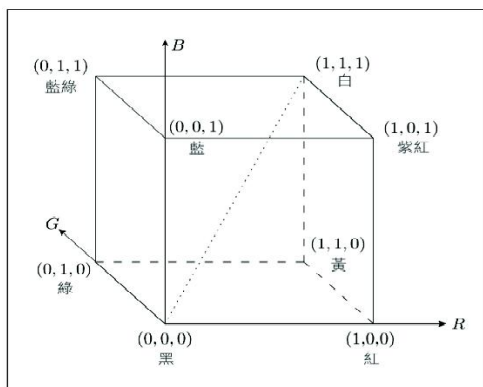


圖1. RGB色彩模型色彩立方體

2.3 YCbCr 色彩空間

YCbCr 色彩空間是修改YUV 色彩空間所得到的，分為亮度元素Y與Cb、Cr兩個彩度元素，對亮度的分離性高，方便與彩度分開操作，適合影像處理使用。

YCbCr色彩空間與RGB色彩空間的轉換關係如下：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

在此使用YCbCr 色彩空間，YCbCr 色彩空間對於光線變化的反應較不敏感，而且對於不同人的膚色也有較好的收斂性。

2.4 膚色區域偵測

採用YCbCr 色彩空間作為膚色區域的判斷，其原因為YCbCr 色彩空間對於光線變化的反應比HSV色彩空間不敏感，且對於不同人的膚色也有較好的收斂性。在文獻[21]中即針對YCbCr 及HSV的色彩空間對膚色分割作過詳盡之探討，膚色成份在YCbCr 及HSV 色彩空間中的分布分別如圖2所示

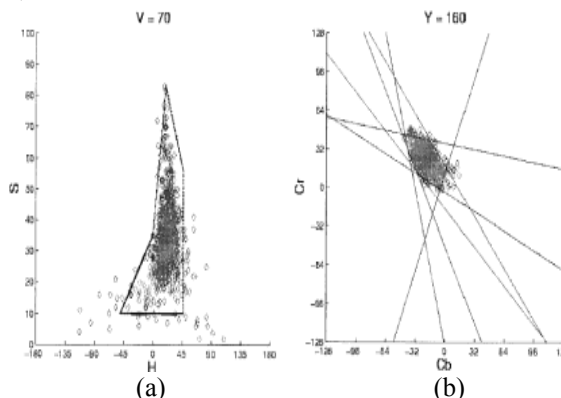


圖2. 膚色在HSV色彩空間(a)及YCbCr色彩空間(b)的2D圖

由圖2可以看出，膚色部分在YCbCr 色彩空間中的分布比在HSV色彩空間中的分布較為集中，容易將膚色部分與非膚色部分區分開來。

本研究即是採用此種膚色判斷模式：

$$\begin{cases} R > 80 \\ 20 < (R - G) < 46 \\ 38 < (R - B) < 93 \end{cases}$$

R表紅色、G表綠色、B表藍色，圖三為膚值偵測的結果。

$$\begin{cases} Y \geq 60 \ \& \& \ Y \leq 250 \\ Cr \geq 135 \ \& \& \ Cr \leq 170 \\ Cb \geq 90 \ \& \& \ Cb \leq 135 \end{cases}$$



圖3. 膚值偵測結果

3. 形狀辨識理論

3.1 二值化影像

二值化(binary)又稱灰階分割(Threshold)，是將一張影像(值介於0~255)轉換成彩色影像只有兩種顏色值的影像，其目的是將複雜影像簡單化以利接下來處理，圖4為未二值化處理的灰度圖，圖5為經二值化處理的灰度圖，可以清楚看出兩者的差異，則圖6是把膚值偵測的結果做二值化。

