

網球視訊之場線偵測與重建

Court Line Detection and Reconstruction of Tennis Videos

姜耀川

義守大學資工系

m9503023@stmail.isu.edu.tw

謝朝和

銘傳大學資傳系

hsiehch@mcu.edu.tw

摘要

本文主要研究如何從網球視訊中偵測出場線並據以重建球場影像。網球場線一般都為白色，然而白色特徵會隨著不同場地以及天候變動而改變，我們提出適應性的強度門檻值，以增加白色場線偵測的可靠度。由於 Camera 拍攝時會有各種可能的變動，如平移(panning), 縮放(zooming)等，會使得部份場線消失；除外，比賽環境的外在因素，如紅土球場會因紅土覆蓋而影響場線偵測。為了解決這些問題，本文提出在 Radon transform 空間下之估測隱藏線的方法，以精確的重建球場影像。

關鍵詞：運動視訊分析，事件偵測，Radon 轉換，場線偵測

Abstract

This paper presents an efficient algorithm to detect court lines of tennis videos and reconstruct tennis court image accordingly. Court lines are usually white. However, the luminance feature of white lines may vary with different stadiums and weather change. In this paper, we propose an adaptive thresholding scheme to improve the robustness of white line detection. In a video sequence, some court lines may disappear due to the camera motion such as panning and zooming. In addition, the court lines are

often corrupted by noise such that they cannot be detected completely. To solve the so-called hidden lines, we present an effective Radon-space method to estimate the hidden lines and thus to recover the court images exactly.

Keywords: sport video analysis, event detection, Radon transform, court line detection

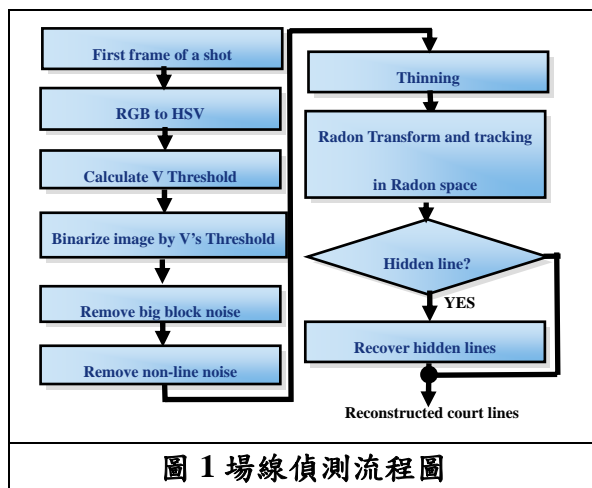
一、前言

在一個網球比賽場地中，場線是必然存在的資訊。場線偵測可以獲得許多有用資訊，如場線之直線方程式、場線之交點座標、場線 Map 等，其結果可做為後續高階處理的依據，例如球員偵測、追蹤以及高階事件識別(event recognition)等，因此場線偵測的研究是相當重要的。在現有網球視訊分析的文獻中，也都將場線偵測列為必需的前處理。

在早先的文獻中，場線偵測可以分成幾種不同方式，白色點辨識[1]、主色(dominant color)辨識[3]、梯度(gradient)辨識[6]等。其中主色辨識在其他運動視訊中也被廣範運用[5][7][8]。主色辨識與梯度辨識皆為運用色彩差異所進行的辨識方法。由於影像上並不僅僅只有場線資訊，因此梯度辨識結果將無法進一步精鍊

(refinement)。主色辨識可獲得主色的範圍與結構、色彩的屬性等多樣低階特徵，可使場線偵測更為精確。故本文將整合上述白色與主色兩種特徵，設計一個有效的場線偵測方法。

網球球場線一般都為白色，然而白色特徵會隨著不同場地、照明變化以及天候變動而改變，因此利用單一固定門檻(threshold)，將無法適應這些變化，而將場線完整切割出來。本文中我們提出適應性的強度門檻值(adaptive threshold)，以增加白色場線偵測的可靠度。由於 Camera 拍攝時會有各種可能的操作，如平移(panning)，縮放(zooming)等，這會使得部份場線在影像上消失；除外，比賽環境的外在因素，如紅土球場會因紅土覆蓋而影響場線偵測。這些都會造成所謂的隱藏線(hidden lines)，為了解決這些問題，本文也提出在 Radon transform 空間下之估測隱藏線的方法，以提高球場影像重建的精確度。



