

第二章 先前的研究

這一章分成兩個部分：第一部分包含第一節，簡介空間資訊選擇在影像壓縮上的應用；第二部分包含第二、三、四節，先介紹 MPEG4 編碼規格；然後簡介空間解析度可調層次式編碼的大宗 - 金字塔型空間解析度可調層次式；接著討論視訊編碼在空間解析度可調層次式上重要的議題 - 移動補償。最後再簡介另一種空間解析度可調層次式的方式 - 小波轉換(Wavelet Transform)空間解析度可調層次式編碼。

2.1 JPEG 影像編碼



頻率資訊選擇對於影像編碼來說非常重要，因為人類視覺對於低頻成分比較敏感，所以壓縮影像的過程中，會保留較多的低頻成分。以下我們簡介 JPEG 影像壓縮編碼，並指出幾個與頻率資訊選擇相關的部分說明。

2.1.1 JPEG 基本系統

下圖為 jpeg 基本系統之概略方塊圖。以下一一說明之

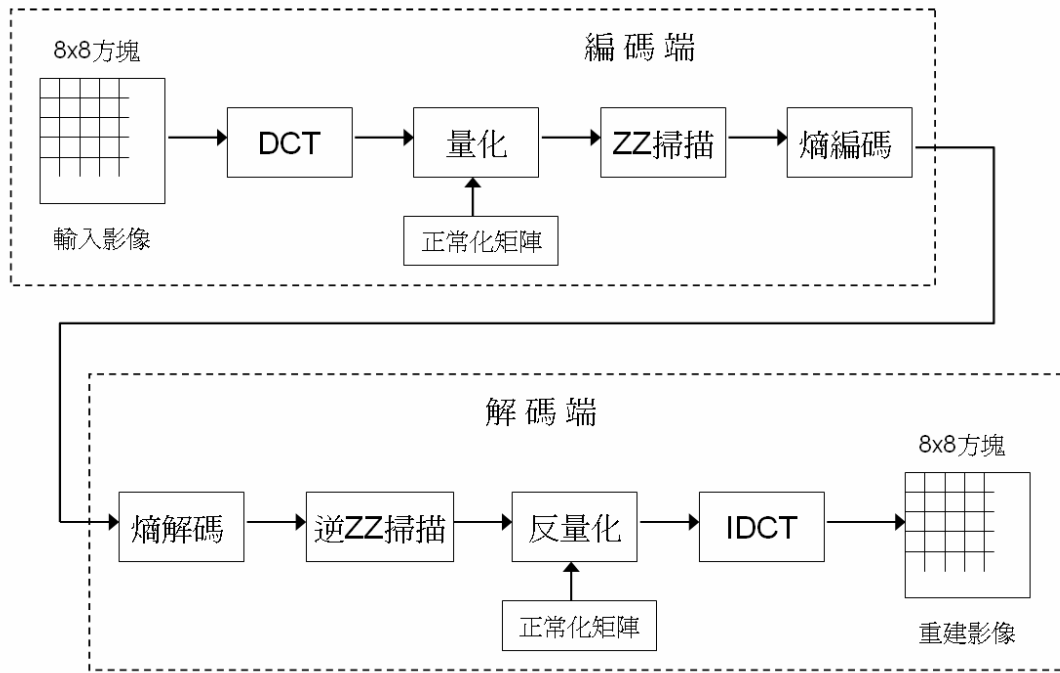


圖 2.1 JPEG 系統流程圖

重點：

- 首先把原影像切割成不重疊的 8x8 方塊，每個方塊各自做 DCT 轉換
- 參考正常化矩陣，對於 64 個 DCT 係數做不同量化位階之量化
- 以鋸齒狀掃描(Zigzag Scan)把 64 個係數串連為一維向量
- 利用設計過的編碼方式，不失真的壓縮一維向量
- 解碼端則反向操作上述步驟

由於本篇論文提出的方法架構在 MPEG-4 之上，所以上面我們對於 JPEG 編碼只稍做介紹；而且，JPEG 編碼與解碼過程中的步驟，在 MPEG 編碼解碼過程中也有；有些完全一樣，有些非常相似；故在下一節我們介紹 MPEG 與 JPEG 的細節。接下來，針對 JPEG 規格中，與「頻率資訊選擇」相關的部分 - 正常化矩陣與漸進式傳送，做進一步的說明。

2.1.2 正常化矩陣 (Normalization Matrix)

使用一個自行選定的正常化矩陣(通用於所有的方塊)將所有的 DCT 係數都正常化(Normalized)，即除以縮小倍數。正常化矩陣的每一個元素都是 8 位元的正整數。這 64 個 8 位元的正整數當作額外資訊先送給接收端。正常化以後的係數接著被量化為整數。正常化矩陣也可以視為量化器之放大縮小，以便於控制每個係數所可能產生的量化誤差值。我們可以使用 HVS(Human Visual System)對比敏感係數來協助正常化矩陣的設計。如此正常化矩陣會依各係數之視覺重要性而定出放大縮小之倍數。

2.1.3 漸進式傳送 (Progressive Transmission)

使用漸進式傳送時，影像資料會分成好幾次傳送；每次的傳送都可以幫助接收端得到更接近於原影像的重建訊號。使用漸進式傳送的動機起源於我們需要透過低頻寬的傳輸線(相對於資料量)，例如電話線，來傳送影像資料，尤其是當我們需要做到快速辨認或者整個傳送時間可能很有限的情況下。典型的例子包括電子瀏覽(Telebrowsing)：透過遠距離的終端機從一個資料庫中搜尋某張影像；及軍事用途：使用者只能使用到很有限的通訊設備。當然，傳統的壓縮演算法也會降低傳送時間，但是它們一般都得等到影像的大部分都已經重建起來後才能做辨認的工作，而且也只允許一種固定的重見影像品質。

使用漸進式傳送很有可能會讓我們省下一大筆資料不需要傳送：一旦目前收到的中間版本影像已經足以讓人滿意或者已經可以看出來並不是我們所要的影像，那麼傳送的程序就可以馬上中止。這種由使用者終止所導致的送資料之減少我們稱之為有效壓縮(Effective Compression)。

JPEG 提供三種漸進式傳送方法：

一. 頻譜選擇法(Spectral Selection)

它的做法是將轉換係數分成幾次來送。依照鋸齒狀掃描的順序，先送低頻係數而後送高頻係數。例如，我們可以如下圖所示的第一次只送 DC 係數，第二次送 AC1 及 AC2，依此類推。由於高頻係數都留到後面才送，因此早先所收到的影像很模糊。

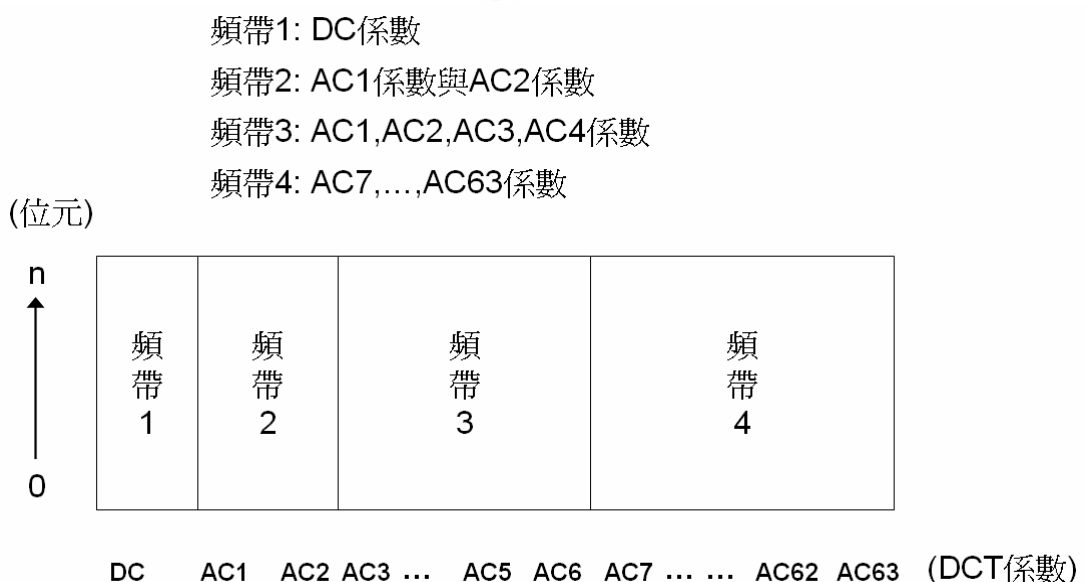


圖 2.2 JPEG 漸進式傳送 - 頻譜選擇法

二. 逐漸接近法(Successive Approximation)

