

การจำแนกท่าตีในวิถีทัศนกีฬาเทนนิสแบบอัตโนมัติ

โกยิต เคชะปริกิติ¹, ธนารัตน์ ชลิดาพงศ์¹ และ อาริต ธรรมโน¹

¹คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

²ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ

Emails: kosit35185@hotmail.com, thanarat@cp.eng.chula.ac.th, arit@it.kmitl.ac.th

บทคัดย่อ

บทความนี้กล่าวถึงขั้นตอนวิธี ในการจำแนกท่าตีในวิถีทัศนกีฬาเทนนิส ที่ได้จากวิถีทัศนที่เผยแพร่บนอินเทอร์เน็ต ซึ่งมักเป็นวิถีทัศนที่มีความละเอียดต่ำ มีสัญญาณรบกวนสูง และมีมุมมองจากกล้องที่ตั้งในระยะไกลเพียงมุมเดียว เริ่มจากการติดตามตำแหน่งผู้เล่นทั้งสอง ด้วยขั้นตอนวิธีย้ายค่าเฉลี่ย ร่วมกับการหาความแตกต่างระหว่างเฟรม กระทำกับฮิสโทแกรมสี จากนั้นทำการตรวจจับเส้นของสนามเทนนิสโดยวิเคราะห์จากค่าความเข้มของจุดภาพ และทำการตรวจจับตำแหน่งของลูกบอล โดยการวิเคราะห์หาบริเวณจุดภาพที่มีการเคลื่อนไหว สุดท้ายนำตำแหน่งผู้เล่นทั้งสอง และ ตำแหน่งลูกบอล เทียบกับตำแหน่งเส้นสนามเทนนิส มาใช้ในการจำแนกท่าตีเทนนิสออกเป็น โฟร์แฮนด์ แบคแฮนด์ และ เซิร์ฟ

คำสำคัญ— การวิเคราะห์วิถีทัศนกีฬา; การติดตามวัตถุ; การจำแนกกิจกรรม

1. บทนำ

ในปัจจุบัน งานวิจัยทางด้านเทคโนโลยีคอมพิวเตอร์วิชันและการประมวลผลภาพมีความก้าวหน้าเป็นอย่างมาก และได้ถูกนำมาประยุกต์ใช้ในหลายด้าน เช่น ระบบตรวจการฉ้อโกงจากวิถีทัศน ระบบวิเคราะห์สภาพจราจร ระบบวิเคราะห์วิถีทัศนแบบอัตโนมัติ ระบบค้นหาวิถีทัศนแบบอัตโนมัติ รวมทั้งการวิเคราะห์วิถีทัศนกีฬา เพื่อบรรยายเหตุการณ์ที่น่าสนใจหรือเหตุการณ์สำคัญที่เกิดขึ้นในวิถีทัศนนั้นๆ บทความนี้จะกล่าวถึงการวิเคราะห์วิถีทัศนกีฬาเทนนิสแบบอัตโนมัติ ซึ่งถึงแม้มีหลายคณะได้ทำงานวิจัยในเรื่องนี้ ดังปรากฏใน [1] [2] [3] แต่หลายระบบยังพบปัญหาในการทำงาน เช่น ขาดความแม่นยำในการติดตามตำแหน่งผู้เล่นเทนนิส เนื่องจากจากกล้องมีความซับซ้อน หรือ ปัญหาในการจำแนกประเภทของเหตุการณ์สำคัญในการเล่นเทนนิส เช่น การเซิร์ฟ การตีลูก โฟร์แฮนด์ แบคแฮนด์ วอลเลย์ เป็นต้น งานวิจัย [4] ได้พัฒนาขั้นตอนวิธีในการบรรยายการแข่งขันเทนนิส จากวิถีทัศนแบบอัตโนมัติ โดยทำการตัดแยกผู้เล่นออกจากฉากหลัง โดยใช้คุณสมบัติสีและการเทียบแบบสามมิติ จากนั้นเรียนรู้และจำแนกเหตุการณ์สำคัญในการเล่นด้วยแบบจำลองฮิดเดนมาร์คอฟ (HMMs) โดยดึงลักษณะเด่นจากโครงร่างของผู้เล่นที่ตัดแยกได้ ผลลัพธ์สุดท้ายคือการจำแนกเหตุการณ์ โดยแบ่งออกเป็น โฟร์แฮนด์ แบค

