



(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 I496090 B

(45)公告日：中華民國 104 (2015) 年 08 月 11 日

(21)申請案號：101132330

(22)申請日：中華民國 101 (2012) 年 09 月 05 日

(51)Int. Cl. : G06K9/00 (2006.01)

G06K9/62 (2006.01)

G06T7/00 (2006.01)

(71)申請人：財團法人工業技術研究院(中華民國) INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE (TW)

新竹縣竹東鎮中興路 4 段 195 號

(72)發明人：李家昶 LI, CHIA CHANG (TW)；廖惇利 LIAO, DUAN LI (TW)

(74)代理人：洪堯順

(56)參考文獻：

TW 201044322A

TW 201142719A

US 2009/0213219A1

US 2011/0311108A1

審查人員：高元良

申請專利範圍項數：19 項 圖式數：11 共 33 頁

(54)名稱

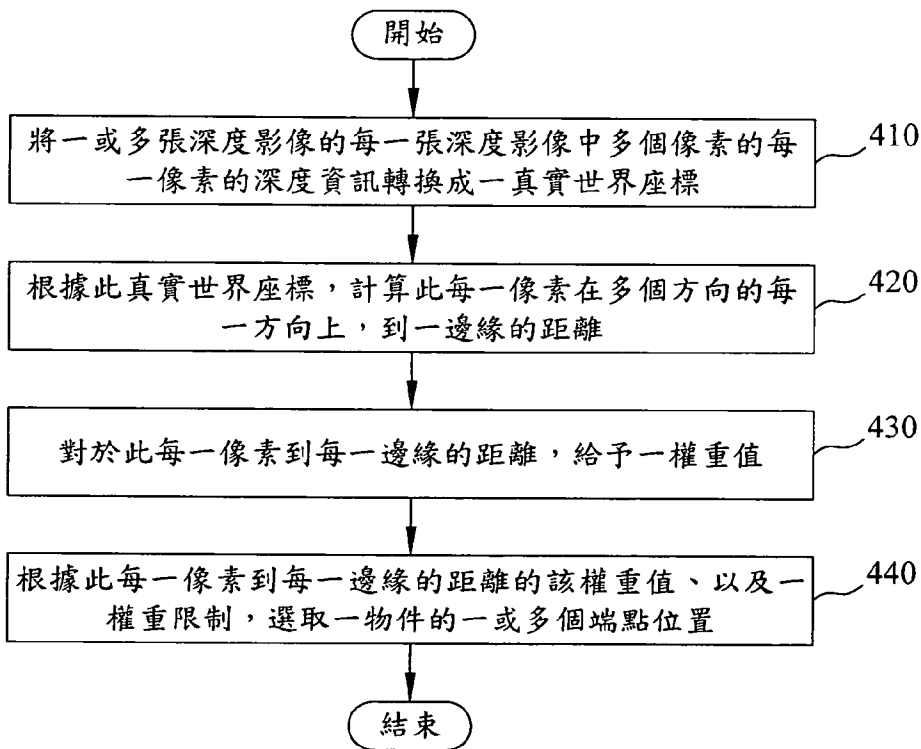
使用深度影像的物件定位方法與裝置

METHOD AND APPARATUS FOR OBJECT POSITIONING BY USING DEPTH IMAGES

(57)摘要

根據一種使用深度影像的物件定位方法的一實施例，藉由一硬體處理器來執行：將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標；根據此真實世界座標，計算此每一像素在多個方向的每一方向上，到一邊緣的距離；對於此每一像素到每一邊緣的距離，給予一權重值；以及根據此每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置。

According to an exemplary embodiment, a method for object positioning by using depth images is executed by a hardware processor as following: converting depth information of each of a plurality of pixels in each of one or more depth images into a real world coordinate; based on the real world coordinates, computing a distance of each pixel to an edge in each of a plurality of directions; assigning a weight to the distance of each pixel to each edge; and based on the weight of the distance of each pixel to each edge and a weight limit, selecting one or more terminal positions of an object.



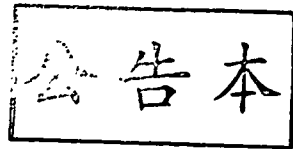
410 . . . 將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標

420 . . . 根據此真實世界座標，計算此每一像素在多個方向的每一方向上，到一邊緣的距離

430 . . . 對於此每一像素到每一邊緣的距離，給予一權重值

440 . . . 根據此每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置

第四圖



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： 101132330

※ 申請日： 101. 9. 0 5

※IPC 分類： G06K 9/00 (2006.1)
G06K 9/62 (2006.1)
G06T 7/00 (2006.1)

一、發明名稱：(中文/英文)

使用深度影像的物件定位方法與裝置/

METHOD AND APPARATUS FOR OBJECT POSITIONING BY
USING DEPTH IMAGES

二、中文發明摘要：

根據一種使用深度影像的物件定位方法的一實施例，藉由一硬體處理器來執行：將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標；根據此真實世界座標，計算此每一像素在多個方向的每一方向上，到一邊緣的距離；對於此每一像素到每一邊緣的距離，給以一權重值；以及根據此每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置。

三、英文發明摘要：

According to an exemplary embodiment, a method for object positioning by using depth images is executed by a hardware processor as following: converting depth information of each of a plurality of pixels in each of one or more depth images into a real world coordinate; based on the real world coordinates, computing a distance of each pixel to

an edge in each of a plurality of directions; assigning a weight to the distance of each pixel to each edge; and based on the weight of the distance of each pixel to each edge and a weight limit, selecting one or more terminal positions of an object.

四、指定代表圖：

(一)本案指定代表圖為：第(四)圖。

(二)本代表圖之元件符號簡單說明：

410 將一或多張深度影像的每一張深度影像中多個像素的
每一像素的深度資訊轉換成一真實世界座標

420 根據此真實世界座標，計算此每一像素在多個方向的
每一方向上，到一邊緣的距離

430 對於此每一像素到每一邊緣的距離，給予一權重值

440 根據此每一像素到每一邊緣的距離的該權重值、以及
一權重限制，選取一物件的一或多個端點位置

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

六、發明說明：

【發明所屬之技術領域】

本揭露係關於一種使用深度影像(depth image)的物件定位(object positioning)方法與裝置。

【先前技術】

以往，以視覺互動為基礎的人機互動系統技術使用單一攝影機或是利用色彩資訊方式來分析影像。在一些情況下，例如使用者與背景色彩相近、或是環境光源變化、或是人多時的複雜背景等，此類技術容易造成影像的辨識率不足。現有的一些技術已使用深度影像的資訊來輔助影像的分析，例如利用深度影像來追蹤使用者局部區域、或是擷取並追蹤使用者的端點位置、或是進行人體端點偵測等。有的技術如利用膚色與深度資訊來找出手部位置、或是手部區域及臉部區域等。

利用深度影像來追蹤使用者局部區域的技術例如第一圖所示，從一深度影像中找出目標物體(target)106 的邊緣，從邊緣找出最符合預先定義的輪廓形狀，例如目標物體 106 的右手輪廓形狀 104。其中，一深度影像具有一相對應的樣本邊緣位能轉換圖(sample edge potential transform map)100。例如，樣本邊緣位能轉換圖 100 包括在位能場(potential field)裡被模型化的欲追蹤的目標物體，樣本邊緣位能轉換圖 100 中每一格子點的值指出此格子點距離目標物體 106 的邊緣 101 有多遠。當目標物體與

攝影機距離變動時，輪廓形狀的大小也會隨之不同，此技術需要多個預先定義的輪廓形狀。

利用深度影像來擷取並追蹤使用者的端點(extremity)位置的技術是從深度影像產生出網格立體像素(Voxels)後，移除背景網格立體像素並區隔出使用者，再從區隔出的使用者中找出使用者的端點位置。也就是說，此技術藉由建立 3-維網格、以及移除背景區隔人體，來找出使用者的端點位置。

