



(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 201016016 A1

(43)公開日：中華民國 99 (2010) 年 04 月 16 日

---

(21)申請案號：098133797 (22)申請日：中華民國 98 (2009) 年 10 月 06 日

(51)Int. Cl. : **H04N7/26 (2006.01)**

(30)優先權：2008/10/07 美國 61/103,362

(71)申請人：歐幾里德探索公司 (美國) EUCLID DISCOVERIES, LLC (US)  
美國

(72)發明人：佩斯 查爾斯 保羅 PACE, CHARLES PAUL (US)

(74)代理人：桂齊恆；閻啟泰

申請實體審查：無 申請專利範圍項數：35 項 圖式數：8 共 78 頁

---

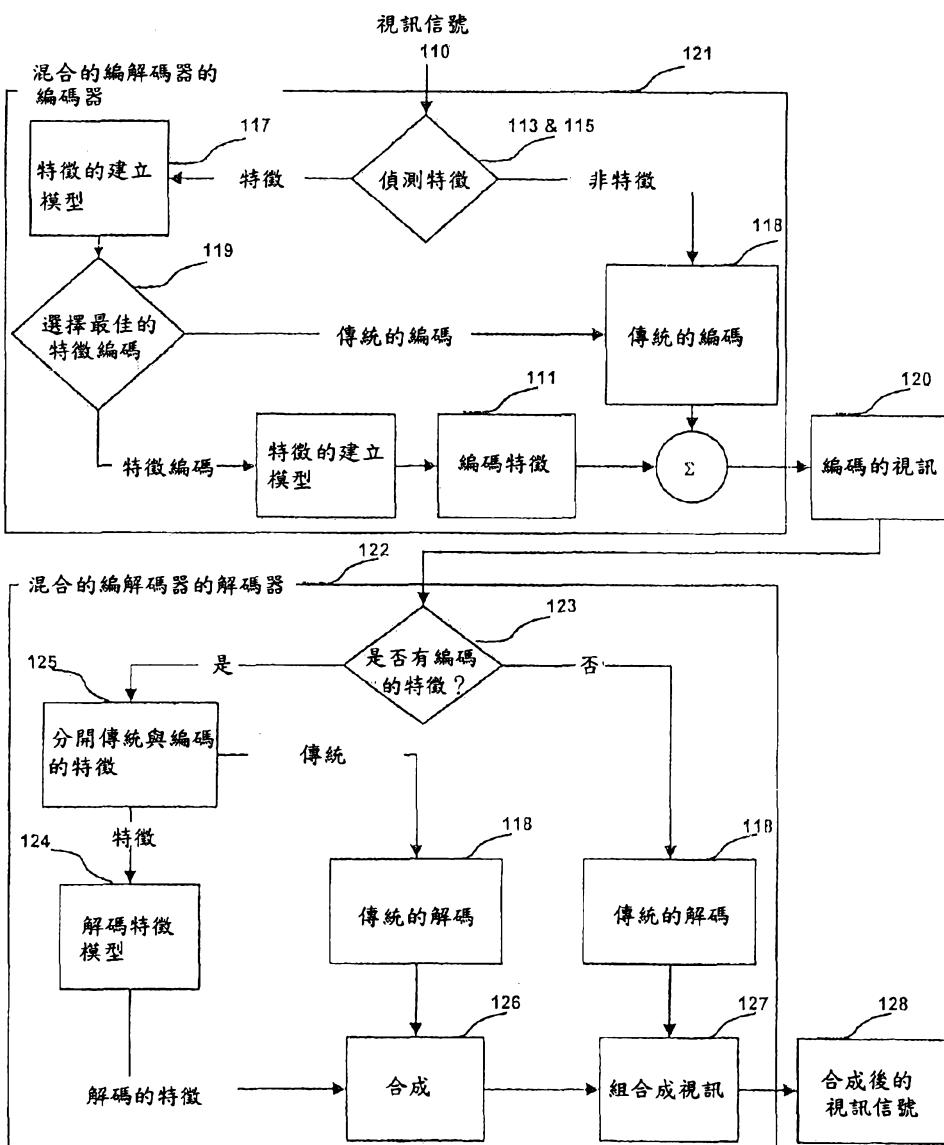
(54)名稱

基於特徵的視訊壓縮

FEATURE-BASED VIDEO COMPRESSION

(57)摘要

處理視訊資料的系統及方法係被提出。具有一系列視訊幀的視訊資料係被接收並且處理。一個候選特徵的一或多個實例係在該些視訊幀中被偵測出。藉由採用在一個運動補償的預測過程中所用的參考幀處理，該候選特徵的實例係與該些視訊幀中的非特徵切割開。該運動補償的預測過程係選出具有對應於該候選特徵的一或多個實例之特徵的先前解碼過的視訊幀。該些先前解碼過的視訊幀係被處理以識別出該候選特徵之可能的符合者。當先前解碼過的視訊幀中有大量的部份包含該候選特徵的實例時，該候選特徵的實例係被聚集成為一組。該候選特徵組係被用來產生一個基於特徵的模型。該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型。該外觀變化模型是藉由對於該候選特徵的實例的變化建立模型而被產生的。該系統判斷出與利用該基於特徵的模型以對於該候選特徵組建立模型相關連之壓縮效率的程度、以及與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連之壓縮效率的程度。該基於特徵的模型之壓縮效率係和傳統的視訊模型之壓縮效率做比較，並且該系統係判斷哪一個是具有較大的壓縮值。根據哪一個是具有較大的壓縮值，該視訊資料係利用該基於特徵的模型或傳統的視訊編碼而被編碼。



- 110：視訊信號輸入
- 111：編碼特徵之手段
- 113：偵測特徵之手段
- 115：分開特徵之手段
- 117：建立特徵模型之手段
- 118：習知的編解碼器
- 119：比較器
- 120：偏碼後的視訊
- 121：混合的編解碼器的編碼器
- 122：混合的編解碼器的編碼器
- 123：判斷之手段
- 124：解碼特徵之手段
- 125：分開之手段
- 126：合成之手段
- 127：結合之手段
- 128：合成的視訊信號



(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 201016016 A1

(43)公開日：中華民國 99 (2010) 年 04 月 16 日

(21)申請案號：098133797

(22)申請日：中華民國 98 (2009) 年 10 月 06 日

(51)Int. Cl. : **H04N7/26 (2006.01)**

(30)優先權：2008/10/07 美國 61/103,362

(71)申請人：歐幾里德探索公司 (美國) EUCLID DISCOVERIES, LLC (US)  
美國

(72)發明人：佩斯 查爾斯 保羅 PACE, CHARLES PAUL (US)

(74)代理人：桂齊恆；閻啟泰

申請實體審查：無 申請專利範圍項數：35 項 圖式數：8 共 78 頁

(54)名稱

基於特徵的視訊壓縮

FEATURE-BASED VIDEO COMPRESSION

(57)摘要

處理視訊資料的系統及方法係被提出。具有一系列視訊幀的視訊資料係被接收並且處理。一個候選特徵的一或多個實例係在該些視訊幀中被偵測出。藉由採用在一個運動補償的預測過程中所用的參考幀處理，該候選特徵的實例係與該些視訊幀中的非特徵切割開。該運動補償的預測過程係選出具有對應於該候選特徵的一或多個實例之特徵的先前解碼過的視訊幀。該些先前解碼過的視訊幀係被處理以識別出該候選特徵之可能的符合者。當先前解碼過的視訊幀中有大量的部份包含該候選特徵的實例時，該候選特徵的實例係被聚集成為一組。該候選特徵組係被用來產生一個基於特徵的模型。該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型。該外觀變化模型是藉由對於該候選特徵的實例的變化建立模型而被產生的。該系統判斷出與利用該基於特徵的模型以對於該候選特徵組建立模型相關連之壓縮效率的程度、以及與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連之壓縮效率的程度。該基於特徵的模型之壓縮效率係和傳統的視訊模型之壓縮效率做比較，並且該系統係判斷哪一個是具有較大的壓縮值。根據哪一個是具有較大的壓縮值，該視訊資料係利用該基於特徵的模型或傳統的視訊編碼而被編碼。

## 六、發明說明：

### 【發明所屬之技術領域】

#### 【先前技術】

##### 預測切割 [主要的]

例如是 MPEG-4 與 H.264 的傳統視訊壓縮係具有能力來指明一些參考幀以在運動補償的預測過程期間使用，以便於預測目前幀。這些標準典型是限制參考幀為一或多個連續的過去的幀，並且在某些情形中是限制為先前已經解碼過的幀的任何集合。通常，參考幀的數目是有限制的，並且該選擇過程在解碼過的幀的串流中可取出多久以前的幀也是有限制的。

##### 壓縮感測 (CS)

影像及視訊壓縮技術一般是嘗試利用資料中的冗餘，而容許該資料中最重要的資訊能夠用“小”數目的參數加以捕捉。“小”是相對於原始未處理的資料大小所定義的。對於一個特定的資料組而言，事先並不知道哪些參數將會是重要的。因為此緣故，習知的影像/視訊壓縮技術在選出那些將會產生最小編碼的參數之前，都會計算(或量測)相當大數目的參數。例如，JPEG 及 JPEG 2000 影像壓縮標準是根據線性轉換(典型是離散餘弦轉換 [DCT] 或是離散子波轉換 [DWT])，其係轉換影像像素成為轉換係數，產生數目上是等於原始像素數目的轉換係數。在轉換空間中，重要的係

數可接著藉由各種技術選出。一個例子是純量量化。當其推到極致而言，這等同於大小的臨界值法(thresholding)。儘管 DCT 及 DWT 可以有效率地被計算，但是在資料縮減前需要計算完整的轉換會造成效率不佳。這兩種轉換的計算所需要的量測數目是等於輸入資料大小。此習知的影像/視訊壓縮技術的特性使得它們在需要高的計算效率時的利用變成是不切實際的。

習知的壓縮容許混合來自多個幀的多個相符處以預測目前幀的區域。該混合通常是線性的、或者是相符處的一個取對數大小的線性組合。此種雙預測方法是有效的情形之一個例子是當有從一個影像至另一影像隨著時間而淡入淡出(fade)時。該淡入淡出的過程是兩個影像的線性混合，並且該過程有時可利用雙預測而有效地建立模型。再者，MPEG-2 內插的模式容許線性參數的內插，以在許多幀上合成該雙預測模型。

習知的壓縮容許指明用於目前幀的編碼的預測所能夠抓取的一或多個參考幀。儘管參考幀與目前的幀典型是時間上相鄰的，但也可以包含該組時間上相鄰的幀以外的參考幀。

相對於以習知的轉換為基礎的影像/視訊壓縮演算法，壓縮感測(CS)演算法在量測(“感測”)步驟期間是直接利用資料中大部份的冗餘。在時域、空間域、以及頻譜域中的冗餘是較高壓縮率的一項主要的貢獻因素。所有壓縮感測演算法的重要結果是一可壓縮的信號可以在相當小數目的

隨機量測且遠小於習知的壓縮演算法所需的數目下被感測出。影像於是可以在正確且可靠地被重建。在給定已知的統計特徵之下，視覺資訊的一個子集合係被用來推出資料的其它部份。

在一個特定的 CS 演算法中所需的量測之精確數目會是依據信號的類型以及從該量測(係數)重建該信號的“回復演算法”而定。請注意的是，一個 CS 演算法在某種確定程度下重建信號所需的量測數目並非直接相關於該演算法的計算複雜度。例如，一種使用 L1-最小化以回復信號的 CS 演算法類型需要相當小數目的量測，但是該 L1-最小化的演算法是非常慢的(非即時的)。因此，實際的壓縮感測演算法會尋求所需的量測數目與重建的正確性及計算的複雜度之間的平衡。相較於習知的編解碼器，CS 係提供一種完全不同的編解碼器設計的模型。

一般而言，在一個典型的 CS 演算法中有三個主要步驟：(1)產生量測矩陣  $M$ ；(2)利用該量測矩陣以進行資料的量測，亦以產生該資料的編碼著稱；以及(3)從該編碼回復原始的資料，亦以解碼步驟著稱。該回復演算法(解碼器)可能是複雜的，並且因為在接收器之計算能力的限制較少，所以整個 CS 演算法通常是以其解碼器來命名。在習知技術中有三個受關注的 CS 演算法的實際應用：正交匹配追蹤(OMP)、L1 最小化(L1M)以及鏈結追蹤(CP)。一般而言，L1M 實際上對於大多數的視訊處理應用而言是計算上過於效率差的。較有效率的 OMP 及 CP 演算法提供大致與 L1M

相同的益處，並且就此而論，它們正是可用於大部份 L1M 的應用之兩個 CS 演算法的選項。

經由逆合成 (Inverse Compositional) 演算法的影像對齊 Basri 及 Jacobs (“朗伯 (Lambertian) 反射及線性子空間”，IEEE 模式分析與機器智能匯刊，2/03)，以下稱為 LRLS，已經展示朗伯物件 (其表面在所有方向上反射光) 可藉由 LRLS “基礎影像”的一個小的 (9 維的) 線性子空間根據球面調和函數而良好地予以近似。該 LRLS 基礎影像可被想像成物件在不同的照明條件及紋理下的版本。LRLS 基礎影像因此是依據物件的結構 (透過其表面法線)、物件在其不同的反射點處之反照率 (albedo)、以及照明模型 (其係遵循朗伯餘弦定律，在方向上積分以產生球面調和函數) 而定的。在該模型的假設下，該 9D 的子空間係捕捉物件影像中超過 99% 的能量強度。外觀子空間的低維度係表示資料中有比習知的壓縮方式可得的更大的冗餘。

逆合成演算法 (IC) 是首先被提出作為用於 2D 運動估計及影像套合的 Lucas-Kanade 演算法之一種有效率的實施方式。後續的實施方式已使用該 IC 演算法以將例如是主動外觀模型及 3D 可形變模型 (3DMM) 的 3D 模型契合至臉部影像。

## 增量奇異值分解 (ISVD) 演算法的應用

一般的維度縮減技術都牽涉到利用在保範的 (norm preserving) 基礎下的線性轉換。一個 SVD 表示的縮減係指

刪除該 SVD 中的某些奇異值/奇異向量對，以產生資料的一個在計算及表現上較有效率的表示。最普遍的是，SVD 分解係藉由將低於某個臨界值的所有奇異值歸零並且刪除對應的奇異向量而有效地予以簡化。此種大小的臨界值法係產生一個具有  $r$  個奇異值( $r < N$ )之簡化的 SVD，從一個 L2-範數(norm)角度來看，其係資料矩陣  $\mathbf{D}$  之最佳的  $r$  維近似。該簡化的 SVD 係得自

$$\widehat{\mathbf{D}} = \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^T, \quad \text{方程式 1}$$

其中  $\mathbf{U}_r$  是  $M \times r$ ， $\mathbf{S}_r$  是  $r \times r$  對角線，並且  $\mathbf{V}_r$  是  $N \times r$ 。

該奇異值分解(SVD)是一個資料矩陣的分解，其係自然導致有資料的小(小型)的描述。給定一個大小  $M \times N$  的資料矩陣  $\mathbf{D}$ ，該 SVD 分解係由  $\mathbf{D} = \mathbf{U}^* \mathbf{S}^* \mathbf{V}'$  得出，其中  $\mathbf{U}$  是(左)奇異向量的一個  $M \times N$  行正交矩陣， $\mathbf{S}$  是一個具有沿著對角線的奇異值( $s_1, s_2, \dots, s_N$ )之  $N \times N$  對角線矩陣，並且  $\mathbf{V}$  是(右)奇異向量的一個  $N \times N$  正交矩陣。

### 緊緻分歧(compact manifold)預測

匹配追蹤(MP)是一種用於導出有效率的信號表示的遞迴演算法。給定用基礎函數(不必然是正交的)的函數庫  $\mathbf{D}$  來代表一個信號向量  $\mathbf{s}$  的問題，MP 係經由在此所述的遞迴過程來選擇用於該表示的函數。在該表示中的第一基礎函數(標示為  $\mathbf{d1}$ )係被選為與信號向量有最大關聯的函數。接著，一個殘餘向量  $\mathbf{r1}$  係藉由將  $\mathbf{d1}$  到該信號上的投射從該信號本身減去而計算出： $\mathbf{r1} = \mathbf{s} - (\mathbf{d1}' * \mathbf{s}) * \mathbf{d1}$ 。接著，在該表示( $\mathbf{d2}$ )

中的下一個函數係被選為與該殘餘的  $r_1$  有最大關聯的函數。 $d_2$  到  $r_1$  上的投射係從  $r_1$  減去以形成另一殘餘  $r_2$ 。相同的過程接著被重複，直到該殘餘的範數下降到低於一特定的臨界值為止。

正交匹配追蹤(OMP)係依循與 MP 相同的遞迴程序，除了會採取一個額外的步驟以確定該殘餘與已經在該表示的集團(ensemble)中的每個函數都是正交的。儘管該 OMP 遞迴是比 MP 中的複雜，但是其額外的計算確保 OMP 在不超過  $Nd$  個步驟下收斂至一個解，其中  $Nd$  是函數庫 D 中的函數數目。

## 【發明內容】

本發明係延伸傳統的視訊壓縮，尤其是其中視覺現象的冗餘超出習知的視訊編解碼器的建立模型功能的情形。本發明係延伸並且可以完全取代傳統的視訊壓縮之現有的方法，其藉由採用強健的電腦視覺及圖形識別演算法。明確地說，本發明係包含特徵的建立模型方法及系統，其專注於出現在視訊的先前解碼過的幀的一或多個幀中的特徵的切割、正規化(normalization)以及整合。相較於習知的壓縮僅考量較少的幀、較小的區域、以及較少的像素，基於特徵的視訊壓縮係考量更多數量的先前解碼過的幀，並且在那些幀的每個幀內考量更大的區域以及更多數目的像素。

習知的壓縮係藉由利用多個參考幀、大區塊

(macroblock)分割、次大區塊分割、以及運動補償的預測以提供在大區塊層級下的一種隱含的切割形式。再者，習知的壓縮係利用運動補償的預測以對於發生在視訊中的空間變形建立模型，並且轉換編碼以對於外觀變化建立模型。本發明係將不同的信號元素的這些建立模型的技術延伸到更複雜的模型，其包含空間切割遮罩、規律網格變形、特徵仿射(affine)運動、三維的特徵運動、三維的照明、以及其他其他的電腦視覺及圖形識別之建立模型的技術。請注意的是，在本文中，“個別的模式”以及“不同的信號元素”是等同的。

本發明係促進視訊信號之個別的模式的識別及切割。在習知的運動補償的預測中所用的參考幀處理的概念係被利用在本發明中以助於此識別及切割。該習知的運動補償的預測過程係在大區塊層級下從一或多個參考幀選擇信號的部份。請注意的是，該習知的運動補償的預測過程通常是根據某種率-失真(rate-distortion)尺度來進行此種選擇。本發明能夠對於過去的幀施以分析，以判斷出將會具有最高機率能夠提供與目前幀的相符處之幀。此外，參考幀的數目可以是遠大於習知的壓縮中典型可見的一至十六個參考幀最大值。根據系統資源而定，參考幀數目可以盡量的高，而達到系統記憶體的極限為止；此係假設在那些幀中有足夠數目之有用的相符處。再者，由本發明產生的資料的中間形式可以減少儲存同樣數目的參考幀所需要的記憶體量。

在一個實施例中，本發明係根據此種參考幀處理來推斷出該視訊信號的切割。在目前幀中的大區塊(像素的區塊)可以透過該運動補償的預測過程，從先前解碼過的幀選擇像素的影磚(tile)，使得該些影磚在空間及時間上都是分開的，此表示用在該運動補償的預測過程的影磚來源可以是來自不同的幀。從不同的幀選擇來源影磚以用於預測目前幀的大區塊所隱涵的分開是指出不同的信號模式被識別出的可能性。當識別出的個別的信號模式可用一種更小的方式編碼時，此進一步證實個別的模式已經被識別出。在本發明中，這些個別的模式係被稱為“特徵(feature)”。當這些特徵在視訊的許多幀上是持續不斷的並且該些特徵可以相關聯時，在該視訊中一種新類型的冗餘已經被識別出。本發明透過外觀及變形模型的產生來利用此冗餘，以便於產生超過習知壓縮所可得的進一步壓縮。再者，由於特徵是在參考幀內被識別出，因此參考幀處理會受影響而朝向利用包含特徵的參考幀；此係導致提高該參考幀處理將會進一步產生該信號中存在的模式切割之可能性。

可提供系統及方法以用於處理視訊資料。由一系列的視訊幀所構成的視訊資料可被接收及編碼。一個候選特徵的一或多個實例可在一或多個視訊幀中被偵測出。候選特徵的偵測係牽涉到判斷實例在該一或多個先前解碼過的視訊幀中之位置資訊。該位置資訊係包含一個幀號碼、在該幀中的一個位置、以及該實例的一個空間的周圍。該候選特徵可以是一組一或多個偵測出的實例。一個運動補償的

預測過程可被用以利用一或多個先前解碼過的視訊幀來預測在該系列中的一個目前視訊幀的一部份。該運動補償的預測過程可用位置的預測而被初始化。該位置的預測可從在先前解碼過的視訊幀中之偵測出的特徵實例來提供位置資訊。該些實例中的一或多個實例可藉由增加該運動補償的預測過程而被轉換。一個特徵以及該轉換後的實例可被界定出來。該些實例中的該一或多個實例可利用一種線性轉換而被轉換。包含該些轉換後的實例之界定出的特徵可被利用以產生一個第一基於特徵的(feature-based)模型。該第一基於特徵的模型可以使得在該目前幀中能夠預測一個實質相符的特徵實例之出現以及一個來源位置。較佳的是，該實質相符的特徵是利用一個率-失真的尺度判斷出之最佳的相符處。該實質相符的特徵實例可以是一個關鍵的特徵實例。該關鍵的特徵實例可以是該第一基於特徵的模型之目前幀的特徵實例接合(synthesis)。該第一基於特徵的模型可以和該一或多個界定出的特徵之一習知的視訊編碼模型相比較，並且該比較可被利用來決定哪一個模型能夠達成較大的編碼壓縮。該比較及判斷步驟的結果可被利用來導引該編碼過程來施加基於特徵的編碼至該些視訊幀中的一或多個視訊幀的部份，並且施加習知的視訊編碼至該一或多個視訊幀的其它部份。

一個候選特徵的一個實例可藉由識別出具有實質接近且空間上鄰近的像素的一個空間上連續的群組而被偵測出。該些識別出的像素可被用來定義該一或多個視訊幀中

的一個視訊幀的一部份。該群組的像素可包含一或多個大區塊、或是一或多個大區塊的部份。

該運動補償的預測過程可被用來從複數個候選特徵實例中選擇一或多個被預測可提供編碼效能的實例。該候選特徵的目前的實例之切割可從該目前視訊幀中之其它特徵及非特徵來決定之。該切割可以是根據該運動補償的預測過程從獨特的先前解碼過的視訊幀所得的預測之選擇而定。該運動補償的預測過程可利用屬於一或多個特徵(此種特徵在該目前幀中具有與該視訊部份符合的實例)的特徵實例之位置資訊而被初始化，其中該視訊部份是在該目前幀中，並且該位置資訊係對應於與在先前解碼過的視訊幀中的相同特徵相關連的特徵實例。

一個第二基於特徵的模型可被形成。該第二基於特徵的模型的形成可利用該第一基於特徵的模型作為從一或多個特徵實例得到一或多個運動補償的預測之一個預測的目標。此第二基於特徵的模型產生該第一基於特徵的模型的一組預測。一旦該組預測與該第一基於特徵的模型結合，則該組預測可以成為該第二基於特徵的模型。該第二基於特徵的模型可被利用以對於來自第一基於特徵的模型之殘餘(residual)建立模型。結構的變化及外觀的變化可以從該第二基於特徵的模型相對於該殘餘來建立模型。該殘餘可利用該特徵實例來加以編碼，此係產生外觀及變形的參數。該些參數可被利用以縮小該殘餘的編碼大小。

一或多個特徵可包含一或多個聚集的(aggregate)特

徵。該些聚集的特徵是根據該候選特徵的實例中之一或多個實例而定的。該些聚集的特徵可藉由聚集不同的候選特徵的實例成為一個聚集的候選特徵而被產生。該些聚集的候選特徵的實例組可被利用以形成一個實質大於未聚集的候選特徵之原始實例的區域。該較大的區域可透過在該組中該候選特徵的實例之間的連貫性(coherency)的識別而被形成。連貫性可被定義為藉由一個較少的參數運動模型來實質近似的該些實例中的外觀一致。該第二基於特徵的模型可提供一個與該解碼過的幀中的該實例相對於該空間的位置相關連之選配的矩形像素區域範圍。該第二基於特徵的模型可藉由對於該特徵之先前正規化的實例建立模型而被導出。該些先前正規化的實例可以是下列的任一個：在該目前幀中的實例；來自一個時間上實質最近之先前解碼過的幀的一個實例；或是來自該些先前解碼過的視訊幀的實例之平均。

該外觀模型可藉由該些正規化的第二基於特徵的模型實例的 PCA 分解來表示。一個變形模型可利用每組的特徵實例中之一致處相較於其第二基於特徵的模型實例的空間變化而被決定。對於在該組中的每個特徵實例而言，下列的一或多個可被利用來對於該變形模型近似該些變形實例中的變化：一個運動補償的預測過程；網格變形；以及一個具有實質減少參數化的運動模型。該些變形實例可被整合成為該變形模型。在該變形模型中的變化可藉由 PCA 分解來表示。

外觀參數及變形參數可被預測。該些預測出的參數可在該目前的實例利用一個基於特徵的模型的接合期間被利用。該外觀及變形模型與時間上最近的參數可被利用來從該基於特徵的模型內插及外推(extrapolate)參數，以預測該目前幀中的像素。該些時間上最近的特徵實例之接合的值可根據已產生那些實例最正確的近似之方法而被線性內插或是線性外推。該模型的實際參數可以是選配地相對於該些預測出的參數來做不同的編碼。

相較於習知的視訊資料編碼，該運動補償的預測過程可以在實質較大數目的先前解碼過的視訊幀之一個選集上做運算。先前解碼過的視訊幀的選集並不需要依賴使用者的監督。

習知的視訊編碼可增加一個實例預測過程，其係在形成該目前幀的部份的預測時，在記憶體中致能對於該些視訊幀中的一或多個視訊幀的部份之更大的壓縮。該實例預測過程可使用該基於特徵的模型以判斷該界定出的特徵出現在一個正被編碼的目標大區塊的一或多個實例。以此種方式，該實例預測過程可以產生該目前幀之預測出的部份。該基於特徵的模型可被利用以合成像素來預測該目前幀的部份。

可對於該些先前解碼過的視訊幀指定一個機率。該機率可以是根據針對該幀之組合預測出的編碼效能改善而定的，而該效能改善則是利用來自該運動補償的預測過程之位置的預測來加以決定的。該機率可被定義為運動補償的

預測過程之組合的編碼效能，其係在針對該目前幀分析該第一基於特徵的模型以及一個第二基於特徵的模型期間被利用。一項從最佳到最差的根據分類該些先前解碼過的視訊幀而定之指標可依據其機率而被產生。該分類後的表列可根據計算及記憶體的條件而被截短。

一個基於特徵的模型可利用該些界定出的特徵中之一或多個而被形成。該基於特徵的模型可包含該些界定出的特徵之位置資訊。該位置資訊可包含來自該些先前解碼過的視訊幀之界定出的特徵的位置及空間的周圍。例如，該位置資訊可包含有關在一個特定幀之內的區域的空間的位置、以及該區域在該幀中的一個矩形範圍的資訊。該基於特徵的模型可以指明哪些先前解碼過的視訊幀(或是其部份)是與該界定出的特徵相關連。

