



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I482501 B

(45) 公告日：中華民國 104 (2015) 年 04 月 21 日

(21) 申請案號：101100464

(22) 申請日：中華民國 101 (2012) 年 01 月 05 日

(51) Int. Cl. :

*H04N19/513 (2014.01)**H04N19/517 (2014.01)*

(71) 申請人：安霸股份有限公司 (中華民國) AMBARELLA TAIWAN LTD. (TW)

新竹市科學園區力行一路 1 號 1 樓 C1

國立交通大學 (中華民國) NATIONAL CHIAO TUNG UNIVERSITY (TW)

新竹市大學路 1001 號

(72) 發明人：彭文孝 PENG, WEN HSIAO (TW)；蔣迪豪 CHIANG, TIHAO (TW)；蔡閏旭 TSAI, RUEN SHIU (TW)

(74) 代理人：蔡坤旺

(56) 參考文獻：

US 6798833B2

US 7613345B2

US 2005/0185621A1

US 2009/0129474A1

US 2011/0170592A1

Liquan Shen, Zhi Liu, Tao Yan, Zhaoyang Zhang, and Ping An, "View-Adaptive Motion Estimation and Disparity Estimation for Low Complexity Multiview Video Coding", IEEE Transaction on Circuits and Systems for Video Technology, Vol 20, No. 6, 2010.

審查人員：謝瑞航

申請專利範圍項數：6 項 圖式數：9 共 22 頁

(54) 名稱

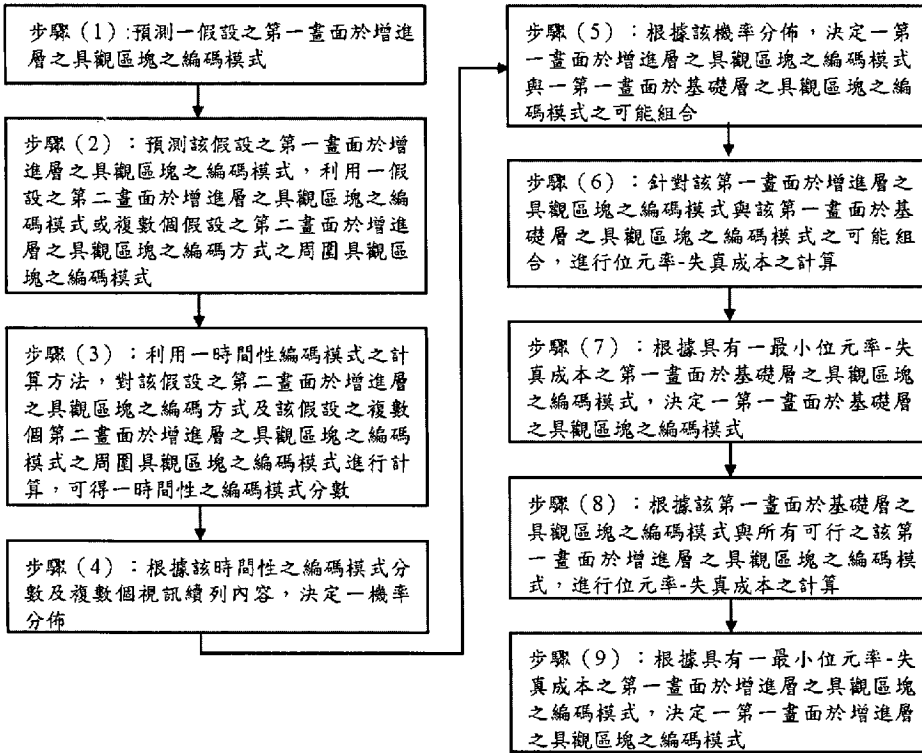
可調性視訊多層編碼控制快速決策模式演算法

FAST MODE DECISION METHOD FOR SCALABLE VIDEO CODING MULTI-LAYER ENCODER CONTROL

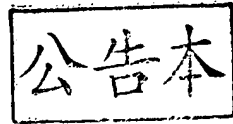
(57) 摘要

本發明揭示一種可調性視訊多層編碼控制快速決策模式演算法。該方法係使用時間性編碼模式分數(Temporal Mode Score, TMS)之概念，可以準確的預測出增進層的編碼模式，並且找出基礎層和增進層的編碼模式並可隨視訊序列的性質做不同的配對計算。

The present invention discloses a fast mode decision method for scalable video coding multi-layer encoder control. The method uses a concept of temporal mode score to precisely a coding model of the enhanced layer, to fine the coding model of the base layer and the enhancement layer, and to process corresponding calculation according to the properties of the scalable video coding.



第 4 圖



# 發明專利說明書

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：101100464

※ 申請日：101. 1. 05

※IPC 分類：

H04N 19/513 2014.01

H04N 19/517 2014.01

一、發明名稱：(中文/英文)

可調性視訊多層編碼控制快速決策模式演算法/Fast Mode

Decision Method for Scalable Video Coding Multi-Layer Encoder

Control

二、中文發明摘要：

本發明揭示一種可調性視訊多層編碼控制快速決策模式演算法。該方法係使用時間性編碼模式分數(Temporal Mode Score, TMS)之概念，可以準確的預測出增進層的編碼模式，並且找出基礎層和增進層的編碼模式並可隨視訊序列的性質做不同的配對計算。

三、英文發明摘要：

The present invention discloses a fast mode decision method for scalable video coding multi-layer encoder control. The method uses a concept of temporal mode score to precisely a coding model of the enhanced layer, to fine the coding model of the base layer and the enhancement layer, and to process corresponding calculation according to the properties of the scalable video coding.

