

ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

การจำแนกท่าตีเทนนิสจากวีดิทัศน์
โดยใช้วิธีการเคลื่อนที่สัมพัทธ์ของผู้เล่นและลูกบอล

TENNIS STROKE CLASSIFICATION FROM VIDEO
USING RELATIVE TRAJECTORIES OF PLAYERS AND A BALL



เลขหมู่.....
เลขทะเบียน.....
วัน, เดือน, ปี.....

b. 12 กย 60 24
i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2554

KMITL-2011-IT-M-001-003

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**TENNIS STROKE CLASSIFICATION FROM VIDEO
USING RELATIVE TRAJECTORIES OF PLAYERS AND A BALL**



**A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

2011

KMITL-2011-IT-M-001-003

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2011

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์

การจำแนกท่าตีเทนนิสจากวิดีโอที่ใช้วิธีการเคลื่อนที่
สัมพันธ์ของผู้เล่นและลูกบอล

นักศึกษา

นายโกษิต เตชะบริกิติ

รหัสประจำตัว

50066525

ปริญญา

วิทยาศาสตรมหาบัณฑิต

สาขาวิชา

เทคโนโลยีสารสนเทศ

พ.ศ.

2554

อาจารย์ผู้ควบคุมวิทยานิพนธ์

รศ.ดร.อาริต ธรรมโน

อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม

ผศ.ดร.ธนารัตน์ ชลิดาพงศ์

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้ นำเสนอขั้นตอนวิธีการจำแนกท่าตีกีฬาเทนนิสจากวิดีโอที่ใช้วิธีการเคลื่อนที่สัมพันธ์ของผู้เล่นและลูกบอล ที่เผยแพร่บนอินเทอร์เน็ต ซึ่งมักเป็นวิดีโอที่มีความละเอียดต่ำ มีสัญญาณรบกวนสูง และมีมุมมองจากกล้องที่ตั้งในระยะไกลเพียงมุมเดียว เริ่มจากการติดตามตำแหน่งผู้เล่นทั้งสอง ด้วยขั้นตอนวิธีย้ายค่าเฉลี่ย ร่วมกับการหาความแตกต่างระหว่างเฟรม กระทำกับฮิสโทแกรมสี จากนั้นทำการตรวจจับเส้นของสนามเทนนิส โดยวิเคราะห์จากค่าความเข้มของจุดภาพ และทำการตรวจจับตำแหน่งของลูกเทนนิส โดยการวิเคราะห์หาบริเวณจุดภาพที่มีการเคลื่อนไหว รวมถึงนำเสนอวิธีในการลดความผิดพลาดที่ได้จากการตรวจจับตำแหน่งลูกเทนนิสที่ผิดพลาดให้กับระบบ โดยใช้ตำแหน่งเส้นของสนามเทนนิส สุดท้ายนำตำแหน่งผู้เล่นทั้งสอง และตำแหน่งลูกบอล มาใช้ในการจำแนกท่าตีเทนนิสออกเป็น เสิร์ฟ แบคแฮนด์ โฟร์แฮนด์ แบคแฮนด์ หน้าเน็ต และโฟร์แฮนด์หน้าเน็ต

Thesis Title	Tennis Stroke Classification from Video Using Relative Trajectories of Players and a Ball
Student	Mr. Kosit Teachabarikiti
Student ID.	50066525
Degree	Master of Science
Program	Information Science
Year	2011
Thesis Advisor	Assoc. Prof. Dr. Arit Thammano
Thesis Co-Advisor	Asst. Prof. Dr. Thanarat Chalidabhongse

ABSTRACT

This thesis presents algorithm for tennis stroke classification from video published in the Internet using relative trajectories of players and a ball. These videos are often low-resolution with noisy and captured by single far-view camera. Algorithm begin by tracking two players using frame differencing and mean shift tracking algorithm on color histogram. Then, tennis court line detection is performed by analyzing the intensity of the image. To detect and locate the tennis ball, our approach analyze moving of pixel as well as offering techniques to eliminate false detection by using the tennis court line. Based on both players and tennis ball positions, the system can precisely classify the players' action into serve, backhand ground stroke, forehand ground stroke, half court backhand, and half court forehand.

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จได้ด้วยความกรุณาและความช่วยเหลืออย่างดียิ่งจาก รศ.ดร.อาริต ธรรมโน อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผศ.ดร. ธนารัตน์ ชลิตาพงษ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ที่ได้ตลอดเวลาในการให้คำปรึกษา คำแนะนำอันมีประโยชน์ และประสบการณ์ด้านการทำงานอันมีค่า ตลอดจนช่วยตรวจทานแก้ไขวิทยานิพนธ์ และบทความตีพิมพ์ต่างๆจนวิทยานิพนธ์ฉบับนี้ลุล่วงไปด้วยดี

ขอขอบพระคุณ ผศ.พงษ์ศักดิ์ วิสูตรกาญจนชัย และ รศ.สมชาติ รุ่งเรืองสรการ จากคณะวิทยาศาสตร์ประยุกต์ ภาควิชาคณิตศาสตร์ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ ที่ให้คำปรึกษาและชี้แนะ

ขอขอบพระคุณอาจารย์ทุกท่าน จากคณะเทคโนโลยีสารสนเทศ ที่ได้ประสิทธิ์ประสาทวิชาความรู้ให้ตลอดมาตั้งแต่เข้าเรียนในคณะนี้

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ทุกท่าน ที่กรุณาช่วยแนะนำงานวิจัยในส่วนต่างๆ และวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ คุณย่า, หนิง และญาติพี่น้องทุกท่านที่คอยสนับสนุนและให้กำลังใจ

ขอขอบพระคุณบิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ ให้การสนับสนุนในทุกๆ เรื่อง ทำให้ข้าพเจ้าทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดี

สำหรับคุณงามความดีอันใดที่เกิดจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบให้กับบิดามารดา ซึ่งเป็นที่รักและเคารพยิ่ง และครูอาจารย์ที่เคารพทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้และถ่ายทอดประสบการณ์ที่ดีให้แก่ข้าพเจ้า ตลอดจนผู้มีพระคุณทุกท่าน

ได้รับทุนสนับสนุนการทำวิทยานิพนธ์จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

โกยิต เตชะบริกิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	VIII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	2
1.3 สมมุติฐานของการศึกษา.....	2
1.4 ทฤษฎีหรือแนวความคิดในงานวิจัย.....	2
1.5 ขอบเขตของการดำเนินงานวิจัย.....	2
1.6 ขั้นตอนการศึกษา.....	3
1.7 โครงสร้างวิทยานิพนธ์.....	3
บทที่ 2 ทฤษฎีพื้นฐาน และงานวิจัยที่เกี่ยวข้องกับการจำแนกทำดี กีฬาเทนนิสจากวิดีโอ.....	4
2.1 งานวิจัยที่เกี่ยวข้องกับการจำแนกทำดีเทนนิสแบบอัตโนมัติ.....	4
2.2 วิธีย้ายค่าเฉลี่ย.....	7
2.3 การติดตามโดยใช้วิธีย้ายค่าเฉลี่ย.....	10
2.3.1 การแทนตำแหน่งเป้าหมาย.....	10
2.3.2 แบบจำลองเป้าหมาย.....	11
2.3.3 แคนดิเดตเป้าหมาย.....	11
2.3.4 หาค่าน้ำหนักและตำแหน่งใหม่.....	12
2.4 การถอดอยเชิงเส้น.....	13
2.5 การหาความแตกต่างระหว่างเฟรม.....	15
บทที่ 3 การจำแนกทำดีกีฬาเทนนิสจากวิดีโอ.....	17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
3.1 การติดตามผู้เล่นเทนนิส.....	19
3.2 การตรวจหาเส้นของสนามเทนนิส.....	24
3.3 การตรวจหาตำแหน่งของลูกเทนนิส.....	28
3.3.1 การหาความแตกต่างระหว่างเฟรม.....	29
3.3.2 มอร์ฟอโลยี.....	30
3.3.3 กรองบล็ออบโดยใช้ขนาด.....	30
3.4 วิธีลดปริมาณความผิดพลาดของการตรวจหาลูกเทนนิส.....	32
3.4.1 กรองบริเวณสนามเทนนิส.....	33
3.4.2 ลบเส้นสนามเทนนิส ป้ายโฆษณา และ บล็ออบที่อยู่ในผู้เล่นเทนนิส.....	34
3.5 การจำแนกท่าตีเทนนิส.....	34
3.5.1 ตรวจหาเหตุการณ์ของการตีลูกเทนนิส.....	35
3.5.2 จำแนกท่าตีเทนนิส.....	36
บทที่ 4 วิธีการทดลองและผลการทดลอง.....	41
4.1 วัตถุประสงค์ที่ใช้ในการทดลอง.....	41
4.2 การกำหนดค่าและวิธีการทดลอง.....	42
4.2.1 ขั้นตอนการติดตามผู้เล่นเทนนิส.....	42
4.2.2 ขั้นตอนการติดตามเส้นสนามเทนนิส.....	43
4.2.3 ขั้นตอนการหาลูกเทนนิส.....	44
4.3 การทดลองและผลลัพธ์ของการติดตามผู้เล่น.....	45
4.4 การทดลองและผลลัพธ์ของการติดตามเส้นสนามเทนนิส.....	46
4.5 การทดลองและผลลัพธ์ของการตรวจจับลูกเทนนิส.....	48
4.6 การทดลองและผลลัพธ์ของการจำแนกท่าตีเทนนิส.....	52
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	59
5.1 สรุปผลการวิจัย.....	59
5.2 ข้อเสนอแนะ.....	60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
เอกสารอ้างอิง.....	61
ภาคผนวก	
ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่.....	63
ประวัติผู้เขียน.....	77



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 ตัวอย่างรูปแบบฟังก์ชันเคอร์เนลแบบต่างๆ จาก [13]	8
4.1 กราฟด์ทฤษฎี	52
4.2 ผลลัพธ์ของการตีลูกเทนนิส	52
4.3 ผลลัพธ์ของการจำแนกท่าตีเทนนิส	53
4.4 Confusion Matrix	53



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 (ก) แบบจำลองเป้าหมาย (\hat{q}) (ข) แคนดิเดทเป้าหมาย ($\hat{p}(y)$)	10
2.2 กราฟเพื่อจะประมาณการถดถอยเชิงเส้น	13
2.3 ขั้นตอนวิธีความแตกต่างระหว่างเฟรม.....	16
3.1 มุมกล้องระยะไกล.....	17
3.2 แผนภาพแสดงขั้นตอนของการจำแนกท่าตีเทนนิสจากวิดีโอ.....	18
3.3 แผนภาพแสดงวิธีการติดตามผู้เล่นเทนนิส.....	19
3.4 (ก) จุดศูนย์กลางและขนาดวงรีที่กำหนด (ข) ภาพหลังจากใช้วิธีหาความแตกต่างระหว่างเฟรม (ค) ภาพจากการทำตัวดำเนินการแบบปิด (ง) ภาพหลังจากคืนค่าสีจุดภาพ	21
3.5 ตัวอย่างผลลัพธ์จากการติดตามตำแหน่งของผู้เล่นทั้งสองฝั่ง	22
3.6 กระบวนการติดตามด้วยวิธี การติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย	23
3.7 แผนภาพแสดงวิธีตรวจหาเส้นสนามเทนนิส	24
3.8 (ก) (ข) จุดที่ถูกกำหนด 17 จุดในสนามเทนนิส.....	25
3.9 (ก) นำจุดจากที่กำหนดมาสร้างเส้น (ข) ภาพการขยายขอบเขตพิกัด (ค) ภาพขยายขนาดเส้น พิกัด x (ง) ภาพขยายขนาดเส้น พิกัด y	27
3.10 (ก) (ข) ตัวอย่างผลลัพธ์ของการหาเส้นสนาม	27
3.11 (ก) (ข) จุดตัด C1 และระยะห่างที่เป็นแท่งสีดำ D1	28
3.12 แผนภาพแสดงการตรวจหาตำแหน่งของลูกเทนนิส	29
3.13 ตัวอย่าง ผลของการหาความแตกต่างระหว่างเฟรม (ก) ด้วยค่าขีดเริ่มเปลี่ยน = 15 (ข) ด้วยค่าขีดเริ่มเปลี่ยน = 25	29
3.14 ตัวอย่างภาพที่ได้ (ก) ก่อนทำตัวดำเนินการแบบปิด (ข) หลังทำตัวดำเนินการแบบปิด.....	30
3.15 (ก) ตัวอย่าง ภาพของลูกเทนนิสเมื่อมีการเคลื่อนที่ (ข) ทำการขยายภาพในตำแหน่งลูกเทนนิส.....	31
3.16 ผลลัพธ์ของการกรองบลิบโดยใช้ขนาด.....	31
3.17 ตัวอย่าง เฟรมที่กล้องมีการแพนตามผู้เล่น	32
3.18 แผนภาพแสดงการลดความผิดพลาดของลูกบอล	33
3.19 ค่าของจุดภาพที่อยู่ในสนามเทนนิส	33
3.20 ผลลัพธ์ของการตรวจหาลูกเทนนิส.....	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.21 แผนภาพแสดงวิธีจำแนกท่าตีเทนนิส	35
3.22 กราฟที่ใช้หาเหตุการณ์ตีลูกเทนนิส.....	35
3.23 ตัวอย่างเหตุการณ์ที่มีช่วงการตี 3 ช่วง.....	36
3.24 กราฟที่ใช้เพื่อการจำแนกท่าตีเทนนิส	37
3.25 ตำแหน่งบอลที่ค้ำนอกของผู้เล่นเทนนิส	37
3.26 ตำแหน่งบอลที่อยู่ระหว่างกลางของผู้เล่นเทนนิส.....	38
3.27 ตำแหน่งพิจารณาการตีหน้าเน็ต	39
3.28 กราฟเส้นรัศมีของผู้เล่นเทนนิสฝั่งใกล้	39
3.29 กราฟเส้นรัศมีของผู้เล่นเทนนิสฝั่งไกล	40
4.1 ตัวอย่างของวิดิทัศน์การแข่งขันกีฬาเทนนิสที่มีการบันทึกมาจากการถ่ายทอดสด.....	41
4.2 ตัวอย่างของภาพนิ่งที่ใช้เป็นอินพุต ซึ่งตัดมาจากวิดิทัศน์.....	42
4.3 ตัวอย่างของการกำหนดตำแหน่ง และขนาดเพื่อทำการติดตามผู้เล่นเทนนิส	43
4.4 ตัวอย่างการกำหนด 17 จุดเพื่อใช้ในการหาเส้นสนามเทนนิส	44
4.5 ตัวอย่างของการหาตำแหน่งลูกเทนนิส	44
4.6 กราฟแสดง ผลของการเรนของชุดข้อมูล 192 เฟรม.....	45
4.7 (ก) (ค) (จ) ผลลัพธ์ของการติดตามโดยใช้วิธีย้ายค่าเฉลี่ย ร่วมกับวิธีการหาความแตกต่าง ระหว่างเฟรม (ข) (ง) (ฉ) ผลลัพธ์ของการติดตามโดยใช้วิธีย้ายค่าเฉลี่ยอย่างเดียว.....	45
4.8 (ก) (ข) (ค) ผลลัพธ์ที่ได้จากการติดตามหาเส้นสนามเทนนิส.....	47
4.9 (ก) (ข) (ค) (ง) (จ) (ฉ) (ช) (ฉ) ผลลัพธ์ที่ได้จากการติดตามลูกเทนนิส.....	49
4.10 ภาพตัวอย่างกรณีจำแนกท่าตีผิด (ก) (ข) (ค) (ง) ลูกเทนนิสเคลื่อนที่อย่างรวดเร็วและประชิด.....	55
4.11 ภาพตัวอย่างกรณีจำแนกท่าตีผิด (ก) (ข) ลูกเทนนิสเคลื่อนที่อยู่นอกสนามเทนนิส รวมทั้งหายไปจากเฟรม	57

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ด้วยปัจจุบันเทคโนโลยีของการบันทึกข้อมูลมีความเจริญก้าวหน้า จึงทำให้มีจำนวนวิดีโอที่สร้างขึ้นอย่างมากมาย และจากจำนวนวิดีโอที่เพิ่มขึ้นนั้น ได้ทำให้เกิดงานวิจัยทางด้านคอมพิวเตอร์วิชัน (Computer Vision) เพื่อช่วยในการจัดเรียง จัดเก็บ หรือเป็นการใช้ประโยชน์จากข้อมูลที่อยู่ในวิดีโอ เป็นต้น โดยตัวอย่างของงานวิจัยประเภทนี้ เช่น การบรรยายวิดีโอ (Video Annotation) ที่ซึ่งในการบรรยายวิดีอนั้นมีเป้าหมายหรือโจทย์หลักคือ การหาใจความสำคัญในวิดีโอ โดยหนึ่งในงานที่นำโจทย์นี้มาใช้คือระบบการบรรยายกีฬาอัตโนมัติ

หากเรามีเทคโนโลยีที่ช่วยในการบรรยายกีฬาอัตโนมัติ ก็จะทำให้ทราบถึงรายละเอียดของการแข่งขันกีฬานั้นๆว่ามีเหตุการณ์อะไรเกิดขึ้นบ้างของในแต่ละเวลา โดยหนึ่งในกีฬาที่คนนิยมรับชมในปัจจุบันคือ กีฬาเทนนิส ซึ่งเป็นกีฬาที่แบ่งผู้เล่นออกเป็นสองฝ่ายโดยมีเน็ตกั้นกลางระหว่างผู้เล่นและมีขอบเขตของสนามที่แน่นอน

อย่างที่ได้อ้างไว้ข้างต้นว่า การบรรยายกีฬาอัตโนมัติ จำเป็นที่ต้องหาใจความสำคัญ โดยหนึ่งในใจความสำคัญของการบรรยายกีฬาเทนนิสนั้นคือ ท่าตีเทนนิส อย่างไรก็ตามไม่ใช่เรื่องง่ายเลยที่จะทำการจำแนกท่าตี ซึ่งมีปัญหาที่ต้องพิจารณา อาทิเช่น ลูกเทนนิสที่มีรูปทรงกลมแต่เมื่อเคลื่อนที่รูปทรงจะถูกเปลี่ยนเป็นเส้นตรง เพราะ เกิดจากอัตราเฟรมของภาพไม่สัมพันธ์กับการเคลื่อนที่ เป็นปัญหาที่เกิดจากโมชันเบลล (motion blur) หรือจะเป็นปัญหาที่เกิดจากคุณภาพของวิดีโอที่มีรายละเอียดต่ำ จึงทำให้รายละเอียดข้อมูลของแต่ละส่วนของภาพไม่ชัดเจน หรือมุกกล้องมีการเปลี่ยนแปลงอยู่ตลอดเวลา จากปัญหาที่ยกมาข้างต้นนั้นจะเห็นได้ว่าไม่ง่ายเลยที่จะทำการจำแนกท่าตีกีฬาเทนนิส

ผลลัพธ์ที่ได้จากการจำแนกท่าตีกีฬาเทนนิส นอกจากนำไปใช้ประโยชน์ในเรื่องของการบรรยายกีฬาเทนนิสแบบอัตโนมัติแล้ว ยังสามารถนำไปใช้ประโยชน์ในเรื่องของ การทำระบบค้นคืนวิดีโอ (Content based video retrieval) ซึ่งเป็นวิธีสืบค้นวิดีโอจากการใช้ลักษณะเฉพาะของวิดีโอ เช่น ลักษณะรูปร่างหรือองค์ประกอบสี ของวัตถุภายในวิดีโอ ซึ่งจากเดิมการสืบค้นวิดีอนั้นใช้การป้อนชื่อไฟล์หรือคำอธิบายถึงวิดีโอ และจะเห็นได้ว่าเป็นการใช้ข้อมูลที่ไม่ใช้ลักษณะของวิดีอนั้นๆ เพราะการตั้งชื่อไฟล์หรือคำอธิบายนั้นมาจากการทำด้วยมนุษย์ ซึ่งแต่ละคนสามารถให้คำจำกัดความของวิดีโอแตกต่างกันออกไป โดยที่เราอาจจะค้นหาท่าตีเทนนิสจากวิดีโอได้ หรือนำวิธีการในงานวิจัยนี้ไปใช้ในการประยุกต์ลักษณะของกีฬาที่มีความใกล้เคียงกัน เช่น แบดมินตัน หรือวอลเลย์บอล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

วิทยานิพนธ์ฉบับนี้มีความมุ่งหวัง เพื่อพัฒนาวิธีการจำแนกท่าตีของผู้เล่นทั้ง 2 ฝ่าย และสามารถจำแนกท่าตีได้ทันทีโดยไม่ต้องผ่านกระบวนการรู้จำ โดยมุมมองของกล้องมาจากระยะไกลเพียงมุมมองเดียวและใช้วิถีทัศน์ที่เผยแพร่จากอินเทอร์เน็ต ซึ่งระบบการจำแนกท่าตีกีฬาเทนนิสที่สมบูรณ์นั้น สามารถนำไปใช้ได้หลากหลายด้าน อาทิเช่น สามารถนำไปสร้างระบบบรรยายวิถีทัศน์การแข่งขันเทนนิส ซึ่งทำให้เราสามารถรู้ถึง เหตุการณ์ที่เกิดขึ้นของในแต่ละช่วงเวลา หรือจะเป็นการสร้างระบบค้นคืนวิถีทัศน์ เพื่อช่วยในการค้นหาวิถีทัศน์การแข่งขันเทนนิสจากระบบ

1.3 สมมุติฐานของการศึกษา

เราสามารถที่จะใช้วิธีการเคลื่อนที่สัมพันธ์ระหว่างผู้เล่นและลูกเทนนิส เพื่อที่จะหาเหตุการณ์ของการตีลูกเทนนิส โดยที่วิถีของลูกเทนนิสจะต้องมีลักษณะพุ่งเข้าหาผู้เล่น จากนั้นพิจารณาถึงเวลาที่มีการตีลูกเทนนิส เพื่อใช้ทำการจำแนกท่าตี โดยที่หากเราทราบถึงตำแหน่งของลูกเทนนิสว่าอยู่ในตำแหน่งใดของผู้เล่น กล่าวคือถ้าเป็นการจับไม้ด้วยมือขวา เมื่อลูกเทนนิสเคลื่อนที่มาทางด้านซ้ายก็จะจำแนกท่าตีคือแบคแฮนด์ และถ้าเคลื่อนที่มาทางด้านขวาก็จะจำแนกเป็นท่าตีโฟร์แฮนด์

1.4 ทฤษฎีหรือแนวคิดในงานวิจัย

จากสมมุติฐานที่ว่าถ้าเราทราบความสัมพันธ์ของตำแหน่งของลูกเทนนิสและตำแหน่งผู้เล่นเทนนิส เราจะสามารถจำแนกท่าตีเทนนิสได้ ซึ่งโดยปกติแล้วลูกเทนนิสในเวลาที่มีการแข่งขันเทนนิสจะมีการเคลื่อนที่อยู่ตลอดเวลา ดังนั้นในการจะแยกแยะลูกเทนนิสออกจากฉากหลัง (Background) ในวิทยานิพนธ์ฉบับนี้จึงใช้วิธีการ หาความแตกต่างระหว่างเฟรม (Frame Differencing) เพื่อทำการหาบริเวณของสิ่งที่มีการเคลื่อนไหวในภาพ เมื่อพิจารณาถึงส่วนที่เป็นลูกเทนนิสจะพบว่า สามารถใช้การกรองขนาดลูกเทนนิสเพื่อทำการหาวัตถุที่มีการเคลื่อนไหว กล่าวคือขนาดของลูกเทนนิสที่ปรากฏในวิถีทัศน์จะมีขนาดที่เท่าๆกัน จากนั้นทำการติดตามตำแหน่งของผู้เล่นเทนนิส ด้วยวิธีการติดตามโดยใช้วิธีย้ายค่าเฉลี่ย (Mean Shift Tracking) ที่สามารถใช้ข้อมูลสีในการกำหนดตัวแบบ ที่ใช้ในการติดตามเฉพาะได้ ซึ่งข้อมูลที่ี้ใช้จะประกอบไปด้วยสีเสื้อและสีกางเกงของผู้เล่นเทนนิส สุดท้ายนำตำแหน่งลูกเทนนิสและผู้เล่นเทนนิสมาวิเคราะห์เพื่อประมาณท่าตีเทนนิส

1.5 ขอบเขตของการดำเนินงานวิจัย

ในวิทยานิพนธ์ฉบับนี้นำเสนอการบรรยายการตีเทนนิสซึ่งเป็นการนำวิถีทัศน์การแข่งขันเทนนิสจากอินเทอร์เน็ต มาทำการตัดออกเป็นคลิป โดยแต่ละคลิปจะเริ่มจากการเสิร์ฟโดยต้องไม่
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นการเสิร์ฟเสียดและจบด้วยการได้หรือเสียดคะแนน 1 คะแนน โดยมุมมองของกล้องมาจาก ระยะไกลเพียงมุมมองเดียว ต้องทำการกำหนด ตำแหน่ง,ขนาด,มือที่จับไม้เทนนิสของผู้เล่น โดยที่ ตำแหน่งของสนามเทนนิส ต้องถูกกำหนดด้วยมือ และสนามเทนนิสเป็นแบบพื้นแข็ง (Hard Court) จากนั้นทำการประมวลผลด้วยโปรแกรม MATLAB โดยเป็นการประมวลผลแบบออฟไลน์ ทำทาง การตีเทนนิสที่เราได้ทำการแบ่งประกอบไปด้วย เสิร์ฟ แบคแฮนด์ โฟร์แฮนด์ แบคแฮนด์หน้าเน็ต และ โฟร์แฮนด์หน้าเน็ต

1.6 ขั้นตอนการศึกษา

ในขั้นตอนของการศึกษานี้ ได้แสดงลำดับการทำงานตั้งแต่เริ่มต้นจนถึงสิ้นสุดการทำงานวิจัย ดังรายละเอียดต่อไปนี้

1. ศึกษาทฤษฎีและงานวิจัยจากเอกสาร บทความต่าง ๆ ที่เกี่ยวข้องกับการทำงานวิจัย
2. กำหนดหัวข้อ เป้าหมาย วัตถุประสงค์ และขอบเขตการทำงานวิจัย
3. ออกแบบวิเคราะห์อัลกอริทึมที่ใช้ติดตามผู้เล่น และ อัลกอริทึมที่ใช้ในการตรวจหาลูก เทนนิส
4. พัฒนาโปรแกรมจากอัลกอริทึมที่ได้ออกแบบ
5. เตรียมข้อมูลเพื่อทำการทดสอบการทำงานของอัลกอริทึม
6. ทำการทดลอง พร้อมทั้งปรับปรุงประสิทธิภาพ
7. สรุปผลการทดลองและเขียนวิทยานิพนธ์

1.7 โครงสร้างวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้จึงแบ่งเป็นส่วนที่สำคัญดังนี้

บทที่ 1 บทนำ กล่าวถึงความเป็นมาและความสำคัญของปัญหา ความมุ่งหมายและ วัตถุประสงค์ของการศึกษา สมมุติฐานของการศึกษา ขอบเขตของการดำเนินงานวิจัย ขั้นตอนการ ดำเนินงานวิจัย และประโยชน์ที่คาดว่าจะได้รับ และ โครงสร้างวิทยานิพนธ์

บทที่ 2 ทฤษฎีที่เกี่ยวข้องในงานวิจัยและงานวิจัยที่เกี่ยวข้อง อธิบายถึงทฤษฎีพื้นฐานและ งานวิจัยที่ผ่านมาในอดีตถึงปัจจุบัน

บทที่ 3 ออกแบบอัลกอริทึมที่ใช้ในการจำแนกท่าตีเทนนิส

บทที่ 4 วิธีการทดลองและผลการทดลองในการบรรยายการตีเทนนิส

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะในการพัฒนางานวิจัยขั้นต่อไป

เอกสารนี้สงวนลิขสิทธิ์โดยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีพื้นฐาน และงานวิจัยที่เกี่ยวข้องกับการจำแนกท่าดีกีฬาเทนนิสจากวีดิทัศน์

ในบทนี้จะกล่าวถึงงานวิจัยต่างๆที่เกี่ยวข้องกับการจำแนกท่าดีกีฬาเทนนิสจากวีดิทัศน์ รวมถึงทฤษฎีพื้นฐานต่างๆ

2.1 งานวิจัยที่เกี่ยวข้องกับการจำแนกท่าดีเทนนิสแบบอัตโนมัติ

ผลงานวิจัยของ Guangyu Zhu, Changsheng Xu, Qingming Huang and Wen Gao [1] เรื่อง "Action Recognition in Broadcast Tennis Video" งานวิจัยนี้ได้นำเสนอ การรู้จำของวีดิทัศน์ ถ่ายทอดสดการแข่งขันเทนนิส โดยเริ่มจากการติดตามผู้เล่นเทนนิสโดยใช้วิธี Support vector regression particle filter จากนั้นทำการบรรยายถึงการเคลื่อนไหว ซึ่งเริ่มจากการกำจัดสิ่งรบกวนรอบๆบริเวณผู้เล่นโดยใช้การปรับเรียบของเกาส์ (Gaussian smoothing) จากนั้นทำออปติคอลโฟว์แบบแบ่งส่วน (Slice-based optical flow histogram) ในบริเวณของผู้เล่นเทนนิสที่ได้จากการติดตาม สุดท้ายทำการจำแนกท่าดีเทนนิสด้วยวิธี Support Vector Machine (SVM) และได้ผลลัพธ์ของการตีเทนนิสแบบโฟร์แฮนด์ และ แบคแฮนด์ อย่างไรก็ตามงานวิจัยชิ้นนี้มีข้อจำกัดคือ สามารถทำการจำแนกท่าดีได้แค่เฉพาะผู้เล่นเทนนิส ที่อยู่ใกล้กับกล้อง เพราะว่าจำนวนจุดภาพของผู้เล่นที่อยู่ไกลจากกล้อง มีจำนวนที่น้อย จึงทำให้ขั้นตอนของการทำด้วยออปติคอลโฟว์แบบแบ่งส่วนได้ผลลัพธ์ที่ไม่เพียงพอต่อการจำแนกท่าดีเทนนิส

ผลงานวิจัยของ Hitesh Shah, Prakash Chokalingam, Balamanohar Paluri, Nalin Pradeep, Balasubramanian Raman [2] เรื่อง "Automated Stroke Classification in Tennis" งานวิจัยนี้ได้แนะนำ การจำแนกท่าดีเทนนิสแบบอัตโนมัติ ในกรณีที่พื้นฉากหลังมีความซับซ้อน โดยเริ่มจากการทำออปติคอลโฟว์ของ M.J. Black, P. Anandan [3] เพื่อใช้ในการตัดแยกผู้เล่นเทนนิสออกจากฉากหลัง จากนั้นทำการสร้างฮิสโทแกรมสีเพื่อเป็นแบบจำลอง การที่ต้องใช้ออปติคอลโฟว์เพื่อเป็นข้อมูลในการคัดแยก เพราะว่าบริเวณใดที่มีวัตถุเคลื่อนที่จะทำให้มีจำนวนของออปติคอลโฟว์สูง ซึ่งโดยทั่วไปแล้วผู้เล่นเทนนิสมักจะมีการเคลื่อนที่มากที่สุดในบริเวณของสนามเทนนิส และจากการศึกษาได้พบว่าถ้าใช้วิธีทางการเข้าคู่แม่แบบ (template matching) และ วิธีทางการเข้าคู่คอนทัวร์ (contour matching) จะไม่สามารถเปรียบเทียบความเหมือนกับโมเดลได้โดยตรง และยังจะเกิดปัญหาในกรณีที่ผู้เล่นเทนนิสไม่ใช่คนในแม่แบบ จึงต้องใช้เทคนิค Yatharth Saraf, R. Balasubramanian, K. Swaminathan [4] ของในบริเวณผู้เล่นเทนนิส เพื่อจะทำการปรับเปลี่ยน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปทรงของผู้เล่นเทนนิส จากภาพปกติไปเป็นภาพแบบโครงกระดูก เพราะว่ามีความยืดหยุ่นในเรื่องของการนำไปใช้ จากนั้นใช้รูปทรงของผู้เล่นเทนนิสมาใช้ในการรู้จำด้วย SVM ผลลัพธ์คือการจำแนกการตีแบบโฟร์แฮนด์ แบคแฮนด์ หรือไม่ใช่ทั้งสองประเภท

ผลงานวิจัยของ M. Petkovic, W. Jonker, and Z. Zivkovic [5] เรื่อง "Recognizing strokes in tennis videos using Hidden Markov Models" ได้นำเสนอการพัฒนาขั้นตอนวิธีในการบรรยายการแข่งขันเทนนิส จากวิดิทัศน์แบบอัตโนมัติ โดยทำการตัดแยกผู้เล่นออกจากฉากหลัง โดยใช้คุณสมบัติสีและการเทียบแบบสามมิติ จากนั้นเรียนรู้และจำแนกเหตุการณ์สำคัญในการเล่นด้วยแบบจำลองฮิดเดนมาร์คอฟ โดยดึงลักษณะเด่นจากโครงร่างของผู้เล่นที่ตัดแยกได้ ผลลัพธ์สุดท้ายคือการจำแนกเหตุการณ์ โดยแบ่งออกเป็น โฟร์แฮนด์ แบคแฮนด์ เซิร์ฟ วอลเลย์โฟร์แฮนด์ วอลเลย์แบคแฮนด์

จากงานวิจัยที่ [1,2,5] เป็นการนำข้อมูลเชิงพื้นที่หรือตำแหน่งของผู้เล่นในการวิเคราะห์ แต่มีบางงานวิจัย [6,7,9,10] ที่ใช้ข้อมูลของ ลูกเทนนิส สนามเทนนิส หรือข้อมูลเสียงเพื่อช่วยในการจำแนกท่าตีเทนนิส

ผลงานวิจัยของ F.Yan, W.Christmas and J.Kittler [6] เรื่อง "A Tennis Ball Tracking Algorithm for Automatic Annotation of Tennis Match " งานวิจัยนี้ได้นำเสนอ อัลกอริทึมการติดตามลูกเทนนิส เพื่อใช้ในการบรรยายการแข่งขันเทนนิสในวิดิทัศน์คุณภาพต่ำโดยใช้กล้องเพียง 1 ตัว โดยเริ่มจากการติดตามผู้เล่นเทนนิสโดยใช้วิธีการลบฉากหลัง (background subtraction) และติดตามบอลที่ได้จากการลบฉากหลัง ผลลัพธ์ที่ได้คือตำแหน่งของผู้เล่นเทนนิส จากนั้นหาตำแหน่งของลูกเทนนิสด้วยวิธี multi-cue foreground blob classification โดยดูความเข้มของจุดภาพที่ซึ่งในตำแหน่งที่เป็นลูกเทนนิสนั้น โดยส่วนมากจะมีความเข้มของจุดภาพที่สูง จากนั้นหาขอบด้วยวิธีโซเบลของในบริเวณที่มีความเข้มของจุดภาพสูง นำขอบที่ได้มาหาทิศทางความลาดชัน สุดท้ายใช้ Particle Filter เพื่อทำการติดตามตำแหน่งของลูกเทนนิส

