

## บทที่ 3

### กรรมวิธีที่นำเสนอ

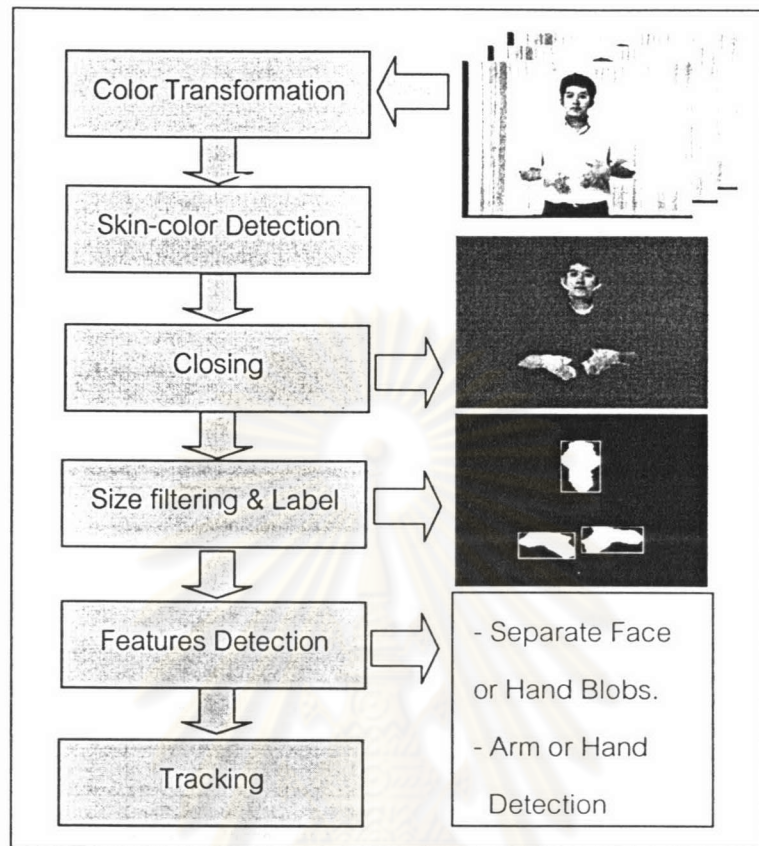
#### 3.1 โครงสร้างของระบบตรวจหาและติดตามใบหน้าและมือ

แนวทางในการพัฒนาระบบตรวจหาและติดตามใบหน้าและมือของผู้แปลภาษามือไทย เมื่อพิจารณาจากลำดับวิธีทัศน์ทดสอบแล้ว ในส่วนการตรวจหา คุณลักษณะสำคัญที่เลือกใช้ คือ คุณสมบัติของสีผิวซึ่งมีความแตกต่างเด่นชัดจากองค์ประกอบอื่น ๆ ภายในเฟรม และโดยปกติแล้วสัญญาณที่ได้จากลำดับวิธีทัศน์ หรือ กล้องวิธีทัศน์ จะแสดงเป็นค่า RGB ซึ่งไม่เหมาะสมสำหรับการแยกบริเวณส่วนสีผิวภายใต้เงื่อนไขของแสงที่ไม่แน่นอน ตัวอย่างของกรณีที่เกิดปัญหานี้ เช่น มีการเคลื่อนไหวของใบหน้าและมือขณะกำลังแสดงท่าทาง ทำให้ค่าแสงตกกระทบบนวัตถุเปลี่ยนแปลงอยู่ตลอดเวลา หรือ มีแหล่งกำเนิดแสงที่ไม่สามารถคาดเดาได้เข้ามารบกวนในลำดับวิธีทัศน์ เพื่อความถูกต้องของการตรวจหาและติดตามในทุกเฟรม จึงต้องทำการแปลงค่าแต่ละจุดภาพจากปริภูมิสี RGB ไปสู่ปริภูมิสีอื่น ๆ ที่มีการแยกกันระหว่างความสว่างและสี

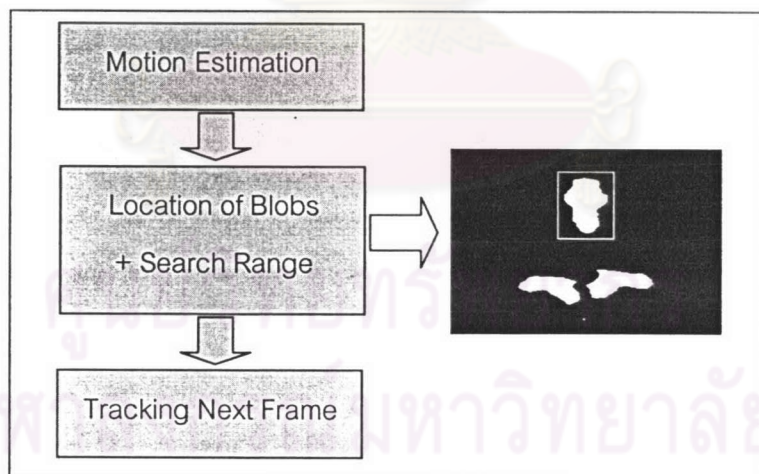
ในวิทยานิพนธ์นี้จะเลือกใช้การแปลงไปสู่ปริภูมิสี YCbCr เนื่องจากเป็นปริภูมิสีที่ใช้ในมาตรฐานการบีบอัดข้อมูลวิธีทัศน์สำหรับการสื่อสาร ซึ่งจะสามารถนำระบบนี้ไปผนวกร่วมใช้งานได้ง่าย และเหตุผลอีกประการที่เลือก คือ ในปริภูมิสี YCbCr (พิจารณาเฉพาะองค์ประกอบของสี Cb และ Cr) จะให้ผลการแยกส่วนสีผิวที่ดีกว่าปริภูมิสี RGB นอร์แมลไลซ์ (พิจารณาแบบ r และ g) และ HSI (พิจารณาเฉพาะองค์ประกอบของสี H และ S) เมื่อทำการทดลองโดยใช้แบบจำลองชนิดไม่มีตัวแปร (Non Parametric Model) ซึ่งได้แสดงในส่วนของผลการทดลอง หรือ จากงานวิจัยที่ผ่านมา เช่น การเปรียบเทียบการตรวจหาสีผิวในปริภูมิสี 5 [17] และ 9 ปริภูมิ [18] งานวิจัยของ M.J. Jones [20] พบว่าค่าในปริภูมิสี YCbCr ซึ่งค่า Cr องค์ประกอบสีแดงนั้นจะมีความเหมาะสมต่อการตรวจพบเฮโมโกลบิน (Haemoglobin) หรือบริเวณผิวหนังของมนุษย์มากกว่าปริภูมิสีอื่น เป็นต้น หลังจากขั้นตอนการคัดเลือกปริภูมิสีที่เหมาะสม ซึ่งพบว่า YCbCr มีข้อได้เปรียบกว่าปริภูมิสีอื่นแล้ว ในการทำงานจริงนั้น จำเป็นที่จะต้องคำนึงถึงความซับซ้อนในการทำงานของแบบจำลองที่นำมาใช้ในการแยกส่วนสีผิว เพื่อประโยชน์ในการพัฒนาระบบให้ทำงานทันตามเวลาจริง (real time) จึงเสนอวิธีการแยกสีผิวโดยใช้แบบจำลองชนิดวงรี (Elliptical Model) บนปริภูมิสีย่อย CbCr