它的做法是每一次都送所轉換係數的的幾個位元。依照最高有效位元(Most Significant Bit)到最低有效位元(Least Significant Bit)的順序，先送重要的位元，然後送次重要位元。例如，我們可以如下圖所示的第一次只送所有係數的前四個位元，第二次送接著的兩個位元，以此類推



圖 2.3 JPEG 漸進式傳送 - 逐漸接近法

三. 聯合法

顧名思義，它是前兩種方法的混合型。根據 Pennebaker 的實驗報告顯示，將 DC 係數完整送出並且降低所有 AC 係數的精密度(使用較少位元來表示 AC 係數)使得整體位元率為 0.24 位元/像素，其結果比位元率為 0.36 位元/像素的頻譜選擇法(送出完整的 DC 係數以及前五個 AC 係數)還好一些。

這個結果說明以頻率順序來傳送資料並不有效；而低頻的重要性更是早已經被認定；所以綜合兩者，我們可以把低頻資料傳得精細些，高頻部分傳得粗略些，然後再慢慢補足。如此一來將會有非常多種傳輸的順序，下圖所示為聯合法的一種安排方式。

傳輸順序:

[DC 1] → [AC 1] → [AC 2]
 → [AC 3] → [AC 4] → [AC 5]
 → [DC 2] → [AC 6]

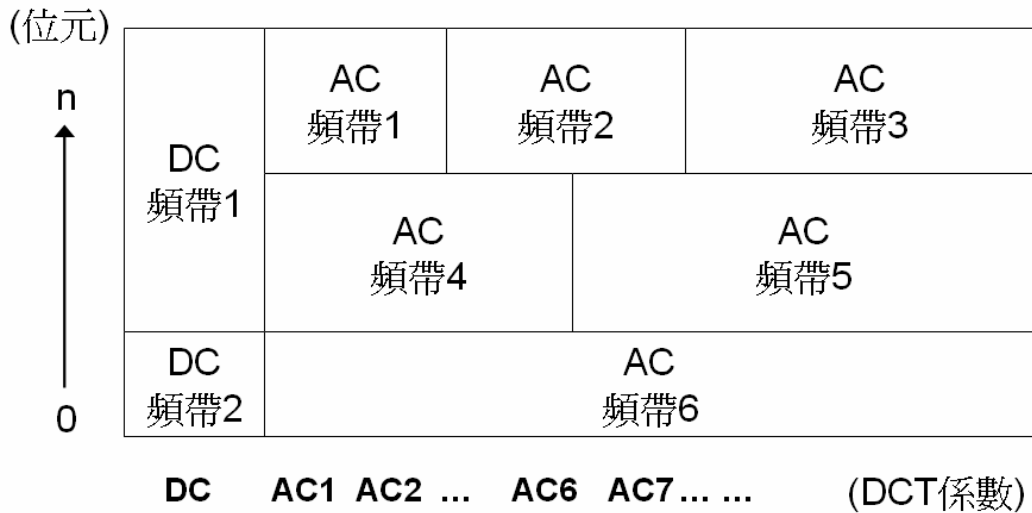


圖 2.4 JPEG 漸進式傳送 - 聯合法

2.1.4 階級式編碼



在許多情況下，壓縮過的資料必須同時提供給不同解析度或者不同品質要求的顯示器及傳送設備，例如一台高品質數位電視機，一台高解析度的列印設備，一台透過電話線接收資料的低解析度傳送器等等。為了應付這一類的應用，JPEG 標準也提供了階級式編碼：將影像表示成不同解析度的階級，換句話說，其編碼過後的串流可以還原成不同解析度的重建影像。這種方法一般便導致所謂的金字塔型結構(Pyramid Structure)。其做法與金字塔型空間解析度可調層次式視訊編碼相同，在第三節中將有詳細說明。

2.1.5 JPEG2000

JPEG2000 是目前影像壓縮規格中效率最好的，跟 JPEG 比起來，JPEG2000 具有以下特性：

- 使用小波編碼
- 超低位元率的壓縮
- 同時可以做灰階與二階影像之壓縮

- 較能容忍網路上傳送時所面對的雜訊
- 靜態與動態 ROI (Region Of Interest)
- 可以依雜訊比或解析度做漸進式傳送

JPEG2000 因為採用小波轉換，小波轉換的過程中即提供了階級影像架構，故 JPEG2000 不需要階級式編碼；也因此，JPEG2000 提供的漸進式傳送方法是依解析度(另外一種是依照雜訊比)來做分割，因為在小波轉換的係數中，不同頻段即可用不同解析度還原。

2.2 MPEG4 視訊編碼

以前的 MPEG 格式大多數只與壓縮技術有關，MPEG-4 則填加了很多新的特色，例如位元率的可伸縮性 (Bit Rate Scalability)、動態子畫面 (Animated Sprites)、互動性、甚至還有版權保護。

下圖為 MPEG 編碼器之方塊圖：

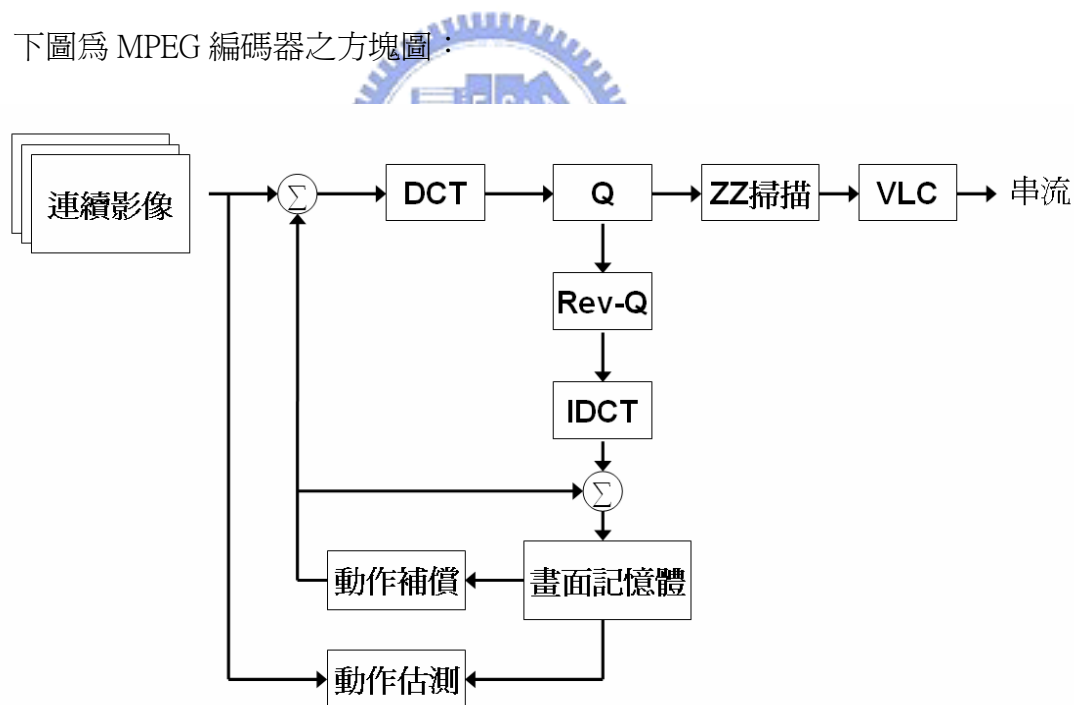


圖 2.5 MPEG4 編碼流程圖

下面將簡介 MPEG-4 編碼中比較重要、同時對空間解析度可調層次式也有影響的運算。

2.2.1 移動補償(Motion Compensation)

