

# 國立交通大學

多媒體工程研究所

碩士論文

可調視訊編碼多層編碼控制的快速演算法:理論的位  
元與失真關係函數初步分析



Fast SVC Multi-Layer Encoder Control: Preliminary Results on  
Theoretical Rate-Distortion Function

研究生：蔡閏旭

指導教授：彭文孝 教授

中華民國 九十九年九月

可調視訊多層編碼控制快速演算法:位元與失真關係函數初步理論分析

Fast SVC Multi-Layer Encoder Control: Preliminary Results on  
Theoretical Rate-Distortion Function

研究生：蔡閔旭

Student：Ruen-Shiu Tsai

指導教授：彭文孝

Advisor：Wen-Hsiao Peng

國立交通大學

多媒體工程研究所



Submitted to Multimedia Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年九月

# 可調視訊多層編碼控制快速演算法:位元與失真關係函數初步理論分析

研究生：蔡閔旭

指導教授：彭文孝

國立交通大學多媒體工程研究所 碩士班

## 摘 要

以可調整視訊編碼(SVC)之架構為基礎，本論文闡釋一個多層編碼控制下的快速演算法。由於傳統由下至上編碼控制的增進層所利用的基層資訊，其適用性不一，相較於單一層編碼會有不對稱的編碼效能損失，而多層編碼控制能在基層和增進層之間的編碼效能做權衡，此編碼方法在先前就被提出。而我們先前的研究，修正了多層編碼控制的公式，使編碼效能更具有可預測性。由於多層式編碼控制的編碼速度過於緩慢，本論文沿用之前的公式，利用時間性資訊和隨編碼內容更新的特性，提出了一個兩段式的快速演算法。根據實驗結果，本快速演算法和徹底式搜尋的比較之下，減少了平均 71~79% 的時間，且在編碼品質上無太多失真。此外，實驗結果也顯示，本論文提出的快速決策演算法和過去的快速決策演算法相比較，更接近徹底式搜尋的編碼效能，顯得更具有穩定性。

另外，鑑於過去的快速決策演算法通常以統計學為基礎進行分析，缺乏理論依據。我們以 H.264/AVC 和 SVC 為基礎，提出了位元配置和失真之間的關係公式，並對 Mode Decision 的影響因素進行理論分析。期待未來可以應用到多層編碼控制。本研究目前還在探索階段。

# Fast SVC Multi-Layer Encoder Control: Preliminary Results on Theoretical Rate-Distortion Function

Student : Ruen-Shiu Tsai

Advisor : Wen-Hsiao Peng

Institute of Multimedia Engineering  
National Chiao Tung University

## ABSTRACT

In this thesis, a fast mode decision for SVC multi-loop encoder control is proposed based on statistics analysis. In conventional bottom-up encoder control, enhancement layer (EL) coding is characterized by its uneven distribution of rate-distortion loss relative to single-layer coding. For trading off the coding efficiency of the base layer and the enhancement layer, a multi-layer encoder control method was proposed. In our previous study, a modified scheme of multi-layer encoder control is presented to provide a new decision criterion which are more predictable and consistent with different choices of the weighting factor. Continuing from previous study, we proposed an two-stage and context-adaptive algorithm with the help of temporal information. Experimental results show that proposed fast mode decision algorithm can save approximately 71~79% of total encoding time compared to conventional multi-layer encoder control, with only little degradation in the rate-distortion performance. The results also show that the proposed algorithm is more stable and, because the RD- performance is closer to exhaustive search.

Described below and above are independent, in this thesis, we also study the theoretical analysis of mode decision. Previous fast mode decision is usually based on statistics analysis rather than theoretical analysis. We study the theoretical analysis of mode decision for H.264/AVC and SVC, propose a formulation between bit allocation and distortion, and add the factors of mode decision. Look forward to extending to multi-layer encoder control. This research is in progress.

## 誌 謝

首先我要感謝我的指導教授—彭文孝 博士，一次次給予我在學問研究上的精闢指導，讓這篇論文得以漸趨完整。彭老師凡事追求卓越的精神，對於研究問題深入剖析的嚴謹態度，以及細心與耐心的指導方式，讓我在這兩年的研究生涯中受益良多。在此謹向我的老師致上無限的敬意。

其次，這篇論文可以完成，也要感謝交大這個大環境，我要感謝林哲永學長在我的研究上擔任了啟蒙的指導，不僅在 H.264 和 SVC 上的專業領域，不辭辛勞的與我討論，更在一開始的研究實作上給予許多珍貴的意見，並且能適時從旁給予建議修正我已偏差的研究方向，使我在這兩年的碩士生涯，不再舉步維艱。謹此致上由衷的謝意。

有榮幸進入 MAPL 實驗室，能夠有熱心與親切的實驗室成員們的切磋與討論，是我在碩士時期最充實的時光。我要感謝其餘的學長姐們—陳漪紋 博士、陳俊吉 博士、與詹家欣 博士 陳建穎，一步步帶領我進入這個專業的領域；感謝我的好同學們王澤璋、楊復堯、與吳思賢，不論是研究上或是生活上，他們總是給予我最直接的協助以及苦樂分享；感謝我的學弟妹吳崇豪、曾于真、黃嘉彥、李宗霖與陳孟傑，在論文撰寫和實驗數據整理方面上，幫了我很大的忙，在研究思維上我最要感謝的是王澤璋和吳崇豪，給了我許多寶貴意見，在實驗和整理資料上最感謝曾于真和黃嘉彥，口試前這兩個月內給予了許多無私的協助。

最後，我要感謝我的父母—蔡智暉 先生、賴惠丹 女士的栽培，在爭取碩士學位的路上，給予精神上的支持和適度的溫暖，能夠專心的在研究領域上打拼。感謝我的姊姊—蔡明秀，分擔家計使我無後顧之憂。感謝我的女友—陳郁欣，畢業前夕分擔我的情緒、對我體諒與關心，讓我在孤獨的研究之路上的最低潮時，還能夠感受到一絲溫暖。感謝你們一路的陪伴打氣與支持，在此僅將這篇論文獻給各位，謝謝你們！

# 目錄

目錄	II
I	
圖表目錄	IV
表格目錄	V
CHAPTER 1. INTRODUCTION	1
1.1. BACKGROUND	1
1.2. PROBLEM STATEMENT	5
1.3. CONTRIBUTION	6
1.4. ORGANIZATION	6
CHAPTER 2. SVC ENCODER CONTROL	7
2.1. INTRODUCTION TO SVC	7
2.1.1 Concept	7
2.1.2 Inter-Layer Prediction	8
A. Inter-Layer Motion Prediction	9
B. Inter-Layer Intra-Prediction	9
C. Inter-Layer Residual Prediction	9
2.2. SVC ENCODER CONTROL	10
2.2.1 Bottom-up Encoder Control	10
2.2.2 Multi-layer Encoder Control	11
A. Schwarz's MLEC	11
B. Lin's MLEC	12
2.3. COMPARISON AND SUMMARY	13
CHAPTER 3. FAST MODE DECISION ALGORITHM FOR SVC MULTI-LAYER ENCODER CONTROL	14
3.1. ANALYSIS OF MODE DISTRIBUTION	16
3.1.1. Mode Distribution in MLEC	16
3.1.2. Update of Mode Distribution Probability	24
3.2. PROPOSED FAST MODE DECISION ALGORITHM	25
3.3. SUMMARY	27
CHAPTER 4. EXPERIMENTS AND ANALYSES	28
4.1. TEST CONDITIONS	28

4.2.	COMPARISON OF FAST MODE DECISION ALGORITHMS.....	29
4.2.1	Experiment 1.....	29
4.2.2	Experiment 2.....	33
4.2.3	Experiment 3.....	36
<b>CHAPTER 5. THEORETICAL ANALYSIS ON SVC RATE-DISTORTION</b>		
	<b>FUNCTION .....</b>	<b>38</b>
5.1.	THEORETICAL MODEL .....	38
5.1.1.	Theoretical R-D functions of Single-Layer Hybrid Coding .....	38
5.1.2	Theoretical R-D functions of Multi-Layer Hybrid Coding .....	46
5.2.	THEORETICAL DATA AND MODEL ANALYSIS.....	50
5.2.1	Experiment 1.....	51
5.2.2	Experiment 2.....	53
5.2.3	Experiment 3.....	55
5.2.4	Experiment 4.....	56
5.2.5	Experiment 5.....	57
<b>CHAPTER 6. CONCLUSIONS.....</b>		
	<b>BIBLIOGRAPHY .....</b>	<b>60</b>



## 圖表目錄

Figure 1.1	The definition of MB Pair (from Lin's thesis) .....	1
Figure 1.2	Schwarz's 的快速演算法在選擇 Dominant Mode Pair 時的規則 ..	2
Figure 1.3	Lin 的快速演算法在選擇 Dominant Mode Pair 時的規則 .....	3
Figure 1.4	所提出快速演算法在選擇 Paired mode 時的規則 .....	5
Figure 2.1	Multi-layer structure with Inter-Layer Prediction for spatial and temporal scalable coding [1] .....	7
Figure 2.2	The flowchart of Inter-Layer Prediction Mechanism (from Lin's thesis) .....	8
Figure 3.1	演算法概觀流程圖 .....	15
Figure 3.2	Directly stored modes in the enhancement layer .....	17
Figure 3.3	EL 在不同的 TMS 時 Mode Decision 為 Skip mode 的機率 .....	19
Figure 3.4	EL 在不同的 TMS 時 Mode Decision 為 Inter16x16 的機率 .....	20
Figure 3.5	EL 在不同的 TMS 時 Mode Decision 為 Inter16x8 的機率 .....	20
Figure 3.6	EL 在不同的 TMS 時 Mode Decision 為 Inter8x16 的機率 .....	21
Figure 3.7	EL 在不同的 TMS 時 Mode Decision 為 Inter8x8 的機率 .....	20
Figure 3.8	EL 在不同的 TMS 時 Mode Decision 為 BLSkip 的機率 .....	22
Figure 3.9	EL 在不同的 TMS 時各種 mode 的機率 .....	22
Figure 3.10	Macroblock accumulate probability in total sequence of each temporal mode score .....	23
Figure 3.11	更新假想個數示意圖 .....	24
Figure 3.12	演算法的流程示意圖 .....	25
Figure 4.1(a)(b)(c)(d)	各種 weight 之下的 hit ratio .....	35
Figure 4.2	The Rate-PSNR curve of QCIF sequence .....	37
Figure 4.3	The Rate-PSNR curve of 4CIF sequence .....	37
Figure 5.1	Inter Prediction 的 residual signal 編碼流程圖 .....	39
Figure 5.2	H.264/AVC 中 residual signal 取得方法示意圖 .....	40
Figure 5.3	倒水定理概念圖 .....	43
Figure 5.4	SVC 中 Inter-Layer Residual Prediction 示意圖 .....	46
Figure 5.5	Rate-PSNR curve .....	51
Figure 5.6(a)(b)(c)	在 BL mode 固定之下不同 EL mode 的 R-D 效能 .....	53
Figure 5.7(a)(b)(c)(d)	BL 和 EL mode 在各種組合之下實行 Inter-Layer Residual Prediction 的 Rate-PSNR curve .....	55
Figure 5.8(a)(b)(c)(d)	不同 QPD 之下的 Rate-PSNR curve .....	56
Figure 5.9(a)(b)(c)	QPD 不同之下的 Rate-PSNR curve .....	57

## 表格目錄

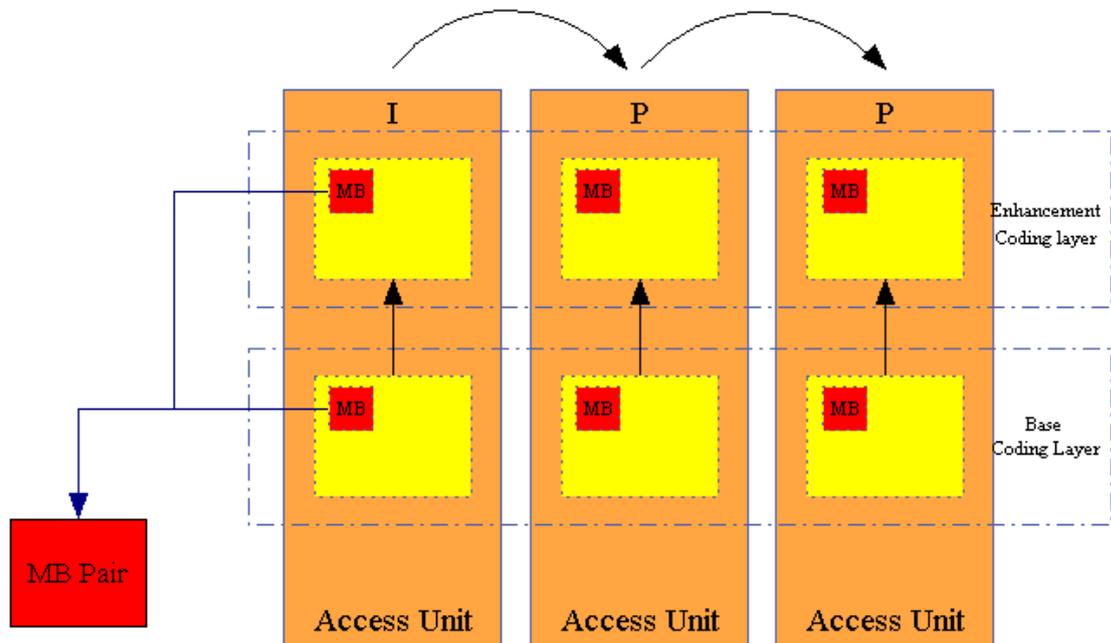
Table 3.1 test mode condition .....	15
Table 3.2 test video sequence .....	16
Table 3.3 training sequence .....	16
Table 3.4 location score in reference frame.....	17
Table 3.5 mode score .....	17
Table 3.6 Example .....	17
Table 3.7 BL&EL mode 的配對 .....	25
Table 4.1 :測試條件 .....	28
Table 4.2 frame rate = 3.75 BL, EL QP = 28, 22 average .....	30
Table 4.3 frame rate = 3.75 BL, EL QP = 34, 28 average .....	31
Table 4.4 frame rate = 3.75 BL, EL QP = 40, 34 average .....	32
Table 4.5 frame rate = 15 BL, EL QP = 34, 28 .....	33
Table 4.6 frame rate = 7.5 BL, EL QP = 34, 28 .....	34
Table 4.7:各種條件之下，testing sequence 的平均 Time Saving.....	36
Table 5.1 symbol definition.....	39
Table 5.2 :實際資料測試條件.....	50
Table 5.3 :理論分析測試條件.....	50



# Chapter 1. Introduction

## 1.1. Background

考慮多變容量的不同設備，和各類型網路所傳遞的視訊內容，在許多應用中可調性已經是很重要的性質，基於這個理由，Joint Video Team (JVT)以 H.264/AVC 為基礎，標準化了 Scalable Video Coding (SVC)[1]，SVC 支援了在單一的 Bitstream 中擁有 spatial，temporal，SNR 和它們組合起來的可調性，且此 Bitstream 可只具備其中幾樣特性，如只有解析度或畫質的差別。空間和畫質的不同可以在不同層間展現，而時間性是以一種有等級制的方法表現，這點和 H.264/AVC 幾乎是一樣的。因為一次編碼多層可以使 Rate 有適當的減少，且依然維持理想的畫質，SVC 提供了具適應性層與層間猜測的方法 (Inter-Layer Prediction Mechanism)，此方法可以盡可能的重複使用較低層資訊。



MB Pair are co-located MBs

A frame is a GOP

Figure 1.1 The definition of MB Pair (from Lin's thesis)

為了決定基層(Base Layer, BL)和增進層(Enhancement Layer, EL)的mode，Joint Scalable Video Model (JSVM)[2]採用一種由下到上編碼方法(Bottom-up Encoder Control, BUEC)。這方法是在決定BL的編碼方式時，只有考慮BL的內容而不考慮較高的EL，同理，較低的EL在決定編碼方式時，不會考慮到較高的EL。由於BUEC具有比單一層編碼(Single-Layer Coding)較無法預期或不規則的Rate-Distortion (R-D)，通常較高的EL會比較低層有更多的編碼損失，像是Distortion，是由於BUEC以循序的方式進行Mode Decision而導致這樣的結果。

想解決上一段所描述的問題，Schwarz *et al.* [3]提供了一種可以在BL和EL間給予適當權重的方法，稱之多層編碼控制(Multi-loop Encoder Control, MLEC)，在做分割分式的決定時，BL使用權重式的Lagrangian成本決策方法。就如Figure 1.1顯示的，在MLEC做分割方式決策時的基本單位是一對Macroblock (MB)的組合，分別來自BL和EL相對位置的MB，對於每一對MB Pair，它們的分割方式有BL和EL分割方式的Cartesian Product組，一共有56組Mode Pair，比起BUEC這明顯是很大的計算量，所以對MLEC來說，一個好的快速決策演算法就顯得必要。

為了減少編碼時間，在H.264/AVC和SVC中已經有為數不少的快速演算法被提出。H.264中，Kannangara *et al.*[12]基於所提出的Lagrangian Cost函數判斷是否及早決定編碼為Skip mode，Kim *et al.*[13]利用current MB的資訊計算出三種特性，用這些特性及早決定編碼為Intra或Inter mode，Kim *et al.*[14]以reference frame中co-located mode決定current mode的candidate mode，Ri *et al.*[15]同時以spatial和temporal的mode決定current mode的candidate mode。而SVC中，像是Yeh *et al.*[16]中利用BL co-located mode和EL周圍mode來預測current mode，並提出一套及早決定mode的判斷方式，Kim *et al.*[17]用BL中co-located MB和其周圍MB，對不同的mode和所在位置給於適當的權重計算出一個值，並同時利用這值和BL mode一同決定candidate mode。

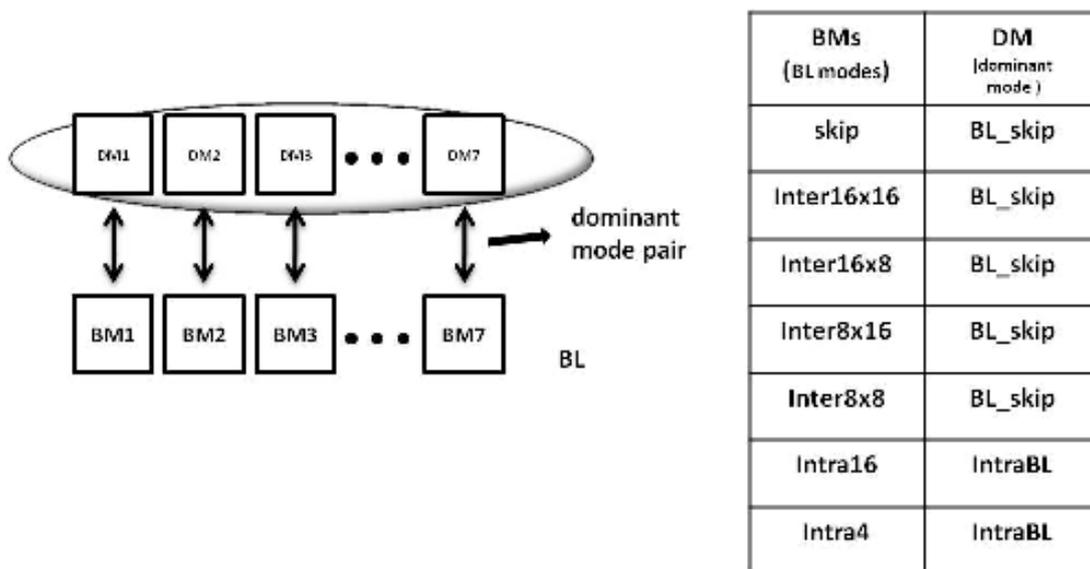


Figure 1.2 Schwarz's 的快速演算法在選擇 Dominant Mode Pair 時的規則

目前根據 MLEC 所設計的快速演算法非常少，其中以 Schwarz et al.[3]所提出的快速演算法為主。其做法是當 BL 的 MB 做 Mode Decision 時，每一個 mode 都會有一個 EL 的 mode 做為配對，此種配對稱為 Dominant Mode Pair (DMP)。配對方法是當 BL 選擇 Inter/Intra 的 mode 時，EL 選擇 BLskip/IntraBL 做為預測(如 Figure 1.2)，然後 BL 利用這組配對去計算 R-D cost，並選擇 R-D cost 最小的做為 BL 的 mode。最後在 EL 中對應的 MB，再根據 BL 選定好的 mode 去做徹底式搜尋，找出 R-D cost 最小的 mode 做壓縮。但此方法存在一些嚴重的問題：

- I. 在[3]中所提的MLEC多目標最佳化問題中，其限制會因為權重的設定而改變。而這限制的改變將會在權重較高的情況下，造成不理想的R-D performance。詳細內容會在2.2.2(A)提及。
- II. 在選擇BL中每一個MB的mode時，都是使用Figure 1.2中的7組配對組合去做計算，並不會依照sequence的特性做調整。這種配對方式太過於單調，不具任何可調性。
- III. 此7組配對只是所有配對模式中的一部分，所以此配對模式與實際的編碼模式可能並不相同。若我們利用此方法去預測EL的mode，可能與實際的mode差異太大，因此會導致BL在做Mode Decision時出現錯誤，若再用此錯誤的mode去找EL中的，將會造成不可預期的嚴重失真。
- IV. 此演算法雖然是根據MLEC去設計，但卻沒有考慮權重(weighting)對Mode Decision的影響。因為在不同權重時，並不會改變BL在選擇mode時所需計算的7組配對。