โครงสร้างของระบบที่นำเสนอแสดงขั้นตอนการทำงานในรูปที่ 3.1 ซึ่งถูกแยกออกเป็น 2 ส่วน คือ ส่วนเริ่มต้น และส่วนการติดตาม ในส่วนเริ่มต้น รูปที่ 3.1 (ก) วิดีทัศน์รับเข้า ซึ่งแสดงเป็นค่า RGB จะถูกแปลงไปสู่ปริภูมิสี YCbCr และนำจุดภาพทุก ๆ จุดภายในเฟรมมาแยกส่วนสีผิวด้วยแบบจำลองวงรี จากนั้นมาสกที่ได้จะผ่านตัวดำเนินการแบบปิดเพื่อลดค่าสัญญาณรบกวนประเภทหลุมขนาดเล็ก (small holes noise) ต่อมาจะทำการวิเคราะห์บริเวณเชื่อมต่อกันเพื่อจำแนกตำแหน่งของแต่ละกลุ่ม (บล็อบ) และกำจัดสัญญาณรบกวนซึ่งเกิดขึ้น ในกรณีที่ฉากหลังมีความซับซ้อนมาก กรณีนี้อาจมีกลุ่มของจุดภาพบางกลุ่มที่ไม่ใช่สีผิวหลุดรอดจากการตรวจสอบโดยแบบจำลองวงรี ทำให้มีความผิดพลาดเกิดขึ้น คุณลักษณะที่นำมาใช้ในการแก้ไขข้อผิดพลาดนี้ คือ การพิจารณาขนาดของแต่ละบล็อบ ซึ่งขั้นตอนวิธีที่ใช้จะตัดสินใจเลือกบล็อบที่ใหญ่ที่สุด 3 บล็อบเป็นใบหน้าและมือ พร้อมทั้งเก็บค่าจุดศูนย์กลางไว้ในหน่วยความจำ กระบวนการถัดไปคือการตรวจหาลักษณะเด่นของแต่ละบล็อบ ขั้นตอนนี้ระบบจะแยกการตัดสินใจเป็น 2 กรณี คือ กรณีที่ผู้แปลภาษามือสวมเสื้อแขนยาว จะแยกความแตกต่างของบล็อบบริเวณใบหน้าและมือ โดยตัดสินใจจากขนาด ซึ่งขนาดที่ใหญ่กว่าจะถูกกำหนดเป็นใบหน้า แต่หากผู้แปลภาษามือสวมเสื้อแขนสั้นการตัดสินใจจากขนาดอย่างเดียวอาจทำให้เกิดความไม่แน่นอนจึงต้องพิจารณาเงื่อนไขเพิ่มเติม คือ การหาค่าประกอบสำคัญบนใบหน้า โดยบล็อบบริเวณใบหน้าก็จะประกอบไปด้วย ตา คิ้ว จมูก และปาก และในกรณีที่ผู้แปลสวมเสื้อแขนสั้นนี้ บล็อบบริเวณแขนจำเป็นที่จะต้องมีการแยกส่วนเชื่อมต่อกันระหว่างข้อมือกับแขน เนื่องจากเป็นข้อต่อที่มีการเคลื่อนที่คนละทิศทาง อันจะเป็นประโยชน์ต่อส่วนการติดตามที่จะประมาณค่าตำแหน่งและสามารถทำงานได้อย่างมีประสิทธิภาพ จากส่วนเริ่มต้นจะทำให้เราทราบตำแหน่งของใบหน้าและมือในเบื้องต้น โดยการประมวลผลด้วยภาพเพียงภาพเดียวหรือเชิงพื้นที่ ก่อนที่จะนำข้อมูลนี้ไปใช้ในส่วนการติดตามต่อไป

ส่วนการติดตาม รูปที่ 3.1 (ข) จะอาศัยสมมติฐานที่ว่าลำดับวิดีโอที่คนในเฟรมที่อยู่ติดกัน จะมีการเปลี่ยนตำแหน่งของวัตถุเคลื่อนที่ในระยะทางไม่มาก ซึ่งเราจะใช้สมมติฐานนี้มากำหนดอาณาบริเวณในการหาวัตถุที่สนใจ (ส่วนสีผิว) ของเฟรมถัดไป เพื่อประโยชน์ในการลดความซับซ้อนในการประมวลผลทางโดเมนของเวลา และจากการศึกษาวิดีโอที่คนในเฟรมพบว่า การเปลี่ยนตำแหน่งของใบหน้าและมือมีความแตกต่างกัน โดยใบหน้าที่มีการเคลื่อนไหวน้อยกว่ามือ ในการประมาณตำแหน่งของใบหน้า จึงสามารถทำได้โดยง่ายเพียงขยายขอบเขตจากตำแหน่งอ้างอิงเดิมออกไปเพียงเล็กน้อย และกำหนดเอาบริเวณนั้นในการค้นหาสำหรับเฟรมถัดไป แต่การประมาณตำแหน่งของมือ จำเป็นต้องใช้สมการประมาณการเคลื่อนที่ที่ซับซ้อนกว่าและข้อมูลจุดศูนย์กลางของวัตถุในเฟรมก่อนหน้า จะถูกนำมาใช้หาอนุพันธ์อันดับหนึ่ง (ความเร็ว) และ อนุพันธ์อันดับสอง (ความเร่ง) เพื่อประมาณการเคลื่อนที่ต่อไป



(ก)



(ข)

รูปที่ 3.1 โครงสร้างของระบบตรวจหาและติดตามใบหน้าและมือ

(ก) ส่วนเริ่มต้น (ข) ส่วนการติดตาม



ในส่วนสมการประมาณการเคลื่อนที่ของมือ วิทยานิพนธ์ฉบับนี้ได้นำเสนอวิธีที่ให้ค่าความผิดพลาดน้อยกว่า รวมถึงการปรับตัวได้ของระบบที่ใช้เมื่อมีความผิดพลาดเกิดขึ้น ซึ่งสามารถรายละเอียดในหัวข้อถัดไป ในส่วนการติดตามนี้จะสามารถนำข้อมูลเชิงเวลามาทำการจำแนกใบหน้าหรือมือได้ โดยจะตรวจสอบจากการเคลื่อนที่ของจุดศูนย์กลางถ่วงของแต่ละวัตถุในทุก ๆ เฟรม เมื่อระบบประมวลผลไประยะเวลาหนึ่ง หากจุดศูนย์กลางถ่วงวัตถุใดมีการเปลี่ยนตำแหน่งไปมาก ระบบจะตัดสินใจเป็นมือ และจุดศูนย์กลางถ่วงวัตถุที่มีการเปลี่ยนตำแหน่งน้อยจะถูกตัดสินให้เป็นใบหน้า หากเป็นกรณีที่ภาษามือค่านั้นมีการเคลื่อนไหวของมือเพียงข้างเดียวและอีกข้างอยู่นิ่ง จะได้ค่าการเปลี่ยนตำแหน่งของจุดศูนย์กลางถ่วงที่น้อยใกล้เคียง ๆ กัน กรณีนี้จะจำแนกใบหน้าหรือมือ โดยพิจารณาจากความสูงต่ำของตำแหน่งจุดศูนย์กลางถ่วง

สำหรับเนื้อหาในบทที่ 3 ตั้งแต่หัวข้อถัดไป คือ หัวข้อที่ 3.2 – 3.5 จะอธิบายขั้นตอนในส่วนเริ่มต้น และหัวข้อที่ 3.6 เป็นส่วนของการติดตาม

