

國立交通大學

多媒體工程研究所

碩士論文

The logo of National Tsing Hua University is a circular emblem with a gear-like border. Inside the circle, there are stylized representations of a book, a building, and a ship. The letters 'E', 'S', and 'A' are prominently displayed in the center. Below the ship, the year '1956' is visible.

紅外線影像中之前景物偵測

A Novel Method of Foreground Object Detection in Infrared Images

研究生：陳證中

指導教授：莊仁輝 教授

中華民國九十七年六月

紅外線影像中之前景物偵測

A Novel Method of Foreground Object Detection in Infrared Images

研究生：陳證中

Student : Cheng-Chung Chen

指導教授：莊仁輝

Advisor : Jen-Hui Chuang

國立交通大學

多媒體工程研究所

碩士論文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

January 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

紅外線影像中之前景物偵測

學生：陳證中

指導教授：莊仁輝 教授

國立交通大學

多媒體工程研究所

摘要

在本篇論文中，我們針對紅外線影像，提出一個新的前景物偵測方法。此方法推廣傳統的高斯混合模型，加入了位置變數以對於整張影像而非每一像素，建立數個「區域高斯模型」，因此所建立的高斯模型數量會比傳統的高斯混合模型少很多。建立起始的背景模型後，之後影像中每一像素是以一個 5×5 的鄰近區塊，來對前一張影像做區域高斯模型的比對，再以最符合的區域高斯模型做更新。實驗結果會看到在攝影機移動不大的情況下，利用本論文之方法在區分紅外線影像的前景跟背景確實能獲得較佳的結果。

A Novel Method of Foreground Object Detection in Infrared Images

Student : Cheng-Chung Chen

Advisor : Jen-Hui Chuang

Institute of Computer Science and Engineering
National Chiao Tung University

ABSTRACT

In this thesis, we propose a novel method of foreground object detection for infrared images. We generalize the Gaussian Mixture Model (GMM) to construct a new Regional Gaussian Mixture Model (RGMM), by adding two random variables of image coordinates. Since the models are built for the whole image, not for every image pixel, the number of RGMM is much smaller than that of GMM for common videos. After an initial background construction, the RGMMs are updated by examining the existence of previous RGMMs in a 5×5 neighborhood for each image pixel, followed by the identification of the best-fit model which is then used in the update process. Experimental results show that better separation of foreground object from background can be achieved by using RGMM for infrared images obtained by a camera with small movements.

致謝

本篇論文得以完成，中間歷經許多題材蒐集、研究工具的選定、實驗方法的摸索、以及論文內容的撰寫與修正，深深體會一個研究從無到有，是需要許多的心力去研究。這兩年來在研究與學習的過程中有很多的收穫，首先要感謝指導教授莊仁輝博士，他在研究上認真的態度給予我很大的啟發，並且讓我的研究生涯中時時刻刻都給予寶貴的意見及指導，也感謝楊谷洋教授、顏嗣鈞教授與雷欽隆教授在口試時針對我的論文給予寶貴的意見，並對論文不足之處提出改善的方法，讓本論文得以更加完善。

在這兩年的研究生活中，感謝實驗室博士班學長林泓宏與羅國華在學業及研究上的指導，讓我的論文有了方向。此外，感謝諾白、寧玄、俊宏、以及忠衛，在一起讀碩士班的時候，彼此互相鼓勵以及幫助，這一切得以讓我進步及成長。最後感謝我的家人，因為有他們在背後的支持，我才能取得碩士學位，感謝所有曾經幫助過我的人。

目錄

致謝	III
目錄	IV
圖目錄	V
一、 序論.....	1
1.1 研究動機.....	1
1.2 相關研究.....	2
1.3 系統流程.....	4
1.4 各章簡介.....	6
二、 背景模型的建立	7
2.1 紅外線影像顯示.....	7
2.2 基於高斯混合模型之背景分析.....	9
2.3 基於連通區域標記法之背景分析.....	12
2.4 模型完整性.....	14
三、 連續影像的處理	17
3.1 基於連通區域模型的區域比較.....	18
3.2 基於影像強度的區域比較.....	19
3.3 區域高斯模型的更新.....	20
四、 實驗結果與探討	22
4.1 靜態連續影像的偵測結果.....	22
4.2 動態連續影像的偵測結果.....	30
五、 結論與未來展望	38
參考文獻	39

圖目錄

圖 1-1	高斯混合模型示意圖	3
圖 1-2	(a) 起始場景, (b) 最終背景。	5
圖 1-3	系統流程圖。	5
圖 2-1	紅外線影像轉換灰階圖	9
圖 2-2	高斯建立流程圖。	10
圖 2-3	掃描方法的差異圖	12
圖 2-4	連接的方法圖。	13
圖 2-5	物體連接標示圖	13
圖 2-6	連接區域標記圖 (8-neighbor) 。	14
圖 2-7	填回剩餘像素流程圖。	15
圖 2-8	剩餘的像素填回	16
圖 3-1	連續影像的處理流程圖。	17
圖 3-2	5×5 區域的比較示意圖。	18
圖 3-3	區域比較後的結果圖。	19
圖 3-4	強度區域比較後結果。	19
圖 3-5	更新示意圖。	20
圖 4-1	複雜背景連續影像偵測結果。	22
圖 4-2	複雜背景連續影像偵測結果。(續)	23
圖 4-3	單純背景連續影像偵測結果。	23
圖 4-4	單純背景連續影像偵測結果。(續)	24
圖 4-5	單純背景連續影像偵測結果。(續)	25
圖 4-6	靜態連續影像比較。	25
圖 4-7	靜態連續影像比較。(續)	26
圖 4-8	靜態連續影像比較。(續)	27
圖 4-9	靜態連續影像比較。(續)	28
圖 4-10	靜態連續影像比較。(續)	29
圖 4-11	動態連續影像偵測結果。	30
圖 4-12	動態連續影像偵測結果。(續)	31
圖 4-13	動態連續影像偵測結果。(續)	32
圖 4-14	動態連續影像比較。	33
圖 4-15	動態連續影像比較。(續)	34
圖 4-16	動態連續影像比較。(續)	35
圖 4-17	動態連續影像比較。(續)	36
圖 4-18	動態連續影像比較。(續)	37

一、序論

1.1 研究動機

有關於影像偵測的研究在學術領域中已經有很多，其中有數種方法，都有不錯的結果，也因為有這些方法可供利用，所衍生出來的應用方面也就越來越廣泛。目前已經有相當多利用物體偵測的應用，例如安全監控、人體行為分析、保全物品監測、智慧型機器人等等，尤其在保全監控方面的應用更是熱門，利用電腦視覺來協助保全，不僅節省人力成本，也降低了人為疏失的因素。

而無論是哪個方面的應用，在電腦視覺方面都會利用到攝影機，而攝影機的種類包括 DV 攝影機、PTZ 攝影機、全景攝影機、以及紅外線攝影機等等。大多數的電腦視覺應用所使用的攝影機，都是可見光的攝影機，較少著重於紅外線部份的研究，因為可以量測溫度資訊的紅外線攝影機價格較高並非隨手可得，所以也比較少被拿來做研究。

但是夜晚的犯罪率比白天高，如果只利用可見光攝影機，來做安全監控，在夜晚會因光線不夠充足，所得到物體的資訊會不足夠，導致在夜晚所拍攝的場景，會發生物體無法辨認的情況出現，如此便會有安全上的漏洞。在夜晚的時候紅外線能偵測到物體的溫度而不受亮度的影響，假如能利用紅外線攝影機，便可以彌補可見光攝影機的缺點，使得紅外線攝影機能進行在夜晚的監控，因此更顯出基於紅外線之電腦視覺的重要。

在本論文中我們用紅外線攝影機拍攝室內的場景，當作實驗環境，然後利用我們所拍攝影片的第一張影像建立背景模型。背景模型建立的方法，是參考高斯混合模型的觀念，我們是希望能藉由建立良好的背景模型，來做前景跟背景的區分，但是高斯混合模型只適用於固定式的攝影機，所以我們做了一些不同於高斯

混合模型的改變，希望能改善高斯混合模型的方法，並可套用在可移動的攝影機上。

由於紅外線影像資料格式常已大於十個位元來表示一個像素，在紅外線影像的顯示上，必須要用到一些灰階轉換的方法，才能轉變成我們能看見的影像。一般常見的轉換的方法有 Min-max contrast stretch、Histogram equalization (HE)[1]、Binary occupied histogram projection (BOHP) [2]，然而這三個轉換方法都有其優點跟缺點，為了能讓影像中的一些細節能更好的顯示，我們是採用本實驗室所發展的一個影像對比強化的方法[3]，讓我們的紅外線影像顯示上較可以看的出物體明暗層次的分別。

1.2 相關研究

在電腦視覺中物體的偵測的方法多半需要建立良好的背景模型，才能有好的結果。在物體即時監控上，最常被使用的是高斯混合模型 (Gaussian mixture model, GMM) [4][5]之技術，圖 1-1 為高斯混合模型的示意圖，這方法利用數學機率的概念去對彩色影像每一個像素建立數個高斯模型 ($\mu_1, \sigma_1, \mu_2, \sigma_2, \dots$)，之後的每一張影像都會與已經建立的模型去做比對，如果像素值有在 2.5σ 範圍內的話，就會更新高斯模型，否則就會新增模型，如果模型數量已經滿了，就會把出現頻率最小的高斯模型予以取代。最後取一個物體偵測門檻值，取出最有可能的背景模型後，前景也跟著被切割出來。雖然物體偵測最後的結果看起來很穩定，但缺點是前景物如果存在模型中很久的話，有可能會被視為背景，除此之外高斯混合模型不適用於移動式的攝影機。

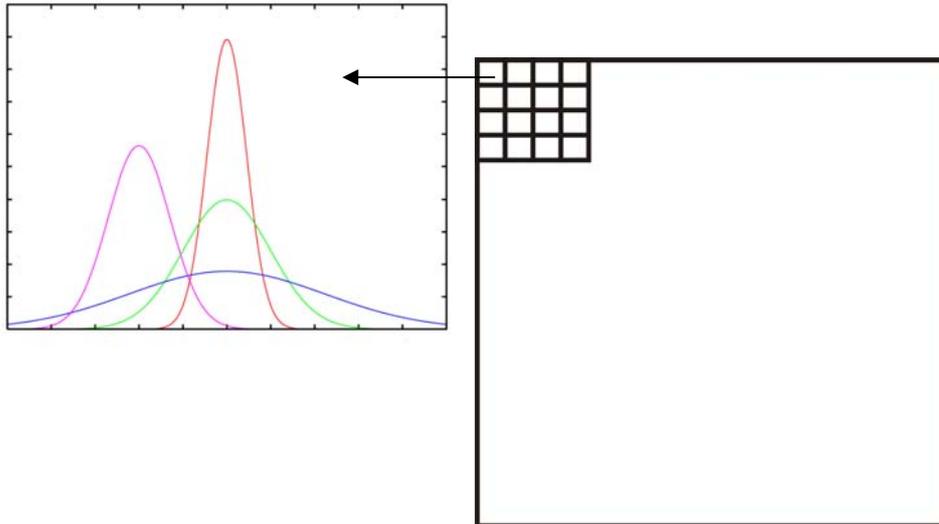


圖 1-1 高斯混合模型示意圖

文獻[6]利用動態偵測及膚色分析實現人臉的偵測。一開始會先把兩張連續影像相減，並且設定一個閾值，當兩影像的差值大於此閾值為 1，否則為 0。得到差值的影像後，再利用膚色特徵在群集 (grouping) 的範圍裡面判斷此區域是否為人臉，雖然此方法結合了動態及膚色可以快速找出人臉，不過一但光線是在不理想的情況下，造成膚色有所偏差的話，此方法就有可能會失敗。為了得到良好的差值影像，同樣的此方法也只能適用於固定式的攝影機。

文獻[7]中利用背景影像跟他其影像相減，得到前景的部份，藉由形態學 (morphology) 提取前景的邊界 (boundary)，再利用連通區域 (connected component) 對連結的邊界群分別給予標記 (labeling) [8]。接著配合透視轉換 (perspective transformation) 判斷邊界的形狀是否跟人體的形狀有類似，就可以判斷是否是我們想要的前景物體。

文獻[9]提出的方法，考慮到使用移動式攝影機來做偵測，利用高斯混合模型與馬可夫隨機域，來達到攝影機移動下物體能準確追蹤。首先他們假設可以知道攝影機的移動量，知道移動量便能推斷相鄰影像的前景物是否一樣，而利用高斯混合模型對所有的場景建立背景模型，之後就能對前景做切割，兩個方法互補來達到移動式攝影機的物體偵測。

文獻[10]利用三種方法來做移動物體的偵測，首先利用高斯混合模型找出前

景，再利用光學流動方法判斷物體的移動是否一致，最後利用時間差的方法，利用相鄰兩張影像前景與背景的移動差異不同，找出移動中的物體。結合三種方法之目的，則為提升物體偵測的準確率。

在紅外線的偵測方面，文獻[11][12]利用一個 $4*4*3$ 的三維空間，這三維是時間 T 跟空間 X 及 Y 位置資訊，來對連續的影像做處理。在這個區域所擁有的三維資料必須先做 PCA 處理，把資料量減少，以降低計算量。這些資料就會當成一個向量，每隔一段時間便會把同一區塊位置中的多個向量，計算出物體的移動量，如果移動距離很短就可能是背景，距離很長的可能為前景，但也有可能距離都差距不大，造成前景跟背景很難分辨。但由於紅外線影像的雜訊比一般可見光影像多，使用這個方法可以不受雜訊的影響。

