

การสร้างภาพใบหน้าคน 3 มิติ จากภาพหลายมุมมอง



นางสาวมนัสนันท์ ศรีคำ

ศูนย์วิทยุทรัพยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2552

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

3D FACE RECONSTRUCTION FROM MULTIVIEW IMAGES



Miss Manassanan Srikham

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 2009

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การสร้างภาพใบหน้าคน 3 มิติในหลายมุมมอง

โดย

นางสาวมนัสนันท์ ศรีคำ

สาขาวิชา

วิศวกรรมไฟฟ้า


อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร.ชาญชัย ปลื้มปิติวิริยะเวช

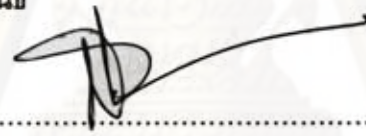
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม


ดร.ธิติพร จันทรวินเมลิอง

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นักศึกษานิพนธ์ฉบับนี้เป็นส่วน
หนึ่งของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

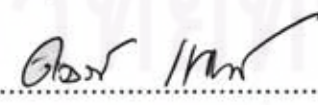

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศศิริวรงค์)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)


..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.ชาญชัย ปลื้มปิติวิริยะเวช)


..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ดร.ธิติพร จันทรวินเมลิอง)


..... กรรมการภายนอกมหาวิทยาลัย
(ดร.ดวงรัตน์ แก่นสวัสดิ์)

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

มนัสนันท์ ศรีคำ : การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง. (3D FACE RECONSTRUCTION FROM MULTIVIEW IMAGES) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ.ดร.ชาญชัย ปลื้มปิติวิริยะเวช,อ. ที่ปรึกษาวิทยานิพนธ์ร่วม: ดร.ธิดิพร จันทร์วิเนลียง, 80 หน้า.

การสร้างภาพ 3 มิติเป็นเทคนิคหนึ่งที่มีความสำคัญในด้านการประมวลผลภาพดิจิทัล การสร้างภาพ 3 มิติเป็นการฉายจุดบนระนาบ 2 มิติลงบนระนาบ 3 มิติด้วยเทคนิคต่างๆ ซึ่งในปัจจุบันมีการสร้างภาพใบหน้าคน 3 มิติด้วยการใช้เลเซอร์สแกนซึ่งมีค่าความถูกต้องแม่นยำแต่มีอุปกรณ์ที่ใช้มีราคาสูงมาก และการสร้างใบหน้าคน 3 มิติจากภาพ 2 มิติเป็นวิธีการที่มีค่าใช้จ่ายน้อยแต่มีความถูกต้องน้อยกว่า ดังนั้นนักวิจัยจึงพยายามพัฒนาการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติเพื่อให้มีประสิทธิภาพใกล้เคียงกับเลเซอร์สแกน

วิทยานิพนธ์ฉบับนี้ นำเสนอการสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองโดยกระบวนการสร้างภาพใบหน้าคน 3 มิติแบ่งออกเป็น 2 ส่วนคือการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบเอกทิวเป็นวิธีการที่มีการจำลองรูปร่างจากกลุ่มตัวอย่างภาพใบหน้าคน เพื่อใช้เป็นแบบจำลองในการแบ่งส่วนภาพของใบหน้า มีกระบวนการปรับแนว การวิเคราะห์องค์ประกอบแกนหลักและการฟิตรูปร่างของแบบจำลองเข้ากับใบหน้าคนในภาพ จากวิธีการจำลองรูปร่างแบบเอกทิวจะได้จุดบนภาพใบหน้า 2 มิติ เรานำจุดที่ได้มาสร้างภาพใบหน้าคน 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรงซึ่งเป็นวิธีการที่ไม่ยุ่งยากและมีความซับซ้อนน้อย

ภาควิชา.....วิศวกรรมไฟฟ้า.....
สาขาวิชา.....วิศวกรรมไฟฟ้า.....
ปีการศึกษา.....2552.....

ลายมือชื่อนิสิต.....มนัสนันท์ ศรีคำ.....
ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

.....
.....
.....

ศูนย์วิทยุวิทยุวิทยุ

จุฬาลงกรณ์มหาวิทยาลัย

#5070599521 : MAJOR ELECTRICAL ENGINEERING

KEYWORDS : 3D FACE RECONSTRUCTION / MULTIVIEW / ACTIVE SHAPE MODEL
/ CAMERA CALIBRATION / DIRECT LINEAR TRANSFORMATION

MANASSANAN SRIKHAM: 3D FACE RECONSTRUCTION FROM MULTIVIEW
IMAGES. THESIS ADVISOR: ASST. PROF. CHARNCHAI PLUEMPITWIRIYAJEJ,
Ph.D., THESIS CO-ADVISOR: THITIPORN CHANWIMALUANG, Ph.D., 80 pp.

3D reconstruction is one of the most important techniques in digital image processing. Points in 2D plane are projected into 3D space by using various reconstruction approaches. In the present time, there are two types of 3D face reconstruction, laser scan technique and 2D images. Although 3D face reconstruction using laser scan technique can provide high precision and accuracy, the instruments and installation are very expensive. In contrast, 3D face reconstruction from 2D images provides low cost, but at the same time, it yields low accuracy and precision. Thus, many researches have been devoted to 3D reconstruction from 2D images in order to gain performance as good as we can get from laser scan.

This thesis proposed 3D face reconstruction from multiview images by those from 2 parts. First, images are segmented by using active shape models which is obtained from face training data for face image segmentation. An active shape models consist of alignment shape, principle component analysis, and fitting shape models. The second part is 3D face reconstruction using direct linear transformation which is a strategy used to find the camera parameter self-calibrations. The algorithm retains simplicity, and at the same time, it can achieve acceptable results.

Department : Electrical Engineering..... Student's Signature Manassanan Srikham
Field of Study : Electrical Engineering..... Advisor's Signature Charnchai Pluempitwiriya
Academic Year : 2009..... Co-Advisor's Signature T. Chanwimaluang

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของอาจารย์ที่ปรึกษา วิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ชาญชัย ปลื้มปิติวิริยะเวช และอาจารย์ที่ปรึกษาวิทยานิพนธ์ ร่วม ดร.ธิติพร จันทรวีเมธียง ซึ่งให้คำปรึกษา ความช่วยเหลือและข้อคิดเห็นต่างๆ อย่างดียิ่งแก่ ผู้วิจัยมาโดยตลอด ขอขอบพระคุณอาจารย์ทุกท่านที่ช่วยประสิทธิ์ประสาทวิชาความรู้อันมีค่า ให้แก่ผู้วิจัยเป็นผลให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณทุนจากโครงการทุนสถาบันบัณฑิตวิทยาศาสตร์และเทคโนโลยีไทย (Thailand Graduate Institute of Science and Technology, TGIST) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) ที่สนับสนุน ค่าเล่าเรียน ค่าทำวิจัย และค่าใช้จ่ายส่วนตัวแก่ผู้วิจัย ระยะเวลาในการรับทุน 2 ปี ตั้งแต่วันที่ 1 มิถุนายน 2551 ถึง 31 พฤษภาคม 2553

ขอขอบคุณผู้ช่วยวิจัยจากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระยะเวลา ในการรับทุน 4 เดือน ตั้งแต่วันที่ 1 มิถุนายน ถึง กันยายน 2552

สุดท้าย ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และพี่น้องทุกคนที่เป็นกำลังใจ ดูแล เอาใจใส่ และให้ความช่วยเหลืออย่างดียิ่ง และทำยที่สุดขอขอบคุณเพื่อน ๆ ทุกคนที่ให้กำลังใจ และช่วยเหลือผู้วิจัยตลอดมา

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
สารบัญภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของวิทยานิพนธ์	4
1.3 ขอบเขตของวิทยานิพนธ์.....	4
1.4 ขั้นตอนและวิธีการดำเนินงาน	4
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	5
บทที่ 2 หลักการและทฤษฎีทั่วไป.....	6
2.1 แนวคิดพื้นฐาน	6
2.2 ระบบพิกัด.....	7
2.3 การแปลงเรขาคณิต	7
2.3.1 การเลื่อนขนาน.....	9
2.3.2 การสเกล.....	10
2.3.3 การหมุนวัตถุ.....	10
2.3.4 การแปลงแบบผสม	11
2.4 รูปทรงเรขาคณิต 1 มุมมอง.....	13
2.4.1 รูปแบบการแปลง	13
2.2.4 แบบจำลองกล้อง.....	15
2.5 เรขาคณิตหลายมุมมอง.....	22
2.5.1 เรขาคณิตอิมพีโพลาร์	22
2.6 การสร้างภาพ 3 มิติ	25
2.6.1 การสร้างภาพ 3 มิติในภาพ 3 มุมมอง	26
2.6.2 การสร้างภาพ 3 มิติในภาพ N มุมมอง	29

บทที่ 3 วิธีการที่นำเสนอ.....	30
3.1 การแบ่งส่วนภาพ.....	30
3.1.1 การกำหนดจุดสำคัญของวัตถุบนภาพ.....	30
3.1.2 แบบจำลองการแจกแจงจุด.....	31
3.1.3 กระบวนการปรับแนวของกลุ่มตัวอย่าง.....	32
3.1.4 ความแปรปรวนรูปร่างแบบจำลอง.....	34
3.1.5 การพิตแบบจำลอง.....	37
3.1.6 ฟังก์ชันการพิต.....	38
3.2 การแปลงเชิงเส้นโดยตรง.....	40
3.2.1 การเทียบวัตถุ.....	44
3.2.2 การสร้างภาพ 3 มิติ.....	45
บทที่ 4 ผลการทดลอง.....	47
4.1 อุปกรณ์ที่ใช้ในการทดลอง.....	47
4.2 การแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบเอกทิฟ.....	48
4.2.1 ฐานข้อมูลภาพ.....	48
4.2.2 การสร้างจุดของกลุ่มตัวอย่าง.....	50
4.2.3 กระบวนการปรับแนวรูปร่าง.....	50
4.2.4 การพิตรูปร่างของแบบจำลอง.....	51
4.3 การสร้างภาพใบหน้าคน 3 มิติ.....	52
4.3.1 การเทียบวัตถุ.....	52
4.3.2 การสร้างภาพ 3 มิติ.....	54
4.4 การทดลองสร้างภาพ 3 มิติ.....	55
4.4.1 การสร้างภาพ 3 มิติหลายมุมมองของกล่องสี่เหลี่ยม.....	55
4.4.2 การสร้างภาพ 3 มิติหลายมุมมองของโอริกามีรูปดาว.....	56
4.4.3 การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง.....	57
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	68
5.1 สรุปผลการวิจัย.....	68
5.2 ข้อเสนอแนะ.....	69
รายการอ้างอิง.....	70

