

บทที่ 4

ผลการทดลอง

ในบทนี้จะเป็นการกล่าวถึงรายละเอียดของผลการทดลอง ซึ่งนำเสนอทั้งในรูปแบบของวัตถุประสงค์ (Objective) และ วัตถุประสงค์ (Subjective) แบ่งออกเป็นส่วนใหญ่ 2 ส่วน ตามโครงสร้างของระบบที่นำเสนอในบทที่ 3 คือ ส่วนการตรวจหาและการติดตาม

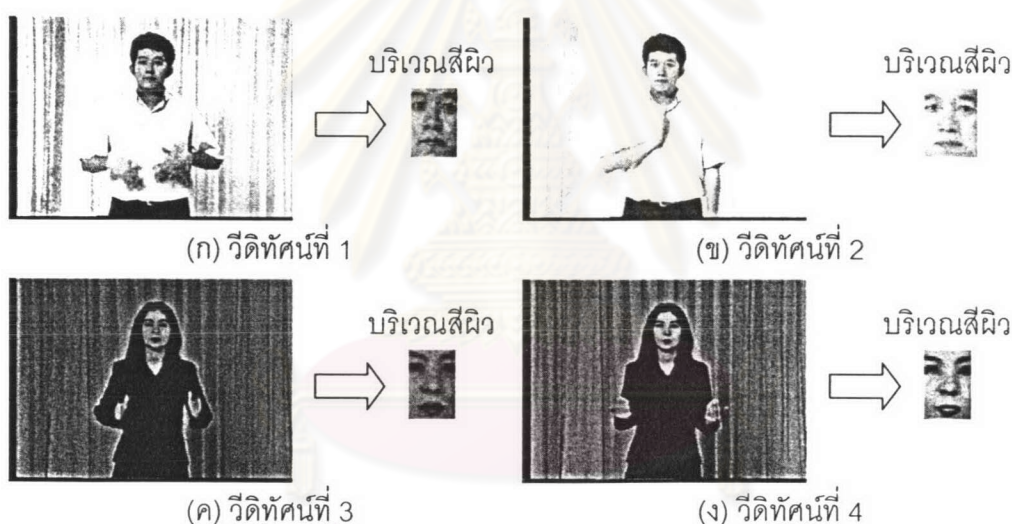
ในส่วนการตรวจหาซึ่งคุณลักษณะที่นำมาใช้คือการแยกส่วนสีผิวดังที่กล่าวมาแล้วในเบื้องต้น แสดงผลของกรรมวิธีที่ใช้วัดสัญญาณรบกวนในแต่ละปริภูมิสีโดยใช้แบบจำลองชนิดไม่มีตัวแปร (Non parametric model) เพื่อคัดเลือกปริภูมิสีที่เหมาะสมที่สุด ผลการแยกส่วนสีผิวเมื่อใช้แบบจำลองที่เสนอโดย S.L. Phung [13] และการแยกสีผิวเมื่อใช้แบบจำลองวงรี รวมถึงกรรมวิธีในการหาขีดแบ่ง (threshold) ที่ดีที่สุด การเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองแบบเดิม [13] กับแบบจำลองวงรี การเรียนรู้สีผิวของกล้องโซนี่ (Sony) รุ่น EVI-D100P ผลจากการใช้ตัวดำเนินการแบบปิดกับโครงสร้างส่วนย่อยที่เลือกใช้ ผลของการแยกองค์ประกอบสำคัญต่าง ๆ อาทิเช่น แยกองค์ประกอบสำคัญบนใบหน้า การทำอินนิง เป็นต้น ผลของการกำหนดจุดศูนย์กลางแต่ละล็อบเพื่อใช้ประโยชน์ในการติดตาม

ส่วนการติดตาม ในเบื้องต้นจะทำการแยกระหว่างใบหน้าและมือเชิงพื้นที่ (Spatial) โดยใช้คุณลักษณะเด่นทางขนาด หรือ อาจต้องพิจารณาองค์ประกอบสำคัญเพิ่มเติมกรณีที่มีการตัดสินจากขนาดไม่เหมาะสม เมื่อแยกระหว่างล็อบใบหน้าและมือได้แล้ว ก็จะทำทำการกำหนดการประมาณการเคลื่อนที่ที่เหมาะสม ซึ่งจะแสดงผลการวัดความผิดพลาดของแบบจำลองที่ใช้ในการติดตามการเคลื่อนที่ของมือ [15] เปรียบเทียบความผิดพลาดที่เกิดขึ้นเมื่อใช้ขั้นตอนวิธีที่นำเสนอ แสดงเวลาในการคำนวณที่เพิ่มขึ้นกับค่าความผิดพลาดที่ลดลงไป และสุดท้ายเป็นผลการทดลองในส่วนที่ได้ศึกษาถึงแนวทางแก้ปัญหากรณีที่มีการผสมและแยกกันระหว่างใบหน้าและมือ

วิดิทัศน์ภาษาไทยที่นำมาใช้ในการทดลอง ส่วนหนึ่งได้มาจากวิดิทัศน์ประกอบการเรียนภาษาไทยจัดทำโดยกระทรวงศึกษาธิการ [21] ซึ่งเป็นการแสดงท่าทางโดยผู้เชี่ยวชาญด้านการใช้ภาษาไทย และความหมายของคำศัพท์ทั้งหมดเป็นคำที่มักนำมาใช้บ่อยสำหรับการสื่อสารในชีวิตประจำวัน การจัดเก็บข้อมูลทำโดยบันทึกเข้าเครื่องคอมพิวเตอร์โดยผ่านการรับภาพ (Capture card) ยี่ห้อวินโนฟ (Winnov) มีความกว้าง (width) และความสูง (height) ขนาด 320x240 จุดภาพ

ในอัตรา 15 เฟรมต่อวินาที แสดงลำดับวีดิทัศน์ภาษามือไทยในรูปแบบที่ 4.1 มีความยาวของจำนวนเฟรม และลักษณะเด่นของแต่ละวีดิทัศน์ดังนี้

- วีดิทัศน์ที่ 1 จำนวน 117 เฟรม ในรูปที่ 4.1 (ก) เป็นวีดิทัศน์ที่ผู้แสดงภาษามือเป็นผู้ชาย สวมเสื้อแขนสั้น มีการเคลื่อนไหวของมือและแขนทั้งสองข้างในแนวราบและแนวตั้ง
- วีดิทัศน์ที่ 2 จำนวน 74 เฟรม ในรูปที่ 4.1 (ข) เป็นวีดิทัศน์ที่ผู้แสดงภาษามือเป็นผู้ชาย สวมเสื้อแขนสั้น มีการเคลื่อนไหวของมือและแขนขวาเพียงข้างเดียว โดยมือและแขนซ้ายอยู่นิ่งข้างลำตัว
- วีดิทัศน์ที่ 3 จำนวน 95 เฟรม ในรูปที่ 4.1 (ค) เป็นวีดิทัศน์ที่ผู้แสดงภาษามือเป็นผู้หญิง สวมเสื้อแขนยาว มีการเคลื่อนไหวของมือและแขนทั้งสองข้างในแนวตั้งเป็นส่วนมาก
- วีดิทัศน์ที่ 4 จำนวน 95 เฟรม ในรูปที่ 4.1 (ง) เป็นวีดิทัศน์ที่ผู้แสดงภาษามือเป็นผู้หญิง สวมเสื้อแขนยาว มีการเคลื่อนไหวของมือและแขนทั้งสองข้างในแนวราบเป็นส่วนมาก



รูปที่ 4.1 รูปภาพที่ใช้ทดลอง และ บริเวณสีผิวที่ถูกตัดมาใช้เป็นตัวอย่าง ในแต่ละวีดิทัศน์

อีกส่วนหนึ่งของวีดิทัศน์ทดสอบได้จากการเก็บข้อมูลจากกล้องชนิดนี้ รุ่น EVI-D100P แสดงในรูปที่ 4.2 รูปภาพส่วนนี้จะถูกนำไปใช้ทดลองสำหรับการเรียนรู้ของกล้อง (การทดลองที่ 4.2.3) เพื่อสร้างแบบจำลองในการแยกส่วนสีผิว และรูปภาพที่บันทึกมีความกว้าง (width) และความสูง (height) ขนาด 360x288 จุดภาพ การบันทึกภาพกระทำบริเวณห้องวิจัย ซึ่งฉากหลังในภาพจะมีความซับซ้อนเป็นอย่างมาก ทั้งนี้เพื่อใช้ตรวจสอบแบบจำลองที่ใช้ว่ามีความทนทานต่อการรบกวนอย่างไร



รูปที่ 4.2 รูปภาพบางส่วนที่บันทึกจากกล้องโซนี่ รุ่น EVI-D100P

4.1 การตรวจหาบริเวณส่วนสีผิวโดยใช้แบบจำลองชนิดไม่มีตัวแปร (Non Parametric Model)

ในหัวข้อนี้เป็นการทดลองหาบริเวณส่วนสีผิวโดยใช้แบบจำลองชนิดไม่มีตัวแปร โดยมีวัตถุประสงค์เพื่อวัดประสิทธิภาพของแต่ละปริภูมิสีว่า ปริภูมิสีใดมีความเหมาะสมในการแยกส่วนสีผิวมนุษย์มากกว่ากัน โดยเปรียบเทียบใน 3 ปริภูมิสี คือ RGB นอร์แมลซ์ (พิจารณาค่า r และ g) HSI (พิจารณาค่า H และ S) และ YCbCr (พิจารณาค่า Cb และ Cr)

ผลการทดลองแสดงในรูปที่ 4.3 คอลัมน์ที่ 1 คือ การแยกส่วนสีผิวโดยใช้สายตา และใช้เป็นค่าอ้างอิงเพื่อเปรียบเทียบความถูกต้องกับผลของมาส์กจากการหาส่วนสีผิวในปริภูมิสีทั้งสาม ดังแสดงในตารางที่ 4.1-4.4 ผลของมาส์กจากการแยกโดยแบบจำลองชนิดไม่มีตัวแปรในปริภูมิสีย่อย rg (รูปที่ 4.3 คอลัมน์ที่ 2) HS (รูปที่ 4.3 คอลัมน์ที่ 3) และ $CbCr$ (รูปที่ 4.3 คอลัมน์ที่ 4)

วิเคราะห์ผลการทดลองการหาบริเวณส่วนสีผิวโดยใช้แบบจำลองชนิดไม่มีตัวแปร ผลของมาส์กที่ได้ ในปริภูมิสีย่อย $CbCr$ จะให้ผลที่ดีกว่า HS และ rg (ค่า Detection rate ที่สูง และ False alarm rate ต่ำ) สำหรับค่าความผิดพลาดที่เกิดขึ้นในปริภูมิสีย่อย rg แสดงตัวอย่างการรบกวนกันเมื่อมีค่าใกล้เคียงสีดำ ตัวอย่างเช่น ในปริภูมิสี RGB ค่า $(1,0,0)$ และ $(0,0,0)$ เป็นค่าที่ใกล้เคียงกัน แต่เมื่อทำการนอร์แมลไลซ์ไปสู่ rg จะแสดงเป็นค่า $(1,0)$ และ $(0,0)$ ตามลำดับ ซึ่งเป็นค่าที่แตกต่างกันสูงที่สุด

หมายเหตุ ในการทดลองค่าในปริภูมิสีย่อย rg และ HS จะถูกนอร์แมลไลซ์ด้วยค่า 224 เพื่อให้เท่ากับค่าในปริภูมิสีย่อย $CbCr$



รูปที่ 4.3 การแยกบริเวณส่วนสีผิวด้วยสายตา (ก-1) - (ง-1) และผลจากการแยกโดยใช้แบบจำลองชนิดไม่มีตัวแปรในปริภูมิสีย่อย rg (ก-2) - (ง-2), HS (ก-3) - (ง-3) และ CbCr (ก-4) - (ง-4) ตามลำดับ

ตารางที่ 4.1 ค่าความถูกต้องในแต่ละปริภูมิสีย่อยของแบบจำลองชนิดไม่มีตัวแปรเทียบกับ การแยกส่วนสีผิวด้วยสายตาของลำดับวิธีทัศน์ที่ 1

ปริภูมิสีย่อย	True Positive	True Negative	False Positive	False Negative	DR (%)	FAR (%)
rg	3510	71415	265	1610	68.55	7.02
HS	2895	71406	274	2225	56.54	8.65
CbCr	3525	71614	66	1595	68.85	1.84

ตารางที่ 4.2 ค่าความถูกต้องในแต่ละประมุขีย่อยของแบบจำลองชนิดไม่มีตัวแปรเทียบกับ
การแยกส่วนสีผิวด้วยสายตาของลำดับวิธีทัศน์ที่ 2

ประมุขีย่อย	True Positive	True Negative	False Positive	False Negative	DR (%)	FAR (%)
rg	3331	72018	236	1215	73.27	6.62
HS	2799	72118	136	1747	58.86	4.29
CbCr	3087	72202	52	1459	67.91	1.66

ตารางที่ 4.3 ค่าความถูกต้องในแต่ละประมุขีย่อยของแบบจำลองชนิดไม่มีตัวแปรเทียบกับ
การแยกส่วนสีผิวด้วยสายตาของลำดับวิธีทัศน์ที่ 3

ประมุขีย่อย	True Positive	True Negative	False Positive	False Negative	DR (%)	FAR (%)
rg	1486	68326	6173	815	64.58	80.60
HS	1334	68134	6365	967	57.97	82.67
CbCr	1571	74484	15	730	68.27	0.95

