



(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 201220246 A1

(43)公開日：中華民國 101 (2012) 年 05 月 16 日

(21)申請案號：100133354

(22)申請日：中華民國 100 (2011) 年 09 月 16 日

(51)Int. Cl. : **G06T1/60 (2006.01)**

(30)優先權：2010/10/15 美國

12/905,595

(71)申請人：威盛電子股份有限公司 (中華民國) VIA TECHNOLOGIES, INC. (TW)

新北市新店區中正路 533 號 8 樓

(72)發明人：焦 陽 JIAO, YANG (JEFF) (US)

(74)代理人：洪澄文；顏錦順

申請實體審查：有 申請專利範圍項數：20 項 圖式數：14 共 45 頁

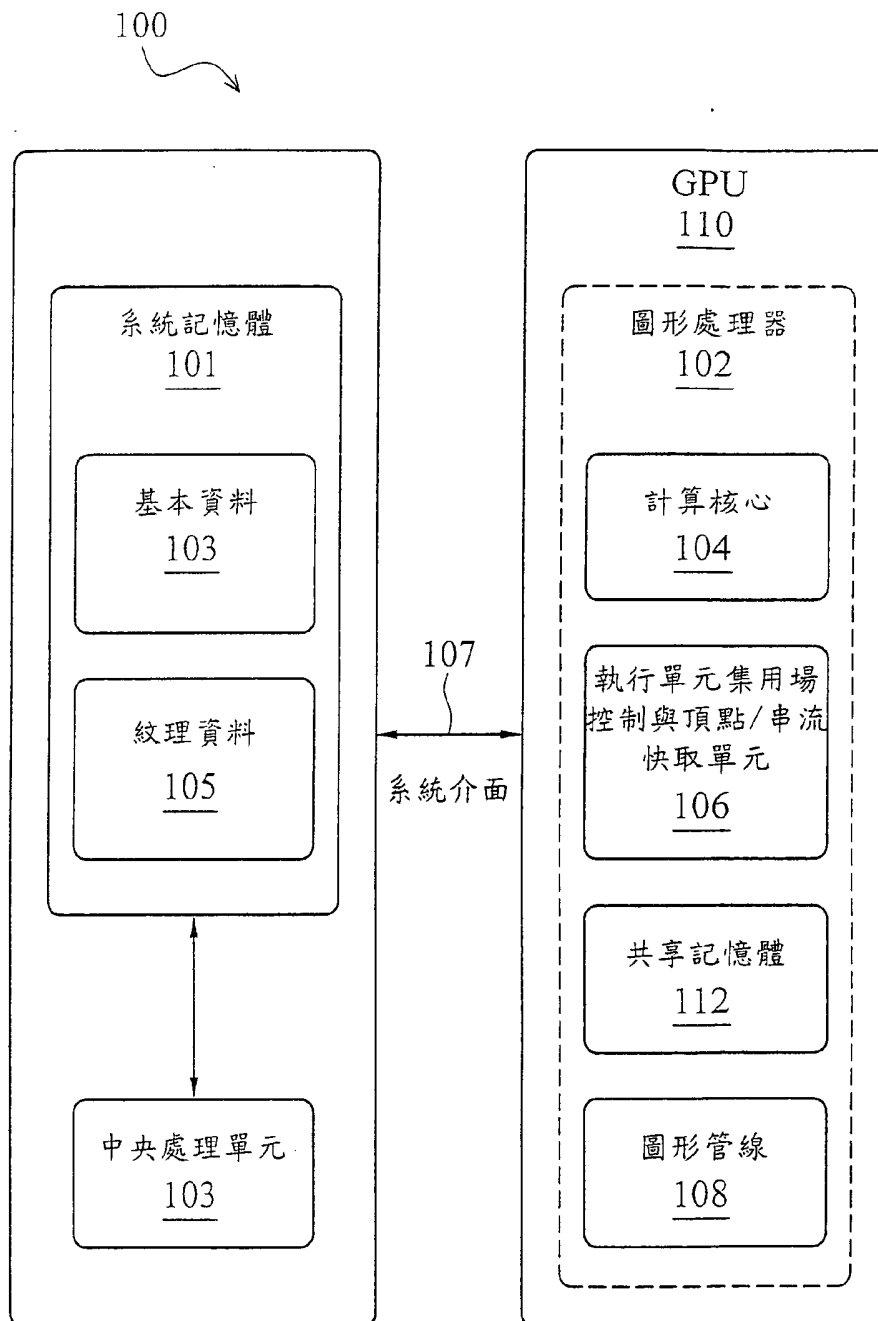
(54)名稱

存取方法以及共享記憶體系統

ACCESSING METHODS AND SHARED MEMORY SYSTEMS USING THE SAME

(57)摘要

本發明提供各種系統與方法來存取圖形處理單元中的共享記憶體。其中一實施例包括判斷將讀取自共享記憶體之資料是否對齊共享記憶體之邊界，其中，資料包括複數資料區塊。判斷即將被安排來處理之資料區塊所採用的拌和型樣。根據資料是否對齊共享記憶體之邊界且根據判斷獲得之拌和型樣，決定執行一或多個環繞功能之順序。藉由執行一或多個環繞功能以及讀取複數資料區塊來存取共享記憶體，以根據拌和型態來建立資料。



- 100：圖形處理器系統
- 101：系統記憶體
- 102：圖形處理器
- 104：計算核心
- 106：執行單元集用場
控制與頂點/串流快取
單元
- 108：圖形管線
- 110：圖形處理單元
(GPU)
- 112：共享記憶體



(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 201220246 A1

(43)公開日：中華民國 101 (2012) 年 05 月 16 日

(21)申請案號：100133354

(22)申請日：中華民國 100 (2011) 年 09 月 16 日

(51)Int. Cl. : **G06T1/60 (2006.01)**

(30)優先權：2010/10/15 美國

12/905,595

(71)申請人：威盛電子股份有限公司 (中華民國) VIA TECHNOLOGIES, INC. (TW)

新北市新店區中正路 533 號 8 樓

(72)發明人：焦 陽 JIAO, YANG (JEFF) (US)

(74)代理人：洪澄文；顏錦順

申請實體審查：有 申請專利範圍項數：20 項 圖式數：14 共 45 頁

(54)名稱

存取方法以及共享記憶體系統

ACCESSING METHODS AND SHARED MEMORY SYSTEMS USING THE SAME

(57)摘要

本發明提供各種系統與方法來存取圖形處理單元中的共享記憶體。其中一實施例包括判斷將讀取自共享記憶體之資料是否對齊共享記憶體之邊界，其中，資料包括複數資料區塊。判斷即將被安排來處理之資料區塊所採用的拌和型樣。根據資料是否對齊共享記憶體之邊界且根據判斷獲得之拌和型樣，決定執行一或多個環繞功能之順序。藉由執行一或多個環繞功能以及讀取複數資料區塊來存取共享記憶體，以根據拌和型態來建立資料。

六、發明說明：

【發明所屬之技術領域】

本發明係有關於一種在圖形處理單元（graphics processing unit，GPU）中的圖形管線(pipeline)，特別是有關於一種在圖形處理單元內執行共享記憶體存取之系統與方法。

【先前技術】

電腦圖形處理系統處理大量的資料，其中包括紋理資料（texture data）。紋理是一個數位影像，通常為矩形，其具有一個(u, v)座標空間。紋理的最小可尋址單元是一紋理影像單元（texel），根據其位置分配到一特定(u, v)座標。在一紋理映射（texture mapping）操作中，當提供一圖形模型以建立一目標影像時，紋理被映射在此圖形模型的表面。在此目標影像中，畫素設置在(x, y)座標系統上之特定座標。紋理映射的目的是在物件的表面上提供逼真的外觀。

MICROSOFT®所提出的 Direct3D 11 規格定義了新類型的著色器/渲染器（shader），即已知的運算著色器（compute shader），其提供多執行緒（multi-thread）支援以幫助圖形進階應用的發展，例如視訊遊戲。運算著色器提供了對於例如在圖形處理器中所使用的單指令多資料（single instruction/multiple data，SIMD）核心之運算功率與頻寬的存取。運算著色器被調用作為多執行緒，其中，這些執行緒被劃分為多個執行緒組。當存在同步化操作以提供執行緒執行的安排時，此同步化手段可以是頻寬與記憶體的加強行為。

【發明內容】

本發明提供一種存取方法，用以存取圖形處理單元中的共享記憶體。此存取方法包括判斷將讀取自共享記憶體之資料是否對齊共享記憶體之邊界，其中，資料包括複數資料區塊。此存取方法更包括判斷即將被安排來處理之資料區塊所採用的拌和型樣。根據資料是否對齊共享記憶體之邊界且根據判斷獲得之拌和型樣，決定執行一或多個環繞功能之順序。此存取方法還包括藉由執行一或多個環繞功能以及讀取複數資料區塊來存取共享記憶體，以根據拌和型態來建立資料。

本發明也提供一種共享記憶體系統。此共享記憶體系統包括共享記憶體與記憶體控制器。共享記憶體其耦接運算著色器以及在圖形處理單元（graphics processing unit，GPU）管線中的複數元件。記憶體控制器判斷關於共享記憶體而將被存取之資料是否對齊共享記憶體之邊界，其中，此資料包括複數區塊。記憶體控制器包括拌和模組與記憶體存取模組。拌和模組判斷即將被安排來處理之複數資料區塊所採用的型樣。記憶體存取模組則根據型態來存取共享記憶體，其中，記憶體存取模組使用一或多個環繞操作來存取該共享記憶體。

本發明另提供一種存取方法，用以存取圖形處理單元中之共享記憶體。此存取方法包括判斷在共享記憶體中將被處理之資料的型樣，其中，資料被劃分成複數存取槽、複數通道、以及複數跨步。根據此型樣，決定執行複數存

取槽環繞、複數通道環繞、以及複數跨步環繞中一者或多個之順序。此存取方法更包括根據此順序來執行複數存取槽環繞、複數通道環繞、以及複數跨步環繞中一者或多個。

【實施方式】

為使本發明之上述目的、特徵和優點能更明顯易懂，下文特舉一較佳實施例，並配合所附圖式，作詳細說明如下。

習知設備在存取圖形管線上之共享記憶體方面的已知缺點是，當典型地逐個擷取資料的區塊以建立需要的區塊時，存取非對齊之資料可能需要多個記憶體存取週期。在本發明中將敘述存取共享記憶體以在圖形管線中實行執行緒的系統與方法，其提供了在存取與運用可變之資料尺寸的適應性，而減少所需的記憶體週期數量，其中此可變之資料尺寸是跨越儲存於圖形處理單元（graphics processing unit，GPU）的共享記憶體內之多個通道(lane)。共享記憶體系統之示範實施例係配置給在執行緒實行期間的記憶體讀取/寫入操作，且包括劃分成複數記憶庫與偏移量的一隨機存取記憶體（RAM）。

在執行緒實行時，所述之共享記憶體系統可在如同執行緒啟動及執行緒同步的操作期間中被使用。在視訊解碼期間，例如資料可能需要根據特定型樣來被讀取且配置，使得可能需要拌和(swizzling)操作。此外，非對齊的資料區塊可能被擷取，其偏移量與記憶庫邊界或是配置在共享記憶體內的畫素區塊沒有對齊。為了說明，參閱第 7 圖，其顯示一個 4x4 畫素區塊的 2 維陣列。在一些情況下，將被

存取之資料區塊 702 可能與多個 4x4 畫素區塊部份重疊，因此其被描述成一非對齊資料區塊，這是因為資料區塊 702 沒有完全地位於一特定 4x4 畫素區塊中。

在第 7 圖之非限制的例子中，非對齊的 4x4 畫素區塊 702 可能透過一系列的儲存槽、通道、以及跨步環繞操作來存取，這些操作於後文說明。藉由使用此處所述之各種共享記憶體存取手段，可存取不同尺寸的非對齊資料區塊。例如，一個 16x1 資料區塊可被存取，且跨越多個通道的 1x16 區塊也可被存取。透過使用所述之記憶體存取手段，可用有效的方法來執行不同的尺寸資料的讀取及寫入。