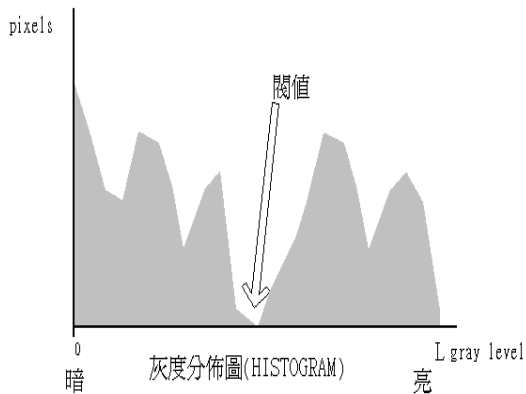


圖4. 未二值化處理的灰度分布圖

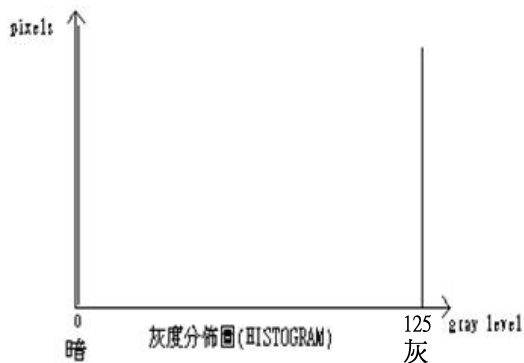


圖5. 經二值化處理的灰度分布圖

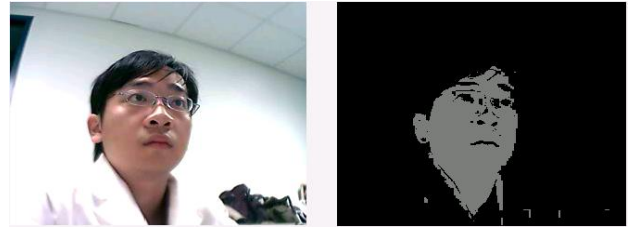


圖6. 將膚值取二值化結果

3.2 形態學處理

型態學(morphology)是影像處理中常應用到之技術，是抽取對表示和描述區域形狀有用之影像分量的一種工具。型態學提供一個唯一和有效的方法處理許多影像處理問題，如影像分割(Segmentation)，邊緣偵測(Edge detection)，凸形封包(convex hull)，細線化(Thinning)及骨架抽取(Skeletonizing)等皆是應用型態學所能達到之功能，圖7(a)為A集合(b)為B集合。以下介紹運用到運算。

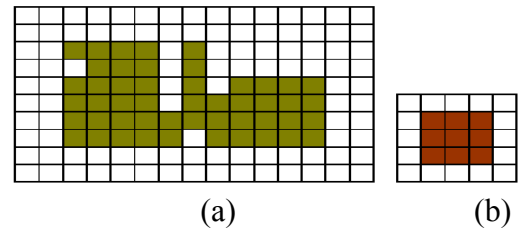


圖7. (a)為A集合(b)為b集合

3.2 侵蝕(Erosion)跟膨脹(Dilation)

侵蝕定義為 $A \ominus B = \bigcap_{b \in B} (A)_{-b}$ ，意思是將A集合向內平移B集合中每個元素的距離，然後取這些集合的交集，所以侵蝕會使A集合縮小，且縮小的範圍是由B集合之結構來決定，侵蝕的使用之一是除去二元影像裡不相干的細節。

膨脹定義為 $A \oplus B = \bigcup_{b \in B} (A)_b$ ，意思是將A集合向外平移B集合中每個元素的距離，然後把這些集合聯集起來，所以膨脹會使A集合擴大，且擴大的範圍是由B集合之結構來決定，膨脹的應用之一是在橋接縫隙，圖8.為Erosion跟Dilation。

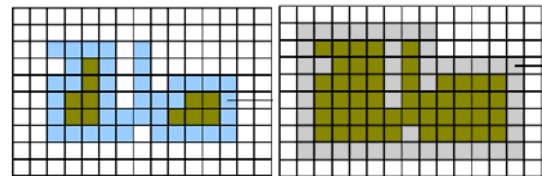


圖8. Erosion跟Dilation

3.4 斷開運算(OPENING)跟閉合運算(CLOSING)

