



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I699993 B

(45) 公告日：中華民國 109 (2020) 年 07 月 21 日

(21) 申請案號：108109902 (22) 申請日：中華民國 108 (2019) 年 03 月 20 日
 (51) Int. Cl. : *H04N19/167 (2014.01)* *H04N19/176 (2014.01)*
 (30) 優先權：2018/09/30 美國 16/147,849
 (71) 申請人：奇景光電股份有限公司 (中華民國) HIMAX TECHNOLOGIES LIMITED (TW)
 臺南市新市區紫棟路 26 號
 (72) 發明人：張珮鈺 JHANG, PEI YU (TW)；楊得煒 YANG, DER WEI (TW)
 (74) 代理人：蔡濱陽
 (56) 參考文獻：
 CN 105531988A CN 105681795A
 審查人員：張長軾
 申請專利範圍項數：11 項 圖式數：4 共 23 頁

(54) 名稱

辨識感興趣的區域

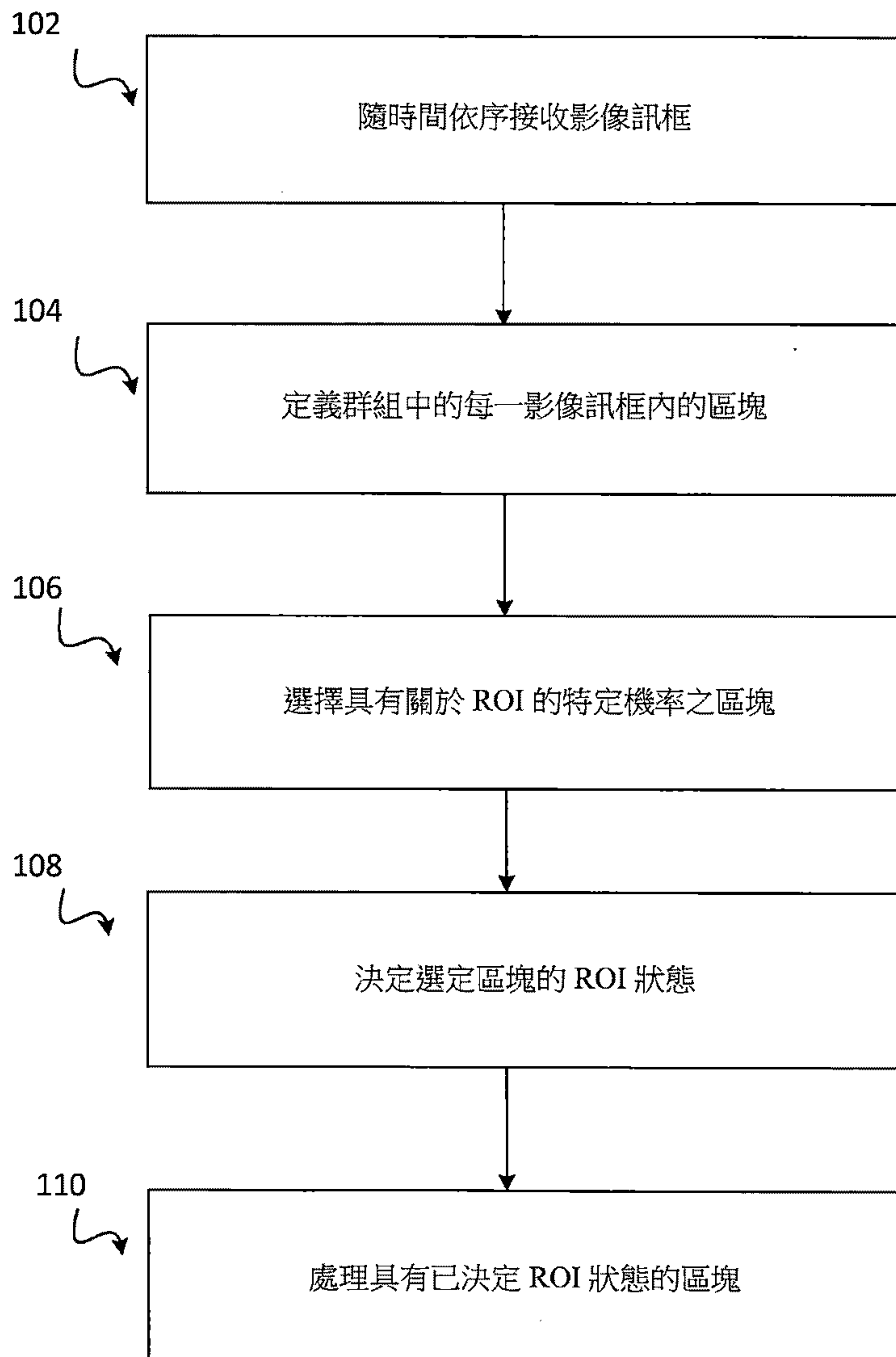
(57) 摘要

本發明揭示一種包括隨時間連續接收一或多個影像訊框之示例方法，其中一第一影像訊框在時間點上是先於一第二影像訊框接收，而且定義該等一或多個影像訊框之任一者中一或多個區塊，其中該第一影像訊框包括一第一區塊，而且該第二影像訊框包括一第二區塊。該方法也包括：隨著該第二區塊具有高於有關一感興趣區域(Region of Interest, ROI)的臨界值之機率，從該第二影像訊框的該等一或多個區塊選擇該第二區塊；及隨著有關該第一區塊的該選定第二區塊，決定該第二區塊的一 ROI 狀態，以根據該第一區塊的一 ROI 狀態來表示該第二區塊是否包括該 ROI 的任何部分。

An example method includes receiving the one or more image frames successively in time, wherein a first image frame is received before a second image frame in time and defining one or more blocks in each of the one or more image frames, wherein the first image frame includes a first block, and the second image frame includes a second block. The method also includes in response to the second block having a probability higher than a threshold to be associated with a region of interest (ROI), selecting the second block from the one or more blocks in the second image frame, and in response to the selected second block being associated with the first block, determining an ROI status of the second block to represent whether the second block includes any part of the ROI based on an ROI status of the first block.

指定代表圖：

100



第一圖

I699993

發明摘要

※ 申請案號： 108109902

H04N 19/167 (2014.01)

※ 申請日： 108年3月20日 ※IPC 分類： H04N 19/176 (2014.01)

【發明名稱】(中文/英文)

辨識感興趣的區域/ REGION OF INTEREST RECOGNITION

【中文】

本發明揭示一種包括隨時間連續接收一或多個影像訊框之示例方法，其中一第一影像訊框在時間點上是先於一第二影像訊框接收，而且定義該等一或多個影像訊框之任一者中一或多個區塊，其中該第一影像訊框包括一第一區塊，而且該第二影像訊框包括一第二區塊。該方法也包括：隨著該第二區塊具有高於有關一感興趣區域(Region of Interest, ROI)的臨界值之機率，從該第二影像訊框的該等一或多個區塊選擇該第二區塊；及隨著有關該第一區塊的該選定第二區塊，決定該第二區塊的一 ROI 狀態，以根據該第一區塊的一 ROI 狀態來表示該第二區塊是否包括該 ROI 的任何部分。

【英文】

An example method includes receiving the one or more image frames successively in time, wherein a first image frame is received before a second image frame in time and defining one or more blocks in each of the one or more image frames, wherein the first image frame includes a first block, and the second image frame includes a second block. The method also includes in response to the second block having a probability higher than a threshold to be associated with a region of interest (ROI), selecting the second block from the one or more blocks in the second image frame, and in response to the selected second block being associated with the first block, determining an ROI status of the second block to represent whether the second block includes any part of the ROI based on an ROI status of the first block.

