

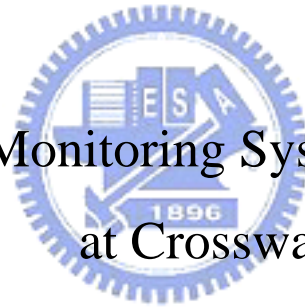
國立交通大學

電機與控制工程研究所

碩士論文

行人穿越道行人監控系統之研究

A Study on Monitoring System for Pedestrians
at Crosswalks



研究生：江啟逞

指導教授：林昇甫 博士

中華民國九十四年七月

行人穿越道行人監控系統之研究

A Study on Monitoring System for Pedestrians at Crosswalks

研究生：江啟逞

Student: Chi-Cheng Chiang

指導教授：林昇甫 博士

Advisor: Dr. Sheng-Fuu Lin

國立交通大學

電機與控制工程研究所

碩士論文

A Thesis

Submitted to Department of Control Engineering
College of Electrical Engineering and Computer Science

National Chiao-Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master

in

Electrical and Control Engineering

July 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年七月

行人穿越道行人監控系統之研究

研究生：江啟暹

指導教授：林昇甫 博士

國立交通大學電機與控制工程研究所

摘 要

使用影像序列作即時的行人偵測被廣泛地應用在智慧型交通系統(intelligent transportation system, ITS)，例如交通監控以及交通號誌控制，目前存在許多行人偵測的方法，但是隨著行人數目以及行人交錯程度的增加，行人偵測的效能將會變的很差。根據行人穿越道上行人的特性，我們提出一個方法可以有效地偵測及追蹤行人穿越道上在人群交錯狀況下的行人。

一般而言，行人會在行人穿越道的兩端穿越道路，因此我們只在行人穿越道的兩端偵測行人，一旦行人被偵測到便開始追蹤行人，如此我們可以得知行人在任意時間的位置並且減少運算複雜度；另外，我們將攝影機鏡頭設置成與道路方向垂直，這樣的好處是行人的影像將會左右對稱，這個特性會被用來辨識行人。行人偵測的方法是先偵測出類似人頭外形的邊緣(edge)，接著使用結合許多行人特性的模糊推論系統(fuzzy inference system, FIS)判斷偵測到的邊緣是否可能是人頭；在某些情形下人頭的邊緣並不明顯，以上的方法將會失效，但對於孤立的行人可以利用行人的輪廓作為偵測行人的依據。追蹤的方法是比對顏色的直方圖(histogram)，預測行進路徑的方法是使用簡易的無限脈衝響應(infinite impulse response, IIR) 濾波器。實驗結果顯示，當光線不是太暗而且行人穿越道不是太長下，本論文所提出的方法對於行人穿越道的行人偵測及追蹤有不錯的效果。

A Study on Monitoring System for Pedestrians at Crosswalks

Student: Chi-Cheng Chiang Advisor: Dr. Sheng-Fuu Lin

Institute of Electrical and Control Engineering
National Chiao-Tung University

Abstract

Real-time pedestrian detection using video sequences has myriad applications in the intelligent transportation system such as traffic monitoring and traffic signal control. Many methods have been proposed for the detection of pedestrian. However, as the number of pedestrian and the amount of their occlusion increases, the system performance degrades. According to the properties of pedestrians at crosswalks we propose an approach to efficiently detect and track pedestrians in the crowded situation at crosswalks.

Generally, the pedestrian will pass through the road from one side to another side of the crosswalk. Therefore we only detect pedestrians on the both sides. Once a pedestrian has been detected, the system will start to track the pedestrian. Then we can know the location of the pedestrian at any time, and the computation complexity can be reduced. In addition, we set up the camera with the direction perpendicular to the road. In this way an advantage is that the pedestrian's image will be about symmetrical. This characteristic will be used for recognizing the pedestrian. For detecting pedestrians the system searches human head shape in the edge image first and then uses the fuzzy inference system which combines several pedestrian's characteristics to judge whether it is human head. That in some situations edges are not apparent makes the above method inactive but the isolated pedestrian can be detected by human shape contour. For tracking pedestrians the system exploit the color histogram match. The next position of every pedestrian can be predicted by infinite impulse response filter. The experimental results show that if the light is not too dim and the crosswalk is not to long, then the proposed method has good performance on detecting and tracking pedestrians.

誌 謝

首先要感謝我的指導教授林昇甫博士，感謝他在研究上的啟迪與指導，以及細心校閱本論文的謬誤之處。同時也要感謝本實驗室的學長建焜、士哲、晉嘉，同學宗漢、旺霖、尚謙、昭佑、俊穎、國育以及學弟們，由於他們熱心的協助與討論，使得研究過程中的困境能夠迎刃而解。

最後要感謝我親愛的父母及兄長，感謝他們的辛勞養育，以及對我的鼓勵與支持，使我得以在精神和經濟上無後顧之憂，順利完成學業。



目錄

中文摘要	i
Abstract	ii
誌謝	iii
目錄	iv
圖目錄	vii
表目錄	xii
	
第一章 緒論	1-1
1.1 相關研究	1-2
1.1.1 行人追蹤	1-2
1.1.2 人群計數	1-5
1.2 研究動機	1-7
1.3 論文架構	1-8
第二章 影像處理技術以及模糊推論系統	2-1
2.1 影像處理技術	2-1
2.1.1 成像幾何學	2-2

2.1.2 邊緣檢測	2-6
2.1.3 色彩空間轉換	2-10
2.1.4 形態學	2-11
2.2 模糊推論系統	2-13
2.2.1 模糊化機構	2-14
2.2.2 模糊規則庫	2-15
2.2.3 模糊推論引擎	2-16
2.2.4 去模糊化機構	2-16

第三章 行人監控系統 3-1

3.1 系統概述	3-1
3.2 相機校正	3-3
3.3 橢圓形邊緣偵測	3-5
3.3.1 前景偵測	3-5
3.3.2 邊緣擷取	3-6
3.3.3 橢圓偵測	3-7
3.4 色彩校正	3-11
3.5 行人識別	3-13
3.5.1 行人特徵	3-13
3.5.1.1 人頭與兩側前景的顏色差異度	3-13
3.5.1.2 身體比例	3-17
3.5.1.3 衣服紋路相似度	3-18
3.5.1.4 膚色量	3-19
3.5.2 模糊推論系統架構	3-21
3.5.2.1 模糊近端行人辨識系統	3-21
3.5.2.2 模糊遠端行人辨識系統	3-23
3.6 輔助的行人偵測	3-25
3.7 行人追蹤	3-27

3.7.1 目標物比對	3-28
3.7.2 行進路徑追蹤	3-30
第四章 實驗結果與分析	4-1
4.1 實驗機制	4-1
4.1.1 實驗設備	4-1
4.1.2 效能評估	4-3
4.2 實驗結果	4-3
4.3 實驗分析	4-10
第五章 結論	5-1
參考文獻	I



圖目錄

1.1. 車輛與行人交會的情形。	1-2
1.2. 利用前景像素的垂直投影偵測行人的圖例。(a) 原始影像；(b) 背景影像；(c) 偵測出為前景的像素；(d) 前景像素垂直投影的直方圖；(e) 偵測出的行人。	1-3
1.3. 利用連通邊界偵測行人的圖例。(a) 邊界影像；(b) 偵測出的行人。	1-3
1.4. 利用膚色及邊緣偵測人臉的圖例。(a) 原始影像；(b) 膚色區域；(c) 邊緣影像；(d) 偵測出的人臉。	1-4
1.5. 利用移動的邊緣偵測人頭的圖例。(a) 原始影像；(b) 移動的影像；(c) 移動的邊緣；(d) 偵測出的人頭。	1-4
1.6. 利用影像處理技術估測人數的圖例。(a) 背景影像及 ROI；(b) 一個前景像素的例子；(c) 一個邊緣影像的例子；(d) 前景像素個數與實際人數的關係；(e) 前景的邊緣像素個數與實際人數的關係。	1-6
2.1. 針孔相機模型。	2-2
2.2 相似三角形模型。	2-2
2.3 WCS 以及 CCS 的相對位置。	2-4
2.4 利用方塊物件求取相機參數的圖例。(a) 原始影像；(b) 影像的消失點、主要點、以及地平線。	2-5
2.5 計算相機位置高度的示意圖。	2-6
2.6 一個應用導數運算子作邊緣檢測的例子。(a) 灰階影像；(b) 沿著水平掃描線的灰階剖面圖；(c) 灰階剖面圖的一階導數。	2-7
2.7 具有任意係數的一個 3×3 遮罩。	2-7
2.8 一個 3×3 的影像區域和 Sobel 遮罩。(a) 影像區域；(b) 用來計算 G_x 的遮罩；(c) 用來計算 G_y 的遮罩。	2-8

2.9	應用邊緣檢測的結果。(a) 原始彩色影像；(b) 灰階影像；(c) 灰階影像的邊緣；(d) 彩色影像的邊緣。	2-9
2.10	RGB 彩色模型。	2-10
2.11	侵蝕運算。(a) 侵蝕前影像；(b) 結構元素；(c) 侵蝕後影像。	2-11
2.12	膨脹運算。(a) 膨脹前影像；(b) 結構元素；(c) 膨脹後影像。	2-12
2.13	邊界提取。(a) 原始影像；(b) 結構元素；(c) 提取的邊界。	2-13
2.14	模糊系統的基本架構。	2-14
2.15	歸屬函數。(a) Z 型；(b) 三角形；(c) S 型。	2-15
2.16	一個模糊推論過程的例子。	2-17
3.1	行人監控系統的架構。	3-2
3.2	橢圓形邊緣的偵測範圍。	3-3
3.3	攝影機架設位置的示意圖。(a) 俯視圖；(b) 側視圖。	3-3
3.4	消失點 V_y 以及 V_z 的座標位置。	3-4
3.5	建立背景影像的圖例。(a)至(c) 用來建立背景影像的部份影像序列；(d) 建立的背景影像。	3-5
3.6	建立前景遮罩影像的圖例。(a) 輸入影像；(b) 背景影像；(c) 前景遮罩。	3-6
3.7	由圖 3.6 偵測出的前景的邊緣。(a) 邊緣強度影像；(b) 邊緣強度大於臨界值 τ_g 的像素。	3-7
3.8	橢圓模型。	3-8
3.9	橢圓形邊緣偵測的圖例。(a) 輸入影像與限定的橢圓中心位置；(b) 行人身長示意圖；(c) 邊緣影像與偵測出的橢圓；(d) 輸入影像與偵測出的橢圓。	3-10
3.10	移動物件的偵測。(a) 目前的輸入影像；(b) 前一刻的輸入影像；(c) 移動的前景；(d) 移動物件的範圍。	3-11
3.11	色彩校正的圖例。(a)輸入影像；(b) 參考影像；(c) 消除色偏後的影像；(d) 調整亮度後的影像。	3-12
3.12	非人頭的橢圓形邊緣。(a) 輸入影像與偵測到的橢圓；(b) 前景的像素；(c) 邊緣影像；(d) 偵測到的橢圓形邊緣。	3-13
3.13	人頭與兩側前景的顏色差異。(a) 兩側皆有前景的人頭；(b) (a)的前景	

像素；(c) 兩側沒有前景的人頭；(d) (c)的前景像素。.....	3-14
3.14 橢圓與兩側偵測顏色的範圍。.....	3-15
3.15 非人頭的直方圖例子。(a) 輸入影像與偵測顏色的範圍；(b) 橢圓內部的直方圖；(c) 橢圓左側前景的直方圖；(d) 橢圓右側前景的直方圖。.....	3-15
3.16 人頭的直方圖例子。(a) 輸入影像與偵測顏色的範圍；(b) 橢圓內部的直方圖；(c) 橢圓左側前景的直方圖；(d) 橢圓右側前景的直方圖。.....	3-16
3.17 人頭與人頭上下兩側的顏色比較。(a) 近端人頭；(b) 遠端人頭。.....	3-17
3.18 身體樣板的示意圖。.....	3-18
3.19 人頭下方左右兩側的示意圖。.....	3-18
3.20 人臉樣板的示意圖。.....	3-19
3.21 膚色模型。(a) 膚色在 YCbCr 色彩空間的分布；(b) 膚色在 CbCr 色彩空間投影的分布。.....	3-20
3.22 人臉顏色與色偏的關係。(a) 黃色路燈下的人臉顏色在 CbCr 色彩空間投影的分布；(b) 色彩校正後的人臉顏色在 CbCr 色彩空間投影的分布。.....	3-21
3.23 NPR 的輸入以及輸出的歸屬函數。(a) <i>CDH</i> 的歸屬函數；(b) <i>RUL</i> 的歸屬函數；(c) <i>TSC</i> 的歸屬函數；(d) <i>PL</i> 的歸屬函數。.....	3-22
3.24 FPR 的輸入以及輸出的歸屬函數。(a) <i>CDH</i> 的歸屬函數；(b) <i>RUL</i> 的歸屬函數；(c) <i>QOS</i> 的歸屬函數；(d) <i>PL</i> 的歸屬函數。.....	3-24
3.25 呈現行人外型前景的圖例。(a) 輸入影像；(b) 前景影像；(c) 邊緣影像。.....	3-25
3.26 以輪廓為基礎的行人偵測系統的架構。.....	3-26
3.27 以輪廓為基礎的行人偵測系統的圖例。(a) 輸入影像；(b) 前景像素；(c) 斷開及閉合後的影像；(d) 邊界影像；(e) 邊界標記；(f) 邊界曲線及找出的人頭中心。.....	3-27
3.28 判斷人型邊界的示意圖。(a) 人型邊界的判斷方法示意圖；(b) 人頭附近的前景偵測範圍。.....	3-27
3.29 新人頭中心的限制範圍。.....	3-28
3.30 建立直方圖的區域。.....	3-29
4.1 不同的實驗環境所建立的背景圖。(a) env1；(b) env2；(c) env3；(d)	

env4 ; (e) env5 ; (f) env6 ; (g) env7 ; (h) env8 ; (i) env9 。	4-2
4.2 env1-1 偵測及追蹤的部份結果影像。(a) 幀 22 ; (b) 幀 79 ; (c) 幀 142 ; (d) 幀 232 ; (e) 幀 328 ; (f) 幀 562 。	4-6
4.3 env2-1 偵測及追蹤的部份結果影像。(a) 幀 52 ; (b) 幀 109 ; (c) 幀 202 ; (d) 幀 310 ; (e) 幀 361 ; (f) 幀 463 。	4-6
4.4 env3-1 偵測及追蹤的部份結果影像。(a) 幀 40 ; (b) 幀 82 ; (c) 幀 136 ; (d) 幀 190 ; (e) 幀 253 ; (f) 幀 307 。	4-7
4.5 env4-1 偵測及追蹤的部份結果影像。(a) 幀 262 ; (b) 幀 421 ; (c) 幀 463 ; (d) 幀 526 ; (e) 幀 622 ; (f) 幀 700 。	4-7
4.6 env5-1 偵測及追蹤的部份結果影像。(a) 幀 64 ; (b) 幀 130 ; (c) 幀 199 ; (d) 幀 244 ; (e) 幀 322 ; (f) 幀 406 。	4-8
4.7 env6-1 偵測及追蹤的部份結果影像。(a) 幀 73 ; (b) 幀 304 ; (c) 幀 484 ; (d) 幀 613 ; (e) 幀 877 ; (f) 幀 982 。	4-8
4.8 env7-1 偵測及追蹤的部份結果影像。(a) 幀 61 ; (b) 幀 145 ; (c) 幀 271 ; (d) 幀 382 ; (e) 幀 505 ; (f) 幀 601 。	4-9
4.9 env8-1 偵測及追蹤的部份結果影像。(a) 幀 205 ; (b) 幀 247 ; (c) 幀 337 ; (d) 幀 514 ; (e) 幀 634 ; (f) 幀 727 。	4-9
4.10 env9-1 偵測及追蹤的部份結果影像。(a) 幀 79 ; (b) 幀 244 ; (c) 幀 496 ; (d) 幀 613 ; (e) 幀 856 ; (f) 幀 1060 。	4-10
4.11 近端行人偵測效能與時段的關係。(a) 行人偵測成功率 ; (b) 行人偵測 精確度 。	4-12
4.12 遠端行人偵測效能與時段的關係。(a) 行人偵測成功率 ; (b) 行人偵測 精確度 。	4-12
4.13 近端行人偵測效能與行人穿越道長度的關係。(a) 行人偵測成功率 ; (b) 行人偵測精確度 。	4-12
4.14 遠端行人偵測效能與行人穿越道長度的關係。(a) 行人偵測成功率 ; (b) 行人偵測精確度 。	4-13
4.15 行人追蹤效能與時段的關係。(a) 近端 ; (b) 遠端 。	4-13
4.16 行人追蹤效能與行人穿越道長度的關係。(a) 近端 ; (b) 遠端 。	4-13
4.17 系統偵測不出的行人。(a) 與前景行人顏色相近的行人 ; (b) 被前面行	

人遮蔽的行人；(c) 沒有顯現正常人頭的行人。	4-15
4.18 系統偵測出的錯誤行人。(a) 機車騎士(編號 54)；(b) 近端形成類似行人的圖樣(編號 12)；(c) 遠端形成類似行人的圖樣(編號 6)。	4-16
4.19 行人突然轉身導致行人追蹤失敗的例子(編號 16)。(a) 幀 139；(b) 幀 163；(c) 幀 187。	4-17
4.20 追蹤中的行人被其它前景持續遮蔽導致行人追蹤失敗的例子(編號 61)。(a) 幀 1786；(b) 幀 1816；(c) 幀 1870。	4-18
4.21 光源變化太大導致行人追蹤失敗的例子(編號 5)。(a) 幀 598；(b) 幀 619；(c) 幀 667。	4-19
4.22 不同方法的行人偵測結果。(a) 輸入影像；(b) 本論文提出的方法；(c) 以輪廓為基礎的方法；(d) 以膚色為基礎的方法。	4-21



表目錄

4.1 實驗環境的代稱。.....	4-2
4.2 影像序列的代稱。.....	4-2
4.3 系統偵測行人的效能。.....	4-4
4.4 系統追蹤行人的效能。.....	4-5



第一章

緒論

隨著都市人口及車輛數目的急遽增加，現代化的交通系統產生了數項嚴重的問題，例如過多的車輛造成了交通壅塞的情形。另外，由於原先的交通系統並未規劃事故發生時因應的相關設備或機制，導致往往在交通事故發生時需要冗長的處理時間而使交通大亂。ITS 就是為了因應上述的問題而提出的構想。根據[1]的論述，ITS 的幾個主要的目標包括：1) 交通流量監控及交通動線最佳化管理[2], [3]；2) 自動駕駛車輛或輔助駕駛系統[4], [5]；以及 3) 其他應用等等。

由於影像具有高搜尋範圍及價格低廉等優點，因此有許多文獻利用影像處理的技術作為交通流量的監控，一般的作法是先偵測移動物件，接著判斷移動物件是否為車輛，並對車輛作追蹤以取得車流量資訊[2]，根據取得的交通流量調整十字路口交通號誌的切換時間，可以對交通動線作最佳化管理，以減少交通擁擠的情況發生[3]。過去的文獻主要是在監控每條道路上車輛的流量，然而要達到交通動線的最佳化，實際上還必須考慮到行人穿越道上的行人流動，這是由於在行人穿越道上通行的人群會與要在十字路口轉彎的車流互相阻滯，如圖 1.1 所示，因此，即時地對行人穿越道的行人作追蹤以及計算行人數目可以提升交通管理的效果。另外，對行人作追蹤可以預測行人的行進路徑，若配合對車輛的行進路徑做預測，當預測出行人與車輛可能發生碰撞的情形時，可以利用通訊技術，對車輛的駕駛人員發出警告訊息以降低交通意外及傷亡人數的產生。

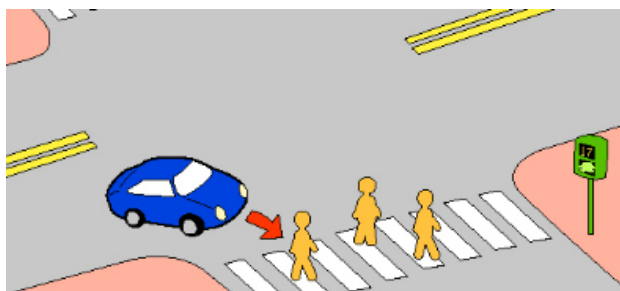


圖 1.1 車輛與行人交會的情形。

1.1 相關研究

與行人相關的研究相當廣泛，大致上可以分為兩類，一類是對行人作偵測及追蹤，這類的研究一般只能對少數的行人做追蹤，通常行人數目為個位數，相關文獻會在 1.1.1 節探討。另一類是估測目標區域(region of interest, ROI)的人群數目，會在 1.1.2 節討論。

1.1.1 行人追蹤

對追蹤物件而言，最簡單的方式是先將物件從背景(background)中偵測出來，也就是取得所謂的前景(foreground)，根據前景中每個像素的相對距離，將距離相近的像素組成一個簇(blob)，再分別對不同的簇作追蹤[6]。這樣的方法相當簡單，但當兩個不同的簇交會後再分離，則這個方法將無法正確的分辨出這兩個不同的簇。為了解決以上問題，必須先辨識出前景中的各個行人，並紀錄每個行人的特徵，作為追蹤不同行人所需的資訊。Haritaoglu 等人[7]設計了一個可以更新的統計式背景模型，利用背景相減法(background subtraction)及中值濾波器(median filter)消去雜訊，得到前景(如圖 1.2(a)到 1.2(c))。對前景的像素做垂直投影得到直方圖(如圖 1.2(d))，若是直方圖的峰值在可能的影像人高範圍內，則代表在該位置有行人(如圖 1.2(e))，但是當輸入影像中某一個行人的下半身與另一名行人的上半身交疊在一起時，這樣的方法將會失敗。Zhao 等人[8]也是利用背景相減法得到前景，藉由形態學(morphology)提取前景的邊界(boundary)，再利用連通分量(connected

component)對連結的邊界群分別作標記(labeling)(如圖 1.3(a))，接著配合透視轉換(perspective transformation)判斷邊界的形狀似否類似人體的形狀以偵測行人(如圖 1.3(b))。

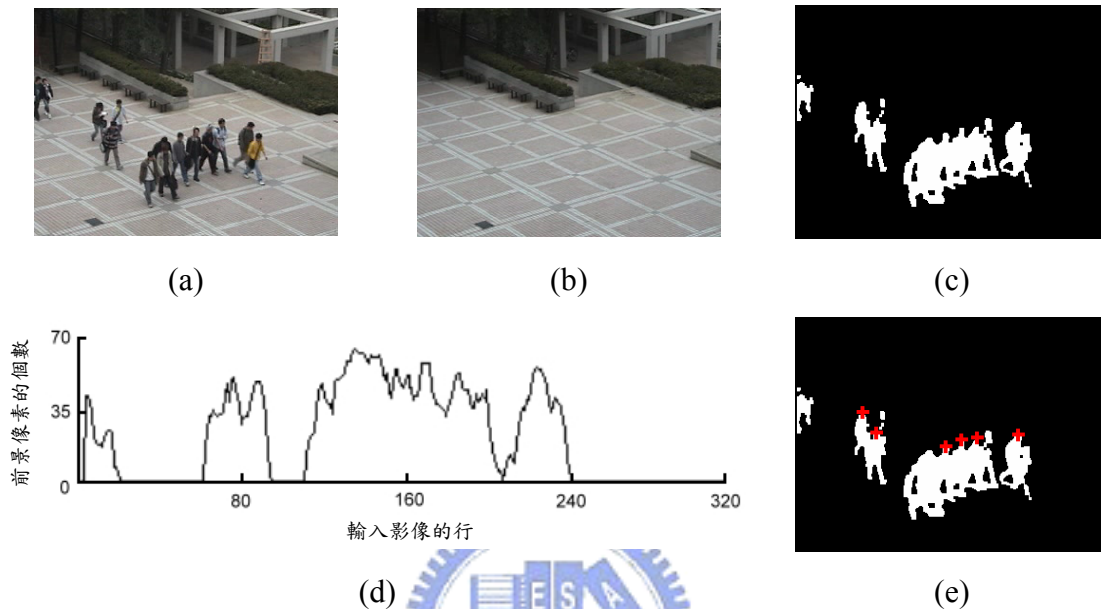


圖 1.2 利用前景像素的垂直投影偵測行人的圖例。(a) 原始影像；(b) 背景影像；(c) 偵測出為前景的像素；(d) 前景像素垂直投影的直方圖；(e) 偵測出的行人。



圖 1.3 利用連通邊界偵測行人的圖例。(a) 邊界影像；(b) 偵測出的行人。

除了上述以整個人的形體偵測行人的方法外，有些文獻則是偵測人的頭部。Hori 等人[9]先偵測出影像中膚色(skin color)的部份並且去除膚色區域小的區塊得到二值化影像(如圖 1.4(b))，再用 Laplacian 濾波器偵測出邊緣 (如圖 1.4(c))，接著對邊緣與下半橢圓形的樣板(template)作比對即可找出與下半橢圓相似性高的人臉

部份(如圖 1.4(d))。其中下半橢圓的比對相當花費時間，為了使系統能作即時的偵測，因此使用**基因演算法**(genetic algorithm)偵測下半橢圓。事實上，影像中膚色的部份通常不多，因此已經大大的減少了下半橢圓的搜尋範圍。為了減少橢圓的偵測時間，Jean 以及 Wu [10]使用**時間差異法**(temporal differencing)找出移動的像素(如圖 1.5(b))。移動的像素和原始影像的邊緣作**交集運算**(AND operation)可得到移動的邊緣(如圖 1.5(c))，對移動邊緣作橢圓偵測即可找出人頭的部份(如圖 1.5(d))，由於移動的邊緣的像素不多，因此可以減少計算量。

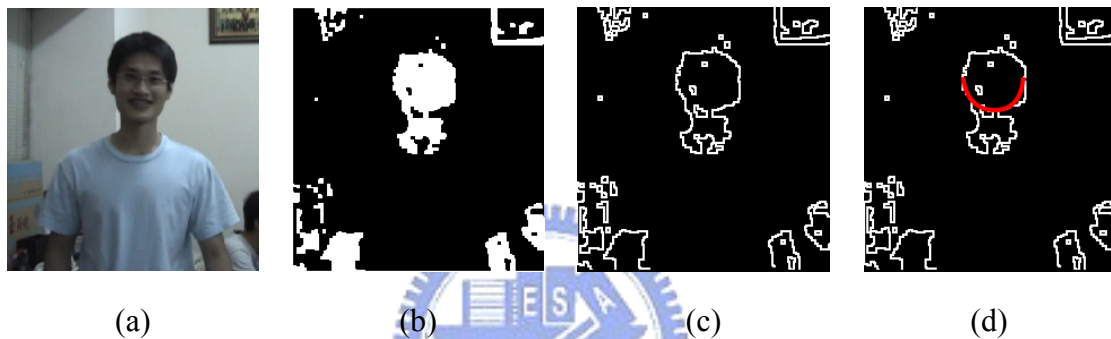


圖 1.4 利用膚色及邊緣偵測人臉的圖例。(a) 原始影像；(b) 膚色區域；(c) 邊緣影像；(d) 偵測出的人臉。

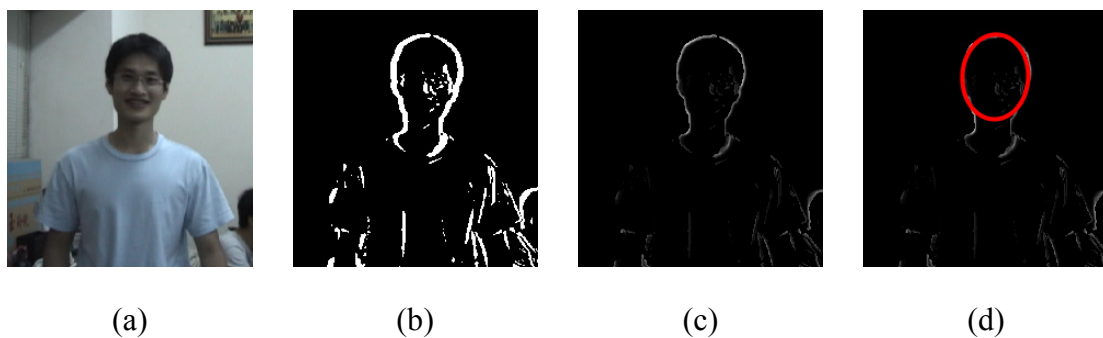


圖 1.5 利用移動的邊緣偵測人頭的圖例。(a) 原始影像；(b) 移動的影像；(c) 移動的邊緣；(d) 偵測出的人頭。

還有許多其它偵測行人的方法，例如使用多支架在不同位置、角度的攝影機拍攝行人以避免後面的行人被前面的行人擋住[11]；使用**立體相機**(stereo camera)

