

第三章 背景濾除演算法

3.1 簡介

高斯混合模型的建立方式已於第二章中詳細的介紹，對於進行背景濾除的背景模型而言，選擇影像中的何種資訊來當特徵向量，是一件重要的工作，顏色和梯度是其中二個最常被拿來當做特徵向量的資訊，但是純粹以顏色向量為基礎的背景模型，容易因為照度上的改變而受影響，相對地，梯度對照度的敏感度較低，所以結合這兩個資訊所建立的 GMM 背景濾除系統，可以降低照度的影響。

一般最基本的背景濾除方式是以像素為單位，像素點各自決定是前景還是背景，但是若只採用像素階層的處理，似乎無法得到很好的結果，所以 Toyama et. Al.[10] 提出了透過像素(pixel)、區域(region)以及整張影像(frame)三種階層的處理，來解決背景濾除的問題，此方法只採用顏色的特徵向量，並用 Weiner Filter 去估測背景模型。而 Javed et. Al.[18]則提出一個階層式的背景濾除方式，並採用顏色與梯度的資訊來建立 GMM 的背景模型，此方法能解決照度的變化、背景物體的移動、初始化時有物體的移動...等問題，但是卻不能偵測出 Shadow 或 Highlight，以及一些顏色和梯度都有劇烈改變的區域，所以我們採用此階層式的架構，並提出能和 GMM 結合的陰影濾除的方法，來達到一個穩健且適應性的背景濾除效果。

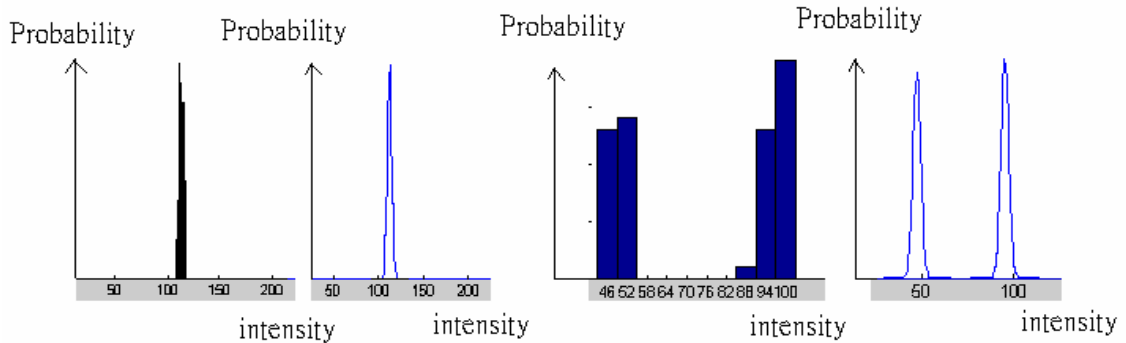
3.2 以顏色為基礎之背景濾除

以像素的顏色(R, G, B)為特徵向量，收集一段時間所得到的影像序列，建立每一個像素的高斯混合模型，像素(i, j)在時間點 t 的第 k 個高斯分佈如下所示：

$$g(x_{i,j}^t | m_{i,j}^{t,k}, \Sigma_{i,j}^{t,k}) = \frac{1}{(2\pi)^{n/2} |\Sigma_{i,j}^{t,k}|^{1/2}} \exp \left[-\frac{1}{2} (x_{i,j}^t - m_{i,j}^{t,k})^T (\Sigma_{i,j}^{t,k})^{-1} (x_{i,j}^t - m_{i,j}^{t,k}) \right] \quad (3.1)$$

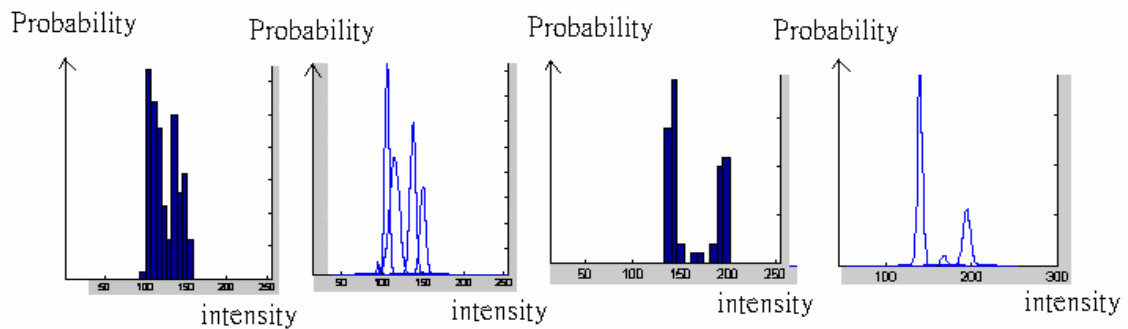
$x_{i,j}^t$ 是 (i, j) 的顏色向量， $m_{i,j}^{t,k}$ 和 $\Sigma_{i,j}^{t,k}$ 則是此分佈的平均向量和變異矩陣，其中因為顏色的三個元素彼此獨立，所以 $\Sigma_{i,j}^{t,k} = \text{diag}[(\sigma_{i,j}^{t,k,R})^2, (\sigma_{i,j}^{t,k,G})^2, (\sigma_{i,j}^{t,k,B})^2]$ ，符合第二章 GMM 建立的假設。

下圖 3-1 是我們觀察某一個顏色元素(R.G.B 任一)，在不同像素時的真實分佈，與我們建立的 GMM 分佈的比較圖，明顯看出不管是單一高斯分佈或是多個高斯分佈，它們的分佈都非常地接近；一般在靜態影像或是室內環境，大部分的像素分佈都如圖(a)所示，是單一的高斯分佈；而圖(b)可能是發生在日光燈閃爍、螢幕掃描、物體週期性搖晃...等的像素點上，所造成在兩個不同顏色的跳動，因此兩個像素分佈的比重相當；而圖(c)可以看出，即使是一個變化不是很巨大的單一分佈，也可以用 GMM 做更精準的模型化；至於圖(d)可能是發生在系統建立背景模型時，有前景物體出現的情況，因為前景物體出現的時間較短，所以其分佈的比重較背景分佈小很多。



(a) 單一高斯分佈

(b) 兩群比重相當的高斯分佈



(c) 精準化模型的分佈

(d) 某一類比重很大的 GMM 分佈

圖 3-1 像素的真實顏色分佈與 GMM 的比較圖

3.2.1 比對方法[3]

建立好高斯混合模型後，當有一張新的影像進來，以像素為單位，與該像素點的高斯混合分佈進行比對，判斷此像素是否為前景。首先要決定哪些高斯能代表目前的背景，不同的高斯代表不同的顏色分佈，其比重的含意表示此顏色停留的時間，而最能代表背景的顏色分佈，即停留的時間越久且變化越小(相當於靜止)，換句話說，就是比重越大、變異數越小的分佈，因此定義一個特性叫 $\text{fitness}=(\text{比重}/\text{變異數})$ ，根據 fitness 的值，由大到小重新排列高斯的分佈，並找出滿足下式的前 B 個分佈，來代表目前的背景：

$$B = \arg \min_b \left(\sum_{k=1}^b w_{i,j}^{t,k} > T \right) \quad (3.2)$$

上式中的 T 代表背景佔整個高斯混合模型的最小比重，當 T 取的太小，則背景可能只是單一模型(uni-modal)，也就是說只取比重最大的高斯來當做背景；相反地， T 取的越大，越多的高斯能併入背景，背景能容許的顏色分佈就越多。決定好背景的分佈後，接下來將新的顏色向量 $X_{i,j}^{t+1}$ 依 fitness 順序與背景分佈一一進行比對，如果此向量代入高斯分佈所得到的機率值，大於它 2.5 個標準差的機率值，表示此向量屬於這群分佈，屬於背景的一部分；如果都沒有比對成功，表示它可能是前景，再將它與其它不屬於背景分佈的高斯進行比對，如果還是沒有比對成功，表示此顏色是第一次出現，所以產生以 $X_{i,j}^{t+1}$ 為平均向量、變異矩陣很大、比重很小的一個高斯分佈。

3.2.2 更新高斯混合模型的參數[3]

高斯混合模型最主要的目的，就是希望能建構出一個最能代表目前狀況的背景模型，因此在系統初始化後，我們必須藉由新的影像資訊，不斷地更新模型的參數，來紀錄所有可能發生的狀況，包括背景的變化、前景的加入、

光影的變動...等。所以新的向量 $X_{i,j}^{t+1}$ 經過比對後，不管是前景或背景，都要透過下面的更新方程式，來更新模型的參數：

$$m_{i,j}^{t+1,k} = (1 - \rho)m_{i,j}^{t,k} + \rho X_{i,j}^{t+1} \quad (3.3)$$

$$\Sigma_{i,j}^{t+1,k} = (1 - \rho)\Sigma_{i,j}^{t,k} + \rho(X_{i,j}^{t+1} - m_{i,j}^{t+1,k})^T(X_{i,j}^{t+1} - m_{i,j}^{t+1,k}) \quad (3.4)$$

$$w_{i,j}^{t+1,k} = (1 - \alpha)w_{i,j}^{t,k} + \alpha \hat{p}(w_{i,j}^{t,k} | X_{i,j}^{t+1}) \quad (3.5)$$

$$\rho = \alpha g(X_{i,j}^{t+1} | m_{i,j}^{t,k}, \Sigma_{i,j}^{t,k}) \quad (3.6)$$

