

# 一種利用特寫鏡頭對數位電影資料進行自動化摘要合成之技術

陳信修  
中華大學資訊工程學系  
hhchen@chu.edu.tw

劉志俊  
中華大學資訊工程學系  
ccliu@chu.edu.tw

## 摘要

隨著生活品質提升，電影在人類精神生活中佔了極大的空間，數位化的電影資料，如 VCD、DVD 以及未來依 MPEG-4 標準所儲存的電影越來越多，所以要如何從影片中擷取出重要的視訊摘要，方便使用者快速的瀏覽與查詢影片的內容，是目前學者積極參與研究的題目。在電影預告片、海報及劇照等傳統影片摘要中存在一共通點，就是這些視訊摘要大量的使用特寫鏡頭來製作，以凸顯該片的特色及演員。在本文中，我們提出一種利用特寫鏡頭為基礎的視訊摘要自動擷取方法，可以取出電影中的特寫鏡頭，再進一步合成電影預告片、電影海報與劇照等電影摘要。

## 關鍵字

視訊摘要、特寫鏡頭、MPEG-4、視訊資料庫

## 1. 序論

在電影宣傳的眾多方法中，以電影海報及劇照較容易引起影迷的注意，它們不但可以表達該片的情境，也提供影迷瞭解該片內容的途徑。一般的電影海報及劇照，通常將電影的主要角色呈現出來，經過刻意的安排合成，以便製作出令人難忘的印象，進而達到吸引人的效果。

在電影資料方面，台灣的国家電影資料庫中，典藏了許多經典影片，但若要透過網際網路查詢視訊資料庫，只能得到相關資料的文字敘述，對於影片的內容無法深入的瞭解。若能將該視訊資料庫中典藏的視訊資料，分別自動化的擷取出視訊摘要，快速呈現此影片的大致內容，則我們就可以透過視訊摘要，確認該視訊

是否就是自己想要的，而不再侷限於文字的描述。

雖然視訊資料是內容最豐富的多媒體格式，但其龐大的資料量導致其並不適於應用在網際網路及無線傳輸上。針對這個問題，MPEG 組織定義了 MPEG-4 標準，以物件為基礎的方式來作視訊資料壓縮，並考慮網際網路與無線網路的傳輸特性特別設計，將龐大的資料量降低，以利於視訊在網際網路及無線傳輸上之應用。在 MPEG-4 視訊壓縮中，每一畫面都是由視訊物件平面(Video Object Plane, VOP)所構成，視訊物件平面則包含了物件形狀和材質的資料，我們可以在 MPEG-4 視訊中擷取出其中所含物件的資料，利用其中的形狀資訊來做出視訊摘要。

在本文第二節中，我們介紹目前電影摘要的基本概念，瞭解基本概念後，在第三節中介紹本文提出的方法，也就是特寫鏡頭的辨識。第四節中，介紹特寫鏡頭做視訊摘要的合成。在第五節中我們展示我們的實驗及結果。最後做本文的結論。

## 2. 相關研究

### 2.1 MPEG-4 物件編碼

MPEG-4 主要是針對視訊所訂定的標準，以物件為單位進行編碼，將物件取出後，減少壓縮的資料量，且在 MPEG-4 中視訊及音訊的壓縮也是不同的，視訊及音訊各自找到最佳的壓縮法，讓整個視訊的資料量達到最小。MPEG-4 視訊壓縮中，每一層由真正影像的實體：視訊物件平面所構成，視訊物件平面則包含了物件形狀(Shape)和材質(Textual)的資料(圖 1)，我們可以在視訊中查詢單一物件，提

---

\*本論文研究為國科會補助之研究成果，

供了在 MPEG-1 及 MPEG-2 標準中沒有的內涵式查詢。

MPEG-4 透過輪廓(Profiles)及層次(Levels)的概念來完成，雖然 MPEG-2 也有類似概念，但其採用較少，在 MPEG-4 中，輪廓是整個工具型態的定義，目前，MPEG 對於 MPEG-4 視訊物件、音訊物件及系統，都定義了輪廓，並期許層次及輪廓之進行相對一致。