移動補償是相當基本的視訊編碼過程，由於一般視訊具有大量的時間性冗餘(Temporal Redundancy)，經由移動補償則可以大幅減低資料量，達到壓縮的目的。移動補償的過程簡述如下：

- 區塊基礎 (Block-based)
把畫面切割為許多小單位，通常為正方塊。
- 移動估測 (Motion Estimation)
對於每一個方塊，在前一張畫面中尋找最相似的方塊，然後記錄其位置所記錄的位置，即是估測出來移動前的位置
- 移動補償
把目前的方塊的影像數值減去移動估測對應的方塊影像數值

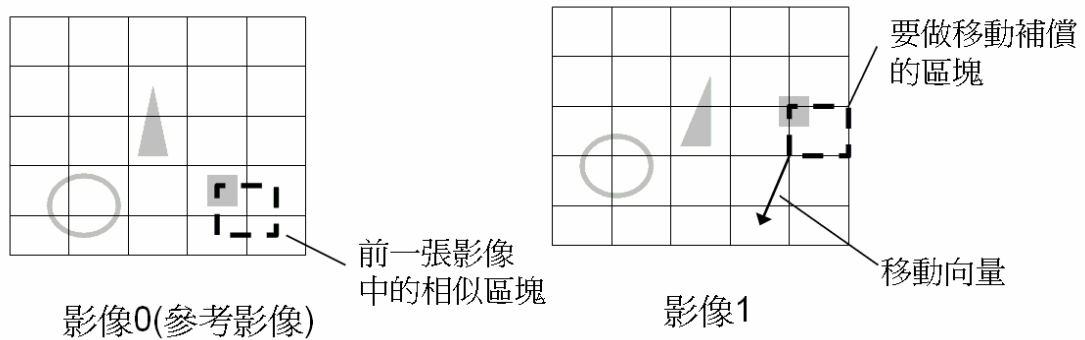


圖 2.6 移動補償示意圖

移動補償的步驟，在 MPEG 規格中並沒有很大的變動，不過在空間解析度可調層次式的架構中，移動補償將會成為重要的議題。

2.2.2 離散餘弦轉換 (Discrete-Cosine-Transform)

離散餘弦轉換是影像壓縮與視訊壓縮常見的方法，由於人眼對於高頻訊號比較不敏感，所以影像經過離散餘弦轉換之後，高頻的係數可以比較粗略的量化，如此一來，可以大幅增加壓縮比，而且對於影像品質的影響很小。

而本篇論文提出的方法，就是直接分割二維離散餘弦轉換各個頻率係數，達到空間解析可調層次式。

2.2.3 鋸齒狀掃描 (Zigzag Scan)

二維的離散餘弦轉換係數，經過量化後，高頻部分數值很小，幾乎皆為零，能量集中在低頻係數。而鋸齒狀掃描將係數由低頻排列至高頻，高頻的零被集中排列，這樣子的排列順序經由變動長度編碼(Run Level Coding)，可達到可觀的壓縮比。

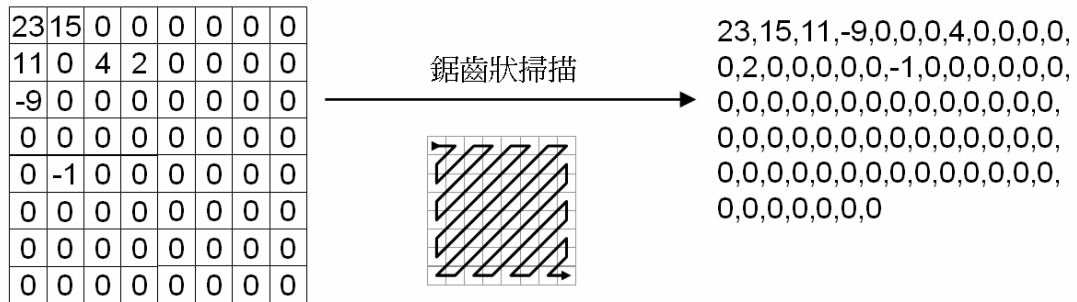


圖 2.7 鋸齒狀掃描

2.2.4 變換長度編碼 (Run-Level-Coding)

鋸齒狀掃描出來的序列，能量集中在前面，後面大多是零，MPEG-4 採用變換長度編碼對序列壓縮。變換長度編碼能非常有效的壓縮鋸齒狀掃描後的序列，延續前一小節的例子，我們只要用 7 個變換長度編碼符號(Symbol)即可完成長度為 64 的序列的編碼：

(Run, Level, End) → (0,23,0) (0,15,0) (0,11,0) (0,-9,0) (3,4,0) (5,2,0) (5,-1,1)

而各個不同的符號，根據機率的統計，再施以霍夫曼編碼(Huffman coding)。

2.2.5 位元平面編碼 (Bit Plane Coding)

位元平面編碼是 MPEG-4 達成「可精緻可調層次式(Fine-Granularity-Scalability)」的主要步驟。離散餘弦轉換的係數，經過量化後，編碼成爲基本層(Base-layer)。而量化誤差則進一步經由位元平面編碼，編碼成加強層(Enhancement-layer)，可以擷取不同資料量解碼，還原出不同的影像品質。位元平面編碼的過程簡述如下：

- 把量化誤差以二進數表示，並且對齊最小有效位(LSB)
- 找出所有數值中的最高有效位元(MSB)

- 從最高有效位元開始，一層一層的對每個位元層編碼
- 每一個位元層，分別做鋸齒狀掃描與變換長度編碼

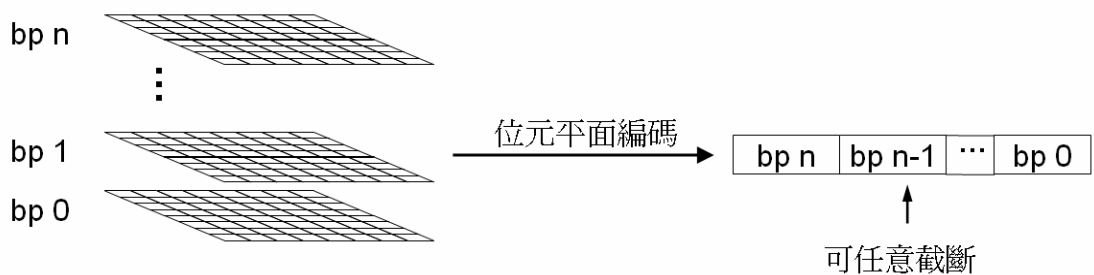


圖 2.8 位元平面編碼

如此一來，接收端若是接收了全部的資料，則可以還原出完整的數值；接收端若是接收了一部分的資料，還是可以還原出所有的數值，但比較不精確。

2.2.6 可調層次式(Scalability)

廣義來說，MPEG-4 的可調層次式有四種，雜訊比可調層次式、時間解析度可調層次式、空間解析度可調層次式、還有資料分割(Data Partitioning)。其中資料分割的目的比較著重資料保護，不如前三者著重於精細可調層次式的實現。接下來我們分別簡述前三項。

2.2.6.1 訊雜比可調層次式(SNR Scalability)

這裡所謂的雜訊，並不是傳輸上的通道雜訊。一般編碼規格所謂的訊雜比，其雜訊(Noise)是指量化雜訊(Quantization Noise)，而 MPEG-4 可精緻可調層次式編碼(Fine-Granularity-Scalability Coding)主要方法是透過位元平面編碼達成，而位元平面編碼所送達的位元平面層數多寡，直接影響到視訊訊雜比，故訊雜比可調層次式經由位元平面編碼，即可達到相當精細的程度。

