

公共空間限制區域非法進入者之偵測

Illegal Entrant Detection in Restricted Regions

石昭玲
中華大學
資工系
sjl@chu.edu.tw

溫敏淦
聯合大學
資管系
mgwen@nuu.edu.tw

韓欽銓*
聯合大學
資工系
cchan@nuu.edu.tw

袁凱群
中央大學
資工所

摘要

隨著取像設備的價格大幅的降低，監控系統目前已經廣泛的應用在日常生活當中。然而目前之監控系統大多僅提供錄影功能，只能提供事後資訊。因此便有人提出了智慧型監控系統的概念，利用電腦視覺的方法，在不需要人為的操作之下，讓監控系統能夠自動對攝影機所擷取的影像進行分析，以具有偵測、追蹤、辨識、分析的功能。因此本論文提出一個限制區域非法進入者偵測系統，利用制服色彩為特徵，用以判斷進入限制區域中的人員是否具有合法之身分。首先利用背景相減法來偵測是否有目標物的存在，並利用目標物的位置、大小與色彩等資訊來追蹤物體。接著使用一個以區塊為主的身體區域分割演算法，將目標物切割為頭部、上半身與下半身區域。最後針對穿著制服的身體區域抽取色彩的特徵並分類，以判斷該進入者是否具有合法之身分。而實驗結果顯示，本論文所提出之方法可以有效的辨別進入者之身分。

關鍵詞：智慧型監控系統、限制區域、非法進入者偵測、身體區域分割演算法、色彩特徵分類

Keywords: Intelligent surveillance system, restricted regions, illegal entrant, body segmentation, color feature classification

一、緒論

經過多年的發展，電腦視覺研究已經有許多成果，在偵測、追蹤和辨識方面，已經有長足的進步，這些研究有著廣泛的應用。例如在視訊監控、智慧型運輸系統、影像資料庫和虛擬實境等。而在智慧型視訊監控之相關研究主要集中於偵測、追蹤、辨識和分析四個部分，而且皆以人類為主要的偵測目標。

首先，有關於偵測的研究部份，Horprasert[1]

等人提出了一個目標物偵測演算法，其利用統計的方法訓練背景模組，將目前畫面中像素點與背景模組的差異分為色彩與亮度的失真，並選擇適當的門檻值判斷出像素點是屬於前景、陰影或背景。Cucchiara[2]等人提出了一個能夠快速偵測移動物與陰影的方法，其利用物體大小、物體像素點變化量、陰影偵測與物體移動量等資訊將物體分為移動物、非真實物體(ghost)、移動物陰影與非真實物體陰影四種類型，並依照物體的類型對背景做不同更新方式，以能夠更準確的找到合適於目前環境的背景資訊。Javed[3]等人提出了一個階層式的背景相減法，此方法結合了色彩與梯度值的偵測方法，主要使用於戶外的環境，在光線突然變化的情形下，此方法仍然能夠將正確的偵測前景物。不過由於此方法須要同時計算色彩與梯度值的資訊，需花費大量的計算時間，因此Stefano[4]等人則對此方法提出改進，將梯度值改為以區塊為主的計算方式，以增進系統的效能。

在追蹤的研究部分，部分的學者[5][6]提出了以區塊為主的追蹤方式，其主要的原理在於將目標物內分為數個色彩相似的區塊，並統計區塊內的色彩資訊，在追蹤時則利用區塊的色彩資訊來將兩張畫面中的區塊對應，這樣做法的好處在於當目標物有合併的情形時，仍然能夠正確的目標物分割。Brémond [7]等人提出了一個可同時追蹤多個物體的方法，其主要是利用物體先前的移動量預測物體在下一個畫面中的位置與大小，並將目前偵測到的物體與預測的物體比較，以找出物體之間的對應關係。Chen[8]等人提出了使用 bipartite graph 來表示物體間對應的方法，在 bipartite graph 中的每一個節點 (node) 即為一個物體，此方法當物體間有交錯、合併與分裂的情形時，仍然能夠正確的找到物體間的對應關係。

在其它方面的研究，如分析與辨識部分，Haritaoglu[9]等人提出一套戶外的即時視訊監控系統 W4，此系統使用灰階的影像或紅外線裝置。W4 能夠力用輪廓的資訊來定位人體的細部模組 (頭、手、腳與軀體)，分析人們的動作和行為。Beynon [10]等人提出了一個結合多架攝影機偵測遺留物的方法，其使用多架攝影機來增加可視的範圍，此系統利用狀態機 (state machine) 來表示遺

* To whom correspondence should be addressed.

留物狀態的改變，並能夠判斷遺留物擁有者與遺留物的關係。Park[11]等人提出了一個人類身體區域切割與追蹤的演算法，其將人體分割為數個區塊，並使用一身體模組找出區塊與各個身體區域的關係，此系統可以用來辨識一些人類之互動情形，如兩個人的握手、擁抱與推擠等動作。Zhao[12]等人利用一對可移動式取像設備，取得影像中目標物的距離資訊，結合立體視覺切割（stereo-based segmentation）與類神經網路辨識，偵測都市街道行人避免發生意外。

在公共空間中如超商櫃檯、飯店櫃檯、銀行櫃檯與醫院護理站等環境，其會含有金錢、貴重物品或是管制性物品，而由於在公共空間中對於進出人員並無特殊的身分限制，再加上進出人員眾多使得這些區域特別容易受到入侵者的侵入。因此本論文提出一套智慧型的監控系統，結合數位影像處理技術和視訊認知技術，對限制區域內人員進行身分的辨別與追蹤。在這些限制區域中之工作人員都具有一個共同的特徵，就是必需要穿著固定的制服，因此本論文便利用制服色彩為特徵，用以判斷進入限制區域中的人員是否具有合法之身分。若進入人員不具有合法之身分，系統便會在第一時間發出警訊，以讓相關人員能夠立即的做出反應。整個系統大略可分為偵測、追蹤、特徵抽取與辨識四個主要的階段，其流程如圖 1 所示。

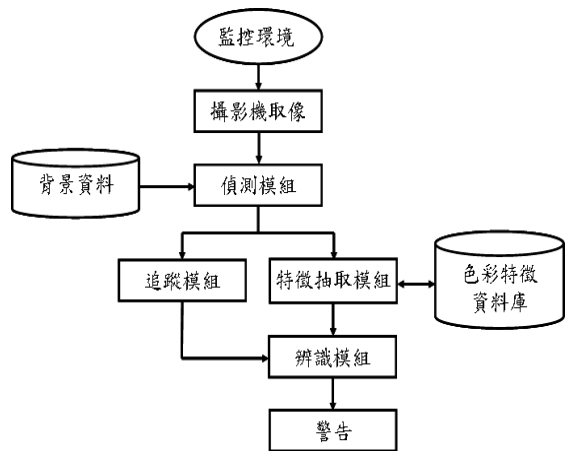


圖 1 系統流程圖

首先利用一個架設於固定位置之攝影機來取得影像，並交由偵測模組處理，偵測模組主要的功能在於將畫面中的前景物與背景分離。在本論文中使用背景相減法，利用事先統計的背景資料與畫面比較，判斷是否存在移動的前景物，若發現有目標物存在時則將目標物交由追蹤模組與特徵抽取模組處理。追蹤模組為偵測模組的後續處理，當偵測到目標物後，必須利用追蹤模組對目標物持續的監控觀察，以分析目標物的行為，並由這些不同的行為，觸發系統中各項不同的事件，再由系統對這些事件做出立即的反應。

特徵抽取模組通常需要取得目標物的各項特徵，在本論文中，特徵抽取模組會利用事先統計的色彩特徵資料庫，針對感興趣的區域做色彩特徵的抽取，並將這些色彩特徵交由辨識模組處理。在辨識模組中利用一個由神經網路訓練得到之分類器，將特徵抽取模組提供之特徵分類，以判斷該目標物是否具有合法之身分，若該目標物不具有合法之身分，系統則會對此非法事件發出警告。

二、目標物偵測與追蹤

目標物的偵測與追蹤是視訊監控系統最重要的部分，其運作的好壞決定整個系統的成敗。傳統上常用以偵測移動目標物之方法大致可以分為三種：背景相減法、連續影像相減法、光流偵測法。

