

行政院國家科學委員會專題研究計畫 成果報告

子計畫二：視訊串流伺服器及視訊資料庫整合(I)

計畫類別：整合型計畫

計畫編號：NSC92-2219-E-009-005-

執行期間：92年08月01日至93年07月31日

執行單位：國立交通大學資訊工程學系

計畫主持人：李素瑛

計畫參與人員：陳敦裕 蕭銘和 陳華總 陳漪紋 郭慧冰 陳昱愷 吳惠君

報告類型：完整報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 93 年 10 月 4 日

行政院國家科學委員會專題研究計畫成果報告

視訊串流伺服器及視訊資料庫整合(I)

計畫編號：NSC 92-2219-E-009-005

執行期限：92年8月1日至93年7月31日

主持人：李素瑛教授 交通大學資訊工程系

計畫參與人員：陳敦裕 蕭銘和 陳漪紋 陳華總 郭慧冰 吳惠君

一、中文摘要

本計畫主要目的在於研究與實踐視訊資料庫以及視訊伺服器。在視訊資料庫方面，除了視訊資料內容之分析與高階與語意特徵擷取之外，並結合音訊資料之特徵加上影片內容字幕之偵測及擷取，利用分析所得之語意資訊加以解譯及推測，以偵測影片資料內容之事件，藉此達成高階特徵之擷取並以此為基礎產生既簡短又具語意意義之視訊資料摘要提供視訊資料預覽之用，除此之外在產生影片內容摘要之過程中，同時產生符合 MPEG-7 之低階至高階特徵之描述 (descriptor) 以及描述結構 (description scheme)，以利資料庫索引之建立並藉此建立互動式影片瀏覽系統。

在視訊伺服器方面，我們希望設計支援單點傳送與多點傳送的視訊串流伺服器。在單點傳送視訊串流伺服器方面，主要有兩項任務，一項是傳輸速率控制(rate control)，另一項則是錯誤控制(error control)。在傳輸速率控制中，不僅要根據網路狀況，更要根據視訊資料的重要性來做網路頻寬的最佳使用；在錯誤控制中，傳統的重送機制必須花費較多的時間才能達成錯誤恢復，因此我們希望透過傳輸錯誤更正(forward error correction)機制，並且搭配不平等的錯誤保護機制(Unequal Error Protection)以降低錯誤恢復所需的時間並且減低頻寬的浪費。

在多點傳送視訊串流伺服器方面，由於多點傳送有效率地使用頻寬，因此我們的重心在於錯誤控制。在多點傳送的錯誤控制中，當封包遺失發生時，有兩個主要的問題需要解決，一個是區域性地錯誤恢復(Local

Error Recovery)，一個是 NACK 爆炸(NACK Implosion)的議題，前者是希望能盡量縮短錯誤恢復所需要的時間，後者則避免造成區域性網路壅塞的問題，因此如何達到區域性地錯誤恢復以及避免 NACK 爆炸所造成的區域性網路壅塞便是我們的目標。

關鍵詞：視訊內容分析；音訊內容分析；互動式影片導覽系統；視訊串流；視訊資料庫

Abstract

The purpose of this project is to develop an efficient and effective video-streaming server. In video database, we will analyze visual, audio features and localize captions in order to support content-based video retrieval based on acquired high-level/semantic features. Furthermore, we will detect video events by inferring from analyzed information of video and audio streams. Besides, video summaries can be obtained by merging semantic video clips that are identified according to the video events. The description/descriptors corresponding to low-level and high-level features and the description schemes can be automatically generated during the process of video summaries generation and will be used for indexing.

In video streaming server, we will implement unicast video streaming server and multicast video streaming server. There are two major tasks for unicast video streaming server: one is rate control and the other is error control. In rate control, unicast video streaming server optimizes the usage of the network bandwidth according to not only

the estimated available bandwidth but also the importance of the video stream. In error control, we will adopt a method called FEC (Forward error correction) with UEP (Unequal Error Protection) to reduce the time spent to recover packet loss and waste of the network bandwidth because traditional retransmission mechanism needs to spend more time to complete the packet loss error recovery.

In multicast video streaming server, we will focus on the error control because multicast is an efficient technique to deliver video content over the Internet. In error control, there are two major issues needed to be resolved: one is local error recovery and the other is NACK implosion. The first issue is to recover the packet loss as soon as possible and the second is to avoid local area network congestion. Therefore, in order to realize the multicast video-streaming server, we will investigate in the achievement of local error recovery while avoiding local area network congestion.

Keywords: Content-Based Video and Audio Analysis, Interactive Video Browsing System, Video Streaming, Video Database

二、緣由與目的

隨著網際網路規模的不斷擴大以及資訊科技進步快速，愈來愈多的多媒體資料以數位的形式藉網路為媒介廣泛地傳播著。資訊種類之多元以及資訊量之豐富對於使用者來說哪些是有用的資訊及哪些是可忽略的資料自然成為一個重要的議題。因此對於使用者而言，在眾多資料之中如何有效且快速地獲得真正所要的內容便是我們所要努力的目標。然而除了提供快速有效的查詢方式之外，還必須讓使用者同時享有自然且友善的查詢環境，而利用資料內容為索引及查詢 (Indexing & Retrieval) 為依據的方式 (Content-Based Access) 正能夠滿足此需求。因為所謂 Content-Based Access 即是以資料內容之特徵當成索引來建立資料庫，因此特徵描述之充分與否則直接影響查詢結果之優劣。而傳統資料內容之描述可為某些關鍵

