

การรู้จำใบหน้า 3 มิติแบบไม่แปรเปลี่ยนตามการแสดงออกโดยใช้บริบทรูปร่าง



นางสาว มนตรา บัวเพชร

สถาบันวิทยบริการ

จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

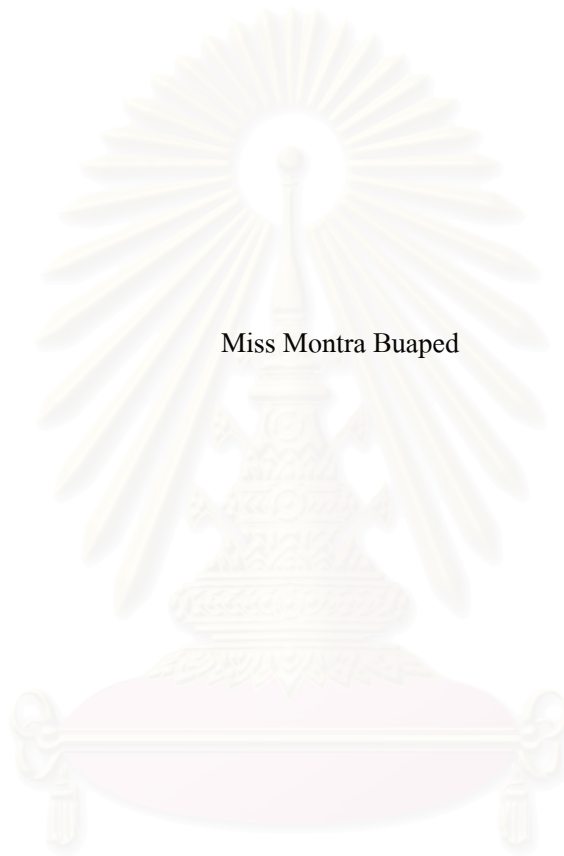
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

EXPRESSION - INVARIANT 3D FACE RECOGNITION USING SHAPE CONTEXT



Miss Montra Buaped

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering
Department of Electrical Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2007

Copyright of Chulalongkorn University

มนตรา บัวเพชร : การรู้จำใบหน้า 3 มิติแบบไม่แปรเปลี่ยนตามการแสดงออกโดยใช้บริบทรูปร่าง.
(EXPRESSION - INVARIANT 3D FACE RECOGNITION USING SHAPE CONTEXT)
อ. ที่ปรึกษา : อ. ดร. ชาญรัช ปลื้มปิติวิริยะเวช, 90 หน้า.

กระบวนการรู้จำใบหน้าเป็นเทคโนโลยีการประมวลผลภาพดิจิทัลที่มีประโยชน์ในการประยุกต์ใช้งานหลายด้าน และมีกระบวนการพัฒนาขั้นตอนวิธีในแบบต่างๆที่หลากหลาย ตั้งแต่การใช้ภาพถ่าย 2 มิติจนเริ่มมีการนำภาพถ่ายใบหน้าแบบ 3 มิติเข้ามาใช้ในปัจจุบัน ข้อดีของภาพ 3 มิติคือสามารถจัดเก็บข้อมูลภาพใบหน้าได้มากกว่าภาพ 2 มิติ แต่ถึงกระนั้นกระบวนการรู้จำใบหน้าแบบ 3 มิติที่เคยมีการนำเสนอกันมาก็ได้รับผลกระทบจากการแปรผันเนื่องจากแสดงออกบนใบหน้า ซึ่งทำให้ความแม่นยำของกระบวนการลดลง อีกทั้งเทคโนโลยีการถ่ายภาพ 3 มิติในปัจจุบัน ยังทำได้ยากกว่าภาพ 2 มิติ ดังนั้นขั้นตอนวิธีที่ต้องใช้ภาพกลุ่มทดลองจำนวนมากเพื่อสร้างการเรียนรู้ให้กับระบบจึงอาจไม่เหมาะสมต่อการนำไปใช้งานนัก

วิทยานิพนธ์ฉบับนี้ นำเสนอขั้นตอนวิธีรู้จำใบหน้า 3 มิติ โดยใช้กระบวนการที่สามารถลดผลกระทบจากการแปรผันเนื่องจากการแสดงออกของใบหน้าในภาพใบหน้า 3 มิติและใช้ภาพกลุ่มทดลองจำนวนน้อยโดยใช้บริบทรูปร่างเป็นตัวดึงคุณลักษณะของภาพ ซึ่งมีคุณสมบัติไม่เปลี่ยนเมื่อภาพใบหน้ามีการแปลงจากผลของการแสดงออกเกิดขึ้น นอกจากนี้ยังมีการพัฒนากระบวนการหาบริบทรูปร่างแบบดั้งเดิมที่มีการคำนวณซับซ้อนให้มีความสะดวกในการประมวลผลมากยิ่งขึ้น ซึ่งประกอบด้วยการใช้ฮิสโตแกรมแบบลอการิทึมทรงกลม (logarithm spherical histogram) การปรับแนวภาพ (alignmnet) และการใช้ตำแหน่งขององค์ประกอบสำคัญบนใบหน้าเพื่อลดขั้นตอนการหาค่าเหมาะที่สุด ซึ่งทำให้กระบวนการคำนวณบริบทรูปร่างมีความเหมาะสมกับการนำไปใช้กับภาพใบหน้า ต่อจากนั้นเป็นการพัฒนาวิธีนำคุณลักษณะที่คำนวณได้ไปใช้ในการพัฒนากระบวนการรู้จำใบหน้า และการเพิ่มประสิทธิภาพให้กับกระบวนการรู้จำใบหน้าแบบ 3 มิติที่ได้ โดยในขั้นตอนสุดท้ายการวิเคราะห์ข้อดี ข้อด้อย ของกระบวนการรู้จำใบหน้าที่ได้นำเสนอ

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อนิสิต.....
สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
ปีการศึกษา.....2550.....

4970511821 : MAJOR ELECTRICAL ENGINEERING

KEY WORD: 3D FACE RECOGNITION/ EXPRESSION INVARIANT/ TRANSFORMATION
INVARIANT/ SHAPE CONTEXT/ CANONICAL IMAGE

MONTRA BUAPED : EXPRESSION - INVARIANT 3D FACE RECOGNITION USING SHAPE
CONTEXT. THESIS ADVISOR : CHARNCHAI PLUEMPITWIRIYAJEJ, Ph.D., 90 pp.

Face recognition is the one of the most important areas in digital image processing framework that has been used in many practical applications. In the past, research in face recognition is usually based on plain photograph or 2D images. As the 3D technology has been rapidly improved, face recognition using 3D facial images is more attractive. The strength of a 3D facial image is that it has more facial information than a 2D one. However, most 3D face recognition algorithms are still effected by various facial expressions that can lower its accuracy. Moreover, the requirement to have many images in the train of system is not so convenient for practical use.

This thesis proposes 3D face recognition algorithm that is immuned to various expressions using a small set of training images. In order to overcome the expression varying problem, it uses 3D shape context, which is invariant to shape transformation, of the 3D faces. To reduce the complexity in traditional shape context method, on improved histogram computation is introduced and the image alignment further reduces the time required in the optimization process, making the process more suitable for practical face recognition system. Then, the features extracted by shape context are used to develop an effective face recognition algorithm. At last, the analysis and conclusion about the proposed algorithm are shown.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department.....Electrical Engineering..... Student's signature.....
Field of study.....Electrical Engineering..... Advisor's signature.....
Academic year2007.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของท่านอาจารย์ที่ปรึกษาวิทยานิพนธ์ อาจารย์ ดร.ชาญชัย ปल्लीปิติวิริยะเวช ซึ่งให้คำชี้แนะและข้อคิดเห็นต่างๆ พร้อมทั้งแรงกระตุ้นและแรงบันดาลใจในการทำวิทยานิพนธ์ด้วยดีตลอดมา

ขอขอบคุณคณะกรรมการสอบวิทยานิพนธ์ทุกท่าน ที่ได้ตรวจสอบวิทยานิพนธ์และให้คำแนะนำที่เป็นประโยชน์อย่างยิ่ง เพื่อปรับปรุงแก้ไขวิทยานิพนธ์ให้มีความถูกต้องสมบูรณ์ยิ่งขึ้น

ขอขอบคุณอาจารย์ทุกๆท่านที่ให้ความรู้ในการศึกษา รวมทั้งคำแนะนำต่างๆ ซึ่งเป็นพื้นฐานของความรู้สำหรับประยุกต์ใช้กับวิจัยได้เป็นอย่างดี

ขอขอบคุณ รุ่นพี่ทุกคน ที่ให้ความช่วยเหลือและคำแนะนำอันเป็นประโยชน์อย่างยิ่ง รุ่นน้องทุกคน ที่คอยเป็นกำลังใจในการทำงานวิจัย และเพื่อนทุกคน ที่คอยดูแลและอยู่เคียงข้างกันเสมอ

ขอขอบคุณทุน โครงการศิษย์ก้นกุฏิ ที่ช่วยสนับสนุนค่าใช้จ่ายสำหรับการศึกษา

สุดท้ายขอกราบขอบคุณ คุณพ่อสาคร บัวเพชร และ คุณแม่จุฬา บัวเพชร ตลอดจนญาติพี่น้องทุกคนที่คอยสนับสนุนในทุกๆด้านและเป็นแรงใจให้ตลอดมา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ	ญ
ดัชนีคำศัพท์	ฐ
บทที่ 1 บทนำ.....	1
1.1. ความเป็นมาและความสำคัญของปัญหา.....	1
1.2. วัตถุประสงค์.....	3
1.3. ขอบเขตของวิทยานิพนธ์	3
1.4. ขั้นตอนและวิธีการดำเนินการ	4
1.5. ประโยชน์ที่คาดว่าจะได้รับ.....	4
บทที่ 2 หลักการและทฤษฎีทั่วไป	5
2.1. กระบวนการรู้จำใบหน้าแบบ 2 มิติ (2D face recognition).....	5
2.1.1 ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม (appearance-based algorithm).....	9
2.1.2 ขั้นตอนวิธีอิงแบบจำลองภาพ (model-based algorithm).....	15
2.2. กระบวนการรู้จำใบหน้าแบบ 3 มิติ (3D face recognition).....	18
2.2.1 ลักษณะของข้อมูลภาพแบบ 3 มิติ.....	18
2.2.2 การแปลงเรขาคณิต (geometric transformation).....	19
2.2.3 การสร้างภาพ 3 มิติ.....	21
2.2.4 การรู้จำใบหน้าแบบ 3 มิติ (3D face recognition)	23
2.3. การรู้จำใบหน้าแบบ 3 มิติโดยไม่แปรผันตามการแสดงออกของบุคคลในภาพ.....	26
2.3.1 การหาระยะทางจีโอเดสิกบนพื้นผิว.....	27
2.3.2 มัลติไดเมนชันแนล สเกลลิง (multidimensional scaling).....	28
2.4. บริบทรูปร่าง (shape context)	31
2.4.1 บริบทรูปร่างสำหรับภาพ 2 มิติ	31
2.4.2 บริบทรูปร่างสำหรับภาพ 3 มิติ	33
2.4.3 การทดสอบไค-สแควร์ (chi-squared test)	35
2.4.4 การหาจุดที่มีความสัมพันธ์กันด้วยการหาค่าเหมาะที่สุด.....	36
บทที่ 3 วิธีดำเนินการวิจัย.....	38
3.1. การปรับปรุงบริบทรูปร่างสำหรับภาพ 3 มิติเพื่อใช้กับภาพใบหน้า	38

3.1.1 การปรับปรุงการหาขั้วของดาวารางในทิศตั้งฉากของพื้นผิวที่จุดอ้างอิงด้วยการปรับแนวโดยการวิเคราะห์ห่องค์ประกอบหลัก (PCA alignment).....	39
3.1.2 การปรับขนาดภาพเพื่อหลีกเลี่ยงขั้นตอนการหาค่าถ่วงน้ำหนักที่เปลี่ยนแปลงตามความหนาแน่นของจุดภาพและขนาดของช่อง.....	42
3.1.3 การกำหนดจุดอ้างอิงที่ตำแหน่งองค์ประกอบสำคัญบนใบหน้า.....	43
3.2. การแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ	44
3.3. การแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ	45
บทที่ 4 ผลการทดลองและการวิเคราะห์ผล.....	47
4.1. รายละเอียดฐานข้อมูลภาพ.....	47
4.1.1 กลุ่มฐานข้อมูลภาพสังเคราะห์	47
4.1.2 กลุ่มฐานข้อมูลภาพใบหน้าบุคคล	49
4.2. การทดลองแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ	52
4.2.1 การทดลองการแยกแยะภาพสังเคราะห์เมื่อจำลองการแปลงในภาพสังเคราะห์ด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ	52
4.2.2 การทดลองการแยกแยะภาพใบหน้าที่มีการแสดงออกด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ	59
4.3. การทดลองแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ	69
4.3.1 การทดลองแยกแยะภาพสังเคราะห์เมื่อจำลองการแปลงในภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอดบนภาพ	69
4.3.2 การทดลองแยกแยะภาพใบหน้าที่มีการแสดงออกด้วยฮิสโทแกรมของจุดยอดบนภาพ.....	70
4.4. การเพิ่มประสิทธิภาพของกระบวนการ แยกแยะภาพใบหน้าด้วยฮิสโทแกรมของจุดยอด.....	72
4.5. การทดลองผลความแม่นยำเมื่อจำนวนจุดยอดในภาพหายไป.....	75
4.6. การทดลองเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีที่นำเสนอ	79
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	80
5.1. สรุปผลการวิจัย.....	80
5.2. ข้อเสนอแนะสำหรับงานวิจัยในอนาคต	81
รายการอ้างอิง	82
ภาคผนวก	84
บทความที่ได้รับการเผยแพร่	85
ประวัติผู้เขียนวิทยานิพนธ์.....	90

สารบัญตาราง

ตาราง	หน้า
ตาราง 2. 1 การเปรียบเทียบข้อดีข้อด้อยของกระบวนการรู้จำใบหน้าแบบต่างๆ	25
ตาราง 4. 1 ผลการแยกแยะภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอดบนภาพ	58
ตาราง 4. 2 ผลการแยกแยะภาพบุคคลด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ	68
ตาราง 4. 3 ผลการแยกแยะภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอดบนภาพ	69
ตาราง 4. 4 ผลการแยกแยะภาพบุคคลด้วยฮิสโทแกรมของจุดยอดบนภาพ	71
ตาราง 4. 5 ผลการเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีแบบต่างๆ	79



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพประกอบ	หน้า
รูปที่ 2.1 ภาพบุคคลเดียวกันที่มีการแปรผันทางตำแหน่ง (ก) ภาพมุมมองหน้าตรง (ข) ภาพมุมมองเอียงซ้าย (ค) ภาพมุมมองเอียงขวา	6
รูปที่ 2.2 ภาพบุคคลเดียวกันที่มีการแปรผันของแสง (ก) ภาพที่มีแสงปกติ (ข) ภาพที่มีแสงมืดเกินไป (ค) ภาพที่มีแสงแปลกไปจากปกติ	7
รูปที่ 2.3 ภาพบุคคลเดียวกันที่มีการแปรผันเนื่องจากการแสดงออกของใบหน้า (ก) ภาพอารมณ์ปกติไม่มีการ แสดงออกทางอารมณ์ (ข) ภาพใบหน้าขณะยิ้ม (ค) ภาพใบหน้าขณะโกรธ	7
รูปที่ 2.4 ภาพบุคคลเดียวกันที่มีวัตถุมาบังพื้นที่บริเวณใบหน้า (ก) ภาพที่ไม่มีวัตถุมาบังใบหน้า (ข) ภาพที่ มีแว่นดำมาบังใบหน้า (ค) ภาพที่มีที่มีทรงผมมาบังใบหน้า	8
รูปที่ 2.5 โครงสร้างการทำงานของกระบวนการรู้จำใบหน้าเบื้องต้น	9
รูปที่ 2.6 แผนภาพการหาเวกเตอร์คุณลักษณะโดยขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม	10
รูปที่ 2.7 ความแตกต่างของเวกเตอร์คุณลักษณะในปริภูมิภาพหลังทำการแปลงเมทริกซ์ภาพด้วย (ก) การ วิเคราะห์องค์ประกอบหลัก เทียบกับ (ข) การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น	13
รูปที่ 2.8 ความแตกต่างของเวกเตอร์คุณลักษณะในปริภูมิภาพเมื่อทำการแปลงเมทริกซ์ภาพด้วย (ก) การ วิเคราะห์องค์ประกอบหลัก เทียบกับ (ข) การวิเคราะห์องค์ประกอบอิสระ	13
รูปที่ 2.9 การเปรียบเทียบความแตกต่างของการแปลงเวกเตอร์ด้วย (ก) กระบวนการวิเคราะห์ องค์ประกอบหลักแบบปกติ เทียบกับ (ข) การวิเคราะห์เคอร์เนลขององค์ประกอบหลัก.....	14
รูปที่ 2.10 การกำหนดปมและกิ่งตามขั้นตอนวิธีเปรียบเทียบกราฟแบบยึดหยุ่นโดยอิงโครงหน้า	15
รูปที่ 2.11 การกำหนดเจ็ทหลายแบบในแต่ละตำแหน่งปม	16
รูปที่ 2.12 การกำหนดจุดบนองค์ประกอบสำคัญบนใบหน้าด้วยขั้นตอนวิธี แบบจำลองลักษณะปรากฏแบบ แอคทีฟเมื่อภาพมีการแปรผันทางตำแหน่ง (ก) เมื่อใบหน้ามีการหันข้าง (ข) เมื่อใบหน้ามีการหันเอียง (ค) เมื่อเป็นภาพใบหน้าตรง	16
รูปที่ 2.13 ข้อมูลภาพ 3 มิติในขั้นตอนวิธีแบบจำลอง 3 มิติแบบเปลี่ยนสภาพได้	17
รูปที่ 2.14 โครงสร้างภาพวัตถุ 3 มิติ ลักษณะของจุดยอดและแผ่นผิวบนวัตถุ 3 มิติ (ก) ภาพวัตถุ 3 มิติ (ข) ข้อมูลรูปร่างของวัตถุ 3 มิติ (ค) จุดยอดและแผ่นผิวแบบ 3 เหลี่ยม	18
รูปที่ 2.15 (ก) ภาพก่อนเกิดการแปลง (ข) ภาพเกิดการเปลี่ยนแปลงตำแหน่งของจุดภาพ.....	19
รูปที่ 2.16 (ก) ภาพก่อนเกิดการแปลง (ข) ภาพเกิดการแปลงค่าระดับสีเทาของภาพ.....	19
รูปที่ 2.17 การแปลงวัตถุแบบแข็งเกร็ง (ก) รูปทรงของวัตถุเดิม (ข) การเลื่อนขนานจากวัตถุเดิม (ค) การ หมุนจากวัตถุเดิม.....	20
รูปที่ 2.18 การแปลงวัตถุแบบไม่แข็งเกร็ง (ก) รูปทรงของวัตถุเดิม (ข) การเอียงจากวัตถุเดิม (ค) การบิดงอ จากวัตถุเดิม	20
รูปที่ 2.19 ภาพที่ถ่ายต่างมุมกันจะมีข้อมูลภาพแตกต่างกัน	21
รูปที่ 2.20 การสร้างภาพ 3 มิติ จากกล้อง โครงสร้างแสง (ก) กล้องถ่ายภาพ 3 มิติแบบกล้องโครงสร้างแสง (ข) การถ่ายภาพแบบกล้องโครงสร้างแสง.....	22

รูปที่ 2.21 การสร้างภาพ 3 มิติ จากตัวรับภาพแบบเลเซอร์ (ก) ตัวรับภาพแบบเลเซอร์ (ข) การถ่ายภาพด้วย
ตัวรับภาพแบบเลเซอร์22

รูปที่ 2.22 ภาพแสดงการแบ่งส่วนพื้นที่ใบหน้าตามความโค้งของพื้นผิว24

รูปที่ 2.23 ภาพพิสัยที่สร้างจากขั้นตอนวิธีของ C. Heshner และคณะ24

รูปที่ 2.24 โครงสร้างการหาภาพแกนหลักของใบหน้า26

รูปที่ 2.25 การเปรียบเทียบความแตกต่างของระยะทางจีโอเดสิก(เส้นทึบ)และระยะทางยูคลิด(เส้นประ).....27

รูปที่ 2.26 การเคลื่อนที่ของหน้าคลื่นด้วยกระบวนการพาสท์มาร์ชชิง.....28

รูปที่ 2.27 (ก) แกนภาพด้วยปริภูมิคาร์ทีเซียน (ข) แกนภาพด้วยแกนลอการิทึมเชิงขั้ว32

รูปที่ 2.28 (ก) โครงสร้างของบริบทรูปร่าง (ข) ฮิสโทแกรมของจุดยอดที่จุดอ้างอิง p_i 33

รูปที่ 2.29 การเปรียบเทียบโครงสร้างของ (ก) ดาวตารางแบบ 2 มิติ (ข) ดาวตารางแบบ 3 มิติ34

รูปที่ 2.30 (ก) โครงสร้างของบริบทรูปร่างแบบ 3 มิติ (ข) ฮิสโทแกรมของจุดยอดที่จุดอ้างอิง.....34

รูปที่ 2.31 (ก) ทิศชี้ของดาวตาราง (ข) ทิศชี้ของดาวตารางเปลี่ยนแปลงตามเวกเตอร์ตั้งฉากที่จุดอ้างอิงใดๆ.....35

รูปที่ 2.32 คู่จุดที่มีความสัมพันธ์กันระหว่างภาพวัตถุ 2 ภาพแสดงด้วยลำดับเลขลำดับเดียวกัน36

รูปที่ 3.1 แกนของภาพใบหน้าจุดยอดมีการกระจายตัวมากไปน้อย 3 แกน ตามลำดับ39

รูปที่ 3.2 แสดงแนวทางการปรับแนวโดยใช้การวิเคราะห์ห้วงค์ประกอบหลักในภาพ 2 มิติ (ก) การกระจายตัว
ของข้อมูลในระนาบ 2 มิติ (ข) การหาแกนที่มีการกระจายตัวของข้อมูลสูงสุด หรือ แกนลักษณะเฉพาะ
(ค) การแสดงผลข้อมูลบนแกนลักษณะเฉพาะ41

รูปที่ 3.3 การเปรียบเทียบระหว่าง (ก) ทิศชี้ของดาวตารางเปลี่ยนแปลงตามเวกเตอร์ตั้งฉากที่จุดอ้างอิง (ข) ทิศ
ชี้ของดาวตารางคงที่ในทิศของเวกเตอร์แกนหลัก42

รูปที่ 3.4 ภาพบุคคลคนละคน มีตำแหน่งประกอบสำคัญบนใบหน้าตำแหน่งเดียวกันเสมอ44

รูปที่ 3.5 แผนผังกระบวนการการแยกแยะภาพด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ.....44

รูปที่ 3.6 แผนผังการแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ45

รูปที่ 4.1 รูปสังเคราะห์ที่ใช้ในการทดลอง 6 แบบ (ก) ตัวหมากรุกราชา (ข) ตัวหมากรุกราชินี (ค) ตัวหมากรุก
ขุนนาง (ง) ตัวหมากรุกอัศวินแบบรายละเอียดสูง (จ) ตัวหมากรุกอัศวินแบบรายละเอียดน้อย และ(ฉ)
ตัวหมากรุกป้อมปราการ47

รูปที่ 4.2 (ก) ภาพสังเคราะห์ที่เกิดการเลื่อนขนานจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไปใน 4 ทิศทาง
(ข) ภาพสังเคราะห์ที่เกิดการหมุนจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไปใน 4 มุม (ค) ภาพ
สังเคราะห์ที่เกิดการเอียงจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไป 4 ทิศทาง48

รูปที่ 4.3 ภาพกลุ่มทดสอบเป็นบุคคลที่มีการแสดงอารมณ์ปกติจำนวน 8 ภาพ (ก) ภาพบุคคลทดสอบที่ 1 (ข)
ภาพบุคคลทดสอบที่ 2 (ค) ภาพบุคคลทดสอบที่ 3 (ง) ภาพบุคคลทดสอบที่ 4 (จ) ภาพบุคคลทดสอบคน
ที่ 5 (ฉ) ภาพบุคคลทดสอบที่ 6 (ช) ภาพบุคคลทดสอบที่ 7 (ซ) ภาพบุคคลทดสอบที่ 849

รูปที่ 4.4 ภาพใบหน้าที่มีการแสดงภาพอารมณ์โกรธ 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ.....50

รูปที่ 4.5 ภาพใบหน้าที่มีการแสดงอารมณ์รังเกียจ 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ50

รูปที่ 4.6 ภาพใบหน้าที่มีการแสดงอารมณ์หวาดกลัว 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ.....50

รูปที่ 4.7 ภาพใบหน้าที่มีการแสดงอารมณ์สุข 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ.....51

รูปที่ 4.8 ภาพใบหน้าที่มีการแสดงอารมณ์เสียใจ 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ51

รูปที่ 4.9 ภาพใบหน้าที่มีการแสดงอารมณ์แปลกใจ 4 ระดับจากน้อยไปมาก(ก)-(ง) ตามลำดับ51

รูปที่ 4.10 ภาพเมทริกซ์ความคล้ายของตัวหมากรูกราชาติที่มีการแปลงแบบต่างๆ.....52

รูปที่ 4.11 ภาพเมทริกซ์ความคล้ายของตัวหมากรูกราจีนีที่มีการแปลงแบบต่างๆ.....53

รูปที่ 4.12 ภาพเมทริกซ์ความคล้ายของตัวหมากรูขุนนางที่มีการแปลงแบบต่างๆ54

รูปที่ 4.13 ภาพเมทริกซ์ความคล้ายของตัวหมากรูอัศวินแบบรายละเอียดสูงที่มีการแปลงแบบต่างๆ55

รูปที่ 4.14 ภาพเมทริกซ์ความคล้ายของตัวหมากรูอัศวินแบบรายละเอียดน้อยที่มีการแปลงแบบต่างๆ56

รูปที่ 4.15 ภาพเมทริกซ์ความคล้ายของตัวหมากรูปกป้องปราการที่มีการแปลงแบบต่างๆ57

รูปที่ 4.16 ภาพตำแหน่งองค์ประกอบบนใบหน้าซึ่งประกอบด้วย ตา คิ้ว จมูก ปาก และ โครงหน้า59

รูปที่ 4.17 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 1 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ60

รูปที่ 4.18 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 2 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ61

รูปที่ 4.19 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 3 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ62

รูปที่ 4.20 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 4 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ63

รูปที่ 4.21 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 5 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ64

รูปที่ 4.22 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 6 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ65

รูปที่ 4.23 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 7 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ66

รูปที่ 4.24 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 8 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ67

รูปที่ 4. 25 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสรัสมิ เมื่อกำหนดจำนวนช่องใน
 ทิสมุมงย มีจำนวน 18 ช่อง และจำนวนช่องในทิสมุมเอซิมัทมีจำนวน 18 ช่อง72

รูปที่ 4. 26 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสมุมงย เมื่อกำหนดจำนวนช่อง
 ในทิสรัสมิ มีจำนวน 12 ช่อง และจำนวนช่องในทิสมุมเอซิมัทมีจำนวน 18 ช่อง.....73

รูปที่ 4. 27 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสมุมเอซิมัท เมื่อกำหนดจำนวน
 ช่องในทิสรัสมิ มีจำนวน 12 ช่อง และจำนวนช่องในทิสมุมงยมีจำนวน 6 ช่อง.....74

รูปที่ 4. 28 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสมุมเอซิมัท เมื่อกำหนดจำนวน
 ช่องในทิสรัสมิ มีจำนวน 12 ช่อง และจำนวนช่องในทิสมุมงยมีจำนวน 18 ช่อง.....74

รูปที่ 4. 29 ภาพที่จุดยอดของใบหน้าที่ด้านข้างหายไปส่วนหนึ่ง.....75

รูปที่ 4. 30 (ก) ภาพใบหน้าที่ซีกขวามีการแบ่งช่วง (ข) ภาพใบหน้าที่จุดยอดด้านขวามีระยะมากกว่า 0.5 หายไป
 (ค) ภาพใบหน้าที่จุดยอดด้านขวามีระยะมากกว่า 0 หายไป76

รูปที่ 4. 31 กราฟผลความแม่นยำเมื่อจำนวนจุดยอดด้านข้างของภาพใบหน้าที่หายไป.....76

รูปที่ 4. 32 (ก) ภาพใบหน้าที่ด้านบนที่มีการแบ่งช่วง (ข) ภาพใบหน้าที่จุดยอดด้านบนมีระยะมากกว่า 0.5
 หายไป (ค) ภาพใบหน้าที่จุดยอดด้านบนมีระยะมากกว่า 0 หายไป.....77

รูปที่ 4. 33 กราฟผลความแม่นยำเมื่อจำนวนจุดยอดด้านบนของภาพใบหน้าที่หายไป77

รูปที่ 4. 34 (ก) ภาพใบหน้าที่ด้านล่างที่มีการแบ่งช่วง (ข) ภาพใบหน้าที่จุดยอดด้านล่างมีระยะน้อยกว่า -0.5
 หายไป (ค) ภาพใบหน้าที่จุดยอดด้านล่างมีระยะน้อยกว่า 0 หายไป78

รูปที่ 4. 35 กราฟผลความแม่นยำเมื่อจำนวนจุดยอดด้านล่างของภาพใบหน้าที่หายไป.....78

ดัชนีคำศัพท์

3D face recognition	กระบวนการรู้จำใบหน้าแบบ 3 มิติ
3D Morphable Model	ขั้นตอนวิธีแบบจำลอง 3 มิติแบบเปลี่ยนสภาพได้
Active Appearance Model	แบบจำลองลักษณะปรากฏแบบแอคทีฟ
affine transform	การแปลงสัมพรรค
algorithm	ขั้นตอนวิธี
alignment	กระบวนการปรับแนว
ambient light	แสงจากสภาพแวดล้อม
appearance-based algorithm	ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม
automatic	อัตโนมัติ
background subtraction	การกำจัดพื้นหลังของภาพ
bending	การบิดงอ
between-class scatter matrix	เมทริกซ์แสดงการกระจายตัวระหว่างกลุ่ม
biometrics	ชีวมิติ
branch	กิ่ง
brand recognition	การรู้จำตราสินค้า
canonical image	ภาพแบบบัญญัติของใบหน้า
cartesian coordinate	พิกัดคาร์ทีเซียน
centering matrix	เมทริกซ์ศูนย์กลาง
chi-squared test	การทดสอบไค-สแควร์
cluster scene	ฉากหลัง
computer vision	การรับรู้ภาพของคอมพิวเตอร์
correlation matrix	เมทริกซ์สหสัมพันธ์
correspondence point	จุดภาพที่มีความสัมพันธ์
cost function	ฟังก์ชันต้นทุน
covariance matrix	เมทริกซ์ความแปรปรวนร่วม
cross product matrix	เมทริกซ์ผลคูณไขว้
curvature	ความโค้ง
curvature-based algorithm	ขั้นตอนวิธีอิงความโค้งจากพื้นผิวของภาพวัตถุ

depth map	ภาพความลึก
digital image processing	การประมวลผลภาพดิจิทัล
direct light	แสงโดยตรง
discrete grid	ตารางแบบเต็มหน่วย
disparity map	ภาพความเหลื่อม
display	การแสดงผล
distortion	การบิดเบี้ยว
double centering	การหาศูนย์กลางซ้อน
edge map	ภาพขอบของวัตถุ
eigen decomposition	การแยกเวกเตอร์ลักษณะเฉพาะ
eigenvalue	ค่าลักษณะเฉพาะ
eigenvector	เวกเตอร์ลักษณะเฉพาะ
eigonal equation	สมการอีโกนอล
euclidean distance	ระยะทางยูคลิด
extrinsic property	คุณสมบัติภายนอก
face	แผ่นผิว
face identification	การระบุใบหน้า
face recognition	การรู้จำใบหน้า
face verification	การพิสูจน์ใบหน้า
facial expression variation	การแปรผันเนื่องจากการแสดงออกของใบหน้า
facial feature segmentation	การดึงข้อมูลขององค์ประกอบบนใบหน้า
facial occlusion	การมีวัตถุมาบังพื้นที่บริเวณใบหน้า
fast marching	กระบวนการฟาสต์ มาร์ชชิง
Feature – based elastic bunch graph matching	ขั้นตอนวิธีเปรียบเทียบกราฟแบบยืดหยุ่นโดยอิงโครงหน้า
feature extraction	การสกัดคุณลักษณะ
feature vector	เวกเตอร์คุณลักษณะ
fourier descriptor	ตัวบอกฟูรีเยร์
gabor kernels	กาบอร์เคอร์เนล
gaussian distribution	การกระจายตัวแบบเกาส์เซียน

geodesic distance	ระยะทางจีโอเดสิก
geometric space	ปริภูมิเรขาคณิต
geometric transformation	การแปลงเรขาคณิต
global property	คุณสมบัติครอบคลุม
graph theory	ทฤษฎีกราฟ
gray-level interpolation	การแปลงค่าระดับสีเทาของภาพ
grid	ตาราง
handwriting recognition	การรู้จำตัวอักษรจากลายมือ
hypothesis testing	การทดสอบสมมุติฐาน
illumination variation	การแปรผันของแสง
image enhancement	การปรับปรุงภาพ
image pre-processing	กระบวนการเตรียมภาพเบื้องต้น
image segmentation	การแบ่งส่วนภาพ
image space	ปริภูมิภาพ
Independent Component Analysis	การวิเคราะห์องค์ประกอบอิสระ
intensity level	ระดับความเข้ม
interaction	การโต้ตอบ
internetworking	การเชื่อมโยงระหว่างเครือข่าย
intrinsic property	คุณสมบัติภายใน
inverse transformation	การแปลงกลับ
Karhunen – Loeve transformation	การแปลงของคาร์เฮนเนน โลฟ
kernel function	เคอร์เนลฟังก์ชัน
Kernel Principle Component Analysis	การวิเคราะห์เคอร์เนลองค์ประกอบหลัก
landmark	จุดกำหนด
laser sensor	การสร้างภาพ 3 มิติ จากตัวรับภาพแบบเลเซอร์
level set	เซตระดับ
Linear Discriminant Analysis	การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น
linear subspace analysis	การวิเคราะห์ปริภูมิย่อยเชิงเส้น
local property	คุณสมบัติเฉพาะที่
logarithm spherical histogram	ฮิสโทแกรมแบบลอการิทึมทรงกลม

manual	ทำด้วยมือ
mapping	การจับคู่
Mereer's theorem	ทฤษฎีของเมอร์เร่
model-based	ขั้นตอนวิธีอิงแบบจำลองภาพ
monotonic	ทิศทางเดียว
multidimensional scaling	มัลติไดเมนชันแนล สเกลลิง
node	ปม
non-linear manifold analysis	การวิเคราะห์ปริภูมิย่อยแบบไม่เชิงเส้น
non-rigid body transform	การแปลงวัตถุแบบไม่แข็งเกร็ง
null hypothesis	สมมุติฐานว่าง
optimization	การหาค่าเหมาะที่สุด
PCA-based analysis	การวิเคราะห์เวกเตอร์แกนหลักของภาพ
pixel	จุดภาพ
polar histogram	ฮิสโทแกรมเชิงขั้ว
pole	ขั้ว
pose variation	การแปรผันทางตำแหน่ง
projection	การฉาย
projection function	ฟังก์ชันการฉาย
recognition using 3D shape and texture	กระบวนการรู้จำใบหน้าโดยใช้ข้อมูลรูปร่างและพื้นผิวภาพ 3 มิติ
recognition using 3D shape assisted 2D	กระบวนการรู้จำใบหน้าแบบ 2 มิติ ร่วมกับข้อมูลภาพ 3 มิติ
recognition using 3D shape only	กระบวนการรู้จำใบหน้าโดยใช้ข้อมูลรูปร่างของภาพ 3 มิติเท่านั้น
reference point	จุดอ้างอิง
rigid body	วัตถุเกร็ง
rotation	การหมุนของวัตถุ
shape	รูปร่าง
shape context	บริบทรูปร่าง
shape descriptor	ตัวบอกรูปร่าง
shape vector	เวกเตอร์รูปร่าง

shearing	การเฉือน
silhouette image	ภาพโครงร่างของวัตถุ
simple topology	ทอพอโลยีแบบพื้นฐาน
sophisticated topology	ทอพอโลยีแบบซับซ้อน
spatial transformation	การแปลงตำแหน่งของจุดภาพ
stereo acquisition	การสร้างภาพ 3 มิติ จากภาพ 2 มิติหลายภาพ
stress function	ฟังก์ชันความเครียด
structural light camera	กล้องโครงสร้างแสง
testing image	ภาพทดสอบ
texture	พื้นผิว
texture vector	เวกเตอร์พื้นผิว
topology	ทอพอโลยี
training image	ภาพทดลอง
transformation	การแปลง
translation	การเลื่อนขนาน
transpose operator	ตัวดำเนินการสลับเปลี่ยน
uncorrelation	ไม่มีความสัมพันธ์กัน
vectorization	การแปลงให้เป็นเวกเตอร์
vertex	จุดยอด
virtual world	โลกเสมือน
within-class scatter matrix	เมทริกซ์แสดงการกระจายตัวภายในกลุ่ม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1. ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันนี้ เทคโนโลยีการประมวลผลภาพดิจิทัล (digital image processing) ได้เข้ามามีบทบาทในชีวิตประจำวันมากขึ้น หนึ่งในเทคโนโลยีที่มีความโดดเด่นและเป็นที่น่าสนใจอย่างมากได้แก่ การรู้จำใบหน้า (face recognition) ซึ่งเป็นการวิเคราะห์ภาพใบหน้าเพื่อแยกแยะและระบุตัวบุคคลในภาพ ทั้งนี้เนื่องจากเป็นเทคโนโลยีที่มีประโยชน์และใช้งานได้ในหลายประเภท อาทิเช่น งานด้านความปลอดภัยและมั่นคงของประเทศ การตั้งกล้องจับภาพคนเข้าเมืองเพื่อตรวจหาผู้ก่อการร้าย งานรักษาความปลอดภัยในอาคารและสถานที่สำคัญต่างๆ เป็นต้น และในปัจจุบันยังนำการรู้จำใบหน้ามาใช้ในงานในเชิงพาณิชย์มากขึ้น เช่น การนำมาใช้แทนระบบชุดตัวขและรหัสผ่าน เป็นต้น ซึ่งความน่าเชื่อถือได้ของระบบมีน้อย เพราะหากมีมีบุคคลอื่นล่วงรู้ชุดรหัสนั้นก็สามารใช้บริการต่างๆและสืบค้นข้อมูลส่วนตัวของบุคคลนั้นๆได้ ดังนั้นจึงเริ่มมีการนำข้อมูลทางชีวมิติ (biometrics) ซึ่งเป็นข้อมูลเฉพาะตัวบุคคลแต่ละคนเข้ามาใช้แทนระบบรหัสผ่านมากขึ้น ซึ่งนอกจากจะมีความแม่นยำและน่าเชื่อถือได้สูงกว่าแล้วยังมีความสะดวกในการใช้งานมากกว่าอีกด้วย [1]

อย่างไรก็ตาม ในการนำกระบวนการรู้จำใบหน้ามาประยุกต์ใช้งานจริงนั้นยังมีปัจจัยหลายอย่างที่ทำให้ความแม่นยำของกระบวนการรู้จำใบหน้าไม่ดีเท่าที่ควร เนื่องจากการนำระบบรู้จำใบหน้าไปใช้งานจริง ภาพที่เข้ามาอาจมีปัจจัยแปรผันหลายอย่าง เช่น การแปรผันทางตำแหน่ง (pose variation) การแปรผันของแสง (illumination variation) การแปรผันเนื่องจากการแสดงออกของใบหน้า (facial expression variation) หรือปัญหาที่เกิดจากการมีวัตถุมาบังพื้นที่บริเวณใบหน้า (facial occlusion) ดังนั้นกระบวนการรู้จำใบหน้าแบบที่ใช้อยู่ในปัจจุบันมักจะจำกัดเงื่อนไขของภาพที่จะนำไปวิเคราะห์ให้เป็นภาพที่ดีเท่านั้น เช่น ต้องเป็นภาพถ่ายหน้าตรงเท่านั้นจึงจะได้ผลลัพธ์ที่ถูกต้องแม่นยำ เป็นต้น ดังนั้นเพื่อเพิ่มประสิทธิภาพของกระบวนการรู้จำใบหน้าให้สามารถแยกแยะบุคคลในสภาพที่มีการแปรผันได้อย่างแม่นยำและน่าเชื่อถือมากขึ้น จึงมีการศึกษาวิจัยและนำเสนอขั้นตอนวิธี (algorithm) หลากหลายรูปแบบเพื่อเพิ่มความแม่นยำในการรู้จำใบหน้าภายใต้เงื่อนไขที่แตกต่างกันไป [2] การแยกแยะความแตกต่างของใบหน้าอาศัยหลักการประมวลผลตามขั้นตอนวิธีแบบต่างๆ ในช่วงเริ่มต้นการวิจัยพัฒนาขั้นตอนวิธีนั้นส่วนมากมุ่งเน้นไปที่การใช้ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม (appearance-based) ซึ่งวิเคราะห์ลักษณะทั่วไปที่ปรากฏในภาพจากการสร้างแบบจำลองทางสถิติของระดับความเข้ม (intensity level) ในแต่ละจุดภาพ (pixel) แต่ขั้นตอนวิธีนี้มีข้อจำกัดอยู่มากเมื่อต้องเผชิญปัจจัยแปรผัน โดยเฉพาะอย่างยิ่งการแปรผันทางตำแหน่งซึ่งทำให้ตำแหน่งองค์ประกอบต่างๆบนใบหน้าในภาพเปลี่ยนไป เนื่องจากทำให้ข้อมูลความเข้มในแต่ละจุดภาพมีค่าเปลี่ยนแปลงไปมากและมีผลทำให้ความแม่นยำในการแยกแยะภาพลดลง ต่อมาจึงได้มีการนำขั้นตอนวิธีอิงแบบจำลองภาพ (model-based) เข้ามาเพื่อแก้ปัญหาที่เกิดกับขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวมเนื่องจากขั้นตอนวิธีอิงแบบจำลองภาพอาศัยการหาความสัมพันธ์ของระยะห่างระหว่างตำแหน่งขององค์ประกอบต่างๆ บนใบหน้า เช่น ตา คิ้ว จมูก มุมปาก ริมฝีปาก และ โกรงหน้า ดังนั้นแม้ว่าภาพใบหน้ามีการเปลี่ยนมุมมองไปในมุมต่างๆ ข้อมูลความสัมพันธ์ระหว่างตำแหน่งองค์ประกอบบนใบหน้าจะไม่เปลี่ยนแปลงไป จึงทำให้การวิเคราะห์ข้อมูลสามารถรองรับผลกระทบจากการแปรผันมุมมองของภาพใบหน้าได้ดีขึ้น

การพัฒนาขั้นตอนวิธีอิงแบบจำลองภาพมีหลายรูปแบบ ทั้งการหาข้อมูลจากความสัมพันธ์ระหว่างตำแหน่งขององค์ประกอบสำคัญๆบนใบหน้าโดยตรง หรือการประยุกต์ทฤษฎีกราฟ (graph theory) มาใช้วิเคราะห์ลักษณะโครงสร้างขององค์ประกอบต่างๆบนใบหน้า แต่ขั้นตอนวิธีอิงแบบจำลองภาพยังมีข้อดีตรงที่ต้องอาศัยกระบวนการแบ่งส่วนภาพ (image segmentation) ที่ดีพอเพื่อหาตำแหน่งขององค์ประกอบต่างๆบนใบหน้าที่เหมาะสม จึงจะได้ผลการแยกแยะที่มีความถูกต้องสูง จนในช่วงไม่กี่ปีที่ผ่านมา ด้วยเทคโนโลยีการเก็บข้อมูลที่มีประสิทธิภาพมากขึ้น จึงเริ่มมีการนำภาพ 3 มิติเข้ามาใช้ในกระบวนการรู้จำใบหน้ากันอย่างแพร่หลาย[3] เนื่องจากภาพในปริภูมิ 3 มิติให้ข้อมูลของภาพได้มากกว่าข้อมูลภาพในปริภูมิ 2 มิติ รวมทั้งสามารถแสดงตำแหน่งขององค์ประกอบต่างๆบนใบหน้าได้ชัดเจนกว่าภาพในปริภูมิ 2 มิติ นอกจากข้อมูลภาพ 3 มิติจะใช้ในกระบวนการรู้จำใบหน้าที่ร่วมกับภาพ 3 มิติอื่น ๆ ด้วยกันเองได้แล้ว ยังสามารถนำมาใช้ร่วมกับข้อมูลภาพแบบ 2 มิติได้อีกด้วย กล่าวคือ ข้อมูลภาพ 3 มิติใช้ประโยชน์ได้กว้างขวางกว่าข้อมูลภาพ 2 มิติ นอกจากนี้ภาพ 3 มิติเพียงภาพเดียวยังสามารถแสดงรายละเอียดของใบหน้าในทุกมุมมองได้ อีกทั้งข้อมูลภาพยังไม่ได้รับผลจากการเปลี่ยนแปลงของแสง จึงสามารถแก้ปัญหาการแปรผันทางตำแหน่งและการแปรผันจากแสงได้ดี