該些界定出的特徵可從該視訊資料利用大區塊運動補償的預測而被正規化並且切割出。該些界定出的特徵可利用該基於特徵的模型而被正規化。該大區塊運動補償的預測可使用在該先前解碼過的影像幀中之特徵位置作為一個位置的預測。該所產生的正規化係提供該特徵在該目前視訊幀中的預測。

該基於特徵的模型可以和產生自相同的視訊資料之習知編碼的另一模型相比較。該比較可被利用以判斷哪個模型能夠得到較大的編碼壓縮效率。不同的編碼技術可根據該編碼比較的結果而被應用到該視訊資料不同的部份。以此種方式可提供差異性的編碼，以使得該系統能夠對於視

訊資料的每個部份都根據基於特徵的編碼或是習知為主的編碼中的何者是提供較大壓縮效率來選擇不同的視訊編碼方式。

一個界定出的特徵可被表示為該特徵在一或多個視訊幀中的一組實例。每個實例可包含：指向一個其中出現該實例的幀之一個參照；一個與該實例在該幀中相關連之空間的位置；以及一個選配的與該實例在該幀中相關於該空間的位置相關連之矩形像素區域範圍。該空間的位置可以提供用於編碼該些視訊幀中的一或多個視訊幀的部份之相符處的預測。可對於每個界定出的特徵都提供一個外觀模型，以對於該界定出的特徵在該組中的不同實例之間的變化建立模型。該外觀模型可藉由對於該特徵之先前正規化的實例建立模型而被導出。該先前正規化的實例可利用運動補償的預測過程、網格變形、以及參數減少的運動模型建立(例如仿射)的任意組合而被正規化。

該正規化可被利用以建立一個變形模型，該變形模型可被利用以對於在每組的特徵實例中的一致處之空間變化建立模型。對於在該組中的每個特徵實例而言，下列的一或多個可被利用以判斷出該變形模型的變形實例：一個運動補償的預測過程、網格變形、以及參數減少的運動模型建立。該些變形實例可被整合成為該變形模型。該變形模型可藉由利用主要成份分析(PCA)的分解而被表示。該變形模型可藉由利用任意分解的演算法之分解而被表示。相較於習知的視訊資料編碼，該運動補償預測過程可以在不需

監督下，在實質較大數目的先前解碼過的視訊幀上運算。

該習知的視訊編碼可包含運動補償的區塊為主的壓縮。該習知的視訊編碼可增加一個殘餘縮減過程，其係在形成一個殘餘幀時，能夠達到在記憶體中該些視訊幀的部份的更大壓縮。該殘餘縮減過程可包含該基於特徵的模型以判斷該界定出的特徵出現在一個正被編碼的目標大區塊中的一或多個實例，以形成該殘餘幀。像素可利用該基於特徵的模型而被合成，以預測該殘餘幠。該基於特徵的模型可被利用於參考幠的索引(index)預測。該些合成後的像素可響應於判斷出該界定出的特徵的一或多個實例在該目前幠中重疊超過一個大區塊，而被再利用於其它殘餘縮減。當該界定出的特徵的一或多個實例實質相符在該目前幠中的一個大區塊之位置資訊時，該些合成後的像素可響應於判斷出該界定出的特徵的一或多個實例係代表一個大區塊，而被再利用於其它殘餘縮減。外觀及變形可根據該基於特徵的模型來加以建立模型。該外觀模型及變形模型以及在該些模型中的一組歷史的參數可被利用以從該基於特徵的模型內插及外推參數，以預測在該目前幠中的像素。再者，更高階的二次以及甚至擴張型卡爾曼(Kalman)濾波器模型都可被利用以預測該些外觀及變形參數。從該基於特徵的模型預測該些參數係能夠達成在該些殘餘參數大小上的縮減，產生較低的精確度，且因此對於預測在該目前幠中的像素所需的參數產生較低的位元率表示。

來自一或多個幠的一或多個大區塊可利用該運動補償

的預測過程來加以選擇。在一個 PCA 模型中，來自大區塊的像素可以是線性組合的像素，並且該些 PCA 模型參數可被內差得出。任何分解的演算法也可等同地被利用來取代 PCA，並且根據其實質相對的益處而被利用。

實質很小的空間區域可在該些視訊幀中被識別出。連貫性的標準可被利用以識別出可被結合成為實質較大的空間區域之空間區域。對於一個較大的空間區域而言，該較大的空間區域成為一個界定出的特徵之合適性可藉由編碼該較大的空間區域的一個基於特徵的模型來加以判斷。該較小區域可以是一個界定出的特徵，而且該較大的區域也可以是一個界定出的特徵。

基於特徵的壓縮可包含物件為主的壓縮過程。物件為主的偵測、追蹤及切割都可在該目前幀中或是在先前解碼過的幀中被應用到一個特徵實例。該特徵實例的一個中間形式可利用空間的切割而被導出。例如，該空間的切割過程可從非物件的背景切割出一個前景物件。當物件存在於一個幀中，到其出現在下一個幀中時，該所產生的切割可提供在一個特徵實例中的一個特定的物件之像素層級的一致。與該物件相關連的像素資料被再取樣，並且接著該被再取樣出的像素資料的空間位置係利用模型而被回復。該再取樣係有效地正規化從一個幀到下一個幀的物件像素資料，並且產生提供一種中間形式的視訊資料，其具有在視訊處理上計算及分析的優點。以此種方式，物件為主的正規化以及建立模型的過程可在該基於特徵的編碼過程期間

被應用到在該目前幀中或是在先前解碼過的幀中之一個特徵實例(或是其部份)。一致處的建立模型、變形的建立模型、外觀的建立模型、輪廓的建立模型、以及結構的建立模型都可被利用來對於在該目前幀中或是在先前解碼過的幀中的一個特徵實例(或其部份)建立模型。

一個界定出的特徵可能會沒有與顯著的實體(物件、子物件)一致處。例如，該些顯著的實體可透過對於偵測出的特徵監督標示為屬於或是不屬於一個物件而被判斷出。該些界定出的特徵可包含兩個或多個顯著的物件的元素、背景、或是該些視訊幀的其它部份。一或多個特徵可構成一個物件。再者，一個界定出的特徵可能不會對應於一個物件。一個界定出的特徵可能不內含在任何物件中。以此種方式，基於特徵的壓縮可以比物件為主的偵測更為有彈性及變化。儘管界定出的特徵可包含物件且內含在物件中，但是界定出的特徵並不必要是物件為主的，而是可採任意形式。

在另一實施例中，壓縮感測(CS)係被應用至該基於特徵的編碼技術。CS 係被應用至該些具有有效的或是界定出的特徵之視訊幀中的像素。CS 亦可被應用到對於該些視訊幀的其它像素的習知編碼。該視訊資料可被做成是稀疏的，以增加 CS 應用的有效性。在模型的形成(外觀及變形模型)期間，CS 可被應用以解析來自部分參數量測的模型參數。

CS 可被應用到該第二基於特徵的模型預測的殘餘。CS 的應用可利用該平均外觀作為一項量測，並且從該量測預

測該視訊信號。與該 CS 預測相關連的變異可從該第二基於特徵的模型被移除。該基於特徵的模型可被利用來專注於其餘的部份之更小的編碼。CS 編碼可被應用到該一或多個視訊幀中之其餘的像素且應用到其餘的視訊幀。

一個混合的編解碼器的解碼器可被設置，其係使用基於特徵的解壓縮以用於解碼視訊資料。編碼的視訊資料可藉由在大區塊層級判斷在該編碼的視訊資料中是否有一個編碼的特徵而被解碼。該編碼的特徵可包含基於特徵的模型。於不存在編碼的特徵之情形中，在該編碼的視訊資料中之大區塊可利用習知的視訊解壓縮而被解碼。在一個編碼的特徵確實存在之情形中，該解碼器可藉由從該編碼的視訊資料分出該些特徵編碼的部份來響應在該編碼的視訊資料中偵測到一個編碼的特徵。藉由分出該些特徵編碼的部份，該系統係能夠與該視訊流中傳統編碼的部份分開地合成該編碼的特徵。來自該編碼的特徵部份之特徵參數可以與該編碼特徵內含的特徵模型相關連。該些特徵參數可被該解碼器利用以合成該編碼的特徵。該視訊資料之傳統壓縮的部份可以與該合成後的特徵結合，以重建該原始的視訊幀。

在另一實施例中，一個視訊編解碼器係能夠處理複數個壓縮的視訊信號模式。在其中一個視訊信號模式中，一個編解碼器的編碼器係提供基於特徵的視訊壓縮。在另一模式中，該編解碼器的編碼器係提供傳統的視訊壓縮。類似地，一個編解碼器的解碼器係響應於不同的視訊信號模

式且能夠根據該視訊信號的內容(例如該視訊信號模式)提供基於特徵的視訊解壓縮以及傳統的視訊壓縮。

該編解碼器可根據基於特徵的編碼或是習知為主的編碼中的哪一種會提供該視訊信號的視訊幀中之一或多個特徵較大的壓縮效率來決定哪一種類型的視訊壓縮是適當的。

## 【實施方式】

### 序言段落

本發明的範例實施例的說明係如下。

所有在此引用的專利、公開的申請案以及參考文獻的教示都以其整體而被納入作為參考。

### 數位處理環境及網路

較佳的是，本發明是實施在一個軟體或硬體環境中。一個此類的環境係被展示在圖 3 中，其描繪一個其中本發明可被實施的電腦網路或是類似的數位處理環境。

客戶電腦/裝置 350 以及伺服器電腦 360 係提供執行應用程式與類似者的處理、儲存以及輸入/輸出裝置。客戶電腦/裝置 350 亦可以透過通訊網路 370 連結至其它計算裝置，其包含其它的客戶裝置/處理 350 以及伺服器電腦 360。通訊網路 370 可以是一個遠端存取的網路、一個全球的網路(例如，網際網路)、一個遍及全球的電腦集合、本地區域或廣域網路、以及目前使用個別的協定(TCP/IP、藍芽、等

等)之閘道的部份，以彼此通訊。其它的電子裝置/電腦網路架構也是適當的。

圖 4 是圖 3 的電腦系統中之一部電腦(例如，客戶處理器/裝置 350 或伺服器電腦 360)的內部結構圖。每個電腦 350、360 係包含一個系統匯流排 479，其中匯流排是一組實際或虛擬的硬體線用於在電腦或處理系統的構件之間做資料傳輸。匯流排 479 實質是一個共用的通道，其連接電腦系統之不同的元件(例如，處理器、碟片儲存、記憶體、輸入/輸出埠、等等)，而能夠達成元件間的資訊傳輸。附接至系統匯流排 479 是用於連接各種輸入與輸出裝置(例如，鍵盤、滑鼠、顯示器、印表機、揚聲器、等等)至電腦 350、360 的 I/O 裝置介面 482。網路介面 486 係容許該電腦能夠連接至各種其它連接到一個網路(例如在圖 3 的 370 處所繪的網路)的裝置。記憶體 490 係提供被用來實施本發明的一個實施例(例如，以上詳述之混合的編解碼器、視訊編碼器壓縮碼、以及解碼器的碼/程式的常式(routine))之電腦軟體指令 492 及資料 494 之揮發性的儲存。碟片儲存 495 係提供被用來實施本發明的一個實施例之電腦軟體指令 492(等同的是“OS 程式”)及資料 494 之非揮發性的儲存。中央處理器單元 484 亦附接至系統匯流排 479，並且設置用於電腦指令的執行。請注意的是，在本文中，“電腦軟體指令”以及“OS 程式”是等同的。

在一個實施例中，該處理器常式 492 及資料 494 是一個電腦程式產品(概括以 492 來參照)，其包含一電腦可讀取

的媒體(例如，一可移除的儲存媒體，例如是一或多個DVD-ROM、CD-ROM、磁碟、磁帶、等等)，其提供用於本發明系統的軟體指令之至少一部份。電腦程式產品 492 可藉由任意適當的軟體安裝程序，即如同此項技術中眾所週知的安裝程序來加以安裝。在另一實施例中，該些軟體指令的至少一部份也可透過電纜線、通訊及/或無線的連線來下載。在其它實施例中，本發明的程式是一個電腦程式傳播的信號產品 307，其係體現在一傳播媒體上的一種傳播的信號(例如，在像是網際網路的全球網路或其它網路上傳播的無線電波、紅外線波、雷射波、聲波或電波)。此種載波媒體或信號係提供用於本發明常式/程式 492 的軟體指令之至少一部份。

在替代的實施例中，該傳播的信號是在該傳播的媒體上承載之一類比載波或數位信號。例如，該傳播的信號可以是在一個全球的網路(例如，網際網路)、一個電訊網路、或是其它網路上傳播的一數位化的信號。在一個實施例中，該傳播的信號係在該傳播媒體上被傳送一段期間，例如，一個軟體應用程式的指令是以封包在一個網路上被傳送一段長為毫秒、秒、分鐘、或是更長的期間。在另一實施例中，電腦程式產品 492 的電腦可讀取的媒體是一個電腦系統 350 可接收及讀取的傳播媒體，例如，藉由接收該傳播媒體且識別一體現在該傳播媒體中之傳播的信號，即如上述用於電腦程式傳播的信號產品。

大致來說，該用語“載波媒體”或是暫態載波係包含以上

的暫態信號、傳播的信號、傳播的媒體、儲存媒體與類似者。

### 概要 - 基於特徵的視訊壓縮

本發明係提供一種混合的(基於特徵的及習知的)編解碼器方法(圖 1)，其具有偵測(113)、分開(115)、建立模型(117)、編碼(111)、以及解碼(124)視訊中的特徵之手段，同時容許習知的編解碼器 118 能夠編碼及解碼非特徵與無法有利地透過該特徵的編碼器/解碼器處理的特徵。圖 1 係描繪一主題視訊信號輸入(由一系列的影像幀所構成的視訊資料)110 係藉由本發明混合的編解碼器 121 而被編碼。該混合的編解碼器係包含該編碼決策的啟發法(*heuristics*)，並且如下處理該視訊信號：在步驟 113，特徵的偵測主要是透過在附近且呈現複雜度的像素群組的識別而達成的。複雜度通常是定義為任何能夠指出像素的編碼是超出原本藉由傳統的視訊壓縮會有效率地編碼的程度之尺度。此種在附近的像素分組係提供偵測出的特徵(在 115 處)從背景及其它的特徵切割出來。該分組係接著被分析以判斷該複雜度是否可以有利地利用本發明的特徵模型來加以建立模型(117)。

一旦特徵被偵測出且加以追蹤，並且該些特徵的模型被產生(在 117 處)之後，該特徵的建立模型以及習知的建立模型係被比較(在比較器 119)以判斷哪一個是具有較大的益處。採用目前幀之運動補償的預測所用的參考幀處理之習

知的視訊編碼機構(在 118 處)係在此過程中被利用。因為該比較(119)已經採用習知的參考幀處理，目前幀的切割係(根據來自不同的參考幀的預測選擇)被產生。在一個參考幀相對於另一參考幀中的像素(更典型為大區塊)之選擇係指出在該幀中以及在該主題視訊 110 本身中的該些特徵的切割。該主題視訊信號輸入 110 之產生的編碼 120 係包含一習知的視訊編碼串流(習知的編碼器 118 的輸出)伴隨著在該些參考幀中再現該些特徵所需之額外的編碼資訊。

該混合的編解碼器的解碼器 122 係描繪解碼該編碼的視訊以便於合成(近似)該輸入視訊信號 110。當檢查內含在該編碼的視訊中之串流資訊時，該混合的編解碼器的解碼器係在子幀的層級、亦即大區塊層級做一項判斷(123)，在該編碼的視訊中是否有一個編碼的特徵。若不存在編碼的特徵，則該習知的大區塊或是非特徵的大區塊係按傳統方式被解碼。若在該編碼的視訊串流中遇到一個編碼的特徵，則該混合的編解碼器係分開(125)該些特徵編碼的部份與傳統編碼的部份，以便於分開個別地合成，在發生接合之後組合該些部份。該混合的編解碼器係使用該編碼的特徵參數以及已由該解碼器所產生之特徵模型(與該編碼器中所做的模型一樣且平行所做的模型)以合成該特徵(124)。接著該些傳統編碼的特徵部份以及該些特徵編碼的部份係被合成(126)以產生一個完整的特徵接合。接著該混合的編解碼器係在 127 處結合該特徵接合與該非特徵接合以產生一個完整合成的視訊信號(128)。

圖 7 是本發明的一個實施例的表示圖，其係利用基於特徵的編碼作為習知編碼的部份替代，並且在某些情形中是全部替代。偵測(710)、追蹤(720)、比較(730)、建立模型(740)、編碼(750)、以及解碼(760)一視訊信號中的特徵係被描繪。

在步驟 710，如同在 113 中，特徵的偵測主要是透過呈現複雜度之空間接近的像素群組的識別來達成的，使得像素群組可以比習知的方式更有效率地編碼/建立模式。這些像素群組有效地分開偵測出的特徵(710)與它附近之非特徵的像素，如同亦在 115 中指出者。偵測出的特徵、或是特徵實例候選、或就稱做是特徵候選係進一步被分析，以關聯兩個或多個幀上的像素群組。此關聯確認特徵實例是代表該些視訊幀中的一個可被追蹤的離散實體(720)，藉此確認該視訊中有額外的冗餘，該冗餘可透過對該特徵建立模型而可被降低(740)。在步驟 720 中，該特徵係經由在目前幀內的特徵的實例(等同於區域)以及該特徵在一或多個其它幀中的實例的識別而被追蹤，如同亦在 117 中指出者。請注意的是，在整個本文中，“特徵實例”係等同於“區域”。再者，當“實例”是參照“特徵實例”及“區域”時，“實例”係等同於“特徵實例”及“區域”。

每一個別特徵的實例都被視為候選特徵，並且透過將它們群組成為特徵組或就是特徵，而變為組成一個完全合格的特徵。在步驟 730 中，這些實例被分析、比較並且透過在該些實例之間的一致處的識別而被分類成為特徵組。

在本文中，特徵候選以及特徵實例是等同的。特徵組係被分析以獲得特徵實例的一個變形變化及外觀變化的模型。在特徵實例之間的變形變化係透過一個變形的建立模型過程而被判斷出。該變形的建立模型過程係比較兩個或多個實例，以決定減少實例間的每個像素的差異所將會需要的空間像素的再取樣。

特徵候選係在步驟 740 中被建立模型，其係應用多個分析技術以精確調整該些取樣出的區域。視訊串流的特徵編碼(750)係利用該些特徵模型，並且部份或全部在不使用習知的視訊編碼之下編碼該視訊串流。該解碼(760)係利用該些特徵模型以建立模型的動作之倒轉(750)來合成該些特徵以解碼該些編碼的特徵成為每個特徵實例的像素的接合(770)，近似出該特徵為如同原先在該視訊中出現者。

### 預測切割 [主要的]

圖 6 展示藉由利用內含在置於一或多個參考幀中的一或多個過去的幀之內的資訊來預測在目前視訊幀中的元素的過程。在一個實施例中，該預測(方法 1 640)係複製來自一或多個先前解碼過的幀 610 的區域到一個參考幀 620 之中。方法 2 650 額外將特徵區域 630-1、630-2、…630-n 所構成的特徵實例 660 置於該參考幀中。直接插入特徵實例到參考幀中係代表本發明的一種簡單形式，其中，在另一個實施例中，該切割就是一個矩形區域，並且該特徵的模型是該特徵實例。額外的壓縮利益可被實現，因為進一步

的建立模型技術係被應用至該些識別出的特徵 660 並且用在該些參考幀之內。

預測切割是習知壓縮的運動補償的預測方法擴展到容許有更正確的預測所藉由的方法。習知的壓縮係使用利用本發明的特徵建立模型方法所產生之額外的參考幀以增高正確性。當這些特徵參考幀的部份被習知的壓縮方式利用時，若特徵編碼是小於原本的習知編碼，則在壓縮上得到利益。

在一個實施例中，特徵係被表示為一組元素或特徵實例。在一個實施例中，該些特徵實例係被實現為矩形區域，每個矩形區域提供一項參照至唯一的幀、在該幀中之一空間的位置、以及在該幀中一個矩形區域範圍。該特徵的每個實例係代表該特徵的一個取樣的影像。在實例之間的特徵外觀上的變化係藉由該特徵的建立模型方法而被建立模型。

在一個實施例中，該些參考幀係存在一或多個來自先前合成的幀之子幀樣本。該些子幀的樣本是根據在先前合成的(解碼過的)幀以及目前幀中之該些子幀的區域之間的特徵實例的一致而定的。

在另一個實施例中，多個影像平面被合併成為較少的影像平面。這些較少的影像平面係使得位在接近該幀中預期的位置處之特徵被預測出。幀的縮減是根據合併非重疊的或是接近零的空間重疊的特徵到相同的平面中。

申請人在實務上的縮減亦已經利用此而進一步發展，

其藉由根據特徵資訊(先前的相符處、追蹤資訊、建立模型的資訊)來估計該特徵的一個界限區塊。

在另一非限制性的實施例中，每個合併的幀係等於被預測的幀的大小，並且該些特徵若不是剛好在習知的運動補償的預測機構所預期的位置、就是在空間上接近於該位置。

### 特徵的偵測

圖 5 係描繪已經在一或多個幀的視訊 520-1、520-2、… 520-n 中被偵測出的特徵 510-1、510-2、… 510-n。典型的是，此種特徵將會利用根據從像素導出的結構資訊以及指出習知的壓縮是利用不相稱的資源量來編碼該特徵的複雜度標準的數個不同的標準而被偵測出。相較於特徵編碼，每個特徵可進一步藉由一個在該圖中顯示為一個“區域”530-1、530-2、… 530-n 之對應的空間範圍(周圍)而在一個幀 520-1、520-2、… 520-n 中在空間上被識別出。

這些區域 530-1、530-2、… 530-n 例如可被抽取為一個簡單的像素資料矩形區域，並且置入一個集團 540 中，整個集團係代表一個特徵。

在一個幀中的一個特徵的每個實例是該特徵的外觀的一個樣本。請注意的是，當足夠數目的這些樣本被聯合到一個集團中時，其可被利用以對該特徵在那些幀中以及在該特徵並未從其取樣的其它幀中的外觀建立模型。此種模型係能夠轉換該外觀成為一個編碼的參數組，其可進一步

透過逆(inverse)模型而被解碼，以產生該特徵的接合。

小的空間區域係被識別且分析，以判斷其是否可根據某些連貫性標準而被組合成較大的空間區域。這些較大的空間區域係接著被分析，以判斷其作為候選特徵的合適性。假設該區域的特徵建立模型並未提供有利的編碼，則該候選特徵不是被拋棄、就是被保留用於對該特徵於後續的幀之未來的實例建立模型。該偵測過程係繼續直到僅剩下該些會呈現有利的建立模型的候選特徵為止。

空間區域在大小上會有所變化，從小群組的像素或子像素，到可能對應於實際的物件或是該些物件的部份之較大的區域，即為易於透過傳統的視訊壓縮演算法所決定之大區塊或次大區塊的分割步驟而被隱含分割的區域。然而，重要的是注意到偵測出的特徵可能不會對應於例如是物件及子物件之離散唯一且可分開的實體。並不要求特徵對應於此種實體。單一特徵可包含兩個或多個物件的元素、或是完全沒有物件的元素。重要的因數是本發明具有可用高於習知方法的效率來處理這些信號成分的可能性，並且這些信號成分純粹是根據其藉由基於特徵的視訊壓縮技術而被有效率地建立模型而充分滿足一個特徵的定義。

小的空間區域可被集合成為較大的區域，以便於將這些較大的區域識別為特徵。小區域係透過識別出區域之間的連貫性而被聚集成為較大的區域。有數種方式可識別出連貫性，其包含連貫的運動、運動補償的預測、以及編碼複雜度。

連貫的運動可透過較高階的運動模型而被發現。例如，每一個別的小區域之平移的運動係被整合成為一個仿射運動模型，其能夠對於該些小區域的每個區域近似出一個較簡單的運動模型。

若小區域的運動可在一致的基礎下被整合成為較複雜的模型，則此隱含表示該些區域之間有相依性，其潛在可提供一個優於習知的運動補償的預測方法之優點，並且此亦指出該些小區域之間的連貫性，其可透過特徵的建立模型而被利用。

編碼複雜度可透過習知的壓縮要代表一或多個該些小區域所需的頻寬之分析來加以判斷。在習知的編碼無法有效率地壓縮並且可能無法關聯為幀與幀之間的冗餘之一特定組的小區域有一個不相稱的頻寬配置之情形中，這些區域潛在可被集合成為一個特徵，其編碼複雜度可指出存在一個現象為特徵的建立模型將會較佳地表示。

一組已知的幀係分別完全被劃分成為以一種非重疊的模式配置之一致的影磚。每個影磚係被分析為一個獨立取樣的像素區域，其實際上被判斷為包含足夠用以描繪特徵的資訊。本發明使用這些取樣的區域以產生多個分類，其接著被使用在訓練一個分類器上。請注意的是，任何特徵的最後位置都可能與此最初的定位不同。

另一個實施例係從定義的影磚以及一個重疊該些影磚的磚面(tiling)來產生取樣的區域。重疊的取樣可被抵消，因而重疊的影磚的中心會出現在每四個基本的影磚之角落

的交叉處。此過完備的(over-complete)分割意謂著增加一個最初的取樣位置將會產生一個偵測出的特徵之可能性。其它可能更為複雜的拓撲學的分割方法亦為預期可用的。

一個特徵建立模型的預測器係分類取樣的區域成為一個叢集(cluster)，其具有相當大的機率是一個區域將會具有與同一叢集中的其它區域的某種一致性。該特徵建立模型的預測器係使用從取樣的區域導出的模式例子。

在一個較佳實施例中，該些特徵係在來自頻譜資料搜集(描述在以下的頻譜資料搜集(profiling)章節中)的幫助下被偵測出。頻譜資料搜集係提供幀中可能是單一特徵的部份之區域。此係被使用作為一個組合取樣的區域成為一個特徵的手段。

在一個實施例中，一個模式(pattern)特徵係被定義為一個頻譜特徵。該頻譜特徵係藉由將該區域從一個色彩空間轉換到一個HSV色彩空間而被找出。該轉換後的區域係接著反覆地向下被次取樣，直到導出的區域的影像向量空間具有遠小於原始區域的影像向量空間的大小為止。這些導出的區域被視為頻譜特徵。該些頻譜特徵係利用一種改良的K-means演算法而加以群集起來。K-means叢集係被用來根據叢集的頻譜分類來標示原始的區域。

在一個實施例中，一個分類器係根據取樣的區域之邊緣內容而被建立。每個區域都被轉換至DCT空間。導出的特徵之DCT係數係接著被加總用於上三角矩陣以及下三角矩陣。這些加總係接著被用以形成一個邊緣特徵空間。該

特徵空間係接著利用 K-means 而加以群集起來，並且原始的區域係根據其導出的區域叢集的分類來加以標示。

在另一實施例中，該頻譜特徵以及邊緣模式特徵分類器係被用來產生每個區域的多個分類。

## 特徵的追蹤

一個實施例係使用新偵測出的以及先前被追蹤的特徵之組合作為基礎以用於判斷在目前幀中具有相同對應的特徵之實例。此特徵在目前幀中的實例的識別以及包含此實例與該區域的先前發生的實例係構成特徵的追蹤。

圖 8 係展示一個特徵追蹤器 830 以及新偵測出的特徵與先前被追蹤的特徵 810 的組合之使用，以追蹤及分類特徵 820-1、820-2、…820-n。最初，一個一般的特徵偵測器 850 係被用來識別特徵。一致的判斷是根據目前幀 840 與先前偵測出的特徵 810 比對而得。該些被追蹤的特徵係被組織成為特徵組、或是分類為屬於一個先前組合的特徵組或是屬於一個新的特徵組。