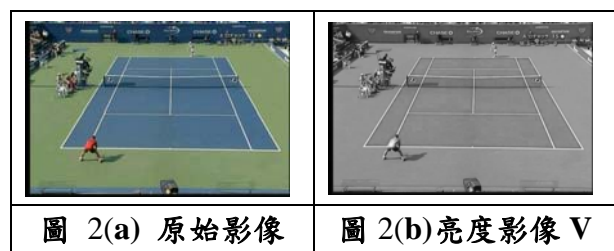
二、場線偵測與重建

本文提出的場線偵測的完整流程如圖 1 所示。首先，將輸入 video 切割成許多 shot，再由 shot 取出第一個 frame 做偵測。利用場線於 HSV 色彩空間下的特性產生一個適應性 V 值 Threshold，此 threshold 適用於該 shot 的所有 frames。使用此

Threshold 產生二值化影像，接著利用形態學運算，濾除一些非場線的雜訊。再透過細線化處理找出唯一的場線資訊並投影到 Radon space 下搜尋出出場線的亮點資訊。利用這些亮點資訊判斷是否有 hidden lines，若有則重新估算並重建隱藏線，以及求解得場線的直線方程式，最後重建球場影像。

2.1 適應性門檻機制

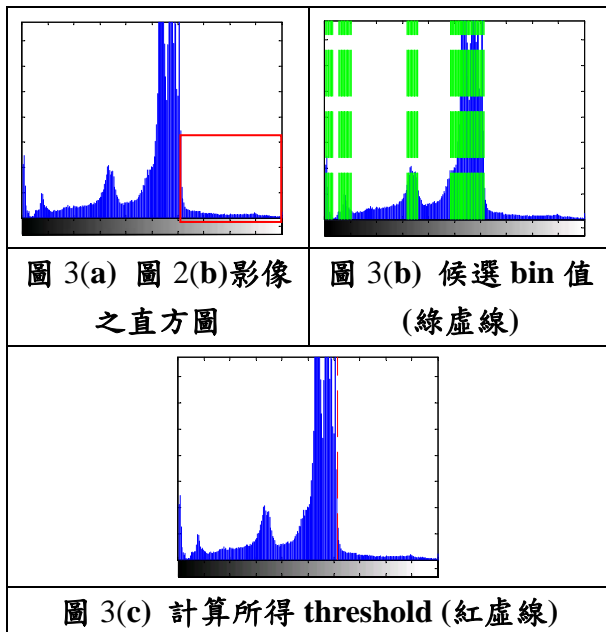
由於場線具有某些特性。在大部份球賽中所使用的場線均以白色為主而網球比賽亦是如此。針對白色於 HSV 色彩空間下有低飽和度 S (low saturation) 且高強度值 V (high value) 的特徵。然而不同比賽環境(如不同場地、同場地不同時間或不同照明等)，色彩特徵的表現是有所差異的。故如何定義一個能隨著比賽環境變化而自動調整的門檻值(adaptive threshold)，是場線偵測好壞的關鍵問題。根據我們的經驗，S 影像資訊可靠度較差，且僅由 V 即可得到不錯的場線偵測效果，故本文僅考慮 V 於場線的偵測。圖 2(a)及 2(b)為原始彩色影像與 V 影像。在此，我們提出了一個簡易的 V 值適應性門檻機制，它自動的產生一個適當的 V 門檻值以解決此問題。



當 camera 以全域角度(global angle)拍攝場地時，非場線像素點必然佔 80% 以上的比例，故場線(白色)所佔比例是很低的。換而言之，場線在 V 影像的 histogram 應為高 V 值但像素點較少(如圖 3(a)紅色框部份)。反之，場地像素應是像素點較多且 V 值不偏高。我們利用公式(1)給予各個 Bin 一個場地像素權重值(weight; W_n)， W_n 越大越有可能是場地像素。將 weight 由大

