

即時睡意辨識系統

Real Time Drowsiness Dection

林學儀
國立勤益科技大學
資訊工程系
hyl@ncut.edu.tw

陳奕宇
國立勤益科技大學
資訊工程系
z7172930@yahoo.com.tw

莊驊
國立勤益科技大學
資訊工程系
Dryadd44651@hotmail.com.tw

蕭力誠
國立勤益科技大學
資訊工程系
a22551162@hotmail.com.tw

曾李銘
國立勤益科技大學
資訊工程系
ronald7856@hotmail.com

競賽類別: 資訊類

摘要

本文中隨手可得之 Webcam 作為擷取影像資訊，採用非接觸式來判別駕駛者的精神狀態，讓駕駛者無使用上的感覺，即可達到安全駕駛之功能，並透過演算法來處理影像常見詬病之亮度曝光問題，可在強光及暗處下使用。

關鍵詞：影像辨識、睡意辨識。

1. 前言

國定假日與親人朋友旅遊，往往是一件快樂的事，使用汽車到達目的地是主要的移動方式之一。但往往玩樂之中又要小心駕駛，往往無法左右逢源，所以車禍是國人前幾名死因之一，疲勞駕駛又是造成此意外的主因之一，俗話說：司機一疏忽，親人兩行淚就是提醒駕駛這小心駕駛。眼睛的閉合速度就是最大的警示，所以本文中利用 Webcam 擷取影像來當作輸入，透過演算法來篩選出所需的特徵值來判別駕駛者的精神狀態，讓駕駛可在精神不濟之時得到提醒，避免發生意外。

2. 相關技術與研究探討

根據統計資料顯示，疲勞駕駛是導致交通事故主要原因之一，疲勞駕駛容易造成精神狀態不佳產生睡意。

因此需要一個即時的監控系統、安全、高可信度來預測睡意的發生，並警告駕駛者以減少意外的發生。目前在研究上有幾種可以偵測駕駛者睡意的方式。

1. Electrooculographic, EOG：直接由影像的擷取或是利用眼動圖來觀察睜眼的頻率，當睜眼速度變慢可能就是有睡意之產生。
2. Electroencephalogram, EEG：當精神不濟或是疲勞時，腦波會有一些特徵訊號，藉此來偵測睡意。

另外還有利用心電圖(Electrocardiogram, ECG)、血壓(Pressure)與呼吸(Respiration)等...

腦波主要是以非侵入式腦波擷取，再經由穩態小波(Stationary wavelet)拆解出不同頻率成分的訊號，再藉由分離出的訊號中找出訊號特徵，將它特徵化後再以類神經網路進行分類辨識[1]。

此文使用影像辨識之原因，因腦波是接觸式之感應，需接觸駕駛者才可得腦波訊號，而影像辨識是屬非接觸式之感應，無須接觸使用者也可得其訊號，對於使用者而言舒適、無障礙。

3. 系統架構

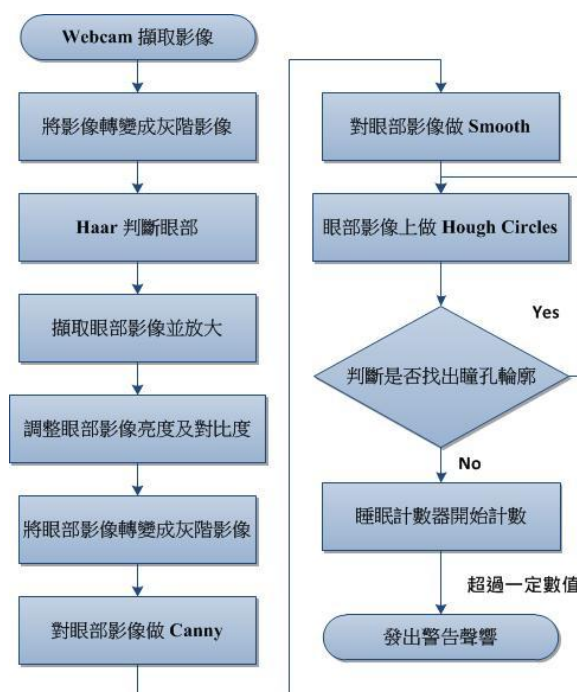


圖 3. 系統流程圖

3.1 影像灰階化

灰階(Grayscale)影像是以亮度為基礎，將彩色影像轉換到黑色至白色之間，其計算公式如下：

$$\text{Grayscale} = 0.299R + 0.587G + 0.114B$$

3.2 Canny 邊緣偵測

為了能判斷未知區域是否在物體邊緣上及對影像中不同紋理的交接處先做處理，必須對影像中物體進行邊緣化，常見的邊緣偵測技術有 Sobel、Laplacian、Canny 等，我們使用 Canny 做邊緣偵測。Canny 邊緣偵測的介紹[2] [3] [4]。