1.3 系統流程

在本論文中的系統，主要的場景是在室內環境中進行拍攝，由於我們另外會用到移動式攝影機的影像，所以場景會隨著攝影機變動，如圖 1-2 所示，前景物體會從圖 1-2 (a) 所示場景走到如 (b) 所示之靠近門口區域。利用拍攝出來的紅外線影像，我們會先對沒有前景物的起始影像建立背景模型，之後的每一張紅外線影像都會在 $5*5$ 的範圍內對建立好的背景模型去做比對動作，來找出前景物體，最後會做更新把前一張的模型跟現在的模型依比例結合，整個系統的流程如圖 1-3 所示。

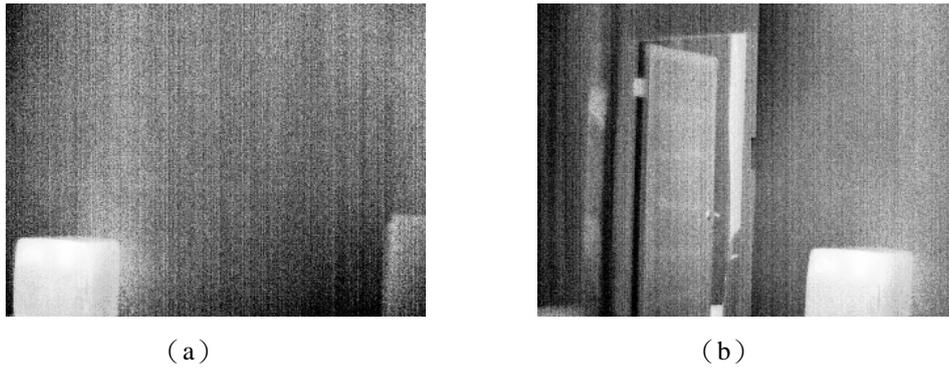


圖 1-2 (a) 起始場景, (b) 最終背景。

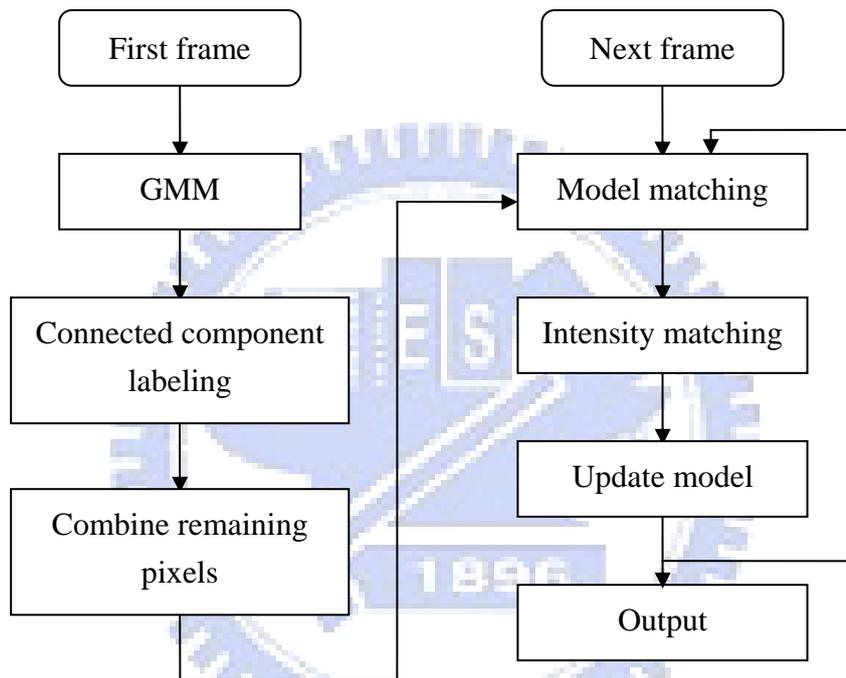


圖 1-3 系統流程圖。

圖 1-3 主要分為兩個部份，左邊是建立背景模型的部份，右邊則說明如何進行前景物的區分。背景模型的建立分為三個步驟，使用改良高斯混合模型的方法來建立背景模型之後，再利用連通區域標記法，對每一個物體都標記自己的模型，再把一些原本屬於同一個模型但並未連接的像素予以補回。前景的區分部份，則是利用一個 5×5 區塊，分析現在的像素是歸屬於之前畫面對應的 5×5 內鄰近像素的哪一個模型，背景模型找出來後，前景物也就能跟著區分出來。

1.4 各章簡介

本論文全部共分為五章，除了本章序論之外，第二章將會介紹起始的背景影像的建立，利用不同的方法，期望能將背景物體各自擁有自己的模型。第三章則是介紹對已經建立好的背景，我們如何利用它來處理連續的紅外線影像，最後得到處理後的結果，是希望能將前景與背景有很大的區分。第四章就是動態攝影機跟靜態攝影機所拍的紅外線影像的實驗結果。第五章為結論並說明未來研究的方向。



二、背景模型的建立

在本章中，我們先對紅外線攝影機所拍攝的影像作顯示處理，而後作是背景模型建立。在背景的建立中，我們會做一些不同於高斯混合模型的改變。由於這些模型為整張影像而非各個像素所有，我們將模型的數量比起傳統的高斯模型來的少，而大幅節省記憶體的空間，並把高斯模型推廣到三維，增加位置變數來達到三維的高斯模型。

2.1 紅外線影像顯示

一開始紅外線的顯示方面，我們雖然有紅外線攝影機附加的軟體，可以拿來作為顯示用，但環境的溫度容易受到高溫物體的影響，使得在顯示紅外線影像轉換灰階上，環境的一些物體會因為由高位元的強度（ $0\sim 2^{14}$ ）轉到低位元的灰階（ $0\sim 255$ ）會有同樣的灰階值，讓整個影像中只有高溫物體能比較有明確的顯示，其餘的物體有些被忽略，導致有些模糊不清，基於這個因素我們說明增強影像對比的方法。

現今的紅外線影像對比強化方法，常見的有三個：

1. Min-Max Contrast stretch
2. Histogram Equalization (HE)
3. Binary Occupied Histogram Projection (BOHP)

第一個方法是直接把強度分佈範圍，直接對應到灰階值 $0\sim 255$ 的範圍，這是最容易執行的方法。第二個方法則是把強度值建立一個直方圖，每個強度值都會有自己的數量，最後利用直方圖等化，從第一個強度開始累加數量並乘以自己佔有的

比例，直到最後累加完會是 1，並把這些值乘以 255 就是各個強度的灰階值。第三個方法是先建立強度的直方圖，計算出直方圖強度數量不為 0 的強度分佈總數，之後直方圖中每個強度乘以 255 除以強度分佈的總數，即為灰階值。

不管使用哪種方法都有他的優點跟缺點，第一個方法雖然最容易執行，但轉換的結果是最差的，第二個轉換的效果比第一個好，但是必須對影像中的所有強度建立直方圖等化，執行的效率會比第一個差一點，最後第三個方法，則是利用第二個方法加以改善，但缺點是若有一個物體其強度的分佈範圍很大的時候，會造成其他物體被壓縮的很嚴重，甚至造成轉換成灰階影像上難以分辨的情況發生。

由於以上的三個方法都有其優點跟缺點，為了能讓影像的一些細節得到更好的顯示，所以我們採用一個[3]所提出的方法，來加強紅外線影像的對比。強化的方法如下列步驟：

1. 針對紅外線影像的強度建立直方圖，每一個強度會紀錄自己的數量。
2. 計算直方圖強度分佈的個數，只有數量大於 0 才會被加進去。
3. 然後會取一個門檻值，超過此門檻的都會被設成此值。
4. 做直方圖等化把強度直方圖低的到高的開始累加，只要有遇到有強度分佈的就會把它加進來並且將該強度佔有的比例乘以 255，直到最後一個強度分佈的欄位等於 255。
5. 建完直方圖等化後，便可以利用來做灰階轉換的對照表。
6. 最後原始的紅外線影像在轉換的時候，便會參考對照表，轉換成對應的灰階。

根據以上的步驟，我們去做灰階轉換，最後的結果如圖 2-1 所示，很明顯的一些低溫的物體可以被對比強化的方法加以顯示出來，但高溫的物體的一些小細節也會被強化。



圖 2-1 紅外線影像轉換灰階圖，(a)一般紅外線轉換灰階方法，(b)紅外線影像對比強化方法。

2.2 基於高斯混合模型之背景分析

由於常見的高斯混合模型[4][5]是利用數學中的常態分佈方法，所以我們必須要計算模型的平均值跟變異量，首先高斯混合模型中每個像素都要建出自己的模型，每個模型都有自己的平均值跟變異量，之後每一個影像中的每一個像素會對應到前一張影像的像素，去做模型的比較，如果符合就更新模型，如果不符合就會新增，之後取出前幾個最穩定模型，來當作背景的背景模型，這樣便可以找出前景物體。

上述高斯混合模型建出來的模型數量需要很大的記憶體空間，所以我們希望能減少模型數量來減少計算量。我們的目的是在影像中各個物體都會有它自己的模型，因此，模型的數量是以物體為單位。此外，我們考慮到紅外線攝影機是會移動的情況下，高斯混合模型並不完全適用動態攝影機，利用我們的方法期望能解決場景移動下物體偵測的問題。建立高斯模型的流程圖如下圖 2-2 所示。

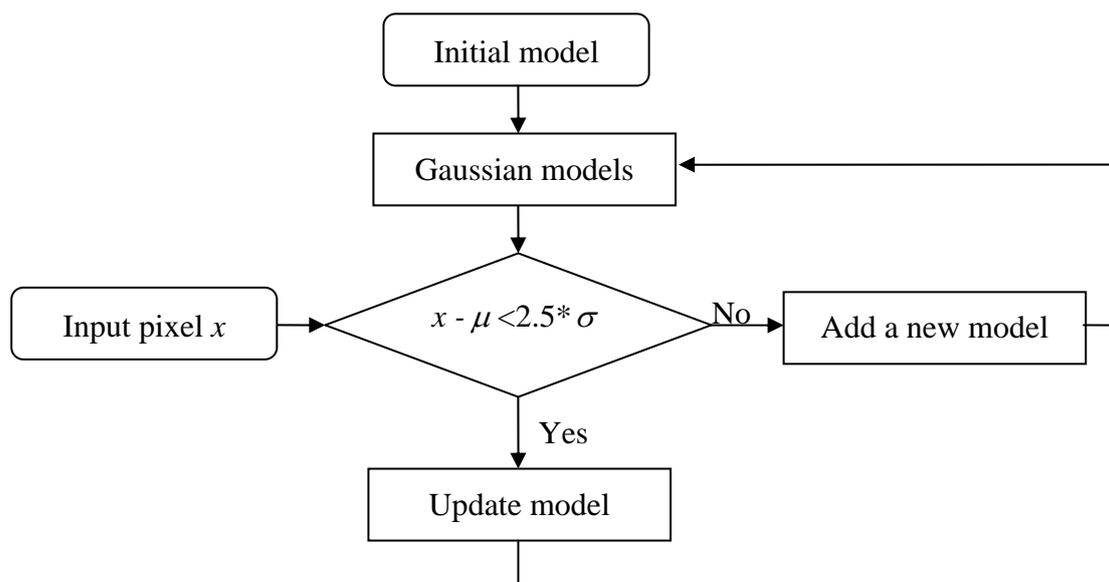


圖 2-2 高斯建立流程圖。

一開始我們先用起始影像中左上特定區域的統計當作起始的高斯模型計算，計算出起始高斯模型的平均值跟變異量 (2.1)。計算完起始模型後，之後起始影像中每個像素都會對已有的模型去做比較，看是否要新增或更新 (2.2)：

$$\mu_0 = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.1)$$

$$\sigma_0^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_0)^2$$

$$T = \begin{cases} 1, & \text{if } (x_i - \mu_i) < 2.5 \times \sigma_i \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

其中 x_i 為現在的影響中某像素的強度， μ_i 跟 σ_i 是第 i 個高斯模型，如果 $T=1$ 表示該像素符合第 i 個高斯模型，則此像素便可以加入第 i 個高斯，並更新平均值跟變異量：

$$\mu_{i+1} = \frac{1}{t+1} (\mu_i * t + x_{i+1}) \quad (2.3)$$

$$\sigma_{i+1}^2 = \frac{1}{t+1} [t * (\sigma_i^2 + \mu_i^2) + x_{i+1}^2] - \mu_{i+1}^2$$

其中 t 為某一個模型從第一個像素累加到第 t 個像素的總數，使用 (2.3) 更新公式，可使得高斯模型在逐點更新的過程中，估計出精確的平均值跟變異量。又如

果有發生某個像素符合多個模型的話，我們會用公式 (2.4) 來判斷此像素顏色與哪一個高斯模型最接近，利用公式 (2.4) 計算的結果，一定會有像素跟某一個模型計算後得到的數值是最小的，所以此像素一定會被歸為某一個模型。

$$\frac{x_i - \mu_i}{\sigma_i} \quad (2.4)$$

如果 $T=0$ 表示沒有符合已有的高斯模型，就必須新增一個高斯模型 (2.5)，模型的平均值我們會用新進來像素的強度值當作平均值，而新模型的變異量我們則設為一初使的變異量 σ_0^2 ，當作新增模型變異量的起始值，而此 σ_0^2 的設定，則可由 (2.1) 得出。作完上述步驟，整張影像每一個像素都會有自己的模型，但是我們發現掃描紅外線影像方法的不同，會造成建自高斯模型的結果不同，如下圖 2-3 所示。