(一) 目標物與陰影偵測

本論文所提出的目標物偵測演算法，其主要原理是利用統計的方法訓練背景模組，並選擇適當的門檻值判斷出像素點是屬於前景、陰影或背景。考慮 RGB 色彩空間中，如圖 2 所示。其步驟分別如下：

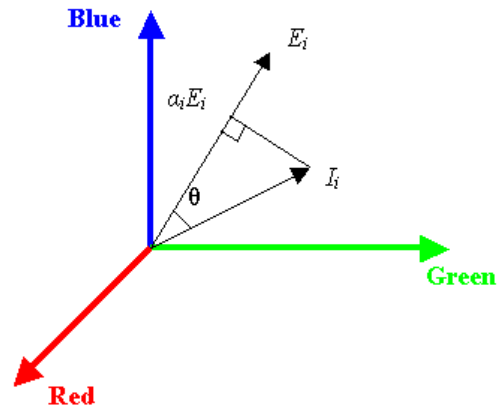


圖 2 RGB 色彩空間中的背景模組

首先定義畫面中某一像素點 i ，讓向量 $E_i = [E_R(i), E_G(i), E_B(i)]$ 表示背景模組中 RGB 色彩的期望值，而 $I_i = [I_R(i), I_G(i), I_B(i)]$ 代表畫面中像素點 i 所量測到的 RGB 色彩值。然後可以從 I_i 中分解包含亮度 (brightness) 失真和色彩 (chromaticity) 失真兩個失真元素，其定義如下：亮度失真 (α)，是一個縮放純量值，讓測量的色彩盡量靠近期望色調線 (expected chromaticity line) 上 (E_i 向量)，其計算公式表示如下方程式：

$$\alpha_i = \min \left\langle \left(\vec{I}_i - \alpha_i \vec{E}_i \right)^2 \right\rangle \quad (1)$$

換句話說，而 $\alpha_i E_i$ 表示向量 I_i 在 E_i 上的投影點。而色彩失真 (CD) 則被定義為測量顏色 I_i 到期望色調線 E_i 之間的正交距離，如下方程式所示：

$$CD_i = \left\| \vec{I}_i - \alpha_i \vec{E}_i \right\| \quad (2)$$

由於較暗的像素點顏色於 RGB 色彩空間中會很接近各種顏色，因此顏色點會被認定為靠近或接近所有的色調線，使得當像素點較暗時，其色彩失真也會相對的較小。[10]則針對此缺點提出了一個改良的色彩失真計算方式，即是以測量顏色 I_i 與期望色調線 E_i 的夾角角度作為其色彩失真，色彩失真可以改寫為如下所示：

$$CD_i = \cos(\theta) = \frac{I_i E_i}{|I_i| |E_i|} \quad (3)$$

其次，利用背景相減演算法，先統計一段靜態背景畫面資料，建立背景模組，每一個畫面中，相同位置的像素點持續累積其統計量，每一個像素點 i 皆包含四種元素組成 $\langle E_i, s_i, a_i, b_i \rangle$ ，其中 E_i 表示顏色的期望值， s_i 表示顏色值的標準差， a_i 表示亮度失真 α_i 的變異數， b_i 表示彩色失真 CD_i 的變異數，分別定義如下：

$$E_i = [\mu_R(i), \mu_G(i), \mu_B(i)] \quad (4)$$

$$s_i = [\sigma_R(i), \sigma_G(i), \sigma_B(i)] \quad (5)$$

但是仍然需要考慮攝影機因光線、雜訊所受到的干擾，為了正規化和平衡 RGB channel，方程式中 α_i 的計算修正成式(6)：

$$\alpha_i = \min \left[\sum_{c=R,G,B} \left(\frac{I_c(i) - \alpha_i \mu_c(i)}{\sigma_c(i)} \right)^2 \right] \quad (6)$$

由於在建立背景模組時，統計 N 張畫面後，發現每張畫面像素的亮度失真和顏色失真皆不相同，所以分別計算其均方根 (Root Mean Square, RMS) 來代表其變異數，所以亮度失真和彩色失真的變異數分別定義如下：

$$\alpha_i = RMS(\alpha_i) = \sqrt{\frac{\sum_{i=0}^N (\alpha_i - 1)^2}{N}} \quad (7)$$

$$b_i = RMS(CD_i) = \sqrt{\frac{\sum_{i=0}^N (1 - CD_i)^2}{N}} \quad (8)$$

接下來將像素點分類，利用背景模組把每一個像素點分類成以下四種類型：

- (1) 背景 (Background, B)：像素點的亮度和顏色近似於背景模組。
- (2) 陰影背景 (Shadow background, S)：像素點的顏色近似於背景模組，但是亮度偏暗。
- (3) 反光背景 (Highlight background, H)：像素點的顏色近似於背景模組，但是亮度偏亮。
- (4) 前景 (Foreground, F)：像素點的亮度和顏色皆異於背景模組。

像素點分類的依據，主要是靠每一個像素點 i ，其本身的亮度失真 (α_i) 和色彩失真 (CD_i) 兩個參數。為了選擇單一的門檻值做為分類的依據，重新正規化亮度失真和色彩失真兩個參數，如下方程式所示：

$$\hat{\alpha}_i = \frac{\alpha_i - 1}{\alpha_i} \quad (9)$$

$$\hat{CD}_i = \frac{1 - CD_i}{b_i} \quad (10)$$

其 $\hat{\alpha}_i$ 中表示正規化後的亮度失真， \hat{CD}_i 表示正規化後的色彩失真，計算出上述兩個參數結果後，並設計判斷規則式(11)，藉此判斷所有像素點的類別。

$$\begin{cases} F: \hat{CD}_i > \tau_{CD}, \text{ else} \\ B: \hat{\alpha}_i \geq \tau_{a1} \text{ and } \hat{\alpha}_i \leq \tau_{a2}, \text{ else} \\ S: \hat{\alpha}_i < \tau_{a2}, \text{ else} \\ H: \text{ otherwise} \end{cases} \quad (11)$$

在上述規則中，門檻值扮演著重要角色，因此利用統計的方式來設定門檻值，以增加判斷規則的強健性，根據門檻值可以根據的預期偵測比率 (detection rate, r) 自動選定，正規化後的 α_i 直方統圖中，兩個門檻值 τ_{a1} (3.0) 和 τ_{a2} (-3.0) 定義了亮度的範圍，從 τ_{a1} 為 $\hat{\alpha}_i$ 的值於偵測比率 r ($r = 0.9999$) 到 τ_{a2} 為 $\hat{\alpha}_i$ 的值於偵測比率 $(1-r)$ 。正規化後的彩色失真 (\hat{CD}_i) 直方統計圖，門檻值 τ_{CD} (3.0) 為彩色失真 \hat{CD}_i 的值於偵測比率 r 。

當使用背景相減法得到目標物的區塊後，接著利用形態學運算 (morphology operation) 來去除畫面中的雜訊。再來使用連通元件 (connected-component) 將二元化影像中相連結的區塊標示成相同的標記。如此可以得到畫面的每一個前景物的

詳細的像素點分布資訊。最後考慮目標物的合理限制條件，如像素點數目，目標物之長寬，並經由條件式判斷的方式將不合理之目標物刪除。

(二) 目標物追蹤

本論文利用[7]所提出的追蹤演算法，主要是利用目標物先前的移動向量 (motion vector) 預測出目標物的最小邊界矩形 (boundary box) 在下一個畫面中的位置與大小，並將目前畫面中偵測到的目標物與預測的目標物比較，以找出目標物之間的對應關係。當目標物間的位置過於相近時，只使用位置與大小做為對應的特徵容易產生錯誤的對應，因此在該演算法中加入色彩特徵，在發生對應混淆的情形下仍能夠正確的找到目標物間的對應。

(A) 目標物對應：

此方法是利用目標物最小邊界矩形的的位置與大小，求得前後畫面中目標物間的關係。首先求出目標物最小邊界矩形的左上角點與右下角點，並利用此兩個點在前面數個畫面的移動向量，來預在下一個畫面中此兩個點的位置，該位置即為下一個畫面中目標物可能的位置，而該位置所形成的最小邊界矩形則可以視為一個預測的目標物。