因為影像中的位置可能在任何位置，任何方向，且強度也不同，所以為了找出邊的位置、方向及強度 Canny 使用一維的高斯函數（公式 3.2.1）

$$G(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \quad (3.2.1)$$

其一階導函數（公式 3.2.2）

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^3}} e^{\frac{-x^2}{2\sigma^2}} \quad (3.2.2)$$

其二階導函數（公式 3.2.3）

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^3}} e^{\frac{-x^2}{2\sigma^2}} \left[1 - \frac{x^2}{\sigma^2} \right] \quad (3.2.3)$$

其中 σ 為標準差，而其分別對應的圖型分別為下圖所示(圖 3.2.1.)。

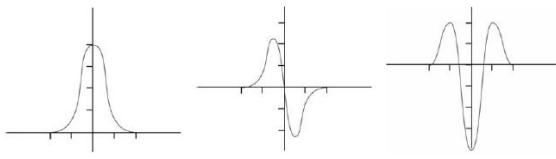


圖 3.2.1. 高斯函數和其一階，二階導函數。

這時我們考慮完整的邊在影像中將呈現出何種情形，當我們將理想的邊和高斯函數作旋積(convolution)，它將呈現出如(圖 3.2.2.)的樣子，如果我們找出影像中邊的位置所在，我們可以透過將原始影像和一階導函數作旋積之後，找出圖形中波峰所在，或是將原始影像和二階導函數作旋積之後找出 zero crossing 的點，即可以找出我們要找的邊，而影像和高斯函數作旋積可藉由(公式 3.2.4)計算求得

$$g(x, y) = D[Guass(x, y) * f(x, y)] \quad (3.2.4)$$

此式又等價於(式 2.5)

$$g(x, y) = D[Guass(x, y)] * f(x, y) \quad (3.2.5)$$

其 D 代表對其作微分， $Guass(x, y)$ 為高斯函數。

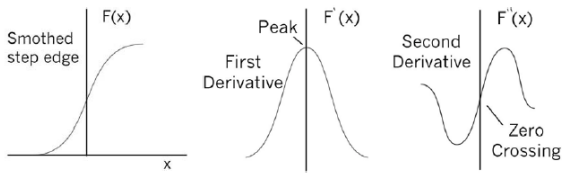


圖 3.2.2. 邊與其一階和二階的導函數關係圖

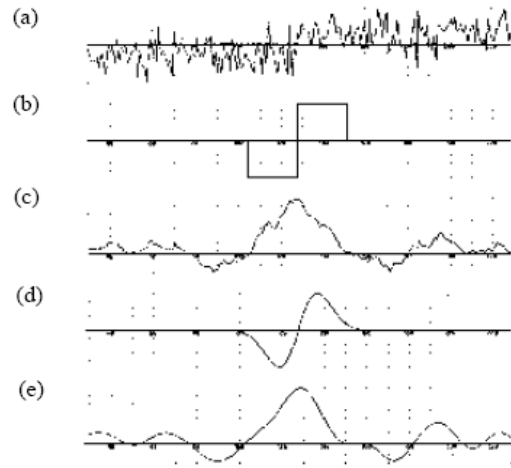


圖 3.2.3. (a)受雜訊影響的邊，(b)和(c)為梯度和高斯的濾鏡，(c)和(e)為作旋積之後的結果。（資料來源：A Computational Approach to Edge Detection.）

Canny 是由以上想法來找出邊的所在，但是不夠的，因一般邊常會受到雜訊影響，使得找出的邊可能有幾種情形發生，第一種情形是受雜訊影響使得是同一條邊的線段，分成了幾部分，第二種情形是把不是邊的線段當成了邊偵測出來，最後一種情形是偵測出來的線段被雜訊影響而變形，以上幾種情形都必須在完成線段偵測之後，要經過後處理來改善這些情形，但是 Canny 偵測線段的演算法在偵測過程中將解決這些問題，圖 3.2.3.為 Canny 偵測邊的方法和利用梯度的偵測方法所作的比較圖。