ภาคผนวก.....	73
ภาคผนวก ก การวิเคราะห์องค์ประกอบแกนหลัก.....	74
ภาคผนวก ข ดัชนีคำศัพท์.....	76
ภาคผนวก ค บทความที่ได้รับการเผยแพร่.....	79
ประวัติผู้เขียนวิทยานิพนธ์	80



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

	หน้า
รูปที่ 2.1 ลักษณะความแตกต่างระหว่างพิกเซลกับวอกเซล.....	6
รูปที่ 2.2 ระบบพิกัดสามมิติ.....	7
รูปที่ 2.3 ส่วนประกอบต่างๆ ของระบบพิกัดโฮโมจีเนียส	9
รูปที่ 2.4 การแปลงยูคลิด	13
รูปที่ 2.5 การแปลงคล้าย.....	14
รูปที่ 2.6 การแปลงสัมพรรค.....	15
รูปที่ 2.7 การแปลงแบบฉาย	15
รูปที่ 2.8 จุดศูนย์กลางการฉายของจุดในบริเวณ P เข้าไปในระนาบภาพ	16
รูปที่ 2.9 การฉายด้วยการชดเชยของจุดศูนย์กลาง.....	17
รูปที่ 2.10 ความสัมพันธ์ของ 2 ระบบด้วยการแปลงและการหมุน.....	19
รูปที่ 2.11 ตัวอย่างลำดับของกล้อง CCD ลักษณะความแตกต่างระหว่างพิกเซลกับวอกเซล	20
รูปที่ 2.12 การบิดเบือนแสดงบนตาราง.....	21
รูปที่ 2.13 ระนาบของอทิโพลาร์จีโอเมทรี	22
รูปที่ 2.14 การสร้างภาพในอุดมคติด้วยจุดตัดของรังสีการฉายกลับ.....	25
รูปที่ 2.15 การสร้างภาพที่ไม่อยู่ในอุดมคติ.....	25
รูปที่ 2.16 ความไม่แน่นอนของจุดในระบบการสร้างภาพ 3 มิติ.....	26
รูปที่ 2.17 ความสัมพันธ์ของไตรโฟคัล	27
รูปที่ 2.18 การแปลงจุดด้วย 2 เมตริกซ์พื้นฐาน.....	28
รูปที่ 3.1 การกำหนดจุดสำคัญของวัตถุ	31
รูปที่ 3.2 ตัวอย่างของใบหน้าที่มีลักษณะต่างกัน.....	32
รูปที่ 3.3 PAC ในเวกเตอร์ 2 มิติ.....	36
รูปที่ 3.4 ขอบของวัตถุและจุดแบบจำลอง	38
รูปที่ 3.5 ตัวอย่างเส้นค้นหาจุดแต่ละจุดแบบจำลอง.....	39
รูปที่ 3.6 การเมทของจุดบนวัตถุเข้าไปในจุดภาพ.....	40
รูปที่ 3.7 ความสัมพันธ์การฉายภาพระบบจริงกับระบบภาพ	40
รูปที่ 3.8 ความสัมพันธ์ตำแหน่งการฉายและปริภูมิวัตถุ.....	41
รูปที่ 3.9 ระนาบภาพ 3 มิติ.....	41

รูปที่ 4.1 อุปกรณ์ที่ใช้ในการทดลอง.....	48
รูปที่ 4.2 ตัวอย่างรูปที่ใช้ในเป็นกลุ่มตัวอย่าง.....	49
รูปที่ 4.3 ภาพถ่ายที่ได้จากกล้องเว็บแคม.....	50
รูปที่ 4.4 จุดบนใบหน้า 300 จุด.....	50
รูปที่ 4.5 ชุดข้อมูลจุดที่ได้จากใบหน้า 20 หน้า.....	51
รูปที่ 4.6 การแบ่งส่วนภาพด้วยแบบจำลองรูปร่างแบบแอกทิฟ.....	52
รูปที่ 4.7 ชุดภาพจากกล้องแต่ละมุมมอง.....	53
รูปที่ 4.8 การอ้างอิงพิกัดในระบบ 3 มิติ.....	53
รูปที่ 4.9 ตัวอย่างพิกัดอ้างอิงในระบบ 2 มิติ.....	54
รูปที่ 4.10 การสร้างภาพ 3 มิติของกล่องสี่เหลี่ยม.....	56
รูปที่ 4.11 สร้างภาพโอริกามีรูปดาวจากภาพ 2 มุมมอง.....	57
รูปที่ 4.12 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมอง.....	58
รูปที่ 4.13 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 2 มุมมอง.....	59
รูปที่ 4.14 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 3 มุมมอง.....	60
รูปที่ 4.15 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 3 มุมมอง.....	61
รูปที่ 4.16 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 4 มุมมอง.....	62
รูปที่ 4.17 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 4 มุมมอง.....	63
รูปที่ 4.18 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 5 มุมมอง.....	64
รูปที่ 4.19 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 5 มุมมอง.....	65
รูปที่ 4.20 ภาพใบหน้าคน 3 มิติ.....	66

ศูนย์วิทยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันนี้ เทคโนโลยีการประมวลผลภาพดิจิทัล (Digital Image Processing) ได้เข้ามามีบทบาทและความสำคัญในชีวิตประจำวันมากขึ้น ทั้งทางด้านการแพทย์ที่ได้นำเทคโนโลยีภาพ 3 มิติเข้ามาช่วยในการวางแผนการผ่าตัดและการออกแบบวัสดุฝังใน การแก้ไขความผิดปกติของโครงสร้างต่าง ๆ ในร่างกาย เครื่องมือทางการแพทย์ก็ได้มีการนำเทคโนโลยีภาพ 3 มิติ เข้ามาช่วยด้วยเช่นกัน อย่างเช่นเครื่องเอกซเรย์ เครื่องสแกนคอมพิวเตอร์ (CT Scan) เครื่องอัลตราซาวนด์ 3 มิติ เครื่องฉายรังสีภาพ 3 มิติ เพื่อให้ได้ภาพที่ง่ายต่อการวินิจฉัยโรคของแพทย์ ทำให้ได้รับความถูกต้องแม่นยำมากขึ้นและทำให้ง่ายต่อการวางแผนการรักษาซึ่งเป็นผลดีต่อผู้ป่วย ทางด้านความปลอดภัยได้มีการพยายามพัฒนาการสร้างภาพ 3 มิติจากกล้องวงจรปิดเพื่อให้ง่ายต่อการสังเกตการณ์และตรวจจับสิ่งผิดปกติรวมถึงการพยายามสร้างภาพ 3 มิติของคนร้ายจากกล้องวงจรปิดด้วยเพื่อใช้เป็นหลักฐานในการระบุตัวคนร้ายได้ถูกต้องมากขึ้น และทางด้านความบันเทิงได้นำเทคโนโลยีภาพ 3 มิติเข้ามาเพื่อให้ได้ภาพที่ออกมาดูเหมือนจริงมากขึ้น ซึ่งปัจจุบันได้พยายามพัฒนาและออกแบบให้มีการสร้างภาพ 3 มิติให้มีความง่ายมากขึ้นและความซับซ้อนน้อยลง การสร้างภาพใบหน้าคน 3 มิติ (3D Face Reconstruction) ก็เป็นเทคโนโลยีที่น่าสนใจอย่างยิ่ง

อย่างไรก็ตามการนำเทคโนโลยีการสร้างภาพ 3 มิติ เข้ามาประยุกต์ใช้กับภาพใบหน้าคนนั้น มีปัจจัยหลายอย่างที่ทำให้ความแม่นยำของกระบวนการสร้างภาพใบหน้าคนได้ไม่ดีเท่าที่ควร ทั้งการเปลี่ยนแปลงของรูปร่างของภาพใบหน้า และตำแหน่งของใบหน้าในภาพ ในปัจจุบันการสร้างภาพใบหน้าคน 3 มิติได้มาจาก 2 ทางคือจากการใช้เลเซอร์สแกน (Laser Scanner) ซึ่งเป็นเทคนิคที่ใช้เครื่องมือในการทำที่มีราคาสูงมาก และการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติ เป็นวิธีการสร้างภาพที่ใช้ต้นทุนต่ำกว่าการใช้เลเซอร์สแกนมาก ถึงแม้ว่าการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติจะให้ความถูกต้องน้อยกว่าการใช้เลเซอร์สแกนก็ตาม แต่ก็ยังถือว่าคุ้มค่ามากกว่า ดังนั้นปัจจุบันจึงมีความพยายามที่จะสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติให้มีประสิทธิภาพมากขึ้น ซึ่งในวิทยานิพนธ์ฉบับนี้ก็ได้พยายามสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติ โดยกระบวนการสร้างแบบจำลองภาพ 3 มิติในตอนแรกถูกคิดค้นโดย Parke [1] การสร้างภาพใบหน้าคน 3 มิตินั้นมีการสร้างมาจากข้อมูลภาพใบหน้าคน 3 มิติที่ได้จากการใช้เลเซอร์สแกนและการใช้ข้อมูลภาพใบหน้าคน 2 มิติร่วมกัน ในงานวิจัยของ Yuankui และคณะ[2] โดยใช้คุณสมบัติคลัสเตอร์ของจุดในภาพใบหน้าคน 3 มิติและ 2 มิติด้วยวิธี Shape From Shading (SFS)

