

公告本

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：95135898

※ 申請日期：95.9.27

※IPC 分類：H04N 7/16  
H04B 1/66  
H04N 7/12  
H04N 7/173  
H04N 11/02  
H04N 11/04

一、發明名稱：(中文/英文)

用於與時域邊界作資料對齊之方法及裝置

METHODS AND DEVICE FOR DATA ALIGNMENT WITH TIME  
DOMAIN BOUNDARY

二、申請人：(共 1 人)

姓名或名稱：(中文/英文)

美商高通公司

QUALCOMM INCORPORATED

代表人：(中文/英文)

喬治 A 懷坦

WHITTEN, GEORGE A.

住居所或營業所地址：(中文/英文)

美國加州聖地牙哥市摩豪斯大道 5775 號

5775 MOREHOUSE DRIVE SAN DIEGO, CA 92121-1714 U. S. A.

國 籍：(中文/英文)

美國 U.S.A.

## 三、發明人：(共 8 人)

姓名：(中文/英文)

1. 福拉米 R 雷福德蘭  
RAVEENDRAN, VIJAYALAKSHMI R.
2. 格登 肯特 沃克  
WALKER, GORDON KENT
3. 田濤  
TIAN, TAO
4. 法尼庫馬 巴哈米迪帕提  
BHAMIDIPATI, PHANIKUMAR
5. 徐方  
SHI, FANG
6. 陳培松  
CHEN, PEISONG
7. 席特曼 加帕西 薩巴曼尼亞  
SUBRAMANIA, SITARAMAN GANAPATHY
8. 塞福勒 亥特 歐茲  
OGUZ, SEYFULLAH HALIT

國籍：(中文/英文)

- |            |        |
|------------|--------|
| 1. 印度      | INDIA  |
| 2. 美國      | U.S.A. |
| 3. 中華人民共和國 | P.R.C. |
| 4. 印度      | INDIA  |
| 5. 加拿大     | CANADA |
| 6. 中華人民共和國 | P.R.C. |
| 7. 印度      | INDIA  |
| 8. 土耳其     | TURKEY |

#### 四、聲明事項：

主張專利法第二十二條第二項  第一款或  第二款規定之事實，其事實發生日期為： 年 月 日。

申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

有主張專利法第二十七條第一項國際優先權：

1. 美國；2005 年 09 月 27 日；60/721,416
2. 美國；2005 年 10 月 17 日；60/727,643
3. 美國；2005 年 10 月 17 日；60/727,644
4. 美國；2005 年 10 月 17 日；60/727,640
5. 美國；2005 年 10 月 24 日；60/730,145
6. 美國；2006 年 03 月 10 日；11/373,579 ~~本案優先權之主張應不予受理~~
7. 美國；2006 年 04 月 03 日；60/789,048
8. 美國；2006 年 04 月 04 日；60/789,377

無主張專利法第二十七條第一項國際優先權：

- 1.
- 2.

主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

主張專利法第三十條生物材料：

須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

## 九、發明說明：

### 【發明所屬之技術領域】

本申請案係針對用以視訊轉碼用於即時串流之視訊資料的設備及方法，且更特定而言，本申請案係針對在行動廣播應用中轉碼用於即時串流之視訊資料。

### 【先前技術】

由於有限之頻寬資源及可用頻寬之可變性，使得有效視訊壓縮可用於許多多媒體應用中，諸如無線視訊串流及視訊電話。諸如MPEG-4 (ISO/IEC)、H.264 (ITU)之一些視訊編碼標準或類似視訊編碼提供了非常適用於諸如無線廣播之應用的高效率編碼。某些多媒體資料(例如數位電視顯示)一般根據諸如MPEG-2之其他標準來編碼。因此，轉換編碼器用於在無線廣播之前將根據一個標準(例如，MPEG-2)進行編碼之多媒體資料轉碼或轉換成另一標準(例如，H.264)。

改良速率最佳化之編解碼器可提供誤差回彈、誤差恢復及可調性方面之優點。此外，使用自多媒體資料自身所判定之資訊亦可提供編碼之額外改良，包括誤差回彈、誤差恢復及可調性。因此，需要一提供多媒體資料之高效處理及壓縮的轉換編碼器，該轉換編碼器使用自多媒體資料自身所判定之資訊、具有可調性且具有誤差回彈性以用於包括串流多媒體資訊之行動廣播的許多多媒體資料應用中。

### 【發明內容】

基於所描述及說明之轉碼設備及方法的本發明內容之每一者具有若干態樣，其中該等態樣中之單一者均無法單獨

負責達成其所要之屬性。在不限制此揭示內容之範疇的情況下，現將簡短地論述其更多突出特徵。在研究此論述之後，且尤其是在閱讀名稱為"實施方式"之部分後，將能理解受此內容驅動的轉碼之特徵係如何提供多媒體資料處理設備及方法之改良。

本文所描述之發明態樣係關於使用內容資訊以用於編碼多媒體資料之各種方法以及用於編碼器(例如，用於轉換編碼器中之編碼器)之各種模組或組件中。轉換編碼器可對使用內容資訊來轉碼多媒體資料進行協調。可自另一源接收內容資訊，例如，藉由視訊所接收之中繼資料。轉換編碼器可經組態以經由各種不同處理操作而產生內容資訊。在某些態樣中，轉換編碼器產生多媒體資料之內容分類，該內容分類繼而用於一或多個編碼過程中。在某些態樣中，受內容驅動之轉換編碼器可判定多媒體資料之空間及時間內容資訊，並使用內容資訊以用於跨通道之內容感知均衡品質編碼及基於內容分類之壓縮/位元分配。

在某些態樣中，獲得或計算多媒體資料之內容資訊(例如，中繼資料、內容量度及/或內容分類)，且繼而將內容資訊提供至轉換編碼器之組件以用於處理用以編碼之多媒體資料。舉例而言，預處理器可使用一些內容資訊以用於場景改變偵測、執行反向電視電影處理("IVTC")、解交錯、移動補償及雜訊抑制(例如，2維子波變換)以及時空雜訊降低(例如，假影移除、解振鈴、解塊及/或去雜訊)。在某些態樣中，預處理器亦可使用內容資訊以用於空間解析率向

下取樣，例如，當自標準定義(SD)向下取樣至四分之一視訊圖形陣列(QVGA)時，判定適當之"安全"及"動作處理"區。

在某些態樣中，編碼器包括一經組態以計算內容資訊之內容分類模組。編碼器可使用內容分類以：用於在判定每一MB之量化參數(QP)中之位元速率控制(例如，位元分配)；用於移動估計，例如，執行色彩移動估計(ME)、執行移動向量(MV)預測、提供一基礎層及一增強層之可調性；及用於誤差回彈，此係藉由使用內容分類來影響預測階層架構及誤差回彈方案(包括(例如)自適應訊框內再新、邊界對齊過程)且在增強層中提供冗餘I訊框資料。在某些態樣中，轉換編碼器與資料多工器配合而使用內容分類以用於跨通道保持最佳化多媒體資料品質。在某些態樣中，編碼器可使用內容分類資訊以用於強迫I訊框週期性地出現於已編碼之資料中從而允許快速通道切換。此等實施亦可利用在用於誤差回彈之已編碼之資料中可能需要的I區塊，使得可經由預測階層架構來有效地組合隨機存取切換及誤差回彈(基於(例如)內容分類)，以改良編碼效率同時增加對誤差之強韌性。

在一個態樣中，一種處理多媒體資料之方法包括：獲得多媒體資料之內容資訊；及編碼多媒體資料，以便將資料邊界與時域中之訊框邊界對齊，其中該編碼係基於內容資訊。該內容可包含內容分類。獲得內容資訊可包括計算來自多媒體資料之內容資訊。在某些情形中，資料邊界包含I訊框資料邊界。資料邊界亦可為多媒體資料之獨立可解碼

的已編碼之資料之邊界。在某些情形中，資料邊界包含片邊界。資料邊界亦可為訊框內編碼之存取單元邊界。資料邊界亦可為P訊框邊界或B訊框邊界。內容資訊可包括多媒體資料之複雜性，且該複雜性可包含時間複雜性、空間複雜性或時間複雜性與空間複雜性。

在另一態樣中，一種用於處理多媒體資料之設備包括：一內容分類器，其經組態以判定多媒體資料之內容分類；及一編碼器，其經組態以編碼多媒體資料，以便將資料邊界與一時域中之訊框邊界對齊，其中該編碼係基於內容資訊。

在另一態樣中，一種用於處理多媒體資料之設備包括：用於獲得多媒體資料之內容資訊的構件；及用於編碼多媒體資料以便將資料邊界與一時域中之訊框邊界對齊的構件，其中該編碼係基於內容資訊。

在另一態樣中，處理器經組態以：獲得多媒體資料之內容資訊；及編碼多媒體資料，以便將資料邊界與一時域中之訊框邊界對齊，其中該編碼係基於內容資訊。

另一態樣包括一包含指令之機器可讀取媒體，該等指令一經執行便導致機器：獲得多媒體資料之內容資訊；及編碼多媒體資料，以便將資料邊界與一時域中之訊框邊界對齊，其中該編碼係基於內容資訊。

另一態樣包含一種處理多媒體資料之方法，其包含：獲得多媒體資料之內容分類；及基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以

增加已編碼之多媒體資料之誤差回彈。

在又一態樣中，一種用於處理多媒體資料之設備包含：一內容分類器，其經組態以獲得多媒體資料之內容分類；及一編碼器，其經組態以基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加已編碼之多媒體資料之誤差回彈。

另一態樣包含一處理器，其經組態以：獲得多媒體資料之內容分類；及基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加已編碼之多媒體資料之誤差回彈。

在另一態樣中，一用於處理多媒體資料之設備包含：用於獲得多媒體資料之內容分類的構件；及用於基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊以增加已編碼之多媒體資料之誤差回彈的構件。

另一態樣包含一包含指令之機器可讀取媒體，該等指令一經執行便導致機器：獲得多媒體資料之內容分類；及基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加已編碼之多媒體資料之誤差回彈。

### 【實施方式】

下列實施方式係針對在此揭示內容中所論述之一些態樣。然而，可以許多種不同方法來體現本發明。在此說明書中對"一個態樣"或"一態樣"之參考意謂在至少一個態樣



中包括了結合該態樣所描述之一特定特徵、結構或特性。在本說明書中之不同位置中短語"在一個態樣中"、"根據一個態樣"或"在某些態樣中"的出現既非必定全部指代相同態樣，亦非為互斥其他態樣之獨立或替代態樣。此外，描述了可藉由某些態樣且非藉由其他態樣顯示之各種特徵。類似地，描述了對於某些態樣而言可為要求但對於其他態樣而言非為要求之各種要求。

下列描述包括用以提供對實例之全面理解的細節。然而，一般熟習此項技術者應理解，即使本文未描述或說明一實例或態樣中之過程或裝置的每一細節，仍可實踐該等實例。舉例而言，可在方塊圖中展示電力組件，該等方塊圖並未說明組件之每一電連接或每一電元件，以避免在不必要之細節中使該等實例晦澀難懂。在其他例子中，可詳細地展示此等組件、其他結構及技術以進一步解釋該等實例。

本揭示內容係關於使用被編碼之多媒體資料之內容資訊來控制編碼及轉碼設備及方法。(多媒體資料之)"內容資訊"或"內容"為廣義術語，其意謂關於多媒體資料之內容的資訊，且可包括(例如)中繼資料、自多媒體資料計算而得之量度及與一或多個量度相關聯之關於內容的資訊(例如，內容分類)。內容資訊可提供至一編碼器或藉由一編碼器來判定，此視特定應用而定。內容資訊可用於多媒體資料編碼之許多態樣，包括場景改變偵測、時間處理、時空雜訊降低、向下取樣、判定量化之位元速率、可調性、誤差回彈、

跨廣播通道保持最佳化多媒體品質及快速通道切換。使用此等態樣中之一或多個態樣，轉換編碼器可協調處理多媒體資料並產生與內容相關之已編碼之多媒體資料。本文中之描述轉碼態樣的描述及圖亦可應用於編碼態樣及解碼態樣。

轉化編碼器設備及方法係關於自一格式轉碼至另一格式，且其在本文中被特定描述為係關於將MPEG-2視訊轉碼為增強之可調性H.264格式以用於在無線通道上傳輸至行動裝置(其說明了某些態樣)。然而，將MPEG-2視訊轉碼為H.264格式之描述並不意欲為限制本發明之範疇，而僅為本發明之某些態樣的例示。所揭示之設備及方法提供一高效架構，其支持具有隨機存取及分層能力之誤差回彈編碼，且亦可應用於轉碼及/或編碼除MPEG-2及H.264之外的視訊格式。

如本文中所使用之"多媒體資料"或僅"多媒體"為一廣義術語，其包括視訊資料(其可包括音訊資料)、音訊資料或視訊資料與音訊資料兩者。如本文中所使用之"視訊資料"或"視訊"為一廣義術語，其指代基於訊框或基於場之資料，該資料包括一或多個影像或相關之影像序列(包含本文、影像資訊及/或音訊資料)，且可用以指代多媒體資料(例如，可互換地使用該等術語)(除非另有規定)。

以下描述了一轉換編碼器之各種組件的實例及可使用內容資訊以用於編碼多媒體資料之過程的實例。

圖1A為說明一多媒體資料廣播系統100之某些態樣之資

料流的方塊圖。在系統100中，多媒體資料提供者106將已編碼之多媒體資料104傳達至一轉換編碼器200。已編碼之多媒體資料104藉由轉換編碼器200接收，該轉換編碼器200在區塊110中將多媒體資料104處理為原始多媒體資料。區塊110中之處理對已編碼之多媒體資料104進行解碼及剖析，且進一步處理多媒體資料以將其預備用於編碼成另一格式。已解碼之多媒體資料被提供至區塊112，在區塊112中多媒體資料被編碼成一預定多媒體格式或標準。多媒體資料一旦已編碼，則在區塊114處其預備用於經由(例如)一無線廣播系統(例如，蜂巢式電話廣播網路或經由另一通信網路)之傳輸。在某些態樣中，已根據MPEG-2標準來編碼已接收之多媒體資料104。在已轉碼之多媒體資料104已被解碼後，轉換編碼器200將多媒體資料編碼成一H.264標準。

圖1B為一轉換編碼器130之方塊圖，該轉換編碼器130可經組態以執行圖1A之區塊110及112中的處理。轉換編碼器130可經組態以接收多媒體資料、將多媒體資料解碼及剖析成分封化基本流(例如，副標題、音訊、中繼資料、"原始"視訊、CC資料及顯示時間戳)、將其編碼成一所要之格式及提供已編碼之資料以用於進一步之處理或傳輸。轉換編碼器130可經組態從而以兩個或兩個以上之資料組來提供已編碼之資料(例如，已編碼之第一資料組及已編碼之第二資料組)，此稱為分層編碼。在某些態樣實例中，一分層編碼方案中之各種資料組(或層)可以不同之品質水平加以編碼，且進行格式化，使得在第一資料組中所編碼之資料具

有比在第二資料組中所編碼之資料低的品質(例如，在顯示時提供一較低視覺品質水平)。

圖 1C 為一處理器 140 之方塊圖，該處理器 140 可經組態以轉碼多媒體資料，且可經組態以執行圖 1A 之區塊 110 及 112 中所描繪之處理的一部分或全部。處理器 140 可包括執行本文所描述之一或多個轉碼過程(包括解碼、剖析、預處理及編碼)的模組 124a...n，且其使用內容資訊以用於處理。處理器 140 亦包括內部記憶體 122，且該處理器 140 可經組態以直接或經由另一裝置間接地與外部記憶體 120 通信。處理器 140 亦包括一通信模組 126，其經組態以與處理器 140 外部之一或多個裝置通信，包括接收多媒體資料及提供已編碼之資料(諸如在第一資料組中所編碼之資料及在第二資料組中所編碼之資料)。在某些態樣實例中，在一分層編碼方案中之各種資料組(或層)可以不同之品質水平加以編碼，且進行格式化，使得在第一資料組中所編碼之資料具有比在第二資料組中所編碼之資料低的品質(例如，在顯示時提供一較低視覺品質水平)。

可藉由硬體、軟體、韌體、中間體、微碼或其之任一組合來實施其之轉換編碼器 130 或預處理器 140(經組態以用於轉碼)組件及其中所含有之過程。舉例而言，剖析器、解碼器、預處理器或編碼器可為獨立組件、作為硬體、韌體、中間體而被併入另一裝置之組件中或實施於在一處理器上執行之微碼或軟體中，或其之組合。當實施於軟體、韌體、中間體或微碼中時，執行移動補償、鏡頭分類及編碼過程

之程式碼或碼段可儲存於諸如儲存媒體之機器可讀取媒體中。碼段可表示程序、函數、子程式、程式、常用程式、子常用程式、模組、軟體套件、類別或指令、資料結構或程式語句之任一組合。可藉由傳遞及/或接收資訊、資料、自變量、參數或記憶體內容將一碼段耦接至另一碼段或一硬體電路。

### 轉換編碼器架構之說明性實例

圖2說明了可用於在圖1之多媒體廣播系統100中所說明之轉換編碼器200的轉換編碼器之實例的方塊圖。轉換編碼器200包含一剖析器/解碼器202、一預處理器226、一編碼器228及一同步層240(下文將予以進一步描述)。如本文中所描述，轉換編碼器200經組態以使用多媒體資料104之內容資訊從而用於轉碼過程之一或多個態樣。內容資訊可經由多媒體中繼資料自轉換編碼器200外部之源而獲得，或藉由轉換編碼器(例如，藉由預處理器226或編碼器228)計算而得。圖2中所示之組件說明了可被包括於一使用內容資訊以用於一或多個轉碼過程之轉換編碼器中的組件。在一特定實施中，可排除轉換編碼器200之一或多個組件，或可包括額外組件。此外，描述了轉換編碼器之若干部分及轉碼過程，以便允許熟習此項技術者即使在本文可能未描述一過程或一裝置之每一細節的情況下仍可實踐本發明。

圖5說明了作為轉換編碼器200之各種組件及/或過程之操作的時間關係之圖解說明的時序圖。如圖5中所示，藉由剖析器205(圖2)而在任意時間零點(0)首先接收諸如

MPEG-2視訊的已編碼之串流視訊104(已編碼之多媒體資料)。接著，諸如藉由剖析器205協同解碼器214來對視訊流進行剖析501、解多工502及解碼503。如所說明，此等過程可並行發生而具有稍微之計時偏差，以便將處理器資料之流輸出提供至預處理器226(圖2)。在時間 $T_i$  504，一旦預處理器226已自解碼器214接收足夠之資料以開始輸出處理結果，則剩餘之處理步驟本質上變為順次的，其中在預處理之後順次發生第一遍編碼505、第二遍編碼506及重編碼507，直至在時間 $T_f$  508完成重編碼為止。

本文中所描述之轉換編碼器200可經組態以轉碼各種多媒體資料，且許多該等過程應用於被轉碼之任一類型的多媒體資料。儘管本文中所提供之若干實例係特定關於將MPEG-2資料轉碼成H.264資料，但並不意謂此等實例將本揭示內容限制於此資料。以下所描述之編碼態樣可應用於將任何適合之多媒體資料標準轉碼成另一適合之多媒體資料標準。

### 剖析器/解碼器

再次參看圖2，剖析器/解碼器202接收多媒體資料104。剖析器/解碼器202包括一傳送流剖析器("剖析器")205，其接收多媒體資料104，且將該資料剖析成一視訊基本流(ES)206、一音訊ES 208、顯示時間戳(PTS)210及諸如副標題212之其他資料。ES載運來自單個視訊或音訊編碼器之一種類型的資料(視訊或音訊)。舉例而言，視訊ES包含用於一序列資料之視訊資料，其包括序列標頭及該序列之所有

子部分。分封化基本流或PES係由已形成為封包之單個ES組成，每一封包通常以一添加之封包標頭開始。一PES流僅含有來自一個源(例如，來自一視訊或音訊編碼器)之一種類型的資料。PES封包具有不對應於傳送封包之固定封包長度的可變長度，且可比一傳送封包長得多。當由一PES流形成傳送封包時，可在緊接著傳送封包標頭之傳送封包有效負載的開始置放該PES標頭。剩餘之PES封包內容填充連續傳送封包之有效負載，直至PES封包被全部使用為止。可將最後之傳送封包填充至一固定長度，例如，藉由以位元組(例如，位元組=0xFF(所有者))進行填滿。

剖析器205將視訊ES 206傳達至一解碼器214，該解碼器214為此處所示之剖析器/解碼器202的部分。在另一組態中，剖析器205及解碼器214為獨立組件。將PTS 210發送至轉換編碼器PTS產生器215，該轉換編碼器PTS產生器215可產生為轉換編碼器200所特有之獨立顯示時間戳以用於配置待自轉換編碼器200發送至一廣播系統的資料。轉換編碼器PTS產生器215可經組態以將資料提供至轉換編碼器200之同步層240，從而協調資料廣播之同步。

圖3說明了過程300之一實例的流程圖，剖析器205在剖析以上所描述之各種分封化基本流時可遵循該過程300。當自一內容提供者106(圖1)接收多媒體資料104時，過程300始於區塊302。過程300進行至區塊304，在區塊304中執行剖析器205之初始化。可藉由一經獨立產生之擷取命令306來觸發初始化。舉例而言，獨立於剖析器205且基於一外部接收

之TV時間表及通道同相對齊資訊之過程可產生擷取命令306。此外，可輸入即時傳送流(TS)緩衝描述符308以幫助初始化及用於主處理兩者。

如區塊304中所說明，初始化可包括：獲取一命令語法核對；執行第一遍PSI/PSIP/SI(節目特定資訊/節目及系統資訊協定/系統資訊)處理；執行特定關於擷取命令或PSI/PSIP/SI一致性核對之處理；分配用於每一PES之PES緩衝器；及設定計時(例如，用於與所要之擷取開始瞬間作對齊)。PES緩衝器保持已剖析之ES資料，且將每一已剖析之ES資料傳達至一對應之音訊解碼器216、測試編碼器220、解碼器214或轉換編碼器PTS產生器215。

初始化之後，過程300進行至用於已接收之多媒體資料104之主處理的區塊310。區塊310中之處理可包括目標封包識別符(PID)過濾、連續PSI/PSIP/SI監控及處理以及一計時過程(例如，用於達成一所要之擷取持續時間)，使得將傳入之多媒體資料傳遞至適當之PES緩衝器。由於在區塊310中處理多媒體資料，使得產生了一程式描述符及PES緩衝器'讀取'之指示，其將如下文中所述與解碼器214(圖2)建立介面。

在區塊310之後，過程300進行至區塊314，其中發生剖析操作之終止，包括產生一計時器中斷及PES緩衝器之釋放(因其消耗而引起)。應注意，對於在程式之描述符中所引用的該程式之所有相關基本流(諸如，音訊、視訊及副標題流)而言，PES緩衝器將存在。

再次參看圖2，剖析器205將音訊ES 208發送至一音訊解



碼器 216 以對應於轉換編碼器實施，且將已編碼之本文 216 提供至同步層 240 並對音訊資訊進行解碼。將副標題資訊 212 遞送至一本文編碼器 220。可將來自解碼器 214 之隱藏式字幕 (CC) 資料 218 提供至本文編碼器 220，該本文編碼器 220 以一受轉換編碼器 200 影響之格式來編碼副標題資訊 212 及 CC 資料 218。

剖析器/解碼器 202 亦包括解碼器 214，該解碼器 214 接收視訊 ES 206。解碼器 214 可產生與視訊資料相關聯之中繼資料、將已編碼之視訊分封化基本流解碼成原始視訊 224 (例如，以標準定義格式) 及處理視訊 ES 流中之視訊隱藏式字幕資料。

圖 4 展示了說明可藉由解碼器 214 執行之解碼過程 400 之一實例的流程圖。過程 400 以在區塊 402 處輸入一視訊基本流資料 206 而開始。過程 400 進行至區塊 404，在區塊 404 中初始化該解碼器。初始化可包括許多任務，包括偵測視訊序列標頭 (VSH)、執行第一遍 VSH、視訊序列 (VS) 及 VS 顯示擴展處理 (包括視訊格式、色原及矩陣係數) 以及分配資料緩衝器以分別緩衝已解碼之圖片、相關聯之中繼資料及隱藏式字幕 (CC) 資料。此外，輸入由剖析器 205 所提供之視訊 PES 緩衝器 '讀取' 資訊 406 (例如，其可由在圖 3 之區塊 310 中的過程 300 產生)。

在區塊 404 處之初始化之後，過程 400 進行至區塊 408，在區塊 408 中由解碼器 214 執行視訊 ES 之主處理。主處理包括針對新資料可用性而輪詢視訊 PES 緩衝器 '讀取' 資訊或 "介

面"；在圖片邊界處解碼視訊ES、重建並儲存像素資料；在圖片邊界處同步視訊及a/v、產生中繼資料並進行儲存；及在圖片邊界處儲存CC資料。主處理408之結果區塊410包括產生序列描述符、產生已解碼之圖片緩衝描述符、產生中繼資料緩衝描述符及產生CC資料緩衝描述符。

在主處理408之後，過程400進行至區塊412，在區塊412中其執行一終止過程。該終止過程可包括判定終止條件，該等終止條件包括在一預定臨限值上的一特定持續時間中無新資料出現、一序列結束碼之偵測及/或一顯式終止信號之偵測。該終止過程可進一步包括釋放已解碼之圖片、相關聯之中繼資料及CC資料緩衝器(因由下文將描述之預處理器而致使其消耗而引起)。過程400結束區塊414，在區塊414中其可輸入等待待被接收為輸入之視訊ES的狀態。

### 預處理器

圖2及更詳細之圖6說明了一可使用內容資訊以用於一或多個預處理操作之預處理器226的樣本態樣。預處理器226自剖析器/解碼器202接收中繼資料222及已解碼之"原始"視訊資料224。預處理器226經組態以對視訊資料224及中繼資料222執行一些類型之處理，且將已處理之多媒體(例如，基礎層參考訊框、增強層參考訊框、頻寬資訊、內容資訊)及視訊提供至編碼器228。多媒體資料之此預處理可改良資料之視覺清晰度、抗混淆及壓縮效率。一般而言，預處理器226接收由剖析器解碼器202中之解碼器214所提供之視訊序列，且將該等視訊序列轉換成漸進視訊序列以由編碼

器228予以進一步處理(例如，編碼)。在某些態樣中，預處理器226可經組態而用於眾多操作，包括反向電視電影處理、解交錯、過濾(例如，假影移除、解振鈴、解塊及/或去雜訊)、重定大小(例如，自標準定義至四分之一視訊圖形陣列(QVGA)之空間解析率向下取樣)及GOP結構產生(例如，計算複雜性映射產生、場景改變偵測及衰落/閃光偵測)。

預處理器226可使用來自解碼器之中繼資料以影響一或多個預處理操作。中繼資料可包括關於、描述或分類多媒體資料之內容的資訊("內容資訊")，詳言之，中繼資料可包括一內容分類。在某些態樣中，中繼資料並不包括為編碼操作所要之內容資訊。在此等情形中，預處理器226可經組態以判定內容資訊並使用內容資訊以用於預處理操作及/或將內容資訊提供至轉換編碼器200之其他組件(例如，解碼器228)。在某些態樣中，預處理器226可使用此內容資訊來影響GOP分割、判定適當類型之過濾及/或判定可傳達至一編碼器之編碼參數。

圖6展示了可被包括於預處理器226中且說明可由預處理器226執行之處理的各種過程區塊之說明性實例。在此實例中，預處理器226接收中繼資料及視訊222、224，且將包含(經處理之)中繼資料及視訊的輸出資料614提供至編碼器228。通常，存在三種類型的可接收之視訊。第一，已接收之視訊可為漸進視訊，其中無需解交錯。第二，視訊資料可為自24fps電影序列轉換之電視電影處理之視訊、交錯視訊(在此情形中之視訊)。第三，視訊可為非電視電影處理之

交錯視訊。預處理器226可如以下所描述來處理此等類型之視訊。

在區塊601處，預處理器226判定已接收之視訊資料222、224是否為漸進視訊。在某些情形中，此可自中繼資料加以判定(若中繼資料包含此資訊)，或藉由處理視訊資料自身而加以判定。舉例而言，以下所描述之反向電視電影處理過程可判定已接收之視訊222是否為漸進視訊。若其為漸進視訊，則過程進行至區塊607，在區塊607中對視訊執行過濾(例如，雜訊抑制器)操作以降低雜訊(諸如，白高斯雜訊)。若視訊資料222、224並非為漸進視訊，則在區塊601處過程進行至區塊604，即進行至一相位偵測器604。

相位偵測器604在源於電視電影處理中之視訊與以標準廣播格式開始之視訊之間進行區分。若已作出決策該視訊為經電視電影處理的(退出相位偵測器604之YES決策路徑)，則在反向電視電影處理606中將經電視電影處理之視訊恢復至其原始格式。識別並消除冗餘訊框，且將源自相同視訊訊框之場重編成一完整影像。因為以一秒之1/24的規則時間間隔攝影記錄了已重建之膠捲影像的序列，所以使用反向電視電影處理之影像而非電視電影處理之資料(其具有不規則時基)，在GOP分割器612或解碼器228中所執行之移動估計過程更為精確。

在一個態樣中，相位偵測器604在接收一視訊訊框後做出一些決策。此等決策包括：(i)來自電視電影處理輸出及3:2下拉相位之當前視訊是否為圖38中所示之五個相位 $P_0$ 、

$P_1$ 、 $P_2$ 、 $P_3$ 及 $P_4$ 中之一者；及(ii)視訊是否經產生為習知NTSC。彼決策被指示為相位 $P_5$ 。此等決策呈現為圖2中所示之相位偵測器604的輸出。自相位偵測器604之標有"YES"(是)的路徑致動反向電視電影處理606，指示其已具備正確下拉相位，使得其可選出由相同攝影影像所形成之場且將該等場組合。自相位偵測器604之標有"NO"(否)的路徑同樣地致動解交錯器605以將一表觀NTSC訊框分成用於最佳化處理之場。相位偵測器604可連續分析視訊訊框，此係因為可在任一時間接收不同類型之視訊。作為一例證，可將符合NTSC標準之視訊插入該視訊而作為廣告。在反向電視電影處理之後，將所得之漸進視訊發送至一可用以降低白高斯雜訊之雜訊抑制器(過濾器)607。

當習知NTSC視訊被確認時(自相位偵測器601之NO路徑)，將其傳輸至解交錯器605以進行壓縮。解交錯器605將已交錯之場變換成漸進視訊，且繼而可對該漸進視訊執行去雜訊操作。以下描述瞭解交錯處理之一說明性實例。

如電視之傳統類比視訊裝置以一交錯之方式再現視訊，亦即，此等裝置傳輸偶數掃描線(偶數場)及奇數掃描線(奇數場)。自信號取樣觀點而言，此等於以由以下方程式所描述之圖案進行的時空子取樣：

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{對於偶數場而言, 若 } y \bmod 2 = 0 \\ \Theta(x, y, n), & \text{對於奇數場而言, 若 } y \bmod 2 = 1 \\ \text{抹除, 否則,} & \end{cases} \quad [1]$$

其中 $\Theta$ 表示原始訊框圖片、 $F$ 表示已交錯之場，且 $(x, y, n)$ 分

別表示一像素之水平位置、垂直位置及時間位置。

在不失一般性之情況下，貫穿此揭示內容可假定  $n=0$  為偶數場，使得以上之方程式 1 簡化為：

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{若 } y \bmod 2 = n \bmod 2, \\ \text{抹除, 否則,} & \end{cases} \quad [2]$$

因為在水平維度中不進行抽取，所以可在下一  $n \sim y$  座標中描繪子取樣圖案。

解交錯器之目標為將已交錯之視訊(一序列場)變換成非交錯之漸進訊框(一序列訊框)。換言之，內插偶數場及奇數場以"恢復"或產生全訊框圖片。此可由方程式 3 表示：

$$F_o(x, y, n) = \begin{cases} F(x, y, n), & y \bmod 2 = n \bmod 2, \\ F_i(x, y, n), & \text{否則,} \end{cases} \quad [3]$$

其中  $F_i$  表示遺漏像素之解交錯結果。

圖 40 為說明解交錯器 605 之一實例之一些態樣的方塊圖，該解交錯器 605 使用 Wmed 過濾及移動估計以自己交錯之多媒體資料產生一漸進訊框。圖 40 之上部分展示了一移動強度映射 4002，其可使用來自一"當前場"、兩個"先前場"(PP 場及 P 場)及兩個"後續場"("下個場"及"下下個場")之資訊而產生。移動強度映射 4002 將當前訊框分類或分割成兩個或兩個以上之不同的移動位準，且該移動強度映射 4002 可藉由時空過濾而產生(下文將對其予以進一步詳細描述)。在某些態樣中，如以下參考方程式 4-8 所描述，產生移動強度映射 4002 以識別靜態區、慢動作區及快動作區。一時空過濾器(例如，Wmed 過濾器 4004)使用基於移動強度映

射之標準來過濾已交錯之多媒體資料，且產生一時空臨時解交錯訊框。在某些態樣中，Wmed過濾過程涉及 $[-1, 1]$ 之水平鄰域、 $[-3, 3]$ 之垂直鄰域及五個相鄰場之時間鄰域，該等相鄰場係由圖40中所說明之五個場("PP場"、"P場"、"當前場"、"下個場"、"下下個場")而表示，其中 $Z^{-1}$ 表示一個場之延遲。相對於"當前場"，"下個場"及"P場"為非同位場，且"PP場"及"下下個場"為同位場。用於時空過濾之"鄰域"指代在過濾操作期間實際使用之場及像素的空間及時間位置，且其說明為如(例如)圖6及圖7中所示之"孔徑"。

解交錯器605亦可包括一雜訊抑制器(去雜訊過濾器)4006，該雜訊抑制器4006經組態以過濾由Wmed過濾器4004所產生的時空臨時解交錯訊框。去雜訊該時空臨時解交錯訊框使得後續移動搜尋過程更為精確，尤其在源交錯多媒體資料序列被白雜訊污染時。其亦可至少部分地移除Wmed圖片中之偶數列與奇數列之間的混淆。可將雜訊抑制器4006實施為各種過濾器，包括基於子波收縮及子波文納濾波器之雜訊抑制器。雜訊抑制器可用以在使用移動補償資訊對雜訊進行進一步處理之前將該雜訊自候選Wmed訊框移除，且可移除存在於Wmed訊框中之雜訊且保持信號存在而不考慮信號之頻率內容。可使用各種類型之包括子波過濾器的去雜訊過濾器。子波為用以在空間域及尺度域兩者中定位一給定信號的一類函數。子波背後之基本觀念為以不同尺度或解析率來分析信號，使得子波表示中之較小改變在原始信號中產生一對應較小之改變。

亦可將子波收縮或子波文納濾波器應用為雜訊抑制器。子波收縮由下列組成：有雜信號之子波變換，接著為較小子波係數至零(或更小值)之收縮，同時保持較大係數不受影響。最後，執行一反向變換以獲取所估計之信號。

