



(21)申請案號：109126181

(22)申請日：中華民國 109 (2020) 年 08 月 03 日

(51)Int. Cl. : H04N19/176 (2014.01)

H04N19/50 (2014.01)

(71)申請人：緯創資通股份有限公司 (中華民國) WISTRON CORPORATION (TW)

新北市汐止區新台五路一段 88 號 21 樓

(72)發明人：李元兵 LEE, YUAN-BIN (TW)；丁筱雯 TIN, HSIAO-WEN (TW)

(74)代理人：葉璟宗；詹東穎；劉亞君

(56)參考文獻：

CN 106791828A

CN 108965892A

審查人員：張長軾

申請專利範圍項數：18 項 圖式數：7 共 47 頁

(54)名稱

視訊處理裝置及視訊串流的處理方法

(57)摘要

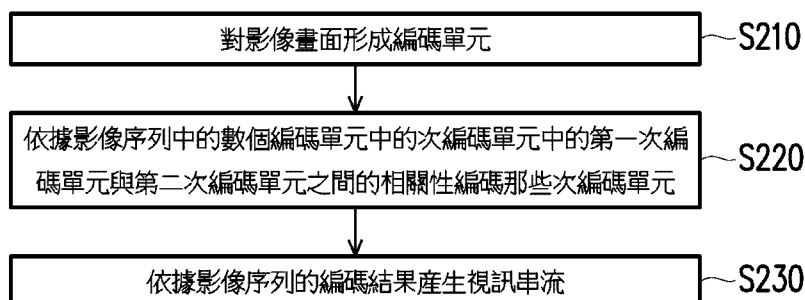
本發明實施例提出視訊處理裝置及視訊串流的處理方法。在此方法中，對影像畫面形成編碼單元，依據影像序列中的數個編碼單元中的次編碼單元中的第一次編碼單元與第二次編碼單元之間的相關性編碼那些次編碼單元，並依據影像序列的編碼結果產生視訊串流。影像畫面是影像序列中的一張畫面，且各編碼單元用於記錄所屬區塊的編碼資訊。藉此，可改善編碼效率。

A video processing apparatus and a processing method of video stream are provided. In the method, at least one encoding unit is formed in an image frame, the sub-encoding units in multiple encoding units of an the image sequence are encoded according to the relation between a first sub-encoding unit and a second sub-encoding unit in the sub-encoding units, and a video stream is generated according to the encoded result of the image sequence. The image frame is an image of the image sequence, and the encoding unit records encoding information of the belonging block. Therefore, the encoding efficiency may be improved.

指定代表圖：

符號簡單說明：

S210~S230:步驟



【圖2】



I743919

【發明摘要】

【中文發明名稱】

視訊處理裝置及視訊串流的處理方法

【英文發明名稱】

VIDEO PROCESSING APPARATUS AND PROCESSING METHOD
OF VIDEO STREAM

【中文】本發明實施例提出視訊處理裝置及視訊串流的處理方法。在此方法中，對影像畫面形成編碼單元，依據影像序列中的數個編碼單元中的次編碼單元中的第一次編碼單元與第二次編碼單元之間的相關性編碼那些次編碼單元，並依據影像序列的編碼結果產生視訊串流。影像畫面是影像序列中的一張畫面，且各編碼單元用於記錄所屬區塊的編碼資訊。藉此，可改善編碼效率。

【英文】 A video processing apparatus and a processing method of video stream are provided. In the method, at least one encoding unit is formed in an image frame, the sub-encoding units in multiple encoding units of an the image sequence are encoded according to the relation between a first sub-encoding unit and a second sub-encoding unit in the sub-encoding units, and a video stream is generated according to the encoded result of the image sequence. The image frame is an image of the image sequence, and the encoding unit

records encoding information of the belonging block. Therefore, the encoding efficiency may be improved.

【指定代表圖】圖2。

【代表圖之符號簡單說明】

S210~S230:步驟

【發明說明書】

【中文發明名稱】

視訊處理裝置及視訊串流的處理方法

【英文發明名稱】

VIDEO PROCESSING APPARATUS AND PROCESSING METHOD
OF VIDEO STREAM

【技術領域】

【0001】 本發明是有關於一種視訊編碼技術，且特別是有關於一種視訊處理裝置及視訊串流的處理方法。

【先前技術】

【0002】 近年來，機器學習(machine learning)技術在許多領域中已得到許多顯著的成果，更尤其是在電腦視覺和影像處理。然而，基於機器學習的視頻編碼(video coding)仍處於起步階段。視頻編碼(video coding)是網路電話(VoIP)的基礎，其效率直接影響到使用者的感受。相較於其他的背景雜訊抑制(background noise suppression)、回音消除(echo cancellation)及語音及臉部辨識(voice and facial recognition)等視頻領域應用，視頻編碼目前發展最快，並可透過機器學習提升或改善通話/視訊品質。值得注意的是，第五代(5G)行動通訊時代的來臨，在行動及高效傳輸的需求浮現後，影像傳輸品質的問題勢必再度被關注，且高效率的視頻編碼更不

斷試圖突破網路傳輸的限制。

【0003】 VoIP 在遠距離語音會議的應用方面，不但能節省費用，還能同時允許多人上線會談，且透過各種會議工具(例如，表達工具(presentation tool)、直播串流(live streaming)等)提供身臨其境的會議體驗。然而，隨著上線人數的增加，網路的頻寬需求與連線數量相對倍增，影像及語音延遲時間也因頻道擁塞而難以控制。此時，會議的品質(例如，服務品質(QoS)/體驗品質(OoE))往往大打折扣。VoIP 通常會經由有線網路連線至網際網路(Internet)上使用，但近年來有越來越多需求是經由行動通訊(例如，4G 或 5G 行動網路)提供服務，使得管理和部署更多資源不僅會增加成本，有時更成為不可能的任務(例如，行動環境中，頻寬不能超過一個上限)。隨著視頻技術的進步，尤其是超高解析度(Ultra-High Definition, UHD)視頻的普及，迫切需要進一步提高壓縮效率，以便將 UHD 視頻容納在有限的儲存空間和有限的傳輸頻寬。因此，如何有效地管理和控制資源，是業界期待解決的問題之一。

【0004】 此外，隨著網路頻寬不斷地提升，也相對帶動即時會議視訊服務的需求。在頻寬有限或分配不足的情況下，容易造成封包傳送延遲(Delay)、抖動(Jitter)或是封包遺失等網路效能降低的狀況，使得 VoIP 視頻的品質低落。因此，如何維持即時視訊傳輸品質，亦是業界期待解決的問題之一。

【發明內容】

【0005】 有鑑於此，本發明實施例提供一種視訊處理裝置及視訊串流的處理方法，基於影像畫面中的不同區塊的差異來編碼，從而降低運算複雜度，並達到影像傳輸需求。

【0006】 本發明實施例的視訊串流的處理方法包括(但不僅限於)下列步驟：對影像畫面形成編碼單元，依據影像序列中的各編碼單元中的次編碼單元中的第一次編碼單元與第二次編碼單元之間的相關性編碼那些次編碼單元，並依據影像序列的編碼結果產生視訊串流。影像畫面是影像序列中的一張畫面，且各編碼單元用於記錄所屬區塊的編碼資訊。

【0007】 本發明實施例的視訊處理裝置包括(但不僅限於)通訊收發器及處理器。通訊收發器用以傳送或接收資料。處理器耦接通訊收發器，並經配置用以對影像畫面形成編碼單元，依據影像序列中的各編碼單元中的次編碼單元中的第一次編碼單元與第二次編碼單元之間的相關性編碼那些次編碼單元，並依據影像序列的編碼結果產生視訊串流。影像畫面是影像序列中的一張畫面，且各編碼單元用於記錄所屬區塊的編碼資訊。

【0008】 基於上述，本發明實施例的視訊處理裝置及視訊串流的處理方法，對影像畫面的編碼單元處理，並依據編碼單元中的不同次編碼單元之間在空間及時間上的相關性來決定這些次編碼單元的決策，使差異較小的次編碼單元可忽略，從而減少編碼的複雜度，進而提升使用者體驗。

【0009】 為讓本發明的上述特徵和優點能更明顯易懂，下文特舉

實施例，並配合所附圖式作詳細說明如下。

【圖式簡單說明】

【0010】

圖 1 是依據本發明一實施例的視訊串流系統的方塊圖。

圖 2 是依據本發明一實施例的視訊串流的處理方法的流程圖。

圖 3 是一範例說明單元。

圖 4 是一範例說明各單元對應處理步驟的流程圖。

圖 5 是依據本發明一實施例的編碼處理方法的流程圖。

圖 6A 是一範例說明自動編碼器(AutoEncoder, AE)。

圖 6B 是依據本發明一實施例的應用 AE 決策單元大小的示意圖。

圖 6C 是一範例說明四分樹的結構。

圖 7 是依據本發明一實施例的影像品質評估方法的流程圖。

【實施方式】

【0011】 圖 1 是依據本發明一實施例的視訊串流系統 1 的方塊圖。請參照圖 1，視訊串流系統 1 包括(但不僅限於)一台或更多台客戶裝置 50 以及視訊處理裝置 100。

【0012】 客戶裝置 50 可以是桌上型電腦、筆記型電腦、智慧型手機、智慧型電視、平板、或機上盒等裝置。客戶裝置 50 包括(但不

僅限於)通訊收發器 51、儲存器 53、顯示器 55 及處理器 57。

【0013】 通訊收發器 51 可以是支援諸如乙太網路、光纖網路、行動網路、Wi-Fi 或其他通訊技術的傳送電路及接收電路。在一實施例中，通訊收發器 51 用以連線至網際網路或區域網路，進而與視訊處理裝置 100 相互傳送資料。

