

AC-3 環場音效與電影劇情關聯之資料探勘模型

吳智偉 劉志俊

中華大學資訊工程學系

ccliu@chu.edu.tw

摘要

DVD 為目前最主要的電影儲存媒介之一。DVD 與以往錄影帶或 VCD 相比，除了視訊的品質大幅提升之外，其音訊資料亦由立體聲改進為多聲道環場音效，而 Dolby AC-3(簡稱 AC-3) 為最主要之環場音效格式。

本篇論文提出了一種針對 AC-3 電影音效與電影劇情關聯之資料探勘模型。我們透過對 AC-3 電影音效進行內涵分析，由 AC-3 解碼過程中擷取出其特徵係數，再根據 AC-3 的環繞音場特性以及人類聽覺感知性質，計算出環場音量字串，再透過一組樣板規則，找到隱藏在環境音量字串中的音效事件，並利用關聯法則分析演算法，進而對電影劇情與音效事件進行關聯分析，找到其間的關聯法則。最後，透過對電影劇情的關聯法則，便可以對未知電影進行自動摘要與劇情自動分析。

1. 序論

有鑒於電子科技與網際網路的蓬勃發展，數位化的音樂、影像以及視訊漸漸的成為關注的焦點。近年來，人們對於多媒體品質的要求越來越嚴格，因此，有許多更先進地多媒體數位化技術的研究成果之產生，也促使了電影工業技術的快速成長。

為了能夠更有效率的運用各種多媒體的資料，多媒體資料的內涵式(Content-based)分析相關技術為目前多媒體與資料庫領域的重要研究項目之一。因為唯有透過對多媒體內涵的了解，才能夠針對各種媒體資料特殊的隱含意義做出適當的處理。電影是屬於視訊與音訊結合(audio-visual)的一種資訊。換言之，要針對電影資訊做分析，勢必會使用到視訊或是音訊資料的分析技術。

在視訊內涵分析方面，已有許多的研究成果。例如尋找視訊中的關鍵框架(key frame)，針對目前的視訊做摘要的動作[15]；也有學者是將重要的或有趣的視訊片段做為視訊的摘要[7]。

在音訊方面，也有許多內涵式的分析方法陸續被提出。在 [14] 一文中，Wold 等人針對了內涵式的查詢，提出了利用音量(Loudness)、音高(Pitch)、明亮度(Brightness)以及頻寬(Bandwidth)等音訊的內涵資料來進行分析，這跟以往的分析方式是不同的，對於音訊的分類也是一大突破。由於各種不同的音效擁有不同的特徵，利用內涵式的概念，便可以將不同類型的音效分為許多不同的群組，進而將其分類；同樣的，在 [5] 一文中，Lu 等人也提出了 Silent ratio、Spectral centroid、Harmonicity 以及 Pitch 等四種特徵來針對音訊資料進行分類。相關的理論也可應用於電影中，在 [16] 一文中，先針對音效的做出有效的分類，再利用各類音效在電影中的分布狀況進行電影類型的分類。在 [4] 一文中，Santo 等人便利用類神經網路所建置的分類系統來進行電影音效分類。

但是，在上述所提及的視訊或音訊的分析，都是獨立運作的。因此，有許多的研究便結合了視訊以及音訊的分析方法，亦及所謂的多重模式分析(multi-modal analysis) [8][10][11]。目的是為了讓分析的結果更為精確。在 [10] 一文中，便利用音訊資料來判斷目前的場景是否為對話的鏡頭；在 [11] 一文中，則是提出一種利用音訊資料來辨識對話鏡頭中對話人的身分，進而找出整個視訊片段中某個人的所有對話鏡頭的分析方法；而在 [8] 一文中，更是利用音樂與語音來揣摩製片者的心境，以建立一部電影的劇情結構。

然而，電影的內涵是相當豐富且複雜的。與一般的音訊分析不同的是，電影的音效是由數種不同類型的音效所共同組成，甚至音效與音效之間是互相重疊的，這在音訊的分析上是一個極為棘手的工作。電影能夠呈現給人們一個是十分趨於真實的感官世界。它是目前最生動的多媒體資訊表達方式，透過電影所產生的視訊與音訊訊號，使用者能夠更精確的了解它所要呈現的意境。