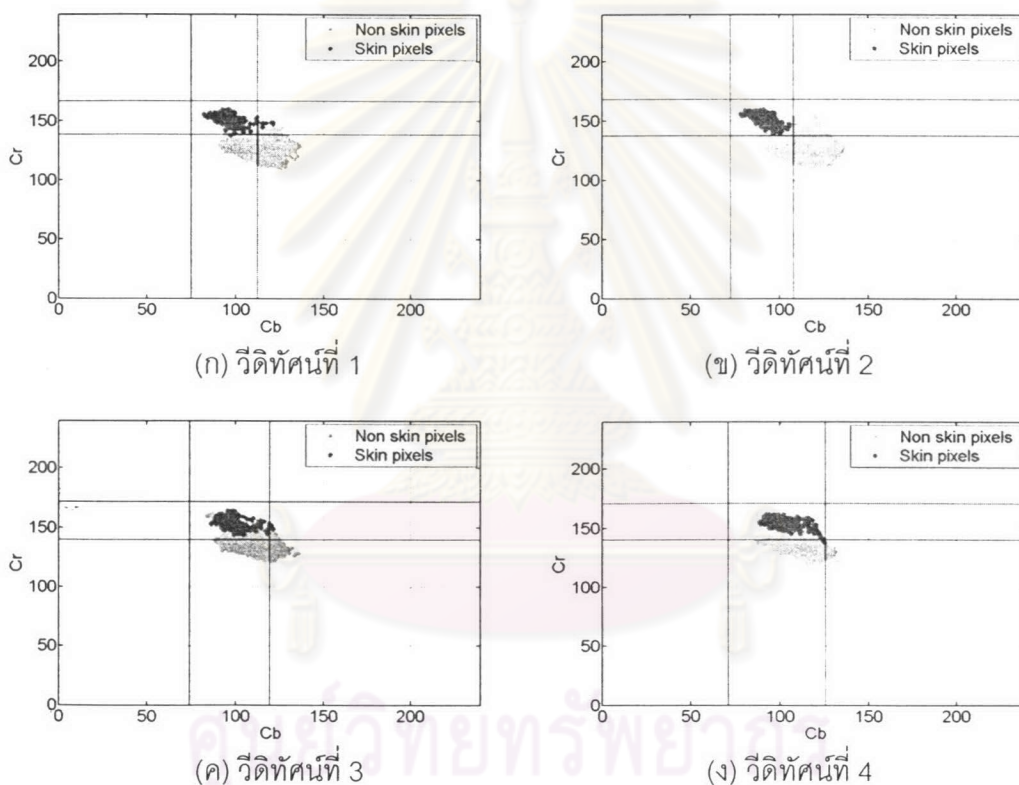
ตารางที่ 4.4 ค่าความถูกต้องในแต่ละประมุขีย่อยของแบบจำลองชนิดไม่มีตัวแปรเทียบกับ
การแยกส่วนสีผิวด้วยสายตาของลำดับวิธีทัศน์ที่ 4

ประมุขีย่อย	True Positive	True Negative	False Positive	False Negative	DR (%)	FAR (%)
rg	1639	72691	1684	786	67.59	50.68
HS	1444	72606	1769	981	59.55	55.06
CbCr	1788	74295	80	637	73.73	4.287

4.2 การแยกบริเวณสีผิวในปริภูมิสีย่อย CbCr

4.2.1 แบบจำลองค่าสูงสุดต่ำสุดของแต่ละองค์ประกอบ

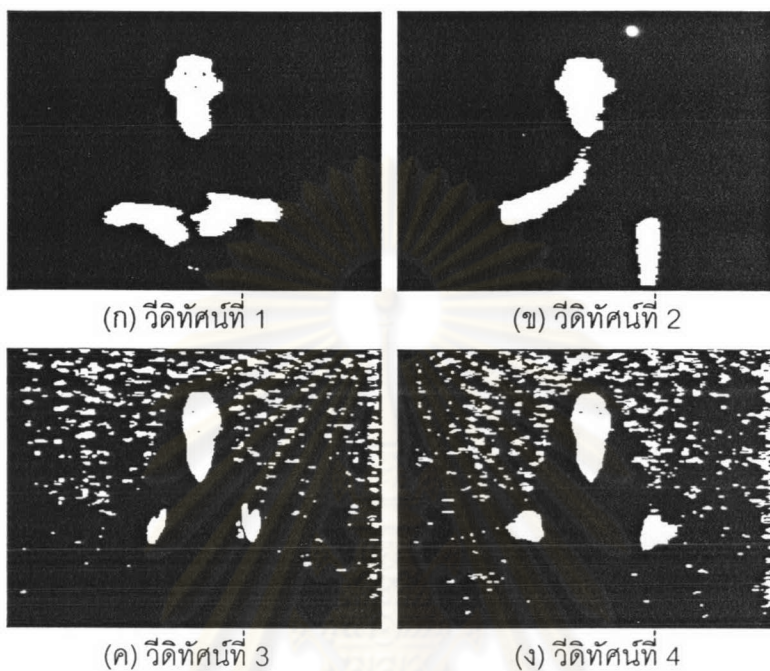
การแยกส่วนสีผิวซึ่งนำเสนอโดย S.L. Phung [13] จะพิจารณาแยกกันระหว่างองค์ประกอบ CbCr เพื่อทำการกำหนดขีดแบ่ง ซึ่งขอบเขตของแบบจำลองในระนาบ CbCr แสดงในรูปที่ 4.4 และผลลัพธ์ของมาสก์ที่ได้แสดงในรูปที่ 4.5



รูปที่ 4.4 การแยกส่วนสีผิวด้วยแบบจำลองซึ่งนำเสนอโดย S.L. Phung [13]

วิเคราะห์ผลการทดลอง จากรูปที่ 4.5 และตารางที่ 4.5 จะเห็นว่ารูปภาพทดสอบจากวิดีโอที่ 1 และ 2 ให้ค่า DR และ ค่า FAR ที่พอยอมรับได้ ส่วนรูปภาพจากวิดีโอที่ 3 และ 4 ค่า FAR มีค่าผิดพลาดมากเกินไปที่จะนำแบบจำลองนี้ไปใช้ในการแยกส่วนสีผิวอย่างมีประสิทธิภาพ เนื่องจากมีฉากหลังบางบริเวณที่มีค่า Cb และ Cr ใกล้เคียงกับค่าของสีผิว ทำให้แบบจำลองตีความหมายผิดพลาด และผลกระทบนี้จะยิ่งเพิ่มขึ้นกรณีที่ฉากหลังมีความซับซ้อนมาก ๆ เพราะมีโอกาสที่จะพบ

จุดภาพที่ใกล้เคียงสีผิวมากขึ้น ซึ่งแนวทางในการแก้ปัญหาคือสร้างแบบจำลองให้ใกล้เคียงกับข้อมูลสีผิวมากที่สุด เพื่อหลีกเลี่ยงความผิดพลาดในส่วนนี้



รูปที่ 4.5 ผลการแยกบริเวณส่วนสีผิวโดยแบบจำลองค่าสูงสุดต่ำสุด [13] เมื่อกำหนดขีดแบ่งเท่ากับ 4

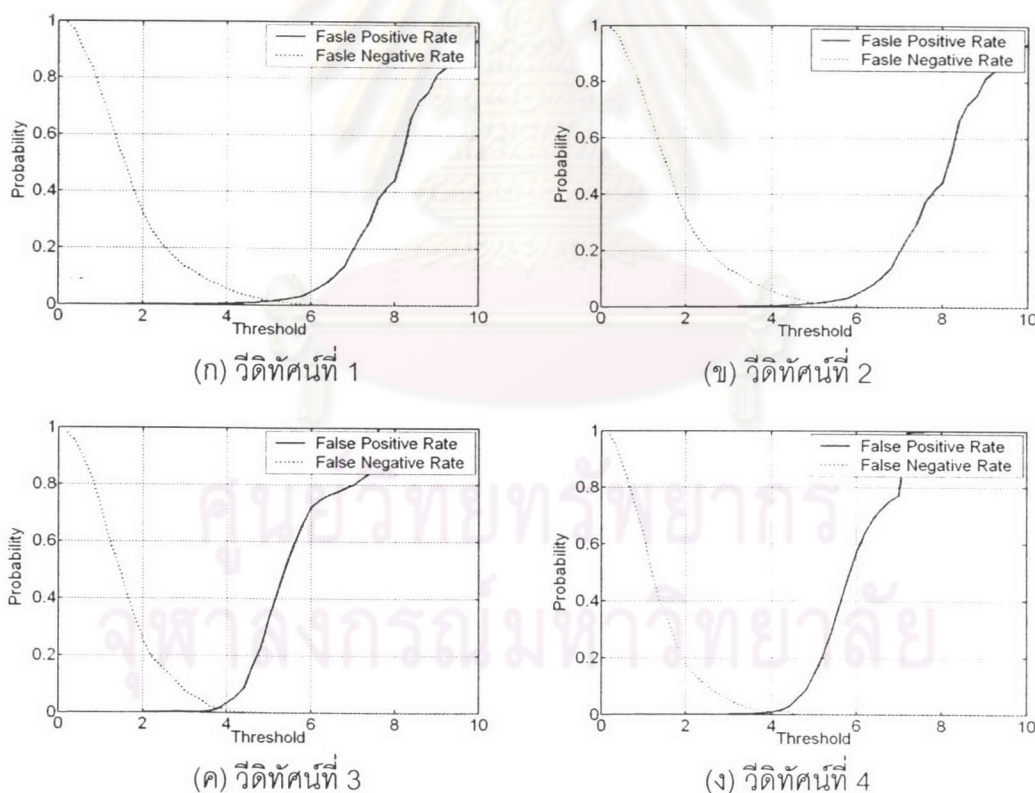
ตารางที่ 4.5 ค่าความถูกต้องของแบบจำลองค่าสูงสุดต่ำสุด [13] เทียบกับการแยกส่วนสีผิวด้วยสายตา

ลำดับ วิตทัศน์	True Positive	True Negative	False Positive	False Negative	DR (%)	FAR (%)
1	4882	71179	501	238	95.35	9.31
2	4200	71924	330	346	92.39	7.28
3	2286	69816	4683	15	99.35	67.20
4	2142	68449	5926	13	99.46	71.07

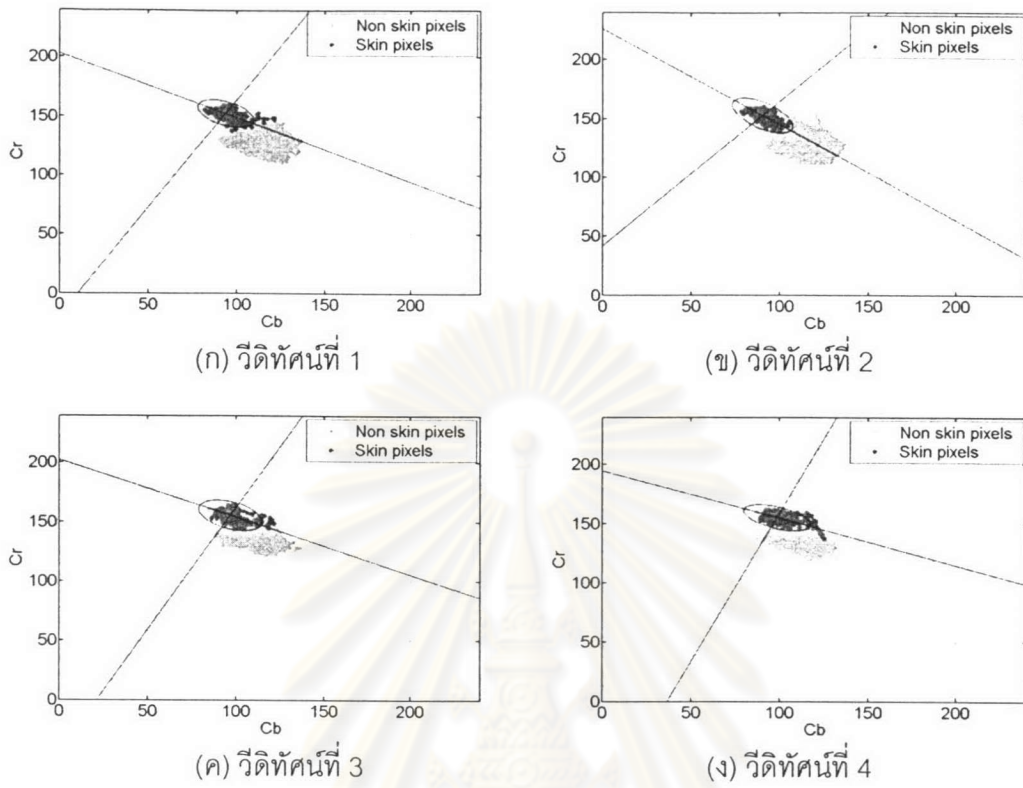
4.2.2 แบบจำลองวงรีในปริภูมิสี่อ้อย CbCr

กรรมวิธีที่นำมาใช้ในวิทยานิพนธ์นี้ คือ การหาแบบจำลองวงรีที่ครอบคลุมบริเวณสีผิวในปริภูมิสี่อ้อย CbCr โดยรายละเอียดได้กล่าวไว้ในบทที่ 3 และเพื่อให้ได้มาซึ่งแบบจำลองวงรีที่เหมาะสมที่สุดในการแยกส่วนสีผิวตัวแปรสำคัญอันหนึ่งที่ใช้ในการกำหนดอาณาบริเวณของวงรี คือ ขีดแบ่ง η ดังนั้นขั้นตอนเบื้องต้นของการสร้างแบบจำลองวงรีจึงได้แสดงวิธีการที่ใช้ในการหาขีดแบ่ง η ที่เหมาะสม โดยมีตัวแปรอีก 2 ตัวที่ถูกนำมาพิจารณา คือ ค่า False Positive Rate และ ค่า False Negative Rate ในการทดลองจะปรับขีดแบ่ง η เพิ่มขึ้นครั้งละ 0.2 ตั้งแต่ค่า 0.2 ถึง 10 ที่จุดตัดของกราฟทั้งสองคือจุดซึ่งเส้นโค้ง (curve) ของฉากหน้า (ในที่นี้คือส่วนสีผิว) และฉากหลัง (ส่วนอื่น ๆ) ตัดกันนั่นเอง

