

# 應用立體視覺與 HOG 及幾何特徵於行人偵測之研究

陳昱宏

長庚大學電機所

m9721041@stmail.cgu.edu.tw

李建德

長庚大學電機所

jdlee@mail.cgu.edu.tw

鄭國祥

工研院機械與系統研究所智慧車輛組安全感測控制技術部

KHCheng@itri.org.tw

**摘要** – 行人偵測功能是智慧型駕駛輔助系統不可或缺的功能之一。本論文以立體視覺為基礎，結合 Geometry-based 特徵與 HOG 特徵，並以改良型 Adaboost 分類器完成行人偵測任務，並可同時得出行人之距離遠近，作為行車判斷之依據。首先，本系統以 MIT 與 INIRA 的行人資料庫進行特徵轉換，並使用事前訓練好的改良型 Ada-Boost 分類器進行分類判斷。經由實驗證明，經由多種行人特徵(HOG 及 Geometry)及多階分類器組合，行人偵測正確率可達 85%，效能優於單獨只用 Harr-Like 特徵(準確率 78.37%)、HOG 特徵(準確率 80.64%)。且處理速度每張影像僅需 70ms，將可滿足實際應用之要求。

**關鍵詞** – 立體視覺、圖像校正、HOG、AdaBoost、V-Disparity

## 一、前言

目前在行人偵測的研究中，根據所使用的感測器，大致可分為三類，第一種是利用可見光攝影機取像，但其缺點為夜間會因為光源不足而造成無法成像的缺點。第二種是利用紅外線攝影機取像[1-4]，它可以應用於夜間偵測行人；其成像原理是使用不可見光的紅外線當作夜間補充光源，使得在夜間光源不足造成無法成像的問題獲得改善，相較於可見光攝影機，紅外線攝影機在白天不開啟紅外線補充光源的時候，與一般可見光攝影機並無差異，而夜間可以藉由不可見光的紅外線作為補充。第三種是用主動式感測器取像，例如：雷達[5]或雷射[6]；這種主動式感測器，提供物體跟攝影機的距離，所以有利於執行車子行進中路況的規劃

或是駕駛安全性的判斷；而且它不受光線或天氣的影響，即使在晚上或下雨天也可以偵測行人，但是缺點是比較難區分行人和其他物體。因此，考慮上述各項取像的優劣，本研究採用紅外線攝影機作為行人偵測系統的取像裝置。

在行人偵測演算法部分，目前雖已有不同的演算法被提出，例如：經由機器學習獲得的分類器：Adaboosting[7-9]、類神經網路(Neural Network)[10-11]、支援向量機(Support vector Machine) [12-13]、馬可夫模型(Markov)[14]等。但這些方法大都無法提供行人對車子的相對位置資訊，這對智慧型駕駛系統而言，將無法提供避免撞擊行人的閃躲策略開發。有鑑於此，本研究使用雙攝影機，依據立體視覺之基本特性架構出三維場景。並在此三維場景之中偵測出行人與攝影機之相對距離。本研究採用雙攝影機的外極線校準，使攝影機取得的影像資訊具有一定的對應關係。對應的左右影像因對應位置相同，因此在對應的位置週遭會具有相同之特徵，以此特徵為基準進行視差計算，可算出一張完整的視差圖。

從視差圖提供的資訊可以計算出該圖像中場景與攝影機的相對距離位置，以此為基準進行三維重建。三維場所提供的資訊可以彌補二維場中不足的資訊，以此為基準進行行人偵測。

## 二、系統與方法

### (一)系統流程

本研究之系統流程如圖 1 所示。首先以校正後的左右圖像作為運算的基底，由於校正後

的左右圖像具有相同起始與結束位置。因此在左右圖像尋找各個可能封閉區塊，進行視差值計算，將可快速判定是否為左右對應區塊。

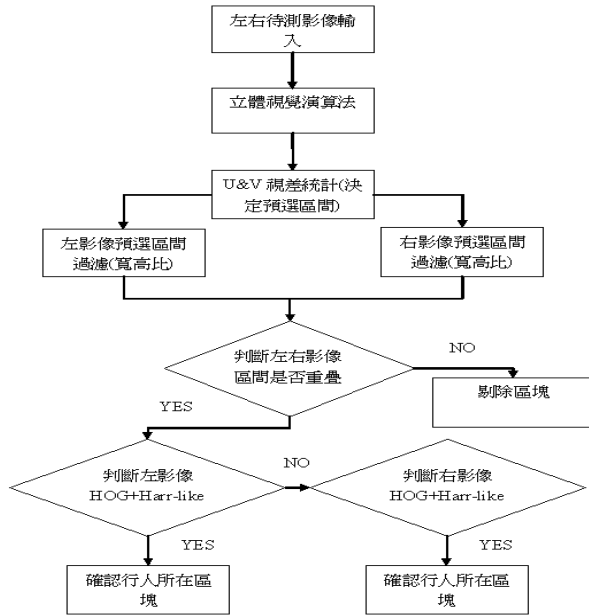


圖 1 系統流程圖

此外，為減少分類器必須處理的圖像資料，使運算速度得以提升。本研究將左右影像輸入在立體視覺基礎上進行 U&V 的視差圖統計，對預選區間進行檢測，再以左右影像的預選區間是否重疊做一次的雜訊濾除，確保預選區間的正确性。