ผลงานวิจัยของ Xinguo Yu, Chern-Horng Sim, Jenny Ran Wang and Loong Fah Cheong [7] เรื่อง "A trajectory-based ball detection and tracking algorithm in broadcast tennis video" งานวิจัยนี้ได้นำเสนอ การใช้ข้อมูลเสียงร่วมกับการค้นหาลูกบอลจากการพล็อตกราฟที่มีลักษณะวิถีโค้งของในวิดิทัศน์ถ่ายทอดสดการแข่งขันเทนนิส โดยเริ่มจากการตรวจหาและติดตามผู้เล่นเทนนิส ด้วยการลบเส้นของส่วนที่เป็นสนามเทนนิส จากนั้นใช้การขยายกลุ่มจุดภาพ (pixel growing procedure) เพื่อแยกวัตถุในแต่ละเฟรมออกโดยมีเงื่อนไขที่ว่า วัตถุที่มีขนาดใหญ่ 2 บอลคือผู้เล่นเทนนิสที่ใกล้และไกลจากกล้อง และมีเงื่อนไขเพิ่มเติม ได้แก่ เนตต้องคั่นระหว่างผู้เล่น 2 ฝ่าย ผู้เล่นที่อยู่ข้างล่าง (ใกล้กล้อง) คือผู้เล่นใกล้ ส่วนผู้เล่นที่อยู่ข้างบนคือผู้เล่นไกล (ไกลกล้อง) และในแต่ละฝั่งของเน็ตจะมีผู้เล่นได้แค่เพียง 1 คน จากนั้นหาตำแหน่งของลูกบอล โดยการกรองด้วยขนาดของจุดภาพ สีลูกบอล และกรองเส้นออก ผลลัพธ์ที่ได้คือตำแหน่งที่เป็นไปได้ว่าคือลูกบอล จากนั้นทำการหา

ช่วงเวลาที่มีการตีลูกบอลโดยใช้เสียง ในเวลาที่ลูกกระทบไม้ใช้วิธีของ M. Xu, N. C. Maddage, C. Xu, M. Rankanhalli, Q. Tian [8] นำตำแหน่งลูกบอลและช่วงเวลาที่มีการตีลูกเทนนิส มาสร้างกราฟ สุดท้ายใช้ kalman filter เพื่อกรองหาวิถีของลูกบอลที่แท้จริงซึ่งจะได้กราฟเป็นวิถีโค้ง ที่ซึ่งในงานวิจัยนี้ไม่มีการจำแนกท่าตีเทนนิส

ผลงานวิจัยของ E. Kijak, G. Gravier, P. Gros, L.Oisel, and F. Bimbot [9] เรื่อง “HMM based structuring of tennis videos using visual and audio cues” งานวิจัยนี้ได้นำเสนอวิธีการดึงลักษณะเด่นทั้งภาพและเสียงที่ได้จากวิดีโอและรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟ ระบบนี้ได้ทำการจำแนกเหตุการณ์ออกเป็น 4 ประเภทคือ การเสิร์ฟพลาต การตีลูกกลับ รีเพลย์ และหยุด

ผลงานวิจัยของ C.O. Conaire, P. Kelly, D. Connaghan, and N.E. O’Connor [10] เรื่อง “TENNISENSE: A platform for extracting semantic information from multi-camera tennis data” งานวิจัยนี้ได้นำเสนอระบบ TennisSense โดยใช้วิดีโอจากกล้องที่อยู่เหนือสนามเทนนิส และมีจำนวนกล้องทั้งหมด 9 ตัว โดยเริ่มจากการติดตามตำแหน่งของลูกเทนนิส โดยใช้วิธีการตรวจจับบริเวณภาพที่มีการเคลื่อนที่ แล้วทำการเชื่อมตำแหน่งที่น่าจะเป็นลูกเทนนิสให้มีลักษณะแบบเชิงเส้น เพราะว่าลูกเทนนิสจากงานวิจัยนี้ใช้มุมกล้องแบบเหนือสนาม จึงทำให้ลูกเทนนิสมีการเคลื่อนที่คล้ายกับลักษณะกราฟแบบเชิงเส้น จากนั้นทำการรู้จำประเภทการตีโดยพิจารณาจากการเคลื่อนที่ของตำแหน่งลูกเทนนิส

จากงานวิจัยที่ได้กล่าวถึงนั้น ยังมีปัญหาในเรื่องของการจำแนกท่าตีเทนนิส เช่น [1,2,5] จำเป็นที่จะต้องทราบถึงรายละเอียดในตำแหน่งของผู้เล่นอย่างเพียงพอ กล่าวคือ แขนหรือมือที่จับไม้ต้องมีความชัดเจน ซึ่งในความเป็นจริงแล้วขนาดของผู้เล่นเทนนิสที่อยู่ฝั่งไกลนั้นมีขนาดเล็ก ประมาณ 35 พิกเซล อีกทั้งวิดีโอยังมีสัญญาณรบกวนสูง จึงเป็นการยากที่จะใช้รายละเอียดของผู้เล่น เพื่อทำการจำแนก จากงานวิจัยที่ [6,7,9,10] จำเป็นจะต้องใช้ข้อมูลหลายองค์ประกอบเพื่อทำการจำแนกท่าตี ก็จะมีปัญหาในเรื่องของการที่จะได้มาซึ่ง ตำแหน่ง (ผู้เล่น, ลูกเทนนิส, สนามเทนนิส) กล่าวคือ อย่างกรณีที่ผู้เล่นเทนนิสฝั่งไกลมีการเคลื่อนที่ใกล้กับกรรมการหลังเส้นทำให้ผลจากการติดตามผิดพลาดได้ หรือการหาเส้นสนามเทนนิสที่ใช้วิธีของ Hough transform ที่เป็นการคำนวณซับซ้อน อีกทั้งในงานวิจัยที่ [10] จำเป็นที่จะต้องใช้กล้องมากกว่า 1 ตัวเพื่อจะทำการจำแนกท่าตี

จากปัญหาที่ได้กล่าวไปนั้น ในงานวิจัยนี้จึงใช้วิธีการจำแนกท่าตีเทนนิสจากวิดีโอโดยใช้วิธีการเคลื่อนที่สัมพันธ์ของผู้เล่นและลูกบอล โดยใช้กล้องเพียง 1 ตัว ซึ่งเป็นการดูแลวิถีระหว่างผู้เล่นและลูกบอล โดยที่สามารถจำแนกท่าตีผู้เล่นเทนนิสฝั่งใกล้และไกลได้ สามารถเพิ่มประสิทธิภาพของการติดตามผู้เล่นโดยใช้วิธีการติดตามโดยวิธีย้ายค่าเฉลี่ย อีกทั้งได้เพิ่มประสิทธิภาพการหาเส้นสนามเทนนิส ด้วยวิธีการพิจารณาความเข้มของจุดภาพผนวกกับวิธีถดถอยเชิงเส้น ทำการหาตำแหน่งของลูกเทนนิสโดยใช้วิธีการหาความแตกต่างระหว่างเฟรม ซึ่งทฤษฎีพื้นฐานเหล่านี้ได้กล่าวไว้ในหัวข้อถัดไป

2.2 วิธีย้ายค่าเฉลี่ย (Mean shift)

ขั้นตอนวิธีย้ายค่าเฉลี่ย เป็นวิธีการที่ใช้ฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) ซึ่งมีทฤษฎีดังนี้

ให้ X เป็นตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable) และมีฟังก์ชันความหนาแน่นของความน่าจะเป็น $f(x)$ เราสามารถหาความน่าจะเป็นได้จาก

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad \text{เมื่อ } -\infty < a < b < \infty \quad (2.1)$$

เมื่อเราไม่ทราบรูปแบบของ $f(x)$ ว่าเป็นรูปแบบใด เราต้องทำการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็น การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็น จะเรียกว่า การประมาณความหนาแน่น (Density Estimation)

การประมาณความหนาแน่นจากตัวอย่างสุ่ม (Random Sample) มีทั้งที่เป็นแบบพารามตริก (Parametric) และ นอนพารามตริก (Non parametric) ซึ่งวิธีย้ายค่าเฉลี่ย เป็นการประมาณความหนาแน่นแบบนอนพารามตริก (Non parametric Density Estimation)

วิธีการประมาณความหนาแน่นแบบพารามตริก นั้นต้องเป็นไปตามเงื่อนไขดังนี้

1. กลุ่มตัวอย่างที่ได้จากการสุ่มต้องมีการแจกแจงปกติ
2. ข้อมูลต้องเป็นชนิดอินเทอร์วอล (Interval) หรือเรโซ (Ratio)

ดังนั้นการประมาณความหนาแน่นแบบนอนพารามตริก จะใช้การประมาณฟังก์ชันความหนาแน่นจากชุดข้อมูล โดยตรง มีหลายวิธีเช่นการประมาณความหนาแน่นโดย ฮิสโตแกรม (Histogram) หรือตัวประมาณความหนาแน่นแบบง่าย (The native estimator) เป็นต้น ส่วนวิธีย้ายค่าเฉลี่ยใช้วิธีการประมาณความหนาแน่นแบบเคอร์เนล (Kernel Density Estimation)

การประมาณความหนาแน่นแบบเคอร์เนล

เคอร์เนล [11][12] คือ ฟังก์ชันของค่าถ่วงน้ำหนัก ที่ถูกใช้ในการประมาณความหนาแน่นแบบนอนพารามตริก เป็นวิธีการที่นิยมทำมากที่สุด เป็นวิธีการที่สังเกตแต่ละค่าจะมีน้ำหนักถ่วง เสนอเป็นครั้งแรกโดย Rosenblatt ในปี ค.ศ. 1956

กำหนดให้ $\hat{f}_h(x)$ คือ ตัวประมาณความหนาแน่นแบบเคอร์เนล

$$\hat{f}_h(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (2.2)$$

โดยที่ x_i คือ ตัวแปรสุ่ม ซึ่ง $i = 1, 2, \dots, n$

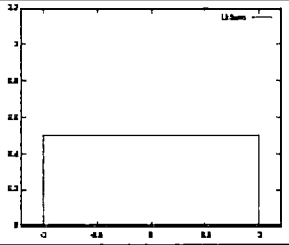
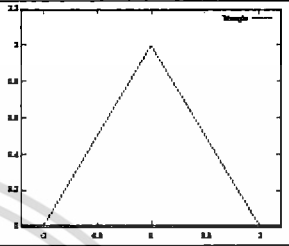
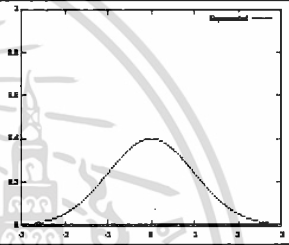
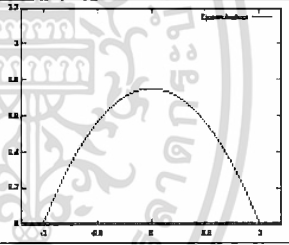
x คือ ตัวแปรที่ไม่ทราบค่า

K คือ ฟังก์ชันเคอร์เนล

h คือ พารามิเตอร์ปรับเรียบหรือแบนด์วิดท์ (Bandwidth)

d คือ จำนวนของมิติ

ตาราง 2.1 ตัวอย่างรูปแบบฟังก์ชันเคอร์เนลแบบต่างๆ จาก [13]

Kernel Functions		
Uniform	$\frac{1}{2}I_{\{ u \leq 1\}}$	
Triangular	$(1- u)I_{\{ u \leq 1\}}$	
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$	
Epanechnikov	$\frac{3}{4}(1-u^2)I_{\{ u \leq 1\}}$	

โดยที่ $u = \left(\frac{x-x_i}{h}\right)$, $I_{\{...\}}$ คือ ฟังก์ชันตัวชี้บอก (Indicator Function)

จากสมการที่ (2.2) เคอร์เนลที่ทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรวม (MISE) [14][15] มีค่าต่ำ ของในการเปรียบเทียบระหว่าง ตัวประมาณการกับตัวความหนาแน่นที่แท้จริง คือ เคอร์เนลของ Epanechnikov [16]

$$K_E(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-\|x\|^2) & \text{if } \|x\| \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

โดยที่ c_d คือ ปริมาตรของมิติทรงกลม

d คือ จำนวนของมิติ

ให้โพรไฟล์ (profile) ของเคอร์เนล K คือฟังก์ชันที่ $k: [0, \infty) \rightarrow R$ โดยที่

$$K(x) = k(\|x\|^2) \quad (2.4)$$

ความแตกต่างของเคอร์เนลฟังก์ชันและโพรไฟล์ฟังก์ชัน คือ เคอร์เนลฟังก์ชันได้ทำการนิยามในพจน์ของเวกเตอร์ ส่วนโพรไฟล์ฟังก์ชันได้ทำการนิยามในพจน์ของตัวแปร ดังนั้น โพรไฟล์ของ Epanechnikov คือ

$$k_E(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-x) & \text{if } x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.5)$$

จากสมการที่ (2.2) สามารถเขียน ตัวประมาณความหนาแน่นในรูปของโพรไฟล์ได้

$$\hat{f}_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (2.6)$$

นิยามฟังก์ชัน

$$g(x) = -k'(x) \quad (2.7)$$

จากสมการที่ (2.7) อนุพันธ์ของ k จะเป็นจริงก็ต่อเมื่อมีค่า $x \in [0, \infty)$ ยกเว้นในเซตจำกัด จากนั้นนิยามเคอร์เนล G ได้

$$G(x) = c_{g,d} g(\|x\|^2) \quad (2.8)$$

โดยที่ $c_{g,d}$ คือ ค่าคงที่นอร์มอลไลซ์

จากสมการ (2.6) ทำการใส่เกรเดียนต์เพื่อ ทำการประมาณความหนาแน่น

$$\begin{aligned} \hat{\nabla} f_{h,K}(x) &\equiv \nabla \hat{f}_{h,K}(x) = \frac{2}{nh^{d+2}} \sum_{i=1}^n (x-x_i) k'\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \\ &= \frac{2}{nh^{d+2}} \sum_{i=1}^n (x_i-x) g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \\ &= \frac{2}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \right] \end{aligned} \quad (2.9)$$

ในสมการที่ (2.9) กำหนดให้ $\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)$ ต้องมีค่าที่ไม่ใช่ศูนย์ จาก Yizong Cheng [17]

จึงทำให้ทราบว่าการทำอนุพันธ์ของโพรไฟล์ Epanechnikov จะทำให้ได้ โพรไฟล์เอกรูป (uniform)

ในวงเล็บหลังของสมการ (2.9) แสดงถึงเวกเตอร์ย้ายค่าเฉลี่ย (mean shift vector) นำมาเขียนใหม่ได้ดังนี้

$$m_{h,G}(x) = \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \right] \quad (2.10)$$

เทอมแรกของในสมการที่ (2.10) เป็นการหาค่าตำแหน่งใหม่ของเวกเตอร์ย้ายค่าเฉลี่ย (Mean shift vector) ซึ่งเราจะนำไปใช้ใน ส่วนของการติดตามโดยใช้วิธีย้ายค่าเฉลี่ยในขั้นตอนของการทำนายตำแหน่งใหม่ในหัวข้อที่ 2.3.4

2.3 การติดตามโดยใช้วิธีย้ายค่าเฉลี่ย (Mean shift Tracking)

การติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย [18][19] เริ่มจากการกำหนดบริเวณของสิ่งที่ต้องการติดตาม เพื่อเป็นการสร้างแบบจำลอง หลังจากนั้นในรอบถัดๆ ไป ทำการค้นหาบริเวณที่ใกล้เคียงและบริเวณเดิม เพื่อเปรียบเทียบว่ามีความเหมือน กับแบบจำลองมากน้อยเพียงไร โดยใช้ค่าสัมประสิทธิ์วัดความเหมือนในการเปรียบเทียบ

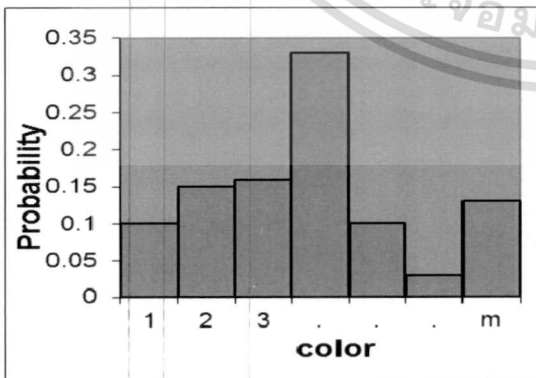
2.3.1 การแทนตำแหน่งเป้าหมาย (Target Representation)

เพื่อที่จะทำการแสดงลักษณะพิเศษของเป้าหมาย อันดับแรกต้องทำการเลือก ปริภูมิลักษณะเด่น (Feature Space) ซึ่งเราสามารถใช่ PDF ของสีเป็นลักษณะเด่นได้ กำหนดให้ที่ตำแหน่งอ้างอิง \hat{q} เป็นฮิสโทแกรมสีของแบบจำลองเป้าหมาย (Target Model) มีตำแหน่งศูนย์กลางอยู่ที่ 0 จากนั้นในเฟรมถัดมากำหนดให้ $\hat{p}(y)$ เป็นฮิสโทแกรมสีของแคนดิเดตเป้าหมาย (Target Candidate) อยู่ที่ตำแหน่ง y จะได้

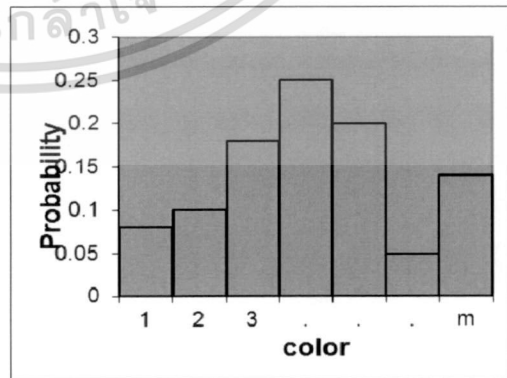
$$\text{แบบจำลองเป้าหมาย : } \hat{q} = \{\hat{q}_u\}_{u=1..m} \quad \text{โดยที่ } \sum_{u=1}^m \hat{q}_u = 1 \quad (2.11)$$

$$\text{แคนดิเดตเป้าหมาย : } \hat{p}(y) = \{\hat{p}_u(y)\}_{u=1..m} \quad \text{โดยที่ } \sum_{u=1}^m \hat{p}_u = 1 \quad (2.12)$$

โดยที่ m คือ จำนวนถังสี (bin) ของ ฮิสโทแกรม
 \hat{q}_u คือ ความน่าจะเป็นของลักษณะเด่นที่ของถังสีที่ u (u^{th} bin)



(ก)



(ข)

รูปที่ 2.1 (ก) แบบจำลองเป้าหมาย (\hat{q}) (ข) แคนดิเดตเป้าหมาย ($\hat{p}(y)$)

รูปที่ 2.1 (ก) (ข) แสดงถึงสมการที่ (2.11) และ (2.12) ว่าใช้ฮิสโทแกรมเพื่อประมาณความหนาแน่นแบบนอนพาราเมตริก

$$\hat{\rho}(y) \equiv \rho[\hat{p}(y), \hat{q}] \quad (2.13)$$

สมการที่ (2.13) คือฟังก์ชันที่ใช้เปรียบเทียบความเหมือนกันระหว่างแบบจำลองเป้าหมายและแคนดิเดตเป้าหมาย หรือสามารถอธิบายเพิ่มได้ว่าเป็นการเปรียบเทียบความเหมือนกันระหว่างฮิสโทแกรมในรูปที่ 2.1 (ก) กับรูปที่ 2.1 (ข)

2.3.2 แบบจำลองเป้าหมาย (Target Model)

กำหนดให้บริเวณของแบบจำลองเป้าหมายแสดงเป็นพื้นที่แบบวงรี ที่มีจุดศูนย์กลางอยู่ที่ 0 จากนั้นทำการ นอร์มอลไลซ์ขนาดของวงรีให้อยู่ในรูปแบบของวงกลมหนึ่งหน่วย (unit circle) ซึ่งจะเป็นการขจัดปัญหาในเรื่องของความแตกต่างในแต่ละขนาดที่นำมาใช้

ให้ $\{x_i\}_{i=1, \dots, n}$ เป็น ตำแหน่งของจุดภาพนอร์มอลไลซ์ ในพื้นที่วงรี ของแบบจำลองเป้าหมาย ที่ได้กำหนดไว้ $k(x)$ คือ เคอร์เนล ซึ่งเอาไว้ใช้ในการกำหนดค่าน้ำหนัก ของความใกล้เคียงหรือไกลจากจุดศูนย์กลาง จะมีค่าสูงก็ต่อเมื่ออยู่ใกล้กับจุดศูนย์กลางและมีค่าต่ำเมื่ออยู่ไกล

ดังนั้นสามารถคำนวณหาแบบจำลองเป้าหมายได้ดังนี้

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i\|^2) \delta_{ib} \quad (2.14)$$

โดยที่ n คือ จำนวนของจุดภาพทั้งหมดที่อยู่ในพื้นที่ของโมเดลเป้าหมาย

b คือ จำนวนแบนด์วิธของฮิสโทแกรม

δ คือ Kronecker delta function จะมีค่าเป็น 1 ถ้าค่าของ x_i มีค่าอยู่ในจำนวนถึงสี่ นอกนั้นมีค่า เป็น 0

C คือ ค่าที่ได้จากนอร์มอลไลซ์ คำนวณได้จาก

$$C = \frac{1}{\sum_{i=1}^n k(\|x_i\|^2)} \quad (2.15)$$

2.3.3 แคนดิเดตเป้าหมาย (Target Candidates)

ให้ $\{x_i\}_{i=1, \dots, n}$ เป็น ตำแหน่งของจุดภาพนอร์มอลไลซ์ ในพื้นที่วงรีของแคนดิเดตเป้าหมาย แต่มีจุดศูนย์กลางอยู่ที่ y และมีขนาดแบนด์วิธ h ใช้เคอร์เนลตัวเดียวกับแบบจำลองเป้าหมาย $k(x)$

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta_{ib} \quad (2.16)$$

ที่ซึ่ง

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right)} \quad (2.17)$$

ความแตกต่างของสมการ (2.14) และ (2.16) คือ ในสมการที่ (2.14) ทำที่จุดศูนย์กลางอยู่ที่ 0 แต่ใน สมการ (2.16) จุดศูนย์กลางจะอยู่ที่ตำแหน่ง y

2.3.4 หาค่าน้ำหนักและตำแหน่งใหม่

วิธีการติดตามโดยวิธีย้ายค่าเฉลี่ย ใช้ค่าสัมประสิทธิ์ Bhattacharyya [20] เป็นตัววัดความเหมือนกันระหว่าง ชุดของข้อมูลที่มีการกระจายตัว 2 ชุด ซึ่งมีข้อดีคือ สามารถเปรียบเทียบในกรณีที่มีจำนวนหรือขนาดของชุดข้อมูลมีขนาดไม่เท่ากัน ทำให้วิธีการนี้รองรับการเปรียบเทียบวัตถุที่มีการเปลี่ยนแปลงรูปร่างและขนาดได้

กำหนดให้ $\hat{\rho}(y)$ คือฟังก์ชันการเปรียบเทียบของตำแหน่งที่ตั้ง y ระหว่าง ฮิสโทแกรมของแบบจำลองเป้าหมายกับแคนดิเดทเป้าหมาย ด้วยวิธีค่าสัมประสิทธิ์ Bhattacharyya

$$\hat{\rho}(y) = \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (2.18)$$

จากสมการที่ (2.18) ถ้าข้อมูลใดๆ 2 ชุดที่นำมาเปรียบเทียบกันนั้น มีความเหมือนกันมากก็จะทำให้ $\hat{\rho}(y)$ มีค่าที่สูงแต่ถ้ามีความเหมือนน้อยก็จะได้ $\hat{\rho}(y)$ มีค่าต่ำ

ในการหาตำแหน่งใหม่นั้น เริ่มต้นจากการใช้ตำแหน่ง \hat{y}_0 ที่ซึ่งได้มาจากเฟรมก่อนหน้านี้ จากนั้นในเฟรมปัจจุบัน ทำการหาค่าของ $\{\hat{p}_u(\hat{y}_0)\}_{u=1..m}$ ที่เป็นแคนดิเดทเป้าหมายของตำแหน่ง \hat{y}_0 ทำการกระจายอนุกรมเทย์เลอร์ รอบจุด $\hat{p}_u(\hat{y}_0)$ เพื่อเป็นการประมาณแบบเชิงเส้น (Linear Approximation) ของค่าสัมประสิทธิ์ Bhattacharyya ในสมการ (2.18) จะได้

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} + \frac{1}{2} \sum_{u=1}^m \hat{p}_u(y) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \quad (2.19)$$

เงื่อนไขที่สำคัญของการติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย คือแคนดิเดทเป้าหมายของเฟรมถัดไป $\{\hat{p}_u(y)\}_{u=1..m}$ ต้องไม่มีการเปลี่ยนแปลงมากจนเกินไปจากในแคนดิเดทเป้าหมายของเฟรมปัจจุบัน $\{\hat{p}_u(\hat{y}_0)\}_{u=1..m}$ และเงื่อนไขอีกข้อคือ $\hat{p}_u(\hat{y}_0) > 0$

ในสมการที่ (2.19) แทนค่า $\hat{p}_u(y)$ ของสมการที่ (2.16) และจัดรูปสมการใหม่ จะได้

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (2.20)$$

ที่ซึ่ง

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \delta_{ib} \quad (2.21)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สังเกตในเทอมแรกของสมการที่ (2.20) จะไม่มี y อยู่ ดังนั้น y จึงเป็นอิสระในเทอมแรก ส่วนในเทอมที่ 2 เป็นการประมาณความหนาแน่นด้วยเคอร์เนล $k(x)$ ที่ตำแหน่ง y ในเฟรมปัจจุบัน โดยมี w_i คือค่าถ่วงน้ำหนัก

$$\frac{C_h}{2} \sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \tag{2.22}$$

ในสมการที่ (2.22) จะแสดงถึง การแสวงหาค่าสูงสุด (Sought Maximum) ซึ่งมีรูปแบบตัวประมาณความหนาแน่นแบบเคอร์เนล ดังนั้นในการหาค่าตำแหน่งใหม่สามารถทำได้จากกระบวนการวิธีย้ายค่าเฉลี่ยในหัวข้อที่ 2.2 จะได้ผลลัพธ์ดังนี้

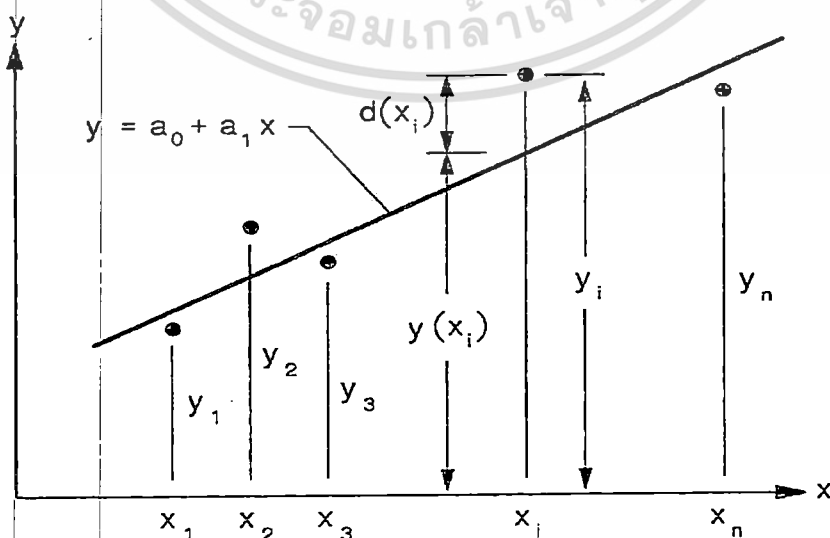
$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)} \tag{2.23}$$

จากที่ได้กล่าวข้างต้นแล้วว่า การหาอนุพันธ์ของโพรไฟต์ Epanechnikov จะทำให้ได้ โพรไฟต์เอกกรุป (uniform) จึงทำให้ค่าเคอร์เนล $g(x) = -k'(x) =$ ค่าคงที่ ดังนั้นสมการที่ (2.23) สามารถตัดค่าเคอร์เนล $g(x)$ ได้ จึงเขียนในรูปใหม่ที่ง่ายขึ้นได้ดังนี้

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \tag{2.24}$$

2.4 การถดถอยเชิงเส้น (Linear Regression) [21]

เป็นระเบียบวิธีในการสร้างฟังก์ชันเส้นตรงที่เหมาะสม (Best fit) สำหรับในแต่ละชุดข้อมูล



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ โดยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
รูปที่ 2.2 กราฟเพื่อจะประมาณการถดถอยเชิงเส้น
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กำหนดให้ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ เป็นชุดของข้อมูลที่เราต้องการสร้างฟังก์ชันเส้นตรงที่เหมาะสม

$$y = a_0 + a_1x \quad (2.25)$$

โดยที่ a_0 และ a_1 คือค่าคงที่เราต้องหาเพื่อที่จะสร้างสมการเส้นตรงที่มีค่าความผิดพลาด โดยเฉลี่ยที่น้อยที่สุดจากทุกข้อมูลที่กำหนด ด้วยการใช้วิธีกำลังสองน้อยที่สุด (Least square)

พิจารณาที่ตำแหน่ง x_i ของข้อมูลใดๆ จะได้ค่าคือ y_i แต่เมื่อนำค่า x_i แทนในสมการ $y = a_0 + a_1x$ จะได้ $y(x_i)$ ดังนั้นค่าของ $y(x_i)$ ที่ได้จากสมการ จะมีค่าที่แตกต่างไปจาก y_i เท่ากับ $d(x_i)$

ดังนั้นค่าความผิดพลาด E ทั้งหมด ที่เกิดจากชุดข้อมูล n ข้อมูล เมื่ออยู่ในรูปกำลังสองของค่าแตกต่าง $d(x_i)$ จะได้

$$E = \sum_{i=1}^n d(x_i)^2 \quad (2.26)$$

เพราะว่า $d(x_i) = y_i - y(x_i)$ แทนค่าลงในสมการที่ (2.26) จะได้ว่า

$$E = \sum_{i=1}^n [y_i - y(x_i)]^2 \quad (2.27)$$

แทนค่า $y(x_i)$ จากสมการ (2.25) ในสมการ (2.27) จะได้ว่า

$$E = \sum_{i=1}^n [y_i - (a_0 + a_1x_i)]^2 \quad (2.28)$$

จากสมการที่ (2.28) จะเห็นได้ว่ามีตัวแปรค่า 2 ตัวคือ a_0 และ a_1 ซึ่งเราสามารถหาค่าของ a_0 และ a_1 ด้วยวิธีกำลังสองน้อยที่สุด โดยการหาค่าต่ำสุดของค่าความผิดพลาดจะได้ $\frac{\partial E}{\partial a_0} = 0$; และ

$$\frac{\partial E}{\partial a_1} = 0;$$

$$\frac{\partial E}{\partial a_0} = \sum_{i=1}^n 2(y_i - a_0 - a_1x_i)(-1) = 0 \quad (2.29)$$

$$\frac{\partial E}{\partial a_1} = \sum_{i=1}^n 2(y_i - a_0 - a_1x_i)(-x_i) = 0 \quad (2.30)$$

จัดรูปสมการ (2.29) จะได้ว่า

$$\begin{aligned} \sum_{i=1}^n a_0 + \sum_{i=1}^n a_1x_i &= \sum_{i=1}^n y_i \\ na_0 + a_1 \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i \end{aligned} \quad (2.31)$$

จัดรูปสมการ (2.30) จะได้ว่า

$$\begin{aligned} \sum_{i=1}^n a_0x_i + \sum_{i=1}^n a_1x_i^2 &= \sum_{i=1}^n x_i y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i y_i \end{aligned} \quad (2.32)$$

แก้สมการ (2.31) และ สมการ (2.32) จะได้ว่า

$$\begin{aligned}
 a_0 &= \frac{\left(\sum_{i=1}^n y_i\right)\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i y_i\right)\left(\sum_{i=1}^n x_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2} \\
 &= \frac{\sum_{i=1}^n y_i}{n} - a_1 \frac{\sum_{i=1}^n x_i}{n} \\
 &= \bar{y} - a_1 \bar{x}
 \end{aligned} \tag{2.33}$$

$$a_1 = \frac{n\left(\sum_{i=1}^n x_i y_i\right) - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2} \tag{2.34}$$

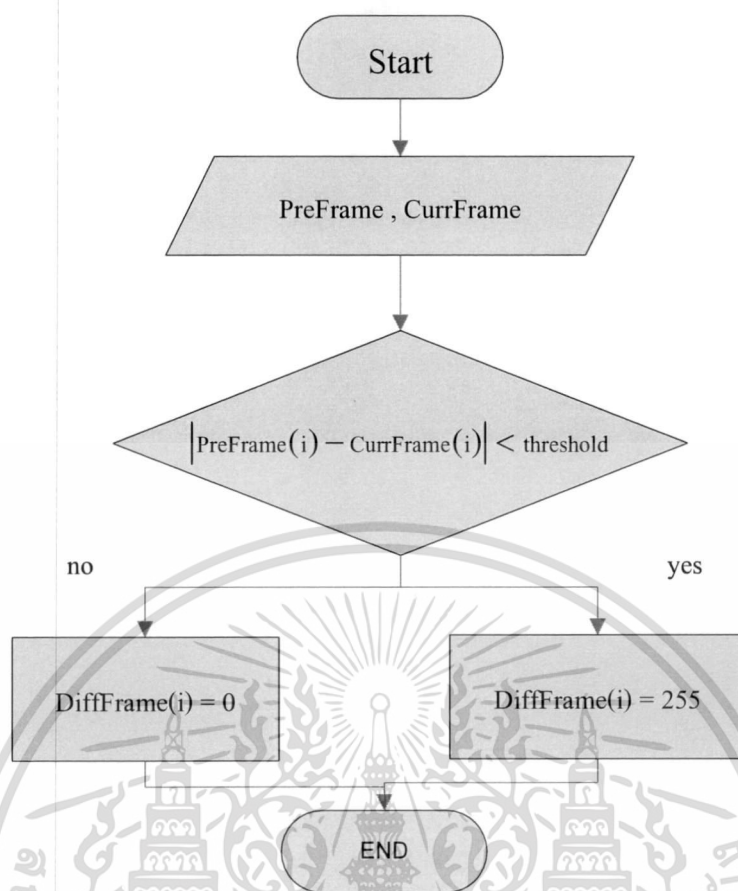
จากนั้นแทนค่า a_0 และ a_1 ลงใน $y = a_0 + a_1 x$ ก็จะได้สมการถดถอยแบบเชิงเส้นที่ต้องการ

2.5 การหาความแตกต่างระหว่างเฟรม (Frame Differencing) [22]

เป็นระเบียบวิธีการที่ทำโดย การนำเฟรมภาพหนึ่งกับอีกเฟรมหนึ่งมาทำการลบ ซึ่งส่วนที่มีความแตกต่างเป็นส่วนที่เราสนใจ สามารถแสดงขั้นตอนได้ดัง รูปที่ 2.3

สมมติให้ภาพระดับเทา (Gray level Image) ได้แก่ภาพ PreFrame CurrFrame และ DiffFrame ซึ่ง PreFrame คือภาพที่ได้จากเฟรมก่อนหน้า ส่วน CurrFrame คือภาพในเฟรมปัจจุบันและ DiffFrame คือภาพผลลัพธ์ที่ได้ จากการใช้วิธีการหาความแตกต่างระหว่างเฟรม

ระบบจะนำ PreFrame และ CurrFrame มาทำการลบ และ ดูความเข้มของจุดภาพ (Pixel intensity) ถ้าค่าที่ได้จากการลบ มากกว่าขีดเริ่มเปลี่ยน (threshold) ความเข้มของจุดภาพใน DiffFrame ตำแหน่งนั้น จะมีค่าเป็น 0 (จุดสีดำ) แต่ถ้าน้อยกว่า จะได้ค่าเป็น 255 (จุดสีขาว)