【代表圖】

【本案指定代表圖】：第（一）圖。

【本代表圖之符號簡單說明】：

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

辨識感興趣的區域/ REGION OF INTEREST RECOGNITION

【背景技術】

【0001】 除非有其他指示，此處說明的方法並不屬於本申請案的申請專利範圍之先前技術，而且不因此處的包含而認為是先前技術。

【0002】 目前已經開發多種技術來識別影像信號中感興趣區域(region of interest, ROI)。雖然一些方式，例如卷積神經網路(convolutional neural network, CNN)、區域 CNN (R-CNN)、更快的 R-CNN 等，已證明有時在檢測數位影像及/或視頻中語義物件的實例方面勝過人類，不過其計算成本仍然很高。

【0003】 因此，希望提供可改善效率來識別 ROI 之方法及系統。

【圖式簡單說明】

【0004】 從下列描述和文後申請專利範圍連同附圖，將可更完整了解到本發明的前述與其他特徵。這些圖式只描述根據本發明的數個具體實施例，因此不應認為限制其範圍。透過附圖更特別及詳細描述本發明。

【0005】 第一圖例示根據本發明的至少某些具體實施例之例示辨識一影像信號的連續影像訊框中物體的感興趣區域(ROI)之方法流程圖；

第二圖為例示根據本發明的至少某些具體實施例之一影像信號中複數個影像訊框之示意圖；

第三圖為例示根據本發明的至少某些具體實施例之決定一影像信號的連續影像訊框中區塊的 ROI 狀態之方法流程圖；及

第四圖為例示根據本發明的至少某些具體實施例之設置辨識一影像信號的連續影像訊框中物體的該 ROI 之一系統區塊圖。

【實施方式】

【0006】 下列說明中揭示的技術細節可讓熟習該項技藝者實施實施

本發明的一或多個具體實施例。

【0007】 由於多層神經網路，例如卷積神經網路(Convolutional Neural Network, CNN)及其變化(例如，區域 CNN (R-CNN)、快速 R-CNN 等)，其檢測和識別物體的錯誤率很低，因此其越來越多用於影像識別系統甚至視頻分析系統中。然而，考慮到在 CNN 的不同層(例如，卷積層、匯集層、整流線性層等)之中，對輸入影像的每個像素執行所需操作的數量以及 CNN 施加的各種限制，已經提出許多方式進一步改善 CNN。

【0008】 在一範例中，第 9,135,553 號美國專利(此後稱為「553 號專利」)提出一種特殊設計的卷積工作電路來執行高速卷積操作。

【0009】 在另一範例中，第 9,424,493 號美國專利說明有關需要 CNN 的固定大小輸入影像之「人為」要求的缺點，而且提出一種從整個影像計算特徵映射並將這些特徵匯集以產生固定長度訓練表示之方法。

【0010】 如所述，這些已知方式的每一者都試圖通過修改這種網路中需要的某些計算來解決 CNN 的限制。然而，不同於本申請案中揭示的該等方式，該等已知方式未能探索在某些情況下減少 CNN 使用率以減輕其限制所帶來衝擊之方法。

【0011】 此外，如前述，由於在多層神經網路中對每個像素進行操作所需的操作次數，使用這種網路的影像識別可能難以實施，尤其是在會需要考慮熱度的行動系統中。有智慧地使用這種多層神經網路執行物體辨識，而不是不加選擇地，通常可以幫助節省電力。

【0012】 第一圖例示根據本發明的至少某些具體實施例之辨識一影像信號的連續影像訊框中物體的感興趣區域(ROI)之方法 100 的流程圖。方法 100 可包括一或多個運算、函數或動作，如一或多個操作 102、104、106、108 及/或 110 所例示。許多操作可根據所要的實施連同成更少操作、分成額外操作及/或取消。

【0013】 方法 100 的程序可從步驟 102 開始，「在時間上依序接收影像訊框」，例如，在時間上於第二影像訊框之前接收第一影像訊框，而且在時間上於第三影像訊框之前接收第二影像訊框。

【0014】 步驟 102 之後接著步驟 104，「定義群組中每一影像訊框中

區塊」。在某些具體實施例中，該處理器設置在該等已接收影像訊框每一者中定義一或多個區塊。在某些具體實施例中，於任何影像訊框中，該等已定義區塊不重疊，而且每一已定義區塊大體上大小相同。該等影像訊框也可由群組中一處理器接收並處理。在某些具體實施例中，該處理器構成將在時間上依序接收的預定數量影像訊框分組，例如，第一群組可包括在時間 t 、 $t+1$ 和 $t+2$ 上接收的三個影像訊框，而且第二群組可包括在時間 $t+3$ 、 $t+4$ 和 $t+5$ 上接收的另外三個影像訊框。該預定數量可根據該處理器的處理效率及/或能力而變。在由較少數量影像訊框組成的群組中，給定這些影像訊框的時間接近度，該群組中訊框之間的相似性及/或相關性可能更高，使得該群組中一訊框的處理結果可用來當成處理相同群組中其他訊框的基礎，以提高處理效率。但是，用於處理具有較少影像訊框的群組之這種效率增益應與增加的總處理時間相平衡，以處理更多的群組總數。群組中此預定影像訊框數量的一範例範圍可為從 2 至 8 的整數。

【0015】 在某些具體實施例中，一較大區塊可定義在該群組中在時間上較該等其他影像訊框之前的一影像訊框中，一較小區塊可定義在該群組中在時間上較該等其他影像訊框之後的另一影像訊框中。底下連同第二圖揭示額外解釋與範例。

【0016】 步驟 104 接著步驟 106，「選擇具有特定機率有關於 ROI 的區塊」。在某些具體實施例中，處理步驟 104 定義的任何區塊來決定有關於該 ROI 的機率。此機率反映出該區塊包括 ROI 任何部分(例如人臉、車輛等)的可能性，該決定可基於任何技術上可行的方式。在某些具體實施例中，該決定係根據某些快速並且相關的資源節省參數，諸如有關於一或多個紋理內容(例如人臉、車輛等等的紋理)及/或一或多個邊緣(例如像素值的急遽變化)的等級。隨著高於一預定臨界值的機率，該處理器構成選擇該區塊以在步驟 108、110 進一步處理。隨著小於或等於該預定臨界值的機率，該處理器構成不選擇該區塊進行進一步處理，如此諸如 CNN 及其變化的多層網路將不用來處理這種區塊。

【0017】 步驟 106 接著步驟 108，「決定選定區塊的 ROI 狀態」。如上所揭示，該等選定區塊具有較高機率有關於該 ROI。在步驟 108，該處理器

設置為決定是否有任何該等選定區塊確實包括該 ROI 的任何部分。在某些具體實施例中，該處理器構成透過諸如 CNN 及其變化的多層網路執行物體辨識方式從而處理第一組選定區塊，以決定是否有任何該第一組選定區塊包括該 ROI 的一部分。隨著一選定區塊確實包括至少某些該 ROI 的決定，該處理器設置成將該選定區塊的該 ROI 狀態設定為「真」或「1」。否則，該選定區塊的該 ROI 狀態設定為「偽」或「0」。

【0018】 在某些具體實施例中，該處理器設置成運用該第一組選定區塊的該已決定 ROI 狀態，而非運用 CNN 及其變化來決定是否有任何第二組選定區塊包括該 ROI 的一部分。在某些具體實施例中，一影像訊框中較大區塊的該已決定 ROI 狀態可用來決定在另一影像訊框中較小區塊的該 ROI 狀態，以避免使用到與執行該 CNN 相關操作相關聯的資源。該較小區塊在其影像訊框中實際位置可類似於該較大區塊在其影像訊框中實際位置。在某些其他具體實施例中，第一群組的一影像訊框中區塊之該已決定 ROI 狀態稍後也用來決定第二群組的另一影像訊框中區塊之該 ROI 狀態。底下揭示額外解釋與範例。