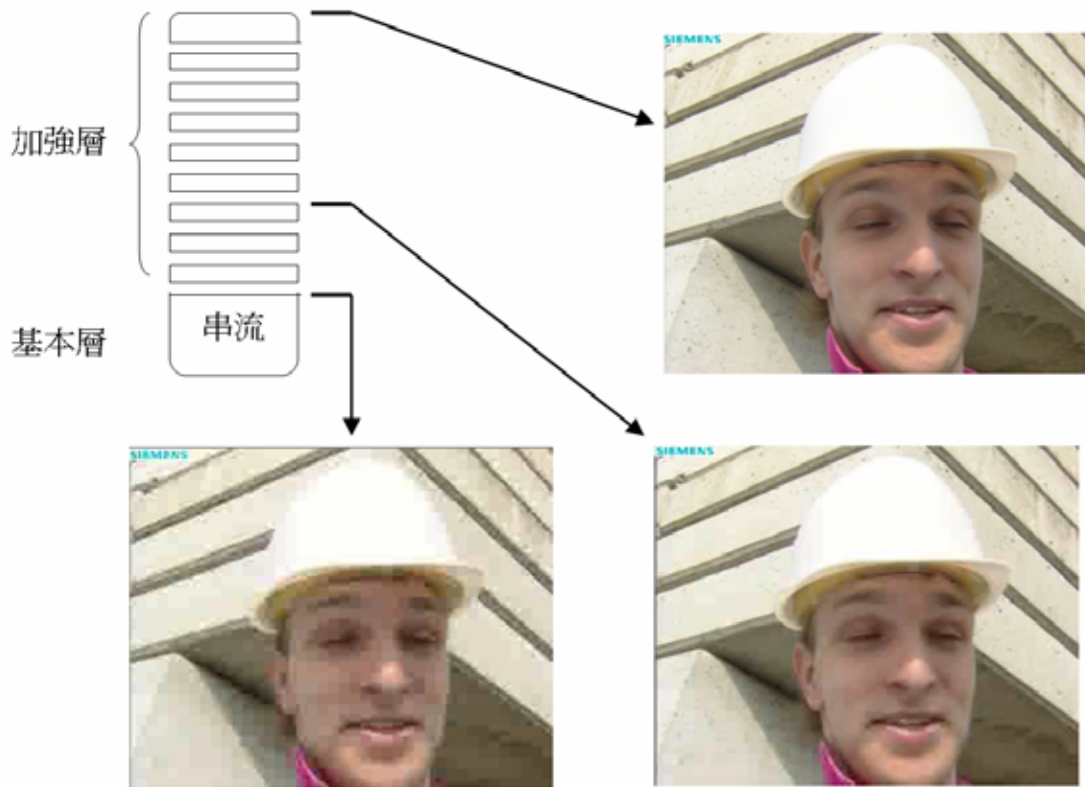


圖 2.9 雜訊比可調層次式

2.2.6.2 時間解析度可調層次式(Temporal Scalability)

一般視訊的時間取樣頻率為每秒 24~30 張，而時間解析度可調層次式即是調整時間取樣頻率。例如：視訊資料編碼成爲兩個串流(stream)，只接收第一個串流可解碼還原出每秒 15 張畫面，如再接收第二個串流就可以還原出每秒 30 張畫面，而從 15 張畫面到 30 張畫面之間，也可以存在許多等級的時間解析度，端看編碼規格支援與否。



圖 2.10 時間解析度可調層次式

2.2.6.3 空間解析度可調層次式(Spatial Scalability)

空間解析度可調層次式所調整的，是視訊檔案的解析度。例如：一個 CIF 格式的視訊檔案，其畫面解析度為 352 X 288，經過具有空間解析度可調層次式能力的編碼器編碼，形成兩個串流，只接收第一個串流可解碼還原出 QCIF 格式的視訊(176X144)，再接受第二個串流就可以還原出原來 CIF 格式的畫面。其可調層次式能力的精細程度，端看編碼規格支援與否。

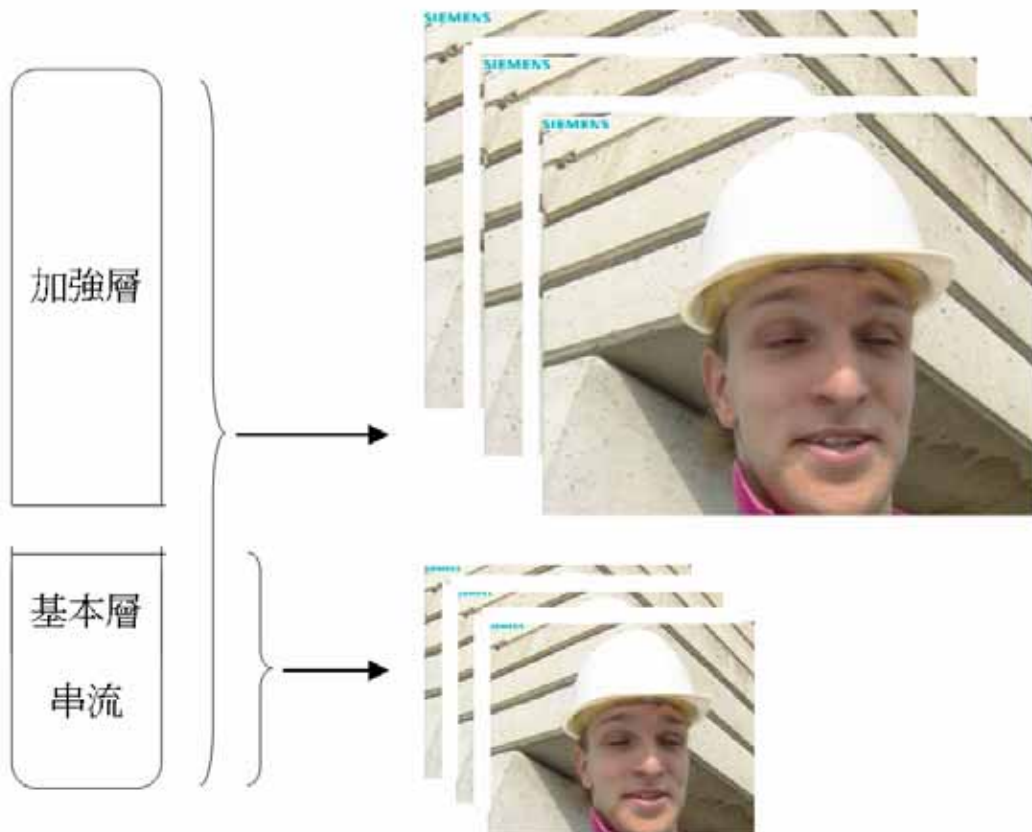


圖 2.11 空間解析度可調層次式

一般來說，時間解析度可調層次式與空間解析度可調層次式的可調層次式程度，不及雜訊比可調層次式來得精細(Fine-Granularity)。一方面是因為三種可調層次式在本質上的難易度不同，另一方面因為目前這兩種可調層次式所採用的方法，其可調層次式能力皆不及雜訊比可調層次式所採用的位元平面編碼。

2.3 金字塔型空間解析度可調層次式

(Pyramid Spatial Scalability)

金字塔型架構，在靜態影像處理中已經被提出，而金字塔型空間解析度可調層次式視訊編碼則是利用金字塔型架構，將連續的視訊畫面一張一張以金字塔型架構完成空間解析度可調層次式，再各自合成不同解析度的連續畫面，然後獨立編碼，形成獨立的串流。

2.3.1 金字塔型影像架構 (Pyramid Image Structure)

影像金字塔即是透過數層的縮減取樣(Down-Sampling)，形成解析度以 2 倍遞減的一連串影像，應用在影像處理與影像壓縮。而應用在影像壓縮時，較高解析度的影像，將會除去已經存在於較低解析度影像的冗餘(Redundancy)，方法是增升取樣(Up Sampling)低解析度影像，並減去之。