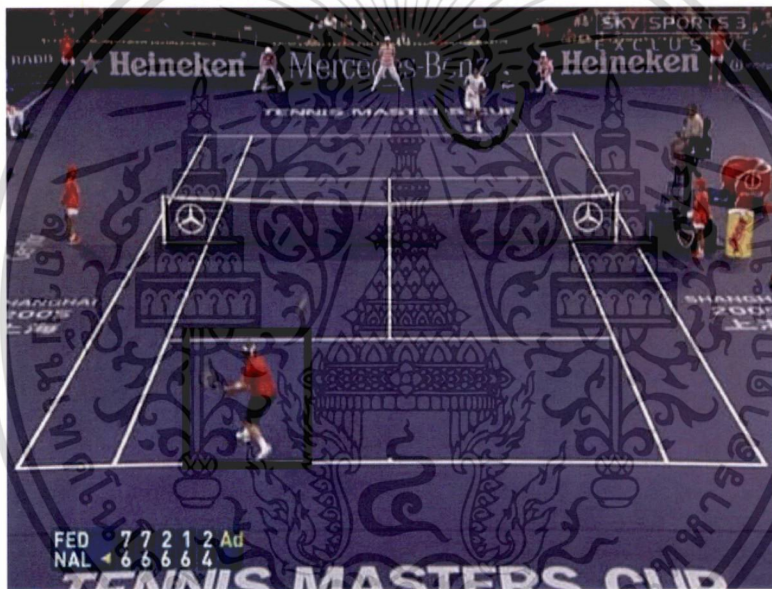
รูปที่ 2.3 ขั้นตอนวิธีความแตกต่างระหว่างเฟรม

บทที่ 3

การจำแนกทำตีกีฬาเทนนิสจากวิดีโอ

บทนี้กล่าวถึง งานวิจัยที่ผู้วิจัยนำเสนอ คือ วิธีการจำแนกทำตีกีฬาเทนนิส โดยใช้วิดีโอที่ได้นำมาจากอินเทอร์เน็ต เฉพาะที่ใช้มุกกล้องที่เป็นแบบระยะไกลเพียงอย่างเดียว

ดังรูปที่ 3.1 ซึ่งเป็นมุกกล้องที่ใช้ถ่ายทอดสดในระหว่างการแข่งขันเทนนิสโดยส่วนใหญ่จากรูปจะเห็นผู้เล่นเทนนิสซึ่งมีเน็ตคั่นกลางอยู่ และตำแหน่งผู้เล่นที่อยู่ติดกับมุกกล้อง (กรอบสี่เหลี่ยมสีดำ) เราจะเรียกว่า “ผู้เล่นใกล้” ส่วนผู้เล่นที่อยู่ไกลจากมุกกล้อง (วงรีสีดำ) เราจะเรียกว่า “ผู้เล่นไกล” โดยที่ในวิทยานิพนธ์นี้จะทำการจำแนกทำตีเทนนิสของผู้เล่นเทนนิสทั้งใกล้และไกล



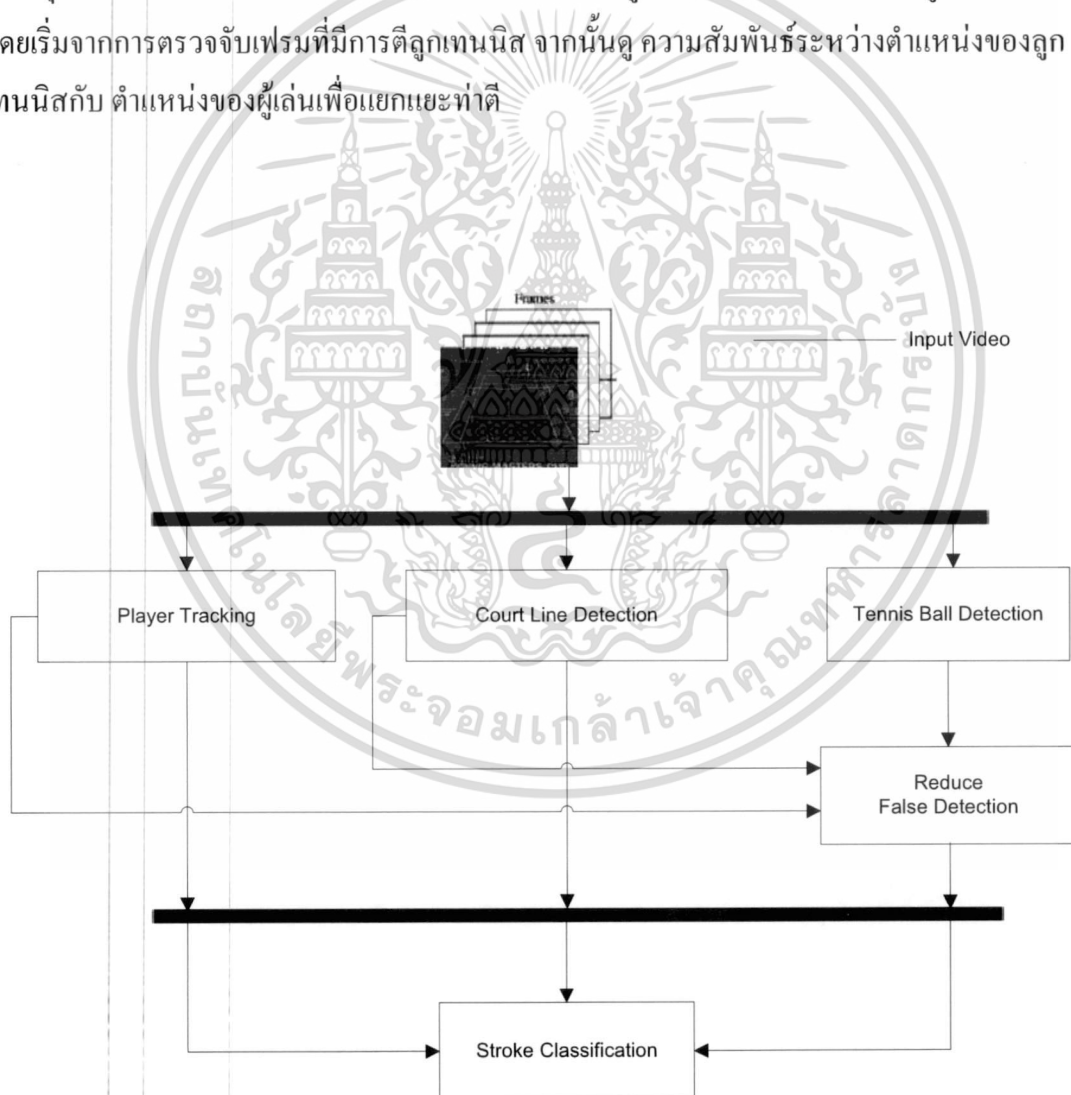
รูปที่ 3.1 มุกกล้องระยะไกล

ระบบการจำแนกทำตีเทนนิสแบบอัตโนมัติ ประกอบไปด้วยส่วนของขั้นตอนหลัก 5 ขั้นตอน คือ ส่วนแรกการติดตามผู้เล่นเทนนิส ของผู้เล่นที่อยู่ฝั่งใกล้และฝั่งไกล จากการสังเกตจะได้ว่าสิ่งที่ใช้เป็นตัวแบบเฉพาะของการติดตามผู้เล่นเทนนิส คือ สีของบริเวณผู้เล่นเทนนิสทั้งใกล้และใกล้ที่เกิดจากสีเสื้อและสีกางเกง ในวิทยานิพนธ์ฉบับนี้เราติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย ในการติดตามผู้เล่นเทนนิส ซึ่งจะได้ผลลัพธ์คือตำแหน่งของผู้เล่นเทนนิส ซึ่งเราจะนำไปใช้ในส่วนของการจำแนกทำตีเทนนิสต่อไป ส่วนที่สอง คือ การหาเส้นของสนามเทนนิส อย่างที่ทราบกันดีอยู่แล้วว่าสีของเส้นสนามเทนนิส คือสีขาวซึ่งมีความเข้มของจุดภาพสูง และมีลักษณะเส้น เป็นแบบเส้นตรง เราจะใช้ความรู้เหล่านี้ในการหาเส้นสนามเทนนิส ซึ่งผลจากส่วนนี้จะถูกนำไปใช้ในการลดปริมาณความ

ผิดพลาดของการตรวจหาลูกเทนนิส ส่วนที่สาม เป็นการหาตำแหน่งของลูกเทนนิส โดยใช้วิธีการเอกสารเป็นเอกสารที่ส่งงานไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกไปเผยแพร่ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หาความแตกต่างระหว่างเฟรม เพื่อหาบริเวณของจุดภาพที่มีการเคลื่อนไหว แต่เนื่องจากความเร็วในการเคลื่อนที่ของลูกเทนนิสมีความเร็วกว่าอัตราเฟรม (frame rate) ของกล้อง ทำให้ลูกเทนนิสไม่เป็นรูปทรงกลม ประกอบกับวิถีทัศน์ที่ใช้ นำมาจากอินเทอร์เน็ต ซึ่งมีที่มาหลากหลายและเป็นวิถีทัศน์ที่ถูกบีบอัดจึงทำให้ สีของลูกเทนนิสมีความกลมกลืนไปกับฉากหลัง ดังนั้น เราจึงไม่สามารถใช้สีและรูปร่างของลูกเทนนิสในการตรวจจับตำแหน่งของเทนนิสได้ เราจึงหาตำแหน่งลูกเทนนิสโดยพิจารณาจากการเคลื่อนที่และขนาดของลูกเทนนิส ส่วนที่สี่ คือ ส่วนที่ลดปริมาณความผิดพลาดของการตรวจหาลูกเทนนิส เนื่องจากผลลัพธ์จากการทำส่วนที่สาม อาจจะได้ตำแหน่งที่ไม่ใช่เทนนิสออกมาด้วย จากการทดลองพบว่าความผิดพลาดเหล่านั้นส่วนมากแล้ว คือ ส่วนที่เป็นเส้นสนาม และ ส่วนของผู้เล่นเทนนิส เราจึงทำการกำจัดตำแหน่งที่ตรงกับส่วนดังกล่าวออกไป และส่วนสุดท้าย คือ การจำแนกท่าตีเทนนิส ที่ใช้ตำแหน่งของผู้เล่นเทนนิส และตำแหน่งลูกเทนนิส โดยเริ่มจากการตรวจจับเฟรมที่มีการตีลูกเทนนิส จากนั้นดู ความสัมพันธ์ระหว่างตำแหน่งของลูกเทนนิสกับ ตำแหน่งของผู้เล่นเพื่อแยกแยะท่าตี

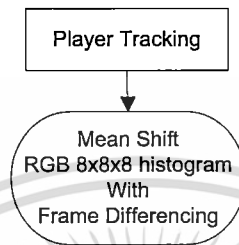


รูปที่ 3.2 แผนภาพแสดงขั้นตอนของการจำแนกท่าตีเทนนิสจากวิถีทัศน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1 การติดตามผู้เล่นเทนนิส (Player Tracking)

การติดตามผู้เล่นเทนนิส มีกระบวนการดังรูปที่ 3.3 โดยเริ่มจาก การกำหนดบริเวณตำแหน่งเริ่มต้นของผู้เล่นเทนนิสเพื่อเข้าสู่กระบวนการติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย โดยใช้ ฮิสโทแกรมขนาด $8 \times 8 \times 8$ ของสี RGB ของบริเวณที่ต้องการจะติดตามเป็นลักษณะเด่น ซึ่งเป็นการนำข้อมูลสี RGB มาใช้ โดยวิธีการหาความแตกต่างระหว่างเฟรมจะถูกนำมาใช้ร่วมด้วย



รูปที่ 3.3 แผนภาพแสดงวิธีการติดตามผู้เล่นเทนนิส

การติดตามผู้เล่นเทนนิสด้วยวิธีย้ายค่าเฉลี่ยมีเงื่อนไขที่ว่า บริเวณที่ทำการติดตามจะต้องสามารถแยกให้เห็นถึงความแตกต่างของฉากหลังได้อย่างเพียงพอ ซึ่งในทางเป็นจริงแล้ว เป็นการยากที่จะทำให้ได้ตามเงื่อนไขนี้ เพราะกรณีที่ผู้เล่นฝั่งไกลสวมเสื้อสีเดียวกับกรรมการ หรือ สีเสื้อของผู้เล่นเทนนิสเป็นสีเดียวกับสนามเทนนิส เพื่อทำการ แก้ปัญหาเหล่านี้ เราได้นำวิธีการหาความแตกต่างระหว่างเฟรมพร้อมด้วยการใช้ RGB แบบ $8 \times 8 \times 8$ ฮิสโทแกรมมาช่วยในการแก้ปัญหา เพราะว่าหลังจากที่ใช้วิธีการหาความแตกต่างระหว่างเฟรมผลที่ได้คือ บริเวณที่มีแต่การเคลื่อนไหว ซึ่งส่วนมากผู้เล่นเทนนิสจะต้องมีการเคลื่อนไหวอยู่ตลอดเวลา ดังนั้นจึงสามารถกำจัดบริเวณที่อยู่นิ่งออกไปได้ อีกทั้งยังใช้แม่สี RGB ซึ่งเป็นการนำข้อมูลสีมาช่วยในการดูความแตกต่างระหว่างผู้เล่นเทนนิสกับบริเวณฉากหลัง

การใช้ RGB แบบ $8 \times 8 \times 8$ ฮิสโทแกรมเป็นการแปลงมิติ RGB จากที่มี 3 มิติไปเป็น 1 มิติโดยใช้ความสัมพันธ์ของ สีแดง สีเขียว และสีน้ำเงิน เพื่อช่วยในการติดตามผู้เล่นเทนนิส โดยเราสามารถแปลง RGB ให้อยู่ใน RGB แบบ $8 \times 8 \times 8$ ฮิสโทแกรมด้วยวิธีดังนี้

$$Scalar = \lfloor R / 32 \rfloor + 8 * \lfloor G / 32 \rfloor + 64 * \lfloor B / 32 \rfloor \quad (3.1)$$

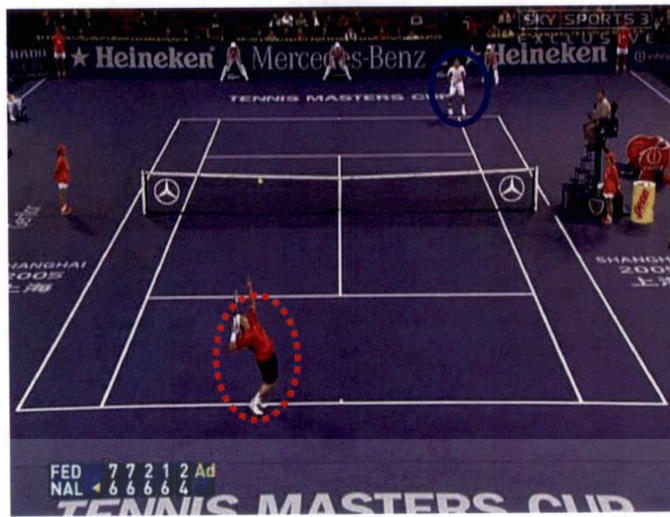
R = ค่าความเข้มของจุดภาพ ในมิติสีแดง

G = ค่าความเข้มของจุดภาพ ในมิติสีเขียว

B = ค่าความเข้มของจุดภาพ ในมิติสีน้ำเงิน

ในสมการ (3.1) จะถูกนำไปใช้ ในส่วนของการสร้างแบบจำลอง ที่ในการติดตามผู้เล่นเทนนิส โดยใช้กระบวนการติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย สามารถสรุปได้ดังรูป ปที่ 3.6 ซึ่งมีขั้นตอนดังต่อไปนี้

- 1 กำหนดจุดศูนย์กลางและขนาดของวงรีให้กับผู้เล่นเทนนิส (y_0) จากนั้นทำการสร้างแบบจำลองโดยใช้สมการ (2.16) ที่ใช้ RGB แบบ $8 \times 8 \times 8$ สีโทแกรมในบริเวณของวงรีเพื่อเก็บเป็นแบบจำลองของผู้เล่นเทนนิส (q)
- 2 ในเฟรมถัดมา ทำวิธีหาความแตกต่างระหว่างเฟรม ได้ผลลัพธ์เป็นภาพระดับเทา จากนั้นแปลงภาพระดับเทา ให้เป็นภาพแบบไบนารีพร้อมทั้งทำมอร์ฟอโลยี (Morphology) ด้วยตัวดำเนินการแบบปิด (Closing Operation) คำนวณค่าสีของจุดภาพที่เป็นพอร์กราวด์ (ผู้เล่น) สุดท้ายคำนวณ RGB แบบ $8 \times 8 \times 8$ สีโทแกรมโดยใช้จุดศูนย์กลางจากเฟรมก่อนหน้า (y_0) ที่ใช้สมการ (2.16) ได้ผลลัพธ์ของแคนดิเดทเป้าหมาย (p_y0)
- 3 หากค่าวงน้ำหนกจากสมการ (2.21) เพื่อทำการหาตำแหน่งจุดศูนย์กลางใหม่ โดยที่ ใช้แบบจำลอง (q) แคนดิเดทเป้าหมาย (p_y0) และค่าถ่วงน้ำหนักของแต่ละจุดภาพในเฟรมเดียวกับที่ได้ทำการหาแคนดิเดทเป้าหมาย ผลลัพธ์คือ ค่าถ่วงน้ำหนักของแต่ละจุดภาพ (w)
- 4 นำค่าถ่วงน้ำหนัก (w) เพื่อใช้ในการหาตำแหน่งใหม่ ในสมการ (2.23) ได้จุดศูนย์กลางใหม่คือ (y_1)
- 5 หาแบบจำลองในตำแหน่งจุดศูนย์กลาง (y_1) โดยใช้สมการ (2.16) ได้ผลลัพธ์ (p_y1)
- 6 ทำการเปรียบเทียบความเหมือน โดยใช้ Bhattacharyya เป็นฟังก์ชันการเปรียบเทียบระหว่าง (q, p_y0) กับ (q, p_y1) ที่ใช้สมการ (2.18)
- 7 ถ้าค่าของ Bhattacharyya (q, p_y0) มีค่ามากกว่า (q, p_y1) แสดงว่าผู้เล่นไม่มีการเคลื่อนที่ ดังนั้นเราจะทำการเพิ่มเฟรมเพื่อทำในเฟรมถัดไปและทำในขั้นตอนที่ 2 จนกว่าจะครบทุกเฟรม
- 8 ถ้าค่าของ Bhattacharyya (q, p_y1) มีค่ามากกว่า (q, p_y0) แต่ระยะการเคลื่อนที่ในตำแหน่ง (y_1) เมื่อเทียบกับ (y_0) แล้ว ได้ระยะทางที่ยอมรับไม่ได้ เราก็จะทำการเพิ่มเฟรมเพื่อทำในเฟรมถัดไป และทำในขั้นตอนที่ 2 จนกว่าจะครบทุกเฟรม
- 9 ถ้าค่าของ Bhattacharyya (q, p_y1) มีค่ามากกว่า (q, p_y0) แต่มีระยะการเคลื่อนที่ในตำแหน่ง (y_1) เมื่อเทียบกับ (y_0) แล้ว ได้ระยะทางที่ยอมรับได้เราจะกลับไปทำในขั้นตอนที่ 2 แต่ยังคงไม่เพิ่มเฟรมคือยังทำที่เฟรมปัจจุบัน ทำการคำนวณค่าตำแหน่งของแบบจำลองคู่แข่งแต่ใช้จุดศูนย์กลาง (y_1) จากนั้นคำนวณหาค่าถ่วงน้ำหนักในขั้นตอนที่ 3 และจุดศูนย์กลางใหม่ในขั้นตอนที่ 4-5 และทำการเปรียบเทียบ Bhattacharyya ในขั้นตอนที่ 6

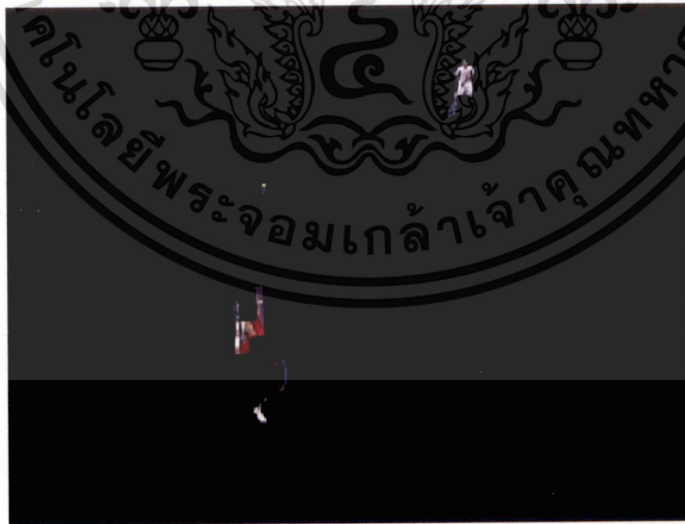


(ก)



(ข)

(ค)

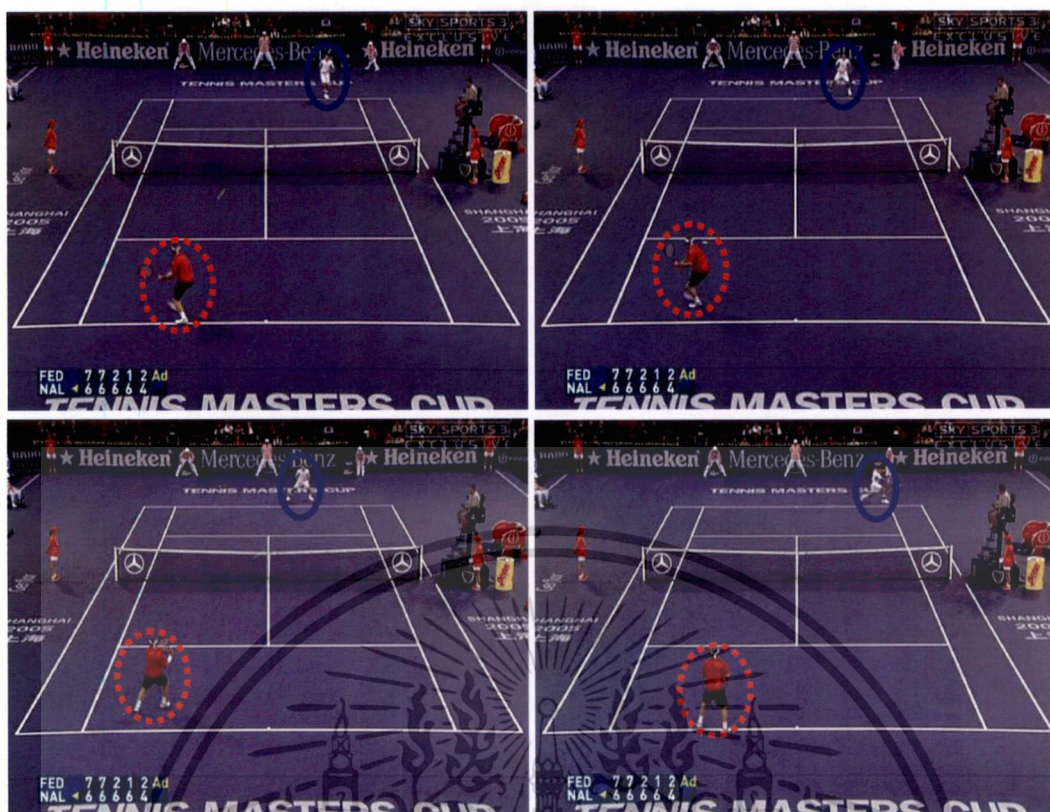


(ง)

รูปที่ 3.4 (ก) จุดศูนย์กลางและขนาดวงรีที่กำหนด (ข) ภาพหลังจากใช้วิธีหาความแตกต่างระหว่าง

เฟรม (ค) ภาพจากการทำตัวดำเนินการแบบปิด (ง) ภาพหลังจากคืนค่าสีจุดภาพ

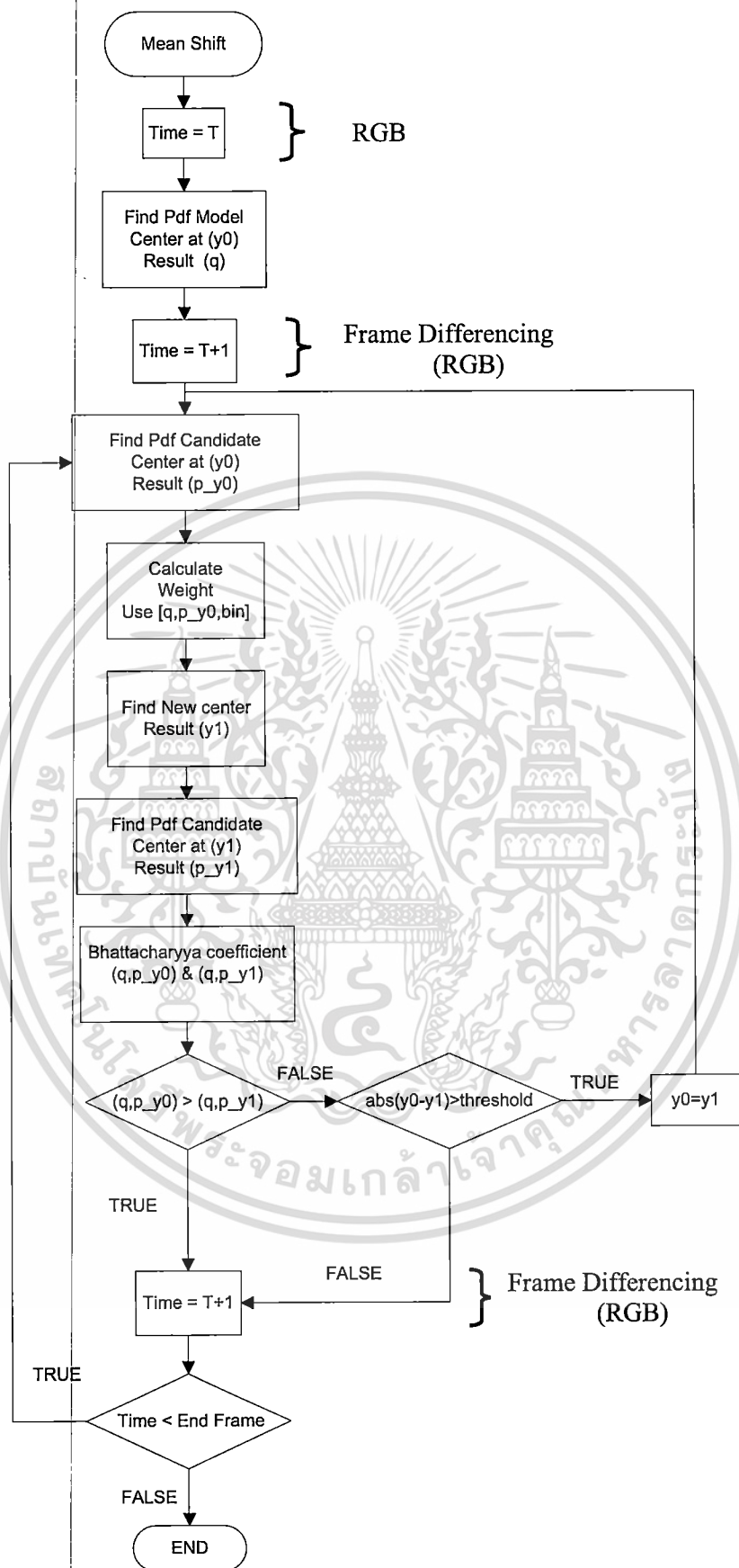
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 ตัวอย่างผลลัพธ์จากการติดตามตำแหน่งของผู้เล่นทั้งสองฝั่ง

รูปที่ 3.4 (ก) แสดงถึง ตัวอย่างของการกำหนดบริเวณวงรีเพื่อที่จะทำการติดตามผู้เล่นใกล้ของ
 ในขั้นตอนที่ 1 ส่วนรูปที่ 3.4 (ข) เป็นการใช่วิธีการหาความแตกต่างระหว่างเฟรม เพื่อหาบริเวณที่มี
 การเคลื่อนไหว จากนั้นในรูปที่ 3.4 (ค) ทำมอร์ฟอโลยีด้วย ตัวดำเนินการแบบปิดเพื่อทำการ
 เชื่อมต่อจุดภาพ และ รูปที่ 3.4 (ง) คำนวณค่าสีของจุดภาพในบริเวณที่เป็นฟอร์กรานด์ ซึ่งรูปที่ 3.4
 (ข)(ค)(ง) เป็นผลจากการทำในขั้นตอนที่ 2 และแสดงผลลัพธ์ของการติดตามผู้เล่นทั้งใกล้และไกล
 ดังรูปที่ 3.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

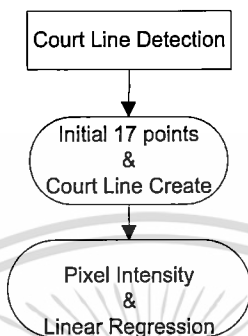


รูปที่ 3.6 กระบวนการติดตามด้วยวิธี การติดตามวัตถุโดยวิธีย้ายค่าเฉลี่ย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

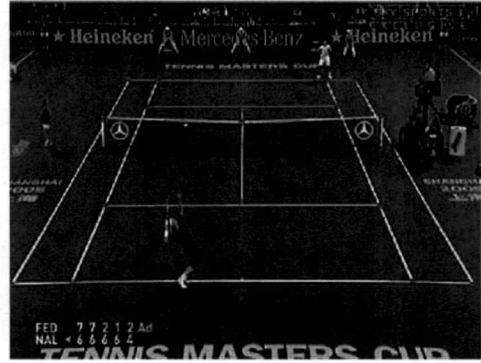
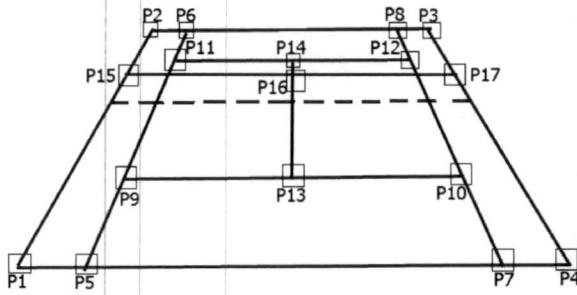
3.2 การตรวจหาเส้นของสนามเทนนิส (Court Line Detection)

คุณลักษณะเด่นของเส้นสนามที่ผู้วิจัยใช้ในวิทยานิพนธ์ฉบับนี้ คือสีของเส้นสนามซึ่งเป็นสีขาว และเป็นเส้นตรง ดังนั้นในการตรวจจับหาเส้นสนามเรา จะได้จากบริเวณที่มีความเข้มจุดภาพสูง และใช้กระบวนการประมาณเส้นตรงด้วยวิธีถดถอยเชิงเส้น



รูปที่ 3.7 แผนภาพแสดงวิธีตรวจหาเส้นสนามเทนนิส

การหาเส้นของสนามเทนนิส ดังรูปที่ 3.7 เริ่มจากการกำหนดจุดในบริเวณขอบของสนามเทนนิสทั้ง 17 จุด จากนั้นเรานำ 17 จุดจากที่กำหนด มาทำการเชื่อมโยงกันระหว่างคู่ของแต่ละจุด เพื่อสร้างเป็นสนามเทนนิสดังในรูปที่ 3.8 จากนั้นทำการหาค่าความเข้มของจุดภาพในบริเวณที่กำหนด เมื่อได้ตำแหน่งที่มีความเข้มของจุดภาพสูงสุดของในแต่ละเส้น นำพิกัดที่มีค่าความเข้มของจุดภาพสูงสุดลบกับพิกัดของเส้นที่ได้จากการสร้างในตอนแรก หลังจากนั้นหาค่าฐานนิยม และหาพิกัดที่อยู่ในช่วงของฐานนิยมบวกลบด้วยซิดเริ่มเปลี่ยน โดยซิดเริ่มเปลี่ยน ได้มาจาก ขนาดของเส้นสนามเทนนิส ของในเฟรมที่มีการกำหนดจุดขอบสนามเทนนิส เหตุผลในการหาค่าฐานนิยม เพราะว่าจากการที่เส้นของสนามเทนนิสเป็นเส้นตรงจึงทำให้ผลต่างของพิกัดมีค่าที่ซ้ำกันเป็นส่วนใหญ่นำพิกัดที่อยู่ในช่วงของฐานนิยมทำวิธีการถดถอยเชิงเส้น



(ก)

(ข)

รูปที่ 3.8 (ก)(ข) จุดที่ถูกกำหนด 17 จุดในสนามเทนนิส

ขั้นตอนของการหาเส้นสนาม มีดังนี้

1. ที่ $t = 1$ แปลงภาพสีเป็นภาพระดับเทา กำหนด 17 จุด ตามรูปที่ (3.8) (ข)
2. สร้างเส้นตรงจากจุด ในขั้นตอนที่ 1 ได้ดังรูปที่ 3.9 (ก) 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 กลุ่ม คือ กลุ่มที่อยู่ในแนวนอนและกลุ่มที่อยู่ในแนวตั้ง

$$LineX_t = \text{เส้นที่ 2, เส้นที่ 4, เส้นที่ 7, เส้นที่ 8, เส้นที่ 10, เส้นที่ 11}$$

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

$$L_t = [LineX_t, LineY_t]$$

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

3. $t = 2$

4. ขยายขนาดเส้นใน $LineX_{t-1}$ และ $LineY_{t-1}$ โดยขนาดที่ขยายได้จาก ความหนาของเส้น สนามเทนนิส ในภาพจากขั้นตอนที่ 1 เส้น $LineX_{t-1}$ จะขยายแกนในพิกัด y ดังรูปที่ 3.9 (ง) ส่วนเส้นที่อยู่ใน $LineY_{t-1}$ จะขยายแกนในพิกัด x ดังรูปที่ 3.9 (ค)

$$ExpandLine_t = [(LineX_{t-1} \pm y) (LineY_{t-1} \pm x)]$$

$ExpandLine_t$ คือ เส้นทั้งหมดที่ได้ถูกขยายความหนา

5. หาพิกัดของความเข้มในจุดภาพสูงสุด ของ $ExpandLine_t$ โดยที่ถ้าเป็นเส้นที่ขยายในพิกัด y จะพิจารณาความเข้มของจุดภาพในแนวตั้งในทางกลับกันถ้าเป็นเส้นที่ขยายในพิกัด x จะพิจารณาความเข้มของจุดภาพในแนวนอน

$$MaxLine_t(x, y) = MaxIntensity(ExpandLine_t)$$

$MaxLine_t(x, y)$ คือ พิกัด (x, y) ของตำแหน่งที่มีค่าความเข้มของจุดภาพสูงสุดของในแต่ละเส้น