Dolby AC-3 是美國高解析度電視(HDTV) 標準所採用的音訊壓縮模式，也是目前 DVD 影片中所廣泛採用的標準模式[1][12]。AC-3 雖然是目前最重要的音訊格式之一，但以往對於 AC-3 的相關研究大部份集中在編碼器與解碼器的較佳化相關探討[12][6]，並沒有研究是針對環場音效的特性進行內涵式分析。本篇論文提出了一種針對 AC-3 電影音效與電影劇情關聯之資料探勘模型。我們透過對 AC-3 電影音效進行內涵分析，由 AC-3 解碼過程中擷取出其特徵係數，再根據 AC-3 的環繞音場特性以及人類聽覺感知性質，計算出環場音量字串，再透過一組樣板規則，找到隱藏在環境音量字串中的音效事件，並利用關聯法則分析演算法，進而對電影劇情與音效事件進行關聯分析，找到其間的關聯法則。最後，透過對電影劇情的關聯法則，便可以對未知電影進行自動摘要與劇情自動分析。

本篇論文結構說明如下：第 2 節簡介 AC-3 環場音效；第 3 節介紹本篇論文電影劇情探勘系統的整體架構；第 4 節說明本篇論文所採用的音訊特徵值；第 5 節會進一步提出整個音訊資料探勘系統的模型；最後說明我們未來工作與研究方向。

2. AC-3 環場音效簡介

在電影工業發展初期，由於當時電影工業面臨電視業者的強力衝擊，再加上磁性儲存媒介的成本影響，不得不發展出更新穎的技術以吸引消費者的青睞。因此，多聲道的技術才被應用在電影工業上。近年來，由於家庭環繞劇院的普及，也促使更多新的技術被提出。使得原本只儲存立體音的數位音樂，也配合電影的播放而製造出立體空間的效果。

多聲道音訊資料發展出許多的多聲道音訊編碼技術，如 AC-3、MPEG AAC [13]以及 DTS。本篇論文所論述的多聲道音訊分析則是採用 AC-3 的編碼技術。AC-3 是目前最受歡迎的多聲道碼技術，雖然 DTS 是較新的產品，但是目前市面上販售的 DVD 尚未完全採用這個新的技術，因此，我們選擇了使用最廣泛的 AC-3 來進行環境音效的內涵分析。

環境音效的目的是為了產生數個數位化的聲道資訊，來還原錄製時來自四面八方之原始的音訊資料(類比的電子脈衝 PCM)。而數位音樂的壓

縮技術則是降低這些數位化資訊的資料量。AC-3 數位壓縮技術的目的也是如此，但與其他壓縮技術不同的地方是，它支援了八種不同的聲道結構。如圖 1 所示，除了傳統式的立體音(Stereo) 與單聲道(Mono) 之外，其支援的環場音效聲道有左前(Front Left)、中置(Center)、右前(Front Right)、左環繞(Surround Left)、右環繞(Surround Right) 與重低音(LFE)等六個分離的聲道，其中，LFE 表示的是極低頻的音訊資料。