利用深度影像來找出使用者身體各部位端點的技術如第二圖所示，此技術從深度影像的資訊 202 產生 3-維表面網格集合(set of surface meshes)210 後，計算此表面網格集合中各網格點路徑的測地線距離長度(geodesic distances)，再根據不同路徑長度來分類此集合中的表面網格，每一同類的網格對應身體的同一部位，例如頭、手、腳、肩膀、或其他部位等。也就是說，此技術藉由產生 3-維表面網格集合、以及計算此表面網格集合中各點路徑的測地線距離，來找出使用者身體各部位的端點位置。

利用膚色與深度資訊來找出手部區域與臉部區域的技術是透過深度影像，先切割出人物區域，並透過膚色偵測切割出人物膚色區域；再透過 k-means 分類膚色區域；最後，透過深度資訊分辨其區域為手部區域或是臉部區域。利用膚色與深度資訊來找出手部位置的技術是採用 LUV

色彩空間(color space)，且搭配與高斯模型混合(Mixture of Gaussians Model)，來找出膚色區域；並以深度資訊輔助，移除背景膚色區域。於前景膚色區域中，比較任兩個區域的大小、高度、深度等資訊，找出手部位置。

還有一種技術是從深度影像中以凸點特徵(Convex Degree Feature, CDF)來分析人物上下手臂。如第三圖所示，此技術從偵測出的頭與肩膀位置，來偵測上半臂(upper arm detection)310，以及計算深度圖(depth map)中每一畫素(pixel)的凸點特徵 320。利用固定大小的內圓(inner region)與外圓(outer region)，統計深度分配的像素比例，並透過這些凸點特徵來偵測手(hand detection)330 及偵測下半臂(lower arm detection)340。由上下手臂的偵測結果來決定影像中的全手臂(full arm)位置 350。

上述技術之影像分析中，有的技術可能會因使用者離攝影裝置距離遠近不一，使得影像上局部區域的輪廓形狀大小都不相同而無法建立單一模型來進行比對。有的技術可能會因使用者前會有遮蔽物，而無法取得使用者完整骨架資訊。有的技術使用膚色資訊，會因為環境光源影響而造成辨識率降低。

因此，如何設計一種僅利用深度影像資訊、不需建立使用者骨架、以及使用真實距離資訊作為擷取特徵，的物件定位技術，來達到不受環境光線影響、不受遮蔽物影

響、只須建立單一模型即可定位不同遠近的物件，是一個重要的議題。

【發明內容】

本揭露實施例提供一種使用深度影像的物件定位方法與裝置。

所揭露的一實施例是關於一種使用深度影像的物件定位方法，應用在一物件定位裝置上。此方法藉由一硬體處理器來執行：將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標；根據此真實世界座標，計算此每一像素在多個方向的每一方向上，到一邊緣的距離；對於此每一像素到每一邊緣的距離，給予一權重值；以及根據此每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置。

所揭露的另一實施例是關於一種使用深度影像的物件定位方法，應用在一物件定位裝置上。此方法藉由一硬體處理器來執行：對於一或多張深度影像的每一張深度影像中多個像素的每一像素，計算此每一像素在多個方向的每一方向上，到一邊緣的距離；將此每一像素到每一邊緣的距離的深度資訊轉換成一真實世界座標；對於該每一像素到每一邊緣的距離，給予一權重值；以及根據該每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物

件的一或多個端點位置。

所揭露的又一實施例是關於一種使用深度影像的物件定位裝置。此裝置可包含一硬體處理器，連接至一深度影像擷取裝置。此硬體處理器還包括一真實世界座標運算模組及一計算裝置。此真實世界座標運算模組將收到的多筆深度資訊的每筆深度資訊轉換成一真實世界座標。此計算裝置被配置來計算來自此深度影像擷取裝置之多張深度影像的每一張深度影像中多個像素的每一像素在多個方向的每一方向上，到一邊緣的距離、並給予一權重值，以及根據這些距離的權重值及一權重限制，選取一物件的一或多個端點位置。

茲配合下列圖示、實施例之詳細說明及申請專利範圍，將上述及本發明之其他優點詳述於後。

【實施方式】

本揭露實施例之物件定位技術是先將接收的深度影像資訊轉換成每一像素的真實世界座標(Real-World Coordinates)，然後在每一像素之周圍的多個方向上，計算此像素到多個邊緣的距離。一邊緣例如是此像素在一方向上所找出的深度差大於一預定閾值的位置。再對此像素到每一邊緣的距離，給予一權重值；根據這些距離的權重值、以及一權重限制，選取一物件的一或多個端點位置。

也就是說，本揭露之使用深度影像的物件定位方法，如第四圖所示，先將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標(步驟 410)後，根據此真實世界座標，計算此每一像素在多個方向的每一方向上，到一邊緣的距離(步驟 420);然後，對於此每一像素到每一邊緣的距離，給予一權重值(步驟 430)，再根據此每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置(步驟 440)。此物件定位方法係應用在一物件定位裝置上，並且由一硬體處理器來執行步驟 410 至步驟 440。此硬體處理器例如是，但不限定於，一具有計算能力的硬體電路、或是備有至少一中央處理單元與記憶體的一電腦系統等。

在步驟 410 中，每一像素的深度資訊例如代表左右影像的視差(此視差單位為像素距離)，此左右影像的視差是由雙攝影機所擷取的一第一影像與一第二影像的像素距離;並且此左右影像的視差被轉換成真實世界座標(此真實世界座標的單位為 cm)。在實際應用上，取得深度影像有多種方式，例如以雙攝影機架構來取得深度影像。以雙攝影機架構為例，第五圖說明如何將一像素 P 的深度資訊轉換成真實世界座標，其中假設 x^l 與 x^r 是對應在左右兩張影像的 x 座標， T 為左右兩攝影機之間的距離， f 為攝影機的焦距; X 、 Y 、 Z 分別為以一參考攝影機(例如右攝影機)為原點的真實世界座標。 Z 為像素 P 相對於攝影機的 z 軸座標。

根據第五圖，因此 Z 軸座標換算方式如下：

$$\frac{T-(x^l-x^r)}{Z-f} = \frac{T}{Z} \Rightarrow Z = \frac{fT}{x^l-x^r}$$

根據 Z 座標，可換算 X 與 Y 座標如下：

$$\begin{aligned} \frac{x^r}{f} = \frac{X}{Z} &\Rightarrow X = Z \frac{x^r}{f} \\ Z = \frac{fT}{x^l-x^r} &\Rightarrow X = \frac{x^r T}{x^l-x^r} \\ \frac{y^r}{f} = \frac{Y}{Z} &\Rightarrow Y = \frac{y^r T}{x^l-x^r} \end{aligned}$$

從一深度影像中得到每一像素的真實世界距離後，根據步驟 420，計算此每一像素在多個方向的每一方向上，到一邊緣的距離。首先，從每一像素周圍的 N 個方向找出深度落差大於預定閾值(threshold)的位置當作邊緣上的端點，然後計算且記錄此像素到這些端點的真實距離。第六圖是該像素在八個方向(即此例為 N 等於 8)上，離邊緣的距離的一範例示意圖，其中 P 點與 Q 點分別代表影像中任一像素，而其箭頭線則代表該像素在八個方向上離邊緣的距離。

換句話說，像素到邊緣的距離為，計算此每一像素在此張深度影像上，沿著一方向直線搜尋，直到搜尋到至少一像素與此每一像素之深度資訊差異超過一預定閾值為止，然後根據前述之真實世界座標來計算此每一像素與至少一像素之間的距離。以第六圖為例，從每一像素之周圍的 8 個方向(N=8)的每一方向找出深度落差大於一預定閾

值，例如設定此預定閾值為 7 公分，的像素點當作一邊緣(也就是說，可從每一像素之周圍的每一方向找出深度落差大於一預定閾值的像素點來決定該邊緣)，並且計算此每一像素與此像素點之間的距離。在本揭露中，物件的 3 維凸點的凸出程度可用來作為設定閾值的參考。例如，若待辨識(to-be-identified)物件為手部，當預定閾值為 7cm 時，代表手部的像素與其周圍的背景像素深度落差超過 7cm 時，該手部的像素可視為凸點。