圖 2-3、2-4 的掃描方法不同，會影響模型數量以及模型的分佈，最後我們只取圖 2-3 (b) 的結果，因為它的模型數量較少，而且執行時也比較簡單。但是這樣的結果，並不如我們所預期的結果，可以看到不同的物體會同樣的模型出現，為了要讓不同的物體，能有不同的模型，我們進一步考慮區域模型在影像空間中連接性的關係，故我們利用連通區域標記法來做進一步的模型分割。

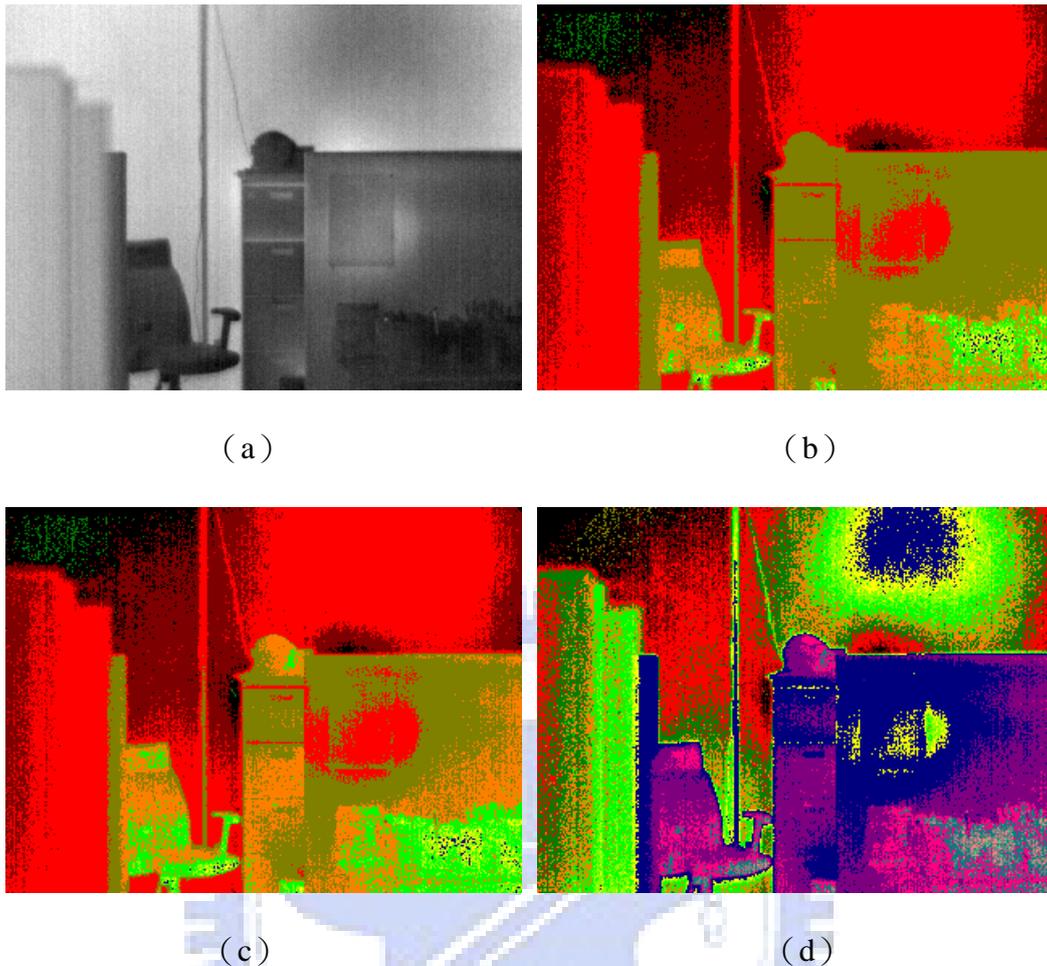


圖 2-3 掃描方法的差異圖 (a) 原始影像，(b) 由左至右一系列一列掃描，(c) 偶數列由左至右奇數列由右至左，(d) 斜向的掃描。

2.3 基於連通區域標記法之背景分析

雖然我們使用高斯混合模型對第一張紅外線影像建出背景模型，結果並不能完全把不同物體完全區分開來，相隔很遠的物體卻會有同一個高斯模型，與我們想要的不同物體都有各自的高斯模型不太符合，所以便想應用區域連接性的方法，希望把不同的物體歸為不同的模型，在此我們就選擇了連通區域標記法。連通區域標記法是演算法中的一種應用，基本上是利用連通單元的方法把相鄰且相似的像素連成一個區域，而連接的方法分為兩個種類 4-neighbor(假設一點座標 (i, j) ，其鄰近的座標必為 $(i-1, j)$ 、 $(i, j-1)$ 、 $(i, j+1)$ 、 $(i+1, j)$) 或 8-neighbor

(假設一點座標 (i, j) ，其鄰近的座標必為 $(i-1, j-1)$ 、 $(i-1, j)$ 、 $(i-1, j+1)$ 、 $(i, j-1)$ 、 $(i, j+1)$ 、 $(i+1, j-1)$ 、 $(i+1, j)$ 、 $(i+1, j+1)$)，如下圖 2-4 所示。

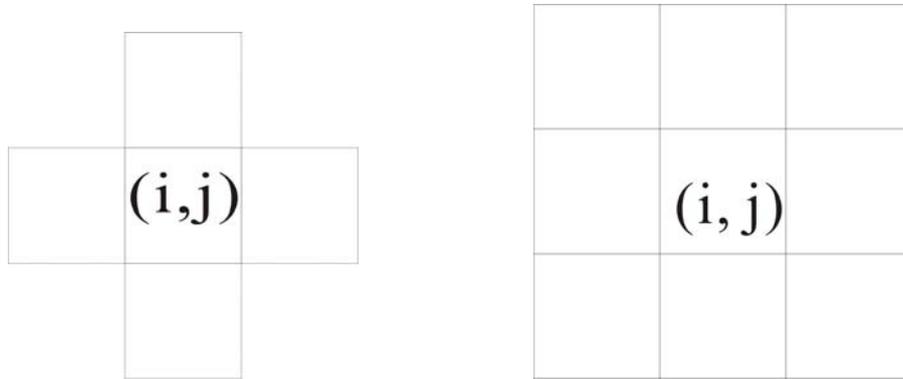


圖 2-4 連接的方法圖。

我們會從圖上任一點開始，去對鄰近的像素看是否有接近的模型，是就會將之連接，直到遇到不同的模型為止，每個連接的區域都會給予一個標籤，到最後每個連接物體都會有不同的標籤，如圖 2-5 所示，每個獨立的連接區域都給予一個標籤，依此反覆做到整張影像所有的像素都已經掃描過，並且每個物體都有標籤為止。

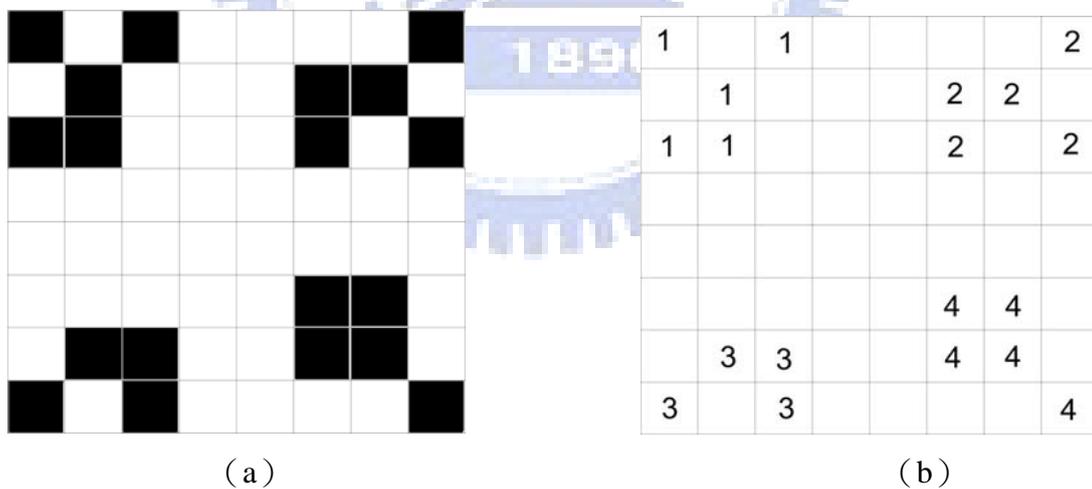


圖 2-5 物體連接標示圖，(a) 二質化影像，(b) 連通區域標記圖。

依據 8-neighbor 的連接方式，我們可以得到如圖 2-6 所示的結果，其中大塊獨立的區域（例：大於 70 個像素），以不同的顏色表示，而細碎的區域（例：小於 70 個像素），則以白色代表，而每個區域皆計算出一個區域高斯模型。

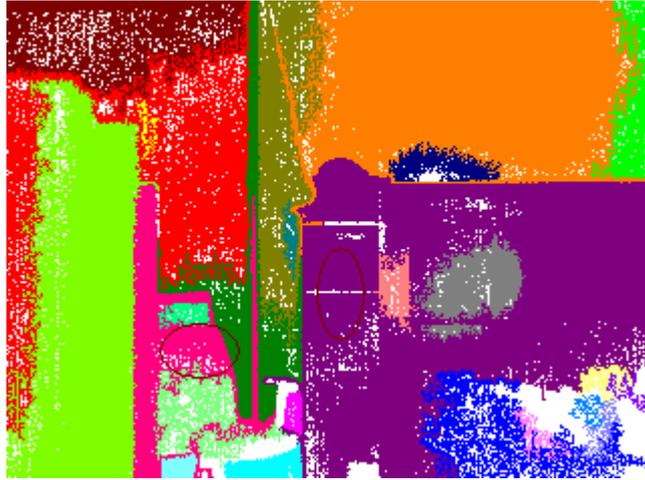


圖 2-6 連接區域標記圖 (8-neighbor)。

在圖 2-6 中有兩個圓框，一個框住椅子，一個框住櫃子，原先在圖 2-3 (b) 中它們會有同一個高斯模型，在做了連通區域標記法後，看的出已經被分為兩個不同的模型。關於區域高斯模型的計算，對於每個區域我們都會去計算一個三維的高斯模型，這三維高斯模型包含了影像中 X 、 Y 位置及強度的隨機變數， X 跟 Y 的高斯模型計算跟強度的高斯模型計算一樣是利用公式 (2.1)。

2.4 模型完整性

2.3 節我們雖然做了連通區域標記法，使得不同的區域，會有不同的模型，但仍有一些白色的像素可能跟某些區域高斯模型位置接近，而且原先在建立高斯模型後也可能是同一個高斯模型，卻因為連接條件的不滿足而被捨棄，所以我們會做一個填回的動作，把白色的像素補回區域高斯模型。整個填回的動作流程表如下圖 2-7 所示。

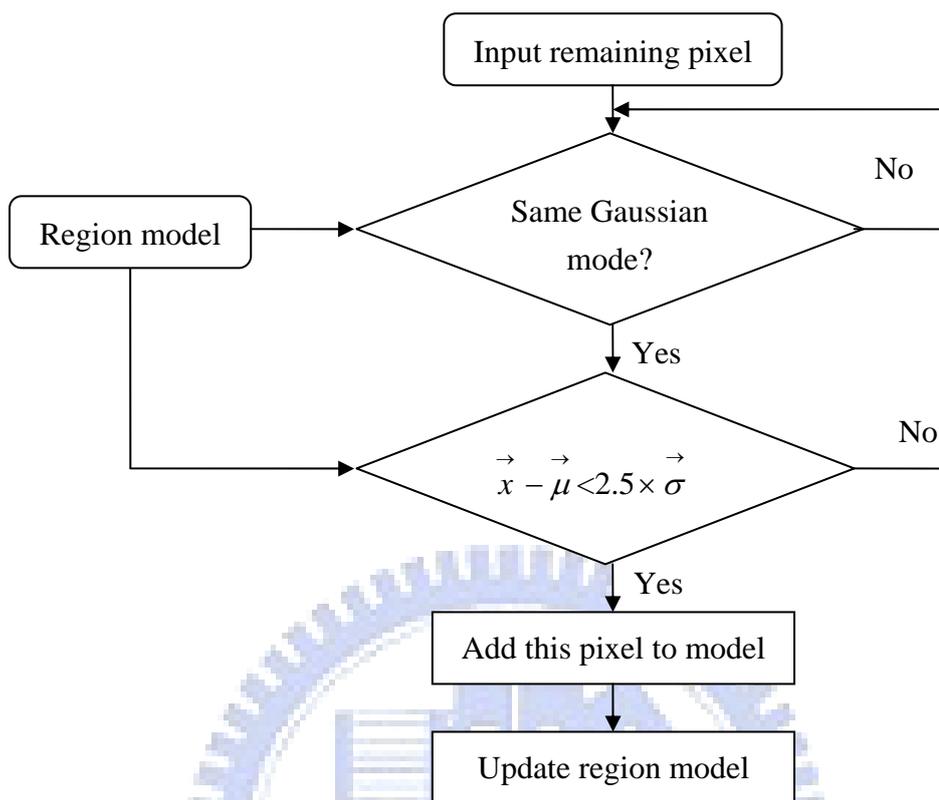
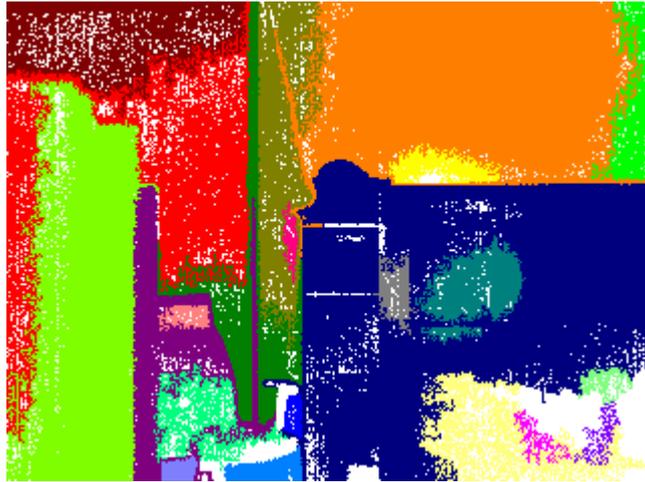
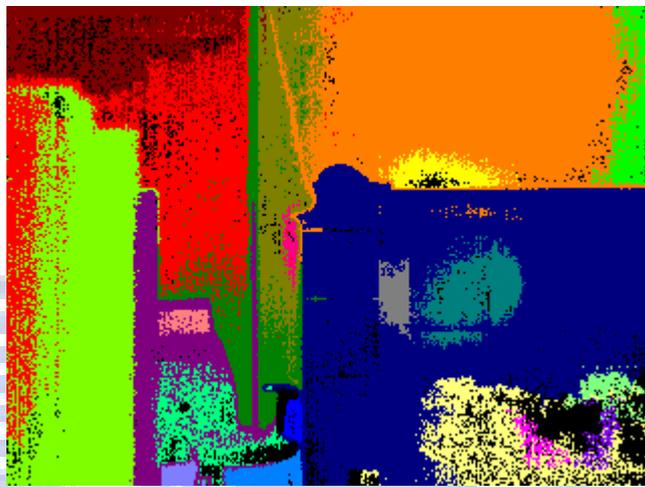