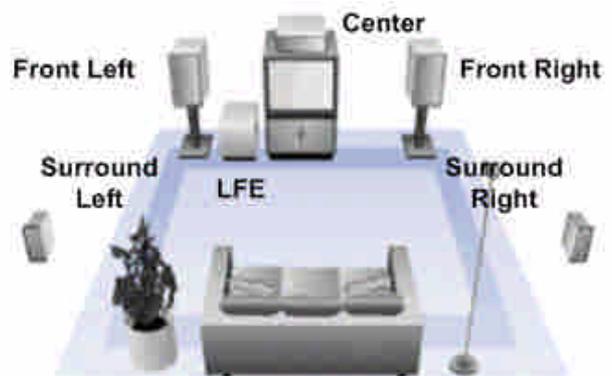


圖 1 5.1 聲道示意圖

3. 電影劇情自動探勘系統整體架構

3.1. AC-3 環場音效與電影劇情關聯分析系統

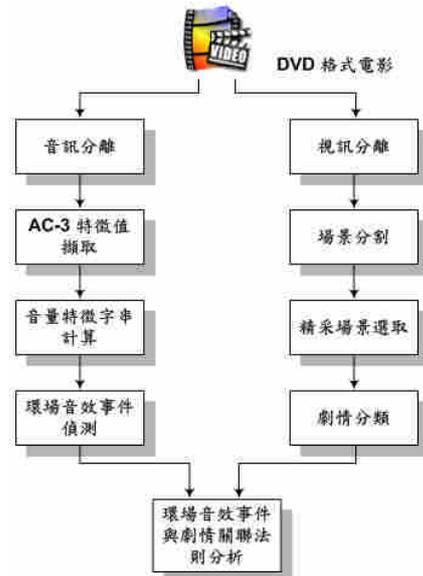


圖 2 AC-3 環場音效與電影劇情關聯分析系統架構圖

圖 2 是 AC-3 環場音效與電影劇情關聯分析系統架構圖。我們先將原始的 DVD 影片中的音訊與視訊兩個部份各自分離出來，再以不同的程序各

自獨立進行處理。首先，在視訊部分，我們以人工的方式尋找精采的電影劇情，並進一步將劇情分類。在音訊部分，將每個劇情所對應的音訊特徵值取出，並計算出音量特徵字串。再利用一組預先定義的字串樣板，來偵測隱藏在音量特徵字串中的環場音效事件。

資料探勘的目的，是在大量的資料中找出某些隱含的資訊或規則。因此，我們將同一類型的劇情所對應的所有環場音效事件進行關聯法則分析，試圖找出某一類型的劇情中共有的音效事件模型。此音效事件模型可應用在未知影片的劇情自動分析以及電影摘要的自動建立。

3.2. 環場音效之內涵

當我們在欣賞一部電影時，往往因為它的劇情，心情也會隨之起伏不定，例如驚悚的、感性的、緊張的劇情。在這些精采的片段中，往往都會配合許多的音效，例如戰爭片的爆破效果、喜劇片的笑聲、文藝片的抒情音樂等等，透過這些音效，讓觀眾的心境也不自覺的投入這部影片。

當導演想要表達出不同的情境時，他會利用適合的音訊來進行搭配，使得整體的效果更為彰顯。換言之，我們可以利用對音訊的分析，進而了解導演的內心世界。如圖 3 所示，導演會將所要詮釋的情境傳達給攝影師以及剪輯的人員，也因為導演情境的不同，視訊的內容也會隨之改變。導演是電影的靈魂人物，換句話說，電影是導演內心情境的一種衍生物。透過電影，可以展現出導演內心的構思與想法。因此，我們認為電影視訊摘要必須以導演的中心思想為出發點，使得電影摘要能夠有效且快速的表達出整部電影的菁華內容。上述電影理論為本篇論文尋找 AC-3 環場音效與電影劇情關聯之基礎理論根源。

3.3. 環場音效與劇情探勘

AC-3 音效最大的特色在於它所支援的環場模式，能夠讓電影更深切的表達出所要呈現的訊息。透過環場音效的呈現，讓觀眾有一種身歷其境的感官效果。也因為如此，電影製作人員往往會在他們認為重要的或精采的電影片段中加入環場音效，藉由環場音效的效果，讓觀眾更能夠加深對整部電影的印象。

因此，我們針對 AC-3 的環場音效特徵來進行分析。如上文所述。具有環場特徵的電影片段，往往是製作人員想要呈現的精采片段，因此，透

過環場音效的分析，不同類型或劇情的電影片段中，會有不同的環場音效表達方式，倘若能夠對不同的環場音效做分析並加以分類，便能夠對整部電影的劇情進行分類或預測，使我們能夠取得整部電影的菁華片段，這也是視訊摘要的本質。

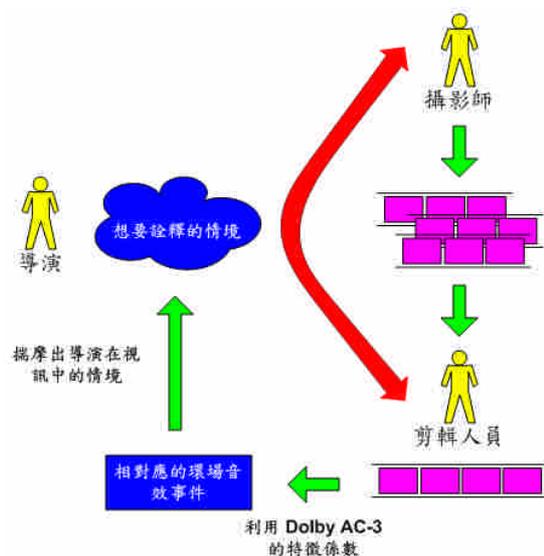


圖 3 導演情境與音訊的關聯圖

4. 環場音效特徵值擷取

在本文中，我們採用環場音效的特徵值，也就是藉由分析 AC-3 中各個聲道的資訊來判斷電影片段中是否有環場音效事件的存在。在本節中，我們將說明所採用環場音效特徵值擷取與計算方式。