到小排序，然後 weight 值大的 Bin 優先挑選，直到累積的像素達總像素的 80% 為止，這些被挑到的 Bin 稱為候選 Bin(candidates)。進而從候選的 Bin 值中，挑選最大的 Bin 值做為強度的門檻值 (threshold)，如圖 3(c) 紅色虛線所示。

$$W_n = \frac{C_n}{(n+1)}, n=0,1,2,\dots,255 \quad (1)$$



2.2 場線的精煉

利用前述所產生的 V 值適應性門檻產生二值化影像，如圖 4(a) 所示，其中暗的 pixel 代表場景 pixel，剩餘亮的 pixel 則可能為場線 pixel。明顯可看出可能為場線 pixel 中含不少的雜訊。因此，我們利用以下三個處理步驟去消除這些雜訊進而獲得一個精確的場線。

步驟 1—大區塊的消除

球員身體、觀眾席等未被過濾之場地的像素通常呈現較大面積的區塊特性。在此利用形態學(morphology)的矩形結構元素(structure element)做為濾除的遮罩加以濾除，結果如圖 4(b)

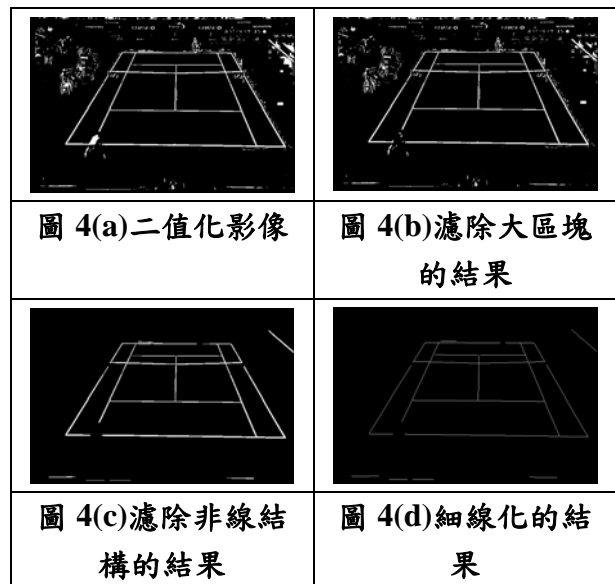
所示。

步驟 2—線結構的擷取

由於場線是一種線性結構，故不具線結構之像素點極可能為雜訊像素。利用形態學的線結構元素濾除非線結構之像素，結果如圖 4(c) 所示。此步驟可建立出場線 Map，可延伸應用在球員偵測時會是一個很有用的資訊。

步驟 3—細線化

由於場線的寬度在影像上的表現不可能為 1 個像素點，故利用細線化的運算取得一個精確且唯一場線資訊，結果如圖 4(d) 所示。

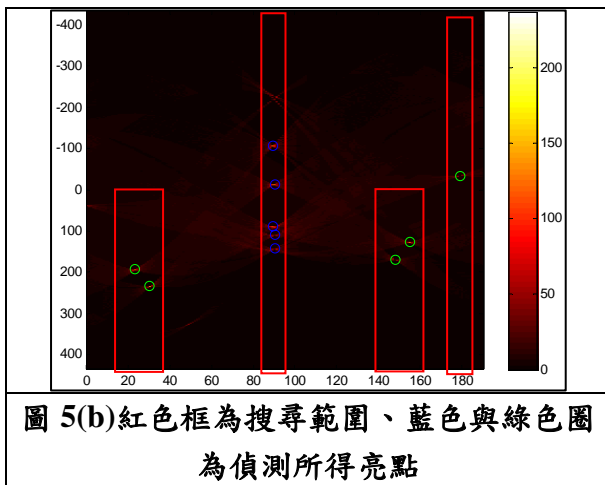
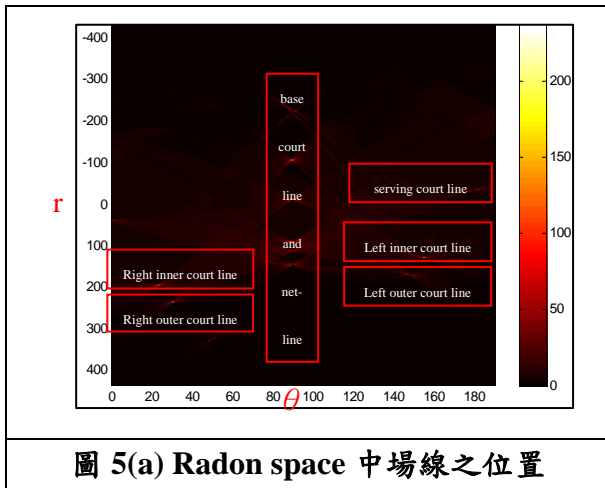