### 3.2 การแยกส่วนสีผิวมนุษย์

คุณลักษณะเด่นที่นำมาใช้เป็นหลักในการตรวจหาใบหน้าและมือมนุษย์ คือ สีผิวของมนุษย์ โดยการแยกส่วนสีผิวนั้นระบบจำเป็นที่จะต้องมีการเรียนรู้ค่าของจุดภาพส่วนสีผิว ซึ่งทำการเก็บไว้ก่อนล่วงหน้าพร้อมทั้งสร้างแบบจำลองที่มีความเหมาะสมขึ้นมา ความแตกต่างทางเชื้อชาติ ระหว่างคนแอฟริกัน (ผิวดำ) คนยุโรป (ผิวขาว) หรือ คนเอเชีย (ผิวเหลือง) ดังรูปที่ 3.2 ความแตกต่างเหล่านี้มีผลโดยตรงกับแบบจำลองที่นำมาใช้ ในงานวิจัยนี้วิธีทัศน์ที่นำมาประมวลผลมี 2 ส่วนด้วยกัน ซึ่งทั้งสองส่วนมีผู้แสดงแบบเป็นคนไทย ดังนั้นแบบจำลองที่สร้างขึ้นจึงมีความเหมาะสมกับสีผิวของคนไทย รวมถึงคนเอเชียเป็นอย่างยิ่ง ส่วนแรก คือ วิธีทัศน์ภาษามือของกระทรวงศึกษาธิการที่แสดงโดยผู้เชี่ยวชาญและบันทึกไว้เพื่อใช้ในการศึกษาภาษามือไทย [21] โดยแต่ละวิธีทัศน์มีสภาพแวดล้อมของแสงที่ต่างกัน ในส่วนนี้จึงเก็บค่าสีผิวแยกเฉพาะแต่ละวิธีทัศน์ อีกส่วนหนึ่งเป็นการบันทึกโดยกล้องไซน์รุ่น EVI-D100 P ซึ่งมีความสามารถในการสาย (pan) ก้มเงย (tilt) และ ซูม (zoom) และจะนำไปใช้พัฒนาระบบแปลความหมายภาษามือต่อไปในอนาคต ทำการถ่ายภาพจากสภาพแวดล้อมภายในห้องวิจัย เก็บค่าสีผิวมนุษย์จำนวน 20 ตัวอย่างสำหรับสร้างแบบจำลอง ในส่วนนี้ได้มีการทดสอบแยกสีผิวของบุคคลที่นอกเหนือจากกลุ่มตัวอย่างที่เก็บค่าไว้ เพื่อชี้ให้เห็นว่าแบบจำลองยังคงสามารถทำงานได้อย่างดี ซึ่งผลการทดลองแสดงในบทที่ 4

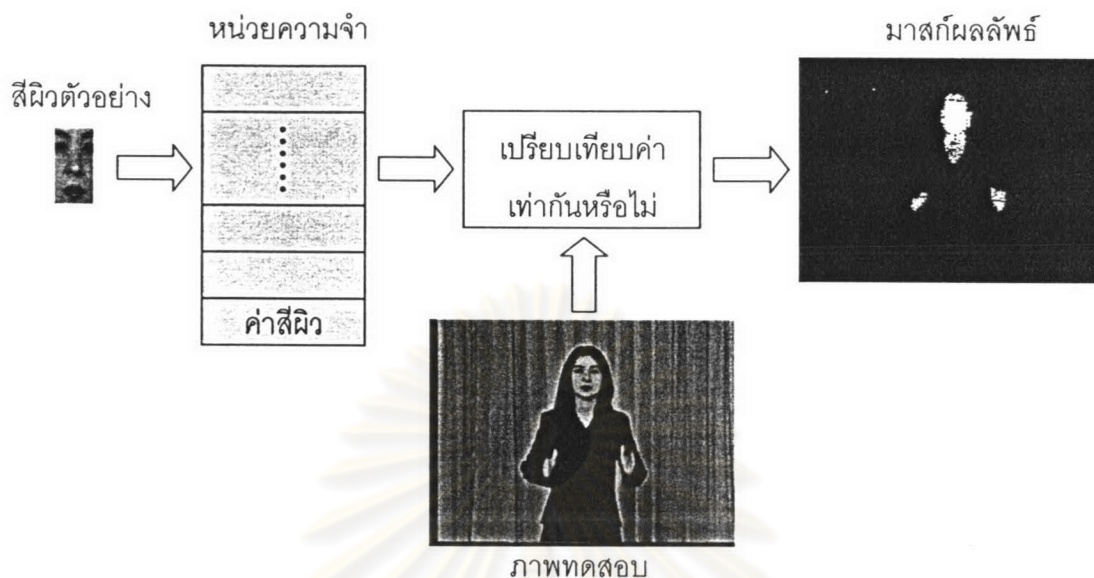


รูปที่ 3.2 สีผิวของมนุษย์ที่มีเชื้อชาติต่าง ๆ กัน

ปัจจุบันมีปริภูมิสีมากมายที่ถูกนำเสนอขึ้น แต่ละปริภูมิสีก็มีลักษณะเด่นเฉพาะด้าน ดังนั้น การเลือกปริภูมิสีที่เหมาะสมสำหรับการแยกส่วนสีผิวถือเป็นประเด็นหนึ่งที่สำคัญ ในการพัฒนาระบบ ตรวจสอบและติดตามใบหน้าและมือที่มีประสิทธิภาพ งานวิจัยเกี่ยวกับการเปรียบเทียบประสิทธิภาพ ของการแยกสีผิวมนุษย์เมื่อใช้ขั้นตอนวิธีหรือปริภูมิสีต่าง ๆ กัน ได้ถูกนำเสนอไว้บ้างแต่ยังมีงานวิจัยใน ส่วนนี้น้อยมาก [17,18,22,23] ดังนั้นวิทยานิพนธ์ฉบับนี้จึงเสนอวิธีการหนึ่ง que แสดงประสิทธิภาพของ การแยกสีผิวในแต่ละปริภูมิสี คือ วิธีของแบบจำลองชนิดไม่มีตัวแปร

### 3.2.1 การแยกส่วนสีผิวแต่ละปริภูมิสีโดยใช้แบบจำลองชนิดไม่มีตัวแปร

เนื่องจากความไม่เหมาะสมหากจะทำการแยกสีผิวในปริภูมิสี RGB ซึ่งเป็นสัญญาณที่ได้จาก กล้องวิดีโอ เพราะมีสหสัมพันธ์กันระหว่างองค์ประกอบของแสงและสี ดังนั้นในหัวข้อนี้จะอธิบายวิธี คัดเลือกปริภูมิสีที่เหมาะสมในการแยกส่วนสีผิวมนุษย์ โดยใช้แบบจำลองชนิดไม่มีตัวแปรในการวัด ประสิทธิภาพ ซึ่งปริภูมิสีที่ถูกเลือกมาทำการเปรียบเทียบมีด้วยกัน 3 ปริภูมิสี คือ HSI (พิจารณา H และ S) นอร์แมลไลซ์ RGB (พิจารณา r และ g) และ YCbCr (พิจารณา Cb และ Cr) ใช้เฉพาะ องค์ประกอบของสีที่จะถึงความสว่าง และเพื่อให้เกิดความยุติธรรมแก่ทุกปริภูมิสีที่นำมาเปรียบเทียบค่า ในปริภูมิสีย่อย rg และ HS จะถูกนอร์แมลไลซ์ให้มีช่วงเท่ากับ 224 เพื่อให้เท่ากับช่วงกว้างในปริภูมิสี ย่อย CbCr (ค่าระหว่าง 16 ถึง 240) ค่าสีผิวตัวอย่างจะถูกนำมาจัดเก็บไว้ในหน่วยความจำ จากนั้นแต่ ละจุดภาพของรูปภาพทดสอบจะถูกนำไปเปรียบเทียบกับทุก ๆ ค่าในหน่วยความจำ หากมีค่าตรงกันพอดีจุดภาพนั้นก็จะถูกกำหนดเป็นสีผิว ซึ่งวิธีการของแบบจำลองชนิดไม่มีตัวแปรนี้แสดงดังรูปที่ 3.3



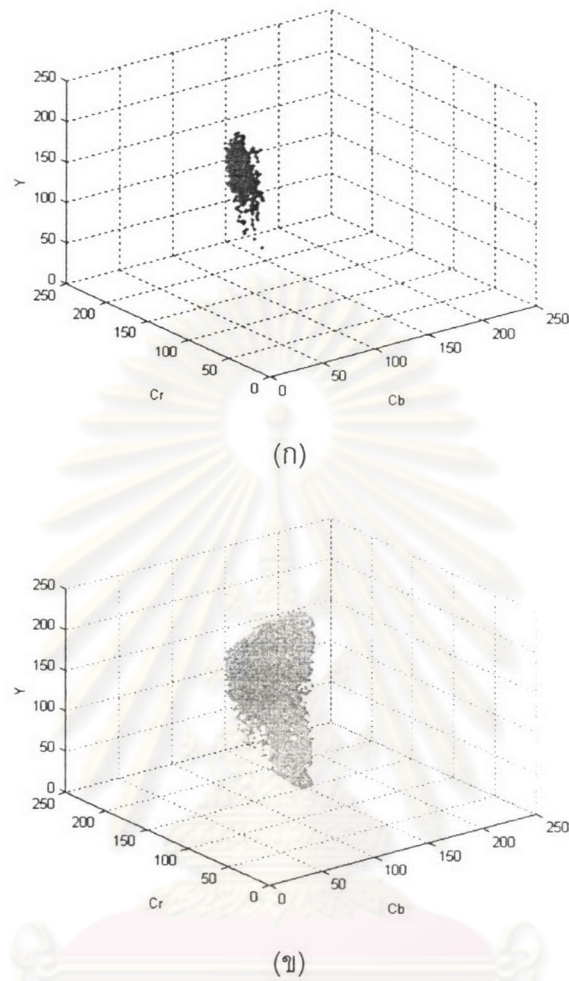
รูปที่ 3.3 วิธีการของแบบจำลองชนิดไม่มีตัวแปร