【0019】 步驟 108 接著步驟 110，「處理具有已決定 ROI 狀態的區塊」。在某些具體實施例中，該處理器構成處理具有在步驟 108 決定為「真」或「1」狀態的該區塊，以偵測及辨識該 ROI。在某些具體實施例中，可使用某些技術可行的方式，例如非最大抑制(NMS，non-maximum suppression)。在該 NMS 方式中，最高得分窗格包括具有狀態已經決定選取「真」或「1」的第一組區塊，然後假設確實覆蓋該 ROI。抑制太靠近所選窗格的其他窗格(每一者包括具有狀態已經決定為「真」或「1」的複數個區塊)。在其餘的窗格中，選擇下一最高得分窗格，而且重複上述步驟，直到僅剩下最高得分窗格。

【0020】 第二圖為例示根據本發明的至少某些具體實施例之一影像信號中複數個影像訊框之示意圖。在某些具體實施例中，該影像信號可包括影像訊框 211、221、231、241、251 和 261。影像訊框 211、221、231、241、251 和 261 分別為在時間點 t 、 $t+1$ 、 $t+2$ 、 $t+3$ 、 $t+4$ 和 $t+5$ 上接收到的連續影像訊框。

【0021】 在某些具體實施例中，假設有預設規則，其中每一群組都由 3 個影像訊框構成。訊框 211、221、231 屬於第一組影像訊框，而且訊框 241、251、261 屬於第二組影像訊框。

【0022】 為了說明為目的，並非限制，連同第一圖中步驟 104「定義群組中每一影像訊框中區塊」，大的非重疊區塊(例如 4 區塊)都定義在每一群組中最早的影像訊框中(例如群組 1 中的影像訊框 211 及群組 2 中的影像訊框 241)。此外，中的非重疊區塊(例如 9 區塊)都定義在接在每一群組中最早影像訊框之後該影像訊框中(例如群組 1 中的影像群組 221 及群組 2 中的影像群組 251)。此外，小的非重疊區塊(例如 36 區塊)都定義在每一群組中最晚影像訊框中(例如群組 1 中的影像群組 231 以及群組 2 中的影像群組 261)。

【0023】 在某些具體實施例中，僅供說明，並非限制，影像訊框 211 和 241 中的 4 個區塊的大小相同。例如，4 個區塊之任一者可由 81 個像素乘 81 個像素來定義。影像訊框 221、251 中的 9 個區塊的大小也可相同。例如，9 個區塊之任一者可由 36 像素乘 36 像素來定義。影像訊框 231、261 中的 36 個區塊的大小也可相同。例如，36 個區塊之任一者可由 9 像素乘 9 像素來定義。

【0024】 在某些具體實施例中，連同步驟 106「選擇具有特定機率會有關於 ROI 的區塊」，在每一區塊上使用某些快速並相對節省資源的方式來決定有關於該 ROI 的該區塊之機率。隨著高於一臨界值的機率，將選取該區塊。為了例示，在第二圖中的影像訊框 211 已定義的區塊之間選擇區塊 212、214。同樣地，在影像訊框 241 中已定義的區塊之間選擇區塊 242 和 244。在影像訊框 221 中已定義的區塊之間選擇區塊 222、223、224、225、226、227。在影像訊框 251 中已定義的區塊之間選擇區塊 252、253、254、255、256。在影像訊框 231 中已定義的區塊之間選擇區塊 231'、232、232'、233、233'、234、234'、235、235'、236、236'、237、237'、238、238'、239、239'、240。在影像訊框 261 中已定義的區塊之間選擇區塊 261'、262、262'、263、263'、264、264'、265、265'、266、266'、267、267'、268、268'、269。

【0025】 在某些具體實施例中，連同步驟 108「決定選定區塊的 ROI 狀態」，使用通過多層神經網路(例如 CNN 相關操作)的物體辨識方式來決定該等選定區塊的該 ROI 狀態。

【0026】 然而，如上所揭示，這些方式相當耗資源。因此，在某些情況下，利用第一影像訊框中第一區塊已決定的 ROI 狀態來及時決定該第二影像訊框中第二區塊的 ROI 狀態，避免執行這些耗費資源的方法。在某些具體實施例中，該第二影像訊框可與該第一影像訊框在相同群組中。在某些具體實施例中，該第二影像訊框與該第一影像訊框可在不同群組中。

【0027】 **案例 I-時間上連續的該等影像訊框都在相同群組中**

【0028】 **A. 影像訊框 211 中的選定區塊**

【0029】 假設在時間點 t 上，影像訊框 211 中的選定區塊 212、214 的該 ROI 狀態已經由 CNN 相關操作決定為 C_t 。

【0030】 **B. 影像訊框 221 中的選定區塊**

【0031】 為了決定影像訊框 221 中的選定區塊 222、223、224、225、226、227 的 ROI 狀態，在一具體實施例中，當影像訊框 221 中的一區塊(例如位於左上角的 36 個像素乘 36 個像素區塊 222)被在時間點是在影像訊框 221 之前的影像訊框 211 中的一空間對應第一區塊(例如也位於左上角的 81 個像素乘 81 個像素區塊 212)所覆蓋，在不用執行該 CNN 相關操作之下，就可決定區塊 222 的該 ROI 狀態為相同於區塊 212 的該 ROI 狀態。因此，決定區塊 222 的該 ROI 狀態也是 C_t ，如第二圖所示。同樣地，影像訊框 221 中的區塊 227 也被影像訊框 211 中的一空間對應範圍區塊 214 所覆蓋。因此，決定區塊 227 的該 ROI 狀態也是 C_t 。

【0032】 在某些具體實施例中，對於影像訊框 221 內選取的區塊 223、224、225、226，因為其並未遭影像訊框 211 的空間對應區塊 212、214 完全覆蓋，類似於在時間點 t 上決定區塊 212、214 的該 ROI 狀態，在時間點 $t+1$ 上也使用該 CNN 相關操作來決定區塊 223、224、225、226 的該 ROI 狀態。在 $t+1$ 時間點的此 ROI 狀態係例示如第二圖中的 C_{t1} 。

【0033】 **C. 影像訊框 231 中的選定區塊**

【0034】 為了決定影像訊框 231 中的區塊 231'、232、232'、233、233'、

234、234'、235、235'、236、236'、237、237'、238、238'、239、239'、240 的該 ROI 狀態，在一具體實施例中，當影像訊框 231 中的選定區塊(例如任何 9 個像素乘 9 個像素區塊 213'、223、232'、233、233'、234、234'、235、235')由影像訊框 211 (即時間上最早並且與影像訊框 221、231 在相同群組中該影像訊框)中的一空間對應區塊(例如 81 個像素乘 81 個像素區塊 212)完全覆蓋，在不用執行該 CNN 相關操作之下，可決定該選取區塊的該 ROI 狀態係相同於區塊 212 的該 ROI 狀態。因此，決定區塊 213'、232、232'、233、233'、234、234'、235、235'的該 ROI 狀態為 C_t ，如第二圖所示。同樣地，當影像訊框 231 中的任何區塊 238'、239、239'、240 都被影像訊框 211 中的一空間對應區塊 214 完全覆蓋，則決定區塊 238'、239、239'、240 的該 ROI 狀態也為 C_t 。