นอกจากการหาความสอดคล้องของภาพใบหน้าคน 3 มิติและ 2 มิติแล้วยังสามารถที่จะสร้างภาพใบหน้าคน 3 มิติด้วยพื้นผิวใบหน้าคน (Surface Reconstruction) ของภาพ 2 มิติได้ด้วย งานวิจัยของ Seitz และ Dyer [3] สร้างภาพ 3 มิติจากพื้นผิวภาพ 2 มิติด้วยวิธีการ Voxel Coloring, Kolmogorov และ Zabih [4] นำเสนอสร้างภาพ 3 มิติจากพื้นผิวดังด้วยวิธีการ Graph Cuts, Faugeras และ Keriven [5] นำเสนอสร้างภาพ 3 มิติด้วยวิธีการ Level Set นอกจากการสร้างภาพ 3 มิติจากพื้นผิวของภาพ 2 มิติ ยังสามารถสร้างภาพใบหน้าคน 3 มิติจากพื้นผิวของภาพ 3 มิติได้ด้วย งานวิจัยของ Clurless และ Levog [6] สร้างภาพ 3 มิติด้วยพื้นผิวภาพ 3 มิติด้วยวิธีการ Volumetric, Tang และ Medioni [7] อธิบายด้วยขั้นตอนวิธีการ Tensor Voting, Zhao et al. [8] นำเสนอด้วยวิธีการ Level Set คล้ายคลึงกับงานวิจัยของ Whitaker [9] ที่นำเสนอด้วยวิธีการ Level Set เช่นกัน จากข้อดีของการสร้างภาพ 3 มิติและ 2 มิติด้วยพื้นผิวนั้นทำให้ Lhuillier และ Quan [10] นำเสนอวิธีการสร้างภาพใบหน้าคน 3 มิติจากข้อมูลภาพ 3 มิติและ 2 มิติร่วมกัน ด้วยวิธีการ Level Set ซึ่งคล้ายคลึงกับงานวิจัยของ Zeng และคณะ [11] ที่ใช้ข้อมูลภาพใบหน้าคน 3 มิติร่วมกับ 2 มิติ นำเสนอด้วยวิธีการ Propagation งานวิจัยของ Yuxiao และคณะ [12] ก็ได้ใช้ข้อมูลภาพใบหน้าคน 3 มิติร่วมกับข้อมูลภาพ 2 มิติในการสร้างภาพ 3 มิติด้วยวิธีการ Kriging Interpolation เช่นเดียวกัน นอกจากการสร้างภาพใบหน้าคน 3 มิติด้วยพื้นผิวแล้วยังมีงานวิจัยของ Volker และ Thomas [13] ที่ใช้ความสอดคล้องกันของภาพใบหน้าคน 3 มิติและ 2 มิติ สร้างภาพใบหน้าคน 3 มิติด้วยวิธีการ Morphable งานวิจัยที่กล่าวมาเป็นการสร้างภาพใบหน้าคน 3 มิติที่เกิดจากการใช้ข้อมูลภาพ 3 มิติและ 2 มิติเข้าด้วยกัน เนื่องจากข้อมูลภาพใบหน้าคน 3 มิติได้จากการใช้เลเซอร์สแกน เราจึงไม่สามารถนำวิธีการดังกล่าวมาใช้กับวิทยานิพนธ์นี้

จากงานวิจัยของ Mark และคณะ [14] ที่ใช้การสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมองได้นำเสนอวิธีการสร้างภาพด้วย Binocular Stereo using Stereo Correspondence เป็นวิธีการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมอง โดยมีขั้นตอนเริ่มจากการคำนวณค่าพารามิเตอร์กล่องด้วยการกำหนดจุดความสอดคล้องของภาพทั้ง 2 ภาพ แล้วคำนวณค่าเมตริกซ์พื้นฐานของกล่องด้วยการประมาณค่าเชิงเส้น และหาค่าที่เหมาะสมของเมตริกซ์อีกครั้งด้วยวิธีการ Levenberg-Marquardt จากขั้นตอนนี้จะทำให้ได้ค่าพารามิเตอร์กล่อง จากนั้นทำการเข้าคู่ (Match) กันของภาพ 2 ภาพด้วยวิธีการ Sum of Absolute Difference (SAD) เป็นวิธีการเข้าคู่กันของจุดบนภาพ 2 มิติที่ได้กำหนดไว้ระหว่างภาพด้านซ้ายและภาพด้านขวา วิธีการนี้เป็นวิธีการที่ง่ายในการเข้าคู่กันของภาพ 2 ภาพตามจุดที่กำหนดไว้

การสร้างใบหน้าคน 3 มิติจากภาพ 2 มุมมองด้วยวิธีการ Binocular Stereo using Manual Triangulation เป็นอีกวิธีการหนึ่งในการสร้างภาพ 3 มิติจากภาพ 2 มุมมองวิธีการนี้เป็นวิธีการที่หาความสอดคล้องของภาพด้วยการกำหนดจุดความสอดคล้องของภาพ 2 ภาพเอง โดยมีขั้นตอนคือ ทำการเทียบวัตถุซึ่งเป็นการประมาณค่าพารามิเตอร์ภายในกล้องและพารามิเตอร์ภายนอกกล้องโดยใช้วิธีการเทียบวัตถุด้วย Tsai's Method [15] และขั้นตอนของการสร้างภาพ 3 มิติได้ใช้ความสอดคล้องของภาพ 2 มิติที่ได้จากการกำหนดจุดความสอดคล้องของภาพไว้ และได้ค่าพารามิเตอร์กล้องจากการเทียบวัตถุแล้วสร้างภาพ 3 มิติด้วยวิธีการ Triangulation

จากวิธีการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมองที่ได้กล่าวมาเป็นการสร้างภาพ 3 มิติที่ใช้จุดความสอดคล้องของภาพ 2 ภาพที่ได้จากการกำหนดจุดค่าความสอดคล้องขึ้นมาในภาพ 2 ภาพแล้วค่อยนำมาสร้างภาพ 3 มิติ และจากการใช้ภาพ 2 มิติ 2 ภาพในการสร้างภาพใบหน้าคน 3 มิตินั้นทำให้มีข้อมูลของภาพมีน้อยเกินไปดังนั้นการสร้างภาพ 3 มิติจากภาพ 2 มุมมองมีค่าความถูกต้องน้อยตามไปด้วย จากงานวิจัยที่ได้กล่าวมาได้สร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมองโดยมีการกำหนดจุดความสอดคล้องของภาพขึ้นมาทำให้เรามีความคิดที่จะสร้างจุดความสอดคล้องระหว่างภาพหลายมุมมองด้วยวิธีการอัตโนมัติ และจากการเทียบวัตถุเพื่อหาค่าพารามิเตอร์มีขั้นตอนที่ค่อนข้างยุ่งยากเราจึงได้หาวิธีการเทียบวัตถุภายในตัวเอง และจากการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมองนั้นเรามองเป็นการสร้างภาพใบหน้าคน 3 มิติที่ใช้ข้อมูลภาพ 2 มิติน้อยเกินไปทำให้ค่าความถูกต้องน้อยลงด้วย เราจึงได้ใช้ภาพหลายมุมมองในการสร้างภาพใบหน้าคน 3 มิติ ซึ่งในภาพหลายมุมมองจะทำให้เราได้ข้อมูลของภาพมากขึ้นซึ่งจะทำให้การสร้างภาพมีค่าความถูกต้องเพิ่มขึ้นด้วย

วิทยานิพนธ์ฉบับนี้ได้นำเสนอขั้นตอนวิธีการสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง โดยมีแนวคิดในการสร้างภาพใบหน้าคน 3 มิติ เริ่มจากการแบ่งส่วนภาพใบหน้าคน (Image Face Segmentation) ด้วยวิธีการจำลองรูปร่างแบบแอคทีฟ (Active Shape Model) หรือ ASM ซึ่งจากการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอคทีฟจะทำให้ได้จุดสำคัญ (Landmarks) ของภาพใบหน้าคนและวิธีการจำลองรูปร่างแบบแอคทีฟยังสามารถหาจุดภาพบนใบหน้าให้มีความสอดคล้องกัน (Correspondence) ของแต่ละมุมมองได้ซึ่งช่วยแก้ปัญหาการหาจุดความสอดคล้องของภาพ เมื่อเราได้จุดของภาพใบหน้าคนในแต่ละมุมมองมาแล้วเราจะสร้างภาพใบหน้าคน 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรง (Direct Linear transformation) หรือ DLT วิธีการนี้เป็นวิธีการที่มีการเทียบวัตถุภายในตัวเองที่ไม่มีความยุ่งยากและซับซ้อนเลย

1.2 วัตถุประสงค์ของวิทยานิพนธ์

1. เพื่อศึกษาความสำคัญของการสร้างภาพใบหน้าคน 3 มิติ
2. เพื่อออกแบบและปรับปรุงขั้นตอนวิธีการแบ่งส่วนภาพใบหน้าคน 2 มิติ
3. เพื่อศึกษาความสัมพันธ์ระหว่างภาพ 3 มิติและภาพ 2 มิติ
4. เพื่อศึกษาวิธีการหาความสอดคล้องของภาพใบหน้าคน 2 มิติหลายมุมมอง
5. เพื่อออกแบบและปรับปรุงขั้นตอนวิธีการสร้างภาพใบหน้าคนใน 3 มิติ