แฮนด์ เซิร์ฟ วอลเลย์ โฟร์แฮนด์ วอลเลย์แบคแฮนด์ งานวิจัย [5] ได้นำเสนอวิธีการติดตามผู้เล่นเทนนิสด้วยวิธีการกรองแบบ SVR particle filter จากนั้นแทนการเคลื่อนที่ของผู้เล่นด้วยฮิสโทแกรมฮิสโทแกรมแบ่งส่วน (Slice-based optical flow histogram) ทำการรู้จำโดยใช้ Support Vector Machine (SVM) ผลลัพธ์ที่ได้คือการจำแนก การตีเทนนิสแบบโฟร์แฮนด์ และ แบคแฮนด์ ส่วนงานวิจัย [6] ใช้วิธีสร้างแบบจำลองและตัดแยกส่วนภาพของผู้เล่นเทนนิสโดยใช้ฮิสโทแกรมสี จากนั้นใช้รูปทรงของผู้เล่นที่ตัดแยกได้มาใช้ในการรู้จำด้วย SVM ผลลัพธ์คือการจำแนกการตีแบบ โฟร์แฮนด์ แบคแฮนด์ หรือ ไม่ใช่ทั้งสองประเภท งานวิจัยที่ได้ยกตัวอย่างข้างต้นล้วนแต่เป็นการใช้ข้อมูลเชิงพื้นที่หรือตำแหน่งของผู้เล่นในการวิเคราะห์ งานวิจัย [7] ได้นำเสนอวิธีการตรวจหาตำแหน่งลูกบอลในภาพ โดยการสร้างกราฟแนววิถีการเคลื่อนที่ของลูกบอลที่มีลักษณะวิถีโค้งพร้อมด้วยการใช้ข้อมูลเชิงขณะที่ลูกบอลกระทบไม้ ส่วนงานวิจัย [8] พัฒนาระบบที่มีชื่อว่า TennisSense โดยใช้วิถีทัศนจากกล้องที่ติดตั้งอยู่บนเสื้อสนามเทนนิส ในการติดตามตำแหน่งของลูกเทนนิส โดยใช้วิธีการตรวจจับบริเวณภาพที่มีการเคลื่อนที่ จากนั้นทำการรู้จำประเภทการตีโดยพิจารณาจากการเคลื่อนที่ของตำแหน่งลูกบอล งานวิจัย [9] นำเสนอวิธีการดึงคุณลักษณะเด่นทั้งภาพและเสียงที่ได้จากวิถีทัศน และรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟ ระบบนี้ได้จำแนกเหตุการณ์ออกเป็น 4 ประเภทคือ การเซิร์ฟพลาต การตีลูกกลับ รีเพลย์ และหยุด งานวิจัย [10] เสนอวิธีการตรวจจับตำแหน่งของลูกบอลด้วยวิธีการจำแนกส่วนภาพโดยใช้หลายคุณสมบัตินี้ โดยดูจากความเข้มของจุดภาพ จากนั้นใช้ตัวกรอง Particle Filter เพื่อทำการติดตามตำแหน่งลูกบอล

งานวิจัยที่กล่าวมาข้างต้นนั้น ส่วนมากใช้วิถีทัศนนำเข้าที่มีคุณภาพสูง บางงานใช้มุมมองที่มีการวางตำแหน่งมาอย่างเฉพาะ บางงานใช้มุมมองหลายมุมมองประกอบกันในการวิเคราะห์ งานวิจัยที่นำเสนอในบทความนี้ เราพัฒนาขั้นตอนวิธีในการจำแนกท่าตีในวิถีทัศนการแข่งขันกีฬาเทนนิสประเภทเดี่ยว ที่เผยแพร่บนอินเทอร์เน็ต ซึ่งมักเป็นวิถีทัศนที่มีความละเอียดต่ำ มีสัญญาณรบกวนสูง และไม่สามารถกำหนดมุมมองที่เฉพาะเจาะจงได้ ดังนั้น ในงานวิจัยนี้ เราจะใช้มุมมองที่พบส่วนใหญ่ คือมุมมองจากกล้องระยะไกลเพียงอย่างเดียว

2. ขั้นตอนวิธี

ในส่วนนี้จะกล่าวถึงขั้นตอนวิธีที่ในการจำแนกเหตุการณ์สำคัญในวิดีโอที่บันทึกภาพเทนิส โดยเริ่มจากการติดตามตำแหน่งผู้เล่นทั้งสอง ด้วยขั้นตอนวิธีย้ายค่าเฉลี่ย (Mean Shift) ร่วมกับการหาความแตกต่างระหว่างเฟรม (Frame Differencing) กับฮิสโทแกรมสี RGB ขนาด 8x8x8 หลังจากนั้น ทำการตรวจจับหาเส้นของสนามเทนนิสโดยวิเคราะห์จากค่าความเข้มของจุดภาพ และทำการตรวจจับตำแหน่งของลูกบอล โดยการวิเคราะห์หาบริเวณจุดภาพที่มีการเคลื่อนไหว สุดท้ายนำตำแหน่งผู้เล่นทั้งสอง และตำแหน่งลูกบอล เทียบกับตำแหน่งเส้นสนามเทนนิส มาใช้ในการจำแนกท่าตีเทนนิสออกเป็น โฟร์แฮนด์ หรือ แบคแฮนด์

2.1. การติดตามตำแหน่งผู้เล่น

ในการติดตามผู้เล่นในวิดิทัศน์การแข่งขันเทนนิสจากมุมกล้องระยะไกล ซึ่งจะเห็นผู้เล่น ทั้งผู้เล่นที่อยู่ฝั่งใกล้กล้อง (near player) และผู้เล่นที่อยู่ฝั่งไกลกล้อง (far player) เราใช้ขั้นตอนวิธีย้ายค่าเฉลี่ย (Mean Shift) [11] [12] ซึ่งเป็นขั้นตอนวิธีที่มีประสิทธิภาพในการติดตามวัตถุ โดยการประมาณความหนาแน่นการกระจายแบบนอนพารามตริก (non parametric density estimation) ซึ่งมีสมมติฐานว่าวัตถุที่จะติดตามนั้น จะต้องมีความแตกต่างเพียงพอที่สามารถแยกออกจากพื้นหลังได้ การใช้วิธีย้ายค่าเฉลี่ยแบบมาตรฐาน เพื่อที่จะใช้ติดตามผู้เล่นเทนนิสในวิดิทัศน์การแข่งขันกีฬาเทนนิสนั้น พบว่าปัญหาคือฉากหลังของผู้เล่นโดยเฉพาะผู้เล่นฝั่งไกล มีความซับซ้อน เนื่องจากปรากฏภาพของ กรรมการกำกับเส้น และ ผู้ชม ที่มีการเคลื่อนไหวตลอดเวลา จึงเป็นการยากที่จะติดตามตำแหน่งผู้เล่นได้อย่างแม่นยำ เราจึงแก้ปัญหานี้ด้วยวิธีการใช้ขั้นตอนวิธีติดตามวัตถุด้วยวิธีย้ายค่าเฉลี่ยรวมกับการหาความแตกต่างระหว่างเฟรม (Frame Differencing) โดยใช้ฮิสโทแกรมบริเวณสี ในบริเวณส่วนภาพที่ทำการติดตาม