在預測目標物部分，令 PMV_{t-1} 表示第 $t-1$ 個畫面中預測的目標物移動向量， MV_{t-1} 表示第 $t-1$ 個畫面中目標物真實的移動向量，則第 t 個畫面中預測的目標物移動向量 PMV_t 則可經由下列公式計算求得：

$$PMV_t = \alpha MV_{t-1} + (1 - \alpha) PMV_{t-1} \quad (12)$$

目標物與預測目標物間的對應關係可分為下面 4 種情形：

- (1) 一個目標物對應到一個預測目標物：表示此兩個目標物即為正確的對應。
- (2) 一個預測目標物沒有對應到任何的目標物：表示目標物從目前畫面起離開此環境。
- (3) 一個目標物沒有對應到任何的預測目標物：表示目標物從目前畫面起進入此環境。
- (4) 數個目標物對應到數個預測目標物：表示目標物間可能合併、分裂或位置過近，而產生對應混淆的情形。

當發生對應混淆時則可將目標物與預測目標物間的關係再細分為下面 3 種情形。

- (1) 一個目標物對應到數個預測目標物：表示前一個畫面中的數個目標物在目前畫面中合併為一個目標物。當發生合併的情形時，將被合併目標物的色彩特徵儲存，使得目標物分裂時能夠對應到未合併前的相同目標物。
- (2) 一個預測目標物對應到數個目標物：表示前一

個畫面中的一個目標物在目前畫面中分裂成數個目標物。當發生分裂的情形時，利用先前合併時所儲存的色彩特徵來產生目標物間的對應。

- (3) n 個預測目標物對應到 n 個目標物 ($n \geq 2$)：表示目標物間的位置與大小皆相近而產生了混淆的對應，此時利用色彩的特徵來產生目標物間的對應。

經由前面所描述的目標物對應的關係，便可利用目標物的位置、大小與色彩等特徵來定訂一些規則，以找出正確的目標物對應關係。

(B) 目標物色彩特徵表示：

在本論文中，目標物之色彩特徵使用 MPEG-7 的 CSD[13] (Color Structure Descriptor)，CSD 類似於傳統的色彩直方圖 (color histogram)，可以表示一張影像中色彩的分布，並改良了色彩直方圖沒有空間關係的缺點。如圖 3 所示，傳統的色彩直方圖會將此兩張影像視為是完全相同，而 CSD 則可以正確的表示出此兩張影像在色彩結構上的不同。

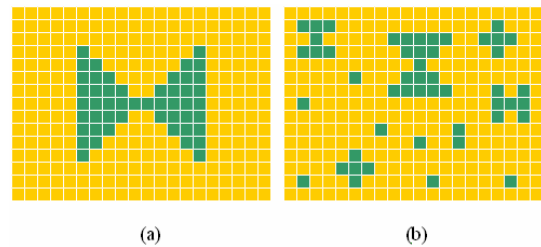


圖 3 兩個不同色彩結構之影像

將目標物影像的色彩資訊從 RGB 色彩空間轉換到 HMMD (Hue-Max-Min-Diff) 色彩空間，在轉換到 HMMD 色彩空間後，將目標物影像內的色彩量化到 64 個 bins，依照 Diff 的大小可分為五個子空間，分別為： $[0, 6)$ ， $[6, 20)$ ， $[20, 60)$ ， $[60, 110)$ 與 $[110, 255]$ ，其中每個子空間中可以依照 Hue 與 Sum 的大小分割為數個區間，分割的區間數目如下： $(1,8)$ ， $(4,4)$ ， $(4,4)$ ， $(8,2)$ ， $(8,1)$ 。

由於所偵測的目標物影像通常小於 256×256 個像素點，因此在本研究中將結構元素 (structuring element) 大小固定設定為 8×8 。在色彩統計的步驟中，利用結構元素由左上往右下水平掃瞄，若色彩在結構元素中出現，則該色彩對應的 bin 則加 1，但結構元素每移動一次每個 bin 至多只能增加 1，CSD 便是利用這樣的統計方式來表示色彩在空間上分布的特性。如圖 4 所示，色彩 C1、C3 與 C7 在同一個結構元素中出現多次，但其對應的 bin 只會增加 1。

在特徵比對的部份，CSD 是利用 L1 distance 來計算兩張影像的色彩相似度。若兩張影像分別為 A 與 B，color structure 分別為 h_A 與 h_B ，相似度公式為：

$$\text{dist}(A, B) = \sum_i |h_A(i) - h_B(i)| \quad (13)$$

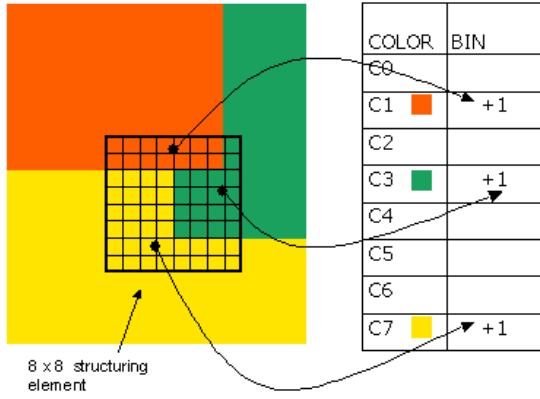


圖 4 CSD 的色彩統計方式

三、頭與身體區域偵測

在偵測出目標物後便可以針對目標物進行身分合法性的辨識。由於在本論文中是以限制區域的進入者是否有穿著特定的制服作為身分辨別的依據，因此針對進入者做色彩特徵的抽取，以進而判斷該進入者是否具有合法之身分。在部分的環境當中工作人員所穿著的制服可能只有上衣部分(如超商或大型賣場等環境)，此時若同時抽取上衣與褲子的色彩特徵，並以此特徵來辨識該人員的身分，容易受到褲子色彩的影響而發生辨識錯誤。因此為了要使所提出的方法能夠更廣泛的應用在各種不同的環境當中，並且增加辨識的準確度，在本論文中將針對所偵測到的目標物做進一步的分析，以找出進入者上半身區域與下半身區域。當制服只有上衣時，對偵測出的上半身區域進行特徵的抽取，當制服同時包含上衣與褲子時，則分別對偵測出的上半身區域與下半身區域進行顏色特徵的抽取。

(一) 區塊產生

在[11]所提出的演算法中是利用最大期望值演算法 (Expectation-Maximization Algorithm)，估測出目前目標物中色彩分布之混合高斯模型參數，利用此參數將目標物影像中的色彩分類，並將相同類別中不具有連接性之像素點分割成一個新的類別，如此一來，每一個類別便可視為是目標物上的一個區塊。不過由於在求得參數的過程中須要大量的計算時間，無法滿足即時系統的需求，因此在本論文中使用分水嶺 (watershed) 分割法來產生區塊，以減少產生區塊所需的時間。利用分水嶺分割法產生區塊的執行步驟如下: (1) 影像簡化, (2) 梯度計算, (3) 泛流處理, (4) 區塊合併。

步驟一：影像簡單化：

當對目標物影像分割時，為了避免雜訊的干擾(如頭髮或衣服紋理)與易於在空間上分割，因此必須對影像做簡單化的處理。形態學濾波器 (morphological filter) 是常用簡單化濾波器，在使用簡單化濾波器前先做以下的定義。

假設 $f(x)$ 為輸入信號， M_n 為結構元素大小為 n 之視窗，經由 M_n 所得到的侵蝕 (erosion) 與膨脹 (dilation) 分別為：

$$\varepsilon_n(f(x)) = \min\{f(x+y) : y \in M_n\} \quad (14)$$

$$\delta_n(f(x)) = \max\{f(x-y) : y \in M_n\} \quad (15)$$

其中侵蝕為結構元素內最小值，膨脹則為結構元素內的最大值。斷開 (opening) 定義為先侵蝕後膨脹，閉合 (closing) 定義為先膨脹後侵蝕，其中斷開可以消除影像中較亮的雜點，閉合則可以消除影像中較暗的雜點，其方程式定義如下：

$$\gamma_n(f(x)) = \delta_n(\varepsilon_n(f(x))) \quad (16)$$

$$\varphi_n(f(x)) = \varepsilon_n(\delta_n(f(x))) \quad (17)$$

在此利用部份重建的閉合 (closing by partial reconstruction) [14]來簡化目標物影像並保留物件的輪廓，定義如下：

$$\varphi^{res} = (\delta_n(f), \varphi_n(f)) \quad (18)$$

其主要的做法是將經過一次膨脹運算後的影像不斷的侵蝕，每做一次侵蝕便與經過閉合運算的影像比較取最大值，直到收斂 (沒有任何變化) 為止，最後所得到的影像即為一張簡單化的影像，如圖 5 所示。