1.3 ขอบเขตของวิทยานิพนธ์

การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองของวิทยานิพนธ์ฉบับนี้ได้ศึกษาการสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองเราได้ภาพ 2 มิติจากการใช้กล้องเว็บแคมจำนวน 5 ตัวโดยได้วางกล้องอยู่ในระนาบเดียวกัน ภาพที่ใช้ในหาสร้างภาพ 3 มิติเราใช้ภาพ 5 มุมมองซึ่งเป็นมุมมองที่สามารถมองเห็นข้อมูลมาก และได้มีการแบ่งการทดลองออกเป็น การแบ่งส่วนภาพและการสร้างภาพใบหน้าคน 3 มิติ โดยมีขอบเขตดังนี้

1. ออกแบบขั้นตอนวิธีการแบ่งส่วนภาพใบหน้าคน 2 มิติที่สามารถใช้กับภาพหลายมุมมองได้
2. หาจุดความสอดคล้องกันของภาพใบหน้าคนของภาพหลายมุมมองโดยใช้การหาจุดความสอดคล้องด้วยวิธีอัตโนมัติ
3. วิเคราะห์หาความสัมพันธ์ของภาพใบหน้าคน 3 มิติและภาพ 2 มิติ
4. วิเคราะห์หาความสัมพันธ์ของภาพใบหน้าคน 2 มิติในหลายมุมมอง
5. หาวิธีการคำนวณค่าพารามิเตอร์กล้องที่มีความยุ่งยากซับซ้อนน้อย
6. การสร้างใบหน้า 3 มิติ โดยใช้ภาพใบหน้า 2 มิติในหลายมุมมอง

1.4 ขั้นตอนและวิธีดำเนินการ

1. ศึกษาวิเคราะห์ถึงกระบวนการของภาพใบหน้าคน 2 มิติ ที่จะนำมาสร้างภาพใบหน้าคน 3 มิติ
2. ศึกษาคุณสมบัติของวิธีการวิเคราะห์รูปร่างของภาพใบหน้าคน 2 มิติ
3. ศึกษาความสัมพันธ์ระหว่างภาพใบหน้าคน 2 มิติหลายมุมมอง
4. ศึกษาความสัมพันธ์ระหว่างภาพใบหน้าคน 2 มิติกับภาพ 3 มิติ
5. ออกแบบและศึกษาคุณสมบัติของการแบ่งส่วนภาพใบหน้าคน 2 มิติ
6. ออกแบบความสอดคล้องของภาพใบหน้าคน 2 มิติหลายมุมมอง
7. ออกแบบและศึกษาคุณสมบัติของการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติ

8. ออกแบบและศึกษาคุณสมบัติของการสร้างภาพไบหน้าคน 3 มิติจากภาพ 2 มิติหลายมุมมอง
9. ทดลองวิธีการที่ออกแบบ และเก็บผลการทดลอง รวบรวมและสรุปผลการวิจัยทั้งหมดเพื่อเขียนวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้เรียนรู้วิธีการแบ่งส่วนภาพไบหน้าคน 2 มิติ
2. ปรับปรุงวิธีการแบ่งส่วนภาพไบหน้าในภาพให้ดีขึ้น
3. ได้ศึกษาถึงความสัมพันธ์ระหว่างภาพไบหน้าคน 2 มิติในหลายมุมมอง
4. ได้ศึกษาถึงความสัมพันธ์ระหว่างภาพไบหน้าคน 3 มิติและ 2 มิติ
5. ได้ศึกษาการวิเคราะห์ความสอดคล้องของข้อมูลภาพไบหน้าคน 2 มิติ
6. ได้ออกแบบและพัฒนากระบวนการวิธีการสร้างภาพไบหน้าคน 3 มิติจากหลายมุมมอง
7. มีผลงานบทความทางวิชาการตีพิมพ์อย่างน้อย 1

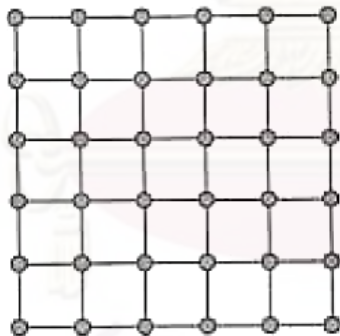
บทที่ 2

หลักการและทฤษฎีทั่วไป

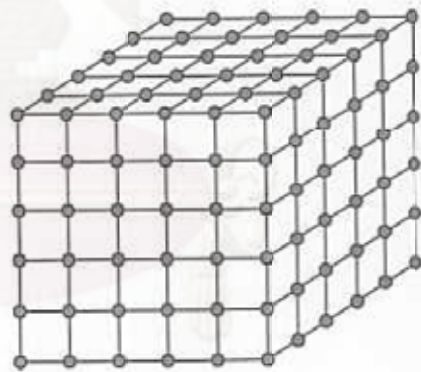
การสร้างภาพ 3 มิติ ในปัจจุบันมีการพัฒนาไปมาก ทั้งทางด้านอุปกรณ์การรับภาพ ค่ายภาพของคอมพิวเตอร์ที่ใช้ในการประมวลผลภาพ และรวมทั้งวิธีการสร้างภาพ 3 มิติ ซึ่งต้องใช้ความรู้พื้นฐานทางด้านคอมพิวเตอร์กราฟิกเข้ามาประยุกต์ใช้ ดังนั้นจึงจำเป็นอย่างยิ่งที่จะต้องศึกษาทฤษฎีทางคอมพิวเตอร์กราฟิกขั้นพื้นฐาน

2.1 แนวคิดพื้นฐาน

โดยทั่วไปภาพดิจิทัลจะประกอบไปด้วยอะเรย์สองมิติที่เป็นข้อมูลแสดงสีหรือความเข้มของแสง ซึ่งองค์ประกอบของข้อมูลเหล่านี้ถูกเรียกว่าพิกเซล (Pixel) หรือก็คือองค์ประกอบของภาพ ในทำนองเดียวกันกรณีการสร้างภาพ 3 มิติก็จะใช้อะเรย์สามมิติเป็นองค์ประกอบของภาพ ซึ่งองค์ประกอบของข้อมูลนี้เรียกว่า วอกเซล (Voxel) นิยามของวอกเซลคือจุดในปริภูมิ สามมิติซึ่งไม่มีขนาด ซึ่งเปรียบได้กับรูปทรงเรขาคณิตที่เป็นองค์ประกอบพื้นฐานในการสร้างภาพทางด้านคอมพิวเตอร์กราฟิก



(ก)



(ข)

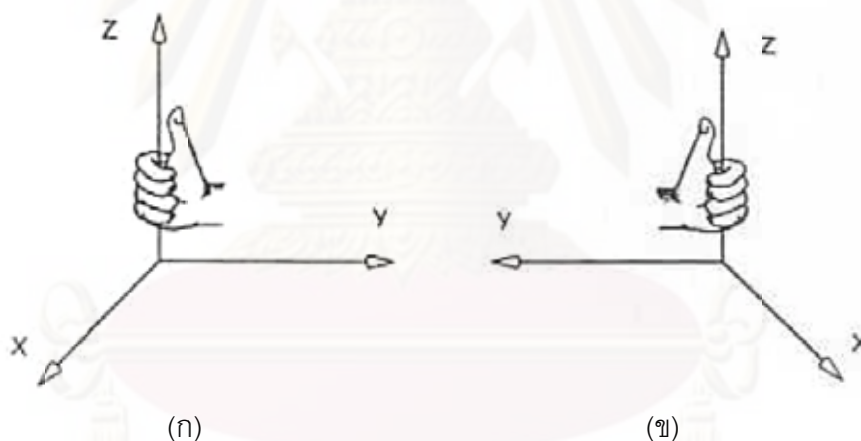
รูปที่ 2.1 ลักษณะความแตกต่างระหว่างพิกเซลกับวอกเซล

(ก) พิกเซล (ข) วอกเซล

2.2 ระบบพิกัด (Coordinate System)

เนื่องจากการระบุตำแหน่งของสิ่งใดๆ ล้วนต้องมีจุดอ้างอิงเสมอ ในทางคอมพิวเตอร์ กราฟฟิกก็เช่นกัน ระบบพิกัดเป็นสิ่งจำเป็นที่ถูกใช้ในการอ้างอิงถึงตำแหน่งของวัตถุที่อยู่ในปริภูมิ (Space) ในระบบสองมิตินั้นระบบพิกัดจะประกอบด้วยแกนอ้างอิง 2 แกน คือ แกน X และแกน Y ส่วนในระบบสามมิตินั้นแกนจะมีแกน Z เพิ่มขึ้นมาอีกหนึ่งแกน นั่นคือระบบพิกัดของระบบสามมิติ จะมีแกนอ้างอิง 3 แกน คือ แกน X แกน Y และแกน Z ระบบพิกัดสามมิติมี 2 ระบบตามทิศทางของ แกน Z ซึ่งได้แก่ ระบบพิกัดมือขวาและระบบพิกัดมือซ้าย

ระบบพิกัดมือขวามักเป็นที่นิยมใช้กันอย่างแพร่หลายในงานทางคณิตศาสตร์ ภูมิศาสตร์ วิศวกรรมศาสตร์ และการอ้างอิงตำแหน่งของวัตถุทั่วไป ส่วนระบบพิกัดมือซ้ายมักนิยมใช้กับงานทางคอมพิวเตอร์กราฟฟิกโดยจะกำหนดให้แกน Z เป็นแกนที่แสดงระยะความลึกที่เข้าไปในจอภาพและมีค่าเป็นบวก และระนาบของจอภาพเป็นระนาบ XY



รูปที่ 2.2 ระบบพิกัดสามมิติ (ก) ระบบพิกัดมือขวา (ข) ระบบพิกัดมือซ้าย

2.3 การแปลงเรขาคณิต

ในงานทางด้านคอมพิวเตอร์กราฟฟิกนั้น บางครั้งจำเป็นต้องมีการเคลื่อนย้ายภาพไปยังตำแหน่งใหม่, ย่อ-ขยายภาพ หรือการหมุนภาพให้อยู่ในมุมมองที่เหมาะสม ซึ่งสามารถทำได้โดยใช้การแปลงเรขาคณิต (Geometric Transformation)