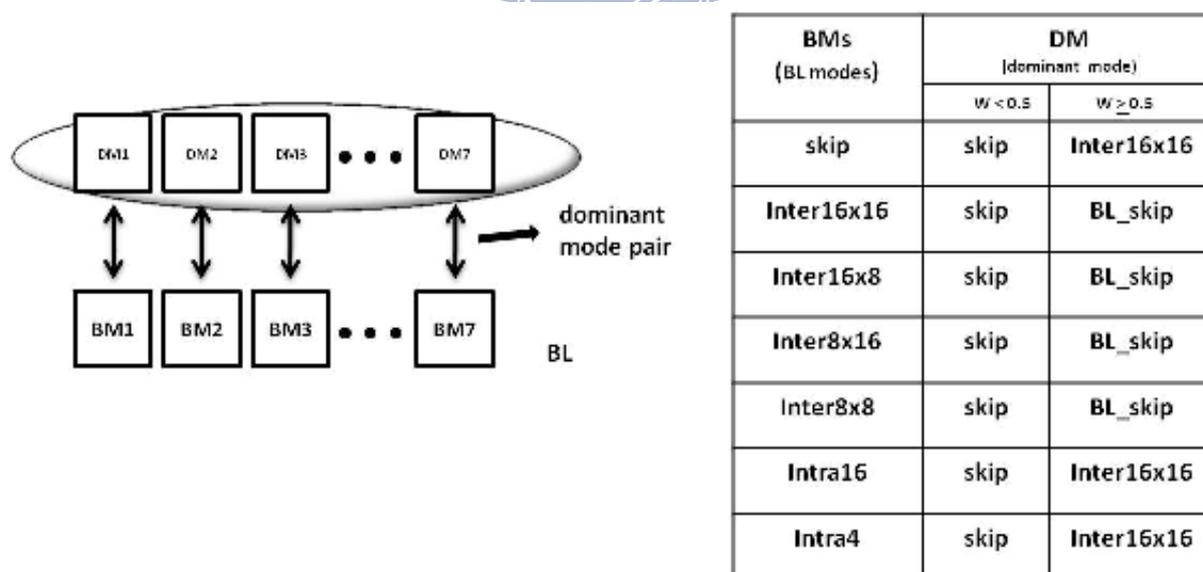


Figure 1.3 Lin 的快速演算法在選擇 Dominant Mode Pair 時的規則

為了解決Schwarz *et al.*[3]所提出快速演算法的問題，Lin *et al.*也提出了一套快速演算法。然而在Lin的演算法中，針對上述的問題修正以下幾點：

- I. 修正了[3]中所提出的MELC多目標最佳化問題，改善了權重因子對限制的影響，詳細內容將在2.2.2.的(A)提及。
- II. 修正了[3]中BL在做Mode Decision時，不會隨權重改變而有不同配對計算的問題。而且也增加DMP的配對組合，使演算法稍具可調性。此修正如Figure 1.3所示，BL在權重小於0.5時，其在EL的配對都為skip；在權重大於等於0.5時，其配對有Inter 16x16和BLskip。而且配對的mode不僅只有BLskip和IntraBL兩種。

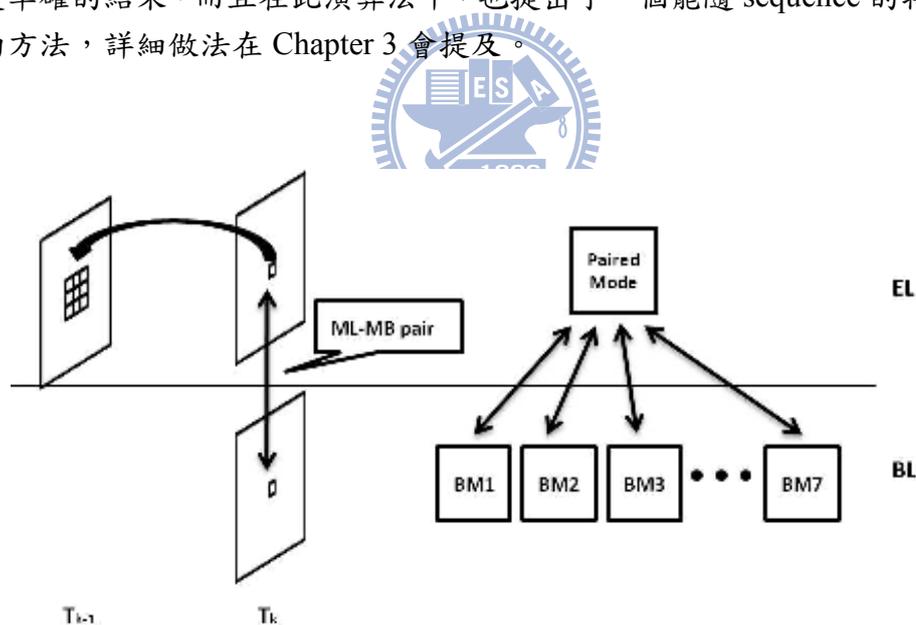


## 1.2. Problem Statement

雖然 Lin 所提出的演算法改善了[3]所存在的問題，並且擁有較好的 R-D 效能，但上述部分的問題仍然存在著。沒解決的問題是：

- I. 在選擇BL mode時，仍然只有少數的配對組合可以做計算。
- II. 依然不能隨sequence的特性，做不同的配對組合。
- III. 在分析時依然缺乏理論依據。

因此在本論文中，繼續針對這些問題，再次提出一套新的快速演算法來解決。然而新的演算法主要是 EL 利用時間(temporal)的特性，先將 EL 中 MB 的 mode 預測出來，再利用此預測的 mode 做為 BL 中 MB 的配對。如 Figure 1.4 所示，Paired mode 是利用前一張畫面的資訊，預測出一個最可能實施的 mode 做為 BL modes 的配對。此種做法和先前 Schwarz 和 Lin 所提出的方法有很大的差異，沒有所謂的 DMP 概念，因為不同的 MB 所要計算的配對組合並不會一樣，不用很單調的只計算少數的配對組合。利用此演算法能提高預測 EL mode 的正確性，再用這個 EL 的預測結果去選擇 BL 的 mode，也可得到更準確的結果。而且在此演算法中，也提出了一個能隨 sequence 的特性而選擇不同配對的方法，詳細做法在 Chapter 3 會提及。



(a)利用temporal資訊示意圖

(b)選擇配對的mode示意圖

Figure 1.4 所提出快速演算法在選擇 Paired mode 時的規則

另外，由於目前大多數演算法在進行分析時都是以統計學為立論，缺乏理論依據，所以試圖模擬Block-Based Hybrid Video Coding的編碼過程，以各種有理論根據的model進行假設，加入各式各樣可能影響Mode Decision的因素進行分析，觀察model是否能夠產生和實際情況類似的趨勢，如果在大多數情況都可以符合實際情況，那將來就可以利用這些model進行Mode Decision的因素分析，產生新的Mode Decision因素。

### 1.3. Contribution

具體的說明在這次工作中我們主要的貢獻，有以下幾點：

- 提出了一套新方法決定配對的 mode，當 BL 要決定 mode 時，在 EL 利用 temporal 資訊計算出一個值，並在這值對應的 mode 機率中，找出機率最大的 mode 成為配對的 mode。
- 提出了一套更新機制，可依編碼內容調整 EL 配對的 mode，在 frame level 進行更新，使 R-D 效能更接近徹底式搜尋，更具備穩定性。
- 推導出 Single-Layer 和 Multi-Layer 的 Rate-Distortion Function，並對這 model 進行觀察分析，試圖找出這些因素變化時對 Mode Decision 的影響，並和實際資料比較，幫助未來在設計演算法時可以有新思維，其內容在 Chapter 5，此研究目前還在探索階段。

實驗結果顯示我們所提供的快速演算法，除了能維持和 BUEC 差不多的時間外，和徹底式搜尋(Exhaustive Search)比較，平均約 72%~78%的 time saving，和幾乎一致的 R-D 效能，而 Lin's 快速演算法約 73%~81%的 time saving，Lin's 演算法第二個步驟改用完整式搜尋約 71%~79% time saving，Schwarz's 快速演算法約 74~82 % time saving，在時間上來說差不多。但是觀察 R-D 時，Schwarz's 快速演算法明顯 Rate 所需增加許多，而本論文提出之演算法和 Lin's 演算法的平均 Rate 都增加 1%以內,PSNR 減少 0.1db 之內，但本論文所提出之演算法明顯更具穩定性,實驗數據可參考 Chapter 4.

### 1.4. Organization

本篇論文的組成如下，Chapter 2 描述 RDO 的問題公式，包含 BUEC 和 MLEC。Chapter 3 呈現本論文所提供快速演算法，Chapter 4 提供了軟體模擬的運算結果和與其他演算法比較，Chapter 5 內容為 Mode Decision 理論分析，但目前還在探索階段，最後，Chapter 6 給一些總整理及觀察結論，和未來可以實行的工作。

# Chapter 2. SVC Encoder Control

## 2.1. Introduction to SVC

### 2.1.1 Concept

Scalable Video Coding (SVC)是Joint Video Team (JVT)以H.264/AVC為基礎的延伸發展出來的標準，使其具有可調性。一個SVC Bitstream是由一個BL和一個或一個以上的ELs所組成，其中包含了幾種可調性，且可以截取一部份的Bitstream產生較低的畫質。SVC支援三種型態的可調性：

- I. 空間可調性(Spatial Scalability):不同解析度的層之間的相關性是經由具適應性的Inter-Layer Prediction (ILP)技術來實現。
- II. 時間可調性(Temporal Scalability):在每個層都提供了多層次的時間預測結構。
- III. 畫質可調性(Quality Scalability)在SVC中提供了兩種方法。
  - 粗質畫面可調性(Coarse-grain Quality Scalable Coding, CGS): 可以被視為一種空間可調性的特例，BL和EL具有相同的解析度。
  - 中質畫面可調性(Medium-grain Quality Scalable Coding, MGS):在每個空間層中提供了畫質精細層，且允許以封包為基礎的畫質可調性。

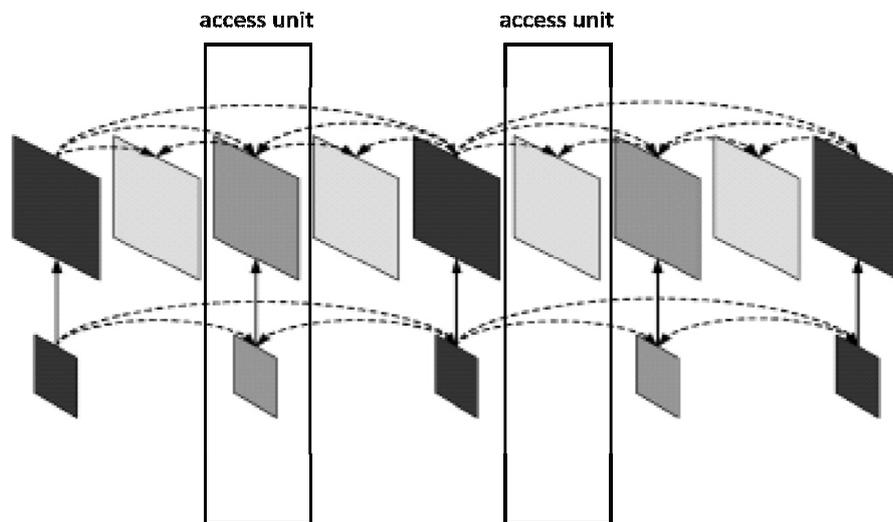


Figure 2.1 Multi-layer structure with Inter-Layer Prediction for spatial and temporal scalable coding [1]

SVC 跟隨同樣用在 H.261 MPEG-2 Video, H.263, 和 MPEG-4 Visual 的傳統多層編碼方法，每一個層都被交付一個相依的識別碼 D<sub>o</sub>BL 的相依識別碼被設定為 0，每多一個層，識別碼就增加 1。為了改善編碼效能，添加了 Inter-Layer Prediction Mechanisms 的方法，在 Figure 2.1.有介紹。一個 access unit 被一段時間內的內容所組成，可能涵蓋多個層的內容，可能包含不同的解析度。因較高層和較低層的 frame rate 可能不同，所以較低層 frame 並不需要包含在所有的 access units 中。

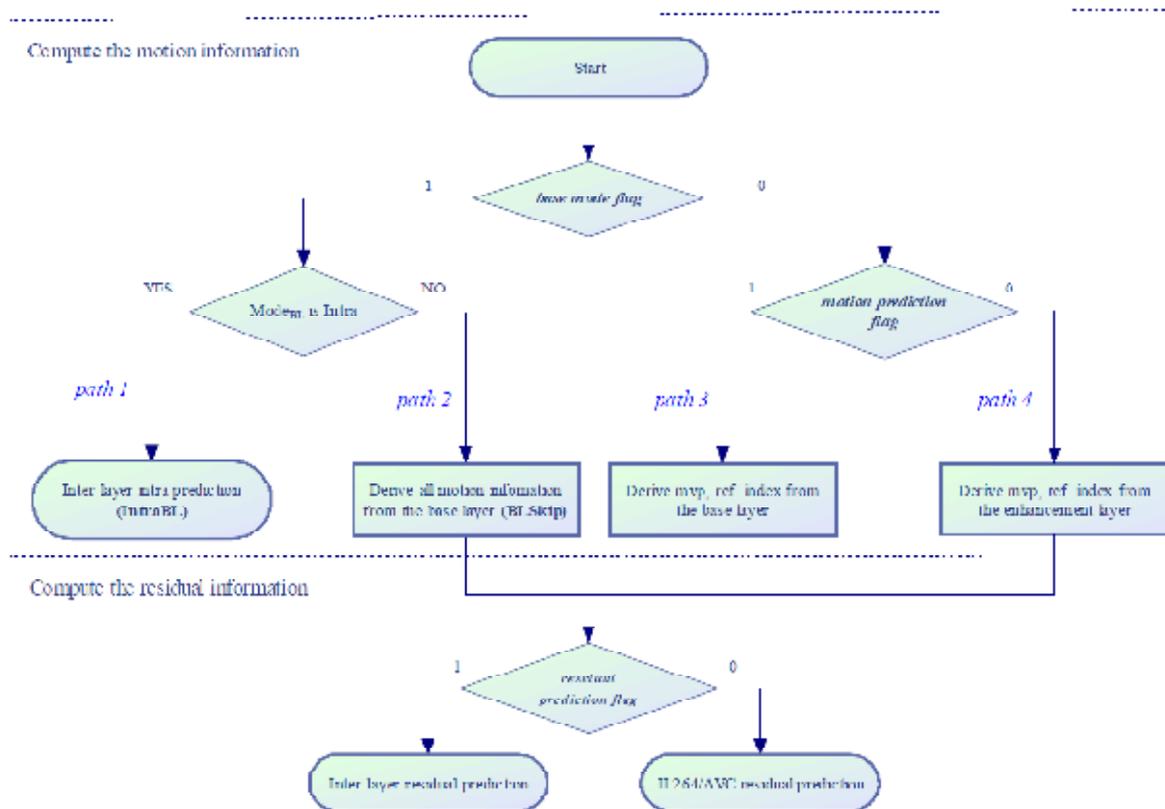


Figure 2.2: The flowchart of Inter-Layer Prediction Mechanism (from Lin's thesis)

## 2.1.2 Inter-Layer Prediction

為了改善EL的Rate-Distortion效能，發展了Inter-Layer Prediction，其目的是為了盡可能的利用較低層的資訊，一般而言，為了改善EL編碼效能，Inter-Layer Prediction依然需要和H.264/AVC中已有的mode比較。兩個Inter-Layer Prediction的觀念被加入在SVC中

- 預測MB的mode和動態參數。
- 預測剩餘訊號。

Fig. 2.2顯示了Inter-Layer Prediction機制的流程。在接下來的內容中，會以流程圖的方式介紹三個Inter-Layer Predictions機制。

## A. Inter-Layer Motion Prediction

當被參考層的MB是Inter-coded，且EL MB也是Inter-coded。在這案例，EL MB可能直接利用BL相對位置的MB的MV做為自己的MV，稱為Inter-Layer Motion Prediction。在Inter-Layer Motion Prediction，SVC包含兩種不同實作方式：

### BLSkip Mode:

是SVC中的一個新的mode type，在流程圖Figure 2.2中的path 2。在BL也是使用Inter coded的情況下，一個被稱為base mode flag的syntax element能表示是否使用這種mode。對於此mode type，所有的動態資訊(motion information)都從BL取得；只有剩餘訊號需要被傳輸，其他的資訊像是Inter Prediction modes或動態參數則直接取用BL的。

### Inter-Layer Motion Prediction:

當base mode flag等於0，且motion prediction flag等於1，就是流程圖Figure. 2.2. 中的path3。對於此mode type，EL的Motion Vector Predictors (MVP)是從BL而來，EL的參考索引也是從BL而來。



## B. Inter-Layer Intra-Prediction

當base mode flag等於1，且相對應的參考層是Intra-coded，這時的MB預測方式是流程圖Figure. 2.2的path 1，使用了參考層重建後的Intra-signal，做為現在的MB的預測值。此方法在Inter-Layer Prediction中被稱為IntraBL。

## C. Inter-Layer Residual Prediction

有一種flag被稱為residual prediction flag，被添加在EL的MB syntax，在流程圖Figure. 2.2的底部，表示了是否使用Inter-Layer Residual Prediction。它可以被運用在H.264/AVC本來就有的MB型態，也適用在SVC中新的MB型態，當residual prediction flag等於1時，參考層中對應位置MB的剩餘訊號可以幫助現在MB的預測，使只有剩餘訊號的差值需要被EL所編碼。

## 2.2. SVC Encoder Control

SVC中的Inter-Layer Prediction支援一次動態搜尋(Motion Compensation, MC)迴圈的單一迴圈的編碼，較低層圖片可以不用完全重建。而以往SVC中，編碼通常需要多迴圈的編碼，相較於單一迴圈的編碼，多了將參考層的residual signal解碼再實行Residual Prediction的動作，來避免編碼和解碼重建時照成的錯誤累積。

SVC的編碼控制，是以bottom-up process的方法，也就是BL會先於EL編碼。和單一層的H.264/AVC比較起來，EL通常會有不穩定的失真和編碼損失。bottom-up process將在Section 2.2.1.中介紹。在MLEC的R-D最佳化問題，可以在BL和EL之間的編碼效能做權衡，通常可以改善畫質可調性的效能，之後在Section 2.2.2中介紹，雖然MLEC的最佳化可以在BL和EL之間做權衡，但是它的計算時間明顯更多。

### 2.2.1 Bottom-up Encoder Control

JSVM [2]編碼控制內所描述的，就是在決定如何編碼時，先從最底層的開始決定，循序一層一層的往上。而該如何編碼的問題可以被公式化成以下多目標最佳化問題：

$$\begin{aligned} \min & D_i(p_i | p_{i-1} \dots p_0) \\ \text{s. t.} & \\ & R_0(p_0) + R_1(p_1 | p_0) + \dots + R_i(p_i | p_{i-1} \dots p_0) \leq R_{ci} \end{aligned} \quad (2.1)$$

$D_i$ 和 $R_i$ 分別表達JSVM中第 $i$ 層的Distortion和Rate， $R_{ci}$ 表示第 $i$ 層的最大限制Rate。對於每個access unit，首先，BL的編碼參數 $p_0$ 決定了接下來的widely-used Lagrangian approach，

$$p_0 = \arg \min_{\{p_0\}} D_0(p_0) + \lambda_0 \cdot R_0(p_0) \quad (2.2)$$

不會受到ELs的影響，各自的表達了在選擇 $p_0$ 參數向量後的Distortion和Rate， $\lambda_0$ 是Lagrange multiplier，它的選擇是基於所選擇的量化參數 $QP_0$ ，和BL相同，在每個EL的編碼參數 $p_i$ 是經由下面式子決定

$$\text{Min}_{\{p_i | p_{i-1} \dots p_0\}} D_i(p_i | p_{i-1} \dots p_0) + \lambda_i \cdot R_i(p_i | p_{i-1} \dots p_0) \quad (2.3)$$

若要決定ELi的參數，需要較低層參數 $p_{i-1}$  to  $p_0$ 都已經給定。當BL的編碼效能和單一層編碼一樣的時候，ELs的編碼通常會有些失真，因為選擇BL的編碼方式，通常只是針對BL做最佳化而不見得適用於EL的編碼。EL重複利用BL的編碼方式的效果有限，基於這個理由而權衡EL和BL的編碼效能，而這個方法將會在下面小節所提及。

## 2.2.2 Multi-layer Encoder Control

### A. Schwarz's MLEC

Explicit Bit Allocation (EBA)，是可以權衡動態資訊(motion information)和剩餘資訊(residual information)的方法，可以在[5]閱讀到所有資料。EBA的特性是位元可以在不同的區域根據關係強度作位移，使得整個視覺畫質可以被最佳化。從此觀點，在BL和EL之間的R-D效能可以視為兩個不同的區域，而在這兩個區域之間做權衡可以視為在SVC中的一個EBA問題，EBA問題將在以下介紹，也會討論發展MLEC的公式的理由。

為了克服BUEC的缺點，有一種編碼方法稱為EBA，而EBA在做畫質可調性的情況下，同時選擇BL和EL最佳化編碼參數的方法已經在[3]中被提出了。在不失一般性的前提下，修正了編碼的控制方法，這裡的公式只以兩個層型態來描述;但是仍然可以很輕易的把兩個層的型態轉成很多層的型態。在兩個層的型態，所有BL的決策都是以下面的式子為基礎，最小化下面的權重函數：

$$\begin{aligned} \text{Min}_{\{p_0, p_1|p_0\}} & (1 - w) \cdot (D_0(p_0) + \lambda_0 \cdot R_0(p_0)) \\ & + w \cdot (D_1(p_1|p_0) + \lambda_1 \cdot (R_0(p_0) + R_1(p_1|p_0))) \end{aligned} \quad (2.4)$$

Eq. (2.4)的第一項和第二項是從BUEC調整過來的，表示在BL和EL之間給於權重。權重因子 $w \in [0; 1]$ ，控制著BL和EL之間的權衡。為了讓R-D效能在 $w=0$ 時可以表現出BUEC的效能，BL再做決策時是以最小化Eq. (2.4)為基礎，但是EL在做決策時卻是以Eq. (2.3)為基礎，也就是回歸了BUEC。當 $w = 1$ 時，BL的參數只針對EL的參數最佳化，而沒有考慮到BL的資訊。 $P_i$ 可以視為在第 $i$ 層中相對位置的MB在決定好mode後的編碼參數，換個角度看，也可以被視為在做Motion Estimation後第 $i$ 層Mode Decision的MV參數。MLEC的動態偵測過程在Schwarz *et al.* in [3]中已有簡單的描述。對於Eq. (2.4)，要最小化Cartesian product space的參數 $P_0$ 和 $P_1$ 。MLEC的問題可以被轉換成以下敘述的多目標最佳化問題：

$$\begin{aligned} & \min(1 - w)D_0(p_0) + wD_1(p_1|p_0) \\ & s. t. \\ & (1) (1 - w) \times (R_0(p_0)) \leq R_B \\ & (2) (w) \times (R_0(p_0) + R_1(p_1|p_0)) \leq R_E \end{aligned} \quad (2.5)$$

$R_B$ 是BL所限制的最大Rate，而 $R_E$ 是EL所限制的最大Rate。在Eq. (2.5)中(2)的限制在 $w = 0$ 時沒有極限，但在 接近1時收斂到 $R_E$ 。這目標和限制會隨權重因子的不同而改變。但是我們直覺的認為應該修正此目標函數。因為這樣的限制並不直觀，使得在一些權重較高的例子得不到預期中的R-D效能，照成了不可預期的結果。