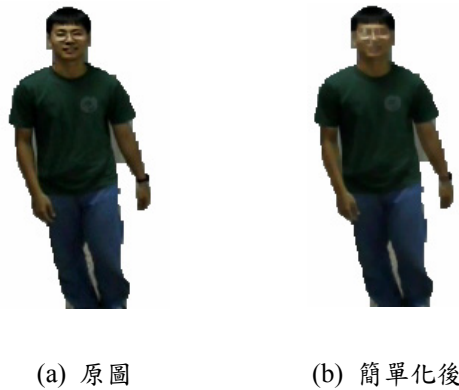


圖 5 簡單化處理後的影像

步驟二：梯度計算：

在本論文中利用形態學梯度 (morphological gradient) 來計算影像中的梯度值，形態學梯度是將影像膨脹的結果減掉影像侵蝕的結果得到：

$$G(f) = \delta(f) - \varepsilon(f) \quad (19)$$

但是有些區域的強度很小，以至於梯度值也很小，照成在使用分水嶺分割法時產生錯誤的邊界。在此參考[15]的作法將色彩的資訊增加權重，對於 G_L 、 G_a 和 G_b 分別取權重值 $w_L=1$ 、 $w_a=2$ 和 $w_b=2$ ，影像的梯度值為加權後的最大值，表示如下：

$$g = \max(w_L \cdot G_L, w_a \cdot G_a, w_b \cdot G_b) \quad (20)$$

其中 G_L 、 G_a 和 G_b 分別為 CIELab 中的 L、a 與 b。因為 CIELab 為 uniform 的色彩空間，即在 CIELab 上所測量的色彩差異與人眼的的視覺較為類似，因此在區塊的分割與合併時會有較好的結果。CIELab 色彩空間之色彩距離測量公式為：

$$\Delta E_{ab} = (\Delta L^2 + \Delta a^2 + \Delta b^2)^{1/2} \quad (21)$$

在計算出目標物影像的梯度值後，為了避免因為過度分割的問題，因此使用預先標記的方式，將梯度過小的點先標記區塊的編號。這樣的前處理做法有兩個好處，一則是解決因梯度值微小變化所產生的過度分割，二則是如此處理後標記過的點就不用再做泛流處理，縮短了分水嶺分割法所需的處理時間。在本實驗中，是選擇目標物影像平均梯度值的 0.6 倍來做為預先標記之門檻值。

步驟三：泛流處理：

分水嶺分割法的泛流處理便是將影像中不同大小的梯度資訊當做高低不平的地勢，利用水流泛濫淹沒地勢低位的原理來進行影像分割，因此必須先依影像中每一點的梯度資訊來決定它之後的流向。在實做上以一個 3x3 的遮罩依序檢示每個點周圍的八個位置，並將點的流向指向周圍八個位置中梯度最小的那個點，如圖 6 所示。

0	1	5	7	9
1	2	7	10	8
10	15	10	9	5
12	13	9	8	4
14	15	7	5	3

(a) 梯度影像

			←	←
		↖	↖	↓
↑	↖	↖	→	
↑	→	↘	↘	
↑	→	→		

(b) 流向資訊

圖 6 依梯度值計算流量

接著依序對影像中尚位標記區塊邊號的點做泛流處理，從每個未被標記的點開始，依循著之前所決定的流量資訊進行泛流，在泛流處理的過程中每個流經的點都視為屬於同一區塊，直到最後流向終點，這個未被標記的點才算結束。最後流向終點的可能性有兩種，一種是梯度極小值的點，另一種是已經有區塊編號的點，如果是前者的話，可以將這整個水流所經過的所有點視為是一個全新的區塊，並給定一個編號，但若是後者，終點為一個已經有區塊編號點，那麼這個水流被認為只是終點所屬區塊的一個分支，因此將這支水流上的所有點都給予和終點相同的區塊編號。

步驟四：區塊合併

由於受到雜訊或影像中紋理的影響，在使用分水嶺影像分割影像後往往會形成許多小區塊，這些小區塊可能是由一個有意義的大區塊所分割出來，因此必需對這些屬於相同大區塊的小區塊做合併的處理。在此利用[11]所提出的區塊合併的方法，將影像中的區塊依照色彩、大小和連接性將區塊做適度的合併，以得到較好的分割結果。區塊 (A_i 與 A_j 分別表示第 i 個與第 j 個區塊) 的屬性與關係描述如下：

1. 單一區塊的屬性

- A_i 的大小： A_i 內像素點的個數，以 $|A_i|$ 表示。
- A_i 的中心點： A_i 內像素點 L、a、b 的平均值，以 $\Phi_i = [\mu_L, \mu_a, \mu_b]$ 表示。
- A_i 邊界點集合： A_i 邊界上的點所形成的集合，以 $\Psi(A_i)$ 表示。

2. 區塊間的關係

- A_i 對 A_j 的連接性： $(\Psi(A_i)$ 中與 A_j 有相連的點數) / ($\Psi(A_i)$ 中所有的點數)，以 $\beta_i(A_j)$ 表示。

在訂定區塊的屬性與關係後便可以藉由區塊的屬性與關係訂定三個區塊合併的規則，將小的區塊合併為較有意義的大區塊，以解決過度分割的問題。規則分別描述如下：

1. **相連性規則：**若 A_i 與 A_j 的相連性大於一門檻值 T_1 (在本實驗中 T_1 為 0.7)，則可能是小的區塊連接著大的區塊，或是小區塊被大區塊包覆，因此將這兩個區塊合併，表示如下：

$$(\beta_i(A_j) > T_1) \vee (\beta_j(A_i) > T_1) \quad (22)$$

2. **色彩相似度規則：**若兩個相連的區塊 A_i 與 A_j 的色彩相似度小於一門檻值 T_2 (在本實驗中 T_2 為 15)，則表示此兩個區塊具有相似的颜色，

因此將這兩個區塊合併，條件表示如下：

$$\|\Phi_i - \Phi_j\| < T_2 \quad (23)$$

3. **小區塊規則**：若 A_i 的大小 $|A_i|$ 小於一門檻值 T_3 (在本實驗中 T_3 為 50)，則將此區塊合併到與其連接且色彩相似度最高的區塊，條件表示如下：

$$|A_i| < T_3 \quad (24)$$

在合併後目標物內的區塊便只剩下數個較大的區塊，因此可建立一個身體模組將區塊對應到身體的部位。

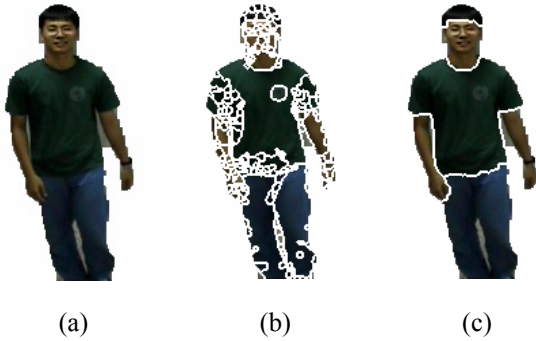


圖 7 區塊產生(a)原圖(b)分水嶺分割(c)區塊合併

(二) 身體模組

在使用分水嶺分割法切割影像中的區塊後，為了要避免過度分割的問題，必須利用區塊的合併盡量將小區塊合併成較有意義的大區塊。不過由於受到衣服中色彩的影響，在區塊合併後仍然無法保證屬於上半身的所有區塊會被合併到一個大的區塊，上半身區域可能被分割成數個較小的區塊，使得難以直接的判斷上半身區域所在的位置。相較於身體的其它區域，頭部區域具有較明顯且穩定的分割特徵[16]，因此利用頭部區域類似橢圓形的特徵，先偵測頭部區域，並利用頭部區域與上半身區域相連且具有固定比例的特性，利用頭部區域的資訊來幫助判斷有那些區塊應屬於上半身區域。