【0014】 儲存器 53 可以是任何型態的固定或可移動隨機存取記憶體(Random Access Memory, RAM)、唯讀記憶體(Read Only Memory, ROM)、快閃記憶體(flash memory)、傳統硬碟(Hard Disk Drive, HDD)、固態硬碟(Solid-State Drive, SSD)或類似元件。在一實施例中，儲存器 53 用以儲存軟體模組、應用程式及其程式碼、以及其他暫存或永久資料或檔案(例如，影像序列、影像畫面、編碼資訊、視訊串流、品質評估模型)，且這些資料將待後續實施例詳述。

【0015】 顯示器 55 可以是液晶顯示器(Liquid-Crystal Display, LCD)、發光二極體(Light-Emitting Diode, LED)顯示器或有機發光二極體(Organic Light-Emitting Diode, OLED)顯示器。在一實施例中，顯示器 55 用以呈現影像畫面或使用介面。

【0016】 處理器 57 可以是中央處理單元(Central Processing Unit, CPU)、圖形處理單元(Graphic Processing Unit, GPU)、微控制單元(Micro Control Unit, MCU)、或特殊應用積體電路(Application-Specific Integrated Circuit, ASIC)等運算單元。在一實施例中，處理器 57 用以執行客戶裝置 50 的所有運作，並可載入且執行儲存器 53 所儲存的軟體模組或應用程式，其詳細運作待後續實施例詳

述。

【0017】 視訊處理裝置 100 可以是桌上型電腦、筆記型電腦、伺服器、雲端平台、或後台主機等裝置。視訊處理裝置 100 包括(但不僅限於)通訊收發器 110、儲存器 130 及處理器 170。

【0018】 通訊收發器 110、儲存器 130 及處理器 170 的實施態樣可分別參照通訊收發器 51、儲存器 53 及處理器 57 的說明，於此不再贅述。

【0019】 在一實施例中，通訊收發器 110 用以連線至網際網路或區域網路，進而與客戶裝置 50 相互傳送資料。

【0020】 在一實施例中，儲存器 130 用以儲存軟體模組、應用程式及其程式碼、以及其他暫存或永久資料或檔案(例如，影像序列、影像畫面、編碼資訊、視訊串流、異常偵測模型)，且這些資料將待後續實施例詳述。

【0021】 在一實施例中，處理器 170 用以執行視訊處理裝置 100 的所有運作，並可載入且執行儲存器 130 所儲存的軟體模組或應用程式，其詳細運作待後續實施例詳述。

【0022】 下文中，將搭配視訊串流系統 1 中的各項裝置、元件及模組說明本發明實施例所述之方法。各個流程可依照實施情形而隨之調整，且並不僅限於此。

【0023】 圖 2 是依據本發明一實施例的視訊串流的處理方法的流程圖。請參照圖 2，處理器 170 對影像畫面形成一個或更多個編碼單元(步驟 S210)。具體而言，影像畫面是影像序列(即，由多張畫

面形成的序列)中的一張畫面。影像序列可能是由相機、錄影機或監視器拍攝所形成數位視訊資料，也可能是經任何運算裝置(例如，電腦、或智慧型手機等)編輯所形成的動畫、電影、廣告或電視劇等視訊資料。原始的影像序列可能包括相當多的資訊(例如，各畫素的色彩、明度等資訊)，導致資料量過大而不便於傳輸或儲存。因此，需要對影像序列進行視訊壓縮處理。

【0024】 在一實施例中，針對高效率視訊編碼(High Efficiency Video Coding, HEVC)或進階視訊編碼(MPEG-4 Part 10, Advanced Video Coding, MPEG-4 AVC)等視訊編碼標準，影像序列中的一個或更多張影像畫面會被進一步分割成一個或更多個處理單元(例如，編碼樹單元(Coding Tree Block, CTU)、巨集區塊(macroblock)或其他基本編碼單元)，且各處理單元還可能進一步對應分割成一個或更多個編碼單元。各編碼單元用於記錄所屬區塊的編碼資訊(例如，編碼模式、亮度、色彩、或語法等)。

【0025】 以 HEVC 為例，HEVC 技術在輸入影像序列進行編碼之前，會先對影像序列中的影像畫面分別切割成相等大小的編碼樹單元後再輸入至編碼器進行編碼。區塊(Block)是儲存資料之處，且許多區塊加上語法元素組成一個單元(Unit)。若輸入的影像畫面採用 YUV(明度(Luma)-色度(Chroma)-濃度)的顏色編碼格式，則一個編碼樹單元是由一個明度編碼樹區塊(Coding Tree Block, CTB)、兩個色度編碼樹區塊及語法元素(Syntax Element)所組成。在其他實施例中，顏色編碼格式可能是 RGB(紅色-綠色-藍色)或其他格式。

【0026】 在一實施例中，編碼單元還能進一步被分割而形成一個或更多個次編碼單元。在部分視訊編碼機制中，透過記錄區塊的實際值與畫面間(inter)或畫面內(intra)預測的預測值之間的殘量(residual)(即，實際值與預測值之間的差異)來壓縮影像畫面。而次編碼單元即是用於記錄所屬區塊的預測值與實際值之間的殘量。

【0027】 以 HEVC 為例，編碼單元還能進一步被分割形成一個或更多個預測單元(Prediction Unit, PU)(由亮度和色度預測塊組成)及轉換單元(Transform Unit, TU)(作為次編碼單元)。圖 3 是一範例說明單元。請參照圖 3，編碼樹單元可能包括不同大小的編碼單元 CU_1 , CU_2 , CU_3 (其大小例如分別是 32×32 、 16×16 及 8×8 像素)。編碼單元 CU_2 還可能進一步包括 8×8 像素的預測單元 PU、 8×8 像素的轉換單元 TU_1 及 4×4 像素的轉換單元 TU_2 。

【0028】 圖 4 是一範例說明各單元對應處理步驟的流程圖。請參照圖 4，處理器 170 輸入影像序列(步驟 S405)之後，將針對各影像畫面所形成的各編碼單元 CU 處理。針對編碼流程 410，對於預測單元 PU 而言，處理器 170 可進行畫面內預測(intra prediction)模式(步驟 S411)或畫面間預測(inter prediction)模式(步驟 S414)。畫面內預測模式包括畫面內預測估測(步驟 S412)及畫面內預測(步驟 S413)。而畫面間預測模式包括動態(motion)估測(步驟 S415)及動態補償(compensation)(步驟 S416)。畫面內預測模式或畫面間預測模式的預測值與原影像畫面的實際值之間的差異量即是殘量。對於轉換單元 TU 而言，殘量經轉換(步驟 S417)及量化(步驟 S418)

所形成的係數(coefficients)資料可與預測模式(畫面間預測及畫面內預測中擇一)、估測所得的動態資料(例如，動態向量)及濾波控制資料形成標頭(header)並據以熵(Entropy)編碼(步驟 S425)，即可形成位元串流(bit stream)(可進一步透過通訊收發器 110 轉換成封包並經由網路傳送到客戶裝置 50(可能會經過其他伺服器轉送))。

【0029】 針對解碼流程 430，客戶裝置 50 的處理器 57 可對轉換單元 TU 反量化(步驟 S431)及反轉換(步驟 S432)以還原出殘量。殘量與預測值合併後將經濾波控制分析(例如，確認環路(loop)濾波器)(步驟 S433)、以及解區塊(deblocking)及取樣自適應偏移(Sample Adaptive Offset, SAO)濾波(步驟 S434)來還原影像畫面，並據以儲存在畫面緩衝器(步驟 S435)以供顯示器 55 播放。

【0030】 須說明的是，在一些實施例中，依據不同視訊編碼規格，次編碼單元可能是不同於轉換單元的其他名稱。

【0031】 處理器 170 依據影像序列中的數個編碼單元中的那些次編碼單元中的第一次編碼單元與第二次編碼單元之間的相關性編碼那些次編碼單元(步驟 S220)。具體而言，雖然 HEVC 可節約流量，但卻增加了壓縮時間及解壓縮所需的運算量，進而大幅提高壓縮複雜性。研究指出，HEVC 的編碼時間(encoding time)平均比 H.264 / AVC 高 253%，這對於實現多媒體應用是不切實際的，尤其是即時視頻的 VoIP 應用，即時串流的延遲可能會造成許多損失及誤解。編碼單元分割預測方法是根據編碼中的一些特徵(例如，影像內容複雜度、動態向量(motion vector)等)及人為製定的決策規

則，來決定如何對編碼單元分割，且編碼單元再遞迴(Recursion)劃分出不同層次的樹結構(例如，四叉樹結構)以形成轉換單元(即，次編碼單元)。

【0032】 針對次編碼單元的編碼，本發明實施例對時間及空間維度分析。圖 5 是依據本發明一實施例的編碼處理方法的流程圖。請參照圖 5，在一實施例中，相關性是時間相關性。處理器 170 可依據時間相關性決定次編碼單元的單元大小(步驟 S510)。具體而言，處理器 170 依據第一次編碼單元與第二次編碼單元之間的差異決定第一次編碼單元或第二次編碼單元的單元大小，且第二次編碼單元是影像序列中不同於影像畫面的另一張畫面中的對應次編碼單元。處理器 170 利用“時間(temporal)”概念判斷不同影像畫面之間的差異(形成時間差異序列)，並透過時間上的差異大小來決定對應的單元大小。其中，若差異越大(代表像素之間越不相似)，則單元大小越小，且對應資訊量較多；若差異越小(代表像素之間越相似)，則單元大小越大，且對應資訊量較少。