決定的預選區間後再以左影像為主體做 HOG 與 Harr-Like 特徵空間轉換，並用事前訓練好的 Ada-Boost 分類器進行分類判斷。若左影像偵測無行人則進入右影像偵測，左右影像皆無人則確定無行人，若其中一邊影像發現有行人則提出通報。

本研究在行人偵測部分，主要以行人特徵為辨識之依據，故為有效提升系統之辨識率，必須先以大量行人圖像訓練本系統的分類器。因此本研究所用之行人影像資料庫包含 MIT, DaimlerChrysler, 與 INRIA 等知名資料庫，如

圖 2 所示，作為行人正樣本的分類器資料庫。另考量 HOG 對於圖像轉置的強健性不佳，因此本研究以上述三類資料庫為基底進行 HOG 與 Harr-Like 特徵空間轉換，並結合在相同的特徵空間，以產生更高維度的特徵向量。用以解決 HOG 對於圖像轉置強健性不佳的缺點。



圖 2 MIT 資料庫

## (二) 圖像校正

圖像校正目的在計算攝影像之外部參數與內部參數。根據 David Fern'andez Llorca [4] 所提出的立體視覺硬體系統，考量該硬體架設方式與地面有一距離，且與地面有一夾角關係，因此本研究以放置於地面的棋盤格樣本對其進行高度校準，如圖 3 所示。

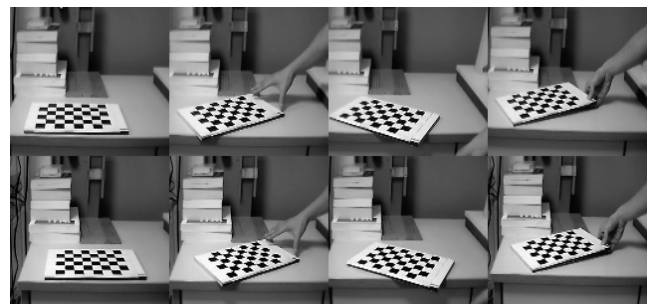


圖 3 左右影像示意圖

此外，為了使樣本具有精確的深度資訊，必須使樣本具有不同的轉角，再對該樣本進行連續取像作為校準之依據影像。以連續的樣本影像對攝影機進行校準，如圖 4 所示。

設定樣本的正規化向量  $\vec{n}_i$ 。考慮仰角內積計算如式(1)：

$$\vec{o} \cdot \vec{n}_i = |\vec{o}| |\vec{n}_i| \cos \alpha_i \quad (1)$$

$$\Rightarrow \cos^{-1} (\vec{o} \cdot \vec{n}_i)$$

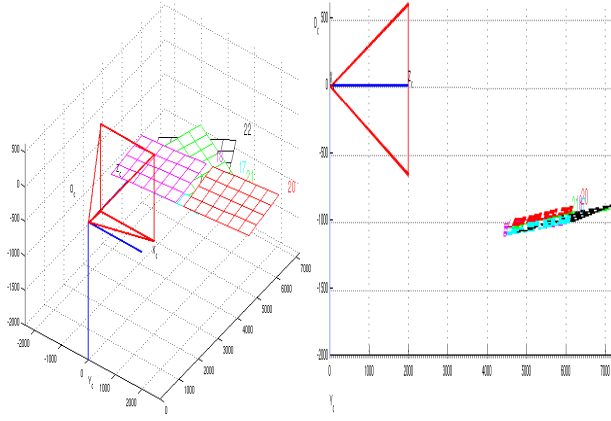


圖 4 棋盤格取像在不同三維位置圖

總計 N 張影像，則平均所有圖像後，取得如下較為穩定的數值：

$$\alpha = \frac{1}{N} \sum_{i=1}^N \alpha_i \quad (2)$$

$$\Rightarrow \frac{1}{N} \sum_{i=1}^N (\cos^{-1} (\vec{o} \cdot \vec{n}_i))$$

提出的高度校準計算方式定義如下。假設  $T_i = (t_{xi}, t_{yi}, t_{zi})$ ，則轉換後的向量則為  $\vec{n}_i = (a_i, b_i, c_i)$  表示。正規化之後的向量用來表示三維點座標定義如式(3)：

$$a_i(x-t_{xi}) + b_i(y-t_{yi}) + c_i(z-t_{zi}) = 0 \quad (3)$$

$$\Rightarrow a_i x + b_i y + c_i z = a_i t_{xi} + b_i t_{yi} + c_i t_{zi}$$

影像平面定義表示  $z=0$ 。則根據式(3)定義線方程式如下：

$$a_i x + b_i y = a_i t_{xi} + b_i t_{yi} + c_i t_{zi} \quad (4)$$

$$\Rightarrow a_i x + b_i y = d_i$$

當  $d_i = a_i t_{xi} + b_i t_{yi} + c_i t_{zi}$  根據式(5)導出兩線相交的平面光學中心點定義為  $p=(0,0,0)$ 。

$$h_i = \frac{d_i}{\sqrt{a_i^2 + b_i^2}} \quad (5)$$

$$\Rightarrow \frac{a_i t_{xi} + b_i t_{yi} + c_i t_{zi}}{\sqrt{a_i^2 + b_i^2}}$$