偵測行人與攝影機的距離以辨識行人[12]。另外紅外線影像(infrared image)等特殊的設備[13]可以用來偵測夜間的行人。

目前即時追蹤的方法主要是以**匹配為基礎的追蹤(matching-based tracking)**，對偵測出的物件萃取某些特徵，利用**相似性函數(likelihood function)**測量不同影像中兩物件特徵的相似度，如果相似度高代表在不同影像中的兩物件在真實世界是同一物件。Comaniciu 等人[14]利用顏色的直方圖作為特徵追蹤每塊區域的影像；Koller 等人[15]利用邊緣特徵追蹤車輛；Cham 等人[16]使用可變大小的樣板作比對，可以追蹤行人以及人的四肢；ISard 等人[17]使用**主動式輪廓模型(snake deformable model)**追蹤物件的輪廓。一種特別的追蹤方法則是根據**銜接式物件(articulated object)**的運動特性作追蹤，在[18]中顯示出不錯的效果。

1.1.2 人群計數

計算人群的行人數目是個充滿困難的議題，由於當行人數目過多時，部分的人會被遮蔽，因此不太可能得出真正的人群數目，又當背景相當複雜時，對於人的偵測更是困難。估測人群數目的方法在過去已被提出，一般的估測系統的架構通常可以分成兩部份，一部分是從輸入影像萃取出某些特徵值，另一部份則是根據取得的特徵值來估計人數或是對人群密度作分類。兩個早期的系統[19]以及[20]利用**閉路電視攝影機(closed circuit television camera, CCTV)**及影像處理的技巧實現即時人數估測，系統的方法步驟概述如下：

步驟一：計算 ROI(如圖 1.6(a)的紅色方框內)的前景(也就是影像中行人遮蔽背景的部份)的總像素個數作為特徵值(如圖 1.6(b))，或是萃取前景的邊緣總像素個數作為特徵值(如圖 1.6(c))。

步驟二：根據特徵值與實際人數的相對關係，得到兩者之間的關係函數(如圖 1.6(d)以及 1.6(e))。為了得到更好的線性關係函數，可以利用 Kalman 濾波器結合上述兩種特徵值求得最佳化的結果。

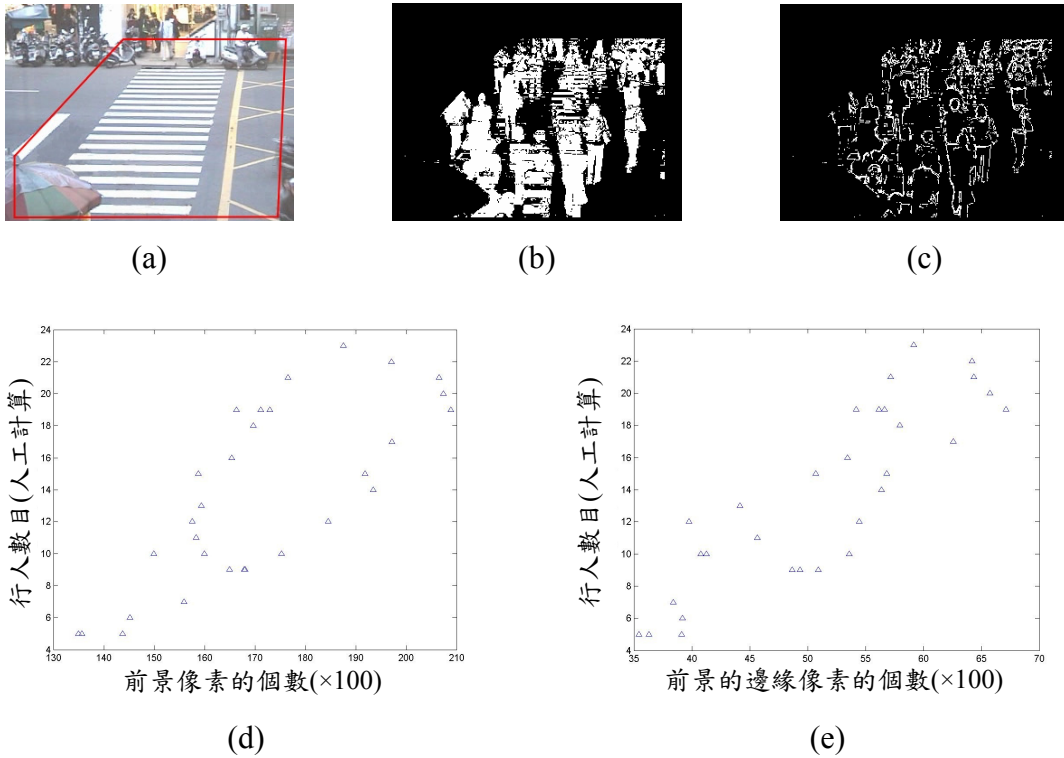


圖 1.6 利用影像處理技術估測人數的圖例。(a) 背景影像及 ROI；(b) 一個前景像素的例子；(c) 一個邊緣影像的例子；(d) 前景像素個數與實際人數的關係；(e) 前景的邊緣像素個數與實際人數的關係。

有許多研究利用其他的特徵，含有高密度人群的影像通常有精緻的圖樣 (pattern)，而低密度人群的影像則通常是粗糙的圖樣，這個技巧在[21]被使用，Marana 等人[21]使用灰階相依矩陣(gray level dependence matrix)對影像的紋路 (texture)作分析。在估測人數或分類人群多寡方面也有許多不同的方法，[22]使用模糊推論系統對行人數目作分類；[23]結合神經網路(neural networks)以及混合的全域學習(hybrid global learning)的演算法對行人數目作分類。

另外有一些不同於以上的方法。有些文獻的方法是基於直接對人作偵測，然後對偵測到的人作計數，類似 1.1.1 節提到的行人偵測；如前文提到的，這樣的方法通常無法偵測出太多人。Lin 等人 [24]使用小波轉換(wavelet transform)提取人臉的特徵以及使用支援向量機(support vector machine, SVM)作分類器可以較精確的偵測出較大量的人頭數目，並且只需要單張的影像，而不須用到背景；配合透視

轉換則可以估測幾百人甚至千人的行人數目；但這個方法由於計算量非常大，因此對目前的硬體設備而言是不可能用在即時的行人偵測與追蹤的系統。

1.2 研究動機

為了能夠偵測及追蹤行人穿越道的行人，一個具備即時與穩健特性的系統是被要求的。行人穿越道的行人在人多的情況下，通常會互相遮蔽，因此使用垂直投影的直方圖[7]以及行人邊界[8]只能偵測到少數分散在人群周圍的行人；[9]先提取影像中膚色的部分，再辨別是否是人臉，當攝影機由較人群高的位置往下拍攝時，人臉是較不易遮蔽的，但是在行人背對著攝影機時將無法用膚色偵測出人臉；[10]使用時間差異法減少橢圓形邊緣的偵測時間，但是當拍攝視野大的場景且有許多的前景物件時，相減的結果會是雜亂的圖樣而無法獲得有效的橢圓形邊緣，因此這個方法在拍攝行人穿越道的場景是不適合的。傳統用來估測行人數目的方法[19], [20]可以即時快速地完成，但無法有效地追蹤行人。另外，傳統的估測人群系統用在行人穿越道的場景會有三個主要的缺點：第一個缺點是由於特徵值與實際人數的關係函數需要由許多資料訓練而得，所以要得到一個關係函數相當的耗時，而每當換了一個場景或是拍攝角度不同就得重新訓練，這是相當不方便的；第二個缺點是由於交通環境充滿了許多非人的物件，而這些非人的物件會影響到特徵值，例如當車輛通過行人穿越道時會增加許多前景或邊緣；最後一個缺點實際上是原始系統原有的缺點，當鏡頭近端的行人的上半部位與鏡頭遠端的行人的下半部位分別投影在影像 ROI 的下端及上端時，系統會判斷出有行人，但在真實的世界中這些行人實際上可能不在行人穿越道上。

根據以上討論，一個即時與穩健的行人穿越道行人監控系統的確是面臨了許多挑戰。為了克服上述問題，本論文設計出結合了許多行人特性的模糊推論系統以辨識行人，並且提出一個區域偵測行人以及全域追蹤行人的構想以符合即時處理的要求，實驗結果顯示，在光線不是太暗而且行人穿越道不是太長的情形下，本論文所提出的方法對於行人穿越道的行人偵測及追蹤有不錯的效果。

1.3 論文架構

本篇論文剩餘章節的安排如下：論文使用到的一些影像處理技術以及模糊推論系統會在第二章作簡要的描述，行人穿越道行人偵測與追蹤系統會在第三章提出並且作詳細的說明，實驗結果及分析會詳列在第四章，最後第五章會作一些結論。



第二章

影像處理技術以及模糊推論系統

本篇論文使用到的影像處理技術以及模糊推論系統將會在此章節作說明，我們將以這些基礎理論發展出一個即時且穩健的行人穿越道行人監控系統，首先利用影像處理技術找出可能是行人的前景影像並萃取其特徵，再將特徵輸入一個模糊推論系統以辨識是否為行人。

一些影像處理的相關技術會在 2.1 節作個別的敘述，模糊推論系統會在 2.2 作簡單的介紹。



2.1 影像處理技術

包括**成像幾何學**(imaging geometry)、**邊緣檢測**、**色彩空間轉換**(color space conversion)、以及**形態學**的影像處理相關技術將會在此節說明。如何將相機影像平面上的點對應到真實世界的三維點上，對真實世界物件的追蹤而言是相當重要的，為了達到這個目的，必須找出相機的內、外部參數，這些計算可由成像幾何學完成。邊緣檢測可以用來找出人頭的邊緣。利用色彩空間轉換可以得到適當的膚色表示方式。形態學運算可以找出行人的輪廓。

成像幾何學會在 2.1.1 節探討，邊緣檢測在 2.1.2 節講解，色彩空間轉換在 2.1.3 節敘述，形態學在 2.1.4 節說明。

2.1.1 成像幾何學

最簡單的相機成像原理可用**針孔相機模型**(pinhole camera model)來說明[25]，它是一個理想的相機模型，如圖 2.1。藉由相似三角形，可以找出三維**世界座標系統**(world coordinate system, WCS)與相機**影像平面**(image plane)座標之間的關係，如圖 2.2，假設一個三維座標點為 (x, y, z) ，而它在影像平面上的投影座標為 (u, v) ，則三維及二維座標之間的關係可表示如下：

$$\begin{aligned}\frac{u}{f} &= \frac{x}{z}, \\ \frac{v}{f} &= \frac{y}{z},\end{aligned}\tag{2.1}$$

其中 f 為相機的**焦距**(focal length)。

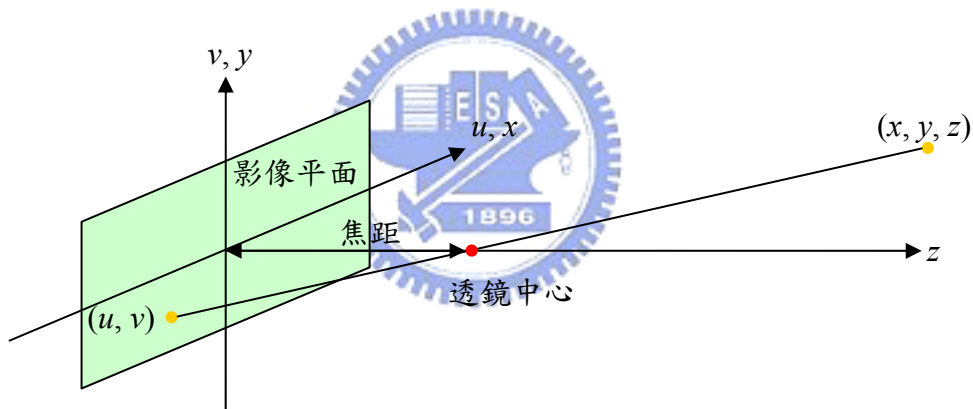


圖 2.1 針孔相機模型。

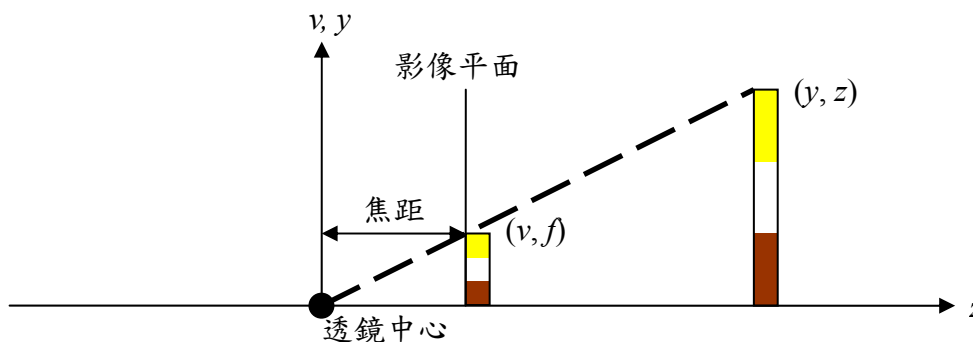


圖 2.2 相似三角形模型。

相機校正(camera calibration)的主要工作是要找出相機的內、外部參數。內部參數包含相機的焦距、主要點(principal point)、歪斜係數(skew coefficient)、以及失真[25]，這些參數值在拍攝時是固定的。外部參數包含 WCS 與相機座標系統(camera coordinate system, CCS)之間的位移及旋轉，這些參數值會隨著相機位置的不同而改變。三維及二維座標之間的關係可由一個 3×4 的透視投影矩陣(perspective transformation matrix) M 表示，令 m_1 、 m_2 、 m_3 為 M 的三個列向量，令 $P=[x \ y \ z \ 1]$ 為 WCS 中一點的齊次座標(homogeneous coordinate)的向量形式，則它在影像平面上的座標(u, v)可以表示為：

$$\begin{aligned} u &= (m_1 \cdot P) / (m_3 \cdot P), \\ v &= (m_2 \cdot P) / (m_3 \cdot P). \end{aligned} \quad (2.2)$$

M 是一個由十一個相機參數所決定的矩陣，在此介紹一個簡化的模型[26]，只使用七個相機參數：焦距 f 、主要點 $p(u_0, v_0)$ 、平掃(pan)的角度 α 、俯仰(tilt)的角度 β 、平擺(yaw)的角度 γ 、以及 C_O 與 O 之間的距離 H_c ，如圖 2.3 所示，假設系統使用的是針孔相機模型，亦即歪斜係數為零以及無失真，令 CCS(正交座標軸 C_X 、 C_Y 、以及 C_Z)的原點 C_O (即透鏡中心)位於 WCS(正交座標軸 X 、 Y 、以及 Z)的原點 O 的垂直正上方，也就是 C_O 在 Y 軸上，其中 X - Z 平面我們稱為地平面(ground plane)，而 α 為對著 Y 軸逆時鐘旋轉的角度， β 為對著 X 軸順時鐘旋轉的角度， γ 為對著 Z 軸順時鐘旋轉的角度， M 可以被簡化為：

$$\begin{aligned} M &= \begin{bmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}, \\ R &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} c\alpha \cdot c\gamma + s\alpha \cdot s\beta \cdot s\gamma & -c\beta \cdot s\gamma & -s\alpha \cdot c\gamma + c\alpha \cdot s\beta \cdot s\gamma \\ c\alpha \cdot s\gamma - s\alpha \cdot s\beta \cdot c\gamma & c\beta \cdot c\gamma & -s\alpha \cdot s\gamma - c\alpha \cdot s\beta \cdot c\gamma \\ s\alpha \cdot c\beta & s\beta & c\alpha \cdot c\beta \end{bmatrix}, \quad (2.3) \\ T &= [t_1 \ t_2 \ t_3]^T = [0 \ H_c \ 0]^T, \end{aligned}$$

其中 $c\alpha$ 、 $s\alpha$ 、 $c\beta$ 、 $s\beta$ 、 $c\gamma$ 、以及 $s\gamma$ 分別代表 $\cos(\alpha)$ 、 $\sin(\alpha)$ 、 $\cos(\beta)$ 、 $\sin(\beta)$ 、 $\cos(\gamma)$ 、以及 $\sin(\gamma)$ 。

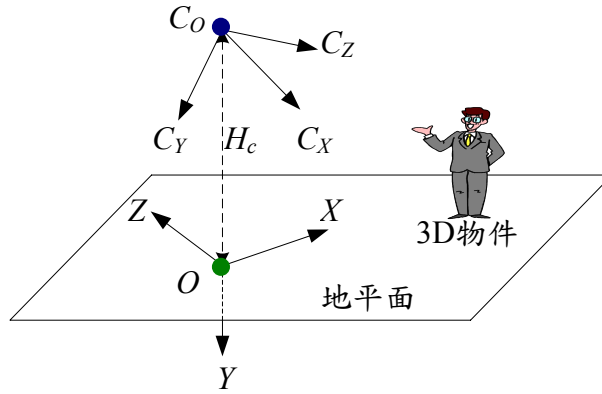


圖 2.3 WCS 以及 CCS 的相對位置。

根據[26]，我們可以得到四個與成像幾何學相關的特性：

特性一：在 WCS 中的任意一組互相平行的直線其投影在影像平面上的影像會是通過一個共同點的直線群，這個點稱為它們的**消失點**(vanishing point)。由式子(2.2)以及(2.3)，可以得到消失點在影像平面的座標：

$$\left(f \cdot \frac{r_{11}a + r_{12}b + r_{13}c}{r_{31}a + r_{32}b + r_{33}c} + u_0, f \cdot \frac{r_{21}a + r_{22}b + r_{23}c}{r_{31}a + r_{32}b + r_{33}c} + v_0 \right), \quad (2.4)$$

其中 $[a \ b \ c]$ 是這組平行線的方向向量(direction vector)。假設我們以 WCS 的三個正交座標軸 X 、 Y 、以及 Z 分別作為三組平行線的方向向量，亦即令方向向量 $[a \ b \ c]$ 分別為 $[1 \ 0 \ 0]$ 、 $[0 \ 1 \ 0]$ 、以及 $[0 \ 0 \ 1]$ ，則三個方向的消失點座標 V_x 、 V_y 、以及 V_z 由式子(2.4)可以分別得到為：

$$\begin{aligned} V_x = (u_{V_x}, v_{V_x}) &= \left(f \cdot \frac{r_{11}}{r_{31}} + u_0, f \cdot \frac{r_{21}}{r_{31}} + v_0 \right), \\ V_y = (u_{V_y}, v_{V_y}) &= \left(f \cdot \frac{r_{12}}{r_{32}} + u_0, f \cdot \frac{r_{22}}{r_{32}} + v_0 \right), \\ V_z = (u_{V_z}, v_{V_z}) &= \left(f \cdot \frac{r_{13}}{r_{33}} + u_0, f \cdot \frac{r_{23}}{r_{33}} + v_0 \right), \end{aligned} \quad (2.5)$$

我們稱這三個消失點為互相正交的消失點。

特性二：以三個互相正交的消失點為端點形成的三角形，其**垂心**(orthocenter)即為主要點，如圖 2.4(b)所示。

特性三：在 WCS 中任意一平面上的任意許多組平行線，其消失點會在影像平面上形成一條直線，稱為**消失線**(vanishing line)。如果這些平行線都平行地平面，則其形成的消失線稱為**地平線**(horizon line)，如圖 2.4(b)所示。

特性四：水平線與地平線的夾角(圖 2.4(b)的 θ)等於平擺的角度 γ 。

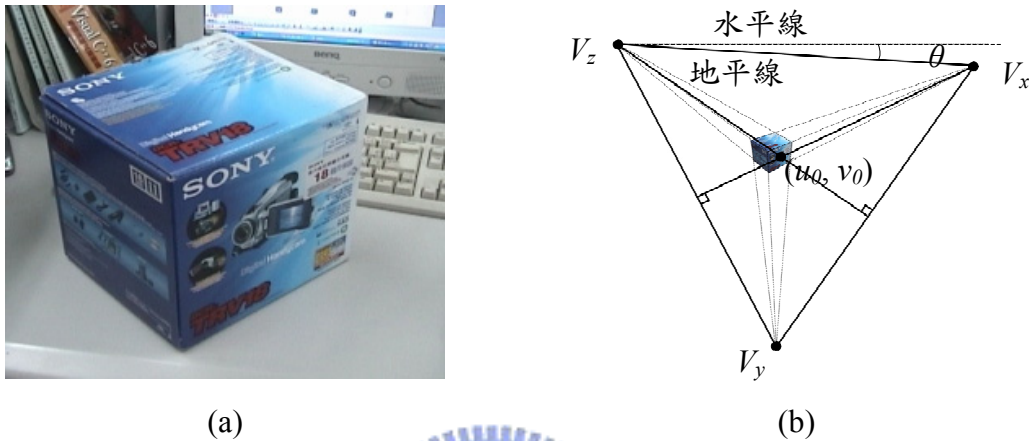


圖 2.4 利用方塊物件求取相機參數的圖例。(a) 原始影像；(b) 影像的消失點、主要點、以及地平線。



以下逐步列出求取七個相機參數的步驟：

步驟一：利用 WCS 中的三組互相正交的平行線，求出三個互相正交的消失點 V_x 、 V_y 、以及 V_z 。

步驟二：由上述的特性二算出主要點座標。

步驟三：由上述的特性四算出平擺的角度 γ 。

步驟四：如果 γ 等於零，由式子(2.3)以及(2.5)可以得到：

$$u_{V_x} = f \cdot \cot(\alpha) / \cos(\beta) + u_0, \quad (2.6.1)$$

$$v_{V_x} = -f \cdot \tan(\beta) + v_0, \quad (2.6.2)$$

$$u_{V_y} = u_0, \quad (2.6.3)$$

$$v_{V_y} = f \cdot \cot(\beta) + v_0, \quad (2.6.4)$$

$$u_{V_z} = -f \cdot \tan(\alpha) / \cos(\beta) + u_0, \quad (2.6.5)$$

$$v_{V_z} = -f \cdot \tan(\beta) + v_0, \quad (2.6.6)$$

(2.6)

由式子(2.6.2)以及(2.6.4)，求得 f 為：

$$f = \sqrt{-(v_{V_x} - v_0)(v_{V_y} - v_0)}, \quad (2.7)$$

由式子(2.6.2)以及(2.7)，求得 β 為：

$$\beta = \tan^{-1}((v_0 - v_{V_x}) / f), \quad (2.8)$$

由式子(2.6.1)、(2.7)、以及(2.8)，求得 α 為：

$$\alpha = \cot^{-1}((u_{V_x} - u_0) \cos(\beta) / f), \quad (2.9)$$

根據交叉比不變性(cross ratio invariance)[27]，可以得到

$$\frac{H}{H_c} = 1 - \frac{d(x', c) \cdot d(x, v)}{d(x, c) \cdot d(x', v)}, \quad (2.10)$$

如圖 2.5，其中 v 是垂直地平面方向的消失點， H 是參考物件在 WCS 的高度， $d()$ 為兩個座標之間的距離，由式子(2.10)可求得 H_c ；如果 γ 不等於零，則將影像繞著影像原點旋轉 γ 角度，使水平線和地平線互相平行，得到的新影像平擺角即為零，因此可以利用以上步驟求得 f' 、 u'_0 、 v'_0 、 α' 、 β' 、 $\gamma' = 0$ 、以及 H'_c ，則原始影像的參數為 $f = f'$ 、 $u_0 = u'_0 \cos \gamma - v'_0 \sin \gamma$ 、 $v_0 = u'_0 \sin \gamma - v'_0 \cos \gamma$ 、 $\alpha = \alpha'$ 、 $\beta = \beta'$ 、 $\gamma = \gamma$ 、以及 $H_c = H'_c$ 。

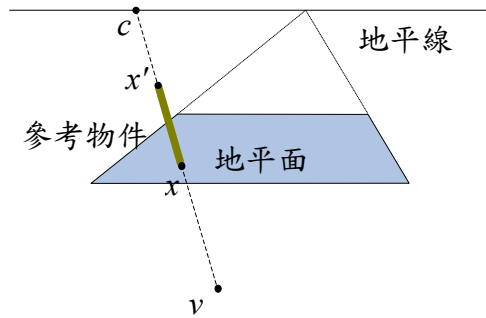


圖 2.5 計算相機位置高度的示意圖。

2.1.2 邊緣檢測

邊緣是二個區域之間有相當不同的灰階特徵的邊界。大多數邊緣檢測技術的基本思想是利用導數運算子的運算[28]，圖 2.6 說明了這個概念。圖 2.6(a)是一個

在暗色背景下有亮條紋的影像，圖 2.6(b)是沿著水平掃描線的灰階剖面圖，圖 2.6(c)是剖面圖的一階導數，在一階導數急遽變化的位置顯示出邊緣的所在。

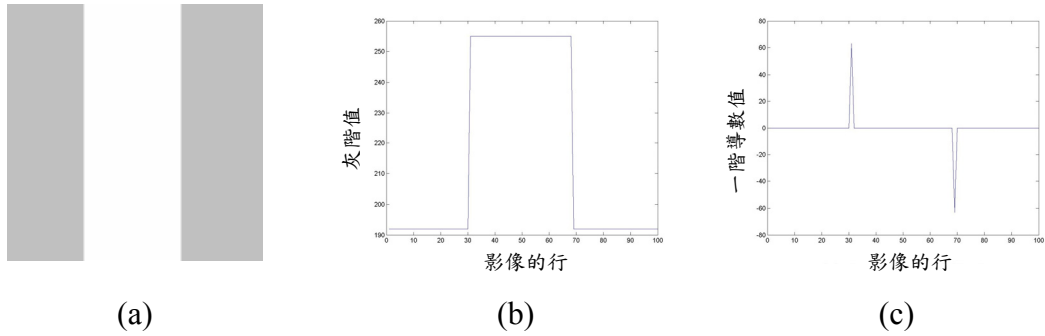


圖 2.6 一個應用導數運算子作邊緣檢測的例子。(a) 灰階影像；(b) 沿著水平掃描線的灰階剖面圖；(c) 灰階剖面圖的一階導數。

導數運算一般是利用空間濾波(spatial filtering)來實現。空間濾波是空間遮罩(spatial mask)在影像處理中的應用(對應的是使用傅立葉轉換等的頻率域濾波器)，而遮罩本身稱為空間濾波器(spatial filter)。空間濾波器可分類為線性濾波器以及非線性濾波器，線性濾波器的基本方法是求遮罩係數和影像中遮罩下面特定位置上像素灰階乘積之和。圖 2.7 給出一個常用的 3×3 遮罩，若用 g_1, g_2, \dots, g_9 表示遮罩下面像素的灰階，則線性遮罩的響應是：

$$Res_s = w_1 g_1 + w_2 g_2 + \dots + w_9 g_9, \quad (2.11)$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

圖 2.7 具有任意係數的一個 3×3 遮罩。

影像處理應用中最常使用來求導數的方法是梯度(gradient)，對於一個函數 $f(x, y)$ ， f 在 (x, y) 處的梯度定義為一個向量：

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}, \quad (2.12)$$

該向量的大小為：

$$|\nabla f| = [G_x^2 + G_y^2]^{1/2}, \quad (2.13)$$

在實際的應用上，向量的大小通常近似為絕對值的和：

$$|\nabla f| = |G_x| + |G_y|, \quad (2.14)$$

梯度 ∇f 在 (x, y) 處的方向角，由向量分析可知：

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_x}{G_y}\right), \quad (2.15)$$

其中 α 角是指和 x 軸的夾角。式子(2.12)中的導數可以用數種數位形式的線性濾波器來近似求得，本篇論文採用 Sobel 空間濾波器。Sobel 運算子具有能提供既有差值又有平滑效果的優點，因為導數會增強雜訊，Sobel 運算子的平滑效應的特點就特別具有吸引力。考慮圖 2.8(a)所示影像區域，位於 z_5 基於 Sobel 遮罩的導數為：

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3), \quad (2.16.1) \quad (2.16)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7), \quad (2.16.2)$$

圖 2.8(b)以及 2.8(c)分別為 x 以及 y 方向的 Sobel 遮罩。

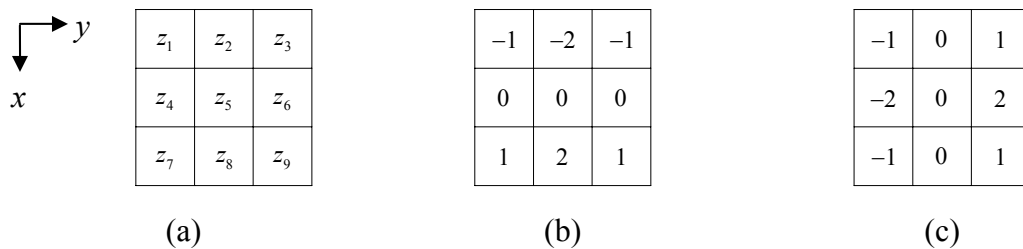


圖 2.8 一個 3×3 的影像區域和 Sobel 遮罩。(a) 影像區域；(b) 用來計算 G_x 的遮罩；(c) 用來計算 G_y 的遮罩。

以上討論的是灰階影像的邊界提取，當以上的方法要使用在彩色影像時必須先轉成灰階影像，圖 2.9 顯示出一個邊緣檢測的例子，圖 2.9(a)是原始彩色影像，

圖 2.9(b)是圖 2.9(a)轉成灰階後的影像，圖 2.9(c)是利用 Sobel 遮罩找出的邊緣影像。圖 2.9(c)顯示出原本在彩色影像中的垂直邊緣無法被檢測出來，這是由於彩色影像轉成灰階影像後，垂直的邊緣消失了。為了解決以上的問題，必須將 Sobel 遮罩分別應用在 R、G、以及 B 的色彩空間。由於對彩色影像做邊緣檢測的研究相當少，因此根據 R、G、以及 B 色彩空間分別得到的導數以求得最後合成的導數方向及大小的方法並沒有一個常用的標準，在此介紹[29]的方法，首先定義

$$J = \begin{bmatrix} Gr_x & Gr_y \\ Gg_x & Gg_y \\ Gb_x & Gb_y \end{bmatrix} = [C_x^T \quad C_y^T], \quad (2.17)$$