若 $X_{i,j}^{t+1}$ 與第 K 個高斯比對成功，則令 $\hat{p}(w_{i,j}^{t,k} | X_{i,j}^{t+1})=1$ ，使用上列四個式子來更新此高斯分佈的所有參數；反之，若未比對成功，則令 $\hat{p}(w_{i,j}^{t,k} | X_{i,j}^{t+1})=0$ ，只更新比重的參數。 α 和 ρ 都算是新向量 $X_{i,j}^{t+1}$ 拿來更新參數的一種權重， α 值取的越大， $X_{i,j}^{t+1}$ 對背景參數的影響越大，相對地，新的前景資訊也越快被併入背景模型中，此權重 α 一般稱它為背景模型的學習比率 (learning rate)。

這種更新方法有一個很大的優點，就是原本已建立好的背景高斯分佈，並不會因背景有所改變，而遭受到破壞，它仍保有它原本的顏色分佈，只是比重的大小會隨出現次數的多寡而有所變動，舉例來說，當一個背景物體被移走，會產生一群新的高斯分佈，但原本物體顏色的高斯分佈並不會消失，它的平均向量與共變異矩陣保持不變，只是比重會慢慢降低，所以一旦物體被放回原位，它很快地又會被併入背景模型中。

3.3.3 學習比率(Learning Rate)的調整

在某些情況下，一些靜態的變化所造成的前景(如椅子的移動)，我們希望它盡快地被併入到背景模型，但在某些情況下，一些屬於真正前景的物體，雖然它保有長時間相對靜止的特性(如兩個人停在原地對話，或穿著均勻顏色衣服的人做緩慢地移動...等)，我們反而希望降低它併成背景的速度。在傳統的更新方程式中，所有的像素都是採用相同的學習比率，所以要對上面敘述的兩種情況取得一個平衡點(trade-off)，似乎是很困難地，因此我們加入了可以自動調整學習比率的機制。

我們可以利用兩張連續影像，計算出影像中所有像素點的活動程度(Activity-Level) A ，若 A 的值大於某一個很低的門檻值 H ，表示此像素的活動並非完全靜止，則將此像素的學習比率降低某個比例 ξ ，使其併入背景模型的速度降低；相反的，若像素的 A 小於 H ，表示其活動幾乎是靜止，則保持原學習比率。活動程度 A 的計算，由下列(3.7)所示：

$$A_{i,j}^{t,k} = (1-\lambda)A_{i,j}^{t-1,k} + \lambda|X_{i,j}^t - X_{i,j}^{t-1}| \quad (3.7)$$

$|X_{i,j}^t - X_{i,j}^{t-1}|$ 代表目前與前一張影像像素的顏色差距，差距越大使得 A 的值越大，表示像素的活動性越大， λ 代表此差距在時間平滑上的權重大小， A 的初始值是 0，藉由不斷地遞迴運算，得到目前的 A 。透過這樣的機制，每個像素可以根據自己分佈的狀況，調整屬於自己的學習比率，來決定自己併入背景的速度。

綜合了 3-2 節的敘述，以顏色為基礎的背景濾除其整體的架構如圖 3-2 所示，包含建立高斯混合模型、新特徵向量的比對、模型參數的更新，以及我們提出可調整學習比率的機制(黃色框框)四個主要部分。

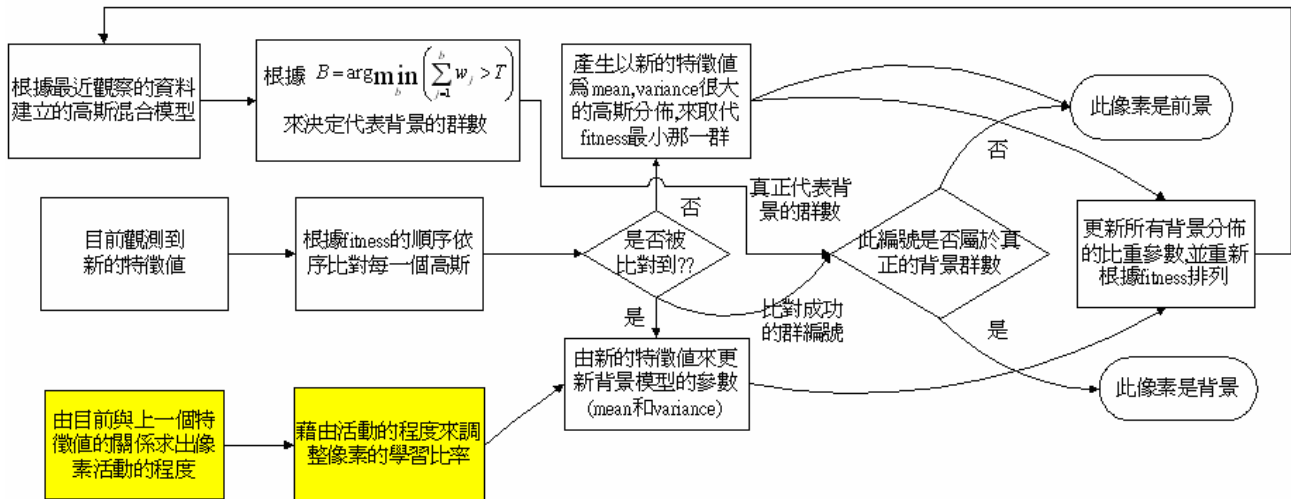


圖 3-2 以顏色為基礎之背景濾除的整體架構圖

3.3 以梯度為基礎之背景濾除[18]

以顏色為基礎的背景濾除，我們採用的特徵向量是(R, G, B)向量，至於在以梯度為基礎的背景濾除，當然就是採用梯度的大小 Δ_m 和梯度的方向 Δ_d ，所組成的梯度向量 Δ 為特徵向量，梯度的大小可由 $\sqrt{f_x^2 + f_y^2}$ 得到，梯度的方向則是 $\tan^{-1} \frac{f_y}{f_x}$ ，其中 f_x 和 f_y 分別代表灰階值(gray-level)影像中，x和y方向的梯度值。

與以顏色為基礎的背景濾除相似，也需要建立 Δ 的高斯背景模型，時間點t像素位置(i, j)的顏色向量 $x_{i,j}^t = [R, G, B]$ ，其灰階值 $g_{i,j}^t$ 透過 $g_{i,j}^t = \alpha R + \beta G + \gamma B$ 獲得，若已知第k群的背景參數(R, G, B三個平均值和變異數)，且在(R, G, B)三個顏色的分佈是互相獨立的前提下，因為R, G, B的分佈是高斯，所以三個高斯分佈做線性組合的分佈亦是高斯分佈，我們可以得到灰階值的分佈如下：

$$g_{i,j}^{t,k} \sim N(\mu_{i,j}^{t,k}, (\sigma_{i,j}^{t,k})^2)$$

$$\mu_{i,j}^{t,k} = \alpha m_{i,j}^{t,k,R} + \beta m_{i,j}^{t,k,G} + \gamma m_{i,j}^{t,k,B} \quad (3.8)$$

$$(\sigma_{i,j}^{t,k})^2 = \alpha^2 (\sigma_{i,j}^{t,k,R})^2 + \beta^2 (\sigma_{i,j}^{t,k,G})^2 + \gamma^2 (\sigma_{i,j}^{t,k,B})^2 \quad (3.9)$$

在(3.8)與(3.9)式中, $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$, 接下來定義灰階影像的梯度值, $f_x^k = g_{i+1,j}^{t,k} - g_{i,j}^{t,k}, f_y^k = g_{i,j+1}^{t,k} - g_{i,j}^{t,k}$, 並假設影像中鄰近二個像素點, 其灰階值的分佈彼此是不相關的前提下, 得到第 k 群的兩個梯度分佈如下:

$$f_x^k \sim N(\mu_{f_x}^k, (\sigma_{f_x}^k)^2), f_y^k \sim N(\mu_{f_y}^k, (\sigma_{f_y}^k)^2)$$

$$\mu_{f_x}^k = \mu_{i+1,j}^{t,k} - \mu_{i,j}^{t,k}, \mu_{f_y}^k = \mu_{i,j+1}^{t,k} - \mu_{i,j}^{t,k} \quad (3.10)$$

$$(\sigma_{f_x}^k)^2 = (\sigma_{i+1,j}^{t,k})^2 + (\sigma_{i,j}^{t,k})^2, (\sigma_{f_y}^k)^2 = (\sigma_{i,j+1}^{t,k})^2 + (\sigma_{i,j}^{t,k})^2 \quad (3.11)$$

雖然假設像素的灰階值彼此是獨立的, 但是像素的 f_x 和 f_y 彼此之間卻不是獨立的, 可由它們兩個的共變異數(Covariance)不為零得知:

$$\begin{aligned} Cov(f_x^k, f_y^k) &= Cov(g_{i+1,j}^{t,k} - g_{i,j}^{t,k}, g_{i,j+1}^{t,k} - g_{i,j}^{t,k}) \\ &= Cov(g_{i,j}^{t,k}, g_{i,j}^{t,k}) = (\sigma_{i,j}^{t,k})^2 \end{aligned} \quad (3.12)$$