參閱第 1 圖，其表示根據本發明實施例之一圖形平台(platform)之功能方塊圖，於其中實施根據本發明實施之一共享記憶體系統。在一些實施中，圖形處理系統 100 可配置為一電腦系統。圖形處理器-系統 100 包括圖形處理單元 (GPU) 110，GPU 110 更包括圖形處理器 102。圖形處理器 102 也包括一多執行單元 (multiple execution unit (EU)) 之計算核心 104，計算核心 104 可以為一軟體可編程核心處理器，且在一實施例中，計算核心 104 包括一個一般目的 (general propose, GP) 之著色器。

圖形處理器 102 包括一執行單元集用場 (execution unit pool, EUP) 控制與頂點/串流 (vertex/stream) 快取單元 106 (在此稱為執行單元集用場控制單元 106) 以及一圖形管線 108。計算核心 104 包括複數執行單元之集用池，以滿足與不同著色器程式相關聯之著色器任務所利用之計算需求，包括頂點著色器 (vertex shader)、幾何著色器

(geometry shader)、以及/或畫素著色器(pixel shader)，用來處理給圖形管線 108 之資料。圖形處理器 102 更包括共享記憶體 112，其與由 GP 著色器對執行緒的執行有關，GP 著色器將於文後描述。

第 2 圖係表示第 1 圖中圖形管線 108 之各種元件。具體而言，第 2 圖所表示之元件包括圖形管線 108 之元件，其全部或部分元件與 GP 著色器 214 配合使用。圖形管線 108 包括輸入組譯器(assembler)202，其使用索引緩衝區來提取或一般地接收來自系統記憶體 101 之頂點資料，並組譯基元(primitive)以進一步在計算核心 104 中處理。需注意，系統記憶體 101 可包括其他記憶體次系統中的本地記憶體(local memory)，例如串流及頂點快取列(cache)。經過組譯的基元被傳送至頂點著色器 204。

頂點著色器 204 藉由執行例如轉換(transformation)、癒合(skinning)、與打亮(lighting)來處理頂點。圖形管線 108 更包括外殼著色器(hull shader) 205、鑲嵌單元(tessellator) 206、以及域著色器(domain shader) 207。一般而言，這些著色器 205、207、及 207 係利用來增加特定表面之平滑度。簡單來說，外殼著色器 205 是針對曲線控制點的選擇，且是每一片面(patch)所需要的，其使用來自頂點著色器 204 之片面控制點來作為輸入。在其他的功能中，外殼著色器 205 計算鑲嵌係數，其被傳送至鑲嵌單元 206。這提供了適合的鑲嵌，可用於連續的視點相關精確度等級(view-dependent level of detail (LOD))。來自域著色器 207 之資料被傳送至幾何著色器 208。

幾何著色器 208 接收關於一全基元之複數頂點以作為輸入，且能輸出多個頂點以形成單一拓撲(topology)，例如三角片帶(triangle strip)、直線片帶(line strip)、以及點列表(point list)等等。幾何著色器 208 執行各種功能(包括陰影錐產生(shadow volume generation))，並輸出資訊至光柵處理器(rasterizer) 210。光柵處理器 210 執行剪輯與基元安排，並判斷何時以及/或如何去調用畫素著色器 212。畫素著色器 212 調用被光柵處理器 210 輸出之基元所包含的每一畫素。GP 著色器 214 執行例如模糊效果之後端處理，且通常被調用作為複數執行緒，而這些執行緒被劃分為複數執行緒群組。最後，表示在一場景中物體之表面(或體積)的幾何被轉譯為儲存在訊框緩衝器之畫素(圖片元素)，接著顯示裝置 216 上顯示。

參閱第 3A 圖，共享記憶體 112 可利用來儲存擷取自系統記憶體 101 之資料以在圖形管線 108 中處理，且同樣地，可自共享記憶體 112 讀取資料並將讀取獲得之資料儲存回系統記憶體 101。在一些實施例中，此共享記憶體系統可實施作為 GP 著色器 214 所使用之一執行單元(EU)共享記憶體緩衝器。定義記憶體存取的指令可由一執行單元來調用。一執行緒處理路徑包括共通暫存器檔案(common register file, CRF) 302 以及一執行單元(EU)資料路徑。EU 資料路徑包括複數算數邏輯單元(arithmetic logic unit, ALU)(未顯示)。如第 3A 圖所示，共享記憶體 112 可包括複數記憶體通道(lane)，也稱為在水平方向上的 RAM 記憶庫。為了說明，假設共享記憶體 112 包括總共 16

條的通道，編號由 0 至 15。共享記憶體 112 也包括在垂直方向上的複數跨步（stride）或 RAM 偏移量（offset）。

一特定執行單元（EU）指令可被調用以指示一記憶體存取，而此記憶體存取包括對共享記憶體 112 內一特定位址的讀取以及/或寫入操作。由此指令所指示之讀取/寫入位址指向共享記憶體內通道為基礎的純量資料（lane-based scalar data）的一起始位置。在一些實施例中，此位址可包括一 RAM 偏移量組件以及一 RAM 記憶體庫組件。舉例來說，考慮一位址 0b0110-0011。依據此特定的實施，此位址可關於共享記憶體內等於 6 的 RAM 偏移量（可稱為 RAM 偏移量 6）與等於 3 的記憶體庫數值（可稱為 RAM 記憶體庫 3）。對於接續的資料，記憶體根據以 1 增加的通道偏移來移動。此外，對於舉例所示的共享記憶體 112 架構，根據被調用之特定記憶體存取指令，RAM 記憶體庫朝著到達記憶體庫 15 而環繞，而 RAM 偏移量可維持不變。各種使用環繞操作的記憶體存取架構在後文說明。

參閱第 3B 圖之示意圖以及在第 4 圖之流程圖，共享記憶體系統可包括記憶體控制器 304，用以控制對共享記憶體 112 的存取以及執行資料流控制。記憶體控制器 304 用來判斷將讀取自共享記憶體 112 之資料或將寫入至共享記憶體 112 之資料是否對齊（align）共享記憶體 112 之一特定畫素區塊或邊界（第 4 圖之步驟方塊 402）。在一些實施例中，共享記憶體 112 之邊界包括由共享記憶體 112 本身的尺寸所定義的邊界以及/或由共享記憶體 112 內複數畫素區塊所定義之邊界。記憶體控制器 304 更包括拌和

(swizzle) 模組 306，用來判斷在即將安排來處理之複數資料區塊所採用的拌和型樣 (swizzle pattern) (步驟方塊 404)。根據資料如何對齊共享記憶體 112 之邊界，此拌和模組 306 更用來判斷執行一或多個環繞功能之順序 (步驟方塊 406)。環繞功能包括一存取槽環繞功能 (slot warp function)、一通道環繞功能 (lane warp function)，以及一跨步環繞功能 (stride warp function)，而執行的記憶體存取可包括任何這些功能的結合。

記憶體控制器 304 更包括記憶體存取模組 308，其執行對應共享記憶體 112 之實際讀取/寫入操作。記憶體存取模組 308 藉由執行一或多個環繞功能以及讀取資料區塊來存取共享記憶體 112，以根據拌和模組 306 決定之拌和型樣來建立資料 (步驟方塊 408)。

第 5 圖係表示根據存取共享記憶體 112 之另一實施例。尤其是第 5 圖是表示存取在圖形處理單元 (GPU) 中一共享隨機存取記憶體 (RAM) 之資料的方法。開始於步驟方塊 502，判斷在共享記憶體 112 中將被處理之資料的型樣，其中，此資料被劃分成複數處理槽、通道、以及跨步。根據此型樣，在步驟方塊 504 中決定執行一或多個處理槽環繞、通道環繞、以及跨步環繞之順序。在步驟方塊 506 中，根據決定的順序來執行一或多個處理槽環繞、通道環繞、以及跨步環繞。

應注意到，假使在軟體中實現，上述的每一步驟方塊可表示包括程式指令的一個模組、資料段以及部分編碼，以實現特定邏輯功能。此程式指令可用源編碼之型態來實

現，而此源編碼包括由程式語言來編寫之敘述；或者是此程式指令可用機器編碼來實現，而此機器編碼包括可由一適當執行系統（例如在電腦系統或其他如第 1 圖所示之系統內之處理器）所識別之數值指令。機器編碼可由源編碼等等轉換獲得。假使在硬體中時間，每一步驟方塊可表示一電路或者是一些互相連接之電路，以實踐實現特定邏輯功能。

第 6 圖係表示將來自系統記憶體 101 之一對齊之 4×4 畫素區塊載入至共享記憶體 112。對於對齊複數 4×4 畫素方塊中之一者的資料而言，資料可由系統記憶體 101 載入至共享記憶體 112，而不需使用環繞操作。如在一非限制的例子中，系統記憶體 101 包括複數 4×4 畫素區塊之 2 維陣列。應注意到，非只可使用 4×4 區塊，資料而也可使用 8×8 區塊、 12×12 區塊、或 16×16 區塊而由系統記憶體 101 載入至共享記憶體 112。對於對齊一 4×4 區塊之複數資料區塊而言，不需要環繞操作來將資料由系統記憶體 101 轉移至共享記憶體 112。

第 7 圖係表示使用存取槽環繞、通道環繞、以及跨步環繞之組合將來自系統記憶體 101 之一非對齊之 4×4 畫素區塊載入至共享記憶體 112。如圖所示，系統記憶體 101 可包括複數 4×4 畫素方塊，以 $Q[0,0]$ 、 $Q[1,0]$ 、 $Q[2,0]$ 、 $Q[3,0]$ 、 $Q[0,1]$ 、 $Q[1,1]$ 、 $Q[2,1]$ 、及 $Q[3,1]$ 來標示。在一非限制的例子中，將擷取自系統記憶體 101 以及寫入至共享記憶體 112 之非對齊資料不會符合這些 4×4 畫素方塊中之一者。關於這一點，資料可能以在一非對齊位置上的一 4×4

區塊來讀取自或寫入至共享記憶體 112。對於一些實施而言，左上方畫素之位址被指定，另外，對於在 4x4 畫素區塊中的一 4-位元偏移量，其中，2 個最大有效位元表示一存取槽環繞數值，而 2 個最小有效位元表示一通道環繞數值。舉例來說，開始於在共享記憶體 112 中基礎位址 A、數值等於 1 的存取槽環繞、數值等於 2 的通道環繞、以及數值等於 4 的跨步環繞之畫素區塊 Q[0,0]，在 4x4 區塊中的資料根據一特定順序來存取。

如此技術領域中具有通常知識者將瞭解，能在系統記憶體 101 與共享記憶體 112 之間轉移儲存於非對齊記憶體位置之資料對於各種圖形處理操作而言是有幫助的。藉由 LUMA 之視訊解碼訊框間預測 (inter-prediction)，例如，資料載入至共享記憶體 112 且被處理，且 GP 著色器 214 將參考資料載入至畫素區塊。舉例來說，這可包括 4x4 或 8x8 區塊。為了詳細說明，假設參考資料包括一 8x8 畫素區塊，根據移動向量 (motion vector) 這需要一參考資料之 13x13 區塊。

如前所述並參閱第 7 圖，一”非對齊”區塊是與跨越或環繞共享記憶體邊界之資料相關聯，而”對齊”之複數區塊是指符合在共享記憶體中一或多個畫素區塊之鄰近區域的區塊。在一些例子中，執行單元將對齊之資料載入記憶體，且 4x4 畫素區塊可越過一 RAM 進入點而被載入至 16 個記憶體庫。在載入操作期間可實施一存取槽環繞以及/或通道環繞操作，以確保任何的 8x8 區塊、12x12 區塊、或 16x16 區空可在行方向與列方向上被存取。然而，對於 4x4 區塊

為基礎的處理，一般不需要實施載入/儲存存取槽/通道環繞。

為了進一步說明，考慮對在一畫素區塊（例如 4x4 或 8x8）中每一區塊的圖形操作，二分之一或四分之一取樣是根據鄰近畫素來計算獲得。為了讀取在 8x8 或 16x16 畫素區塊中行方向與列方向上的資料，共享記憶體透過一系列的資料旋轉與資料移動來存取。此外，之後共享記憶體之讀取/寫入可使用各種共享記憶體讀取/寫入指令，來處理在行方向與列方向上的資料。關於此點，可載入來自共享記憶體中的裝置記憶體（例如 4x4、8x8、12x12、或 16x16 區塊）之資料。

