

行政院國家科學委員會專題研究計畫成果報告

互動式多媒體服務之先進視訊編碼技術研究(三)

Advanced Video Coding Techniques for Interactive Multimedia (III)

計畫編號：NSC 88-2218-E-009-027

執行期限：87年8月1日至88年7月31日

主持人：林大衛 交通大學電子工程系所 教授

一、摘要

我們研究網路視訊編解碼技術，含視訊壓縮編碼方法與真時視訊編解碼二方面。在視訊壓縮編碼法方面，我們提出一個變率傳輸下之近最佳率差編碼的理論及演算法。此外又提出一個多視訊共同率差編碼的途徑；模擬結果顯示，較之不考慮視訊率差特性的一般方法，此途徑可獲致相當顯著的品質增益。在真時視訊編解碼方面，我們使用一個 TI TMS320C62xx DSP 的發展系統，實現了一組符合 H.263 標準的視訊編解碼程式。除作視訊編解碼器之軟體實現技術之探討外，並可為後續網路視訊通訊研究之用。

關鍵詞：率差理論、視訊傳輸率控制、多視訊共同編碼、軟體視訊編解碼器

Abstract

We research into network video coding technology, including video compression techniques and real-time video coding and decoding. Concerning video compression techniques, we obtain near-optimal rate-distortion theory and computational algorithm for variable-bit-rate transmission. We also obtain an approach to joint rate-distortion coding of multiple videos, for which the simulation results show that the approach can obtain significant performance gain in comparison to one which does not pay such regard to the rate-distortion characteristics of the videos. Concerning real-time video coding and decoding, we employ a TI TMS320C62xx DSP development system to implement a set of H.263 video coding and decoding programs. Aside from investigation into software realization techniques of video codecs, the results are applicable to subsequent research in network video communication.

Keywords: Rate-Distortion Theory, Video Rate Control, Joint Multiple-Video Coding, Software Video Codecs

二、計畫緣由與目的

本計畫研究網路視訊編解碼技術。一方面探討視訊壓縮編碼的方法，另一方面探討真時視訊編解碼的實現。

在視訊壓縮編碼的方法方面，現有之一般國際數位視訊壓縮標準，均設有一些編碼參數可供編碼器自行選擇使用。因此，如何選擇適當的參數組合，以達優良的視訊品質，就成為一個很重要的題目。我們於近幾年來，已從率差編碼(rate-distortion coding)的角度，進行了一系列的研究。本研究考慮在變率傳輸(variable-rate transmission)下之最佳率差編碼理論與演算法，以及多個視訊共同傳輸時之率差編碼法。前者之應用領域，包括在 ATM 或 Internet 等變率通道上的傳輸。後者之應用領域，則包括衛星傳輸、地面廣播、與有線傳輸中，共通道多視訊之共同編碼；視訊會議中，共端點多視訊之共同編碼；以及 MPEG4 中，多個視訊物件平面(VOP)之共同編碼。

在真時視訊編解碼的實現方面，近年來數位硬軟體的發展，已使此能使用個人電腦的 CPU 或 DSP 插板達成。我們採用一個 TI TMS320C62xx DSP 的發展系統，依據 H.263 標準實現一個真時視訊編解碼程式。一方面可探討此種真時實現之技術關鍵所在，另一方面可供後續網路視訊通訊研究之用。

三、結果與討論

A. 變率傳輸下之近最佳視訊編碼

由於視訊率差特性之時變性，一般認為，變率傳輸可較定率傳輸為佳。但變率傳輸下之最佳率差編碼理論與演算法，前此僅有次最佳解存在，最佳解則尚未為人導出。我們導出一個近最佳解(near-optimal solution)及演算法。詳參[1]。

本問題之傳輸系統架構如圖一所示。在遲滯編碼的環境下，我們導出一組由傳輸率限制與緩衝器大小所造成的視訊位元率限制。我們考慮 Lagrange 乘數最佳化方法，並根據上述限制導出一些近最佳解的特性。且導出一個求取近最佳解的演算法，步驟如下：

1. 考慮連續 N 張圖框(frames) [n,n+N]。在其所宜使用之總位元限制下執行以下之 Lagrange 乘數最佳化：

$$\sum_{i=n}^{n+N} \min_{Q(i)} [D(i) + \lambda b(i)].$$

其中 D(i) 為第 i 圖框的差(distortion), b(i) 為其位元率, Q(i) 為所有可能的對該圖框的量子化方式, λ 為 Lagrange 乘數。若是此最佳化之結果可以找到一組傳輸率不違反變率通道傳輸率之限制, 則移至下一組圖框; 否則進行下一步。

