

(21)申請案號：101150858

(22)申請日：中華民國 101 (2012) 年 12 月 28 日

(51)Int. Cl.：

G06F3/042 (2006.01)

G06F3/03 (2006.01)

(71)申請人：碟王科技開發股份有限公司 (中華民國) DISK KING TECHNOLOGY CO., LTD
(TW)

臺中市南屯區公益路 2 段 51 號 14 樓

國立交通大學 (中華民國) NATIONAL CHIAO TUNG UNIVERSITY (TW)

新竹市大學路 1001 號

(72)發明人：王韋弘 WANG, WEI HONG (TW)；王聖智 WANG, SHENG ZHI (TW)；曾禎宇 ZENG, ZENG YU (TW)；陳柏翔 CHEN, BO XIANG (TW)

申請實體審查：有 申請專利範圍項數：10 項 圖式數：11 共 36 頁

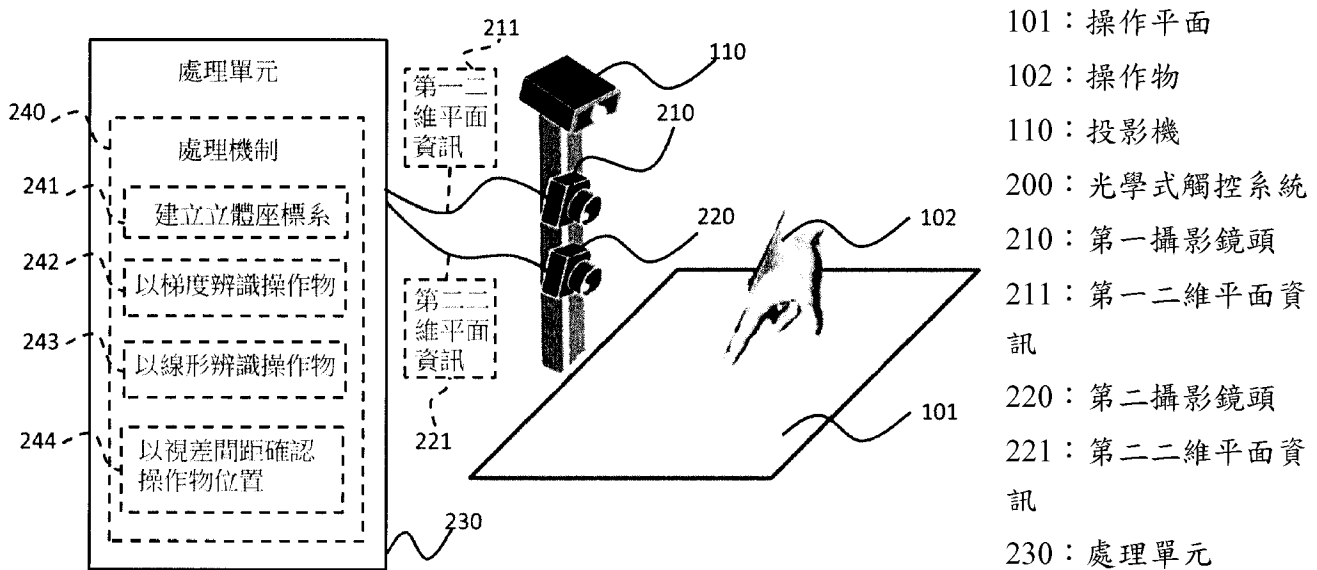
(54)名稱

光學式觸控系統

OPTICAL TOUCH SYSTEM

(57)摘要

一種光學式觸控系統，用以辨識一操作物於三維操作空間中的位置，包括一第一攝影鏡頭、一第二攝影鏡頭與一運算單元。第一攝影鏡頭與第二攝影鏡頭分別自三維操作空間中取得第一二維平面資訊與第二二維平面資訊，運算單元再以之辨識操作物的位置。運算單元係執行如下處理機制來辨識操作物的位置：首先，利用兩個二維平面資訊建立一立體座標系；然後，以兩個二維平面資訊中的梯度變化辨識操作物；接下來，以兩個二維平面資訊中的線形資訊辨識操作物；最後，以操作物於第一二維平面資訊與第二二維平面資訊中的視差間距，來定義操作物的位置。



第 2 圖

101：操作平面

102：操作物

110：投影機

200：光學式觸控系統

210：第一攝影鏡頭

211：第一二維平面資訊

220：第二攝影鏡頭

221：第二二維平面資訊

230：處理單元

240：處理機制

241-244：步驟

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：10/150 858

G06F 3/042 (2006.01)

※申請日：101.12.28

※IPC 分類：

G06F 3/03 (2006.01)

一、發明名稱：(中文/英文)

光學式觸控系統 / Optical Touch System

二、中文發明摘要：

一種光學式觸控系統，用以辨識一操作物於三維操作空間中的位置，包括一第一攝影鏡頭、一第二攝影鏡頭與一運算單元。第一攝影鏡頭與第二攝影鏡頭分別自三維操作空間中取得第一二維平面資訊與第二二維平面資訊，運算單元再以之辨識操作物的位置。運算單元係執行如下處理機制來辨識操作物的位置：首先，利用兩個二維平面資訊建立一立體座標系；然後，以兩個二維平面資訊中的梯度變化辨識操作物；接下來，以兩個二維平面資訊中的線形資訊辨識操作物；最後，以操作物於第一二維平面資訊與第二二維平面資訊中的視差間距，來定義操作物的位置。

三、英文發明摘要：

An optical touch system is disclosed to identify the location of an operating object in a three dimensional space. The system includes a first camera, a second camera and a processing unit. The first camera and the second camera obtains a first two dimensional data and a second two dimensional data from the three dimensional space

respectively, and the processing unit identifies the location of the operating object by using the data. In detail, the processing unit executes a processing mechanism described as following to calculate the location. First, a three dimensional coordination is established by using the two dimensional data. Second, each of the gradient variation in the two dimensional data is applied to identify the location of the operating object. Third, each of the shape information in the two dimensional data is applied to identify the location of the operating object. Forth, the parallax pitch of the operating objects in the two dimensional data is applied to identify the actual location of the operating object in the three dimensional space.

四、指定代表圖：

(一)本案指定代表圖為：第(2)圖。

(二)本代表圖之元件符號簡單說明：

110：投影機

101：操作平面

102：操作物

200：光學式觸控系統

210：第一攝影鏡頭

211：第一二維平面資訊

220：第二攝影鏡頭

221：第二二維平面資訊

230：處理單元

240：處理機制

241-244：步驟

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】

本發明是有關於一種觸控技術，特別是關於一種光學觸控技術。

【先前技術】

請參考第 1 圖，第 1 圖是習知之桌上型光學觸控裝置的操作原理示意圖。第 1 圖中，桌上型光學觸控裝置 100 主要包括一投影機 110、一紅外線接收鏡頭 120 與一紅外線發射器 130。其原理如下：

首先，投影機 110 在桌面上定義出一個操作平面 101，操作平面 101 上可以顯示出投影機 110 所要展現的多媒體內容。然後，紅外線發射器 130 在操作平面 101 上鋪設出一層紅外線光幕；接下來，當操作物 102，例如手指或筆，在操作平面 101 上動作時，紅外線會擊中操作物 102 然後反射到紅外線接收鏡頭 120。因此，紅外線接收鏡頭 120 可以取得一個平面上有一個特定位置是紅外線亮點的資訊，此特定位置即為操作物 102 的所在位置。然而，這種以紅外光偵測操作物 102 位置的觸控技術，容易受到外在環境光中的紅外線干擾。

【發明內容】

故，本發明之發明人有鑑於上述缺失，乃蒐集相關資料，經由多方評估及考量，並以從事於此行業累積之多年經驗，經由不斷試作及修改，始設計出此一光學式觸控系統。

依據本發明一實施方式，提出一種光學式觸控系統，係用以辨識一操作物於一三維操作空間中的位置，其包括一第一攝影鏡頭、一第二攝影鏡頭與一運算單元。第一攝影鏡頭係用以自三維操作空間中取得一第一二維平面資訊；第二攝影鏡頭係用以自三維操作空間中取得一第二二維平面資訊。運算單元利用第一二維平面資訊與第二二維平面資訊，辨識操作物的位置；其中，運算單元係執行一處理機制來辨識操作物的位置，處理機制包括下列步驟：

步驟一，利用第一二維平面資訊與第二二維平面資訊建立一立體座標系；步驟二，以第一二維平面資訊與第二二維平面資訊中的梯度變化辨識操作物；步驟三，以第一二維平面資訊與第二二維平面資訊中的線形資訊辨識操作物；步驟四，以操作物於第一二維平面資訊與第二二維平面資訊中的視差間距，來定義操作物的位置。

值得注意的是，在本發明其他實施方式中，運算單元於執行處理機制時，在以第一二維平面資訊與第二二維平面資訊中的梯度變化辨識操作物中，係事先建立操作物之一方向梯度圖(HOG)模型，且將方向梯度圖(HOG)模型當作過濾器，分別在第一二維平面資訊與第二二維平面資訊中辨識操作物。

舉例來說，上述方向梯度圖(HOG)模型可由一本過濾器(root filter)與多個部分過濾器(part filters)來組成。在可選擇的設計上，多個部分過濾器(part filters)可各自有相對應之變形成本，以定錨

(anchor)在本過濾器(root filter)之周邊。

另一方面，運算單元在執行處理機制時，以第一二維平面資訊與第二二維平面資訊中的線形資訊，辨識操作物時，可事先建立操作物之一線形模型來描述操作物的輪廓，進而分別在第一二維平面資訊與第二二維平面資訊中辨識該操作物。例如，運算單元在處理第一二維平面資訊與第二二維平面資訊中的線形資訊時，可將多個破碎的邊緣鍊以至少一段橋，修補連接成至少一輪廓段，進而與基於操作物的輪廓所預設的線形模型進行相似權重的投射比對。