已知 f_x 和 f_y 的分佈皆是高斯分佈, 且兩個不是互相獨立地, 可以根據標準的分佈轉換方式[17], 找出相互高斯(Jointly Gaussian)分佈:

$$\begin{aligned} F_{f_x^k, f_y^k}(f_x^k, f_y^k) &= F_{f_x^k, f_y^k}(f_x^k)(f_y^k) \quad (3.13) \\ &= \frac{1}{2\pi \sigma_{f_x}^k \sigma_{f_y}^k \sqrt{1-\rho^2}} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[\left(\frac{f_x^k - \mu_{f_x}^k}{\sigma_{f_x}^k} \right)^2 - 2\rho \frac{(f_x^k - \mu_{f_x}^k)(f_y^k - \mu_{f_y}^k)}{\sigma_{f_x}^k \sigma_{f_y}^k} + \left(\frac{f_y^k - \mu_{f_y}^k}{\sigma_{f_y}^k} \right)^2 \right] \right\} \end{aligned}$$

其中 ρ 是相關係數(Correlation Coefficient), 它的定義是:

$$\rho = \frac{\Delta m_{11}}{\sqrt{m_{20}m_{02}}} = \frac{Cov(f_x^k, f_y^k)}{\sqrt{Cov(f_x^k, f_x^k)Cov(f_y^k, f_y^k)}} = \frac{(\sigma_{i,j}^{t,k})^2}{\sigma_{f_x}^k \sigma_{f_y}^k} \quad (3.14)$$

但是我們的目的是希望能找出特徵向量 $[\Delta_m, \Delta_d]$ 的分佈，所以透過

$f_x = \Delta_m \cos \Delta_d$ 和 $f_y = \Delta_m \sin \Delta_d$ 兩個關係，將(3.13)式改寫成下式：

$$F^k(\Delta_m, \Delta_d) = \frac{\Delta_m}{2\pi \sigma_{f_x}^k \sigma_{f_y}^k \sqrt{1-\rho^2}} \exp\left(-\frac{z}{2(1-\rho^2)}\right) \quad (3.15)$$

$$z = \left(\frac{\Delta_m \cos \Delta_d - \mu_{f_x}^k}{\sigma_{f_x}^k}\right)^2 - 2\rho \frac{(\Delta_m \cos \Delta_d - \mu_{f_x}^k)(\Delta_m \sin \Delta_d - \mu_{f_y}^k)}{\sigma_{f_x}^k \sigma_{f_y}^k} + \left(\frac{\Delta_m \sin \Delta_d - \mu_{f_y}^k}{\sigma_{f_y}^k}\right)^2$$

背景梯度的分佈如(3.15)式所示，式子中的所有參數皆可由之前建立的顏色背景高斯分佈的平均向量和共變異矩陣得到，當有一個新的影像，先計算其特徵向量 $\Delta = [\Delta_m, \Delta_d]$ ，再將此特徵向量代入(3.15)式的背景梯度分佈方程式，若得到的機率值小於 T_g ，則表示此像素不屬與該群，可能是前景，反之，機率值大於 T_g ，則是背景。

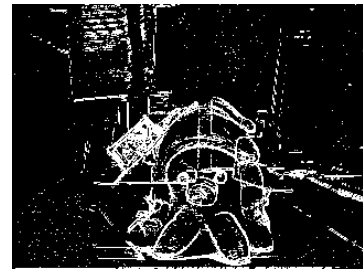
因為建立顏色背景模型是採用高斯混合模型，所以我們選擇其中比重最大的高斯分佈，並結合短期模型的觀念，求出相對應的梯度分佈，至於為什麼選取最大比重的高斯以及如何結合短期模型，將在後面 3.6 節作詳細地說明。在以梯度為基礎的背景濾除，因為所有的參數都是由顏色的背景參數計算而來，所以根本無須對參數作更新的動作。圖 3-3 是做梯度濾除的實驗，(a)是背景影像，(b)是一張有前景出現的測試影像，(c)是做完梯度濾除的結果。



(a)背景影像



(b)前景的進入



(c)梯度濾除的結果

圖 3-3 梯度濾除的實驗

3.4 區域階層處理(Region Level Processing)

前三節都是以像素為單位，分別作顏色和梯度的背景濾除，可以得到各自的前景，而本節所做得區域階層處理，顧名思義就是加入了區域的觀念，首先將顏色濾除的前景像素點巨集成一塊塊的前景區域，再由梯度濾除的前景結果，來確認這些前景區域的有效性(validity)，簡單說就是結合了顏色和梯度的資訊，再做一次前景與背景的判斷。

在做完背景濾除後，與作區域階層處理之前，我們有一些前處理必須先做，首先是在作顏色濾除後的二值影像中，會發現有很多雜訊點的產生，這些雜訊不但對之後前景的區塊化，或是最終的前景背景決定，甚至於背景濾除後的應用，如追蹤或辨識，都有很大的影響，所以我們採用一些形態學的濾波方式來消除它；至於前景物件的區塊化，我們可以透過連結區塊標記的方法，來紀錄共有幾個前景區域，以及每個前景像素屬於哪一個區域。

3.4.1 形態學的濾波(Morphological Filtering)

形態學(Morphology) 原本係研究動植物形態的生物學門，但在影像處理方面，則用來專指抽取出二值影像中特定成份的技術，如影像分割(Segmentation)，邊緣檢測 (Edge detection)，細線化 (Thinning)，骨架抽出(Skeletonizeing)皆可說是形態學上的運算之一。在這裡我們採用兩個基本的形態學運算-膨脹(Dilation)和侵蝕(Erosion)，它們的定義如下：

$$B \oplus X = \{b | (X)_b \cap B \neq \emptyset\} \quad (3.16)$$

$$B \ominus X = \{b | (X)_b \subseteq B\} \quad (3.17)$$

上式中的 \oplus 和 \ominus 二個運算子分別代表膨脹與侵蝕，B 是欲做濾波的二值影像，X 則是結構元素，這裡我們採用 3x3 的矩形遮罩(Mask)，圖 3-4 是一張二值影像 B 經過結構遮罩 X 膨脹和侵蝕後的結果。

1	1	1	1	1	1	1	
			1	1	1	1	
			1	1	1	1	
		1	1	1	1	1	
			1	1	1	1	
		1	1				

(a)二值影像 B

1	1	1
1	<u>1</u>	1
1	1	1

(b)3*3 結構遮罩 X

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1

(c)膨脹後的影像

				1	1		
				1	1		
				1	1		

(d)侵蝕後的影像

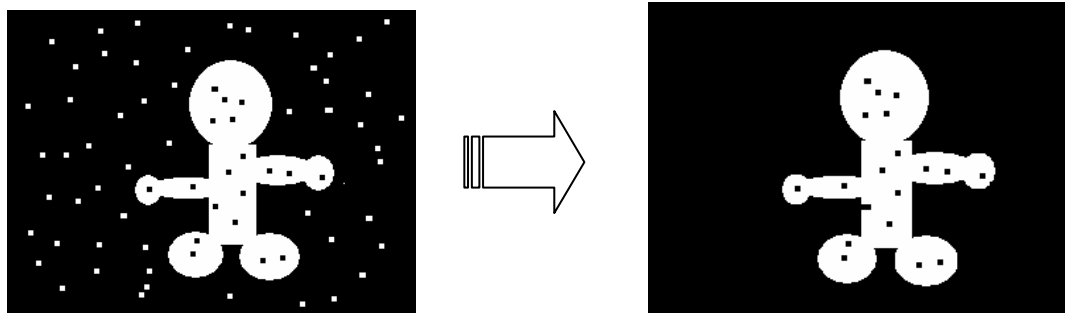
圖 3-4 二值影像做侵蝕與膨脹的例子

組合 \oplus 和 \ominus ，可得到兩個新的操作斷開(Opening)與閉合(Closing)，其定義如下：

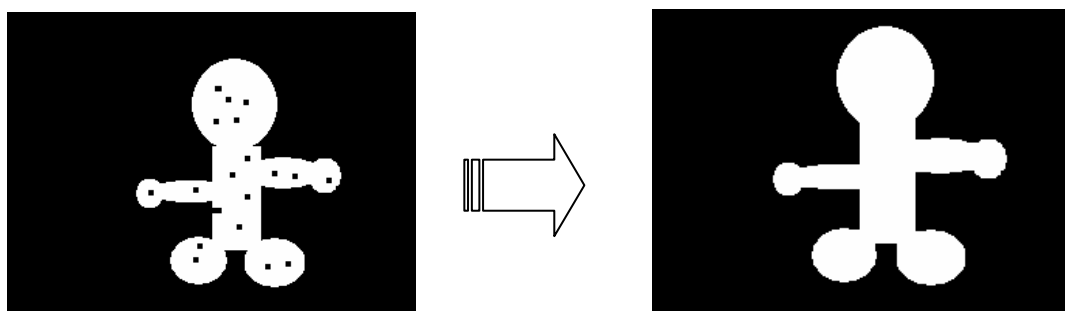
$$X \circ B = (X \ominus B) \oplus B \tag{3.18}$$

$$X \bullet B = (X \oplus B) \ominus B \tag{3.19}$$

下圖 3-5 是一個斷開與閉合的例子，先用斷開運算子來消除前景之外的一些孤立(Isolated)雜點，再用閉合運算子來填補前景物件內部的空洞(Holes)。



(a) 斷開運算



(b) 閉合運算

圖 3-5 做斷開與閉合的例子

3.4.2 連結區塊標記(Connected Component Labeling)

背景濾除後，前景的部分已經從背景分離出來，若前景的物體有很多個，則必須對這些分開的物體做標記(Labeling)，做標號的目的是要讓每一個連結區塊有不同的記號，屬於相同區塊的像素有相同的記號，並由標記的個數推知前景物件的個數，這裡我們採用一種遞迴(Recursive)標記的演算法 [19]來完成之，其步驟說明如下，而整個演算法的流程見圖 3-8：

