

# 行政院國家科學委員會專題研究計畫成果報告

## 智慧型多媒體資訊處理系統的研究(一)

### The Study of Intelligent Multimedia Information Processing System(I)

計畫編號：NSC 88-2213-E-009-065

執行期限：87年8月1日至88年7月31日

主持人：傅心家 交通大學資訊工程系

#### 一、中文摘要

隨著電腦速度的提昇以及網路的普及，如何在浩瀚的網路多媒體資訊中快速及正確找尋想要的多媒體資訊，已是一個非常重要的課題。目前已有的影像搜尋系統均著重於如何有效的在已知資料庫中搜尋靜態影像，其比對之特徵均以整個影像為基礎。而對於環球網際網路(WWW)上的動態影像的搜尋目前少有成熟的研究出現。本計畫提出了一個以類神經網路為基礎的智慧型多媒體資訊處理系統(Intelligent Multimedia Information Processing Systems, IMIPS)。

關鍵詞：智慧型多媒體資訊處理系統、類神經網路、動態影像搜尋引擎

#### Abstract

Retrieving multimedia information efficiently among the world wide web becomes more and more demand. Many images search engines have been developed, so far their retrieved image were based on the image database constructed by themselves, and the developed of video system are seldom to be announced. In this project, We intend to design an intelligent multimedia information processing system (IMIPS) based on neural networks.

**Keywords:** IMIPS, neural networks, video search engine

#### 二、緣由與目的

隨著電腦速度的提昇以及網路的普及，如何在浩瀚的網路多媒體資訊中快速及正確找尋想要的資訊，已是一個非常重要的課題。對使用者來說多媒體資訊比傳統資訊更豐富，更多元化。但對於電腦而言如何存取搜尋卻是個艱深的任務。因此如何使電腦認識這些多媒體資訊成為當今所需解決的問題。由於對於資料表現方法的不同，多媒體資訊處理有著不同於一般傳統資訊處理方式。傳統的文字資訊處理方式已不能滿足多媒體所含的豐富資訊。

隨著網路速度的提升及記憶體單位成本的下降，動態影像資訊預期將快速增加(例如：電影，影集...等)。如何更有效率且更方便的對WWW上的視訊做搜尋將是一個很重要的課題。目前以靜態影像為基礎的處理方式已無法有效滿足使用者對視訊query的要求，因為視訊除了靜態影像外，還包含了動態影像。而動態影像除了有位置空間的關係外，尚還有時間的關係。所以必須針對動態影像的特性建立一方法架構，以滿足使用者對視訊query的要求。基於這些要求，本計畫提出了一個以類神經網路為基礎的智慧型多媒體資訊處理系統(Intelligent Multimedia Information Processing Systems, IMIPS)。

#### 三、結果與討論

本年度我們完成了以網際網路為搜尋資料庫的智慧型多媒體資訊處理系統[5]。茲就此系統的組織架構及工作原理詳述如下。

### (一) 組織架構

IMIPS 的組織架構如下圖所示。我們將 IMIPS 分成以下主要的六個部分，分別為

- 服務者 (Server) 端部分
  1. Video Spider
  2. Keyframes Selector
  3. Feature Extractor
  4. DataBase Management System
- 顧客(Client) 端部分
  5. User Interface
  6. Query Processor

現就此六部分的工作原理敘述如下。

### (二) 工作原理

**Intelligent Video Spider** 主要的工作是在 WWW 上搜集視訊檔案。為了能確實反應 WWW 上視訊資料的變動，Spider 必須定時在 WWW 搜尋視訊資料。如何有效率的在 WWW 上搜尋是這部分所應探討的地方。本系統的 Spider 主要分成二個部份 [1]

#### 1. travel-spider

#### 2. maintain-spider

其中 **travel-spider** 負責尋找網路上不同網頁的位址 (URL)；**maintain-spider** 則負責監督已找到的網頁，如果某個網頁的內容有變動過，要重新對該網頁進行分析，以確保資料庫的資料不會過時。

**Keyframes Selector** 主要的工作是將 Spider 所搜集的視訊做前處理。由於視訊的資料量比起靜態影像要大的多，因此對於視訊，我們必須做一些處理，以利後面的運作。我們將抽取一些代表性的畫面 (keyframe) 來代表視訊，以減少儲存的空間及往後做特徵抽取所花費的速度。我們可將視訊經由 MPEG 壓縮來擷取代表性的畫面。由於視訊經 MPEG 壓縮後，不僅本身的資料量小很多，且其壓縮後的 I-frame 正可用做代表性的畫面。除此之外，我們還可比較 I-frame 間的差異來選擇出代表性的畫面，進一步的使資料量減至最小。由於現存於 WWW 上的視訊多為 MPEG 檔，因此我們代表性的畫面的抽取方式非常的適用於這個環境。且若視訊不為 MPEG 檔，我們仍可用已有的壓縮軟體將其做成 MPEG 檔 [2]。

