

行政院國家科學委員會專題研究計畫 成果報告

MPEG-7 影像索引及辨識處理硬體架構設計

計畫類別：個別型計畫

計畫編號：NSC93-2622-E-002-010-CC3

執行期間：93年05月01日至94年04月30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，黃毓文，張靖瑩，張育璋

報告類型：精簡報告

處理方式：本計畫為提升產業技術及人才培育研究計畫，不提供公開查詢

中 華 民 國 94 年 8 月 1 日

國科會補助提升產業技術及人才培育研究計畫成果精簡報告

學門領域：微電子學門工程處 (E6)

計畫名稱：MPEG-7 影像索引及辨識處理硬體架構設計

計畫編號：NSC92-2622-E-002-010-CC3

執行期間：民國 93 年 05 月 01 日起至民國 94 年 04 月 30 日

執行單位：國立台灣大學電子工程學研究所

主持人：陳良基

參與學生：

姓名	年級 (大學部、碩士班、博士班)	已發表論文或已申請之專利 (含大學部專題研究論文、碩士論文)	工作內容
黃毓文	博士班		演算法之分析設計， 模組架構設計，程式 設計
張靖瑩	博士班	Architecture and Analysis of Color Structure Descriptor for Real-Time Video Indexing and Retrieval (PCM2004) Architecture of MPEG-7 Color Structure Description Generator for Real-time Video Applications (ICIP2004) Architecture and Analysis of Color Structure and Scalable Color Descriptor for Real-Time Video Indexing and Retrieval (ISCE2004)	演算法之分析設計， 模組架構設計，程式 設計
張育璋	博士班		模組架構設計，程式 設計

合作企業簡介

合作企業名稱：晶睿通訊

計畫聯絡人：馬仕毅

資本額：20000 萬元

產品簡介：MPEG-4 網路攝影機、MPEG-4 迴轉式網路攝影機、MPEG-4/MJPEG 光學變焦迴轉式網路攝影機

網址：<http://www.vivotek.com>

電話：02-82455282

研究摘要(500 字以內)：

隨著數位影像科技的快速發展，一般大眾自行製作數位多媒體的速度也大幅提升。越來越多的媒體檔案產生，代表了未來在搜尋特定視訊檔案也更加困難。因此 MPEG-7 對於聲音、影像、影片訂定了許多用來判別不同特質的方法，以便在未來提供大家快速搜尋的服務。在靜態影像方面，包含了以色彩、材質、形狀為基礎的描述子。在動態視訊方面，包含了以物體位置和移動估計為基礎的描述子。另外，在人臉辨識這個重要的技術，也有獨立的描述子。要產生這些描述子所要付出的代價即為非常高的運算量，若能在產生視訊資料的同時，同時獲得以上的描述資料，將可以大幅降低後製處理的時間。如何針對每一種描述子的特性設計最佳模組，進而共用硬體降低成本，構思高效能的架構，是我們希望瞭解的課題。

人才培育成果說明：

本計畫產生三篇國外論文與一顆晶片，不僅學到演算法分析與架構設計考量的技巧，而且讓學生能出國與世界的研究人員討論最先進的想法與創意。另外還激發研究生思考一個國際標準從未想過的可能應用。

技術研發成果說明：

從本計畫中，在軟體方面，我們利用原先用於搜尋與擷取的 MPEG-7 標準中的色彩結構描述子(Color Structure Descriptor)和可縮放色彩描述子(Scalable Color Descriptor)，將它們運用在監視系統中對物件分割以及骨架化上。在硬體方面，我們分析了這兩個描述子的運算形態，以及在達到即時產生時所需的運算量，了解哪個部分所占的運算時間、運算量和記憶體搬運動作最龐大，繼而推導出適合的平行演算法。我們提出了第一個色彩結構描述子和可縮放色彩描述子加速器的硬體架構。它能即時產生每秒 30 張 256x256 畫面大小的兩種描述子。這套硬體架構與 2.54 GHz 的微處理器平台相較，約可

提供 12 倍的運算加速，以滿足有即時運算需求的應用。這兩個描述子的共同硬體已由國家晶片系統設計中心(CIC)梯次 U18-93B 下線，我們以 Verilog HDL 實現設計，並遵循 CIC 所提供的 cell-based 設計流程，經 Synopsys Design Compiler 以 Avanti 0.18um cell library 合成、和 Apollo 擺設繞線後，設計可通過 Verilog XL 模擬。最後由聯電以 0.18um CMOS 1P6M 製程製作我們的晶片。這顆晶片經由測式機台所測得的結果顯示，操作在 25MHz 和 1.8V 下，功能運算完全正常，消耗的功率為 36.6mW。

技術特點說明：

這兩個描述子之所以能以單一晶片來實現，是因為它們兩個最主要的差別是在一個 Haar 濾波器，其他部分經由控制電路來完成功能選擇進而達到共用的目的。

色彩結構描述子的概念是從傳統的色票統計值 (Color Histogram) 演化而來，差別在於它利用了一種結構框的做法，將小塊區域內出現過的顏色，記錄在統計表中。兩張色系類似的影像，色彩分佈散亂和分佈集中的兩種狀況，描述子的結果就會明顯反應出來。由於每張來源影像的大小並不固定，為了解決影像相同但是為不同解析的現象，MPEG-7 規定所有影像在進到色彩結構描述子產生器之前，都必須縮放(Scale)至 256x256 像素大小以下，之後才開始色彩結構描述子的計算。整體硬體包含四大部分：色彩座標轉換、結構框色票統計、全域色票統計、以及非線性量化色票。