計算出每一像素在 N 個方向的邊緣距離後，根據待辨識的物件，例如手部，以每一方向的真實世界座標的距離為基準，定義每一方向 n 的權重 f_n 為此方向 n 之真實世界座標的距離的函數，使其滿足像素在一正確距離時，其權重值越高，而距離差距越大時，權重值越小的效果。也就是說，以像素到邊緣的距離在一特定真實距離時，給予權重值為最大，與此特定真實距離差距越大時，則給予權重值為越小。每一像素之權重值的給予可根據待辨識物件的凸點特徵，對每一像素之周圍的不同方向給予不同的權重值。

以待辨識物件為手部為例，如第七 A 圖所示，可將手部 710 視為一個半徑 6cm 的凸面球體 720，因此可定義每一方向 n 的權重 f_n 如下，其中 $n=1,2,\dots,8$ ，代表八個方向(如凸面球體 720 之八個箭頭)的權重函數

$$f_n(d) = \begin{cases} 1 - (d-6)^3 / \text{Nom}^3, & \text{若 } d > 6 \\ 1 - (d-6)^4 / \text{Nom}^3, & \text{其他} \end{cases}$$

其中 d 為此方向 n 之像素到邊緣的距離， N_{orn} 等於 6 且為一正規化參數，當像素到邊緣的距離超過 $(6+N_{\text{orn}})$ 或小於 $(6-N_{\text{orn}})$ 時，權重值最小。權重函數 $f_n(d)$ 的分佈如第七 B 圖所示，其中橫座標代表像素到邊緣的距離 d ，縱座標代表權重 $f_n(d)$ ；此權重函數 $f_n(d)$ 的含意為，當像素到邊緣的距離 d 大於 6 公分時，權重 $f_n(d)$ 以三次方計算；當 d 等於 6 公分時，權重值最大；當 d 小於 6 公分時，權重以四次方計算，使其滿足 d 於 0~6 公分的區段時較小，並以四次方計算來增加不同距離之權重的差異。給予每一像素在周圍的每一方向的權重值後，可再加總此像素在周圍的每一方向的權重值的總和 S_w 。例如，但不限定於，八個方向之距離的權重值的總和為

$$S_w = \sum_{i=1}^8 f_n(d)$$

根據步驟 440，藉由加總後的權重值 S_w 及一權重限制，就可找出待辨識物件的位置，說明如下。首先，根據上述步驟計算出的每一像素的權重值，可將深度影像 810 轉換成權重影像 820，如第八圖所示。根據一實施例，可將每一像素的權重值正規化(normalize)，例如將每一權重值轉為 0~255 的範圍，來產生一張權重影像圖，也可再透過平滑化處理(例如以一影像積分方法計算如 20×20 區域的平均特徵權重)，以去除小區域的雜訊，而得到平滑化後權重影像 830。將每一像素的權重值正規化的計算公式的範例如下： $(S_w/N) \times 2^N$ ，其中此像素的周圍有 N 個方向。

得到平滑化後權重影像 830 後，設定一權重限制(例如預設為 100)，並且在平滑化後權重影像 830 中，選取在一特定區域範圍內有一最大權重值的像素當作待辨識物件的候選端點位置。舉例來說，由左上至右下掃描一次平滑化後權重影像 830，檢查有超過該權重限制的任一像素則列為一候選端點位置，並檢查已存在的候選端點是否在該特定區域範圍(例如 50×50)區域內;若有此候選端點，則選取其中權重值較大者為候選端點位置。第九 A 圖至第九 D 圖是一範例示意圖，說明根據像素到邊緣的距離的權重來選取待辨識物件的候選端點位置。

如第九 A 圖至第九 D 圖的範例所示，假設在掃描過程中，A 點已被列為候選端點位置(如第九 A 圖所示)，往下掃描時找到 B 點，此 B 點權重較大並且距離 A 點在該特定區域(如 50×50 區域)以內(如第九 B 圖所示)，則保留 B 點為候選端點位置並去除 A 點(如第九 C 圖所示)，最後，選取具有最大權重值的 B 點做為待辨識物件的端點位置(如第九 D 圖所示)。如果 A、B 兩點之間的距離範圍超過該特定區域(如 50×50 區域)，則 A、B 兩點皆為候選端點位置，也就是皆為待辨識物件的端點位置。也就是說，根據本揭露實施例，在一特定區域大小中，可選取具有一最大權重值的像素來決定物件的端點位置。

上述中，將深度影像中的像素從深度資訊轉換成真實

世界座標的步驟與根據真實世界座標來計算像素到邊緣的距離的步驟，此兩步驟的處理順序也可以在計算像素到邊緣的距離時，再轉換成真實世界座標的距離。也就是說，根據本揭露的另一實施例的使用深度影像的物件定位方法，如第十圖所示，可先對於一或多張深度影像的每一張深度影像中多個像素的每一像素，計算此像素在多個方向的每一方向上，到一邊緣的距離(步驟 1010)，然後再將每一像素到每一邊緣的距離的深度資訊轉換成一真實世界座標(步驟 1020)。之後再執行步驟 430 與步驟 440。

在步驟 1020 中，以一目標像素座標 (x_1, y_1, d_1) 為例，到一邊緣像素座標 (x_2, y_2, d_2) 的一真實世界距離 R_d 可透過先前描述的轉換公式計算來取得此目標像素座標的真實世界座標 (X_1, Y_1, Z_1) 、以及此邊緣像素座標的真實世界座標 (X_2, Y_2, Z_2) ，再透過如歐式距離(Euclidean distance)公式來計算其真實世界距離 R_d 。也就是下列的計算公式。

$$\begin{aligned} (X_1, Y_1, Z_1) &= \left(\frac{x_1 T}{d_1}, \frac{y_1 T}{d_1}, \frac{f T}{d_1} \right) \\ (X_2, Y_2, Z_2) &= \left(\frac{x_2 T}{d_2}, \frac{y_2 T}{d_2}, \frac{f T}{d_2} \right) \\ R_d &= \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2} \end{aligned}$$

承上述，第十一圖是根據本揭露一實施例，說明一種使用深度影像的物件定位裝置。參考第十一圖，此物件定位裝置 1100 包含一硬體處理器 1120，並且此硬體處理器 1120 連接至一深度影像擷取裝置 1110。硬體處理器 1120

還包括一真實世界座標運算模組 1122、以及一計算裝置 1124。真實世界座標運算模組 1122 分別連接至深度影像擷取裝置 1110 及計算裝置 1124，並且將收到的多筆深度資訊的每筆深度資訊轉換成一真實世界座標 1122a。計算裝置 1124 被配置來計算來自深度影像擷取裝置 1110 之多張深度影像的每一張深度影像中多個像素的每一像素在多個方向的每一方向上，到一邊緣的距離、並給予一權重值，以及根據這些距離的權重值、及一權重限制 1134，選取一物件的一或多個端點位置 1124a。

物件定位裝置 1100 還可包括深度影像擷取裝置 1110，用來擷取該多張深度影像。深度影像擷取裝置 1110 例如是，但不限定於，深度感測器(depth sensor)、或是具有雙攝影機架構的影像擷取裝置，其深度影像的像素資訊為與深度感測器的真實距離、或是由雙攝影機擷取之一第一影像與一第二影像的像素距離。每筆深度資訊係來自深度影像擷取裝置 1110 所擷取的多張深度影像、或是計算裝置 1124 算出的每一像素到每一邊緣的距離。

真實世界座標運算模組 1122 可先將深度影像擷取裝置 1110 所擷取之深度影像中的像素從深度資訊轉換成真實世界座標後，輸出至計算裝置 1124。也可以先由計算裝置 1124 算出像素到邊緣的距離後，再由真實世界座標運算模組 1122 轉換成真實世界座標的距離。如前述所載，從每一像素之周圍的每一方向找出深度落差大於一預定

閾值的像素點當作一邊緣。計算裝置 1124 如何計算每一像素在每一方向上到一邊緣的距離、如何給予一權重值，以及根據這些距離的權重值及一權重限制，如何選取一物件的一或多個端點位置等，皆已描述於前述本揭露實施例中，此處不再重述。

