

第三章 空間資訊選擇性 (Frequency Information selection)

本章說明本篇論文提出的方法，從基本概念，推展至實作上的重點。

3.1 概念



影像中不同頻率的成分，其重要性也不同。無論是由視覺實驗得來的對比敏感度函式/對比敏感度曲線(Contrast Sensitivity Function)，或是由實際編碼規格中所用的量化矩陣，我們可以得知低頻的分量比高頻分量重要。

而在 MPEG-4 的基本層的離散餘弦係數有經過不平均量化(un-uniform quantization)，故已經加強了基本層的視覺效果；但是加強層的係數只有單純的作位元平面編碼，意即把每個頻率的分量依照位元層順序作編碼並送出。而空間資訊選擇性的功能即是在加強層也提供不同頻段的獨立位元層次編碼，故在傳送時能自由地選擇各頻段的資訊量，進而在有限資料傳輸量下達到最佳視覺效果。

由於加強層採用的位元平面編碼與後續的可變長度編碼達到了良好的壓縮效果，於是我們提出的空間資訊選擇的方法架構在原有的位元平面編碼之上，我們把同一個區塊的離散餘弦係數分割成數個頻段，而希望不同頻段的位元層能獨立傳輸，換句話說，在低頻成分比較重要的情況下，低頻段係數可以多傳幾個位元層，高頻段係數則少傳一些。

與基本層的不平均量化比較，如此的方式將產生兩個限制：

- 每個頻段內的係數仍需以相同的位元層順序傳送
由於編碼效率的考量，我們把 8×8 離散餘弦係數分成三個頻段，故只能提供三個自由度的空間資訊選擇，這個限制使得我們無法像基本層一般對 64 的係數作

- 不同的量化，調整視覺效果的差異
- 不同頻段之間的最小差異將會是一個位元層
等效於其量化位階只能相差 2 的整數次方倍，這是位元平面編碼的先天限制，也是數位資訊的限制。致使三個頻段各自的選擇能力變得較粗略。

以上兩個限制與編碼效率無關，然而編碼效率對於視覺品質的影響更為直接。由於 MPEG-4 規格經過長時間的改進，其編碼效率佳；在我們調整架構之後，編碼架構降低是可預期的，而我們也採取一些措施來提升編碼效率。但是在微調之前，我們先展示在這兩個先天限制之下，空間資訊選擇性仍具有其利益(benefit)，然後我們再進行調整與補強。

3.1.1 效果預測

我們在 baboon 影像上模擬 MPEG-4 對於 I-畫面的編解碼。

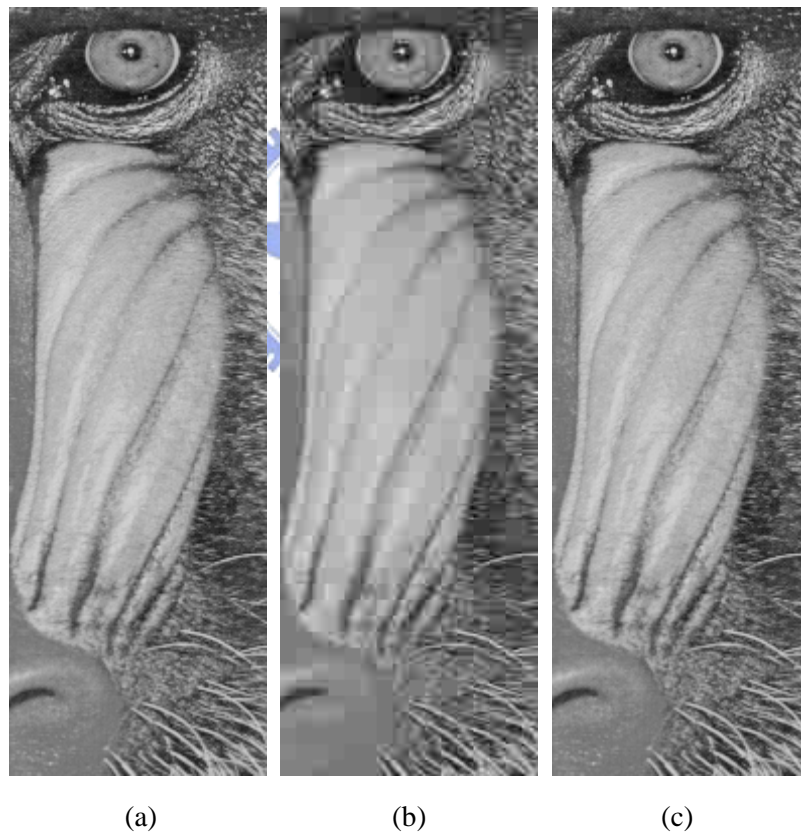


圖 3.1 MPEG4 I 畫面模擬 (a) 原始影像 (b) 還原基本層 (c) 還原基本層與加強層

在選定固定 $QP=28$ 的情況下，加強層共有 6 層位元層。我們將每個區塊中的 64 個係數分成三個頻段：L0、L1、L2

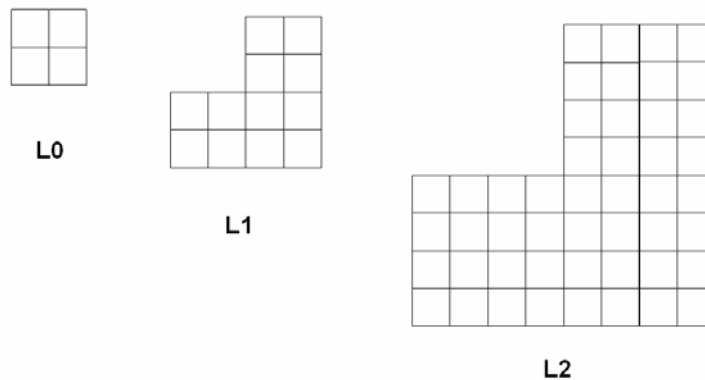


圖 3.2 頻段分割

表 3.1 位元層中'1'個數的統計

	第一層	第二層	第三層	第四層	第五層	第六層	一層係數個數
L0	150	1029	1243	1457	1529	1557	4608
L1	841	2406	3561	4150	4468	4630	13824
L2	1405	4968	9742	14625	17360	18334	55296

由於不同頻段包含的係數個數不等，所以每個位元層內'1'的數目分部不均，總合來說：

1. 高頻段包含的'1'較低頻段為多
2. 越低有效位元平面，高頻段與低頻段的'1'數目比例更懸殊。

每個頻段作獨立的位元平面取捨，而最高位元層為6，故每一頻段有0~6七種選擇；三個頻段具有343種組合。組合繁多，故我們針對一些假定設計比較方式，並請N位測試者比較，以下是我們測試與表現的方式：(N=6)

以兩張影像的比較為基礎，請測試者同時觀察原始影像與兩張影像之一，任意切換兩張影像比較，以三個等級(優 等 劣)評等相互關係，分別以[1 0 -1]表示[優 等 劣]紀錄兩張影像的相互關係。累計所有測試者的評分，並以不同符號表示：

表 3.2 評比符號

積分 x	$x > (2N/3)$	$(2N/3) > x > (N/3)$	$(N/3) > x > (-N/3)$	$(-N/3) > x > (-2N/3)$	$(-2N/3) > x$
符號	>>	>	~	<	<<

觀察結果與推論歸納如下：

1. 假設：

整體的位元層中，較高位元的幾層重要性較大。

觀察：(請見圖 3. (a-1 ~ a-4))

表 3.3 測試紀錄一

測試紀錄
[1 1 1] >> [0 0 0]
[2 2 2] >> [1 1 1]
[3 3 3] >> [2 2 2]
[4 4 4] >> [3 3 3]
[5 5 5] ~ [4 4 4]

歸納：

a 方式中(與原本的位元層編碼相同)加強層取到第三層以後，視覺效果改善漸緩。其原因在於區塊效應的消失。

推論：

前三層的視覺影響較大，故我們稍後的比較，應該著重於每個頻段在前面三層位元層的取捨；另外，區塊效應的程度，對於中低品質的影像品質很重要。(區塊效應的消除為中低品質影像視覺上的主要改善)