4.1. AC-3 音效特徵值

要取得 AC-3 環場音效的資訊，我們必須先取得各個聲道的音訊特徵值。本文使用的特徵值來自於 AC-3 的解碼程序中的 6 聲道解耦合(decoupling)模組之輸出，如圖 4 所示。

首先，從以 AC-3 壓縮的音訊資料中取出代表六個聲道的音訊框架中各個子頻帶(sub-band) 的資料 $sb_i(n)$ ，計算其平方值：

$$sb_i^2(n) = Mant_i Exp_i \quad (1)$$

其中 $Mant$ 與 Exp 分別為第 n 個框架(frame) 離散餘弦轉換係數之假數(Mantissa)與指數(Exponent)， i 為子頻帶的頻帶編號。接下來計算其每秒平均的能量 σ_{RMS} ：

$$\sigma_{RMS} = \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{256} \sum_{i=0}^{255} sb_i^2(n) \quad (2)$$

其中 N 為每秒鐘之音訊框架(frame) 個數； sb_i^2 為第 i 個子頻帶的係數值，取平方計算其能量。

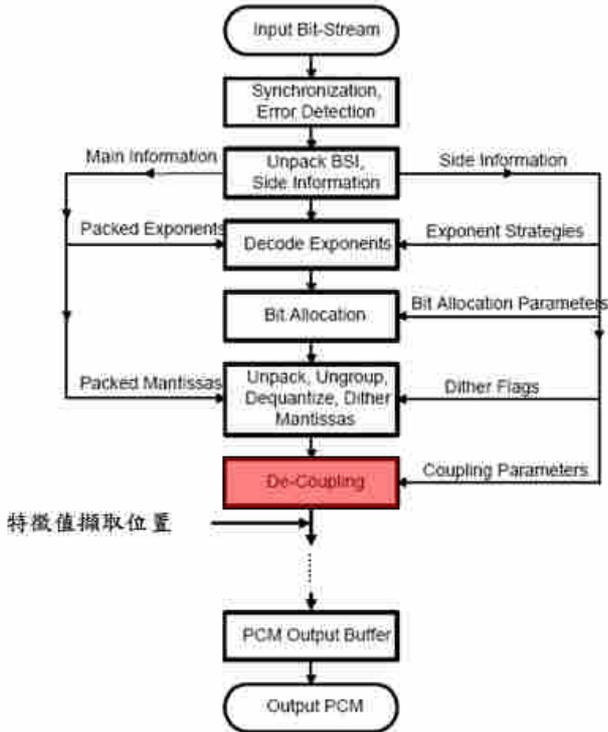


圖 4 Dolby AC-3 解碼流程圖[1][12]

4.2. 音量特徵字串

為了讓一般數值化的環場音效特徵更接近人類感知的特性，並使特徵值符號化以便於分析，我們將數值化的特徵值轉換為音量特徵字串。

所謂的音量特徵字串，是指在一個連續時間的電影片段中，先依照固定的單位時間(例如 0.5 秒)進行切割，切割的片斷稱之為子片段；再根據子片段中的音量大小給定一個代表性字元符號。根據我們的實驗得知，大部分的具有環場音效事件的片段，音量都集中在 55dB ~ 90dB 之間，因此，我們根據實際電影音效，統計其音量分布並進行實驗，以測量各種音量範圍對電影欣賞者的實際感受效果，將感受接近之音量範圍以一個音量字元取代，如表 1 所示。所以一個電影片段中的一段音訊訊號可以轉換成一個由音量字元所組成之字串，我們稱之為音量特徵字串。

表 1 音量字元

音量範圍	音量字元	音量範圍	音量字元
0~10	0	70~75	5
10~40	1	75~80	6
40~55	2	80~85	7
55~65	3	85~90	8
65~70	4	90~100	9