หลักการของวิธีย้ายค่าเฉลี่ย คือการกำหนดบริเวณส่วนภาพเริ่มต้นที่จะทำการติดตาม จากนั้นทำการหาบริเวณที่มีความหนาแน่นของการกระจายสีที่เหมือนกับต้นแบบมากที่สุดในแต่ละรอบที่กำหนด ในขอบเขตแบนด์วิดท์ ซึ่งในแต่ละรอบถ้ามีการย้ายจุดศูนย์กลาง แสดงว่าค่าแห่งจุดศูนย์กลางที่ย้ายนั้นมีความหนาแน่นของบริเวณสูงกว่าจุดศูนย์กลางเดิม แต่ถ้าไม่มีการย้ายจุดศูนย์กลางก็ให้ถือว่าจุดนั้นเป็นจุดที่มีความหนาแน่นของบริเวณใกล้เคียงสูงสุด

ปัญหาที่เกิดจากสภาพแวดล้อมในวิทัศน์ เช่น ผู้เล่นฝั่งไกลซึ่งเป็นไปได้ที่จะสวมเสื้อสีเขียวกับการกระโดด หรือผู้ชมที่เคลื่อนไหวไปมา สามารถทำการแก้ปัญหาโดยใช้การหาความแตกต่างระหว่างเฟรม ผลลัพธ์ที่ได้คือบริเวณส่วนภาพที่มีการเคลื่อนไหวในเฟรมนั้นๆ จากนั้นทำการประมวลส่วนภาพด้วยตัวกระทำมอร์โฟโลยี (Morphology Operation)

เพื่อเป็นการจัดสัญญาณรบกวนในภาพ จากนั้นทำขั้นตอนวิธีย้ายค่าเฉลี่ย โดยใช้อัลกอริทึมการปริภูมิสี RGB แบบ 8x8x8 บิต ของส่วนภาพดังกล่าว

เราสามารถแปลงค่าจุดภาพในปริภูมิสี RGB ให้อยู่ในรูปฮิสโทแกรม
ปริภูมิ RGB แบบ 8x8x8 บิต ด้วยวิธีดังนี้

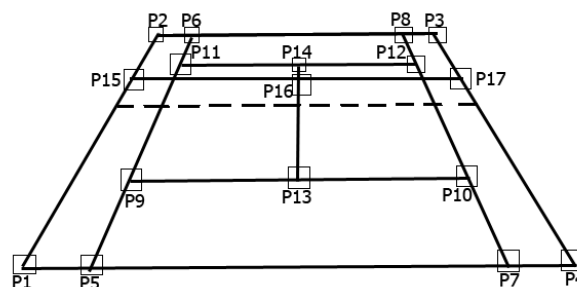
$$S = \text{floor}[R/32] + 8 * \text{floor}[G/32] + 64 * \text{floor}[B/32] \quad (1)$$

โดยที่ R,G,B คือค่าความเข้มจกภาพในปริภูมิสีแดง เขียว น้ำเงิน ตามลำดับ

2.2. การหาเส้นสนาม

ในการหาเส้นสนาม เราใช้หลักการจากข้อสังเกตที่ว่าความเข้มข้นของ
คุณภาพบริเวณเส้นสนามจะสูงกว่าบริเวณอื่น และเส้นสนามเป็นเส้นตรง
ดังนั้น การหาส่วนที่เป็นเส้นสนามเทนนิส เราเริ่มจากการกำหนดจุดใน
บริเวณขอบของสนามเทนนิสทั้ง 17 จุด จากนั้นเชื่อมเส้นตรงระหว่างจุดดัง
แสดงในรูปที่ 1 จากนั้นทำการหาค่าความเข้มข้นของคุณภาพในบริเวณที่
กำหนด เมื่อได้ตำแหน่งที่มีความเข้มข้นสูงสุดของในแต่ละเส้น นำพิกัดที่มีค่า
ความสว่างสูงสุดลบกับเส้นที่ได้จากการสร้างในตอนแรก จากนั้นหาฐาน
นิยมและหาพิกัดที่อยู่ในช่วงของฐานนิยม นำพิกัดที่อยู่ในช่วงของฐานนิยม
ทำการถดถอยเชิงเส้น (Linear Regression)

โดยมีขั้นตอนดังนี้



รูปที่ 1. จุดที่ถักกำหนด 17 จุดในสนามเทนนิส

1. ที่ $t = 1$ แปลงภาพสี่เป็นภาพเทา กำหนด 17 จุด ตามรูปที่ 1
2. สร้างเส้นตรงจากจุด ในขั้นตอนที่ 1 ได้แก่ เส้นที่ 1 P1-P2 ;
เส้นที่ 2 P2-P3 ; เส้นที่ 3 P3-P4 ; เส้นที่ 4 P4-P1 ; เส้นที่ 5 P5-P6 ; เส้นที่ 6
P7-P8 ; เส้นที่ 7 P9-P10 ; เส้นที่ 8 P11-P12 ; เส้นที่ 9 P13-P14 ; เส้นที่ 10
P15-P16 ; เส้นที่ 11 P16-P17 ; นำเส้นทั้ง 11 เส้นแบ่งกลุ่มเป็น 2 กลุ่มคือ
กลุ่มที่อยู่บนแนวนอนและกลุ่มที่อยู่บนแนวตั้ง

$LineX_i$ = เส้นที่2, เส้นที่4, เส้นที่7, เส้นที่8, เส้นที่10, เส้นที่11

$LineY = \text{เส้นที่}1, \text{เส้นที่}3, \text{เส้นที่}5, \text{เส้นที่}6, \text{เส้นที่}9$

$$L_i = [LineX_i \quad LineY_i]$$

$LineX_t$ คือ เส้นในแนวนอนทั้งหมดในเฟรมที่ t , $LineY_t$ คือ เส้นในแนวตั้งทั้งหมดในเฟรมที่ t , L คือเส้นทั้งหมดของในเฟรมที่ t