ในการพิจารณาการแปลงนั้นจะพิจารณาได้ว่าเป็นการแปลงจุดในระนาบสามมิติ และเพื่อความเข้าใจที่ง่ายจึงมักจะเขียนอยู่ในรูปของเมตริกซ์ โดยกำหนดให้จุดที่ต้องการแปลงเป็นจุด P มีพิกัดเป็น (x, y, z) ใดๆ เมื่อต้องการเลื่อนขนานตำแหน่งของจุด P นี้ไปยังพิกัดใหม่คือ P' ซึ่งมีพิกัดเป็น (x', y', z') ก็ทำได้โดยการนำพิกัดของจุด P ไปบวกกับเมตริกซ์การแปลงซึ่งก็คือ T สามารถแสดงการแปลงได้ดังนี้

$$P' = P + T = [x \quad y \quad z] + [T_x \quad T_y \quad T_z] \quad (2.1)$$

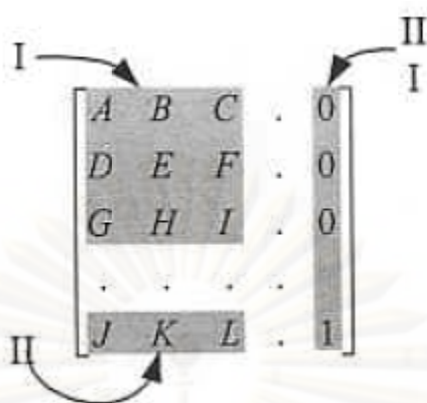
นอกจากนี้เมื่อนำพิกัดของจุด P มาคูณกับเมตริกซ์การแปลง T ให้ได้พิกัดใหม่คือ P' นั้น จะเป็นการสเกล, การหมุน, การสะท้อน หรืออื่นๆ ขึ้นอยู่กับค่าพารามิเตอร์ภายในเมตริกซ์การแปลงนั้น สามารถเขียนเป็นเมตริกซ์การแปลงได้ดังนี้

$$P' = PT = [x \quad y \quad z] \cdot \begin{bmatrix} a & b & c \\ d & f & g \\ h & i & j \end{bmatrix} \quad (2.2)$$

สำหรับการแปลงเรขาคณิตในทางคอมพิวเตอร์กราฟิกนั้นโดยทั่วไปแล้วจะประกอบด้วย การแปลงหลายๆ ชนิดประกอบกัน และเนื่องจากเป็นการคำนวณในรูปของเมตริกซ์จึงทำให้ยากในการคำนวณเมตริกซ์ที่มีรูปแบบแตกต่างกัน จึงได้มีการนำระบบพิกัดโฮโมจีเนียส (Homogeneous Coordinate System) เข้ามาใช้เพื่อแก้ปัญหา

ในระบบพิกัดโฮโมจีเนียส จุดในระบบสามมิติ (x, y, z) จะถูกแทนที่ด้วยระบบสี่มิตินั้นคือ (x, y, z, h) โดยจะมี h เป็นค่าแพกเตอร์ที่มีหน้าที่สเกลซึ่งจะมีค่าไม่เท่ากับศูนย์ จุด (x, y, z, h) จะถูกนอร์มอลไลซ์ (Normalize) เป็น $(x/h, y/h, z/h, 1)$ โดยทั่วไปแล้วค่า h จะมีค่าเท่ากับ 1 ดังนั้นจุด (x, y, z) ใดๆ จะมีค่าในระบบพิกัดโฮโมจีเนียสเป็น $(x, y, z, 1)$ เพื่อความเข้าใจในระบบพิกัดโฮโมจีเนียส รูปที่ 2.3 จะบอกถึงส่วนประกอบต่างๆ ของระบบนี้

ศูนย์วิทยุทรพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.3 ส่วนประกอบต่างๆ ของระบบพิกัดโฮโมจีเนียส

จากรูปที่ 2.3 เมตริกซ์ย่อย I เป็นส่วนที่กำหนดว่าเป็นการแปลงแบบใด เช่น การหมุน, การสเกล หรือการสะท้อน เป็นต้น ส่วนเมตริกซ์ย่อย II เป็นส่วนของการย้ายตำแหน่งแบบเชิงเส้น และในส่วนสุดท้ายเมตริกซ์ย่อย III ใช้ในการรวมระหว่างการเลื่อนขนานและการแปลงที่อยู่ในรูปการคูณ โดยมีรายละเอียดของการแปลงแบบต่างๆ ดังนี้

2.3.1 การเลื่อนขนาน (Translation)

เมตริกซ์การแปลงที่ใช้สำหรับการแปลงที่ต้องการเลื่อนขนานตำแหน่งมีลักษณะดังสมการที่ (2.3) ดังนี้

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_y & T_z & 0 \end{bmatrix} \quad (2.3)$$

จะได้พิกัดของจุดในระบบโฮโมจีเนียส (x', y', z', h) ดังสมการที่ (2.4)

$$[x' \ y' \ z' \ h] = [x \ y \ z \ 1] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_y & T_z & 0 \end{bmatrix} \quad (2.4)$$

เขียนใหม่ได้ดังนี้

$$[x' \ y' \ z' \ h] = [x + T_x \ y + T_y \ z + T_z \ 1] \quad (2.5)$$

โดยที่ค่า T_x, T_y และ T_z แสดงถึงระยะทางที่ต้องการเลื่อนขนานตำแหน่งไปในแนวแกน X แกน Y และแกน Z ตามลำดับ

2.3.2 การสเกล (Scaling)

การสเกลเป็นการแปลงเพื่อปรับเปลี่ยนขนาดของวัตถุที่ต้องการ ซึ่งมีเมตริกซ์การแปลงดังนี้

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.6)$$

จะได้พิกัดของจุดในระบบโฮโมจีเนียส (x', y', z', h) ดังสมการที่ (2.7)

$$[x' \ y' \ z' \ h] = [x \ y \ z \ 1] \cdot \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.7)$$

ได้ผลลัพธ์เป็นดังสมการที่ (2.8)

$$[x' \ y' \ z' \ h] = [xS_x \ yS_y \ zS_z \ 1] \quad (2.8)$$

โดยที่ S_x, S_y และ S_z เป็นสัมประสิทธิ์การสเกลสำหรับพิกัด X, Y และ Z ตามลำดับ ค่าของสัมประสิทธิ์ทั้งสามจะเป็นตัวบอกว่าเป็นการย่อหรือขยายภาพ กล่าวคือถ้า S มีค่ามากกว่า 1 จะเป็นการขยายภาพและถ้า S มีค่าน้อยกว่า 1 จะเป็นการย่อขนาดภาพ

2.3.3 การหมุนวัตถุ (Rotation)

การหมุนวัตถุจะเป็นการหมุนวัตถุรอบแกนต่างๆ ทั้งสามแกน ซึ่งได้แก่แกน X แกน Y และแกน Z โดยจะอ้างอิงการมองในลักษณะมองออกจากจุดกำเนิดไปตามแนวแกน เมตริกซ์การแปลงของการหมุนรอบแกนต่างๆ ในทิศทางตามเข็มนาฬิกามีดังนี้

เมตริกซ์การแปลงของการหมุนรอบแกน X ไปเป็นมุม θ คือ

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.9)$$

เมตริกซ์การแปลงของการหมุนรอบแกน Y ไปเป็นมุม ϕ คือ

$$R_y = \begin{bmatrix} \cos \phi & 0 & -\sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ \sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.10)$$

เมตริกซ์การแปลงของการหมุนรอบแกน Z ไปเป็นมุม φ คือ

$$R_z = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 & 0 \\ -\sin \varphi & \cos \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.11)$$

2.3.4 การแปลงแบบผสม (Multiple Transformations)

โดยทั่วไป ในการใช้งานทางด้านคอมพิวเตอร์กราฟฟิก การหมุนภาพสามมิติมักจะใช้ในหลายๆ เมตริกซ์ร่วมกัน ซึ่งเป็นการแปลงหลายๆ ชนิดร่วมกันนั่นเอง สำหรับการหมุนในสองแกนก็สามารถหาเมตริกซ์การแปลงของการหมุนรอบสองแกนได้โดยการนำเมตริกซ์การแปลงของการหมุนรอบแกนเดียวมาคูณกันแต่ต้องพึงระวังว่าการคูณเมตริกซ์นั้น ไม่มีคุณสมบัติการสลับที่ ดังตัวอย่างต่อไปนี้

เมื่อกำหนดให้เมตริกซ์ M เป็นเมตริกซ์การแปลงแบบผสม

$$M = R_x \times R_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \phi & 0 & -\sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ \sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} \cos \phi & 0 & -\sin \phi & 0 \\ \sin \theta \sin \phi & \cos \theta & \sin \theta \cos \phi & 0 \\ \cos \theta \sin \phi & -\sin \theta & \cos \theta \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.12)$$

$$M' = R_Y \times R_X = \begin{bmatrix} \cos \phi & 0 & -\sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ \sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M' = \begin{bmatrix} \cos \phi & \sin \theta \sin \phi & -\cos \theta \sin \phi & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ \sin \phi & -\sin \theta \cos \phi & \cos \theta \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.13)$$

จะเห็นว่าเมตริกซ์ M เป็นการหมุนแกน X ก่อนแล้วจึงหมุนรอบแกน Y ส่วนเมตริกซ์ M' จะหมุนรอบแกน Y ก่อน แล้วค่อยหมุนรอบแกน X ซึ่งผลลัพธ์ที่ได้จะไม่เท่ากัน

ส่วนในกรณีที่มีการหมุนหลายแกนและมีการเลื่อนขนานตำแหน่งด้วยก็สามารถหาเมตริกซ์การแปลงได้ดังตัวอย่างนี้ โดยมีลำดับการแปลงคือจะย้ายจุดก่อนแล้วจึงหมุนในแนวแกน X และ Y ตามลำดับ จะเป็นดังนี้

$$M = [T][R_X][R_Y]$$

$$M = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_y & T_z & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \phi & 0 & -\sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ \sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.14)$$

