

行政院國家科學委員會專題研究計畫成果報告

MPEG-4 系統及 3D 視訊編碼技術之研究

A Study on MPEG-4 System and 3D Video Coding

計畫編號：NSC 89-2213-E-009-132

執行期限：88 年 8 月 1 日至 89 年 7 月 31 日

主持人：杭學鳴 國立交通大學電子工程學系教授

計畫參與人員：陳思聖，黃揚智，趙子毅，陳景川，孫逢佑

國立交通大學電子研究所

中文摘要

MPEG-4 標準採用以物件為基礎的影音表示模型。本計畫的目標是研究並且模擬 MPEG-4 Systems 標準。我們特別專注在 MPEG-4 標準中的互動性特色。這部分的工作可分三項。第一項是研究 MPEG-4 Systems 部分的標準文獻，其次是追蹤及編譯 MPEG-4 提供的編寫工具及播放程式。最後，使用 MPEG-4 的編寫工具來建構互動式的景來展現 MPEG-4 提供的互動能力。本報告中亦含 2D/3D 物件導向視訊表示法之初步成果。

關鍵詞：MPEG-4 標準，多媒體系統，物件導向視訊

Abstract

This report is dedicated to the study and simulation of the MPEG-4 Systems specifications, particularly on the interactive support features of MPEG-4. Three tasks are involved in this work. The first task is studying the MPEG-4 Systems specifications, especially on the scene description and its interactive support definitions. Secondly, we trace and rebuild the MPEG-4 authoring tools and players. To demonstrate the MPEG-4 interactive capabilities, the final last task is using the MPEG-4 authoring tools to construct interactive scenes as shown at the end of this report. Also included are some preliminary results on 2D/3D object-based representation.

Keywords: MPEG-4 Standards, Multimedia Systems, Object-based representation

第一部份 MPEG-4 系統的研究

A. 背景與目的

ISO (國際標準組織) MPEG (Moving Picture Expert Group, 動態影像專家群) 已經制定完成 MPEG-1、MPEG-2 國際標準，並成功帶動 VCD、DTV、DVD 等產業的起飛。也因此下一代多媒體傳輸系統的標準 MPEG-4 (ISO/IEC 14496) 備受各方注目。MPEG-4 應用範圍至少包含 DTV、Interactive Graphics Applications 及 Interactive Multimedia (如 World Wide Web) 等三大領域。我們可以發現領導軟體產業、電腦產業、通訊產業甚至娛樂產業的各大公司幾乎無一缺席 MPEG-4 標準的制訂。而這些公司加入標準的制訂更加強化 MPEG-4 做為下一代多媒體標準的地位。

MPEG-4 標準發展主要目標有三：一、提供多媒體製作者便利、可再生 (reusability) 的製作方式，並提供智慧財產權的相關管理與保護。二、給網路服務提供者 (Network Service Provider) 更多的資訊。如此網路服務提供者能因應各種媒體特性，而給予適當的 QoS (Quality of Service, 服務品質)。三、對於使用者，MPEG-4 則提供即時通訊 (Real Time Communication)、高度互動 (High Level Interaction) 及移動式多媒體 (Mobile Multimedia) 等諸多功能。

MPEG-4 標準採用以物件為基礎的影音表

示模型。此模型具有超鍊結及互動性的能力，並且支援自然及合成的內容。MPEG-4 符合了許多應用領域的需求，諸如在網際網路上傳送動態影像、多媒體廣播、以內容為基礎的影音資料庫的存取、遊戲、透過行動網路傳送進階影音通訊、電視購物以及遠端監控。MPEG-4 與在這標準之前的其他同類型標準最大的不同的特色，就在於它所提供的這個影音表示模型。一個以物件為基礎的景，建構在具有時間與空間關係的各種物件群上，它允許和諧的整合多種不同類型的資訊，同時並存在一個景中。本篇報告的目標是研究並且模擬 MPEG-4 Systems 標準。我們特別專注在 MPEG-4 標準中的互動性特色。