【0035】 在已經套用影像訊框 211 中的所有已決定 ROI 狀態之後，則影像訊框 221 中決定的該 ROI 狀態也可用來決定影像訊框 231 中的該等選定區塊的該 ROI 狀態。在某些具體實施例中，在已決定影像訊框 231 中的區塊 213'、232、232'、233、233'、234、234'、235、235'、238'、239、239'、240 的該 ROI 狀態為 C_t 之後，則仍舊需要決定區塊 236、236'、237、237'、238 的該 ROI 狀態。為了利用在影像訊框 221 中決定的該 ROI 狀態，當影像訊框 231 中的一選定區塊(例如任何 9 個像素乘 9 個像素區塊 236、236'、237'、238)由影像訊框 221 (即時間點晚於影像訊框 211 但時間點早於影像訊框 231 的該影像訊框)中的一空間對應區塊(例如分別為 36 個像素乘 36 個像素區塊 223 和 226)完全覆蓋，在不用執行該 CNN 相關操作之下，就可決定區塊 236、236'、237'、238 的該 ROI 狀態分別係相同於影像訊框 221 中的區塊 223 和 226 之該 ROI 狀態。因此，決定區塊 236、236'、237'、238 的該 ROI 狀態為 C_{t1} 。

【0036】 在某些具體實施例中，在已經決定影像訊框 231 中的選定區塊 237 並未由相同群組中的影像訊框 211 或 221 內的任何空間對應區塊完全覆蓋之後，則使用該 CNN 操作來決定影像訊框 231 中區塊 237 的該 ROI 狀態。如第二圖所示，在時間點 $t+2$ 上決定區塊 237 的該 ROI 狀態為 C_{t2} 。因此，明顯減少決定影像訊框 221、231 中選定區塊的該 ROI 狀態所需之總

資源。

【0037】 案例 II – 隨時間連續的影像訊框在不同群組中

【0038】 在某些具體實施例中，影像訊框 211、221、231 在該第一群組中，而且影像訊框 241、251、261 在該第二群組中。在影像訊框的該第二群組之前先接收影像訊框的該第一群組。同樣，在該第二群組中，影像訊框 241 在時間框上為最早的，而且影像訊框 261 在該第二群組中是時間框最晚的。雖然該等影像訊框是在不同群組中，尤其是在時間上接近，其仍舊彼此相關。在某些具體實施例中，第一群組的影像訊框 211 可對應至第二群組的影像訊框 241，而且在影像訊框 211 中定義的該等區塊可具有與對應影像訊框 241 中定義的該等區塊相同大小(例如 81 像素乘 81 像素)。同樣地，第一群組的影像訊框 221 可對應至第二群組的影像訊框 251，而且在影像訊框 221 中定義的該等區塊可具有與對應影像訊框 251 中定義的該等區塊相同大小(例如 36 像素乘 36 像素)。第一群組的影像訊框 231 可對應至第二群組的影像訊框 261，而且在影像訊框 231 中定義的該等區塊可具有與對應影像訊框 261 中定義的該等區塊相同大小(例如 9 個像素乘 9 個像素)。

【0039】 A. 影像訊框 241 中的選定區塊

【0040】 在某些具體實施例中，雖然影像訊框 211、241 在兩不同群組中，影像訊框 211 (在時間上於影像訊框 241 之前)中選定區塊 214 的該已決定 ROI 狀態可用來決定影像訊框 241 中的空間對應區塊 244 的該 ROI 狀態，因為區塊 214 和區塊 244 兩者在其個別影像訊框的右下角，而且區塊 244 是被區塊 214 完全覆蓋。因此如第二圖所示，在不用執行該 CNN 相關操作之下，可決定區塊 244 的該 ROI 狀態為 C_t 。

【0041】 然而，如第二圖所示，在空間上，對應至影像訊框 241 中區塊 242 的影像訊框 211 中的區塊並沒有已決定的 ROI 狀態。因此為了決定區塊 242 的該 ROI 狀態，在某些具體實施例中，運用該等 CNN 相關操作。如第二圖所示，在時間點 $t+3$ 上決定選定區塊 242 的該 ROI 狀態為 C_{t3} 。

【0042】 B. 影像訊框 251 中的選定區塊

【0043】 在某些具體實施例中，為了決定該第二群組中影像訊框 251 和 261 中的該等選定區塊的該 ROI 狀態，可使用上面關於決定該第一群組

中影像訊框 221、231 中該等選定區塊的該 ROI 狀態所揭示之相同方式。換言之，當影像訊框 251 中選定區塊 252、256 (例如 36 個像素乘 36 個像素) 分別遭到影像訊框 241 中的空間對應區塊 242、244 (例如 81 個像素乘 81 個像素) 完全覆蓋時，在不用執行該 CNN 相關操作之下，就可決定區塊 252、256 的該 ROI 狀態分別相同於區塊 242、244 的該已決定 ROI 狀態成為 $Ct3$ 和 Ct 。

【0044】 為了決定影像訊框 251 中的其餘選定區塊 253、254 和 255 的該 ROI 狀態，雖然區塊 253 沒有被影像訊框 241 中的任何空間對應區塊完全覆蓋，但區塊 253 在該第一群組的影像訊框 221 中具有一對應區塊 226。如第二圖所示，區塊 226 具有以 $Ct1$ 表示的一已決定 ROI 狀態。因為區塊 253 是被區塊 226 完全覆蓋，因此，在不用執行該 CNN 相關操作之下，就可決定區塊 253 的該 ROI 狀態也為 $Ct1$ 。

【0045】 針對影像訊框 251 中的選定區塊 254 和 255，其並沒有被具有相同第二群組中的影像訊框 241 內的一已決定 ROI 狀態的任何區塊、或被第一群組中的影像訊框 221 內的一已決定 ROI 狀態的任何區塊所完全覆蓋。如此，在時間點 $t+4$ 上執行該等 CNN 相關操作來決定選定區塊 254、255 的該 ROI 狀態，而且以 $Ct4$ 代表該已決定的 ROI 狀態，如第二圖所示。

【0046】 C. 訊框 261 中的選定區塊

【0047】 當影像訊框 261 中的一選定區塊(例如任何 9 像素乘 9 像素區塊 264'、265 和 265')由影像訊框 241 (即該第二群組中時間上最早的該影像訊框)中的一空間對應區塊(例如 81 個像素乘 81 個像素區塊 242)完全覆蓋，則不用執行該 CNN 相關操作就可決定該選定區塊的該 ROI 狀態與區塊 242 的該 ROI 狀態相同。如第二圖所示，決定區塊 264'、265、265' 的該 ROI 狀態為 $Ct3$ 。同樣地，因為影像訊框 261 中的選定區塊 266、266'、267、267'、268、268'、269 由影像訊框 241 中的空間對應區塊 244 完全覆蓋，所以決定區塊 266、266'、267、267'、268、268'、269 的該 ROI 狀態係相同於區塊 244 的該已決定 ROI 狀態或 Ct 。

【0048】 在已經套用影像訊框 241 中的所有已決定 ROI 狀態，則影像訊框 251 中決定的該 ROI 狀態也可用來決定影像訊框 261 中的該等選定

區塊的該 ROI 狀態。在某些具體實施例中，決定區塊 264'、265、265'、266、266'、267、267'、268'、268、269 的該 ROI 狀態為 $Ct3$ 和 Ct 之後，仍舊需要決定區塊 261'、262、262'、263、263'、264 的該 ROI 狀態。為了利用在訊框 251 中決定的該 ROI 狀態，當影像訊框 261 中的一選定區塊(例如 9 個像素乘 9 個像素區塊 262、263'、263'、264 之任一者)由影像訊框 251 (即時間點上晚於影像訊框 241 但時間點上早於影像訊框 261 的該影像訊框)中的一空間對應區塊(例如分別為 36 像素乘 36 像素區塊 254 和 255)完全覆蓋，在不用執行該 CNN 相關操作之下，就可決定區塊 262、263'、263'、264 的該 ROI 狀態分別相同於影像訊框 251 中的區塊 255 和 254 之該 ROI 狀態。因此，決定區塊 262、263'、263'和 264 的該 ROI 狀態為 $Ct4$ 。

