

音訊編解碼標準MPEG-AAC之研究

Study on Audio CODEC-- MPEG-AAC

計畫編號: NSC89-2213-E-009-065

計畫期限: 88/8/1 - 89/7/31

主持人: 劉啟民

交通大學資訊工程系教授

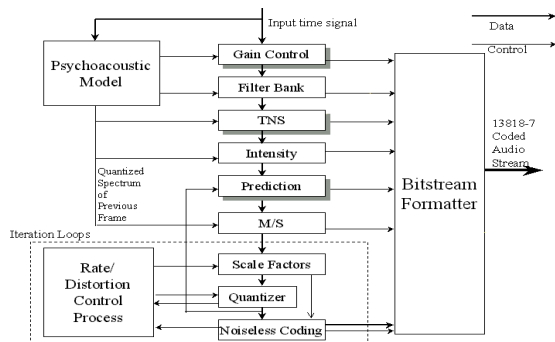
一、中文摘要(關鍵詞: Audio CODEC, MPEG—Audio, AAC)

MPEG-AAC 反應的是計算硬體的進步; 因此, MPEG 也遠比 MPEG1, MPEG2, 和 AC-3 更複雜, 如下圖所示, 音訊壓縮標準 MPEG-AAC 的壓縮包含有以下八模組: (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)。本計畫對此八模組進行理論探討和實做研究。

英文摘要 (Keywords: Speaker Recognition, Hidden Markov Model, Speech Processing)

There are mainly eight 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. This project will study the theoretical analysis and the associated real-time implementation system.

二、計畫之緣由與目的



錯誤! 找不到參照來源。 [1]

MPEG-2 AAC 音訊編碼標準制定完成於一九九七年四月。不同於 MPEG-2 BC (Backward Compatible), MPEG-2 AAC 不再受到向下相容 MPEG-1 的限制, 而更改了一些原有模組的編碼方式, 並且加入了一些新的模組以達到更高的壓縮率以及品質。我們透過錯誤! 找不到參照來源。來概略說明 MPEG-2 AAC 的編碼方式:

AAC 編碼標準中提供了三種 profile: Main profile, Low-complexity profile, 以及 Scalable sampling rate profile, 每種 profile 代表不同的編碼模組組合, 在錯誤! 找不到參照來源。中有陰影的三個模組即表示可以依照 profile 來設定內部的複雜度或者是

否採用該模組。使用者可以依照所需要的聲音品質、系統所擁有的計算能力, 來選擇適當的 profile。各個 profile 內的模組設定列於錯誤! 找不到參照來源。

Profile	模組設定
Main	} 不使用 Gain control 模組 } TNS 濾波器的級數最高 (20 級) } 使用 Prediction 模組
Low-Complexity (LC)	} 不使用 Gain control 模組 } 減少 TNS 濾波器的級數 (12 級) } 不使用 Prediction 模組
Scalable Sampling Rate (SSR)	} 使用 Gain control 模組 } 減少 TNS 濾波器的級數 (12 級) } 不使用 Prediction 模組

表 1. AAC 中三種 profiles 的模組設定

MPEG-2 AAC 所加入的新模組包含:

增益控制 (Gain control)

在 SSR profile 中, 加入了一個 Gain control 這個新的模組, 由 PQF (Polyphase Quadrature Filter)、gain detector 以及 gain modifier 組成。這個模組利用 PQF 將時間上的訊號切割成四個相同寬度的頻帶, 以達到 scalable 的效果 [2], 並經由增益 (gain) 的調整來減輕 preecho 的現象。

Temporal Noise Shaping (TNS)

以往在處理單一區塊內的訊號若在時間上有能量瞬間變化 (transient) 現象出現時, 往往會因為 preecho 現象造成解碼結果品質不佳的情形。在 AAC 中為了達到編碼效率的改進, 因此時頻轉換區塊的大小是 MPEG-1 層次三的兩倍, 但是這同時也造成產生 preecho 現象的機會增加, 所以如何減輕 preecho 現象在 AAC 中是個重要的問題。

TNS 是感官式音訊編碼 (Perceptual Audio Coding) 中的一個新的觀念, 目的在於使單一區塊在經過時間-頻率轉換、量化及編碼之後, 音訊在區塊內時間上的噪音遮蔽效果仍然能夠維持。在加入 TNS 這個模組後, 將可以發現 preecho 的現象被有效地改善。

預測模組 (Prediction)

預測模組的效果在於減少相鄰音框之間的資訊重複性 (redundancy)。由於聲音訊號在一般情形下, 短時間內具有相當程度的相似性, 所以如果採用相鄰音框間互相預測的方法, 可以去除這些相似的多餘資料, 以達到更好的壓縮效果。由於 AAC 可以提供最高到 96 kHz 取樣頻率的編碼, 在這麼高的取樣頻率下預測模組將扮演增加編碼效率的重要角色。

我們發現除了新加入的 TNS 模組外，在 AAC 之前的設計通常使用區塊切換的方法來減輕 preecho 的現象，所以在本計畫中探討這兩種方法的優缺點，以及應該如何選擇適當的方式來減輕 preecho。另一個問題則是當 TNS 和預測模組結合時，在後者中有部分的運算其實是不必要的，所以我們提出新的方法來改進 AAC 的預測模組。

三、研究方法與成果

本計畫對 MPEG-AAC 八模組提出實驗結果

- (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)。

並完成展示系統。

就創新性而言，本計畫從 preecho 現象開始討論，然後進一步解釋如何使用 TNS 來減少 preecho，最後提出我們認為最有效的解決 preecho 現象的模組組合。