去雜訊過濾提高了有雜環境中之移動補償的精確度。子波收縮去雜訊可涉及在子波變換域中之收縮，且通常包含三個步驟：一線性正向子波變換、一非線性收縮去雜訊及一線性反向子波變換。文納濾波器為一MSE最佳化線性過濾器，其可用以改良因相加雜訊及模糊而被降級的影像。此等過濾器一般在此項技術中已為吾人所熟知，且其描述於(例如)以上引用之"Ideal spatial adaptation by wavelet shrinkage"及S. P. Ghael、A. M. Sayeed及R. G. Baraniuk之"Improvement Wavelet denoising via empirical Wiener filtering"(*SPIE*記錄，第3169卷，第389-399頁，San Diego，1997年7月)中，其以引用之方式全部明確地併入本文中。

在某些態樣中，去雜訊過濾器係基於一(4,2)雙正交三次B雲規函數子波過濾器之一態樣。可藉由下列正向及反向變換來界定一個此過濾器：

$$h(z) = \frac{3}{4} + \frac{1}{2}(z + z^{-1}) + \frac{1}{8}(z + z^{-2}) \quad (\text{正向變換}) \quad [4]$$

及

$$g(z) = \frac{5}{4}z^{-1} - \frac{5}{32}(1 + z^{-2}) - \frac{3}{8}(z + z^{-3}) - \frac{3}{32}(z^2 + z^{-4}) \quad (\text{反向變換}) \quad [5]$$

去雜訊過濾器之應用可增加在一有雜環境中之移動補償的精確度。此等過濾器之實施進一步描述於D.L. Donoho及



I.M. Johnstone 之 "Ideal spatial adaptation by wavelet shrinkage"(生物統計學，第8卷，第425-455頁，1994)，其以引用之方式全部明確地併入本文中。

圖40之下部分說明了用於判定已交錯之多媒體資料之移動資訊(例如，移動向量候選、移動估計、移動補償)的一態樣。詳言之，圖40說明了一移動估計及移動補償方案，其用以產生所選訊框之移動補償臨時漸進訊框，且繼而其與Wmed臨時訊框進行組合以形成一所得之"最後"漸進訊框，其展示為解交錯之當前訊框4014。在某些態樣中，已交錯之多媒體資料之移動向量("MV")候選(或估計)自外部移動估計器而被提供至解交錯器，且用以提供雙向移動估計器及補償器("ME/MC")4018之一起點。在某些態樣中，MV候選選擇器4022使用先前判定之MV以用於經處理之區塊之MV候選的鄰近區塊，諸如先前處理之區塊(例如，在解交錯之先前訊框4020中之區塊)的MV。基於先前之解交錯訊框70及下個(例如，未來)Wmed訊框4008，可雙向地進行移動補償。藉由組合器4012來合併或組合一當前Wmed訊框4010及一移動補償("MC")當前訊框4016。將所得之現為一漸進訊框的解交錯當前訊框4014提供返回至ME/MC 4018以待用作一解交錯之先前訊框4020，且其亦在解交錯器605外部進行通信以用於後續處理。

可去耦解交錯預測方案，包含使用Wmed+MC解交錯方案自場內內插之場間內插。換言之，可為場內內插之目的而主要使用時空Wmed過濾，同時可在移動補償期間執行場間

內插。此降低了Wmed結果之峰值信雜比比，但在應用移動補償後視覺品質更令人滿意，此係因為因不精確場間預測模式決策而產生之壞像素將自Wmed過濾過程而被移除。

在適當反向電視電影處理或解交錯處理之後，在區塊608處處理漸進視訊以用於混淆抑制及重取樣(例如，重定大小)。在某些重取樣態樣中，實施多相重取樣器以用於圖片尺寸之重定大小。在向下取樣之一實例中，在原始圖片與經重定大小之圖片之間的率比可為 $p/q$ ，其中 $p$ 及 $q$ 為互質數。相位之總數為 $p$ 。多相位過濾器之截止頻率在某些態樣中為0.6以用於將係數重定大小為0.5左右。截止頻率並不精確地匹配該重定大小比率以便提高經重定大小之序列的高頻率回應。此不可避免地容許某些混淆。然而，已熟知人類眼睛較模糊而無混淆之圖片而言更青睞清晰但有稍微混淆之圖片。

圖41說明了多相位重取樣之一實例，其展示若重定大小比率為 $3/4$ 時的相位。圖41中所說明之截止頻率亦為 $3/4$ 。藉由垂直軸而在上圖中說明了原始像素。辛克函數亦以該等軸為中心而繪製以表示過濾器波形。因為將截止頻率選擇為完全與重取樣比率相同，所以在重定大小後辛克函數之零重疊該等像素之位置(在圖41中以交叉說明)。為了在重定大小後找到一像素值，可如下列方程式中所示自該等原始像素合計基值：

$$v(x) = \sum_{i=-\infty}^{\infty} u(i) \times \text{sinc}(\pi f_c(i-x)) \quad [6]$$

其中  $f_c$  為截止頻率。可將以上 1 維多相位過濾器應用於水平維度及垂直維度兩者。

重取樣(重定大小)之另一態樣說明了過度掃描。在 NTSC 電視信號中，一影像具有 486 條掃描線，且在數位視訊中一影像在每一掃描線上可具有 720 個像素。然而，由於尺寸與螢幕格式之間的失配，所以並非完整影像之全部在電視上均為可見的。將不可見之影像部分稱為過度掃描。

為了幫助廣播台將有用之資訊置於盡可能多之電視可見的區域中，移動圖片與電視工程協會(SMPTE)界定了稱為安全動作區及安全標題區之動作訊框的特定尺寸。見關於 *Specifications for Safe Action and Safe Title Areas Test Pattern for Television Systems* 之 SMPTE 推薦實踐 RP 27.3-1989。由 SMPTE 將安全動作區界定為其中"所有重要動作必須發生"之區域。將安全標題區界定為其中"可限制所有有用資訊以確保在大多數家庭電視接收器上之可見性"的區域。

舉例而言，參看圖 25，安全動作區 2510 佔用了螢幕之中心 90%，周圍給出 5% 的邊界。安全標題區 2505 佔用了螢幕之中心 80%，給出了 10% 的邊界。現參看圖 26，因為安全標題區如此小，所以為了在影像中添加更多內容，某些台將把本文置於安全動作區，其在白色矩形窗口 2615 內部。

通常，在過度掃描中可見黑色邊界。舉例而言，在圖 26 中，黑色邊界出現於影像之上面 2620 及下面 2625 處。因為 H.264 視訊在移動估計中使用邊界擴展，所以可在過度掃描

中移除此等黑色邊界。擴展之黑色邊界可增加剩餘。適當地，可切割2%之邊界，且繼而對其實施重定大小。因此，可產生用於重定大小之過濾器。在多相位向下取樣前執行截斷以移除過度掃描。

再次參看圖6，漸進視訊繼而進行至區塊610，在區塊610中執行解塊器及解振鈴操作。在視訊壓縮應用中通常出現兩種類型之假影："成塊"及"振鈴"。因為壓縮演算法將每一訊框劃分成區塊(例如，8x8個區塊)，所以出現成塊假影。每一區塊經重建而具有某些較小誤差，且在一區塊之邊緣處的誤差與在鄰近區塊之邊緣處的誤差常形成對比，從而使區塊邊界可見。相反，振鈴假影在影像特徵之邊緣周圍呈現為失真。因為編碼器在量化高頻率DCT係數時廢除過多資訊，所以出現振鈴假影。在某些說明性實例中，解塊及解振鈴兩者均可使用低通FIR(有限脈衝回應)過濾器來隱藏此等可見假影。

在解塊處理之一實例中，可將一解塊過濾器應用於一訊框之所有4x4區塊邊緣，除了在訊框之邊界處的邊緣及禁止解塊過濾器過程之任何邊緣外。應在完成訊框建構過程後基於巨集區塊來執行此過濾過程，其中在一訊框中之所有巨集區塊以增加巨集區塊位址之次序加以處理。對於每一巨集區塊而言，首先自左至右過濾垂直邊緣，且繼而自頂部至底部過濾水平邊緣。如圖39中所示，對於水平方向及對於垂直方向而言，對四個16樣本邊緣執行亮度解塊過濾過程，且對兩個8樣本邊緣執行針對每一色度分量之解塊過

濾器過程。在當前巨集區塊上方及左側之樣本值(可能已因對先前巨集區塊之解塊過程操作而被修改)應用作對當前巨集區塊上之解塊過濾器過程的輸入，且可在當前巨集區塊之過濾期間進一步修改。在垂直邊緣之過濾期間所修改之樣本值可用作用於相同巨集區塊之水平邊緣之過濾的輸入。可獨立地針對亮度及色度分量來調用解塊過程。

在解振鈴處理之一實例中，可自適應地應用2維過濾器以使邊緣附近之區域平滑。邊緣像素經受很少量之過濾或不經受過濾以避免模糊。

## **GOP分割器**

在解塊及解振鈴之後，藉由GOP分割器612來處理漸進視訊。GOP定位可包括偵測鏡頭改變、產生複雜性映射(例如，時間、空間頻寬映射)及自適應GOP分割。下文對此等之每一者進行了描述。

### **A. 場景改變偵測**

鏡頭偵測係關於判定在一組圖片(GOP)中之一訊框何時顯示指示已發生場景改變之資料。一般而言，在一GOP內，訊框可在任何兩個或三個(或三個以上)之相鄰訊框中無顯著改變，或可存在緩慢改變或快速改變。當然，若需要，則此等場景改變分類可視一特定應用而定而進一步被細分為更大程度之改變。

偵測鏡頭或場景改變對於視訊之有效編碼而言至關重要。通常，當一GOP並未顯著改變時，在GOP之開始的I訊框之後，眾多預測訊框可充分地編碼視訊，使得該視訊之

後續解碼及顯示為視覺上可接受的。然而，當一場景突然或緩慢改變時，可能需要額外I訊框及較少之預測編碼(P訊框及B訊框)來產生後續解碼之視覺上可接受的結果。

以下描述了改良現有編碼系統之效能的鏡頭偵測與編碼系統及方法。此等態樣可實施於預處理器226(圖7)之GOP分割器612中，或包括於一可在有預處理器或無預處理器之情況下進行操作的編碼器裝置中。此等態樣利用包括視訊資料之相鄰訊框之間的統計比較的統計(或量度)來判定是否發生了一突然場景改變、一場景是否正在緩慢改變或在該場景中是否存在可使視訊編碼尤為複雜之攝影機閃光。該等統計可自一預處理器獲得，且繼而被發送至一編碼裝置，或可在一編碼裝置中產生該等統計(例如，藉由一經組態以執行移動補償的處理器)。所得之統計幫助場景改變偵測決策。在實施轉碼之系統中，常存在一適當之預處理器或可組態之處理器。若該預處理器執行幫助移動補償之解交錯，則移動補償統計為可用的且為現成的。在此等系統中，鏡頭偵測演算法可稍微增加系統複雜性。

本文所描述之鏡頭偵測器之說明性實例僅需要利用來自先前訊框、當前訊框及下個訊框之統計，且因此具有極低之潛伏時間。鏡頭偵測器區分若干不同類型之鏡頭事件，包括突然場景改變、交叉衰落及其他緩慢場景改變以及攝影機閃光。藉由在編碼器中利用不同策略來判定不同類型之鏡頭事件，編碼效率及視覺品質得以增強。

場景改變偵測可用於任一用於其之視訊編碼系統以藉由

以一固定時間間隔插入一I訊框來巧妙地節省位元。在某些態樣中，藉由預處理器所獲得之內容資訊(例如，被併入中繼資料中或由預處理器226計算而得)可用於場景改變偵測。舉例而言，視內容資訊而定，可針對不同類型之視訊內容動態地調整以下所描述之臨限值及其他標準。

通常對結構化圖片組(GOP)操作視訊編碼。GOP通常始於一訊框內編碼之訊框(I訊框)，接著為一系列P(預測)或B(雙向)訊框。通常，I訊框可儲存顯示該訊框所需之所有資料，B訊框依賴於先前及隨後之訊框中的資料(例如，僅含有改變自先前訊框之資料或與下個訊框中之資料不同)，且P訊框含有改變自先前訊框之資料。在共同使用中，在編碼視訊中I訊框與P訊框及B訊框交替。就尺寸(例如，用以編碼該訊框之位元的數目)而言，I訊框通常比P訊框大得多，P訊框又大於B訊框。為了進行有效編碼、傳輸及解碼處理，GOP之長度應足夠長以降低自較大I訊框之有效損耗，且應足夠短以對抗編碼器與解碼器之間的失配或通道損壞。此外，因相同原因，可訊框內編碼P訊框中之巨集區塊(MB)。

場景改變偵測可用於視訊編碼器以判定適當之GOP長度，且基於該GOP長度插入I訊框，而非以固定時間間隔插入一通常不需要之I訊框。在一實用之串流視訊系統中，通信通道通常因位元誤差或封包損耗而被損壞。置放I訊框或I MB之地點可顯著影響已解碼之視訊品質及觀測體驗。一編碼方案為使用訊框內編碼之訊框以用於具有自並置之先前圖片或圖片部分之顯著改變的圖片或圖片之部分。通

常，此等區域無法藉由移動估計來有效地進行預測，且若根據訊框間編碼技術(例如，使用B訊框及P訊框之編碼)來免除此等區域，則可更為有效地實施編碼。在通道損壞之情形中，彼等區域可能遭受誤差傳播，該誤差傳播可藉由訊框內編碼來降低或消除(或幾乎如此)。

可將GOP視訊之部分分類成兩個或兩個以上之類別，其中每一區域可具有不同之可視特定實施而定的訊框內編碼標準。作為一實例，可將視訊分類成三個類別：突然場景改變、交叉衰落及其他緩慢場景改變以及攝影機閃光。

突然場景改變包括顯著不同於先前訊框之訊框，其通常由攝影機操作而導致。因為此等訊框之內容與先前訊框之內容不同，所以應將該等突然場景改變訊框編碼為I訊框。

交叉衰落及其他緩慢場景改變包括場景之緩慢切換，其通常由攝影機鏡頭之電腦處理而導致。兩種不同場景之逐漸混合在人類眼睛看來可更為令人滿意，但對視訊編碼提出了挑戰。移動補償無法有效地降低彼等訊框之位元速率，且可為此等訊框而更新更多之訊框內MB。

當訊框之內容包括攝影機閃光時，發生攝影機閃光或攝影機閃光事件。此等閃光在持續時間上(例如，一個訊框)相對較短，且極為明亮，使得描繪該等閃光之訊框中的像素相對於一相鄰訊框上之對應區域顯示異乎尋常地高的亮度。攝影機閃光突然且立即移位一圖片之亮度。通常，攝影機閃光之持續時間比人類視覺系統(HVS)之時間遮罩持續時間短，其通常被界定為44 ms。人類眼睛對亮度之此等



短叢發的品質並不敏感，且因此可粗略地對其進行編碼。因為閃光訊框無法藉由移動補償進行有效地處理，且該等閃光訊框對於未來訊框而言為差之預測候選，所以此等訊框之粗略編碼並不會降低未來訊框之編碼效率。由於"人工"高亮度，所以不應將被分類為閃光之場景用於預測其他訊框，且由於相同原因，其他訊框無法有效地用於預測此等訊框。一旦經識別，則可取出此等訊框，此係因為該等訊框需要相對較高之處理量。一種選擇為移除該等攝影機閃光訊框且在其位置中編碼一DC係數，此解決方案較簡單，計算快速且節省許多位元。

當偵測到訊框之以上類別中之任一者時，則宣告一鏡頭事件。鏡頭偵測不僅適用於改良編碼品質，其亦可幫助識別視訊內容搜尋及檢索。下文中描述了場景偵測過程之一說明性態樣。在此實例中，對於為鏡頭偵測而被處理之一選定訊框而言，鏡頭偵測過程首先計算資訊或量度。量度可包括來自視訊之雙向移動估計及補償處理之資訊以及其他基於亮度之量度。

為了執行雙向移動估計/補償，可藉由一雙向移動補償器來預處理一視訊序列，該雙向移動補償器將當前訊框之每 $8 \times 8$ 區塊與該等訊框(最相鄰之鄰近訊框)之兩者(過去之訊框及未來之訊框)中的區塊匹配。移動補償器產生每一區塊之移動向量及差分量度。圖29為一說明，其展示了當前訊框C與過去訊框P及未來(或下個)訊框N之匹配像素的一實例，且描繪了至該等匹配像素之移動向量(過去移動向量

$MV_p$ 及未來移動向量 $MV_N$ )。下文中參看圖32而通常描述了雙向移動向量產生及相關編碼的一般描述。

在判定雙向移動資訊(例如, 識別對應之相鄰訊框中之MB(最匹配)的移動資訊)後, 可藉由當前訊框與下個訊框及先前訊框之各種比較來產生額外量度(例如, 藉由GOP分割器612中之移動比較器或另一適當組件)。移動比較器可產生每一區塊之差分量度。該差分量度可為均方差分之總和(SSD), 或絕對差分之總和(SAD)。在不失一般性之情況下, 在此將SAD用作一實例。

對於每一訊框而言, 將亦稱為"對比率"之SAD比率計算如下:

$$\gamma = \frac{\varepsilon + SAD_p}{\varepsilon + SAD_N} \quad [6]$$

其中 $SAD_p$ 及 $SAD_N$ 分別為正向差分量度及反向差分量度之絕對差分的總和。應注意, 分母含有一較小正數 $\varepsilon$ 以防止"用零除"誤差。分子亦含有 $\varepsilon$ 以平衡分母中1之效應。舉例而言, 若先前訊框、當前訊框及下個訊框相同, 則移動搜尋應得出 $SAD_p = SAD_N = 0$ 。在此情形中, 以上計算產生結果 $\gamma = 1$ 而非零或無窮大。

可為每一訊框而計算亮度直方圖。通常, 多媒體影像具有八個位元之亮度深度(例如, "箱"之數目)可將用於根據某些態樣計算亮度直方圖之亮度深度設定至16以獲得直方圖。在其他態樣中, 可將亮度深度設定至一適當數目, 其可視進行處理之資料的類型、可用之計算功率或其他預定

標準而定。在某些態樣中，可基於一計算或接收之量度(諸如資料之內容)來動態地設定亮度深度。

以下之方程式說明了計算一亮度直方圖差分( $\lambda$ )的一實例：

$$\lambda = \frac{\sum_{i=1}^{16} |N_{Pi} - N_{Ci}|}{N} \quad [7]$$

其中  $N_{Pi}$  為在先前訊框之第  $i$  個箱中之區塊的數目，且  $N_{Ci}$  為在當前訊框之第  $i$  個箱中之區塊的數目，且  $N$  為在一訊框中之區塊的總數目。若先前及當前訊框之亮度直方圖差分完全不同(或不相交)，則  $\lambda = 2$ 。

使用此資訊，將一訊框差分量度( $D$ )計算如下：

$$D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \quad [8]$$

其中  $A$  為由應用所選擇之常數， $\gamma_C = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N}$ ，且  $\gamma_P = \frac{\varepsilon + SAD_{PP}}{\varepsilon + SAD_C}$ 。

若訊框差分量度滿足方程式 9 中所示之標準，則將所選(當前)訊框分類為一突然場景改變訊框：

$$D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \geq T_1 \quad [9]$$

其中  $A$  為由應用所選擇之常數，且  $T_1$  為一臨限值。

在一實例中，模擬展示了設定  $A = 1$  及  $T_1 = 5$  可達成優良之偵測效能。若當前訊框為一突然場景改變訊框，則  $\gamma_C$  應較大且  $\gamma_P$  應較小。可使用比率  $\frac{\gamma_C}{\gamma_P}$  而非單獨之  $\gamma_C$ ，使得將量度標準化至該情形之活動水平。

應注意，以上標準以非線性之方式使用亮度直方圖差分( $\lambda$ )。圖16說明 $\lambda * (2\lambda + 1)$ 為一凸函數。當 $\lambda$ 較小時(例如，接近於零)，其幾乎不為預強調。 $\lambda$ 變得愈大，則該函數進行之強調愈多。藉由此預強調，對於任何大於1.4之 $\lambda$ 而言，若將臨限值 $T_1$ 設定於5，則偵測到一突然場景變化。

若場景強量度 $D$ 滿足方程式5中所示之標準，則將當前訊框判定為一交叉衰落或緩慢場景變化：

$$T_2 \leq D < T_1 \quad [10]$$

對於一定數目之連續訊框而言，其中 $T_1$ 為以上所使用之相同臨限值，且 $T_2$ 為另一臨限值。

閃光事件通常導致亮度直方圖移位至較亮側。在此說明性態樣攝影機中，將亮度直方圖統計用於判定當前訊框是否包含攝影機閃光。如方程式11及12中所示，鏡頭偵測過程可判定當前訊框之負亮度是否比先前訊框之亮度大某一臨限值 $T_3$ ，及當前訊框之亮度是否比下個訊框之亮度大該臨限值 $T_3$ ：

$$\bar{Y}_C - \bar{Y}_P \geq T_3 \quad [11]$$

$$\bar{Y}_C - \bar{Y}_N \geq T_3 \quad [12]$$

若不滿足以上標準，則不將當前訊框分類為包含攝影機閃光。如以下方程式中所說明，若滿足該標準，則鏡頭偵測過程判定反向差分量度 $SAD_P$ 及正向差分量度 $SAD_N$ 是否大於某一臨限值 $T_4$ ：

$$SAD_P \geq T_4 \quad [13]$$

$$SAD_N \geq T_4 \quad [14]$$

其中  $\bar{Y}_C$  為當前訊框之平均亮度， $\bar{Y}_P$  為先前訊框之平均亮度， $\bar{Y}_N$  為下個訊框之平均亮度，且  $SAD_P$  及  $SAD_N$  為與當前訊框相關聯之正向差分量度及反向差分量度。

鏡頭偵測過程藉由首先判定當前訊框之亮度是否大於先前訊框之亮度及下個訊框之亮度來判定攝影機閃光事件。若非如此，則該訊框非為攝影機閃光事件，但若如此，則其可為攝影機閃光事件。鏡頭偵測過程繼而可估計反向差分量度是否大於一臨限值  $T_3$  及正向差分量度是否大於一臨限值  $T_4$ ，若此等兩個條件均滿足，則鏡頭偵測過程將當前訊框分類為具有攝影機閃光。若不滿足該標準，則不將該訊框分類為任何類型之鏡頭事件，或可對其給定一預設分類，其識別待對該訊框實施之編碼(例如，遺漏型訊框，編碼為 I 訊框)。

以上展示了  $T_1$ 、 $T_2$ 、 $T_3$  及  $T_4$  之某些例示性值。通常，經由測試鏡頭偵測之特定實施來選擇此等臨限值。在某些態樣中，預定該等臨限值  $T_1$ 、 $T_2$ 、 $T_3$  及  $T_4$  中之一或多者，且將此等值併入編碼裝置中之鏡頭分類器中。在某些態樣中，可在處理期間(例如，動態地)基於使用被供應至鏡頭分類器之資訊(例如，中繼資料)或基於藉由鏡頭分類器自身所計算之資訊來設定該等臨限值  $T_1$ 、 $T_2$ 、 $T_3$  及  $T_4$  中之一或多者。

使用鏡頭偵測資訊來編碼視訊通常在編碼器中執行，但為了鏡頭偵測揭示之完整性而在此加以描述。參看圖 30，編碼過程 301 可使用鏡頭偵測資訊，以基於該序列之訊框中

的偵測鏡頭來編碼視訊。過程301進行至區塊303，且核對以查看是否將當前訊框分類為突然場景改變。若如此，則在區塊305處，可將當前訊框編碼為I訊框，且可判定GOP邊界。若非如此，則過程301進行至區塊307，若當前訊框被分類為緩慢改變場景之一部分，則在區塊309處可將緩慢改變場景中之當前訊框及其他訊框編碼為一預測訊框(例如，P訊框或B訊框)。過程301繼而進行至區塊311，在區塊311中其核對該當前訊框是否被分類為一包含攝影機閃光之閃光場景。若如此，則在區塊313處，可識別該訊框以用於特殊處理(例如，移除或編碼該訊框之DC係數)，若非如此，則不實施當前訊框之分類，且當前訊框可根據其他標準來編碼(編碼為I訊框)，或遺漏。

在以上描述之態樣中，藉由訊框差分量度D來指示待壓縮之訊框與其相鄰之兩個訊框之間的差分量。若偵測到顯著的單向亮度改變，則其表示訊框中之交叉衰落效應。交叉衰落愈突出，則可藉由使用B訊框達成之增益愈多。在某些態樣中，如以下方程式中所示來使用一已修改之訊框差分量度：

$$D_1 = \begin{cases} \left(1 - \alpha + 2\alpha \frac{|d_P - d_N|}{d_P + d_N}\right) \times D, & \text{若 } Y_P - \Delta \geq Y_C \geq Y_N + \Delta \text{ 或 } Y_P + \Delta \leq Y_C \leq Y_N - \Delta, \\ D, & \text{否則,} \end{cases}$$

[15]

其中  $d_P = |Y_C - Y_P|$  及  $d_N = |Y_C - Y_N|$  分別為當前訊框與先前訊框之間的亮度差分及當前訊框與下個訊框之間的亮度差

分， $\Delta$ 表示可在正常實驗中判定之常數(因為其可視實施而定)，且 $\alpha$ 為具有在0與1之間的值的加權變量。

## B. 頻寬映射產生

預處理器 226(圖 6)亦可經組態以產生一可用於編碼多媒體資料之頻寬映射。在某些態樣中，編碼器 228(圖 7)中之內容分類模組 712 代替產生頻寬映射。

人類視覺品質  $V$  可為編碼複雜性  $C$  及分配位元  $B$  (亦稱為頻寬) 兩者之函數。圖 15 為說明此關係之圖表。應注意，編碼複雜性量度  $C$  自人類視覺之觀點來研究空間及時間頻率。因為對於人類眼睛失真更為敏感，所以複雜性值相應地更高。通常可假定  $V$  在  $C$  中單調減少且在  $B$  中單調增加。

為了達成恆定視覺品質，將頻寬 ( $B_i$ ) 指派給待編碼之第  $i$  個對象 (訊框或 MB)，其滿足在緊接於下文之兩個方程式中所表示的標準：

$$B_i = B(C_i, V) \quad [16]$$

$$B = \sum_i B_i \quad [17]$$

在緊接於上文之兩個方程式中， $C_i$  為第  $i$  個對象之編碼複雜性， $B$  為總的可用頻寬，且  $V$  為一對象之達成的視覺品質。難以將人類視覺品質公式化為一方程式。因此，以上方程式組並未精確地加以界定。然而，若假定 3 維模型在所有變量中為連續的，則可認為頻寬比 ( $B_i/B$ ) 在  $(C, V)$  對之鄰域內未改變。在以下所示之方程式中界定頻寬比  $\beta_i$ ：

$$\beta_i = B_i/B \quad [18]$$

繼而可如下列方程式中所表示來界定位元分配：

$$\begin{aligned} \beta_i &= \beta(C_i) \\ 1 &= \sum_i \beta_i \quad (\text{對於 } (C_i, V) \in \delta(C_0, V_0) \text{ 而言}) \end{aligned} \quad [19]$$

其中  $\delta$  指示 "鄰域"。

編碼複雜性受人類視覺敏感性影響(時間及空間兩者)。吉洛德(Girot)人類視覺模型為可用於界定空間複雜性之模型的一實例。此模型研究局部空間頻率及環境照明。所得之量度稱作  $D_{csat}$ 。在過程中之預處理點，未知將對一圖片進行訊框內編碼還是訊框間編碼，且產生兩者之頻寬比。根據不同視訊對象之  $\beta_{INTRA}$  之間的比率來分配位元。對於訊框內編碼之圖片而言，在下列方程式中表示頻寬比：

$$\beta_{INTRA} = \beta_{0INTRA} \log_{10} (1 + \alpha_{INTRA} Y^2 D_{csat}) \quad [20]$$

在以上方程式中， $Y$  為巨集區塊之平均亮度分量， $\alpha_{INTRA}$  為亮度均方及其後之  $D_{csat}$  項的加權係數， $\beta_{0INTRA}$  為用以保證  $1 = \sum_i \beta_i$  的標準化係數。舉例而言， $\alpha_{INTRA} = 4$  之值達成優良之視覺品質。內容資訊(例如，內容分類)可用於將  $\alpha_{INTRA}$  設定至一值，該值對應於視訊之特定內容的所要之優良視覺品質水平。在一實例中，若視訊內容包含 "會說話的頭" 新聞廣播，則可將視覺品質水平設定較低，此係因為可認為資訊影像或視訊之可顯示部分不如音訊部分重要，且可分配較少位元來編碼資料。在另一實例中，若視訊內容包含一運動事件，則內容資訊可用於將  $\alpha_{INTRA}$  設定至一值，該值對應於一較高視覺品質水平，此係因為顯示之影像對於觀測者



而言可能更為重要，且因此可分配更多位元來編碼資料。

為了理解此關係，應注意，頻寬對數地分配有編碼複雜性。亮度均方項  $Y^2$  反映了具有較大量值之係數使用更多位元來編碼之事實。為了防止對數獲得負值，在括號中將 1 添加至該項。亦可使用具有其他底數之對數。

時間複雜性係藉由訊框差分量度之量測(其量測兩個連續訊框之間的差分)加以判定，此將移動量(例如，移動向量)連同諸如絕對差分之總和(SAD)的訊框差分量度考慮在內。

訊框間編碼之圖片的位元分配可研究空間複雜性及時間複雜性兩者。將此表示於下文：

$$\beta_{INTER} = \beta_{0INTER} \log_{10} \left( 1 + \alpha_{INTER} \cdot SSD \cdot D_{csat} \exp \left( -\gamma \|MV_P + MV_N\|^2 \right) \right) \quad [21]$$

在以上方程式中， $MV_P$  及  $MV_N$  為當前 MB 之正向及反向移動向量(見圖 29)。可注意，藉由均方差分(SSD)之總和來替代訊框內編碼頻寬公式中的  $Y^2$ 。為了理解以上方程式中之  $\|MV_P + MV_N\|^2$  的作用，注意人類視覺系統之下一特性：經受平滑、可預測移動(較小  $\|MV_P + MV_N\|^2$ )之區域引起了注意，且可由眼睛追蹤，且通常無法忍受比固定區域再多的失真。然而，經受快速或不可預測移動(較大  $\|MV_P + MV_N\|^2$ )之區域無法進行追蹤，且可忍受顯著之量化。實驗展示  $\alpha_{INTER} = 1$ 、 $\gamma = 0.001$  達成優良的視覺品質。

### C. 自適應 GOP 分割

在可由預處理器 226 執行之處理的另一說明性實例中，圖 6 之 GOP 分割器 612 亦可自適應地改變被編碼在一起之一組

圖片的組成，且將參考使用MPEG2之實例對該GOP分割器612加以論述。某些較舊之視訊壓縮標準(例如，MPEG2)無需GOP具有一規則結構(儘管可強加為如此)。MPEG2序列始終始於I訊框，亦即，在未參看先前圖片之情況下已被編碼之一者。通常藉由固定P之GOP或在I訊框後之預測圖片中的間距，可在編碼器處預配置MPEG2 GOP格式。P訊框為已自先前I或P圖片進行部分預測之圖片。將開始之I訊框與後續之P訊框之間的訊框編碼為B訊框。"B"訊框(B代表雙向)可將先前及其次之I或P圖片個別或同時用作參考。編碼一I訊框所需要之位元的數目平均超過編碼一P訊框所需要之位元的數目，同樣，編碼一P訊框所需要之位元的數目平均超過編碼一B訊框所需要的數目。一跳躍訊框(若使用其)將無需位元用於其表示。

在P訊框及B訊框之使用中及在較近之壓縮演算法(降低表示視訊所需要之資料之速率之訊框的跳躍)中所隱含的概念為時間冗餘的消除。當時間冗餘較高(亦即，幾乎不存在自圖片至圖片之改變)時，P、B或跳躍圖片之使用有效地表示視訊流，此係因為較早解碼之I或P圖片隨後被用作解碼其他P或B圖片的參考。

自適應GOP分割係基於自適應地使用此概念。量化訊框之間的差分，且在對已量化之差分執行適當測試後，自動做出藉由I、P、B或跳躍訊框表示圖片之決策。自適應結構具有在固定GOP結構中無法獲得之優點。固定結構將忽略已發生極小內容改變之可能性，自適應程序將允許多得多

之B訊框插入於每一I與P之間，或插入於兩個P訊框之間，藉此減少充分表示該訊框序列所需要之位元的數目。相反地，當視訊內容中之改變顯著時，P訊框之效率大大降低，此係因為預測訊框與參考訊框之間的差分過大。在此等條件下，匹配對象可掉出移動搜尋區域，或匹配對象之間的相似性由於攝影機角度之改變所導致之失真而被降低。此時，應將P訊框或I訊框及其相鄰P訊框選擇為彼此更接近，且應插入較少B訊框。固定GOP不可實施彼調整。

在此處所揭示之系統中，自動感測此等條件。該GOP結構具有靈活性且使其適應內容之此等改變。該系統估計一訊框差分量度，其可認為係訊框間之距離量測(具有距離之相同附加性質)。概念上，給定訊框 $F_1$ 、 $F_2$ 及 $F_3$ 具有訊框間距離 $d_{12}$ 及 $d_{23}$ ，認為 $F_1$ 與 $F_3$ 間之距離為至少 $d_{12} + d_{23}$ 。基於此距離類量度來實施訊框指定。

GOP分割器藉由當接收圖片類型時將該等圖片類型指派至該等訊框來進行操作。圖片類型指示在編碼每一區塊中所需要之預測方法：

在不參看其他圖片之情況下編碼I圖片。因為該等I圖片獨立，所以其在資料流中提供存取點，其中解碼可開始。可將I編碼類型指派至一訊框(若至其前任訊框之"距離"超過一場景改變臨限值)。

P圖片可使用先前I或P圖片以用於移動補償預測。該等圖片將先前場或訊框(其可自所預測之區塊移位)中之區塊用作編碼之基礎。在自所研究之區塊減去參考區塊後，通常

使用用於空間冗餘之消除的離散餘弦變換來編碼剩餘之區塊。將P編碼類型指派至一訊框(若其與被指定為一P訊框之上一訊框間的"距離"超過一通常小於第一臨限值之第二臨限值)。

如以上所描述，B訊框圖片可使用先前及其次之P或I圖片以用於移動補償。可正向、反向或雙向預測B圖片中之區塊，或可在不參考其他訊框之情況下訊框內編碼該區塊。在H.264中，參考區塊可為來自32個訊框之多達32個區塊的線性組合。若該訊框無法被指定為一I或P類型，則將其指定為一B類型(若自該訊框至其直接前任的"距離"大於一通常小於第二臨限值的第三臨限值)。

若該訊框無法被指定成為一編碼之B訊框，則將其指派至"跳躍訊框"狀態。可跳過此訊框，此係因為其實質上為一先前訊框之複本。

以顯示次序估計一量化相鄰訊框間之差分的量度為所發生之此處理的第一部分。此量度為以上所提及之距離，藉由該量度，針對每一訊框之適當類型來估計該訊框。因而，在I訊框與相鄰P訊框之間或兩個連續P訊框之間的間距可為變化的。計算量度係藉由利用基於區塊之移動補償器處理視訊訊框而開始的，其中區塊為視訊壓縮之基本單元，儘管諸如8x8、4x4及8x16之其他區塊尺寸可行，但其通常包括16x16個像素。對於由兩個解交錯之場所組成的訊框而言，可基於場來實施移動補償，對參考區塊之搜尋發生於場中而非訊框中。對於在當前訊框之第一場中之區塊而