而 Canny 的偵測方式還有兩個閾值必須決定，配合著這兩個閾值便可以解決上述的三種邊緣偵測的問題，第一個門檻值的作用為，當一段邊緣受雜訊影響的時候，會產生出像是圖 3.2.3.(e)中一般的波峰，因此如果將所有波峰所在位置當成是邊緣的話，將會產生出誤判的情形，所以第一個門檻值便是用來過濾這種情形，當某一波峰的值大於此門檻值時才是真正的邊，其餘的波峰則當成是雜訊。

而第二個閾值則是要解決當一條邊緣受雜訊影響，可能被斷成好幾斷的情形，解決方式為一種遞回的做法，對影像中的每一點觀察其 8 方向的鄰點，看是否有邊的點存在，如果有且大於這個門檻值，那麼此鄰點可能是跟此點是屬於一條線段，因此就將此鄰點加入同一條線段之中，之後不斷的對影像中的每一點作同一偵測和判斷，如此即可將影像中被截斷的線連接起來。

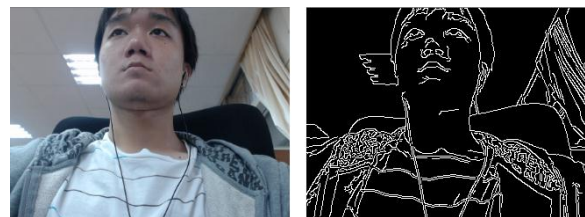


圖 3.2.4. (a)原始影像 (b)Canny 邊緣偵測結果

3.3 平滑線性濾波器

一個空間濾波器包括(1)一個鄰域(neighborhood)(通常是一個小矩形)，以及(2)作用在鄰域所圍繞之影像像素上的預定運算(predefined operation) [3] [4] [5][6]。

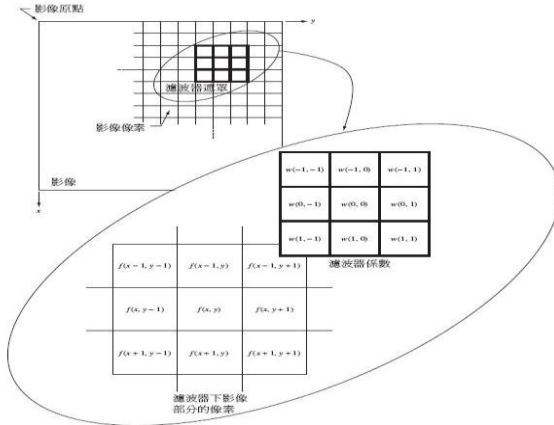


圖 3.3.1. 用一 3×3 濾波器遮罩的線性空間濾波的技術。所選用來表示濾波器遮罩係數之座標的形式使線性濾波表示式的書寫簡化。

平滑空間濾波器用於模糊化和減少雜訊，濾波器的輸出是濾波器遮罩的鄰域所含像素的平均。

平滑濾波器定義為鄰域中強度準位的平均值取代影像中的每一個像素，在強度上「銳利」變化降低的影像。因為隨機雜訊通常在強度準位上含有銳利的變化，所以平滑的最明顯運用是在減少雜訊。平均濾器的主要應用是把影像中「不相干」的細節減少。所謂「不相干」我們是指小於濾波器遮罩大小的像素區域。

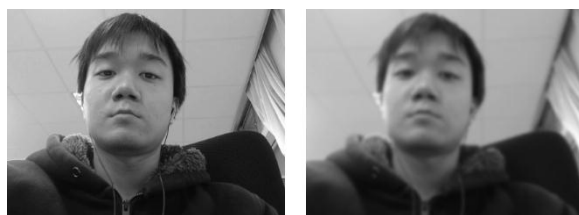


圖 3.3.1. (a)灰階影像 (b)影像平滑濾波結果

3.4 亮度及對比對調整

在 RGB 模型中，每種彩色以紅、綠、藍來顯現。利用此特點，將 RGB 三通道，分別取出 R、G、B 之各別通道，並定義彩色增強的亮度及對比度[2] [3][7]，公式如下：