$ExpandLine_t$ ของภาพในเวลา t

6. คำนวณค่าพิกัดของเส้น L_{t-1}

$$Line_t(x, y) = \text{returncodinate}(L_{t-1})$$

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

7. หาผลต่างของพิกัดจากขั้นตอนที่ 5 และ 6

$$(Line_t(x, y) - MaxLine_t(x, y)) = Pix_i(x, y)$$

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

8. นำพิกัด x และ y ของทุกตำแหน่ง i ใน Pix_i มารวมกันได้

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

9. หาฐานนิยมของ $Mode(D_i) = m$

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

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

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

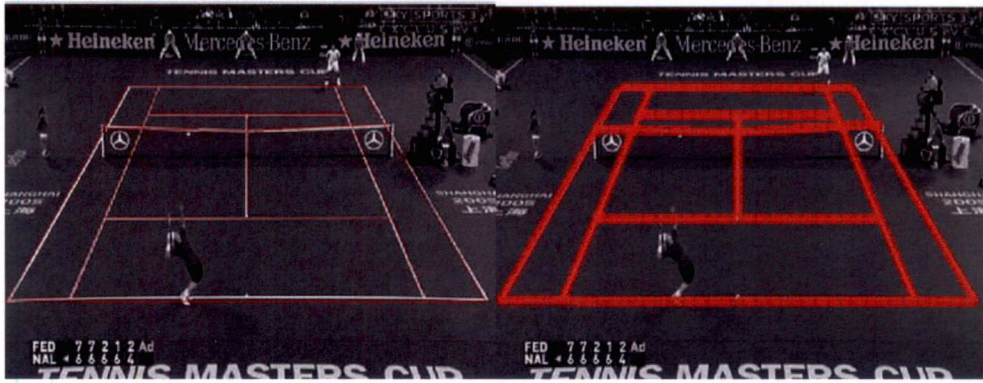
11. ทำวิธีการถดถอยเชิงเส้น จะทำให้ได้ผลลัพธ์คือเส้นสนาม

$$LineCourt_t = \text{linearregression}(Y_i)$$

$LineCourt_t$ ทำการเก็บผลลัพธ์จากการหาเส้นสนามในเวลา t

12. กำหนด $Line_t = LineCourt_t$ เพื่อใช้เป็นข้อมูลในการทำเฟรม ถัดไป

13. ทำซ้ำขั้นตอนที่ 4-12 จนครบเฟรม

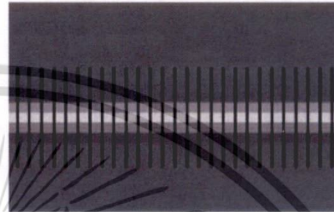


(ก)

(ข)



(ค)

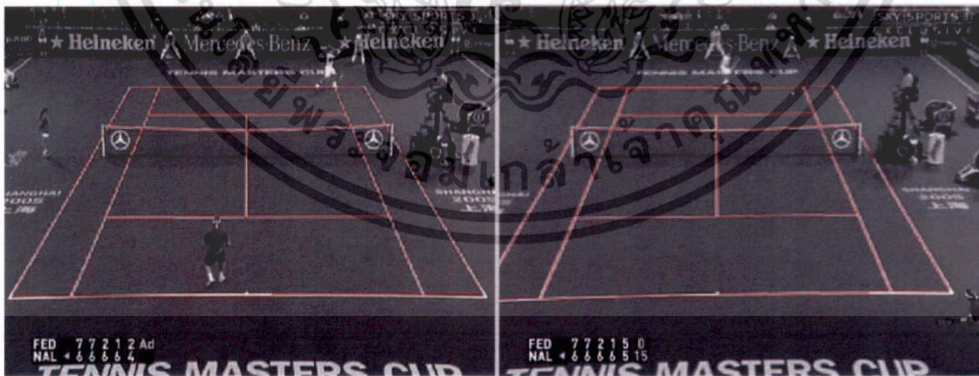


(ง)

รูปที่ 3.9 (ก) นำจุดจากที่กำหนดมาสร้างเส้น (ข) ภาพการขยายขอบเขตพิกัด

(ค) ภาพขยายขนาดเส้น พิกัด x (ง) ภาพขยายขนาดเส้น พิกัด y

รูปที่ 3.9 แสดงรูปตัวอย่างขั้นตอนที่ 2 และ 3 ของ การหาเส้นสนามเทนนิส และได้ผลลัพธ์ในรูปที่ 3.10 (ก)(ข)



(ก)

(ข)

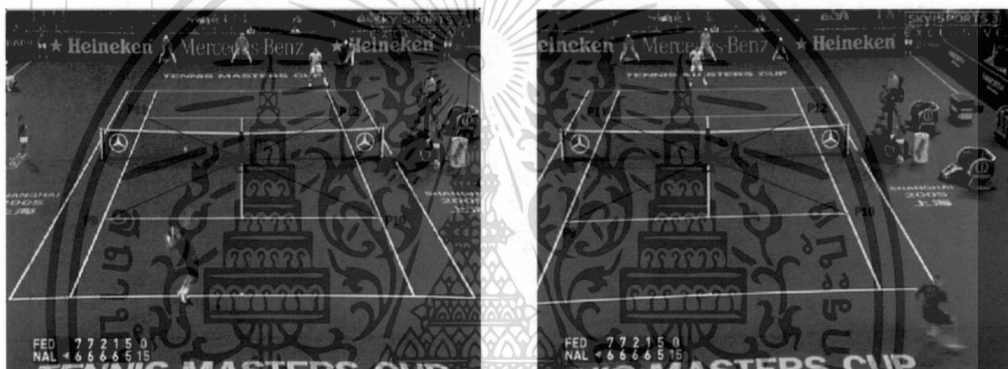
รูปที่ 3.10 (ก)(ข) ตัวอย่างผลลัพธ์ของการหาเส้นสนาม

เราจะทำการหาเส้นปะ จากในรูปที่ 3.8(ก) ที่เป็นส่วนของเน็ตด้านล่าง ซึ่งเราไม่สามารถใช้วิธีการหาความเข้มของจุดภาพสูงได้ เนื่องจากสีของบริเวณตำแหน่งของเน็ตด้านล่างสามารถมีได้หลายสี ดังนั้นในการหาส่วนที่เป็นเน็ตด้านล่าง เราสามารถทำได้โดยเริ่มจาก การหาจุดกึ่งกลางของเอกสารเป็นเอกสารที่สวอนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สนามเพราะ ผลลัพธ์ที่ได้คือตำแหน่งกึ่งกลางส่วนของเน็ตด้านล่าง จากนั้นหาระยะห่างในตำแหน่งจุดกึ่งกลางของเน็ตด้านล่างกับเน็ตด้านบน โดยที่ตำแหน่งของเน็ตด้านบนได้จากขั้นตอนในการหาเส้นสนามเทนนิส หลังจากได้ค่าระยะห่างแล้วก็จะนำมาบวกกับตำแหน่งของเน็ตด้านบนทุกตำแหน่งก็จะได้ส่วนของเน็ตด้านล่าง

ทำการหาโดย

1. สร้างเส้นระหว่าง $P11-P10$ และ $P12-P9$
2. หาจุดตัดของเส้นทั้งสอง จะได้จุดกึ่งกลางของเน็ต (CI) ซึ่งแสดงในรูปที่ 3.11 (ก) (ข) หาระยะห่างระหว่าง CI กับ $P16$ ได้ระยะห่างคือ ($D1$)
3. สูดท้ายจะได้เส้นปะในรูปที่ 3.8 ที่เป็นส่วนของเน็ตด้านล่าง โดยนำเส้นที่สร้างระหว่างจุด $P15$ กับ $P16$ และ จุด $P16$ กับ $P17$ มาบวกกับค่า $D1$



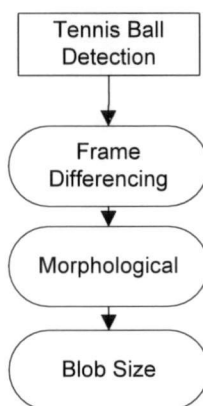
(ก)

(ข)

รูปที่ 3.11 (ก)(ข) จุดตัด CI และระยะห่างที่เป็นแท่งสีดำ $D1$

3.3 การตรวจหาตำแหน่งของลูกเทนนิส (Tennis Ball Detection)

ด้วยความเร็วของลูกเทนนิสอันเกิดจากการตีของผู้เล่น หรือ ประสิทธิภาพของกล้องวีดิทัศน์ที่ใช้ในการบันทึกการแข่งขันเทนนิสที่เป็นชนิดทั่วไปคือมีความสามารถในการบันทึกประมาณ 25-30 เฟรม/วินาที อีกทั้งวีดิทัศน์ก็มีรายละเอียดของภาพต่ำ จึงเป็นการยากที่เราจะใช้สีของลูกเทนนิสหรือลักษณะรูปร่างของลูกเทนนิสเพื่อใช้ในการตรวจหาตำแหน่งของลูกเทนนิส



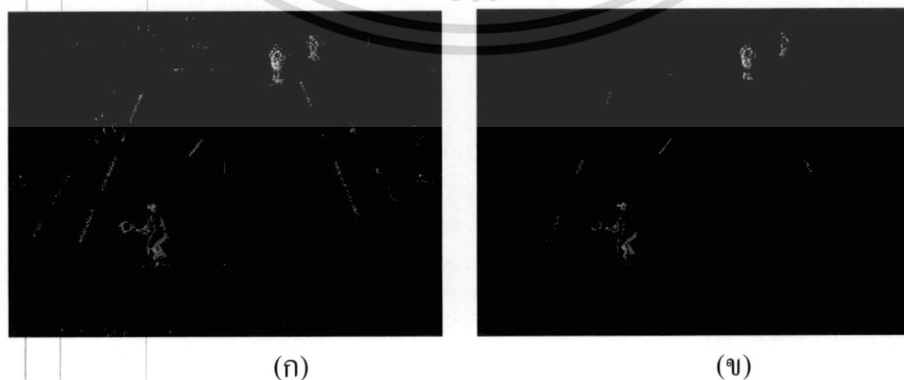
รูปที่ 3.12 แผนภาพแสดงการตรวจหาตำแหน่งของลูกเทนนิส

ขั้นตอนดังรูปที่ 3.12 เริ่มจาก การหาความแตกต่างระหว่างเฟรม เพื่อทำการหาบริเวณของสิ่งที่มีการเคลื่อนไหว จากนั้นทำการปรับปรุงภาพด้วยการตัวกระทำมอร์ฟอโลยี และสุดท้ายทำการกรองบล็อบด้วยขนาด เพื่อหาตำแหน่งของลูกเทนนิส

3.3.1 การหาความแตกต่างระหว่างเฟรม

วิธีการหาความแตกต่างระหว่างเฟรม เป็นวิธีการอย่างหนึ่งของการตรวจจับการเคลื่อนไหวที่เราได้กล่าวไว้แล้วใน บทที่ 2 ซึ่งเป็นวิธีที่สามารถทำได้ง่ายและมีความรวดเร็วในการทำงาน เราจึงนำวิธีการหาความแตกต่างระหว่างเฟรม มาใช้ในการหาบริเวณของสิ่งที่มีการเคลื่อนไหวในเฟรมของวิดีโอ เพื่อที่จะหาตำแหน่งของลูกเทนนิส

วิธีการหาความแตกต่างระหว่างเฟรม มีสิ่งที่จะต้องคำนึงถึงคือการเลือกค่าขีดเริ่มเปลี่ยนให้มีความเหมาะสม เพราะถ้าเราเลือกค่า ขีดเริ่มเปลี่ยนที่สูงเกินไป ก็จะได้รายละเอียดที่ต่ำ แต่ถ้าเลือกค่าขีดเริ่มเปลี่ยนที่ต่ำไปก็จะได้รายละเอียดของการเคลื่อนไหวที่มากเกินไป ดังตัวอย่างในรูปที่ 3.13 (ก) และ 3.13 (ข)



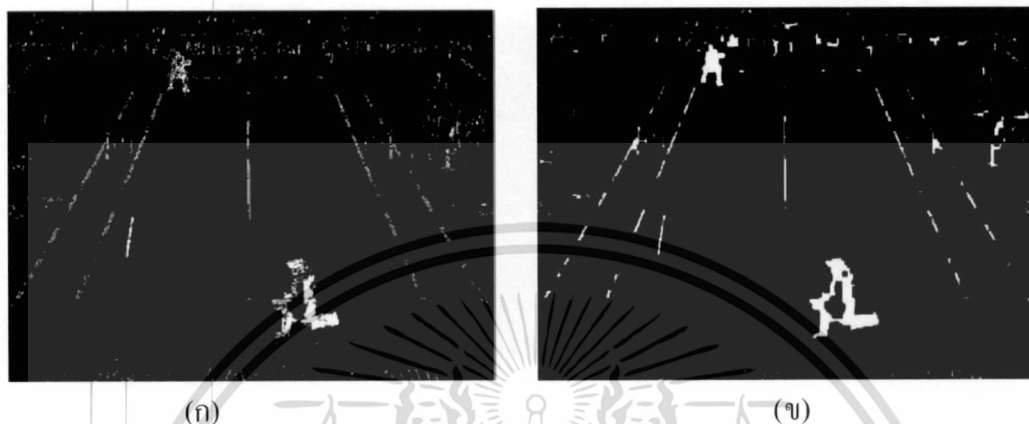
รูปที่ 3.13 ตัวอย่าง ผลของการหาความแตกต่างระหว่างเฟรม

(ก) ด้วยค่าขีดเริ่มเปลี่ยน = 15 (ข) ด้วยค่าขีดเริ่มเปลี่ยน = 25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 มอร์ฟอโลยี

ภายหลังจากที่ได้ทำวิธีการหาความแตกต่างระหว่างเฟรม ขั้นตอนถัดมาคือทำมอร์ฟอโลยีด้วยตัวดำเนินการแบบปิด เพื่อทำการลบรอยขีด และ ทำการเชื่อมต่อบริเวณที่เป็นช่องว่าง เป็นการปรับปรุงภาพให้พร้อมสำหรับการทำในขั้นตอนถัดไป



รูปที่ 3.14 ตัวอย่างภาพที่ได้ (ก) ก่อนทำตัวดำเนินการแบบปิด (ข) หลังทำตัวดำเนินการแบบปิด

ในรูปที่ 3.14 (ก) คือภาพก่อนการทำ ด้วยตัวดำเนินการแบบปิด จะเห็นได้ว่าบางช่วงบริเวณยังเกิดช่องว่าง และหลังจากการทำด้วยตัวดำเนินการแบบปิด ในรูปที่ 3.14 (ข) จะได้ภาพที่สมบูรณ์มากขึ้น คือ บริเวณของจุดภาพ ได้ถูกเชื่อมต่อกันและพร้อมที่จะนำไปใช้ในขั้นตอนถัดไป

3.3.3 กรองบล็ออบโดยใช้ขนาด

ภายหลังจากการตรวจจับสิ่งที่เคลื่อนไหว ขั้นตอนถัดมาเราจะทำการหาบริเวณใดบ้างที่เราตรวจจับมาได้คือลูกเทนนิส โดยในการหาว่าลูกเทนนิสอยู่ที่ตำแหน่งใด เราใช้ช่วงขนาดของบล็ออบในการกรองที่ซึ่งช่วงขนาดที่ใช้ในการกรองได้มาจาก การเทรน (train) ด้วยขนาดของลูกเทนนิสจากกลุ่มตัวอย่าง



(ก)



(ข)

รูปที่ 3.15 (ก) ตัวอย่าง ภาพของลูกเทนนิสเมื่อมีการเคลื่อนที่ (ข) ทำการขยายภาพในตำแหน่งลูกเทนนิส

จากที่ได้กล่าวมาข้างต้นแล้วว่า คุณสมบัติของลูกเทนนิสเมื่อใช้กับวิถีทัศน์ที่มีรายละเอียดต่ำ จะทำให้คุณสมบัติรูปทรงและสีเปลี่ยนไป ดังรูปที่ 3.15 (ก) โดยในวงกลมสีดำคือลูกเทนนิสที่มีการเคลื่อนที่ จะเห็นได้ว่ารูปทรงของลูกเทนนิสจากเดิมที่เป็นทรงกลมก็จะถูกเปลี่ยนไปเป็นลักษณะที่เป็นเส้นตรงหรือเรียกว่า โมชันเบลต ดังรูปที่แสดงใน 3.15 (ข) คือการนำบริเวณของวงกลมสีดำจากในรูป 3.15 (ก) มาขยายจะเห็นได้ว่าคุณสมบัติสีมีการเปลี่ยนไป จากที่สีเหลืองของลูกเทนนิสกลายเป็นสีขาว ดังนั้นสิ่งที่สังเกตได้คือช่วงขนาดลูกเทนนิส เราจึงทำการกรองโดยใช้ช่วงขนาด ของบล็อบ แสดงผลลัพธ์ที่ได้ ดังรูปที่ 3.16

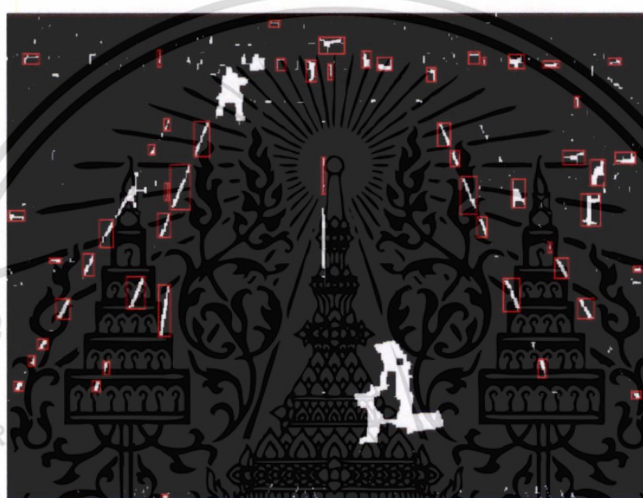


รูปที่ 3.16 ผลลัพธ์ของการกรองบล็อบโดยใช้ขนาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

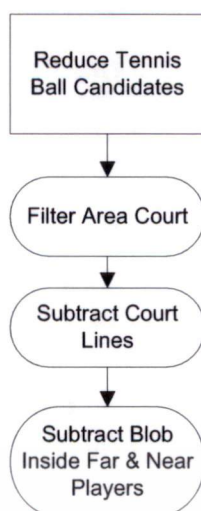
3.4 วิธีลดปริมาณความผิดพลาดของการตรวจหาลูกเทนนิส (Reduce Tennis Ball Candidates)

ในการแข่งขันกีฬาเทนนิสโดยใช้มุมกล้องระยะไกล ในบางช่วงจังหวะที่กล้อง มีการแพนตามผู้เล่น หรือสภาพแสงที่เปลี่ยนไป จะทำให้เกิดปัญหาในการตรวจหาลูกเทนนิส ดังตัวอย่างในรูปที่ 3.17 เพราะจากวิธีการตรวจหาตำแหน่งของลูกเทนนิส เราได้เลือกใช้วิธีการหาความแตกต่างระหว่างเฟรม ซึ่งเป็นการใช้เฟรมต่อเฟรมในการทำ ดังนั้นจึงทำให้เกิดบริเวณที่มีการเคลื่อนไหวในภาพเป็นอันมาก ประกอบกับเราใช้วิธีการกรองบล็อบโดยใช้ขนาด เพื่อที่ทำการหาตำแหน่งของลูกเทนนิส เพียงอย่างเดียว จึงทำให้เกิดผลลัพธ์ที่ผิดพลาดเป็นจำนวนมาก



รูปที่ 3.17 ตัวอย่าง เฟรมที่กล้องมีการแพนตามผู้เล่น

จากรูปที่ 3.17 จะเห็นได้ว่า ผลจากการกรองบล็อบโดยใช้ขนาดจะทำให้ได้ตำแหน่งที่ไม่ใช่ลูกเทนนิสเป็นจำนวนมาก อันเนื่องมาจากการแพนของกล้อง จากการสังเกตในรูปที่ 3.17 ตำแหน่งที่ไม่ใช่ลูกเทนนิส ส่วนมากแล้วจะเกิดจาก เส้นของสนามเทนนิส หรือส่วนที่เป็นกรรมการข้างสนาม เป็นต้น ดังนั้นเราจึงหาวิธีเพื่อลดความผิดพลาดของการตรวจหาลูกเทนนิส ได้ดังรูปที่ 3.18

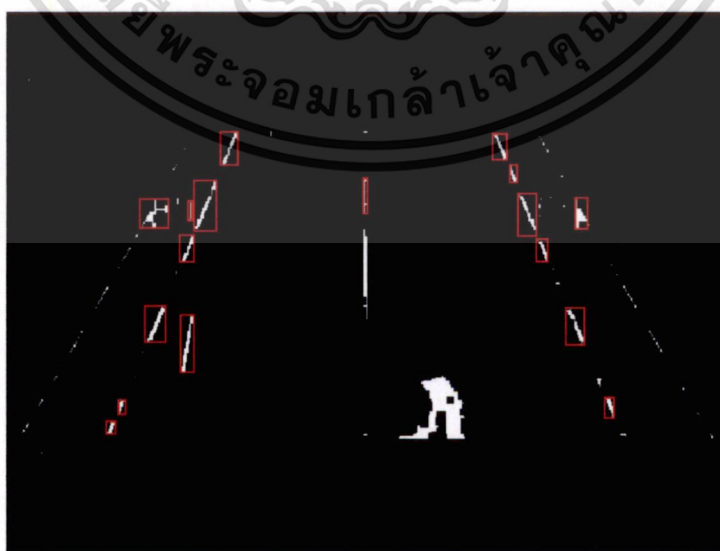


รูปที่ 3.18 แผนภาพแสดงการลดความผิดพลาดของลูกเทนนิส

โดยในขั้นแรกนำภาพและผลลัพธ์จากการกรองบอล ที่ได้จากส่วนงานการตรวจหาลูกเทนนิสมาเข้าสู่วิธีการกรองบอลที่อยู่ในบริเวณสนามเทนนิส จากนั้นลบเส้นสนามเทนนิส ป้ายโฆษณา และบอลที่อยู่ในผู้เล่น ไกลและใกล้ออก

3.4.1 กรองบริเวณสนามเทนนิส

นำผลลัพธ์จากการหาเส้นสนามเทนนิส เพื่อช่วยในการหาบริเวณลูกเทนนิสที่อยู่ในสนามเทนนิส เพราะส่วนมากแล้วลูกเทนนิสจะอยู่ในบริเวณของสนามเทนนิส เราจึงกรอง บอลที่อยู่ในสนามเทนนิส



รูปที่ 3.19 ค่าของจุดภาพที่อยู่ในสนามเทนนิส

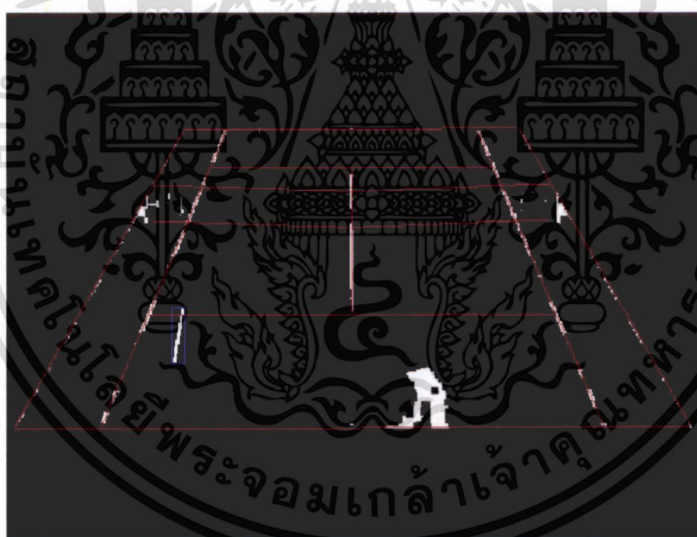
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เฉพาะในโครงการที่ขอเท่านั้น มิใช่ผู้จัดทำให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.19 แสดงถึงการกรองบอลที่อยู่ในสนามเทนนิส ซึ่งจะเห็นได้ว่าจะสามารถกำจัดบอลที่อยู่นอกสนามเทนนิส แต่ยังคงได้ตำแหน่งลูกเทนนิสที่ผิดพลาด อันเนื่องมาจากเส้นของสนามเทนนิส

3.4.2 ลบเส้นสนามเทนนิส ป้ายโฆษณาและบอลที่อยู่ในผู้เล่นเทนนิส

อย่างที่ได้อธิบายไว้แล้วว่าเมื่อกล้องมีการแพนจะทำให้เกิดการตรวจจับที่ผิดพลาด การกรองโดยใช้ขนาด ส่วนมากผลลัพธ์ที่กรองออกมา จะได้รับบริเวณที่เป็นเส้นสนามเทนนิส เรานำผลลัพธ์ที่ได้จากการหาเส้นสนามมาช่วยในการลบเส้นสนาม กับช่วยในการลบบริเวณป้ายโฆษณาที่อยู่ในบริเวณเนต

ทำการลบบอลที่อยู่ในบริเวณของผู้เล่นเทนนิสฝั่งไกลและใกล้เพื่อให้ได้ผลลัพธ์จากการกรองด้วยบอลได้ผลที่ถูกต้อง ดังรูปที่ 3.20 เส้นสีแดงคือผลลัพธ์ที่ได้จากขั้นตอนการหาเส้นสนามเทนนิส ส่วนกรอบสี่เหลี่ยมสีน้ำเงินแสดงถึงผลลัพธ์ที่ได้จากการทำในส่วนของวิธีลดปริมาณความผิดพลาดของการตรวจหาลูกเทนนิส

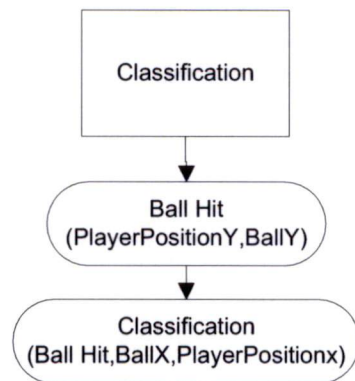


รูปที่ 3.20 ผลลัพธ์ของการตรวจหาลูกเทนนิส

3.5 การจำแนกท่าตีเทนนิส (Tennis Stroke Classification)

ผลลัพธ์ของในแต่ละส่วนงานข้างต้นที่ได้ทำกล่าวมานั้น จะถูกนำมาใช้ในการจำแนกท่าตีเทนนิส โดยจากการสังเกตผู้วิจัยพบว่า เหตุการณ์ของการตีเทนนิสจะเกิดขึ้นก็ต่อเมื่อ ลูกเทนนิสเคลื่อนตัวเข้าใกล้กับตำแหน่งของผู้เล่นเทนนิส ในการตรวจจับหาเฟรมที่มีการตีเทนนิส หาได้โดยการหาบริเวณที่วิถีการเคลื่อนที่ (Trajectory) ของลูกเทนนิสเข้าใกล้ผู้เล่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

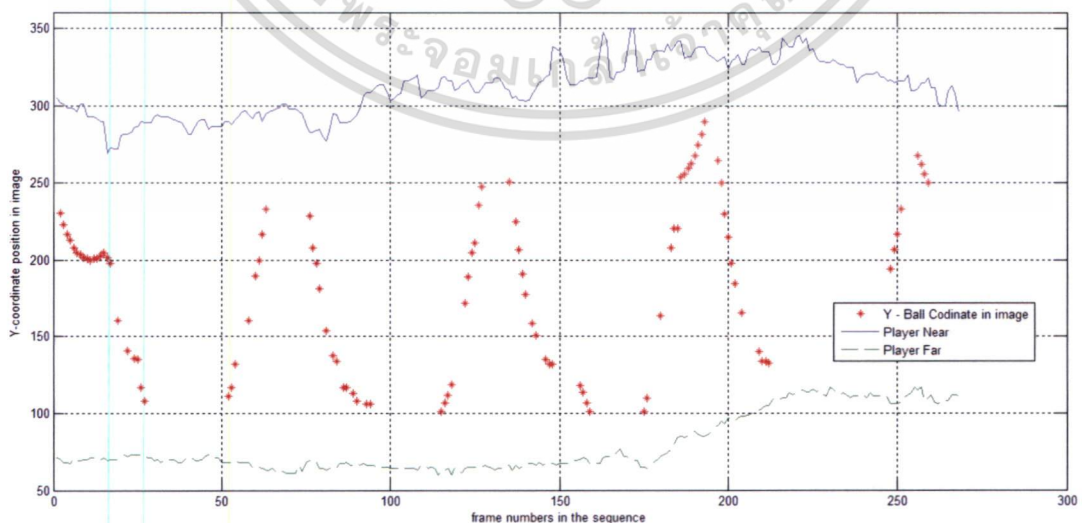


รูปที่ 3.21 แผนภาพแสดงวิธีจำแนกท่าตีเทนนิส

โดยขั้นตอนในรูปที่ 3.21 เริ่มจากใช้ตำแหน่งของผู้เล่นและตำแหน่งของลูกเทนนิส เพื่อทำการหาว่าในเวลาใดที่ผู้เล่นเทนนิส ได้มีการตีลูกเทนนิส จากนั้นทำการจำแนกท่าการตีเทนนิส โดยใช้ผลลัพธ์จากเวลาที่มีการตีลูกเทนนิส ร่วมกับตำแหน่งของผู้เล่นเทนนิสและลูกเทนนิส เพื่อดูว่าลูกเทนนิสอยู่ทางด้านใดของผู้เล่น

3.5.1 ตรวจสอบเหตุการณ์ของการตีลูกเทนนิส

เราจะทำการสร้างกราฟ โดยที่ในแนวตั้งกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรม โดยที่ข้อมูลที่ใช้ในการวาดกราฟ จะประกอบไปด้วย ตำแหน่งของลูกเทนนิสและตำแหน่งที่เป็นจุดศูนย์กลางของผู้เล่นเทนนิส ดังรูป 3.22

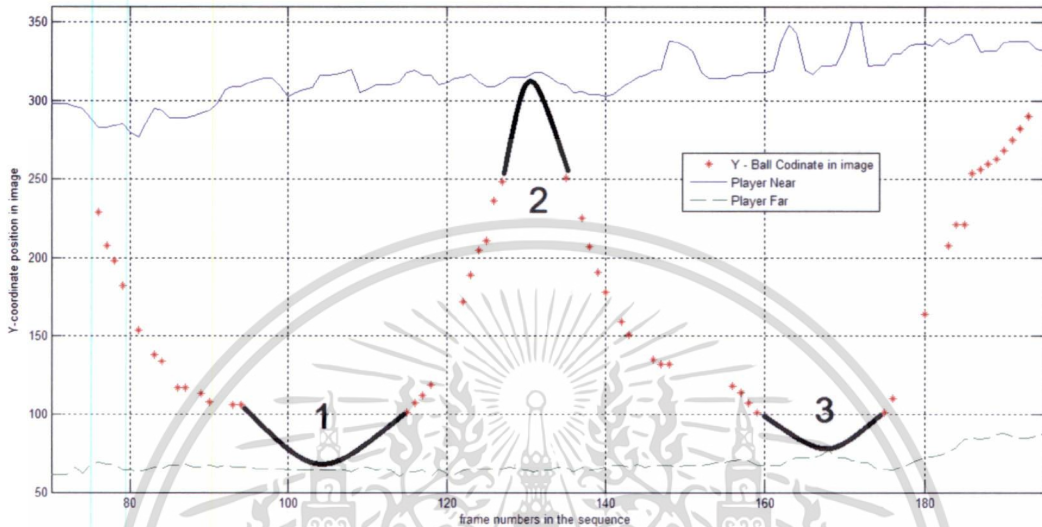


รูปที่ 3.22 กราฟที่ใช้หาเหตุการณ์ตีลูกเทนนิส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กราฟของรูปที่ 3.22 จุดสีแดงคือ ตำแหน่งของลูกเทนนิส ส่วนเส้นสีน้ำเงินด้านบนคือ ตำแหน่งของผู้เล่นฝั่งใกล้ และเส้นสีเขียวด้านล่างคือตำแหน่งผู้เล่นฝั่งไกล

ในการหาเหตุการณ์ของการตีลูกเทนนิส สามารถตรวจสอบได้จากการหาแนววิถี (trajectory) ของลูกเทนนิสที่เข้าใกล้กับแนววิถีของผู้เล่นเทนนิส ดังตัวอย่างในรูปที่ 3.23



รูปที่ 3.23 ตัวอย่างเหตุการณ์ที่มีช่วงการตี 3 ช่วง

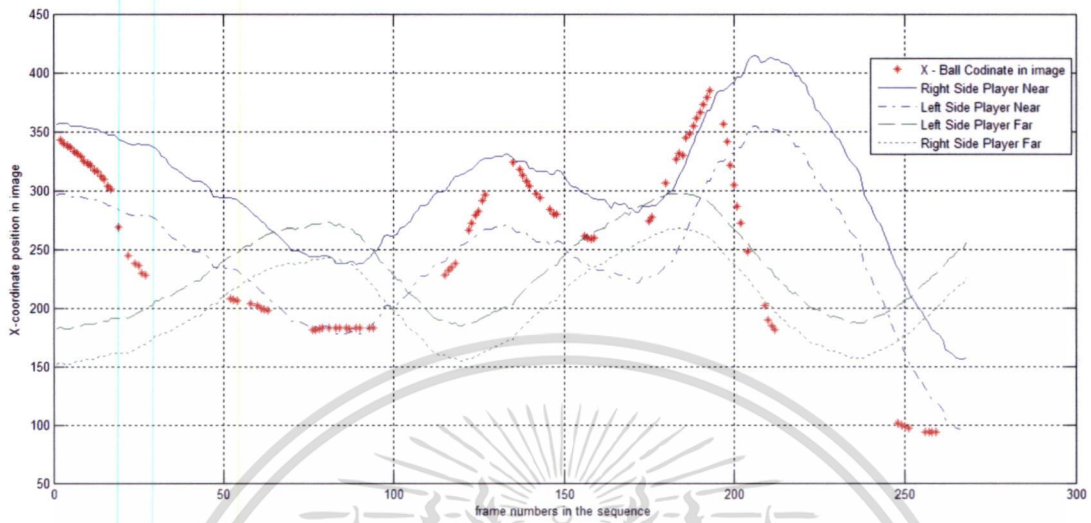
โดยในรูปที่ 3.23 เมื่อทำการตรวจสอบเหตุการณ์ของการตีลูกเทนนิสจะพบได้ว่า มีเหตุการณ์เกิดขึ้นทั้งหมด 3 ครั้ง เป็นเหตุการณ์ที่ผู้เล่นฝั่งใกล้ตี 1 ครั้ง ส่วนผู้เล่น ฝั่งไกลตี 2 ครั้ง ซึ่งได้แสดงตัวเลขกำกับไว้ในรูปที่ 3.23 พิจารณาสีเส้นสีดำของเหตุการณ์การตีเทนนิส ของผู้เล่นฝั่งใกล้ซึ่งที่แสดงไว้คือเลข 1 จะเห็นได้ว่ามีแนววิถีของลูกเทนนิสได้เข้าใกล้กับแนววิถีของผู้เล่นเทนนิสอยู่ด้วยกัน 2 แนว ซึ่งในแนววิถีแรก จะมีเฟรมที่ต่ำกว่าของแนวที่สอง เราจะทำการเรียกแนววิถีแรกนี้ว่า เฟรมก่อนตีลูกเทนนิส ส่วนแนววิถีที่สองเราจะเรียกว่า เฟรมหลังตีลูกเทนนิส ซึ่งเราสนใจแค่เพียงในแนวแรกเท่านั้นคือเฟรมก่อนตีลูกเทนนิส ที่เราจะทำการเก็บเฟรมของในแนวนี้นี้ไว้เพื่อใช้ในการจำแนกท่าตีเทนนิส ส่วนของผู้เล่นฝั่งใกล้เราก็จะพิจารณาด้วยวิธีเดียวกับที่ได้กล่าวไว้

