

# 行政院國家科學委員會補助專題研究計畫成果報告

## 音訊編解碼標準 MPEG4-TFC 之研究

計畫類別： 個別型計畫          整合型計畫

計畫編號：NSC 89-2218-E-009-014

執行期間： 89 年 7 月 1 日至 90 年 7 月 31 日

計畫主持人：劉啟民 教授

共同主持人：

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：交通大學 資訊工程系

中華民國 91 年 1 月 10 日

# 音訊編解碼標準 MPEG4-TFC 之研究

## The Study on Audio Coding Standard—MPEG4-TF Coding

計畫編號：NSC 89-2218-E-009-014

執行期限：民國 89 年 8 月 1 日起至民國 90 年 7 月 31 日止

主持人：劉啟民 交通大學資訊工程系

主要參與研究生：李文傑，李思瑋，楊存中，姚錦樹，林春發

### 一、中文摘要

壓縮標準 T/F Coding 共包含十模組：(一)增益控制，(二)濾波和頻譜轉換(Filterbank)，(三)波形雜訊控制(Temporal Noise Shaping, TNS)，(四)頻道間耦合(Intensity/Coupling)，(五)預測(Prediction)，(六)立體聲調整(M/S)，(七)位元分配(Bit-Allocation)，(八)人耳聽覺模式(Perceptual Model)，(九)偶合向量量化(Twin VQ)，(十)位元式數學化編碼(Bit-Sliced Arithmetic Coding)。而另外此壓縮流程並定義以下工具模組，(十一)感官雜訊取代(Perceptual Noise Substitution)，(十二)頻率選擇性轉換(Frequency Selective Switch Module)，和(十三)取樣上升濾波器(Upsampling Filter Tool)。其中，前八模組與 AAC 的模組相同理論，也是上年度國科會計劃的研究內容，本年度延續上年度前八模組的研究成果，並對另五模組加以討論。

### Abstract

The objective of the project is to study and implement the new audio standard. There are mainly thirteen modules in MPEG—AAC: (1) Gain Control, (2) Filterbank, (3) Temporal noise shaping, (4) Intensity/Coupling, (5) Prediction, (6) M/S adjustment, (7) Bit allocation, (8) Perceptual Model, (9) Twin Vector Quantization, (10) Bit-Sliced Arithmetic Coding, (11) Perceptual Noise Substitution (12) Frequency-Selective Switch Module, and (13) Upsampling Filter Tools.

Keywords: MPEG4, T/F Coding, Audio Coding.

### 二、緣由與目的

MPEG-4 之時頻轉換編碼，可當較低層的處理，同時亦是最重要的品質增進技術。其以 MPEG-2 AAC 為基礎，加入了數個新的模組，增強原本的編碼架構，提供更好更有效率的壓縮表現。在整個音訊編碼中，扮演舉足輕重的角色，是提升品質不可缺少的處理技術。

時頻轉換編碼採用 AAC 的架構，故兩者主要流程相當類似。不過為了增進壓縮的效能與品質，其加入了四個新的模組，包含：長期預測(Long Term Prediction)、知覺上類雜訊取代(Perceptual Noise Substitution)、位元切分的算術編碼(Bit-sliced Arithmetic Coding)與轉換域上加權交錯的向量量化(Transform-domain weighted interleave Vector Quantization)。

長期預測模組(LTP)：用於移除訊號中長期的重複性，以獲得編碼增益為主要目的。其操作是在時間軸上做預測，而在頻譜上取差值。觀察變化緩和或穩定的訊號，或是人聲部分具有明顯音高性質(pitch property)者，其前後時間之波形相當類似、關連性很大，利用預測的方法，可以移除這樣的重複性，而獲得編碼上的增益。LTP 便是為了這樣的目的。其在時間軸上計算預測值，兩者皆轉至頻譜上後再做相減，看其結果是否有好處，而決定該訊號採不採用 LTP。

知覺上類雜訊取代(PNS)：作雜訊取代，以更精簡的方式傳送資料。由於人耳對於雜訊感覺不靈敏，察覺不出頻譜分析結果與波形細緻的分佈，雜訊聽起來就像另一個。因此以雜訊取代方式，可以節省許多的位元。當在編碼時，若偵測到訊號為類雜訊，便不須傳遞頻譜值，而改以傳送雜訊指示旗標以及相關資訊(如：雜訊能量)，如此解碼端得知該訊號為類雜訊，自行產生虛造的訊號還原即可。故可節省大量的編碼位元。因為是利用人耳聽覺模式對雜訊感覺的特性，特稱之為『知覺上類雜訊取代』。