【0033】 為了得出時間上的差異，在一實施例中，處理器 170 將第二次編碼單元輸入至異常偵測模型，以得出還原單元。異常偵測模型是將第一次編碼單元作為訓練資料並基於機器學習演算法所訓練。機器學習演算法可以是自動編碼器(AutoEncoder, AE)、降噪自動編碼器(Denoising Autoencoder)、主成分分析(Principal components analysis, PCA)或其他演算法。在一實施例中，處理器 170 可將一張或更多張場景影像(例如，會議室、辦公室的視訊會

議影像、或直播串流影像等，並視為正常影像)輸入到機器學習演算法的初始模型進行訓練，以形成異常偵測模型。在一些實施例中，處理器 170 還可使用測試資料(即，特定影像資料，例如是特定場景影像的集合)輸入到異常偵測模型，來測試異常偵測模型的準確率及失敗率，進而修正異常偵測模型。

【0034】 以自動編碼器為例，圖 6A 是一範例說明自動編碼器 AE。請參照圖 6A，自動編碼器 AE 的架構中可細分為編碼器(encoder)和解碼器(decoder)兩部分，它們分別進行壓縮與解壓縮的動作。在訓練過程中，讓輸出值和輸入值表示相同意義(損失函數(lose function)為輸出值和輸入值越接近越好)。自動編碼器 AE 的編碼器可進行降維(dimension reduction)，而解碼器可進行還原，使得自動編碼器 AE 可解釋為使用較低維度的特徵來近似原始輸入。原始輸入 OI 經編碼器處理可得到壓縮的嵌入編碼(embedding code)，這個編碼即是原始輸入 OI 的一個壓縮表示 CI。為了取得中間的編碼的輸入表示，壓縮表示 CI 的編碼經解碼器處理可得到重建輸入 RI。若此輸出的重建輸入 RI 與原始輸入 OI 相近，則代表這個壓縮表示 CI 與原始輸入 OI 之間存在特定映射關係。

【0035】 將自動編碼器 AE 概念應用在切割次編碼單元的決策中，在訓練階段中，處理器 170 可使用參考區塊(reference block)訓練自動編碼器 AE 的異常偵測模型。例如，參考區塊可以是將前述場景影像或其他影像分割成對應次編碼單元的數個訓練區塊。此外，假設異常偵測模型將參考區塊作為輸入資料可輸出相同或相近的

區塊。換句而言，可預期自動編碼器 AE 將參考區塊編碼後再解碼所得的估測區塊幾乎相同於參考區塊。接著，處理器 170 基於訓練資料中的輸入物件(例如，參考區塊)及預期的輸出結果(例如，相同或相近的參考區塊)訓練異常偵測模型，使輸入物件輸入到異常偵測模型可得出預期的輸出結果(幾乎相同於參考區塊的估測區塊)。例如，處理器 170 將輸入物件作為異常偵測模型對應函數的輸入變數，而輸出結果作為此函數的輸出變數，並據以找出此函數。

【0036】 處理器 170 可依據第一次編碼單元與第二次編碼單元的還原單元之間的差異形成時間差異序列。具體而言，處理器 170 可將影像序列中的一張影像畫面的某一個或更多個次編碼單元(例如，第一次編碼區塊)作為訓練所用的參考區塊，且透過已訓練的異常偵測模型推論此影像序列中不同時間點的另一張影像畫面的對應次編碼單元(例如，第二次編碼單元)的還原單元是否相近於參考區塊。若將第二次編碼單元輸入到自動編碼器 AE 所得出的估測區塊(即，其還原單元)相同或相近於訓練所用的參考區塊，則表示此次編碼單元的動態差異較低。若將第二次編碼單元輸入到自動編碼器 AE 得出的估測區塊(即，其還原單元)不同於訓練所用的參考區塊，則異常偵測模型可判斷此次編碼單元具有動態差異較高，並需要進一步對其編碼。

【0037】 圖 6B 是依據本發明一實施例的應用自動編碼器 AE 決策單元大小的示意圖。請參照圖 6B，待測的次編碼單元 TUT 的單元

大小例如是 32x32、16x16、8x8 或 4x4 像素。假設次編碼單元 TUT 已被分割成特定單元大小，次編碼單元 TUT 的值 x (對應到第一次編碼單元)經分析函數 g_a 轉換可得出新編碼 y (位於編碼空間 CS)，編碼 y 經量化得出量化值 q ，且量化值 q 被壓縮成位元後即可用於測量編碼率 R 。接著，量化值 q 經反量化後形成編碼 \hat{y} ，編碼 \hat{y} 經合成函數 g_s 反轉換後可獲得重建的次編碼單元 TUT(其值 \hat{x})。接著，將第二次編碼單元輸入到值 \hat{x} 。值為 x 及 \hat{x} 的次編碼單元 TUT 同時透過相同的函數 g_p (例如，自動編碼器的異常偵測模型)轉換，以分別得到還原單元的值為 z 及 \hat{z} (位於表示空間 PS)。處理器 170 可由公式(1)：

$$\min_{g_a g_s / q} D(z, \hat{z}) + \lambda R(q(y)) = V \dots (1)$$

得出兩者之間的差異 V (其中， λ 是常數， D 是用於得出讓值 z ， \hat{z} 相等的函數並可用諸如均方誤差(MSE)或差值等方式實現)。

【0038】 由於考慮的差異時間序列(對應到畫面間的像素之間的時間相關性)，因此處理器 170 會對影像序列中不同影像畫面中的對應次編碼單元依據時間順序產生一連串的差異 V ，並據以形成時間差異序列。處理器 170 即可判斷時間差異序列對應的單元大小。

【0039】 在一實施例中，處理器 170 可設定多個時間細度群組(分別對應到不同時間細粒度(grain fineness)(對應到不同單元大小))，並判斷時間差異序列所對應的時間細度群組。例如，表(1)是不同時間細度群組 T1~T4 對應的單元大小：

表(1)

時間細度群組	單元大小
T1	32×32
T2	16×16
T3	8×8
T4	4×4

假設差異 V 屬於時間細度群組 T1，則表示此第一次編碼單元的單元大小被設定為 32×32 像素，即平面區域或細節資訊較少。假設差異 V 屬於時間細度群組 T4，則表示此第一次編碼單元的單元大小被設定為 4×4 像素，即細節資訊較多。其餘依此類推，於此不再贅述。此外，圖 6C 是一範例說明四分樹的結構。請參照圖 6C，四個時間細度群組將形成四個階層，並可依照此階層進行編碼，例如以 1 代表分割為更小的單元。

【0040】 須說明的是，表(1)中的時間細度群組數量及其對應的單元大小僅是作為範例說明，應用者可視需求自行變化。此外，在一些實施例中，基於不同訓練機制，異常偵測模型還可能直接將第一次編碼單元分類到特定時間細度群組。

【0041】 請返回圖 5，在一實施例中，相關性是空間相關性。處理器 170 可依據空間相關性決定是否對次編碼單元編碼(步驟 S530)。具體而言，處理器 170 依據第一次編碼單元與第二次編碼單元之間的差異決定是否編碼。此時，第二次編碼單元是與第一次編碼單元相同的影像畫面中的另一次編碼單元。處理器 170 利用“空間”概念判斷相同影像畫面中不同區塊之間的差異(即，空間差異或空

間相關性)。在一實施例中，第一次編碼單元與第二次編碼單元之間在空間上的差異是第一次編碼單元與第二次編碼單元的畫素之間的差的絕對值總合((Sum of Absolute Difference, SAD)，其數學表示式(2)為：

$$SAD_{1,2} = \sum_{i=1}^N \sum_{j=1}^M |B_1(i,j) - B_2(i,j)| \dots (2)$$

其中， $SAD_{1,2}$ 是第一次編碼單元 B_1 與第二次編碼單元 B_2 在各像素坐標(i, j)之間的差的絕對值總合，i 為 1 至 N 之間的整數，j 是 1 至 M 之間的整數，N、M 為正整數。

【0042】 在其他實施例中，空間差異也可以是第一次編碼單元與第二次編碼單元的畫素之間的平均差、均方根誤差或其他差異的量度。在一些實施例中，圖 5 所示實施例可應用在圖 4 中畫面內編碼模式 S411 及畫面間編碼模式 S414。例如，在畫面內編碼模式 S411 中，使用自動編碼器 AE 及時間差異序列作為轉換單元 TU 的切割大小的判斷。又例如，在畫面間編碼模式 S414 中，使用轉換單元 TU 在差的絕對值總合來決定是否編碼此轉換單元 TU。

【0043】 空間差異得出之後，處理器 170 可依據此差異大小來決定是否對第一次編碼單元或第二次編碼單元(例如，對應到影像畫面中的某一區塊)編碼。其中，若差異越大(例如，未小於空間差異門檻值，即代表像素之間越不相似)，則處理器 170 將執行此第一次編碼單元或第二次編碼單元的編碼；若差異越小(例如，小於空間差異門檻值，即代表像素之間越相似，例如都是辦公室的牆(即，相同背景))，則處理器 170 將禁能此第一次編碼單元或第二次編碼