其中

$$Gr_x = \frac{\partial R}{\partial x}, \quad (2.18)$$

$$Gr_y = \frac{\partial R}{\partial y},$$

同理可得 Gg_x 、 Gg_y 、 Gb_x 、以及 Gb_y ，則梯度的方向滿足下列式子：

$$\tan(2\theta) = \frac{2 \cdot C_x \cdot C_y}{\|C_x\|^2 - \|C_y\|^2}, \quad (2.19)$$

梯度的大小滿足下列式子：

$$m^2 = \|C_x\|^2 \cos^2(\theta) + 2 \cdot C_x \cdot C_y \cdot \sin(\theta) \cos(\theta) + \|C_y\|^2 \sin^2(\theta), \quad (2.20)$$

圖 2.9(d)是根據[29]求得的邊緣影像。



圖 2.9 應用邊緣檢測的結果。(a) 原始彩色影像；(b) 灰階影像；(c) 灰階影像的邊緣；(d) 彩色影像的邊緣。

2.1.3 色彩空間轉換

一張彩色影像可以使用許多不同的彩色模型來呈現，常見的彩色模型有 RGB、HSV、以及 YCbCr[28]，本篇論文將會使用到從 RGB 色彩模型轉換到 HSV 以及 YCbCr 色彩模型的色彩空間轉換。

RGB 彩色模型是由三種原色組成，分別為紅色、綠色、以及藍色，且 RGB 彩色模型是架構在直角座標系統(Cartesian coordinate system)，如圖 2.10。一般由感光耦合元件(charge coupled device, CCD)擷取得到的影像即是用 RGB 彩色模型。

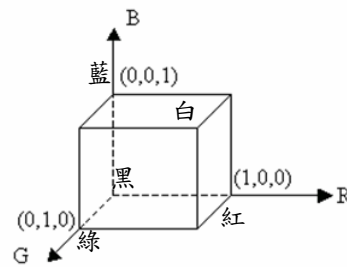


圖 2.10 RGB 彩色模型。

HSV 彩色模型很適合人類描述一種顏色，因為人類描述一種顏色是以色調 (hue)、飽和度(saturation)、以及強度(brightness)來表示。本篇論文會用到強度 V 以及飽和度 S ，定義如下：

$$V = (R + B + G) / 3, \quad (2.21.1)$$

$$S = 1 - \frac{3}{(R + B + G)} \cdot \min(R, G, B), \quad (2.21.2) \quad (2.21)$$

YCbCr 彩色模型將色彩資訊分離出亮度(luminance)與彩度(chrominance)，其中 Y 為亮度， Cb 以及 Cr 是兩個彩度值。給定一個 RGB 像素，則根據 CCIR601 標準的定義，它的 Y 、 Cb 、以及 Cr 分別由下列式子算出：

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B, \quad (2.22.1)$$

$$Cb = -0.172 \cdot R - 0.339 \cdot G + 0.511 \cdot B + 128, \quad (2.22.2) \quad (2.22)$$

$$Cr = 0.511 \cdot R - 0.428 \cdot G - 0.083 \cdot B + 128, \quad (2.22.3)$$

以上 R 、 G 、以及 B 代表像素的紅色、綠色、以及藍色，且值的大小正規化到 $[0,255]$ 。

2.1.4 形態學

形態學[28]有許多各式各樣的運算可以用來處理物體的外形及結構。這些運算是使用在二值化影像。**侵蝕(erosion)**及**膨脹(dilation)**是兩種最基本的運算，**斷開(opening)**及**閉合(closing)**運算是侵蝕及膨脹這兩種運算的延伸，這四種運算會在此章節作說明，另外也會對邊界提取以及連通分量的提取作說明。

首先介紹一些基本定義，設 A 和 B 是 Z^2 中的兩個集合，它們的分量分別為 $a=(a_1,a_2)$ 以及 $b=(b_1,b_2)$ ，則 A 平移 $x=(x_1, x_2)$ ，用 $(A)_x$ 表示之，定義為

$$(A)_x = \{c | c = a + x, a \in A\}, \quad (2.23)$$

B 的反射，用 \hat{B} 表示之，定義為

$$\hat{B} = \{x | x = -b, b \in B\}, \quad (2.24)$$

A 的補集是

$$\bar{A} = \{x | x \notin A\}, \quad (2.25)$$

A 和 B 的差集表示為 $A-B$ ，定義為

$$\begin{aligned} A-B &= \{x | x \in A, x \notin B\} \\ &= A \cap \bar{B}. \end{aligned} \quad (2.26)$$

假設一影像的集合為 A ，且有一結構元素的集合為 B ，則 A 被 B 侵蝕記為 $A \ominus B$ ，定義為

$$A \ominus B = \{x | (B)_x \subseteq A\}, \quad (2.27)$$

圖 2.11 是一個侵蝕的例子，圖 2.11(a) 中灰色部分是影像的集合 A ，圖 2.11(b) 中灰色部分是結構元素的集合 B ，圖 2.11(c) 中灰色部份是侵蝕後的集合。

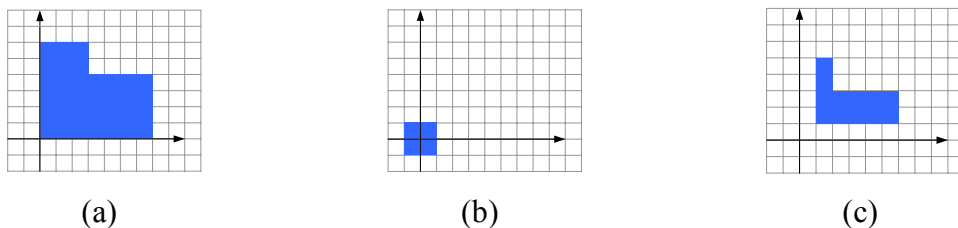


圖 2.11 侵蝕運算。(a) 侵蝕前影像；(b) 結構元素；(c) 侵蝕後影像。

假設一影像的集合為 A ，且有一結構元素的集合為 B ，則 A 被 B 膨脹記為 $A \oplus B$ ，定義為

$$A \oplus B = \{x | (\hat{B})_x \cap A \neq \Phi\}, \quad (2.28)$$

圖 2.12 是一個膨脹的例子，圖 2.12(a) 中灰色部分是影像的集合 A ，圖 2.12(b) 中灰色部分是結構元素的集合 B ，圖 2.12(c) 中灰色部份是膨脹後的集合。

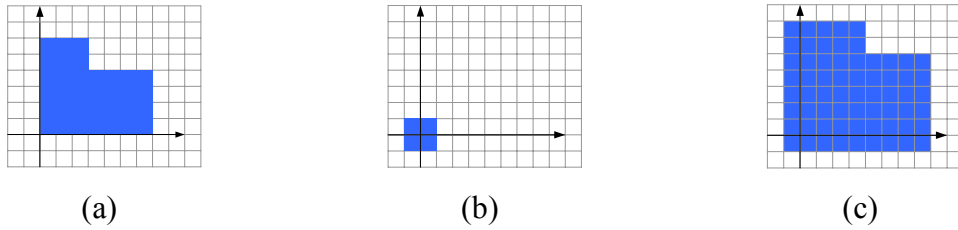


圖 2.12 膨脹運算。(a) 膨脹前影像；(b) 結構元素；(c) 膨脹後影像。

影像的斷開運算是由侵蝕運算及膨脹運算所組成。斷開運算可以定義成：

$$A \circ B = (A \ominus B) \oplus B, \quad (2.29)$$

其中 A 是影像的集合， B 是結構元素，它代表的意思為影像先經過侵蝕後再膨脹。斷開運算可以用來消除物體突出的地方及雜訊。

閉合運算剛好與斷開運算相反，它可以定義成：

$$A \bullet B = (A \oplus B) \ominus B, \quad (2.30)$$

其中 A 是影像的集合， B 是結構元素，它代表的意思為影像先經過膨脹後再侵蝕。閉合運算可以用來消除影像中的空洞及填滿空隙。

一個影像的集合 A 的邊界集合，用 $\beta(A)$ 表示之，可以這樣得到：先用結構元素的集合 B 侵蝕 A ，然後形成 A 和它的侵蝕之間的差集，即

$$\beta(A) = A - (A \ominus B), \quad (2.31)$$

圖 2.13 是一個邊界提取的例子，圖 2.13(a) 中灰色部分是影像的集合 A ，圖 2.13(b) 中灰色部分是結構元素的集合 B ，圖 2.13(c) 中灰色部份是提取的邊界。

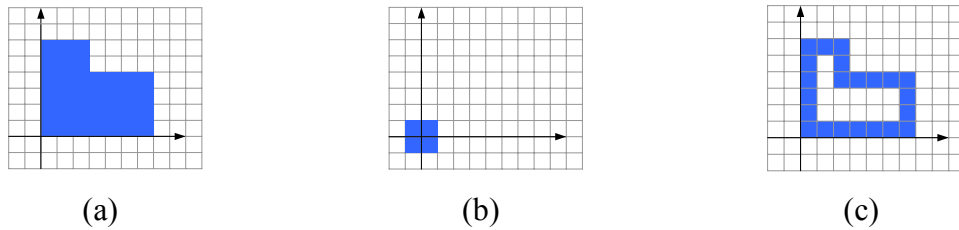


圖 2.13 邊界提取。(a) 原始影像；(b) 結構元素；(c) 提取的邊界。

假設 p 以及 q 是影像子集 A 中的像素，如果存在一條從 p 到 q 的路徑，其像素全部由 A 中的像素組成，那麼 p 與 q 在 A 中是連通的。對 A 中的任意像素 p ，與 p 相連通的 A 中的像素集合稱為 A 的連通分量。假設 Y 表示包含於集合 A 的一個連通分量，並假定已知 Y 的一個點 p ，則下面的疊代表示式給出 Y 的所有元素：

$$X_k = (X_{k-1} \oplus B) \cap A, \quad k = 1, 2, 3, \dots, \quad (2.32)$$

其中 $X_0 = p$ ， B 是一個合適的結構元件，不同的連通路徑定義使用不同的 B 。如果 $X_k = X_{k-1}$ ，則演算法收斂，並令 $Y = X_k$ 。



2.2 模糊推論系統

模糊推論系統已廣泛地應用在於許多領域，如自動控制、圖形識別、以及決策分析，使用模糊系統有兩個主要的優點，第一個優點是不需要精確的數學模型，第二個優點是可以將人類專家的知識結合至系統的判斷上。一般來說，模糊系統的基本架構如圖 2.14 所示，可以分為四個主要的功能方塊[30]，包括：1) **模糊化機構**(fuzzifier)；2) **模糊規則庫**(fuzzy rule base)；3) **模糊推論引擎**(fuzzy inference engine)；以及 4) **去模糊化機構**(defuzzifier)。模糊化機構的功能為將明確的外界輸入資料轉換成適當的語意式模糊資訊；模糊規則庫存放解決相關問題所需的知識及規則；模糊推論引擎藉由模糊推論的方式來模擬人類的思考決策模式以解決面對的問題；去模糊化機構的功能是將模糊推論引擎所推出的模糊資訊轉回外界的明確數值。以上四個功能方塊的詳細解說如後。

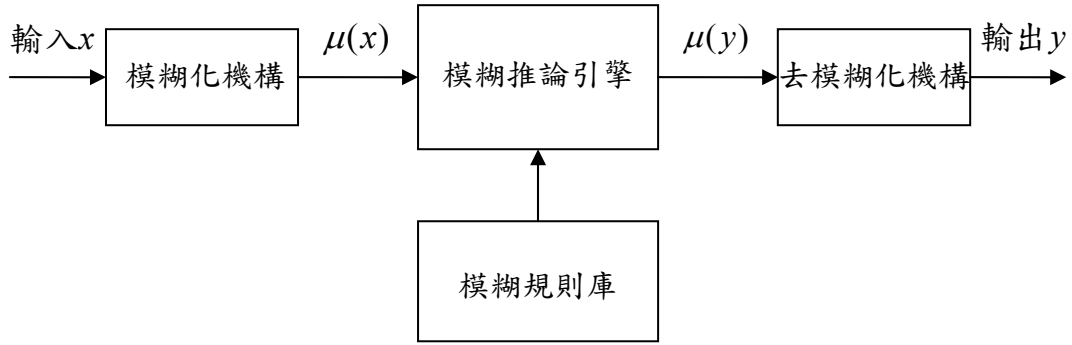


圖 2.14 模糊系統的基本架構。

2.2.1 模糊化機構

模糊化機構可以被視為一種映射，由明確的輸入空間映射至特定模糊集合空間[30]。我們可以定義在論域(universe of discourse) U 中的一個模糊集合(fuzzy set) A 為：

$$A = \{(x, \mu_A(x)) \mid x \in U\}, \quad (2.33)$$

其中 $\mu_A(x)$ 是模糊集合 A 的歸屬函數(membership function)， $\mu_A(x)$ 代表論域 U 中元素 x 對模糊集合 A 的歸屬程度。通常歸屬函數的值是介於0與1之間，三個常見的歸屬函數 $\mu_A(x)$ 、 $\mu_B(x)$ 、以及 $\mu_C(x)$ 如圖 2.15，並定義如下：

1. Z型歸屬函數：

$$\mu_A(x) = \begin{cases} 1, & \text{if } x \leq x_0, \\ \frac{x-x_a}{x_0-x_a}, & \text{if } x_0 \leq x \leq x_a, \\ 0, & \text{if } x \geq x_a, \end{cases} \quad (2.34)$$

其中輸入為 x_0 時的歸屬值(membership value)為 $\mu_A(x_0)=1$ ，如圖 2.15(a)所示。

2. 三角形歸屬函數：

$$\mu_B(x) = \begin{cases} 0, & \text{if } x \geq x_a \text{ or } x \leq x_b, \\ \frac{x-x_b}{x_0-x_b}, & \text{if } x_b \leq x \leq x_0, \\ \frac{x-x_a}{x_0-x_a}, & \text{if } x_0 \leq x \leq x_a, \end{cases} \quad (2.35)$$

其中 $\mu_B(x_0)=1$ ，如圖 2.15(b)所示。

3. S 型歸屬函數：

$$\mu_C(x) = \begin{cases} 0, & \text{if } x \leq x_b, \\ \frac{x-x_b}{x_0-x_b}, & \text{if } x_b \leq x \leq x_0, \\ 1, & \text{if } x \geq x_0, \end{cases} \quad (2.36)$$

其中 $\mu_C(x_0)=1$ ，如圖 2.15(c)所示。

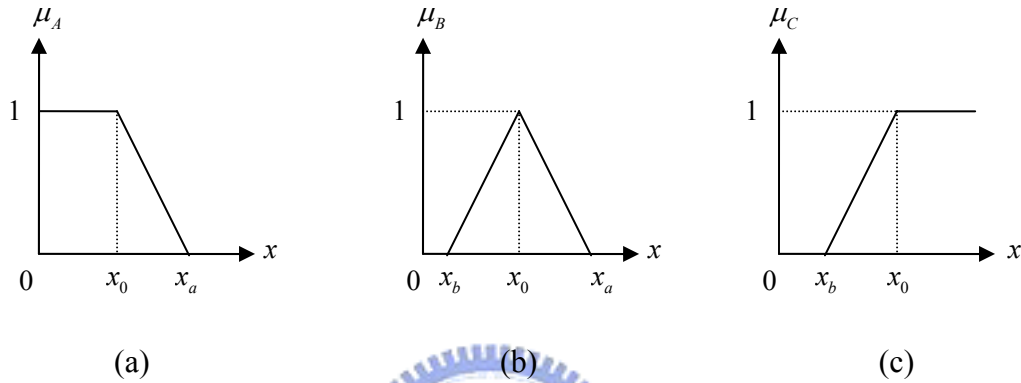


圖 2.15 歸屬函數。(a) Z 型；(b) 三角形；(c) S 型。

2.2.2 模糊規則庫

模糊規則庫是由一組以 IF-THEN 形式的模糊規則所組成，這組模糊規則是用以描述系統的輸入與輸出的關係[30]。為了說明模糊規則，先定義語意式變數 (linguistic variable)： $(x, U, \{T^1, T^2, \dots, T^k\}, \{\mu^1, \mu^2, \dots, \mu^k\})$ ，其中 x 是變數的名稱； U 是論域； T^i 是語意式變數的語意值 (linguistic value)，令 $T(x) = \{T^1, T^2, \dots, T^k\}$ ，則 $T(x)$ 稱為 x 的措詞集 (term set)； μ^i 是相對於 T^i 的歸屬函數。如果 x 是溫度，論域 U 定義於 $[0, 60]$ ，則措詞集可以是 $T(x) = \{\text{低}, \text{適中}, \text{高}\}$ 。

最常見的模糊規則型式有三種：語意式模糊規則、函數式模糊規則、以及 Tsukamoto 模糊規則，以下列出本論文使用的語意式模糊規則：

$$R^i: \text{IF } x \text{ is } A_i, \dots, \text{AND } y \text{ is } B_i, \text{ THEN } z \text{ is } C_i, \quad i=1, 2, \dots, n, \quad (2.37)$$

其中 x, \dots, y 以及 z 是語意式變數， A_i, \dots, B_i 以及 C_i 是相對於 x, \dots, y 以及 z 的語意

值， x, \dots, y 以及 z 的論域分別是 U, \dots, V 以及 W 。

2.2.3 模糊推論引擎

模糊推論引擎是模糊系統的核心，它可以藉由近似推論或模糊推論的進行得到一推論出來的模糊資訊，用來模擬人類的思考決策[30]。以下藉由一個簡單的例子說明模糊推論的過程：


輸入： x is A and y is B

模糊規則 R^1 ：IF x is A_1 AND y is B_1 , THEN z is C_1

模糊規則 R^2 ：IF x is A_2 , AND y is B_2 , THEN z is C_2 (2.38)

輸出： z is C

假設兩輸入的模糊集合皆為**模糊單點**(fuzzy singleton)，兩個模糊規則的**啟動強度**(firing strength)可以分別表示為：


$$\begin{aligned}\alpha_1 &= \mu_{A_1}(x_0) \wedge \mu_{B_1}(y_0) \quad \text{and} \\ \alpha_2 &= \mu_{A_2}(x_0) \wedge \mu_{B_2}(y_0),\end{aligned} \quad (2.39)$$

根據 Mamdani 所提出的**最小值推論法**(minimum implication rule)，最後推論出模糊集合 C ，其歸屬函數為：

$$\mu_c(z) = [\alpha_1 \wedge \mu_{C_1}(z)] \vee [\alpha_2 \wedge \mu_{C_2}(z)], \quad (2.40)$$

以上的 \wedge 及 \vee 在此分別代表取最小值運算及取最大值運算。以上模糊推論過程如圖 2.16。

2.2.4 去模糊化機構

去模糊化機構是將經過模糊推論之後產生的結論轉換為一明確數值的過程[30]。採用不同的模糊規則，推論出來的結果型式會不同，語意式模糊規則經由模

模糊推論引擎推論出來的結果是一模糊集合。常見的去模糊化的方法是**中心面積法**(center of area defuzzifier, COA)以及**最大平均法**(mean of maximum defuzzifier, MOM)。本篇採用中心面積法，當論域為離散時，其明確輸出值由以下式子得出：

$$z_{COA} = \frac{\sum_{i=1}^N \mu_C(z_i) \cdot z_i}{\sum_{i=1}^N \mu_C(z_i)}, \quad (2.41)$$

其中 N 代表輸出的**量化數**(quantization level)， z_i 代表第 i 個量化值，以及 $\mu_C(z_i)$ 代表 z_i 屬於模糊集合 C 的歸屬值。

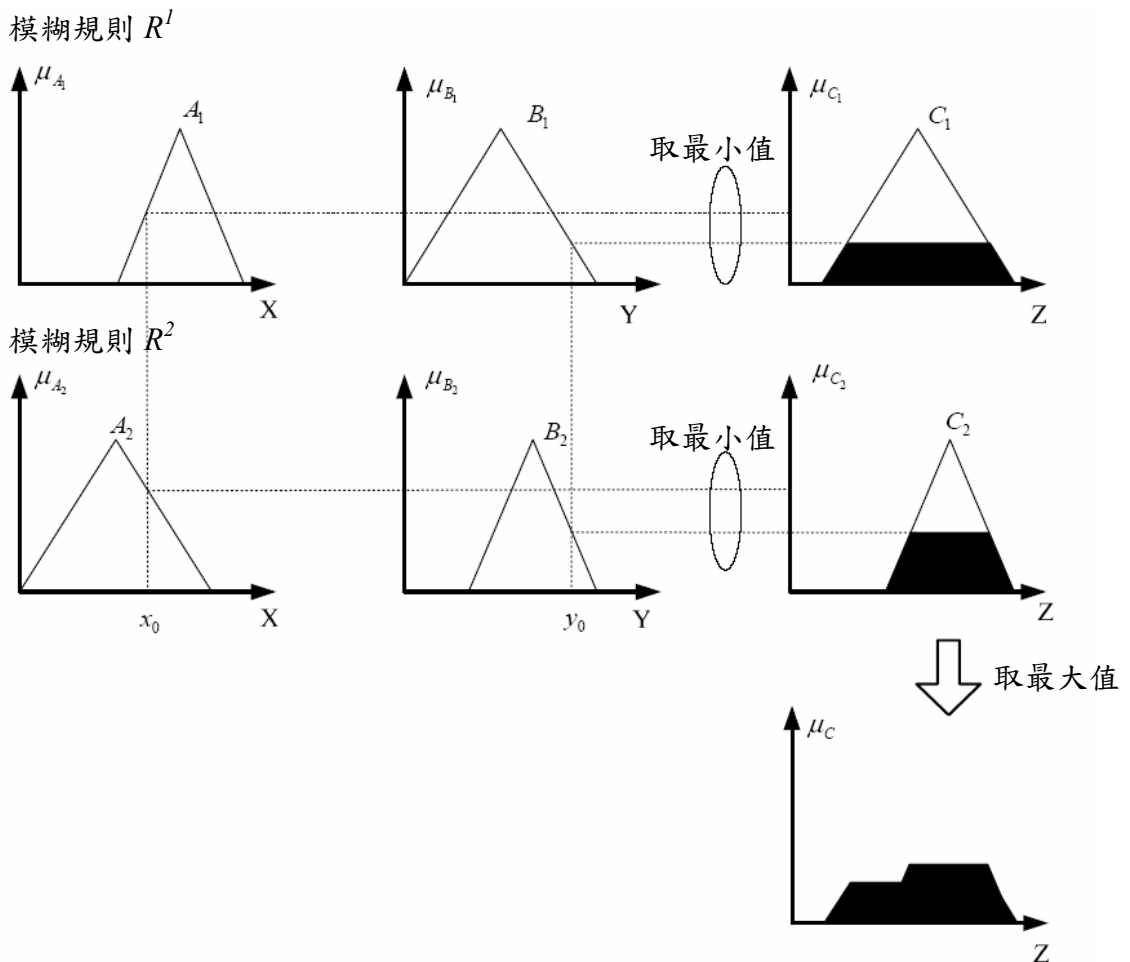


圖 2.16 一個模糊推論過程的例子。

第三章

行人監控系統

即時的行人偵測與追蹤是相當受到重視的論題，由第一章的探討可以知道目前存在許多的相關文獻，但是大部分的方法無法有效處理互相交疊的行人，對於行人交錯嚴重的行人穿越道場景而言這些方法是不可行的，因此本論文設計出一個行人監控系統，可以有效地處理這樣的問題，詳細的內容將在此章節說明。

在章節 3.1 中會概述整個行人監控系統的流程與原理。在 3.2 節中會介紹相機校正的方法。在 3.3 節會探討橢圓形邊緣的偵測方法。在 3.4 節中會介紹**色彩校正** (color correction)。在 3.5 節會講解用來辨識行人的模糊推論系統。在 3.6 節說明輔助的行人偵測系統。行人的追蹤會在 3.7 節敘述。

3.1 系統概述

圖 3.1 為行人監控系統的架構，整個系統的流程可以分成六個步驟。步驟一為相機校正，利用行人穿越道的圖形特性可以很容易地得到相機的內、外部參數，藉由相機的內、外部參數可以求得真實世界中的物件投影在影像平面上的長度，也可以求得影像中的物件在真實世界中的位置及長度，這些資訊將會用在行人的辨識與追蹤。對於每一個行人穿越道的場景而言，相機校正只需要手動地建構一次，整個行人監控系統有關於影像長度的參數即可根據相機的內外部參數自動建立。步驟二是偵測出可能是人頭的橢圓形邊緣，首先偵測出前景，再由前景偵測邊緣，最後利用橢圓形樣板找出符合人頭影像大小的橢圓形邊緣，這樣的邊緣將

會是人頭的候選者。步驟三是色彩校正，色彩校正後的影像會在步驟四的行人識別及步驟六的行人追蹤用到。由步驟二找出的橢圓形邊緣不一定是人頭，經由步驟四的行人識別可以辨識出真正的人頭，這個步驟是本篇論文的一個重點，我們使用結合了許多行人特性的模糊推論系統以辨識行人。由以上步驟可以偵測出大部分的行人，但在某些情況下人頭的邊緣不會清楚地呈現，因此步驟五會利用以輪廓為基礎的行人偵測方法以檢測出邊緣不明顯的行人。步驟六是對偵測出的行人作追蹤並且預測行人的行進路徑，追蹤的方式是以顏色配對為主，邊緣形狀為輔，行進路徑是使用 IIR 濾波器作預測。這些步驟的詳細內容會在以下小節做說明。

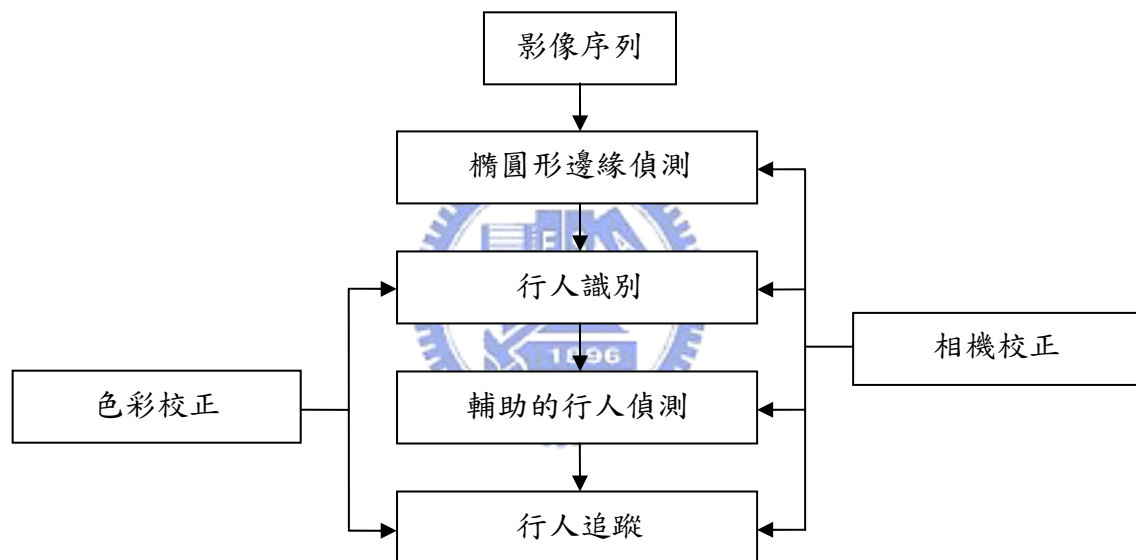


圖 3.1 行人監控系統的架構。

我們的行人監控系統有兩個特別的設定，一個是橢圓形邊緣的偵測範圍被限定在行人穿越道的兩端(如圖 3.2 的紅色方框內)，另一個設定是攝影機的鏡頭方向與道路方向垂直(如圖 3.3)。由於橢圓形邊緣偵測會耗費許多運算時間，為了使系統具備有即時處理的能力，偵測的範圍必須是有限的區域，而不是整張影像，所以我們將橢圓形邊緣的偵測範圍限定在行人穿越道的兩端，這樣的設定是合理的，因為大部分的行人會經由行人穿越道的兩端穿越道路，如果我們能在行人穿越道兩端成功地偵測到行人並持續不斷地對行人作追蹤，那麼我們便可知行人