言，在其後之訊框的場中找到一正向參考區塊，同樣，在緊接在當前場前之訊框的場中找到一反向參考區塊。可將該等當前區塊組合成一補償場。該過程繼續該訊框之第二場。該等兩個補償場經組合而形成一正向補償訊框及一反向補償訊框。

對於在反向電視電影處理606中所產生之訊框而言，因為僅重建之膠捲形成訊框，所以對參考區塊之搜尋係僅基於訊框。找到兩個參考區塊及兩個差分(正向及反向)，其亦導致一正向補償訊框及一反向補償訊框。總之，移動補償器產生每一區塊之移動向量及差分量度，但在解交錯器605之輸出正被處理之情形中，區塊為NTSC場之一部分，且若反向電視電影處理之輸出被處理，則區塊為膠捲訊框之一部分。注意，視是否正估計一正向差分還是反向差分而定，在一先前場或訊框，或在緊接於其後之場或訊框中，正研究之場或訊框中之區塊與一最匹配該區塊的區塊之間來估計量度之差分。僅亮度值參與此計算。

移動補償步驟因而產生兩組差分。此等差分在當前亮度值之區塊與參考區塊中的亮度值之區塊之間，該等參考區塊自時間上緊接於當前訊框之前及緊接於其後的訊框所取得。為每一像素而判定每一正向及每一反向差分之絕對值，且每一者在整個訊框上被獨立總計。當處理包含一訊框的已解交錯之NTSC場時，兩個場均被包括於兩個總和中。以此方式，找到正向及反向差分之總和絕對值 $SAD_P$ 及 $SAD_N$ 。

對於每一訊框而言，使用該關係來計算SAD比率。

$$\gamma = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N} \quad [22]$$

其中SAD<sub>P</sub>及SAD<sub>N</sub>分別為正向及反向差分的總和絕對值。將一較小正數添加至分子ε來防止"除以零"誤差。將類似之ε項添加至分母，其進一步降低了在SAD<sub>P</sub>或SAD<sub>N</sub>接近於零時γ的敏感性。

在一替代態樣中，該差分可為SSD(均方差分之總和)及SAD(絕對差分之總和)或SATD，其中藉由在取得區塊元素中之差分前將二維離散餘弦變換應用於該等區塊來變換像素值的該等區塊。儘管在其他態樣中可使用一較小區域，但該等總和係在活動視訊之區域上估計而得。

亦計算如接收(非移動補償)之每一訊框的亮度直方圖。直方圖影響DC係數，亦即，在係數之16x16陣列中的(0,0)係數，其為將二維離散餘弦變換應用於亮度值之區塊(若其為可用的)的結果。同等地，可在直方圖中使用16x16區塊中之亮度之256個值的平均值。對於亮度深度為八個位元之影像而言，箱之數目被設定於16。下一量度估計直方圖差分。

$$\lambda = \frac{1}{N} \sum_{i=1}^{16} |N_{Pi} - N_{Ci}| \quad [23]$$

在以上方程式中，N<sub>Pi</sub>為來自第i個箱中之先前訊框之區塊的數目，且N<sub>Ci</sub>為來自屬於第i個箱之當前訊框之區塊的數目，N為一訊框中之區塊的總數目。

組合此等中間結果，以將當前訊框差分量度形成為

$$D = \frac{\gamma_C}{\gamma_P} + \lambda(2\lambda + 1), \quad [24]$$

其中  $\gamma_C$  為基於當前訊框之 SAD 比率，且  $\gamma_P$  為基於先前訊框之 SAD 比率。若一場景具有平滑移動且其亮度直方圖幾乎未改變，則  $D \approx 1$ 。若當前訊框顯示一突然場景改變，則  $\gamma_C$  將較大且  $\gamma_P$  應較小。使用比率  $\frac{\gamma_C}{\gamma_P}$  而非單獨之  $\gamma_C$ ，使得將量度

標準化至該情形之活動水平。

圖 42 說明了將壓縮類型指派至訊框的過程。方程式 19 中所界定之當前訊框差分  $D$  為關於訊框指定而做出決策的基礎。如決策區塊 4202 指示，若研究中之訊框為一序列中之第一者，則遵循標記有 YES 之決策路徑至區塊 4206，藉此宣告該訊框為 I 訊框。在區塊 4208 中將累計之訊框差分設定至零，且該過程返回(在區塊 4210 中)至開始區塊。若正研究之訊框非為一序列中之第一訊框，則自做出決策之區塊 4202 遵循標記有 NO 之路徑，且在測試區塊 4204 中針對場景改變臨限值來測試當前訊框差分。若當前訊框差分大於彼臨限值，則遵循標記有 YES 之決策路徑至區塊 4206，其再次導致 I 訊框之指定。

若當前訊框差分小於場景改變臨限值，則遵循 NO 路徑至區塊 4212，在區塊 4212 中將當前訊框差分添加至累計之訊框差分。在決策區塊 4214 處繼續該流程圖，將累計之訊框差分與臨限值  $t$  相比較，該臨限值  $t$  一般而言小於場景改變臨限值。若累計之訊框差分大於  $t$ ，則控制轉移至區塊 4216，且將訊框指定為 P 訊框，繼而在步驟 4218 中將累計之訊框差

分重設至零。若累計之訊框差分小於 $t$ ，則控制自區塊4214轉移至區塊4220。在此將當前訊框差分與小於 $t$ 之 $\tau$ 相比較。若當前訊框差分小於 $\tau$ ，則在區塊4222中將該訊框指定為跳躍的，且繼而該過程返回，若當前訊框差分大於 $\tau$ ，則在區塊4226中將該訊框指定為B訊框。

### 編碼器

返回參看圖2，轉換編碼器200包括一編碼器228，該編碼器228自預處理器226接收已處理之中繼資料及原始視訊。中繼資料可包括在源視訊104中最初接收之任何資訊及藉由預處理器226計算而得之任何資訊。編碼器228包括一第一遍編碼器230、一第二遍編碼器232及一重編碼器234。編碼器228亦自轉換編碼器控制231接收輸入，該轉換編碼器控制231可將資訊(例如，中繼資料、誤差回彈資訊、內容資訊、已編碼之位元速率資訊，基礎層及增強層平衡資訊以及量化資訊)自第二遍編碼器232提供至第一遍編碼器230、重編碼器234以及預處理器226。編碼器228使用自預處理器226所接收之內容資訊及/或藉由編碼器228自身(例如，藉由內容分類模組712(圖7))所產生之內容資訊，來編碼已接收之視訊。

圖7說明了可包括於一可用於圖2中所說明之編碼器228之例示性兩遍式編碼器中之功能模組的方塊圖。儘管圖7及本文之描述不必陳述可併入一編碼器中之所有功能，但圖7中展示了功能模組之各種態樣。因此，在以下之基礎層及增強層編碼的論述後，下文描述了該等功能模組之一些態



樣。

### 基礎層及增強層編碼

編碼器 228 可為一 SNR 可調性編碼器，其可將來自預處理器 226 之原始視訊及中繼資料編碼成一第一組編碼資料(本文亦稱為一基礎層)及一或多個額外組之編碼資料(本文亦稱為增強層)。編碼演算法產生基礎層及增強層係數，當對該等係數解碼時，其在該等兩個層均可用於解碼時可在解碼器處加以組合。當等兩個層均不可用時，基礎層之編碼允許其被編碼為一單層。

參看圖 31 描述了此多層編碼過程的一個態樣。在區塊 321 處，藉由完全訊框內編碼之巨集區塊(訊框內編碼 MB)編碼一 I 訊框。在 H.264 中，藉由充分利用之空間預測(其提供一顯著量之編碼增益)來編碼 I 訊框中之訊框內編碼 MB。存在兩個子模式：Intra4x4 及 Intra16x16。若基礎層將利用藉由空間預測所提供之編碼增益，則該基礎層需要在編碼及解碼該增強層之前而被編碼及解碼。使用 I 訊框之兩遍編碼及解碼。在基礎層中，基礎層量化參數  $QP_b$  提供變換係數之一粗略量化步長。將在增強層處編碼原始訊框與重建基礎層訊框間之逐像素差分。增強層使用一量化參數  $QP_e$ ，其提供一更精密之量化步長。諸如圖 2 之編碼器 228 的編碼構件可執行在區塊 321 處之編碼。

在區塊 323 處，編碼器編碼正處理之 GOP 中之 P 及 / 或 B 訊框的基礎層資料及增強層資料。諸如編碼器 228 之編碼構件可執行在區塊 323 處之編碼。在區塊 325 處，編碼過程核對

是否存在更多待編碼之P或B訊框。諸如SNR可調性編碼器228之編碼構件可執行動作325。若剩餘更多之P或B訊框，則重複步驟323，直至完成對GOP中之所有訊框的編碼。儘管如以下將論述在P及B訊框中可能存在訊框內編碼MB，但P及B訊框包含訊框間編碼之巨集區塊(訊框間編碼MB)。

為了使解碼器在基礎層資料與增強層資料之間進行區別，編碼器228編碼額外資訊(區塊327)。該等類型之額外資訊包括(例如)識別層之數目的資料、將層識別為一基礎層之資料、將層識別為一增強層之資料、識別層間之層間關係的資料(諸如，層2為用於基礎層1之增強層，或層3為用於增強層2之增強層)或將層識別為一串增強層中之最後增強層的資料。額外資訊可包含於與其相關之基礎層資料及/或增強層資料連接的標頭中，或包含於獨立資料訊息中。諸如圖2之編碼器228的編碼構件可執行在區塊327處之過程。

為了具有單層解碼，必須在反向量化前組合兩個層之係數。因此，必須互動地產生該等兩個層之係數，否則，此可能引入顯著量之額外資訊。額外資訊增加的一個原因為基礎層編碼及增強層編碼可使用不同時間參考。需要演算法來產生基礎層及增強層係數，當該等兩個層為可用時，在解量化之前於解碼器處組合該等係數。同時，演算法應在增強層不可用時提供可接受之基礎層視訊，或解碼器因為(諸如)功率節省之原因而決定不解碼該增強層。在緊接於下文之標準可預測編碼之簡短論述的內容中，以下進一步論述了此過程之說明性實例的細節。

P 訊框(或任何訊框間編碼區段)可利用一當前圖片中之區域與一參考圖片中之最匹配預測區域之間的時間冗餘。可以一移動向量來編碼該參考訊框中之最匹配預測區域的位置。將當前區域與最匹配參考預測區域間之差分稱為剩餘誤差(或預測誤差)。

圖 32 為在(例如)MPEG-4 中之 P 訊框建構過程之一實例的說明。過程 331 為可發生於圖 31 之區塊 323 中之一實例過程之更為詳細的說明。過程 331 包括由 5 x 5 個巨集區塊組成之當前圖片 333，其中在此實例中巨集區塊之數目為任意的。一巨集區塊由 16 x 16 個像素組成。可藉由 8 位元亮度值(Y)及兩個 8 位元色度值(Cr 及 Cb)來界定像素。在 MPEG 中，可以 4:2:0 格式來儲存 Y、Cr 及 Cb 分量，其中在 X 及 Y 方向中將 Cr 及 Cb 分量向下取樣 2。因此，每一巨集區塊將由 256 個 Y 分量、64 個 Cr 分量及 64 個 Cb 分量組成。在與當前圖片 333 不同之時間點自參考圖片 337 預測當前圖片 333 之巨集區塊 335。在參考圖片 337 中實施搜尋以定位在 Y、Cr 及 Cb 值方面最接近於正編碼之當前巨集區塊 335 的最匹配之巨集區塊 339。以移動向量 341 來編碼參考圖片 337 中之最匹配巨集區塊 339 的位置。參考圖片 337 可為解碼器將在當前圖片 333 之建構前進行重建之 I 訊框或 P 訊框。自當前巨集區塊 335(計算 Y、Cr 及 Cb 分量之每一者的差分)減去最匹配巨集區塊 339，從而導致剩餘誤差 343。以 2 維離散餘弦變換(DCT)345 來編碼剩餘誤差 343，且繼而對該剩餘誤差 343 進行量化 347。可執行量化 347，以藉由(例如)將較少之位元分

配至高頻率係數同時將較多之位元分配至低頻率係數來提供空間壓縮。剩餘誤差343之量化係數連同移動向量341及參考圖片333識別資訊為表示當前巨集區塊335之已編碼資訊。該已編碼資訊可儲存於記憶體中以供未來使用，或可為了(例如)誤差校正或影像增強之目的而加以運用，或在網路349上進行傳輸。

剩餘誤差343之已編碼量化係數連同已編碼移動向量341可用於在編碼器中重建當前巨集區塊335，以用作後續之移動估計及補償之參考訊框的一部分。編碼器可仿效用於此P訊框重建之解碼器的程序。解碼器之仿效將導致編碼器及解碼器兩者均對相同參考圖片起作用。在此呈現了無論實施於一編碼器中(用於進一步訊框間編碼)或一解碼器中之重建過程。在重建參考訊框(或正參考之一圖片或訊框的一部分)後，可開始P訊框之重建。對已編碼之量化係數進行解量化351，且繼而執行2維反向DCT或IDCT 353，從而導致解碼或重建之剩餘誤差355。將已編碼之移動向量341解碼且將其用於在已重建之參考圖片337中定位已重建之最匹配巨集區塊357。繼而將重建之剩餘誤差355添加至重建之最匹配巨集區塊357，以形成重建之巨集區塊359。重建之巨集區塊359可儲存於記憶體中、獨立顯示或與其他重建之巨集區塊一起在一圖片中或為影像增強而進一步加以處理。

B訊框(或藉由雙向預測所編碼之任一區段)可利用一當前圖片中之區域、一先前圖片中之最匹配預測區域、一後

續圖片中之最匹配預測區域間的時間冗餘。將後續最匹配預測區域及先前最匹配預測區域組合而形成一組合之雙向預測區域。當前圖片區域與最匹配組合之雙向預測區域間之差分為剩餘誤差(或預測誤差)。可以兩個移動向量來編碼在後續參考圖片中之最匹配預測區域及在先前參考圖片中之最匹配預測區域的位置。

圖33說明了一可藉由編碼器228執行之用於編碼基礎層及增強層係數之編碼器過程的一實例。編碼基礎層及增強層以提供一SNR可調性位元流。圖33描繪了諸如將在圖31之步驟323中執行之編碼訊框間MB剩餘誤差係數的一實例。然而，類似方法亦可用於編碼訊框內MB係數。諸如圖2之編碼器組件228的編碼構件可執行圖33中所說明之過程及圖32之步驟323。將原始(待編碼)視訊資料406(在此實例中視訊資料包含亮度及色度資訊)輸入至一基礎層最匹配巨集區塊迴路302及一增強層最匹配巨集區塊迴路365。迴路363及365兩者之目標為最小化分別在加法器367及369處所計算之剩餘誤差。可平行(如所示)或順次執行迴路363及365。迴路363及365包括分別用於搜尋含有參考訊框之緩衝器371及373的邏輯，以識別最小化最匹配巨集區塊與原始資料361間之剩餘誤差的最匹配巨集區塊(緩衝器371及373可為相同緩衝器)。因為基礎層迴路363將通常利用一比增強層迴路365粗略之量化步長(較高QP值)，所以迴路363及迴路365之剩餘誤差將不同。變換區塊375及377變換每一迴路之剩餘誤差。

繼而在選擇器 379 中將已變換之係數剖析為基礎層係數及增強層係數。如以下所論述，選擇器 379 之剖析可呈現若干形式。剖析技術之一共同特徵為計算增強層係數  $C'_{enh}$ ，使其為對基礎層係數  $C'_{base}$  之差分改進。計算增強層(待為對基礎層之改進)允許解碼器藉由其自身解碼基礎層係數且具有影像之合理表示，或將基礎層及增強層係數組合且具有影像之改進的表示。繼而，藉由量化器 381 及 383 來量化選擇器 379 所選擇之係數。已量化之係數  $\tilde{C}'_{base}$  及  $\tilde{C}'_{enh}$  (分別藉由量化器 381 及 383 計算而得) 可儲存於記憶體中，或在一網路上傳輸至一解碼器。

為了匹配一解碼器中之巨集區塊之重建，解量化器 385 解量化基礎層剩餘誤差係數。將已解量化之剩餘誤差係數反向變換 387 並將其添加 389 至緩衝器 371 中所找到之最匹配巨集區塊，從而導致匹配解碼器中將重建之內容的重建巨集區塊。量化器 383、解量化器 391、反向變換器 393、加法器 397 及緩衝器 373 如在基礎層迴路 363 中所實施般在增強迴路 365 中執行類似的計算。此外，將加法器 393 用於組合已解量化之用於增強層重建的增強層及基礎層係數。增強層量化器及解量化器將通常利用比基礎層更精密之量化器步長(較低 QP)。

圖 34、35 及 36 展示了可用於圖 33 之選擇器 379 中之基礎層及增強層係數選擇器過程的實例。諸如圖 2 之編碼器 228 的選擇構件可執行在圖 34、35 及 36 中所描繪之過程。將圖 34 用作一實例，如下列方程式中所示，將已變換之係數剖析

為基礎層係數及增強層係數。

$$C'_{base} = \begin{cases} 0, & \text{若 } C_{base} \text{ 及 } C_{enh} \text{ 為相反符號} \\ \min(C_{base}, C_{enh}), & \text{否則} \end{cases} \quad [25]$$

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base})) \quad [26]$$

其中 "min" 函數可為一數學最小值或該等兩個自變量之最小量值。在圖 34 中，將方程式 25 描繪為區塊 401，且將方程式 26 經描繪為加法器 510。在方程式 26 中， $Q_b$  代表基礎層量化器 381，且  $Q_b^{-1}$  代表基礎層之解量化器 385。方程式 26 將增強層係數轉換成藉由方程式 25 計算而得之基礎層係數之一差分改進。

圖 35 為基礎層及增強層係數選擇器 379 之另一實例的說明。在此實例中，區塊 405 中所包含之方程式 (.) 表示下列內容：

$$C'_{base} = \begin{cases} C_{base}, & \text{若 } |Q_b^{-1}Q_b(C_{base}) - C_{enh}| < |C_{enh}| \\ 0, & \text{否則} \end{cases} \quad [27]$$

加法器 407 如下列兩個方程式中所示計算增強層係數：

$$C'_{enh} = C_{enh} - Q_b^{-1}(Q_b(C'_{base})) \quad [28]$$

其中藉由方程式 27 給出  $C'_{base}$ 。

圖 36 為基礎層及增強層選擇器 379 之另一實例的說明。在此實例中，基礎層係數未改變，且增強層等於量化/反向量化基礎層係數與原始增強層係數間之差分。

除了基礎層及增強層剩餘誤差係數外，解碼器需要識別如何編碼 MB 之資訊。諸如圖 2 之編碼器組件 228 之編碼構件可編碼額外資訊，該額外資訊可包括訊框內編碼及訊框間

編碼部分之映射，諸如一MB映射，其中巨集區塊(或子巨集區塊)被識別為經訊框內編碼或訊框間編碼(亦識別訊框間編碼之類型，包括(例如)正向、反向或雙向)及訊框訊框間編碼部分參看哪一(些)訊框。在一實例態樣中，在基礎層中編碼MB映射及基礎層係數，且在增強層中編碼增強層係數。

P訊框及B訊框可含有訊框內編碼MB以及訊框間MB。混合視訊編碼器通常使用速率失真(RD)最佳化來決定將P或B訊框中之一些巨集區塊編碼為訊框內編碼MB。為了具有單層解碼(其中訊框內編碼MB並不取決於增強層訊框間MB)，任何鄰近之訊框間MB不用於基礎層訊框內編碼MB之空間預測。為了針對增強層解碼而保持計算複雜性不改變，對於基礎層P或B訊框中之訊框內編碼MB而言，可跳過增強層處之改進。

P或B訊框中之訊框內編碼MB需要比訊框間MB多得多之位元。為此原因，僅可以較高QP在基礎層品質下來編碼P或B訊框中之訊框內編碼MB。此將引入視訊品質之一定下降，但若如以上論述藉由基礎層及增強層中之訊框間MB係數而在隨後之訊框中對此下降加以改進，則此下降應為不顯明的。兩個原因使此下降為不顯明的。第一個原因為人類視覺系統(HVS)之特徵，且另一原因為訊框間MB改進訊框內MB。藉由將位置自一第一訊框改變至一第二訊框之對象，在第一訊框中之某些像素在第二訊框中為不可見的(待被覆蓋之資訊)，且在第二訊框中之某些像素在第一時間內



為可見的(未被覆蓋之資訊)。人類眼睛對於未被覆蓋及待被覆蓋之視覺資訊並不敏感。因此，對於未被覆蓋之資訊而言，即使以較低品質對其加以編碼，眼睛亦不可分辨該差異。若相同資訊保留於隨後之P訊框中，則將存在增強層處之隨後之P訊框可對其進行改進的較高可能性，此係因為增強層具有較低QP。

將在P或B訊框中引入訊框內編碼MB之另一通用技術稱作訊框內再新。在此情形中，儘管標準R-D最佳化將指定該等MB應為訊框間編碼MB，但仍將某些MB編碼為訊框內編碼MB。包含於基礎層中之此等訊框內編碼MB可藉由 $QP_b$ 或 $QP_e$ 加以編碼。若將 $QP_e$ 用於基礎層，則在增強層處無需改進。若將 $QP_b$ 用於基礎層，則可能需要改進，否則在增強層處，品質之下降將為顯明的。因為在編碼效率方面訊框間編碼比訊框內編碼更為有效，所以將訊框間編碼增強層處之此等改進。因此，基礎層係數將不用於增強層。因此，在不引入新操作之情況下，品質在增強層處得以改良。

B訊框通常使用於增強層中，此係因為其提供之壓縮品質較高。然而，B訊框可能必須參考P訊框之訊框內編碼MB。如以上所論述，若將在增強層品質下編碼B訊框之像素，則可能由於P訊框訊框內編碼MB之較低品質而需要過多位元。如以上所論述，藉由利用HVS之品質，當參考P訊框之較低品質之訊框內編碼MB時，可在較低品質下編碼B訊框MB。

P或B訊框中之訊框內編碼MB的極端情形為當P或B訊框

中之所有MB由於在正編碼之視訊中存在場景改變而在訊框內模式下加以編碼時。在此情形中，可在基礎層品質下編碼整個訊框且增強層處無改進。若在B訊框處發生場景改變且假定僅在增強層中編碼B訊框，則B訊框可在基礎層品質下編碼或僅被遺漏。若在P訊框處發生場景改變，則可無需改變，但可在基礎層品質下遺漏或編碼該P訊框。可調性層之編碼進一步描述於名稱為"SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING"之美國專利申請案第[代理人案號/參考號050078]號中。

#### 編碼器第一遍部分

圖7展示了圖2之編碼器228的說明性實例。所展示之區塊說明了可包括於編碼器228中之各種編碼器處理。在此實例中，編碼器228包括在一分界線704上之第一遍部分706及在線704下之第二遍部分706(包括圖2中之第二遍編碼器232及重編碼器234的功能)。

編碼器228自預處理器226接收中繼資料及原始視訊。中繼資料可包括所接收或由預處理器226計算而得之任何中繼資料，其包括關於視訊之內容資訊的中繼資料。編碼器228之第一遍部分702說明了可包括於第一遍編碼702中(以下關於其功能將進行描述)之例示性過程。如熟習此項技術者將知道，可以各種形式(例如，硬體、軟體、韌體或其組合)體現此功能。

圖7說明了一自適應訊框內再新(AIR)模組。AIR模組710將一輸入提供至一I訊框實例化模組708，該模組708基於中

繼資料實例化—I訊框。第一遍部分702亦可包括一內容分類模組712，其經組態以接收中繼資料及視訊且判定關於該視訊之內容資訊。可將內容資訊提供至一速率控制位元分配模組714，該模組714亦接收中繼資料及視訊。控制位元分配模組714判定速率位元控制資訊，且將其提供至模式決策模組715。可將內容資訊及視訊提供至一訊框內模型(失真)模組716，該模組716將訊框內編碼失真資訊提供至模式決策模組715及基礎層及增強層模組718之可調性速率失真。將視訊及中繼資料提供至一移動估計(失真)模組720，該模組720將訊框間編碼失真資訊提供至基礎層及增強層模組718之可調性速率失真。基礎層及增強層模組718之可調性速率失真使用自移動估計模組720及訊框內模型失真模組716之失真估計而判定可調性速率失真資訊，其被提供至模式決策模組715。模式決策模組715亦自片/MB排序模組722接收輸入。片/MB排序模組722自一誤差回彈模組740(展示於第二遍部分706中)接收輸入，且將關於將視訊之獨立可編碼部分(片)與用於誤差回彈之存取單元邊界進行對齊的資訊提供至模式決策模組715。模式決策模組715基於其輸入來判定編碼模式資訊，且將"最佳"編碼模式提供至第二遍部分706。以下描述了此第一遍部分702編碼之某些實例的說明性解釋。

如上文所陳述，內容分類模組712接收藉由預處理器226所提供之中繼資料及原始視訊。在某些實例中，預處理器226計算來自多媒體資料之內容資訊，且將內容資訊提供至

內容分類模組712(例如，在中繼資料中)，該模組712可使用內容資訊來判定用於多媒體資料之內容分類。在某些其他態樣中，內容分類模組712經組態以自多媒體資料判定各種內容資訊，且亦可經組態以判定一內容分類。

內容分類模組712可經組態以判定具有不同類型之內容之視訊的不同內容分類。不同內容分類可導致不同參數，該等參數用於編碼多媒體資料之態樣中，例如，判定一用於判定量化參數、移動估計、可調性、誤差回彈的位元速率(例如，位元分配)，跨通道保持最佳多媒體資料品質，及用於快速通道切換方案(例如，週期性地強迫I訊框允許快速通道切換)。根據一實例，編碼器228經組態以基於內容分類來判定速率失真(R-D)最佳化及位元速率分配。判定內容分類允許基於一內容分類將多媒體資料壓縮至一對應於一所要之位元速率之給定品質水平。又，藉由分類多媒體資料之內容(例如，基於人類視覺系統判定一內容分類)，取決於視訊內容，產生了在一接收裝置之顯示器上之所傳達的多媒體資料之所得感知品質。

作為內容分類模組712所經受之用以分類內容的程序的實例，圖9展示了說明一例示性過程之過程900，藉由該過程內容分類模組712可進行操作。如所示，過程900始於輸入區塊902，在區塊902中內容分類模組712接收原始多媒體資料及中繼資料。過程900繼而進行至區塊904，在區塊904中內容分類模組712判定多媒體資料之空間資訊及時間資訊。在某些態樣中，藉由空間遮罩及時間遮罩(例如，過

濾)來判定空間資訊及時間資訊。可基於包括場景改變資料及移動向量(MV)平滑的中繼資料來判定空間資訊及時間資訊。過程900繼而進行至區塊912，其執行空間複雜性、時間複雜性及敏感性估計。過程900繼而進行至區塊916，在區塊916中基於在區塊904及912中判定之空間、時間及敏感性資料之結果來分類多媒體資料之內容。亦在區塊916中，可選擇特定速率失真(R-D)曲線，及/或可更新R-D曲線資料。過程900繼而進行至輸出區塊918，在區塊918中該輸出可包括一指示空間及時間活動(例如，內容分類)之複雜性失真映射或值及/或所選擇之R-D曲線。返回參看圖7，內容分類模組712將一輸出提供至以上所論述之速率控制位元分配模組714、訊框內模型(失真)模組716且亦提供至I訊框實例化模組708。

### 內容資訊

內容分類模組712可經組態以計算來自多媒體資料之各種內容資訊，包括各種關於內容之量度，包括空間複雜性、時間複雜性、對比率值、標準偏差及訊框差分量度(以下對其進一步加以描述)。

內容分類模組712可經組態以判定多媒體資料之空間複雜性及時間複雜性，且亦將一紋理值與空間複雜性相關聯，且將一移動值與時間複雜性相關聯。內容分類模組712自預處理器226接收已預處理之關於進行編碼之多媒體資料之內容的內容資訊，或者，預處理器226可經組態以計算內容資訊。如以上所描述，內容資訊可包括(例如)一或多個

$D_{csat}$ 值、對比率值、移動向量(MV)及絕對差分之總和(SAD)。

一般而言，多媒體資料包括一或多個序列之影像或訊框。可將每一訊框分解成用於處理之像素區塊。空間複雜性為一廣義術語，其通常描述在一訊框內之空間細節之程度的量測。具有亮度及色度之主要平面或未改變或較少改變之區域的場景將具有低空間複雜性。空間複雜性與視訊資料之紋理相關聯。在此態樣中，空間複雜性係基於稱作 $D_{csat}$ 之人類視覺敏感性量度，為每一區塊而計算該量度作為局部空間頻率及環境照明的函數。一般熟練之技工知道使用空間頻率圖案及視覺影像之照明及對比特性來利用人類視覺系統的技術。許多敏感性量度因利用人類視覺系統之透視限制而眾所周知，且可用於本文所描述之方法。

時間複雜性為一廣義術語，其用於通常描述如在一序列訊框中之訊框間參考之多媒體資料中之移動程度的量測。具有極小移動或無移動之場景(例如，視訊資料之訊框之序列)具有低時間複雜性。可為每一巨集區塊而計算時間複雜性，且其可基於 $D_{csat}$ 值、移動向量及一訊框與另一訊框(例如，一參考訊框)間之絕對像素差分的總和。

訊框差分量度給出兩個連續訊框間之差分的量測，其將移動量(例如，移動向量或MV)連同表示為預測子與當前巨集區塊間之絕對差分之總和(SAD)的剩餘能量考慮在內。訊框差分亦提供雙向或單向預測效率之量測。

基於自一預處理器(其潛在執行移動補償解交錯)所接收之移動資訊的訊框差分量度之一實例為如下。解交錯器執

行一雙向移動估計，且因而雙向移動向量及SAD資訊為可用的。可如下所得出藉由每一巨集區塊之SAD\_MV所表示之訊框差分：

$$\text{SAD\_MV} = \log_{10} [\text{SAD} * \exp(-\min(1, \text{MV}))] \quad [29]$$

其中  $\text{MV} = \text{Square\_root}(\text{MV}_x^2 + \text{MV}_y^2)$ 、 $\text{SAD} = \min(\text{SAD}_N, \text{SAD}_P)$ ，其中  $\text{SAD}_N$  為自反向參考訊框計算而得之SAD，且  $\text{SAD}_P$  為自正向參考訊框計算而得之SAD。

以上參考方程式6-8描述了估計一訊框差分之另一方法。可如以上在方程式6中較早描述來計算SAD比率(或對比率) $\gamma$ 。亦可判定每一訊框之亮度直方圖，直方圖差分 $\lambda$ 係使用方程式7計算而得。可如方程式8中所示計算訊框差分量度D。

在一說明性實例中，以下列方式利用對比率及訊框差分量度以獲得視訊內容分類，其可在一給定視訊序列中可靠地預測該等特徵。儘管在此被描述為發生於編碼器228中，但預處理器226亦可經組態以判定一內容分類(或其他內容資訊)，且將內容分類經由中繼資料而傳遞至編碼器228。在以下實例中所描述之過程將內容分類成八個可能之類別，其類似於自基於R-D曲線之分析所獲得之分類。分類過程為每一超訊框輸出一在0與1間之範圍內的值，其視場景之複雜性及在彼超訊框中場景改變發生之數目而定。在預處理器中之內容分類模組可為每一超訊框執行下列步驟(1)-(5)，以自訊框對比及訊框差分值獲得內容分類量度。

1. 自巨集區塊對比值來計算平均訊框對比及訊框對比偏

差。

2. 使用自模擬所獲得之值(分別為40及5)來標準化訊框對比及訊框差分值。

3. 使用(例如)廣義方程式來計算內容分類量度：

$$\begin{aligned} \text{CCMetric} = & \text{CCW1} * \text{I\_Frame\_Contrast\_Mean} + \text{CCW2} \\ & * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}^2 * \\ & \exp(\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \quad [30] \end{aligned}$$

其中CCW1、CCW2、CCW3及CCW4為加權係數。在此實例中，將該等值選擇為CCW1為0.2、CCW2為0.9、CCW3為0.1且CCW4為-0.00009。

4. 判定在超訊框中場景改變之數目。一般而言，超訊框指代在一特定時間週期中可顯示之一組圖片或訊框。通常，該時間週期為一秒。在某些態樣中，一超訊框包含30個訊框(30/fps視訊)。在其他態樣中，一超訊框包含24個訊框(24/fps視訊)。視場景改變之數目而定，執行下列情形中之一者。

(a) 無場景改變：當在超訊框中不存在場景改變時，量度完全僅視訊框差分值而定，如下列方程式中所示：

$$\begin{aligned} \text{CCMetric} = & (\text{CCW2} + (\text{CCW1}/2)) * \text{Frame\_Difference\_Mean} - (\text{CCW3} - \\ & (\text{CCW1}/2)) * 1 * \exp(-\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \end{aligned}$$

[31]

(b) 單場景改變：當存在在超訊框中所觀測之單場景改變訊框時，將使用預設方程式來計算量度，如以下所示：



$$\text{CCMetric} = \text{CCW1} * \text{I\_Frame\_Contrast\_Mean} + \text{CCW2} * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}^2 * \exp(\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \quad [32]$$

(c) 兩個場景改變：當觀測到在給定超訊框中存在至多 2 個場景改變時，相較於第一超訊框，最後之超訊框被給予了更多權重，此係因為該第一超訊框無論如何將很快藉由後者加以再新，如下列方程式中所示：

$$\text{CCMetric} = 0.1 * \text{I\_Frame\_Contrast\_Mean}_1 + \text{CCW1} * \text{I\_Frame\_Contrast\_Mean}_2 + (\text{CCW2} - 0.1) * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}_1^2 * \text{I\_Contrast\_Deviation}_2^2 * \exp(\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \quad [33]$$

(d) 三個或三個以上之場景改變：若觀測到給定超訊框具有 3 個以上之 I 訊框 (比如 N 個)，則最後之 I 訊框給予更多權重，且所有其他 I 訊框給予 0.05 之權重，如下列方程式中所示：

$$\text{CCMetric} = 0.05 * \text{I\_Frame\_Contrast\_Mean}_{(1\dots N-1)} + \text{CCW1} * \text{I\_Frame\_Contrast\_Mean}_{(N)} + (\text{CCW2} - (0.05 * (N-1))) * \text{Frame\_Difference\_Mean} - \text{CCW3} * \text{I\_Contrast\_Deviation}_{(N)}^2 * \text{I\_Contrast\_Deviation}_{(1\dots N-1)}^2 * \exp(\text{CCW4} * \text{Frame\_Difference\_Deviation}^2) \quad [34]$$