2. 找出無法符合變率通道限制的第一組圖框序列, 假設為 [p,q]。執行如上式之 Lagrange 乘數最佳化, 其可用之位元率, 依其所無法符合之限制為上限或下限, 為在此序列時間內所容許之最大或最小值。回步驟 1。但在執行該步驟時將 [p,q] 排除在外, 並將其所用去之位元率自總位元率中減去。此排除與減去之動作係累計執行。

B. 多視訊之共同編碼

本問題之架構如圖二所示。圖三所示為 M=3 時的一個例子。其中我們將同時間的圖框總稱為超框(super frame)。在超框的觀念下, 可見多視訊之共同編碼, 與單一視訊的編碼相當類似。但一重要之點是, 由於不同超框的複雜度可能極不相同, 好的編碼法所造成的位元率可能隨時間有極大的上下變化。我們需要找到一個好的位元分配方法。其實 MPEG1/2 之視訊編碼, 因 I, P, B 圖框之率差關係不同, 亦面對類似的問題。

以下簡述我們所提出, 在超框結構呈周期性變化時 (如圖三的例子) 的基本編碼法。在可容許之編碼遲滯較低、或超框結構不呈周期性變化時, 可引申修改使用。詳參[2]。

1. 於時間 n 時, 獲取自其時起之一個超框周期中, 所有超框的率差關係。
2. 在此超框周期中, 所宜使用之總位元限制下, 使用 Lagrange 乘數法求取最佳之

位元分配。

3. 檢視此解會否造成緩衝器上溢或下溢。若否, 則依此解對時間 n 之超框作編碼。不然, 視其所造成者為上溢或下溢, 將超框 n 能用之位元率設為剛好不造成上溢或下溢之值, 並依此對之作最佳編碼。
4. 將時間前進一單位。

以上演算法, 可證明為非最佳[3], 但模擬編碼結果顯示: 較之不考慮視訊率差特性的一般方法, 此方法可獲致相當顯著的品質增益。

以下呈示一些模擬結果。其中我們考慮 QCIF 大小之 Foreman, Salesman, 及 Claire 三視訊之共同 H.263 編碼。三者的圖框率各為每秒 10 張、5 張、及 10/3 張。在共同編碼時, 我們以調整 QP (quantization parameter) 值來作率與差間之相對取捨, 其中我們考慮每圖框調一次 QP 及每區塊組 (GOB) 調一次 QP 等兩種方式。我們並將結果與公開軟體 tmn-2.0 之編碼結果比較。該軟體有二個運作模式: 給定位元率與給定 QP。由於該軟體並無多視訊共同編碼之能力, 我們在此二模式之編碼中係設定其輸入參數, 使各視訊之位元率與其相對圖框率成正比。

每圖框調一次 QP 的數據顯示: 在平均差方面, 共同編碼較之給定位元率的編碼, 對每一視訊之品質均有相當增益。較之給定 QP 的編碼, 共同編碼之總差值亦較低。圖四呈示各視訊各圖框之峰訊雜比值。每一區塊組調一次 QP 的結果, 較每圖框調一次並無太大的改善。其原因可能與上述共同編碼演算法之非最佳性及典型視訊之率差特性有關。

C. 真時軟體視訊編解碼器之實現

我們在 Blue Wave Systems 之 TI TMS320C62xx DSP 發展系統上遂行 H.263 之軟體編解碼器實現。由於係第一版實現, 我們僅考慮基本之 H.263, 而不考慮其 options。我們採用公開軟體 tmn-2.0 為基礎, 加以修改而成。以下先簡介所採用之 DSP, 再說明我們的 H.263 實現。詳可參[4]。

C62xx 的主要架構示於圖五。其處理器部分之主要特點如下: 256-bit very-long-instruction-word (VLIW) architecture, instruction-level parallelism, 及 software pipelining。其配置有 8 個各自獨立

的 functional units (含 6 個 ALU 及 2 個乘法器), 有 2 個 data paths 及 32 個暫存器。這些均可在圖五中窺知。TI 文獻所載之最高 clock rate 為 200 MHz, 但我們在計畫執行期間所能購得者僅為 160 MHz。因此, 在 8 個 functional units 完全平行使用的狀況下, 可得 1280 MIPS。

C62xx 的記憶體特點如下: 具 64 kB 之 data RAM 及 64 kB 之 program RAM、二個 DMA ports、32-bit 之 external bus 可用以與外接記憶體溝通。其記憶體空間分為若干群組(banks); 各群組本身都是 single-port, 但不同的群組可同步擷取(access)。