-色彩座標轉換

色彩座標轉換是由傳統紅-綠-藍維度(R-G-B)轉換成彩度-最大值-最小值-差值維度 (Hue - Max - Min-Difference)。其中在彩度的運算方面需要一個乘法和除法。最後所需的彩度精確位元數只要四位 (4 bits)，若是直接將這個乘法和除法轉換成硬體，則除了需要大量的邏輯閘數(Logic Gate Counts)以及高位的精確度，運算時間也會成為瓶頸(Bottleneck)。於是根據彩度的條件，比較紅、綠、藍三值的大小分成六個群組，彩度的計算可以用查表的方式獲得，不僅所占面積較小，運算時間也短非常多。根據實測結果，不儘大幅降低關鍵運算路徑時間，所需邏輯匣數也降為原先由除法和乘法實作的 40%。

-結構框色票統計

色彩結構描述子有別於傳統色票統計的地方是在它會去計算影像結構的完整性，或是說影像結構的複雜程度，而統計出不同的色票結果。但是由於這個計算會使得每個像素都要被重複讀取六十四次，這造成了龐大的記憶體存取。由於原本的演算法所需要的頻寬太大，所以為了實現它，概念上先利用一塊內部記憶體，記錄一塊結構框裡所有的 HMMD 值，再利用另一塊內部記憶體，記錄這個結構框的色彩統計圖。有了結構框的色彩統計圖之後，要再取得下一結構框的統計圖時，就只要多讀一排像素。由於利用了內部緩衝區(Buffer)的做法，每掃描一新的結構框就不必重複讀取該結構框中所有的像素顏色，再利用平行概念，一次處理多塊相鄰的結構框，這樣能使得外部的記憶體頻寬可以大幅下降。要平行多少塊相同的架構，取決於所需的運算速度和硬體空間的平衡。在此選擇平行三倍，一來是面積考量，二來是考慮合理運算週期數，三來是為了與可縮放色彩描述子所需要的內部記憶體用量相容。

利用到統計表的內部記憶體都必須面臨一個問題，在更新一個色票，都需要讀寫各一個週期，而且為了在兩個週期內完成更新，無法同時在內部記憶體的輸入與輸出各擋一個暫存器，以提高測試含蓋率(Fault Coverage)。如果兩邊都擋暫存器，會多一個氣泡週期(Bubble Cycle)，讓原先的硬體利用率(Utilization)從 50%再下降到 33%。這裡並不能採用雙接口記憶體(Dual Port SRAM)，因為在一個週期內讀寫的位置可能相同。如果只在記憶體的一邊擋暫存器，並不用擔心訊號的穩定時間(Hold Time)不夠，因為在更新統計表的值的運算只有一個位元很短的加法器。

-全域色票統計

由於色彩結構描述子是每掃描完結構框的一行後，有新的結構框的值產生。在更新後段的色彩結構描述子的 256 種色票統計值的時間，就必須在完成一個結構框掃描時間內做完。且由於平行三倍的硬體，相鄰三個結構框要讀進 10x8 個像素(每個結構框有 8x8 個像素)，掃完一行的時間為 10x4 個週期(10 像素 x 4(週期/像素))，256 種色票統計值必須在 40 個週期內完成更新。在使用內部記憶體一讀一寫需要兩個週期，所以安排四塊 16x64 (address x wordlength) 的內部記憶體，以提供每兩週期同時處理 16 種色票的頻寬。在此一樣有無法在記憶體兩端同時擋暫存器的問題。

-非線性量化色票

在完成所有的色票統計表之後，最後一步是量化輸出。每個色票統計值最後都會被量化成 0 到 255 的值。由於輸入影像大小是定值，所以原來在標準化 (Normalization) 所需的除法器，可以用 255 個 16 位元比較器來解決。255 個比較可以在一個週期內完成一個色票描述子，但是這樣的比較器實在太多，由於可用的週期非常充裕，所以改採用摺疊架構，每個週期都只與合理範圍內的中間值相比，八個週期輸出一個色票描述子，共需時 256x8 個週期。

可利用之產業及可開發之產品：

MPEG-7 主要是制定用來描述多媒體內容的標準，以便使用者可以比之前使用的文字描述搜尋法，更方便且更有效率地搜尋、瀏覽、擷取所需的多媒體。這樣的應用可適用於大眾廣播業、數位典藏、家庭娛樂等等。除了用在搜尋功能外，它也可用在影像的前處理，像自動白平衡、色彩增益控制、飽合度調整、自動對比等等，在影像壓縮上，可由統計結果產生機率模型來估計壓縮率。當然，更可以分析影像特性而用在監視系統保全上。

推廣及運用的價值：

有了每個檔案所具有的多媒體特質，我們可以利用這些工具，應用在以下服務：

數位資料庫(影像、音樂字典、生醫影像分類……)

多媒體編輯(個人化電子新聞服務、媒體編寫……)

文化保存服務(歷史博物館、藝術畫廊……)

多媒體字典服務(黃頁、旅遊指南、地理資訊系統……)

廣播媒體選擇(收音機、電視選臺……)

新聞業(搜尋特定政治人物……)

電子商務(個人化購物喜好……)

監視系統(交通控制……)

家庭娛樂(個人化多媒體收藏系統、家庭錄影帶編輯、卡拉 OK)