單元的編碼。

【0044】 在一實施例中，針對被禁能編碼的第一次編碼單元或第二次編碼單元，處理器 170 可將其合併到與其空間中像素差異度低的其他第三次編碼單元，以形成合併單元，且第三次編碼單元與第一次編碼單元或第二次編碼單元的差異小於空間差異門檻值。即，將差異較小的次編碼單元合併。接著，處理器 170 即可執行對合併單元的編碼。例如，第一次編碼單元是參考第三次編碼單元的預測值得出殘量。又例如，第一次編碼單元與第三次編碼單元結合在一起並同時進行預測及編碼。

【0045】 在另一實施例中，處理器 170 可設定多個空間等級(分別對應到不同空間細粒度)，並判斷空間差異所對應的空間等級。這些空間等級分別對應到不同像素差異程度，並依據差異程度大小排序。例如，表(2)是不同空間等級 S1~S3 對應的值域：

表(2)

空間等級	值域
S1	0 到第一等級上限
S2	第一等級上限至第二等級上限
S3	第二等級上限至第三等級上限

假設空間差異的值在 0 到第一等級上限之間，則歸類到空間等級 S1，且表示此第一次編碼單元與第二次編碼單元之間的像素差異較小。假設空間差異的值在第一等級上限到第二等級上限之間，則歸類到空間等級 S2，且表示此第一次編碼單元與第二次編碼單元

之間的像素差異相較大些(可能具有較粗糙的紋理)。假設空間差異的值在第二等級上限到第三等級上限之間,則歸類到空間等級 S3,且表示此第一次編碼單元與第二次編碼單元之間的像素差異更大(可能是影像畫面的邊緣像素)。

【0046】 在一些實施例中,處理器 170 可將前述判斷是否編碼的空間差異門檻值設定成第一等級上限、第二等級上限或第三等級上限,以決定那些空間等級的次編碼單元需要進一步編碼。

【0047】 須說明的是,表(2)中的空間等級數量及其對應的值域僅是作為範例說明,應用者可視需求自行變化。

【0048】 又一實施例中,處理器 170 可對整張影像畫面決定相關於自然影像的統計分佈。次編碼單元經切割之後,處理器 170 分別決定第一次編碼單元及第二次編碼單元對應的統計分佈,且判斷兩統計分佈之間的差異,並據以決定是否需要對第一次編碼單元或第二次編碼單元進一步編碼。於另一實施例中,處理器 170 可判斷第一次編碼單元或第二次編碼單元對應的統計分佈中是否具有重要特徵,並決定對具有重要特徵的第一次編碼單元或第二次編碼單元進行編碼。具體而言,自然影像的統計分佈例如是基於自然場景統計(Natural Scene Statistics, NSS)。自然場景統計能描述自然影像中的統計性質。處理器 170 可應用小波或拉普拉斯分別計算統計分佈。透過對整張影像畫面決定統計分佈後再分割成區塊的手段,處理器 170 可得知此區塊是否包括重要特徵。例如,假設一張人臉是很重要的,整張畫面的統計分佈中,對應於此張人臉

的部分的統計分佈會顯示出具有重要特徵，但是分割後的區塊卻是一個眼睛，此時統計分佈可表現出此區塊重要(即，眼睛對應的統計分佈為人臉的統計分佈中的一部分，故可從對應眼睛的統計分佈中判斷眼睛為重要特徵)。基於自然場景統計對應的重要程度，若不同次編碼單元的兩統計分佈相近(例如使用相似度門檻值比較)，則代表空間上的變化是平坦區域(即，重要程度相近，故不需要編碼或合併編碼)；若兩統計分佈不相近，則表示空間上的變化較大(即，重要程度的差異大，故需要編碼)。於一實施例中，若差的絕對值總合(SAD)所得的空間差異較小，則處理器 170 可進一步比對第一次編碼單元及第二次編碼單元的統計分佈。若兩者的統計分佈相近或相同，即可禁能編碼。藉由兩種比對手段(即，差的絕對值總合及統計分佈的比對)，可確保空間差異評估的正確性。

【0049】 在另一實施例中，影像可被切割成較小的評估區塊，並應用小波或拉普拉斯計算各評估區塊的統計分佈，以作為各評估區塊的特徵值。這些評估區塊的特徵值可做為編碼過程中特徵提取的確認。此特徵值代表 NSS 的規律性。在一些實施例中，前述場景影像也可被切成較小的評估區塊，並應用小波或拉普拉斯計算各評估區塊的統計分佈，以進行 NSS 規律評估。

【0050】 請回到圖 2，基於前述次編碼單元的單元大小決策及編碼與否的決策，處理器 170 可依據影像序列的編碼結果產生視訊串流(步驟 S230)。例如，處理器 170 使用設定的單元大小來承載預測資訊與實際資訊之間的殘量、對次編碼單元合併編碼、或者對設

定的次編碼單元編碼所產生的位元串流即是編碼結果(可參照圖 4 的編碼流程 410，但不限於特定視訊編碼標準)。視訊串流可進一步透過通訊收發器 110 傳送而出。另一方面，客戶裝置 50 的處理器 57 可透過通訊收發器 51 直接或間接地自視訊處理裝置 100 接收到此視訊串流，並對視訊串流解碼(與編碼作業相反的動作(例如，反量化、反轉換、加上預測值等))即可得到重建後的影像畫面，且可進一步透過顯示器 55 播放。

【0051】 除了編碼複雜度的降低，本發明實施例更進一步考慮影像品質，以避免減損使用者的觀看體驗。影像品質的評估方法可概括分為客觀及主觀二種，常見的客觀評估數學模型有均方根誤差(Root-Mean-Square Error, RMSE)，平均誤差(MSE)及峰值訊雜比(Peak Signal to Noise Ratio, PSNR)。以均方根誤差為例，均方根誤差的值越小，代表著兩張影像相似性愈高，但並不代表此影像的品質越好。為了能呈現使用者對即時串流的影像品質的主觀視覺感受，本發明實施例將位元率(bitrate)、串流內容及使用者體驗三種資訊作為影像品質的評估因素。

【0052】 在與「解析度」相配性的前提下，較高的位元率可容納更高的視訊品質。因此，當傳輸資料的位元率增加時，人們的感知品質會提高。為了提升使用者的即時串流的品質滿意度及有效的利用頻寬，傳輸的資料量應該在位元率較低時提高。本發明實施例將「編碼等級」用於調整位元率的機制可滿足使用者主觀對即時串流的品質滿意度，而不是廣泛性的提升或降低位元率。

【0053】 在一實施例中，處理器 170 可依據傳送視訊串流的傳輸頻寬決定數個編碼等級中的一者。這些編碼等級即是分別對應前述那些空間等級。例如，編碼等級 b1 對應到表(2)的空間等級 S1，編碼等級 b2 對應到空間等級 S2，且編碼等級 b3 對應到空間等級 S3。須說明的是，編碼等級不限於三個，並可視需求而變化。

【0054】 在保持影像品質的前提下，處理器 170 可依據當前的傳輸頻寬設置最接近但不超過的可用傳輸頻寬的編碼層級，以增加或減少資料傳輸流量。主要概念是將即時串流中的影像畫面中相較複雜或有較多變化的部份進行處理(例如，編碼及傳輸)，而簡單的部份(例如，沒有變化或較少變化的部份)則不進行處理，解碼時使用前面已經編碼過的畫面進行預測即可。透過此等動態調整方式可進一步取得編碼效率及節約流量之間的平衡，從而達到使用者對即時串流的品質(例如是 QoE)的期望。

【0055】 當傳輸頻寬小於頻寬門檻值時，代表可用傳輸頻寬不夠，且需要減少資料傳輸量，因此編碼等級須提升。例如，將當前編碼等級升高一個等級或視可用傳輸頻寬調整。反之，當傳輸頻寬未小於頻寬門檻值時，代表可用傳輸頻寬足夠，且可進一步增加資料傳輸量，因此編碼等級可降低。例如，將當前編碼等級降低一個等級或降到最低等級。舉例而言，假設當前編碼等級為 b2，當傳輸頻寬小於 40 Kbps(即，頻寬門檻值)時，則處理器 170 將編碼等級修改成 b3。

【0056】 此外，處理器 170 可依據決定的編碼等級改變前述空間

差異門檻值。若編碼等級增加(即，傳輸頻寬降低)，則空間差異門檻值也增加；若編碼等級降低(即，傳輸頻寬增加)，則空間差異門檻值也降低。舉例而言，當傳輸頻寬降低時，編碼等級被提升至 b3，且只對邊緣像素的區塊進行編碼及傳輸，而其他部份是使用前面已經編碼過的畫面進行預測，即可減少傳輸資料量。相反而言，當傳輸頻寬增加時，編碼等級被降低到 b2 或 b1，則需要進一步對紋理(texture)或平坦(flat)的區塊進行編碼及傳輸。藉此，可因應於傳輸頻寬的變化而重新分配編碼機制，進而改善了編碼效率及用戶對即時串流的體驗。

【0057】 以實際情境說明，假設目前編碼等級為對應空間等級 S2 的 b2，且頻寬門檻值設為 40 Kbps。當傳輸頻寬小於頻寬門檻值時，視頻會議應用程序將無法正常工作(可能會有延遲或封包遺失等情況發生)，因此無法滿足用戶的即時視頻期望。

【0058】 一旦傳輸頻寬小於頻寬門檻值，處理器 170 可將空間差異門檻值及編碼等級提升(但不可高於最高的空間等級的上限及編碼等級的上限)。例如，編碼等級提升到為 b3，空間差異門檻值設為空間等級 S3，代表次編碼單元是邊緣(即，空間差異大於空間等級 S2 的上限)才需編碼。即，只對有即時串流進行重點(此處將空間差異較大的次編碼單元視為重點)的編碼，以降低整體的位元率，提高編碼效率，並維持即時串流品質。