以下將更詳細地說明已敘述基本的架構以實施共享記憶體系統，執行共享記憶體存取之方法的各種例子。尤其是，將說明包含牽涉第 3A 圖之共享記憶體 112 之一系列存取槽、通道、以及跨步環繞功能的各種讀取/寫入操作。應注意到，對於下文所述之共享記憶體操作，根據一特定型樣來利用一或多個環繞操作以讀取或寫入資料。對於一些實施例而言，環繞操作包括下列中的一或多個：存取槽環繞、通道環繞、以及跨步環繞。根據由 RAM 偏移量數值、存取槽環繞數值、通道環繞數值、以及跨步環繞數值所指定之一特定型樣，資料可讀取自共享記憶體或寫入至共享記憶體。

回來參閱第 3A 圖，對於示範的實施例，組成共享記憶體 112 之複數 RAM 記憶庫分組為複數存取槽，例如，使得一獲得的存取槽包括四個 RAM 記憶庫，且一共享記憶體操

作之起始點由 RAM 偏移量與一指令所提供之位址來決定。根據此，將被存取之下一資料係由通道環繞數值、存取槽環繞數值、以及跨步環繞數值來決定。尤其是，執行一或多個存取槽環繞、通道環繞、以及跨步環繞之程序可由指明通道環繞數值、存取槽環繞數值、以及跨步環繞數值之一或多個指令來開始實施。關於此點，關於以一圖形管線中不同之元件來存取共享記憶體 112 的適應性可藉由提供將被指定之不同存取模式來達成。這可由根據處理所需之資料型樣對環繞的順序進行調整來實現。現在將參閱第 8-14 圖，其說明利用存取槽環繞、通道環繞、以及跨步環繞以存取第 1 圖中共享記憶體的各種實施例。

連續讀取/寫入操作

一共享記憶體存取技術係針對有關於共享記憶體 112 的連續讀取/寫入操作。由 EU 指令所提供之讀取/寫入操作指向在共享記憶體 112 內通道為基礎的純量資料的起始位置。對於連續的共享記憶體讀取操作而言，自跨越一有效位址之記憶體位址的範圍讀取內容至一較高範圍，其中，此共享記憶體位址一般是純量數值。

參閱第 8 圖，被讀取之資料包括複數個資料區塊，且儲存在共享記憶體 112 內的連續記憶體位置。對於一些實施來說，連續的讀取操作可被一指令所調用，藉以使得運算元（operand）對應上述有效位址以及較高範圍。假使資料將被寫入至共通暫存器檔案（CRF）302，基底位址設置在由運算元 Rs1 之最小有效通道數來的第一個有效通道。在一些實施例中，此位址係以下式來計算：執行緒群組基

底位址 (thread group base address) + Rs1[0] + 目前位址 $\ll 4$ + 記憶體庫。由 EU 指令所提供讀取/寫入位址指向共享記憶體 112 內通道為基礎的純量資料的起始位置。

如一非限制之實施例，考慮位址 0b0110-0011 的情況，其指向共享記憶體 112 內 RAM 偏移量 6 以及記憶體庫 3。對於後續的每一 32-位元純量資料，位址 (包括 RAM 偏移量以及通道偏移量，此兩者都可是整數數值) 將以 1 來增加。如第 8 圖所示之非限制的例子，來自 RAM 偏移量 6 及記憶體庫 3 至 15 的資料以及來自 RAM 偏移量 7 及記憶體庫 0 至 2 的資料將被取出，且寫入至目標通道 (DST) 0 至 15。為了表明牽涉讀取/寫入連續操作，此位址被提供為源位址 1 (暫存通道 0, Rs.0)，且加上目前 (IMM) 範圍以產生最後位址 (ADDR)。

讀取/寫入環繞操作

參閱第 9 圖，另一共享記憶體存取技術係針對有關於讀取/寫入環繞操作，其中，RAM 偏移量維持不變。關於此實施例，將被讀取或寫入之資料配置成跨越共享記憶體 112 中的數個 RAM 記憶體庫或通道。讀取/寫入位址由一 EU 指令所指示，且指向共享記憶體 112 內通道為基礎的純量資料的起始位置。舉例來說，位址 0b0110-0011 指向共享記憶體 112 中的 RAM 偏移量 6 與記憶體庫 3。對於後續的每一 32-位元純量資料，通道或記憶體庫偏移量以 1 來增加，且朝著第 16 個 RAM 記憶體庫而環繞。對於此讀取/寫入操作，RAM 偏移量維持不變。在第 9 圖所示之例子中，32-位元資料的 16 個記憶體庫跨越 RAM 偏移量 6 與記憶體庫 3 至 15，

且環繞至記憶體庫 0 至 2。在這些共享記憶體位址之資料被取出，且寫入目標通道 (DST) 0 至 15。在一些實施例中，此位址被提供為源位址 1(暫存通道 0, Rs.0)，且加上 (IMM) 範圍以產生最後位址 (ADDR)。

讀取/寫入存取槽環繞操作

參閱第 10 圖，其表示讀取/寫入存取槽環繞操作。對於讀取寫入存取槽環繞操作，資料根據出於一存取槽觀點及一通道觀點的環繞操作而讀取自或寫入至共享記憶體。尤其是，在一已知存取槽中，執行通道環繞。例如，對於包括 4 個記憶體庫之存取槽 1002，一通道環繞包括在一已知存取槽中在到達第 4 個通道後而環繞回第 1 個通道。例如，涉及一特定存取槽（例如存取槽 2）的一通道環繞可包括存取通道 3 及 4 且接著存取 1 及 2。同樣地，一存取槽環繞包括在到達在共享記憶體 112 中的最後一個存取槽後而環繞回至第一個存取槽。參閱第 10 圖，其表示 4 個存取槽。在到達第 4 個存取槽 1002 後而向左移動，接著存取第 1 個存取槽。一跨步環繞包括在到達共享記憶體 112 中的最大 RAM 偏移量後而環繞回第一個跨步或者 RAM 偏移量。根據一實施，寫入位址係由一 EU 指令以運算元型態來指示，且指向共享記憶體 112 中的 RAM 偏移量。

在此實施例中，實施兩個拌和檔案以提供一個 4-通道存取槽為基礎的循環位移量（2 位元，0~3）以及在存取槽循環位移量（2 位元，0~3）中的通道，以提供能指定關於資料之特定拌和型樣的適應性。假設 RAM 偏移量數值等於 6，等於 1 的存取槽環繞數值與通道環繞數值導致數

值”S”寫入至第 10 圖所示的位置。因此，對於此共享記憶體操作，資料的環繞具有兩個觀點，一個是關於存取槽環繞，而另一個是關於通道環繞。所導致的型樣也如圖所示。

讀取/寫入環繞存取槽增量操作

讀取/寫入環繞存取槽操作的一種變化是涉及一增量操作，在此稱為讀取/寫入環繞存取槽增量操作，參閱第 11 圖，由 EU 指令所提供之讀取/寫入位址指向共享記憶體內之 RAM 偏移量。如上述讀取/寫入環繞存取槽操作，包含兩個額外的拌和檔案以提供一個 4-通道存取槽為基礎的循環位移量（2 位元，0~3）以及在存取槽循環位移量（2 位元，0~3）中的通道。此外，也包含一跨步數值，以允許在 RAM 偏移量方向上存取共享記憶體 112。在所示之一非限制的例子，指定等於 1 的存取槽環繞數值、等於 2 的通道環繞數值、以及等於 2 的跨步環繞數值。即是，對於在共享記憶體 112 中每一純量數值而言，RAM 偏移係以跨步環繞數值所指的量來增加。

讀取環繞存取槽環繞增量操作

參閱第 12 圖，另一操作為伴隨滑動增量的讀取環繞/存取槽環繞操作，其調用來存取共享記憶體 112。關於此操作，由一 EU 指令提供讀取操作，其指向在共享記憶體中的一 RAM 偏移量。同樣地，包含兩個額外的拌和檔案以提供一個 4-通道存取槽為基礎的循環位移量（2 位元，0~3）以及在在一已知的存取槽（2 位元，0~3）中的通道為基礎的循環位移量。在此操作中，跨步環繞數值指示一 RAM 偏移量以進行往返改變。參閱第 1 圖之共享記憶體

112，為了詳細說明，假設共享記憶體 112 以一 2D 陣列來製作。

跨步數值係用來提供在共享記憶體 112 之垂直或 Y 方向上的資料讀取或寫入。應注意到此操作提供對資料之任何非對齊區塊的存取。例如，當執行像移動補償的這種處理時，則可涉及存取非對齊區塊。關於此點，一資料區塊（例如不符合在共享記憶體 112 內一畫素區塊之一連續部分的 4x4 區塊）可有效地讀取或寫入至共享記憶體而不需以一連續方式來逐個處理每一資料區塊。透過存取槽環繞操作、通道環繞操作、以及跨步環繞操作的使用，可存取在共享記憶體內的資料。如所示之非限制的例子，其指明等於 1 的存取槽環繞數值、等於 2 的通道環繞數值、以及等於 4 的跨步環繞數值。

當自共享記憶體 112 讀取資料且入至一共通暫存器檔案（common register file，CRF）302 時，可能需要額外的拌和來進一步配置資料以在共通暫存器檔案 302 中處理。關於這點，一些實施例更包括第 3B 圖之一輸出拌和模組 310，其支援三種拌和模式：垂直模式、水平模式、水平拌和模式。對於如同視訊解碼轉換的操作，可能必須在相鄰的四個 4x4 區塊上執行 4x4 或 8x8 轉移。第 13-14 圖說明由輸出拌和模組 310 所執行的水平操作以及水平拌和操作。

如圖所示，資料由系統記憶體 101 載入至共享記憶體 112，以作為包括對準之 4x4 資料區塊的 1D 陣列。資料可藉由配置輸出拌和模組 310 以在水平模組下操作而讀取自

共享記憶體 112 以作為一對準的 4x4 區塊。此資料接著寫入至共通暫存器檔案 302。關於此點，輸出拌和模組 310 允許任何四個 4x1 或 1x4 區塊被擷取且寫入至共通暫存器檔案 302 而不會發生衝突。同樣地，任何兩個 8x1 或 1x8 區塊可被擷取而不會發生衝突。參閱第 14 圖，輸出拌和模組 310 可支援一水平拌和模式，藉此使得讀取自系統記憶體 101 且儲存至共享記憶體 112 之資料在寫入至共通暫存器檔案 302 之前經歷了一水平拌和。

這裡所述之任何元件包括軟體或編碼，且可收錄於由一指令執行系統所使用或與一指令執行系統所連接的任何電腦可讀媒介，例如在一電腦系統或任何系統內的處理器。在本說明書的內容中，一電腦可讀取媒介可以是任何實體媒體，其可包含、儲存、或維持由一指令執行系統所使用或與一指令執行系統所連接的軟體或編碼。例如，一電腦可讀媒體可儲存由一處理裝置（例如第 1 圖之圖形處理器 102）所執行的一或多個程式。

例如但並非限制，電腦可讀媒體可以是電子、磁性、光學、或半導體系統、設備、或裝置。電腦可讀媒體之更具體例子可包括下列：具有一或多導線的電子連接、可攜式電腦磁盤、隨機隨取記憶體（RAM）、唯讀記憶體（ROM）、可抹除可編程唯讀記憶體（EPROM、EEPROM、或快閃記憶體）、以及可攜式光盤唯讀記憶體（CDROM）。