3.5.2 จำแนกท่าตีเทนนิส

ดังที่ได้กล่าวมาแล้วว่า เราจะใช้ผลลัพธ์จากการหาเหตุการณ์ของการตีเทนนิส มาใช้เพื่อการจำแนกท่าตีเทนนิส เพราะเมื่อมีเหตุการณ์ตีเทนนิสก็จะต้องมีท่าตีเกิดขึ้น

ทำการสร้างกราฟ ที่ซึ่งในแนวดิ่งของกราฟ คือ ตำแหน่งพิกัด x ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรม โดยข้อมูลที่ใช้ในการวาดกราฟ จะประกอบไปด้วย ตำแหน่งของลูก

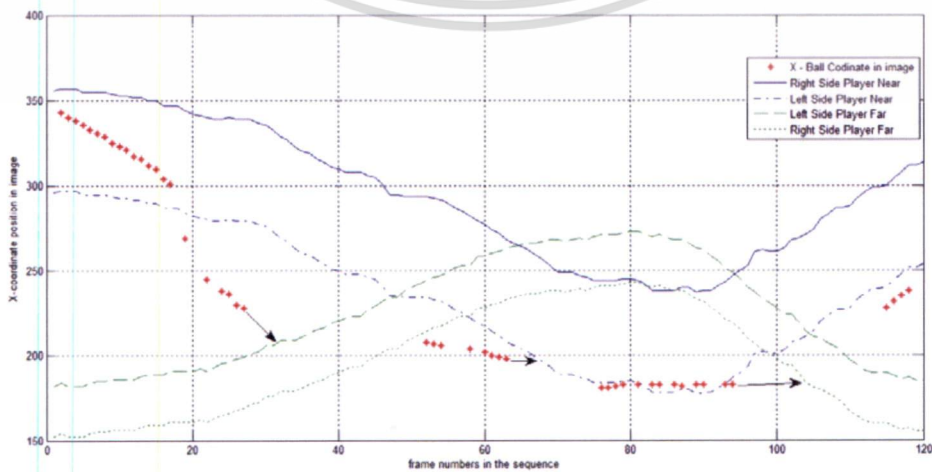
เทนนิสและตำแหน่งทางด้านซ้ายและขวาของผู้เล่นเทนนิส ดังรูปที่ 3.24 ตำแหน่งด้านซ้ายและขวา
 ของ ผู้เล่นได้มาจากกรีสมิของวงรีในส่วนของกราฟติดตาม



รูปที่ 3.24 กราฟที่ใช้เพื่อการจำแนกท่าตีเทนนิส

กราฟของรูปที่ 3.24 เส้นสีน้ำเงินจะแสดงถึงตำแหน่งของผู้เล่นเทนนิสฝั่งใกล้ที่ซึ่งมี 2 เส้น โดยเส้นด้านบนคือตำแหน่งทางด้านขวามือของผู้เล่นส่วนเส้นล่างคือตำแหน่งทางด้านซ้าย ส่วนเส้นสีเขียวแสดงถึงผู้เล่นเทนนิสฝั่งไกล โดยเส้นบนคือตำแหน่งทางด้านซ้ายมือของผู้เล่นส่วนเส้นล่างคือตำแหน่งทางด้านขวามือ จุดสีแดงคือ ตำแหน่งของลูกเทนนิส

นำเฟรมที่ได้จากเหตุการณ์ของการตีลูกเทนนิส ทำการพิจารณาว่าตำแหน่งลูกเทนนิสอยู่
 ด้านซ้ายหรือด้านขวาของผู้เล่นเทนนิส ถ้าอยู่ทางด้านขวาของผู้เล่นฝั่งใกล้ผลลัพธ์ของการตีที่ได้คือ
 “โฟร์แฮนด์” แต่ถ้าอยู่ทางด้านซ้ายผลลัพธ์ของการตีที่ได้คือ “แบคแฮนด์” ส่วนถ้าเป็นผู้เล่นฝั่งไกล
 ผลลัพธ์ของการตีที่ได้คือ แบคแฮนด์ และ โฟร์แฮนด์ ตามลำดับ ดังรูปที่ 3.25

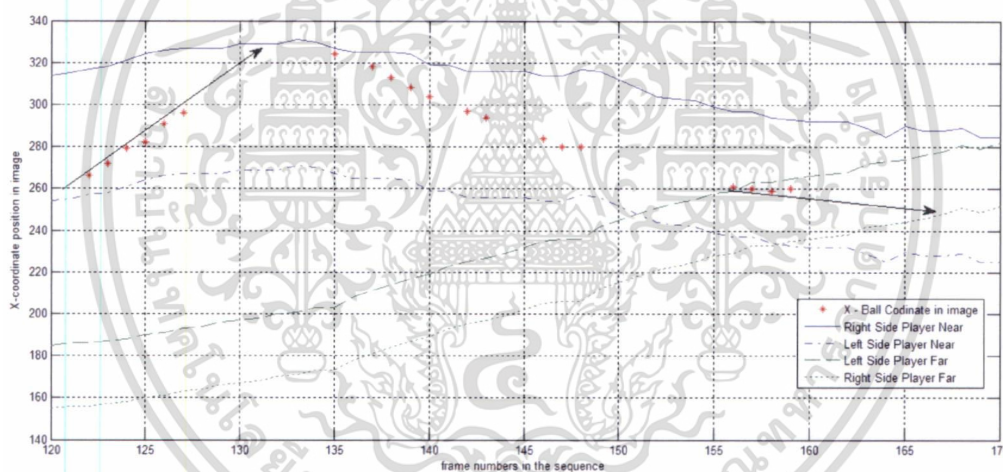


รูปที่ 3.25 ตำแหน่งบอลที่ด้านนอกของผู้เล่นเทนนิส

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อให้นักวิจัยและผู้สนใจอื่น ๆ ที่เห็นนำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.25 เราจะทำการจำแนกท่าตีเทนนิส โดยสมมติว่ามีเหตุการณ์ของการตีเทนนิส 3 ครั้ง โดยเป็นการตีของผู้เล่นฝั่งไกล 2 ครั้ง และ ผู้เล่นฝั่งใกล้ 1 ครั้ง โดยที่ปลายของลูกศร จะหมายถึงเฟรมหลังตีลูกเทนนิส ส่วนหัวลูกศรแสดงถึงระยะทางจากลูกเทนนิสที่ใกล้ที่สุดของ ตำแหน่งผู้เล่นเทนนิสทางด้านซ้ายหรือขวา เราจะทำการพิจารณาลูกศรสีดำที่อยู่ซ้ายสุดของกราฟ ซึ่งเป็นเหตุการณ์การตีเทนนิสของผู้เล่นฝั่งไกล ที่หัวลูกศรชี้ไปทางด้านซ้ายของผู้เล่น ดังนั้นจึงเป็นการตีท่าแบคแฮนด์ ส่วนลูกศรตรงกลาง เป็นเหตุการณ์การตีเทนนิสของผู้เล่นฝั่งใกล้ ที่หัวลูกศรชี้ไปทางด้านซ้ายของผู้เล่นฝั่งใกล้ ดังนั้นจึงเป็นการตีท่าแบคแฮนด์ สุดท้ายลูกศรที่อยู่ขวาสุดเป็นการตีท่าโฟร์แฮนด์ของผู้เล่นฝั่งไกล

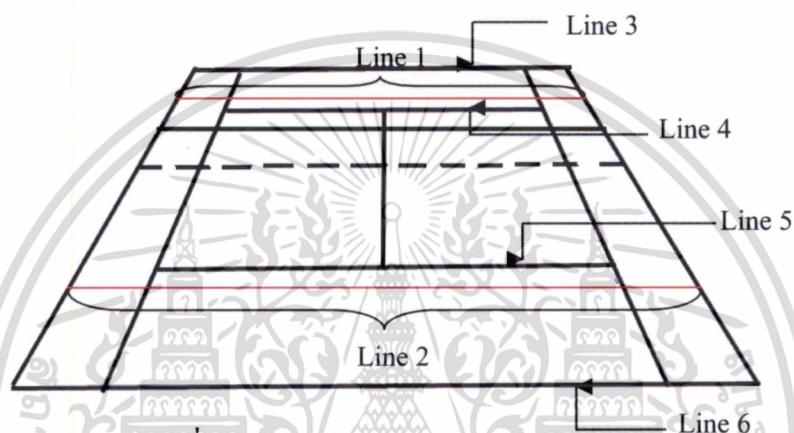
แต่ถ้าตำแหน่งลูกเทนนิสเข้าไปอยู่ระหว่างตำแหน่งด้านซ้ายและขวาของผู้เล่น เราทำการพิจารณาโดยถ้าทิศทางของลูกเทนนิสเริ่มเข้าจากทางด้านขวาของผู้เล่น ผลลัพธ์ที่ได้คือลูกเทนนิส จะอยู่ทางด้านซ้ายของผู้เล่น แต่ถ้าลูกเทนนิสเข้ามาทางด้านซ้ายของผู้เล่น ผลลัพธ์ที่ได้คือลูกเทนนิส จะอยู่ทางด้านขวาของผู้เล่น ดังรูปที่ 3.26



รูปที่ 3.26 ตำแหน่งบอลที่อยู่ระหว่างกลางของผู้เล่นเทนนิส

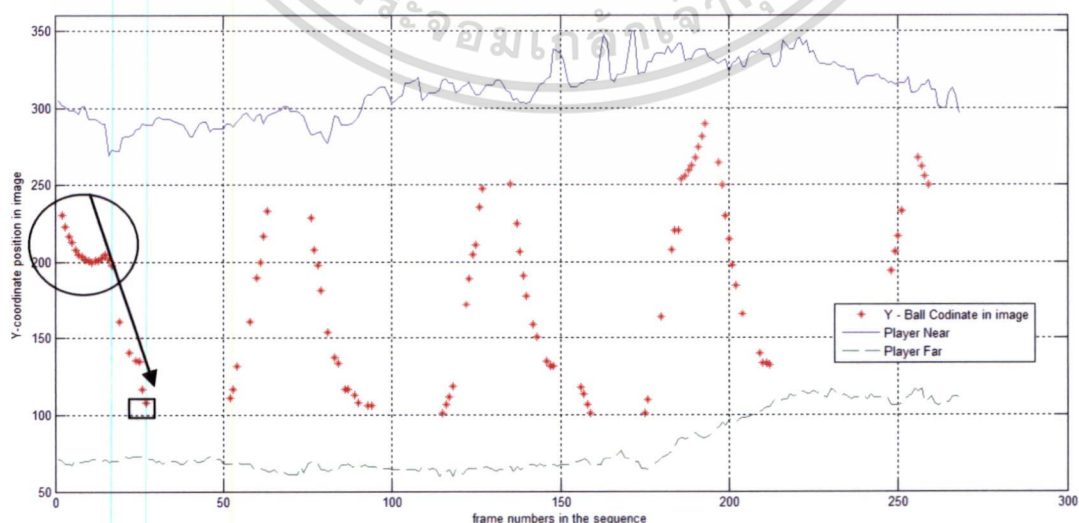
จากรูปที่ 3.26 สมมติมีเหตุการณ์การตีเทนนิส 2 ครั้ง โดยเป็นการตีของผู้เล่นฝั่งใกล้และไกล ตามลำดับ พิจารณาลูกศรทั้ง 2 อันจะพบว่า ลูกเทนนิสในตำแหน่ง เฟรมก่อนตีลูกเทนนิสจะอยู่ระหว่างทางด้านซ้ายและขวาของผู้เล่น ดังนั้นเราจะดูทิศทางของลูกเทนนิสที่เริ่มเข้า ทำการพิจารณาลูกศรสีดำซ้ายสุดจะเห็นได้ว่าลูกเทนนิสจะเข้ามาทางด้านซ้ายของผู้เล่นฝั่งใกล้ ดังนั้นผลลัพธ์ที่ได้ลูกเทนนิสจะเคลื่อนที่ไปทางด้านขวาซึ่งได้ทำตีแบบโฟร์แฮนด์ ส่วนลูกศรสีดำขวาสุดที่ ลูกเทนนิสได้เข้ามาจากทางด้านซ้ายของผู้เล่นฝั่งไกล ดังนั้นผลลัพธ์ที่ได้คือลูกเทนนิสจะต้องไปทางด้านขวา จึงได้ทำตีโฟร์แฮนด์

ทำการพิจารณาตำแหน่งของการตีเทนนิสว่า เป็นการตีหน้าเน็ตหรือไม่ โดยใช้ข้อมูลจาก เส้นสนามเทนนิส และตำแหน่งของผู้เล่นในช่วงเหตุการณ์ ที่มีลูกเทนนิสเข้าใกล้ มาพิจารณา โดยจากรูปที่ 3.27 ถ้าตำแหน่งความสูงด้านล่าง (บริเวณเท้าผู้เล่น) ของผู้เล่นเทนนิสไกลและใกล้อยู่ระหว่างเส้น Line1 ถึง Line2 ระบบจะจำแนก ว่าเป็นการตีหน้าเน็ต ที่ซึ่งเส้น Line4 และ Line5 หาได้จากรูปที่ 3.8 (ก) คือเส้นระหว่างจุด $P9-P10$ และ $P11-P12$ ตามลำดับ ซึ่งเส้น Line1 และ Line2 หาได้จาก $(Line4 - ค่าขีดเริ่มเปลี่ยน)$ และ $(Line5 + ค่าขีดเริ่มเปลี่ยน)$



รูปที่ 3.27 ตำแหน่งพิจารณาการตีหน้าเน็ต

ในการพิจารณาว่าผู้เล่นเทนนิสฝั่งใดเสิร์ฟนั้น จะทำการพิจารณาลักษณะกราฟ ของช่วง ก่อนที่ลูกเทนนิสได้เข้าใกล้เป็นครั้งแรก หรือทิศทางที่ลูกเทนนิสเข้าสู่ผู้เล่น โดยใช้กราฟจาก ขั้นตอนการตรวจหาเหตุการณ์ของการตีลูกเทนนิสมาพิจารณา

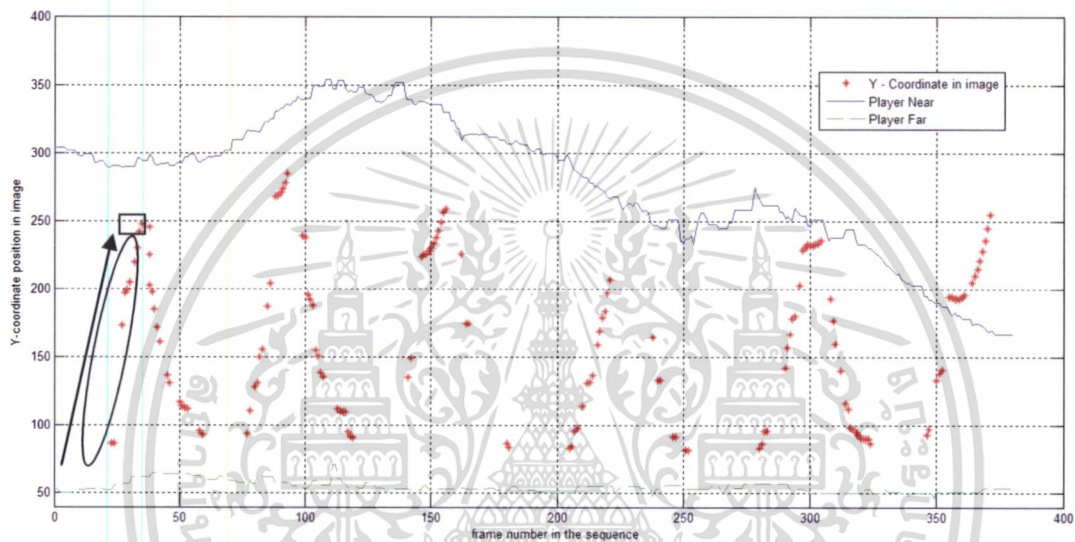


รูปที่ 3.28 กราฟเสิร์ฟของผู้เล่นเทนนิสฝั่งใกล้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.28 แนวโค้งกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวราบ คือ ลำดับของเฟรม โดยที่เป็นการเสิร์ฟของผู้เล่นใกล้จะ ได้กราฟที่มีลักษณะเหมือนในวงกลม มีทิศทางแนววิถีเคลื่อนที่จากผู้เล่นใกล้ไปสู่ผู้เล่นไกล และจากกราฟจะเห็นได้ว่าเหตุการณ์ที่ลูกเทนนิสได้เข้าใกล้เป็นครั้งแรก (กรอบสี่เหลี่ยม) ได้แก่ ผู้เล่นไกล

สาเหตุที่ทำให้ได้กราฟมีลักษณะเหมือนรูปทรง แบบข้างในวงกลม เพราะว่าการเสิร์ฟนั้นผู้เล่นเทนนิส ต้องมีการ โยนลูกเทนนิสขึ้นในแนวโค้ง จากนั้นจะมีช่วงจังหวะที่ลูกเทนนิสเคลื่อนที่ตกลง และผู้เล่นถึงจะตีลูกเทนนิส จึงทำให้ได้รูปทรงกราฟ เหมือนแบบข้างในวงกลม



รูปที่ 3.29 กราฟเสิร์ฟของผู้เล่นเทนนิสฝั่งไกล

จากรูปที่ 3.29 แนวโค้งกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวราบ คือ ลำดับของเฟรม โดยที่เป็นการเสิร์ฟของผู้เล่น ใกล้ ซึ่งจะ ได้กราฟที่มีลักษณะเหมือนในวงรี ซึ่งมีทิศทางแนววิถีเคลื่อนที่จากผู้เล่นใกล้ไปสู่ผู้เล่นใกล้ และจากกราฟจะเห็นได้ว่าเหตุการณ์ที่ลูกเทนนิสได้เข้าใกล้เป็นครั้งแรก (กรอบสี่เหลี่ยม) ได้แก่ ผู้เล่นใกล้

การเสิร์ฟของผู้เล่น ใกล้จะมีการ โยนลูกเทนนิสขึ้นในแนวโค้งเหมือนผู้เล่นใกล้ แต่เนื่องมาจากเราใช้วิธีลดปริมาณความผิดพลาดของการตรวจหาลูกเทนนิส ซึ่งเป็นการ กรองบลิบที่ อยู่ในสนามเทนนิสจึงทำให้ไม่สามารถตรวจพบลูกเทนนิสที่ผู้เล่น ใกล้ โยนขึ้นได้ จึงทำให้ได้กราฟที่เป็นแนววิถีผู้เล่นใกล้ไปสู่ผู้เล่นใกล้ และมีลักษณะกราฟดังวงรี

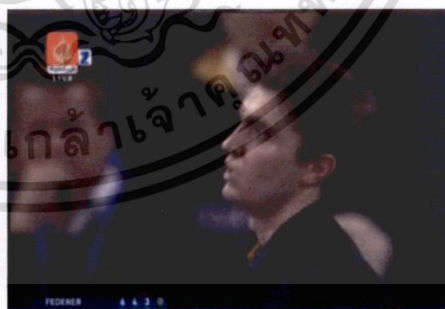
บทที่ 4

วิธีการทดลองและผลการทดลอง

ในบทนี้จะกล่าวถึงแหล่งที่มาพร้อมทั้งวิธีการเตรียมวิดิทัศน์ รวมถึงการกำหนดค่าที่จะใช้ในการทดลอง สุดท้ายแสดงผลการทดลองของการติดตามผู้เล่นเทนนิส การติดตามหาเส้นสนามเทนนิส การตรวจจับลูกเทนนิส และผลของการจำแนกท่าตีเทนนิส

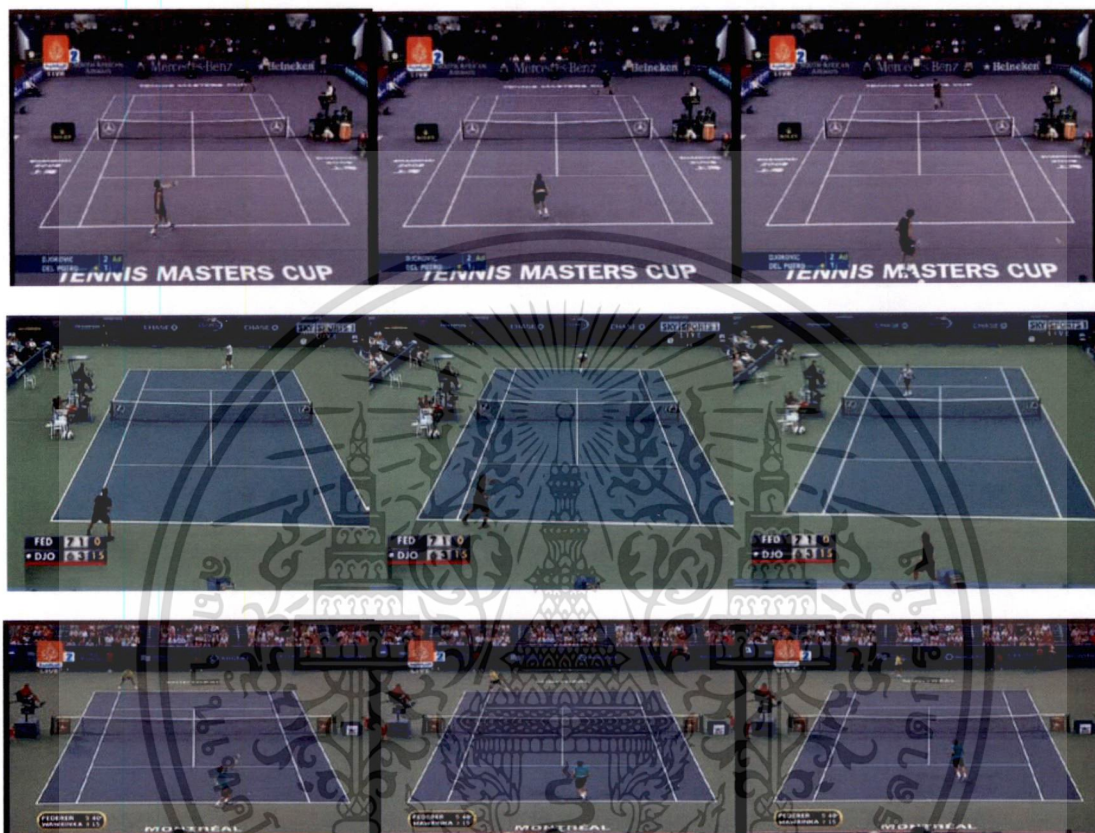
4.1 วิดิทัศน์ที่ใช้ในการทดลอง

ขั้นตอนในวิทยานิพนธ์ฉบับนี้เริ่มจากการหาวิดิทัศน์การแข่งขันกีฬาเทนนิส โดยแหล่งที่มาของวิดิทัศน์นั้น ได้มาจากการค้นหาในอินเทอร์เน็ต ทั้งนี้เนื่องจากผู้ทำวิทยานิพนธ์สังเกตเห็นว่าจำนวนของวิดิทัศน์การแข่งขันเทนนิสที่อยู่ในอินเทอร์เน็ตนั้นมีจำนวนมาก อีกทั้งด้วยปัจจุบันแล้วอินเทอร์เน็ตเป็นเทคโนโลยีที่ทุกคนสามารถเข้าถึงได้ ผมนวกกับความเร็วที่เพิ่มสูงขึ้นจากเมื่อก่อน จึงทำให้สามารถดาวน์โหลดได้จากอินเทอร์เน็ต โดยวิดิทัศน์ที่ได้มานั้นมีความละเอียดอยู่ที่ 512X384 , 640X380, 640X480 พิกเซล มีอัตราเฟรมอยู่ที่ 25 เฟรม/วินาที และสนามเทนนิสเป็นแบบพื้นแข็ง ซึ่งวิดิทัศน์ที่ได้ทำการหา นี้ เป็นการแข่งขันเทนนิสที่มีการบันทึกมาจากการถ่ายทอดสดให้ผู้ชมทางบ้านได้รับชม โดยมีความต่างกันของผู้เล่นเทนนิสทั้งหมด 6 คู่ เป็นการแข่งกันของประเภทชายเดี่ยวทั้งหมด อันมาจากรายการ การแข่งขัน TENNIS MASTER, US OPEN และ ATP



รูปที่ 4.1 ตัวอย่างของวิดิทัศน์การแข่งขันกีฬาเทนนิสที่มีการบันทึกมาจากการถ่ายทอดสด เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากที่หาวิดีโอที่สนใจได้แล้ว ขั้นตอนต่อไปจะทำการตัดแบ่งวิดีโอที่สนใจออกเป็น คลิป (Clip) โดยที่คลิปแต่ละอันนั้นจะถูกตัดด้วยมือโดย เริ่มตัดแบ่ง จากการ ที่ผู้เล่นเทนนิสเริ่ม เสรฟ โดยที่ จะต้องไม่เป็นการเสิร์ฟเสียและจะจบคลิปด้วยการ ได้หรือเสียบะแนบของผู้เล่นเทนนิส โดยมุมมอง ของกล้องมาจากกระยะไกลเพียงมุมมองเดียว



รูปที่ 4.2 ตัวอย่างของภาพนิ่งที่ใช้เป็นอินพุต ซึ่งตัดมาจากวิดีโอ

4.2 การกำหนดค่าและวิธีการทดลอง

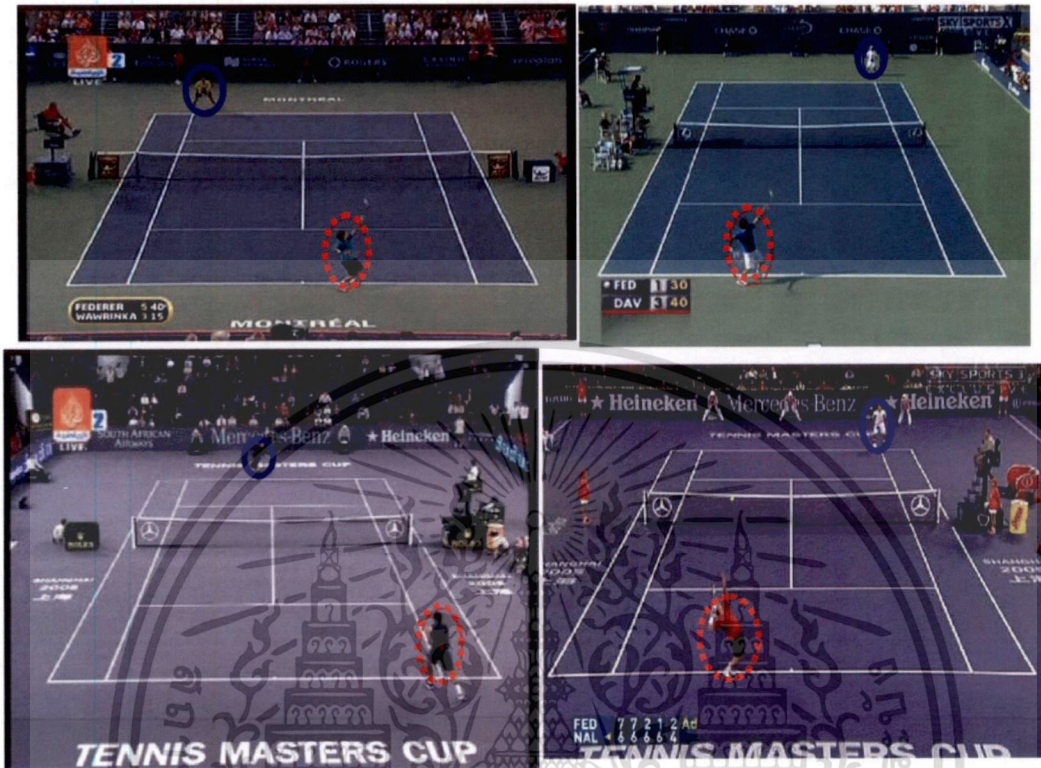
ในการทดลองนั้นต้องมีการกำหนดตำแหน่งของผู้เล่นเทนนิสทั้งฝั่งใกล้และไกล ตำแหน่งของ สนามเทนนิสในส่วนของกรติดตามหาเส้นสนามเทนนิส หรือขนาดของจุดภาพเพื่อใช้ในการหา ลูกเทนนิส รวมทั้งการกำหนดค่าต่างๆที่จะใช้ในการทดลอง

4.2.1 ขั้นตอนการติดตามผู้เล่นเทนนิส

ตำแหน่งที่จะกำหนดให้ในส่วนของการติดตามผู้เล่นเทนนิส จะกำหนดตำแหน่งที่อยู่ตรงกลางของผู้เล่นเทนนิสในเฟรมแรก เพื่อที่จะเป็นจุดศูนย์กลางของวงรีจากนั้นขนาดของวงรี ได้มาจากความกว้างและสูงของผู้เล่นเทนนิส ดังรูปที่ 4.3 วงรีสีน้ำเงิน (เส้นทึบ) และวงรีสีแดง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(เส้นประ) แสดงถึงบริเวณที่ได้ทำการกำหนดเพื่อใช้ในการติดตามผู้เล่นเทนนิสของคนไกล และไกลตามลำดับ



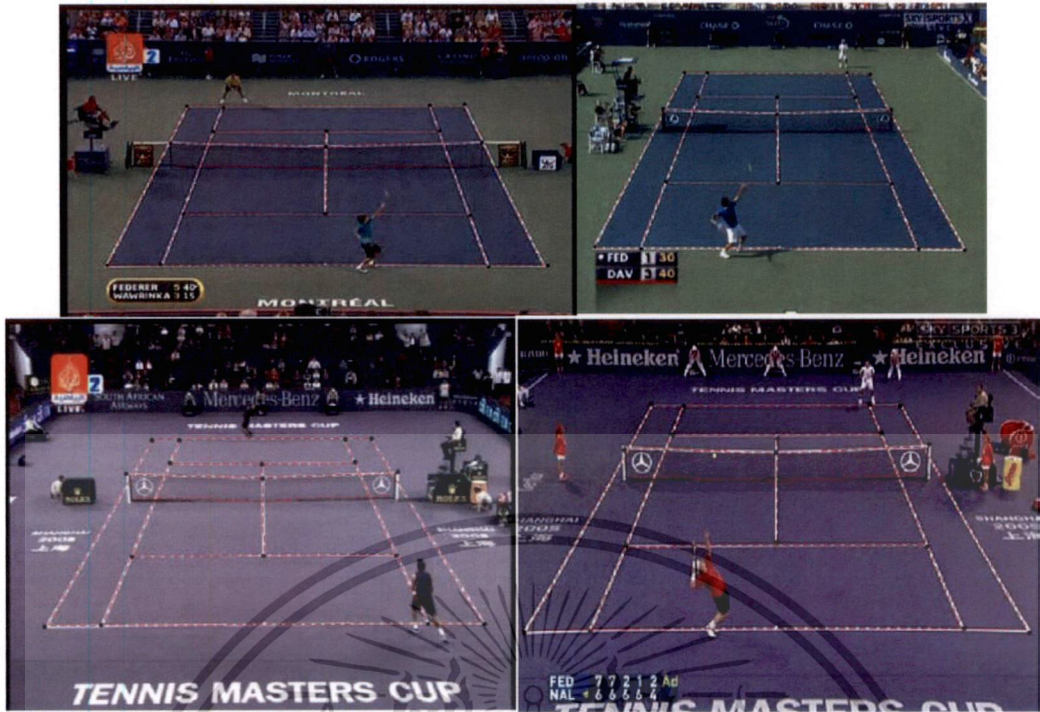
รูปที่ 4.3 ตัวอย่างของการกำหนดตำแหน่ง และขนาดเพื่อทำการติดตามผู้เล่นเทนนิส

จากในขั้นตอนดังรูปที่ 3.6 จะต้องมีการกำหนดค่าขีดเริ่มเปลี่ยน ที่ใช้ในขั้นตอนของการหาความแตกต่างระหว่างเฟรม เพื่อจะลดข้อผิดพลาดอันเกิดจากนากหลัง โดยค่าขีดเริ่มเปลี่ยนที่ใช้มีค่าเท่ากับ 30 และขนาดเมทริกซ์เพื่อเชื่อมต่อจุดภาพ ของการทำด้วยตัวดำเนินการแบบปิดคือ 9X9

4.2.2 ขั้นตอนการติดตามเส้นสนามเทนนิส

การตรวจหาเส้นสนามเทนนิสขั้นตอนแรกคือการกำหนด 17 จุดดังรูปที่ 3.8 ก ของในเฟรมแรกเพื่อที่จะได้ทราบถึงตำแหน่งของเส้น ดังรูปที่ 4.4 ที่จุดสีดำแสดงถึงการกำหนดจุดทั้ง 17 จุด จากนั้นดูความหนาของเส้นสนาม เพื่อเก็บเป็นค่าขีดเริ่มเปลี่ยน ของในขั้นตอนการหาเส้นสนาม ในหัวข้อที่ 3.2 ของขั้นตอนที่ 10

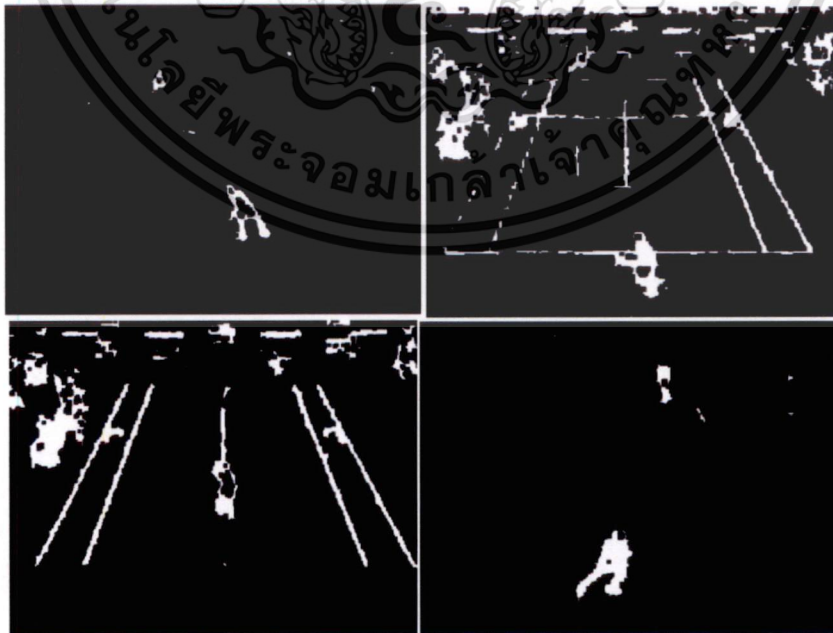
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 ตัวอย่างการกำหนด 17 จุดเพื่อใช้ในการหาเส้นสนามเทนนิส

4.2.3 ขั้นตอนการหาลูกเทนนิส

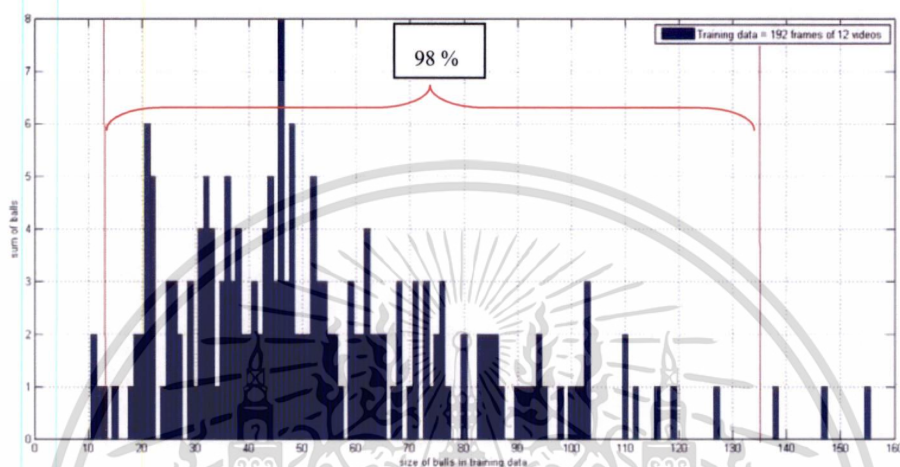
จากในหัวข้อ 3.3.1 เราใช้วิธีการหาความแตกต่างระหว่างเฟรม ซึ่งในการทดลองนี้เราใช้ค่าขีดเริ่มเปลี่ยนเท่ากับ 10 และทำมอร์ฟอโลยี ด้วยตัวดำเนินการแบบปิด โดยใช้ขนาดเมทริกซ์เท่ากับ 5X5 ดังรูปที่ 4.5