特徵的一致最初可透過習知的梯度下降法最小化均方誤差的估計來加以判斷。該所產生的空間位移係得出該特徵在目前幀中之大約的位置。在該搜尋中使用的樣版(template)不必是該特徵的單一區域，而可以是與該特徵關連的區域中之任一個。最後的相符係接著以一種強健的方式評估作為非重疊的區域影磚中符合一個 MSE 臨界值的一個計數。

藉由施加一項空間的限制在兩個或多個區域的一致上，追蹤器能夠減少被追蹤的特徵數目，因此增加追蹤的計算效率。兩個或多個特徵之空間的一致亦可以指出額外的特徵情形，其中該特徵在過去實際上可能曾經是兩個特徵、或是某種其它複雜的特徵拓撲。追蹤器的模式容許有暫時簡併的(degenerate)追蹤狀態而讓特徵得以被追蹤，但會降低被追蹤的區域優先順序。

### 特徵的建立模型

在某個分析階段期間，將被預測的區域係被用來通過區域一致模型(RCM)，以便於決定在該模型之中將被用來建構一個區域預測模型的區域。

在一個實施例中，目標區域係被用來更新該 RCM，藉此產生在內含於該 RCM 之中的其它區域與該目標區域之間平移及中點正規化的一致。該所產生的成對的區域一致係識別出最可能對於該目標區域產生一個預測模型的其它區域。

本發明係將一個特定的目標區域之最佳的一致中的一或多個一致的組合內含到一組稱為一個區域集團(ensemble)中。該區域集團可被空間正規化朝向一個在該集團中之關鍵的(key)區域。在一個實施例中，時間上最接近該目標區域的區域係被選出作為該關鍵的區域。執行這些正規化所需的變形係被收集成為一個變形集團，並且該所產生的正規化的影像係被收集成為一個外觀集團，即如在美國專利

號 7,508,990、7,457,472、7,457,435、7,426,285、7,158,680、7,424,157 及 7,436,981 以及美國申請案號 12/522,322 中所述者，所有案件均為本案受讓人所有。上列的專利及申請案之整體教示係被納入作為參考。

該外觀集團係被處理以產生一個外觀模型，並且該變形集團係被處理以產生一個變形模型。該外觀及變形模型的組合係變成該目標區域的特徵模型。在一個實施例中，模型形成的方法是該集團的主要成份分析(PCA)的分解，接著是所產生的基礎向量的截斷。在另一實施例中，用於截斷的標準可以是集團內的重建。

在另一實施例中，模型(外觀及變形模型)形成的方法是在本文別處所述的壓縮感測(CS)，其中模型參數是從部分的參數量測解析而來的。

該目標區域被投射到該特徵模型上，以產生特徵參數。特徵參數係用於該區域的變形及外觀的建立模型。再者，特徵參數是該目標區域的編碼。

兩個或多個集團內的區域之特徵模型參數係利用時間的標準而被選出。給定在該些區域本身與該些目標區域之間已知的間隔，這些參數係被用來預測該目標區域的狀態。狀態模型的一個例子是在給定時間的步階下，兩個或多個特徵參數的線性外推。該線性模型係被用來預測該目標區域的特徵參數。

若外推的值提供該目標區域的適當接合(解碼)，則載明該目標區域的特徵參數變成不必要的、或是特徵參數可相

對於外推的參數而有差別地予以載明。

用於外推的狀態模型可以是比簡單的線性模型高階的。在一個實施例中，一個擴張型卡爾曼濾波器係被用來估計該特徵參數狀態。

### 區域一致模型

分類、登記以及變形分析的組合係提供一組資訊指出兩個或多個區域可被結合成為一個外觀及變形的聯合模型，稱為區域一致模型(RCM)的機率。

在一個較佳實施例中，該特徵偵測方法(如上所述)係遞增地分析新的特徵。此分析的一項結果是較高的機率為一個區域將會對應於被用來建構其中一個特徵偵測器的其它區域。

一旦區域係如同在以上討論的特徵偵測而被分類成為叢集，並且給定其個別的叢集標示後，則叢集間的區域係被分析以決定在區域對之間的每區域的一致性。

在一個較佳實施例中，上述的分類器係被用來定義取樣的區域的叢集，其區域來源像素係透過區域平移微調(在以下說明)而進一步被分析及定義。

此外，在一個較佳實施例中，在平移區域微調之後，區域的一致可進一步就其區域變形分析(在以下說明)的方面被定義。

在一個實施例中，該 RCM 的建構係遞增地被達成。兩個或多個區域係被用來對於組合的分類器/變形分析機構提

供最初的輸入。該 RCM 接著是利用改變分類器及變形分析元素的新區域而被更新。

在一個實施例中，上述的 RCM 遞增更新係被建構成使得一個特定的模型之區域的一致係根據詳述於下的基礎 (base) 複雜度分析，而以一種遍歷的 (traversal) 順序而被處理。

在一個實施例中，如上所論述之根據一種基礎複雜度分析 (在以下說明) 而定之遍歷的順序是一個利用遍歷的終止標準更新該 RCM 的遞迴過程之部份。當外觀 / 變形模型是從該些一致導出時，該終止標準會使得該處理完成到一個程度是最大化該 RCM 以具有最大降低複雜度的可能性來代表一致的能力。

## 區域平移微調

在一個實施例中，取樣出的區域被收集在一起而成為一組訓練的取樣區域。這些區域在每個幀中的空間位置係被微調。

微調係包含每個取樣的區域與其它各個取樣的區域完全的比較。此比較是由兩個影磚套合 (registration) 所構成。一個套合是一個第一區域對於一個第二區域的比較。第二個套合是該第二區域對於該第一區域的比較。每個套合是在該些區域於其個別的影像中的位置處被執行。所產生的套合偏移量以及對應的位置偏移量係被保留且被稱為關聯性 (correlation)。

該些關聯性係被分析以判斷多個套合是否指示一個取樣的區域位置應該被微調。若在來源幀中的微調的位置將會對於一或多個其它區域產生較低的錯誤相符，則該區域位置係被調整成該微調的位置。

該區域在來源幀中之微調的位置是透過其它區域一致的位置的線性內插法而被決定出，該些區域一致是在時間上橫跨來源幀中的該區域。

## 頻譜的資料搜集(profiling)

頻譜的資料搜集方法是一種統計的“均值追蹤及擬合”方法。此種方法的其它例子係被描述在如同應用到出現在影像及視訊幀中的空間機率分布之偵測、追蹤及建立模型的 CAMSHIFT、平均位移、中心點(medoid)位移、以及其衍生方法的文獻中。本發明之頻譜的資料搜集方法開始是分析橫跨一或多個幀的一個影像平面的一個區域的頻譜(色彩)平面的強度成分、像素。強度成分首先係透過該些值經由一個直方圖分箱(binning)方法的離散化而被處理。接著一個區域的直方圖係和一種追蹤機構一起使用以識別出在後續的幀中具有類似的直方圖之更多的對應區域。該區域的成分(位置、離散化標準及直方圖)組係反覆地微調，因此其收斂在這些成分的一個共同組。該微調後的成分組是該頻譜輪廓。該頻譜的資料搜集方法是一種特徵的偵測方法。

利用一維的 K-means 分類有一項優點，因而 HSV 色彩空間的色度通道係以分類器的形式被利用。此外，該些像

素係被分類並且直方圖的分箱係被填入，因而空間的不變矩 (invariant moment) 係被決定出。

用於本發明的核心基礎功能係利用已存在的資料來導出用於新資料的模型。該已存在的資料可透過任何編碼/解碼的方式而獲得並且被假設是可得到的。本發明係分析此資料以決定一組候選模式資料，其被稱為特徵資料，可包含視訊信號的一個空間局部成份的外觀及變形的資料。

給定一特定組已存在的特徵資料以及一個新的目標資料點之下，執行分析以決定建立用於代表該目標資料點的一個模型所需的特徵資料之最少量的描述。在不失一般性的意義下，該已存在的特徵資料係被稱為候選特徵向量，並且該目標資料點係被稱為目標向量。再者，該過程可應用到一或多個目標向量。

給定一個目標向量及一組候選特徵向量(全都視為相同特徵的部份)之下，該些候選特徵向量的一個最小子集合係被選出以在低誤差下合成該目標向量，此係產生一個小型且正確的分歧表示。

本發明聚集一組候選特徵向量成為所謂的特徵集團 (ensemble)。在一個實施例中，產生該特徵集團的第一步驟是選出一個關鍵的向量，亦即一個被判斷是該目標向量的一個良好近似的特徵向量。該關鍵的向量是該特徵集團中的第一向量。該特徵集團的其它候選特徵向量係依照其與該關鍵的向量之關聯性的順序而被選出(因而該特徵集團中的第二向量是具有下一個最高的與該關鍵的向量之關聯性

的特徵向量)。以此種方式排序一個特徵集團係稱為關鍵關聯性排序的(KCO)。

在另一實施例中，該特徵集團係利用該目標向量本身而被產生。該特徵集團的候選特徵向量係根據其與該目標向量的關聯性而被選出。任何使用目標向量關聯性的排序方法都被稱為目標關聯性排序的(TCO)。在一個 TCO 特徵集團中的第一特徵向量是具有與該目標向量最大的關聯性的候選特徵。在一個較佳實施例中，每次一個特徵向量“進入”該集團時，經由現有的集團( $U_r$ )之近似的目標重建係計算為  $U_r^*U_r^*t$ ，接著從該目標向量  $t$  減去以形成一個殘餘向量。該集團的下一個特徵向量接著被選擇為具有與該殘餘向量最大的關聯性的候選特徵。此種計算該殘餘向量並且接著選擇與該殘餘向量最佳的相符的遞迴過程因此被稱為順序的目標關聯性排序(STCO)。對於一個特定的集團大小而言，STCO 係確保該目標向量最有效率的表示。對於小的集團大小而言，其在功能上等同於正交匹配追蹤(請參見習知技術)，但在計算上更有效率。

在另一實施例中，並未計算殘餘向量，而且該特徵集團所有的候選特徵向量都是根據其與該目標向量本身的關聯性而被選出。此 TCO 方法，稱為全域的目標關聯性排序(GTCO)是比 STCO 快且簡單，但可能在該集團中產生冗餘。然而，對於選出該集團而言，兩種 TCO 方法都大致上遠優於該 KCO 方法。

一個位元遮罩(bitmask)係被用來傳送已被選用於該特

徵集團的特徵向量。

在一個實施例中，在 SVD 為基礎的編碼之前，該特徵集團中的特徵向量以及該目標向量本身係通過一種離散子波轉換(DWT)。此係使得該目標向量中的資訊更小型且更容易藉由 SVD 向量的一個小子空間來表示。該 DWT 是一種眾所週知的方法用於多尺度縮小信號資訊。在一個較佳實施例中，該 DWT 係以 Daubechies 9-7 雙正交的子波而被應用。該 DWT 係個別地應用至每個成分，因為特徵向量係在 YUV 色彩空間中。例如，長度 384 的 YUV 向量需要在 Y 成分上的一個長度 256 的 DWT 以及在 U 及 V 成分上的長度 64 的 DWT。

### 壓縮感測(CS)

在本發明的一個實施例中，壓縮感測(CS)係被採用作為在該特徵的建立模型(在別處描述)過程中之模型形成(外觀及變形模型)的方法。

在本發明中有三種所關注的 CS 演算法實際的應用：正交匹配追蹤(OMP)、L1 最小化(L1M)以及鏈結追蹤(CP)。每一種演算法都具有其本身的強項與弱點，但是該 L1M 對於大多數視訊處理應用而言都是難以接受的慢，因而在此領域中，OMP 及 CP 是兩種供選擇的 CS 演算法，而 L1M 則較不常被使用。

CS 演算法的有效性實際上是受限於計算時間、記憶體限制、或是量測的總數。為了對抗這些限制且實際改善 CS

演算法的效能，本發明使用數種可能的方法中的一或多種。簡言之，該些方法係透過以下而達到益處：(1)減少該文獻中為了精確的重建所指定的量測數目；(2)藉由一或多種特定的資料縮減技術來增加在輸入資料中的稀疏度(sparsity)；(3)分割該資料以減輕記憶體的限制；以及(4)適應性地將誤差的預估內建到重建演算法中。

一個實施例係利用一項事實為通常用於重建的數學要求是比必要的要求更為嚴格。利用比該文獻中所指定的較少量測來經常一致地達成“良好的”的影像資料重建是可能的。“良好的”重建是表示對於人眼而言，在視覺上相較於“完全”重建有很小的差別。例如，以一半所指定的量測數目應用鏈結追蹤(CP)仍然可達成“良好的”重建。

在另一實施例中，該輸入資料係“被縮減”以使得其更為稀疏，此係減少所需的量測數目。資料縮減技術包含將資料通過離散子波轉換(DWT)，因為資料在子波域中經常是較為稀疏的；藉由截斷來實際減少輸入資料的總大小，亦已知為下取樣(down-sampling)；以及對該資料進行臨界值法(除去所有小於某個臨界值的成分)。該些資料縮減技術中，DWT 轉換是最不會“侵入性的(invasive)”，並且理論上容許該輸入資料能夠完全回復。另外兩種縮減技術是“有損失的(lossy)”，因而無法讓信號完全回復。DWT 與 CP 配合良好，但是與正交匹配追蹤(OMP)或是 L1 最小化(L1M)配合的不好。因此，用於此資料縮減實施例的理想組合是鏈結追蹤演算法與該離散子波轉換的資料縮減技術。

在另一特別是極適合於平行處理的架構之實施例中，輸入資料是被劃分成區段(或是 2D 影像被劃分成影磚)，並且每個區段係個別地利用較小數目之所需的量測而被處理。此種方式良好適用在通常會受限於記憶體限制的 OMP 及 L1M。所需的量測矩陣的大小是造成 OMP 及 L1M 記憶體限制的原因。吾人可計算出記憶體矩陣超出系統記憶體的量。此額外的記憶體需求是一項“超取樣(oversampling)”因數。其係設定將該信號分割成區段數目的一個下限。

在另一實施例中，該過程將某種誤差的預估內建到該重建演算法中。該預估的誤差可能是由於超出正常的雜訊或是不準確的量測所引起的。該過程藉由放寬最佳化的限制、或是藉由在完成該重建過程之前停止遞迴來補償之。該重建於是為該資料的一個近似的擬合(fit)，但是此近似的解決方式可能是足夠的、或者當輸入資料是充滿雜訊或是不準確時此近似的解決方式可能是唯一可行的解決方式。

圖 2 係顯示一個概念上的視訊壓縮架構，其係在編碼器實施壓縮感測的量測。原始的視訊串流(200)係被傳送通過一個運動補償的預測演算法(202)以套合該資料(203)，藉此建立在多個幀中的像素群組之間的一致，使得由於運動所造成的冗餘可被分解出。接著，預處理(204)係被應用以使得該資料儘可能的稀疏(在 205 處)，因而 CS 量測以及之後的重建將會是儘可能有效的。CS 量測係被取得(206)並且變成該 CS 編碼(207)(已可發送)。之後在接合期間，該 CS 演算法係被用來解碼該些量測。

本發明識別、分開及預處理來自原始的視訊串流之信號成分，而成為極適於 CS 處理的稀疏信號。CS 演算法當然是與本發明的實施例相容。應注意到的是，圖 2 的某些觀點係相關於在美國專利號 7,508,990、7,457,472、7,457,435、7,426,285、7,158,680、7,424,157、以及 7,436,981 以及美國申請案號 12/522,322 中所論述的實施例，所有美國專利案都是本案受讓人的。以上所列的專利案及專利申請案的整體教示係被納入在此作為參考。

在視訊壓縮的背景中，當輸入影像具有某種稀疏性或是可壓縮性時，CS 係給予一項重要的益處。若輸入影像是密集的，則 CS 不是用於壓縮或重建的正確解決方式。CS 演算法可用少於習知的壓縮演算法所需的量測(需要相當於影像中的像素數目的量測)來壓縮及重建稀疏的輸入影像。請注意的是，信號稀疏性或可壓縮性是大多數壓縮技術所假設有的，因此 CS 所能夠提供改良的影像是設計大多數壓縮技術所針對的影像。

同樣要注意的是，將雜訊加入稀疏的影像會使得其在數學上更密集，但是並不會使得其“在資訊上”較不稀疏的。其仍然是一個稀疏的信號，並且以上述實際的實施方式中的一或多個方式來利用 CS 可產生此類的信號之有用的重建。

### 基礎複雜度分析

代表性的取樣的視訊區域可利用一種基礎方法而被分

析。一種此類的方法將會是習知以區塊為主的壓縮，例如 MPEG-4。

## 透過逆合成演算法的影像對齊

Xu 及 Roy-Chowdhury (“整合在連串視訊中的運動、照明及結構...”，IEEE 模式分析與機器智能匯刊，2007 年 5 月)係將該 LRLS 架構延伸至移動的物件(例如，在連串視訊中)，顯示此種物件是可藉由一種具有 9 個照明函數(原始的 LRLS 基礎影像)以及 6 個反映在 LRLS 基礎影像上的運動效果的運動函數之 15 維度的雙線性的基礎來加以良好的近似。

由 Xu 及 Roy-Chowdhury 近來提出的 IC 實施 (“在動態畫面中的 3D 姿態及照明的逆合成估計”，IEEE 模式分析與機器智能匯刊，其將被發表)係使用該逆合成(IC)演算法以從一序列的視訊幀中估計 3D 運動及照明參數。一個 2D 至 3D 至 2D 的翹曲(warping)函數係被用來在一個標準的(canonical)姿態下對齊來自不同的幀之(目標)影像與一個“關鍵的”幀(樣版)。給定一個幀的影像資料以及被成像的物件的一個基本的 3D 模型下，該 2D 至 3D 對映係決定在該 3D 模型中的哪些 3D 點(面/頂點)對應於哪些影像像素。一旦該 2D 至 3D 對映已經界定之後，該物件的姿態係在 3D 中位移先前的幀的姿態估計，藉此對齊目前幀與該關鍵的幀。在 3D 中位移後的物件係接著利用該 3D 至 2D(投射)的對映而對映回到 2D，以形成一個“姿態正規化的”影像幀。

一旦目標幀已經利用該 2D 至 3D 至 2D 的對映而套合到該樣版(關鍵的幀)之後，該所產生的姿態正規化的幀(PNF)係被用來估計 15 個運動參數，其對應於 9 個照明以及 6 個運動的變數。該些照明變數係經由該 PNF 至該些 LRLS(照明)基礎影像的最小平方擬合(least-squares fit)來加以預估出。在一個實施例中，藉由該些 LRLS 基礎影像預估出的照明成分係接著從該 PNF 減去，並且殘餘者係被用來經由最小平方擬合至該些運動函數以估計出 6 個運動參數(3 個平移以及 3 個旋轉)。該 PNF 接著可從該 15 維度的“雙線性的”照明/運動基礎以及其對應的參數向量而被重建。

本發明使用 Xu/Roy-Chowdhury 的 IC 實施的觀點，以幫助影像套合的應用。在一個實施例中，該 2D 至 3D 至 2D 的對映係被使用作為特徵區域的中點正規化之一個計算上有效率的替代方式。對於其中正確的 3D 模型(例如用於臉的 Vetter 模型)存在的特徵而言，該對映的過程是特別有用的。在此實施例中，該些模型點係在某種姿態(該“模型姿態”)下被指定，並且該關鍵的幀(該樣版)以及目前幀(或目標幀)係被套合至該模型姿態。

### 增量的奇異值分解(ISVD)演算法的應用

在本發明中，該 SVD 係利用共同大小的臨界值方法的一種變化，在此稱為百分比臨界值法而被縮減。在一個實施例中，在一個特定的 SVD 分解中的奇異值的總能量 E 係計算為該些奇異值的總和。當奇異值依序地(以大小遞減的

順序，從最大至最小)加總直到在縮減組中的奇異值的總和超過 E 的某個百分比臨界值為止，該些奇異值的一個分組(在本文中稱為一個“縮減組”)係被產生。除了該大小臨界值並不需要是事先已知的以外，此縮減方法係等同於大小臨界值法(請參見習知技術)。

在本發明中，該奇異值分解(SVD)係如下所述地應用至特徵資料。該  $M \times N$  資料矩陣  $\mathbf{D}$  是由一個特定的視訊影像幀的區域(影磚)所導出的特徵向量的一個集團所組成。該  $M \times 1$  特徵向量係從 2D 影像影磚而被行向量化的，並且被連鎖以形成該資料矩陣  $\mathbf{D}$  的行。在一個實施例中，該資料矩陣係接著被因式分解成其 SVD，並且接著被簡化， $\mathbf{D}_r = \mathbf{U}_r * \mathbf{S}_r * \mathbf{V}_r'$ ，其中該縮減是透過百分比臨界值法。留下來的奇異向量係接著被用來編碼該  $M \times 1$  目標向量  $\mathbf{t}$ (將被傳送的特徵)，其中最後的編碼是由  $\mathbf{U}_r' * \mathbf{t}$  而得。典型的大小可能是  $M=384$ ， $N=20$ ，且  $r=10$ ，因而一個長度 384 的目標向量係以 10 個係數被壓縮(編碼)。

因為並非所有在該集團資料矩陣  $\mathbf{D}$  中的特徵向量都是立即可利用的，所以該增量的 SVD(ISVD)係被用來根據現有的奇異值分解以及資料更新來更新該 SVD。在一個實施例中，小數目的特徵向量係分組在一起以形成一個最初的資料矩陣  $\mathbf{D}_0$ ，針對該最初的資料矩陣  $\mathbf{D}_0$ ，習知的 SVD 可輕易地算出。接著，隨著額外的特徵資料向量被加到該集團資料矩陣，該 ISVD 係為了該增加後的資料矩陣而被用來更新該 SVD。在另一個實施例中，因為在該子空間已經在

該集團資料矩陣中被表示之下，新的特徵資料向量可能有時候是冗餘的，所以在新的資料向量被加到現有的集團之前，一個線性獨立(independence)測試係被施加至該新的資料向量。一旦整組的特徵資料向量已經都加到該集團之後，該 SVD 係被更新且縮減(經由百分比臨界值法)以提供最後的 SVD 為基礎的編碼。

在另一實施例中，該 SVD 係利用留下來的奇異向量( $U_r$ 的行)與該目標向量  $t$  的關聯性而被降低。總關聯性能量  $CE$  係計算為該些關聯性的總和。當關聯性依序地(以大小遞減的順序，從最大至最小)加總直到在該縮減組中的關聯性的總和超出  $CE$  的某個百分比臨界值為止，該些奇異值的一個分組(在本文中稱為一個“縮減組”)係被產生。此縮減該 SVD 的方法，稱為目標-關聯性百分比臨界值法，係依循與基本的百分比臨界值的 SVD 縮減方法相同的方法，除了(留下來的奇異向量與該目標向量的)目標關聯性係被使用於計算，而不是奇異值用於計算。

### 轉換為基礎的處理

本發明係在轉換空間中的視訊幀資料上執行經驗的特徵分類。在一個實施例中，一組來自一個參考幀的  $N_t$  個特徵係提出作為該分類器的輸入。每個特徵係利用可選擇的線性轉換(可行的轉換包含離散子波轉換[DWT]以及曲波(curvelet)轉換[CuT])，從像素空間轉換至轉換空間。接著，對應於每個特徵之最大的  $P$  個係數的索引係被表列，並且

對於每個特徵而言，在所有的係數表列中  $P$  個最常出現的係數係被用來產生一個 ( $P \times 1$ ) 分類向量 (CV) (總共  $N_t$  個“參考”CV)。接著，每個新的特徵向量  $v$  係被分類，此係藉由轉換該向量、抽出  $v$  的 CV 索引、以及計算在  $v$  的 CV 以及每個參考 CV 之間的相似性量測。該測試的特徵係被分類為該具有參考 CV 會最大化該相似性量測的特徵。

來自兩個或多個具有不同的強項及弱點的線性轉換的資訊可利用正交匹配追蹤來加以結合，以改善該經驗的轉換為基礎的特徵分類器的效能。在一個實施例中，來自可有效的代表紋理的 DWT 的基礎向量以及來自可有效的代表邊緣的 CuT 的基礎向量係結合成為一個函數庫  $D$ 。接著，OMP 係被用來利用在  $D$  中的函數計算  $N_t$  個特徵的每個特徵的一個信號表示、以及該“測試”的特徵向量的一個表示。該分類器接著如同上述的基本的轉換為基礎的分類器繼續進行。以此種方式組合來自多種轉換的資訊可以改善分類器效能而優於每一個別的分類器所能達成的效能。

線性轉換(例如，DWT 及 CuT)亦可被利用於特徵的壓縮及編碼。在一個實施例中，一旦一個特徵被轉換，該些轉換係數係依據大小來排序，並且根據一個能量保留準則(例如，保持足夠的係數以使得 99% 的特徵能量被保持)來分劃。典型的是，保持 99% 的信號能量所需的轉換係數是遠少於在像素空間中保持 99% 的信號能量所需的像素。該些轉換係數值係代表該特徵的編碼，並且該壓縮增益係得自於所保留的轉換係數相對於該特徵中的像素數目的百分

比。在另一個實施例中，來自多種轉換的資訊可再次利用OMP來結合，以改善壓縮增益。

儘管本發明已經特別參照其範例實施例來展示與描述，但是所屬技術領域中具有通常知識者將會理解各種在形式及細節上的變化都可以在實施例中做改變，而不脫離由所附的申請專利範圍涵蓋的本發明的範疇。

### 【圖式簡單說明】

前述內容從上述本發明的範例實施例之更特定的說明來看(如同在所附的圖式中所描繪者)將會是明顯，其中類似的參考符號是指不同的圖式中相同的元件。該圖式並不必然按照比例的，而是強調描繪本發明的實施例。

圖 1 是本發明的一個實施例(混合的編解碼器)的概要圖。

圖 2 是在本發明的編碼器中體現的一個視訊壓縮架構的方塊圖。

圖 3 是其中利用到本發明的實施例之電腦網路環境的概要圖。

圖 4 是在圖 3 的網路中之電腦節點的方塊圖。

圖 5 是描繪代表本發明的一個實施例之特徵的建立模型之圖。

圖 6 是描述根據本發明的一個實施例的預測過程之圖。

圖 7 是本發明的一個實施例(編解碼器)的方塊圖。

圖 8 是描繪根據本發明的一個實施例的特徵追蹤之圖。

201016016

【主要元件符號說明】

無

201016016

## 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：98133191

※ 申請日：98.10.6

※IPC 分類：

H04N 7/26

(2006.01)

### 一、發明名稱：(中文/英文)

基於特徵的視訊壓縮

FEATURE-BASED VIDEO COMPRESSION

### 二、中文發明摘要：

處理視訊資料的系統及方法係被提出。具有一系列視訊幀的視訊資料係被接收並且處理。一個候選特徵的一或多個實例係在該些視訊幀中被偵測出。藉由採用在一個運動補償的預測過程中所用的參考幀處理，該候選特徵的實例係與該些視訊幀中的非特徵切割開。該運動補償的預測過程係選出具有對應於該候選特徵的一或多個實例之特徵的先前解碼過的視訊幀。該些先前解碼過的視訊幀係被處理以識別出該候選特徵之可能的符合者。當先前解碼過的視訊幀中有大量的部份包含該候選特徵的實例時，該候選特徵的實例係被聚集成為一組。該候選特徵組係被用來產生一個基於特徵的模型。該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型。該外觀變化模型是藉由對於該候選特徵的實例的變化建立模型而被產生的。該系統判斷出與利用該基於特徵的模型以對於該候選特徵組建立模型相關連之壓縮效率的程度、以及與利用傳統的視訊壓縮以對於該候選特徵組建立

模型相關連之壓縮效率的程度。該基於特徵的模型之壓縮效率係和該傳統的視訊模型之壓縮效率做比較，並且該系統係判斷哪一個是具有較大的壓縮值。根據哪一個是具有較大的壓縮值，該視訊資料係利用該基於特徵的模型或傳統的視訊編碼而被編碼。

### 三、英文發明摘要：

Systems and methods of processing video data are provided. Video data having a series of video frames is received and processed. One or more instances of a candidate feature are detected in the video frames. The instances of the candidate feature are segmented from non-features in the video frames by employing reference frame processing used in a motion compensated prediction process. The motion compensated prediction process selects previously decoded video frames, which have features corresponding to one or more instances of the candidate feature. The previously decoded video frames are processed to identify potential matches of the candidate feature. When a substantial amount of portions of previously decoded video frames include instances of the candidate feature, the instances of the candidate feature are aggregated into a set. The candidate feature set is used to create a feature-based model. The feature-based model

includes a model of deformation variation and a model of appearance variation of instances of the candidate feature. The appearance variation models are created by modeling variation of instances of the candidate feature. The system determines the extent of compression efficiency associated with using the feature-based model to model the candidate feature set, as well as the extent of compression efficiency associated with using conventional video compression to model the candidate feature set. The feature-based model compression efficiency is compared with the conventional video modeling compression efficiency, and the system determines which one is of greater compression value. The video data is encoded using the feature-based models and conventional video encoding based on which one is of greater compression value.