斷開通常是平滑影像輪廓，截斷窄的細頸，消除細的突支。閉合也傾向於平滑輪廓部分，不過與

斷開相反，它把窄的中斷部分和長細缺口連接起來，消除小洞，填補輪廓上的缺口，圖9跟圖10為Open運算。



圖9. 侵蝕後做膨脹為斷開運算

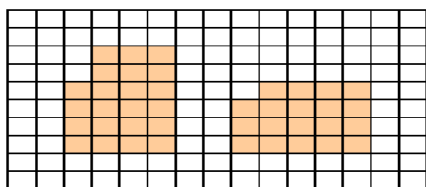


圖10. Open

斷開運算之基本操作是先對影像做侵蝕運算，再對影像做膨脹運算，定義為 $A \circ B = (A \ominus B) \oplus B$ ，其功用在於除去影像中小的亮細節(前景點)，而其他較大的亮度區域則不受干擾。閉合運算則和斷開運算恰好相反，先對影像做膨脹運算，再對影像做侵蝕運算，定義為 $A \cdot B = (A \oplus B) \ominus B$ ，可以去除小的暗細節，而亮細節及其他較大的暗細節區域則不受影響。因此影像先作侵蝕運算再作膨脹運算之結果和先作膨脹運算在作侵蝕運算之結果是完全不同，圖11跟圖12為Close運算，圖13為二值化的結果作Erosion的結果，圖14為圖13做Dilation的結果。



圖11. 膨脹後做侵蝕為閉合運算

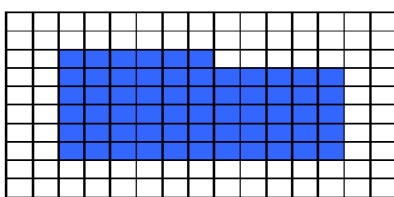


圖12. Close

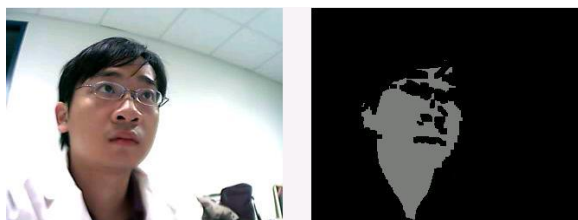


圖13. Erosion的結果



圖14. 對圖13做Dilation的結果

3.5 中值濾波

經過侵蝕膨脹把遠離人臉的雜訊去掉之後，在經由中值濾波把人臉內的雜訊給去除掉以獲得完整的二值化人臉，圖15為中值濾波流程圖。

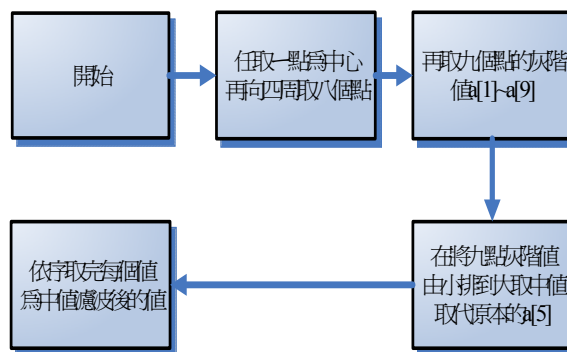


圖15. 中值濾波流程圖

以圖10為例，將數值取出並將其排序後為0,0,0,1,1,1,1,1,1，並把九宮格正中位置以中間值(1)取代，圖16為舉例的3*3空間。

0	1	1
0	0	1
1	1	1

圖16. 3 * 3的遮罩

3.6 邊緣偵測

邊緣偵測的方法有許多種，最常用的方法是利用遮罩的方式對整張影像進行銳化，處理後的影像凡屬於邊界區域會成為亮區，其它則變為暗區；如此便可得到物體邊緣資訊，如 Sobel 梯度運算子。我們可以善加利用影像中背景為黑色及前景為灰階的結果，直接搜尋像素點的四鄰點位置，若有發生從黑色變為灰階或從灰階變為黑色的情況，便將此像素點記錄，其為邊緣點，即可快速求得所要的目標物體邊緣，圖17為邊緣偵測結果。