有了細線化影像後，由於為了獲得場線的精確位置，我們使用 Radon transform 將細線化影像轉到參數空間中(距離與角度)，取得亮點資訊，結果如圖 5(a) 所示，一個亮點即表示一條(場)線。獲得亮點資訊即可求解場線之直線方程式，如公式(2) 所示。進一步利用兩條方程式(水平與垂直各一條)計算出交點座標。

$$x \cos \theta + y \sin \theta = r \quad (2)$$

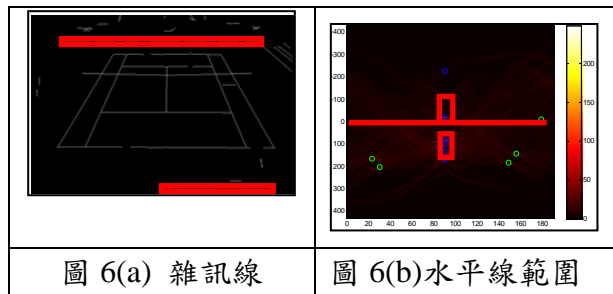
θ 表示角度; r 表示距離;

在 Radon space 下的亮點資訊透過限制取得較有可能的亮點資訊而非錯誤資訊(利用如較亮、可能位置、個數等)。一個全域角度拍攝時，場線於 Radon space 下的角度是在某個範圍內，就算是攝影機做運動時所造成的偏移角度也不可能太大(可能範圍 3~5 度且攝影機位置固定於中間)，故可以定出一個合理的搜尋亮點的範圍，如圖 5(b)所示。

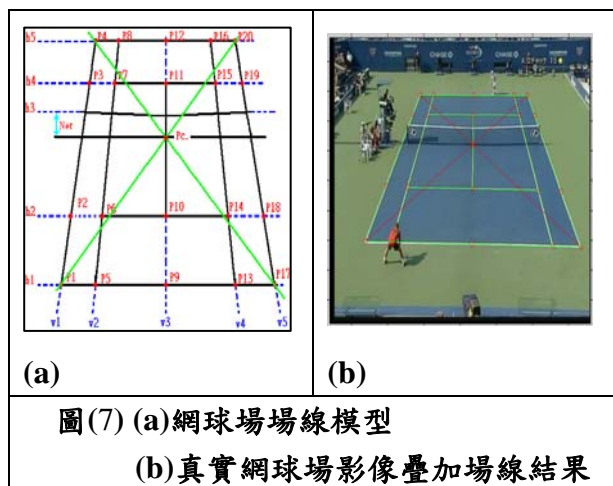


但偵測亮點時可能誤判非場線之亮點為場線之亮點，此稱為雜訊線，最常出現在水平線，因為其角度相似且長度較長，較容易造成誤判，如圖 6(a)所示，紅色框

部份為易與水平場線混淆之雜訊線。由於場地是一致色彩，故此問題常發生於場景中的牆邊、logo 等地方。經由我們實驗觀察發現大部份時間在 Radon space 下的水平(場)線(即 h1~5)，如圖 7(a)所示，會以 0(距離, r)為基準以 2:3 的方式呈現，如圖 6(b)所示。故正值會有兩個靠近 0 的亮點為場線資訊，反之，負值亦有三個亮點，利用此特性即可濾除雜訊線。



利用場線的位置與交點座標，如圖(7)所示，可延伸應用於定義一個球員初始偵測範圍。如此一來，球員不至於因搜尋視窗定義不良而導至失誤，並且此範圍有可以減少不必要的搜尋計算，如太偏僻的左、右、上、下端等影像邊緣。若考慮著影像的大小與場線四個角落(P1、P4、P17、P20)之間的關係則這個搜尋視窗是具有彈性的。



