

行政院國家科學委員會專題研究計畫 期中進度報告

多人環境下的 MPEG-4 媒體播放系統(1/3)

計畫類別：個別型計畫

計畫編號：NSC91-2213-E-002-090-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學資訊工程學系暨研究所

計畫主持人：陳文進

計畫參與人員：黃奕勤. 楊書旻. 林彥光. 蔡鈞傑

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 5 月 30 日

多人環境下的MPEG-4媒體播放系統(I)

國科會補助專題研究計畫

計畫編號：NSC 91-2213-E-002-090

全程計畫：民國 91 年 08 月 01 日至民國 94 年 07 月 31 日

本年度計畫：民國 91 年 08 月 01 日至民國 92 年 07 月 31 日

計畫主持人：陳文進 台灣大學資訊工程學系教授

一、摘要

(中文摘要)

本計畫中文名稱為“多人環境下的 MPEG-4 媒體播放系統”，目標在完成在網路環境下符合 MPEG-4 架構的多人互動系統。多媒體系統的發展繼 VCD(MPEG-1)、DVD(MPEG-2)之後，接下來最大的發展與突破便是 MPEG 制定群於 1998 年制定的國際標準—MPEG-4。這個標準，攬括了目前所有的多媒體領域，試圖利用一個標準統一所有多媒體的技術，並利用物件概念徹底顛覆目前以畫面為主的傳統媒體，取而代之的，是一個可以整合現今 3D/2D Graphic、Animation、Video Codec、Streaming、Interactive 與 Programmatic 的複合環境。

然而現有 MPEG-4 媒體在多人使用環境下的能力非常的薄弱，使得 MPEG-4 的能力備受質疑，也使得當初 MPEG-4 規格在制定時的遠景被大大打了折扣。有鑑於此，本計畫試圖完成 MPEG-4 在多人使用環境下的機制，包括有 MPEG-4 異質媒體傳輸機制、MPEG-4 異質媒體同步機制、MPEG-4 伺服器端與接受端多人使用環境機制，以及最後為了驗證整套系統架構的可行性以及正確性而設計的幾項 MPEG-4 網路應用系統，如果此項計畫成功將是直接驗證 MPEG-4 系統確實擁有當初制定時所聲稱的擁有高彈性、一般性，也將把 MPEG-4 從研究階段引領至可實用性的層級。

(英文摘要)

The name of this project is “MPEG-4 based Multimedia Presentation Systems for Multi-Users Environment”. Its goal aims at implementing the multi-users interactive systems under MPEG-4 architecture. MPEG-4 is the succeeding standard of MPEG-1 and MPEG-2 video standards. This standard is designed by ISO standard committee in 1998. Instead of using the current frame-based video technology, it adopts the object-oriented concept, which integrates the existing multimedia technologies, such as 3D/2D graphic、animation、video codec、multimedia streaming、interactive and programmatic environment into single architecture.

However, current MPEG-4 standard is weak in supporting the multi-users applications. Therefore,

the goal of this project is to design the multi-users mechanism, including MPEG-4 heterogeneous transport protocol, MPEG-4 heterogeneous synchronization mechanism, and MPEG-4 client-server multi-users environment. This project will also develop several MPEG-4 application systems to verify the correctness and feasibility of the above mechanism.

二、計畫緣由與目的

MPEG-4 標準是由 MPEG 組織(Moving Picture Experts Group)所制定,是由數百個當今學術界以及工程界中的傑出研究人員以及工程師共同制定。MPEG 組織過去曾經發展了眾所皆知的 MPEG-1 以及 MPEG-2 標準,分別為目前膾炙人口的 VCD 以及 DVD 規格所使用的技術,並榮獲了 Emmy Award 大獎,可以見得 MPEG 組織在學術界以及業界的深遠地位。