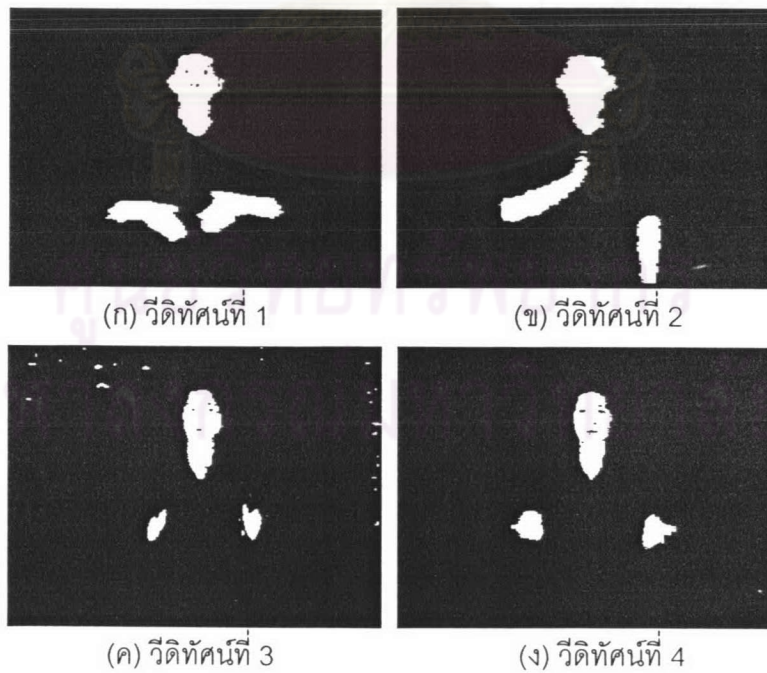
จากการทดลองพบว่าขีดแบ่ง η ที่ทำให้กราฟตัดกันของแต่ละวิถีทัศนียภาพ มีค่าอยู่ในช่วงระหว่าง 3.8 – 5.4 ดังรูปที่ 4.6



รูปที่ 4.6 การหาขีดแบ่งที่เหมาะสมในแต่ละรูปภาพ



รูปที่ 4.7 การกระจายตัวของจุดภาพสีผิวในปริภูมิสีย่อย CbCr ของแต่ละวิธีที่ศนทดสอบ



รูปที่ 4.8 ผลการแยกบริเวณสีผิวโดยใช้แบบจำลองวงรี เมื่อกำหนดขีดแบ่ง เท่ากับ 4

จากผลการทดลองจะเห็นว่าแบบจำลองวงรีที่เราใช้ ให้ค่าความถูกต้องในการแยกส่วนสีผิวได้ดีพอสมควร รูปที่ 4.8 แสดงผลของมาสก์ที่ได้เมื่อกำหนดขีดแบ่งเท่ากับ 4 ซึ่งจะมีจุดภาพเพียงส่วนน้อยเท่านั้นที่แบบจำลองตีความหมายผิดไป ซึ่งสามารถแก้ไขได้โดยใช้ ตัวดำเนินการแบบปิด (ไดเลชันและอีโรชัน) ส่วนค่าความถูกต้องของแบบจำลองวงรีเทียบกับการแยกส่วนสีผิวด้วยสายตา แสดงในตารางที่ 4.6

ตารางที่ 4.6 ค่าความถูกต้องของแบบจำลองวงรีเทียบกับการแยกส่วนสีผิวด้วยสายตา

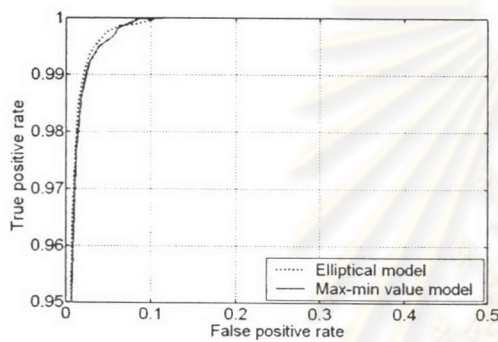
ลำดับ วีดิทัศน์	True	True	False	False	DR (%)	FAR (%)
	Positive	Negative	Positive	Negative		
1	4598	71503	177	522	89.80	3.71
2	4111	71989	265	435	90.43	6.06
3	2178	74331	168	123	94.65	7.16
4	2299	74238	137	126	94.80	5.62

วิเคราะห์ผลการทดลองการเปรียบเทียบระหว่างแบบจำลองทั้งสอง การใช้แบบจำลองวงรีจะครอบคลุมข้อมูลค่าจุดภาพบริเวณส่วนสีผิวในปริภูมิสีย่อย CbCr มากกว่าวิธีเดิมซึ่งพิจารณาค่า Cb และค่า Cr แต่ละองค์ประกอบแยกกัน ในตารางที่ 4.7 แสดงผลของค่า DR และค่า FAR ที่ลดลงเปรียบเทียบระหว่างขั้นตอนวิธีดั้งเดิม กับ ขั้นตอนวิธีแบบจำลองวงรีที่นำเสนอ เมื่อกำหนดขีดแบ่งเท่ากัน คือ 4 ซึ่งหากพิจารณาโดยรวมทั้ง 4 รูปภาพทดสอบ ขั้นตอนวิธีของแบบจำลองวงรี จะให้ผลของ ค่า DR ลดลงเพียงเล็กน้อยเท่านั้น และสามารถทำให้ค่า FAR ซึ่งเป็นความผิดพลาดที่เราไม่ต้องการลดลงไปอย่างมาก

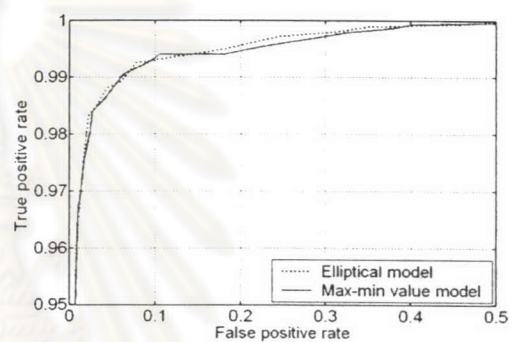
ตารางที่ 4.7 ค่า DR และค่า FAR ที่ลดลง เปรียบเทียบระหว่างแบบจำลองทั้งสอง

ลำดับวีดิทัศน์	ค่า DR (%) ลดลง	ค่า FAR (%) ลดลง
1	5.55	5.60
2	1.96	1.22
3	4.70	60.04
4	4.66	65.45

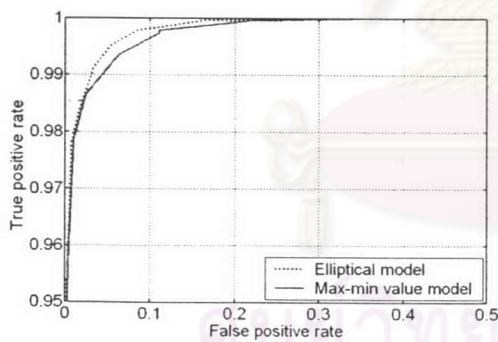
การเปรียบเทียบประสิทธิภาพของแบบจำลองวงรีกับแบบจำลองค่าสูงสุดต่ำสุด [13] ด้วยกราฟ ROC แสดงในรูปที่ 4.9 จากผลการทดลองพบว่า ในวิถีทัศนที่ 1 และ 2 เส้นโค้ง ROC มีค่าใกล้เคียงกันมากแต่หากพิจารณาอย่างถี่ถ้วนแล้วแบบจำลองวงรีที่เสนอมีเส้นโค้ง ROC ที่เหนือกว่าแบบจำลองค่าสูงสุดต่ำสุดเป็นส่วนใหญ่ ส่วนในวิถีทัศนที่ 3 และ 4 เส้นโค้ง ROC ของแบบจำลองวงรีมีค่าเหนือกว่าอย่างเห็นได้ชัด เพราะการใช้แบบจำลองค่าสูงสุดต่ำสุดสำหรับ 2 วิถีทัศนนี้จะเกิดความผิดพลาดในการตีความหมายของแบบจำลอง เนื่องจากจากหลังมีค่า C_b และ C_r ใกล้เคียงกับสี่ผิวด แต่แบบจำลองวงรีที่เสนอสสามารถขจัดความผิดพลาดในส่วนนี้ได้



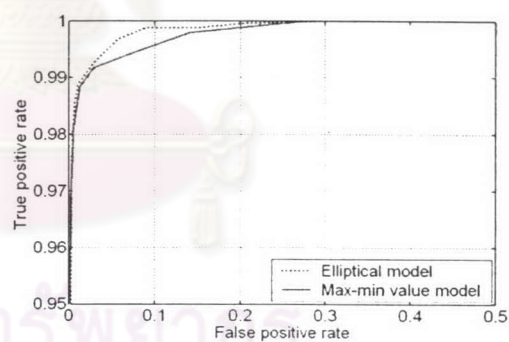
(ก) วิถีทัศนที่ 1



(ข) วิถีทัศนที่ 2



(ค) วิถีทัศนที่ 3



(ง) วิถีทัศนที่ 4

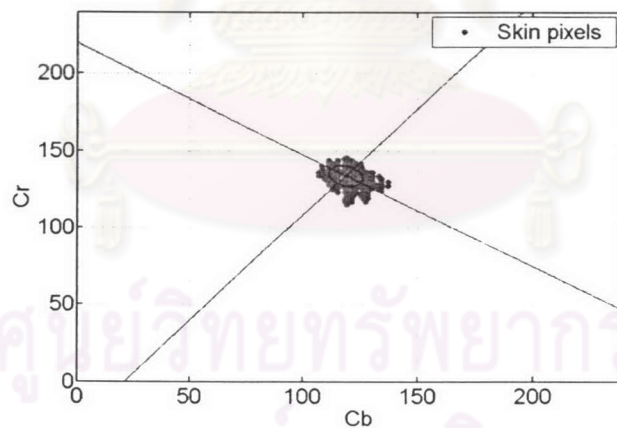
รูปที่ 4.9 การเปรียบเทียบประสิทธิภาพของแบบจำลองค่าสูงสุดต่ำสุด [13] กับแบบจำลองวงรีด้วยกราฟ ROC

4.2.3 การเรียนรู้สีผิวของกล้องโซนี่ (Sony) รุ่น EVI-D100P



รูปที่ 4.10 บริเวณสีผิวที่ใช้เป็นตัวอย่าง

การทดลองในหัวข้อนี้จะนำรูปภาพทดสอบที่ได้จากการบันทึกโดยกล้องโซนี่ รุ่น EVI-D100P กล้องมีความสามารถในการส่าย (pan) และ ก้มเงย (tilt) การทดลองจะแสดงถึงประสิทธิภาพของแบบจำลองวงรีเมื่อนำไปใช้ในการแยกส่วนสีผิว โดยมีขั้นตอนของการเรียนรู้ค่าสีผิวของกล้อง คือ เก็บค่าสีผิวบริเวณใบหน้าจากกลุ่มตัวอย่างทั้งหมด 18 ตัวอย่าง ในบริเวณใกล้เคียงกันและมีสภาพแวดล้อมที่ค่าแสงแตกต่างกันไม่มาก เพื่อนำสร้างแบบจำลองวงรีที่ใช้แยกส่วนสีผิว แต่กรณีที่ต้องนำไปใช้งานจริงอาจต้องมีการเก็บตัวอย่างที่มากกว่านี้เพื่อให้แบบจำลองมีความแม่นยำมากขึ้น



รูปที่ 4.11 การกระจายตัวบนปริภูมิสีย่อย CbCr ของบริเวณสีผิวตัวอย่าง

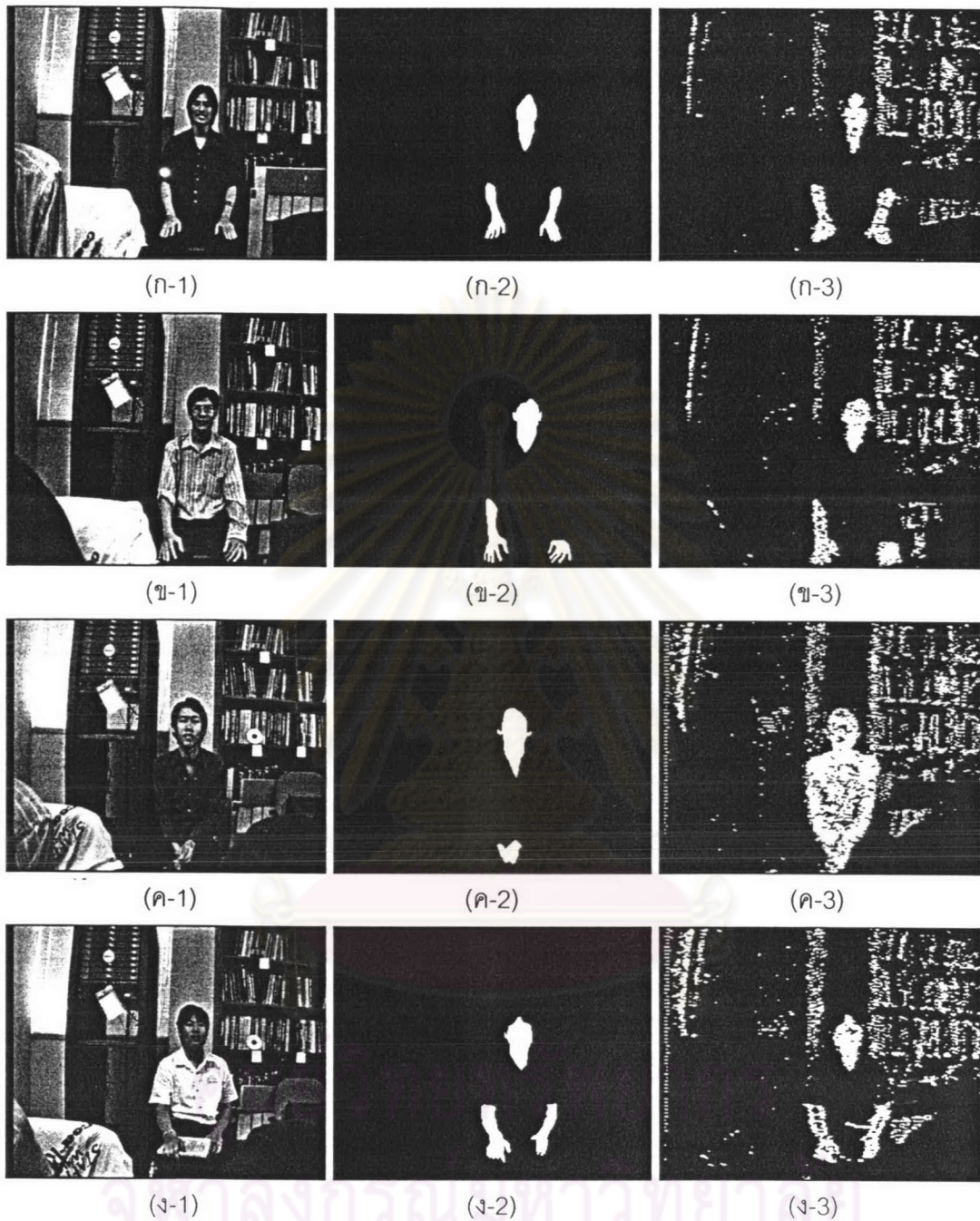
เมื่อกำหนดค่า $\eta = 2$ ในรูปที่ 4.11 แสดงวงรีที่ครอบคลุมจุดภาพบริเวณส่วนสีผิว