2.3 消失場線的重建

由於 Camera 拍攝時會有各種可能的

變動，如 panning, zooming 等，會使得部份場線消失；除外，比賽環境的外在因素，如紅土球場會因紅土覆蓋而影響場線偵測。為了讓後續場線追蹤時有一個完整的初始場線資訊。故消失場線的重建必然是初始的重要步驟之一。而消失線從產生的時間點來分會有二。一為初始偵測時場線已消失、另一為追蹤時場線消失。而追蹤時可用前幾張畫框的資訊即可估測得一個有效的場線位置，但初始偵測時無此類資訊可使用。因此，必須透過某些處理才可獲得估測的位置。

針對第一張畫框的場線偵測會遇到下列兩種不同問題：

1) 偵測失誤產生的隱藏線

在某些情況下，場線因外在的因素影響(如紅土球場，場線被紅土覆蓋)，使得白色資訊不顯著，以致於造成偵測時部份場線資訊遺失，使得投影到 Radon space 下的亮點消失或不顯著，而無法偵測出此線的亮點位置，如圖(8)所示，這種情形通常發生於因場線形狀失真造成資訊量較少的水平場線 h4 與 h5 上，參照圖(7)。

文獻[2]曾提到影像上的垂直場線至少存在兩條、水平場線四或五條，另外我們觀察得知下半場地的底線較不會扭曲失真，且存在於影像上的可能性很高。我們先取出水平場線 h1 上任意兩個場線交點。利用這影像交點距離與模型影像(圖 9(a)) 上相同兩點的交點距離做比較，取得一差異值(縮放差值)來校正模型與目前影像的場線大小差異。

假設影像上 v_1 與 v_3 存在，取得 $\overline{p_1 p_9}$ (參照圖(6))與模型上的 $\overline{p_1^* p_9^*}$ 之比例為 d 值，則線間距 $\overline{h_1 h_5}$ 會等於模型的 $\overline{h_1 h_5}$ 乘以 d ，如公式(3)。利用此關係可以重建遺失的場線 h4, h5，如圖 9。圖 9(a)為任意一張待偵測的影像，經過處理得到的場線

資訊。圖 9(b)及圖 9(d)為圖 9(c)與圖 9(e)的原圖之細線化二值影像。由細線化二值影像上可清楚發現紅色框所圍的正是偵測失誤造成的隱藏線位置。

$$d = \overline{p_1 p_9} / \overline{p_1^* p_9^*};$$

$$\overline{h_1 h_5} = \overline{h_1^* h_5^*} \times d;$$

‘*’表示模型上的資訊。

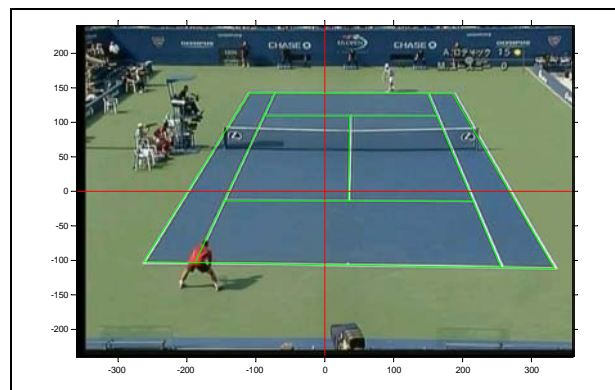
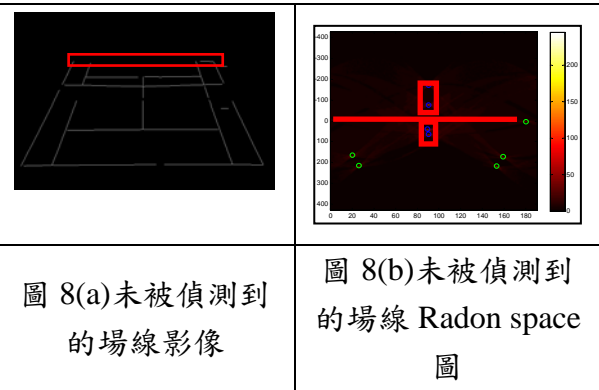