字、屬性的集合、或是某些以文字為主的抽象表示法。

過去以文字為基礎的資料庫系統已發展了一段時間並且已經相當成熟，但相反地以多媒體內容為索引及抽取依據的資料庫管理系統發展卻相當落後。最近將 meta-data 的觀念延伸至其他應用已經引起廣泛的探討，其應用可包含形形色色且較為複雜的資料型態，如壓縮的影像、audio、video 等等。因此對於大量的音樂資料、新聞、紀錄片、電影以及影像提供充分適當的描述對於不同領域及不同應用的使用者而言相當重要。在應用上比如，數位圖書館、隨選視訊以及網路多媒體傳播等等。

因此 MPEG-7 試著定義多媒體特徵的描述方法使得以內容為基礎的資料索引及資料搜尋更為容易。此外 MPEG-7 包含更多的資料型態並且將內容的描述 (descriptor) 和描述法 (description scheme) 標準化藉以達成能夠描述更多種類之多媒體資料。然而特徵產生之方式並非 MPEG-7 欲標準化的部份，除此之外由於多媒體資料型態及內容的多樣性使得更有效、更具高階意義的特徵亟待產生。另一方面為了提供系統間之相互作用 (interoperability)，特徵抽取以及搜尋引擎兩大議題皆非 MPEG-7 欲標準化之部分。

本計畫主要目的為視訊資料內容之分析與擷取，並利用分析所得之資訊加以推理 (inference)，以偵測視訊資料內容之事件，藉此達成高階特徵之擷取並以此為基礎產生既簡短又具語意意義之視訊資料摘要 (summary)，除此之外在產生視訊內容摘要之過程中，同時產生低階至高階特徵之描述 (descriptor) 以及描述結構 (description scheme) 以符合 MPEG-7 之標準。

三、實驗結果與討論

1. 傳輸速率控制模組 (Rate Control Module)

在傳輸速率控制模組中，多媒體伺服器便根據多媒體資料的重要性來做相對應的調整，如此不僅可以充分使用現有的網路頻寬，更可以好好利用接收端有限且受限制的緩衝區來達到多媒體品質的最佳化。此

模組之輸入包含多媒體資料位元流以及估量的網路頻寬，輸出為經過調整的多媒體資料位元流，核心部分則包含資料傳輸率配置演算法 (Bit-rate Allocation Algorithm)、分層丟棄策略(Layer Dropping Policies)、虛擬的 FS 緩衝區模型(Virtual Forward-shifting Buffer Model)以及洞悉多媒體資料內容的傳輸速率控制器(Content-aware Rate Controller)

2. 傳輸錯誤更正模組 (Forward Error Correction Module)

錯誤更正模組將輸入的封包依其資料的重要性來決定是否要輸出重複的封包，例如，輸入的封包中，倘若其屬於基本層的資料封包，則其會被複製一份並且輸出；倘若其屬於加強層的資料封包，則其並不會有被複製的動作，而僅僅只是被輸出，因此，由於我們只對基本層做封包保護的動作，所以可以降低頻寬的花費，再者，上述的做法屬於傳輸層次，所以也不會造成編碼端以及解碼端的延遲，最後，即使網路狀況不佳，導致封包遺失率提高，但接收端依然可以根據收到資料的多寡，來做解碼的動作。

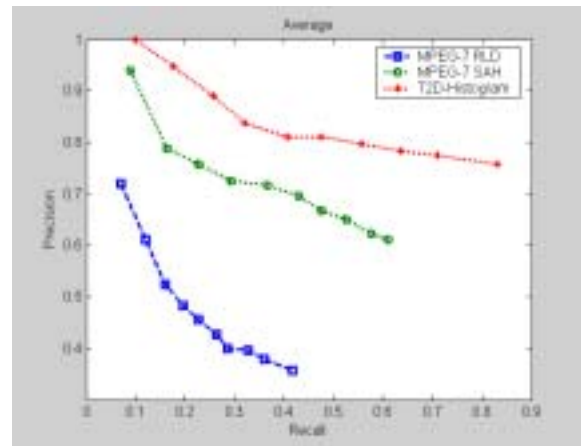
3. 網路模組 (Network Module)

在網路模組中，我們使用的是 RTP (Real-Time Transport Protocol)/RTCP(RTP control protocol)，RTP 被用來傳輸多媒體資料，RTCP 則被用來監視資料傳輸的狀況，要注意的是，RTP/RTCP 本身並不提供 QoS 的能力，也就是它不保證資料能正確無誤的被接收，也不保證資料能依序的到達，但是它提供 QoS 相關的資訊，在 RTP 的標頭欄位，其中的序列號碼(Sequence Number)可以用來重組接收封包的順序，也可以用來做封包遺失的偵測；時間戳記可以被用來做多媒體資料的同步；另外，RTCP 則會週期性地傳送控制封包，讓通訊的雙方可以知道彼此接收資料的品質狀況。

4. 視覺資料特徵擷取

我們設計了一個以移動物體為基礎之高階視訊內容描述方法 (object-based T2D-histogram descriptor)，此描述方法不僅精簡且在協助視訊片段搜尋上更是有效率，為衡量此描述方法之效能，我們與 MPEG-7 之兩個 motion activity

descriptors, Run-Length Descriptor (RLD)以及 Shot Activity Histogram (SAH) 比較,如圖一所示:



圖一、視訊擷取效能比較

由實驗結果可知，在視訊擷取的準確性方面，我們所提出的 T2D- histogram 比 RLD 好約為 40%，另外亦比 SAH 好將近 20%。