รูปที่ 4.5 ตัวอย่างของการหาดำแหน่งลูกเทนนิส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังที่ได้กล่าวไว้แล้วว่า เราใช้การกรองด้วยขนาดของจุดภาพ ดังนั้นขนาดที่ใช้ในการกรอง บล็อกเพื่อหาตำแหน่งของลูกเทนนิส จะได้มาจากการเทรนของชุดข้อมูล 192 เฟรมจากการสุ่ม จำนวน 12 คลิป โดยในแต่ละเฟรม จะทำการนับขนาดของจุดภาพที่ซึ่งเป็นตำแหน่งของลูกเทนนิส จากนั้นทำการหาช่วงของความเชื่อมั่นที่ 98% ซึ่งจะได้ผลลัพธ์ของการหาคือ 13-136 ที่เราจะใช้ค่าของความเชื่อมั่นนี้ในการกรองบล็อก

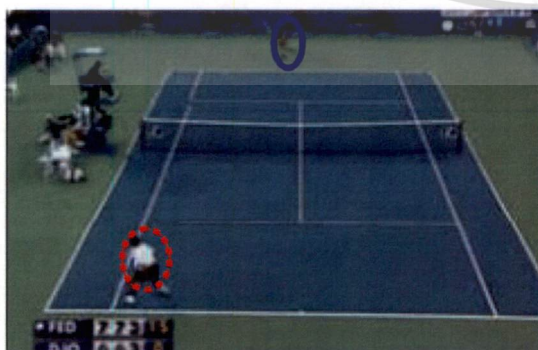


รูปที่ 4.6 กราฟแสดง ผลของการเทรนของชุดข้อมูล 192 เฟรม

ในแกนนอนของรูปที่ 4.6 แสดงถึงขนาดของจุดภาพ ส่วนในแกนตั้ง จะแสดงถึงผลรวมของจำนวนลูกเทนนิส และเส้นสีแดง แสดงถึงช่วงของความเชื่อมั่นที่ 98%

4.3 การทดลองและผลลัพธ์ของการติดตามผู้เล่น

โดยในส่วนนี้เราจะทำการแสดงผลการทดลอง จากการที่ได้ภาพที่จะนำมาเป็นอินพุตให้กับระบบ และกำหนดค่าต่างๆ ไปผ่านกระบวนการติดตามผู้เล่นเทนนิส



(ก)

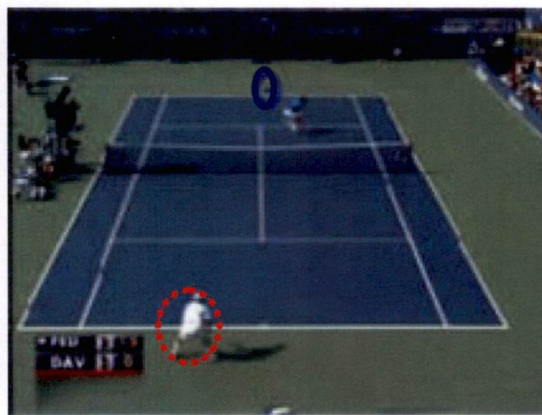


(ข)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ง)



(จ)



(ฉ)

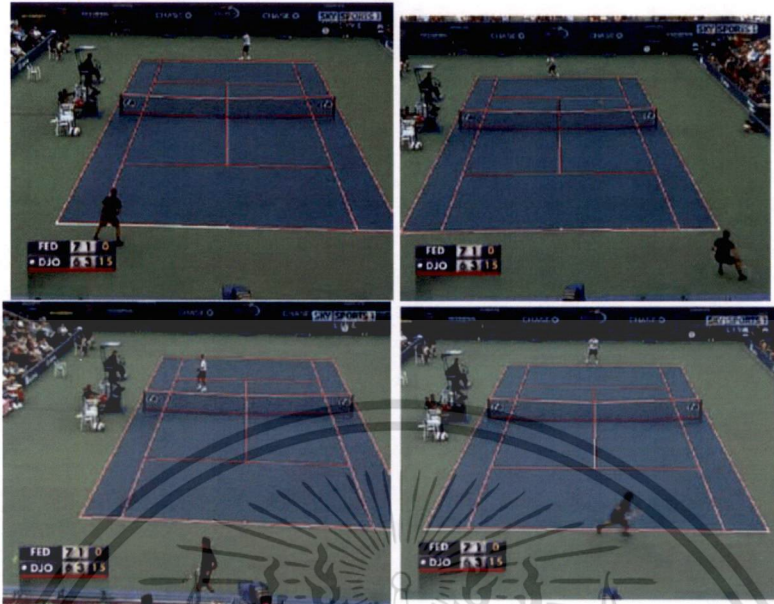
รูปที่ 4.7 (ก) (ค) (จ) ผลลัพธ์ของการติดตามโดยใช้วิธีย้ายค่าเฉลี่ย ร่วมกับวิธีการหาความแตกต่างระหว่างเฟรม (ข) (ง) (ฉ) ผลลัพธ์ของการติดตามโดยใช้วิธีย้ายค่าเฉลี่ยอย่างเดียว

จากรูปที่ 4.7 (ข) ผลลัพธ์ที่ได้จากการติดตามโดยที่ไม่ใช้วิธีหาความแตกต่างระหว่างเฟรมมารวมด้วยจะทำให้เกิดการผิดพลาด เพราะ ผู้เล่นเทนนิสฝั่งโกล์นั้น ได้สวมเสื้อที่เป็นสีขาวที่มีสีเหมือนกับป้ายบอกคะแนน จึงทำให้ระบบสับสนได้ จากรูปที่ 4.7 (ง) เกิดความผิดพลาดเพราะผู้เล่นเทนนิสฝั่งโกล์สวมเสื้อสีเดียวกับสนามเทนนิส และจากรูปที่ 4.7 (ฉ) ผู้เล่นฝั่งโกล์สวมเสื้อสีดำซึ่งเป็นสีเดียวกับกรรมการกำกับเส้น ซึ่งทั้งหมดนี้สามารถแก้ปัญหาก็ได้โดยวิธีการในบทที่ 3 หัวข้อ 3.1 คือการใช้การติดตามโดยใช้วิธีย้ายค่าเฉลี่ย ร่วมกับวิธีการหาความแตกต่างระหว่างเฟรม จะได้รูปที่ 4.7 (ก) (ค) (จ)

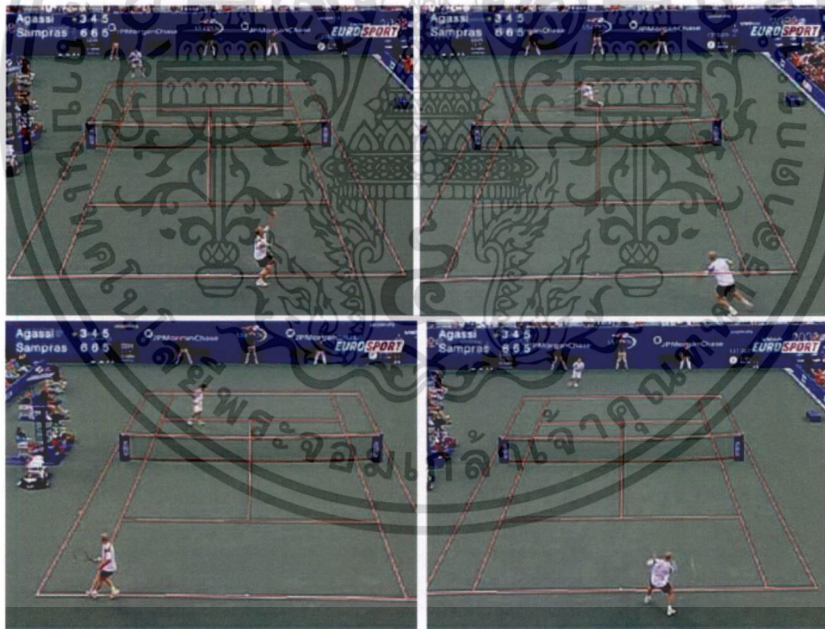
4.4 การทดลองและผลลัพธ์ของการติดตามเส้นสนามเทนนิส

ทำการทดลองและกำหนดค่าอินพุตให้กับระบบเพื่อหาเส้นสนามเทนนิส จะได้ผลลัพธ์ตัวอย่างดังรูปที่ 4.8 (ก) (ข) (ค) ซึ่งในแต่ละรูปคือเฟรมของวิดีโอทัศน์ และเส้นสีแดงคือผลลัพธ์ที่ได้จากระบบ โดยสามารถเห็นได้ว่าระบบทำงานได้อย่างถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

รูปที่ 4.8 (ก) (ข) ผลลัพธ์ที่ได้จากการติดตามหาเส้นสนามเทนนิส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ค)

รูปที่ 4.8 (ค) ผลลัพธ์ที่ได้จากการติดตามหาเส้นสนามเทนนิส (ต่อ)

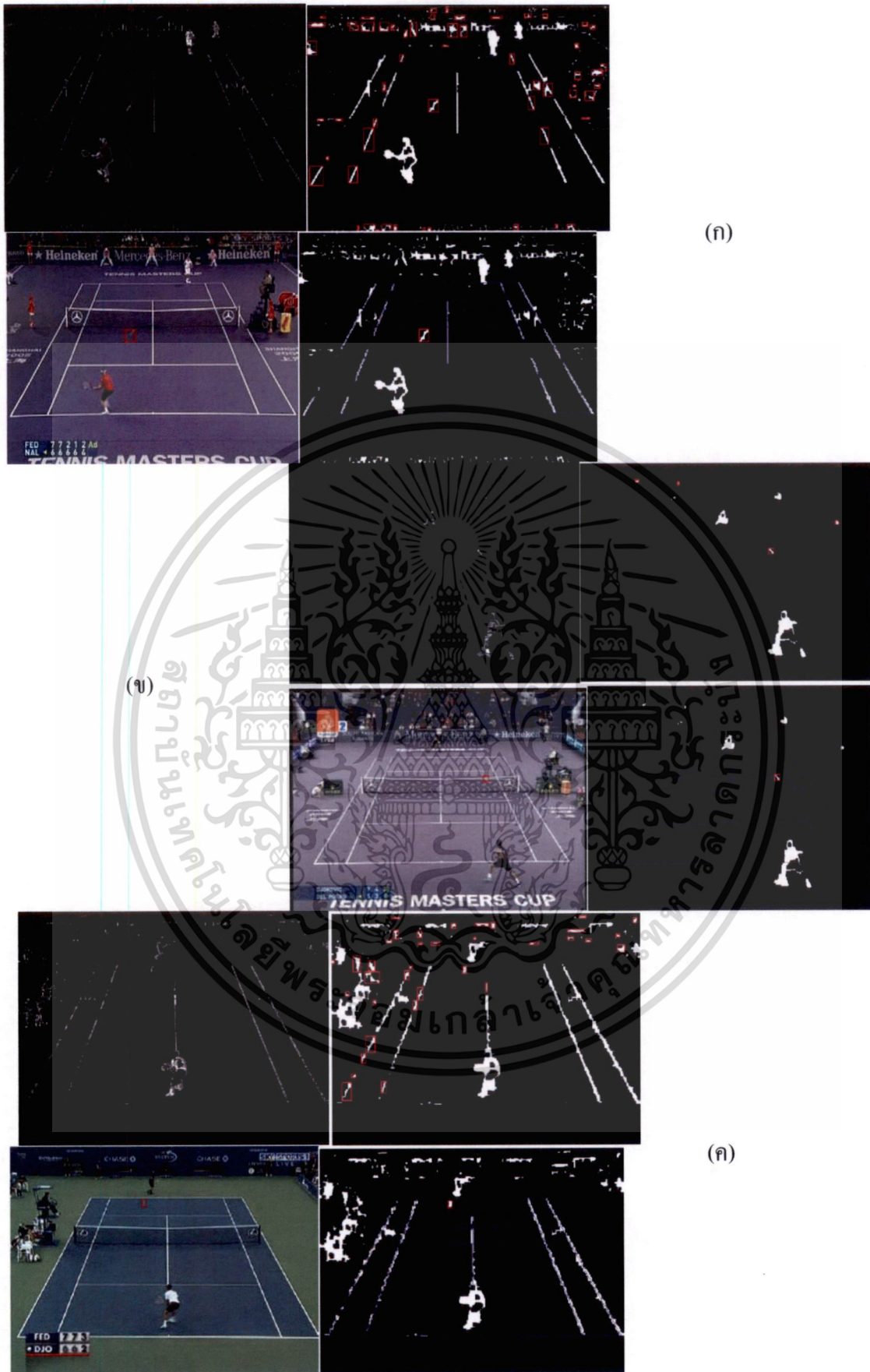
4.5 การทดลองและผลลัพธ์ของการตรวจจับลูกเทนนิส

ในส่วนนี้จะเป็นการแสดงผลการทดลองการตรวจจับลูกเทนนิส โดยใช้ผลลัพธ์ที่ได้จากการหาคำแหน่งผู้เล่น และสนามเทนนิส

จากรูปที่ 4.9 กรอบทางซ้ายบนคือผลลัพธ์ที่ได้จากการทำวิธีการหาความแตกต่างระหว่างเฟรม กรอบทางขวาบนคือผลลัพธ์ที่ได้จากการกรองบลิบ กรอบทางขวาล่างเป็นส่วนที่ได้จากการกรองบลิบและลดปริมาณความผิดพลาด และกรอบทางซ้ายล่างคือผลลัพธ์จากการตรวจจับลูกเทนนิส

ในรูปที่ 4.9 จะเห็นได้ว่า ถ้าระบบใช้วิธีการกรองโดยใช้ช่วงขนาดบลิบเพียงอย่างเดียว จะทำให้ได้ผลลัพธ์ที่เป็นลูกเทนนิสเป็นจำนวนมาก ซึ่งส่วนมากแล้วบลิบที่ทำให้เกิดความผิดพลาด คือ บริเวณของเส้นสนามและฉากหลัง จึงแก้ปัญหาดังกล่าวด้วยวิธีในหัวข้อ 3.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

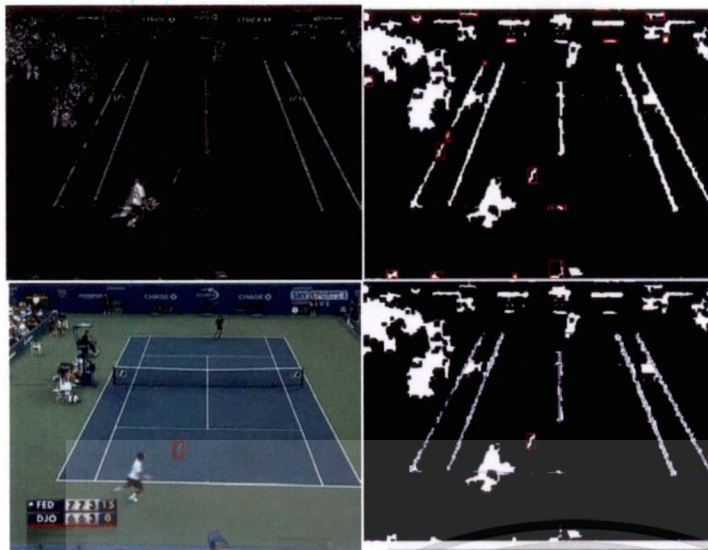


(ข)

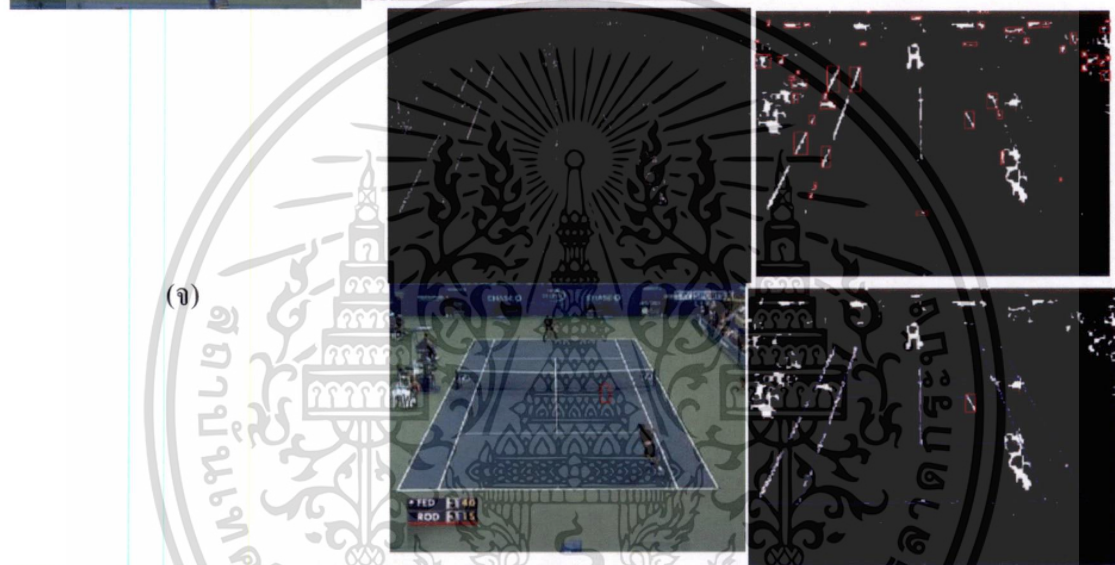
(ก)

(ค)

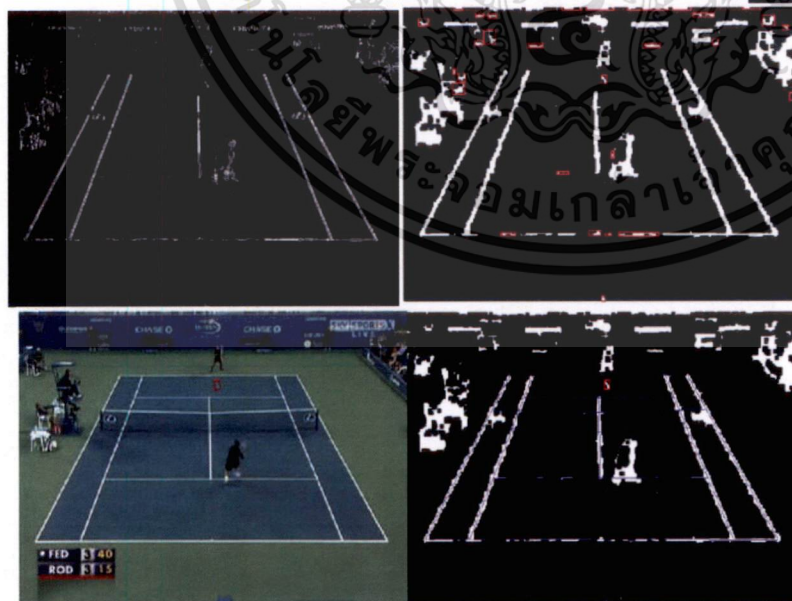
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของ บริษัท เทคโนโลยีการกีฬาแห่งประเทศไทย ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ง)



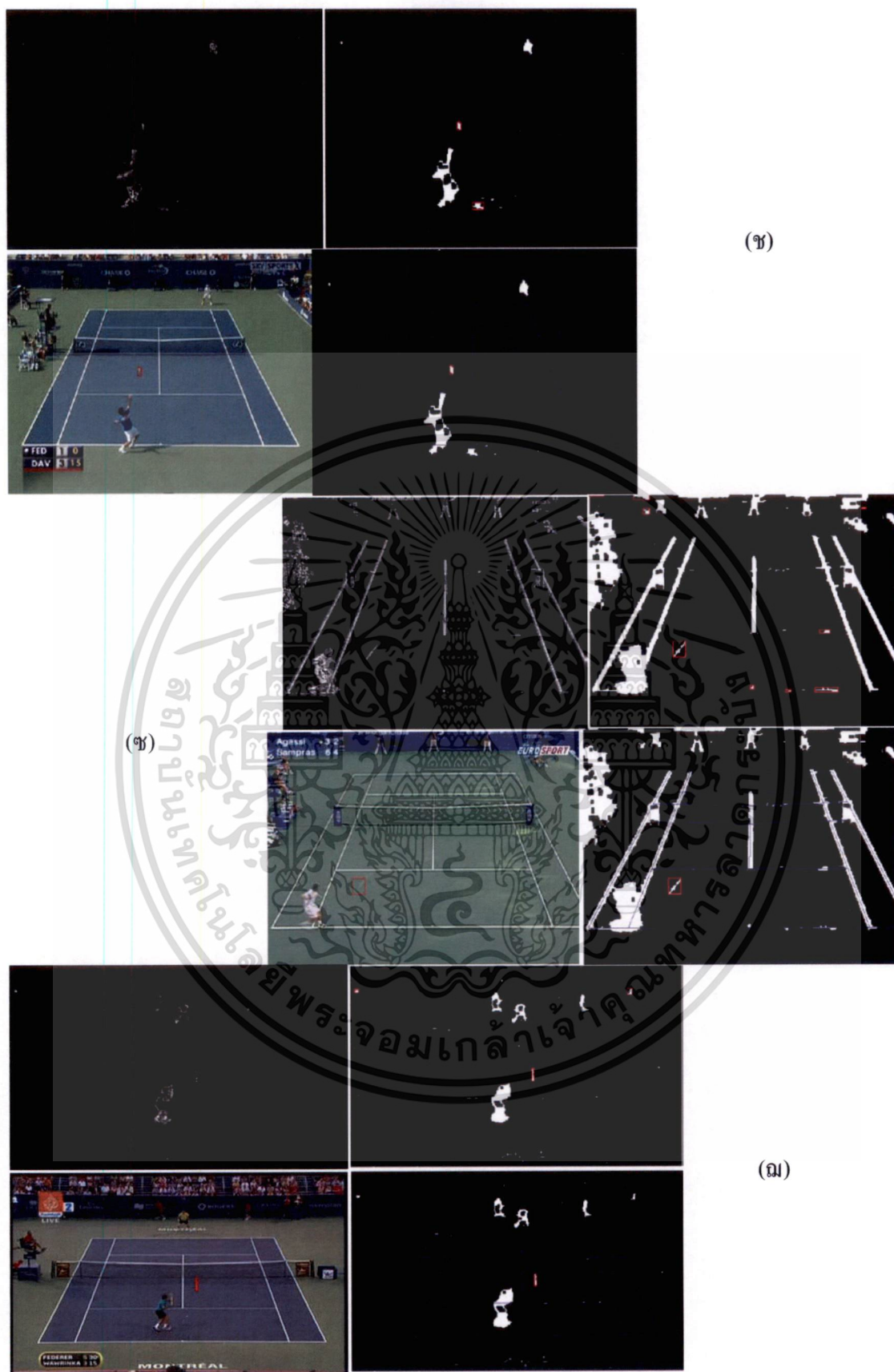
(จ)



(ฉ)

รูปที่ 4.9 (ง),(จ),(ฉ), ผลลัพธ์ที่ได้จากการติดตามลูกเทนนิส (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 (ข),(ค),(ง) ผลลัพธ์ที่ได้จากการติดตามลูกเทนนิส (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6 การทดลองและผลลัพธ์ของการจำแนกท่าตีเทนนิส

หลังจากที่ได้ตำแหน่งของผู้เล่นเทนนิส สนามเทนนิส และลูกเทนนิส เราจะนำผลลัพธ์ที่ได้ไปผ่านขบวนการจำแนกท่าตีเทนนิส โดยผลการทดลองในวิทยานิพนธ์ฉบับนี้ สามารถแยกออกได้เป็น 3 ส่วน โดยมีตาราง 4.1 แสดงถึงข้อมูลจริง (Ground truth) ของวิดีโอทั้งหมดที่ได้นำมาทำการทดสอบ โดยในการทดสอบนี้ใช้จำนวนเฟรมทั้งหมด 12,498 เฟรม

1. ผลลัพธ์ของลูกเทนนิสเข้าใกล้ผู้เล่น เป็นการหาเฟรมที่ลูกเทนนิสเข้าใกล้ผู้เล่นเทนนิส ซึ่งผลลัพธ์ที่ได้จากในขั้นตอนนี้จะนำไปใช้ในการจำแนกท่าตีเทนนิส ดังตาราง 4.2 โดยที่ผลจากการเสิร์ฟ ระบบจะไม่ทำการนับว่ามีลูกเทนนิสเข้าใกล้ผู้เล่น
2. ผลลัพธ์ของการแบ่งประเภทการตี หลังจากที่ได้ตำแหน่งที่ลูกเทนนิสเข้าใกล้ผู้เล่นเทนนิสแล้ว เราจะทำการจำแนกท่าตีเทนนิสออกเป็น 5 คลาส ได้แก่ เสิร์ฟ ท่าตีโฟร์แฮนด์ ท่าตีแบคแฮนด์ ท่าตีโฟร์แฮนด์หน้าเน็ต ท่าตีแบคแฮนด์หน้าเน็ต ดังตาราง 4.3
3. ตาราง Confusion Matrix เป็นการตรวจสอบเพื่อดูความถูกต้องของในแต่ละท่าตีเทนนิส ที่ได้ทำการจำแนก ดังตาราง 4.4

จากตารางที่ 4.1-4.4 SERVE = เสิร์ฟ, FH = ท่าตีโฟร์แฮนด์, BH = ท่าตีแบคแฮนด์, FH NET = ท่าตีโฟร์แฮนด์ หน้าเน็ต, BH NET = ท่าตีแบคแฮนด์ หน้าเน็ต

ตารางที่ 4.1 กราวด์ทรูธ

กราวด์ทรูธ									
ผู้เล่นไกล (Far)					ผู้เล่นใกล้ (Near)				
SERVE	FH	BH	FH NET	BH NET	SERVE	FH	BH	FH NET	BH NET
11	86	68	11	8	22	65	84	5	6

ตารางที่ 4.2 ผลลัพธ์ของการตีลูกเทนนิส

ผู้เล่นเทนนิส	จำนวนการตี	ตรวจพบการตี (True Positive)	ตรวจไม่พบการตี (False Negative)	ตรวจจับผิดพลาด (False Positive)
ผู้เล่นไกล	173	163 (94.22)%	10 (5.78)%	4
ผู้เล่นใกล้	160	158 (98.75)%	2 (1.25)%	2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 ผลลัพธ์ของการจำแนกท่าตีเทนนิส

	จำนวนท่าตี	TP	FN	FP
ผู้เล่นไกล				
SERVE	11	11 (100.00)%	0 (0)%	0
FH	86	74 (86.05)%	12 (13.95)%	6
BH	68	61 (89.70)%	7 (10.30)%	10
FH NET	11	5 (45.45)%	6 (54.55)%	1
BH NET	8	5 (62.50)%	3 (37.50)%	1
ผู้เล่นใกล้				
SERVE	22	22 (100.00)%	0 (0)%	0
FH	65	63 (96.92)%	2 (3.08)%	9
BH	84	75 (89.28)%	9 (10.72)%	2
FH NET	5	5 (100.00)%	0 (0)%	0
BH NET	6	4 (66.67)%	2 (33.33)%	0

ตารางที่ 4.4 Confusion Matrix

input/output	SERVE	FH	BH	FH NET	BH NET	LOST	SUM	% ความ ถูกต้อง
SERVE	33	0	0	0	0	0	33	100.00
FH	0	137	12	0	0	2	151	90.73
BH	0	15	136	0	0	1	152	89.47
FH NET	0	0	0	10	1	5	16	62.50
BH NET	0	0	0	1	9	4	14	64.29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

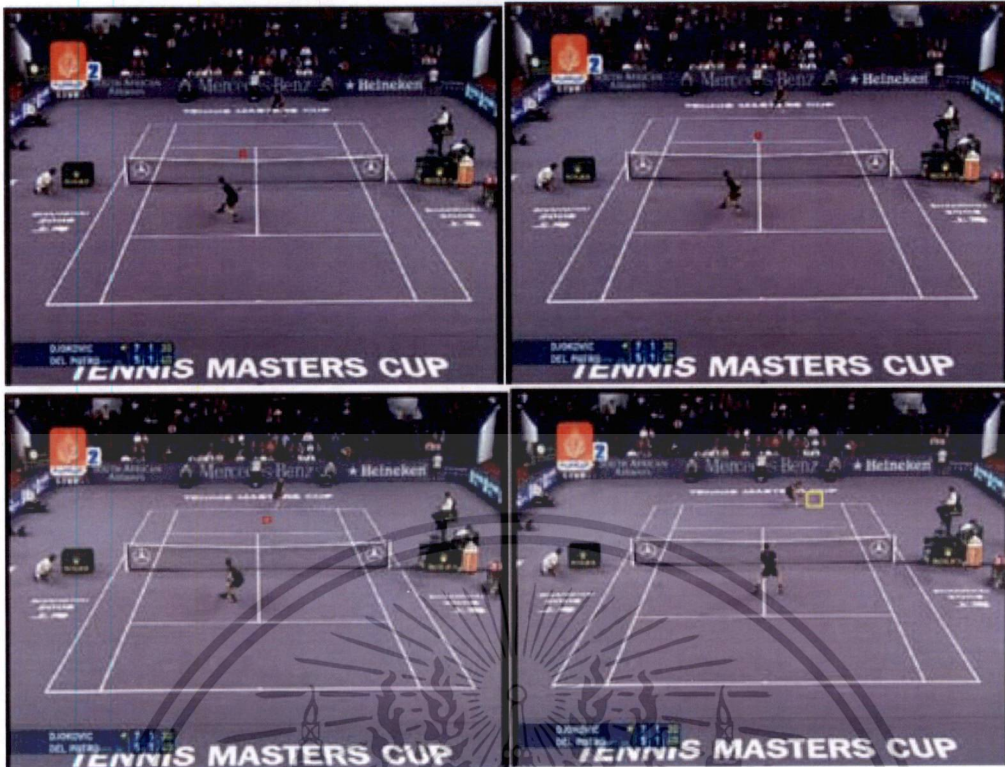
ผลการทดลองในตารางที่ 4.2 สามารถเห็นได้ว่าจำนวน กราวด์ทรูชของลูกเทนนิสที่เข้าใกล้ ผู้เล่นไกลและใกล้คือ 173, 160 ครั้ง ระบบมีความถูกต้องของการตรวจพบลูกเทนนิสเข้าใกล้ผู้เล่น 163, 158 ครั้ง คิดเป็นความถูกต้องของในการตรวจสอบถึง 94.22 % และ 98.75% โดยที่ตรวจไม่พบการตี 10, 2 ครั้ง และตรวจจับผิดพลาด 4, 2 ครั้ง ของผู้เล่นไกลและใกล้

จากนั้นนำผลลัพธ์ที่ถูกต้อง จากตารางที่ 4.2 มาทำการจำแนกท่าตี เป็นท่า โฟร์แฮนด์, แบคแฮนด์, โฟร์แฮนด์หน้าเน็ต และแบคแฮนด์หน้าเน็ต ซึ่งมีความถูกต้องของผู้เล่นไกล (ทรูโพสซิทีฟ) 74, 61, 5, 5 ครั้ง ส่วนผู้เล่นใกล้ 63, 75, 5, 4 ครั้ง เกิดการจำแนกท่าตีผิดพลาด (ฟอลส์โพสซิทีฟ) ในส่วนของผู้เล่นไกล 6, 10, 1, 1 ครั้ง และผู้เล่นใกล้ 9, 2, 0, 0 โดยที่ระบบตรวจไม่พบการจำแนก (ฟอลส์เนกาทีฟ) ของผู้เล่นไกล 12, 7, 6, 3 ครั้ง และผู้เล่นใกล้ 2, 9, 0, 2 ครั้ง ทั้งผู้เล่นไกลและใกล้ มีค่าทรูโพสซิทีฟ ของการเสิร์ฟได้แก่ 11 และ 22 ครั้ง ตามลำดับ โดยที่ค่าฟอลส์โพสซิทีฟ และ ฟอลส์เนกาทีฟมีค่าเป็น 0

จากในตารางที่ 4.4 จะทำการพิจารณาค่าฟอลส์เนกาทีฟของในตารางที่ 4.3 ซึ่งผลรวมของท่าตีโฟร์แฮนด์ ผู้เล่นไกลรวมกับใกล้ คือ 14 ครั้ง ระบบมีความผิดพลาด โดยเข้าใจผิดว่าเป็น ท่าตีแบคแฮนด์ 12 ครั้ง กับระบบตรวจสอบไม่เจอท่าตี 2 ครั้ง พิจารณาท่าตีแบคแฮนด์ ผู้เล่นไกลรวมกับใกล้ คือ 16 ครั้ง ระบบมีความผิดพลาดโดยเข้าใจผิดว่าเป็น ท่าตีโฟร์แฮนด์ 15 ครั้ง กับระบบตรวจสอบไม่เจอท่าตี 1 ครั้ง จากนั้นพิจารณาท่าตีโฟร์แฮนด์หน้าเน็ต ของผู้เล่น ไกลรวมกับใกล้ คือ 6 ครั้ง ซึ่งระบบมีความผิดพลาด โดยเข้าใจผิดว่าเป็น ท่าตีแบคแฮนด์หน้าเน็ต 1 ครั้ง กับระบบตรวจสอบไม่เจอท่าตี 5 ครั้ง สุดท้ายพิจารณาท่าตีแบคแฮนด์หน้าเน็ต ผู้เล่นไกลรวมกับใกล้ คือ 5 ครั้ง ซึ่งระบบมีความผิดพลาด โดยเข้าใจผิดว่าเป็น ท่าตีโฟร์แฮนด์หน้าเน็ต 1 ครั้ง กับระบบตรวจสอบไม่เจอท่าตี 4 ครั้ง

จากผลการทดลองทำให้เห็นได้ว่า ผลของการจำแนกท่าตีเทนนิสของผู้เล่นฝั่งใกล้ส่วนมากนั้น จะมีความถูกต้องมากกว่าผู้เล่นฝั่งไกล เพราะความละเอียด (resolution) ของสนามเทนนิสทางด้านผู้เล่นฝั่งใกล้มีมากกว่า จึงทำให้ได้ตำแหน่งของผู้เล่นเทนนิสและบริเวณการเคลื่อนที่ของลูกเทนนิสมากกว่าทางผู้เล่นเทนนิสฝั่งไกล ซึ่งระบบจะมีความถูกต้องสูง ในกรณีที่ผู้เล่นทั้งสองฝั่ง อยู่บริเวณท้ายของสนามเทนนิส เพราะว่าระบบจะสามารถเห็นแนววิถีของลูกได้ดี แต่ถ้าผู้เล่นทั้งสองฝั่งอยู่บริเวณหน้าเน็ตจะทำให้ระบบเห็นแนววิถีลูกได้ยาก จึงส่งผลให้ การตรวจสอบลูกเทนนิสเข้าใกล้ผู้เล่นไม่พบ