【0049】 在某些具體實施例中，雖然已決定影像訊框 261 中的選定區塊 261'並未由具備該第二群組的影像訊框 241 和 251 中的該已決定 ROI 狀態之一空間對應區塊完全覆蓋，區塊 261'具有一空間對應區塊 236'於該第一群組的影像訊框 231 中，並有已決定的 ROI 狀態 $Ct1$ 。因此，類似於根據區塊 214 的該已決定 ROI 狀態來決定區塊 244 的該 ROI 狀態，在不用執行該 CNN 相關操作之下，可根據區塊 236'的該已決定 ROI 狀態或 $Ct1$ 來決定區塊 261'的該 ROI 狀態。

【0050】 在某些具體實施例中，在已經決定影像訊框 261 中的選定區塊 262 並未由相同群組中影像訊框 241 和 251 中的任何空間對應區塊完全覆蓋，也未由具有該第一群組的對應影像訊框 231 中的該已決定 ROI 狀態的任何區塊完全覆蓋之後，則使用該 CNN 操作來決定影像訊框 261 中區塊 262 的該 ROI 狀態。如第二圖所示，區塊 262 的該 ROI 狀態決定為是在時間點 $t+5$ 上的 $Ct5$ 。

【0051】 總結來說，在某些具體實施例中，用來決定一影像訊框中選定區塊(例如影像訊框 261 中的區塊 262)的該 ROI 狀態之該處理可遵循下列順序：1) 參考具有相同群組中(例如該第二群組)該較早接收到的影像訊框中的該已決定 ROI 狀態之該等區塊，從最早接收的訊框(例如該第二群組中影像訊框 241)開始；2) 在相同群組中該等較早影像訊框中(例如影像訊框 241 和 251)所有該等相關區塊已經檢查之後，參考具有不同群組中該對應影

像訊框中(例如該第一群組的影像訊框 231)該已決定 ROI 狀態之該等區塊；及 3) 若仍舊無具有該已決定 ROI 狀態的對應區塊，則執行該 CNN 相關操作來決定此選取區塊的該 ROI 狀態。因為已在時間上連續接收該等影像訊框，相同群組中一對影像訊框之間的時間差小於不同群組中一對影像訊框之間的時間差。換言之，使用第二圖當成範例，該第二群組中影像訊框 241 和 251 之間的時間差小於影像訊框 251 與影像訊框 221 之間的時間差。如此，相較於影像訊框 221 中的區塊 227 與影像訊框 251 中的區塊 256 的該已決定 ROI 狀態，影像訊框 241 中的區塊 244 與影像訊框 251 中的區塊 256 的該已決定 ROI 狀態非常有可能維持相同。

【0052】 第三圖為例示根據本發明的至少某些具體實施例之決定一影像信號的連續接收影像訊框中該等選定區塊的該 ROI 狀態之方法 300 之流程圖。在某些具體實施例中，該等連續接收的影像訊框可在相同或不同群組中。方法 300 可包括一或多個運算、函數或動作，如一或多個操作 302、304、306、308 及/或 310 所例示。許多操作可根據所要的實施連同成更少操作、分成額外操作及/或取消。在某些具體實施例中，方法 300 執行有關步驟 108 所例示「決定選定區塊的 ROI 狀態」的操作。

【0053】 方法 300 的處理可從步驟 302「選取的區塊由第一影像訊框中對應的第一區塊完全覆蓋？」開始。如上揭示，為了決定第二影像訊框中選定區塊的該 ROI 狀態，步驟 302 決定該選定區塊是否由第一影像訊框中該對應的第一區塊完全覆蓋，其中該第一影像訊框與該第二影像訊框屬於相同群組。在某些具體實施例中，在該群組中該第一影像訊框在時間上早於該第二影像訊框。

【0054】 對於該選定區塊由該第一影像訊框中該對應的第一區塊完全覆蓋，步驟 302 可接著步驟 304「根據第一區塊的已決定 ROI 狀態來決定選定區塊的 ROI 狀態」。例如在與第二圖連同之下，影像訊框 251 中的區塊 252、256 分別由影像訊框 241 的空間對應區塊 242、244 完全覆蓋。因此，區塊 252 的該 ROI 狀態係根據區塊 242 的該已決定 ROI 狀態(即 $Ct3$)來決定。同樣地，區塊 256 的該 ROI 狀態係根據區塊 244 的該已決定 ROI 狀態(即 Ct)來決定。因此，在步驟 304，決定該 ROI 狀態時不執行該等 CNN

相關操作。

【0055】 在決定該選定區塊並未由一第一影像訊框中對應的第一區塊完全覆蓋之回應上，步驟 302 可接著步驟 306「不同群組中對應區塊具有已決定 ROI 狀態？」。在某些具體實施例中，該選定區塊具有在另一群組中一對應影像訊框中的一對應區塊。若另一群組中該對應影像訊框中的該對應區塊具有該已決定 ROI 狀態，步驟 306 可接著步驟 308「根據對應區塊的該已決定 ROI 狀態來決定該選定區塊的 ROI 狀態」。在某些具體實施例中，可在該選定區塊與該對應區塊之間執行一相似度比較(例如絕對差操作的總和)來決定該對應區塊是否相似於該第二區塊。在某些其他具體實施例中，例如連同第二圖，影像訊框 251 中的區塊 253 並未由影像訊框 241 中的區塊 242 和 244 完全覆蓋。然而，區塊 253 有一對應區塊 226 於對應影像訊框 221 中，其有該已決定 ROI 狀態，而且影像訊框 221 屬於不同群組。因此，區塊 253 的該 ROI 狀態係根據區塊 226 的該已決定 ROI 狀態(即 *Ct1*)來決定。因此，在步驟 308，不執行 CNN 相關操作來決定該 ROI 狀態。

【0056】 若步驟 306 不同群組中沒有對應區塊具有已決定 ROI 狀態，步驟 306 可接著步驟 310「透過多層神經網路使用物體辨識方式來決定 ROI 狀態」。例如在與第二圖連同之下，影像訊框 251 中的選定區塊 254 和 255 並未由影像訊框 241 中具有該已決定 ROI 狀態的影像訊框 241 中的該等空間對應區塊完全覆蓋。不同群組中的影像群組 221 內的其對應區塊也不具有該已決定 ROI 狀態。因此在步驟 310，透過像是 CNN 及其變化的多層神經網路使用一物體辨識方式來決定區塊 254 和 255 的該 ROI 狀態。

【0057】 第四圖顯示構成辨識一影像信號的連續已接收影像訊框中物體的該 ROI 之一系統例示具體實施例之方塊圖。尤其是，ROI 辨識系統 420 可設置成處理視訊序列輸入 410，並產生包括該已決定 ROI 的視訊序列輸出 465，讓顯示器 470 顯示。ROI 辨識系統 420 可構成包括但不限制於視訊解碼器 430、區塊定義引擎 440、區塊選擇引擎 450、ROI 狀態決定引擎 460、區塊處理引擎 480、處理器 421 及/或記憶體 422。

【0058】 在某些具體實施例中，視訊序列輸入 410 可對應至一視訊串流。視訊解碼器 430 可解壓縮並解碼該視訊串流，並擷取連續影像訊框 431。

一或多個影像訊框 431 可包括該 ROI。

【0059】 然後影像訊框 431 可由區塊定義引擎 440、區塊選擇引擎 450 和 ROI 狀態決定引擎 460 來處理。連同第一圖，區塊定義引擎 440 可執行步驟 104。在某些具體實施例中，區塊選擇引擎 450 構成執行步驟 106。在某些其他具體實施例中，ROI 狀態決定引擎 460 構成執行步驟 108，而且區塊處理引擎 480 構成執行步驟 110。