四、指定代表圖：

(一)本案指定代表圖為：第 ( 4 ) 圖。

(二)本代表圖之元件符號簡單說明：

482501-1A

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

## 六、發明說明：

### 【發明所屬之技術領域】

本發明係有關於一種視訊編碼演算法，更特別指一種可調性視訊多層編碼控制快速決策模式演算法。藉由使用時間性編碼模式分數(Temporal Mode Score, 簡稱 TMS)之概念，以更準確的預測出增進層的編碼模式，並且準確的找出基礎層和增進層的編碼模式並可隨視訊序列的性質做不同的配對計算，具有較低編碼失真、減少編碼時間、可調性及具較佳位元率-失真效能(Rate-Distortion performance)之編碼模式等優點。

### 【先前技術】

由於多變容量的不同設備以及網路上的視訊撥放服務越來越多，單一視訊編碼已不足以應付這些項目。基於這個理由，Joint video Team (JVT)以 H.264/AVC 為基礎，制定了一套新的標準-可調性視訊編碼(Scalable Video Coding, 簡稱 SVC)，其提供了單一位元串流擁有在空間性、時間性、品質性以及三種之間組合的可調性，利用這種可調性性質對單一位元串流去做編碼，可得到不同解析度以及不同品質的服務。

可調性視訊編碼擁有單一基礎層和一個或多個增進層，其增進層盡可能的利用基礎層的資訊，且一次編碼多層可以達到減少位元率並且也可達到預期中的畫面品質。傳統的可調性視訊編碼在決定基礎層和增進層的編碼模式時，共用式可調性視訊編碼

(Joint Scalable video Model, 簡稱 JSVM)是採取由下而上的編碼方法(Bottom up encoder control)。在這種編碼方法中，只考慮到兩層之間較低層的效能。故傳統的編碼方法，將會造成兩層之間較高層出現較嚴重的失真，造成在選擇基礎層(Base Layer, 簡稱 BL)的編碼模式時，完全沒有考慮到增進層(Enhancement Layer, 簡稱 EL)的編碼效能，因此會導致增進層有不可預期的失真錯誤產生。

Schwarz et al.在早期提供了一種可以在基礎層和增進層之間給予適當權重的方法，稱做多層式編碼控制(Multilayer Encoder Control)。在做編碼模式的決策時，基礎層使用一種權重式的Lagrangian 成本決策方法。如圖 1，在多層式編碼模式決策時，其基本單位是一對具觀區塊 (Macroblock, MB)組合，分別來自基礎層和增進層相對位置的具觀區塊。對於兩層之間的編碼模式，其配對組合一共有 56 組，若計算這 56 組的編碼模式，並在其中找出一組最好的，比起傳統由下而上的編碼方法，明顯有很大的計算量，由此可見一個好的快速決策模式方法顯得非常重要。

Schwarz et al.又提出了一個兩段式快速編碼模式決策演算法。其做法是當基礎層的具觀區塊做編碼模式決策時，每一個編碼模式都會有一個增進層的編碼模式做為配對，此種配對稱為主導性配對(Dominant Mode Pair)。如圖 2 所示，其配對方法是當基礎層選擇 畫面間預測(Inter Prediction)或畫面內預測(Intra Prediction)的編碼模式時，增進層對應選擇 BL\_skip 或 IntraBL 做為預測，然後基礎層利用這組配對去計算位元率-失真效能成本

(Rate-Distortion-cost, 簡稱 RD-cost), 並選擇 RD-cost 最小的編碼模式做為基礎層的編碼模式, 最後在增進層中對應的具觀區塊再根據選定好的基礎層編碼模式去做完整搜尋, 找出最小的 RD-cost 編碼模式做壓縮。但此方法存在一些嚴重的問題:

(1)在選擇基礎層中每一個具觀區塊的編碼模式時, 都是使用圖 2 中的 7 組配對組合去做計算, 並不會依照視訊序列的特性做調整。這種配對方式太過於單調, 不具任何可調性。

(2)此配對模式因為只是 56 組配對模式中的 7 組, 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估增進層的編碼模式, 可能與實際增進層的編碼模式相差太大, 因此會導致基礎層在做編碼模式的選擇時出現錯誤, 若再用此錯誤的基礎層編碼模式去找增進層中的編碼模式, 將會造成不可預期的嚴重失真。

鑒於先前技術之缺點, 有必要提出一種可調整性影像處理方法, 藉由時間性之編碼模式分數之使用, 可以減少其編碼時所需之運算量並可解決上述所提及之問題。

#### 【發明內容】

鑒於上述先前技術之缺點, 本發明之主要目的即在於提供一種可調性視訊多層編碼控制快速決策模式演算法, 可以判斷小波轉換的次頻帶是否適合方向性拆解的方法, 再根據所有方向次頻帶上的訊號特性, 提出新的相鄰表, 來達成更佳的壓縮效果。