本發明雖以較佳實施例揭露如上，然其並非用以限定本發明的範圍，任何所屬技術領域中具有通常知識者，在不脫離本發明之精神和範圍內，當可做些許的更動與潤

飾，因此本發明之保護範圍當視後附之申請專利範圍所界定者為準。

【圖式簡單說明】

第 1 圖表示根據本發明實施例之圖形系統，於其中實施共享記憶體；

第 2 圖係表示第 1 圖之圖形管線之各種元件；

第 3A 及 3B 圖表示第 1 圖之共享記憶體之各種元件；

第 4 圖表示根據本發明一實施例，對應第 1 圖之共享記憶體而執行記憶體存取之方法；

第 5 圖表示根據本發明另一實施例，對應第 1 圖之共享記憶體而執行記憶體存取之方法；

第 6 圖表示將一對準的 4x4 畫素區塊由系統記憶體載入共享記憶體之示意圖；

第 7 圖表示使用存取槽環繞、通道環繞、以及跨步環繞來將一非對準的 4x4 畫素區塊由系統記憶體載入共享記憶體之示意圖；以及

第 8-14 圖表示利用存取槽環繞、通道環繞、以及跨步環繞來存取第 1 圖之共享記憶體之各種實施例。

【主要元件符號說明】

第 1 圖：

100～圖形處理器系統； 101～系統記憶體；

102～圖形處理器； 104～計算核心；

106～執行單元集用場控制與頂點/串流快取單元；

108～圖形管線；

110～圖形處理單元（GPU）；

112～共享記憶體；

第 2 圖：

202～輸入組譯器； 204～頂點著色器；

205～外殼著色器； 206～鑲嵌單元；

207～域著色器； 208～幾何著色器；

210～光柵處理器； 212～畫素著色器；

214～一般目的（GP）著色器；

216～顯示裝置；

第 3A-3B 圖：

302～共通暫存器檔案（CRF）；

304～記憶體控制器； 306～拌和模組；

308～記憶體存取模組； 310～輸出拌和模組；

第 4 圖：

402、404、406、408～步驟方塊；

第 5 圖：

502、504、506～步驟方塊；

第 6-7 圖：

101～系統記憶體； 112～共享記憶體；

702～資料區塊 702；

第 8-14 圖：

201220246

101～系統記憶體； 112～共享記憶體；

310～輸出拌和模組；。

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：100(2020)000000000

※申請日：100. 9. 16

※IPC 分類：

G06T 1/60 2006.01

一、發明名稱：(中文/英文)

存取方法以及共享記憶體系統/Accessing Methods And Shared Memory Systems Using The Same

二、中文發明摘要：

本發明提供各種系統與方法來存取圖形處理單元中的共享記憶體。其中一實施例包括判斷將讀取自共享記憶體之資料是否對齊共享記憶體之邊界，其中，資料包括複數資料區塊。判斷即將被安排來處理之資料區塊所採用的拌和型樣。根據資料是否對齊共享記憶體之邊界且根據判斷獲得之拌和型樣，決定執行一或多個環繞功能之順序。藉由執行一或多個環繞功能以及讀取複數資料區塊來存取共享記憶體，以根據拌和型態來建立資料。

三、英文發明摘要：

Various systems and methods are described for accessing a shared memory in a graphics processing unit (GPU). One embodiment comprises determining whether data to be read from a shared memory aligns to a boundary of the shared memory, wherein the data comprises a plurality of data blocks, and wherein the shared memory comprises a plurality of banks and a plurality of offsets. A swizzle pattern in

which the data blocks are to be arranged for processing is determined. Based on whether the data aligns with the boundary of the shared memory and based on the determined swizzle pattern, an order for performing one or more wrapping functions is determined. The shared memory is accessed by performing the one or more wrapping functions and reading the data blocks to construct the data according to the swizzle pattern.

七、申請專利範圍：

1.一種存取方法，用以存取一圖形處理單元中之一共享記憶體，該存取方法包括：

判斷將讀取自該共享記憶體之一資料是否對齊該共享記憶體之一邊界，其中，該資料包括複數資料區塊；

判斷即將被安排來處理之該等資料區塊所採用的一拌和型樣；

根據該資料是否對齊該共享記憶體之該邊界且根據判斷獲得之該拌和型樣，決定執行一或多個環繞功能之一順序；以及

藉由執行該或該等環繞功能以及讀取該等資料區塊來存取該共享記憶體，以根據該拌和型態來建立該資料。

2.如申請專利範圍第 1 項所述之存取方法，其中，該共享記憶體包括複數記憶體庫以及複數偏移量。

3.如申請專利範圍第 2 項所述之存取方法，其中，該等記憶體庫被分組成複數存取槽，且該等存取槽根據複數存取槽數值來存取。

4.如申請專利範圍第 3 項所述之存取方法，其中，該等偏移量根據複數跨步數值而往返改變。

5.如申請專利範圍第 3 項所述之存取方法，其中，藉由執行該或該等環繞功能以及讀取該等資料區塊來存取該共享記憶體之步驟包括以下一或多個：

一存取槽環繞；

一通道環繞；以及

一跨步環繞。

6.一種共享記憶體系統，包括：

一共享記憶體，耦接一運算著色器以及在一圖形處理單元（graphics processing unit，GPU）管線中的複數元件；以及

一記憶體控制器，判斷關於該共享記憶體而將被存取之資料是否對齊該共享記憶體之一邊界，其中，且該資料包括複數區塊，該記憶體控制器包括；

一拌和模組，用以判斷即將被安排來處理之該等資料區塊所採用的一型樣；以及

一記憶體存取模組，用以根據該型樣來存取該共享記憶體，其中，該記憶體存取模組使用一或多個環繞操作來存取該共享記憶體。

7.如申請專利範圍第 6 項所述之共享記憶體系統，其中，該共享記憶體之該邊界包括以下之一者：

由該共享記憶體之尺寸所定義之邊界；以及

由該共享記憶體內複數畫素區塊所定義之邊界。

8.如申請專利範圍第 6 項所述之共享記憶體系統，其中，該共享記憶體包括複數水平記憶庫以及複數垂直偏移量。

9.如申請專利範圍第 8 項所述之共享記憶體系統，其中，該等水平記憶庫被劃分為複數存取槽，且每一該存取槽包括一既定數量之該等水平記憶庫。

10.如申請專利範圍第 8 項所述之共享記憶體系統，其中，該等垂直偏移量被劃分成複數跨步。

11.如申請專利範圍第 6 項所述之共享記憶體系統，其

中，該記憶體存取模組藉由以一順序執行該或該等環繞操作來存取該共享記憶體，且該順序係以該資料之該型樣為依據。

12.如申請專利範圍第 6 項所述之共享記憶體系統，其中，該等環繞操作包括：

- 一存取槽環繞；
- 一通道環繞；以及
- 一跨步環繞。

13.如申請專利範圍第 6 項所述之共享記憶體系統，其中，存取該共享記憶體包括讀取自該共享記憶體以及寫入至該共享記憶體中之一者。

14.如申請專利範圍第 6 項所述之共享記憶體系統，其中，該記憶體存取模組更用來將讀取自該共享記憶體之資料寫入至一共通暫存器檔案，且根據將資料寫入至該共通暫存器檔案，該拌和模組對被寫入之資料執行一拌和功能。

15.如申請專利範圍第 14 項所述之共享記憶體系統，其中，該拌和功能包括以下之一者：

- 一水平拌和操作；以及
- 一垂直拌和操作。

16.一種存取方法，用以存取一圖形處理單元中之一共享記憶體，該存取方法包括：

判斷在該共享記憶體中將被處理之一資料的一型樣，其中，該資料被劃分成複數存取槽、複數通道、以及複數跨步；

根據該型樣，決定執行複數存取槽環繞、複數通道環

繞、以及複數跨步環繞中一者或多個之一順序；以及
根據該順序來執行該等存取槽環繞、該等通道環繞、
以及該等跨步環繞中一者或多個。

17.如申請專利範圍第 16 項所述之存取方法，更包括：
判斷在該共享記憶體中被實施將執行之記憶體操作的
該資料是否對齊該共享記憶體之一邊界。

18.如申請專利範圍第 17 項所述之存取方法，其中，反
應於判斷出該資料非對齊該共享記憶體之該邊界，根據該
順序來執行該等存取槽環繞、該等通道環繞、以及該等跨
步環繞中一者或多個。