在任何時間的位置，我們把這個方法稱為「區域偵測、全域追蹤」；紅色方框的大小及位置由使用者設定，但 l_f 及 l_n 的長度不能太短，其長度設定必須使得以一般速度行走的行人至少會出現在擷取的影像中的紅色方框內一次，而 l_f 及 l_n 的長度亦不能太長，以避免無法在限定時間內完成整個行人監控系統的運作。攝影機的鏡頭方向與道路方向垂直的好處則是可以利用到行人影像對稱的特性，這些特性將會在 3.5.1 節說明。

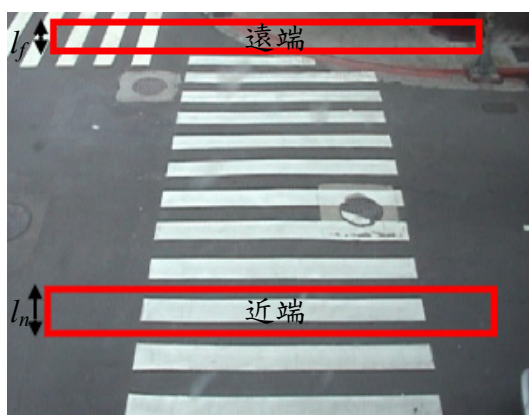
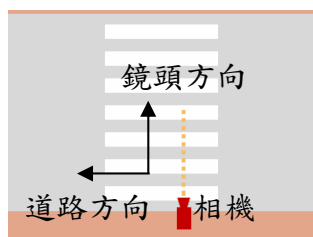
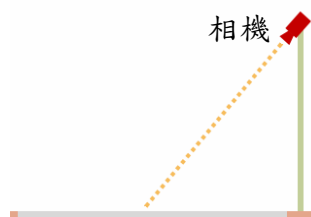


圖 3.2 橢圓形邊緣的偵測範圍。



(a)



(b)

圖 3.3 攝影機架設位置的示意圖。(a) 俯視圖；(b) 側視圖。

3.2 相機校正

章節 2.1.1 中說明了相機校正的方法，並且介紹了一個用到七個相機參數(焦距 f 、主要點 $p(u_0, v_0)$ 、平掃角 α 、俯仰角 β 、平擺角 γ 、以及相機所在位置的高度 H_c)的模型，其 CCS 及 WCS 的相對關係可參考圖 2.3。基於上述模型，此章節提出

一個合理且更簡化的模型。我們作了以下假設：1) 設定 $p(u_0, v_0)$ 為影像平面中心，且 v_0 位於影像高度的一半，如圖 3.4；2) $\gamma=0$ ；3) $\alpha=0$ 。根據行人穿越道的圖形特性可以很容易地得到消失點 V_y 以及 V_z ，如圖 3.4，利用行人穿越道兩側的平行線延伸可以得到消失點 $V_z(u_{V_z}, v_{V_z})$ ，且由式子(2.6.5)可以得到 u_0 等於 u_{V_z} ，利用不同行人的身體傾斜線可以得到另一個消失點 $V_y(u_{V_y}, v_{V_y})$ 。由式子(2.6.4)以及(2.6.6)可以求得焦距 f 為：

$$f = \sqrt{-v_{V_y} \cdot v_{V_z}}. \quad (3.1)$$

由式子(3.1)以及(2.6.6)可以求得為：

$$\beta = \tan^{-1}(-v_{V_z} / f). \quad (3.2)$$

為了求得 H_c ，可以令參考物件為影像平面中的任意一位行人，則由影像中行人頭頂的座標、影像中行人腳底的座標、以及估測的行人身高，由式子(2.10)即可求得 H_c 。經由以上假設及推導得到簡化後的透視投影矩陣 M 為：

$$M = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta & H_c \\ 0 & \sin \beta & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.3)$$

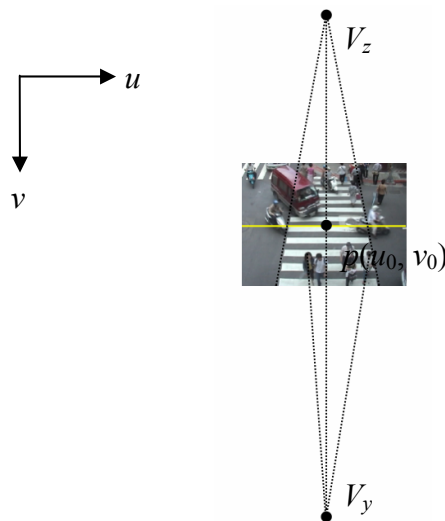


圖 3.4 消失點 V_y 以及 V_z 的座標位置。

3.3 橢圓形邊緣偵測

3.3.1 節會說明前景的取得方式。3.3.2 節敘述藉由前景擷取邊緣的方法，先偵測前景再藉由前景擷取邊緣是為了避免得到非前景的邊緣。3.3.3 節會討論橢圓形邊緣的偵測方法。

3.3.1 前景偵測

本篇論文擷取前景的方法為背景相減法，這樣的方法必須先建立背景影像，背景影像 I_b 的建立可以使用以下演算法：

$$I_{b_k}(u, v) = \text{median}(I_k^{(i)}(u, v)), i = 0, 1, 2, \dots, n-1, \quad (3.4)$$

其中 k 是 r 或 g 或 b ，分別代表 RGB 彩色模型的 R、G、以及 B 色彩空間， $I_{b_k}(u, v)$ 代表背景影像上 (u, v) 點的顏色在 k 色彩空間的值， $I_k^{(i)}(u, v)$ 代表第 i 個輸入影像在 (u, v) 點的顏色在 k 色彩空間的值， n 是要建立背景影像所需要的影像序列數目， $\text{median}()$ 代表對 n 個輸入影像在 (u, v) 點的顏色取中值。一個建立背景影像的例子如圖 3.5。



圖 3.5 建立背景影像的圖例。(a)至(c) 用來建立背景影像的部份影像序列；(d) 建立的背景影像。

對輸入影像與背景影像作比較的結果可以用前景遮罩影像 I_m 表示， I_m 是一個二值化的影像，可由下列式子求得：

$$Im(u, v) = \begin{cases} 1, & \text{if } \|I(u, v) - Ib(u, v)\| > \tau_f, \\ 0, & \text{otherwise,} \end{cases} \quad (3.5)$$

$$\|I(u, v) - Ib(u, v)\| = \sqrt{(I_r(u, v) - Ib_r(u, v))^2 + (I_g(u, v) - Ib_g(u, v))^2 + (I_b(u, v) - Ib_b(u, v))^2}, \quad (3.6)$$

其中 τ_f 是一個適當的**臨界值**(threshold)，式子(3.5)的意義是當輸入影像與背景影像在 (u, v) 點的**範數**(norm)大於某一臨界值，則該點為前景，並標示為 1，一個取得前景遮罩影像的例子如圖 3.6 所示。由 Im 可以得到前景 Ip 的定義如下：

$$Ip(u, v) = \begin{cases} I(u, v), & \text{if } Im(u, v) = 1, \\ \text{NA}, & \text{otherwise,} \end{cases} \quad (3.7)$$

其中 NA 代表無意義，亦即 Ip 只有定義在 Im 為 1 的像素點。

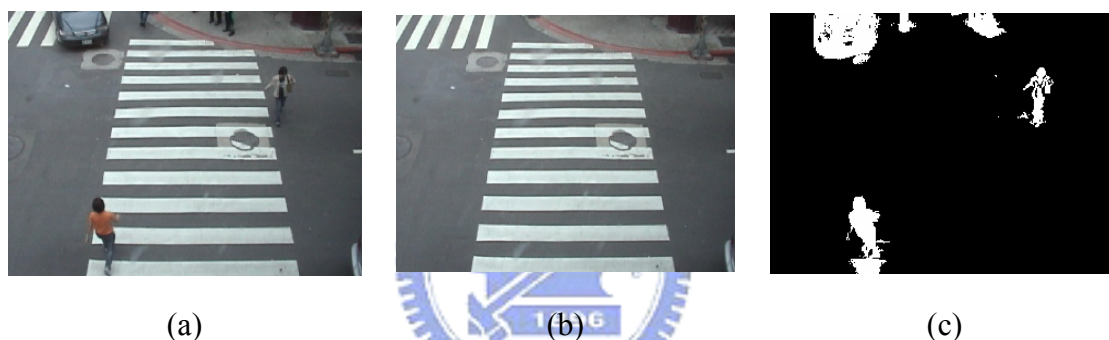


圖 3.6 建立前景遮罩影像的圖例。(a) 輸入影像；(b) 背景影像；(c) 前景遮罩。

物體反射的顏色在不同光源的情況下會不同，因此一個固定的場景在不同的時刻的背景影像也會不同，為了解決這個問題可以建立幾個光源強度不同的背景影像，另外一種常見的方式則是採用即時更新背景的技巧，可以參考[31]。隨著光源強度以及物體顏色的不同，上述的 τ_f 的最佳值也會不同，我們可以將 τ_f 視為光源強度 V 的函數，經由統計可以得到兩者的關係式 $\tau_f(V)$ 。本篇論文的 τ_f 設為 20，實驗結果顯示在日間或是夜間的效果都可以令人接受。

3.3.2 邊緣擷取

章節 2.1.2 介紹了對彩色影像作邊緣偵測的方法，式子(2.17)到(2.20)可以求得

邊緣的強度大小，由於計算較繁雜，在此提出一個簡易的邊緣計算方法如下：

$$G(u, v) = \frac{\sqrt{(Gr_u(u, v))^2 + (Gg_u(u, v))^2 + (Gb_u(u, v))^2} + \sqrt{(Gr_v(u, v))^2 + (Gg_v(u, v))^2 + (Gb_v(u, v))^2}}{2} \quad (3.8)$$

其中 $G(u, v)$ 代表影像在 (u, v) 點的邊緣強度大小， $Gr_u(u, v)$ 、 $Gg_u(u, v)$ 、以及 $Gb_u(u, v)$ 代表像素 (u, v) 分別在 R、G、以及 B 色彩空間影像的水平方向導數，可由式子(2.16.2)求得， $Gr_v(u, v)$ 、 $Gg_v(u, v)$ 、以及 $Gb_v(u, v)$ 代表 (u, v) 點分別在 R、G、以及 B 色彩空間影像的垂直方向導數，可由式子(2.16.1)求得。圖 3.7(a)是由圖 3.6(a)求得的邊緣，其中只有在 $Im(u, v)$ 的值為 1 的像素位置作邊緣偵測，邊緣的強度大小已被正規化到 $[0, 255]$ 。

邊緣的強度大小太小代表在該位置的邊緣並不明顯，我們只擷取明顯的邊緣，若是邊緣強度大於某適當的臨界值 τ_g 則對該像素位置標示 1，否則為 0，如此可以得到一個二值化的邊緣影像 I_e ，如圖 3.7(b)所示。 τ_g 類似 τ_f ，隨著光源強度的不同， τ_g 的最佳值也會不同，將 τ_g 視為光源強度 I 的函數，經由統計可得到 $\tau_g(I)$ 。本篇論文的 τ_g 設為 40，實驗結果顯示在日間或是夜間的效果都可以令人接受。

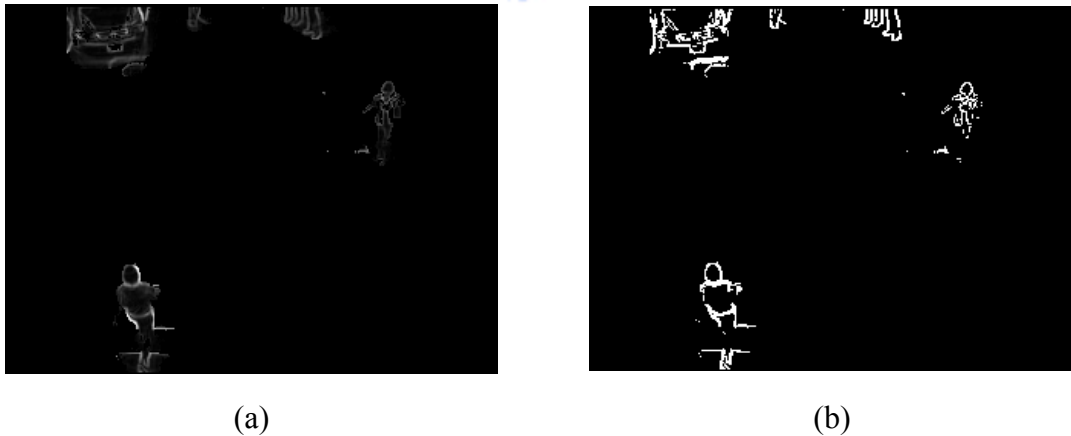


圖 3.7 由圖 3.6 偵測出的前景的邊緣。(a) 邊緣強度影像；(b) 邊緣強度大於臨界值 τ_g 的像素。

3.3.3 橢圓偵測

橢圓可以由五個參數表示，如圖 3.8 所示 (u_e, v_e) 為橢圓中心、 A_e 以及 B_e 是橢

圓的兩個主軸(principal axis)的長度、 θ_e 是方位(orientation)，橢圓上的任意一點 (u, v) 滿足下列方程式：

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos \theta_e & \sin \theta_e \\ -\sin \theta_e & \cos \theta_e \end{bmatrix} \begin{bmatrix} u - u_e \\ v - v_e \end{bmatrix}, \quad (3.9)$$

$$\frac{(u')^2}{A_e^2} + \frac{(v')^2}{B_e^2} = 1,$$

如果在橢圓上的邊緣總個數大於某個適當的臨界值 τ_e ，則稱邊緣在該位置形成橢圓形邊緣，本篇論文的 τ_e 值設為 $\pi \cdot A_e$ 。

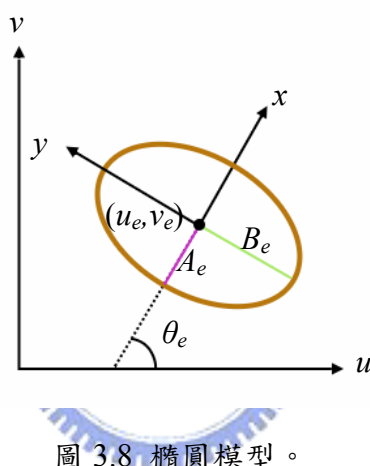


圖 3.8 橢圓模型。

偵測橢圓形邊緣是為了找出形成類似人頭的邊緣，由於一般人頭的大小大約相同，因此人頭投影在影像平面上的人頭一半寬度 A_e 以及人頭投影在影像平面上的人頭一半長度 B_e 的變化量可以設的比較小，但對整張影像搜尋橢圓仍然相當花費時間，傳統偵測橢圓的方法如用 Hough 轉換[32]或是基因演算法[33]就目前的硬體而言並無法即時偵測出橢圓，因此本篇論文對 E 的五個參數給定一些限制條件：

- 1) $\theta_e = 0$ ；
 - 2) 橢圓中心 (u_e, v_e) 只能存在限制的範圍內，如圖 3.9(a) 的紫色方框內；
 - 3) $A_e = 0.01 \cdot (8 + \delta_a) \cdot h_p$ ，其中 $\delta_a \in \{0, 1\}$ ， h_p 會在以下敘述；
 - 4) $B_e = (1 + 0.1 \cdot \delta_b) \cdot A_e$ ，其中 $\delta_b \in \{0, 1, 2, 3\}$ 。
- 行人與攝影機的距離會影響行人投影在影像平面上的影像大小，為了求得 A_e ，必須先算出行人在 WCS 的人頭及人腳座標，再算出行人投影在

影像平面上的身長 h_p ，將 h_p 乘上適當的係數後便可以得到近似的 A_e ；如圖 3.9(b)，假設橢圓形邊緣的中心在 (u_h, v_h) (亦即人頭中心在影像平面的投影)，人頭中心在 WCS 的座標為 (x_p, H_p, z_p) ，其中 H_p 在本篇論文中設定為 150，式子(3.3)簡化後的透視投影矩陣 M 表示為

$$M = \begin{bmatrix} m_{11} & 0 & 0 & 0 \\ 0 & m_{22} & m_{23} & m_{24} \\ 0 & m_{32} & m_{33} & 0 \end{bmatrix}, \quad (3.10)$$

則由式子(2.2)以及(3.10)可以得到 x_p 以及 z_p 如下列所示：

$$\begin{bmatrix} x_p \\ z_p \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

$$\begin{aligned} b_{11} &= -m_{11}, \\ b_{12} &= u_h \cdot m_{33}, \\ b_{21} &= 0, \\ b_{22} &= v_h \cdot m_{33} - m_{23}, \\ c_1 &= -u_h \cdot m_{32} \cdot H_p, \\ c_2 &= (-v_p \cdot m_{32} + m_{22}) \cdot H_p + m_{24}, \end{aligned} \quad (3.11)$$

得到 x_p 以及 z_p 後即可得到行人的腳在 WCS 的座標為 $(x_p, 0, z_p)$ ，代入式子(2.2)得到行人的腳投影在影像平面的座標 (u_f, v_f) ，計算 (u_h, v_h) 與 (u_f, v_f) 之間的距離即可求得 h_p 。一個橢圓形邊緣偵測的例子如圖 3.9 所示，圖 3.9(c)中總共偵測到三個橢圓形邊緣，由圖 3.9(d)可以看出在行人穿越道的近端找到的橢圓形邊緣為人頭，而在遠端找到的兩個橢圓形邊緣並非真正的人頭，為了辨識找出的橢圓形邊緣是否為人頭，必須再經過 3.5 節的模糊推論系統作判斷。

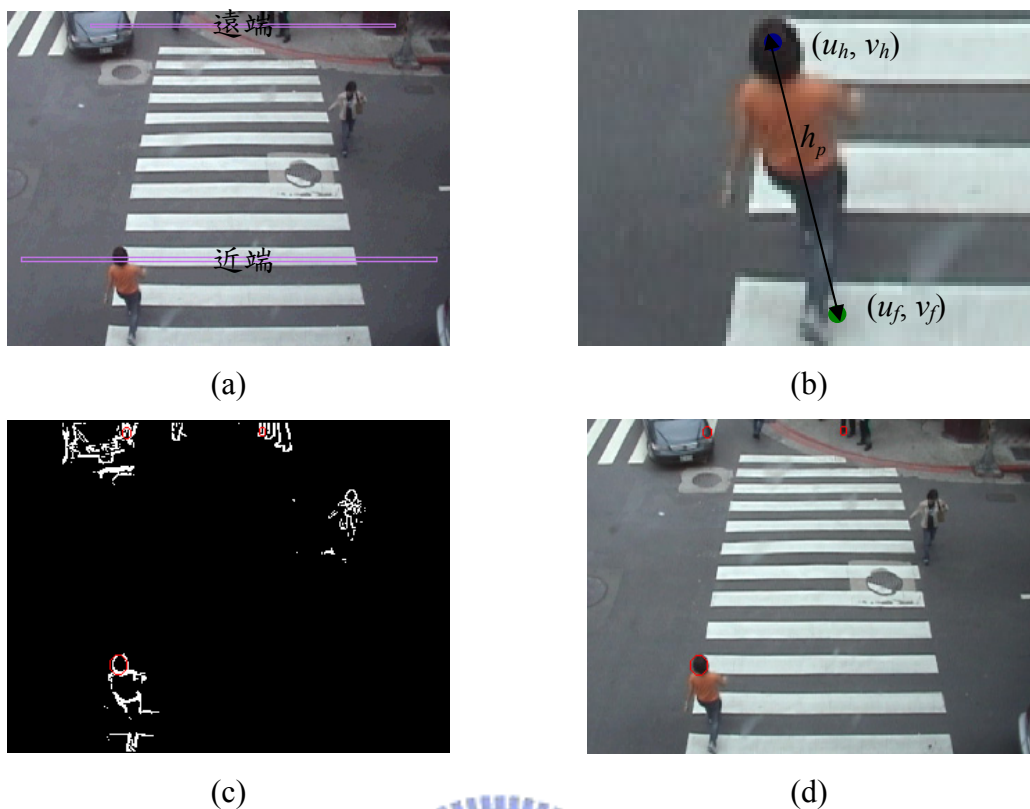


圖 3.9 橢圓形邊緣偵測的圖例。(a) 輸入影像與限定的橢圓中心位置；(b) 行人身長示意圖；(c) 邊緣影像與偵測出的橢圓；(d) 輸入影像與偵測出的橢圓。

橢圓形邊緣是建構在偵測到的前景上，我們使用背景相減法偵測前景，Jean 以及 Wu [10]則是使用時間差異法找出移動的前景，以避免偵測到不是移動的行人的前景，使用下列式子：

$$Id^{(i)}(u, v) = \begin{cases} 1, & \text{if } \|I^{(i)}(u, v) - I^{(i-1)}(u, v)\| > \tau_d, \\ 0, & \text{otherwise,} \end{cases} \quad i = 1, 2, 3, \dots, \quad (3.12)$$

可以得到二值化的差異影像 $Id^{(i)}$ ， $Id^{(i)}(u, v)=1$ 的像素位置代表有移動的前景，若將 $Id^{(i)}$ 分成上半部以及下半部，並分別搜尋 $Id^{(i)}$ 中上半部以及下半部的移動的前景的最上、最下、最左、以及最右的出現位置，則可以分別得到行人穿越道近端與遠端的移動物件的範圍。以圖 3.10 為例，圖 3.10(a)是目前的輸入影像，圖 3.10(b)是前一刻的輸入影像，由式子(3.12)得到圖 3.10(c)的差異影像 $Id^{(i)}$ ，圖 3.10(d)中的方框是偵測出的移動物件的範圍，如果要以方框內的每一點為橢圓中心偵測橢圓形邊緣，則以圖 3.10(d)為例，總共的橢圓偵測以及行人識別時間約為 0.35 秒，這

樣的運算時間並不適合用在即時的處理系統。

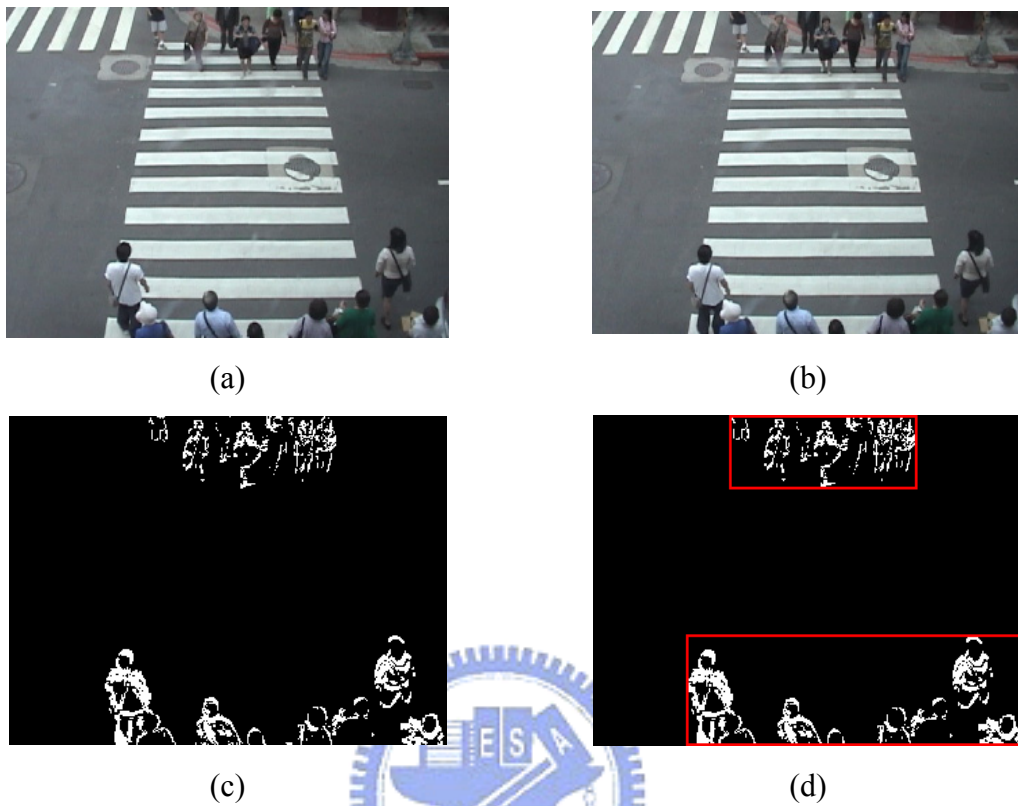


圖 3.10 移動物件的偵測。(a) 目前的輸入影像；(b) 前一刻的輸入影像；(c) 移動的前景；(d) 移動物件的範圍。

3.4 色彩校正

色彩校正後的影像會在 3.5 節的行人識別以及 3.7 節的行人追蹤用到，作色彩校正的原因會在使用到色彩校正後的影像的小節作說明，在此先介紹色彩校正的功能及演算法。

在不同光源的照射下相同的物體會反射出不同的顏色，為了使物體的顏色不受光源的改變而改變，必須使用到色彩校正的方法，使得物體能呈現一致性的顏色。一個簡單的色彩校正是灰界理論法(gray world assumption)[34]，假設 \bar{R} 、 \bar{G} 、以及 \bar{B} 是輸入影像所有像素分別在 R、G、以及 B 色彩空間的平均值，輸入影

像的平均灰階值為 $W = (\bar{R} + \bar{G} + \bar{B})/3$ ，則輸入影像 I 上任一點 (u, v) 消去色偏後的顏色可以由下列式子求得：

$$\begin{aligned} I_r'(u, v) &= I_r(u, v) \cdot (W / \bar{R}), \\ I_g'(u, v) &= I_g(u, v) \cdot (W / \bar{G}), \\ I_b'(u, v) &= I_b(u, v) \cdot (W / \bar{B}), \end{aligned} \quad (3.13)$$

其中 $I_r'(u, v)$ 、 $I_g'(u, v)$ 、以及 $I_b'(u, v)$ 是消去色偏後的影像 I' 在 (u, v) 點的 R 、 G 、以及 B 值，圖 3.11(c) 是圖 3.11(a) 經由式子(3.13) 運算後得到的影像，可以看到黃色色偏已經被消除，但是亮度偏低，調整亮度的方法是先將 I' 的色彩模型由 RGB 轉成 YCbCr，再將 I' 的每一個像素的亮度 $I_y'(u, v)$ 乘上一個權重，使得調整後影像的亮度平均值等於參考影像的亮度平均值，而彩度值不變，這個過程可以由下列式子表示：

$$\begin{aligned} I_y''(u, v) &= I_y'(u, v) \cdot (\bar{Y}_r / \bar{Y}'), \\ I_{cb}''(u, v) &= I_{cb}'(u, v), \\ I_{cr}''(u, v) &= I_{cr}'(u, v), \end{aligned} \quad (3.14)$$

其中 $I_y''(u, v)$ 、 $I_{cb}''(u, v)$ 、以及 $I_{cr}''(u, v)$ 是調整完亮度的影像 I'' 在 (u, v) 點的 Y 、 Cb 、以及 Cr 值， \bar{Y}' 是 I' 的亮度平均值， \bar{Y}_r 是參考影像的亮度平均值，圖 3.11(d) 是由圖 3.11(c) 調整亮度後得到的影像，圖 3.11(b) 則是參考影像。我們在本篇論文定義 I_w 為作完色彩校正後的影像，在接下來的章節將會被使用。

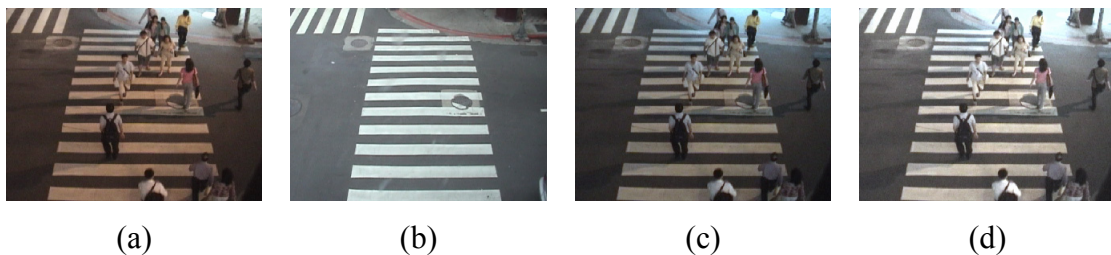


圖 3.11 色彩校正的圖例。(a) 輸入影像；(b) 參考影像；(c) 消除色偏後的影像；(d) 調整亮度後的影像。

3.5 行人識別

為了判斷 3.3 節偵測到的橢圓形邊緣是否為行人的頭部，必須作進一步的識別，根據行人的特性我們建立辨識行人的模糊推論系統，詳細內容會在此章節說明。3.5.1 節定義行人的特徵並說明萃取的方法，3.5.2 節敘述模糊推論系統的架構。