MPEG-1 以及 MPEG-2 主要的應用都是在播放儲存在 CD-ROM 或 DVD-ROM 裡面的電影,或是遠端傳送的高畫質數位電視,其主要的功能是以畫面為主的影像壓縮技術。然而, MPEG-4 標準試圖將當今目前所有的多媒體技術通通整合起來,以一個統一的標準去規範當前甚至未來的多媒體技術,並建立起新的多媒體整合環境。然而,各種多媒體技術的細節可能差異頗大,本質上也可能差異頗多,所以為了整合這些多媒體元素, MPEG-4 突破了傳統做法,以物件導向(Object-Oriented)的精神去整合這些元素,讓各種本質上差異頗大的多媒體技術也能夠整合在一個畫面內,使得各種多媒體元件不用再轉換成相同形式,可以使用其最自然的型態並加以整合。

為了符合在各種場合下的應用, MPEG-4 所需的頻寬可以從 5 Kbps 有彈性的至 10Mbps,不管是高畫質電視的應用或是掌上型數位助理的無線通訊應用,都在 MPEG-4 的設計考量內,所以不只是以畫面為主的應用,也不只是侷限於某些軟硬體下的應用環境, MPEG-4 都試圖將他們整合。目前 MPEG-4 最被看好的應用領域有以下幾類:

1. 高畫質互動電視應用
2. 消費性網路娛樂平台
3. 跨平台視訊通訊服務

將以上的應用分類之後,可以整理出幾項重要的技術項目:

1. 網路串流(Streaming)技術

在高畫質互動電視的應用領域裡,這項技術需要的是向網路高畫質的影片傳輸技術,而在跨平台的視訊通訊服務裡面,需要的是能夠適合各種低頻寬平台以及各種寬頻寬平台的影像傳輸技術,而在消費性網路平台裡,需要的則是各種媒體素材的串流技術,在這個例子裡面,各種媒體素材的種類有圖片、聲音、2D/3D 動畫(Animation)資料、JavaScript 程式片段、即時性影片等等,這些媒體素材也是必須藉由網路串流技術來傳輸,所以在這個技術領域裡,要解決的問題就是各種異質媒體的串流技術。

2. 伺服器與多人互動環境技術

在消費性網路娛樂平台的應用領域裡,最具挑戰的項目就是使用者如何與伺服器以及其他的使用者之間互相互動,所謂的互動指的是播放端的畫面能夠根據當時真實的情況反映出符合使用者感官的畫面。所以多人互動的目的可以分為兩種種類,其一是畫面的同步性,指的是每個使用者能夠即時的看到當時正確的畫面,每個使用者也許當時正在觀看的角度與環境位置不同,但都必須看到正確的畫面結果;其二是畫面的邏輯性正確,指的是也許每個使用者可能會看到沒

有同步或是有誤差的畫面，但是不會影響每個使用者參予互動的邏輯性。這兩種類別都有其使用上的適當場合，目前在其他的娛樂平台上都可以看見他們的應用實例。

3. 接受端媒體即時合成技術

以上三種應用裡面，都需要使用畫面合成的技術，因為 MPEG-4 的場景是以物件導向為基本概念，所有媒體都是以物件形式存在，而在接受端就需要即時的將所有媒體素材合成，MPEG-4 之所以能夠成為育樂娛樂平台，之所以能夠即時的對使用的參予而有所反映，這些都跟它擁有即時合成技術很有關係，所以沒有了即時合成技術，也就使得 MPEG-4 立即喪失之前提及的一些好處。

以上各項技術皆為本計劃所欲發展的核心技術，也是完成多人互動平台最重要的元素，沒有以上的技術將無法完成 MPEG-4 各種不同的應用領域的需求，所以本計劃將一一的實現他們。