จะได้ จุดศูนย์กลางวงรี $(h,k) = (118.54, 133.85)$

ค่าความยาวแกนเอก $a = 13.69$

ค่าความยาวแกนโท $b = 7.55$

ความชันของแกนเอก $m = -0.73$



รูปที่ 4.12 ผลของมาสก์ที่ได้เมื่อแยกสีผิวด้วยแบบจำลองวงรีเมื่อกำหนดค่า $\eta = 2$

ในรูปที่ 4.12 คอลัมน์ที่ 1 คือ รูปภาพที่นำมาทำการทดสอบการแยกส่วนสีผิว คอลัมน์ที่ 2 คือ มาสก์การแยกสีผิวเมื่อพิจารณาด้วยสายตาของผู้วิจัย และ คอลัมน์ที่ 3 คือ มาสก์การแยกสีจากแบบจำลองวงรีที่ใช้

รูปที่ 4.12 (ก), (ข) และ (ค) เป็นรูปของบุคคลที่อยู่ในกลุ่มตัวอย่างซึ่งได้มีการเก็บค่าสีผิว ซึ่งค่าความผิดพลาดที่เห็นได้ชัดเจนจะเกิดกับรูปที่ 4.12 (ค) สาเหตุเพราะสีเสื้อบุคคลที่อยู่ในภาพมีโทนสีแดงเข้ม ซึ่งค่อนข้างใกล้เคียงกับค่า CbCr ของสีผิว จึงถูกตัดสินใจให้อยู่ภายในวงรีและแบบจำลองกำหนดให้เป็นสีผิว ส่วนรูปที่ 4.12 (ง) เป็นการนำรูปซึ่งไม่ได้มีการเก็บค่าตัวอย่างของสีผิวมาก่อนหน้าทำการทดสอบ แบบจำลองวงรีที่ใช้ก็ยังให้ค่าของบริเวณสีผิวเป็นส่วนใหญ่

หมายเหตุ มาสก์ของผลการทดลองในรูปที่ 4.12 นี้ ยังไม่ผ่านตัวดำเนินการแบบปิด และการใช้คุณลักษณะของขนาด (Size feature) มาช่วยในการตัดสินใจ ดังนั้นในรูปที่ 4.12 (ก) (ข) และ (ง) จึงสามารถให้ค่าความถูกต้องที่ดีกว่านี้ได้ หากใช้ตัวดำเนินการแบบปิด (เติมเต็มในส่วนบริเวณใบหน้าและมือ) และการตัดสินใจเลือกขนาดที่เหมาะสม (กำจัดค่ารบกวนของฉากหลัง)

ตารางที่ 4.8 ค่าความถูกต้องของแบบจำลองวงรีเทียบกับการแยกส่วนสีผิวด้วยสายตา
เมื่อวีดิทัศน์รับเข้าได้จากการบันทึกด้วยกล้องชนิด EVI-D100P

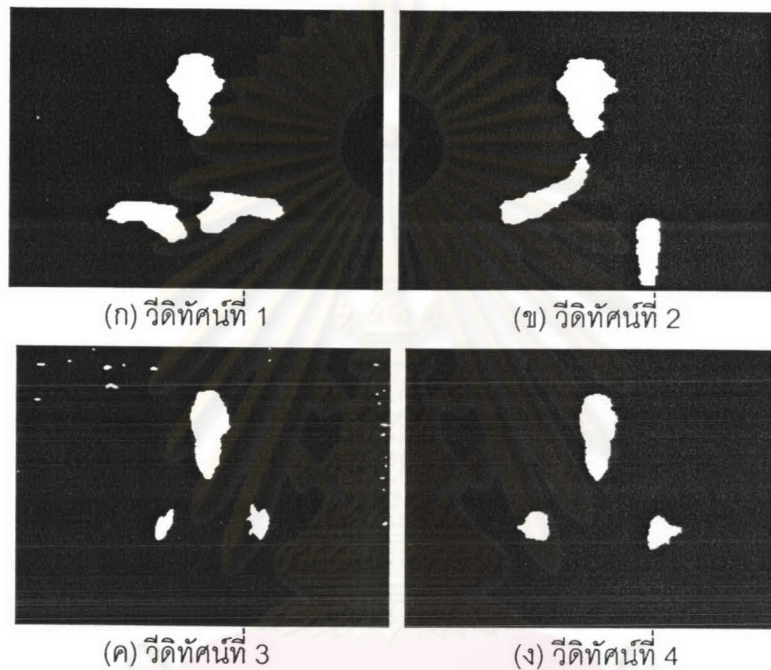
รูปภาพ	True	True	False	False	DR (%)	FAR (%)
	Positive	Negative	Positive	Negative		
4.11 (ก)	2032	97200	4044	404	83.42	66.56
4.11 (ข)	2030	97382	3855	413	83.09	65.51
4.11 (ค)	1667	89299	12421	293	85.05	88.17
4.11 (ง)	2312	94482	6349	537	81.15	73.31

ปัญหาของการแยกส่วนสีผิวด้วยแบบจำลองวงรี เมื่อนำไปใช้พัฒนาระบบแปลความหมายภาษามือจริง ซึ่งการแก้ปัญหาจำเป็นที่จะต้องนำคุณลักษณะอื่นมาพิจารณาเพิ่มเติมนอกเหนือจากสีผิว ตัวอย่างของปัญหาที่พบ ได้แก่

1. ปัญหาเนื่องจากความซับซ้อนของฉากหลัง ทำให้เกิดค่าผิดพลาดแบบ FP มาก ส่งผลให้ค่า FAR ที่คำนวณออกมามีค่าสูงมาก
2. ปัญหาเรื่องสีของเครื่องแต่งกาย หากเกิดกรณีที่เสื้อผ้าที่สวมใส่ของบุคคลในภาพมีค่า CbCr ที่ใกล้เคียงกับบริเวณสีผิวแบบจำลองที่นำมาใช้ก็จะตีความหมายผิดพลาดไป
3. ปัญหาเรื่องค่าแสงรบกวน การรบกวนของแสงยังคงมีผลกระทบกับแบบจำลองที่ใช้อยู่บ้าง แม้จะละทิ้งองค์ประกอบความสว่าง (Y) ไปแล้วก็ตาม

4.3 ผลจากตัวดำเนินการแบบปิด

ตัวดำเนินการแบบปิด หรือ การทำไดเลชันและอีโรชันตามลำดับ สำหรับรูปภาพทดสอบจะใช้โครงสร้างส่วนย่อยจากสมการที่ 3.13 ในบทที่ 3 ซึ่งเป็นโครงสร้างส่วนย่อยที่ให้ผลดีที่สุดจากการทดลองเมื่อเทียบกับแบบอื่น ๆ ผลที่ได้แสดงในรูปที่ 4.13



รูปที่ 4.13 ผลจากการใช้ตัวดำเนินการแบบปิด

วิเคราะห์ผลการทดลอง การใช้ตัวดำเนินการแบบปิด จะทำให้ได้มาส์กที่สมบูรณ์มากยิ่งขึ้น ส่วนค่าความผิดพลาดที่เกิดขึ้นในรูปที่ 4.13 (ค) ที่บริเวณฉากหลัง มีสาเหตุมาจากค่าสีบริเวณนั้นมีค่าที่ตรงกับบริเวณส่วนสีผิว ทำให้แบบจำลองในการแยกสีผิวตีความหมายผิดไป และ บริเวณส่วนนั้นใหญ่เกินกว่าที่ตัวดำเนินการแบบปิด จะแก้ไขความผิดพลาดได้ แต่ปัญหานี้จะสามารถแก้ไขได้โดยใช้การตัดสินใจเลือกขนาดที่ใหญ่ที่สุด 3 บล็อกเพื่อกำหนดเป็นใบหน้าและมือ ฉะนั้นในเฟรมถัดไปแบบจำลองการแยกส่วนสีผิวก็จะสามารถตีความหมายได้ถูกต้อง

4.4 ผลการแยกองค์ประกอบสำคัญต่าง ๆ

ในหัวข้อนี้จะแสดงผลการแยกองค์ประกอบสำคัญต่าง ๆ ซึ่งประกอบไปด้วย ผลการแยกบริเวณส่วนใบหน้าจากคอ ผลการแยกองค์ประกอบสำคัญบริเวณใบหน้า และผลการแยกจุดเชื่อมต่อระหว่างแขนกับข้อมือ

4.4.1 ผลการแยกบริเวณส่วนใบหน้าจากคอ



รูปที่ 4.14 การแยกบริเวณส่วนใบหน้าจากคอ โดยกรรมวิธีตรวจหาขอบแบบไขว้ศูนย์

วิเคราะห์ผลการทดลอง จากกรรมวิธีในการแยกส่วนใบหน้าจากคอที่นำเสนอ แสดงผลดังรูปที่ 4.14 ซึ่งสามารถแยกบริเวณใบหน้าของผู้แปลภาษามือทั้งสองท่านที่นำมาทดสอบได้อย่างถูกต้อง รูปขั้วมือสุดเป็นผลของการตรวจหาขอบด้วยวิธีไขว้ศูนย์ ถัดมาเป็นการเติมเต็มในส่วนที่ถูกล้อมรอบ และทำการเลือกบริเวณที่ใหญ่ที่สุดเป็นใบหน้า

4.4.2 ผลการแยกองค์ประกอบสำคัญบริเวณใบหน้า



รูปที่ 4.15 ผลการแยกองค์ประกอบสำคัญบนใบหน้าโดยใช้ความแตกต่างของความสว่าง

เนื่องจากบริเวณที่เป็นส่วนสำคัญบนใบหน้า เช่น ตา จมูก ปาก และ คิ้ว มีองค์ประกอบของความสว่างที่ต่ำ การใช้ความแตกต่างขององค์ประกอบความสว่าง (Y) จะทำให้ได้ผลลัพธ์ในส่วนเหล่านี้ของผู้แปลภาษามือออกมาทั้งหมด ดังแสดงในรูปที่ 4.15

4.4.3 ผลการแยกจุดเชื่อมต่อระหว่างแขนกับข้อมือ

กรณีผู้แปลภาษามือสวมเสื้อแขนสั้น กระบวนการอินนิงจะถูกเสนอให้ใช้กับบล็อปที่เป็นมือทั้งสองข้าง เพื่อนำมากำหนดบริเวณมือหรือแขนของผู้แปลภาษามือ ซึ่งให้ผลลัพธ์ดังแสดงในรูปที่ 4.16



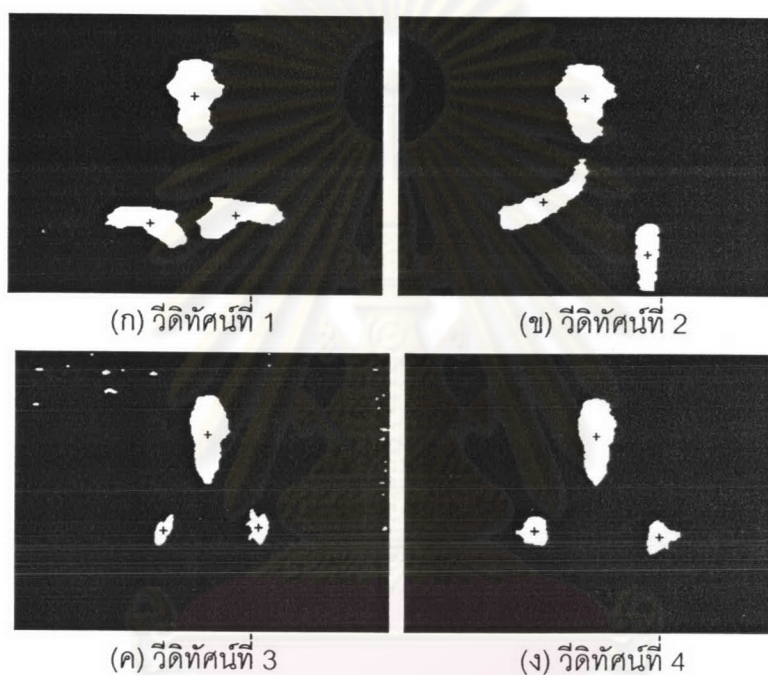
รูปที่ 4.16 ผลจากกระบวนการอินนิง

วิเคราะห์ผลการทดลอง วิดีทัศน์ที่ 1 แต่ละเฟรมมีรายละเอียดส่วนมือที่พอเพียงเมื่อทำอินนิงแล้วจะเกิดรอยแยกอย่างชัดเจน แต่ในวิดีโอทัศน์ที่ 2 เฟรมที่ 1 และ 40 จะไม่เห็นรายละเอียดของมืออย่างชัดเจนเมื่อใช้กระบวนการอินนิงแล้วยังยากต่อการตัดสินใจส่วนมือ ส่วนในเฟรมที่ 30 สังเกตได้ว่าที่มือขวาของผู้แปลมีรอยแยกมากกว่าส่วนอื่นในบล็อปเดียวกัน ทำให้สามารถกำหนดบริเวณมือได้อย่างแม่นยำ