3. $t = 2$

4. ขยายขนาดเส้นใน $LineX_{i-1}$ และ $LineY_{i-1}$ โดยขนาดที่ขยายได้มาจาก ความหนาของเส้นสนามเทนนิส ในภาพจากขั้นตอนที่ 1 เส้น $LineX_{i-1}$ จะขยายแกนในพิกัด y ดังรูปที่ 2a ส่วนเส้นที่อยู่ใน $LineY_{i-1}$ จะขยายแกนในพิกัด x ดังรูปที่ 2b

$$ExPandLine_i = \left[(LineX_{i-1} \pm y) (LineY_{i-1} \pm x) \right]$$

$ExPandLine_i$ คือ เส้นทั้งหมดที่ได้ถูกขยายความหนา

5. หาพิกัดของความเข้มในจุดภาพสูงสุด ของ $ExPandLine_i$ โดยที่ถ้าเป็นเส้นที่ขยายในพิกัด y จะพิจารณาความเข้มของจุดภาพในแนวตั้งในทางกลับกันถ้าเป็นเส้นที่ถูกขยายในพิกัด x จะพิจารณาความเข้มของจุดภาพในแนวนอน

$$MaxLine_i(x, y) = MaxIntensity(ExPandLine_i)$$

$MaxLine_i(x, y)$ คือ พิกัด (x, y) ของตำแหน่งที่มีค่าความเข้มของจุดภาพสูงสุดของในแต่ละเส้น $ExPandLine_i$ ของภาพในเวลา t

6. คืนค่าพิกัดของเส้น L_{i-1}

$$Line_i(x, y) = returncodinate(L_{i-1})$$

$Line_i(x, y)$ คือ พิกัดจากเส้น L_{i-1} ในตำแหน่ง (x, y) ของภาพในเวลา $t-1$

7. หาผลต่างของพิกัดจากขั้นตอนที่ 5 และ 6

$$(Line_i(x, y) - MaxLine_i(x, y)) = Pix_i(x, y)$$

$Pix_i(x, y)$ คือผลต่างของพิกัด (x, y) ในตำแหน่งที่ i

8. นำพิกัด x และ y ของทุกตำแหน่ง i ใน Pix มารวมกันได้

$$D_i = Pix_i(x + y)$$

9. หาฐานนิยมของ $Mode(D_i) = m$

10. หาค่า D_i ที่อยู่ระหว่าง $m \pm th$ เพื่อทำการเก็บพิกัด โดยที่ th คือค่าขีดแบ่ง

$$Y_i = \begin{cases} \text{keep Codinate } MaxLine_i(x, y) & \text{if } m - th \leq D_i \leq m + th \\ \text{eliminate} & \text{Otherwise} \end{cases}$$

Y_i คือพิกัดจาก $MaxLine_i$ ที่อยู่ระหว่าง $m \pm th$ โดยมีพิกัด x และ y ที่ตรงกับตำแหน่ง D_i

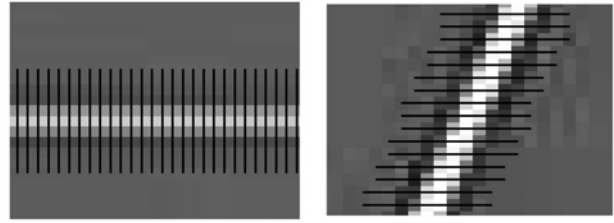
11. ทำวิธีการถดถอยเชิงเส้น จะทำให้ได้ผลลัพธ์คือเส้นสนาม

$$LineCourt_i = linearregression(Y_i)$$

$LineCourt_i$ ทำการเก็บผลลัพธ์จากการหาเส้นสนามในเวลา t

12. กำหนด $Line_i = LineCourt_i$ เพื่อใช้เป็นข้อมูลในการทำเฟรมถัดไป

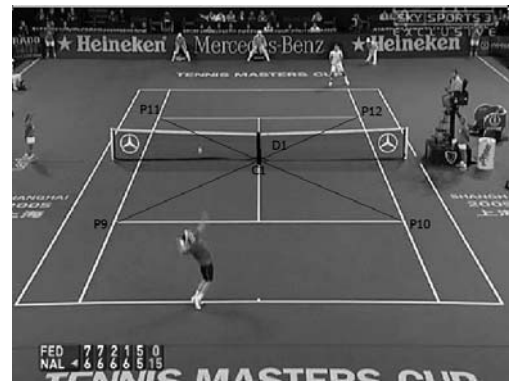
13. ทำซ้ำขั้นตอนที่ 4-12 จนครบเฟรม



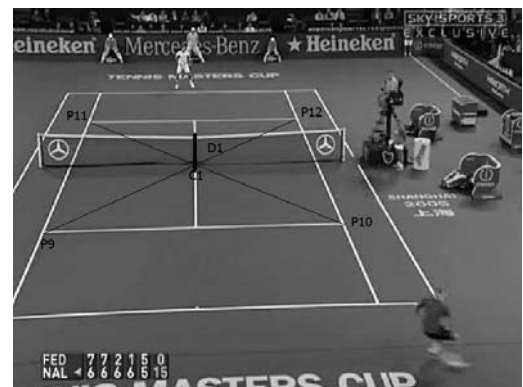
รูปที่ 2 แสดงถึงการขยายขนาดเส้น รูปซ้าย 2a ขวา 2b