圖17. 邊緣偵測的影像

4. 研究目標

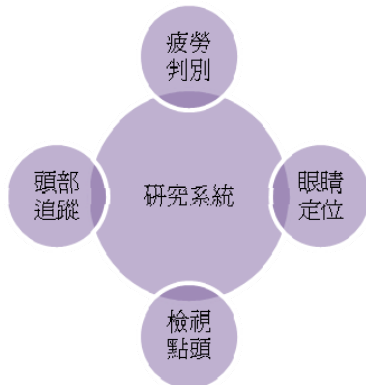


圖 18. 系統架構

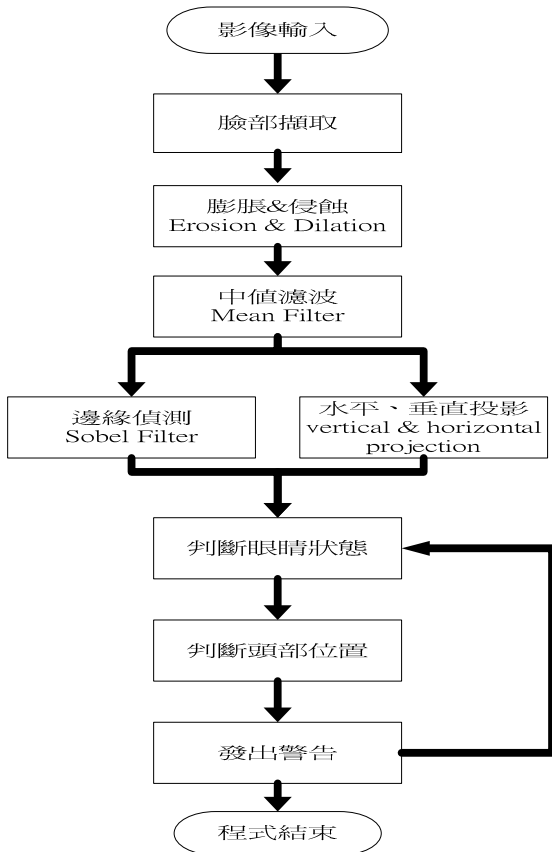


圖 19. 系統流程圖

4.1 頭部追蹤

在人臉追蹤系統中，適應性對實際情況運作時很重要，適應性就是指系統在不同的環境之下〈如

目標、環境的變化〉，不會受其影響，還可以正常操作，所以在設計時，以下的因素就必須要考慮到：

A、背景的濾除

背景對於人臉追蹤而言就是雜訊，所以必須要作影像的前置處理，盡量把背景去掉，對於攝影機靜止不動時，較好的方式就是將兩張連續影像相減，把沒有在移動的背景去除，但對於移動式攝影機就不能使用這個方法，因為背景和主體對攝影機來說都在移動，無法分辨出靜止的背景和移動的物體。

B、人臉的大小

人臉在影像中的大小，會因位置及方向的不同而有大小的變化，追蹤系統必須在人臉變化有一定大小範圍內才能夠追蹤。

C、亮度變化

由於周圍環境的亮度不同，會使得影像當中的顏色深淺產生變化，這種情形會直接影響到人臉捕捉的準確性，所以亮度參數的部分就必須要考慮到人臉追蹤，這也是我們所遇到的困難。

4.2 眼睛定位

將膚色的範圍做填充以及適度的侵蝕之後再與以灰階化做閾值二值化留下的範圍做交集，可獲得臉部範圍中為黑色的部份，在以此範圍取水平與垂直投影可獲得兩眼的座標位置，並且以水平或垂直投影量來計算是否為睜眼或是閉眼的判斷。