จากรูปที่ 4.10-4.11 เป็นภาพตัวอย่างของในกรณีที่จำแนกท่าตีผิดพลาดและตรวจไม่พบการจำแนกท่าตี ซึ่งคือค่าฟอลส์โพสซิทีฟและค่าฟอลส์เนกาทีฟ ของตารางที่ 4.3

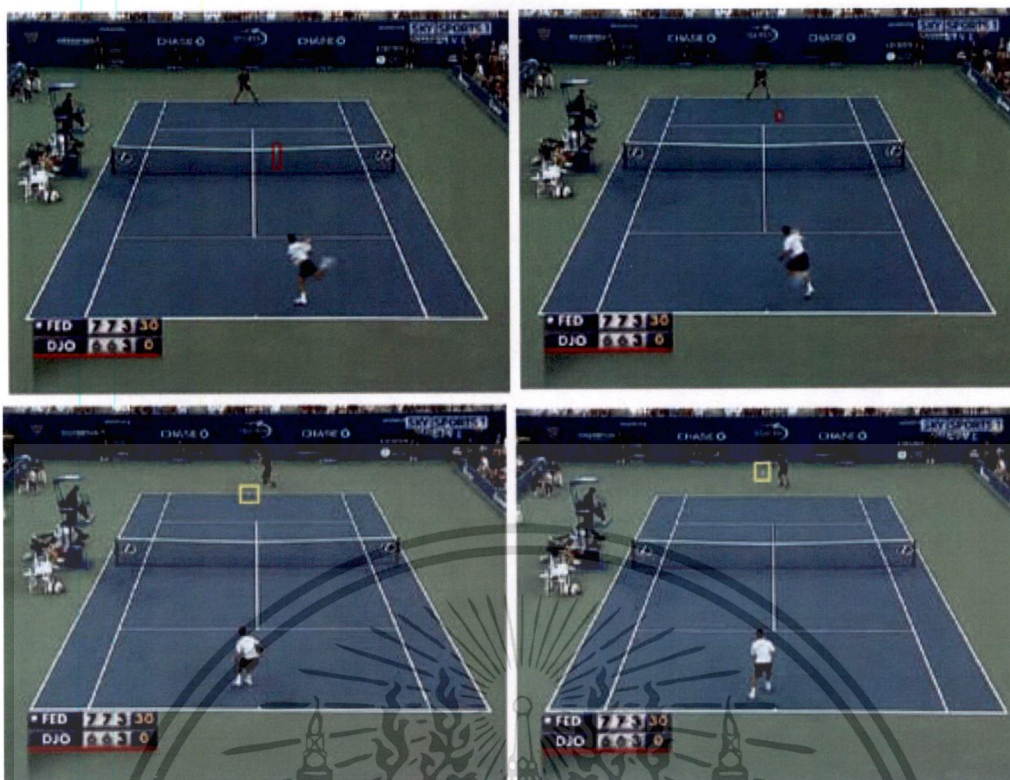


(ก)

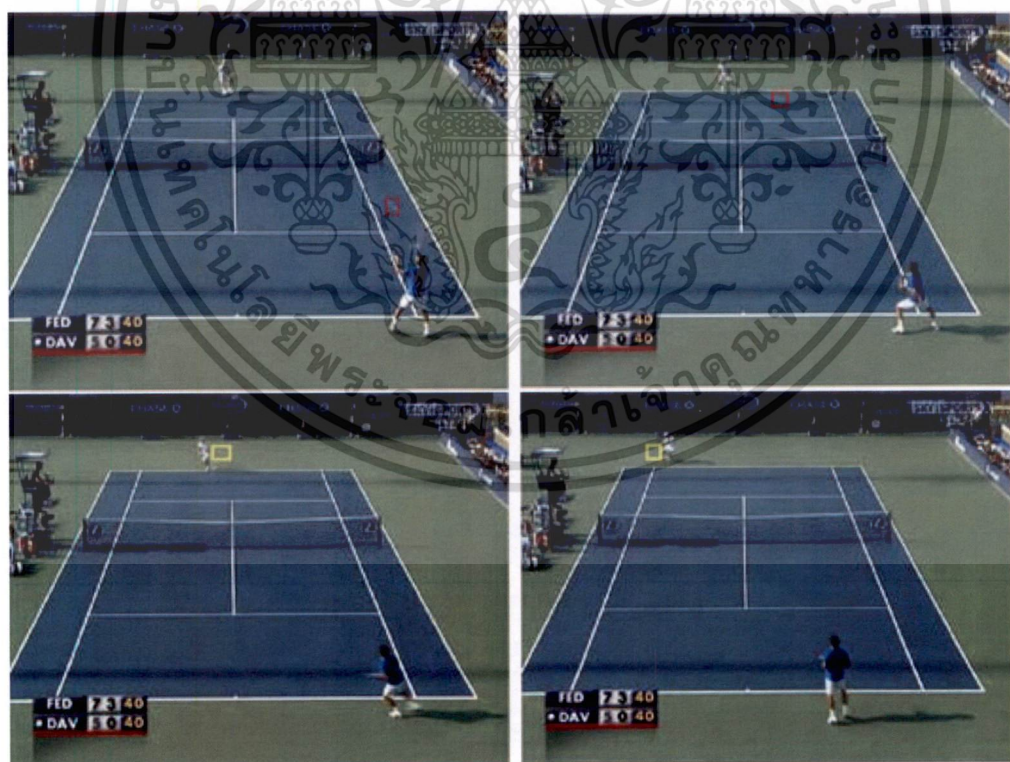


(ข)

รูปที่ 4.10 ภาพตัวอย่างกรณีจำแนกทำผิด (ก),(ข) ลูกเทนนิสเคลื่อนที่อย่างรวดเร็วและประชิด
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ค)



(ง)

รูปที่ 4.10 ภาพตัวอย่างกรณีงานแรกทำผิดพลาด (ค),(ง) ลูกเทนนิสเคลื่อนที่อย่างรวดเร็วและประชิด (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

รูปที่ 4.11 ภาพตัวอย่างกรณีจำแนกทำผิดผิด (ก), (ข) ลูกเทนนิสเคลื่อนที่อยู่นอกสนามเทนนิส

รวมทั้งหายไปจากเฟรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในรูปที่ 4.10-4.11 กรอบสี่แดงแสดงถึงระบบสามารถตรวจพบลูกเทนนิส ส่วนกรอบสี่เหลืองเป็นตำแหน่งลูกเทนนิสที่ระบบไม่สามารถตรวจพบได้ โดยที่มีลำดับของภาพเริ่มจากกรอบบนซ้าย, บนขวา, ล่างซ้าย, ล่างขวา

ส่วนของการจำแนกท่าตีที่ผิดพลาดสามารถเกิดขึ้นดังตัวอย่างที่ 4.10 ได้หลายปัจจัย เช่น รูปที่ 4.10 (ก) (ข) นั้นในกรอบบนซ้ายจะเห็นได้ว่าตำแหน่งของผู้เล่นเทนนิสอยู่ที่หน้าเน็ต และได้ทำการตีลูกเทนนิสระบบสามารถตรวจพบลูกเทนนิสได้จนถึงกรอบล่างซ้ายจึงได้จำแนกท่าตี แต่เมื่อลูกเทนนิสตกกระทบสนาม เป็นผลให้แนววิถีเปลี่ยน ท่าตีที่ได้จำแนกไปตอนต้นนั้นเลยผิด ซึ่งผนวกกับระยะห่างของผู้เล่นมีความใกล้เคียงอีกด้วย ส่วนในกรณีรูปที่ 4.10 (ค) (ง) ความเร็วการเคลื่อนที่ของลูกเทนนิสในกรอบบนซ้ายและขวามีมาก โดยในกรอบล่างซ้ายเมื่อลูกเทนนิสตกในสนามเทนนิส มีระยะประชิดกับผู้เล่นมากจึงทำให้ไม่สามารถตรวจพบลูกเทนนิส ในการจำแนกเลยต้องใช้ตำแหน่งของลูกเทนนิสที่ได้จากกรอบบนซ้ายแทน จึงทำให้เกิดความผิดพลาด ดังนั้นสาเหตุที่ทำให้การจำแนกท่าตีผิดพลาด จากตัวอย่าง คือระยะวิถีของลูกเทนนิสสั้นหรือ ลูกเทนนิสเคลื่อนที่อย่างรวดเร็วและประชิดตัวของผู้เล่น

กรณีในรูปที่ 4.11 (ก) และ (ข) ลูกเทนนิสมีการเคลื่อนที่อยู่นอกสนามเทนนิสและทับเส้นสนามเทนนิส ทำให้ระบบตรวจสอบไม่พบลูกเทนนิส จึงทำให้ตรวจไม่พบท่าตีเทนนิส

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้เราได้นำเสนอการจำแนกท่าตีกีฬาเทนนิสจากวิดีโอ ที่ใช้มุมมองระยะไกลเพียงอย่างเดียว ซึ่งสามารถแบ่งเป็นส่วนของกระบวนการได้ 5 ขั้นตอนหลัก ได้แก่ การติดตามผู้เล่นเทนนิส การหาเส้นของสนามเทนนิส การตรวจหาลูกเทนนิส การลดความผิดพลาดของการตรวจหาลูกเทนนิส และการจำแนกท่าตี โดยที่การติดตามหาตำแหน่งผู้เล่นเทนนิส อาจเกิดปัญหาในกรณีที่ผู้เล่นเทนนิสสวมชุดเหมือนกรรมการ หรือสีของชุดเป็นสีเดียวกับสนาม เพื่อแก้ปัญหาในงานวิจัยนี้ได้ใช้อัลกอริทึม การติดตามโดยวิธีย้ายค่าเฉลี่ย ร่วมกับ ความแตกต่างระหว่างเฟรม อีกทั้งยังใช้ปริภูมิสี RGB เพื่อเพิ่มความถูกต้องของการติดตาม การหาเส้นของสนามเทนนิสใช้ความรู้ที่ว่าเส้นของสนามเทนนิสมีสีขาว จึงทำให้งานวิจัยนี้ใช้ค่าความเข้มของจุดภาพที่สูง เพื่อประโยชน์ในการติดตามและหาเส้นสนาม ผนวกกับคุณสมบัติเส้นสนามเทนนิสเป็น แบบเส้นตรงจึงใช้การถดถอยเชิงเส้นเข้ามาาร่วมด้วย นอกจากนี้การหาตำแหน่งของลูกเทนนิสได้ใช้วิธีการหาความแตกต่างระหว่างเฟรม จากนั้นทำมอร์ฟอโลยี ด้วยตัวดำเนินการแบบปิด เพื่อทำการลบรอยขีด และเชื่อมต่อบริเวณที่เป็นช่องว่าง สุดท้ายกรองบิลเล็บโดยใช้ขนาด ผลลัพธ์ของการหาตำแหน่งลูกเทนนิสที่ได้ออกมา ยังได้ส่วนที่มีความผิดพลาดมาก ซึ่งเกิดจากเส้นสนามและบริเวณฉากหลัง ทำให้ต้องมีการใช้เทคนิคบางอย่างเพื่อกำจัดสิ่งที่ไม่ใช่ลูกเทนนิส ซึ่งเทคนิคนั้นคือการกรองบิลเล็บที่อยู่ในสนามเทนนิส พร้อมทั้งการนำข้อมูลเส้นสนามที่ได้จากการทำก่อนหน้ามามาทำการลบ รวมถึงบิลเล็บที่อยู่ในผู้เล่นเทนนิสอีกด้วย สุดท้ายในการจำแนกท่าตีเทนนิส ได้ใช้ข้อมูลการเคลื่อนที่ของผู้เล่นเทนนิสกับตำแหน่งของลูกเทนนิสมาทำการสร้างกราฟจำนวน 2 กราฟ โดยกราฟแรกในแนวตั้งกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรมส่วนกราฟที่สองประกอบไปด้วย แนวตั้ง คือ ตำแหน่งพิกัด x ของในภาพ ส่วนในแนวนอน คือ ลำดับของเฟรม จากนั้นใช้กราฟแรกเพื่อทำการหาช่วงเวลาที่ถูกเทนนิสเข้าใกล้ผู้เล่น จากนั้นนำช่วงที่ได้ไปพิจารณาว่าอยู่ทางด้านใดของผู้เล่นโดยใช้กราฟที่สอง ถ้าลูกเทนนิสอยู่ทางด้านขวาของผู้เล่นฝั่งใกล้ผลลัพธ์ของการตีที่ได้คือ โฟร์แฮนด์ แต่ถ้าอยู่ทางด้านซ้ายผลลัพธ์ของการตีที่ได้คือ แบคแฮนด์ ส่วนถ้าเป็นผู้เล่นฝั่งไกลผลลัพธ์ของการตีที่ได้คือ แบคแฮนด์ และ โฟร์แฮนด์ตามลำดับ จากนั้นพิจารณาถึงตำแหน่งจุดยืนของผู้เล่นในสนามเทนนิส เพื่อทำการจำแนกต่อว่าเป็นการตีหน้าเน็ตหรือไม่ วิเคราะห์ผู้เล่นเทนนิสฝั่งใดที่มีการเสิร์ฟโดยสังเกตได้จากรูปทรงของกราฟ สรุปผลลัพธ์จากการจำแนกที่ได้มีทั้งหมด 5 คลาสซึ่งได้แก่ เสิร์ฟ แบคแฮนด์ โฟร์แฮนด์ แบคแฮนด์หน้าเน็ต โฟร์

แฮนด์หน้าเน็ต ที่ซึ่งวิธีการในงานวิจัยของเราสามารถจำแนกท่าตีเทนนิส ได้ทั้งผู้เล่นไกลและใกล้ โดยเป็นการใช้กล้องเพียง 1 ตัวและสามารถจำแนกท่าตีได้โดยไม่ต้องทำการเรียนรู้

เมื่อพิจารณาถึงการจำแนกท่าตีที่ผิดพลาด อันเนื่องมาจากการขัดต่อสมมติฐาน เช่นกรณีที่ถูกเทนนิสพุ่งเข้าหาผู้เล่นเทนนิสฝั่งใกล้ทางด้านซ้าย จากสมมติฐานต้องเป็นการจำแนกท่าตีได้เป็นท่าแบ็คแฮนด์ แต่ผู้เล่นเทนนิสต้องการตีลูกกลับด้วยความแรงและรวดเร็ว จึงได้ทำการเอี้ยวตัวเพื่อทำการตีท่าโฟร์แฮนด์ หรือกรณีที่ผู้เล่นเทนนิสมีความถนัดในการจับไม้ทั้ง 2 ข้างจึงทำให้ระบบผิดพลาดได้

5.2 ข้อเสนอแนะ

ในการกำหนดบริเวณผู้เล่นเทนนิสควรให้ครอบคลุมแค่เพียงชุดของผู้เล่นเทนนิส เพื่อความถูกต้องและประหยัดเวลาในการรันโปรแกรม อีกทั้งในส่วนของการเลือกใช้ค่าขีดเริ่มเปลี่ยน ของขั้นตอนการหาลูกเทนนิสควรจะสามารถเห็นการเคลื่อนที่ของลูกเทนนิสได้ สุดท้ายควรหาตัวอย่างการแข่งขันกีฬาเทนนิส ที่มีการตีหน้าเน็ตมาทดลองเพิ่ม

ส่วนในการพัฒนาขั้นต่อไปอาจจะมีการนำรูปร่างของผู้เล่นเทนนิส เพื่อช่วยในเรื่องของการจำแนกท่าตี หรือช่วยในการหาตำแหน่งเริ่มต้นของผู้เล่นเทนนิส ซึ่งในงานวิจัยนี้กำหนดตำแหน่งด้วยมือ ในส่วนของการหาเส้นสนามเทนนิส อาจจะนำวิธีการหาขอบเพื่อใช้ในการกำหนดจุดเริ่มต้น นอกจากนี้ควรมีการปรับปรุงวิธีการหาลูกเทนนิส เพื่อให้ครอบคลุมในกรณีที่ลูกเทนนิสอยู่นอกสนามเทนนิสหรือหายไปจากเฟรม รวมถึงกรณีที่ลูกเส้นสนามเทนนิสบัง ในเรื่องของการจำแนกท่าตีอาจจะมีการเพิ่มคลาสบอลเลย์และคลาสลูกตาย โดยใช้นุ่มกล้องเพิ่มเติมเพื่อที่จะสามารถสังเกตถึงการตกกระทบระหว่างลูกเทนนิสกับสนามเทนนิส

เอกสารอ้างอิง

- [1] 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, Sep 2006.
- [2] H. Shah, P. Chokalingam, B. Paluri, N. Pradeep, and R. Balasubramanian. "Automated stroke classification in tennis". In **ICIAR07**, 2007, pages 1128-1137.
- [3] M.J. Black and P. Anandan. "A framework for the robust estimation of optical flow." in **ICCV '93**, May 1993, pp. 231–236.
- [4] Y. Saraf, R. Balasubramanian and K. Swaminathan. "A classical approach for thinning of binary images using divergence of the potential field." **International Journal of Computer Mathematics**, Vol. 82, No.6, 2005, pp. 673-684.
- [5] M. Petkovic, W. Jonker, and Z. Zivkovic. "Recognizing strokes in tennis videos using Hidden Markov Models." **Proc. of IASTED VIIP**, 2001.
- [6] 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.
- [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] M. Xu, N. C. Maddage, C. Xu, M. Rankanhalli, and Q. Tian. "Creating audio keywords for event detection in soccer video." In **Proc. Of ICME 2003**, Vol.2, pp281-284.
- [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] C.O. Conaire, P. Kelly, D. Connaghan, and N.E. O'Connor, "TENNISENSE: A platform for extracting semantic information from multi-camera tennis data." **Proc. Intl Conf. on Digital Signal Processing (DSP'09)**, 2009.
- [11] B.W. Silverman. **Density Estimation for Statistics and Data Analysis**. New York: Chapman And Hall, 1986
- [12] ปิยะฉัตร ติลาศิตประศาสน์. "การเลือกแบบจำลองสำหรับการประมาณความหนาแน่นแบบเคอร์เนลของฟังก์ชันตัวแปรสุ่ม." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต. สาขาวิชาสถิติประยุกต์, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2550.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [13] Wolfgang Hardle. **Smoothing Techniques with Implementation in S**, New York: Springer-Verleg, 1990
- [14] D. Comaniciu, P. Meer. "Mean shift: A robust approach toward feature space Analysis." **IEEE Trans. Pattern Anal. Machine Intell.**, vol. 24, no. 5, May 2002, pp. 603–619.
- [15] D. Comaniciu, V. Ramesh, and P. Meer. "Real-time tracking of non-rigid objects using mean shift." **Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition**, Hilton Head, SC, volume II, June 2000, pp. 142–149.
- [16] D.W. Scott. **Multivariable Density Estimation**, New York: Wiley, 1992.
- [17] Y. Cheng. "Mean shift, mode seeking, and clustering." **IEEE Transactions on PAMI**, Vol. 17, Aug 1995, pp. 790-799.
- [18] D. Comaniciu, V. Ramesh, and P. Meer. "Kernel-based object tracking." **IEEE Trans on Pattern Analysis and Machine Intelligence (PAMI)**, May 2003, pp. 564-577.
- [19] S.T. Birchfield and S. Rangarajan. "Spatiograms versus histograms for region-based Tracking." **Proceedings of the IEEE International Conference on IEEE CVPR'05** , June 2005, pp.1158-1163.
- [20] T. Kailath. "The divergence and Bhattacharyya distance measures in signal selection." **IEEE Trans. Commun. Tech.**, vol. 15, Feb 1967, pp. 52–60.
- [21] สรกานต์ ศรีตองอ่อน. "Linear Regression ." [Online]. Available : www.sorakarn-kmutnb.com/subject/LinearRegression.pdf.
- [22] นัฐพงษ์ หนูสิงห์. "Multiple Peoples Detection and Tracking Designed for Video Surveillance System." ครงงาน. ภาควิชาวิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์. 2551.



ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

บทความและผลงานวิจัยที่ได้รับการตีพิมพ์

1. โกยิต เตชะบริกิติ, ธนรัตน์ ชลิดาพงศ์ และ อาริต ธรรมโน. “การจำแนกท่าตีในวิดีโอที่สนักกีฬาเทนนิสแบบอัตโนมัติ.” **The 3rd National Conference on Information Technology(NCIT2010)**, pp. 128-133, Thailand, October 28-29, 2010.
2. Kosit Teachabarikiti, Thanarat H. Chalidabhongse, Arit Thammano. “Players Tracking and Ball Detection for an Automatic Tennis Video Annotation.” **11th International Conference on Control, Automation, Robotics and Vision(ICARCV2010)**, pp.2491-2494, Singapore, December 5-8.



การจำแนกท่าเดินวิถีทัศนศึกษาเทนนิสแบบอัตโนมัติ

โกษิต เตชะบริกิติ¹, ธนารัตน์ ชลิตาพงศ์¹ และ อาริต ธรรมโม¹

¹คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพมหานคร
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ
E-mails: kosit35185@hotmail.com, thanarat@cp.eng.chula.ac.th, arit@it.kmitl.ac.th

บทคัดย่อ

บทความนี้กล่าวถึงขั้นตอนวิธี ในการจำแนกท่าเดินในวิถีทัศนศึกษาเทนนิส ที่ได้จากวิถีทัศนศึกษาที่เผยแพร่บนอินเทอร์เน็ต ซึ่งมักเป็นวิถีทัศนศึกษาที่มีความละเอียดต่ำ มีสัญญาณรบกวนสูง และมีมุมมองจากกล้องที่ตั้งในระยะไกลเพียงมุมเดียว เริ่มจากการคิดค้นตำแหน่งผู้เล่นทั้งสอง ด้วยขั้นตอนวิธีย้ายค่าเฉลี่ย ร่วมกับการหาความแตกต่างระหว่างพรม กระทำกับฮิสโทแกรมสี จากนั้นทำการตรวจจับเส้นของสนามเทนนิสโดยวิเคราะห์จากค่าความเข้มของจุดภาพ และทำการตรวจจับตำแหน่งของลูกบอล โดยการวิเคราะห์หาบริเวณจุดภาพที่มีการเคลื่อนไหว สุดท้ายนำตำแหน่งผู้เล่นทั้งสอง และ ตำแหน่งลูกบอล เทียบกับตำแหน่งเส้นสนามเทนนิส มาใช้ในการจำแนกท่าเดินเทนนิสออกเป็น โฟร์แฮนด์ แบคแฮนด์ และ เซิร์ฟ

คำสำคัญ— การวิเคราะห์วิถีทัศนศึกษา; การคิดค้นวัตถุ; การจำแนกกิจกรรม

1. บทนำ

ในปัจจุบัน งานวิจัยทางด้านเทคโนโลยีคอมพิวเตอร์วิชั่นและการประมวลผลภาพมีความก้าวหน้าเป็นอย่างมาก และได้ถูกนำมาประยุกต์ใช้ในหลายๆด้าน เช่น ระบบตรวจการผิดจากรีติทัศน์ ระบบวิเคราะห์สภาพจราจร ระบบวิเคราะห์วิถีทัศนศึกษาแบบอัตโนมัติ ระบบค้นหาวิถีทัศนศึกษาอัตโนมัติ รวมทั้งการวิเคราะห์วิถีทัศนศึกษา เพื่อบรรยายเหตุการณ์ที่น่าสนใจหรือเหตุการณ์สำคัญที่เกิดขึ้นในวิถีทัศนศึกษา บทความนี้จะกล่าวถึงงานวิจัยในเรื่องนี้ ดังปรากฏใน [1] [2] [3] แต่หลายระบบยังพบปัญหาในการทำงาน เช่น ขาดความแม่นยำในการติดตามตำแหน่งผู้เล่นเทนนิส เนื่องจากภาพลางมีความซับซ้อน หรือ ปัญหาในการจำแนกประเภทของเหตุการณ์สำคัญในการเล่นเทนนิส เช่น การเซิร์ฟ การตีลูกโฟร์แฮนด์ แบคแฮนด์ วอลเลย์ เป็นต้น งานวิจัย [4] ได้พัฒนาขั้นตอนวิธีในการบรรยายการแข่งเทนนิส จากวิถีทัศนศึกษาแบบอัตโนมัติ โดยทำการคัดแยกผู้เล่นออกจากภาพหลัง โดยใช้คุณสมบัติสีและการเทียบแบบสามมิติ จากนั้นเรียนรู้และจำแนกเหตุการณ์สำคัญในการเล่นด้วยแบบจำลองฮิดเดนมาร์คอฟ (HMMs) โดยดึงลักษณะเด่นจากโครงร่างของผู้เล่นที่คัดแยกได้ผลลัพธ์สุดท้ายคือการจำแนกเหตุการณ์ โดยแบ่งออกเป็น โฟร์แฮนด์ แบคแฮนด์ เซิร์ฟ วอลเลย์ โฟร์แฮนด์ วอลเลย์แบคแฮนด์ งานวิจัย [5] ได้นำเสนอวิธีการติดตามผู้เล่นเทนนิสด้วยวิธีการกรองแบบ SVR particle filter จากนั้นแทนการเคลื่อนที่ของผู้เล่นด้วยฮิสโทแกรมของปติกอลโฟร์แบคแฮนด์ ส่วน (Slice-based optical flow histogram) ทำการรู้จำโดยใช้ Support Vector Machine (SVM) ผลลัพธ์ที่ได้คือการจำแนก การตีเทนนิสแบบโฟร์แฮนด์ และ แบคแฮนด์ ส่วนงานวิจัย [6] ใช้วิธีสร้างแบบจำลองและคัดแยกส่วนภาพของผู้เล่นเทนนิสโดยใช้ฮิสโทแกรมสี จากนั้นใช้รูปทรงของผู้เล่นที่คัดแยกได้มาใช้ในการรู้จำด้วย SVM ผลลัพธ์คือการจำแนกการตีแบบ โฟร์แฮนด์ แบคแฮนด์ หรือ ไม่ใช่ทั้งสองประเภท งานวิจัยที่ได้ยกตัวอย่างข้างต้นล้วนแต่เป็นการใช้ข้อมูลเชิงพื้นที่หรือตำแหน่งของผู้เล่นในการวิเคราะห์ งานวิจัย [7] ได้นำเสนอวิธีการตรวจจับตำแหน่งลูกบอลในภาพ โดยการสร้างกราฟแนววิถีการเคลื่อนที่ของลูกบอลที่มีลักษณะวิถีโค้งพร้อมด้วยการใช้ข้อมูลเชิงขณะทีลูกบอลกระทบไม้ ส่วนงานวิจัย [8] พัฒนาระบบที่มีชื่อว่า TennisSense โดยใช้วิถีทัศนศึกษาจากกล้องที่ติดตั้งอยู่เหนือสนามเทนนิส ในการติดตามตำแหน่งของลูกบอลเทนนิส โดยใช้วิธีการตรวจจับบริเวณภาพที่มีการเคลื่อนไหว จากนั้นทำการรู้จำประเภทการตีโดยพิจารณาจากการเคลื่อนที่ของตำแหน่งลูกบอล งานวิจัย [9] นำเสนอวิธีการดึงคุณลักษณะเด่นทั้งภาพและเสียงที่ได้จากวิถีทัศนศึกษา และรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟ ระบบนี้ได้จำแนกเหตุการณ์ออกเป็น 4 ประเภทคือ การเซิร์ฟที่พลาด การตีลูกกลับ รีเพลย์ และหยุด งานวิจัย [10] เสนอวิธีการตรวจจับตำแหน่งของลูกบอลด้วยวิธีการจำแนกส่วนภาพโดยใช้หลายๆคุณสมบัติ โดยดูจากความเข้มของจุดภาพ จากนั้นใช้ตัวกรอง Particle Filter เพื่อทำการติดตามตำแหน่งลูกบอล

งานวิจัยที่กล่าวมาข้างต้นนั้น ส่วนมากใช้วิถีทัศนศึกษาที่มีคุณภาพสูง บางงานใช้มุมกล้องที่มีการวางตำแหน่งอย่างเฉพาะ บางงานใช้มุมมองหลายมุมมองประกอบกัน ในการวิเคราะห์ งานวิจัยที่นำเสนอในบทความนี้ เราพัฒนาขั้นตอนวิธีในการจำแนกท่าเดินในวิถีทัศนศึกษาแข่งขันเทนนิสประเภทเดี่ยว ที่เผยแพร่บนอินเทอร์เน็ต ซึ่งมักเป็นวิถีทัศนศึกษาที่มีความละเอียดต่ำ มีสัญญาณรบกวนสูง และไม่สามารถกำหนดมุมมองที่เฉพาะเจาะจงได้ ดังนั้นในงานวิจัยนี้ เราจะใช้มุมมองที่พบส่วนใหญ่ คือมุมมองจากกล้องระยะไกลเพียงอย่างเดียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ขั้นตอนวิธี

ในส่วนนี้จะกล่าวถึงขั้นตอนวิธีในการจำแนกเหตุการณ์สำคัญในวิดีโอที่นักกีฬาเทนนิส โดยเริ่มจากการติดตามตำแหน่งผู้เล่นทั้งสอง ด้วยขั้นตอนวิธีย้ายค่าเฉลี่ย (Mean Shift) ร่วมกับการหาความแตกต่างระหว่างเฟรม (Frame Differencing) กับวิธีโทแกรมสี RGB ขนาด 8x8x8 หลังจากนั้น ทำการตรวจนับหาเส้นของสนามเทนนิสโดยวิเคราะห์จากค่าความเข้มของจุดภาพ และทำการตรวจนับหาตำแหน่งของลูกบอล โดยการวิเคราะห์หาบริเวณจุดภาพที่มีการเคลื่อนไหว สุดท้ายนำตำแหน่งผู้เล่นทั้งสอง และตำแหน่งลูกบอล เทียบกับตำแหน่งสนามเทนนิส มาใช้ในการจำแนกท่าตีเทนนิสออกเป็น โฟร์แฮนด์ หรือ แบคแฮนด์

2.1. การติดตามตำแหน่งผู้เล่น

ในการติดตามผู้เล่นในวิดีโอที่นักกีฬาเทนนิสจากมุมมองกล้องระยะใกล้ ซึ่งจะเห็นผู้เล่น ทั้งผู้เล่นที่อยู่ฝั่งใกล้กล้อง (near player) และผู้เล่นที่อยู่ฝั่งไกลกล้อง (far player) เราใช้ขั้นตอนวิธีย้ายค่าเฉลี่ย (Mean Shift) [11] [12] ซึ่งเป็นขั้นตอนวิธีที่มีประสิทธิภาพในการติดตามวัตถุ โดยการใช้การประมาณความหนาแน่นการกระจายแบบนอนพาราเมตริก (non-parametric density estimation) ซึ่งมีสมมติฐานว่าวัตถุที่จะติดตามนั้น จะต้องมีรูปร่างที่คงที่เพื่อที่จะสามารถแยกออกจากพื้นหลังได้ การใช้วิธีย้ายค่าเฉลี่ยแบบมีกรฐาน เพื่อที่จะใช้ติดตามผู้เล่นเทนนิสในวิดีโอที่นักกีฬาเทนนิสที่สนามเทนนิสนั้น พบว่าปัญหาที่เจอจากหลังของผู้เล่น โดยเฉพาะผู้เล่นฝั่งไกลมีความซับซ้อน เนื่องจากปรากฏภาพของ กรรมาการกำกับเส้น และ ผู้เล่น ที่มีการเคลื่อนไหวตลอดเวลา จึงเป็นการยากที่จะติดตามตำแหน่งผู้เล่น ได้อย่างแม่นยำ เราจึงแก้ปัญหาเหล่านี้ด้วยวิธีการใช้ขั้นตอนวิธีติดตามวัตถุด้วยวิธีย้ายค่าเฉลี่ยร่วมกับการหาความแตกต่างระหว่างเฟรม (Frame Differencing) โดยใช้โทแกรมสี RGB ในบริเวณส่วนภาพที่ทำการติดตาม

หลักการของวิธีย้ายค่าเฉลี่ย คือการกำหนดบริเวณส่วนภาพเริ่มต้นที่จะทำการติดตาม จากนั้นทำการหาบริเวณที่มีความหนาแน่นของการกระจายสีที่เหมือนกับต้นแบบมากที่สุดในแต่ละรอบที่กำหนด ในขอบเขตเบนด์วิดท์ ซึ่งในแต่ละรอบก็มีการย้ายจุดศูนย์กลาง แสดงว่าตำแหน่งจุดศูนย์กลางที่ย้ายนั้นมีความหนาแน่นของบริเวณสูงกว่าจุดศูนย์กลางเดิมแต่ถ้าไม่มีการย้ายจุดศูนย์กลางก็ให้ถือว่าจุดนั้นเป็นจุดที่มีความหนาแน่นของบริเวณใกล้เคียงสูงสุด

ปัญหาที่เกิดจากสภาพแวดล้อมในวิดีโอ เช่น ผู้เล่นฝั่งไกลซึ่งเป็นไปได้ที่จะสวมเสื้อสีเดียวกับกรรมาการสนาม หรือผู้เล่นที่เคลื่อนไหวไปมาสามารถทำการแก้ปัญหาโดยใช้การหาความแตกต่างระหว่างเฟรม ผลลัพธ์ที่ได้คือบริเวณส่วนภาพที่มีการเคลื่อนไหวไปเฟรมนั้นๆ จากนั้นทำการประมวลผลส่วนภาพด้วยตัวกระทำมอร์ฟอโลยี (Morphology Operation)

เพื่อเป็นการกำจัดสัญญาณรบกวนในภาพ จากนั้นทำขั้นตอนวิธีย้ายค่าเฉลี่ย

โดยใช้โทแกรมสี RGB แบบ 8x8x8 บิต ของส่วนภาพดังกล่าว

เราสามารถแปลงค่าจุดภาพในปริภูมิสี RGB ให้อยู่ในรูปสีโทแกรมปริภูมิ RGB แบบ 8x8x8 บิต ด้วยวิธีดังนี้

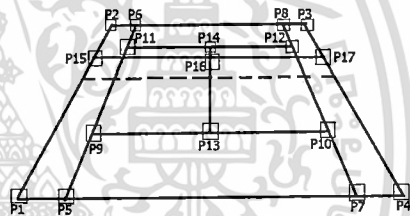
$$S = \text{floor}[R/32] + 8 * \text{floor}[G/32] + 64 * \text{floor}[B/32] \quad (1)$$

โดยที่ R,G,B คือค่าความเข้มจุดภาพในปริภูมิสีแดง เขียว น้ำเงิน ตามลำดับ

2.2. การหาเส้นสนาม

ในการหาเส้นสนาม เราใช้หลักการจากข้อสังเกตที่ว่าค่าความเข้มของจุดภาพบริเวณเส้นสนามจะสูงกว่าบริเวณอื่น และเส้นสนามเป็นเส้นตรง ดังนั้น การหาส่วนที่เป็นเส้นสนามเทนนิส เราเริ่มจากการกำหนดจุดในบริเวณขอบของสนามเทนนิสทั้ง 17 จุด จากนั้นเชื่อมเส้นตรงระหว่างจุดดังกล่าวในรูปที่ 1 จากนั้นทำการหาค่าความเข้มของจุดภาพในบริเวณที่กำหนด เมื่อได้ตำแหน่งที่มีความเข้มสูงสุดของแต่ละเส้น นำพิกัดที่มีความสว่างสูงสุดลบกับเส้นที่ได้จากการสร้างในคอนเทรสต์ จากนั้นหาฐานนิยามและหาพิกัดที่อยู่ในช่วงของฐานนิยาม นำพิกัดที่อยู่ในช่วงของฐานนิยามทำการถดถอยเชิงเส้น (Linear Regression)