【0060】 在某些具體實施例中，連同第一圖，ROI 辨識系統 420 可運用處理器 421 與區塊定義引擎 440、區塊選擇引擎 450、ROI 狀態決定引擎 460 及/或區塊處理引擎 480 互動來執行步驟 102、104、106、108 及/或 110。處理器 421 可為一微處理器、圖形處理單元或根據可程式指令來執行命令的任何處理單元。在某些具體實施例中，處理器 421 可運用記憶體 422 來執行該可程式指令，並儲存區塊定義引擎 440、區塊選擇引擎 450、ROI 狀態決定引擎 460 及/或區塊處理引擎 480 的中間處理結果，以使用上面 ROI 辨識系統 420 中所揭示的任何引擎來進行進一步處理。記憶體 422 可為任何形式的非暫態電腦可讀取儲存媒體，包括但不受限於隨機存取記憶體 (RAM, random access memory)、唯讀記憶體 (ROM, read-only memory)、快閃記憶體、傳統磁性或光學碟片、磁帶機或這些裝置的組合。

【0061】 顯示器 470 的某些範例可包括但不受限於電腦監視器、裝置螢幕、電視或投影機。

【0062】 雖然已經參考特定示範具體實施例來描述本發明，不過應了解，本發明並不限於所描述之具體實施例，而是可採用文後申請專利範圍的精神和範疇內進行修改和更改來實踐。因此，說明書與圖式僅供參考而不是限制。

【符號說明】

100 方法	222-227 區塊
102-110 操作	231 影像訊框
211 影像訊框	231' 區塊
212 區塊	232-240 區塊
221 影像訊框	232'-239' 區塊

- 241 影像訊框
- 214 區塊
- 242 區塊
- 244 區塊
- 251 影像訊框
- 252-256 區塊
- 261 影像訊框
- 261'-268' 區塊
- 262-269 區塊
- 300 方法
- 302-310 操作
- 410 視訊序列輸入
- 420 ROI 辨識系統
- 421 處理器
- 422 記憶體
- 430 視訊解碼器
- 431 影像訊框
- 440 區塊定義引擎
- 450 區塊選擇引擎
- 465 視訊序列輸出
- 460 ROI 狀態決定引擎
- 470 顯示器
- 480 區塊處理引擎

【發明申請專利範圍】

【請求項1】 一種非暫態電腦可讀取儲存媒體，其包括由一電腦系統的一處理器所執行之一組指令，使該處理器執行一方法來辨識一影像信號中一或多個影像訊框的感興趣區域(ROI)，該方法包括：

 隨時間連續接收該等一或多個影像訊框，其中一第一影像訊框在時間點上是先於一第二影像訊框接收；

 定義該等一或多個影像訊框之每一者中一或多個區塊，其中該第一影像訊框包括一第一區塊，而且該第二影像訊框包括一第二區塊，其中該第一影像訊框中該第一區塊之大小係大於該第二影像訊框中該第二區塊之大小；

 若該第二區塊具有一機率高於有關該 ROI 的一臨界值，則從該第二影像訊框的該等一或多個區塊選擇該第二區塊；

 若該選定第二區塊有關於該第一區塊，則根據該第一區塊的一 ROI 狀態決定該第二區塊的一 ROI 狀態，來表示該第二區塊是否包括該 ROI 的任何部分；及

 若該選定第二區塊無關於該第一區塊，則透過一多層神經網路，使用一物體辨識方式來決定該第二區塊的該 ROI 狀態。

【請求項2】 如申請專利範圍第 1 項之非暫態電腦可讀取儲存媒體，其中基於有關該第二區塊內紋理內容和邊緣的一或多的層次來決定該第二區塊有關該 ROI 的機率。

【請求項3】 如申請專利範圍第 1 項之非暫態電腦可讀取儲存媒體，其中在一或多個影像訊框之任一者中定義的該等一或多個區塊並未重疊，而且具有大體上相同的大小。

【請求項4】 如申請專利範圍第 1 項之非暫態電腦可讀取儲存媒體，其中該第一影像訊框與該第二影像訊框都屬於一第一群組。

【請求項5】 如申請專利範圍第 4 項之非暫態電腦可讀取儲存媒體，其中根據該選定第二區塊在該第二影像訊框中實際位置匹配該第一區塊在該第一影像訊框中實際位置並且該第二區塊由該第一區塊完全覆蓋，決定該選定第二區塊有關於該第一區塊。

【請求項6】 如申請專利範圍第 1 項之非暫態電腦可讀取儲存媒體，其中根據該選定第二區塊與該第一區塊間的相似度，決定該選定第二區塊有關於該第一區塊。

【請求項7】 如申請專利範圍第 1 項之非暫態電腦可讀取儲存媒體，其包括額外指令，隨著由該處理器的執行，使該處理器：

接收一第三影像訊框與一第四影像訊框，其中該第二影像訊框在時間點上是先於該第三影像訊框接收，而且該第三影像訊框在時間點上是先於該第四影像訊框接收，該第三影像訊框包括一第三區塊，該第四影像訊框包括一第四區塊，而且該定義更包括：

定義該第一影像訊框中該第一區塊的大小大於該第二影像訊框中該第二區塊；

定義該第三影像訊框中該第三區塊的大小大體上等於該第一影像訊框中該第一區塊；及

定義該第四影像訊框中該第四區塊的大小大體上等於該第二影像訊框中該第二區塊，其中該第一影像訊框與該第二影像訊框都在一第一群組中，而且該第三影像訊框與該第四影像訊框都在一第二群組中。

【請求項8】 如申請專利範圍第 7 項之非暫態電腦可讀取儲存媒體，其包括額外指令，隨著由該處理器的執行，使該處理器：

在透過該多層神經網路執行該物體辨識方式之前，決定該選定第二區塊是否有關於該第四區塊。

【請求項9】 一種構成辨識一影像信號中一或多個影像訊框中感興趣區域(ROI)之設備，包括：

一處理器；

一視訊解碼器，其構成隨著時間連續擷取一或多個影像訊框，其中一第一影像訊框在時間點上是先於一第二影像訊框接收；及

一非暫態電腦可讀取儲存媒體儲存指令，隨著由該處理器的執行，使該處理器：

定義該等一或多個影像訊框之每一者中一或多個區塊，其中該第一影像訊框包括一第一區塊，而且該第二影像訊框包括一第二區塊，其中該第一影像訊框中該第一區塊之大小係大於該第二影像訊框中該第二區塊之大小；