แนวทางการวิจัยด้านการวิเคราะห์ภาพใบหน้า 3 มิตินั้นเน้นไปที่การวิเคราะห์รูปร่าง (shape) ของภาพใบหน้า ซึ่งมีลักษณะของข้อมูลเป็นจุดยอด (vertex) ในปริภูมิ 3 มิติ ในช่วงแรกนั้นได้มีการนำเสนอขั้นตอนวิธีนำข้อมูลความโค้ง (curvature) ของพื้นผิวใบหน้ามาวิเคราะห์ เนื่องจากมีลักษณะความโค้งนูนและเว้าตามโครงสร้างของอวัยวะส่วนต่างๆ เช่น จมูกมีความโค้งนูน หรือ เบ้าตาจะมีลักษณะโค้งเว้า โดยลักษณะเหล่านี้จะมีความแตกต่างกันไปในแต่ละบุคคล ส่วนแนวคิดอีกแบบจะเป็นการสร้างภาพความลึก(depth map) แล้วนำมาแสดงผลเป็นภาพ 2 มิติ จากนั้นใช้การวิเคราะห์เวกเตอร์ในการแยกแยะความแตกต่างของบุคคลแต่ละคน [3] แต่การวิเคราะห์ดังกล่าวจะได้รับผลกระทบของภาพที่เกิดการแสดงออกของใบหน้าได้ เนื่องจากวิธีการดังกล่าวตั้งอยู่บนเงื่อนไขว่าใบหน้าเป็นวัตถุแข็ง (rigid body) ที่รูปร่างของพื้นผิวไม่มีการแปลง (transformation) ซึ่งไม่เป็นความจริงกับใบหน้าที่มีการแสดงออกที่พื้นผิวของใบหน้าที่มีการแปลงเกิดขึ้นได้

ดังนั้น จึงเริ่มมีแนวทางการวิจัยเพื่อให้สามารถครอบคลุมถึงผลของการแปรผันซึ่งเกิดจากการแสดงออกของใบหน้า จนมีงานวิจัยซึ่งนำเสนอขั้นตอนวิธีที่สามารถรองรับผลกระทบที่เกิดจากการแสดงออกของบุคคล[4] โดยตั้งอยู่บนสมมติฐานที่ว่า การแสดงออกของใบหน้าที่มีลักษณะเหมือนการเกิดการแปลงวัตถุแบบไม่แข็งเกร็ง (non-rigid body transform) บนพื้นผิว 3 มิติ ซึ่งวัตถุ 3 มิตินั้นมีคุณสมบัติหลักอยู่ 2 ลักษณะ คือ คุณสมบัติภายนอก (extrinsic property) ซึ่งจะเปลี่ยนค่าไปเมื่อวัตถุเกิดการแปลงและคุณสมบัติภายใน (intrinsic property) ซึ่งจะมีค่าคงที่เสมอไม่ว่าวัตถุจะเกิดการเปลี่ยนแปลงแบบใดๆ ดังนั้นไม่ว่าภาพใบหน้าที่มีการแสดงออกอย่างไร คุณสมบัติภายในของภาพจะยังคงอยู่เหมือนกันทุกการแสดงผล ซึ่งในงานวิจัยขั้นนี้จะถือว่าระยะทางจีโอเดสิก (geodesic distance) หรือระยะทางบนพื้นผิวโค้งระหว่างจุดต่าง ๆ ของวัตถุในภาพเป็นคุณสมบัติภายในของวัตถุ เมื่อนำเมตริกซ์ระยะทางจีโอเดสิกระหว่างจุดทุกจุดบนพื้นผิวไปสร้างภาพแบบบัญญัติของใบหน้า (canonical image) จะได้ภาพใบหน้าที่มีคุณสมบัติไม่แปรเปลี่ยนตามการแสดงผลของบุคคลในภาพ

การดึงคุณสมบัติภายในของภาพในงานวิจัยกระทำได้โดยหาระยะทางจีโอเดสิกระหว่างจุดต่างๆบนพื้นผิวใบหน้า ด้วยขั้นตอนของกระบวนการฟาสต์ มาร์ชชิง (fast marching) [5] ซึ่งเป็นการคำนวณโดยใช้เซตระดับ (level set) ในทิศทางเดียว (monotonic) แต่การดึงคุณสมบัติภายในด้วยวิธีนี้ยังคงมีความซับซ้อนในการคำนวณสูงและเมื่อนำมาใช้กับกระบวนการรู้จำใบหน้าซึ่งมีข้อมูลภาพจำนวนมากอยู่แล้วทำให้ระบบล่าช้าได้ ในขณะที่เดียวกันมีงานวิจัยด้านการวิเคราะห์ภาพในระนาบ 2 มิติได้มีการนำเสนอบริบทรูปร่าง (shape context) [6]

เพื่อประยุกต์ใช้ในงานวิเคราะห์ภาพที่มีรูปร่างเฉพาะ เช่น การรู้จำตัวอักษรจากลายมือ (handwriting recognition) หรือ การแยกแยะตราสินค้า (brand recognition) ซึ่งผลการทดลองพบว่าบริบทรูปร่างมีคุณสมบัติทนทานต่อการแปลงภาพรูปร่างแบบ 2 มิติได้ดี และยังได้ผลลัพธ์ที่ดีกับภาพที่เกิดการแปลงสัมพรรค (affine transform) อีกด้วย จึงมีการนำเสนอบริบทรูปร่างมาใช้ในปริภูมิ 3 มิติเพื่อวิเคราะห์วัตถุในฉากหลัง (cluster scene) ซึ่งแสดงให้เห็นถึงคุณสมบัติของบริบทรูปร่างที่สามารถดึงคุณสมบัติภายในของภาพที่มีการแปลงได้

วิทยานิพนธ์ฉบับนี้ นำเสนอขั้นตอนวิธีการลดผลกระทบที่เกิดจากการแปรผันเนื่องจากการแสดงออกของใบหน้าในภาพใบหน้า 3 มิติ โดยมีแนวคิด คือ การแสดงออกของบุคคลในภาพถ่าย 3 มิติ ทำให้เกิดการแปลงรูปร่างของภาพใบหน้าในลักษณะเดียวกับการเกิดการแปลงวัตถุแบบไม่แข็งแรง จึงนำเสนอการใช้บริบทรูปร่างเป็นตัวดึงข้อมูลคุณสมบัติภายในของภาพ แทนวิธีการหาระยะทางจีโอเดสิกบนพื้นผิว 3 มิติ โดยพัฒนาการใช้ฮิสโตแกรมแบบลอการิทึมทรงกลม (logarithm spherical histogram) ให้มีความเหมาะสมกับการนำไปใช้กับภาพใบหน้า นอกจากนี้จะมีการวิเคราะห์ข้อดี ข้อด้อย และคุณลักษณะของใบหน้าที่สร้างจากการใช้บริบทรูปร่าง [6] แทนการหาระยะทางระหว่างจุดบนพื้นผิว รวมถึง การนำไปประยุกต์ใช้งานเพื่อเพิ่มประสิทธิภาพให้กับกระบวนการรู้จำใบหน้าแบบ 3 มิติ ที่มีการเปลี่ยนแปลงด้านการแสดงออกของภาพใบหน้า

1.2. วัตถุประสงค์

เพื่อออกแบบและปรับปรุงขั้นตอนวิธีการลดผลกระทบจากการแปรผันเนื่องจากการแสดงออกของใบหน้าของภาพใบหน้าในปริภูมิภาพแบบ 3 มิติ โดยการปรับปรุงขั้นตอนวิธีบริบทรูปร่างเพื่อใช้สำหรับดึงคุณลักษณะของภาพใบหน้า ซึ่งมีคุณสมบัติไม่แปรเปลี่ยน (invariant) ตามการแปลงรูปร่างของวัตถุที่เกิดขึ้นในภาพ รวมถึงการนำไปประยุกต์ใช้สำหรับกระบวนการรู้จำใบหน้า และศึกษาแนวทางเพิ่มประสิทธิภาพให้กับกระบวนการรู้จำใบหน้าด้วยบริบทรูปร่าง

1.3. ขอบเขตของวิทยานิพนธ์

1. เพื่อออกแบบขั้นตอนวิธีการลดผลกระทบจากการแปลงเรขาคณิตของภาพวัตถุ 3 มิติแบบการแปลงวัตถุแบบแข็งแรง และการแปลงวัตถุแบบไม่แข็งแรง
2. เพื่อพัฒนาและปรับปรุงขั้นตอนวิธีที่ออกแบบเพื่อหาคุณลักษณะที่ไม่แปรเปลี่ยนจากการแสดงออกของบุคคลของภาพใบหน้าในปริภูมิภาพแบบ 3 มิติ
3. เพื่อพัฒนากระบวนการรู้จำใบหน้า 3 มิติแบบไม่แปรเปลี่ยนตามการแสดงออก โดยใช้ภาพใบหน้า 3 มิติจากฐานข้อมูลภาพใบหน้า BU-3DFE
4. เพื่อเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีที่นำเสนอกับการสร้างภาพแบบบัญญัติจากระยะทางจีโอเดสิก
5. เพื่อเปรียบเทียบผลความแม่นยำของกระบวนการรู้จำใบหน้าแบบ 3 มิติ ของภาพใบหน้าที่มีการแสดงออกของบุคคลกับขั้นตอนวิธีอื่นๆที่มีการนำเสนอ

1.4. ขั้นตอนและวิธีการดำเนินการ

1. ศึกษากระบวนการวิเคราะห์ภาพใบหน้า 3 มิติที่ครอบคลุมผลกระทบจากการเปลี่ยนแปลงการแสดงผลของบุคคลโดยการสร้างภาพแทนหลักของใบหน้า
2. ศึกษาความสามารถและคุณสมบัติของวิธีการวิเคราะห์รูปร่างของวัตถุในปริภูมิ 2 มิติ และ 3 มิติที่มีการแปลงในรูปแบบต่างๆ โดยการใช้บริบทรูปร่าง
3. ศึกษาความสัมพันธ์ระหว่างการเกิดการแปลงเรขาคณิตของวัตถุในปริภูมิ 2 มิติและ 3 มิติ
4. ออกแบบและศึกษาคุณสมบัติของบริบทรูปร่างในการวิเคราะห์โครงสร้างของวัตถุในปริภูมิ 3 มิติ
5. ออกแบบกระบวนการรู้จำใบหน้าโดยการใช้บริบทรูปร่างในปริภูมิ 3 มิติ
6. ทดลองวิธีการที่ออกแบบ และเก็บผลการทดลอง
7. รวบรวมและสรุปผลการวิจัยทั้งหมด เพื่อเขียนวิทยานิพนธ์

1.5. ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ทราบถึงการเกิดการแปลงเรขาคณิตแบบต่างๆ ของข้อมูลภาพวัตถุ 3 มิติ
2. ได้ทราบถึงแนวทางการวิเคราะห์ข้อมูลภาพแบบ 3 มิติ
3. ได้ขั้นตอนวิธีการที่มีความสามารถในการลดผลของการแปลงเรขาคณิตที่เกิดขึ้นในภาพ
4. ได้ทราบถึงแนวทางการออกแบบกระบวนการรู้จำใบหน้าโดยใช้บริบทรูปร่าง
5. ได้ตีพิมพ์ในวารสารทางวิชาการ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

หลักการและทฤษฎีทั่วไป

ในปัจจุบัน มีการใช้งานคอมพิวเตอร์เพื่อเก็บข้อมูลประเภทต่างๆกันอย่างแพร่หลาย ทั้งการเก็บเอกสารฐานข้อมูล และสื่อประสม จึงจำเป็นต้องมีการพัฒนากระบวนการที่สามารถดึงข้อมูลเหล่านั้นออกมาใช้ประโยชน์ได้อย่างมีประสิทธิภาพ การรับรู้ภาพของคอมพิวเตอร์ (computer vision) เป็นงานวิจัยซึ่งทำการวิเคราะห์ข้อมูลประเภทสื่อประสม ซึ่งประกอบด้วย ภาพและวิดีโอ เพื่อนำข้อมูลที่นำไปใช้งานตามความต้องการ เทคโนโลยีการประมวลผลภาพดิจิทัล เป็นส่วนย่อยของงานด้านการรับรู้ภาพของคอมพิวเตอร์ซึ่งเน้นไปที่การวิเคราะห์ข้อมูลที่เป็นภาพ ซึ่งอาจเป็นภาพถ่ายทั่วไป ภาพถ่ายทางการแพทย์ ภาพจากดาวเทียม หรือภาพ 3 มิติ กระบวนการรู้จำใบหน้าเป็นหนึ่งในกระบวนการประมวลผลภาพที่ได้รับความสนใจอย่างมากเนื่องจากสามารถนำไปประยุกต์ใช้งานได้หลากหลาย อาทิเช่น การใช้งานด้านการรักษาความปลอดภัย การใช้เพื่อความมั่นคง หรือการใช้งานในเชิงพาณิชย์ เป็นต้น

กระบวนการรู้จำใบหน้า เป็นขั้นตอนวิธีประมวลผลภาพใบหน้าบุคคลเพื่อเป้าหมายในการแยกแยะและระบุตัวบุคคลในภาพ โดยกระบวนการรู้จำใบหน้าแบบต่างๆถูกพัฒนาขึ้นมาอย่างต่อเนื่อง เพื่อแก้ไขปัญหาจากปัจจัยแปรผันที่ทำให้ความแม่นยำของกระบวนการรู้จำใบหน้าลดลง ได้แก่ การแปรผันของตำแหน่ง การแปรผันจากแสง การแปรผันเนื่องจากการแสดงออกของใบหน้า และการมีวัตถุมาบังพื้นที่บริเวณใบหน้า แนวทางวิจัยกระบวนการรู้จำใบหน้าเริ่มตั้งแต่กระบวนการรู้จำใบหน้าแบบ 2 มิติ โดยการวิเคราะห์ภาพโดยใช้ค่าระดับความเข้มของจุดภาพมาเป็นคุณลักษณะในการวิเคราะห์ข้อมูล ด้วยกระบวนการวิเคราะห์เวกเตอร์ จากนั้นได้พัฒนาจนเกิดกระบวนการรู้จำใบหน้าด้วยการวิเคราะห์ตำแหน่งและรูปร่างขององค์ประกอบบนใบหน้า เช่น ตา จมูก และโครงหน้า เป็นต้น ด้วยทฤษฎีกราฟและแบบจำลองใบหน้า จนปัจจุบันกระบวนการรู้จำใบหน้าแบบ 3 มิติเข้ามาเป็นที่นิยมและมีการวิจัยกันอย่างแพร่หลาย เนื่องจากมีการพัฒนาด้านการประมวลผลและอุปกรณ์สำหรับเทคโนโลยีภาพ 3 มิติ โดยการนำข้อมูลรูปร่างและพื้นผิวของใบหน้ามาเป็นคุณลักษณะในการวิเคราะห์ภาพ

2.1. กระบวนการรู้จำใบหน้าแบบ 2 มิติ (2D face recognition)

การวิเคราะห์ภาพใบหน้าเพื่อแยกแยะบุคคลหรือกระบวนการรู้จำใบหน้าเป็นหนึ่งในกระบวนการประมวลผลภาพดิจิทัลที่ได้รับความสนใจอย่างมากในปัจจุบัน เนื่องจากสามารถนำมาประยุกต์ใช้งานได้หลากหลาย เช่น การใช้งานในด้านการรักษาความปลอดภัยภายในอาคารหรือสถานที่ การใช้ติดตามข้อมูลของบุคคลจากภาพถ่ายหรือกล้องวิดีโอ และการใช้งานแทนระบบบัตรหรือรหัสผ่าน เป็นต้น

การทำงานของกระบวนการรู้จำใบหน้าแบ่งออกได้เป็น 2 ขั้นตอน ได้แก่

1. การพิสูจน์ใบหน้า (face verification)

การพิสูจน์ใบหน้า คือ กระบวนการตรวจสอบภาพใบหน้าของบุคคลบุคคลที่อยู่นอกเหนือฐานข้อมูลว่าเป็นบุคคลเดียวกับบุคคลที่ปรากฏในฐานหรือไม่ โดยการเปรียบเทียบระหว่างภาพใบหน้าเพียง 2 ภาพคือ ภาพใบหน้าของบุคคลที่ปรากฏในฐานข้อมูลและภาพใบหน้าของบุคคลที่อยู่นอกเหนือฐานข้อมูล และพิสูจน์ว่าใช่ภาพคนเดียวกันหรือไม่ โดยขั้นตอนนี้จะไม่มีการชี้ชัดเจาะจงว่า ภาพนั้นๆเป็นบุคคลใด ดังนั้นผลลัพธ์ที่ได้จากกระบวนการส่วนนี้จะมีเพียง 2 แบบ คือ ภาพใบหน้าของบุคคลที่อยู่นอกเหนือฐานข้อมูลเป็นภาพเดียวกันกับคือภาพใบหน้าของบุคคลที่ปรากฏในฐานข้อมูลหรือไม่ใช่บุคคลเดียวกัน

2. การระบุใบหน้า (face identification)

การระบุใบหน้า คือ การระบุเจาะจงว่าภาพภาพนั้นเป็นภาพของใครในฐานข้อมูล โดยจะให้ผลลัพธ์เป็นชื่อหรือข้อมูลต่างๆของภาพบุคคลที่อยู่นอกเหนือฐานข้อมูล ซึ่งขั้นตอนนี้แตกต่างจากการพิสูจน์ใบหน้าที่บอกเพียงภาพใบหน้าที่ทำการทดสอบเป็นภาพเดียวกันหรือไม่เท่านั้น โดยกระบวนการนี้จะต้องอาศัยผลของการพิสูจน์ใบหน้าที่ระหว่างภาพทดสอบกับทุกๆภาพในฐานข้อมูลเพื่อค้นหาภาพทดสอบมีความใกล้เคียงกับภาพใบหน้าในฐานข้อมูลภาพใดมากที่สุด แล้วทำการดึงชื่อ ข้อมูล หรือประวัติของบุคคลในฐานข้อมูลออกมา

งานวิจัยที่ผ่านมาแนะนำเสนอกระบวนการรู้จำใบหน้าโดยใช้ข้อมูลภาพถ่ายของบุคคลซึ่งเกิดจากการฉายภาพวัตถุหรือใบหน้าของบุคคลในปริภูมิ 3 มิติไปยังปริภูมิภาพ 2 มิติ ซึ่งมีการพัฒนาขั้นตอนวิธีเพื่อเพิ่มประสิทธิภาพและความแม่นยำในกระบวนการรู้จำใบหน้า แต่ถึงกระนั้นก็ยังมียังมีปัจจัยที่ทำให้ความแม่นยำของกระบวนการรู้จำใบหน้านั้นไม่ดีเท่าที่ควรเมื่อนำไปใช้งานจริง ทั้งนี้เนื่องมาจาก ในการทำงานจริงนั้น ภาพที่รับมาเป็นสัญญาณเข้ามามีปัจจัยต่างๆมากมาย ที่ส่งผลกระทบต่อความแม่นยำของกระบวนการรู้จำใบหน้า ซึ่งปัจจัยสำคัญที่มีผลกระทบต่อความแม่นยำของกระบวนการรู้จำใบหน้า มีดังต่อไปนี้

1. การแปรผันทางตำแหน่ง (pose variation)

การแปรผันทางตำแหน่ง เกิดจากทั้งการหมุนของใบหน้าหรือการเปลี่ยนแปลงมุมกล้องขณะถ่ายภาพก็ได้ มุมที่เกิดขึ้นอาจมีทั้งการหันในมุมราบซ้าย-ขวา หรืออาจเป็นลักษณะของมุมก้ม-เงย ทำให้ภาพที่รับได้มีความแตกต่างไปจากเดิม ซึ่งมีผลต่อกระบวนการรู้จำใบหน้าโดยเฉพาะอย่างยิ่งในกระบวนการรู้จำใบหน้าแบบอิงลักษณะที่ปรากฏโดยรวมซึ่งใช้ค่าระดับความเข้มของจุดภาพเป็นคุณลักษณะในการวิเคราะห์ภาพ ซึ่งวิธีแบบนี้มักไม่มีการหาตำแหน่งขององค์ประกอบต่างๆของใบหน้าในภาพไว้ก่อน หรือการคำนวณการเอียงของภาพ ผลที่ได้คือ สัญญาณของภาพบุคคลเดียวกันที่มีตำแหน่งในภาพแตกต่างกันอาจถูกแยกแยะว่าเป็นบุคคลคนละคนกันก็ได้

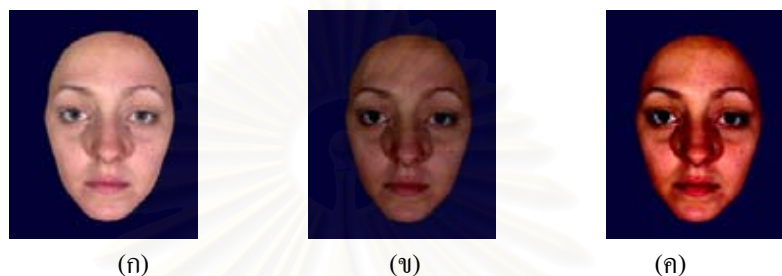


รูปที่ 2.1 ภาพบุคคลเดียวกันที่มีการแปรผันทางตำแหน่ง

(ก) ภาพมุมมองหน้าตรง (ข) ภาพมุมมองเอียงซ้าย (ค) ภาพมุมมองเอียงขวา

2. การแปรผันของแสง (illumination variation)

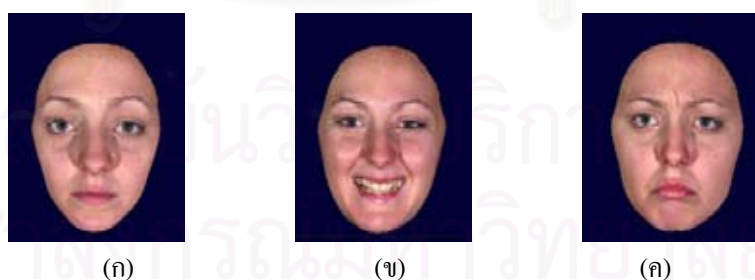
การแปรผันของแสง เกิดจากปริมาณแสงในภาพใบหน้าในภาพมีการเปลี่ยนแปลงซึ่งอาจมีสาเหตุมาจากแสงจากสภาพแวดล้อม (ambient light) เช่น ภาพที่ถ่ายกลางแจ้งจะมีปริมาณแสงมากกว่าภาพที่ถ่ายในที่ร่ม และแสงที่เกิดจากการฉายแสงลงบนใบหน้าโดยตรง (direct light) เช่น ภาพที่ถ่ายโดยใช้ไฟแฟลชจากกล้องถ่ายภาพ จะมีปริมาณแสงในภาพมากกว่าภาพที่ไม่ใช้ไฟแฟลช ซึ่งการแปรผันของแสงที่เกิดขึ้นนี้ จะทำให้ค่าระดับความเข้มแสง (illumination) ในภาพนั้นเปลี่ยนแปลงไปในแต่ละภาพ ภาพบุคคลคนเดียวกันแต่แสงในภาพแตกต่างกัน ดังแสดงในรูปที่ 2.2 อาจมีผลให้กระบวนการรู้จำใบหน้าแยกแยะบุคคลนั้นผิดพลาดได้



รูปที่ 2.2 ภาพบุคคลเดียวกันที่มีการแปรผันของแสง
(ก) ภาพที่มีแสงปกติ (ข) ภาพที่มีแสงมืดเกินไป (ค) ภาพที่มีแสงแปลกไปจากปกติ

3. การแปรผันเนื่องจากการแสดงออกของใบหน้า (facial expression variation)

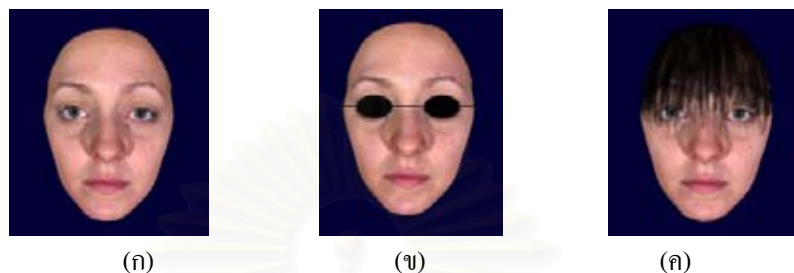
การแปรผันเนื่องจากการแสดงออกของใบหน้าบุคคล เกิดจากการแสดงอารมณ์ของบุคคลในภาพ เช่น ภาพที่บุคคลมีการแสดงอารมณ์ยิ้ม หรือ โกรธ ซึ่งการแสดงอารมณ์ที่เกิดขึ้นอาจทำให้ลักษณะของภาพของบุคคลคนเดียวกันมีความแตกต่างกันได้ ตัวอย่างเช่น ภาพคนที่มีลักษณะโครงหน้าแคบ เมื่อยิ้ม จะทำให้โครงหน้ากว้างขึ้น และรูปร่างของปากเปลี่ยนแปลงไป การแสดงออกทางอารมณ์เหล่านี้ดังแสดงในรูปที่ 2.3 มีผลต่อลักษณะขององค์ประกอบหลักของส่วนต่างๆ ในภาพใบหน้า ซึ่งอาจทำให้กระบวนการรู้จำใบหน้ามีการแยกแยะผิดพลาดได้



รูปที่ 2.3 ภาพบุคคลเดียวกันที่มีการแปรผันเนื่องจากการแสดงออกของใบหน้า
(ก) ภาพอารมณ์ปกติไม่มีการแสดงออกทางอารมณ์ (ข) ภาพใบหน้าขณะยิ้ม (ค) ภาพใบหน้าขณะโกรธ

4. การมีวัตถุมาบังพื้นที่บริเวณใบหน้า (facial occlusion)

การมีวัตถุมาบังพื้นที่บริเวณใบหน้า เป็นปัญหาที่เกิดจากการมีสิ่งใดๆมาปิดบังพื้นที่บางส่วนของใบหน้า เช่น การใส่แว่นสายตา แว่นกันแดด หมวก หรือการที่เส้นผม หนวด ลงมาปิดบางส่วนของใบหน้า ทำให้มองเห็นองค์ประกอบบนใบหน้าไม่ครบทุกส่วน ซึ่งมีผลให้การทำงานของกระบวนการรู้จำใบหน้าผิดพลาดไปได้



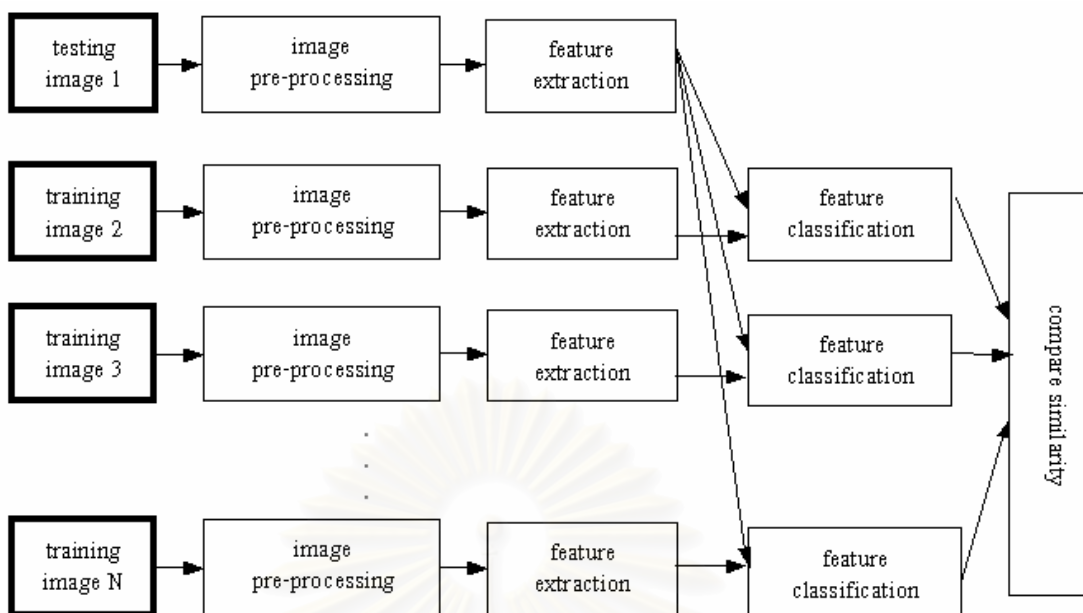
รูปที่ 2.4 ภาพบุคคลเดียวกันที่มีวัตถุมาบังพื้นที่บริเวณใบหน้า

(ก) ภาพที่ไม่มีวัตถุบังใบหน้า (ข) ภาพที่มีแว่นดำมาบังใบหน้า (ค) ภาพที่มีที่มียางผมมาบังใบหน้า

กระบวนการรู้จำใบหน้าแม้จะมีรูปแบบการคำนวณที่แตกต่างกัน แต่ก็มีหลักการทำงานเบื้องต้นที่คล้ายคลึงกัน กล่าวคือ เริ่มจากการรับข้อมูลภาพจากอุปกรณ์รับภาพ โดยอาจมีการใช้กระบวนการเตรียมภาพเบื้องต้น (image pre-processing) เช่น การปรับปรุงภาพ (image enhancement) การกำจัดพื้นหลังของภาพ (background subtraction) เป็นต้น เพื่อให้ภาพที่ได้มีคุณภาพดีขึ้นและเหมาะสมสำหรับนำไปประมวลผลข้อมูลต่อไป

ขั้นตอนถัดมา คือการดึงข้อมูลที่เป็นองค์ประกอบสำคัญของภาพหรือกระบวนการสกัดคุณลักษณะ (feature extraction) เนื่องจากสัญญาณภาพโดยปกติจะมีสัญญาณที่เป็นข้อมูลที่มีความสำคัญน้อยและสัญญาณรบกวนจำนวนหนึ่ง ดังนั้นเพื่อลดความผิดพลาดในกระบวนการแยกแยะภาพจึงต้องมีการดึงสัญญาณภาพเฉพาะส่วนที่มีความสำคัญหรือคุณลักษณะสำคัญของภาพออกมาใช้ในการวิเคราะห์เท่านั้น ซึ่งขั้นตอนวิธีการสกัดคุณลักษณะและรูปแบบของคุณลักษณะที่ได้จะมีลักษณะแตกต่างกันไปซึ่งจะกล่าวในหัวข้อต่อไป โดยอาจแบ่งย่อยได้เป็นวิธีการหาคุณลักษณะที่เหมาะสม ทำให้ข้อมูลสามารถแยกแยะได้ชัดเจนและการเลือกคุณลักษณะที่ดี เพื่อให้ได้จำนวนคุณลักษณะที่พอเพียงแก่การนำไปใช้งาน

หลังจากได้สกัดคุณลักษณะของภาพใบหน้ามาได้แล้ว จึงทำการหาความแตกต่างของคุณลักษณะ (feature classification) เหล่านี้ โดยเปรียบเทียบระหว่างคุณลักษณะของภาพที่ต้องการทดสอบ (testing image) กับคุณลักษณะของภาพทดลอง (training image) เพื่อเปรียบเทียบความคล้ายคลึงและค้นหาภาพทดลองที่มีความเหมือนกับภาพที่ต้องการทดสอบมากที่สุด ซึ่งแสดงโครงสร้างการทำงานได้ดังรูปที่ 2.5

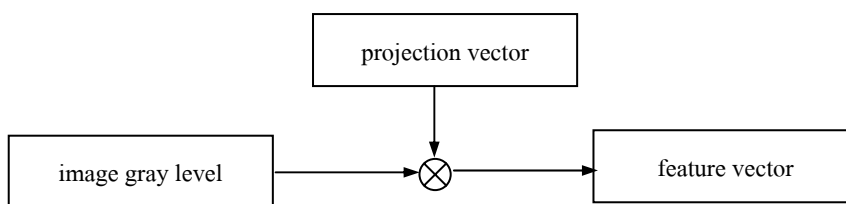


รูปที่ 2.5 โครงสร้างการทำงานของกระบวนการรู้จำใบหน้าเบื้องต้น

ในอดีต กระบวนการรู้จำใบหน้าเป็นการวิเคราะห์ภาพถ่ายของบุคคลซึ่งเป็นข้อมูลภาพในแบบ 2 มิติ ซึ่งการแยกแยะความแตกต่างของใบหน้าบุคคลจะอาศัยหลักการประมวลผลตามขั้นตอนวิธีแบบต่างๆ โดยแบ่งออกเป็น 2 ขั้นตอนวิธีหลัก คือ แบบขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม ซึ่งจะวิเคราะห์ลักษณะทั่วไปที่ปรากฏในภาพ ส่วนขั้นตอนวิธีการแยกแยะความแตกต่างของใบหน้าบุคคลอีกวิธีหนึ่ง คือ ขั้นตอนวิธีอิงแบบจำลองภาพ ซึ่งเป็นกระบวนการนำภาพมาหาตำแหน่งหรือความสัมพันธ์กันของอวัยวะสำคัญๆ บนใบหน้า เช่น ตา ริมฝีปาก เป็นต้น เพื่อนำมาวิเคราะห์เปรียบเทียบกับแบบจำลองของใบหน้าตัวอย่าง

2.1.1 ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม (appearance-based algorithm)

ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวมเป็นรูปแบบขั้นตอนวิธีที่มองรูปภาพทั้งภาพเป็นเมทริกซ์ โดยค่าของแต่ละพิกัดในเมทริกซ์แทนค่าระดับความเข้มในแต่ละจุดภาพ ทำการพิจารณาลักษณะการกระจายตัวของค่าระดับสีเทาในแต่ละจุดภาพ และทำการฉาย (projection) เมทริกซ์ภาพไปยังปริภูมิภาพ (image space) ซึ่งขั้นตอนนี้จะทำโดยการแปลงผ่านเวกเตอร์การฉาย (projection vector) เพื่อหาเวกเตอร์คุณลักษณะ (feature vector) ของภาพ ดังแสดงด้วยแผนภาพในรูปที่ 2.6 การแปลงภาพไปยังปริภูมิภาพนั้นมีจุดประสงค์เพื่อลดมิติของเมทริกซ์ภาพ โดยจะแสดงผลเฉพาะข้อมูลของภาพในส่วนที่มีความสำคัญเท่านั้น ส่วนองค์ประกอบของภาพที่มีความสำคัญน้อยซึ่งจัดเป็นสัญญาณที่ไม่ต้องการใช้ในการวิเคราะห์จะถูกกำจัดทิ้งไป ในขั้นตอนการแยกแยะความแตกต่างของบุคคลนั้น จะทำการคำนวณความแตกต่างจากเวกเตอร์คุณลักษณะที่ถูกสร้างมาจากภาพ ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม ขั้นตอนวิธีนี้มีการวิเคราะห์ข้อมูลใน 2 ลักษณะ คือ การวิเคราะห์ปริภูมิย่อยเชิงเส้น (linear subspace analysis) และการวิเคราะห์ปริภูมิย่อยแบบไม่เชิงเส้น (non-linear manifold analysis)



รูปที่ 2.6 แผนภาพการหาเวกเตอร์คุณลักษณะโดยขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม

2.1.1.1 การวิเคราะห์ปริภูมิย่อยเชิงเส้น (linear subspace analysis)

การวิเคราะห์ปริภูมิย่อยเชิงเส้นเป็นการวิเคราะห์ภาพใบหน้า โดยทำการฉายเมทริกซ์ภาพที่ต้องการไปยังปริภูมิภาพโดยอาศัยการแปลงผ่านเวกเตอร์การฉาย ตามสมการเชิงเส้น ดังแสดงในสมการที่ (2.1)

$$\mathbf{U} = \mathbf{A}_i^T \mathbf{V} \quad (2.1)$$

โดยที่	\mathbf{A}_i	คือ เมทริกซ์ภาพที่ i ขนาด $n \times m$
	\mathbf{U}	คือ เวกเตอร์คุณลักษณะขนาด $m \times d$
	\mathbf{V}	คือ เวกเตอร์การฉายขนาด $n \times d$
	T	คือ ตัวดำเนินการสลับเปลี่ยน (transpose operator)

การวิเคราะห์ปริภูมิย่อยเชิงเส้นมีข้อเสีย คือ เมื่อภาพมีการแปรผันทางตำแหน่งหรือการแปรผันของแสง ผลความแม่นยำของระบบจะลดลง เนื่องจากการวิเคราะห์ปริภูมิย่อยเชิงเส้นทำการคำนวณโดยใช้ค่าระดับสีเทาในแต่ละจุดภาพในเมทริกซ์ภาพเป็นข้อมูลสำคัญ ดังนั้นเมื่อเกิดการแปรผันที่ทำให้ค่าระดับสีเทาในแต่ละจุดภาพเปลี่ยนแปลงไปมาก จึงส่งผลให้การแยกแยะภาพบุคคลมีความคลาดเคลื่อนได้ ตัวอย่างขั้นตอนวิธีในการวิเคราะห์ปริภูมิย่อยเชิงเส้น ได้แก่ การวิเคราะห์องค์ประกอบหลัก การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น (Linear Discriminant Analysis, LDA) การวิเคราะห์องค์ประกอบอิสระ (Independent Component Analysis, ICA) เป็นต้น โดยแต่ละขั้นตอนวิธีจะมีวิธีการหาเวกเตอร์การฉาย (\mathbf{V}) ที่แตกต่างกันไป ดังนี้

1. การวิเคราะห์องค์ประกอบหลัก (Principle Component Analysis, PCA)

การวิเคราะห์องค์ประกอบหลักเป็นขั้นตอนวิธีซึ่งทำการหาเวกเตอร์การฉาย เพื่อให้เมทริกซ์ภาพแต่ละภาพมีเวกเตอร์คุณลักษณะในปริภูมิภาพที่ไม่มีความสัมพันธ์กัน (uncorrelation) เพื่อให้เห็นความแตกต่างของข้อมูลสำคัญของแต่ละภาพอย่างชัดเจน โดยใช้หลักการแปลงของคาร์เฮอเนน โลฟ (Karhunen - Loeve transformation) โดยเวกเตอร์คุณลักษณะสามารถคำนวณได้จากการหาเวกเตอร์ลักษณะเฉพาะ (eigenvector) ของเมทริกซ์ความแปรปรวนร่วม (covariance matrix) ของกลุ่มภาพทดลองทั้งหมด ดังแสดงในสมการที่ (2.2)

$$\mathbf{G} = \frac{1}{M} \sum_{j=1}^M (\mathbf{A}_j - \mu_{A_j})^T (\mathbf{A}_j - \mu_{A_j}) \quad (2.2)$$

โดยที่ \mathbf{G} คือ เมทริกซ์ความแปรปรวนร่วม
 M คือ จำนวนภาพทั้งหมดในกลุ่มทดลอง
 \mathbf{A}_j คือ ข้อมูลภาพที่ j
 μ คือ ค่าเฉลี่ยของข้อมูลภาพในกลุ่มทดลองทั้งหมด

โดยค่าเฉลี่ยของกลุ่มข้อมูลภาพ μ สามารถคำนวณได้ตามสมการที่ (2.3)

$$\mu = \frac{1}{M} \sum_{j=1}^M \mathbf{A}_j \quad (2.3)$$

การลดขนาดของเมทริกซ์ความแปรปรวนร่วมทำได้โดยการแยกเวกเตอร์ลักษณะเฉพาะ (eigen decomposition) ของเมทริกซ์ความแปรปรวนร่วมดังแสดงในสมการที่ (2.4) และเลือกใช้เวกเตอร์ลักษณะเฉพาะ โดยเลือกจากค่าลักษณะเฉพาะ (eigenvalue) ที่มีค่าสูงสุดจำนวน d ค่า โดยค่า d มีค่าน้อยกว่าขนาดมิติของข้อมูลภาพ

$$\lambda_i \mathbf{v}_i = \mathbf{C} \mathbf{v}_i \quad (2.4)$$

โดยที่ \mathbf{v}_i คือ เวกเตอร์ลักษณะเฉพาะค่าที่ i เมื่อ $i = 1, \dots, d$
 λ_i คือ ค่าลักษณะเฉพาะค่าที่ i

ในขั้นตอนการแยกเวกเตอร์ลักษณะเฉพาะ นอกจากจะช่วยลดมิติของข้อมูลภาพที่ต้องคำนวณลงได้แล้ว เวกเตอร์การฉายที่หาได้ จะมีคุณสมบัติในการดึงข้อมูลของภาพส่วนที่มีความแปรปรวนของกลุ่มภาพสูงสุดออกมา ซึ่งถือเป็นส่วนข้อมูลที่มีความสำคัญของภาพเหมาะสมแก่การนำมาใช้วิเคราะห์ต่อไป โดยข้อมูลภาพที่มีความสำคัญน้อยจะถูกจัดว่าเป็นสัญญาณที่ไม่ต้องการในการประมวลผลจึงถูกตัดทิ้งไป เมื่อทำการแปลงตามสมการที่ (2.1) จะได้เวกเตอร์คุณลักษณะซึ่งมีเฉพาะข้อมูลส่วนที่มีความสำคัญของภาพแต่ละภาพไปทำการวิเคราะห์ความแตกต่างของภาพต่อไป

การวิเคราะห์องค์ประกอบหลัก เป็นขั้นตอนวิธีที่มีการวิจัยกันอย่างกว้างขวางและมีการพัฒนาในหลายรูปแบบเพื่อเพิ่มศักยภาพในการทำงาน ซึ่งแนวทางการพัฒนาของการวิเคราะห์องค์ประกอบหลักที่เป็นที่รู้จักกันอย่างแพร่หลาย ได้แก่

1. การวิเคราะห์องค์ประกอบหลักแบบ 1 มิติ (1D PCA)

เริ่มจากปี ค.ศ. 1991 เมื่อ M. Turk และ A. Pentland [7] ได้นำเสนอกระบวนการรู้จำใบหน้าด้วยการวิเคราะห์องค์ประกอบหลัก โดยอิงหลักการวิเคราะห์เวกเตอร์ ก่อนเริ่มการคำนวณภาพใบหน้าที่มีลักษณะของเมทริกซ์จึงถูกนำมาเรียงต่อกันแถวต่อแถวหรือหลักต่อหลัก เพื่อทำการแปลงให้เป็นเวกเตอร์ (vectorization) ดังสมการที่ (2.5)

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{bmatrix} \rightarrow \mathbf{A}_{PCA} = [a_{11} \ a_{12} \ \cdots \ a_{1n} \ a_{21} \ \cdots \ a_{2n1} \ a_{31} \ \cdots \ a_{mn}] \quad (2.5)$$

ข้อเสียของการแปลงเมทริกซ์ภาพให้เป็นเวกเตอร์ก่อนที่จะทำการคำนวณ จะทำให้ขนาดของเมทริกซ์ความแปรปรวนร่วม (\mathbf{C}) ที่คำนวณได้ตามสมการที่ (2.2) มีขนาดใหญ่มากเมื่อเทียบกับขนาดของภาพ นั่นคือ ถ้าภาพที่ต้องการมีขนาด $n \times m$ มิติ เมื่อแปลงเป็นเวกเตอร์ตามสมการที่ (2.5) เวกเตอร์ภาพจะมีขนาด $1 \times nm$ มิติ และจะได้เมทริกซ์ความแปรปรวนร่วมขนาด $nm \times nm$ มิติ ซึ่งทำให้การคำนวณในขั้นตอนถัดไป ในสมการที่ (2.4) เป็นไปได้ยาก

2. การวิเคราะห์ห้วงค์ประกอบหลักแบบ 2 มิติ (2D PCA)

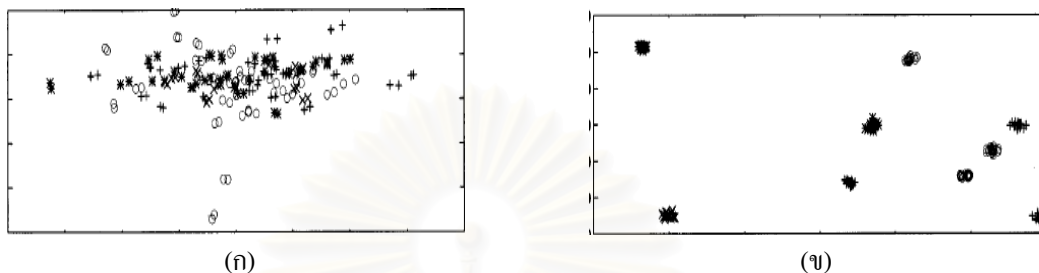
การคำนวณด้วยวิธีการวิเคราะห์ห้วงค์ประกอบหลักแบบ 1 มิติมักเกิดปัญหาเนื่องจากขนาดของเมทริกซ์ความแปรปรวนร่วมมีขนาดใหญ่ขึ้น J.Yang [8] ได้นำเสนอขั้นตอนวิธี การวิเคราะห์ห้วงค์ประกอบหลักแบบ 2 มิติ ในปี ค.ศ. 2004 ซึ่งอาศัยการคำนวณโดยใช้ภาพทั้งภาพในการคำนวณหาเมทริกซ์ความแปรปรวนร่วม ซึ่งวิธีนี้จะช่วยเพิ่มความแม่นยำได้มากยิ่งขึ้น เพราะข้อมูลความสัมพันธ์ด้านตำแหน่งของจุดภาพไม่ถูกทำลายจากการแปลงเมทริกซ์ให้เป็นเวกเตอร์

การคำนวณการวิเคราะห์ห้วงค์ประกอบหลักแบบ 1 มิติ ทำให้จุดภาพที่อยู่ใกล้กันในภาพถูกเรียงให้อยู่ในตำแหน่งที่ห่างกัน แต่การวิเคราะห์ห้วงค์ประกอบหลักแบบ 2 มิติ ซึ่งคำนวณภาพใบหน้าเป็นเมทริกซ์ในลักษณะ 2 มิติ จะรักษาข้อมูลของตำแหน่งค่าของจุดภาพเอาไว้ ดังนั้น ค่าความแปรปรวนร่วมที่คำนวณได้ นอกจากจะแสดงความสัมพันธ์ของค่าระดับสีเทาในแต่ละจุดภาพแล้ว ยังมีข้อมูลความสัมพันธ์ของตำแหน่งจุดภาพอีกด้วย ซึ่งช่วยให้การคำนวณมีความแม่นยำเพิ่มมากขึ้น ส่วนปัญหาเรื่องขนาดของเมทริกซ์ความแปรปรวนร่วมจะหมดไป กล่าวคือ คือ ถ้าเมทริกซ์ภาพที่ต้องการมีขนาด $n \times m$ มิติ เมทริกซ์ความแปรปรวนร่วมจะมีขนาดเพียง $m \times m$ มิติ และเมื่อทำการแยกเวกเตอร์ลักษณะเฉพาะ เวกเตอร์การฉายจะถูกลดขนาดลงเป็น $m \times d$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2. การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น (Linear Discriminant Analysis, LDA)

การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น [9] เป็นขั้นตอนวิธีที่ต้องการหาเวกเตอร์การฉายที่แสดงเวกเตอร์คุณลักษณะของเมทริกซ์ภาพในปริภูมิภาพที่แสดงความแตกต่างระหว่างกลุ่มสูง และแสดงความแตกต่างภายในกลุ่มเดียวกันน้อย ผลจากการแปลงภาพผ่านเวกเตอร์การฉาย คือ ภาพแต่ละกลุ่มจะมีการเกาะตัวภายในกลุ่มชัดเจน และภาพต่างกลุ่มกันจะกระจายตัวห่างกันมาก ซึ่งทำให้ผลการแยกแยะภาพมีความชัดเจนมากขึ้น ดังรูปที่ 2.7



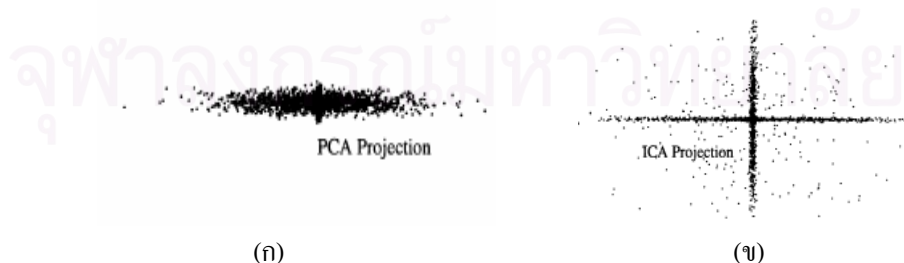
รูปที่ 2.7 ความแตกต่างของเวกเตอร์คุณลักษณะในปริภูมิภาพหลังทำการแปลงเมทริกซ์ภาพด้วย

(ก) การวิเคราะห์ห้อยค์ประกอบหลัก เทียบกับ (ข) การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น

จาก รูปที่ 2.7 จะเห็นได้ว่าเวกเตอร์คุณลักษณะที่หาได้จากการแปลงด้วยการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นจะเห็นความแตกต่างระหว่างกลุ่มข้อมูลชัดเจนกว่าการวิเคราะห์ห้อยค์ประกอบหลัก ซึ่งถ้าข้อมูลทดลองที่ใช้ได้เวกเตอร์คุณลักษณะในรูปแบบนี้การวิเคราะห์ด้วยการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นจะมีความแม่นยำในการแยกแยะบุคคลสูงกว่าการวิเคราะห์ห้อยค์ประกอบหลัก ทั้งนี้ ขึ้นอยู่กับลักษณะของภาพที่ใช้ในการทดลองด้วย

3. การวิเคราะห์ห้อยค์ประกอบอิสระ (Independent Component Analysis, ICA)

หลักการของการวิเคราะห์ห้อยค์ประกอบหลัก [10] คือการหาเวกเตอร์แปลงภาพที่ให้ผลลัพธ์เป็นสัญญาณที่ไม่มีความสัมพันธ์กัน โดยมีแนวคิดว่าคุณภาพทั้งหมดจะมีการกระจายตัวแบบเกาส์เซียน (gaussian distribution) แต่ถ้าข้อมูลภาพนั้นมีการกระจายตัวต่างออกไป ข้อมูลที่ไม่มีความสัมพันธ์กันจะอยู่ในช่วงของ lower-order เท่านั้น โดยข้อมูลในส่วน higher-order จะยังมีความสัมพันธ์กันอยู่ ดังนั้น การวิเคราะห์ห้อยค์ประกอบอิสระจะหาแกนที่ทำให้ข้อมูลในส่วน higher-order ไม่มีสัมพันธ์กันด้วยนั้น คือ จะต้องหาเวกเตอร์การฉายที่ให้เวกเตอร์คุณลักษณะที่อิสระต่อกัน



รูปที่ 2.8 ความแตกต่างของเวกเตอร์คุณลักษณะในปริภูมิภาพเมื่อทำการแปลงเมทริกซ์ภาพด้วย

(ก) การวิเคราะห์ห้อยค์ประกอบหลัก เทียบกับ (ข) การวิเคราะห์ห้อยค์ประกอบอิสระ

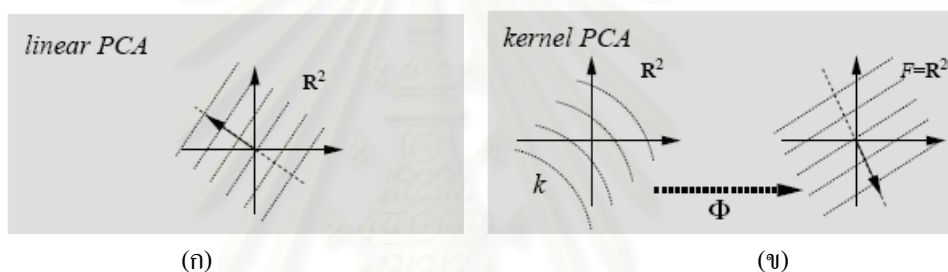
จากรูปที่ 2.8 จะเห็นได้ว่าเวกเตอร์คุณลักษณะที่หาได้จากการแปลงด้วยการวิเคราะห์องค์ประกอบอิสระ จะแสดงการกระจายตัวของข้อมูลในปริภูมิภาพที่อิสระต่อกัน คือ เวกเตอร์ลักษณะเฉพาะมีการกระจายตัวบนชัดเจนกว่าการหาเวกเตอร์การแปลงด้วยการวิเคราะห์องค์ประกอบหลัก

2.1.1.2 การวิเคราะห์ภาพไม่เชิงเส้น (non-linear manifold analysis)

การวิเคราะห์ภาพไม่เชิงเส้นเป็นขั้นตอนวิธีที่มีการวิเคราะห์ภาพคล้ายคลึงกับการวิเคราะห์ปริภูมิย่อยเชิงเส้น แต่มีความซับซ้อนกว่าเนื่องจากการแปลงภาพด้วยฟังก์ชันการฉาย (projection function) ซึ่งเป็นกระบวนการแบบไม่เชิงเส้น ตัวอย่างขั้นตอนวิธีแบบนี้ ได้แก่

1. การวิเคราะห์เคอร์เนลองค์ประกอบหลัก (Kernel Principle Component Analysis)

การวิเคราะห์เคอร์เนลองค์ประกอบหลัก [11] จะใช้การฉายภาพเป็นแบบไม่เชิงเส้นด้วย ทฤษฎีของเมอร์เร่ (Mercer's theorem) โดยการสร้างเคอร์เนลฟังก์ชัน (kernel function) เพื่อหาปริภูมิของเวกเตอร์คุณลักษณะคล้ายการใช้ขั้นตอนวิธีการวิเคราะห์องค์ประกอบหลัก โดยลักษณะที่แตกต่างกันอธิบายได้ดังรูปที่ 2.9



รูปที่ 2.9 การเปรียบเทียบความแตกต่างของการแปลงเวกเตอร์ด้วย

(ก) กระบวนการการวิเคราะห์องค์ประกอบหลักแบบปกติ เทียบกับ (ข) การวิเคราะห์เคอร์เนลองค์ประกอบหลัก

จากรูปที่ 2.9 จะเห็นได้ว่า การวิเคราะห์เคอร์เนลองค์ประกอบหลักคำนวณโดยที่การแปลงสัญญาณผ่านเคอร์เนลฟังก์ชัน (Φ) ซึ่งเป็นกระบวนการแบบไม่เชิงเส้น ดังนั้นข้อมูลภาพที่ใช้จะต้องถูกแปลงให้เป็นข้อมูลไม่เชิงเส้นก่อน ข้อได้เปรียบของการวิเคราะห์เคอร์เนลองค์ประกอบหลักคือจะสามารถวิเคราะห์ข้อมูลภาพที่ higher-dimension ได้ดีกว่าการวิเคราะห์องค์ประกอบหลัก แต่ปัญหาในการใช้การวิเคราะห์เคอร์เนลองค์ประกอบหลักคือปริภูมิของเวกเตอร์คุณลักษณะที่ได้ออกมาจะมีขนาดใหญ่กว่าขนาดของข้อมูลรูปภาพ

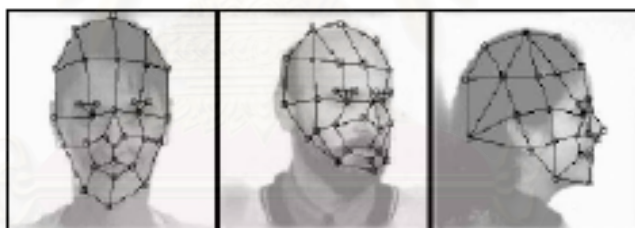
การใช้ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวมเป็นวิธีที่ได้รับผลกระทบจากการแปรผันของปัจจัยต่างๆ ทำให้ความถูกต้องในการคำนวณลดลง ซึ่งการจะทำให้ระบบมีความแม่นยำสูงก็ต่อเมื่อต้องใช้ภาพทดสอบหลายภาพให้ระบบมีการเรียนรู้ เพื่อลดผลกระทบที่มาจากปัจจัยภายนอก เช่น แสง มุมภาพ และการแสดงอารมณ์ เป็นต้น ดังนั้น ข้อด้อยอีกประการหนึ่งคือ ถ้ามีจำนวนภาพทดสอบน้อยเกินไป จะได้ผลการทดสอบไม่แม่นยำ

2.1.2 ขั้นตอนวิธีอิงแบบจำลองภาพ (model-based algorithm)

กระบวนการรู้จำใบหน้าด้วยขั้นตอนวิธีอิงแบบจำลองภาพ ถูกพัฒนาเพื่อเข้ามาแก้ปัญหาของ ขั้นตอนวิธีอิงลักษณะที่ปรากฏโดยรวม โดยเฉพาะอย่างยิ่งในเรื่องของการเปลี่ยนแปลงมุมของภาพใบหน้า ขั้นตอนวิธีกลุ่มนี้จะเป็นการหาความสัมพันธ์ของระยะห่างระหว่างตำแหน่งองค์ประกอบต่างบนใบหน้า เช่น ตำแหน่งหู ตา จมูก เป็นต้น และอาจรวมไปถึงการวิเคราะห์ลักษณะพื้นผิว (texture) ของใบหน้า จากนั้นนำข้อมูลไปเปรียบเทียบกับข้อมูลของภาพใบหน้าในฐานะข้อมูลเพื่อแยกแยะความแตกต่างของภาพใบหน้าแต่ละภาพ จะเห็นได้ว่าการใช้ขั้นตอนวิธีในกลุ่มนี้ได้ต้องมีประสิทธิภาพจะต้องมีการดึงข้อมูลขององค์ประกอบบนใบหน้า (facial feature segmentation) ออกมาได้แม่นยำ ซึ่งเป็นกระบวนการส่วนนี้อาจทำได้ด้วยมือ (manual) หรือใช้ขั้นตอนวิธีเพื่อหาแบบอัตโนมัติ (automatic) ตัวอย่างขั้นตอนวิธีอิงแบบจำลองภาพ ได้แก่

2.1.2.1 ขั้นตอนวิธีเปรียบเทียบกราฟแบบยืดหยุ่นโดยอิงโครงหน้า (feature – based elastic bunch graph matching, FBG)

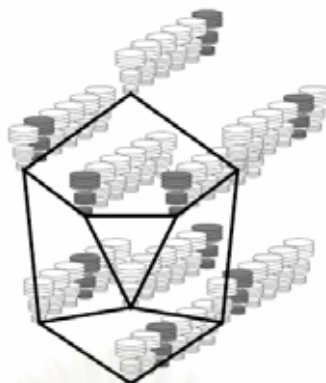
ขั้นตอนวิธีเปรียบเทียบกราฟแบบยืดหยุ่นโดยอิงโครงหน้า [12] จะแทนโครงหน้าด้วยกราฟที่ประกอบด้วยกิ่ง (branch) และ ปม (node) โดยแทนปมที่ตำแหน่งองค์ประกอบบนใบหน้า เช่น ดวงตา ปาก จมูก โครงหน้า เป็นต้น โดยเชื่อมโยงแต่ละปมด้วยกิ่งเพื่อแสดงโครงสร้างของพื้นผิวใบหน้า ดังแสดงในรูปที่ 2.10 จากนั้นทำการแยกแยะแล้วจะเปรียบเทียบแต่ละองค์ประกอบของใบหน้าของภาพทดสอบกับทุกๆภาพในกลุ่มทดลองเพื่อแยกแยะภาพบุคคลที่คล้ายคลึงกันออกมา



รูปที่ 2.10 การกำหนดปมและกิ่งตามขั้นตอนวิธีเปรียบเทียบกราฟแบบยืดหยุ่นโดยอิงโครงหน้า

วิธีการนี้จะไม่ได้รับผลกระทบจากการเปลี่ยนแปลงมุมภาพเนื่องจากปมจะมีจุดตำแหน่งที่แน่นอนบนองค์ประกอบบนใบหน้า และกิ่งจะเปลี่ยนแปลงตำแหน่งได้ตามตำแหน่งของปมซึ่งขึ้นกับการหมุนของภาพ ที่ตำแหน่งปมแต่ละปมของภาพจะถูกแทนด้วยค่าสัมประสิทธิ์ (coefficient) ของกาบอร์เคอร์เนล (gabor kernels) โดยกำหนดค่าเจ็ท (jet, J) แทนเซตของค่าสัมประสิทธิ์ที่ได้จากการแปลงกาบอร์เวฟเล็ท (gabor wavelet)

โดยกราฟต้นแบบของภาพกลุ่มทดลองจะถูกเก็บเป็นค่าเจ็ทที่หาได้จากปมที่กำหนดตำแหน่งแน่นอนบนใบหน้า โดยตำแหน่งปมแต่ละปมจะประกอบด้วยค่าเจ็ทหลากหลายรูปแบบ เช่น ปมที่ตำแหน่งดวงตา จะมีต้นแบบของตาหลายแบบเช่น ตาเปิด ตาปิด ตาสองชั้น ตาชั้นเดียว เป็นต้น ซึ่งสามารถอธิบายโครงสร้างของเจ็ทในกราฟออกมาได้ ดังรูปที่ 2.11



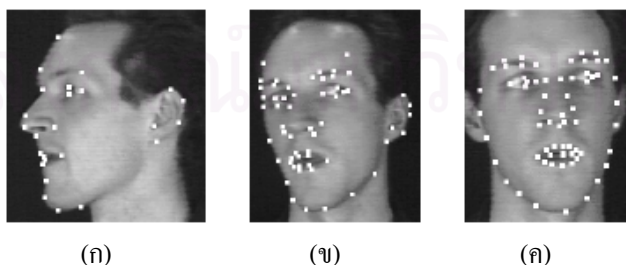
รูปที่ 2.11 การกำหนดเจ็ทหลายแบบในแต่ละตำแหน่งปม

ในการวัดความคล้ายคลึงกันเพื่อแยกแยะความแตกต่างของบุคคลในภาพจะเทียบหาความแตกต่างระหว่างภาพทดสอบกับต้นแบบกราฟในกลุ่มทดลอง โดยคำนวณความแตกต่างของค่าเจ็ทซึ่งแทนรายละเอียดของภาพที่ปม ซึ่งเทียบค่าเจ็ทที่ตำแหน่งปมจุดเดียวกัน

2.1.2.2 แบบจำลองลักษณะปรากฏแบบแอคทีฟ (active appearance model, AAM)

แบบจำลองลักษณะปรากฏแบบแอคทีฟ [13] ทำงานโดยวางตำแหน่งจุดกำหนด (landmark) ตามองค์ประกอบต่างๆบนใบหน้า และเก็บข้อมูลภาพแต่ละจุดกำหนดเพื่อหาเวกเตอร์รูปร่าง (shape vector) ของภาพใบหน้า จากนั้นจะหาความแตกต่างของภาพจากพิกัดของเวกเตอร์รูปร่างระหว่างภาพทดสอบกับภาพทดลอง นอกจากนี้ เรายังสามารถเก็บรายละเอียดของพื้นผิวของภาพใบหน้าได้ด้วยการเก็บเวกเตอร์พื้นผิว (texture vector) ขึ้นตอนวิธีวิธีแบบจำลองลักษณะปรากฏแบบแอคทีฟสามารถใช้งานได้กับภาพ 2 มิติ หรืออาจนำไปใช้กับภาพ 3 มิติได้ ในขั้นตอนการวิเคราะห์ข้อมูลของภาพอาจประยุกต์ใช้การวิเคราะห์องค์ประกอบหลักหรือการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นมาใช้ในการคำนวณได้

เมื่อมีการรับข้อมูลภาพทดสอบเข้ามา ขั้นตอนวิธีจะคำนวณหาเวกเตอร์รูปร่างและเวกเตอร์พื้นผิวของภาพทดสอบ ในรูปของความแตกต่างระหว่างภาพทดสอบกับค่าเฉลี่ยของกลุ่มข้อมูลภาพทดลอง โดยเราสามารถแสดงจุดกำหนดบนใบหน้าที่ใช้ในการสร้างเวกเตอร์รูปร่างและเวกเตอร์พื้นผิวได้ดังรูปที่ 2.12



รูปที่ 2.12 การกำหนดจุดบนองค์ประกอบสำคัญของใบหน้าด้วยขั้นตอนวิธีแบบจำลองลักษณะปรากฏแบบแอคทีฟเมื่อภาพมีการแปรผันทางตำแหน่ง
(ก) เมื่อใบหน้ามีการหันข้าง (ข) เมื่อใบหน้ามีการหันเอียง (ค) เมื่อเป็นภาพใบหน้าที่ตรง

เมื่อน้ำมีการหันหรือเอียงจุดกำหนดที่ได้จะเปลี่ยนไปและทำให้เวกเตอร์รูปร่างที่หาได้เปลี่ยนแปลงไป เราสามารถแสดงการแปรผันของภาพที่หันหน้าด้วยมุมใดๆ โดยการคำนวณหาค่าพารามิเตอร์ควบคุมค่าหนึ่ง ซึ่งค่าพารามิเตอร์ที่คำนวณได้นี้ทำให้เราสามารถแสดงข้อมูลภาพที่ต้องการได้ในทุกมุม นั่นคือระบบสามารถรองรับปัญหาความผิดพลาดที่เกิดจากการแปรผันทางตำแหน่งของใบหน้าได้ดี อีกทั้งยังนำแนวคิดนี้ไปใช้กับข้อมูลภาพแบบ 3 มิติได้อีกด้วย

2.1.2.3 แบบจำลอง 3 มิติแบบเปลี่ยนสภาพได้ (3D morphable model)

แบบจำลอง 3 มิติแบบเปลี่ยนสภาพ [14] ได้เป็นวิธีที่มีความซับซ้อนสูงกว่าขั้นตอนวิธีที่ได้กล่าวมา เนื่องจากทำงานโดยการเก็บข้อมูลเป็นข้อมูล 3 มิติ ซึ่งประกอบด้วยรูปร่างของใบหน้าและลักษณะพื้นผิวของรูป นำมาคำนวณเป็นพารามิเตอร์และทำการแยกแยะภาพจากค่าพารามิเตอร์นั้น เราสามารถกำหนดเวกเตอร์รูปร่าง \tilde{S} ด้วยพิกัดบนแกนคาร์ทีเซียน (X, Y, Z) และเวกเตอร์พื้นผิว \tilde{T} ด้วยค่าระดับสีเทาที่พื้นผิวด้วยค่าสีในระบบสีแดง-เขียว-น้ำเงิน (R, G, B) โดยถ้าภาพมีจำนวนจุดภาพจำนวน n จุด เวกเตอร์รูปร่างและเวกเตอร์พื้นผิว แสดงได้ดังสมการที่ (2.6) และสมการที่ (2.7) ตามลำดับ

$$\tilde{S} = (X_1, Y_1, Z_1, \dots, X_n, Y_n, Z_n)^T \quad (2.6)$$

$$\tilde{T} = (R_1, G_1, B_1, R_2, \dots, G_n, B_n)^T \quad (2.7)$$

โดยใช้ความแตกต่างของรูปร่างและพื้นผิวระหว่างภาพทดสอบและกับแบบจำลองใบหน้า 3 มิติ เพื่อทำการเปลี่ยนแปลงรูปร่างของแบบจำลองให้มีรูปร่างใกล้เคียงกับภาพทดสอบมากที่สุดทำการคำนวณหาพารามิเตอร์ที่ใช้ในการเปลี่ยนแปลงรูปร่างแบบจำลอง และใช้ข้อมูลพารามิเตอร์ที่คำนวณได้ในการแยกแยะภาพบุคคลแต่ละคน



รูปที่ 2.13 ข้อมูลภาพ 3 มิติในขั้นตอนวิธีแบบจำลอง 3 มิติแบบเปลี่ยนสภาพได้

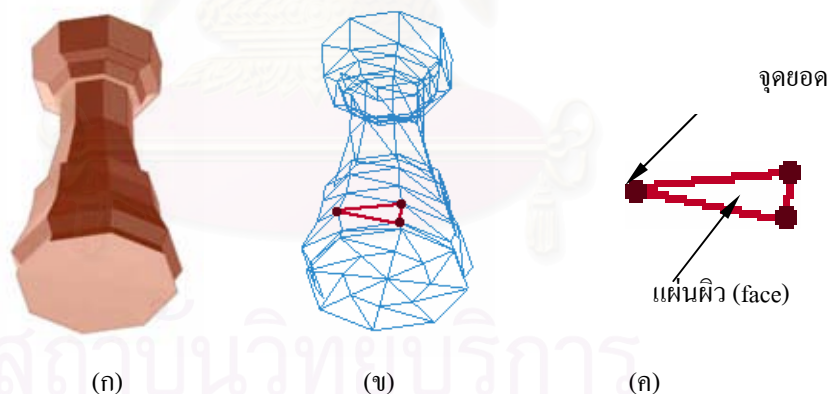
จะเห็นได้ว่า กระบวนการรู้จำใบหน้าเริ่มมีแนวโน้มในการวิเคราะห์ข้อมูลภาพเริ่มต้นที่ภาพแบบ 2 มิติ และพัฒนามาจนในปัจจุบันกระบวนการรู้จำใบหน้าแบบ 3 มิติได้เข้ามามีบทบาทในงานวิจัยจำนวนมาก โดยในหัวข้อต่อไป จะเป็นการแสดงถึงรายละเอียดของกระบวนการรู้จำใบหน้าแบบ 3 มิติ โดยประกอบด้วย ลักษณะของข้อมูลภาพแบบ 3 มิติ ขั้นตอนวิธีของกระบวนการรู้จำใบหน้าแบบ 3 มิติที่มีการวิจัยกันในปัจจุบัน รวมถึงปัญหาของกระบวนการรู้จำใบหน้าแบบ 3 มิติ

2.2. กระบวนการรู้จำใบหน้าแบบ 3 มิติ (3D face recognition)

เนื่องจากการพัฒนาทางด้านเทคโนโลยีของภาพ 3 มิติที่มีการพัฒนาไปมากรวมทั้งภาพแบบ 3 มิติยังมีลักษณะใกล้เคียงกับการรับรู้ข้อมูลภาพใบหน้าของมนุษย์มากกว่าข้อมูลภาพแบบ 2 มิติ ทำให้สามารถเห็นรายละเอียดของข้อมูลได้มาก ซึ่งช่วยให้กระบวนการรู้จำใบหน้าแบบ 3 มิติได้เข้ามามีบทบาทในงานวิจัยจำนวนมากและมีการพัฒนาไปอย่างรวดเร็ว แม้การนำไปใช้งานจริงนั้นยังไม่แพร่หลายนัก เนื่องจากยังต้องใช้เทคโนโลยีที่มีความก้าวหน้าสูงและยังเป็นงานวิจัยที่อยู่ในช่วงพัฒนา

2.2.1 ลักษณะของข้อมูลภาพแบบ 3 มิติ

โดยทั่วไปแล้ว วัตถุ 3 มิติประกอบด้วยรูปร่างและพื้นผิว ซึ่งข้อมูลของจุดยอดและแผ่นผิวจะถูกนำมาประกอบกันเพื่อสร้างรูปร่างของวัตถุ โดยนำภาพพื้นผิวมาแสดงผลบนรูปร่างของวัตถุอีกที ในที่นี้จะนำเสนอเพียงรายละเอียดและการวิเคราะห์เฉพาะของรูปร่างของวัตถุเท่านั้น ข้อมูลรูปร่างของภาพ 3 มิติ ถูกกำหนดด้วยมาตรฐาน VRML (Virtual Reality Modeling Language) ซึ่งเป็นมาตรฐานสำหรับการแสดงผล (display) การโต้ตอบ (interaction) และการเชื่อมโยงระหว่างเครือข่าย (internetworking) ในโลกเสมือน (virtual world) ข้อมูลภาพ 3 มิติถูกกำหนดให้มีรูปแบบไฟล์สกุล wrf ซึ่งแสดงผลภาพด้วยจุดยอด (vertex) และแผ่นผิว (face) โดยจุดยอดของวัตถุมีลักษณะเป็นพิกัดบนแกนคาร์ทีเซียน 3 มิติ ส่วนแผ่นผิวมีลักษณะเป็นตำแหน่งของจุดยอดที่สัมพันธ์กันซึ่งใช้สร้างแผ่นผิวนั้นๆ โดยแผ่นผิวทำหน้าที่กำหนดโครงสร้างของรูปร่างวัตถุ ดังแสดงในรูปที่ 2.14 (ง) และแสดงภาพขยายของแผ่นผิวแบบ 3 เหลี่ยม ซึ่งสร้างจากจุดยอด 3 จุดในรูปที่ 2.14 (ค)



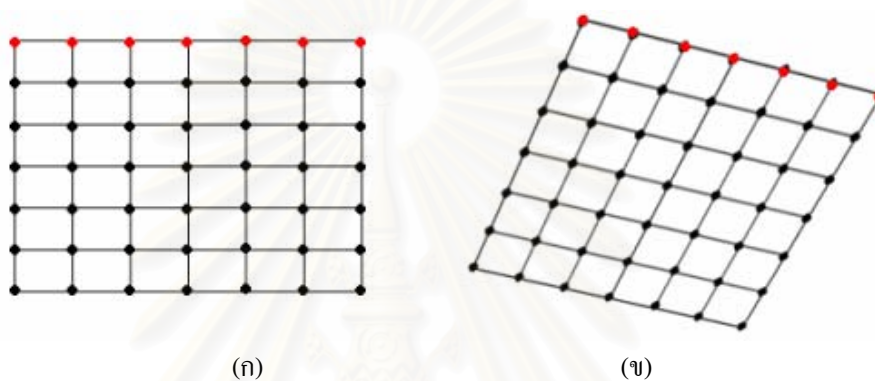
รูปที่ 2.14 โครงสร้างภาพวัตถุ 3 มิติ ลักษณะของจุดยอดและแผ่นผิวบนวัตถุ 3 มิติ

(ก) ภาพวัตถุ 3 มิติ (ง) ข้อมูลรูปร่างของวัตถุ 3 มิติ (ค) จุดยอดและแผ่นผิวแบบ 3 เหลี่ยม

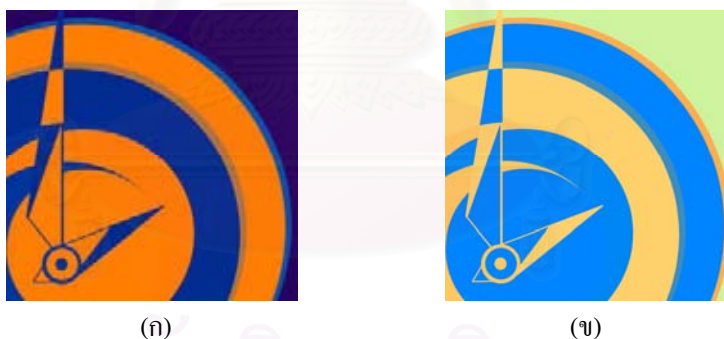
ภาพวัตถุ 3 มิติสามารถมีรูปร่างเปลี่ยนไปจากเดิมได้ ถ้ามีการแปลงเกิดขึ้นกับภาพ ภาพวัตถุ 1 ภาพ อาจมีการแปลงหลายแบบเกิดขึ้นกับภาพ ซึ่งรูปร่างของภาพอาจมีเค้าโครงคล้ายคลึงภาพเดิม หรือ อาจไม่เหลือลักษณะเดิมของภาพเลยก็ได้ซึ่งภาพวัตถุจะมีลักษณะแตกต่างไปจากเดิมมากน้อยเพียงใด ขึ้นอยู่กับระดับและชนิดของการแปลงที่เกิดขึ้น โดยการแปลงที่เกิดขึ้นนี้ เรียกว่า การแปลงเรขาคณิต (geometric transformation)

2.2.2 การแปลงเรขาคณิต (geometric transformation)

การแปลงเรขาคณิต [15] เป็นกระบวนการที่ทำให้องค์ประกอบทางเรขาคณิตในภาพ เช่น มุมระหว่างเส้นตรง ความยาวเส้นตรง หรือ ความขนานของเส้นตรง มีการเปลี่ยนแปลงไปซึ่งอาจแบ่งได้ 2 ขั้นตอนคือ การแปลงตำแหน่งของจุดภาพ (spatial transformation) ซึ่งทำให้รูปทรงของวัตถุเกิดการเปลี่ยนแปลงไปดังรูปที่ 2.15 และการแปลงค่าระดับสีเทาของภาพ (gray-level interpolation) ซึ่งมีผลให้ค่าความเข้มของสีในแต่ละตำแหน่งของจุดภาพเปลี่ยนแปลงไป ดังแสดงในรูปที่ 2.16 โดยในที่นี้ จะนำเสนอเฉพาะส่วนของการวิเคราะห์ภาพที่มีลักษณะการเกิดการเปลี่ยนของตำแหน่งของจุดภาพเท่านั้นเนื่องจากการแปลงรูปร่างของวัตถุ 3 มิติจะมีผลให้ตำแหน่งของจุดยอดต่าง ๆ ในภาพเปลี่ยนตำแหน่งไปเท่านั้น



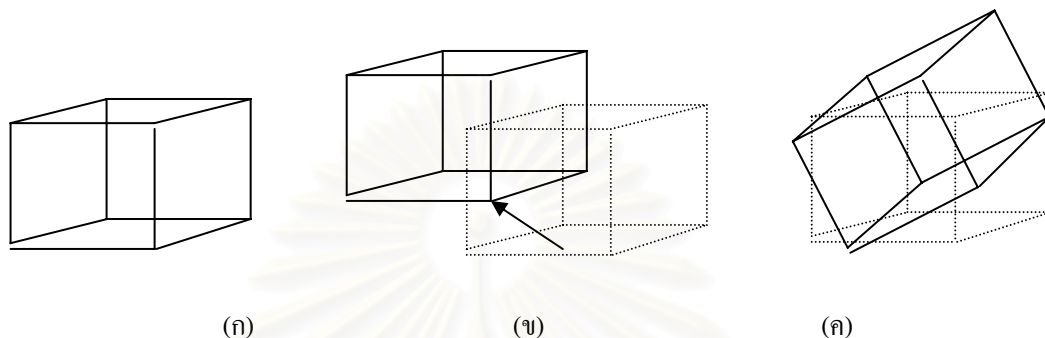
รูปที่ 2.15 (ก) ภาพก่อนเกิดการแปลง (ข) ภาพเกิดการเปลี่ยนแปลงตำแหน่งของจุดภาพ



รูปที่ 2.16 (ก) ภาพก่อนเกิดการแปลง (ข) ภาพเกิดการแปลงค่าระดับสีเทาของภาพ

2.2.2.1 การแปลงวัตถุแบบแข็งเกร็ง (rigid body transformation)

การแปลงวัตถุแบบแข็งเกร็ง เป็นการแปลงภาพโดยที่ภาพยังคงรักษารูปร่างเดิมของวัตถุเอาไว้ เช่น การเลื่อนขนาน (translation) ดังแสดงในรูปที่ 2.17 (ข) และการหมุนของวัตถุ (rotation) ดังแสดงรูปที่ 2.17 (ค) เป็นต้น จะเห็นได้ว่าการเปลี่ยนแปลงในลักษณะนี้จุดยอดมีการเปลี่ยนตำแหน่งไปในรูปแบบเดียวกันทั้งหมดโดยแผ่นผิวไม่เปลี่ยนแปลง ทำให้ระยะทางระหว่างจุดยอดและมุมระหว่างแผ่นผิวของภาพไม่เปลี่ยนแปลง จึงสามารถทำการแปลงกลับ (inverse transformation) เพื่อหารูปร่างเดิมของภาพวัตถุก่อนการแปลงได้ง่าย

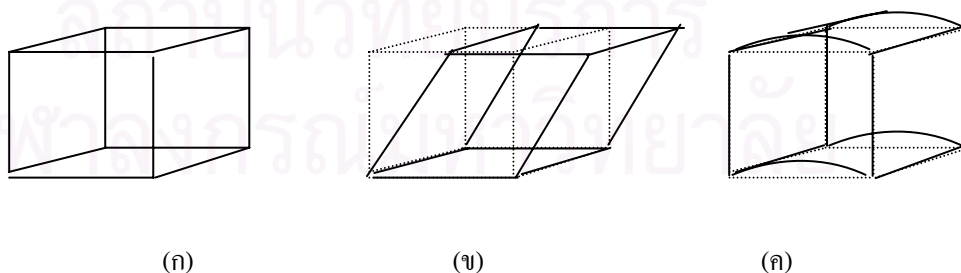


รูปที่ 2.17 การแปลงวัตถุแบบแข็งเกร็ง

(ก) รูปทรงของวัตถุเดิม (ข) การเลื่อนขนานจากวัตถุเดิม (ค) การหมุนจากวัตถุเดิม

2.2.2.2 การแปลงวัตถุแบบไม่แข็งเกร็ง (non-rigid body transformation)

การแปลงวัตถุแบบไม่แข็งเกร็งเป็นการแปลงภาพโดยที่รูปร่างเดิมของวัตถุเปลี่ยนแปลงไปด้วย เนื่องจากจุดยอดบางส่วนหรือทั้งหมดในภาพมีการเปลี่ยนแปลงโดยมีรูปแบบแตกต่างกัน ในขณะที่แผ่นผิวยังมีลำดับของจุดยอดคงเดิม แต่รูปร่างของแผ่นผิวอาจไม่เหมือนเดิม ทำให้ระยะทางระหว่างจุดยอดและมุมระหว่างแผ่นผิวของภาพมีการเปลี่ยนแปลงไป เช่น การเฉือน (shearing) ดังแสดงในรูปที่ 2.18 (ข) หรือ การเกิดการบิดงอ (bending) ดังแสดงในรูปที่ 2.18 (ค) หากภาพวัตถุเกิดการแปลงในลักษณะนี้ การแปลงกลับเพื่อหารูปร่างแบบเดิมของภาพวัตถุมีความซับซ้อนกว่ากรณีเกิดการแปลงวัตถุแบบแข็งเกร็ง



รูปที่ 2.18 การแปลงวัตถุแบบไม่แข็งเกร็ง

(ก) รูปทรงของวัตถุเดิม (ข) การเฉือนจากวัตถุเดิม (ค) การบิดงอจากวัตถุเดิม

2.2.3 การสร้างภาพ 3 มิติ

การสร้างภาพ 3 มิติ [16] ในปัจจุบันนี้มีการพัฒนาไปมาก ทั้งอุปกรณ์รับภาพ และศักยภาพของคอมพิวเตอร์ที่ใช้เพื่อประมวลผลภาพ ซึ่งการรับและสร้างข้อมูลภาพ 3 มิติก็มีหลากหลายวิธี โดยแต่ละวิธีต่างมีข้อดีข้อเสียแตกต่างกันไป ซึ่งสามารถแบ่งได้เป็น

2.2.3.1 การสร้างภาพ 3 มิติ จากภาพ 2 มิติหลายภาพ (stereo acquisition)

การสร้างภาพ 3 มิติโดยใช้ข้อมูลภาพ 2 มิติจำนวนมากกว่า 1 ภาพที่มุมกล้องต่างกันดังแสดงในรูปที่ 2.19 โดยการเปรียบเทียบตำแหน่งของจุดภาพที่มีความสัมพันธ์ (correspondence point) ระหว่างภาพที่ทำการประมวลผล เพื่อหาภาพความเหลื่อม (disparity map) เพื่อคำนวณหาข้อมูลความลึกของภาพ ซึ่งจะบอกรายละเอียดความลึกของภาพที่ต้องการ จากนั้นทำการ เก็บข้อมูลภาพ 3 มิติที่ได้เพื่อใช้ในการประมวลผลภาพที่ต้องการต่อไป การสร้างภาพด้วยวิธีนี้สามารถใช้กล้องรับภาพแบบที่มีใช้กันอยู่ในปัจจุบันสร้างข้อมูลภาพ 3 มิติขึ้นมาได้ ซึ่งทำให้สะดวกและประหยัดต่อการนำไปใช้งาน แต่ภาพที่ได้จากวิธีนี้จะมีรายละเอียดและถูกต้องแม่นยำต่ำเนื่องจากข้อมูลที่ได้ทั้งจากการหาเปรียบเทียบตำแหน่งของจุดภาพที่มีความสัมพันธ์ การหาภาพความเหลื่อม หรือ การประมวลผลภาพอาจมีความผิดพลาดของข้อมูลเกิดขึ้นได้



รูปที่ 2.19 ภาพที่ถ่ายต่างมุมกันจะมีข้อมูลภาพแตกต่างกัน

2.2.3.2 การสร้างภาพ 3 มิติ จากกล้องโครงสร้างแสง (structural light camera)

การสร้างภาพ 3 มิติ จากกล้องโครงสร้างแสง เป็นเทคโนโลยีที่ปรับปรุงโครงสร้างของระบบกล้องที่ใช้รับภาพให้สามารถรับข้อมูลความลึกของพื้นผิวภาพได้จาก การฉายแสงที่มีรูปแบบแน่นอนไปยังพื้นผิววัตถุ แล้วทำการตรวจจับการบิดเบี้ยว (distortion) ที่เกิดขึ้นเพื่อนำมาประมวลหาข้อมูลความลึกของพื้นผิวที่ตำแหน่งต่างๆ และนำมาสร้างภาพ 3 มิติของวัตถุที่ต้องการได้ โดยการสร้างภาพ 3 มิติจากวิธีนี้ นอกจากจะได้รายละเอียดของข้อมูลที่แม่นยำมากขึ้น อีกทั้งยังสามารถทำการถ่ายภาพได้ไวโดยใช้อุปกรณ์ที่มีราคาไม่สูงนัก



(ก)



(ข)

รูปที่ 2.20 การสร้างภาพ 3 มิติ จากกล้องโครงสร้างแสง

(ก) กล้องถ่ายภาพ 3 มิติแบบกล้องโครงสร้างแสง (ข) การถ่ายภาพแบบกล้องโครงสร้างแสง

2.2.3.3 การสร้างภาพ 3 มิติ จากตัวรับภาพแบบเลเซอร์ (laser sensor)

เทคโนโลยีการถ่ายภาพ 3 มิติด้วยตัวรับภาพแบบเลเซอร์จะได้ข้อมูลที่มีความแม่นยำสูง เนื่องจากการใช้เลเซอร์ตรวจวัดความลึกของพื้นผิวจะได้ข้อมูลที่มีความละเอียดสูงมากและได้ค่าที่แน่นอน แต่วิธีนี้ยังไม่สะดวกในการทำงานในปัจจุบันมากนัก เนื่องจากต้องใช้เทคโนโลยีขั้นสูงรวมถึงอุปกรณ์ที่มีราคาแพงอีกทั้งยังใช้เวลาในการรับภาพแต่ละภาพค่อนข้างนาน



(ก)



(ข)

รูปที่ 2.21 การสร้างภาพ 3 มิติ จากตัวรับภาพแบบเลเซอร์

(ก) ตัวรับภาพแบบเลเซอร์ (ข) การถ่ายภาพด้วยตัวรับภาพแบบเลเซอร์

2.2.4 การรู้จำใบหน้าแบบ 3 มิติ (3D face recognition)

กระบวนการรู้จำใบหน้าเป็นงานวิจัยที่ได้รับความสนใจและมีการพัฒนาอย่างต่อเนื่องจนถึงปัจจุบัน ในช่วงไม่กี่ปีมานี้ แนวทางวิจัยกระบวนการรู้จำใบหน้าเริ่มหันมาสนใจที่การวิเคราะห์ภาพในแบบ 3 มิติ เนื่องจากสามารถแก้ไขปัญหาค่าเป็นขีดจำกัดของการวิเคราะห์ภาพในแบบ 2 มิติที่ผ่านมา โดยเฉพาะการแก้ปัญหาในเรื่องของการแปรผันของตำแหน่งภาพและการแปรผันของแสง เนื่องจากการประมวลผลในแบบ 3 มิติ จะสามารถแสดงรายละเอียดของภาพได้ทุกมุมมอง รวมทั้งการเก็บข้อมูลความลึกของภาพจะไม่ได้รับผลกระทบจากการเปลี่ยนแปลงปริมาณของแสงในบริเวณที่ทำการถ่ายภาพวัตถุ ในการวิจัยด้านการรู้จำใบหน้าแบบ 3 มิติ [3] สามารถแบ่งขั้นตอนวิธีที่ใช้การวิเคราะห์ออกได้ดังรายละเอียดต่อไปนี้

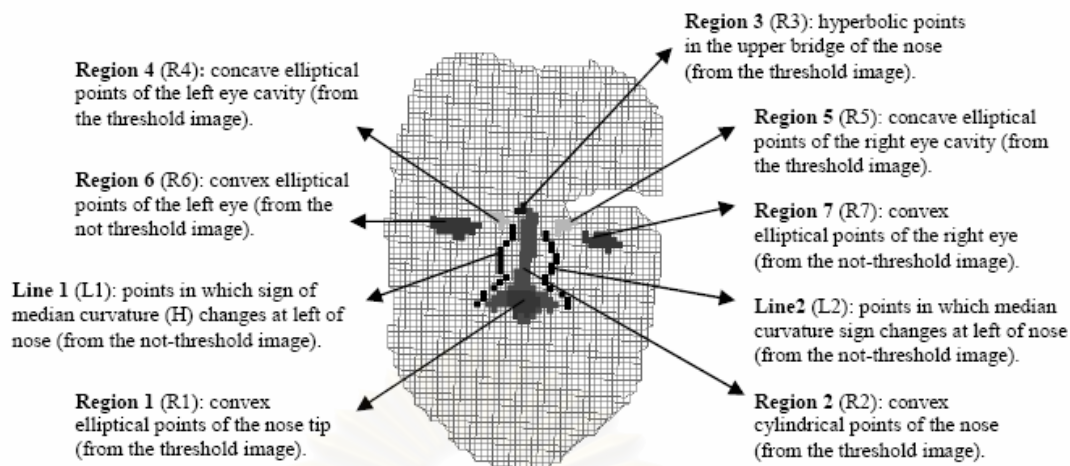
2.2.4.1 กระบวนการรู้จำใบหน้าโดยใช้ข้อมูลรูปร่างของภาพ 3 มิติเท่านั้น (recognition using 3D shape only)

เนื่องจากข้อมูลภาพแบบ 3 มิติ จะมีรายละเอียดของรูปร่างใบหน้าทั้งตำแหน่งและความลึกของภาพ จึงมีการนำข้อมูลส่วนนี้มาใช้ในกระบวนการรู้จำใบหน้า โดยการวิเคราะห์ข้อมูลภาพอาจแบ่งตามลักษณะการคำนวณได้ 2 วิธี คือ

1. การวิเคราะห์ข้อมูลด้วยวิธีอิงความโค้งจากพื้นผิวของภาพวัตถุ (curvature-based)

การวิเคราะห์ข้อมูลด้วยวิธีอิงความโค้งจากพื้นผิวของภาพวัตถุวิเคราะห์ลักษณะของอวัยวะบนใบหน้ามีความโค้งที่แตกต่างกัน เช่น จมูกมีความโค้งยื่นออกมาจากใบหน้า ในขณะที่กระบอกตาจะมีความโค้งเว้าเข้าไปจากผิวหนัง และทำการหาแบบจำลองความโค้งของใบหน้าเพื่อใช้ในการแยกแยะความแตกต่างของบุคคล ซึ่งขั้นตอนวิธีในกลุ่มนี้มักเกิดปัญหาเมื่อภาพใบหน้ามีขนาดที่แตกต่างกันเนื่องจากค่าความโค้งของพื้นผิวภาพอาจมีค่าไม่คงที่ได้ แต่ขั้นตอนวิธีแบบนี้จะรองรับผลกระทบการแปรผันจากการแสดงออกทางใบหน้าได้ดี ตัวอย่างขั้นตอนวิธีในกลุ่มนี้ได้แก่ กระบวนการรู้จำใบหน้าที่น่าเสนอโดย A. B. Moreno และคณะ [17]

ในงานวิจัยชิ้นนี้ นำเสนอการวิเคราะห์ภาพโดยการแบ่งพื้นที่ใบหน้าด้วยข้อมูลความโค้งเกาส์เซียนของพื้นผิวภาพใบหน้าออกเป็น 7 พื้นที่และเส้นจำนวน 2 เส้น คือ ปลายจมูก สันจมูก ค้างจมูก เบ้าตาซ้าย เบ้าตาขวา ตาซ้าย ตาขวา เส้นข้างจมูกซ้าย และเส้นข้างจมูกขวา ซึ่งภาพแสดงการแบ่งส่วนพื้นที่ใบหน้าตามความโค้งของพื้นผิวได้ดังรูปที่ 2.22 โดยคุณลักษณะของภาพถูกกำหนดด้วยข้อมูลจากความโค้งของพื้นผิวที่อิสระต่อกันจำนวน 86 ค่า ซึ่งมีข้อมูลจำนวนเพียง 35 ค่าที่เหมาะสมต่อการใช้เป็นคุณลักษณะสำหรับกระบวนการรู้จำใบหน้า