三、研究方法及成果

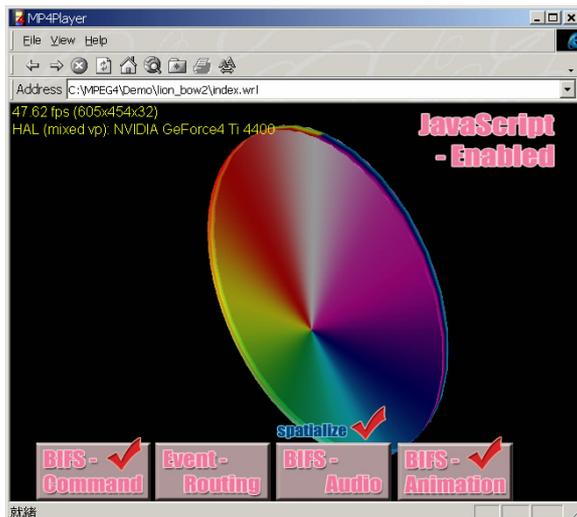
本計畫預計以三年時間完成，第一年主要目標在完成 MPEG-4 播放端無網路環境下互動機制以及伺服器端媒體串流機制，第二年主要目標在完成 MPEG-4 播放端與伺服器在多人網路環境下的同步機制，而第三年主要目標在驗證之前所完成的各種機制的可行性以及正確性，驗證方法是利用完成各種 MPEG-4 網路環境下的應用來驗證是否整個機制有任何的缺點。

目前本計劃接近第一年之完成階段，預期完成目標均已達成，以下簡介這些成果。

1. 完成 MPEG-4 播放端 JavaScript 機制：

若在一個遊戲場景中，想要表示以下邏輯“若使用者拿到正確的鑰匙，而且門現在是關著的，那麼就打開門。”的話，用現有的機制是很困難而且也非常不直覺的，所以，MPEG-4 沿襲 VRML97 的基本方法並加上了 ECMAScript (ECMA: The European Computer Manufacturers Association, 定義在 ISO/IEC 16262) 的規格作為 MPEG-4 內嵌 JavaScript 的語言規格。

使用 JavaScript 可以很簡單的解決需要擁有邏輯的場景，因為邏輯本身就比較適合使用程式碼的方式去呈現它。讓場景作者直接在場景中放入程式碼，並在媒體播放時即時執行，並根據當時的使用者的輸入以及環境即時的執行產生結果，這種方式絕對是最具有彈性而且受限最少的，然而，這樣的方式使得 MPEG-4 場景的編輯者必須具備程式設計的能力，我們要知道，並非所有的人都有能力撰寫程式。所以，MPEG-4 也定義了一些額外的互動機制，例如 Event-Route、BIFS-Command、BIFS-Audio 以及 BIFS-Anim。在我們的實作中，我們成功的將 JavaScript 整合以上 MPEG-4 所定義的互動機制中，並讓媒體作者可以充分的發揮他們的創意。



圖表 1：本計劃所開發的 MPEG-4 瀏覽器，並且成功將 JavaScript 技術用以控制 MPEG-4 Scene Description 所定義的各項互動機制。

2. MPEG-4 播放端各媒體的無網路環境下媒體同步機制：

在 MPEG-4 場景中，雖然媒體的類型有很多種，但是基於同步的考量而分類的話，我們可以將媒體分為三個類型，這三個類型是根據每個媒體時間單位（如果是影片，則約是 1/30 秒，如果是 CD 品質的聲音，則是 1/44100 秒）與電腦能夠處理的時間而定，可分為：

1. 離散時間型媒體：

這類的媒體的定義就是媒體的時間單位相對電腦能夠處理的時間而言非常的大，也就是說，並不需要 CPU 時時刻刻一直處理這個媒體，而對人類的感官而言，也不會覺得不順暢。