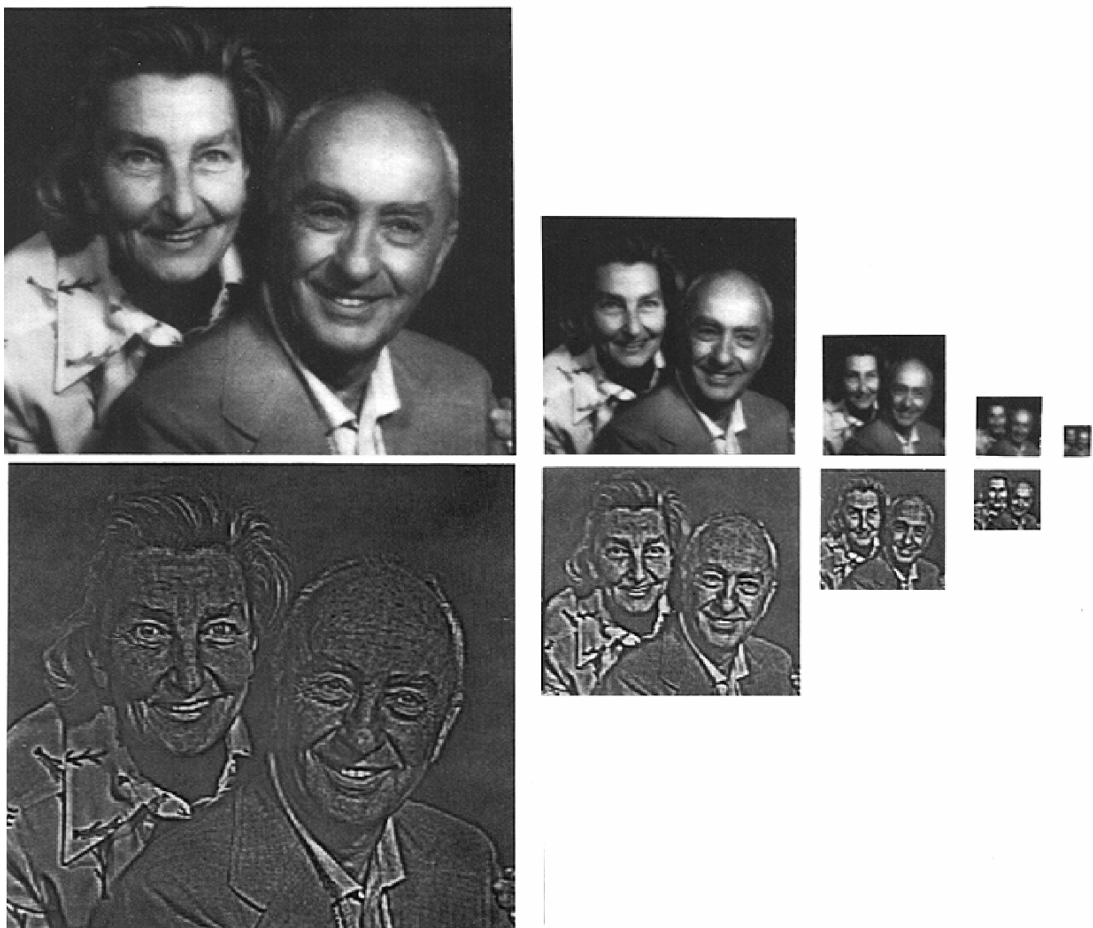


圖 2.12 金字塔影像群

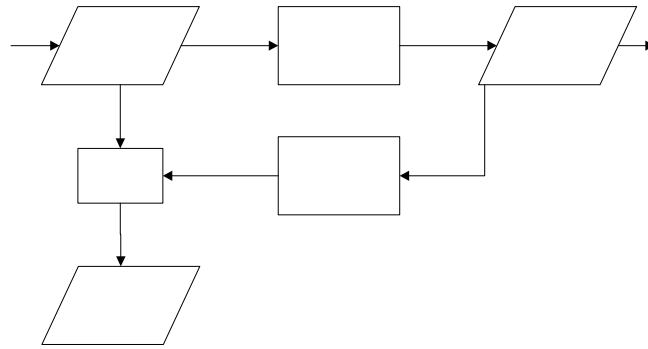


圖 2.13 形成金字塔影像架構之基本過程

2.3.2 縮減取樣 (Down-sampling)

最簡單的縮減取樣是單純的移除多餘的取樣點，但是會造成干擾失真(Aliasing)，所以一般會在移除多餘取樣點之前，會先做反失真濾波(Anti-aliasing filtering)。其步驟為：

反失真濾波 → 移除多餘取樣點

或是進一步簡化為

在保留的取樣點上做反失真濾波

對於數位影像來說，理想的反失真濾波很難定義，因為數位影像為有限長度(Finite Duration)，有限精確度(Finite Precision)，頻譜的定義難統一；即使選用特定的頻譜定義，在目前的運算複雜度要求下，幾乎無法達成完美的反失真濾波。一般採用的反失真濾波，是利用低通濾波遮罩(Low-pass Filter Mask)與影像作二維離散迴旋積分(Discrete Convolution)，而其濾波遮罩的大小，與濾波品質成正相關。

2.3.3 金字塔型空間解析度可調層次式

金字塔型空間解析度可調層次式，即是把金字塔型影像架構應用在連續影像上，再分別編碼，解碼還原後，再依照金字塔型影像架構重建。

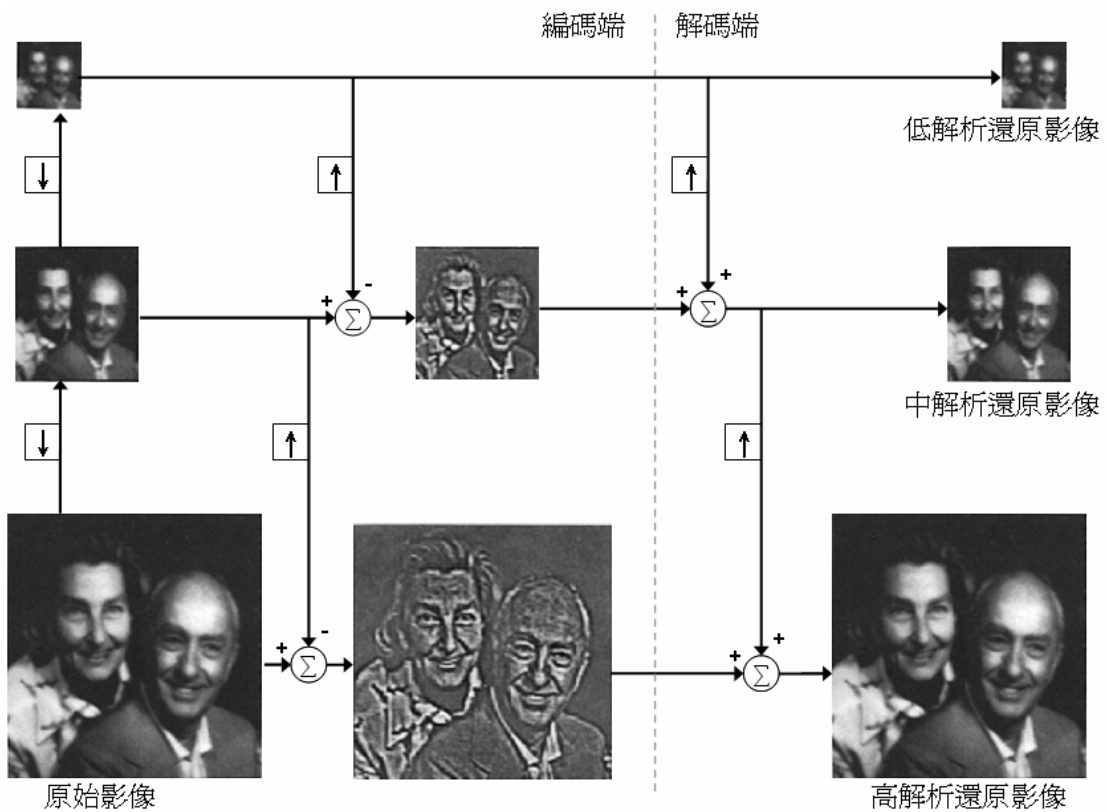


圖 2.14 金字塔空間解析度可調層次式

2.3.4 空間性資訊分離 (Spatial Partition)