則計算後真實的  $h'$  從式(4)，(5)，可導出如下：

$$h_i' = h_i \cos \alpha_i \quad (6)$$

最後可以獲得 N 個物體真實的平均高度  $h'$  如下：

$$h' = \frac{1}{N} \sum_{i=1}^N h_i' \quad (7)$$

$$\Rightarrow \frac{1}{N} \sum_{i=1}^N \left( \frac{d_i}{\sqrt{a_i^2 + b_i^2}} \cos \alpha_i \right)$$

$$\begin{pmatrix} x_m \\ y_m \end{pmatrix} = \frac{f}{s_x} \frac{\begin{pmatrix} \cos \phi & 0 & \sin \phi & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cos \Psi & \sin \Psi \\ -\sin \phi & 0 & \cos \phi & 0 & -\sin \Psi & \cos \Psi \end{pmatrix} \begin{pmatrix} X^w [i] \\ Y^w [i] \\ Z^w [i] \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}}{\begin{pmatrix} \cos \phi & 0 & \sin \phi & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cos \Psi & \sin \Psi \\ -\sin \phi & 0 & \cos \phi & 0 & -\sin \Psi & \cos \Psi \end{pmatrix} \begin{pmatrix} X^w [i] \\ Y^w [i] \\ Z^w [i] \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}} + \alpha_i + 0.5 \quad (8)$$

最後再以式(8)進行後續的實際大小換算，以獲得更精準的實際數值。

定義內參數如下：

- 焦長： $f$ ，點大小： $s_x, s_y$ ，
- 影像中心： $o_x, o_y$ ，
- 非線性失真係數： $k_1, k_2 \dots$ 。

定義外參數如下：

- 旋轉係數： $\phi, \square, \Psi$ 。
- 轉換矩陣： $T$ 。

根據式(3)假設的  $T_i = (t_{xi}, t_{yi}, t_{zi})$  假設的座標點位置應用於連續 N 張影像解內外參數的聯立方程式(8)。

### (三)視差影像

本研究採用的視差圖計算法則是以 Graph-based 的方式對左右影像同一區間的視差數值進行計算。其示意圖如圖 5 所示。

正常假設目標物的色彩值與週遭環境有所差異的情況下。使用左影像中同一區間的色彩值作為一類別，與右影像對應，可以在灰階度當中計算其視差數值。其流程如圖 6 所示

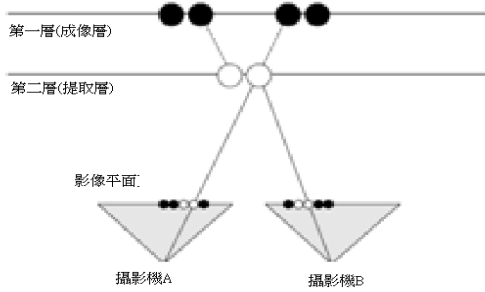


圖 5 雙攝影機視差計算示意圖

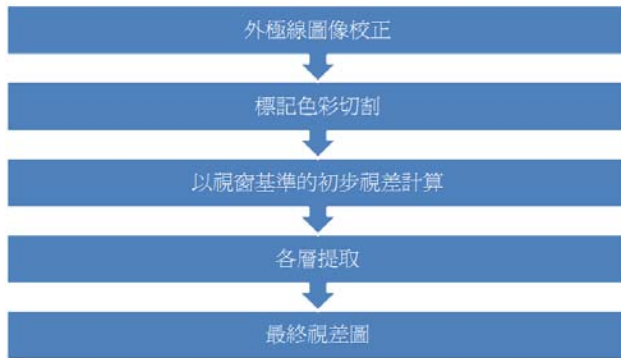


圖 6 Graph-based 視差圖流程

首先初始化視差圖區間且設定掃描的視窗大小為 3x3 或 7x7，將左影像中色彩值分割出

來，且進行標記用於對應右影像之用。

在此，被標記的區塊以集合方式表述如下：

設  $f()$  為標籤函數， $p$  為點資訊， $k$  為標記。則集合為式(9)所示。

$$\forall p \in I : f(p) = f(x,y) = k, k \in \{0 \ 1 \dots \ N\} \quad (9)$$

令  $S$  為提取層之左影像集合，則式(9)可轉為式(10)。

$$\forall s \in S : f(s) = k, k \in \{0 \ 1 \dots \ N\} \quad (10)$$

左右對應之色彩分隔標記圖用以提取相同區間暫存。

由於在色彩分界處會有數值變異的高頻雜訊問題。在此對提取層的視差圖做低通濾波，將其數值平滑降低邊緣造成的雜訊，其低通濾波核結構如式(11)。

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ or } \frac{1}{25} \begin{bmatrix} 3 & 3 & 3 \\ 3 & 1 & 3 \\ 3 & 3 & 3 \end{bmatrix} \quad (11)$$

將相同色彩值對應的區間轉入灰階圖像中做左右相減，計算其區塊的視差值(如式 12)，且相減後的數值圖經過低通濾波將數值平滑之後，即為該視差圖。

$$S \ A \ D = \sum_{i=1}^N \sum_{j=1}^M | I_{Lij} - I_{Rij} | \quad (12)$$

由於產生的數值在-127 至 128 之間，因此配合正常成像格式為 0 至 255 之間，因此將數值正規化至 0 至 255 之間。如式(13)所示。

$$I(x, y) = \begin{cases} \text{if } n \in 2 * k, I(x, y) = g(n_x, n_y) \\ \text{if } n \in k, I(x, y) = g(n_x, n_y) \end{cases} \quad (13)$$