2. 立即時間型媒體：

這類的媒體其媒體資料可由目前時間計算而得，由於這類媒體永遠可以在播放的剎那取得最精準的媒體資料，所以在各類媒體中是最不需要刻意同步的。

3. 連續時間型媒體：

這類的媒體由於其單位時間極小，一旦中斷其媒體的處理或是處理媒體不夠連續均勻就會產生使用者的感官難以接受的結果。

我們針對以上媒體設計不同的同步機制：

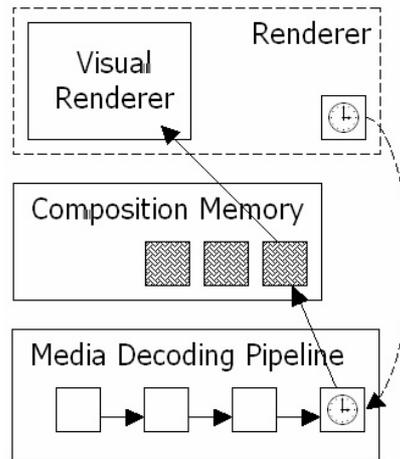
1. 立即時間型媒體：

這類的媒體由於可以在播放的剎那取得最精準的媒體資料，所以並不需要刻意考慮同步的問題，只要系統畫面高於 25fps ($1000\text{ms} / 40\text{ms} = 25$)，這類的媒體就可以達到同步的要求。

2. 離散時間型媒體：

這類的媒體主要的同步準則就是儘早解壓縮完畢，並且等待直到系統的絕對時間到達媒體播放時間，然後再將媒體播放出去。在這裡，需要考慮的因素是系統也有可能發生忙碌的情形，使得系統來不及向 Decoder 拿取解壓縮好的媒體，如果 Decoder 會等待 Renderer 來拿取媒體的話，

將使得 Decoder 發生越來越嚴重的惡性循環，所以，一個間接的架構必須被設計出來。



圖表 2：離散時間媒體同步架構。

圖表 2 為離散時間型媒體的同步架構，為了分開解壓縮端與 Renderer 之間的控制，使得兩邊的流程能夠彼此獨立，我們在兩者之間放置 Composition Memory 模組。當 Decoder 解出一個 Sample 後，便會進入等待，直到這個媒體應該被播出的時間再扣除一個 Sample 的時間時就會將之置於 Composition Memory，並替換帶之前舊的媒體 Sample。而 Renderer 會在該 Rendering Frame 時至 Composition Memory 取出該媒體最新的 Sample，並拿去合成之用。

3. 連續時間型媒體：

由於 MPEG-4 目前只有聲音是連續時間型媒體，所以我們將針對聲音為主要的討論對象。聲音由於每個 Sample 的取樣時間極短，所以必須要有音效卡這項硬體的輔助，傳統處理聲音的流程是將聲音置於一個 Buffer 中，並告知音效卡這個 Buffer 的取樣 Bits 數、Channels 數、每秒取樣頻率，然後音效卡就能為我們撥放這個聲音。

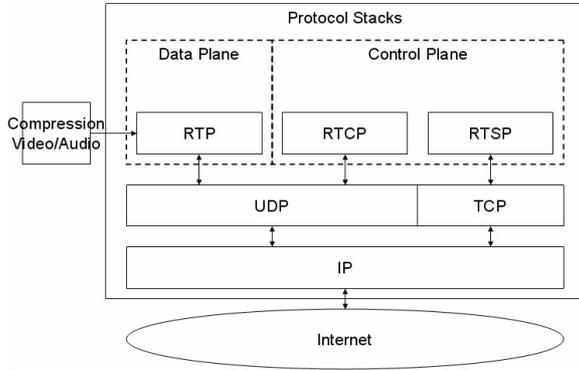
通常聲音的播放速度會隨著軟硬體的設計而有不同的誤差，這些誤差來源有以下原因：第一，硬體不一定會支援跟我們要撥放的聲音一樣頻率的 Sampling Rate，所以硬體會選擇較接近的頻率播放，這兩個頻率由於相差甚小，所以人類聽覺不一定會有感知。第二，系統可能突然處於忙碌狀態，使得記憶體中的 Buffer 來不及搬到音效卡的記憶體中，使得聲音出現一小段的誤差，通常這種偶而發生一次的錯誤人類聽覺有可能以為聲音本來就是如此而忽略。