我們以三層影像金字塔為例，分析空間性資訊如何在金字塔影像架構中分離。下圖中，中解析度與低解析度影像分別做了

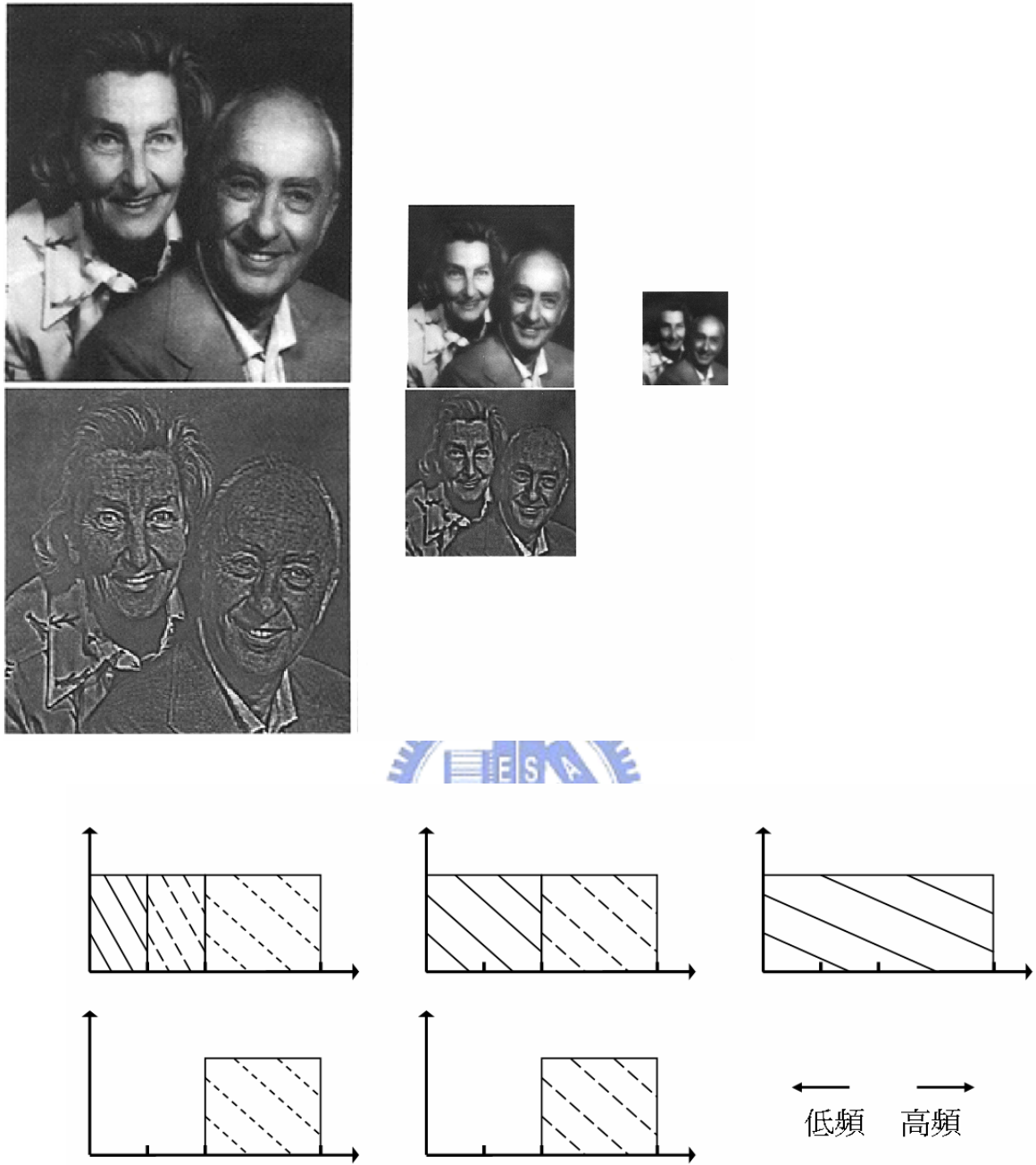


圖 2.15 金字塔影像結構頻譜對照之一

由於金字塔影像結構具有縮減取樣的步驟，而空間域的縮減取樣會致使頻率域的頻譜延展，所以不容易在頻譜上看出空間資訊的分離。於是我們把所有的影像調整(增升取樣)至相同的大小，這樣可以清楚的看出頻率資訊的分離，而增升取樣不會造成頻率資料的遺失，故不會影響我們的觀察。

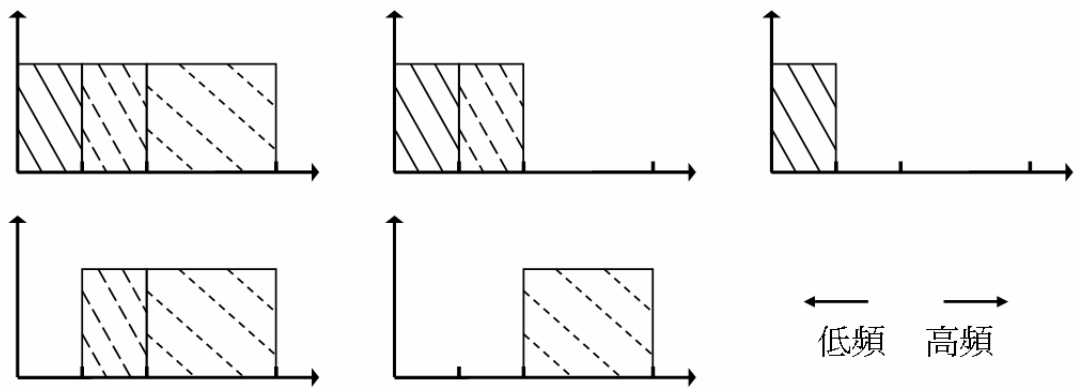
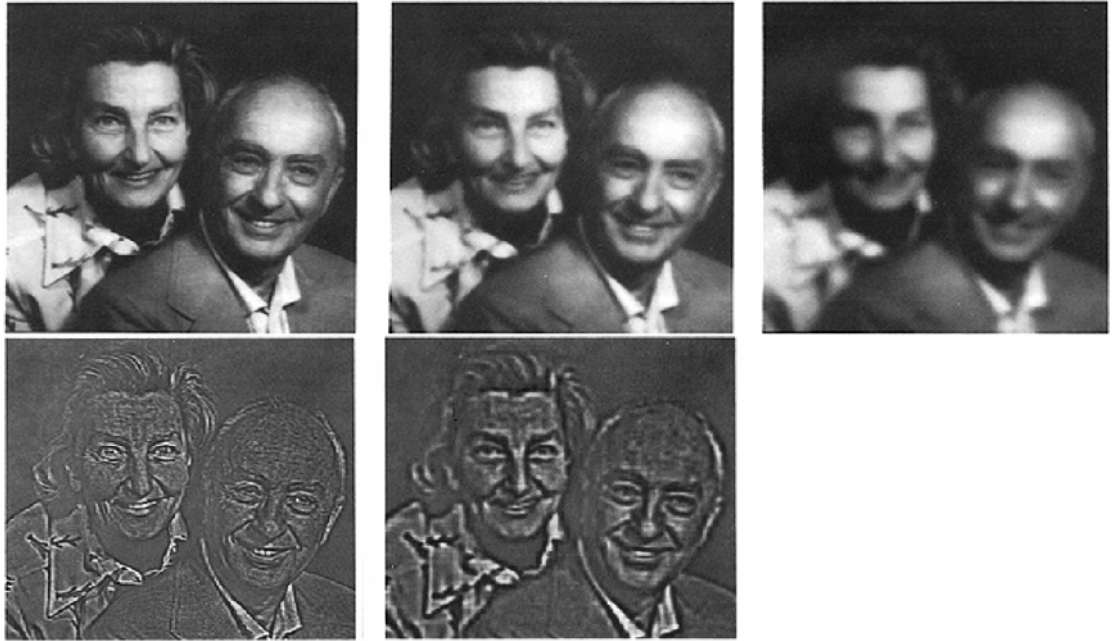