### 3.2.2 การกระจายตัวของจุดภาพบริเวณสีผิวในปริภูมิสีย่อย CbCr

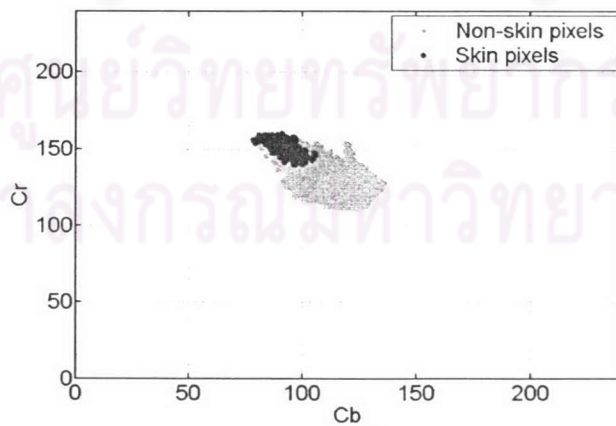
ในงานวิจัยที่ผ่านมาจะทำการเก็บข้อมูลค่าของสีผิวตัวอย่าง แล้วแยกด้วยกรรมวิธีต่าง ๆ กัน เช่น กำหนดค่าสูงสุดต่ำสุดเป็นค่าในการตัดสินใจบริเวณส่วนสีผิว [13] ใช้เส้นตรงหลาย ๆ เส้นในการแยกบริเวณส่วนสีผิว [9] ใช้พาราโบลาคกรอบคลุมข้อมูลส่วนสีผิว [6] ใช้สมมติฐานของเกาส์ [8] หรือใช้วงรีในการแบ่งขีดแบ่ง (threshold) บริเวณส่วนสีผิว [12]

ในปริภูมิสี YCbCr ค่าองค์ประกอบทั้งสามของจุดภาพสีผิวแสดงในรูปที่ 3.4 (ก) และค่าบริเวณอื่น ๆ ของรูปภาพแสดงในรูป 3.4 (ข) หากพิจารณาในปริภูมิสีย่อย CbCr โดยละทิ้งองค์ประกอบความสว่าง (Y) เมื่อเรานำค่า CbCr ของจุดภาพแต่ละจุดมาเขียนลงบนระนาบ CbCr จะได้ลักษณะการกระจายตัวดังรูปที่ 3.5 โดยแยกส่วนบริเวณสีผิว กับบริเวณอื่น ๆ ของรูปภาพ จะสังเกตเห็นได้ว่าค่า CbCr ของจุดภาพที่เป็นสีผิวจะเกาะกลุ่มกันอยู่บริเวณหนึ่ง เป้าหมาย คือ ต้องทำการแยกบริเวณส่วนนี้ออกมา ซึ่งวิธีการที่น่าเสนอจะใช้แบบจำลองที่เป็นวงรีเช่นเดียวกัน แต่ระบบตรวจหาบริเวณส่วนสีผิวที่สร้างขึ้นจะสามารถนำไปใช้งานร่วมกับระบบที่มีการปรับตัวได้ เนื่องจากได้มีการนำเสนอกรรมวิธีที่ให้นาขอบเขตของวงรีเพื่อกระทบกับข้อมูลส่วนที่เป็นสีผิวเพิ่มเติมอีกด้วย ซึ่งจะแสดงรายละเอียดในหัวข้อถัดไป



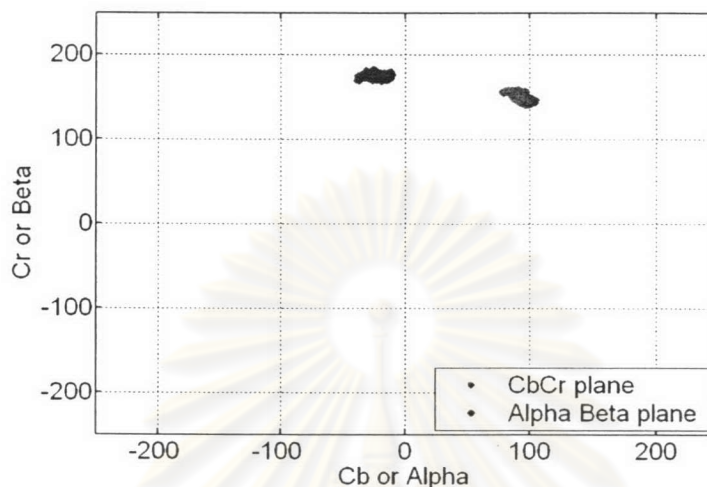


รูปที่ 3.4 การกระจายของ (ก) ส่วนสีผิว (ข) ส่วนที่ไม่ใช่สีผิว ในปริภูมิสี YCbCr



รูปที่ 3.5 การกระจายของสีผิวและส่วนที่ไม่ใช่สีผิว ในปริภูมิสีย่อย CbCr

### 3.2.3 กรรมวิธีการหาขอบเขตของวงรีด้วยวิธีหาความเยื้องศูนย์กลาง



รูปที่ 3.6 ค่าการกระจายตัวของจุดภาพสีผิว บนระนาบ CbCr และ  $\alpha\beta$

เมื่อพิจารณาการกระจายตัวของสีผิวที่เลือกมาเป็นค่าตัวอย่างในระนาบ CbCr ตามรูปที่ 3.5 แล้ว พบว่ามีการเกาะกลุ่มกันอยู่บริเวณหนึ่ง และบริเวณนี้มีลักษณะคล้ายรูปทรงทางเรขาคณิต คือ รูปวงรีมากที่สุด จึงเลือกที่จะแยกส่วนสีผิวกลุ่มนี้ด้วยแบบจำลองวงรี แต่ปัญหาที่เกิดขึ้นคือแกน Cb และ Cr ไม่ใช่แกนที่ขนานกับแกนเอกและแกนโทของวงรี จึงต้องทำการแปลงค่าไปสู่แกน  $\alpha$  และ  $\beta$  โดยการหาความเยื้องศูนย์กลางของกลุ่มจุดภาพสีผิวนี้

กรรมวิธีในการหาวงรีล้อมรอบบริเวณสีผิวมีขั้นตอนดังนี้  
การแปลงคู่อันดับ  $Cb(i, j)$  และ  $Cr(i, j)$  ของจุดภาพที่เป็นสีผิวไปสู่แกน  $\alpha$  และ  $\beta$  ทำได้โดยสมการที่ (3.1)

$$\begin{bmatrix} \alpha(i, j) \\ \beta(i, j) \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} Cb(i, j) \\ Cr(i, j) \end{bmatrix} \quad (3.1)$$

เมื่อ  $Cb(i, j)$  คือ ค่า Cb ของจุดภาพใด ๆ

$Cr(i, j)$  คือ ค่า Cr ของจุดภาพใด ๆ



การหาทิศทางของเส้นตรงที่ยาวที่สุด คือ  $\theta$  สามารถหาได้จากสมการที่ (3.2)

$$\theta = \frac{1}{2} \tan^{-1} \left( \frac{2\mu_{11}}{\mu_{20} + \mu_{02}} \right) \quad (3.2)$$