以上誤差在一般應用中並不是需要關心的問題，因為在一般應用中聲音都是時間的參考者，反正全部媒體都是參考聲音，就算誤差累積也並不破壞整體的同步性。但是這些誤差在 MPEG-4 中卻破壞同步極為嚴重。因為 MPEG-4 場景中存在許多的聲音，任何聲音播放的快慢都將破壞整體的同步性，再加上人類對聲音的敏感度極高，使得這個問題顯的極為困難。

3. 伺服器端與 MPEG-4 播放端之間一般媒體(Video、Audio)的串流(Streaming)架構：

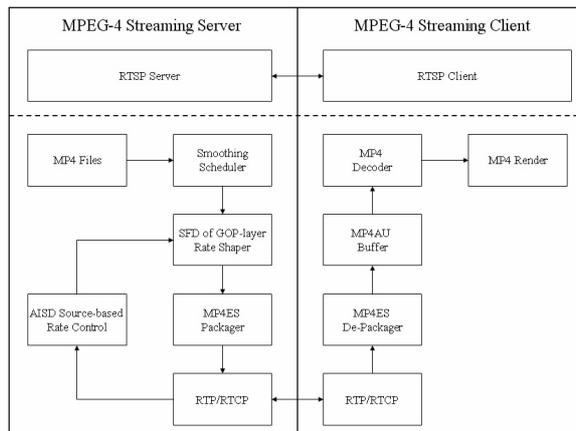
MPEG-4 系統可以提供使用者一個完整且有效率的內容描述與同步呈現的架構，然而這樣的架構卻無法對應到現在的網際網路的傳輸協定上，因此 ISO 組織另外制定了一個 MPEG-4 over IP 的傳輸標準。

在 MPEG-4 標準的第八部份(Carriage of ISO/IEC 14496 contents over IP networks)，定義了在網際網路上傳輸 MPEG-4 資料的系統架構。在此架構中，指出了使用 RTP/RTCP 作為網路傳輸層的傳輸協定(transport protocol)，以 RTSP 為其網路會議層的控制協定(session control protocol)，並以 SDP 為其會議描述協定(session description protocol)。如下圖所示：



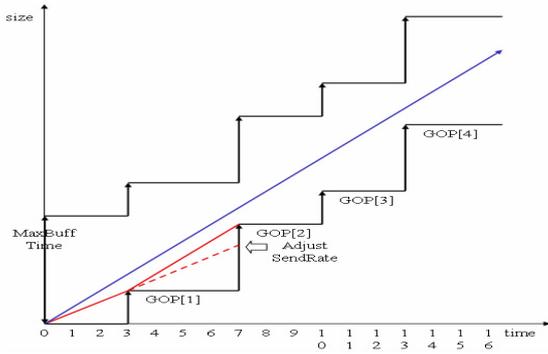
圖表 3：MPEG-4 over IP 傳輸層協定

在符合 MPEG-4 over IP 與串流傳輸的架構下，我們提出了 MPEG-4 平順傳輸串流架構。在我們的傳輸架構中，希望除了符合標準外，並且能夠有適當的流量控制與平順的傳輸排程。下圖表示 MPEG-4 平順傳輸串流架構的系統架構：