圖 2.16 金字塔影像結構頻譜對照之二

由上面的對應可發現，金字塔影像結構，本質上等於把頻率資訊分割。

2.4 移動補償(Motion Compensation)

空間解析度可調層次式編碼的移動補償必須多份，因為每個解析度之間必須有獨立的移動補償，否則會存在錯誤飄移(Error Drift)。以下我們分別詳述一致的移動補償與獨立的移動補償，更清楚的分析錯誤由何處產生。

2.4.1 一致的移動補償

由於最高解析階層具有最精確的移動向量，故一致的移動補償即是以最高解析階層的移動補償為主，再將移動補償的殘數(residue)分成三個階層的串流。

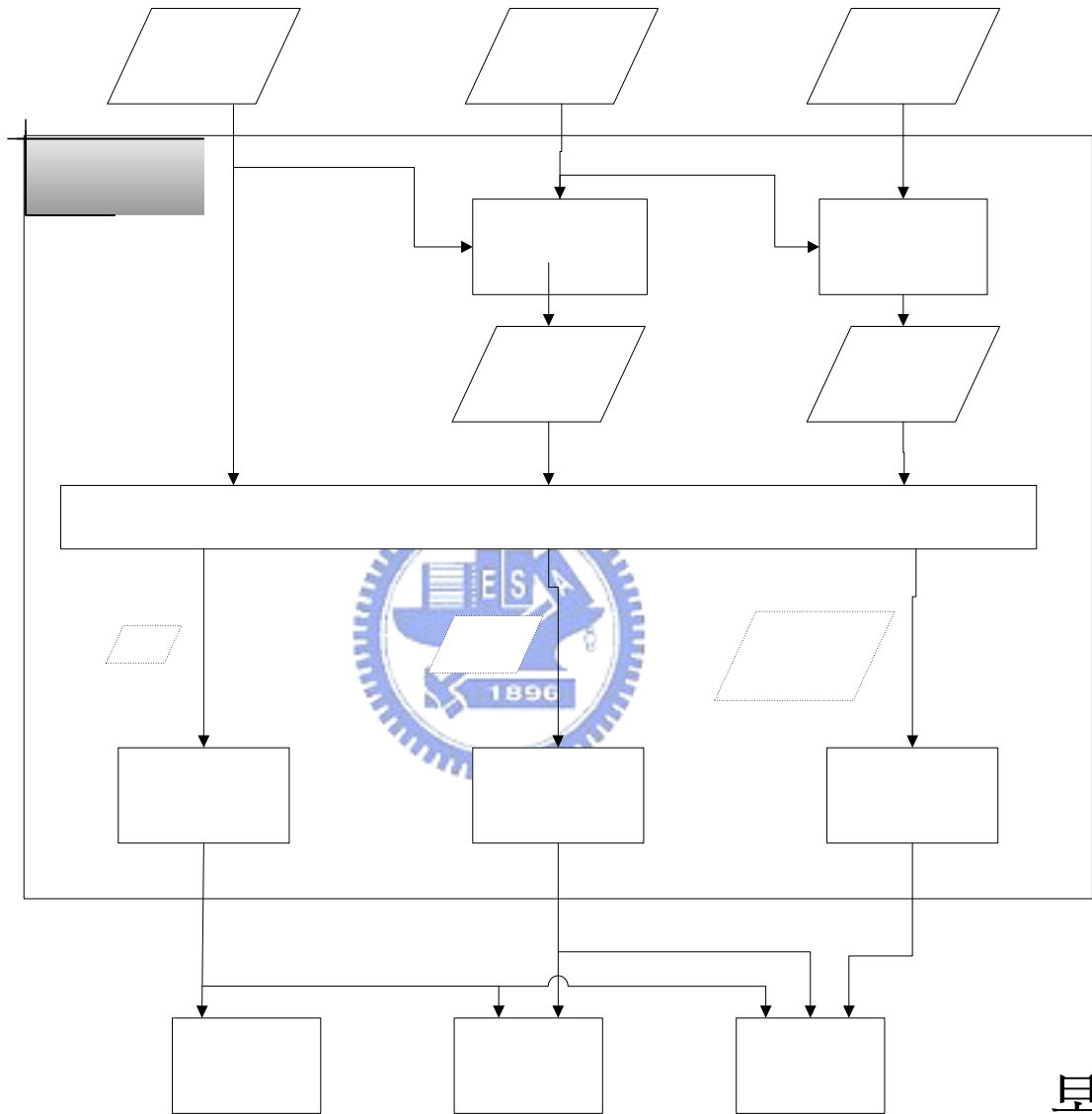


圖 2.17 一致的移動補償

影像 0

而在低解析解碼器與中解析解碼器中，移動補償將無法與編碼器完全匹配(Match)，原因有二：(以中解析解碼器為例)

- 解碼器畫面較小

中解析解碼器的畫面長寬是編碼器的一半，移動向量也減半，在此種條件下，無法做到與編碼器相匹配的移動補償。

- 移動補償為非線性運算

編碼器

空間解析度可調層次式是個線性的運算，但是移動補償則否。編碼器端做移動補償時，只做一次，勢必將低解析、中解析、高解析參考畫面一併做補償，也就是相減的動作。相減的動作原屬於線性運算，但加上移動向量的位移後，將變成非線性運算。編碼器編碼過程中參考了三個解析度的資訊，而且經過非線性運算；而在中解析解碼器中，不會有高解析度資訊的加入，這樣將會影響到中解析解碼器對於低、中解析度資訊的還原，而造成錯誤，因為移動補償不匹配造成的錯誤。