### (四)V-Disparity 與 U-Disparity 影像

以往 Labayrade and Aubert[14]曾提出利用 V-Disparity 觀念偵測障礙物候選者。其研究主要在立體照像機的平面中，把四周環境的情況，分割成水平平面、垂直平面和任何物件平面(傾斜或垂直於地面)，如圖 7 所示。在攝影機的光學軸鏡頭中的剖面圖，任何一個平面都可以投影成一直線。定義平面是為了找出一個最有效的特徵，例如：地平面或是障礙物平面。偵測到這些特徵就為 V-Disparity 影像和 U-Disparity 影像萃取出垂直、水平線段。

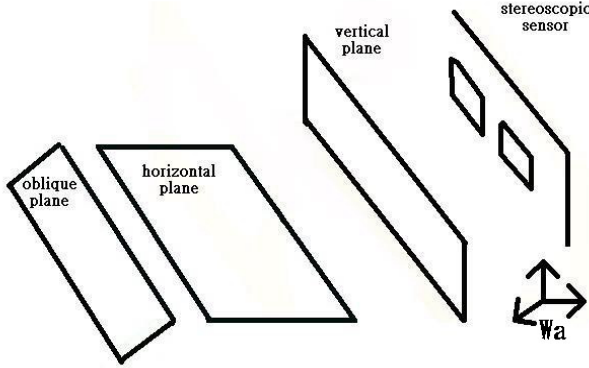


圖 7 空間平面圖

由於計算影像座標(u, v)視差圖 v 軸上的每一個像素值，所以稱之為 V-Disparity[14]。圖 8 經由式(14)，我們可以求出 V-Disparity 影像，它是累加沿著視差圖 v 軸上相同的視差值而產生，所以 V-Disparity 可提供行人候選者最頂端和最底端的邊界，如圖九所示。

$$I_u(v, d) = \sum_{u=0}^{image \rightarrow width} g(d) \quad (14)$$

$$其中 g(d) \begin{cases} 1 & \text{whem } d = I_{DM}(u, v) \\ 0 & \text{otherwise} \end{cases}, I_v(v, d) 為$$

V-Disparity 影像； $I_{DM}(u, v)$  為視差圖； $d$  為視差值； $image \rightarrow width$  為影像的寬度； $image \rightarrow height$  為影像的高度。由圖 9 所示，即可找出行人候選者及其高度。

圖 8 經由式(15)，即可求出 U-Disparity 影像，它是累加沿著視差圖 u 軸上相同的視差值產生的，所以 U-Disparity 可提供行人候選者最左側和最右側的邊界，如圖 10 所示。

$$I_u(v, d) = \sum_{u=0}^{image \rightarrow width} g(d) \quad (15)$$

$$其中 g(d) \begin{cases} 1 & \text{whem } d = I_{DM}(u, v) \\ 0 & \text{otherwise} \end{cases}, I_u(u, d) 為$$

U-Disparity， $I_{DM}(u, v)$  為視差圖； $d$  為視差值； $image \rightarrow width$  為影像的寬度； $image \rightarrow height$  為影像的高度。



圖 8 視差示意圖

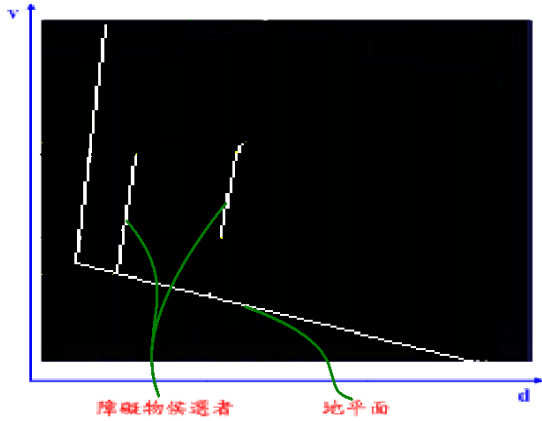


圖9 V-Disparity 示意圖

由圖 10 所示，即可得知行人候選者及其寬度。

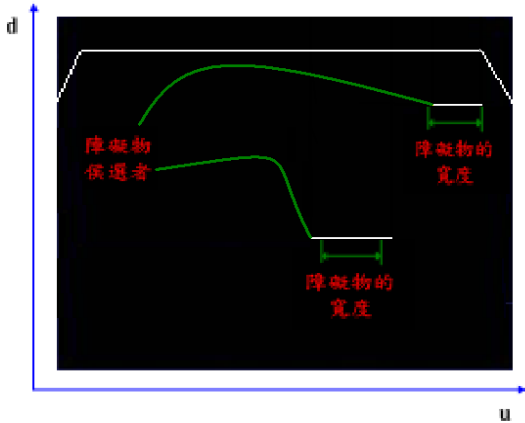


圖 10 U-Disparity 示意圖

綜上所述，V-Disparity 影像中的垂直線代表為行人候選者的高度；U-Disparity 影像中的水平線代表為行人候選者的寬度，結合這些資訊後，我們就可以用紅色線把行人候選者框出來，如圖 11 所示。

