

國立交通大學

多媒體工程研究所

碩士論文

視訊背景次像素平移量分析與分割

Sub-pixel Motion Analysis and Segmentation of Background in
Video

研究生：李宗穎

指導教授：莊仁輝 教授

中華民國九十六年六月

視訊背景次像素平移量分析與分割

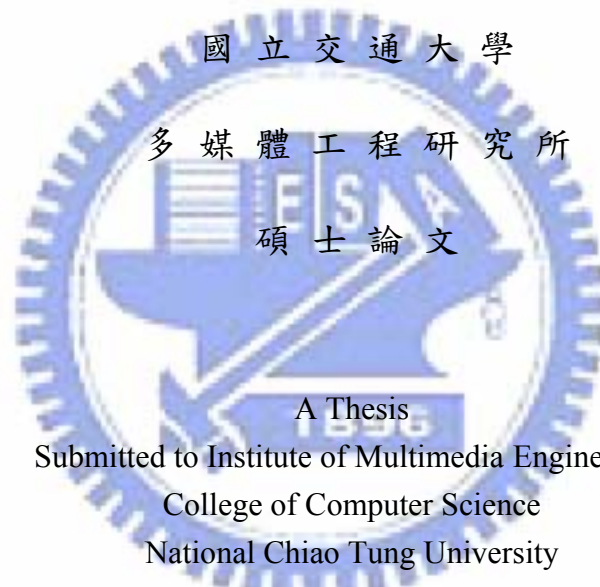
Sub-pixel Motion Analysis and Segmentation of Background in
Video

研究生：李宗穎

Student : Lee, Tsung-Ting

指導教授：莊仁輝

Advisor : Jen-Hui Chuang



A Thesis
Submitted to Institute of Multimedia Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Computer Science

June 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年六月

視訊背景次像素平移量分析與分割

學生：李宗穎

指導教授：莊仁輝 博士

國立交通大學

多媒體工程研究所

摘要

在此篇論文中，我們提出了一種對於視訊背景次像素平移量的估計方法，此方法結合了影像空間以及頻率空間中的資訊。首先，我們利用影像空間的資訊作背景區域的比對，找出部分背景區域以及粗略估計背景平移量。接著利用行與列的一維傅立葉轉換，在頻率空間，利用影像轉換後的相位關係，分別分析背景區域的垂直及水平方向的次像素平移量。透過漸進式由粗而細的平移量分析，我們可以分割出背景區域。實驗結果顯示，利用上述演算法可以得到比像素為單位的平移量更為精準的次像素平移量。此外，本論文的作法也避免了利用頻率空間進行前景、背景分割時，無法處理前景物體形變的問題。

Sub-pixel Motion Analysis and Segmentation of Background in Video

Student : Lee, Tsung-Ying

Advisor : Dr. Jen-Hui Chuang

Institute of Multi-media
National Chiao Tung University

The logo of National Chiao Tung University is a circular emblem. It features a central shield with a book, a lamp, and the letters 'ES A'. Below the shield is a banner with the year '1896'. The entire emblem is surrounded by a gear-like border with the university's name in Chinese characters.

ABSTRACT

In this thesis, we propose an approach to sub-pixel background motion analysis in a video by both spatial and frequent information. First, rough translation of the background is obtained by partial background matching using image space information. Subsequently, more accurate sub-pixel motion analysis of background is carried out in one dimensional Fourier domain of image rows and columns using the phase information. Through the above coarse-to-fine procedure of motion analysis, major and significant background area in the video can be segmented. Experimental results show that the sub-pixel motion analysis is indeed more accurate than pixel-based methods, while the problem of shape changing of foreground objects, often associated with segmentation in Fourier domain, is avoid.

誌謝

本論文承蒙指導教授莊仁輝老師的細心指導與教誨，終於讓這篇論文得以付梓。在平時的討論上，莊老師認真作研究及生活處世的態度，讓本人獲益良多。我要感謝智慧型系統實驗室的高肇宏學長、林泓宏學長等的指教，陪我一同度過兩年時光的同學們，和其他學長、學弟妹們，在這段時間的協助與關心。也感謝口試委員們對這篇論文的批評、指教，使得論文的內容更添豐富。此外，對於家人與朋友們的支持與鼓勵，我始終銘記於心，衷心感謝大家對我的關懷。



目錄

摘 要.....	I
ABSTRACT.....	II
誌謝.....	III
目錄.....	IV
圖例目錄.....	VI
第一章 研究動機.....	1
第二章 背景.....	2
2.1 次像素平移量分析.....	2
2.1.1 基於影像空間的分析.....	2
2.1.2 基於頻率空間的分析.....	3
2.2 前景與背景分割.....	5
2.2.1 靜止的背景.....	5
2.2.2 移動中的背景.....	8
第三章 背景區域的比對及平移量計算.....	10
3.1 以平均值與變異量做對應背景區域的比對.....	11
3.1.1 背景區域的粗略估計.....	11
3.1.2 背景平移量的粗略估計.....	16
3.2 頻率空間相位線性分析.....	17
3.2.1 平移的性質.....	17
3.2.2 振幅比對與相位誤差.....	23
3.2.3 位移量的估計.....	25
3.2.4 帶狀區域的位移量估計.....	26
第四章 背景區域的分割.....	28
4.1 利用平均與變異量進行行／列的比對.....	28

4.2 相位線性分析.....	30
4.3 與背景平移量的比較、背景區域的估計結果.....	32
4.4 邊界效應的處理.....	33
第五章 實驗結果與問題討論.....	35
5.1 人造影像分析.....	35
5.2 自然影像之分析與模擬.....	39
5.2.1 自然影像的背景區域分割.....	39
5.2.2 自然影像的次像素平移量討論與應用.....	44
第六章 結論.....	51
參考文獻.....	52



圖例目錄

圖 3-1 利用行的平均值計算垂直方向平移量計算的流程圖。.....	10
圖 3-2 (a) 前後兩張影像。 (b) 兩張影像各行的灰階平均值。 (c) 兩張影像各行的灰階變異量。 紅色區域為影像中的純背景區域。.....	12
圖 3-3 包含移動物體的行在前後兩個影像中呈現不一樣的平均值。(a) 前後兩張影像。 (b) 兩張影像各行的灰階平均值。 (c) 兩張影像各行的灰階變異量。.....	13
圖 3-4 當選擇的區域的變異量過小時，容易造成的錯誤。(a) 前後兩張影像。 (b) 兩張影像各行的灰階平均值。 (c) 兩張影像各行的灰階變異量。.....	14
圖 3-5 比 圖 3-2 有更寬的純背景區域被以紅色標示。(a) 前後兩張影像。 (b) 兩張影像各行的灰階平均值。 (c) 兩張影像各行的灰階變異量。.....	15
圖 3-6 (a)、(b)兩張影像為影片中前後兩張影格。紫紅色區域為垂直帶狀區域，經過比對之後被判定最相似的兩對應區域。.....	16
圖 3-7 (a), (b)為兩張只有垂直平移關係的影像。(c)是兩張影像中央行的相位差關係。.....	18
圖 3-8 還原相位差線性關係的演算法。.....	19
圖 3-9 圖 3-7 (c) 經過轉換之結果.....	19
圖 3-10 (a)、(b) 為影片中前、後兩個影像 (c) 表示白色方塊由上向下移動。(d) 使用頻率空間來處理數位影像時，則可視為影像為週期性的，也就是在影像邊界外的部份，是由同樣的影像所構成，而形成另外一個白色方塊，因此，我們可以將方塊的移動視為由下向上移動。且此兩平移量值總和為影像長度。.....	20
圖 3-11 兩張僅存在垂直向上平移關係的影像。(a) 紅色區域表示在另一張影像中將隱沒的部分 (b) 綠色區域表示在影像中新增加的部分.....	21

圖 3-12 (a)、(b)為兩張僅存在平移關係的真實影像。(c)為(a)、(b)影像之中央行(以紅色線段標示)的相位差關係。(d)為(c)中的折線,經過轉換後所得的近似線性相位差關係。.....	22
圖 3-13 (a)、(b)為兩張只有存在垂直平移關係的影像。(c)兩色分別為兩張影像中央紅色線段標示之行的振幅,可以明顯觀察到兩曲線不完全重疊。雖然兩張影像的內容大部分是相同,背景移動造成的在影像邊界部分的增減,使振幅存在少量的差異。.....	23
圖 3-14 (a)、(b)中的黑色箭號表示兩個相同角度,但是大小不同的向量。在黑色箭號端點的圓則是假設的誤差範圍。紅色箭號則是估計值。.....	24
圖 3-15 圖 3-12 (c) 利用 PCA 求得相位差的線性關係。藍色線為原始的相位差的低頻率部分,紅線為 PCA 估計所得的近似直線,原點在圖的中央。....	26
圖 4-1 前景與背景區域分割的流程圖。.....	29
圖 4-2 經過比對,最相似的區域。(a)、(b)為影像的前後兩張影格,紫紅色標示的區域為兩張影像最相似的垂直區域。(c)、(d)是與(a)、(b)同樣的前後兩張影格,青色標示的區域為兩張影像最相似的水平區域。.....	30
圖 4-3 圖 4-2(a)、(b)兩張影像相位差 PCA 分析結果。藍色線段為相位差關係,紅色線段為直線估計結果。.....	31
圖 4-4 圖 4-2(c)、(d)兩張影像相位差 PCA 分析結果。藍色線段為相位差關係,紅色線段為直線估計結果。.....	31
圖 4-5 利用背景水平方向的平移量找出對應的垂直背景區域。.....	33
圖 4-6 利用背景垂直方向的平移量找出對應的水平背景區域。.....	34
圖 5-1 (a)、(b)為兩張合成的影像,兩張影像擷取自同一張照片,其擷取位置相差在 X 方向及 Y 方向分別為 12、10 像素。(c)、(d)分別為兩張影像估計的背景區域。估計的背景平移量為 12.0053 及 10.0047。.....	36
圖 5-2 (a)圖 5-1 中各行的變異量。(b)圖 5-1 中各列的變異量。.....	36

圖 5-3	(a)、(b)為兩張合成的影像，兩張影像擷取自同一張照片，其模擬的背景 平移量為(-2.5, -3.5)。(c)、(d)分別為在兩張影像所估計的背景區域。估計的 背景平移量為(-2.2047, -3.5386)。	37
圖 5-4	(a)圖 5-3(a)各行的變異量。(b)圖 5-3(a)各列的變異量。	37
圖 5-5	(a)、(b)為兩張合成的影像，兩張影像擷取自同一張照片，其模擬的背景 平移量為(6.5, 3.5)。(c)、(d)分別為在兩張影像所估計的背景區域。估計的 背景平移量為(6.2852, 3.5693)。	38
圖 5-6	(a)、(b)為 Coastguard 影片的兩張影格。(c)、(d)分別為兩張影像的背景 區域估計結果。估計的背景平移量為(7.5186, 0.0460)	39
圖 5-7	(a)、(b)為連續的兩張真實影像。(c)、(d)分別為兩張影像的背景區域 估計結果。估計的背景平移量為(-0.0528, -2.9177)	40
圖 5-8	圖 5-7 (a) 影像每行的變異量。	40
圖 5-9	(a)、(b)為連續的兩張停車場上行人走過的真實影像。(c)、(d)分別為 兩張影像的背景區域估計結果。估計的背景平移量為(-12.9296, 3.6552)	41
圖 5-10	紅色線段為圖 5-9(a)各列的灰階平均值，藍色線段為圖 5-9(b)各列的灰 階平均值。	42
圖 5-11	圖 5-9(a)各列的灰階變異量。	42
圖 5-12	停車場中兩行人併行的影片中的六張影像。	43
圖 5-13	(a)、(b)是圖 5-12 (a)、(b)影像背景區域分割的結果。(c)、(d)是圖 5-12 (c)、(d)影像背景區域分割的結果。(e)、(f)是圖 5-12 (e)、(f)影像背景區域 分割的結果。	44
圖 5-14	將圖 5-9 (b)影像以不同的平移量還原移動後與圖 5-9 (a)比較差異的結 果。(a)平移量為(-13,3)。(b)平移量為(-12,3)。(c)平移量為(-13,4)。(d) 平移量為(-12,4)。	45
圖 5-15	分別對圖 5-14 四張影像的背景區域以及非背景區域所作的分佈圖。藍 色線段為估計背景區域，紅色線段為扣除邊界及背景區域的其他部分	46

圖 5-16 連續的四張影像，第二、三、四張影像分別跟第一張影像的次像素平移量估計，及背景區域估計的擷果。(a)、(c)、(e)皆為第一張影像，與第二、三、四張影像的進行背景區域估計。(b)為第二張影像。(c)為第三張影像。(d)為第四張影像。(a)與(b)，(c)與(d)。(e)與(f)的估計平移量分別為(-5.3604, 0.9742)、(-8.1320, 2.1695)、(-12.9269, 3.6552)。..... 47

圖 5-17 (a)-(d)為圖 5-16(b)還原整數平移量與圖 5-16(a)比較的差異。(e)-(h)為圖 5-16(d)還原整數平移量與圖 5-16(c)比較的差異。(i)-(l)為圖 5-16(f)還原整數平移量與圖 5-16(e)比較的差異。..... 48

圖 5-18 圖 5-17 各張影像差異的分佈圖。..... 49



第一章 研究動機

背景의 平移量在有攝影機移動的影片中是一項很重要的資訊。背景의 平移量可以讓我們模擬攝影機的移動，以便於前景物體的分析，或是例用影像的拼貼達到場景的建立。然而背景의 平移量並非離散，在需要高準確性的研究範圍中，例如遙測系統或是生醫影像，我們需要更精確的次像素 (sub-pixel) 平移量分析，來提高背景平移量分析的準確性。

次像素的背景平移量分析可以從影像空間及頻率空間著手。利用影像空間的資訊是直觀的做法，但是通常卻需要花費大量的計算來模擬出連續的影像平面。而在頻率空間，利用相位關係可以反應出影像平移量的特性來計算次像素的平移量，在計算量減少許多，而隨著快速傅立葉轉換的發明，省卻許多從影像空間轉換到頻率空間的計算，頻率空間的次像素平移量分析一般而言比從影像空間分析中利用建立複雜的連續曲面方程式快速很多。

我們在此篇論文中提出一個新的方法：利用影像的次像素背景平移量分析，來達到影像的前景、背景區域的分割。一般的次像素影像背景平移量分析的演算法，著重在影像的拼貼，影像中並沒有前景物體，例如衛星的空照圖。而我們則利用背景區域的尋找，排除前景區域，讓有前景物體的影像，也可以進行背景平移量分析。並可進一步對前景、背景區塊分割。

至於在頻率空間利用平移量來分割前景、背景物體之作法方面，因為相位分析需要建立在物體形狀的一致之前提上，因此對於可能產生形變的物體，例如走動的行人，則無法利用相位分析其平移量。所以我們透過先僅計算背景的平移量，再利用背景的平移量來分割前、背景。

接下來，在第二章會先介紹相關的次像素平移量分析及影像分割的研究。第三章描述我們的平移量分析演算法，包含粗略的平移量分析，以及較精確的次像素的平移量分析。之後在第四章，則是探討利用背景平移量分析的結果，來進行影像分割之作法。第五章會討論我們的實驗結果，最後則是我們的結論。

第二章 背景

2.1 次像素平移量分析

在影片的拍攝時，如果攝影機移動，我們可以從前後兩張影格中觀察到背景部分有平移。如果影像中不存在移動的前景物體，則可使用影像拼貼技術中所使用的技巧，利用背景區域分析出相對的背景平移量。從影像空間中分析，大多的拼貼技術均侷限在以像素為單位的精確度，僅有少部分則可達到次像素的精確度；從頻率空間分析，則往往可以估算更精確的次像素平移量。

2.1.1 基於影像空間的分析

影像的拼貼技術有相當久的發展歷史，並且有發展出多種方法。假設兩張影像間僅存在平移關係，從影像空間著手是最直觀的想法。基於影像空間之方法中，影像之間的相關性（correlation）是個重要的統計數據[1]。我們可以使用兩張影像 I_1 及 I_2 ，二維的相關性，如 2-1 式，來衡量相似度。如果相似度很高，則 $C(u, v)$ 會存在一個高峰（peak）。一般而言，2-1 式中之相關性均經過標準化以避免因影像整體亮度變化所產生的影響。

$$C(u, v) = \frac{\sum_x \sum_y I_1(x, y) I_2(x - u, x - y)}{\sqrt{\sum_x \sum_y I_2^2(x - u, x - y)}} \quad (2-1)$$

差值平方總合（sum of difference square）和相關性的觀念相近，如 2-2 式：

$$D(u, v) = \sum_x \sum_y [I_1(x, y) - I_2(x - u, x - y)]^2 \quad (2-2)$$

$D(u, v)$ 值越小，則表示相似度越高。此種計算的好處，是可以將公式簡化成三個項目：兩張影像的個別能量（energy terms）以及兩張影像相乘的結果，如此一來，在比對的過程中固定一張影像，則可以簡化計算量。此外，為了增加利用相

關性之次像素方法的比對準確性，可在相關性中加上以平均值和變異數，如在 2-3 式中是分別以 μ 、 σ^2 代表。

$$\frac{Cov(I_1, I_2)}{\sigma_1 \sigma_2} = \frac{1}{N} \sum_x \sum_y [I_1(x, y) - \mu_1][I_2(x-u, y-v) - \mu_2] \quad (2-3)$$

其中 N 為標準化項

$$N = \left[\sum_x \sum_y [I_1(x-u, y-v) - \mu_1]^2 \sum_x \sum_y [I_2(x-u, y-v) - \mu_2]^2 \right]^{\frac{1}{2}}$$

雖然有不同的統計量來表示影像平移後的相似程度，但是這些都是使用像素為單位的方式進行比對，而不夠精確。因此，P. R. Hill [2]等人提出基於絕對差值總合（sum of absolute difference, SAD），利用八個鄰近點，建立二元二次曲面方程式，則可以使用曲面方程式求出非整數點的值。使用 $\frac{1}{4}$ 、 $\frac{1}{8}$ 、 $\frac{1}{16}$ 像素為基本單位的方式因為以二的倍數作為分割像素的依據，可以增加計算速度。Suh 和 Jeong [3]則提出更精準的 $\frac{1}{2}$ 像素（half-pixel）計算方式，使用了周圍八個像素建立二元四次曲面方程式，來計算出八個 $\frac{1}{2}$ 像素位置的值，並且使用矩陣的方式增加計算速度。

2.1.2 基於頻率空間的分析

Kuglin 和 Hines [4]在 1975 年提出使用相位相關性（phase correlation）對齊經過平移的兩張影像。影像經過傅立葉轉換（Fourier transform），以複數呈現，並可分為振幅以及相位兩個部分，如 2-4 式所示：

$$\begin{aligned} F(\omega_x, \omega_y) &= R(\omega_x, \omega_y) + iI(\omega_x, \omega_y) \\ &= |F(\omega_x, \omega_y)| e^{j\phi(\omega_x, \omega_y)} \end{aligned} \quad (2-4)$$

$R(\omega_x, \omega_y)$ 、 $I(\omega_x, \omega_y)$ 分別表示實數與虛數， $|F(\omega_x, \omega_y)|$ 是頻率空間的振幅， $\phi(\omega_x, \omega_y)$ 是頻率空間的相位。經過平移的影像與原先的影像在相位的部分呈現線性關係，可以由此線性關係推算出平移量。

Stone [5]等人，利用影像平移後在頻率空間呈現的相位特性，來處理真實影像，並針對因影像失真所造成的錯誤進行改善。Stone 將兩張影像 $f_1(x, y)$ 、 $f_2(x, y)$ 經過傅立葉轉換得到 $F_1(u, v)$ 及 $F_2(u, v)$ ，求得比率 $R(u, v)$ ，忽略影像高失真的部分，決定次像素的精確度與所允許的偏差後，使用最小平方近似法(least square) 計算兩張影像中頻率的相位近似線性關係，也就是近似的平移量。