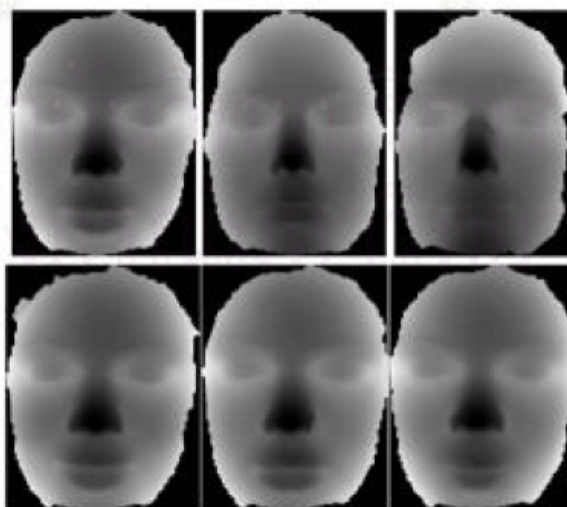


รูปที่ 2.22 ภาพแสดงการแบ่งส่วนพื้นที่ใบหน้าตามความโค้งของพื้นผิว

2. วิธีการวิเคราะห์เวกเตอร์แกนหลักของภาพ (PCA-based)

วิธีการวิเคราะห์เวกเตอร์แกนหลักของภาพมีหลักการคำนวณคล้ายคลึงกับการวิเคราะห์เวกเตอร์ของข้อมูลภาพแบบ 2 มิติ แต่ความแตกต่าง คือ กรณีนี้จะเป็นการคำนวณจากข้อมูลภาพ 3 มิติ นั่นเอง ซึ่งขั้นตอนวิธีนี้ มักไม่ได้รับผลกระทบจากขนาดภาพที่ไม่เท่ากันแต่จะมีความผิดพลาดได้เมื่อภาพมีการแสดงออกทางใบหน้า ตัวอย่างขั้นตอนวิธีในกลุ่มนี้ ได้แก่ กระบวนการรู้จำใบหน้าที่น่าสนใจโดย C. Heshier และคณะ [18]

ในงานวิจัยมีการนำเสนอการใช้ข้อมูลภาพ 3 มิติในการสร้างภาพพิสัย (range image) โดยการฉายข้อมูลจุดยอดของภาพในปริภูมิ 3 มิติ ไปยังระนาบ 2 มิติ ซึ่งภาพที่ได้จะแตกต่างจากการถ่ายภาพใบหน้าปกติ เนื่องจากภาพที่ได้เป็นภาพแสดงข้อมูลความลึกของใบหน้า ดังแสดงในรูปที่ 2.23 จากนั้นนำภาพที่ได้ไปวิเคราะห์โดยแยกขั้นตอนออกเป็น 2 วิธี วิธีหนึ่งคือการดึงคุณลักษณะของภาพพิสัยด้วยกระบวนการวิเคราะห์ภาพจากองค์ประกอบหลักและ อีกวิธีเป็นการวิเคราะห์องค์ประกอบอิสระ เพื่อดึงคุณลักษณะของภาพออกมา



รูปที่ 2.23 ภาพพิสัยที่สร้างจากขั้นตอนวิธีของ C. Heshier และคณะ

2.2.4.2 กระบวนการรู้จำใบหน้าแบบ 2 มิติ ร่วมกับข้อมูลภาพ 3 มิติ (recognition using 3D shape assisted 2D)

กระบวนการรู้จำใบหน้าแบบ 2 มิติ โดยใช้ข้อมูลภาพ 3 มิติเสริมการทำงานเป็นการรวมการวิเคราะห์ข้อมูลภาพใบหน้าโดยใช้ข้อมูลร่วมกันระหว่างข้อมูลภาพ 2 มิติและ 3 มิติ โดยการนำแบบจำลองภาพ 3 มิติมาใช้ร่วมกับการวิเคราะห์ข้อมูลภาพ 2 มิติ ในการวิเคราะห์ภาพลักษณะนี้จะเน้นไปที่การแก้ปัญหาการแปรผันของแสงในภาพ 2 มิติร่วมกันกับการวิเคราะห์ภาพ 3 มิติ โดยมีการนำแบบจำลองการสะท้อนของแสงในพื้นที่ผิวลักษณะต่างๆมาใช้ เพื่อลดผลกระทบด้านแสงที่มีต่อความแม่นยำในการแยกแยะบุคคลของภาพ

2.2.4.3 กระบวนการรู้จำใบหน้าโดยใช้ข้อมูลรูปร่างและพื้นผิวภาพ 3 มิติ (recognition using 3D shape and texture)

การวิเคราะห์ภาพ 3 มิติ โดยใช้ข้อมูลทั้งหมดที่ภาพมี ซึ่งประกอบด้วยรูปร่างของใบหน้า และ พื้นผิวของใบหน้าในการวิเคราะห์แยกแยะความแตกต่างระหว่างบุคคล เนื่องจากเป็นการใช้ประโยชน์จากข้อมูลภาพทั้งหมดในการคำนวณ การวิเคราะห์ภาพในขั้นตอนวิธีลักษณะนี้จะสามารถแสดงรายละเอียดของภาพได้ทั้งการแปรผันของตำแหน่งภาพ การแปรผันของแสง และการแสดงอารมณ์ของบุคคลในภาพใบหน้า ซึ่งอาจมีการใช้การวิเคราะห์ภาพแบบเวกเตอร์มาร่วมกับข้อมูลภาพ 2 มิติ เพื่อเพิ่มประสิทธิภาพของการคำนวณมากยิ่งขึ้น

จากที่ได้กล่าวมาทั้งหมด สามารถสรุปกระบวนการรู้จำใบหน้าแบบต่างๆเป็นตารางเพื่อเปรียบเทียบข้อดีข้อด้อยของกระบวนการรู้จำใบหน้าแบบต่างๆ โดยคร่าว่ได้ดังตาราง 2. 1

ตาราง 2. 1 การเปรียบเทียบข้อดีข้อด้อยของกระบวนการรู้จำใบหน้าแบบต่างๆ

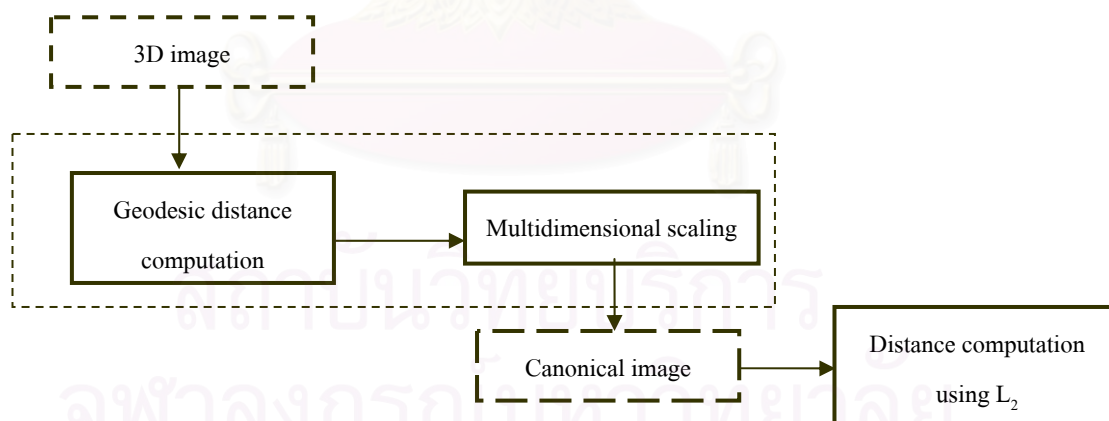
	ข้อดี	ข้อด้อย
กระบวนการรู้จำใบหน้าแบบ 2 มิติ		
วิธีอิงลักษณะที่ปรากฏโดยรวม	มีการคำนวณรวดเร็ว และมีความแม่นยำทนทานต่อการแปรผันจากปัจจัยภายนอกได้บ้าง	ความแม่นยำจะลดลง เมื่อเกิดการแปรผันจากปัจจัยภายนอก ซึ่งต้องใช้กลุ่มภาพทดลองขนาดใหญ่ เพื่อลดผลกระทบที่เกิดขึ้น
วิธีอิงแบบจำลองภาพ	ทนทานต่อการแปรผันของตำแหน่ง และ การแปรผันจากแสงได้ดี	ความแม่นยำขึ้นอยู่กับกระบวนการหาตำแหน่งองค์ประกอบบนใบหน้าที่ดี
กระบวนการรู้จำใบหน้าแบบ 3 มิติโดยใช้ข้อมูลรูปร่างเท่านั้น		
วิธีอิงความโค้งจากพื้นผิว	มีการคำนวณละเอียด และมีความแม่นยำ	มีปัญหาจากการคำนวณเมื่อภาพมีจำนวนจุดยอด หรือขนาดของภาพแตกต่างกัน
วิธีการวิเคราะห์เวกเตอร์แกนหลัก	มีการคำนวณรวดเร็ว และทนทานต่อการแปรผันเนื่องจากการแสดงออกของใบหน้าได้บ้าง	การคำนวณจะคลาดเคลื่อนเมื่อเกิดการแปรผันเนื่องจากการแสดงออกของใบหน้า เนื่องจากใช้คำนวณแบบวัตถุแข็งเกร็ง

กระบวนการรู้จำใบหน้าแบบ 3 มิติ มีข้อดีในการลดปัญหาจากการแปรผันทางตำแหน่งและการแปรผันของแสงได้ดีกว่ากระบวนการรู้จำใบหน้าแบบ 2 มิติ อย่างไรก็ตามกระบวนการรู้จำใบหน้าแบบ 3 มิติเองก็ได้รับผลกระทบจากการแปรผันเนื่องจากการแสดงออกของใบหน้าเป็นอย่างมาก โดยเฉพาะกับการวิเคราะห์ด้วยวิธีการวิเคราะห์เวกเตอร์แกนหลัก จึงมีการนำเสนอกระบวนการรู้จำใบหน้าแบบ 3 มิติโดยไม่แปรผันตามการแสดงออกของบุคคลในภาพ โดยใช้ภาพแบบบัญญัติ ซึ่งใช้ระยะทางจีโอเดสิกของจุดยอดในภาพเป็นคุณลักษณะของใบหน้า

2.3. การรู้จำใบหน้าแบบ 3 มิติโดยไม่แปรผันตามการแสดงออกของบุคคลในภาพ

เนื่องจากกระบวนการรู้จำภาพแบบ 3 มิติจะได้รับผลของการแปรผันเนื่องจากการแสดงออกของใบหน้า Bronstein และคณะ ได้นำเสนองานวิจัยโดยใช้สมมุติฐานว่า การแปรผันเนื่องจากการแสดงออกของใบหน้านั้นทำให้ภาพใบหน้า 3 มิติเกิดการแปลงวัตถุแบบไม่แข็งเกร็งในภาพ โดยใช้ข้อมูลของระยะทางจีโอเดสิกบนพื้นผิวของภาพใบหน้าในการดึงข้อมูลภาพที่ไม่แปรผันตามการแสดงออกของบุคคลเพื่อสร้างภาพแบบบัญญัติซึ่งมีคุณสมบัติไม่แปรเปลี่ยนเมื่อภาพใบหน้าเกิดการแปลงแบบไม่แข็งเกร็งอันเนื่องมาจากการแปรผันเนื่องจากการแสดงออกของใบหน้า

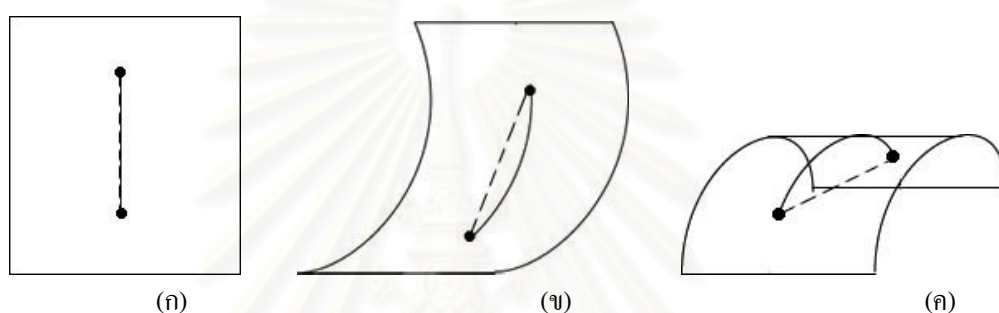
ภาพแบบบัญญัติของใบหน้า [19] คือภาพ 3 มิติที่มีคุณสมบัติต่างๆที่ แม้ว่าภาพเริ่มต้นจะเกิดการแปลงแบบต่างๆ ขึ้น โดยสร้างจากการนำพื้นผิวของภาพในปริภูมิ 3 มิติมาคำนวณหาคุณสมบัติภายในของพื้นผิวซึ่งมีคุณสมบัติไม่แปรเปลี่ยนเมื่อเกิดการแปลงของภาพ โดยในงานวิจัยของ Bronstein และคณะ ได้เลือกใช้ระยะทางจีโอเดสิกหรือระยะทางบนพื้นผิววัตถุระหว่างจุดยอดของภาพเป็นคุณสมบัติภายในของภาพ จากนั้นทำการแปลงข้อมูลคุณสมบัติภายในของภาพหรือระยะทางจีโอเดสิกระหว่างจุดยอดของภาพไปยังปริภูมิอื่นด้วยขั้นตอนวิธีมิติลดแบบหลายมิติ (multidimensional scaling) ซึ่งทำการแสดงผลข้อมูลของวัตถุออกมาเป็นภาพในปริภูมิเรขาคณิต (geometric space) โดยสามารถแสดงขั้นตอนการทำงานได้ดังรูปที่ 2.24



รูปที่ 2.24 โครงสร้างการหาภาพแกนหลักของใบหน้า

2.3.1 การหาระยะทางจีโอเดสิกบนพื้นผิว

คุณสมบัติของพื้นผิวอาจแบ่งได้เป็น 2 ประเภท คือ คุณสมบัติภายนอกซึ่งเมื่อวัตถุเกิดการแปลง แล้วจะมีการเปลี่ยนรูปร่างไปตามลักษณะการแปลงที่เกิดขึ้น ส่วนคุณสมบัติภายในนั้นยังมีค่าคงที่แม้ว่าวัตถุจะเกิดการแปลงไปในรูปแบบอื่นๆแล้วก็ตาม การหาระยะทางจีโอเดสิกระหว่างจุดบนพื้นผิวเป็นการคำนวณคุณสมบัติภายในของวัตถุวิธีหนึ่ง โดยการหาระยะห่างระหว่างจุดบนพื้นผิวโค้งแสดงด้วยเส้นทึบในรูปที่ 2.25 ซึ่งจะเห็นว่าแม้พื้นผิวจะเกิดการบิดงอไป ระยะทางจีโอเดสิก(เส้นทึบ)ระหว่างจุด A และ B ยังมีค่าเท่าเดิมไม่เปลี่ยนแปลงไป ซึ่งแตกต่างจากระยะทางยูคลิด (euclidean distance) ซึ่งแสดงด้วยเส้นประที่ค่าไม่คงที่เมื่อวัตถุเกิดการแปลงในรูปแบบต่างๆ วิธีการคำนวณหาระยะทางจีโอเดสิกที่มีประสิทธิภาพวิธีหนึ่งคือ กระบวนการฟาสท์มาร์ชชิง[20]



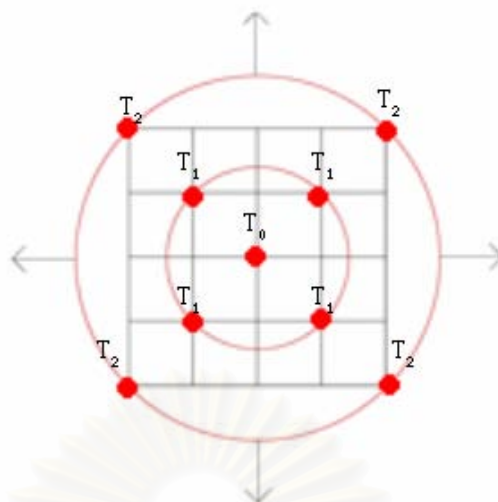
รูปที่ 2.25 การเปรียบเทียบความแตกต่างของระยะทางจีโอเดสิก(เส้นทึบ)และระยะทางยูคลิด(เส้นประ)

(ก) ระยะทางระยะทางจีโอเดสิกและระยะทางยูคลิดบนพื้นผิวที่ไม่มีการแปลง

(ข) และ(ค) ระยะทางระยะทางจีโอเดสิกและระยะทางยูคลิดบนพื้นผิวที่การแปลงเกิดขึ้น

การหาระยะทางจีโอเดสิกจะกระทำโดยอาศัยกระบวนการฟาสท์มาร์ชชิง ซึ่งการคำนวณหาระยะทางระหว่างจุดใดๆบนพื้นผิวจากกระบวนการฟาสท์มาร์ชชิง มีแนวคิดที่ เมื่อสมมุติหน้าคลื่นวิ่งผ่านพื้นผิวใด ๆ ด้วยแรง F คงที่ โดยนับจากจุดเริ่มต้นที่ T_0 หน้าคลื่นเริ่มเคลื่อนที่บนผิววัตถุออกไปรอบทิศทาง ดังนั้นจุดยอดจุดอื่น ๆ ที่อยู่รอบจุดเริ่มต้นจะมองเห็นหน้าคลื่นเคลื่อนที่ผ่านตำแหน่งพิกัดของแต่ละจุดที่ T ที่แตกต่างกันออกไป โดยจุดที่อยู่ใกล้จุดเริ่มต้นจะมองเห็นหน้าคลื่นเคลื่อนที่ผ่านที่ T น้อยกว่าจุดที่อยู่ไกลจุดเริ่มต้น โดยระยะทางจีโอเดสิกจะคำนวณได้จากความแตกต่างของเวลา T ที่จุดยอดแต่ละจุด

กระบวนการฟาสท์มาร์ชชิงเป็นหนึ่งในวิธีการหาเซตระดับ แต่แตกต่างจากเซตระดับโดยทั่วไปที่ให้น้ำคลื่นเคลื่อนที่ไปได้ในทุกทิศทางรวมถึงการเคลื่อนที่ย้อนกลับในทิศทางที่หน้าคลื่นวิ่งผ่านไปแล้วได้ด้วย และแรง F อาจไม่คงที่ [21] ดังแสดงลักษณะการเคลื่อนที่ของหน้าคลื่นในกระบวนการฟาสท์มาร์ชชิงได้ดังรูปที่ 2.26



รูปที่ 2.26 การเคลื่อนที่ของหน้าคลื่นด้วยกระบวนการมาร์ชชิง

โดยลักษณะการเคลื่อนที่ของหน้าคลื่นในกระบวนการมาร์ชชิงสามารถแสดงได้ด้วยสมการอีโกนอล (eigonal equation) ซึ่งเขียนได้ดังสมการที่ (2.8)

$$|\nabla T| = F(x, y) \quad (2.8)$$

โดยที่ T คือ เวลาที่หน้าคลื่นเคลื่อนที่ผ่านจุดยอด
 F คือ แรงที่หน้าคลื่นใช้ในการเคลื่อนที่

จากสมการที่ (2.8) สามารถเขียนได้ในรูปการประมาณค่าบนตารางแบบเต็มหน่วย (discrete grid) เมื่อหน้าคลื่นมีการเคลื่อนที่ในทิศทางเดียวโดยกำหนดให้ $D_{ij}^{-x}, D_{ij}^{+x}, D_{ij}^{-y}, D_{ij}^{+y}$ คือเกรเดียนต์ในทิศทาง $-x, +x, -y, +y$ ตามลำดับ ได้ดังสมการที่ (2.9) โดยระยะทางจีโอเดสิกระหว่างคู่จุดใด ๆ บนพื้นผิววัตถุ 3 มิติ นั้นหาจากเวลา T ที่แตกต่างกันที่หน้าคลื่นวิ่งผ่านคู่จุดนั้น ๆ

$$\left[\max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0)^2 + \max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0)^2 \right]^{1/2} = F_{ij} \quad (2.9)$$

2.3.2 มัลติไดเมนชันแนล สเกลลิง (multidimensional scaling)

จากงานวิจัยเกี่ยวกับกระบวนการมัลติไดเมนชันแนล สเกลลิง [22]-[23] กระบวนการนี้ใช้ในการแก้ปัญหาการแยกแยะความเหมือนหรือความต่างของกลุ่มวัตถุ โดยใช้ระยะทางระหว่างกลุ่มวัตถุเป็นตัวชี้วัดความต่างต่างนั้นๆ ซึ่งตัวชี้วัดนี้อาจมีความคลาดเคลื่อนจากค่าความเหมือนหรือความต่างที่แท้จริงของกลุ่มวัตถุได้ ขั้นตอนวิธีมัลติไดเมนชันแนล สเกลลิงใช้หาตัวชี้วัดของข้อมูลที่ทำให้ค่าความผิดพลาดระหว่างตัวชี้วัดกับค่าจริงของข้อมูลมีค่าน้อยที่สุด

จากนั้นมีการนำมัลติไดเมนชันแนล สเกลลิงมาใช้แสดงความสัมพันธ์ของกลุ่มข้อมูลในรูปข้อมูลทางเรขาคณิต ซึ่งในที่นี้ความสัมพันธ์ของกลุ่มข้อมูล คือระยะทางจีโอเดสิกระหว่างจุดยอดของวัตถุ และการแสดงความสัมพันธ์ของกลุ่มข้อมูลออกมาเป็นปริภูมิข้อมูลทางเรขาคณิตคือการจับคู่ (mapping) ข้อมูลระยะทางจีโอเดสิกระหว่างจุดยอดไปยังปริภูมิเรขาคณิต ตัวอย่างเช่น ปริภูมิยูคลิด ปริภูมิพาราโบล่า ปริภูมิทรงกลม เป็นต้น โดยกำหนดค่าฟังก์ชันความเครียด (stress function) ของการจับคู่เป็นค่าความคลาดเคลื่อนของการจับคู่ข้อมูล จุดประสงค์ของวิธีมัลติไดเมนชันแนล สเกลลิง คือการจับคู่ข้อมูลที่ทำให้ฟังก์ชันความเครียดมีค่าน้อยที่สุด โดยฟังก์ชันความเครียดสามารถแสดงได้ดังสมการที่ (2.10)

$$\sigma_p(\Psi; \mathbf{D}, \mathbf{W}) = \left(\frac{1}{\sum_{j>i} w_{ij}} \sum_{j>i} (w_{ij} (d_{ij} - \Psi_{ij}))^p \right)^{\frac{1}{p}} \quad (2.10)$$

โดยที่ σ_p คือ ฟังก์ชันความเครียดของปริภูมิลำดับที่ p
 \mathbf{D} คือ เมทริกซ์ของระยะทางจีโอเดสิก d_{ij} ระหว่างจุดยอด i และ j
 Ψ คือ เมทริกซ์ข้อมูลของปริภูมิทางเรขาคณิต x_{ij} ระหว่างจุดยอด i และ j
 \mathbf{W} คือ เมทริกซ์ค่าถ่วงน้ำหนัก w_{ij} ระหว่างจุดยอด i และ j

เมื่อต้องการหาฟังก์ชันความเครียดเมื่อต้องการแสดงข้อมูลระยะทาง \mathbf{D} ไปยังปริภูมิยูคลิด Ψ โดยไม่มีการถ่วงน้ำหนักของข้อมูลจะเขียนสมการที่ (2.10) ได้ใหม่ดังสมการที่ (2.11)

$$\sigma(\Psi; \mathbf{D}) = \frac{1}{n} \sum_{j>i} (d_{ij} - \Psi_{ij})^2 \quad (2.11)$$

ขั้นตอนวิธีมัลติไดเมนชันแนล สเกลลิงแสดงการหาปริภูมิข้อมูลใหม่ของเมทริกซ์ระยะทาง \mathbf{D} ซึ่งทำให้ค่าฟังก์ชันความเครียดมีค่าน้อยสุดโดยวิธีการการแยกองค์ประกอบหลักของเมทริกซ์ระยะทาง แต่เนื่องจากเมทริกซ์ระยะทางอาจขาดคุณสมบัติ positive semi definite ซึ่งมีในเมทริกซ์สหสัมพันธ์ (correlation matrix) เมทริกซ์ความแปรปรวนร่วมเกี่ยว และเมทริกซ์ผลคูณไขว้ (cross product matrix)

ดังนั้นก่อนทำการแยกองค์ประกอบหลักของเมทริกซ์ระยะทางจะต้องทำการหาศูนย์กลางซ้อน (double centering) ของเมทริกซ์ระยะทางเพื่อหาผลคูณไขว้ของเมทริกซ์เพื่อให้สามารถทำการแยกองค์ประกอบหลักของเมทริกซ์ระยะทางได้ ซึ่งการหาศูนย์กลางซ้อนของเมทริกซ์ระยะทางแสดงได้ดังสมการที่ (2.12)

$$B = -\frac{1}{2} \mathbf{E} \mathbf{D} \mathbf{E}^T \quad (2.12)$$

โดยที่ B คือ ผลคูณไขว้ของเมทริกซ์ระยะทาง
 \mathbf{E} คือ เมทริกซ์ศูนย์กลางซ้อน (centering matrix)

โดย เมทริกซ์ศูนย์กลางซ้อน Ξ ได้จากสมการที่ (2.13)

$$\Xi = \mathbf{I} - \frac{1}{n} \mathbf{1}^T \mathbf{1} \quad (2.13)$$

โดยที่ \mathbf{I} คือ เมทริกซ์เอกลักษณ์ขนาด $n \times n$
 $\mathbf{1}$ คือ เวกเตอร์ความยาว $1 \times n$ ซึ่งสมาชิกทุกตัวมีค่าเป็น 1

จากนั้นทำการหาผลคูณไขว้ของเมทริกซ์ปริภูมิโดยกำหนดให้มีค่าดังสมการที่ (2.14)

$$B = \Psi \Psi^T \quad (2.14)$$

โดยที่ Ψ คือ ปริภูมิที่ต้องการจับคู่ข้อมูลของเมทริกซ์ระยะทางไป

และทำการแยกองค์ประกอบหลักของเมทริกซ์ผลคูณไขว้ จะได้ผลดังสมการที่ (2.15)

$$B = \Pi \Lambda \Pi^T \quad (2.15)$$

โดยที่ Π คือ เวกเตอร์องค์ประกอบหลัก
 Λ คือ เมทริกซ์ซึ่งแสดงค่าองค์ประกอบหลักในแนวแกนทแยงมุม

ดังนั้นจะแสดงค่าปริภูมิที่ต้องการส่งผ่านข้อมูล Ψ ได้ดังสมการ(2.16)

$$\Psi = \Pi \Lambda^{\frac{1}{2}} \quad (2.16)$$

โดยการเลือกปริภูมิ Ψ เพื่อทำการส่งผ่านนั้นจะเลือกจากทอพอโลยี (topology) ของข้อมูลระยะทาง \mathbf{D} คือ ถ้าข้อมูลระยะทาง \mathbf{D} เป็นข้อมูลของรูปร่างที่มีทอพอโลยีแบบพื้นฐาน (simple topology) การส่งผ่านไป ยังพื้นผิวที่เป็นปริภูมิยูคลิดก็เพียงพอแล้ว แต่ถ้าข้อมูลระยะทาง \mathbf{D} เป็นข้อมูลของรูปร่างที่มีทอพอโลยีแบบ ซับซ้อน (sophisticated topology) การส่งผ่านอาจต้องใช้ปริภูมิที่มีความซับซ้อนขึ้นเช่น ปริภูมิพาราโบลา ปริภูมิ ทรงกลม เพื่อให้การส่งผ่านมีความผิดเพี้ยนน้อยที่สุดจากที่กล่าวมา การสร้างภาพแบบบัญญัติ เริ่มจากการหา ความสัมพันธ์ของจุดยอดของภาพวัตถุด้วยระยะทางจีโอเดสิกระหว่างจุด และส่งผ่านข้อมูลระยะทางจีโอเดสิก ระหว่างจุดไปยังปริภูมิยูคลิดเพื่อแสดงผลข้อมูลออกมาเป็นภาพทางเรขาคณิตด้วยขั้นตอนวิธีมัลติไดเมนชันแนล สเกลลิง

จากขั้นตอนที่ได้กล่าวมา ข้อมูลของระยะทางจีโอเดสิกระหว่างจุด \mathbf{D} ที่แสดงผลด้วยค่าในปริภูมิ X จะถูกนำมาสร้างภาพแบบบัญญัติซึ่งมีคุณสมบัติไม่เปลี่ยนแปลงไปตามการแปรผันเนื่องจากการแสดงออกของ ใบหน้าเพื่อใช้ในการวิเคราะห์แยกแยะบุคคล

2.4. บริบทรูปร่าง (shape context)

บริบทรูปร่างเป็นตัวบอกรูปร่าง (shape descriptor) ชนิดหนึ่ง ซึ่งมีลักษณะการคำนวณง่ายและไม่แปรผันตามการเปลี่ยนแปลงของรูปร่างวัตถุ คุณลักษณะที่คำนวณได้จากขั้นตอนวิธีบริบทรูปร่างมีข้อมูลเป็นฮิสโทแกรมของจุดยอดในภาพ ซึ่งบอกลักษณะการกระจายตัวของตำแหน่งจุดยอดต่างๆรอบๆจุดอ้างอิง (reference point) ที่กำหนด โดยการกำหนดว่าจุดยอดใดควรเป็นจุดอ้างอิงนั้นอาจใช้ทุกจุดเป็นจุดอ้างอิงหรือเลือกเฉพาะจุดที่มีความสำคัญมาเป็นจุดอ้างอิงตามลักษณะการใช้งาน จากนั้นนำฮิสโทแกรมที่คำนวณจากแต่ละตำแหน่งจุดอ้างอิงมาใช้เป็นคุณลักษณะของภาพวัตถุนั้นๆ ซึ่งการคำนวณด้วยวิธีนี้มีข้อดีคือคุณลักษณะที่ได้แสดงทั้งคุณสมบัติครอบคลุม (global property) และคุณสมบัติเฉพาะที่ (local property) ของวัตถุ โดยในหัวข้อนี้ เราจะกล่าวถึงรายละเอียดของการคำนวณบริบทรูปร่างแบบดั้งเดิมที่เป็นการคำนวณสำหรับภาพ 2 มิติ และบริบทรูปร่างที่มีการพัฒนาให้ทำงานกับภาพ 3 มิติ รวมทั้งบริบทรูปร่างแบบ 3 มิติที่นำมาใช้ในวิทยานิพนธ์ฉบับนี้

2.4.1 บริบทรูปร่างสำหรับภาพ 2 มิติ

การใช้งานบริบทรูปร่างสำหรับภาพ 2 มิตินั้น ภาพวัตถุที่ใช้ควรเป็นภาพที่บอกรูปร่างของวัตถุนั้น เช่น ภาพโครงร่างของวัตถุ (silhouette image) หรือ ภาพขอบของวัตถุ (edge map) เป็นต้น ด้วยแนวคิดที่คล้ายคลึงกับตัวบอกรูปร่างฟูรีเยร์ (fourier descriptor) ซึ่งใช้ค่าสัมประสิทธิ์ของภาพในเชิงความถี่มาเป็นคุณลักษณะของภาพวัตถุ ส่วนบริบทรูปร่างเป็นการนำค่าสัมประสิทธิ์ของภาพในเชิงตำแหน่งหรือความถี่ที่จุดยอดของภาพวัตถุปรากฏในตาราง (grid) มาใช้แทนคุณลักษณะของภาพวัตถุ บริบทรูปร่างทำการคำนวณหาการกระจายตัวของพิกัดของจุดยอดในปริภูมิ 3 มิติรอบจุดอ้างอิงโดยการหาฮิสโทแกรมเชิงขั้ว (polar histogram) ของจุดบนขอบของภาพ โดยจุดอ้างอิงอาจเป็นจุดบนขอบภาพทุกจุดหรืออาจเป็นเพียงบางจุดก็ได้ขึ้นอยู่กับความเหมาะสมกับภาพที่ต้องการ เช่น ถ้าภาพมีขนาดเล็ก จำนวนจุดยอดที่เป็นขอบภาพมีจำนวนน้อย ทุกจุดสีบนขอบภาพควรจะถูกกำหนดให้เป็นจุดอ้างอิงและทำการคำนวณหาฮิสโทแกรมเชิงขั้วทุกตำแหน่งอ้างอิงเพื่อนำมาใช้เป็นคุณลักษณะแทนข้อมูลภาพ แต่หากภาพที่ได้มีขนาดใหญ่และภาพขอบมีความละเอียดสูง อาจทำการสุ่มเพียงบางจุดมาใช้เป็นจุดอ้างอิงเพื่อหลีกเลี่ยงการคำนวณที่มากเกินไปจนความจำเป็น เป็นต้น

การหาบริบทรูปร่างของภาพวัตถุ เริ่มต้นจากการคำนวณฮิสโทแกรมของจุดที่ตำแหน่งจุดอ้างอิง p_i ซึ่งฮิสโทแกรมที่คำนวณได้จะบอกลักษณะการกระจายตัวของจุดยอดจุดอื่นๆรอบๆจุดอ้างอิง p_i นั้น โดยค่าฮิสโทแกรมแต่ละค่าจะแสดงค่าที่จุดยอดปรากฏในตาราง (grid) ช่องที่ k ($bin(k)$) เมื่อ p คือจุดยอดทุกจุดบนขอบภาพวัตถุ ค่าฮิสโทแกรมของจุดอ้างอิงที่ i แสดงจำนวนจุดยอดที่มีพิกัดอยู่ในอาณาบริเวณของช่องที่ k คำนวณได้จากสมการที่ (2.17)

$$h_i(k) = \#\{p \neq p_i : (p - p_i) \in bin(k)\} \quad (2.17)$$

โดยที่ $h_i(k)$ คือ ค่าฮิสโทแกรม ของจุดยอดที่ i ในอาณาบริเวณช่องที่ k
 $\#$ คือ จำนวนจุดยอดในอาณาบริเวณช่องที่ k

โดยลักษณะพิเศษของฮิสโทแกรมเชิงขั้วที่ใช้ในการคำนวณหาปริมาตรรูปร่าง คือมีการแบ่งตารางในแนวรัศมีด้วยแกนแบบลอการิทึม เพื่อให้ตารางที่มีตำแหน่งอยู่ใกล้เคียงจุดอ้างอิงให้มีความละเอียดมากกว่า ตารางบริเวณที่อยู่ไกลจากจุดอ้างอิง p_i โดยการหารระยะระหว่างจุดยอด p ใดๆกับจุดยอดอ้างอิง p_i จะต้องแปลงค่าพิกัด (x, y) ของจุดยอดบนปริภูมิคาร์ทีเซียน 2 มิติ ไปยังพิกัด (r, θ) บนแกนลอการิทึมเชิงขั้ว ดังสมการที่ (2.18) และสมการที่ (2.19)

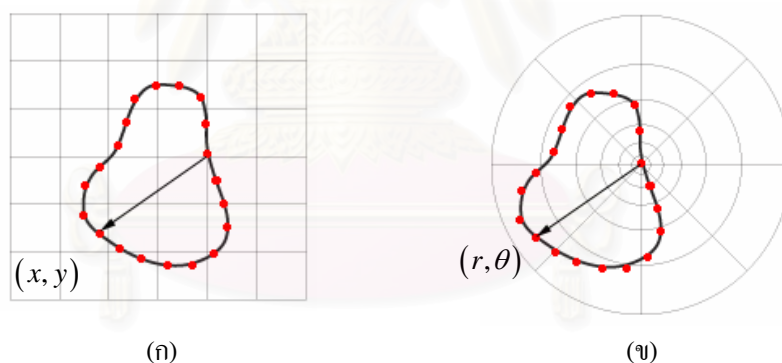
$$r = \log \sqrt{x^2 + y^2} \quad (2.18)$$

โดยที่ r คือ ระยะระหว่างจุดยอดกับจุดยอดอ้างอิง p_i

$$\theta = \tan^{-1} \left(\frac{y}{x} \right) \quad (2.19)$$

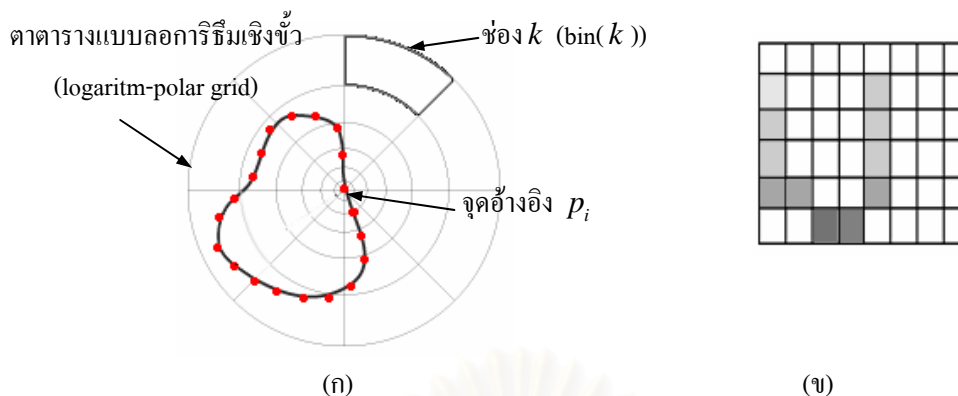
โดยที่ θ คือ มุมที่จุดยอดทำกับแกน ณ ตำแหน่งของจุดอ้างอิง p_i

เราสามารถแสดงลักษณะของการแบ่งช่องด้วยแกนลอการิทึมเชิงขั้วเปรียบเทียบกับกรแบ่งช่องด้วยจุดยอดบนปริภูมิคาร์ทีเซียน 2 มิติได้ดังรูปที่ 2.27 โดยรูปที่ 2.27 (ก) เป็นรูปของวัตถุบนตารางในพิกัดคาร์ทีเซียน ซึ่งแสดงตำแหน่งของจุดยอดด้วยพิกัด (x, y) และรูปที่ 2.27 (ข) เป็นรูปของวัตถุบนตารางแบบลอการิทึมเชิงขั้วซึ่งแสดงตำแหน่งจุดยอดด้วยพิกัด (r, θ)



รูปที่ 2.27 (ก) แกนภาพด้วยปริภูมิคาร์ทีเซียน (ข) แกนภาพด้วยแกนลอการิทึมเชิงขั้ว

ในรูปที่ 2.28 (ก) แสดงโครงสร้างของปริมาตรรูปร่างที่ประกอบด้วยตารางแบบลอการิทึมเชิงขั้วซึ่งประกอบไปด้วยช่อง ซึ่งมีการแบ่งช่องทั้งในแนวรัศมี r และการแบ่งช่องตามแนวของมุม θ ซึ่งฮิสโทแกรมของจุดยอดที่จุดอ้างอิง p_i รูปที่ 2.28 (ข) แสดงความถี่ที่จุดปรากฏในช่อง โดยถ้ามีจำนวนจุดปรากฏในช่องสูงจะแสดงด้วยสีเข้มและถ้ามีจำนวนจุดที่ปรากฏในช่องน้อยจะแสดงด้วยสีอ่อน



รูปที่ 2.28 (ก) โครงสร้างของบริบทรูปร่าง (ข) ฮิสโทแกรมของจุดยอดที่จุดอ้างอิง p_i

บริบทรูปร่าง คือ ฮิสโทแกรมของของจุดยอดที่ทุกๆตำแหน่งจุดอ้างอิง p_i ซึ่งฮิสโทแกรมของจุดยอดมีคุณสมบัติครอบคลุม คือเป็นการคำนวณจากภาพรวมของทั้งภาพ ทำให้การแปลงใดๆที่เกิดกับภาพมีผลต่อการเปลี่ยนแปลงของค่าฮิสโทแกรมที่ได้ไม่มากนัก แต่ฮิสโทแกรมจะแสดงรายละเอียดปลีกย่อยของภาพได้ไม่ดัดนัก ดังนั้นในการหาบริบทรูปร่างของวัตถุจะคำนวณฮิสโทแกรมของทุกๆจุดอ้างอิงซึ่งเป็นส่วนย่อยของภาพ ทำให้บริบทรูปร่างมีคุณสมบัติการวิเคราะห์แบบเฉพาะที่ด้วย นั่นคือบริบทรูปร่างเป็นตัวบอกรูปร่างที่มีคุณสมบัติผสมผสานระหว่างคุณสมบัติครอบคลุมและคุณสมบัติเฉพาะที่ ทำให้บริบทรูปร่างเป็นคุณลักษณะมีความสามารถในการบอกรายละเอียดของภาพวัตถุได้ดียิ่งขึ้น

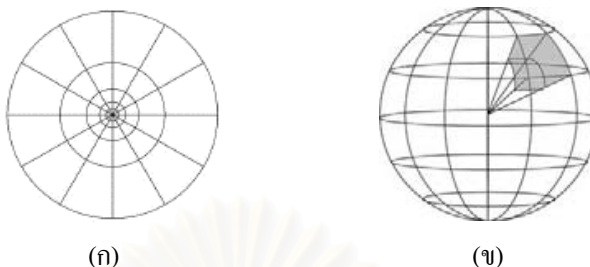
2.4.2 บริบทรูปร่างสำหรับภาพ 3 มิติ

จากหัวข้อที่ผ่านมา จะเห็นได้ว่า โครงสร้างแบบเดิมของบริบทรูปร่างทำการคำนวณฮิสโทแกรมจากดาวตารางแบบแบบลอการิทึมเชิงขั้ว โดยมีศูนย์กลางที่จุดอ้างอิงแต่ละจุดบนรูปร่างที่ต้องการในปริภูมิ 2 มิติ ดังนั้นเมื่อข้อมูลที่ใช้เป็นพื้นผิวแบบ 3 มิติ จึงต้องมีการปรับเปลี่ยนดาวตารางให้มีลักษณะเป็นดาวตารางแบบ 3 มิติ และฮิสโทแกรมที่ได้จะมีมิติของข้อมูลเพิ่มขึ้น ในงานวิจัยของ Andrea Fromel และคณะ [24] ได้นำเสนอการสร้างฮิสโทแกรมของจุดแบบ 3 มิติ โดยกำหนดให้มีการแบ่งดาวตารางในแนวรัศมี R จำนวน $J+1$ ช่อง ซึ่งจะได้ว่า $R = \{R_0, R_1, \dots, R_J\}$ แบ่งดาวตารางในทิศของมุมเงย (elevation, Θ) จำนวน $K+1$ ช่อง ซึ่งจะได้ว่า $\Theta = \{\Theta_0, \Theta_1, \dots, \Theta_J\}$ และ แบ่งดาวตารางในทิศมุมแอซิมัท (azimuth, Φ) จำนวน $L+1$ ช่อง ซึ่งจะได้ว่า $\Phi = \{\Phi_0, \Phi_1, \dots, \Phi_J\}$ ซึ่งจะได้ขอบของการแบ่งช่องในแนวรัศมี R ช่องที่ j ดังแสดงในสมการที่ (2.20)

$$R_j = \exp \left\{ \ln(r_{\min}) + \frac{j}{J} \ln \left(\frac{r_{\max}}{r_{\min}} \right) \right\} \quad (2.20)$$

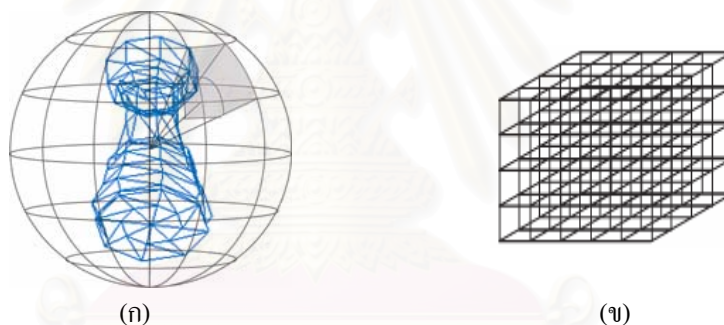
โดยที่ R_j คือ รัศมี R ช่องที่ j ดังแสดงใน
 r_{\min} คือ รัศมีที่มีค่าน้อยสุด หรือ R_0
 r_{\max} คือ รัศมีที่มีค่ามากที่สุด หรือ R_J

ส่วนการแบ่งดาวารางในทิศของมุมเงยและมุมแอสซิมาททำการแบ่งขนาดเท่ากัน โดยในรูปที่ 2.29 ทำการเปรียบเทียบโครงสร้างของดาวารางแบบ 2 มิติ และ โครงสร้างดาวารางแบบ 3 มิติ ซึ่งมีลักษณะของช่องดังส่วนที่มีสีเทาในรูปที่ 2.29 (ข)



รูปที่ 2.29 การเปรียบเทียบ โครงสร้างของ (ก) ดาวารางแบบ 2 มิติ (ข) ดาวารางแบบ 3 มิติ

โดยโครงสร้างดาวารางแบบ 3 มิติ มีลักษณะของช่องดังส่วนที่มีสีเทาในรูปที่ 2.30 (ก) โดยฮิสโทแกรมของจุดอ้างอิงจะมีลักษณะดังรูปที่ 2.30 (ข) ซึ่งแต่ละช่องของฮิสโทแกรมจะมีค่าของจำนวนจุดยอดที่ปรากฏในแต่ละช่องของดาวารางที่สัมพันธ์กัน



รูปที่ 2.30 (ก) โครงสร้างของบริบทรูปร่างแบบ 3 มิติ (ข) ฮิสโทแกรมของจุดยอดที่จุดอ้างอิง

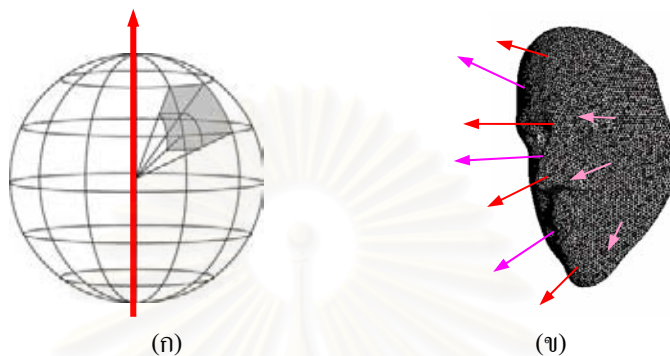
จากนั้นทำการถ่วงน้ำหนักให้กับค่าฮิสโทแกรมแต่ละช่องด้วยปริมาตรของช่อง เพื่อลดความแปรปรวนเนื่องจากขนาดของช่องที่ไม่เท่ากัน รวมทั้งมีการนอร์มอลไลซ์ด้วยความหนาแน่นของจุดยอดรอบจุดอ้างอิงซึ่งช่วยลดสัญญาณรบกวนของภาพ ดังแสดงในสมการที่ (2.22)

$$\omega(p_i) = \frac{1}{\rho_i \sqrt[3]{V(j,k,l)}} \quad (2.21)$$

โดยที่ $\omega(p_i)$ คือ ค่าถ่วงน้ำหนักของจุดอ้างอิงที่ p_i
 ρ_i คือ ความหนาแน่นของจุดรอบจุดอ้างอิง
 $V(j,k,l)$ คือ ปริมาตรของช่องที่ (j,k,l)