2. 假設：

每個頻段的重要性不等，其較重要的位元層數目不等。

觀察：(請見圖 3. (b-1 ~ b-8))

表 3.4 測試紀錄二

測試紀錄		
[1 0 0] ~ [0 0 0]	[0 1 0] >> [0 0 0]	[0 0 1] ~ [0 0 0]
[2 1 1] >> [1 1 1]	[1 2 1] >> [1 1 1]	[1 1 2] > [1 1 1]
[3 2 2] >> [2 2 2]	[2 3 2] >> [2 2 2]	[2 2 3] > [2 2 2]
[4 3 3] ~ [3 3 3]	[3 4 3] ~ [3 3 3]	[3 3 4] ~ [3 3 3]

歸納：

以下兩兩比較，在某方具有客觀資料優勢下，品質差別非常微小：

- [4 3 3] vs. [3 3 3] (但是[3 2 2] vs. [2 2 2] 差異相當明顯)
- [3 4 3] vs. [3 3 3] (但是[2 3 2] vs. [2 2 2] 差異相當明顯)
- [3 3 4] vs. [3 3 3] (但是[2 2 3] vs. [2 2 2] 差異明顯)

推論：

「L0」高於第四層的位元層對於影像品質改善難以察覺。

「L1」高於第四層的位元層對於影像品質改善難以察覺。

「L2」高於第四層的位元層對於影像品質改善難以察覺。

以上三點程度上稍有差異，但是分界點都落在第三、四個位元層之間，故三個頻段的第四位元層的絕對重要性難以決定。這就是整數位元層帶來的效應 - 可調

層次式精細度不佳。

3. 假設：

相同位元層中的不同頻段，重要性不等。

觀察：(請見圖 3. (c-1 ~ c-9))

表 3.5 測試紀錄三之一

測試紀錄		
[1 0 0] < [0 1 0]	[0 1 0] >> [0 0 1]	[1 0 0] > [0 0 1]
[2 1 1] >> [1 1 2]	[2 1 1] >> [1 2 1]	[1 2 1] < [1 1 2]
[3 2 2] < [2 2 3]	[3 2 2] << [2 3 2]	[2 3 2] ~ [2 2 3]
[4 3 3] < [3 3 4]	[4 3 3] ~ [3 4 3]	[3 4 3] ~ [3 3 4]
[5 4 4] ~ [4 4 5]	[5 4 4] ~ [4 5 4]	[4 5 4] ~ [4 4 5]

表 3.6 測試紀錄三之二

測試紀錄		
[1 1 0] ~ [0 1 1]	[1 1 0] > [1 0 1]	[1 0 1] << [0 1 1]
[2 2 1] >> [1 2 2]	[2 2 1] > [2 1 2]	[2 1 2] >> [1 2 2]
[3 3 2] << [2 3 3]	[3 3 2] < [3 2 3]	[3 2 3] << [2 3 3]
[4 4 3] < [3 4 4]	[4 4 3] < [4 3 4]	[4 3 4] ~ [3 4 4]
[5 5 4] ~ [4 5 5]	[5 5 4] ~ [5 4 5]	[5 4 5] ~ [4 5 5]

歸納：

以下的比較，在資料優勢無客觀優勝者情況下，有可認知差別：

- [0 1 0] > [1 0 0] > [0 0 1] and [1 1 0] ~ [0 1 1] > [1 0 1]
- [2 1 1] >> [1 1 2] > [1 2 1] and [2 2 1] > [2 1 2] >> [1 2 2]
- [2 3 2] ~ [2 2 3] > [3 2 2] and [2 3 3] > [3 2 3] > [3 3 2]
- [3 4 4] ~ [4 3 4] > [4 4 3]

推論：

在位元層數目小，影像品質改善明顯時，在同一個位元平面內，低頻段的資訊對於影像品質較重要，能減低區塊效應的不適感；在位元層數目大，區塊效應消失，影像品質達到中上的時候，在同一個位元平面內，高頻段的資訊對於影像品質較重要，能改善影像的模糊感。

4. 假設：

低頻段的重要性很重要，值得比其他頻段多傳兩層以上的位元層。

觀察：(請見圖 3. (d-1 ~ d-8))

表 3.7 測試紀錄四

測試紀錄

[2 0 0] > [1 1 0]
[3 1 1] > [2 2 1]
[4 2 2] << [3 3 2]
[5 3 3] < [4 4 3]
[6 4 4] ~ [5 5 4]

歸納：

平均位元層數小於 2 時，低頻段多傳兩層的效果稍微優於平均分配；平均位元層數大於 3 以上時，平均分配的影像品質明顯較佳。

推論：

在平均位元層數低時，低頻段資訊具有大過一層位元層的影響力；待平均位元層升高之後，即影像品質較佳、區塊效應淡化時，過高比例的低頻段成分反而明顯的拉低了影像品質。

5. 假設：

在「沒有人眼可感知差異」的條件下減少資料量，高頻段也許可以減少多一些。

觀察：(請見圖 3. (e-1~e-3))

表 3.8 測試紀錄五

測試紀錄		
[3 3 3] <<(7/7) [6 6 6]	[4 4 4] <<(7/7) [6 6 6]	[5 5 5] ~(0) [6 6 6]
[3 3 2] <<(7/7) [6 6 6]	[4 4 3] <<(7/7) [6 6 6]	[5 5 4] ~(1/7) [6 6 6]
[3 2 2] <<(7/7) [6 6 6]	[4 3 3] <<(7/7) [6 6 6]	[5 4 4] <(4/7) [6 6 6]
[2 3 3] <<(7/7) [6 6 6]	[3 4 4] <<(7/7) [6 6 6]	[4 5 5] ~(1/7) [6 6 6]

註：此項實驗的環境設定與之前有一處不同→影像以原尺寸顯示，

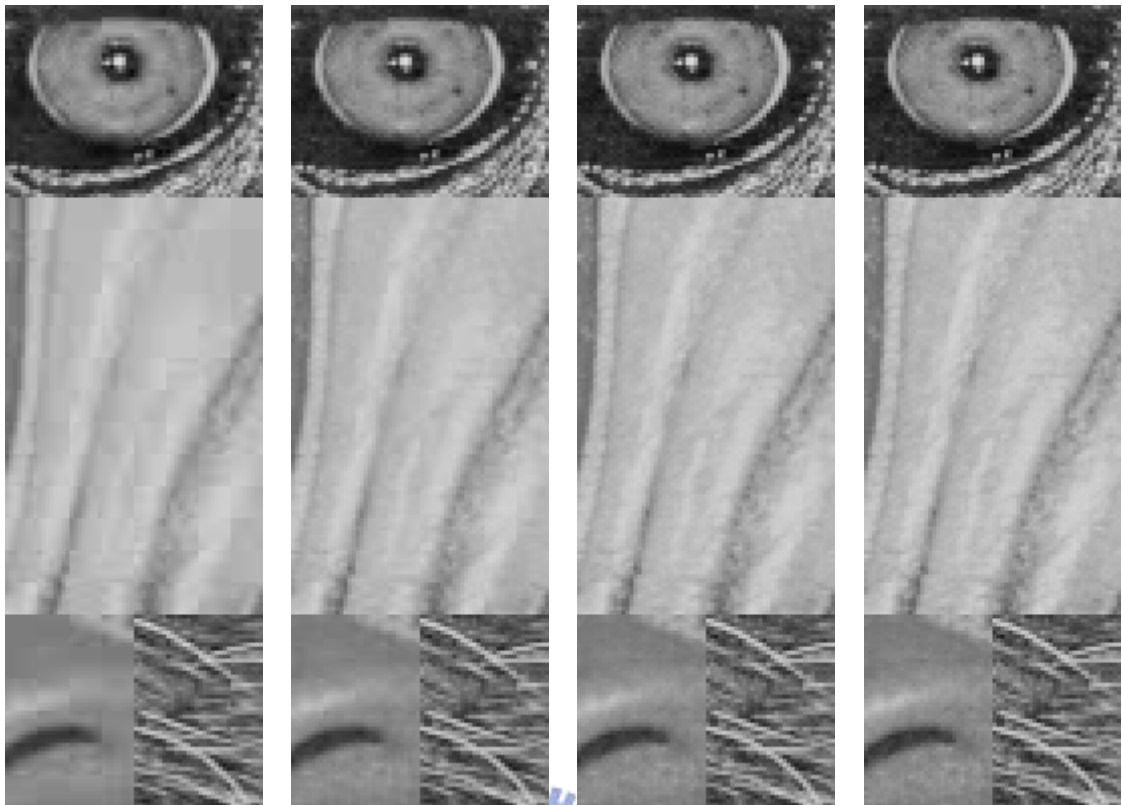
歸納：

最後一層位元層的影響沒有人能辨別；倒數第二層位元層可以部分移除，視覺上仍難以辨別差異。

推論：

拿了五層的加強層以後，影像的品質已臻完美；此時，大部分的人對於少拿第五層的低頻段(L0)或高頻段(L2)無法察覺，甚至同時少拿中高頻段(L1&L2)也有人無法察覺；然而所有人都能辨別少拿整個第五層的差別。