B. MPEG-4 標準與工具

我們先研究 MPEG-4 Systems 部分的標準文獻，特別是著重於景描述及它的互動性支援。其次是追蹤及編譯 MPEG-4 提供的編寫工具及播放程式。

MPEG-4 系統主要由 Scene Description 與 Object Description Framework 所構成。Scene Description 定義物件如何構成景 (Scene)。在 MPEG4 V1 中包含大部分 VRML 97 (ISO/IEC 14772-1:97) 的內容，同時也加入一些 MPEG 4 對聲音、臉部模擬等的特有節點。MPEG4 V2 中則會包含 VRML97 的全部內容，同時也強化對聲音、身體、臉部的模擬以及與多重使用者互動的功能定義。Object Description Framework 提供元件資料流 (Elementary Streams) 與景描述 (Scene Description) 之間的連結。Object Description Framework 是由一組能夠鑑別、描述並適當連結資料流與媒體物件 (Media Object) 的描述子 (Descriptor) 所構成。它並提供額外 OCI (物件內容資訊, Object Content Information) 與 IPMP (智慧財產權管理與保護, Intellectual Property Management and Protection)。

MPEG-4 中互動性的支援繼承自 VRML 的 event model。其中包括了 Sensors, Interpolators, Script 等節點，及 Route 語法。

Sensors 節點用來偵測使用者或因為景的變化而產生的 events。Interpolator 節點依照時變性的參數作內插用以產生動畫的效果。Script 節點可用來儲存其他節點的狀態、作些計算並操縱景。Route 語法則是將各種節點間 event 傳送的路徑連結起來。

IM-1 (Implementation 1) 團體釋放出了一些與 MPEG-4 標準相容的編寫工具 (authoring tools) 及播放程式 (Players) 的原始碼。我們將這些軟體套件進行除錯與修改部分原始碼之後，成功的編譯成執行檔以供使用。這些軟體套件都源自於一個核心 (Core)，此核心的主要類別及資料流的關係如圖一所示。

C. 模擬與實驗

IM-1 團體所釋放出的編寫工具有 BIFS/OD Encoder，用來將 textual format 的 Scene description 與 Object Description Commands 編碼成二進位的格式 (BIFS stream 及 OD stream)。TRIF/MP4 Multiplexer 將 BIFS stream、OD stream 及其他對應的 Media Streams 合成為 Trivial 或 MPEG-4 File Format 以供各式各樣的播放程式進行播放，這些編寫工具與播放程式之間的關係如圖二所示。

我們使用 MPEG-4 的編寫工具來建構互動式的景來展現 MPEG-4 提供的互動能力，如圖三、圖四、圖五所示。圖三所展現的是一個初始為綠色的圓盤，當使用者拖曳這個圓盤，則此圓盤的顏色會隨著圓盤本身所在位置不同而不同。圖四中有一白一紅大小不同的兩個圓球，當使用者點取白球時，其中一球會變成藍色一球變成綠色。若再點取較大的球則顏色又將互換。圖四則是有三個盒子，一黃一藍一紅。初始時，紅盒子進行逆時針旋轉。使用者點取黃盒子時，紅盒子變成綠角錐仍舊逆時針旋轉。若點取藍盒子則紅盒子變成紫色的圓柱，放開後又還原成紅色盒子。

D. 結論

雖然 MPEG-4 標準組織提供了一些基本的

工具用以對 MPEG-4 系統進行模擬，但這些工具與播放程式仍舊有許多缺漏及錯誤。此外，這些播放程式的原始碼非常龐大，其中任何一個程式都由超過十萬行的 C++ 語言的程式碼所構成。我們努力的追蹤並修改這些原始碼，終於成功的將他們編譯完成。最終，我們也回饋了一些重要的程式錯誤報告給 IM1 這個團體。