2.4.2 獨立的移動補償

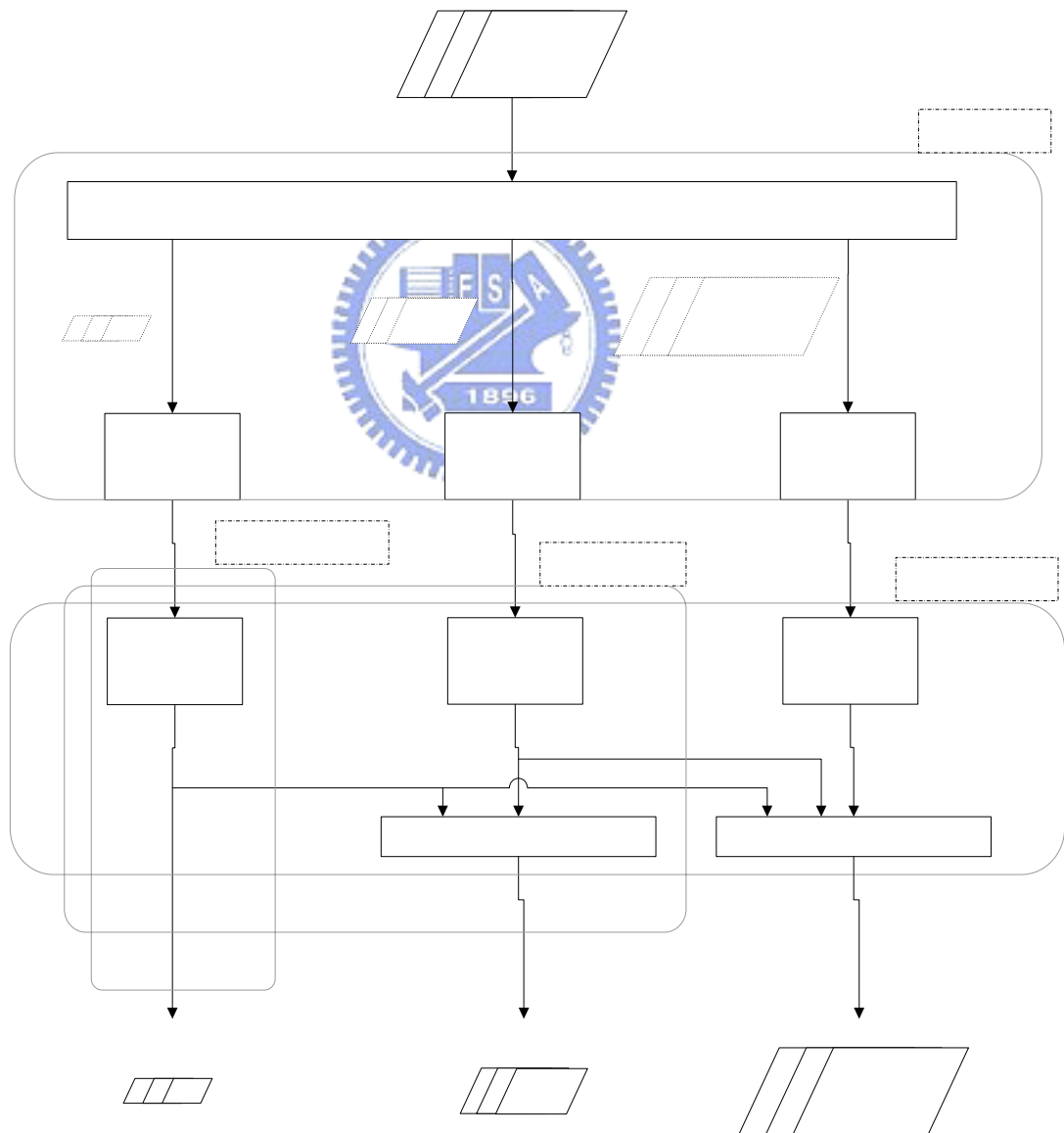


圖 2.18 獨立的移動補償

在此架構中，每個解析度的資訊有獨立的移動補償，空間資訊的分離與組合在前後包著編碼與解碼，故空間資訊的分離與組合維持本身的線性特性；換句話說，每個解析度的編碼與解碼之間，沒有空間資訊的分離，故兩者的移動補償能完全匹配，如同不具有空間可調層次式的編碼器與解碼器。而每個解析度之間的移動補償雖然獨立，但是低解析度移動向量可作為高解析度移動向量的預測，進而減低資料量。

2.5 其他空間解析度可調層次式編碼

空間解析度可調層次式視訊編碼另一個常見的方法是小波轉換。事實上小波轉換與離散餘弦轉換是影像編碼的兩條主流，離散餘弦轉換發展較悠久，相關研究相當成熟，小波轉換發表較慢，但表現不俗，例如 JPEG2000 採用小波轉換，而達到了目前最佳的影像壓縮品質；而目前視訊編碼規格幾乎都採用離散餘弦轉換來壓縮資料，小波轉換在視訊編碼的地位比起 DCT 影像編碼稍低，但仍有大量相關研究。而由於小波轉換良好的分頻特性，小波轉換也可以達成空間解析度可調層次式。

2.5.1 小波轉換空間解析度可調層次式編碼

小波轉換視訊編碼的原始演算法來自於預測性嵌入式零樹小波轉換(Predictive embedded Zerotree Wavelet - PEZW)，其步驟包括：執行小波轉換、建立小波零樹、量化以及將小波係數進行熵編碼(Entropy Coding)。在此我們不詳述每個步驟的細節，只稍微介紹小波轉換的概念

2.5.1.1 離散小波轉換

小波轉換的特色歸納如下：

- 出發點：
達到兼顧資料的頻率性與空間性的頻域轉換。
- 要求：
完美匹配(Perfect match)的分頻與組合，一般還會要求分頻的過程中能縮減取樣，減低資料量。
- 做法：
利用小波函數(Wavelet function)與縮放函數(Scaling function)與輸入做迴旋積分得到分頻後的資料；反轉換則是利用對應的函數進行完美重建。

- 優點：
 1. 兼顧資料的頻率性與區域性，不會有區塊效應。
 2. 資料量較金字塔影像結構低。

轉換結果如下圖：



圖 2.19 一次小波轉換

分頻後的訊號，可以再做進一步的小波轉換。一般會針對低低部分做多層的小波轉換，以達到空間解析度可調層次式。下圖的影像做了兩層小波轉換：

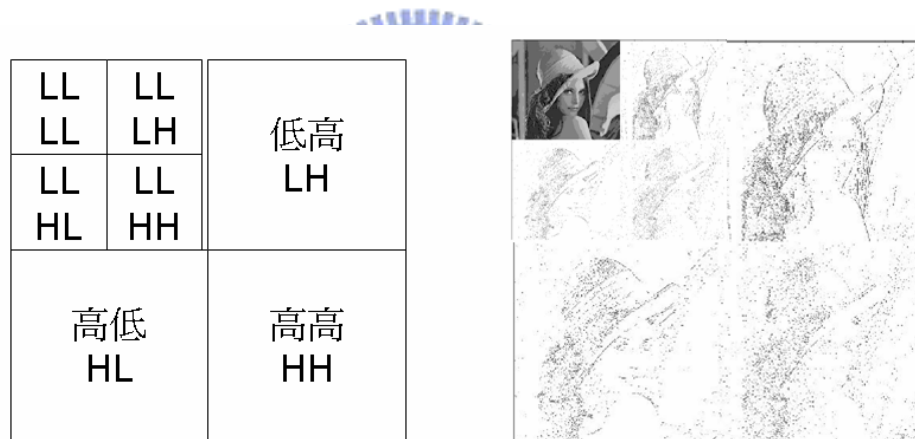


圖 2.20 二次小波轉換

2.5.1.2 小波轉換空間解析度可調層次式編碼

小波轉換編碼應用在空間解析度可調層次式非常適合，如下圖所示：

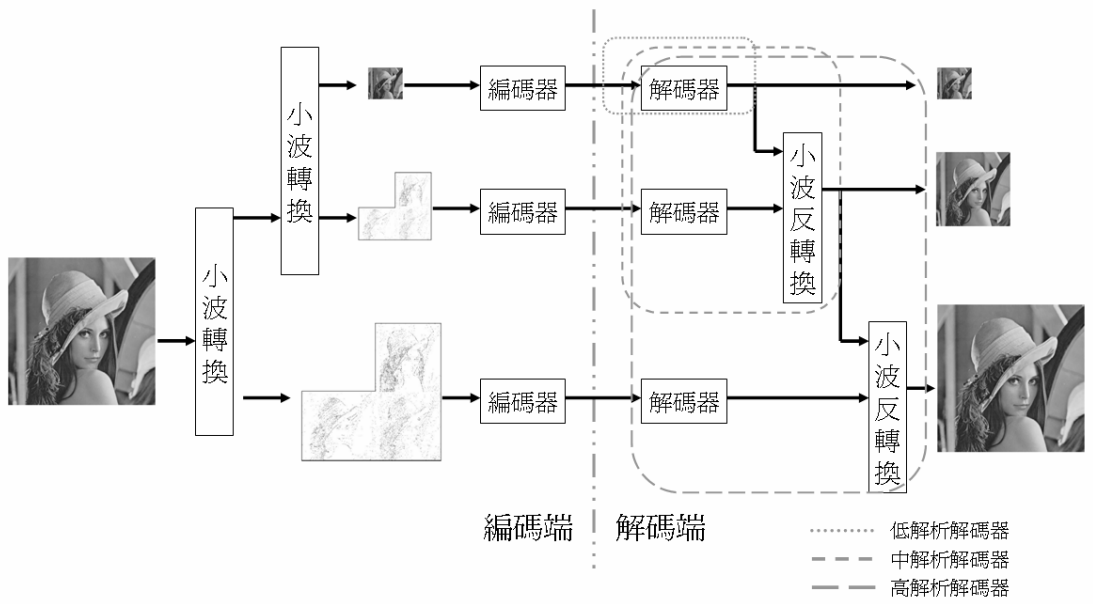


圖 2.21 小波轉換空間解析度可調層次式