位元切分算術編碼(BSAC)：與霍夫曼編碼(Huffman Coding)一樣，也是不失真、不引進雜訊(noiseless)的編碼方法。其將資料作位元切分，組成一個個向量，由位元高往低做算術編碼，提供更有效率的壓縮，並具有細緻位元率調整的彈性(fine granule scalability)。首先將資料以二進位表示，由最重要位元(most significant bit, MSB)開始處理，組成 4 位元的向量，再以算術編碼方式加以壓縮，從低頻往高頻操作，再從 MSB 到 LSB(least significant bit)。同時引入多層次架構(multi-layer structure)的觀念，針對每一層定義所要處理的最高頻帶(band)，層次愈高，處理的範圍愈大，包含的頻帶也愈多。而且每高一層，位元率便提高 1kbps，以更多的資源編碼訊號，在可用的位元與音訊品質間達到一個平衡，提升壓縮效果與表現。

轉換域上加權交錯之向量量化(Twin-VQ)：以向量量化取代原本的單一數值量化(scalar quantization)，在較低位元率下提供更精簡、更有效率的壓縮。其包含兩次的平坦處理，以及最後的向量量化：將頻譜係數以交錯方式(interleave)分割成一個個子向量，並以比重計算取最佳的編碼向量。而量化採用共軛結構(conjugate

structure)，即有兩個碼書（codebook），分別從中選出一個向量，再組合起來，表示解碼後的頻譜係數。尋找過程有三階段：首先為事先選擇，分別從二碼書中，選出最接近原值的數個候選者。其次是主要尋找，從兩組候選者中，找出最佳的組合以表示輸入訊號。最後資料包裝，傳遞索引值。

### 三、研究報告應含的內容

本計畫成果已整理成三碩士論文[1]- [3]。對 MPEG4 T/F Coding 十三模組提出研究成果：

1. 增益控制，
  2. 濾波和頻譜轉換（Filterbank），
  3. 波形雜訊控制（Temporal Noise Shaping，TNS），
  4. 頻道間耦合（Intensity/Coupling），
  5. 預測（Prediction），
  6. 立體聲調整（M/S），
  7. 位元分配（Bit-Allocation），
  8. 人耳聽覺模式(Perceptual Model)。
  9. 偶合向量量化（Twin VQ），
  10. 位元式數學化編碼(Bit-Sliced Arithmetic Coding)，
  11. 感官雜訊取代(Perceptual Noise Substitution)，
  12. 頻率選擇性轉換（Frequency Selective Switch Module），和
  13. 取樣上升濾波器（Upsampling Filter Tool）。
- 已完全達到原計畫預期目標。

尤其對 MPEG4 四個新加進來的模組，我們深入地研究與分析，提出議題與我們的建議方法：

在 LTP（Long Term Prediction）模組裡，針對延遲時間的計算，我們提出兩個搜尋方法以減少複雜度，就所做實驗為例，計算量約剩下 15%，其正確率高達 94%以上。另外，我們也仔細分析 LTP 採用之決定，對不同的區塊轉換，研究 LTP 使用情況。並計算因加入 LTP 模組，真正獲得的編碼增益，接著評斷其編碼空間。

對於 PNS（Perceptual Noise Substitution）模組，我們由雜訊特性、語音壓縮技術與人耳聽覺的遮蔽效應，研究出三種類雜訊訊號的偵測方法，並以實驗驗證方法的可行性。結果顯示使用雜訊偵測，品質確實優於不用偵測或不採用 PNS 的情況，而且三種方法在不同位元率下，各自擁有最好的表現，因此我們做一個結論，提出因時因地採用的雜訊偵測方法。

在 BSAC（Bit-Sliced Arithmetic Coding）模組中，我們詳細地探討 BSAC 整個編碼架構與流程步驟，並研究其採用的層次結構，分析此結構之目的與運用。另外，我們將 BSAC 和同樣為無雜訊編碼的 Huffman coding 做一番比較，由實驗數據分析兩者之差異與優劣。

對於最後一個模組 TwinVQ，我們不做深入的研究探討，只簡略討論向量量化的部分。主要對向量量化與 AAC 舊有的量化編碼詳細地比較分析，看其差異與適宜的編碼空間。

本計畫分析與改良 MPEG-4 時頻轉換編碼中四個新加入的模組，不僅提供一個完整的架構分析與時頻轉換編碼的實作技巧，進一步希望能有效地提升音訊品質，努力達到低位元率、高品質的聽覺享受。

### 四、參考文獻

- [1] 姚錦樹，MPEG-4 時頻轉換編碼之模組分析與改良，八十九年六月交通大學資工系碩士論文。
- [2] 林春發，MPEG-4 音訊壓縮中 TwinVQ 之探討，八十九年六月交通大學資工系碩士論文。
- [3] 楊存中，MPEG-2 AAC 中 TNS 及預測模組之探討，八十八年六月交通大學資工系碩士論文。