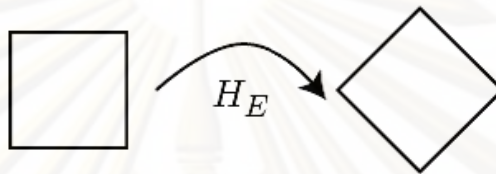
สำหรับกรณีที่มีการสเกลหรือหมุนรอบแนวแกนอื่น ก็สามารถหาเมตริกซ์การแปลงได้ในทำนองเดียวกันนี้

2.4 รูปทางเรขาคณิต 1 มุมมอง (Single View Geometry)

การแปลงเรขาคณิต (Geometric Transformations) และรูปแบบการแปลงสามารถเขียนในแบบจำลองทางคณิตศาสตร์ของกระบวนการจาก 3 มิติ เข้าไปใน ระนาบภาพ 2 มิติ ได้ดังนี้

2.4.1 รูปแบบของการแปลง

การแปลงยูคลิด (Euclidean Transformation) เป็นการแปลงที่มีขนาดเท่ากันของความยาว พื้นที่ และไม่แปรผันต่อการเลื่อนขนาน ดังรูปที่ 2.4



รูปที่ 2.4 การแปลงยูคลิด

สามารถเขียนได้ดังสมการ

$$x' = H_E x = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} x \quad (2.15)$$

ซึ่ง R คือ เมตริกซ์การหมุน (Rotation Matrix) กำหนดค่าใน 2 มิติ และ 3 มิติ ได้ดังนี้

$$R = R(\theta) \quad (2.16)$$

$$R = R_z(\theta_z)R_y(\theta_y)R_x(\theta_x) \quad (2.17)$$

และ t คือ เวกเตอร์การแปลง

เมตริกซ์การหมุนใน 2 มิติ เป็นดังนี้

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (2.18)$$

และการหมุนในระบบ 3 มิติ

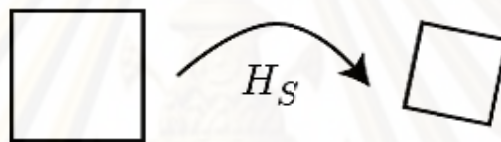
$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \quad (2.19)$$

$$R_y(\theta_y) = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \quad (2.20)$$

$$R_z(\theta_z) = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.21)$$

ที่ θ_x, θ_y และ θ_z คือ มุมของแต่ละแกน

การแปลงคล้าย (Similarity Transformation) เป็นการแปลงที่เหมือนการแปลงยูคลิด เพียงแต่มีการเพิ่มสเกล (Scaling) เข้าไปดังรูปที่ 2.5

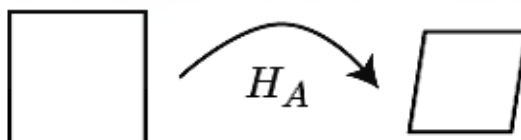


รูปที่ 2.5 การแปลงคล้าย

เขียนเมตริกซ์การแปลงคล้ายได้ดังนี้

$$x' = H_S x = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} x \quad (2.22)$$

การแปลงสัมพรรค (Affine Transformation) เป็นการแปลงแบบไม่เป็นเชิงเส้น (Non-Singular Linear) มีขนาด การสเกล การหมุนของความยาว และพื้นที่ไม่เท่ากันดังรูปที่ 2.6



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

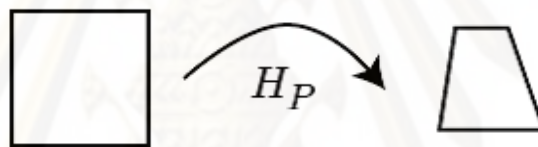
รูปที่ 2.6 การแปลงสัมพรรค

เมตริกซ์การแปลงสัมพรรคเป็นดังนี้

$$x' = H_A x = \begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix} x \quad (2.23)$$

การแปลงแบบฉาย (Projective Transformation) แสดงดังรูปที่ 2.7 และเขียนสมการเป็นดังนี้

$$x' = H_P x = \begin{bmatrix} A & t \\ v^T & 1 \end{bmatrix} x \quad (2.24)$$



รูปที่ 2.7 การแปลงแบบฉาย

2.4.2 แบบจำลองกล้อง (Camera Model)

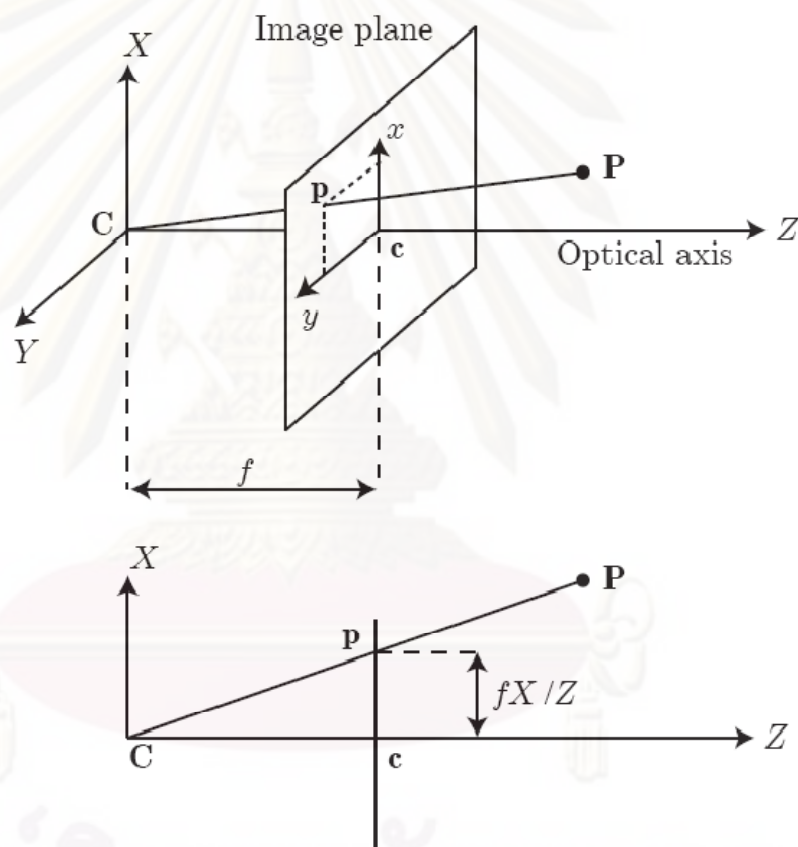
กล้องในประมวลผลภาพดิจิทัลไม่ได้เป็นแค่อุปกรณ์การบันทึกภาพเท่านั้นแต่เป็นการแมป (Mapping) ของวัตถุจริงใน 3 มิติ เข้าไปในภาพ 2 มิติ ความแตกต่างของแบบจำลองแต่ละชนิดขึ้นอยู่กับคุณสมบัติของกล้อง กล้องที่ใช้ 2 มิติทั่วไปส่วนมากใช้แบบจำลองพื้นฐานเหมือนกัน ในวิทยานิพนธ์นี้จะอธิบายถึงแบบจำลองกล้องรูเข็ม (Pinhole Camera Model)

แบบจำลองพื้นฐานกล้องรูเข็ม (Basic Pinhole Model)

การเมท (Maps) ของจุด 3 มิติ $P = (X, Y, Z)^T$ ดูได้ดังรูปที่ 2.8 เข้าไปในจุดภาพ $p = (fX/Z, fY/Z, f)^T$ การแสดงการฉายภาพของจุดเข้าไปในระนาบภาพ เราจะให้จุดศูนย์กลางการฉาย (Central Projection) จับคู่จากจุดจริง (World Point) เข้าไปในพิกัดของภาพ

$$(X, Y, Z)^T = (fX/Z, fY/Z)^T \quad (2.25)$$

รูปแบบการเมทในระบบ 3 มิติ ไประบบ 2 มิติ $\mathbb{R}^3 \mapsto \mathbb{R}^2$



รูปที่ 2.8 จุดศูนย์กลางการฉายของจุดในบริเวณ P เข้าไปในระนาบภาพ

จากรูปที่ 2.8 จุดศูนย์กลางของการฉาย C คือ จุดศูนย์กลางของกล้องหรือจุดศูนย์กลางออปติก (Optical Center) เส้นตรง Z จากจุดศูนย์กลางกล้อง C ตั้งฉากกับระนาบภาพเป็นแกนออปติก (Optical Axis) และจุดตัดบนแกนออปติกที่จุด c เรียกว่าจุดศูนย์กลาง (Center Point)

การฉายโดยใช้พิกัดโฮโมจีเนียส (Projection using Homogeneous Coordinates)

ในจุดจริง (World Points) และจุดบนภาพ (Image Points) สามารถกำหนดด้วยพิกัดโฮโมจีเนียสได้ดังสมการ

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX \\ fY \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} f & & 0 \\ & f & 0 \\ & & 1 \\ & & & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2.26)$$

การฉายของจุด 3 มิติ เข้าไปในระนาบภาพ 2 มิติได้ดังนี้

$$\lambda x = PX \quad (2.27)$$

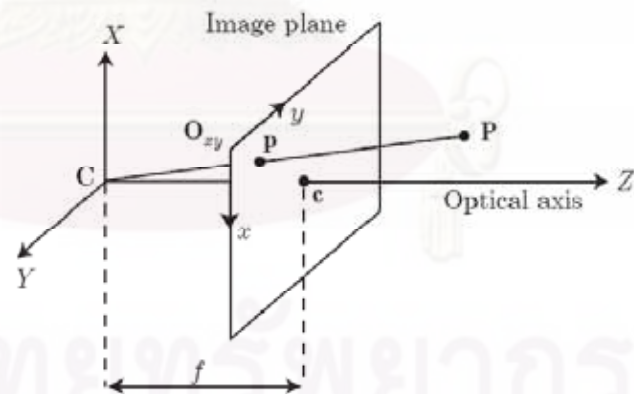
ซึ่ง $x = (x, y, 1)^T$ คือ พิกัดโฮโมจีเนียสใน 2 มิติ