C62xx 定義了若干可由 C 語言程式呼叫的內建函式(intrinsic functions)。這些內建函式係對應於一些 C62xx 特定的指令, 可以有效利用 C62xx 的架構特點。一般而言, DSP 軟體發展可循以下之三步驟:

1. 發展良好的 C 程式碼, 但暫時不必特別顧及 DSP 的硬體特徵。
2. 根據 DSP 的硬體特徵來修改 C 程式碼(包括使用其內建函式), 使能有效運用 DSP。
3. 如果以上的程式碼仍不夠有效率, 則選擇其中較重要的部分做組合語言碼之修改。

本計畫主要的工作層次在前二步驟。

鑑於我們所用之發展環境之輸出入功能之特性, 本計畫主要考慮 sub-QCIF 格式之視訊編解碼。我們發現: 若直接使用 tmn-2.0 程式做 C compilation, 則其結果極無效率而不能真時執行; 一個主要因素是其需要太多的 external memory accesses, 此外當然也因沒能有效地運用 DSP 的平行處理能力。因此我們重組了編解碼程式的資料結構, 以減少 external memory access 的數量。在運動估計方面, 我們改用了 H.263 TMN8 的 fast diamond search, 而飛 tmn-2.0 中的 full search。修改後的運動估計, 每一大區(macroblock) 只需二次 external memory accesses: 一次為 fast diamond search, 一次為其後的 half-pel refinement。運動估計之後執行的 DCT/IDCT 及量子化與反量子化就不需要再做 external memory access。

實驗結果顯示: 目前實現之編解碼器, 在框內編碼時約耗 2.94 MIPS, 框際編碼則約 7.46 MIPS。因此, 160 MHz 的 DSP 每秒可作 54.4 張框內編碼或 21.4 張框際編碼, 200 MHz 的 DSP 每秒可作 68.0 張框內

編碼或 26.8 張框際編碼。

以上實現之程式, 尚有許多可以改進與後續發展之處。例如: 強化其平行處理或 DMA 功能以加快其編解碼速度, 擴展其能處理的視訊規格, 與攝影機及監視器之真時連線, 以及二點或多點間的真時網路傳輸。

四、計畫成果自評

研究內容與原計畫相符程度: 大體相符。相符程度約 80%。

達成預期目標情況: 創新之發現、理論之推導與模式建立、計算機模擬軟體之建立、實驗原型之建立、人才培育。

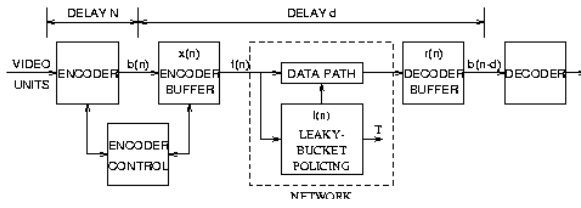
成果之學術與應用價值等: 學術價值高, 部分已發表於國際與國內會議、或已投稿於國際會議及期刊。應用價值中等, 尚需進一步研究。可應用之成果為單視訊變率傳輸時之近最佳遲滯編碼法、多視訊之次最佳共同編碼法、及 H.263 真時編解碼軟體。三者均有進一步研究的價值。前二者或亦可以考慮申請專利。

綜合評估: 本計畫獲得一些不錯的成果, 具相當之學術與應用價值, 並達成人才培育之目標。自評為「佳」。

五、參考文獻

- [1] D. W. Lin, "Jointly optimal video coding and rate control for variable-bit-rate transmission," in *Proc. IEEE Int. Conf. Image Processing*, paper no. WP10.7, 1998.
- [2] C.-W. Hung and D. W. Lin, "Joint bit allocation for multiple videos over multiple frames," in *Proc. Workshop Consumer Electronics*, pp. 85-90, Taipei, Taiwan, ROC, Oct. 1999.
- [3] W.-C. Gu and D. W. Lin, "Joint rate-distortion coding of multiple videos," *IEEE Trans. Consumer Electron.*, vol. 45, no. 1, pp. 159-164, Feb. 1999.
- [4] S.-W. Chen and D. W. Lin, "H.263 video codec implementation on a TMS320C62xx digital signal processor," in *Proc. Workshop Consumer Electronics*, pp. 1-4, Taipei, Taiwan, ROC, Oct. 1999.

六、圖表



圖一：變率視訊傳輸系統

(b) Salesman

圖二：多視訊之共同位元率分配

(c) Claire

圖三：三個視訊共同編碼的例子

圖四：多視訊不同編碼法之峰訊雜比曲線

(a) Foreman

圖五：TMS320C62xx 主要結構 (取自 TI 文件)