MPEG-4 採用新一代的視訊壓縮方法，利用人眼的視覺特性，針對視訊傳輸的本質，以輪廓及紋路為基礎，支援視訊內容交互的功能。而實現內容交互功能的關鍵在於視訊物件(Video Object,VO)的編碼上，為此 MPEG-4 使用了視訊物件平面的觀念。根據人眼有感覺的一些特性，如形狀、運動、紋路等，將視訊中每一個場景視為不同之視訊物件平面所組成[1][2]。



圖 1.MPEG-4 物件形狀編碼

## 2.2 視訊摘要相關研究

在多媒體領域中，雖然有許多媒體標準及格式，但真正使用摘要的媒體並不多，視訊在摘要的應用較為成熟。從電影預告片及電視節目預告之普遍，就可知道摘要在視訊上的應用並非一二日而成的。雖然有許多專家專研於視訊摘要的研究，但視訊摘要還是沒有一個固定的標準。

有人研究以心理結構方式，改變舊有之框架結構的方式，讓視訊摘要能夠滿足人類心理的需求，作者將人類心理反應與視訊序列上的每一個軌跡作搭配，形成具有心理結構成分的視訊摘要。此方法改變了以往處理多媒體的觀點[3]。

在[4]論文中提到，目前研究視訊摘要有兩種方法，第一種方法是將一個場景中重要的物件形狀或場景取出，編輯成為一段視訊摘要；另

一種方法是將視訊中”重要的”或”有趣的”場景擷取出來，代表這個視訊的摘要。

## 3. 電影摘要系統的基本概念

### 3.1 現行電影摘要的種類

以目前電影現有視訊摘要種類來說，可區分為電影預告片、電影海報(劇照)、劇情簡介及電影寫真集等。

#### 3.1.1 電影預告片

一般的電影預告片(previews)均為短片，通常是將電影的精華片段，經過刻意安排剪輯而成，以便製造出令人難忘的印象，而達到吸引人的效果。

#### 3.1.2 電影海報、劇照



(a)特寫版面海報



(b)演員群版面海報



(c)摘要版面海報



(d)綜合版面海報

圖 2. 特寫版面、演員群版面

摘要版面、綜合版面海報

每一部即將要上映的影片，都有其宣傳照片，宣傳照片有分海報(posters)及劇照(stage photos)，這些都是在電影攝製期間或前後，為了打廣告、作公關或做電影院櫥窗展示而拍攝的照片。其內容以特寫鏡頭最多，以顯示片中角色之重要性。而電影海報的製作，有分特寫

版面、演員群版面、摘要版面及綜合版面；特寫版面是將片中的主要角色搭配電影背景，獨自完成一副宣傳海報(圖 2(a))。演員群版面是將該片中的所有演員群角色，利用合成方式，完成一副電影海報(圖 2(b))。摘要版面是在該部電影所有劇照中，選出要呈獻給影迷的照片，將精華劇照結合成一副海報，希望藉由海報吸引更多觀眾(圖 2(c))。綜合版面屬於就沒規則性的海報，它可以讓設計者自由排版，不論特寫、主角群或摘要，它都可以應用在其中，當然還包含了場景等等，發揮的空間非常大，但它並非主力強打的海報(圖 2(d))。

### 3.1.3 劇情簡介

將該部電影作劇情簡介(synopsises)、劇情解說及敘述，通常以敘述者的角度直接用語言或文字來介紹影片內容、交代劇情或發表議論的一種方式。目前國家電影資料庫就是以劇情描述為主。

### 3.1.4 電影寫真集

每一部電影在拍攝過程時，都會將每一個鏡頭、排練過程、使用道具及 NG 鏡頭等作記錄，整理成一部電影劇照本或電影寫真集 [7]，配合該部電影的上映，讓喜愛這部電影的觀眾，能將這些喜愛的劇照帶回家收藏。寫真照片也是大部分取材於特寫鏡頭。