在完成區塊合併的步驟後，頭部區塊通常包含著數個較小區塊，如頭髮與臉部等區塊，由於這些區塊的色彩差異較大且區塊間不具有包覆的特性，因此在區塊合併的過程中並不會被合併為一個較大的區塊，因此必須先將這些區塊合併後才能得到一個完整的頭部區塊。在此利用人類頭部類似橢圓形的特性，利用圓形的測量公式 (circularity measure) 判斷那些區塊在合併後會最類似圓形，該合併後的區塊即為頭部區塊。區塊的合併方法則

是利用人的頭部是在身體最上方的特性，將區塊中心點依照 y 軸座標由大至小 (由高至低) 排列，並依序將區塊合併。

頭部區塊偵測的演算法描述如下：

步驟1：將目標物內區塊中心點的位置由高往低排列。令 R_0 為空集合， N 為目標物內像素點數目， $k=1$ 。

步驟2：令 A 為未被選取過且位置最高的區塊， $R_k = R_{k-1} \cup A$ 。

步驟3：若 $|R_k| < \frac{N}{3}$ ，則結束此演算法。

步驟4：利用圓形的測量公式計算出 R_k 的圓形相似度 C_k ，並跳至步驟 2。

為了減少演算法的計算量並增加判斷的正確性，在此加入一個頭部區域大小的限制條件：人類頭部區域的大小應在身體區域大小的 1/3 以內，所以當合併區塊的像素點數目大於目標物像素點數目的 1/3 時，則終止此演算法。在一般的情況下，最大值所對應的合併區塊即為頭部區塊，不過當所得到的目標物影像是進入者的側面或背面影像時，容易受到頭髮的影響，使得最大值所對應的合併區塊可能只包含了頭髮區塊，而無法正確的找出完整的頭部區塊。在此加入膚色的條件以解決此問題，判斷那些合併區塊包含了膚色的區塊，並優先從這些包含膚色區塊的合併區塊中，選擇值最大的區塊作為頭部區塊。當所有的合併區塊皆不包含膚色的區塊時，則選擇最大值所對應的合併區塊為頭部區塊。在膚色的區塊的判斷方面，使用統計與型態學的運算方法，找出膚色區域在 CIELab 色彩空間中的分布範圍，當區塊的中心點落在此分布範圍時，將此區塊視為膚色之區塊。在第四節中使用相同的方法來統計制服色彩的分布範圍，並對此方法有詳細的描述。

在[17]中提出了兩個圓形的測量公式 $C1$ 與 $C2$ ，公式分別描述如下：

$$C1 = \frac{|P|^2}{K}, \quad C2 = \frac{\mu_R}{\sigma_R} \quad (25)$$

$C1$ 是計算區塊的緊密 (compactness) 程度，當區塊越相似圓形時其緊密度會越高，但是 $C1$ 會受到區塊邊界平滑程度的影響而產生錯誤的判斷。 $C2$ 是計算區塊邊界點到中心點距離的平均值與區塊邊界點到中心點距離的標準差，並以此兩個值的比值來做圓形的測量。在本論文中是以 $C2$ 來作為圓形的測量公式。

在圖 8 中，(a)與(b)的右圖為利用區塊產生後的結果。(c)為(a)中區塊合併後之 C2 值，可以看出區塊 A₀與區塊 A₁合併後的 C2 值最高，表示此兩區塊合併後即為頭部區域。(d)為(b)中區塊合併後之 C2 值，其中區塊 A₀雖然 C2 值最高，但由於會優先選擇膚色的區塊，因此選擇 A₀與 A₁合併後的區塊為頭部區域。知道頭部區域後便可以利用[17]所提出計算橢圓長軸的公式計算出頭部的長度。

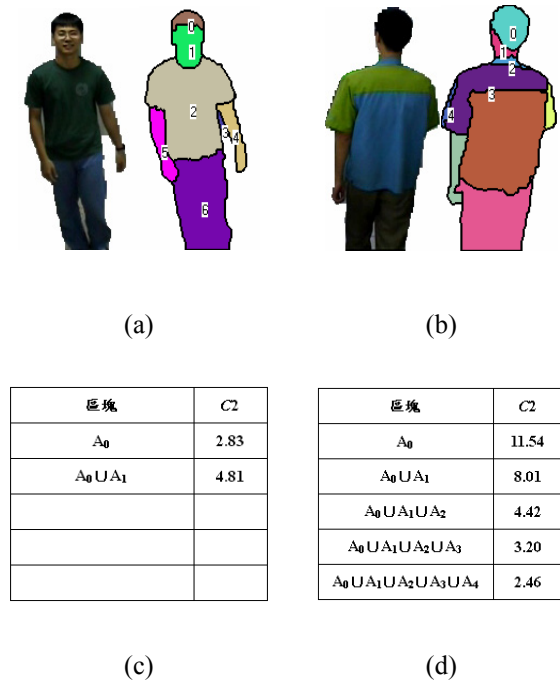


圖 8 利用區塊圓形測量公式偵測頭部

因為頭部與上半身區域有一定的比例關係，因此可以利用頭部的長度估計出上半身區域的長度。在此訂定頭部與上半身區域的比例為 1 比 2。在知道上半身區域的長度後可將中心點高度落在頭部最低邊界 (lowest boundary) 與上半身區域最低邊界間的區塊視為是上半身區塊的一部分，並將這些區塊合併，如此一來便可以獲得完整的上半身區塊，而下半身部位則是將目標物內剩下的區塊 (非頭部區塊與上半身區塊) 合併求得。當目標物的衣服與褲子的色彩相似時，上半身區域與下半身區域會被包含在同一區塊中，此時則可利用上半身區域的最低邊界來評估，是否此區塊有一定數量的像素點落在其它的區域，若有此種情形發生，則利用上半身區域的最低邊界將此區塊切割為兩個區塊，這兩個被分割之區塊即分別屬於目標物的上半身區域與下半身區域。

四、色彩特徵抽取與分類

將目標物的身體區域分割為上半身與下半身

區域後，便可針對穿著制服的身體區域來做色彩特徵的抽取與分類，以藉此辨別目標物的身分。在本論文中，提出了一個自動統計制服色彩的方法，在訓練階段中輸入適當的樣本資料，系統便能自動的找出制服中有幾種主要的色彩，並能夠自動的判斷每種主要色彩在 CIELab 色彩空間上的分布範圍。在特徵抽取的階段，利用制服色彩分布範圍來找出制服色彩在身體區域中的分布。在特徵的訓練階段，利用類神經網路中的倒傳遞演算法訓練一分類器，以將特徵分類，並以此分類結果來判斷目標物之身分。在此先對於感興趣的區域作定義 (Region Of Interesting, ROI)，在本論文中 ROI 即是制服出現的區域。若是限制區域中人員所穿著的制服只有上衣，則 ROI 為上半身區域。若人員所穿著的制服包含了上衣與褲子，則 ROI 為上半身與下半身區域。

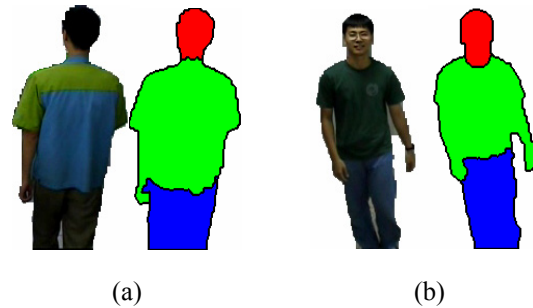


圖 9 身體區塊偵測

(一) 制服色彩統計

為了要讓系統能夠自動的統計制服色彩在 CIELab 色彩空間中的分布範圍，必需在訓練的階段中，輸入數個穿著合法制服人員之視訊資料。

首先利用前一章節中所提出身體區域分割演算法，將目標物的身體區域切割為上半身與下半身區域，由於在此所輸入的視訊資料皆為合法進入人員，因此在 ROI 中的區塊即屬於制服區塊的一部分，ROI 中區塊的色彩即為制服的色彩，因此便可以藉由這些區塊來瞭解制服色彩的分布範圍。此外加入一個區塊大小的條件，若 ROI 中的區塊與 ROI 的大小比例小必須大於一門檻值 T_{size} (在本實驗中設定為 ROI 大小的 5%)，因為過小的區塊大多位於區塊間的交界處，屬於雜訊的部份，因此將這些區塊去除，避免影響分類的結果。如圖 10(a)所示，在超商的例子中，ROI 即為上半身區域。