นอกจากนี้ จะเห็นได้ว่าภาพวัตถุในปริภูมิ 3 มิติมีทิศทางไม่แน่นอน ภาพวัตถุชิ้นเดียวกันอาจมีการวางตัวต่างกัน ส่งผลให้บริบทรูปร่างของวัตถุนั้นแตกต่างกันด้วย ดังนั้นในงานวิจัยของ Andrea Frome และคณะ [24] จึงมีการกำหนดทิศของขั้วทิศขั้ว (pole) ของดาวทรงกลม ดังแสดงในรูปที่ 2.31 (ก) ให้มีทิศทางชี้ไปในทิศของเส้นตั้งฉากที่ตำแหน่งจุดอ้างอิงดังรูปที่ 2.31 (ข) ซึ่งช่วยแก้ปัญหาการที่ภาพของวัตถุมีการเรียงตัวไม่เหมือนกันได้

ทิศของขั้ว (pole)



รูปที่ 2.31 (ก) ทิศขั้วของดาวทรงกลม (ข) ทิศขั้วของดาวทรงกลมเปลี่ยนแปลงตามเวกเตอร์ตั้งฉากที่จุดอ้างอิงใดๆ

จากที่ได้กล่าวมา บริบทรูปร่างเป็นตัวบอกรูปร่างที่มีคุณสมบัติผสมผสานทั้งการแสดงข้อมูลเฉพาะที่และข้อมูลแบบครอบคลุม เนื่องจากการแสดงแต่ละจุดยอดของตำแหน่งข้อมูลด้วยฮิสโทแกรมเชิงขั้วแบบลอการิทึมของจุดยอดทั้งหมดในภาพ เพื่อแสดงความสัมพันธ์ของการกระจายตัวของข้อมูลทั้งหมด จากนั้นในกระบวนการแยกแยะความแตกต่างของภาพจะทำการหาฟังก์ชันต้นทุนจากการหาค่าไค-สแควร์ระหว่างฮิสโทแกรมของแต่ละจุดยอดเพื่อบอกความแตกต่างระหว่างฮิสโทแกรมของจุด โดยถ้าค่าไค-สแควร์ที่ได้มีค่าแตกต่างกันมาก แสดงว่า ความสัมพันธ์ของจุดคู่หนึ่งๆมีความแตกต่างกันสูง

2.4.3 การทดสอบไค-สแควร์ (chi-squared test)

การทดสอบสมมติฐาน (hypothesis testing) เป็นกระบวนการที่ใช้ทดสอบผลการทดลองที่เกิดขึ้นว่า เป็นผลที่มีนัยยะสำคัญหรือไม่ (significant or accidental) ตัวอย่างเช่น ในการทดลองที่ผลการทดลองที่จะเกิดขึ้นได้นั้นมีรูปแบบการกระจายตัวที่รู้ค่าอยู่รูปแบบหนึ่ง หรือเรียกว่า สมมติฐานว่าง (null hypothesis) เมื่อทำการทดลองแล้วผลที่ได้มีความน่าเชื่อถือได้มากน้อยเพียงใด จะขึ้นอยู่กับว่า ผลการทดลองที่เกิดขึ้นมีความสัมพันธ์ใกล้เคียงกับรูปแบบการกระจายตัวที่คาดการณ์ไว้เบื้องต้นมากน้อยเพียงไร

การทดสอบไค-สแควร์ [25] เป็นหนึ่งในกระบวนการทดสอบสมมติฐาน กับผลการทดลองที่มีลักษณะเป็นค่าไม่ต่อเนื่อง (discrete) จำนวน I ค่า ค่าไค-สแควร์หาได้จากสมการที่ (2.22)

$$\chi^2 = \sum_{i=1}^I \frac{(n_{i_ex} - n_{i_null})^2}{n_{i_null}} \quad (2.22)$$

โดยที่ χ^2 คือ ค่าไค-สแควร์
 n_{i_null} คือ ความเป็นไปได้ที่ของรูปแบบผลการทดลองที่เกิดขึ้น
 n_{i_ex} คือ รูปแบบของผลการทดลองที่เกิดขึ้นจริง

จากสมการที่ (2.22) จะเห็นได้ว่า ถ้าค่าที่ได้มีค่าน้อยแสดงว่าผลการทดลอง n_{i_ex} มีค่าสอดคล้องกับรูปแบบการกระจายตัวที่รู้ค่า n_{i_null} แต่ถ้าค่า χ^2 ที่ได้มีค่ามากแสดงว่าผลการทดลอง n_{i_ex} มีค่าแตกต่างจากรูปแบบการกระจายตัวที่รู้ค่า n_{i_null} อยู่มาก

เนื่องจากบริบทรูปร่าง มีคุณสมบัติเป็นฮิสโทแกรมของจุดภาพซึ่งแสดงการกระจายตัวของจุดยอด ดังนั้นจึงมีการนำการทดสอบไค-สแควร์มาทำการเปรียบเทียบความแตกต่างของบริบทรูปร่างของภาพวัตถุ โดยการใช้ค่าไค-สแควร์ในการหาฟังก์ชันต้นทุน (cost function) ระหว่างรูปร่างที่ทดสอบ ซึ่งในกรณีนี้ค่า χ^2 จะเป็นค่าที่แสดงถึงความแตกต่างระหว่างฮิสโทแกรมของจุดยอดแต่ละจุดระหว่างรูป 2 รูปที่นำมาทดสอบเปรียบเทียบความแตกต่างกัน เมื่อนำรูปทดสอบ $P = \{p_i\}$ และ $Q = \{q_j\}$ มาผ่านกระบวนการหาบริบทรูปร่างจะได้ค่าฮิสโทแกรมช่องที่ k ของจุดอ้างอิง p_i และ q_j มีค่าเป็น $h_i(k)$ และ $h_j(k)$ ตามลำดับ เมื่อ K คือจำนวนช่องที่มีทั้งหมด ค่าต้นทุนระหว่าง p_i และ q_j จะคำนวณได้จากสมการที่ (2.23)

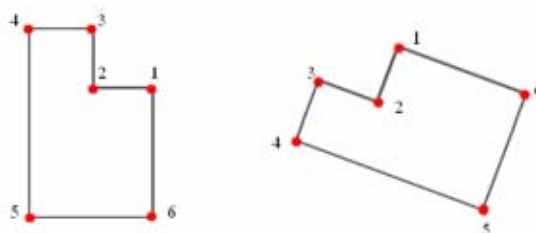
$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (2.23)$$

โดยที่ C_{ij} คือ ค่าต้นทุนแสดงความแตกต่างของฮิสโทแกรมระหว่างจุดอ้างอิงที่ i ของภาพ P และ จุดอ้างอิงที่ j ของภาพ Q

ในการแยกแยะความแตกต่างระหว่างบริบทรูปร่างของภาพวัตถุ 2 ภาพ จะทำการหาเมตริกซ์ต้นทุน $C = \{C_{ij}\}$ ระหว่างทุกคู่จุดอ้างอิงของภาพทั้ง 2 ภาพ แล้วหาค่าความแตกต่างของภาพวัตถุจากค่าต่ำสุดของผลรวมค่าต้นทุนระหว่างคู่จุดที่มีความสัมพันธ์กัน (correspondence point) ซึ่งขั้นตอนนี้หาได้จากกระบวนการหาค่าที่เหมาะสมที่สุด (optimization) โดยในงานวิจัยของ Belongi และคณะ [6] ได้นำเสนอวิธี Hungarian Method [28]

2.4.4 การหาจุดที่มีความสัมพันธ์กันด้วยการหาค่าที่เหมาะสมที่สุด

กระบวนการแยกแยะความแตกต่างระหว่างภาพวัตถุ 2 ภาพนั้นจะทำการคำนวณหาจุดที่มีความสัมพันธ์กันดังแสดงในรูปที่ 2.32 ซึ่งเป็นจุดที่มีผลรวมของค่าต้นทุนต่ำสุด โดยกระบวนการหาจุดที่มีความสัมพันธ์กันนั้นจะคำนวณจากกระบวนการหาค่าที่เหมาะสมที่สุดด้วยวิธี Hungarian Method ซึ่งทำการคำนวณรอบเพื่อคำนวณค่าต้นทุนรวมจากการจับคู่จุดในแต่ละแบบ เพื่อหาจุดที่เหมาะสมที่สุดซึ่งมีค่าต้นทุนรวมต่ำสุด โดยกระบวนการหาค่าที่เหมาะสมที่สุดด้วยวิธี Hungarian Method มีลักษณะการคำนวณแบบวนรอบโดยค่าที่ต้องการหาได้จากการคำนวณ $O(N^3)$ รอบ เมื่อ N คือ จำนวนข้อมูลที่มี



รูปที่ 2.32 จุดที่มีความสัมพันธ์กันระหว่างภาพวัตถุ 2 ภาพแสดงด้วยลำดับเลขลำดับเดียวกัน

จากที่ได้กล่าวมา พบว่าในปัจจุบันการวิจัยด้านกระบวนการรู้จำใบหน้าในแบบ 3 มิติเริ่มเข้ามามีบทบาทมากยิ่งขึ้น เนื่องจากการพัฒนาทางด้านเทคโนโลยีด้านภาพ 3 มิติมากขึ้น แต่ถึงกระนั้นกระบวนการรู้จำใบหน้าแบบ 3 มิติเองยังมีปัญหาในการใช้งานจากปัจจัยภายนอก เช่น ผลกระทบของภาพการแสดงออกของใบหน้า เป็นต้น นอกจากนี้กระบวนการรู้จำใบหน้าแบบ 3 มิติควรมีคุณสมบัติในการแยกแยะภาพโดยใช้กลุ่มทดลองภาพน้อย เนื่องจากการรับภาพ 3 มิติแต่ละภาพต้องใช้เวลาาน อีกทั้งควรมีกระบวนการคำนวณที่ไม่ซับซ้อน เนื่องจากภาพ 3 มิติเป็นข้อมูลที่มีขนาดใหญ่ ถ้าการคำนวณที่ใช้ซับซ้อนมากจะใช้ทรัพยากรในการคำนวณสูง ซึ่งในบทที่ 3 ได้นำเสนอแนวทางการใช้บริบทรูปร่างในกระบวนการรู้จำใบหน้าแบบ 3 มิติ ซึ่งบริบทรูปร่างนั้น มีคุณสมบัติเด่นหลายประการ คือ มีการคำนวณไม่ซับซ้อน สามารถคำนวณภาพที่มีจำนวนจุดยอดไม่เท่ากันได้ ต้องการภาพกลุ่มทดลองเพื่อสร้างการเรียนรู้ให้ระบบน้อย และทนทานต่อการผลกระทบจากปัจจัยภายนอกต่างๆได้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

จากที่ได้กล่าวไปในบทที่ 2 กระบวนการที่ Bronstein และคณะ[4] ได้นำเสนอการดึงข้อมูลของระยะทางจีโอเดสิกระหว่างจุดยอดนั้นมีการคำนวณที่ซับซ้อน อีกทั้งถ้าจำนวนจุดยอดของภาพมีจำนวนมาก จะยิ่งทำให้การคำนวณเกิดการล่าช้าได้ ดังนั้นจึงมีแนวทางที่จะหาคุณลักษณะของภาพที่มีคุณสมบัติไม่แปรเปลี่ยนเมื่อภาพเกิดการแปลงผันจากการแสดงออกของบุคคลแทนการคำนวณหาระยะทางจีโอเดสิกระหว่างจุดยอด ซึ่งในวิทยานิพนธ์ฉบับนี้ได้นำเสนอการใช้ความแตกต่างของบริบทรูปร่างมาแทนขั้นตอนดังกล่าว ซึ่งการนำบริบทรูปร่างมาใช้กับภาพหน้า ได้มีการปรับเปลี่ยนบางขั้นตอนที่มีความซับซ้อนเพื่อให้มีการคำนวณมีประสิทธิภาพมากยิ่งขึ้น โดยประกอบด้วย การปรับปรุงการหาขั้วของตาตารางในทิศตั้งฉากของพื้นผิวที่จุดอ้างอิงด้วยการปรับแนวโดยการวิเคราะห์องค์ประกอบหลัก การปรับขนาดภาพเพื่อหลีกเลี่ยงขั้นตอนการหาค่าถ่วงน้ำหนักที่เปลี่ยนแปลงตามความหนาแน่นของจุดภาพและขนาดของช่อง และการกำหนดจุดยอดบนตำแหน่งองค์ประกอบสำคัญบนใบหน้าเพื่อลดขั้นตอนการหาความแตกต่างของภาพจากขั้นตอนวิธีการหาค่าที่เหมาะสมที่สุด

จากนั้นเป็นการนำเสนอขั้นตอนวิธีในหาคุณลักษณะของภาพใบหน้า 2 แบบ คือ การนำบริบทรูปร่างไปผสมผสานกับแนวคิดของงานวิจัยด้านกระบวนการรู้จำใบหน้าที่น่าเสนอโดย Broanstein และคณะ [4] ซึ่งนำเสนอเรื่องกระบวนการรู้จำใบหน้าแบบไม่แปรผันตามการแสดงออกของบุคคล โดยใช้เมทริกซ์ของระยะทางจีโอเดสิกระหว่างจุดยอดบนภาพใบหน้าเป็นคุณลักษณะที่ไม่แปรเปลี่ยนเมื่อภาพใบหน้ามีการแสดงออกแบบต่างๆ ในวิทยานิพนธ์ฉบับนี้ได้นำเสนอกระบวนการที่มีลักษณะคล้ายคลึงกับงานวิจัยดังกล่าว แต่ได้ปรับเปลี่ยนการหาคุณลักษณะของใบหน้าจากการหาเมทริกซ์ของระยะทางจีโอเดสิกระหว่างจุดยอด เป็นการหาเมทริกซ์ความคล้ายของบริบทรูปร่าง โดยใช้ความแตกต่างของฮิสโทแกรมของจุดยอดบนภาพที่คำนวณจากค่าไอ-สแควร์ระหว่างฮิสโทแกรม แทนการคำนวณหาระยะทางจีโอเดสิก และการใช้ฮิสโทแกรมของจุดยอดระหว่างภาพใบหน้ามาเป็นคุณลักษณะในการคำนวณความแตกต่างของภาพโดยตรง

3.1. การปรับปรุงบริบทรูปร่างสำหรับภาพ 3 มิติเพื่อใช้กับภาพใบหน้า

จากหัวข้อที่ 2.4 จะเห็นได้ว่า บริบทรูปร่างเกิดจากการคำนวณฮิสโทแกรมจากตาตารางแบบลอการิทึมเชิงขั้วโดยมีศูนย์กลางที่จุดอ้างอิงแต่ละจุดบนรูปร่างที่ต้องการในปริภูมิ 2 มิติ [6] และเป็นฮิสโทแกรมจากตาตารางแบบแบบลอการิทึมทรงกลมสำหรับภาพ 3 มิติ [24] ซึ่งจากการคำนวณฮิสโทแกรมจากตาตารางแบบลอการิทึมทรงกลมนั้น มีขั้นตอนการคำนวณที่ซับซ้อนและใช้เวลาในการประมวลผลนานอยู่ 3 ขั้นตอน คือ การหาขั้วของตาตารางในทิศตั้งฉากของพื้นผิวที่จุดอ้างอิง การหาค่าถ่วงน้ำหนักที่เปลี่ยนแปลงตามความหนาแน่นของจุดภาพและขนาดของช่อง และการหาความแตกต่างของภาพจากขั้นตอนวิธีการหาค่าที่เหมาะสมที่สุด ในหัวข้อนี้เรานำเสนอการคำนวณหาบริบทรูปร่างสำหรับภาพ 3 มิติเพื่อใช้กับภาพใบหน้าซึ่งปรับปรุงขั้นตอนที่ซับซ้อนให้มีการคำนวณรวดเร็วยิ่งขึ้น เนื่องจากในการใช้งานบริบทรูปร่างกับภาพใบหน้า นั้น มีข้อได้เปรียบการใช้งานโดยทั่วไป คือ แม้ว่าภาพใบหน้าจะเป็นภาพของบุคคลต่างบุคคลกัน แต่ลักษณะโดยคร่าวของภาพทุกภาพจะมีโครงสร้างใกล้เคียงกัน ซึ่งทำให้สามารถแทนที่ขั้นตอนที่มีการคำนวณยุ่งยากด้วยกระบวนการที่ไม่ซับซ้อนได้

3.1.1 การปรับปรุงการหาข้อของตาตารางในทิศตั้งฉากของพื้นผิวที่จุดอ้างอิงด้วยการปรับแนวโดยการวิเคราะห์ห้อยค์ประกอบหลัก (PCA alignment)

ในการหาบริบทรูปร่างต้องทำการหาทิศของข้อของตาตารางเนื่องจากข้อมูลภาพในปริภูมิ 3 มิติจะมีทิศทางการวางตัวของวัตถุที่ไม่แน่นอน ดังนั้นเพื่อให้ฮิสโทแกรมของจุดอ้างอิงจุดเดียวกันของวัตถุเดียวกันที่มีการวางตัวในทิศที่แตกต่างกันมีค่าเท่ากัน จึงต้องกำหนดทิศตั้งฉากกับพื้นผิวนั้นเป็นทิศอ้างอิง แต่ในขั้นตอนนี้อาจสามารถทำการหาทิศที่แน่นอนของวัตถุได้ก่อน ก็จะสามารถกำหนดทิศทางที่แน่นอนของตาตารางได้ ซึ่งในหัวข้อนี้นำเสนอการใช้การวิเคราะห์ห้อยค์ประกอบหลักเพื่อทำการปรับแนวภาพให้มีทิศทางที่แน่นอนก่อนทำการหาบริบทรูปร่าง

แนวคิดของการวิเคราะห์ห้อยค์ประกอบหลักสามารถนำมาปรับใช้เพื่อให้ภาพมีการวางตัวในทิศที่ต้องการได้ โดยการกำหนดให้ จุดยอดของภาพวัตถุคือจุดข้อมูลที่มีการกระจายตัวในปริภูมิ 3 มิติ ทิศทางที่ข้อมูลมีความแปรปรวนสูงหรือทิศทางของเวกเตอร์ลักษณะเฉพาะจะมีการวางตัวไปในทิศทางที่ข้อมูลมีการกระจายตัวสูง ซึ่งเป็นทิศที่เป็นแกนของภาพ นั่นคือ ในวัตถุที่มีรูปร่างคล้ายคลึงกัน ลักษณะการกระจายตัวของจุดยอดจะมีลักษณะไม่แตกต่างกันมากนัก แกนของภาพจะมีทิศทางเดียวกัน

การปรับแนวภาพใบหน้าโดยการวิเคราะห์ห้อยค์ประกอบหลัก เริ่มจากการหาทิศทางที่ข้อมูลมีการกระจายตัวสูงสุด โดยเลือกทิศทางที่เวกเตอร์แกนหลักมีค่าแกนหลักสูงสุด 3 ค่า เพื่อหาทิศของความยาวของใบหน้าหรือทิศทางจากหน้าผากถึงคางซึ่งโดยทั่วไปจะมีการกระจายตัวของจุดยอดสูงสุด ทิศของความกว้างของใบหน้าหรือทิศทางระหว่างโครงหน้าซ้าย ขวา ซึ่งมีการกระจายตัวของจุดยอดรองลงมา และทิศของความหนาของใบหน้า หรือทิศทางจากจมูกถึงโครงหน้าส่วนหลัง ซึ่งมีการกระจายตัวของจุดยอดน้อยสุด ดังที่แสดงในรูปที่ 3.1 จะเห็นได้ว่า แม้ใบหน้าแต่ละหน้าจะมีลักษณะส่วนย่อยแตกต่างกันไป แต่ลักษณะโดยรวมจะคล้ายคลึงกัน ดังนั้นเมื่อทำการฉายภาพใบหน้าไปยังเวกเตอร์แกนลักษณะเฉพาะหรือแกนภาพที่หาได้ของแต่ละภาพ จะได้ภาพใบหน้าที่ถูกปรับแนวแล้วหันไปในทิศทางเดียวกัน



รูปที่ 3.1 แกนของภาพใบหน้าที่จุดยอดมีการกระจายตัวมากไปน้อย 3 แกน ตามลำดับ

จากที่ได้กล่าวไว้ในหัวข้อย่อย 1 ของหัวข้อ 2.1.1.1 กระบวนการวิเคราะห์ห้อยค์ประกอบหลัก [8] เป็นขั้นตอนวิธีที่ใช้สำหรับกระบวนการรู้จำใบหน้าโดยหาปริภูมิภาพที่ข้อมูลมีค่าความแปรปรวนร่วมสูง หรือมีการ

ลักษณะกระจายตัวของกลุ่มข้อมูลสูงสุด โดยจะแสดงผลของกลุ่มข้อมูลด้วยสัมประสิทธิ์ที่แสดงปริมาณข้อมูลบนปริภูมิภาพนั้นๆ โดยเรียกปริภูมิภาพที่มีการกระจายตัวข้อมูลสูงสุดว่า เวกเตอร์ลักษณะเฉพาะและค่าสัมประสิทธิ์ของข้อมูลบนปริภูมิภาพนี้ว่า ค่าลักษณะเฉพาะ ซึ่งจะนำไปใช้หาคุณลักษณะของข้อมูล เพื่อใช้แสดงเฉพาะรายละเอียดที่สำคัญของข้อมูลแทนการใช้ข้อมูลทั้งหมดโดยตรง

เมื่อแทนตำแหน่งของจุดยอดจำนวนทั้งหมด N จุดบนภาพ 3 มิติด้วยเมตริกซ์ขนาด $3 \times N$ มิติ $\mathbf{S} = [\mathbf{X}^T \ \mathbf{Y}^T \ \mathbf{Z}^T]$ เมื่อ \mathbf{X} , \mathbf{Y} และ \mathbf{Z} คือ พิกัดบอกตำแหน่งจุดยอดตามแกน x , y และ z บนพิกัดคาร์ทีเซียน (cartesian coordinate) ตามลำดับ แกนที่มีการกระจายตัวสูงสุดของข้อมูลหาได้จากเมตริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix) ของข้อมูลดังสมการที่ (3. 1)

$$\mathbf{G} = (\mathbf{S} - \boldsymbol{\mu}_s)^T (\mathbf{S} - \boldsymbol{\mu}_s) \quad (3. 1)$$

เมื่อ \mathbf{S} คือ เมตริกซ์ความแปรปรวนร่วมเกี่ยว
 \mathbf{V} คือ เมตริกซ์จุดยอด
 $\boldsymbol{\mu}_s$ ค่าเฉลี่ยของตำแหน่งจุดยอดของภาพดังแสดงในสมการที่ (3. 2)

$$\boldsymbol{\mu}_s = \left(\frac{1}{N} \sum_{i=1}^N S_i \right) \mathbf{1} \quad (3. 2)$$

เมื่อ S_i คือ จุดยอดที่ i ของภาพ
 $\mathbf{1}$ คือ เวกเตอร์ขนาด $1 \times N$ ซึ่งสมาชิกทุกตัวมีค่าเป็น 1

โดยแทนที่ได้จากกระบวนการแยกองค์ประกอบหลักนี้หาจากระบวนการแยกลักษณะเฉพาะหาจากการแทนค่า \mathbf{G} ไปในสมการที่ (2.4) ดังแสดงในสมการที่ (3. 3)

$$\mathbf{G}L_j = \lambda_j L_j \quad ; j = 1, 2, \dots, d \quad (3. 3)$$

เมื่อ L_j คือ เวกเตอร์ลักษณะเฉพาะ
 λ_j คือ ค่าลักษณะเฉพาะ

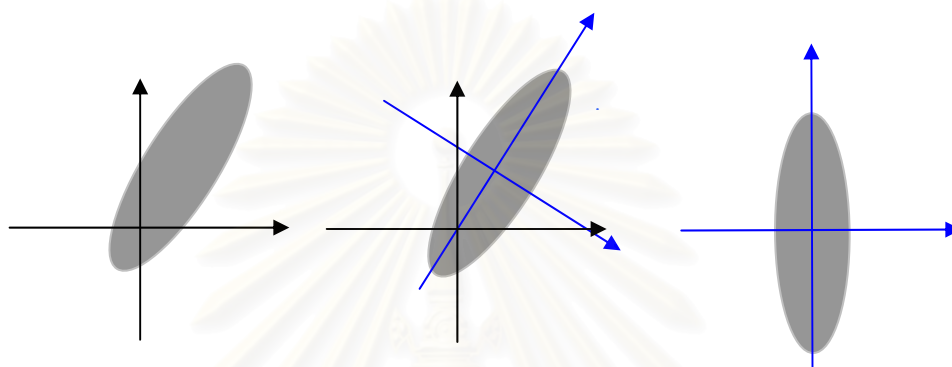
โดยเลือกคู่เวกเตอร์ลักษณะเฉพาะ และ ค่าลักษณะเฉพาะ ที่มีค่าสูงสุด d ค่า มาใช้หาคุณลักษณะของข้อมูล B_j ดังแสดงในสมการที่ (3. 4)

$$B_j = \mathbf{S}L_j \quad ; j = 1, 2, \dots, d \quad (3. 4)$$

จากขั้นตอนดังกล่าว จะได้หาแกนของข้อมูลที่มีทิศทางไปยังทิศที่ข้อมูลมีการกระจายตัวสูงสุด และค่าสัมประสิทธิ์แสดงปริมาณข้อมูลบนแกนนั้น โดยนำข้อมูลที่ได้มาทำการฉายภาพจะได้ภาพใหม่ของวัตถุบนแกนลักษณะเฉพาะที่มีค่าลักษณะเฉพาะสูงสุด ดังแสดงในสมการที่ (3. 5)

$$\tilde{\mathbf{S}} = \mathbf{B}\mathbf{L}^T \quad ; j = 1, 2, \dots, d \quad (3.5)$$

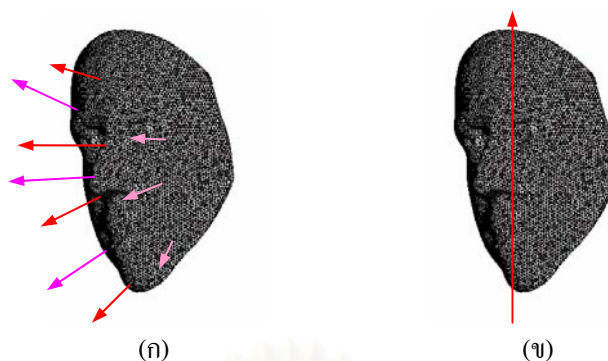
เมื่อ	$\tilde{\mathbf{S}}$	คือ ภาพวัตถุที่ผ่านการปรับแนว
	\mathbf{B}	คือ เมทริกซ์คุณสมบัติลักษณะ โดยมีค่าเป็น $[B_1, \dots, B_j]$
	\mathbf{L}	คือ เมทริกซ์ลักษณะเฉพาะ โดยที่ค่าเป็น $[L_1, \dots, L_j]$



รูปที่ 3.2 แสดงแนวทางการปรับแนวโดยใช้การวิเคราะห์องค์ประกอบหลักในภาพ 2 มิติ
(ก) การกระจายตัวของข้อมูลในระนาบ 2 มิติ (ข) การหาแกนที่มีการกระจายตัวของข้อมูลสูงสุด
หรือ แกนลักษณะเฉพาะ (ค) การแสดงผลข้อมูลบนแกนลักษณะเฉพาะ

เพื่อให้ง่ายต่อการเข้าใจ เราแสดงแนวทางการปรับแนวโดยใช้การวิเคราะห์องค์ประกอบหลักในภาพ 2 มิติ โดยในรูปที่ 3.2 (ก) แสดงลักษณะการกระจายตัวของกลุ่มข้อมูลบนแกน (x, y) ในปริภูมิใดๆ แล้วทำการวิเคราะห์องค์ประกอบหลักของภาพโดยหาแกนที่ข้อมูล มีการกระจายตัวสูงสุดหรือเวกเตอร์ลักษณะเฉพาะซึ่งแสดงได้ด้วยแกน (x', y') รูปที่ 3.2 (ข) โดย y' เป็นเวกเตอร์ลักษณะเฉพาะที่มีค่าลักษณะเฉพาะสูงสุดหรือ มีการกระจายตัวมากที่สุด ส่วน x' เป็นเวกเตอร์ลักษณะเฉพาะที่มีค่าลักษณะเฉพาะน้อยกว่าคือมีการกระจายตัวของข้อมูลน้อยกว่าในแกน y' นั้นเอง จากนั้นนำข้อมูลที่ได้อามาสร้างภาพใหม่ของวัตถุบนแกนลักษณะเฉพาะที่มีค่าลักษณะเฉพาะสูงสุดซึ่งแสดงดังรูปที่ 3.2 (ค)

การปรับแนวภาพโดยใช้การวิเคราะห์องค์ประกอบหลัก สามารถแก้ปัญหาการเรียงตัวของข้อมูลที่มีทิศทางไม่แน่นอนได้ หรืออาจกล่าวได้ว่า ภาพถูกสร้างบนแกนเวกเตอร์ลักษณะเฉพาะจะมีลักษณะไม่แปรเปลี่ยนต่อการแปลงของวัตถุแข็งเกร็งที่เกิดขึ้น เนื่องจากภาพใดๆที่เกิดการเลื่อนตำแหน่งและการหมุนจะถูกปรับแนวให้กลับมาอยู่ในทิศทางที่ต้องการ ดังนั้นในการคำนวณหาบริบทรูปร่างของภาพใบหน้าจึงมีการหาทิศทางที่แน่นอนของวัตถุก่อน และทำการวางแกนหลักของภาพวัตถุให้มีทิศทางที่ต้องการ โดยกำหนดทิศชี้ของตาตารางให้มีทิศทางคงที่ ซึ่งในรูปที่ 3.3 แสดงการเปรียบเทียบระหว่างการกำหนดทิศชี้ของตาตารางที่เปลี่ยนแปลงตามเวกเตอร์ตั้งฉากที่จุดอ้างอิง และแบบคงที่คงที่ในทิศของเวกเตอร์แกนหลัก



รูปที่ 3.3 การเปรียบเทียบระหว่าง (ก) ทิศชี้ของดาวหางเปลี่ยนแปลงตามเวกเตอร์ตั้งฉากที่จุดอ้างอิง
(ข) ทิศชี้ของดาวหางคงที่ในทิศของเวกเตอร์แกนหลัก

จะเห็นได้ว่า แม้ว่าการปรับแนวโดยการวิเคราะห์ห้วงค์ประกอบหลักจะช่วยแก้ไขการเกิดการแปลงของวัตถุแข็งเกร็งเช่น การเลื่อนที่ หรือการหมุนได้ แต่การสร้างภาพใหม่ของวัตถุโดยวิธีวิเคราะห์ห้วงค์ประกอบหลักไม่สามารถแก้ปัญหาจากการแปลงของวัตถุไม่แข็งเกร็งได้ เนื่องจากวิเคราะห์เฉพาะทิศทางการวางตัวของวัตถุในภาพไม่ได้รวมถึงการเปลี่ยนแปลงระยะทางระหว่างจุดยอดจุดต่างๆ หรือมุมระหว่างแผ่นผิวแต่ละแผ่น ซึ่งมีผลให้รูปร่างของวัตถุเปลี่ยนแปลงไป ดังนั้นการแปลงแบบวัตถุไม่แข็งเกร็งต้องใช้การวิเคราะห์ที่มีรายละเอียดย่อยของข้อมูลวัตถุ โดยในที่นี้แนะนำการใช้บริบทรูปร่างดึงข้อมูลส่วนย่อยของภาพออกมา โดยบริบทรูปร่างในแบบ 2 มิติที่ได้นำเสนอไปในบทที่ 2 จะมีการปรับปรุงเพิ่มเติมเพื่อให้ใช้กับภาพ 3 มิติได้

3.1.2 การปรับขนาดภาพเพื่อหลีกเลี่ยงขั้นตอนการหาค่าถ่วงน้ำหนักที่เปลี่ยนแปลงตามความหนาแน่นของจุดภาพและขนาดของช่อง

จากสมการ (2.21) จะเห็นว่า มีการคำนวณค่าถ่วงน้ำหนักเพื่อลดความแปรปรวนของจากขนาดของช่องที่ไม่เท่ากัน และการนอร์มอลไลซ์ด้วยความหนาแน่นจุดยอดรอบจุดอ้างอิง ซึ่งถ้ามีการปรับขนาดของภาพให้ใกล้เคียงกันก่อนบริบทรูปร่างที่คำนวณได้จากภาพที่แตกต่างกันจะมีความแปรปรวนอันเกิดจากขนาดของช่องลดลงโดยคำนวณตำแหน่งจุดยอดหลังการปรับขนาดภาพตามสมการที่ (3. 6)

$$(\tilde{x}, \tilde{y}, \tilde{z}) = \left(\frac{x}{\|\max(x) - \min(x)\|}, \frac{y}{\|\max(y) - \min(y)\|}, \frac{z}{\|\max(z) - \min(z)\|} \right) \quad (3. 6)$$

โดยที่ $(\tilde{x}, \tilde{y}, \tilde{z})$ คือ ตำแหน่งของจุดยอดหลังการปรับขนาด
 (x, y, z) คือ ตำแหน่งของจุดยอดก่อนการปรับขนาด

โดยการคำนวณหาปริมาตรรูปร่างยังคงลักษณะการคำนวณแบบเดิมไว้ โดยทำการแปลงค่าพิกัดของข้อมูล ซึ่งเป็นตำแหน่งอยู่ในปริภูมิคาร์ทีเซียน 3 มิติ (x, y, z) ไปยังตารางลอการิทึมทรงกลม (logarithm-spherical grid) (ρ, ϕ, θ) ตามสมการที่ (3. 7)-(3. 9)

$$\rho = \log \sqrt{x^2 + y^2 + z^2} \quad (3. 7)$$

โดยที่ ρ คือ ระยะระหว่างจุดยอดกับจุดยอดอ้างอิง p_i

$$\phi = \tan^{-1} \left(\frac{\sqrt{x^2 + y^2}}{z} \right) \quad (3. 8)$$

โดยที่ ϕ คือ มุมระหว่างจุดยอดทำกับแกน z

$$\theta = \tan^{-1} \left(\frac{y}{x} \right) \quad (3. 9)$$

โดยที่ θ คือ มุมระหว่างจุดยอดทำกับแกน x

ความแปรปรวนที่เกิดจากความแตกต่างของจำนวนจุดภาพจะทำการนอร์มอลไลซ์ด้วยจำนวนจุดยอดทั้งหมดของภาพ ทำให้ฮิสโทแกรมที่ได้จากการคำนวณตามสมการที่ (2.17) สามารถเขียนใหม่ได้ดังสมการที่ (3. 10)

$$\tilde{h}_i(k) = \frac{h_i(k)}{n} \quad (3. 10)$$

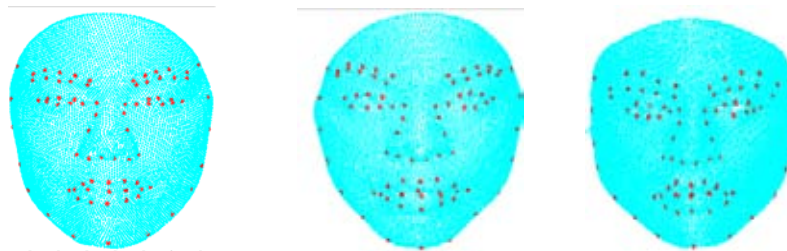
โดยที่ $\tilde{h}_i(k)$ คือ ค่าฮิสโทแกรม ของจุดยอดผ่านการนอร์มอลไลซ์

$h_i(k)$ คือ ค่าฮิสโทแกรม ของจุดยอดที่ i ในอาณาบริเวณช่องที่ k

n คือ จำนวนจุดยอดทั้งหมดในภาพ

3.1.3 การกำหนดจุดอ้างอิงที่ตำแหน่งองค์ประกอบสำคัญบนใบหน้า

ในการหาจุดยอดที่มีความสัมพันธ์กันด้วยการหาค่าเหมาะที่สุดนั้น จะต้องทำการคำนวณแบบวนรอบ ซึ่งใช้เวลาในการคำนวณสูง ในหัวข้อนี้นำเสนอแนวทางเพื่อลดการคำนวณในขั้นตอนนี้ลงได้ โดยการกำหนดจุดยอดที่เป็นจุดอ้างอิงในการคำนวณฮิสโทแกรมไว้ที่ตำแหน่งองค์ประกอบสำคัญบนใบหน้า ซึ่งประกอบด้วย ตา คิ้ว จมูก ปาก และโครงหน้า เนื่องจากเมื่อภาพที่ใช้ในการคำนวณเป็นภาพใบหน้าทั้งหมด ดังนั้นทุกภาพจึงมีตำแหน่งองค์ประกอบสำคัญที่คล้ายคลึงกันดังแสดงในรูปที่ 3.4 ดังนั้นจุดยอดเหล่านี้จึงเป็นจุดที่มีความสัมพันธ์กันระหว่างภาพอยู่แล้ว จึงไม่จำเป็นต้องคำนวณหาจุดที่มีความสัมพันธ์กันด้วยการหาค่าเหมาะที่สุดได้



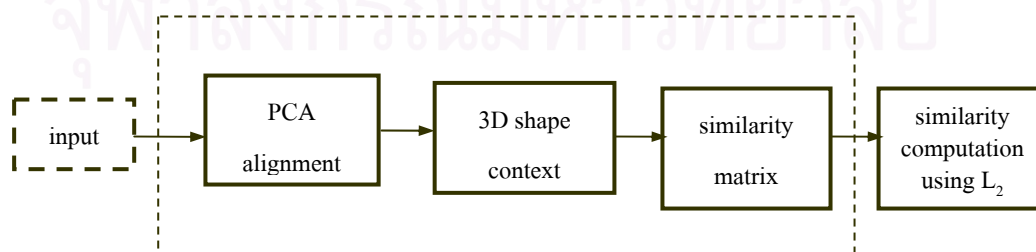
รูปที่ 3.4 ภาพบุคคลคนละคน มีตำแหน่งประกอบสำคัญบนใบหน้าตำแหน่งเดียวกันเสมอ

3.2. การแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ

Broanstein และคณะ [4] ได้นำเสนอกระบวนการรู้จำใบหน้าแบบไม่แปรผันตามการแสดงออกของบุคคล โดยสร้างภาพแบบบัญญัติของใบหน้าจากการหาระยะทางจีโอเดสิกระหว่างจุดบนพื้นผิวด้วยกระบวนการพาสท์มาร์ชชิง ซึ่งขั้นตอนวิธีนี้สามารถลดผลของกระทบจากการแสดงออกของบุคคลในกระบวนการรู้จำภาพใบหน้าแบบ 3 มิติได้ ซึ่งมีกระบวนการคำนวณมีความซับซ้อนสูง ดังนั้นในหัวข้อนี้ เรานำเสนอขั้นตอนวิธีที่ผสมผสานระหว่างการใช้บริบทรูปร่างกับขั้นตอนวิธีแบบเดิม โดยทำการหาฮิสโทแกรมของจุดภาพด้วยกระบวนการคำนวณบริบทรูปร่างเพื่อหาเมตริกซ์ความคล้ายคลึงของฮิสโทแกรมของจุดภาพมาเป็นคุณลักษณะของใบหน้าแทนเมตริกซ์ของระยะทางจีโอเดสิก

เมื่อภาพใบหน้ามีการแปรผันจากการแสดงออก ระยะทางจีโอเดสิกระหว่างจุดภาพ นำเสนอแนวทางในการหาสมบัติภายในของพื้นผิว 3 มิติโดยการนำบริบทรูปร่างมาใช้ร่วมกับการวิเคราะห์องค์ประกอบหลักในการคำนวณแทน ซึ่งบริบทรูปร่างนั้นมีการพัฒนาคุณสมบัติเพิ่มจากงานวิจัยเดิม [6] เนื่องจากโครงสร้างการทำงานของบริบทรูปร่างแบบเดิมเป็นการทำงานบนรูปร่างในปริภูมิ 2 มิติเท่านั้น โดยข้อดีของบริบทรูปร่างคือมีความซับซ้อนในการคำนวณน้อยกว่าการหาระยะทางจีโอเดสิกระหว่างจุดบนพื้นผิว และมีการวิเคราะห์ภาพด้วยการคำนวณฮิสโทแกรมกับจุดยอดทุกจุดซึ่งทำให้บริบทรูปร่างมีทั้งคุณสมบัติครอบคลุมและคุณสมบัติเฉพาะที่ ซึ่งวิธีการที่นำเสนอนี้ทำให้ระบบสามารถรองรับผลจากการแปรผันทางการแสดงออกในภาพใบหน้าแบบ 3 มิติได้อีกทั้งนำไปใช้เพื่อเพิ่มความแม่นยำให้กับกระบวนการรู้จำใบหน้าได้ต่อไป

ในการสร้างภาพแบบบัญญัติจากการคำนวณหาระยะทางจีโอเดสิกระหว่างจุดบนพื้นผิวมีความซับซ้อนสูง ดังนั้นจึงได้นำเสนอแนวทางการนำบริบทรูปร่างแบบ 3 มิติที่คำนวณได้ง่าย และมีความยืดหยุ่นสูงกว่า ร่วมกับการวิเคราะห์องค์ประกอบหลักในการหาเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพที่มีความทนทานต่อการแปรผันซึ่งเกิดจากการแสดงออกของบุคคลในภาพ โดยเราแสดงโครงสร้างของขั้นตอนวิธีได้ดังรูปที่ 3.5



รูปที่ 3.5 แผนผังกระบวนการการแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ

การคำนวณหาเมตริกซ์ความคล้ายคำนวณจากการหาค่าโค-สแควร์ของฮิสโทแกรมระหว่างจุดยอดของภาพวัตถุเดียวกัน เนื่องจาก ดังนั้นเราสามารถแทนค่าในสมการที่ (2.23) ได้ดังสมการที่ (3. 11)

$$C_{i\bar{i}} = C(p_i, p_{\bar{i}}) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_{\bar{i}}(k)]^2}{h_i(k) + h_{\bar{i}}(k)} \quad (3. 11)$$

โดยที่ C_{ij} คือ ค่าต้นทุนแสดงความแตกต่างของฮิสโทแกรมระหว่างจุดอ้างอิงที่ i ของภาพ P และ จุดอ้างอิงที่ \bar{i} ของภาพ P

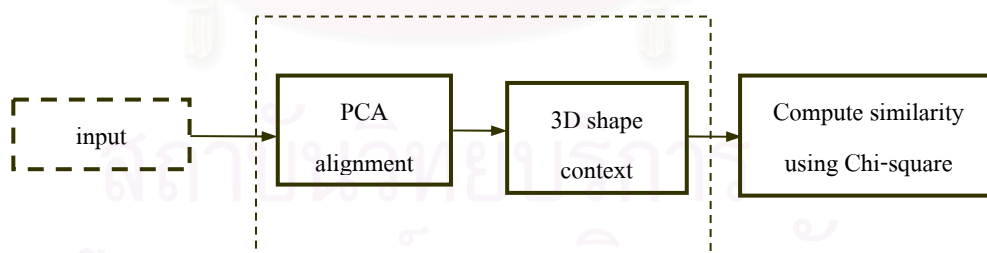
เมื่อนำขั้นตอนวิธีการสร้างเมตริกซ์ความคล้ายของวัตถุมาใช้เพื่อเปรียบเทียบความแตกต่างของภาพคำนวณได้จากการหาระยะทางยูคลิดระหว่างเมตริกซ์ดังสมการที่(3. 12)

$$L_2(C_P, C_Q) = \|C_P - C_Q\|^2 \quad (3. 12)$$

โดยที่ C_P คือ ค่าเมตริกซ์ความคล้ายของวัตถุ P
 C_Q คือ ค่าเมตริกซ์ความคล้ายของวัตถุ Q

3.3. การแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ

จากขั้นตอนวิธีในหัวข้อ 3.2 เป็นการใช้เมตริกซ์ความคล้ายที่สร้างจากฮิสโทแกรมของภาพมาเป็นคุณลักษณะของภาพ ดังนั้นในหัวข้อนี้จึงนำเสนอวิธีการแยกแยะภาพโดยใช้บริบทรูปร่างโดยตรง ซึ่งการคำนวณหาบริบทรูปร่างของภาพวัตถุใช้บริบทรูปร่างที่มีการปรับปรุงให้เหมาะสมกับการคำนวณสำหรับภาพใบหน้าตามที่กล่าวใน 3.1 โดยโครงสร้างของขั้นตอนวิธีสามารถแสดงได้ดังรูปที่ 3.6



รูปที่ 3.6 แผนผังการแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ

จากรูปที่ 3.6 จะเห็นว่า ภาพใบหน้าแบบ 3 มิติจะผ่านกระบวนการปรับแนวภาพก่อนทำการหาบริบทรูปร่างเพื่อให้ได้คุณลักษณะของภาพที่ไม่แปรผันต่อการแสดงออกของภาพ จากนั้นในการแยกแยะความแตกต่างของภาพระหว่างภาพ 2 ภาพ จะเป็นการใช้การทดสอบโคห์-สแควร์ระหว่างฮิสโทแกรมของจุดยอดของภาพวัตถุทั้งสองเพื่อหาค่าความแตกต่างของภาพออกมา