第二部份 物件導向視訊表示法

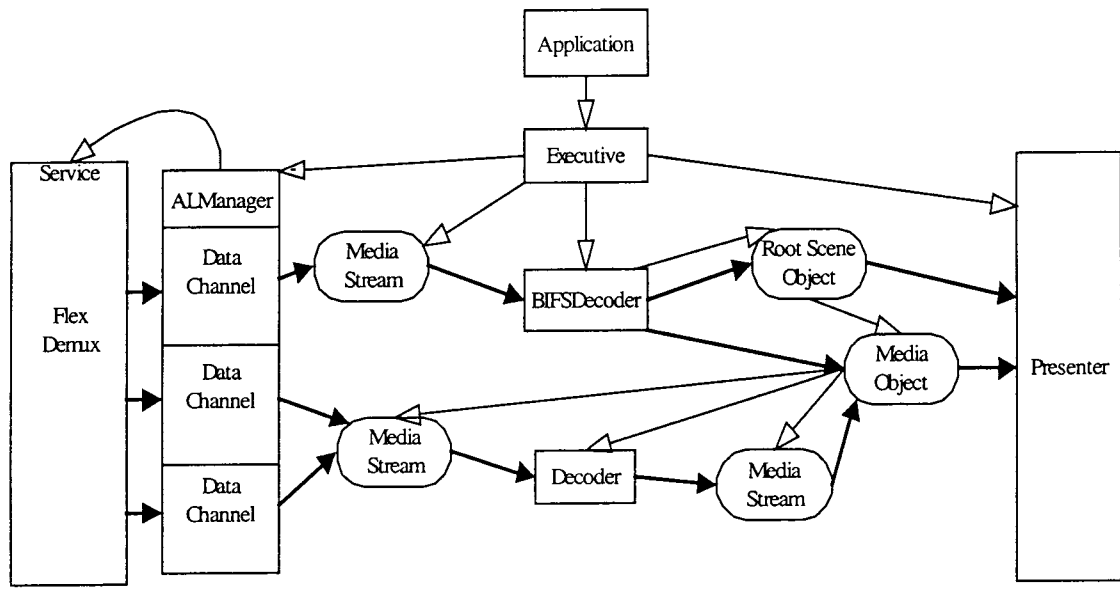
我們的目標是從一連串의 影像序列中，攝取其中物件的 2D/3D 資訊。較具體的說，我們希望能將物件分割出來，計算其 3D 結構資訊，轉換成適當的 2D/3D 表示法或編碼。在過去的研究中，我們採用 block matching 方式找出照片的 motion vectors。圖六是針對 *Flower Garden Sequence* 所求得 motion vectors 之一例。圖中 pixel 的灰度深淺代表 motion vector 值的大小，愈明

亮的顏色表示移動愈快，黑點是無法求得正確 motion vector 的點 (遮蔽區域)。

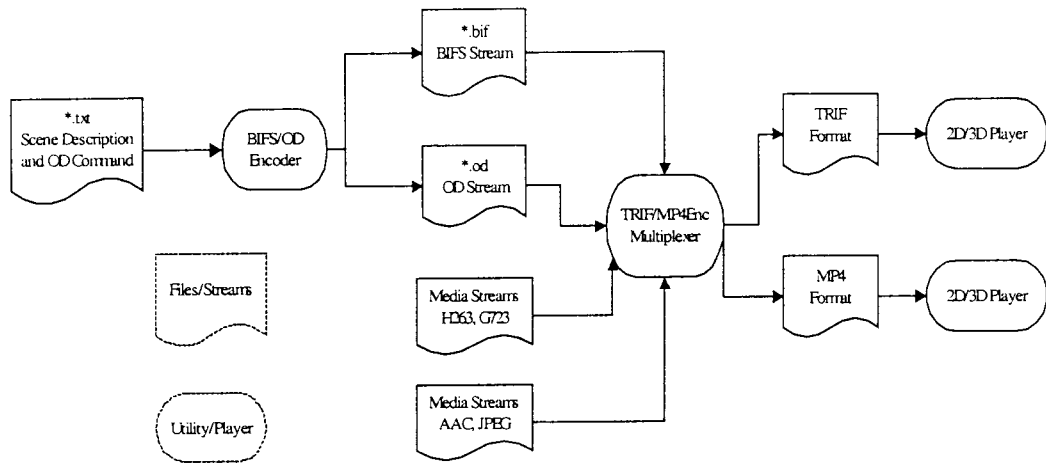
然後將照片中 pixels 依其 motion vectors 切割成不同區域。類似 motion vector 的鄰近 pixels 結合成一個區域，並同時注意其幾何上的地緣關係。同時，我們可以推測物件在幾何空間中的 depth。每一個區域，我們視為一個物件。