ซึ่งสามารถหา  $\mu_{pq}$  ได้จากสมการที่ (3.3)

$$\mu_{pq} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (i - \bar{Cb})^p (j - \bar{Cr})^q f(i, j) \quad (3.3)$$

โดยที่

$$\bar{Cb} = \frac{m_{10}}{m_{00}} \quad (3.4)$$

$$\bar{Cr} = \frac{m_{01}}{m_{00}} \quad (3.5)$$

และสามารถคำนวณ  $m_{pq}$  ได้จากสมการที่ (3.6)

$$m_{pq} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} i^p j^q f(i, j) \quad (3.6)$$

โดยที่  $\bar{Cb}$  คือ ค่าเฉลี่ยของ  $Cb$

$\bar{Cr}$  คือ ค่าเฉลี่ยของ  $Cr$

การหาความยาวแกนเอกและแกนโทของวงรี คำนวณจากความกว้างของค่าเฉลี่ยบวกและลบ ด้วยค่าคงที่คูณกับค่าเบี่ยงเบนมาตรฐาน ซึ่งจะมีความทนทานต่อสัญญาณรบกวนมากกว่าการนำค่าต่ำสุดและสูงสุดมาใช้ ในกรณีที่มีค่าจุดภาพใดจุดภาพหนึ่งมีค่าแตกต่างไปจากกลุ่มข้อมูลส่วนใหญ่

$$V_{\pm\alpha} = \bar{\alpha} \pm \eta_1 * std(\alpha) \quad (3.7)$$

$$V_{\pm\beta} = \bar{\beta} \pm \eta_1 * std(\beta) \quad (3.8)$$

โดยที่  $\bar{\alpha}$  คือ ค่าเฉลี่ยของ  $\alpha$   
 $\bar{\beta}$  คือ ค่าเฉลี่ยของ  $\beta$   
 $\eta_1$  คือ ค่าถ่วงน้ำหนัก  
 $std(\alpha)$  คือ ค่าเบี่ยงเบนมาตรฐานของ  $\alpha$   
 $std(\beta)$  คือ ค่าเบี่ยงเบนมาตรฐานของ  $\beta$

ในระนาบ  $\alpha\beta$  จุดยอดบนแกนเอกของวงรี คือ  $(V_{\pm\alpha}, \bar{\beta})$  และ จุดยอดบนแกนโทของวงรี คือ  $(\bar{\alpha}, V_{\pm\beta})$  สำหรับการหาขีดแบ่ง  $\eta_1$  (Threshold) ที่ดีที่สุด ได้อธิบายและแสดงขั้นตอนไว้ในบทที่ 4 หลังจากนั้นจะนำค่าจุดยอดที่ได้แปลงกลับไปสู่ระนาบ CbCr เพื่อกำหนดขอบเขตวงรีที่ได้ การแปลงกลับทำโดยใช้สมการที่ (3.9) และ (3.10)

$$\begin{bmatrix} V_{\pm Cb} \\ Cr \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} V_{\pm\alpha} \\ \bar{\beta} \end{bmatrix} \quad (3.9)$$

$$\begin{bmatrix} \bar{Cb} \\ V_{\pm Cr} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \bar{\alpha} \\ V_{\pm\beta} \end{bmatrix} \quad (3.10)$$

ความยาวแกนเอกของวงรี ( $a$ ) ในระนาบ CbCr คือ ระยะทางระหว่าง  $(V_{\pm Cb}, Cr)$

ความยาวแกนโทของวงรี ( $b$ ) ในระนาบ CbCr คือ ระยะทางระหว่าง  $(\bar{Cb}, V_{\pm Cr})$

สมการวงรี ที่นำมาใช้ในการแยกบริเวณสีผิว ถูกแสดงในสมการที่ (3.11)

$$\frac{L_1^2}{a^2} + \frac{L_2^2}{b^2} = 1 \quad (3.11)$$

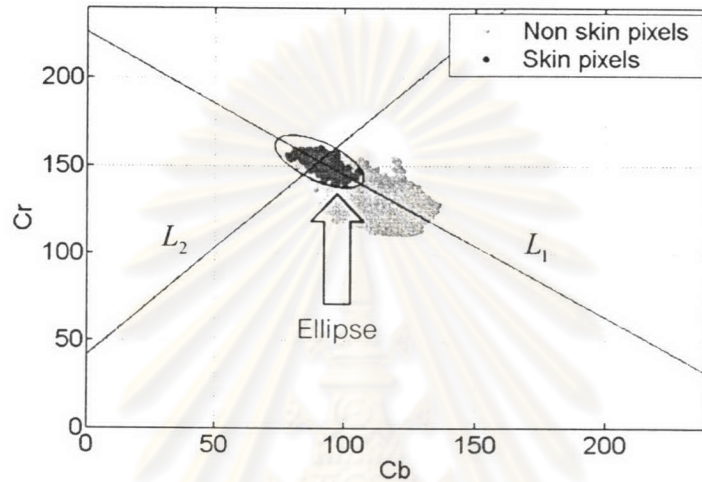
โดยที่  $L_1 : ACb + BCr + C = 0$  คือ เส้นตรงที่เป็นแกนเอกของวงรี

$L_2 : BCb - ACr + D = 0$  คือ เส้นตรงที่เป็นแกนโทของวงรี

$a$  คือ ความยาวแกนเอกของวงรี

$b$  คือ ความยาวแกนโทของวงรี

การคำนวณหาสมการเส้นตรง  $L_1$  ทำได้โดยใช้ค่าความชันเท่ากับ  $\tan \theta$  และจุดที่เส้นตรงผ่าน 1 จุด คือ  $(\bar{C}_b, \bar{C}_r)$  ส่วนสมการ  $L_2$  คือเส้นที่ตั้งฉากกับ  $L_1$  และตัดกันที่จุด  $(\bar{C}_b, \bar{C}_r)$  สร้างได้ในทำนองเดียวกัน (ความชันเท่ากับ  $-\frac{1}{\tan \theta}$ ) ในที่นี่เราจะใช้สมการที่ (3.11) ซึ่งเป็นสมการวงรี ในการกำหนดขอบเขตล้อมรอบบริเวณส่วนที่เป็นสีผิว



รูปที่ 3.7 ขอบเขตของวงรีเมื่อกำหนดขีดแบ่งเท่ากับ 4

ในการตัดสินใจลักษณะว่าจุดภาพนั้นว่าเป็นสีผิวหรือไม่ จะตรวจสอบจากค่า CbCr ของจุดภาพ หากพบว่าอยู่ภายในวงรี จะกำหนดจุดภาพนั้นเป็นสีผิว ดังแสดงในสมการที่ (3.12)

$$MSkin(i, j) = \begin{cases} 1, & \text{if (enclose with ellipse)} \\ 0, & \text{else} \end{cases} \quad (3.12)$$

เมื่อ  $MSkin(i, j)$  คือ มาสก์การแยกส่วนสีผิว

### 3.3 ตัวดำเนินการแบบปิด

การลดสัญญาณรบกวนประเภทหลุมขนาดเล็กสามารถทำได้โดยผ่านกระบวนการประมวลผลภาพดิจิทัล คือ การทำไคเลชันและอีโรชัน หากนำ 2 กระบวนการนี้มาใช้ร่วมกัน ในลำดับของไคเลชันและตามด้วยอีโรชัน จะเรียกว่า ตัวดำเนินการแบบปิด ซึ่งเป็นขั้นตอนที่นำมาใช้หลังจากแยกส่วนสีผิว