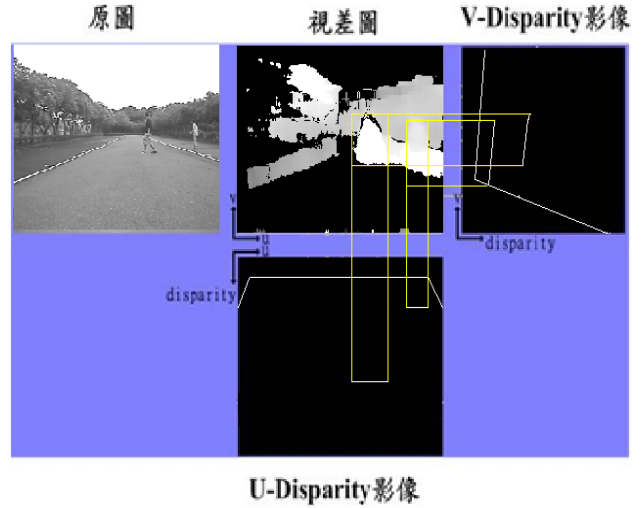


圖 11 行人候選者區塊選定

### (五)改良型 HOG 演算法

以 Navneet Dalal[15]與藤吉弘亘[16]所提出之行人特徵描述演算法 HOG (Histograms of Oriented Gradients)為本研究描述行人之主體特徵。

所謂 HOG 是指對於一個樣本資料取微分統計量，如式(16)所表。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (16)$$

將式(16)所取得的樣本圖像數值以式(17)做一比值轉換，令其數值較具安定性。

$$\theta(u, v) = \tan^{-1} \frac{f_v(u, v)}{f_u(u, v)} \quad (17)$$

其中  $f_u(u, v)$  與  $f_v(u, v)$  個別為式(18)。

$$\begin{cases} f_u(u, v) = I(u+1, v) - I(u-1, v) \\ f_v(u, v) = I(u, v+1) - I(u, v-1) \end{cases} \quad (18)$$

考量模組化的角度為 0~180 度，因此必須將轉換後的角度做一正規化之動作。如式(19)所示。

$$v = \frac{f}{\sqrt{\|v_k\|^2 + \varepsilon^2}} \quad (19)$$

將資料庫中之樣本圖像做 HOG 轉換，其結果如圖 12 所示。圖 13 則表示該樣本的區域 HOG 統計量。

另考量 HOG 特徵描述本身雖對於光亮變化具有強健性，但是對影像旋轉卻不具強健性。以往 Tomoki Watanabe[17]等人曾提出 Co-HOG 的高維 HOG 特徵配對，改善此問題，該 Co-GOH 轉換如圖 14 所示。但是其运算速度不盡理想，無法達到即時應用需求。

有鑑於此，本研究提出一種改良型 HOG 演算法，以 Geometry-based 特徵描述與 HOG 特徵描述結合為同一特徵空間，並採用改良型 Ada-Boost 演算法進行樣本的分類訓練。

由於 HOG 轉換後的特徵值對於光亮變化具有不錯的強健性。因此，本研究採用以此種轉換為行人辨識的最主要特徵值。部分樣本圖像經由此轉換所得的 HOG 特徵值如圖 15 所示。

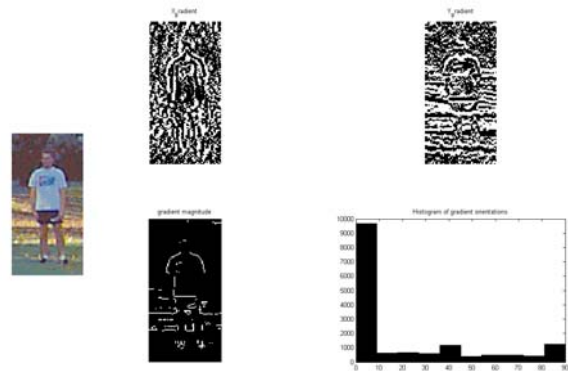


圖 12 全域 HOG 統計量示意圖

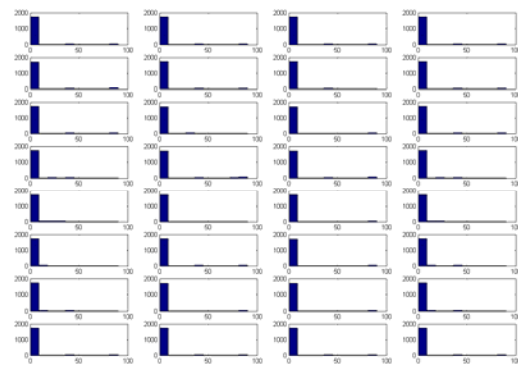


圖 13 區域 HOG 統計量示意圖

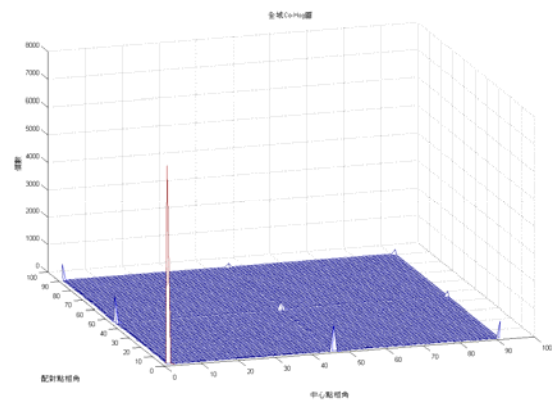


圖 14 全域 Co-HOG 統計量示意圖

所謂 Geometry-based 特徵描述是指對於式(16)