$X = (X, Y, Z, 1)^T$ คือ พิกัดโฮโมจีเนียสใน 3 มิติ

λ คือ สเกลแฟกเตอร์โฮโมจีเนียส

พื้นฐานการขดเซยจุด (Principal point offset)

ในการจำลองระบบจุดศูนย์กลางที่ได้มีแนวโน้มของระบบพิกัดแบบดั้งเดิมแต่ในความเป็นจริงแล้วอาจจะไม่เกิดขึ้น



รูปที่ 2.9 การฉายด้วยการขดเซยของจุดศูนย์กลาง

จากรูปที่ 2.9 จุดศูนย์กลางมีพิกัด $c = (c_x \ c_y)$ ในระบบพิกัดภาพเราสามารถจับคู่ได้ดังนี้

$$(X, Y, Z)^T \mapsto (fX/Z + c_x, fY/Z + c_y)^T \quad (2.28)$$

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zc_x \\ fY + Zc_y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} f & c_x & 0 \\ & f & c_y \\ & & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2.29)$$

กำหนดให้

$$K = \begin{pmatrix} f & c_x \\ & f & c_y \\ & & 1 \end{pmatrix} \quad (2.30)$$

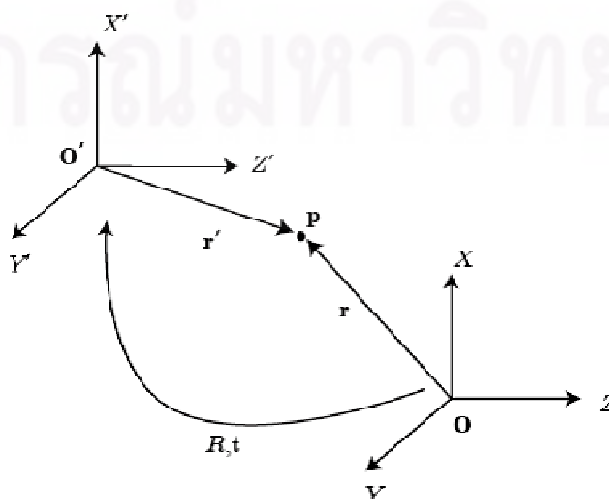
การปรับเทียบเมตริกซ์ในสมการที่ (2.29) เขียนได้ดังนี้

$$\lambda x = K[I \ 0]X \quad (2.31)$$

โดยที่ส่วนประกอบของการปรับเทียบเมตริกซ์เป็นพารามิเตอร์ภายใน (Intrinsic or Internal Parameters) และเมตริกซ์การฉายกลายเป็น $K[I \ 0]$

การหมุนและการเลื่อนขนานของกล้อง (Camera Rotation and Translation)

ในระบบพิกัดของระบบจริงและพิกัดของกล้องบางครั้งอาจจะไม่ได้เกิดขึ้นพร้อมกัน ดังนั้นจะต้องมีค่าพารามิเตอร์บางตัวมาทำให้ระบบของกล้องดีขึ้น ความแตกต่างของระบบพิกัดยูคลิด (Euclidean Coordinate) ใน $\mathbb{R}^3(0,0')$ มีความสัมพันธ์กันด้วยการเลื่อนขนานและการหมุน แสดงดังภาพที่ 2.10



รูปที่ 2.10 ความสัมพันธ์ของ 2 ระบบด้วยการแปลงและการหมุน

จุด P ที่แสดงในพิกัดของระบบแรกที่เวกเตอร์ตำแหน่ง r สามารถแสดงพิกัดในระบบที่สองได้ด้วยสมการดังนี้

$$r' = Rr + t \quad (2.32)$$

ซึ่ง R เป็น เมตริกซ์การหมุน และ t เป็นเวกเตอร์การเลื่อนขนาน

ในระบบพิกัดโฮโมจีเนียสจะสามารถเปลี่ยนรูปแบบการแปลงเป็น

$$r' = Hr \quad (2.33)$$

เมื่อมีการฉาย (Projection) ของระบบสามารถเขียนใหม่ได้ดังนี้

$$\lambda x = K [I \ 0] (RX + t) \quad (2.34)$$

เราสามารถรวมเอารูปแบบการฉายได้รูปแบบใหม่คือ

$$\lambda x = K [R \ t] X = PX \quad (2.35)$$

เมื่อ X คือ โครงสร้างพิกัดของระบบจริง เมตริกซ์ K เป็นเมตริกซ์ตัวเทียบวัด (Calibration Matrix) ที่ประกอบด้วยค่าพารามิเตอร์ภายใน (Intrinsic Parameters) ของกล้อง และ $[R \ t]$ เป็นค่าพารามิเตอร์ภายนอก (Extrinsic Parameters) ของกล้อง

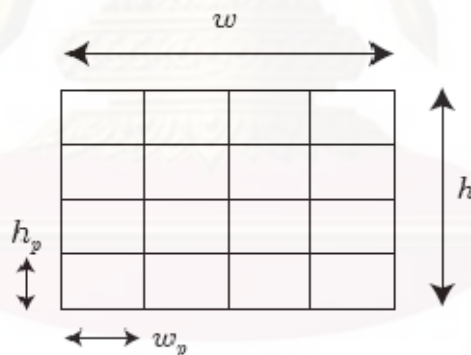
การจำแนกกล้อง (Discrete Cameras)

โดยทั่วไปภาพที่ปรากฏในกล้องมีการจำแนกของภาพในพิกเซล (Pixel) โดยสมมติให้บางสเกลในภาพทั้งทิศทางของกล้องจริง เช่น ถ้าความสูง h_p ของพิกเซลไม่เท่ากับความกว้าง w_p ดังนั้นจะต้องเพิ่มพารามิเตอร์เข้าไปด้วย

ขนาด $w \times h$ มีจำนวนของพิกเซลในแต่ละทิศทาง $n_x = w/w_p$ และ $n_y = h/h_p$ จำนวนพิกเซลต่อขนาดหน่วยคือ $m_x = \frac{1}{w_p}$ และ $m_y = \frac{1}{h_p}$ ในทิศทาง x และทิศทาง y ดังนั้นการแปลงจากพิกัดระบบจริงไปเป็นพิกัดพิกเซลได้จากการคูณด้านซ้ายของสมการที่ (2.30) ด้วยเมตริกซ์ $diag(m_x, m_y, 1)$

$$K = \begin{bmatrix} \alpha_x & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad (2.36)$$

ซึ่ง $\alpha_x = fm_x, \alpha_y = fm_y$ และ α_y/α_x ตามลำดับ



รูปที่ 2.11 ตัวอย่างลำดับของกล้อง CCD

พิกเซลที่ไม่เป็นสี่เหลี่ยมมุมฉาก (Non-rectangular Pixels)

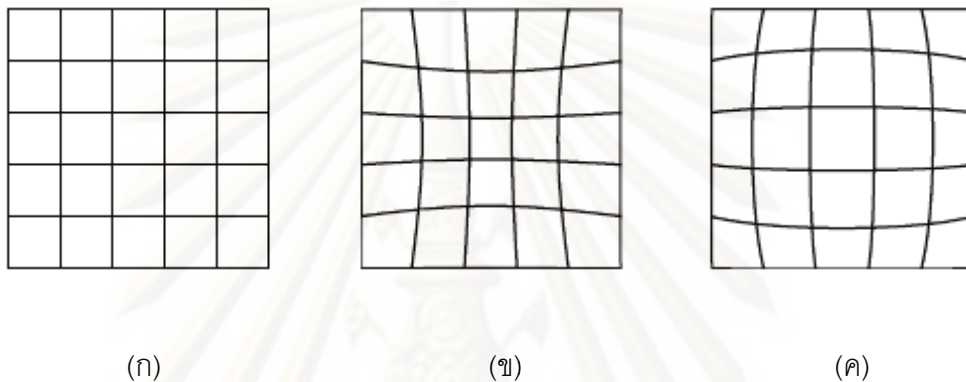
จะมีการเพิ่มค่าการเอียง (Skew) เข้าไปในค่าพารามิเตอร์ด้วยดังนั้นรูปแบบของเมตริกซ์การเทียบวัดกล้องเป็นดังสมการ

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad (2.37)$$

ซึ่ง s คือค่าการเอียง

การบิดเบือนของภาพ (Distortion)

รูปแบบของกล้องรูเข็ม ถ้าเส้นออปติคคอลลของกล้องมีการบิดเบือนจะทำให้เลนส์ไม่มีการปรับแนวของเลนส์ด้วยทำให้รูปแบบของกล้องที่ได้รับผิดเพี้ยนไป โดยทั่วไปการบิดเบือนมี 2 ชนิด ดังรูปที่ 2.12 ส่วนมากการบิดเบือนของภาพที่ได้เกิดจากเลนส์ของกล้อง



รูปที่ 2.12 การบิดเบือนแสดงบนตาราง (Grid) (ก) ตารางภาพ (Grid Image) ที่ไม่มีการบิดเบือน (ข) ตารางภาพที่มีการบิดเบือนแบบ Pincushion (ค) ตารางภาพที่มีการบิดเบือนแบบ Barrel

การแมทเชิงเส้น (Linear Mapping) ของเลนส์และการเขียนทางคณิตศาสตร์ของการบิดเบือนได้ดังนี้

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = L(r) \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.38)$$