為達上述目的，本發明提供一種可調性視訊多層編碼控制快速決策模式演算法，用於一影像訊號之壓縮與編碼，該方法至少包含下列步驟：步驟（1）：預測一假設之第一畫面於增進層之具觀區塊之編碼模式；步驟（2）：預測該假設之第一畫面於增進層之具觀區塊之編碼模式，利用一假設之第二畫面於增進層之具觀區塊之編碼模式或複數個假設之第二畫面於增進層之具觀區塊之編碼方式之周圍具觀區塊之編碼模式；步驟（3）：利用一時間性編碼模式之計算方法，對該假設之第二畫面於增進層之具觀區塊之編碼方式及該假設之複數個第二畫面於增進層之具觀區塊之編碼模式之周圍具觀區塊之編碼模式進行計算，可得一時間性之編碼模式分數；步驟（4）：根據該時間性之編碼模式分數及複數個視訊續列內容，決定一機率分佈；步驟（5）：根據該機率分佈，決定一第一畫面於增進層之具觀區塊之編碼模式與一第一畫面於基礎層之具觀區塊之編碼模式之可能組合；步驟（6）：針對該第一畫面於增進層之具觀區塊之編碼模式與該第一畫面於基礎層之具觀區塊之編碼模式之可能組合，進行位元率-失真成本之計算；步驟（7）：根據具有一最小位元率-失真成本之第一畫面於基礎層之具觀區塊之編碼模式，決定一第一畫面於基礎層之具觀區塊之編碼模式；步驟（8）：根據該第一畫面於基礎層之具觀區塊之編碼模式與所有可行之該第一畫面於增進層之具觀區塊之編碼模式，進行位元率-失真成本之計算；步驟（9）：根據具有一最小位元率-失真成本之第一畫面於增進層之具觀區塊之編碼模



式，決定一第一畫面於增進層之具觀區塊之編碼模式。

綜上所述，本發明之一種可調性視訊多層編碼控制快速決策模式演算法將具有以下之功效：

1. 利用時間(Temporal)的性質去預測增進層中的編碼模式，再將這個預測和基礎層的編碼模式做配對。先去預測增進層的編碼模式，因此沒有固定的 Dominant Mode Pair 概念，在選擇基礎層的編碼模式時，不同的具觀區塊可能會搭配不同的增進層編碼模式做計算，可減少編碼時間。
2. 提供時間性編碼模式分數(Temporal Mode Score, 簡稱 TMS) 的概念，TMS 可以更準確的預測出增進層的編碼模式，並且準確的找出基礎層和增進層的編碼模式。
3. 根據視訊序列的性質做調整，每一個具觀區塊都先去增進層預測出機率最大的編碼模式做為基礎層的配對，因此具有可調性。

為讓本發明之上述和其他目的、特徵、和優點能更明顯易懂，下文特舉數個較佳實施例，並配合所附圖式，作詳細說明如下。

### 【實施方式】

雖然本發明可表現為不同形式之實施例，但附圖所示者及於下文中說明者係為本發明可之較佳實施例，並請了解本文所揭示者係考量為本發明之一範例，且並非意圖用以將本發明限制於圖示及/或所描述之特定實施例中。

現請參考第 3 圖，其顯示本發明之快速決策模式方法之架構示意圖。圖中的 A 是基礎層中目前畫面的具觀區塊、B 是增進層中相對於 A 位置的具觀區塊而 C 是在增進層中的前一張畫面相對於 B 的九個具觀區塊位置。

現請搭配參考第 4 圖，其顯示本發明之可調性視訊多層編碼控制快速決策模式演算法之步驟流程圖，主要包含九個步驟。

首先於步驟(1)中，其為預測一假設之第一畫面於增進層之具觀區塊之編碼模式，主要由於在尋找基礎層之具觀區塊之編碼模式時，必須先預測於增進層之具觀區塊之編碼模式。

於步驟(2)中，預測該假設之第一畫面於增進層之具觀區塊之編碼模式，其利用一假設之第二畫面於增進層之具觀區塊之編碼模式或複數個假設之第二畫面於增進層之具觀區塊之編碼方式之周圍具觀區塊之編碼模式。

於步驟(3)中，其利用一時間性編碼模式之計算方法，對該假設之第二畫面於增進層之具觀區塊之編碼方式及該假設之複數個第二畫面於增進層之具觀區塊之編碼模式之周圍具觀區塊之編碼模式進行計算，可得一時間性之編碼模式分數(Temporal Mode Score, 簡稱 TMS)。

現請參考第 5 圖，其顯示本發明之增進層(EL)中之

一個具觀區塊於第二畫面之所對應九格位置所選擇的編碼模式，該圖顯示第一畫面基礎層的一個具觀區塊在第一畫面增進層所對應到的具觀區塊位置為 E。亦即，若要決策基礎層一個具觀區塊的編碼模式時，需要先找出此該 E 位置的具觀區塊所預測之編碼模式，作為第一畫面基礎層的一個配對。此外，第一畫面 E 位置假設之編碼模式為利用第二畫面於增進層之具觀區塊之編碼方式及該假設之複數個第二畫面於增進層之具觀區塊之編碼模式之周圍具觀區塊之編碼模式。亦即，圖中 E 所假設之編碼模式就是利用圖中 A'~I' 位置之編碼模式，搭配圖 6(a) 之具觀區塊相對位置之時間性編碼模式之計算方法所計算出來，進而可得到圖 6(b) 之各種編碼模式之時間性之編碼模式分數示意圖。