หมายเหตุ การทำอินนิงจะให้ประสิทธิภาพที่ดีหากวิดีโอทัศน์ที่นำมาใช้มีความคมชัดสูง เพราะจะเห็นรายละเอียดของมือที่ชัดเจนซึ่งจะเกิดรอยแยกที่มากและสามารถกำหนดขีดแบ่งได้ง่าย

4.5 ผลจากการวิเคราะห์หาบริเวณที่เชื่อมต่อกันและหาจุดศูนย์กลางถ่วงของแต่ละวัตถุ

ในหัวข้อนี้ จะทำการวิเคราะห์หาบริเวณที่เชื่อมต่อกัน เพื่อทราบให้ถึงตำแหน่งของแต่ละบล็อบ ด้วยวิธีเข้ารหัสแบบลูกโซ่ (Chain code) โดยพิจารณาความสัมพันธ์กับจุดภาพรอบข้าง 8 จุดภาพ และหาจุดศูนย์กลางถ่วงของแต่ละบล็อบ เพื่อนำมาเป็นตัวแทนสำหรับการพิจารณาการเคลื่อนที่และใช้ประโยชน์ในการติดตามต่อไป



รูปที่ 4.17 ผลจากการวิเคราะห์หาบริเวณที่เชื่อมต่อกัน และหาจุดศูนย์กลางถ่วงของแต่ละวัตถุ

การวิเคราะห์หาบริเวณที่เชื่อมต่อกันกระทำเพื่อจัดกลุ่มแต่ละบล็อบเชิงพื้นที่ เพื่อนำตำแหน่งของจุดภาพสีผิวเหล่านั้นมาคำนวณหาจุดศูนย์กลางถ่วง แทนด้วยเครื่องหมาย "+" ในรูปที่ 4.17 ผลของจุดศูนย์กลางถ่วงที่ได้ของบล็อบบริเวณใบหน้าจะอยู่กึ่งกลางใกล้เคียงกับจมูกหรือปาก ส่วนบล็อบบริเวณมือ ในกรณีผู้แปลภาษามือสวมเสื้อแขนสั้นจะอยู่ส่วนแขนจะค่อนข้างมาทางมือ แต่ในกรณีสวมเสื้อแขนยาวจะอยู่กลางมือพอดี

สำหรับตั้งแต่หัวข้อที่ 4.6 เป็นต้นไป จะแสดงผลการทดลองในส่วนการติดตาม ซึ่งจะได้แสดงถึงขั้นตอนวิธีที่นำเสนอว่าให้ผลลัพธ์ที่มีประสิทธิภาพมากขึ้นจากเดิมเพียงใด

4.6 การเคลื่อนที่ไปยังตำแหน่งต่าง ๆ ของจุดศูนย์กลางแต่ละบริเวณ

แนวทางการเคลื่อนที่ของจุดศูนย์กลาง (Trajectory) ของแต่ละวิถีทัศนียภาพแสดงในรูปที่ 4.18 ซึ่งกราฟแสดงการเปลี่ยนตำแหน่งของจุดศูนย์กลางของใบหน้าและมือในทุก ๆ เฟรม

รูปที่ 4.18 ในทุกลำดับวิถีทัศนียภาพบริเวณใบหน้า จะมีการเปลี่ยนตำแหน่งน้อยมาก

รูปที่ 4.18 (ก) มีการเคลื่อนที่ของมือทั้งสองข้าง จุดศูนย์กลางของวัตถุจะเปลี่ยนแปลงในแนวตั้งมากกว่าแนวราบ จากการวิเคราะห์กราฟมือทั้งสองข้างจะเคลื่อนที่ในลักษณะคล้ายกัน เพราะมีตำแหน่งใกล้เคียงกันในตามแนวความสูง และมีความสมมาตรตามแนวความกว้าง

รูปที่ 4.18 (ข) มีการเคลื่อนที่เฉพาะมือขวาเพียงอย่างเดียว มือซ้ายอยู่นิ่ง จากกราฟแสดงให้เห็นว่าไม่มีการเปลี่ยนตำแหน่ง

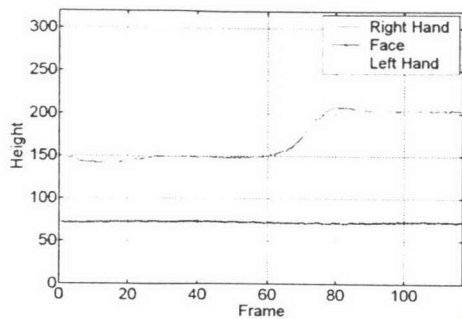
รูปที่ 4.18 (ค) มือทั้งสองมีการเคลื่อนที่ในลักษณะคล้ายกัน มีการเปลี่ยนตำแหน่งตามแนวความสูงของภาพเป็นส่วนมาก

รูปที่ 4.18 (ง) มือทั้งสองมีการเคลื่อนที่ในลักษณะคล้ายกัน มีการเปลี่ยนตำแหน่งตามแนวความกว้างของภาพเป็นส่วนมาก

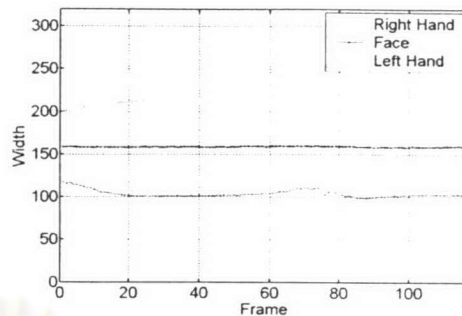


รูปที่ 4.18 แนวทางการเคลื่อนที่ของจุดศูนย์กลาง

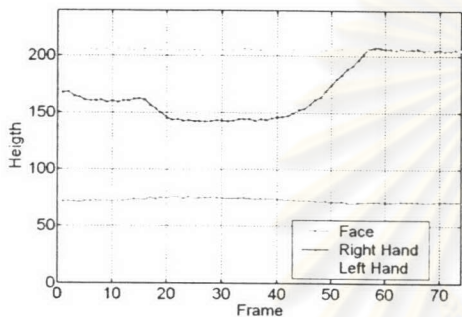
ส่วนรูปที่ 4.19 พิจารณาตามความสูง (คอลัมน์ที่ 1) และความกว้าง (คอลัมน์ที่ 2) ของภาพ



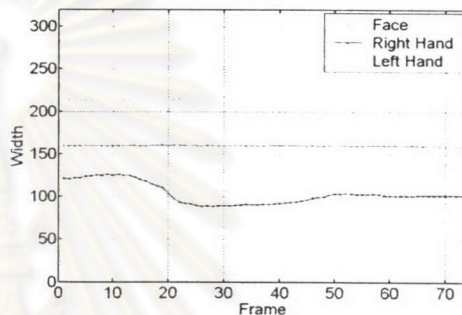
(ก-1)



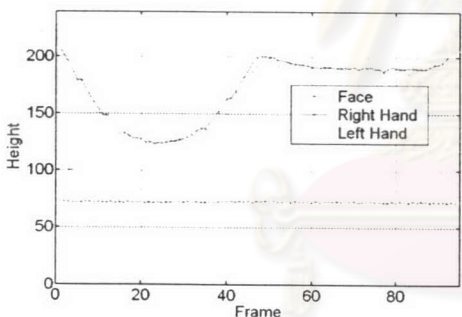
(ก-2)



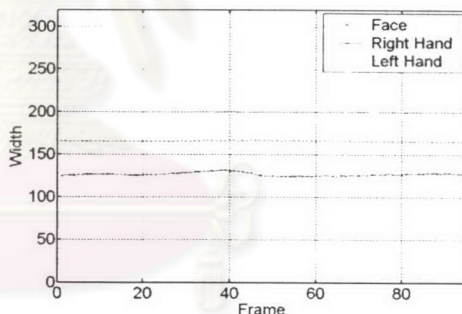
(ข-1)



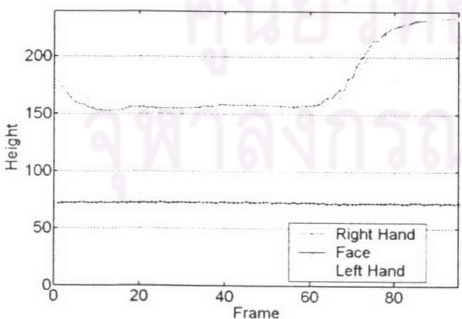
(ข-2)



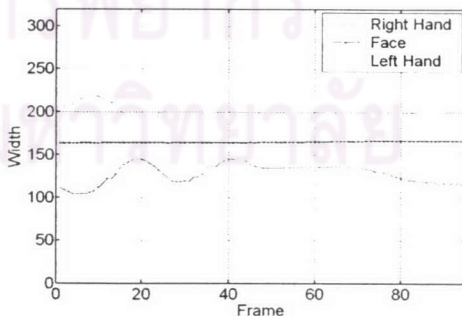
(ค-1)



(ค-2)



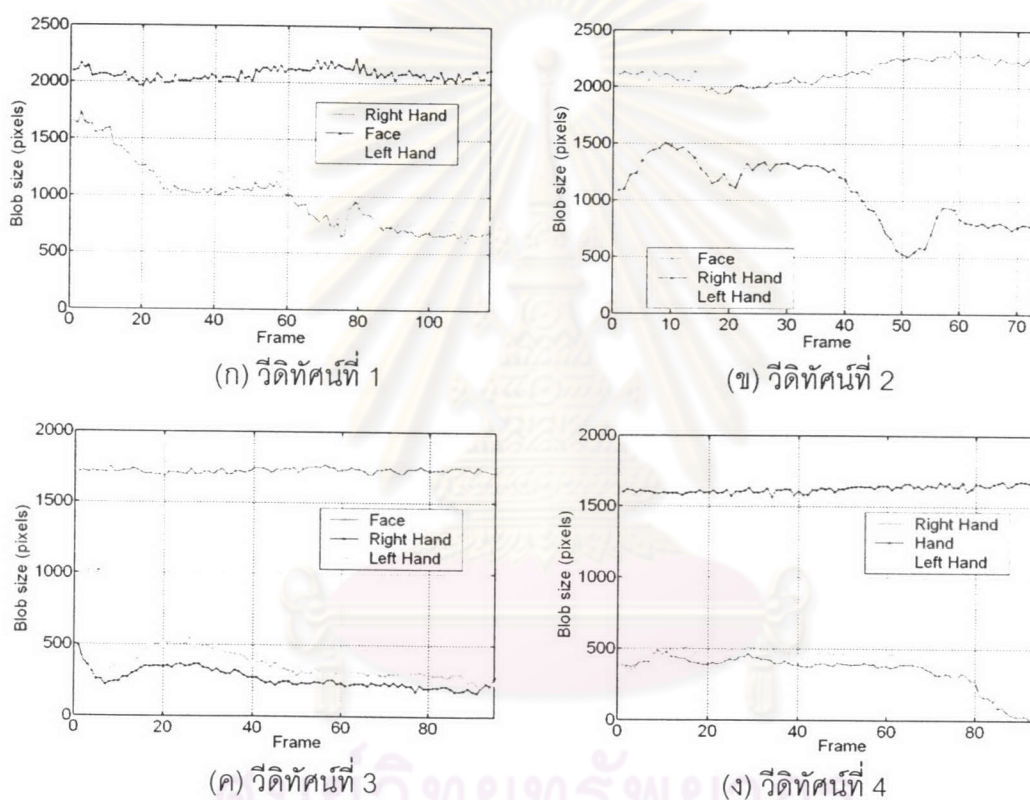
(ง-1)



(ง-2)

รูปที่ 4.19 การเปลี่ยนตำแหน่งของจุดศูนย์กลางถ่วงของแต่ละบลิบ

จากการวิเคราะห์การเปลี่ยนตำแหน่งของใบหน้าและมือในแนวความสูงและแนวความกว้างของภาพ (รูปที่ 4.19 คอลัมน์ที่ 1 และ 2 ตามลำดับ) เราสามารถนำตำแหน่งที่ได้ไปใช้ในการจำแนกใบหน้าและมืออย่างคร่าว ๆ ได้ โดยพิจารณาการเปลี่ยนตำแหน่งของจุดศูนย์กลางของแต่ละวัตถุ หากวัตถุได้มีการเปลี่ยนตำแหน่งมากจะกำหนดเป็นมือ และ หากวัตถุได้มีการเปลี่ยนตำแหน่งน้อยจะเป็นใบหน้า หากมีมากกว่าหนึ่งวัตถุที่เปลี่ยนตำแหน่งน้อยใกล้เคียงกัน ก็สามารถจำแนกจากตำแหน่งที่วัตถุนั้น ๆ อยู่ กล่าวคือ บริเวณใบหน้าควรจะอยู่กึ่งกลางของภาพและอยู่เหนือมือทั้งสองข้าง



รูปที่ 4.20 ขนาดของบล็อบในแต่ละเฟรม

วิเคราะห์ผลการทดลอง การตรวจสอบขนาดของแต่ละบล็อบในทุก ๆ เฟรม จากวิดีโอที่ 3 และ 4 พบว่าส่วนของใบหน้าและมือมีขนาดที่ต่างกันมาก เนื่องจากผู้แปลภาษามือสวมเสื้อแขนยาว จึงสามารถใช้คุณลักษณะนี้ทำจำแนกได้ว่าส่วนใดเป็นใบหน้าและส่วนใดเป็นมือ แต่ในกรณีของวิดีโอที่ 1 และ 2 ซึ่งผู้แปลภาษามือสวมเสื้อแขนสั้น การจำแนกโดยใช้คุณลักษณะของขนาดอย่างเดียว อาจทำให้เกิดความผิดพลาดได้ อาจจำเป็นต้องอาศัยคุณลักษณะอื่นมาช่วยในการตัดสินใจเพิ่มเติม