圖 9(a) 模型影像

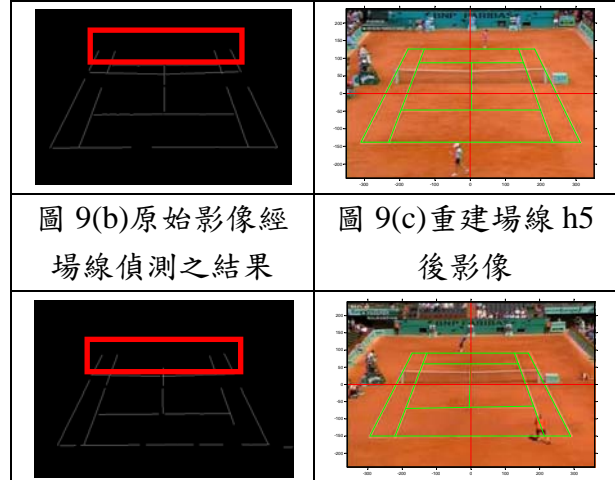


圖 9(d) 原始影像經
場線偵測之結果

圖 9(e)重建場線 h4
與 h5 後影像

2) 拍攝角度產生的隱藏線

此問題發生於垂直場線，如圖(10)所示，通常存在於左右側場線(如圖 7(a)的 $v1$ 、 $v2$ 、 $v4$ 、 $v5$)。而隱藏線的顯著特性有 1)外場垂直場線會比內場垂直場線先消失或一同消失，場地左右兩側場線同時消失可能性極少。2)發球中線極多數時間(包含發球時)會存在於影像上。故利用場地的對稱性與發球中線的資訊即可估測出影像上隱藏線所在位置，文獻[2]提出利用場線置中的影像 (center frame)做為模型影像來估測隱藏線，但此方法有兩個問題存在。一為場線置中的影像如何取得進而建立所謂的場線模型；另一問題是當攝影機 panning 時，此方法僅能獲得接近值。但場線對稱性的誤差值在越大的影像中相對越大，因此對稱性方式僅能做參考，而無法與場線完全匹配。在容許的誤差下如何找出影像上存在的剩餘資訊或是無剩餘資訊時直接給定一個估測位置變成此方法的延伸問題。

我們提出公式(4-1)與公式(4-2)將非置中場線影像轉換到置中場線，解決第一個問題。此轉換是依據三角形的定理與 Radon transform 投影的方式取得一個模擬對稱的置中場線資訊。其中由於影像的中間至發球中線大部份時間近似垂直。因此，公式上未知的 β 值趨近於中間發球線的 r 值。

有了模擬對稱的置中場線資訊($r1$)接著利用文獻[2]所提出的公式(公式 4-4)即可重建近似的隱藏線(如圖 10(d))所示，再進行校正的動作即能解決第二個問題，其解說如圖 11，而校正方法在下段中做說明。



圖 10(a)原始影像

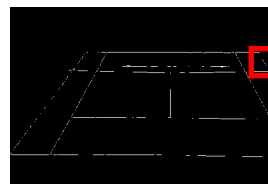


圖 10(b)細線化二值
影像

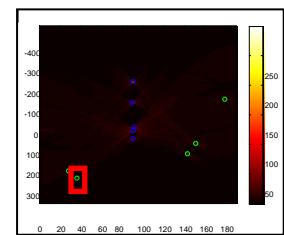


圖 10(c)估測隱藏線
在 Radon Space 的位
置

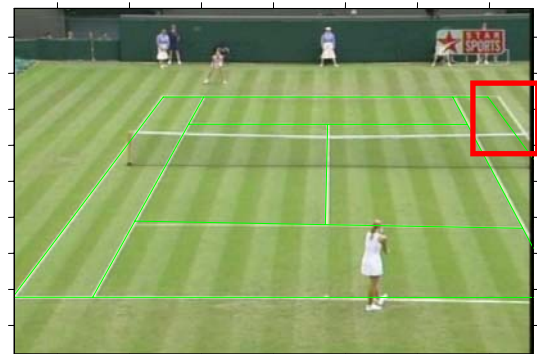


圖 10(d)利用對稱性的估測方式重建的隱
藏線

$$\begin{cases} \alpha = \frac{r}{\cos(\theta)}; & (1) \\ r_1 = (\alpha + \beta) \times \cos(\theta); r_1 \approx r_2; & (2) \\ \theta_3 = \theta_2 \cong \theta_1 = \theta & (3) \\ r_3 = r_2 + |r_1 - r|; & (4) \end{cases}$$