需注意的是，其中該時間性編碼模式之計算方法為  $TMS = A' \times 1 + B' \times 2 + C' \times 1 + D' \times 2 + E' \times 4 + F' \times 2 + G' \times 1 + H' \times 2 + I' \times 1$ ，此外，其 A'、B'、C'、D'、E'、F'、G'、H'、I' 為該假設之第二畫面於增進層之具觀區塊之編碼模式所對應九格位置所選擇的編碼模式。該時間性之編碼模式分數之範圍介於 0 至 64。若 A'、B'、C'、D'、E'、F'、G'、H'、I' 出現畫面內預測 (Intra Prediction) 之編碼模式，其值為 0。

於步驟 (4) 中，其根據該時間性之編碼模式分數及複數個視訊續列內容，決定一機率分佈。該複數個視訊

續列內容係包含 Akiyo、Carphone、City、Coastguard、Highway、Stefan 及 Walk 視訊續列。

於步驟 (5) 中，其根據該機率分佈，決定一第一畫面於增進層之具觀區塊之編碼模式與一第一畫面於基礎層之具觀區塊之編碼模式之可能組合。

於步驟 (6) 中，其針對該第一畫面於增進層之具觀區塊之編碼模式與該第一畫面於基礎層之具觀區塊之編碼模式之可能組合，進行位元率-失真成本之計算。

於步驟 (7) 中，其根據具有一最小位元率-失真成本之第一畫面於基礎層之具觀區塊之編碼模式，決定一第一畫面於基礎層之具觀區塊之編碼模式。

於步驟 (8) 中，其根據該第一畫面於基礎層之具觀區塊之編碼模式與所有可行之該第一畫面於增進層之具觀區塊之編碼模式，進行位元率-失真成本之計算。

於步驟 (9) 中，其根據具有一最小位元率-失真成本之第一畫面於增進層之具觀區塊之編碼模式，決定一第一畫面於增進層之具觀區塊之編碼模式。

為更了解本發明之精神，此處以一第一實施例進行快速決策模式方法之說明。若假設之第一畫面於增進層之具觀區塊之編碼模式為 Inter 8x8，請同時參考第 6(a)圖及第 7 圖，並計算出假設之第二畫面於增進層之具觀區塊之編碼方式及該假設之複數個第二畫面於增進層之具觀

區塊之編碼模式之周圍具觀區塊之編碼模式之 TMS，計算方式如下：

$$TMS = 0 \times 1 + 1 \times 2 + 1 \times 1 + 1 \times 2 + 4 \times 4 + 2 \times 2 + 2 \times 1 + 2 \times 2 + 1 \times 1 = 32$$

接著，請參考第 8 圖依據  $TMS = 32$  所對應的機率分佈，可找出機率最高的編碼模式做為第一畫面於增進層之具觀區塊之編碼模式之可能組合。此處則以 BL\_skip 做為第一畫面於增進層之具觀區塊之編碼模式之可能組合。

現請參考第 9 圖，根據第一畫面於增進層之具觀區塊之編碼模式之可能組合為 BL\_skip，如圖中實線框所示。接著，針對實線框中之各組可能之編碼模式進行位元率-失真成本之計算，並以具有最小之位元率-失真成本做為第一畫面於基礎層之具觀區塊之編碼模式。

請再次參考第 9 圖，根據第一畫面於基礎層之具觀區塊之編碼模式進行位元率-失真成本之計算，此處以  $16 \times 16$  做為第一畫面於基礎層之具觀區塊之編碼模式，如虛線框所示，針對虛線框中之各組可能之編碼模式進行位元率-失真成本之計算，並以具有最小之位元率-失真成本做為第一畫面於增進層之具觀區塊之編碼模式。

現請再次參考第 7 圖，現請參考圖中括號之部分，亦即出現畫面內預測的情況下，其 TMS 的計算方式如下：

$$TMS = 0 \times 1 + 1 \times 2 + 1 \times 1 + 0 \times 2 + 4 \times 4 + 2 \times 2 + 2 \times 1 + 0 \times 2 + 0 \times 1 = 25$$

由於有畫面內預測的編碼模式存在，故須進行正規化的動作。其方法為：

$Round[(25/11) \times 16] = 36$ ，其中 *Round* 為四捨五入之運算，除以 11 是因其於九格的位置中，畫面間預測編碼模式位置的分數加總為 11。故此處以  $TMS=36$  所對應的機率分布，找出機率最高的編碼模式做為基礎層的配對。

綜上所述，本發明之一種可調性視訊編碼之多層編碼控制之快速決策模式方將具有以下之功效：

1. 利用時間(Temporal)的性質去預測增進層中的編碼模式，再將這個預測和基礎層的編碼模式做配對。先去預測增進層的編碼模式，因此沒有固定的 Dominant Mode Pair 概念，在選擇基礎層的編碼模式時，不同的具觀區塊可能會搭配不同的增進層編碼模式做計算，可減少編碼時間。
2. 提供時間性編碼模式分數(Temporal Mode Score ,TMS)的概念，TMS 可以更準確的預測出增進層的編碼模式，並且準確的找出基礎層和增進層的編碼模式。
3. 根據視訊序列的性質做調整，每一個具觀區塊都先去增進層預測出機率最大的編碼模式做為基礎層的配對，因此具有可調性。