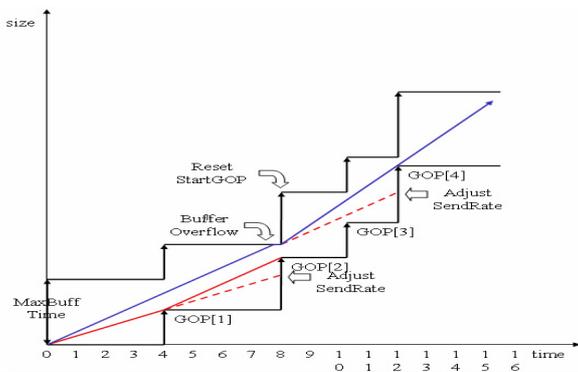
圖表 4：MPEG-4 平順傳輸串流架構圖

如圖表 5, 我們的 MPEG-4 平順傳輸串流技術可以調整傳送者的傳送速率，使傳送者能夠以適當的傳送速率，達到即時傳輸的要求。當傳送者的傳送速率無法達到即時傳輸的要求時，我們需要調整傳送速率。在調整傳送速率後，我們需要重新計算傳送過程，使我們的統計能夠正確無誤。



圖表 5：平順傳輸排程示意圖(一)

如圖表 6, 如果發生接收者目前的緩衝區滿溢, 則重新計算傳送過程的起始點可以由此開始。因為發生緩衝區滿溢之前的傳輸速率, 已經不影響其後的傳送速率。如此可以減輕計算出適當傳送速率的時間。



圖表 6：平順傳輸排程示意圖(二)

使用我們的平順排程將可以大為降低所需播放的頻寬使用量, 如下圖所示, 可以看出我們的設計更具有有效利用有限頻寬的特性。



圖表 7: 左圖為使用非平順排程的使用頻寬變化圖; 右圖為使用我們實作的平順排程演化法的使用頻寬變化圖。

四、結論與討論

本計劃將完成 MPEG-4 網路環境下完整的互動機制、異質媒體串流技術、多人使用環境下的互動機制，預期對於多媒體互動教學領域、以及以 CD-ROM 或是 DVD 形式存在的媒體形式、單向網路形式(衛星或是 Cable)形式存在的媒體應用領域、Internet 上的網路連線遊戲、互動電視領域，都將有直接的刺激以及幫助，並帶動許多新的商機。

由於 MPEG-4 試圖以統一的架構去整合現有的多媒體技術，包括有 2D/3D Graphic、Animation、Video Codec、Streaming、Interactive 與 Programmatic 的複合環境，所以發展 MPEG-4 技術將使得整個多媒體領域技術提昇，提高國家多媒體領域的競爭能力，對於現在重視以及強調軟體產業的我國而言，絕對是不容忽視的領域技術。

目前產業界的一致共識是，與其發展一個個單獨且互不相同的技術，不如整體考量，讓這一個個的技術可以像堆積木般發揮整合的效果，進而提高整體的實用價值。MPEG-4 就是為了此項目的而存在，尤其它更是國際知名的多媒體標準，去提高國內 MPEG-4 技術的水準可以大大提昇國內的多媒體競爭能力，對於國人將來競爭國際互動電視平台、網路互動資訊服務、線上遊戲等等都有極大的幫助，因此，MPEG-4 絕對是一個我們不可放棄的技術領域。

五、相關論文之發表狀況

1. 台大資工所 禹智鏗碩士論文，“MPEG-4 平順傳輸串流架構設計與實作”，民國 92 年 6 月。
2. 台大資工所 楊鈞傑碩士論文，“MPEG-4 多人使用環境下的架構設計與實作”，民國 92 年 6 月。
3. 台大資工所 謝孟吉博士論文，“MPEG-4 系統設計與實作”，民國 92 年 6 月。

六、參考文獻

- [1] ISO/IEC 14496-1, Information Technology – Generic Coding of Audio-Visual Objects: Systems, Final Committee Draft of International Standard, May 1998.
- [2] ISO/IEC 14496-2, Information Technology - Coding of Audio-Visual Objects: Visual, Committee Draft of International Standard, May 1998.
- [3] ISO/IEC 14496-6, Delivery Multimedia Integration Framework, Committee Draft of International Standard, March 1998.
- [4] “VRML 97: The Virtual Reality Modeling Language”, International Standard ISO/IEC 14772-1:1997.
- [5] 其他參考文獻請參考計畫書(NSC 91-2213-E-002-090)。