從另一個角度觀之，在本發明其他實施方式中，當前述立體座標系、梯度變化辨識及線形資訊辨識功能完備後，當一物件介入第一二維平面資訊與第二二維平面資訊，且偵測此物件之視差間距介於一上限值與一下限值間，即可執行梯度變化辨識及線形資訊辨識功能，以判定物件是否為操作物。

申言之，若物件之視差間距低於下限值，可直接判定物件非操作物；若物件之視差間距高於下限值，可直接判定物件未觸發一觸控事件，而不論其為操作物與否。

【實施方式】

為達成上述目的及功效，本發明所採用之技術手段及構造，茲繪圖就本發明各實施方式詳加說明其特徵與功能如下，俾利完全了解。

請參閱第 2 圖，第 2 圖是本實施方式之光學式觸控系統 200 的結構示意圖。第 2 圖中，光學式觸控系統 200 包括第一攝影鏡頭 210、第二攝影鏡頭 220 與一處理單元 230。第一攝影鏡頭 210 從操作平面 101 與操作物 102 所在的三維立體空間中，取得一第一二維平面資訊 211，第二攝影鏡頭 220 則從相同的三維立體空間中取得一第二二維平面資訊 221。然後，處理單元 230 再根據第一二維平面資訊 211 與第二二維平面資訊 221，偵測出操作物 102 的具體位置。更進一步的說明之，處理單元 230 必須先建立一個三維系統來對應操作平面 101 與操作物 102 所在的立體操作空間，然後再分析第一二維平面資訊 211 與第二二維平面資訊 221 的資訊來確認操作物 102 的位置，進而達到觸控的效果。

具體而言，處理單元 230 必須執行一處理機制 240，處理機制 240 至少包括下列步驟：

首先執行步驟一，建立立體座標系程序 241：利用第一二維平面資訊 211 與第二二維平面資訊 221 間的差異，來建立一個三維的立體座標系。舉例來說，對三維直角坐標系 xyz 而言，第一二維平面資訊 211 與第二二維平面資訊 221 是從不同的 z 軸高度位置，取得的相同 $x-y$ 兩軸平面資訊。因此，利用第一二維平面資訊 211 與第二二維平面資訊 221 的像差，即可建立上述的立體座標系。更詳細的立體座標系建立技術，將於後面舉例說明之。

然後執行步驟二，以梯度辨識操作物程序 242：在三維的立體座標系已被建立妥善後，分別從有操作物 102 介入的第一二維平

面資訊 211 與第二二維平面資訊 221 中，辨識出操作物 102 在三維的立體座標系中的位置。此處，操作平面 101 上的背景畫面可以是主動式的畫面，例如液晶顯示屏，也可以是被動式的投影畫面，例如由投影機 110 所界定具有多媒體內容的操作平面 101。以第一攝影鏡頭 210 取得的第一二維平面資訊 211 為例，當操作物 102 出現在操作平面 101 上時，處理單元 230 在本程序中，需於第一二維平面資訊 211 中，利用物件邊框的顏色梯度變化，界定出個別物件的位置與形狀；藉此，符合預設操作物 102 形狀之物件，即可被鎖定進而取得位置資訊。當然，第二二維平面資訊 221 亦比照辦理；更詳細的以梯度辨識操作物技術，將於後面舉例說明之。

接下來執行步驟三，以線形辨識操作物程序 243：在上述以梯度辨識操作物程序 242 中，純以梯度變化來界定出符合預設操作物 102 形狀之物件時，有相當程度的誤判可能性，亦即將操作平面 101 上的背景畫面中形似操作物 102 的其他物件，誤認為操作物 102。因此，處理單元 230 在本程序中，進一步利用在前述程序 242 中各個疑似操作物 102 形狀之物件的線形，作進一步的判定，利用預設操作物 102 線形資訊，排除基於梯度資訊而類似操作物 102 的物件。當然，第二二維平面資訊 221 亦比照辦理；更詳細的以線形辨識操作物技術，將於後面舉例說明之。

最後執行步驟四，以視差間距確認操作物位置程序 244：請一併參考第 3 圖，第 3 圖是本實施方式中，以視差間距確認操作物

位置的原理示意圖。從第一二維平面資訊 211 與第二二維平面資訊 221 中分別取出操作物 102 位置後，處理單元 230 在本程序中，即可利用兩者間的視差間距界定出操作物 102 的真實位置。具體而言，如第 3 圖所繪示，以手指作為操作物 102 時，左邊為指尖碰觸投影面，右邊為指尖尚未碰觸投影面。無論指尖是否碰觸投影面，對上方的第一攝影鏡頭 210 的俯視視野來說，指尖中心幾乎會落在投影面相同的位置，但是對下方的第二攝影鏡頭 220 而言，指尖碰觸投影面與尚未碰觸投影面會造成視野觀測的指尖中心分別落在投影面上不同的位置；兩者間產生的距離即為視差間距，若操作物 102 觸及平面，視差間距 201 是小於一個臨界值的；反之，若操作物 102 遠離平面，亦即非觸控狀態，視差間距 202 必大於上述臨界值。因此，本程序即可由此判斷操作物 102 是否處於觸控狀態，以決定是否利用操作物 102 的位置執行各種觸控功能。

值得注意的是，從第 3 圖所繪示之左右兩邊的狀態，可以發現指尖碰觸投影面時，上方的第一攝影鏡頭 210 所觀測的指尖中心和下方的第二攝影鏡頭 220 所觀測的指尖中心落於投影面的位置還是有些許的差異，事實上這是為何定義觸控點為指尖中心而不採用指尖最前端的原因。指尖中心是有一定的厚度，所以當指尖碰觸投影面時，上方的第一攝影鏡頭 210 與下方的第二攝影鏡頭 220 所決定的指尖中心經過映射運算(Homography)之後還是有預期的落差存在，這是因為指尖中心並不是完全的落在平面上。

由於本實施方式只採用可見光攝影機，所以投影畫面的內容對指尖偵測有直接的影響，當投影的影像出現類似手指的東西或是影像中包含真正的手指，會造成系統誤判，所以利用有厚度的指尖中心當作觸控點，可以有效避免畫面中出現的物體造成的誤判，亦即畫面中的手指影像反而會因為對應得太完美而不被當成操作物 102。

易言之，處理機制 240 還可包括步驟五如下：利用第一攝影鏡頭 210 與第二攝影鏡頭 220 所取得之影像間的落差，可以建議一個立體系統；雖然，在後續的使用上，兩鏡頭對同一物件取值仍有落差，但可以糾偏校正之。經校正後，若一物件之視差間距 201 為零，反而表示其非操作物 102，而是背景畫面上的影像物件。考量糾偏校正無法百分之百，本實施方式可定義操作物 102 之視差間距 201 必須大於一下限值。

另一方面，隨著操作平面 101 與操作物 102 所在的立體操作空間縱深變化，會因為第一攝影鏡頭 210 與第二攝影鏡頭 220 對指尖中心斜投影角度的不同，而無法使用固定的臨界值來判斷觸碰事件。易言之，雖然指尖離投影面的高度是相同的，卻因為第一攝影鏡頭 210 與第二攝影鏡頭 220 的位置不同而出現不同的視差，因此本實施方式另提出了動態臨界值方案來解決這樣的問題；亦即，判定指尖是否碰觸到投影面是由下列方程式(1)所決定：

$$TH_{low}(x, y) < D_{is}(x, y) < TH_{high}(x, y), \forall (x, y) \in \Gamma \quad (1)$$

其中， $TH_{low}(x, y)$ 和 $TH_{high}(x, y)$ 為動態臨界值的上下界，會隨著指尖

中心出現在投影面不同的位置而改變。理想上， $TH_{low}(x, y)$ 在任何位置應該都是 0，因為映射運算(Homography)可以把平面上出現的點，幾乎沒有視差的在兩個視野之間對應；但實際上，卻必須考量尋找指尖中心時可能有些許誤差。若定義指尖碰觸投影面時指尖中心在兩個視野之間的視差為 $D_{touch}(x, y)$ ，本實施方式採用

$TH_{low}(x, y) = 0.5D_{touch}(x, y)$ 及 $TH_{high}(x, y) = 1.5D_{touch}(x, y)$ 。事實上，

$TH_{high}(x, y)$ 的值影響了觸控的精準度，也就是指尖離投影面多近會被判定為觸碰事件；若指尖中心的位置愈精準則可以將 $TH_{high}(x, y)$ 改為愈接近 $D_{touch}(x, y)$ 的值，以提供更高精準度的觸碰判定。因此，當上方視野所找到的指尖中心經過映射運算(Homography)映射到下方視野時，和下方視野找到的指尖中心相比，若差距 $D_{is}(x, y)$ 介於臨界值之間，則可以判定指尖碰觸投影面。

在實作上，建立立體座標系程序 241 已經建立了第一攝影鏡頭 210 與第二攝影鏡頭 220 之間在一個共同平面的對應關係；所以只需要在第一攝影鏡頭 210 的視野當中尋找操作物 102 可能的位置並定義可能的操作點，例如指尖，透過映射運算(Homography)即可輕易的決定當指尖碰觸投影面時，第二攝影鏡頭 220 中的指尖會落在這些可能的操作點的相對位置上，進而只需要在第二攝影鏡頭 220 視野中有興趣的位置尋找是否出現指尖中心並判斷該指尖是否碰觸投影面即可。因為實務上只在意觸碰的事件，所以當指尖尚未碰觸投影面時，由第一攝影鏡頭 210 決定的指尖中心對應到第二攝影鏡頭 220 的相對位置是不會出現指尖的，故可直