[r、θ]表示目前影像 Radon 空間的 r 值與 θ 值;
[r1、r2、θ1、θ2]表示模擬模型(置中場線)的 r 值與 θ 值;
[r3、θ3]表示求得的預估隱藏線位置的 r 值與 θ 值;
[β]表示影像中間點至發球中線的垂直距離;

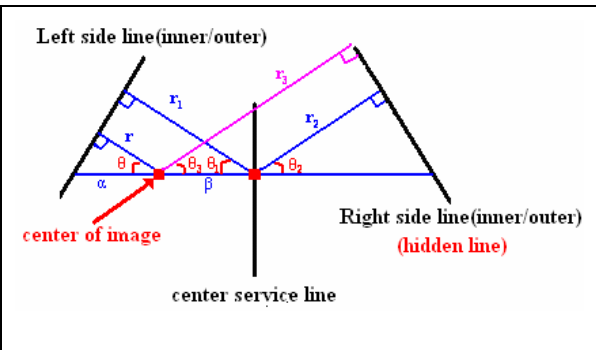


圖 11 利用對稱特性估測隱藏線解說圖

由於水平場線與垂直場線有交點存在，故若與隱藏垂直線產生的交點存在於影像內必有部份隱藏線的場線資訊顯示於影像中。反之，則完全隱藏。因此，若希望在視覺上避免重建的失誤則必須進一步做校正處理。利用下面表 1 的演算法即可達到一個良好的校正效果。圖 12 即為校正前後之隱藏線重建結果，很明顯地，此校正方法確實改善了隱藏線的重建精確性。

表 1 隱藏線校正演算法
初始化: 輸入影像為細線化的二值影像; w : 影像寬度;

假設 v5 為隱藏線;v 為估測隱藏線;
v(1,1)為估測隱藏線 r 值;v(2,1)為估測隱藏線 θ 值;已知水平場線 h5 的直線方程式。

- Step1:
計算 v 與 h5 的交點座標 p(x,y);
若(p 點落在影像內)，則執行 Step 2;
否則執行 Step 6;
- Step 2 :
若(p 點位於影像邊緣)，則執行 Step 6;
否則執行 Step 3;
- Step 3 :
若(p 點的右(左)方為 1 且 p 點的右(左)下方為 0) 則
p 點位移至右(左)方;
執行 Step 2;
否則執行 Step 4;
- Step 4 :
若(p 點的右(左)下方為 1) 則
p 點位移至右(左)下方;
執行 Step 5;
若(p 點的右(左)上方為 1) 則
p 點位移至上方;
執行 Step 2;
否則執行 Step 6;
- Step 5 :
若(隱藏線為外線) 則
若(p 點的右(左)方為 0)且(p 點的右(左)下方為 0)且(p 點下方開一個 3*4 區塊，區塊內為 1 的像素點要多於 3) 則
p 點位移回至左(右)上方;
執行 Step 6;
否則執行 Step 2;

否則
 若(p點的右(左)方為0)且(p點下方開一個3*4區塊, 區塊內為1的像素點要多於3)則
 p點位移回至左(右)上方;
 執行 Step 6;
 否則執行 Step 2;
 Step 6:
 若 (p點座標更新)則 利用 p點座標校正 $v(1, 1)$ 值;