จากรูปที่ 1 เส้นประคือเส้นที่เป็นส่วนของเน็ตด้านล่างเราทำการหาโดย

1. สร้างเส้นระหว่าง P11-P10 และ P12-P9
2. หาจุดตัดของเส้นทั้งสองเส้นจะได้จุดกึ่งกลางของเน็ต (C1) ซึ่งแสดงในรูปที่ 3 และ รูปที่ 4
3. หาระยะห่างระหว่าง C1 กับ P16 ได้ระยะห่างคือ (D1)
4. ได้เส้นประ ดังรูปที่ 1 ที่เป็นส่วนของเน็ตด้านล่าง โดยสร้างเส้นระหว่างจุด P15 กับ P16 และ จุด P16 กับ P17 นำเส้นทั้งสองที่สร้างบวกกับค่า D1



รูปที่ 3 แสดงจุดตัด C1 และระยะห่างที่เป็นแท่งสีดำ D1



รูปที่ 4 แสดงจุดตัด C1 และระยะห่างที่เป็นแท่งสีดำ D1

2.3. การตรวจหาลูกบอล

เมื่อใช้วีดิทัศน์การแข่งขันเทนนิสที่มีรายละเอียดต่ำโดยเฉพาะที่หามาจากอินเทอร์เน็ตผนวกกับการเคลื่อนที่ของลูกเทนนิสอย่างรวดเร็วเป็นการยากที่ใช้อุปกรณ์หรือสีของลูกเทนนิสเพื่อใช้ในการหาลูกเทนนิส

วิธีการที่ใช้ในการตรวจจับลูกเทนนิสเป็นดังนี้

เริ่มจากใช้วิธีความแตกต่างระหว่างเฟรมซึ่งเป็นวิธีการที่คำนวณได้ง่าย และรวดเร็วในการหาบริเวณที่มีการเคลื่อนที่หลังจากนั้นเราจะใช้เทคนิคความสัมพันธ์บางอย่างเพื่อจะขจัดบริเวณที่ไม่ใช่ลูกเทนนิส

1. ความแตกต่างระหว่างเฟรม

เป็นวิธีอย่างหนึ่งของการหาบริเวณที่มีการเคลื่อนไหว การคำนวณทำได้ ง่ายและรวดเร็วโดย นำภาพเฟรมที่ i ลบด้วยภาพเฟรมที่ $i-1$ $X_i - X_{i-1}$ ถ้าผลต่างมีค่าเกินกว่าขีดแบ่ง แสดงว่าจุดภาพมีความต่าง

2. ทำมอร์ฟอโลยี

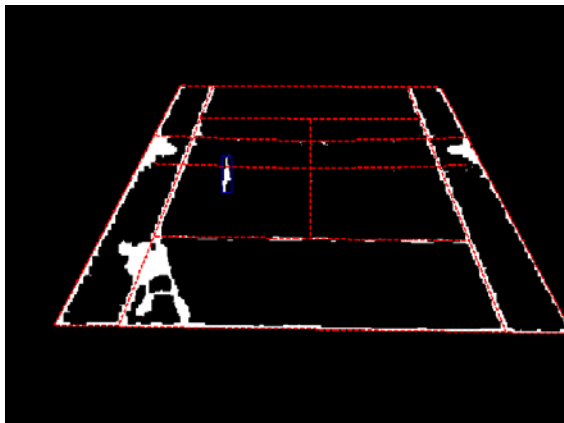
หลังจากทำความเข้าใจความแตกต่างระหว่างเฟรม เราจะทำการกระบวนการแบบปิด (Operation Closing) เพื่อทำการประติดประต่อส่วนที่เล็กๆของจุดภาพ

3. กรองลูกบอลโดยใช้ขนาด

กระบวนการหาลูกบอลเราใช้การกรองโดยใช้ช่วงขนาดลูกบอลที่ซึ่ง ช่วงของลูกบอลได้จากกระบวนการเทรนนิ่ง

4. กรองลูกบอลในขอบเขตของสนามโดยใช้ระยะห่างเส้นสนาม

เมื่อใดที่กล้องมีการเคลื่อน ผลลัพธ์ของการตรวจหาลูกบอลจะมีจำนวน ที่สูงอันเนื่องมาจากเส้นสนาม ดังนั้นเราจะแก้ไขข้อผิดพลาดนี้โดยการหา ระยะห่างเส้นสนาม โดยนำข้อมูลเส้นสนามที่ได้จากการทำในขั้นตอนที่ 2.2 ทำการหาขอบเขตในสนามจากนั้นหาระยะห่างเส้นสนามทุกเส้นกับ ผลลัพธ์ของการตรวจหาลูกบอล ดังรูปที่ 5 เส้นสีแดงจะแสดงถึงสนาม เทนนิสส่วนกรอบสี่เหลี่ยมที่เส้นที่บ่งชี้คือผลลัพธ์ของการตรวจจับลูกบอล



รูปที่ 5 ผลลัพธ์ของการตรวจหาลูกบอลด้วยวิธีลบเส้นสนาม

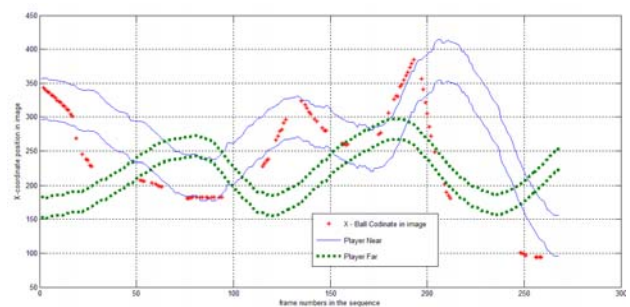
5. ลบจุดภาพที่เป็นบริเวณส่วนภาพของผู้เล่นทั้งฝั่งใกล้ฝั่งไกล

2.4. จำแนกท่าตีเทนนิส

เนื่องจากวิดิทัศน์ที่ใช้มีความละเอียดต่ำเราจึงใช้ข้อมูลเชิงเวลาในการ จำแนกท่าตีโดยข้อมูลที่นำมาใช้นั้นได้แก่ ตำแหน่งของผู้เล่นเทนนิสและ ข้อมูลลูกบอล ซึ่งข้อมูลทั้ง 2 นี้มีตำแหน่งพิกัดประกอบไปด้วย (x,y) ซึ่ง