由於無法事先得知制服中色彩的數目，為了簡化分類的過程以產生較好的分類結果，在此只針對區塊的平均色彩分類，即每個區塊只用一個色彩為代表，並利用 maximin-distance 演算法將這些區塊的平均色彩分類。圖 10(b)為超商制服色彩在 CIELab 色彩空間中分布的情形，左圖為區塊平均值的分布，右圖為區塊平均值分類的結果。

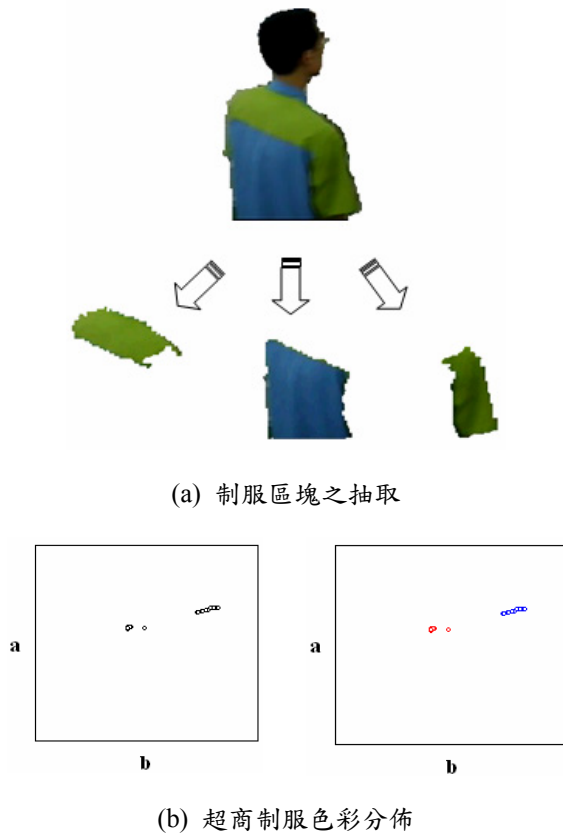


圖 10 超商制服色彩分類

在完成上述步驟後可以得到數個不同的類別，每個類別皆為制服中的一種色彩，不過這些分類結果只代表了區塊平均值的分布範圍，因此將區塊中的所有像素點加入到區塊平均值所屬的類別中，利用統計與型態學的運算找出每個類別主要的分布範圍，並利用一個最小凸多邊形（Convex Hull）將該類別主要的分布範圍包圍起來。演算法描述如下：

- 步驟1： 利用統計的方式去除雜訊點，先統計每個類別的平均值與標準差，並使用標準差的倍數作為一門檻值，將離平均值過遠的點去除。
- 步驟2： 在使用統計的方法去除雜訊點後有部分的孤立點仍然會存在，因此使用型態學的侵蝕和膨脹去除這些孤立點。
- 步驟3： 利用最小凸多邊形將剩下的點給包圍起來，並用表格紀錄多邊型內的範圍以便查詢。

圖 11(a),11(b) 為超商制服色彩的統計結果。圖 11(c)為利用統計的超商制服範圍來偵測超商的實驗結果，其中紅色的區域即為偵測出屬於制服的區域。

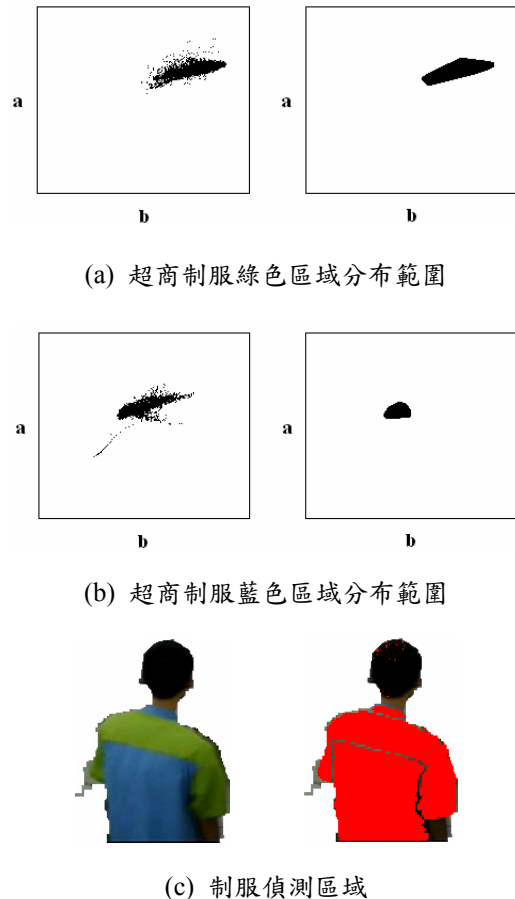


圖 11 利用統計的色彩分布範圍來偵測制服區域

(二) 色彩特徵抽取

由於已經知道制服色彩的分布範圍，因此可以找出在 ROI 中那些像素點的色彩落在此範圍之中，這些像素點在 ROI 中的分布，即是要抽取之色彩特徵。為了要表示制服色彩在 ROI 中的分布，並增加制服色彩在空間上的特性，在此利用與 color structure 類似的方法，使用一個 8x8 的結構元素由左上往右下水平掃瞄，不過由於只須計算制服色彩的分布特徵，因此每個 bin 所對應的色彩即為先前所統計出的制服色彩。

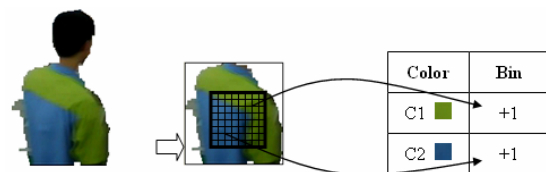


圖 12 超商制服色彩特徵

(三) 色彩特徵分類

當抽取出制服色彩的特徵後，利用類神經網路來做特徵的辨識，所使用的是倒傳遞演算法

(backpropagation networks, 或稱為多層感知機)。隱藏層的層數與每一層類神經元數目的選擇上至今仍沒有一個經過嚴謹分析後而得到的解答, 通常都是以嘗試錯誤法來尋找最佳結構, 但是理論上感之機只需要兩層架構, 再加上隱藏層神經元夠多的話, 此感之機的輸出便可以逼進任一連續函數 [18]。

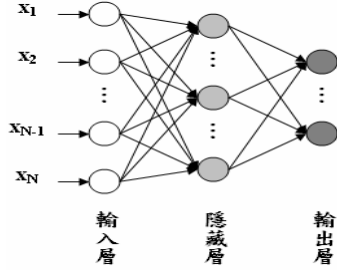


圖 13 一層隱藏層的倒傳遞神經網路架構圖

倒傳遞演算法網路訓練方式包含兩個階段: 前饋處理階段 (forward phase) 以及倒傳遞階段 (backward learning phase)。在前饋處理階段為當網路有輸入值時, 會將節點的值與隱藏層的節點權重值 (weight) 作計算, 然後在輸出層輸出計算結果, 此一階段是由輸入層往輸出層方向前進。倒傳遞學習階段為輸入層往輸出層方向前進, 目的是在修正隱藏層的節點權重值, 讓輸出結果跟預期輸出誤差越小。學習的終止條件可以設定為: (1) 當訓練到達一定次數就停止學習。(2) 誤差小於一定的閾值就停止學習。(3) 當學習的循環結束後對網路的能力進行測試, 以瞭解及推論能力是否達到所要求的目標, 若其推論能力達到目標則則停止學習。

倒傳遞演算法的學習流程如下 [18]:

