



## 感知導向視訊編碼器：人眼感知分析引擎之硬體架構設計及其於H.264視訊編碼器之應用

### 作品名稱

Perception-Aware Video Encoder: Hardware Architecture Design of Bio-Inspired Human Eyes Perception Evaluation Engine for H.264 Video Encoder

### 隊伍名稱

視覺多媒體 **Visual Multimedia**

### 隊長

吳東興 台灣大學 電子工程學研究所

### 隊員

吳冠林 台灣大學 電子工程學研究所

### 指導老師

簡韶逸 台灣大學 電子工程學研究所

### 作品摘要

多媒體已深深影響現在人類的生活，由於計算量與傳輸頻寬的限制，高品質視訊資料仍存在很大的改善空間。目前視訊壓縮標準H.264雖提供數十至百倍的壓縮率，但視訊最終端是人眼，傳統視訊編碼只考慮以Peak Signal-to-noise Ratio來當作品質指標，並沒有仔細考慮人眼視覺系統(Human Visual System, HVS)來做最佳化。適當的分配位元，可使位元率降低，相對提昇視覺品質。然而一個模擬人眼視覺特性和提供相同主觀視訊品質情況下並降低位元傳輸率的系統，需要大量的系統頻寬及運算複雜度，無法滿足即時視訊編碼系統，因此我們提出人眼感知分析引擎之實做，加強一般視訊編碼晶片在這方面的功能。

本作品的目標是透過一感知分析引擎分析視訊內容並模擬人眼模型以決定視訊內容的位元分配，我們結合結構相似度模型，視覺注視模型，視覺感知模型等來得到視訊畫面中每個區塊對於人眼感知的重要程度，結合H.264視訊編碼器，我們提出一套決定Quantization Parameter (QP)的方法，進一步發展適合硬體實做之架構設計。

提出之演算法藉由與H.264參考軟體之模擬，在不降低主觀視覺品質的情況下與QP 24~36的範圍中節省了5~40%的位元率。在硬體實做上，我們的晶片使用TSMC 0.18  $\mu$  m技術，面積大小為3.3x3.3mm<sup>2</sup>，功率消耗為83.9mW，硬體最大可處理之解析度為HDTV720p (1280x720)。

### Abstract

Multimedia has been affecting the human lives. Current H.264 video compression standard provides a dozen to a hundred times of compression ratio. The traditional video coding only use Peak Signal-to-noise Ratio as quality indicators without considering the human visual system. Proper bit allocation can reduce the bit rate, relative to enhance visual quality. However, models which simulate visual properties require a large system bandwidth and computing complexity; it cannot satisfy the real-time issue, so we proposed a perception evaluation engine, strengthen the general video encoder functions in this regard.

The goal is to work through a perceptual evaluation engine to determine the bit allocation of video contents; we combined the structural similarity model, visual attention/perception model to get the importance of perception for each block. Combined with H.264 video encoder, we propose Quantization Parameter (QP) selection scheme. Algorithm proposed is combined with H.264, without lowering the subjective quality, saving 5-40% of the bit rate in the QP range 24-36. For hardware implementation, we use TSMC 0.18  $\mu$  m technology, chip area is 3.3x3.3mm<sup>2</sup>, power consumption is 83.9mW, and the processing capability is HDTV720p (1280x720) in 30fps.