接判定為指尖尚未碰觸投影面，進而加快整體運算速度。

換句話說，在整體運作上，若第一攝影鏡頭 210 與第二攝影鏡頭 220 對一物件取得之視差間距 201 大於一上限值，表示其無論為操作物 102 與否，皆尚未接觸操作平面 101，故亦不需於循環時序中，再執行較耗費運算資源之以梯度辨識操作物程序 242 與以線形辨識操作物程序 243。整體觀之，在系統架設之初，需要先透過建立立體座標系程序 241、以梯度辨識操作物程序 242 與以線形辨識操作物程序 243 從第一攝影鏡頭 210 與第二攝影鏡頭 220 所取得之資訊中，辨識用以建立立體系統及校正各種參數之操作物 102，藉此濾除繁冗的背景資訊；待經以視差間距確認操作物位置程序 244 後，確切定義操作物 102 位置，立體系統即呈完備。後續使用時，則先執行以視差間距確認操作物位置程序 244，對第一二維平面資訊 211 與第二二維平面資訊 221 中顯然存在視差間距的物件(背景畫面的物件不會顯然存在視差間距，除非系統數據偏移)，粗估其視差間距是否介於下限值與上限值之間；若其低於下限值，表示定義為背景物件或其他干擾物(類如蚊蠅飛蟲)，而非操作物 102；若其高於上限值，表示無論其為操作物 102 與否，皆尚未發生觸控事件(Event)，無後續的觸摸操控功能可言。

更具體的說明之，以一隻手臂手掌握拳伸出食指，介入已經建立好的立體操作系統為例。從第一攝影鏡頭 210 與第二攝影鏡頭 220 取得之第一二維平面資訊 211 與第二二維平面資訊 221，可以粗略鑑別出幾個具有視差間距的物件，兩個影像間手掌或拳頭

部的視差間距會大於預設上限值，故不予處理；而兩個影像間的食指部位則應於觸摸操作平面 101 時，粗估其視差間距介於上限值與下限值間，故進一步啟動以梯度辨識操作物程序 242 與以線形辨識操作物程序 243，確認其為預設之操作物 102。反之，若手臂握拳，不伸出食指，落在操作平面 101，則其必有至少一視差間距也會介於上限值與下限值間，故須進一步由以梯度辨識操作物程序 242 與以線形辨識操作物程序 243，排除其為適格之操作物 102。

接下來，逐一介紹前述建立立體座標系程序 241、以梯度辨識操作物程序 242 與以線形辨識操作物程序 243 的技術細節。請參考第 4 圖，第 4 圖是建立立體座標系程序 241 的運算步驟示意圖。首先，左上小圖的第一步驟中，C 點與 C' 點為兩台已知的攝影機中心，其代表的物理意義包括但不限於上述的第一攝影鏡頭 210 與第二攝影鏡頭 220。平面 E 和平面 E' 分別為與 C 點與 C' 點對應的影像平面。X 點為三維空間中的一個參考點，其與攝影機中心 C 點與 C' 點共同組成一個對軸平面 π 。而且攝影機中心 C 點與 C' 點的連線組成一條基準線，基準線分別穿過影像平面 E 與 E'，形成對軸點 e 與 e'。另，攝影機中心 C 點與 C' 點分別連到參考點 X，各自通過影像平面 E 與 E'，形成另一組平面參考點 x 與 x'。對軸點 e 與 e' 各自與平面參考點 x 與 x' 形成對軸線 I 與 I'。

接下來，右上小圖的第二步驟中，當 X 在三維空間中移動時，對軸平面 π 只是繞著由攝影機中心決定的基準線做旋轉而已。對

攝影機中心 C 點與 C' 點而言， X 落在 E 平面的 x 位置及 E' 平面的 x' 位置。所以，當三維空間中的 X 沿著 $c \rightarrow x$ 方向移動時，對攝影機中心 C 點來說，參考點 X 還是落在相同的 x 位置，但對於攝影機中心 C' 點而言，參考點 X 會在對軸平面 π 與影像平面 E' 交集的對軸線 I' 上移動；也就是說，當三維空間中的一個點與兩個攝影機中心共同決定一個平面時， X 如果在對軸平面 π 上移動，則只要在攝影機 C 的視野 E 決定一個點即可以在攝影機 C' 的視野 E' 對應到一條線，這樣的關係稱為對軸幾何關係。

因此，假設目前只從攝影機 C 看到一個點 x 而不知道它在三維空間中真正的位置 X ，由對軸幾何關係可以知道 X 會對應到攝影機 C' 視野的影像平面 E' 中的一條線 I' ；接下來，立體座標系程序 241 利用限制 X 在三維空間移動的位置來建立影像平面 E 與 E' 之間點與點的對應關係。藉此便可以將第一二維平面資訊 211 與第二二維平面資訊 221 整理成一個三維立體座標。

攝影機 C 的視野無法分辨 X 沿著 $c \rightarrow x$ 方向移動的位置，因此立體座標系程序 241 先利用一個不通過攝影機中心點 C 與 C' 的平面 Γ 切過 $c \rightarrow x$ 。在平面的 Γ 限制下，當攝影機 C 看到了一個三維空間的點 X 在影像平面 x 的位置時，由於 X 被限制在平面 Γ 上無法沿著 $c \rightarrow x$ 自由移動，而可以得知該點會落在攝影機 C' 視野中的 x' 而不再是一條對軸線 I' 。易言之，透過一個平面 Γ 即可建立兩台攝影機之間的點對點關係。

請再參考右下小圖的第三步驟，由三維空間中的一個平面 Γ

建立兩台攝影機之間的對應，這樣的關係稱為平面對應關係。在平面對應關係中，當攝影機 C 的視野看到 x ，將射線 $c \rightarrow x$ 延伸和平面 Γ 相交可以得知 x 是由 Γ 平面上的 x_Γ 產生的；意即其可由 $x = H_{\Gamma C} x_\Gamma$ 描述之。而另一方面， x_Γ 對攝影機 C' 的視野而言，是落在 x' ，意即其可由 $x' = H_{\Gamma C'} x_\Gamma$ 描述之。結合這兩台攝影機對一共同平面的關係，即可建立兩個影像平面之間的平面對應關係為 $x' = H_{\Gamma C'} H_{\Gamma C} x = Hx$ ；而且，因為 $H_{\Gamma C}$ 和 $H_{\Gamma C'}$ 皆是 1 對 1 映射的關係，所以 H 是可逆的矩陣，即 $x = H^{-1} x'$ 。

最後，如左下小圖的第四步驟所示，藉由平面對應關係 H 建立兩台攝影機於三維空間中一個平面的對應關係之後，便可進一步探討當三維空間中的參考點 X 遠離先前所定義的平面 Γ 時，所產生的視差。申言之，透過平面 Γ 建立攝影機 C 與 C' 之間的平面對應關係 H 之後，對於任何一個出現在影像平面 E 的點 x 都可以透過平面對應關係 H 找到它在另一個影像平面 E' 的對應點 $\tilde{x}' = Hx$ 。當 X 不在平面 Γ 上且仍然落在射線 $c \rightarrow x$ 上，則對於 C 來說 x 和 x_Γ 依然對應到相同的點 X 上。因此，透過平面對應關係 H 會將 X 映射到 \tilde{x}' 上；但是，對於 C' 而言 X 事實上應該落在對軸線 Γ' 的另一個點 x' 上， \tilde{x}' 與 x' 之間的差距即為由平面 Γ 所衍生的視差。

接下來，本實施方式以第 5-9 圖介紹以梯度辨識操作物程序 242 的技術細節。請先參考第 5 圖，第 5 圖是操作物 102 在背景畫面的操作示意圖。從第 5 圖中可觀察到，操作平面 101 上有多媒

體背景畫面存在；此時，若多媒體背景畫面係以投影方式形成，則其勢必直接在操作物 102 上產生光影干擾 103，影響色塊的判斷。更具體的來說，要偵測的操作物 102，即手指，其受到了投影機投影的畫面所干擾，因此除了背景的變化之外，操作物 102 也會隨著投影畫面的內容不同而有色彩上的改變。在動態的背景以及隨背景而改變色彩的前景中，較穩定的資訊為邊緣；因此，接下來教示如何由邊緣資訊準確的找出手指的位置。

在實作中，除了環境光源之外，投影機也產生一個動態的光源，因此會遭遇到高動態範圍的亮度變化，為了有效的使用邊緣資訊，此處採用特徵描述算子(feature descriptor)建立方向梯度圖(Histogram of Orientation Gradients, HOG)，藉以利用物體的邊緣會隨著對比的不同而在強度上有明顯差異的特徵。方向梯度算子(HOG descriptor)將影像視為許多小區域的組合並收集這些區域的特徵，透過方向的量化來處理物體產生的簡單形變，透過每一個小區域之間的正交歸化(normalization)來減輕高對比影像所產生不同邊緣的強度，詳述如下。

請一併參考第 6 圖與第 7 圖，其皆係方向梯度特徵描述運算的說明圖。方向梯度算子(HOG descriptor)主要是在計算影像的梯度，再針對方向量化之後統計每一個小區塊中的值，藉以透過量化及局部的統計來有效描述物體局部的外表和形狀。在方向梯度特徵描述運算中，首先利用差分濾波器(difference filters)在影像的每一個像素(pixel)都計算出梯度(gradient)。令 $r(x,$