## 七、申請專利範圍：

1. 一種處理視訊資料的電腦方法，其係包括以電腦實施的步驟：

接收由一系列的視訊幀所構成的視訊資料；以及  
藉由以下來編碼該些視訊幀的部份：

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例；

該偵測係判斷在該一或多個先前解碼過的視訊幀中的實例之位置資訊，該位置資訊係包含一個幀號碼、在該幀中的一個位置、以及該實例的一個空間的周圍；

該候選特徵是一組一或多個偵測出的實例；

藉由一個運動補償的預測過程以利用一或多個先前解碼過的視訊幀來預測在該系列中的一個目前視訊幀的一部份；

該運動補償的預測過程係以位置的預測初始化，其中該些位置的預測係提供來自偵測出的特徵實例在先前解碼過的視訊幀中之位置資訊；

利用該些藉由加入該運動補償的預測過程轉換的實例中的一或多個實例，界定一個特徵以及該些轉換後的實例以產生一個第一基於特徵的模型，該第一基於特徵的模型係使得在該目前幀中預測一個實質相符的特徵實例的出現及一個來源位置成為可能的，其中該實質相符的特徵實例是一個關鍵的特徵實例；

比較該一或多個界定出的特徵的第一基於特徵的模型

與一個習知的視訊編碼模型；

以及從該比較判斷出哪一個模型可致能較大的編碼壓縮；以及

利用該比較及判斷的步驟的結果，施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，以及施加習知的視訊編碼至該一或多個視訊幀的其它部份。

2.如申請專利範圍第1項之方法，其中在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例係進一步包含：

藉由識別具有實質接近的空間上鄰近的像素之一個空間上連續的群組以偵測一個候選特徵的至少一個實例；以及

該些識別出的像素係界定該一或多個視訊幀中的一個視訊幀的一部份。

3.如申請專利範圍第2項之方法，其中在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例進一步包含：

利用該運動補償的預測過程，從複數個候選特徵實例選擇被預測能提供編碼效率的一或多個實例；以及

根據該運動補償的預測過程從獨特的先前解碼過的視訊幀之預測的選擇以判斷該候選特徵之目前的實例與在該目前視訊幀中的其它特徵及非特徵的切割。

4.如申請專利範圍第2項之方法，其中該運動補償的預測過程係進一步利用屬於一或多個特徵的特徵實例來初始

化，此等特徵係在該目前幀中具有實例為符合該視訊部份，其中該視訊部份是在該目前幀中。

5.如申請專利範圍第2項之方法，其中該群組的像素進一步包含以下的一或多個：大區塊或是大區塊的部份。

6.如申請專利範圍第1項之方法，其進一步包含藉由以下來形成一個第二基於特徵的模型：

利用該第一基於特徵的模型作為來自一或多個特徵實例的一或多個運動補償的預測之一個預測的目標，產生該第一基於特徵的模型之一組預測；以及

在組合之後，該組預測變成該第二基於特徵的模型。

7.如申請專利範圍第6項之方法，其中該第二基於特徵的模型係被用來對第一基於特徵的模型的殘餘建立模型，其包含：

對該第二基於特徵的模型相對於該殘餘之結構變化以及外觀變化建立模型；

利用該模型來編碼該殘餘以產生外觀及變形參數；以及

利用該些參數以縮減該殘餘的編碼大小。

8.如申請專利範圍第1項之方法，其中界定一或多個特徵進一步包含藉由以下來根據該候選特徵的實例中的一或多個實例界定一或多個聚集的特徵：

聚集不同的候選特徵的實例成為一個聚集的候選特徵；以及

利用該聚集的候選特徵之該組實例以形成一個實質大

於未聚集的候選特徵之原始的實例的區域，其中該較大的區域係透過在該組中該候選特徵的實例間的連貫性的識別而被形成。

9.如申請專利範圍第 8 項之方法，其中該連貫性係被定義為在該些實例中藉由一個較少參數運動模型實質近似出的外觀一致。

10.如申請專利範圍第 7 項之方法，其中該第二特徵的模型係提供一個與該實例在該解碼過的幀中相對於該空間的位置相關連之選配的矩形像素區域範圍。

11.如申請專利範圍第 10 項之方法，其中該第二特徵的模型係藉由對該特徵之先前正規化的實例建立模型而導出的；以及

其中該些先前正規化的實例是以下所列的任一個：在該目前幀中的該實例、來自一個實質最近的先前解碼過的幀之一個實例、或是來自該些先前解碼過的視訊幀的該些實例的平均。

12.如申請專利範圍第 11 項之方法，其中該外觀模型係藉由該些正規化的第二基於特徵的模型實例的 PCA 分解來表示。

13.如申請專利範圍第 10 項之方法，其進一步包括判斷在每組的特徵實例中相較於其第二基於特徵的模型實例的一致之空間變化的一個變形模型；

為了該變形模型，對於在該組中的每個特徵實例，利用下列的一或多個來近似在該些變形實例中的變化：一個

運動補償的預測過程、網格變形、以及一個具有實質縮減的參數化之運動模型；

將該些變形實例整合到該變形模型中；以及  
其中在該變形模型中的變化係藉由一 PCA 分解來表示。

14.如申請專利範圍第 1 項之方法，其中該運動補償的預測過程係在一個數量上實質大於習知的視訊資料編碼之先前解碼過的視訊幀之一個選集上做運算；以及

其中先前解碼過的視訊幀的該選集並不依賴使用者的監督。

15.如申請專利範圍第 1 項之方法，其中響應於該比較及判斷的步驟來施加習知的視訊編碼係進一步包含當形成該目前幀的部份的一預測時，藉由一個致能該些視訊幀的一或多個視訊幀的部份在記憶體中有更大的壓縮的實例預測過程來加入該習知的視訊編碼；以及

其中該實例預測過程進一步包含：

利用該基於特徵的模型以決定該界定出的特徵出現在一個正被編碼的目標大區塊的一或多個實例，以形成該目前幀之預測的部份；以及

利用該基於特徵的模型以合成像素來預測該目前幀的部份。

16.如申請專利範圍第 15 項之方法，其中響應於該比較及判斷的步驟以施加習知的視訊編碼至該些視訊幀的一或多個視訊幀的部份係進一步包含：

指定一個機率給該些先前解碼過的視訊幀，其中該機率是根據對於該幀利用來自該運動補償的預測過程之位置的預測而判斷出的組合的預測編碼效能的改善而定的；

界定該機率為在用於該目前幀的第一基於特徵的模型以及一個第二基於特徵的模型的分析期間所用之運動補償的預測過程的該組合的編碼效能；

根據該些先前解碼過的視訊幀之機率以分類該些先前解碼過的視訊幀以決定一個從最佳至最差的指標；以及

根據計算及記憶體上的需求來截短該索引表列。

17.如申請專利範圍第 15 項之方法，其進一步包含響應於判斷出下列情形時，再次使用該特徵實例之預測的像素於預測在該目前幀中的其它特徵實例：

在該目前幀中，該界定出的特徵的一或多個實例係重疊超過一個大區塊；或是

當該界定出的特徵的一或多個實例實質相符在該目前幀中的一個大區塊的位置資訊時，該界定出的特徵的一或多個實例係代表一個大區塊。

18.如申請專利範圍第 10 項之方法，其進一步包括預測該些外觀參數以及變形參數以用於一個基於特徵的模型之目前的實例的接合，以及利用該外觀模型及變形模型以及時間上最近的參數以從該基於特徵的模型內插及外推參數來預測在該目前幀中的像素的步驟，其包含：

根據哪一種方法已產生該些時間上最近的特徵實例之最正確的近似來決定用於該些實例的接合的值是線性內插

或是線性外推；

利用較高階的二次方法來偵測該些線性內插及外推的方法之實質減小的效果；

偵測該些二次方法之實質減小的效果並且採用包括擴張型卡爾曼濾波器之更高等的以狀態為基礎的方法以預測該些外觀及變形參數；以及

其中該模型之實際的參數係選配地相對於該些預測出的參數做差分編碼。

19.如申請專利範圍第18項之方法，其中來自該基於特徵的模型之參數係使得在預測該目前幀中的像素所需的計算資源上能夠減少，使得當利用傳統的視訊壓縮以利用該些先前解碼過的視訊幀的一或多個部份來預測該目前幀中的像素時需要更多的計算資源。

20.如申請專利範圍第1項之方法，其中該基於特徵的編碼係被嵌入在習知的視訊編碼內。

21.如申請專利範圍第1項之方法，其中該一或多個界定出的特徵並沒有與在該一或多個視訊幀中之不同的顯著的實體(物件、子物件)的一致。

22.如申請專利範圍第1項之方法，其中該些顯著的實體係透過使用者對於偵測出的特徵之監督的標示為屬於或不屬於一個物件而被決定。

23.如申請專利範圍第1項之方法，其中該些界定出的特徵係包含該些視訊幀的兩個或多個顯著的實體、背景或其它部份的元素。

24.如申請專利範圍第 1 項之方法，其中一個界定出的特徵並不對應於一個物件。

25.如申請專利範圍第 11 項之方法，其中施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，以及施加習知的視訊編碼至該一或多個視訊幀的其它部份的步驟：

施加壓縮感測至該第二基於特徵的模型預測的殘餘；

其中該壓縮感測的施加係利用該平均外觀作為一量測並且從該量測來預測該信號；

其中與該壓縮感測的預測相關連的變異係從該第二基於特徵的模型移除；

其中基於特徵的建立模型係專注在該剩餘的殘餘之更小的編碼；以及

施加習知的視訊編碼至該一或多個視訊幀之剩餘的像素以及剩餘的視訊幀。

26.如申請專利範圍第 25 項之方法，其進一步包括使得該視訊資料變為稀疏以增加該施加壓縮感測的步驟的效果之步驟。

27.如申請專利範圍第 1 項之方法，其中該些實例中的一或多個實例係利用一種線性轉換來加以轉換。

28.如申請專利範圍第 1 項之方法，其中該實質相符的特徵是一個利用一種率-失真的尺度判斷出之最佳的相符。

29.如申請專利範圍第 1 項之方法，其進一步包含藉由以下來解碼該編碼的視訊資料：

在一大區塊層級上判斷在該編碼的視訊資料中是否有

一個編碼的特徵；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該些合成的特徵實例以重建原始的視訊資料。

30.如申請專利範圍第1項之方法，其中該基於特徵的編碼係包含對於該一或多個視訊幀的部份施加物件為主的編碼。

31.一種用於處理具有一或多個視訊幀的視訊資料之數位處理系統，其係包括：

一或多個執行一個編碼器的電腦處理器；

該編碼器係利用基於特徵的編碼來編碼該些視訊幀的部份，此係藉由：

偵測一個候選特徵在該些視訊幀的一或多個視訊幀中的一或多個實例；

利用一個運動補償的預測過程，分割在該一或多個視訊幀中的該候選特徵的一或多個實例與非特徵，該運動補

償的預測過程係選擇具有對應於該候選特徵的該一或多個實例的特徵之先前解碼過的視訊幀；

利用該候選特徵的實例中的一或多個實例以界定一或多個特徵實例，其中該一或多個界定出的特徵實例係被預測為能夠在該基於特徵的編碼中提供相對於習知的視訊編碼增加相當大的緊密性；

從該一或多個先前解碼過的視訊幀判斷位置資訊，該位置資訊係包含該一或多個界定出的特徵實例在該一或多個先前解碼過的視訊幀中的一個位置以及一個空間的周圍；

利用該一或多個界定出的特徵實例以形成一個基於特徵的模型，該基於特徵的模型係包含來自該些先前解碼過的視訊幀的位置資訊；

利用該基於特徵的模型以正規化該一或多個界定出的特徵實例，該正規化係利用來自該一或多個先前解碼過的視訊幀的位置資訊作為一個位置的預測，所產生的正規化是該一或多個界定出的特徵實例在該目前視訊幀中的預測；

對於該些界定出的特徵中的一或多個特徵來比較該基於特徵的模型與一個習知的視訊編碼模型，並且從該比較來判斷哪一個模型會得到較大的編碼壓縮；以及

利用該比較及判斷的步驟的結果，施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，並且施加習知的視訊編碼至該一或多個視訊幀的其它部份。

32. 一種處理視訊資料的方法，其係包括：

接收具有一系列視訊幀的視訊資料；

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵；

藉由採用在一個運動補償的預測過程中所用的參考幀處理以分割在該些視訊幀中的該候選特徵與非特徵；

處理先前解碼過的視訊幀的一或多個部份以識別出該候選特徵之可能的符合者；

判斷先前解碼過的視訊幀中包含該候選特徵的實例之大量的部份；

將該候選特徵的實例聚集成為該候選特徵的一組實例；

處理該候選特徵組以產生一個基於特徵的模型，其中該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型，該些外觀變化模型是藉由對該候選特徵的實例的像素變化建立模型來加以產生的，該些結構的變化模型是藉由對該候選特徵的實例的像素一致變化建立模型來加以產生的；

判斷與利用該基於特徵的模型以對於該候選特徵組建立模型相關連的壓縮效率；

判斷與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連的壓縮效率；

比較該基於特徵的模型的壓縮效率與該習知的視訊建立模型的壓縮效率，並且判斷哪一個是具有較大的壓縮值；

根據哪一個是具有較大的壓縮值來利用該基於特徵的模型以及習知的視訊編碼以編碼該視訊資料。

33. 一種用於處理具有一或多個視訊幀的視訊資料之數位處理系統，其係包括：

執行一個編碼器的一或多個電腦處理器；

該編碼器係利用基於特徵的編碼以藉由以下來編碼該些視訊幀的部份：

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵；

藉由採用在一個運動補償的預測過程中所用的參考幀處理以分割在該些視訊幀中的該候選特徵與非特徵；

處理先前解碼過的視訊幀的一或多個部份以識別出該候選特徵之可能的符合者；

判斷先前解碼過的視訊幀中包含該候選特徵的實例之大量的部份；

將該候選特徵的實例聚集成為該候選特徵的一組實例；

處理該候選特徵組以產生一個基於特徵的模型，其中該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型，該些外觀變化模型是藉由對該候選特徵的實例的像素變化建立模型來加以產生的，該些結構的變化模型是藉由對該候選特徵的實例的像素一致變化建立模型來加以產生的；

判斷與利用該基於特徵的模型以對於該候選特徵組建

立模型相關連的壓縮效率；

判斷與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連的壓縮效率；

比較該基於特徵的模型的壓縮效率與該習知的視訊建立模型的壓縮效率，並且判斷哪一個是具有較大的壓縮值；

根據哪一個是具有較大的壓縮值來利用該基於特徵的模型以及習知的視訊編碼以編碼該視訊資料。

#### 34. 一種處理視訊資料的方法，其係包括：

在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該些合成的特徵實例以重建原始的視訊資料。

#### 35. 一種用於處理視訊資料的資料處理系統，其係包括：

執行一個混合的編解碼器的解碼器之一或多個電腦處

理器，其係能夠藉由以下來利用視訊資料解碼：

藉由在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料以解碼一個編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該視訊資料之該些合成的特徵以重建一原始的視訊資料。

## 八、圖式：

(如次頁)

理器，其係能夠藉由以下來利用視訊資料解碼：

藉由在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料以解碼一個編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該視訊資料之該些合成的特徵以重建一原始的視訊資料。

## 八、圖式：

(如次頁)