Step 0：二值影像 B，前景像素值是 1，背景像素值是 0。

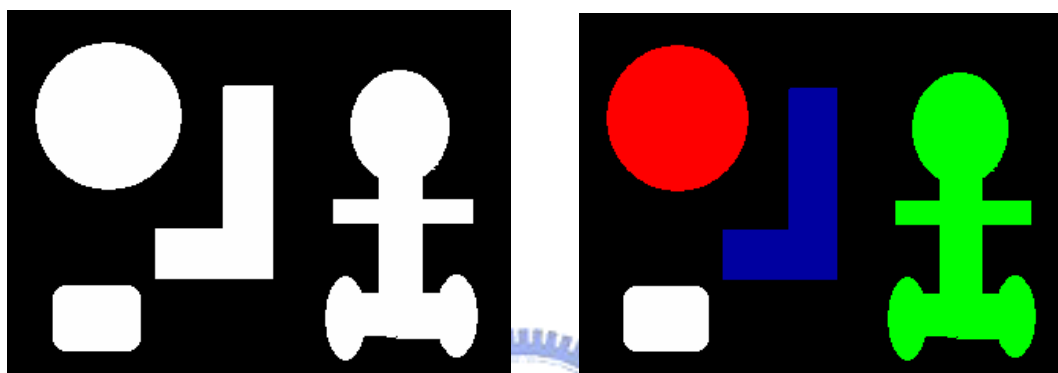
Step 1：將 B 送進 Negate()函數，將像素值是 1 的變成-1，得到一個標記的影像輸出 LB。

Step 2：影像從左到右、從上到下進行掃描，若有 $LB(x, y) = -1$ 的座標點，標記加 1，做 Step 3；反之，若整張影像都找不到 $LB(x, y) = -1$ ，表示已做完所有標記動作，做 Step 5。

Step 3: 做 Search 函數，就是將新的標記給 $LB(x, y)$ ，找出屬於 (x, y) 的鄰近區域(neighborhood)座標點 (x', y') 。

Step 4: 如果找到滿足 $LB(x', y')=1$ ，將 (x', y') 當作輸入，回到 Step 3；反之，回到 Step 2，如此不斷地遞迴收尋。

Step 5: 根據不同的標記，代表不同的前景物體，我們可以用不同的顏色來表示之，圖 3-6 是一個有四個連結區塊的例子。



(a)有 4 群前景的二值影像 (b)不同區域以不同顏色標記

圖 3-6 影像中有四個連結區塊的例子

Step 3 提到的鄰近區域，一般來說有二種，4-neighborhood 和 8-neighborhood，如下圖 3-7 所示，在這裡我們採用 8-neighborhood。

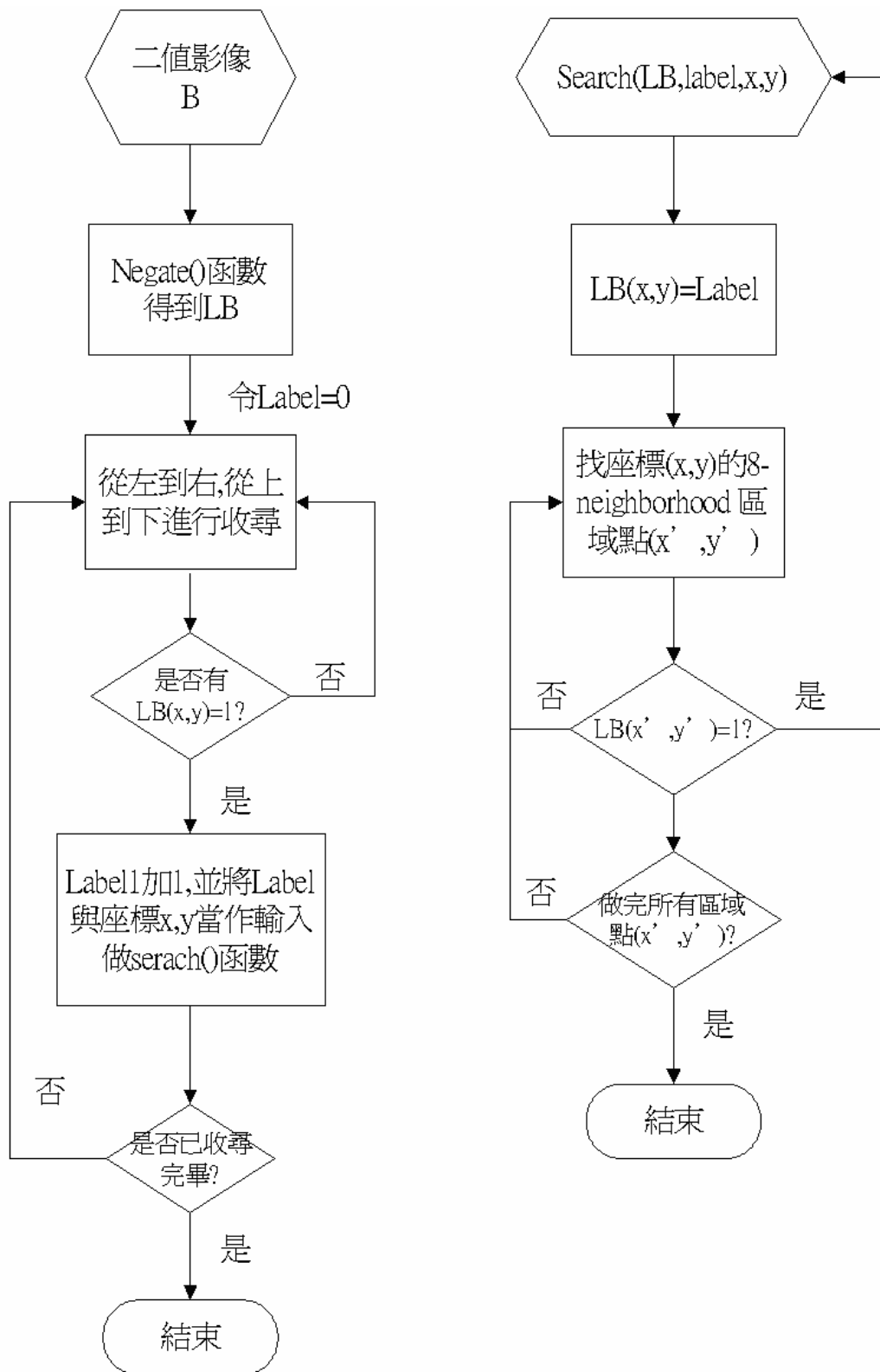
	1	
2	*	3
	4	

1	2	3
4	*	5
6	7	8

(a) 4-neighborhood

(b) 8-neighborhood

圖 3-7 兩種尋找鄰近區域的遮罩



3.4.3 前景/背景的決定

一張新的影像 I 進來，經過 3-1 節與 3-2 節以顏色與梯度為基礎的濾除後，可以分別得到 $C(I)$ 與 $G(I)$ 兩張二值影像，在結合這兩各資訊之前，先定義一些之後做判斷會用到的符號：

- (1) $C(i, j) = 1 \rightarrow$ 顏色濾除的前景像素點 (i, j) 。
- (2) $G(i, j) = 1 \rightarrow$ 梯度濾除的前景像素點 (i, j) 。
- (3) $R_a, 1 \leq a \leq M \rightarrow$ 在顏色濾除下，以連結區塊法找出的 M 個前景區域。
- (4) $\partial R_a \rightarrow$ 前景區域 R_a 的所有邊界像素點 (i, j) 的集合
- (5) $|\partial R_a| \rightarrow$ 屬於 ∂R_a 集合的個數
- (6) $\nabla I \rightarrow$ 用梯度的方法找出影像真實的邊緣

以數學的形式來看，只要滿足下面式子的前景區域，皆判斷成真正的前景：

$$\frac{\sum_{(i, j) \in \partial R_a} (\nabla I(i, j) G(i, j))}{|\partial R_a|} \geq P_B \quad (3.20)$$

它的基本觀念就是任何的前景區域相對於一個真實的物體，所以在它的邊界點上會有很大的梯度差，所以真正屬於前景的區域 R_a ，其邊界點 ∂R_a 在梯度濾除中判斷成前景的機率很大；相對地，一些可能因為照度的改變，所造成誤判成前景的區域 R_b (例如手電筒的照射)，其邊界點 ∂R_b 在梯度模型中的改變是很平滑地，所以它判斷成前景的機率就很低。下圖 3-9 是一個簡單的範例，圖(a)中手電筒的燈光與人同時出現在白牆前，圖(b)畫出第 90 行(紅色線)與第 150 行(藍色線)的梯度值，由圖(b)明顯地可以看出，在人的邊界上有很大大梯度值，但是在手電筒的燈光邊界上，梯度值較小。

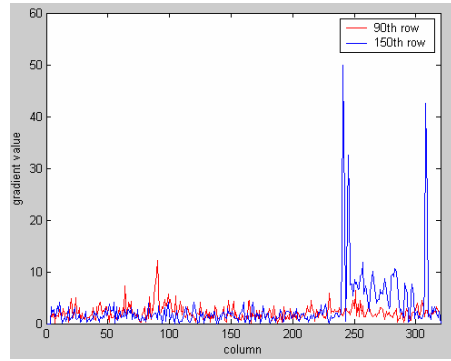
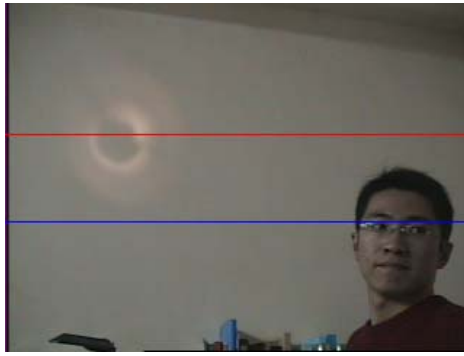