Foroosh [6]等人在 2002 年提出以交叉功率譜 (cross-power spectrum) 來進行相位變化量的分析研究。假設兩張影像為 $f_1(x, y)$ 及 $f_2(x, y)$ ，且 $f_2(x, y) = f_1(x - x_0, y - y_0)$ ，經過傅立葉轉換後可得 $F_1(u, v)$ 及 $F_2(u, v)$ ，並將影像作取樣，取樣數為 $M \times N$ ，可分別得 $F_{d1}(u, v)$ 及 $F_{d2}(u, v)$ ，如 2-5 式所示。交錯功率譜是兩張具有平移關係的影像經過傅立葉轉換、降頻後，相除的結果，如 2-6 式所示。利用交叉功率譜，則可找出因位移而在相位差中呈現的高峰，而推算出位移量，如 2-7 式所示，其中 $c(x, y)$ 是 2-6 式中 $C(u, v)$ 經過反傅立葉轉換的結果。

$$\begin{aligned} F_{d1}(u, v) &= \frac{1}{MN} \sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} F_1\left(\frac{u+2\pi m'}{M}, \frac{v+2\pi n'}{N}\right) \\ F_{d2}(u, v) &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F_2\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right) \end{aligned} \quad (2-5)$$

$$\begin{aligned} C(u, v) &= \frac{F_{d2}(u, v)}{F_{d1}(u, v)} \\ &= \frac{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F_2\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right) \exp\left(-i\left(\frac{u+2\pi m}{M}x_0, \frac{v+2\pi n}{N}y_0\right)\right)}{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F_1\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right)} \end{aligned} \quad (2-6)$$

$$c(x, y) \approx \frac{\sin(\pi(Mx - x_0))}{\pi(Mx - x_0)} \frac{\sin(\pi(Ny - y_0))}{\pi(Ny - y_0)} \quad (2-7)$$

2006 年，Foroosh 及 Balci [7]則提出估計相位差的線性關係的方法。他們的作法是先將兩張 $M \times N$ 影像經過傅立葉轉換後得到相位差，再使用選定濾波器之

N 個有限脈衝響應 (finite impulse response) h 所組成的矩陣 H 估計相位差的線性關係，如 2-8 式所示。Feroosh 使用一般相互驗證法 (generalized cross validation, GCV) 求出 H 的最佳解，則可用 H 推算出相位差的線性關係，即兩張影像間的水平量。

$$h = [h_1, \dots, h_{2L+1}]^T$$

$$H = \begin{bmatrix} h^T & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & h^T \end{bmatrix} \quad (2-8)$$

2.2 前景與背景分割

前景物體與背景的分割，有許多不種的方式，其中一種普遍的方式就是利用水平量的不同來區別前、背景。如果在攝影機靜止的情形下，背景物體的水平量視為零，有移動的東西都將被視為前景物體；但在攝影機本身就有移動的狀況下，就必須計算出全部景物的水平量，再行判斷。

2.2.1 靜止的背景

在固定式、無水平旋轉的攝影機所拍攝出來的影像，其背景若為靜止，則其前景物體可建立各自的移動機率模型。基於條件隨機場 (conditional random field, CRF)，Wang 及 Ji 利用動態條件隨機場 (dynamic conditional random field, DCRF) 建立前景物體的移動機率模型 [8]。其作法是在影片中取 k 個連續的影格，每個前、背景分割區域 s_k 的事後機率 (posterior probability) 可以用條件隨機場來表示，並利用條件隨機場遵守馬可夫模型 (Markov model) 的特性，將 s_k 的事後機率分佈及狀態改變機率 (state transition probability) 用 2-9 式估計：

$$\begin{aligned}
p(s_k | z_k) &\propto \exp \left\{ - \sum_{x \in X} \left[V_x(s_k(x) | z_{1:k}) + \sum_{y \in N_x} V_{x,y}(s_k(x), s_k(y) | z_{1:k}) \right] \right\} \\
p(s_{k+1} | s_k) &\propto \exp \left\{ - \sum_{x \in X} \left[V_x(s_{k+1}(x) | s_k(M_x)) + \sum_{y \in N_x} V_{x,y}(s_k(x), s_k(y)) \right] \right\}
\end{aligned} \tag{2-9}$$

其中， $V_x(s_k(x) | z_{1:k})$ 是單一像素的位能場（one-pixel potential）， $V_{x,y}(s_k(x), s_k(y) | z_{1:k})$ 則是兩像素的位能場（two-pixel potential），分別代表已知影像的單一像素觀察值及各像素間的互動關係。 $z_{1:k}$ 是 k 個影格的影像組成的序列， N_x 是像素 x 的鄰近像素集合， M_x 則是像素 x 在連續影格中，存在對應關係的集合。換句話說，Wang 利用單一像素隨時間的變化，以及此單一像素與鄰近隨時間變化的關係，推算出分割區域的事後機率；再利用已有的分割區域隨時間的變化趨勢，推算出下一個時間點的分割區域位置的機率，也就是狀態改變機率。最後即可從事後機率分佈和狀態改變分佈，得到分割區域及位移變化的機率。

2005 年，Wang 等人[9]又再提出使用貝式網路建立圖像模型（graphic model），用以分割各前景物體與背景區域。他們將機率參數分成移動向量 d_k 、像素亮度分割區域 s_k 、物體分割區域 z_k ，以及影片中三個連續影格 g_{k-1} 、 g_k 、 g_{k+1} ，再利用貝氏定理（Bayes' rule）表示機率分佈，如 2-10 式，加上空間、時間的條件限制，即可估計出機率分佈。

$$\begin{aligned}
&p(d_k, s_k, z_k, g_{k-1}, g_k, g_{k+1}) \\
&= p(g_{k-1}, g_{k+1} | g_k, d_k) p(g_k, s_k) p(s_k) p(d_k | z_k) p(z_k | s_k)
\end{aligned} \tag{2-10}$$

有別於針對物體建立移動機率模型，影像中物體的平移量也可利用影像平移造成的相位差來求得。Briassouli 及 Ahuja 分別在 2004、2005、2006 年發表在頻率空間分析影像中相位關係，來分割影像中物體的研究[10,11,12]。將 N 個連續影格中物體分為 M 塊區域及靜止的背景，則第 k 張影像經過傅立葉轉換 X_k 可以表示成各物體的傅立葉轉換 $S_1 \cdots S_M$ 總合，如 2-11 式所示：

$$\begin{aligned}
X_1 &= S_b + S_1 + S_2 + \dots + S_M + V_{noise,1} \\
X_k &= S_b + S_1 e^{-j(k-1)\omega^T r_1} + \dots + S_M e^{-j(k-1)\omega^T r_M} + S_{un}^{1,k} - S_{occ}^{1,N} + V_{noise,k}
\end{aligned} \tag{2-11}$$

其中 S_b 表示背景區域經過傅立葉轉換後的結果， V_{noise} 是影像中的雜訊， $S_{un}^{1,k}$ 、 $S_{occ}^{1,k}$ 則分別表示影像中物體遮蔽及被遮蔽的區域經過傅立葉轉換的結果。但是 2-11 式中包含一個假設：前景物體的振幅並無改變，也就是說 $S_1 \dots S_M$ ，在每一張影像中都必須存在且一致。如此，也表示前景物體不能有形體上的改變，因為形體上的改變同時會改變經過傅立葉轉換後的振幅，而無法使用 2-11 式描述頻率空間的影像。

如果影像中的前景物體沒有任何改變，則在傅立葉轉換後，振幅也不會改變。而影像中物體的平移關係，會在頻率中的相位中呈現，也就是 2-11 式中的指數項。用矩陣方式表達，則可寫成 2-12 式。

$$X = AS + V_{noise} + V_{bck} \tag{2-12}$$

其中 A 包含了移動資訊，每一個列可表示成 $A_k = [1, e^{-j(k-1)\omega^T r_1}, \dots, e^{-j(k-1)\omega^T r_M}]$ 。 $S_{un}^{1,k}$ 和 $S_{occ}^{1,k}$ 可視為背景估算的誤差項，並合併成一個項目，為了有別於影像上的雜訊，是以 V_{bck} 表示。

為了讓影像的分割結果更準確，Briassouli 及 Ahuja 在 2004 年的研究中加入了影像空間的遮罩。而其方法是使用最小平方法估計出 2-12 式中的 S ，經過反傅立葉轉換 (inverse Fourier transform) 後得到各分割區域的像素集合 \hat{s} ，並使用 2-13 式針對每個像素的區域分類作精確調整，再重新計算 2-12 式，經過幾個重複步驟，得到較精確的分割結果。

$$D_{mask,l}(x, y) = \frac{D_{max,l} - D_l(x, y)}{D_{max,l}} \tag{2-13}$$

其中 l 表示分割物體的標號； $D_l(x, y) = |s(x, y) - \hat{s}(x, y)|$ 為原像素值與分割物體的差距，而 $D_{\max, l} = \max_{x, y} \{D_l(x, y)\}$ 則是 $D_l(x, y)$ 的最大值。如果 $D_l(x, y)$ 越接近 0， $D_{\max, l}(x, y)$ 越接近 1，則像素點就會被歸類到第 l 個分割物體。

之後，Briassouli 及 Ahuja 改用交錯相關性 (cross-correlation) 來修正在頻率空間中使用最小平方法估計的位移量。求得各分隔區域的像素集合 \hat{s} 後，對每個像素估計結果 $s(x, y)$ 計算與原始影像中對應點 $f(x, y)$ 的附近相鄰區域 $N_b(x, y)$ 的交錯相關性 $c_b(x, y)$ ，建立相關性地圖 (correlation map)，並依照 2-14 式作為分類依據：

$$Prob(x \in object) = P(c_b > \eta) = Q\left(\frac{\eta - \mu}{\sigma}\right) \quad (2-14)$$

其中 $Q(x)$ 表示一個常態機率分布，如 2-15 式所示：

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \quad (2-15)$$

在實驗中，Briassouli 及 Ahuja 取 $P(c_b > \eta) = 0.1$ 作為判斷此像素是否屬於一個移動物體的依據。

2.2.2 移動中的背景

Wang 和 Adelson 將有背景平移量的影片，以平移量的不同，分成不同的圖層 (layers) [13]。相對於靜止背景的分析，背景的移動會影響到影片中物體的分析，將背景與物體分成不同的圖層，則可一併分析背景的分析。在計算影像中的光流 (optical flow) 後，Wang 和 Adelson 將影像分為幾個小區域，並針對小區域使用仿射運動模型 (affine motion model) 重複的測量平移量，如 2-16 式與 2-17 式所示，包含平移、旋轉、變焦等六個參數。

$$V_x(x, y) = a_{x0} + a_{x1}x + a_{x2}y \quad (2-16)$$

$$V_y(x, y) = a_{y0} + a_{y1}x + a_{y2}y \quad (2-17)$$

同樣使用圖層的方式分隔前、後景物體，Torr、Szeliski 與 Anandan 使用貝氏機率模型 (Bayesian model) 建立各圖層的模型[14]，並使用 EM 演算法估算事後機率分布。此外，Xiao 與 Shah 則是使用圖切割 (graph cut) 來找出不同的動態圖層[15]。



第三章 背景區域的比對及平移量計算

為了要計算背景區域的次像素平移量，首先我們兩張影像中找到一個對應的背景區域，再對此區域計算次像素的平移量。我們使用影像中區域的相似性進行比對，找出最相似的對應背景區域。再利用對應的背景區域的相位差估計次像素平移量。為了簡化計算，我們利用行與列的平均值將影像非成兩個一維度的數列計算。

為了避免混淆，我們將針對影像各行來說明。圖 3-1 是我們利用影像各行的平均值估計垂直方向的平移量。首先利用平均值與變異量對兩張輸入影像進行背景區域的比對 (correspondence using mean and variance)，再對背景區域進行頻率空間相位線性分析 (linear phase estimation)，得到垂直方向的次像素平移量。

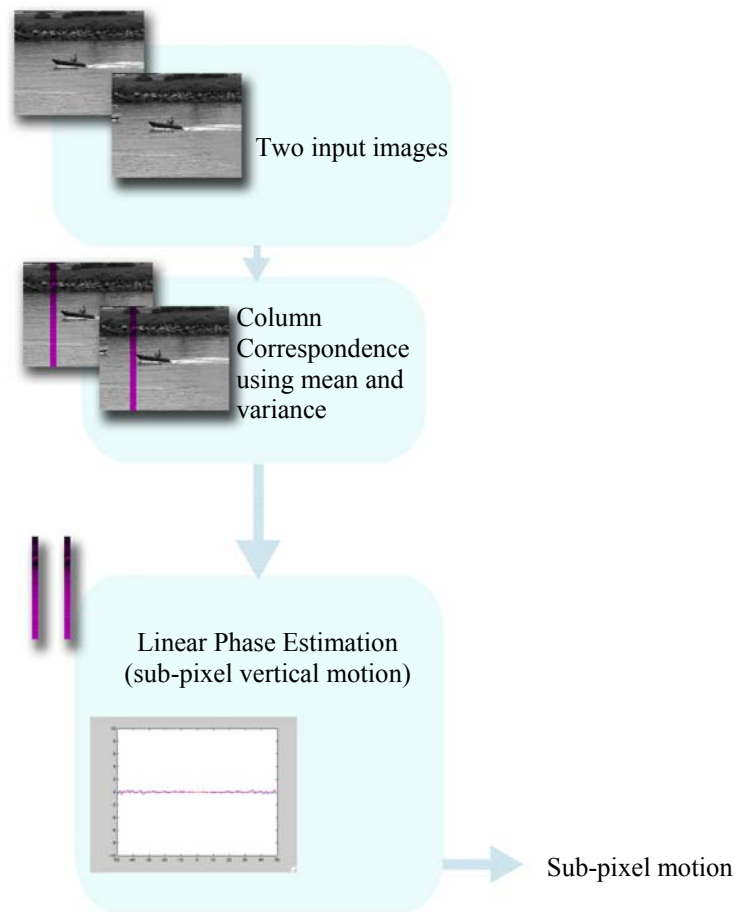


圖 3-1 利用行的平均值計算垂直方向平移量計算的流程圖。

接下來我們將說明「以平均值與變異量做對應區域的比對」及「頻率空間相位線性分析」的方法。在第四章中將利用此兩方法分別計算出垂直與水平兩個方向的平移量後，再進行背景區域的分割，並合併垂直與水平方向的結果。

3.1 以平均值與變異量做對應背景區域的比對

此步驟中，我們希望可以得到兩個重要的資訊供之後使用。一是純背景區域，也就是在兩張影像中，同時都不存在前景移動物體的區域。一是利用各區域的相對應關係粗略估算前景與背景區域的平移量。

3.1.1 背景區域的粗略估計

在本論文中，我們假設影像中的背景，在實際空間中是不會移動、沒有變化的區域。雖然整體的影像中，因為攝影機的移動使得全部的影像內容都有移動，所謂的背景無移動是指在相對於兩張影像忽略相機造成全域的平移量後，沒有移動的部份。據此，我們用比對的方式找出兩張影像中，完全相同的帶狀部份，而且其中不包含相對於背景有移動的物體，或因為移動物體所造成的遮蔽區域。

首先我們對兩張相連續影像簡略分析其物體間的相對應關係。為了簡化計算量，僅使用各行的平均值，使二維影像簡化成一維陣列。使用平均值可以減少影像整體移動所造成的邊界效應，因為將一整行取平均值時，在此行中的背景少量移動，可以因平均而簡化。

圖 3-2(a)為前後兩張影格，以紅色標示的垂直帶狀區域並不包含前景物體，僅包含背景的區域。而圖 3-3(a)中的紅色帶狀區域，因為包含了前景的移動物體，所以圖 3-3(b)中各行的平均值在兩張影像中呈現比圖 3-2(b)中紅色帶狀區域大的差異。圖 3-2(c)與圖 3-3(c)是兩張影像各行的變異量。利用圖 3-2 與圖 3-3 各行平均值變化的差異，假設 t 時間點的部分背景區域可以在 $t+1$ 時間點中找到對應的

背景區域，背景區域各行平均值的變化會比包含前景移動物體的區域來的小，如此就可以找到不包含前景移動物體的純背景區域。

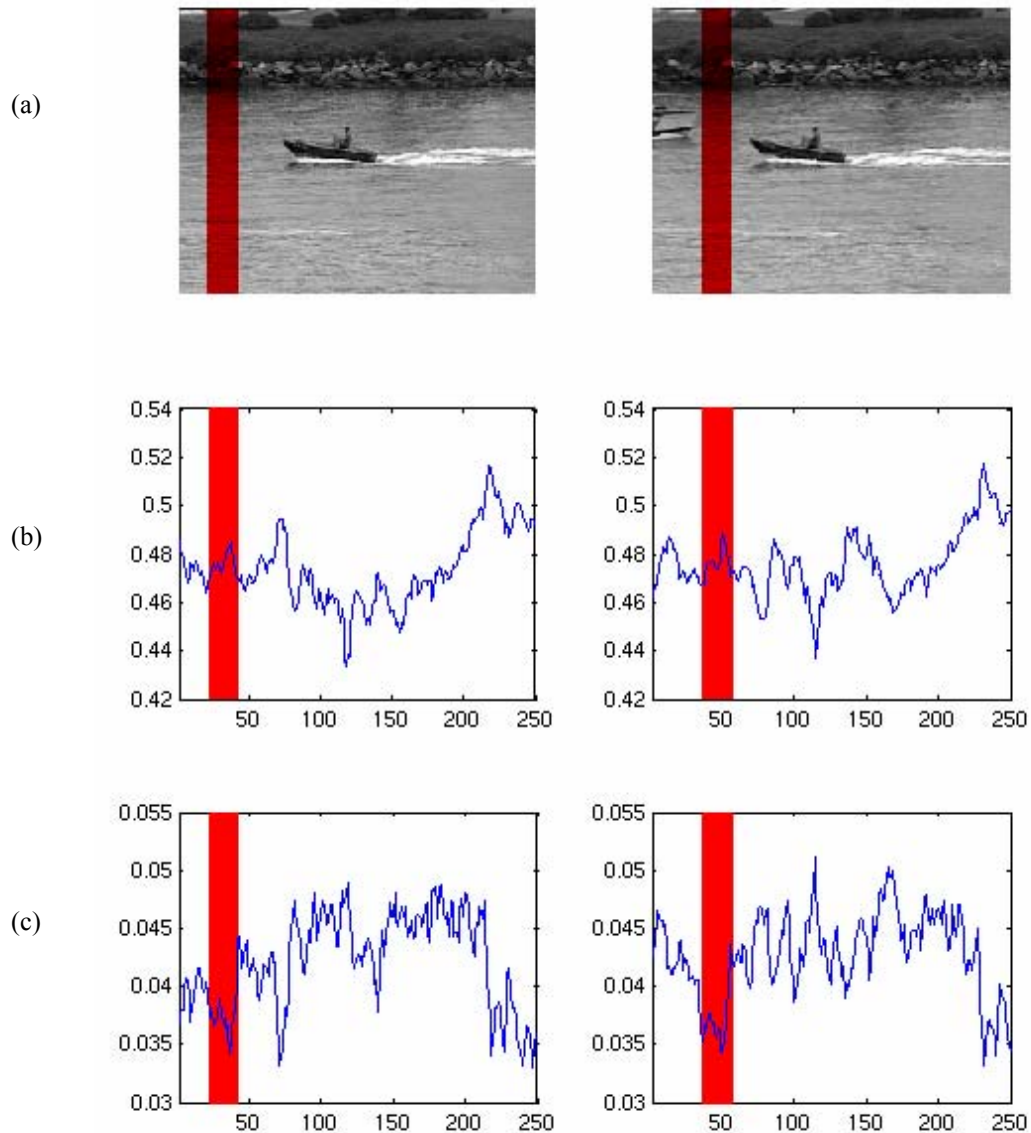


圖 3-2 (a) 前後兩張影像。(b) 兩張影像各行的灰階平均值。(c) 兩張影像各行的灰階變異量。紅色區域為影像中的純背景區域。

另外在後續的討論中，我們再加上帶狀區域中每一行的變異量來代表此行的灰階變化程度。如果在一平順無變化的區域時，因為沒有特徵點來作為判斷平移量的區域，無法斷定平移量。例如在完全平坦光滑的結冰湖面溜冰時，單看湖面的冰時，因為完全相同，所以會讓溜冰的人忽略自己的位移變化，而必須靠著其他非湖面冰塊的物體，才有辦法判斷自己的速度及位置。如果在圖 3-2(a)中，我