圖 2-7 填回剩餘像素流程圖。

首先在填回的動作裡我們會先找出某一個白色的像素，找出是否有在原先建立高斯模型時是和某一個區域高斯模型為同一個高斯模型，然後會用到公式 (2.2) 去比較 X 跟 Y 的連通區域模型是否有符合 2.5σ 範圍內，比較完位置後再來就是比較強度的連通區域模型是否有符合 2.5σ 範圍內，都符合的話就會將這個像素歸入這個區域高斯模型，並且用公式 (2.3) 來更新 X 、 Y 及強度三個變數的連通區域模型。如果有發生某一個剩餘的像素有符合多個區域的時候，我們會用公式 (2.5) 來找出最接近的，依據公式 (2.5) 的計算，每一個模型都會計算出一個值，只要找到最小的值就會該像素會歸為此模型，所以某一個的像素也必定會被歸到一個連通區域，最後還是會有白色的像素並沒有符合上述的條件時，我們會再給予另一個顏色，在圖 2-8 我們是給予黑色。



(a)



(b)

圖 2-8 剩餘的像素填回，(a) 未做填回前的結果，(b) 做完填回的結果。

做完填回的動作，背景模型的建立到此也就大致完成，從圖 2-8 結果可以看出模型的分佈會比單純用高斯模型建立的模型分佈來的好，只單用高斯模型建立背景，整張影像的模型分佈是比較雜亂，而且也離我們想要的不同物體有不同的模型，也有些差距，最後我們加了一些方法，讓整張的背景模型變得更好，也比較接近我們想要的結果。

三、連續影像的處理

在第二章我們已經先針對一開始的影像，來建立背景模型，接下來的我們引進區域高斯模型遞移的概念，對連續的影像來做處理。我們會利用一個 5×5 的區塊，去對前一張來做一維高斯模型的比較，而不是像常見的高斯混合模型，以個別獨立的像素作處理。使用 5×5 區塊的目的就是為了能在攝影機移動時，假設影像移動速度沒有太快，我們還是從 5×5 範圍中找出有比較符合的區域高斯模型，同時我們希望新進來的像素只找其相鄰的模型，這樣才能減少處理不相干的模型。本章整個流程如圖 3-1 所示。

第一節我們把下一張影像的某一個像素，去找出在 5×5 範圍內是否有相似的區域高斯模型以便給予相同的區域高斯模型標籤，再來第二節則是把未歸屬於任何一個區域高斯模型的像素，能去找出 5×5 範圍內，是否有強度接近之背景像素，用來降低雜訊的影響，最後所有的像素都比對完後，我們會做模型的更新。

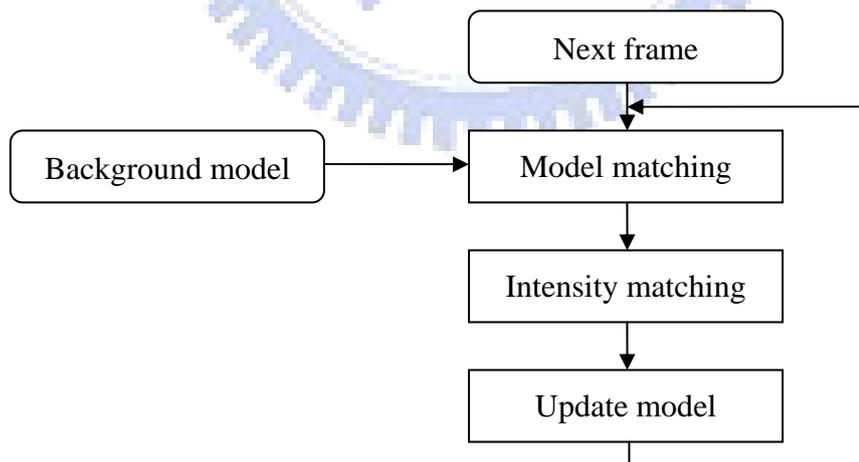


圖 3-1 連續影像的處理流程圖。

3.1 基於連通區域模型的區域比較

處理連續影像處理的第一步，就是要對前一張影像作模型的比較，如圖 3-2 所示。我們雖然建立三維的區域高斯模型，但因為位置的變數與影像強度的變數多不為線性獨立，所以在模型的比較上，會比較不適用，所以我們是利用 5×5 的區塊，在第 $T+1$ 張影像中的某一個像素，對第 T 張影像中 5×5 範圍內針對強度的區域高斯模型來做比較，以公式 (2.2) 來做比對，如果該像素有符合就歸屬為某一個區域高斯模型，之後第 $T+1$ 張影像中每個像素都會做此動作。

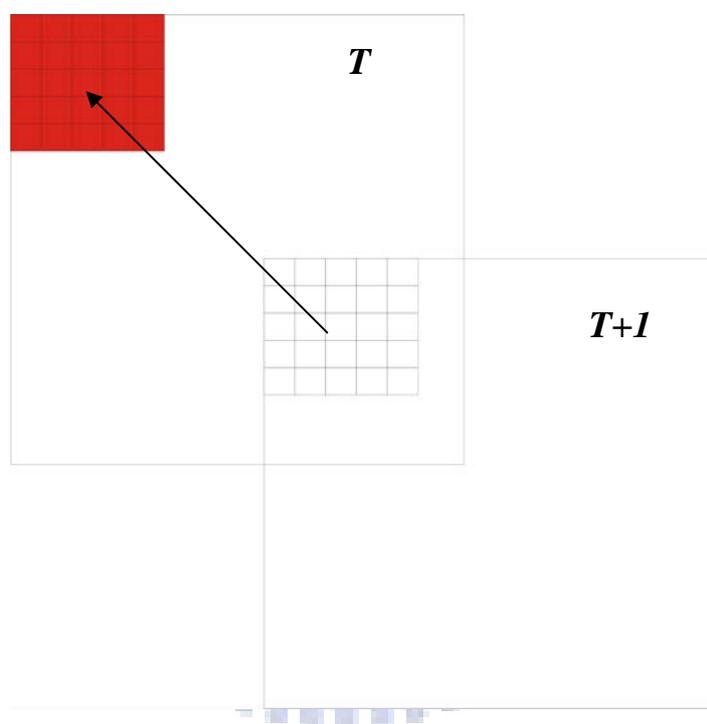


圖 3-2 5×5 區域的比較示意圖。

作完上述的步驟後，第 $T+1$ 張影像大部分的像素，都會歸屬到某一個區域高斯模型。其結果我們可以從圖 3-3 看到，大致上的物體都有對應的某一個模型，但還是有一些像素無法對應到，此類像素皆以白色表示，其沒有對應的原因乃是可能本身為一個雜訊點，或者是場景因為攝影機移動有新進來的物體，造成像素並不歸屬任何一個區域高斯模型，所以我們接下來就是想把沒對應到的像素，再做強度的比較，使雜訊的影響變少。

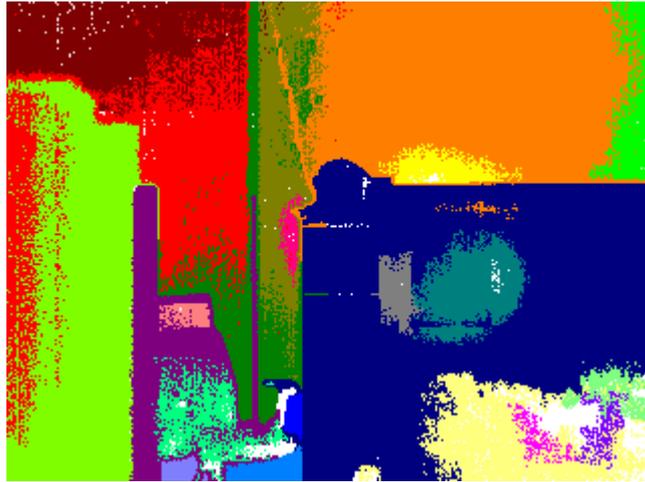


圖 3-3 區域比較後的結果圖。

3.2 基於影像強度的區域比較

此節的作法會先等區域高斯模型的區域比較後，如果沒有找到一個符合的區域高斯模型，我們才會做影像強度的比較，同樣的我們還是會在 5×5 區域內做比較，只要在 5×5 區域內有一個接近的強度（如：強度差距在 10 以下），我們就會給予黑色，代表強度有接近的部份，做出來的結果如圖 3-4 所示。在圖 3-4 中強度接近的，便會給予黑色，這樣的結果可以看出白色明顯變少了。做完模型比較跟強度比較後，我們會把白色跟黑色以外的結果記錄下來，做區域高斯模型的更新，現在紀錄的結果會把前一張的模型做時序上的更新結合，在下一張作區域模型比較時，便會用到此更新統計結果。

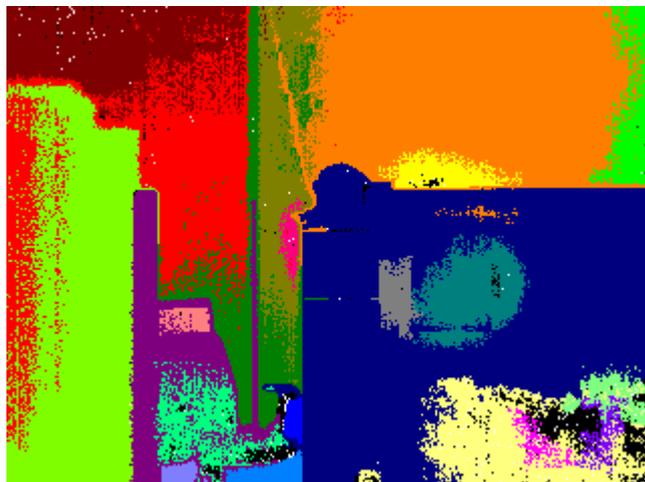


圖 3-4 強度區域比較後結果。

3.3 區域高斯模型的更新

做完上面兩個章節的步驟後，我們想要依據時序的變化，來更新區域高斯混合模型，如圖 3-5 所示。首先我們現在區域比對完後，每一個像素都會有對應的區域高斯模型編號，對相同編號的所有像素，計算其平均值跟變異量如（公式 2.1）。新得出的區域高斯模型，便會跟前一張的高斯模型做結合，得到更新後真正的平均值跟變異量。

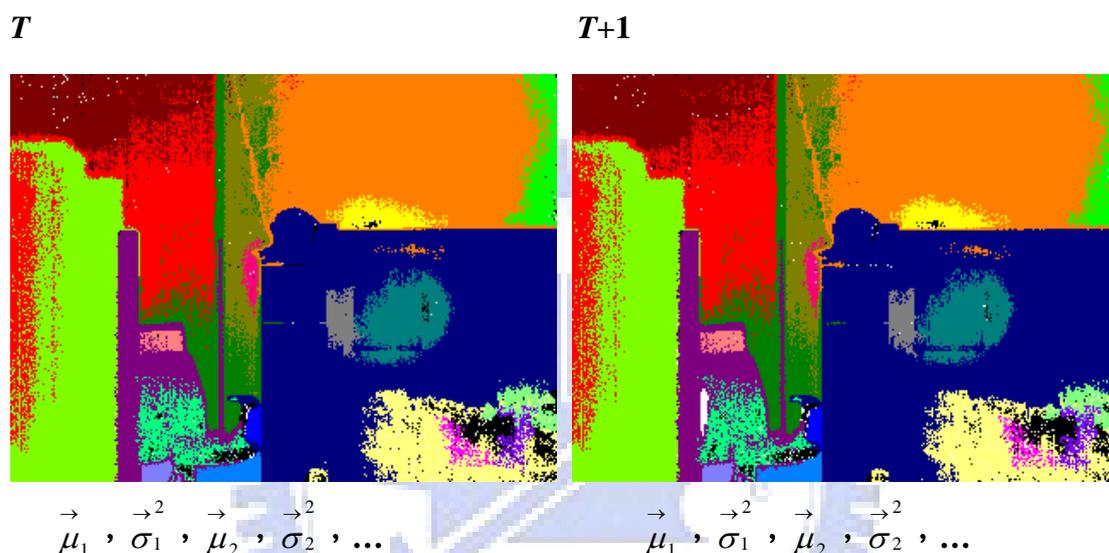


圖 3-5 更新示意圖。

前後影像之區域高斯模型結合方法，如公式 (3.1) 所示：

$$\begin{aligned}
 \vec{\mu}_{t+1} &= (1-\rho)\vec{\mu}_t + \rho\vec{\mu}_{t+1} \\
 \vec{\sigma}_{t+1} &= (1-\rho)\vec{\sigma}_t + \rho\vec{\sigma}_{t+1} \\
 \rho &= \alpha \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{1}{2} \frac{(\mu_{t+1} - \mu_t)^2}{\sigma_t^2}\right)
 \end{aligned} \tag{3.1}$$