若該第二區塊具有高於有關該 ROI 的臨界值的機率，則從該第二影像訊框的該等一或多個區塊選擇該第二區塊；

若該選定第二區塊有關於該第一區塊，則根據該第一區塊的一 ROI 狀態決定該第二區塊的一 ROI 狀態，來表示該第二區塊是否包括該 ROI 的任何部分；及

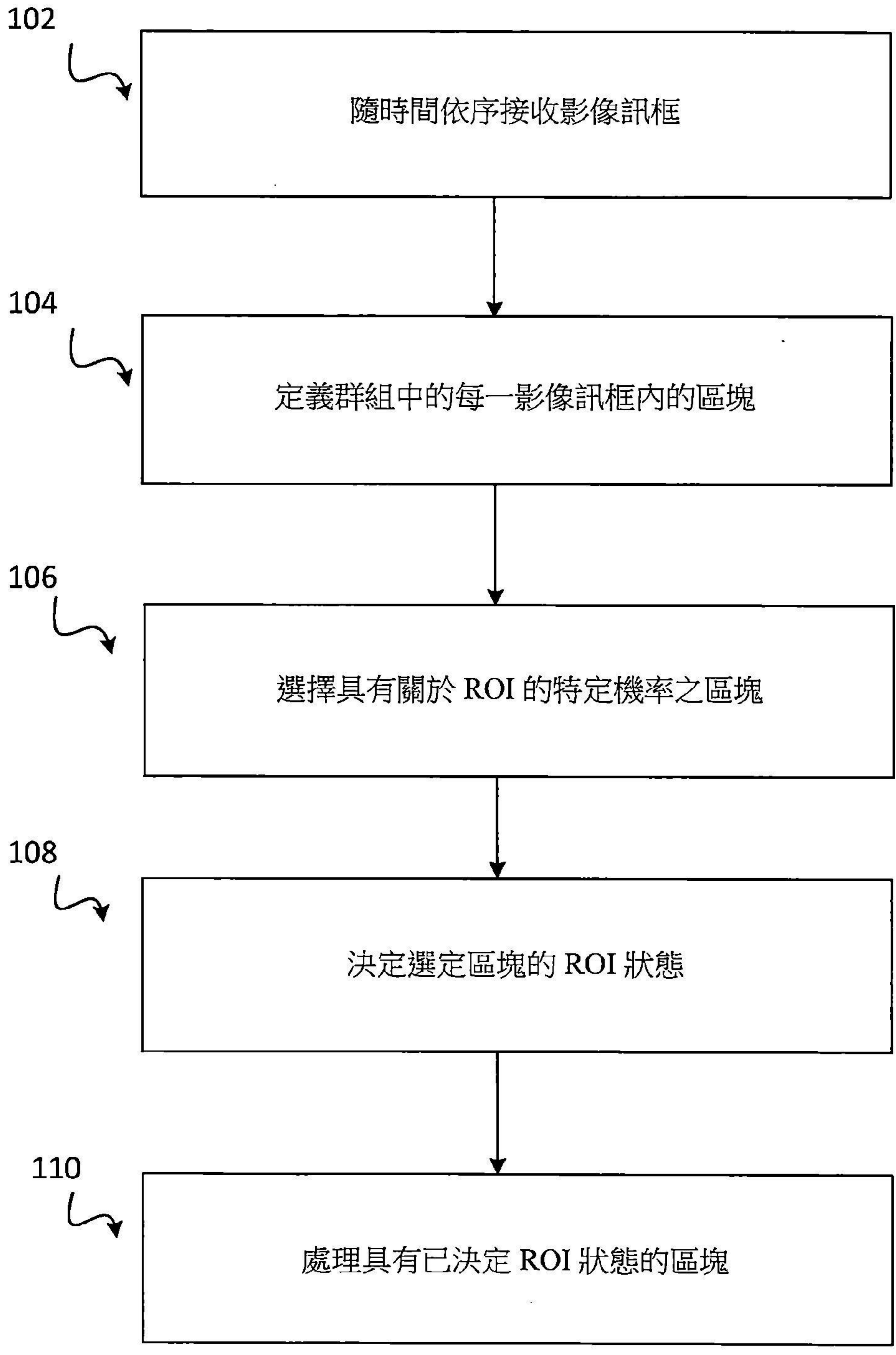
若該選定第二區塊無關於該第一區塊，透過一多層神經網路，使用一物體辨識方式來決定該第二區塊的該 ROI 狀態。

【請求項10】 如申請專利範圍第 9 項之設備，其中該第一影像訊框與該第二影像訊框都屬於一第一群組。

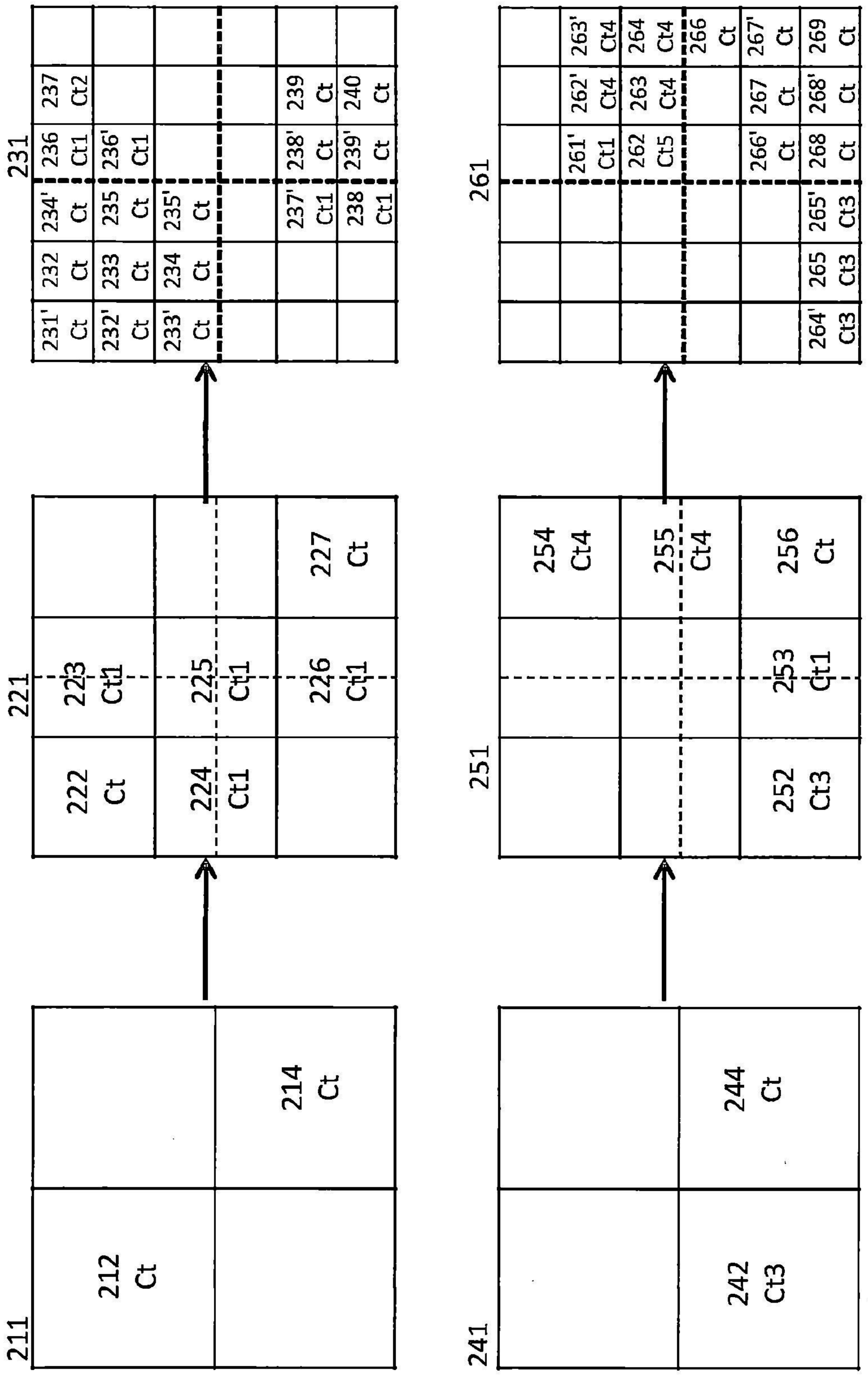
【請求項11】 如申請專利範圍第 10 項之設備，其中該處理器構成根據該選定第二區塊在該第二影像訊框中實際位置匹配該第一區塊在該第一影像訊框中實際位置並且該第二區塊由該第一區塊完全覆蓋，決定該選定第二區塊有關於該第一區塊。