4.7 ผลการทดลองการวัดความเป็นวงกลม

ในกรณีที่ผู้แปลภาษามือสวมเสื้อแขนสั้น ซึ่งการใช้ขนาดในตัดสินค้าแยกบริเวณมือและแขนเชิงพื้นที่เพียงอย่างเดียว อาจเกิดความผิดพลาดได้ กรณีที่นำไปใช้งานกับระบบจริง เพื่อให้ระบบตัดสินค้าได้แม่นยำมากยิ่งขึ้น อาจต้องนำคุณลักษณะความเป็นวงกลมมาร่วมพิจารณาด้วย โดยอาศัยสมมติฐานที่ว่าบริเวณใบหน้าจะมีความเป็นวงกลมมากกว่าส่วนแขนและมือ หัวข้อนี้จะแสดงผลการวัดความเป็นวงกลมของแต่ละบริเวณ โดยเลือกรูปภาพบางเฟรมซึ่งผู้แปลภาษามือแสดงท่าทางที่ต่าง ๆ กันมาทดสอบจากวิดีโอที่ 1 และ 2 (ผู้แปลภาษามือสวมเสื้อแขนสั้น)

กำหนดให้ F (Face) คือ บริเวณใบหน้า

RH (Right Hand) คือ บริเวณมือขวา

LH (Left Hand) คือ บริเวณมือซ้าย

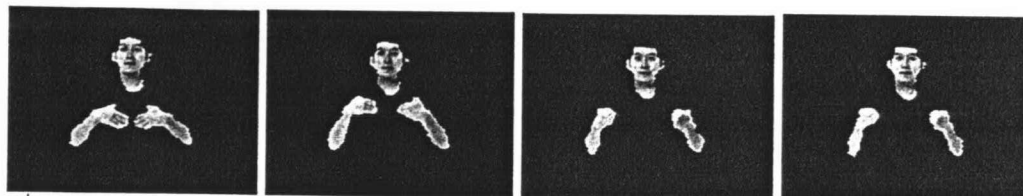
μ_R คือ ค่าระยะทางรัศมีเฉลี่ย

σ_R คือ ค่าเบี่ยงเบนมาตรฐานของระยะทางรัศมี

C คือ ค่าความเป็นวงกลม

ตารางที่ 4.9 ขนาดและค่าความเป็นวงกลมในแต่ละบริเวณของวิดีโอที่ 1

เฟรมที่	บล็อบ	ขนาด (จุดภาพ)	μ_R	σ_R	C
1	F	2100	18.961	8.024	2.363
	RH	1656	21.825	9.722	2.245
	LH	1548	20.238	9.549	2.12
10	F	2071	18.675	7.833	2.384
	RH	1588	22.879	10.129	2.259
	LH	1652	21.257	10.193	2.086
30	F	2014	18.333	7.654	2.395
	RH	1055	16.794	7.793	2.155
	LH	1145	16.145	7.499	2.153
50	F	2014	18.399	7.733	2.379
	RH	1039	17.551	8.308	2.113
	LH	1126	17.139	8.148	2.104

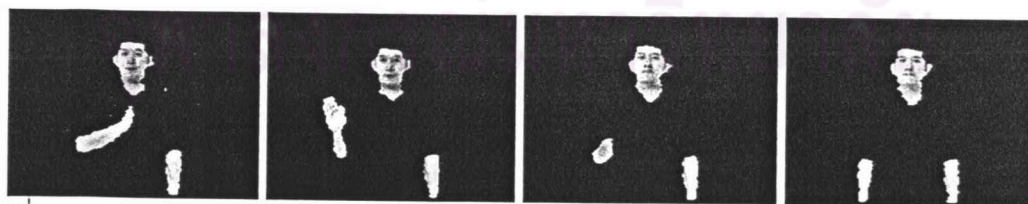


เฟรมที่ 1 10 30 50

รูปที่ 4.21 รูปภาพทดสอบจากวีดิทัศน์ 1 ที่นำมาวัดค่าความเป็นวงกลม (เฟรมที่ 1,10,30 และ 50)

ตารางที่ 4.10 ขนาดและค่าความเป็นวงกลมในแต่ละบริเวณของวีดิทัศน์ที่ 2

เฟรมที่	บล็อก	ขนาด (จุดภาพ)	μ_R	σ_R	C
10	F	2100	18.434	7.550	2.442
	RH	1476	21.747	11.663	1.865
	LH	929	14.607	7.249	2.015
30	F	2038	18.212	7.408	2.458
	RH	1331	19.316	10.272	1.881
	LH	848	13.738	6.872	1.999
50	F	2243	19.188	7.898	2.429
	RH	533	9.276	3.818	2.430
	LH	879	14.060	7.007	2.007
70	F	2234	19.191	7.994	2.401
	RH	773	13.856	7.095	1.953
	LH	854	14.197	7.2	1.972



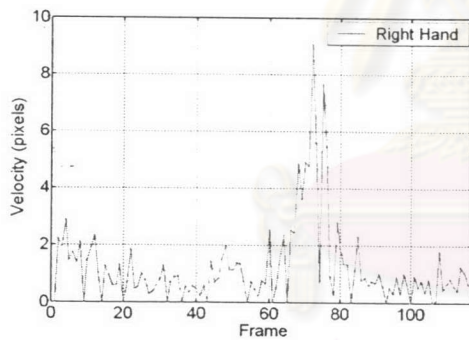
เฟรมที่ 10 30 50 70

รูปที่ 4.22 รูปภาพทดสอบจากวีดิทัศน์ 2 ที่นำมาวัดค่าความเป็นวงกลม (เฟรมที่ 10,30,50 และ 70)

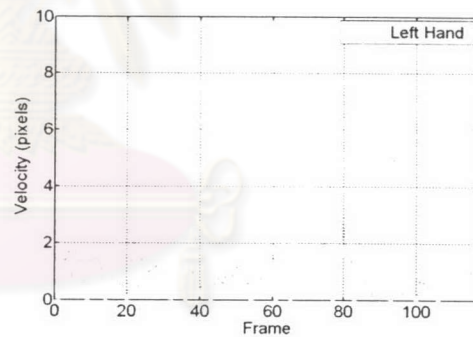
วิเคราะห์ผลการทดลอง จะเห็นได้ว่าค่าความเป็นวงกลม (C) ของรูปภาพที่นำมาทดสอบเกือบทั้งหมดบล็อบบริเวณใบหน้าจะมีค่ามากกว่ามือทั้งสอง ยกเว้นเฟรมที่ 50 ของวิดีโอคนที่ 2 บล็อบบริเวณใบหน้ามีค่าความเป็นวงกลมน้อยกว่ามือขวา เพราะผู้แปลเคลื่อนมือลงทำให้มีการบังกันระหว่างมือและแขน เมื่อสังเกตจากรูปที่ 4.22 บริเวณส่วนนี้ก็มีความคล้ายคลึงกับวงกลมมากเช่นกัน แนวทางแก้ไขกรณีของปัญหานี้ คือ นำคุณลักษณะความเป็นวงกลมและขนาดมาพิจารณาร่วมกัน เนื่องจากขนาดที่แสดงในตารางที่ 4.10 เฟรมที่ 50 บริเวณใบหน้ามีค่า 2243 จุดภาพ ส่วนบริเวณมือขวามีค่า 533 จุดภาพ หากใช้ขนาดในการกำหนดแต่ละบริเวณก็สามารถทำได้ถูกต้องเพราะมีความแตกต่างกันมากระหว่าง 2 บริเวณนี้

4.8 ความเร็วในการเคลื่อนที่ของมือ

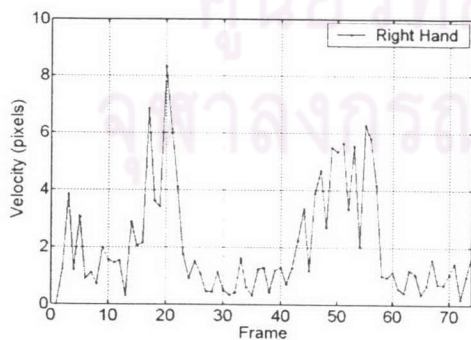
ในรูปที่ 4.23 แสดงความเร็วในการเคลื่อนที่ของมือขวา (คอลัมน์ที่ 1) และมือซ้าย (คอลัมน์ที่ 2) ของลำดับวิดีโอคนที่นำมาใช้ทดสอบ



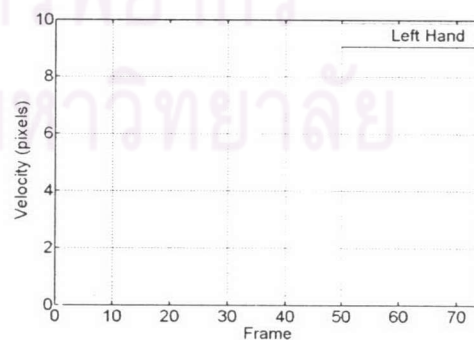
(ก-1) วิดีโอคนที่ 1 (มือขวา)



(ก-2) วิดีโอคนที่ 1 (มือซ้าย)

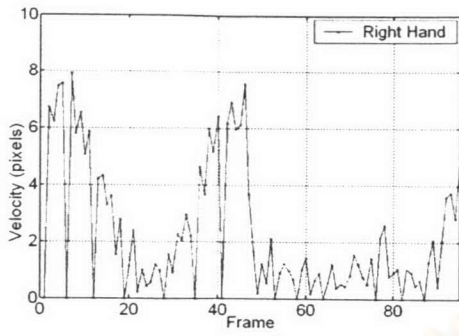


(ข-1) วิดีโอคนที่ 2 (มือขวา)

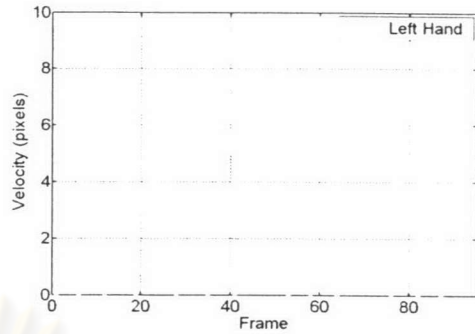


(ข-2) วิดีโอคนที่ 2 (มือซ้าย)

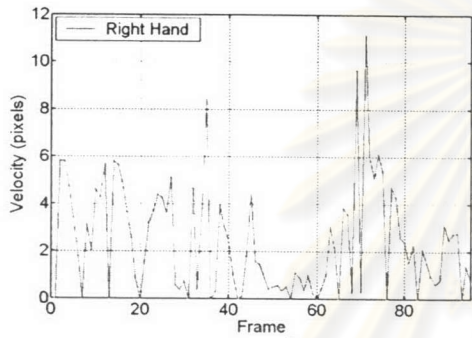
รูปที่ 4.23 ความเร็ว (จุดภาพต่อเฟรม) ในการเคลื่อนที่ของแต่ละบล็อบบ



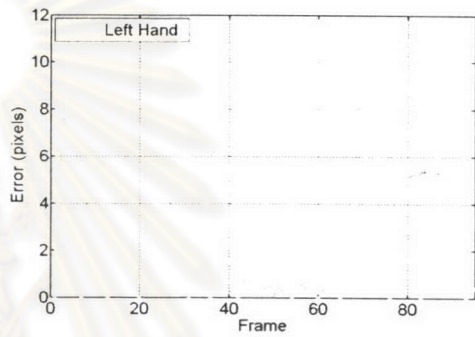
(ค-1) วิดีทัศน์ที่ 3 (มือขวา)



(ค-2) วิดีทัศน์ที่ 3 (มือซ้าย)



(ง-1) วิดีทัศน์ที่ 4 (มือขวา)

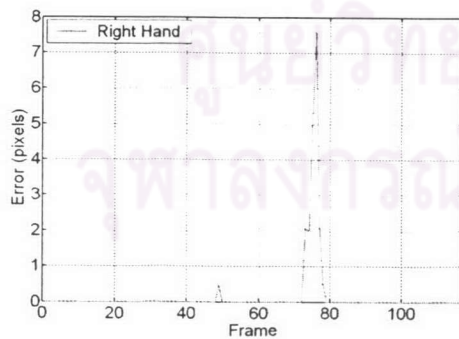


(ง-2) วิดีทัศน์ที่ 4 (มือซ้าย)

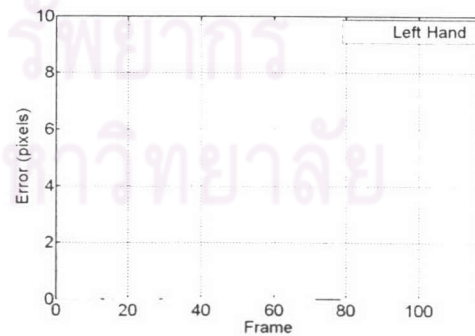
รูปที่ 4.23 (ต่อ) ความเร็ว (จุดภาพต่อเฟรม) ในการเคลื่อนที่ของแต่ละบลิบ

4.9 ความผิดพลาดของตำแหน่งมือทั้งสอง

4.9.1 การประมาณการเคลื่อนที่แบบที่ 2

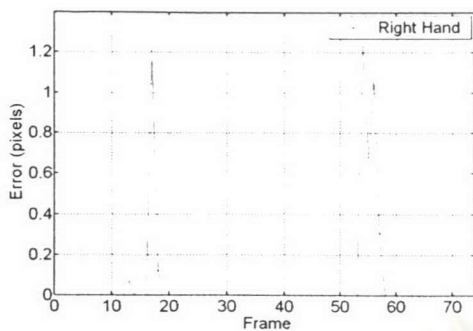


(ก-1) วิดีทัศน์ที่ 1 (มือขวา)

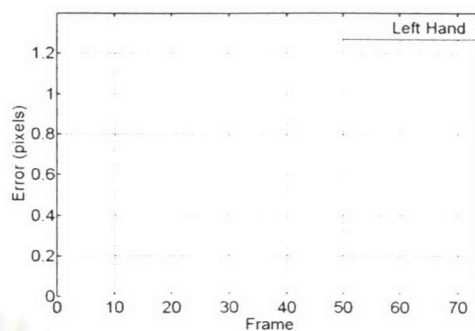


(ก-2) วิดีทัศน์ที่ 1 (มือซ้าย)