為了避免照成不可預期的R-D效能，Lin以原始公式為基礎，做適度的修正，以另一個公式表達MLEC想解決的問題，提出一個新的決策標準。在下一個小節中將會詳細的說明這些細節；另外同樣是MLEC的[7]中，也有對Motion Estimation/Mode Decision的式子做出修正的例子。

## B Lin's MLEC

因為Eq. (2.5)的多目標限制會因為 $w$ 的變化而有改變，為了避免此問題，*Lin et al.*為了MLEC發展了另一個RDO式子。在這個兩層的機制中，所提出的MLEC問題可用接下來的多目標最佳化問題公式化：

$$\begin{aligned}
 & \min(1 - w)D_0(p_0) + wD_1(p_1|p_0) \\
 & s. t. \\
 & (1) R_0(p_0) \leq R_B \\
 & (2) R_0(p_0) + R_1(p_1|p_0) \leq R_E
 \end{aligned} \tag{2.6}$$

$D_i$ 和 $R_i$ 分別是是第 $i$ 層的失真度和Rate， $R_B$ 是BL的目標Rate， $R_E$ 是EL的目標Rate，權重因子 $w \in [0;1]$ 控制BL和EL間的編碼效能。在此方程式中，權重因子 $w$ 並不包含在Rate的最佳化問題的限制中，但是權衡BL和EL的特性依然是存在的。如此一來，這MLEC的問題將可以避免因權重因子改變而影響到BUEC的R-D效能限制。因此，在決定BL和EL的mode時就以下面修正過的R-D cost函數為基礎：

$$\begin{aligned}
 & \text{Min}_{\{p_0, p_1|p_0\}} (1 - w) \cdot D_0(p_0) + \tilde{\lambda}_0 \cdot R_0(p_0) \\
 & + w \cdot D_1(p_1|p_0) + \tilde{\lambda}_1 \cdot (R_0(p_0) + R_1(p_1|p_0))
 \end{aligned} \tag{2.7}$$

在這最小化的過程中，EL並不用Eq. (2.3).來進行mode決策， $D_i(p_i)$ 和 $R_i(p_i)$ 表示了所選擇參數向量 $P_i$ 的Distortion和Rate。 $p_i$ 的編碼參數可以被視為第 $i$ 層現在的MB型態mode的參數，也可以被看成ME時在第 $i$ 層現在MB的動態向量。我們用以下公式簡化BL ME的過程，假設 $V_0 = V_1$ ：

$$\min_{\{v_0, v_1|v_0\}} (1 - w) \cdot D_0(v_0) + w \cdot D_1(v_0) + (\tilde{\lambda}_0 + \tilde{\lambda}_1) \cdot R_0(v_0) \tag{2.8}$$

在這方程式， $v_0$ 和 $v_1$ 分別表達BL和EL的MV， $D_i(v_0)$ 表示在第i層的原始訊號和預測訊號的Sum of Absolute Error (SAE)。當base mode flag 等於0時，且motion prediction flag 等於1，EL的MV是直接BL的MV，若以公式表示 $R_1(p_1|p_0) = R_1(v_1 - v_0)$ ，表示直接假設  $v_1 = v_0$ ，使得 $R_1(p_1|p_0) \approx R_1(0) \approx 0$ 。

在Eq. (2.7)中，一組Mode Pair一次決定一組MLEC的動態資訊。經由設定 $w = 0$ ，編碼控制只針對BL的編碼效能做最佳化，而完全不考慮EL的失真;也就是說， $R_1(p_1|p_0)$ 被選擇為0(EL mode是skip mode)，當 $w$ 等於1時，BL的參數只針對EL的編碼效能做最佳化而不考慮BL的重建失真，所以 $R_0(p_0)$ 被選擇為0(BL mode為Skip mode)。

## 2.3. Comparison and Summary

以下針對BLEC和MLEC的特性比較做總結:

- 在BLEC中，BL的編碼效能是以單一層為基礎，但EL總是會有些編碼效能的損失。
- 在BLEC中，EL只能有限的重複利用BL的資訊，因為BL所選擇的編碼參數只對BL做最佳化，不見得適合EL。
- BLEC是以frame等級的編碼過程，但MLEC是一對MB等級的編碼過程。
- 在MLEC，可以權衡BL和EL的編碼效能，但明顯是更為複雜編碼過程。

經由以上的總結，我們知道”BUEC”和”MLEC”間的不同，接下來我們對Schwarz’s MLEC和Lin’s MLEC做一些觀察，發現兩種MLEC的多目標問題公式雖不同，但解決方法卻是一樣的，我們做以下假設:

$$\begin{aligned}\tilde{\lambda}_0 &= (1 - w) \lambda_0 \\ \tilde{\lambda}_1 &= w \lambda_1\end{aligned}$$

由於 $\lambda$ 的最佳解唯一，上述式子表示(2.6)和(2.4)雖然不同，但 $\lambda$ 存在互通性。改成Schwarz’s MLEC在做Motion Estimation時，BL採用式子(2.4)來決定參數，而EL是採取(2.3)這種只考慮單一層編碼的式子，Lin’s MLEC的BL和EL都是以(2.8)決定編碼參數，看起來不同，但都是先決定BL mode，表示在決定EL mode時BL mode資訊都已經確定，所以在式子(2.8)就不必考慮BL mode的資訊，剩下部分剛好是(2.4)這式子，只需考慮EL的單一層編碼。根據以上觀察，雖然多目標問題公式不同，但我們可以將這兩種解決方法視為相同。

另外，因為Schwarz’s和Lin’s的演算法雖然可以有不錯的time saving，但是演算法穩定性卻是個問題，在Chapter 3中我們所提出的演算法將可以解決這個問題。

# Chapter 3. Fast Mode Decision Algorithm for SVC Multi-layer Encoder Control

本篇論文所提出的快速演算法建立在SNR可調性，兩個CGS層，沒有Inter8x4 Inter4x8，Inter4x4以及Intra8，這些mode如Table 3.1的內容，BL 共有7種mode，EL共有9種mode，在徹底式搜尋下共有56組Mode Pair。而所提出的演算法概觀大致上如Figure 3.1所示，我們會先取得EL的時間性mode分數(Temporal Mode Score, TMS)，依TMS選出最適當的mode做為Paired mode，然後進行兩段式演算法。當一對MB pair搜尋完成後，更新TMS的假想個數，一張frame搜尋完後再更新TMS的mode機率。我所提出之演算法架構，沒有dominant mode，而是利用EL資訊取得TMS並選擇配對的mode，而且當一張frame搜尋完後，更新mode機率，因此能隨sequence性質修正所選擇配對的mode，讓演算法具有適應性。

關於TMS的計算方法和mode機率分佈以及更新假想個數的方法將在Section 3.2.1 和Section 3.2.2.更詳細的介紹，Section 3.2.3除了介紹演算法內容外，也有對快速演算法的設計做出更詳細的說明。對於我即將提出的快速演算法，還有以下幾個關鍵的問題。

1. 如何選擇最適當的mode的initial probability?
2. 如何發展一套演算法可以符合所有的權重和sequence性質

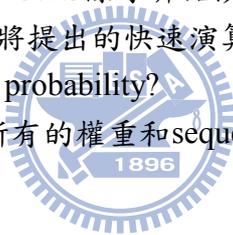


Table 3.1 test mode condition

Mode Types	Base Layer Modes	Enhancement Layer Modes
Inter MB Type	Skip Inter16x16 Inter16x8 Inter8x16 Inter8x8	Skip Inter16x16 Inter16x8 Inter8x16 Inter8x8
Intra MB Type	Intra16 Intra4	Intra16 Intra4
Inter-layer Prediction Type		BLSkip IntraBL

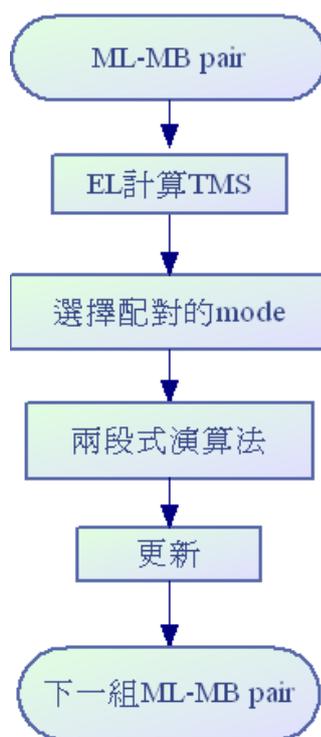


Figure 3.1 演算法概觀流程圖

### 3.1. Analysis of Mode Distribution

#### 3.1.1. Mode Distribution in MLEC

首先，MLEC在徹底式搜尋時是非常花時間的，雖然在我們給定權重因子之下，它使兩個SNR層之間有很好的編碼效能，在編碼後產生了一個包含兩個層的權重式的SVC Bitstream。測試的mode集合在Table 3.1. 從這個表格可以知道，MLEC在BL和EL組合起來的徹底式搜尋一共有56組Mode Pair。

Table 3.2 test video sequence

QCIF Sequence	4CIF Sequence
Soccer	Soccer
Foreman	Harbour
Footman	Crew
Mobile	
Crew	
Harbour	
Ice	

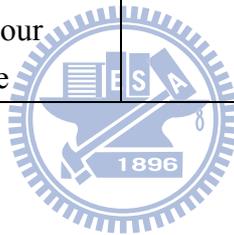


Table 3.3 training sequence

QCIF
Akiyo
Carphone
City
Coastguard
Highway
Stefan
Walk

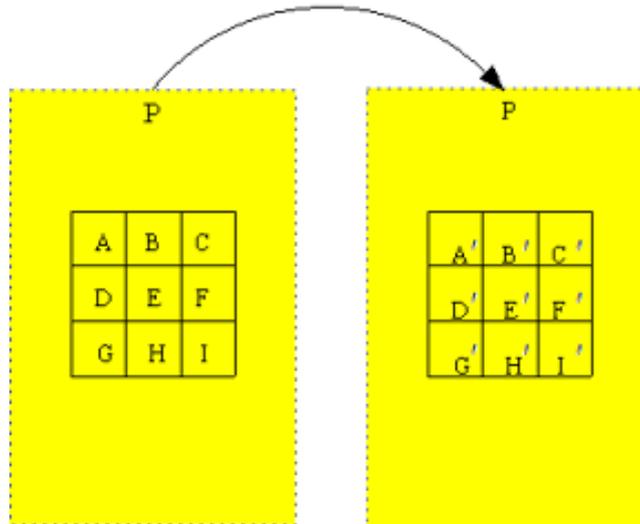


Figure 3.2 Directly stored modes in the enhancement layer

Table 3.4 location score in reference frame

1	2	1
2	4	2
1	2	1

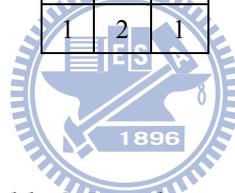


Table 3.5 mode score

mode	score
Skip mode	0
Inter16x16	1
Inter 16x8	2
Inter 8x16	2
Inter 8x8	4
BLSkip	Follow BL mode

Table 3.6 Example

Skip mode	16x16	BLSkip And BL mode 16x16
16x16(Intra4)	8x8	16x8
8x16	8x16(IntraBL)	16x16(Intra16)

第二，前一張frame的畫質特性和現在要編碼frame的畫質特性可能非常類似，如果討論到更局部，現在要編碼的MB和它co-located及其周圍MB的特性必定更為接近，基於這樣的原因，我們提出時間性mode分數(Temporal Mode Score, TMS)，其計算方式是利用前一張frame的co-located mode和其周圍的mode，計算TMS，請看Table3.4顯示了位置和權重的關係，Table3.5顯示不同mode的權重，如果要計算Figure 3.2的右邊那張frame的E的TMS，計算方法如下：

$$\text{TMS}(E') = 1 * A + 2 * B + 1 * C + 2 * D + 4 * E + 2 * F + 1 * G + 2 * H + 1 * I$$

其中前後兩張不同frame的A~I和A'~I'表示相對位置的MB，也就是前一張frame的E是現在這張frame的E' co-located MB。但如果有Intra mode的話，則需要做正規化(normalize)。這裡舉一個例子，請看Table 3.6，先不看括號內的mode，TMS的計算方式，就是

$$0 * 1 + 1 * 2 + 1 * 1 + 1 * 2 + 4 * 4 + 2 * 2 + 2 * 1 + 2 * 2 + 1 * 1 = 32$$

這時的TMS就是32，另一個例子，將位置D、H和I的MB改成使用括號內的Intra mode。

$$0 * 1 + 1 * 2 + 1 * 1 + 0 * 2 + 4 * 4 + 2 * 2 + 2 * 1 + 0 * 2 + 0 * 1 = 25$$

$$\text{Round}[25/11 * 16] = 36$$

這時的TMS就是36，是因為不考慮Intra mode的結果，直接對Inter mode的部分做正規化。而11的部分就是Inter mode部分的權重加總。以上述方法計算之下，我們可直覺的猜測，TMS的數字大小會影響mode分佈。為了證實這樣的想法，我們做了實驗，計算當TMS為多少時，某個mode的機率分佈，所以請看Figure 3.3。

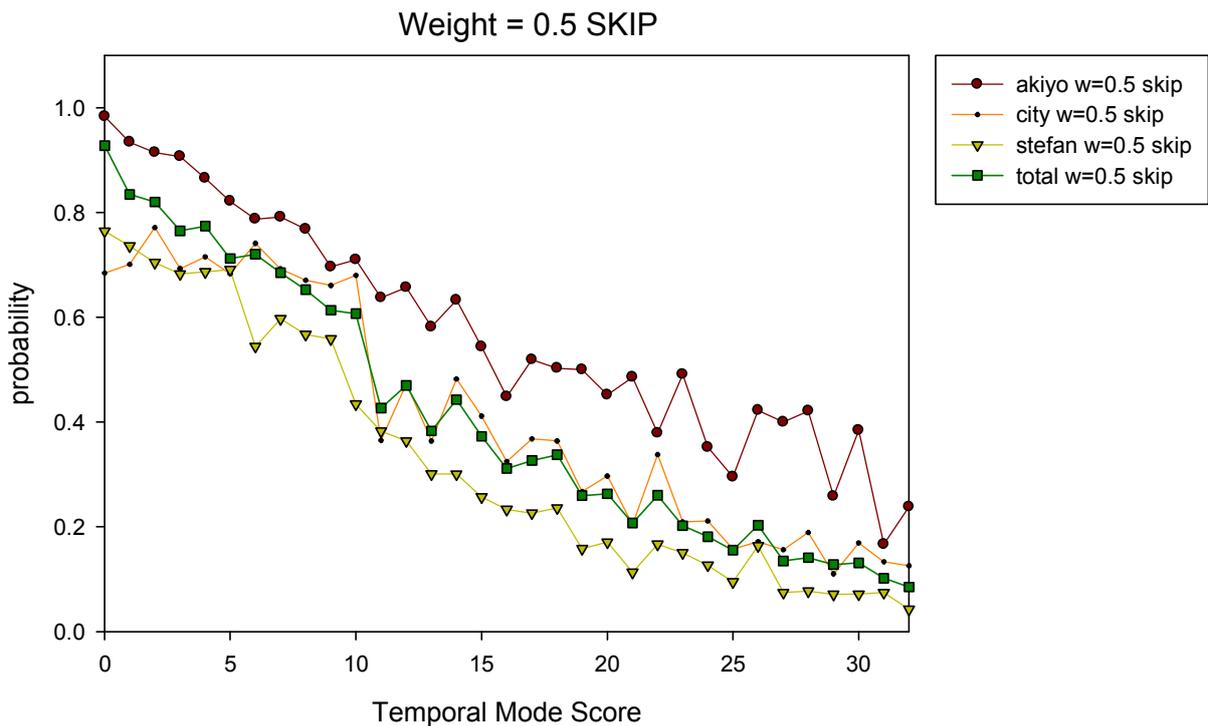


Figure 3.3 EL在不同的TMS時Mode Decision為Skip mode的機率

橫軸是TMS，縱軸是機率，這圖是顯示Skip mode在不同TMS時，Skip mode發生的機率，可以很容易的發現，當TMS漸漸變大時，Skip mode的機率也越來越小了；Figure 3.4是介紹Inter 16x16在不同的TMS時的機率分佈，觀察total的那條曲線，可看出Inter16x16在TMS由小到大的機率變化過程，先有一些起伏，然後再漸漸的降低，在Figure 3.4~7中，同樣觀察total 那條曲線，Inter16x8和Inter 8x16約在TMS = 32左右時有最高的機率，而Inter 8x8是隨TMS越高機率漸漸提升，可以看圖觀察之。BLSkip也是隨TMS越高，機率越來越高；而在這些個數機率分佈可以看Figure 3.10.在TMS在32~64時的mode個數只占了0~64個數的5%，表示了當TMS大於35後的機率參考價值不夠高,所以可以不討論這不部分的機率分佈。

而TMS介於[0~64]，屬於整數，又因為63完全不會出現，所以共有64種可能的temporal分數，每個TMS都會對應到一組mode的機率。觀察Figure 3.3~3.9後發現每種mode的趨勢都有隨TMS在變，所以我們可以加入TMS來設計演算法。