圖 3-9 (a)手電筒的燈光與人同時出現 (b)不同排的梯度分佈圖

(3.20)式多考慮了影像的真實邊緣 ∇I 資訊，這是因為當背景物體移動時，原來位置的區域 R_x 和移位後的位置區域 R_y ，它們的邊緣在梯度模型上判斷成前景的機率都很高，兩個都會當作是前景，所以增加區域的邊界是否符合影像邊緣的判斷，則因為 R_x 並沒有真正的邊緣，而不屬於真正的前景。

綜合前面的數學公式與實際上的涵義，總而言之，一個真正的前景區域，它的邊界點不只是在 $G(I)$ 佔有較大的比率，且也必須是屬於真實影像的邊緣。

有些可能是前景的區域，經過(3.20)式的判斷後，知道它並非真正的前景，那屬於這個區域的所有像素點，都應該是屬於背景，所以我們藉由調高學習比率的方式(我們採取的方式是調高一個倍數 δ)，讓這些像素能快速併入背景。

3.5 陰影濾除(Shadow Removal)

由於移動的前景物體所造成的光影變化，在視覺監視系統中是很常發生的，尤其是在室內環境下，因為前景相對於整張影像所佔的比例較大，其光影的影響(特別是 Highlight)特別明顯，這些變化往往經過背景濾除的處理後，會被誤判成前景，而與實際的物體結合在一起，所以陰影的偵測與濾除是必要的。在我們的論文中，所討論有關陰影濾除的部分，不單單只針對物體投影到背景所造成的真實陰影，凡是在『色彩』的特性上變化不大，且只有在『亮度』上有所改變的光影變化，我們統稱它為陰影，都將它們偵測出

來並濾除之。

在我們所採用的背景濾除演算法中，最後判斷前景或背景的依據，是透過前景區域的邊界點 ∂R_a ，與梯度濾除的結果 $G(i, j)$ 、影像真實的邊緣 ∇I 來進行比對，一旦因為光影的變化，使得 Shadow 或 Highlight 與物體結合，造成前景區域的邊界 ∂R_a ，與 $G(i, j)$ 、 ∇I 比對成功的機率降低，甚至小於門檻值 P_B ，導致前景區域容易被誤判成背景。下面舉兩個例子來說明之，圖 3-10 是大量的 Highlight 併入前景，圖 3-11 則是手投影到椅背上的陰影併入前景，圖(a)是原始影像，圖(b)是顯示 Highlight 與 Shadow 的分佈，其中藍色代表前景，綠色代表 Highlight，紅色代表 Shadow，圖(c)是顯示因為 Highlight、Shadow 併入前景，導致前景區域變大，其中白色代表前景，黑色代表背景，圖(d)和圖(e)分別代表 $G(i, j)$ 與 ∇I 的影像，明顯地看出圖(c)中前景區域的邊界點，已經很難和圖(d)、圖(e)比對成功，所以最後娃娃與手就會消失變成背景。

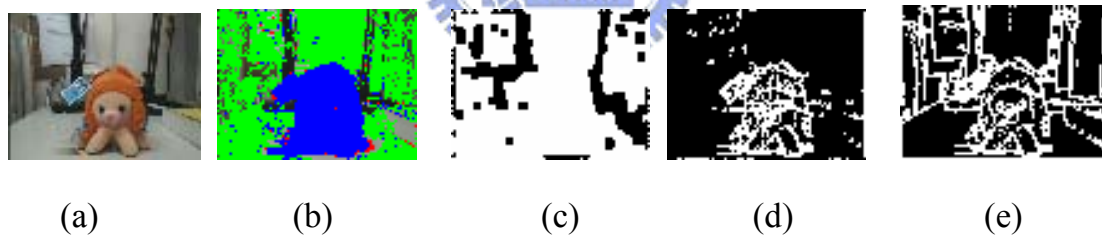


圖 3-10 因前景物體產生的 Highlight 對系統的影響



圖 3-11 因前景物體產生的 Shadow 對系統的影響

前景物體的消失對於視覺監視系統來說，是一件很嚴重且不被允許發生的事，所以我們一定要加入陰影濾除的機制，將 Highlight 和 Shadow 從前景中去除，至於陰影濾除的方法，前面 1.2.2 節已介紹了很多種，但那些方

法大多只針對 Shadow 的偵測，並未考慮所有光影的變化，加上背景模型的顏色都是決定性(Deterministic)地已知，並非如我們的模型是統計上的分佈，所以我們不採用它們。在我們的系統中，採用類似 T. Hoprasert etc. al 的統計上陰影偵測 [21]方法中，判斷 Shadow 或 Highlight 的觀念，提出一個與 GMM 結合的陰影濾除方法。

3.5.1 統計上非參數(s_{NP} ，Statistical Non-Parametric)方法[21]

s_{NP} 之陰影濾除方法是建立在人的視覺具有對顏色不變性(Color Constancy)的特性，以及利用朗伯假說(Lambertian Hypothesis)，對於完全無光澤表面的物體來說，其感知的顏色是透過照度(Illumination)與表面的反射係數(Surface Reflectance)乘積而得到，因而設計一個三度空間顏色 RGB 的模型，將任一顏色向量分解成亮度(Brightness)與色彩(Chromaticity)兩個成份，下圖 3-12 是三度空間 RGB 的模型， $E_i = [\mu_R(i), \mu_G(i), \mu_B(i)]$ 是收集了 N 張影像，所得到的參考或背景影像中，第 i 個像素的平均向量， α_i 和 CD_i 分別是新的影像 $I_i = [I_R(i), I_G(i), I_B(i)]$ 與 E_i 之間亮度和色彩的失真(Distortion)。

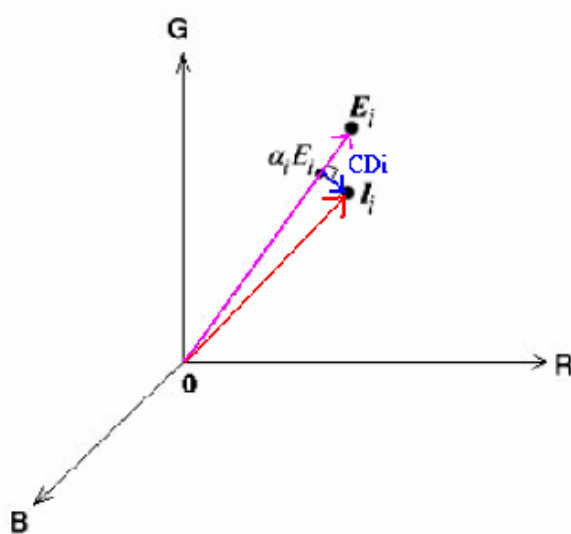


圖 3-12 三度空間 RGB 的顏色模型

亮度的失真 α_i ，是滿足使得新的顏色向量 I_i 與 E_i 最靠近時的一個純量，也就是滿足下式最小時的 α_i ：

$$\phi(\alpha_i) = (I_i - \alpha_i E_i)^2 \quad (3.21)$$

而 CD_i 則是定義 I_i 與 E_i 之間的垂直距離，可以直接將經由(3.21)求得的 α_i 代入下式得之：

$$CD_i = \|I_i - \alpha_i E_i\| \quad (3.22)$$

因為 RGB 三個頻帶的變化量都不同，所以必須透過 $S_i = [\sigma_R(i), \sigma_G(i), \sigma_B(i)]$ 將顏色做正規化，而 α_i 和 CD_i 的公式變成如下：

$$\alpha_i = \frac{\left(\frac{I_R(i)\mu_R(i)}{\sigma_R^2(i)} + \frac{I_G(i)\mu_G(i)}{\sigma_G^2(i)} + \frac{I_B(i)\mu_B(i)}{\sigma_B^2(i)} \right)}{\left(\left[\frac{\mu_R(i)}{\sigma_R(i)} \right]^2 + \left[\frac{\mu_G(i)}{\sigma_G(i)} \right]^2 + \left[\frac{\mu_B(i)}{\sigma_B(i)} \right]^2 \right)} \quad (3.23)$$

$$CD_i = \sqrt{\left(\frac{I_R(i) - \alpha_i \mu_R(i)}{\sigma_R(i)} \right)^2 + \left(\frac{I_G(i) - \alpha_i \mu_G(i)}{\sigma_G(i)} \right)^2 + \left(\frac{I_B(i) - \alpha_i \mu_B(i)}{\sigma_B(i)} \right)^2} \quad (3.24)$$

最後再對 α_i 和 CD_i 在背景影像中像素 i 的均方根值(RMS)進行正規化得到 $\hat{\alpha}_i$ 和 \hat{CD}_i ，藉由這兩個值與下面四個判斷的準則，來將像素進行分類：

$$C(i) = \begin{cases} \textit{Foreground} & ; \hat{CD}_i > \tau_{CD} \textit{ or } \hat{\alpha}_i < \tau_{alow} , \textit{ else} \\ \textit{Background} & ; \hat{\alpha}_i < \tau_{\alpha 1} \textit{ and } \hat{\alpha}_i > \tau_{\alpha 2} , \textit{ else} \\ \textit{Shadow} & ; \hat{\alpha}_i < 0 , \textit{ else} \\ \textit{Highlight} & ; \textit{ otherwise} \end{cases} \quad (3.25)$$