$y)$ 和 $\theta(x, y)$ 分別表示強度梯度(intensity gradient)在像素 pixel (x, y) 的大小(magnitude)和取向(orientation)。每一個像素的梯度方向(gradient orientation)都利用下列方程式(2)或(3)量化成 p 個 bin(此處例示 $p=9$)：

$$B_1(x, y) = \text{round} \left(\frac{p\theta(x, y)}{2\pi} \right) \bmod p \quad (2)$$

$$B_2(x, y) = \text{round} \left(\frac{p\theta(x, y)}{\pi} \right) \bmod p \quad (3)$$

其中， B_1 代表對比敏感度(contrast sensitive)的量化方式，而 B_2 代表對比不敏感度(contrast insensitive)的量化方式。然後，令 $b \in \{0, 1, \dots, p-1\}$ ，則影像的特徵圖(feature map)由下列方程式(4)來完成(B 值表示 B_1 或 B_2)：

$$F(x, y)_b = \begin{cases} r(x, y) & \text{if } b = B(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

承上所述，函數 F 表示一個邊緣(edge)會被看成某一個特定的方向依照邊緣(edge)的強度 voting 到對應的 bin。處理完每一個像素的圖之後，再統整這些資訊，將影像分割成許多小區域(cell)的組合。透過這些邊緣的統計能夠有效的描述物體的特性。舉例來說，小區域(cell)是 8×8 的像素，每一個區域可以看成是一個維度 p 的向量。為了減少物體恰巧出現在所切割的區域之間，每一個像素的邊緣強度除了貢獻給自己位置所在的區域，也會插值(interpolation)到其他鄰近的四個區域。由於梯度的值和影像的對比有很大的關係，因此區域之間還會做正交歸化(normalization)，亦即一個小區域(cell)和鄰近的三個小區域看成是一個區

塊(block)。

接下來，令 $\delta, \gamma \in \{-1, 1\}$ 則正交歸化運算子(normalization factors)如下列方程式(5)所示：

$$N_{\delta,\gamma}(i, j) = (\|C(i, j)\|^2 + \|C(i + \delta, j)\|^2 + \|C(i, j + \gamma)\|^2 + \|C(i + \delta, j + \gamma)\|^2)^{\frac{1}{2}} \quad (5)$$

其中，每一個區域和不同的幾個鄰近區域都做一次歸化(normalize)，因此產生了四種歸化之後的值，如下列方程式(6)所示：

$$H(i, j) = \begin{pmatrix} T_{\alpha}(C(i, j)/N_{-1,-1}(i, j)) \\ T_{\alpha}(C(i, j)/N_{+1,-1}(i, j)) \\ T_{\alpha}(C(i, j)/N_{+1,+1}(i, j)) \\ T_{\alpha}(C(i, j)/N_{-1,+1}(i, j)) \end{pmatrix} \quad (6)$$

式中的 $T_{\alpha}(v)$ 代表向量 v 中的每個元素和 α 比較之後選擇較小的值。

透過這些流程，對任一張輸入的影像都能得到對應的特徵圖。最後只要利用支持向量機制(Support Vector Machine, SVM)進行監督學習運算(supervised learning method)即可得到操作物 102 的 HOG 模型 104，如第 6 圖所示。

最後，利用已知的 HOG 模型 104 在任何一張新輸入的影像是採用滑移視窗法(sliding-window)偵測操作物 102 的位置，其過程是將 HOG 模型 104 當作是一個過濾器(filter F)，這個過濾器包含不同權重的向量。而偵測的過程是利用這個過濾器(filter F) 在影像特徵圖 H 當中從左上角表示為 (x, y) 的子視窗(sub-window) 開始，採用內積的方式計算其分數，如下列方程式(7)所示：

$$\sum_{x', y'} F[x', y'] \cdot H[x + x', y + y'] \quad (7)$$

在計算完整張影像每一個位置的分數之後，即可透過簡單的門檻

值來找出欲偵測的物體。

然而，雖然 HOG 模型 104 可以偵測操作物 102 的位置，但是其對於物體的旋轉是相當敏感的。以第 7 圖為例，當背景畫面中有許多類似外型 105 存在時，在 HOG 模型 104 對物體旋轉敏感的特性下，都容易出現誤判。具體說明之，在將梯度取向(gradient orientation)voting 到最鄰近的 bin 當中時，物體旋轉時產生的 voting 結果會和原先預期的不同。現實應用中，操作物 102 常會有輕微的旋轉，這些旋轉產生的特徵圖(feature map)幾乎都不同。在計算分數時，只有和 HOG 模型 104 同樣方向的物體可以得到較高的分數，進而鑑別其為操作物 102。當旋轉的角度和 HOG 模型 104 差異越大，計算分數的結果也會得到越低的分數。因此只要操作物 102 的旋轉角度過大，計算完的分數就不足以將操作物 102 從背景中分離出來，進而在類似外型 105 處產生許多誤動作(false alarms)。

為了克服上述缺陷，本實施方式於以梯度辨識操作物程序 242 中進一步引入基於部分模型的 HOG 運算方式；細論之，由於操作物 102 通常是會有微幅變動的，因此除了原本的 HOG 模型 104 之外，另外加入一些可移動的部分模型，使之隨著操作物 102 的部分變動來改變位置。具體請參考第 8 圖，第 8 圖是基於部分模型的 HOG 運算方式的示意圖，其係利用原本的 HOG 模型 104 當做本過濾器 106(root filter)，另外加入數個部分過濾器 107(part filters)來描述物體變動較大的部分，而部分過濾器 107 的移動並

不是毫無限制的，每個部分過濾器 107 有對應的變形成本 108(deformation cost)。

具體而言，本過濾器 106(root filter)和部分過濾器 107(part filters)分別在不同的解析度下計算，也就是本過濾器 106(root filter)找到大概的物體位置再透過高解析度的部分過濾器 107(part filters)去覆蓋物體的小部分變動。此時，雖然操作物 102 的外型會改變，但是可變動的部位並不是毫無限制的移動，每一個部分過濾器 107(part filters)會有一個定錨(anchor)的位置來確保這些部分的移動是合乎常理的。

換句話說，定錨(anchor)就是可變動部位最有機會出現的位置，當部位離 a 定錨(anchor)越遠，對應的成本(cost)就會越大，如下列方程式(8)所述：

$$\text{scope}(P_0, \dots, P_n) = \sum_{i=0}^n F_i' \cdot \phi(H, p_i) - \sum_{i=1}^n d_i \cdot (dx_i, dy_i, dx_i^2, dy_i^2)$$

$$\text{where } (dx_i, dy_i) = (x_i, y_i) - (2(x_0, y_0) + v_i) \quad (8)$$

其中， p_0 表示本過濾器 106，而 p_1, \dots, p_n 表示 n 個部分過濾器 107，函數 $F_i' \cdot \phi(H, p_i)$ 表示將特徵圖 H 中左上角為 p 的子視窗(sub-window)中的特徵向量排成一行；而 F_i' 則是 F 排成一行之後的轉置。每一個部份模型由三個部分組成，以 (F_i, v_i, d_i) 表示，其中 F_i 是第 i 個部分的過濾器； v_i 是一個二維的向量，用以表示第 i 個部分的定錨位置； d_i 則是一個四維的向量用來表示二次函數的係數，其目的是用來表示部分離開定錨時所對應的成本函數。當 $d_i = (0, 0, 1, 1)$

則第 i 個部份的變形成本只是部分和定錨之間單純二維空間的距離。

接下來，為了計算本體在每一個位置所對應到最好的部分位置，可以透過下列方程式(9)表示之：

$$\text{scope}(P_0) = \max_{p_1, \dots, p_n} \text{score}(p_0, \dots, p_n) \quad (9)$$

剩下的步驟，就跟前述 HOG 運算的步驟相同，透過門檻值的選取即可偵測出物體的位置。

最後，本實施方式以第 9-11 圖介紹以線形辨識操作物程序 243 的技術細節。承前所述，雖然採用本過濾器 106(root filter) 和部分過濾器 107(part filters) 的 HOG 模型可以準確找到操作物 102 的位置，但是以小區域(cell)為基礎的運算觀念，可能不足以將操作物 102 資訊處理到像素等級(pixel level)。

以操作物 102 為手指作例子，指尖中心的精準度會直接影響觸控點位置的準確度及觸碰的判定；為了提高系統運作的精確度，實有將指尖中心的判斷精準度提升的需求。因此，本實施方式於以線形辨識操作物程序 243 中，以線形模型(shape model)來詳細描述手指的輪廓，利用手指已知的輪廓將前述包含部分模型的 HOG 模型 104 所偵測之後的誤動作(false alarms)移除，並且利用手指的輪廓來找到指尖的中心位置。線形模型(shape model)的特色是除了可以偵測物體之外，還可以將物體的外型描述出來。由於線形模型(shape model)建構在局部的輪廓特徵(pairs of adjacent segments, PAS)，因此需先教示說明局部輪廓特徵(P

AS)的偵測基礎原理，即輪廓網段(contour segment network, CSN)如下：

首先，因為邊緣(edge)的偵測容易受到影像亮度及對比的影響，因此一條邊緣(edge)經常會被切成好幾段破碎的邊緣鍊(edge chains)。因此，連接邊緣鍊的運算方式，就是將這些被切斷的地方記錄起來。舉例來說，如果邊緣鍊 C2 中有一個點落在邊緣鍊 C1 的搜尋範圍內，則邊緣鍊 C1 會和邊緣鍊 C2 連起來；此處請參考第 9 圖，第 9 圖是兩條邊緣鍊的連接示意圖。值得注意的是，如第 9 圖所示，搜尋範圍 109 採用梯型的意義在於隨著離邊緣鍊的端點距離越遠，另一條邊緣鍊出現的位置的不確定性也隨著增加。