承上述，本揭露實施例之物件定位方法與裝置僅利用深度影像資訊，所以不受環境光線的影響；也不需建立使用者骨架，所以不受遮蔽物的影響；以真實距離資訊作為擷取特徵，只須建立單一模型即可估測與追蹤不同遠近的待辨識物件。本揭露實施例可應用在例如手勢控制系統、家電控制、互動廣告看板、3C 產業等領域的物件定位上。

綜上所述，本揭露實施例提供一種使用深度影像的物件定位方法與裝置。其技術將收到的深度影像的像素資訊轉換成真實世界三維座標資訊，並計算每一像素數個方向到邊緣的距離，再根據此數個方向的邊緣距離給予權重值，最後根據像素的權重值來決定待辨識物件的位置。此技術具有不受環境光線的影響、不受遮蔽物的影響、以及只須建立單一模型即可估測與追蹤不同遠近的待辨識物件等特點。

以上所述者僅為本揭露實施例，當不能依此限定本揭露實施之範圍。即舉凡本發明申請專利範圍所作之均等變化與修飾，皆應仍屬本發明專利涵蓋之範圍。

【圖式簡單說明】

第一圖是利用深度影像來追蹤使用者局部區域之技術的一範例示意圖。

第二圖是利用深度影像來找出使用者身體各部位端點之技術的一範例示意圖。

第三圖是從深度影像中以凸點特徵來分析人物上下手臂之技術的一範例示意圖。

第四圖是根據本揭露一實施例，說明一種使用深度影像的物件定位方法。

第五圖是根據本揭露一實施例，說明如何將一像素的深度資訊轉換成真實世界座標。

第六圖是根據本揭露一實施例，說明一像素在八個方向上，離邊緣的距離的一範例示意圖。

第七 A 圖是根據本揭露一實施例，說明以手部為辨識物件，並視為一凸面球體，以及像素在其八個方向上計算離邊緣的距離。

第七 B 圖是根據本揭露一實施例，說明一權重函數 $f_n(d)$ 的分佈。

第八圖是根據本揭露一實施例，將深度影像透過轉換及平滑化後的影像。

第九 A 圖至第九 D 圖是一範例示意圖，說明根據像素到邊緣的距離的權重來選取待辨識物件的候選端點。

第十圖是根據本揭露另一實施例，說明一種使用深度影像的物件定位方法。

第十一圖是根據本揭露一實施例，說明一種使用深度影像

的物件定位裝置。

【主要元件符號說明】

- 100 樣本邊緣位能轉換圖 106 目標物體
 101 邊緣 104 右手輪廓形狀
- 202 深度影像的資訊 210 3-維表面網格集合
- 310 偵測上半臂
 320 計算深度圖中每一畫素的凸點特徵
 330 偵測手 340 偵測下半臂
 350 決定影像中的全手臂位置
- 410 將一或多張深度影像的每一張深度影像中多個像素的
 每一像素的深度資訊轉換成一真實世界座標
 420 根據此真實世界座標，計算此每一像素在多個方向的
 每一方向上，到一邊緣的距離
 430 對於此每一像素到每一邊緣的距離，給予一權重值
 440 根據此每一像素到每一邊緣的距離的該權重值、以及
 一權重限制，選取一物件的一或多個端點位置
- x^l 與 x^r 對應在左右兩張影像的 x 座標
 f 攝影機的焦距 T 左右兩攝影機之間的距離
 P 像素
 X 、 Y 、 Z 以一參考攝影機為原點的真實世界座標
 P 、 Q 分別代表影像中任一像素

710 手部 720 凸面球體

d 像素到邊緣的距離 $f_n(d)$ 權重

810 深度影像 820 權重影像

830 平滑化後權重影像

A、B 候選端點位置

1010 對於一或多張深度影像的每一張深度影像中多個像素的每一像素，計算此像素在多個方向的每一方向上，到一邊緣的距離

1020 將每一像素到每一邊緣的距離的深度資訊轉換成一真實世界座標

1100 物件定位裝置 1110 深度影像擷取裝置

1120 硬體處理器 1122 真實世界座標運算模組

1122a 真實世界座標 1124 計算裝置

1124a 一或多個端點位置 1134 權重限制

七、申請專利範圍：

1. 一種使用深度影像的物件定位方法，應用在一物件定位裝置上，該方法藉由一硬體處理器來執行：
將一或多張深度影像的每一張深度影像中多個像素的每一像素的深度資訊轉換成一真實世界座標；
根據該真實世界座標，計算該每一像素在多個方向的每一方向上，到一邊緣的距離；
對於該每一像素到每一邊緣的距離，給予一權重值；以及
根據該每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置。
2. 如申請專利範圍第 1 項所述之方法，其中該每一像素的深度資訊代表左右影像的視差，該左右影像的視差是由雙攝影機所擷取的一第一影像與一第二影像的像素距離。
3. 如申請專利範圍第 1 項所述之方法，其中計算該每一像素在多個方向的每一方向上，到一邊緣的距離還包括：
計算該每一像素在該張深度影像上，沿著一方向直線搜尋，直到搜尋到至少一像素與該每一像素之深度資訊差異超過一預定閾值為止；以及
根據該真實世界座標，計算該每一像素與該至少一像素之間的距離。
4. 如申請專利範圍第 1 項所述之方法，其中該方法對於該每一像素在不同方向上，到該邊緣的距離給予不同的權重值。
5. 如申請專利範圍第 1 項所述之方法，其中該權重值的給

予係以該每一像素到邊緣的距離在一特定真實距離時，給予該權重值為最大，與該特定真實距離差距越大時，則給予該權重值為越小。

6. 如申請專利範圍第 1 項所述之方法，其中該選取該物件的該一或多個端點位置還包括：

在一特定區域大小中，選取具有一最大權重值的像素來決定該物件的該一或多個端點位置。

7. 如申請專利範圍第 1 項所述之方法，其中該選取該物件的該一或多個端點位置還包括：

在一平滑化後權重影像中，檢查有超過該權重限制的任一像素，則列為該物件的一候選端點位置。

8. 一種使用深度影像的物件定位方法，應用在一物件定位裝置上，該方法藉由一硬體處理器來執行：

對於一或多張深度影像的每一張深度影像中多個像素的每一像素，計算該每一像素在多個方向的每一方向上，到一邊緣的距離；

將該每一像素到每一邊緣的距離的深度資訊轉換成一真實世界座標；

對於該每一像素到每一邊緣的距離，給予一權重值；以及

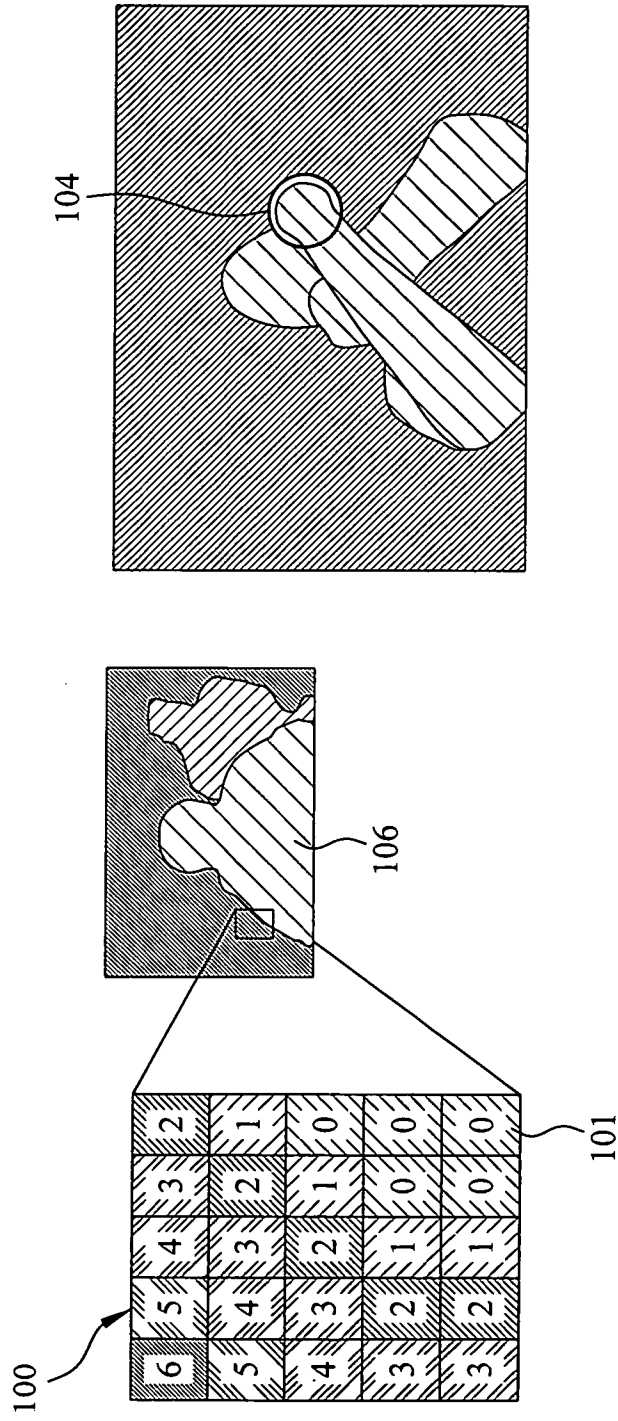
根據該每一像素到每一邊緣的距離的該權重值、以及一權重限制，選取一物件的一或多個端點位置。