們忽略的上方的沿岸部分，如圖 3-4(a)所示，單靠觀察幾乎沒有灰階變化的水面，是無法判斷出背景的水平移量。縱使圖 3-4(b)中水面部分各行平均值相差微小，但因為圖 3-4(c)可以看到水面的變異量太低，造成無法判斷的情況。

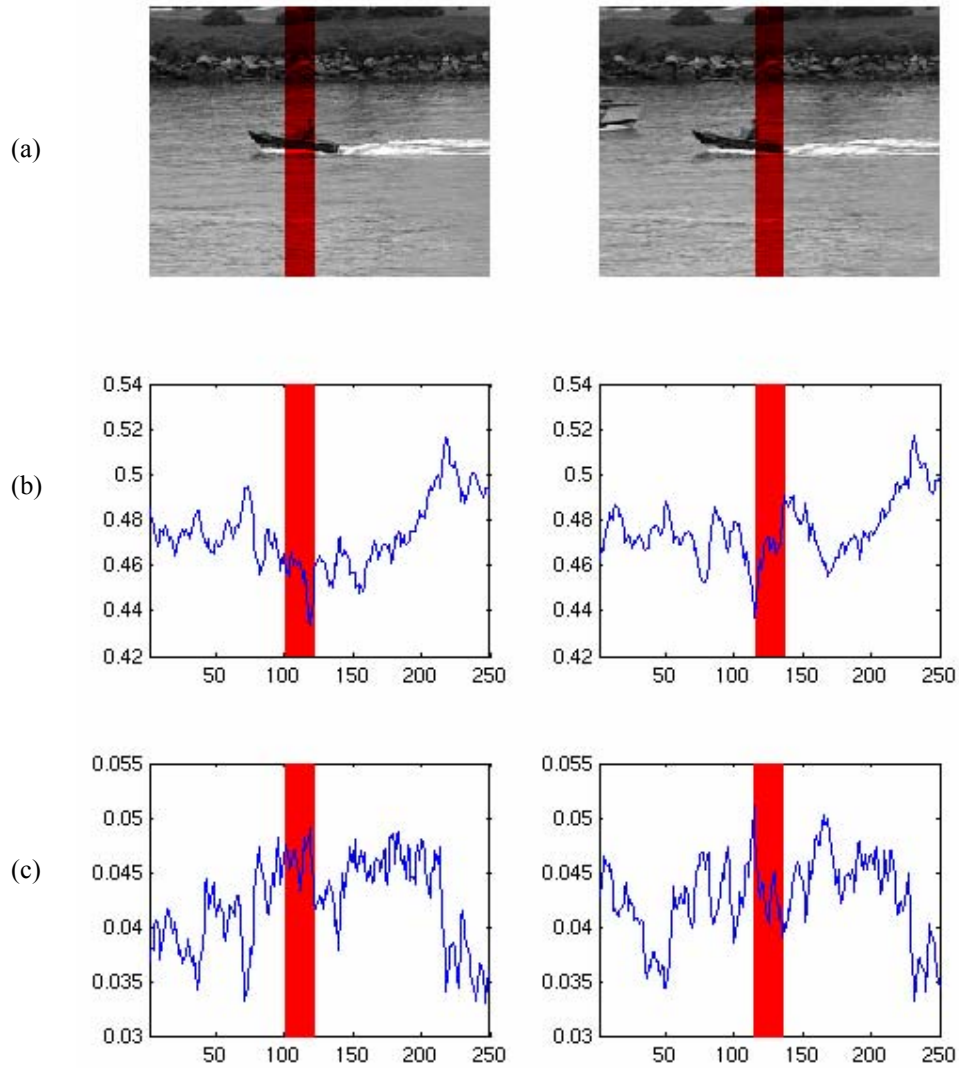


圖 3-3 包含移動物體的行在前後兩個影像中呈現不一樣的平均值。(a) 前後兩張影像。(b) 兩張影像各行的灰階平均值。(c) 兩張影像各行的灰階變異量。

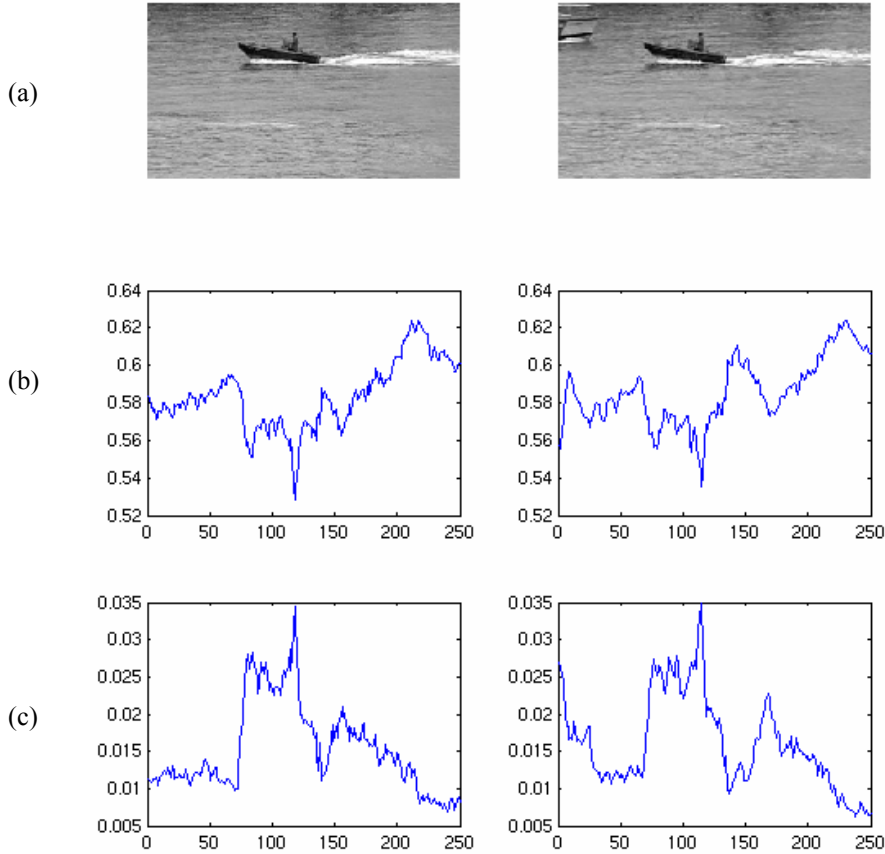


圖 3-4 當選擇的區域的變異量過小時，容易造成的錯誤。(a) 前後兩張影像。(b) 兩張影像各行的灰階平均值。(c) 兩張影像各行的灰階變異量。

而後，各個帶狀區域的比對是依照 Fisher's Linear Discriminant (FLD) 概念來評分的：

$$C(X_t^{k,m}, X_{t+1}^{k,n}) = \frac{1}{k} \sum_{i=0}^{k-1} \left(\frac{(\mu_t^{i+m} - \mu_{t+1}^{i+n})^2}{(\sigma^2)_t^{i+m} + (\sigma^2)_{t+1}^{i+n}} \right) \quad (3-1)$$

$$X_t^{k,m} = [\mu_t^m \mu_t^{m+1} \mu_t^{m+2} \dots \mu_t^{m+k-1}]$$

$$X_{t+1}^{k,n} = [\mu_{t+1}^n \mu_{t+1}^{n+1} \mu_{t+1}^{n+2} \dots \mu_{t+1}^{n+k-1}]$$

其中， $X_t^{k,m}, X_{t+1}^{k,n}$ 分別為前後兩個影格中寬度為 k 的帶狀區域，也就是包含從第 m 行到 $m+k-1$ 行及第 n 行到第 $n+k-1$ 行的平均值的一維陣列， μ^i 為影像中第 i 行的平均值， $(\sigma^2)_t^i$ 為影像中第 i 行的變異量。使用帶狀區域是為了考量到影像之間的相對關係，單使用一行的平均值來比對，其資訊量過於簡略，且會完全忽略了各行間之關係。

為了能夠找出不同寬度的對應區域， k 在比對的過程中允許被改變。帶狀區域的寬度越大，則所包含的平均值越多，這表示有更多的資訊量提供比對使用。另外，也必須考量到移動物體之間、移動物體與影像邊界中的背景區域也有可能很小，所形成的帶狀區域很窄，所以必須考量到不同的 k 值。我們採用的 k 值範圍為 11 像素至影像寬度的一半。圖 3-5 紅色標示的區域比圖 3-2 更寬，雖然紅色標示區域都不包含前景，我們允許不同寬度帶狀區域，進行比對，來找到最相似的背景區域。在此階段，兩張連續影像中經比對後被判定為最相似的區域，可視為純背景區域，此即為其背景區域的粗略估計。

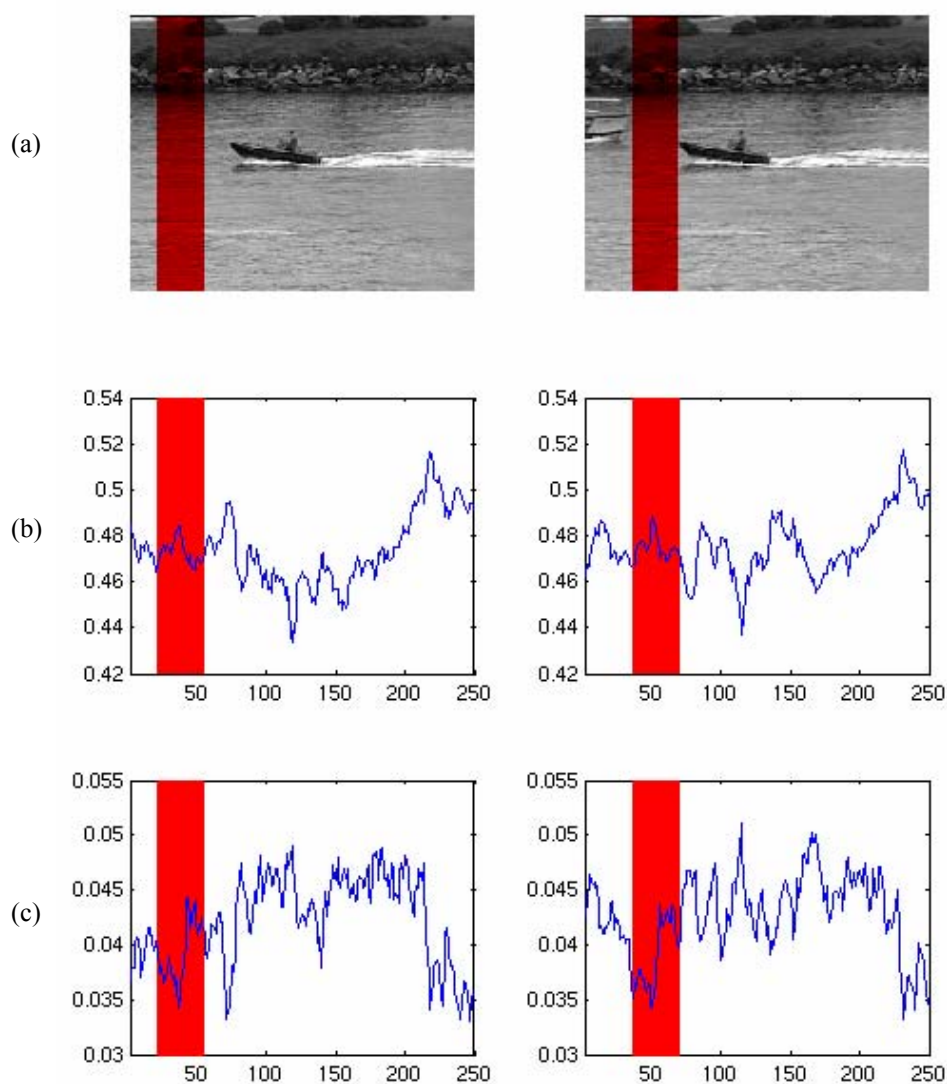


圖 3-5 比 圖 3-2 有更寬的純背景區域被以紅色標示。(a) 前後兩張影像。(b) 兩張影像各行的灰階平均值。(c) 兩張影像各行的灰階變異量。

3.1.2 背景平移量的粗略估計

我們以 3-1 式中的 $C(X_t^{k,m}, X_{t+1}^{k,n})$ 為評分標準，計算各個對應的帶狀區域的相似度，找出其中相似度最高的對應帶狀區域，如 3-2 式所示。因為此組帶狀區域是影片中前後兩張影像相對應的背景區域，所以這兩個相對應帶狀區域的位置差值 $n-m$ ，可作為背景平移量的粗略估計。

圖 3-6 中兩張影像經過比對，將最相近區域標示為紫紅色，兩條垂直帶狀區域的水平位置差值，即為背景的水平方向平移量的粗略估計。在實驗中，為了省略計算，我們僅使用寬度為偶數的帶狀區域進行比對，並限制 k 最大為四分之一的影像寬度。

$$(X_t^{k,m}, X_{t+1}^{k,n}) = \arg \min_{X_t^{k,m}, X_{t+1}^{k,n}} (C(X_t^{k,m}, X_{t+1}^{k,n})) \quad (3-2)$$

(a)



(b)



圖 3-6 (a)、(b)兩張影像為影片中前後兩張影格。紫紅色區域為垂直帶狀區域中，經過比對之後被判定最相似的兩對應區域。

雖然這兩條帶狀區域並非全部的背景區域，因為這兩個區域存有兩個特性：一、相似性，由各行的平均值所組成的陣列相似；二、影像內容的起伏變化較大，因為變異量大，所以可當作背景的代表區域，以供之後計算時減少計算量之用。也就是說之後的步驟中對背景區域進行相位線性分析估計背景平移量時，我們只需要使用此對應區域，而不需要整張影像作計算。但因為真實影像的平移量非整數，所以影像之間不存在完全一致的整數平移關係，所以需要計算次像素平移量。

3.2 頻率空間相位線性分析

在取得了兩張影像相對應的部份背景區域後。我們將接著對此區域作較為精確的平移量分析。假設我們已經使用 3.1 節中背景區域的比對方式取得垂直的帶狀背景區域，所以僅針對帶狀區域，分析此帶狀區域方向上的影像平移量，即可得垂直方向的背景平移量。而利用頻率空間的相位變化來分析影像的平移量，將可以達到次像素的精確度。

3.2.1 平移的性質

我們可以將影像的一行轉換到頻率空間，由 $F(v) = |F(v)|e^{-j\phi(v)}$ 可得到其相位函數 $\phi(v)$ ，如果兩張影像的內容之間存在單純垂直平移關係，則可得到兩張影像相對應的兩行在頻率空間的關係式，如 3-3 式所示：

$$f(y + y_0) \Leftrightarrow F(v)e^{-j2\pi(vy_0/M)} \quad (3-3)$$

其中 $f(y + y_0)$ 為影像內容平移後的某行， $F(u)$ 為影像內容平移前的該行在頻率空間的函式， y_0 為兩張影像內容之間的水平平移量， M 為影像的寬。 $f(y)$ 與 $f(y + y_0)$ 在頻率空間的相位差為 $2\pi(vy_0/M)$ ，是一線性函數。圖 3-7 (a)、(b) 為兩張僅具有水平平移關係的影像，(c) 則是兩張影像經過傅立葉轉換的相位差關係，雖然 3-3 式說明了此相位差為一線性函數，但是因為在計算兩張影像一列

的相位時，該相位的數值只能介於 $-\pi$ 到 π 之間，所以求得的相位差，呈現一鋸齒狀。

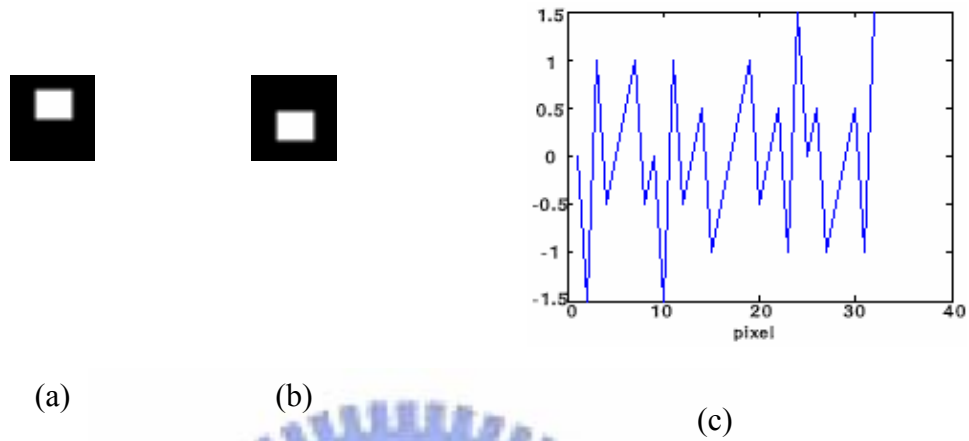


圖 3-7 (a), (b)為兩張只有垂直平移關係的影像。(c)是兩張影像中央行的相位差關係。

假設兩張只有垂直平移關係的影像分別為 $f_1(x, y)$, $f_2(x, y) = f_1(x, y + y_0)$ 。為了簡化計算，我們只看其中一行，則此行在兩張影像的關係為 $f_2(y) = f_1(y + y_0)$ ，經過傅立葉轉換（Fourier Transform）可得 3-4 式。

$$\begin{aligned} f_1(y) &\Leftrightarrow G_1(v)e^{-j\phi_1(v)} \\ f_2(y) &\Leftrightarrow G_2(v)e^{-j\phi_2(v)} \end{aligned} \quad (3-4)$$

因為 $f_2(y) = f_1(y + y_0)$ ， $G_2(v)e^{-j\phi_2(v)} = G_1(v)e^{-j\phi_1(v)}e^{-j2\pi(vy_0/M)}$ ，所以我們可以得到相位差 $\phi(v) = \phi_2(v) - \phi_1(v) = 2\pi(vy_0/M)$ 。因為經過傅立葉將影像轉換到頻率空間後，相位被侷限在 $-\pi$ 至 π 之間，相位差也同時被侷限在 $\pm 2\pi$ 之間，而形成如圖 3-7 (c)中的鋸齒狀。因此，兩張影像的相位差關係必須再經過轉換以符合線性關係的假設。還原相位差線性關係的轉換方式為：因為假設位移量不超過影像的一半，所以相位差不超過 $\pm \pi$ ，如果 $\phi'(v) - \phi'(v-1)$ 大於 π ，則將 $\phi'(v)$ 減去 2π ；反之，如果 $\phi'(v) - \phi'(v-1)$ 小於 $-\pi$ ，則將 $\phi'(v)$ 加上 2π ，直到 $\phi'(v)$ 和

$\phi'(v-1)$ 的差符合假設，在 $\pm\pi$ 之間。演算法如圖 3-8 所示，圖 3-7(c) 經過此演算法轉換後之結果如圖 3-9 所示。

```

for each  $v$  in  $\phi'(v)$ 
  if  $\phi'(v) - \phi'(v-1) > \pi$  then  $\phi'(v) \leftarrow \phi'(v) - 2\pi$ 
  else if  $\phi'(v) - \phi'(v-1) \leq -\pi$  then  $\phi'(v) \leftarrow \phi'(v) + 2\pi$ 
until  $-\pi < \phi'(v) - \phi'(v-1) \leq \pi$ 

```

圖 3-8 還原相位差線性關係的演算法。

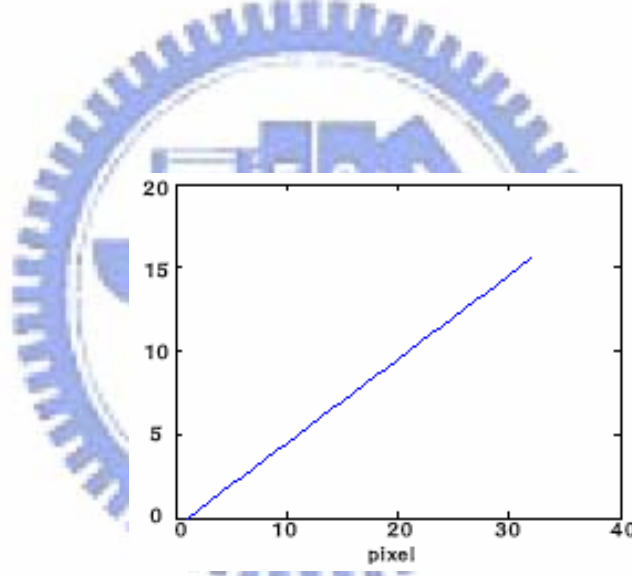


圖 3-9 圖 3-7(c) 經過轉換之結果

假設兩張影像之間的背景平移關係不超過影像的長或寬的一半。如果超過影像長或寬的一半，因為離散傅立葉轉換 (discrete Fourier transform) 的週期特性，可視為反向的移動關係，也就是說，其相位差線性函數的斜率不超過 $\pm\pi$ ，如 3-5 式所示。

$$\begin{aligned}
 -M/2 < y_0 \leq M/2 \\
 \Rightarrow -\pi < 2\pi y_0 / M \leq \pi
 \end{aligned}
 \tag{3-5}$$