【0059】 在另一實施例中，視訊處理裝置 100 可依據客戶裝置 50 對於影像品質評估(例如，針對前述串流內容及使用者體驗因素)的

回饋來改變編碼機制。圖 7 是依據本發明一實施例的影像品質評估方法的流程圖。請參照圖 7，客戶裝置 50 的處理器 57 對視訊串流中的影像形成數個待評估區塊(步驟 S710)。具體而言，有別於傳統技術對於整張影像的所有像素進行影像品質評估，本發明實施例是以區塊為基本單位來評估。而待評估區塊可以是前述次編碼單元的單元大小或是其他大小。

【0060】 處理器 57 可將待評估區塊輸入品質評估模型(步驟 S730)。具體而言，品質評估模組是基於機器學習演算法所訓練。此機器學習演算法可以是膠囊網路(Capsule network, CapsNet)、卷積神經網路(Convolutional neural network, CNN)、緊密卷積網路(Dense Convolutional Network, DenseNet)或其他演算法。

【0061】 以膠囊網路為例，在影像識別上和 CNN 不同之處在於，使用接近生物神經組織(biological neural organization)的概念。當視角(viewpoint)改變時，部分(part)或物件(object)具有線性影響，而像素則是非線性影響。膠囊網路是帶有方向的向量，並可根據統計訊息進行特徵檢測。例如，根據五官辨識出一隻「貓」。此外，膠囊網路還可以對特徵進行理解。例如，若五官不在其特定的區域(例如，嘴在額頭、或眼睛在耳朵上方)，則膠囊網路能透過檢測出物體的不同方向(例如，順時針旋轉)，從而了解處於特定位置的貓五官才能被稱作是「貓」。

【0062】 在一實施例中，處理器 57 可將前述場景影像切割成數個待學習區塊，並將這些場景影像的那些待學習區塊作為學習資料，

且將此學習資料作為訓練品質評估模型的訓練樣本。例如，會議影像作為場景影像，則訓練品質評估模型後即可將影像中的人臉作為重要資訊。即，取得影像特徵，並據以作為人臉感興趣區域。而影像特徵是指將待學習區塊輸入機器學習演算法的神經網路中，可針對待學習區塊學習出局部影像的特徵。例如，紋理和梯度信息。而在膠囊網路中可以向量記錄紋理和梯度訊息、以及相對位置等資訊。此外，膠囊網路的作用可以確保重要資訊的位置正確，進而輕易地分辨出重要資訊。

【0063】 在另一實施例中，處理器 57 可將例如是平均意見分數 (Mean Opinion Score, MOS) 或差分平均意見分數 (Differential Mean Opinion Score, DMOS) 影像資料庫中的資料作為學習資料，對學習資料中的影像切割成數個待學習區塊，並將這些待學習區塊作為訓練品質評估模型的訓練樣本。

【0064】 在又一實施例中，處理器 57 可統計一位或更多位人員對那些待學習區塊主觀評斷的觀感評估分數。觀感評估分數是基於人類觀感實際的評分結果所得出。以國際標準化的協議電信聯盟 (ITU) 測試為例，ITU 測試分為幾個段落，至少由 15 位觀察人員組成，觀察人員受要求以 1 到 5 的分數對待學習區塊評分。每個人的主觀影像質量評分標準可能不同，使得不同人通常給出的分數會不同。每位觀察人員給出的分數取決於自身經驗及喜好，甚至可能根據當天的心情及事物而變化。處理器 57 可將那些待學習區塊對應的觀感評估分數作為訓練品質評估模型的訓練樣本。而透過

對區塊的觀感評估分數輸入至機器學習演算法的神經網路，可建立品質評估模型，並讓品質評估模型可用於推論其他情境下的視訊串流的影像品質。

【0065】 再一實施例中，處理器 57 可依據那些待學習區塊對應的影像特徵、觀感評估分數及/或統計分佈(基於自然場景統計)決定數個目標函數中的一者，並作為訓練品質評估模型的訓練樣本。具體而言，目標函數可以是粒子群最佳化 (Particle Swarm Optimization, PSO) 的適應性函數(fitness function)、模糊決策的成本函數計算或其他最佳化函數。這些目標函數具有上限及下限，即目標函數的最佳解在上下及下限之間的可行區域(feasible region)(或稱容許範圍)內變動。在編解碼過程中，資訊減少或遺失是無法避免的。因此，求取目標函數的解可在一個區間內，目標函數的上限及下限的目的是指影像品質提高或減損可在某個容忍範圍內。求得最佳解(即，和原始影像對應區塊相同的影像品質)是主要目標，其次可找尋其他相關的可行解(即，在容忍範圍內的影像品質提高或減損)。

【0066】 以影像特徵、觀感評估分數及統計分佈三者皆作為目標函數的參數為例：

$$\min_f = \min \text{block}(t, m, n) \dots (3)$$

$$\max_f = \max \text{block}(t, m, n) \dots (4)$$

目標函數 $\text{block}(t, m, n)$ 包括待評估區塊的影像特徵 t 、觀感評估分數 m 及統計分佈 n 。 \max_f 代表上限， \min_f 代表下限，即求解目標

可在上限 max_f 和下限 min_f 間變動。

【0067】 須說明的是，方程式(3)、(4)僅是作為範例說明，其他實施例也可能是將影像特徵與觀感評估分數及統計分佈中的任一者作為目標函數的參數。

【0068】 處理器 57 可依據品質評估模型的輸出結果決定影像的品質(步驟 S750)。具體而言，經訓練的品質評估模組可將待評估區塊作為輸入資料，並據以推論其對應的結果。在一實施例中，若場景影像及/或訓練資料的影像特徵、觀感評估分數、統計分佈及目標函數作為品質評估模組的學習資料，則品質評估模型的輸出結果可包括特定的影像特徵、觀感評估分數、統計分佈及目標函數。例如，品質評估模組對視訊串流經重建後的影像畫面中的待評估區塊擷取影像特徵，並辨識出對應此影像特徵所屬類別(例如，人臉、眼睛、或嘴)的區塊，再依據此類別的區塊取得對應的觀感評估分數、統計分佈及目標函數。

【0069】 接著，處理器 57 可依據輸出結果中的影像特徵、觀感評估分數、統計分佈及目標函數，決定各待評估區塊對應的目標函數的計算值(即，同時依據影像特徵、觀感評估分數及統計分佈決定對應待評估區塊的計算值)。處理器 57 可依據此計算值判斷對應待評估區塊的品質。即，計算值為影像品質的指標。

【0070】 在一實施例中，處理器 57 可判斷計算值是否介於對應的目標函數的上限及下限之間(即，容許範圍)。若此計算值位於最佳解的上限及下限內，即代表使用者對此待評估區塊的感觀接受度

在容許範圍內(差值代表容忍度，並可經由經驗而來或動態修正最佳解的上下限)，並可進一步透過顯示器 55 播放重建的影像畫面(即，視訊串流經解碼後的影像畫面)。若計算值未位於上限及下限內或無解，則代表目標函數無法得出最佳解、使用者對此待評估區塊的感觀接受度不在容許範圍內或畫面失真，並可進一步透過通訊收發器 51 發出對影像重新編碼的要求。而若計算值等於最佳解，則代表使用者對此待評估區塊的感觀接受度相同，並可進一步透過顯示器 55 播放重建的影像畫面。

【0071】 在另一實施例中，若場景影像及/或訓練資料的影像特徵、觀感評估分數及目標函數作為品質評估模組的學習資料，則品質評估模型的輸出結果可包括特定的影像特徵、觀感評估分數及目標函數(此實施例是依據影像特徵及觀感評估分數形成目標函數)。接著，處理器 57 可將輸出結果中的影像特徵及觀感評估分數輸入至輸出結果中的目標函數，並據以決定各待評估區塊對應的目標函數的計算值(即，依據影像特徵及觀感評估分數決定對應待評估區塊的計算值)，並據以作為影像品質的指標。

【0072】 又一實施例中，若場景影像及/或訓練資料的影像特徵、統計分佈及目標函數作為品質評估模組的學習資料，則品質評估模型的輸出結果可包括特定的影像特徵、統計分佈及目標函數(此實施例是依據影像特徵及統計分佈形成目標函數)。接著，處理器 57 可將輸出結果中的影像特徵及統計分佈輸入至輸出結果中的目標函數，並據以決定各待評估區塊對應的目標函數的計算值(即，

依據影像特徵及統計分佈決定對應待評估區塊的計算值)，並據以作為影像品質的指標。

【0073】 由於影像畫面中的畫素分佈並非均勻分佈，因此影像畫面中的單一待評估區塊不一定可以反映出整張影像畫面的所有畫素，且各待評估區塊得到使用者關注的程度也許不同。一般而言，重要資訊(例如，會議室中的人比會議室背景更重要)或變化較大的資訊(例如，切換場景或換人發言)是使用者所關注的部份。