**Feature Extractor** 為組成 IMIPS 重要的部份之一。其工作是將視訊暫存器中的視訊特徵抽取出來，若 Feature Extractor 無法很正確的將視訊的特徵抽取出來，則整個系統的命中率一定不高。我們的方法是先將整張靜態影像切割成許多方塊，再從這些方塊組合中找出顏色的特徵。我們再根據這些特徵用 Multiple-Instance 的觀念用 EM 的方法訓練出樣本代表的特徵 [3]。

**Database** 主要是將分析好的資料作有效的管理。在這個系統中，從抓取資料到分析資料，都會產生大量資訊，而這些資訊必須記錄下來，作為後續工作的依據，如果沒有好的方法來管理這些資訊，將來在存取時，會成為系統的負擔，為了減輕

系統負擔，我們利用資料庫的特性，對這些需要儲存的資訊作有效的管理，並依照不同的需要及用途，每個類別使用一個獨立的 table，以利將來查詢時，能夠快速的檢索出所需的資料[1]。

**User Interface** 主要是 IMIPS 與使用者溝通的橋樑。其設計的重點在於如何讓使用者能很容易的與 IMIPS 溝通，所以友善的界面是 User Interface 設計的目標。基於這個目標，我們擬對每個可 query 的不同特徵建立所屬的樣本資料庫[3]，使用者可在樣本資料庫中點選適合的樣本做 query。

**Query Processor** 則是將 User Interface 所送進來的 query 經過一些處理，抽取 query 的特徵。然後送進服務者端的 Similarity Comparator 做比較[3]。

此外，為了加強 IMIPS 在資訊處理方面的能力，我們以 Java Applet 建構一個類似視窗環境下『所見即所得』的編輯環境，提供媒體編輯者簡易的剪接、旁白、字幕等後製作功能。使用者可在此環境下對於搜尋過程中得到的短片加上註解及相關索引，供未來其他使用者利用；或是透過此系統將 IMIPS 搜尋所得的媒體作進一步的編輯利用[4]。

#### 四、計畫成果自評

本年度為此計畫的第一年。在本年度中，我們實作了一個以動態視訊為搜尋來源的搜尋引擎雛型[5]。在現今的網路頻寬下，視訊的資料並不像文字資料那樣普及，而且目前網路上的視訊資料絕大多數都不會很龐大，但科技進步的腳步日新月異，在不久的將來，網路上的視訊資料，會和文字資料一樣發達，然而，文字的搜尋引擎已不虞匱乏，當視訊和文字一樣普及時，也需要搜尋引擎來幫助使用者在浩瀚的網路資源中，快速找到及取得想要的

資訊。

本計畫著重在有效的取得視訊資料，正確的分析處理視訊資料，以及便利的搜尋環境界面。不管是那一類型的搜尋引擎，資料的蒐集是無法避免的，但不同目的的搜尋引擎，蒐集的資料也不相同。在目前的搜尋引擎，絕大部份是以文字為主，視訊的搜尋引擎尚屬少見，即使有，也很少能像文字搜尋引擎般的令人滿意，而我們的目標是希望使用者在使用過我們的系統後，都能得到想要的結果，為了這個目標，我們分析影片內容，得到的處理結果，就好像有人親自看過影片，並對影片作了評論。希望在不久的未來，視訊資料大量流通於網路上時，透過本系統的幫忙，讓使用者得到他想要的資訊。

#### 五、參考文獻

- [1] 陳偉修, "*The study of intelligent spider for video search engine*", 國立交通大學資訊工程研究所碩士論文, June 1999。
- [2] 李信緯, "*The Study of Scene Change Detection and Moving Object Detection in MPEG Video*", 國立交通大學資訊科學研究所碩士論文, June 1999。
- [3] 黃立吾, "*Image classification and indexing by Multiple-Instance Learning*", 國立交通大學資訊工程研究所碩士論文, June 1999。
- [4] 王斌銳, "*A SMIL-based Authoring Environment for Internet Video*", 國立交通大學資訊工程研究所碩士論文, June 1999。
- [5] The Intelligent Multimedia Information Processing System (IMIPS), "<http://www.csie.nctu.edu.tw/~mmlab/Demo/IMIPS>"。