由於 AC-3 具有六個聲道，因此會有六個彼此之間互相獨立的音量字串計算流程，如圖 5 示。

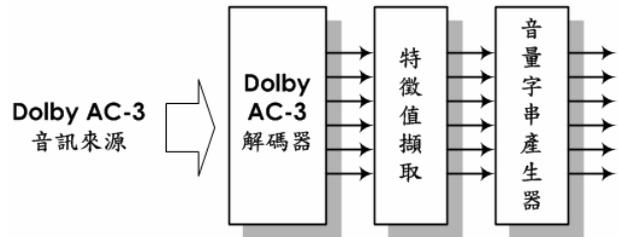


圖 5 音量字串擷取流程圖

4.3. 音效事件模型

在說明音效事件模型之前，我們先舉一個例子說明音效事件的涵義：假設我們正在觀賞一部影片，影片中有一枚巡弋飛彈由我們的左前方飛到右後方，並且在右後方引爆。根據這個例子，我們可以得到如圖 6 所示的音量字串。

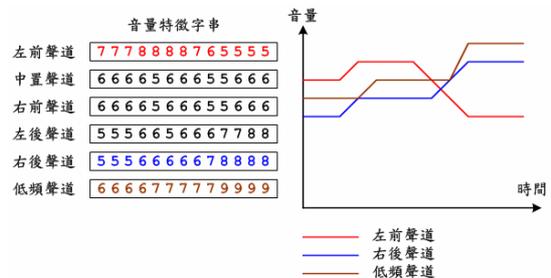


圖 6 音量特徵字串與事件關聯模型

AC-3 是以觀眾的位置為中心，進而模擬出四周的環場音效。因此，當某個物件在觀眾的周圍移動時，相對位置的聲道將會做出適當的音量變化反應，這就是音效事件模型。如圖 6 所示，由於物件是由左前方往右後方移動，因此剛開始的音量集中在左前方，後半段的音量才集中在右後方，由音量變化圖便可以看出端倪。

圖 9 為一個典型的音量字串與劇情的關聯示意圖，我們以人工的方式挑選出電影中具有環場音效的劇情片段，再將這些片段所對應的音量特徵字串取出。

為了能夠更確切的表達音效事件，我們針對環場的模式做出了如圖 7 的分類樹。前四個分類我們都明確的定義它的規則(3)(4)(5)(6)，剩餘的兩個分類則是其他音效分類的延伸規則，其中 α_i 、 β_i 代表音量字元，而每個音量字元的時間單位為 0.5 秒。

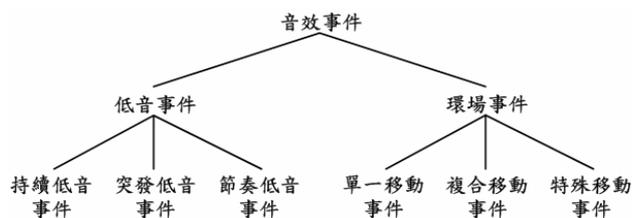


圖 7 音效事件分類樹

在低音事件方面，我們細分為三個類型：

■ 持續低音事件

在連續的時間內，低頻聲道中持續有強烈的低頻訊號產生。持續低音事件偵測之音量字串比對樣板 $P_{\text{持續低音}}$ 為

$$\begin{cases} P_{\text{持續低音}} = \beta_0 \beta_1 \dots \beta_n \\ n \geq k, \beta_i \geq B, |\beta_i - \beta_{i+1}| \leq W \end{cases} \quad (3)$$

其中， k 為持續低音事件最小的持續時間； B 為最小低音能量； W 為連續兩個音量字元的最大能量差距。根據實驗分析，系統參數設為 $k = 4, B = 5, W = 2$ 。

■ 突發低音事件

低頻聲道在某個時間點突然的發出強烈訊號，而持續的時間不會太長。突發低音事件偵測之音量字串比對樣板 $P_{\text{突發低音}}$ 為

$$\begin{cases} P_{\text{突發低音}} = \beta_0 \beta_1 \dots \beta_n \\ k_{\max} \geq n \geq k_{\min}, \beta_i \geq B \end{cases} \quad (4)$$

其中， k_{\max} 為突發低音事件的最長持續時間； k_{\min} 為突發低音事件的最短持續時間； B 為最小突發低音能量。根據實驗的分析，系統參數設為 $k_{\min} = 1, k_{\max} = 3, B = 7$ 。