雖然本發明已以前述較佳實施例揭示，然其並非用以限定本發明，任何熟習此技藝者，在不脫離本發明之精神和範圍內，當可作各種之更動與修改。如上

述的解釋，都可以作各型式的修正與變化，而不會破壞此創作的精神。因此本發明之保護範圍當視後附之申請專利範圍所界定者為準。

### 【圖式簡單說明】

第 1 圖係為習之技術之基礎層 (BL) 及增進層 (EL) 之編碼模式；

第 2 圖係為習之技術之由 Schwarz et al. 提出之演算法編碼模式之配對規則；

第 3 圖係為本發明之可調性視訊多層編碼控制快速決策模式演算法之架構示意圖；

第 4 圖係為本發明之快速決策模式方法之步驟流程圖；

第 5 圖係為本發明之增進層 (EL) 中之一個具觀區塊於第二畫面之所對應九格位置所選擇的編碼模式；

第 6 圖係為本發明之 (a) 具觀區塊相對位置之時間性編碼模式之計算分數及 (b) 各種編碼模式的時間性之編碼模式分數示意圖；

第 7 圖係為本發明之第一實施例之九格對應編碼模式；

第 8 圖係為本發明之第一實施例之不同時間性之編碼模式分數對應複數個視訊續列之機率分佈圖；以及

第 9 圖係為本發明之第一實施例之基礎層 (BL) 及增進層 (EL) 之編碼模式之組合示意圖。

**【主要元件符號說明】**



## 七、申請專利範圍：

1. 一種可調性視訊多層編碼控制快速決策模式演算法，用於一影像訊號之壓縮與編碼，該方法至少包含下列步驟：

步驟（1）：預測一假設之第一畫面於增進層之具觀區塊之編碼模式；

步驟（2）：利用一假設之第二畫面於增進層之具觀區塊之編碼模式或複數個假設之第二畫面於增進層之具觀區塊之編碼方式之周圍具觀區塊之編碼模式；

步驟（3）：利用一時間性編碼模式之計算方法，對該假設之第二畫面於增進層之具觀區塊之編碼方式及該假設之複數個第二畫面於增進層之具觀區塊之編碼模式之周圍具觀區塊之編碼模式進行計算，可得一時間性之編碼模式分數；

步驟（4）：根據該時間性之編碼模式分數及複數個視訊續列內容，決定一機率分佈；

步驟（5）：根據該機率分佈，決定一第一畫面於增進層之具觀區塊之編碼模式與一第一畫面於基礎層之具觀區塊之編碼模式之可能組合；

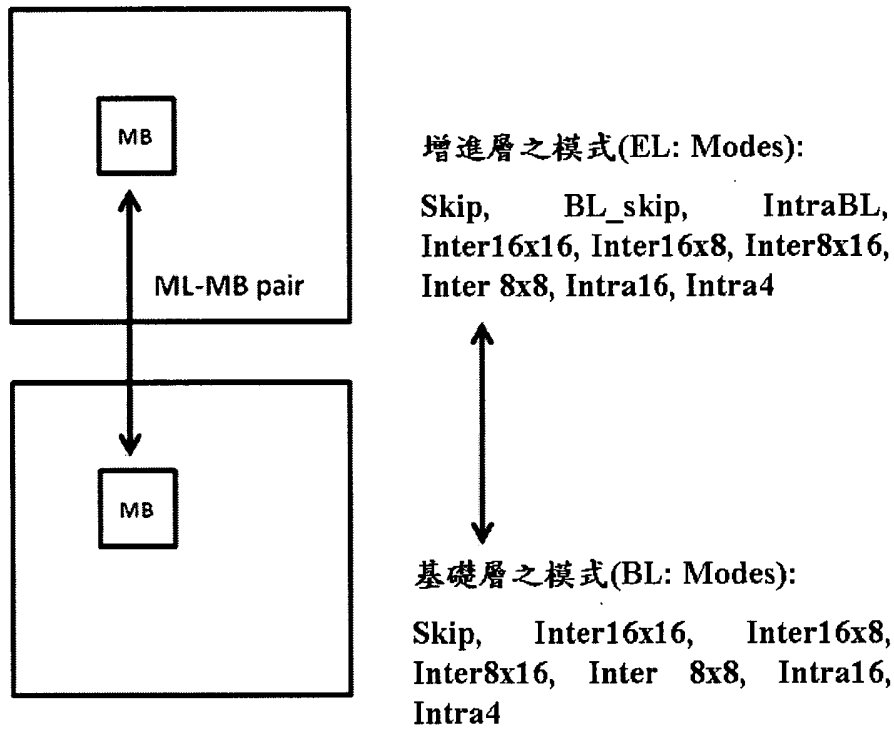
步驟（6）：針對該第一畫面於增進層之具觀區塊之編碼模式與該第一畫面於基礎層之具觀區塊之編