19.如申請專利範圍第 17 項所述之存取方法，其中，該
共享記憶體之該邊界包括以下一或多個：

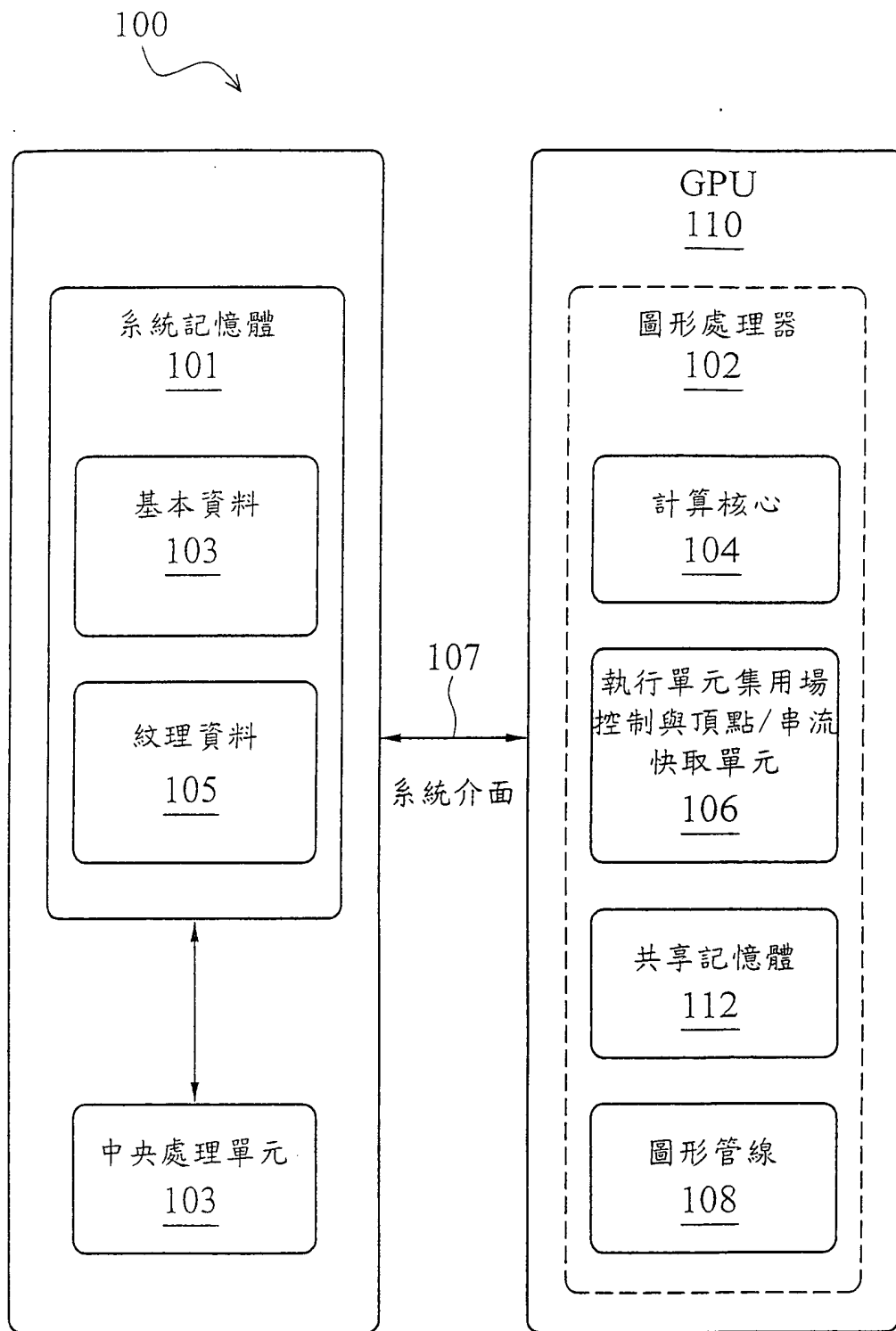
由該共享記憶體之尺寸所定義之邊界；以及

由該共享記憶體內複數畫素區塊所定義之邊界。

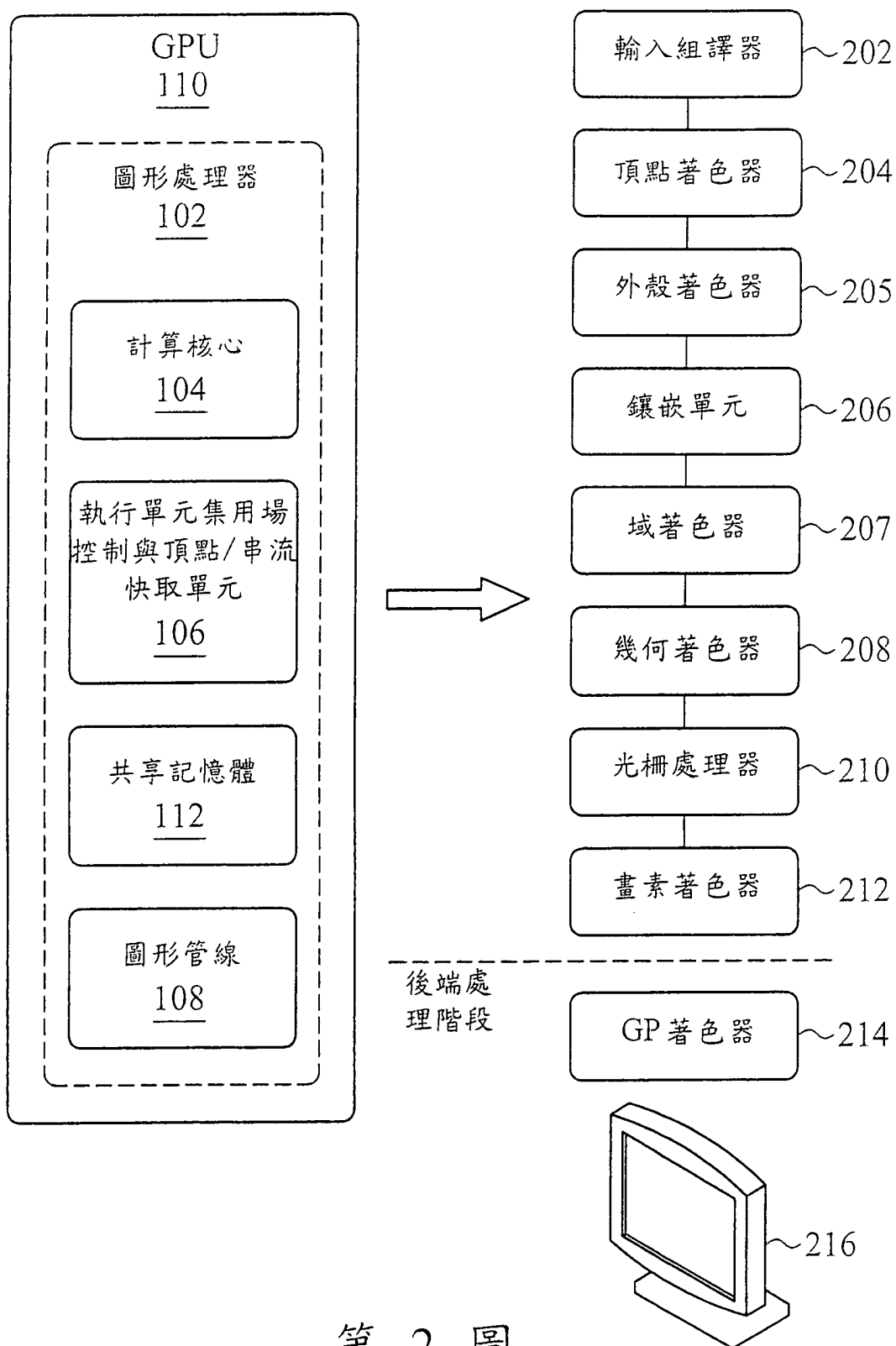
20.如申請專利範圍第 16 項所述之存取方法，其中，根
據該順序來執行該等存取槽環繞、該等通道環繞、以及該
等跨步環繞中一者或多個之步驟，由指定複數存取槽環繞
數值、複數通道環繞數值、以及複數跨步環繞數值的一或
多個指令來開始實施。