【0074】 為了考慮不同待評估區塊對影像品質評估的影響，本發明實施例使用編碼等級做為待評估區塊的加權平均所用的權重。處理器 57 可決定各待評估區塊的編碼等級，且不同的編碼等級對應到不同的權重。以表(2)的編碼等級為例，編碼等級 b3、b2 及 b1 分別對應到權重 w3、w2 及 w1。相較於編碼等級 b1，編碼等級 b3 會讓使用者有更多注意力，因此權重 w3 的值大於權重 w1 的值。即，編碼等級越高，權重的值越高。相反而言，編碼等級越低，權重的值越低。處理器 57 可依據那些待評估區塊的品質(例如是前述計算值)及其對應的權重決定整張影像畫面的品質。影像畫面的品質的計算公式如下：

$$Q = \frac{\sum_{k=1}^L q_k w_k}{\sum_{k=1}^L w_k} \dots (5)$$

Q 為整張影像畫面的品質， q_k 為第 k 待評估區塊的品質， w_k 是第 k 待評估區塊的權重，假設影像畫面分割成 L 個待評估區塊，且 L 為正整數。

【0075】 以影像畫面被分割成三個待評估區塊為例，其編碼等級

b3、b2 及 b1 對應到的權重分別為 60%、30% 及 10%。整張影像畫面的品質為 $(q_1 * 0.6 + q_2 * 0.3 + q_3 * 0.1) / (0.6 + 0.3 + 0.1)$ 。

【0076】 須說明的是，若影像畫面的品質的分數是以 DMOS 或 MOS 來計分，則 DMOS 的分數越高表示品質越差，或 MOS 的分數越高表示品質越高。

【0077】 此外，若得出品質不佳的回饋，視訊處理裝置 100 可依據影像品質評估結果來改變編碼機制。在一實施例中，處理器 170 更依據視訊串流經解碼後的影像品質(例如，前述整張影像畫面的品質、或部分待評估區塊的品質)改變空間差異門檻值。例如，若影像品質不佳(例如，其計算值不位於目標函數的容許範圍內)，則處理器 170 可降低空間差異門檻值，使降低空間等級(即，值域較低)的次編碼單元以取得更多編碼資訊(原先這些較低空間等級的次編碼單元不執行編碼，解碼時將參考其他預測值，且降低空間差異門檻值後若這些次編碼單元符合較低空間等級也可執行編碼)，並重新編碼那些次編碼單元。

【0078】 在另一實施例中，處理器 170 可依據視訊串流經解碼後的影像品質改變單元大小的決定。若影像品質不佳，則處理器 170 可在編碼流程中對單元大小的決定放寬條件。例如，處理器 170 可降低時間細度群組對應的單元大小，使細節資訊增加。

【0079】 綜上所述，本發明實施例視訊處理裝置及視訊串流的處理方法，將影像畫面中的空間及時間特徵分為不同細粒度，使用異常偵測模型作為編碼單元結構切割次編碼單元的決策依據，並判

斷各次編碼單元屬於那個空間等級的區間門檻值，進而判斷次編碼單元是否需要進行編碼，從而減少編碼複雜度。此外，使用 NSS 統計特性可助於提供更精確的分割決策。另一方面，本發明實施例整合傳輸頻寬、觀感評估分數及自然場景統計分佈三種資訊作為影像品質的評估因素。基於編碼等級動態調整位元率以滿足使用者主觀對即時串流的品質滿意度，而不是廣泛性的提升或降低位元率。此外，本發明實施例採用無參考(no-reference, NR)影像品質評估概念，不需參考影像的比對，也不需要對影像中的所有畫素擷取影像特徵。由於參考 NSS 統計特性，因此本發明實施例機器學習演算法的神經網路擷取待評估區塊的影像特徵，並使用編碼等級對待評估區塊的像素進行加權，從而對整個影像進行品質評估。影像品質評估還能進一步回饋到編碼機制，讓調整後的編碼機制能符合人類視覺觀感的需求。

【0080】 雖然本發明已以實施例揭露如上，然其並非用以限定本發明，任何所屬技術領域中具有通常知識者，在不脫離本發明的精神和範圍內，當可作些許的更動與潤飾，故本發明的保護範圍當視後附的申請專利範圍所界定者為準。

【符號說明】

【0081】

1:視訊串流系統

50:客戶裝置

51、110:通訊收發器

53、130:儲存器

55:顯示器

57、170:處理器

100:視訊處理裝置

S210~S230、S405~S435、S510~S530、S710~S750:步驟

CU、CU₁、CU₂、CU₃:編碼單元

PU:預測單元

TU、TU₁、TU₂:轉換單元

410:編碼流程

430:解碼流程

AE:自動編碼器

OI:原始輸入

CI:壓縮表示

RI:重建輸入

CS:編碼空間

PS:表示空間

TUT:次編碼單元

x 、 \hat{x} 、 z 、 \hat{z} :值

y 、 \hat{y} :編碼

g_p :函數

q :量化值

D:函數

R:編碼率

【發明申請專利範圍】

【請求項1】 一種視訊串流的處理方法，包括：

對一影像畫面形成至少一編碼單元，其中該影像畫面是一影像序列中的一張畫面，且每一該編碼單元用於記錄所屬區塊的編碼資訊；

依據該影像序列中的多個該編碼單元中的多個次編碼單元中的一第一次編碼單元與一第二次編碼單元之間的相關性編碼該些次編碼單元；以及

依據該影像序列的編碼結果產生一視訊串流，

其中依據該影像序列中的該些編碼單元中的該些次編碼單元中的該第一次編碼單元與該第二次編碼單元之間的相關性編碼該些次編碼單元的步驟包括：

依據該第一次編碼單元與該第二次編碼單元之間的差異決定該第一次編碼單元或該第二次編碼單元的單元大小，其中該第二次編碼單元是該影像序列中不同於該影像畫面的另一張畫面中的對應次編碼單元；或

依據該第一次編碼單元與該第二次編碼單元之間的空間差異決定是否編碼該第一次編碼單元或該第二次編碼單元，其中該第二次編碼單元是與該第一次編碼單元相同的該影像畫面中的另一次編碼單元。

【請求項2】 如請求項1所述的視訊串流的處理方法，其中依據該影像序列中的該些編碼單元中的該些次編碼單元中的該第一次編

碼單元與該第二次編碼單元之間的相關性編碼該些次編碼單元的步驟更包括：

反應於該空間差異小於一空間差異門檻值，禁能該第一次編碼單元或該第二次編碼單元的編碼；以及反應於該空間差異未小於該空間差異門檻值，執行該第一次編碼單元或該第二次編碼單元的編碼。

【請求項3】 如請求項1所述的視訊串流的處理方法，其中依據該第一次編碼單元與該第二次編碼單元之間的差異決定該第一次編碼單元或該第二次編碼單元的單元大小的步驟更包括：

將該第二次編碼單元輸入至一異常偵測模型，以得出一還原單元，其中該異常偵測模型是將該第一次編碼單元作為訓練資料並基於一機器學習演算法所訓練；

依據該第一次編碼單元與該第二次編碼單元的該還原單元之間的差異形成一時間差異序列；以及

判斷該時間差異序列對應的單元大小。

【請求項4】 如請求項1所述的視訊串流的處理方法，其中該第一次編碼單元與該第二次編碼單元之間的該空間差異是該第一次編碼單元與該第二次編碼單元的畫素之間的差的絕對值總合(Sum of Absolute Difference, SAD)。

【請求項5】 如請求項1所述的視訊串流的處理方法，其中依據該第一次編碼單元與該第二次編碼單元之間的該空間差異決定是否編碼該第一次編碼單元或該第二次編碼單元的步驟更包括：

決定該影像畫面相對於自然影像的統計分佈；

依據該影像畫面的統計分佈分別決定該第一次編碼單元及該第二次編碼單元對應的統計分佈；以及

依據該第一次編碼單元及該第二次編碼單元對應的統計分佈決定是否編碼該第一次編碼單元或該第二次編碼單元。

【請求項6】 如請求項2所述的視訊串流的處理方法，其中禁能該第一次編碼單元或該第二次編碼單元的編碼的步驟包括：

將該第一次編碼單元或該第二次編碼單元合併到一第三次編碼單元，以形成一合併單元，其中該第三次編碼單元與該第一次編碼單元或該第二次編碼單元的該空間差異小於該空間差異門檻值；以及

執行對該合併單元的編碼。

【請求項7】 如請求項1所述的視訊串流的處理方法，其中依據該第一次編碼單元與該第二次編碼單元之間的空間差異決定是否編碼該第一次編碼單元或該第二次編碼單元的步驟包括：

依據傳送該視訊串流的一傳輸頻寬決定多個編碼等級中的一者；以及

依據決定的該編碼等級改變該空間差異門檻值。

【請求項8】 如請求項1所述的視訊串流的處理方法，更包括：

依據該視訊串流經解碼後的影像品質改變該空間差異門檻值。

【請求項9】 如請求項1所述的視訊串流的處理方法，更包括：

依據該視訊串流經解碼後的影像品質改變該單元大小。

【請求項10】 一種視訊處理裝置，包括：

一通訊收發器，用以傳送或接收資料；以及

一處理器，耦接該通訊收發器，並經配置用以：

對一影像畫面形成至少一編碼單元，其中該影像畫面是一影像序列中的一張畫面，且每一該編碼單元用於記錄所屬區塊的編碼資訊；

依據該影像序列中的多個該編碼單元中的多個次編碼單元中的一第一次編碼單元與一第二次編碼單元之間的相關性編碼該些次編碼單元；以及

依據該影像序列的編碼結果產生一視訊串流，其中該視訊串流透過該通訊收發器傳送，

其中該處理器更經配置用以：

依據該第一次編碼單元與該第二次編碼單元之間的差異決定該第一次編碼單元或該第二次編碼單元的單元大小，其中該第二次編碼單元是該影像序列中不同於該影像畫面的另一張畫面中的對應次編碼單元；或