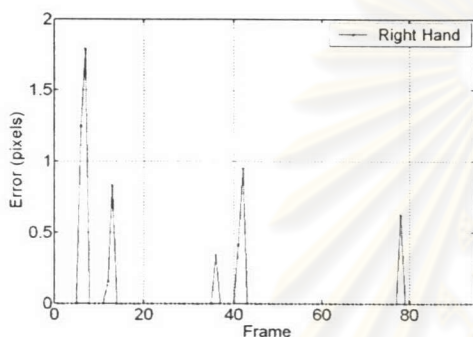
รูปที่ 4.24 ค่าความผิดพลาดของตำแหน่งมือทั้งสอง [15]



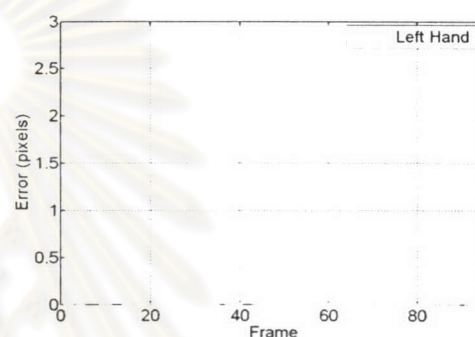
(ข-1) วิดีทัศน์ที่ 2 (มือขวา)



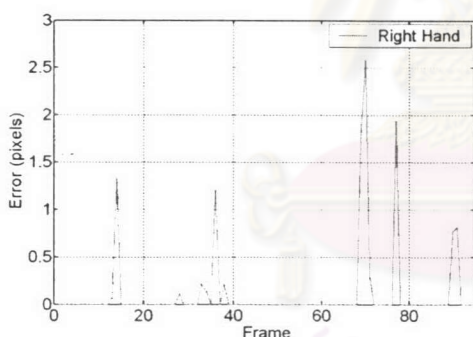
(ข-2) วิดีทัศน์ที่ 2 (มือซ้าย)



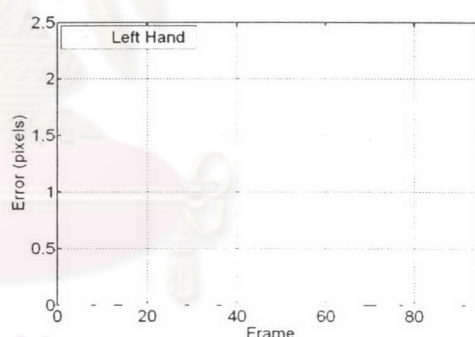
(ค-1) วิดีทัศน์ที่ 3 (มือขวา)



(ค-2) วิดีทัศน์ที่ 3 (มือซ้าย)



(ง-1) วิดีทัศน์ที่ 4 (มือขวา)



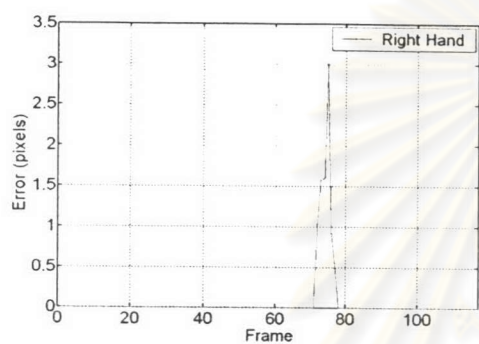
(ง-2) วิดีทัศน์ที่ 4 (มือซ้าย)

รูปที่ 4.24 (ต่อ) ค่าความผิดพลาดของตำแหน่งมือทั้งสอง [15]

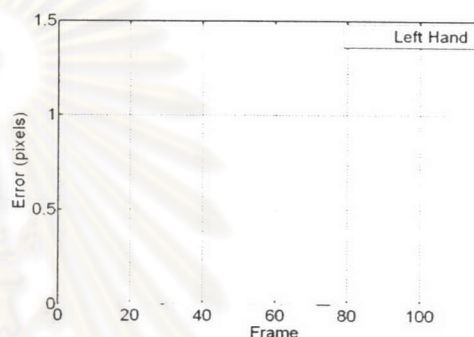
วิเคราะห์ผลการทดลอง ค่าความผิดพลาดที่เกิดขึ้นจากการประมาณการเคลื่อนที่นั้น จะเกิดขึ้นระหว่างเฟรมที่มีการเปลี่ยนแปลงความเร็วอย่างทันทีทันใด (พิจารณารูปที่ 4.23 ร่วมกับรูปที่ 4.24 พบว่าระหว่างเฟรมที่มีการเปลี่ยนแปลงความเร็วสูงจะเกิดค่าผิดพลาดทางตำแหน่งที่สูงเช่นกัน) ซึ่งค่าความผิดพลาดสูงสุดประมาณ 8 จุดภาพ โดยเกิดขึ้นที่มือขวาและซ้ายของวิดีโอที่ 1 (รูปที่ 4.24 (ก)) ส่วนในวิดีโอที่ 2 มือซ้าย (รูปที่ 4.24 (ข-2)) ไม่มีความผิดพลาดเนื่องจากมือซ้ายไม่มีการเคลื่อนที่ หัวข้อถัดไปจะแสดงผลของค่าผิดพลาดเมื่อใช้ขั้นตอนวิธีที่เสนอ

4.9.2 การประมาณการเคลื่อนที่แบบที่ 1 และปรับขอบเขตค้นหาให้สอดคล้องกับความเร็ว

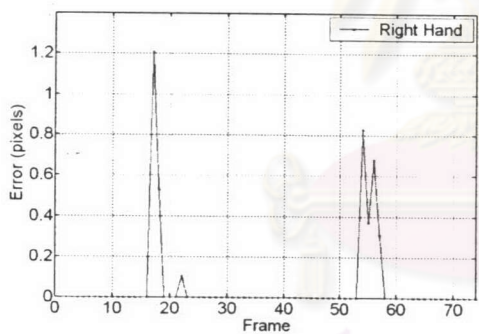
วิเคราะห์ผลการทดลอง เมื่อใช้ขั้นตอนวิธีที่นำเสนอ คือ การประมาณเคลื่อนที่ดั้งเดิมแบบที่ 1 (พิจารณาโดยละทิ้งความเร็ว) รวมถึงปรับขอบเขตการค้นหากรณีที่มีการเปลี่ยนแปลงความเร็วเกินขีดแบ่งที่กำหนดไว้ ทำให้ช่วยชดเชยค่าความผิดพลาดของตำแหน่งมือทั้งสองข้างลดลง หรือบางวีดิทัศน์แทบจะไม่มีค่าความผิดพลาดเลย เนื่องจากวีดิทัศน์ทดสอบไม่มีผลต่อความเร็ว ซึ่งผลของค่าความผิดพลาดของวิธีที่นำเสนอแสดงในรูปที่ 4.25



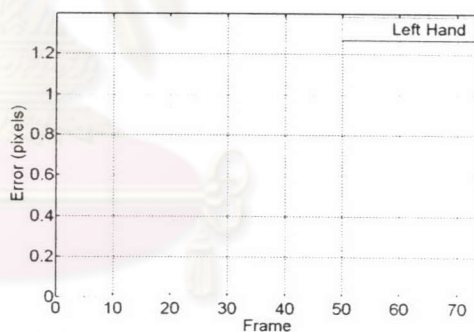
(ก-1) วีดิทัศน์ที่ 1 (มือขวา)



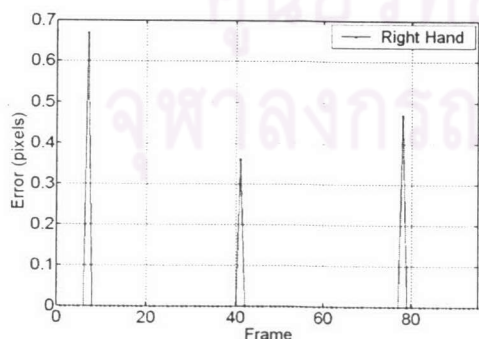
(ก-2) วีดิทัศน์ที่ 1 (มือซ้าย)



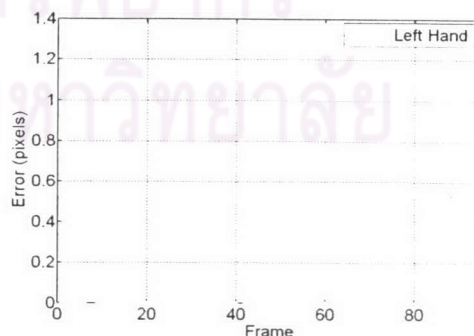
(ข-1) วีดิทัศน์ที่ 2 (มือขวา)



(ข-2) วีดิทัศน์ที่ 2 (มือซ้าย)

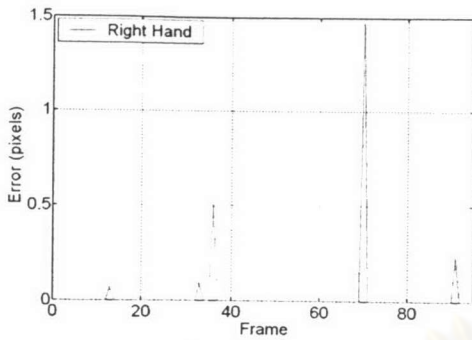


(ค-1) วีดิทัศน์ที่ 3 (มือขวา)

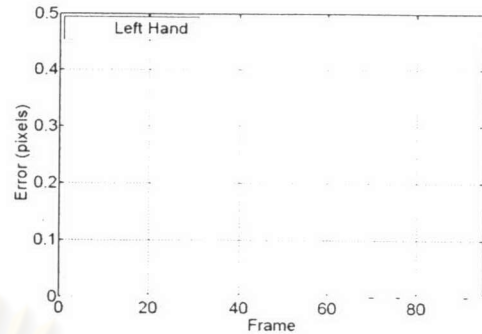


(ค-2) วีดิทัศน์ที่ 3 (มือซ้าย)

รูปที่ 4.25 ค่าความผิดพลาดของตำแหน่งมือทั้งสอง (นำเสนอ)



(ง-1) วิดีทัศน์ที่ 4 (มือขวา)

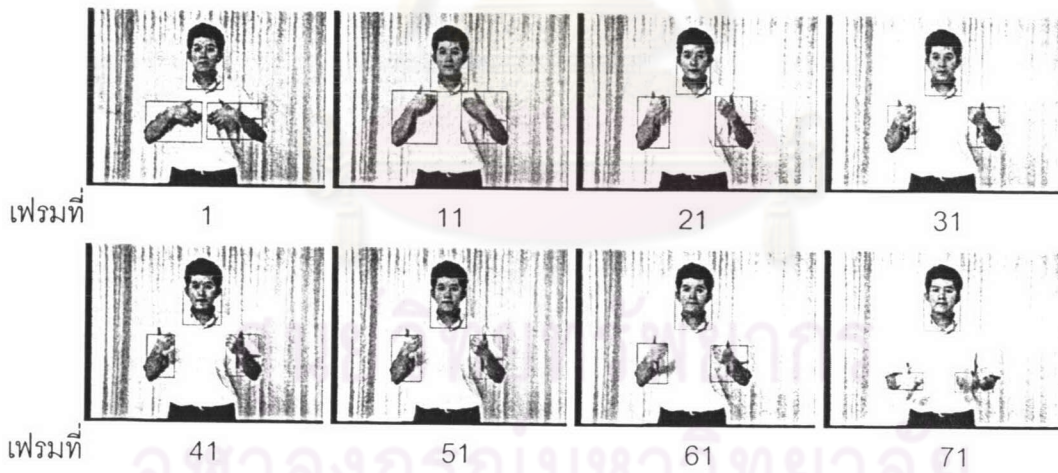


(ง-2) วิดีทัศน์ที่ 4 (มือซ้าย)

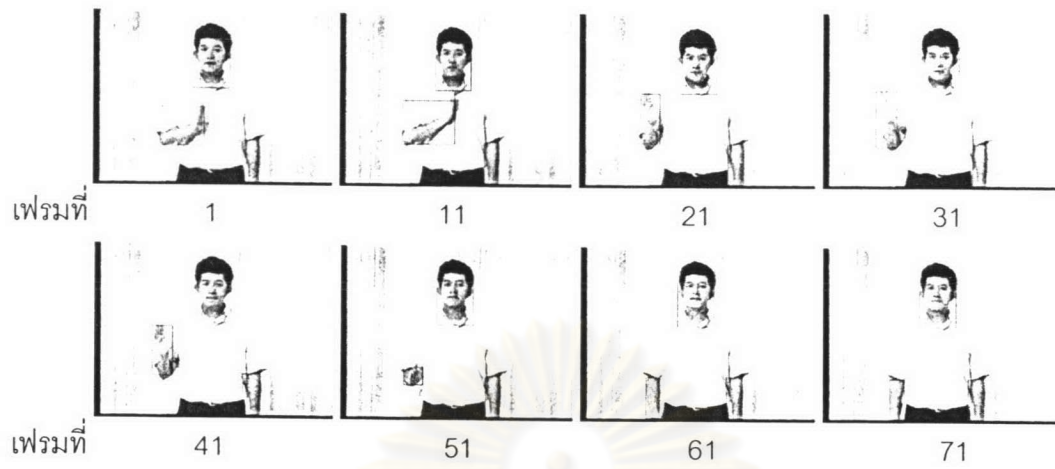
รูปที่ 4.25 (ต่อ) ค่าความผิดพลาดของตำแหน่งมือทั้งสอง (นำเสนอ)