原始的亮度對比數值範圍為 $-100 \sim 100$ 之間。

$$\begin{aligned} -100 &\leq Contrast \leq 100 \\ -100 &\leq Brightness \leq 100 \end{aligned} \quad (3.4.1)$$

對於對比率(Contrast ratio)來講， Δ 範圍應該

落在 $0 \sim 255$ ，這邊將對比率的公式做重新的調整。

$$\begin{cases} \text{if } Contrast > 0 & \Delta_1 = 127 * \frac{Contrast}{100} \\ \text{else } Contrast \leq 0 & \Delta_2 = -128 * \frac{Contrast}{100} \end{cases}$$

$$0 < \Delta_1 \leq 127$$

$$0 < \Delta_2 \leq 128 \quad (3.4.2)$$

對比率代表著斜率的 α 值，而亮度則是決定線性公式位移的情況，也就是 β 值，而 $y = \alpha x + \beta$ 這個線性公式所表達的情況如下：

$$\begin{cases} \text{if } Contrast > 0 & y_i = \left(\frac{255}{255 - 2\Delta_1} \right) x_i + \left(\frac{255}{255 - 2\Delta_1} \right) Brightness - \left(\frac{255}{255 - 2\Delta_1} \right) \Delta_1 \\ \text{else } Contrast \leq 0 & y_i = \left(\frac{256 - 2\Delta_2}{255} \right) x_i + \left(\frac{256 - 2\Delta_2}{255} \right) Brightness + \Delta_2 \end{cases} \quad (3.4.3)$$

α 值的範圍落在 $0 \sim 255$ 之間，而它的情況如下：

$$y = \alpha x + \beta \quad (3.4.4)$$

$$\begin{cases} \text{if } Contrast > 0 & 1 < \alpha \leq 255 \\ \text{else } Contrast \leq 0 & 0 \leq \alpha \leq 1 \end{cases}$$

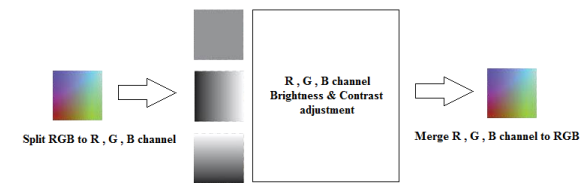


圖 3.4.1. 亮度及對比度調整流程

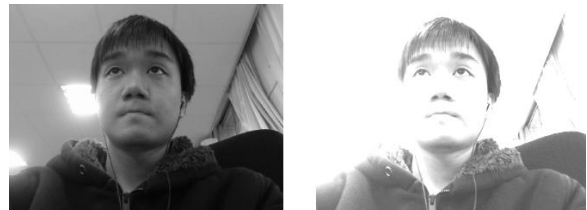


圖 3.4.2. (a)灰階影像 (b)亮度及對比度調整結果

3.5 Haar 人臉檢測

Viola 提出了積分影像(Integral Image)概念和一個基於 Adaboost 方法訓練人臉偵測分類器的方法，建立了一個即時的人臉偵測系統 [3] [8][9][10]。

(1) 積分影像(Integral Image)

對於一個輸入影像 I ，像素 (x, y) 處的積分影像值定義為 $ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$ 其中 $i(x, y)$ 表示像素 (x, y) 的灰階值，如圖 3.5.1.所表示積分影像值等於陰影部分所有灰階值的和。

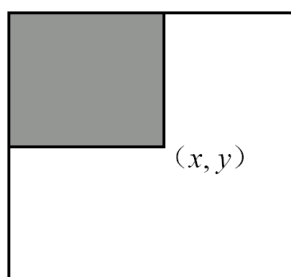


圖 3.5.1. 點(x,y)處的積分影像值

如此定義，如果要得到一個輸入影像 I 的積分影像時，僅需對原圖逐點掃描一次就可計算出來。計算方式如下

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (3.5.1)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (3.5.2)$$

其中 $s(x, y) = \sum_{y_1 \leq y} i(x, y_1)$ 是原影像 (x, y) 點的列累積和 (Cumulative Row Sum)， $s(x, -1) = 0$ 和 $ii(x - 1, y) = 0$ 。