■ 節奏低音事件

低頻聲道發出有節奏性的訊號。節奏低音事件偵測之音量字串比對樣板 $P_{\text{節奏低音}}$ 為

$$\begin{cases} P_{\text{節奏低音}} = (\alpha_0 \alpha_1 \dots \alpha_m \beta_0 \beta_1 \dots \beta_n) \times R \\ \alpha_i \leq B_{\min}, \beta_i \geq B_{\max} \\ k_{\min} \leq m \leq k_{\max}, k_{\min} \leq n \leq k_{\max}, R \geq 2 \end{cases} \quad (5)$$

其中， k_{\max} 為節奏低音事件的最長持續時間； k_{\min} 為節奏低音事件的最短持續時間； B_{\min} 為較低的低音能量之臨界值； B_{\max} 為較高的低音能量之臨界值； R 為節奏低音字串重複的次數。根據實驗分析，系統參數設定為 $k_{\min} = 1, k_{\max} = 4, B_{\min} = 5, B_{\max} = 7$ 。

環場事件方面，我們也細分為三個類型：

■ 單一移動事件

聲音由前左、前右、後左、後右的任意一個聲道移動到另一個聲道，如圖 8 所示。

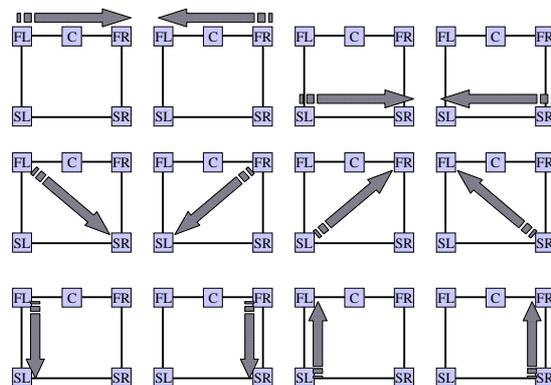


圖 8 12 種音效單一移動模式

偵測單一移動事件必須參考到兩個聲道的音量字串，所以音量字串比對樣板 $P_{\text{單一移動}}$ 亦有兩組

$$\begin{cases} P_{\text{單一移動}} = \begin{pmatrix} \alpha_0 \alpha_1 \dots \alpha_m \beta_0 \beta_1 \dots \beta_n \\ \beta'_0 \beta'_1 \dots \beta'_n \alpha'_0 \alpha'_1 \dots \alpha'_m \end{pmatrix} \\ |n - m| \leq 2 \\ \alpha_i, \alpha'_i \leq B_{\min} \\ \beta_i, \beta'_i \geq B_{\max} \\ k_{\min} \leq m \leq k_{\max}, k_{\min} \leq n \leq k_{\max} \end{cases} \quad (6)$$

其中， k_{\max} 為物件在某一個聲道中的最長持續時間； k_{\min} 為物件在某一個聲道中的最短持續時間； B_{\min} 為較低的能量之臨界值； B_{\max} 為較高的能量之臨界值。根據實驗分析，系

統參數設定為 $k_{min} = 1, k_{max} = 3, B_{min} = 4, B_{max} = 6$ 。

■ 複合移動事件

在短時間 T_c 內有多個連續的單一移動事件連續發生，例如左右盤旋。根據實驗 $T_c = 20$ 個音量字元，兩事件最大間隔時間 $T_{gap} = 2$ 個音量字元。

■ 特殊移動事件

在物件移動的過程中，伴隨著強烈的低頻訊號；或是前方的聲道為靜音時，後方的聲道卻發出聲音，表示主角後方有事件發生。

表 2 電影劇情與音效事件實例

劇情與音量字串	劇情描述	音效事件
	飛彈爆炸會產生連續而強烈的低頻訊號	持續低音事件
	突然的汽車追撞會使得低頻聲道出現一小段的強烈訊號	突發低音事件
	軍隊行軍時的踏步聲	節奏低音事件
	球體由左前方往左後方移動	單一移動事件
	球體由左後方移動到右前方後，再由右前方移動到左後方	複合移動事件
	炸彈在汽車的後方引爆，而震波由汽車後方移動至前方	特殊移動事件

5. 劇情與音效事件資料探勘

不同類型的事件或劇情，都會對應到一個或多個不同的音效事件，也會對應到不同的音量字串。因此，藉由音量字串的分析，我們便可以對各種事件或劇情進行關聯法則分析以找出其間的規則。