## 3.2 視訊摘要系統架構

我們所提出的電影摘要系統之架構如圖 3 所示。對於一部 MPEG-4 格式之數位電影，首先經過鏡頭偵測模組(shot change detection)切割成一連串之電影鏡頭。對於每一電影鏡頭，透過形狀資訊擷取模組(shape extraction)，取出每一個物件之形狀。利用這些資訊，特寫鏡頭偵測模組(close-up shot detection)得以自動判斷此鏡頭是否為特寫鏡頭。最後，我們利用視訊摘要合成模組(video summary synthesis)，對特寫鏡頭依照角色進行分類，找出各個角色之特寫鏡頭總數，來衡量其重要性(戲份)，訂定劇照及海報的輸出樣版，將特寫鏡頭自動的合成摘要劇照、摘要海報與電影預告片。

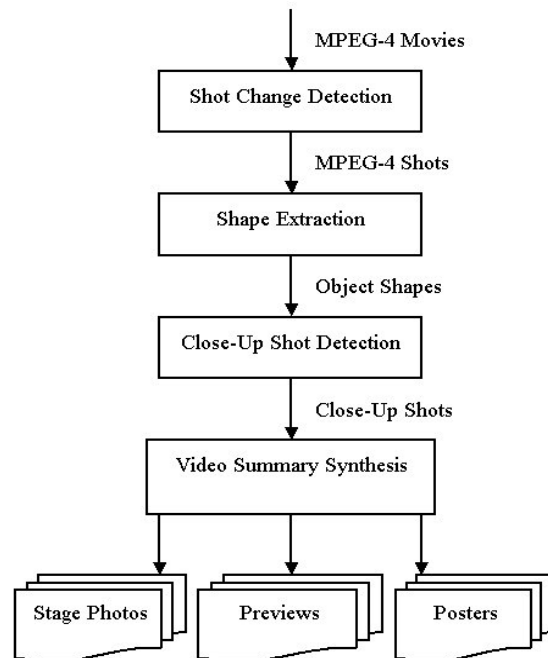


圖 3. 電影視訊摘要系統架構圖

## 4. 特寫鏡頭的辨識

### 4.1 電影鏡頭的種類

在本論文第二節中提到了許多電影摘要的宣傳方式，其中的預告片、劇照及海報，都大量應用特寫鏡頭，所以特寫鏡頭的自動辨識是本系統最核心的技術。

表 1. 電影術語對照表

Close-up	特寫	經由近距離所拍攝的人或物的放大或細節描繪的鏡頭，如拍攝對象是人，則指肩部以上的拍攝範圍。
Close shot	近景	介於中景與特寫之間的鏡頭。
Medium shot	中景	在視界與視覺角度方面介於特寫鏡頭與遠景鏡頭之間的一種鏡頭。以中景鏡頭來表示一個人時，最典型的視覺範圍是從該人的膝蓋以上來拍攝。
Medium close shot	中近景	表現文物腰部以上的活動情形，介於中景和特寫間的一種取景方式。中近景在拍攝劇情片人物時經常使用。
Full shot	全景	被攝體佔滿整個螢幕的鏡頭。若被攝體是一個人，則他或她的身體會全部容納在鏡頭中。

鏡頭是經由攝影機所拍攝到的人或物之放大或細節描繪，如拍攝對象是人，則依據拍攝到的範圍來區分不同的鏡頭。以臥虎藏龍 DVD 視訊為例，我們在固定間格所擷取出來的場景裡，將每一個場景內所包含的物件取出，使物件一一獨立出來，做為我們對鏡頭定義之實

例，參考電影鏡頭相關術語表中的定義[表 1]，分別判斷其所屬鏡頭如下(圖 4)。



(a)特寫鏡頭



(b)近景鏡頭



(c)中近景鏡頭



(d)中景鏡頭



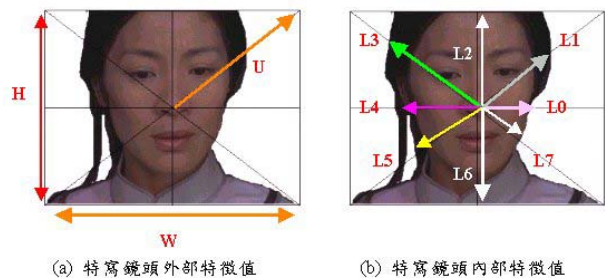
(e)全景鏡頭

圖 4. 鏡頭的種類

## 4.2 特寫鏡頭的偵測方式

特寫鏡頭是經由近距離所拍攝到的人或物的放大或細節描繪的鏡頭，若拍攝的對象是人物，則拍攝的範圍從這個人的肩部以上，包含整個頭部，也就是說在一個場景中，人物的面積佔據了整個場景的大部分空間。在特寫鏡頭的偵測方法中，我們依據電影辭典[6]對電影名詞的詮釋，配合本論文中提出的特徵值擷取方法，來作特寫鏡頭的判斷。首先我們先為特寫鏡頭下定義，定義如下：