3.5.1 行人特徵

模糊推論系統的輸入為四個行人特徵，分別為：1) 人頭與兩側前景的顏色差異度；2) 身體比例；3) 衣服紋路相似度；4) 膚色量，分別在下列小節說明。

3.5.1.1 人頭與兩側前景的顏色差異度

根據 3.3.3 節的定義，只要橢圓上的邊緣像素個數大於 $\pi \cdot A_e$ ，則判定該位置有形成橢圓形的邊緣，並成為人頭的候選者，但這樣的定義會造成許多非人頭的橢圓形邊緣，如圖 3.12 的肩膀邊緣。根據觀察，當攝影機的鏡頭方向與道路方向垂直時(如圖 3.3)，非人頭的橢圓內部顏色與橢圓左右某一側前景顏色的差異度會很小(如圖 3.12(a))，對人頭橢圓而言橢圓內部顏色與橢圓左右兩側前景顏色的差異度則相對較大(如圖 3.13(a)以及 3.13(b))，甚至人頭周圍幾乎沒有前景 (如圖 3.13(c)以及 3.13(d))，則此時代表人頭橢圓與左右兩側的顏色差異度最大。

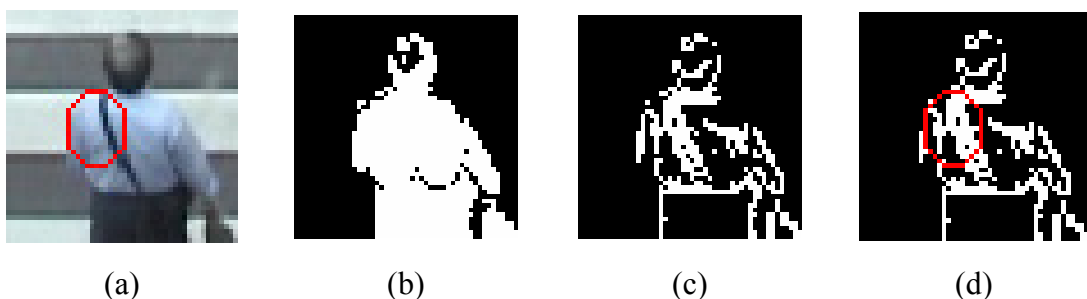


圖 3.12 非人頭的橢圓形邊緣。(a) 輸入影像與偵測到的橢圓；(b) 前景的像素；(c) 邊緣影像；(d) 偵測到的橢圓形邊緣。

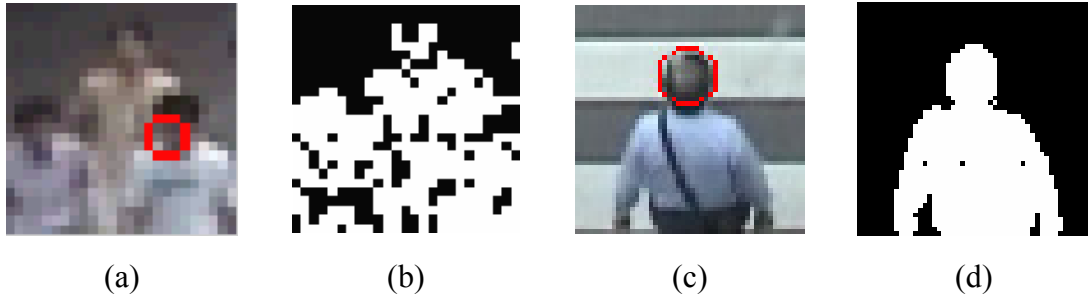


圖 3.13 人頭與兩側前景的顏色差異。(a) 兩側皆有前景的人頭；(b) (a)的前景像素；(c) 兩側沒有前景的人頭；(d) (c)的前景像素。

為了評估偵測出的橢圓的內部顏色與橢圓左右兩側前景的顏色的差異，我們使用顏色的直方圖比對的技巧[35]，首先介紹直方圖的建立方式，假設像素的顏色分別由三個色彩空間的 k_1 、 k_2 、以及 k_3 個色階表示，則直方圖相對會有 $k_1+k_2+k_3$ 個倉(bin)，每個像素的顏色在三個色彩空間的值會分別對應到三個不同的倉，每個倉會記錄影像對應到該倉的像素個數；接著定義人頭與兩側前景顏色差異(color difference between head and both sides, CDH)如下：

$$\begin{aligned}
 CDH &= \min(CDH_l, CDH_r), \\
 CDH_l &= ((\sum_{k=1}^N \max(H_h(k), H_{hl}(k)) + 3 \cdot n_{hl}) - 3 \cdot n) / (3 \cdot n), \\
 CDH_r &= ((\sum_{k=1}^N \max(H_h(k), H_{hr}(k)) + 3 \cdot n_{hr}) - 3 \cdot n) / (3 \cdot n),
 \end{aligned} \tag{3.15}$$

其中 $H_h(k)$ 、 $H_{hl}(k)$ 、以及 $H_{hr}(k)$ 分別為橢圓內部、橢圓左側前景、以及橢圓右側前景建立的直方圖的第 k 個倉的像素個數， n_{hl} 為橢圓左側非前景的像素個數， n_{hr} 為橢圓右側非前景的像素個數， n 為橢圓內部或是橢圓左側或是橢圓右側用來統計直方圖的像素總個數， N 為直方圖的倉的個數，式子(3.15)的意義是當兩個直方圖的交疊部分愈少，則顏色差異愈大， CDH 值會愈大， CDH 的最大值發生在兩個直方圖完全沒有交疊時，一個特殊的例子是當橢圓兩側皆無前景的情形(如圖 3.13(c) 以及 3.13(d))，此時 $CDH=1$ ，若兩個直方圖完全交疊則 $CDH=0$ ，為 CDH 的最小值。本篇論文作顏色比對的三個色彩空間根據[36]採用 $B-G$ 、 $G-R$ 、以及 $B+G+R$ ，色階數目分別為 8、8、以及 4， N 為 20，這樣的色彩空間可以將彩度與亮度分離且容易計算，圖 3.14 顯示橢圓內部與橢圓左右兩側的定義範圍，橢圓

內部為綠色方框內，橢圓左側為藍色方框內，橢圓右側為黃色方框內，三個方框大小相同，其中 (u_h, v_h) 是橢圓形邊緣的中心， h_p 為 3.3.3 節求得的行人投影在影像平面上的長度。橢圓與兩側直方圖建立的例子如圖 3.15 以及圖 3.16，其中 1-8 的倉對應 $B-G$ 的色階、9-16 的倉對應 $G-R$ 的色階、而 17-20 的倉對應 $B+G+R$ 的色階，圖 3.15 是非人頭的橢圓形邊緣，其 $CDH=0.053$ ，圖 3.16 是人頭的橢圓形邊緣，其 $CDH=0.481$ ，在這兩個例子中人頭的 CDH 值較非人頭的 CDH 值大。

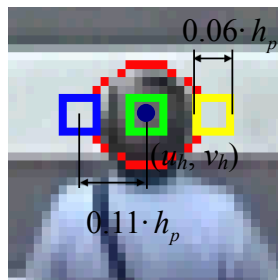
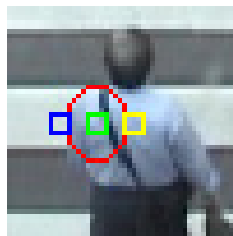
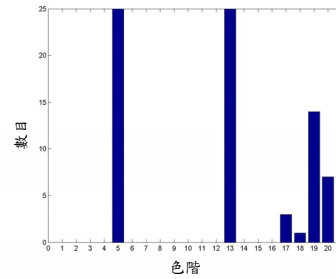


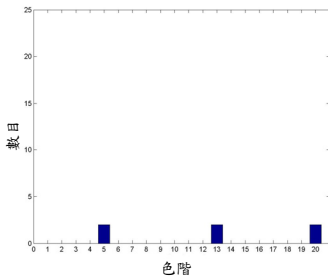
圖 3.14 橢圓與兩側偵測顏色的範圍。



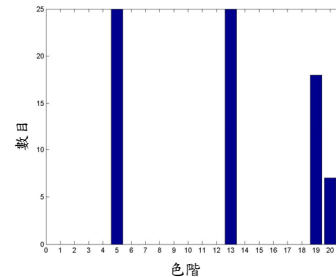
(a)



(b)



(c)



(d)

圖 3.15 非人頭的直方圖例子。(a) 輸入影像與偵測顏色的範圍；(b) 橢圓內部的直方圖；(c) 橢圓左側前景的直方圖；(d) 橢圓右側前景的直方圖。

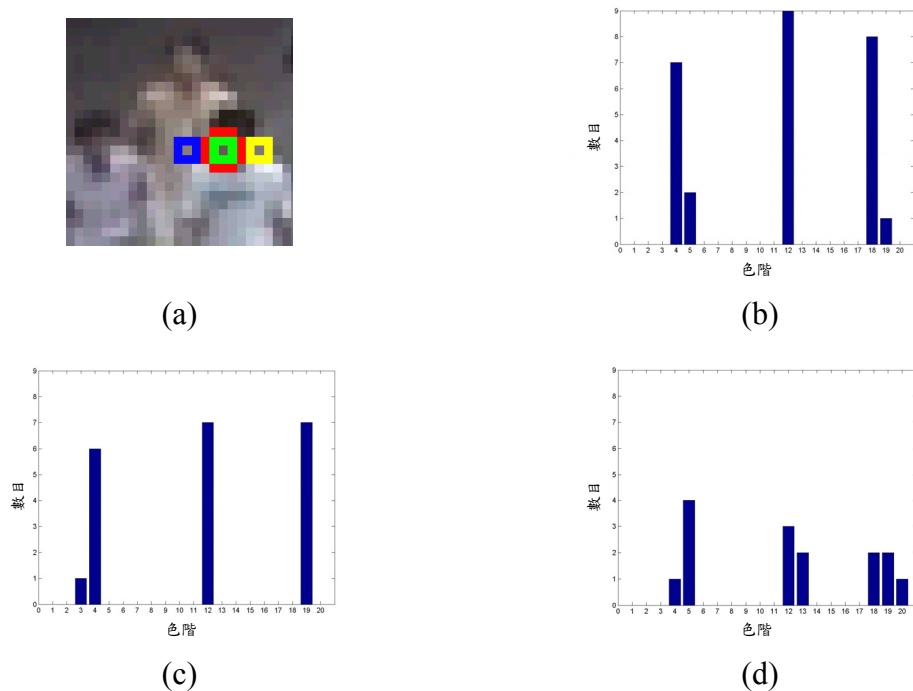


圖 3.16 人頭的直方圖例子。(a) 輸入影像與偵測顏色的範圍；(b) 橢圓內部的直方圖；(c) 橢圓左側前景的直方圖；(d) 橢圓右側前景的直方圖。

在光源較強的情形下色彩比較鮮明，因此可以較容易地檢測出色彩的差異度，但在光線昏暗下 CDH 的值會偏低，而在路燈下還會造成色偏(如圖 3.11(a))，所以計算 CDH 所建立的直方圖是根據作完色彩校正後的影像 I_w 而非攝影機直接擷取的輸入影像 I ，在這裡可以看到作色彩校正的理由。

對於橢圓內部與橢圓周圍顏色的比較，我們只考慮橢圓內部與橢圓左右兩側前景的顏色差異度，而不考慮橢圓內部與橢圓上下兩側前景的顏色差異度，這樣設定的原因可以參考圖 3.17，其中的紅色橢圓是我們要討論的人頭，綠色方框是人頭上下兩側的定義範圍，由圖 3.17(a)可以看出近端人頭與人頭上下兩側前景的顏色差異度可能會很小，由圖 3.17(b)可以看出遠端人頭與人頭上下兩側前景的顏色差異度可能會很小，因此我們可以得知偵測到的橢圓是人頭的可能性與人頭上下兩側前景的顏色差異度並沒有較明顯的正相關性，所以我們只以橢圓內部與橢圓左右兩側前景的顏色差異度評估橢圓是人頭的可能性。



圖 3.17 人頭與人頭上下兩側的顏色比較。(a) 近端人頭；(b) 遠端人頭。

3.5.1.2 身體比例

假設 (u_h, v_h) 是人頭影像的中心， (u_f, v_f) 是行人的腳在影像平面的座標，則兩點之間應該存在屬於行人身軀的前景影像，若是兩點間的前景的像素太少則代表該位置有行人的機率很小，反之則該位置有行人的機率較大，此外，行人上半身大小與下半身大小的比例會接近一個常數值。我們定義身體比例(ratio of up body to low body, RUL)表示存在身體的可能性，定義如下：

$$RUL = \frac{N_p}{N_b} \cdot r_{ul},$$

$$r_{ul} = \begin{cases} \frac{N_l}{N_u/2}, & \text{if } N_u/2 \geq N_l, \\ \frac{N_u/2}{N_l}, & \text{if } N_u/2 < N_l, \end{cases} \quad (3.16)$$

其中 N_b 是上半身身體樣板以及下半身身體樣板上的總像素個數，圖 3.18 的黃色平行四邊形是上半身身體樣板，橙色平行四邊形是下半身身體樣板，其中 h_p 與 3.3.3 節的定義相同， N_p 是上半身身體樣板以及下半身身體樣板上的前景像素個數， N_u 是上半身身體樣板的前景像素個數， N_l 是下半身身體樣板的前景像素個數，且 $N_p = N_u + N_l$ 。

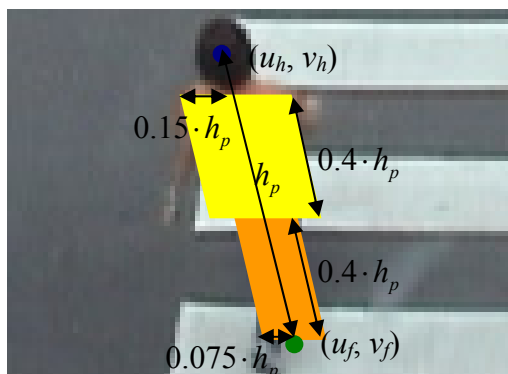


圖 3.18 身體樣板的示意圖。

3.5.1.3 衣服紋路相似度

根據觀察行人的衣服紋路通常是左右對稱，且當鏡頭架設在高處以及鏡頭與行人距離相近時行人身體的上半部不易被其他行人遮蔽(如圖 3.19)，因此如果在行人穿越道的近端(如圖 3.2 的近端)偵測到的橢圓形邊緣是人頭，則橢圓下方左右兩側(如圖 3.19 的青色和紫紅色方框內)的衣服紋路相似度應該會很高。

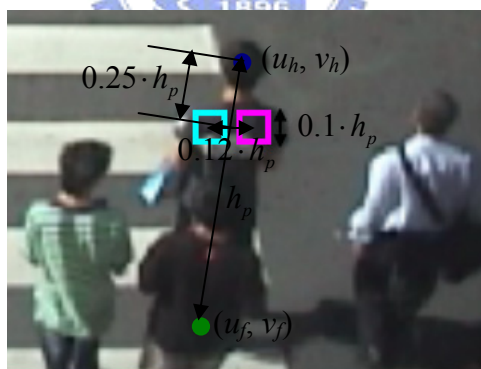


圖 3.19 人頭下方左右兩側的示意圖。

為了評估橢圓形邊緣下方左右兩側的前景紋路相似度，我們使用類似 3.5.1.1 節直方圖比對的方法，定義衣服紋路相似度(texture similarity of clothes, TSC)如下：

$$TSC = \sum_{k=1}^N \min(H_{ll}(k), H_{rr}(k)) / (3 \cdot n), \quad (3.17)$$

其中 $H_{ll}(k)$ 為橢圓下方左側直方圖第 k 個倉的像素個數， $H_{rr}(k)$ 為橢圓下方右側直方圖第 k 個倉的像素個數， n 為橢圓下方左側或是橢圓下方右側用來統計直方圖的

像素總個數， N 為直方圖的倉的個數，式子(3.17)的意義是當兩個直方圖的交疊部分愈多，則顏色愈相似、 TSC 值愈大， TSC 的最大值會發生在兩個直方圖完全交疊時，此時 $TSC=1$ ，當兩個直方圖完全無交疊時則 $TSC=0$ 為最小值。三個色彩空間採用 $B-G$ 、 $G-R$ 、以及 $B+G+R$ ， $N=20$ ，與 3.5.1.1 節相同，圖 3.19 顯示橢圓下方左側及右側的定義範圍，橢圓下方左側為青色方框內，橢圓下方右側為紫紅色方框內，兩個方框的大小相同，其中 h_p 與 3.3.3 節的定義相同。

與 3.5.1.1 節的 CDH 類似地，在光線昏暗下 TSC 的值會偏低，所以計算 TSC 所建立的直方圖是根據作完色彩校正後的影像 I_w 而非輸入影像 I 。

3.5.1.4 膚色量

假設 (u_h, v_h) 是人頭影像的中心並且行人正對著鏡頭，則 (u_h, v_h) 點周圍像素的顏色在一般的情形下會是膚色，如果偵測到的橢圓形邊緣內的膚色像素愈多代表該橢圓愈可能是人臉，根據這個特性我們定義膚色量(quantity of skin pixels, QOS) 表示存在人臉的可能性，定義如下：

$$QOS = \frac{\text{人臉樣板上的膚色像素個數}}{\text{人臉樣板上的總像素個數}}, \quad (3.18)$$

人臉樣板如圖 3.20 的紅色正方形，其中 (u_h, v_h) 是偵測到的橢圓中心， h_p 與 3.3.3 節的定義相同。

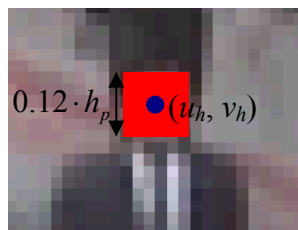


圖 3.20 人臉樣板的示意圖。

為了偵測顏色屬於膚色的像素個數，必須先建立一個適當的膚色模型，根據[37]的討論，膚色使用 YCbCr 色彩模型表示是較佳的選擇，除了計算容易的優點外，

另一個優點是膚色在 Cb 以及 Cr 色彩空間的最大值及最小值不太受 Y 值的影響，膚色在 YCbCr 的顏色分佈會類似一個直筒形，如圖 3.21(a)所示。圖 3.21 是我們採用 100 張在日間由攝影機拍得的人臉所建立的膚色在 YCbCr 空間的顏色分布圖，每個點代表一個人臉上像素的顏色在 YCbCr 空間的值，我們使用六個限制式限制膚色在 YCbCr 色彩空間的範圍，如下列所示：

$$Y - 16 > 0, \quad (3.19.1)$$

$$Y - 235 < 0, \quad (3.19.2)$$

$$10 \cdot Cb - 7 \cdot Cr - 130 > 0, \quad (3.19.3)$$

$$10 \cdot Cb - 7 \cdot Cr - 460 < 0, \quad (3.19.4)$$

$$Cb + Cr - 242 > 0, \quad (3.19.5)$$

$$Cb + Cr - 270 < 0, \quad (3.19.6)$$

(3.19)

圖 3.21(b)中的四條直線 y_1 、 y_2 、 y_3 、以及 y_4 分別為式子(3.19.3)、(3.19.4)、(3.19.5)、以及(3.19.6)的限制邊界。

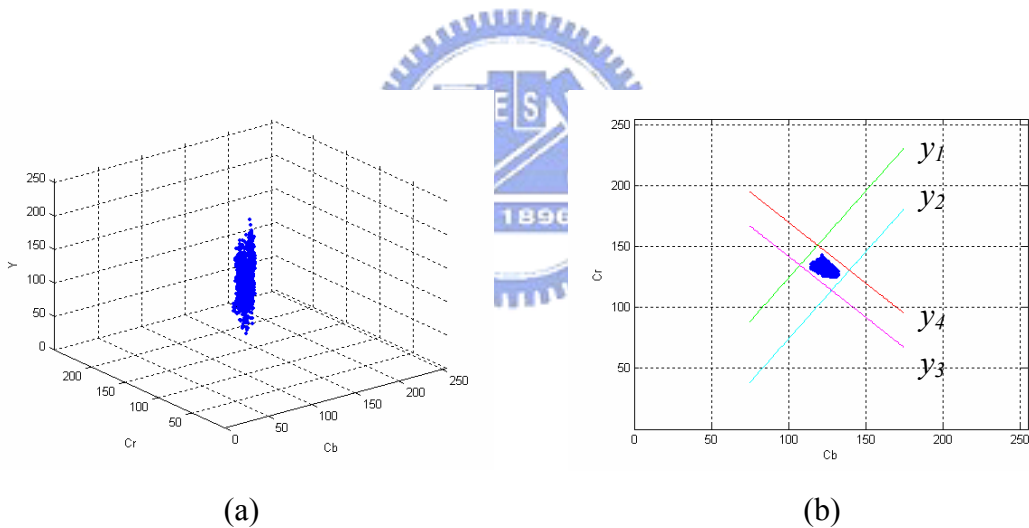


圖 3.21 膚色模型。(a) 膚色在 YCbCr 色彩空間的分布；(b) 膚色在 CbCr 色彩空間投影的分布。

在有色偏的光源下人臉的顏色在 YCbCr 空間的分布會與在正常日光照射下的情形不同，圖 3.22(a)是我們採用 50 張在夜間路燈照明下的人臉所建立的顏色分布圖，其中四條直線 y_1 、 y_2 、 y_3 、以及 y_4 的定義與圖 3.21(b)相同，圖 3.22(a)人臉顏色分布的範圍較圖 3.21(b)小且 Cr 值較大而 Cb 值較小，圖 3.22(b)則是人臉作完色彩校正後的顏色分布，其中四條直線 y_1 、 y_2 、 y_3 、以及 y_4 的定義與圖 3.21(b)相同，

圖 3.22(b)人臉顏色分布範圍較圖 3.22(a)大且與圖 3.21(b)的範圍類似，所以計算 QOS 前必須對輸入影像作色彩校正，在此看到對輸入影像作色彩校正的必要性。

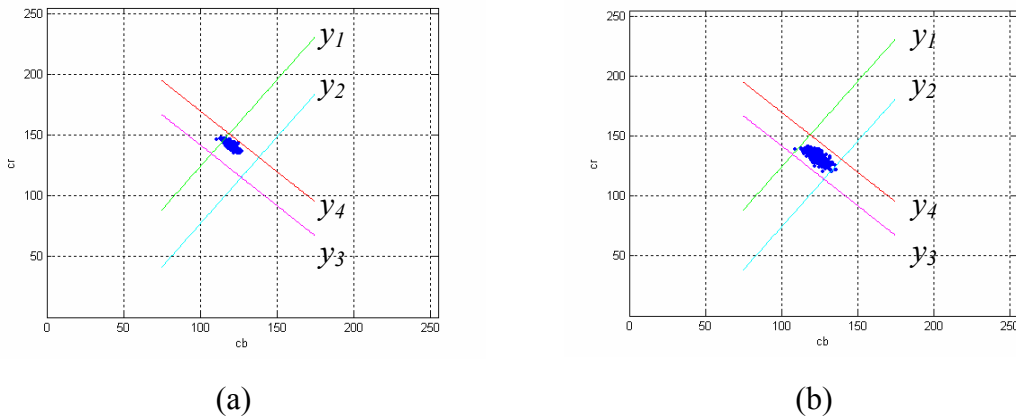


圖 3.22 人臉顏色與色偏的關係。(a) 黃色路燈下的人臉顏色在 CbCr 色彩空間投影的分布；(b) 色彩校正後的人臉顏色在 CbCr 色彩空間投影的分布。

3.5.2 模糊推論系統架構

根據以上討論的行人特徵我們建立模糊推論系統以辨識行人，由於鏡頭近端行人的特性與鏡頭遠端行人的特性有些不同，因此我們分別建立兩個模糊推論系統，一個是模糊近端行人辨識系統會在 3.5.2.1 節說明，另一個是模糊遠端行人辨識系統在 3.5.2.2 節說明。

3.5.2.1 模糊近端行人辨識系統

模糊近端行人辨識系統(fuzzy near pedestrian recognition system, NPR)是要辨識由鏡頭近端要進入行人穿越道的行人，這樣的行人應該是背對著鏡頭的，所以無法偵測人臉的膚色，因此無法使用到 QOS ，但是我們可以藉由行人背部上衣的紋路計算 TSC 作為 NPR 的輸入，NPR 共有三個輸入以及一個輸出，三個輸入為 CDH (值域為 $[0,1]$)、 RUL (值域為 $[0,1]$)、以及 TSC (值域為 $[0,1]$)，唯一的輸出為行人程度(pedestrian level, PL)(值域為 $[0,100]$)。如果 CDH 的值愈大，代表偵測到的橢圓

內部與兩側前景的顏色差異度愈大，則該橢圓愈可能是人頭；如果 RUL 的值愈大，代表偵測到的橢圓下側愈有可能有行人的身體，則該位置愈有可能是行人；如果 TSC 的值愈大，代表偵測到的橢圓的左下與右下的紋路相似度愈高，則該位置愈有可能是行人； PL 用來判斷在鏡頭近端偵測到的橢圓形邊緣是否為行人的頭，如果 PL 大於 48 代表該位置有行人。模糊化機構須藉由歸屬函數將明確的輸入映射至模糊量值，本篇論文的三個輸入的歸屬函數的設定分別如圖 3.23(a)、3.23(b)、以及 3.23(c)所示，輸出的歸屬函數則如圖 3.23(d)，歸屬函數的設定與模糊規則的設計有相當密切的關係，我們根據實驗的觀察與經驗設定歸屬函數以及相關的模糊規則。

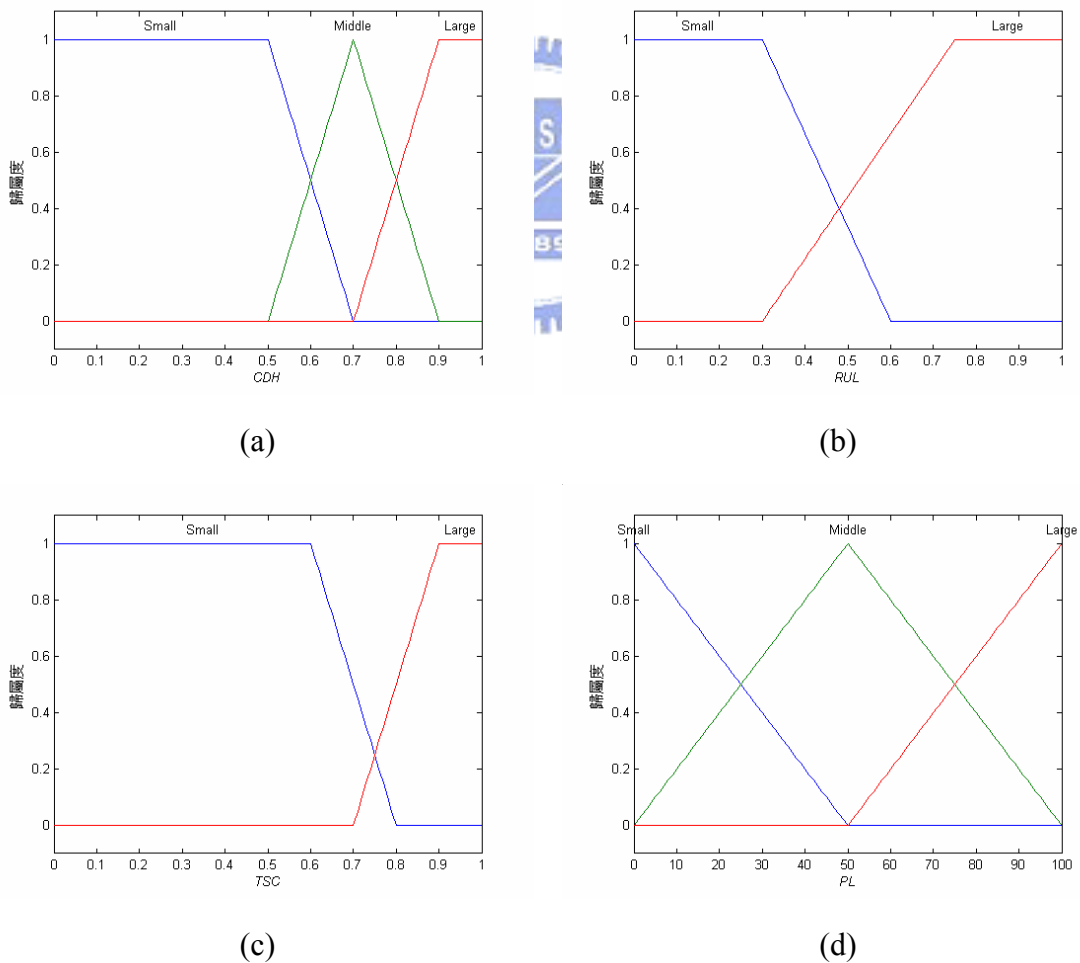


圖 3.23 NPR 的輸入以及輸出的歸屬函數。(a) CDH 的歸屬函數；(b) RUL 的歸屬函數；(c) TSC 的歸屬函數；(d) PL 的歸屬函數。

模糊推論系統辨識行人的效能取決於模糊規則庫的設計，一個最大化的模糊規則庫的設計通常是採用下列的形式[38]：

$$R^i : \text{IF } X_1 \text{ is } A_1^i \text{ AND } \dots \text{ AND } X_n \text{ is } A_n^i \text{ THEN } Y \text{ is } LY^i, \quad (3.20)$$