中的  $m(u, v)$  的二維座標位置計算質心，且納入寬高比等元素，使得 HOG 可以獲得幾何上的特徵向量。

對於所有圖像預選區間中的強健邊緣相對於質心的統計量與各自的寬高比統計量，我們可以將其視為一種幾何性質的特徵，其描述方式如式(19) 所示。

其中  $e_s$  表示邊緣佇列， $m$  表示質心， $V$  為邊緣與質心的方向向量。

$$V_{em} = |e(i, j) - m(x, y)| \quad (19)$$

經上述法則運算後的特徵如圖 16 所示。

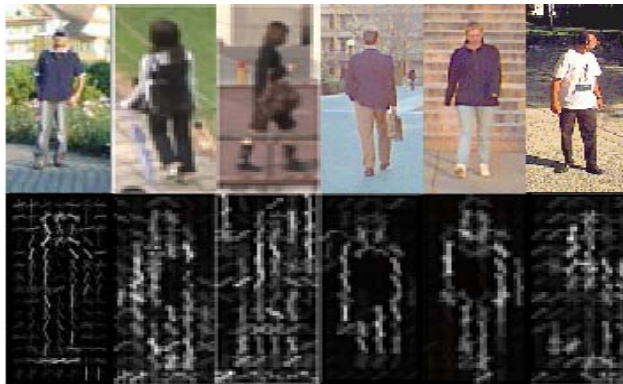


圖 15 HOG 樣本示意圖

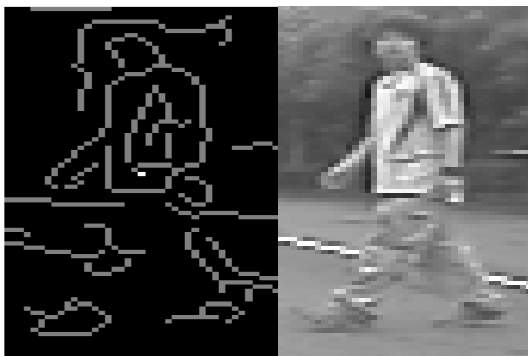


圖 16 幾何特徵轉換示意圖

圖 16 中的高亮白點為質心的表示點，其餘的為強邊緣的表示點，再將其點套用式(19)計算

出方向向量，並與寬高比結合成為一特徵向量。

### 三、實驗結果

參考 Bo Ling [1], Ignacio Parra Alonso [2], Bertozzi [3]... 等人所提出之立體視覺架構，本研究採用光軸平行的架構作為立體視覺的主要硬體架構如圖 17。

以光軸平行的雙攝影機進行取像，其中兩台攝影機需要具備有相同的焦長，鏡面係數，光圈大小，以及相同的編碼模式。此外，考慮兩攝影機間之基線須有一定距離，方能使其影像具有較多的可對應區間，同時考量攝影機鏡頭的成像大小與鏡像失真效應，故本研究設定基線長為 15cm。其實際硬體如圖 15 所示。接著，本研究採用棋盤格樣本作為攝影機校準之標準圖像。



圖 17 本研究所建構之立體視覺系統

另外，在本研究的硬體配置條件下，考量攝影機規格參數(如表 1)，其焦長為 4.3mm 且最遠照射距離約為 20m。但最遠照射距離並不同於最遠有效距離，因此實際測試後發現，該攝影機之最遠有效成像距離為 15m，對應校準後的視差值具有一定的誤差。視差值對應回相對距



離的公式如式(20)所示。

$$R = f \times \left( \frac{b}{d} \right) \quad (20)$$

表 1 攝影機規格

型號	焦長	編碼	有效距離	最低照度
YST-65 E	4.3mm	NTSC:5 37*505	15m	0.01Lux/F1.2 LUX(IRON)

如圖 18 所計算出的視差圖，各點所對應的數值換算回原始距離的對應如表 2。

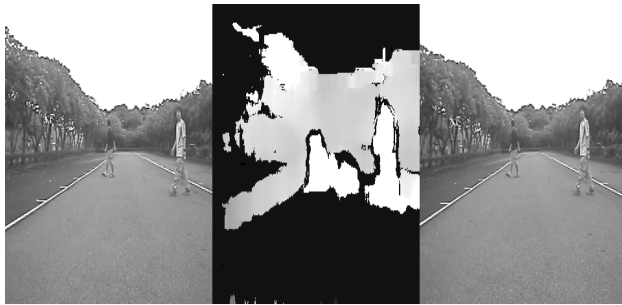


圖 18 左右影像與視差圖

另考量硬體架設時基線為 15cm，攝影機規格焦長為 4.3mm，計算出最遠可視距離為 14.6773m。最小為 0.0043m。其測試的平均誤差如表 3 及表 4 所示。