1. 首先，我們將每個從場景中擷取出來的特寫物件，利用矩形將整個物件包圍起來，再以對角線及各邊的中心線，將物件區分成八個區域，產生”米”字形的效果(圖 5)。



(a) 特寫鏡頭外部特徵值

(b) 特寫鏡頭內部特徵值

圖 5. 特寫鏡頭的特徵值擷取圖

2. 在圖 5(a)中，我們先定義特寫鏡頭的外部變數，矩形的寬為 W，高為 H，參考線為 U。其中的參考線 U 為對角線的一半，用來作內部定義線段的基準。
3. 在圖 5(b)中，我們定義特寫鏡頭的內部變數，面對特寫鏡頭的右臉，我們定義它為線段 L0，右眉為線段 L1，頭頂為線段 L2，左眉為線段 L3，左臉為線段 L4，左頰為線段 L5，胸口為線段 L6，右頰為線段 L7，八條內部線段。線段 L0 為中心點至該物件邊緣之長度，在圖 5 中，就是從鼻子到右臉邊緣的長度；其他線段依此類推，唯有圖 5(b)中的線段 L5 測量方式不同，其測量方式是從鼻子至第一次接觸到物件的邊緣止，也就是圖 5(b)中其左下方的一束頭髮，因為與物件主體無緊貼在一起，所以我們就不測量它。
4. 我們將定義好的線段，套用至特寫鏡頭之物件上。首先測量該矩形的寬、高及參考線之長度，將其記錄於表格中；其次，我們再測量線段 L0 至 線段 L7 的長度，並分別記錄下來。
5. 在測量出各線段的原始數據後，我們就將它們正規化，依據每一個物件的參考線段為基礎，分別將各線段作正規化的動作， $L_I$  為內部線段值，U 為參考線段值， $L'_I$  為正規化後各線段的值，公式如下：
$$L_I / U = L'_I, I=0.1.2...7 \quad (1)$$
6. 我們以 20 組特寫鏡頭物件，經過公式(1)計算後得到正規化值, 如表 2 所示。

表 2. 特寫鏡頭正規值表

圖號	W(cm)	H(cm)	L'0	L'1	L'2	L'3	L'4	L'5	L'6	L'7
特寫 01	8	6.3	0.3529	0.6078	0.6176	0.7059	0.5098	0.4902	0.6176	0.3529
特寫 02	7.38	6.35	0.3093	0.5567	0.6546	0.3814	0.2371	0.2268	0.6546	0.3093
特寫 03	6.8	6.38	0.3441	0.8387	0.686	0.3656	0.2903	0.3419	0.686	0.3441
特寫 04	9.4	6.3	0.292	0.3717	0.5575	0.3982	0.2566	1	0.5575	1
特寫 05	8.4	6.3	0.3048	0.5714	0.6	0.6857	0.4095	0.4667	0.6	0.419
特寫 06	9.5	6.2	0.2632	0.4386	0.5439	0.6491	0.3947	0.4386	0.5439	0.2895
特寫 07	6.8	6.1	0.2637	0.5714	0.6703	0.4374	0.3297	0.3516	0.6703	0.1978
特寫 08	8.4	6.38	0.5283	0.5094	0.6019	0.5849	0.4434	0.5283	0.6019	1
特寫 09	6.99	6.35	0.3617	0.9362	0.6755	0.6596	0.3085	0.3191	0.6755	0.4681
特寫 10	7.5	6.3	0.3452	0.5381	0.6396	0.3553	0.264	0.2437	0.6396	0.3046
特寫 11	6.4	6.22	0.5393	0.6629	0.6989	0.7079	0.4719	0.6517	0.6989	0.7865
特寫 12	7.25	6.4	0.4352	0.6425	0.6632	0.5078	0.3731	0.3109	0.6632	0.4456
特寫 13	8	6.2	0.495	0.7723	0.6139	0.5842	0.301	0.4257	0.6139	0.4653
特寫 14	12.45	6.25	0.3022	0.5468	0.4496	0.4748	0.3871	1	0.4496	1
特寫 15	7.6	6.2	0.5612	0.6531	0.6327	0.4694	0.4082	1	0.6327	1
特寫 16	6.52	4.4	0.3823	0.4304	0.557	0.4051	0.4076	0.4937	0.557	0.4684
特寫 17	13	6.2	0.4306	0.4375	0.4306	0.5347	0.4847	1	0.4306	1
特寫 18	7.9	6.2	0.3149	0.4356	0.6139	0.4158	0.2733	0.2772	0.6139	0.2574
特寫 19	7.6	6.38	0.4242	0.6061	0.6444	0.4404	0.303	0.303	0.6444	0.4404
特寫 20	8	6.19	0.335	0.4532	0.6099	0.532	0.3941	0.4335	0.6099	0.333