依據該第一次編碼單元與該第二次編碼單元之間的空間差異決定是否編碼該第一次編碼單元或該第二次編碼單元，其中該第二次編碼單元是與該第一次編碼單元相同的該影像畫面中的

另一次編碼單元。

【請求項11】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

反應於該空間差異小於一空間差異門檻值，禁能該第一次編碼單元或該第二次編碼單元的編碼；以及

反應於該空間差異未小於該空間差異門檻值，執行該第一次編碼單元或該第二次編碼單元的編碼。

【請求項12】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

將該第二次編碼單元輸入至一異常偵測模型，以得出一還原單元，其中該異常偵測模型是將該第一次編碼單元作為訓練資料並基於一機器學習演算法所訓練；

依據該第一次編碼單元與該第二次編碼單元的該還原單元之間的差異形成一時間差異序列；以及

判斷該時間差異序列對應的單元大小。

【請求項13】 如請求項10所述的視訊處理裝置，其中該第一次編碼單元與該第二次編碼單元之間的該空間差異是該第一次編碼單元與該第二次編碼單元的畫素之間的差的絕對值總合。

【請求項14】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

決定該影像畫面相對於自然影像的統計分佈；

依據該影像畫面的統計分佈分別決定該第一次編碼單元及該

第二次編碼單元對應的統計分佈；以及

依據該第一次編碼單元及該第二次編碼單元對應的統計分佈決定是否編碼該第一次編碼單元或該第二次編碼單元。

【請求項15】 如請求項11所述的視訊處理裝置，其中該處理器經配置用以：

將該第一次編碼單元或該第二次編碼單元合併到一第三次編碼單元，以形成一合併單元，其中該第三次編碼單元與該第一次編碼單元或該第二次編碼單元的該空間差異小於該空間差異門檻值；以及

執行對該合併單元的編碼。

【請求項16】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

依據傳送該視訊串流的一傳輸頻寬決定多個編碼等級中的一者；以及

依據決定的該編碼等級改變該空間差異門檻值。

【請求項17】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

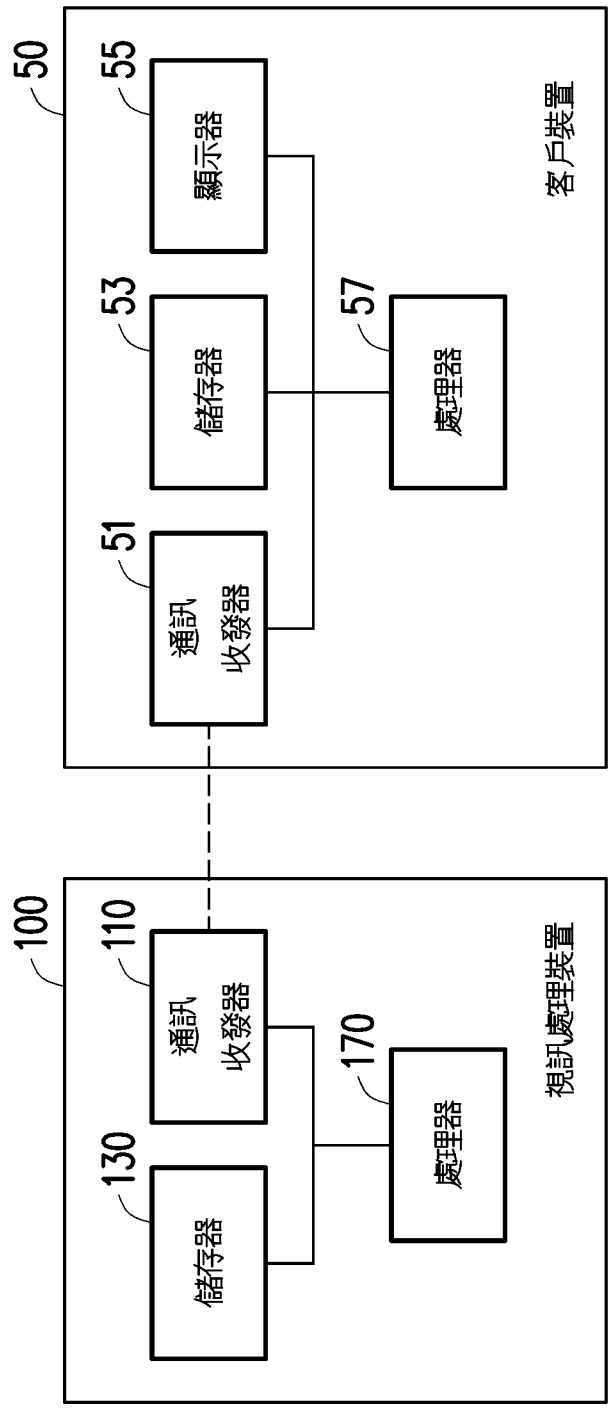
透過該通訊收發器接收該視訊串流經解碼後的影像品質；以及

依據該影像品質改變該空間差異門檻值。

【請求項18】 如請求項10所述的視訊處理裝置，其中該處理器經配置用以：

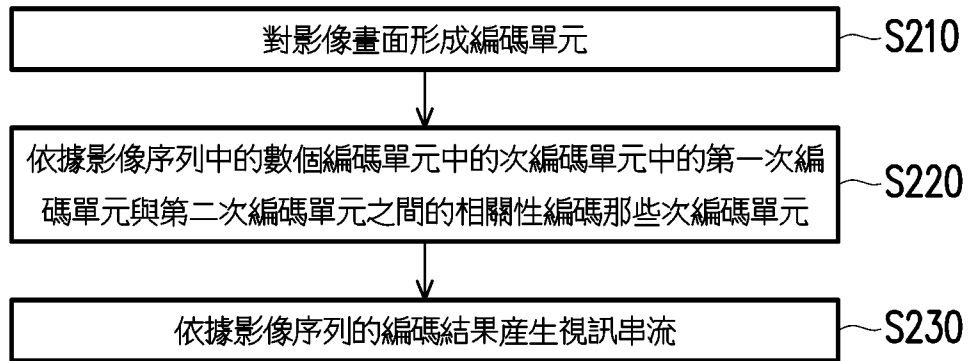
透過該通訊收發器接收該視訊串流經解碼後的影像品質；以
及
依據該影像品質改變該單元大小。