เมื่อเปรียบเทียบระหว่างการแยกแยะภาพใบหน้าทั้ง 2 ขั้นตอน จะพบว่าผลการแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพมีการเก็บข้อมูลภาพสูงกว่าการแยกแยะภาพด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ แต่การแยกแยะภาพด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมจะได้เปรียบที่มีการเก็บข้อมูลสำหรับประมวลผลได้ดีกว่าเนื่องจากคุณลักษณะของภาพ 3 มิติจะเหลือเพียงเมทริกซ์ความคล้ายเท่านั้น โดยเราได้ทำการทดลองเพื่อเปรียบเทียบคุณสมบัติของขั้นตอนวิธีแบบต่างๆ ในบทที่ 4



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการทดลองและการวิเคราะห์ผล

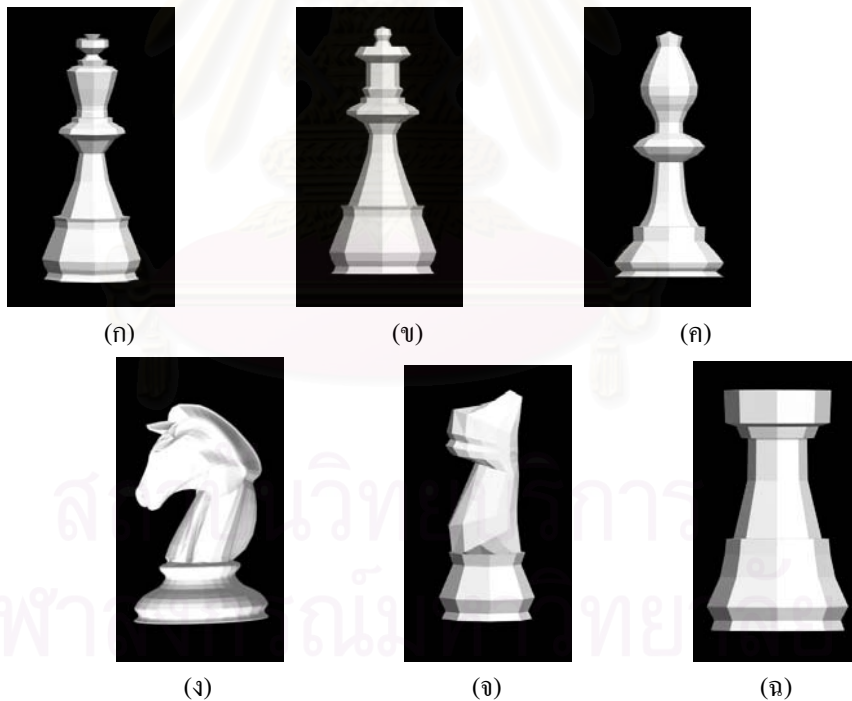
ในบทนี้จะแสดงผลการทดลองตามขั้นตอนวิธีที่ได้กล่าวไว้ในบทที่ 3 เมื่อทำการทดสอบกับภาพ 3 มิติที่มีการแปลงแบบต่างๆ รวมไปถึงการทดสอบกับภาพใบหน้าที่มีการแสดงออก เพื่อทดสอบว่าขั้นตอนวิธีสามารถลดผลจากการจำลองการแปลงแบบต่างๆหรือการแสดงออกบนใบหน้าที่ได้ และแยกแยะความแตกต่างของภาพแต่ละกลุ่มที่แตกต่างกันได้

4.1. รายละเอียดฐานข้อมูลภาพ

ฐานข้อมูลภาพที่ใช้ในการทดสอบประกอบด้วยฐานข้อมูล 2 กลุ่ม คือ ฐานข้อมูลรูปวัตถุสังเคราะห์ และฐานข้อมูลภาพใบหน้าจริง

4.1.1 กลุ่มฐานข้อมูลภาพสังเคราะห์

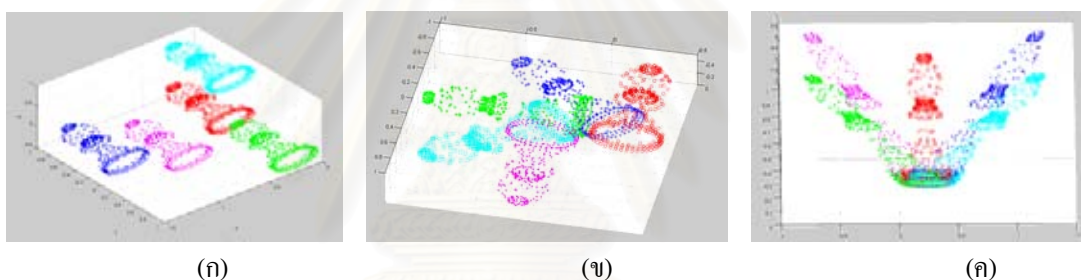
ฐานข้อมูลภาพสังเคราะห์เป็นข้อมูลภาพของตัวหมากรุกที่สร้างขึ้นในรูปแบบไฟล์ .wrl จากฐานข้อมูล INRIA Gamma team research database Website Collections [26] จำนวน 6 แบบ ดังแสดงในรูปที่ 4.1 ซึ่งแสดงผลภาพด้วยโปรแกรม 3DS Max.8 เพื่อให้สามารถแสดงผลรูปร่างได้ชัดเจน



รูปที่ 4.1 รูปสังเคราะห์ทั้ง 6 แบบที่ใช้ในการทดลอง (ก) ตัวหมากกรักราชา (ข) ตัวหมากกรักราชินี (ค) ตัวหมากกรุกขุนนาง (ง) ตัวหมากกรุกอัศวินแบบรายละเอียดสูง (จ) ตัวหมากกรุกอัศวินแบบรายละเอียดต่ำ และ (ฉ) ตัวหมากกรุกป้อมปราการ

รูปที่ 4.1 เป็นภาพสังเคราะห์ของตัวหมากรุกทั้ง 6 แบบที่ใช้ในการทดลอง คือ ตัวหมากรุกราชา (king) ซึ่งประกอบด้วยจุดยอดจำนวน 314 จุด ตัวหมากรุกราชินี (queen) ซึ่งประกอบด้วยจุดยอดจำนวน 282 จุด ตัวหมากรุกขุนนาง (bishop) ซึ่งประกอบด้วยจุดยอดจำนวน 250 จุด ตัวหมากรุกอัศวิน (knight) แบบรายละเอียดสูง ซึ่งประกอบด้วยจุดยอดจำนวน 766 จุด ตัวหมากรุกอัศวินแบบรายละเอียดต่ำ ซึ่งประกอบด้วยจุดยอดจำนวน 214 จุด และตัวหมากรุกป้อมปราการ (rook) ซึ่งประกอบด้วยจุดยอดจำนวน 130 จุด

ในการทดลองโดยใช้ภาพสังเคราะห์ ภาพจะถูกแบ่งออกเป็น 2 กลุ่มคือ กลุ่มภาพทดลอง ซึ่งเป็นกลุ่มภาพที่กำหนดไว้ล่วงหน้า โดยกลุ่มภาพทดลองเป็นภาพปกติกของตัวหมากรุกที่ไม่มีการแปลงใดๆ ซึ่งมีจำนวน 6 ภาพ และ กลุ่มภาพทดสอบ ซึ่งเป็นกลุ่มภาพที่ถูกนำมาคำนวณเพื่อเทียบความคล้ายคลึงกับภาพในกลุ่มทดลอง เพื่อแยกแยะภาพไปในกลุ่มให้ถูกต้อง และนำไปใช้หาความแม่นยำในขั้นตอนวิธีต่อไป กลุ่มภาพทดสอบมีจำนวน 78 ภาพ โดยเป็นภาพตัวหมากรุก 6 ชนิด แต่ละชนิดแบ่งเป็นภาพที่แตกต่างกัน 13 ภาพ ซึ่งประกอบด้วยภาพที่ไม่มีการแปลงใดๆ 1 ภาพ กับภาพที่มีการแปลงเรขาคณิต 12 ภาพ คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ จำนวน 4 ภาพ ดังแสดงตัวอย่างในรูปที่ 4.2(ก) การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° จำนวน 4 ภาพ ดังแสดงตัวอย่างในรูปที่ 4.2 (ข) และ การเกิดการเอียงไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ จำนวน 4 ภาพ ดังแสดงตัวอย่างในรูปที่ 4.2 (ค)



รูปที่ 4.2 (ก) ภาพสังเคราะห์ที่เกิดการเลื่อนขนานจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไปใน 4 ทิศทาง
 (ข) ภาพสังเคราะห์ที่เกิดการหมุนจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไปใน 4 มุม
 (ค) ภาพสังเคราะห์ที่เกิดการเอียงจากตำแหน่งเริ่มต้นแสดงด้วยภาพวัตถุสีแดงไป 4 ทิศทาง

สาเหตุที่เลือกใช้ข้อมูลภาพสังเคราะห์เป็นภาพตัวหมากรุก เนื่องจากมีรูปแบบคล้ายคลึงกับลักษณะของใบหน้าที่มีการแสดงออก คือ ลักษณะรูปร่างโดยรวมมีความใกล้เคียงกันแต่รายละเอียดปลีกย่อยมีความแตกต่างกัน ซึ่งข้อมูลตัวหมากรุกจะมีความแตกต่างของรายละเอียดปลีกย่อยที่เห็นได้ชัดเจนกว่าภาพใบหน้า

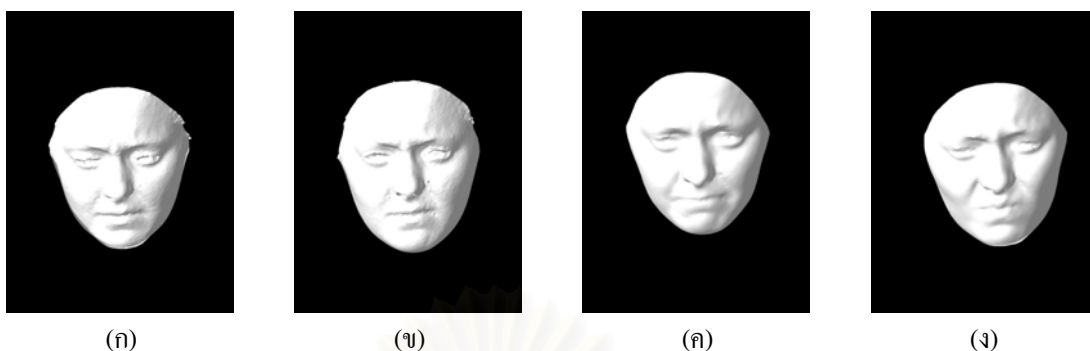
4.1.2 กลุ่มฐานข้อมูลภาพใบหน้าบุคคล

ฐานข้อมูลภาพใบหน้าบุคคล เป็นข้อมูลภาพ 3 มิติที่มีการแสดงออกของ Binghamton University 3D Facial Expression Database (BU-3DFE) [27] โดยในที่นี้ได้เลือกทำการทดลองกับข้อมูลบุคคลจำนวน 8 คน โดยมีการแสดงออกแตกต่างกันคนละ 25 แบบ ได้แก่ อารมณ์ปกติ (neutral) อารมณ์โกรธ (angry) อารมณ์รังเกียจ (disgust) อารมณ์หวาดกลัว (fear) อารมณ์สุข (happy) อารมณ์เสียใจ (sad) และอารมณ์แปลกใจ (surprise) โดยอารมณ์ปกติจะมี 1 แบบ ส่วนอารมณ์แบบอื่นจะมีระดับการแสดงออกแบบละ 4 ระดับ โดยแสดงออกจกน้อยไปมากตามลำดับ ซึ่งภาพมีรูปแบบไฟล์ .wrl โดยแต่ละภาพจะประกอบด้วยจุดยอดประมาณ 6,000-8,000 จุด

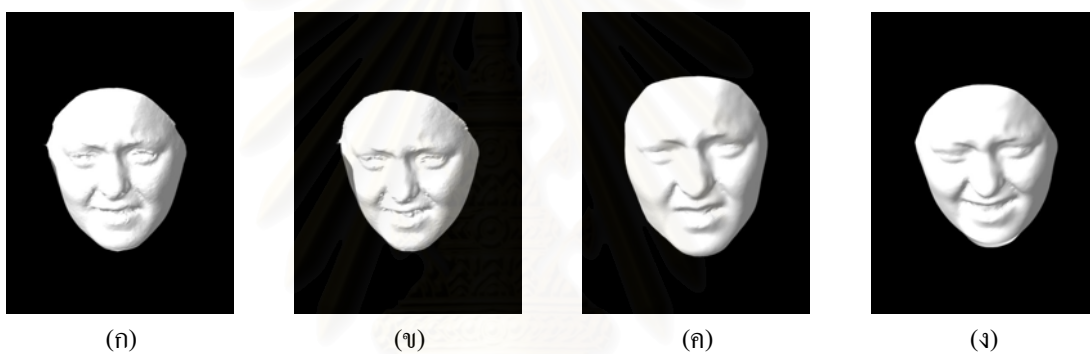
ในการทดลองกระบวนการรู้จำใบหน้า โดยใช้ภาพใบหน้าบุคคล ภาพจะถูกแบ่งออกเป็น 2 กลุ่มคือ กลุ่มภาพทดลอง ซึ่งเป็นภาพที่มีการแสดงอารมณ์ปกติของทุกบุคคลจำนวน 8 ภาพ ดังแสดงในรูปที่ 4.3 และ กลุ่มภาพทดสอบ ซึ่งเป็นกลุ่มภาพที่มีการแสดงอารมณ์แบบต่างๆ กลุ่มภาพทดสอบมีจำนวนทั้งหมด 200 ภาพ โดยเป็นภาพใบหน้าบุคคล 8 คน แต่ละคนแบ่งเป็นภาพที่มีการแสดงออกแตกต่างกัน 25 ภาพ ดังแสดงในรูปที่ 4.4 ถึง รูปที่ 4.9 ซึ่งการทดสอบใช้เพียงข้อมูลรูปร่างของใบหน้าเท่านั้น โดยไม่นำส่วนข้อมูลภาพพื้นผิวมาใช้ ซึ่งรูปที่ 4.3 และ รูปที่ 4.4 ถึง รูปที่ 4.9 แสดงภาพด้วยการประมวลผลภาพโดยโปรแกรม 3DS Max.8 เพื่อให้สามารถแสดงผลรูปร่างได้ดีกว่าภาพจากโปรแกรม MATLAB



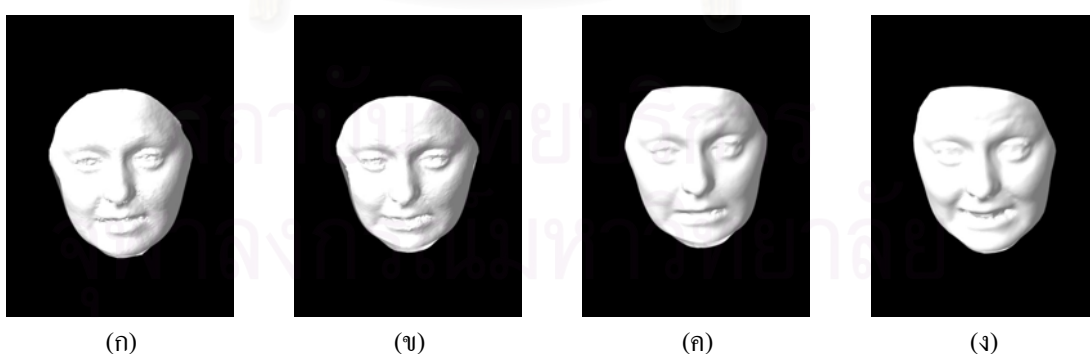
รูปที่ 4.3 ภาพกลุ่มทดสอบเป็นบุคคลที่มีการแสดงอารมณ์ปกติจำนวน 8 ภาพ (ก) ภาพบุคคลทดสอบที่ 1 (ข) ภาพบุคคลทดสอบที่ 2 (ค) ภาพบุคคลทดสอบที่ 3 (ง) ภาพบุคคลทดสอบที่ 4 (จ) ภาพบุคคลทดสอบคนที่ 5 (ฉ) ภาพบุคคลทดสอบที่ 6 (ช) ภาพบุคคลทดสอบที่ 7 (ช) ภาพบุคคลทดสอบที่ 8



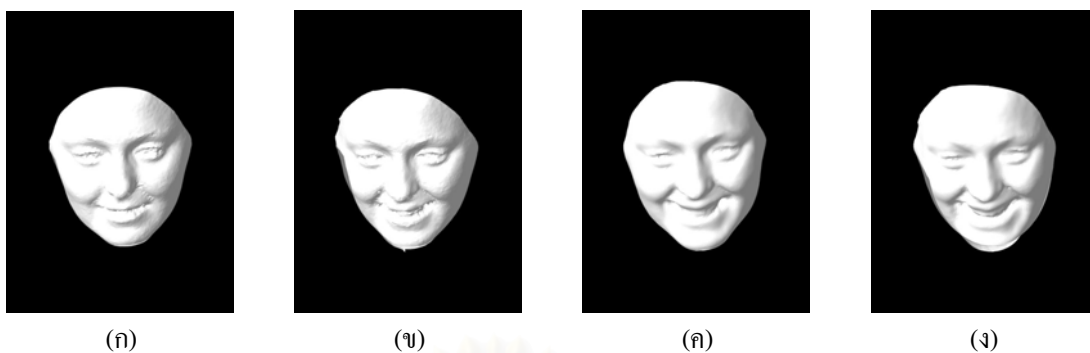
รูปที่ 4.4 ภาพใบหน้าที่มีการแสดงภาพอารมณ์โศรก 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ



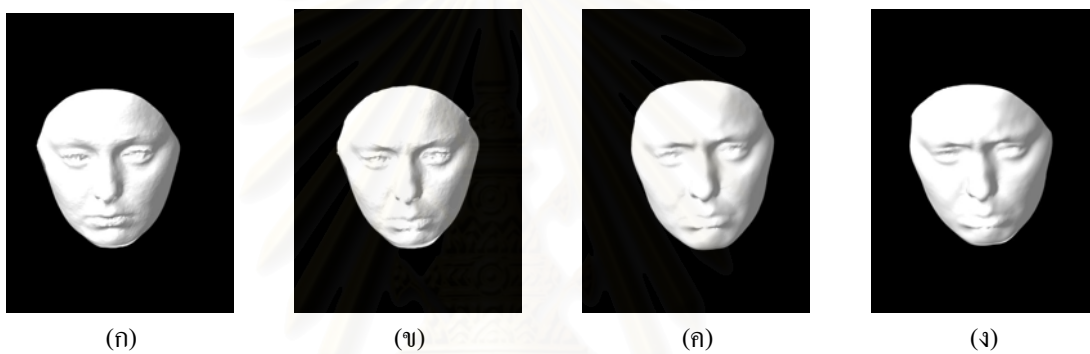
รูปที่ 4.5 ภาพใบหน้าที่มีการแสดงอารมณ์รังเกียจ 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ



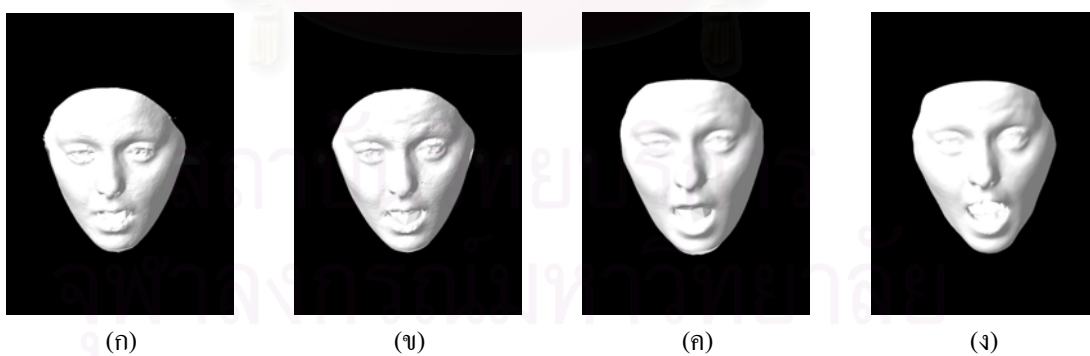
รูปที่ 4.6 ภาพใบหน้าที่มีการแสดงอารมณ์หวาดกลัว 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ



รูปที่ 4.7 ภาพใบหน้าที่มีการแสดงอารมณ์สุข 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ



รูปที่ 4.8 ภาพใบหน้าที่มีการแสดงอารมณ์เสียใจ 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ



รูปที่ 4.9 ภาพใบหน้าที่มีการแสดงอารมณ์แปลกใจ 4 ระดับจากน้อยไปมาก (ก)-(ง) ตามลำดับ

4.2. การทดลองแยกแยะภาพด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ

4.2.1 การทดลองการแยกแยะภาพสังเคราะห์เมื่อจำลองการแปลงในภาพสังเคราะห์ด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ

ในขั้นแรก เราทำการทดลองกับกลุ่มฐานข้อมูลภาพสังเคราะห์ โดยใช้ขั้นตอนวิธีการแยกแยะภาพด้วยเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพที่ได้กล่าวในหัวข้อ 3.2 โดยกำหนดจำนวนช่องของฮิสโทแกรมในแนว (ρ, θ, ϕ) เป็น (10, 18, 18) ช่อง ตามลำดับ และทำการสุ่มจุดยอดจำนวน 83 จุด เพื่อทำการหาเมทริกซ์ความคล้ายของฮิสโทแกรมของจุดยอด ผลภาพเมทริกซ์ความคล้ายของแต่ละภาพที่ใช้ในการทดลองดังรูปที่ 4.10 ถึง รูปที่ 4.15



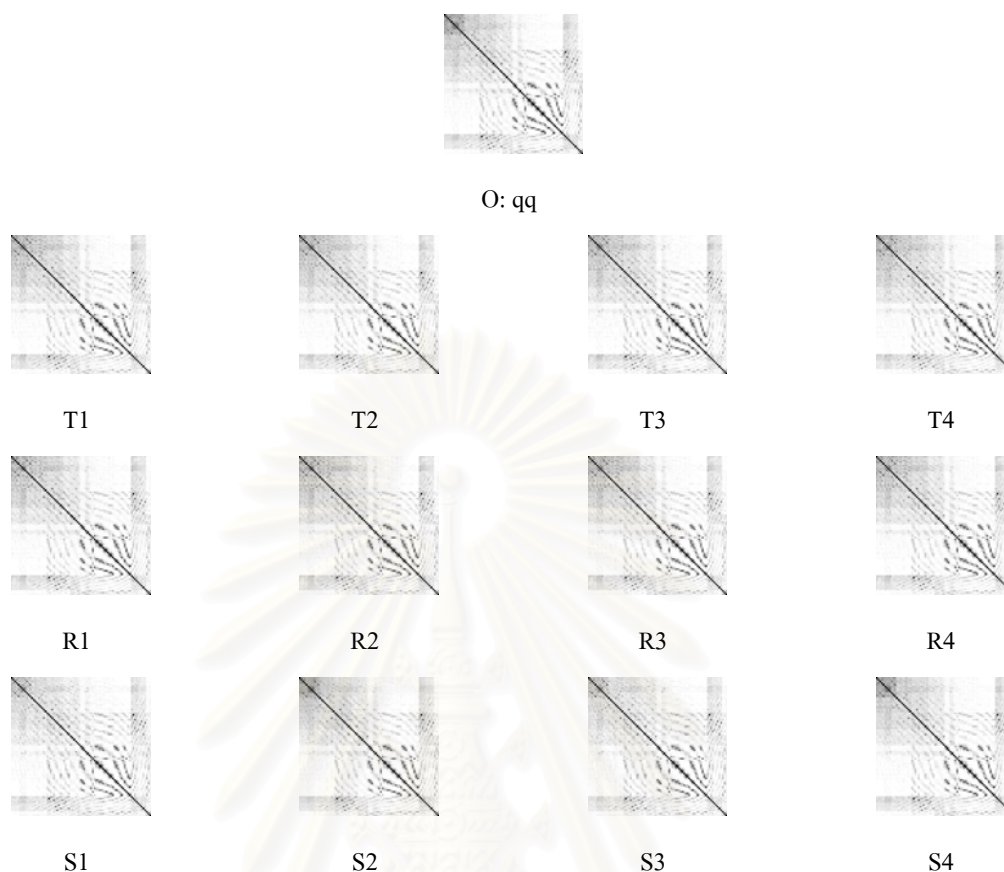
รูปที่ 4.10 ภาพเมทริกซ์ความคล้ายของตัวหมากรุกราชาที่มีการแปลงแบบต่างๆ

โดยที่ O: kk คือภาพเมทริกซ์ความคล้ายของตัวหมากรุกราชาที่ไม่เกิดการแปลงใดๆ

T1 – T4 คือ การเลื่อนขนานไปตามทิศ (1, 0, 0), (0, 1, 0), (0, 0, 1) และ (1, 1, 1) ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเฉือนไปตามทิศ (1, 0, 1), (-1, 0, -1), (1, 0, -1) และ (-1, 0, 1) ตามลำดับ



รูปที่ 4.11 ภาพเมทริกซ์ความคล้ายของตัวหมากรูกราซึนึที่มีการแปลงแบบต่างๆ

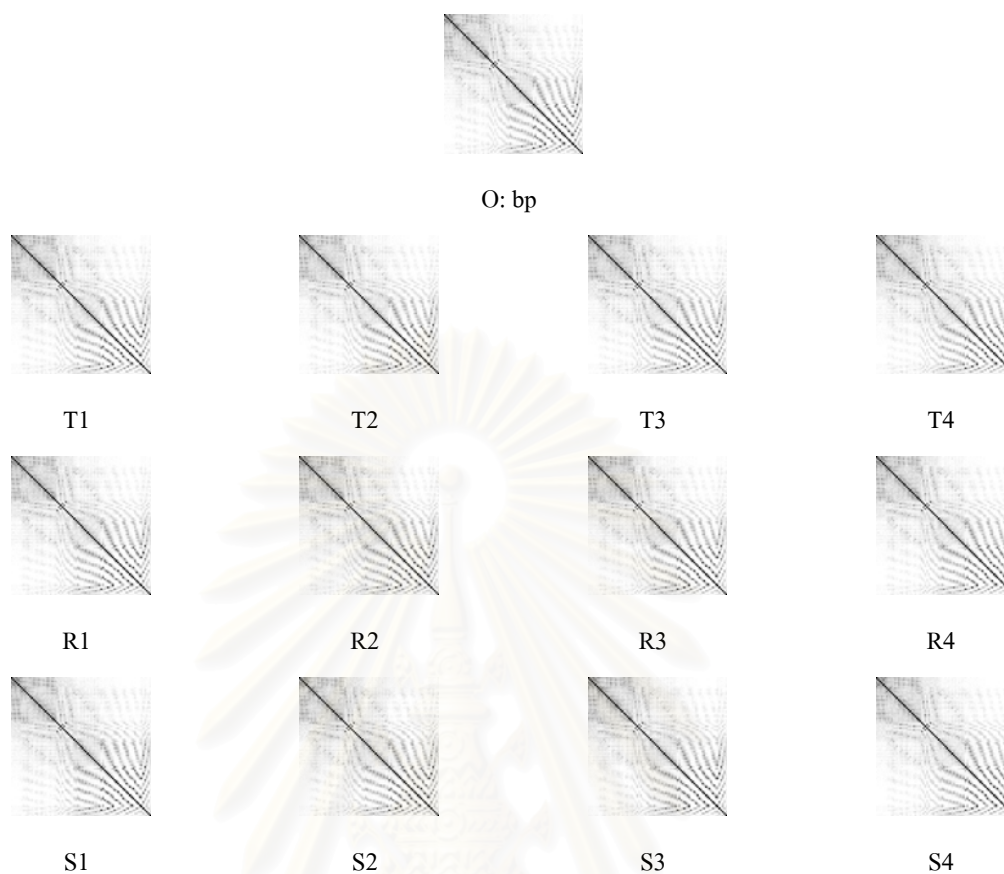
โดยที่ O: qq คือภาพเมทริกซ์ความคล้ายของตัวหมากรูกราซึนึที่ไม่เกิดการแปลงใดๆ

T1 – T4 คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเฉือนไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.12 ภาพเมทริกซ์ความคล้ายของตัวหมากรุกขุนนางที่มีการแปลงแบบต่างๆ

โดยที่ O: bp คือภาพเมทริกซ์ความคล้ายของตัวหมากรุกขุนนางที่ไม่เกิดการแปลงใดๆ

T1 – T4 คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทางเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเลื่อนไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.13 ภาพเมทริกซ์ความคล้ายของตัวหมากรุกอัศวินแบบรายละเอียดสูงที่มีการแปลงแบบต่างๆ

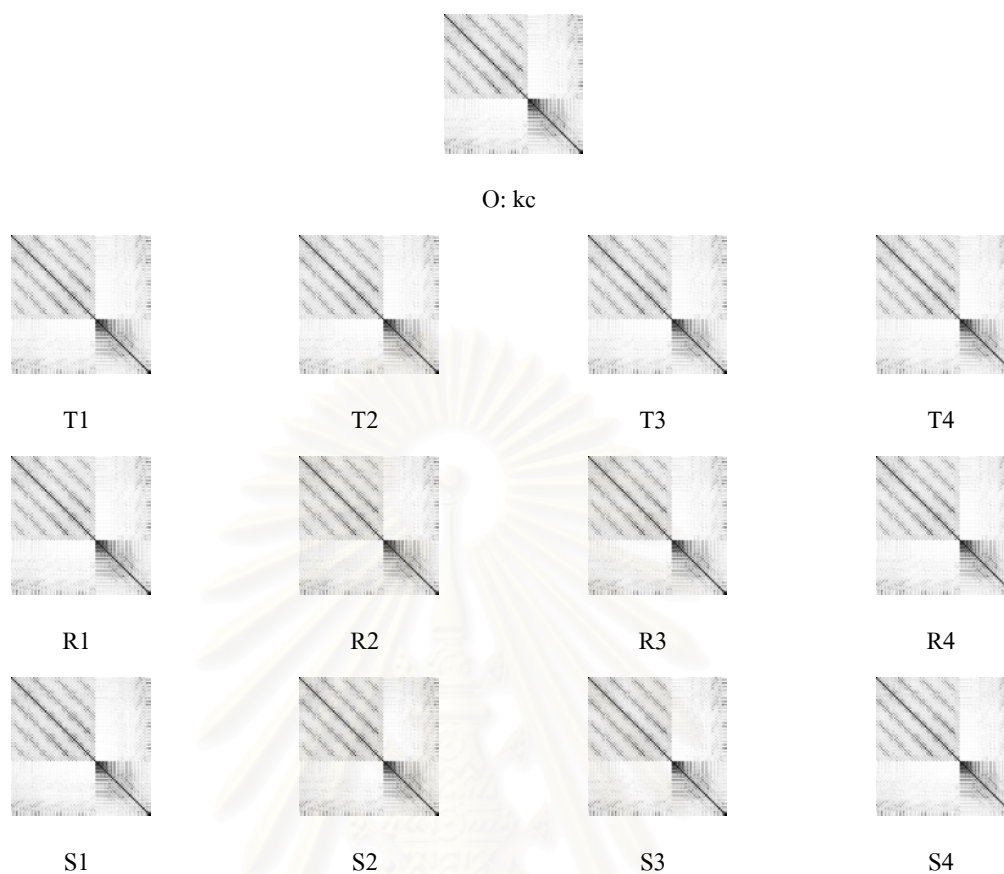
โดยที่ O: kf คือภาพเมทริกซ์ความคล้ายของตัวหมากรุกอัศวินแบบรายละเอียดสูงที่ไม่เกิดการแปลงใดๆ

T1 – T4 คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเฉือนไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.14 ภาพเมทริกซ์ความคล้ายของตัวหมากรุกอัศวินแบบรายละเอียดน้อยที่มีการแปลงแบบต่างๆ

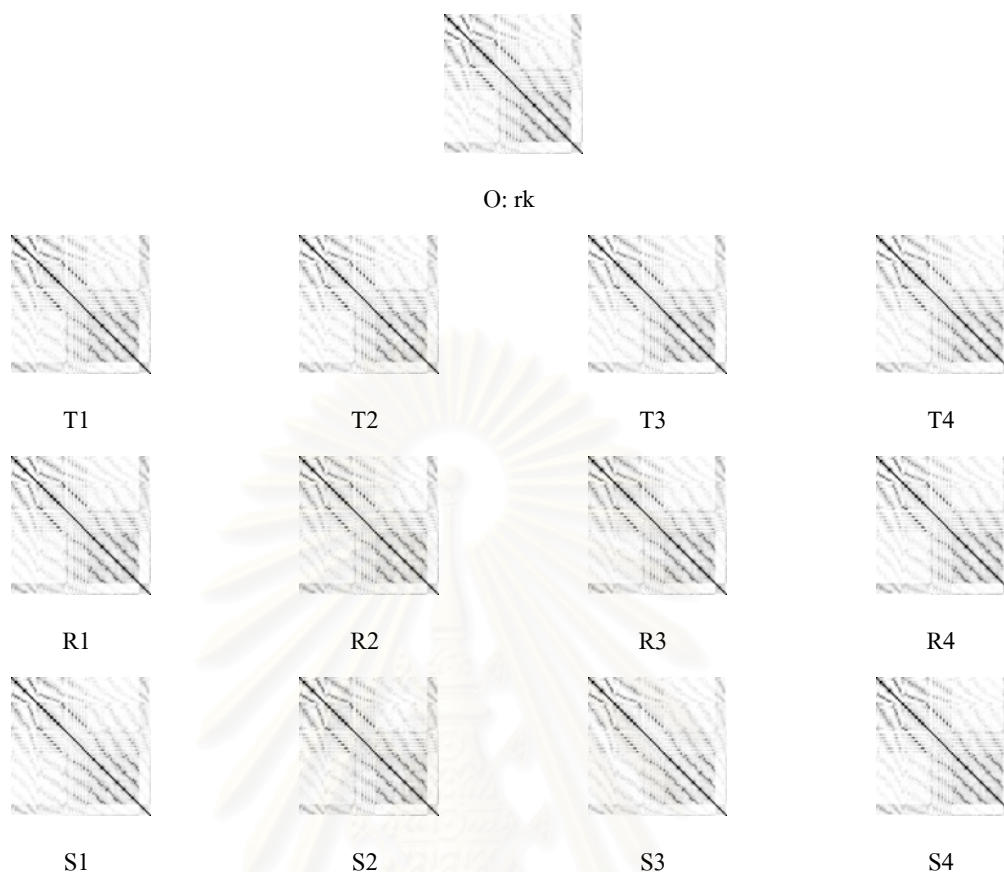
โดยที่ O: kc คือภาพเมทริกซ์ความคล้ายของตัวหมากรุกอัศวินแบบรายละเอียดต่ำที่ไม่เกิดการแปลงใดๆ

T1 – T4 คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเฉือนไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.15 ภาพเมทริกซ์ความคล้ายของตัวหมากรุกป้อมปราการที่มีการแปลงแบบต่างๆ

โดยที่ O: rk คือภาพเมทริกซ์ความคล้ายของตัวหมากรุกป้อมปราการที่ไม่เกิดการแปลงใดๆ
 T1 – T4 คือ การเลื่อนขนานไปตามทิศ $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ และ $(1, 1, 1)$ ตามลำดับ
 R1 – R4 คือ การหมุนของวัตถุไปตามทิศทางเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ
 S1 – S4 คือ การเกิดการเฉือนไปตามทิศ $(1, 0, 1)$, $(-1, 0, -1)$, $(1, 0, -1)$ และ $(-1, 0, 1)$ ตามลำดับ

การเปรียบเทียบความแตกต่างระหว่างเมทริกซ์ความคล้ายของตัวหมากรุกกับข้อมูลกลุ่มภาพทดลองโดยใช้ระยะทางยูคลิดระหว่างเมทริกซ์ความคล้ายได้ผลดังตาราง 4.1

จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.1 ผลการแยกแยะภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอดบนภาพ

เมื่อ T1 – T4 คือ การเลื่อนขนานไปตามทิศ (1, 0, 0), (0, 1, 0), (0, 0, 1) และ (1, 1, 1) ตามลำดับ

R1 – R4 คือ การหมุนของวัตถุไปตามทิศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ

S1 – S4 คือ การเกิดการเลื่อนไปตามทิศ (1, 0, 1), (-1, 0, -1), (1, 0, -1) และ (-1, 0, 1) ตามลำดับ

	Obj1:kk	Obj2:qq	Obj3:bp	Obj4:kf	Obj5:kc	Obj6:rk	ACC.
T1	1	2	3	4	5	6	6
T2	1	2	3	4	5	6	6
T3	1	2	3	4	5	6	6
T4	1	2	3	4	5	6	6
R1	1	2	3	4	5	6	6
R2	1	2	3	4	5	6	6
R3	1	2	3	4	5	6	6
R4	1	2	3	4	5	6	6
S1	1	2	3	4	5	6	6
S2	1	2	3	4	5	6	6
S3	1	2	3	4	5	6	6
S4	1	2	3	4	5	6	6
ACC.	12	12	12	12	12	12	72

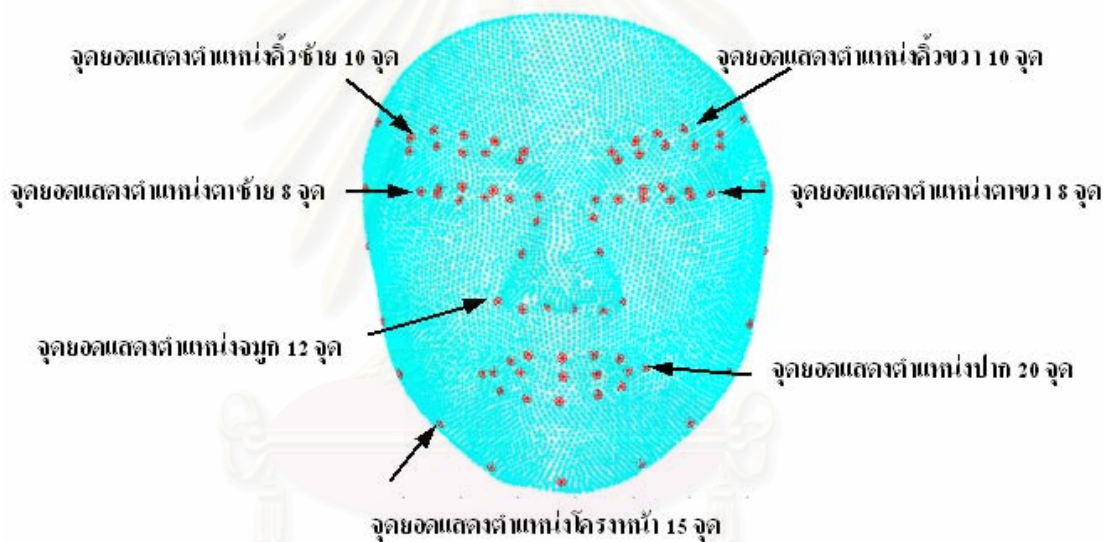
จากตาราง 4.1 จะเห็นได้ว่า ขั้นตอนวิธีที่ใช้มีความแม่นยำในการแยกแยะภาพตัวหมากรุกที่เกิดการแปลงแบบต่างๆ ได้แม่นยำถึง 100 เปอร์เซ็นต์ คือ วัตถุชิ้นที่ 1 ซึ่งเป็นหมากรุกพระราช (Obj1:kk) ที่มีการแปลงทั้ง 12 แบบ (T1-T4, R1-R4, S1-S4) ถูกแยกแยะได้ว่าเป็นวัตถุกลุ่ม 1 ทั้งหมด เช่นเดียวกับภาพวัตถุทุกแบบ ทั้งนี้ จะสังเกตได้จากภาพเมทริกซ์ความคล้ายของตัวหมากรุกแต่ละแบบ ที่แม้จะเกิดการแปลงแบบต่างๆ แต่ผลที่ได้ก็มีลักษณะที่แตกต่างกันอย่างชัดเจนนั่นเอง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4.2.2 การทดลองการแยกแยะภาพใบหน้าที่มีการแสดงออกด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ

ในขั้นตอนถัดมา เราทำการทดลองกับภาพใบหน้าบุคคลที่มีการแสดงออกแบบต่างๆ 7 แบบ คือ อารมณ์ปกติ อารมณ์โกรธ อารมณ์รังเกียจ อารมณ์หวาดกลัว อารมณ์สุข อารมณ์เสียใจ และอารมณ์แปลกใจ โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับ

โดยทำการเลือกจุดยอดในตำแหน่งองค์ประกอบบนใบหน้าจำนวน 83 จุด ได้แก่ ตำแหน่งจุดยอดที่ตาซ้ายและขวารวม 16 จุด คิ้วซ้ายและขวารวม 20 จุด จมูกจำนวน 12 จุด ริมฝีปากบนและล่างรวม 20 จุด และโครงหน้าอื่นๆอีกจำนวน 15 จุด ซึ่งแสดงตำแหน่งองค์ประกอบบนใบหน้าได้ดังรูปที่ 4.16 เพื่อนำมาหาเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดโดยกำหนดจำนวนช่องของฮิสโทแกรมในแนว (ρ, θ, ϕ) เป็น (10, 18, 18) ช่อง ตามลำดับ เราสามารถแสดงภาพเมตริกซ์ความคล้ายของแต่ละภาพที่ใช้ในการทดลองได้ดังรูปที่ 4.17 ถึง รูปที่ 4.24 เพื่อทำการแยกแยะความแตกต่างด้วยระยะทางยูคลิดต่อไป



รูปที่ 4.16 ภาพตำแหน่งองค์ประกอบบนใบหน้าซึ่งประกอบด้วย ตา คิ้ว จมูก ปาก และ โครงหน้า

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.17 ภาพเมทริกซ์ความคล้ายของไบทอน้ำบุคคลคนที่ 1 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.18 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 2 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)

อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.19 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 3 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.20 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 4 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.21 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 5 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.22 ภาพเมทริกซ์ความคล้ายของไบทอนูคลคอนที่ 6 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.23 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 7 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก



รูปที่ 4.24 ภาพเมทริกซ์ความคล้ายของใบหน้าบุคคลคนที่ 8 เมื่อมีการแสดงออกแบบต่างๆ 7 แบบ

คือ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)
 อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก

ตาราง 4.2 ผลการแยกแยะภาพบุคคลด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอคบนภาพ

เมื่อ อารมณ์ปกติ (N) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)

อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก

	คนที่1	คนที่2	คนที่3	คนที่4	คนที่5	คนที่6	คนที่7	คนที่8	ACC.
AN1	6	6	6	2	2	2	8	1	0
AN2	2	2	2	6	2	1	6	2	1
AN3	2	2	1	6	2	6	7	2	3
AN4	1	2	6	2	2	8	2	6	2
DI1	2	4	6	8	2	4	8	8	1
DI2	2	6	2	2	2	6	8	8	2
DI3	6	2	1	6	5	1	6	2	2
DI4	2	2	1	6	5	8	8	8	3
FE1	6	2	2	2	4	2	8	2	1
FE2	2	2	2	1	4	6	6	8	3
FE3	2	2	6	2	2	8	7	2	2
FE4	6	2	4	6	2	6	2	2	2
HA1	2	2	1	6	2	2	8	2	1
HA2	2	2	8	2	5	6	2	2	3
HA3	6	2	8	6	6	2	8	6	1
HA4	2	1	6	2	2	8	2	2	0
SA1	2	2	2	4	1	2	2	8	3
SA2	2	2	1	4	5	2	7	2	4
SA3	6	2	1	6	5	2	2	6	2
SA4	2	2	2	1	2	6	2	1	2
SU1	2	4	8	2	6	2	6	8	1
SU2	8	6	2	4	4	3	2	6	1
SU3	1	2	2	2	6	2	2	8	3
SU4	2	2	4	8	4	2	4	2	1
ACC.	2	18	0	3	5	6	3	7	44

จากตาราง 4.2 จะเห็นได้ว่า ขั้นตอนวิธีที่ใช้มีความแม่นยำในการแยกแยะภาพบุคคลที่มีการแสดงออกได้เพียง 22.92 เปอร์เซ็นต์เท่านั้น คือ ใบหน้าบุคคลที่ 1 มีการแยกแยะได้ถูกต้อง 2 ภาพ ได้แก่ ภาพที่มีการแสดงออกด้วยอารมณ์โกรธระดับ 4 (AN4) และ อารมณ์แปลกใจระดับ 3 (SU3) และเช่นเดียวกับกับภาพบุคคลอื่นๆ โดย

สามารถแยกแยะได้ถูกต้อง 44 ภาพ จากภาพกลุ่มทดสอบทั้งหมด 192 ภาพ ซึ่งเมื่อสังเกตจากภาพเมทริกซ์ความคล้ายของใบหน้าบุคคล จะเห็นได้ว่าข้อมูลที่ได้ของแต่ละบุคคลมีความคล้ายคลึงกันมาก เพราะการลดข้อมูลภาพจากข้อมูลฮิสโทแกรมของจุดภาพลงมาเป็นเมทริกซ์ความคล้าย อาจทำให้ข้อมูลบางส่วนหายไป

4.3. การทดลองแยกแยะภาพด้วยฮิสโทแกรมของจุดยอคนภาพ

4.3.1 การทดลองแยกแยะภาพสังเคราะห์เมื่อจำลองการแปลงในภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอคนภาพ

เช่นเดียวกับการทดลองในหัวข้อ 4.2 คือใช้ภาพสังเคราะห์ตัวหมากรุก 6 แบบ โดยกลุ่มภาพทดลองเป็นภาพที่ไม่มีมีการแปลงใดๆ และกลุ่มภาพทดสอบใช้ภาพที่ไม่มีมีการแปลงใดๆเช่นเดียวกับกลุ่มภาพทดลองร่วมกับภาพที่มีการแปลงเรขาคณิต 3 แบบ คือ การเลื่อนขนาน การหมุน และ การเอียง แต่ในการทดลองนี้ใช้ขั้นตอนวิธีดังที่ได้กล่าวในหัวข้อ 3.3 คือการใช้บริบทรูปร่างที่ได้จากภาพวัตถุมาเป็นคุณลักษณะในการแยกแยะภาพ และใช้การทดสอบไค-สแควร์ในการหาความแตกต่างของภาพ โดยกำหนดจำนวนช่องของฮิสโทแกรมในแนว (ρ, θ, ϕ) เป็น (10, 18, 18) ช่อง ตามลำดับ ซึ่งแสดงผลการแยกแยะภาพได้ดังตาราง 4.3