經過以上特寫鏡頭的定義後，我們得到了每一個特寫鏡頭的數值，這些數值可以用作我們對特寫鏡頭的判斷。本文中，我們對特寫鏡頭的判斷提出了兩階段的判斷方法，其整體判斷結構說明如下：

**第一階段：**我們利用物件之頭胸位置及臉部對稱性，來過濾從場景中取出之物件，將符合條件的物件交由第二階段作特寫鏡頭的判斷。以下我們就兩種條件作個說明：

1. 頭胸位置：因特寫鏡頭是從一個人的肩部以上拍攝入鏡，相當於頭部幾乎佔了大部分場景的空間，所以我們依據該物件的線段 L'2 及線段 L'6，來過濾物件在場景中是否有特寫鏡頭的條件，其公式如下：

$$\begin{aligned} (ABS(L'2-0.6)) &\leq 0.2 \\ (ABS(L'6-0.6)) &\leq 0.2 \end{aligned} \quad (2)$$

其中 L'2 及 L'6 的計算方式如公式(1)，我們將特寫鏡頭正規化後，將得到之線段值

驗證後，我們將它列為特寫鏡頭過濾的第一種方法。以圖 6 的例圖來看，頭胸位置法會將(a)、(c)及(f)過濾出來。

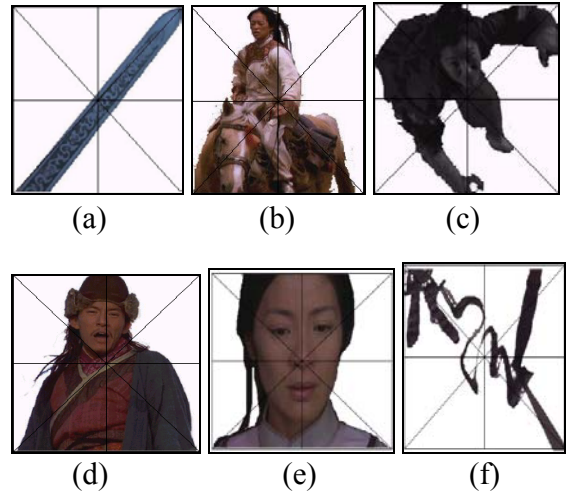


圖 6. 鏡頭判斷例圖

2. 臉部對稱性：以特寫鏡頭拍攝到的人物臉部，若以”米”字去切割畫面，通常可以將人的臉部一分為二，所以我們依據該物件的線段 L'0 及線段 L'4，來判斷物件在場景中是否符合臉部對稱的條件，其判斷公式如下：

$$\begin{aligned} (ABS(L'0-0.4)) &\leq 0.2 \\ (ABS(L'4-0.4)) &\leq 0.2 \end{aligned} \quad (3)$$

若以此方法單獨來判斷，在圖 6 中，我們會將(a)、(b)、(c)及(f)濾出。若經過頭胸位置法過濾後，會再將(b)圖過濾出來，留(d)及(e)圖繼續做第二階段的判斷。