4.10 ผลของการติดตามใบหน้าและมือ

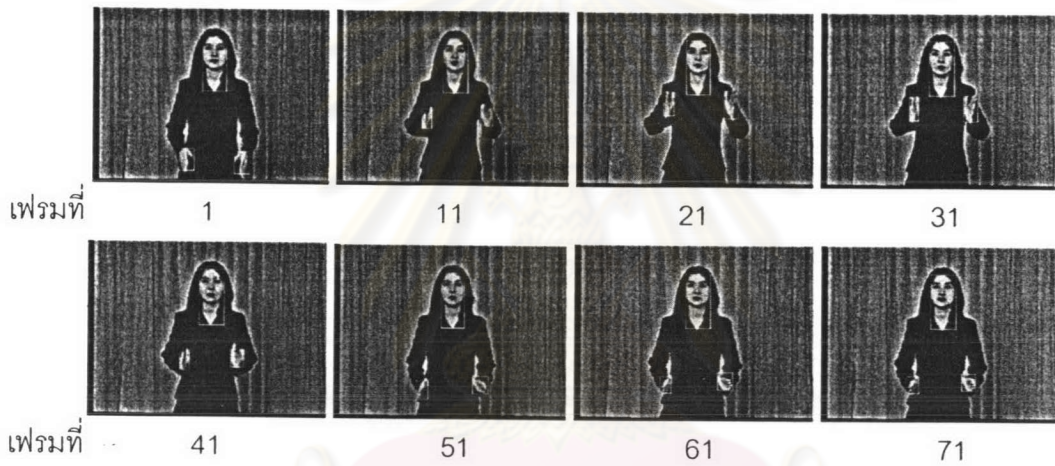
ในหัวข้อนี้จะแสดงผลที่ได้จากการติดตามใบหน้าและมือของวิดีโอที่นำมาทดสอบ เมื่อใช้ขั้นตอนวิธีที่เสนอ แสดงในรูปที่ 4.26 -4.29 (วิดีโอที่ 1-4 ตามลำดับ) กรอบที่ล้อมรอบ คือ ตำแหน่งต่ำสุดและสูงสุดของแต่ละบลิบ



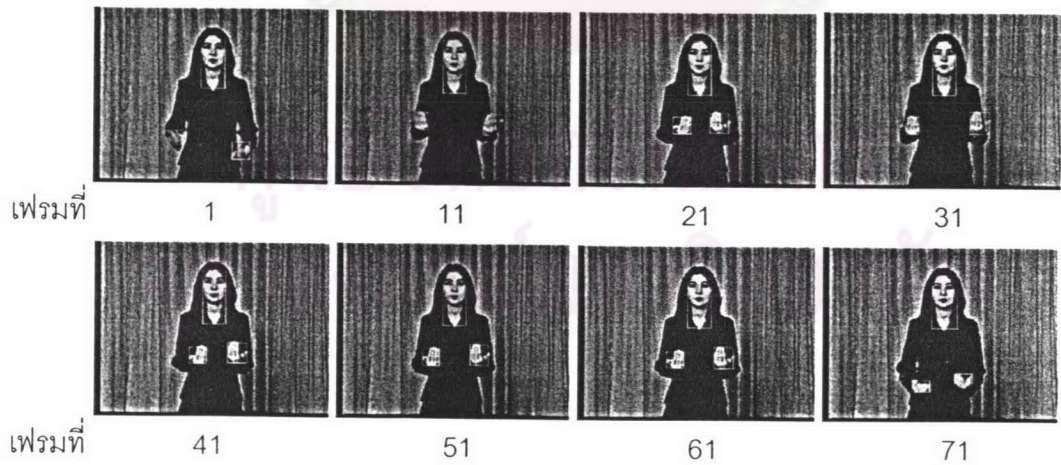
รูปที่ 4.26 ผลลัพธ์ที่ได้จากการติดตามในแต่ละเฟรมของวิดีโอที่ 1



รูปที่ 4.27 ผลลัพธ์ที่ได้จากการติดตามในแต่ละเฟรมของวีดิทัศน์ที่ 2



รูปที่ 4.28 ผลลัพธ์ที่ได้จากการติดตามในแต่ละเฟรมของวีดิทัศน์ที่ 3



รูปที่ 4.29 ผลลัพธ์ที่ได้จากการติดตามในแต่ละเฟรมของวีดิทัศน์ที่ 4

ตารางที่ 4.11 ค่าผิดพลาดที่เกิดขึ้นและเวลาประมวลผลที่ใช้ระหว่างขั้นตอนวิธีของ [15]

เปรียบเทียบกับวิธีปรับขอบเขตการค้นหาให้สอดคล้องกับความเร็ว

ลำดับ วิธีทัศน์	ขั้นตอน วิธี	ค่าผิดพลาด (จุดภาพ)		เวลา (วินาที)	จำนวนรอบ คำนวณซ้ำ
		RH	LH		
1	ดั้งเดิม	19.59	22.69	95.34	-
	นำเสนอ	8.6	3.61	120.31	4
2	ดั้งเดิม	4.69	0	71.95	-
	นำเสนอ	4.03	0	72.59	0
3	ดั้งเดิม	6.36	5.58	94.22	-
	นำเสนอ	1.5	1.7	96.59	8
4	ดั้งเดิม	11.46	7.9	94.16	-
	นำเสนอ	2.35	0.95	99.74	16

วิเคราะห์ผลการทดลองจากตารางที่ 4.11 ซึ่งจำลองโดยโปรแกรม MATLAB รุ่น 6.5

วิธีทัศน์ที่ 1 บริเวณมือขวาและมือซ้ายค่าผิดพลาดรวมลดลงประมาณ 11 และ 19 จุดภาพ ตามลำดับ ใช้เวลาในการประมวลผลเพิ่มขึ้นประมาณ 25 วินาที มีการคำนวณซ้ำ 4 รอบ เหตุผลที่วิธีทัศน์นี้ใช้เวลาในการประมวลผลมากกว่าวิธีทัศน์อื่นเพราะจำนวนเฟรมที่มากกว่า และผู้แปลภาษามือสวมเสื้อแขนสั้น ทำให้ขอบเขตในการค้นหามีบริเวณกว้างกว่า แต่ก็สามารถชดเชยความผิดพลาดได้มากเช่นกัน

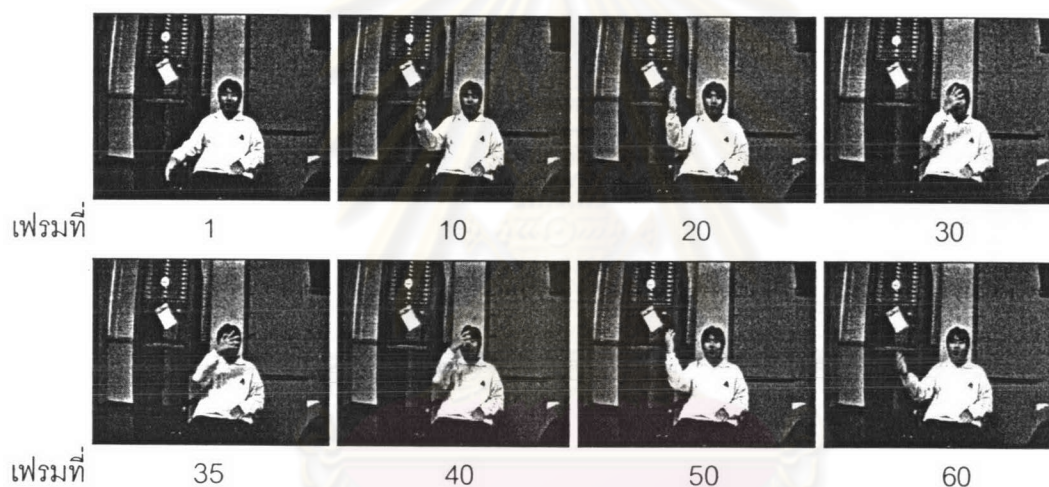
วิธีทัศน์ที่ 2 บริเวณมือขวาค่าผิดพลาดรวมลดลงประมาณ 0.6 จุดภาพ ส่วนมือซ้ายไม่เกิดค่าผิดพลาดเพราะไม่มีการเคลื่อนไหวมือของผู้แปลภาษามือ ใช้เวลาเพิ่มขึ้น 1 วินาที มีการชดเชยความผิดพลาดได้น้อยเพราะไม่มีการคำนวณซ้ำ เนื่องจากความแรงในการเคลื่อนที่ของมือไม่เกินขีดแบ่งที่ได้กำหนดไว้

วิธีทัศน์ที่ 3 บริเวณมือขวาและมือซ้ายค่าผิดพลาดรวมลดลงประมาณ 5 และ 4 จุดภาพ ตามลำดับ ใช้เวลาในการประมวลผลเพิ่มขึ้นประมาณ 2.4 วินาที มีการคำนวณซ้ำ 8 รอบ ซึ่งเป็นเวลาที่เพิ่มขึ้นเพียงเล็กน้อยหากเทียบกับวิธีทัศน์ที่ 1 เหตุที่เป็นเช่นนี้เพราะมีขอบเขตของบริเวณมือที่เล็กกว่านั่นเอง (ผู้แปลสวมเสื้อแขนยาว)

วิธีทัศนที่ 4 บริเวณมือขวาและมือซ้ายค่าผิดพลาดรวมลดลงประมาณ 9 และ 7 จุดภาพตามลำดับ ใช้เวลาในการประมวลผลเพิ่มขึ้นประมาณ 5.6 วินาที มีการคำนวณซ้ำ 16 รอบ วิธีทัศนนี้จะมีค่าความเร่งในการเคลื่อนที่ของมือเกินขีดแบ่งที่กำหนดไว้มากที่สุด

4.11 แนวทางแก้ปัญหาการผสานและแยกกันระหว่างใบหน้าและมือ

วิธีทัศนที่นำมาทดสอบแนวทางแก้ปัญหาการผสานและแยกจากกันของใบหน้าและมือทำการบันทึกด้วยกล้องโซนี่ รุ่น EVI-D 100P มีความยาวจำนวน 66 เฟรม ซึ่งจากหลังคอนข้างซับซ้อน โดยผู้แสดงจะเคลื่อนมือขวามาบังใบหน้าและเคลื่อนออก ส่วนมือซ้ายจะอยู่นิ่ง ดังแสดงในรูปที่ 4.30

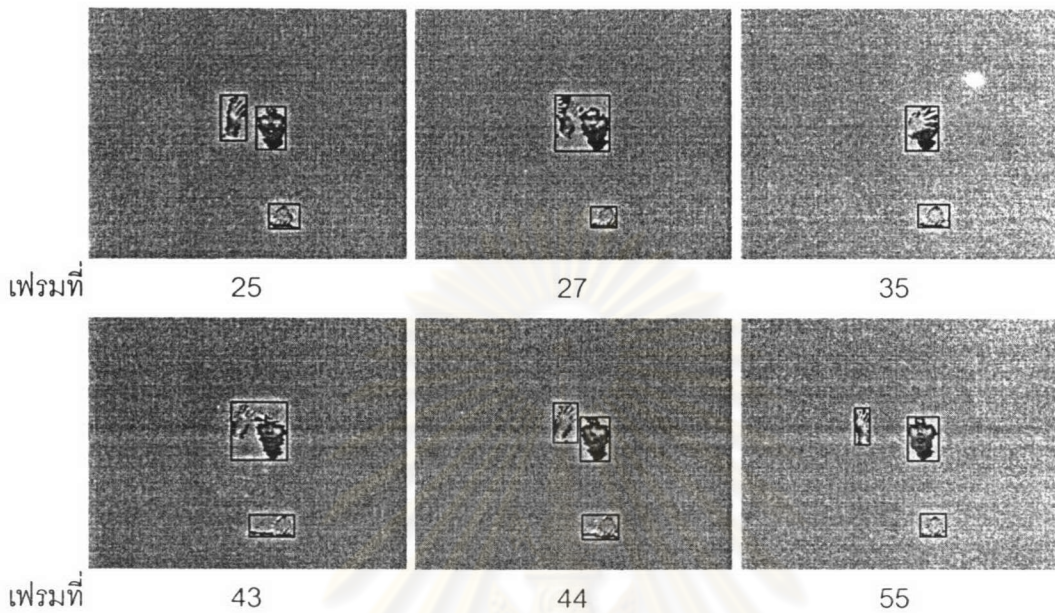


รูปที่ 4.30 วิธีทัศนที่มีการผสานและแยกกันระหว่างใบหน้าและมือของผู้แปลภาษามือ

การแยกส่วนสีผิวจะใช้แบบจำลองวงรี โดยค่าคงที่ต่าง ๆ ได้จากการเรียนรู้สีผิวของกล้องโซนี่รุ่น EVI-D 100P (หัวข้อที่ 4.2.3) และโครงสร้างส่วนย่อยที่นำมาใช้สำหรับตัวดำเนินการแบบปิดถูกแสดงในสมการที่ (4.1) ซึ่งทดสอบแล้วให้ผลดีที่สุด

$$SE = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.1)$$

เมื่อ SE คือ มาสก์ของโครงสร้างส่วนย่อย



รูปที่ 4.31 ผลของแนวทางแก้ปัญหาคารณสานและแยกกันระหว่างใบหน้าและมือ

วิเคราะห์ผลการทดลอง จากรูปที่ 4.31 วิดีทัศน์เริ่มมีการผสานกันของมือขวาและใบหน้าในเฟรมที่ 27 เนื่องจากขั้นตอนวิธีตรวจพบว่ามีตำแหน่งของ 2 บล็อกนี้ซ้ำกัน ซึ่งเมื่อมี 2 วัตถุผสานกัน จะถูกกำหนดเป็นวัตถุกลุ่มเดียวกัน จนกระทั่งถึงเฟรมที่ 44 จึงพบว่า 2 ส่วนนี้มีการแยกกัน ซึ่งทราบจากการวิเคราะห์การเชื่อมต่อกันพบว่ามี 2 บริเวณที่มีขนาดเพียงพอที่จะยอมรับว่าเป็นวัตถุ
 หมายเหตุ หากต้องการแยกขอบเขตของใบหน้าและมือ ขณะที่ผสานกันจำเป็นต้องใช้คุณลักษณะอื่นเพิ่มเติมนอกเหนือจากสีผิว เช่น ขอบ รูปร่าง เวกเตอร์การเคลื่อนที่ ฯลฯ

ศูนย์วิจัยทรัพยากร
 จุฬาลงกรณ์มหาวิทยาลัย