9. 如申請專利範圍第 8 項所述之方法，其中該方法對於該每一像素在不同方向上，到該邊緣的距離給予不同的權重值。

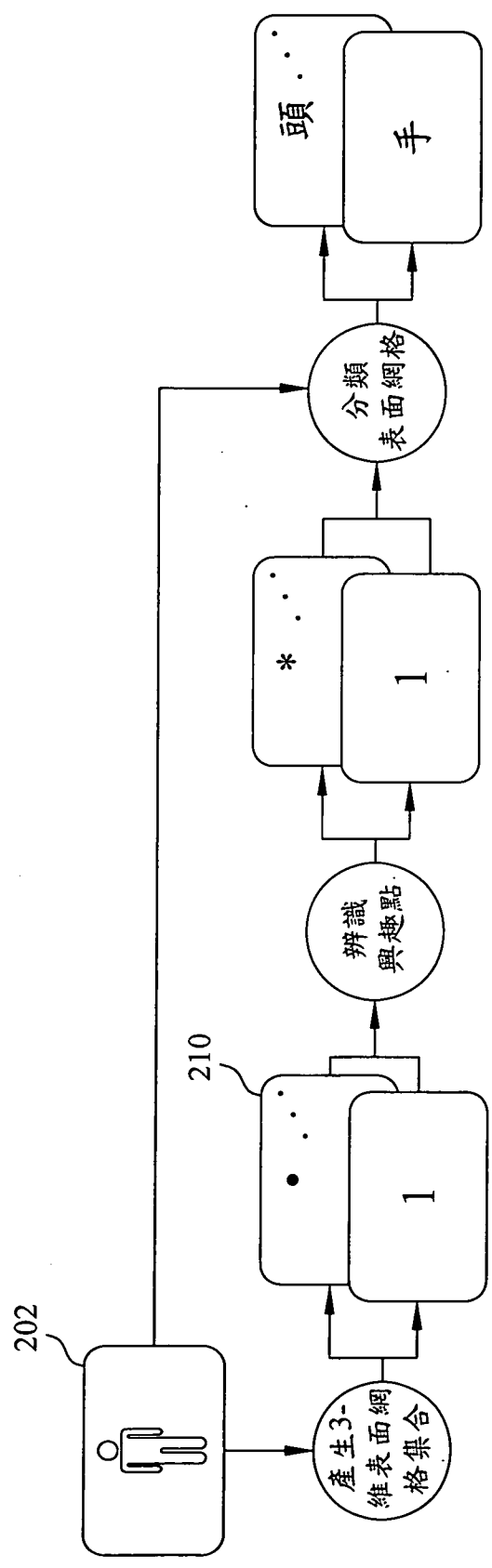
10. 如申請專利範圍第 8 項所述之方法，其中該權重值的給予係以該每一像素到邊緣的距離在一特定真實距離時，給予該權重值為最大，與該特定真實距離差距越大時，則給予該權重值為越小。
11. 如申請專利範圍第 8 項所述之方法，其中從該每一像素之周圍的每一方向找出深度落差大於一預定閾值的像素點來決定該邊緣。
12. 一種使用深度影像的物件定位裝置，包含：
 - 一硬體處理器，連接至一深度影像擷取裝置，且該硬體處理器還包括：
 - 一真實世界座標運算模組，將收到的多筆深度資訊的每筆深度資訊轉換成一真實世界座標；及
 - 一計算裝置，被配置來計算來自該深度影像擷取裝置之多張深度影像的每一張深度影像中多個像素的每一像素在多個方向的每一方向上，到一邊緣的距離並給予一權重值，以及根據該些距離的權重值及一權重限制，選取一物件的一或多個端點位置。
13. 如申請專利範圍第 12 項所述之裝置，其中該多筆深度資訊的每筆深度資訊係來自該深度影像擷取裝置所擷取的該多張深度影像、或是該計算裝置算出的該每一像素到每一邊緣的距離。
14. 如申請專利範圍第 12 項所述之裝置，其中該深度影像擷取裝置是一深度感測器、或是具有雙攝影機架構的一影像擷取裝置。
15. 如申請專利範圍第 14 項所述之裝置，其中該每一張深

- 度影像的像素資訊為與該深度感測器的真實距離。
16. 如申請專利範圍第 14 項所述之裝置，其中該每一張深度影像的像素資訊是由雙攝影機所擷取之一第一影像與一第二影像的像素距離。
 17. 如申請專利範圍第 12 項所述之裝置，其中該硬體處理器是一具有計算能力的硬體電路、以及備有至少一中央處理單元與記憶體的一電腦系統之前述兩者之其中一種。
 18. 如申請專利範圍第 12 項所述之裝置，其中該裝置還包括該深度影像擷取裝置，用來擷取該多張深度影像。
 19. 如申請專利範圍第 12 項所述之裝置，其中從該每一像素之周圍的每一方向找出深度落差大於一預定閾值的像素點來決定該邊緣。