201220246

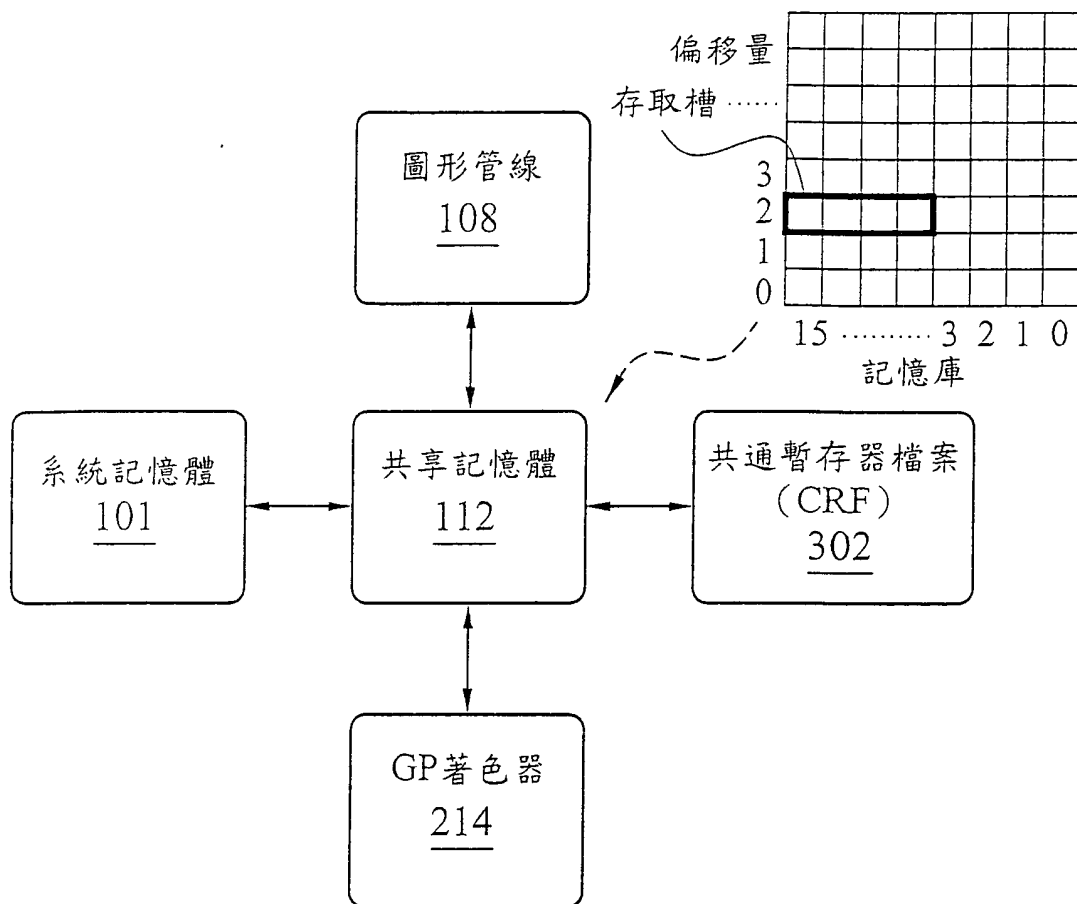
八、圖式：



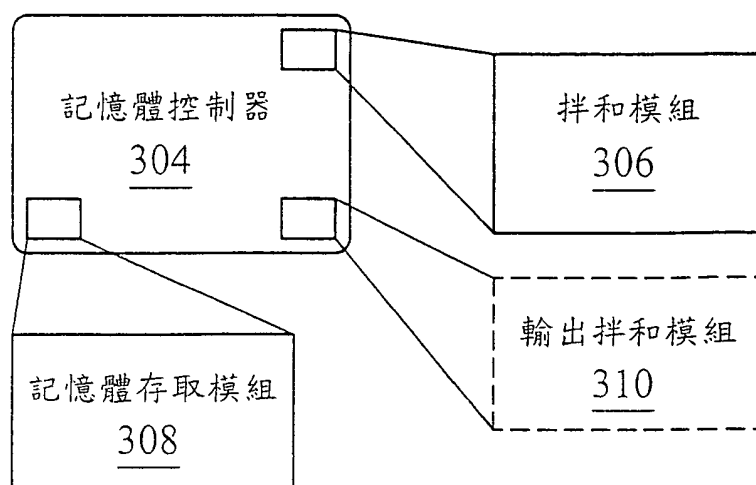
第 1 圖



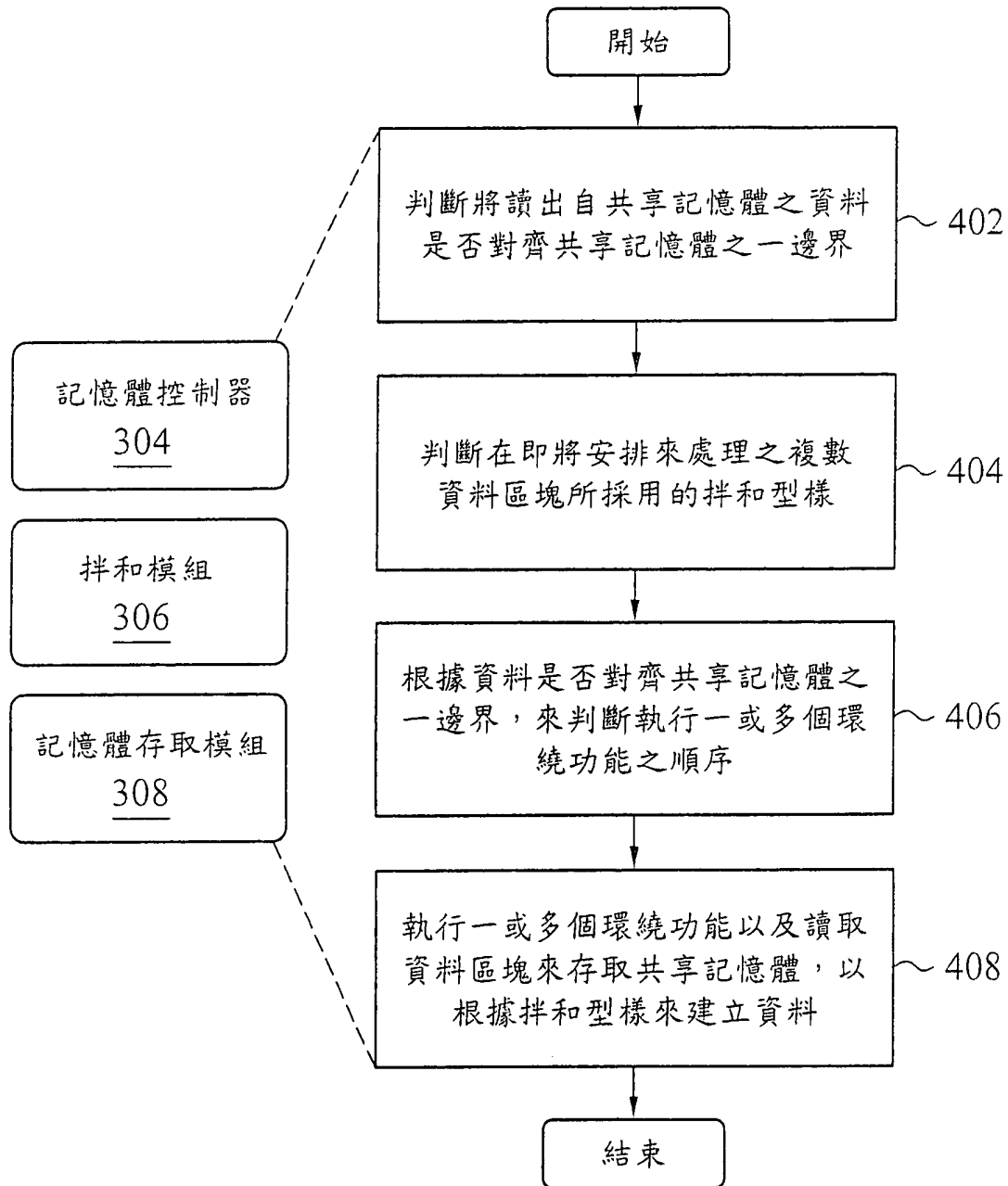
第 2 圖



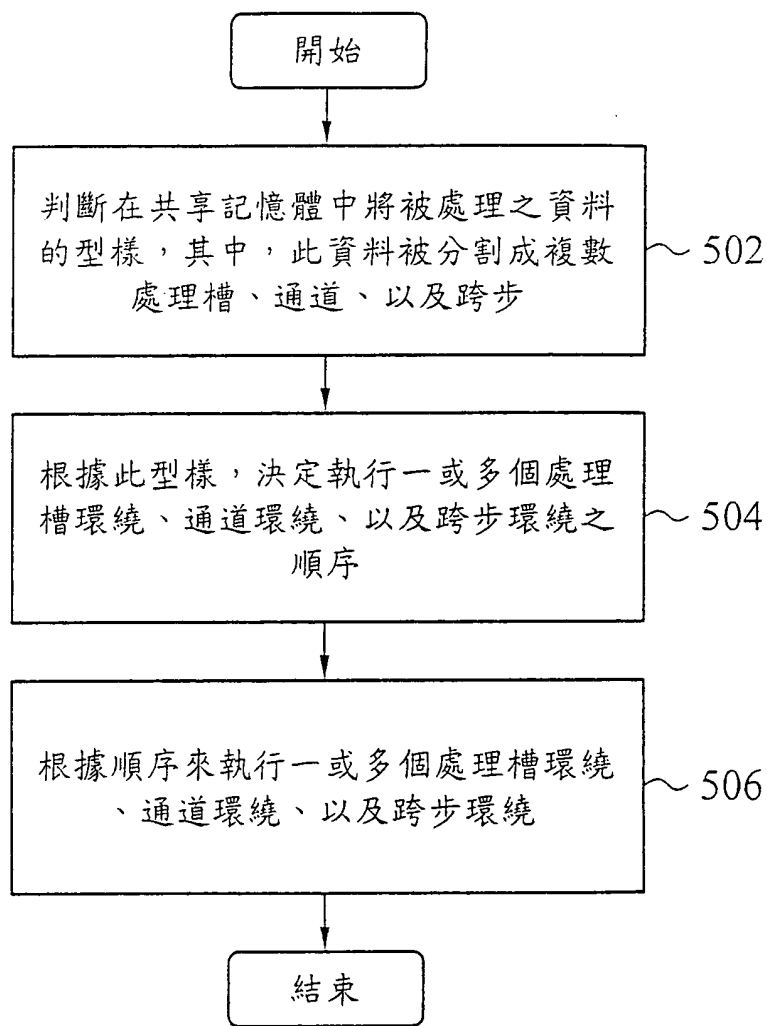
第 3A 圖



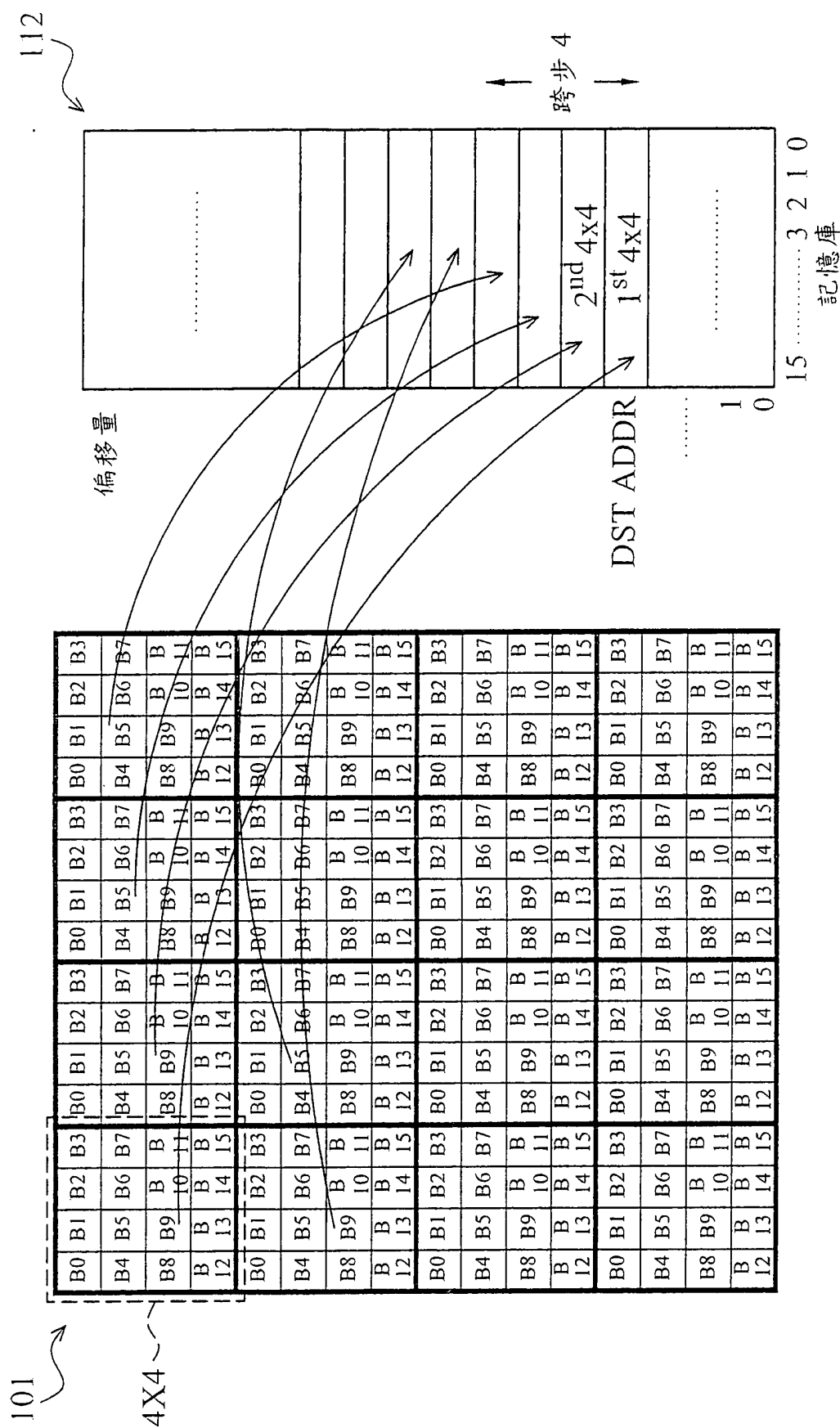
第 3B 圖



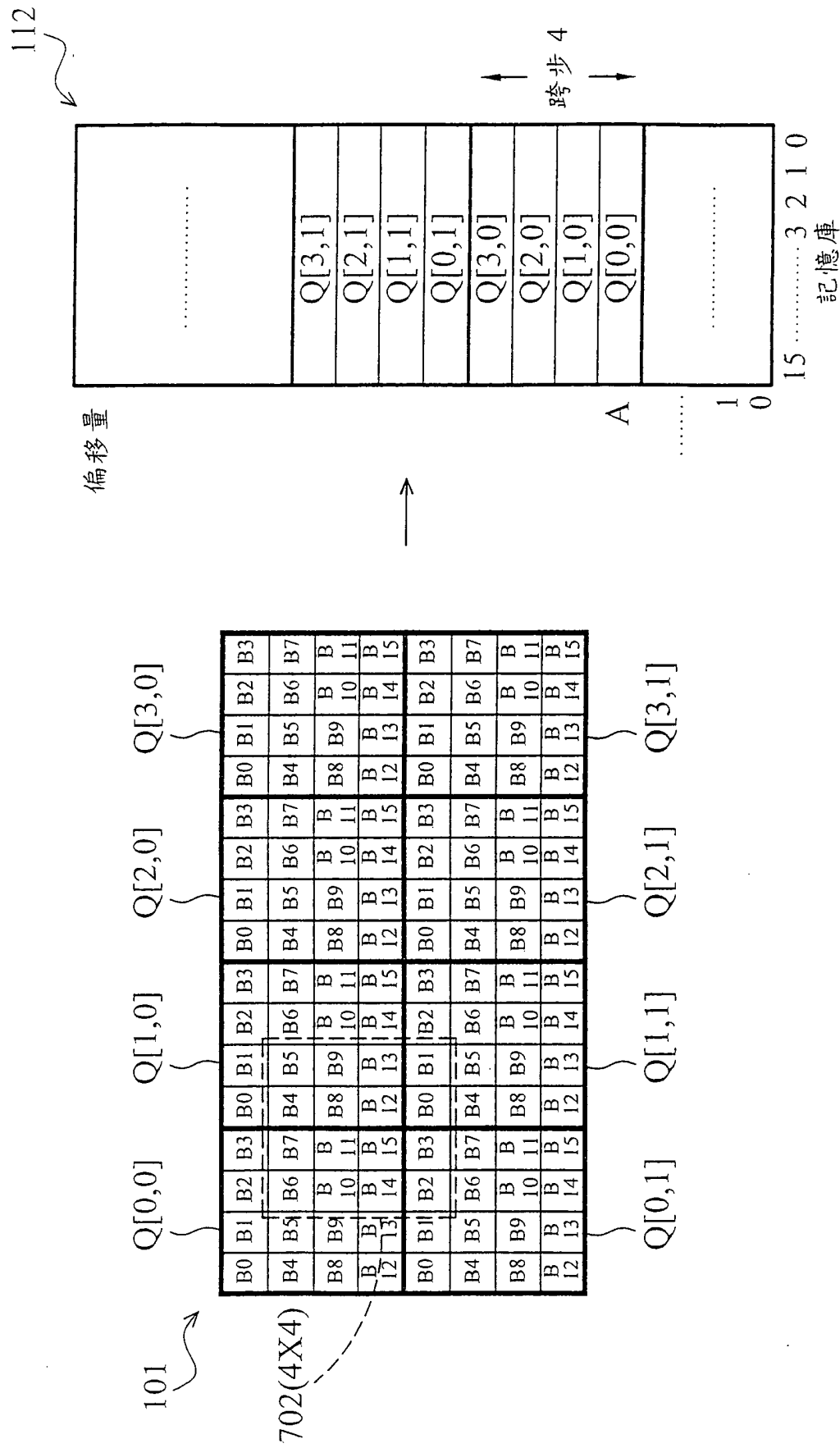