接下來，在連接邊緣鍊之後，須將邊緣鍊分割成近似於直線的輪廓段(contour segments)。此時，輪廓段的段橋(segments bridging)定義了其本身切線連續(tangent-continuous)的連接，段橋(segments bridging)可以有效的將斷掉的邊緣修復。輪廓網段(CSN)即主要建構在連接邊緣鍊(edge chain links)和輪廓段(contour segments)的利用上，其係以連接邊緣鍊(edge chain links)將可能遺失的邊緣資訊記錄下來，再基於輪廓段(contour segments)利用近似於直線的線段來重組整個網段(network)。具體請參考第 10 圖，第 10 圖是輪廓網段的運作示意圖，其係繪示下列六種修補的規則：

(1)如第 10 圖中規則一(Rule 1)所示：在同一個邊緣鍊上的

其中一個線段前端點會連接到另一個線段的後端點。

(2) 如第 10 圖中規則二(Rule 2)所示：當兩個邊緣鍊(C1, C2)在端點連接時，邊緣鍊 C1 的線段會連接到邊緣鍊 C2 的線段。

(3) 如第 10 圖中規則三(Rule 3)所示：考慮 T 型接面的連接時，邊緣鍊 C1 的線段會連接到邊緣鍊 C2 兩個線段的最近端點。

(4) 如第 10 圖中規則四(Rule 4)所示：令 s 為連接邊緣鍊 C1 與邊緣鍊 C2 的段橋，段橋 s 會在邊緣鍊 C1 線段的前端點之後，且在邊緣鍊 C2 線段的後端點之前。

(5) 如第 10 圖中規則五(Rule 5)所示：同一個邊緣鍊上的兩個線段如果有連續的端點，則連接在一起。

(6) 如第 10 圖中規則六(Rule 6)所示：考慮一個沒有前端點連接的段橋 s ，因其覆蓋了邊緣鍊 C2 的前端點，且延續到邊緣鍊 C2 的後段，故若邊緣鍊 C2 連接到另一個邊緣鍊 C3，則段橋 s 亦連接到邊緣鍊 C3 的線段。

承上所述，因為每一個邊緣鍊都可能連接到其他數個邊緣鍊，故輪廓網段(CSN)是一個相當複雜的網狀結構，而本體特徵(local features)可描述一對連接線段之間的關係，進而簡化整個問題的複雜度。亦即，兩兩之間的線段會被看成一個局部輪廓特徵(PAS)。接下來，說明如何利用數學方程來描述局部輪廓特徵(PAS)如下：

局部輪廓特徵(PAS) $P = (x, y, s, e, d)$ 是由一個位置 (x, y) 、一個級距參數 s 、一個強度參數 e 及一個描述子 $d = (\theta_1, \theta_2, l_1, l_2, \bar{r})$ 所組成；

請參考第 11 圖，第 11 圖是局部輪廓特徵的運算示意圖。第 11 圖中，因為局部輪廓特徵(PAS)係以線段(segments)連接邊緣鍊(edge chains)之間的空隙，因此局部輪廓特徵(PAS)可以很穩定的偵測物體。由於局部輪廓特徵(PAS)只包含單獨的兩個線段，其可以有效且單純的還原物體的邊界並去除雜亂的背景。

局部輪廓特徵(PAS)之間的差異性透過 $D(P, Q)$ 來計算， $D(P, Q)$ 表示兩個 PAS 描述子 d_p, d_q 之間的相異程度，如下列方程式(10)：

$$D(d^p, d^q) = \omega_r \|r_i^p - r_i^q\| + \omega_\theta \sum_{i=1}^2 D_\theta(\theta_i^p, \theta_i^q) + \sum_{i=1}^2 |\log(l_i^p / l_i^q)| \quad (10)$$

方程式(10)的第一項係計算兩個線段相對位置的差異，而第二項 $D_\theta \in [0, \pi/2]$ 則係計算線段之間的取向(orientation)差異，第三項則計算他們的長度差異。舉例來說，一個操作中的手指可能會有輕微的旋轉，因此可將方程式(10)依需求改寫為下列方程式(11)，以解決輕微的旋轉造成 PAS 之間的差異加大：

$$D(d^p, d^q) = \omega_r \|r_i^p - r_i^q\| + \omega_\theta [D_\theta(\theta_1^p - \theta_2^p) - D_\theta(\theta_1^q - \theta_2^q)] + \sum_{i=1}^2 |\log(l_i^p / l_i^q)| \quad (11)$$

在方程式(11)中，進一步考慮了兩個 PAS 特徵之間的相對取向差距。

因此，以線形辨識操作物程序 243 的線形模型(shape model)即是藉由事先定義一個想偵測的物體，例如操作物 102，利用其所包含的 PAS 特徵會重複的出現在訓練資料中，透過收集這些 PAS 特徵來建立一個碼書(code book) $C = \{t_i\}$ 。這個碼書(code book)包含了格式 $T\{t_i\}$ 的 PAS 描述子，而 V_i 為對應到 t_i 的投射空間(voti

ng space)。每一個投射空間(voting space)都是一個三維的空間，其中兩個維度做為位置，另外一個維度為特徵大小。每一個 PAS 會分配給許多格式 T，只要他們的 PAS 差異性足夠小，如下列方程式(12)所示：

$$\gamma: T = \{t_j | D(d, t_j) < \gamma\} \quad (12)$$

而投射的權重則由 PAS 和該格式的差異性決定，差異越小則權重越高，如下列方程式(13)所描述：

$$e \cdot (1 - D(d, t_j) / \gamma) \quad (13)$$

惟，以上所述僅為本發明之各種實施方式而已，非因此即侷限本發明之專利範圍，故舉凡運用本發明說明書及圖式內容所為之簡易修飾及等效結構變化，均應同理包含於本發明之專利範圍內。

【圖式簡單說明】

- 第 1 圖 係為習知之桌上型光學觸控裝置的操作原理示意圖。
- 第 2 圖 係為本實施方式之光學式觸控系統 200 的結構示意圖。
- 第 3 圖 係為本實施方式以視差間距確認操作物位置的原理示意圖。
- 第 4 圖 係為建立立體座標系程序 241 的運算步驟示意圖。
- 第 5 圖 係為操作物 102 在背景畫面的操作示意圖。
- 第 6 圖 係為方向梯度特徵描述運算的說明圖。
- 第 7 圖 係為方向梯度特徵描述運算的說明圖。

第 8 圖 係為基於部分模型的 HOG 運算方式的示意圖。

第 9 圖 係為兩條邊緣鍊的连接示意圖。

第 10 圖 係為輪廓網段的運作示意圖

第 11 圖 係為局部輪廓特徵的運算示意圖。

【主要元件符號說明】

100：桌上型光學觸控裝置	108：變形成本
110：投影機	109：搜尋範圍
120：紅外線接收鏡頭	200：光學式觸控系統
130：紅外線發射器	210：第一攝影鏡頭
101：操作平面	211：第一二維平面資訊
102：操作物	220：第二攝影鏡頭
103：光影干擾	221：第二二維平面資訊
104：HOG 模型	201、202：視差間距
105：類似外型	230：處理單元
106：本過濾器	240：處理機制
107：部分過濾器	241-244：步驟

七、申請專利範圍：

1、一種光學式觸控系統，係用以辨識一操作物於一三維操作空間中的位置，包括：

一第一攝影鏡頭，係用以自該三維操作空間中，取得一第一二維平面資訊；

一第二攝影鏡頭，係用以自該三維操作空間中，取得一第二二維平面資訊；以及

一運算單元，係用以利用該第一二維平面資訊與該第二二維平面資訊，辨識該操作物的位置；

其中，該運算單元係執行一處理機制來辨識該操作物的位置，該處理機制包括下列步驟：

步驟一：利用該第一二維平面資訊與該第二二維平面資訊建立一立體座標系；

步驟二：以該第一二維平面資訊與該第二二維平面資訊中的梯度變化，辨識該操作物；

步驟三：以該第一二維平面資訊與該第二二維平面資訊中的線形資訊，辨識該操作物；以及

步驟四：以該操作物於該第一二維平面資訊與該第二二維平面資訊中的視差間距，定義該操作物的位置。

2、如申請專利範圍第1項所述之光學式觸控系統，其中該運算單元係於執行該步驟二時，事先建立該操作物之一方向梯度圖(HOG)

模型，且將該方向梯度圖(HOG)模型當作過濾器，分別在該第一二維平面資訊與該第二二維平面資訊中辨識該操作物。

3、如申請專利範圍第 2 項所述之光學式觸控系統，其中該方向梯度圖(HOG)模型係由一本過濾器(root filter)與複數個部分過濾器(part filters)所組成。

4、如申請專利範圍第 3 項所述之光學式觸控系統，其中該複數個部分過濾器(part filters)係各自有相對應之變形成本，以定錨(anchor)在該本過濾器(root filter)之周邊。

5、如申請專利範圍第 1 項所述之光學式觸控系統，其中該運算單元係於執行該步驟三時，事先建立該操作物之一線形模型來描述該操作物的輪廓，進而分別在該第一二維平面資訊與該第二二維平面資訊中辨識該操作物。

6、如申請專利範圍第 5 項所述之光學式觸控系統，其中該運算單元在處理該第一二維平面資訊與該第二二維平面資訊中的線形資訊時，係將複數個破碎的邊緣鍊以至少一段橋，修補連接成至少一輪廓段。