201016016

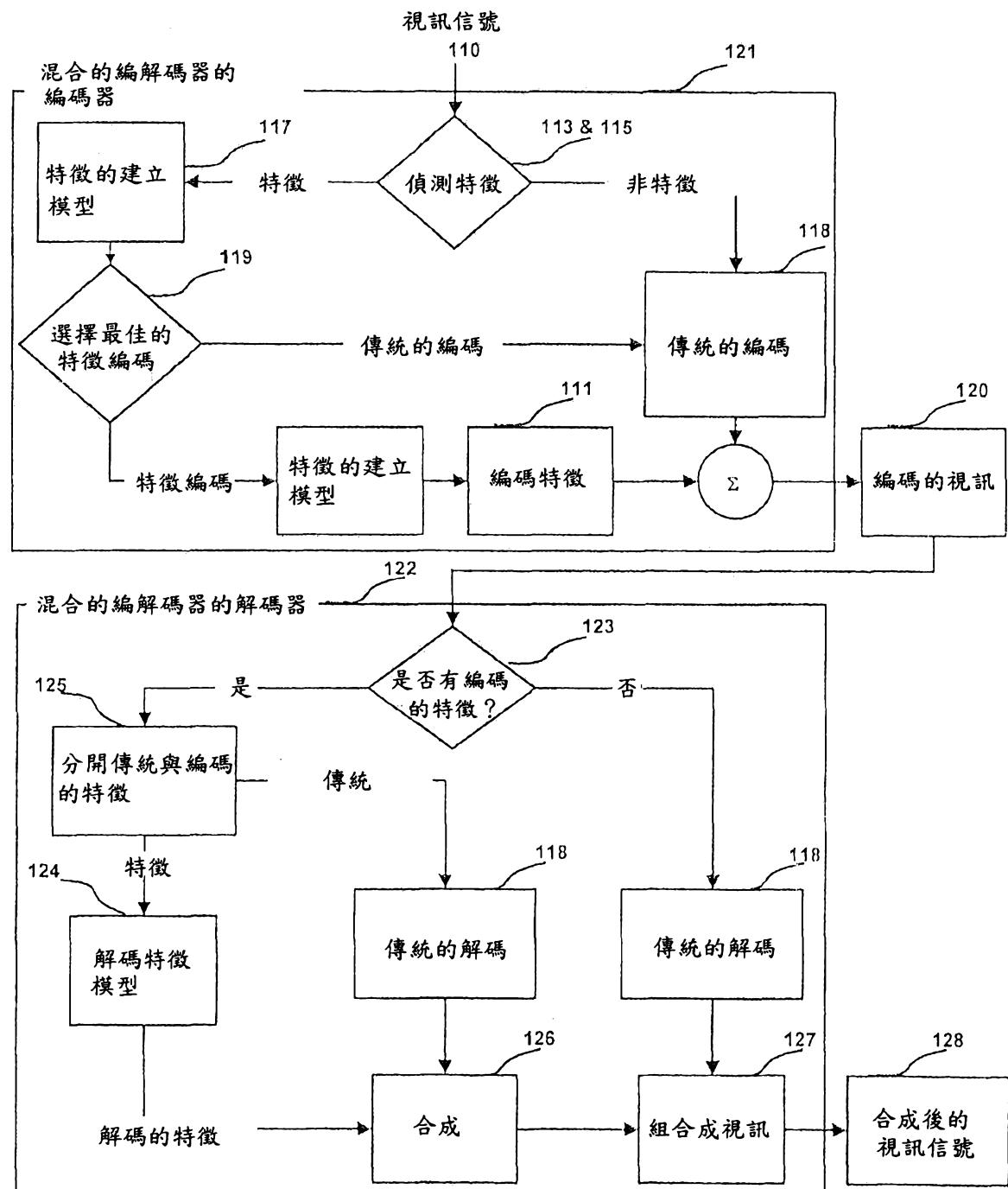


圖 1

201016016

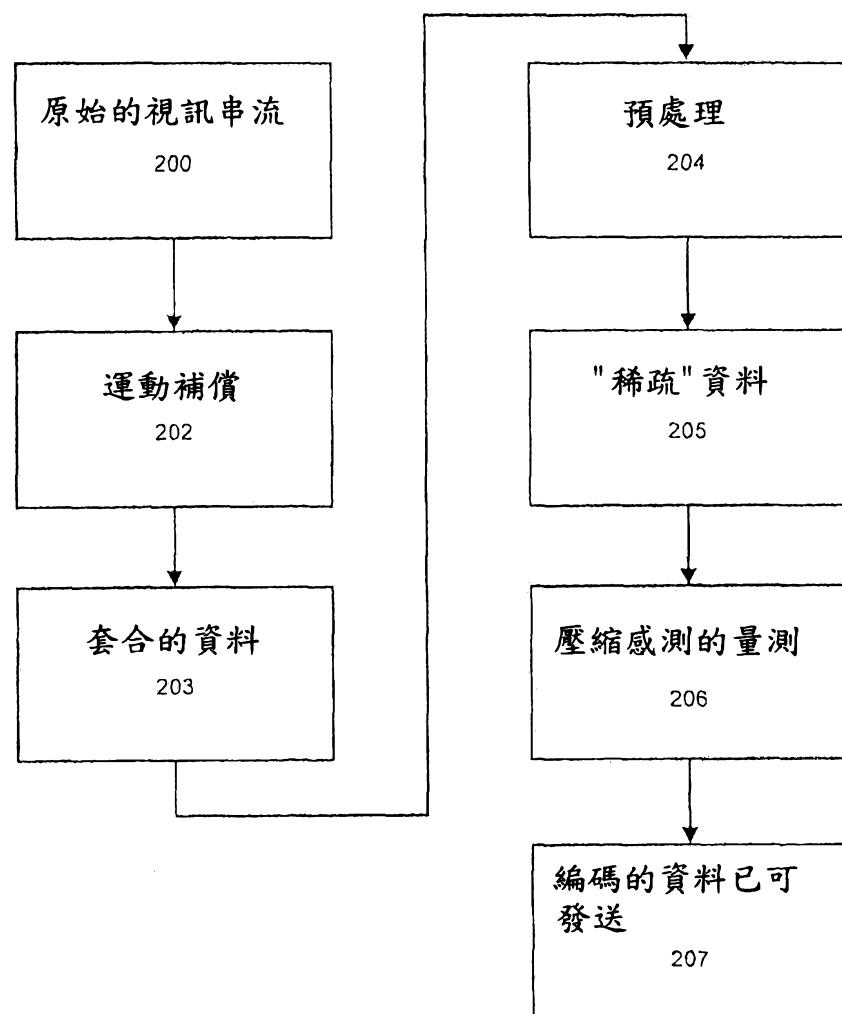


圖2

201016016

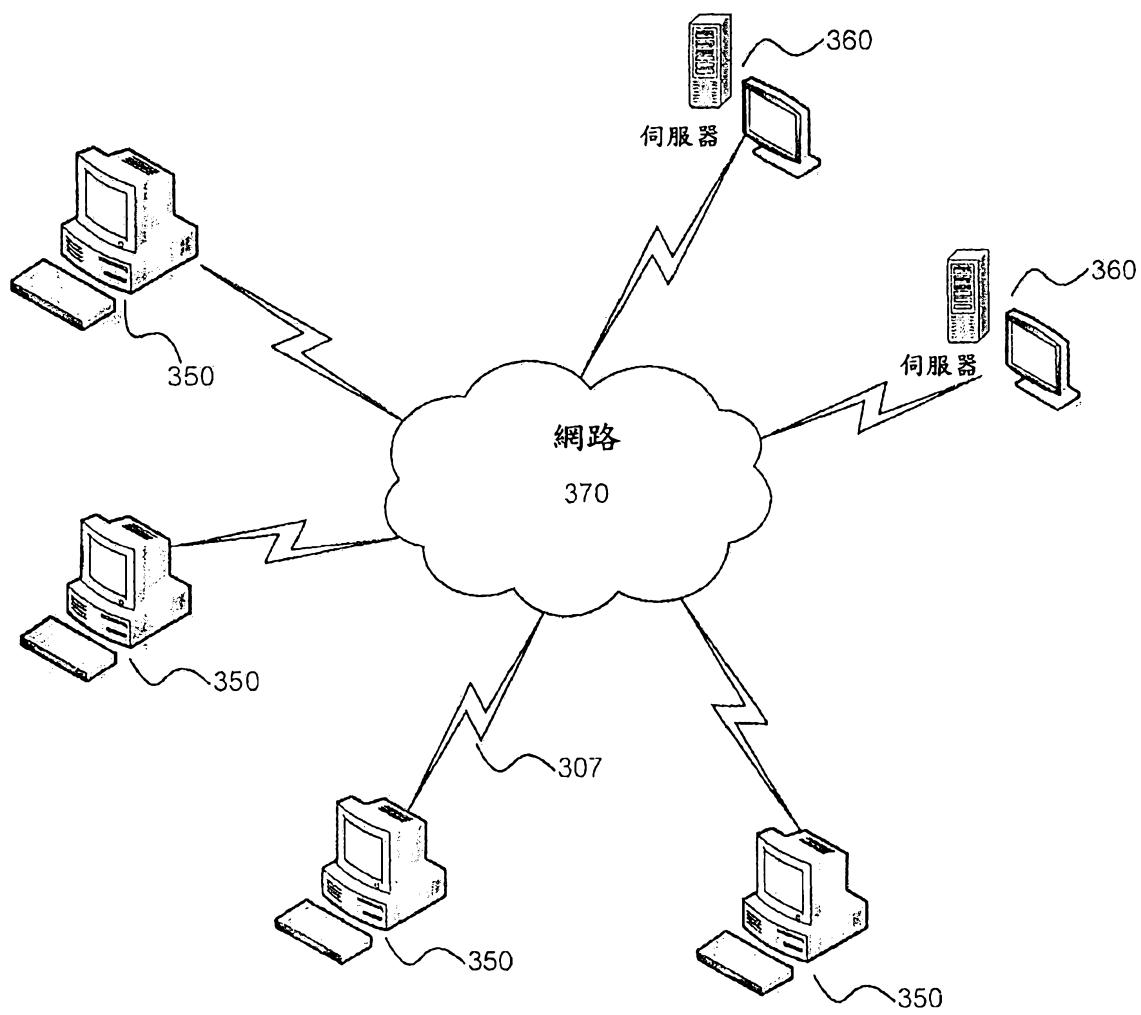


圖3

201016016

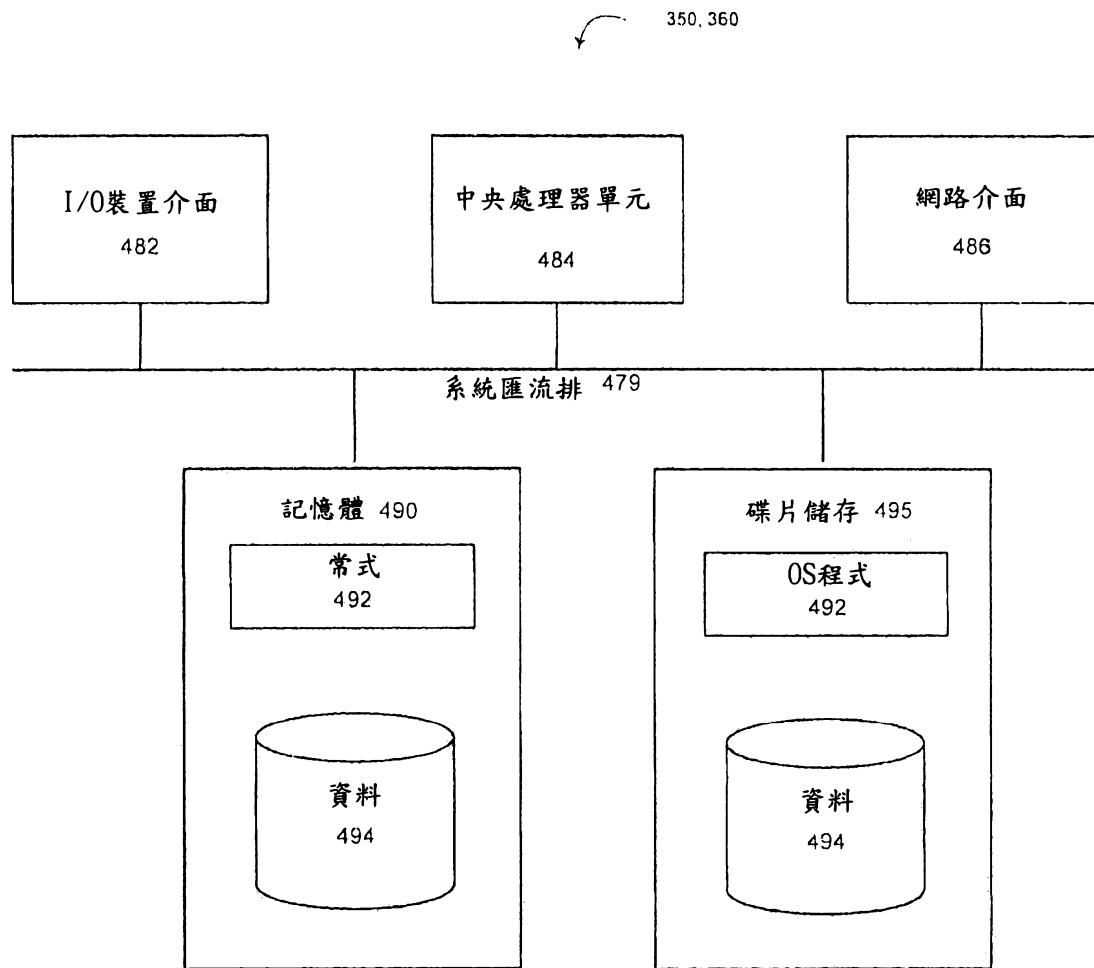


圖4

201016016

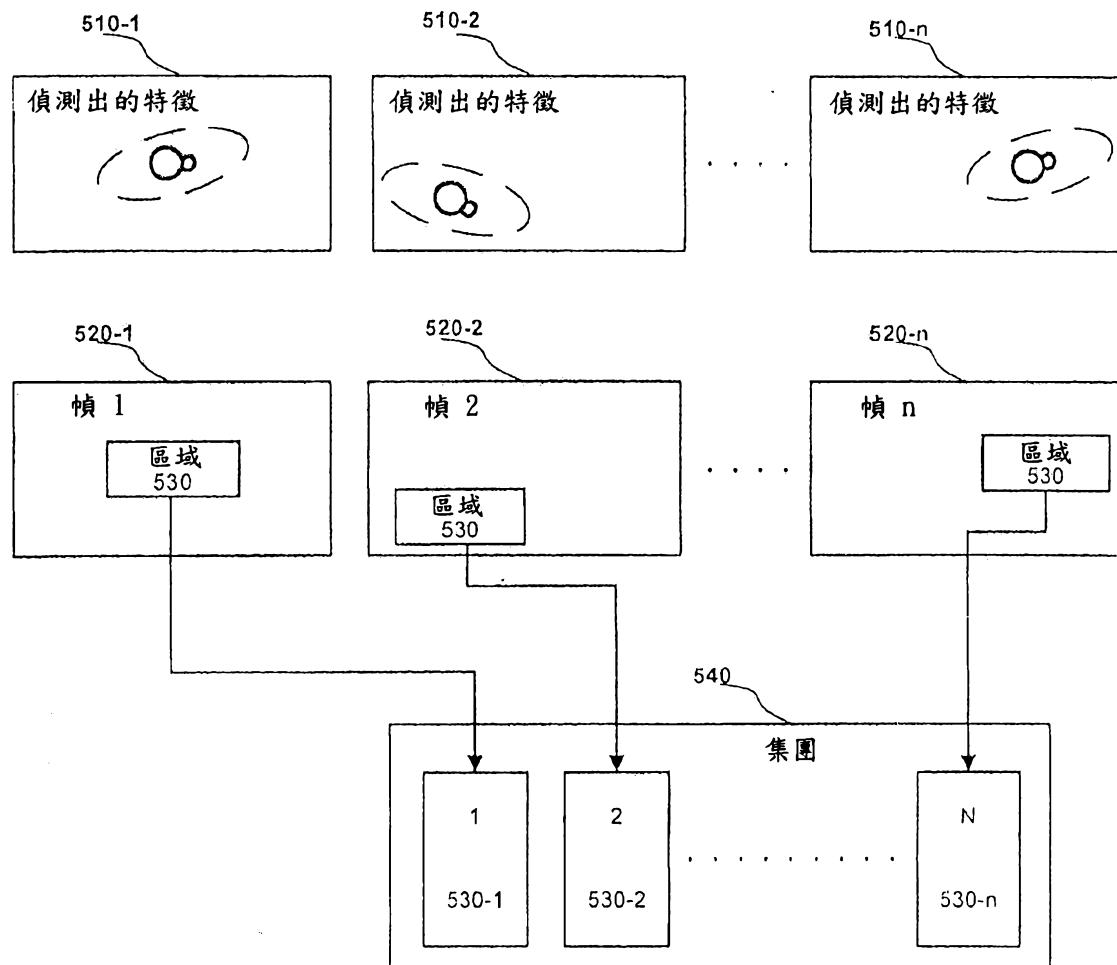


圖 5

201016016

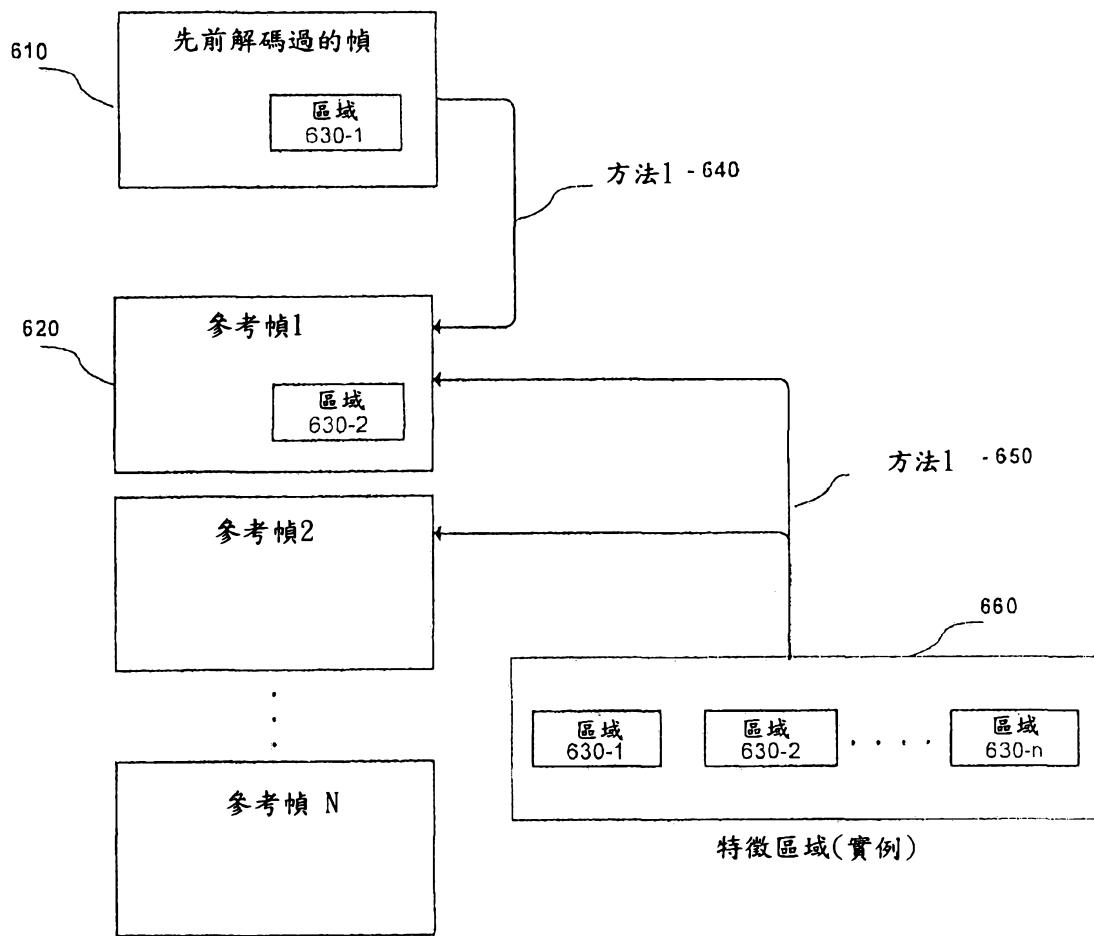


圖6

201016016

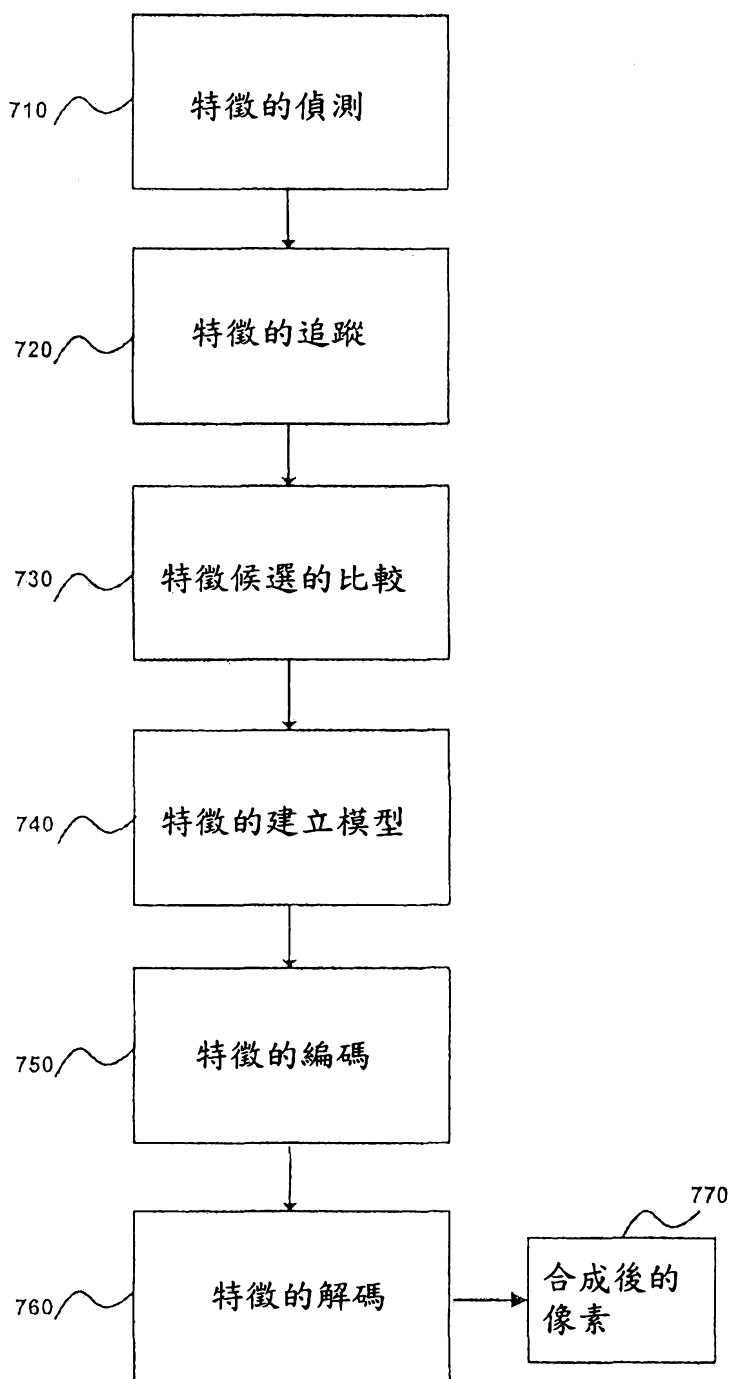


圖 7

201016016

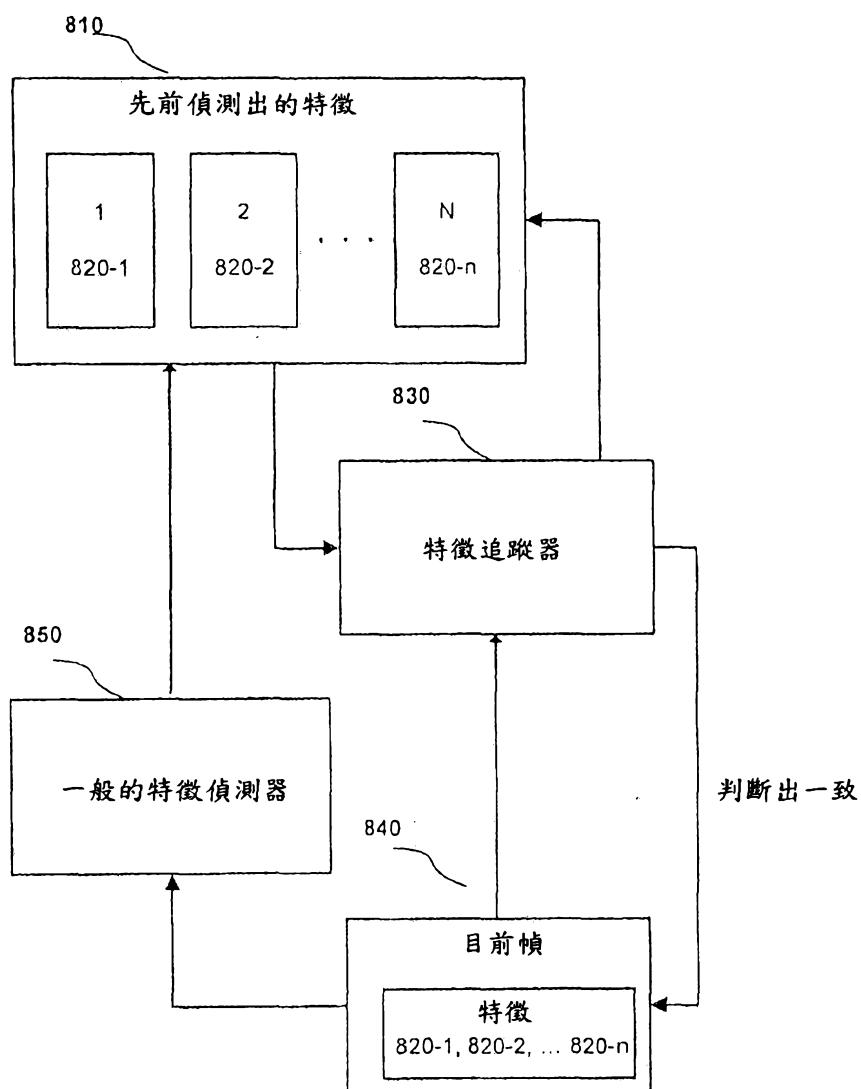


圖8

201016016

四、指定代表圖：

(一)本案指定代表圖為：第（無）圖。

(二)本代表圖之元件符號簡單說明：

無

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

201016016

98年1月26  
開光

## 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：98133799

※申請日：98.10.6      ※IPC 分類：H04N 7/26

(2006.01)

### 一、發明名稱：(中文/英文)

基於特徵的視訊壓縮

FEATURE-BASED VIDEO COMPRESSION

### 二、中文發明摘要：

處理視訊資料的系統及方法係被提出。具有一系列視訊幀的視訊資料係被接收並且處理。一個候選特徵的一或多個實例係在該些視訊幀中被偵測出。藉由採用在一個運動補償的預測過程中所用的參考幀處理，該候選特徵的實例係與該些視訊幀中的非特徵切割開。該運動補償的預測過程係選出具有對應於該候選特徵的一或多個實例之特徵的先前解碼過的視訊幀。該些先前解碼過的視訊幀係被處理以識別出該候選特徵之可能的符合者。當先前解碼過的視訊幀中有大量的部份包含該候選特徵的實例時，該候選特徵的實例係被聚集成為一組。該候選特徵組係被用來產生一個基於特徵的模型。該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型。該外觀變化模型是藉由對於該候選特徵的實例的變化建立模型而被產生的。該系統判斷出與利用該基於特徵的模型以對於該候選特徵組建立模型相關連之壓縮效率的程度、以及與利用傳統的視訊壓縮以對於該候選特徵組建立

## 六、發明說明：

### 【發明所屬之技術領域】

本發明大致上係關於數位信號處理之領域，且特別是關於用於信號或影像資料、以及特別是視訊資料之有效率的表示和處理之電腦設備和電腦實施方法。

### 【先前技術】

#### 預測切割 [主要的]

例如是 MPEG-4 與 H.264 的傳統視訊壓縮係具有能力來指明一些參考幀以在運動補償的預測過程期間使用，以便於預測目前幀。這些標準典型是限制參考幀為一或多個連續的過去的幀，並且在某些情形中是限制為先前已經解碼過的幀的任何集合。通常，參考幀的數目是有限制的，並且該選擇過程在解碼過的幀的串流中可取出多久以前的幀也是有限制的。

#### 壓縮感測 (CS)

影像及視訊壓縮技術一般是嘗試利用資料中的冗餘，而容許該資料中最重要な資訊能夠用“小”數目的參數加以捕捉。“小”是相對於原始未處理的資料大小所定義的。對於一個特定的資料組而言，事先並不知道哪些參數將會是重要的。因為此緣故，習知的影像/視訊壓縮技術在選出那些將會產生最小編碼的參數之前，都會計算(或量測)相當大數目的參數。例如，JPEG 及 JPEG 2000 影像壓縮標準是根據

線性轉換(典型是離散餘弦轉換[DCT]或是離散子波轉換[DWT])，其係轉換影像像素成為轉換係數，產生數目上是等於原始像素數目的轉換係數。在轉換空間中，重要的係數可接著藉由各種技術選出。一個例子是純量量化。當其推到極致而言，這等同於大小的臨界值法(thresholding)。儘管DCT及DWT可以有效率地被計算，但是在資料縮減前需要計算完整的轉換會造成效率不佳。這兩種轉換的計算所需要的量測數目是等於輸入資料大小。此習知的影像/視訊壓縮技術的特性使得它們在需要高的計算效率時的利用變成是不切實際的。

習知的壓縮容許混合來自多個幀的多個相符處以預測目前幀的區域。該混合通常是線性的、或者是相符處的一個取對數大小的線性組合。此種雙預測方法是有效的情形之一個例子是當有從一個影像至另一影像隨著時間而淡入淡出(fade)時。該淡入淡出的過程是兩個影像的線性混合，並且該過程有時可利用雙預測而有效地建立模型。再者，MPEG-2內插的模式容許線性參數的內插，以在許多幀上合成該雙預測模型。

習知的壓縮容許指明用於目前幀的編碼的預測所能夠抓取的一或多個參考幀。儘管參考幀與目前的幀典型是時間上相鄰的，但也可以包含該組時間上相鄰的幀以外的參考幀。

相對於以習知的轉換為基礎的影像/視訊壓縮演算法，壓縮感測(CS)演算法在量測(“感測”)步驟期間是直接利用

資料中大部份的冗餘。在時域、空間域、以及頻譜域中的冗餘是較高壓縮率的一項主要的貢獻因素。所有壓縮感測演算法的重要結果是一可壓縮的信號可以在相當小數目的隨機量測且遠小於習知的壓縮演算法所需的數目下被感測出。影像於是是可以正確且可靠地被重建。在給定已知的統計特徵之下，視覺資訊的一個子集合係被用來推出資料的其它部份。

在一個特定的 CS 演算法中所需的量測之精確數目會是依據信號的類型以及從該量測(係數)重建該信號的“回復演算法”而定。請注意的是，一個 CS 演算法在某種確定程度下重建信號所需的量測數目並非直接相關於該演算法的計算複雜度。例如，一種使用 L1-最小化以回復信號的 CS 演算法類型需要相當小數目的量測，但是該 L1-最小化的演算法是非常慢的(非即時的)。因此，實際的壓縮感測演算法會尋求所需的量測數目與重建的正確性及計算的複雜度之間的平衡。相較於習知的編解碼器，CS 係提供一種完全不同的編解碼器設計的模型。

一般而言，在一個典型的 CS 演算法中有三個主要步驟：(1)產生量測矩陣  $M$ ；(2)利用該量測矩陣以進行資料的量測，亦以產生該資料的編碼著稱；以及(3)從該編碼回復原始的資料，亦以解碼步驟著稱。該回復演算法(解碼器)可能是複雜的，並且因為在接收器之計算能力的限制較少，所以整個 CS 演算法通常是以其解碼器來命名。在習知技術中有三個受關注的 CS 演算法的實際應用：正交匹配追

蹤(OMP)、L1 最小化(L1M)以及鏈結追蹤(CP)。一般而言，L1M 實際上對於大多數的視訊處理應用而言是計算上過於效率差的。較有效率的 OMP 及 CP 演算法提供大致與 L1M 相同的益處，並且就此而論，它們正是可用於大部份 L1M 的應用之兩個 CS 演算法的選項。

經由逆合成(Inverse Compositional)演算法的影像對齊 Basri 及 Jacobs(“朗伯(Lambertian)反射及線性子空間”，IEEE 模式分析與機器智能匯刊，2/03)，以下稱為 LRLS，已經展示朗伯物件(其表面在所有方向上反射光)可藉由 LRLS“基礎影像”的一個小的(9 維的)線性子空間根據球面調和函數而良好地予以近似。該 LRLS 基礎影像可被想像成物件在不同的照明條件及紋理下的版本。LRLS 基礎影像因此是依據物件的結構(透過其表面法線)、物件在其不同的反射點處之反照率(albedo)、以及照明模型(其係遵循朗伯餘弦定律，在方向上積分以產生球面調和函數)而定的。在該模型的假設下，該 9D 的子空間係捕捉物件影像中超過 99%的能量強度。外觀子空間的低維度係表示資料中有比習知的壓縮方式可得的更大的冗餘。

逆合成演算法(IC)是首先被提出作為用於 2D 運動估計及影像套合的 Lucas-Kanade 演算法之一種有效率的實施方式。後續的實施方式已使用該 IC 演算法以將例如是主動外觀模型及 3D 可形變模型(3DMM)的 3D 模型契合至臉部影像。

### 增量奇異值分解(ISVD)演算法的應用

一般的維度縮減技術都牽涉到利用在保範的 (norm preserving)基礎下的線性轉換。一個 SVD 表示的縮減係指刪除該 SVD 中的某些奇異值/奇異向量對，以產生資料的一個在計算及表現上較有效率的表示。最普遍的是，SVD 分解係藉由將低於某個臨界值的所有奇異值歸零並且刪除對應的奇異向量而有效地予以簡化。此種大小的臨界值法係產生一個具有  $r$  個奇異值 ( $r < N$ ) 之簡化的 SVD，從一個 L2-範數 (norm) 角度來看，其係資料矩陣  $\mathbf{D}$  之最佳的  $r$  維近似。該簡化的 SVD 係得自

$$\widehat{\mathbf{D}} = \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^T, \quad \text{方程式 1}$$

其中  $\mathbf{U}_r$  是  $M \times r$ ， $\mathbf{S}_r$  是  $r \times r$  對角線，並且  $\mathbf{V}_r$  是  $N \times r$ 。

該奇異值分解 (SVD) 是一個資料矩陣的分解，其係自然導致有資料的小(小型)的描述。給定一個大小  $M \times N$  的資料矩陣  $\mathbf{D}$ ，該 SVD 分解係由  $\mathbf{D} = \mathbf{U}^* \mathbf{S}^* \mathbf{V}'$  得出，其中  $\mathbf{U}$  是(左)奇異向量的一個  $M \times N$  行正交矩陣， $\mathbf{S}$  是一個具有沿著對角線的奇異值 ( $s_1, s_2, \dots, s_N$ ) 之  $N \times N$  對角線矩陣，並且  $\mathbf{V}$  是(右)奇異向量的一個  $N \times N$  正交矩陣。

### 緊緻分歧 (compact manifold) 預測

匹配追蹤 (MP) 是一種用於導出有效率的信號表示的遞迴演算法。給定用基礎函數 (不必然正交的) 的函數庫  $\mathbf{D}$  來代表一個信號向量  $\mathbf{s}$  的問題，MP 係經由在此所述的遞迴過程來選擇用於該表示的函數。在該表示中的第一基礎函數

(標示為  $d_1$ )係被選為與信號向量有最大關聯的函數。接著，一個殘餘向量  $r_1$  係藉由將  $d_1$  到該信號上的投射從該信號本身減去而計算出： $r_1 = s - (d_1' * s) * d_1$ 。接著，在該表示( $d_2$ )中的下一個函數係被選為與該殘餘的  $r_1$  有最大關聯的函數。 $d_2$  到  $r_1$  上的投射係從  $r_1$  減去以形成另一殘餘  $r_2$ 。相同的過程接著被重複，直到該殘餘的範數下降到低於一特定的臨界值為止。

正交匹配追蹤(OMP)係依循與 MP 相同的遞迴程序，除了會採取一個額外的步驟以確定該殘餘與已經在該表示的集團(ensemble)中的每個函數都是正交的。儘管該 OMP 遞迴是比 MP 中的複雜，但是其額外的計算確保 OMP 在不超過  $Nd$  個步驟下收斂至一個解，其中  $Nd$  是函數庫 D 中的函數數目。

## 【發明內容】

本發明係延伸傳統的視訊壓縮，尤其是其中視覺現象的冗餘超出習知的視訊編解碼器的建立模型功能的情形。本發明係延伸並且可以完全取代傳統的視訊壓縮之現有的方法，其藉由採用強健的電腦視覺及圖形識別演算法。明確地說，本發明係包含特徵的建立模型方法及系統，其專注於出現在視訊的先前解碼過的幀的一或多個幀中的特徵的切割、正規化(normalization)以及整合。相較於習知的壓縮僅考量較少的幀、較小的區域、以及較少的像素，基於特徵的視訊壓縮係考量更多數量的先前解碼過的幀，並且

在那些幀的每個幀內考量更大的區域以及更多數目的像素。

習知的壓縮係藉由利用多個參考幀、大區塊(macroblock)分割、次大區塊分割、以及運動補償的預測以提供在大區塊層級下的一種隱含的切割形式。再者，習知的壓縮係利用運動補償的預測以對於發生在視訊中的空間變形建立模型，並且轉換編碼以對於外觀變化建立模型。本發明係將不同的信號元素的這些建立模型的技術延伸到更複雜的模型，其包含空間切割遮罩、規律網格變形、特徵仿射(affine)運動、三維的特徵運動、三維的照明、以及其他其他的電腦視覺及圖形識別之建立模型的技術。請注意的是，在本文中，“個別的模式”以及“不同的信號元素”是等同的。

本發明係促進視訊信號之個別的模式的識別及切割。在習知的運動補償的預測中所用的參考幀處理的概念係被利用在本發明中以助於此識別及切割。該習知的運動補償的預測過程係在大區塊層級下從一或多個參考幀選擇信號的部份。請注意的是，該習知的運動補償的預測過程通常是根據某種率-失真(rate-distortion)尺度來進行此種選擇。本發明能夠對於過去的幀施以分析，以判斷出將會具有最高機率能夠提供與目前幀的相符處之幀。此外，參考幀的數目可以是遠大於習知的壓縮中典型可見的一至十六個參考幀最大值。根據系統資源而定，參考幀數目可以盡量的高，而達到系統記憶體的極限為止；此係假設在那些幀中