依據 3D 立體幾何一些假設，我們可以推測物件的 3D 資料。使用這些資料，在某種程度上可以重建新的角度與拍攝距離的物件。重塑而成新角度影像。重塑而成的例子如圖七。

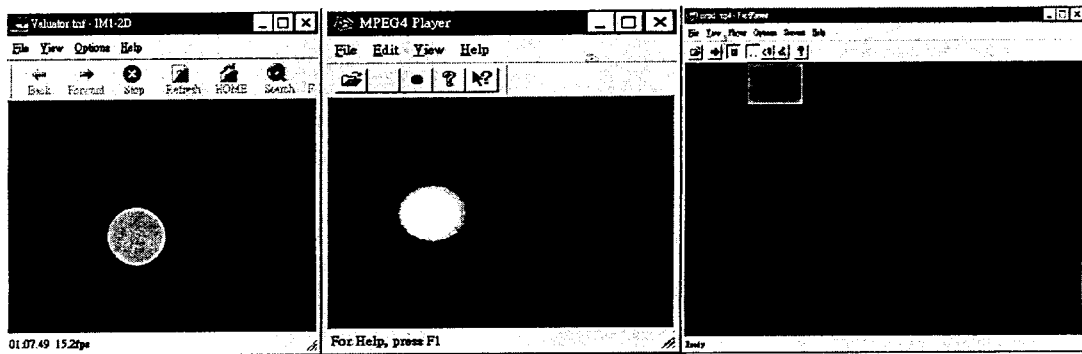
Publication: Y.-C. Huang and H.-M. Hang, "A Study on Interactive Support of MPEG-4," WCE2000, pp.27-32, Oct. 2000.



圖一：MPEG-4 Core 的類別與資料流示意圖



圖二：各編寫工具與播程式之關係



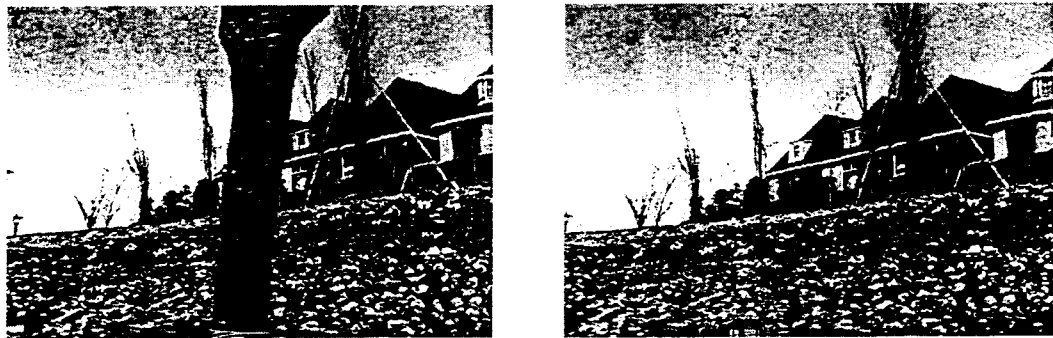
圖三：Valuator

圖四：Spheres

圖五：Change Shapes



圖六：Motion map of the *Flower Garden Sequence* 之一例。



(a)

(b)

圖七：組合不同 Layer 的重建影像