

行政院國家科學委員會專題研究計畫 期中進度報告

子計畫二：智慧型視訊處理及硬體架構設計(2/3)

計畫類別：整合型計畫

計畫編號：NSC92-2219-E-002-017-

執行期間：92年08月01日至93年07月31日

執行單位：國立臺灣大學電子工程學研究所

計畫主持人：陳良基

計畫參與人員：陳良基

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 6 月 1 日

智慧型音視訊技術及多媒體應用—子計劃二

智慧型視訊處理及硬體架構設計(2/3)

Intelligent Video Signal Processing and Hardware Architecture Design

計劃編號：92-2219-E-002-017

執行期限：92/08/01 93/07/31

子計劃主持人：陳良基 教授 Email: lgchen@cc.ee.ntu.edu.tw

執行機構：國立台灣大學電子工程學研究所

一、中文摘要

多媒體技術的演進，已經由原本以壓縮技術為重心，逐漸轉為朝向傳輸、互動、資料庫，及智財管理及保護等多方面發展，而多媒體資料的處理，也由以往傳統的以波形為基礎的處理轉變成以多媒體內容為基礎的智慧型信號處理。本計畫將提出一完整的智慧型視訊壓縮及傳送系統，其內容包括了智慧型的前處理、智慧型的視訊編碼、智慧型的通訊技術，以及能同時兼顧高運算能力和彈性的硬體架構設計，期能進而發展出一個適合智慧型處理的硬體發展平台，從而提供一個從演算法到硬體架構的整合解決方案。

ABSTRACT

Because of the development of large scale storage devices and broadband communication, the technology trend of multimedia signal processing is from compress technology to transmission, interactivity, database, and intellectual property management and protection, and from waveform based signal processing to intelligent signal processing. This trend is also the target of this project. In this project, a complete intelligent video compression and transmission system will be developed. It includes four main parts: intelligent preprocessing, intelligent video coding, intelligent video transmission, and platform-based hardware architecture of video coding system. The new techniques developed in this project can be the key and base technology of the future intelligent

multimedia processing.

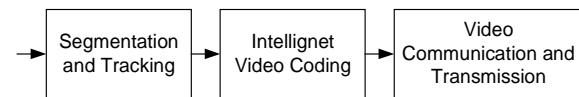
二、緣由與目的

多媒體信號的編碼，已由早期以波形為基礎的編碼，逐漸轉向以物件或是以模型為基礎的編碼方式。而多媒體信號的處理，也由儲存逐漸轉向多功能性方向、傳輸、互動、資料庫、或是智財權管理及保護的方向。也就是從傳統的壓縮技術不管多媒體內容的處理方式，轉變成為先認知多媒體內容進而加以處理的方式，也就是所謂智慧型的處理，其中包括了智慧型的編碼、資料存取、描述、保護，以提供使用者更多的多媒體功能。而本研究計畫的目的，就是建築在以往對於數位視訊技術的演算法及硬體架構設計的基礎上，對於智慧型多媒體處理中智慧型視訊技術的演算法及硬體實現做深入的研究。本計畫的重點，將放在視訊物件的前處理、編碼、傳輸、及描述等較為前段的部分，也是在智慧型視訊信號處理中，最重視即時性也最為重要的基石。

三、研究方法與成果

本計畫將針對智慧型視訊處理中最為重要也是最基礎的智慧型編碼和傳輸發展其相關演算法及硬體架構。一個智慧型視訊壓縮及傳送系統方塊圖如下圖所示。在攝影機拍下視訊資料後，視訊物件前處理單元會做切割及追蹤的動作，把視訊物件切割出來，以利後級做以物件為基礎的智

慧型處理，包括了智慧型的編碼，以及利用在編碼過程中獲得的相關物件資料做描述，最後再把視訊物件傳送出去，而視訊物件切割與追蹤、智慧型的視訊編碼及描述、視訊物件的傳輸的相關演算法及硬體架構設計就是本計畫即將要研究的課題。



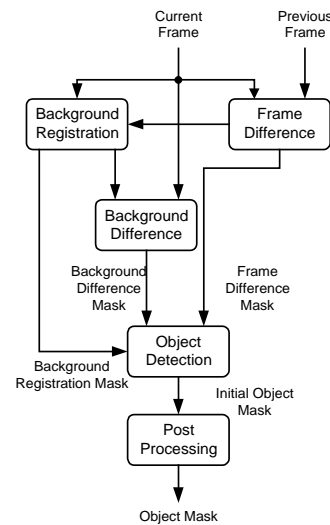
圖一、智慧型視訊壓縮及傳送系統方塊圖

此計畫所涵蓋的範圍，包括前級處理的視訊切割與追蹤、使用場景精靈和二維動態網狀之智慧型的編碼、適用於 MPEG-4 的硬體平台架構設計（以 simple profile 和 FGS profile 做為目標應用）、智慧型的錯誤隱藏技術，以下將分別對各項技術詳細說明研究方法及進行成果。