有足夠數目之有用的相符處。再者，由本發明產生的資料的中間形式可以減少儲存同樣數目的參考幀所需要的記憶體量。

在一個實施例中，本發明係根據此種參考幀處理來推斷出該視訊信號的切割。在目前幀中的大區塊(像素的區塊)可以透過該運動補償的預測過程，從先前解碼過的幀選擇像素的影磚(tile)，使得該些影磚在空間及時間上都是分開的，此表示用在該運動補償的預測過程的影磚來源可以是來自不同的幀。從不同的幀選擇來源影磚以用於預測目前幀的大區塊所隱涵的分開是指出不同的信號模式被識別出的可能性。當識別出的個別的信號模式可用一種更小的方式編碼時，此進一步證實個別的模式已經被識別出。在本發明中，這些個別的模式係被稱為“特徵(feature)”。當這些特徵在視訊的許多幀上是持續不斷的並且該些特徵可以相關聯時，在該視訊中一種新類型的冗餘已經被識別出。本發明透過外觀及變形模型的產生來利用此冗餘，以便於產生超過習知壓縮所可得的進一步壓縮。再者，由於特徵是在參考幀內被識別出，因此參考幀處理會受影響而朝向利用包含特徵的參考幀；此係導致提高該參考幀處理將會進一步產生該信號中存在的模式切割之可能性。

可提供系統及方法以用於處理視訊資料。由一系列的視訊幀所構成的視訊資料可被接收及編碼。一個候選特徵的一或多個實例可在一或多個視訊幀中被偵測出。候選特徵的偵測係牽涉到判斷實例在該一或多個先前解碼過的視

訊幀中之位置資訊。該位置資訊係包含一個幀號碼、在該幀中的一個位置、以及該實例的一個空間的周圍。該候選特徵可以是一組一或多個偵測出的實例。一個運動補償的預測過程可被用以利用一或多個先前解碼過的視訊幀來預測在該系列中的一個目前視訊幀的一部份。該運動補償的預測過程可用位置的預測而被初始化。該位置的預測可從在先前解碼過的視訊幀中之偵測出的特徵實例來提供位置資訊。該些實例中的一或多個實例可藉由增加該運動補償的預測過程而被轉換。一個特徵以及該轉換後的實例可被界定出來。該些實例中的該一或多個實例可利用一種線性轉換而被轉換。包含該些轉換後的實例之界定出的特徵可被利用以產生一個第一基於特徵的(feature-based)模型。該第一基於特徵的模型可以使得在該目前幀中能夠預測一個實質相符的特徵實例之出現以及一個來源位置。較佳的是，該實質相符的特徵是利用一個率-失真的尺度判斷出之最佳的相符處。該實質相符的特徵實例可以是一個關鍵的特徵實例。該關鍵的特徵實例可以是該第一基於特徵的模型之目前幀的特徵實例接合(synthesis)。該第一基於特徵的模型可以和該一或多個界定出的特徵之一習知的視訊編碼模型相比較，並且該比較可被利用來決定哪一個模型能夠達成較大的編碼壓縮。該比較及判斷步驟的結果可被利用來導引該編碼過程來施加基於特徵的編碼至該些視訊幀中的一或多個視訊幀的部份，並且施加習知的視訊編碼至該一或多個視訊幀的其它部份。

一個候選特徵的一個實例可藉由識別出具有實質接近且空間上鄰近的像素的一個空間上連續的群組而被偵測出。該些識別出的像素可被用來定義該一或多個視訊幀中的一個視訊幀的一部份。該群組的像素可包含一或多個大區塊、或是一或多個大區塊的部份。

該運動補償的預測過程可被用來從複數個候選特徵實例中選擇一或多個被預測可提供編碼效能的實例。該候選特徵的目前的實例之切割可從該目前視訊幀中之其它特徵及非特徵來決定之。該切割可以是根據該運動補償的預測過程從獨特的先前解碼過的視訊幀所得的預測之選擇而定。該運動補償的預測過程可利用屬於一或多個特徵(此種特徵在該目前幀中具有與該視訊部份符合的實例)的特徵實例之位置資訊而被初始化，其中該視訊部份是在該目前幀中，並且該位置資訊係對應於與在先前解碼過的視訊幀中的相同特徵相關連的特徵實例。

一個第二基於特徵的模型可被形成。該第二基於特徵的模型的形成可利用該第一基於特徵的模型作為從一或多個特徵實例得到一或多個運動補償的預測之一個預測的目標。此第二基於特徵的模型產生該第一基於特徵的模型的一組預測。一旦該組預測與該第一基於特徵的模型結合，則該組預測可以成為該第二基於特徵的模型。該第二基於特徵的模型可被利用以對於來自第一基於特徵的模型之殘餘(residual)建立模型。結構的變化及外觀的變化可以從該第二基於特徵的模型相對於該殘餘來建立模型。該殘餘可

利用該特徵實例來加以編碼，此係產生外觀及變形的參數。該些參數可被利用以縮小該殘餘的編碼大小。

一或多個特徵可包含一或多個聚集的 (aggregate) 特徵。該些聚集的特徵是根據該候選特徵的實例中之一或多個實例而定的。該些聚集的特徵可藉由聚集不同的候選特徵的實例成為一個聚集的候選特徵而被產生。該些聚集的候選特徵的實例組可被利用以形成一個實質大於未聚集的候選特徵之原始實例的區域。該較大的區域可透過在該組中該候選特徵的實例之間的連貫性 (coherency) 的識別而被形成。連貫性可被定義為藉由一個較少的參數運動模型來實質近似的該些實例中的外觀一致。該第二基於特徵的模型可提供一個與該解碼過的幀中的該實例相對於該空間的位置相關連之選配的矩形像素區域範圍。該第二基於特徵的模型可藉由對於該特徵之先前正規化的實例建立模型而被導出。該些先前正規化的實例可以是下列的任一個：在該目前幀中的實例；來自一個時間上實質最近之先前解碼過的幀的一個實例；或是來自該些先前解碼過的視訊幀的實例之平均。

該外觀模型可藉由該些正規化的第二基於特徵的模型實例的 PCA 分解來表示。一個變形模型可利用每組的特徵實例中之一致處相較於其第二基於特徵的模型實例的空間變化而被決定。對於在該組中的每個特徵實例而言，下列的一或多個可被利用來對於該變形模型近似該些變形實例中的變化：一個運動補償的預測過程；網格變形；以及一

個具有實質減少參數化的運動模型。該些變形實例可被整合成為該變形模型。在該變形模型中的變化可藉由 PCA 分解來表示。

外觀參數及變形參數可被預測。該些預測出的參數可在該目前的實例利用一個基於特徵的模型的接合期間被利用。該外觀及變形模型與時間上最近的參數可被利用來從該基於特徵的模型內插及外推(extrapolate)參數，以預測該目前幀中的像素。該些時間上最近的特徵實例之接合的值可根據已產生那些實例最正確的近似之方法而被線性內插或是線性外推。該模型的實際參數可以是選配地相對於該些預測出的參數來做不同的編碼。

相較於習知的視訊資料編碼，該運動補償的預測過程可以在實質較大數目的先前解碼過的視訊幀之一個選集上做運算。先前解碼過的視訊幀的選集並不需要依賴使用者的監督。

習知的視訊編碼可增加一個實例預測過程，其係在形成該目前幀的部份的預測時，在記憶體中致能對於該些視訊幀中的一或多個視訊幀的部份之更大的壓縮。該實例預測過程可使用該基於特徵的模型以判斷該界定出的特徵出現在一個正被編碼的目標大區塊的一或多個實例。以此種方式，該實例預測過程可以產生該目前幀之預測出的部份。該基於特徵的模型可被利用以合成像素來預測該目前幀的部份。

可對於該些先前解碼過的視訊幀指定一個機率。該機

率可以是根據針對該幀之組合預測出的編碼效能改善而定的，而該效能改善則是利用來自該運動補償的預測過程之位置的預測來加以決定的。該機率可被定義為運動補償的預測過程之組合的編碼效能，其係在針對該目前幀分析該第一基於特徵的模型以及一個第二基於特徵的模型期間被利用。一項從最佳到最差的根據分類該些先前解碼過的視訊幀而定之指標可依據其機率而被產生。該分類後的表列可根據計算及記憶體的條件而被截短。

一個基於特徵的模型可利用該些界定出的特徵中之一或多個而被形成。該基於特徵的模型可包含該些界定出的特徵之位置資訊。該位置資訊可包含來自該些先前解碼過的視訊幀之界定出的特徵的位置及空間的周圍。例如，該位置資訊可包含有關在一個特定幀之內的區域的空間的位置、以及該區域在該幀中的一個矩形範圍的資訊。該基於特徵的模型可以指明哪些先前解碼過的視訊幀(或是其部份)是與該界定出的特徵相關連。

該些界定出的特徵可從該視訊資料利用大區塊運動補償的預測而被正規化並且切割出。該些界定出的特徵可利用該基於特徵的模型而被正規化。該大區塊運動補償的預測可使用在該先前解碼過的影像幀中之特徵位置作為一個位置的預測。該所產生的正規化係提供該特徵在該目前視訊幀中的預測。

該基於特徵的模型可以和產生自相同的視訊資料之習知編碼的另一模型相比較。該比較可被利用以判斷哪個模

型能夠得到較大的編碼壓縮效率。不同的編碼技術可根據該編碼比較的結果而被應用到該視訊資料不同的部份。以此種方式可提供差異性的編碼，以使得該系統能夠對於視訊資料的每個部份都根據基於特徵的編碼或是習知為主的編碼中的何者是提供較大壓縮效率來選擇不同的視訊編碼方式。

一個界定出的特徵可被表示為該特徵在一或多個視訊幀中的一組實例。每個實例可包含：指向一個其中出現該實例的幀之一個參照；一個與該實例在該幀中相關連之空間的位置；以及一個選配的與該實例在該幀中相關於該空間的位置相關連之矩形像素區域範圍。該空間的位置可以提供用於編碼該些視訊幀中的一或多個視訊幀的部份之相符處的預測。可對於每個界定出的特徵都提供一個外觀模型，以對於該界定出的特徵在該組中的不同實例之間的變化建立模型。該外觀模型可藉由對於該特徵之先前正規化的實例建立模型而被導出。該先前正規化的實例可利用運動補償的預測過程、網格變形、以及參數減少的運動模型建立(例如仿射)的任意組合而被正規化。

該正規化可被利用以建立一個變形模型，該變形模型可被利用以對於在每組的特徵實例中的一致處之空間變化建立模型。對於在該組中的每個特徵實例而言，下列的一或多個可被利用以判斷出該變形模型的變形實例：一個運動補償的預測過程、網格變形、以及參數減少的運動模型建立。該些變形實例可被整合成為該變形模型。該變形模

型可藉由利用主要成份分析(PCA)的分解而被表示。該變形模型可藉由利用任意分解的演算法之分解而被表示。相較於習知的視訊資料編碼，該運動補償預測過程可以在不需監督下，在實質較大數目的先前解碼過的視訊幀上運算。

該習知的視訊編碼可包含運動補償的區塊為主的壓縮。該習知的視訊編碼可增加一個殘餘縮減過程，其係在形成一個殘餘幀時，能夠達到在記憶體中該些視訊幀的部份的最大壓縮。該殘餘縮減過程可包含該基於特徵的模型以判斷該界定出的特徵出現在一個正被編碼的目標大區塊中的一或多個實例，以形成該殘餘幀。像素可利用該基於特徵的模型而被合成，以預測該殘餘幀。該基於特徵的模型可被利用於參考幀的索引(index)預測。該些合成後的像素可響應於判斷出該界定出的特徵的一或多個實例在該目前幀中重疊超過一個大區塊，而被再利用於其它殘餘縮減。當該界定出的特徵的一或多個實例實質相符在該目前幀中的一個大區塊之位置資訊時，該些合成後的像素可響應於判斷出該界定出的特徵的一或多個實例係代表一個大區塊，而被再利用於其它殘餘縮減。外觀及變形可根據該基於特徵的模型來加以建立模型。該外觀模型及變形模型以及在該些模型中的一組歷史的參數可被利用以從該基於特徵的模型內插及外推參數，以預測在該目前幀中的像素。再者，更高階的二次以及甚至擴張型卡爾曼(Kalman)濾波器模型都可被利用以預測該些外觀及變形參數。從該基於特徵的模型預測該些參數係能夠達成在該些殘餘參數

大小上的縮減，產生較低的精確度，且因此對於預測在該目前幀中的像素所需的參數產生較低的位元率表示。

來自一或多個幀的一或多個大區塊可利用該運動補償的預測過程來加以選擇。在一個 PCA 模型中，來自大區塊的像素可以是線性組合的像素，並且該些 PCA 模型參數可被內差得出。任何分解的演算法也可等同地被利用來取代 PCA，並且根據其實質相對的益處而被利用。

實質很小的空間區域可在該些視訊幀中被識別出。連貫性的標準可被利用以識別出可被結合成為實質較大的空間區域之空間區域。對於一個較大的空間區域而言，該較大的空間區域成為一個界定出的特徵之合適性可藉由編碼該較大的空間區域的一個基於特徵的模型來加以判斷。該較小區域可以是一個界定出的特徵，而且該較大的區域也可以是一個界定出的特徵。

基於特徵的壓縮可包含物件為主的壓縮過程。物件為主的偵測、追蹤及切割都可在該目前幀中或是在先前解碼過的幀中被應用到一個特徵實例。該特徵實例的一個中間形式可利用空間的切割而被導出。例如，該空間的切割過程可從非物件的背景切割出一個前景物件。當物件存在於一個幀中，到其出現在下一個幀中時，該所產生的切割可提供在一個特徵實例中的一個特定的物件之像素層級的一致。與該物件相關連的像素資料被再取樣，並且接著該被再取樣出的像素資料的空間位置係利用模型而被回復。該再取樣係有效地正規化從一個幀到下一個幀的物件像素資

料，並且產生提供一種中間形式的視訊資料，其具有在視訊處理上計算及分析的優點。以此種方式，物件為主的正規化以及建立模型的過程可在該基於特徵的編碼過程期間被應用到在該目前幀中或是在先前解碼過的幀中之一個特徵實例(或是其部份)。一致處的建立模型、變形的建立模型、外觀的建立模型、輪廓的建立模型、以及結構的建立模型都可被利用來對於在該目前幀中或是在先前解碼過的幀中的一個特徵實例(或其部份)建立模型。

一個界定出的特徵可能會沒有與顯著的實體(物件、子物件)一致處。例如，該些顯著的實體可透過對於偵測出的特徵監督標示為屬於或是不屬於一個物件而被判斷出。該些界定出的特徵可包含兩個或多個顯著的物件的元素、背景、或是該些視訊幀的其它部份。一或多個特徵可構成一個物件。再者，一個界定出的特徵可能不會對應於一個物件。一個界定出的特徵可能不內含在任何物件中。以此種方式，基於特徵的壓縮可以比物件為主的偵測更為有彈性及變化。儘管界定出的特徵可包含物件且內含在物件中，但是界定出的特徵並不必要是物件為主的，而是可採任意形式。

在另一實施例中，壓縮感測(CS)係被應用至該基於特徵的編碼技術。CS 係被應用至該些具有有效的或是界定出的特徵之視訊幀中的像素。CS 亦可被應用到對於該些視訊幀的其它像素的習知編碼。該視訊資料可被做成是稀疏的，以增加 CS 應用的有效性。在模型的形成(外觀及變形模型)

期間，CS 可被應用以解析來自部分參數量測的模型參數。

CS 可被應用到該第二基於特徵的模型預測的殘餘。CS 的應用可利用該平均外觀作為一項量測，並且從該量測預測該視訊信號。與該 CS 預測相關連的變異可從該第二基於特徵的模型被移除。該基於特徵的模型可被利用來專注於其餘的部份之更小的編碼。CS 編碼可被應用到該一或多個視訊幀中之其餘的像素且應用到其餘的視訊幀。

一個混合的編解碼器的解碼器可被設置，其係使用基於特徵的解壓縮以用於解碼視訊資料。編碼的視訊資料可藉由在大區塊層級判斷在該編碼的視訊資料中是否有一個編碼的特徵而被解碼。該編碼的特徵可包含基於特徵的模型。於不存在編碼的特徵之情形中，在該編碼的視訊資料中之大區塊可利用習知的視訊解壓縮而被解碼。在一個編碼的特徵確實存在之情形中，該解碼器可藉由從該編碼的視訊資料分出該些特徵編碼的部份來響應在該編碼的視訊資料中偵測到一個編碼的特徵。藉由分出該些特徵編碼的部份，該系統係能夠與該視訊流中傳統編碼的部份分開地合成該編碼的特徵。來自該編碼的特徵部份之特徵參數可以與該編碼特徵內含的特徵模型相關連。該些特徵參數可被該解碼器利用以合成該編碼的特徵。該視訊資料之傳統壓縮的部份可以與該合成後的特徵結合，以重建該原始的視訊幀。

在另一實施例中，一個視訊編解碼器係能夠處理複數個壓縮的視訊信號模式。在其中一個視訊信號模式中，一

一個編解碼器的編碼器係提供基於特徵的視訊壓縮。在另一模式中，該編解碼器的編碼器係提供傳統的視訊壓縮。類似地，一個編解碼器的解碼器係響應於不同的視訊信號模式且能夠根據該視訊信號的內容(例如該視訊信號模式)提供基於特徵的視訊解壓縮以及傳統的視訊壓縮。

該編解碼器可根據基於特徵的編碼或是習知為主的編碼中的哪一種會提供該視訊信號的視訊幀中之一或多個特徵較大的壓縮效率來決定哪一種類型的視訊壓縮是適當的。

## 【實施方式】

### 序言段落

本發明的範例實施例的說明係如下。

所有在此引用的專利、公開的申請案以及參考文獻的教示都以其整體而被納入作為參考。

### 數位處理環境及網路

較佳的是，本發明是實施在一個軟體或硬體環境中。一個此類的環境係被展示在圖 3 中，其描繪一個其中本發明可被實施的電腦網路或是類似的數位處理環境。

客戶電腦/裝置 350 以及伺服器電腦 360 係提供執行應用程式與類似者的處理、儲存以及輸入/輸出裝置。客戶電腦/裝置 350 亦可以透過通訊網路 370 連結至其它計算裝置，其包含其它的客戶裝置/處理 350 以及伺服器電腦 360。

通訊網路 370 可以是一個遠端存取的網路、一個全球的網路(例如，網際網路)、一個遍及全球的電腦集合、本地區域或廣域網路、以及目前使用個別的協定(TCP/IP、藍芽、等等)之閘道的部份，以彼此通訊。其它的電子裝置/電腦網路架構也是適當的。

圖 4 是圖 3 的電腦系統中之一部電腦(例如，客戶處理器/裝置 350 或伺服器電腦 360)的內部結構圖。每個電腦 350、360 係包含一個系統匯流排 479，其中匯流排是一組實際或虛擬的硬體線用於在電腦或處理系統的構件之間做資料傳輸。匯流排 479 實質是一個共用的通道，其連接電腦系統之不同的元件(例如，處理器、碟片儲存、記憶體、輸入/輸出埠、等等)，而能夠達成元件間的資訊傳輸。附接至系統匯流排 479 是用於連接各種輸入與輸出裝置(例如，鍵盤、滑鼠、顯示器、印表機、揚聲器、等等)至電腦 350、360 的 I/O 裝置介面 482。網路介面 486 係容許該電腦能夠連接至各種其它連接到一個網路(例如在圖 3 的 370 處所繪的網路)的裝置。記憶體 490 係提供被用來實施本發明的一個實施例(例如，以上詳述之混合的編解碼器、視訊編碼器壓縮碼、以及解碼器的碼/程式的常式(routine))之電腦軟體指令 492 及資料 494 之揮發性的儲存。碟片儲存 495 係提供被用來實施本發明的一個實施例之電腦軟體指令 492(等同的是“OS 程式”)及資料 494 之非揮發性的儲存。中央處理器單元 484 亦附接至系統匯流排 479，並且設置用於電腦指令的執行。請注意的是，在本文中，“電腦軟體指令”以及

“OS 程式”是等同的。

在一個實施例中，該處理器常式 492 及資料 494 是一個電腦程式產品(概括以 492 來參照)，其包含一電腦可讀取的媒體(例如，一可移除的儲存媒體，例如是一或多個DVD-ROM、CD-ROM、磁碟、磁帶、等等)，其提供用於本發明系統的軟體指令之至少一部份。電腦程式產品 492 可藉由任意適當的軟體安裝程序，即如同此項技術中眾所週知的安裝程序來加以安裝。在另一實施例中，該些軟體指令的至少一部份也可透過電纜線、通訊及/或無線的連線來下載。在其它實施例中，本發明的程式是一個電腦程式傳播的信號產品 307，其係體現在一傳播媒體上的一種傳播的信號(例如，在像是網際網路的全球網路或其它網路上傳播的無線電波、紅外線波、雷射波、聲波或電波)。此種載波媒體或信號係提供用於本發明常式/程式 492 的軟體指令之至少一部份。

在替代的實施例中，該傳播的信號是在該傳播的媒體上承載之一類比載波或數位信號。例如，該傳播的信號可以是在一個全球的網路(例如，網際網路)、一個電訊網路、或是其它網路上傳播的一數位化的信號。在一個實施例中，該傳播的信號係在該傳播媒體上被傳送一段期間，例如，一個軟體應用程式的指令是以封包在一個網路上被傳送一段長為毫秒、秒、分鐘、或是更長的期間。在另一實施例中，電腦程式產品 492 的電腦可讀取的媒體是一個電腦系統 350 可接收及讀取的傳播媒體，例如，藉由接收該

傳播媒體且識別一體現在該傳播媒體中之傳播的信號，即如上述用於電腦程式傳播的信號產品。

大致來說，該用語“載波媒體”或是暫態載波係包含以上的暫態信號、傳播的信號、傳播的媒體、儲存媒體與類似者。

#### 概要-基於特徵的視訊壓縮

本發明係提供一種混合的(基於特徵的及習知的)編解碼器方法(圖 1)，其具有偵測(113)、分開(115)、建立模型(117)、編碼(111)、以及解碼(124)視訊中的特徵之手段，同時容許習知的編解碼器 118 能夠編碼及解碼非特徵與無法有利地透過該特徵的編碼器/解碼器處理的特徵。圖 1 係描繪一主題視訊信號輸入(由一系列的影像幀所構成的視訊資料)110 係藉由本發明混合的編解碼器 121 而被編碼。該混合的編解碼器係包含該編碼決策的啟發法(heuristics)，並且如下處理該視訊信號：在步驟 113，特徵的偵測主要是透過在附近且呈現複雜度的像素群組的識別而達成的。複雜度通常是定義為任何能夠指出像素的編碼是超出原本藉由傳統的視訊壓縮會有效率地編碼的程度之尺度。此種在附近的像素分組係提供偵測出的特徵(在 115 處)從背景及其它的特徵切割出來。該分組係接著被分析以判斷該複雜度是否可以有利地利用本發明的特徵模型來加以建立模型(117)。

一旦特徵被偵測出且加以追蹤，並且該些特徵的模型

被產生(在 117 處)之後，該特徵的建立模型以及習知的建立模型係被比較(在比較器 119)以判斷哪一個是具有較大的益處。採用目前幀之運動補償的預測所用的參考幀處理之習知的視訊編碼機構(在 118 處)係在此過程中被利用。因為該比較(119)已經採用習知的參考幀處理，目前幀的切割係(根據來自不同的參考幀的預測選擇)被產生。在一個參考幀相對於另一參考幀中的像素(更典型為大區塊)之選擇係指出在該幀中以及在該主題視訊 110 本身中的該些特徵的切割。該主題視訊信號輸入 110 之產生的編碼 120 係包含一習知的視訊編碼串流(習知的編碼器 118 的輸出)伴隨著在該些參考幀中再現該些特徵所需之額外的編碼資訊。

該混合的編解碼器的解碼器 122 係描繪解碼該編碼的視訊以便於合成(近似)該輸入視訊信號 110。當檢查內含在該編碼的視訊中之串流資訊時，該混合的編解碼器的解碼器係在子幀的層級、亦即大區塊層級做一項判斷(123)，在該編碼的視訊中是否有一個編碼的特徵。若不存在編碼的特徵，則該習知的大區塊或是非特徵的大區塊係按傳統方式被解碼。若在該編碼的視訊串流中遇到一個編碼的特徵，則該混合的編解碼器係分開(125)該些特徵編碼的部份與傳統編碼的部份，以便於分開個別地合成，在發生接合之後組合該些部份。該混合的編解碼器係使用該編碼的特徵參數以及已由該解碼器所產生之特徵模型(與該編碼器中所做的模型一樣且平行所做的模型)以合成該特徵(124)。接著該些傳統編碼的特徵部份以及該些特徵編碼的部份係被

合成(126)以產生一個完整的特徵接合。接著該混合的編解碼器係在 127 處結合該特徵接合與該非特徵接合以產生一個完整合成的視訊信號(128)。

圖 7 是本發明的一個實施例的表示圖，其係利用基於特徵的編碼作為習知編碼的部份替代，並且在某些情形中是全部替代。偵測(710)、追蹤(720)、比較(730)、建立模型(740)、編碼(750)、以及解碼(760)一視訊信號中的特徵係被描繪。

在步驟 710，如同在 113 中，特徵的偵測主要是透過呈現複雜度之空間接近的像素群組的識別來達成的，使得像素群組可以比習知的方式更有效率地編碼/建立模式。這些像素群組有效地分開偵測出的特徵(710)與它附近之非特徵的像素，如同亦在 115 中指出者。偵測出的特徵、或是特徵實例候選、或就稱做是特徵候選係進一步被分析，以關聯兩個或多個幀上的像素群組。此關聯確認特徵實例是代表該些視訊幀中的一個可被追蹤的離散實體(720)，藉此確認該視訊中有額外的冗餘，該冗餘可透過對該特徵建立模型而可被降低(740)。在步驟 720 中，該特徵係經由在目前幀內的特徵的實例(等同於區域)以及該特徵在一或多個其它幀中的實例的識別而被追蹤，如同亦在 117 中指出者。請注意的是，在整個本文中，“特徵實例”係等同於“區域”。再者，當“實例”是參照“特徵實例”及“區域”時，“實例”係等同於“特徵實例”及“區域”。

每一個別特徵的實例都被視為候選特徵，並且透過將

它們群組成為特徵組或就是特徵，而變為組成一個完全合格的特徵。在步驟 730 中，這些實例被分析、比較並且透過在該些實例之間的一致處的識別而被分類成為特徵組。

在本文中，特徵候選以及特徵實例是等同的。特徵組係被分析以獲得特徵實例的一個變形變化及外觀變化的模型。在特徵實例之間的變形變化係透過一個變形的建立模型過程而被判斷出。該變形的建立模型過程係比較兩個或多個實例，以決定減少實例間的每個像素的差異所將會需要的空間像素的再取樣。

特徵候選係在步驟 740 中被建立模型，其係應用多個分析技術以精確調整該些取樣出的區域。視訊串流的特徵編碼(750)係利用該些特徵模型，並且部份或全部在不使用習知的視訊編碼之下編碼該視訊串流。該解碼(760)係利用該些特徵模型以建立模型的動作之倒轉(750)來合成該些特徵以解碼該些編碼的特徵成為每個特徵實例的像素的接合(770)，近似出該特徵為如同原先在該視訊中出現者。

## 預測切割 [主要的]

圖 6 展示藉由利用內含在置於一或多個參考幀中的一或多個過去的幀之內的資訊來預測在目前視訊幀中的元素的過程。在一個實施例中，該預測(方法 1 640)係複製來自一或多個先前解碼過的幀 610 的區域到一個參考幀 620 之中。方法 2 650 額外將特徵區域 630-1、630-2、…630-n 所構成的特徵實例 660 置於該參考幀中。直接插入特徵實例

到參考幀中係代表本發明的一種簡單形式，其中，在另一個實施例中，該切割就是一個矩形區域，並且該特徵的模型是該特徵實例。額外的壓縮利益可被實現，因為進一步的建立模型技術係被應用至該些識別出的特徵 660 並且用在該些參考幀之內。