【發明圖式】

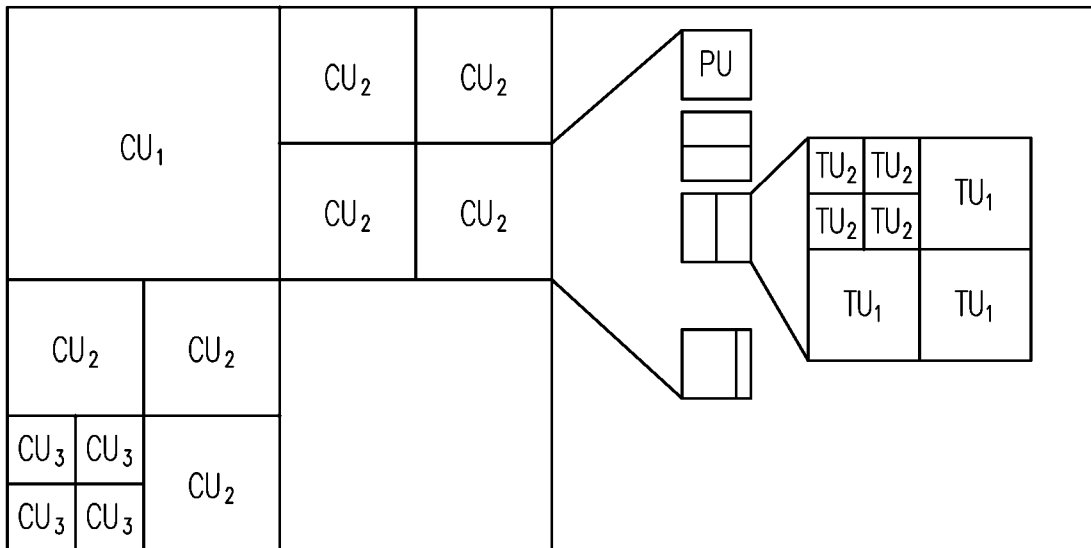


1

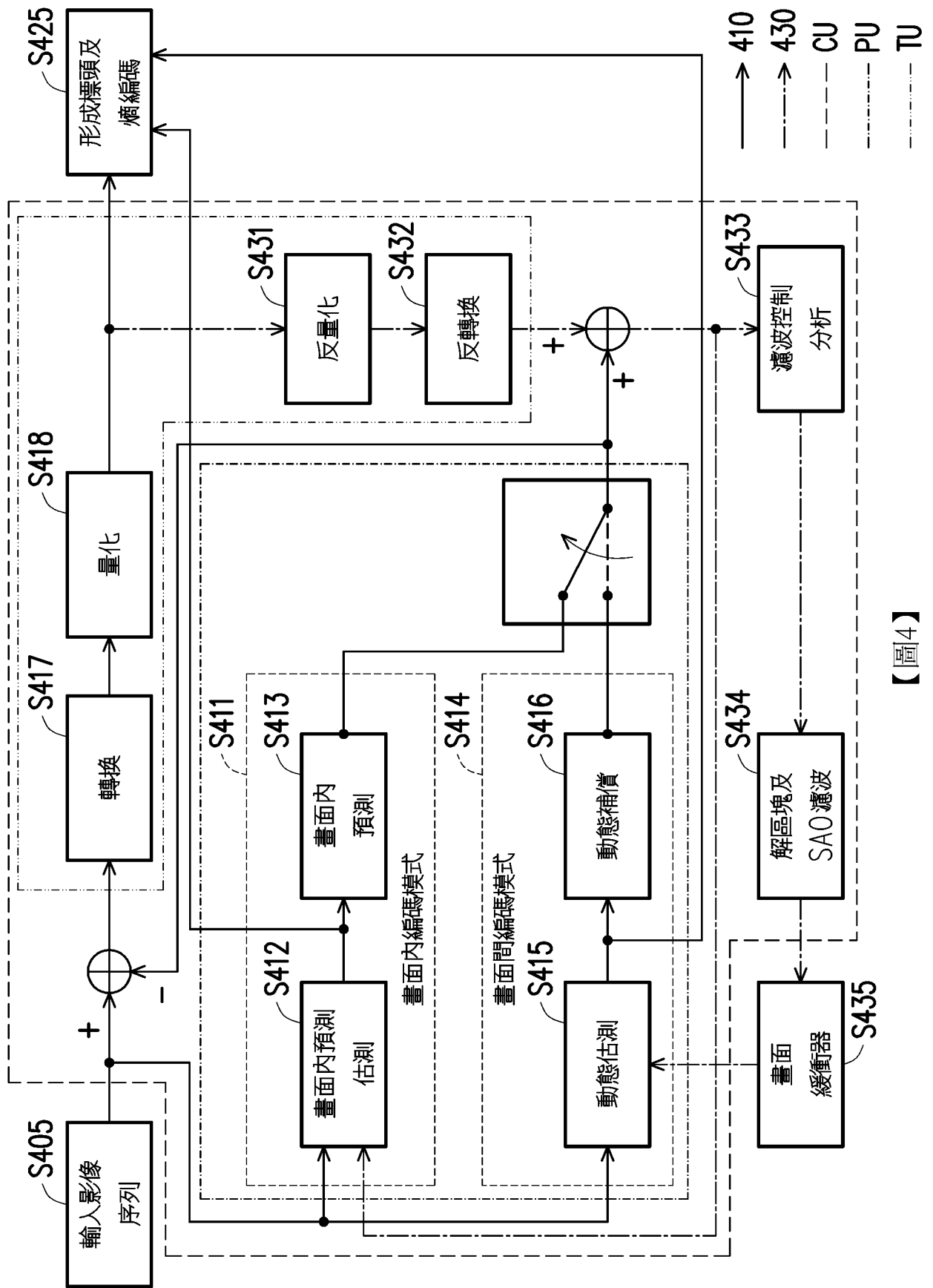
【圖1】



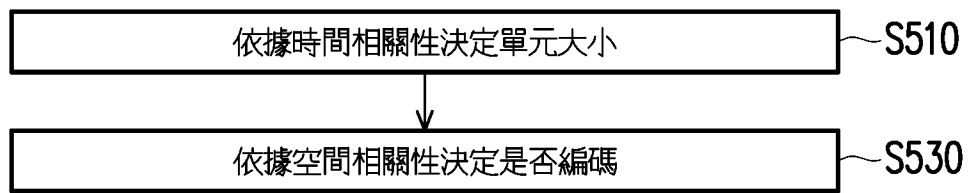
【圖2】



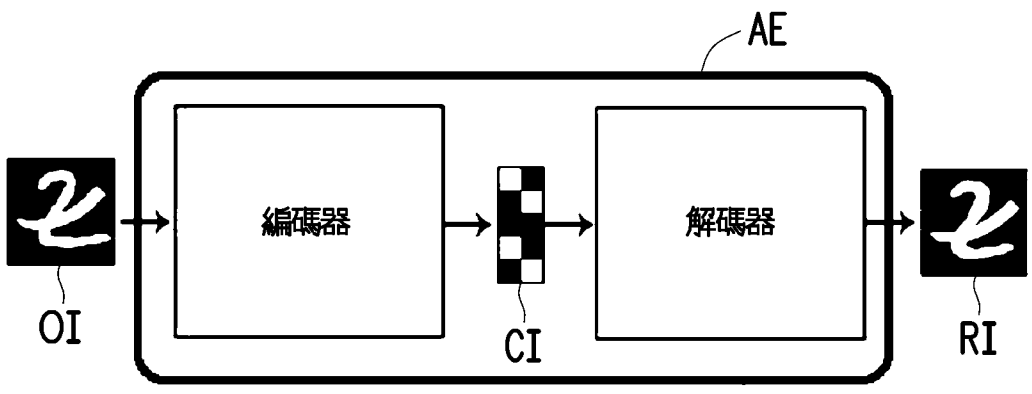
【圖3】



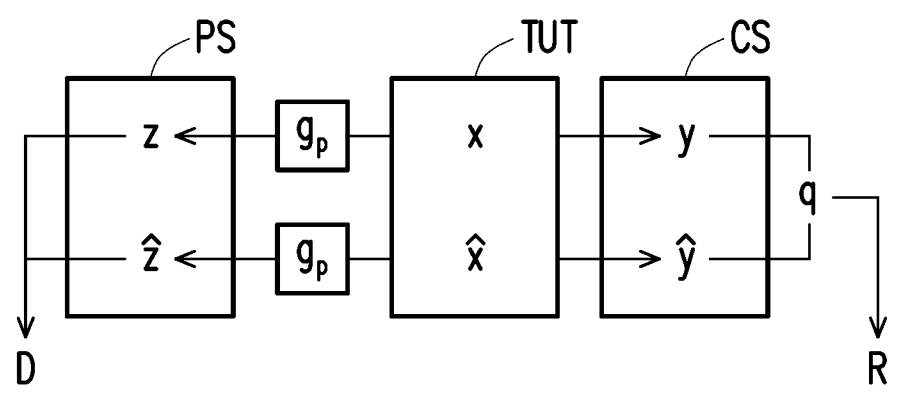
【圖4】



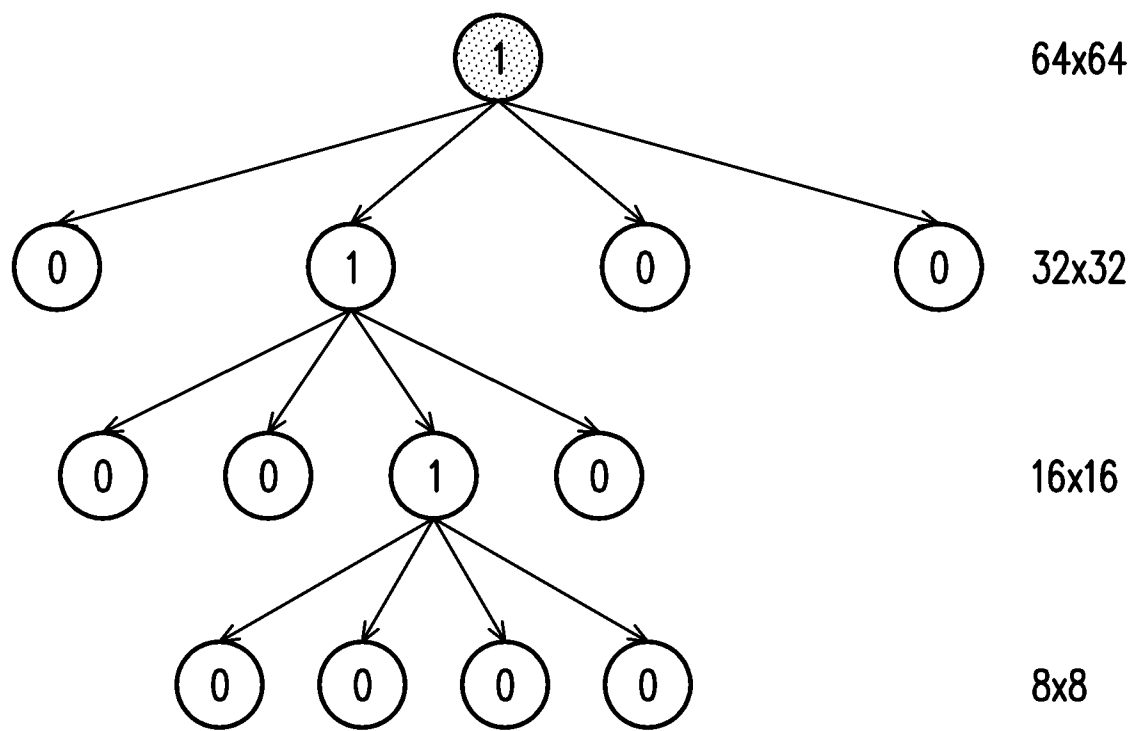
【圖5】



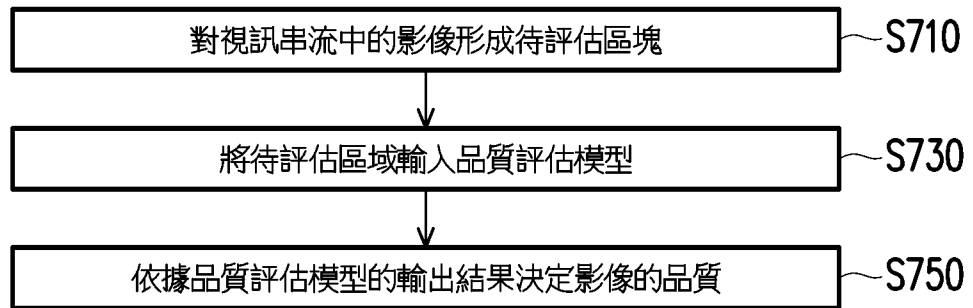
【圖6A】



【圖6B】



【圖6C】



【圖7】