- 碼模式之可能組合，進行位元率-失真成本之計算；
- 步驟(7)：根據具有一最小位元率-失真成本之第一畫面於基礎層之具觀區塊之編碼模式，決定一第一畫面於基礎層之具觀區塊之編碼模式；
- 步驟(8)：根據該第一畫面於基礎層之具觀區塊之編碼模式與所有可行之該第一畫面於增進層之具觀區塊之編碼模式，進行位元率-失真成本之計算；
- 步驟(9)：根據具有一最小位元率-失真成本之第一畫面於增進層之具觀區塊之編碼模式，決定一第一畫面於增進層之具觀區塊之編碼模式。
2. 如申請專利範圍第1項之可調性視訊多層編碼控制快速決策模式演算法，其中該時間性編碼模式之計算方法為
- $$TMS = A' \times 1 + B' \times 2 + C' \times 1 + D' \times 2 + E' \times 4 + F' \times 2 + G' \times 1 + H' \times 2 + I' \times 1 ;$$
- 其中， $A'$ 、 $B'$ 、 $C'$ 、 $D'$ 、 $E'$ 、 $F'$ 、 $G'$ 、 $H'$ 、 $I'$ 為該假設之第二畫面於增進層之具觀區塊之編碼模式所對應九格位置所選擇的編碼模式。
3. 如申請專利範圍第2項之可調性視訊多層編碼控制快速決策模式演算法，其中步驟(1)之動作係執行於尋找基礎層之具觀區塊之編碼模式之前。
4. 如申請專利範圍第2項之可調性視訊多層編碼控制快速決策模式演算法，其中 $A'$ 、 $B'$ 、 $C'$ 、 $D'$ 、 $E'$ 、

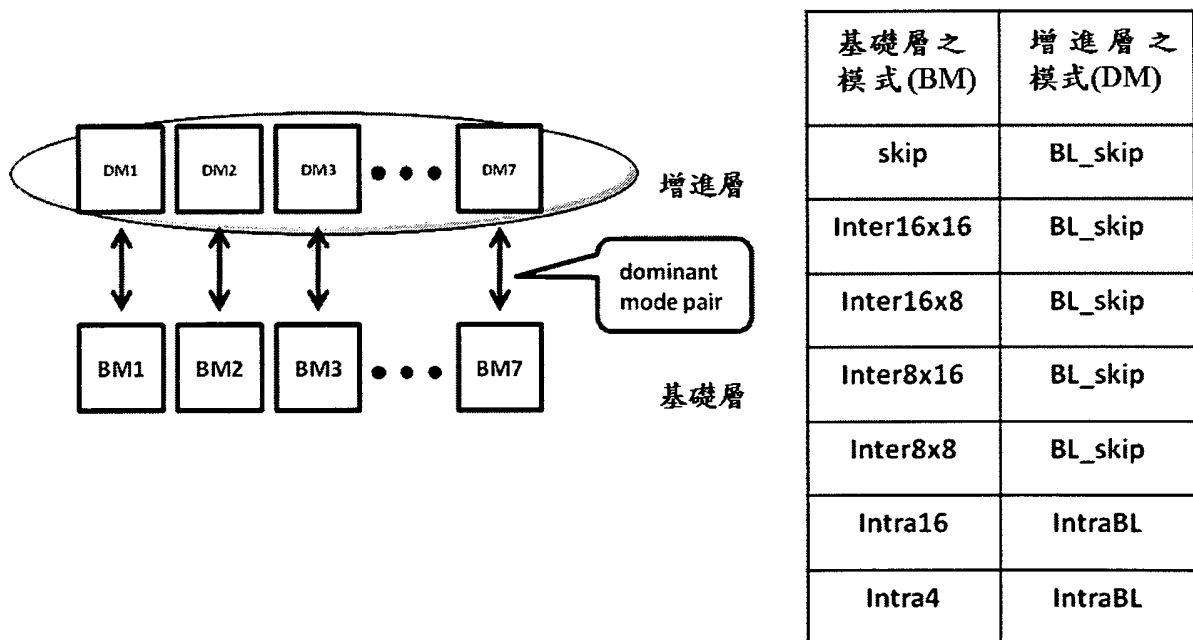
$F'$ 、 $G'$ 、 $H'$ 、 $I'$ 在畫面間預測 (Inter Prediction) 或畫面內預測 (Intra Prediction) 之編碼模式時，其值為 0。

5. 如申請專利範圍第 1 項之可調性視訊多層編碼控制快速決策模式演算法，其中該時間性之編碼模式分數之範圍介於 0 至 64。
6. 如申請專利範圍第 1 項之可調性視訊多層編碼控制快速決策模式演算法，其中該複數個視訊續列內容係包含 Akiyo、Carphone、City、Coastguard、Highway、Stefan 及 Walk 視訊續列。