除了 α 是用來控制 blending factor ρ 的參數，需要手動設定，其他的參數都是計算出來的結果。三維的區域高斯因為比較適用於理想狀況，之後將在未來有更深入的研究。

做完更新後，以後每一張影像，都會從上述的步驟開始，直到更新做完。每一影像做完後，如果有出現一大片白色的區域，可能為前景物體，其餘的細碎的白色，可能是為雜訊。在下節實驗結果中，我們可以看到完整的背景更新結果，

以及前景的對應區分。

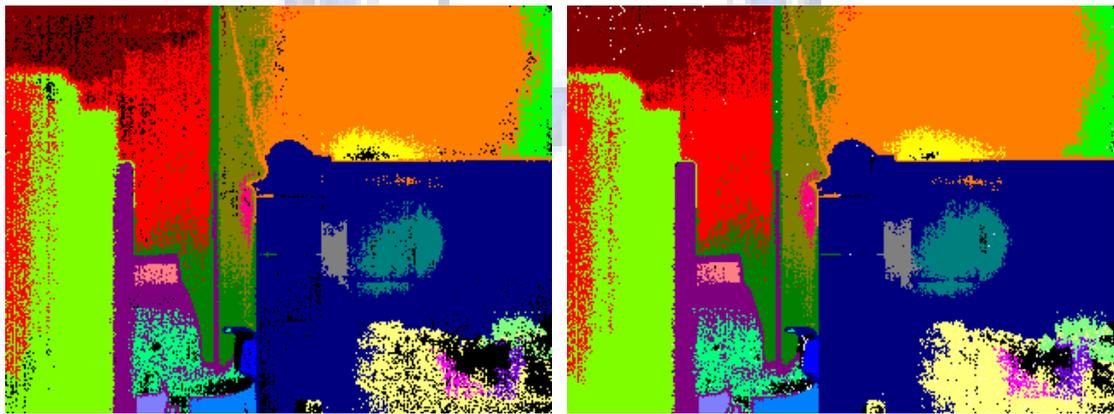


四、實驗結果與探討

我們提出的紅外線前景物的偵測方法，在本章節可以看到更完整的結果，證明我們的方法可以確切找出前景物。而我們不僅在靜態攝影機的影像可以偵測前景物，在動態攝影機的影像也可以有不錯的前景切割結果。我們另外也會跟高斯混合模型來做比較，比較兩者間的差異。

4.1 靜態連續影像的偵測結果

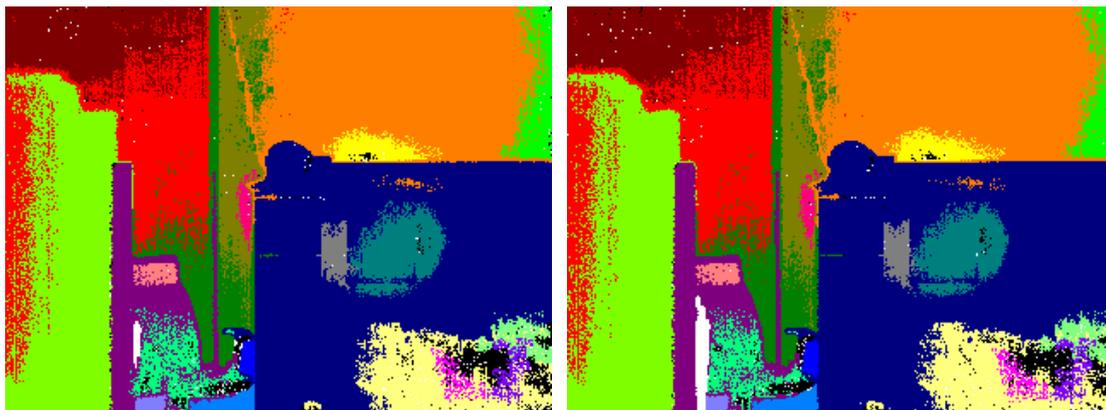
在靜態影像的前景物偵測結果，下面有兩個不同的連續影像，兩者的差別在於環境有很大的不同，一個是背景有很多物體，而另一個就只有少數個物體，圖 4-1 為第一個連續影像的結果，圖 4-2 是另一個連續影像的結果。由兩個結果看出，不同數目背景物體造成的影響。



(a)

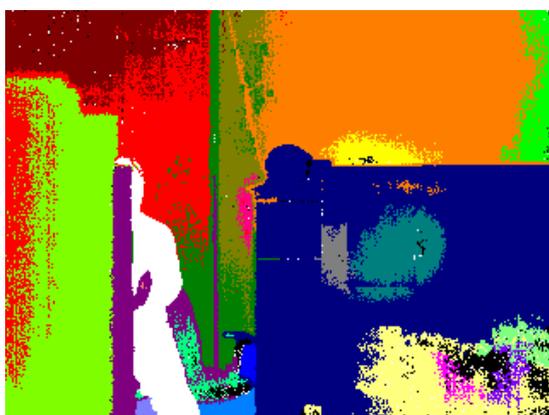
(b)

圖 4-1 複雜背景連續影像偵測結果，(a) 第 1 張影像，(b) 第 2 張影像，(c) 第 3 張影像，(d) 第 4 張影像，(e) 第 5 張影像。



(c)

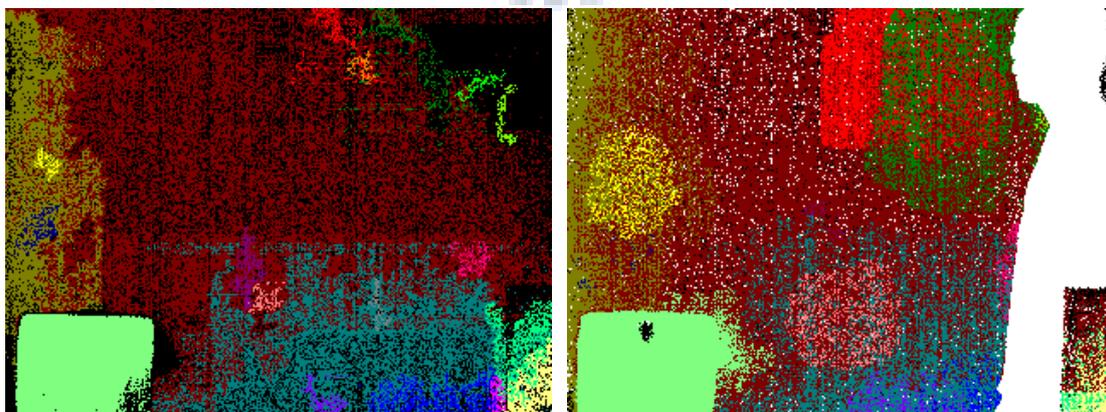
(d)



(e)

圖 4-2 複雜背景連續影像偵測結果，(a) 第 1 張影像，(b) 第 2 張影像，(c) 第 3 張影像，(d) 第 4 張影像，(e) 第 5 張影像。(續)

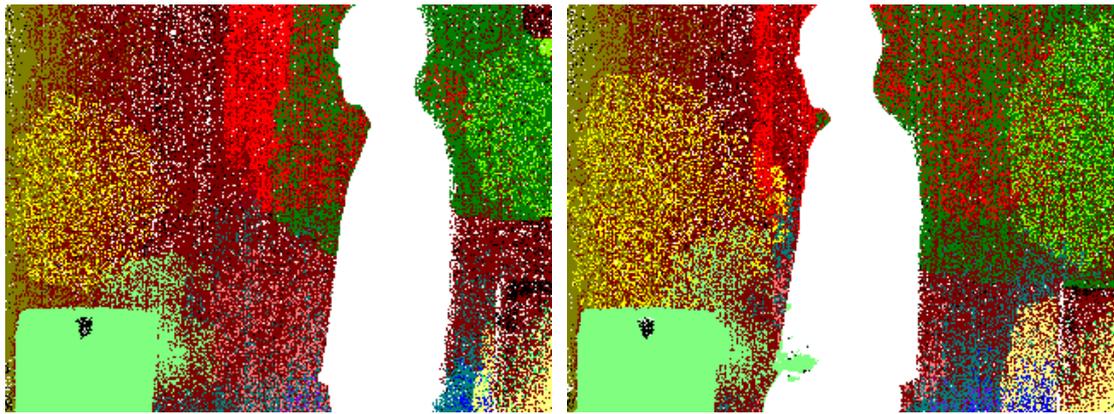
圖 4-1 中是我們原先的實驗結果，人會拿著一個物體，雖然會造成一些影響，但還是可以看出來，人的部份大致上是完整的，而白色零碎的點看的出來也很少。



(a)

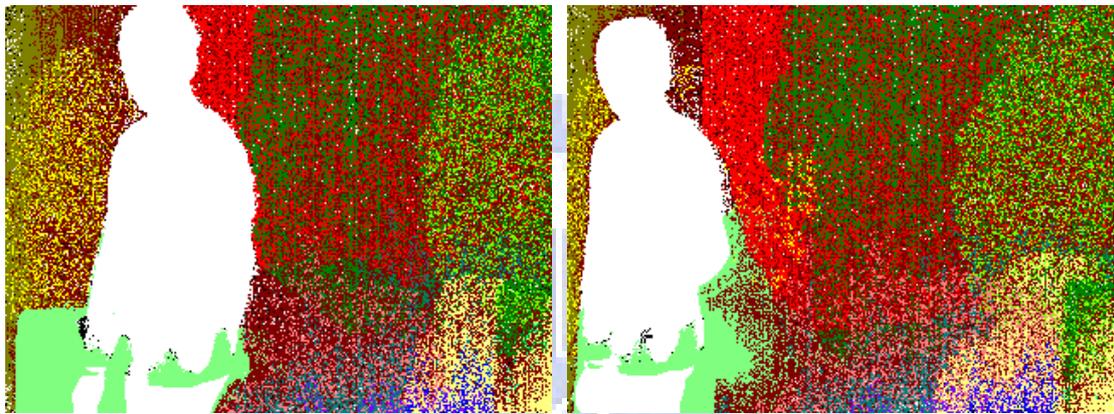
(b)

圖 4-3 單純背景連續影像偵測結果，其中 (a)，(b)，…，(j) 分別是第 1, 20, 40, 60, 80, 100, 120, 140, 160, 180 張影像。



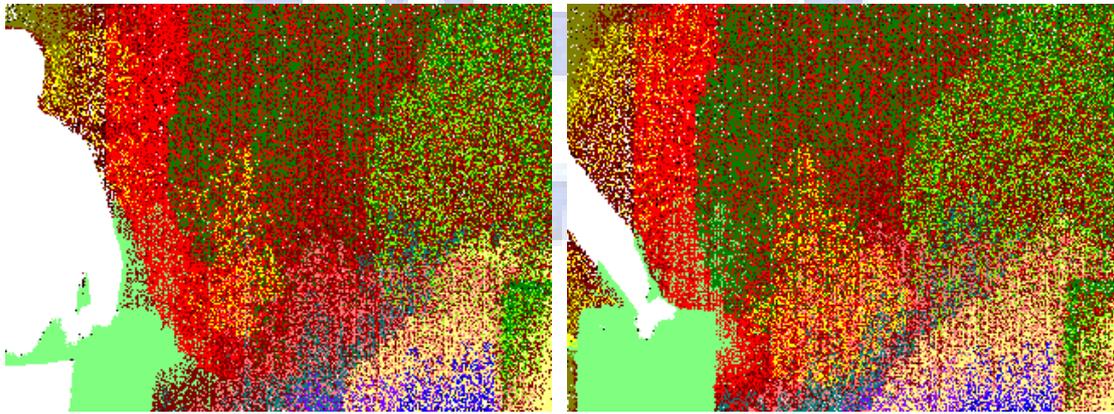
(c)

(d)



(e)