預測切割是習知壓縮的運動補償的預測方法擴展到容許有更正確的預測所藉由的方法。習知的壓縮係使用利用本發明的特徵建立模型方法所產生之額外的參考幀以增高正確性。當這些特徵參考幀的部份被習知的壓縮方式利用時，若特徵編碼是小於原本的習知編碼，則在壓縮上得到利益。

在一個實施例中，特徵係被表示為一組元素或特徵實例。在一個實施例中，該些特徵實例係被實現為矩形區域，每個矩形區域提供一項參照至唯一的幀、在該幀中之一空間的位置、以及在該幀中一個矩形區域範圍。該特徵的每個實例係代表該特徵的一個取樣的影像。在實例之間的特徵外觀上的變化係藉由該特徵的建立模型方法而被建立模型。

在一個實施例中，該些參考幀係存在一或多個來自先前合成的幀之子幀樣本。該些子幀的樣本是根據在先前合成的(解碼過的)幀以及目前幀中之該些子幀的區域之間的特徵實例的一致而定的。

在另一個實施例中，多個影像平面被合併成為較少的影像平面。這些較少的影像平面係使得位在接近該幀中預

期的位置處之特徵被預測出。幀的縮減是根據合併非重疊的或是接近零的空間重疊的特徵到相同的平面中。

申請人在實務上的縮減亦已經利用此而進一步發展，其藉由根據特徵資訊(先前的相符處、追蹤資訊、建立模型的資訊)來估計該特徵的一個界限區塊。

在另一非限制性的實施例中，每個合併的幀係等於被預測的幀的大小，並且該些特徵若不是剛好在習知的運動補償的預測機構所預期的位置、就是在空間上接近於該位置。

## 特徵的偵測

圖 5 係描繪已經在一或多個幀的視訊 520-1、520-2、… 520-n 中被偵測出的特徵 510-1、510-2、… 510-n。典型的是，此種特徵將會利用根據從像素導出的結構資訊以及指出習知的壓縮是利用不相稱的資源量來編碼該特徵的複雜度標準的數個不同的標準而被偵測出。相較於特徵編碼，每個特徵可進一步藉由一個在該圖中顯示為一個“區域”530-1、530-2、… 530-n 之對應的空間範圍(周圍)而在一個幀 520-1、520-2、… 520-n 中在空間上被識別出。

這些區域 530-1、530-2、… 530-n 例如可被抽取為一個簡單的像素資料矩形區域，並且置入一個集團 540 中，整個集團係代表一個特徵。

在一個幀中的一個特徵的每個實例是該特徵的外觀的一個樣本。請注意的是，當足夠數目的這些樣本被聯合到

一個集團中時，其可被利用以對該特徵在那些幀中以及在該特徵並未從其取樣的其它幀中的外觀建立模型。此種模型係能夠轉換該外觀成為一個編碼的參數組，其可進一步透過逆(inverse)模型而被解碼，以產生該特徵的接合。

小的空間區域係被識別且分析，以判斷其是否可根據某些連貫性標準而被組合成較大的空間區域。這些較大的空間區域係接著被分析，以判斷其作為候選特徵的合適性。假設該區域的特徵建立模型並未提供有利的編碼，則該候選特徵不是被拋棄、就是被保留用於對該特徵於後續的幀之未來的實例建立模型。該偵測過程係繼續直到僅剩下該些會呈現有利的建立模型的候選特徵為止。

空間區域在大小上會有所變化，從小群組的像素或子像素，到可能對應於實際的物件或是該些物件的部份之較大的區域，即為易於透過傳統的視訊壓縮演算法所決定之大區塊或次大區塊的分割步驟而被隱含分割的區域。然而，重要的是注意到偵測出的特徵可能不會對應於例如是物件及子物件之離散唯一且可分開的實體。並不要求特徵對應於此種實體。單一特徵可包含兩個或多個物件的元素、或是完全沒有物件的元素。重要的因數是本發明具有可用高於習知方法的效率來處理這些信號成分的可能性，並且這些信號成分純粹是根據其藉由基於特徵的視訊壓縮技術而被有效率地建立模型而充分滿足一個特徵的定義。

小的空間區域可被集合成為較大的區域，以便於將這些較大的區域識別為特徵。小區域係透過識別出區域之間

的連貫性而被聚集成為較大的區域。有數種方式可識別出連貫性，其包含連貫的運動、運動補償的預測、以及編碼複雜度。

連貫的運動可透過較高階的運動模型而被發現。例如，每一個別的小區域之平移的運動係被整合成為一個仿射運動模型，其能夠對於該些小區域的每個區域近似出一個較簡單的運動模型。

若小區域的運動可在一致的基礎下被整合成為較複雜的模型，則此隱含表示該些區域之間有相依性，其潛在可提供一個優於習知的運動補償的預測方法之優點，並且此亦指出該些小區域之間的連貫性，其可透過特徵的建立模型而被利用。

編碼複雜度可透過習知的壓縮要代表一或多個該些小區域所需的頻寬之分析來加以判斷。在習知的編碼無法有效率地壓縮並且可能無法關聯為幀與幀之間的冗餘之一特定組的小區域有一個不相稱的頻寬配置之情形中，這些區域潛在可被集合成為一個特徵，其編碼複雜度可指出存在一個現象為特徵的建立模型將會較佳地表示。

一組已知的幀係分別完全被劃分成為以一種非重疊的模式配置之一致的影磚。每個影磚係被分析為一個獨立取樣的像素區域，其實際上被判斷為包含足夠用以描繪特徵的資訊。本發明使用這些取樣的區域以產生多個分類，其接著被使用在訓練一個分類器上。請注意的是，任何特徵的最後位置都可能與此最初的定位不同。

另一個實施例係從定義的影磚以及一個重疊該些影磚的磚面(tiling)來產生取樣的區域。重疊的取樣可被抵消，因而重疊的影磚的中心會出現在每四個基本的影磚之角落的交叉處。此過完備的(over-complete)分割意謂著增加一個最初的取樣位置將會產生一個偵測出的特徵之可能性。其它可能更為複雜的拓撲學的分割方法亦為預期可用的。

一個特徵建立模型的預測器係分類取樣的區域成為一個叢集(cluster)，其具有相當大的機率是一個區域將會具有與同一叢集中的其它區域的某種一致性。該特徵建立模型的預測器係使用從取樣的區域導出的模式例子。

在一個較佳實施例中，該些特徵係在來自頻譜資料搜集(描述在以下的頻譜資料搜集(profiling)章節中)的幫助下被偵測出。頻譜資料搜集係提供幀中可能是單一特徵的部份之區域。此係被使用作為一個組合取樣的區域成為一個特徵的手段。

在一個實施例中，一個模式(pattern)特徵係被定義為一個頻譜特徵。該頻譜特徵係藉由將該區域從一個色彩空間轉換到一個HSV色彩空間而被找出。該轉換後的區域係接著反覆地向下被次取樣，直到導出的區域的影像向量空間具有遠小於原始區域的影像向量空間的大小為止。這些導出的區域被視為頻譜特徵。該些頻譜特徵係利用一種改良的K-means演算法而加以群集起來。K-means叢集係被用來根據叢集的頻譜分類來標示原始的區域。

在一個實施例中，一個分類器係根據取樣的區域之邊

緣內容而被建立。每個區域都被轉換至 DCT 空間。導出的特徵之 DCT 係數係接著被加總用於上三角矩陣以及下三角矩陣。這些加總係接著被用以形成一個邊緣特徵空間。該特徵空間係接著利用 K-means 而加以群集起來，並且原始的區域係根據其導出的區域叢集的分類來加以標示。

在另一實施例中，該頻譜特徵以及邊緣模式特徵分類器係被用來產生每個區域的多個分類。

## 特徵的追蹤

一個實施例係使用新偵測出的以及先前被追蹤的特徵之組合作為基礎以用於判斷在目前幀中具有相同對應的特徵之實例。此特徵在目前幀中的實例的識別以及包含此實例與該區域的先前發生的實例係構成特徵的追蹤。

圖 8 係展示一個特徵追蹤器 830 以及新偵測出的特徵與先前被追蹤的特徵 810 的組合之使用，以追蹤及分類特徵 820-1、820-2、…820-n。最初，一個一般的特徵偵測器 850 係被用來識別特徵。一致的判斷是根據目前幀 840 與先前偵測出的特徵 810 比對而得。該些被追蹤的特徵係被組織成為特徵組、或是分類為屬於一個先前組合的特徵組或是屬於一個新的特徵組。

特徵的一致最初可透過習知的梯度下降法最小化均方誤差的估計來加以判斷。該所產生的空間位移係得出該特徵在目前幀中之大約的位置。在該搜尋中使用的樣版(template)不必是該特徵的單一區域，而可以是與該特徵關

連的區域中之任一個。最後的相符係接著以一種強健的方式評估作為非重疊的區域影磚中符合一個 MSE 臨界值的一個計數。

藉由施加一項空間的限制在兩個或多個區域的一致上，追蹤器能夠減少被追蹤的特徵數目，因此增加追蹤的計算效率。兩個或多個特徵之空間的一致亦可以指出額外的特徵情形，其中該特徵在過去實際上可能曾經是兩個特徵、或是某種其它複雜的特徵拓撲。追蹤器的模式容許有暫時簡併的(degenerate)追蹤狀態而讓特徵得以被追蹤，但會降低被追蹤的區域優先順序。

## 特徵的建立模型

在某個分析階段期間，將被預測的區域係被用來通過區域一致模型(RCM)，以便於決定在該模型之中將被用來建構一個區域預測模型的區域。

在一個實施例中，目標區域係被用來更新該 RCM，藉此產生在內含於該 RCM 之中的其它區域與該目標區域之間平移及中點正規化的一致。該所產生的成對的區域一致係識別出最可能對於該目標區域產生一個預測模型的其它區域。

本發明係將一個特定的目標區域之最佳的一致中的一或多個一致的組合內含到一組稱為一個區域集團(ensemble)中。該區域集團可被空間正規化朝向一個在該集團中之關鍵的(key)區域。在一個實施例中，時間上最接近該目標區

域的區域係被選出作為該關鍵的區域。執行這些正規化所需的變形係被收集成為一個變形集團，並且該所產生的正規化的影像係被收集成為一個外觀集團，即如在美國專利號 7,508,990、7,457,472、7,457,435、7,426,285、7,158,680、7,424,157 及 7,436,981 以及美國申請案號 12/522,322 中所述者，所有案件均為本案受讓人所有。上列的專利及申請案之整體教示係被納入作為參考。

該外觀集團係被處理以產生一個外觀模型，並且該變形集團係被處理以產生一個變形模型。該外觀及變形模型的組合係變成該目標區域的特徵模型。在一個實施例中，模型形成的方法是該集團的主要成份分析(PCA)的分解，接著是所產生的基礎向量的截斷。在另一實施例中，用於截斷的標準可以是集團內的重建。

在另一實施例中，模型(外觀及變形模型)形成的方法是在本文別處所述的壓縮感測(CS)，其中模型參數是從部分的參數量測解析而來的。

該目標區域被投射到該特徵模型上，以產生特徵參數。特徵參數係用於該區域的變形及外觀的建立模型。再者，特徵參數是該目標區域的編碼。

兩個或多個集團內的區域之特徵模型參數係利用時間的標準而被選出。給定在該些區域本身與該些目標區域之間已知的間隔，這些參數係被用來預測該目標區域的狀態。狀態模型的一個例子是在給定時間的步階下，兩個或多個特徵參數的線性外推。該線性模型係被用來預測該目

標區域的特徵參數。

若外推的值提供該目標區域的適當接合(解碼)，則載明該目標區域的特徵參數變成不必要的、或是特徵參數可相對於外推的參數而有差別地予以載明。

用於外推的狀態模型可以是比簡單的線性模型高階的。在一個實施例中，一個擴張型卡爾曼濾波器係被用來估計該特徵參數狀態。

### 區域一致模型

分類、登記以及變形分析的組合係提供一組資訊指出兩個或多個區域可被結合成為一個外觀及變形的聯合模型，稱為區域一致模型(RCM)的機率。

在一個較佳實施例中，該特徵偵測方法(如上所述)係遞增地分析新的特徵。此分析的一項結果是較高的機率為一個區域將會對應於被用來建構其中一個特徵偵測器的其它區域。

一旦區域係如同在以上討論的特徵偵測而被分類成為叢集，並且給定其個別的叢集標示後，則叢集間的區域係被分析以決定在區域對之間的每區域的一致性。

在一個較佳實施例中，上述的分類器係被用來定義取樣的區域的叢集，其區域來源像素係透過區域平移微調(在以下說明)而進一步被分析及定義。

此外，在一個較佳實施例中，在平移區域微調之後，區域的一致可進一步就其區域變形分析(在以下說明)的方

面被定義。

在一個實施例中，該 RCM 的建構係遞增地被達成。兩個或多個區域係被用來對於組合的分類器/變形分析機構提供最初的輸入。該 RCM 接著是利用改變分類器及變形分析元素的新區域而被更新。

在一個實施例中，上述的 RCM 遞增更新係被建構成使得一個特定的模型之區域的一致係根據詳述於下的基礎 (base) 複雜度分析，而以一種遍歷的 (traversal) 順序而被處理。

在一個實施例中，如上所論述之根據一種基礎複雜度分析 (在以下說明) 而定之遍歷的順序是一個利用遍歷的終止標準更新該 RCM 的遞迴過程之部份。當外觀/變形模型是從該些一致導出時，該終止標準會使得該處理完成到一個程度是最大化該 RCM 以具有最大降低複雜度的可能性來代表一致的能力。

## 區域平移微調

在一個實施例中，取樣出的區域被收集在一起而成為一組訓練的取樣區域。這些區域在每個幀中的空間位置係被微調。

微調係包含每個取樣的區域與其它各個取樣的區域完全的比較。此比較是由兩個影磚套合 (registration) 所構成。一個套合是一個第一區域對於一個第二區域的比較。第二個套合是該第二區域對於該第一區域的比較。每個套合是

在該些區域於其個別的影像中的位置處被執行。所產生的套合偏移量以及對應的位置偏移量係被保留且被稱為關聯性 (correlation)。

該些關聯性係被分析以判斷多個套合是否指示一個取樣的區域位置應該被微調。若在來源幀中的微調的位置將會對於一或多個其它區域產生較低的錯誤相符，則該區域位置係被調整成該微調的位置。

該區域在來源幀中之微調的位置是透過其它區域一致的位置的線性內插法而被決定出，該些區域一致是在時間上橫跨來源幀中的該區域。

## 頻譜的資料搜集 (profiling)

頻譜的資料搜集方法是一種統計的“均值追蹤及擬合”方法。此種方法的其它例子係被描述在如同應用到出現在影像及視訊幀中的空間機率分布之偵測、追蹤及建立模型的 CAMSHIFT、平均位移、中心點 (medoid) 位移、以及其衍生方法的文獻中。本發明之頻譜的資料搜集方法開始是分析橫跨一或多個幀的一個影像平面的一個區域的頻譜 (色彩) 平面的強度成分、像素。強度成分首先係透過該些值經由一個直方圖分箱 (binning) 方法的離散化而被處理。接著一個區域的直方圖係和一種追蹤機構一起使用以識別出在後續的幀中具有類似的直方圖之更多的對應區域。該區域的成分 (位置、離散化標準及直方圖) 組係反覆地微調，因此其收斂在這些成分的一個共同組。該微調後的成分組是該頻譜

輪廓。該頻譜的資料搜集方法是一種特徵的偵測方法。

利用一維的 K-means 分類有一項優點，因而 HSV 色彩空間的色度通道係以分類器的形式被利用。此外，該些像素係被分類並且直方圖的分箱係被填入，因而空間的不變矩 (invariant moment) 係被決定出。

用於本發明的核心基礎功能係利用已存在的資料來導出用於新資料的模型。該已存在的資料可透過任何編碼/解碼的方式而獲得並且被假設是可得到的。本發明係分析此資料以決定一組候選模式資料，其被稱為特徵資料，可包含視訊信號的一個空間局部成份的外觀及變形的資料。

給定一特定組已存在的特徵資料以及一個新的目標資料點之下，執行分析以決定建立用於代表該目標資料點的一個模型所需的特徵資料之最少量的描述。在不失一般性的意義下，該已存在的特徵資料係被稱為候選特徵向量，並且該目標資料點係被稱為目標向量。再者，該過程可應用到一或多個目標向量。

給定一個目標向量及一組候選特徵向量(全都視為相同特徵的部份)之下，該些候選特徵向量的一個最小子集合係被選出以在低誤差下合成該目標向量，此係產生一個小型且正確的分歧表示。

本發明聚集一組候選特徵向量成為所謂的特徵集團 (ensemble)。在一個實施例中，產生該特徵集團的第一步驟是選出一個關鍵的向量，亦即一個被判斷是該目標向量的一個良好近似的特徵向量。該關鍵的向量是該特徵集團中

的第一向量。該特徵集團的其它候選特徵向量係依照其與該關鍵的向量之關聯性的順序而被選出(因而該特徵集團中的第二向量是具有下一個最高的與該關鍵的向量之關聯性的特徵向量)。以此種方式排序一個特徵集團係稱為關鍵關聯性排序的(KCO)。

在另一實施例中，該特徵集團係利用該目標向量本身而被產生。該特徵集團的候選特徵向量係根據其與該目標向量的關聯性而被選出。任何使用目標向量關聯性的排序方法都被稱為目標關聯性排序的(TCO)。在一個 TCO 特徵集團中的第一特徵向量是具有與該目標向量最大的關聯性的候選特徵。在一個較佳實施例中，每次一個特徵向量“進入”該集團時，經由現有的集團( $U_r$ )之近似的目標重建係計算為  $U_r^*U_r^*t$ ，接著從該目標向量  $t$  減去以形成一個殘餘向量。該集團的下一個特徵向量接著被選擇為具有與該殘餘向量最大的關聯性的候選特徵。此種計算該殘餘向量並且接著選擇與該殘餘向量最佳的相符的遞迴過程因此被稱為順序的目標關聯性排序(STCO)。對於一個特定的集團大小而言，STCO 係確保該目標向量最有效率的表示。對於小的集團大小而言，其在功能上等同於正交匹配追蹤(請參見習知技術)，但在計算上更有效率。

在另一實施例中，並未計算殘餘向量，而且該特徵集團所有的候選特徵向量都是根據其與該目標向量本身的關聯性而被選出。此 TCO 方法，稱為全域的目標關聯性排序(GTCO)是比 STCO 快且簡單，但可能在該集團中產生冗

餘。然而，對於選出該集團而言，兩種 TCO 方法都大致上遠優於該 KCO 方法。

一個位元遮罩 (bitmask) 係被用來傳送已被選用於該特徵集團的特徵向量。

在一個實施例中，在 SVD 為基礎的編碼之前，該特徵集團中的特徵向量以及該目標向量本身係通過一種離散子波轉換 (DWT)。此係使得該目標向量中的資訊更小型且更容易藉由 SVD 向量的一個小子空間來表示。該 DWT 是一種眾所週知的方法用於多尺度縮小信號資訊。在一個較佳實施例中，該 DWT 係以 Daubechies 9-7 雙正交的子波而被應用。該 DWT 係個別地應用至每個成分，因為特徵向量係在 YUV 色彩空間中。例如，長度 384 的 YUV 向量需要在 Y 成分上的一個長度 256 的 DWT 以及在 U 及 V 成分上的長度 64 的 DWT。

## 壓縮感測 (CS)

在本發明的一個實施例中，壓縮感測 (CS) 係被採用作為在該特徵的建立模型 (在別處描述) 過程中之模型形成 (外觀及變形模型) 的方法。

在本發明中有三種所關注的 CS 演算法實際的應用：正交匹配追蹤 (OMP)、L1 最小化 (L1M) 以及鏈結追蹤 (CP)。每一種演算法都具有其本身的強項與弱點，但是該 L1M 對於大多數視訊處理應用而言都是難以接受的慢，因而在此領域中，OMP 及 CP 是兩種供選擇的 CS 演算法，而 L1M 則

較不常被使用。

CS 演算法的有效性實際上是受限於計算時間、記憶體限制、或是量測的總數。為了對抗這些限制且實際改善 CS 演算法的效能，本發明使用數種可能的方法中的一或多種。簡言之，該些方法係透過以下而達到益處：(1)減少該文獻中為了精確的重建所指定的量測數目；(2)藉由一或多種特定的資料縮減技術來增加在輸入資料中的稀疏度 (sparsity)；(3)分割該資料以減輕記憶體的限制；以及(4)適應性地將誤差的預估內建到重建演算法中。

一個實施例係利用一項事實為通常用於重建的數學要求是比必要的要求更為嚴格。利用比該文獻中所指定的較少量測來經常一致地達成“良好的”的影像資料重建是可能的。“良好的”重建是表示對於人眼而言，在視覺上相較於“完全”重建有很小的差別。例如，以一半所指定的量測數目應用鏈結追蹤 (CP) 仍然可達成“良好的”重建。

在另一實施例中，該輸入資料係“被縮減”以使得其更為稀疏，此係減少所需的量測數目。資料縮減技術包含將資料通過離散子波轉換 (DWT)，因為資料在子波域中經常是較為稀疏的；藉由截斷來實際減少輸入資料的總大小，亦已知為下取樣 (down-sampling)；以及對該資料進行臨界值法 (除去所有小於某個臨界值的成分)。該些資料縮減技術中，DWT 轉換是最不會“侵入性的 (invasive)”，並且理論上容許該輸入資料能夠完全回復。另外兩種縮減技術是“有損失的 (lossy)”，因而無法讓信號完全回復。DWT 與 CP 配合良好，

但是與正交匹配追蹤(OMP)或是 L1 最小化(L1M)配合的不好。因此，用於此資料縮減實施例的理想組合是鏈結追蹤演算法與該離散子波轉換的資料縮減技術。

在另一特別是極適合於平行處理的架構之實施例中，輸入資料是被劃分成區段(或是 2D 影像被劃分成影磚)，並且每個區段係個別地利用較小數目之所需的量測而被處理。此種方式良好適用在通常會受限於記憶體限制的 OMP 及 L1M。所需的量測矩陣的大小是造成 OMP 及 L1M 記憶體限制的原因。吾人可計算出記憶體矩陣超出系統記憶體的量。此額外的記憶體需求是一項“超取樣(oversampling)”因數。其係設定將該信號分割成區段數目的一個下限。

在另一實施例中，該過程將某種誤差的預估內建到該重建演算法中。該預估的誤差可能是由於超出正常的雜訊或是不準確的量測所引起的。該過程藉由放寬最佳化的限制、或是藉由在完成該重建過程之前停止遞迴來補償之。該重建於是為該資料的一個近似的擬合(fit)，但是此近似的解決方式可能是足夠的、或者當輸入資料是充滿雜訊或是不準確時此近似的解決方式可能是唯一可行的解決方式。

圖 2 係顯示一個概念上的視訊壓縮架構，其係在編碼器實施壓縮感測的量測。原始的視訊串流(200)係被傳送通過一個運動補償的預測演算法(202)以套合該資料(203)，藉此建立在多個幀中的像素群組之間的一致，使得由於運動所造成的冗餘可被分解出。接著，預處理(204)係被應用以使得該資料儘可能的稀疏(在 205 處)，因而 CS 量測以及之

後的重建將會是儘可能有效的。CS 量測係被取得(206)並且變成該 CS 編碼(207)(已可發送)。之後在接合期間，該 CS 演算法係被用來解碼該些量測。

本發明識別、分開及預處理來自原始的視訊串流之信號成分，而成為極適於 CS 處理的稀疏信號。CS 演算法當然是與本發明的實施例相容。應注意到的是，圖 2 的某些觀點係相關於在美國專利號 7,508,990、7,457,472、7,457,435、7,426,285、7,158,680、7,424,157、以及 7,436,981 以及美國申請案號 12/522,322 中所論述的實施例，所有美國專利案都是本案受讓人的。以上所列的專利案及專利申請案的整體教示係被納入在此作為參考。

在視訊壓縮的背景中，當輸入影像具有某種稀疏性或是可壓縮性時，CS 係給予一項重要的益處。若輸入影像是密集的，則 CS 不是用於壓縮或重建的正確解決方式。CS 演算法可用少於習知的壓縮演算法所需的量測(需要相當於影像中的像素數目的量測)來壓縮及重建稀疏的輸入影像。請注意的是，信號稀疏性或可壓縮性是大多數壓縮技術所假設有的，因此 CS 所能夠提供改良的影像是設計大多數壓縮技術所針對的影像。

同樣要注意的是，將雜訊加入稀疏的影像會使得其在數學上更密集，但是並不會使得其“在資訊上”較不稀疏的。其仍然是一個稀疏的信號，並且以上述實際的實施方式中的一或多個方式來利用 CS 可產生此類的信號之有用的重建。

## 基礎複雜度分析

代表性的取樣的視訊區域可利用一種基礎方法而被分析。一種此類的方法將會是習知以區塊為主的壓縮，例如 MPEG-4。

## 透過逆合成演算法的影像對齊

Xu 及 Roy-Chowdhury (“整合在連串視訊中的運動、照明及結構...”，IEEE 模式分析與機器智能匯刊，2007 年 5 月)係將該 LRLS 架構延伸至移動的物件(例如，在連串視訊中)，顯示此種物件是可藉由一種具有 9 個照明函數(原始的 LRLS 基礎影像)以及 6 個反映在 LRLS 基礎影像上的運動效果的運動函數之 15 維度的雙線性的基礎來加以良好的近似。

由 Xu 及 Roy-Chowdhury 近來提出的 IC 實施 (“在動態畫面中的 3D 姿態及照明的逆合成估計”，IEEE 模式分析與機器智能匯刊，其將被發表)係使用該逆合成(IC)演算法以從一序列的視訊幀中估計 3D 運動及照明參數。一個 2D 至 3D 至 2D 的翹曲(warping)函數係被用來在一個標準的(canonical)姿態下對齊來自不同的幀之(目標)影像與一個“關鍵的”幀(樣版)。給定一個幀的影像資料以及被成像的物件的一個基本的 3D 模型下，該 2D 至 3D 對映係決定在該 3D 模型中的哪些 3D 點(面/頂點)對應於哪些影像像素。一旦該 2D 至 3D 對映已經界定之後，該物件的姿態係在 3D

中位移先前的幀的姿態估計，藉此對齊目前幀與該關鍵的幀。在 3D 中位移後的物件係接著利用該 3D 至 2D(投射)的對映而對映回到 2D，以形成一個“姿態正規化的”影像幀。

一旦目標幀已經利用該 2D 至 3D 至 2D 的對映而套合到該樣版(關鍵的幀)之後，該所產生的姿態正規化的幀(PNF)係被用來估計 15 個運動參數，其對應於 9 個照明以及 6 個運動的變數。該些照明變數係經由該 PNF 至該些 LRLS(照明)基礎影像的最小平方擬合(least-squares fit)來加以預估出。在一個實施例中，藉由該些 LRLS 基礎影像預估出的照明成分係接著從該 PNF 減去，並且殘餘者係被用來經由最小平方擬合至該些運動函數以估計出 6 個運動參數(3 個平移以及 3 個旋轉)。該 PNF 接著可從該 15 維度的“雙線性的”照明/運動基礎以及其對應的參數向量而被重建。

本發明使用 Xu/Roy-Chowdhury 的 IC 實施的觀點，以幫助影像套合的應用。在一個實施例中，該 2D 至 3D 至 2D 的對映係被使用作為特徵區域的中點正規化之一個計算上有效率的替代方式。對於其中正確的 3D 模型(例如用於臉的 Vetter 模型)存在的特徵而言，該對映的過程是特別有用的。在此實施例中，該些模型點係在某種姿態(該“模型姿態”)下被指定，並且該關鍵的幀(該樣版)以及目前幀(或目標幀)係被套合至該模型姿態。