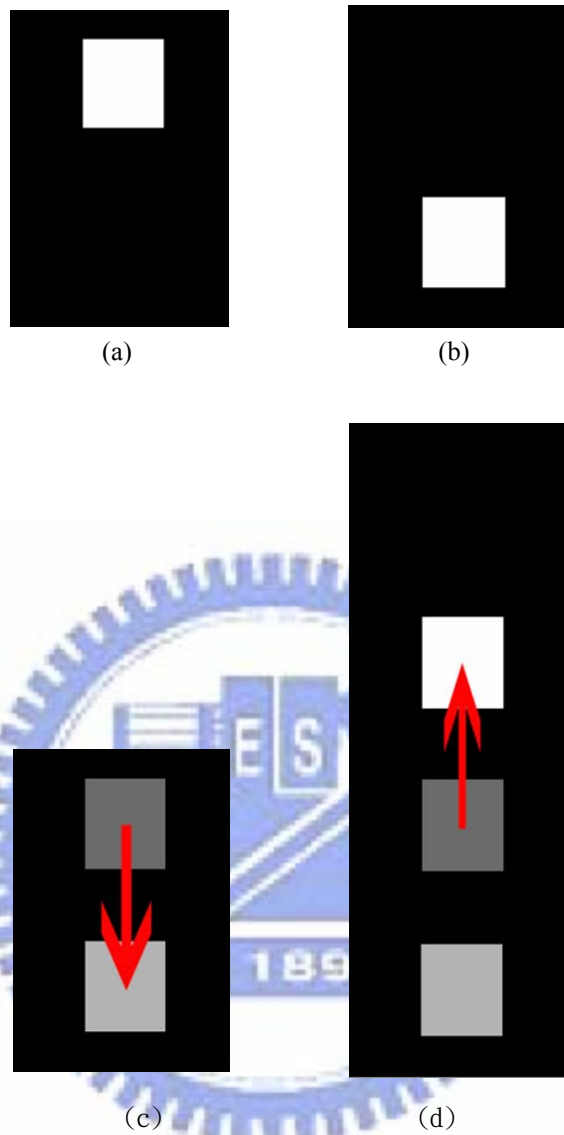


圖 3-10 (a)、(b) 為影片中前、後兩個影像。(c) 表示白色方塊由上向下移動。(d) 使用頻率空間來處理數位影像時，則可視為影像為週期性的，也就是在影像邊界外的部份，是由同樣的影像所構成，而形成另外一個白色方塊，因此，我們可以將方塊的移動視為由下向上移動。且此兩平移量總和為影像長度。

在圖 3-10 中，(a)、(b) 為兩張僅具有垂直平移關係的影像，在一般的判斷下，可以視為由上向下移動，如(c)中的紅色箭頭所示。但是在離散傅立葉轉換所造成的週期性關係之下，我們也可將移動方向視為由下向上，如(d)中的紅色箭頭

所示，並且將(d)中下方的方塊視為上方白色方塊的週期所產生的結果。所以所有的平移量，都存在另外一個反向，且與原先平移量的總和為週期大小的平移量，在本論文研究中，我們將採用值小於半個週期的解。

然而真實影像當中，並不必然有這麼單純的平移關係。在一組影像中，存在前景物體移動以及背景移動時，會產生兩種非平移關係，一是物體與物體之間、物體與背景之間被物體的遮蔽區域，在前景物體移動方向上的背景或是另一個物體將隨著時間而被擋住，物體後方的背景也隨之出現；第二個非平移關係存在於影像的邊界，因為背景的移動，在影像的邊界將也會產生多出來的背景，以及隱沒在邊界之外的區域。因此，所得結果必然不是預期般的單純線性關係。

在我們背景平移量分析的方法中，因為只考慮背景的移動，所以可以忽略物體所造成的遮蔽，而只需考慮影像背景移動所造成的邊界效應。圖 3-11 是兩張只有垂直平移關係的影像，由(a)至(b)可以視為影像由下向上移動。(a)中的紅色區域表示隨著影像平移的部分，將隱沒的邊界之外的區域；(b)中的綠色區域則是因為影像移動，而新增加的部分。

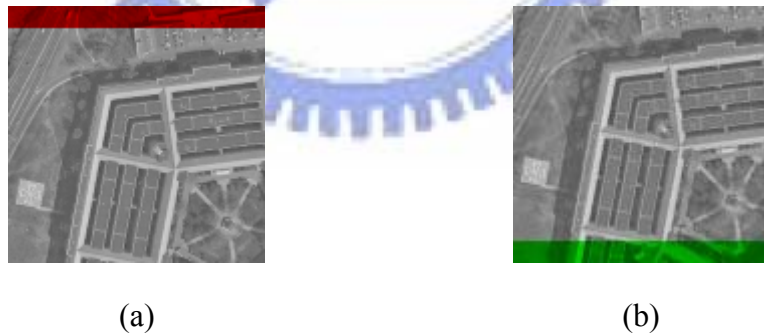


圖 3-11 兩張僅存在垂直向上平移關係的影像。(a) 紅色區域表示在另一張影像中將隱沒的部分 (b) 綠色區域表示在影像中新增加的部分

首先，將相位線性分析的背景平移量估計演算法使用在同一張真實影像的兩個區域，這兩個區域之間只有存在平移關係。影像邊界所造成的影響如圖 3-12 所示，(a)、(b)是由同一張影像中擷取出來，但是因為擷取範圍不同而尺寸相同，

所以產生了平移量。(c)是(a)、(b)兩張影像紅線標示的地方，經過傅立葉轉換後，求得的相位差結果。(d)則是將相位差經過圖 3-8 還原相位差線性關係的演算法轉換結果。雖然(d)結果接近一條直線，但是並不為純粹的直線，而有許多小幅度的曲折，因此會需要使用估計的方式找出近似的直線。

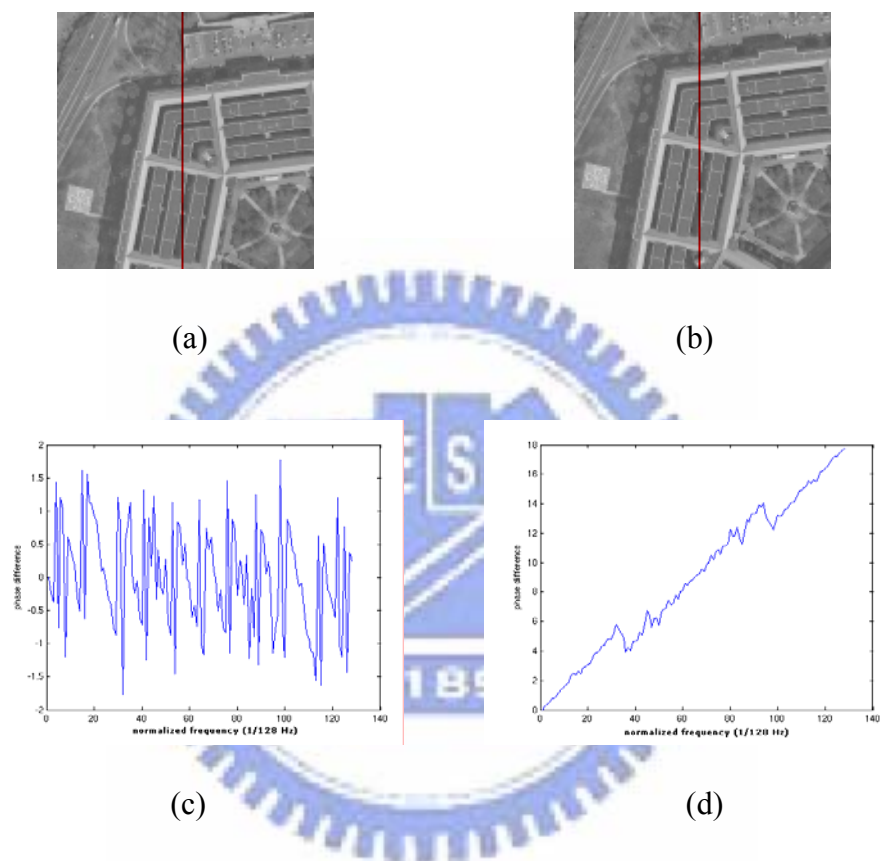


圖 3-12 (a)、(b)為兩張僅存在平移關係的真實影像。(c)為(a)、(b)影像之中央行(以紅色線段標示)的相位差關係。(d)為(c)中的折線，經過轉換後所得的近似線性相位差關係。

但是使用真實影像在比較相位時存在了更多的問題，其中最主要的關鍵是在分析影像各行在頻率空間的相位時，同時要考慮各頻率的振幅。當振幅過小時，必須考慮到相位的誤差放大問題，且當兩張影像相對應各行在頻率空間的振幅相差過大時，比較相位並不具有平移的意義。以下將針對振幅的影響及相位差線性關係的估計說明。

3.2.2 振幅比對與相位誤差

由於利用相位差分析影像某行的平移關係必須建立在此行在各頻率的振幅相同的情況，所以我們必須加上振幅的比對。然而，在影像邊界的部份，背景移動將造成影像內容的增減，並非單純平移的關係，所以振幅並不會完全一致。圖 3-13 的(a)、(b)為兩張僅具有單純水平平移關係的影像，(c)則是兩張影像中央，紅色線段標示處的振幅。我們將兩張影像的振幅重疊在一起觀察，會發現(c)中的兩色線段，並非完全重疊，儘管大致上的相似，還是有少部分的不同，也就是邊界部分影像增減所造成的影響。

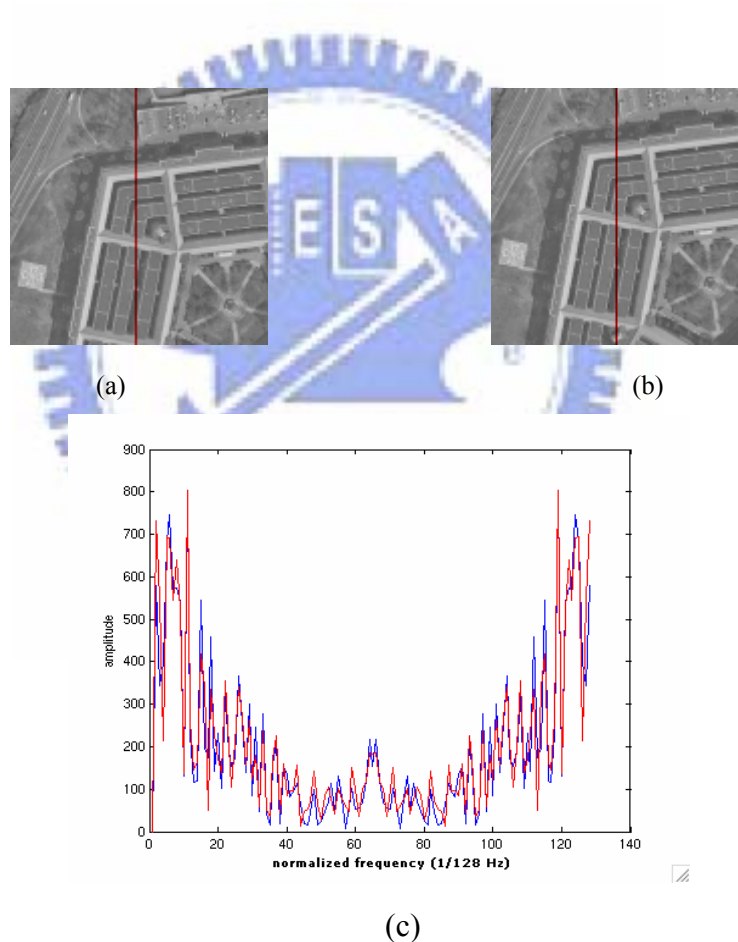


圖 3-13 (a)、(b)為兩張只有存在垂直平移關係的影像。(c) 兩色分別為兩張影像中央紅色線段標示之行的振幅，可以明顯觀察到兩曲線不完全重疊。雖然兩張影像的內容大部分是相同，背景移動造成的在影像邊界部分的增減，使振幅存在少量的差異。

為了確保相位平移關係的成立，我們對振幅的相似度做了限制。在實驗中，經過振幅的比對，我們只採用振幅的平均誤差小於最大值的 10% 的部分，如 3-6 式所示：

$$\frac{1}{M^2} [|F_1(x)| - |F_2(x)|]^2 < \alpha^2 \quad (3-6)$$

其中， $F_1(x)$ 、 $F_2(x)$ 表示兩張影像中某一行經過傅立葉轉換的結果， $|F_1(x)|$ 、 $|F_2(x)|$ 即是振幅。 M 是行的長度， α 則是設定的臨界值。在實驗中我們取 $\alpha = 0.1$ ，也就是必須符合 3-6 式，才會繼續從相位中推算平移量。

除此之外，我們還必須考慮相位的誤差。如果我們將 $F(u)$ 視為一向量，其長度為振幅，角度為相位，如圖 3-14 表示。其中(a)、(b)中的黑色箭號是兩個方向相同、大小不同的向量，紅色箭號則是在黑色圓圈的水平誤差範圍內所求得的向量。 θ_1 、 θ_2 分別代表(a)中黑色及紅色向量的角度。 θ_1' 、 θ_2' 則代表(b)中黑色及紅色向量的角度。我們可以發現 $|\theta_1 - \theta_2| < |\theta_1' - \theta_2'|$ ，因為向量的能量太小，使得相位的誤差影像更大。

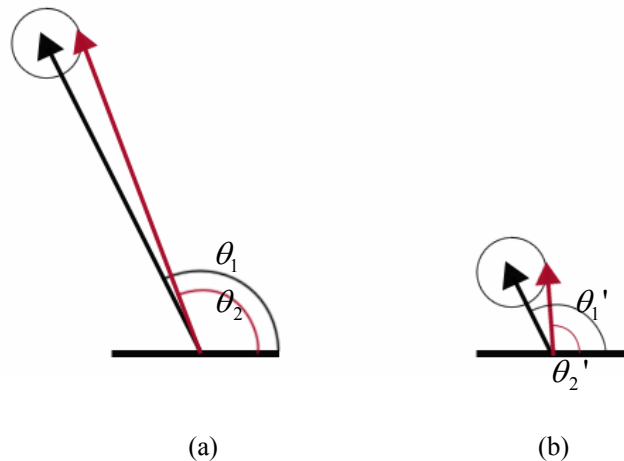


圖 3-14 (a)、(b) 中的黑色箭號表示兩個相同角度，但是大小不同的向量。在黑色箭號端點的圓則是假設的誤差範圍。紅色箭號則是估計值。

為此我們將必須考慮到每個向量振幅的大小。如果振幅過小，則此向量可能因為相位的誤差過大，使得我們的相位差的推算上產生錯誤。

所以為了解決振幅差異性的問題，我們加入了振幅的比較，針對振幅相差懸殊的部份加以排除，而不需要再進行相位差的計算。對於振幅過小的頻率，因為相位可能存在過大的誤差，所以亦不列入 PCA 的資料。

3.2.3 位移量的估計

為了判斷相位的線性關係，我們使用 Principle Component Analysis (PCA) 來找出近似的直線。而為了讓估計的結果不受到誤差以及細部變化所造成的影響，我們另外加入了兩項限制：一是排除振幅過小的向量；二是忽略高頻部分。

首先依據 3.2.1 節中所提，振幅過小使得相位誤差過大的情形，我們在判斷相位之前，先檢視各頻率的振幅大小。在實驗中，我們取最大振幅的 1% 為臨界值。根據圖 3-8 中的演算法，在判斷 $F_1(v)$ 及 $F_2(v)$ 相位差 $\phi_1(v) - \phi_2(v)$ 時，必須參考上一個頻率 $F_1(v-1)$ 及 $F_2(v-1)$ 的振幅，在此階段，如果參考的上個頻率 $F_1(v-1)$ 及 $F_2(v-1)$ 不存在，也就是被我們排除的頻率，那麼將利用 $F_1(v-3)$ 及 $F_2(v-3)$ 與 $F_1(v-2)$ 及 $F_2(v-2)$ ，以外差方式算出 $\hat{F}_1(v-1)$ 及 $\hat{F}_2(v-1)$ 來取代原本的值。

接下來必須考慮到影像中的雜訊與誤差。真實影像中難免存在雜訊，而影片的拍攝是對現實用的連續空間取樣成數位的像素，也無法避免在細部出現的誤差。以頻率空間來看，雜訊和細部都存在於高頻的部分，所以我們必須捨棄高頻處，來降低雜訊和誤差的影響。我們僅考慮在 $[-\pi/2, \pi/2]$ 區間中的頻率，也就是取週期中，頻率較低的一半來觀察相位差。

圖 3-15 是圖 3-12 (c) 之相位差的低頻率部分，在排除振幅過小頻率後，利用 PCA 所估計的結果。藍色線段表示原本的相位差關係，紅色線段則是我們求得的近似直線。原始的影像圖 3-12 (a)、(b) 是兩張從同一張影像中的不同區域擷取出來的結果，僅存在垂直方向的平移關係，我們取兩張影像最中間的一行來計算平移量。圖 3-15 的藍色線段為了方便表示，將兩張影像對應的行在振幅太小的

頻率，用外差值代替，所以看到的線段在每個頻率處都存在點，但是在計算 PCA 時，並沒有列入計算。

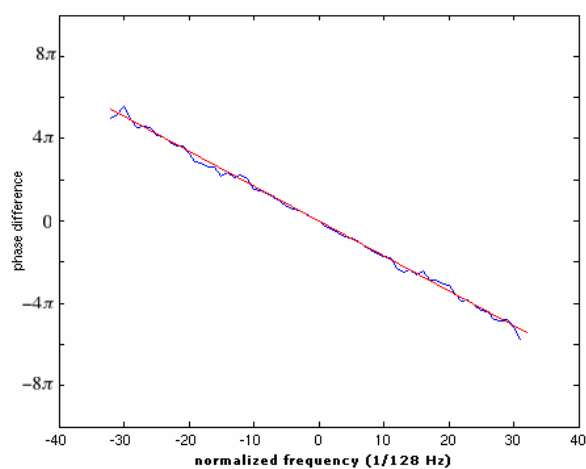


圖 3-15 圖 3-12 (c) 利用 PCA 求得相位差的線性關係。藍色線為原始的相位差的低頻率部分，紅線為 PCA 估計所得的近似直線，原點在圖的中央。

依照求出的直線，我們可以計算直線斜率 s ，其單位為徑度。依照影像寬度 M 對應徑度為 2π ，則直線斜率 s 可用 3-7 式推算出對應的像素差。即可由相位差的線性關係推算出影像的垂直平移量 y_{motion} 。

$$y_{motion} = \frac{s * M}{2\pi} \quad (3-7)$$

3.2.4 帶狀區域的位移量估計

根據以上，我們可以從兩張影像中各取一行，求得垂直的平移量。現在則將此方法套用在兩張影像中的帶狀區域上。因為先前已經利用比對的方式，在兩張影像中求得僅含背景對應的垂直帶狀區域，所以我們只需要針對帶狀區域中使用相位差估算次像素的位移量。

我們將兩張影像的對應垂直帶狀區域的每一行做一維的傅立葉轉換。接著，將兩張影像中對應各行，都分別進行相位差的計算，取得利用 PCA 估計的線性關係及誤差。

因為帶狀區域的對應關係是以像素為單位，為了確保準確性，我們除了針對對應的行進行相位差分析，也將對應行與前、後各平移一、兩個像素的行進行相位差分析。這樣可以避免在相位分析時，因為以像素為單位的比對來估計真實的連續空間位移量的誤差，而造成的錯誤。在取得各行的相位差估計及誤差後，我們選擇誤差最小，也就是相位差最接近線性關係的一組，以其估算的位移量為次像素的位移量估計。



第四章 背景區域的分割

在此階段，我們除了取得次像素的背景區域位移量外，同時也保留之前各區域的利用各行的平均值與變異量進行的背景區域比對資訊。利用這兩項資訊，我們可以將背景區域和前景區域做一個大略上的分割。

在第三章中，我們已經說明了如何完成行的背景區域比對及相位線性分析。接下來，我們則是要同時使用行與列的背景區域比對相位線性分析，來進行背景區域的分割。分割的流程如圖 4-1 所示，共可分為五個部分：利用平均與變異量進行行／列的比對（column／row correspondence using mean and variance）、相位線性分析（linear phase estimation）、與背景位移量的比較（row-band／column-band comparison）、背景區域切割的結果（background segmentation results），以及邊界效應的處理（boundary effect removing）。行／列的比對分別對垂直的平移量與水平的平移量做粗略的估計，並且在兩張影像中找出相似的對應區域。相位線性分析利用部份的背景區域找出更精確的平移量。之後我們則利用準確的平移量來和之前行／列比對出來的對應區域比較，找出垂直及水平的背景區域。最後利用此平移量來移除背景移動造成的邊界效應。以下將針對每個部分做詳細的說明。

4.1 利用平均與變異量進行行／列的比對

首先我們以連續的兩張影像為輸入，並分別對列與行計算平均值及變異量。依照章節 3.1 所提的方式，估計水平及垂直的粗略平移量，找出最相似的背景區域。同時保留比對出來的對應區域。

圖 4-2 是在輸入兩張影像後，經過分別的垂直區域及水平區域比對得到最相似的垂直與水平區域。(a)、(b)中的紫紅色區域，是經過不同寬度的垂直區域，經過比對的最相似區域，因為不包含前景物體的兩艘船，所以是單純的背景區域。(c)、(d)中的青色區域，則是最相似的水平區域，相對於單純的草地及水面，是多紋理變化的背景區域。

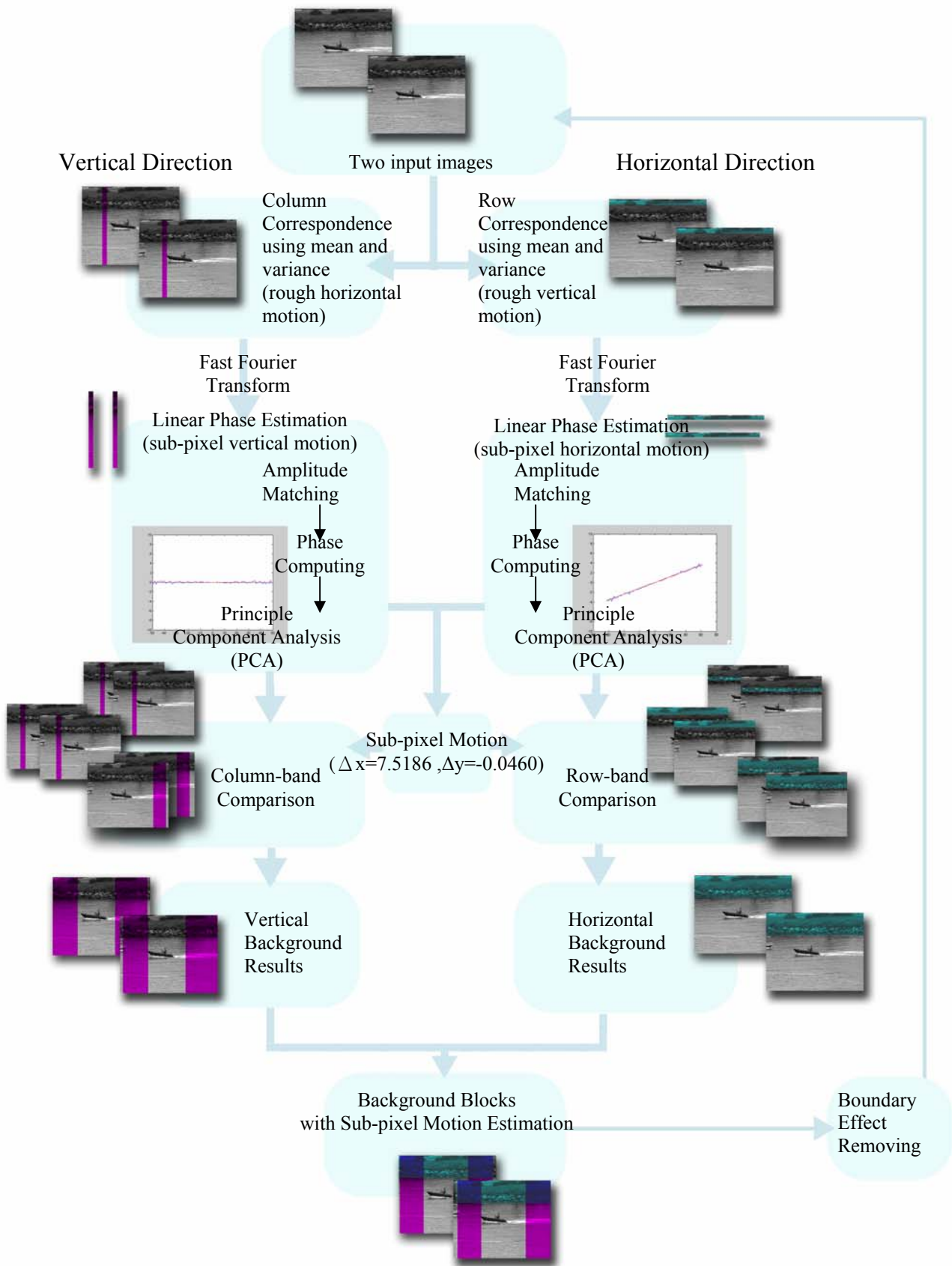


圖 4-1 前景與背景區域分割的流程圖。

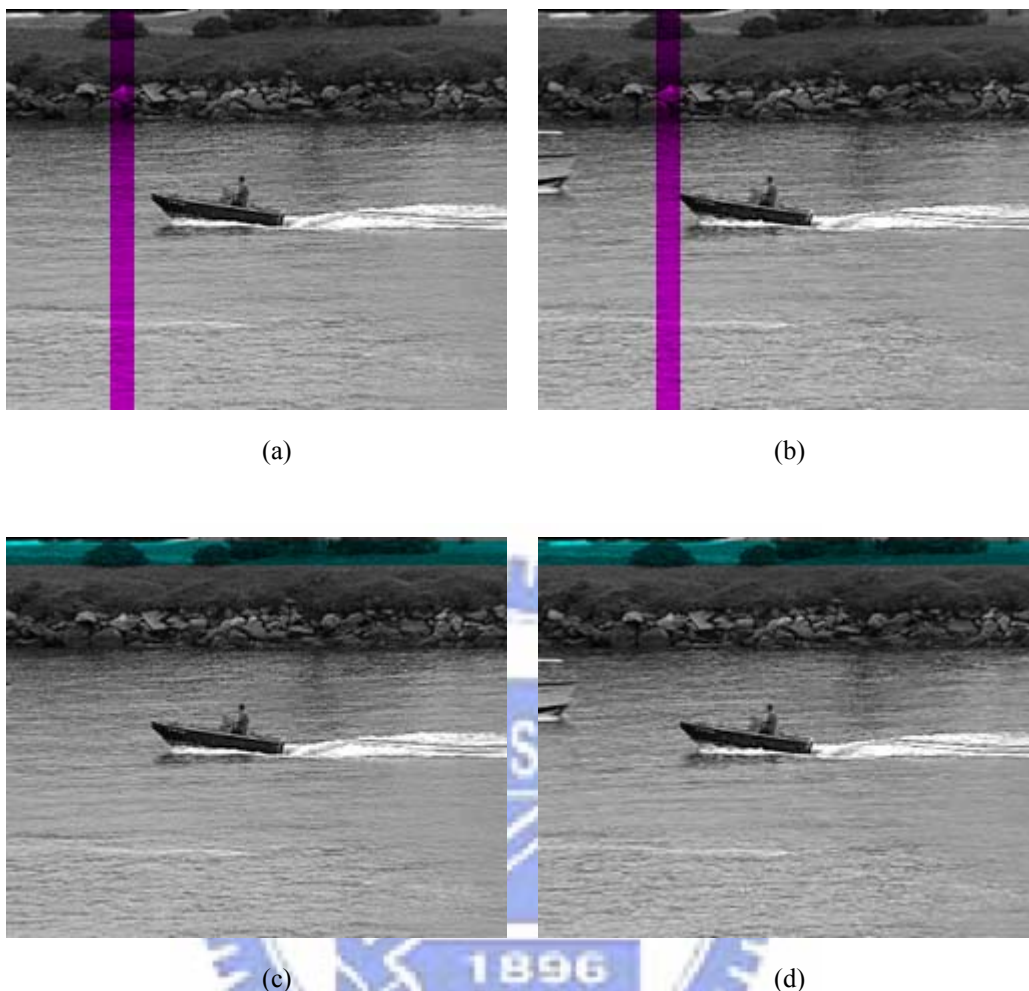


圖 4-2 經過比對，最相似的區域。(a)、(b)為影像的前後兩張影格，紫紅色標示的區域為兩張影像最相似的垂直區域。(c)、(d)是與(a)、(b)同樣的前後兩張影格，青色標示的區域為兩張影像最相似的水平區域。

4.2 相位線性分析

取得比對結果中，相似度最高的部分後，依照章節 3.2 所提的方式，利用相位差估算平移量。其中包含三個步驟：振幅的比對 (amplitude matching)、相位的計算 (phase computing) 及估算線性關係的 PCA (principle component analysis)。從比對區域中篩選出振幅相近行或列的組合，分別對這些組合進行相位差的計算，並用 PCA 估算相位差的線性關係與誤差。

我們對圖 4-2(a)、(b)紫紅色區域中的每行的組合進行相位差估計。從組合中挑選 PCA 對線性關係估計誤差最小的一組，採用其相位差為整張影像的垂直平移量。圖 4-3 的藍色線段是圖 4-2(a)第 62 列及圖 4-2(b)第 69 列的相位差關係，紅色直線線段則是利用 PCA 估計的結果。紅色直線的斜率換算以像素當單位為 0.046，也就是影像在垂直方向的平移量。圖 4-4 則是圖 4-2(c)、(d)中綠色區域的每列的組合經過相位差 PCA 分析的結果。以紅色直線的斜率計算出的影像水平平移量為 7.5182 個像素。

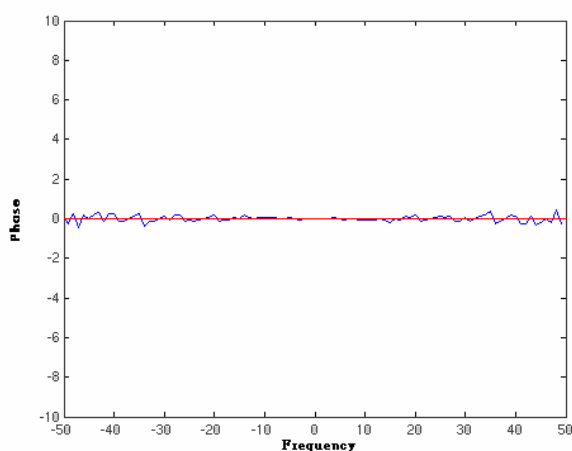


圖 4-3 圖 4-2(a)、(b)兩張影像相位差 PCA 分析結果。藍色線段為相位差關係，紅色線段為直線估計結果。

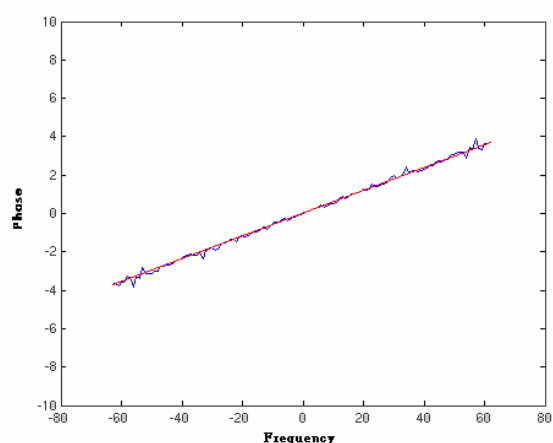


圖 4-4 圖 4-2(c)、(d)兩張影像相位差 PCA 分析結果。藍色線段為相位差關係，紅色線段為直線估計結果。

4.3 與背景平移量的比較、背景區域的估計結果

此部分包含兩個步驟：比較（column-band/row-band comparison）及整合結果（background results）。根據前兩個步驟，我們得到了次像素的背景平移量，以及利用平均值、變異量比對的結果。我們將比對的相似區域組合用次像素的背景平移量對照，找出平移量符合背景平移量的相似區域組合，這表示這些組合的平移量跟背景一致，也就是背景區域。

在對圖 4-2 進行章節 4.1 的比對時，我們除了得到最相似區域，同時利用其他比對的區域作為背景區域的搜尋資料。我們使用相位差估計的平移量選擇符合平移量的對應區域組合，因為符合背景的平移量，表示這些對應區域只包含背景，且都是背景的一部分，所以將這些區域整合起來，即為影像中的背景區域。在實驗中，我們取每個不同帶狀區域寬度的前十組比對相似度最高的作為背景區域與背景平移量進行比較。

圖 4-5 的紫紅色標示區域是經過與背景水平平移量相比，約等於背景水平平移量的對應區域。因為背景平移量是基於次像素的結果，所以我們取對應區域的平移量與背景平移量的整數部分相同或為背景平移量的整數部分加一的部分。將所有符合背景平移量的對應區域取聯集，可得垂直的背景區域。圖 4-6 則為利用背景垂直平移量選出對應的背景區域。青色標示的區域是約符合背景垂直平移量的對應區域，同樣的對這些對應區域取聯集，可得水平的背景區域。

在圖 4-6 中，水面的部分在並沒有被視為背景區域。在進行對應區域比對時，因為加入變異數的考量，我們排除了單一畫面中僅有小幅度變化的區域。所以沒有被標示為背景區域，並非就是前景區域，也包含因為單一畫面中變化太小而不考慮的部分。

4.4 邊界效應的處理

為了讓相位分析的結果更為正確，我們可以將因為影像背景移動，在邊界區域造成的增加或減少的區域排除。雖然用來做相位分析的區域已經不是整張影像，而是一個水平區域和一個垂直區域，但是這兩種區域多少都還是包含了一些許的邊界效應。這些在邊界上增加或減少的區域，因為無法完全地在另一張影像上找到對應的區域，所以會影響振幅比較和相位差分析的結果。

在得到背景平移量時，我們已經可以知道因為背景移動造成邊界效應的區域，所以我們可以將這些區域移除，再次進行背景平移量的分析。以次像素背景平移量的整數部分為準，將兩張比對影像因背景移動增加而無法找到相對應的邊界區域移除，重新進行一次比對及相位差分析，則可減少邊界效應的影響。

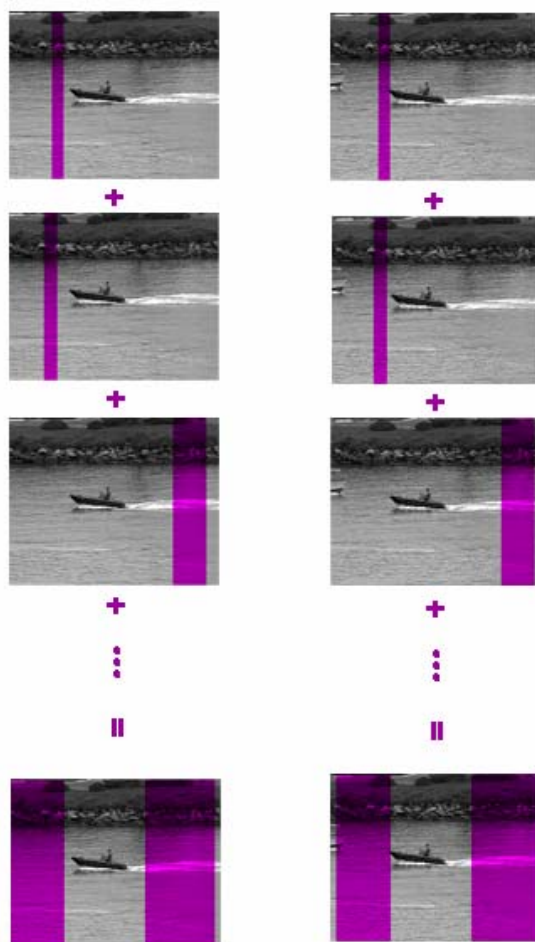


圖 4-5 利用背景水平方向的平移量找出對應的垂直背景區域。



圖 4-6 利用背景垂直方向的平移量找出對應的水平背景區域。



第五章 實驗結果與問題討論

我們分別使用了人工合成的影像以及真實的影像來進行實驗。人造影像可以設定背景平移量，來觀察相位線性分析估計的次像素平移量。真實影像除了可以觀察分割的結果，也可以觀察影像中的雜訊的影響。

5.1 人造影像分析

為了比較次像素演算法的分析結果，我們使用人造的影像來進行實驗。我們使用一張照片，先在此張照片上擷取一小部分作為第一個影像。之後平移一小部分，再在同一張照片上擷取同樣尺寸的影像。如此，我們可以假設擷取區域的平移量是這兩張影像的背景平移量。另外我們再兩張影像上，各放置一小區域的另一張照片，來模擬前景物體。

圖 5-1 為兩張截取自同一張照片的影像。模擬的背景平移量在 X 及 Y 方向為 12 及 10 個像素。經過我們的分析，估計出來的背景平移量為 12.0053 及 10.0047 像素，僅有在小數點下第三位的誤差。在背景區域的分割上，紫紅色為在水平方向估計的背景區域，青色則是在垂直方向估計的背景區域，藍色部分則是在兩個方向都判斷為背景區域。在下方的樹叢以及右方的部分郵筒，沒有被判斷為背景區域，這是因為在比對時，變異量不足的緣故。圖 5-2(a) 是圖 5-1(a) 圖每行的變異量，圖 5-2(b) 是圖 5-1(a) 圖每列的變異量，我們可以發現，在兩張變異量的圖中，靠近右方邊界的變異量比其他部分較低，所以被列入了不被考量的區域。

為了模擬次像素的背景平移量，我們使用尺寸為兩倍大的影像，並在平移後將影像縮小。在同一張照片中的不同位置擷取出兩張同樣尺寸的影像，而後，我們將兩張影像以線性的演算法將影像長寬各縮小為原本的一半，則產生兩張平移量為次像素的影像。

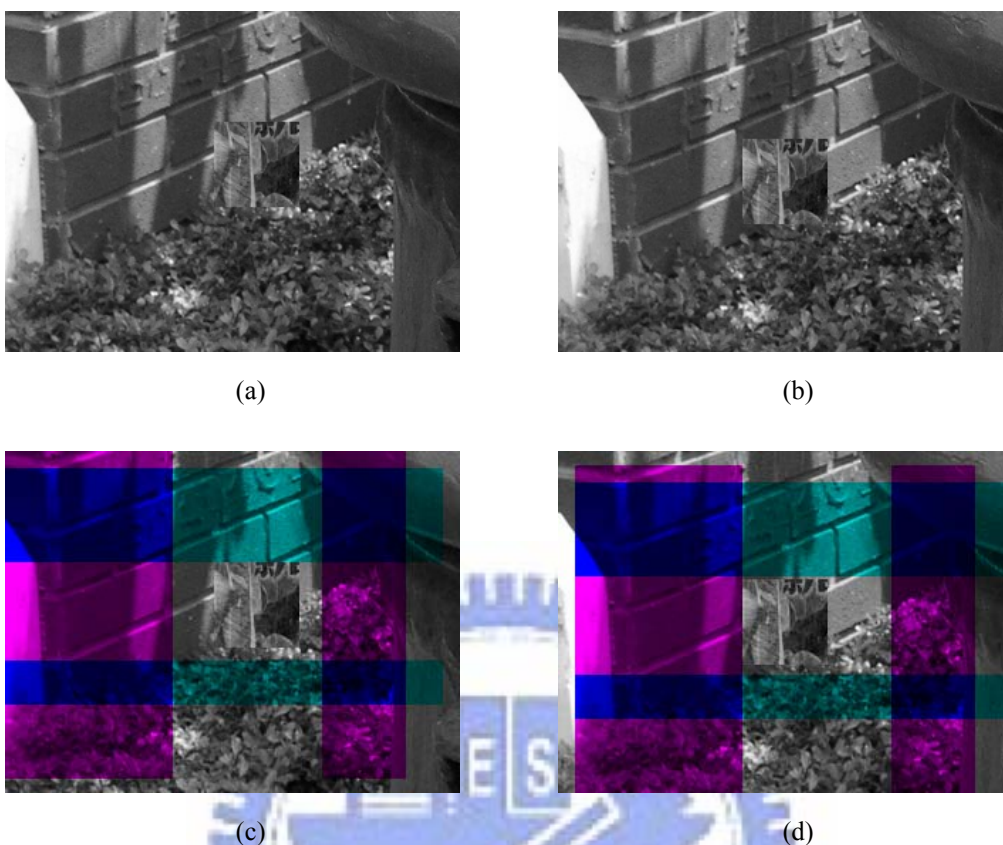


圖 5-1 (a)、(b)為兩張合成的影像，兩張影像擷取自同一張照片，其擷取位置相差在 X 方向及 Y 方向分別為 12、10 像素。(c)、(d)分別為兩張影像估計的背景區域。估計的背景平移量為 12.0053 及 10.0047。

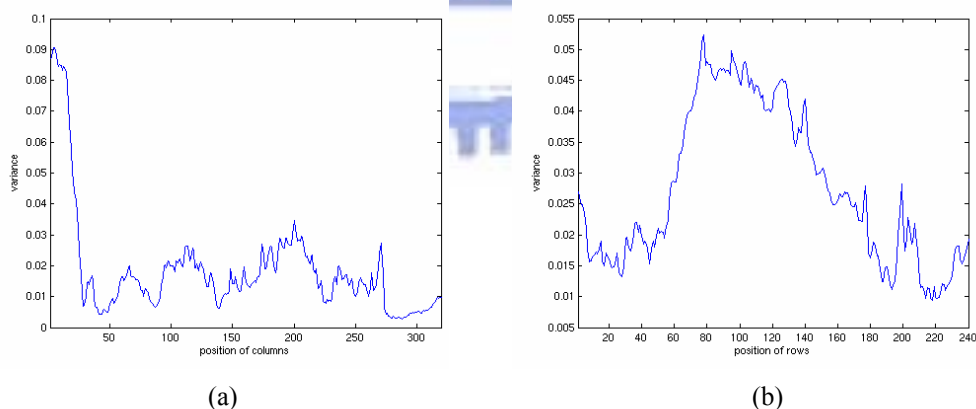


圖 5-2 (a)圖 5-1 中各行的變異量。(b)圖 5-1 中各列的變異量。

圖 5-3 (a)、(b)是模擬次像素背景平移量的兩張影像，原本是兩張尺寸為 640×480 的影像，背景平移量為 $(-5, -7)$ 。我們將影像的尺寸縮小為 320×240 ，則背景平移量變為 $(-2.5, -3.5)$ ，也就是非整數像素的平移量。我們估計出的背景

平移量為 $(-2.2047, -3.5386)$ 。在圖 5-3 (c)、(d)背景區域的分割結果中，在前景物體的右邊邊界及下方邊界有不錯的結果，但是在前景物體的上方及左方，都各有一部分的背景區域，因為黑色過多或是亮度變化過小而沒有被標示。我們可以在圖 5-4 中看到這兩個區域的變異量，較其他部分來的低。

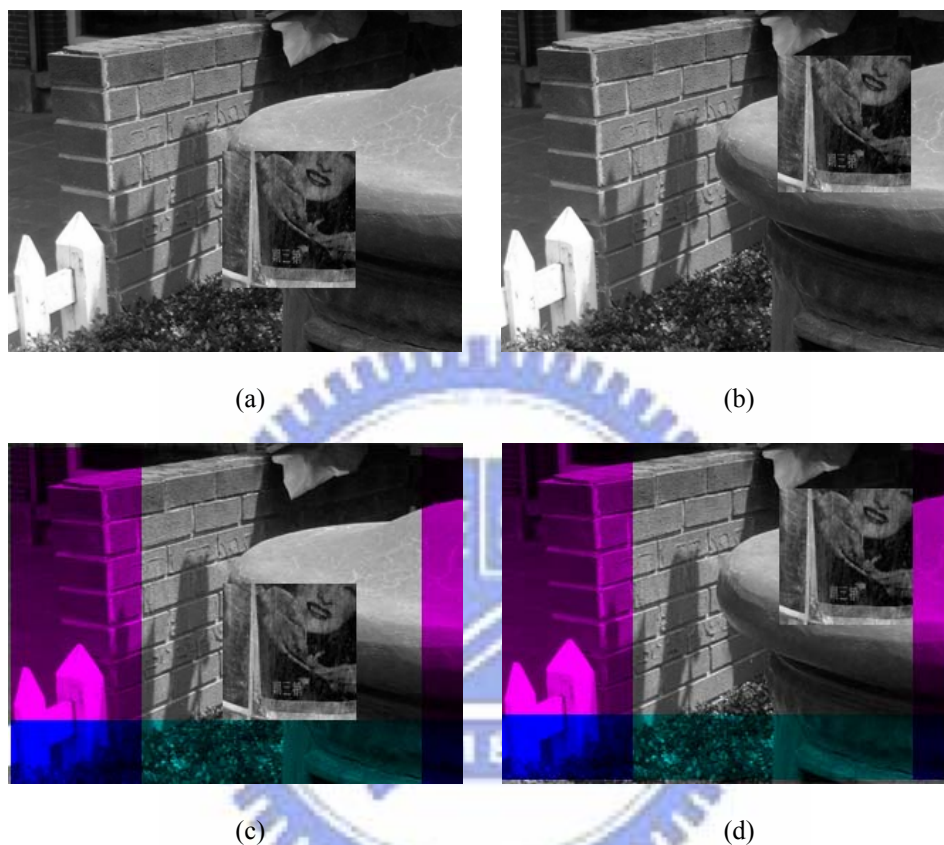


圖 5-3 (a)、(b)為兩張合成的影像，兩張影像擷取自同一張照片，其模擬的背景平移量為 $(-2.5, -3.5)$ 。(c)、(d)分別為在兩張影像所估計的背景區域。估計的背景平移量為 $(-2.2047, -3.5386)$ 。

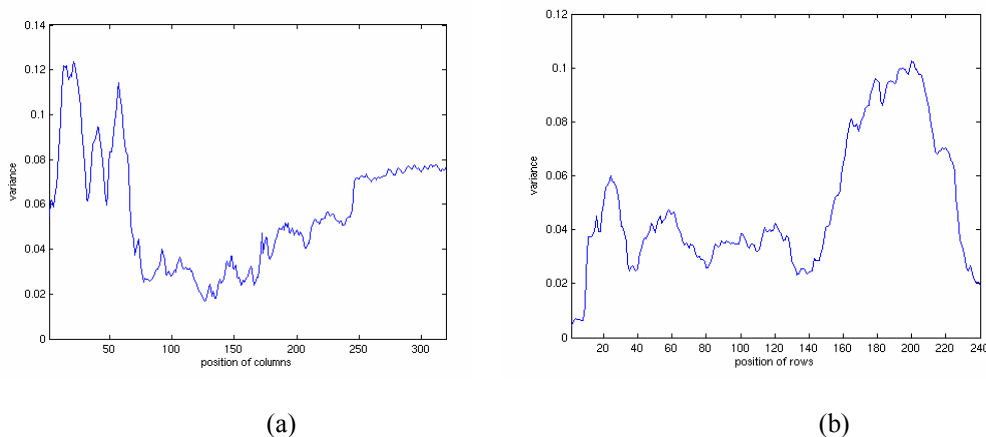


圖 5-4 (a)圖 5-3(a)各行的變異量。(b)圖 5-3(a)各列的變異量。

圖 5-5(a)、(b)則是另一組模擬次像素背景平移量的影像。背景平移量為(6.5, 3.5)，估計結果則是(6.2852, 3.5693)。圖 5-5 (c)、(d)中背景區域的分割，大部分的背景區域都有被標示出來，在右側樹木較多的部分，容易因為變異量較低而不列入背景。前景物體的左邊邊界及下方邊界都有不錯的分割結果，上方區域則有部分的前景區域被列入背景區域。因為前景物體判斷錯誤的部分和被其遮蔽的背景區域，其水平方向的平均值相近，造成判斷的困難。

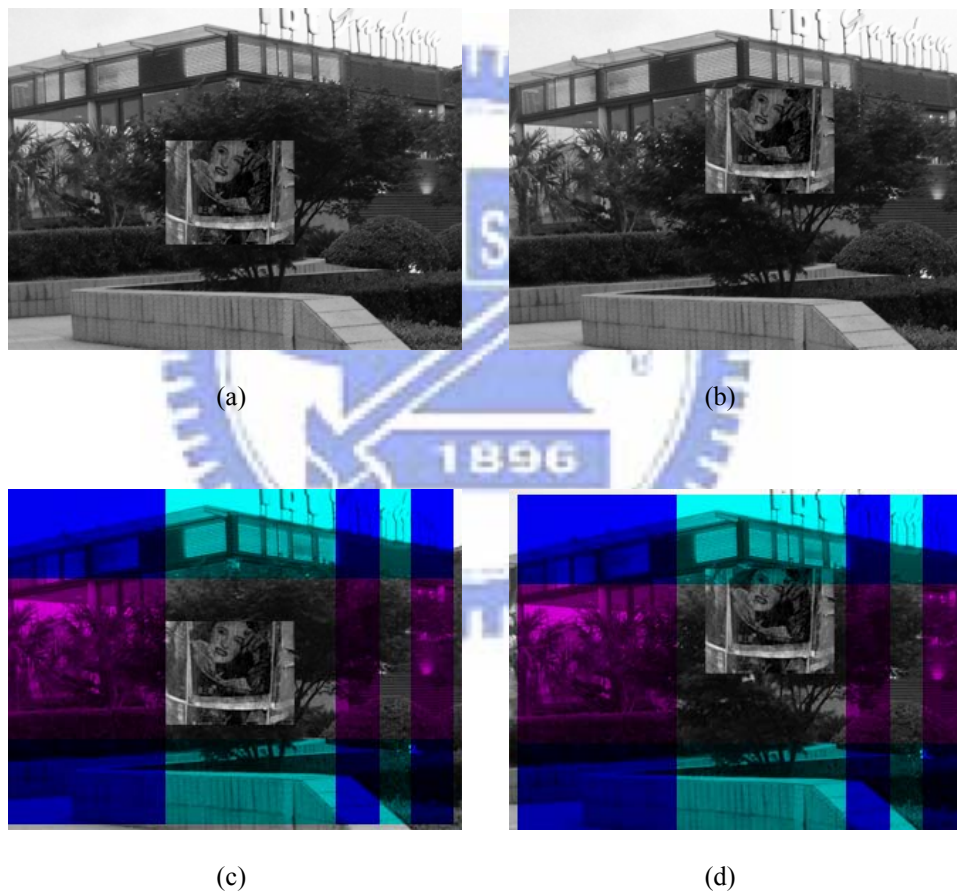


圖 5-5 (a)、(b)為兩張合成的影像，兩張影像擷取自同一張照片，其模擬的背景平移量為(6.5, 3.5)。(c)、(d)分別為在兩張影像所估計的背景區域。估計的背景平移量為(6.2852, 3.5693)。

5.2 自然影像之分析與模擬

接下來，我們使用連續的影像來進行實驗。從影片中擷取出兩個影格，來判斷背景平移量及背景區域。因為無法從影像中得知真正的背景平移量，所以 5.2.1 節僅觀察背景區域分割的結果，5.2.2 節則對次像素的平移量做進一步的討論。

5.2.1 自然影像的背景區域分割

圖 5-6 是 Coastguard 影片中的兩張影格，以及背景區域的分割結果。水平方向的背景區域分隔，僅有影像上方的沿岸及草地被標示出來，其他包含船隻和水面的部分，則沒有被列入背景區域。船隻前景物體，有各自的平移量；水面則是因為變異量太小的緣故，沒有被列入考量。在垂直的分割區域上，可以看出中央的小船後方存在剛剛濺起的水花，變化比更後方的水花來的大，所以不列入背景區域，而左方船隻，則有前方一小部分被列入背景區域，因為船身偏白色與水面本來顏色相差不多，所以在列的平均值上變化不大，而被列入背景區域。

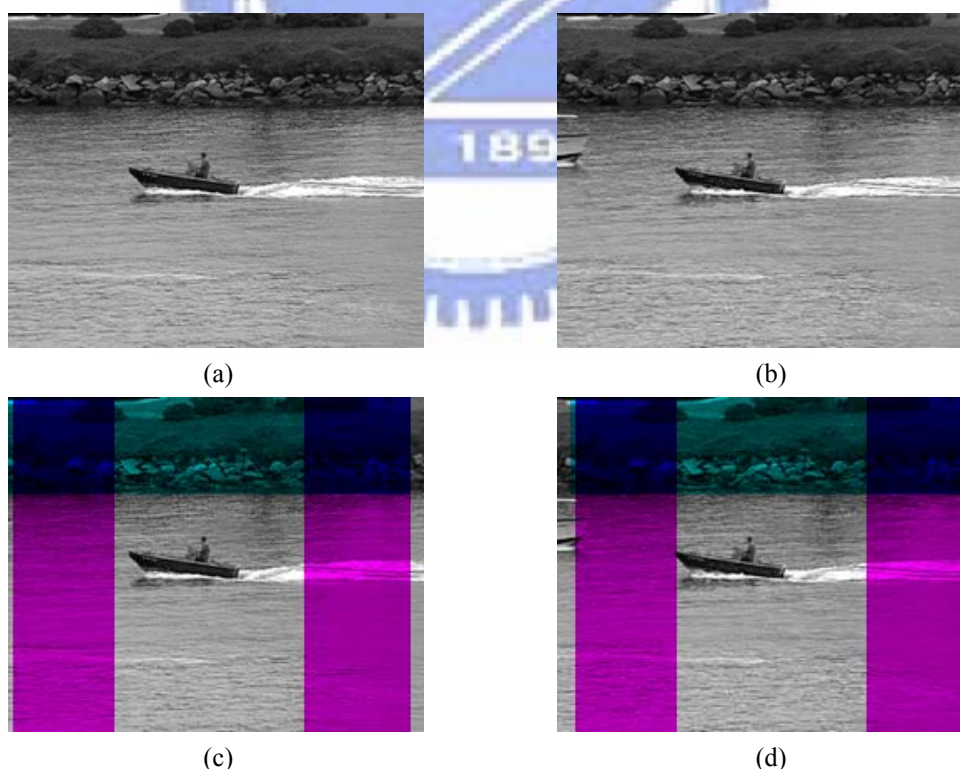


圖 5-6 (a)、(b)為 Coastguard 影片的兩張影格。(c)、(d)分別為兩張影像的背景區域估計結果。估計的背景平移量為(7.5186, 0.0460)

圖 5-7 是兩張連續的影像以及背景區域分割的結果。我們可以發現玩偶和背景區域有被區別出來。在水平方向的背景區域分割有不錯的結果，青色區域標示了大部分的水平方向背景；在垂直方向的背景區域顯得較為破碎，紫紅色區域標示了玩偶的左側兩個地方。我們可以從圖 5-8 中觀察圖 5-7(a) 的變異量，在兩個紫紅色區域中的變異量較小，在玩偶右側的變異量也是呈現較小的情形，使這兩個區域被列入不考量的範圍。同理，在影像的左右邊界，也都有少部分背景區域沒有被分割出來。

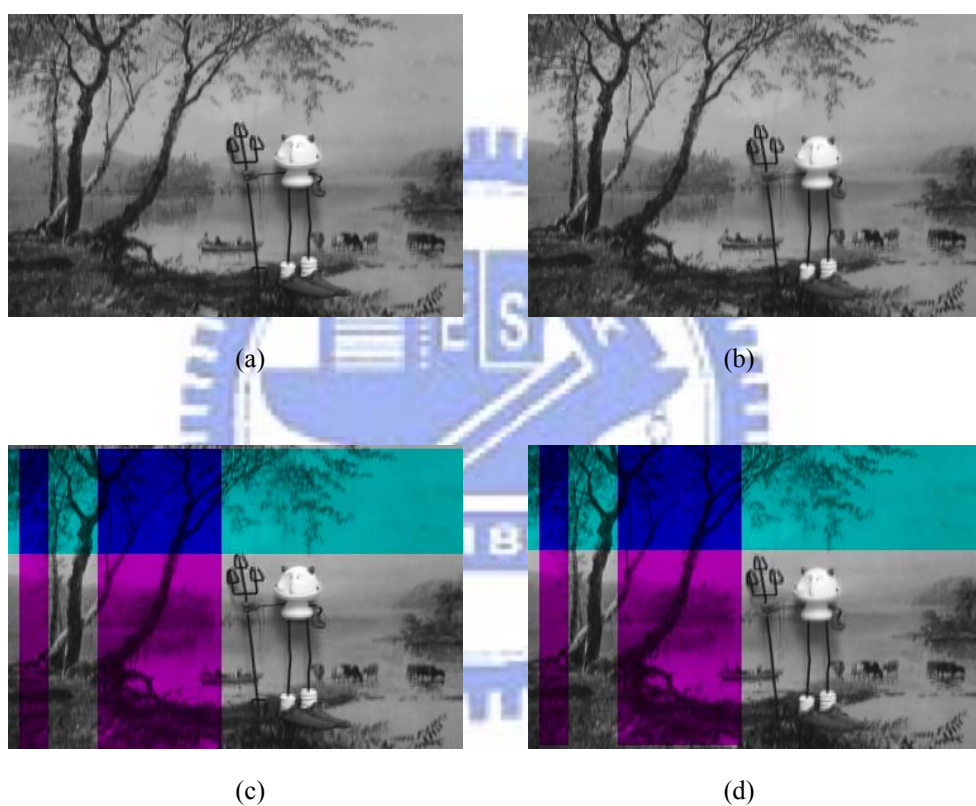


圖 5-7 (a)、(b)為連續的兩張真實影像。(c)、(d)分別為兩張影像的背景區域估計結果。估計的背景平移量為 $(-0.0528, -2.9177)$

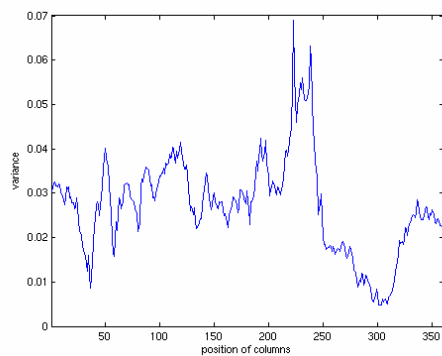


圖 5-8 圖 5-7 (a) 影像每行的變異量。

圖 5-9(a)、(b)是兩張真實影像，影像中有兩個行人，左邊的男性有明顯的移動，右邊的女性雖然沒有明顯的移動，但是動作上有改變。圖 5-9 (c)、(d)則是背景區域分割的結果。右邊男性的部分有不錯的結果，左邊女性則頭部部分因為變動不明顯，加上沒有移動，被列入了背景區域。我們可以發現前景物體除了平移外，外型的改變也可以被判斷出來。但是因為此處並沒有考量影像的內容，所以同一個物體有可能因為小部分沒有平移量，而被切開。另外，上方邊界的部分在水平背景區域分割時沒有被標示出。我們觀察這個區域的平均值與變異量，如圖 5-10 和圖 5-11 所示，扣除邊界效應的部分，在紅色線段表示的圖 5-9 (a)第 10 列的附近以及藍色線段表示的圖 5-9 (b)第 15 列的附近，平均值並不是對應的，而是前者高於後者；而我們可以發現在圖 5-11 第 10 列附近區域的變異量比較小，雖然不足以小到讓我們忽略這個區域，但在平均值的差異和變異量的同時作用下（見 3-1 式），使得我們不將此區域列入背景。

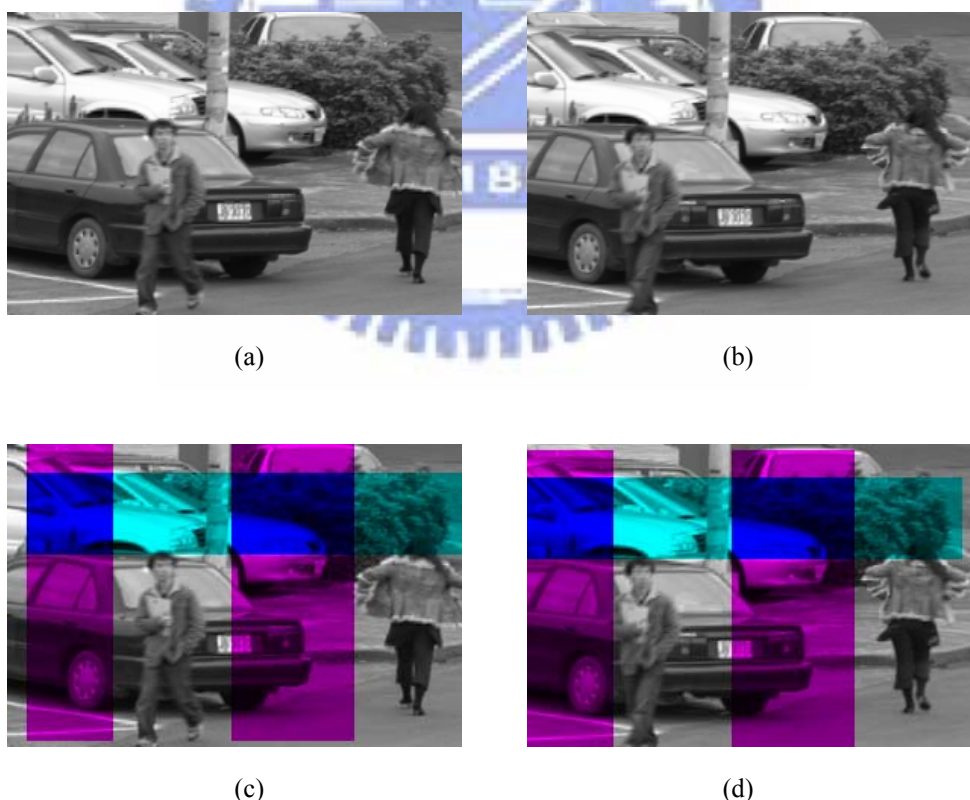


圖 5-9 (a)、(b)為連續的兩張停車場上行人走過的真實影像。(c)、(d)分別為兩張影像的背景區域估計結果。估計的背景平移量為(-12.9296, 3.6552)

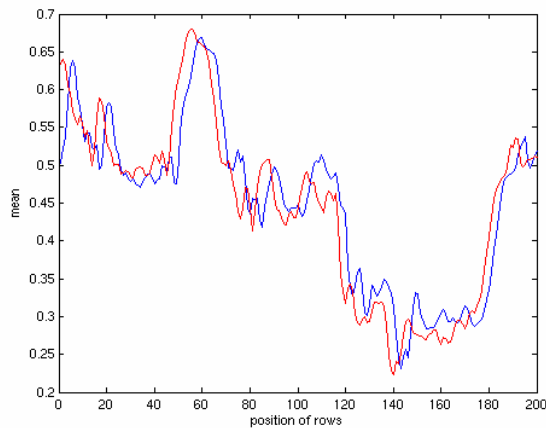


圖 5-10 紅色線段為圖 5-9(a)各列的灰階平均值，藍色線段為圖 5-9(b)各列的灰階平均值。

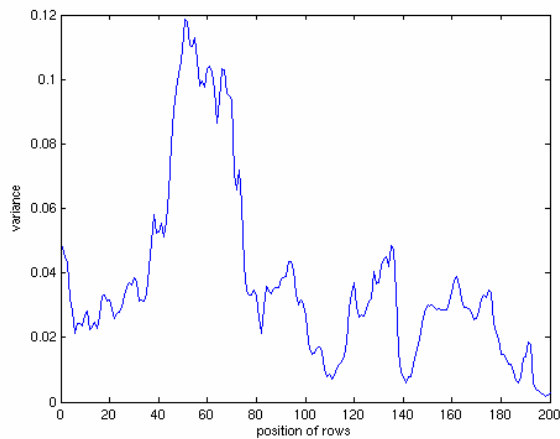


圖 5-11 圖 5-9(a)各列的灰階變異量。

圖 5-12 顯示的是停車場的影像序列，圖 5-13 則是其背景區域分割的結果。圖 5-13(a)、(b)是使用圖 5-12(a)、(b)兩張影像進行分割的結果，圖 5-13(c)、(d)及(e)、(f)則分別是使用圖 5-12(c)、(d)及(e)、(f)等影像進行分割的結果。我們可以發現，(a)、(b)中兩個行人分開較遠的情形下，背景區域分割的結果可以找出兩個行人之間背景區域。之後(c)、(d)中兩個人逐漸接近以至於重疊，兩個行人的前景區域就合併了。最後(e)、(f)中，因為兩個人的距離，相對於背景估計的帶狀區域寬度，仍相隔很近，尤其是(e)中兩個行人的手提袋非常靠近，約為 9 至 10 個像素，所以仍舊沒有被分成兩個前景區域。在這邊我們使用的帶狀區域

的最小寬度是 10 個像素，也就是背景區域必須超過 10 行或是 10 列，才會被分割出來。在圖 5-13(a)、(b)及(c)、(d)中，右邊的道路部份，則是因為變異量太小而不列入背景區域的考量。

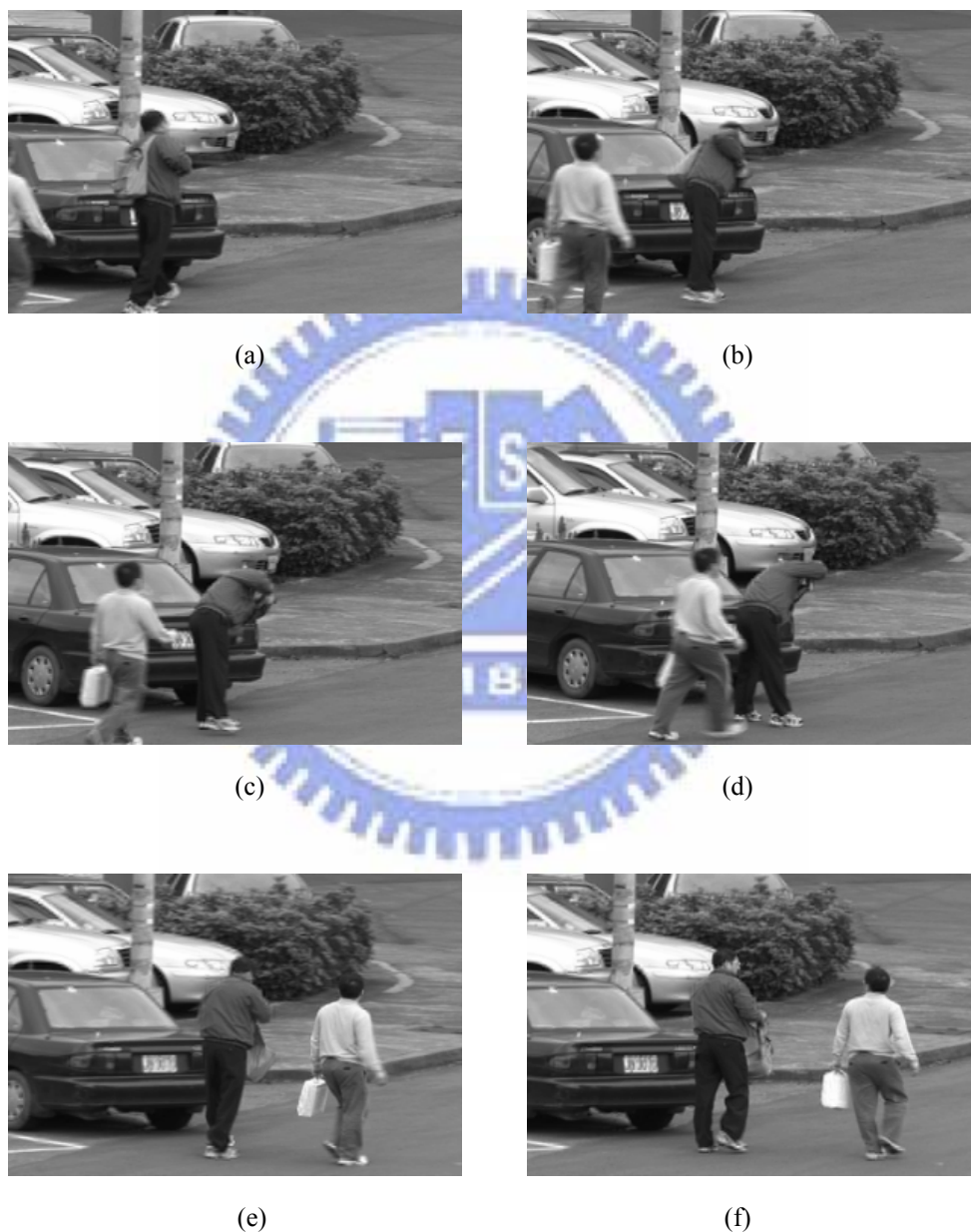


圖 5-12 停車場中兩行人併行的影片中的六張影像。

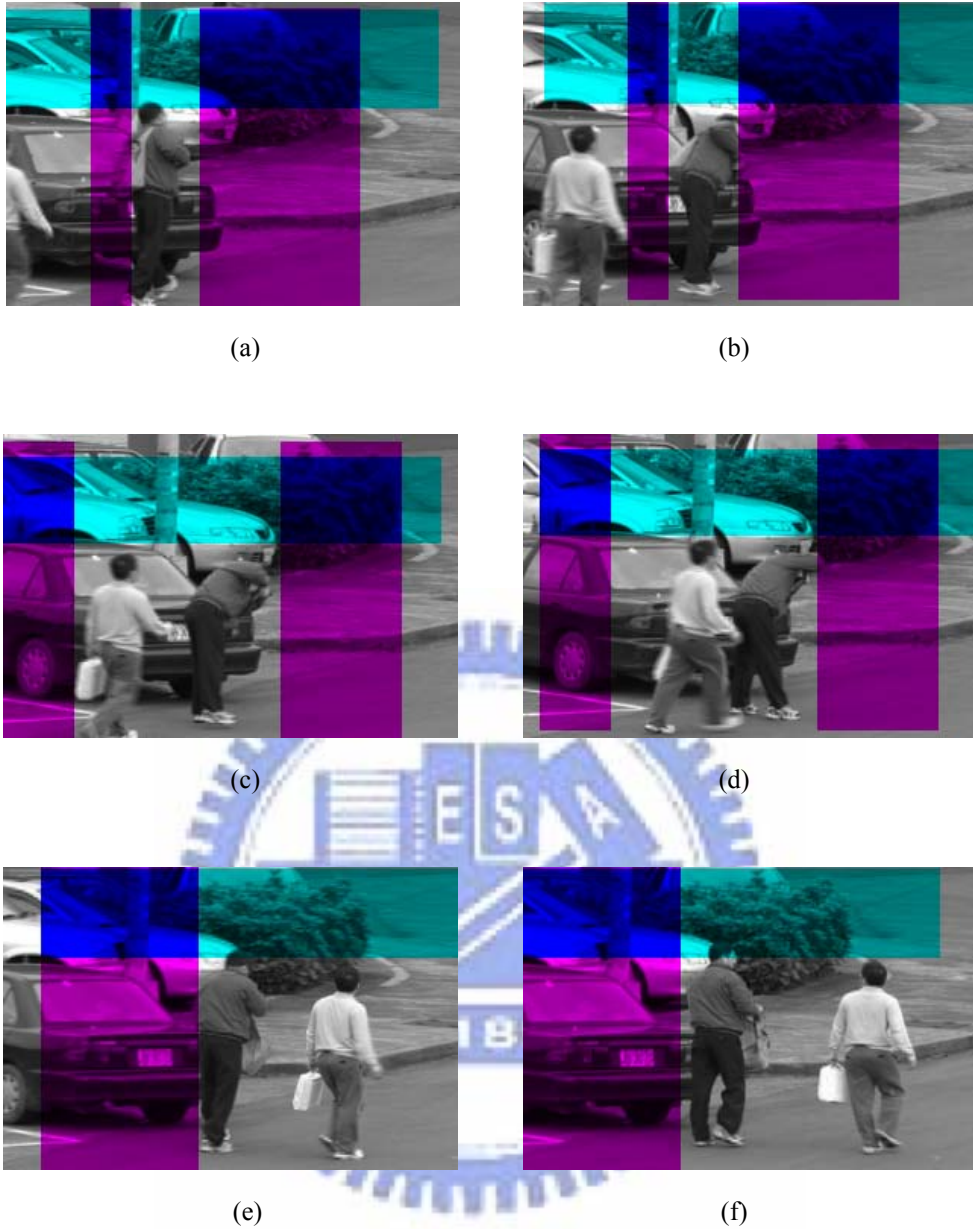


圖 5-13 (a)、(b)是圖 5-12 (a)、(b)影像背景區域分割的結果。(c)、(d)是圖 5-12 (c)、(d)影像背景區域分割的結果。(e)、(f)是圖 5-12 (e)、(f)影像背景區域分割的結果。

5.2.2 自然影像的次像素平移量討論與應用

為了比較次像素的平移量以及像素為單位的平移量，我們利用平移量來還原影像的移動，比較與第一張影像的差異。圖 5-14 是利用圖 5-12 (e)、(f)的影像以及背景區域分割和還原平移量的比較。圖 5-12 的次像素平移量分析結果在水平方向為-25.1245 像素，垂直方向為 0.2651 像素。我們分別取四種以像素為單位的

平移量來比較，分別為(-26,0)、(-25,0)、(-26,1)、(-25,1)像素，也就是最接近次像素平移量的整數像素平移量。將圖 5-12 (f)還原平移量後與圖 5-12 (e)相減取差異的絕對值後所得的結果，分別如圖 5-14 (a)、(b)、(c)、(d)所示。我們可以發現，圖 5-14(b)整體而言，在背景區域顯得較為乾淨。

接著我們將圖 5-14 四張影像分別畫出如圖 5-13(e)、(f)所標示估計背景區域，以及非背景區域的分佈圖，如圖 5-15 所示。圖 5-15(b)圖的藍色線段所表示的估計背景區域，呈現較為接近於零的結果。且和非背景區域的紅色線段，也有比較明顯的區隔。這表示原圖的平移量比較接近於(-25,0)，而其他三種整數平移量。我們分析的結果在水平方向為-25.1245 像素，垂直方向為 0.2651 像素，兩者相符。

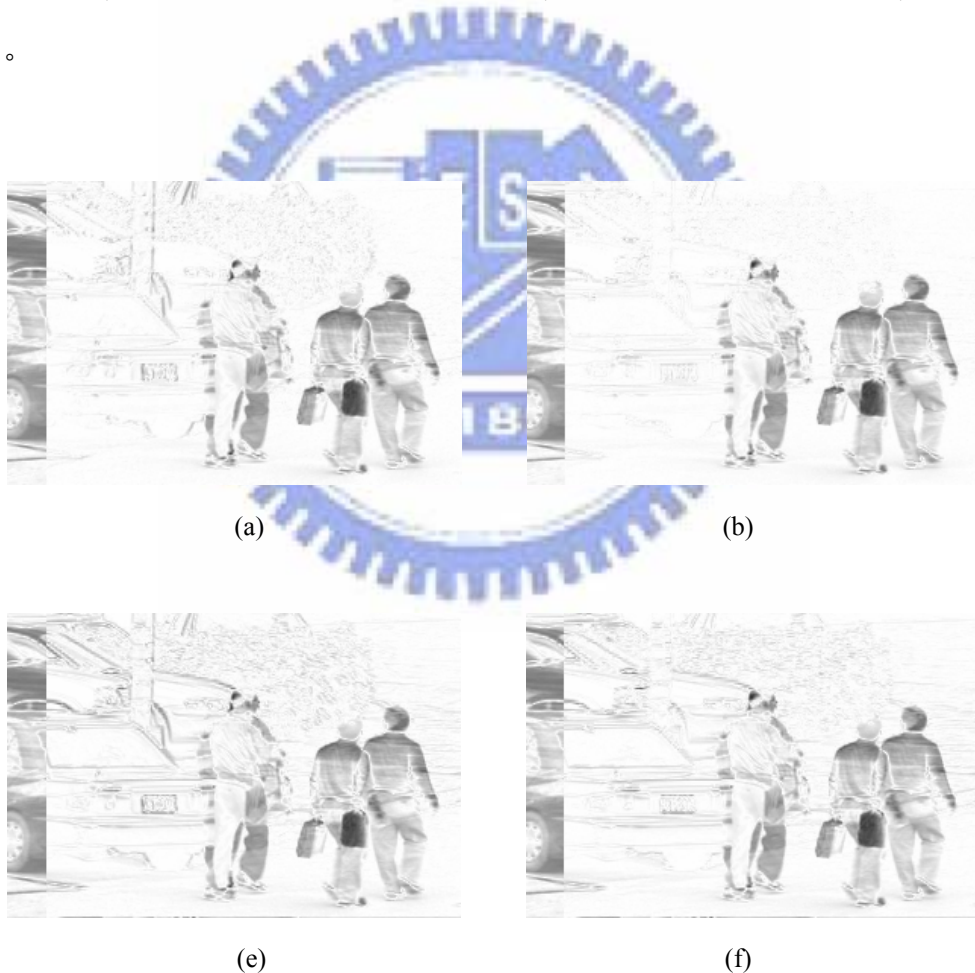


圖 5-14 將圖 5-9 (b)影像以不同的平移量還原移動後與圖 5-9 (a)比較差異的結果。(a)平移量為(-13,3)。(b)平移量為(-12,3)。(c)平移量為(-13,4)。(d)平移量為(-12,4)。

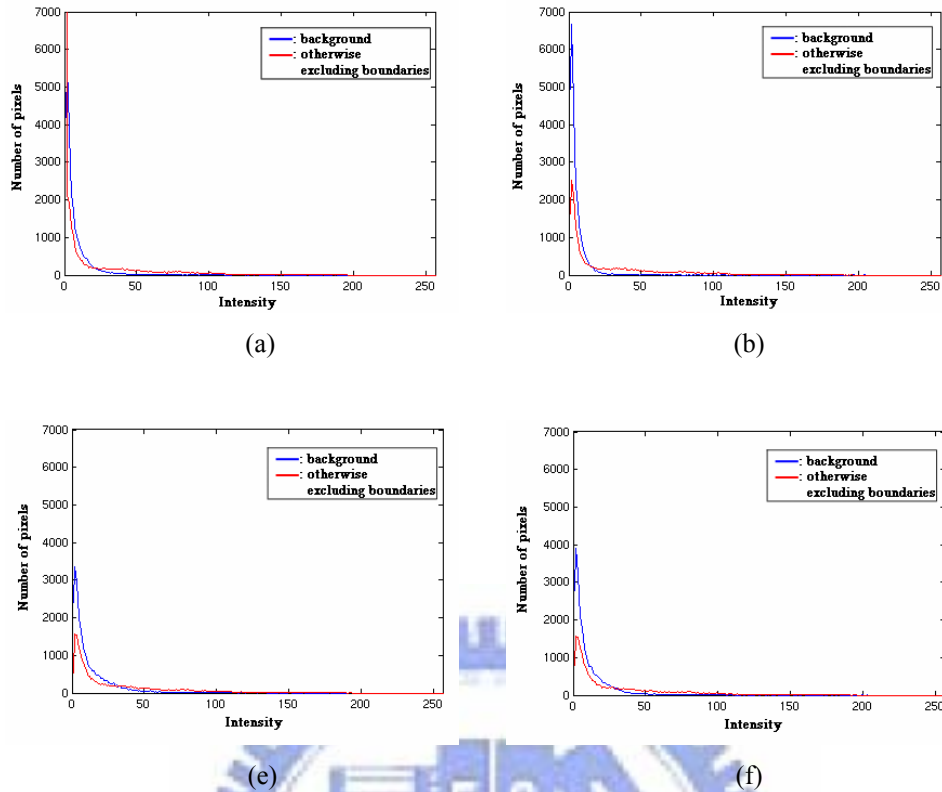


圖 5-15 分別對圖 5-14 四張影像的背景區域以及非背景區域所作的分佈圖。藍色線段估計背景區域，紅色線段為扣除邊界及背景區域的其他部分

利用估計的背景區域資訊，我們也可以在視訊中選擇適當的影格，移除背景區域，以進行前景物體的分割。首先我們使用視訊中的連續四張影像，分別利用第一張影像與第二、三、四張影像進行次像素平移量估計以及背景區域估計。之後分別還原最靠近次像素平移量的整數平移量來比較影像平移前後的差異，並利用估計的背景區域資訊，找出最適合模擬攝影機移動的影格，則可以還原攝影機의 平移，得到去除背景區域的結果。