以上的設定可以得到最多條的模糊規則，以 NPR 為例，它有三個輸入，其中一個輸入有三個語意值，另外兩個輸入皆分別有兩個語意值，則 NPR 最多有 $3 \times 2 \times 2 = 12$ 條模糊規則。一個較有效率的模糊規則庫的設計方法則是使用不在意項(don't care)[38]，如下列的形式：

$$R^i : \text{IF } X_1 \text{ is } A_1^i \text{ AND } X_2 \text{ is } NC \text{ AND } \dots \text{ AND } X_n \text{ is } NC \text{ THEN } Y \text{ is } LY^i, \quad (3.21)$$

其中 *NC* 代表不在意項且使用的個數及位置不受限定，上式可以簡單的表示如下

$$R^i : \text{IF } X_1 \text{ is } A_1^i \text{ THEN } Y \text{ is } LY^i, \quad (3.22)$$

這樣的設計方法可以減少模糊規則的數目，避免考慮繁瑣且可能影響不大的模糊規則，事實上式子(3.22)隱含了其它的模糊規則，但要注意的是設計完成的模糊規則庫必須包含所有的狀況，以 NPR 為例則是要包含 12 個狀況。根據以上討論的行人特徵、實驗的觀察、以及[38]訂定模糊規則，NPR 共有五個模糊規則列出如下：

$$R^1 : \text{IF } CDH \text{ is Small THEN } PL \text{ is Small}, \quad (3.23)$$

$$R^2 : \text{IF } RUL \text{ is Small THEN } PL \text{ is Small}, \quad (3.24)$$

$$R^3 : \text{IF } CDH \text{ is Middle AND } RUL \text{ is Large AND } TSC \text{ is Small} \\ \text{THEN } PL \text{ is Small}, \quad (3.25)$$

$$R^4 : \text{IF } CDH \text{ is Middle AND } RUL \text{ is Large AND } TSC \text{ is Large} \\ \text{THEN } PL \text{ is Middle}, \quad (3.26)$$

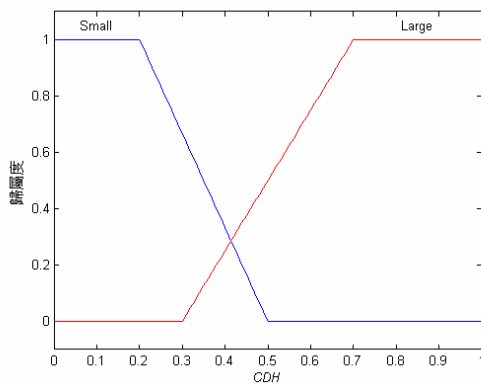
$$R^5 : \text{IF } CDH \text{ is Large AND } RUL \text{ is Large THEN } PL \text{ is Large}, \quad (3.27)$$

NPR 作模糊推論的方法為 Mamdani 的最小值推論法，去模糊化的方法為 COA。

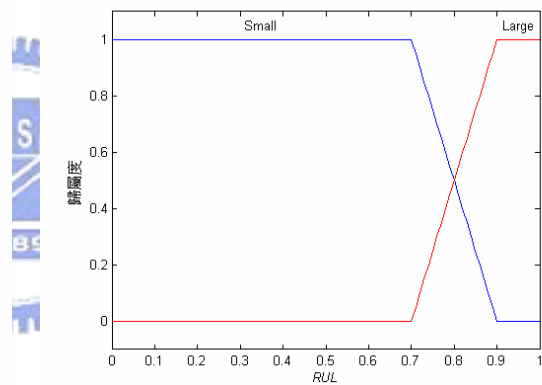
3.5.2.2 模糊遠端行人辨識系統

模糊遠端行人辨識系統(fuzzy far pedestrian recognition system, FPR)是要辨識由鏡頭遠端要進入行人穿越道的行人，這樣的行人應該是正對著鏡頭，我們可以

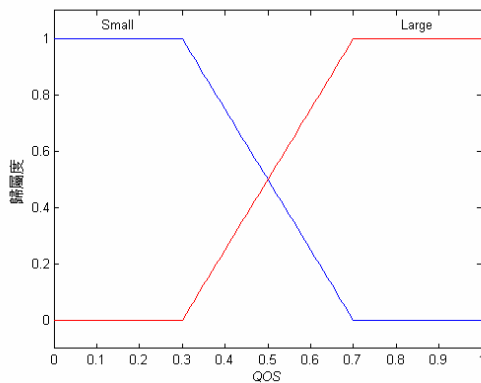
偵測人臉的膚色作為遠端行人的特徵，也就是計算 QOS 作為 FPR 的輸入，不過 FPR 並不適合採用 TSC 作為輸入，因為遠端行人距離鏡頭太遠，上衣被其他行人遮蔽的機會很大。 FPR 共有三個輸入以及一個輸出，三個輸入為 CDH (值域為 $[0,1]$)、 RUL (值域為 $[0,1]$)、以及 QOS (值域為 $[0,1]$)，唯一的輸出為 PL (值域為 $[0,100]$)。與 NPR 類似地，如果 CDH 的值愈大，則該橢圓愈可能是人頭，如果 RUL 的值愈大，則該位置愈有可能是行人；如果 QOS 的值愈大，代表偵測到的橢圓上的膚色像素愈多，則該橢圓愈有可能是人頭； PL 用來判斷在鏡頭遠端偵測到的橢圓形邊緣是否為行人的頭，如果 PL 大於 42 代表該位置有行人，三個輸入的歸屬函數分別如圖 3.24(a)、3.24(b)、以及 3.24(c) 所示，輸出的歸屬函數則如圖 3.24(d)。



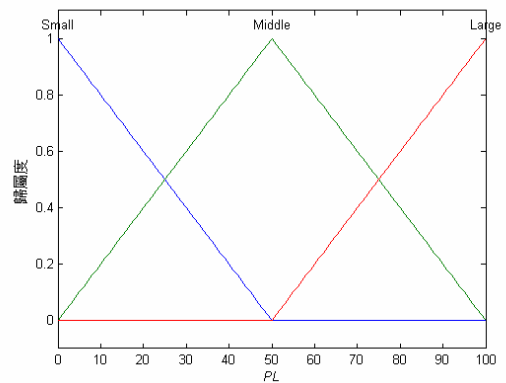
(a)



(b)



(c)



(d)

圖 3.24 FPR 的輸入以及輸出的歸屬函數。(a) CDH 的歸屬函數；(b) RUL 的歸屬函數；(c) QOS 的歸屬函數；(d) PL 的歸屬函數。

我們根據以上討論的行人特徵、實驗的觀察、以及[38]的概念訂定模糊規則，

FPR 共有五個模糊規則列出如下：

$$R^1: \text{IF } CDH \text{ is Small THEN } PL \text{ is Small,} \quad (3.28)$$

$$R^2: \text{IF } RUL \text{ is Small THEN } PL \text{ is Small,} \quad (3.29)$$

$$R^3: \text{IF } QOS \text{ is Small THEN } PL \text{ is Small,} \quad (3.30)$$

$$R^4: \text{IF } QOS \text{ is Large AND } RUL \text{ is Large THEN } PL \text{ is Middle,} \quad (3.31)$$

$$R^5: \text{IF } QOS \text{ is Large AND } CDH \text{ is Large THEN } PL \text{ is Large,} \quad (3.32)$$

FPR 作模糊推論的方法為 Mamdani 的最小值推論法，去模糊化的方法為 COA。

3.6 輔助的行人偵測

以上的行人偵測方法是先偵測橢圓形邊緣再判斷該位置的橢圓是否為人頭，但在某些情況下邊緣並不明顯(如圖 3.25(c))，此時這個方法將會失敗，然而某些孤立的行人雖然沒有明顯的邊緣卻有明顯的行人外型的前景輪廓(如圖 3.25(b))，根據這個特性可以作為孤立行人的偵測系統，用以輔助以邊緣為基礎的行人偵測系統。

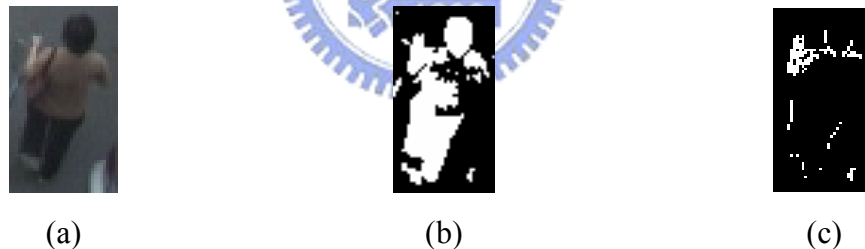


圖 3.25 呈現行人外型的圖例。(a) 輸入影像；(b) 前景影像；(c) 邊緣影像。

以輪廓為基礎的行人偵測系統主要是根據[8]的方法作修改，系統的流程圖如圖 3.26。步驟一的前景偵測在 3.3.1 節已經說明，由於本節的行人偵測系統是輔助 3.5 節的方法，假設在 3.3.1 節得到的前景影像為 I_p ，則扣除在 3.5 節已經偵測到並且持續作追蹤的行人的前景得到新的前景影像 I_p' (如圖 3.27(b))。步驟二是斷開及閉合，斷開是為了消除前景的雜訊點可由式子(2.29)完成，閉合是為了使前景的邊緣更平滑可由式子(2.30)完成，由於接下來會對邊界作標記，要對整張影像作標記非常耗費電腦的記憶體以及時間，因此我們只對影像的部分區域作斷開以及閉合

運算(如圖 3.27(c))。步驟三是邊界提取,利用式子(2.31)可以得到前景的邊界集合(如圖 3.27(d))。步驟四是找出邊界的連通分量並作標記(如圖 3.27(e)),其中不同顏色的邊界彼此並不相連),利用式子(2.32)對每個邊界作標記在步驟五才能分別對不同的邊界作人型邊界識別。步驟五為人型邊界識別,首先判斷藉由步驟四得到的邊界曲線(對每一行邊緣只取最上面的邊緣,如圖 3.27(f))是否類似 Ω 的形狀,參考圖 3.28(a)的示意圖,假設 (u_t, v_t) 為人頭頂點,則 $R2$ 區域的邊界的 v 座標平均分別與 $R1$ 區域的邊界的 v 座標平均以及 $R3$ 區域的邊界的 v 座標平均的差距都要介於 $0.14 \cdot h_p$ 以及 $0.2 \cdot h_p$ 之間,人頭中心的座標 (u_h, v_h) 則等於 $(u_t, v_t + 0.1 \cdot h_p)$;得到的人型邊界不一定是行人的邊界,必須再滿足兩個條件才判斷為行人,第一個條件是 RUL (即 3.5.1.2 定義的 RUL) 的值要大於 0.5,代表人頭下方有身體的前景,另一個條件是人頭左右及上面的前景像素佔圖 3.28(b)的方框內總像素的百分比要小於 25%,這是為了去除由破碎的前景所造成類似人型的邊界。

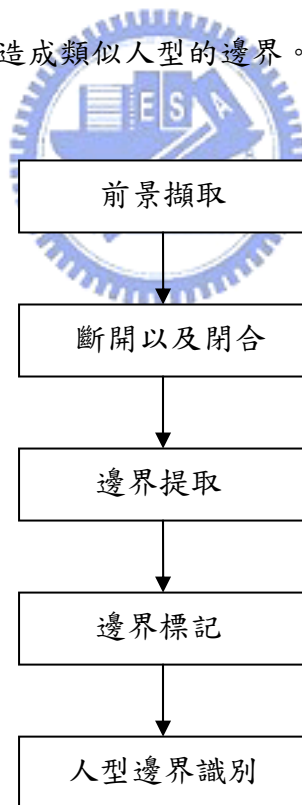


圖 3.26 以輪廓為基礎的行人偵測系統的架構。

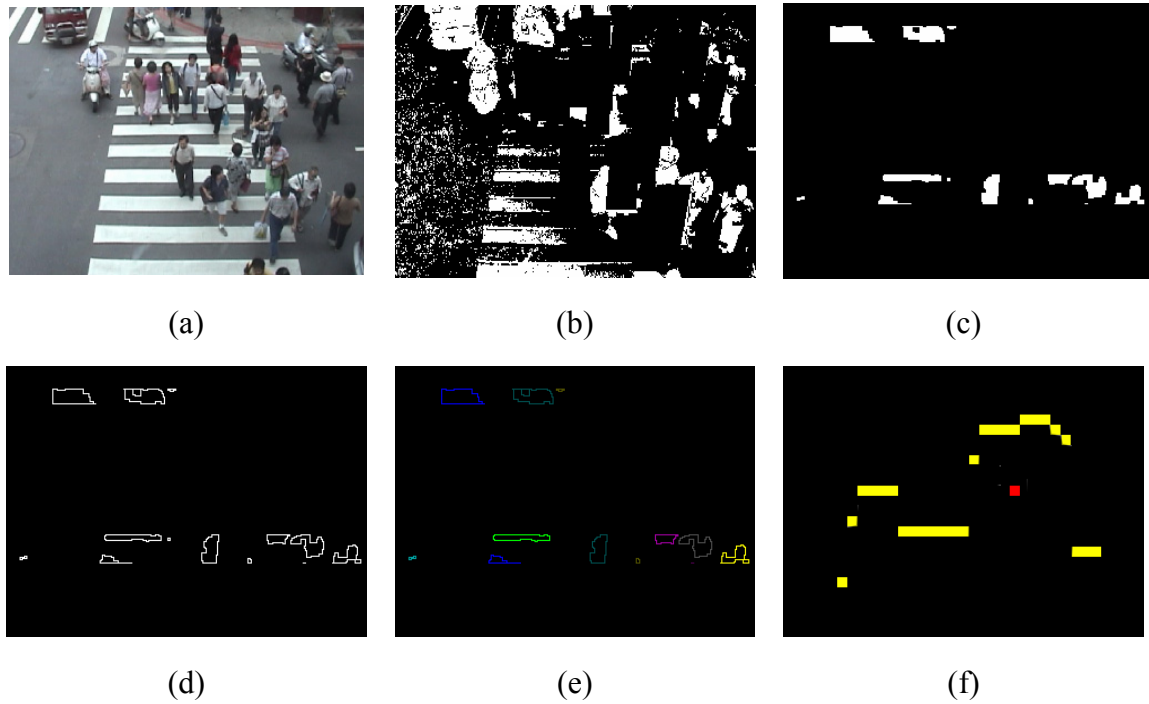


圖 3.27 以輪廓為基礎的行人偵測系統的圖例。(a) 輸入影像；(b) 前景像素；(c) 斷開及閉合後的影像；(d) 邊界影像；(e) 邊界標記；(f) 邊界曲線及找出的人頭中心。

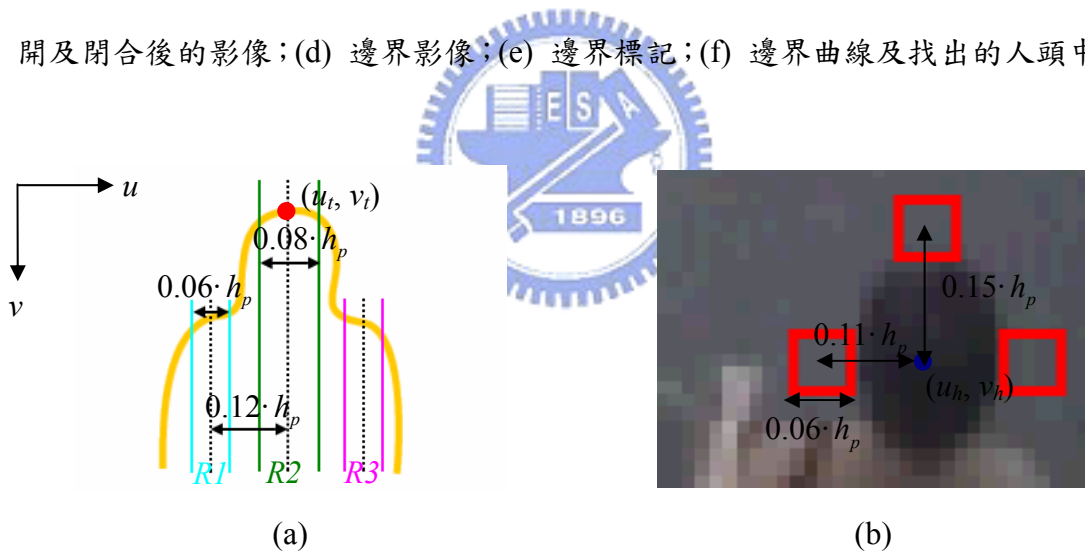


圖 3.28 判斷人型邊界的示意圖。(a) 人型邊界的判斷方法示意圖；(b) 人頭附近的前景偵測範圍。

3.7 行人追蹤

在行人人頭偵測出來之後，接著是對每一個行人目標物作追蹤。追蹤的方法是在接下來的輸入影像中找出與目標物顏色最接近的影像位置，詳細的內容會在 3.7.1 節說明。在目標物被遮蔽或是光線變化劇烈時可能無法在接下來的影像中找

到與目標物顏色匹配的影像位置，此時得根據目標物之前的行進路徑預測接下來的位置，預測的方法是使用簡單的 IIR 濾波器作預測，這個方法會在 3.7.2 節說明。

3.7.1 目標物比對

這個小節討論如何由上一個影像中人頭中心的舊位置 $(u_h^{(i-1)}, v_h^{(i-1)})$ 找出人頭中心在目前影像中最可能的新位置 $(u_h^{(i)}, v_h^{(i)})$ 。一般追蹤的方式是對由上一個影像中偵測到的目標物位置擷取到的特徵與由目前影像中偵測到的目標物位置擷取到的特徵作比對，但是由於我們的行人監控系統只在行人穿越道的近端與遠端作行人偵測，因此我們的追蹤是對由上一個影像中找出的人頭中心位置 $(u_h^{(i-1)}, v_h^{(i-1)})$ 擷取到的特徵與目前影像中舊的人頭中心 $(u_h^{(i-1)}, v_h^{(i-1)})$ 周圍某個範圍內(如圖 3.29 的方框內，方框的大小是根據一般行人經過 0.1 秒所移動的距離而估得，我們定義方框內所有點的集合為 $U_r^{(i)}$)的每個點擷取到的特徵分別作比對，以找出目前影像中人頭中心的新位置。

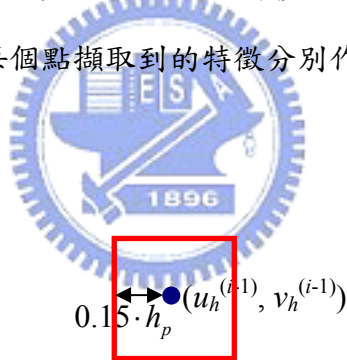


圖 3.29 新人頭中心的限制範圍。

利用目標物的顏色特徵作比對是個很直觀的方法，由於行人的上半身不易被遮蔽因此我們擷取行人的人頭以及上衣的顏色作為比對用的特徵，但是當行人的頭與衣服顏色相同時人頭中心的新位置可能會在整個行人身上游移，因此我們偵測人頭的邊緣作為定位的輔助工具，由於人頭邊緣的 A_e 與 B_e 的比率(參考圖 3.8)隨著行人距離鏡頭遠近會不斷改變，因此橢圓形邊緣的偵測並不容易，因此我們只偵測下半圓形邊緣或上半圓形邊緣，而不是偵測橢圓。

建立初始的目標物特徵的方法如下，假設在行人穿越道的兩端偵測到的人頭中心位置為 $(u_h^{(0)}, v_h^{(0)})$ ，則以某個範圍內(中心為 $(u_h^{(0)}, v_h^{(0)})$ 、邊長為 $0.06 \cdot h_p$ 的方框

內)的每一點作為圓心，偵測半徑為 A_e 的上半圓形邊緣或下半圓形邊緣，修正後的人頭中心位置 (u_h', v_h') 為使得在半圓形上邊緣數目最多的圓心，接著擷取行人上半身顏色的直方圖，擷取顏色的區域如圖 3.30 的上、中、左、以及右的四個方框內(邊長為 $0.08 \cdot h_p$)，共取得四個直方圖，分別為 $H_{bu}^{(0)}$ 、 $H_{bm}^{(0)}$ 、 $H_{bl}^{(0)}$ 、以及 $H_{br}^{(0)}$ ，而三個色彩空間採用 $B-G$ 、 $G-R$ 、以及 $B+G+R$ ， $N=20$ ，與 3.5.1.1 節相同。

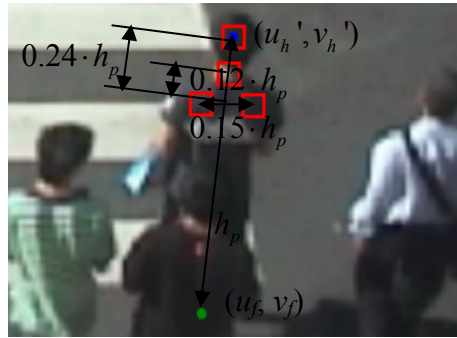


圖 3.30 建立直方圖的區域。

尋找目前影像中最適合的人頭中心的新位置的方法如下，假設 $(u_h^{(i-1)}, v_h^{(i-1)})$ 為上一張影像中人頭中心的位置，則人頭中心新位置只可能會存在 $U_r^{(i)}$ 的範圍內(如圖 3.29)，以 $U_r^{(i)}$ 的每一點作為假設的新的人頭中心並建立其相對的四個直方圖 $H_{buc}^{(i)}$ 、 $H_{bmc}^{(i)}$ 、 $H_{blc}^{(i)}$ 、以及 $H_{brc}^{(i)}$ ，若是分別與 $H_{bu}^{(i-1)}$ 、 $H_{bm}^{(i-1)}$ 、 $H_{bl}^{(i-1)}$ 、以及 $H_{br}^{(i-1)}$ 的相似度都很大則代表該位置才有可能新的人頭位置，我們定義這些直方圖相似度大的位置的集合為 $U_h^{(i)}$ ，為了表示兩個直方圖之間的相似度我們定義函數 D 如下：

$$D = \sum_{k=1}^N \min(H_0(k), H_1(k)) / (3 \cdot n), \quad (3.33)$$

當四個區域的直方圖分別求得的 D 都大於 0.7 才判定該位置有可能是新的人頭位置。 $U_h^{(i)}$ 包含的點可能不只一個，以 $U_h^{(i)}$ 的每一點為圓心，偵測半徑為 A_e 的上半圓形邊緣或下半圓形邊緣(如果在建立初始的目標物特徵時 (u_h', v_h') 是上半圓形的圓心則在此時只偵測上半圓形邊緣，否則只偵測下半圓形邊緣)，則真正新的人頭中心位置 $(u_h^{(i)}, v_h^{(i)})$ 為使得在半圓形上邊緣數目最多的圓心，但是如果所有半圓形

上的邊緣數目都沒有大於 A_e 則真正新的人頭中心位置 $(u_h^{(i)}, v_h^{(i)})$ 為使得四個區域的 D 加起來最大的位置。如果能在目前影像中找到滿足以上條件的新的人頭中心位置 $(u_h^{(i)}, v_h^{(i)})$ ，則更新四個區域的直方圖，更新直方圖的演算法如下：

$$\begin{aligned} H_{bu}^{(i)}(k) &= (1 - \lambda_{bu}) \cdot H_{bu}^{(i-1)}(k) + \lambda_{bu} \cdot H_{bue}^{(i)}(k), & i = 1, 2, 3, \dots, \\ H_{bm}^{(i)}(k) &= (1 - \lambda_{bm}) \cdot H_{bm}^{(i-1)}(k) + \lambda_{bm} \cdot H_{bme}^{(i)}(k), & i = 1, 2, 3, \dots, \\ H_{bl}^{(i)}(k) &= (1 - \lambda_{bl}) \cdot H_{bl}^{(i-1)}(k) + \lambda_{bl} \cdot H_{ble}^{(i)}(k), & i = 1, 2, 3, \dots, \\ H_{br}^{(i)}(k) &= (1 - \lambda_{br}) \cdot H_{br}^{(i-1)}(k) + \lambda_{br} \cdot H_{bre}^{(i)}(k), & i = 1, 2, 3, \dots, \end{aligned} \quad (3.34)$$

其中 $H_{bue}^{(i)}$ 、 $H_{bme}^{(i)}$ 、 $H_{ble}^{(i)}$ 、以及 $H_{bre}^{(i)}$ 是新的人頭中心 $(u_h^{(i)}, v_h^{(i)})$ 所建立的四個區域的直方圖，本篇論文的 λ_{bu} 以及 λ_{bm} 設為 0.2， λ_{bl} 以及 λ_{br} 設為 0.1，實驗顯示有較好的追蹤效果；如果不能在目前影像中找到滿足以上條件的新的人頭中心位置則使用 3.7.2 節介紹的方法預測新的人頭中心在目前影像中的位置，但是若此時 $i=1$ 則代表在上一個影像中被偵測出的人頭可能是機車騎士的人頭以致於移動的速度太快而無法在目前的影像中比對出新的位置，因此我們將刪除該人頭而不作行進路徑預測及後續的追蹤。



3.7.2 行進路徑預測

假設 $(u_h^{(i-1)}, v_h^{(i-1)})$ 為上一張影像中人頭中心的位置，要預測目前影像中新的人頭中心的位置 $(u_h^{(i)}, v_h^{(i)})$ 必須先將 $(u_h^{(i-1)}, v_h^{(i-1)})$ 轉成 WCS 的座標 $(x_p^{(i-1)}, H_p, z_p^{(i-1)})$ ，加上預測的行人目前移動距離得到目前人頭在 WCS 的座標 $(x_p^{(i)}, H_p, z_p^{(i)})$ ，再轉回影像平面中的位置 $(u_h^{(i)}, v_h^{(i)})$ 。預測的目前移動距離可由下列的 IIR 濾波器求得：

$$l_x^{(i)} = \begin{cases} x_p^{(i-1)} - x_p^{(i-2)}, & i = 2, \\ (1 - \lambda_{lx}) \cdot l_x^{(i-1)} + \lambda_{lx} \cdot (x_p^{(i-1)} - x_p^{(i-2)}), & i = 3, 4, 5, \dots, \end{cases} \quad (3.35)$$

$$l_z^{(i)} = \begin{cases} z_p^{(i-1)} - z_p^{(i-2)}, & i = 2, \\ (1 - \lambda_{lz}) \cdot l_z^{(i-1)} + \lambda_{lz} \cdot (z_p^{(i-1)} - z_p^{(i-2)}), & i = 3, 4, 5, \dots, \end{cases} \quad (3.36)$$

其中 $l_x^{(i)}$ 以及 $l_z^{(i)}$ 分別代表在第 i 個影像中行人在 x 以及 z 方向的預測移動距離，本篇論文的 λ_{lx} 以及 λ_{lz} 設為 0.2，實驗顯示有較好的追蹤效果。預測的人頭中心在 WCS

的 $x_p^{(i)}$ 以及 $z_p^{(i)}$ 可由下列式子求得：

$$\begin{aligned}x_p^{(i)} &= x_p^{(i-1)} + l_x^{(i)}, & i = 2, 3, 4, \dots, \\z_p^{(i)} &= z_p^{(i-1)} + l_z^{(i)}, & i = 2, 3, 4, \dots,\end{aligned}\tag{3.37}$$

求得 $x_p^{(i)}$ 以及 $z_p^{(i)}$ 後，即可得到在第 i 個影像中人頭中心在影像平面的預測位置

$(u_h^{(i)}, v_h^{(i)})$ 。



第四章

實驗結果與分析

在此章節中要討論提出的行人穿越道行人監控系統對於行人偵測與追蹤的效能，在 4.1 節中會說明用來測試系統效能所採用的實驗機制，在 4.2 節中會顯示實驗的結果，在 4.3 節中會對實驗結果作分析。

4.1 實驗機制

實驗的設備、環境、以及測試用的影像序列會在 4.1.1 節敘述，用來評估系統效能的方法會在 4.1.2 節說明。



4.1.1 實驗設備

我們的行人穿越道行人監控系統是由 C++ 程式語言實現並在中央處理器為 Pentium 4 1.8G Hz 的個人電腦上執行測試，系統的輸入為每秒 10 張的影像序列，輸入影像的大小為 320x240 像素，CCD 架設的位置如圖 3.3。