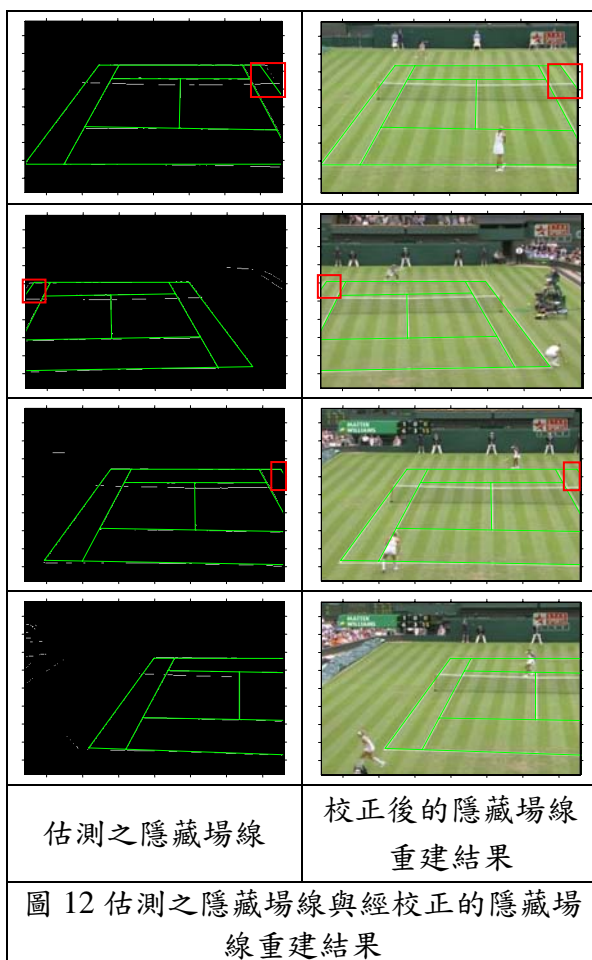
的直方圖, 我們設計了一個適應性門檻值決定法則, 以偵測白色場線, 它對於不同場地及天候變化, 可以自動調整門檻值。偵測結果的場線經由 Radon Transform, 轉到參數空間, 在此空間下計算場線的方程式。除外, 我們也運用了場線模型的概念解決了消失線的問題, 使得即使某些場線不存在於影像中, 仍然可以重建完整的球場影像。本文的場線偵測結果可產生有用資訊包括: 場線之直線方程式、場線之交點座標、球員初始偵測範圍、場線 Map 等, 這些資訊可應用到球員偵測與追蹤, 使得偵測更有效率。

四、誌謝

本研究由國科會補助經費, 計畫編號 NSC 96-2221-E-130-021, 在此致謝。

五、參考文獻

- [1] D. Zhong, S.F. Chang, "Real-time view recognition and event detection for sports video," Journal of Visual Communication and Image Representation, vol.15., pp. 330-347, 3, Sept., 2004
- [2] H. Miyamori and S.I. Iisaku "Video Annotation for Content-based Retrieval using Human Behavior Analysis and Domain Knowledge," Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp.320-325, 2000.
- [3] C.-L. Huang, H.-C. Shih and C.-Y. Chao, "Semantic Analysis of Soccer Video using Dynamic Bayesian Network," IEEE Trans. on Multimedia, Vol. 8, No. 4, pp. 749-760, Aug., 2006.
- [4] J. Han and D. Farin and P. H. N. de With, "Multi-level analysis of Sports Video Sequences," in SPIE Conference on Multimedia Content Analysis, Management, and Retrieval, vol. 1, Jan. 2006, San Jose, USA.
- [5] J. Han, P. H. N. de With "A unified and



三、結論

本文提出了一個網球場線偵測並據以重建球場影像的有效方法, 利用亮度資訊

Efficient framework for Court-Net Sports Videos Analysis Using 3-D Camera Modeling,” SPIE Electronic Imaging, vol. 1, p. 6506-15, Jan. 2007, San Jose, USA.

- [6] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati, "Semantic Adaptation of Sports Video with User-centred Performance Analysis," IEEE Transactions on Multimedia, vol. 8, n. 3, pp. 433-443, June 2006.
- [7] N. Rea, R. Dahyot and A. Kokaram, "Classification and Representation of Semantic Content in Broadcast Tennis Videos," IEEE International Conference on Image Processing, vo3,pp. III-1204-7, Sept., 2005
- [8] Z. Zivkovic, et al, "Two Video Analysis Applications Using Foreground Background Segmentation," Proceedings of the VIE-2003 Conference on Visual Information Engineering, pp. 310-313, Jul. 2003, Surrey, Guildford.