5. 當訊框差分平均值小於 0.05 時，在低移動場景之情形中可將校正用於量度。可將 (CCOFFSET) 0.33 之偏差添加至 CCMetric。

內容分類模組 712 使用  $D_{csat}$  值、移動向量及/或絕對差分之總和以判定一指示巨集區塊之空間複雜性的值(或指定量之視訊資料)。藉由訊框差分量度之量測來判定時間複雜性(在兩個連續訊框間之差分，其將移動量、移動向量及該等訊框間之絕對差分的總和考慮在內)。

在某些態樣中，內容分類模組 712 可經組態以產生一頻寬映射。舉例而言，若預處理器 226 不產生一頻寬映射，則可藉由內容分類模組 712 來執行頻寬映射產生。

### 判定紋理及移動值

對於多媒體資料中之每一巨集區塊而言，內容分類模組 712 將一紋理值與空間複雜性相關聯，且將一移動值與時間複雜性相關聯。紋理值係關於多媒體資料之發光值，其中低紋理值指示資料之鄰近像素之發光值中的較小改變，且高紋理值指示資料之鄰近像素之發光值中的較大改變。一旦計算紋理值及移動值，則內容分類模組 712 藉由研究移動及紋理資訊兩者來判定一內容分類。內容分類模組 712 將被分類之視訊資料的紋理與一相對紋理值相關聯，例如，"低"紋理、"中等"紋理或"高"紋理，其通常指示巨集區塊之亮度值的複雜性。又，內容分類模組 712 將為被分類之視訊資料所計算之移動值與一相對移動值相關聯，例如"低"移動、"中等"移動或"高"移動，其通常指示巨集區塊之移動量。在替代態樣中，可使用用於移動及紋理之較少或較多種類。繼而，藉由研究相關聯之紋理值及移動值來判定內容分類量度。內容分類之一說明性態樣的進一步描述揭示於在

2006年3月10日申請之同在申請中的名稱為"CONTENT CLASSIFICATION FOR MULTIMEDIA PROCESSING"之美國專利申請案第11/373,577號中，該申請案讓渡於此受讓人且藉此以引用之方式明確地併入本文中。

圖8說明了一說明如何將紋理值及移動值與內容分類相關聯之分類圖的一實例。一般熟習此項技術者熟悉用以實施此分類圖之許多方法，例如，在一查找表或一資料庫中。基於視訊資料內容之預定估計來產生該分類圖。為了判定視訊資料分類，"低"、"中等"或"高"(在"x軸"上)之紋理值與"低"、"中等"或"高"(在"y軸"上)之移動值相互參考。將在相交區塊中所指示之內容分類指派至視訊資料。舉例而言，"高"之紋理值及"中等"之移動值導致分類七(7)。圖8說明了在此實例中與八個不同內容分類相關聯之相對紋理及移動值的各種組合。在某些其他態樣中，可使用較多或較少之分類。內容分類之一說明性態樣的進一步描述揭示於在2006年3月10日申請之名稱為"CONTENT CLASSIFICATION FOR MULTIMEDIA PROCESSING"的同在申請中之美國專利申請案第11/373,577號中，其讓渡於此受讓人且藉此以引用之方式明確地併入本文中。

#### 速率控制位元分配

如本文中所描述，多媒體資料內容分類可用於編碼演算法中以有效地改良位元管理，同時保持視訊之恆定感知品質。舉例而言，可將分類量度用於場景改變偵測、編碼位元速率分配控制及訊框速率上轉換(FRUC)之演算法中。壓

縮器/減壓器(編解碼器)系統及數位信號處理演算法通常用於視訊資料通信中，且可經組態以節省頻寬，但存在品質與頻寬節省間之折衷。最佳編解碼器提供最多頻寬節省，同時產生最少之視訊品質降級。

在一說明性實例中，速率控制位元分配模組714使用內容分類來判定位元速率(例如，經分配而用於編碼多媒體資料之位元的數目)，且將位元速率儲存於記憶體中以由編碼器228之其他過程及組件來使用。自視訊資料之分類所判定的位元速率可幫助節省頻寬，同時提供在一致品質水平下之多媒體資料。在一個態樣中，可將不同位元速率與八個不同內容分類中之每一者相關聯，且繼而使用彼位元速率來編碼多媒體資料。所得之效應為儘管多媒體資料之不同內容分類分配有用於編碼之不同數目的位元，但在一顯示器上觀測時，所感知之品質為類似或一致的。

一般而言，具有較高內容分類之多媒體資料指示較高位之移動及/或紋理，且在編碼時分配有較多位元。具有較低分類(指示較少紋理及移動)之多媒體資料分配有較少位元。對於特定內容分類之多媒體資料而言，可基於為觀測多媒體資料而選擇之目標感知品質水平來判定位元速率。判定多媒體資料品質可藉由人類觀測及分級多媒體資料來判定。在某些替代態樣中，可藉由使用(例如)信雜比演算法之自動測試系統來進行多媒體資料品質之估計。在一個態樣中，為每一內容分類之多媒體資料來預定一組標準品質水平(例如，五個)及達成每一特定品質水平所需之對應的位

元速率。為了判定一組品質水平，可藉由產生平均意見分數(MOS)來估計特定內容分類之多媒體資料，該平均意見分數(MOS)在使用某一位元速率編碼多媒體資料時提供多媒體資料之視覺感知品質的數值指示。MOS可表示為在範圍1至5內的一單數字，其中1為最低感知品質，且5為最高感知品質。在其他態樣中，MOS可具有超過五或少於五之品質水平，且可使用每一品質水平的不同描述。

判定多媒體資料品質可藉由人類觀測及分級多媒體資料來判定。在某些替代態樣中，可藉由使用(例如)信雜比演算法之自動測試系統來進行多媒體資料品質之估計。在一個態樣中，為每一內容分類之多媒體資料來預定一組標準品質水平(例如，五個)及達成每一特定品質水平所需之對應的位元速率。

獲知某一內容分類之多媒體資料的視覺感知品質水平與位元速率間之關係可藉由選擇一目標(例如，所要之)品質水平來判定。用以判定位元速率之目標品質水平可預選、由使用者來選擇、經由需要來自使用者或來自另一過程之輸入的自動過程或半自動過程來選擇(或由編碼裝置或系統基於預定標準來動態地選擇。可基於(例如)編碼應用之類型或將接收多媒體資料之用戶端裝置的類型來選擇目標品質水平。

在圖7中之所說明的實例中，速率控制位元分配模組714接收來自內容分類模組712之資料及直接來自預處理器226之中繼資料兩者。速率控制位元分配模組714常駐於編碼器

228之第一遍部分中，且速率控制微調模組738常駐於第二遍部分706中。此二遍式速率控制態樣經組態，使得第一遍(速率控制位元分配模組714)執行具有一超訊框預覽(例如，將256 kbps之長期平均位元速率作為目標)之上下文自適應位元分配，且限制峰值速率，且第二遍(速率控制微調模組738)改進兩層可調性之第一遍結果且執行速率適應。速率控制對四個位準起作用：(1)GOP位準-將I、P、B及F訊框之位元分配控制為在GOP內部係非均衡的；(2)超訊框位準-控制對最大超訊框尺寸之硬極限；(3)訊框位準-根據基於內容資訊(例如，內容分類)之多媒體資料訊框的空間及時間複雜性來控制位元要求；及(4)巨集區塊位準-基於空間及時間複雜性映射(其係基於內容資訊(例如，內容分類)來控制巨集區塊之位元分配。

在圖10中說明了速率控制模組714之操作的例示性流程圖。如圖10中所示，過程1000始於輸入1002區塊。速率控制模組714接收各種輸入，圖7不必說明所有輸入。舉例而言，輸入資訊可包括來自預處理器226之中繼資料、目標位元速率、編碼器緩衝器尺寸(或作為一當量，用於速率控制之最大延遲時間)、初始速率控制延遲及訊框速率資訊。進一步之輸入資訊可包括在圖片組(GOP)位準下之輸入，包括(例如)最大超訊框尺寸、GOP之長度及P/B訊框分配(包括場景改變資訊)、所要之基礎層及增強層的配置、用於未來30個訊框之GOP中之圖片的複雜性失真量度。其他輸入資訊包括在圖片位準下之輸入，包括當前圖片(接收自內容分類

模組 712)之複雜性-失真映射、量化參數(QP)及過去之30個訊框(套在一滑動窗口上)的位元分解。最後，在巨集區塊(MB)位準下之輸入資訊包括(例如)在一參考圖片中之並置巨集區塊(MB)的平均絕對差分(MAD)及量化(無論是否跳躍)後之巨集區塊的編碼區塊圖案(CBP)。

在區塊 1002 處之輸入之後，過程 1000 進行至區塊 1004 以進行初始化從而編碼位元流。同時，執行緩衝器初始化 1006。接著，如區塊 1008 中所示初始化 GOP，其中 GOP 位元分配 1010 作為初始化之一部分而被接收。在 GOP 初始化之後，流程進行至區塊 1012，在區塊 1012 中初始化一片。此初始化包括如由區塊 1014 所示之標頭位元的更新。在執行區塊 1004、1008 及 1012 之初始化之後，如由區塊 1016 所示進行一基礎單元或巨集區塊(MB)的速率控制(RC)。作為區塊 1016 中之巨集區塊之速率控制判定的一部分，經由編碼器 228 中之介面接收輸入。此等輸入可包括巨集區塊(MB)位元分配 1018、二次式模型參數之更新 1020 及自中數("MAD"，分散之穩固估計)參數 1022 之中數絕對偏差的更新。接著，過程 1000 進行至區塊 1024 以執行在編碼一圖片之後的操作 1024。此程序包括如由區塊 1026 所示接收緩衝器參數之一更新。過程 1000 繼而進行至輸出區塊 1028，在輸出區塊 1028 中速率控制模組 714 輸出待由如圖 7 中所示之模式決策模組 715 使用的每一巨集區塊 MB 之量化參數 QP。

### 移動估計

移動估計模組 720 自預處理器 226 接收中繼資料及原始視

訊之輸入，且將可包括區塊尺寸、移動向量失真量度及參考訊框識別符之輸出提供至一模式決策模組715。圖11說明了移動估計模組720之例示性操作。如所示，過程1100始於輸入1102。在訊框位準下，模組720接收參考訊框ID及移動向量之輸入。在巨集區塊位準下，輸入1102包括輸入像素及參考訊框像素。過程1100繼續至步驟1104，其中執行色彩移動估計(ME)及移動向量預測。為了進行此過程，接收各種輸入，包括MPEG-2移動向量及亮度移動向量MV 1106、移動向量平滑1108及非因果移動向量1110。接著，過程1100進行至區塊1112，在區塊1112中執行移動向量搜尋演算法或方法，諸如六邊形或菱形搜尋方法。至區塊1112處之過程的輸入可包括如由區塊1114所示的絕對差分之總和(SAD)、均方差分之總和(SSD)及/或其他量度。一旦執行移動向量搜尋，則過程1100進行至終止區塊1116，在終止區塊1116中執行終止處理。過程1000繼而結束於輸出區塊1118處，該輸出區塊1118產生區塊尺寸、移動向量(MV)、失真量度及參考訊框識別符之輸出。

### 基礎層及增強層之可調性R-D

圖13說明了可藉由可調性R-D模組718執行之可調性過程1300的例示性流程圖。過程1300始於開始區塊1302，且進行至區塊1304，在區塊1304中可調性R-D模組718自移動估計模組720接收一輸入且執行移動估計。移動估計如由區塊1306所指示依賴於基礎層參考訊框、增強層參考訊框及待編碼之原始訊框的輸入。此資訊可藉由GOP分割器612計算



而得，且經由(例如)中繼資料將此資訊傳達至可調性R-D模組718。過程1300進行至區塊1308以判定基礎層資料及增強層資料之可調性資訊。接著，如區塊1310中所示執行基礎層編碼，接著為在區塊1312中之增強層編碼。如由區塊1314所說明，增強層之編碼可將用於層間預測之基礎層編碼結果用作一輸入，因而在時間上增強層編碼在基礎層編碼之後執行。此進一步描述於同在申請中之名稱為"SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING"的美國專利申請案第[代理人案號/參考號050078]號中。在編碼完成後，過程1300結束於區塊1316處。

#### 片/巨集區塊排序

第一遍部分702亦包括一片/巨集區塊排序模組722，該模組722在第二遍部分中自一誤差回彈模組740接收一輸入，且將片對齊資訊提供至模式決策模組715。片為獨立可解碼之(熵解碼)已編碼視訊資料塊。存取單元(AU)為已編碼之視訊訊框，其中每一者包含一組始終精確含有一個初級編碼之圖片的NAL單元。除了初級編碼之圖片外，一存取單元亦可含有一或多個冗餘編碼之圖片或其他未含有編碼圖片之片或片資料分割的NAL單元。存取單元之解碼總是產生一解碼之圖片。

訊框可為提供最高時間分集之實體層封包(稱作TDM囊)的分時多工區塊。超訊框對應於一個時間單位(例如，1秒)且含有四個訊框。將片及AU邊界對齊至時域中之訊框邊界

會導致被破壞之資料的最有效之分隔及局部化。在一強衰落中，TDM囊中之大部分鄰近資料受到誤差影響。由於時間分集，使得剩餘之TDM囊具有較高之完整機率。可利用未被破壞之資料來恢復並隱藏來自受影響之TDM囊的損失資料。類似邏輯應用於頻域多工(FDM)，其中經由在資料符號調變之頻域中之分隔來獲得頻率分集。此外，類似邏輯應用於在無線網路中通常所應用之空間(經由傳輸器及接收器天線中之分隔)及其他形式的分集。

為了將片及AU對齊至訊框，外碼(FEC)碼塊建立及MAC層封裝同樣應對齊。圖20說明了在片及AU中之已編碼的視訊資料或視訊位元流之組織。可在一或多個位元流中構成已編碼之視訊，例如，其中應用了分層視訊編碼之基礎層位元流及增強層位元流。

視訊位元流包含如在圖20中藉由訊框1' 2005、訊框3' 2010及訊框M' 2015所說明之AU。該等AU包含如由片1 2020、片2 2025及片N 2030所說明之資料的片。片之每一開始藉由一起始碼來識別且被提供至網路適應。一般而言，I訊框或訊框內編碼AU較大，接著為P訊框或正向預測訊框，接著為B訊框。將一AU編碼成多個片在已編碼之位元速率方面會引起顯著額外成本，此係因為限制了跨片之空間預測，且片標頭亦引起額外成本。因為片邊界為再同步點，所以將鄰近實體層封包限制至片控制了誤差，因為當一PLP被破壞時，誤差被限定於PLP中之片，而若PLP含有多個片或多個片之部分，則誤差將影響PLP中之所有片或

片之部分。

因為I訊框通常較大(例如，大約幾十個kbit)，所以額外成本歸因於多個片並非占總I訊框尺寸或總位元速率之較大比例。又，在訊框內編碼之AU中具有較多片使得再同步能夠更好且更頻繁，且使得空間誤差隱藏能夠更有效。又，因為P及B訊框經預測為脫離I訊框，所以I訊框載運視訊位元流中之最重要的資訊。I訊框亦充當用於通道擷取之隨機存取點。

● 現參看圖21，仔細地將I訊框對齊至訊框邊界及同樣將具有I AU之片對齊至訊框邊界使得誤差控制能夠最有效，誤差保護(由於若屬於訊框1 2105之一片損失，則屬於訊框2 2110之片具有較高的完整機率，因為訊框2 2110具有與訊框1 2105誤差恢復之顯著時間分隔)可經由再同步及誤差隱藏來執行。

● 因為通常將P訊框之尺寸定為大約若干個kbit，所以在無效率之有害損耗的情況下，將P訊框之片及整數個P訊框對齊至訊框邊界使得誤差回彈成為可能(由於與I訊框之彼等原因類似的原因)。可在此等態樣中使用時間誤差隱藏。或者，分散連續P訊框使得該等P訊框到達不同訊框內可提供P訊框中之添加的時間分集，此可能因為時間隱藏係基於來自先前重建之I或P訊框的移動向量及資料。B訊框可為極小(幾百個位元)至中等大(若干1000位元)。因此，需要將整數個B訊框對齊至訊框邊界以在無效率之有害損耗的情況下達成誤差回彈。

## 模式決策模組

圖 12 說明了模式決策模組 715 之操作的某些實例。如所示，過程 1200 始於輸入區塊 1202。在一說明性實例中，輸入至模式決策模組 715 之各種資訊包括片類型、Intra 4x4cost、Intra 16x16cost、IntraUV 8x8cost、IntraY 16x16 Mode、IntraUV Mode、移動向量資料(MVD)、量化參數(QP)、SpPredMB4x4Y、SpPredMB16x16Y、SpPredMB8x8U、SpPredMB8x8V、速率失真旗標、原始YMB像素、原始UMB像素及原始VMB像素。過程 1200 繼而進行至編碼初始化之區塊 1204，其可藉由引導編碼器初始化之輸入信號或介面來初始化(如由區塊 1206 所指示)。初始化可包括設定容許模式(其包括跳躍、引導)、設定模式權重(若需要，則預設對於所有模式而言將為相等權重)及設定緩衝器。在初始化之後，過程 1200 進行至區塊 1208，在區塊 1208 中執行模式決策之主處理，其包括：計算每一容許模式之巨集區塊(MB)模式成本、藉由一加權係數來加權每一MB模式成本及選擇一最小MB模式成本模式。此等操作所涉及之輸入包括如由區塊 1210 及 1212 所說明之移動估計(例如，MVD及預測)及空間預測(例如，所有訊框內成本及預測)。與模式決策模組 715 建立介面的為為區塊 1214 中之熵編碼，其改良了壓縮速率。過程 1200 繼而進行至區塊 1216，在區塊 1216 中緩衝器被更新以將資訊傳遞至編碼器第二遍部分 706。最後，過程 1200 進行至區塊 1218，在區塊 1218 中可將"最佳"編碼模式傳達至編碼器第二遍部分 706。

### 編碼器第二遍部分

再次參看圖7，編碼器228之第二遍部分706包括一用於執行第二遍編碼之第二遍編碼器模組232。第二遍編碼器232自模式決策模組715接收輸出。第二遍編碼器232包括一MC/變換量化模組726及一曲折(ZZ)/熵編碼器728。將第二遍編碼器232之結果輸出至一可調性模組730及一位元流填充模組731，該位元流填充模組731經由一同步層240(於圖2中說明)而輸出一已編碼之基礎層及增強層以藉由轉換編碼器200進行傳輸。如圖2中所示，應注意，藉由同步層240將來自第二遍編碼器232及重編碼器234之基礎層及增強層組合成一分封化PES 242(其包括基礎層及增強層)、一資料PES 244(例如，CC及其他本文資料)及一音訊PES 246。應注意，音訊編碼器236接收已解碼之音訊資訊218，且又編碼該資訊並將已編碼之資訊238輸出至同步層240。

### 重編碼器

再次參看圖7，編碼器第二遍部分706亦包括一重編碼器234，其對應於圖2中之重編碼器234。重編碼器234亦接收第一遍部分702之輸出，且包括一MC/變換量化726及ZZ/熵編碼728部分。此外，可調性模組730輸出至重編碼器234。重編碼器234將自重編碼所產生的基礎層及增強層輸出至位元流填充模組731，以傳輸至一同步器(例如，圖2中所示之同步層240)。圖7中之編碼器228實例亦包括一速率控制微調模組738，該模組738將位元流填充回饋提供至第二遍編碼器232中之MC/變換量化模組234及重編碼器234中之

ZZ/熵模組 736，以幫助調諧第二遍編碼(例如，以增加壓縮效率)。

### 誤差回彈模組

圖 7 中所說明之編碼器 228 實例亦包括第二遍部分 706 中之誤差回彈模組 740。誤差回彈模組 740 與位元流填充模組 731 及片/MB 排序模組 722 通信。誤差回彈模組 740 自預處理器 228 接收中繼資料，且選擇一誤差回彈方案(例如，將片及存取單元對齊至訊框邊界)、預測階層架構及自適應訊框內再新。誤差回彈方案之選擇可基於在中繼資料中所接收之資訊或自位元流填充模組 731 及片/MB 排序模組 722 傳達至誤差回彈模組的資訊。誤差回彈模組 740 在第一遍部分 702 中將資訊提供至片/巨集區塊(MB)排序模組，以實施所選擇之誤差回彈過程。在易出誤差之環境中的視訊傳輸可使用誤差回彈策略及演算法，其可導致將更清晰及更少誤差填充之資料呈現給觀測使用者。以下之誤差回彈描述可應用於現有或未來應用、傳送及實體層或其他技術之任一個體或組合。有效抗錯強韌性演算法結合通信系統之所要性質(諸如，低潛伏時間及高通量)使對 OSI 層中之誤差敏感性性質及誤差保護能力之理解完整。誤差回彈處理可基於多媒體資料之內容資訊，例如，基於多媒體資料之內容分類。該等主要優點中的一者為自衰落及多路徑通道誤差之可恢復性。以下描述之誤差回彈方法係特定關於可併入編碼器 228 中(例如，詳言之誤差回彈模組 740 及片/MB 排序模組 722 中)且通常可擴展至易出誤差之環境中之資料通信的

過程。

### 誤差回彈

對於基於預測之混合壓縮系統而言，在無任何時間預測之情況下獨立編碼訊框內編碼訊框。可自過去訊框(P訊框)及未來訊框(B訊框)來時間預測訊框間編碼訊框。可經由在參考訊框(一或多個)中之搜尋過程來識別最佳預測子，且使用諸如 SAD 之失真量測來識別最佳匹配。當前訊框之預測編碼區域可為具有變化尺寸及形狀(16x16、32x32、8x4等)之區塊，或經由(例如)分段而被識別為一對象的一組像素。

時間預測通常在許多個訊框(例如，10至幾十個訊框)上擴展，且在一訊框被編碼為I訊框時終止，GOP通常藉由I訊框頻率來界定。對於最大編碼效率而言，GOP為一場景，例如，將GOP邊界與場景邊界對齊，且將場景改變訊框編碼為I訊框。在低移動中，序列包含一相對靜態背景，且移動通常被限制於前景物體。此等低移動序列之內容的實例包括新聞及天氣預報節目，其中超過30%之最多觀測內容屬於此種類。在低移動序列中，大部分地區被訊框間編碼，且已預測之訊框經由中間預測訊框而返回參看I訊框。

參看圖22，I訊框中之訊框內編碼區塊2205為用於已編碼訊框(或AU)P1中之訊框間編碼區塊2210的預測子。在此實例中，此等區塊之區域為背景之固定部分。經由連續時間預測，訊框內編碼區塊2205對誤差之敏感性提高，此係因為其為一"優良"預測子，此亦意味其之"重要性"更高。此外，根據稱為預測鏈之此時間預測鏈，訊框內編碼區塊2205

在顯示中持續較長(持續該圖中之實例中之場景的持續時間)。

將預測階層架構界定為基於此重要性程度或持續性(親本在頂部(訊框內編碼區塊2205)且子在底部)之量測所建立之區塊的樹。注意，在P1中之訊框間編碼區塊2215係在該階層架構之第二級上，等等。葉為終止一預測鏈的區塊。

可不考慮內容類型(諸如，音樂以及運動而並不僅是新聞)而為視訊序列建立預測階層架構，且一般而言該預測階層架構可應用於基於預測之視訊(及資料)壓縮(此應用於在此申請案中所描述的所有發明)。一旦建立預測階層架構，則可更為有效地應用以下所描述之諸如自適應訊框內再新之誤差回彈演算法。此重要性量測可基於一給定區塊自誤差之可恢復性，諸如經由隱藏操作及應用自適應訊框內再新來增強編碼位元流對誤差的回彈。重要性量測之估計可基於將一區塊用作一預測子的次數，亦稱為持續性量度。持續性量度亦用於藉由阻止預測誤差傳播來改良編碼效率。持續性量度亦增加具有較高重要性之區塊的位元分配。

#### 自適應訊框內再新

自適應訊框內再新為一種可基於多媒體資料之內容資訊的誤差回彈技術。在訊框內再新過程中，儘管標準R-D最佳化將指定某些MB應為訊框間編碼MB，但對該等MB進行訊框內編碼。AIR使用移動加權訊框內再新，以將訊框內編碼MB引入P或B訊框中。包含於基礎層中之此等訊框內編碼MB可藉由 $QP_b$ 或 $QP_e$ 而加以編碼。若將 $QP_e$ 用於基礎層，則



在增強層處不應實施改進。若將QP<sub>b</sub>用於基礎層，則改進可為適當的，否則在增強層處，品質之下降將為顯明的。因為在編碼效率方面訊框間編碼比訊框內編碼更為有效，所以將訊框間編碼增強層處之此等改進。藉此，基礎層係數將不用於增強層，且在不引入新操作之情況下，在增強層處改良品質。

在某些態樣中，代替或除了移動加權基礎外，自適應訊框內再新可基於多媒體資料之內容資訊(例如，內容分類)。舉例而言，若內容分類相對較高(例如，具有高空間及時間複雜性之場景)，則自適應訊框內再新可將相對較多之訊框內編碼MB引入P或B訊框中。或者，若內容分類相對較低(其指示具有低空間及/或時間複雜性之弱動態場景)，則自適應訊框內再新可將較少之訊框內編碼MB引入P及B訊框中。用於改良誤差回彈之此等量度及方法可不僅應用於無線多媒體通信之情形中，且一般而言可朝資料壓縮及多媒體處理而加以應用(例如，在圖形再現中)。

### 通道切換訊框

如本文中所界定之通道切換訊框(CSF)為一廣義術語，其描述一被插入於廣播流中之適當位置處的隨機存取訊框，其用於快速通道擷取且因此用於廣播多工器中之流間的快速通道改變。通道切換訊框亦增加抗錯強韌性，因為該等通道切換訊框提供在具有誤差之情況下傳輸主訊框時可使用之冗餘資料。I訊框或漸進I訊框(諸如H.264中之漸進解碼器再新訊框)通常充當一隨機存取點。然而，頻繁之I訊框(或

比場景持續時間短之短GOP)導致壓縮效率顯著降低。因為訊框內編碼區塊可用於誤差回彈，所以可經由預測階層架構來有效地組合隨機存取及誤差回彈以改良編碼效率，同時增加對誤差之強韌性。

隨機存取切換及抗錯強韌性之改良可一致達成，且其可基於諸如內容分類之內容資訊。對於低移動序列而言，預測鏈較長，且用以重建一超訊框或場景所需要之資訊的一顯著部分被包含於在場景開始時出現的I訊框中。通道誤差傾向於具有叢發性，且當觸發一衰落且FEC及通道編碼失效時，存在隱藏失效之大量的剩餘誤差。此對於低移動(且因此低位元速率)序列而言特別嚴重，此係因為編碼資料之量並不顯著足以提供視訊位元流內之優良的時間分集，且因為此等序列為再現適用於重建之每一位元的高度可壓縮序列。高移動序列對誤差更為強韌，此係由於內容之性質(每一訊框中之更多新資訊)增加了訊框內編碼區塊的數目，該等訊框內編碼區塊為獨立可解碼的且固有地對誤差的回彈性更大。對於高移動序列而言，基於預測階層架構之自適應訊框內再新達成一高效能，且對於低移動序列而言，效能改良並不顯著。因此，包含大部分I訊框之通道切換訊框為低移動序列之分集的優良來源。當誤差觸發一超訊框時，連續訊框中之解碼由於達成了預測及誤差回彈而自恢復損失資訊之CSF開始。

在諸如具有相對較高之內容分類(例如，6-8)之序列的高移動序列之情形中，CSF可由堅持SF之區塊(彼等為優良預

測子)組成。不必編碼CSF之所有其他區域，此係因為此等區域為具有較短預測鏈之區塊，此意味其以訊框內區塊而終止。因此，CSF仍用於當一誤差出現時由於預測而自損失資訊來恢復。低移動序列之CSF與I訊框之尺寸相同，且可經由較大量之量化以一較低位元速率來編碼該等CSF，其中高移動序列之CSF比對應之I訊框小得多。

基於預測階層架構之誤差回彈可良好作用於可調性，且可達成一高度有效之分層編碼。在實體層技術中支援階層式調變的可調性可能需要具有特定頻寬比之視訊位元流的資料分割。此等特定頻寬比可不始終為針對最佳可調性之理想比(例如，具有最少額外成本)。在某些態樣中，使用具有1:1頻寬比之2層可調性。將視訊位元流分割成相等尺寸之2個層可能不如對於低移動序列般有效。對於低移動序列而言，包含所有標頭及中繼資料資訊之基礎層大於增強層。然而，因為低移動序列之CSF較大，所以其恰好適於增強層中之剩餘頻寬中。

高移動序列具有充足的剩餘資訊，使得可以最少額外成本達成1:1之資料分割。此外，此等序列之通道切換訊框比高移動序列之通道切換訊框小得多。因此，基於預測階層架構之誤差回彈同樣可良好作用於高移動序列之可調性。基於此等演算法之描述，針對中等移動片段而擴展以上所論述之概念係可行的，且一般而言該等提議之概念適用於視訊編碼。

多工器

在某些編碼器態樣中，多工器可用於編碼藉由編碼器所產生之多個多媒體流，且用於預備用以廣播之編碼位元。舉例而言，在圖2中所示之編碼器228的說明性態樣中，同步層240包含一多工器。多工器可經實施以提供位元速率分配控制。可將已估計之複雜性提供至多工器，該多工器可繼而根據為已多工之視訊通道所預期之編碼複雜性來分配彼等視訊通道之集合的可用頻寬，此繼而允許一特定通道之品質保持相對恆定，儘管用於已多工之視訊流之集合的頻寬相對恆定。此在通道之集合內提供一通道以具有可變位元速率及相對恆定之視覺品質，而非相對恆定之位元速率及可變視覺品質。

圖18為說明編碼多個多媒體流或通道1802之系統的方塊圖。藉由個別編碼器1804來編碼多媒體流1802，該等個別編碼器1804與一多工器(MUX)1806通信，該多工器(MUX)1806又與傳輸媒體1808通信。舉例而言，多媒體流1802可對應於各種內容通道，諸如新聞通道、運動通道、電影通道及其類似物。編碼器1804將多媒體流1802編碼成該系統所特定之編碼格式。儘管在編碼視訊流之內容中進行了描述，但所揭示之技術的原理及優點通常可應用於包括(例如)音訊流之多媒體流。將已編碼之多媒體流提供至多工器1806，該多工器1806組合各種已編碼之多媒體流，且將組合流發送至傳輸媒體1808以用於傳輸。

傳輸媒體1808可對應於各種媒體，諸如(但不限於)數位衛星通信，諸如DirecTV®、數位電纜、有線及無線網際網路

通信、光學網路、蜂巢式電話網路及其類似物。傳輸媒體 1808 可包括(例如)對射頻(RF)之調變。通常，由於頻譜約束及其類似物，傳輸媒體具有一有限頻寬，且自多工器 1806 至傳輸媒體之資料被保持於一相對恆定位元速率(CBR)。

在習知系統中，在多工器 1806 之輸出處使用恆定位元速率(CBR)可能要求輸入至多工器 1806 的已編碼之多媒體或視訊流亦為 CBR。如在先前技術中所描述，在編碼視訊內容時使用 CBR 可導致可變視覺品質，其通常為非吾人所需要的。

在所說明之系統中，兩個或兩個以上之編碼器 1804 中傳達輸入資料之一預期編碼複雜性。一或多個編碼器 1804 可回應地自多工器 1806 接收適應之位元速率控制。此允許期望編碼相對複雜之視訊的編碼器 1804 以准可變位元速率之方式接收彼等視訊訊框之一較高位元速率或較高頻寬(每一訊框更多位元)。此允許多媒體流 1802 以一致之視覺品質而被編碼。由一編碼相對複雜之視訊的特定編碼器 1804 所使用之額外頻寬來自在編碼器經實施以在恆定位元速率下操作的情況下另外已用於編碼其他視訊流 1804 的位元。此將多工器 1806 之輸出保持於恆定位元速率(CBR)。

儘管個別多媒體流 1802 可具有相對"叢發性"，亦即，所使用之頻寬發生變化，但多個視訊流之累積和可具有較少叢發性。可將來自編碼較低複雜之視訊之通道的位元速率藉由(例如)多工器 1806 重分配至編碼相對複雜之視訊的通道，且此可總體上增強組合之視訊流的視覺品質。

編碼器 1804 向多工器 1806 提供一組待編碼及多路傳輸在一起之視訊訊框之複雜性的指示。多工器 1806 之輸出應提供一不高於為傳輸媒體 1808 所指定之位元速率的輸出。複雜性之指示可基於如以上所論述之內容分類，以提供所選擇之品質水平。多工器 1806 分析複雜性之指示，且向各種編碼器 1804 提供所分配之數目的位元或頻寬，且編碼器 1804 使用此資訊來編碼該組中之視訊訊框。此允許一組視訊訊框個別為可變位元速率，且總體而言仍達成恆定位元速率。

一般而言，對於任一通用壓縮器而言，內容分類亦可用於致能多媒體之基於品質的壓縮。內容分類及在此所描述之方法及設備可用於任何多媒體資料之基於品質及/或基於內容的多媒體處理。一般而言，對於任一通用壓縮器而言，一實例為內容分類在多媒體之壓縮中的使用。另一實例在任一減壓器或解碼器或後處理器中之減壓或解碼中，諸如內插、重取樣、增強、恢復及顯示操作。

現參看圖 19，一典型視訊通信系統包括一視訊壓縮系統，其由藉由通信網路所連接之一視訊編碼器及一視訊解碼器組成。無線網路為一類易出誤差之網路，其中通信通道除了路徑損耗外顯示行動情境中之對數正態衰落或陰影及多路徑衰落。為了對抗通道誤差且提供應用層資料之可靠通信，RF 調變器包括正向誤差校正，其包括交錯器及諸如卷積或渦輪編碼之通道編碼。

視訊壓縮降低源視訊中之冗餘，且增加已編碼之視訊資

料之每一位元中所載運的資訊量。此增加了品質影響，甚至在損失了小部分編碼視訊時亦如此。視訊壓縮系統中所固有的空間及時間預測加重了該損耗，且導致誤差傳播，其引起重建視訊中之可見假影。在視訊編碼器處之誤差回彈演算法及在視訊解碼器處之誤差恢復演算法增強了視訊壓縮系統之抗錯強韌性。

通常，視訊壓縮系統對於基本網路而言為不可知的。然而，在易出錯誤之網路中，應用層中之防錯演算法極需與鏈接層/實體層中之FEC及通道編碼相整合，且其提供增強系統之誤差效能的最佳效率。

圖14說明了用以編碼訊框之可出現於編碼器228中之速率失真資料流的一實例。過程1400始於開始1402處，且進行至決策區塊1404，在區塊1404中其自預處理器226接收場景改變偵測器輸入1410(例如，經由中繼資料)，且獲取誤差回彈輸入1406。若該資訊指示所選擇之訊框為一I訊框，則該過程訊框內編碼該訊框。若該資訊指示所選擇之訊框為一P或B訊框，則過程使用訊框內編碼及移動估計(訊框間)編碼來編碼該訊框。

在對於區塊1404之條件產生一肯定條件後，過程1400進行至一預備區塊1414，在區塊1414中速率R被設定至值 $R=R_{qual}$ ，所要之目標品質係基於R-D曲線。自一包含R-D曲線之資料塊1416接收此設定。過程1400繼而進行至區塊1418，在區塊1418中基於來自區塊1420處之內容分類過程的影像/視訊活動資訊(例如，內容分類)來執行速率控制位