第 4 圖



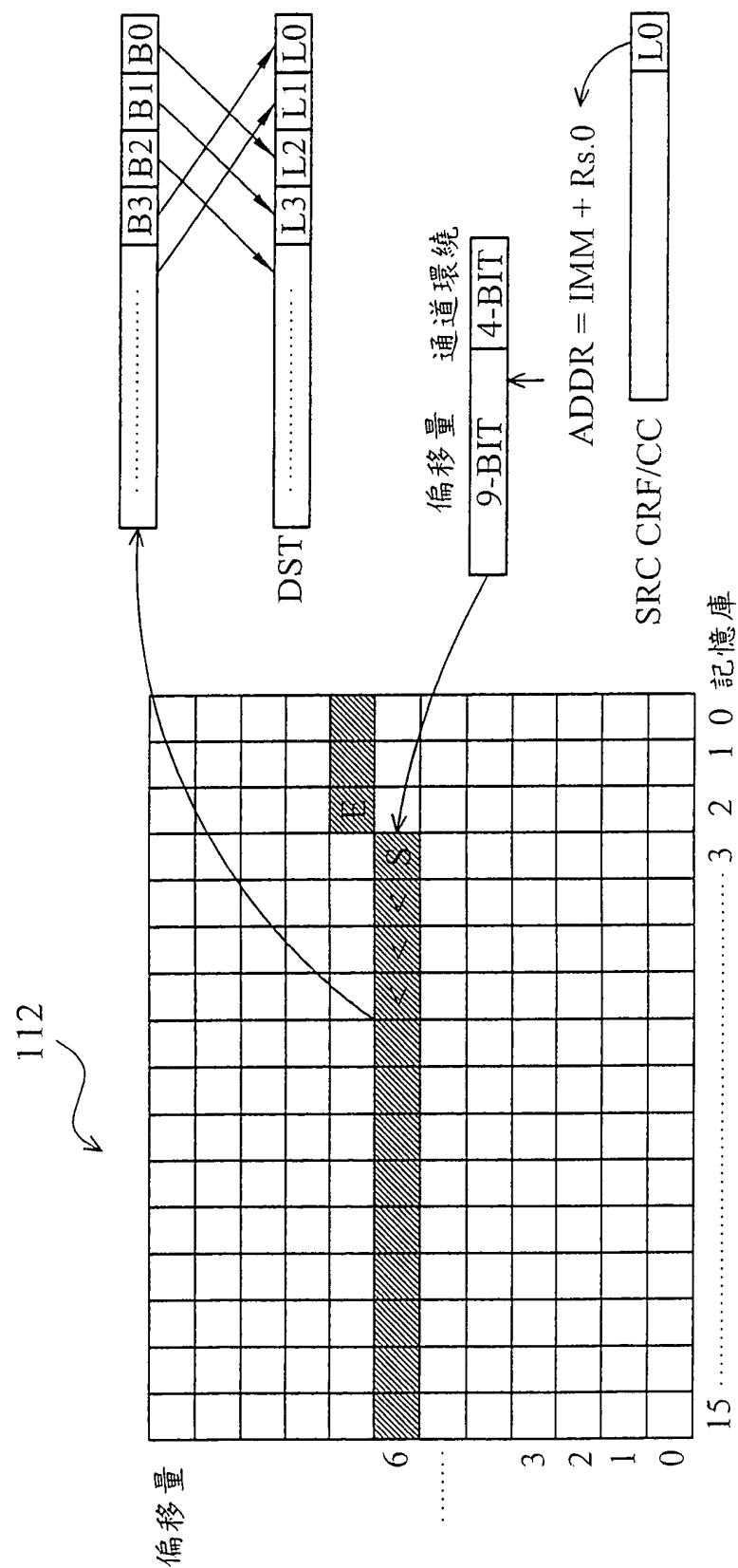
第 5 圖



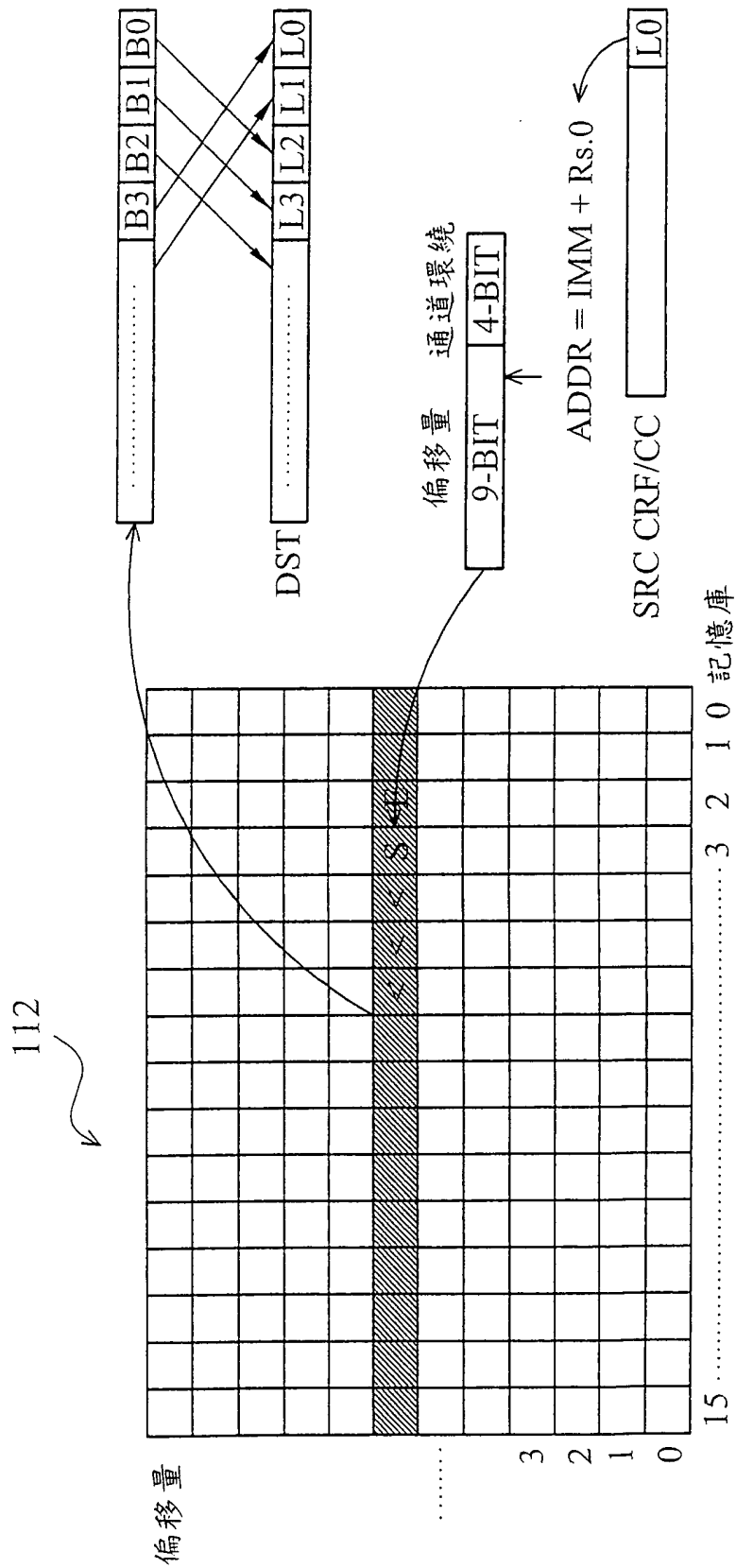
第 6 圖



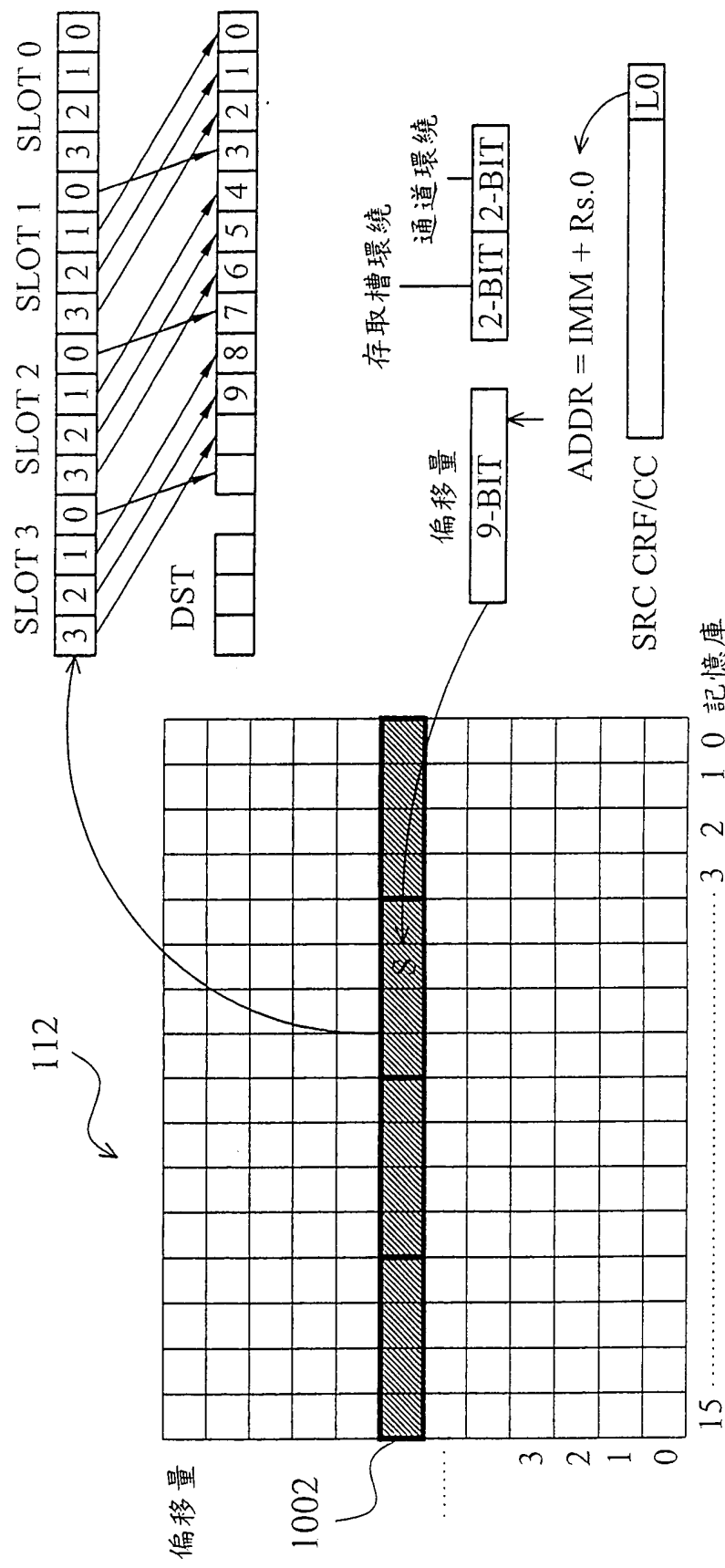
第 7 圖



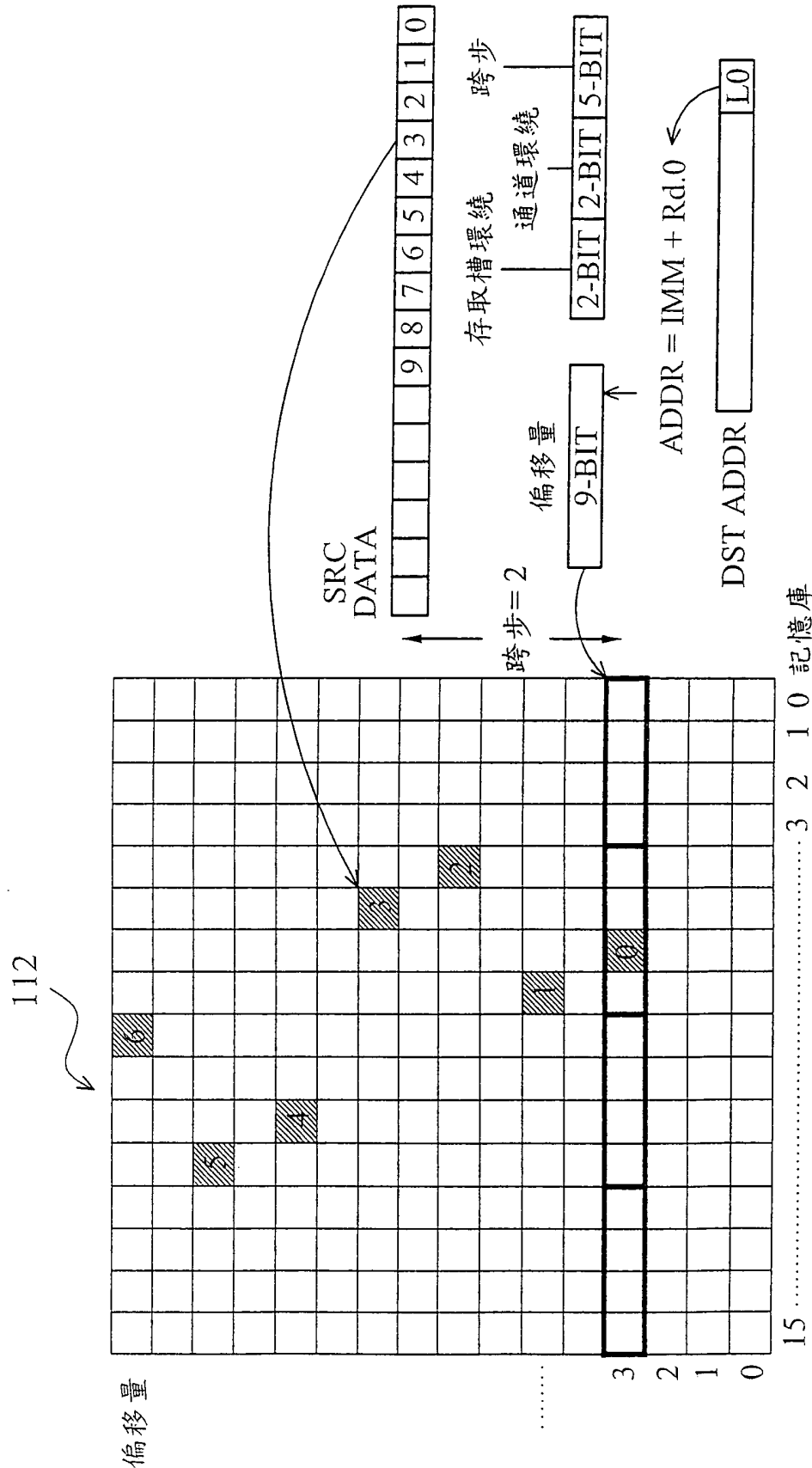
第 8 圖



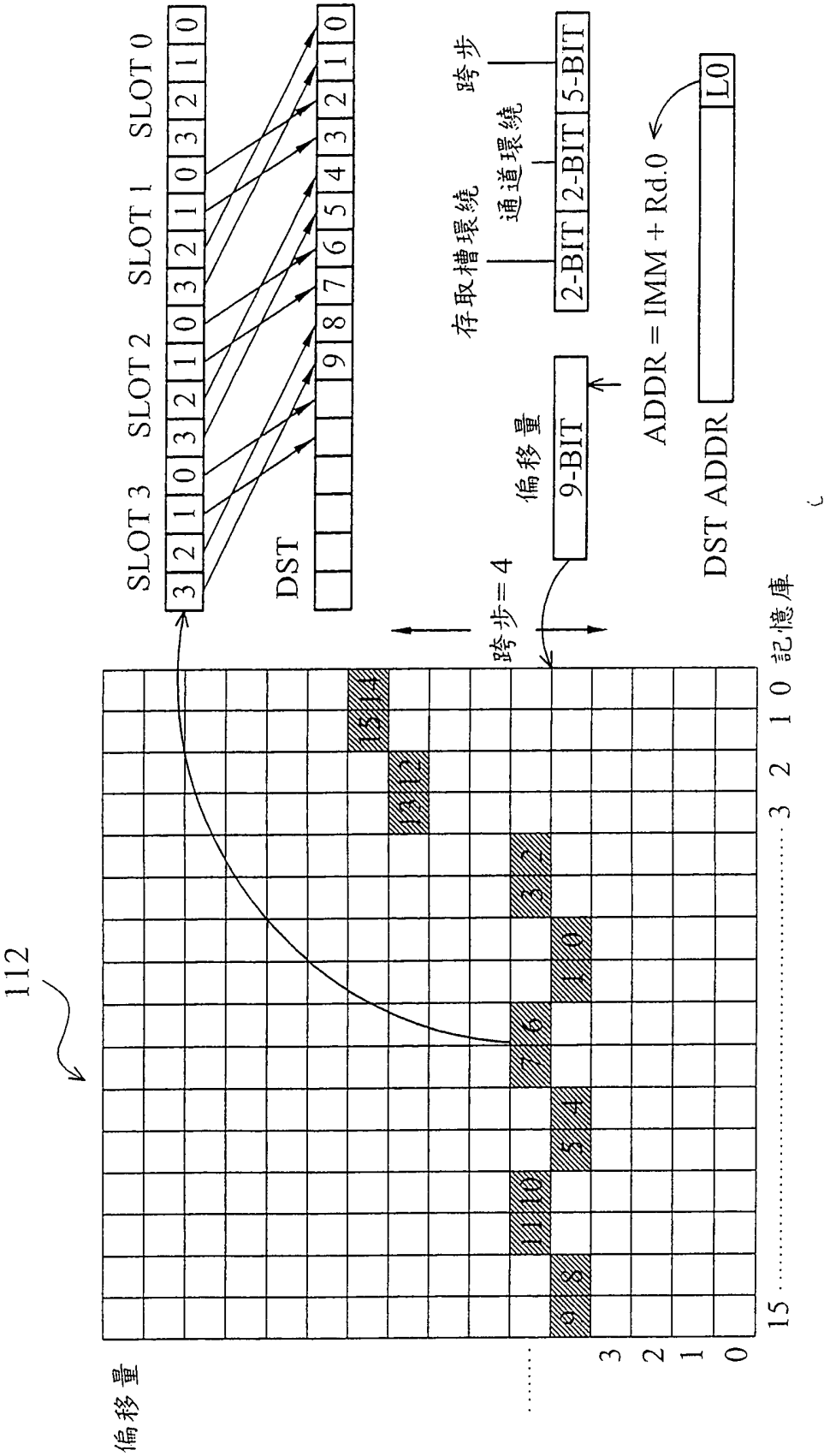
第 9 圖



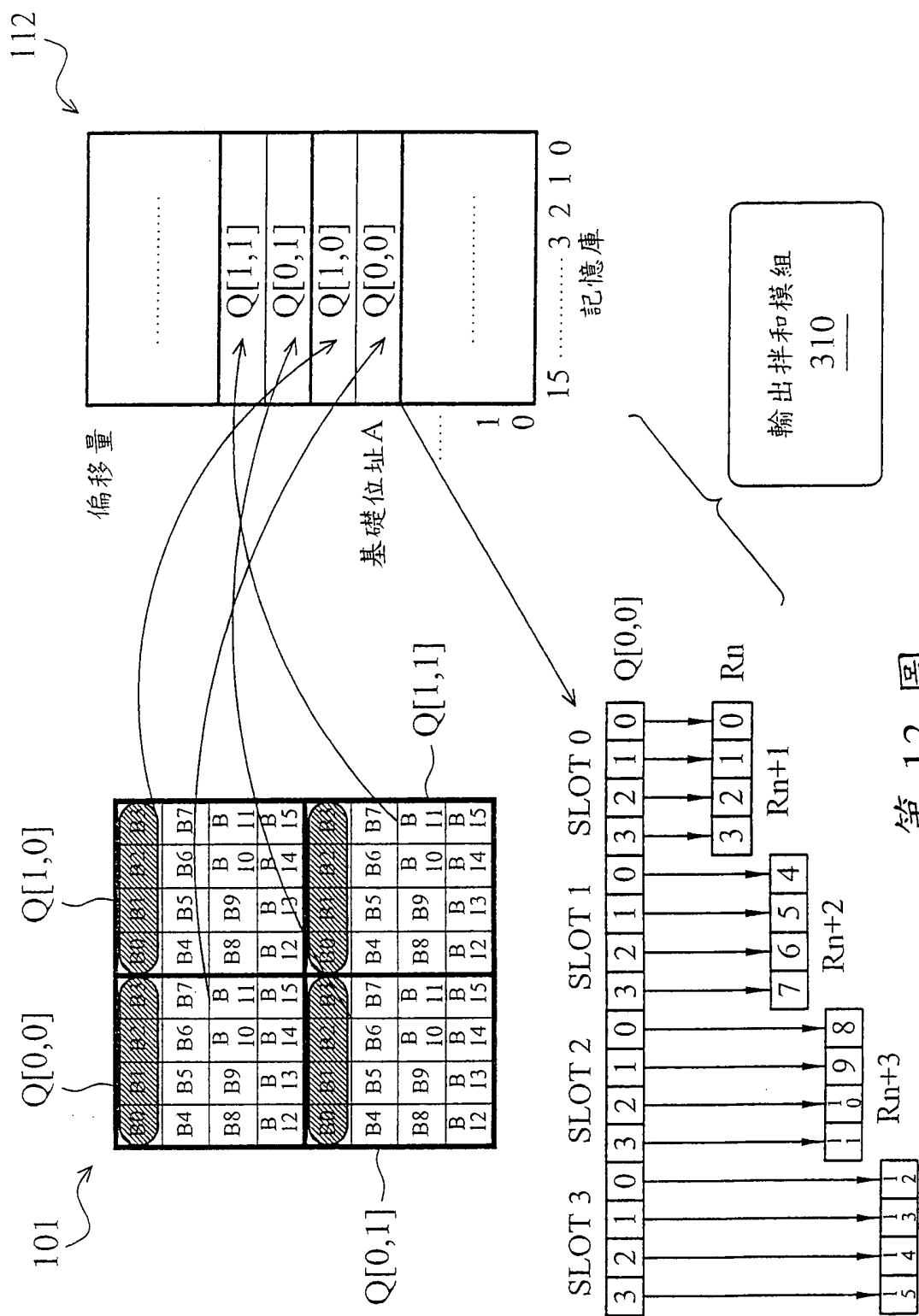
第 10 圖