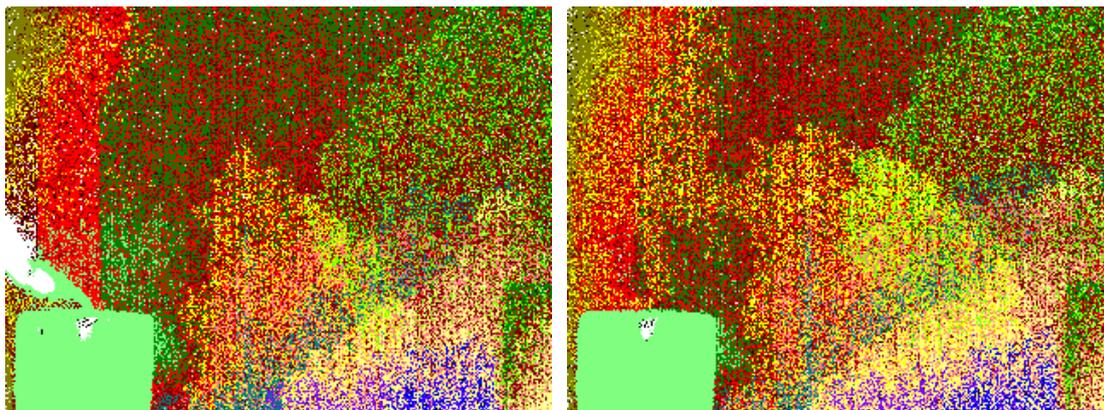
(f)



(g)

(h)

圖 4-4 單純背景連續影像偵測結果，其中 (a) , (b) , ... , (j) 分別是第 1, 20, 40, 60, 80, 100, 120, 140, 160, 180 張影像。(續)



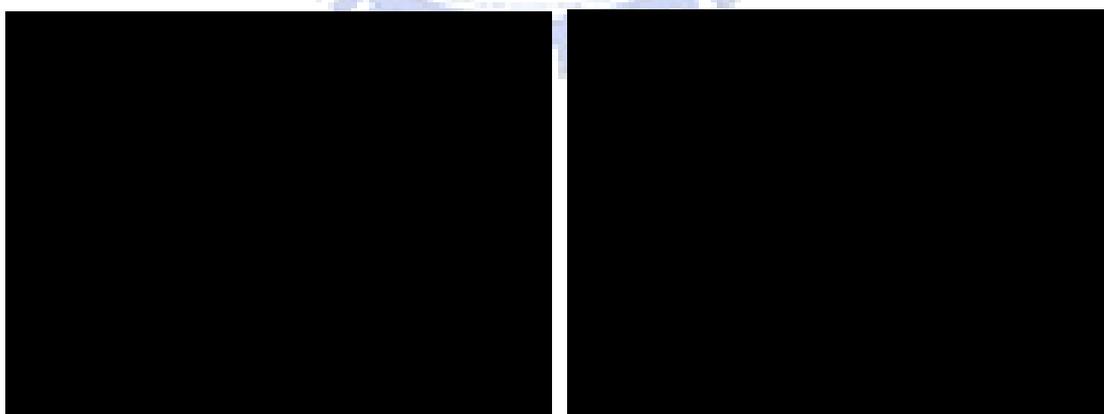
(i)

(j)

圖 4-5 單純背景連續影像偵測結果，其中 (a)，(b)，…，(j) 分別是第 1, 20, 40, 60, 80, 100, 120, 140, 160, 180 張影像。(續)

圖 4-2 的結果很明顯的，背景的部分因為物體較少，所以模型的分佈會比有很多物體的複雜一點，但前景物還是可以有很好的區分，但最後因為微波爐跟人的衣服溫度很接近，使得人有些部份被微波爐影響，而沒有很好的區分。

再來我們把高斯混合模型跟我們的方法分別做比較，針對一樣的連續影像來做前景的區分。圖 4-3 (a1)，(b1)，…，(m1) 是高斯混合模型的物體偵測結果，而圖 4-3 (a2)，(b2)，…，(m2) 則是我們方法的物體偵測結果，兩者間有些微的差異。在執行的效率上，兩個方法會對同樣的連續影像跑十次，然後取平均值，高斯混合模型的執行時間平均是 20.8313 秒，而我們的方法則需要花費 29.6685 秒。



(a1)

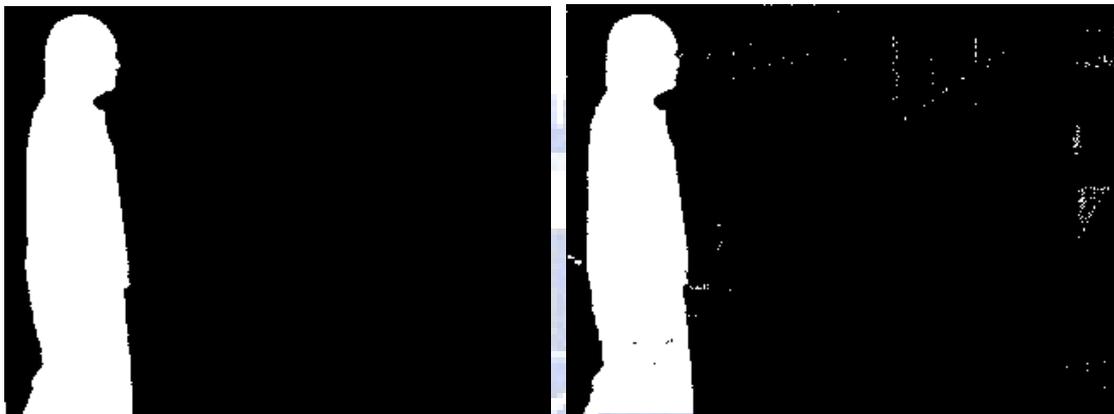
(a2)

圖 4-6 靜態連續影像比較(a1)，(b1)，…，(m1)是高斯混合模型的偵測結果，(a2)，(b2)，…，(m2)是我們方法的偵測結果。(a1)、(a2)是第一張紅外影像偵測結果，之後都是間隔 50 張的偵測結果，最後 (m1)、(m2) 是第 596 張的偵測結果。



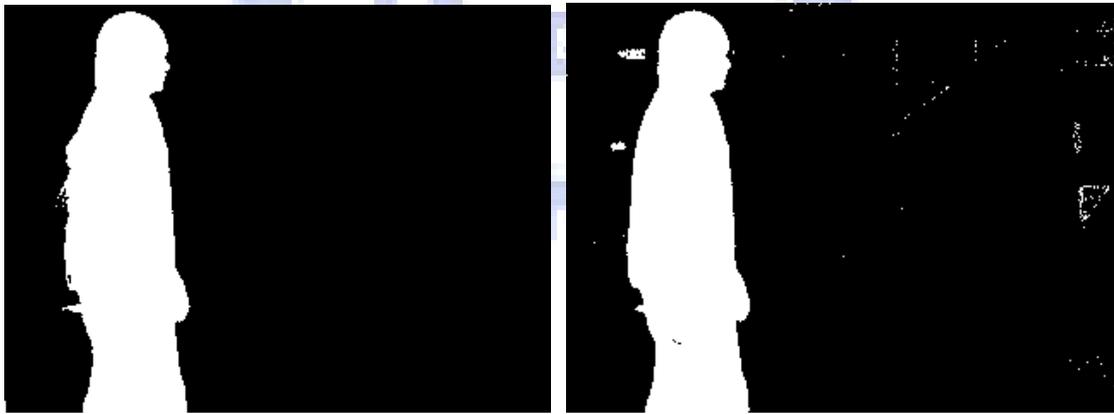
(b1)

(b2)



(c1)

(c2)



(d1)

(d2)

圖 4-7 靜態連續影像比較(a1), (b1), ..., (m1)是高斯混合模型的偵測結果, (a2), (b2), ..., (m2)是我們方法的偵測結果。(a1)、(a2)是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 596 張的偵測結果。(續)

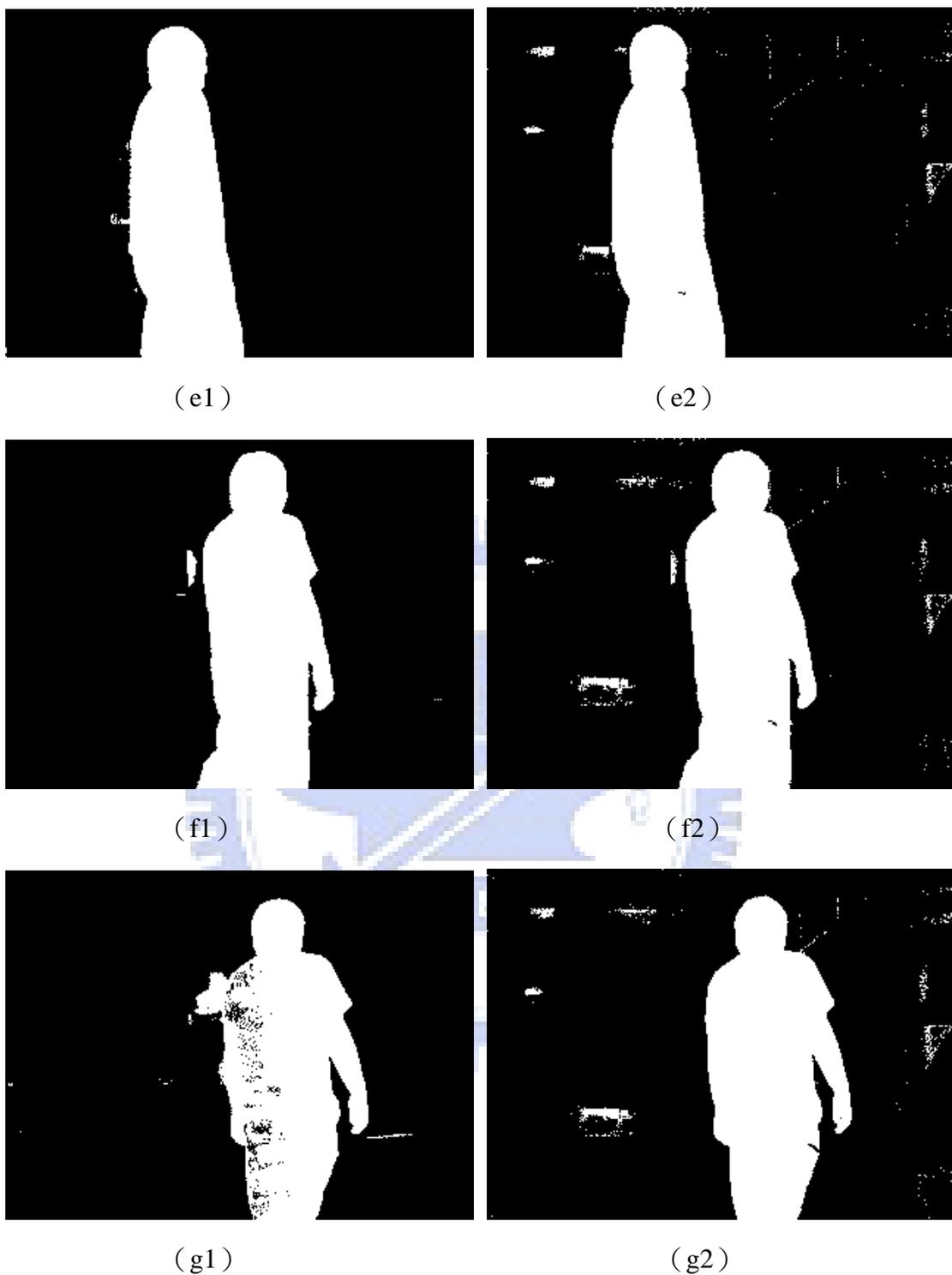
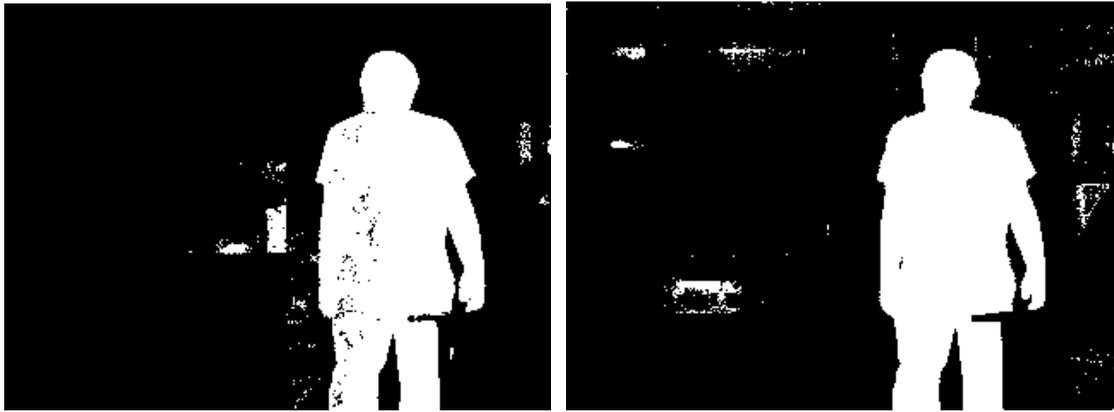
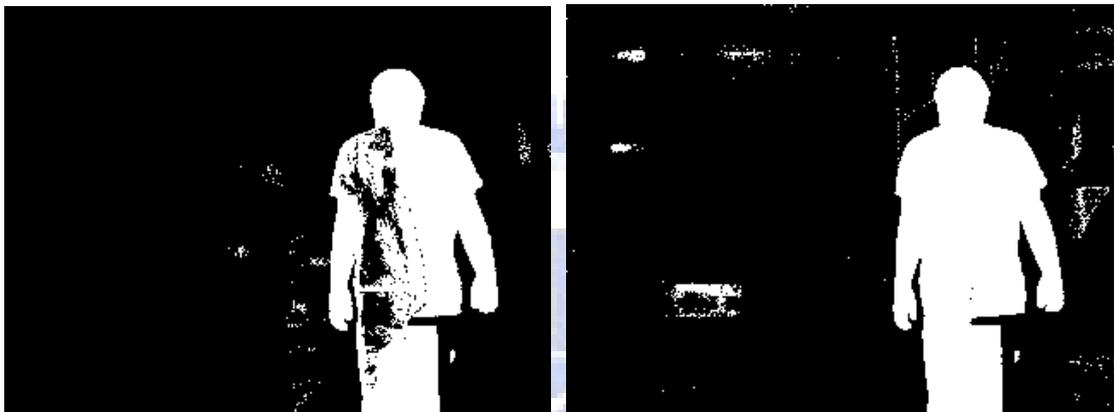