元分配  $\{Q_{pi}\}$ 。

速率控制位元分配區塊 1418 又用於區塊 1422 中之移動估計。移動估計 1422 亦可接收來自預處理器 1412 之中繼資料、來自區塊 1424 之移動向量平滑 (MPEG-2+歷史) 及來自區塊 1426 之多個參考訊框 (因果+非因果巨集區塊 MB) 的輸入。過程 1400 繼而進行至區塊 1428，在區塊 1428 中對於速率控制位元分配 ( $Q_{pi}$ ) 判定訊框內編碼模式之速率計算。過程 1400 接著進行至區塊 1430，在區塊 1430 中判定模式及量化參數。基於區塊 1422 輸入、誤差回彈 1406 輸入及可調性 R-D (在區塊 1432 處加以判定) 之移動估計來作出區塊 1430 之模式決策。一旦決定模式，則流程進行至區塊 1432。應注意，當資料自編碼器之第一遍部分傳遞至第二遍部分時，發生自區塊 1430 至 1432 之流程。

在區塊 1432 處，藉由編碼器 228 之第二遍來執行變換及量化。如藉由區塊 1444 所指示來調整或微調變換/量化過程。可藉由一速率控制微調模組 (圖 7) 來影響此變換/量化過程。過程 1400 繼而進行至區塊 1434 以用於曲折排序及熵編碼從而產生已編碼之基礎層。曲折排序以用於編碼之有效格式來預備量化資料。熵編碼為使用一系列位元編碼來表示一組可能之符號的壓縮技術。如先前參看圖 31-36 所描述，亦將變換/量化區塊 1432 之增強層結果發送至加法器 1436，該加法器 1436 減去基礎層且將該結果發送至用於增強層之 ZZ/熵編碼器 1438。進一步注意，增強層被反饋 (見線 1440 之實際速率更新) 以用於更新實際速率之內容分類



1420及一用於判定速率控制所使用之位元速率的長期及短期歷史之操作。

圖 17A 為一過程之流程圖，該過程用於編碼多媒體資料，以便將一資料邊界與時域中之一訊框邊界對齊，且其中該編碼係基於內容資訊。可藉由圖 17B 中所說明之設備及本文所揭示之其他組件(包括轉換編碼器 200)來執行過程 1700。過程 1700 開始，且在區塊 1702 處其獲得多媒體資料之內容資訊。此可藉由內容分類器 1712(圖 17B)或另一組件(諸如，圖 7 中之內容分類器 712)來執行。過程 1700 繼而進行至區塊 1704，在區塊 1704 中其編碼多媒體資料，以將一資料邊界與時域中之一訊框邊界對齊，其中編碼係基於在區塊 1702 處所獲得之內容資訊。此可藉由編碼器 1714(圖 17B)或在另一實例中藉由轉換編碼器 200(圖 7)來執行。將資料邊界與時間邊界對齊可導致誤差回彈。

圖 17B 為一可執行圖 17A 及圖 17C 中所說明之過程的編碼設備 1710 之高位準方塊圖。設備 1710 可包括內容分類構件，諸如用於獲得多媒體資料之內容資訊的模組 1712。內容分類構件之實例亦參看圖 7 而加以說明及描述，且其可包括(例如)一內容分類模組、一編碼器、一預處理器或一轉換編碼器。設備 1710 亦包括用於編碼多媒體資料之構件，諸如用於編碼多媒體資料以便將一資料邊界與時域中之一訊框邊界對齊的模組 1714。此構件之實例包括一編碼器(例如，編碼器 228)或一轉換編碼器(例如，轉換編碼器 200)。

圖 17C 為用於編碼多媒體資料之過程 1750 的流程圖。過程

1750開始，且在區塊1752處獲得多媒體資料之內容分類。此可藉由內容分類器1712(圖17B)或另一組件(諸如，圖7中之內容分類器712)來執行。過程1750進行至區塊1754，在區塊1754中基於內容分類將多媒體資料編碼為訊框內編碼區塊或訊框間編碼區塊。此可藉由編碼器1714(圖17B)或在另一實例中藉由轉換編碼器200(圖7)來執行。過程1750繼而結束，直至需要進一步多媒體資料處理為止。

圖23、24、27及28為舉例說明體現本文所描述之態樣之編碼多媒體資料的方法的過程流程圖。圖23為說明基於內容資訊來編碼多媒體資料之過程2300的過程流程圖。在區塊2305處，過程2300接收已編碼之多媒體資料，且在區塊2310處，過程2300解碼多媒體資料。在區塊2315處，過程2300判定與已解碼之多媒體資料相關聯的內容資訊。在區塊2320處，過程2300基於內容資訊來編碼多媒體資料。

圖24為說明編碼多媒體資料以便基於內容資訊位準來對齊資料邊界之過程2400的過程流程圖。在區塊2405處，過程2400獲得與多媒體資料相關聯之內容資訊，此可(例如)藉由預處理器226或圖7中所示之內容分類模組712來實施。在區塊2410處，過程2400編碼多媒體資料以便基於內容資訊來對齊資料邊界。舉例而言，基於正編碼之多媒體資料的內容分類，將片邊界及存取單元邊界與訊框邊界對齊。編碼資料繼而可用於後續處理及/或傳輸至行動裝置，且過程2400結束。

圖27為說明使用基於內容資訊之自適應訊框內再新方案

來編碼資料之過程2700的過程流程圖。當過程2700開始時，已獲得多媒體資料。在區塊2705處，過程2700獲得多媒體資料之內容資訊。獲得內容資訊可藉由(例如)如以上所描述之預處理器226或內容分類模組712來執行。過程2700進行至區塊2710，在區塊2710中其使用自適應訊框內再新誤差回彈方案來編碼多媒體資料，其中自適應訊框內再新誤差回彈方案係基於內容資訊。可藉由編碼器228來執行區塊2710之功能。使編碼資料可用於後續處理及傳輸，且過程2700繼而結束。

圖28為說明使用基於多媒體內容資訊之冗餘I訊框來編碼多媒體資料之過程的過程流程圖。當過程2800開始時，多媒體資料可用於處理。在區塊2805處，過程2800獲得多媒體資料之內容資訊。如以上所描述，此可藉由(例如)預處理器226及/或編碼器228來實施。在區塊2810處，過程2800基於內容資訊編碼多媒體資料，以便將一或多個額外I訊框插入編碼資料中。此可藉由如以上所描述之編碼器228結合一誤差回彈方案來實施，此視所使用之誤差回彈方案而定而將I訊框插入基礎層或增強層中。在區塊2810之後，編碼資料可用於後續處理及/或傳輸至行動裝置。

應注意本文所描述之方法可實施於一般熟習此項技術者所知的各種通信硬體、處理器及系統上。舉例而言，對於用戶端如本文所描述般來操作的一般要求為用戶端具有一用以顯示內容及資訊之顯示器、一用以控制用戶端之操作的處理器及一用於儲存關於用戶端操作之資料及程式的記

憶體。在一個態樣中，用戶端為一蜂巢式電話。在另一態樣中，用戶端為具有通信能力之掌上型電腦。在又一態樣中，用戶端為具有通信能力之個人電腦。此外，可將諸如GPS接收器之硬體併入用戶端以實施各種態樣。可藉由經設計以執行本文所描述之功能的通用處理器、數位信號處理器(DSP)、特殊應用積體電路(ASIC)、場可程式化閘陣列(FPGA)或其他可程式化邏輯設備、離散閘或電晶體邏輯、離散硬體組件或其任何組合來實施或執行結合本文所揭示之態樣而描述之各種說明性邏輯、邏輯區塊、模組及電路。通用處理器可為一微處理器，但在替代例中，該處理器可為任一習知處理器、控制器、微控制器或狀態機。亦可將一處理器實施為計算裝置的一組合，例如，一DSP與一微處理器之組合、複數個微處理器之組合、結合一DSP核心之一或多個微處理器之組合或任一其他此組態之組合。

可藉由經設計以執行本文所描述之功能的通用處理器、數位信號處理器(DSP)、特殊應用積體電路(ASIC)、場可程式化閘陣列(FPGA)或其他可程式化邏輯裝置、離散閘或電晶體邏輯、離散硬體組件或其任何組合來實施或執行結合本文所揭示之態樣而描述之各種說明性邏輯、邏輯區塊、模組及電路。通用處理器可為一微處理器，但在替代例中，該處理器可為任一習知處理器、控制器、微控制器或狀態機。亦可將一處理器實施為計算裝置的一組合，例如，一DSP與一微處理器之組合、複數個微處理器之組合、結合一DSP核心之一或多個微處理器之組合，或任一其他此組

態之組合。

所揭示之方法及設備提供將以一格式編碼之視訊資料至以另一格式編碼之視訊資料的轉碼，其中該編碼係基於視訊資料之內容，且該編碼對誤差具有回彈性。結合本文所揭示之實例加以描述的方法或演算法可直接體現於硬體、藉由一處理器所執行之軟體模組、韌體或此等中之兩個或兩個以上的組合中。軟體模組可常駐於RAM記憶體、快閃記憶體、ROM記憶體、EPROM記憶體、EEPROM記憶體、暫存器、硬碟、抽取式碟片、CD-ROM或此項技術中已知之任一其他形式的儲存媒體中。將一例示性儲存媒體耦接至處理器，使得該處理器可自該儲存媒體讀取資訊及將資訊寫入該儲存媒體。在替代例中，該儲存媒體可整合於該處理器。該處理器與該儲存媒體可常駐於一ASIC中。該ASIC可常駐於一使用者終端機中。在替代例中，該處理器及該儲存媒體可作為離散組件常駐於使用者終端機中。

以上所描述之實例僅為例示性的，且熟習此項技術者現可在不脫離本文所揭示之發明概念的情況下實施以上描述之實例的多種使用及變更。熟習此項技術者可易於明白對此等實例之各種修改，且本文所界定之一般原理可在不脫離本文所描述之新穎態樣之精神或範疇的情況下應用於其他實例，例如，在瞬時訊息傳遞服務或任何一般無線資料通信應用中。因而，本揭示內容之範疇並不意欲受限於本文所示之實例，而是應符合於與本文所揭示之原理及新穎特徵一致的最廣泛之範疇。本文專門使用詞語"例示性"來

意謂"用作一實例、例子或說明"。本文描述為"例示性"之任一實例並不一定解釋為比其他實例較佳或有利。因此，將僅藉由下列申請專利範圍之範疇來界定本文所描述之新穎態樣。

### 【圖式簡單說明】

圖1A為一媒體廣播系統之方塊圖，該媒體廣播系統包括一用於在不同視訊格式之間進行轉碼的轉換編碼器。

圖1B為一編碼器之方塊圖，該編碼器經組態以編碼多媒體資料且提供一已編碼之第一資料組及一已編碼之第二資料組。

圖1C為一經組態以編碼多媒體資料之處理器的方塊圖。

圖2為圖1之系統之轉換編碼器之一實例的方塊圖。

圖3為說明用於圖2之轉換編碼器內的剖析器之操作的流程圖。

圖4為說明用於圖2之轉換編碼器內的解碼器之操作的流程圖。

圖5為說明由圖2之轉換編碼器執行的一序列操作之系統時序圖。

圖6為說明一可用於圖2之轉換編碼器中之預處理器的一序列操作及功能之流程圖。

圖7為可用於圖2之轉換編碼器中的例示性2遍式編碼器之方塊圖。

圖8說明了一分類圖之實例，該分類圖說明如何將紋理及移動值與內容分類相關聯的一個態樣。

圖 9 為說明一用於內容分類之例示性操作(諸如用於圖 7 之編碼器中)的流程圖。

圖 10 為說明一速率控制之操作(諸如供圖 7 之編碼器使用)的流程圖。

圖 11 為說明一例示性移動估計器之操作(諸如供圖 7 之編碼器使用)的流程圖。

圖 12 為說明一例示性模式決策編碼器功能之操作(諸如供圖 7 之編碼器使用)的流程圖。

圖 13 為說明一用於圖 7 之編碼器中的實現可調性之例示性操作的流程圖。

圖 14 為說明實現速率失真資料流(例如，如在圖 7 之編碼器中所發生)之例示性操作的流程圖。

圖 15 為說明編碼複雜性、分配之位元及人類視覺品質間之關係的圖表。

圖 16 為說明一非線性場景偵測公式之圖表。

圖 17A 為一過程之流程圖，該過程用於編碼多媒體資料以便將資料邊界與時域中之訊框邊界對齊，且其中該編碼係基於內容資訊。

圖 17B 為一可執行圖 17A 及圖 17C 中所說明之過程之編碼設備的高級方塊圖。

圖 17C 為一用於編碼多媒體資料之過程的流程圖，其中基於內容分類該多媒體資料被編碼為訊框內編碼區塊或訊框間編碼區塊。

圖 18 為說明一使用移動估計/補償之解交錯過程的圖表。

圖 19 為多媒體通信系統之方塊圖。

圖 20 為說明視訊位元流在增強層及基礎層中之組織的圖表。

圖 21 為說明片至視訊訊框邊界之對齊的圖表。

圖 22 為說明預測階層架構之方塊圖。

圖 23 為說明一種基於內容資訊來編碼多媒體資料之方法的過程流程圖。

圖 24 為說明一種編碼多媒體資料以便基於內容資訊位準來對齊資料邊界之方法的過程流程圖。

圖 25 為說明一資料訊框之安全動作區及安全標題區之圖表。

圖 26 為說明一資料訊框之安全動作區的圖表。

圖 27 為說明一使用基於多媒體內容資訊之自適應訊框內更新來編碼多媒體資料之過程的過程流程圖。

圖 28 為說明一使用基於多媒體內容資訊之冗餘 I 訊框來編碼多媒體資料之過程的過程流程圖。

圖 29 說明了在一當前訊框與一先前訊框  $MV_p$  之間及在一當前訊框與一下一訊框  $MV_N$  之間的移動補償向量。

圖 30 為說明鏡頭偵測之過程流程圖。

圖 31 為說明編碼基礎層及增強層之過程流程圖。

圖 32 為說明編碼一巨集區塊之示意圖。

圖 33 為說明用於編碼基礎層及增強層之模組的示意圖。

圖 34 展示了基礎層及增強層係數選擇器過程之實例。

圖 35 展示了基礎層及增強層係數選擇器過程之另一實



例。

圖 36 展示了基礎層及增強層係數選擇器過程之另一實例。

圖 37 為說明基於內容資訊來編碼多媒體資料之過程流程圖。

圖 38 為說明在一反向電視電影處理過程中之可能系統決策的圖表。

圖 39 說明了在一待由一解塊過程過濾之巨集區塊中的邊界。

圖 40 為說明一時空解交錯過程之圖表。

圖 41 說明了 1 維多相重取樣之實例。

圖 42 為說明在視訊串流中之自適應 GOP 結構之實例的流程圖。

應注意，在適當處，相同數字貫穿該等圖式之若干視圖而指代相同部分。

### 【主要元件符號說明】

100	多媒體資料廣播系統
104	多媒體資料
106	多媒體資料提供者
120	外部記憶體
122	內部記憶體
124a...n	模組
126	通信模組
130	轉換編碼器

140	處理器
200	轉換編碼器
202	剖析器/解碼器
205	傳輸流剖析器/剖析器
206	視訊基本流
208	音訊基本流
210	顯示時間戳
212	副標題
214	解碼器
215	顯示時間戳產生器
216	音訊解碼器
218	隱藏式字幕資料
220	測試編碼器
222	中繼資料
224	"原始"視訊資料
226	預處理器
228	編碼器
230	本文編碼器
231	轉換編碼器控制
232	第二遍編碼器
234	重編碼器
236	音訊編碼器
238	編碼資訊
240	同步層

242	分封化基本流
244	資料分封化基本流
246	音訊分封化基本流
308	緩衝描述符
333	當前圖片
335	巨集區塊
337	參考圖片
339	巨集區塊
349	網路
357	巨集區塊
359	巨集區塊
361	原始資料
363	基礎層迴路
365	增強層迴路
367	加法器
369	加法器
371	緩衝器
373	緩衝器
375	變換區塊
377	變換區塊
379	選擇器
381	量化器
383	量化器
385	解量化器

391	解量化器
393	反向變換器
397	加法器
406	視訊 PES 緩衝器 '讀取' 資訊
407	加法器
604	相位偵測器
605	解交錯器
607	去雜訊器/過濾器
612	GOP 分割器
614	輸出資料
702	第一遍部分
704	分界線
706	第二遍部分
708	I 訊框實例化模組
710	AIR 模組
712	內容分類模組
714	速率控制位元分配模組
715	模式決策模組
716	訊框內模型(失真)模組
718	基礎層及增強層模組
720	移動估計(失真)模組
722	片/MB 排序模組
726	MC/變換量化模組
728	曲折(ZZ)/熵編碼器

730	可調性模組
731	位元流填充模組
736	ZZ/熵模組
738	速率控制微調模組
740	誤差回彈模組
1802	多媒體流或通道
1804	編碼器
1806	多工器
1808	傳輸媒體
1710	編碼設備
1712	內容分類器模組
1714	編碼器模組
2005	訊框
2010	訊框
2015	訊框
2020	片
2025	片
2030	片
2105	訊框
2110	訊框
2205	訊框內編碼區塊
2210	訊框間編碼區塊
2215	訊框間編碼區塊
2505	安全標題區

2510	安全動作區
2615	白色矩形窗口
2620	上面
2625	下面
4002	移動強度映射
4004	Wmed過濾器
4006	去雜訊器(去雜訊過濾器)
4008	下一Wmed訊框
4010	當前Wmed訊框
4012	組合器
4014	解交錯當前訊框
4016	移動補償當前訊框
4018	ME/MC
4020	解交錯先前訊框
4022	MV候選選擇器

## 五、中文發明摘要：

本發明描述一種使用內容資訊來編碼多媒體資料之設備及方法。一種處理多媒體資料之方法包括：獲得多媒體資料之內容資訊；及編碼多媒體資料，以便將資料邊界與一時域中之訊框邊界對齊，其中該編碼係基於內容資訊。在另一態樣中，一種處理多媒體資料之方法包括：獲得多媒體資料之內容分類；及基於該內容分類將多媒體資料中之區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加已編碼之多媒體資料之誤差回彈。本發明亦揭示一種可處理此等方法中所描述之多媒體資料的設備。

## 六、英文發明摘要：

七、指定代表圖：

(一)本案指定代表圖為：第 ( 17B ) 圖。

(二)本代表圖之元件符號簡單說明：

1710 編碼設備

1712 內容分類器/模組

1714 編碼器/模組

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)



## 十、申請專利範圍：

1. 一種處理多媒體資料之方法，該方法包含：

獲得該多媒體資料之內容資訊，其中該內容資訊包含用於該多媒體資料之一或多個區塊之每一者之一內容分類，且其中該或該等區塊之每一者的該內容分類係至少部分地基於下述一或多者而被判定：

一與該區塊之一空間複雜性相關聯之紋理值，及

一與該區塊之一時間複雜性相關聯之移動值；及

編碼該多媒體資料以定義一片，該片包含獨立可解碼的視訊資料，以便將一與該片相關聯之資料邊界與一時域中之一訊框邊界對齊，其中該時域中之該訊框邊界包含一與若干實體層封包相關聯之邊界，其中該編碼係基於該或該等內容分類。

2. 如請求項1之方法，其中獲得該內容資訊包含計算來自該多媒體資料之該內容資訊。
3. 如請求項1之方法，其中該資料邊界進一步包含一I訊框資料邊界。
4. 如請求項1之方法，其中該資料邊界進一步包含一訊框內編碼之存取單元邊界。
5. 如請求項1之方法，其中該資料邊界進一步包含一P訊框邊界。
6. 如請求項1之方法，其中該資料邊界進一步包含一B訊框邊界。
7. 一種處理多媒體資料之方法，該方法包含：

獲得用於該多媒體資料之一或多個區塊之一內容分類，其中用於該或該等區塊之該內容分類係至少部分地基於下述一或多者而被判定：

一與該或該等區塊之一空間複雜性相關聯之紋理值，及

一與該或該等區塊之一時間複雜性相關聯之移動值；及

基於該內容分類將該多媒體資料之該或該等區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加該已編碼之多媒體資料之一誤差回彈。

8. 如請求項7之方法，其中該編碼包含增加該或該等區塊之一數目，該或該等區塊係被編碼為對應於該內容分類中之一減少的訊框內編碼區塊。

9. 一種用於處理多媒體資料之設備，該設備包含：

一內容分類器，其經組態以獲得該多媒體資料之內容資訊，其中該內容資訊包含用於該多媒體資料之一或多個區塊之每一者之一內容分類，且其中該或該等區塊之每一者的該內容分類係至少部分地基於下述一或多者而被判定：

一與該區塊之一空間複雜性相關聯之紋理值，及

一與該區塊之一時間複雜性相關聯之移動值；及

一編碼器，其經組態以編碼該多媒體資料以定義一片，該片包含獨立可解碼的視訊資料，以便將一與該片相關聯之資料邊界與一時域中之一訊框邊界對齊，

其中該時域中之該訊框邊界包含一與若干實體層封包相關聯之邊界，其中該編碼係基於該或該等內容分類。

10. 如請求項9之設備，其中該內容分類器經組態以藉由計算來自該多媒體資料之該內容資訊而獲得該內容資訊。
11. 如請求項9之設備，其中該資料邊界進一步包含一I訊框資料邊界。
12. 如請求項9之設備，其中該資料邊界進一步包含一訊框內編碼之存取單元邊界。
13. 如請求項9之設備，其中該資料邊界進一步包含一P訊框邊界。
14. 如請求項9之設備，其中該資料邊界進一步包含一B訊框邊界。
15. 一種用於處理多媒體資料之設備，該設備包含：
  - 一內容分類器，其經組態以獲得用於該多媒體資料之一或多個區塊之一內容分類，其中用於該或該等區塊之該內容分類係至少部分地基於下述一或多者而被判定：
    - 一與該或該等區塊之一空間複雜性相關聯之紋理值，及
    - 一與該或該等區塊之一時間複雜性相關聯之移動值；及
  - 一編碼器，其經組態以基於該內容分類將該多媒體資料之該或該等區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加該已編碼之多媒體資料之一誤差回彈。
16. 如請求項15之設備，其中為了編碼該或該等區塊，該編

碼器經組態以增加該或該等區塊之一數目，該或該等區塊係被編碼為對應於該內容分類中之一減少的訊框內編碼區塊。

17. 一種用於處理多媒體資料之設備，該設備包含：

用於獲得該多媒體資料之內容資訊的構件，其中該內容資訊包含用於該多媒體資料之一或多個區塊之每一者之一內容分類，且其中用於該或該等區塊之每一者的該內容分類係至少部分地基於下述一或多者而被判定：

- 一與該區塊之一空間複雜性相關聯之紋理值，及
- 一與該區塊之一時間複雜性相關聯之移動值；及

用於編碼該多媒體資料以定義一片，該片包含獨立可解碼的視訊資料，以便將一與該片相關聯之資料邊界與一時域中之一訊框邊界對齊的構件，其中該時域中之該訊框邊界包含一與若干實體層封包相關聯之邊界，其中該編碼係基於該或該等內容分類。

18. 一種用於處理多媒體資料之設備，該設備包含：

用於獲得用於該多媒體資料之一或多個區塊之一內容分類的構件，其中用於該或該等區塊之該內容分類係至少部分地基於下述一或多者而被判定：

- 一與該或該等區塊之一空間複雜性相關聯之紋理值，及

- 一與該或該等區塊之一時間複雜性相關聯之移動值；及

用於基於該內容分類將該多媒體資料之該或該等區塊

編碼為訊框內編碼區塊或訊框間編碼區塊，以增加該已編碼之多媒體資料之一誤差回彈的構件。

19. 如請求項18之設備，其中該用於編碼之構件包含用於增加該或該等區塊之一數目之構件，該或該等區塊係被編碼為對應於該內容分類中之一減少的訊框內編碼區塊。

20. 一種處理器，其經組態以：

獲得多媒體資料之內容資訊，其中該內容資訊包含用於該多媒體資料之一或多個區塊之每一者之一內容分類，且其中該或該等區塊之每一者的該內容分類係至少部分地基於下述一或多者而被判定：

一與該區塊之一空間複雜性相關聯之紋理值，及

一與該區塊之一時間複雜性相關聯之移動值；及

編碼該多媒體資料以定義一片，該片包含獨立可解碼的視訊資料，以便將一與該片相關聯之資料邊界與一時域中之一訊框邊界對齊，其中該時域中之該訊框邊界包含一與若干實體層封包相關聯之邊界，其中該編碼係基於該或該等內容分類。

21. 如請求項20之處理器，其中該處理器經組態以藉由計算來自該多媒體資料之該內容資訊而獲得該內容資訊。

22. 如請求項20之處理器，其中該資料邊界進一步包含一I訊框資料邊界。

23. 如請求項20之處理器，其中該資料邊界進一步包含一訊框內編碼之存取單元邊界。

24. 如請求項20之處理器，其中該資料邊界進一步包含一P訊

框邊界。

25. 如請求項20之處理器，其中該資料邊界進一步包含一B訊框邊界。

26. 一種處理器，其經組態以：

獲得用於多媒體資料一或多個區塊之一內容分類，其中用於該或該等區塊的該內容分類係至少部分地基於下述一或多者而被判定：

一與該或該等區塊之一空間複雜性相關聯之紋理值，及

一與該或該等區塊之一時間複雜性相關聯之移動值；及

基於該內容分類將該多媒體資料之該或該等區塊編碼為訊框內編碼區塊或訊框間編碼區塊，以增加該已編碼之多媒體資料的一誤差回彈。

27. 如請求項26之處理器，其中為了編碼該或該等區塊，該處理器經組態增加該或該等區塊之一數目，該或該等區塊係被編碼為對應於該內容分類中之一減少的訊框內編碼區塊。

28. 一種包含指令之機器可讀取媒體，該等指令一經執行便導致一機器：

獲得多媒體資料之內容資訊，其中該內容資訊包含用於該多媒體資料之一或多個區塊之每一者之一內容分類，且其中用於該或該等區塊之每一者的該內容分類係至少部分地基於下述一或多者而被判定：

一與該區塊之一空間複雜性相關聯之紋理值，及

一與該區塊之一時間複雜性相關聯之移動值；及

編碼該多媒體資料以定義一片，該片包含獨立可解碼的視訊資料，以便將一與該片相關聯之資料邊界與一時域中之一訊框邊界對齊，其中該時域中之該訊框邊界包含一與若干實體層封包相關聯之邊界，其中該編碼係基於該或該等內容分類。

29. 如請求項28之機器可讀取媒體，其進一步包含若干指令，該等指令在被執行時導致該機器藉由計算來自該多媒體資料之該內容資訊而獲得該內容資訊。

30. 如請求項28之機器可讀取媒體，其中該資料邊界進一步包含一I訊框資料邊界。

31. 如請求項28之機器可讀取媒體，其中該資料邊界進一步包含一訊框內編碼之存取單元邊界。

32. 一種包含若干指令之機器可讀取媒體，該等指令一經執行便導致一機器：

獲得用於多媒體資料之一或多個區塊之一內容分類，其中用於該或該等區塊之該內容分類係至少部分地基於下述一或多者而被判定：

一與該或該等區塊之一空間複雜性相關聯之紋理值，及

一與該或該等區塊之一時間複雜性相關聯之移動值；及

基於該內容分類將該多媒體資料之該或該等區塊編碼

為訊框內編碼區塊或訊框間編碼區塊，以增加該已編碼之多媒體資料之一誤差回彈。

33. 如請求項32之機器可讀取媒體，其中用以編碼該或該等區塊之該等指令包含用以增加該或該等區塊之一數目之指令，該或該等區塊係被編碼為對應於該內容分類中之一減少的訊框內編碼區塊。



I3 6952 修正 (pp. 1-42)  
10年1月5日 補充

公告本

第 095135898 號專利申請案  
中文圖式替換本(101 年 9 月)