4.3 檢視點頭

是將頭部追蹤的影像，以邊緣偵測的方式處理每一張影像，累計凡是255為一個投影的水平像素。處理完畢後，如果頭往下點，垂直投影會往下聚集，如果頭往上抬，垂直投影會往上聚集，移動軌跡就會出現。

4.4 疲勞偵測

先獲得攝影機所獲得的畫面中膚色是否有站一定的比例（膚色像素數大於六分之一及小於三分之一）來判定是人臉在適當的距離內，若是在此範圍內在本系統判別出使用者的睜眼與閉眼情況後，再加上是否點頭的動作，進行疲勞判別處理，若使用者出現連續閉眼或點頭頻繁且超過預設門檻值範圍的情況，則本系統判定使用者已呈現疲勞的狀態；反之，若使用者未出現連續閉眼或雖閉眼但未超過預設門檻值範圍的情況，則本系統判定使用者未呈現疲勞狀態。

5. 設備與開發環境介紹

5.1 設備：

5.1.1 網路攝影機

產品規格：畫質 9000K Pixels，動態影像 3840×2880 30fps Interpolation，感光處理 CMOS Sensor，傳輸介面 USB1.1/USB2.0 Gold coated USB interface，鏡頭 5-layer Optical Glass Len，焦距 36mm to infinity，視訊功能 自動白平衡、自動曝光、夜間、背光增強模式、色彩色相校正，軟體功能 拍照、錄影、特效、相框、縮放，硬體功能 內建六燈紅外線、麥克風，圖 20 為攝影機外貌。



圖 20 攝影機

5.1.2 桌上型電腦：

電腦規格：處理器 Intel(R) Celeron(R) CPU 3.06GHz，顯示卡 ASUS P5VDC-TVM，記憶體 512MB，硬碟 HDT72251，光碟機 ASUS DVD-E16A2，電源供應器 ASUS S-30FP。

5.2 開發環境：

作業系統為 Windows XP，Borland C++Builder6 程式設計，圖 21 為軟體架構。



圖 21. 軟體架構

6. 研究成果



圖 22. 頭部偵測結果



圖 23. 點頭偵測結果



圖 24. 換有背景的头面部偵測結果

針對這次研究的過程，分有幾個部分說明：一開始要開發在 Altera DE2 實驗版上，經過不斷的摸索、討論跟研究，發現無法繼續進行，因為 DE2 編譯時間過長，記憶體不夠，處理速度太慢，導致無法成功的將人類頭部與疲勞程度辨識出來，所以先使用 Borland C++ 代替開發。

在 Borland C++ 去除雜訊也是花了很多心力，因為在膚質偵測的結果轉化成二值影像後，會發現

人臉周圍會有類似膚色的小雜點，而這些雜點會影響到後面的偵測，而且如果能得到一個完整的臉型，那將在判斷上有很大的幫助。頭部追蹤也很不容易，用了大量的運算，及時判斷出臉部的位置，並將位置框出，眼睛方面，雖然是很好的特徵，但是在亮度方面還不穩定，所以是需要加強的，圖 22 為頭部偵測左圖為原圖、右圖為頭部偵測結果。圖 23 為點頭偵測結果，圖 24 是換成有背景的頭部偵測。

7. 結論

本系統可以二十四小時持續監控，完全不需要休息，可以用在行車、學校、電腦前、辦公室等地方，可以有效降低車禍率，提升學生學習及上班族工作的效率。在點頭方面，因為點頭是人類的行為之一，可用來表示人類意思，研究點頭辨識將有助於人機介面互動的開發，使電腦認識人類點頭行為，也可運用於簡單的資料選擇，例如是和否的選擇。將來可以把整個運算交於硬體來獨立執行取代多工電腦，讓本研究的技術可以產品化與生活化，更加簡化讓一般民眾更易於操作，整合其他生物辨識技術，讓他的應用更加多元更具彈性。應用於醫療或肢障者提供一種新的人機介面，增進殘障同胞的溝通學習能力。