圖 4-8 靜態連續影像比較(a1), (b1), ..., (m1)是高斯混合模型的偵測結果, (a2), (b2), ..., (m2)是我們方法的偵測結果。(a1)、(a2)是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2)是第 596 張的偵測結果。(續)



(h1)

(h2)



(i1)

(i2)



(j1)

(j2)

圖 4-9 靜態連續影像比較(a1), (b1), ..., (m1)是高斯混合模型的偵測結果, (a2), (b2), ..., (m2)是我們方法的偵測結果。(a1)、(a2)是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2)是第 596 張的偵測結果。(續)



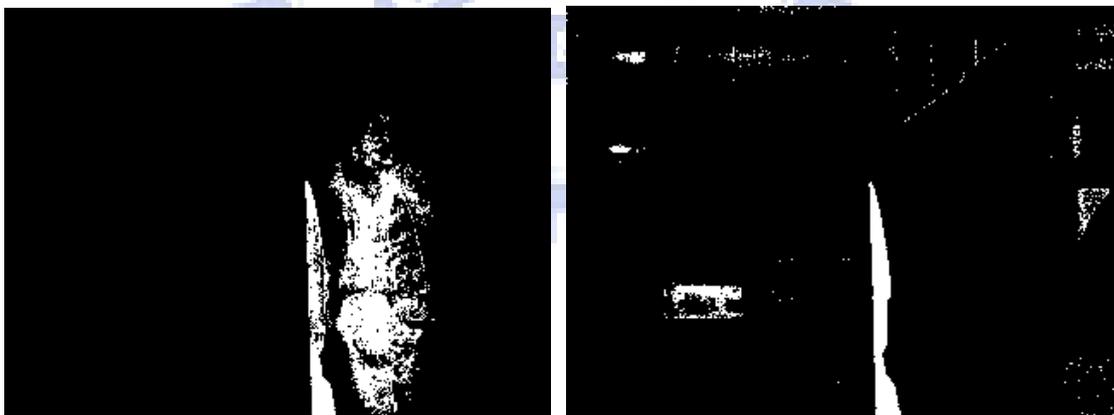
(k1)

(k2)



(l1)

(l2)



(m1)

(m2)

圖 4-10 靜態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的偵測結果, (a2), (b2), ..., (m2) 是我們方法的偵測結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 596 張的偵測結果。(續)

從圖 4-3 我們可以看到我們方法的偵測結果, 在前景物的部份大致上有很穩定的切割。但高斯混合模型在圖 4-3 (g1) 以後, 可以發現前景物體, 會有一些不穩定的結果, 甚至有找到錯的前景物體, 這是由於高斯混合模型會在一段時間內

將一直存在同一位置的前景物歸為背景。最後的高斯混合模型的偵測影像，就會有不完整的前景物體切割出現，相較於我們的偵測結果到最後，前景物還是能有完整的切割。

4.2 動態連續影像的偵測結果

在接下來的動態連續影像中，攝影機會隨著人移動，但偵測的結果還是可以看出前景物大致上可以區分出來，但是背景物體的一些部份溫度跟人的溫度有一些接近的地方，造成人在比對的時候會找到符合的模型，導致人的切割上，會有一些部份受到背景物體的影響。圖 4-4 為動態連續影像的偵測結果。

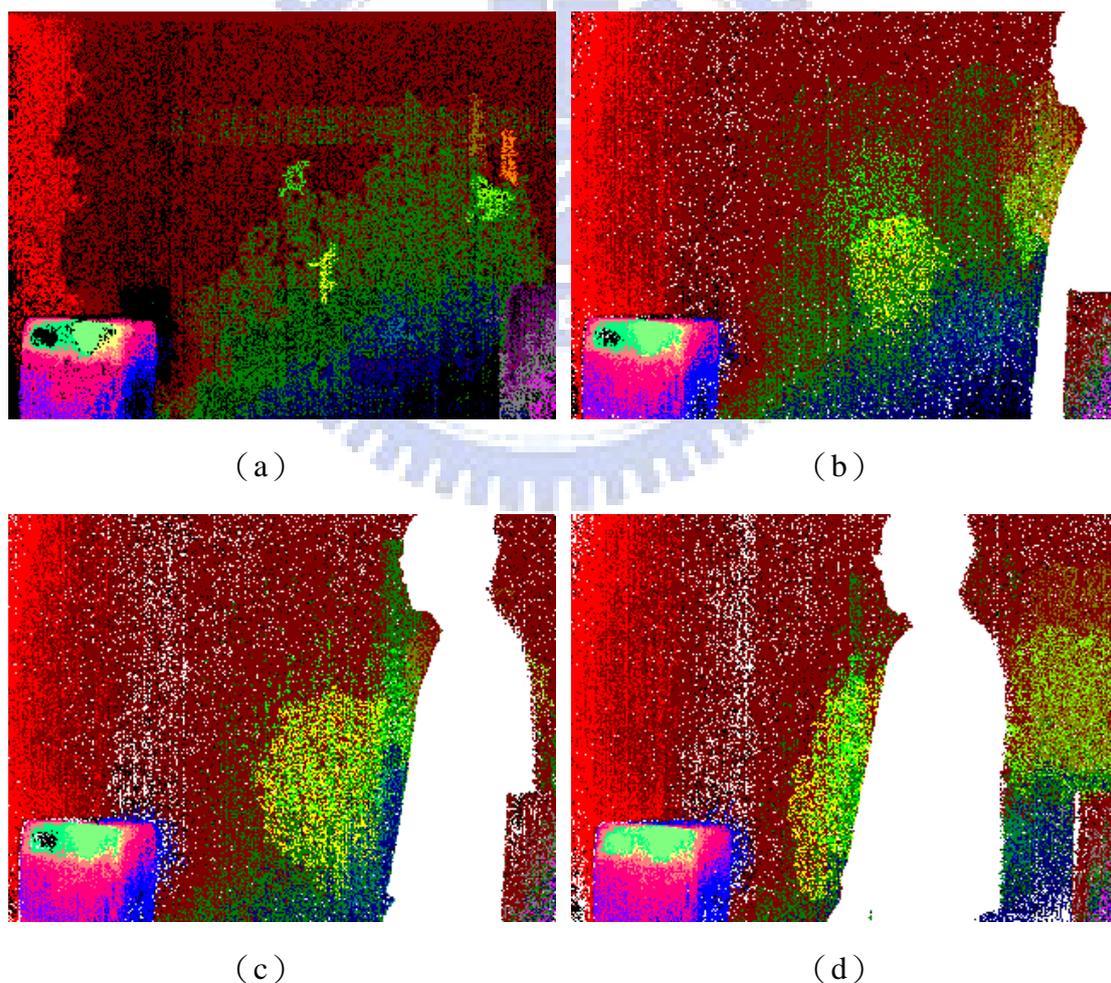
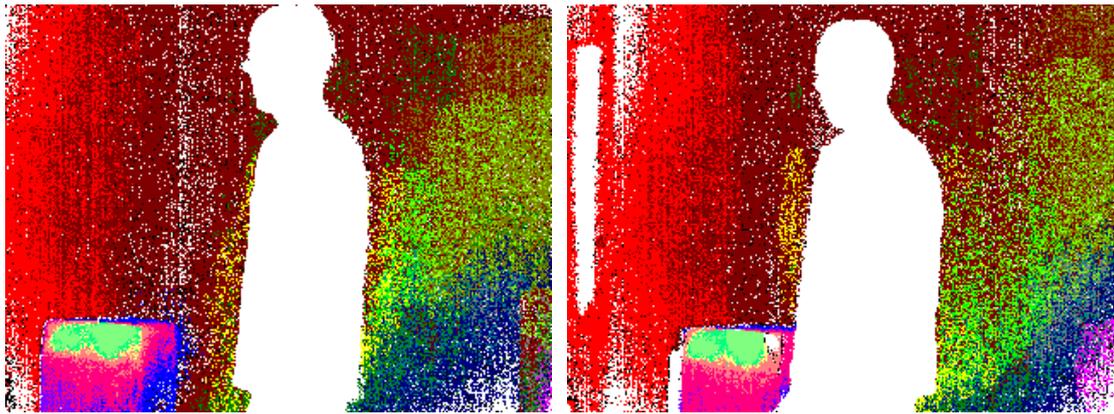
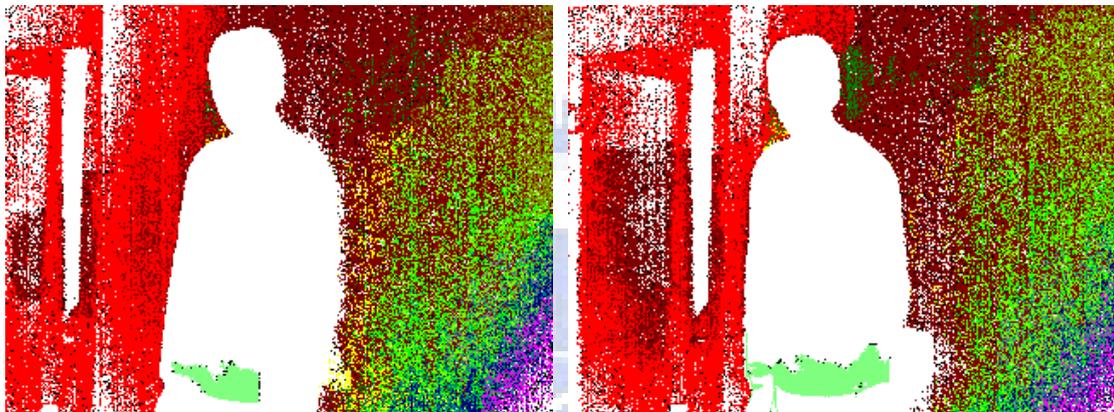


圖 4-11 動態連續影像偵測結果(a),(b),..., (k)分別是第 1,20,40,60,80,100,120,140,160,180,200 張影像。



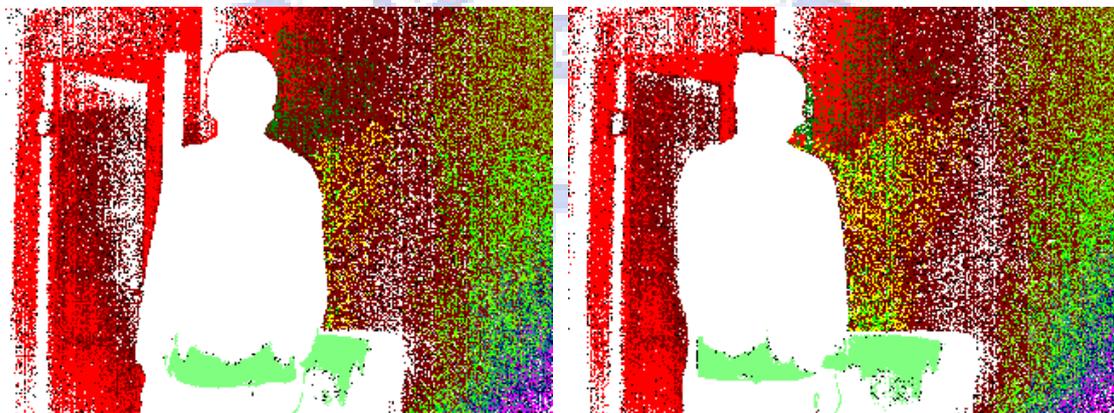
(e)

(f)



(g)

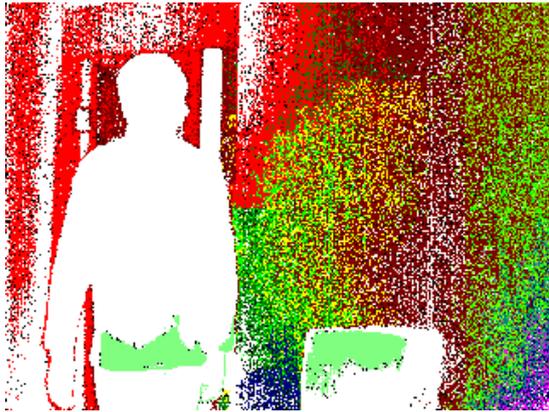
(h)



(h)

(j)

圖 4-12 動態連續影像偵測結果(a),(b),...,(k)分別是第 1,20,40,60,80,100,120,140,160,180,200 張影像。(續)

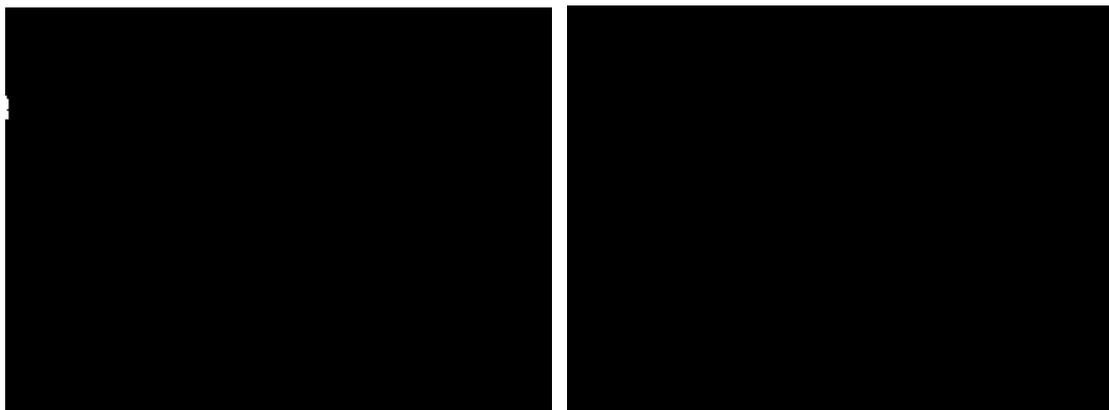


(k)