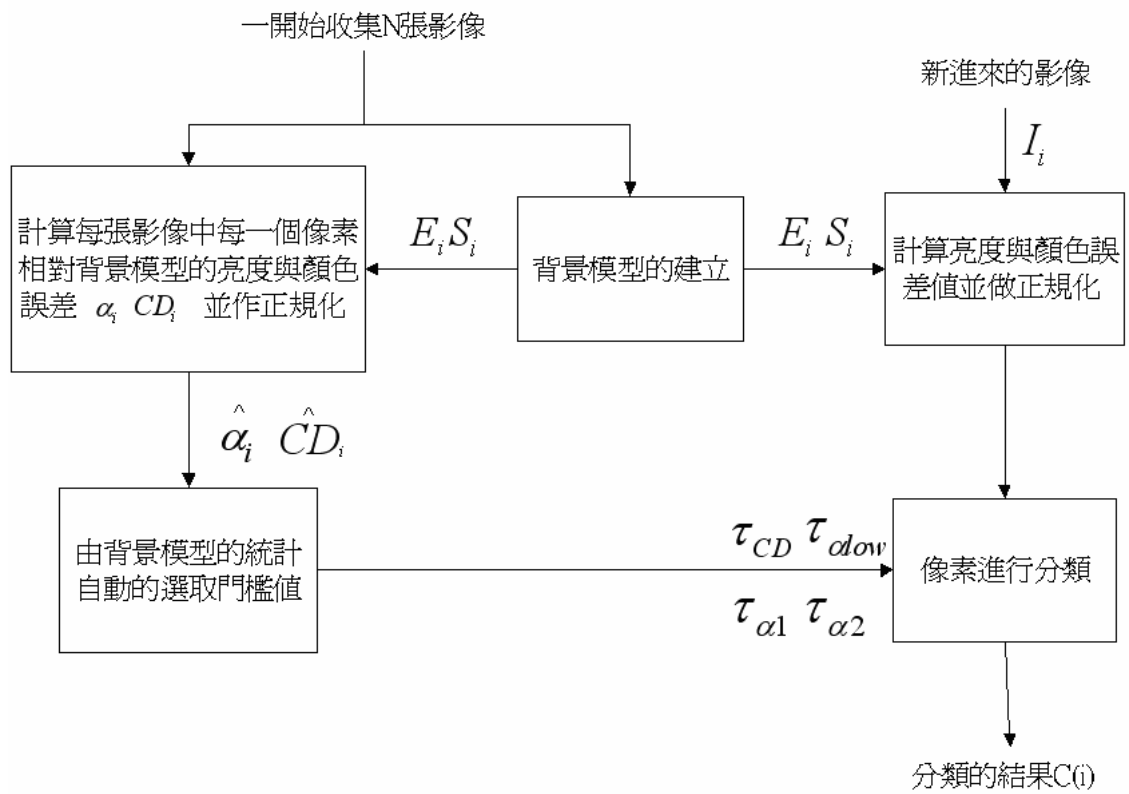


圖 3-13 S_{NP} 的流程圖

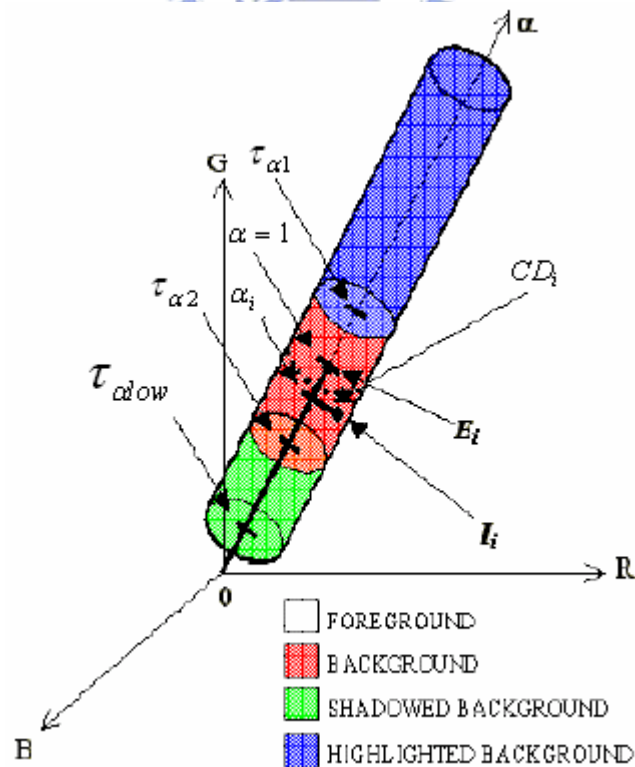


圖 3-14 在顏色 RGB 的三度空間中，像素分類的示意圖

圖 3-13 是 S_{NP} 方法的整個流程圖，主要包括建立背景模型、計算亮度色彩的失真和像素的分類三大部分，(3.25) 式中四個像素分類的判斷，主要的觀念是背景、Shadow 和 Highlight，它們在色彩上與背景變化不大，所以色彩的失真 \hat{CD}_i 較小，因此只要 \hat{CD}_i 大於色彩失真的門檻值 τ_{CD} ，即判斷是前景，但是一些較暗的前景，因其 RGB 向量較小，造成 \hat{CD}_i 的值相對地也小，容易誤判成陰影，所以加入了 τ_{alow} 的條件，對應圖 3-14 的分類示意圖，前景是落在以 τ_{CD} 為半徑， τ_{alow} 為最低邊界的圓柱體外之白色區域，而落在圓柱體內的像素，再根據 $\tau_{\alpha 1}$ 和 $\tau_{\alpha 2}$ 兩個屬於背景的亮度失真門檻值，以及 Shadow 的亮度較暗，Highlight 的亮度較亮的特性，繼續做分類。

(3-25) 式中的 3 個門檻值，可以利用統計學習的方式自動選取，將所有 NXY (影像的大小是 $X*Y$ ，背景的張數是 N) 個取樣點的 $\hat{\alpha}_i$ 和 \hat{CD}_i ，從小到大順序的作統計(如圖 3-15)，並由特定的偵測比率(Detecting Rate)，決定其門檻值：

$$\frac{K_{low}}{NXY} = (1 - rate) \quad \frac{K_{high}}{NXY} = rate$$

$$\hat{CD}(K_{high}) = \tau_{CD} \quad \hat{\alpha}(K_{low}) = \tau_{\alpha 2} \quad \hat{\alpha}(k_{high}) = \tau_{\alpha 1} \quad (3.26)$$

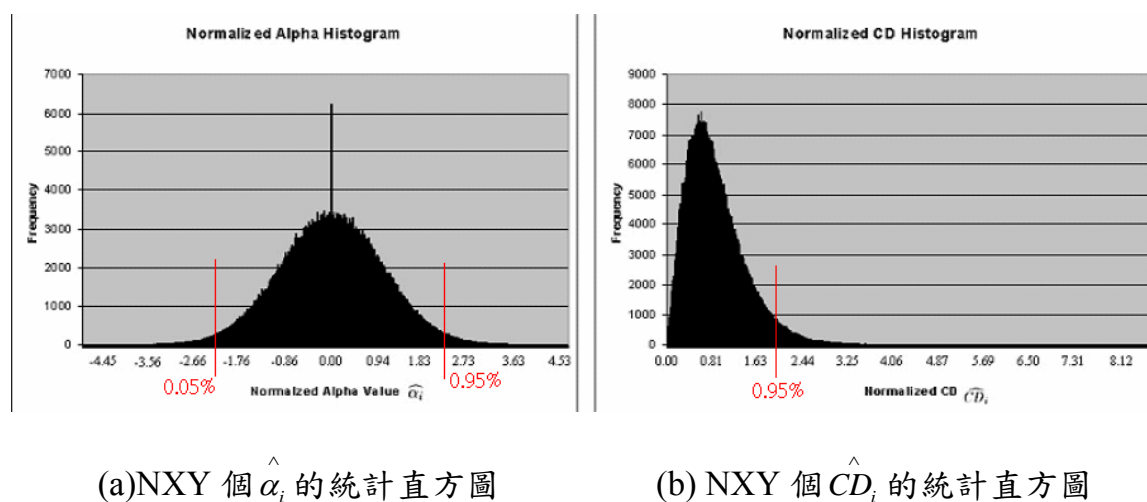


圖 3-15 由統計方式自動選取三個門檻值

3.5.2 結合 GMM 的陰影濾除方法

S_{NP} 的方法在陰影濾除上有不錯的效果，但是它有一些缺點，說明如下：

- 它在建立背景模型的部分，只取其平均值來當做參考的影像，如此方法只能針對靜態影像，如果影像中有不確定性的變化，或是背景物體的移動，則就不能適用 S_{NP} 方法。
- 一旦門檻值由背景統計得到後，所有像素不管其 E 是多少，都採用相同的門檻值，這樣的方式似乎缺少了針對不同的顏色大小，來調整門檻值的適應性，所以往往容易將較暗的前景誤判成陰影。
- 3 個門檻值可由背景的資訊統計自動選取，當背景有所變化，雖然可以將新的背景加入統計，更新門檻值，但是它卻沒有記錄前景資訊的能力。
- 更新後的門檻值適用於新的影像中所有的像素點。