ตาราง 4.3 ผลการแยกแยะภาพสังเคราะห์ด้วยฮิสโทแกรมของจุดยอคนภาพ

เมื่อ T1 – T4 คือ การเลื่อนขนานไปตามทิศ (1, 0, 0), (0, 1, 0), (0, 0, 1) และ (1, 1, 1) ตามลำดับ
R1 – R4 คือ การหมุนของวัตถุไปตามทิศทางเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ
S1 – S4 คือ การเกิดการเอียงไปตามทิศ (1, 0, 1), (-1, 0, -1), (1, 0, -1) และ (-1, 0, 1) ตามลำดับ

	Obj1:kk	Obj2:qq	Obj3:bp	Obj4:kf	Obj5:kc	Obj6:rk	ACC.
T1	1	2	3	4	5	6	6
T2	1	2	3	4	5	6	6
T3	1	2	1	4	5	6	5
T4	1	2	3	4	5	6	6
R1	1	2	1	4	5	6	5
R2	1	2	2	4	5	6	5
R3	1	2	3	4	5	4	6
R4	1	2	2	4	5	6	5
S1	1	2	3	4	5	6	6
S2	1	2	3	4	5	6	6
S3	1	2	3	4	5	6	6
S4	1	2	3	4	5	6	6
ACC.	12	12	9	12	12	11	68

จากตาราง 4.3 จะเห็นได้ว่า ขั้นตอนวิธีที่ใช้มีความแม่นยำในการแยกแยะภาพตัวหมากรุกที่เกิดการแปลงแบบต่างๆ ด้วยฮิสโทแกรมของจุดยอคนภาพได้แม่นยำถึง 94.44 เปอร์เซ็นต์ โดยคิดเป็นจำนวนภาพที่มี

การแยกแยะถูกต้อง 68 ภาพ จากภาพทดสอบทั้งหมด 72 ภาพ โดยเกิดความผิดพลาดที่ภาพวัตถุที่ 3 (Obj3:bp) ที่มีการแปลงแบบเลื่อนขนานในทิศ (0, 0,1) (T3) และการหมุน 45° (R1) และภาพวัตถุที่ 6 (Obj6:rk) ที่เกิดการหมุน 135° (R3)

4.3.2 การทดลองแยกแยะภาพใบหน้าที่มีการแสดงออกด้วยฮิสโทแกรมของจุดยอดบนภาพ

ในขั้นตอนถัดมา เราทำการทดลองกับภาพใบหน้าบุคคลที่มีการแสดงออกแบบต่างๆ 7 แบบ คือ อารมณ์ปกติ อารมณ์โกรธ อารมณ์รังเกียจ อารมณ์หวาดกลัว อารมณ์สุข อารมณ์เสียใจ และอารมณ์แปลกใจ โดยแต่ละการแสดงออกยกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับ โดยกำหนดจำนวนช่องของฮิสโทแกรมในแนว (ρ, θ, ϕ) เป็น (10, 18, 18) ช่อง ตามลำดับ ซึ่งได้ผลการแยกแยะดังแสดงในตาราง 4.4



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.4 ผลการแยกแยะภาพบุคคลด้วยฮิสโทแกรมของจุดยอดบนภาพ

เมื่อ อารมณ์ปกติ (NE) อารมณ์โกรธ (AN) อารมณ์รังเกียจ (DI) อารมณ์หวาดกลัว (FE)

อารมณ์สุข (HA) อารมณ์เสียใจ (SA) และอารมณ์แปลกใจ (SU)

โดยแต่ละการแสดงออกกเว้นการแสดงอารมณ์ปกติ จะมีระดับการแสดงออก 4 ระดับจากน้อยไปมาก

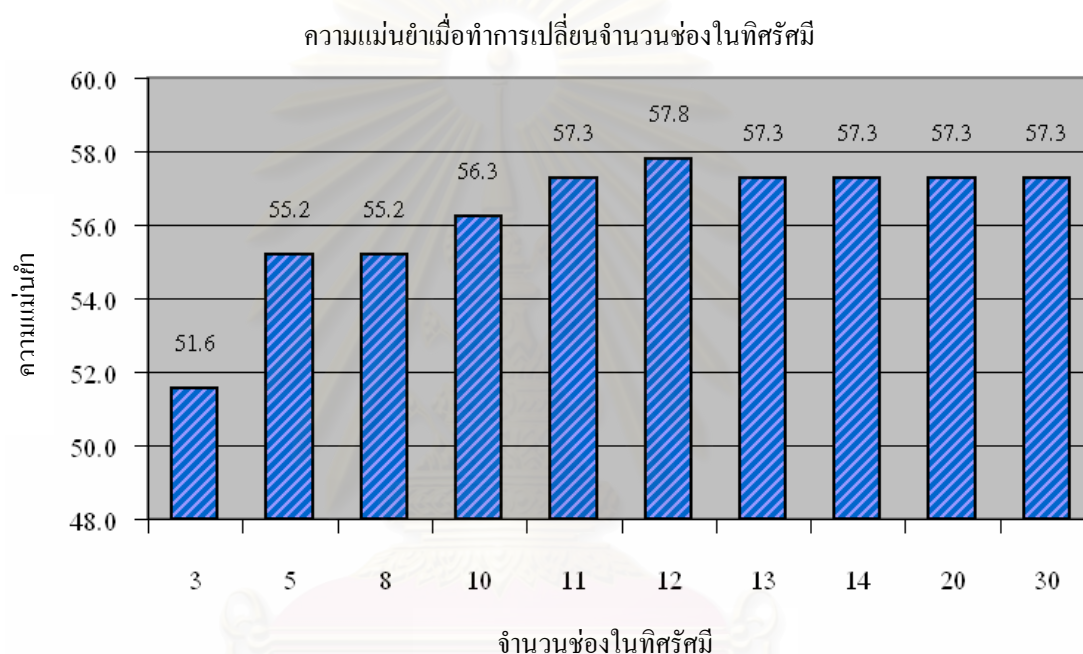
	คนที่1	คนที่2	คนที่3	คนที่4	คนที่5	คนที่6	คนที่7	คนที่8	ACC.
AN1	4	2	6	4	5	6	7	8	6
AN2	3	6	2	4	2	6	7	6	3
AN3	1	6	6	4	5	6	7	6	5
AN4	1	2	6	2	5	6	7	2	5
DI1	2	3	3	4	5	6	6	8	5
DI2	2	2	2	4	5	6	7	8	6
DI3	1	2	6	2	4	6	6	6	3
DI4	1	2	6	4	5	6	6	8	6
FE1	3	2	6	4	5	6	8	2	4
FE2	4	2	3	2	1	6	6	8	4
FE3	1	2	3	4	1	6	7	8	7
FE4	1	2	4	4	5	6	2	6	5
HA1	1	2	3	4	5	6	7	2	7
HA2	2	2	2	2	5	6	6	8	4
HA3	1	2	2	4	5	6	6	2	5
HA4	4	4	2	2	5	6	2	2	2
SA1	2	6	6	4	1	6	7	8	4
SA2	2	2	1	4	5	6	7	6	5
SA3	1	6	6	6	5	6	6	6	3
SA4	1	2	3	2	5	6	7	8	7
SU1	2	2	3	4	1	6	7	8	6
SU2	4	6	1	4	4	6	7	6	3
SU3	1	2	2	4	5	6	7	8	7
SU4	1	2	1	4	4	6	6	8	5
ACC.	12	17	6	17	16	24	13	12	117

จากตาราง 4.4เห็นได้ว่า ขั้นตอนวิธีที่ใช้มีความแม่นยำในการแยกแยะภาพบุคคลที่มีการแสดงออกได้
ถูกต้อง 117 รูป จากภาพกลุ่มทดสอบทั้งหมด 192 รูป ซึ่งคิดเป็น 60.94 %

4.4. การเพิ่มประสิทธิภาพของกระบวนการ แยกแยะภาพใบหน้าด้วยฮิสโทแกรมของจุดยอด

จากการทดลองที่ผ่านมาจะเห็นได้ว่า กระบวนการแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดมีความเหมาะสมในการใช้งานแยกแยะภาพใบหน้าที่มีการแสดงออกมากกว่าเมตริกซ์ความคล้าย ดังนั้นในการทดลองนี้ได้ทดลองเพื่อเพิ่มประสิทธิภาพของกระบวนการรู้จำใบหน้าโดยใช้ฮิสโทแกรมของจุดยอด ด้วยการหาจำนวนตาตารางที่ให้ความแม่นยำกับกระบวนการสูงสุด

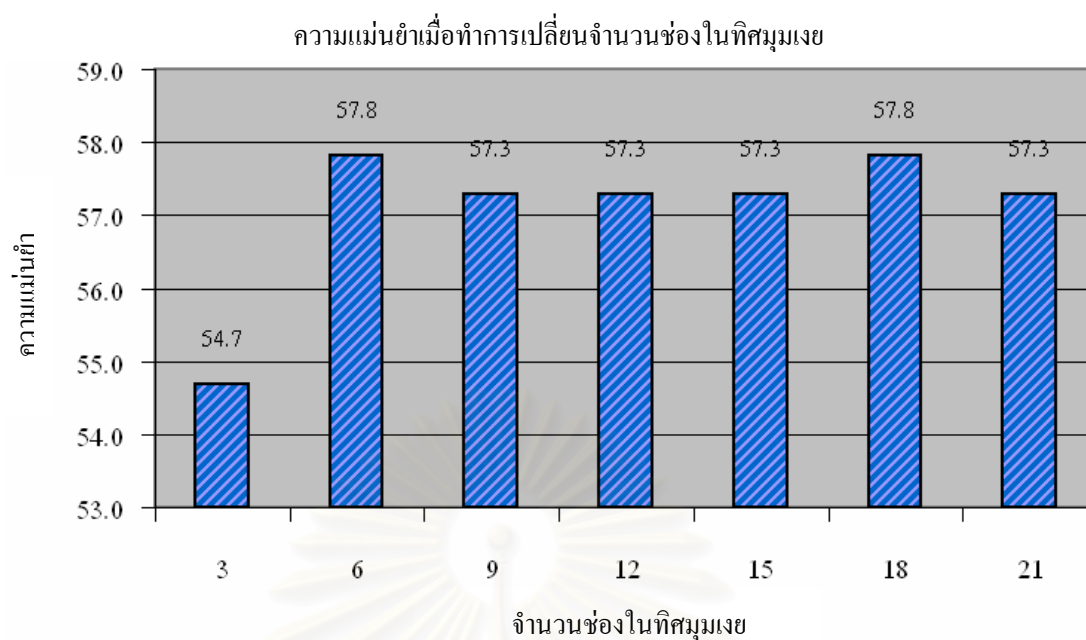
โดยขั้นแรกทำการกำหนดจำนวนช่องของตาตารางในทิศทางของมุมเงย θ และมุมเอซิมัท ϕ ให้มีจำนวนคงที่ คือ 18 ช่อง ทั้งสองทิศทาง แล้วทำการเปลี่ยนจำนวนช่องในแนวนอน ρ ด้วยค่าต่างๆ ผลที่ได้แสดงได้ดังกราฟในรูปที่ 4.25



รูปที่ 4.25 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิศทางแนวนอน เมื่อกำหนดจำนวนช่องในทิศทางมุมเงย มีจำนวน 18 ช่อง และจำนวนช่องในทิศทางเอซิมัทมีจำนวน 18 ช่อง

จากรูปที่ 4.25 เห็นได้ว่าความแม่นยำมีค่าน้อยเมื่อมีจำนวนช่องน้อยและมีค่าเพิ่มขึ้นจนมีค่ามากที่สุดที่จำนวนช่อง 12 ช่อง โดยได้ความแม่นยำ 57.3 % จากนั้นความแม่นยำจะลดลงเล็กน้อยและมีค่าค่อนข้างคงที่

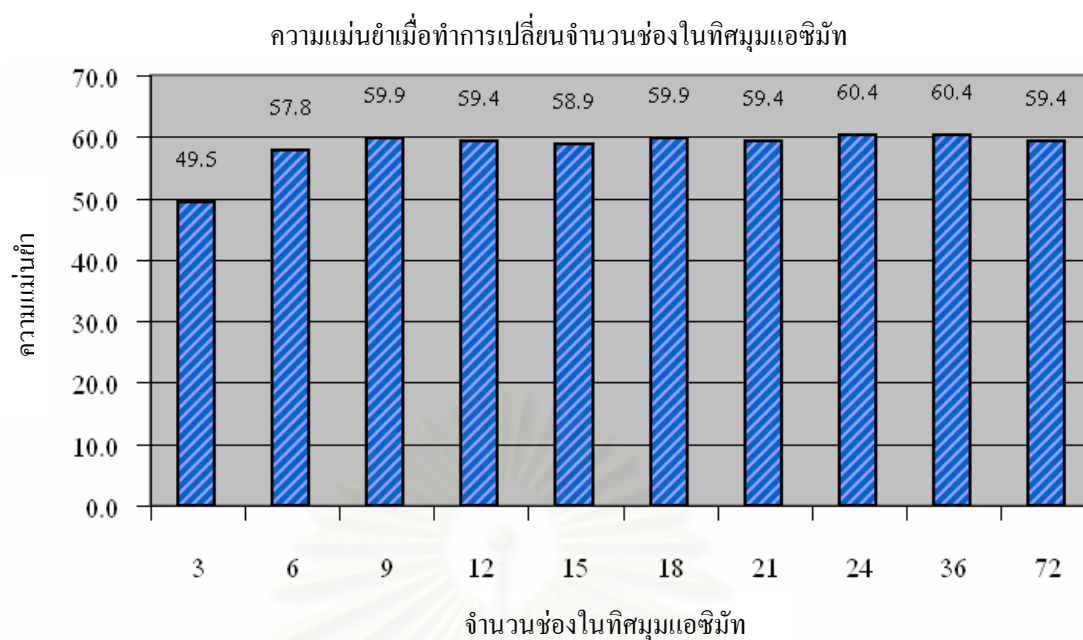
ในขั้นตอนถัดมา ทำการกำหนดจำนวนช่องของตาตารางในทิศทางของแนวนอน ρ ให้มีจำนวนคงที่ 12 ช่อง และมุมเอซิมัท ϕ ให้มีจำนวนคงที่ 18 ช่อง ทั้งสองทิศทาง แล้วทำการเปลี่ยนจำนวนช่องในทิศทางมุมเงย θ ด้วยค่าต่างๆ ผลที่ได้แสดงได้ดังกราฟในรูปที่ 4.26



รูปที่ 4.26 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสุมเมย
เมื่อกำหนดจำนวนช่องในทิสรัศมี มีจำนวน 12 ช่อง และจำนวนช่องในทิสุมเมยแฉิมัทมีจำนวน 18 ช่อง

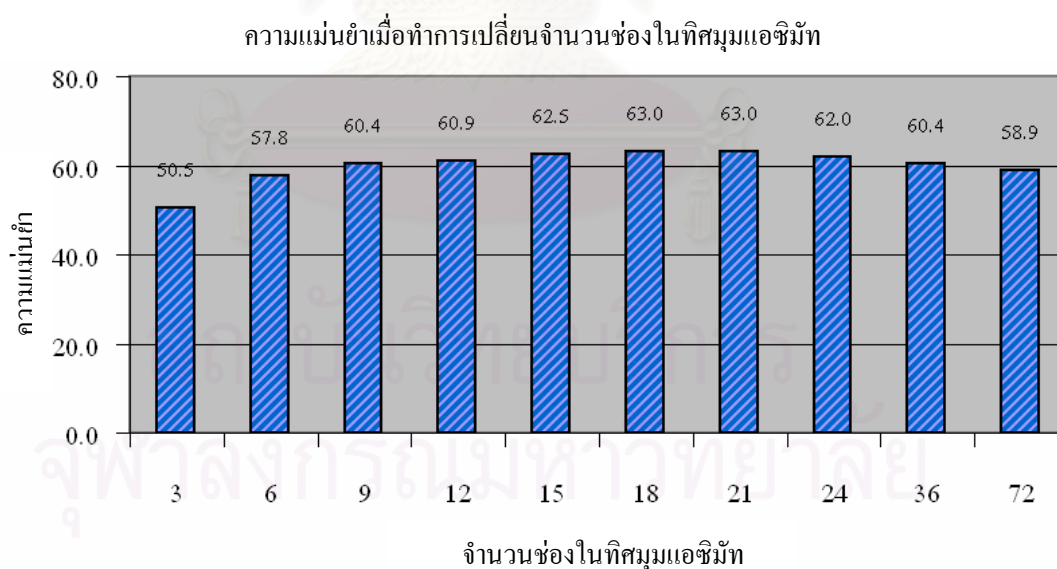
จากรูปที่ 4.26 เห็นได้ว่าความแม่นยำมีค่าน้อยเมื่อมีจำนวนช่อง 3 ช่อง และค่อนข้างคงที่เมื่อจำนวนช่องเพิ่มขึ้น โดยจำนวนช่องที่ให้ความแม่นยำสูงสุด 57.8% มี 2 ค่า คือ 6 ช่องและ 18 ช่อง

ในขั้นตอนถัดมา จึงทำการกำหนดจำนวนช่องของตารางในทิศทางของแนวลำดับ ρ ให้มีจำนวนคงที่ 12 ช่อง และ มุมเมย θ ให้มีจำนวนคงที่ 6 ช่อง และ 18 ช่อง ตามลำดับ แล้วทำการเปลี่ยนจำนวนช่องในทิสุมเมย ϕ ด้วยค่าต่างๆ ผลที่ได้แสดงได้ดังกราฟใน รูปที่ 4.27 และ รูปที่ 4.28 ตามลำดับ



รูปที่ 4. 27 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสุมมเอเชียมีท
เมื่อกำหนดจำนวนช่องในทิสุมมี มีจำนวน 12 ช่อง และจำนวนช่องในทิสุมมเอเชียมีจำนวน 6 ช่อง

จากรูปที่ 4. 27 เห็นได้ว่าความแม่นยำมีค่าน้อยเมื่อมีจำนวนช่องน้อย และความแม่นยำค่อยข้างคงที่เมื่อจำนวนช่องเพิ่มขึ้น โดยจำนวนช่องที่ทำให้ความแม่นยำสูงสุด 50.4 % มี 2 ค่า คือ 24 ช่อง และ 36 ช่อง



รูปที่ 4. 28 กราฟเปรียบเทียบความแม่นยำเมื่อทำการเปลี่ยนจำนวนช่องในทิสุมมเอเชียมีท
เมื่อกำหนดจำนวนช่องในทิสุมมี มีจำนวน 12 ช่อง และจำนวนช่องในทิสุมมเอเชียมีจำนวน 18 ช่อง

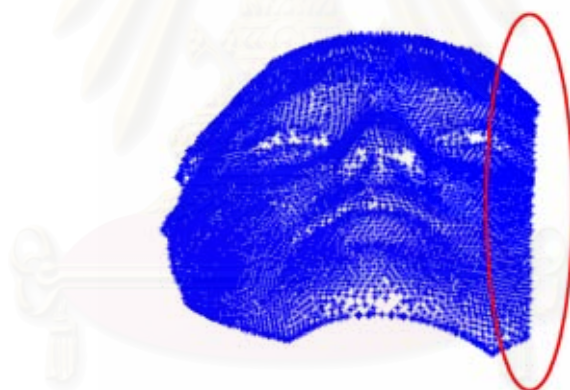
จากรูปที่ 4. 28 เห็นได้ว่าความแม่นยำมีค่าน้อยเมื่อมีจำนวนช่องน้อย และความแม่นยำค่อยข้างคงที่เมื่อจำนวนช่องเพิ่มขึ้น โดยจำนวนช่องที่ให้ความแม่นยำสูงสุด 63% มี 2 ค่า คือ 18 ช่องและ 21 ช่อง

จากกราฟในรูปที่ 4.25 - รูปที่ 4. 28 จะเห็นได้ว่า ถ้าหากช่องตารางมีจำนวนน้อยเกินไป ความแม่นยำจะลดลง โดยการเพิ่มจำนวนช่องตารางจะมีผลให้ความแม่นยำเพิ่มขึ้นและมีลักษณะค่อนข้างคงที่ โดยจำนวนช่องตารางที่ให้ค่าความแม่นยำสูงสุดคือ การแบ่งช่องในแนวรัศมี ρ 12 ช่อง การแบ่งช่องในทิสมุมเมย θ 18 ช่อง และการแบ่งจำนวนช่องในทิสมุมแอซิมัท ϕ 21 หรือ 24 ช่อง

โดยการแบ่งช่องที่มีจำนวนน้อยเกินไปนั้น ทำให้รายละเอียดข้อมูลของจุดยอดมีการรวมตัวกันในช่องใดช่องหนึ่งมากเกินไป และทำให้การแยกแยะมีประสิทธิภาพคือยกลง ในขณะที่การแบ่งช่องจำนวนมาก อาจให้ผลที่แม่นยำขึ้น แต่การแบ่งช่องที่มากเกินไปจนความจำเป็นจะทำให้ใช้เวลาในการการคำนวณช้าลงในขณะที่ผลความแม่นยำที่ได้มีค่าไม่เปลี่ยนแปลงมากนัก

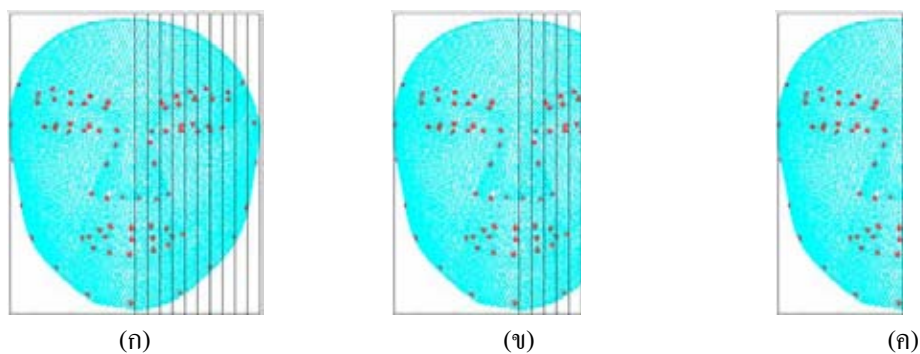
4.5. การทดลองผลความแม่นยำเมื่อจำนวนจุดยอดในภาพหายไป

ในการทดลองนี้ ทำการจำลองกรณีทีภาพใบหน้ากลุ่มทดสอบมีจุดยอดไม่เต็มภาพ โดยเป็นการจำลองกรณีทีภาพมีการแปรผันจากตำแหน่ง หรือภาพใบหน้ามีการหันไปซ้ียงทิศทางต่างๆ ซึ่งทำให้จุดยอดในมุมทีถูกบังหายไปจากภาพทีบันทึกได้ดังแสดงตัวอย่างในรูปที่ 4. 29 โดยเป็นตัวอย่างภาพทีจุดยอดของใบหน้าด้านข้างหายไปส่วนหนึ่ง ซึ่งในการทดลองนี้ทำการเปรียบเทียบความแม่นยำเมื่อจุดยอดของภาพด้านต่างๆหายไป



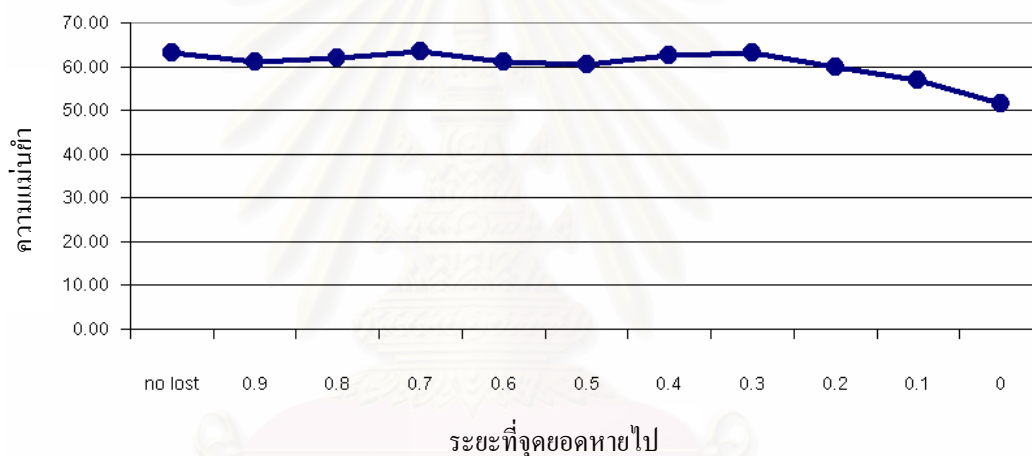
รูปที่ 4. 29 ภาพทีจุดยอดของใบหน้าด้านข้างหายไปส่วนหนึ่ง

ในขั้นตอนแรกเป็นการทดลองเมื่อจุดภาพด้านข้างหายไป โดยสมมุติว่าภาพใบหน้าซ้ายและขวามีความสมมาตรกัน ดังนั้นจะเป็นการทดลองเมื่อจำนวนจุดยอดของใบหน้าหายไปเพียงด้านเดียว เมื่อขนาดภาพมีความกว้างในระยะ $[-1,1]$ ในการทดลองทำการแบ่งระยะภาพใบหน้าด้านขวาเป็น 10 ช่วง และทำการทดลองกรณีเมื่อจุดยอดของภาพใบหน้ามีจำนวนจุดหายไปเมื่อจุดยอดนั้นมีตำแหน่งมากกว่า 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, และ 0 ตามลำดับ ดังแสดงในรูปที่ 4. 30



รูปที่ 4. 30 (ก) ภาพใบหน้าซึ่งขวามีการแบ่งช่วง (ข) ภาพใบหน้าที่ถูกขยอด้้นขวามีระยะมากกว่า 0.5 หายไป
(ค) ภาพใบหน้าที่ถูกขยอด้้นขวามีระยะมากกว่า 0 หายไป

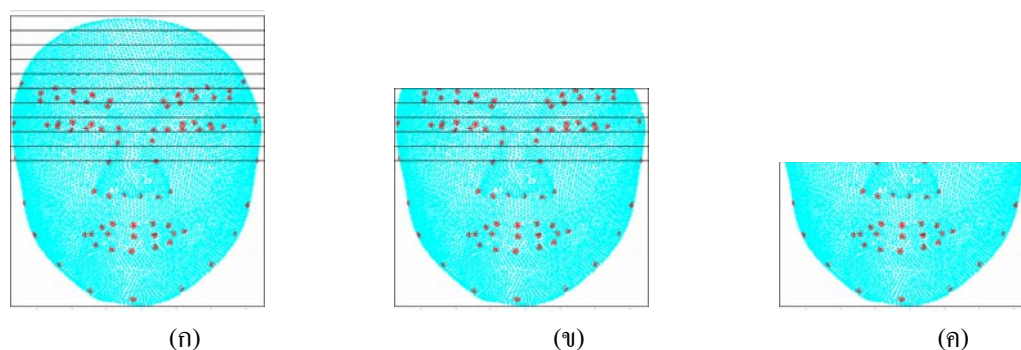
โดยผลการทดลองเมื่อขยอของซีกหน้าฝั่งขวาหายไปสามารถแสดงได้ดังกราฟในรูปที่ 4.31
ผลความแม่นยำเมื่อจำนวนจุดขยอด้้นข้างของภาพใบหน้าหายไป



รูปที่ 4.31 กราฟผลความแม่นยำเมื่อจำนวนจุดขยอด้้นข้างของภาพใบหน้าหายไป

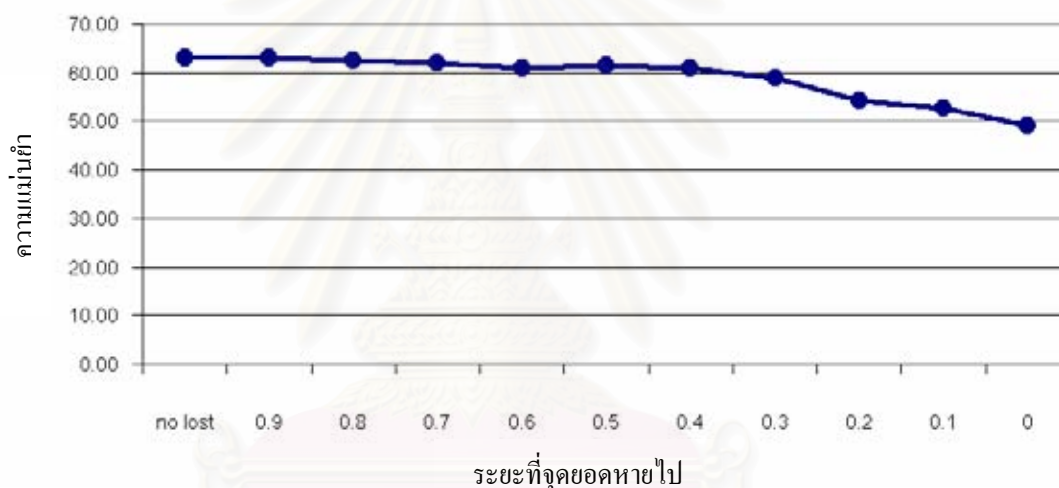
จากกราฟในรูปที่ 4.31 พบว่าความแม่นยำเมื่อเทียบกับภาพที่ไม่มีจุดขยอหายไปแล้วความแม่นยำก่อนข้างคงที่ในระยะแรก และลดลงเล็กน้อยเมื่อจุดขยอหายไปมากขึ้น ซึ่งแสดงให้เห็นว่าการใช้บริบทรูปร่างในกระบวนการรู้จำใบหน้าทนต่อการแปรผันจากตำแหน่งซึ่งทำให้จำนวนจุดขยอด้้นข้างหายไป

ในขั้นตอนถัดไปเป็นการทดลองเมื่อจุดภาพด้้นบนหายไป เมื่อขนาดภาพมีความยาวในระยะ $[-1,1]$ โดยช่วง $[1,0]$ เป็นภาพส่วนบนของใบหน้า ในการทดลองทำการแบ่งระยะภาพใบหน้าด้้นขวาเป็น 10 ช่วง และทำการทดลองกรณีเมื่อจุดขยอของภาพใบหน้ามีจำนวนจุดหายไปเมื่อจุดขยอนั้นมีตำแหน่งมากกว่า 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, และ 0 ตามลำดับ ดังแสดงในรูปที่ 4. 32



รูปที่ 4.32 (ก) ภาพใบหน้าด้านบนที่มีการแบ่งช่วง (ข) ภาพใบหน้าที่จุดยอดด้านบนมีระยะมากกว่า 0.5 หายไป (ค) ภาพใบหน้าที่จุดยอดด้านบนมีระยะมากกว่า 0 หายไป

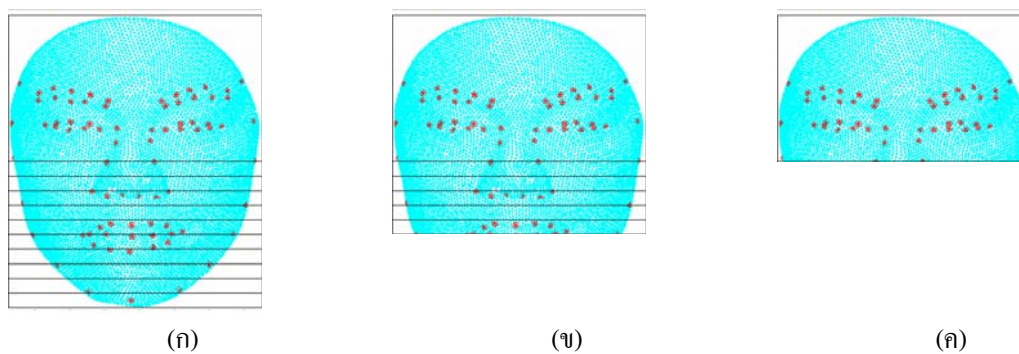
โดยผลการทดลองเมื่อจุดยอดของซีกหน้าฝั่งขวาหายไปสามารถแสดงได้ดังกราฟในรูปที่ 4.33 ผลความแม่นยำเมื่อจำนวนจุดยอดด้านบนของภาพใบหน้าหายไป



รูปที่ 4.33 กราฟผลความแม่นยำเมื่อจำนวนจุดยอดด้านบนของภาพใบหน้าหายไป

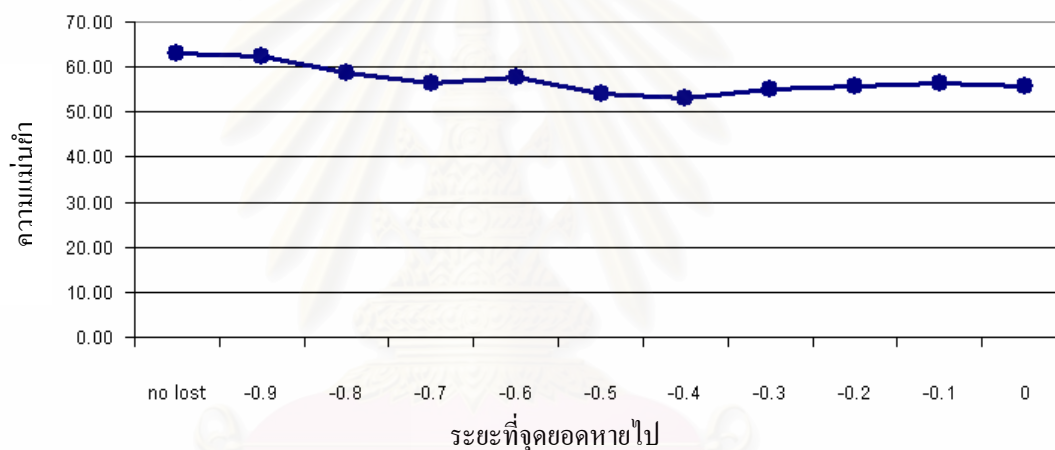
จากกราฟในรูปที่ 4.33 พบว่าความแม่นยำเมื่อเทียบกับภาพที่ไม่มีจุดยอดหายไปแล้วความแม่นยำมีค่าลดลงเล็กน้อยในช่วงแรกและลดลงมากขึ้นเมื่อจุดยอดหายไปมากขึ้น ซึ่งแสดงให้เห็นว่าการใช้บริบทรูปร่างในกระบวนการรู้จำใบหน้าทนต่อการแปรผันจากตำแหน่งซึ่งทำให้จำนวนจุดยอดด้านบนหายไปได้ในระดับหนึ่ง

ในขั้นตอนสุดท้ายเป็นการทดลองเมื่อจุดภาพด้านล่างหายไปเมื่อขนาดภาพมีความยาวในระยะ $[-1,1]$ โดยช่วง $[-1,0]$ เป็นภาพส่วนล่างของใบหน้า ในการทดลองทำการแบ่งระยะภาพใบหน้าด้านขวาเป็น 10 ช่วง และทำการทดลองกรณีเมื่อจุดยอดของภาพใบหน้ามีจำนวนจุดหายไปเมื่อจุดยอดนั้นมีตำแหน่งมากกว่า 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, และ 0 ตามลำดับ ดังแสดงในรูปที่ 4.30



รูปที่ 4.34 (ก) ภาพใบหน้าด้านล่างที่มีการแบ่งช่วง (ข) ภาพใบหน้าที่จุดยอดด้านล่างมีระยะน้อยกว่า -0.5 หายไป
(ค) ภาพใบหน้าที่จุดยอดด้านล่างมีระยะน้อยกว่า 0 หายไป

โดยผลการทดลองเมื่อจุดยอดของซีกหน้าฝั่งขวาหายไปสามารถแสดงได้ดังกราฟในรูปที่ 4.35
ผลความแม่นยำเมื่อจำนวนจุดยอดด้านล่างของภาพใบหน้าหายไป



รูปที่ 4.35 กราฟผลความแม่นยำเมื่อจำนวนจุดยอดด้านล่างของภาพใบหน้าหายไป

จากกราฟในรูปที่ 4.35 พบว่าความแม่นยำเมื่อเทียบกับภาพที่ไม่มีจุดยอดหายไปแล้วความแม่นยำมีค่าลดลงเมื่อจุดยอดหายไปมากขึ้น ซึ่งแสดงให้เห็นว่าการใช้บริบทรูปร่างในกระบวนการรู้จำใบหน้าแทนต่อการแปรผันจากตำแหน่งซึ่งทำให้จำนวนจุดยอดด้านล่างหายไปได้บ้าง

4.6. การทดลองเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีที่นำเสนอ

ในการทดลองนี้ ทำการเปรียบเทียบความแม่นยำของกระบวนการรู้จำใบหน้ากับขั้นตอนวิธีแบบอื่น โดยได้ทำการเปรียบเทียบกับขั้นตอนวิธีที่นำเสนอโดย Moreno และคณะ [17] ซึ่งเป็นการวิเคราะห์ภาพใบหน้าด้วยวิธีอิงความโค้งจากพื้นผิวของภาพวัตถุ และขั้นตอนวิธีที่นำเสนอโดย Heshar และคณะ [18] ซึ่งเป็นวิธีการวิเคราะห์เวกเตอร์แกนหลักของภาพ กับขั้นตอนวิธีการรู้จำใบหน้าโดยฮิสโทแกรมของจุดยอด โดยมีข้อจำกัดในการทดลองเหมือนกัน กล่าวคือ ใช้ภาพใบหน้าจากฐานข้อมูล BU-3DFE 8 คน แบ่งเป็นภาพกลุ่มทดลองจำนวน 1 ภาพต่อบุคคล ซึ่งรวมเป็นภาพกลุ่มทดสอบ 8 ภาพ และเป็นภาพกลุ่มทดสอบจำนวน 24 ภาพต่อบุคคล ซึ่งรวมเป็นภาพกลุ่มทดสอบ 192 ภาพ จากนั้นได้ทำการทดลองโดยเพิ่มภาพในกลุ่มทดลองโดยสุ่มภาพที่มีการแสดงออกทางอารมณ์มาเป็นภาพทดลองจำนวน 3 ภาพ 6 ภาพ และ 12 ภาพต่อบุคคล ซึ่งรวมเป็นภาพกลุ่มทดสอบ 24 ภาพ 48 ภาพ และ 96 ภาพ ตามลำดับ อีกทั้งทำการทดสอบการแยกแยะภาพด้วยสายตาโดยให้บุคคล 10 คน ทำการทดลองแยกแยะภาพจุดยอดของใบหน้าจากฐานข้อมูลทั้ง 200 ภาพ โดยให้เปรียบเทียบกับภาพกลุ่มทดลอง 8 ภาพ ซึ่งผลการเปรียบเทียบแสดงได้ดังตาราง 4.5

ตาราง 4.5 ผลการเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีแบบต่างๆ

วิธีการ	ความแม่นยำ เมื่อมีจำนวนภาพกลุ่มทดลองต่างกัน			
	1 ภาพ	3 ภาพ	6 ภาพ	12 ภาพ
กระบวนการรู้จำใบหน้าของ Moreno และคณะ	36.45%	37.5%	38.8%	38.46%
กระบวนการรู้จำใบหน้าของ Heshar และคณะ	40.63%	53.98%	53.91%	79.81%
กระบวนการรู้จำใบหน้าโดยฮิสโทแกรมของจุดยอด	63.02%	74.43%	84.21%	89.42%
กระบวนการรู้จำใบหน้าโดยการแยกแยะด้วยมนุษย์	32.2%			

จากตาราง 4.5 พบว่า กระบวนการรู้จำใบหน้าของ Moreno และคณะ ซึ่งเป็นกระบวนการแยกแยะภาพโดยใช้ข้อมูลความโค้งของพื้นผิวภาพบริเวณพื้นที่ส่วนกลางใบหน้า คือ บริเวณจมูกและดวงตา มาสร้างเป็นคุณลักษณะสำหรับการแยกแยะความแตกต่าง ดังนั้นในการทดลองกับฐานข้อมูลที่มีการเปลี่ยนแปลงทางอารมณ์ที่มักมีการเปลี่ยนแปลงที่บริเวณปากและโครงหน้า จึงทำให้จำนวนภาพทดสอบเพิ่มมากขึ้นไม่ช่วยเพิ่มความแม่นยำได้มากนัก ขณะที่กระบวนการรู้จำใบหน้าของ Heshar และคณะ ซึ่งเป็นวิธีการแยกแยะภาพโดยใช้ข้อมูลความลึกของภาพทั้งภาพ มีความแม่นยำเพิ่มขึ้นเมื่อมีภาพทดลองมากขึ้น เนื่องจากการเพิ่มจำนวนภาพทดลองช่วยให้ระบบสามารถเรียนรู้ลักษณะการแสดงออกทางใบหน้าที่แตกต่างกันได้ ส่วนการแยกแยะด้วยมนุษย์นั้น ให้ผลความแม่นยำโดยเฉลี่ยที่ 32.2% เนื่องจากการมองภาพของมนุษย์นั้นสามารถแยกแยะภาพที่มีลักษณะเป็นพื้นผิวของภาพได้ดีกว่าภาพที่เป็นลักษณะของจุดยอดซึ่งมีความเหมาะสมกับการประมวลผลด้วยคอมพิวเตอร์ แต่การแยกแยะภาพถึง 200 ภาพนั้นอาจเกิดความล้าได้

ส่วนกระบวนการรู้จำใบหน้าโดยฮิสโทแกรมของจุดยอดให้ผลความแม่นยำสูงกว่าวิธีอื่นๆ ซึ่งแสดงให้เห็นได้ว่า บริบทรูปร่างสามารถดึงข้อมูลภาพใบหน้าเพื่อใช้ในกระบวนการแยกแยะได้เป็นอย่างดี และเมื่อทำการเพิ่มจำนวนภาพในกลุ่มทดสอบเป็น 3 ภาพ 6 ภาพ และ 12 ภาพ ตามลำดับ พบว่ากระบวนการรู้จำใบหน้าโดยฮิสโทแกรมของจุดยอด ยังให้ความแม่นยำในการแยกแยะได้สูงสุด

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1. สรุปผลการวิจัย

ในวิทยานิพนธ์ฉบับนี้นำเสนอกระบวนการรู้จำใบหน้ากับภาพใบหน้า 3 มิติ โดยใช้ข้อมูลรูปร่างของภาพ 3 มิติเท่านั้น ซึ่งในการใช้ข้อมูลเฉพาะรูปร่างของใบหน้านั้น ภาพจะไม่ได้รับผลกระทบจากปัญหาการเปลี่ยนแปลงของแสง แต่จะมีผลกระทบจากการแปรผันจากการแสดงออกเนื่องจากการแสดงอารมณ์ทางใบหน้า จะทำให้ข้อมูลรูปร่างของภาพใบหน้าเปลี่ยนแปลงไป โดยขั้นตอนวิธีที่เคยมีการนำเสนอมามักมีปัญหาเมื่อภาพใบหน้าที่มีการแสดงออก

ดังนั้นในวิทยานิพนธ์ฉบับนี้จึงนำเสนอขั้นตอนวิธีแยกแยะภาพใบหน้าที่บุคคลโดยใช้บริบทรูปร่าง ซึ่งมีลักษณะเป็นฮิสโทแกรมของจุดยอดของภาพวัตถุ โดยใช้สมมุติฐานที่ว่าถ้าหากการเปลี่ยนแปลงของจุดยอดค่าผ่านช่องตารางแล้ว คุณลักษณะของภาพที่มีการแสดงออกจะไม่เปลี่ยนแปลงไปจากเดิม ซึ่งบริบทรูปร่างที่ใช้ในการคำนวณมีการปรับปรุงขั้นตอนวิธีที่มีความซับซ้อนแบบเดิมให้มีการคำนวณที่สะดวกมากขึ้น คือ มีการปรับแนวภาพก่อนและกำหนดให้ค่าของตารางมีทิศทางคงที่ จากนั้นเป็นการปรับขนาดภาพเพื่อลดความแปรปรวนจากขนาดของช่องที่แตกต่างกัน และ การกำหนดจุดบนตำแหน่งองค์ประกอบสำคัญของใบหน้าที่แน่นอนเพื่อลดขั้นตอนวิธีการหาค่าเหมาะที่สุด จากนั้นทำการทดลองกับภาพฐานข้อมูล 2 แบบ คือฐานข้อมูลภาพสังเคราะห์และฐานข้อมูลภาพใบหน้าจริง โดยแบ่งขั้นตอนวิธีการแยกแยะภาพ 2 วิธีคือ การทดลองการแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพ ซึ่งเป็นการนำฮิสโทแกรมของจุดยอดไปใช้สร้างเมตริกซ์ความคล้าย ก่อนนำเมตริกซ์ความคล้ายมาใช้ทดสอบความแตกต่างของภาพ และ การทดลองการแยกแยะภาพด้วยฮิสโทแกรมของจุดยอดบนภาพ ซึ่งเป็นการนำฮิสโทแกรมของจุดยอดของภาพไปใช้คำนวณความแตกต่างระหว่างภาพโดยตรง

จากการทดลองพบว่า การแยกแยะภาพด้วยเมตริกซ์ความคล้ายของฮิสโทแกรมของจุดยอดบนภาพให้ผลที่แม่นยำสำหรับการทดสอบกับภาพสังเคราะห์แต่กลับให้ผลการแยกแยะที่แยกรณีใช้กับภาพใบหน้าจริง ซึ่งเมื่อวิเคราะห์จากภาพเมตริกซ์ความคล้ายของภาพสังเคราะห์แล้ว จะเห็นได้ว่า วัตถุต่างชนิดกันแม้จะมีการแปลงที่แตกต่างกัน แต่ภาพเมตริกซ์ความคล้ายที่ได้จะมีความคล้ายคลึงกันมาก ซึ่งทำให้ผลการแยกแยะออกมาแม่นยำ แต่เมตริกซ์ความคล้ายของภาพใบหน้าจริงเมตริกซ์ความคล้ายของภาพมีลักษณะใกล้เคียงกันหมดกับภาพทุกภาพ ทำให้การแยกแยะภาพทำได้ไม่ดี ซึ่งปัญหานี้เกิดมาจากภาพใบหน้าที่มีลักษณะโดยรวมความคล้ายคลึงกันมาก ดังนั้นการแปลงฮิสโทแกรมของภาพเป็นเมตริกซ์ความคล้ายนั้นจะทำให้ข้อมูลรายละเอียดปลีกย่อยของภาพหายไปจากกระบวนการนี้