表 2 距離與誤差對應表

實際有效距離	0(m)	15(m)
計算有效距離	0(m)	14.67(m)
誤差率	0%	2.22%

表 3 平均距離誤差

測試資料	1800
平均誤差	3.72%
最大誤差	6.91%

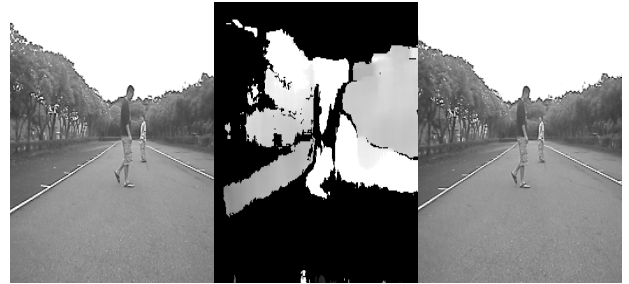


圖 19 輸入之左右影像與視差圖結果

以圖 19 為例，計算出雙攝影機之視差影像後對該視差圖進行 U&V 視差圖計算，並以左右影像的重疊區塊加強確認預選區間。結果如圖 20 所示。取得 U&V 的統計圖表後，以連續的線段區塊表示大量同數值出現的位置，並將其圈選後作出標記，如圖 21 所示。



圖 20 視差圖與 V-Disparity 圖

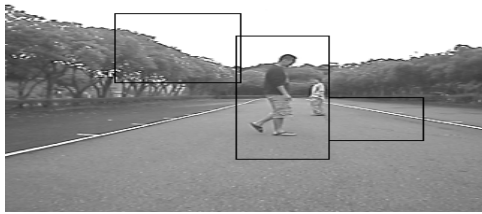


圖 21 U&V 重疊區塊

從圖 21 中我們取得了重疊的預選區塊在二維影像上的座標位置，以此標記好的位置做為分類器迭代的起始位置，開始進行分類。以 MIT 與 INRIA 資料庫訓練好的 HOG 分類器為例進行重疊區塊的篩選，其結果如圖 22 所示。

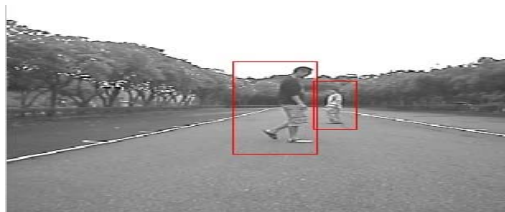


圖 22 左影像重疊區塊確認

此外，本系統為了加快行人辨識速度性，因此判斷順序設定為：若左影像偵測行人成功，則放棄右影像的偵測。這樣既能節省時間亦能提高準確度。系統執行介面如圖 23 所示。



圖 23 運行範例圖

本論文訓練用的分類器是以 MIT 與 INRIA 的樣本資料庫為基準之 HOG, Harr-Like 與 Geometry-based 個別特徵，並且以此分類器對待測的影像進行測試，其結果如表 4、表 5 所示。

表 4 自訂環境測試表

資料庫	Pos	Neg	Features	Test (張)	準確率 (%)	Task (張/ms)
MIT	924	700	HOG	1800(D)	79.53	65.71
MIT	924	700	HOG,Ge	1800(D)	81.26	67.35
MIT	924	700	Harr-Like	1800(D)	73.92	71.74
MIT	924	700	Ge-based	1800(D)	76.71	37.38
MIT	924	700	HOG,Ge, Harr	1800(D)	83.31	86.77
MIT	924	700	HOG	2400(N)	86.94	62.12
MIT	924	700	HOG,Ge	2400(N)	90.41	66.75
MIT	924	700	Harr-Like	2400(N)	81.63	73.91
MIT	924	700	Ge-based	2400(N)	80.11	38.13
MIT	924	700	HOG,Ge, Harr	2400(N)	91.87	88.14
INIRA	2416	700	HOG	1800(D)	80.64	65.93
INIRA	2416	700	HOG,Ge	1800(D)	85.47	69.74
INIRA	2416	700	Harr-Like	1800(D)	78.37	75.67
INIRA	2416	700	Ge-based	1800(D)	72.13	39.27
INIRA	2416	700	HOG,Ge, Harr	1800(D)	89.57	89.38
INIRA	2416	700	HOG	2400(N)	85.11	63.13
INIRA	2416	700	HOG,Ge	2400(N)	92.51	71.24
INIRA	2416	700	Harr-Like	2400(N)	76.98	78.79
INIRA	2416	700	Ge-based	2400(N)	81.48	34.58
INIRA	2416	700	HOG,Ge, Harr	2400(N)	91.18	87.61

表 5 實際環境測試表

Database	Pos	Neg	Features	Test (frame)	Accuracy (%)	Task (frame/ms)
MIT	924	700	HOG	3600(N)	88.37	67.82
MIT	924	700	HOG,Ge	3600(N)	92.74	69.34
MIT	924	700	Harr-Like	3600(N)	83.31	69.56
MIT	924	700	Ge-based	3600(N)	78.69	38.37
MIT	924	700	HOG,Ge, Harr	3600(N)	93.87	83.21
INIRA	2416	700	HOG	3600(N)	87.11	62.93
INIRA	2416	700	HOG,Ge	3600(N)	92.51	69.75
INIRA	2416	700	Harr-Like	3600(N)	79.98	74.72
INIRA	2416	700	Ge-based	3600(N)	81.48	37.66
INIRA	2416	700	HOG,Ge, Harr	3600(N)	93.18	88.32