使用積分影像時，任何影像內部的矩形和 (Rectangular Sum) 可以僅使用四個參考值計算即可取得，如圖 3.5.2. 所示。在 Viola 的系統中，每個矩形特徵值的計算，最多只需要從積分影像中取 9 個元素做加減運算。這是 Viola 方法速度非常快的根本原因。

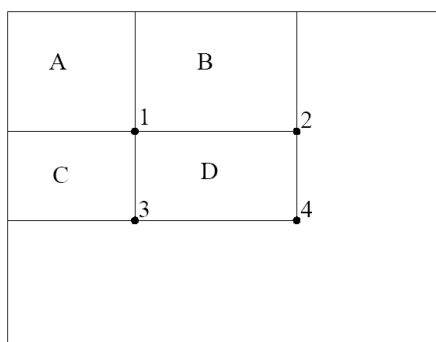


圖 3.5.2. 計算矩形和示意圖，矩形 D 的所有灰階值的和為 $4+1-(2+3)$

(2) 矩形特徵、弱分類器、Adaboost 演算法

Freund 的 Adaboost 演算法是一種疊代方法，主要用途是透過從大量的弱分類器中選取最具有分類意義的那些組合成一個強分類器，其中弱分類器指的是那些性能比隨機分類略好一點的分類器，在 Viola 中也有具體演算法可以參考。在 Viola 的方法中，使用矩形特徵作為分類的依據。矩形特徵的值是指影像上兩個或者多個形狀大小相同的矩形內部所有像素灰階值之和的差值。在 Viola 的系統中，使用了四種矩形特徵，如圖 3.5.3. 所示。

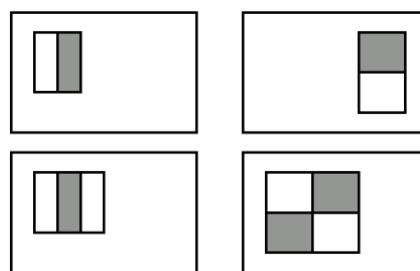


圖 3.5.3. Viola 使用的四種矩形特徵

在這裡，一個弱分類器構造如下。一個矩形特徵 j 對應著一個弱分類器 h_j ，對於一個候選輸入區塊 x ，設該矩形特徵 x 上的值為 $f_j(x)$ ，則弱分類器分類函數表達如下：

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (3.5.3)$$

其中 p_j 為 ± 1 ， θ_j 是一個閾值， x 為一個影像中 24×24 像素的子區塊，對於這樣大小的區塊，可能的矩形特徵的個數超過 180,000。在實際計算中，必須找到那些對於分類比較重要的矩形特徵，Adaboost 演算法是選取這些特徵的有效手段，圖 2.9 就是利用此方法找到的第一和第二特徵。在給定的訓練樣本上，Adaboost 演算法從所有可能的弱分類器中挑選錯誤最低的那一個，然後改變樣本的權值，使得那些被錯分的樣本得到進一步重視，重複上述操作。這樣，每一步操作都得到一個弱分類器，最後的強分類器由這些弱分類器的線性組合構成。Freund 和 Schapire 證明了隨著弱分類器個數的增加，得到的強分類器在訓練樣本上的分類錯誤依指數遞減。

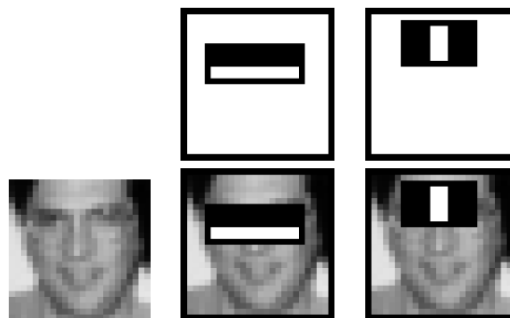


圖 3.5.4. AdaBoost 演算法選出的第一及第二個特徵

(3) 層疊分類器 (Cascade of Classifiers)

層疊分類器是很多強分類器的組合，如圖 3.5.5. 所示。每一層都是 Adaboost 演算法訓練得到的一個強分類器，都經過閾值調整，使得每一層都能讓幾乎全部人臉樣本通過，而排除了大部分非人臉樣本。