5.1. 電影劇情分類

為了針對電影劇情進行分析，我們首先必須對電影劇情進行分類的動作，電影劇情分類樹如圖 9 所示。

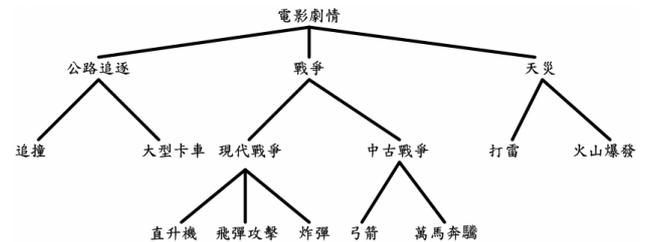


圖 9 電影劇情分類樹

有了圖 9 所示的分類樹，我們便可以開始針對各種劇情進行分析。由於本文所提出的音量字串是屬於連續時間的序列型態，因此，我們首先利用第 4 節所提出之音效事件偵測樣板，經由樣板比對來找出隱藏在音量字串中的音效事件。

舉例來說，如圖 10 之電影劇情，其對應的低音字串為 $\{7755775577557777\}$ ，以節奏低音事件樣板比對之後，我們得到

$$P_{\text{節奏低音}} = (5577) \times 3$$

$$\alpha_i = 5, \beta_i = 7$$

$$m = 2, n = 2$$

由上述分析得知此低音字串符合節奏低音事件樣板的規範。



圖 10 節奏低音事件

5.2. 音效事件與劇情之關聯分析

我們從 8 部電影中以人工的方式擷取了 100 個電影劇情，由這些電影劇情的音效部分擷取出 AC-3 音效特徵值，再計算出每個電影劇情的 6 組音量字串。接著經由音效事件偵測樣板，我們得到每個電影劇情所對應之音效事件，如表 3 所示。

表 3 電影劇情與音效事件

電影劇情	AC-3 音效事件
炸彈爆炸	持續低音、單一移動
槍戰	突發低音、單一移動
飛彈	持續低音、單一移動
打雷	突發低音
火山爆發	持續低音、突發低音
機關槍	持續低音、單一移動

根據表 3 所示，我們將各種音效事件視為項目 (item)，再利用 Agrawal[2][3] 所提出的 Apriori 關聯探勘演算法，找出各種電影劇情所對應的共同音效事件。部份關聯法則如表 4 所示。

表 4 電影劇情與音效事件之探勘結果

AC-3 音效事件與電影劇情關聯	Support	Confidence
{持續低音、突發低音} → 火山爆發	4%	100%
{持續低音、突發低音} → 打雷	8%	87.5%
突發低音 → 槍戰	6%	66.67%
持續低音 → 機關槍	5%	60%
節奏低音 → 機器人行走	2%	100%
突發低音 → 大型怪物	4%	100%
持續低音 → 爆炸	10%	90%
持續低音 → 萬馬奔騰	2%	100%
{持續低音、單一移動} → 飛彈爆炸	2%	50%

6. 結論

本文提出了一種環場音效與電影劇情的關聯模型，藉由 Dolby AC-3 的環場特性，擷取出六個聲道的音量字串。

在環場模式方面，我們進一步的將其分為六大類，目的是為了能夠更確切的表達各種環場模式。

我們利用 6 種環場模式組合出各種電影劇情，再搭配關聯探勘演算法，試圖找出各種電影劇情所對應的環場音效模型。

在未來工作方面，我們目前正在進行大規模的挑選環場音效劇情片段的剪輯工作，以建立大型環場影片資料庫以使關聯法則分析的樣本集更完整。其次，我們也考慮加入以往我們在音色分類的研究成果，以提高劇情自動分類的效能。最後，我們將所得到的關聯法則用於電影摘要的研究，以建立更完善的電影自動摘要系統。

7. 參考文獻

[1] ATSC A/52, "Digital Audio Compression (AC-3) Standard," *United States Advanced Television Systems Committee*.

[2] Agrawal, R., Imielinske, T., and Swami, A. "Mining association rules between sets of items in large database," *Proc. of the ACM SIGMOD 1993 International Conference on Management of Data, Washington, D.C., May 1993*, pp.207-216.

[3] Agrawal, R., and Srikant, R., "Fast Algorithms for

Mining Association Rules," *Proc. of the 20th International Conference on Very Large Database, Santiago, Chile*.