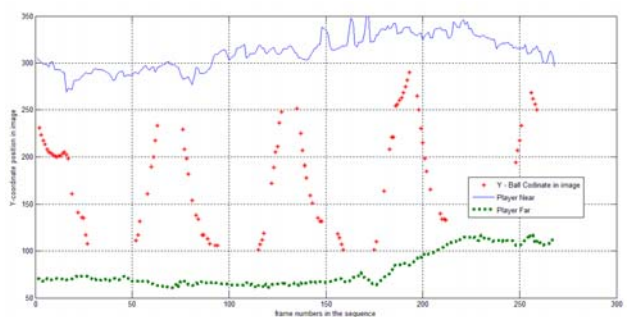
เราจะทำการสร้างกราฟ 2 กราฟ เพื่อใช้ในการจำแนกท่าตีเทนนิสโดยมี ขั้นตอนการทำงานดังนี้

1. แยกพิกัด x และ y ของตำแหน่งบอลและตำแหน่งผู้เล่น
2. สร้าง 2 กราฟโดยในแนวดิ่งของกราฟแรก (รูปที่ 6) คือ ตำแหน่ง พิกัด x ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรม โดย ข้อมูลที่ใช้ในการวาดกราฟแรก จะประกอบไปด้วย ตำแหน่งของ ลูกเทนนิสและตำแหน่งทางด้านซ้ายและขวาของผู้เล่นเทนนิส ส่วนกราฟที่สอง (รูปที่ 7) ในแนวดิ่งของกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรม โดยที่ข้อมูลที่ใช้ในการวาดกราฟ จะประกอบไปด้วย ตำแหน่งของลูกเทนนิส และตำแหน่งที่เป็นจุดศูนย์กลางของผู้เล่นเทนนิส



รูปที่ 6 กราฟที่ใช้เพื่อการจำแนกท่าตีเทนนิส

จากรูปที่ 6 เส้นสีน้ำเงิน (เส้นที่บน) จะแสดงถึงตำแหน่งของผู้เล่นเทนนิสฝั่ง ใกล้ที่ซึ่งมี 2 เส้นโดยเส้นด้านบนคือตำแหน่งทางด้านขวามือของผู้เล่นส่วน เส้นล่างคือตำแหน่งทางด้านซ้าย ส่วนเส้นสีเขียว (เส้นที่บน) แสดงถึงผู้เล่น เทนนิสฝั่งไกล โดยเส้นบนบนคือตำแหน่งทางด้านขวามือของผู้เล่นส่วนเส้นล่าง คือตำแหน่งทางด้านขวามือ จุดสีแดงคือ ตำแหน่งของลูกบอล

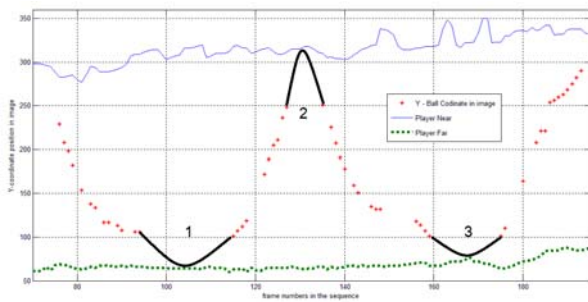


รูปที่ 7 กราฟที่ใช้หาเหตุการณ์ตีลูกเทนนิส

จุดสีแดงในรูปที่ 7 คือตำแหน่งของลูกเทนนิส ส่วนเส้นสีน้ำเงินด้านบน (เส้นที่บน) คือตำแหน่งของผู้เล่นฝั่งใกล้ และเส้นสีเขียวด้านล่าง (เส้นที่บน) คือตำแหน่งผู้เล่นฝั่งไกล

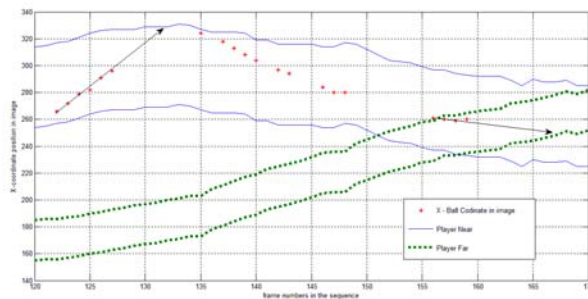
ใช้กราฟที่สอง ที่แสดงในรูปที่ 7 เพื่อช่วยในการหาช่วงเวลาตีเทนนิส โดยช่วงของเวลาที่ได้จะประกอบไปด้วย เฟรมก่อนตีลูกเทนนิส และ เฟรม

หลังตีลูกเทนนิส โดยเฟรมก่อนตีลูกเทนนิสจะมีลำดับเฟรมที่ต่ำกว่าเฟรมหลังตีเทนนิส ดังรูปที่ 8 ที่มีช่วงเฟรมของการตีเทนนิส 3 ช่วง