8. 參考文獻

- [1]. 鐘仁厚，「基於模糊邏輯之臉部表情辨識」，國立中央大學電機工程學系，碩士論文，2008年7月。
- [2]. 陳俊宏，「不當網頁中之色情圖片偵測」，國立中央大學資訊工程研究所，碩士論文，2006年6月。
- [3]. 楊惠琳，「秀拉點描派之電腦模擬」，國立臺灣科技大學資訊管理系，碩士論文，2006年6月。
- [4]. 簡逸聰，「一個以膚色為基礎之互補人臉偵測策略」，國立中央大學資訊工程研究所，碩士論文，2001年6月。
- [5]. 謝耀璋，「主動攝影機即時人臉追蹤之研究」，國立中山大學機械與機電工程學系，碩士論文，2005年6月。
- [6]. 鄭凱方，「Face Recognizability — best face shot candidate for surveillance system」，國立中央大學資訊工程研究所，碩士論文，2005年6月。
- [7]. 王健權，「以數位影像處理搭配動態攝影機作即時人臉追蹤之研究」，國立成功大學電機工程系，碩士論文，2002年6月。
- [8]. 林煌山，「Detection of Facial Occlusions by Skin-Color based and Local-Minimum based Feature Extractor」，國立中央大學資訊工程研究所，碩士論文，2003年6月。
- [9]. 吳穎強，「類人化機器人面部表情裝置之設計與研製」，逢甲大學自動控制工程學系碩士班，碩士論文，2005年7月。
- [10]. M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuro-science*, vol.3, no.1, pp.71-86, 1991.
- [11]. K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection", in *Proc. Image Understanding Workshop*, pp. 843-850, Monterey, Calif., Nov. 1994.
- [12]. H. A. Rowley, S. Baluja, and T. Kanade, "Human face detection in visual scenes", *Tech. Rep. CMU-CS-95-158R*, Carnegie Mellon University, 1995.
- [13]. S. H. Jeng, H. Y. Mark Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, "An efficient approach for facial feature detection using geometrical face model", to appear in *Pattern Recognition*, 1997.
- [14]. K. Sobottka and I. Pitas, "Extraction of facial regions and features using color and shape information", in *Proc. International Conference on Pattern Recognition*, pp.421-425, Vienna, Austria, Aug. 1996.
- [15]. H. Wu, Q. Chen, and M. Yachida, "A fuzzy-theory-based face detector", in *Proc. International Conference on Pattern Recognition*, pp.406-410, Vienna, Austria, Aug. 1996.
- [16]. B. Moghaddam, and A. Pentland, "Probabilistic Visual Learning for Object Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, no. 7, pp. 696-710, July 1997.
- [17]. B. Moghaddam, W. Wahid, and A. Pentland, "Beyond Eigenfaces: Probabilistic Matching for Face Recognition," *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 21-35, 1998.
- [18]. C. Cortes, V. Vapnik, "Support vector networks," *Machine Learning*, Vol. 20, pp. 273-297, 1995.
- [19]. E. Osuna, "Support Vector Machines: Training and Applications," Ph.D. thesis, Dept. of EECS, Massachusetts Institute of Technology, 1998.
- [20]. E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: an Application to Face Detection," *Proc. Computer Vision and Pattern Recognition*, pp. 17-19, 1997.
- [21]. C. Garcia, and G. Tziritas. "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis." in *IEEE Transactions on Multimedia* vol. 1, No. 3, pp. 264-277, 1999.
- [22]. Chiuhsiun Lin, Kuo-Chin Fan "Triangle-based approach to the detection of human face" *Pattern Recognition* Vol.34, pp.1271-1284, 2001.
- [23]. Y. Ishii, H. Hongo, K. Yamamoto, Y. Niwa "Real-Time Face and Head Detection using Four Directional Features" *Automatic Face and Gesture Recognition, Sixth IEEE International Conference* 17-19pp.403 - 408, 2004.