[4] De Santo, *et al.*, "Classifying Audio of Movies by a Multi-expert System," *Image Analysis and Processing, 2001. Proceedings. 11th International Conference on*, pp. 386-391, 26-28 Sept. 2001.

[5] Lu, G.J. and T. Hankinson, "A Technique Towards Automatic Audio Classification and Retrieval," *In Proc. IEEE Intl. Conf. on Signal Processing, Vol. 2*, pp. 1142-1145, 1998.

[6] C.M. Liu, S.W. Lee, and W.C. Lee, "Bit Allocation Method for Dolby AC-3 Encoder", *IEEE Trans. On Consumer Electronics*, 1998

[7] Rainer Lienhart, Silvia Pfeiffer, and Wolfgang Effelsberg, "Video Abstracting," *Communications of the ACM, vol. 40, No. 12*, pp. 55-62, 1997.

[8] Minami K, Akutsu A, Hamada H, Tonomura Y, "Video Handling with Music and Speech Detection", *IEEE Multimedia*, 1998.

[9] Nakajima, Y. *et al.*, "A Fast Audio Classification from MPEG Coded Data," *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing, vol. 6*, pp 3005-3008, 1999.

[10] N. V. Patel and I. K. Sethi, "Audio Characterization for Video Indexing," *SPIE*, vol. 2670, pp. 373-384, 1996.

[11] N. V. Patel and I. K. Sethi, "Video Classification Using Speaker Identification," *SPIE*, vol. 3022, pp. 218-225, 1997.

[12] Vernon, S., "Design and Implementation of AC-3 Coders," *IEEE Transactions on Consumer Electronics, vol. 41, Issue 3*, pp. 754-759, Aug. 1995.

[13] Watson, M.A.; Buettner, P., "Design and Implementation of AAC Decoders," *IEEE Transactions on Consumer Electronics, vol. 46, Issue 3*, pp. 819-824, Aug. 2000.

[14] Wold, E. *et al.*, "Content-based Classification, Search, and Retrieval of Audio," *IEEE Multimedia, vol. 3, no. 2*, pp. 27-36, 1996.

[15] Minerva M. Yeung and Boon-Lock Yeo, "Video Visualization for Compact Presentation and Fast Browsing of Pictorial Content," *IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, No. 5*, October 1997.

[16] 葉億真、劉志俊, "音效資料的內涵式分類及其在電影資料庫的應用," 第二屆數位典藏技術研討會, 2003.

重要劇情



00:02 飛彈爆炸



00:06 飛機起飛



爆炸 00:26

環境音效特徵字串

L 3233667766666666...6676777776634343.....333 3477667655577776633...
 C 3235778777766667...6555676666334443.....333 777767777778776633...
 R 3235667777666666...6677876666333434.....433 3567656655577666433...
 SL 555665665556555...6666776666666333.....436 5466666676777664333...
 SR 224565554456555...6777776666666333.....334 5565666676777664343...
 LFE 22225777777777...6556887666110000.....334 3367767765455553344...

00:00

00:30

重要劇情



公路追逐 00:36



公路追逐 00:38



炸彈 00:45

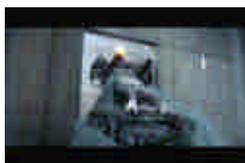
環境音效特徵字串

L 34433355554555676.....7667655553335676....645667766666576.....
 C 333233666666665676.....66777677666656777...6467766777766636.....
 R 34433355553366334.....67676666554446657...6236667766654434.....
 SL 36363355556673334.....3456655544435664....6555677877664533.....
 SR 31111355665673334.....56566665544435545....6555677777766634.....
 LFE 32222355554473333.....56566666653125888....6017777777766666.....

00:30

01:00

重要劇情



機槍 01:05



飛彈 01:09



飛機 01:14



炸彈 01:23

環境音效特徵字串

L 6666666666665...566645346666.....66666675676.....76565555565554
 C 6666666666666...366756653446.....33666676676.....666766545554444
 R 565554556554...566644336666.....36666663365.....76565555555455
 SL 4555454555556...677766766556.....44777775666.....455655565687765
 SR 4555544555564...577776766555.....44777776776.....34465555544357
 LFE 5566555555566...444543211111.....66555467777.....776777765677667

01:00

01:30

L 左前聲道 C 中置聲道 R 右前聲道 SL 左後聲道 SR 右後聲道 LFE 低頻聲道

圖 11 音效字串與重要劇情之關聯示意圖