ดังนั้นในขั้นตอนถัดมา เราได้ทำการทดลองแยกแยะภาพใบหน้าโดยใช้ข้อมูลของฮิสโทแกรมของจุดยอดโดยตรง ซึ่งให้ผลความแม่นยำกับภาพสังเคราะห์ลดลง ซึ่งภาพที่ผิดพลาดนั้นเป็นภาพตัวหมากรุกขุนนางที่เกิดการหมุน โดยเกิดความผิดพลาดจากการปรับแนวที่ทำให้ภาพมีการเหลื่อมเล็กน้อย โดยเป็นการแยกแยะภาพผิดเป็นตัวหมากรุกพระราชา และตัวหมากรุกราชินี ที่มีรูปร่างลักษณะใกล้เคียงกับตัวหมากรุกขุนนาง ซึ่งแสดงให้เห็นว่า ค่าฮิสโทแกรมของจุดยอดนั้นมีการเก็บรายละเอียดของภาพวัตถุได้ละเอียดมากกว่าเมตริกซ์ความคล้าย ซึ่งเห็นผลได้จริงในการทดลองกับภาพจริงซึ่งให้ผลการแยกแยะที่ถูกต้องมากกว่า

จากนั้นทำการทดลองเพื่อเพิ่มประสิทธิภาพของกระบวนการแยกแยะภาพใบหน้าด้วยฮิสโทแกรมของจุดยอด โดยหาผลของการแบ่งช่องที่มีผลต่อค่าความถูกต้องในการแยกแยะซึ่งพบว่า การแบ่งจำนวนช่องของตารางมีผลต่อความแม่นยำในการแยกแยะใบหน้าบุคคล ซึ่งการแบ่งช่องที่มีความละเอียดน้อยเกินไปจะทำให้ประสิทธิภาพของการแยกแยะภาพนั้นลดลงเนื่องจากข้อมูลฮิสโทแกรมของภาพใบหน้ามีรายละเอียดไม่เพียงพอ ขณะที่การเพิ่มจำนวนช่องไม่ค่อยมีผลต่อความแม่นยำมากนัก แต่การแบ่งจำนวนช่องที่มากเกินไปจะทำให้ความเร็วในการคำนวณนั้นลดลง

ในการทดลองถัดมาเป็นการทดลองเมื่อจำนวนจุดยอดของภาพหายไป ซึ่งเป็นการจำลองกรณีการเกิดการแปรผันจากตำแหน่ง ซึ่งเมื่อภาพใบหน้ามีการหันไปในด้านต่างๆจะทำให้จุดยอดบางส่วนถูกบดบัง ซึ่งมีการทดลอง 3 กรณี คือ เมื่อจุดยอดของใบหน้าซีกขวาหายไป เมื่อจุดยอดด้านบนของใบหน้าหายไป และเมื่อจุดยอดด้านล่างของใบหน้าหายไป โดยไม่ได้ทำการทดลองกับกรณีใบหน้าซีกซ้ายหายไปเนื่องจากใบหน้าซ้ายและขวามีลักษณะสมมาตรกัน โดยกระบวนการแยกแยะจะใช้ข้อมูลส่วนที่เหลือมาใช้ในการคิดคำนวณ ซึ่งผลที่ได้แสดงให้เห็นว่าเมื่อจำนวนจุดยอดหายไปไม่มากนัก กระบวนการรู้จำใบหน้าโดยใช้ฮิสโทแกรมของจุดยอดจะมีความแม่นยำค่อนข้างคงที่และเริ่มมีค่าลดลงมากขึ้นเมื่อจุดยอดหายไปมากขึ้น ซึ่งแสดงให้เห็นว่าบริบทรูปร่างมีบทบาทต่อการแปรผันในกรณีจุดยอดบางส่วนหายไปได้ในระดับหนึ่ง

ในการทดลองสุดท้าย เป็นการเปรียบเทียบผลการแยกแยะภาพใบหน้ากับขั้นตอนวิธีแบบอื่นที่มีการนำเสนอ โดยใช้การข้อกำหนดของฐานข้อมูลแบบเดียวกัน คือเป็นฐานข้อมูลที่มีการแปรผันจากการแสดงออกสูง และใช้ภาพกลุ่มทดสอบจำนวนน้อย ซึ่งพบว่า ขั้นตอนวิธีการรู้จำใบหน้าโดยใช้เมตริกซ์ความคล้ายให้ผลการทดลองที่น้อยกว่าวิธีอื่นเนื่องจากข้อมูลภาพใบหน้าถูกลดจำนวนลงจากขั้นตอนการหาเมตริกซ์ความคล้ายจากฮิสโทแกรมของจุดยอด แต่กระบวนการรู้จำใบหน้าโดยใช้ฮิสโทแกรมของจุดยอดกลับให้ผลความแม่นยำสูงกว่าวิธีอื่นๆซึ่งแสดงให้เห็นว่า บริบทรูปร่างมีความสามารถในการสกัดคุณลักษณะของข้อมูลภาพเพื่อใช้ในกระบวนการรู้จำใบหน้าได้

5.2. ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

กระบวนการรู้จำใบหน้าแบบ 3 มิติถือเป็นเทคโนโลยีใหม่ ที่เริ่มมีการวิจัยกันไม่นานนักอันเนื่องมาจากการพัฒนาทั้งด้านอุปกรณ์และการประมวลผลภาพที่ดีมากกว่าในอดีต โดยข้อได้เปรียบของข้อมูลภาพแบบ 3 มิติที่มีจำนวนข้อมูลทั้งรูปร่างและพื้นผิว จากการทดลองจะเห็นได้ว่า บริบทรูปร่างเป็นอีกกระบวนการที่มีความเหมาะสมสำหรับดึงคุณลักษณะของภาพเพื่อใช้วิเคราะห์แยกแยะความแตกต่างสำหรับกระบวนการรู้จำใบหน้า เนื่องจากมีความสามารถในการแยกแยะความแตกต่างของภาพที่มีการแสดงออกได้ โดยใช้จำนวนภาพในกลุ่มทดสอบน้อยอีกทั้งมีการคำนวณที่ไม่ซับซ้อนนัก แต่จำเป็นต้องอาศัยการพัฒนาเพื่อเพิ่มความแม่นยำในการวิเคราะห์ภาพ เช่น การนำข้อมูลพื้นผิวซึ่งเป็นภาพใน 2 มิติเข้ามาช่วยเพิ่มความแม่นยำในการวิเคราะห์ภาพ การใช้ขั้นตอนการ นำฮิสโทแกรมของจุดยอดไปสร้างคุณลักษณะแบบอื่นๆ เป็นต้น

รายการอ้างอิง

- [1] L. Akarun, B. Gökberk, and A. A. Salah. 3D Face Recognition for Biometric Applications. Department of Computer Engineering, Istanbul, Turkey. (Unpublished Manuscript)
- [2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face Recognition : A Literature Survey. ACM Computing Surveys(CSUR) 35, 4 (December 2003): 399 - 458.
- [3] K. W. Bowyer, K. Chang, and P. Flynn. A Survey of Approaches to Three-Dimensional Face Recognition. Proceeding of 17th International Conference on Pattern Recognition (ICPR'04) 1 (2004): 358–361.
- [4] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-Invariant 3D Face Recognition. In J. Kittler and M. S. Nixon (eds.), Audio- and Video- Based Person Authentication (AVBPA2003), pp. 62 - 70. 2003.
- [5] J. A. Sethian, A Fast Marching Level Set Method for Monotonically Advance Fronts. Proceeding of the National Academy of Science 93 (February 1996).
- [6] S. Belongie, J. Malik, and J. Puzicha. Shape Matching and Object Recognition Using Shape Contexts. IEEE Transaction on Pattern Analysis and Machine Intelligence 24, 4 (April 2002): 509 - 522.
- [7] M. Turk and A. Pentland, Eigenfaces for Recognition, The Journal of Cognitive Neuroscience 3, 1 (1991).
- [8] J. Yang. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence 26, 1 (January 2004).
- [9] K. Etemad and R. Chellappa, Discriminant Analysis for Recognition of Human Face Images, Journal of the Optical Society of America A 14, 8 (August 1997).
- [10] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Transaction on Neural Networks 13, 6 (November 2002).
- [11] B. Scholkopf, A. Smola and K. R. Muller, Nonlinear Component Analysis as a Kernel Eigenvalue Problem, Max-Planck Institute, No. 44, December 1996. (Unpublished Manuscript)
- [12] L. Wiskott, J. M. Fellous, N. Kruger and C. Von Der Malsburg, Face Recognition by Elastic Bunch Graph Matching , IEEE Transaction on Pattern Analysis and Machine Intelligence 19, 7 (January 1997).
- [13] T. F. Cootes, K. Walker and C. J. Taylor, View-Based Active Appearance Models, Department of Imaging Science and Biomedical Engineering, University of Manchester, England. (Unpublished Manuscript)
- [14] V. Blanz and T. Vetter, A Morphable Model For The Synthesis Of 3D Faces, Max-Planck Institute, 1999. (Unpublished Manuscript)

- [15] R. C. Gonzalez. Digital Image Processing. Prentice Hall. 2002.
- [16] M. Chan, C. Y. Chen, G. Barton, P. Delmas, G. Gimel'farb, P. Leclercq, T. Fisher. Evaluation of 3D Face Analysis and Synthesis Techniques. IEEE International Conference on Multimedia and Expo(ICME). 2004.
- [17] A. B. Moreno, A. Sanchez, J. F. Velez, and F. J. Diaz. Face Recognition Using 3D Surface-Extracted Descriptors. Irish Machine Vision and Image Processing Conference (IMVIP 2003), September 2003.
- [18] C. Heshner, A. Srivastava, and G. Erlebacher. A Novel Technique for Face Recognition Using Range Images. 7th International Symposium on Signal Processing and Its Applications. 2003.
- [19] A. Elad. Bending Invariant Representation for Surfaces. Department of Communication Science, Technion-Israel Institute of Technology. (Unpublished Manuscript)
- [20] R. Kimmel, and J. A. Sethian. Computing Geodesic Paths on Manifold. Proceeding of the National Academy of Science 95 (July 1998).
- [21] J. A. Sethian. Level Set Methods and Fast Marching Methods [Online]. Available from: http://math.berkeley.edu/~sethian/2006/Explanations/Menu_Expanded_Overview.html
- [22] F. W. Young. Multidimensional Scaling [Online]. University of North Carolina. Available from: <http://forrest.psych.unc.edu/teaching/p208a/mds/mds.html>
- [23] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Calculus of Non-Rigid Surfaces for Geometry and Texture Manipulation. IEEE Transaction on Visualization and Computer Graphics 13,5 (September 2007).
- [24] A. Frome, D. Huber, R. Kolluri, T. BÄulow, and J. Malik. Recognizing Objects in Range Data Using Regional Point Descriptors. University of California Berkeley and Carnegie Mellon University, 2004. (Unpublished Manuscript)
- [25] R. O. Duda. Pattern Classification. Wiley-Interscience.
- [26] INRIA Gamma team research. INRIA Gamma team research database Website Collections [Machine Readable data file]. แหล่งที่มา : <http://www-c.inria.fr/gamma/download/CHESS/index3.php>.
- [27] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato. A 3D Facial Expression Database For Facial Behavior Research. 7th International Conference on Automatic Face and Gesture Recognition (FGR06). April 2006.
- [28] L. Berkovits. The Hypertext Document of the Course Matriisi-Algebra 03019P [Online]. University of Oulu. Available from: http://www.ee.oulu.fi/~mpa/matreng/ematr1_2.htm



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทความที่ได้รับการเผยแพร่

1. มนตรา บัวเพชร และ ชาญชัย ปลื้มปิติวิริยะเวช, การลดผลกระทบจากการแปลงทางเรขาคณิตในภาพวัตถุ 3 มิติโดยใช้บริบทรูปร่าง (Reducing the Effect of Geometric Transform in 3-Dimensional Images Using Shape Context), การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 30 (EECON-30, กาญจนบุรี, ประเทศไทย) , เล่มที่ ,หน้า 773- 776.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การลดผลกระทบจากการแปลงทางเรขาคณิตในภาพวัตถุ 3 มิติโดยใช้บริบทรูปร่าง Reducing the Effect of Geometric Transform in 3-Dimensional Images Using Shape Context

มาตรา นันทพร และ ชานูชัย ปรีดีปิติวิริยะเวช
ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10530
โทร 0-2218-6908 โทรสาร 0-2218-6912 E-mail: stone_course@hotmail.com, chamchai.p@eng.chula.ac.th

บทคัดย่อ

ความก้าวหน้าทางเทคโนโลยีประมวลผลภาพดิจิทัลมีการพัฒนาไปอย่างรวดเร็วเพื่อการประยุกต์ใช้งานในด้านต่างๆ จนสามารถนำเทคโนโลยีภาพ 3 มิติมาใช้แทนกันอย่างแพร่หลาย แต่ปัญหาที่เกิดขึ้นจากการแปลงทางเรขาคณิตที่เกิดขึ้นกับวัตถุทำให้การประมวลผลต่อข้อมูลภาพ 3 มิติ มีความผิดพลาดสูง บทความนี้นำเสนอขั้นตอนวิธีการใช้บริบทรูปร่างร่วมกับการวิเคราะห์องค์ประกอบหลักเพื่อสร้างภาพแบบบัญญัติที่มีความสมบูรณ์ไม่แปรเปลี่ยนต่อการแปลงของวัตถุในแบบต่างๆ ทำให้การวิเคราะห์ภาพมีความถูกต้องมากยิ่งขึ้น

คำสำคัญ: การแปลงทางเรขาคณิต, บริบทรูปร่าง, การวิเคราะห์องค์ประกอบหลัก, ภาพแบบบัญญัติ

Abstract

There is much development in the digital image processing, especially on 3 dimensional (3D) images. Many researches in recognition field pay attention in 3D image analysis because 3D images are more informative than 2D images. However, the problems caused by the geometric transformation may lead to incorrect results for analysis. In this paper, we introduce efficient invariant algorithms that overcome the geometric transformation in 3D images.

Keywords: Geometric transformation, shape context, principle component analysis, canonical image

1. คำนำ

ปัจจุบันได้มีการนำหลักการประมวลผลข้อมูลภาพดิจิทัล (digital image processing) เข้ามารวมกันใช้ในงานด้านต่างๆ อย่างแพร่หลาย ทั้งในเชิงพาณิชย์ ด้านการแพทย์ ด้านการทหารและความมั่นคงของชาติ เทคโนโลยีในปัจจุบันได้ก้าวหน้าไปมากทั้งรูปแบบขั้นตอนวิธีในการประมวลผลข้อมูลและอุปกรณ์ที่ใช้รับภาพ ซึ่งได้พัฒนาจนสามารถรับภาพแบบ 3 มิติและเริ่มมีใช้กันอย่างแพร่หลายแล้ว[1] ทำให้การวิเคราะห์ข้อมูลภาพในปริภูมิ 3 มิติ มีบทบาทสำคัญมากขึ้น เนื่องจากข้อมูลภาพแบบ 3 มิติมีความสามารถในการแสดงรายละเอียดของวัตถุได้สูงกว่าและใกล้เคียงกับวัตถุจริงมากกว่าข้อมูลภาพแบบ 2 มิติ

ในงานวิจัยเพื่อแยกแยะความแตกต่างของวัตถุในปริภูมิ 3 มิติ มักจะเกิดปัญหา เนื่องจากข้อมูลภาพ 3 มิติมีการแปลงทางเรขาคณิต (geometric transform) ได้หลากหลายแบบ ทั้งที่เป็นลักษณะการแปลงแบบวัตถุแข็งเกร็ง (rigid body transform) และ การแปลงแบบวัตถุไม่แข็งเกร็ง (non-rigid body transform) ทำให้การแยกแยะข้อมูลเกิดความผิดพลาดได้

ข้อมูลภาพ 3 มิติหนึ่ง ประกอบด้วย ไปด้วยข้อมูลที่เป็รูปร่าง (shape) ซึ่งแสดงลักษณะรูปร่างของวัตถุ และ รายละเอียดของพื้นผิว (texture) ในการเกิดการเปลี่ยนแปลง รายละเอียดของพื้นผิวมัน สามารถใช้การวิเคราะห์ข้อมูลในแบบ 2 มิติได้ แต่การเปลี่ยนแปลงของข้อมูลในส่วนขงรูปร่าง จะต้องทำการวิเคราะห์แบบ 3 มิติ เนื่องจากข้อมูลรูปร่างจะแสดงเป็นตำแหน่งพิกัดในปริภูมิ 3 มิติ

บทความฉบับนี้ แสดงลักษณะการเกิดการแปลงของวัตถุในรูปแบบต่างๆ ร่วมทั้งนำเสนอแนวทางการออกแบบขั้นตอนวิธีสร้างภาพแบบบัญญัติ (canonical image) จากสมบัติภายใน (intrinsic property) ของวัตถุ [2,3] ซึ่งมีคุณสมบัติไม่แปรเปลี่ยน (invariant) แม้จะเกิดการแปลงทางเรขาคณิตของวัตถุ โดยใช้แนวทางของขั้นตอนวิธีการวิเคราะห์องค์ประกอบหลัก (principle component analysis, PCA) เพื่อลดผลกระทบจากการแปลงแบบวัตถุแข็งเกร็ง ร่วมกับการใช้บริบทรูปร่าง (shape context) เพื่อลดผลกระทบจากการแปลงแบบวัตถุไม่แข็งเกร็ง และแสดงผลการทดลองพร้อมภาพตัวอย่างที่มีการเปลี่ยนแปลงรูปร่างในลักษณะต่างๆ ผลที่ได้นี้สามารถนำไปใช้วิเคราะห์ภาพ 3 มิติ ที่เกิดการแปลงทางเรขาคณิต เพื่อให้ผลการวิเคราะห์ที่แม่นยำและแน่นอนมากยิ่งขึ้น

2. การสร้างภาพแบบบัญญัติด้วยบริบทรูปร่าง

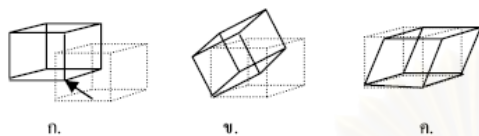
ข้อมูลภาพของวัตถุ 3 มิติแบบดิจิทัล ประกอบด้วยข้อมูลรูปร่างของวัตถุ ซึ่งแสดง รายละเอียดตำแหน่งของจุดยอด (vertex) เป็นพิกัดจุดในปริภูมิ 3 มิติเชื่อมต่อกันด้วยข้อมูลพื้นผิว (face) ซึ่งแสดงสีผิวภาพของวัตถุ และข้อมูลรายละเอียดของพื้นผิวบนผิวภาพของวัตถุนั้น ในการวิเคราะห์ข้อมูลภาพวัตถุ 3 มิติหนึ่ง คุณลักษณะสำคัญที่ใช้วิเคราะห์ภาพวัตถุ 3 มิติจะเป็นข้อมูลจากรูปร่างของวัตถุ อาทิเช่น ตำแหน่งจุดยอดแต่ละจุดในภาพ มุมระหว่างพื้นผิว หรือ ความโค้งระหว่างจุดยอดของภาพ เป็นต้น ซึ่งข้อมูลส่วนนี้จะมีการเปลี่ยนแปลงไปเมื่อเกิดการแปลงทางเรขาคณิต ส่งผลให้เกิดความผิดพลาดในการวิเคราะห์ผลข้อมูลภาพได้

DS28

2.1 การแปลงทางเรขาคณิตของวัตถุ

การแปลงทางเรขาคณิต [4] ของวัตถุเกิดเมื่อตำแหน่งของจุดยอดแต่ละจุด พื้นที่ผิว มุมระหว่างเส้นผิวของวัตถุมีการเปลี่ยนแปลงไปทำให้รูปทรงของวัตถุมีการเปลี่ยนแปลงไปด้วย แบ่งได้เป็น 2 ประเภท ดังนี้

2.1.1 การแปลงภาพของวัตถุแข็งเกร็ง การเปลี่ยนแปลงแบบนี้รูปร่างของวัตถุในภาพจะไม่เปลี่ยนแปลง กล่าวคือ มุมและความยาวของวัตถุจะไม่เปลี่ยนแปลง แต่ตำแหน่งของวัตถุทั้งชิ้นจะเปลี่ยนไป ตัวอย่างการแปลงแบบนี้ ได้แก่ การเกิดการเลื่อนขนาน (translation) และ การหมุน (rotation) ดังแสดงในรูปที่ 1



รูปที่ 1 ก. วัตถุหลังจากการเลื่อนขนาน ข. วัตถุหลังจากการหมุน ค. วัตถุหลังจากการเฉือน

2.1.2 การแปลงภาพของวัตถุไม่แข็งเกร็ง เป็นการแปลงที่ทำให้รูปร่างของภาพวัตถุมีการเปลี่ยนแปลงไปด้วย เป็นผลให้ระยะทางระหว่างจุดยอดจุดต่างๆ หรือมุมระหว่างเส้นผิวแต่ละเส้นเปลี่ยนแปลงไป ตัวอย่างการแปลงแบบนี้ได้แก่ การเฉือน (shearing) ซึ่งคล้ายกับการมีแรงมาดันส่วนหนึ่งของวัตถุ ทำให้รูปร่างของวัตถุเปลี่ยนไปในทิศทางของแรงนั้นๆ โดยตัววัตถุเองอาจไม่เกิดการเลื่อนขนาน ไปจากตำแหน่งเดิม

โดยทั่วไปข้อมูลเดิมของภาพวัตถุจะเกิดความผิดเพี้ยนไปจากเดิม เนื่องจากการแปลงของวัตถุเกิดขึ้น ทำให้ต้องหาค้นคืนวิธีดึงข้อมูลเดิมนั้นกลับมามีค่าช่วยให้การวิเคราะห์ภาพในขั้นต่อไปเป็นไปอย่างถูกต้อง การแปลงภาพของวัตถุแข็งเกร็งจะสามารถดึงข้อมูลภาพเดิมก่อนเกิดการเปลี่ยนแปลงกลับมามีค่าได้ง่าย เนื่องจากรูปร่างของวัตถุไม่ได้เกิดการเปลี่ยนแปลง ตัวอย่างขั้นตอนวิธีที่ใช้ในการแก้ไขผลที่เกิดจากการแปลงภาพของวัตถุแข็งเกร็งได้แก่ การวิเคราะห์องค์ประกอบหลัก แต่เกิดการแปลงภาพของวัตถุไม่แข็งเกร็ง ซึ่งรูปร่างของวัตถุเปลี่ยนแปลงไปด้วย ย่อมดึงข้อมูลก่อนเกิดการแปลงกลับมามีค่ายากกว่า ดังนั้นเราจึงมีแนวคิดการนำบริบทรูปร่างมาใช้ร่วมกับการวิเคราะห์องค์ประกอบหลักเพื่อลดผลกระทบในภาพที่เกิดจากการแปลงภาพของวัตถุไม่แข็งเกร็ง

2.2 การวิเคราะห์องค์ประกอบหลัก

กระบวนการวิเคราะห์องค์ประกอบหลัก [5] เป็นกระบวนการหาแกนที่มีการกระจายตัวสูงสุดของกลุ่มข้อมูล โดยจะแสดงผลของกลุ่มข้อมูลด้วยสัมประสิทธิ์ที่แสดงปริมาณข้อมูลบนแกนนั้นๆ เราเรียกแกนที่มีการกระจายตัวสูงสุดว่า เวกเตอร์ลักษณะเฉพาะ (eigenvector) และค่าสัมประสิทธิ์ของแกนนี้ว่า ค่าลักษณะเฉพาะ (eigenvalue) ซึ่งจะนำไปใช้หาคุณลักษณะ (feature) ของข้อมูล เพื่อใช้แสดงรายละเอียดของข้อมูลแทนการใช้ข้อมูลโดยตรง เนื่องจากคุณลักษณะของข้อมูลแสดงเฉพาะส่วนสำคัญๆ ของข้อมูลเท่านั้น โดยค่าส่วนที่เป็นสัญญาณรบกวนของข้อมูลจะถูกตัดทิ้งไป เมื่อแทนตำแหน่งของจุดยอดจำนวนทั้งหมด N จุดบนภาพ 3 มิติด้วยเมทริกซ์ขนาด $3 \times N$ $V = [X^T Y^T Z^T]$ เมื่อ X, Y และ Z คือพิกัดบอกตำแหน่งจุดยอดตามแกน x, y และ z บนพิกัดคาร์ทีเซียน (cartesian coordinate) ตามลำดับ แกนที่มีการกระจายตัวสูงสุดหาได้จาก

เมตริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix) ของข้อมูลจากสมการ (1)

$$G = (V - \mu)^T (V - \mu) \tag{1}$$

เมื่อให้ G คือ เมตริกซ์ความแปรปรวนร่วมเกี่ยว โดยที่ μ คือ ค่าเฉลี่ยของตำแหน่งจุดยอดของภาพสามารถหาได้จากสมการ (2)

$$\mu = \left(\frac{1}{N} \sum_{i=1}^N V_i \right) \mathbf{1} \tag{2}$$

เมื่อ V_i คือ จุดยอดที่ i ของภาพ 3 มิติ และ $\mathbf{1}$ คือเมทริกซ์ขนาด $1 \times N$ ซึ่งสมาชิกทุกตัวมีค่าเป็น 1 โดยแกนที่ได้จากกระบวนการแยกองค์ประกอบหลักนี้หากกระบวนการแยกลักษณะเฉพาะ (eigen decomposition)

$$G A_j = \lambda_j A_j \quad ; j = 1, 2, \dots, d \tag{3}$$

เมื่อ A_j คือ เวกเตอร์ลักษณะเฉพาะและ λ_j คือค่าลักษณะเฉพาะ โดยสามารถนำค่าเวกเตอร์ลักษณะเฉพาะ และ ค่าลักษณะเฉพาะ ที่มีค่าสูงสุด d ค่า มาใช้หาคุณลักษณะของข้อมูล เนื่องจากค่าเวกเตอร์ลักษณะเฉพาะ และ ค่าลักษณะเฉพาะนี้จะแสดงองค์ประกอบหลักของข้อมูล โดยหาคุณลักษณะของข้อมูล B_j ได้จาก

$$B_j = V A_j \quad ; j = 1, 2, \dots, d \tag{4}$$

จากขั้นตอนดังกล่าว จะหาแกนของข้อมูลที่มีทิศทางไปยังทิศที่ข้อมูลมีการกระจายตัวสูงสุด และค่าสัมประสิทธิ์แสดงปริมาณข้อมูลบนแกนนั้น เราสามารถนำข้อมูลสมมติในการแก้ปัญหาการเกิดการแปลงแบบวัตถุแข็งเกร็งกับภาพ 3 มิติ โดยนำข้อมูลที่ได้มาสร้างภาพใหม่ของวัตถุ (reconstruct) บนแกนลักษณะเฉพาะที่มีค่าลักษณะเฉพาะ สูงสุดแทนด้วย \tilde{V} เมื่อให้ $B = [B_1, \dots, B_d]$ และ $A = [A_1, \dots, A_d]$ ได้จากสมการ

$$\tilde{V} = B A^T \quad ; j = 1, 2, \dots, d \tag{5}$$

การสร้างภาพใหม่ของวัตถุ โดยวิธีวิเคราะห์องค์ประกอบหลักไม่สามารถแก้ปัญหาจากการแปลงของวัตถุไม่แข็งเกร็งได้ เนื่องจากวิเคราะห์เฉพาะทิศทางการวางตัวของวัตถุในภาพไม่ได้รวมถึงการเปลี่ยนแปลงระยะทางระหว่างจุดยอดจุดต่างๆ หรือมุมระหว่างเส้นผิวแต่ละเส้น ซึ่งมีผลให้รูปร่างของวัตถุเปลี่ยนแปลงไป ดังนั้นการแปลงแบบวัตถุไม่แข็งเกร็งต้องใช้การวิเคราะห์ที่รายละเอียดของข้อมูลวัตถุ เช่น การใช้บริบทรูปร่างดึงข้อมูลส่วนย่อยของภาพออกมา

2.3 บริบทรูปร่าง

เนื่องจากการแก้ปัญหารูปร่างของวัตถุเกิดการแปลงแบบวัตถุไม่แข็งเกร็งด้วยการวิเคราะห์องค์ประกอบหลัก ไม่สามารถทำได้ อย่างมีประสิทธิภาพ ดังนั้นจึงต้องนำขั้นตอนวิธีบริบทรูปร่าง [6] เข้ามาช่วยเพิ่มประสิทธิภาพการวิเคราะห์ข้อมูล โดยบริบทรูปร่างนั้นเป็นตัวบอกรูปร่างที่มีคุณสมบัติผสมผสานทั้งการแสดงข้อมูลเฉพาะที่ (local) และข้อมูลแบบครอบคลุม (global) เนื่องจากการเป็นการแสดงแต่ละจุดยอด

ของตำแหน่งข้อมูลด้วยฮิสโทแกรม (histogram) ของจุดยอดทั้งหมดในภาพ เพื่อแสดงความสัมพันธ์ของการกระจายตัวของข้อมูลทั้งหมด

2.3.1 การหาบริบทรูปร่างของจุดยอด

การหาฮิสโทแกรมของจุดยอดของข้อมูล จะใช้ฮิสโทแกรมเชิงขั้วแบบลอการิทึม (log-polar histogram) ซึ่งมีการแบ่งช่อง (bin) ของทรงกลมตามแนวรัศมี (ρ) มุมเงย (θ) และมุมเอซิมัท (ϕ) โดยแกนรัศมีแบ่งด้วยมาตราส่วนลอการิทึม เพื่อให้ความสัมพันธ์กับจุดข้อมูลที่อยู่ใกล้มากกว่าจุดข้อมูลที่อยู่ไกล เราสามารถแสดงการหาฮิสโทแกรมของข้อมูล $h_i(k)$ ช่องที่ k ของแต่ละจุดยอด p_i ได้ดังสมการ

$$h_i(k) = \#\{p \neq p_i : (p - p_i) \in bin(k)\} \quad (6)$$



รูปที่ 2 โครงสร้างของฮิสโทแกรมเชิงขั้วแบบ 3 มิติ

เมื่อ $\#$ หมายถึงจำนวนจุดยอด ในการหาฮิสโทแกรมของข้อมูลจำนวน N จุด โดยแบ่งช่องตามมุมเงยเป็นช่องย่อยจำนวน S ช่อง แบ่งมุมเอซิมัท เป็นช่องย่อยจำนวน T ช่อง และตามแกนรัศมีเป็นช่องย่อยจำนวน U ช่อง ดังนั้น ภาพ 3 มิติแต่ละภาพจะมีข้อมูลฮิสโทแกรมเป็นเมทริกซ์ขนาด $S \times T \times U \times N$ มิติ

2.3.2 การทดสอบโคสเคอร์ (χ^2 -test) ระหว่างฮิสโทแกรม

จากกระบวนการขั้นต้น เราจะได้ฮิสโทแกรมของจุดยอดแต่ละจุด แสดงข้อมูลความสัมพันธ์ระหว่างจุดนั้นๆ กับจุดรอบข้างทั้งหมด ซึ่งฮิสโทแกรม ที่ได้นี้จะมีคุณสมบัติไม่แปรเปลี่ยนไปตามการแปลงของวัตถุไม่เชิงเกร็งกับรูปร่างของวัตถุ แต่เนื่องจากข้อมูลที่ได้มีลักษณะเป็นข้อมูล 4 มิติ การเปรียบเทียบความแตกต่างฮิสโทแกรม โดยตรงจะทำได้ยาก เราจึงใช้การทดสอบโคสเคอร์ [7] ระหว่างแต่ละคู่จุด ในภาพมาหาความแตกต่างแทน ค่าที่ได้นี้จะเปรียบได้กับระยะห่างระหว่างฮิสโทแกรมแต่ละฮิสโทแกรม ซึ่งแสดงการหาได้ดังสมการ

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (7)$$

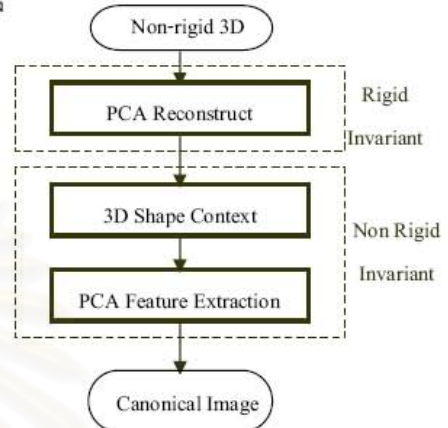
โดยที่ C_{ij} คือ ค่าโคสเคอร์ระหว่างฮิสโทแกรมของจุด p_i และ q_j เมื่อหาระยะห่างระหว่างฮิสโทแกรมของทุกคู่จุดและหาค่าเรียงเป็นเมทริกซ์ เราเรียกเมทริกซ์ C นี้ว่า เมทริกซ์ต้นทุน (cost matrix) หาก C_{ij} มีค่ามากหมายความว่าค่าที่ได้มีค่าแตกต่างกันมาก แสดงว่า ความสัมพันธ์ของจุดคู่นั้นๆ มีความแตกต่างกันสูง ในทางกลับกันค่า C_{ij} ที่มีค่าน้อยหมายถึงจุด p_i และ q_j นั้นแตกต่างกันน้อย

2.3.3 การสร้างภาพแบบบัญญัติจากเมทริกซ์ต้นทุน

ในขั้นตอนนี้ เรานำคุณสมบัติการดึงคุณลักษณะของข้อมูลด้วยกระบวนการวิเคราะห์ห้วงองค์ประกอบหลักมาใช้ เพื่อดึงข้อมูลสมบัติภายในของภาพจากเมทริกซ์ต้นทุน โดยแทนเมทริกซ์ข้อมูลด้วย C และทำการหาคุณลักษณะของข้อมูลตามสมการที่ (1) - (4) โดยเลือกเวกเตอร์ลักษณะเฉพาะที่มีค่าลักษณะเฉพาะมากที่สุด 3 ค่า มาแสดงผลคุณลักษณะที่หาได้บนแกนภาพ 3 มิติเป็นภาพแบบบัญญัติของวัตถุ

การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 30 (EECON-30) 25-26 ตุลาคม 2550 มหาวชิวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

จากที่ได้กล่าวมาสามารถสรุปขั้นตอนต่างๆ ที่ได้ออกแบบดำเนินงานในรูปที่ 4 ในขั้นแรกนำข้อมูลจุดยอดของวัตถุ 3 มิติซึ่งอาจมีการแปลงทางเรขาคณิตแล้วนำไปผ่านกระบวนการวิเคราะห์ห้วงองค์ประกอบหลักเพื่อลดผลกระทบจากการแปลงภาพวัตถุเชิงเกร็ง หลังจากนั้นนำผลลัพธ์ที่ได้มาหาบริบทรูปร่างแล้วหาคุณลักษณะของวัตถุด้วยกระบวนการวิเคราะห์ห้วงองค์ประกอบหลักอีกครั้งหนึ่งเพื่อให้ได้ภาพแบบบัญญัติของวัตถุซึ่งไม่แปรเปลี่ยนแม้จะมีการแปลงภาพวัตถุไม่เชิงเกร็ง



รูปที่ 3 โครงสร้างการทำงานของขั้นตอนวิธี

2.3.4 การหาความแตกต่างด้วยระยะยูคลิด

เมื่อผ่านขั้นตอนต่างๆ เราจะได้สร้างภาพแบบบัญญัติของวัตถุ ซึ่งมีคุณสมบัติไม่แปรเปลี่ยนตามการแปลงรูปร่างของวัตถุ เราหาความแตกต่างระหว่างผลที่ได้ด้วยการหาระยะห่างระหว่างวัตถุโดยใช้ระยะทางยูคลิด (Euclidean distance) L_2 เมื่อ P และ Q คือ เซตของจุดยอดของภาพวัตถุ 3 มิติ 2 ภาพ ตามลำดับดังสมการ

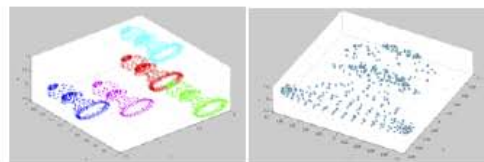
$$L_2(P, Q) = \frac{1}{N} \sum_{i=1}^N \sqrt{(p_i - q_i)^2} \quad (8)$$

3. ผลการทดลอง

ในการทดลอง เราใช้ภาพ 3 มิติของตัวหมากกรุกที่สร้างขึ้นในรูปแบบไฟล์ .wat ที่มีจุดยอดทั้งหมด 416 จุด[8] โดยแบ่งออกเป็นการทดสอบด้วยการวิเคราะห์ห้วงองค์ประกอบหลักกับภาพ 3 มิติ เมื่อเกิดการแปลงทางเรขาคณิตของภาพวัตถุแบบวัตถุเชิงเกร็ง 2 อย่าง ได้แก่ การเลื่อนขนาน และการหมุนของวัตถุ และ ทดสอบผลการทำงานของบริบทรูปร่างร่วมกับการวิเคราะห์ห้วงองค์ประกอบหลักเมื่อเกิดการแปลงภาพของวัตถุไม่เชิงเกร็ง ได้แก่ การเกิดการเลื่อน

3.1 ผลการทดลองเมื่อภาพมีการเลื่อนขนาน

ในการทดลองแบบจำลองการเลื่อนขนาน ของวัตถุไปตามทิศ (1, 0, 0), (0, 1, 0), (0, 0, 1) และ (1, 1, 1) ตามลำดับและใช้การวิเคราะห์ห้วงองค์ประกอบหลักกับภาพทดลอง ซึ่งแสดงได้ดังภาพ

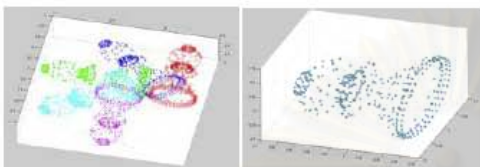


รูปที่ 4 ภาพการเกิดการเลื่อนขนานของวัตถุในทิศต่างๆ (ซ้าย) และ ภาพหลังการผ่านขั้นตอนวิธีวิเคราะห์ห้วงองค์ประกอบหลัก(ขวา)

จากภาพจะเห็นว่าเมื่อนำข้อมูลภาพที่มีการเลื่อนตำแหน่งไปตั้งตำแหน่งต่างๆ มาผ่านขั้นตอนวิธีวิเคราะห์องค์ประกอบหลัก ผลที่ได้คือภาพทั้งหมดเลื่อนมาที่ตำแหน่งเดียวกัน (ภาพทุกภาพซ้อนอยู่ที่จุดเดียวกัน) ทั้งหมด แสดงว่าการวิเคราะห์องค์ประกอบหลักสามารถแก้ปัญหาการเกิดการเลื่อนขนานของวัตถุได้เป็นอย่างดี

3.2 การทดลองเมื่อภาพมีการหมุนของวัตถุในมุมต่างๆ

ในการทดลองแบบจำลองการหมุนของวัตถุไปคามติศทวนเข็มนาฬิกาด้วยมุม 45° 90° 135° และ 180° ตามลำดับ และใช้การวิเคราะห์องค์ประกอบหลักกับภาพทดลองซึ่งแสดงได้ดังรูปที่ 6

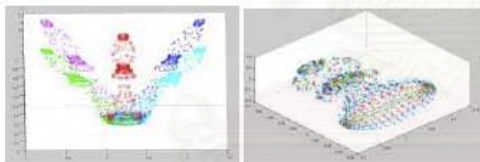


รูปที่ 5 ภาพการเกิดการหมุนของวัตถุในมุมต่างๆ(ซ้าย) และภาพหลังการผ่านขั้นตอนวิธีวิเคราะห์องค์ประกอบหลัก(ขวา)

จะเห็นว่าเมื่อนำข้อมูลภาพที่มีการหมุนไปบันทึกค่าต่างๆมาผ่านขั้นตอนวิธีวิเคราะห์องค์ประกอบหลัก ผลที่ได้คือ ภาพที่มีตำแหน่งอยู่ที่จุดเดียวกันทั้งหมด (ภาพซ้อนอยู่ที่จุดเดียวกัน) ซึ่งจากการทดลองจะเห็นได้ว่าขั้นตอนวิธีวิเคราะห์องค์ประกอบหลักสามารถแก้ปัญหการเกิดการแปลงแบบวัตถุแข็งเกร็งได้ดี

3.3 การทดลองเมื่อภาพมีการเอียง

ในการทดลองแบบจำลองการเกิดเรงเดือนวัตถุไปยังทิศต่างๆ โดยใช้การวิเคราะห์องค์ประกอบหลักเพียงอย่างเดียว ดังรูปที่ 7

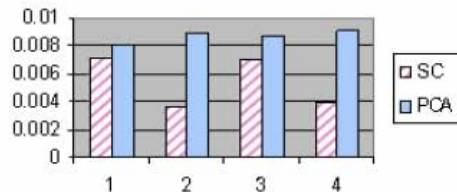


รูปที่ 6 ภาพการเกิดการเอียงของวัตถุในทิศต่างๆ(ซ้าย) และ ภาพหลังการผ่านขั้นตอนวิธีวิเคราะห์องค์ประกอบหลัก(ขวา)

จะเห็นได้ว่า ผลของภาพที่เกิดการเอียงและภาพหลังจากผ่านกระบวนการวิเคราะห์องค์ประกอบหลักนั้นจุดแต่ละจุดไม่ซ้อนกันซึ่งต่างจากภาพที่เกิดการเอียงหรือการหมุน ดังนั้นเราจึงนำวิธีปรับรูปร่างเข้ามาช่วยเพื่อลดผลจากความคลาดเคลื่อนจากการเอียง โดยได้ผลดังแสดงในรูปที่ 8



รูปที่ 7 ภาพหลังผ่านขั้นตอนวิธีการวิเคราะห์องค์ประกอบหลักและปรับรูปร่างตามลำดับ



รูปที่ 8 ความแตกต่างระหว่างวัตถุเมื่อใช้การวิเคราะห์องค์ประกอบหลักเพียงอย่างเดียวเปรียบเทียบกับการใช้ปรับรูปร่างเข้าช่วย

เมื่อนำผลภาพแบบบัญญัติจากการวิเคราะห์องค์ประกอบหลักในรูปที่ 6 มาหารระยะห่างจากภาพเดิมที่ไม่ได้ผ่านการแปลงทางเรขาคณิตใดๆ มาตามสมการ (8) และเปรียบเทียบกับภาพแบบบัญญัติที่ได้จากการหาปรับรูปร่าง พบว่ามีระยะห่างจากภาพเดิมน้อยกว่าการวิเคราะห์องค์ประกอบหลักเพียงอย่างเดียวดังแสดงในรูปที่ 8 นั้นหมายความว่าปรับรูปร่างช่วยลดผลการเอียงที่เกิดขึ้นในภาพได้ดี และคาดการณ์ได้ว่า ขั้นตอนที่น่าเสนอสามารถช่วยลดผลการแปลงทางเรขาคณิตของวัตถุไม่แข็งเกร็งแบบอื่นๆได้เช่นกัน

4. สรุป

เรานำเสนอขั้นตอนวิธีการลดผลของการเกิดการแปลงของวัตถุแบบวัตถุแข็งเกร็งโดยการใช้ขั้นตอนวิธีการวิเคราะห์องค์ประกอบหลัก และการเปลี่ยนแปลงรูปร่างของวัตถุแบบวัตถุไม่แข็งเกร็ง โดยการสร้างภาพแบบบัญญัติจากการหาปรับรูปร่างของวัตถุซึ่งแสดงผลแล้วพบว่าช่วยลดผลการแปลงของวัตถุได้อย่างมีประสิทธิภาพยิ่งขึ้น

เอกสารอ้างอิง

[1] K.W. Bowyer, "A Survey of Approaches to Three-Dimensional Face Recognition", Proc.of ICPR,2004.
 [2] A.M. Bronstein, "Expression-Invariant 3D Face Recognition", Springer-Verlag Berlin Heidelberg, 2003.
 [3] A. Elad, "Bending Invariant Representation for Surfaces", Dept.of Com. Sci.,Technion-Israel Institute of Technology.
 [4] R.C. Gonzalez, "Digital Image Processing", Prentice Hall, 2002, p.270-276.
 [5] J. Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE Tran.on PAMI, Vol.26, No.1, January, 2004.
 [6] S. Belongie, "Shape Matching and Object Recognition Using Shape Context", IEEE Tran.on PAMI, Vol.24, No.24, April, 2002.
 [7] R.O. Duda, "Pattern Classification", Wiley-Interscience, p.628-630.
 [8] <http://www-c.inria.fr/gamma/download/CHESS/index3.php>



มนตรา บัวเพชร สำเร็จการศึกษาระดับปริญญาตรี สาขาวิศวกรรมไฟฟ้าจากจุฬาลงกรณ์มหาวิทยาลัย พ.ศ.2549 และปัจจุบันกำลังศึกษาระดับปริญญาโท สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย งานวิจัยที่สนใจได้แก่ Image Processing



ชาญชัย ปลื้มปิติวิริยะเวช สำเร็จการศึกษาระดับปริญญาตรี สาขาวิศวกรรมไฟฟ้า จาก University of Maryland at College Park สหรัฐอเมริกา พ.ศ.2538 และ สำเร็จการศึกษาระดับปริญญาโท และ เอก สาขา วิศวกรรมไฟฟ้าและคอมพิวเตอร์ จาก Carnegie Mellon University สหรัฐอเมริกา พ.ศ. 2546 ปัจจุบันเป็นอาจารย์ประจำภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย งานวิจัยที่สนใจได้แก่ Medical Image Processing, 3D Image Visualization, Biometric Authentication และ Face Recognition

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวมนตรา บัวเพชร เกิดเมื่อวันที่ 31 ตุลาคม พ.ศ.2527 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า จากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2548 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า สังกัดห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ที่คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2549



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย