

# 國立交通大學

電機資訊學院 資訊學程

## 碩士論文

MPEG-2 影片上的自動化鏡頭轉換偵測法

Automatic Shot Change Detection in MPEG-2 Video



研究生：馮貞慈

指導教授：陳玲慧 教授

中華民國九十四年六月

# MPEG-2 影片上的自動化鏡頭轉換偵測法

## Automatic Shot Change Detection in MPEG-2 Video

研究生：馮貞慈

Student : Chen-Tzu Feng

指導教授：陳玲慧

Advisor : Dr. Ling-Hwei Chen

國立交通大學

電機資訊學院 資訊學程

碩士論文



A Thesis

Submitted to Degree Program of Electrical Engineering and Computer Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Computer Science

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

# MPEG-2 影片上的自動化鏡頭轉換偵測法

學生：馮貞慈

指導教授：陳玲慧 博士

國立交通大學電機資訊學院 資訊學程（研究所）碩士班

## 摘 要

近年來數位影片的應用及重要性與日俱增，在多媒體系統應用中，快速搜尋影片或是存取使用者所須的資訊，第一步就是需要偵測鏡頭的變換來減低處理的資料量，本論文提出一個可依照不同影片型態能自動化設定偵測換鏡的門檻值，直接讀取MPEG-2 視訊影片中的兩兩相鄰的I (P) DC畫面的DC差異值，P 畫面裡編碼方式為intra-coding 的MB個數及P畫面裡具有主要位移向量的MB個數，這三個壓縮資訊的特徵來偵測發生鏡頭變換的位置，得到相當不錯的偵測結果，針對偵測瞬間式鏡頭變換的正確性約 98.7%，針對偵測漸進式鏡頭變換的正確性約 92.3%，可以將此方法應用在各種類型的影片上。同時為了容易研究及增加應用性，我們也開發了一個方便使用的應用系統。

# Automatic Shot Change Detection in MPEG-2 Video


Student : Chen-Tzu Feng

Advisors : Dr. Ling-Hwei Chen

Degree Program of Electrical Engineering Computer Science

National Chiao Tung University

## ABSTRACT

The logo of National Chiao Tung University is a circular emblem with a gear-like border. Inside the circle, there is a stylized representation of a building or a ship, and the year '1896' is inscribed at the bottom.

Recently, importance and application of the digital video is to grow with each passing day. And digital video has also been used widely in multimedia systems. In order to search and access user-desired information quickly, the first step is to partition a digital video into groups of meaningful segments by detecting shot change frame. In this thesis, we develop an algorithm for efficiently detecting shot changes in MPEG-2 compressed video. Since the method is based on the statistics of video difference and the statistics of information in MPEG-2 compressed video, it could be applied to various kinds of digital videos. Furthermore, we provide a useful application program of shot change detection to make use conveniently.

---

## 誌 謝

首先對於我的指導教授陳玲慧教授獻上最真誠的感謝，在她細心的指導下，讓我能夠體驗到學習的樂趣與研究的精神。更感謝平時因工作，無法即時達到老師的要求，老師對我的包容與耐心。

此外還要感謝自動化資訊處理實驗室的學長姊與學弟妹們，在這三年來對我的指導與協助，讓我能夠順利完成研究所學業。以及謝謝公司上司在我求學的這段期間，對我的包容。

最後，由衷地感謝我的父母與家人多年來給我的關懷與栽培，使我得以專心於學業上，僅以我最誠摯的心意將此篇論文獻給我的父母與家人。



# 目 錄

MPEG-2 影片上的自動化鏡頭轉換偵測法 .....	III
<b>AUTOMATIC SHOT CHANGE DETECTION IN MPEG-2 VIDEO</b> .....	IV
誌 謝 .....	V

目 錄.....	VI
----------	----

圖目錄 .....	VIII
-----------	------

表目錄 .....	IX
-----------	----

第一章 簡介.....	1
-------------	---

第二章 MPEG-2 VIDEO 壓縮標準.....	5
----------------------------	---

2.1 MPEG-2 視訊系統的架構 .....	5
--------------------------	---

2.2 MPEG-2 視訊在影像空間上的壓縮理論 .....	11
--------------------------------	----

2.3 MPEG-2 視訊在時間軸上的壓縮理論 .....	13
-------------------------------	----

第三章 相關研究及文獻探討.....	15
--------------------	----

3.1 利用 DC 值偵測換鏡的方法 .....	15
--------------------------	----

3.2 利用大區塊資訊的換鏡偵測法 .....	17
-------------------------	----

3.3 自動設定門檻值的方法 .....	19
<b>第四章 本論文所提的鏡頭變換偵測法 .....</b>	<b>25</b>
4.1 換鏡偵測所需要的三種參考特徵 .....	26
4.1.1 DC 畫面差異值及門檻值設定 .....	28
4.1.2 在 P 畫面中的 intra-coding 的 MB 總數的分布統計 .....	31
4.1.3 在 P 畫面中的具主要位移向量的 MB 數量的分布統計 .....	35
4.2 換鏡偵測的流程 .....	39
4.2.1 偵測瞬間式換鏡之方法 .....	39
4.2.2 偵測漸進式換鏡之方法 .....	42
<b>第五章 實驗結果及分析 .....</b>	<b>45</b>
<b>第六章 結論與未來展望 .....</b>	<b>51</b>
<b>參 考 文 獻 .....</b>	<b>52</b>



# 圖目錄

圖 1.1 瞬間式換鏡示意圖 .....	2
圖 1.2 漸進式換鏡示意圖 .....	3
圖 2.1 Video sequence layer 及 GOP layer 示意圖 .....	6
圖 2.2 畫面參考方式 .....	8
圖 2.3 Macro block 的架構圖 .....	9
圖 2.4 Macro block 的往前，往後參考示意圖 .....	11
圖 2.5 位移估測及位移補償示意圖 .....	14
圖 3.1 DC 值畫面示意圖 .....	16
圖 3.2 換鏡時的前後參考關係 .....	18
圖 3.3 換鏡時的 Changing MB 的關係 .....	19
圖 3.4 一個視訊影片中連續兩畫面的差異值 .....	21
圖 3.5 圖 3.4 的直方圖 .....	22
圖 3.6 林文隆方法的示意圖 .....	23
圖 3.7 一個無法偵測到瞬間式換鏡的例子 .....	24
圖 4.1 我們提出之換鏡偵測法的流程圖 .....	26
圖 4.2 相鄰兩 I 或 P 畫面的 DC 畫面差異值 .....	29
圖 4.3 圖 4.2 的直方圖 .....	29
圖 4.4 DC 畫面差異值分布及相對應門檻值示意圖 .....	30
圖 4.5 圖 4.3 中之三個門檻值 .....	31
圖 4.6 P Frame 裡 intra-coding 的 MB 數的分布圖 .....	32
圖 4.7 圖 4.6 的直方圖 .....	33
圖 4.8 圖 4.7 的示意圖 .....	34
圖 4.9 一個 MPEG-2 Video 中具有主要位移向量的 MB 個數的直方圖 .....	36
圖 4.10 MPEG-2 Video 具主要位移向量的 MB 個數的直方圖及門檻值設定示意圖 .....	37
圖 4.11 是一個 MPEG-2 Video 漸進式換鏡的 Motion Vector number 變化示意圖 .....	38
圖 4.12 偵測瞬間式換鏡的流程圖 .....	42
圖 4.12 偵測漸進式換鏡的流程圖 .....	44
圖 5.1 所用的實驗資料庫影片的代表畫面 .....	46
圖 5.2 瞬間式鏡頭漏判的畫面 .....	48
圖 5.3 漸近式鏡頭漏判的畫面分析之一 .....	49
圖 5.3 由光線造成的瞬間式鏡頭誤判的畫面 .....	49
圖 5.4 由旋轉造成的瞬間式鏡頭誤判的畫面 .....	49
圖 5.5 影片中誤判的漸進式鏡頭 .....	50



# 表目錄

表 5.1 瞬間式鏡頭變換偵測及漸進式鏡頭變換偵測的實驗結果.....47



# 第一章 簡介

隨著科技一日千里的進步以及電腦數位化時代的來臨，數位化的多媒體資料格式及應用也愈來愈廣泛，多媒體的格式也越來越多，如 MPEG-1，MPEG-2 及 MPEG-4 等視訊格式，而這些影片都是由大量的資料組成，而要處理這些大量的視訊資料的第一步就是要將影片依鏡頭(Shot)做視訊資料分割，將連續的鏡頭與有鏡頭轉換(Shot change)做出區隔，而以鏡頭做為視訊資料索引查詢的單位，這樣可以減少在處理視訊影片的分析與索引時的資料量及所需要的時間，而這個將影片做視訊資料分割的動作就需要先將影片做鏡頭變換的偵測(Shot change detection)。



在一段影片中，若拍攝的鏡頭及內容具有連續性及相關性，我們認為這一段影片是屬於同一個鏡頭。同一鏡頭影片的資料具有高度的相關性，因此，要對視訊資料進行分析及檢索時，只要針對每一個鏡頭中具代表性的畫面，進行資料的儲存，搜尋，瀏覽和擷取的處理。屬於同一個鏡頭的一段影片，可以代表記錄同一個事件或同一個主題，而鏡頭轉換，換鏡前的影片內容和換鏡後的影片內容沒有相關性。

鏡頭轉換主要可分為瞬間式轉換(Abrupt Transition)和漸進式轉換(Gradual Transition)二種，瞬間式轉換是指在影片中的鏡頭轉換是立即性的轉換，兩個鏡頭的分界很清楚，如圖 1.1 所示，圖中每一格代表一個畫面(frame)，其中在 frame

i 與 frame i+1 之間存在一個瞬間式換鏡。而漸進式轉換則是在二個鏡頭的轉換間加入了一些特效功能，將換鏡的過程延長並使得換鏡的過程變得更加平滑，如圖 1.2 所示，在鏡頭轉換過程中，是第一個鏡頭逐漸消失，而第二個鏡頭逐漸現並取代原先的鏡頭，漸進式換鏡的型態可依換鏡轉換過程的特效來分類，常見的漸進式換鏡方式有溶解(Dissolve)，淡入(Fade-in)，淡出(Fade-out)和抹去(Wipe)這四種型態，如圖 1.2 所示的漸進式換鏡的型態就屬於溶解型的漸近式換鏡。

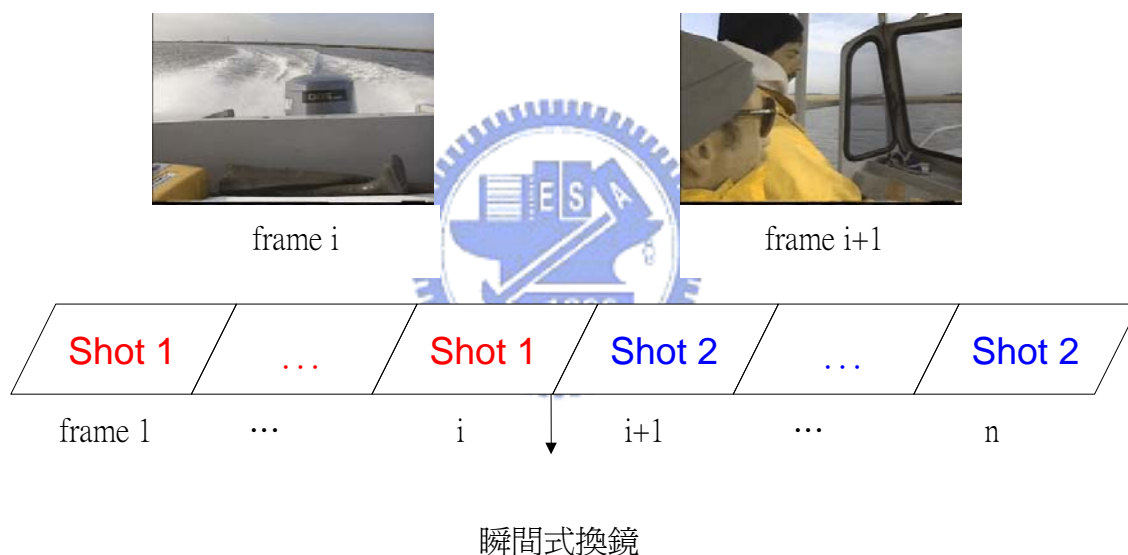


圖 1.1 瞬間式換鏡示意圖

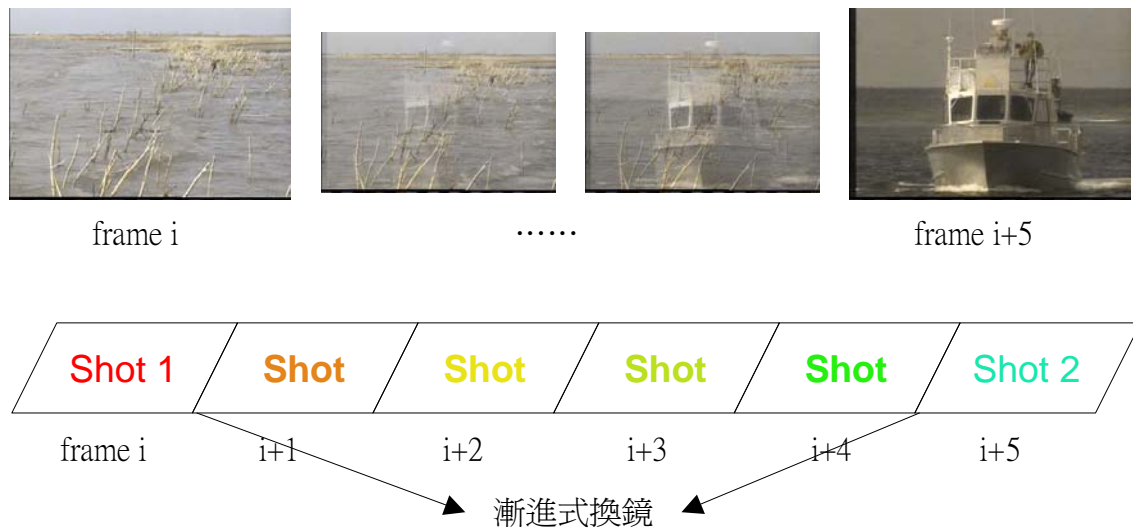


圖 1.2 漸進式換鏡示意圖

有許多學者致力於視訊影片鏡頭變換偵測的研究，也提供了很多的偵測特徵及應用方法，早期鏡頭變換的研究大多針對未壓縮的視訊影片(uncompressed video)來進行偵測[1-2]。使用未壓縮的視訊資料進行換鏡偵測，在處理的資料量十分龐大，且將視訊影片完全解碼再來偵測換鏡所需的計算時間也比較長，且目前普遍使用的 MPEG-1，MPEG-2 視訊影片皆屬於壓縮格式，故近年來研究的重點逐漸專注於如何直接使用已壓縮的視訊影片(compressed video)來進行偵測[3-4]。本篇論文提出的方法就是利用已壓縮的 MPEG-2 視訊影片來進行鏡頭變換的偵測。

不同類型的影片其畫面間變化的程度並不相同，如在運動型態的影片中，運動員的移動速度可能很快，故當相鄰兩畫面的差異很大時，可能只是連續運動的畫面，而非有鏡頭轉換的情況發生；而在偏靜態的影片如新聞播報，每一個畫面的差異都不大，就算是有鏡頭轉換的情況發生，畫面的差異也有可能遠比連續運動的畫面差異小。因此，使用同一個門檻值(Threshold)來偵測所有類型的視訊影

片，並不能得到最佳之結果。我們藉由統計影片的差異性，自動為各種不同型態的影片算出壓縮資訊的參考值並選擇適當之門檻值，以增加換鏡偵測的正確率。並將影片分為數個內容相似的片段，以方便後續的影片檢索，影片搜尋及關鍵畫面選取的功能，並且開發一個方便使用的應用系統。

本篇論文共分為六章，首先第一章說明本論文之研究動機及研究目的，第二章將介紹 MPEG-2 視訊壓縮的標準及壓縮原理，第三章則介紹前人所提出的有關影片鏡頭變換偵測的方法。我們提出之偵測方法及相關流程則於第四章介紹，第五章為我們的實驗結果與分析，最後則於第六章提出本篇論文的結論，未來的展望及努力的方向。



## 第二章 MPEG-2 Video 壓縮標準

MPEG 的全名是” Motion Picture Expert Group” ，是由 ISO/IEC 在 1988 年成立的組織，這個組織的工作是制定數位視訊的壓縮標準，在 1991 年提出 MPEG-1 視訊壓縮標準[5]，主要支援 Bit rate 為 1.5Mbps 的數位儲存媒體如影音 CD(VCD)。為了支援高畫質視訊壓縮的應用，MPEG 又在 1994 制定了 MPEG-2 視訊壓縮標準[6]，MPEG-2 是 MPEG-1 的延伸，基本架構及壓縮的演算法大致相同，但 MPEG-2 比 MPEG-1 多支援了交錯式掃瞄(interlaced scan) ，和更多的 YUV 色彩取樣格式(4:2:0,4:2:2) ，以及其它更進階的編碼技術，最常見的 MPEG-2 格式的視訊應用是衛星電視的傳輸及數位影音光碟(DVD) ，以下將在 2.1 節介紹 MPEG-2 視訊系統的架構，2.2 節介紹 MPEG-2 視訊在影像空間上的壓縮理論，2.3 節介紹 MPEG-2 視訊在時間軸上的壓縮理論。



### 2.1 MPEG-2 視訊系統的架構

MPEG-2 的視訊系統的架構可大致分為六個階層，由上到下依序是 Video Sequence layer，Group of pictures(GOP) layer，Picture layer，Slice layer，Macroblock(MB) layer 及 Block layer，以下將分別介紹之。

1. **Video Sequence layer:** 是指一個完整的視訊影片，在這一層中會記錄這一個影片的畫面大小，畫面播放速率，並包含許多連續的 GOP。
2. **Group of pictures(GOP) layer:** GOP 是由一張或多張的連續畫面所組成，每個畫面利用三種不同編碼方式之其中一種進行編碼。這三種不同編碼方式的畫面分別是 I-Frame(intra-coding frame)，P-Frame(Predictive-coding frame)和

B-Frame(Bidirectionally Predictive-coding frame)，一個 GOP 播放順序的組成包含一個 I 畫面，數個 P 畫面，數個 B 畫面。典型 GOP 組成是 I B B P B B P B B P B B P B B P，Video Sequence layer 及 Group of pictures(GOP) layer 的架構圖如圖 2.1 所示。

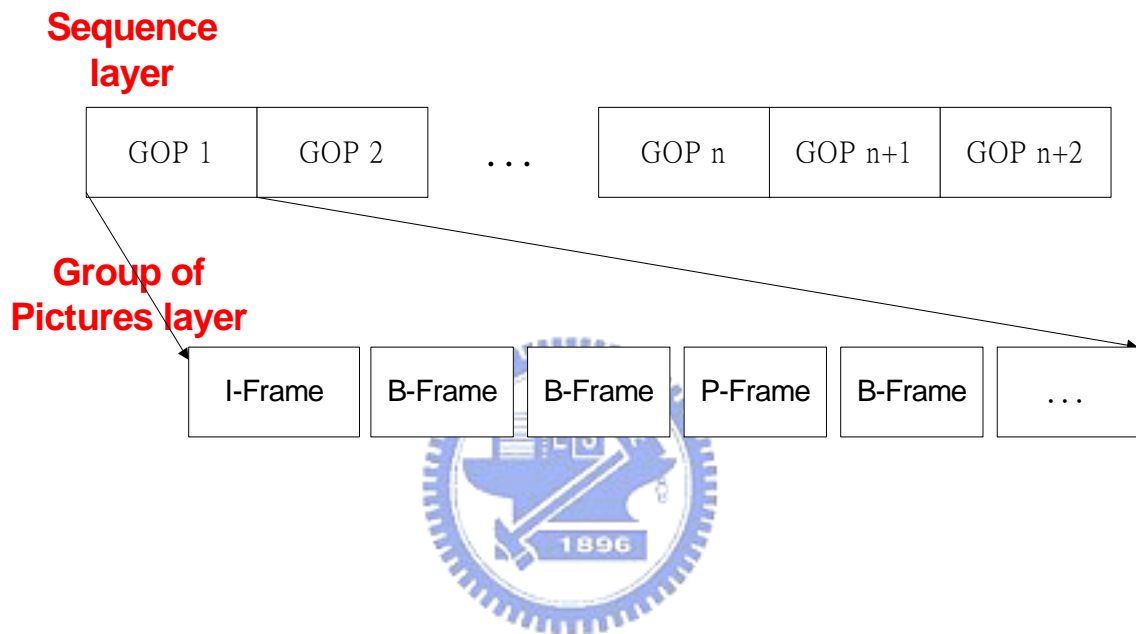


圖 2.1 Video sequence layer 及 GOP layer 示意圖

3. **Picture layer:** 在 MPEG-2 視訊影片中有三種不同編碼方式的畫面

(a) I-Frame (Intra-coding frame)

使用 JPEG Standard 的壓縮標準來編碼，不需要參考其他的畫面就可以進行編碼及解碼，I 畫面可做為視訊影片中的索引點，以提供隨機存取(seek)的功能。

(b) P-Frame (Predicted frame)

使用過去的畫面來當做參考畫面，進行位移估測(Motion Estimation)及位移補償(Motion Compensation)來減少時間軸上的重覆資料也可以減少編碼後的資料量。

### (c) B-Frame (Bi-directional frame)

使用過去和未來的畫面來當做參考畫面，進行雙向的(bi-directional)位移估測(Motion Estimation)及位移補償(Motion Compensation)的計算，所減少的時間軸上的重覆資料量比 P 畫面更多，所以更能有效的減少編碼後的資料量。

圖 2.2 中列出 I-Frame，P-Frame，B-Frame 的參考關係，I-Frame 不需要參考任何畫面就可獨立編解碼，P-Frame 需要參考前面已經解碼的 I 畫面或 P 畫面才能編解碼，B-Frame 需要參考前面已經解碼的 I 畫面或 P 畫面和後面尚未播放但已解碼的 I 畫面或 P 畫面的資料。

在 MPEG-2 影片中，資料存放的順序和影片播放的順序是不同的，因為 B 畫面的雙向參考會需要往前，往後參考的畫面完全解碼之後，這個 B 畫面才能被完全解碼，若是 MPEG-2 的播放順序是

I (1) B(2) B(3) P(4) B(5) B(6) P(7) B(8) B(9) P(10) B(11) B(12) P (13)B(14) B(15) P(16)

則存放在 MPEG-2 視訊資料中的順序是

I (1) P(4) B(2) B(3) P(7) B(5) B(6) P(10) B(8) B(9) P (13) B(11) B(12) P(16) B(14) B(15)



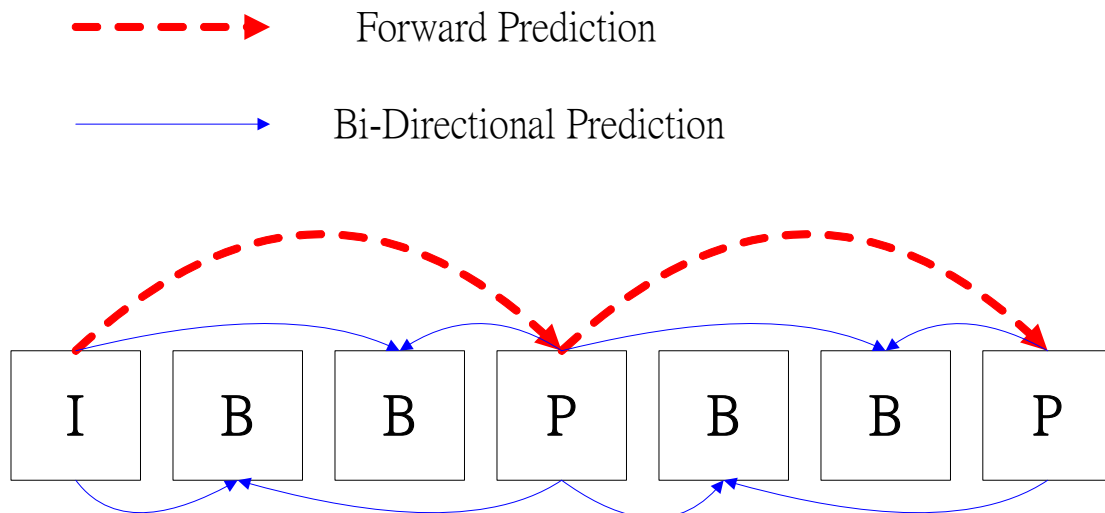
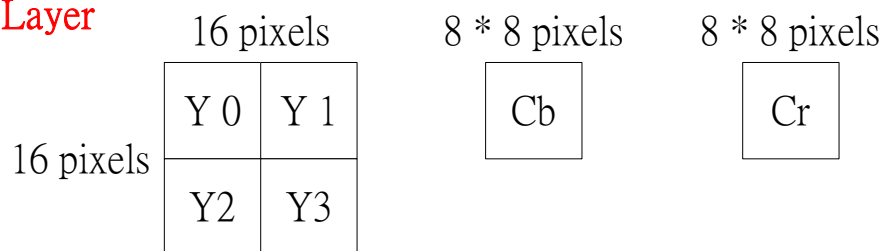


圖 2.2 畫面參考方式

4. **Slice layer:** 一個畫面是由若干個不可重疊的 Slice 組成，而一個 Slice 是由數個不可重疊的 MacroBlock 所組成，每一個 Slice 至少要有一個 MacroBlock 而且每一個 Slice 所包含的 MacroBlock 不同，在編碼時，一個畫面分割成數個 Slice 來編碼，在解碼時也依照 Slice 分開解碼，這樣可以避免解碼的錯誤影響到影片畫面。
5. **Macroblock(MB) layer:** Macroblock layer 是 MPEG-2 視訊資料結構中重要的一層，常見的在 MPEG-2 壓縮領域的鏡頭變換偵測方法[4]大多是運用 Macroblock 的編碼資訊來進行換鏡偵測的研究，Macroblock 是進行位移估測及位移補償的基本單位，MB 是由畫面上  $16 * 16$  的區塊所組成，根據色彩取樣格式的不同，一個 MB 裡所含的 Block 數目也會有所不同，如 YUV 4:2:0 的色彩取樣格式，一個 MB 有 4 個 Y，一個 Cb 和一個 Cr 這六個 block，如圖 2.3 所示。

## Macro Block Layer



## Block Layer

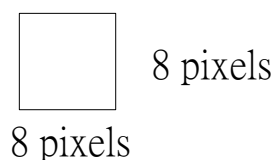


圖 2.3 Macro block 的架構圖

依據 MB 的編碼方式的不同，MPEG-2 定義了四種 MB 的型態。

### 1.IMB (Intra-coding MacroBlock)

若 MB 的編碼型態是 IMB(Intra-coding MacroBlock)，表示這個 MB 的編碼是直接使用這個 MB 的資料進行 DCT 編碼，並沒有參考到其他的畫面上的資料，也沒有位移向量的資訊。

### 2.FMB (Forward-coding MacroBlock)

若 MB 的編碼型態是 FMB(Forward-coding MacroBlock)，表示這個 MB 的編碼有參考到過去播放過的畫面上的資料，這個 MB 可以在向前參考的畫面中進行位移估測及位移補償來產生位移向量及預測誤差。

### 3.BMB (Backward-coding MacroBlock)

若 MB 的編碼型態是 BMB(Backward-coding MacroBlock)，表示這個 MB 的編碼有參考到未來將播放的畫面上的資料，這個 MB 可以在向後參考的畫面中進行位移估測及位移補償來產生位移向量及預測誤差。

### 4.BIMB (Bi-directional Interpolated)

若 MB 的編碼型態是 BIMB(Bi-directional Interpolated)，表示這個 MB 的編碼有參考到過去播放過及未來將播放的畫面上的資料，這個 MB 可以在向前參考及向後參考的畫面中進行位移估測及位移補償來產生位移向量及預測誤差。

而不同類型的畫面中可能出現的 MB 型態也受到畫面的編碼方式的影響，因為 I 畫面只有用本身單張資料壓縮，因此在 I 畫面裡的 MB 只有 IMB(Intra-coding MacroBlock)這種型態，而在 P 畫面中可以利用本身單張資料壓縮或是向前參考的方式編碼，所以在 P 畫面裡的 MB 可能會出現 IMB (Intra-coding MacroBlock)和 FMB (Forward-coding MacroBlock)這二種 MB 型態，而在 B 畫面中可以利用本身單張資料壓縮，向前參考，向後參考，雙向參考的方式編碼，所以這四種型態的 MB 都有可能出現在 B 畫面中。Macroblock 的參考關係如圖 2.4 所示。





空間域方塊中像素個數相同的頻率域係數方塊；如果將此頻率域的係數資料以 (2.2)式之 IDCT 做轉換後，再將每個像素值加上 128，即可還原原來的數位灰階影像。

離散餘弦正轉換和離散餘弦反轉換的公式如下：

$$D(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \dots\dots\dots(2.1)$$

$$f(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j)D(i, j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \dots\dots\dots(2.2)$$

(2.1)式及(2.2)式中的(i,j)係指頻率域之係數位置。(x,y)係指空間域像素的位置。f(x,y)係指(x,y)位置之像素值減去 128 後的值，D(i,j)係指(i,j)位置上的頻率係數值；N 表示方塊的寬度及長度，此處 N=8。其中 C(i)定義為：

$$C(i) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } i = 0 \\ 1 & \text{otherwise} \end{cases} \dots\dots\dots(2.3)$$

在 DCT 轉換後的頻率域的座標 D(0,0)的係數稱為 DC 值代表空間域上 8X8 block 所有像素平均值的 8 倍，其它的 63 個值稱為 AC 值，DC 值是最低頻的部份而頻率由左上至右下逐漸遞增，低頻部份的係數與影像的輪廓有關，而高頻部份則是與影像的細節部份有關，自然影像經過轉換後能量會集中在低頻處，高頻處的能量會趨近於零，經 DCT 轉換後並沒有任何資料損失，故沒有壓縮的效果。利用能量集中的特性，進行量化處理讓高頻處的係數儘量變成零，讓原本非零的係數的分佈範圍儘量縮小，以利壓縮，量化是一種破壞性的壓縮技術，量化後的資料再還原時會和原來的資料不完全相同，因為人眼對影像細節的變化辨識能力並不高，所以可以在高頻處直接忽略，對人眼來說差異性並不大，可因此達到壓

縮的目的，JPEG，MPEG-2，H.263 這些視訊及影像壓縮標準都是採用 DCT 及量化的技術來達成影像空間上的壓縮效果。

## 2.3 MPEG-2 視訊在時間軸上的壓縮理論

在視訊影片中，相鄰兩個畫面的時間僅相差 1/10 至 1/30 秒，故內容相似度很高，如果每張畫面都使用前節的 DCT 及影像儲存處理的話，在空間上會花費許多資源，而且會重覆儲存許多相同的部份，因此，MPEG-2 使用位移估測(Motion Estimation)及位移補償(Motion Compensation)的方法來編碼這些重覆的資料，這稱為 inter-coding 的編碼方法。

在 P 畫面或是 B 畫面中的編碼過程中，會先將目前的畫面分割為 16x16 的 Macroblock，然後在被參考的畫面中以相同座標為中心，在附近尋找相似的區塊，該最相似區塊稱為 best-matched block，並且計算這二個區塊的座標差值以及這二個區塊的內容差值，分別稱為位移向量(Motion Vector)和預測誤差(Prediction Error)，這個過程稱為位移估測(Motion Estimation)，編碼時只要記錄位移向量(Motion Vector)和使用 DCT 編碼，經量化及 run-length encoding(RLE)預測誤差(Prediction Error)，可減少記錄的資料量。解碼時，要先取得參考畫面的資料，由位移向量來取得相對應的 16X16 區塊的資料再加上預測誤差就可以回復原來的畫面，這個過程稱為位移補償(Motion Compensation)，如圖 2.5 所示，是一個 P 畫面往前參考的示意圖。

根據參考畫面的不同，MB 的編碼方式可分為三種。如果參考畫面在目前畫面之前，稱為向前參考(forward prediction)，那這個 MB 的編碼型態就是 FMB(Forward-coding MacroBlock)，如果參考畫面在目前畫面之後，稱為向後參考(backward prediction)，那這個 MB 的編碼型態就是 BMB(Backward-coding

MacroBlock)，如果參考畫面前後都有，稱為雙向參考(Bi-directionally prediction)，那這個 MB 的編碼型態就是 BIMB (Bi-directional Interpolated)，在編碼時如果無法在參考的畫面上找到足夠相似的區塊，那就必需將區塊本身的資料進行 2.2 節所描述的 DCT 編碼，那這個 MB 的編碼型態就是 IMB(Intra-coding MacroBlock)，稱為使用 intra-coding 模式編碼。

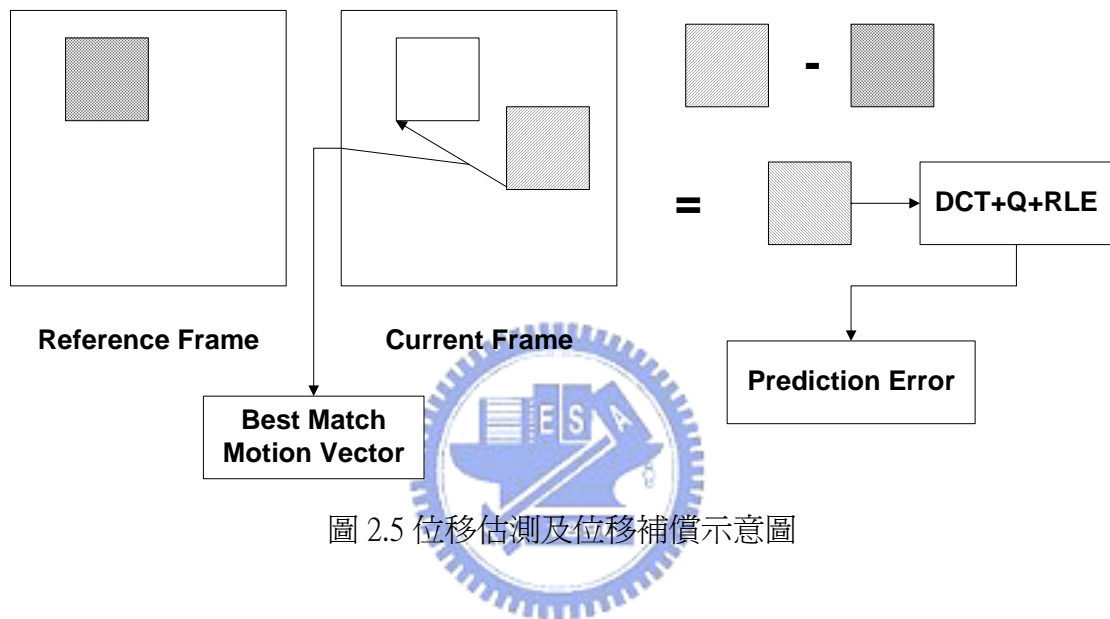


圖 2.5 位移估測及位移補償示意圖

## 第三章 相關研究及文獻探討

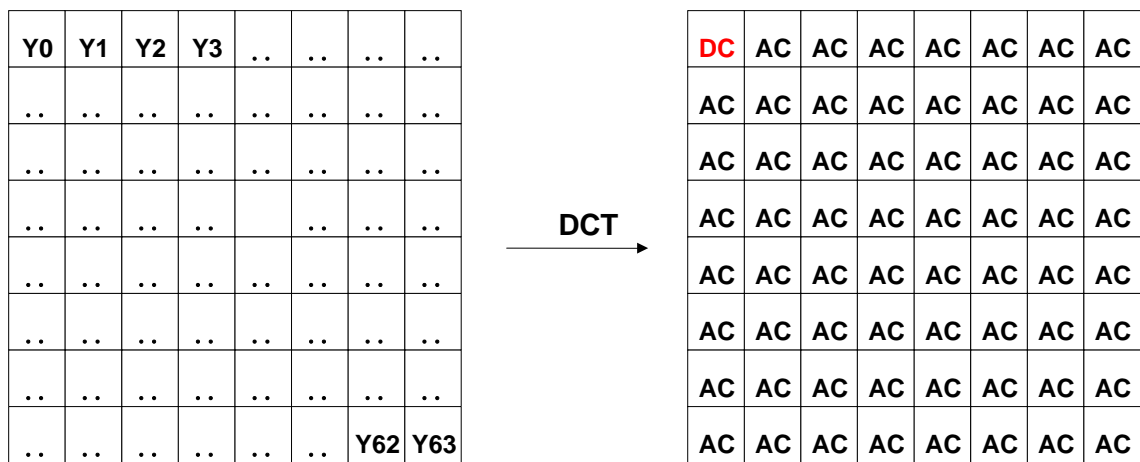
目前已有不少有關視訊影片鏡頭變換偵測的研究[3,4,7]，在本章中，我們將簡介常用之方法。

### 3.1 利用 DC 值偵測換鏡的方法

DCT 編碼會將空間域的資料轉換為頻率域的資料，DCT 編碼處理的基本單位是一個  $8 \times 8$  的 block，block 經過 DCT 編碼之後，能量大部份集中在 DCT 係數中座標如圖 3.1(a)所示，為(0,0)的值上，這個 DCT 係數中座標為(0,0)的值稱為 DC 值，DC 值是原來  $8 \times 8$  的 block 內的所有像素值的平均值的 8 倍，我們可以把每一個方塊的 DC 值除以 8 並當做一個點，組合成一個畫面長度為原圖長度的  $1/8$ ，畫面寬度為原圖寬度的  $1/8$  的縮小圖，稱為 DC 值畫面，DC 值畫面相當於原圖縮小至  $1/64$ ，如圖 3.1(c)所示。







(a) 8X8 block 透過 DCT 轉換



(b) 原始畫面



(c) 原始畫面轉成縮小為 1/64 DC 值畫面

圖 3.1 DC 值畫面示意圖

Yeo 和 Liu[3]提出了用 DC 值畫面來進行換鏡偵測的方法，他們先重建視訊影片中的 DC 值畫面，然後將兩兩相鄰的 DC 值畫面相減算出相鄰畫面的 DC 值差異，再根據 DC 值差異來判斷是否有換鏡狀況的發生。

$$D_{DC}(i, i+1) = \sum_{x,y} (|DC_i(x, y) - DC_{i+1}(x, y)|) \dots\dots\dots(3.1)$$

i, i+1: 兩個相鄰的畫面編號

$DC_i(x, y)$ : 第  $i$  張 DC 值畫面在座標  $(x, y)$  的灰階值

$D_{DC}(i, i+1)$ : 第  $i$  張與第  $i+1$  張 DC 值畫面的絕對差異值

他們提出一個移動窗格(Sliding Window)的概念，就是選用大小為  $2m-1$  的移動窗格，在偵測換鏡時會前往後多檢查  $m$  張 DC 值畫面的差異，來減少因為 zooming 或是 fast panning 所造成的誤判現象。

在下列二項條件都符合時，就判定在第  $i$  個畫面和第  $i+1$  個畫面間是有瞬間式換鏡的情況發生。

1.  $D(i, i+1)$  是大小為  $2m-1$  的移動窗格內的最大值。
2.  $D(i, i+1)$  與  $2m-1$  的移動窗格內的第二大的差異值的比值大於某一個門檻值。



使用第二個條件是因為在有閃光燈的情況下會有連續幾個 DC 值畫面的差異值都很大，而且非常接近，所以在移動窗格內最大的差異值與第二大的差異值會相當接近，而這二個最大及次大的差異值會比移動窗格內的其他差異值大很多，所以可以用第二個條件來排除閃光燈，zooming 以及 fast panning 所造成的誤判。在編碼過的 MPEG-2 影片中，DC 值不需要做 IDCT 轉換就能取得，故本方法之時間複雜度很低。缺點是在漸進式的換鏡正確率並不是很高，依照 Yeo 和 Liu[3] 在論文中提到瞬間式換鏡正確率約 99%，漸進式的換鏡正確率約 89.5%。

### 3.2 利用大區塊資訊的換鏡偵測法

Macroblock 是 MPEG-2 標準裡進行位移估測及位移補償的基本單位，由 MB

的編碼型態和位移向量的資訊以及 MB 的參考關係可得知目前畫面與前後畫面的差異，故利用這些資訊可進行鏡頭變換的偵測。

Pei 和 Chou[4]提出一個利用 Macroblock 編碼類型的資訊來進行偵測鏡頭變換的方法。

他們先將一個 GOP 再細分為 Subgroup of pictures(SGOP)，其中每個 SGOP 可能有重疊，SGOP 依照畫面編碼的型態可以再分類為 IBBP，PBBP，PBBI 這三種組合，並利用 SGOP 中 B 畫面之參考狀況來決定是否存在換鏡。如圖 3.2 所示，在第一個 I(P)畫面和第一個 B 畫面中有一個換鏡的情況發生，由於後三個畫面屬於同一個鏡頭，擁有很高的相似度，因此中間二個 B 畫面中，絕大部份 MB 都會向後參考。

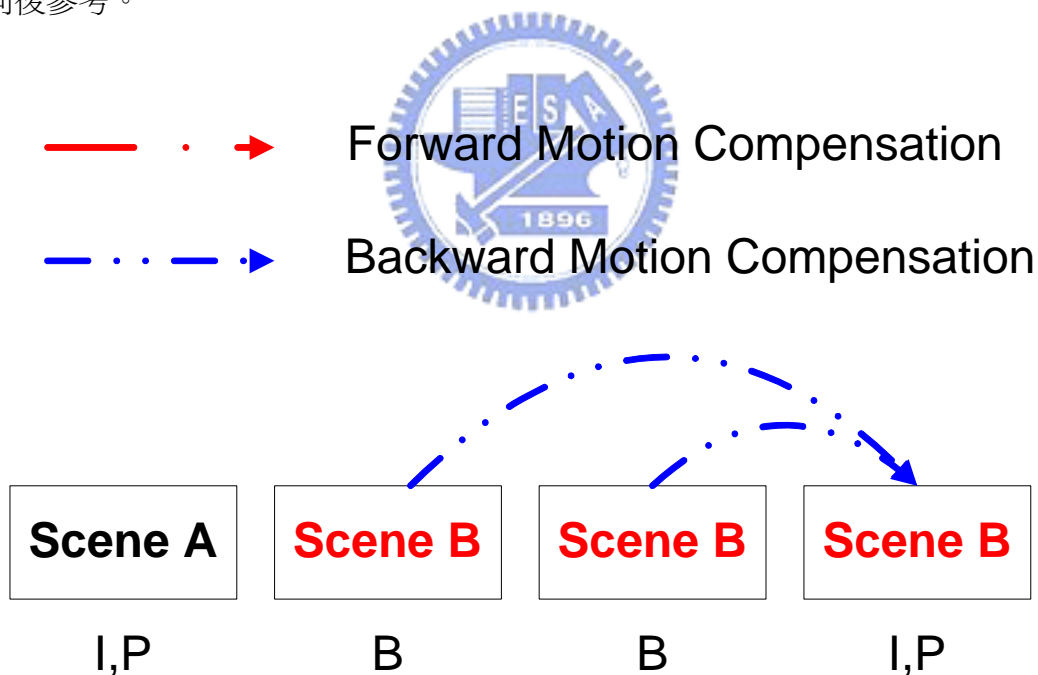


圖 3.2 換鏡時的前後參考關係

而他們有提出一個Changing MB的概念，就是在SGOP的四張圖片中的同一個位置上的MB如果符合Changing MB的式樣，如圖 3.3 所示，第一個B畫面和第二個B畫面的同一個位置上的MB都是屬於向前參考的編碼方式，表示第一個B畫面和第二

個B畫面跟第一個I(P)畫面在同一個位置上的MB相似性很高，但是第二個I(P)畫面中同一位置的MB就不屬於向前參考的編碼方式，這表示第一個I(P)畫面和第二個I(P)畫面在這同一位置的MB相似性比較低，在第二個I(P)畫面中同一位置的MB稱為Changing MB，所以在此張畫面當中的Changing MB的總數若大於門檻值，就判定有瞬間式換鏡的情況發生。

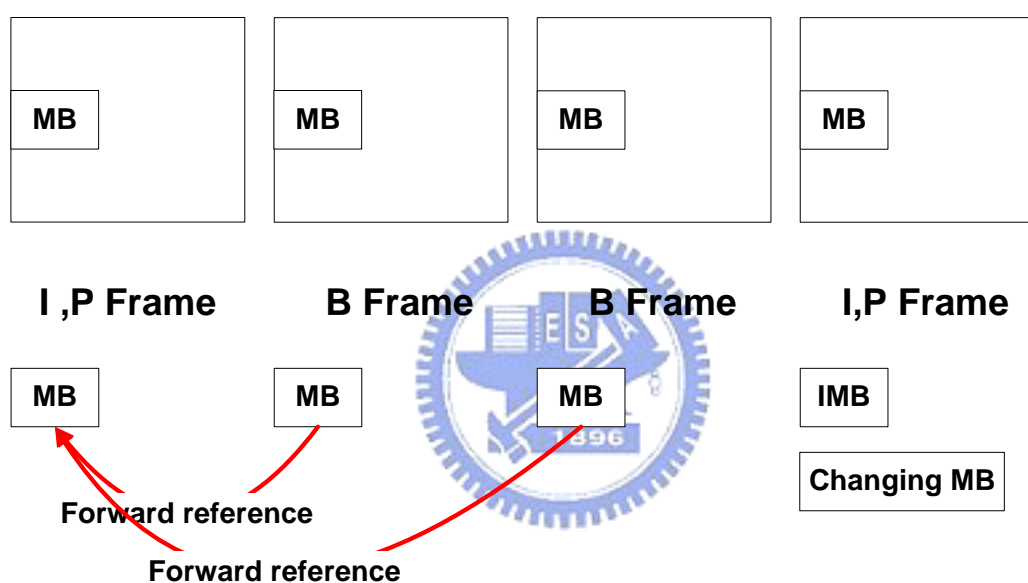


圖 3.3 換鏡時的 Changing MB 的關係

此方法的優點為，只要做可變長度解碼(RLE)即可得到 MB 的編碼型態而進行換鏡偵測，速度會比用 DC 值畫面做比較的方法更快，而且簡單。其缺點依 MB 編碼的方式而有差異,容易影響偵測的結果。

### 3.3 自動設定門檻值的方法

不同類型的 MPEG-2 影片，其畫面間的變化程度是不同的，如在運動型態的

影片中的每一個畫面的差異可能都很大，但是卻不見得是有鏡頭轉換情況發生，而只是連續運動的畫面，而在偏靜態的影片如新聞播報的影片中每一個畫面的差異都不大，就算是有鏡頭轉換的情況發生，畫面的差異也有可能比連續運動的畫面差異小。因此，對所有影片使用同一個門檻值來偵測所有類型的視訊影片是不實際的。林文隆在其碩士論文[7]中提出一個自動決定門檻值之換鏡偵測方法，藉由統計影片之差異性，自動的為各種不同型態的影片算出用以偵測特徵之門檻值，如此能有效的增加瞬間式換鏡偵測及漸進式換鏡偵測的正確率。

對一個未經壓縮的影片，首先透過公式(3.2)計算相鄰兩畫面的差異值。圖 3.4 列出一個視訊影片相鄰兩畫面的差異值。由圖 3.4 可以發現，當發生漸進式轉換時，轉換之畫面其差異值均會上生升，而發生瞬間式轉換時，該畫面差異值會大幅上升，可由此變化來偵測換鏡。這一個視訊影片有 3 個漸進式轉換，2 個瞬間式的鏡頭轉換。



$$D_x(i, i+1) = \sum_{x,y} (|X_i(x, y) - X_{i+1}(x, y)|) \dots\dots\dots(3.2)$$

i: 影片第 i 個畫面

i+1: 影片第 i+1 個畫面

$X_i(x,y)$ : 影片第 i 個畫面在座標(x,y)的灰階值

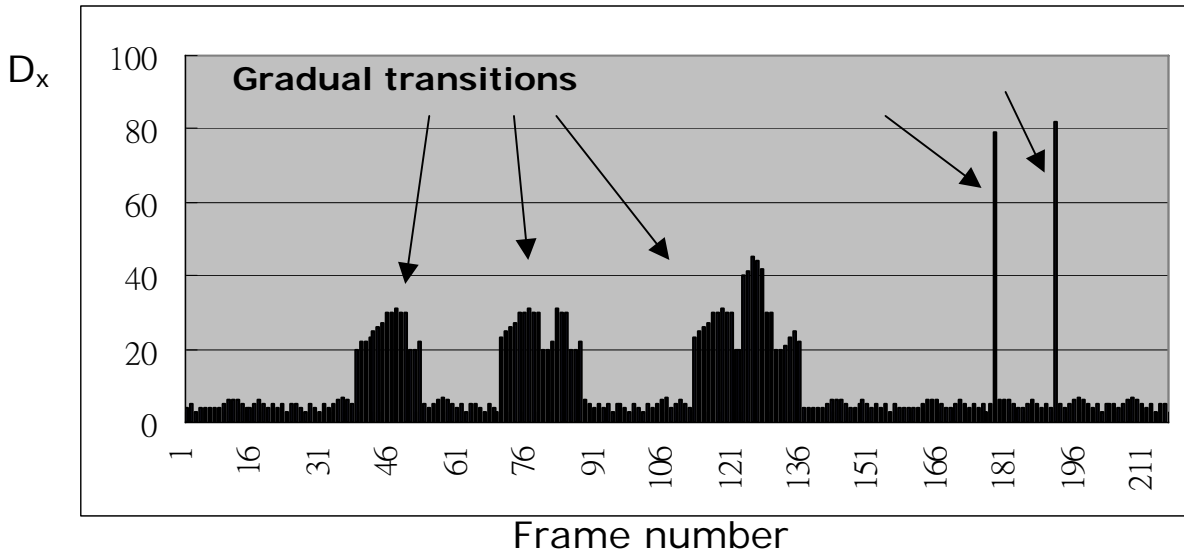


圖 3.4 一個視訊影片中連續兩畫面的差異值

爲了取得一個適當的門檻值，該論文將所有相鄰兩畫面差異值統計爲直方圖 (Histogram)，橫座標表示差異值 $D_x$ ，縱座標表示差異值出現的次數。圖 3.5 是圖 3.4 畫面差異值的直方圖。在圖 3.5 中可以分析出三個值較高的區域，在 X 軸(畫面的差異值)上由小到大的第一個區域是沒有換鏡轉換的區域，因爲在一段影片中的換鏡畫面的數量遠比沒有換鏡轉換的畫面少很多，而沒有換鏡時的畫面內容相似度很高，兩兩畫面間的差異值也很小，這些差異值是集中在第一區域。第二個區域是漸進式轉換分布的區域，因爲在漸進式轉換時畫面間的差異值會比沒有換鏡時的畫面差異值大，而且漸進式轉換時的畫面數量也比沒有換鏡時的畫面數量小，所以會在沒有換鏡轉換的區域的正常分布之後，會再有一個小聚集代表是

漸進式轉換分布的區域，而在瞬間式換鏡時的畫面的差異值比漸進式換鏡時的畫面的差異值大很多，但是瞬間式轉換時的畫面數量也比漸進式換鏡時的畫面數量小很多，故在差異值大的地方，會有數個統計值為 1 或 2 之值，這是代表瞬間式轉換的第三區域。林文隆之方法[7]將漸進式換鏡的門檻值設為在沒有換鏡轉換的區域與漸進式轉換分布的區域中畫面數量最小的點的畫面差異值，因為在沒有換鏡轉換的區域的正常分布之後再出現另一個畫面聚集的區域代表已經不是屬於這個影片的正常無換鏡的分布而是漸進式換鏡的分布。而瞬間式換鏡畫面差異值的門檻值會出現在直方圖中連續數個畫面數量為零之後再次出現非零的畫面差異值，圖 3.6 是該方法的示意圖。

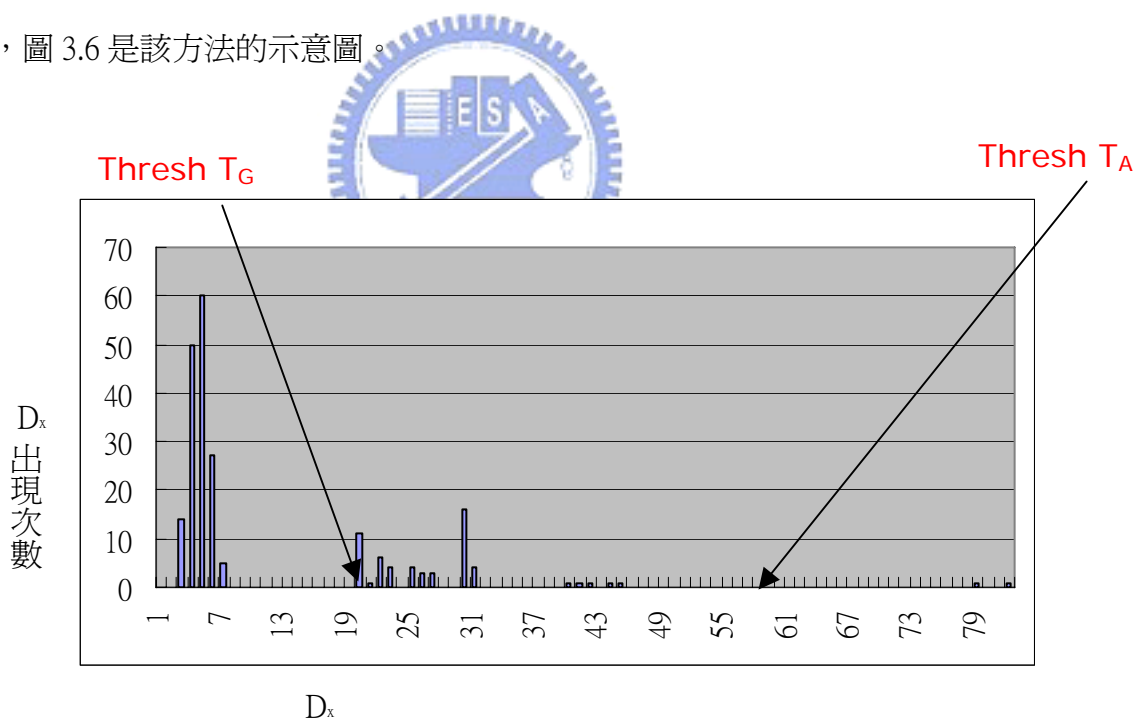


圖 3.5 圖 3.4 的直方圖

Threshold  $T_A$ : Abrupt Transition threshold, 代表瞬間式換鏡的畫面差異值的門檻值。

Threshold  $T_G$ : Gradual Transition threshold, 代表漸進式換鏡的畫面差異值的門檻值。

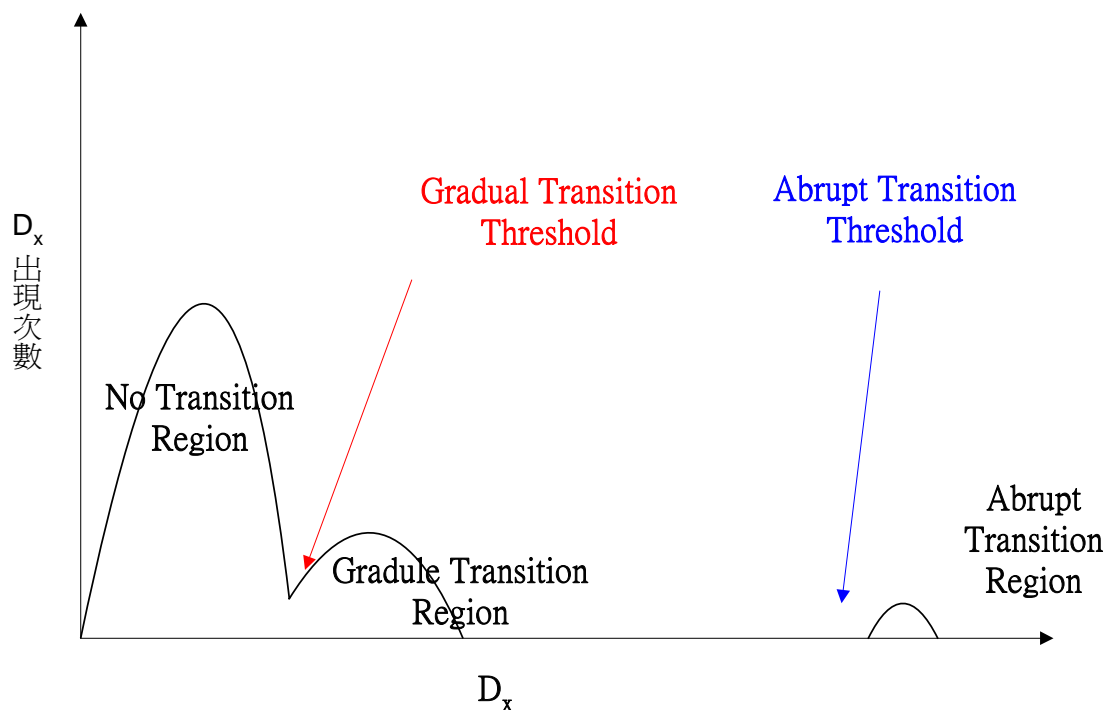
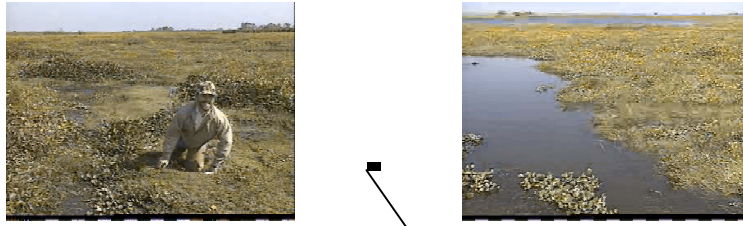


圖 3.6 林文隆方法的示意圖

在判定瞬間式換鏡偵測上，如果兩兩相鄰的畫面的差異值大於 $T_A$ 則認定有瞬間式換鏡的情況發生。在判定漸進式換鏡偵測上，如果兩兩相鄰的畫面的差異值小於 $T_A$ 但是大於 $T_G$ ，則將此畫面認定為可能為漸進式轉換的進入畫面，而後找尋兩兩相鄰的畫面的差異值小於 $T_G$ 的畫面將此畫面認定為可能為漸進式轉換的離開畫面，再求出可能為漸進式轉換的進入畫面與離開畫面的畫面差異值如果大於 $T_A$ 則認定有漸進式換鏡的情況發生。

該方法之優點是能自動求出每段影片適合的畫面差異值的瞬間式換鏡的門檻值及漸進式換鏡的門檻值，這能有效的提高換鏡偵測的正確率及減少誤判率。缺點是此方法利用未壓縮之視訊，故需將 MPEG-2 資料流解開為原始視訊，處理量及時間複雜度都較大，且在某些情形會有無法偵測到瞬間式換鏡的情形，圖 3.5 為該方法實際應用之一例。





瞬間式換鏡

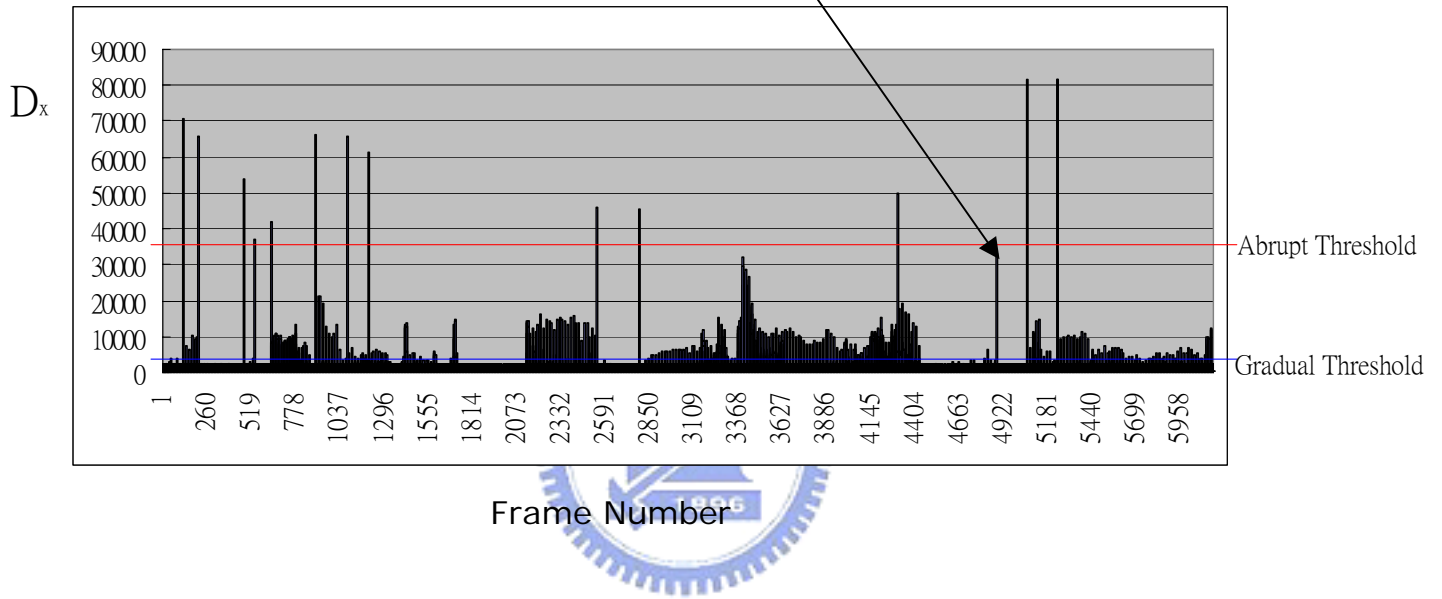


圖 3.7 一個無法偵測到瞬間式換鏡的例子

## 第四章 本論文所提的鏡頭變換偵測法

雖然林文隆[7]所提出具適應性之偵測法能得到較好之偵測率，但需將 MPEG-2 資料流解碼為無壓縮資料所損耗之時間大大影響了其實用性。為了節省計算量，我們修改林文隆的方法，並提出一能不需將 MPEG-2 完整解碼且具適應性的偵測法。我們的方法能依影片型態自動得到適合的門檻值來比較 DC 值畫面，並利用畫面裡的 MB 編碼類型和畫面裡的 MB 的位移向量來進行鏡頭變換的偵測。我們的方法分為兩個階段，第一階段先將 MPEG-2 資料流經過可變長度解碼(Variable length decoding)及變動長度解碼(Run Length decoding)，重建每一個 MB 的 DC 值並組成 DC 值畫面，接著利用 DC 畫面計算 I,P 畫面的 DC 差異值，並統計 P 畫面裡編碼方式為 intra-coding 的 MB 個數,及 P 畫面裡具最主要位移向量的 MB 個數。接著利用這些數據計算該 MPEG-2 影片的瞬間式換鏡及漸進式換鏡的 DC 差異值的門檻值, intra-coding MB 個數的門檻值,和具最主要位移向量的個數門檻值。第二階段則依據第一階段所得到的三種資訊的門檻值, 來偵測影片中瞬間式換鏡及漸進式換鏡的位置，圖 4.1 表示我們提出方法的流程圖。

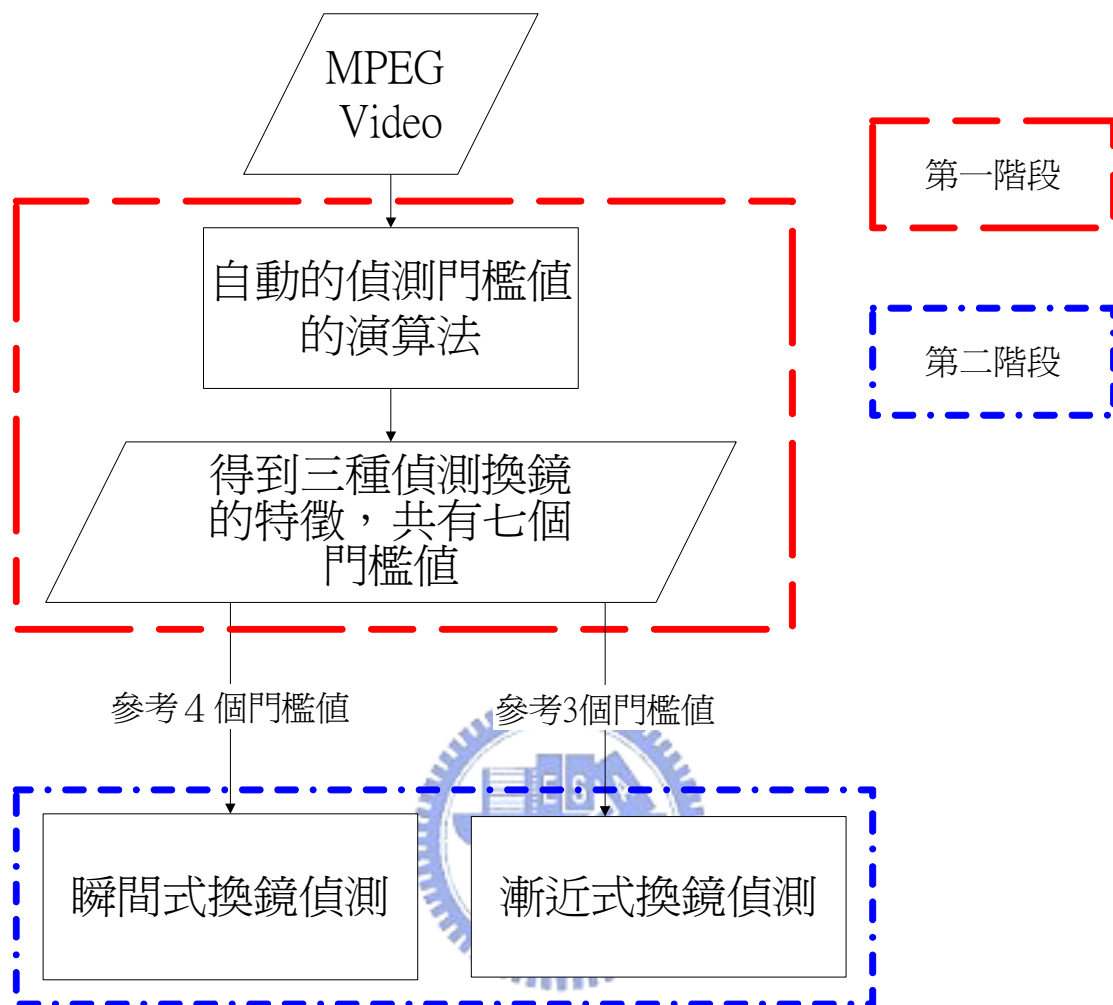


圖 4.1 我們提出之換鏡偵測法的流程圖

#### 4.1 換鏡偵測所需要的三種參考特徵

在第一階段中須先將 MPEG-2 影片每一畫面經過可變長度解碼(Variable length decoding)，以取得每一個 MB 的 DC 值並建立 DC 值畫面，然後擷取三種可用來偵測換鏡時的特徵，第一種特徵是距離最近的二個 I 畫面或是 P 畫面 DC 值畫面的差異值，在影片中 DC 畫面差異值呈現出下面關係，無換鏡畫面之差異值小於漸進式換鏡畫面之差異值，而漸進式換鏡畫面之差異值會小於瞬間式換鏡之

差異值，第二種特徵是每個 P 畫面裡的 intra-coding MB 的個數，根據 P 畫面向前參考的編碼原理，如果 P 畫面和前一個參考畫面有很強的相似性，那麼在 P 畫面裡的 intra-coding MB 數量就會很小，相反的如果在 P 畫面裡的 intra-coding MB 數量很大的話就表示這個 P 畫面和它的前一個參考畫面幾乎找不到相似的地方，而很有可能是有鏡頭變換的情況發生。由於 P 畫面裡的 intra-coding MB 的數量和影片種類有很大的關係，不應由一預設之門檻值來決定。因此我們提出一個門檻值自動計算之方法，藉由統計一段影片在 P 畫面裡的 intra-coding MB 的數量分布情形來決定 intra-coding MB 的數量在瞬間式換鏡及漸進式換鏡的門檻值，詳細方法將在 4.1.2 節介紹。第三種特徵是 P 畫面裡具有主要位移向量的 MB 個數，這裡主要位移向量是指在 P 畫面中，統計所有 MB 的位移向量，出現最多次的位移向量就稱為主要位移向量，一般而言，如果和前一個參考畫面屬於同一鏡頭則兩畫面的背景部份應該類似，因此背景部份的 MB 應有相似的位移向量，此位移向量會在畫面中出現最多次，因此若一 P 畫面裡的具有主要的位移向量的 MB 個數很多，表示前一個參考畫面與目前的 P 畫面屬於同一鏡頭的可能性較高。且根據我們的觀察，不論在瞬間式換鏡或漸進式換鏡時的 P 畫面裡的 MB 的具主要位移向量數量都明顯比沒有換鏡情況下具主要位移向量的 MB 數量少很多，瞬間式換鏡時的 P 畫面裡的具主要位移向量的 MB 數量又比漸進式換鏡時的具主要位移向量的 MB 數量少，然而因每一段影片裡畫面中主要位移向量的 MB 數量會依照影片型態和運動節奏而有很大的不同，我們利用統計 P 畫面裡的 MB 在主要的位移向量的數量來找出適當的門檻值以偵測瞬間式換鏡及漸進式換鏡。

因為在解碼取得 DC 值畫面時也已經得知每個 MB 的編碼方式和 MB 的位移向量，所以取得上述三種特徵花費的時間很少。

### 4.1.1 DC 畫面差異值及門檻值設定

我們利用(3.1)式計算兩個相鄰的 I 或 P 畫面的 DC 畫面差異值。對所有相鄰 I(P) 畫面的 DC 畫面差異值建立一直方圖(histogram)，再接著利用林文隆所提出之自動取得門檻值的方法對瞬間式換鏡及漸進式換鏡各找出一個適當的偵測換鏡的門檻值，由於林文隆的方法是取在統計的直方圖中 Y 軸上連續出現數個零之後的 X 軸的畫面差異值，但是影片中有些瞬間式換鏡的畫面差異值可能會落在這個門檻值與漸進式換鏡的門檻值中間，因此有可能被忽略，如圖 3.5 所示。我們使用第三個門檻值，將直方圖中漸進式換鏡的小聚落之後的第一個在 Y 軸上為 0 的畫面差異值也認定為可能會發生換鏡的候選門檻值，然後再以 P 畫面中的 intra-coding MB 數量及具主要位移向量的 MB 數量來做為換鏡偵測的依據。

在圖 4.2 顯示一影片 DC 畫面差異值可以看出，在換鏡時的前後兩個 I 或 P 畫面的差異值很大而在一般連續的鏡頭裡，前後兩個 I 或 P 畫面的差異值很小。

圖 4.3 為圖 4.2 的直方圖。圖 4.4 為尋找門檻值的示意圖，三個門檻值分別標於圖上。

無換鏡

瞬間式換鏡

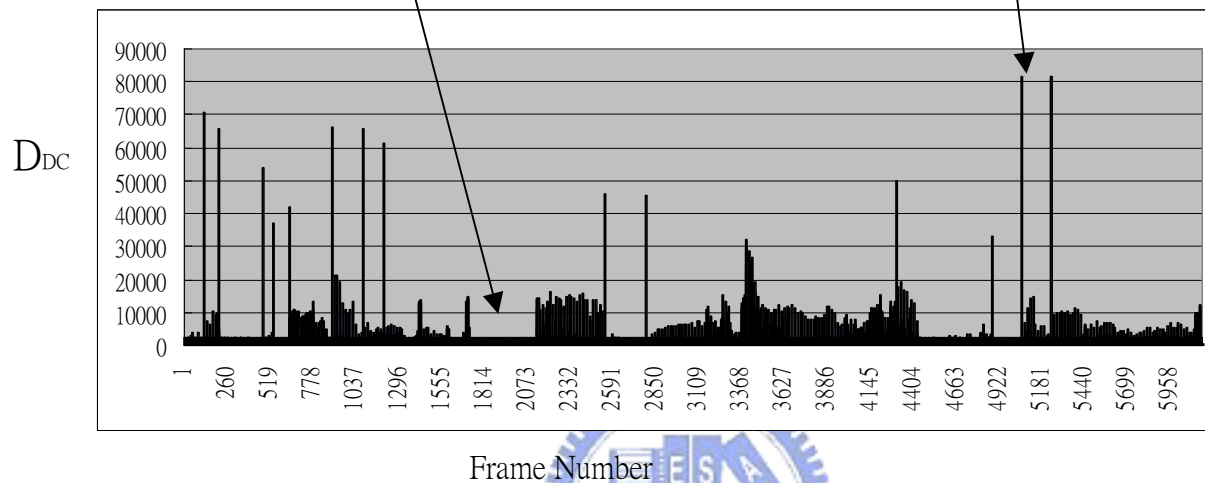
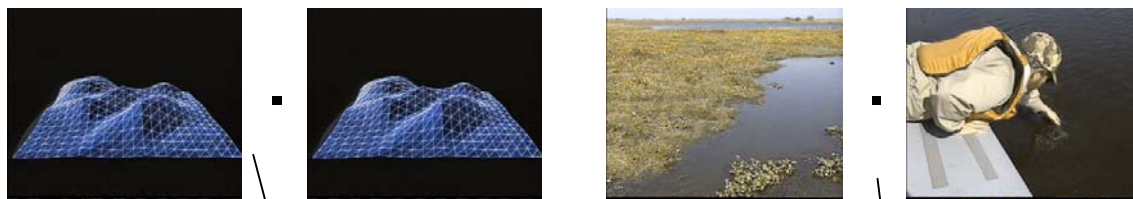


圖 4.2 相鄰兩 I 或 P 畫面的 DC 畫面差異值

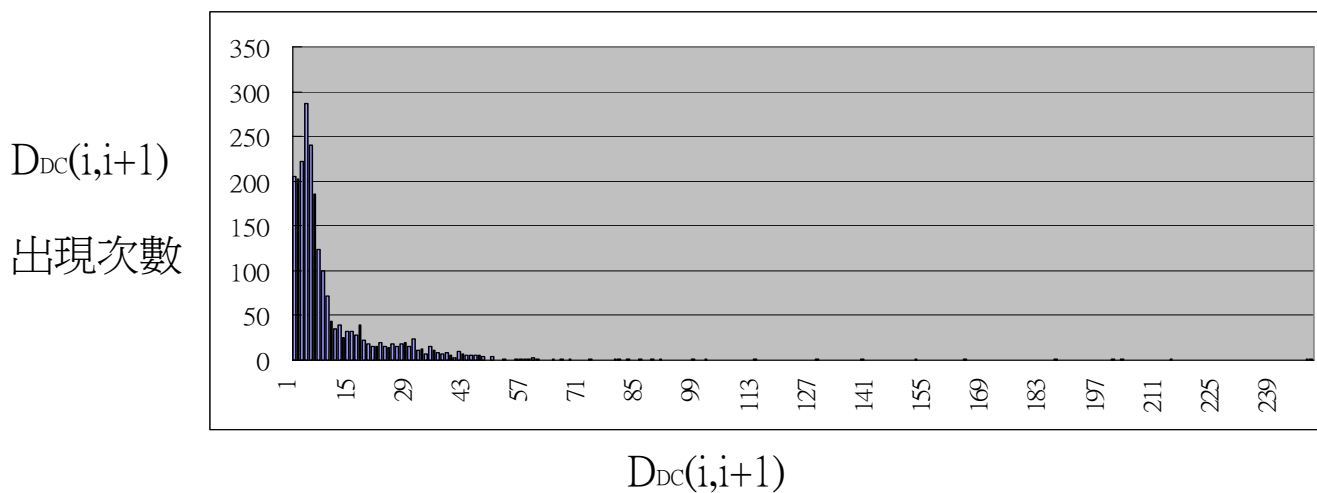


圖 4.3 圖 4.2 的直方圖

$T_{G\_DC}$ : 位於無換鏡區域及漸進式換鏡區域的谷底所代表的DC差異值。

$T_{A1\_DC}$ : 在 $T_{G\_DC}$ 之後，第一個出現0的DC差異值。

$T_{A2\_DC}$ : 為連續的0之後，第一個出現非0的DC差異值。

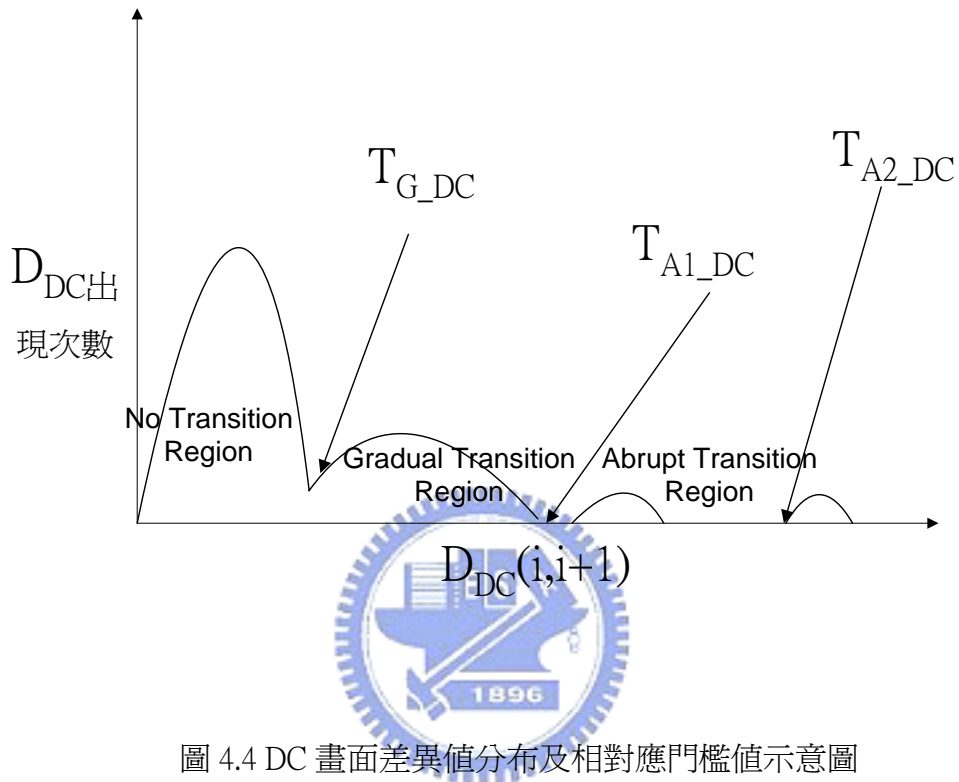


圖 4.5 顯示出圖 4.3 中之三個門檻值。以上所述的三個門檻值，將作為第二階段換鏡偵測的參考。

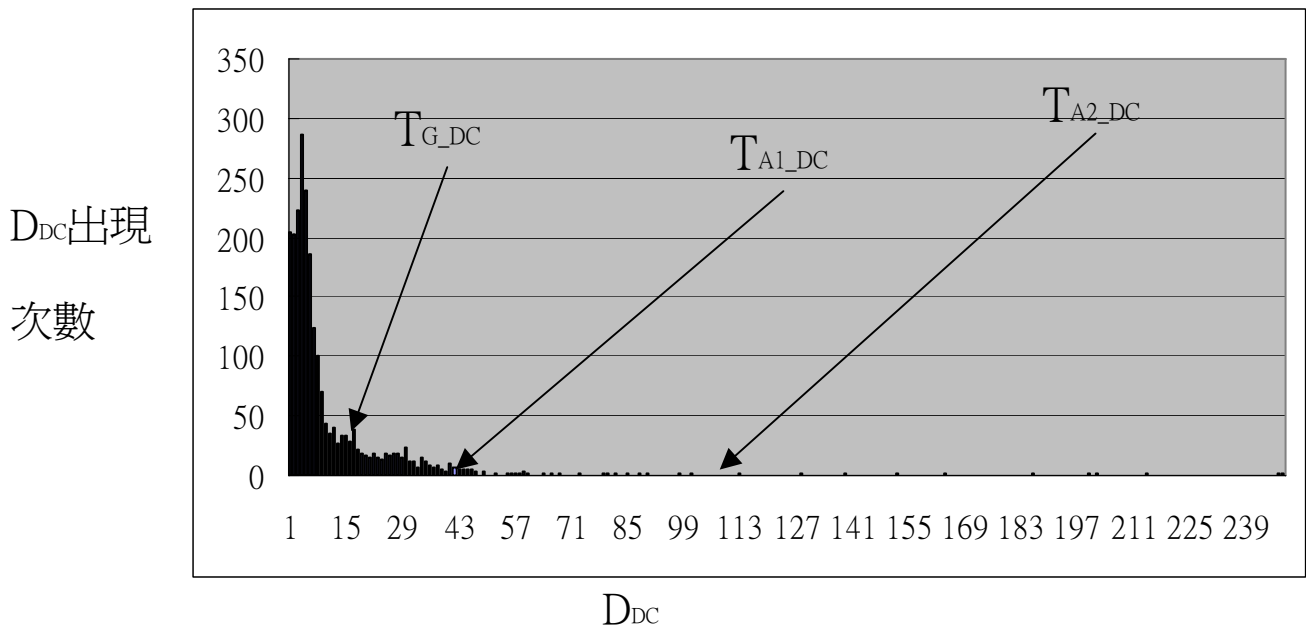


圖 4.5 圖 4.3 中之三個門檻值

#### 4.1.2 在 P 畫面中的 intra-coding 的 MB 總數的分布統計

若於 P 畫面之前出現一個換鏡，會因為無法向前參考的關係，而使得畫面中的 intra-coding 的 MB 個數增多。圖 4.6 顯示一個影片中各個 P 畫面中 intra-coding MB 的個數，由圖中得知在漸進式換鏡時 P 畫面中的 intra-coding MB 的數量會比沒有換鏡時的畫面的 intra-coding 的 MB 的數量大很多，而瞬間式換鏡時 P 畫面中的 intra-coding 的 MB 的數量又會比漸進式換鏡時的畫面的 intra-coding 的 MB 的數量大很多。



### 漸進式換鏡

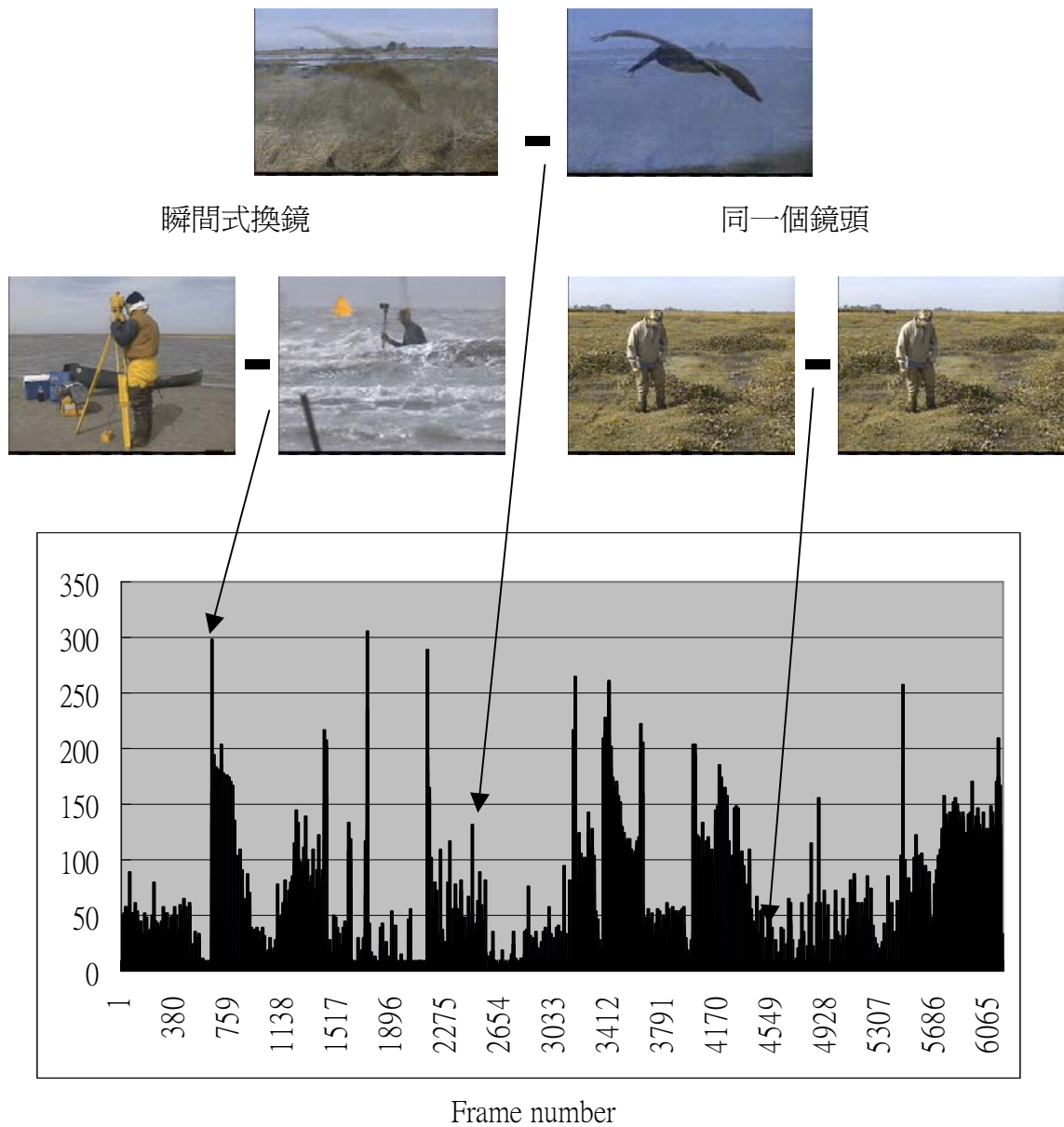


圖 4.6 P Frame 裡 intra-coding 的 MB 數的分布圖

圖 4.7 是圖 4.6 的直方圖(histogram)統計，由此直方圖可知大部份的 P 畫面的 intra-coding 的 MB 的數量都是不多的，這是因為大部份的畫面屬於沒有換鏡轉換的畫面，所以在直方圖上會先出現沒有換鏡轉換的畫面的最大聚集，之後會出現其他一個小聚集，我們認定這些小聚集可能是屬於漸進式轉換畫面的區域，而在 intra-coding 的 MB 的數量很大的區域裡有少數的畫面分布在此區域裡，這些畫面

極有可能是瞬間式換鏡的畫面。

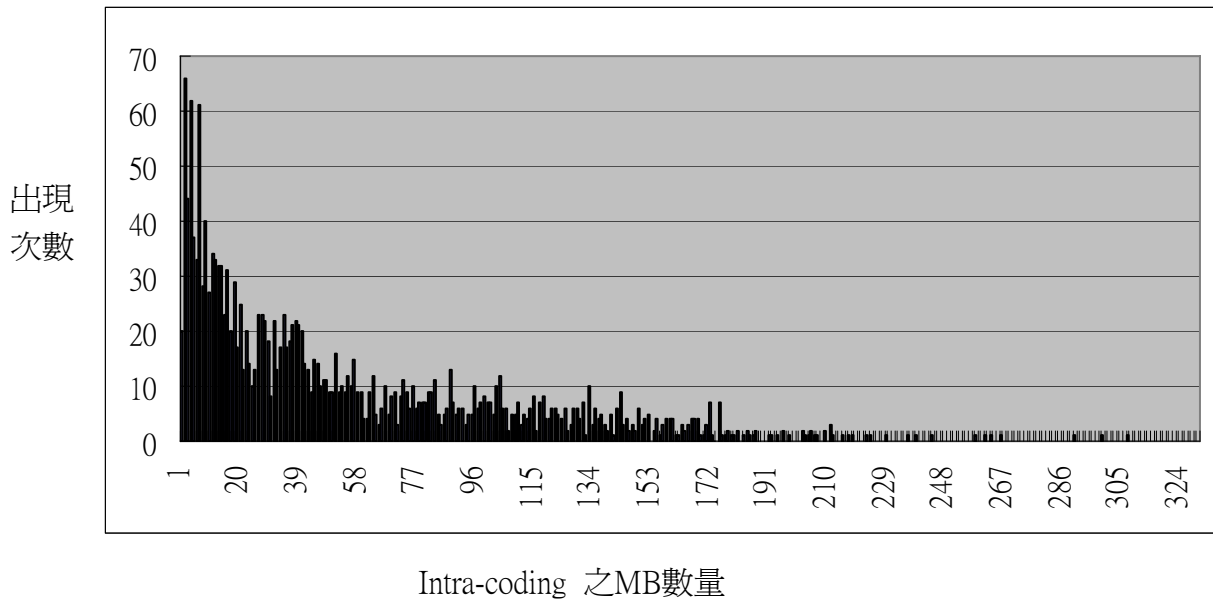
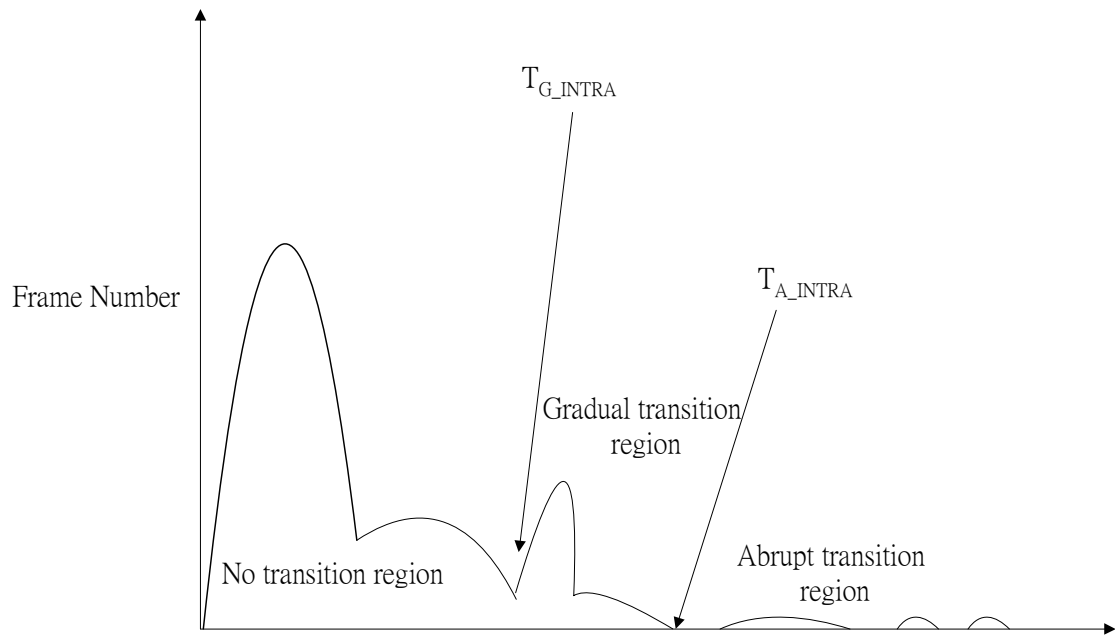


圖 4.7 圖 4.6 的直方圖

如圖 4.8 所示，我們可以針對不同的影片的P畫面的intra-coding MB的數量的直方圖來找出合適的漸進式轉換換鏡及瞬間式換鏡在P畫面的intra-coding MB的數量的門檻值，在直方圖中的代表無換鏡區域的最大聚集之後出現具有最大上昇斜率的谷底值，我們認定為漸進式換鏡在P畫面的intra-coding MB的數量的門檻值 ( $T_{G\_INTRA}$ )，在 $T_{G\_INTRA}$ 之後出現的第一個出現 0 的intra-coding MB的數量，我們認定其為瞬間式換鏡在P畫面的intra-coding MB的數量的門檻值( $T_{A\_INTRA}$ )。



The numbers of Intra-coding MB

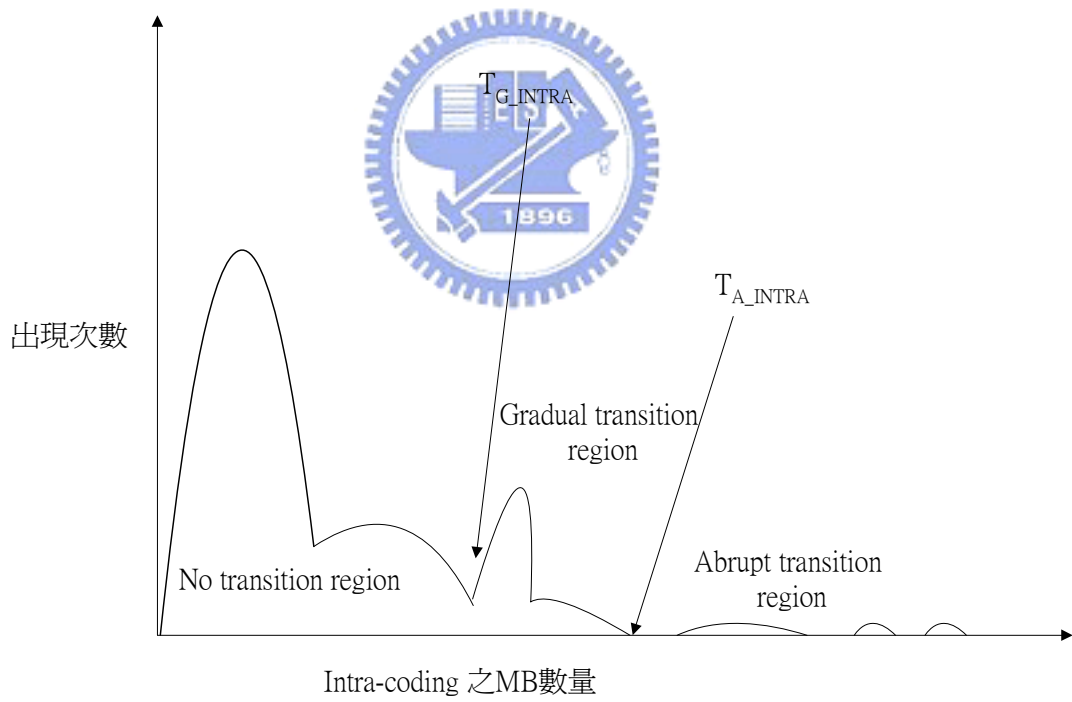


圖 4.8 圖 4.7 的示意圖

$T_{G\_INTRA}$ 和 $T_{A\_INTRA}$ 在第二階段時可做為是否有瞬間式換鏡及漸進式換鏡的

依據。

### 4.1.3 在 P 畫面中的具主要位移向量的 MB 數量的分布統計

因為在 P 畫面中的位移向量是向前參考的，如果有換鏡現象產生，在 P 畫面中有位移向量的 MB 個數會很少，而且在瞬間式換鏡和漸進式換鏡時，在 P 畫面中的位移向量也會有不同的分布表現。同時由於每一段影片的編碼方式，影片節奏及影片型態都會影響 P 畫面中的位移向量的分布，因此我們將提出一方法，對不同的影片自動化的設定在瞬間式換鏡和漸進式換鏡時 P 畫面中的具主要位移向量的 MB 數量的門檻值。

首先對這一段影片的每個 P 畫面算出該畫面中所有 MB 的位移向量，統計所有位移向量之出現次數並將位移向量出現最多次的位移向量訂為此畫面的主要位移向量  $V$ ，並記錄其出現  $N_v$  次數，將  $N_v$  做一統計。畫出直方圖。如圖 4.9 所示。

瞬間式換鏡



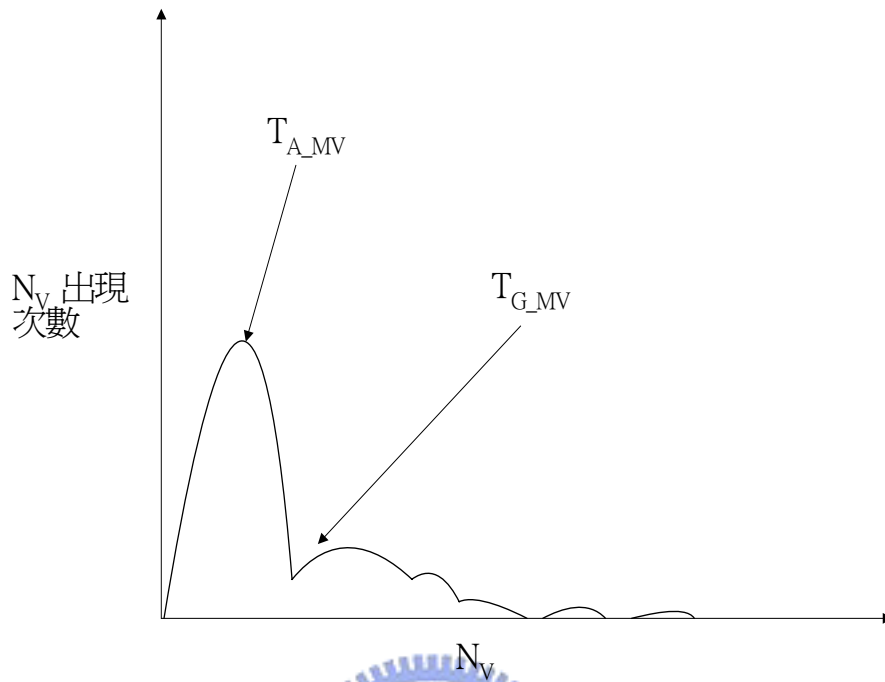
無換鏡





圖 4.9 一個 MPEG-2 Video 中具有主要位移向量的 MB 個數的直方圖

在瞬間式的換鏡畫面中具有主要位移向量的MB數量都很少，而漸近式轉換



的畫面，具有主要位移向量的MB數量也比較少，但比瞬間式的換鏡中的具有主要位移向量的MB數量大，在瞬間式的換鏡畫面中具有主要位移向量的MB數量會比在這一段影片中最多P畫面擁有的具有主要位移向量的MB數量少，所以我們認定最多P畫面擁有的具有主要位移向量的MB數量為瞬間式換鏡在P畫面具有主要位移向量的MB數量的門檻值( $T_{A,MV}$ )，在 $T_{A,MV}$ 之後出現的另一個上升斜率最高的聚集所擁有的具有主要位移向量的MB數量為漸進式換鏡在P畫面具有主要位移向量的MB數量的門檻值( $T_{G,MV}$ )，如圖 4.10 所示。

圖 4.10 MPEG-2 Video 具主要位移向量的 MB 個數的直方圖及門檻值設定示意圖

我們也觀察到在漸近式換鏡過程中，P 畫面具有主要位移向量的 MB 數量會

在漸近式換鏡過程有一個由大變小又變大的過程，而這個具主要位移向量的 MB 數量變化過程在單純的鏡頭移動的非換鏡過程中是沒有的，如圖 4.11 所示，所以我們依據此特徵，在漸近式轉換的過程中偵測是否有這個  $N_V$  由大變小又變大的過程。

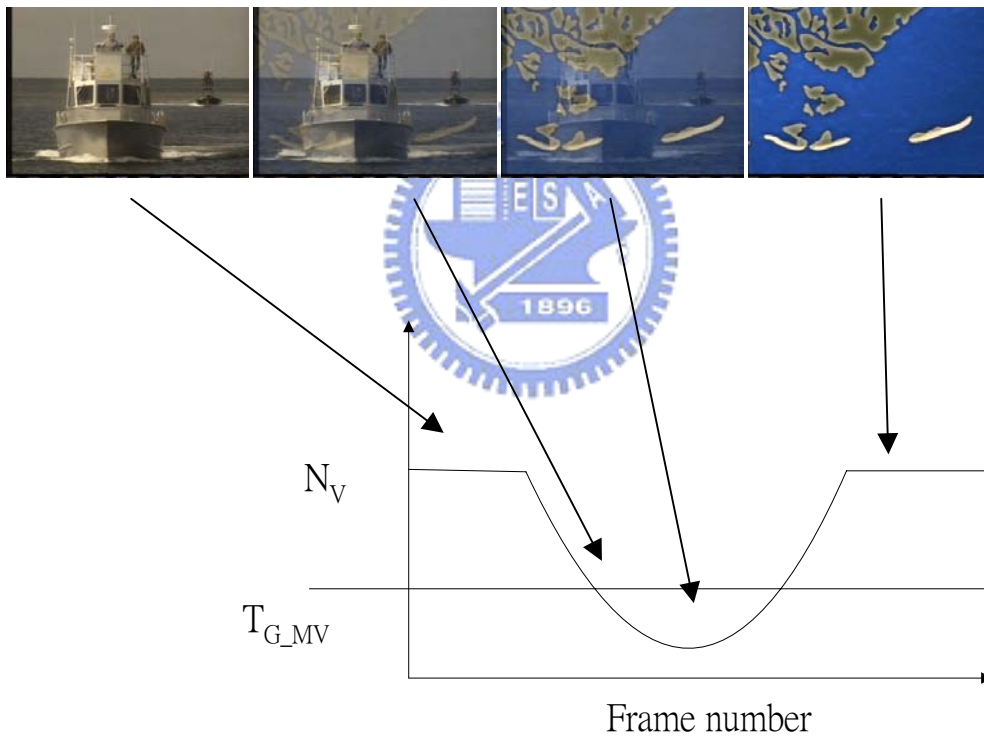


圖 4.11 是一個 MPEG-2 Video 漸進式換鏡的 Motion Vector number 變化示意圖

$T_{A,MV}$ ， $T_{G,MV}$ 這二個在P畫面具有主要位移向量的MB數量的門檻值在第二階段時可做為是否有瞬間式換鏡及漸進式換鏡的依據。

我們在 4.1.1, 4.1.2, 4.1.3 節共取得 7 個特徵值及門檻值，將做為下節偵測換鏡之用。

## 4.2 換鏡偵測的流程

取得 4.1 節所描述的 7 個門檻值之後，本節我們將利用門檻值偵測影片中之換鏡，即圖 4.1 中之第二階段。以下將就瞬間式換鏡及漸進式換鏡之偵測法分別描述。



### 4.2.1 偵測瞬間式換鏡之方法

由於 I 畫面與 P 畫面編碼之方式完全不同，故偵測過程將依畫面種類而有所不同，如圖 4.12 所示。

如果是 I 畫面:

- (1) 如果此 I 畫面與前一個 I 畫面的距離小於正常的 GOP (group of picture) 長度  $L_{GOP}$ ，就認定有一個瞬間式換鏡的發生。因為在一個影片中的 GOP 長度是固定的，會發生 I 畫面與前一個 I 畫面的距離小於  $L_{GOP}$  是因為發生了瞬間式換鏡的關係，導致目前的畫面無法以向前參考的 P 畫面存在，而必需以自己本張畫面的資料編碼。



或是

- (2) 如果此I畫面與前一個I或P畫面的DC值差異大於 $T_{A2\_DC}$ ，就認定有一個瞬間式換鏡的發生。

如果是 P 畫面:

- (1) 如果此P畫面與前一個I或P畫面的DC值差異大於 $T_{A2\_DC}$ ，就認定有一個瞬間式換鏡的發生。

或是

- (2) 如果此P畫面與前一個I或P畫面的DC值差異大於 $T_{A1\_DC}$ ，而且在此P畫面的intra-coding 的MB的數量大於 $T_{A\_INTRA}$ ，而且此P畫面 $N_v$ 小於 $T_{A\_MV}$ 就認定有一個瞬間式換鏡的發生。





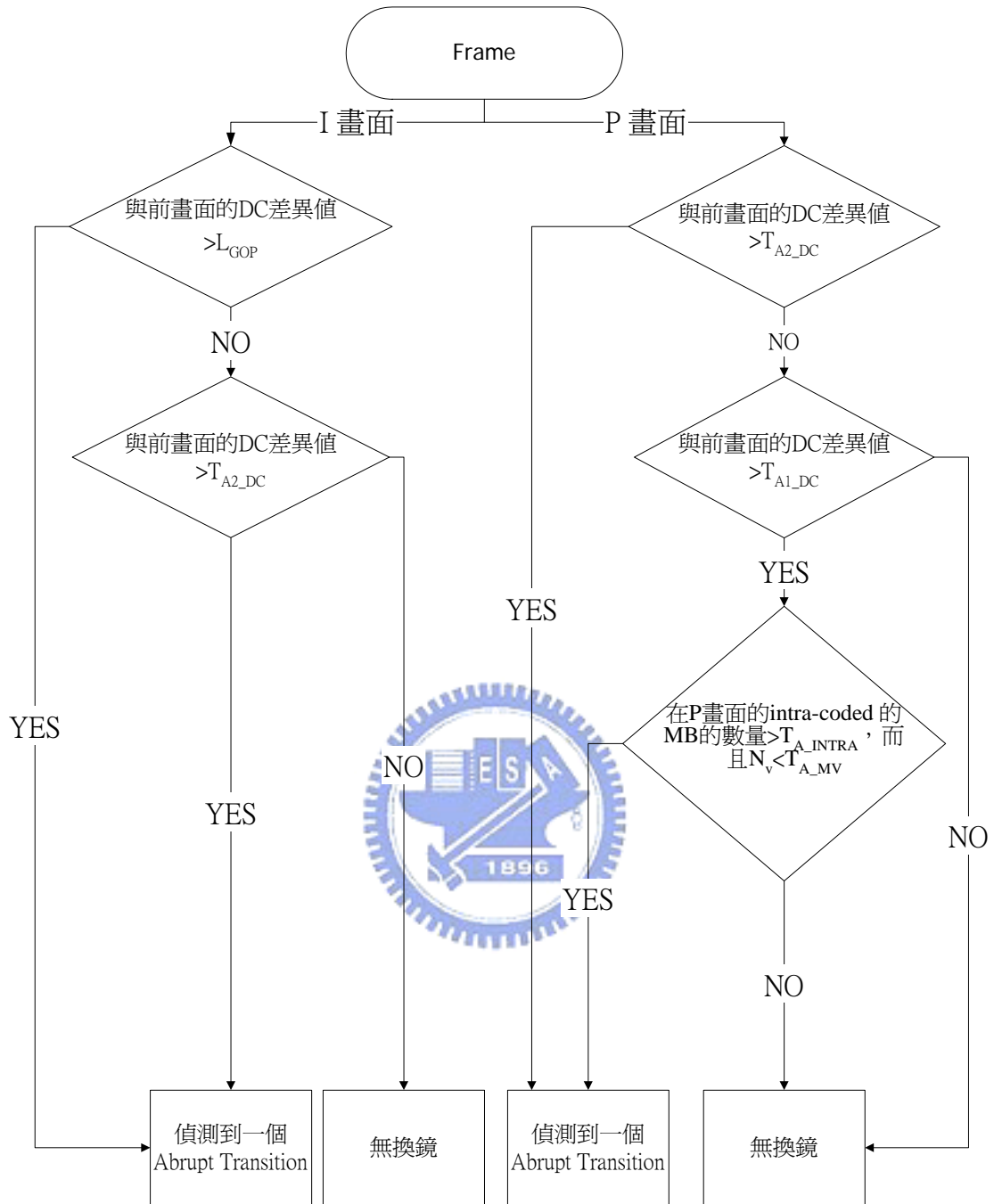


圖 4.12 偵測瞬間式換鏡的流程圖

## 4.2.2 偵測漸進式換鏡之方法

我們先經過第一階段的自動化偵測門檻值的方法來取得這一段影片的偵測

漸進式換鏡的門檻值如下所列。

如果這個I或P畫面在經過瞬間式換鏡偵測之後，認定並沒有發生瞬間式換鏡，就會再做漸進式的換鏡偵測。如果 $D_{DC}(i,i+1) > T_{G\_DC}$ ，則將畫面 $i+1$  認定為可能為漸進式轉換的進入畫面 $F1$ ，並先假設認定目前已進入漸近式換鏡的區間，而後找尋 $D_{DC}(k,k+1)$ 小於 $T_{G\_DC}$ 的畫面，將畫面 $k+1$  認定為可能為漸進式轉換的離開畫面 $F2$ ，當離開可能的漸近式換鏡的區間時會利用以下偵測漸進式換鏡的流程，來偵測此可能的漸近式換鏡的區間是否真的為漸進式換鏡區間，偵測漸進式換鏡的流程圖，如圖 4.13 所示。

(1)  $D_{DC}(F1,F2) > T_{A2\_DC}$  或  $D_{DC}(F1,F2) > \text{轉換區間的畫面數 } L(F1,F2) \times$

$T_{G\_DC}$ ，其中 $L(F1,F2)$ 為轉換區間的畫面數

而且

(2) 在轉換過程中至少有一個P畫面的intra-coding 的MB的數量大於

$T_{G\_INTRA}$ ，而且至少有一個P畫面的 $N_v > T_{G\_MV}$ ，就認定有一個漸進式換鏡

的發生。



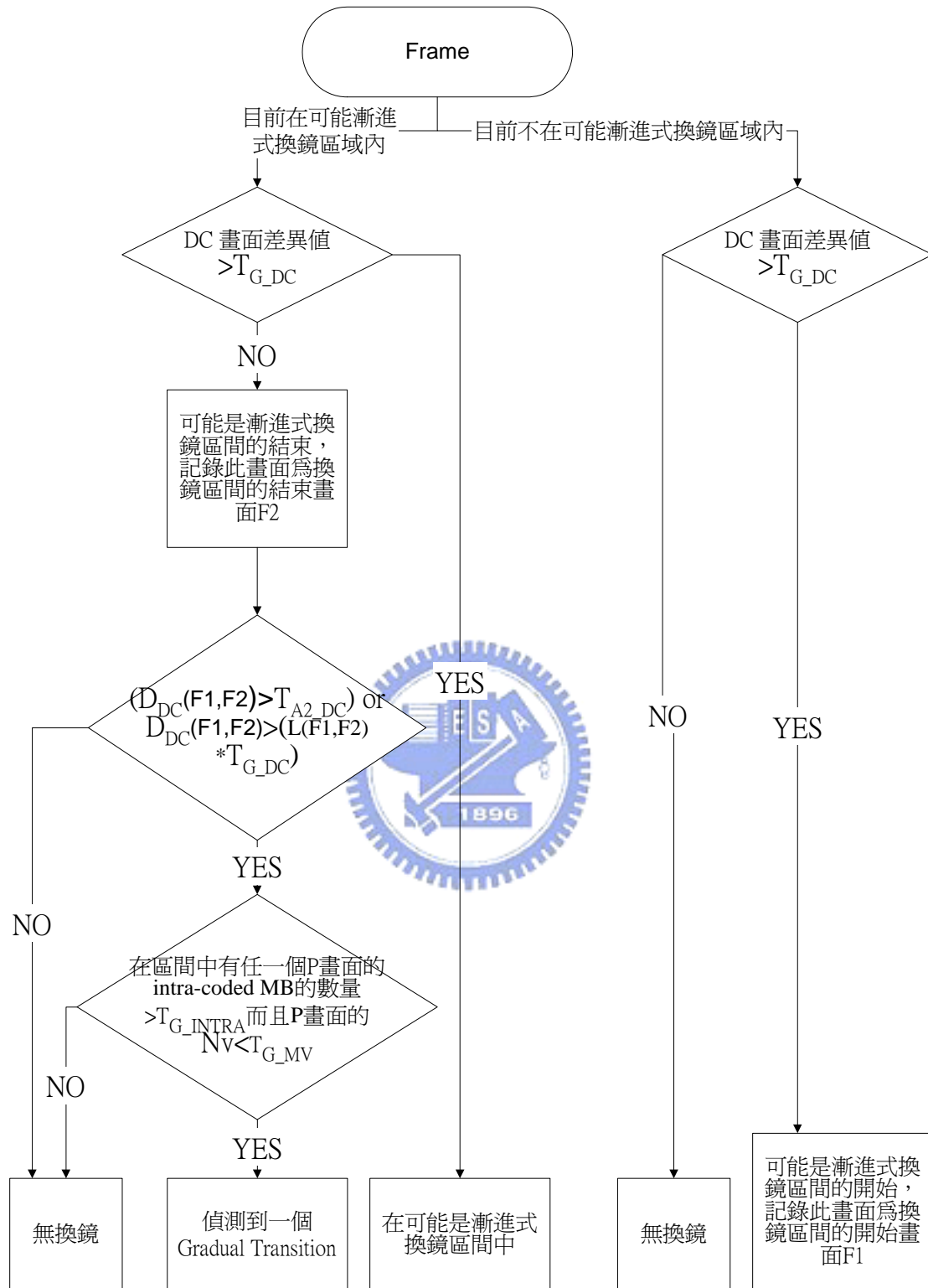


圖 4.12 偵測漸進式換鏡的流程圖

## 第五章 實驗結果及分析

我們在第四章提出了這個依照每段不同影片自動選擇畫面DC值差異的門檻值，在P畫面中以intra-coding型式編碼的MB個數的門檻值，和P畫面中的主要移動向量的數量的門檻值，利用這三種主要參考資訊來偵測瞬間式鏡頭變換及漸進式鏡頭變換，我們的實驗平台是Windows 2000 作業系統，架構在Intel Pentium Processor 1.6GHz，256 MB RAM 的個人電腦上，程式開發工具是Microsoft Visual C++ 6.0，而我們採用的MPEG-2 解碼程式是用網路上下載的mpegplyer的公用程式，它是一個用Microsoft Visual C++ 6.0 撰寫的MPEG-2 解碼程式及方便使用的應用系統，為了使編碼格式統一，我們將所有實驗的影片利用TMPGEnc V2.5 統一編碼成 352\*240 尺寸的畫面大小，每秒鐘播放 30 張的MPEG-2 影片，每段GOP的長度是 16，GOP格式為I B B P B B P B B P B B P B B P，我們實驗的影片來源有從網站<http://www.open-video.org/>的分享的視訊影片資料庫下載的影片及用電視拈取盒錄製的影片，七個影片是從網站<http://www.open-video.org/> 的分享的視訊影片資料庫中下載的影片，三個從電影及電視中錄下的影片，這些影片包含了卡通影片，偏向靜態的室內記錄影片，室外的原野記錄影片，新聞影片，綜藝節目片段的影片和電影中籃球比賽的片段，所用的實驗資料庫影片的代表畫面，列於圖 5.1。



(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)



(i)



(j)



- (a) Boat.mpg從[www.open-video.org](http://www.open-video.org) 下載的戶外原野記錄片
- (b) Cat.mpg從[www.open-video.org](http://www.open-video.org) 下載的室內記錄片
- (c) TV.mpg從[www.open-video.org](http://www.open-video.org) 下載的播報教學影片
- (d) Sport.mpg 從電影”鬼膽神偷”中節錄的籃球比賽影片
- (e) Cartoon.mpg從[www.open-video.org](http://www.open-video.org) 下載的卡通影片
- (f) Hcil.mpg從[www.open-video.org](http://www.open-video.org) 下載的電腦教學影片
- (g) QuietRev.mpg從[www.open-video.org](http://www.open-video.org) 下載的戶外施工的影片
- (h) Dog.mpg從[www.open-video.org](http://www.open-video.org) 下載的全家旅遊的影片
- (i) News1.mpg 從電視節錄的播報新聞的影片
- (j) Pet.mpg 從電視”寵物當家”節錄的一段綜藝節目影片

圖 5.1 所用的實驗資料庫影片的代表畫面

上 10 段影片的值測結果列於表 5.1，其中正確率是指偵測到的換鏡數目除以實際上的換鏡數目，正確率愈高就表示換鏡偵測的效果愈好，如果沒有將換鏡變化偵測出來就稱該畫面為漏判畫面，如果不是鏡頭變換的畫面卻被偵測為換鏡畫

面就稱該畫面為誤判畫面，換鏡偵測結果列於表 5.1 中，

	實際瞬間式 換鏡數目	偵測到的 瞬間式換鏡 數目 (正確率)	實際漸進式 換鏡數目	偵測到的 漸進式換鏡 數目 (正確率)	影片的 長度	誤判的畫 面數目
Boat	15	15 (100%)	9	9 (100%)	6152	3
Cat	12	12 (100%)	5	5 (100%)	2383	3
TV	14	14 (100%)	9	8 (88.8%)	1741	3
Sport	20	19 (95%)	1	1 (100%)	1075	1
Cartoon	12	12 (100%)	3	3 (100%)	1563	14
Hcil	37	37 (100%)	11	9 (81.8%)	7167	9
QuietRev	32	32 (100%)	56	51 (91.1%)	16443	11
Dog	10	10 (100%)	1	1 (100%)	1512	4
News1	13	13 (100%)	2	2 (100%)	1854	0
Pet	58	56 (96.5%)	7	7 (100%)	10984	4
總結	223	220 (98.7%)	104	96 (92.3%)	50874	54

表 5.1 瞬間式鏡頭變換偵測及漸進式鏡頭變換偵測的實驗結果

在我們的研究方法下的換鏡偵測結果，因為有使用自動設定門檻值的方法，所以在不同型態及不同主題的影片中的瞬間式鏡頭變換偵測和漸進式鏡頭變換偵測的正確率上都有不錯表現，但是在有 zooming， fast panning 或打閃光燈的狀況下有時還是會有誤判的現象發生，以下是我們針對漏判畫面的分析，圖 5.2



是一個瞬間式鏡頭漏判的狀況，在影片 Sport 中畫面 725 與畫面 726 中有一個瞬間式換鏡，但是由於畫面 726 是一個 I 畫面，無法參考畫面中的 MB 編碼型態及畫面中主要移動向量的個數，而且二張畫面的 DC 差異值並沒有大於判定換鏡的 DC 差異值的門檻值，所以會有漏判的現象。



(a)畫面 725



(a)畫面 726

圖 5.2 瞬間式鏡頭漏判的畫面

在漸近式鏡頭漏判的狀況，大致可分為二種型式，第一種是漸進式鏡頭的轉變過程很短，只有一個畫面的轉變過程，我們的方法有時會無法有效的處理，如圖 5.3 所示是影片 QuietRev 中漸進式鏡頭的轉變過程很短而造成漏判的畫面，第二種是在漸進式鏡頭的轉變過程之後，鏡頭或是畫面還是在持續的移動，導致畫面間的 DC 差異值還是大於漸進式變換鏡頭的 DC 門檻值，依我們的偵測方法會認定目前的畫面還是在漸進式鏡頭的轉變過程之中，而會造成漏判的情形。



(a)畫面 4801



(b)畫面 4802



(b)畫面 4803

圖 5.3 漸近式鏡頭漏判的畫面分析之一

關於在瞬間式鏡頭誤判的分析，在鏡頭移動太過劇烈，光線明暗變化很大的情況或是畫面有旋轉的狀況，會因為二個畫面裡的 DC 差異值太大而有可能會造成誤判的情況，如圖 5.4 所示影片 Hcil 的畫面 1852，畫面 1853 之間有打閃光燈而造成瞬間式鏡頭誤判的情況，而在圖 5.5 中則是在影片 Cartoon 中的畫面 791，畫面 792，有畫面旋轉狀況時而引起的誤判的情況。



(a)畫面 1852



(b)畫面 1853

圖 5.3 由光線造成的瞬間式鏡頭誤判的畫面



(a)畫面 791



(b)畫面 792

圖 5.4 由旋轉造成的瞬間式鏡頭誤判的畫面

關於在漸進式鏡頭誤判的分析，最常見的誤判的情況是有很大的物體從鏡頭

前面經過，在我們的偵測方法會因為物體經過又離開的現象而造成的漸進式鏡頭誤判，因為在真正漸進式換鏡時的畫面 DC 差異值的特徵和物體經過又離開的現象的畫面 DC 差異值的特徵十分相似，有時會造成誤判的情況如圖 5.5 所示，Dog 影片中的漸進式鏡頭誤判的情況。



圖 5.5 影片中誤判的漸進式鏡頭

## 第六章 結論與未來展望

本論文提出一個可依照不同影片型態和運動節奏而能自動化設定偵測換鏡的門檻值，直接讀取 MPEG-2 視訊影片中的壓縮資訊，偵測發生鏡頭變換的位置，我們先分析在 MPEG-2 視訊影片中發生瞬間式換鏡及漸進式換鏡時，MPEG-2 視訊影片中的壓縮資訊的特徵及參考資料，得到兩相鄰的 I (P) DC 畫面的 DC 差異值，P 畫面裡編碼方式為 intra-coding 的 MB 個數及 P 畫面裡具最主要位移向量的 MB 個數，這三個特徵可做為偵測換鏡的參考依據，並利用自動設定門檻值的方法來得到一段影片中兩兩相鄰的 I (P) DC 畫面的 DC 差異值的瞬間式換鏡及漸進式換鏡的門檻值，P 畫面裡編碼方式為 intra-coding 的 MB 個數的瞬間式換鏡及漸進式換鏡的門檻值，及 P 畫面裡具主要位移向量的 MB 個數的瞬間式換鏡及漸進式換鏡的門檻值，來進行瞬間式換鏡及漸進式換鏡的偵測。由實驗結果可證明，本論文提出的方法，針對偵測瞬間式鏡頭變換在各個不同型態和運動節奏的影片上都有不錯的正确性(約 98.7%)，針對偵測漸進式鏡頭變換在各個不同型態和運動節奏的影片上也都有不錯的正确性(約 92.3%)，但是因為 I 畫面比 P 畫面缺少 intra-coding 的 MB 個數及具有主要位移向量的 MB 個數的參考依據，所以在偵測瞬間式鏡頭變換時，會有極少數漏判及誤判的情況，而在偵測漸進式鏡頭變換時在激烈運動，快速移動鏡頭的情況下則容易有誤判的情況，應該可以再使用分段自動偵測門檻值的方法降低誤判的情況，並再加強我們的演算法。

# 參 考 文 獻

- [1] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-video Search for Object Appearance," *IFIP: Visual Database Systems II*, pp. 113-127, 1995.
- [2] H. J.Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic Partitioning of Full-Motion Video," *Multimedia Systems*, Vol.1, pp 10-28, 1993.
- [3] B. L. Yeo and B. Liu, "Rapid Scene Analysis on Compressed Video," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 5, No. 6, pp. 533-544, 1995.
- [4] S. C. Pei and Y. Z. Chou, "Efficient MPEG Compressed Video Analysis Using Macroblock Type Information," *IEEE Transaction on Multimedia*, Vol. 1, No. 4, pp. 321-333, 1999.
- [5] ISO/IEC IS 11172-2, MPEG-1 Video.
- [6] ISO/IEC IS 13818-2, MPEG-2 Video.
- [7] W. L. Lin, "Automatic Shot Change Detection In Digital Video," Master Thesis, CIS, NCTU, 2001.
- [8] J. Meng, Y. Juan, and S-F. Chang," Scene change detection in MPEG compressed Video sequence," *Digital Video Compression: Algorithms and Tehniques, SPIE*, Vol. 2419, pp. 14-25, 1995.
- [9] W. Wang and Z. Weng, " Scene Abrupt Change Detection," *Canadian Conference on Electrical and Computer Engineering*. Vol. 2 pp. 880-883, 2000
- [10]C. L.Huang, and B.Y. Liao, " A Robust Scene-Change Detection Method for Video Segmentaiton," *IEEE Trans. Circuit and System for Video Technology*, Vol11, No. 12

pp.1281-1288, 2001.

[11]B.T. Truong,C .Dorai, and S. Venkatesh ” Improve Fade and Dissolve Detection for Reliable Video Segmentation,” *International Conference on Image processing Vancouver, Canada* , pp.961-964, 2000.

[12]B. M. Mehtre, M. S. Kankanhalli, A. D. Narasimhalu and G.C. Man ” Color Matching for Image Retrieval,” *Pattern Recongnit .Lett.*, vol ,16, pp.325-331, 1994.