圖 4-13 動態連續影像偵測結果(a),(b),...,(k)分別是第 1,20,40,60,80,100,120,140,160,180,200 張影像。(續)

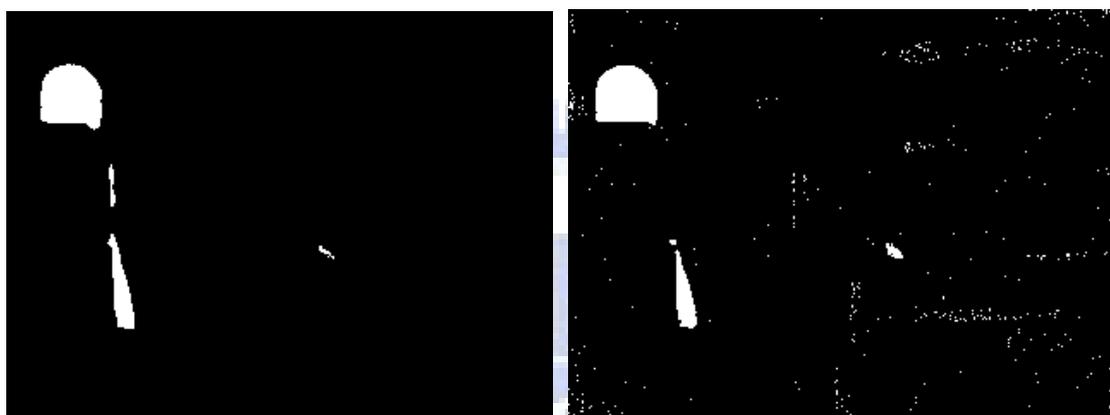
圖 4-4 除了微波爐回影響到人的部份，還有一個問題就是人擋住微波爐的時間太久，當人走過後微波爐便無法在 5×5 的區域內找到符合的模型。所以我們是必須要考量物體的移動和遮蔽問題，還有就是由於攝影機的移動，畫面左側也多了一些可能不是前景（白色）的區域。

再來就是我們與高斯混合模型，在移動式的攝影機偵測結果的比較。我們都知道高斯混合模型比較適用於靜態的攝影機，但套用在移動式攝影機的偵測效果，不會比靜態的偵測效果好。我們的方法雖然有利用到高斯混合模型，但為了能套用在移動式攝影機，所以做了不同於高斯混合模型的改變。圖 4-5 左側部分是高斯混合模型在移動式攝影機的偵測結果，而圖 4-5 的右側部分則是我們方法的偵測結果。



(a1)

(a2)



(b1)

(b2)



(c1)

(c2)

圖 4-14 動態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的結果, (a2), (b2), ..., (m2) 是我們方法的結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 600 張的偵測結果。



(d1)

(d2)



(e1)

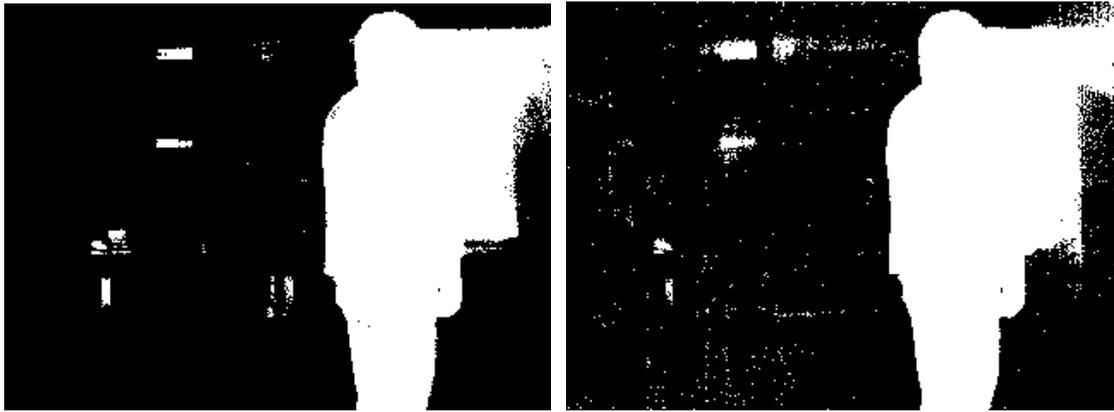
(e2)



(f1)

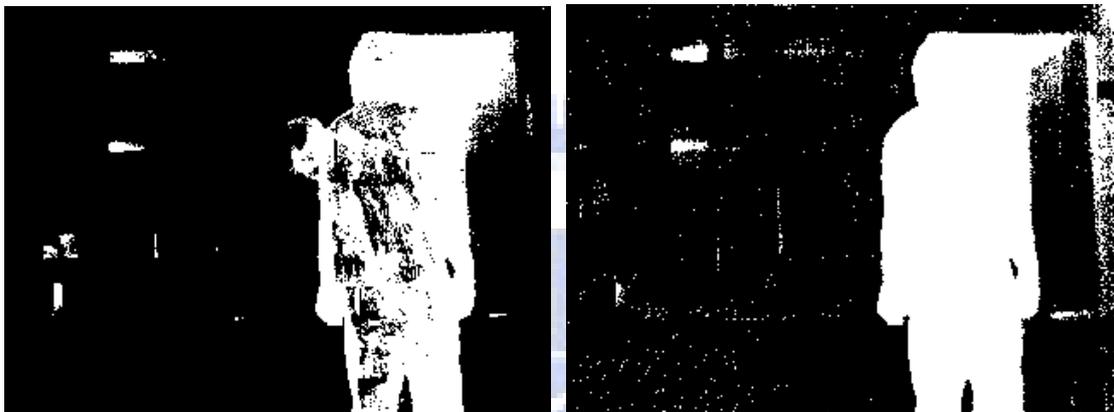
(f2)

圖 4-15 動態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的結果, (a2), (b2), ..., (m2) 是我們方法的結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 600 張的偵測結果。(續)



(g1)

(g2)



(h1)

(h2)



(i1)

(i2)

圖 4-16 動態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的結果, (a2), (b2), ..., (m2) 是我們方法的結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 600 張的偵測結果。(續)



(j1)



(j2)



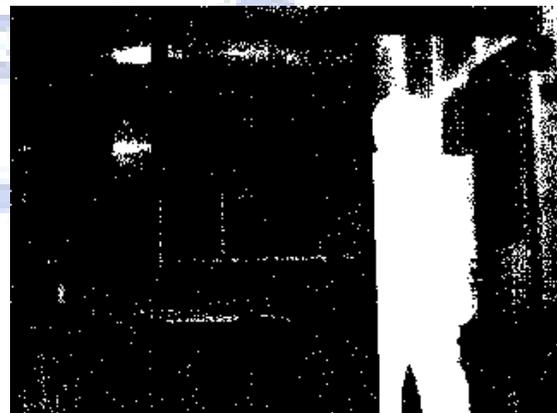
(k1)



(k2)

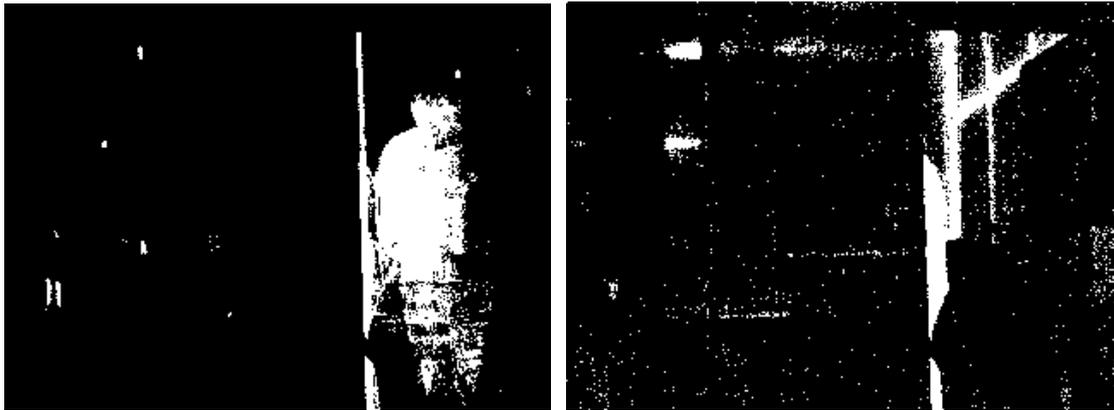


(l1)



(l2)

圖 4-17 動態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的結果, (a2), (b2), ..., (m2) 是我們方法的結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 600 張的偵測結果。(續)



(m1)

(m2)

圖 4-18 動態連續影像比較 (a1), (b1), ..., (m1) 是高斯混合模型的結果, (a2), (b2), ..., (m2) 是我們方法的結果。(a1)、(a2) 是第一張紅外影像偵測結果, 之後都是間隔 50 張的偵測結果, 最後 (m1)、(m2) 是第 600 張的偵測結果。(續)

圖 4-5 兩個方法的執行時間上, 高斯混合模型花費 21.5345 秒, 我們的方法則花費 32.0999 秒。但從圖中可以看出我們方法的偵測結果, 與高斯混合模型的偵測結果差異。從圖 4-5 (h1) 之後可以看到, 前景的切割部分與我們的切割有些差異, 因為高斯混合模型會一段時間內將前景物歸為背景, 但我們的方法就無此問題。例如最後的兩個畫面中前景的中央區域先被高斯混合模型歸為背景, 一旦前景物離開, 又造成高斯混合模型將原本背景的部分當作前景的問題, 所以造成前景部份切割並不是很好。

五、結論與未來展望

在本篇論文中，我們以高斯混合模型當為基礎，稍微改變一下模型的維度，建出起始的高斯模型後，再利用連通區域標記法，對於不同的物體給予不同的模型，最後再修補一些鄰近的相似點，背景模型的建立就可以有良好的結果。之後的影像都會對建好的模型，在 5×5 的區域做比對動作，一旦有符合的模型該像素便會記錄符合的模型，若沒有符合就會再比對影像強度，使與背景不符的像素減少。最後影像中所有的像素都做完比對後，再把紀錄符合的模型，一一重算新的模型，再把現在的模型，跟上一張的模型，做更新的計算。依此方法所偵測得的前景的區域，大致上都可以獲得比傳統高斯混合模型計算所得更為完整的輪廓。

對於未來的研究，要改善的是場景移動時，新的背景可能導致我們不想要的前景區域會出現，所以必須要有一個方法，可以正確的找出我們想要的前景區域。另外就是背景的物體，如果有跟前景物體溫度接近的情況下，是否可以完全區分前景物體跟背景物體，仍有待探討。還有就是去探討如果背景物體有被前景物體遮蔽，遮蔽的時間太久，導致背景物體無法找到接近的模型，是否有其他方法可以解決。

參考文獻

- [1] R. C. Gonzales and R. E. Woods, *Digital Image Processing* 2nd ed., Prentice Hall New Jersey, 2002.
- [2] J. Silverman, J. Mooney, W. Ewing, and D. Saro, “Real-Time Automated Scene Display for Infrared Cameras,” United States Patent 5249241, 1993.
- [3] 王俊昌, 莊仁輝, “一個實用的紅外線影像對比強化方法,” 中華民國專利發明 I255653, 2006.
- [4] C. Stauffer and W. E. L. Grimson, “Adaptive Background Mixture Models for Real-Time Tracking,” *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 2, 1999.
- [5] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee. “Using Adaptive Tacking to Classify and Monitor Activities in a Site,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp.22-29, June 1998.
- [6] S. Birchfield, “An Elliptical Head Tracker,” *Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1710-1714, 1997.
- [7] T. Zhao and R. Nevatia, “Tracking Multiple Humans in Complex Situations,” *Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1208-1212, 2004.
- [8] L. D. Stefano and A. Bulgarelli, “A Simple and Efficient Connected Components Labeling Algorithm,” *International Conference on Image Analysis and Processing*, pp. 322-327, 1999.
- [9] S.A. Berrabah, G. De Cubber, V Enescu, and H. Sahli, “MRF-Based Foreground Detection in Image Sequences from a Moving Camera,” *IEEE International Conference on Image Processing*, pp. 1125 – 1128, 2006
- [10] D. Zhou and H. Zhang, “Modified GMM Background Modeling and Optical Flow for Detection of Moving Objects,” *IEEE International Conference on Systems, Man and Cybernetics*, vol. 3, pp. 2224 – 2229, 2005
- [11] M. Fathy and M. Y. Siyal, “An Image Detection Technique Based on Morphological Edge Detection and Background Differencing for Real-Time Traffic Analysis,” *Pattern Recognition*, vol. 16, pp. 1321-1330, 1995.
- [12] A. J. Lipton, H. Fujiyoshi and R. S. Patil, “Moving Target Classification and Tracking from Real-Time Video,” *Fourth IEEE Workshop on Application of Computer Vision*, pp. 8-4, 1998.