โดยมีขั้นตอนดังนี้



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

$$ExpandLine_t = [(LineX_{t-1} \pm y) (LineY_{t-1} \pm x)]$$

$ExpandLine_t$ คือ เส้นทั้งหมดที่ได้ถูกขยายความหนา

5. หาพิกัดของความเข้มในจุดภาพสูงสุด ของ $ExpandLine_t$ โดยที่ถ้าเป็นเส้นที่ขยายในพิกัด y จะพิจารณาความเข้มของจุดภาพในแนวตั้ง ในทางกลับกันถ้าเป็นเส้นที่ถูกขยายในพิกัด x จะพิจารณาความเข้มของจุดภาพในแนวนอน

$$MaxLine_t(x, y) = MaxIntensity(ExpandLine_t)$$

$MaxLine_t(x, y)$ คือ พิกัด (x, y) ของตำแหน่งที่มีความเข้มของ

จุดภาพสูงสุดของในแต่ละเส้น $ExpandLine_t$ ของภาพในเวลา t

6. คืนค่าพิกัดของเส้น L_{t-1}

$$Line_t(x, y) = returnCoordinate(L_{t-1})$$

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

$t-1$

7. หาผลต่างของพิกัดจากขั้นตอนที่ 5 และ 6

$$(Line_t(x, y) - MaxLine_t(x, y)) = Pix(x, y)$$

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

8. นำพิกัด x และ y ของทุกตำแหน่ง i ใน Pix มารวมกันได้

$$D_i = Pix(x \pm y)$$

9. หาฐานนิยมของ $Mode(D_i) = m$

10. หากค่า D_i ที่อยู่ระหว่าง $m \pm th$ เพื่อทำการเก็บพิกัด โดยที่ th

คือค่าขีดเบ่ง

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

Y_i คือพิกัดจาก $MaxLine_t$ ที่อยู่ระหว่าง $m \pm th$ โดยมีพิกัด x และ y ที่

ตรงกับตำแหน่ง D_i

11. ทำวิธีการลดข้อเชิงเส้น จะทำให้ได้ผลลัพธ์คือเส้นสนาม

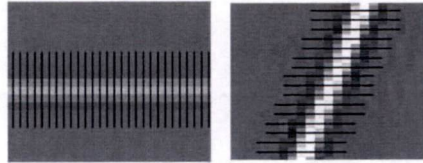
$$LineCourt_t = linearregression(Y_i)$$

$LineCourt_t$ ทำการเก็บผลลัพธ์จากการหาเส้นสนามในเวลา t

12. กำหนด $Line_t = LineCourt_t$ เพื่อใช้เป็นข้อมูลในการทำเฟรม

ถัดไป

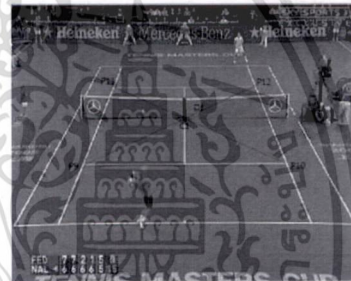
13. ทำซ้ำขั้นตอนที่ 4-12 จนครบเฟรม



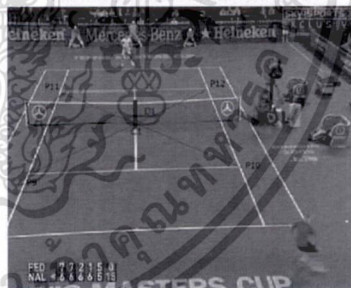
รูปที่ 2 แสดงถึงการขยายขนาดเส้น รูปซ้าย 2a ขวา 2b

จากรูปที่ 1 เส้นประคือเส้นที่เป็นส่วนของแนวด้านล่างเรทำการหาโดย

1. สร้างเส้นระหว่าง P11-P10 และ P12-P9
2. หาจุดตัดของเส้นทั้งสองเส้นจะได้จุดกึ่งกลางของแนต (C1) ซึ่งแสดงในรูปที่ 3 และ รูปที่ 4
3. หาระยะห่างระหว่าง C1 กับ P16 ได้ระยะห่างคือ (D1)
4. ได้เส้นประ ดังรูปที่ 1 ที่เป็นส่วนของแนวด้านล่าง 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

เมื่อใช้วิถีทัศน์การแข่งเทนนิสที่มีรายละเอียดค่าโดยเฉพาะที่เหมาะสม จากอินเทอร์เน็ตผนวกกับการเคลื่อนที่ของลูกเทนนิสอย่างรวดเร็วเป็นการยากที่จะรูปร่างหรือสีของลูกเทนนิสเพื่อใช้ในการหาลูกเทนนิส

วิธีการที่ใช้ในการตรวจจับลูกเทนนิสเป็นดังนี้

เริ่มจากใช้วิธีความแตกต่างระหว่างเฟรมซึ่งเป็นวิธีการที่คำนวณได้ง่าย และรวดเร็วในการหาบริเวณที่มีการเคลื่อนที่หลังจากนั้นเราจะใช้เทคนิค ความสัมพันธ์บางอย่างเพื่อจะจัดบริเวณที่ไม่มีลูกเทนนิส

1. ความแตกต่างระหว่างเฟรม เป็นวิธีอย่างหนึ่งของการหาบริเวณที่มีการเคลื่อนไหว การคำนวณทำได้ ง่ายและรวดเร็วโดย นำภาพเฟรมที่ t ลบด้วยภาพเฟรมที่ $t-1$ $X_t - X_{t-1}$ ถ้าผลต่างมีค่าเกินกว่าขีดเบ่ง แสดงว่าจุดภาพมีความต่าง
2. ทัมอร์ฟอโลยี หลังจากทำการความแตกต่างระหว่างเฟรม เราจะทำการกระบวนการแบบปิด (Operation Closing) เพื่อทำการประติดประต่อส่วนที่เล็กของจุดภาพ

3. กรองลูกบอลโดยใช้ขนาด กระบวนการหาลูกบอลเราใช้การกรองโดยใช้ช่วงขนาดลูกบอลที่ซึ่ง ช่วงของลูกบอลได้จากการกระบวนการเทรนนิ่ง
4. กรองลูกบอลในขอบเขตของสนามโดยใช้ระยะห่างเส้นสนาม เมื่อใดที่ใกล้มีการเคลื่อน ผลลัพธ์ของการตรวจหาลูกบอลจะมีจำนวน ที่สูงขึ้นเนื่องมาจากเส้นสนาม ดังนั้นเราจะแก้ไขข้อผิดพลาดนี้โดยการหา ระยะห่างเส้นสนาม โดยนำข้อมูลเส้นสนามที่ได้จากการทำในขั้นตอนที่ 2.2 ทำการหาขอบเขตในสนามจากนั้นหารระยะห่างเส้นสนามทุกเส้นกับ ผลลัพธ์ของการตรวจหาลูกบอล ดังรูปที่ 5 เส้นสีแดงแปลตรงสนาม

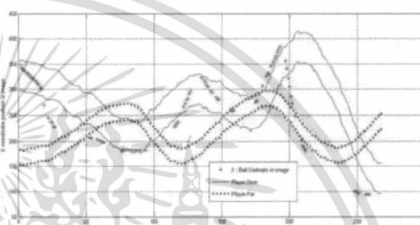


รูปที่ 5 ผลลัพธ์ของการตรวจหาลูกบอลด้วยวิธีขอบเส้นสนาม

5. ลบจุดภาพที่เป็นบริเวณส่วนภาพของผู้เล่นทั้งฝั่งใกล้ฝั่งไกล
- 2.4. จำนวนกัทิตเทนนิส เนื่องจากวิธีที่ค้นพบที่ใช้มีความละเอียดค่าเราจึงใช้ข้อมูลเชิงเวลาในการ จำนวนกัทิต โดยข้อมูลที่นำมาใช้นั้นได้แก่ ตำแหน่งของผู้เล่นเทนนิสและ ข้อมูลลูกบอล ซึ่งข้อมูลทั้ง 2 นี้มีตำแหน่งพิกัดประกอบไปด้วย (x,y) ซึ่ง

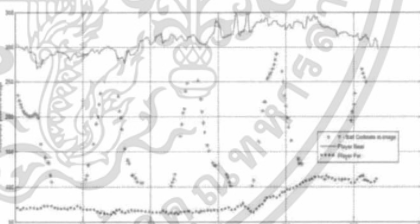
เราจะทำการสร้างกราฟ 2 กราฟ เพื่อใช้ในการจำแนกกัทิตเทนนิสโดยมี ขั้นตอนการทำดังนี้

1. แยกพิกัด x และ y ของตำแหน่งบอลและตำแหน่งผู้เล่น
2. สร้าง 2 กราฟโดยในแนวตั้งของกราฟแรก (รูปที่ 6) คือ ตำแหน่ง พิกัด x ของในภาพ ส่วนในแนวราบ คือ ค่าคัมของเฟรม โดย ข้อมูลที่ใช้ในการวาดกราฟแรก จะประกอบไปด้วย ตำแหน่งของ ลูกเทนนิสและตำแหน่งทางด้านซ้ายและขวาของผู้เล่นเทนนิส ส่วนกราฟที่สอง (รูปที่ 7) ในแนวตั้งของกราฟ คือ ตำแหน่งพิกัด y ของในภาพ ส่วนในแนวราบ คือ ค่าคัมของเฟรม โดยที่ข้อมูลที่ใช้ ในการวาดกราฟ จะประกอบไปด้วย ตำแหน่งของลูกเทนนิส และตำแหน่งที่เป็นจุดศูนย์กลางของผู้เล่นเทนนิส



รูปที่ 6 กราฟที่ใช้สำหรับการจำแนกกัทิตเทนนิส

จากรูปที่ 6 เส้นสีน้ำเงิน (เส้นทึบ) จะแสดงถึงตำแหน่งของผู้เล่นเทนนิสฝั่ง ใกล้ฝั่งซึ่งมี 2 เส้นโดยเส้นด้านบนคือตำแหน่งทางด้านซ้ายมือของผู้เล่น ส่วน เส้นล่างคือตำแหน่งทางด้านซ้าย ส่วนเส้นสีเขียว (เส้นประ) แสดงถึงผู้เล่น เทนนิสฝั่งไกล โดยเส้นบนคือตำแหน่งทางด้านซ้ายมือของผู้เล่นส่วนเส้นล่าง คือตำแหน่งทางด้านขวามือ จุดสีแดงคือ ตำแหน่งของลูกบอล

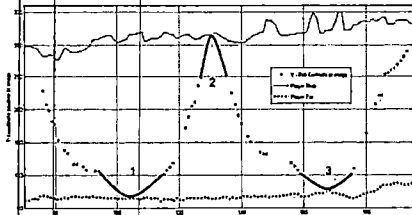


รูปที่ 7 กราฟที่ใช้หาเหตุการณ์ตีลูกเทนนิส

จุดสีแดงในรูปที่ 7 คือตำแหน่งของลูกเทนนิส ส่วนเส้นสีน้ำเงินด้านบน (เส้นทึบ) คือตำแหน่งของผู้เล่นฝั่งใกล้ และเส้นสีเขียวด้านล่าง (เส้นประ) คือตำแหน่งผู้เล่นฝั่งไกล

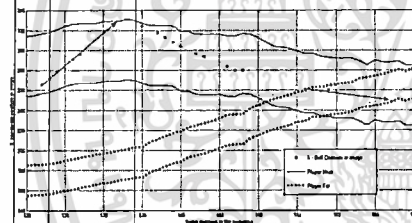
ใช้กราฟที่สอง ที่แสดงในรูปที่ 7 เพื่อช่วยในการหาช่วงเวลาตีเทนนิส โดยช่วงของเวลาที่ได้จะประกอบไปด้วย เฟรมก่อนตีลูกเทนนิส และ เฟรม

หลังตีลูกเทนนิส โดยเฟรมก่อนตีลูกเทนนิสจะมีน้ำหนักเฟรมที่ต่ำกว่าเฟรม
หลังตีเทนนิส ดังรูปที่ 8 ที่มีช่วงเฟรมของกราฟตีเทนนิส 3 ช่วง



รูปที่ 8 ตัวอย่างเหตุการณ์ที่มีช่วงการตี 3 ช่วง

หลังจากได้ช่วงเวลาที่มีการตีเทนนิสเราในช่วงเวลานี้ไปใช้ในการ
แบ่งย่อยประเภทการตีโดยใช้กราฟในรูปที่ 6 พิจารณารวว่าตำแหน่งบอลอยู่
ด้านซ้ายหรือด้านขวาของผู้เล่นเทนนิสตำแหน่งบอลอยู่ทางด้านซ้ายของผู้เล่น
ไกลผลลัพท์คือตีทำโฟลวแมนด์ แต่ถ้าอยู่ทางด้านขวามลลัพท์ที่ได้คือตี
แบ็คแฮนด์ ส่วนตัวเป็นผู้เล่นฝั่งใกล้ตำแหน่งบอลอยู่ทางด้านขวามลลัพท์ที่
ได้คือการตีทำโฟลวแมนด์ ถ้าอยู่ด้านซ้ายมลลัพท์ที่ได้คือการตีทำแบ็ค
แฮนด์ แต่ถ้าตำแหน่งลูกบอลเข้าไปอยู่ระหว่างตำแหน่งด้านซ้ายและขวา
ของผู้เล่น แสดงได้ดังรูปที่ 9 เราจะทำการพิจารณาโดยถ้าวิถีทางของลูก
บอลเริ่มเข้าจากทางด้านขวาของผู้เล่นมลลัพท์ที่ได้คือลูกบอลจะอยู่ทาง
ด้านซ้ายของผู้เล่นแต่ถ้าลูกบอลเข้ามาทางด้านซ้ายของผู้เล่นมลลัพท์ที่ได้คือ
ลูกบอลจะอยู่ทางด้านขวาของผู้เล่น

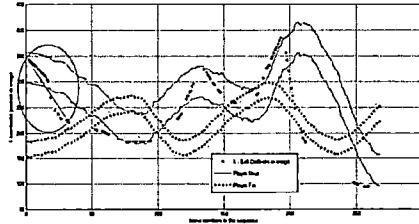


รูปที่ 9 แสดงตำแหน่งบอลที่อยู่ระหว่างกลางของผู้เล่นเทนนิส

จากรูปที่ 9 ทำการพิจารณาลูกครีส์ค่าซ้ายสุดจะเห็นได้ว่าลูกเทนนิสจะ
เข้ามาทางด้านซ้ายของผู้เล่นฝั่งใกล้ ดังนั้นมลลัพท์ที่ได้คือเทนนิสจะ
เคลื่อนที่ไปทางด้านขวาจึงได้ทำดีแบ็คแฮนด์ ส่วนลูกครีส์ค่าขวาสุดที่
ลูกเทนนิสได้เข้ามาจากทางด้านซ้ายของผู้เล่นฝั่งใกล้ ดังนั้นมลลัพท์ที่ได้คือลูก
เทนนิสจะต้องไปทางด้านขวา จึงได้ทำดีโฟลวแมนด์

ในการหาว่าผู้เล่นเทนนิสฝั่งใดเลิฟพื้นั้น ทำการพิจารณาโดยใช้กราฟ
จากรูปที่ 6 เพื่อทำการดูแนวโน้มของตำแหน่งของลูกเทนนิสว่ามีน้ำหนัก
กับผู้เล่นเทนนิสฝั่งใด

จากรูปที่ 10 เป็นการแสดงตัวอย่างของการจำแนกเหตุการณ์เลิฟ จะ
เห็นได้ว่าในบริเวณเวลานั้น พิกัดของลูกบอลมีความใกล้กับตำแหน่งของผู้
เล่นเทนนิสใกล้ และมีแนวโน้มของการเคลื่อนที่ของลูกบอลเข้าไปสู่ผู้เล่น
ใกล้ ดังนั้นจึงเป็นการเลิฟของผู้เล่นฝั่งใกล้



รูปที่ 10 แสดงตำแหน่งบอลเวลาที่ตีลูกเทนนิส

3. ผลการทดลอง

เราได้ทำการทดสอบวิธีการที่นำเสนอโดยใช้วิธีที่ค้นทั้งหมด 8 คลิป ที่ซึ่ง
ได้ทำการคัดมาจากการแข่งขันเทนนิส รายการTENNIS MASTER, US
OPEN และ ATP โดยแต่ละคลิปที่ได้ ทำการทดสอบเป็นการแข่งขันเทนนิสที่
แตกต่างกันทั้งหมด 6 คู่ จำนวนเฟรมที่ทำการทดลองมีทั้งหมด 2926 เฟรม
ในแต่ละวิธีที่ค้นคลิปเริ่มจากการเลิฟ โดยต้องไม่เป็นการเลิฟเสียและจะ
จบลงด้วยการเสียดะแมน 1 คะแนน ทั้งนี้ผู้เล่นเทนนิสทั้งไกลและใกล้
จะต้อง จับไม้เทนนิสด้วยมือขวา

ทำการทดลองทั้ง 8 คลิป โดยใช้วิธีการที่ได้กล่าวมาแล้วในหัวข้อที่ 2
คือ ทำการติดตามหาตำแหน่งผู้เล่นเทนนิสทั้งไกลและใกล้ จากนั้นทำการ
ตรวจหาลูกบอล โดยที่นำผลลัพท์ของการหาเส้นสนามเทนนิสมาใช้ร่วม
ด้วยเพื่อเพิ่มความถูกต้อง สุดท้ายทำการจำแนกท่าตีเทนนิส โดยใช้ผลของ
การตรวจหาลูกบอลร่วมกับตำแหน่งผู้เล่นเทนนิส ได้ผลการจำแนกมา
3 ประเภท นำผลลัพท์ที่ได้มาเปรียบเทียบกับ กราวด์ทรูธ (ground truth) ใน
ตารางที่ 1

ตารางที่ 2 แสดงถึงผลลัพท์ของการตีลูกบอล ส่วนตารางที่ 3 แสดง
ผลลัพท์ที่ได้จากการจำแนกท่าตีเทนนิส และในรูปที่ 11 ในแถวบน
แสดงผลลัพท์ของการใช้ขั้นตอนวิธีแยกค่าเฉลี่ย ร่วมกับ ความแตกต่าง
ระหว่างเฟรม ส่วนแถวล่างเป็นการใช้ขั้นตอนวิธีแยกค่าเฉลี่ย เพียงอย่าง
เดียวจะเห็นได้ว่าถ้าใช้เพียงขั้นตอนวิธีแยกค่าเฉลี่ย ทำให้เกิดความผิดพลาด
ในกรณีตีเสี่ยของผู้เล่นที่มีเหมือนสนับเทนนิสหรือตีเหมือนกรมการ

ตารางที่ 1. แสดงจำนวนกราวด์ทรูธ

กราวด์ทรูธ					
ผู้เล่นไกล (FAR)			ผู้เล่นใกล้ (NEAR)		
SERVE	PH	BH	SERVE	PH	BH
2	21	19	6	16	18

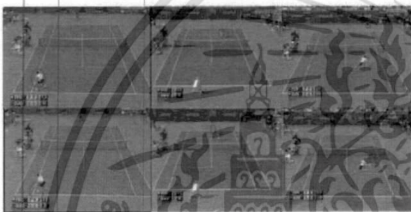
ตารางที่ 2. ผลลัพท์ของการตีลูกบอล

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	PERCISION	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
 PERCISION = TP / (TP+FP), RECALL = TP / (TP+FN)
 BH = ท่าตีแบคแฮนด์, FH = ท่าตีโฟร์แฮนด์, SERVE = เซิร์ฟ



รูปที่ 11 แสดงผลลัพธ์ของการติดตาม

4. สรุปผลของงานที่จะทำต่อไป

ในอนาคตเราได้นำเสนอการติดตามผู้เล่น การหาเลนเนตเทนนิส และการตรวจหาลูกบอลในวิดีโอที่สมัครแข่งขันเทนนิสที่มีความละเอียดวิดีโอที่คมชัด ในมุมมองกล้องระยะไกลเพียงอย่างเดียว ใช้อัลกอริทึม จินตคณิตวิสัย ค่าเฉลี่ย ร่วมกับ ความแตกต่างระหว่างเฟรมโดยใช้ปริภูมิสี RGB หาเส้นของสนามเทนนิส โดยการดูค่าความสว่างของจุด ทดสอบทำการประมวลผลเส้นตรง นอกจากนั้นเรายังตรวจสอบและติดตามตำแหน่งของลูกเทนนิส โดยความแตกต่างระหว่างเฟรม รวมทั้งใช้เทคนิคของความลับที่รับมาอย่าง เพื่อกำจัดสิ่งที่ไม่ใช่ลูกบอล 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. Pauri, 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.

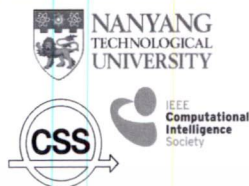


11th International Conference on Control, Automation, Robotics and Vision

ICARCV 2010

<http://www.icarcv.org>

7 - 10 December 2010, Singapore



IEEE Control Systems Society



Organiser

-School of Electrical & Electronic Engineering,
Nanyang Technological University

Technical Co-Sponsors

-IEEE Control Systems Society
-IEEE Computational Intelligence Society
-IEEE Robotics and Automation Society

International Advisory Committee

Baillieul, John, USA
Borne, Pierre E., France
Bose, Bimal K., USA
Cao, Xiren, HK, China
Cassandras, Christos G, USA
Chai, Tianyou, China
Chen, Guanrong (Ron), HK, China
De Luca, Alessandro, Italy
De Silva, Clarence W, Canada
Er, Meng Hwa, Singapore
Fogel, David B., USA
Goodwin, Graham C, Australia
Guo, Lei, China
Hill, David, Australia
Ikeuchi, Katsushi, Japan
Isidori, Alberto, Italy
Ljung, Lennart, Sweden
Kam, Chan Hin, Singapore
Kaynak, M. Okayay, Turkey
Krstic, Miroslav, USA
Kumar, P. R., USA
Laugier, Christian, France
Lee, George C. S., USA
Lee, Kok Meng, USA
Lewis, Frank L., USA
Newcomb, Robert W, USA
Ortega, Romeo, France
Pal, Nikhil R, India
Parent, Michel, France
Polycarpou, Marios M, Cyprus
Siciliano, Bruno, Italy
Slotine, Jean-Jacques E, USA
Takagi, Hideyuki, Japan
Tarn, Tzyh Jong, USA
Wang, Danwei, Singapore
Yen, Gary G., USA
Zurada, Jacek M., USA

Organising Committee

Wen, Changyun (General Chair)
Wang, Danwei (Advisor)
Suganthan, P. N. (Program Chair)
Xie, Lihua (Invited Session Chair)
Sung, Eric (Invited Session Co-Chair)
Cheah, C. C. (Invited Session Co-Chair)
Soh, Yeng Chai (Panel & Keynote Chair)
Wang, Han (Finance Chair)
Zhang, Cishen (Publication Chair)
Chua, Chin Seng (Local Arrangement Chair)
Wong, Patricia (Publicity Chair & Secretary)

CALL FOR PAPERS

The 11th International Conference on Control, Automation, Robotics and Vision, ICARCV 2010, will be held in Singapore from 7 to 10 December 2010. The conference is organised by the Nanyang Technological University of Singapore. ICARCV focuses on both theory and applications mainly covering the topics of control, automation, robotics and vision. In addition to the technical sessions, there will be invited sessions, keynote addresses and plenary panel sessions.

The Proceedings of ICARCV are included in the IEEE Xplore and indexed by EI Compindex as well as ISI Proceedings.

The topics of interest include, but are not limited to:

Control: Adaptive control; Robust control; Process control; Complex systems; Co-operative control; Identification and estimation; Nonlinear systems; Intelligent systems; Discrete event systems; Hybrid systems; Networked control systems; Sensor network systems; Delay systems; Neural networks; Fuzzy systems; Control of biological systems; Precision motion control; Control applications; Control engineering education.

Automation: Man-machine interactions; Process automation; Intelligent automation; Factory modeling and simulation; Computational intelligence in automation; Home, laboratory and service automation; Network-based systems; Planning, scheduling and coordination; Nano-scale automation and assembly; Instrumentation systems; Biomedical instrumentation and applications.

Robotics: Modeling and identification; Robot control; Mobile robotics; Mobile sensor networks; Perception systems; Micro robots and micro-manipulation; Visual servoing; Search, rescue and field robotics; Robot sensing and data fusion; Localization, navigation and mapping; Dexterous manipulation; Medical robots and bio-robotics; Human centered systems; Space and underwater robots; Tele-robotics; Computational intelligence in robotics; Mechanism design and applications.

Vision: Image/video analysis; Feature extraction, grouping and segmentation; Scene analysis; Pattern recognition; Learning in vision; Human-computer interaction; Tracking and surveillance; Biometrics; Biomedical image analysis; Computational intelligence in vision; Activity/behaviour recognition; Applications.

Best Paper Award: Selection of the best paper will be made at the Conference based on both the technical content and presentation. The winner will be chosen by the Technical Program Committee in consultation with the International Advisory Committee.

Paper Submission:

Papers must be written in English and should describe original work in detail. Please download the template from the conference website and submit the full papers online <http://www.icarcv.org> by 15 April, 2010. Upon acceptance, authors will be required to register and present their papers.

Invited Sessions:

Prospective organizers of Invited Sessions are invited to submit proposals focusing on topics of ICARCV 2010 to the Invite Session Chairs, **Professors Lihua Xie (Control)**, **Chien Chern Cheah (Robotics and Automation)** and **Eric Sung (Vision)** by 15 April, 2010. Invited sessions are to provide a forum for focused discussions on new topics, or innovative applications of established approaches.

Keynote Speakers:

Professor Miroslav Krstic, *University of California, San Diego, USA*
Professor Bruno Siciliano, *University of Naples Federico II, Italy*
Professor Zhengyou Zhang, *Microsoft Research, USA*

Important Dates:

Deadline for Full Paper submission	15 April 2010 (Extended)
Notification of Acceptance	1 July 2010
Deadline for Camera Ready Manuscript Submission	1 September 2010
Deadline for Authors' Registration	1 September 2010

Website - <http://www.icarcv.org>

Email - secretariat@icarcv.org



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Players Tracking and Ball Detection for an Automatic Tennis Video Annotation

Kosit Teachabarikiti¹, Thanarat H. Chalidabhongse², Arit Thammano¹

¹Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand 10520

²Faculty of Engineering, Chulalongkorn University
Bangkok, Thailand 10330

E-mail addresses: {kosit35185@hotmail.com, thanarat.c@chula.ac.th, arit@it.kmitl.ac.th}

Abstract— This paper describes our algorithms for players tracking and ball detection for an automatic broadcast tennis video annotation. The system detects and tracks the players using a robust non-parametric procedure for estimating density gradients called the mean shift algorithm. The basic mean shift tracking algorithm assumes that the target object has to separate sufficiently from background, but this assumption is not always true especially when tracking is carried out in dynamic backgrounds such as in sport videos. To cope with this problem, in our proposed system, we embrace the motion segmentation and use the 8x8x8 color histogram to be feature distribution for mean shift tennis players tracking. In order to determine the players' actions precisely, the system also detect and track ball positions using frame differencing as well as applying some correlation techniques to eliminate false detections. Based on both players' motion patterns and ball positions, the system can precisely classify the players' action into backhand ground stroke and forehand ground stroke. Videos of broadcast tennis games downloaded from the Internet have been tested. The results show our system is able to precisely classify the player's actions with 83.7% precision and 82% recall rates.

Keywords— object tracking, action recognition, sport video analysis, tennis videos.

I. INTRODUCTION

Video annotation has long been an active research area in computer vision. The goal is to automatically extract salient information from input videos either in real-time or off-line. One of the applications of this area is automatic sport video analysis and annotation [1][2][3]. In tennis context, the game is typically described by key events such as strokes, including serves, forehands, backhands, volleys, drop shots, etc. In order to detect these strokes from videos, while some researchers classify the strokes using only the player motion [4][5], some found that more cues such as court lines, ball position, and audio are also beneficial [6][7][8][9]. [4] segmented player from the background scene using color and 3D model fitting. Then, six HMMs were constructed and trained using combinations of features extracted from player silhouette. One for each type of recognized events: forehand, backhand, service, smash, forehand volley and backhand volley. [5] proposed motion descriptor that called slice-based optical flow histogram. Instead of trying to measure precise

pixel displacement, they treated this optical flow as spatial pattern of noisy measurement. Then, they employed SVM for player's action training and classification into two classes; left and right swings. In addition to player spatial information extracted from videos, many works incorporate some other information that can be useful in event recognition. [6] detected court lines by means of the Hough Transform, while the players' positions were extracted by finding edge pixels whose orientation is different from the lines of the court. [7] presented a trajectory-based algorithm to detect and track the ball. Their approach did not decide whether an object is the ball, instead it decides whether a candidate trajectory is a ball trajectory. [8] presented their TennisSense system that extracts useful metadata from the overhead court camera. The tennis ball is tracked using motion images for ball candidate detection and then link ball candidates into locally linear tracks. From these tracks, the serve and rally events can be recognized. [9] presented HMM-based structuring of tennis videos using visual and audio cues. Their system labels each segment of a tennis video with one of the following four types: first missed serve, rally, replay, and break. Hidden Markov Models (HMMs) were used to merge audio-visual information, and to represent the hierarchical structure of a tennis match. Domain knowledge is implicitly taken into account through the learning process associated with HMMs.

Many of these works use high quality videos and some uses specific set-up camera angle or even multiple cameras. In this paper, we proposed algorithms for player tracking and ball detection in low resolution broadcast tennis videos from a single far-view camera. Our underlying motivation is to develop an automatic system to annotate broadcast tennis videos that are published on the Internet. This is a challenging problem especially when most of the broadcast videos found on Internet are low-resolution and noisy.

II. SYSTEM OVERVIEW

Fig.1 shows our algorithm for event recognition in broadcast tennis videos. The system first detects and tracks the players using a robust non-parametric procedure for estimating density gradients called the mean shift algorithm on 8x8x8 RGB color

difference image with this template. We then can eliminate most of the false detection due to camera motion. Fig.5 shows the final result of ball detection from moving camera after performed steps as described, while Fig.6 shows the result in normal (non-moving camera) scene.

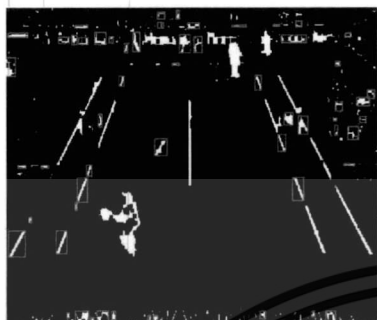


Figure 3. False ball detection due to camera motion

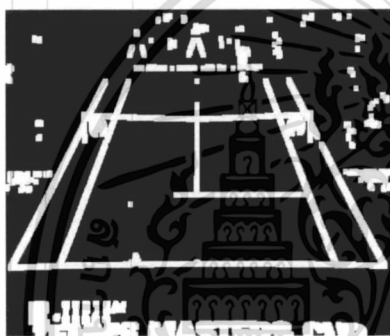


Figure 4. Court line pixel template



Figure 5. Final result of ball detection



Figure 6. Players and ball detection result in normal scene.

C. Stroke classification

After we obtain the tracked locations of players and ball in tennis video sequences, we can use this spatial temporal information to detect and classify the players' action into backhand ground stroke and forehand ground stroke.

1) *Ball hit detection:* Consider trajectory of players and ball in vertical axis of the image in Fig.7, the ball hit event can be detected when the ball trajectory crosses with the players' trajectory.

2) *Stroke classification:* For forehand or backhand strokes, we consider another trajectory of players' position in X-coordinate. Noticing that when the near-player hit forehand, his/her motion is from left to right; while the backhand is vice versa.

IV. EXPERIMENTAL RESULTS

We tested the proposed method on 10 video clips that cut from two matches of broadcast Tennis Masters Cup. The total frame numbers are 1,914 frames. Each video clip starts by serve and ends by 1 ball point.

Table 1 shows the result of ball hit detection. The precision is 96.08% and recall is 98%. Table 2 shows classification results. On average, the stroke classification precision is 83.7% while the recall rate is 82%.

V. CONCLUSION AND FUTURE WORK

We have presented a players tracking and ball detection in low resolution broadcast tennis videos from a single far-view camera. The players are tracked using color-based mean shift algorithm. We also detect and track ball positions using frame differencing as well as applying some correlation techniques to eliminate false detections. Based on both players' motion patterns and ball positions, the system can precisely classify the players' action into forehand and backhand ground stroke.

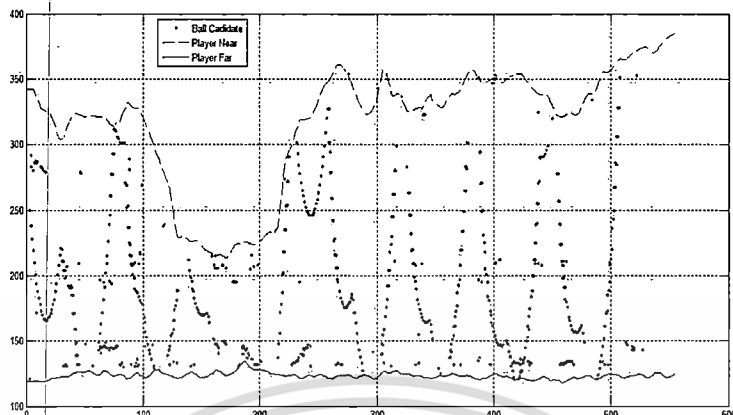


Figure 7. Plots of trajectory of far-player (black), near-players (blue), and the ball (red). The vertical axis is the Y-coordinate position in image. The horizontal axis is frame numbers in the sequence. Ball hit event is detected when the trajectory of the ball cross with the trajectory of the players. The Y-coordinate position of hit event determines whether it is ground stroke or volley.

TABLE I. BALL HIT DETECTION RESULT

FAR/ NEAR	TP	FN	FP	PRECISION	RECALL
FAR	26	1	1	96.3%	96.3%
NEAR	23	0	1	95.8%	100%

TABLE II. STROKE CLASSIFICATION RESULT

	TP	FN	FP	PRECISION	RECALL
FAR-PLAYER					
FH	6	4	2	75.0%	60.0%
BH	15	2	3	83.3%	88.2%
NEAR-PLAYER					
FH	5	1	2	71.4%	83.3%
BH	15	2	1	93.8%	88.2%

In the future, we plan to extend the system to be able to classify more salient events such as serve and idle state.

REFERENCES

- [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. Levisnais-Obadia, and D. Koubaroullis, "Generation of semantic cues for sports video annotation," Proc. IEEE ICIP 2001.
- [4] M. Pelkovic, 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] C. Calvo, A. Micarelli, and E. Sangineto, "Automatic annotation of tennis video sequences," Proc. the 24th DAGM Symposium on Pattern Recognition, 2002.
- [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] 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.
- [11] S.T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," Proc. IEEE CVPR'05, pp.1158-1163, June 2005.

ประวัติผู้เขียน

ชื่อผู้เขียน

นายโกษิต เตชะบริกิติ

วัน เดือน ปีเกิด

วันที่ 29 มิถุนายน พ.ศ. 2527

ที่อยู่

507/370 ซ.สาธุประดิษฐ์ 31 ถ.สาธุประดิษฐ์

แขวงช่องนนทรี เขตยานนาวา กรุงเทพฯ 10120

อีเมล

kosit35185@hotmail.com

วุฒิการศึกษาระดับปริญญาตรี

วิทยาศาสตรบัณฑิต สาขาคณิตศาสตร์ประยุกต์

สถานที่สำเร็จการศึกษา

สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปีที่สำเร็จการศึกษา

2549



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้