รูปที่8 ตัวอย่างเหตุการณ์ที่มีช่วงการตี 3 ช่วง

หลังจากได้ช่วงเวลาที่มีการตีเทนนิสเรานำช่วงเวลานี้ไปใช้ในการแบ่งแยกประเภทการตีโดยใช้กราฟในรูปที่ 6 พิจารณาว่าตำแหน่งบอลอยู่ด้านซ้ายหรือด้านขวาของผู้เล่นเทนนิสถ้าบอลอยู่ทางด้านซ้ายของผู้เล่นไกลผลลัพธ์คือตีทำฟลวแฮนด์ แต่ถ้าอยู่ทางด้านขวาผลลัพธ์ที่ได้คือตีแบ็คแฮนด์ ส่วนถ้าเป็นผู้เล่นฝั่งใกล้ตำแหน่งบอลอยู่ทางด้านขวาผลลัพธ์ที่ได้คือการตีทำฟลวแฮนด์ ถ้าอยู่ด้านซ้ายผลลัพธ์ที่ได้คือการตีทำแบ็คแฮนด์ แต่ถ้าตำแหน่งลูกบอลเข้าไปอยู่ระหว่างตำแหน่งด้านซ้ายและขวาของผู้เล่น แสดงได้ดังรูปที่ 9 เราจะทำการพิจารณาโดยถ้าทิศทางของลูกบอลเริ่มเข้าจากทางด้านขวาของผู้เล่นผลลัพธ์ที่ได้คือลูกบอลจะอยู่ทางด้านซ้ายของผู้เล่น แต่ถ้าลูกบอลเข้ามาทางด้านซ้ายของผู้เล่นผลลัพธ์ที่ได้คือลูกบอลจะอยู่ทางด้านขวาของผู้เล่น

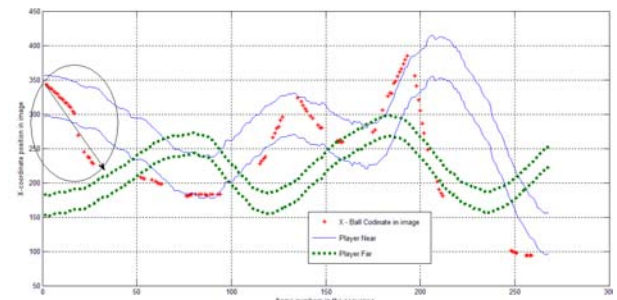


รูปที่9 แสดงตำแหน่งบอลที่อยู่ระหว่างกลางของผู้เล่นเทนนิส

จากรูปที่ 9 ทำการพิจารณาลูกศรีสีดำซ้ายสุดจะเห็นได้ว่าลูกเทนนิสจะเข้ามาทางด้านซ้ายของผู้เล่นฝั่งใกล้ ดังนั้นผลลัพธ์ที่ได้ลูกเทนนิสจะเคลื่อนที่ไปทางด้านขวาซึ่งได้ทำตีแบบโฟว์แฮนด์ ส่วนลูกศรีสีดำขวาสุดที่ลูกเทนนิสได้เข้ามาจากทางด้านซ้ายของผู้เล่นฝั่งไกล ดังนั้นผลลัพธ์ที่ได้คือลูกเทนนิสจะต้องไปทางด้านขวา จึงได้ทำตีโฟว์แฮนด์

ในการหาว่าผู้เล่นเทนนิสฝั่งใดเสิร์ฟนั้น ทำการพิจารณาโดยใช้กราฟจากรูปที่ 6 เพื่อทำการดูแนวโน้มของตำแหน่งของลูกเทนนิสว่ามีความใกล้กับผู้เล่นเทนนิสฝั่งใด

จากรูปที่ 10 เป็นการแสดงตัวอย่างของการจำแนกเหตุการณ์เสิร์ฟ จะเห็นได้ว่าในบริเวณวงรีนั้น พิกัดของลูกบอลมีความใกล้กับตำแหน่งของผู้เล่นเทนนิสใกล้ และมีแนวโน้มของการเคลื่อนที่ของลูกบอลเข้าไปสู่ผู้เล่นไกล ดังนั้นจึงเป็นการเสิร์ฟของผู้เล่นฝั่งใกล้



รูปที่10 แสดงตำแหน่งบอลวิเคราะห์เหตุการณ์เสิร์ฟ

3. ผลการทดลอง

เราได้ทำการทดสอบวิธีการที่นำเสนอโดยใช้วิดีโอทั้งหมด 8 คลิป ที่ซึ่งได้ทำการตัดมาจากการแข่งขันเทนนิส รายการTENNIS MASTER, US OPEN และ ATP โดยแต่ละคลิปที่ได้ ทำการทดสอบเป็นการแข่งขันเทนนิสที่แตกต่างกันทั้งหมด 6 คู่ จำนวนเฟรมที่ทำการทดลองมีทั้งหมด 2926 เฟรม ในแต่ละวิดีโอคลิปเริ่มจากการเสิร์ฟ โดยต้องไม่เป็นการเสิร์ฟเสียและจะจบลงด้วยการเสิร์ฟคะแนน 1 คะแนน ทั้งนี้ผู้เล่นเทนนิสทั้งไกลและใกล้จะต้อง จับไม้เทนนิสด้วยมือขวา

ทำการทดลองทั้ง 8 คลิป โดยใช้วิธีการที่ได้กล่าวมาแล้วในหัวข้อที่ 2 คือ ทำการติดตามหาตำแหน่งผู้เล่นเทนนิสทั้งไกลและใกล้ จากนั้นทำการตรวจหาลูกบอล โดยที่นำผลลัพธ์ของการหาเส้นสนามเทนนิสมาใช้ร่วมด้วย เพื่อเพิ่มความถูกต้อง สุดท้ายทำการจำแนกท่าตีเทนนิส โดยใช้ผลของการตรวจหาลูกบอลร่วมกับตำแหน่งผู้เล่นเทนนิส ได้ผลจากการจำแนกมา 3 ประเภท นำผลลัพธ์ที่ได้มาเปรียบเทียบกับ กราวด์ทรู (ground truth) ในตารางที่ 1