### 增量的奇異值分解(ISVD)演算法的應用

在本發明中，該 SVD 係利用共同大小的臨界值方法的

一種變化，在此稱為百分比臨界值法而被縮減。在一個實施例中，在一個特定的 SVD 分解中的奇異值的總能量  $E$  係計算為該些奇異值的總和。當奇異值依序地(以大小遞減的順序，從最大至最小)加總直到在縮減組中的奇異值的總和超過  $E$  的某個百分比臨界值為止，該些奇異值的一個分組(在本文中稱為一個“縮減組”)係被產生。除了該大小臨界值並不需要是事先已知的以外，此縮減方法係等同於大小臨界值法(請參見習知技術)。

在本發明中，該奇異值分解(SVD)係如下所述地應用至特徵資料。該  $M \times N$  資料矩陣  $\mathbf{D}$  是由一個特定的視訊影像幀的區域(影磚)所導出的特徵向量的一個集團所組成。該  $M \times 1$  特徵向量係從 2D 影像影磚而被行向量化的，並且被連鎖以形成該資料矩陣  $\mathbf{D}$  的行。在一個實施例中，該資料矩陣係接著被因式分解成其 SVD，並且接著被簡化， $\mathbf{D}_r = \mathbf{U}_r * \mathbf{S}_r * \mathbf{V}_r'$ ，其中該縮減是透過百分比臨界值法。留下來的奇異向量係接著被用來編碼該  $M \times 1$  目標向量  $\mathbf{t}$ (將被傳送的特徵)，其中最後的編碼是由  $\mathbf{U}_r' * \mathbf{t}$  而得。典型的大小可能是  $M=384$ ， $N=20$ ，且  $r=10$ ，因而一個長度 384 的目標向量係以 10 個係數被壓縮(編碼)。

因為並非所有在該集團資料矩陣  $\mathbf{D}$  中的特徵向量都是立即可利用的，所以該增量的 SVD(ISVD)係被用來根據現有的奇異值分解以及資料更新來更新該 SVD。在一個實施例中，小數目的特徵向量係分組在一起以形成一個最初的資料矩陣  $\mathbf{D}_0$ ，針對該最初的資料矩陣  $\mathbf{D}_0$ ，習知的 SVD 可

輕易地算出。接著，隨著額外的特徵資料向量被加到該集團資料矩陣，該 ISVD 係為了該增加後的資料矩陣而被用來更新該 SVD。在另一個實施例中，因為在該子空間已經在該集團資料矩陣中被表示之下，新的特徵資料向量可能有時候是冗餘的，所以在新的資料向量被加到現有的集團之前，一個線性獨立(independence)測試係被施加至該新的資料向量。一旦整組的特徵資料向量已經都加到該集團之後，該 SVD 係被更新且縮減(經由百分比臨界值法)以提供最後的 SVD 為基礎的編碼。

在另一實施例中，該 SVD 係利用留下來的奇異向量( $U_r$ 的行)與該目標向量  $t$  的關聯性而被降低。總關聯性能量  $CE$  係計算為該些關聯性的總和。當關聯性依序地(以大小遞減的順序，從最大至最小)加總直到在該縮減組中的關聯性的總和超出  $CE$  的某個百分比臨界值為止，該些奇異值的一個分組(在本文中稱為一個“縮減組”)係被產生。此縮減該 SVD 的方法，稱為目標-關聯性百分比臨界值法，係依循與基本的百分比臨界值的 SVD 縮減方法相同的方法，除了(留下來的奇異向量與該目標向量的)目標關聯性係被使用於計算，而不是奇異值用於計算。

### 轉換為基礎的處理

本發明係在轉換空間中的視訊幀資料上執行經驗的特徵分類。在一個實施例中，一組來自一個參考幀的  $N_t$  個特徵係提出作為該分類器的輸入。每個特徵係利用可選擇的

線性轉換(可行的轉換包含離散子波轉換[DWT]以及曲波(curvelet)轉換[CuT])，從像素空間轉換至轉換空間。接著，對於每個特徵之最大的  $P$  個係數的索引係被表列，並且對於每個特徵而言，在所有的係數表列中  $P$  個最常出現的係數係被用來產生一個( $P \times 1$ )分類向量(CV)(總共  $N_t$  個“參考”CV)。接著，每個新的特徵向量  $v$  係被分類，此係藉由轉換該向量、抽出  $v$  的 CV 索引、以及計算在  $v$  的 CV 以及每個參考 CV 之間的相似性量測。該測試的特徵係被分類為該具有參考 CV 會最大化該相似性量測的特徵。

來自兩個或多個具有不同的強項及弱點的線性轉換的資訊可利用正交匹配追蹤來加以結合，以改善該經驗的轉換為基礎的特徵分類器的效能。在一個實施例中，來自可有效的代表紋理的 DWT 的基礎向量以及來自可有效的代表邊緣的 CuT 的基礎向量係結合成為一個函數庫  $D$ 。接著，OMP 係被用來利用在  $D$  中的函數計算  $N_t$  個特徵的每個特徵的一個信號表示、以及該“測試”的特徵向量的一個表示。該分類器接著如同上述的基本的轉換為基礎的分類器繼續進行。以此種方式組合來自多種轉換的資訊可以改善分類器效能而優於每一個別的分類器所能達成的效能。

線性轉換(例如，DWT 及 CuT)亦可被利用於特徵的壓縮及編碼。在一個實施例中，一旦一個特徵被轉換，該些轉換係數係依據大小來排序，並且根據一個能量保留準則(例如，保持足夠的係數以使得 99% 的特徵能量被保持)來分劃。典型的是，保持 99% 的信號能量所需的轉換係數是遠

少於在像素空間中保持 99% 的信號能量所需的像素。該些轉換係數值係代表該特徵的編碼，並且該壓縮增益係得自於所保留的轉換係數相對於該特徵中的像素數目的百分比。在另一個實施例中，來自多種轉換的資訊可再次利用 OMP 來結合，以改善壓縮增益。

儘管本發明已經特別參照其範例實施例來展示與描述，但是所屬技術領域中具有通常知識者將會理解各種在形式及細節上的變化都可以在實施例中做改變，而不脫離由所附的申請專利範圍涵蓋的本發明的範疇。

#### 【圖式簡單說明】

前述內容從上述本發明的範例實施例之更特定的說明來看(如同在所附的圖式中所描繪者)將會是明顯，其中類似的參考符號是指不同的圖式中相同的元件。該圖式並不必然按照比例的，而是強調描繪本發明的實施例。

圖 1 是本發明的一個實施例(混合的編解碼器)的概要圖。

圖 2 是在本發明的編碼器中體現的一個視訊壓縮架構的方塊圖。

圖 3 是其中利用到本發明的實施例之電腦網路環境的概要圖。

圖 4 是在圖 3 的網路中之電腦節點的方塊圖。

圖 5 是描繪代表本發明的一個實施例之特徵的建立模型之圖。

圖 6 是描述根據本發明的一個實施例的預測過程之圖。

圖 7 是本發明的一個實施例(編解碼器)的方塊圖。

圖 8 是描繪根據本發明的一個實施例的特徵追蹤之圖。

【主要元件符號說明】

110	視訊信號輸入
111	編碼特徵之手段
113	偵測特徵之手段
115	分開特徵之手段
117	建立特徵模型之手段
118	習知的編解碼器
119	比較器
120	編碼後的視訊
121	混合的編解碼器的編碼器
122	混合的編解碼器的解碼器
123	判斷之手段
124	解碼特徵之手段
125	分開之手段
126	合成之手段
127	結合之手段
128	合成的視訊信號
200-207	概念上的視訊壓縮架構之每個方塊
307	電腦程式傳播的信號產品
350	客戶電腦/裝置

360	伺服器電腦
370	網路
479	系統匯流排
482	I/O 裝置介面
484	中央處理器單元
486	網路介面
490	記憶體
492	電腦軟體指令
494	資料
495	碟片儲存
510-1、510-2、…510-n	特徵
520-1、520-2、…520-n	幀
530	區域
530-1、530-2、…530-n	區域
540	集團
610	先前解碼過的幀
620	參考幀
630-1、630-2、…630-n	特徵區域
640	方法 1
650	方法 2
660	特徵實例
710-770	圖 7 的實施例(編解碼器)之每個步驟
810	先前被追蹤的特徵
820-1、820-2、…820-n	特徵

201016016

- 830 特徵追蹤器
- 840 目前幀
- 850 一般的特徵偵測器

201016016

98年1月26  
開光

## 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：98133799

※申請日：98.10.6      ※IPC 分類：H04N 7/26

(2006.01)

### 一、發明名稱：(中文/英文)

基於特徵的視訊壓縮

FEATURE-BASED VIDEO COMPRESSION

### 二、中文發明摘要：

處理視訊資料的系統及方法係被提出。具有一系列視訊幀的視訊資料係被接收並且處理。一個候選特徵的一或多個實例係在該些視訊幀中被偵測出。藉由採用在一個運動補償的預測過程中所用的參考幀處理，該候選特徵的實例係與該些視訊幀中的非特徵切割開。該運動補償的預測過程係選出具有對應於該候選特徵的一或多個實例之特徵的先前解碼過的視訊幀。該些先前解碼過的視訊幀係被處理以識別出該候選特徵之可能的符合者。當先前解碼過的視訊幀中有大量的部份包含該候選特徵的實例時，該候選特徵的實例係被聚集成為一組。該候選特徵組係被用來產生一個基於特徵的模型。該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型。該外觀變化模型是藉由對於該候選特徵的實例的變化建立模型而被產生的。該系統判斷出與利用該基於特徵的模型以對於該候選特徵組建立模型相關連之壓縮效率的程度、以及與利用傳統的視訊壓縮以對於該候選特徵組建立

模型相關連之壓縮效率的程度。該基於特徵的模型之壓縮效率係和該傳統的視訊模型之壓縮效率做比較，並且該系統係判斷哪一個是具有較大的壓縮值。根據哪一個是具有較大的壓縮值，該視訊資料係利用該基於特徵的模型或傳統的視訊編碼而被編碼。

### 三、英文發明摘要：

Systems and methods of processing video data are provided. Video data having a series of video frames is received and processed. One or more instances of a candidate feature are detected in the video frames. The instances of the candidate feature are segmented from non-features in the video frames by employing reference frame processing used in a motion compensated prediction process. The motion compensated prediction process selects previously decoded video frames, which have features corresponding to one or more instances of the candidate feature. The previously decoded video frames are processed to identify potential matches of the candidate feature. When a substantial amount of portions of previously decoded video frames include instances of the candidate feature, the instances of the candidate feature are aggregated into a set. The candidate feature set is used to create a feature-based model. The feature-based model

- includes a model of deformation variation and a model of appearance variation of instances of the candidate feature. The appearance variation models are created by modeling variation of instances of the candidate feature. The system determines the extent of compression efficiency associated with using the feature-based model to model the candidate feature set, as well as the extent of compression efficiency associated with using conventional video compression to model the candidate feature set. The feature-based model compression efficiency is compared with the conventional video modeling compression efficiency, and the system determines which one is of greater compression value. The video data is encoded using the feature-based models and conventional video encoding based on which one is of greater compression value.

## 七、申請專利範圍：

1. 一種處理視訊資料的電腦方法，其係包括以電腦實施的步驟：

接收由一系列的視訊幀所構成的視訊資料；以及  
藉由以下來編碼該些視訊幀的部份：

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例；

該偵測係判斷在該一或多個先前解碼過的視訊幀中的實例之位置資訊，該位置資訊係包含一個幀號碼、在該幀中的一個位置、以及該實例的一個空間的周圍；

該候選特徵是一組一或多個偵測出的實例；

藉由一個運動補償的預測過程以利用一或多個先前解碼過的視訊幀來預測在該系列中的一個目前視訊幀的一部份；

該運動補償的預測過程係以位置的預測初始化，其中該些位置的預測係提供來自偵測出的特徵實例在先前解碼過的視訊幀中之位置資訊；

利用該些藉由加入該運動補償的預測過程轉換的實例中的一或多個實例，界定一個特徵以及該些轉換後的實例以產生一個第一基於特徵的模型，該第一基於特徵的模型係使得在該目前幀中預測一個實質相符的特徵實例的出現及一個來源位置成為可能的，其中該實質相符的特徵實例是一個關鍵的特徵實例；

比較該一或多個界定出的特徵的第一基於特徵的模型

與一個習知的視訊編碼模型；

以及從該比較判斷出哪一個模型可致能較大的編碼壓縮；以及

利用該比較及判斷的步驟的結果，施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，以及施加習知的視訊編碼至該一或多個視訊幀的其它部份。

2.如申請專利範圍第1項之方法，其中在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例係進一步包含：

藉由識別具有實質接近的空間上鄰近的像素之一個空間上連續的群組以偵測一個候選特徵的至少一個實例；以及

該些識別出的像素係界定該一或多個視訊幀中的一個視訊幀的一部份。

3.如申請專利範圍第2項之方法，其中在該些視訊幀的一或多個視訊幀中偵測一個候選特徵的一或多個實例進一步包含：

利用該運動補償的預測過程，從複數個候選特徵實例選擇被預測能提供編碼效率的一或多個實例；以及

根據該運動補償的預測過程從獨特的先前解碼過的視訊幀之預測的選擇以判斷該候選特徵之目前的實例與在該目前視訊幀中的其它特徵及非特徵的切割。

4.如申請專利範圍第2項之方法，其中該運動補償的預測過程係進一步利用屬於一或多個特徵的特徵實例來初始

化，此等特徵係在該目前幀中具有實例為符合該視訊部份，其中該視訊部份是在該目前幀中。

5.如申請專利範圍第2項之方法，其中該群組的像素進一步包含以下的一或多個：大區塊或是大區塊的部份。

6.如申請專利範圍第1項之方法，其進一步包含藉由以下來形成一個第二基於特徵的模型：

利用該第一基於特徵的模型作為來自一或多個特徵實例的一或多個運動補償的預測之一個預測的目標，產生該第一基於特徵的模型之一組預測；以及

在組合之後，該組預測變成該第二基於特徵的模型。

7.如申請專利範圍第6項之方法，其中該第二基於特徵的模型係被用來對第一基於特徵的模型的殘餘建立模型，其包含：

對該第二基於特徵的模型相對於該殘餘之結構變化以及外觀變化建立模型；

利用該模型來編碼該殘餘以產生外觀及變形參數；以及

利用該些參數以縮減該殘餘的編碼大小。

8.如申請專利範圍第1項之方法，其中界定一或多個特徵進一步包含藉由以下來根據該候選特徵的實例中的一或多個實例界定一或多個聚集的特徵：

聚集不同的候選特徵的實例成為一個聚集的候選特徵；以及

利用該聚集的候選特徵之該組實例以形成一個實質大

於未聚集的候選特徵之原始的實例的區域，其中該較大的區域係透過在該組中該候選特徵的實例間的連貫性的識別而被形成。

9.如申請專利範圍第 8 項之方法，其中該連貫性係被定義為在該些實例中藉由一個較少參數運動模型實質近似出的外觀一致。

10.如申請專利範圍第 7 項之方法，其中該第二特徵的模型係提供一個與該實例在該解碼過的幀中相對於該空間的位置相關連之選配的矩形像素區域範圍。

11.如申請專利範圍第 10 項之方法，其中該第二特徵的模型係藉由對該特徵之先前正規化的實例建立模型而導出的；以及

其中該些先前正規化的實例是以下所列的任一個：在該目前幀中的該實例、來自一個實質最近的先前解碼過的幀之一個實例、或是來自該些先前解碼過的視訊幀的該些實例的平均。

12.如申請專利範圍第 11 項之方法，其中該外觀模型係藉由該些正規化的第二基於特徵的模型實例的 PCA 分解來表示。

13.如申請專利範圍第 10 項之方法，其進一步包括判斷在每組的特徵實例中相較於其第二基於特徵的模型實例的一致之空間變化的一個變形模型；

為了該變形模型，對於在該組中的每個特徵實例，利用下列的一或多個來近似在該些變形實例中的變化：一個

運動補償的預測過程、網格變形、以及一個具有實質縮減的參數化之運動模型；

將該些變形實例整合到該變形模型中；以及  
其中在該變形模型中的變化係藉由一 PCA 分解來表示。

14.如申請專利範圍第 1 項之方法，其中該運動補償的預測過程係在一個數量上實質大於習知的視訊資料編碼之先前解碼過的視訊幀之一個選集上做運算；以及

其中先前解碼過的視訊幀的該選集並不依賴使用者的監督。

15.如申請專利範圍第 1 項之方法，其中響應於該比較及判斷的步驟來施加習知的視訊編碼係進一步包含當形成該目前幀的部份的一預測時，藉由一個致能該些視訊幀的一或多個視訊幀的部份在記憶體中有更大的壓縮的實例預測過程來加入該習知的視訊編碼；以及

其中該實例預測過程進一步包含：

利用該基於特徵的模型以決定該界定出的特徵出現在一個正被編碼的目標大區塊的一或多個實例，以形成該目前幀之預測的部份；以及

利用該基於特徵的模型以合成像素來預測該目前幀的部份。

16.如申請專利範圍第 15 項之方法，其中響應於該比較及判斷的步驟以施加習知的視訊編碼至該些視訊幀的一或多個視訊幀的部份係進一步包含：

指定一個機率給該些先前解碼過的視訊幀，其中該機率是根據對於該幀利用來自該運動補償的預測過程之位置的預測而判斷出的組合的預測編碼效能的改善而定的；

界定該機率為在用於該目前幀的第一基於特徵的模型以及一個第二基於特徵的模型的分析期間所用之運動補償的預測過程的該組合的編碼效能；

根據該些先前解碼過的視訊幀之機率以分類該些先前解碼過的視訊幀以決定一個從最佳至最差的指標；以及

根據計算及記憶體上的需求來截短該索引表列。

17.如申請專利範圍第 15 項之方法，其進一步包含響應於判斷出下列情形時，再次使用該特徵實例之預測的像素於預測在該目前幀中的其它特徵實例：

在該目前幀中，該界定出的特徵的一或多個實例係重疊超過一個大區塊；或是

當該界定出的特徵的一或多個實例實質相符在該目前幀中的一個大區塊的位置資訊時，該界定出的特徵的一或多個實例係代表一個大區塊。

18.如申請專利範圍第 10 項之方法，其進一步包括預測該些外觀參數以及變形參數以用於一個基於特徵的模型之目前的實例的接合，以及利用該外觀模型及變形模型以及時間上最近的參數以從該基於特徵的模型內插及外推參數來預測在該目前幀中的像素的步驟，其包含：

根據哪一種方法已產生該些時間上最近的特徵實例之最正確的近似來決定用於該些實例的接合的值是線性內插

或是線性外推；

利用較高階的二次方法來偵測該些線性內插及外推的方法之實質減小的效果；

偵測該些二次方法之實質減小的效果並且採用包括擴張型卡爾曼濾波器之更高等的以狀態為基礎的方法以預測該些外觀及變形參數；以及

其中該模型之實際的參數係選配地相對於該些預測出的參數做差分編碼。

19.如申請專利範圍第18項之方法，其中來自該基於特徵的模型之參數係使得在預測該目前幀中的像素所需的計算資源上能夠減少，使得當利用傳統的視訊壓縮以利用該些先前解碼過的視訊幀的一或多個部份來預測該目前幀中的像素時需要更多的計算資源。

20.如申請專利範圍第1項之方法，其中該基於特徵的編碼係被嵌入在習知的視訊編碼內。

21.如申請專利範圍第1項之方法，其中該一或多個界定出的特徵並沒有與在該一或多個視訊幀中之不同的顯著的實體(物件、子物件)的一致。

22.如申請專利範圍第1項之方法，其中該些顯著的實體係透過使用者對於偵測出的特徵之監督的標示為屬於或不屬於一個物件而被決定。

23.如申請專利範圍第1項之方法，其中該些界定出的特徵係包含該些視訊幀的兩個或多個顯著的實體、背景或其它部份的元素。

• 24.如申請專利範圍第 1 項之方法，其中一個界定出的特徵並不對應於一個物件。

25.如申請專利範圍第 11 項之方法，其中施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，以及施加習知的視訊編碼至該一或多個視訊幀的其它部份的步驟：

施加壓縮感測至該第二基於特徵的模型預測的殘餘；

其中該壓縮感測的施加係利用該平均外觀作為一量測並且從該量測來預測該信號；

其中與該壓縮感測的預測相關連的變異係從該第二基於特徵的模型移除；

其中基於特徵的建立模型係專注在該剩餘的殘餘之更小的編碼；以及

施加習知的視訊編碼至該一或多個視訊幀之剩餘的像素以及剩餘的視訊幀。

26.如申請專利範圍第 25 項之方法，其進一步包括使得該視訊資料變為稀疏以增加該施加壓縮感測的步驟的效果之步驟。

27.如申請專利範圍第 1 項之方法，其中該些實例中的一或多個實例係利用一種線性轉換來加以轉換。

28.如申請專利範圍第 1 項之方法，其中該實質相符的特徵是一個利用一種率-失真的尺度判斷出之最佳的相符。

29.如申請專利範圍第 1 項之方法，其進一步包含藉由以下來解碼該編碼的視訊資料：

在一大區塊層級上判斷在該編碼的視訊資料中是否有

一個編碼的特徵；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該些合成的特徵實例以重建原始的視訊資料。

30.如申請專利範圍第1項之方法，其中該基於特徵的編碼係包含對於該一或多個視訊幀的部份施加物件為主的編碼。

31.一種用於處理具有一或多個視訊幀的視訊資料之數位處理系統，其係包括：

一或多個執行一個編碼器的電腦處理器；

該編碼器係利用基於特徵的編碼來編碼該些視訊幀的部份，此係藉由：

偵測一個候選特徵在該些視訊幀的一或多個視訊幀中的一或多個實例；

利用一個運動補償的預測過程，分割在該一或多個視訊幀中的該候選特徵的一或多個實例與非特徵，該運動補

- 償的預測過程係選擇具有對應於該候選特徵的該一或多個實例的特徵之先前解碼過的視訊幀；

利用該候選特徵的實例中的一或多個實例以界定一或多個特徵實例，其中該一或多個界定出的特徵實例係被預測為能夠在該基於特徵的編碼中提供相對於習知的視訊編碼增加相當大的緊密性；

從該一或多個先前解碼過的視訊幀判斷位置資訊，該位置資訊係包含該一或多個界定出的特徵實例在該一或多個先前解碼過的視訊幀中的一個位置以及一個空間的周圍；

利用該一或多個界定出的特徵實例以形成一個基於特徵的模型，該基於特徵的模型係包含來自該些先前解碼過的視訊幀的位置資訊；

利用該基於特徵的模型以正規化該一或多個界定出的特徵實例，該正規化係利用來自該一或多個先前解碼過的視訊幀的位置資訊作為一個位置的預測，所產生的正規化是該一或多個界定出的特徵實例在該目前視訊幀中的預測；

對於該些界定出的特徵中的一或多個特徵來比較該基於特徵的模型與一個習知的視訊編碼模型，並且從該比較來判斷哪一個模型會得到較大的編碼壓縮；以及

利用該比較及判斷的步驟的結果，施加基於特徵的編碼至該些視訊幀的一或多個視訊幀的部份，並且施加習知的視訊編碼至該一或多個視訊幀的其它部份。

32. 一種處理視訊資料的方法，其係包括：

接收具有一系列視訊幀的視訊資料；

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵；

藉由採用在一個運動補償的預測過程中所用的參考幀處理以分割在該些視訊幀中的該候選特徵與非特徵；

處理先前解碼過的視訊幀的一或多個部份以識別出該候選特徵之可能的符合者；

判斷先前解碼過的視訊幀中包含該候選特徵的實例之大量的部份；

將該候選特徵的實例聚集成為該候選特徵的一組實例；

處理該候選特徵組以產生一個基於特徵的模型，其中該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型，該些外觀變化模型是藉由對該候選特徵的實例的像素變化建立模型來加以產生的，該些結構的變化模型是藉由對該候選特徵的實例的像素一致變化建立模型來加以產生的；

判斷與利用該基於特徵的模型以對於該候選特徵組建立模型相關連的壓縮效率；

判斷與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連的壓縮效率；

比較該基於特徵的模型的壓縮效率與該習知的視訊建立模型的壓縮效率，並且判斷哪一個是具有較大的壓縮值；

根據哪一個是具有較大的壓縮值來利用該基於特徵的模型以及習知的視訊編碼以編碼該視訊資料。

33. 一種用於處理具有一或多個視訊幀的視訊資料之數位處理系統，其係包括：

執行一個編碼器的一或多個電腦處理器；

該編碼器係利用基於特徵的編碼以藉由以下來編碼該些視訊幀的部份：

在該些視訊幀的一或多個視訊幀中偵測一個候選特徵；

藉由採用在一個運動補償的預測過程中所用的參考幀處理以分割在該些視訊幀中的該候選特徵與非特徵；

處理先前解碼過的視訊幀的一或多個部份以識別出該候選特徵之可能的符合者；

判斷先前解碼過的視訊幀中包含該候選特徵的實例之大量的部份；

將該候選特徵的實例聚集成為該候選特徵的一組實例；

處理該候選特徵組以產生一個基於特徵的模型，其中該基於特徵的模型係包含該候選特徵的實例的一個變形變化的模型以及一個外觀變化的模型，該些外觀變化模型是藉由對該候選特徵的實例的像素變化建立模型來加以產生的，該些結構的變化模型是藉由對該候選特徵的實例的像素一致變化建立模型來加以產生的；

判斷與利用該基於特徵的模型以對於該候選特徵組建

立模型相關連的壓縮效率；

判斷與利用傳統的視訊壓縮以對於該候選特徵組建立模型相關連的壓縮效率；

比較該基於特徵的模型的壓縮效率與該習知的視訊建立模型的壓縮效率，並且判斷哪一個是具有較大的壓縮值；

根據哪一個是具有較大的壓縮值來利用該基於特徵的模型以及習知的視訊編碼以編碼該視訊資料。

#### 34. 一種處理視訊資料的方法，其係包括：

在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該些合成的特徵實例以重建原始的視訊資料。

#### 35. 一種用於處理視訊資料的資料處理系統，其係包括：

執行一個混合的編解碼器的解碼器之一或多個電腦處

理器，其係能夠藉由以下來利用視訊資料解碼：

藉由在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料以解碼一個編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該視訊資料之該些合成的特徵以重建一原始的視訊資料。

## 八、圖式：

(如次頁)

理器，其係能夠藉由以下來利用視訊資料解碼：

藉由在一大區塊層級上判斷在編碼的視訊資料中是否有一個編碼的特徵以解碼該編碼的視訊資料以解碼一個編碼的視訊資料；

響應於在該編碼的視訊資料中沒有編碼的特徵的判斷，利用習知的視訊解碼來解碼；

響應於在該編碼的視訊資料中有一個編碼的特徵的判斷，分開該編碼的特徵與該編碼的視訊資料，以便於與該視訊資料之傳統編碼的部份分開地合成該編碼的特徵實例；

決定基於特徵的模型以及與該編碼的特徵相關連之特徵參數；

利用該些決定出的基於特徵的模型以及特徵參數以合成該編碼的特徵實例；以及

組合該視訊資料之傳統編碼的部份以及該視訊資料之該些合成的特徵以重建一原始的視訊資料。

## 八、圖式：

(如次頁)

201016016

四、指定代表圖：

(一)本案指定代表圖為：第（1）圖。

(二)本代表圖之元件符號簡單說明：

110	視訊信號輸入
111	編碼特徵之手段
113	偵測特徵之手段
115	分開特徵之手段
117	建立特徵模型之手段
118	習知的編解碼器
119	比較器
120	編碼後的視訊
121	混合的編解碼器的編碼器
122	混合的編解碼器的解碼器
123	判斷之手段
124	解碼特徵之手段
125	分開之手段
126	合成之手段
127	結合之手段
128	合成的視訊信號

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無