十一、圖式：

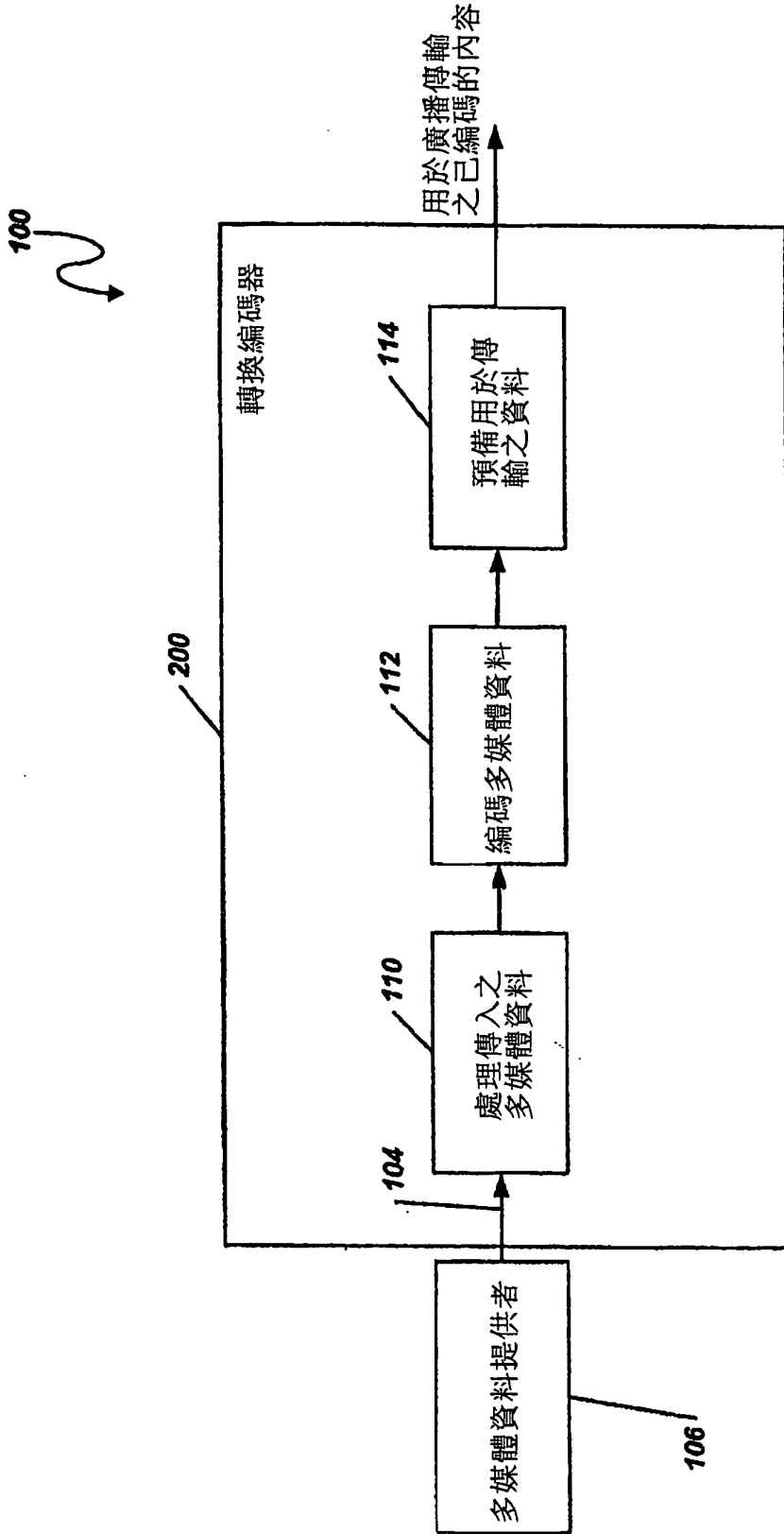


圖1A

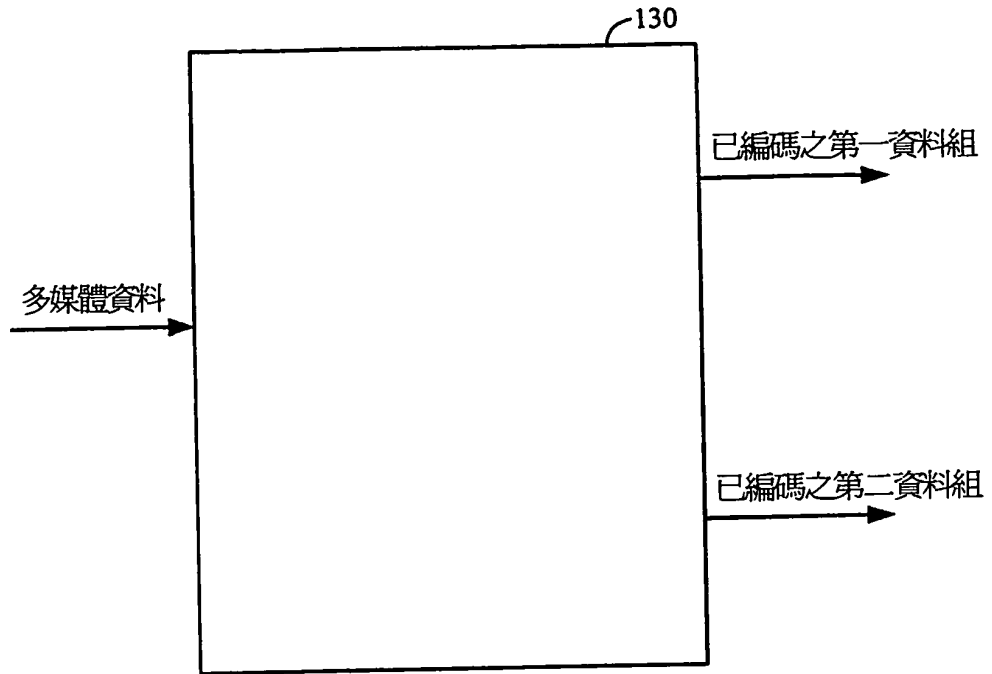


圖1B

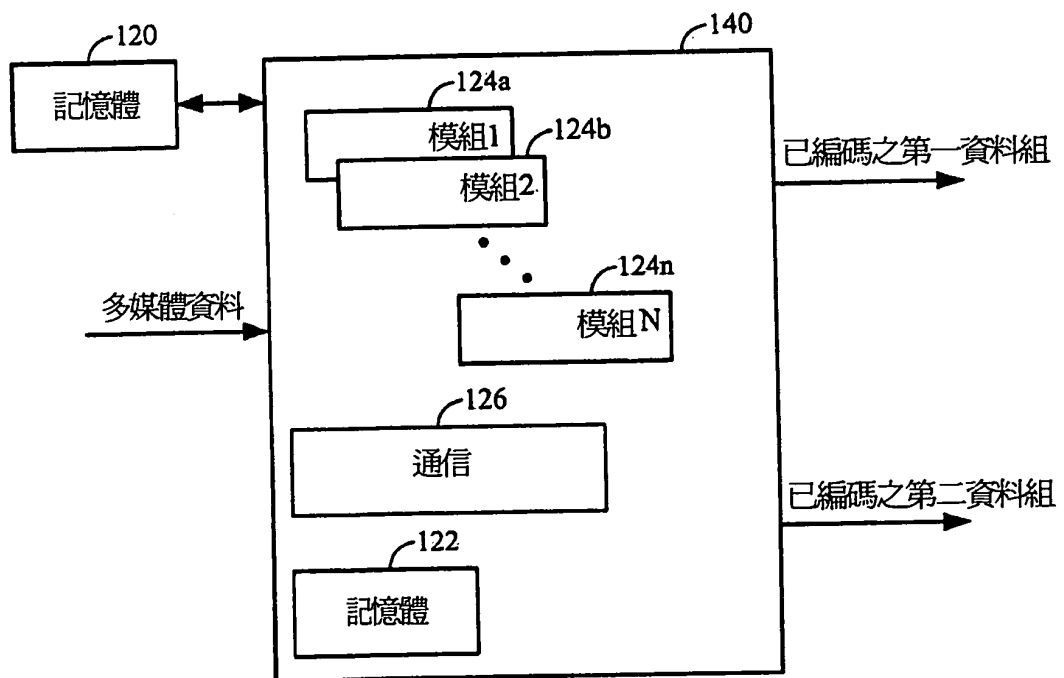


圖1C



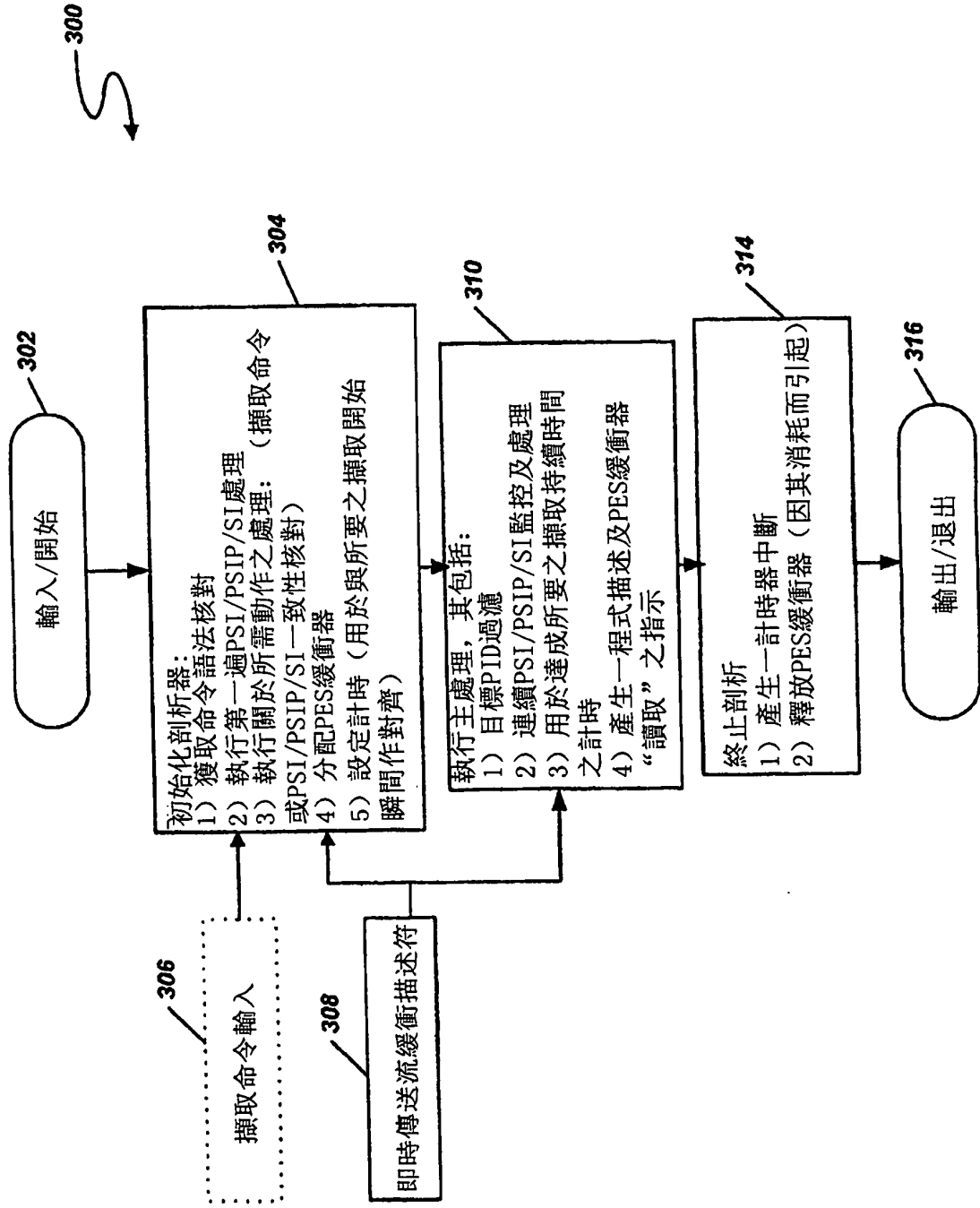


圖3

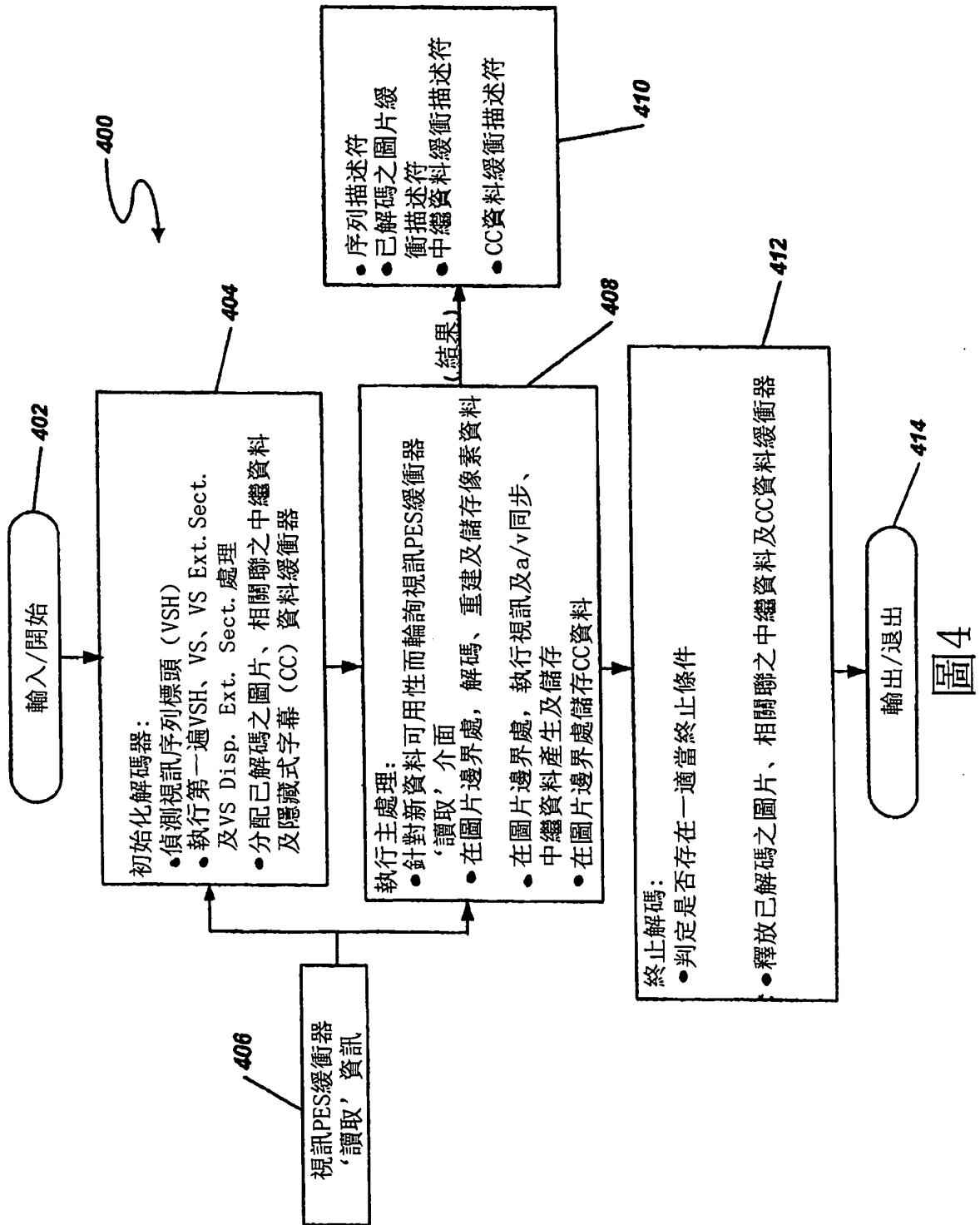


圖4

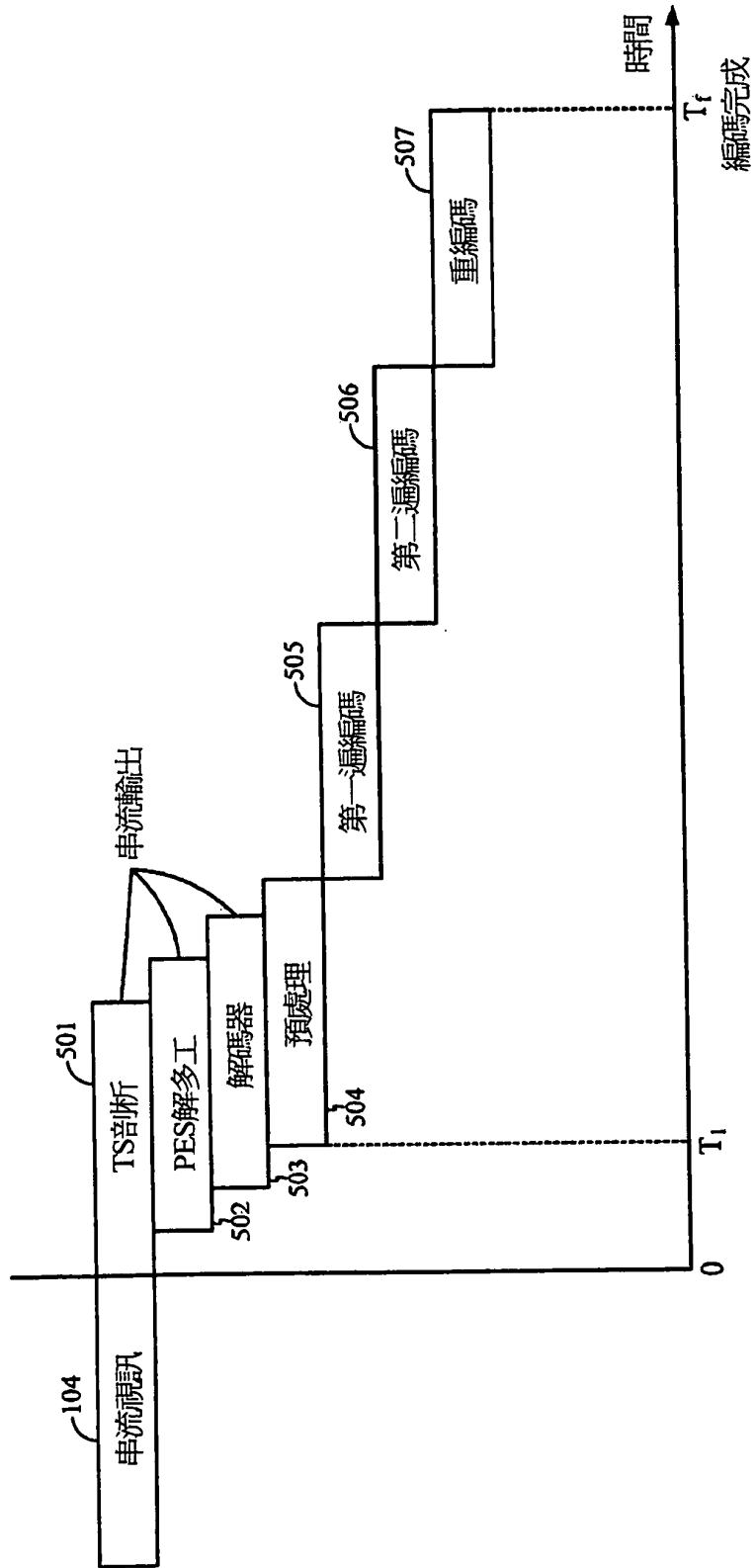


圖5

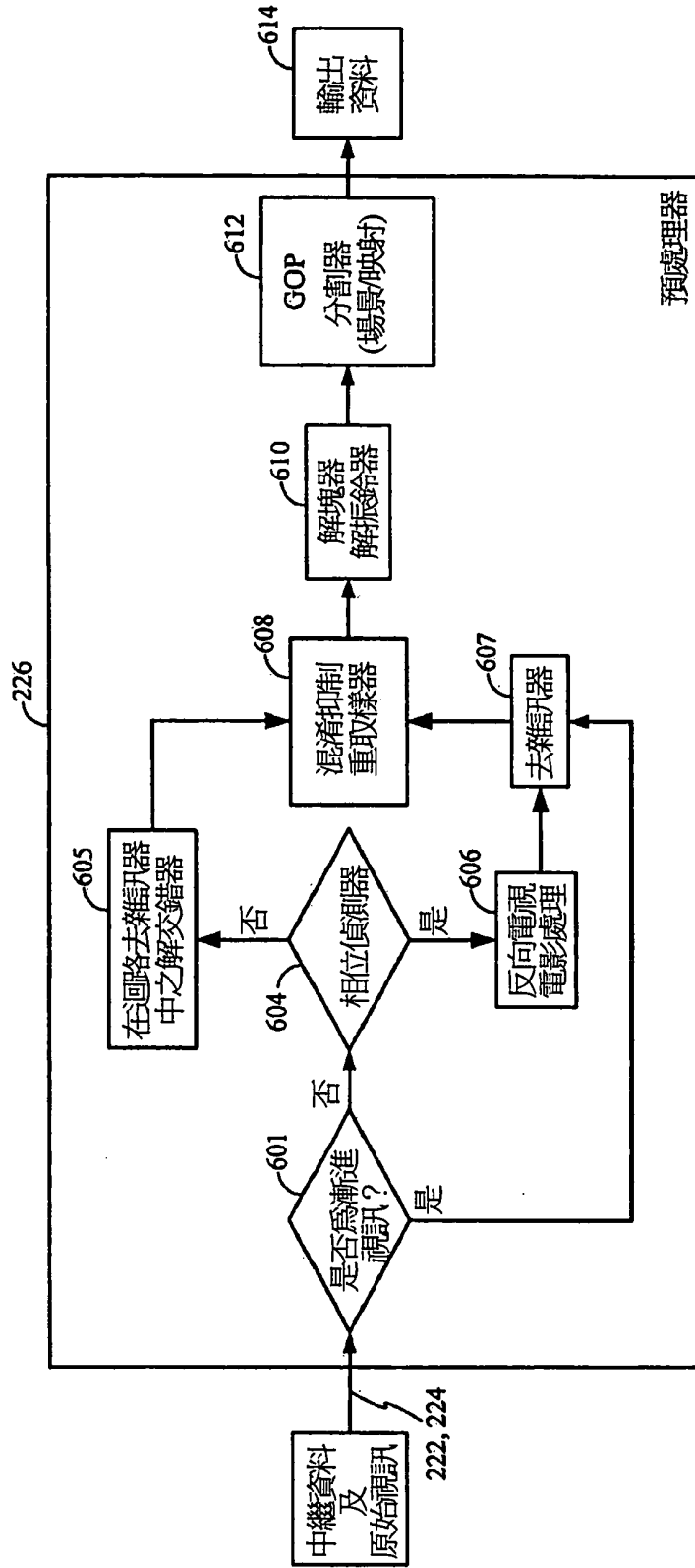


圖6

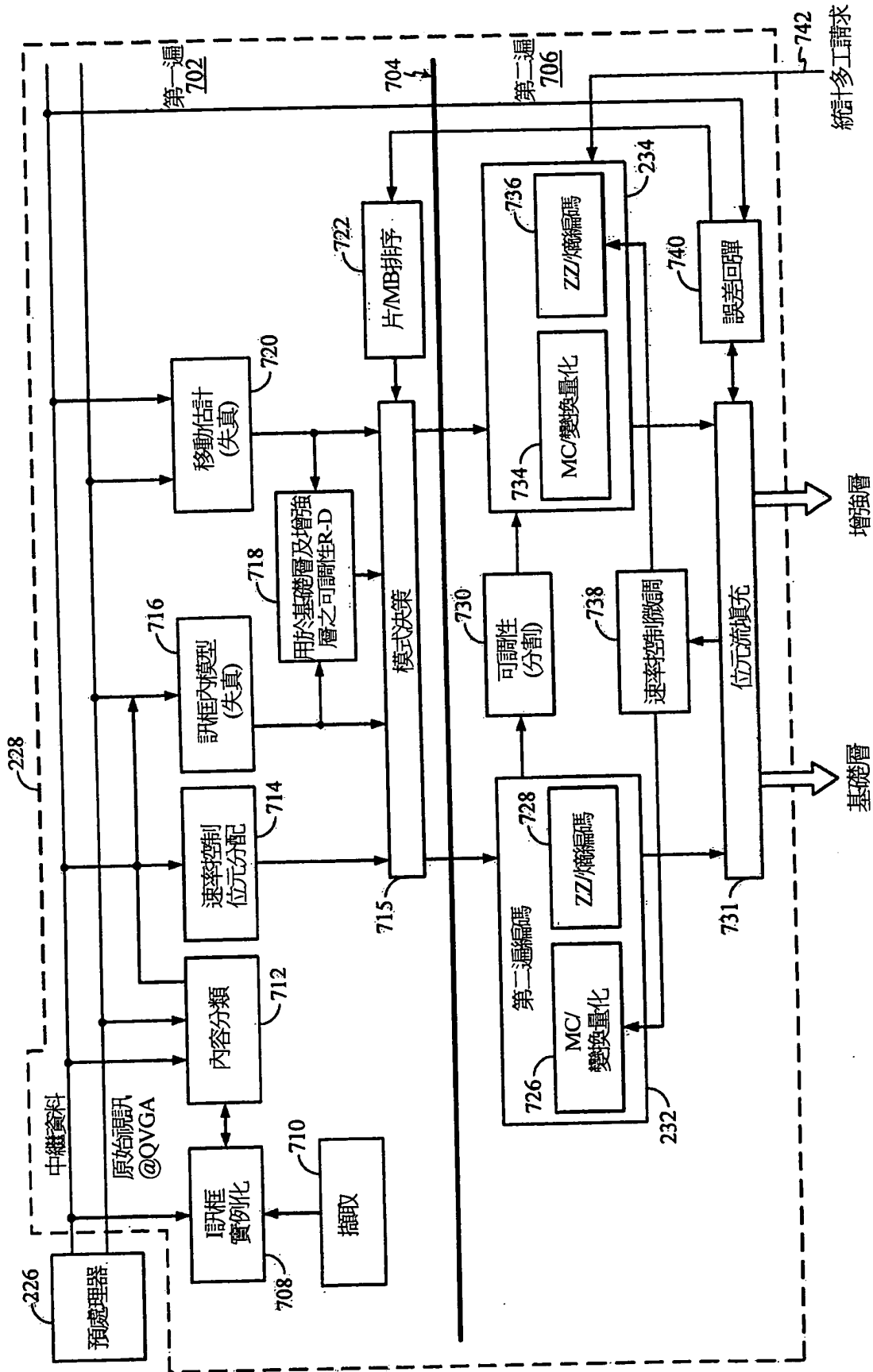
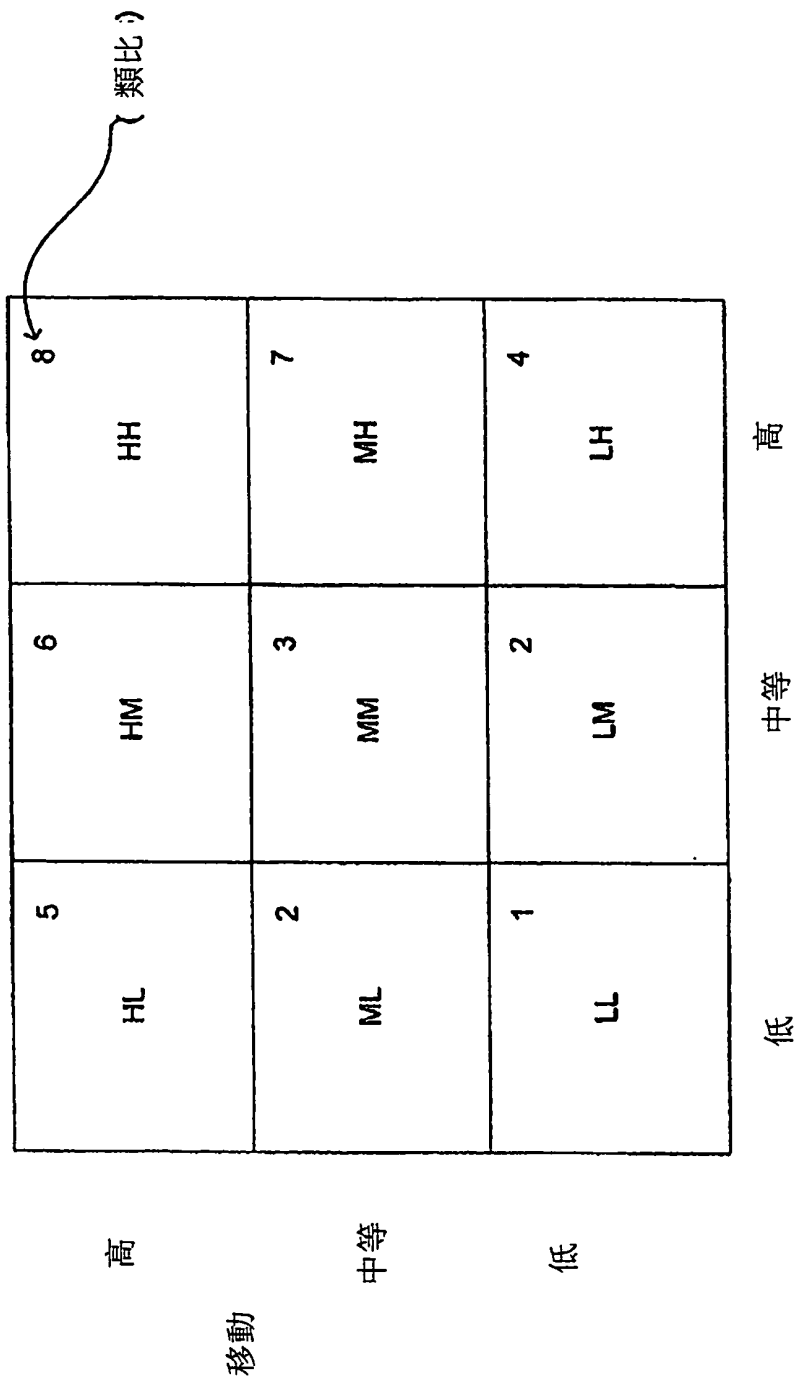