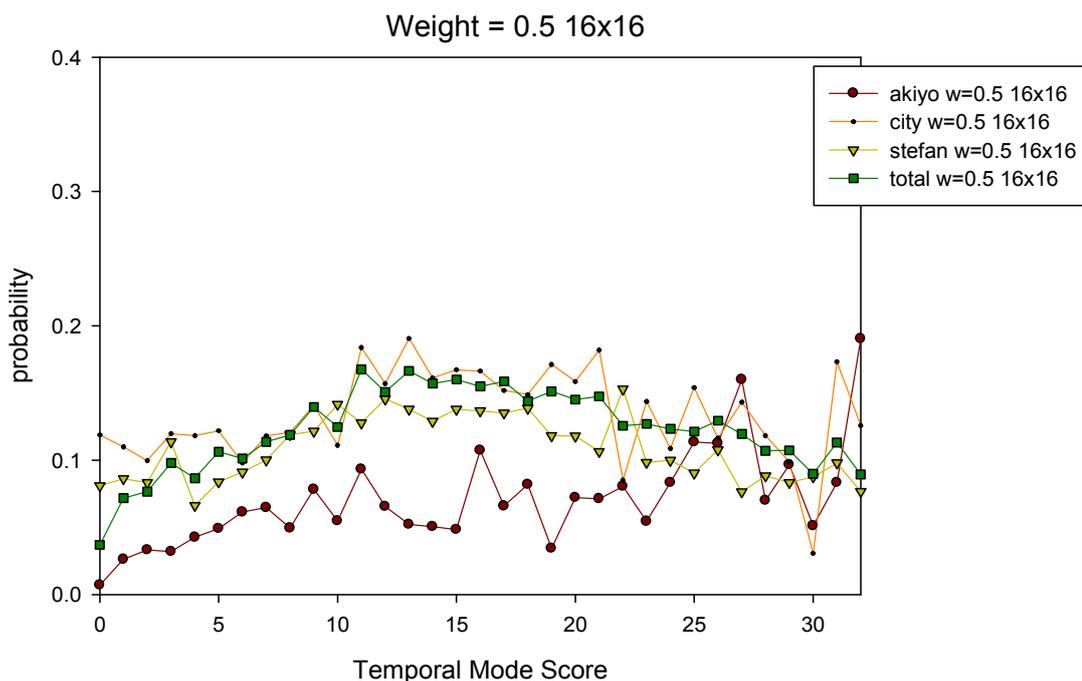


Figure 3.4 EL在不同的TMS時Mode Decision為Inter16x16的機率

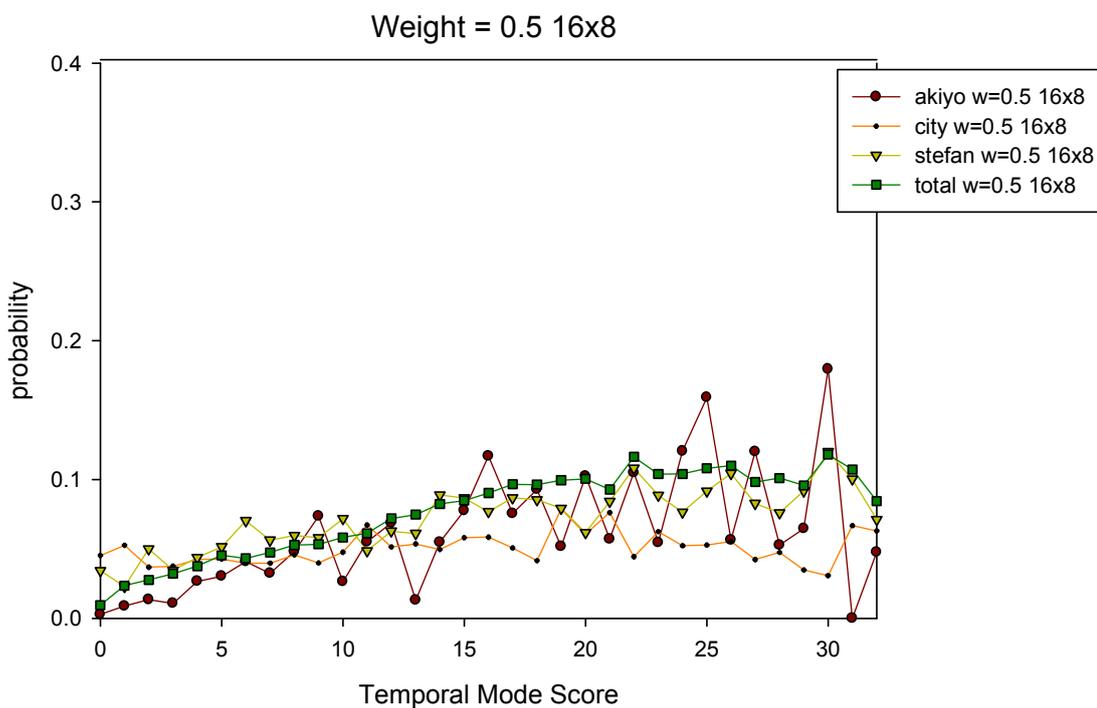


Figure 3.5 EL在不同的TMS時Mode Decision為Inter16x8的機率

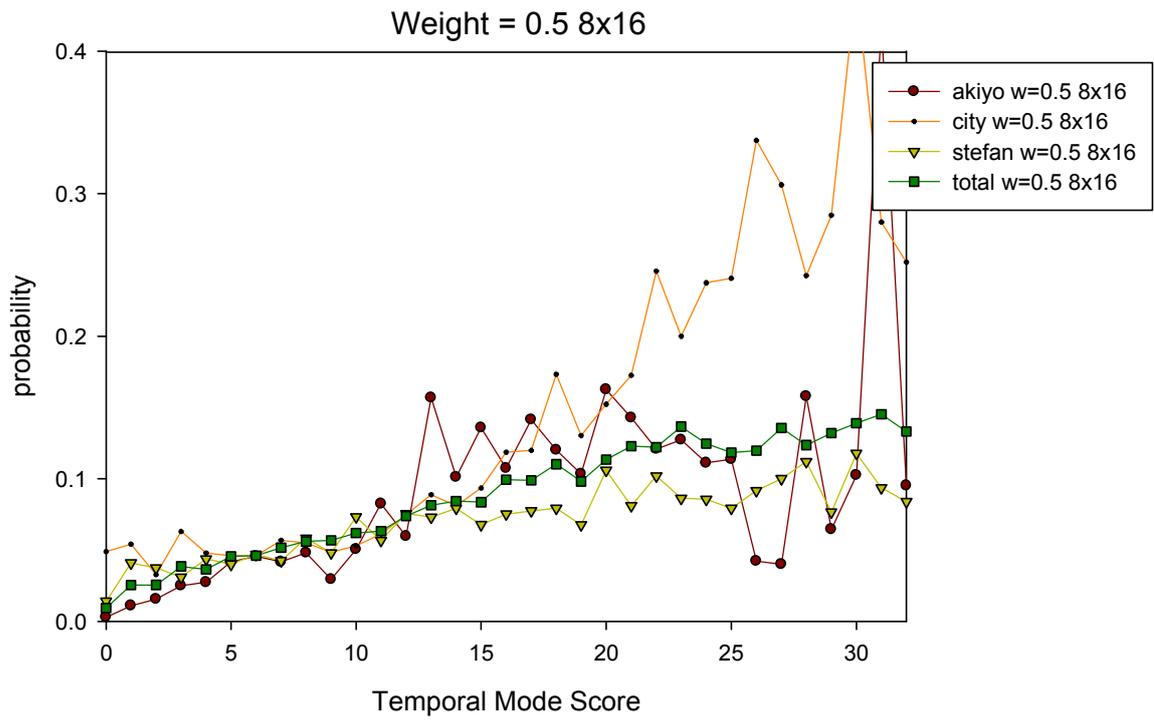


Figure 3.6 EL 在不同的 TMS 時 Mode Decision 為 Inter8x16 的機率

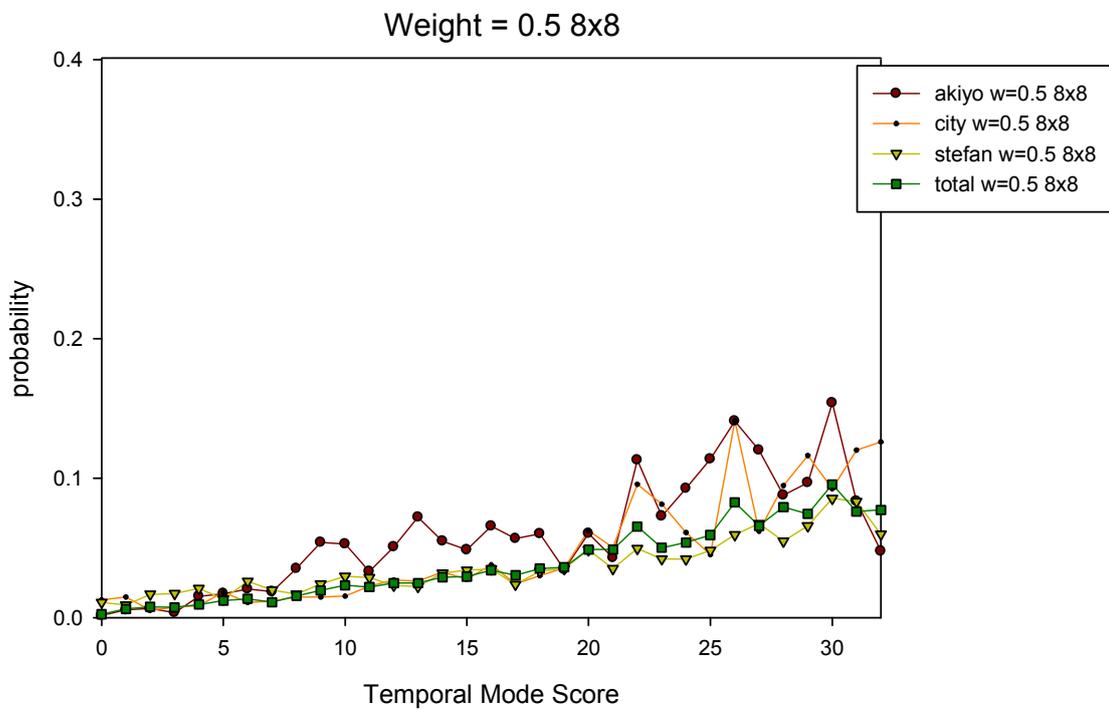


Figure 3.7 EL 在不同的 TMS 時 Mode Decision 為 Inter8x8 的機率

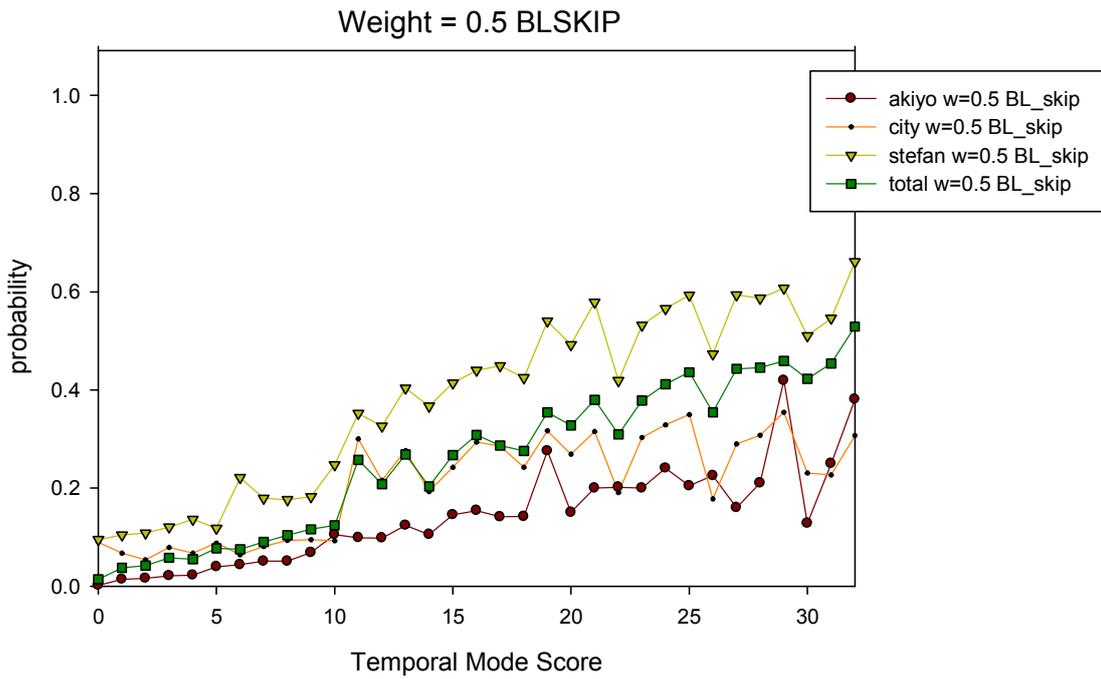


Figure 3.8 EL在不同的TMS時Mode Decision為BLSkip的機率

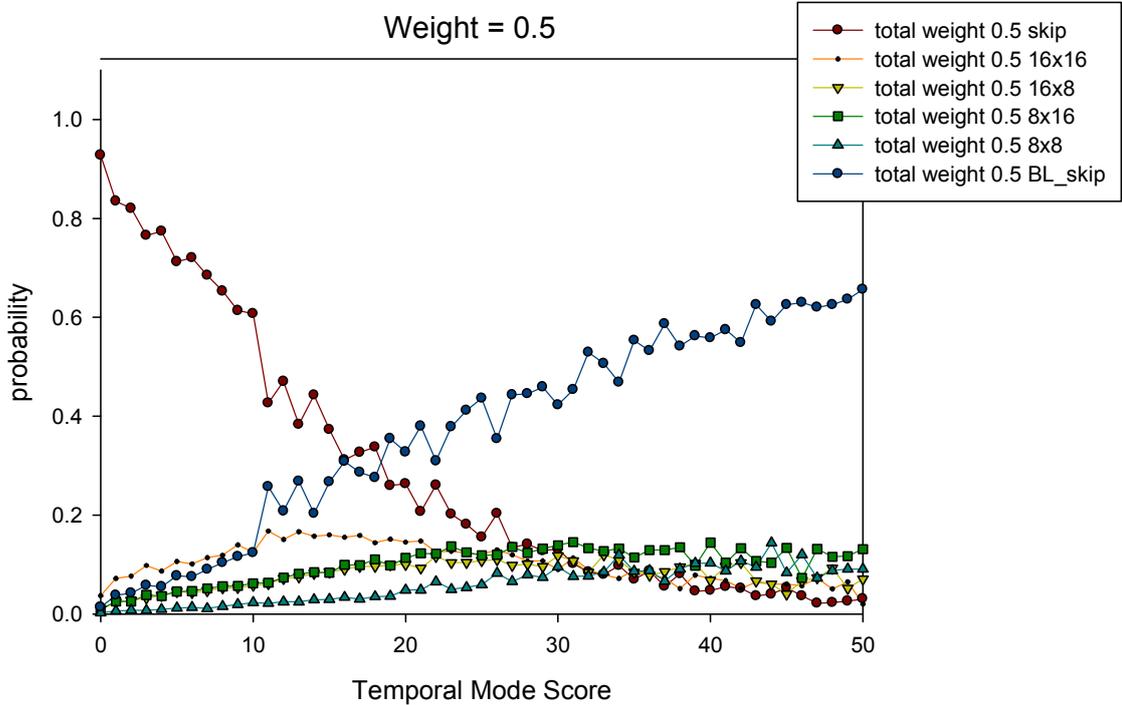


Figure 3.9 EL在不同的TMS時各種mode的機率

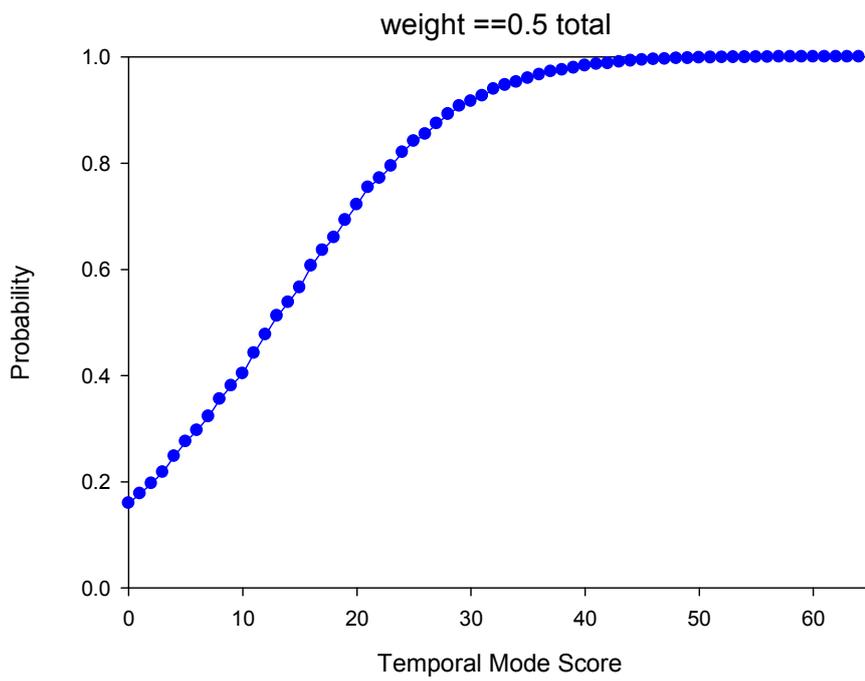
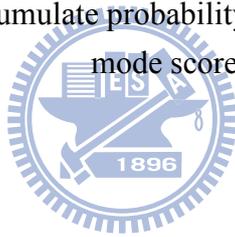


Figure 3.10 Macroblock accumulate probability in total sequence of each temporal mode score



### 3.1.2. Update of Mode Distribution Probability

由於我們的目的是做出一套能符合所有sequence的快速演算法，並能適應sequence性質，使快速演算法的R-D效能夠接近徹底式搜尋。觀察Figure 3.3~3.9，可以看出不同sequence間相同的mode隨TMS由小至大的mode distribution變化大致上是類似的，雖然有一些機率的差異，但是夠接近的，所以把TMS的initial probability設定為所有training sequence的平均機率分佈，Figure 3.9是權重為0.5時的例子，但這樣還不能隨sequence性質而改變mode distribution的特性，所以接下來將介紹如何進行更新mode distribution。

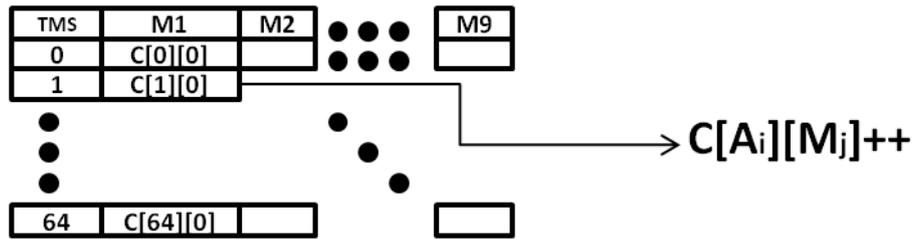


Figure 3.11 更新假想個數示意圖

$$p[A_i][M_j] = \frac{c[A_i][M_j]}{c[A_i][M_1] + c[A_i][M_2] + \dots + c[A_i][M_9]} \quad (3.1)$$

$$New\_p[A_i][M_j] = p[A_i][M_j] \times Initial\_Allocation$$

因為initial mode distribution，不見得貼近每個sequence的mode probability，所以這裡提出了一個方法來更新。除了能適當的維持本來的sequence特性，也可以應sequence性質變化而逐漸的改變了mode probability。這裡提出了”假想個數”的觀念，而詳細的更新步驟如下：

Step 1 :一開始每個TMS給定一個適當的初始配置，假設為50，而mode會被分配到初始配置乘上初始機率的假想個數。

Step 2 :當進行編碼時如Figure 3.11所示， $A_i$ 表示第*i*個TMS， $M_j$ 表示第*j*個mode， $c[A_i][M_j]$ 表示TMS為 $A_i$ 且mode為 $M_j$ 的假想個數，如果此時TMS為 $A_i$ ，編碼完決定是 $M_j$ ， $c[A_i][M_j]$ 就加1。

Step 3:一張frame編碼完時，如公式(3.1)所示，會重新計算所有TMS不同mode的機率分佈，也就是frame等級更新，把一個TMS的假想個數總數回歸到50(或依初始設定)。

這樣的方法，就可以不失去本來sequence的性質，又可以依sequence性質的變化更新。

### 3.2. Proposed Fast Mode Decision Algorithm

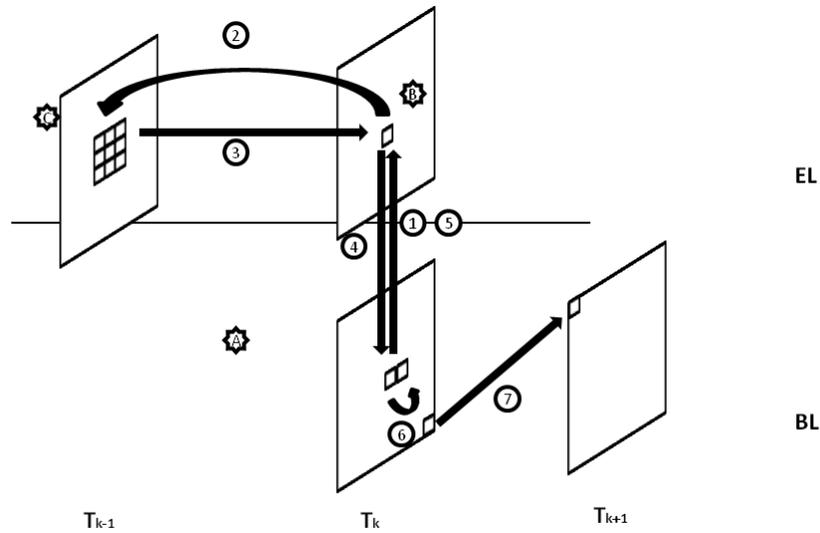


Figure 3.12 演算法的流程示意圖

Table 3.7 BL&EL mode 的配對

	EL	skip	BL_skip	Intra BL	...	Intra 4
BL						
Skip			EL: BL_skip BL: skip			
16x16	EL: skip BL: 16x16	EL: BL_skip BL: 16x16	EL: BL_skip BL: 16x16	EL: IntraBL BL: 16x16	...	EL: Intra 4 BL: 16x16
...			...			
Intra 4			EL: BL_skip BL: Intra 4			

因為SVC中MLEC是BL和EL的mode同時決定。所以再設計演算法時，預測配對組合的準確性，將是決定演算法優劣的重要因素。然而 EL做編碼時，所需的Rate通常較BL來的多，因此直覺地認為 EL mode預測的準確性又比 BL來的重要。而在設計演算法的過程中，的確得到這項證實。因此本論文所提出的快速演算法，是先從EL利用temporal的特性來預測mode。演算法的詳細流程如Figure 3.12，在本小節中將利用條列的方式來說明。首先定義Figure 3.12中的符號，A是BL中目前畫面的MB。B是EL中相對於A位置的MB。C是在EL中前一張畫面相對於B的九個MB位置。而step 1~7分別是演算法的步驟流程，步驟如下：

### Step 1

目前畫面若是第一張 P 畫面，因為前一張 I 畫面並沒有提供 TMS 資訊，所以執行完整搜尋，找出所有配對中 R-D cost 最小的做為編碼；否則在找 BL 中 MB 的 mode 時，必須先預測一個在 EL 相對位置 MB 的 mode 做配對。因此先去 EL 中預測出相對位置 MB 的 mode。

### Step 2、3

而在 EL 的配對是利用前一張畫面相對位置的 MB 以及其周圍 MB 的 mode 資訊去預測出一個最有可能的 mode。因此演算法提出一個時間性的編碼模式分數(Temporal Mode Score, TMS)，再利用此 TMS 去預測出配對的 mode。詳細做法上述已有提到，請參考 3.1.1。

### Step 4

利用上述的 TMS 找出 BL 在 EL 的一組 mode 配對，再利用這個配對去對 BL 所有的 mode type 做完整搜尋，找出 R-D Cost 最小的做為 BL mode。如 Table 3.7 中編號 1 的藍色框所示，若在 EL 利用 TMS 找到的 mode 預測為 BLskip，則去計算藍色框中所有配對的 R-D cost，並找出 R-D cost 最小的一組做為 BL 的 mode。

### Step 5

然後再利用 BL 的 mode 做為 EL 的一組配對。同樣計算 EL 中所有 mode，配對 BL 已選好 mode 的 R-D cost，最後將 R-D cost 最小的做為 EL 的 mode。如 Table 3.7 中編號 2 的紅色框所示。若在 step4 選定 16x16 為 BL 的 mode，再去計算紅色框裡所有 mode 配對的 R-D cost。並選出一組最小的做為 EL 的 mode。

### Step 6

若做完一組 MB pair 的編碼，則執行 TMS 中假想個數的累加，並判斷是否編碼完一張畫面，若是則跳到 step7；否則，回到 step2，繼續做下一對 MB pair。

### Step 7

執行更新的動作。更新所有 mode 在不同 TMS 下的機率分佈，詳細內容在 3.1.2 提及。並回到 step2。

### 3.3. Summary

這章節提供了一個新的快速演算法，以EL reference frame的temporal資訊計算TMS，並利用此TMS的統計特性選擇配對的mode，然後實行兩段式演算法，能達到BUEC的速度和接近徹底式搜尋的編碼效能。在每次的MB pair決定後，更新假想個數，一張frame編碼完後，更新mode機率，使演算法可適應多數特性的sequence，產生具穩定性的效能。為了證明所提出的快速演算法可適應多數特性的sequence且能產生具穩定性的效能。在接下來的章節，和其他在MLEC運作的快速演算法比較，其中包含了Schwarz's快速演算法和Lin's快速演算法。



# Chapter 4. Experiments and Analyses

## 4.1. Test Conditions

本篇論文所提供的快速決策演算法和徹底式搜尋，以及其他快速演算法，都在JSVM 9.12.2[6]中被實作，共有7個QCIF格式和3個4CIF格式的測試用sequence，其中包含了各種不同時間和空間特性。所有的實驗在2.8GHz Intel 虛擬八核心， i7 860 CPU， and 8GB 記憶體，而Table 4.1是測試條件的細節：

Table 4.1 :test condition

Prediction structure	IPPP
Inter-layer coding structures	CGS
resolution: frame per second	QCIF :15、7.5、3.75 4CIF :60
Frame number	QCIF-15-(150、130、120) QCIF-7.5-(75、65、60) QCIF-3.75-(37、32、30) 4CIF-60-30
Intra period	-1
Search range	32
Search mode	Table 3.1
QP	Fixed QP setting $QP_E = 22, 28, 34$ $QP_B = QP_E + 6$
Motion Estimation	Formula (2.6)
Entropy coding	CAVLC
Inter-Layer Residual Prediction	available

## 4.2. Comparison of Fast Mode Decision Algorithms

由於本論文所提出的演算法具可調性，為了證明它可以在大多數條件下進行良好的運作，所以接下來的實驗都在固定 QP 下進行比較，可以包含各種性質的編碼內容，而比較的演算法分別是 Lin 的快速演算法和 Schwarz *et al.* [3] 的快速演算法，這兩種演算法皆在 1.1 小節有詳細的介紹。接下來進行了以下實驗，並將這些比較各自做討論：

1. 在各種固定 QP 之下，設定很低的 frame rate 運行快速演算法
2. 進行大量的實驗，包含各種 QP，解析度和 frame rate，可參考 Table 4.1
3. 各種設定之下平均的 time saving