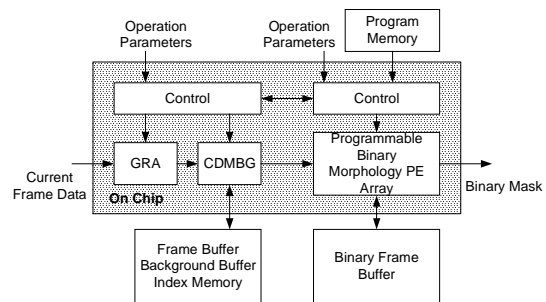
A. 智慧型視訊切割

我們在去年提出了智慧型視訊切割演算法，其方塊流程圖如圖一。這套演算法是以改變偵測(change detection)及背景登記(background registration)為基礎的演算法，其主要的構想就是不同於以往改變偵測找尋畫面中「動」的部分，此方法找尋的是畫面中「不動」的部分，也就是背景的部分，再利用背景的資訊來幫助我們把前景物件切割出來。另外此系統提供另外兩種模式：影子消除模式、全域移動估計模式；影子消除模式則是在輸入此系統前，畫面會先以形態學梯度運算的技術消除畫面中光影的資訊，只留下材質的資訊，如此一來，物件切割就不會再受到影子及光影的影響了。而藉由全域移動估計模式可以使得即使攝影機在晃動下，我們仍然具有很好的切割效果。根據這個演算法，我們提出了相對應的硬體架構如圖三。其中 CDMBA 為主要運算核心、Programmable Binary Morphology PE Array 則負責型態學和後置處理的相關運

算，GRA 則是負責影子消除模式的運算。



圖二、視訊切割系統



圖三、已提出視訊切割演算法之硬體架構

B. 智慧型視訊追蹤

在第一年的計畫之中，我們已對目前已提出的視訊物件追蹤的技術做了全面性的研究，而在本年度的計畫中，則是提出如圖四的演算法流程。我們使用空間上的資訊來做追蹤的演算法是比較有效率的。而視訊物件追蹤在追蹤不同型態的物體時會有極大的差別，我們所要追蹤的視訊物件是人，而膚色是人的重要特徵，因此演算法中針對膚色的部分進行分析。利用在現存的研究中，膚色在 YCbCr 轉換後之 Cb-Cr 空間中有一個特殊的分布來做判斷條件。因此我們所提出的演算法有非常好的效果，可以準確的追蹤到人的位置。



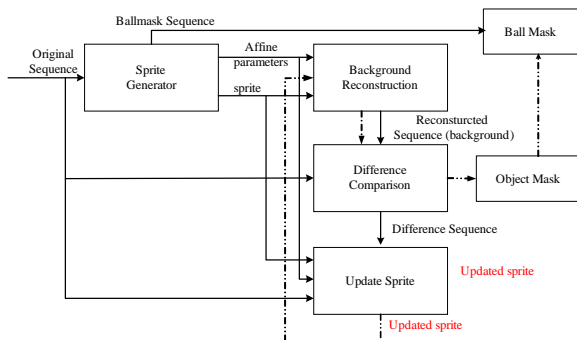
圖四、物體追蹤演算法之流程圖

C. 場景精靈編碼

Sprite 的產生是 sprite 編碼系統中最為困難的一部分，藉由消除連續影像間相對的運動，便可利用拼貼的技巧將攝影過程中拍攝到的背景物件加以組成一個整體的場景稱為 sprite，如圖四所示。



圖五、背景物件 sprite

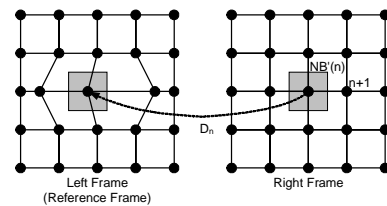


圖六、sprite 產生之系統方塊圖

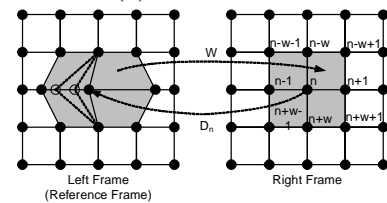
在第一年中，我們以軟體實現了一個完整的編碼系統，圖五則是我們今年提出的一個更完整的系統詳細流程方塊圖，他不需要我們先提供去除前景的影像片段、換句話說它據有自行處理前後景分離的演算法，使得我們可以從一個影片片段中組出一個連續場景，進而可以切割前景物件、進行編碼處理的一系列動作。