於是，我們確定頻率資訊選擇性的確提供我們較大的「不必要資訊」的預算(Budget)。同時，我們認定「不可認知的資料減少」不具有加法不變性，意即“單獨少 A 或單獨少 B 皆無法察覺，不保證同時少了 A 與 B 亦無法察覺”。

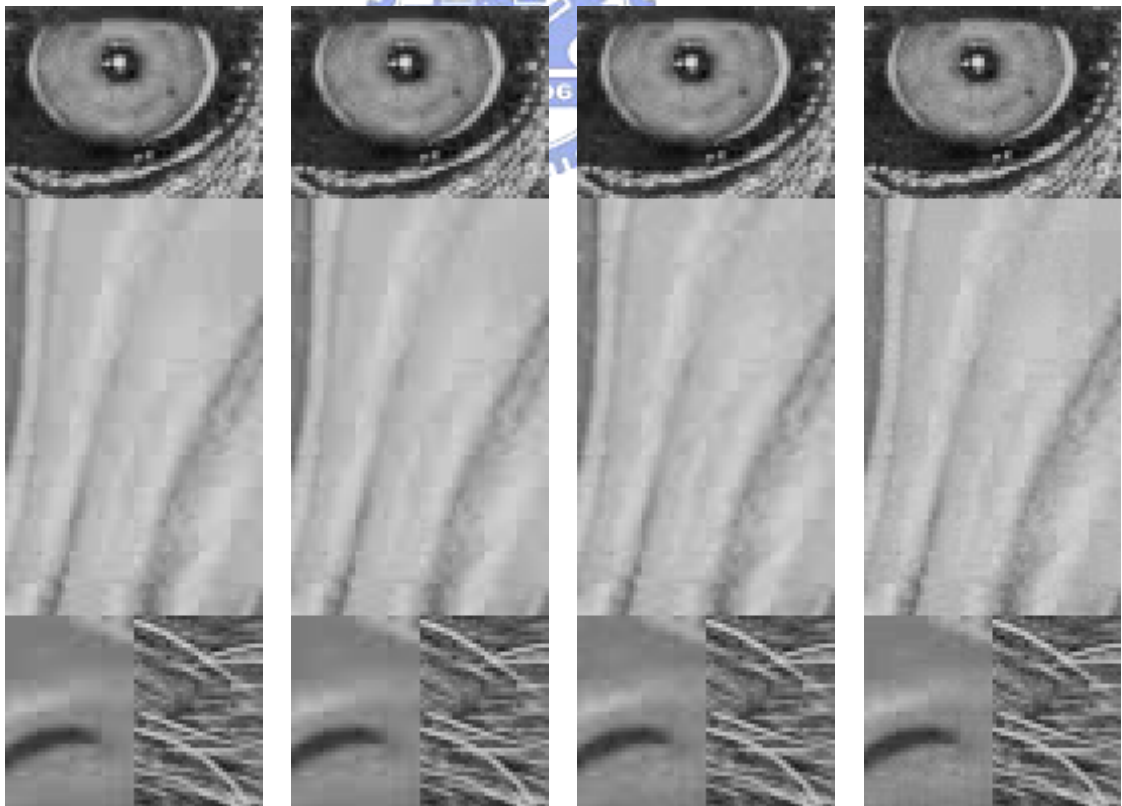


(a-1)[222]

(a-2)[333]

(a-3)[444]

(a-4)[555]

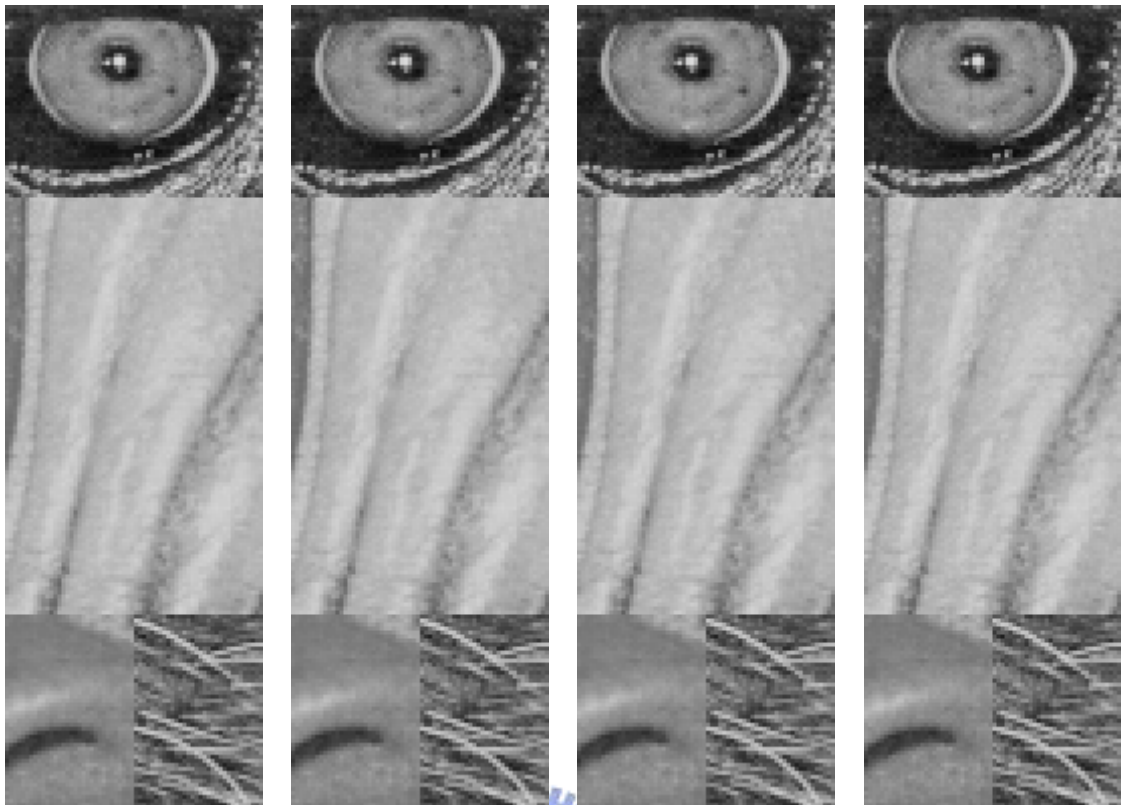


(b-1)[222]

(b-2)[322]

(b-3)[232]

(b-4)[223]