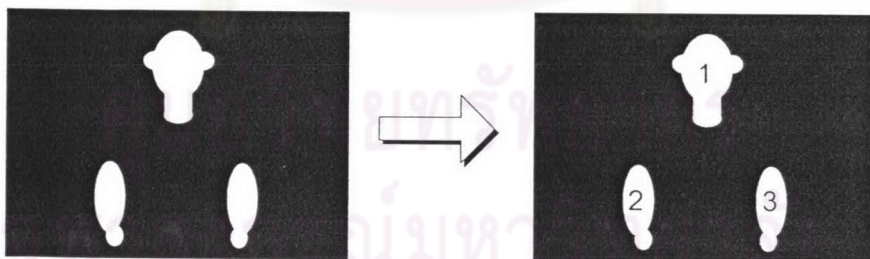
ด้วยแบบจำลองวงรีเพื่อแก้ไขข้อผิดพลาดของการตรวจหาสีผิว ตัวดำเนินการแบบปิดที่นำมาใช้นั้น มีความจำเป็นอย่างยิ่งที่จะต้องเลือกโครงสร้างส่วนย่อย (Structuring element) ที่เหมาะสม โดยการหาโครงสร้างส่วนย่อยที่ให้ผลดีที่สุดซึ่งได้จากการทดลองมีมาสก์ดังแสดงตามสมการที่ (3.13)

$$SE = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (3.13)$$

เมื่อ  $SE$  คือ มาสก์ของโครงสร้างส่วนย่อย

### 3.4 การทำป้ายและหาตำแหน่ง (Labeling and Localization)

การทำป้าย (Labeling) เป็นกระบวนการที่ต่อจากการแยกสีผิวและตัวดำเนินการแบบปิด (Closing operator) การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน เพื่อกำหนดตำแหน่งในส่วนที่เป็นจุดภาพกลุ่มบริเวณใบหน้าและมือ ซึ่งเลือกใช้วิธีการเข้ารหัสลูกโซ่แบบพิจารณาจุดภาพรอบข้าง 8 จุด ตำแหน่งของแต่ละกลุ่มในทุก ๆ เฟรมจะถูกเก็บไว้ในหน่วยความจำ เพื่อประโยชน์ในการนำไปใช้สำหรับการติดตามและแปลความหมายต่อไป



รูปที่ 3.8 หลังจากผ่านการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

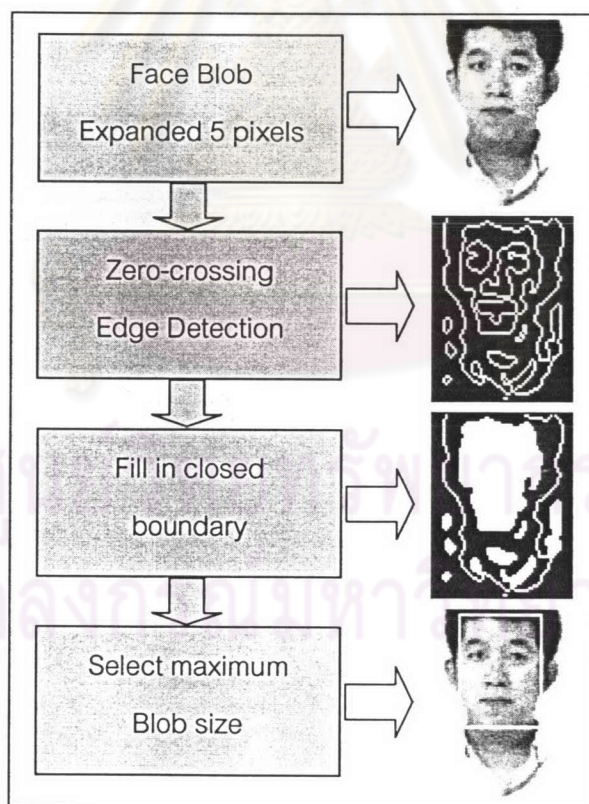
ข้อมูลของตำแหน่งที่จะนำไปใช้ในส่วนของติดตาม จะใช้จุดศูนย์กลางถ่วง (Center of Gravity) เป็นตัวแทนของข้อมูลในแต่ละกลุ่มเพื่อพิจารณาการเปลี่ยนตำแหน่งระหว่างเฟรม (Temporal) เพื่อประโยชน์ในการประมาณค่าในอนาคตและทำการติดตาม (Tracking) ต่อไป

### 3.5 การแยกส่วนสำคัญต่าง ๆ

ในหัวข้อนี้จะกล่าวถึงวิธีที่ใช้ในการแยกส่วนสำคัญได้แก่ ใบหน้า คอ ตา คิ้ว จมูก ปาก และข้อต่อระหว่างมือกับแขน ซึ่งสิ่งเหล่านี้เป็นประโยชน์ต่อการแปลความหมายภาษามือ ประกอบไปด้วยหัวข้อย่อยดังนี้ การแยกบริเวณใบหน้าออกจากส่วนคอ การแยกองค์ประกอบที่สำคัญบริเวณใบหน้า และการแยกระหว่างส่วนมือและแขนเชิงพื้นที่

#### 3.5.1 ขั้นตอนวิธีที่ใช้ในการแยกบริเวณใบหน้าออกจากส่วนคอ

เพื่อประโยชน์ในการนำไปแปลความหมายได้ถูกต้องเที่ยงตรง ผลจากการแยกส่วนสีผิวส่วนบลิบบบริเวณใบหน้า จะได้ส่วนคอ หรือบางครั้งรวมถึงไหล่ของผู้แปลภาษามือ วิทยานิพนธ์นี้จึงเสนอกรรมวิธีที่ใช้ในการแยกเฉพาะส่วนใบหน้าออกจากส่วนอื่น ๆ



รูปที่ 3.9 ขั้นตอนวิธีที่ใช้ในการแยกส่วนใบหน้าออกจากคอ

ขั้นตอนวิธีที่ใช้ในรูปที่ 3.9 มี 4 ขั้นตอน ประกอบด้วย

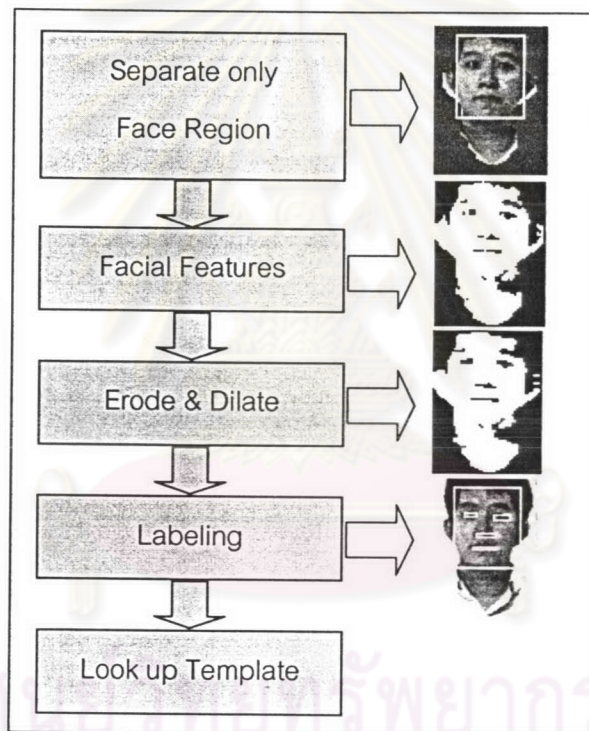
ขั้นที่ 1 ขยายขอบเขตจากตำแหน่งใบหน้าอีก 5 จุดภาพ เพื่อให้ครอบคลุมบริเวณใบหน้าทั้งหมดสำหรับการตรวจหาขอบ

ขั้นที่ 2 ตรวจหาขอบด้วยกรรมวิธีแบบไขว้ศูนย์ (Zero crossing)

ขั้นที่ 3 เติมจุดภาพสีขาวลงในบริเวณที่ถูกล้อมรอบ

ขั้นที่ 4 เลือกบริเวณที่ใหญ่ที่สุดและถือเอาบริเวณนั้นเป็นใบหน้า

### 3.5.2 การแยกองค์ประกอบที่สำคัญบริเวณใบหน้า (ตา คิ้ว จมูก และ ปาก)



รูปที่ 3.10 ขั้นตอนวิธีที่ใช้ในการหาค่าองค์ประกอบสำคัญบนใบหน้า

จากหัวข้อย่อยที่ผ่านมา เมื่อได้ตำแหน่งของส่วนที่เป็นใบหน้าแล้ว หากต้องการรายละเอียดที่มากกว่านั้นสามารถทำได้โดยใช้ขั้นตอนวิธีดังแสดงในรูปที่ 3.10 ซึ่งส่วน ตา คิ้ว จมูก และ ปาก จะมีค่าสีที่แตกต่างกับสีผิวบนใบหน้า เมื่อพิจารณาค่าองค์ประกอบทางแสง (Y) พบว่ามีค่าต่ำเพราะค่าสีใกล้เคียงกับสีดำ กระบวนการในขั้นตอนนี้ คือ หาค่าเฉลี่ยขององค์ประกอบทางแสงเฉพาะจุดภาพ