圖式

100

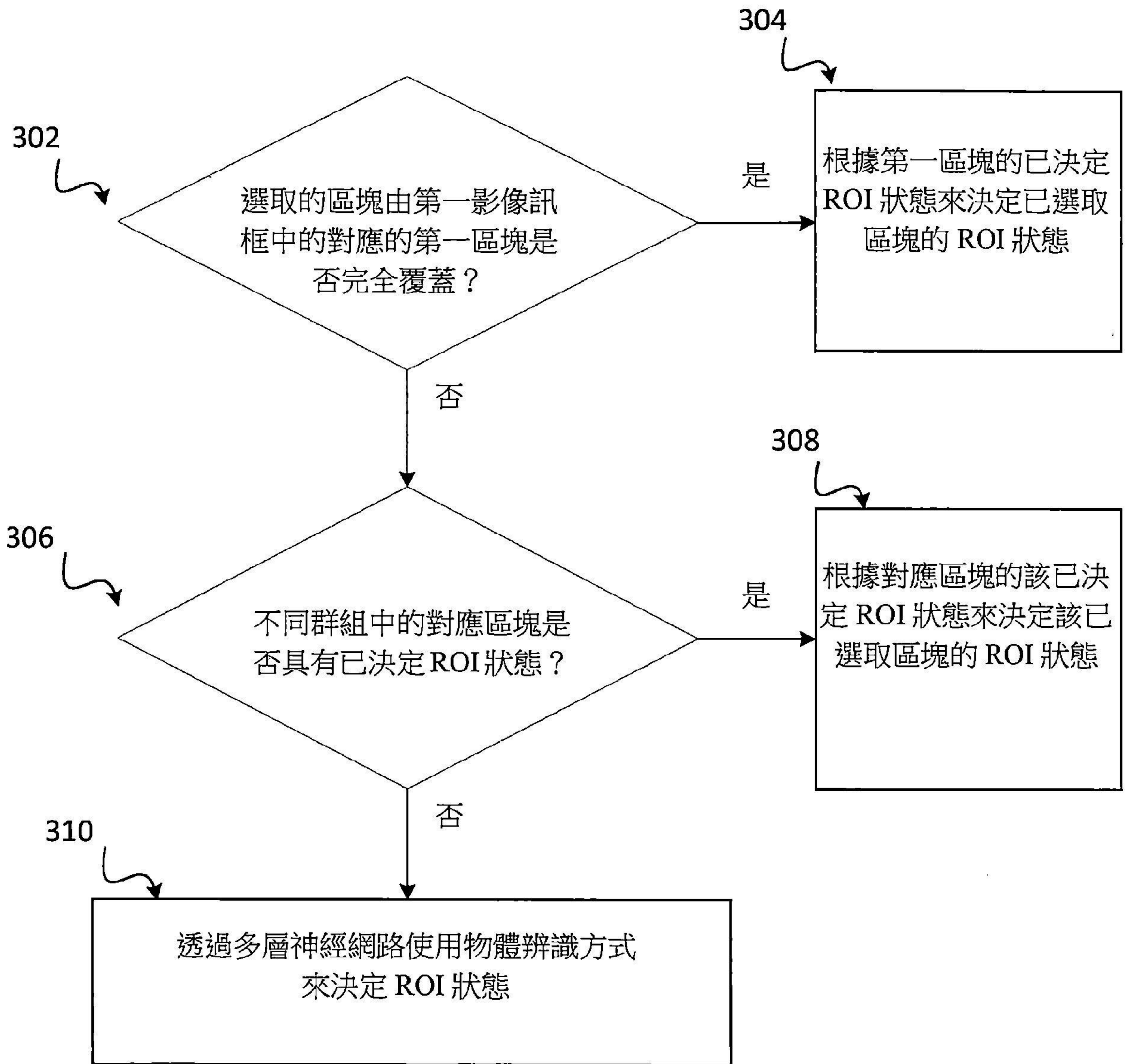


第一圖

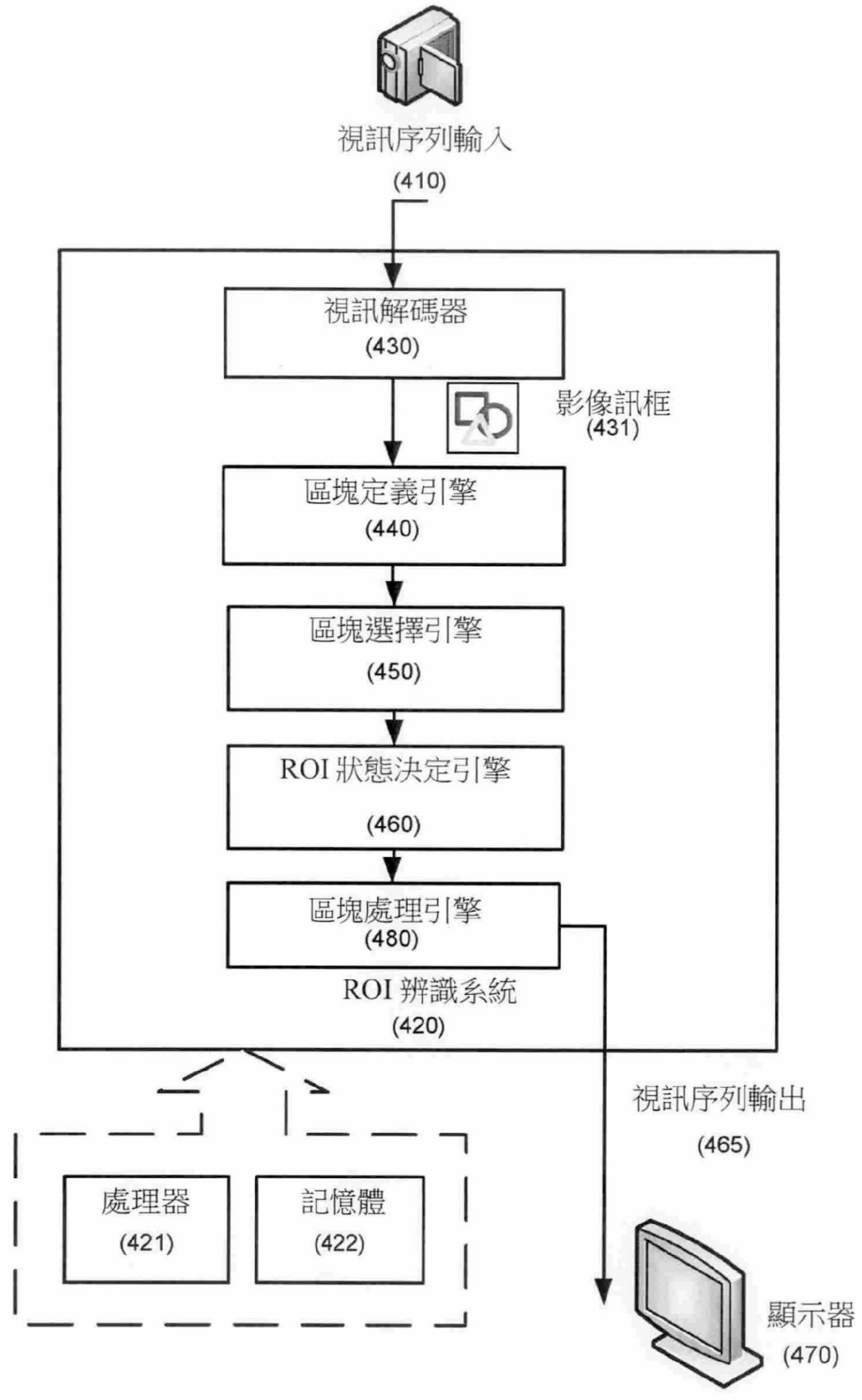


第二圖

300



第三圖



第四圖