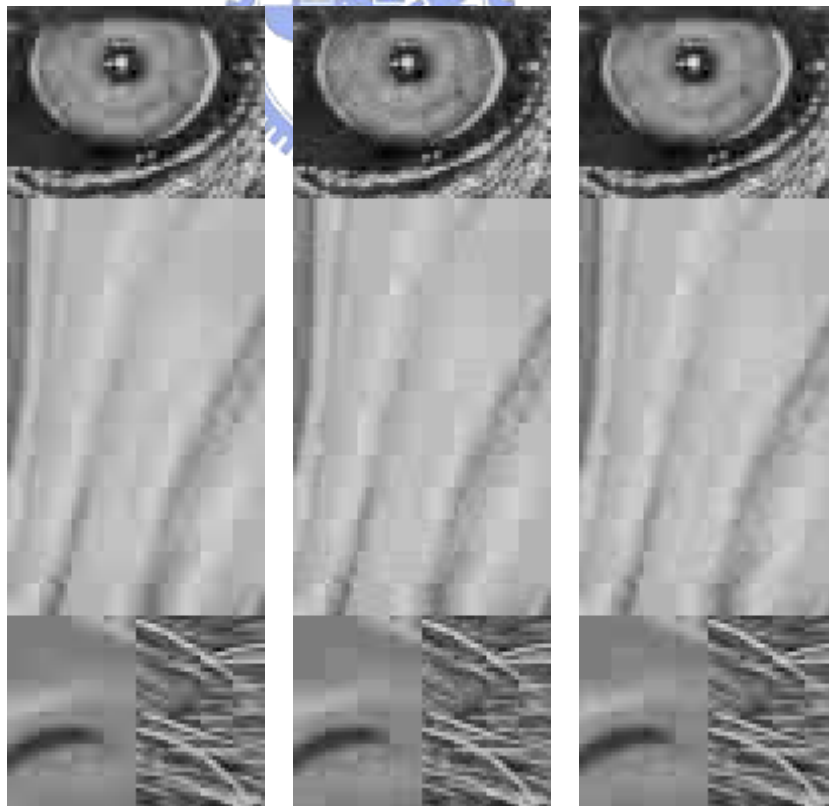


(b-5)[333]

(b-6)[433]

(b-7)[343]

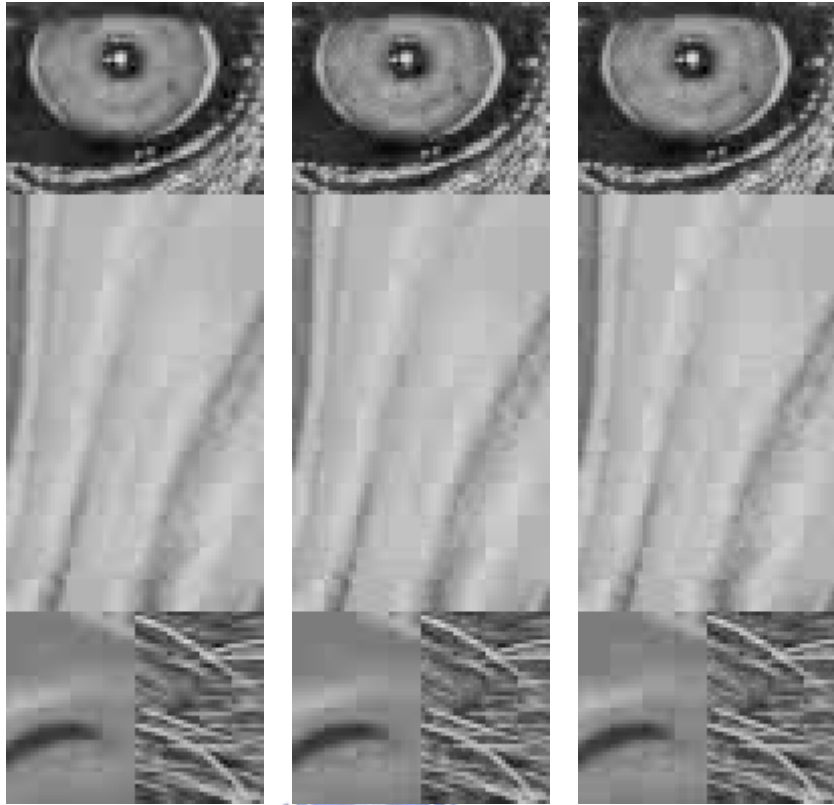
(b-8)[334]



(c-1)[211]

(c-2)[112]

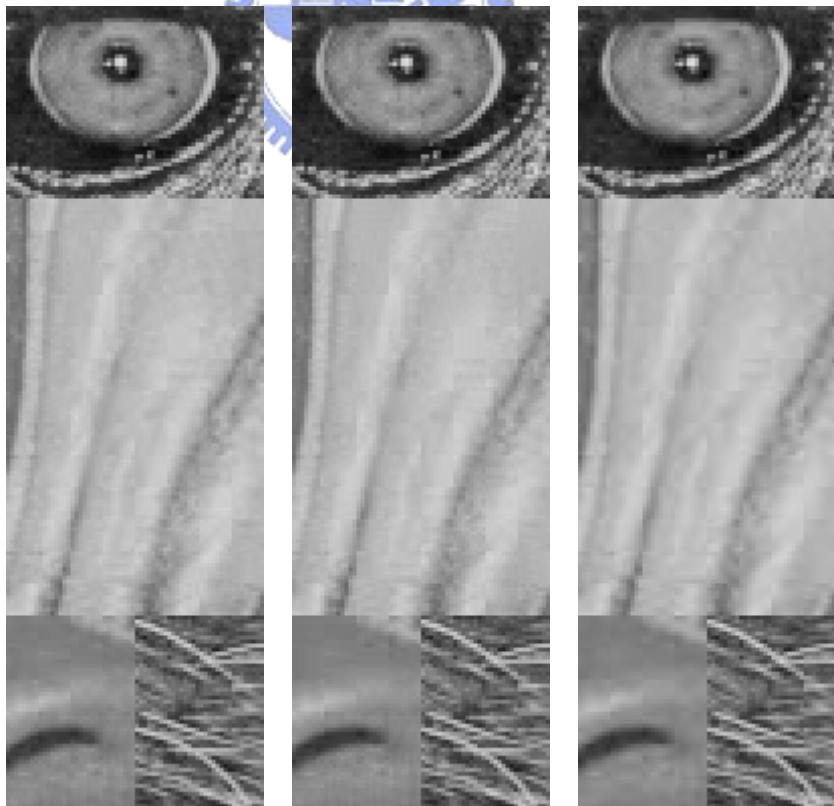
(c-3)[121]



(c-4)[221]

(c-5)[212]

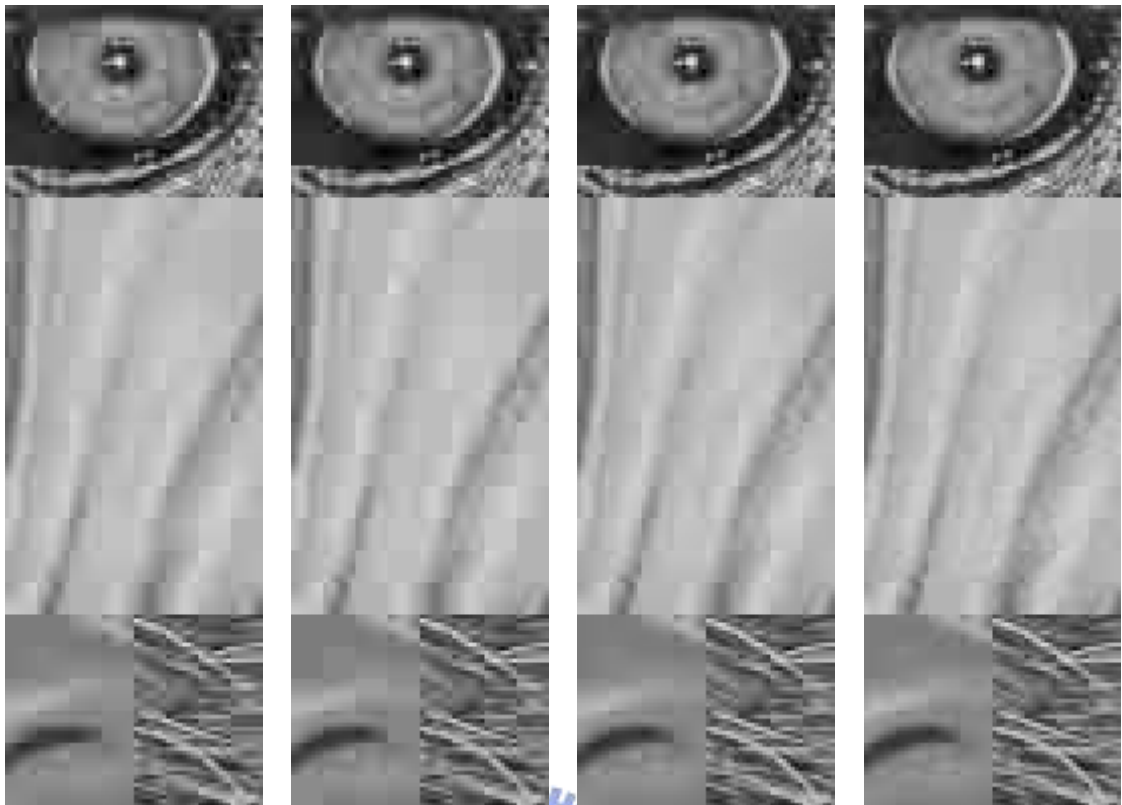
(c-6)[122]