บริเวณส่วนสีผิว เพื่อกำหนดขีดแบ่ง (threshold) ตามสมการที่ (3.14) และนำมาทำการสร้างมาสก์ใหม่ โดยจุดภาพใดมีค่า  $Y$  น้อยกว่าขีดแบ่ง จะถือว่าเป็นส่วนตา คิ้ว จมูก และ ปาก จากนั้นใช้ตัวดำเนินการแบบปิดเพื่อกำจัดสัญญาณรบกวนประเภทหลุมขนาดเล็ก และ วิเคราะห์บริเวณที่เชื่อมต่อกันเพื่อกำหนดตำแหน่งขององค์ประกอบสำคัญบนใบหน้า

$$\text{Threshold} = \bar{Y} - \eta_2 * \text{std}(Y) \quad (3.14)$$

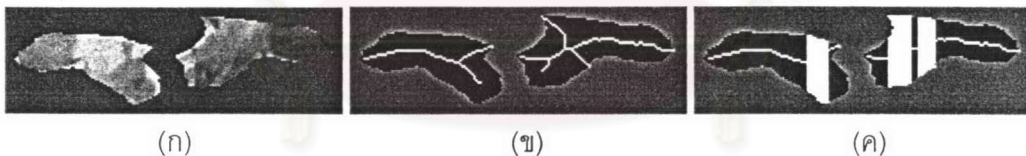
$\bar{Y}$  คือ ค่าเฉลี่ยขององค์ประกอบทางแสง ( $Y$ ) ของจุดภาพบริเวณส่วนสีผิว

$\eta_2$  คือ ค่าถ่วงน้ำหนักในที่นี้ใช้ค่าเท่ากับ 1.5

$\text{std}(Y)$  คือ ค่าเบี่ยงเบนมาตรฐานขององค์ประกอบทางแสง ( $Y$ ) ของจุดภาพบริเวณส่วนสีผิว

### 3.5.3 การแยกระหว่างส่วนมือและแขนเชิงพื้นที่

กรณีผู้แสดงภาษามือสวมเสื้อแขนสั้น การติดตามมือจำเป็นที่จะต้องแยกระหว่างข้อต่อส่วนมือและแขน เพื่อให้เกิดความถูกต้องในการติดตามการเคลื่อนที่ วิธีการหนึ่งในการแยกเชิงพื้นที่ คือ กระบวนการทinning (Thinning) ซึ่งมีวิธีการดังนี้



(ก)

(ข)

(ค)

รูปที่ 3.11 (ก) มาสก์ที่ได้จากการแยกส่วนสีผิว

(ข) มาสก์ที่ผ่านกระบวนการทinning

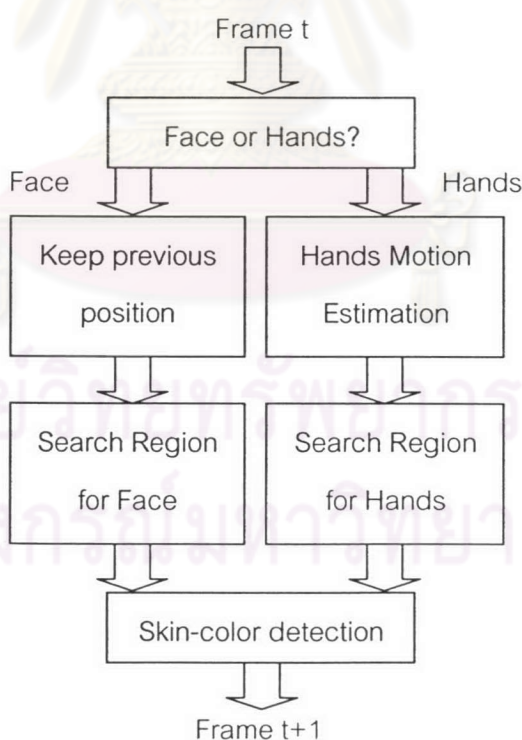
(ค) บริเวณที่ถูกกำหนดเป็นมือ

จากรูปที่ 3.11 (ก) เป็นผลลัพธ์ของบล็อบบริเวณมือทั้งสองที่ได้จากการแยกส่วนสีผิว เมื่อนำมาสก์ของมือผ่านกระบวนการทinningแล้ว จะได้ผลดังแสดงในรูปที่ 3.11 (ข) ซึ่งจะเห็นว่า ส่วนที่มีการหักมุมบริเวณข้อมือหรือนิ้ว จะมีการแยกออกของเส้นจำนวนมาก ซึ่งสามารถนำผลที่ได้จากกระบวนการทinningไปวิเคราะห์แยกส่วนของมือ โดยถือเอากลุ่มที่มีการแตกของเส้นเหล่านี้เป็นมือ ดังรูปที่ 3.11 (ค)

หมายเหตุ กระบวนการอินนิงจะไม่สามารถแยกบริเวณแขนกับมือได้ในวิดีโอที่คนเฟรมที่ไม่มีการหักมุมระหว่างแขนกับมืออย่างชัดเจน หรือ กรณีมาส์กที่ได้จากการแยกส่วนสีผิวมีการสูญเสียข้อมูลส่วนมือไป

### 3.6 การติดตาม (Tracking)

ในส่วนการติดตาม ระบบจะทำการตัดสินใจว่าบล็อบบริเวณสีผิวที่ได้มาบล็อบบนโดเป็นบริเวณใบหน้า โดยใช้คุณลักษณะเด่นของขนาดก่อนในเบื้องต้น หากเกิดความไม่แน่ใจเมื่อพิจารณาจากคุณลักษณะของขนาด (กรณีผู้แปลภาษามือสวมเสื้อแขนสั้น) จะใช้องค์ประกอบสำคัญบนใบหน้ามาช่วยในการตัดสินใจเพิ่มเติม หลังจากทราบว่าบล็อบบนโดเป็นบริเวณใบหน้าหรือมือแล้ว ในส่วนใบหน้า นั้นเนื่องจากการเปลี่ยนตำแหน่งไม่มาก การตรวจหาในเฟรมถัดไปจึงทำได้โดยง่ายเพียงขยายขอบเขตการตรวจหาสีผิวจากตำแหน่งอ้างอิงของเฟรมก่อนหน้าเท่านั้น แต่ในส่วนมือมีการเคลื่อนที่ค่อนข้างมากจึงต้องสร้างสมการในการประมาณการเคลื่อนที่ เพื่อกำหนดขอบเขตในการตรวจหาสีผิวที่เหมาะสมต่อไป ซึ่งขั้นตอนวิธีที่นำมาใช้แสดงในรูปที่ 3.12