7、如申請專利範圍第 1 項所述之光學式觸控系統，該處理機制更

包括：

當一物件介入該第一二維平面資訊與該第二二維平面資訊，且偵測該物件之視差間距介於一上限值與一下限值間，執行該步驟二與該步驟三，以判定該物件是否為該操作物。

8、如申請專利範圍第7項所述之光學式觸控系統，該處理機制更包括：

當該物件之視差間距低於該下限值，直接判定該物件非該操作物。

9、如申請專利範圍第7項所述之光學式觸控系統，該處理機制更包括：

當該物件之視差間距高於該下限值，直接判定該物件未觸發一觸控事件。

10、一種光學式觸控系統，係用以辨識一操作物於一三維操作空間中的位置，包括：

一第一攝影鏡頭，係用以自該三維操作空間中，取得一第一二維平面資訊；

一第二攝影鏡頭，係用以自該三維操作空間中，取得一第二二維平面資訊；以及

一運算單元，係用以利用該第一二維平面資訊與該第二二維平

面資訊，辨識該操作物的位置；

其中，該運算單元係執行一處理機制來辨識該操作物的位置，該處理機制包括下列步驟：

步驟一：利用該第一二維平面資訊與該第二二維平面資訊建立一立體座標系；

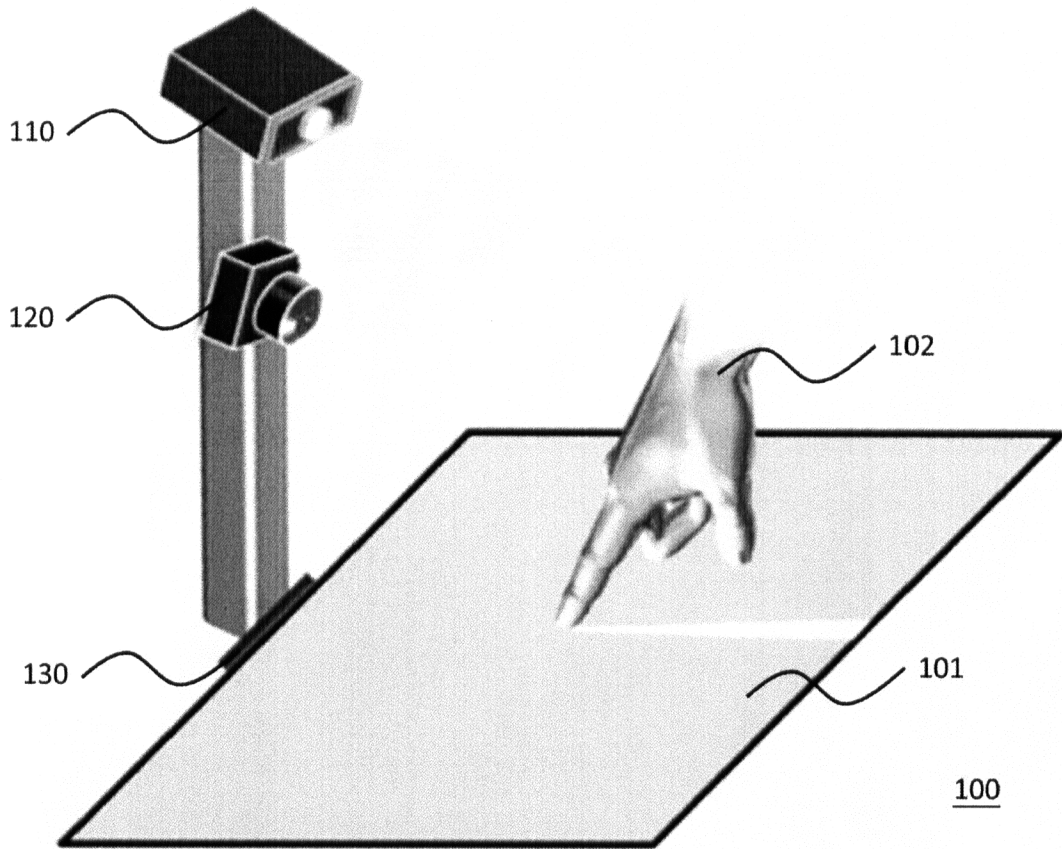
步驟二：事先建立該操作物之一方向梯度圖(HOG)模型，且將該方向梯度圖(HOG)模型當作過濾器，分別在該第一二維平面資訊與該第二二維平面資訊中辨識該操作物；

步驟三：事先建立該操作物之一線形模型來描述該操作物的輪廓，進而分別在該第一二維平面資訊與該第二二維平面資訊中辨識該操作物；

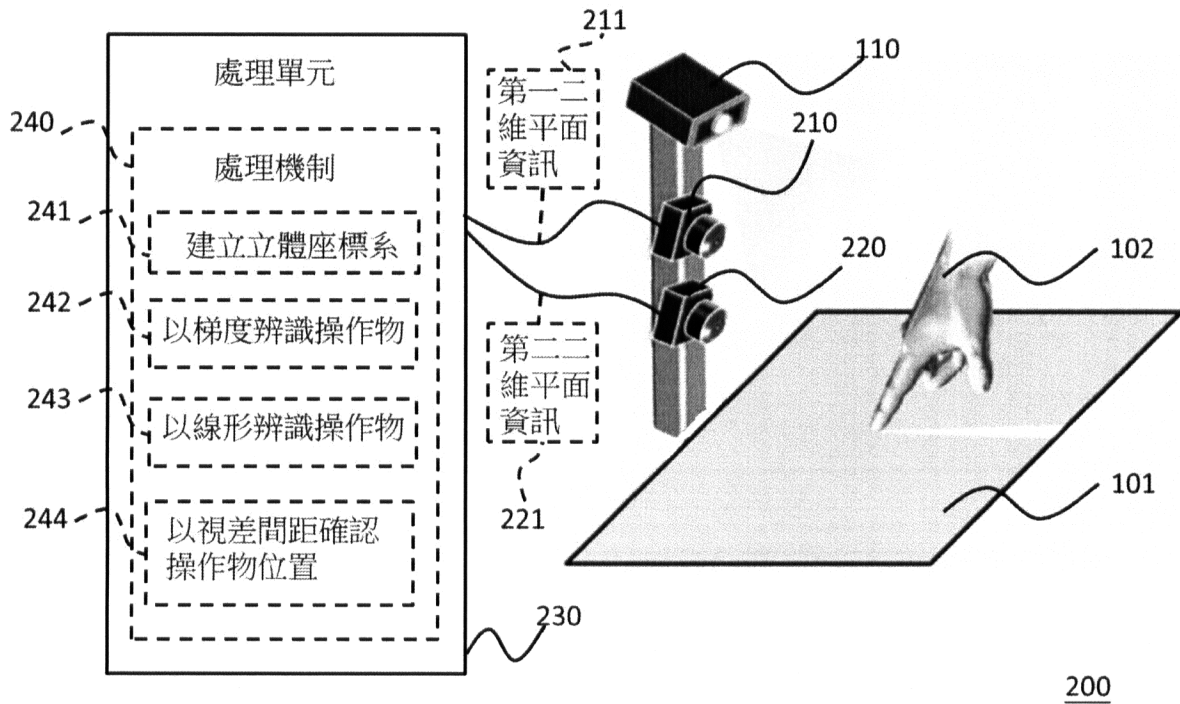
步驟四：以該操作物於該第一二維平面資訊與該第二二維平面資訊中的視差間距，定義該操作物的位置；