(c-7)[233]

(c-8)[323]

(c-9)[332]

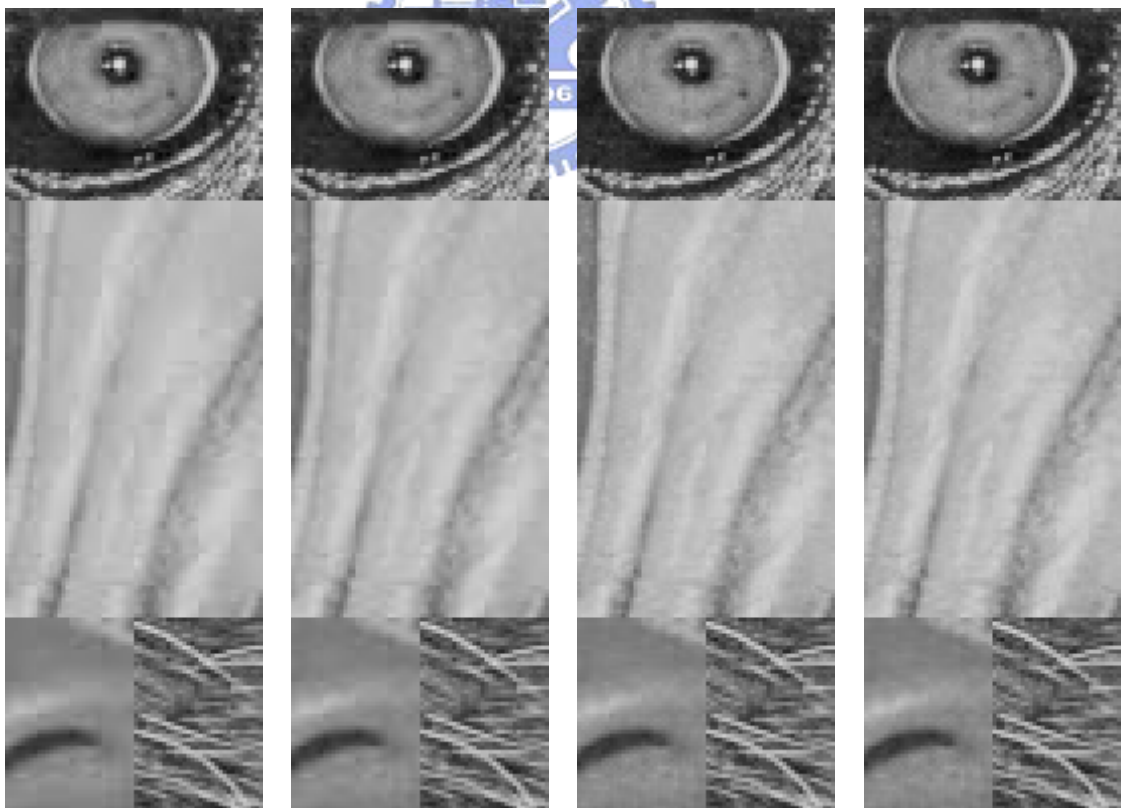


(d-1)[200]

(d-2)[110]

(d-3)[311]

(d-4)[221]

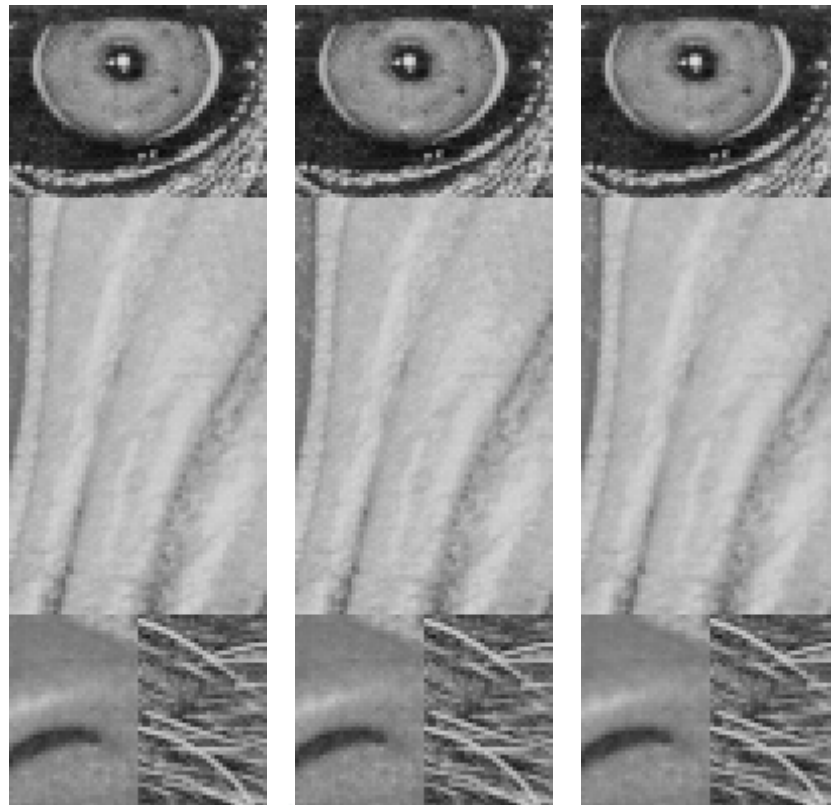


(d-5)[422]

(d-6)[332]

(d-7)[533]

(d-8)[443]



(e-1)[555]

(e-2)[666]

(e-3)[554]

圖 3.3 視覺實驗比較圖 – baboon 部分放大

3.1.2 小結

從視覺實驗的結果得知：

1. 在資料量低時(平均兩個位元層以下)，大抵上，低頻段資訊比較重要
2. 在資料量高時(平均三個位元層以上)，額外的高頻資訊將會獲致較大的品質改善
3. 不同測試者的測試結果具有相當的差別
4. 平均位元層介於 2~3 之間時，不同測試者的結果較一致
5. 最高位元層的資訊其視覺效果差異非常低

由以上的結果來看，架構在目前 MPEG-4 編碼規格上的「頻率資訊選擇」仍具有可認知的視覺改善。而且在我們比較感興趣的位元率區間，同時也是影像品質增進較大的位元率下，其改善效果更明顯。故我們認為「頻率資訊選擇」在 MPEG-4 編碼架構上具有良好的視覺改善效果。

另外，在實驗的過程中，很多測試者有相同的疑問：“要我選擇比較像的，還是比較好看的”（在這次實驗中他們都被告知應該選擇比較像原始影像的）。故壓縮技術的評比除了保留視訊資訊的能力外，主觀喜好的最佳化也能當做評比的標準。然而主觀喜好的評比

方式很難有公正的測量方式來做進一步的驗證，但頻率資訊選擇已經提供針對主觀喜好調整的能力，而且是在編碼過後調整，故可以在僅一次編碼過後，針對不同的主觀要求調整，滿足不同使用者的不同需求。