由於上面的這些缺點，若採用 GMM 的背景模型，則可以獲得改善，所以我們提出類似 S_{NP} 建立 RGB 顏色模型的方式，並與 GMM 結合的陰影濾除方法。它的基本觀念是，背景模型中每一群高斯分佈，在顏色空間中是以高斯的平均向量為中心，高斯的 2.5 倍標準差為三個軸，所形成的一個橢球空間，若新的向量落在此橢球的空間內，表示它屬於背景的一部分，而落在與此橢球相切所形成的圓柱體內的向量，表示它們與背景的顏色變化不大，只有亮度上的變化，所以比背景亮的即 Highlight，比較暗的即 Shadow，其分佈的空間如圖 3-16 所示。

增加 afa_low 邊界的原因，與 S_{NP} 增加 τ_{alow} 的原因一樣，是避免暗的前景誤判成 Shadow，至於多加 afa_high 邊界，是因為在亮度大的地方，圓柱體的範圍變的很大，容易將屬於前景的誤判 Highlight，在這裡這兩個邊界，可由高斯分佈的標準差大小來決定之，後面有詳細的說明。

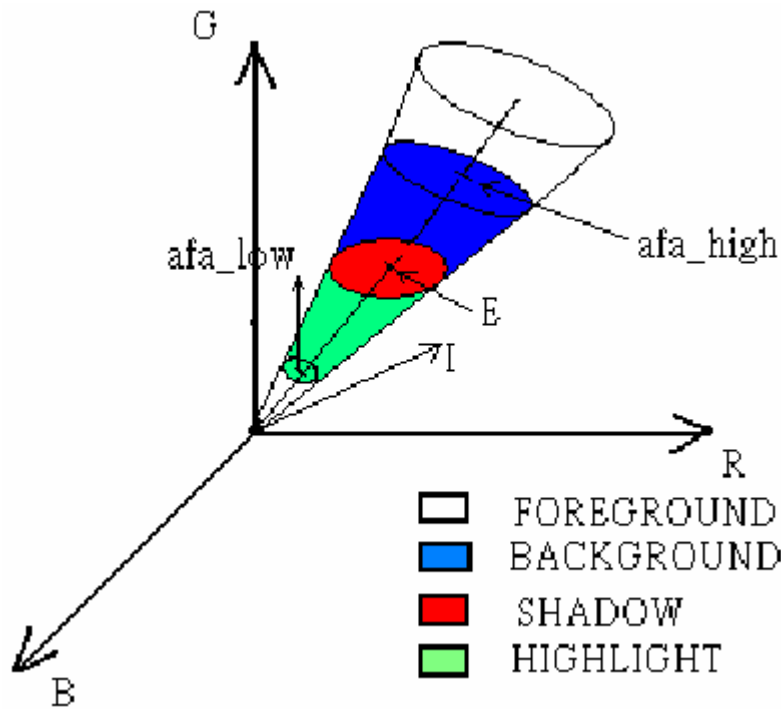


圖 3-16 在我們的陰影濾除方法中，像素分類的示意圖

因為 RGB 各自的標準差不一樣，造成圓柱曲面與橢球中心向量的夾角都不一樣，無法在三度空間中利用夾角來進行分類，所以我們將它們投影到二維的空間，利用橢圓及切線的斜率來進行分類，接下來我們以投影到(R, G)空間來做簡單的說明，示意圖如圖 3-17 所示。

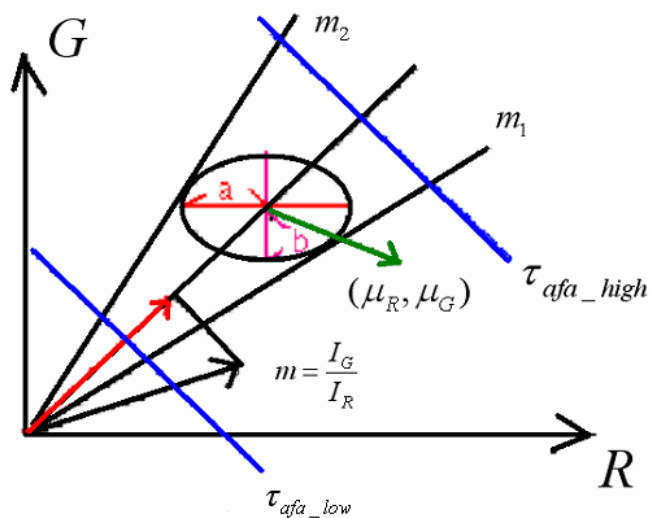


圖 3-17 二維空間中根據斜率的關係做像素的分類

以 (μ_R, μ_G) 為橢圓的中心， $a = 2.5 * \sigma_R$ ， $b = 2.5 * \sigma_G$ ， a 、 b 分別是橢圓的長短軸，可以得到橢圓的幾何方程式如下：

$$\frac{(R - \mu_R)^2}{a^2} + \frac{(G - \mu_G)^2}{b^2} = 1 \quad (3.27)$$

要得到與此橢圓相切的兩條直線，把直線方程式 $G = mR$ 帶入上式，整理後得到一元二次方程式：

$$(b^2 + a^2 m^2)R^2 + (-2b^2 \mu_R - 2a^2 m \mu_G)R + (b^2 \mu_R^2 + a^2 \mu_G^2 - a^2 b^2) = 0 \quad (3.28)$$

因為是切線方程式，橢圓與直線的交點只有一點，所以 R 的解只有一個，令 $A = (b^2 + a^2 m^2)$ 、 $B = (-2b^2 \mu_R - 2a^2 m \mu_G)$ 、 $C = (b^2 \mu_R^2 + a^2 \mu_G^2 - a^2 b^2)$ ，代入方程式 $B^2 - 4AC = 0$ 中，繼續求 m 的解，再經過整理後，得到 m 的一元二次方程式：

$$(a^2 - \mu_R^2)m^2 + (2\mu_R \mu_G)m + (b^2 - \mu_G^2) = 0 \quad (3.29)$$

與橢圓相切的直線有兩條，所以 m 的根是兩個，其值分別是 m_1 和 m_2 ：

$$m_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

$$A = a^2 - \mu_R^2 \quad B = 2\mu_R \mu_G \quad C = b^2 - \mu_G^2 \quad (3.30)$$

新的顏色向量 $I = [I_R, I_G]$ ，其斜率 $m = \frac{I_G}{I_R}$ ，若 I 相對於背景 E 是 Shadow 或 Highlight，則其斜率 m 必介於與橢圓相切的兩條切線斜率 $m_{1,2}$ 之間，但不能只做一次投影就決定之，必須再投影到 $[G,B]$ 、 $[B,R]$ ，三次的投影都符合上述的條件，才能確定此向量 I 是 E 的光影變化，否則都算是前景。確定此向量是光影的變化後，再計算投影到背景向量的長度 afa (下圖 3-17，紅色的向量長)，這個值所代表的函義就是相對背景的亮度，若 afa 小於 1 且大於 $\tau_{\text{afa_low}}$ 則是 Shadow，若 afa 大於 1 且小於 $\tau_{\text{afa_high}}$ 則是 Highlight，都不符

合則是前景。

對高斯分佈而言，若資料點是落在距離平均值很大倍數的標準差之外，得到的機率值幾乎是 0，代表它非常不屬於此分佈，利用這個觀念，若新的向量和 (μ_R, μ_G) 的距離，超過其標準差一個很大倍數，表示它屬於此分佈的機率為 0，則它必定是前景，所以我們以 (μ_R, μ_G) 為基準，往上、往下加減 N 倍標準差(N 取大於 10 的值)得到的位置，來決定 τ_{afa_high} 和 τ_{afa_low} 。下圖 3-18 是我們提出結合 GMM 之陰影濾除的完整流程圖。

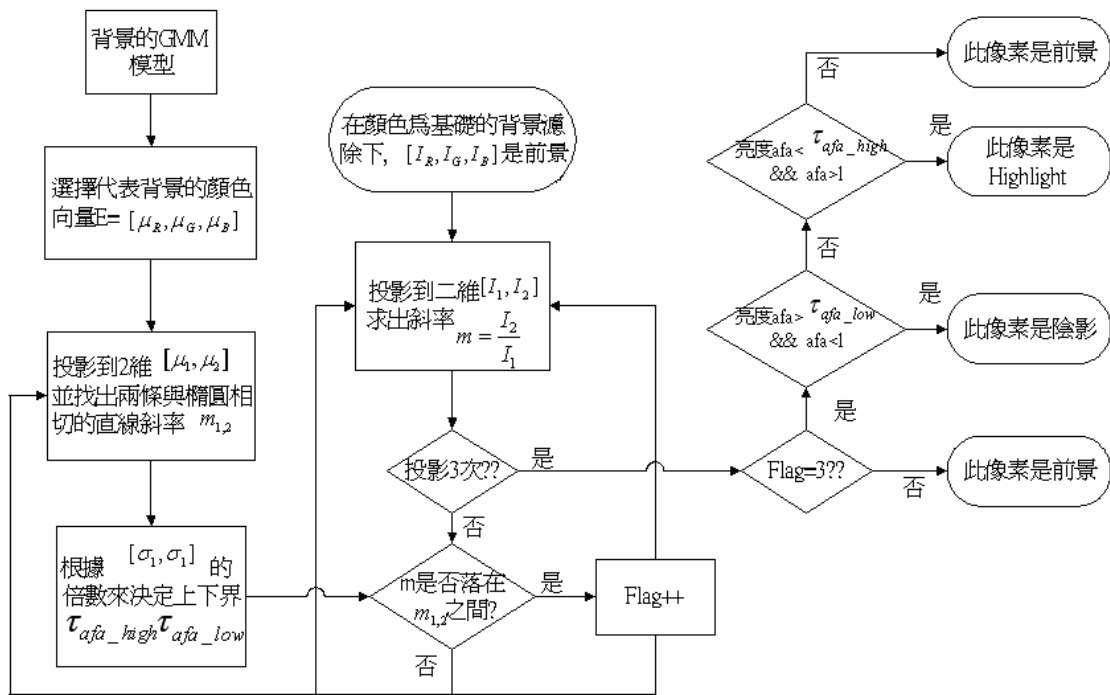


圖 3-18 陰影濾除的流程圖

3.6 短期模型(Short-Term Model)