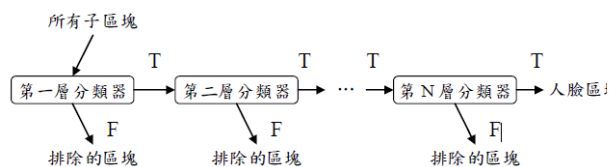


圖 3.5.5. 層疊分類器示意圖

3.6 圓形霍夫轉換

在影像進行霍夫轉換之前，將影像雜濾除，只留所需影像資訊。影像中的亮度、色度及飽合度，只有亮度的資訊對人眼是比較敏感的，所以亮度的差異在人眼上也比較容易辨識。霍夫轉換所要的是直線的資訊，而影像中直線資訊就是影像中物件的邊緣，因此在灰階影像上，需再對其進行邊緣偵測。影像處理中，我們採 Canny 邊緣偵測[3][4][5][6]。

為防止直線斜率無限大發生，變更直線方程式的表示方式，將原本直線方程式參數(m,b)改以(r,θ)的方式來表示，如式(3.6.1)。所以霍夫轉換直線偵測主要是將歐基里德空間的直線方程式中的參數(m,b)以(r,θ)來表示，如圖 3.6.1，其中 r 為直線到原點的垂直距離，θ 為 r 與 x 軸正方向的夾角。

$$r = x \cos \theta + y \sin \theta \quad (3.6.1)$$

霍夫轉換做法是先透過邊緣偵測，二值化...等影像處理找出影像的特徵點，再將每個特徵點代入霍夫轉換公式，在參數空間中繪出其特徵點轉換後的線條，而由直線方程式轉換後產生的是直線線條，但以霍夫轉換直線偵測的方程式以(r,θ)來表示時，將會是一條曲線線條，其表示範圍是在 $0 \sim 2\pi$ 之間。在歐基里德空間中設一個特徵點 A，在參數空間上可以描繪出一條曲線，若在參數空間上產生一點 B 時，則在歐基里德空間上將產生一條直線線條，如圖 3.6.2. 所示，所以在參數空間上將會有影像中所有特徵點所產生的所有曲線。根據每條曲線交叉又重疊的點來累積並記錄曲線交叉的重疊次數，也就是做投票(Voting)動作。找出投票數最高也就是累積次數最多的那一點記錄其(r,θ)座標，該累積次數最多的即最可能為影像中的直線，如圖 3.4 所示。將所有高過一定累積次數的(r,θ)座標代入公式(3.6.1)，即可繪出影像中所求的直線。

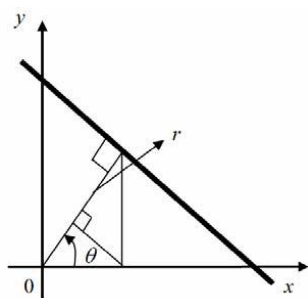


圖 3.6.1. 霍夫轉換(r,θ)參數說明圖

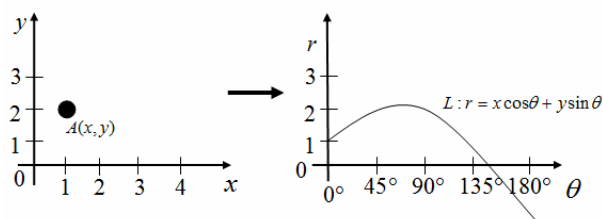


圖 3.6.2. 直線霍夫轉換示意圖

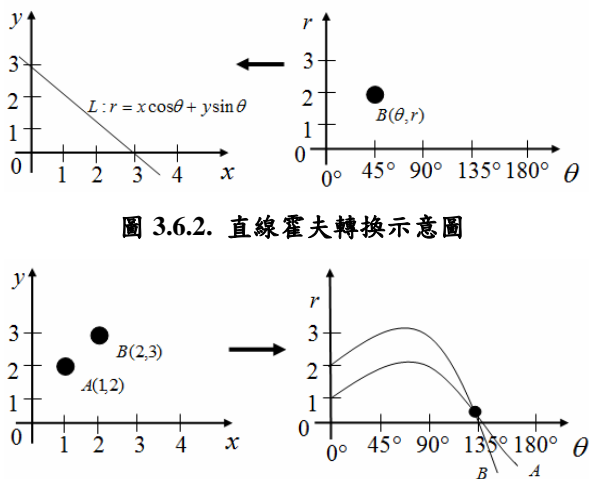


圖 3.6.3. 直線霍夫轉換投票示意圖

圓形偵測與直線偵測的前置做法相同，將灰階化後影像進行邊緣偵測，再對其二值化。接著將影像中的特徵點代入方程式(3.6.2)，其中(x, y)為影像座標點，(a,b)為圓心座標，r 為圓的半徑。