D. 二維動態網編碼

在 MPEG-4 的混合輪廓(hybrid profile)中所提供的二維動態網(2-D dynamic mesh)提供了一個智慧型前景物件編碼的新想法。二維動態網編碼能提供更細緻、更高維度、更複雜的移動補償，在某些情況下，甚至可以不用做補償剩餘材質編碼(residual texture coding)就可以達到人眼可以接受的視覺效果，而且二維動態網能提供視訊物件更高的可編輯性及內差能力，也能提供視訊物件更為細緻的動態描述。在這一年中，我們對目前二維動態網的演算法做分析與研究，並且也找出不同演算法的優缺點，相信在接下來的一年中接下來就可以提出我們自己的演算法架構。



(a) 第一階段



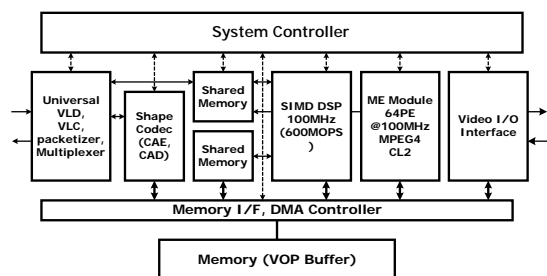
(b) 第二階段

圖七、兩階段之二維動態網編碼

E. 適用於 MPEG-4 的硬體平台架構設計

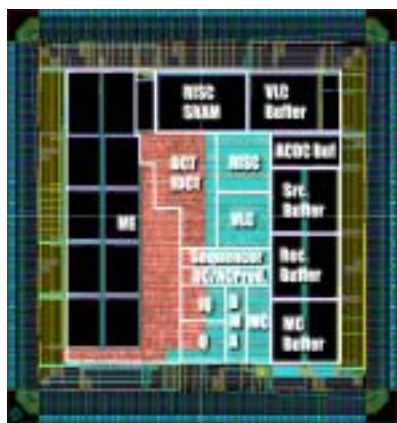
根據我們之前的分析要完成一個

MPEG-4 的系統，所需的硬體架構必須是混合式的。因此我們提出了一個混合式的 MPEG-4 硬體架構，整體架構如圖六。系統中有一個微處理器、一些陣列處理器，其硬體架構如下圖所示：



圖八、預計 MPEG-4 硬體架構

為能支援 MPEG-4 標準的多變性，此硬體中必須要有可程式化的部分，此外，隨著新功能的加入，可能會需要一些必要的加速硬體模組，此系統必須要有開放的架構並有能讓這些模組加入的介面，而此種硬體的設計方式，正是所謂平台式設計 (platform based design) 的概念。整個晶片實現結果如圖九，詳細規格則列於圖十。本晶片其運算處理能力為 MPEG-4 Simple Profile @ level 3, CIF Format, 30 fps。其面積為 5.1 mm x 5.1 mm，採用的是 TSMC 0.35 um 製程，功率消耗為 339.51mW，工作頻率 40MHz 就可以即時運算 MPEG-4 SP@L3。



圖九、實作的 MPEG-4 晶片

Chip	MPEG-4 Video Encoder
Specification	Simple profile Level 3
Encoding Complexity	352 x 288 at 30 fps
Technology	TSMC 0.35 um 1P4M
Die Size	5.1 x 5.1 mm ²
On-chip memory	39,080 bits
Off-chip memory	2,027,527 bits
Transistor count	828692 trans.
Working frequency	40 MHz
Voltage	3.3V
Power Consumption	339.51mW

圖十、MPEG-4 晶片規格

四、結論

本子計劃已達成第二年之預定目標，在各方面都有相當不錯的收穫和成果。在接下來的年度中，將繼續從事第三年之預定目標，期泛用型 object segmentation 之積體電路硬體實作。並且在各方面提出具有創新性和效果較佳的演算法和架構設計，同時進行計畫第三年中的相關系統整合、相關硬體實作部分。

五、參考文獻

- [1] Jim Brailean, Weiping Li, Jorn Ostermann, Ya-Quin Zhang, "Chapter 3 MPEG-4 Video Coding Standard", in "Circuits and Systems in the Information Age"
- [2] Thomas Sikora, "The MPEG-4 Video Standard Verification Model", IEEE Transactions on CSVT, vol 7, No.1, Feb 1997
- [3] ITU-T Recommendation T.4, "Standardisation of Group 3 Facsimile Apparatus for Document Transmission."
- [4] Aljoscha Smolic, Thomas Sikora, and Jens-Rainer Ohm, "Long-term global motion estimation and its application for sprite coding, content description, and segmentation," *IEEE Tran. on Circuit and Systems for Video Technology*, vol. 9, no. 8, pp.1227-1242, December 1999.
- [5] D. Zhang and G. Lu, "Segmentation of moving objects in image sequence: a review," *Circuits Systems Signal Processing*, vol. 20, no. 3, pp. 143-183, 2001.
- [6] M. Kim, J. G. Choi, H. Lee D. Kim, M. H. Lee, C. Ahn, and Y.-S. Ho, "A VOP generation tool: Automatic segmentation of moving objects in image sequences based on spatio-temporal information," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 8, pp.

1216–1226, Dec. 1999.