由表 4 可看出，以 MIT 資料庫而言，可明顯看出使用改良型的 HOG 特徵(準確率 81.26%)，相較於單獨只用 Harr-Like(準確率 73.92%)、HOG 特徵(準確率 79.53%)，可得較好的行人偵測準確率。以 INRIA 資料庫做比較，可明顯看出使用改良型的 HOG 特徵(準確率 85.47%)，相較於單獨只用 Harr-Like(準確率 78.37%)、HOG(準確率 80.64%)特徵，可得較好的行人偵測準確率。

#### 四、 結論

本論文提出一種新的行人偵測架構，利用立體視覺演算法，進行 U&V 視差統計產生候選區域後，結合 HOG 與 Geometry 做特徵空間轉換，並使用事前訓練好的 Ada-Boost 分類器進行分類判斷。本研究採用光軸平行的架構作為立體視覺的主要硬體架構，並利用 MIT 與 INRIA 作為樣本的資料庫測試 1800 張的影像，比較 HOG+Geometry、Harr-Like 和 Geometry-based 的實驗結果後，顯示 HOG+Geometry 具有較高的準確率，且處理速度每張影像僅需 70ms。

#### 五、 參考文獻

- [1] Bo Ling , Michael I. Zeifmana, David R.P. Gibson, “Multiple Pedestrians Detection Using IR LED Stereo Camera,” Proc. of SPIE, Vol. 6764, 67640A-1,2007.
- [2] Ignacio Parra Alonso, David Fernández Llorca, Miguel Ángel Sotelo, “Combination of Feature Extraction Methods for SVM Pedestrian Detection,” IEEE Trans. Intelligent Transportation Systems, Vol. 8, No. 2, pp. 292-307, 2007.
- [3] M. Bertozzi,A. Broggi,C. Caraffi,M. Del Rose,M. Felisa,G. Vezzoni,“Pedestrian detection by means of far-infrared stereo vision,” Computer Vision and Image Understanding, Vol. 106, pp. 194–204, 2007.
- [4] David Fern´andez Llorca, “Stereo Vision Based Pedestrian Detection System For Assisted Driving,” University Of Alcalá Escuela Polit ´EcniaA Superior Department of Electronics PhD THESIS (Summary), 2008.
- [5] Bastian Leibe , Aleš Leonardis , Bernt Schiele, “Robust Object Detection with Interleaved Categorization and Segmentation,” Int. J. Computer Vision, Vol. 77, pp. 259–289, 2008.
- [6] Philip Kelly, B.A. (Mod), “Pedestrian Detection and Tracking Using Stereo Vision Techniques,” Dublin City University School of Electronic Engineering Supervisor: Dr. Noel E. O’Connor, 2007.
- [7] Yingping Huang ,and Ken Young, “Binocular Image Sequence Analysis: Integration of Stereo Disparity and Optic Flow for Improved Obstacle Detection and Tracking,” EURASIP Journal on Advances in Signal Processing, Vol , pp. 1-10, 2008.
- [8] Hern´an Badino and Rudolf Mester “Stereo-based Free Space Computation in Complex Traffic Scenarios,” Proceeding of SSIAl , pp. 189-192, 2008.
- [9] Zhifeng Liu ,and Reinhard Klette ,“Dynamic Programming Stereo on Real-World Sequences,” The .enpeda.. Project, The University of Auckland, New Zealand.
- [10] Rapha,el Labayrade, Didier Aubert, Jean-Philippe Tarel, “Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through V-disparity Represent at

- ion,” Proceeding of SSIAI, pp. 646-651 , 2008
- [11] Philip Kelly, “Pedestrian Detection in Uncontrolled Environments using Stereo and Biometric Information,” Proceeding of VSSN’, pp. 161-169, 2008.
- [12] Hong Wang, Qiang Chen, Wenchao Cai, “Shape-based Pedestrian/Bicyclist Detection via Onboard Stereo Vision,” IMACS Multiconference on "Computational Engineering in Systems Applications,"(CESA) October 4-6, pp. 1776-1780 , 2006.
- [13] Rafael Muñoz-Salinas, Miguel Garcí’a-Silvente, Rafael Medina Carnicer, “Adaptive multi-modal stereo people tracking without background modeling,” J. Vision Communication Image Representation, Vol. 18, pp. 75–91, 2008.
- [14] Xia Liu and Kikuo Fujimura, ”Pedestrian Detection Using Stereo Night Vision,” IEEE Trans. On Vehicular Technology, Vol. 53, No. 6, pp. 1666-1678, 2004.
- [15] Dalal, N. Triggs, B. ,” Histograms of Oriented Gradients for Human Detection,” Computer Vision and Pattern Recognition, CVPR. IEEE Computer Society Conference Vol.1, pp. 886-893, 2005
- [16] 藤吉弘亘. "Gradient 特徴抽出 – SIFT 及 HOG - ," 情報処理學會 研究報告 CVIM 160, pp. 211-224, 2007.
- [17] T. Wada, F. Huang, and S. Lin , ” Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection,”(Eds.): PSIVT, LNCS 5414, pp. 37–47, 2009.