圖 5-16 是四張連續的影像，分別利用第一張影像與第二、三、四張影像進行次像素平移量估計以及背景區域估計的結果。(a)與(b)，(c)與(d)。(e)與(f)的估計平移量分別為(-5.3604, 0.9742)、(-8.1320, 2.1695)、(-12.9269, 3.6552)。接著我們對三個次像素平移量取最靠近的整數平移量，分別為(-6, 0)、(-5, 0)、(-6, 1)、(-5, 1)，和(-9, 2)、(-8, 2)、(-9, 3)、(-8, 3)，以及(-13, 3)、(-12, 3)、(-13, 4)、(-12, 4)，平移影像後與第一張影像比較差異，結果如圖 5-17 所示。並畫出估計背景區域

及非背景區域的差異分佈圖，如圖 5-18 所示。從分佈圖中可以發現，圖 5-18 (c)、(d)、(f)藍色線段表示的背景區域差異大多分佈在接近零的地方，而三者中，(f)中紅色線段表示的非背景區域略高於(c)、(d)的紅色線段。從表 5-1 中觀察也可發現，圖 5-17(f)在估計的背景區域之中，平均值僅 0.0193，是圖 5-17 中最低的。這樣的結果可以幫助我們判斷圖 5-17(f)是背景區域移除最完整的結果。

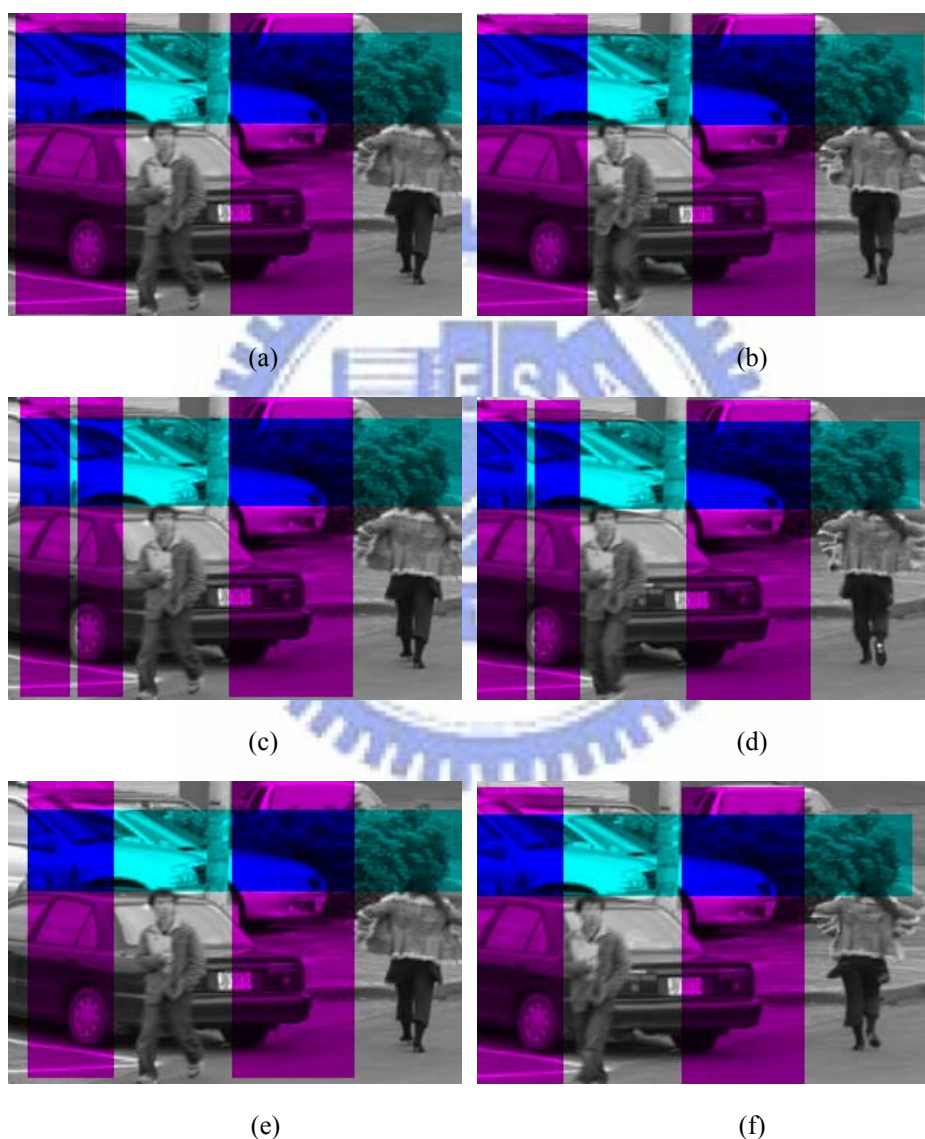


圖 5-16 連續的四張影像，第二、三、四張影像分別跟第一張影像的次像素平移量估計，及背景區域估計的擷果。(a)、(c)、(e)皆為第一張影像，與第二、三、四張影像的進行背景區域估計。(b)為第二張影像。(c)為第三張影像。(d)為第四張影像。(a)與(b)，(c)與(d)。(e)與(f)的估計平移量分別為(-5.3604, 0.9742)、(-8.1320, 2.1695)、(-12.9269, 3.6552)。

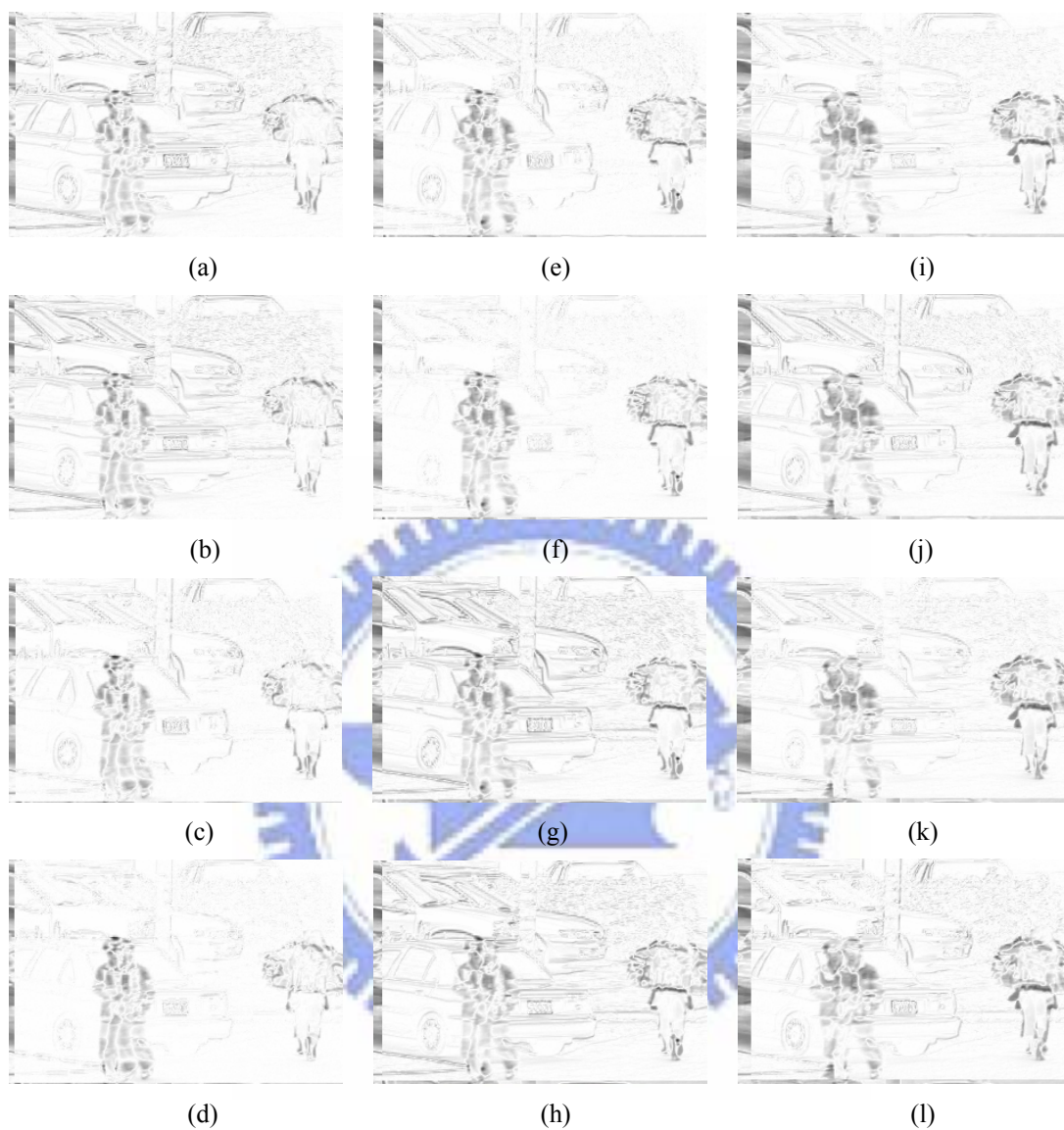


圖 5-17 (a)-(d)為圖 5-16(b)還原整數平移量與圖 5-16(a)比較的差異。(e)-(h)為圖 5-16(d)還原整數平移量與圖 5-16(c)比較的差異。(i)-(l)為圖 5-16(f)還原整數平移量與圖 5-16(e)比較的差異。

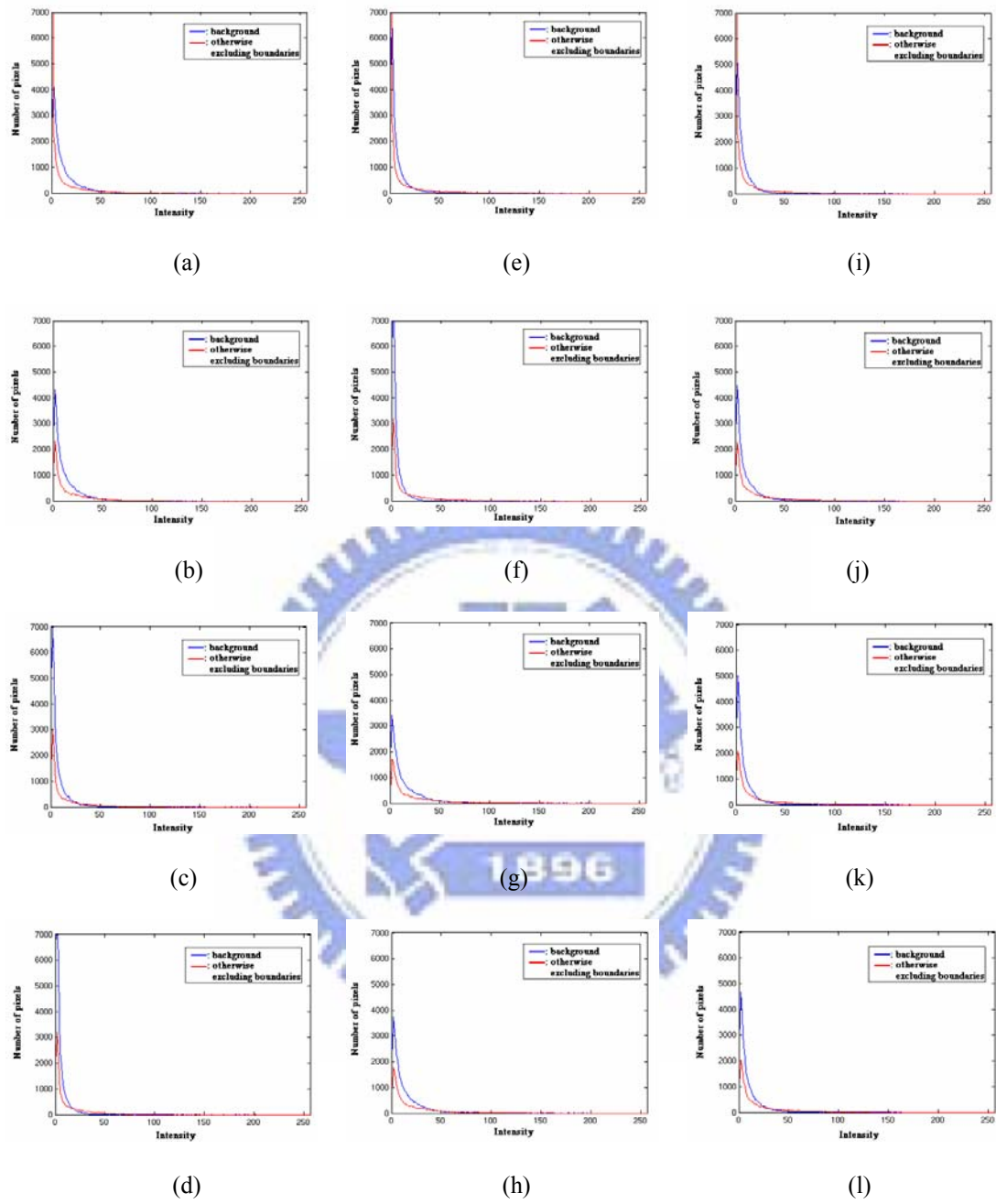
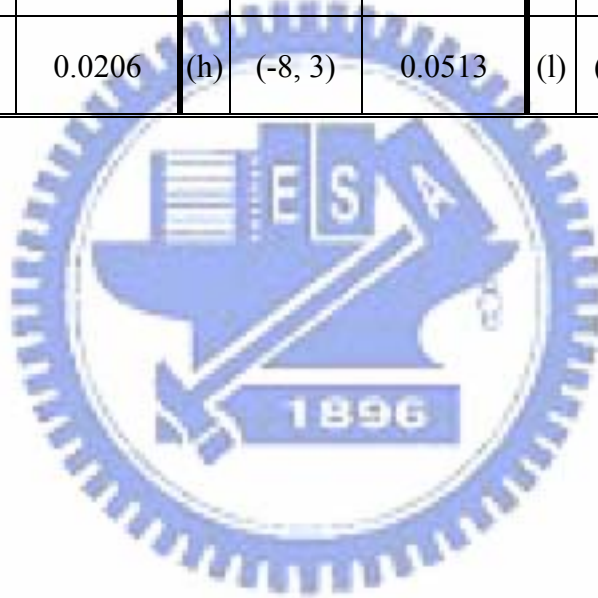


圖 5-18 圖 5-17 各張影像差異的分佈圖。

表 5-1 圖 5-17 各張影像估計背景區域的平均差值。

(-5.3604, 0.9742)			(-8.1320, 2.1695)			(-12.9269, 3.6552)		
Fig. No.	Rounded motion	Mean error	Fig. No.	Rounded motion	Mean error	Fig. No.	Rounded motion	Mean error
(a)	(-6, 0)	0.0491	(e)	(-9, 2)	0.0300	(i)	(-13, 3)	0.0297
(b)	(-5, 0)	0.0478	(f)	(-8, 2)	0.0193	(j)	(-12, 3)	0.0402
(c)	(-6, 1)	0.0297	(g)	(-9, 3)	0.0586	(k)	(-13, 4)	0.0322
(d)	(-5, 1)	0.0206	(h)	(-8, 3)	0.0513	(l)	(-12, 4)	0.0375



第六章 結論

我們在此篇論文提出了一個結合次像素背景平移量分析及背景區域分割的方法。此方法利用頻率空間的相位特性，對於連續的兩張未經過任何前置處理的影像計算出次像素背景平移量，再利用平移量找出背景區域，並與前景區域做分割。

在背景平移量的分析上，此方法結合了影像空間以及頻率空間的做法。先利用影像空間找出少部分的背景區域，並且僅以各行（列）之灰階平均值和變異量來計算，而不是直接使用二維的影像。之後再於頻率空間對少部分的背景區域進行次像素的平移量分析，得到非整數像素的平移量資訊。由於此種做法利用先前使用平均值與變異量作背景區域比對的資訊，再繼續利用平移量進行背景部分的分割，因為並沒有考慮影像內容以及前景物體的形狀，只要物體在影像上的成像未受到干擾，對於有發生形變的物體也是可以進行背景區域分割。

對於未來的研究，加強分割的準確性是一個重要的目標，在分割之後，進一步的對物體的形狀作修正，而不只單使用矩形方格的表示。在影像中對應的各行（列）在頻率空間振幅相似度的比較也是一個值得探討的問題。另外，雖然頻率空間的相位分析所使用的計算量很低，但在結合影像空間的以灰階平均值和變異量作背景區域比對時花費不少計算時間，如果僅使用變異量最大之帶狀區域，或是其他可以減少搜尋範圍，是否有助於背景區域比對速度的提升，也是值得進一步探討的部分。

參考文獻

- [1] L. G Brown, “A survey of image registration techniques,” *ACM Computing Survey*, vol. 24, no. 4, Dec. 1992.
- [2] P. R. Hill, T. K. Chiew, D. R. Bull, and C. N. Canagarajah, “Interpolation free subpixel accuracy motion estimation,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 12, pages 1519-1526, Dec. 2006.
- [3] J. W. Suh, and J. Jeong, “Fast sub-pixel motion estimation techniques having lower computation complexity,” *IEEE Trans. on Consumer Electronics*, vol. 50, no. 3, pages 968-993, Aug. 2004.
- [4] D. G. Kuglin, and D. C. Hines, “The phase correlation image alignment method,” *IEEE International Conf. on Cybernetics and Society*, pages 163-165, Sep. 1975.
- [5] H. Stone, M. Orchard, and E. C. Chang, “Subpixel registration of images,” *Asilomar Conf.e on Signal, Systems, and Computers*, vol. 2, pages 1446-1452, Oct. 1999.
- [6] H. Foroosh, J. B. Zerubia, and M. Berthod, “Extension of phase correlation to subpixel registration,” *IEEE Trans .on Image Processing*, vol. 11, no. 3, pages 188-200, Mar. 2002.
- [7] M. Balci and H. Foroosh, “Subpixel estimation of shifts directly in the Fourier domain,” *IEEE Trans. on Image Processing*, vol. 15, no. 7, pages 1965-1972, Jul. 2006.
- [8] Y. Wang and Q. Ji, “A dynamic conditional random field model for object segmentation in image sequences,” *IEEE International Conf. on Computer Vision and Pattern Recognition*, vol. 1, pages 264-270, Jun. 2005.

- [9] Y. Wang, K. F. Loe, T. Tan, and J. K. Wu, "Spatiotemporal video segmentation based on graphical models," *IEEE Trans. on Image Processing*, vol. 14, no. 7, pages 937-947, Jul. 2005.
- [10] A. Briassouli and N. Ahuja. "Fusion of frequency and spatial domain information for motion analysis," *IEEE International Conf. on Pattern Recognition*, vol. 2, pages 175-178, Aug. 2004.
- [11] A. Briassouli and N. Ahuja. "Integrated spatial and frequency domain 2D motion segmentation and estimation," *IEEE International Conf. on Computer Vision*, vol. 1, pages 244-250, Oct. 2005.
- [12] A. Briassouli and N. Ahuja, "Spatial and Fourier error minimization for motion estimation and segmentation," *IEEE International Conf. on Pattern Recognition*, vol. 1, pages 94-97, 2006.
- [13] J. Y. A. Wang and E. H. Adelson, "Layered representation for motion analysis," *IEEE International Conf. on Computer Vision and Pattern Recognition*, pages 261-266, New York, Jun. 1993.
- [14] P. H. S. Torr, R. Szeliski, and P. Anandan, "An integrated Bayesian approach to layer extraction from image sequences," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pages 297-303, Mar. 2001.
- [15] J. Xiao and M. Shah, "Motion layer extraction in the presence of occlusion using graph cut," *Proc. IEEE International Conf. on Computer Vision and Pattern Recognition*, pages 972-979, 2004.