而在實驗中進行比較的方法如下：

$$\Delta PSNR = (\text{快速演算法的 } Y - PSNR) - (\text{徹底式搜尋的 } Y - PSNR)$$

$$\Delta Rate = [(\text{快速演算法的 Rate} - \text{徹底式搜尋的 Rate}) / \text{徹底式搜尋的 Rate}]$$

$$\times 100\%$$

$$\text{time saving} = [(\text{徹底式搜尋時間} - \text{快速演算法時間}) / \text{徹底式搜尋時間}] \times 100\%$$



### 4.2.1 Experiment 1

本論文所提的快速演算法在選擇 EL 配對的 mode 時，是利用 temporal 資訊進去預測。若當 temporal 資訊的相關性不夠，例如 frame rate 很低時，則此演算法是否還能利用 temporal 的資訊來預測 EL 配對的 mode。為了證明這點，在實驗 1 設定 testing sequence 的 frame rate 為 3.75，並考慮不同權重時的影響。且將本實驗的快速演算法和其他的快速演算法做比較，實驗結果在 Table 4.2 ~ 4.4。

Table 4.2 frame rate = 3.75 BL, EL QP = 28,22 average

Algo.	W	BL		EL	
		Rate	PSNR	Rate	PSNR
proposed	0	0	0	0	0
	0.25	1.36554	0.01032	0.82324	0.04585
	0.5	0.96161	0.038	0.57094	0.02323
	0.75	1.81138	0.27368	0.70137	0.0246
	1	0	0	0.00955	-0.0023
Lin	0	0	0	0	0
	0.25	-0.4855	-0.0998	0.65049	-0.0564
	0.5	1.8746	-0.256	2.60373	-0.224
	0.75	2.37178	-0.1393	2.67437	-0.2466
	1	0.0235	0	0.00275	0
Schwarz	0	0.45795	0.00378	0	0
	0.25	0.27025	-0.0182	-0.0032	0.02672
	0.5	0.35568	0.02162	-0.0069	0.00947
	0.75	0.95922	0.31287	-0.0347	-0.0041
	1	19.0574	0.57242	0.49511	-0.0005

Table 4.2~4.4 中顯示所有 testing sequence 的平均結果。由於實驗 1 是觀察 temporal 資訊相關性不足時，對 EL 所預測 mode 的影響，觀察 EL 的效能，由結果可以發現本論文所提供的演算法，其 $\Delta$ PSNR 和 $\Delta$ Rate 各別在-0.1db 以及 1 %之內，然而其他的演算法皆有超出此範圍。這裡為了讓讀者可以更容易的看出實驗結果的差異，本實驗將 $\Delta$ Rate 大於 1%和 $\Delta$ PSNR 小於 0.1db 的各別標上紅色和綠色的字體。所以在實驗結果可以看出，本演算法在 frame rate 很小時，還是能利用 temporal 的資訊來預測出配對的 mode，並且比其他兩種演算法有更好的效能。

Table 4.3 frame rate = 3.75 BL, EL QP = 34, 28 average

Algo.	W	BL		EL	
		Rate	PSNR	Rate	PSNR
proposed	0	0	0	0	0
	0.25	0.80199	0.03288	0.78872	0.05307
	0.5	1.1745	0.08638	0.39673	-0.004
	0.75	2.12879	0.30258	0.48461	-0.0311
	1	0	0	0.01717	0.00178
Lin	0	0	0	0	0
	0.25	-0.7878	-0.0866	0.87939	-0.0473
	0.5	1.74561	-0.1699	1.7372	-0.1813
	0.75	1.41335	-0.289	2.00933	-0.177
	1	0.09838	0.00067	-0.0089	-0.001
Schwarz	0	0	0	0	0
	0.25	0.72751	0.0721	3.31254	0.52952
	0.5	1.13033	0.17052	0.95834	0.11525
	0.75	2.22873	0.58032	0.37929	0.02445
	1	42.929	1.20445	2.45252	-0.0024

Table 4.4 frame rate = 3.75 BL, EL QP = 40, 34 average

Algo.	W	BL		EL	
		Rate	PSNR	Rate	PSNR
proposed	0	0	0	0	0
	0.25	0.41097	-0.012	0.64901	-0.0235
	0.5	1.93217	0.12255	0.56954	-0.0136
	0.75	0.67418	0.376	0.1672	-0.0159
	1	0	0	-0.0491	-0.0031
Lin	0	0	0	0	0
	0.25	-2.0389	-0.1354	0.98084	-0.0516
	0.5	1.54448	-0.1349	1.00985	-0.1255
	0.75	-0.2268	-0.0272	0.62247	-0.074
	1	0.08037	0.00023	0.00495	0.00042
Schwarz	0	0	0	0	0
	0.25	2.433	0.1402	7.24584	0.6944
	0.5	3.5708	0.34952	2.62716	0.17268
	0.75	10.2432	1.11533	0.97144	0.02358
	1	55.6968	1.62115	3.98235	0.00853

## 4.2.2 Experiment 2

有別於其他演算法，本論文提出的演算法具有可調整性，最重要的目的是使演算法能適應各種性質的 sequence，甚至在受到權重影響下能保有良好的 R-D 效能，所以進行大量的實驗，實驗設定為 Table 4.1，這裡擺出其中幾個具代表性的例子：

Table 4.5 frame rate = 15 BL, EL QP = 34, 28

Algo.	W	BL		EL	
		Rate	PSNR	Rate	PSNR
proposed	0	0	0	0	0
	0.25	0.42565	0.02387	0.5862	-0.0028
	0.5	1.3534	0.1186	0.34711	-0.0001
	0.75	1.21758	0.29581	-0.0772	-0.0083
	1	0	0	-0.001	-0.005
Lin	0	0	0	0	0
	0.25	-0.9715	-0.073	1.07948	0.05939
	0.5	1.03855	-0.0502	0.59915	-0.0242
	0.75	1.377	-0.0512	0.66825	-0.0429
	1	0.17946	-0.0033	0.04071	0.00189
Schwarz	0	0	0	0	0
	0.25	1.2176	0.06876	8.16225	0.97633
	0.5	1.32389	0.16707	2.74839	0.30409
	0.75	1.15166	0.58824	0.28262	0.09091
	1	28.8637	0.6791	2.72067	0.02494

觀察 Table 4.5，這是在運行了大量的數據後，觀察後歸納出最常見的例子。提出的演算法和 Lin's 演算法效能差不多，但 Schwarz's 演算法的 Rate 需求比徹底式搜尋多出不少。

Table 4.6 frame rate = 7.5 BL, EL QP = 34, 28

Algo.	W	BL		EL	
		Rate	PSNR	Rate	PSNR
proposed	0	0	0	0	0
	0.25	0.62801	0.0202	0.56364	0.00969
	0.5	1.00861	0.11049	0.09596	0.00391
	0.75	1.75192	0.33627	0.22185	-0.0058
	1	-0.012	-0.0001	0.12383	0.0025
Lin	0	0	0	0	0
	0.25	-0.7108	-0.0777	1.04037	-0.0456
	0.5	1.1943	-0.0952	1.07327	-0.0835
	0.75	1.94561	-0.1195	1.4452	-0.0945
	1	0.1249	-0.0003	0.07604	0.00167
Schwarz	0	0	0	0	0
	0.25	1.28928	0.07393	5.39084	0.73173
	0.5	1.096	0.18351	1.39064	0.18613
	0.75	2.01323	0.57317	0.37491	0.059
	1	48.1896	1.02899	2.68818	0.02233

Table 4.6 中 Lin's 演算法和 Schwarz's 演算法，在 R-D 效能上表現較差，但我所提出之演算法在 EL 依然有穩定的表現，這例子所產生的結果，在所觀察的數據中也是很常見。

我們分析以上兩個例子的命中率。在 Figure 4.1 中，以 hit ratio 和 R-D 效能比較，明顯 hit ratio 越高的 R-D 效能越接近徹底式搜尋。Figure 4.1 中可以看出演算法在 weight 1 和 0 時命中率幾乎接近百分之百而 weight 0.25~0.75 左右時命中率較低，明顯在 weight 0.25~0.75 之間時，因為編碼內容更難預測，使快速演算法在這設定下較難維持良好的 R-D 效能，但根據之前的觀察可以知道，所提出的演算法在這些情況下 EL 依然維持不錯的表現。

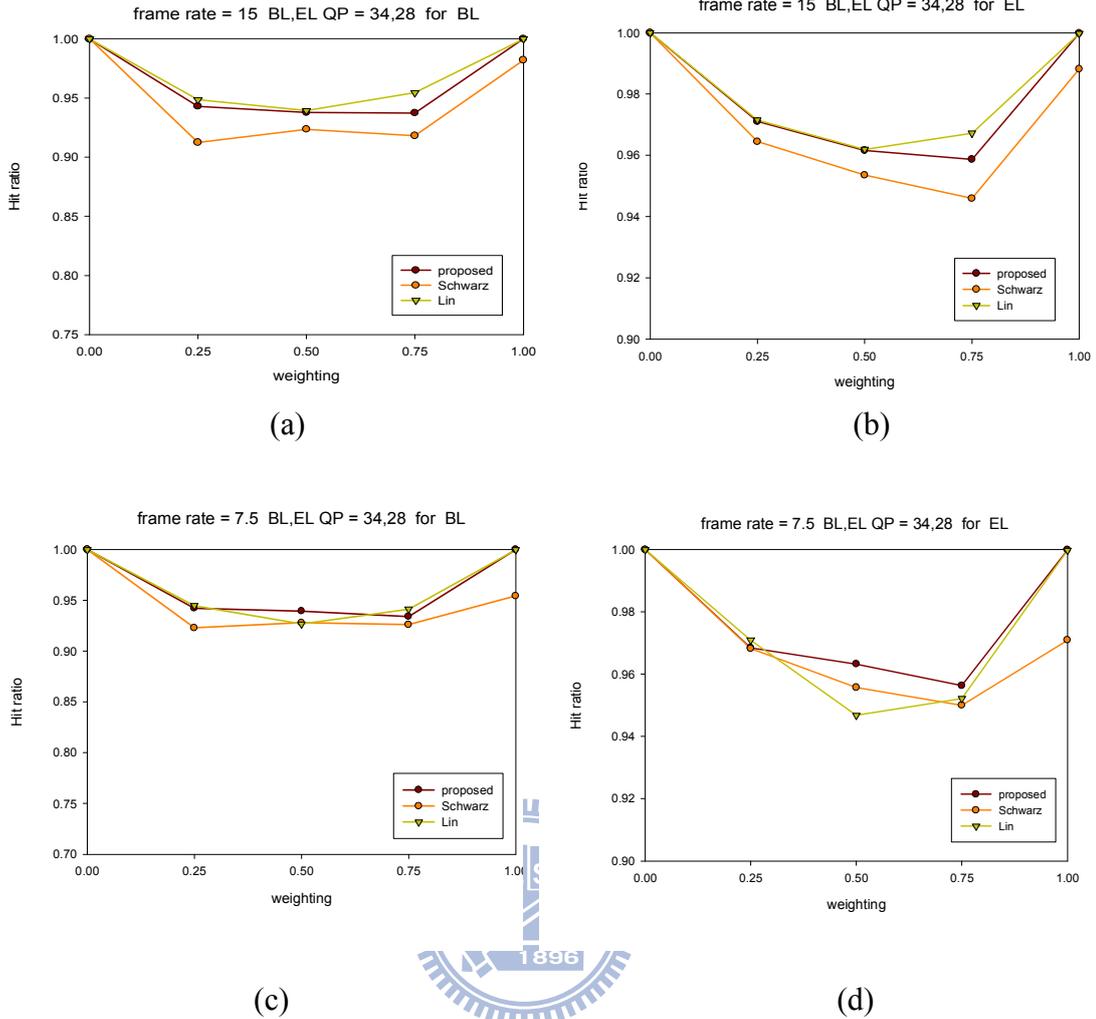


Figure 4.1(a)(b)(c)(d)各種 weight 之下的 hit ratio

但是本論文所提出之演算法，雖然在 EL 能有優良的 R-D 效能，但 BL 的表現依然和其他演算法一樣，有時候會有 Rate 大於 1% 的情況。因為兩段式演算法先決定 BL mode，但 BL mode 所搭配的 EL mode 很可能不是最適合的 mode，因此導致 BL 選擇到錯誤的 mode，使 BL R-D 效能不佳，這是兩段式演算法共同擁有的缺點。

### 4.2.3 Experiment 3

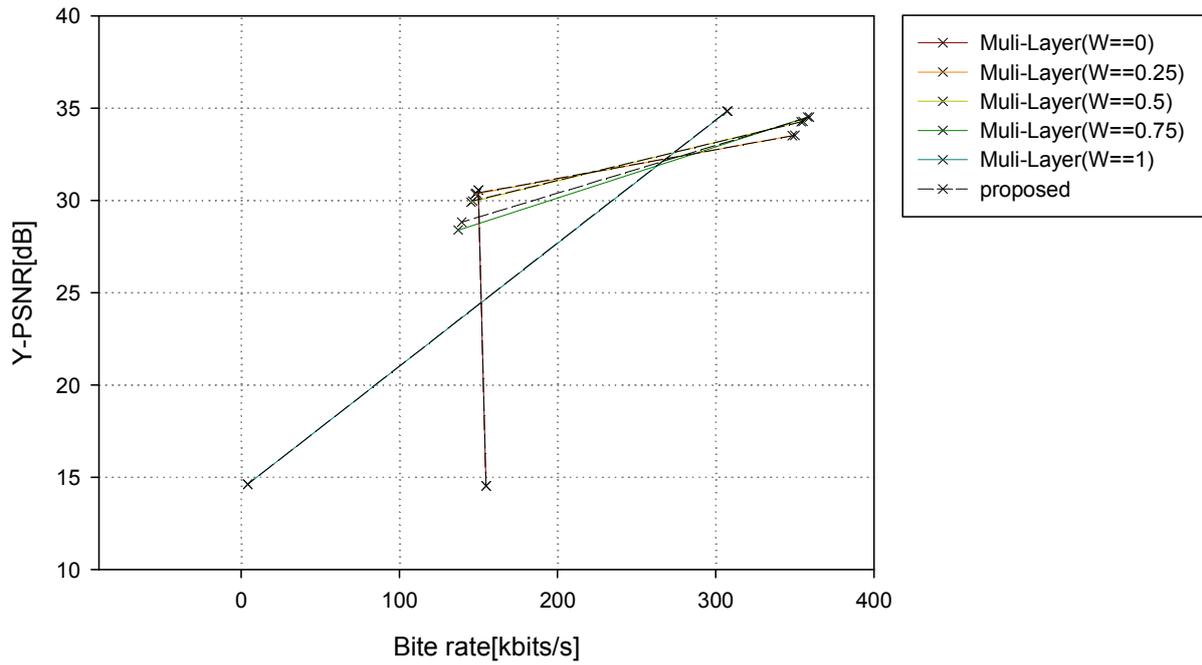
快速演算法除了要維持穩定的 R-D 效能，還必須有足夠的 time saving。Table 4.7 中是各種測試條件下的 time saving:

Table 4.7: 各種條件之下，testing sequence 的平均 Time Saving

Format /Frame Rate	QP <sub>B</sub> /QP <sub>E</sub>	Proposed	Lin	Schwarz
QCIF / 3.75	QP <sub>B</sub> :28, QP <sub>E</sub> :22	77.24514	74.9001	78.24258
	QP <sub>B</sub> :34, QP <sub>E</sub> :28	77.62963	76.48772	78.95085
	QP <sub>B</sub> :40, QP <sub>E</sub> :34	76.95936	77.52308	78.68517
QCIF / 7.5	QP <sub>B</sub> :28, QP <sub>E</sub> :22	76.5156	75.95822	77.884
	QP <sub>B</sub> :34, QP <sub>E</sub> :28	75.99884	76.86205	78.56844
	QP <sub>B</sub> :40, QP <sub>E</sub> :34	78.88088	79.07185	79.20268
QCIF / 15	QP <sub>B</sub> :28, QP <sub>E</sub> :22	74.24251	73.8715	81.25331
	QP <sub>B</sub> :34, QP <sub>E</sub> :28	78.50109	76.37538	82.97111
	QP <sub>B</sub> :40, QP <sub>E</sub> :34	77.67027	78.03382	81.56794
4CIF / 60	QP <sub>B</sub> :28, QP <sub>E</sub> :22	72.91371	71.75903	75.76203
	QP <sub>B</sub> :34, QP <sub>E</sub> :28	72.36085	71.52953	74.16106
	QP <sub>B</sub> :40, QP <sub>E</sub> :34	73.95595	73.46076	75.00302

從 Table 4.7 中，本論文提出之演算法平均 time saving 為 72%~78%，Lin's 演算法 71%~79%，Schwarz's 演算法 74%~82%。而相互比較來看，三者演算法的時間其實差不多，Schwarz's 的演算法稍快於其他兩種，本論文所提出的演算法和 Lin's 差不多。Figure 4.2~3.顯示了 testing sequence QCIF 格式的 football 和 Mobile 的 Rate-PSNR 曲線，包含了在不同的權重因子下，與徹底式搜尋 Rate-PSNR curve 的比較。從此可以證實本論文所提出的快速演算法和徹底式搜尋的結果非常接近。

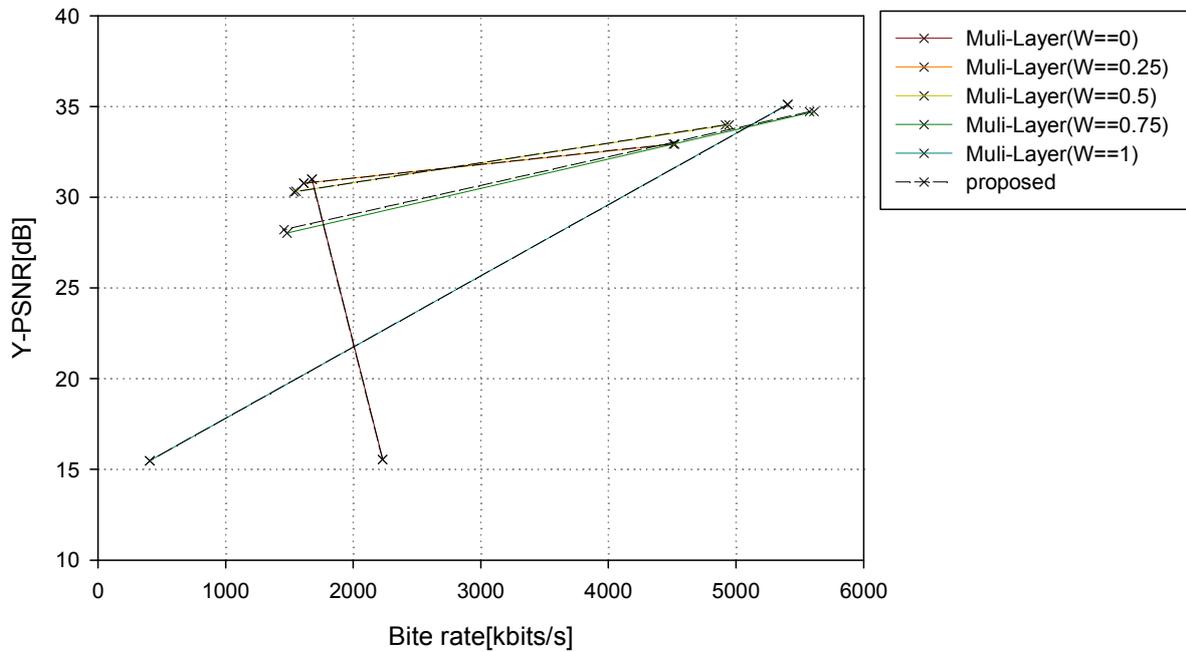
football(QCIF)



Exhaustive Search scheme vs. Fast Mode Decision scheme

Figure 4.2. The Rate-PSNR curve of QCIF sequence

soccer(4CIF)



Exhaustive Search scheme vs. Fast Mode Decision scheme

Figure 4.3. The Rate-PSNR curve of 4CIF sequence

# Chapter 5. Theoretical Analysis on SVC

## Rate-Distortion Function

在上面的章節提出了一套依統計方法而擬定出來的快速決策演算法，可在 R-D 效能接近徹底式搜尋的情況下進行加速，減少編碼的時間。然而不少其他快速決策演算法也能達到這樣的要求，但這些演算法通常缺乏了理論依據，能得其效果卻不知其所以然。所以想跳脫傳統在做快速決策演算法時，幾乎只以統計分析為立論的思考模式。這章節針對分割模式做理論分析，使我們清楚理解影響分割模式決策的因素，如果能完成這個分析，便可運用在 H.264/AVC 和 SVC 的 Mode Decision，甚至將觀念延伸到 MLEC。

### 5.1. Theoretical Model

H.264/AVC 做 Mode Decision 時，以 Lagrangian Cost 為決策標準，選擇  $\min\{D(p) + \lambda R(p)\}$  的 mode 進行編碼。討論普遍 sequence 的 Mode Decision，當 QP 很大時表示畫質上的細節可以省略，而 QP 較小時會保存更多的細微部分，所以同一個 MB 在 QP 不同時可能選擇不一樣的 mode 編碼，我們期待在做理論分析時可以看到這樣的變化。而以 H.264 為基礎延伸的 SVC 加入了對 Mode Decision 有影響的 Inter-Layer Prediction 機制，由於 EL 和 BL 的 QP 變化會改變 residual signal 的相關性，決定是否實行 Inter-Layer Residual Prediction。BL mode 和 EL mode 是否相同也會改變 residual signal 相關性。期待在理論分析中發現 EL 和 BL 的 QP 差值和 BL mode 在實行 Inter-Layer Residual Prediction 時會影響 EL 的 Mode Decision。

#### 5.1.1. Theoretical R-D functions of Single-Layer Hybrid Coding

如 Figure 5.1 所示，H.264/AVC 的 Block-Based Hybrid Video Coding 架構在對 Inter mode 進行編碼時，先從 reference frame 取得預測訊號並計算出 residual signal，然後傳輸 residual signal 經過 DCT 和 quantization 後的 coefficient。我們在理論分析時模擬這些步驟，5.1.1 中的 Step 1 由於理論分析時不能直接使用實際的 residual signal，但需要 residual 的變異程度，所以這步驟推導了 Inter mode 的 spatial domain residual covariance model，因為需要 DCT domain residual variance，Step 2 介紹 covariance 的 DCT 轉換方法，Step 3 建立 R-D 間關係，加入 QP 影響，並說明 R-D 計算方法。以 H.264 為基礎延伸的 SVC 在 Inter Prediction 時多了 Inter-Layer Prediction 機制，將在 5.1.2 討論 Inter-Layer Prediction 對 residual covariance model 的影響。在 Table 5.1 定義了理論分析需要的一些 symbol。