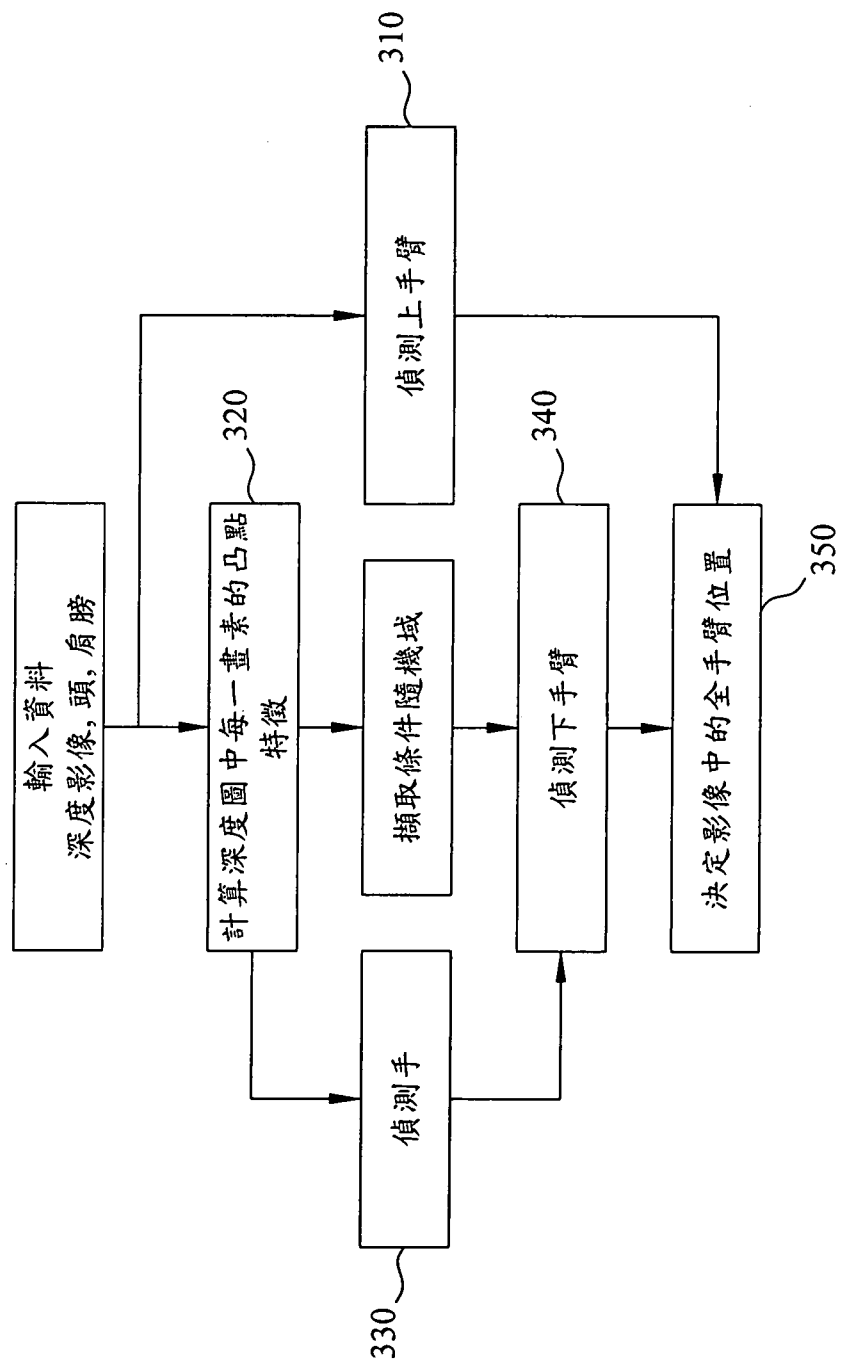
八、圖式



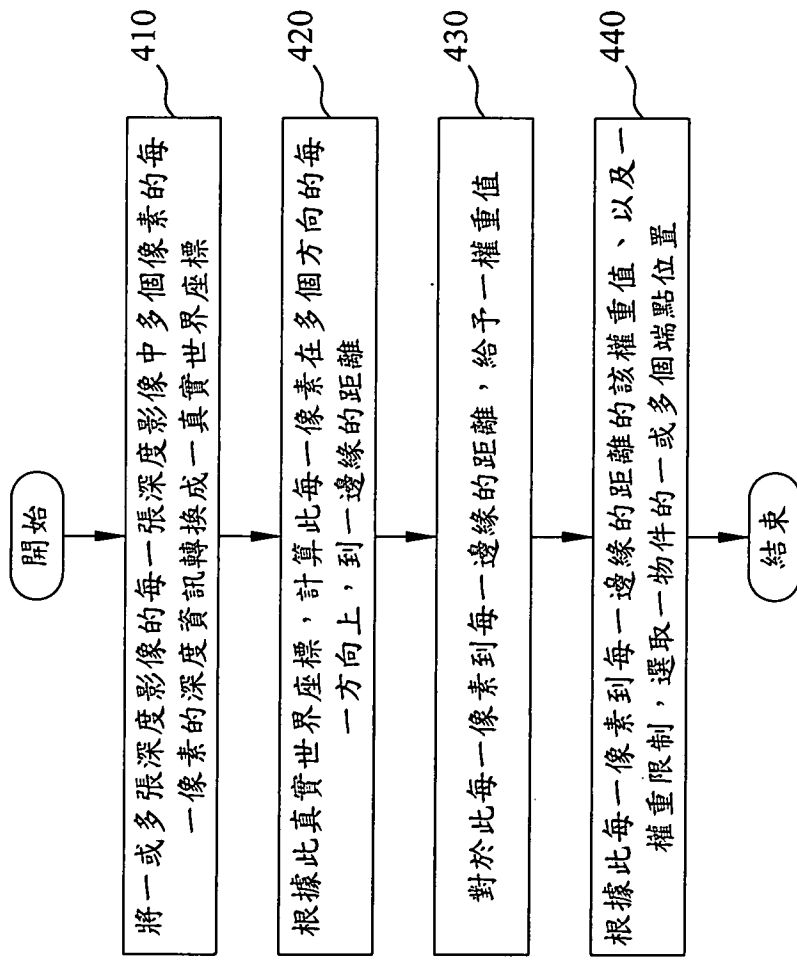
第一圖



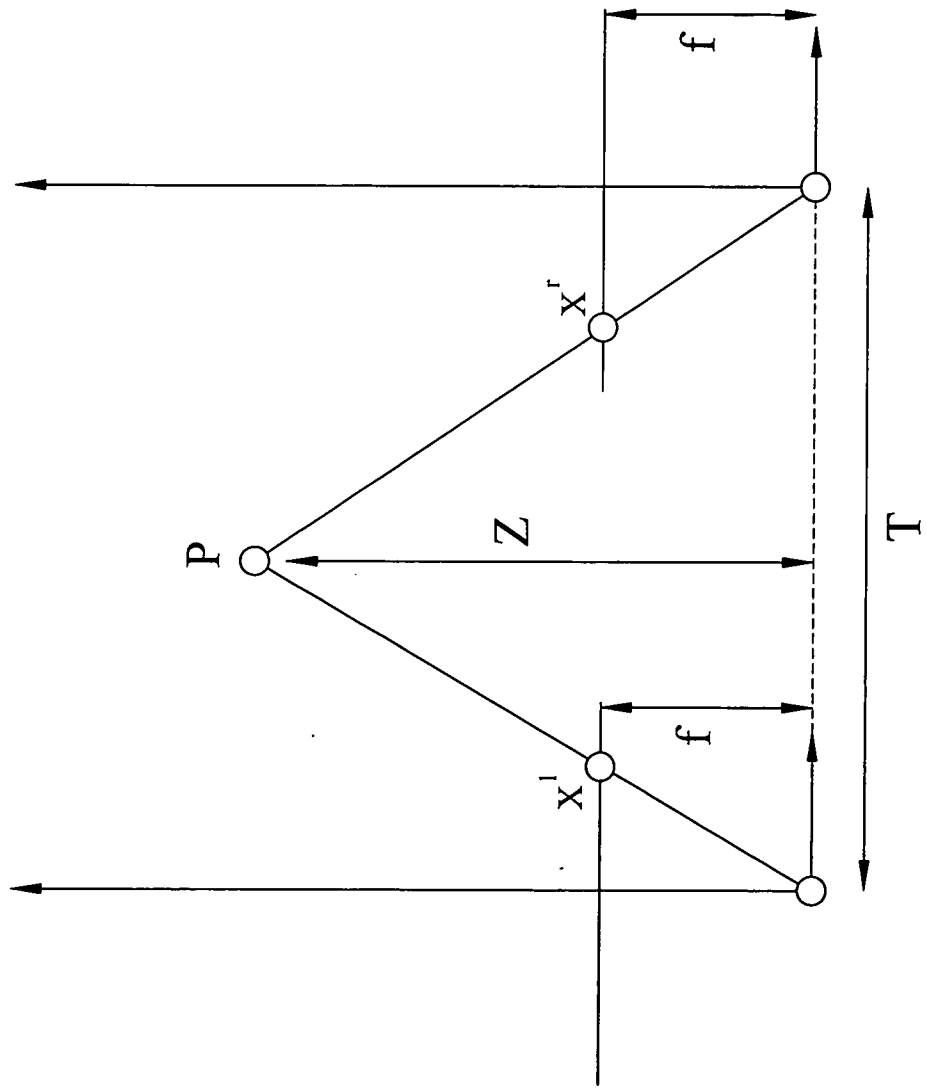
第二圖



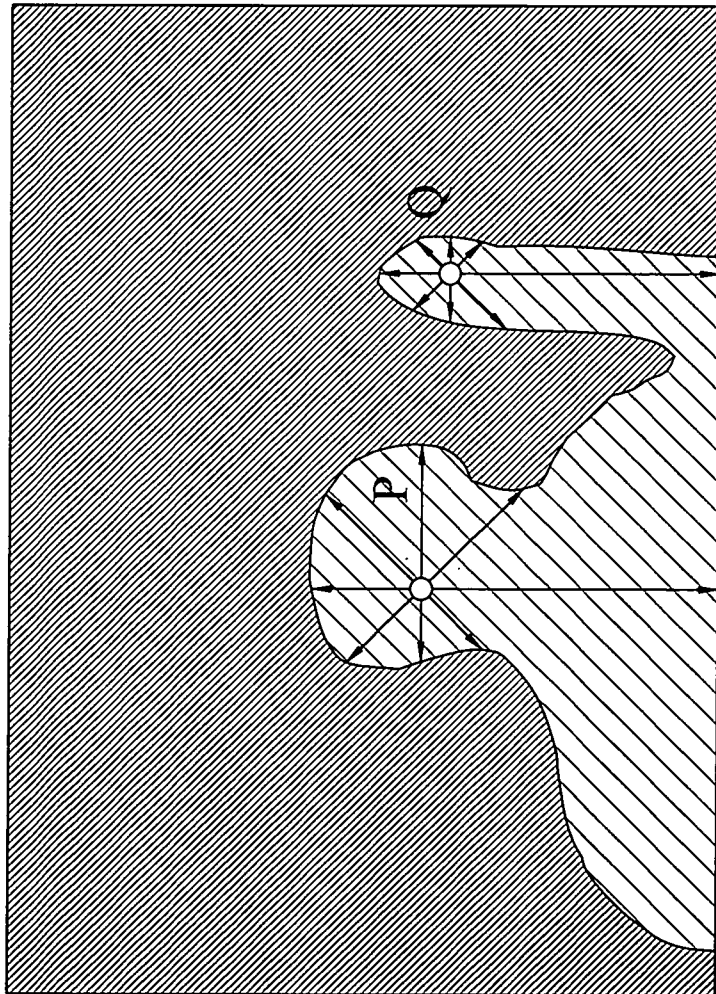
第三圖



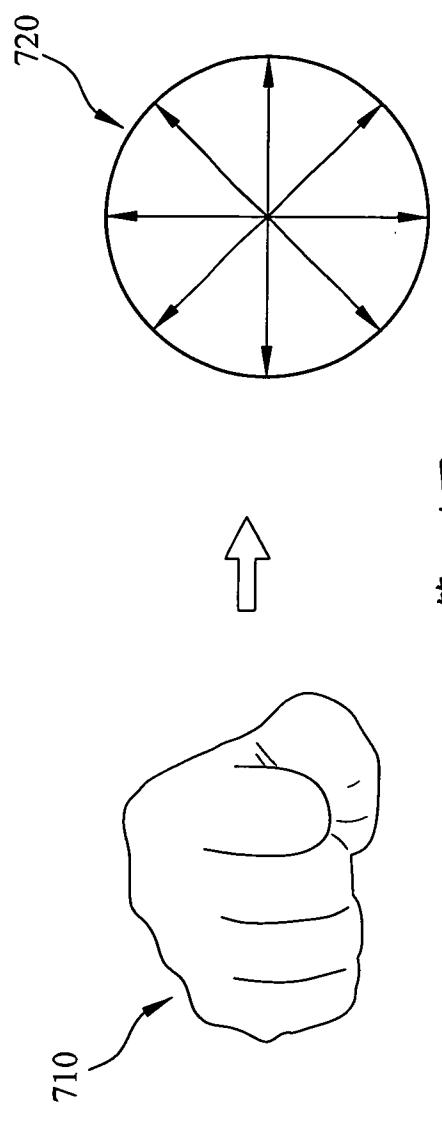