**第二階段：**

1. 左右對稱性：在經過第一階段過濾後，我們可以知道這個被拍攝的人，其頭部大約頂到鏡頭上方，而其胸部剛好在鏡頭的下方，所以我們假設”米”字的中心點大約在人的臉部範圍內，我們透過這個假設將人臉一分為二，也就是它的中心點剛好位於人臉的中心，所以我們依據上述定論，分別去判斷特寫鏡頭的左右對稱性，驗證特

寫鏡頭是否有左右對稱的性質，驗證公式如下：

$$\begin{aligned} &(\text{ABS}(\text{L4}/\text{L0}) - 0.8) \leq 0.2 \text{ if } \text{L4} < \text{L0} \text{ else } (\text{ABS}(\text{L0}/\text{L4}) - 0.8) \leq 0.2 \\ &(\text{ABS}(\text{L3}/\text{L1}) - 0.8) \leq 0.2 \text{ if } \text{L3} < \text{L1} \text{ else } (\text{ABS}(\text{L1}/\text{L3}) - 0.8) \leq 0.2 \quad (4) \\ &(\text{ABS}(\text{L5}/\text{L7}) - 0.8) \leq 0.2 \text{ if } \text{L5} < \text{L7} \text{ else } (\text{ABS}(\text{L7}/\text{L5}) - 0.8) \leq 0.2 \end{aligned}$$

在公式(4)中，0.8為左右對稱性質的基準點，而容許值為正負0.2，也就是我們允許其特徵值在 $0.8 \pm 0.2$ 之間，凡落於此範圍內的值，我們就判斷它有對稱性質。其他線段之驗證法如公式(4)，將特寫鏡頭經過公式(4)判斷後，我們發現大部分的特寫鏡頭都是左右對稱的，只有少部分不對稱的，所以我們將它列入特寫鏡頭判斷方法一。

2. 肩、臉寬度比例：經過了以上兩種判斷方法後，我們針對特寫鏡頭中，探討人的肩寬與臉寬的比例關係，透過這個方法可以判斷出是否為人物的特寫鏡頭。我們利用先前內部定義的線段L0、L4及外部定義之寬W，作為判斷之依據，其判斷公式如下：

$$(\text{ABS}[(\text{L0} + \text{L4}) / \text{W}] - 0.5) \leq 0.2$$

透過公式(5)的判斷後，我們得到了肩、臉寬度比例大約在0.3~0.7之間，而其平均值在0.5比例的地方，所以我們依據其分佈的狀態，作為判斷特寫鏡頭的第二個方法。

經過以上判斷方法後，我們取得了特寫鏡頭的共同特徵點，發現其中有幾個樣本的判斷結果不同，我們知道特寫鏡頭主要拍攝範圍在肩部以上至頭頂，若被拍攝者長髮披肩、頭帶髮飾、帶耳飾或帽子等，都有可能影響判斷的結果，但以整體的判斷結果來說，判斷成功率還算很高的。

## 5. 視訊摘要的合成

### 5.1 劇照的合成

在經過特寫鏡頭辨識與擷取後，我們依據每一個特寫鏡頭出現的頻率、位置及間距，衡量其重要性與代表性，取出代表鏡頭、代表畫

面與代表角色[表3]，利用系統中的視訊摘要合成系統，並將這些取出的特寫鏡頭劇照合成起來(圖7)。



圖7. 按各主角特寫鏡頭，系統挑選的代表性特寫鏡頭畫面(選自電影“臥虎藏龍”)

表3. 主角角色出現特寫鏡頭比例表

主角姓名	出現特寫鏡頭比例
李慕白	49.4%
玉嬌龍	47.0%
余秀蓮	41.1%
羅小虎	8.8%
碧眼狐狸	2.6%

### 5.2 電影海報的合成

代表角色由特寫鏡頭物件擷取出後，利用視訊摘要合成系統，定義出海報的樣版(圖8)，讓視訊摘要合成系統產生我們所要的海報(圖9)。在本文中，我們的樣版是採用對數螺線(Logarithmic spiral)，由笛卡兒發明。它是一種半徑以幾何級數(2,4,6,8,16...)增長的螺線，又稱為等角螺線(Equiangular spiral)[8]。我們依照代表角色的重要性，由樣版中的劇照一、