步驟五：當一物件介入該第一二維平面資訊與該第二二維平面資訊，且偵測該物件之視差間距介於一上限值與一下限值間，執行該步驟二與該步驟三，以判定該物件是否為該操作物。

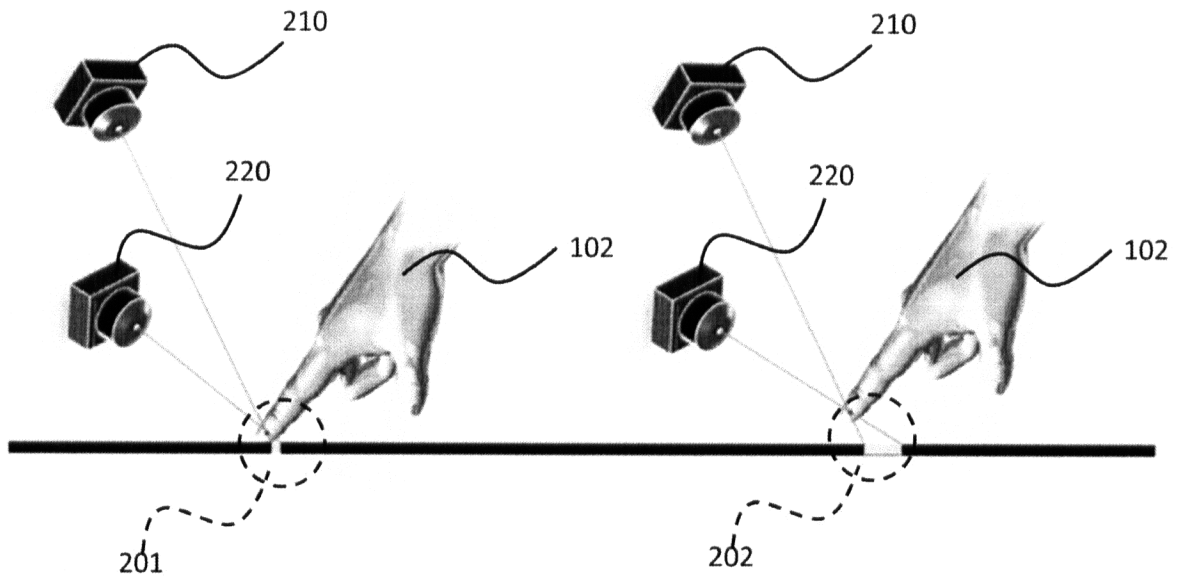
八、圖式：



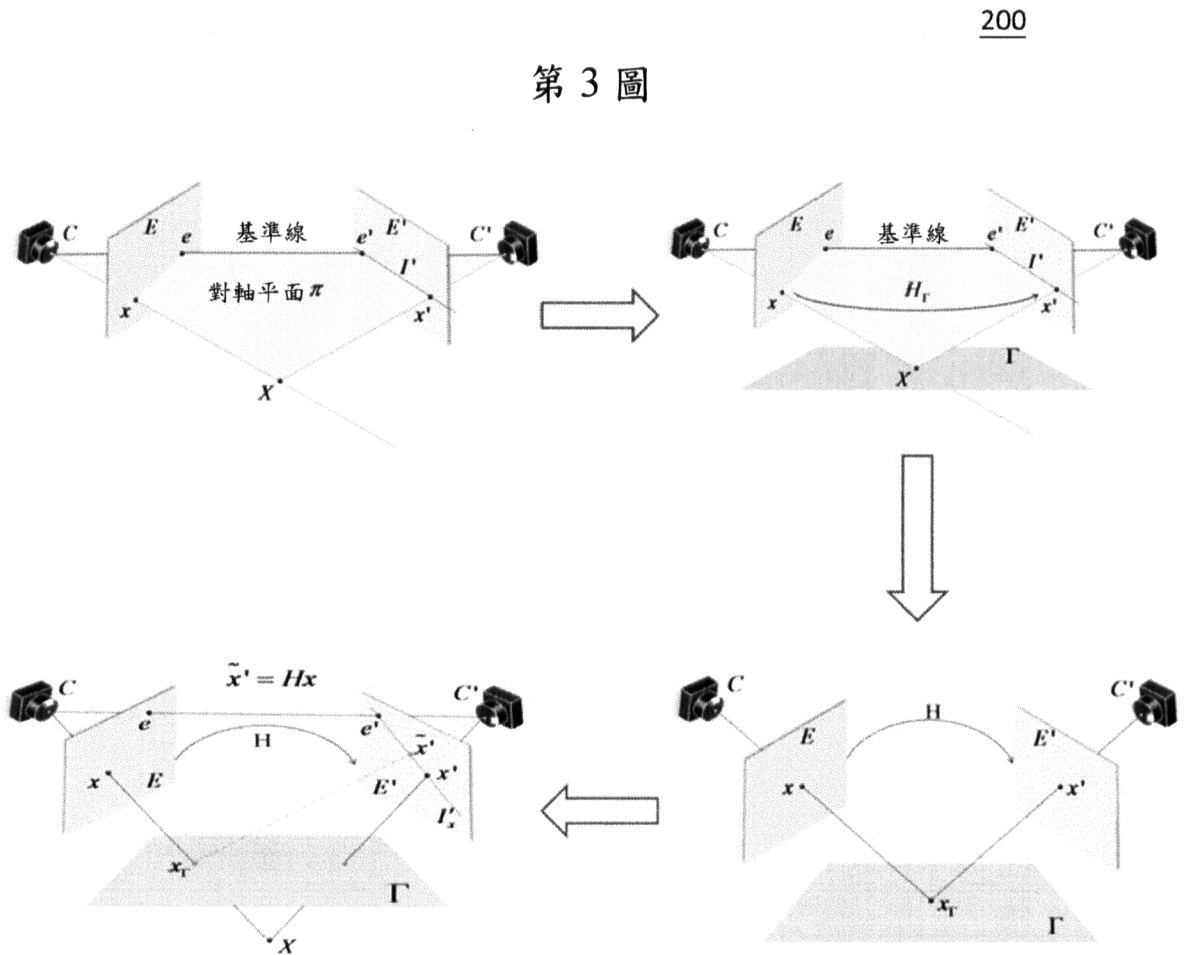
第 1 圖



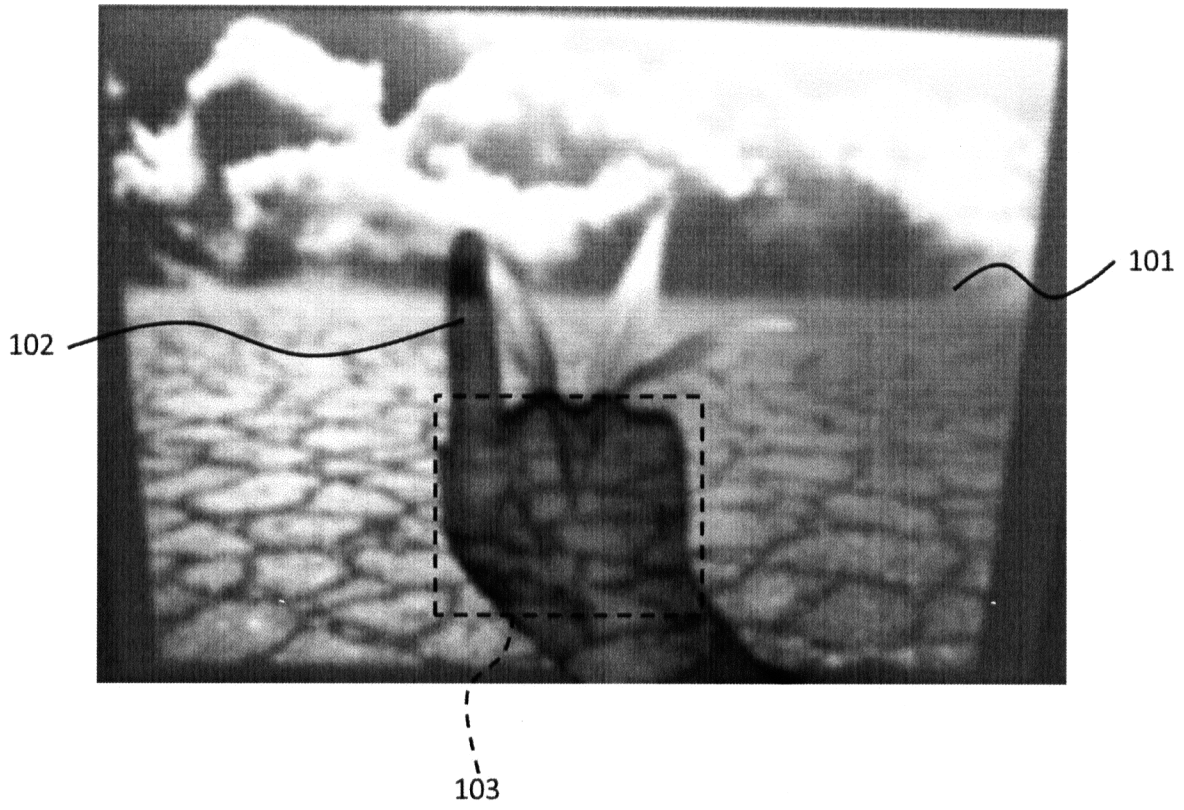
第 2 圖



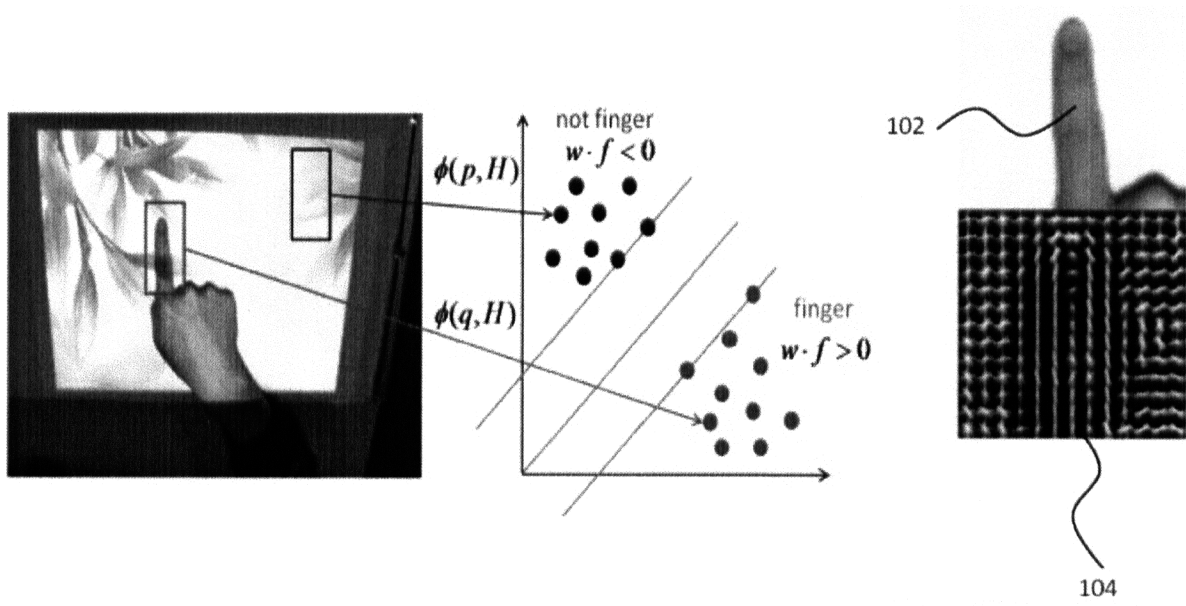
第 3 圖