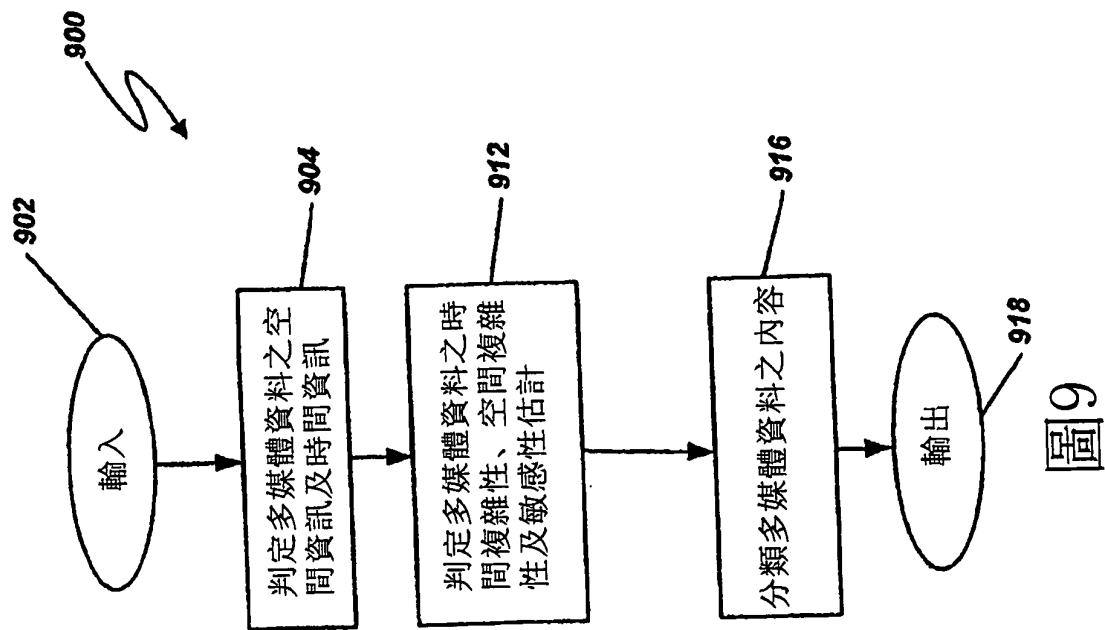
圖7





紋理

圖8



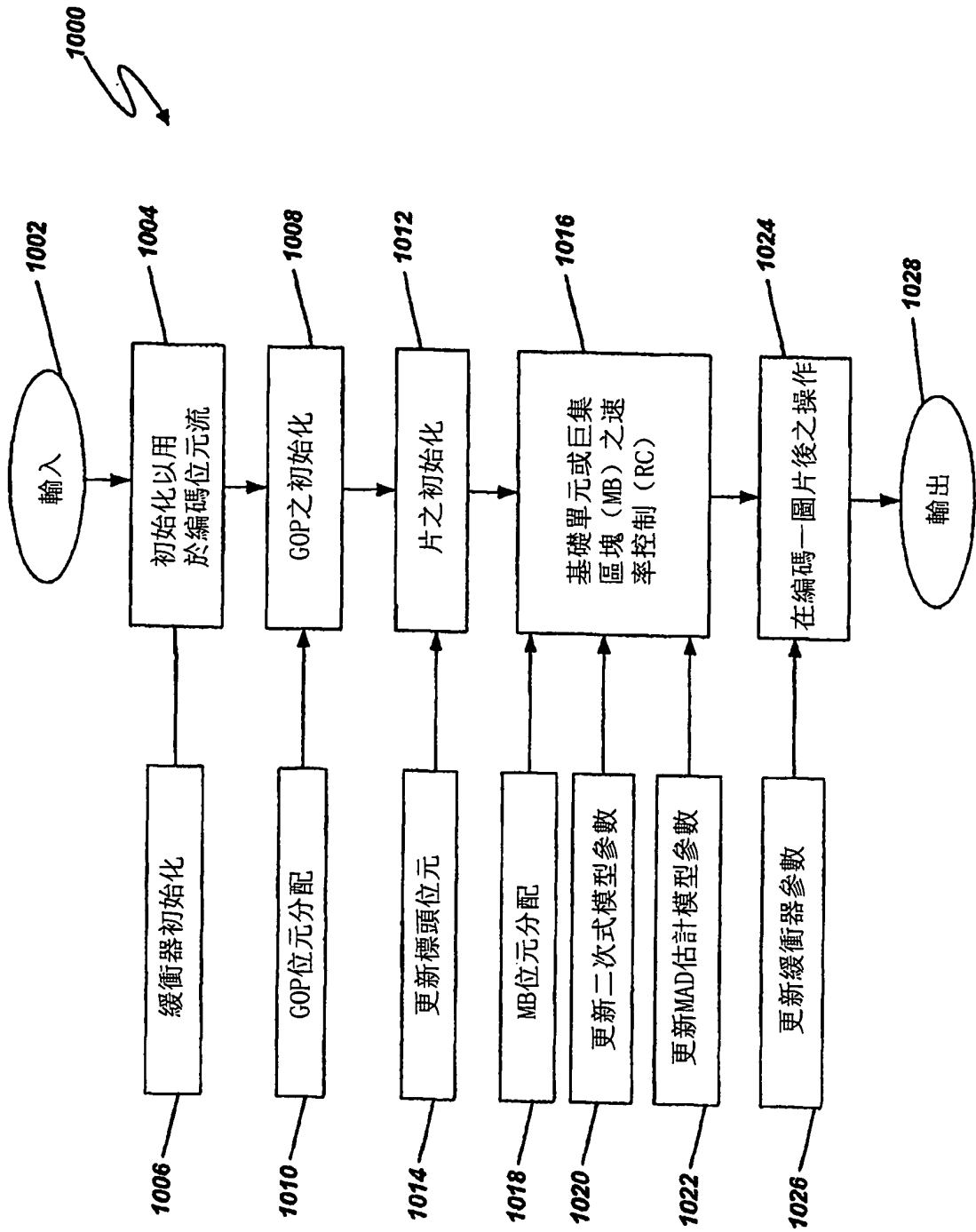


圖10

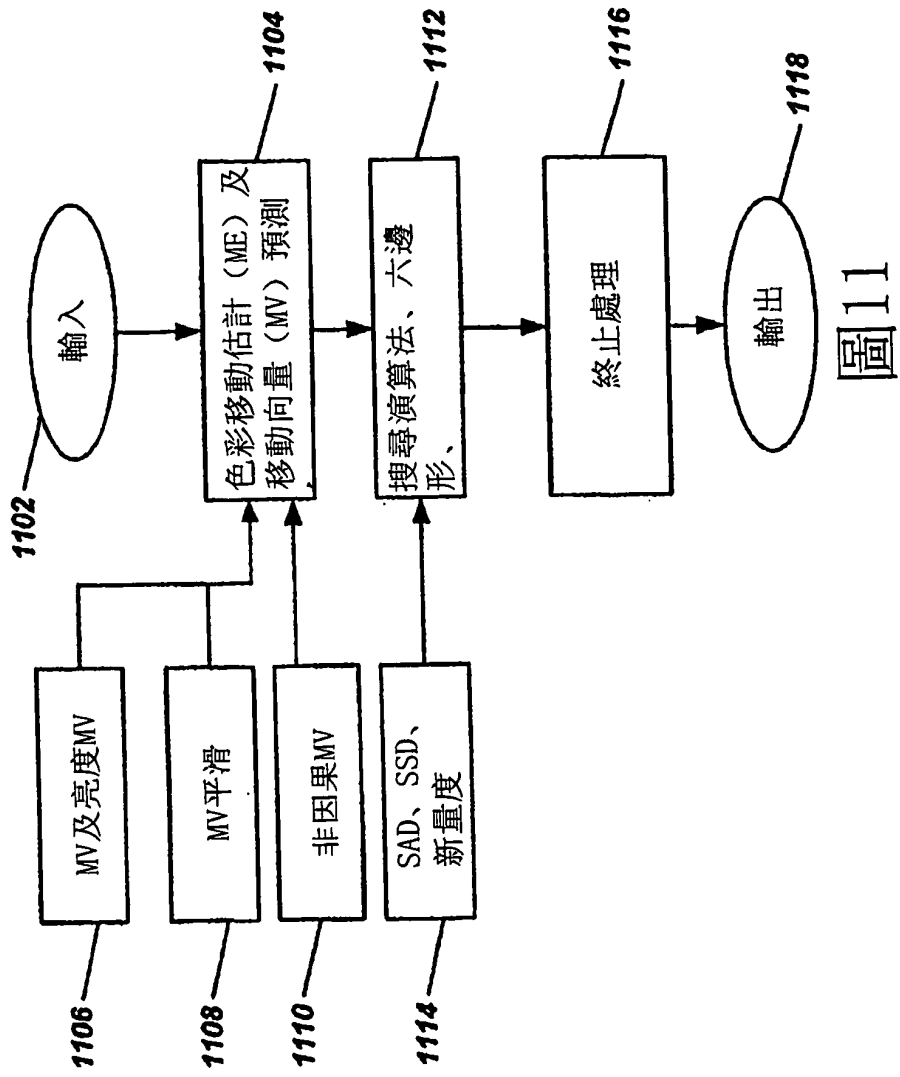


圖11

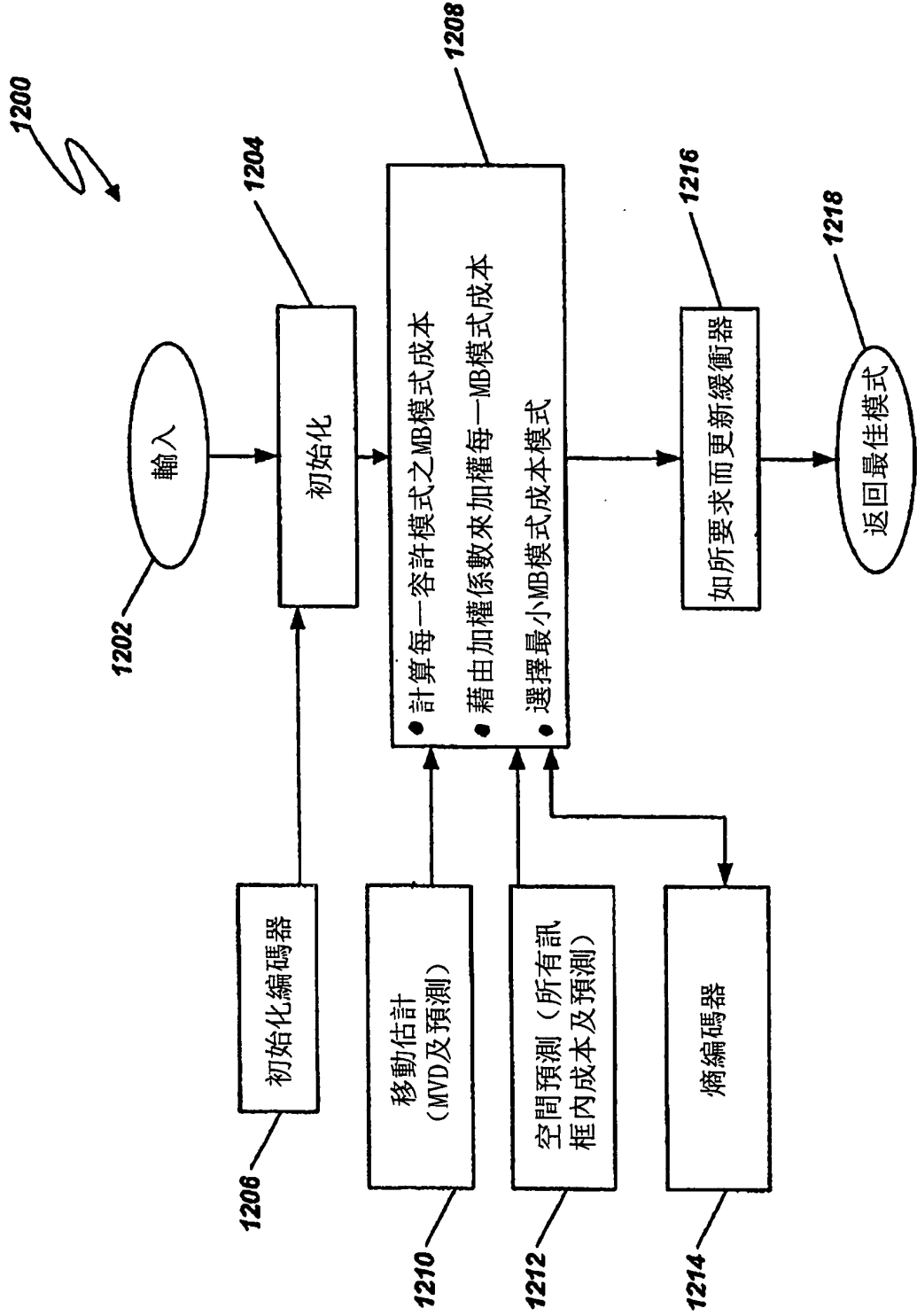


圖12

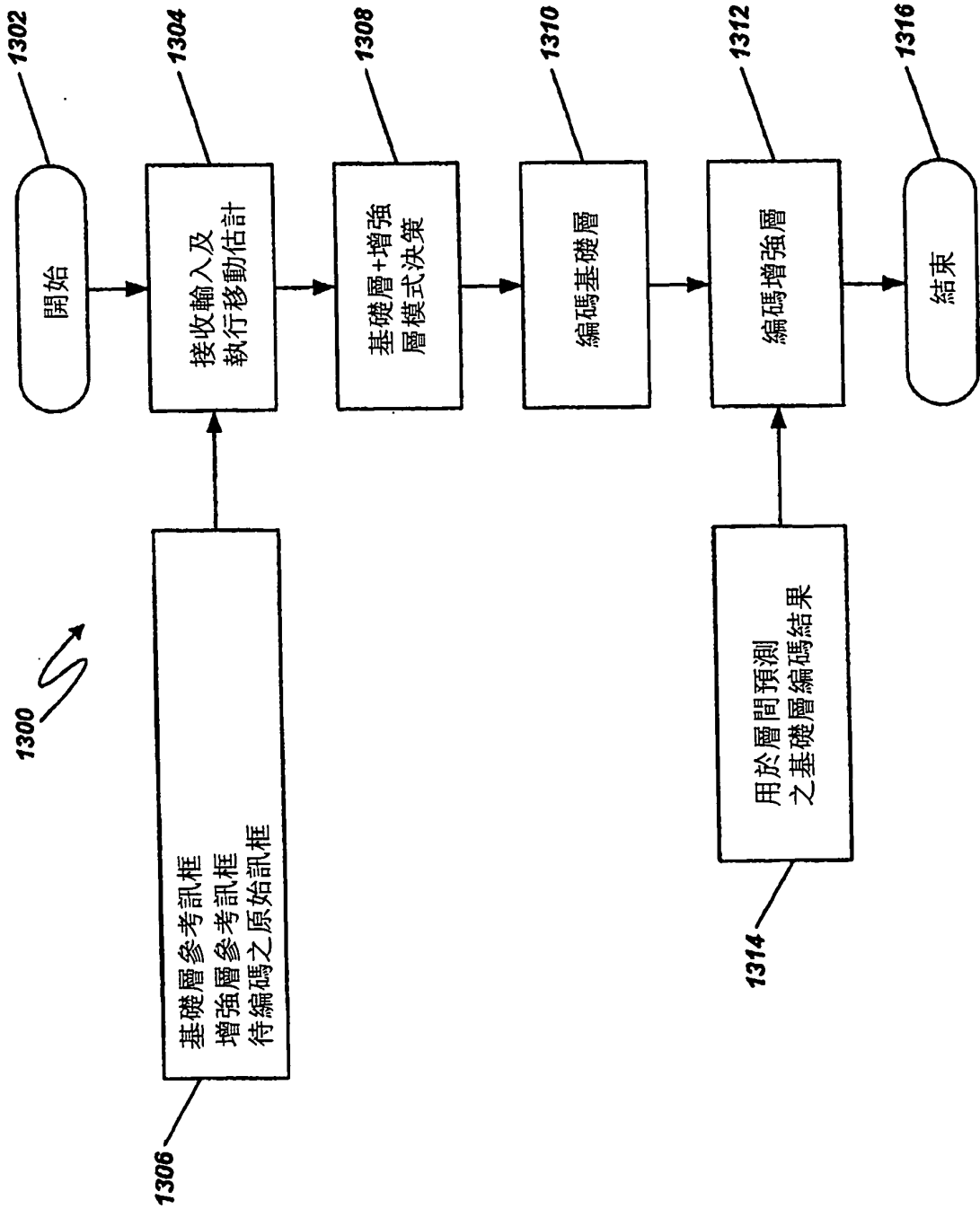


圖13



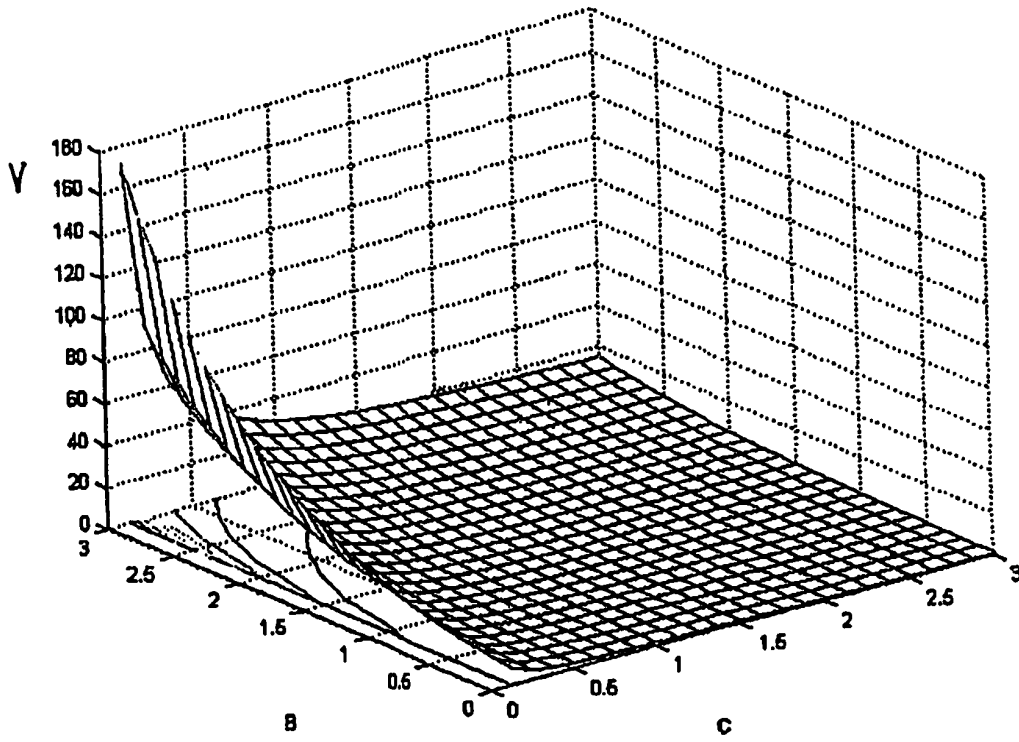


圖15



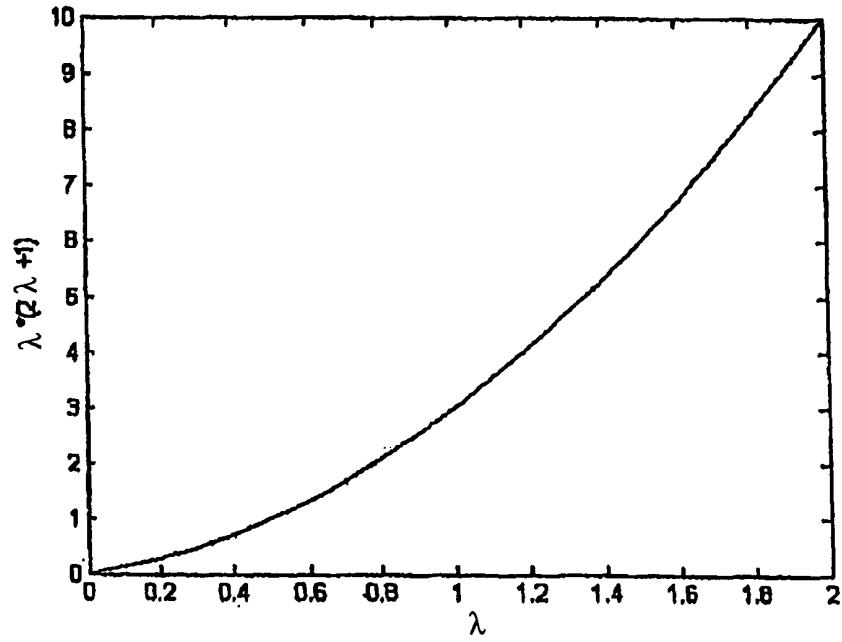


圖16

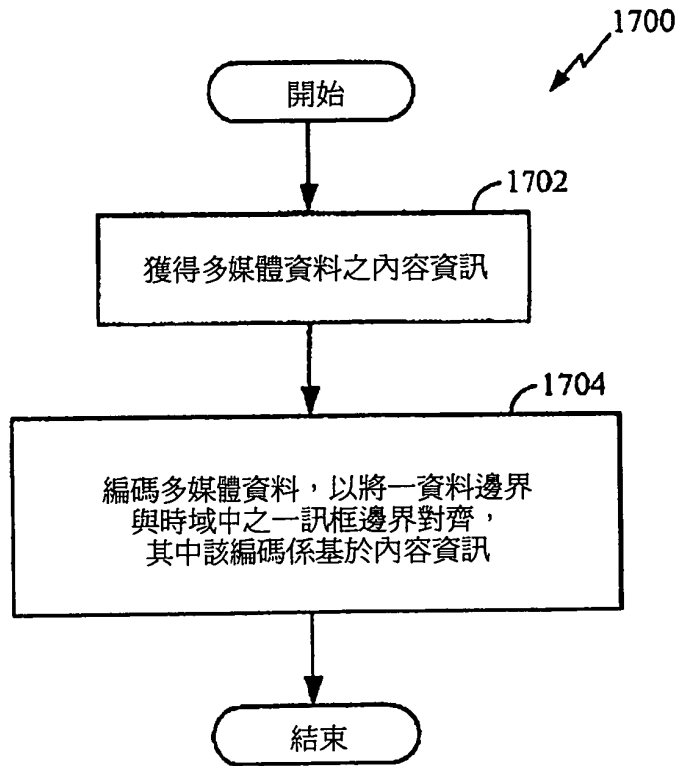


圖17A

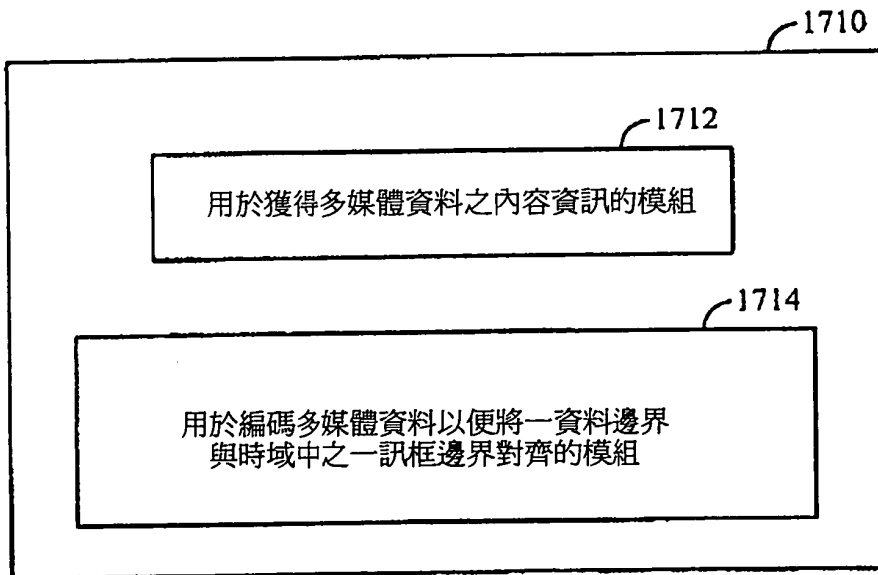


圖17B

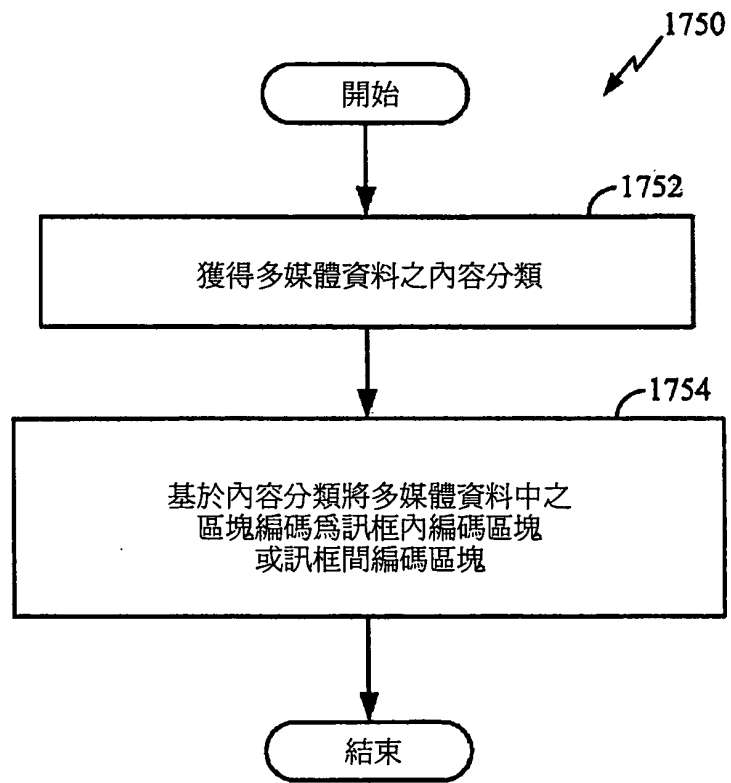


圖17C

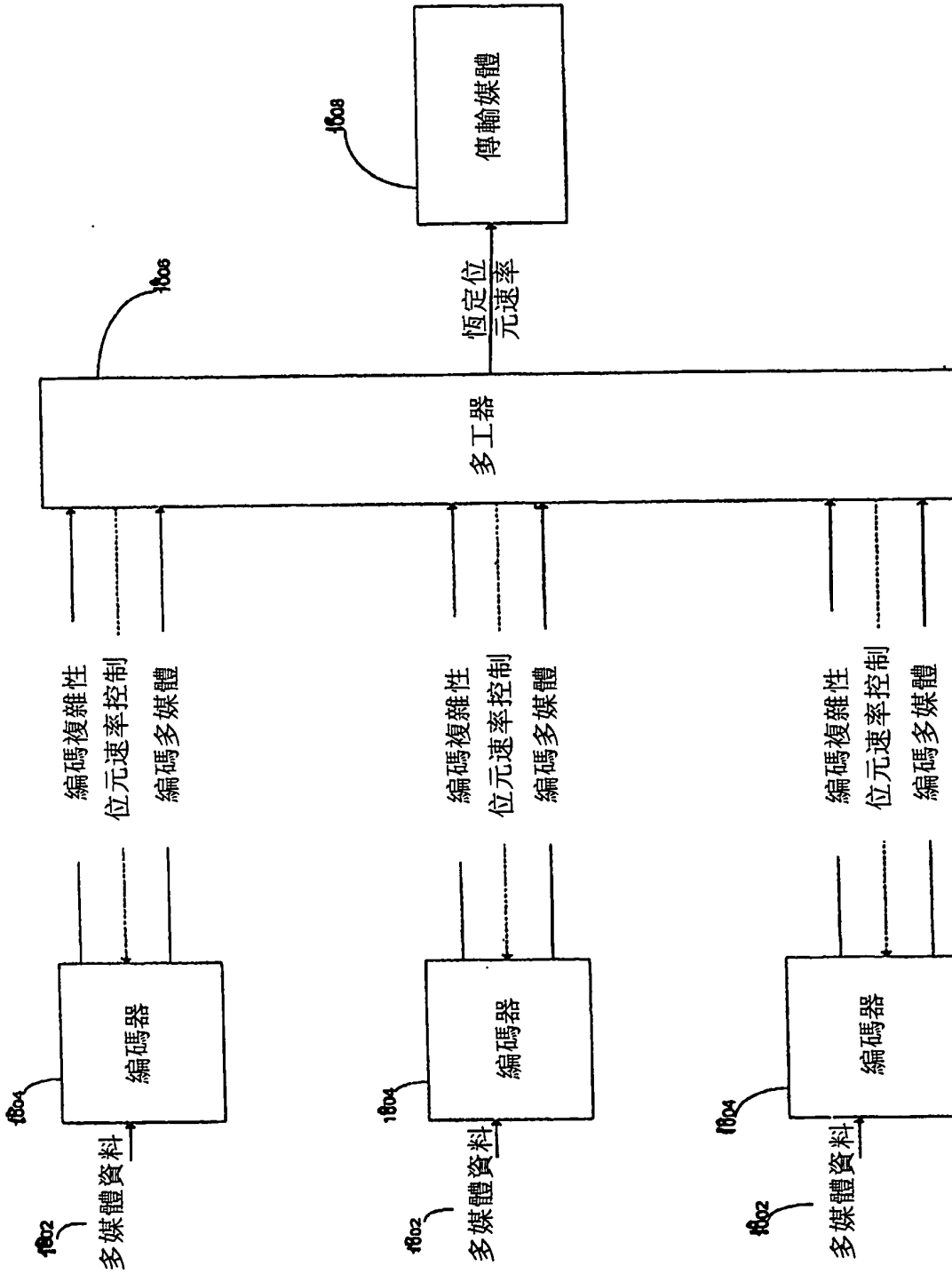


圖18

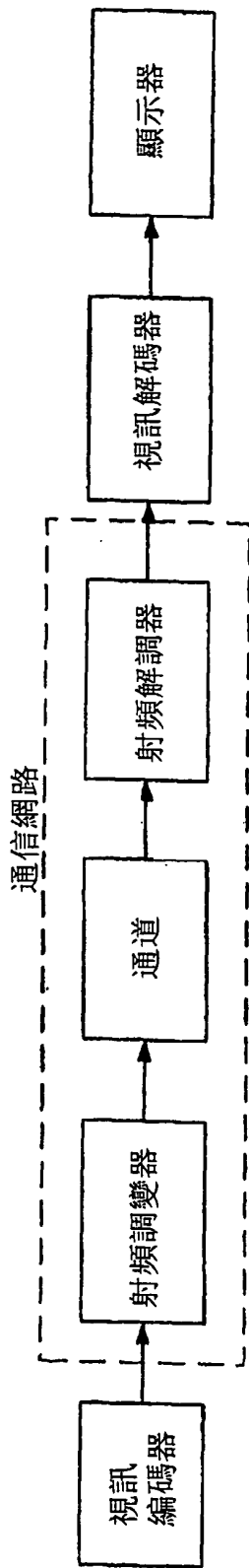


圖19

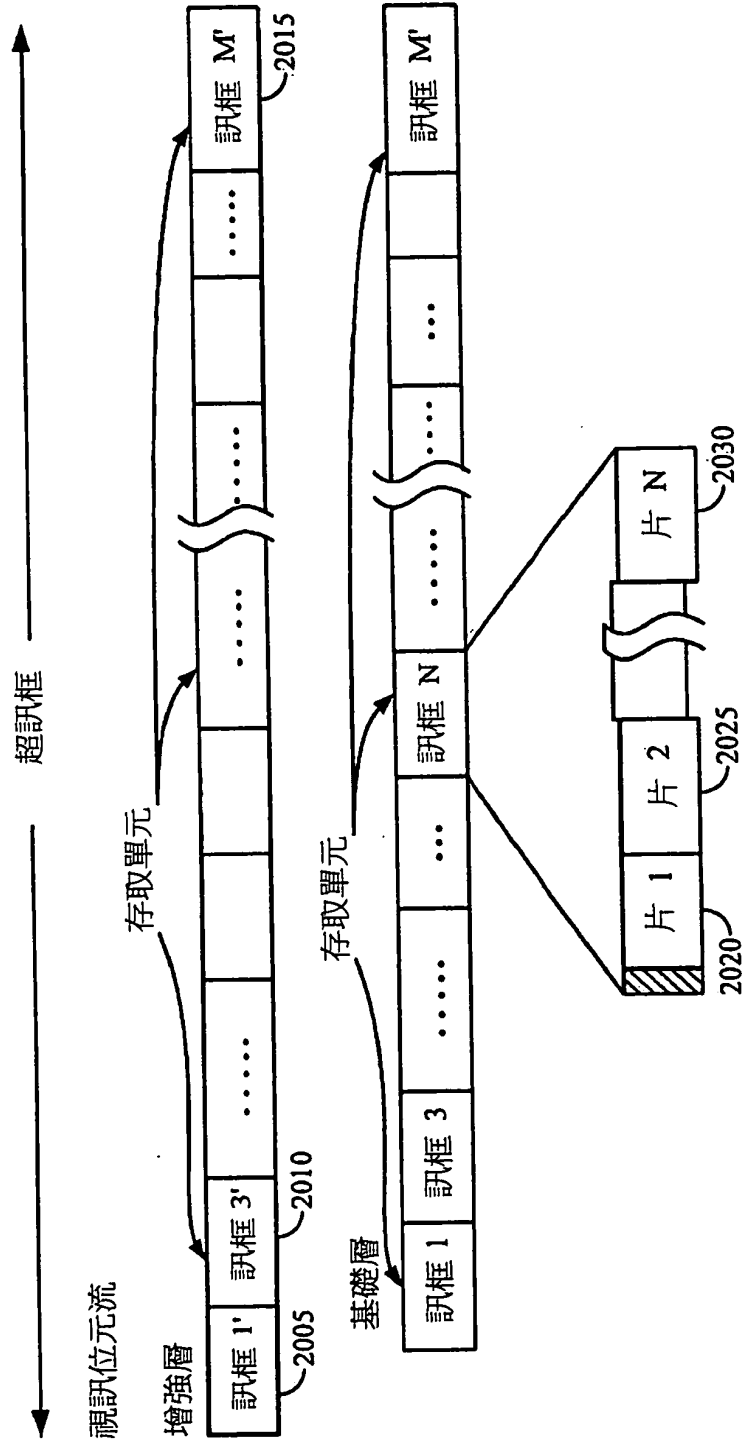


圖20

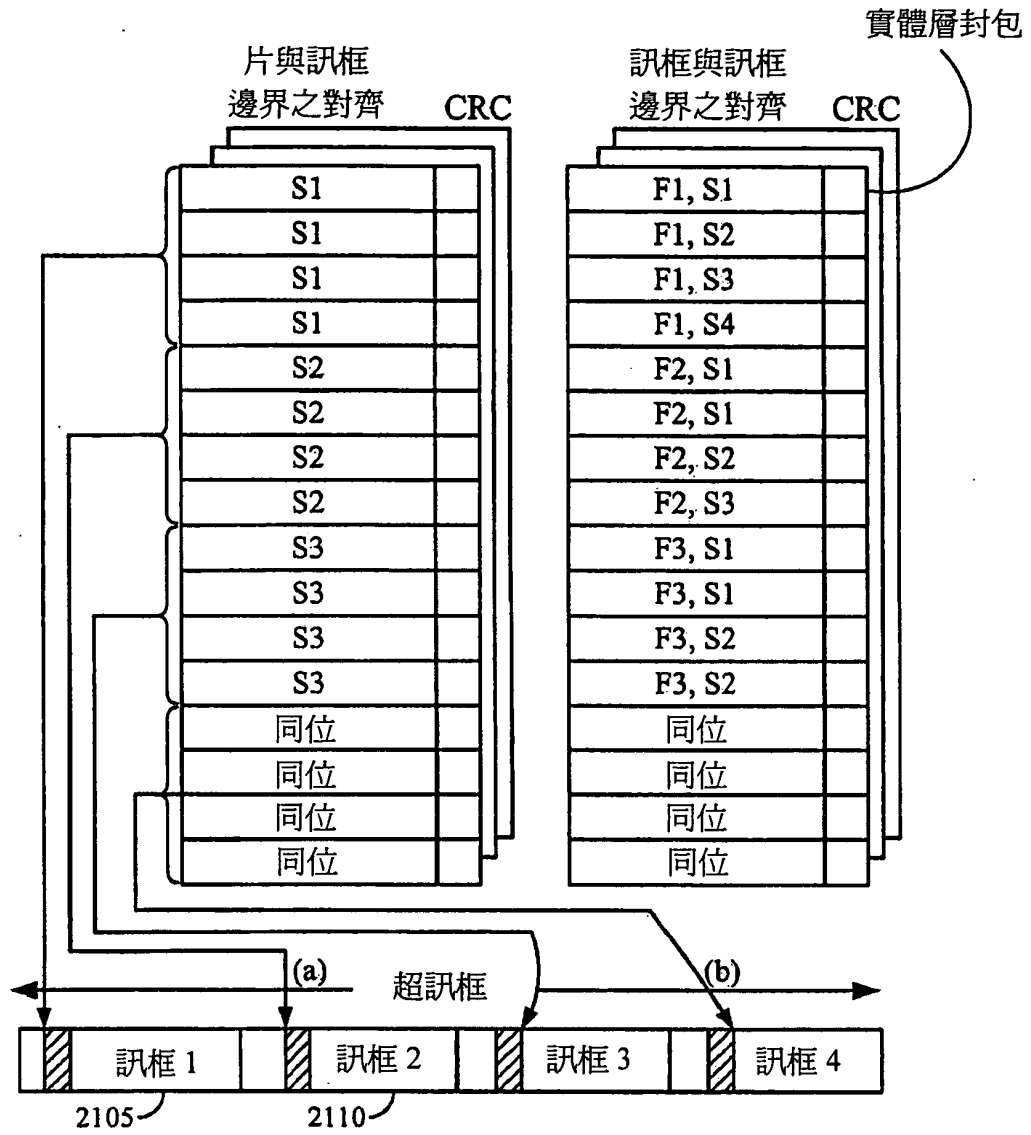


圖21

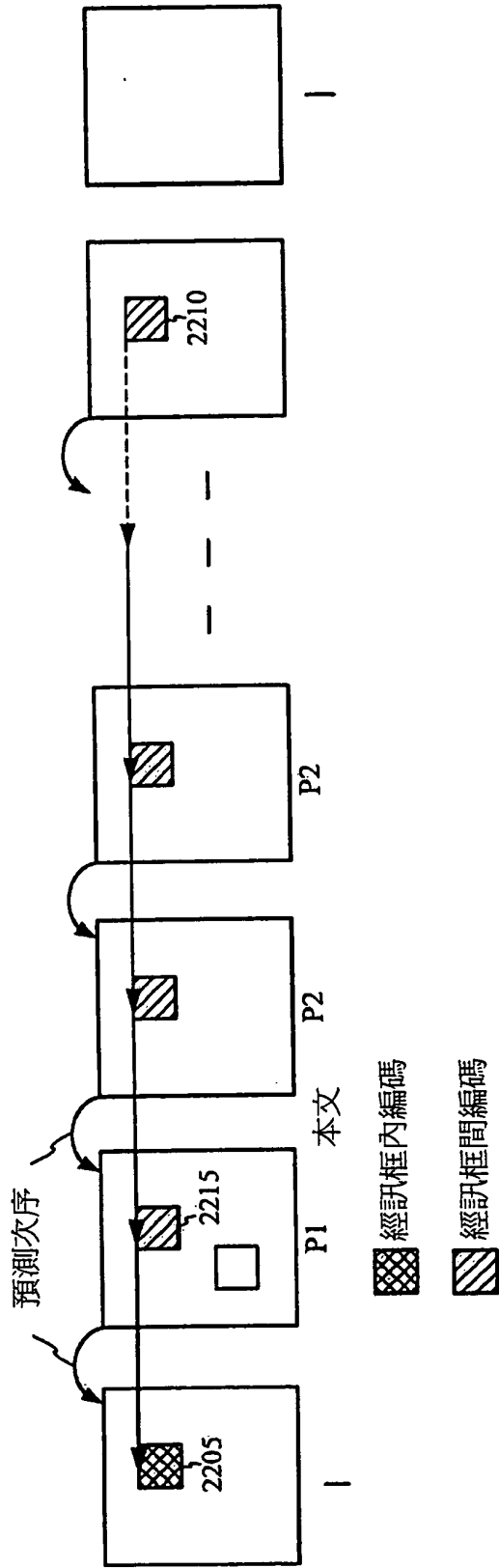


圖22



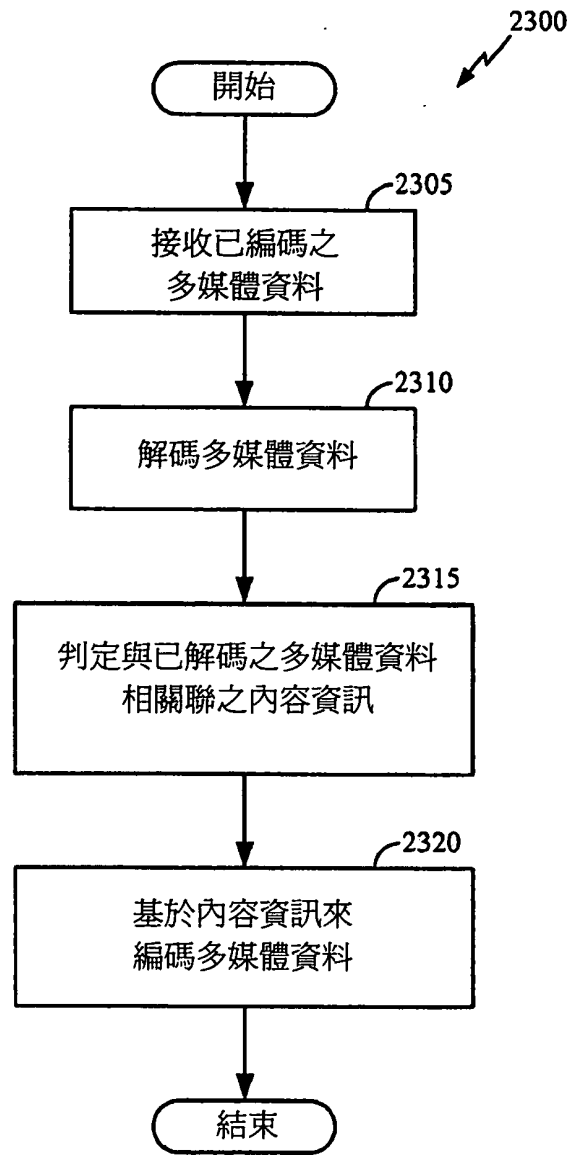


圖23

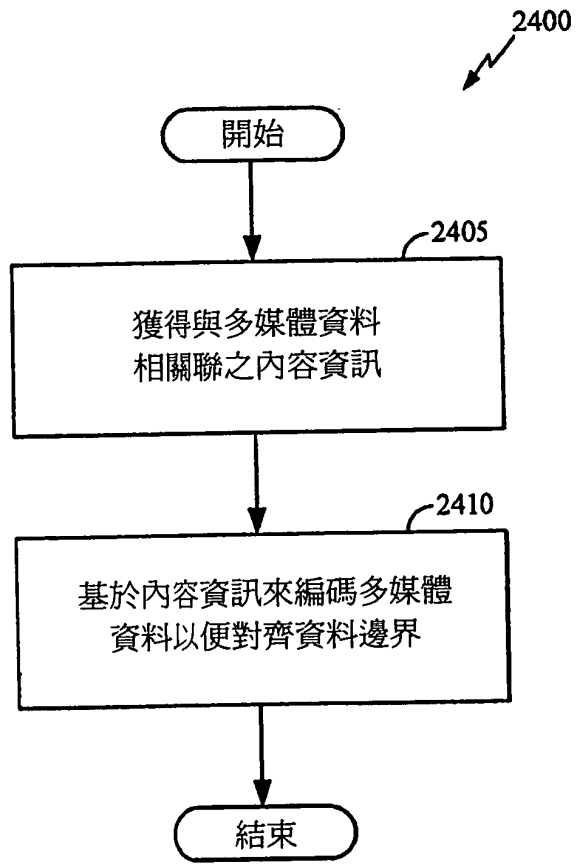
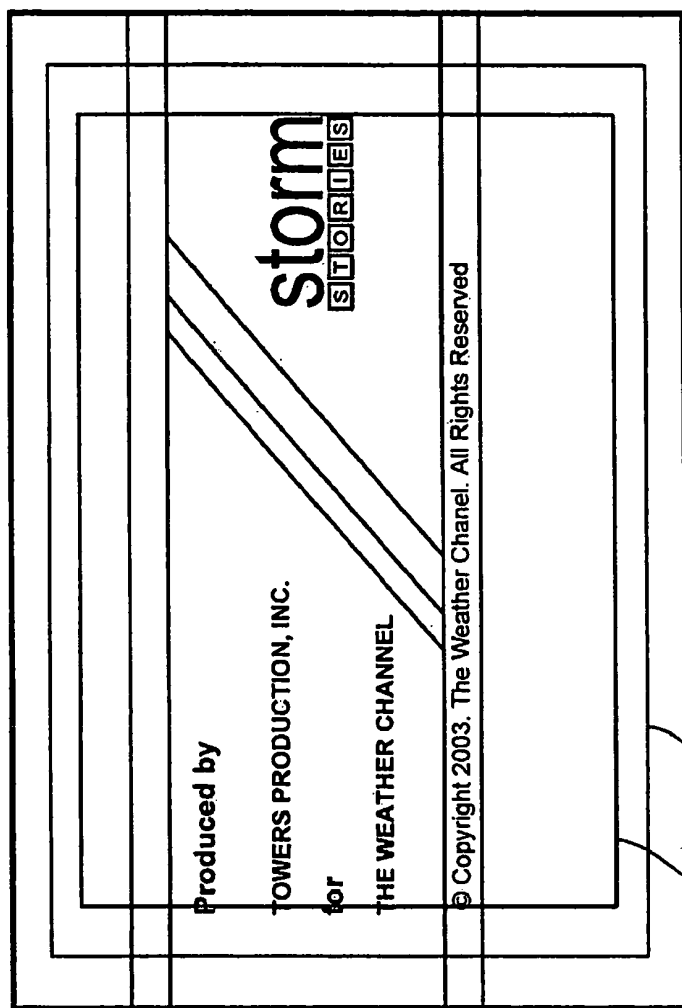