步驟1: 網路初始化, 學習循環 $n=0$, 設定學習率並且使用隨機亂數將權重值初始化為 $-1 \sim 1$ 之間的實數。

步驟2: 讀入一個輸入向量。

步驟3: 計算出輸入向量隱藏層中類神經元的輸出。

$$y_j(n) = \varphi_j(v_j(n))$$

$$v_j(n) = \sum_{i=0}^p w_{ji}(n)y_i(n) \quad (26)$$

步驟4: 若輸出層中類神經元的輸出產生正確的分類, 跳至步驟7。

步驟5: 計算出隱藏層中類神經元的誤差函數

$$\delta_j(n) = (d_j(n) - O_j(n))O_j(n)(1 - O_j(n)) \quad (27)$$

計算出輸出層中類神經元的誤差函數

$$\delta_j(n) = y_j(n)(1 - y_j(n)) \sum_k \delta_k(n)w_{kj}(n) \quad (28)$$

步驟6: 調整所有類神經元的權重值向量

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n)y_i(n) \quad (29)$$

跳至步驟2。

步驟7: 若滿足終止條件則結束, 否則跳至步驟2。

在學習的過程中, 步驟1為初始化網路, 即是設定輸入、隱藏層和輸出層的節點數目、學習率與學習次數等參數, 並以一個隨機的亂數初始化權重值。步驟2將輸入資料讀入。步驟3與步驟4為前饋處理階段, 網路方向由輸入層至輸出層, 並得到一個推論輸出向量。步驟5與步驟6為倒傳遞處理階段, 網路方向由輸出層至輸入層, 並修正每個類神經元的權重值, 以讓推論輸出越來越逼近期望值。步驟7為終止條件的測試。

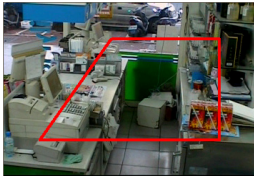
當網路訓練完成後, 便可以使用網路來做辨識的工作。在辨識的階段中只有利用前饋處理的階段, 不會再用到倒傳遞階段, 此時只要輸入色彩特徵, 便可以透過網路的回想過程得到一個推論向量, 那個推論向量就是網路的辨識結果。

五、實驗結果

在本論文中測試了超商與實驗室兩個不同的環境, 在超商環境中, 限制區域即為櫃檯區域, 在實驗室的環境中, 則使用模擬的方式, 假設部分的區域為限制性區域。由於目標物在剛進入攝影機的視範圍內時, 物體的資訊還不夠完整, 因此在限制區域內劃定一偵測區域 (如圖 14 所示), 當目標物之中心點落在此偵測區域時, 才會針對辨別目標物之身分。在超商的環境中, 因為攝影機架設的限制, 使得所得到的目標物影像只包含了上半身的區域。在實驗室的環境中, 為了測試本論文所提出之演算法在有完整的目標物影像時是否仍然能夠正確的運作, 因此刻意的將偵測區域往內縮減, 以確定所得到為包含整個身體區域的目標物影像。

其次, 偵測出目標物, 如圖 15 與圖 16 分別表示超商與實驗室環境偵測之目標物, 值得一提, 圖 15 (b) 與圖 16 (b) 中, 由於目標物部分區域與背景的色彩相似, 因此在使用背景相減法來做目標物偵測時, 目標物會有破碎的情形。

緊接著, 將身體區域劃分, 在圖 17 (b) 中, 此目標物帶著帽子, 在圖 18 (a) 中, 由於與背景色彩相似而使得無法偵測出完整的目標物, 在這些情況下, 本論文所提出之方法仍然能夠正確的將該目標物正確的切割。



(a) 超商環境

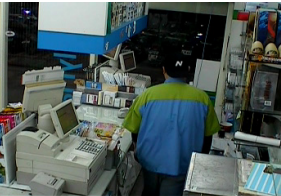


(b) 實驗室環境

圖 14 偵測區域之劃定



(a)



(b)



圖 15 超商環境中目標物偵測



(a)

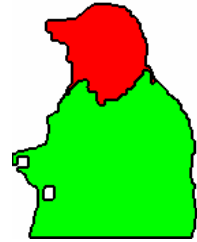


(b)

圖 16 實驗室環境中目標物偵測



(a)



(b)



(c)



(d)

圖 17 超商環境中身體區域區分



(a)



(b)



(c)



(d)

圖 18 實驗室環境中身體區域區分

最後，對進入者進行身份合法性判別，圖 19 與圖 20 為超商環境中的實驗結果。實驗室環境假設此限制區域的人員的合法制服為深藍色的短袖上衣，所此制服所統計出的色彩分布範圍如圖 21。



圖 19 超商環境中合法進入者



圖 20 超商環境中非法進入者

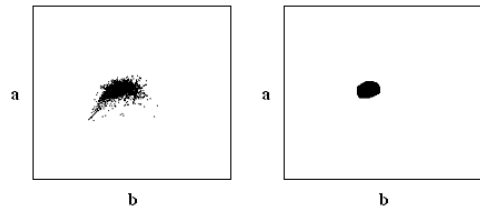


圖 21 實驗室限制區域制服色彩之統計

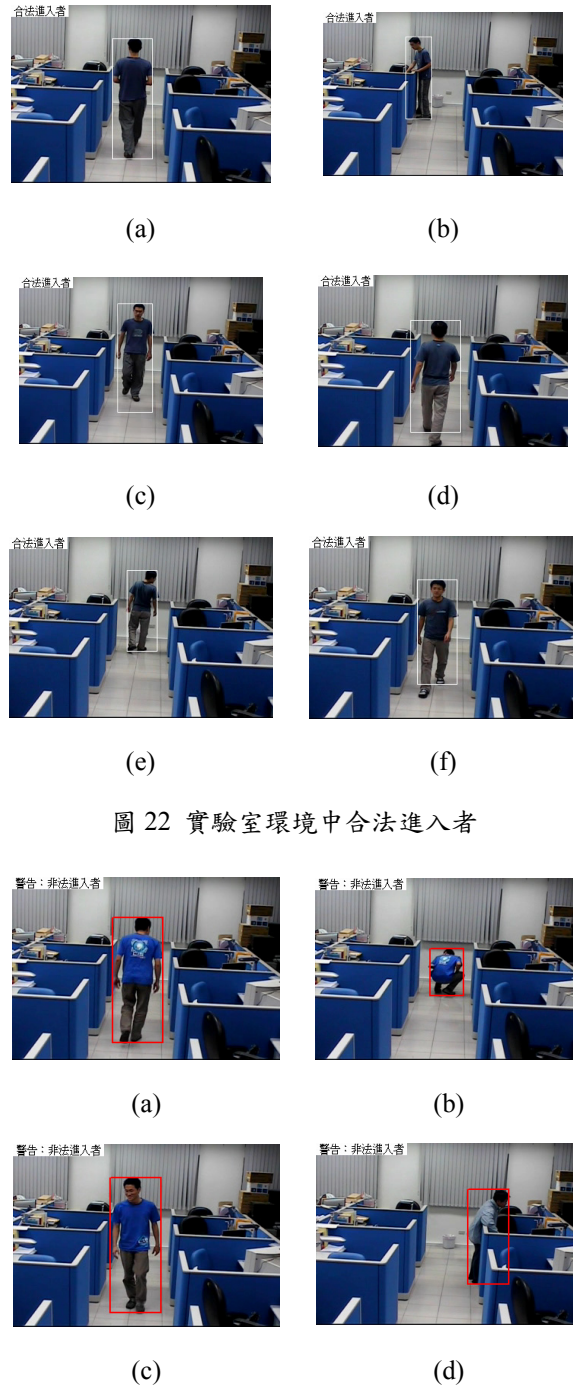
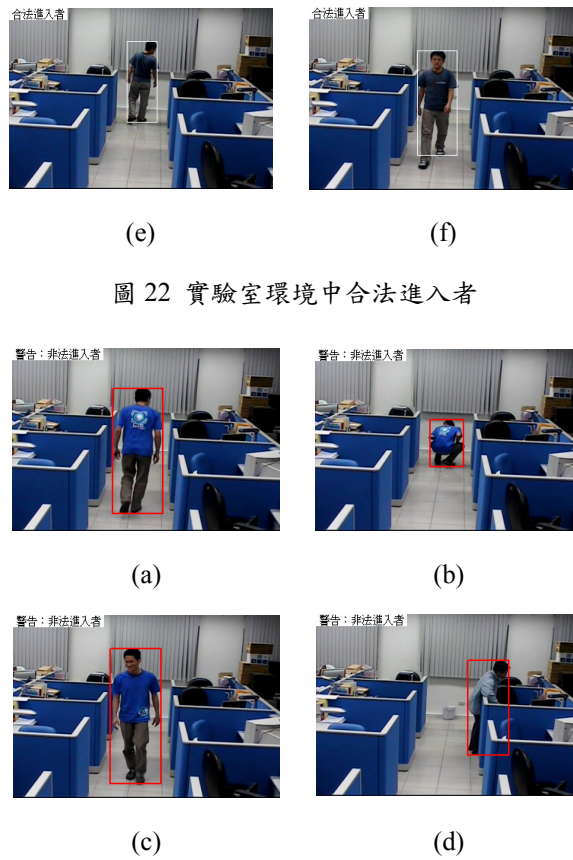