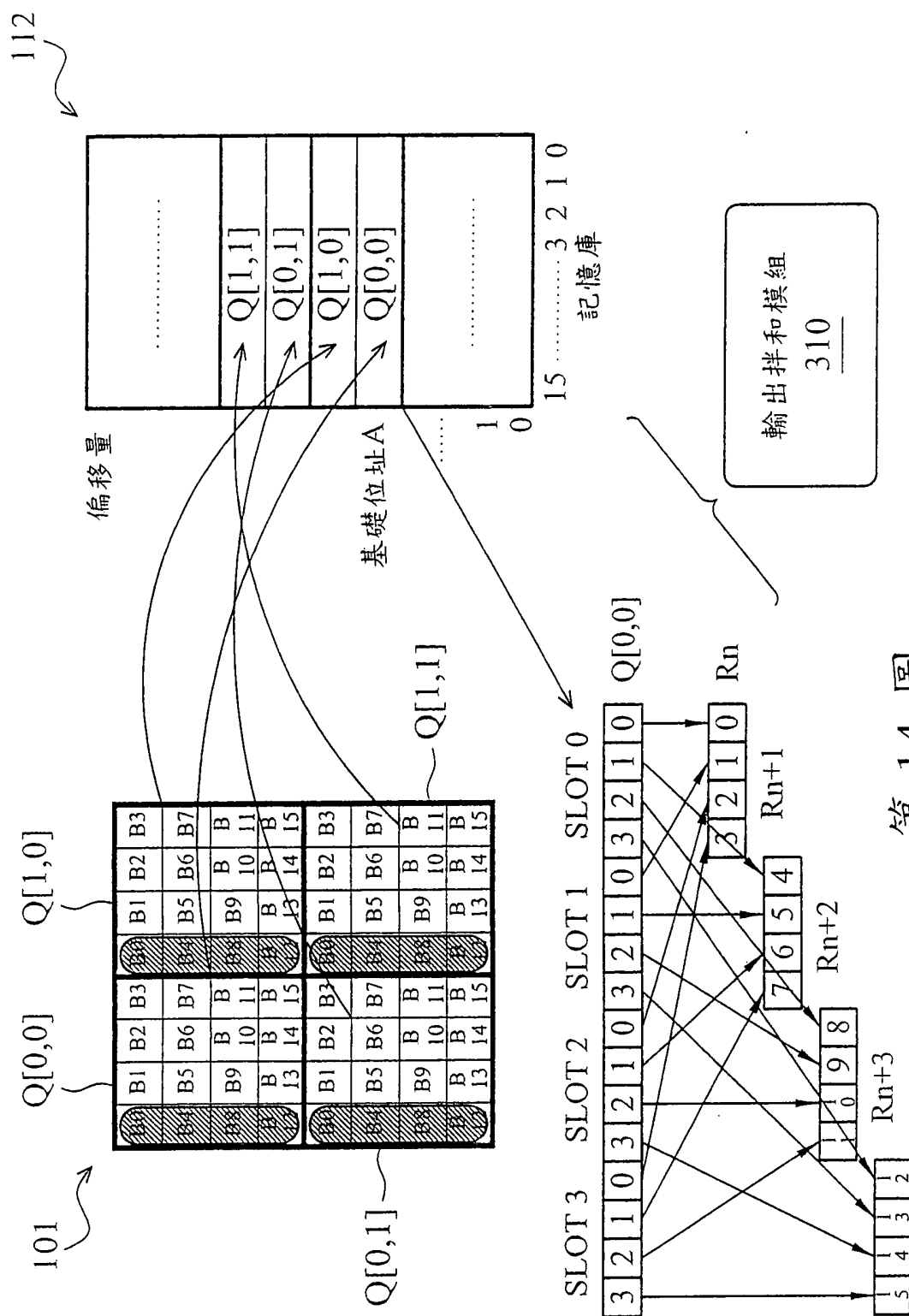
第 11 圖



第 12 圖



第 13 圖



第 14 圖

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：

100～圖形處理器系統；

101～系統記憶體；

102～圖形處理器；

104～計算核心；

106～執行單元集用場控制與頂點/串流快取單元；

108～圖形管線；

110～圖形處理單元（GPU）；

112～共享記憶體。

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

略