本計畫深入解釋 AAC 中預測模組的做法及加入此一模組後對既有模組的影響。更提出解決 TNS 和預測模組結合時造成不必要的多餘運算問題的方法。

本計畫此部分的重要結果已於楊存中碩士論文篇發表。

四、結論與未來展望

在本計畫中，我們討論了 MPEG-2 AAC 中加入的新模組對於整個編碼架構的影響。

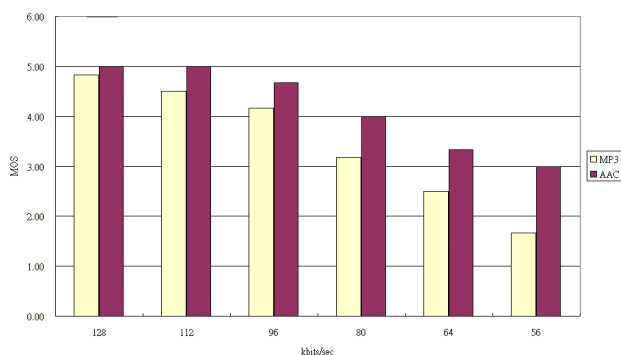
為了達到編碼效率的增加，AAC 中時頻轉換區塊的大小變成 MPEG-1 layer III 的兩倍，但這也造成產生 preecho 現象的機會增加。為了減輕 preecho，AAC 加入了 Temporal Noise Shaping (TNS) 的模組。由於以往常用的區塊大小切換的方法會使用較多位元來編碼，再加上決定方法的問題，有可能造成某些音框品質的降低；而 TNS 的方法則沒有這樣的問題。因此在聽覺實驗的結果中，我們可以歸納出 TNS 模組其實可以取代傳統的區塊切換方法，達到較有效率、品質較好的編碼結果。

AAC 最高可以編碼到 96 kHz 取樣頻率的音樂。為了消除音框之間的資料重複性，AAC 加入了預測模組。我們發現在編碼標準中，重疊使用預測模組和 TNS 模組的頻帶將會有不必要的計算，所以提出了兩個新方法來解決這個問題。而經由實驗的結果發現，使用新方法來編碼可以得到計算能力的節省與編碼增益的提

高兩方面的好處，因此原來的編碼標準是具有改進空間的。

經由 AAC 中設計的改進以及新模組的加入，使得經由 AAC 編碼出來的音樂品質較 MPEG-1 layer III 為佳。根據我們的聽覺實驗可以發現 AAC 使用較少的位元數達到較好的品質，且在低位元數編碼時更為明顯。此系統實做和實驗結果，說明此計畫在理論研究和技術累積均已達到一重要里程碑。

MP3 與 AAC 的編碼品質比較 (44.1 kHz 立體聲)



五、參考文獻

- [1] M. Bosi, K. Brandenburg, J. Johnston, et al., "ISO/IEC 13818-7 MPEG-2 Advanced Audio Standard," April 1997
- [2] M. Bosi, K. Brandenburg, et al., "ISO/IEC MPEG-2 Advanced Audio Coding," Journal of AES, October 1997
- [3] J. D. Johnston, "Audio Coding with Filter Banks," pp.287-307 in "Subband and Wavelet Transforms," Kluwer Academic Publisher, Norwell 1996
- [4] Jurgen Herre, J.D. Johnston, "Exploiting Both Time and Frequency Structure in a System that Uses an Analysis/Synthesis Filterbank with High Frequency Resolution," 103rd AES Convention, Preprint 4519, Sept. 1997
- [5] M. Link, "An Attack Processing of Audio Signals for Optimizing the Temporal Characteristics of a Low Bit-Rate Audio Coding System," 95th AES convention, Preprint 3696, October 1993
- [6] Masahiro Iwadare, Akihiko Sugiyama, et al. "A 128 kb/s Hi-Fi Audio CODEC Based on Adaptive Transform Coding with Adaptive Block Size MDCT," IEEE Journal of Selected Area in Communications, Vol. 10, No.1, Jan 1992
- [7] J. Herre and J. D. Johnston, "Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS)," 101st AES convention, Preprint 4384
- [8] N. Jayant and P. Noll, "Digital Coding of Waveforms," Englewood Cliffs, NJ, Prentice Hall, 1984

- [9] Allen Gersho and Robert M. Gray, "Vector Quantization and Signal Compression," Kluwer Academic Publisher, 1992
- [10] J. Herre and J. D. Johnston, "Continuously Signal-Adaptive Filterbank for High-Quality Perceptual Audio Coding," IEEE Workshop on ASSP, 1997
- [11] Philippe Delsarte and Yves Genin, "On the Splitting of Classical Algorithms in Linear Prediction Theory," IEEE Trans. on ASSP, Vol. 35, No. 5, May 1987
- [12] S. R. Quackenbush, and J. D. Johnston, "Noiseless coding of quantized spectral components in MPEG-2 Advanced Audio Coding," IEEE Workshop on ASSP, Oct. 1997
- [13] Thippar V. Sreenivas and Martin Dietz, "Vector Quantization of Scale Factors in Advanced Audio Coder (AAC)," IEEE Intl. Conference on ASSP, May 1998
- [14] D. H. The and A. P. Tan, "An Improved Stereophonic Coding Scheme Compatible to the ISO/MPEG Audio Coding Algorithm," ICCS/ISITA 1992
- [15] H. Fuchs, "Improving MPEG audio coding by backward adaptive linear stereo coding," 99th AES convention, Preprint 4086, Oct. 1995