再次強調，以上的測試與結論皆針對「架構於 MPEG4 規格上之頻率資訊選擇改善視訊品質的能力」設計，其結論不應當作普遍影像品質的品質參考準則，因為很多變因沒有受到合理妥善的安排，是直接以我們提出的經過調整的編碼架構為參考。例如：雖然高頻段在高位元層時其重要性大於低頻段，但是高頻段的係數較低頻段多，能量也較大，所以不應將兩者的重要性延伸作高頻低頻資訊的重要比例。

3.2 架構

3.2.1 整體架構

本篇論文提出架構在 MPEG-4 上，具有三頻段頻率資訊選擇能力的視訊編碼，其整體架構如下圖：

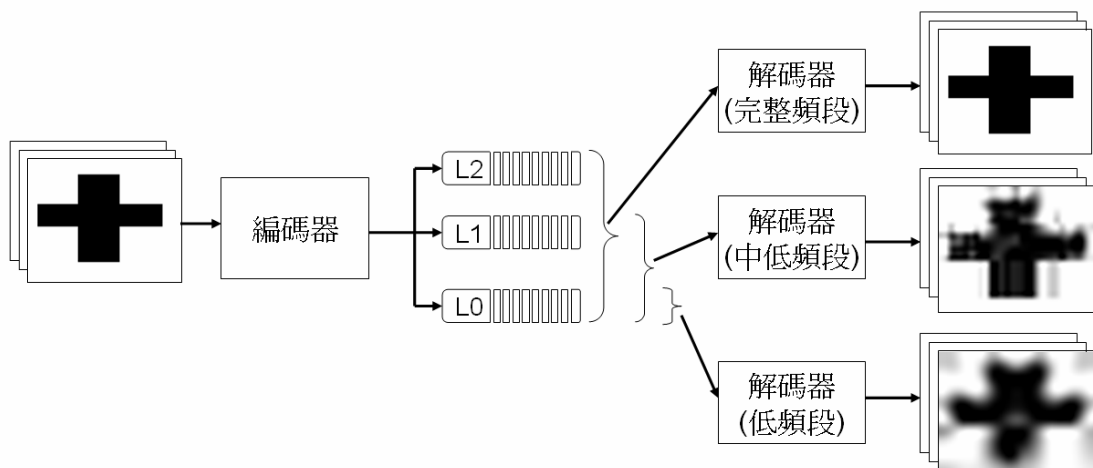


圖 3.4 頻率資訊選擇示意圖

整體架構上，最大的特色在於，編碼過的串流分成三個頻段，每個頻段有各自的基本層與加強層，也就是每個頻段有獨立的雜訊比可調層次式能力。類似 JPEG 的聯合漸進傳送法，但在雜訊比上仍然保有精細的可調層次式(Fine Granularity Scalability)。後面簡述編碼器與解碼器架構上的特徵。

3.2.2 編碼器

3.2.2.1 鋸齒狀掃描

編碼器的架構中，鋸齒狀掃描(Zigzag-scan)與變動長度編碼(Run-Level-Coding)與原先不同；鋸齒狀掃描順序必須調整，而做變動長度編碼時，也必須分為三段式的掃描，我們稱之為分段鋸齒狀掃描(Split-zigzag scan)。

分段鋸齒狀掃描的調整如下：

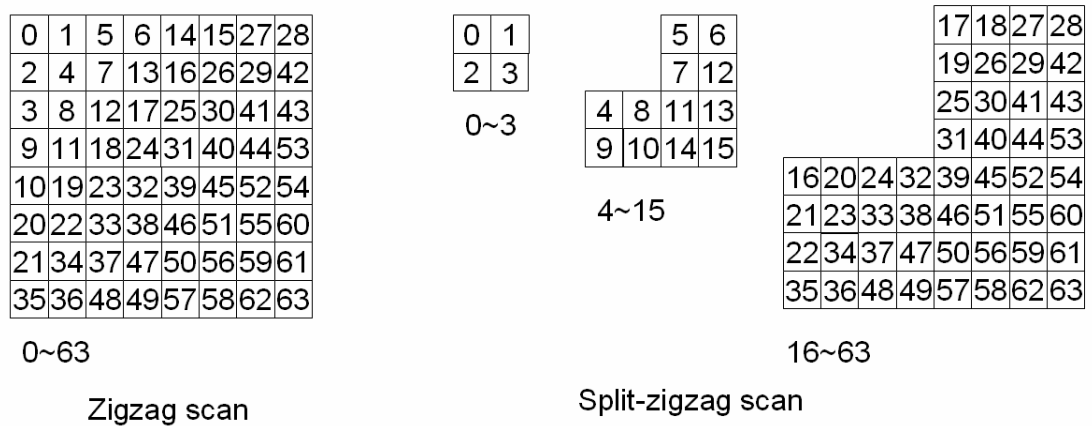


圖 3.5 分段鋸齒狀掃描

重點：

- 把 8x8 離散餘弦係數分成三個頻帶，分別具有 4，12，48 個係數
- 先掃描低頻係數，再掃描中頻係數，最後掃描高頻係數
- 在每一個頻帶中，仍採用鋸齒狀掃描的順序掃描

示意如下圖：

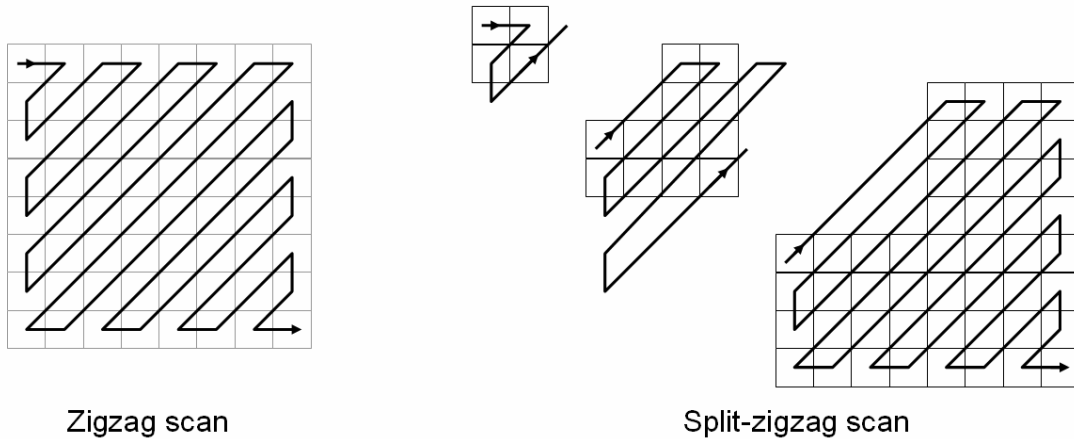


圖 3.6 分段鋸齒狀掃描原理

掃描出來的三個串列經過各自的變動長度編碼，形成三個串流，從低頻、中頻到高频分別稱之為階層-0、階層-1、階層-2(L0、L1、L2)。

分段鋸齒狀掃描應用在兩處：

- 基本層中經過量化的離散餘弦轉換係數
- 加強層中每一層位元平面

故基本層與加強層的係數都能分成三個頻段編碼；而基本層中含有的移動向量及其他解碼所需資料，一併放在基本層的 L0 當中，因為無論解碼端如何取舍資料，基本層 L0 都是必要的。

3.2.2.2 可變長度編碼

因為除了更改鋸齒狀掃描順序，掃描也分為三段式的掃描，故會多了兩個「斷點」，我們在基本層與加強層對於斷點的處理不同，分述如下

在基本層中，因為較低頻段的串流一定在較高頻段的串流之前接收；意即，L0 收完，才會收 L1，不可能不還原 L0 就還原 L1，故我們採取「連續式可變長度編碼符號」，舉例說明之：

表 3.9 基本層離散餘弦係數例子

97	0	8	0	0	0	0	0
-4	-5	0	0	0	0	0	0
0	-1	0	0	0	2	0	0
0	0	0	0	0	0	0	0

0	0	3	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

分段鋸齒狀掃描完的順序為：& 代表頻段分隔處

97,0,-4,-5&0,8,0,0,-1,0,0,0,0,0,0&0,0,0,0,0,0,0,3,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0;