因為我們的背景模型是多個高斯分佈，所以在做以梯度為基礎的背景模型，以及陰影濾除的判斷時，我們要選擇哪一個或是哪些個高斯來當作主要的背景 E 呢？一般來說，選擇的方式有下列兩種：

(1) 選擇 GMM 中比重最大的高斯：

比重的大小代表那群分佈出現的機率，比重越大表示它待的時間越久，拿它來當做最主要的背景判斷依據，是非常合理且直覺地；但是這種選法，很可能發生應該是 Shadow 或 Highlight 部分，被誤判成前景，我們稱這種誤判是 False Positive，其發生的原因由圖 3-19 可以清楚的看出。一開始 E 是由比重最大的分佈(1)來代表，後來新的分佈(2)併入背景中(前景的物體變成背景)，且目前出現的分佈皆是(2)(物體停留在影像中)，但是因為(2)的比重還未超過(1)，所以 E 仍選擇分佈(1)，造成目前相對於分佈(2)是 Shadow 或 Highlight 的，被分類成前景。

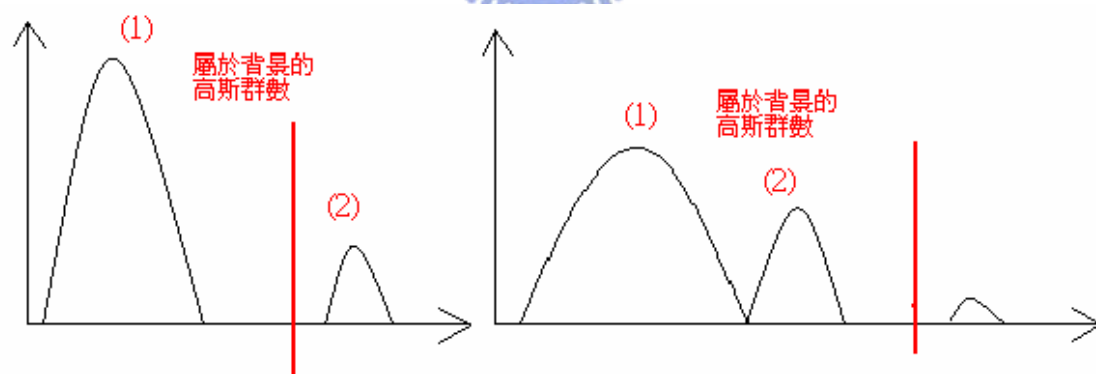


圖 3-19 False Positive 可能發生的狀況

(2) 在 GMM 中，滿足 $B = \arg \min_b (\sum_{k=1}^b w_{i,j}^{t,k} > T)$ 條件的 B 個高斯：

既然選擇單一的分佈來當做 E 是有問題的，那我們取 B 個高斯分佈來代表 E，就能解決 False Positive 的狀況。B 個分佈在三度顏色空間中，相對是 B 個橢球空間，可以找出 B 個與橢球相切的圓錐體，

一旦新的向量落在任一個圓錐體內，即判斷是 Highlight 或 Shadow，若都沒在這 B 個圓錐體內的向量，則屬於前景。但是這種選法，若發生如下圖 3-20 的情況，一個比重很小的分佈(3)併入了背景群數 B，且新的向量剛好落在分佈(3)的圓錐體內，則原本應該屬於前景的變成了 Shadow 或 Highlight，我們稱這種誤判是 False Negative。

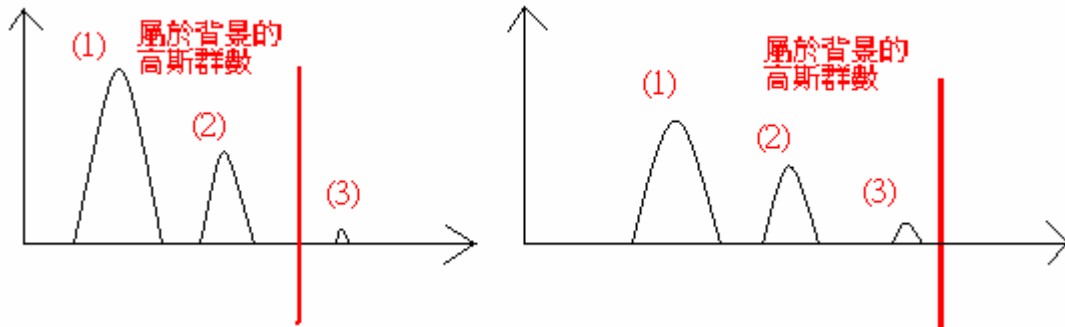


圖 3-20 False Negative 可能發生的狀況

在視覺監視系統中，前景區域是我們關心的重點，False Negative 的情況是我們不希望發生得，所以在我們的演算法中，我們採用選擇 GMM 中比重最大的分佈來當作 E，並加入了紀錄最近每個高斯分佈出現頻率的短期模型(Short-term Model)，來解決上面提到可能發生 False Positive 的狀況。

相對於短期模型的就是長期模型，這裡所謂的長期模型，就是在背景濾除的系統開始執行後，紀錄所有曾經發生過的事情，不管是前景的移動或是背景的改變，皆更新到長期模型的參數中，前面我們建立好的 GMM，就是一種長期模型；至於短期模型，顧名思義，就是去紀錄最近的一段短期時間內(在我們的演算法中，我們取最近 300 張影像)，每一群高斯分佈出現的頻率，出現的頻率越大，代表在最近的時間內，此分佈越能代表目前的背景狀況。

我們結合了這兩種模型，當背景濾除系統初始化後，兩個模型都開始記錄，在還未到達短期時間點(前 300 張影像)之前，我們皆採用長期模型中比重最大的分佈當作 E，當到達短期時間之後的每一個時間點，我們先觀察短

期模型中，是否出現頻率大於一個比例(約 80%，即出現 240 次)的分佈，如果有則以此分佈為 E；如果都沒有，表示最近的分佈是動態不固定的，我們很難由短時間的資訊來決定背景，所以仍採用長期模型中比重最大的分佈為 E。下圖 3-21 是結合了長期模型和短期模型的示意圖，去決定目前所代表的背景高斯分佈 E，並拿此分佈去做梯度濾除的模型建立與陰影濾除的判斷依據。

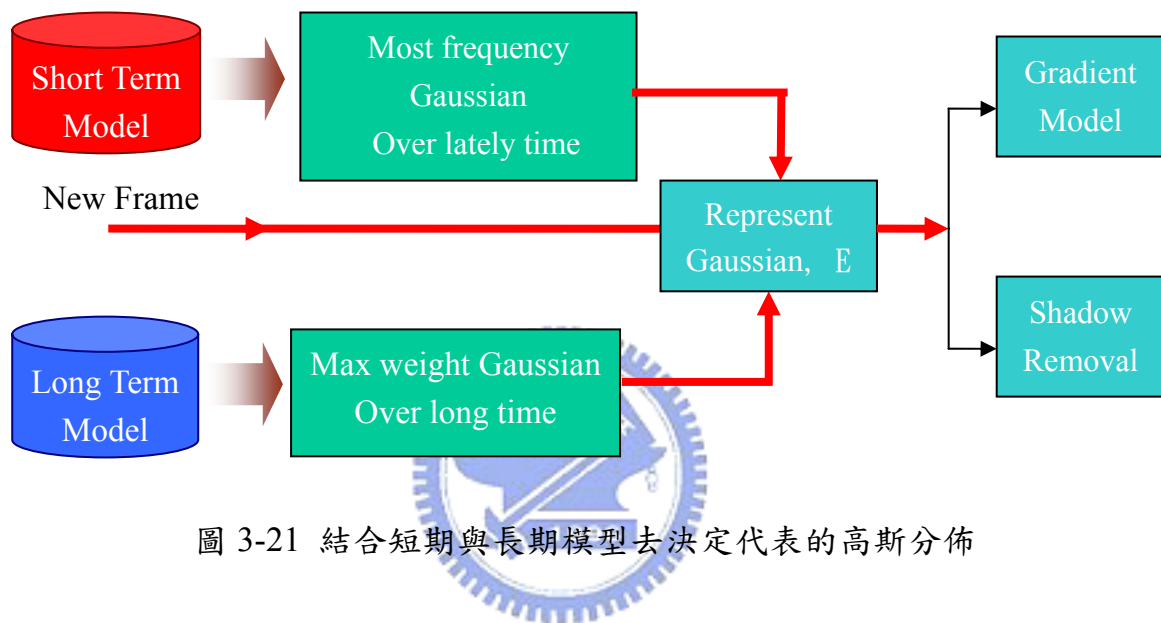


圖 3-21 結合短期與長期模型去決定代表的高斯分佈