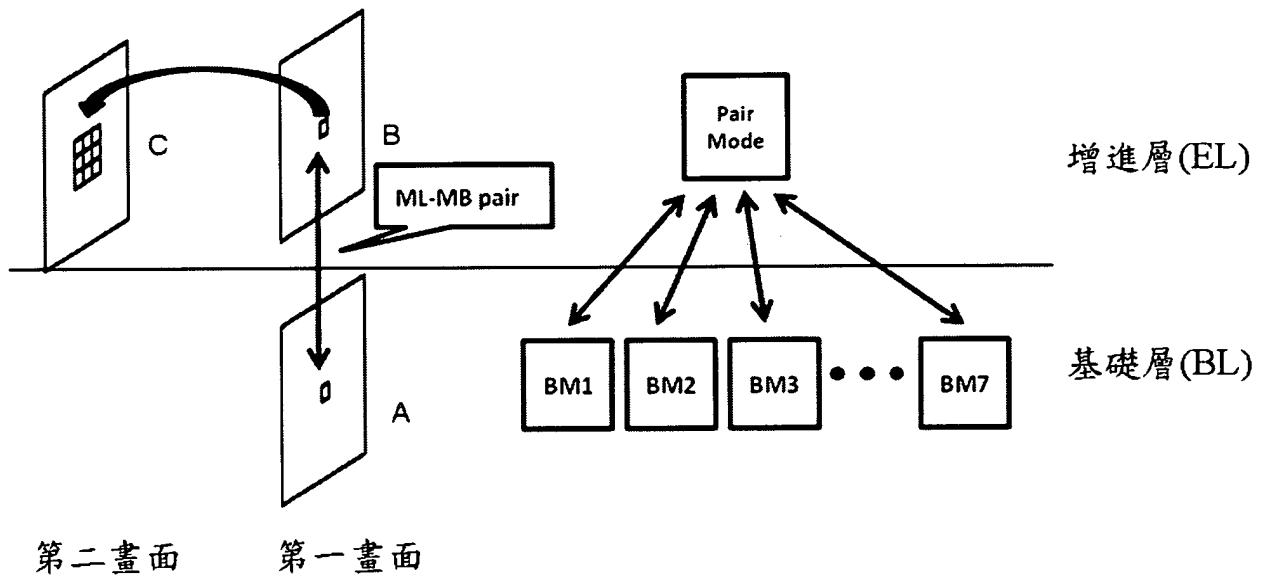
八、圖式：



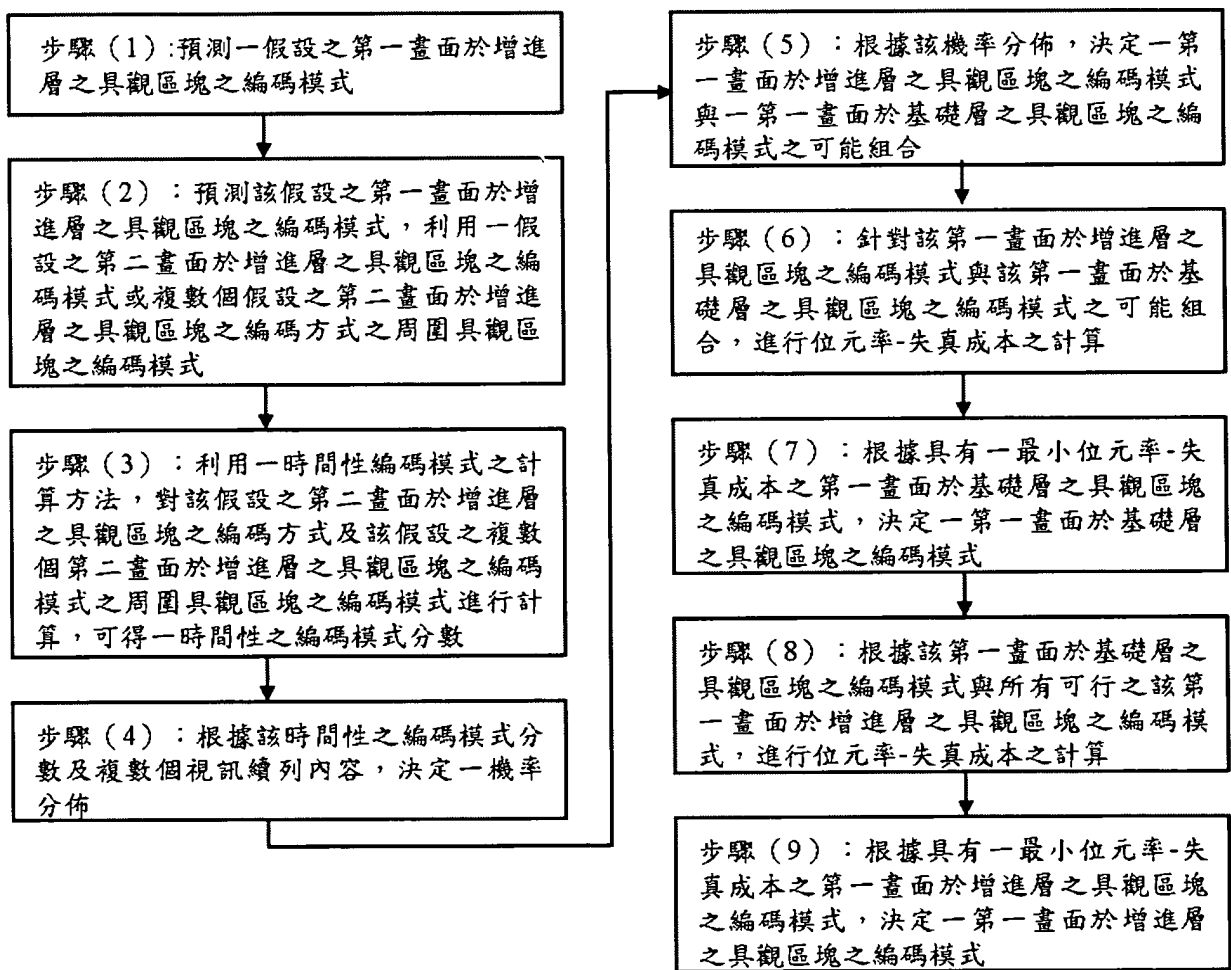
第 1 圖



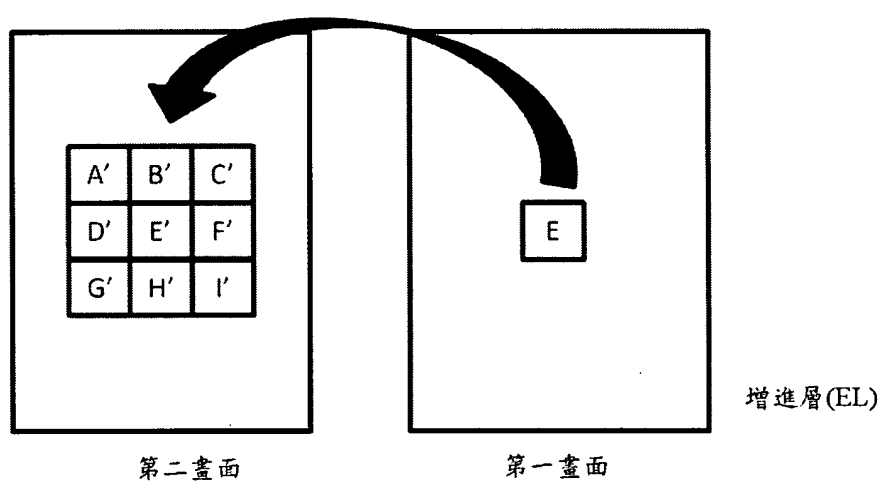
第 2 圖



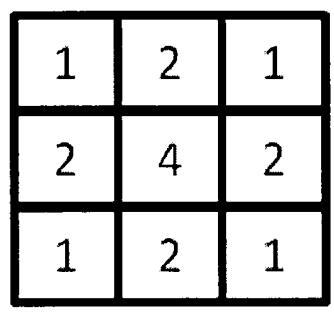
第 3 圖



第 4 圖



第 5 圖



(a)

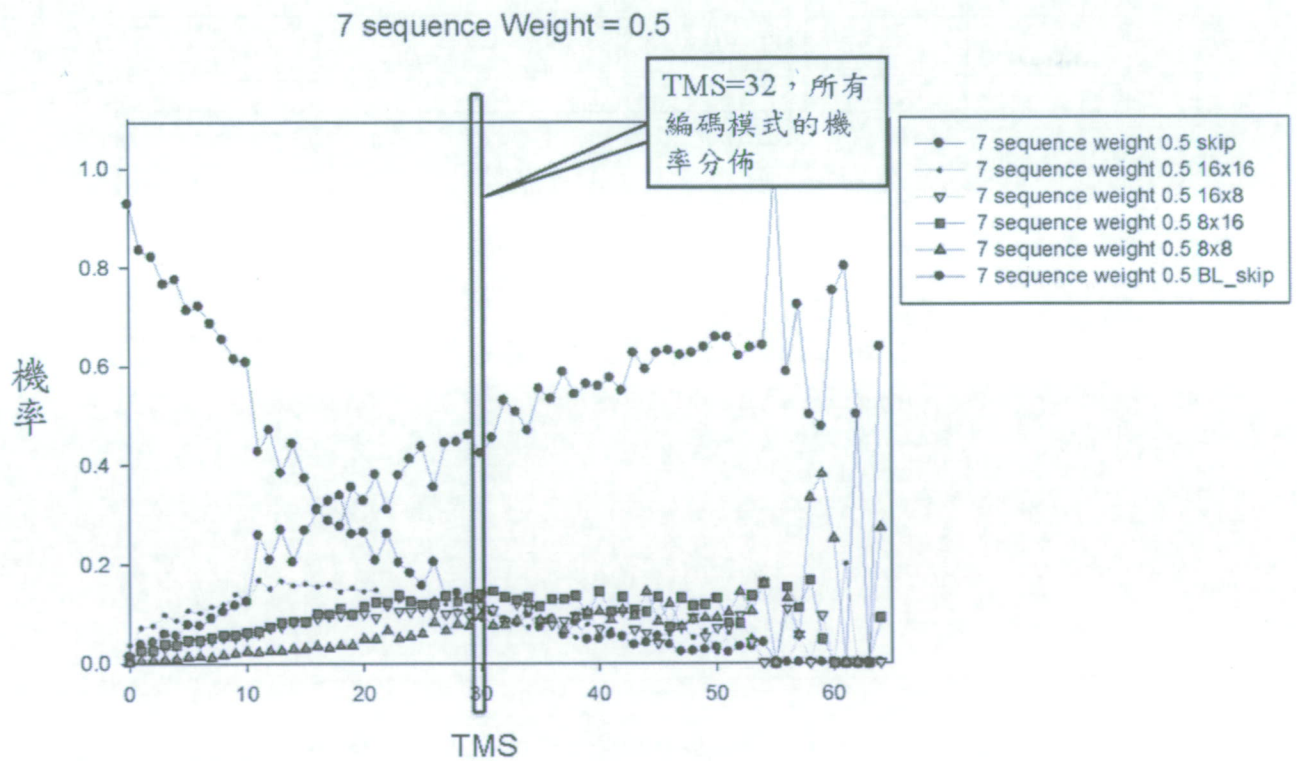
編碼模式	分數
skip	0
Inter16x16	1
Inter16x8	2
Inter8x16	2
Inter8x8	4
BL_skip	Follow BL mode

(b)

第 6 圖

skip	Inter16x16	BL_skip and BL:16x16
Inter16x16 (Intra 4)	Inter8x8	Inter16x8
Inter8x16	Inter8x16 (intra BL)	Inter16x16

第 7 圖



第 8 圖

			①			
BL	EL	skip	BL_skip	Intra BL	...	Intra 4
Skip			EL: BL_skip BL: skip			
②	16x16	EL: 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			

第 9 圖