圖24



2510 2505

圖25

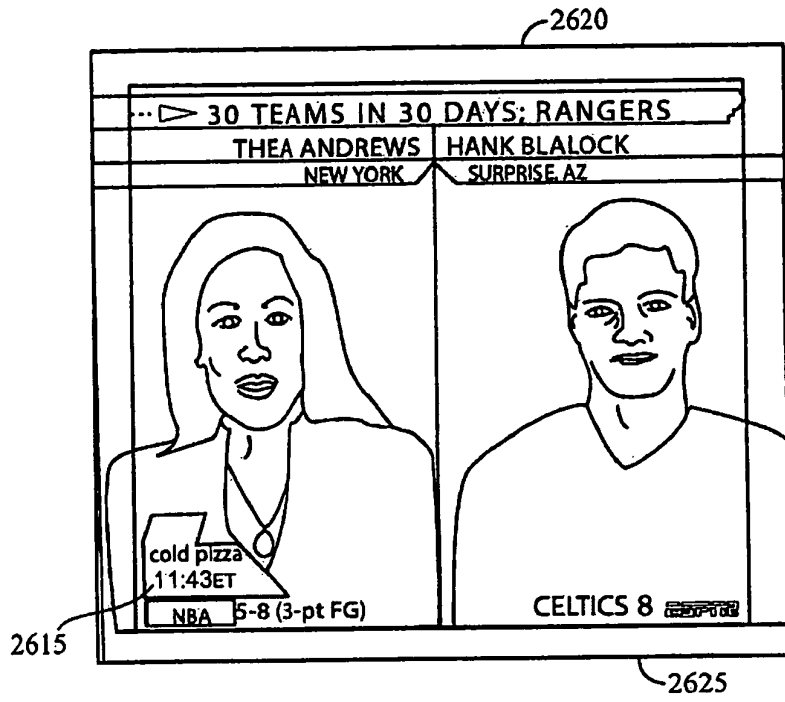


圖26

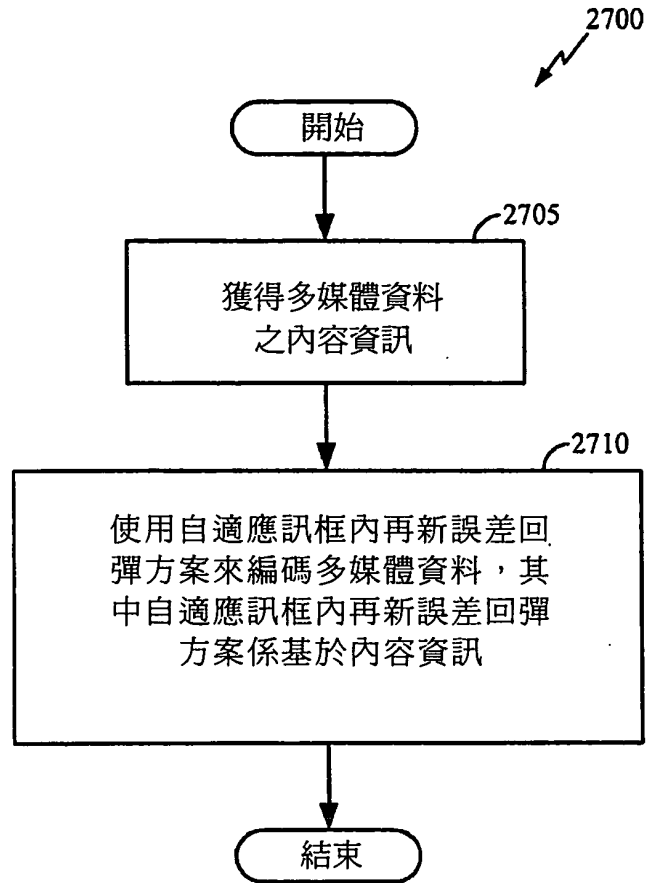


圖27

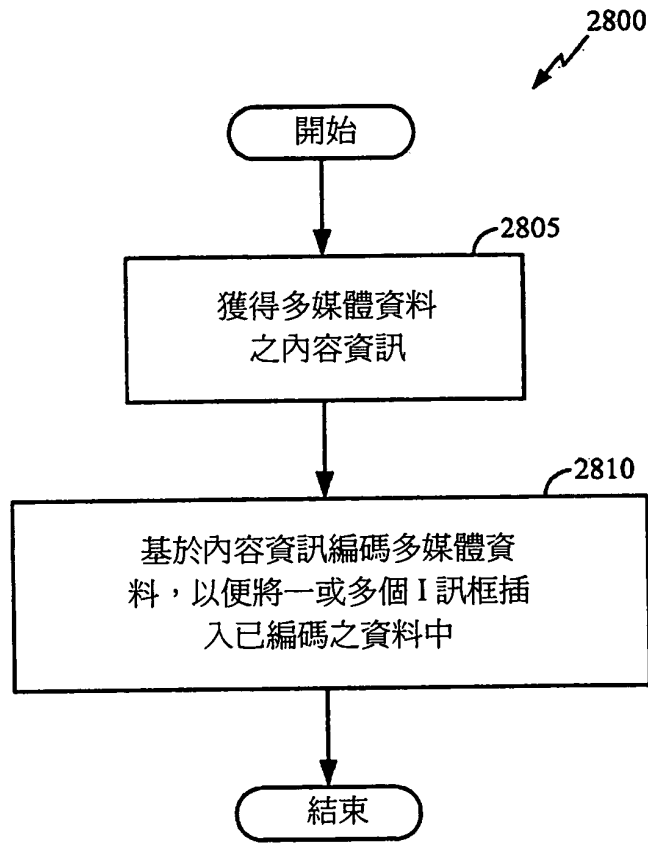


圖28

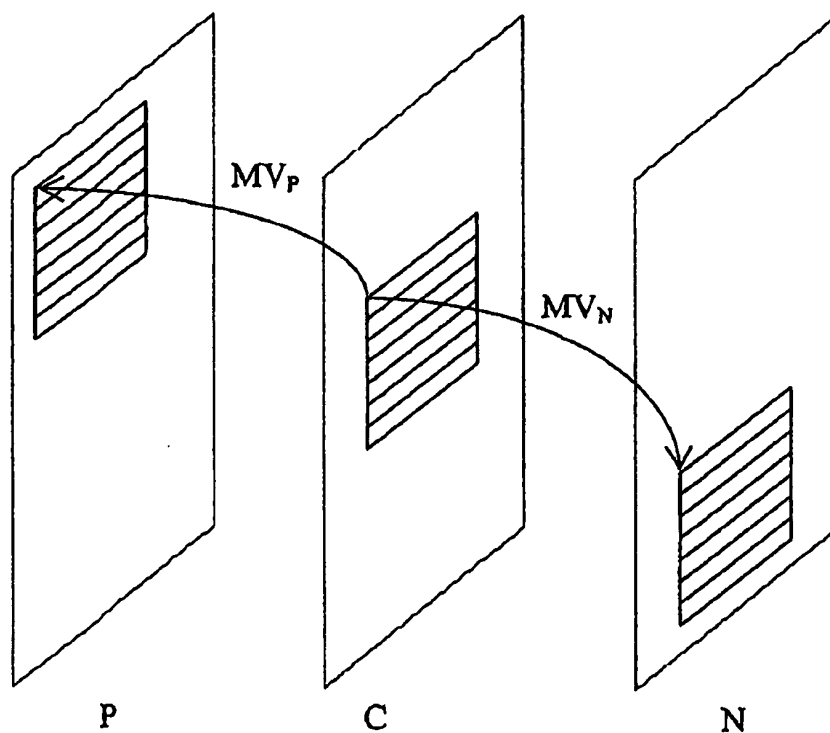


圖29

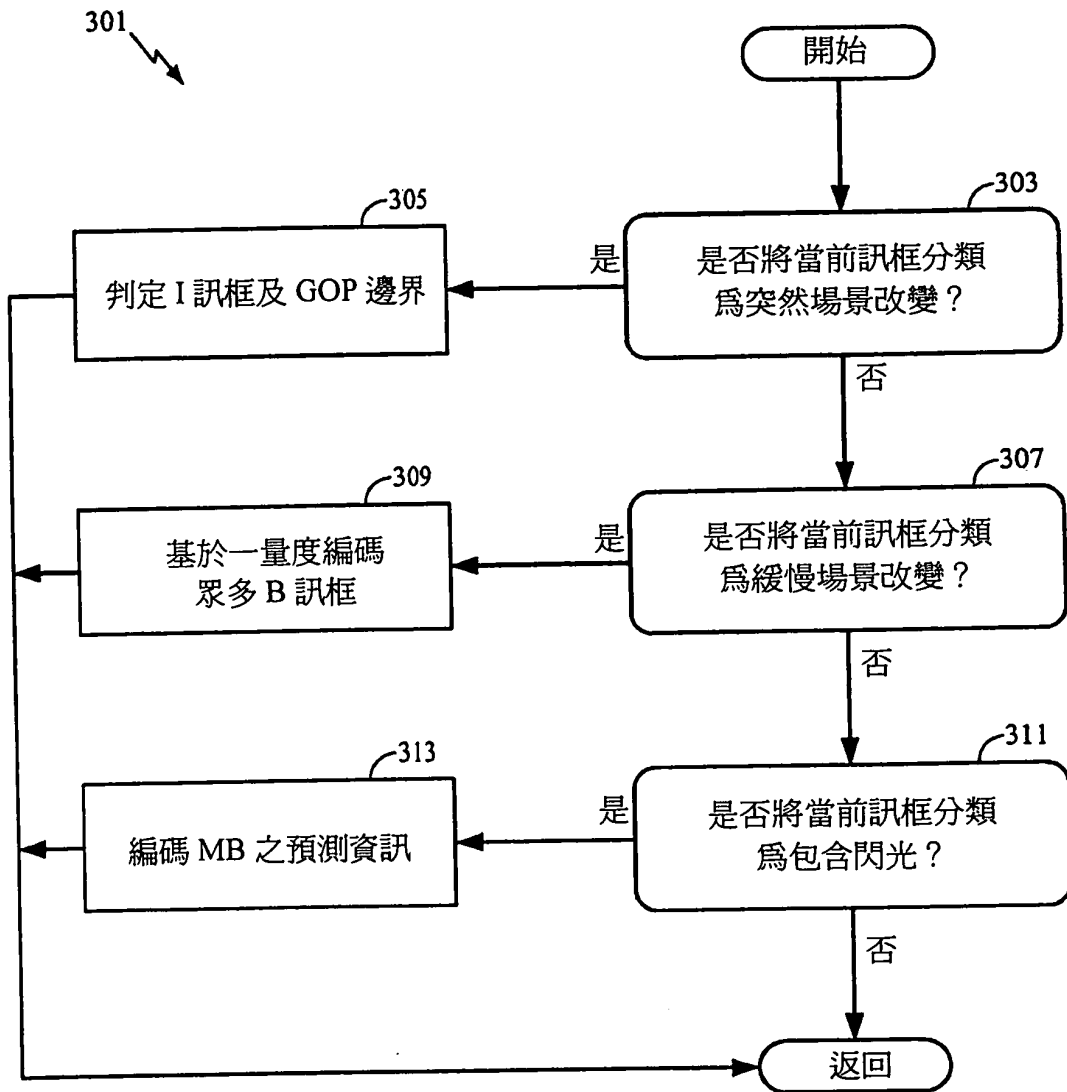


圖30



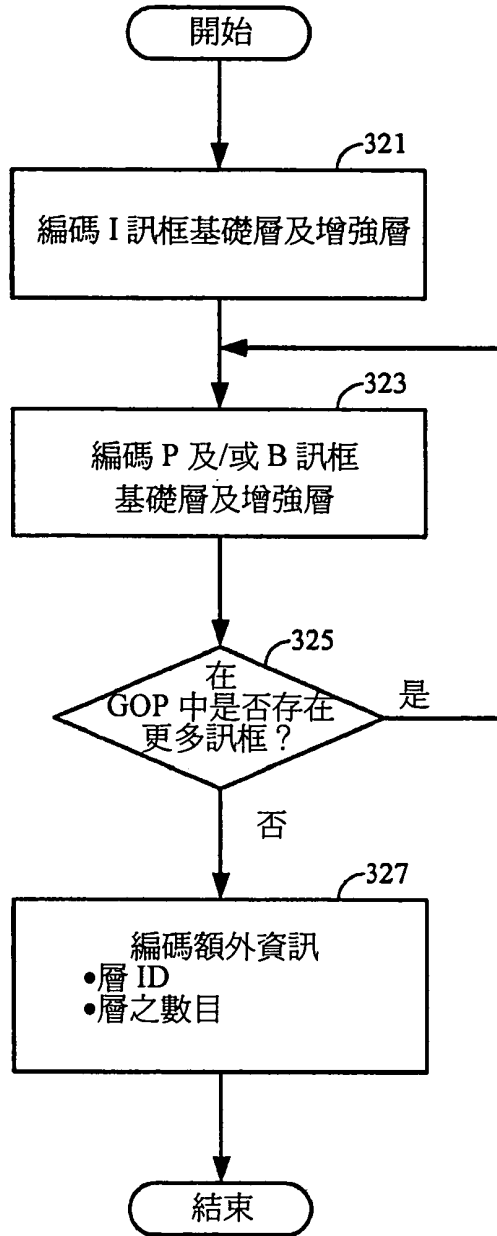


圖31

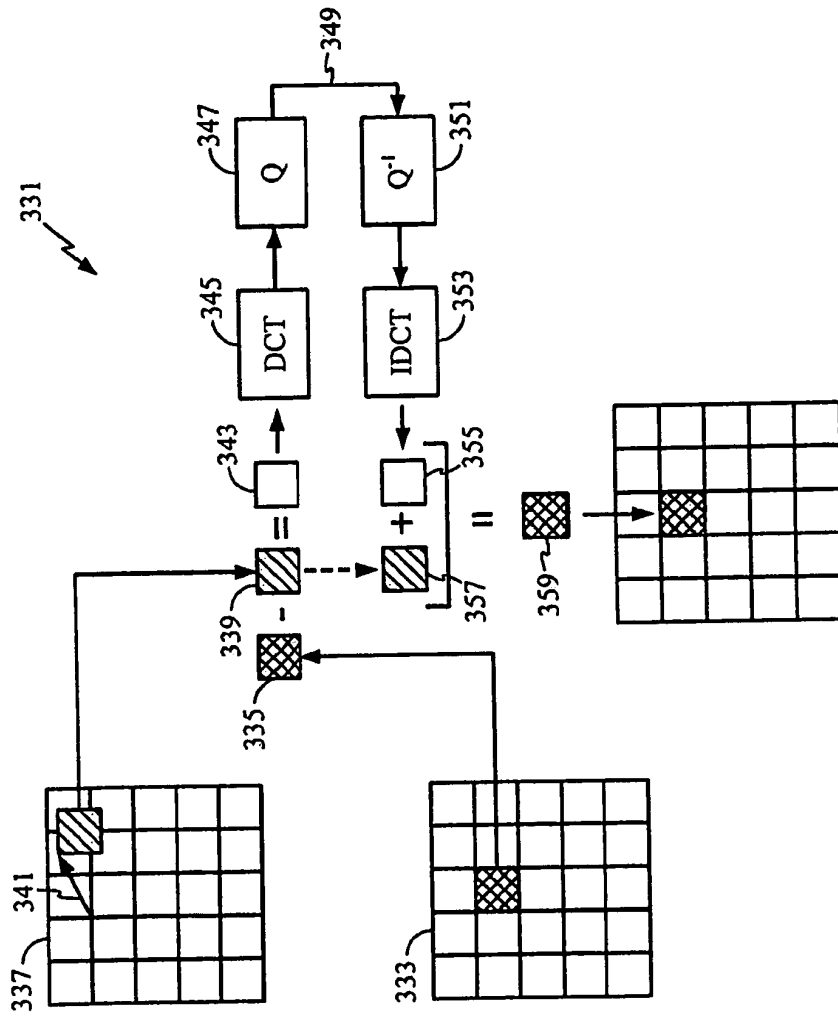


圖32

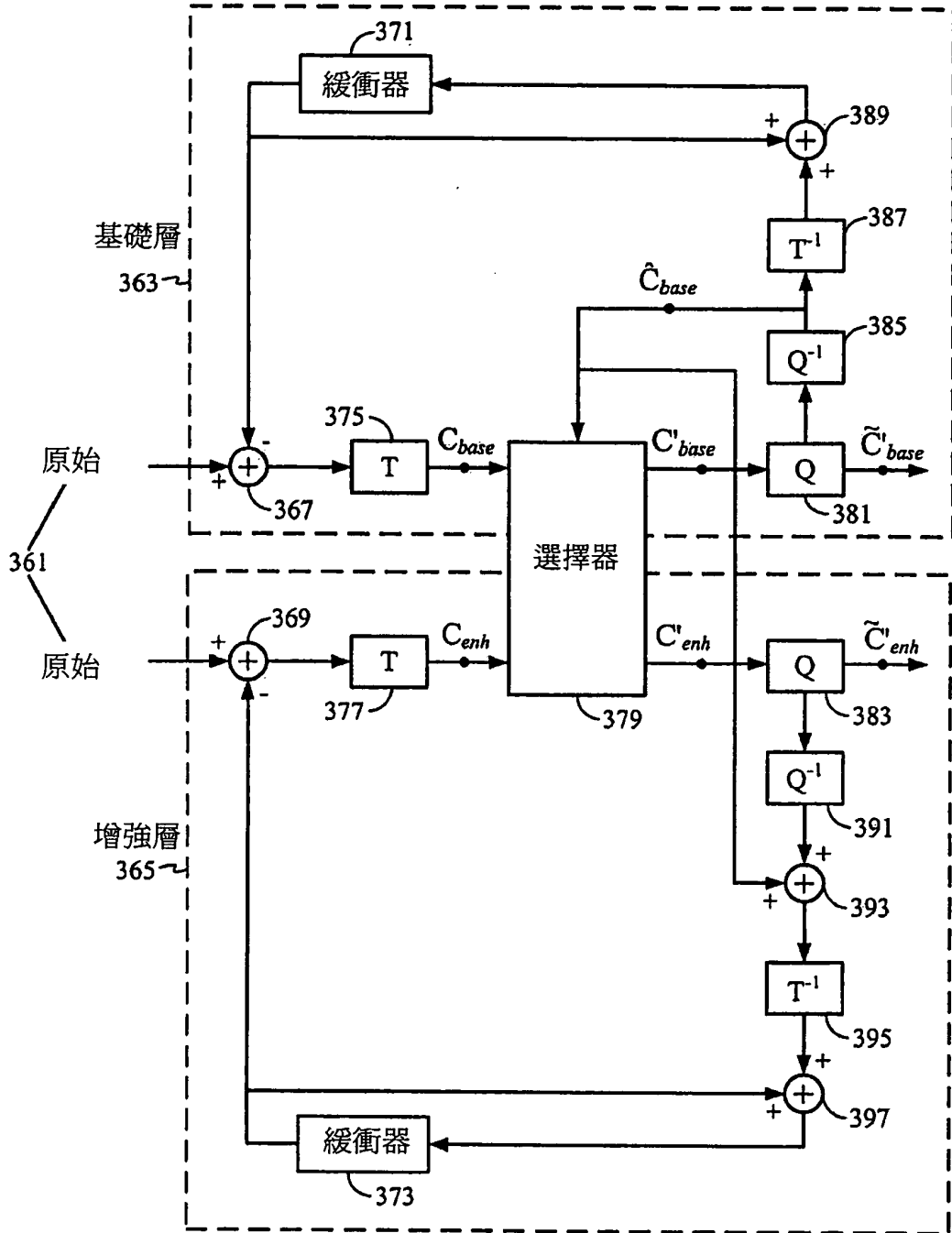


圖33

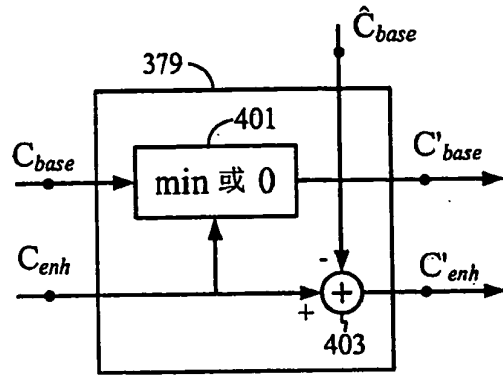


圖34

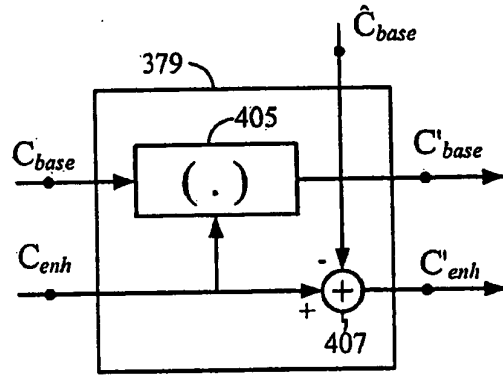


圖35

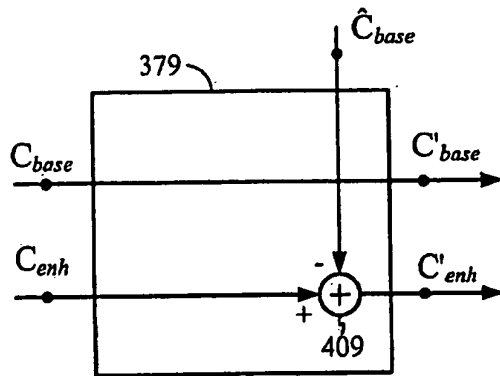


圖36

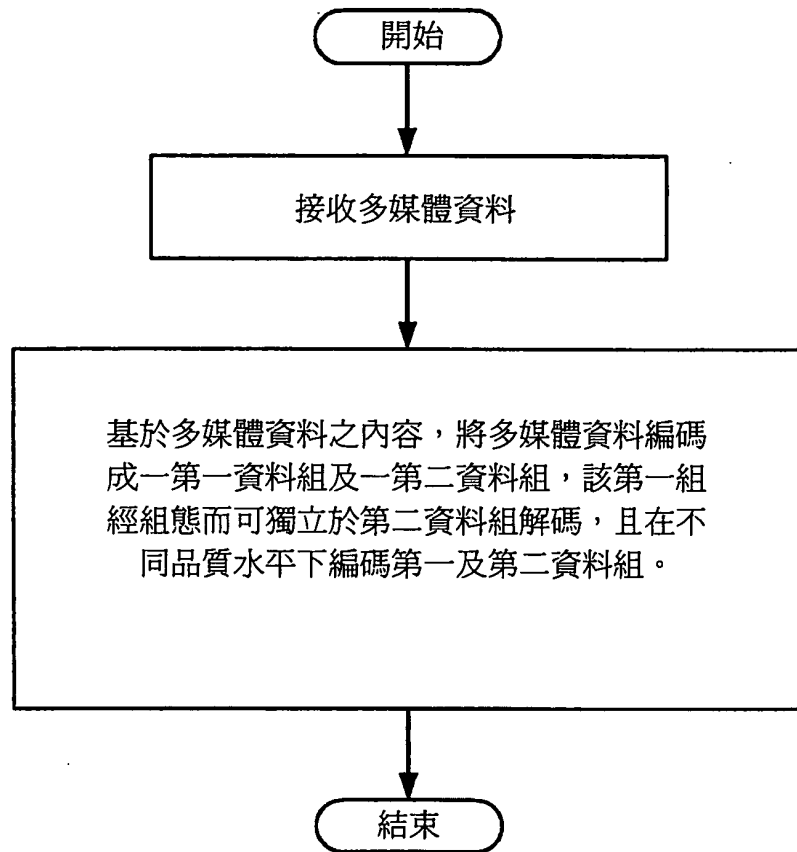
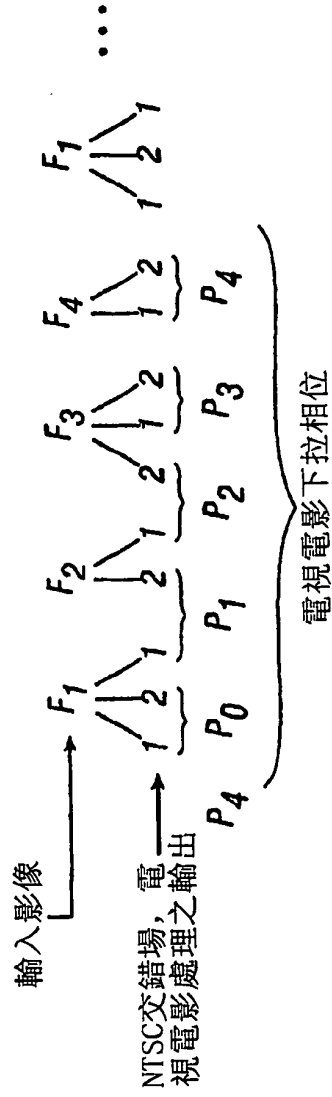


圖37



可能之系統決策

圖38

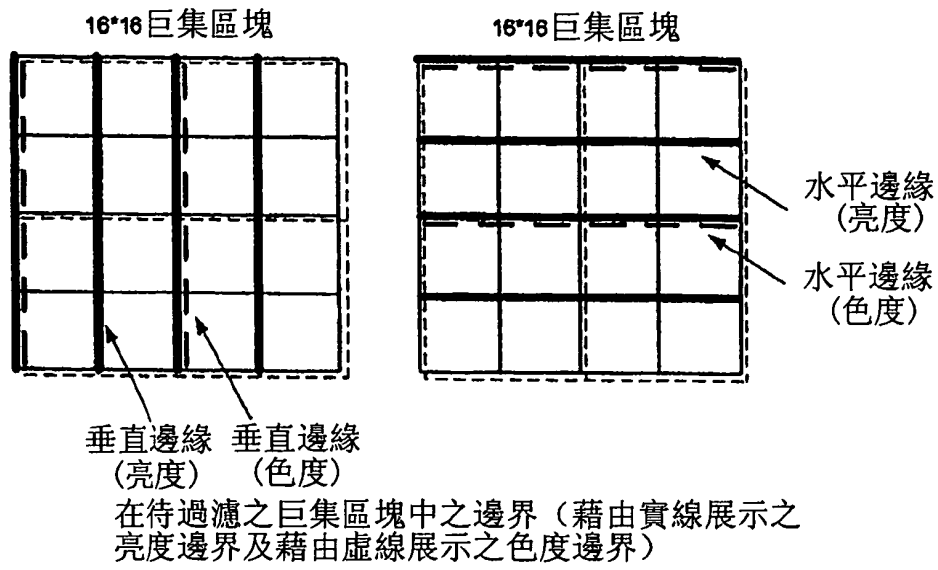


圖39

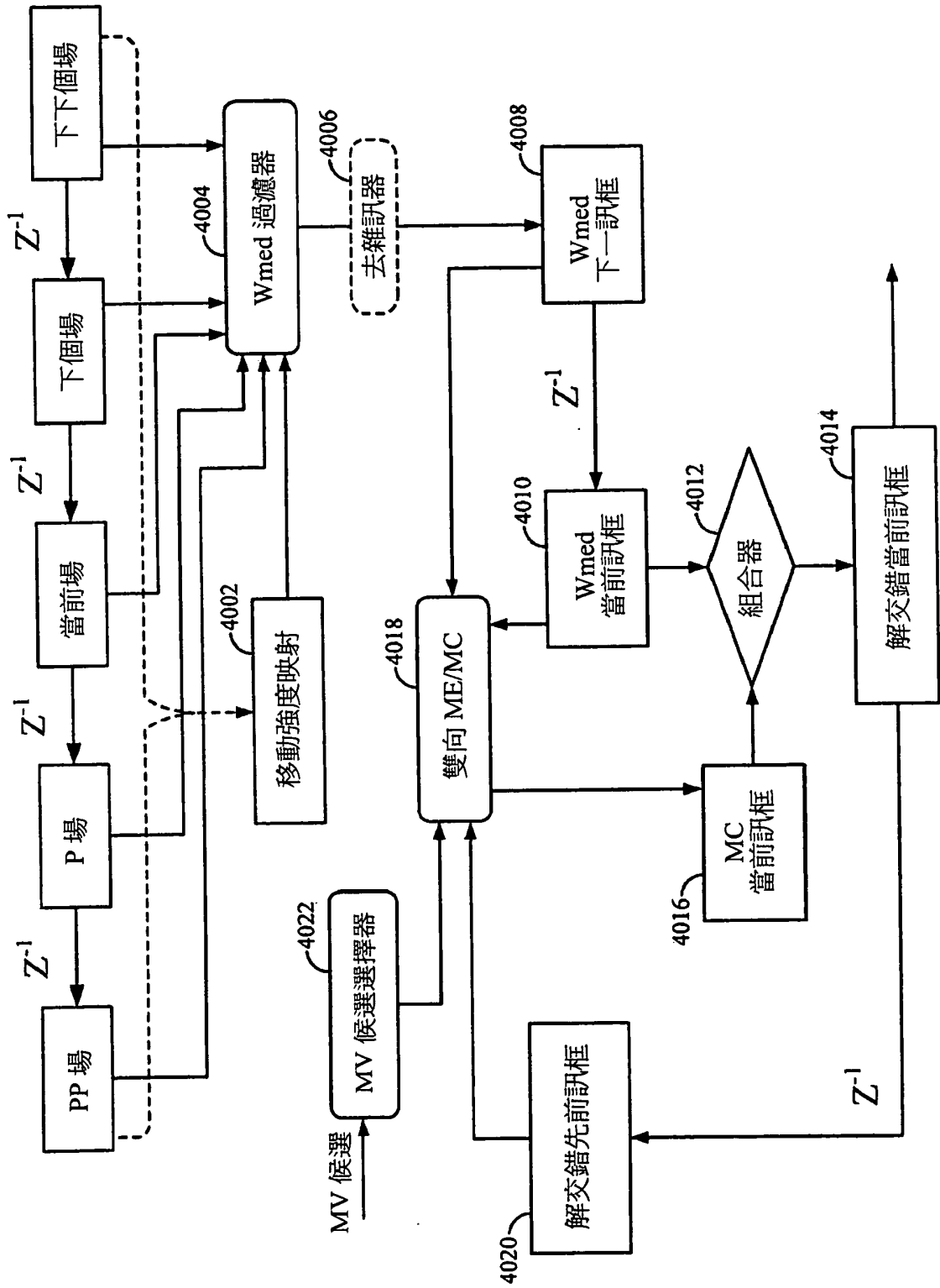
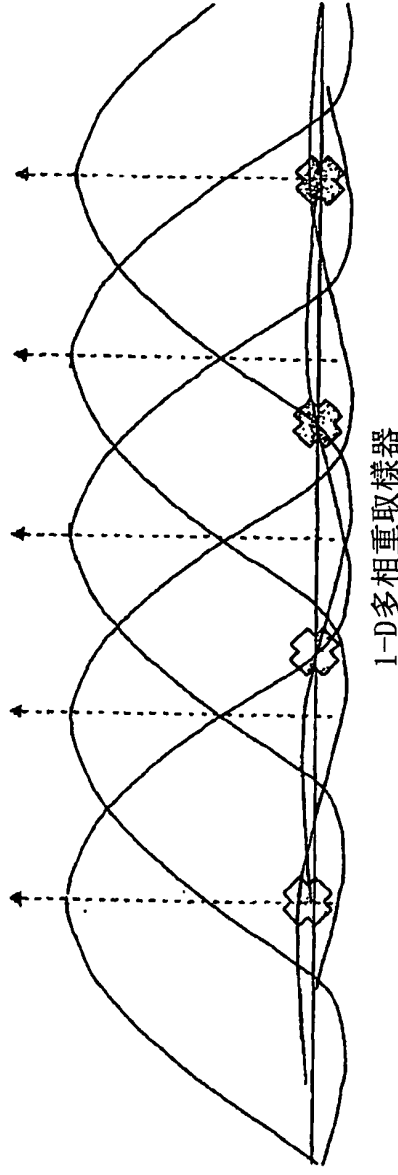


圖40





1-D 多相重取樣器

圖41

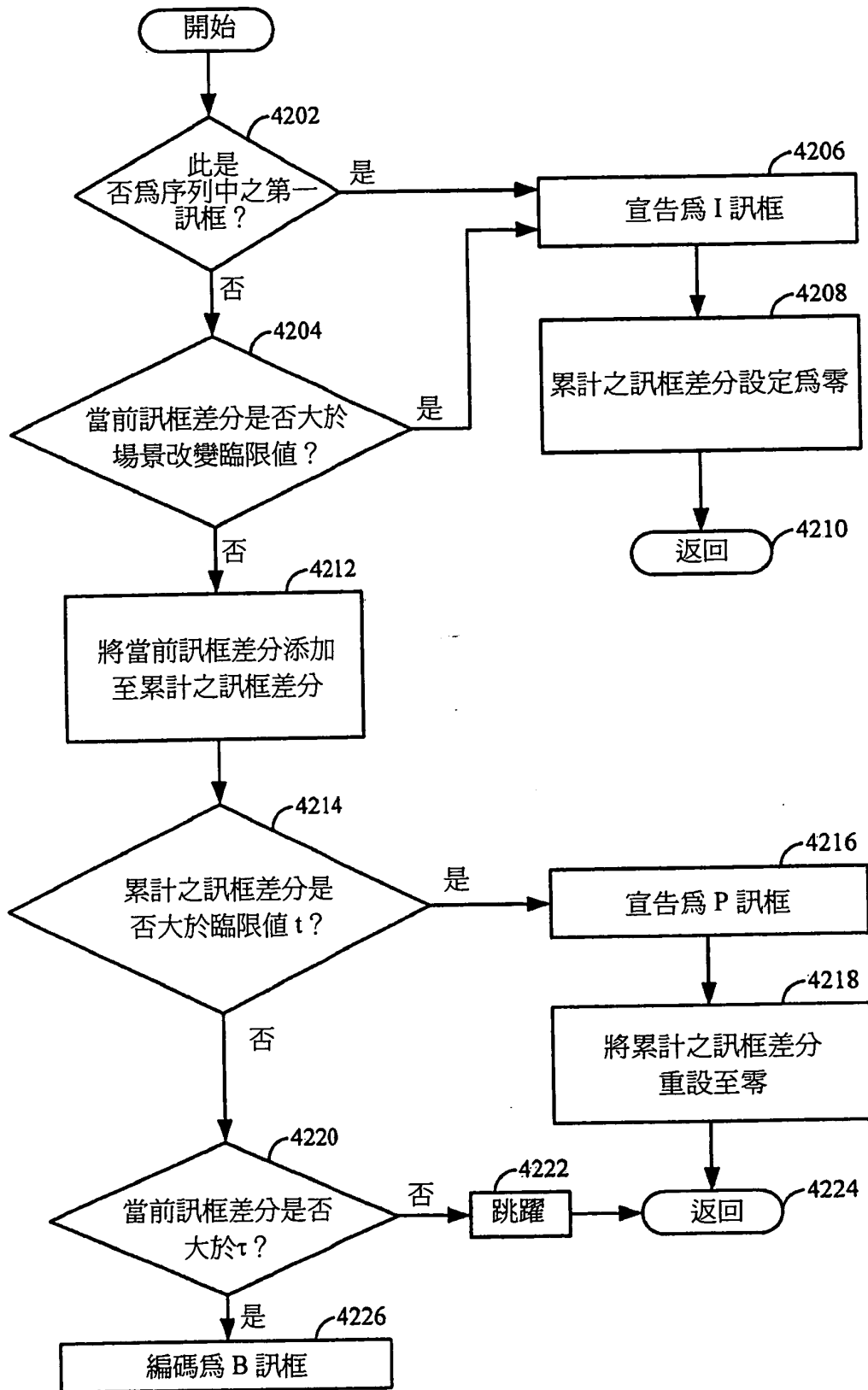


圖42

訊框類型至訊框之自適應指派之流程圖