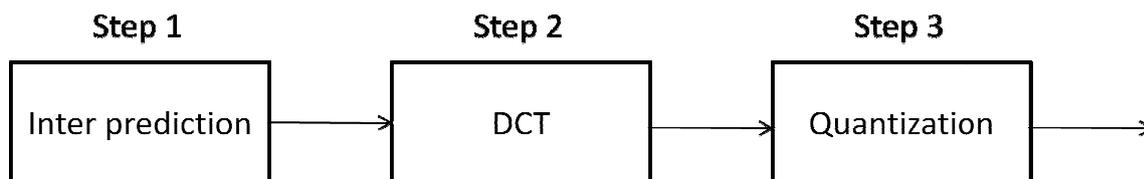


Figure 5.1 Inter Prediction 的 residual signal 編碼流程圖

Table 5.1 symbol definition

$I$	intensity of current frame
$\tilde{I}$	intensity of reference frame
$n_B(s)$	quantization noise at point $s$ in current frame in base layer
$\tilde{n}_E(s)$	quantization noise at point $s$ in reference frame in enhancement layer
$\tilde{n}_B(s)$	quantization noise at point $s$ in reference frame in base layer
$v_B$	block motion of current MB mode in base layer
$v_E$	block motion of current MB mode in enhancement layer
$v(s)$	true motion of $s$ point
$s_i$	two dimensions position, $s_i = (r_i, c_i)$
$s_c$	block center
$\sigma_I^2$	intensity variance
$\sigma_m^2$	motion variance
$\sigma_n^2$	noise variance in Single-Layer
$\sigma_L^2$	noise variance in L Layer
$\rho_m^{\ s_i - s_j\ _1}$	the correlation between true motion of $s_i$ and $s_j$
$\rho_n^{\ s_i - s_j\ _1}$	the correlation between quantization noise of $s_i$ and $s_j$
$\rho_{n'}^{\ s_i - s_j\ _1}$	the correlation between quantization noise of $s_i$ in EL and $s_j$ in BL
$r(s)$	the residual signal of $s$ point
$D(s)$	the distortion of $s$ point

**Step 1:**

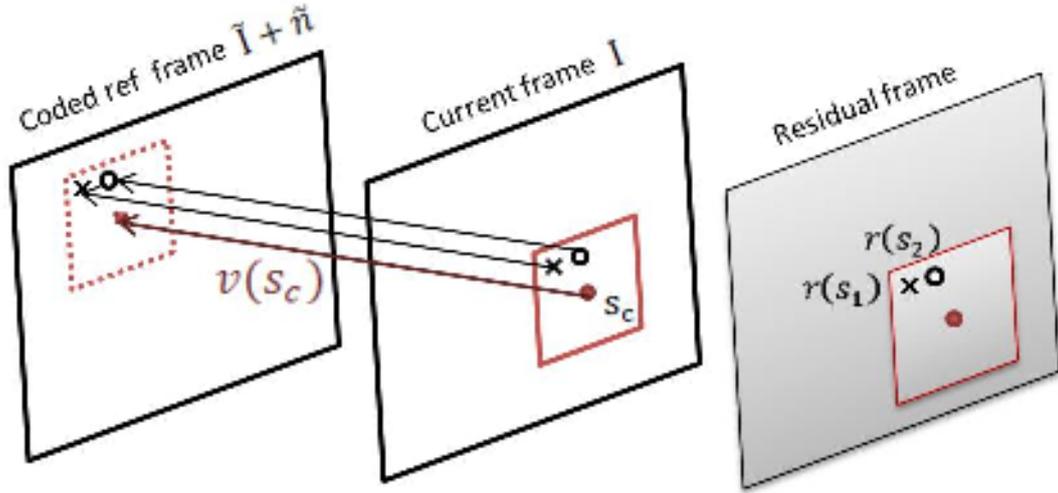


Figure 5.2 H.264/AVC 中 residual signal 取得方法示意圖

由於後面步驟需要計算 Rate-Distortion (R-D) 間的關係，所使用的 R-D function 需要 MB 中不同位置的 residual variance，本論文針對 Inter mode 做理論分析，所以利用 *Tao et al.*[10] 所提供的公式推導了 spatial domain residual covariance

$$\begin{aligned}
 & E[r(s_1)r(s_2)] \\
 &= E[(I(s_1) - \tilde{I}(s_1 + v(s_c)) - \tilde{n}(s_1 + v(s_c)))(I(s_2) - \tilde{I}(s_2 + v(s_c)) - \tilde{n}(s_2 + v(s_c)))] \\
 &= E[(\tilde{I}(s_1 + v(s_1)) - \tilde{I}(s_1 + v(s_c)) - \tilde{n}(s_1 + v(s_c))) \\
 &\quad (\tilde{I}(s_2 + v(s_2)) - \tilde{I}(s_2 + v(s_c)) - \tilde{n}(s_2 + v(s_c)))] \\
 &= E_M[E_I[(\tilde{I}(s_1 + v(s_1)) - \tilde{I}(s_1 + v(s_c)) - \tilde{n}(s_1 + v(s_c))) \\
 &\quad (\tilde{I}(s_2 + v(s_2)) - \tilde{I}(s_2 + v(s_c)) - \tilde{n}(s_2 + v(s_c)))] \\
 &= E_M[R_{\tilde{I}}(s_1 + v(s_1), s_2 + v(s_2)) - R_{\tilde{I}}(s_1 + v(s_c), s_2 + v(s_2)) \\
 &\quad - R_{\tilde{I}}(s_1 + v(s_1), s_2 + v(s_c)) + R_{\tilde{I}}(s_1 + v(s_c), s_2 + v(s_c))] + E[\tilde{n}(s_1 + v(s_c))\tilde{n}(s_2 + v(s_c))] \\
 &= \underbrace{\frac{4\sigma_I^2\sigma_m^2}{K}(1 + \rho_m^{\|s_1-s_2\|_1} - \rho_m^{\|s_1-s_c\|_1} - \rho_m^{\|s_2-s_c\|_1})}_{(A)} + \underbrace{\rho_n^{\|s_1-s_2\|_1}\sigma_n^2}_{(B)}
 \end{aligned} \tag{5.1}$$

公式(5.1)中如 Figure 5.2 中所示，第一和第二個等式是由於我們以下的假設：

$$r(s_i) = I(s_i) - \tilde{I}(s_i + v(s_c)) - \tilde{n}(s_i + v(s_c)) \tag{5.2}$$

$$I(s_i) = \tilde{I}(s_i + v(s_i)) \tag{5.3}$$

公式(5.2)中 residual signal 是原始訊號扣去預測訊號的差值，由於 reference frame 提供的是重建後的 signal，所以需考慮 noise 的影響。公式(5.3)是根據 true motion 的定義。而  $I(s_i), v(s_i)$  是隨機變數；根據 Law of Total Expectation，第三個等式中先對 intensity 計算期望值後再對 motion 計算期望值；第四個等式中假設 intensity 和 noise 之間為獨立並根據 *Tao et al*[10]對兩個 intensity 間相關性的假設

$$E[\tilde{I}(s_i)\tilde{I}(s_j)] = R_{\tilde{I}}(s_i, s_j) = \sigma_I^2 \left(1 - \frac{\|s_i - s_j\|_2^2}{K}\right) \quad (5.4)$$

公式(5.1)第五個等式的計算過程繁複，這裡舉出其中一個例子，其他部分可利用相同觀念計算得出

$$\begin{aligned} & E_M[R_{\tilde{I}}(s_1 + v(s_1), s_2 + v(s_2))] \\ &= E_M\left[\sigma_I^2 \left(1 - \frac{\|s_1 + v(s_1) - s_2 + v(s_2)\|_2^2}{K}\right)\right] \\ &= E_M\left[\sigma_I^2 \left(1 - \frac{(v_x(s_1) - v_x(s_2) - x)^2 + (v_y(s_1) - v_y(s_2) - y)^2}{K}\right)\right] \\ &= \sigma_I^2 \left\{1 - \frac{1}{K} (\|s_1 - s_2\|_2^2 - 2xE_M[v_x(s_1)] + 2xE_M[v_y(s_2)]) \right. \\ &\quad \left. + 2yE_M[v_x(s_1)] - 2yE_M[v_y(s_2)] + 2\sigma_m^2 - 4\rho_m^{\|s_1 - s_2\|} \sigma_m^2 + 2\sigma_m^2\right\} \\ &= \sigma_I^2 \left(1 - \frac{\|s_1 - s_2\|_2^2 + 4\sigma_m^2 - 4\rho_m^{\|s_1 - s_2\|} \sigma_m^2}{K}\right) \end{aligned} \quad (5.5)$$

公式(5.5)第一個等式是根據公式(5.4)，第二個等式將式子改成二維空間表示法以利於計算，假設  $s_1 - s_2 = (x, y)$ ，且  $(x, y)$  為定值， $v(s_i) = (v_x(s_i), v_y(s_i))$  其中  $v_x(s_i), v_y(s_i)$  為隨機變數，第三個等式是根據 *Tao et al.*[10]對兩個 motion 間相關性的假設如下：

$$E[v(s_i)v(s_j)] = R_{vv}(s_i, s_j) = \sigma_m^2 \rho_m^{\|s_i - s_j\|} \quad (5.6)$$

因為 motion vector 的期望值為 0，所以可簡化成第四個等式。在公式(5.1)第五個等式中還有對  $s_1$  和  $s_2$  兩點 quantization error 之間的 correlation 做假設

$$E[\tilde{n}(s_i + v(s_c))\tilde{n}(s_j + v(s_c))] = \rho_n^{\|s_i - s_j\|} \sigma_n^2 \quad (5.7)$$

會做這樣的假設是因為 quantization error 會因為距離而影響相關性，根據以上結果和觀念可以得到公式(5.1)的結論，而其中的 quantization error 的 variance 假設為

$$\sigma_n^2 = \frac{Qstep^2}{12} \quad (5.8)$$

式子(5.8)在[11]中是 scalar quantizer 在 uniform quantization 時 quantization error 和 Qstep 之間的關係式，雖然和實際情況有所差異，但暫時使用這關係式，未來可能對這方面作出一些修正。

## Step 2:

本論文使用 DCT 而不用最理想的 KLT(Karhunen-Loeve Transform)，因為經由公式(5.1)計算出來的 spatial domain residual covariance 經由這兩種方法轉換後的 frequency domain residual variance，其值非常接近，選擇哪種方法並不影響理論分析，所以我們依然和 H.264/AVC 一樣使用 DCT 進行理論分析。將 4x4 的 spatial domain residual 和 DCT domain residual 以 raster scan order 表示如下：

$$S = \begin{bmatrix} r(s_1) \\ r(s_2) \\ \vdots \\ r(s_{16}) \end{bmatrix}, D = \begin{bmatrix} r'(s_1) \\ r'(s_2) \\ \vdots \\ r'(s_{16}) \end{bmatrix} \quad (5.9)$$

$r_i$ 和 $\tilde{r}_i$ 分別表示 spatial domain 和 DCT domain 第 i 位置的 coefficient，S 和 D 表示 16x1 的 matrix，如果我們對 S 做一維的 DCT 轉換會有以下關係：

$$US=D \Rightarrow U \begin{bmatrix} r(s_1) \\ r(s_2) \\ \vdots \\ r(s_{16}) \end{bmatrix} = \begin{bmatrix} r'(s_1) \\ r'(s_2) \\ \vdots \\ r'(s_{16}) \end{bmatrix} \quad (5.10)$$

U 是一個 16x16 matrix，其計算方法是利用 2 維的 4x4 DCT matrix 以 Kronecker Product 計算後取得，而根據[11]中的證明結果，利用以下算式便可將公式(5.1)計算出的 spatial domain residual covariance 轉換成 DCT domain residual covariance

$$U \times E \begin{pmatrix} \begin{bmatrix} r(s_1) \\ r(s_2) \\ \vdots \\ r(s_{16}) \end{bmatrix} \\ [r(s_1) \quad r(s_2) \quad \cdots \quad r(s_{16})] \end{pmatrix} \times U^T = E \begin{pmatrix} \begin{bmatrix} r'(s_1) \\ r'(s_2) \\ \vdots \\ r'(s_{16}) \end{bmatrix} \\ [r'(s_1) \quad r'(s_2) \quad \cdots \quad r'(s_{16})] \end{pmatrix}$$

(5.11)

### Step 3:

Single-Layer Coding 編碼需要計算某 QP 時一個 MB 的 Rate 和 Distortion，所以理論分析也需要一個計算 MB Rate 和 Distortion 的方法，這裡利用倒水定理，倒水定理的內容和證明在[11]中有更詳細的說明，其公式如下：

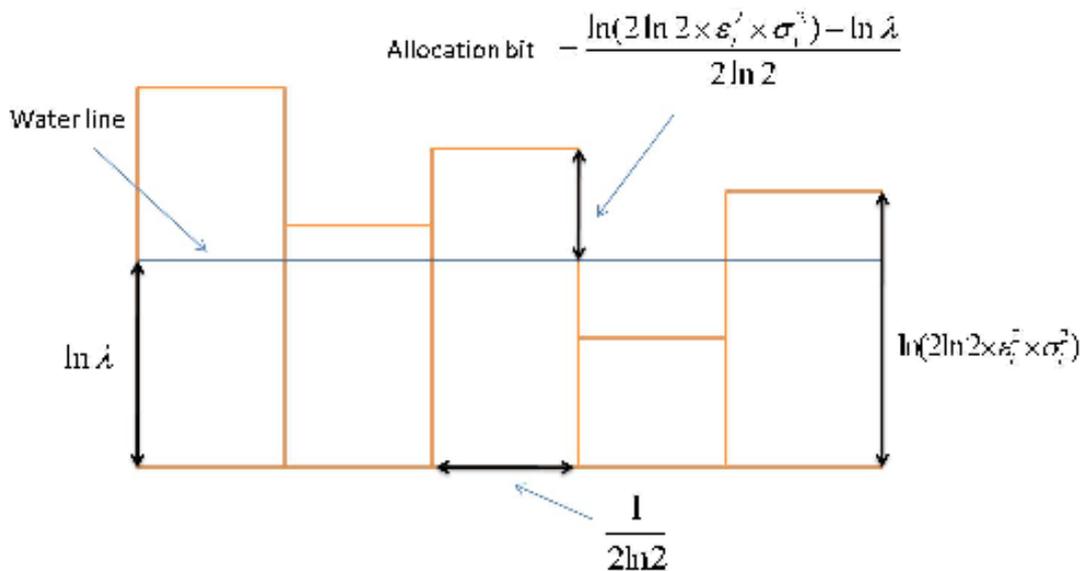


Figure 5.3 倒水定理概念圖

$$r(s_i) = \begin{cases} \frac{\ln(2 \ln 2 \times \sigma_i^2 \times \varepsilon_i^2) - \ln \lambda}{2 \ln 2}, & \text{if } \lambda \leq 2 \ln 2 \times \varepsilon_i^2 \times \sigma_i^2 \\ 0, & \text{if } \lambda > 2 \ln 2 \times \varepsilon_i^2 \times \sigma_i^2 \end{cases} \quad (5.12)$$

假設在公式(5.12)中已經知道所有的 $\sigma_i^2$ ，而調整係數 $\varepsilon_i^2$ 會隨 coefficient 的 distribution 而改變，這時只要調整 $\lambda$ ，就可以改變每個 coefficient 的位元配置，如 Figure 5.3，調整 $\lambda$ 就如同調整水位線，水位以上的 coefficient 才需要分配位元，水位以下的不分配位元。得到每個 coefficient 的 Rate 後利用 Rate-Distortion Function 計算相對應的 Distortion，便可建立 MB Level 的 R-D 關係。

上面那段描述了使用倒水定理的方法，接下來依理論分析的需求做一些說明和假設。理論上，要得到一個 MB 的最佳效能編碼 R-D，可以要求 Distortion 然後對一個 MB 做位元配置最佳化，這時候被編碼的 coefficient 間必須無相關性，而所期待的最佳化 R-D 關係式如下

$$\min R \text{ s.t. } D \leq D_c, \text{ where}$$

$$R = \frac{\sum_{i=1}^{256} R_i(D_i)}{256}, D = \frac{\sum_{i=1}^{256} D_i}{256},$$

期望在要求的平均D之下，用最少的R進行壓縮，並在最後計算出 pixel 所需要的平均R所以在 step 2 中運用 DCT 將 Spatial domain residual variance 轉成近似無相關性的 DCT domain residual variance。得到 Rate 後需要合適的 Rate-Distortion Function 建立 R-D 關係，這裡假設 DCT coefficient 是 Gaussian distribution,其 Rate-Distortion Function 如下：

$$R_i(D_i) = \frac{1}{2} \log_2 \left( \frac{\sigma_i^2}{D_i} \right) \quad (5.13)$$

經由簡單的運算後公式可變成



$$D_i = 2^{-2R_i} \sigma_i^2 \quad (5.14)$$

只要知道 $\sigma_i^2$ 便可在公式(5.13)(5.14)建立單一 DCT coefficient 間的 R-D 關係，所以把倒水定理中求得的 pixel level 的 DCT coefficient Rate 代入公式(5.14)中，就可以得到 pixel level 的 DCT coefficient 的 Distortion，進而求出 MB Level 的 Rate 和 Distortion。

目前已知如何取得 MB Level 的 Rate 和 Distortion，而 Single-Layer Coding 經由編碼後可以得到對應 QP 的 Rate 和 Distortion，要使理論分析模擬這樣的情況，能在給定 QP 後得到相對應的 Rate 和 Distortion。因為 QP 通常決定了 Single-Layer Coding 編碼後的畫質，所以這裡建立 QP 和 Distortion 的關係式，理論上 Distortion 和 quantization error 是一樣的，如以下式子所表示：

$$D = \frac{Qstep^2}{12} \quad (5.15)$$

公式(5.15)知道 QP 和 Distortion 的關係，先指定某 QP 計算 Distortion，然後將倒水定理 pixel level coefficient 求得的 Rate 帶入公式(5.14)得到 pixel level coefficient 的 Distortion，然後計算平均，如果平均 Distortion 可以和先前 QP 計算得到的 Distortion 一

樣，那就表示已經求得這 QP 的 Rate。



### 5.1.2 Theoretical R-D functions of Multi-Layer Hybrid Coding

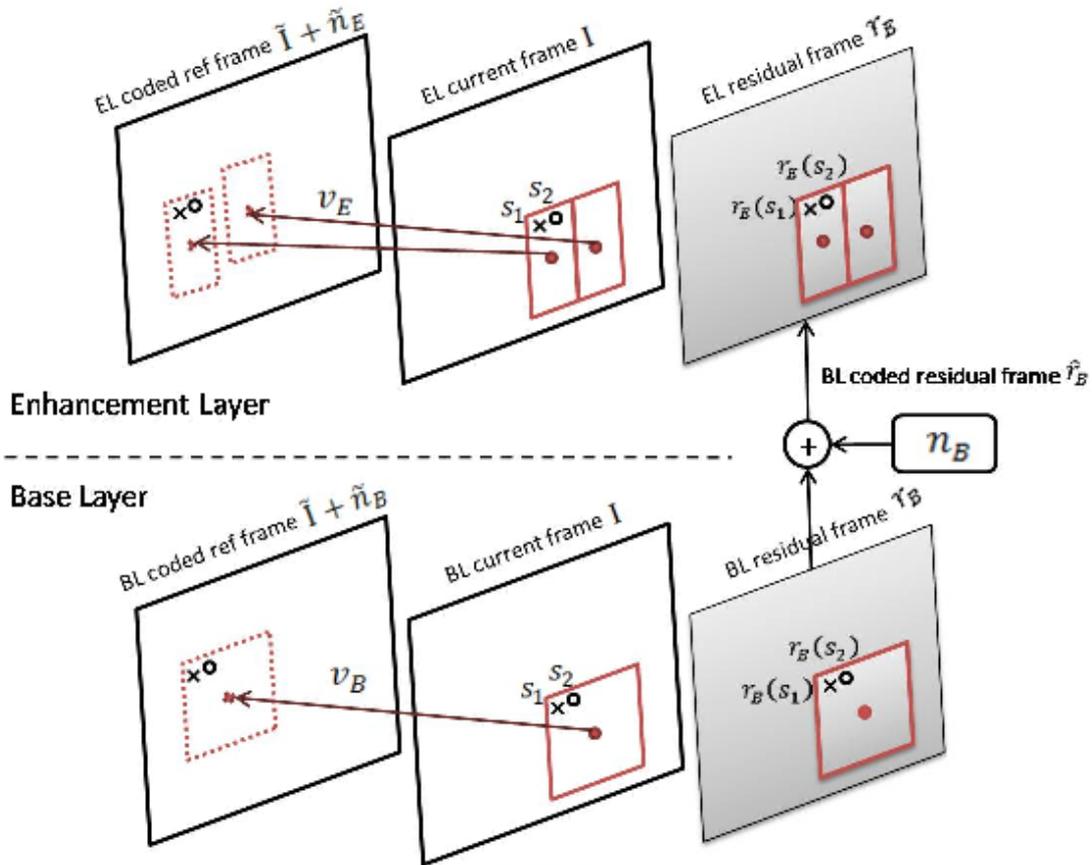


Figure 5.4 SVC 中 Inter-Layer Residual Prediction 示意圖

將理論分析延伸到 Multi-Layer, 和 Single-Layer 差別在於不能只考慮一個層的內容, 當實行 Inter-Layer Prediction 時, 就必須考慮到參考層資訊, 而如果沒有做 Inter-Layer Prediction, 那就同 H.264 不需考慮層與層間的關係。如 Figure 5.4 顯示, 在 EL coded reference frame 做 MC 取得預測訊號, 以 EL current frame 的原始訊號扣去預測訊號, 取得 EL 的 residual signal, 如果不考慮 Inter-Layer Prediction, 只需要傳輸 EL 的 residual signal, 這和 Single-Layer Prediction 雷同; 如果考慮 Inter-Layer Prediction, 將傳輸 EL 的 residual signal 扣去 BL 的 coded residual signal, 這表示在理論分析時需要多考慮 BL current frame 和 reference frame 的造成的 quantization error。

一樣對 Inter mode 做分析, 多考量 Inter-Layer Residual Prediction, 利用 *Tao et al.*[10] 推導出 residual variance 和 covariance, 我們考慮 EL mode 和 BL mode, 分成 EL 和 BL mode 相同和不同兩種情況討論, 如以下公式:

**Different mode:**

$$\begin{aligned}
& E[(r_E(s_1) - r_B(s_1) - n_B(s_1))(r_E(s_2) - r_B(s_2) - n_B(s_2))] \\
&= E[(\tilde{I}(s_1 + v_B) - \tilde{I}(s_1 + v_E) - \tilde{n}_E(s_1 + v_E) + \tilde{n}_B(s_1 + v_B) - n_B(s_1)) \\
& \quad (\tilde{I}(s_2 + v_B) - \tilde{I}(s_2 + v_E) - \tilde{n}_E(s_2 + v_E) + \tilde{n}_B(s_2 + v_B) - n_B(s_2))] \\
&= E_M[E_I[\tilde{I}(s_1 + v_B)\tilde{I}(s_2 + v_B) - \tilde{I}(s_1 + v_E)\tilde{I}(s_2 + v_E) \\
& \quad - \tilde{I}(s_1 + v_B)\tilde{I}(s_2 + v_E) + \tilde{I}(s_1 + v_E)\tilde{I}(s_2 + v_B)]] + E[noise] \tag{5.16} \\
&= E_M[R_{\tilde{I}}(s_1 + v_B, s_2 + v_B) - R_{\tilde{I}}(s_1 + v_E, s_2 + v_E) \\
& \quad - R_{\tilde{I}}(s_1 + v_B, s_2 + v_E) + R_{\tilde{I}}(s_1 + v_E, s_2 + v_E)] + E[noise] \\
&= \underbrace{\frac{8\sigma_I^2\sigma_m^2}{K}(1 - \rho_m^{\|s_E - s_B\|})}_{(A)} + E[noise]
\end{aligned}$$

在 EL 和 BL 不同 mode 時，假設 EL 和 BL 找出的 block motion 都是 true motion，公式(5.16)中第一個等式根據以下假設，舉的是  $s_1$  位置的例子， $s_2$  同理：

$$\begin{aligned}
& r_E(s_1) - r_B(s_1) \\
&= I(s_1) - \tilde{I}(s_1 + v_E) - \tilde{n}_E(s_1 + v_E) - [I(s_1) - \tilde{I}(s_1 + v_B) - \tilde{n}_B(s_1 + v_B)] \tag{5.17} \\
&= \tilde{I}(s_1 + v_B) - \tilde{I}(s_1 + v_E) - \tilde{n}_E(s_1 + v_E) + \tilde{n}_B(s_1 + v_B)
\end{aligned}$$

公式(5.16)中的第一個等式因為 EL 實行 Inter-Layer Residual Prediction 時所接受到的是重建後的 BL residual signal，所以必須多考慮 BL current frame OP 造成 quantization error。第二個等式由於假設 residual signal 和 quantization error 之間獨立，所以兩者間交叉項為零。因 EL 和 BL 的 QP 會影響 quantization error 間相關性，且點和點間距離也是影響因素，所以我們對 quantization error correlation 定義為：

$$R_{L_1 L_2}^{(n)}(s_i, s_j) = \rho_n^{\|s_i - s_j\|} \sigma_{L_1} \sigma_{L_2} \tag{5.18}$$

公式(5.18)的意義是  $L_1$  中  $s_i$  和  $L_2$  中  $s_j$  位置 quantization error 間的 correlation， $L_1$   $L_2$  表示 Layer，可能為 B 或 E 分別表示 BL 或 EL 的 current frame，如果是某 Layer 的 reference frame，則會表示為  $\tilde{B}$  或  $\tilde{E}$ ，是 BL 或 EL 的 reference frame。因為同層與不同層的 noise 相關程度也不同，所以公式(5.18)在同層時 noise 相關係數以  $\rho_n$  表示，如果不同層則以  $\rho_{n'}$  表示，因此可將公式(5.16)中的  $E[noise]$  表示成：

$$\begin{aligned}
& E[\text{noise}] \\
& = R_{\tilde{E}\tilde{E}}^{(n)}(s_1, s_2) + R_{\tilde{B}\tilde{B}}^{(n)}(s_1, s_2) + R_{\tilde{B}\tilde{B}}^{(n)}(s_1, s_2) - R_{\tilde{E}\tilde{B}}^{(n)}(s_1 + v_E, s_2 + v_B) - R_{\tilde{B}\tilde{E}}^{(n)}(s_1 + v_B, s_2 + v_E) \quad (5.19) \\
& = \rho_n^{\|s_1 - s_2\|} \sigma_{\tilde{E}}^2 + \rho_n^{\|s_1 - s_2\|} \sigma_{\tilde{B}}^2 + \rho_n^{\|s_1 - s_2\|} \sigma_{\tilde{B}}^2 - \rho_n^{\|s_1 + v_E - s_2 - v_B\|} \sigma_{\tilde{E}} \sigma_{\tilde{B}} - \rho_n^{\|s_1 + v_B - s_2 - v_E\|} \sigma_{\tilde{E}} \sigma_{\tilde{B}}
\end{aligned}$$

由於在 currnet frame 中某點的 noise 和 reference frame 中某點 noise 的相關性太低，所以直接假設為 0。公式(5.16)的第三個等式是根據公式(5.4)，第四個等式可參考公式(5.5)。

### The same mode:

$$\begin{aligned}
& E[(r_E(s_1) - r_B(s_1) - n_B(s_1))(r_E(s_2) - r_B(s_2) - n_B(s_2))] \\
& = E[(\tilde{I}(s_1 + v_B) - \tilde{I}(s_1 + v_B + v_n) - \tilde{n}_E(s_1 + v_B + v_n) + \tilde{n}_B(s_1 + v_B) - n_B(s_1)) \\
& (\tilde{I}(s_2 + v_B) - \tilde{I}(s_2 + v_B + v_n) - \tilde{n}_E(s_2 + v_B + v_n) + \tilde{n}_B(s_2 + v_B) - n_B(s_2))] \\
& = E_M[E_I[\tilde{I}(s_1 + v_B)\tilde{I}(s_2 + v_B) - \tilde{I}(s_1 + v_B + v_n)\tilde{I}(s_2 + v_B) \\
& - \tilde{I}(s_1 + v_B + v_n)\tilde{I}(s_2 + v_B) + \tilde{I}(s_1 + v_B + v_n)\tilde{I}(s_2 + v_B + v_n)]] + E[\text{noise}]' \quad (5.20) \\
& = E_M[R_{\tilde{I}\tilde{I}}(s_1 + v_B, s_2 + v_B) - R_{\tilde{I}\tilde{I}}(s_1 + v_B + v_n, s_2 + v_B) \\
& - R_{\tilde{I}\tilde{I}}(s_1 + v_B + v_n, s_2 + v_E) + R_{\tilde{I}\tilde{I}}(s_1 + v_B + v_n, s_2 + v_B + v_n)] + E[\text{noise}]' \\
& = \underbrace{\frac{4\sigma_I^2}{K} E_M[x_n^2]}_{(A)} + E[\text{noise}]'
\end{aligned}$$

公式(5.20)中假設 EL 和 BL 相同 mode 之下實行 Inter-Layer Residual Prediction 並計算其 residual covariance，第一個等式根據以下假設，舉的是  $s_1$  位置的例子， $s_2$  同理，並多考慮 BL QP 造成的 quantization error:

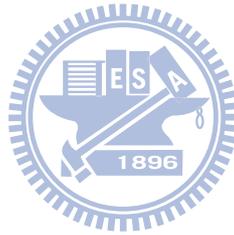
$$\begin{aligned}
& r_E(s_1) - r_B(s_1) \\
& = I(s_1) - \tilde{I}(s_1 + v_E) - \tilde{n}_E(s_1 + v_E) - [I(s_1) - \tilde{I}(s_1 + v_B) - \tilde{n}_B(s_1 + v_B)] \\
& = \tilde{I}(s_1 + v_B) - \tilde{I}(s_1 + v_E) - \tilde{n}_E(s_1 + v_E) + \tilde{n}_B(s_1 + v_B) \\
& = \tilde{I}(s_1 + v_B) - \tilde{I}(s_1 + v_B + v_n) - \tilde{n}_E(s_1 + v_B + v_n) + \tilde{n}_B(s_1 + v_B)
\end{aligned} \quad (5.21)$$

因為 EL 和 BL 的 QP 不同，所以實行 MC 後的 motion 也可能不同， $v_n$  為兩層 motion 之差，是一個隨機變數，假設  $v_n = (x_n, y_n)$ 。第二個等式是假設 residual signal 和 quantization error 獨立，第三個等式是根據公式(5.4)，第四個等式可參考公式(5.5)。而  $E[\text{noise}]'$  的計算過程在以下公式:

$$\begin{aligned}
& E[\text{noise}] \\
&= R_{\tilde{E}\tilde{E}}^{(n)}(s_1, s_2) + R_{\tilde{B}\tilde{B}}^{(n)}(s_1, s_2) + R_{BB}^{(n)}(s_1, s_2) \\
&\quad - R_{\tilde{E}\tilde{B}}^{(n)}(s_1 + v_B + v_n, s_2 + v_B) - R_{\tilde{B}\tilde{E}}^{(n)}(s_1 + v_B, s_2 + v_B + v_n) \\
&= \rho_n^{\|s_1 - s_2\|_1} \sigma_{\tilde{E}}^2 + \rho_n^{\|s_1 - s_2\|_1} \sigma_{\tilde{B}}^2 + \rho_n^{\|s_1 - s_2\|_1} \sigma_B^2 - E[\rho_n^{\|s_1 - s_2 + v_n\|_1}] \sigma_{\tilde{E}} \sigma_{\tilde{B}} - E[\rho_n^{\|s_1 - s_2 - v_n\|_1}] \sigma_{\tilde{E}} \sigma_{\tilde{B}}
\end{aligned}
\tag{5.22}$$

公式(5.22)中， $v_n = (x_n, y_n)$ ， $x_n, y_n$ 都是隨機變數，在運算時會代入一個合理的值。

本小節敘述理論分析中實行 Inter Prediction 時 Multi-Layer Scalable Coding 和 Single-Layer Coding 的差別，在於 Multi-Layer Scalable Coding 必須多考慮 Inter-Layer Prediction。和模擬 Single-Layer Coding 編碼過程的差別只有 Figure 5.2 的 step 1，然後循序的進行 step2 和 step3。



## 5.2. Theoretical Data and Model Analysis

因為假設在相同 sequence 性質上運作，所以理論分析時給於每個 mode 相同的參數而實際資料的實驗設定請看 Table 5.2，理論分析的實驗設定請看 Table 5.3。接下來的實驗會針對 Mode Decision 的因素進行分析，觀察理論分析的結果和實際情況是否符合，最後利用 model 進行更詳細的解釋。

Table 5.2 : 實際資料測試條件

	H.264/AVC
解析度:每秒 frame 數	QCIF :15、7.5、3.75 CIF :30、15、7.5
mode 選擇	16x16、16x8、8x16、8x8
Intra 週期	-1
搜尋範圍	32
搜尋型態	Full Search

Table 5.3 : 理論分析測試條件

	Single-Layer Coding	Multi-Layer Coding
intensity variance $\sigma_I^2$	100	1000
motion variance $\sigma_m^2$	25	100
Positive number K	150	150
correlation between two motion $\rho_m$	0.9	0.9
correlation between two noise in EL $\rho_n$	0.3	0.3
correlation between two noise between EL and BL $\rho'_n$	0.6	0.6
Motion Rate	rational value	No Motion Rate

## 5.2.1 Experiment 1

此實驗觀察理論分析中 QP 對 Mode Decision 的影響，並舉 Single-Layer 為例子和實際資料比較。理論分析上參數上的設定為 Table 5.3 的設定。而實際 R-D 的取得方法，是在做 Mode Decision 時實行徹底式搜尋，只針對 sequence 中某一張 frame 記錄不同 mode 時每個 pixel 的平均 R-D。

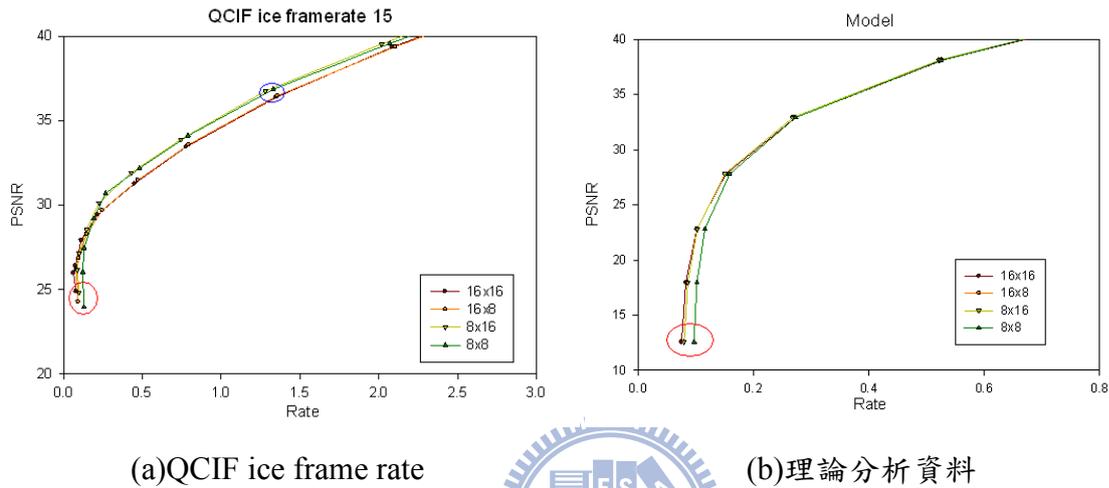


Figure 5.5 Rate-PSNR curve

Figure 5.5(a)中每個點表示某 QP 的 Rate-PSNR，為了方便比較，同一條 Rate-PSNR curve 的 QP 間隔為 5。觀察實際資料紅色圈內，當 QP 很大時，16x16 mode 的 PSNR 最高且 Rate 需要最少，所以這時候選擇 16x16 能有最佳的編碼效能。相同的道理，觀察藍色圈內，PSNR 接近 37 時會選擇 8x16，表示當 QP 由大變小時，會漸漸的選擇其他分割較細的 mode。理論分析中也可以觀察出類似趨勢，因為公式(5.15)中 Distortion 隨 QP 而改變，所以 Figure 5.5(b)中相同 QP 下的點，所需 Rate 最少的 mode 就是編碼效能最佳的 mode，觀察紅色圈內，明顯 QP 很大時會選擇 16x16，QP 由大變小時，漸漸的改選了其他 mode。所以本論文所提出之 model 可以在 QP 改變時模擬出類似實際情況的 Mode Decision。

以 model 分析 QP 對 Mode Decision 的影響，觀察公式(5.1)的(A)部分，因為 block center 位置會隨 mode 不同而有差別。在不同 mode 時設定參數  $\rho_m, \sigma_m^2, \sigma_l^2, K$  一樣，則 16x16 在(A)部分的值會最大，8x8 的值最小，(B)部分則受到 reference frame QP 影響其大小，(A)和(B)部分的相互比例明顯影響著 Mode Decision。當 QP 很小時，(A)部分值較大，(B)較小，不同 mode 時的 covariance 比例上差異較大；QP 較大時，(B)部分值很大，顯得不同 mode 的 covariance 比例差異變小。在進行 DCT 轉換後，(B)的影響會集中在 DC 值，如以下的例子所示：

$$U \times \begin{pmatrix} s_{1,1} + a & s_{1,2} + a & s_{1,3} + a & \cdots & s_{1,16} + a \\ s_{2,1} + a & s_{2,2} + a & s_{2,3} + a & & \\ s_{3,1} + a & s_{3,2} + a & s_{3,3} + a & & \vdots \\ \vdots & & & \ddots & \\ s_{16,1} + a & & \cdots & & s_{16,16} + a \end{pmatrix} \times U^T = \begin{pmatrix} d_{1,1} + a & d_{1,2} & d_{1,3} & \cdots & d_{1,16} \\ d_{2,1} & d_{2,2} & d_{2,3} & & \\ d_{3,1} & d_{3,2} & d_{3,3} & & \vdots \\ \vdots & & & \ddots & \\ d_{16,1} & & \cdots & & d_{16,16} \end{pmatrix} \quad (5.23)$$

在(5.23)中  $s_{a,b}$ 、 $d_{a,b}$  表示的就是公式(5.1)的(A)部分，不考慮 noise 的 spatial domain residual covariance 和 DCT domain residual covariance，為了簡化表達過程，將(B)部分都假設為  $a$ 。進行 DCT 後能量集中在 DC 值，表示當 reference frame QP 很大時，DCT domain 的 DC 值會更大，如果想達到相同的 Distortion，DC 值更大會使 DC 值需要更多的 Bit-Allocation，進而造成 MB 需要更多的 Rate。



## 5.2.2 Experiment 2

在理論分析中實行 Inter-Layer Residual Prediction，觀察 BL mode 對 EL 做 mode Decision 的影響，是否和實際情況類似。當 BL mode 不同時，通常有不一樣的 BL residual signal，由於 EL 實行 Inter-Layer Residual Prediction 時必須扣去 BL residual signal，表示會因 BL mode 不同而扣去不一樣的 BL residual signal，進而影響 Mode Decision。因 SVC 採用 BUEC，先決定 BL mode，再決定 EL mode，所以在 BL mode 固定之下觀察實行 Inter-Layer Residual Prediction 的 EL mode R-D 效能，而本實驗中的 Rate 表示 Residual Rate，不包含 Motion Rate。

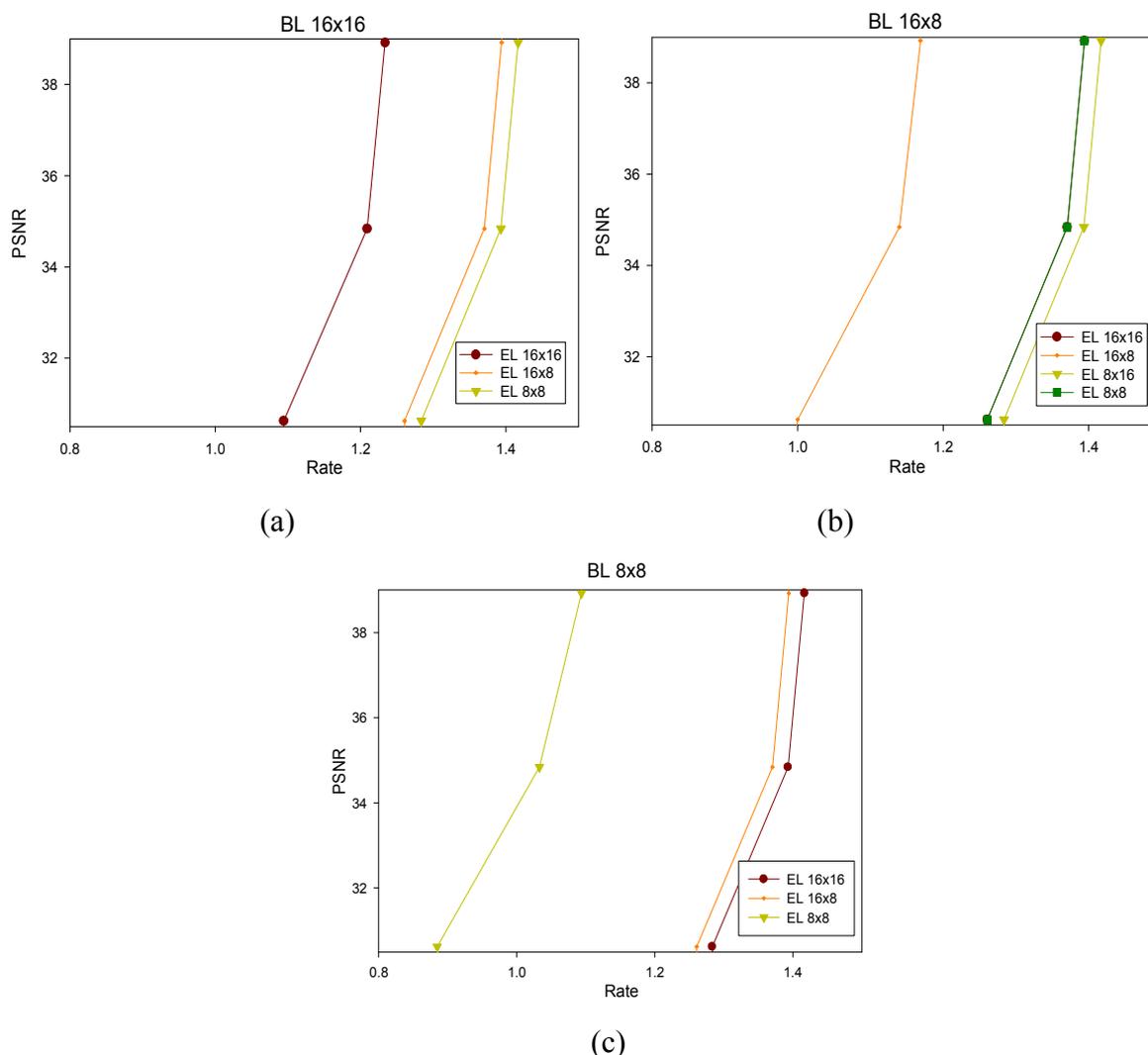


Figure 5.6(a)(b)(c)在 BL mode 固定之下不同 EL mode 的 R-D 效能

Figure 5.6 中，BL mode 已經固定，顯示 EL 在不同的 Inter mode 下實行 Inter-Layer Residual Prediction 的 Rate-PSNR 曲線，figure 中顯示了當 BL mode 不一樣時，Mode Decision 的順序明顯不同，和實際的數據有類似的趨勢。

對 model 進行分析，從 model 中解釋 Figure 5.7 的結果。觀察公式(5.16)的(A)部分，可以看出 EL 和 BL mode 的 block center 位置影響(A)的值，當 mode 的差異性越大，像是 16x8 和 8x16 值就越大，若是 16x16 和 16x8，值就相對較小了些，而相同 mode 的情況，必須參看公式(5.20)， $x_n$ 會代入適當的值，其值常理來說會小於 EL 和 BL mode 不同時 EL mode 和 BL mode 的 block center 的距離。



### 5.2.3 Experiment 3

QP Difference (QPD)表示 EL 和 BL 間 QP 的差值，此實驗要觀察本論文所提出的理論分析中 QPD 對 Inter-Layer Residual Prediction 的影響。因為 QPD 和 residual signal 的相關性有一定影響，當 QPD 小的時候，EL 和 BL 的 residual signal 相關性強，這時實行 Inter-Layer Residual Prediction 可能會得到更好的 R-D 效能。當 QPD 很大時，EL 和 BL 的 residual signal 相關性弱，這時候實行 Inter-Layer Residual Prediction 的 R-D 效能可能不比 Single-Layer Coding。所以本實驗實行 Inter-Layer Residual Prediction，分別取 QPD 等於 2、4、6 來比較，其他參數設定都跟 Table 5.3 內容一樣，而本實驗中的 Rate 表示 Residual Rate，不包含 Motion Rate。