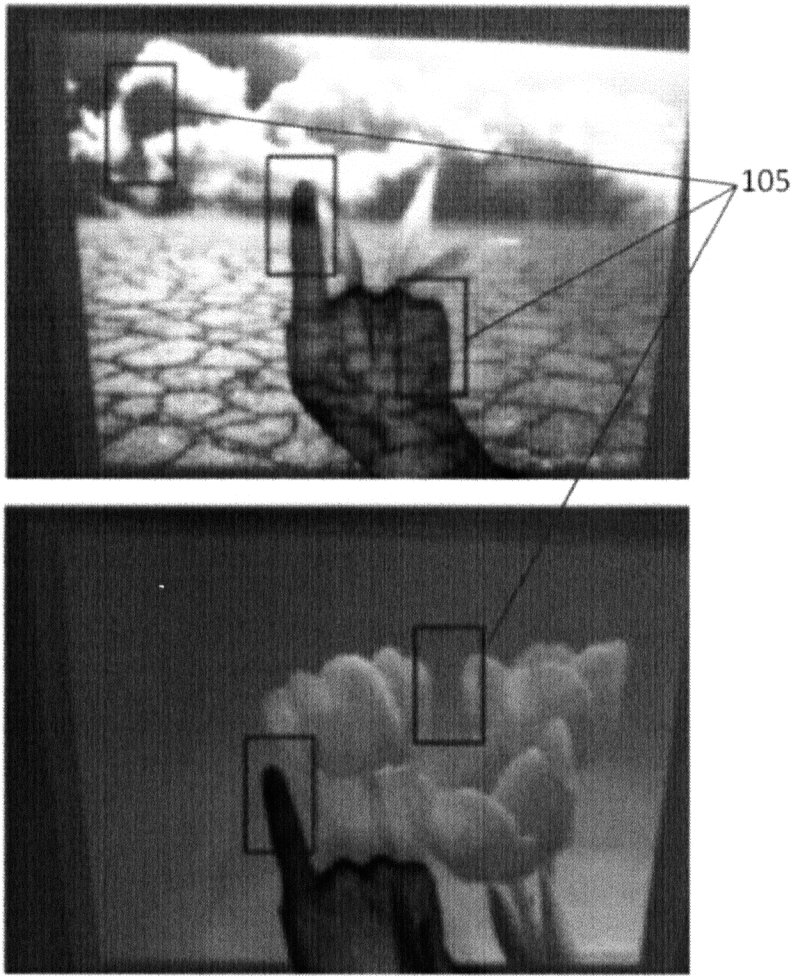
第 4 圖



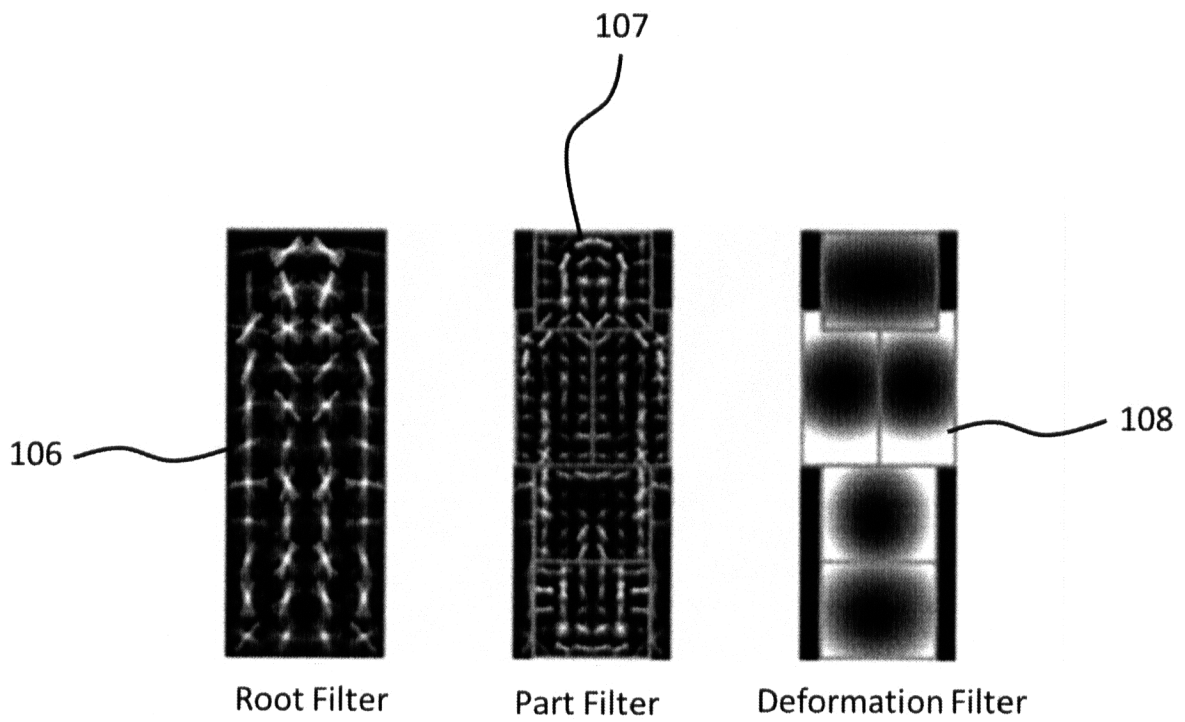
第 5 圖



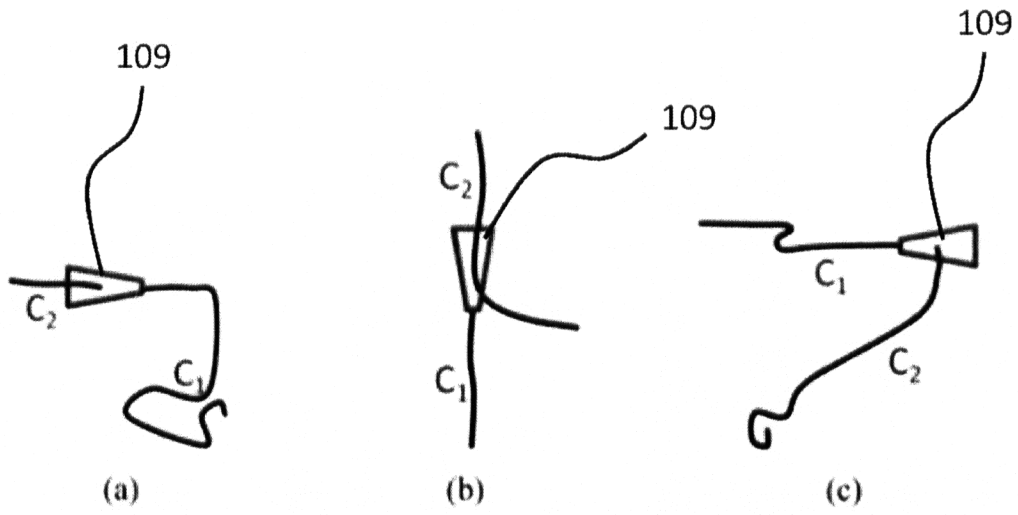
第 6 圖



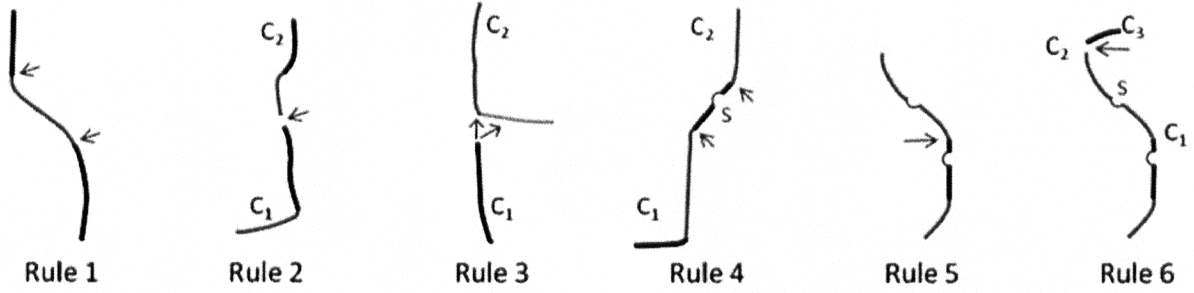
第 7 圖



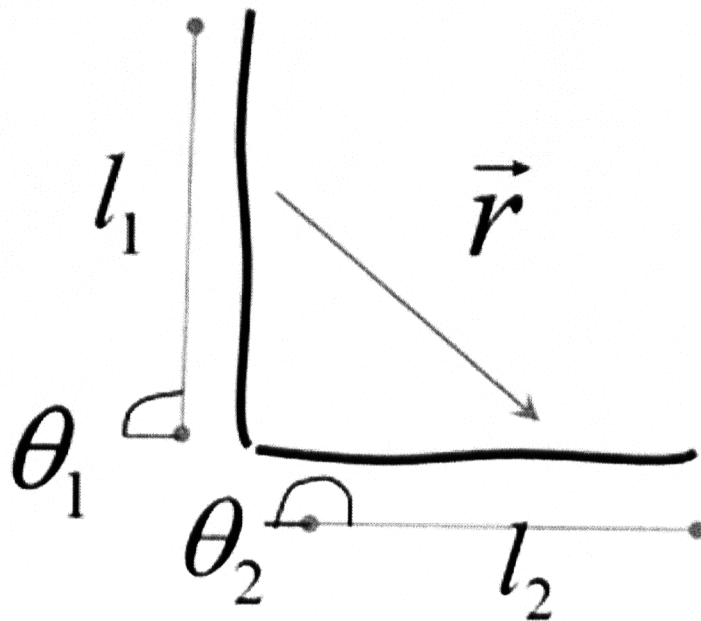
第 8 圖



第 9 圖



第 10 圖



第 11 圖