$$(x - a)^2 + (y - b)^2 = r^2 \quad (3.6.2)$$

圓形偵測的霍夫轉換主要是在三維的參數空間上表示，所有在影像中的圓，轉換到參數空間上會成為一個點，而根據半徑的不同，在參數空間上會因此成為一個圓錐體的形狀，如圖 3.6.4，與直線偵測的霍夫轉換的參數空間有所不同。

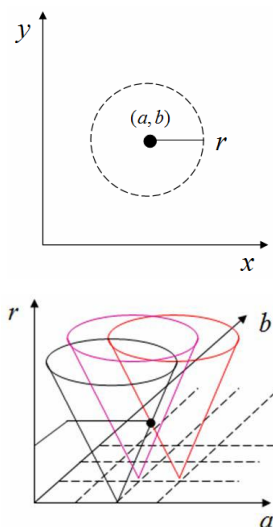


圖 3.6.4. 圓形霍夫轉換示意圖

而圓形偵測的做法與直線偵測相似，將邊緣偵測及二值化後所得到的特徵點，也就是影像的座標點代入公式(3.6.2)，設定半徑 r 值的範圍，再將圓心的座標點代入公式(3.6.2)，依據設定的圓心與半徑，會在參數空間上匯出一個圓錐體，依據不同的圓心就能得到各個圓心在參數空間上匯出的圓錐體。和直線偵測一樣，對參數空間中的圓錐體進行投票，並設門檻值對投票數做判斷。當投票數大於或等於設定門檻時，便記錄該組(a,b)座標及半徑 r ，則該組座標即為影像中最有可能的圓，因此可在原影像中繪出圓形。

3.7 實驗結果



圖 3.7.1. Webcam 擷取之影像



圖 3.7.2. Haar 取出眼部特徵值並作亮度及對比度調整



圖 3.7.3. 將影像作邊緣化偵測及平滑處理



圖 3.7.4. 使用霍夫圓形抓取出眼部及瞳孔



圖 3.7.5. 時間內未抓取到圓形便發出警告

4. 結論

透過影像處來判斷駕駛者的精神狀態是一個非常有效用的方法，可在不同的光源底下進行，不會受到干擾是本文的特色之一，外加使用容易取的 Webcam 當作感測器，大幅度的降低成本，並使用 C 語言為開發藍本，日後應用於車用電子上，可達到快速整合開發之應用，不須使用特殊的硬體即可達到所想要的成果，讓行車安全又更上一層樓。

5. 參考文獻

- [1] 徐良育、蔡百原，以數位訊號處理單晶片實現可攜式睡意辨識系統，中原大學生物醫學工程學系碩士學位論文，桃園，2007。
- [2] 繆紹綱，數位影像處理，白灣培生教育出版股份有限公司，2009。
- [3] GARY BRADSKI, ADRIAN KAEHLER, LEARNING OPENCV, O'REILLY MEDIA, 2008.
- [4] 鄭志宏、林義隆、陳韋任，使用霍夫神經網路於車道線偵測之研究，義守大學資訊工程研究所碩士論文，高雄，2009。
- [5] [HTTP://OPENCV.WILLOWGARAGE.COM/WIKI/](http://OPENCV.WILLOWGARAGE.COM/WIKI/)
- [6] [HTTP://WWW.OPENCV.ORG.CN/INDEX.PHP](http://WWW.OPENCV.ORG.CN/INDEX.PHP)
- [7] [HTTP://YESTER-PLACE.BLOGSPOT.COM/](http://YESTER-PLACE.BLOGSPOT.COM/)
- [8] 蘇木春、楊煒達，簡易方法之少量人臉辨識系統，國立中央大學資訊工程研究所碩士論文，桃園，2007。
- [9] 蘇木春、熊昭岳，行車安全偵測系統，國立中央大學資訊工程研究所碩士論文，桃園，2005。
- [10] 陳勝勇、劉盛，基於 OPENCV 的電腦視覺技術實現，科學出版社，2008。