第四圖



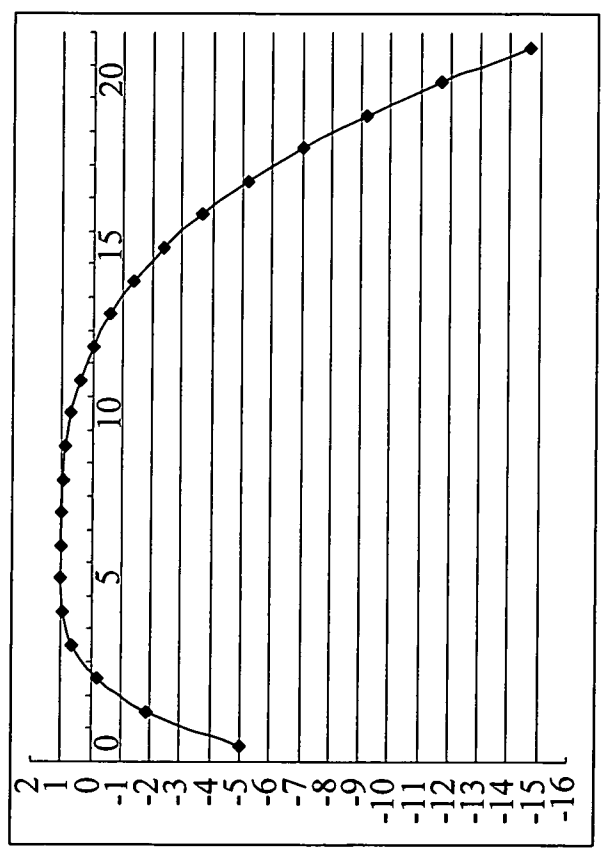
第五圖



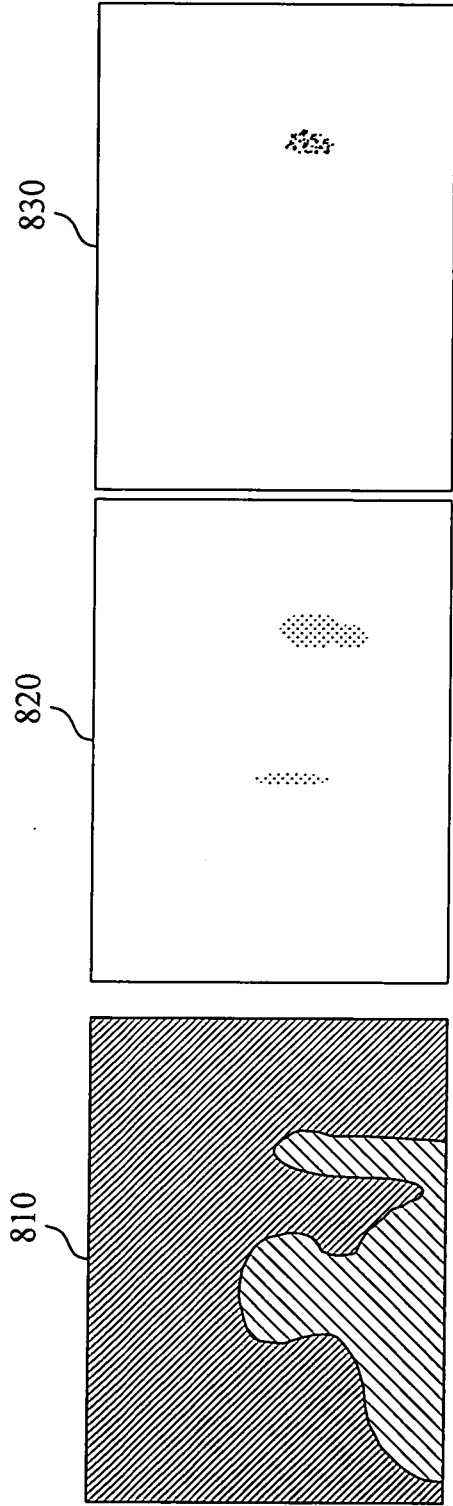
第六圖



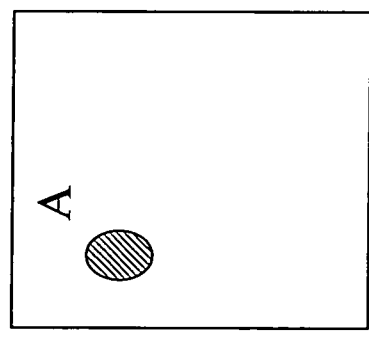
第七A圖



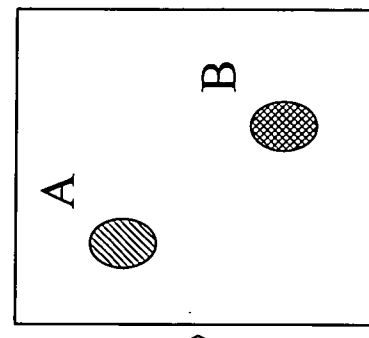
第七B圖