劇照二……依序由大而小排列，而主要演員姓名則採列代表角色之前四位。背景可以由場景中任選其一，並做模糊化的動作。

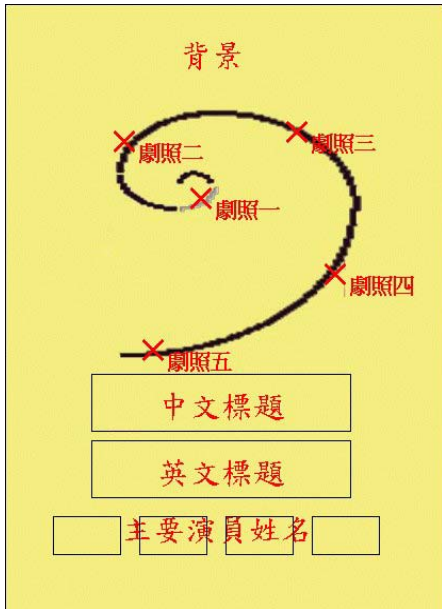


圖 8. 電影海報的合成樣版

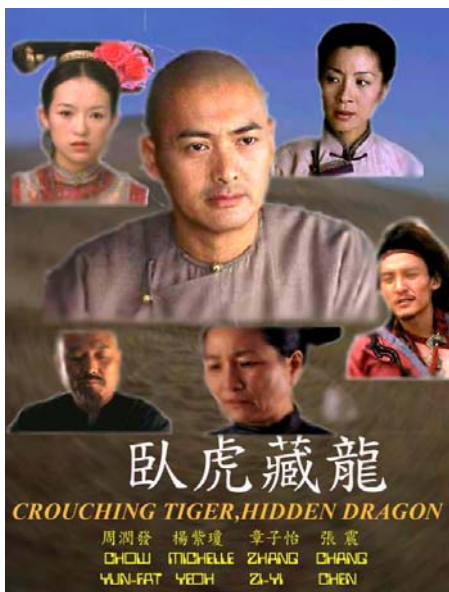


圖 9. 系統依代表角色與海報樣版所合成之”臥虎藏龍”電影海報

## 6. 實驗

### 6.1 實驗環境

在實驗過程中，我們的實驗樣本區分兩部分，第一部份我們以 MPEG-4 標準範例做為實驗樣本。第二部分，我們我們使用了四部電影視訊，分別為臥虎藏龍、駭客任務、鐵達尼號及神鬼戰士，臥虎藏龍是武俠片的代表，駭客任務是動畫特效片的代表，鐵達尼號是文藝片的代表，而神鬼戰士是戰爭格鬥片的代表，我們依據其不同的製片風格，來驗證我們所提出的方法。

### 6.2 實驗結果與分析

而鏡頭偵測方面，我們統計偵測出的總鏡頭數及特寫鏡頭總數，發現其比例大約每 100 個場景中，只有 15 個場景中有特寫鏡頭出現，如圖 10 所示：

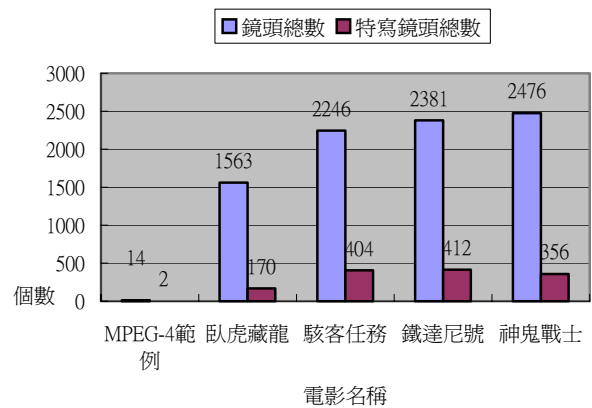


圖 10. 鏡頭總數與特寫鏡頭總數比例圖

在系統的偵測後，我們得到了我們實驗的 Precision 及 Recall，如表 4。

表 4. MPEG-4 範例、電影特寫鏡頭偵測數據 (選自電影”臥虎藏龍、鐵達尼號、神鬼戰士、駭客任務”)