圖 22 實驗室環境中合法進入者





(e) (f)

圖 23 實驗室環境中非法進入者

在此我們嘗試改變環境中的亮度，以測試亮度對於之前所提出之制服色彩統計的影響。實驗結果顯示，當亮度輕微的改變時，所提出之方法可以正確的判斷制服色彩。不過當環境中亮度劇烈的變化時（如圖 24），部分的制服色彩並不會落在所統計之範圍內，不過由於限制區域大多是在室內的環境中，具有穩定的光線來源，因此這樣的結果並不會過度的影響本論文所提出的方法。



(a) (b)

圖 24 不同光線下的偵測結果

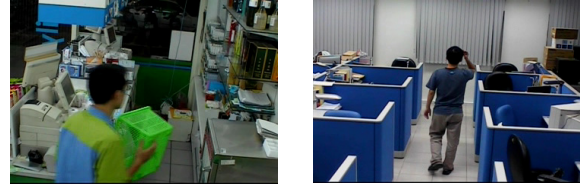
表 1 實驗準確度

	FRR	FAR	準確度
超商	10%	0%	93.8%
實驗室	6.7%	6.7%	93.3%

在超商環境中，實驗資料為 65 個進入櫃檯的人員，其中 40 個為合法進入者，25 個為非法進入者，各取 15 個進入者作為訓練資料。在實驗室環境中，實驗資料為 30 個進入限制區域的人員，其中 15 個為合法進入者，15 個非法進入者，各取 8 個進入者作為訓練資料。

實驗數據如表 1 所示，本論文所提出之方法在超商環境中準確度為 93.8%，實驗室環境中準確度為 93.3%。在超商環境中 FRR 為 10%，其中部分錯誤原因在於身體區域分割錯誤，使得抽取到非制服之色彩特徵。另一部分錯誤原因則在於由於合法進入者手中拿了物品，使得 ROI 變大，進而影響

制服色彩在 ROI 中的比例，產生錯誤的分類（如圖 25(a)）。在實驗室的環境中 FRR 為 6.7%，錯誤的原因為身體區域分割錯誤，使得抽取到非制服之色彩特徵（如圖 25 (b)，因頭部與手部相連在一起，使得頭部區域判斷錯誤）。在實驗室環境中 FAR 為 6.7%，錯誤的原因在於該非法進入者的褲子與制服之色彩相同，當身體區域分割錯誤時，使得抽取到與制服色彩相同之特徵，產生分類的錯誤。



(a) (b)

圖 25 產生錯誤分類之影像

在執行時間部分，本論文所提出之目標物偵測與追蹤之方法，在超商與實驗室環境中，平均每張 frame 需要 0.11 秒。在身分辨識的部分則包含了身體區域的分割與色彩特徵的抽取，由於此兩個步驟之執行時間會受到目標物大小與目標物內區塊數目的影響。在超商的環境中，目標物較大且目標物的色彩較多，因此平均每張 frame 需要 0.37 秒。在實驗室的環境中，平均每張 frame 則只需要 0.14 秒。

六、結論與未來方向

在本論文中，提出了一個以制服色彩為特徵的限制區域非法進入者偵測系統，對於限制區域內的目標物，進行偵測、追蹤並判別其身分合法性。在使用的場所方面，此系統主要應用於公共空間中的限制區域，如超商櫃檯、飯店櫃檯、銀行櫃檯與醫院護理站等區域環境，分別偵測、追蹤移動目標物，為了使提出方法能夠更廣泛的應用在各種不同的環境當中，並且增加辨識的準確度，對目標物做進一步分析，找出進入者上半身區域與下半身區域，並針對有穿著制服的區域進行色彩特徵的抽取。在特徵抽取的階段中，利用統計的制服色彩分布範圍來求得制服色彩在身體區域中分布的情形。在特徵的訓練階段，利用類神經網路中的倒傳遞演算法訓練一分類器，以將色彩特徵分類，並以此分類結果來判斷目標物之身分。

本論文的未來工作可以分為下列幾項：

1. 使用更準確的目標物偵測方法：以能夠解決環境亮度變化或目標物與背景色彩相似而使得目標物破碎的問題。

2. 增加目標物的分類器：設計一分類器，以判斷進入限制區域內的物體是否為所要處理的目標物。
3. 增加事件的定義：將非法進入者的行為視為是一個非法的事件，針對此事件的發生建立索引以供查詢。
4. 更穩定的身體區域分割演算：解決分水嶺分割法過度分割與過度合併的問題。
5. 更強健的身體模組：能夠更準確分割身體區域，不受身體部分區域破碎的影響。

七、致謝

The work was supported by Technology Development Program for Academia of DOIT, MOEA, Taiwan under grant no. 94-EC-17-A-02-S1-032.

八、參考文獻

- [1] T. Horprasert, D. Harwood and L.S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Proc. IEEE Frame-Rate Workshop*, 1999.
- [2] R. Cucchiara, C. Grana, M. Piccardi and A. Prati, "Detecting moving objects, ghosts and shadows in video streams," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1337-1342, Oct. 2003.
- [3] O. Javed, K. Shafiqu and M. Shah, "A hierarchical approach to robust background subtraction using color and gradient information," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 22-27, 2001.
- [4] L.D. Stefano, S. Mattoccia and M. Mola, "A change-detection algorithm based on structure and colour," in *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, pp.252-259, 2003.
- [5] S. Khan and M. Shah, "Tracking people in presence of occlusion," in *Proc. Asian Conf. Computer Vision, Taipei, Taiwan*, 2000.
- [6] A. Cavallaro, O. Steiger and T. Ebrahimi, "Tracking video objects in cluttered background," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 4, Apr. 2005.
- [7] F. Brémond and M. Thonnat, "Tracking multiple nonrigid objects in video sequences," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 575-584, Sep. 1998 .
- [8] H.T. Chen, H.H. Lin and T.L. Liu, "Multi-object tracking using dynamical graph matching," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 210-217, 2001.
- [9] I. Haritaoglu, D. Harwood and L.S. Davis, "W⁴: real-time surveillance of people and their activities," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, Aug. 2000.
- [10] M.D. Beynon, D.J. Van Hook, M. Seibert, A. Peacock and D. Dudgeon, "Detecting abandoned packages in a multi-camera video surveillance system," in *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, pp. 221-228, 2003.
- [11] S. Park and J.K. Aggarwal, "Segmentation and Tracking of Interacting Human Body Parts under Occlusion and Shadowing," in *Proc. IEEE Workshop Motion and Video Computing*, pp. 105-111, 2002.
- [12] L. Zhao and C.E. Thorpe, "Stereo- and neural network-based pedestrian detection," *IEEE Trans. Intelligent Transportation Systems*, vol. 1, no. 3, pp. 148-154, Sep. 2000.
- [13] ISO/IEC 15938-3:2001, "Multimedia Content Description Interface-Part 3: Visual," Version 1.
- [14] P. Salembier and M. Pardàs, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. Image Processing*, vol. 3, no. 5, 639-651, Sep. 1994.
- [15] M. Kim, J.G. Choi, D. Kim, H. Lee, M.H. Lee, C. Ahn and Y.S. Ho, "A VOP generation tool: automatic segmentation of moving objects in image sequences based on spatio-temporal information," *IEEE Trans. Circuits and System for Video Technology*, vol. 9, pp. 1216-1226, Dec. 1999.
- [16] S. Birchfield, "An elliptical head tracker," in *Proc. IEEE Conf. Signal, System and Computers*, vol. 2, pp.1710-1714, 1997.
- [17] L.G. Shapiro and G.C. Stockman, "Computer Vision," *New Jersey: Prentice Hall*, 2001.
- [18] 蘇木春, 張孝德, "機器學習：類神經網路、模糊系統以及基因演算法則", 全華科技圖書股份有限公司, 2003.