實驗環境分為三個場景，分別為雙線道、四線道、以及八線道，並且各包含三個時段，分別為日間、傍晚、以及夜間，總共有九種實驗環境，每種實驗環境有兩段測試用的影像序列，每段影像序列是在行人號誌顯示行人通行時所拍攝得到，所以總計共有十八段影像序列，表 4.1 顯示各個實驗環境在本篇論文中的代稱，表 4.2 顯示各段影像序列在本篇論文中的代稱，圖 4.1 則是在各種實驗環境所分別建立的背景圖。

表 4.1 實驗環境的代稱。

時段	雙線道	四線道	八線道
日間	env1	env4	env7
傍晚	env2	env5	env8
夜間	env3	env6	env9

表 4.2 影像序列的代稱。

時段	雙線道	四線道	八線道
日間	env1-1, env1-2	env4-1, env4-2	env7-1, env7-2
傍晚	env2-1, env2-2	env5-1, env5-2	env8-1, env8-2
夜間	env3-1, env3-2	env6-1, env6-2	env9-1, env9-2

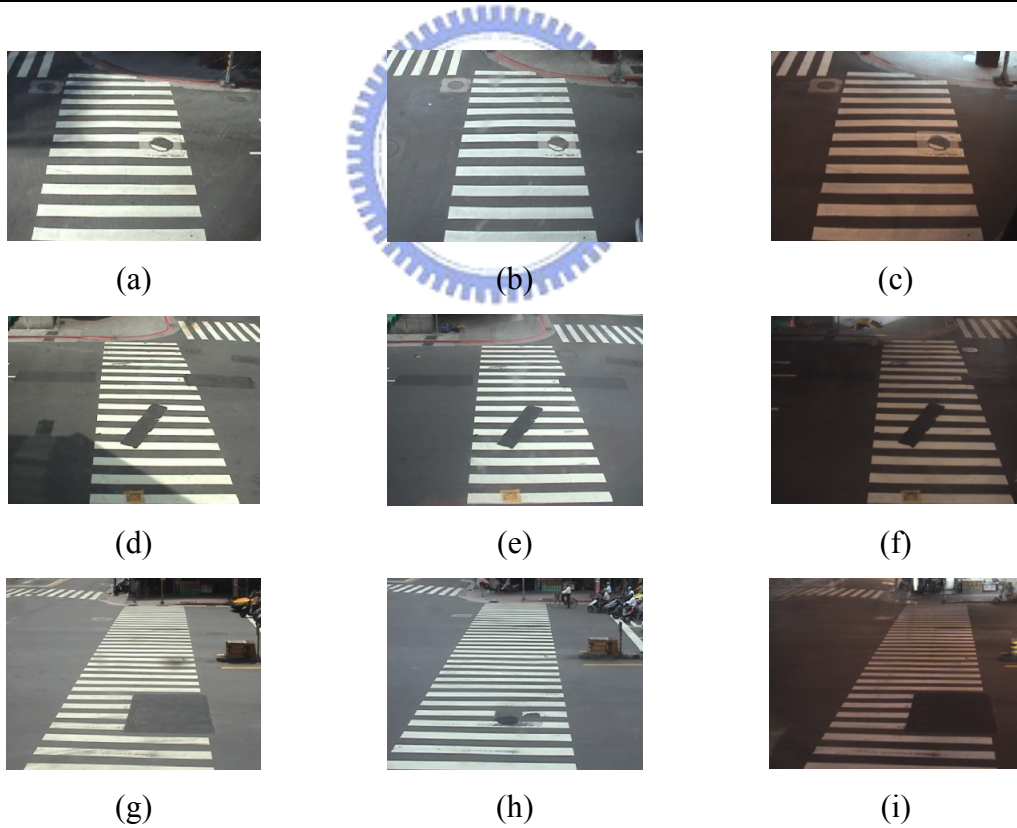


圖 4.1 不同的實驗環境所建立的背景圖。(a) env1；(b) env2；(c) env3；(d) env4；(e) env5；(f) env6；(g) env7；(h) env8；(i) env9。

4.1.2 效能評估

系統的效能評估可以分成兩個主要的部份，一個是行人偵測，另一個部份則是行人追蹤。為了適當地評估行人偵測的效能，我們定義行人偵測成功率以及行人偵測精確度如下列所示：

$$\text{行人偵測成功率} = \frac{\text{偵測到正確行人的數目}}{\text{偵測到正確行人的數目} + \text{未被偵測到的行人數目}} \times 100\%, \quad (4.1)$$

$$\text{行人偵測精確度} = \frac{\text{偵測到正確行人的數目}}{\text{偵測到正確行人的數目} + \text{偵測到錯誤行人的數目}} \times 100\%, \quad (4.2)$$

偵測到正確行人代表偵測到的人頭位置有真正的人頭，如果偵測到的位置並非是人頭而是手臂或是其他物件則代表偵測到錯誤行人，當行人偵測成功率等於 100% 代表所有的行人都被系統偵測到，當行人偵測精確度等於 100% 代表所有偵測到的物件都是正確的行人，一個理想的行人偵測系統必須行人偵測成功率以及行人偵測精確度都接近 100%。行人追蹤的效能評估使用以下的定義：

$$\text{行人追蹤成功率} = \frac{\text{追蹤成功的行人數目}}{\text{偵測到的正確行人數目}} \times 100\%, \quad (4.3)$$

追蹤成功代表經由系統運算得到的行人人頭位置一直維持在實際的行人影像上。

4.2 實驗結果

行人穿越道近端行人與遠端行人的特性有些不同，為了比較近端與遠端的行人偵測以及行人追蹤的效能差異，我們對每種實驗環境分別計算近端與遠端的行人偵測成功率、行人偵測精確度、以及行人追蹤成功率，行人偵測的實驗結果如表 4.3 所示，行人追蹤的實驗結果如表 4.4 所示，對於近端行人而言，正確行人、錯誤行人、以及未被偵測的行人皆不包括正對著鏡頭的行人，因為系統在行人穿越道近端主要是要偵測背對鏡頭要進入行人穿越道的行人，類似地，對於遠端行人而言並不考慮背對著鏡頭的行人。每種實驗環境的一些偵測及追蹤的結果影像如圖 4.2 至 4.10 所示，每個被追蹤的物件皆標示個別的編號以作分別。根據實驗的結果，所有影像序列的平均每張影像處理時間為 0.045 秒，而最長的單張影像處

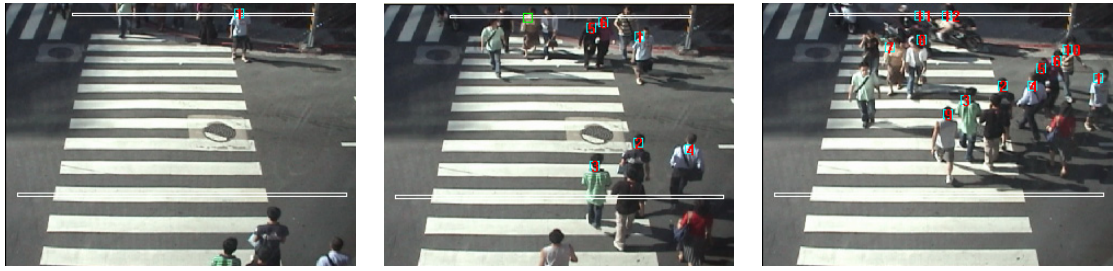
理時間為 0.078 秒，小於 0.1 秒，因此我們的系統能達到每秒處理十張的運算速度。

表 4.3 系統偵測行人的效能。

實驗環境	影像序列	偵測到正確行人的數目	未被偵測到的行人數目	偵測到錯誤行人的數目	行人偵測成功率(%)	行人偵測精確度(%)	
env1	近端	env1-1	9	3	0	84.21	100
		env1-2	7	0	0		
	遠端	env1-1	13	2	5	88.46	79.31
		env1-2	10	1	1		
env2	近端	env2-1	21	3	2	88.89	92.31
		env2-2	3	0	0		
	遠端	env2-1	17	4	6	84.38	81.82
		env2-2	10	1	0		
env3	近端	env3-1	10	2	1	84.62	91.67
		env3-2	1	0	0		
	遠端	env3-1	18	1	3	86.96	86.96
		env3-2	2	2	0		
env4	近端	env4-1	3	0	0	80	100
		env4-2	1	1	0		
	遠端	env4-1	8	3	3	76.19	84.21
		env4-2	8	2	0		
env5	近端	env5-1	17	1	0	78.13	100
		env5-2	9	1	0		
	遠端	env5-1	23	4	7	83.87	78.79
		env5-2	3	1	0		
env6	近端	env6-1	4	3	0	72.22	100
		env6-2	9	2	0		
	遠端	env6-1	6	4	1	64.70	84.62
		env6-2	5	2	0		
env7	近端	env7-1	7	2	2	85.71	85.71
		env7-2	5	0	0		
	遠端	env7-1	5	2	0	72.73	88.89
		env7-2	3	1	1		
env8	近端	env8-1	9	1	0	92.86	92.86
		env8-2	4	0	1		
	遠端	env8-1	3	0	1	77.78	87.5
		env8-2	4	2	1		
env9	近端	env9-1	4	1	0	83.33	83.33
		env9-2	1	0	1		
	遠端	env9-1	5	4	1	61.54	80
		env9-2	3	1	1		

表 4.4 系統追蹤行人的效能。

實驗環境		影像序列	偵測到正確行人的數目	追蹤成功的行人數目	行人追蹤成功率(%)
env1	近端	env1-1	9	9	100
		env1-2	7	7	
	遠端	env1-1	13	12	91.30
		env1-2	10	9	
env2	近端	env2-1	21	20	95.83
		env2-2	3	3	
	遠端	env2-1	17	16	92.59
		env2-2	10	9	
env3	近端	env3-1	10	10	100
		env3-2	1	1	
	遠端	env3-1	18	17	95
		env3-2	2	2	
env4	近端	env4-1	3	3	100
		env4-2	1	1	
	遠端	env4-1	8	7	93.75
		env4-2	8	8	
env5	近端	env5-1	17	15	92
		env5-2	9	9	
	遠端	env5-1	23	19	84.62
		env5-2	3	3	
env6	近端	env6-1	4	4	92.31
		env6-2	9	8	
	遠端	env6-1	6	5	72.72
		env6-2	5	3	
env7	近端	env7-1	7	6	91.67
		env7-2	5	5	
	遠端	env7-1	5	4	75
		env7-2	3	2	
env8	近端	env8-1	9	7	84.62
		env8-2	4	4	
	遠端	env8-1	3	2	71.43
		env8-2	4	3	
env9	近端	env9-1	4	3	80
		env9-2	1	1	
	遠端	env9-1	5	3	62.5
		env9-2	3	2	



(a)

(b)

(c)



(d)

(e)

(f)

圖 4.2 env1-1 偵測及追蹤的部份結果影像。(a) 幀 22；(b) 幀 79；(c) 幀 142；
(d) 幀 232；(e) 幀 328；(f) 幀 562。



(a)

(b)

(c)



(d)

(e)

(f)

圖 4.3 env2-1 偵測及追蹤的部份結果影像。(a) 幀 52；(b) 幀 109；(c) 幀 202；
(d) 幀 310；(e) 幀 361；(f) 幀 463。



(a)

(b)

(c)



(d)

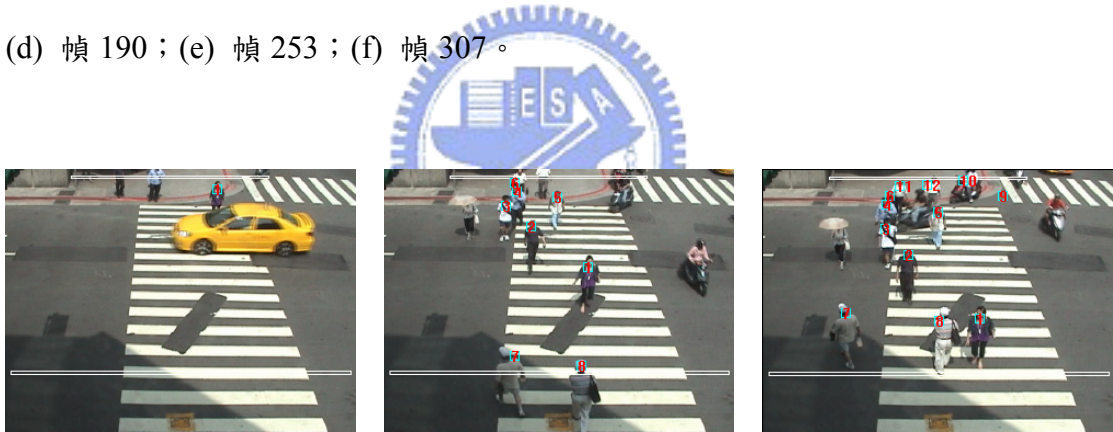


(e)



(f)

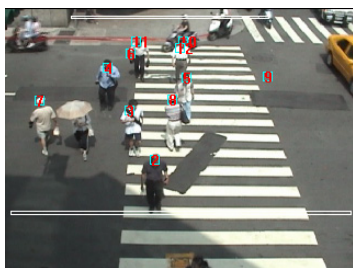
圖 4.4 env3-1 偵測及追蹤的部份結果影像。(a) 幀 40；(b) 幀 82；(c) 幀 136；
(d) 幀 190；(e) 幀 253；(f) 幀 307。



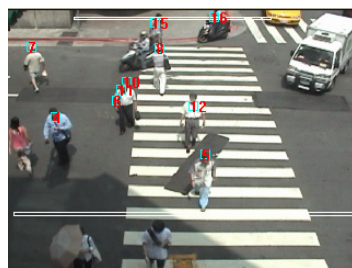
(a)

(b)

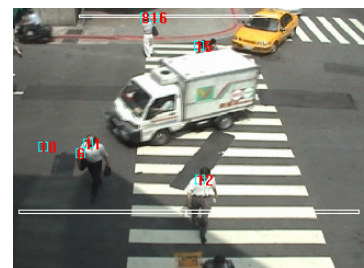
(c)



(d)

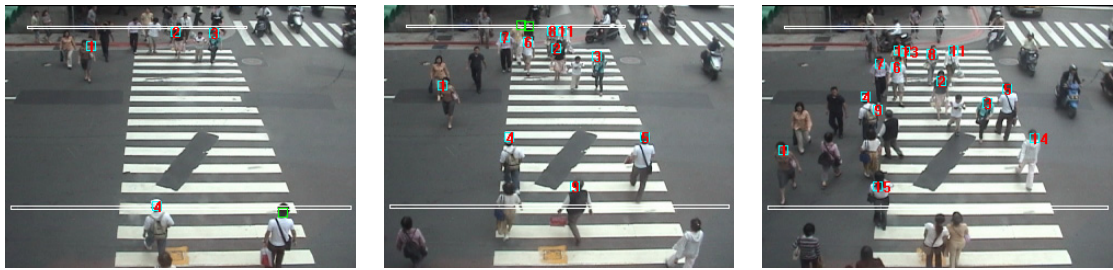


(e)



(f)

圖 4.5 env4-1 偵測及追蹤的部份結果影像。(a) 幀 262；(b) 幀 421；(c) 幀 463；
(d) 幀 526；(e) 幀 622；(f) 幀 700。



(a)

(b)

(c)

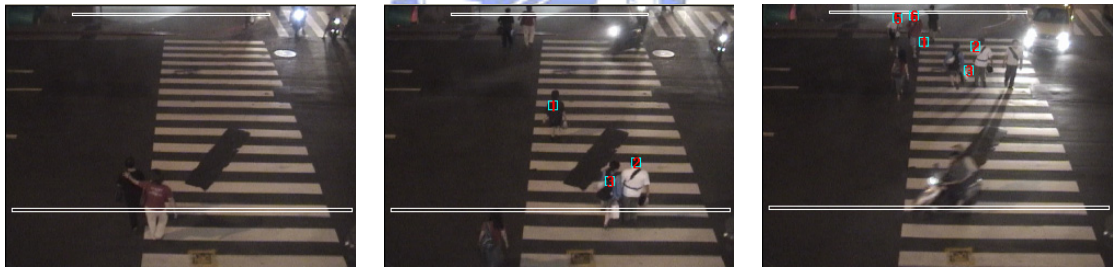


(d)

(e)

(f)

圖 4.6 env5-1 偵測及追蹤的部份結果影像。(a) 幀 64；(b) 幀 130；(c) 幀 199；
(d) 幀 244；(e) 幀 322；(f) 幀 406。



(a)

(b)

(c)

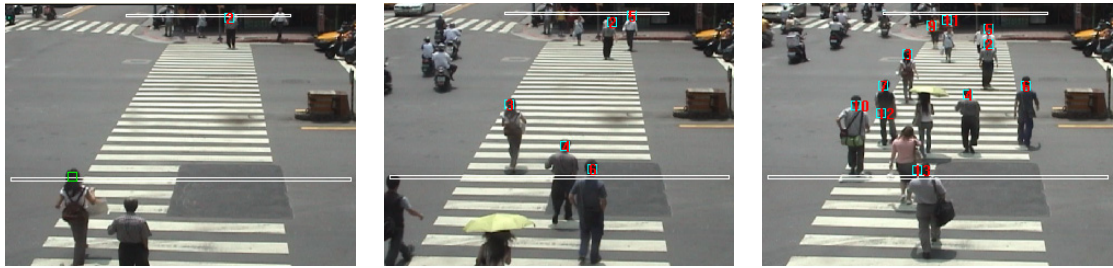


(d)

(e)

(f)

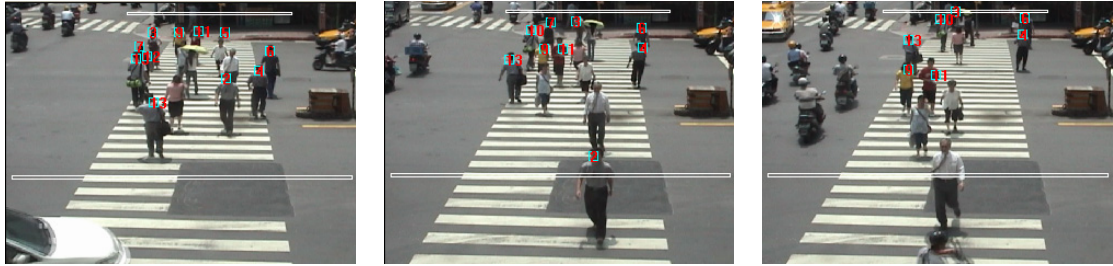
圖 4.7 env6-1 偵測及追蹤的部份結果影像。(a) 幀 73；(b) 幀 304；(c) 幀 484；
(d) 幀 613；(e) 幀 877；(f) 幀 982。



(a)

(b)

(c)

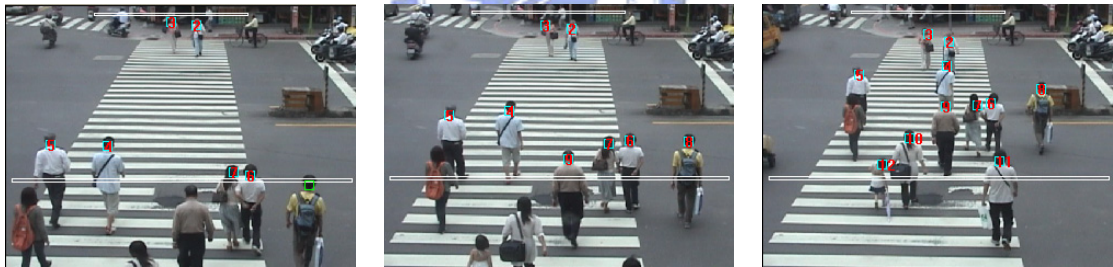


(d)

(e)

(f)

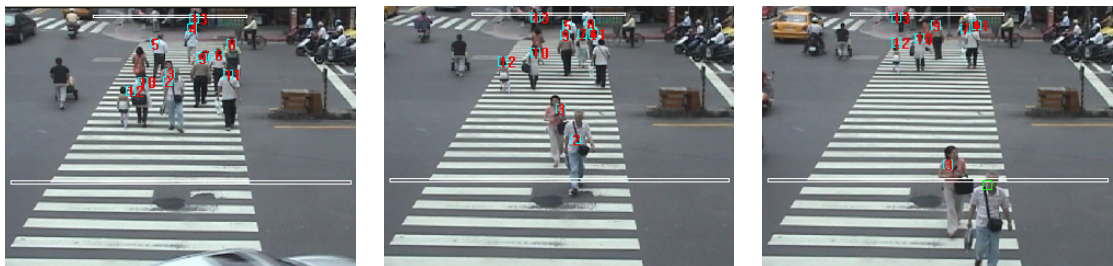
圖 4.8 env7-1 偵測及追蹤的部份結果影像。(a) 幀 61；(b) 幀 145；(c) 幀 271；
(d) 幀 382；(e) 幀 505；(f) 幀 601。



(a)

(b)

(c)



(d)

(e)

(f)

圖 4.9 env8-1 偵測及追蹤的部份結果影像。(a) 幀 205；(b) 幀 247；(c) 幀 337；
(d) 幀 514；(e) 幀 634；(f) 幀 727。

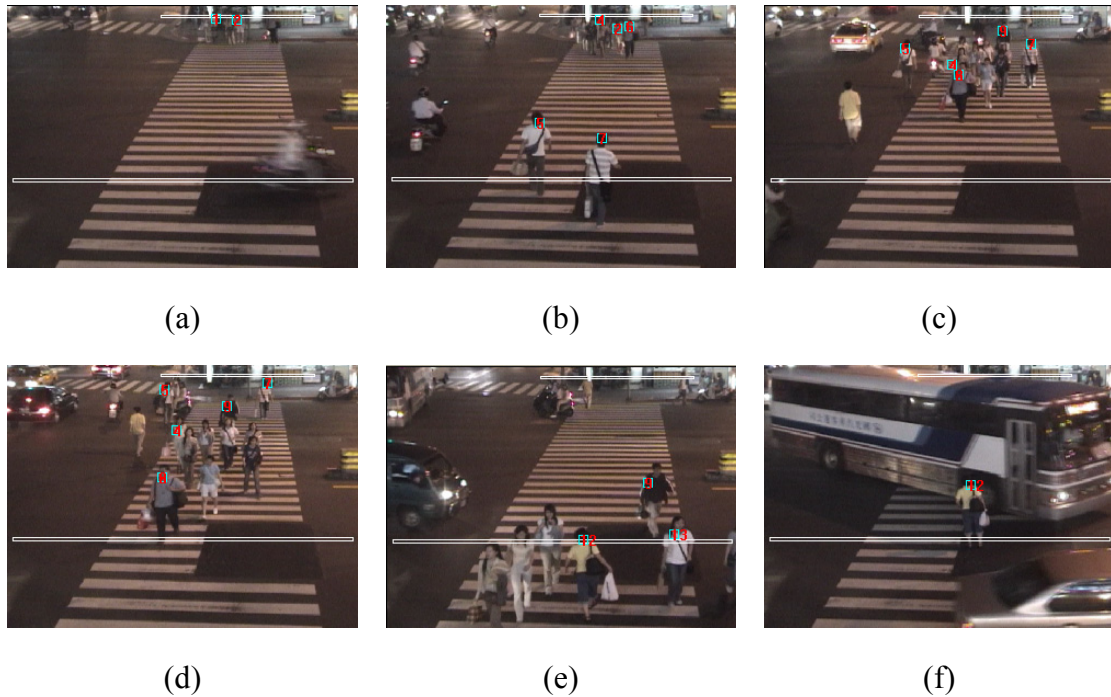


圖 4.10 env9-1 偵測及追蹤的部份結果影像。(a) 幀 79；(b) 幀 244；(c) 幀 496；(d) 幀 613；(e) 幀 856；(f) 幀 1060。

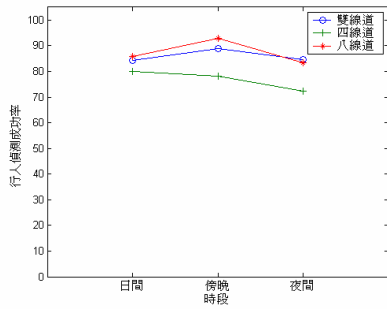
4.3 實驗分析



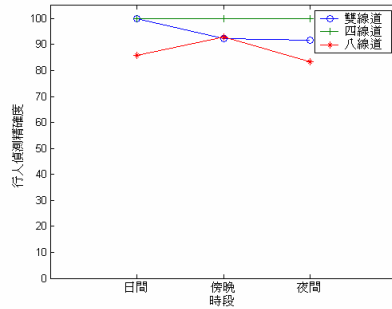
由表 4.3 可以看出除了遠端夜間四線道、以及遠端夜間八線道外，在其它的情況下系統的行人偵測成功率都在 70%以上，而行人偵測精確度在所有的影像序列中則都在 70%以上，由表 4.4 可以看出除了遠端夜間四線道、遠端日間八線道、遠端傍晚八線道、以及遠端夜間八線道外，在其它的情形下系統的行人追蹤成功率都在 80%以上。為了分析光源與行人穿越道長度對於系統效能的影響，我們對不同時段的實驗結果以及不同場景的實驗結果分開作討論。圖 4.11(a)是近端行人偵測成功率與時段的關係圖，雙線道、四線道、以及八線道場景的曲線都類似水平線，顯示出近端行人偵測成功率不太受時段影響；圖 4.11(b)是近端行人偵測精確度與時段的關係圖，該圖顯示出近端行人偵測精確度似乎不受時段影響；圖 4.12(a)是遠端行人偵測成功率與時段的關係圖，該圖顯示出四線道以及八線道場景在夜間時段的遠端行人偵測成功率急遽變差；圖 4.12(b)是遠端行人偵測精確度與時段

的關係圖，該圖顯示出遠端行人偵測精確度不太受時段的影響；圖 4.13(a)是近端行人偵測成功率與行人穿越道長度的關係圖，顯示近端行人偵測成功率與行人穿越道長度似乎沒有相關性；圖 4.13(b)是近端行人偵測精確度與行人穿越道長度的關係圖，顯示近端行人偵測精確度與行人穿越道長度之間似乎沒有相關性；圖 4.14(a)是遠端行人偵測成功率與行人穿越道長度的關係圖，該圖顯示出隨著行人穿越道的長度變長遠端行人偵測成功率變得愈差；圖 4.14(b)是遠端行人偵測精確度與行人穿越道長度的關係圖，顯示遠端行人偵測精確度與行人穿越道長度之間似乎沒有相關性；圖 4.15 是行人追蹤效能與時段的關係圖，顯示四線道與八線道的場景隨著光源變暗遠端行人追蹤成功率會變得愈差；圖 4.16 是行人追蹤效能與行人穿越道長度的關係圖，顯示隨著行人穿越道的長度變長行人追蹤成功率變得愈差。根據以上討論可以得知幾個較明顯的關係：當光源愈昏暗或是行人穿越道愈長，則遠端的行人偵測成功率會愈低；當行人穿越道愈長，則不論近端或遠端的行人追蹤成功率都會愈低；另外，由表 4.3 以及 4.4 可以看出近端的行人偵測精確度以及行人追蹤成功率大致上都比遠端的高。



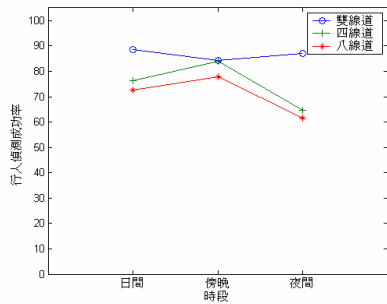


(a)

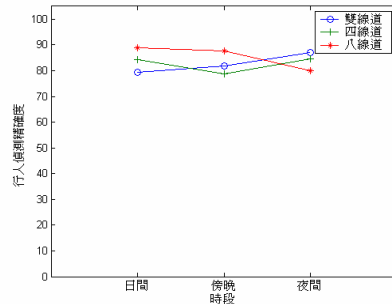


(b)

圖 4.11 近端行人偵測效能與時段的關係。(a) 行人偵測成功率；(b) 行人偵測精確度。

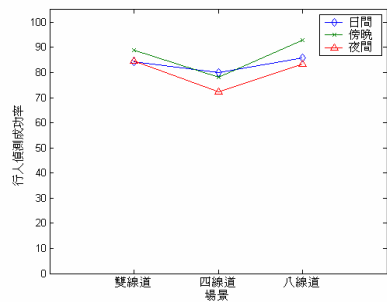


(a)

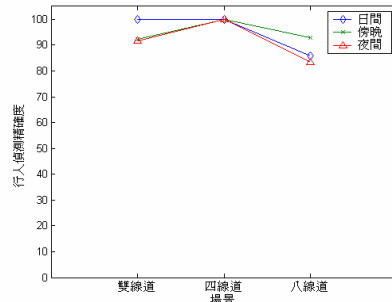


(b)