ตารางที่ 2 แสดงถึงผลลัพธ์ของการตีลูกบอล ส่วนตารางที่ 3 แสดงผลลัพธ์ที่ได้จากการจำแนกท่าตีเทนนิส และในรูปที่ 11 ในแถบบนแสดงผลลัพธ์ของการใช้ขั้นตอนวิธีแยกค่าเฉลี่ย ร่วมกับ ความแตกต่างระหว่างเฟรม ส่วนแถวล่างเป็นการใช้ขั้นตอนวิธีแยกค่าเฉลี่ย เพียงอย่างเดียวจะเห็นได้ว่าถ้าใช้เพียงขั้นตอนวิธีแยกค่าเฉลี่ย ทำให้เกิดความผิดพลาดในกรณีตีเสิร์ฟของผู้เล่นที่มีลีลาเหมือนสนามเทนนิสหรือลีลาเหมือนกรรมการ

ตาราง 1. แสดงจำนวนกราวด์ทรู

กราวด์ทรู					
ผู้เล่น ไกล (FAR)			ผู้เล่น ใกล้ (NEAR)		
SERVE	FH	BH	SERVE	FH	BH
2	21	19	6	16	18

ตาราง 2. ผลลัพธ์ของการตีลูกบอล

Far/Near	TP	FN	FP	PERCISION	RECALL
FAR	38	2	0	100.00%	95.00%
NEAR	33	1	1	97.06%	97.06%

ตาราง 3. ผลลัพธ์ของการจำแนกท่าตีเทนนิส

	TP	FN	FP	PRECISION	RECALL
FAR-PLAYER					
FH	19	1	1	95.00%	95.00%
BH	17	1	1	94.44%	94.44%
SERVE	2	0	0	100.00%	100.00%
NEAR-PLAYER					
FH	16	0	1	94.12%	100.00%
BH	16	1	0	100.00%	94.12%
SERVE	6	0	0	100.00%	100.00%

TP = true positive, FN = false negative, FP = false positive

PRECISION = $TP / (TP + FP)$, RECALL = $TP / (TP + FN)$

BH = ท่าตีแบ็คแฮนด์, FH = ท่าตีโฟลวแฮนด์, SERVE = เซิร์ฟ



รูปที่11 แสดงผลลัพธ์ของการติดตาม

4. สรุปผลและงานที่จะทำต่อไป

ในบทความนี้เราได้นำเสนอการติดตามผู้เล่น การหาสนามเทนนิส และการตรวจหาลูกบอลในวิดีโอการแข่งขันเทนนิสที่มีความละเอียดวิดีโอที่ต่ำ ในมุมมองกล้องระยะไกลเพียงอย่างเดียว ใช้อัลกอริทึม ขั้นตอนวิธีย้ายค่าเฉลี่ย ร่วมกับ ความแตกต่างระหว่างเฟรมโดยใช้ปริภูมิสี RGB หาเส้นของสนามเทนนิสโดยการดูค่าความสว่างของจุดภาพและทำการประมาณเส้นตรง นอกจากนี้เรายังตรวจสอบและติดตามตำแหน่งของลูกเทนนิสโดยความแตกต่างระหว่างเฟรม รวมทั้งใช้เทคนิคของความสัมพันธ์บางอย่างเพื่อกำจัดสิ่งที่ไม่ใช่ลูกบอล ใช้การเคลื่อนที่ของผู้เล่นและตำแหน่งของลูกบอลเพื่อจำแนกท่าตีเทนนิสเป็น โฟลวแฮนด์ แบ็คแฮนด์ และ เซิร์ฟ

ในอนาคตเราวางแผนที่จะทำให้ระบบมีความสามารถในการจำแนกท่าตีเทนนิสเพิ่มมากขึ้น เช่น วอลเลย์

เอกสารอ้างอิง

- [1] J.R. Wang and N. Parameswaran, "Survey of sports video analysis: research issues and applications," Proc. Pan-Sydney Area Workshop on Visual Information Processing (VIP2003), Sydney, Australia.
- [2] X. Yu and D. Farin, "Current and emerging topics in sports video processing," Proc. of IEEE ICME, 2005.

- [3] J. Kittler, K. Messer, W.J. Christmas, B. Levenaise-Obadia, and D. Koubaroulis, "Generation of semantic cues for sports video annotation," Proc. IEEE ICIP 2001.
- [4] M. Petkovic, W. Jonker, and Z. Zivkovic, "Recognizing strokes in tennis videos using Hidden Markov Models," Proc. of IASTED VIIP, 2001.
- [5] G. Zhu, C. Xu, Q. Huang, and W. Gao, "Action recognition in broadcast tennis video," Proc. the 18th Intl Conf. on Pattern Recognition (ICPR'06), Vol.1, 2006.
- [6] H. Shah, P. Chokalingam, B. Paluri, N. Pradeep, and R. Balasubramanian. Automated stroke classification in tennis. In ICIAR07, pages 1128--1137, 2007.
- [7] X. Yu, C.H. Sim, J. R. Wang, and L.F.Cheong, "A trajectory-based ball detection and tracking algorithm in broadcast tennis video," Proc. Intl Conf. on Image Processing (ICIP'04), 2004.
- [8] C.O. Conaire, P. Kelly, D. Connaghan, and N.E. O'Connor, "TENNISSENSE: A platform for extracting semantic information from multi-camera tennis data," Proc. Intl Conf. on Digital Signal Processing (DSP'09), 2009
- [9] E. Kijak, G. Gravier, P. Gros, L.Oisel, and F. Bimbot, "HMM based structuring of tennis videos using visual and audio cues," Proc. of IEEE ICME, 2003.
- [10] F. Yan, W. Christmas, and J. Kittler, "A tennis ball tracking algorithm for automatic annotation of tennis match," British Machine Vision Conference, 2005, pp. 619--628.
- [11] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," IEEE Trans on Pattern Analysis and Machine Intelligence (PAMI), pp.564-577, May 2003.
- [12] S.T. Birchfield and S. Rangarajan, "SpatioGrams versus histograms for region-based tracking," Proc. IEEE CVPR'05, pp.1158-1163, June 2005.