如果採用原來的可變長度編碼，將獲得如下的符號：(Run,Level,End)

(0,97,0) (1,-4,0) (0,-5,0) (1,8,0) (3,-1,0) (9,3,0) (15,2,0) (15,1,1)

所謂連續式可變長度編碼符號，就是不因頻段分隔而截斷符號，若是有符號跨越多個頻段，則此符號將歸屬於其中最低的頻段。上例將會被分成三個頻段：

L0：(0,97,0) (1,-4,0) (0,-5,0)

L1：(1,8,0) (3,-1,0)

L2：(9,3,0) (15,2,0) (15,1,1)

如此，綜合 L0、L1、L2 三個頻段的基本層，其資料量將與原來未分割式的掃描一樣。

在加強層中，由於每個頻段的雜訊比可調層次式是獨立的，所以每個頻段的資料必須能獨立的解碼還原，故不能如同基本層採用連續式可變長度符號編碼。我們採取的方法是，把三個頻段當成獨立的序列分別做可變長度編碼，意即，會有三個區塊結束符號(EOB, End Of Block)。考慮對如下的位元層進行編碼：

表 3.10 離散餘弦轉換係數位元層例子

0	1	0	0	0	0	0	0
0	0	0	0	0	0	1	0
0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0

分段鋸齒狀掃描完的順序為：& 代表頻段分隔處

0,1,0,0&0,0,0,0,1,0,0,0,0,0,1,0&0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0;

如果採用原來的可變長度編碼，將獲得如下的符號：(Run, End)

(1,0) (6,0) (5,0) (13,0) (2,0) (14,1)

爲了使每個頻段能獨立編碼解碼，我們必須先把掃描好的序列分成獨立的三段：

L1 : 0,1,0,0 &

L2 : ,0,0,1,0,0,0,0,0,1,0 &

L3 : ,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,
0,0,0,0,0;

然後各自獨立做可變長度編碼：

L0 : (1,1)

L1 : (4,0) (5,1)

L2 : (12,0) (2,0) (14,1)

如此一來，分段做可變長度編碼，其編碼符號總數並沒有增加，但是內容變動了。但在整個頻段都爲 0 的情況下編碼符號總數會增加，舉例如下：

分段鋸齒狀掃描完的順序爲：& 代表頻段分隔處

0,1,0,0&0,0,0,0,0,0,0,0,0,0,0,0,0&0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0;

如果採用原來的可變長度編碼，將獲得如下的符號：(Run,,End)

(1,0) (26,0) (2,0) (14,1)

如果獨立做可變長度編碼：

L0 : (1,1)

L1 : (15,0) → 一個全爲零的頻段

L2 : (12,0) (2,0) (14,1)

註：(15,0)在長度爲 16 的序列中，即代表「全零」的符號

可得：一個全爲 0 的頻段，將會帶給我們額外的一個編碼符號。



3.2.1 解碼器

解碼器的架構也只有變動長度解碼與逆鋸齒狀掃描需要變更。其變更完全對應於編碼器的動作。

3.3 表現改進 (Performance Improvement)

由於編碼過程的鋸齒狀掃描順序更動，同時也分成三個頻段獨立掃描，故可變長度編碼的符號也將與原來不同。接下來我們將探討鋸齒狀掃描的變動，會帶來什麼影響，而我們將如何調整，來提高壓縮率。最後比較改進前後的表現。

3.3.1 更動的掃描帶來的影響

我們用實際的例子來觀察更動的掃描對於可變長度編碼符號帶來的影響：

某一位元層採用原來的可變長度編碼獲得如下的符號：(Run,,End)

(1,0) (6,0) (5,0) (13,0) (2,0) (14,1)

如果各自獨立做可變長度編碼：

L0 : (1,1)

L1 : (4,0) (5,1)

L2 : 12,0) (2,0) (14,1)



很明顯的，有兩個符號變動，而在原來的霍夫曼編碼表中(Huffman Codebook)，(1,1)符號比(1,0)符號位元數還要多，(5,1)符號也比(5,0)符號位元數還要多。所以雖然符號數目一樣是六個，可是後者的編碼總位元數比前者多，其原因在於編碼符號的統計特性改變。進一步觀察，我們可以發現分段鋸齒狀掃描與獨立可變長度編碼帶來下列的影響：

- 整體的靜態統計特性改變
 - 符號所跑的長度(Run)平均變短
 - 區塊終止符號(EOB)出現頻率增加
 - 對於每個頻段的影響不盡相同
- L0 只有四個係數，做可變長度編碼不夠有效率，也稍嫌複雜

3.3.2 調整方式

3.3.2.1 霍夫曼編碼表

由於每個頻段的靜態統計特性相差甚遠，我們決定分別調整其熵編碼。綜合我們所觀察到的現象，我們針對三個頻段做出不同的調整：

- L0：

由於 L0 頻段只有四個係數，包含 16 種情況，所以我們直接統計 16 個情況的機率，然後施以霍夫曼編碼。

- L1：

由於 L1 掃描序列長度為 16，若是利用原來的霍夫曼編碼表，有一些符號根本不會用到(Run>15 的符號)，所以我們統計了可能出現的符號的靜態統計特性，重新建立了霍夫曼編碼表。

- L2：

L2 掃描序列長度為 48，原先的霍夫曼編碼表可以完全應用，我們嘗試修改霍夫曼編碼表，但其改進不明顯，故仍沿用原來的霍夫曼編碼表。



3.3.2.1 位元層編碼

原先的位元層編碼，同一層 64 個位元為一個單位，然後定出最高有效位元位元層，開始編碼，直到最低有效位元層。舉例說明之：

表 3.11 加強層離散餘弦係數例子

17	0	8	0	0	2	0	0
-4	-5	0	0	0	0	0	0
0	-1	0	4	0	2	0	0
0	0	0	0	0	0	-2	0
0	0	3	0	0	0	0	0
0	3	0	0	1	0	0	1
0	0	0	0	0	0	0	0
	0	0	0	0	-1	0	0

上圖的 8x8 矩陣作位元層編碼，由於絕對值最大為 17，需要 5 個位元來表示；將最高有效位元位元層編號為 0，則最低有效位元位元層編號為 4，總共五層。但是在上述例子中，L0 裡最大的絕對值是 17，需要 5 個位元來表示；L1 裡最大的絕對值是 8，需要四個位元來

表示；L2 裡最大的絕對值是 3，需要兩個位元來表示。故每一個頻段的最高有效位元不同，若是從同一層開始編碼，將會在 L1 的第一位元層，L2 的前三個位元層傳送不必要的「全零符號」。

所以我們最後採取的方式，三個頻段的離散餘弦轉換係數經過分段矩齒狀掃描以後，就各自獨立進行編碼，示意如下圖：

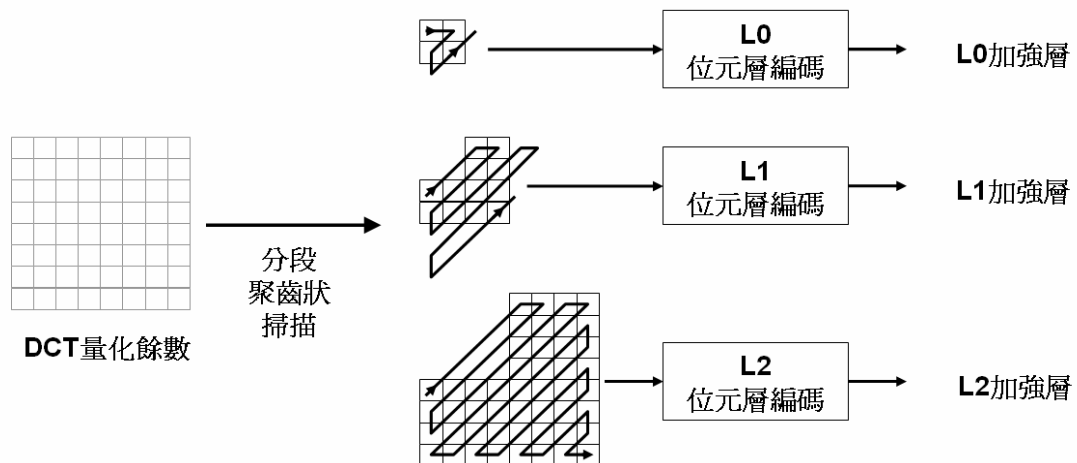


圖 3.7 更改過後之位元層編碼

3.3.3 效果

我們首先比較整體的壓縮效率，即比較編碼出來的加強層的資料量；然後比較在我們有興趣的位元率範圍裡，頻率資訊選擇所能提供的影像品質增強；最後提出高位元率的壓縮品質提升的可能。