影片名稱	鏡頭總數	特寫鏡頭數	Precision	Recall
MPEG-4	14	2	100%	100%
臥虎藏龍	228	36	77.8 %	62.2 %
鐵達尼號	372	37	78.3 %	60.4 %
神鬼戰士	290	40	80.0 %	59.2 %
駭客任務	259	32	78.1 %	62.5 %

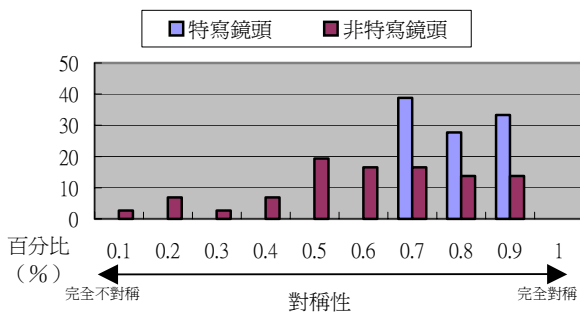


圖 11. 特寫鏡頭與非特寫鏡頭  
左右對稱比較圖

在統計特寫與非特寫鏡頭左右對稱後，我們由圖 11 中得知，特寫鏡頭其對稱性在 0.7 至 0.9 之間，也就是說我們在要求特寫鏡頭時，對稱性是非常重要的條件，比起非特寫鏡頭的對稱性要求更高。因此，藉由左右對稱性的性質，我們可以把 50% 以上的非特寫鏡頭偵測出來，減少我們處理的時間。

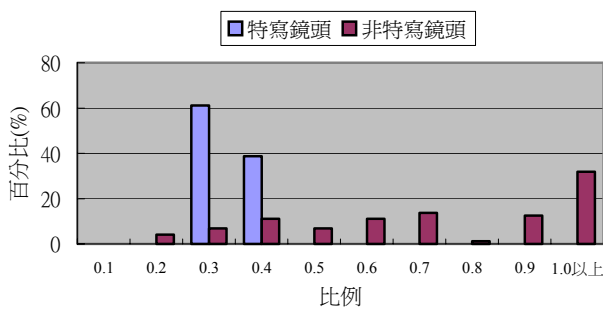


圖 12. 特寫鏡頭與非特寫鏡頭  
肩臉寬度比較圖

在本論文所提的方法中，我們主要在擷取出人物的特寫鏡頭，所以肩寬與臉寬比例是我們辨識人物特寫鏡頭的關鍵，由我們實驗結果(圖 12)得知，人的臉寬大約是肩寬的三分之一左右，所以藉由這個原則所訂定的條件判斷，可以將物件中是特寫鏡頭的場景擷取出來，得到本文所要的結果。

## 7. 結論

在本文所提出的特寫鏡頭擷取方法，透過實驗的數據顯示，此方法可以有效率的將特寫物件取出。不僅如此，我們所提出的方法，也可以利用在不同物件上，找出其共通點及特徵值，即可將所需的物件找出。最後，我們希望此方法能充分的使用於視訊摘要上，提升視訊摘要的品質及製作。

## 8. 參考文獻

- [1] "ISO/IEC 14496-2 Information Technology-Generic Coding of Audio-Visual Objects." *ISO/IEC* 1998.
- [2] Erol, B. and F. Kossentini. "Automatic Key Video Object Plane Selection Using the Shape Information in the MPEG-4 Compressed," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 2, No. 2, pp. 129-138, 2000.
- [3] Moriyama, T. and M. Sakauchi, "Video Summarization Based on the Psychological Content in the Track Structure." *ACM Multimedia Workshop*, pp. 191-194, 2000.
- [4] Oh, J. H. and K.A. Hua, "An Efficient Technique for Summarizing Videos Using Visual Contents," *IEEE International Conference on Multimedia and Expo*, Vol. 2, pp. 1167-1170, 2000.
- [5] Kim, C. and J. N. Hwang, "An Integrated Scheme for Object-based Video Abstraction." *ACM Multimedia*, pp. 303-311, 2000.
- [6] "電影辭典" 國家電影資料館 1997.
- [7] "臥虎藏龍寫真集" 東販出版社, 2000.
- [8] Howard Eves, 歐陽絳譯 "數學史概論" 曉園出版社 1993