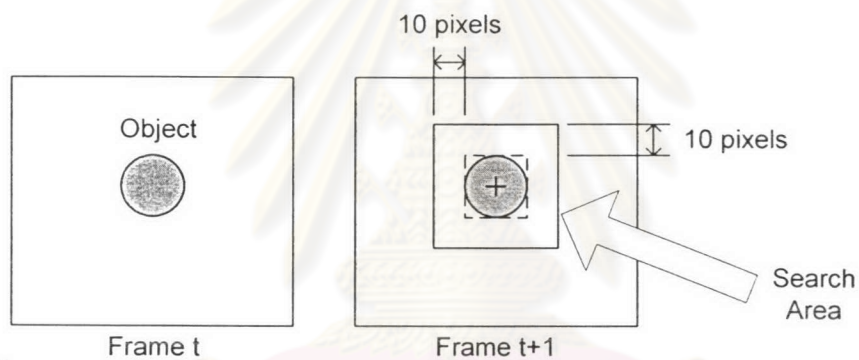


รูปที่ 3.12 ขั้นตอนวิธีในการติดตามใบหน้าและมือ

### 3.6.1 การหาตำแหน่งที่ใช้เป็นค่าบรรทัดฐาน

ในขั้นตอนแรก จะทำการหาตำแหน่งที่ใช้เป็นค่าบรรทัดฐาน โดยในเฟรมแรกจะทำการตรวจหาและกำหนดตำแหน่งของใบหน้าและมือด้วยแบบจำลองสีผิว โดยพิจารณาจากทั้งเฟรม ส่วนในเฟรมถัด ๆ ไป วิธีตรวจหาสีผิว ทำโดยใช้ตำแหน่งสูงสุดและต่ำสุดของแต่ละบล็อก ที่ได้อ้างอิงจากเฟรมก่อนหน้า รวมกับขอบเขตในการค้นหาอีก 10 จุดภาพ (ตามสมมติฐานที่ว่าวิถีทัศน์เฟรมที่อยู่ติดกันจะไม่เปลี่ยนแปลงอย่างทันทีทันใด) ดังสมการที่ (3.15) และวิธีการกำหนดขอบเขตตามรูปที่ 3.13

$$\begin{bmatrix} x_{Search}(t+1) \\ y_{Search}(t+1) \end{bmatrix} = \begin{bmatrix} x(t) \pm 10 \\ y(t) \pm 10 \end{bmatrix} \quad (3.15)$$



รูปที่ 3.13 การกำหนดขอบเขตในการตรวจหาวัตถุเฟรมถัดไป

### 3.6.2 การประมาณการเคลื่อนที่ของมือ

หากนำสมการการเคลื่อนที่ของวัตถุในแนวเส้นตรงมาประยุกต์ใช้ในการประมวลผลลำดับวิดีโอทัศน์ เพื่อติดตามการเคลื่อนที่ของมือ เนื่องจากเวลาระหว่างเฟรมแต่ละเฟรมมีค่าเท่ากันจึงสามารถละทิ้งการพิจารณาองค์ประกอบทางเวลาได้ ดังนั้นความเร็วในการเปลี่ยนตำแหน่งของจุดศูนย์กลางในแต่ละแกนสามารถหาได้ตามสมการที่ (3.16)

$$\begin{bmatrix} V_x(t) \\ V_y(t) \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} - \begin{bmatrix} x(t-1) \\ y(t-1) \end{bmatrix} \quad (3.16)$$



เมื่อ  $x(t)$  และ  $y(t)$  คือ ตำแหน่งของจุดศูนย์กลางวงวัตถุในเฟรมที่  $t$  ทางแกนอนและแกนตั้งตามลำดับ

การหาความเร่งในการเคลื่อนที่ได้จากการเปลี่ยนแปลงความเร็วของเฟรมที่อยู่ติด ๆ กัน ตามสมการที่ (3.17)

$$\begin{bmatrix} A_x(t) \\ A_y(t) \end{bmatrix} = \begin{bmatrix} V_x(t) \\ V_y(t) \end{bmatrix} - \begin{bmatrix} V_x(t-1) \\ V_y(t-1) \end{bmatrix} \quad (3.17)$$

### การติดตามมือเมื่อใช้ค่าความเร็วและความเร่งมาร่วมพิจารณา

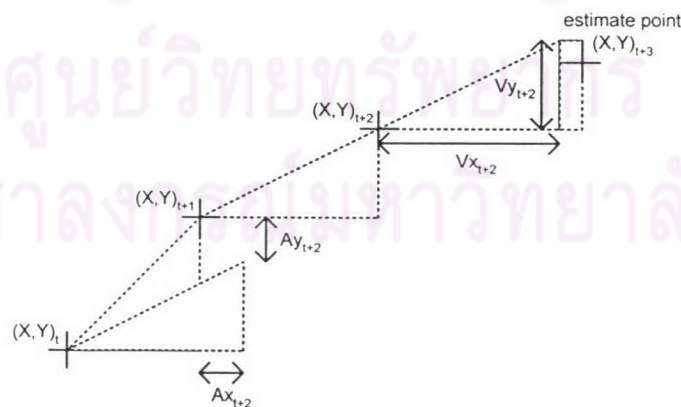
การประมาณค่าตำแหน่งของมือสำหรับเฟรมถัดไป โดยพิจารณาจากค่าความเร็วและความเร่งในการเคลื่อนที่แต่ละบลิบ จากข้อมูลปัจจุบันและก่อนหน้า โดยในที่นี้จะแสดงสมการประมาณการเคลื่อนที่ 2 แบบ แบบที่ 1 ตามสมการที่ (3.18) และแบบที่ 2 [15] แสดงในสมการที่ (3.19) ซึ่งมีกรรมวิธีดังรูปที่ 3.14

แบบที่ 1

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} V_x(t) \\ V_y(t) \end{bmatrix} \quad (3.18)$$

แบบที่ 2

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} x(t) + V_x(t) \\ y(t) + V_y(t) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} A_x(t) \\ A_y(t) \end{bmatrix} \quad (3.19)$$



รูปที่ 3.14 การประมาณการเคลื่อนที่แบบที่ 2 [15]

ขั้นตอนวิธีที่ใช้ในเฟรมที่ 1, 2 และ 3 จะใช้วิธีการเดียวกับ การหาตำแหน่งที่เป็นค่าบรรทัดฐาน เพื่อเก็บข้อมูลของความเร็วและความเร่งแต่ละบลิบ ส่วนเฟรมถัด ๆ ไปจะประมวลผลจากสมการที่ (3.18) หรือ (3.19) ร่วมกับขยายขอบเขตในการค้นหาตามสมการที่ (3.20) เพื่อทำการค้นหาตำแหน่งใหม่ของแต่ละบลิบในเฟรมนั้น

การกำหนดบริเวณที่จะนำมาใช้ตรวจหาสีผิวในเฟรมถัดไป

$$\begin{bmatrix} x_{Search}(t+1) \\ y_{Search}(t+1) \end{bmatrix} = \begin{bmatrix} x(t+1) \pm W \\ y(t+1) \pm W \end{bmatrix} \quad (3.20)$$

$x_{search}(t+1)$  และ  $y_{search}(t+1)$  คือ บริเวณในการตรวจหาวัตถุในแนวนอนและแนวตั้ง

$x(t+1)$  และ  $y(t+1)$  คือ ตำแหน่งในแนวนอนและแนวตั้งที่ประมาณได้

$W$  คือ ขนาดหน้าต่างที่ถูกขยายขึ้นเพื่อการตรวจหา

### 3.6.3 การปรับขนาดขอบเขตการค้นหาให้สอดคล้องกับความเร็ว

การปรับตัวของขั้นตอนวิธีเพื่อให้เกิดความถูกต้องในการหาตำแหน่ง จากการทดลองพบว่า สมการการประมาณการเคลื่อนที่ ที่นำมาใช้จะเกิดความผิดพลาดระหว่างเฟรมที่มีการเปลี่ยนตำแหน่งของวัตถุมากหรือมีการเปลี่ยนแปลงความเร็วในการเคลื่อนที่อย่างทันทีทันใดนั่นเอง เงื่อนไขในการตัดสินใจเลือกขอบเขตการค้นหาพิจารณาได้ตามสมการที่ (3.21)

$$W_i = \begin{cases} W & \text{if } A_{x,y} < \varepsilon \\ W_{exp} & \text{if } A_{x,y} \geq \varepsilon \end{cases} \quad (3.21)$$

$W$  คือ ขนาดหน้าต่างที่ถูกขยายขึ้นเพื่อการตรวจหา

$W_{exp}$  คือ ขนาดหน้าต่างที่ถูกขยายขึ้นมากกว่าเดิมเพื่อการตรวจหา

$\varepsilon$  คือ ขีดแบ่งซึ่งได้จากการทดลองมีค่าเท่ากับ 4