圖 4.12 遠端行人偵測效能與時段的關係。(a) 行人偵測成功率；(b) 行人偵測精確度。

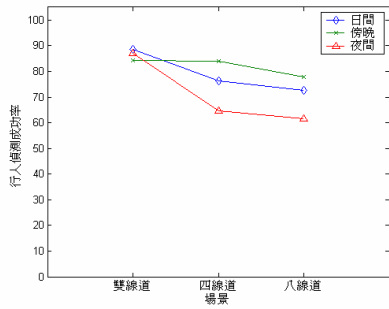


(a)

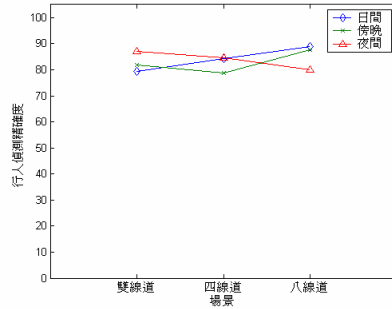


(b)

圖 4.13 近端行人偵測效能與行人穿越道長度的關係。(a) 行人偵測成功率；(b) 行人偵測精確度。

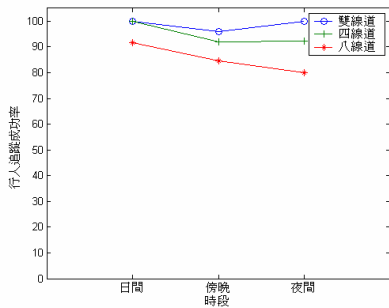


(a)

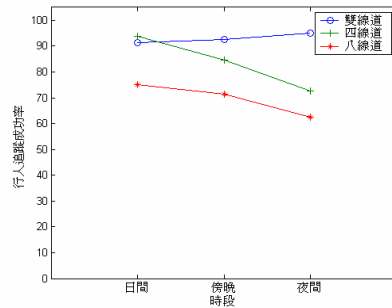


(b)

圖 4.14 遠端行人偵測效能與行人穿越道長度的關係。(a) 行人偵測成功率；(b) 行人偵測精確度。

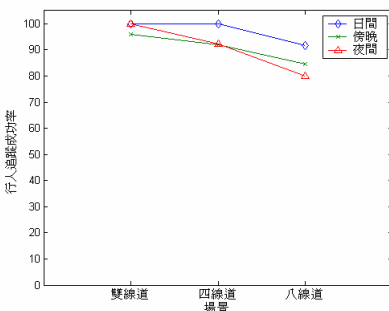


(a)

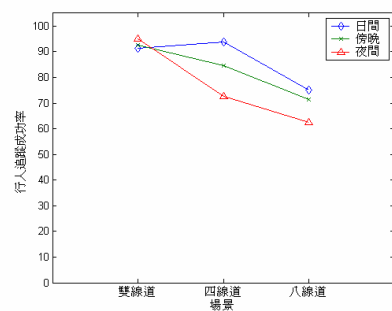


(b)

圖 4.15 行人追蹤效能與時段的關係。(a) 近端；(b) 遠端。



(a)

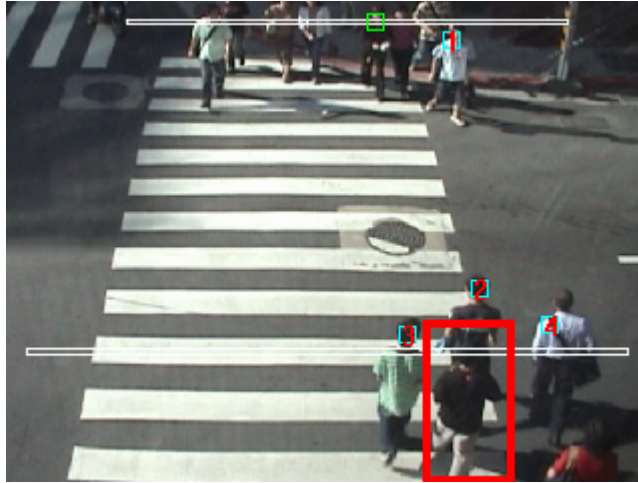


(b)

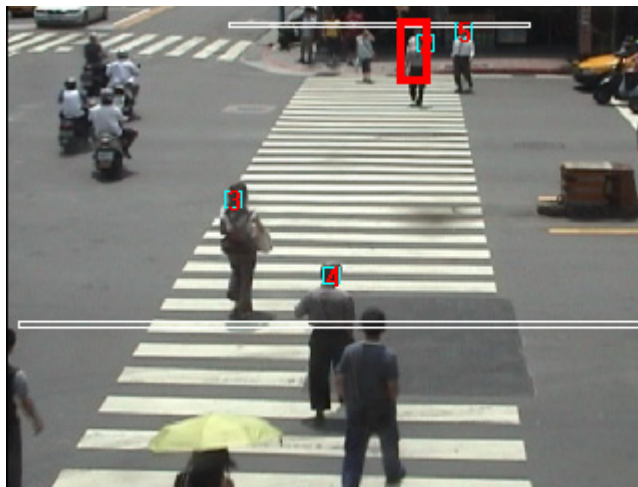
圖 4.16 行人追蹤效能與行人穿越道長度的關係。(a) 近端；(b) 遠端。

系統偵測不出的行人主要可分成三類：1) 與背景或前景顏色相近的行人(如圖 4.17(a))；2) 被前面行人遮蔽的行人(如圖 4.17(b))；3) 沒有顯現正常人頭的行人，例如拿傘的行人(如圖 4.17(c))，對於這類的行人未來必須再發展一個偵測的子系統。系統偵測出的錯誤行人則是外形類似行人的物件，例如機車騎士(如圖 4.18(a)的編號 54)，這是實驗時最常出現的錯誤行人，另一種錯誤行人則是恰巧形成類似行人的圖樣(如圖 4.18(b)的編號 12 以及 4.18(c)的編號 6)。在追蹤的部份，出現追蹤失敗的情形主要有三類：1) 行人動作變化太大，例如行人突然轉頭或是突然轉身(如圖 4.19 的編號 16)；2) 追蹤中的行人被其它前景持續遮蔽(如圖 4.20 的編號 61)；3) 光源變化太大，使得追蹤中的行人顏色變化太大導致追蹤失敗(如圖 4.21 的編號 5)。從實驗的觀察可以知道偵測出錯誤行人的一個主要原因是由機車騎士所造成，追蹤失敗的原因則有追蹤目標被其它車輛遮蔽以及光源變化太大，為了提升行人穿越道行人監控系統的行人偵測及追蹤的效能，未來必須增加偵測以及追蹤車輛的能力，並且要能偵測光源變化以減少光源劇烈變化時對追蹤的影響。

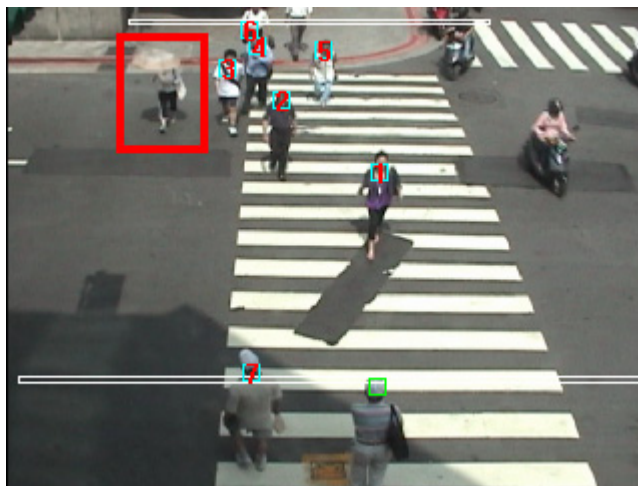




(a)

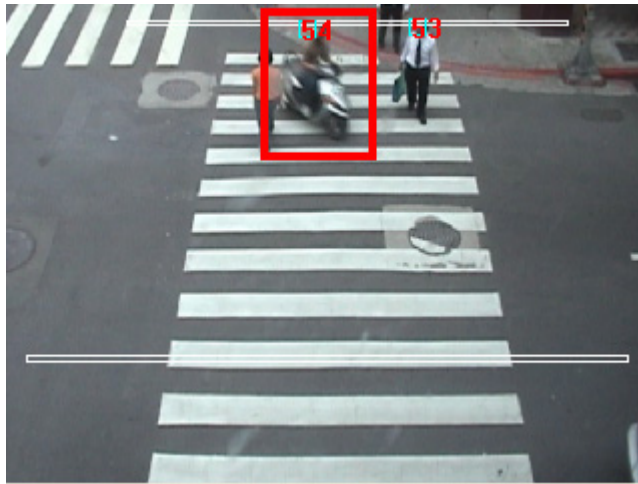


(b)

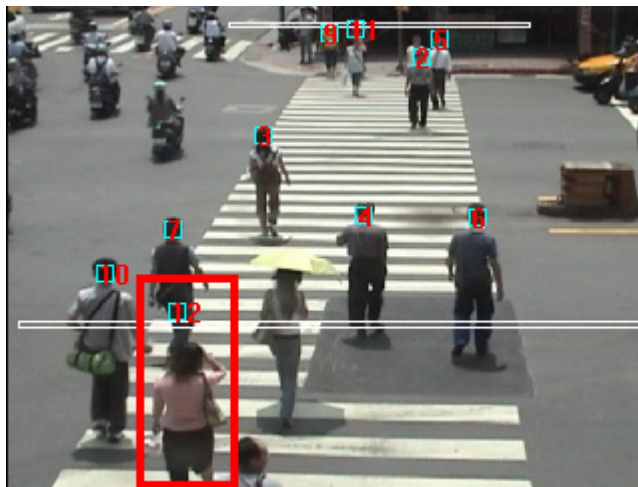


(c)

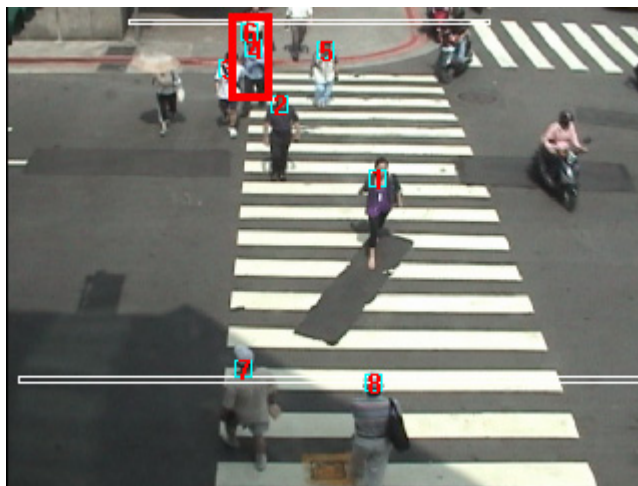
圖 4.17 系統偵測不出的行人。(a) 與前景行人顏色相近的行人；(b) 被前面行人遮蔽的行人；(c) 沒有顯現正常人頭的行人。



(a)



(b)



(c)

圖 4.18 系統偵測出的錯誤行人。(a) 機車騎士(編號 54)；(b) 近端形成類似行人的圖樣(編號 12)；(c) 遠端形成類似行人的圖樣(編號 6)。



(a)

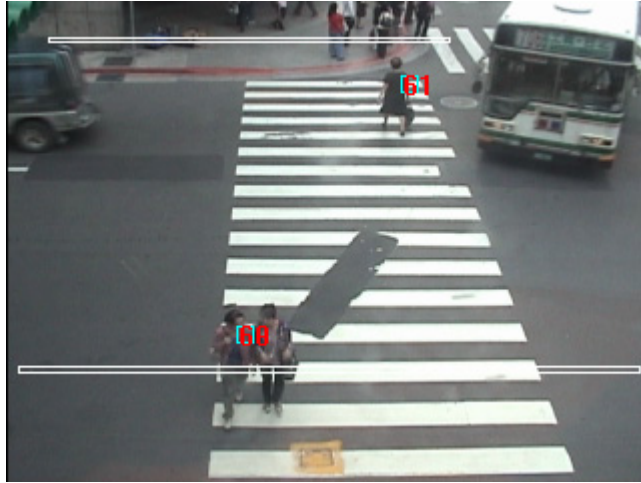


(b)



(c)

圖 4.19 行人突然轉身導致行人追蹤失敗的例子(編號 16)。(a) 幀 139；(b) 幀 163；(c) 幀 187。



(a)



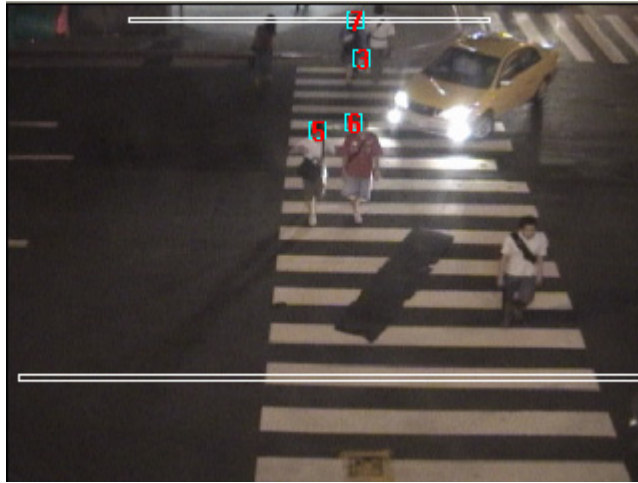
(b)



(c)

圖 4.20 追蹤中的行人被其它前景持續遮蔽導致行人追蹤失敗的例子(編號 61)。

(a) 幀 1786 ; (b) 幀 1816 ; (c) 幀 1870 。



(a)



(b)



(c)

圖 4.21 光源變化太大導致行人追蹤失敗的例子(編號 5)。(a) 幀 598;(b) 幀 619;
(c) 幀 667。

為了顯示出目前的行人偵測及追蹤方法並不適用於行人穿越道場景，我們以兩個較具代表性的方法與本論文提出的方法作比較，這兩個方法分別為以輪廓為基礎的方法[8]以及以膚色為基礎的方法[9]，兩者的系統架構都是對每張輸入影像偵測行人，再對由相鄰的輸入影像中偵測出的行人作匹配及追蹤，因此系統效能的好壞主要是取決於行人偵測的效果，接下來我們會比較這兩個方法以及本論文提出的方法的行人偵測結果。圖 4.22(a)是測試用的輸入影像，行人偵測的範圍(即人頭的可能位置)為兩條綠線之間，其中範圍內的行人數目共 29 人；圖 4.22(b)是本論文提出的方法的結果，利用區域偵測、全域追蹤的作法，最後偵測到正確行人的數目為 21 人，偵測到錯誤的行人數目為 2 人；圖 4.22(c)是以輪廓為基礎的方法的結果，偵測到正確行人的數目為 6 人，偵測到錯誤的行人數目為 0 人；圖 4.22(d)是以膚色為基礎的方法的結果，偵測到正確行人的數目為 4 人，偵測到錯誤的行人數目為 1 人。在圖 4.22(c)中可以看出以輪廓為基礎的方法只能偵測出行人上半身影像沒有與其它前景交疊在一起的行人，這是由於以輪廓為基礎的方法是利利用行人的外型偵測行人，這樣的行人大部分是在人群周圍的行人或是孤立的行人；在圖 4.22(d)中可以看出以膚色為基礎的方法只能偵測人臉對著相機鏡頭的行人，此外如果行人的影像太小則在消去膚色雜訊區塊時人臉的部份影像可能會被去除而導致系統無法偵測出人臉，另外一個無法偵測出人臉的情況則是當衣服顏色與膚色相近時，此時人臉與衣服之間不會產生邊緣，因此無法以下半橢圓形邊緣偵測出人臉；由圖 4.22 的例子可以看出以輪廓為基礎的方法以及以膚色為基礎的方法的確不適合用在行人影像交疊嚴重且行人影像太小的行人穿越道場景。

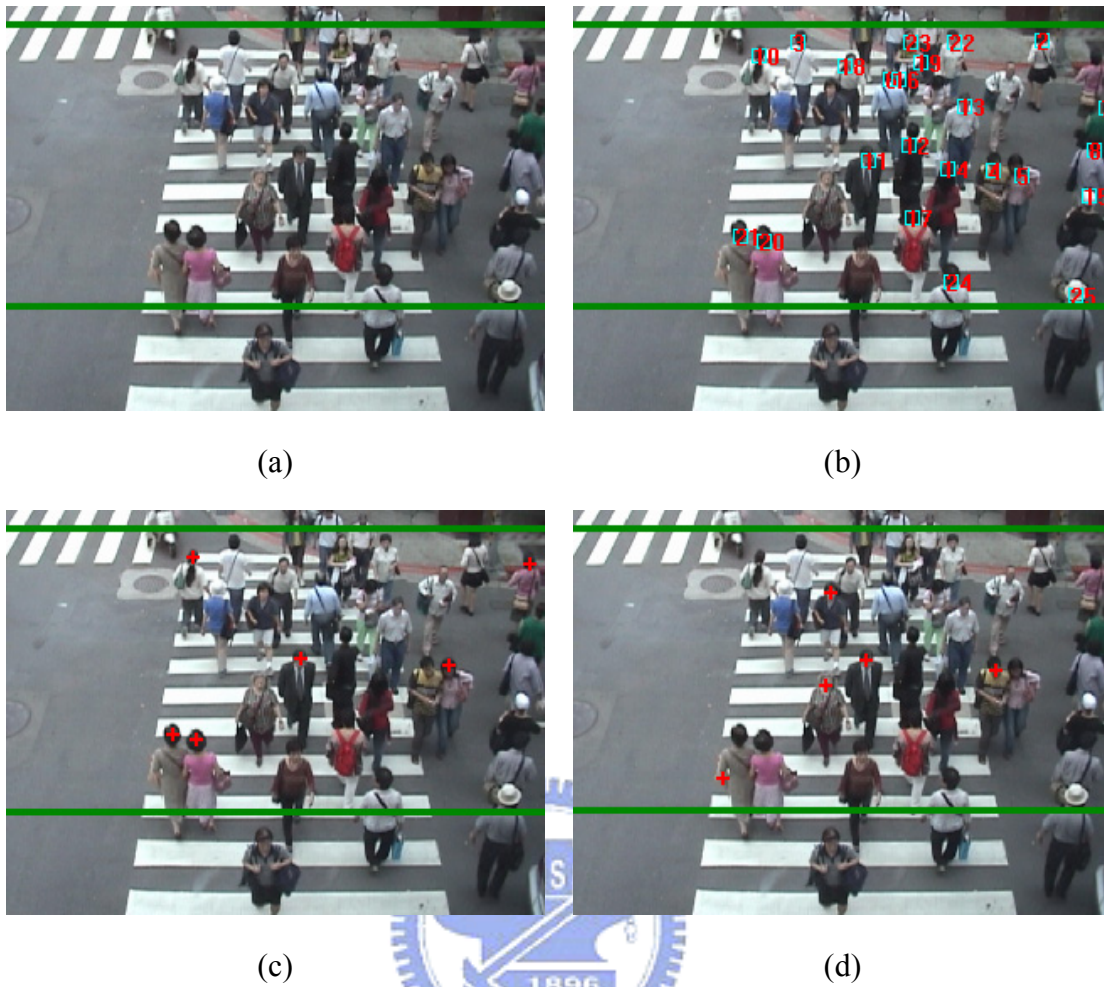


圖 4.22 不同方法的行人偵測結果。(a) 輸入影像；(b) 本論文提出的方法；(c) 以輪廓為基礎的方法 [8]；(d) 以膚色為基礎的方法 [9]。

第五章

結論與未來發展

即時的行人偵測與追蹤是相當受到重視的論題，目前存在許多的相關文獻，但是大部分的方法僅適合用在行人數目不多的情形，而且隨著行人數目以及行人交錯程度的增加，行人偵測的效能將會變的很差，對於行人交錯嚴重的行人穿越道場景而言這些方法是不可行的。根據行人穿越道上行人的特性，本篇論文提出一個方法可以有效地偵測及追蹤行人穿越道上交錯的行人，我們只在行人穿越道的近端與遠端偵測行人，一旦行人被偵測到便開始追蹤行人，如此可以得知行人在所有時間的位置並且能減少系統運算的時間，另外，攝影機鏡頭設置成與道路方向垂直，行人的影像將會左右對稱，行人偵測的方法是先偵測出類似人頭外形的邊緣，接著使用結合許多行人特性的模糊推論系統判斷偵測到的邊緣是否可能是人頭，對於孤立的行人則可以利用行人的輪廓偵測行人，而行人追蹤則是使用行人上半身的顏色，因此不易被遮蔽。實驗結果顯示本論文提出的行人穿越道行人監控系統的行人偵測成功率以及行人偵測精確度大致上都在 70%以上，而行人追蹤成功率則在 80%以上，當光線不是太暗而且行人穿越道不是太長的情形下系統的行人偵測以及追蹤的效能顯示出不錯的結果。

一般攝影機在夜間擷取行人影像的能力較差，可以使用紅外線攝影裝備輔助系統作行人偵測以及追蹤，而在八線道的場景可以嘗試在行人穿越道的兩端都架設攝影機且分別偵測及追蹤該端的行人，並作同步連線以完成全域追蹤，如此可以避免由於行人穿越道長度太長而導致系統效能變得太差。目前系統並沒有在雨天作測試，將來會進一步探討雨天對系統的影響，並增加適當的處理機制。未來

我們將會利用我們的行人穿越道行人監控系統整合整個交通監控系統，以達到交通動線的最佳化管理。



參考文獻

- [1] F. Y. Wang, P. B. Mirchandani, and N. Zheng, “Advances and Trends in Research and Development of Intelligent Transportation Systems: An Introduction to the Special Issue,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 222-223, Dec. 2004.
- [2] B. Maurin, O. Masoud, and N. P. Papanikolopoulos, “Tracking All Traffic: Computer Vision Algorithms for Monitoring Vehicles, Individuals, and Crowds,” *IEEE Trans. on Robotics & Automation*, vol. 12, no. 1, pp. 29-36, March 2005.
- [3] L. Y. Deng, N. C. Tang, D. L. Lee, C. T. Wang, and M. C. Lu, “Vision Based Adaptive Traffic Signal Control System Development,” *Int’l Conf. Advanced Information Networking and Applications*, vol. 2, pp. 25-30, Taipei, Taiwan, 2005.
- [4] J. E. Naranjo, C. Gonzalez, J. Reviejo, R. Garcia, and T. Pedro, “Adaptive Fuzzy Control for Inter-Vehicle Gap Keeping,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 4, no.3, pp. 132-142, Sept. 2003.
- [5] A. Escalera, J. M. Armingol, J. M. Pastor, and F. J. Rodriguez, “Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 5, no. 2, pp. 57-68, June 2004.
- [6] S. Hongeng and R. Nevatia, “Multi-Agent Event Recognition,” *Proc. Int’l Conf. Computer Vision*, vol. 2, pp. 84-91, Vancouver, Canada, 2001.
- [7] I. Haritaoglu, D. Harwood, and L. S. Davis, “W4: Real-Time Surveillance of

- People and Their Activities,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, Aug. 2000.
- [8] T. Zhao and R. Nevatia, “Tracking Multiple Humans in Complex Situations,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1208-1221, Sept. 2004.
- [9] Y. Hori, K. Shimizu, Y. Nakamura, and T. Kuroda, “A Real-Time Multi Face Detection Technique Using Positive-Negative Lines-of-Face Template,” *Int’l Conf. Pattern Recognition*, vol. 1, pp. 765-768, Cambridge, UK, 2004.
- [10] J. H. Jean and R. Y. Wu, “Adaptive Visual Tracking of Moving Objects Modeled with Unknown Parameterized Shape Contour,” *IEEE int’l Conf. Networking, Sensing and Control*, vol. 1, pp. 76-81, Taipei, Taiwan, 2004.
- [11] D. B. Yang, H. H. Gonzalez, and L. J. Guibas, “Counting People in Crowds with a Real-Time Network of Simple Image Sensors,” *Proc. IEEE Int’l Conf. Computer Vision*, vol. 1, pp. 122-129, Nice, France, 2003.
- [12] L. Zhao and C. E. Thorpe, “Stereo and Neural Network-Based Pedestrian Detection,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 1, no. 3, pp. 148-154, Sept. 2000.
- [13] F. Xu, X. Liu, and K. Fujimura, “Pedestrian Detection and Tracking with Night Vision,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 6, no. 1, pp. 63-71, March 2005.
- [14] D. Comaniciu, V. Ramesh and P. Meer, “Kernel-Based Object Tracking,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 564-577, May 2003.
- [15] D. Koller, K. Daniilidis, and H. H. Nagel, “Model-Based Object Tracking with Monocular Image Sequences of Road Traffic Scenes,” *Int’l Journal Computer Vision*, vol. 10, no. 3, pp. 257-281, 1993.

- [16] T. Cham and J. M. Rehg, "A Multiple Hypothesis Approach to Figure Tracking," *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 2239-2245, CO, USA, 1999.
- [17] M. Isard and A. Blake, "Condensation-Conditional Density Propagation for Visual Tracking," *Int'l Journal Computer Vision*, vol. 29, no. 1, pp. 5-28, 1998.
- [18] C. Sminchisescu and B. Triggs, "Kinematic Jump Processes for Monocular 3D Human Tracking," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 69-76, Wisconsin, USA, 2003.
- [19] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd Monitoring Using Image Processing," *Electronics and Communication Engineering Journal*, vol. 7, pp. 37-47, Feb. 1995.
- [20] C. S. Regazzoni, A. Tesei, and V. Murino, "A Real-Time Vision System for Crowding Monitoring," *Proc. IECON Int'l Conf. Industrial Electronics, Control, and Instrumentation*, vol. 3, pp. 1860-1864, Tokyo, Japan, 1993.
- [21] A. N. Marana, S. A. Velastin, L. F. Costa, and R. A. Lotufo, "Estimation of Crowd Density Using Image Processing," *IEE Colloquium Image Processing for Security Applications*, pp. 11/1-11/8, 1997.
- [22] T. Coianiz, M. Boninsegna, and B. Caprile, "A Fuzzy Classifier for Visual Crowding Estimates," *IEEE Int'l Conf. Neural Networks*, vol. 2, pp. 1174-1178, Nagoya, Japan, 1996.
- [23] S. Y. Cho, T. W. S. Chow, and C. T. Leung, "A Neural-Based Crowd Estimation by Hybrid Global Learning Algorithm," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 29, no. 4, pp. 535-541, Aug. 1999.
- [24] S. F. Lin, J. Y. Chen, and H. X. Chao, "Estimation of Number of People in Crowded Scenes Using Perspective Transformation," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 31, no. 6, pp. 645-654, Nov. 2001.

- [25] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Upper Saddle River, NJ: Prentice Hall, 2003.
- [26] F. Lv, T. Zhao, and R. Nevatia, "Self-Calibration of a Camera from Video of a Walking Human," *Proc. Int'l Conf. Pattern Recognition*, vol. 1, pp. 562-567, Quebec, Canada, 2002.
- [27] A. Criminisi, I. Reid, and A. Zisserman, "Single View Geometry," *Int'l Journal Computer Vision*, 40(2), pp. 123-148, 2000.
- [28] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Upper Saddle River, NJ: Prentice Hall, 2002.
- [29] A. Koschan and M. Abidi, "Detection and Classification of Edges in Color Images," *IEEE Trans. on Signal Processing Magazine*, vol. 22, no. 1, pp. 64-73, Jan. 2005.
- [30] C. T. Lin and C. S. George Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Upper Saddle River, NJ: Prentice Hall, 1996.
- [31] S. M. Bhandarkar and X. Luo, "An Efficient Background Updating Scheme for Real-Time Traffic Monitoring," *IEEE Int'l Conf. Intelligent Transportation Systems*, vol. 3, no. 6, pp. 859-864, Parma, Italy, 2004.
- [32] R. A. McLaughlin, "Randomized Hough Transform: Better Ellipse Detection," *Proc. IEEE TENCON. Digital Signal Processing Applications*, vol. 1, pp. 409-414, WA, Australia, 1996.
- [33] Y. Jie, N. Kharma, and P. Grogono, "Fast Robust GA-Based Ellipse Detection," *Proc. Int'l Conf. Pattern Recognition*, vol. 2, pp. 859-862, Cambridge, UK, 2004.
- [34] H. K. Lam, O. S. Au, and C. W. Wong, "Automatic White Balancing Using Standard Deviation of RGB Components," *Proc. Int'l Symposium, Circuits and Systems*, vol. 3, pp. 921-924, Vancouver, Canada, 2004.
- [35] M. Swain and D. Ballard, "Color Indexing," *Int'l Journal Computer Vision*, 7(1), pp. 11-32, 1991.

- [36] S. Birchfield, "Elliptical Head Tracking Using Intensity Gradients and Color Histograms," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 232-237, Santa Barbara, California, 1998.
- [37] R. L. Hsu, M. Abdel, and A. K. Jain, "Face Detection in Color Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no.5, pp. 696-706, May 2002.
- [38] P. Carmona, J. L. Castro, and J. M. Zurita, "FRIwE: Fuzzy Rule Identification with Exceptions," *IEEE Trans. on Fuzzy Systems*, vol. 12, no. 1, pp. 140-151, Feb. 2004.