為了維持測試資料的一貫性，我們必須使用長寬皆為 64 倍數的視訊(下一章會說明原因)，故我們將 CIF(352x288)大小的測試視訊擷取左上角的 320x256 部分當作原始視訊。為了避免混淆，我們將測試視訊的名稱作縮短以示區別：「akiyo」→「aki」、「foreman」→「fman」、「mobile」→「mob」、[stefan] → [stef]。

3.3.3.1 壓縮效率

表 3.12 壓縮效率的比較

(Kbyte)	L0	L1	L2	3-level	MPEG-4	increase(%)

aki	1,520	3,756	8,651	13,927	13,496	3.19
fman	1,638	4,559	11,741	17,938	17,344	3.42
mob	1,858	6,158	17,337	25,353	24,597	3.07
stef	1,759	5,522	13,320	20,601	20,026	2.87
total	6,775	19,995	51,049	77,819	75,463	3.12

參數：PbetweenI=19,BbetweenP=2;300frames;QP=28(const)

以四個常見的測試連續影像，畫面大小皆為 320x256，共計 1,200 張畫面編碼的結果，我們提出的具有頻率資訊選擇能力 MPEG4 架構增加了 3.12%的資料量而已，以壓縮效率的觀點來看，新架構的壓縮能力緊跟原本 MPEG4。

3.3.3.2 低位元率視訊品質

MPEG-4 提供了相當廣闊的位元率可調層次式範圍，而頻率資訊選擇更進一步增加可調層次式的維度，故存在非常多的位元率組合。如何系統性的決定最有效的頻率選擇方式，是另一個深且廣的議題，在此我們尚未討論。故我們僅列出某些的位元率下，頻率資訊選擇可以達到的視訊品質改善。

在還沒有完整的可調層次式方式之前，我們採用一個簡單的方式 - 只取 L0 頻段。下面列出 fman 測試連續影像的還原品質：(左為原本架構；右具有頻率選擇能力-只選取 L0 頻段)

參數設定：QP=28 的基本層；取 640Kbps 加強層(2666Kbytes/frame)；BbetweenP=2，Pbetween=19。



frame 00 (I-frame)



frame 01 (B-frame)



frame 03 (P-frame)



frame 04 (B-frame)



frame 12 (P-frame)



frame 13 (B-frame)



frame 21 (P-frame)



frame22 (B-frame)



frame 42 (P-frame)



frame 43 (B-frame)



frame 57 (The last P-frame in the IPB period)

圖 3.8 低位元率模擬結果 -fman

左：原本架構；右：具有頻率選擇能力-只選取 L0 頻段

「只取 L0 頻段」對於 fman 測試連續影像的效果特別好，加強層大小由 256Kbps 到 768Kbps 時，所有測試者一致認為「頻率資訊選擇」提供了明顯的品質提升，尤其是在每個 IPB 週期的開始；也就是當區塊效應嚴重的時候，加強低頻段資訊能明顯提升 fman 的視訊品質。

仔細探討其影像品質，會發現兩者並沒有絕對的優劣。以 fman 最主要的人臉來說，原本的編碼架構有著極明顯的區塊效應，而加強 L0 低頻段以後，區塊效應減輕，看起來較舒服；如果注意後方建築物的稜線，原本的架構還原影像具有清楚的邊界，而右圖犧牲高頻段資訊的結果是模糊的邊界，就這點來說，「只取 L0 頻段」並沒有加強畫面品質(雖然也有測試者偏好模糊邊界)。總合來說，在這個測試影像中，人臉的重要性遠大過於背後建築物，故所有的測試者皆認為右圖品質較佳。

根據以上的討論，我們可以粗略歸納出「只取 L0 頻段」適用的條件：

1. 區塊效應明顯時 → 通常是加強層位元率不高的時候
2. 影像的主要部分沒有銳利邊界時

然而同樣的方式對其他測試影像的效果就不見得那麼好，下面是 aki 的測試結果：

參數設定：QP=28 的基本層；取 256Kbps 加強層(1066Kbytes/frame)；BbetweenP=2，Pbetween=19。



frame 00 (I-frame)



frame 01 (B-frame)



frame 03 (P-frame)



frame 04 (B-frame)



frame 12 (P-frame)



frame 13 (B-frame)



frame 21 (P-frame)



frame 22 (B-frame)



frame 42 (P-frame)



frame 43 (B-frame)



frame 57 (The last P-frame of the IPB period)

圖 3.9 低位元率模擬結果 - aki

左：原本架構；右：具有頻率選擇能力-只選取 L0 頻段

由於 aki 測試影像的壓縮率較高，在 512Kbps 的加強層時就沒有區塊效應了，故在 256Kbps 區塊效應還在，加強低頻段資訊便有其功用。與 fman 相同的，aki 的主題也是人臉；不同的是，aki 中的主播臉部特徵很明顯，無論是眉毛或是嘴巴，都有著銳利且深刻的邊界，故捨棄高頻段的結果早成明顯的漣波(Ripple)，對於影像品質的傷害很大。但另一方面，加強低頻段資訊也明顯的消除了臉部平滑區域與衣服上的區塊效應，致使較佳的影像品質。在取 256Kbps 加強層的情況下，超過一半的測試者認為「頻率資訊選擇」提供了較佳的影像品質，在更高的位元率時，由於區塊效應消失，使得我們提供的改善不明顯，但犧牲的高頻段資訊造成的漣波仍在，故我們提出的方法便無法提升影像品質。

請參考下圖：

參數設定：QP=28 的基本層；取 512Kbps 加強層(1066Kbytes/frame)；BbetweenP=2，Pbetween=19。



frame 00 (I-frame)

圖 3.10 中位元率模擬結果 -aki

左：原本架構；右：具有頻率選擇能力-只選取 L0 頻段

3.3.3.3 高位元率壓縮品質

由於人眼對於高頻段分量較不敏感，基本層的編碼也有考慮到這點，而採用不均勻量化。然而加強層卻只是單純的把所有離散餘弦轉換係數以整數精確度編碼，這對於失真編碼(MPEG4 即是失真編碼)來說不太正確。失真編碼不注重完整的保留資訊，注重的是如何在品質與壓縮率下取得最佳平衡，故我們認為加強層中較不重要的高頻，可以犧牲其精確度來換取壓縮率。

以下我們在 MPEG4 與我們的架構所形成的加強層上作任意的可調層次式，然後解碼，企圖在「沒有人眼可感知差異」的條件下減少資料量。根據第一節所做的視覺實驗可知，MPEG-4 架構在捨去最低有效位元層的情況下人眼無法發現差異，而頻率資訊選擇可以在第二層額外捨去 L0 或 L1 仍然無法察覺；同時，L2 頻段的資料量較大。故為了增加壓縮效果，我們多捨去 L2 一層位元層。以下是捨去超過人眼感知程度的資訊後，加強層壓縮效率的比較：

表 3.13 高壓縮率考量隻下的壓縮結果

	L0	L1	L2	3-level	MPEG-4	decrease(%)
aki	1,189	2,417	1,609	5215	7,269	28.26
fman	1,321	3,317	3,704	8342	11,244	25.81
mob	1,559	5,036	8,911	15506	18,988	18.34
stef	1,490	4,347	5,381	11218	14,156	20.75
total	5559	15117	19605	40281	51657	22.02

平均來說，在高效率壓縮的考量下，「頻率資訊選擇」可以提供我們超過 20%的額外壓縮效率。