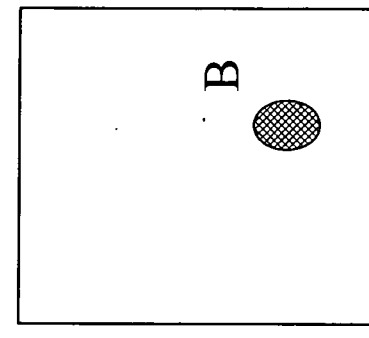
第八圖



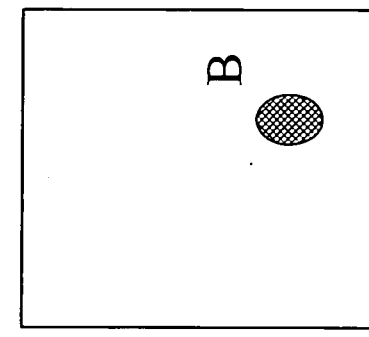
第九A圖



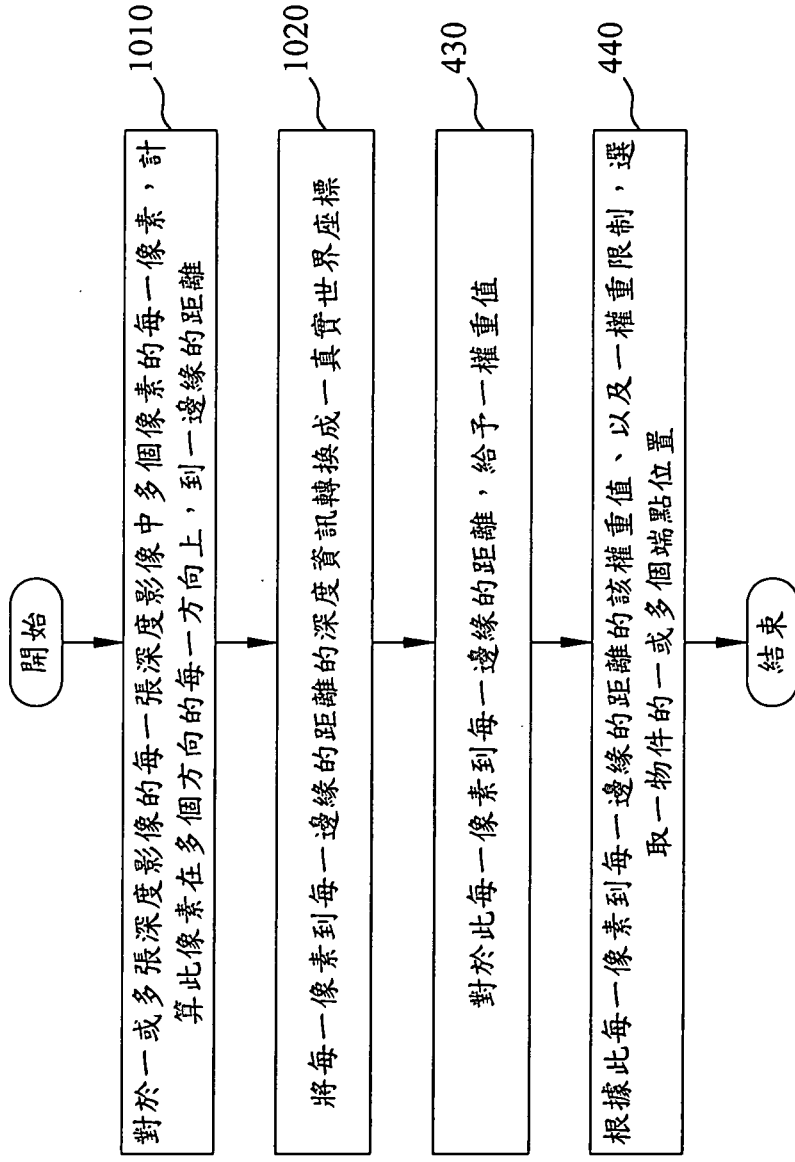
第九B圖



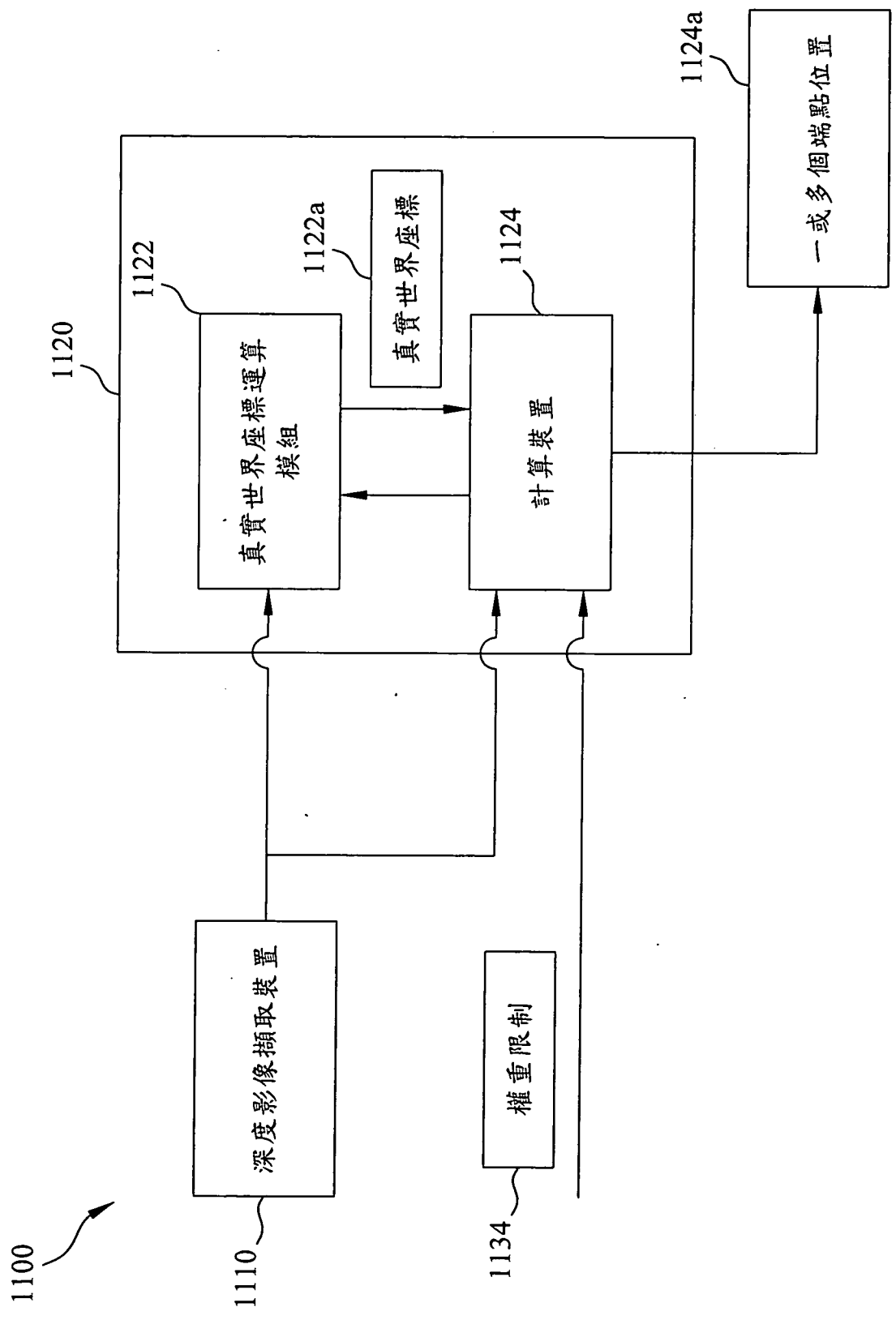
第九C圖



第九D圖



第十圖



第十一圖