觀察 Figure 5.7，明顯當 QPD 的值越小，要求 EL 達到相同的 PSNR，則 Rate 所需越少，反之亦然。這和實際情況相符合。

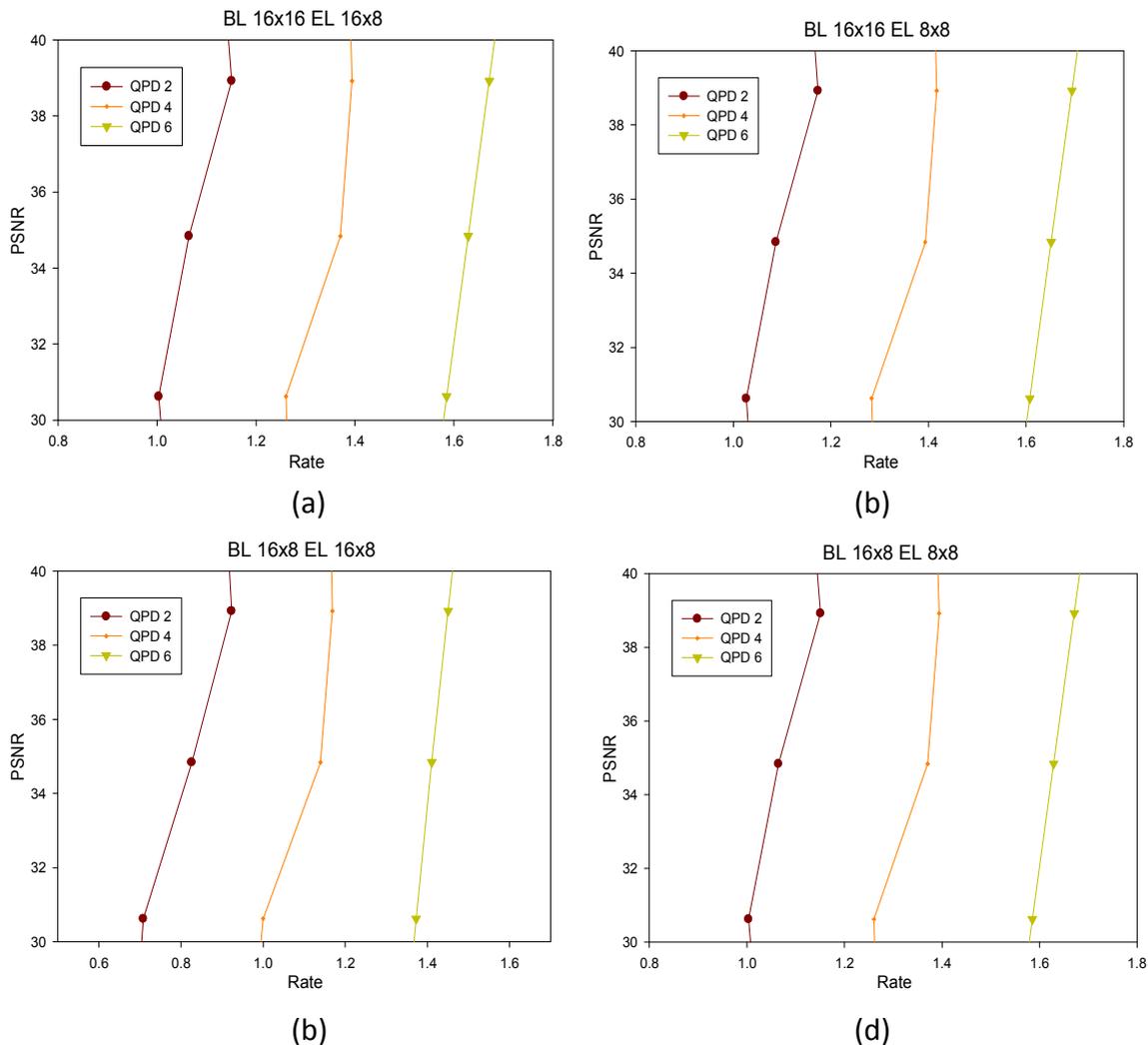


Figure 5.7 (a)(b)(c)(d)BL 和 EL mode 在各種組合之下實行 Inter-Layer Residual Prediction 的 Rate-PSNR curve

## 5.2.4 Experiment 4

Experiment 3 觀察 QPD 對 EL Rate 的影響，而此實驗要觀察所提出的理論分析在實行 Inter-Layer Residual Prediction 之下 QPD 對 Total Rate 的影響。實驗設定在 EL QP 相同時，QPD 為 2、4、6，實行 Inter-Layer Residual Prediction，比較 Total Rate 的變化。實驗設定請參考 Table 5.3，而本實驗中的 Rate 表示 Residual Rate，不包含 Motion Rate。

Figure 5.8 中顯示當 QPD 差距較小時，Total Rate 會更少，和實際數據的情況相同。

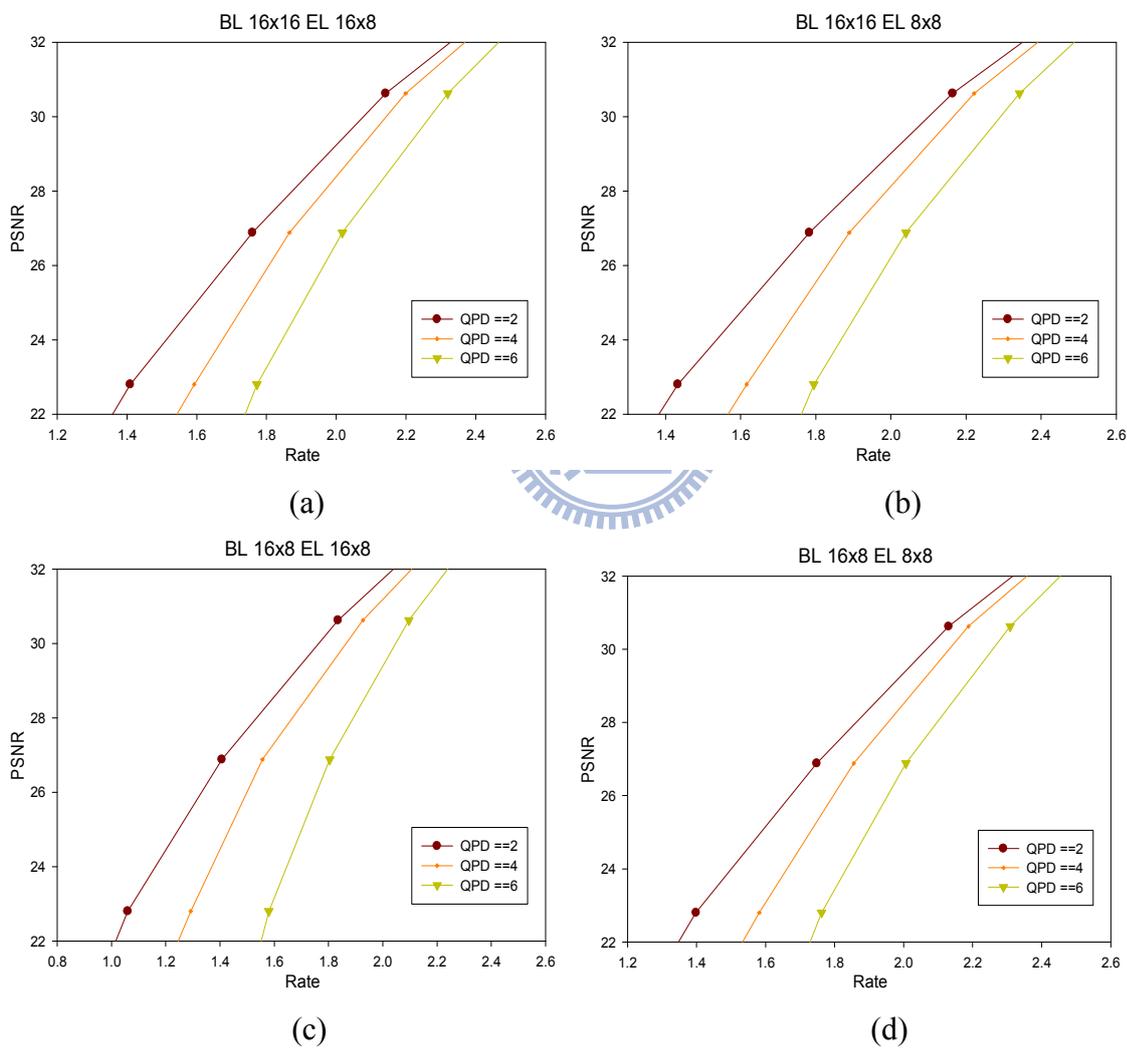


Figure 5.8(a)(b)(c)(d)不同 QPD 之下的 Rate-PSNR curve

## 5.2.5 Experiment 5

本實驗觀察在 QPD 不同時，實行 Inter-Layer Residual Prediction 時 EL 的 BL usage，BL usage 的定義是 EL 對 BL 的資訊利用度，就是 EL 在使用 BL 參數之下的能得到多少好處。所以實驗設定實行 Inter-Layer Residual Prediction，加入 EL Single-Layer Coding 和 SVC simulcast 的情況，EL Single-Layer Coding 是在考慮 EL QP 之下計算 Single-Layer Coding 的 Rate-PSNR，simulcast 是指 SVC 中 EL 不實行 Inter-Layer Residual Prediction 的情況計算 Rate-PSNR，需加上 BL Rate，觀察在不同 QPD 時的 BL usage，QPD = 2、4、6，實驗設定請參考 Table 5.3，而本實驗中的 Rate 表示 Residual Rate，不包含 Motion Rate。一般來說，EL 和 BL 的 QP 越接近，BL usage 會越高，期望在理論分析中看到這樣的現象。

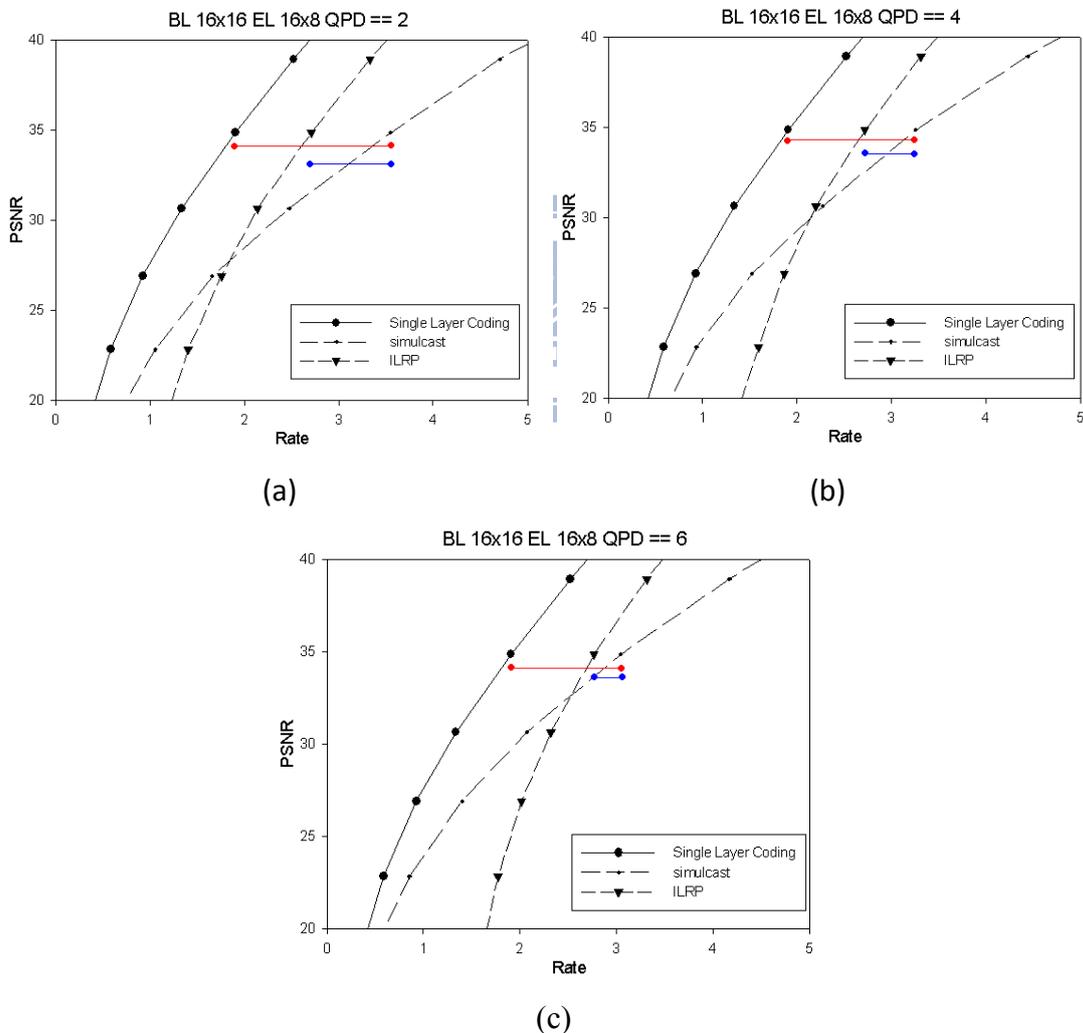


Figure 5.9 (a)(b)(c)QPD 不同之下的 Rate-PSNR curve

在 Figure 5.9(a)(b)(c)中針對三種條件的某一 QP 分析，一般來說，如果 Inter-Layer Residual Prediction 的 Rate-PSNR curve 越接近 Single-Layer Coding，離 simulcast 越遠，表示其 BL usage 更好。如 Figure 5.9(a)(b)(c)中，把 Single-Layer Coding 和 simulcast 的點距離用紅色連線顯示，Inter-Layer Residual Prediction 和 simulcast 的點距離用藍色連線顯示，藍色線段佔紅色線段的比例就是 BL usage 程度比例越高表示 BL usage 越高，反之亦然，經由計算 Figure 5.9 所舉的例子 QPD == 2 時的 BL usage 是 51%，QPD == 4 時的 BL usage 是 40%，QPD == 6 時的 BL usage 是 24%，所以本論文的理論分析可以顯示 QPD 對 BL usage 的影響，QPD 越小時 BL usage 通常是越高的，這和實際的情況相符合。

以分析 model 的方式來解釋 QPD 的影響，因為固定 Mode Pair 的情況下公式(5.16)和(5.20)的(A)部分都一樣，主要是 QP 對 noise 的影響，所以對公式(5.19)和公式(5.22)進行分析：

$$\begin{aligned}
& E[noise] \\
&= R_{\bar{E}\bar{E}}^{(n)}(s_1, s_2) + R_{\bar{B}\bar{B}}^{(n)}(s_1, s_2) + R_{BB}^{(n)}(s_1, s_2) - R_{\bar{E}\bar{B}}^{(n)}(s_1 + v_E, s_2 + v_B) - R_{\bar{B}\bar{E}}^{(n)}(s_1 + v_B, s_2 + v_E) \\
&= \rho_n^{\|s_1-s_2\|_1} \sigma_{\bar{E}}^2 + \rho_n^{\|s_1-s_2\|_1} \sigma_{\bar{B}}^2 + \rho_n^{\|s_1-s_2\|_1} \sigma_B^2 - \rho_{n'}^{\|s_1+v_E-s_2-v_B\|_1} \sigma_{\bar{E}} \sigma_{\bar{B}} - \rho_{n'}^{\|s_1+v_B-s_2-v_E\|_1} \sigma_{\bar{E}} \sigma_{\bar{B}} \\
&= \underbrace{\rho_n^{\|s_i-s_j\|_1} (\sigma_{\bar{E}} - \sigma_{\bar{B}})^2}_{(A)} + \rho_n^{\|s_i-s_j\|_1} \sigma_B^2 + (2\rho_n^{\|s_i-s_j\|_1} - E[\rho_{n'}^{\|s_1+v_E-s_2-v_B\|_1}] - E[\rho_{n'}^{\|s_1+v_B-s_2-v_E\|_1}]) \sigma_{\bar{E}} \sigma_{\bar{B}}
\end{aligned} \tag{5.24}$$



$$\begin{aligned}
& E[noise]' \\
&= R_{\bar{E}\bar{E}}^{(n)}(s_1, s_2) + R_{\bar{B}\bar{B}}^{(n)}(s_1, s_2) + R_{BB}^{(n)}(s_1, s_2) - R_{\bar{E}\bar{B}}^{(n)}(s_1 + v_B + v_n, s_2 + v_B) - R_{\bar{B}\bar{E}}^{(n)}(s_1 + v_B, s_2 + v_E + v_n) \\
&= \rho_n^{\|s_1-s_2\|_1} \sigma_{\bar{E}}^2 + \rho_n^{\|s_1-s_2\|_1} \sigma_{\bar{B}}^2 + \rho_n^{\|s_1-s_2\|_1} \sigma_B^2 - E[\rho_{n'}^{\|s_1-s_2+v_n\|_1}] \sigma_{\bar{E}} \sigma_{\bar{B}} - E[\rho_{n'}^{\|s_1-s_2-v_n\|_1}] \sigma_{\bar{E}} \sigma_{\bar{B}} \\
&= \underbrace{\rho_n^{\|s_i-s_j\|_1} (\sigma_{\bar{E}} - \sigma_{\bar{B}})^2}_{(A)} + \rho_n^{\|s_i-s_j\|_1} \sigma_B^2 + (2\rho_n^{\|s_i-s_j\|_1} - E[\rho_{n'}^{\|s_1-s_2+v_n\|_1}] - E[\rho_{n'}^{\|s_1-s_2-v_n\|_1}]) \sigma_{\bar{E}} \sigma_{\bar{B}}
\end{aligned} \tag{5.25}$$

式子中(A)部分可明顯可看出固定 EL QP 時，QPD 差越小時，(A)部分的值越小，反之亦然，因為進行 DCT 後，能量會集中在左上角，可以看公式(5.23)的例子，表示當 a 的值越小時，所需要的 Rate 數變少，反之亦然。

## Chapter 6. Conclusions

因為SVC Bottom-up Encoder Control (BUEC)在EL的R-D效能相較於單一層編碼有不可預期的失真。因此有Schwarz *et al.* [3]提出了一個Multi-loop Encoder Control (MLEC)的觀念和快速決策演算法。而Lin修正了這篇paper的最佳化限制問題和提出了新的Mode Decision/ME公式，而我沿用此公式。在這篇論文裡，提供了一個新的兩段式快速演算法，加入了Temporal Mode Score (TMS)的觀念，model了一個初始機率公式，和新增了可以隨sequence性質修正搭配的mode的方法。模擬的實驗結果證明了此方法確實提供了非常好的編碼效能，整體的結果也比以往MLEC的快速決策演算法來的穩定，而三個演算法的time saving差不多，約71%~82%。

理論分析部分，能在調整一些參數後，模擬出和實際數據類似的趨勢，期許未來能夠在實際sequence中取得適合的資訊，直接代入model中，便可以根據model的趨勢來判斷該如何進行Mode Decision。此研究目前為止依然是停留在較早的階段，未來的研究可以繼續延伸到以下幾點

- (1) 將我們提出的演算法實做到Hierarchical B的架構上
- (2) 最佳化問題可以被延升為解決R-D最佳化的相依性問題[8][9]，針對MLEC發一個新的Lagrange-multiplier的計算法
- (3) 試圖連結model和實際資料的一些參數,使model能模擬出和實際數據相同的趨勢
- (4) 理論分析延伸到和MLEC



# BIBLIOGRAPHY

- [1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103-1120, 2007.
- [2] J. Reichel, H. Schwarz, and M. Wien, "Joint Scalable Video Model 10 (JSVM 10)," *Joint Video Team*, JVT07/W202, 2007.
- [3] H. Schwarz and T. Wiegand, "R-D Optimized Multi-layer Encoder Control for SVC," *IEEE Int'l Conf. on Image Processing*, 2007.
- [4] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688-703, 2003.
- [5] D.S. Taubman and M. W. Marcellin, "JPEG2000 Image Compression, Fundamentals, Standards and Practice," 2002.
- [6] T. France, H. Schwarz, and M. Wien "JSVM Software (version 9.12.2)," *Joint Video Team*, JVT08/Z203, 2008.
- [7] X. Li, P. Amon, A. Hutter, and K. Andre, "One-pass multi-layer rate-distortion optimization for quality scalable video coding," *IEEE International Conference on ICASSP 2009*.
- [8] K. Ramchandran, A. Ortega and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. Image Process*, vol. 3, no.5, pp. 533-545, 1994.
- [9] S. Liu and C.-C. J. Kuo "Joint temporal-spatial bit allocation for video coding with dependency," *IEEE Trans. Circuits, Syst., Video Technol.*, vol. 15, no.5, pp. 15-26, 2005.
- [10] B. Tao and M. Orchard, "A parametric solution for optimal overlapped block motion compensation," *IEEE Transaction on Image Processing*, vol. 10, no. 3, pp. 341-350, 2001.
- [11] Y. Wang, J. Ostermann, and Y-Q.Zhang, "Video processing and communications," *Prentice Hall*, 2001.
- [12] C. S. Kannangara, I. E. G. Richardson, M. Bystrom, J. R. Solera ,Y. Zhao ,A. MacLennan and R. Cooney "Low-complexity skip prediction for H.264 through lagrangian cost estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2 pp. 202-208, 2006.
- [13] C. Kim and C.-C. J. Kuo, "Feature-based intra/inter coding mode selection for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 4 pp. 441-453, 2007.

- [14] B. G. Kim "Novel inter-mode decision algorithm based on macroblock (MB) tracking for the P-slice in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 2, pp. 273-279, 2008.
- [15] S. H. Ri, Y. Vatis and J. Ostermann "Fast inter-mode decision in an H.264/AVC encoder using mode and lagrangian cost correlation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 302-306, 2009.
- [16] C.-H. Yeh, K.-J. Fan, M.-J. Chen, and G.-L. Li, "Fast mode decision algorithm for scalable video coding using Bayesian theorem detection and Markov process," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 4, pp. 563-574, 2010.
- [17] S.-T. Kim, K. Konda, C.-S. Park, C.-S. Cho, and S.-J. Ko, "Fast mode decision algorithm for inter-layer coding in scalable video coding," *IEEE Trans. Consumer Electronics*, vol. 55, no. 3, pp. 1572-1580, 2009.