ซึ่ง $(x_d, y_d)^T$ คือพิกัดของภาพที่วัดการบิดเบือน

$(x, y)^T$ คือพิกัดในภาพอุดมคติ

$r = \sqrt{x^2 + y^2}$ คือรัศมีระยะทางจากจุดศูนย์กลางการบิดเบือน

$L(r)$ คือฟังก์ชันการลดลงของการบิดเบือน

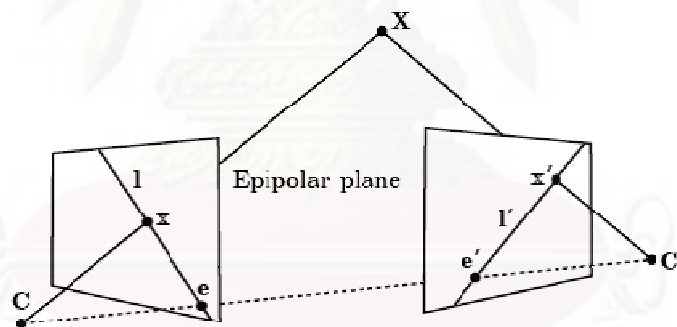
2.5 เรขาคณิตหลายมุมมอง (Multiple View Geometry)

โดยทั่วไปค่าต่างๆ ที่ได้จากภาพในหลายมุมมองจะมีค่าพารามิเตอร์แตกต่างกันไปหัวข้อนี้ อธิบายถึงความสัมพันธ์ระหว่างภาพ 2 ภาพหรือมากกว่า ซึ่งอธิบายในส่วนของภาพ มุมมอง (View) และกล้อง (รูปแบบของค่าพารามิเตอร์ภายในและภายนอกของกล้อง)

2.5.1 เรขาคณิตอีพิโพลาร์ (Epipolar Geometry)

เรขาคณิตอีพิโพลาร์อธิบายถึง การฉายแบบเรขาคณิตที่มีความเกี่ยวข้องกันระหว่างภาพ 2 มุมมอง ซึ่งขึ้นอยู่กับค่าพารามิเตอร์ภายในของกล้องและลักษณะท่าทาง (Pose) ของภาพ เท่านั้น และเรขาคณิตอีพิโพลาร์สามารถจะแสดงได้ด้วยเมตริกซ์ขนาด 3×3 เรียกว่าเมตริกซ์พื้นฐาน (Fundamental Matrix) F เราสามารถอธิบายคุณสมบัติของเรขาคณิตอีพิโพลาร์และเมตริกซ์พื้นฐานได้ โดยที่เรารู้ค่าพารามิเตอร์ภายในของกล้องที่ได้จากการเทียบวัดซึ่งเราจะหาค่าของเมตริกซ์จำเป็น (Essential Matrix) ได้เช่นเดียวกัน

ในรูปที่ 2.13 แสดงถึงพื้นฐานของเรขาคณิตอีพิโพลาร์ที่เป็นระนาบเรขาคณิตอีพิโพลาร์ที่มีเส้นตัดกันและจุดที่มีความสอดคล้องกัน (Corresponding)



รูปที่ 2.13 ระนาบของเรขาคณิตอีพิโพลาร์

เรขาคณิตอีพิโพลาร์เป็นการกำหนดค่าโดยระนาบของอีพิโพลาร์ บนเส้นของ $CC'X$ ซึ่ง จะตัดกันบนระนาบของภาพในเส้นอีพิโพลาร์ l และ l' ซึ่งประกอบด้วยอีพิโพล e และ e' และการฉายภาพมีความสอดคล้องกันด้วย x และ x'

เส้นฐาน (Baseline) เป็นเส้นร่วมกันของจุดศูนย์กลางกล้อง C และ C'

อีพิโพล (Epipole) e และ e' เป็นจุดตัดของเส้นฐานในระนาบภาพ

ระนาบอีพิโพลาร์ (Epipolar Plane) ระนาบที่ประกอบของเส้นฐาน

เส้นอีพิโพลาร์ (Epipolar Line) เป็นเส้นที่ตัดกันของระนาบอีพิโพลาร์อยู่บนระนาบภาพ ซึ่งจุด x ในภาพแรกมีความสอดคล้องกับเส้นอีพิโพลาร์ I' ด้วยจุด x' ในภาพที่สอง

เมตริกซ์พื้นฐาน (Fundamental Matrix)

เราสามารถอธิบายได้ด้วยพีชคณิตจากเรขาคณิตอีพิโพลาร์ ดังแสดงในรูปที่ 2.13 ที่จุด x ในภาพที่ 1 มีความสอดคล้องกับภาพที่ 2 ด้วยเส้นอีพิโพลาร์ I' และจุด x' ในภาพที่ 2 มีความสอดคล้องอยู่บนเส้นอีพิโพลาร์นี้เพราะเป็นการฉายของรังสี (Ray) จากจุด x ของจุด C ในกล้องที่ 1

$$x \mapsto I' \quad (2.39)$$

การแมป (maps) ของจุดบนภาพมีความสอดคล้องกับภาพอื่นด้วยเส้นอีพิโพลาร์ และการแมปแสดงได้ด้วย เมตริกซ์พื้นฐาน

$$I' = Fx \quad (2.40)$$

เงื่อนไขของความสอดคล้องเป็นดังสมการ

$$x'^T Fx = 0 \quad (2.41)$$

โดยที่ $x \leftrightarrow x'$ เป็นคู่ของความสอดคล้องบนจุดภาพและ x' เป็นจุดบนเส้นอีพิโพลาร์ $I' = Fx$ และผลการคูณ (Dot Product) ของสมการที่ (2.41) เท่ากับศูนย์ และคุณสมบัติของเมตริกซ์พื้นฐานมีความอิสระจากเมตริกซ์ของกล้อง คุณสมบัติของเมตริกซ์พื้นฐานเป็นดังนี้

1. ถ้า F เป็นเมตริกซ์พื้นฐานสำหรับคู่ของกล้อง (P, P') และ F^T เป็นเมตริกซ์พื้นฐานสำหรับคู่ของกล้อง (P', P) ที่จะเปลี่ยนไป
2. สำหรับจุด x ใดๆ ที่ไม่เป็นอีพิโพลภายในตัวเองแล้วเส้นอีพิโพลาร์ $I' = Fx$ ประกอบด้วยอีพิโพล e' ดังนั้น $e'^T (Fx) = (e'^T F)x = 0 \forall x$ นั่นคือ $e'^T F = 0$ และ e เป็นเวกเตอร์ศูนย์ทางซ้ายของ F เหมือนกับ e เป็นเวกเตอร์ศูนย์ทางขวาของ F
3. F มีดีกรีความอิสระเท่ากับ 7 เมตริกซ์โฮโมจีเนียส 3×3 และ $\det(F) = 0$

เมตริกซ์จำเป็น (Essential Matrix)

เมตริกซ์จำเป็น E เป็นกรณีพิเศษของเมตริกซ์พื้นฐาน F เมื่อกล้องมีการเทียบวัด(เมื่อรู้เมตริกซ์ของการเทียบวัด) มีกล้อง 2 กล้องโดยที่ กล้องที่ 1 เป็น $P = K[R \ t]$ และกล้องที่ 2 เป็น $P' = K'[R' \ t']$ ซึ่งเราจะรู้เมตริกซ์ของการเทียบวัดของ K และ K' เราสามารถนอร์มอลไลซ์พิกัดภาพได้ดังสมการ

$$x_n = K^{-1}x \quad (2.42)$$

$$x'_n = K'^{-1}x' \quad (2.43)$$

การฉายของภาพเป็นความสอดคล้องกันของ 2 กล้อง $P_n = [R \ t]$ และ $P'_n = [R' \ t']$ เรียกว่า เมตริกซ์การนอร์มอลไลซ์กล้อง เมื่อพิจารณาคู่ของเมตริกซ์การนอร์มอลไลซ์ $P_n = [I \ 0]$ และ $P'_n = [R \ t]$ ดังนั้นเมตริกซ์พื้นฐานเป็นดังสมการ

$$E = [t]_x R = R [R^T t]_x \quad (2.44)$$

ซึ่ง $[\cdot]_x$ เป็นเมตริกซ์ความเอียง (Skew) ของเวกเตอร์

$$[a]_x = \left[(a1, a2, a3)^T \right]_x = \begin{bmatrix} 0 & -a3 & a2 \\ a3 & 0 & -a1 \\ -a2 & a1 & 0 \end{bmatrix} \quad (2.45)$$

โดยทั่วไปข้อมูลของกล้องที่ 2 มีความสัมพันธ์กับกล้องที่ 1 และคู่จุดใดๆ ของกล้องอยู่ในรูปแบบของการแปลงของกล้องที่ 1 ด้วย $[R^T \ -R^T t]$ อยู่ในรูปแบบบัญญัติ (Canonical Form) และเรียกว่า กล้องแบบบัญญัติ (Canonical Camera) คุณสมบัติของเมตริกซ์พื้นฐานแสดงในสมการที่ (2.46) ซึ่งคล้ายคลึงกับสมการที่ (2.41)

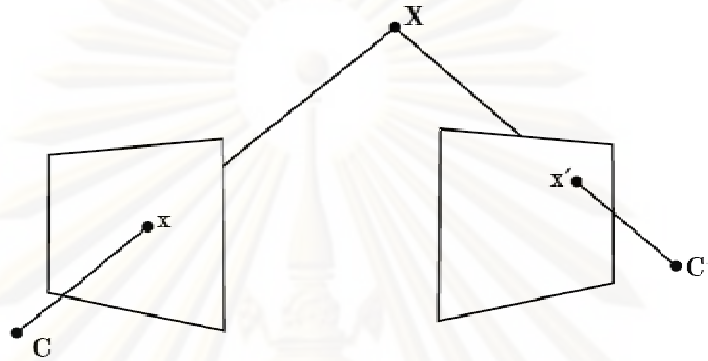
$$x'_n E x_n = 0 \quad (2.46)$$

เมตริกซ์พื้นฐานมีความสัมพันธ์โดยตรงกับเมตริกซ์จำเป็นดังสมการที่ (2.47)

$$E = K'^T F K \quad (2.47)$$

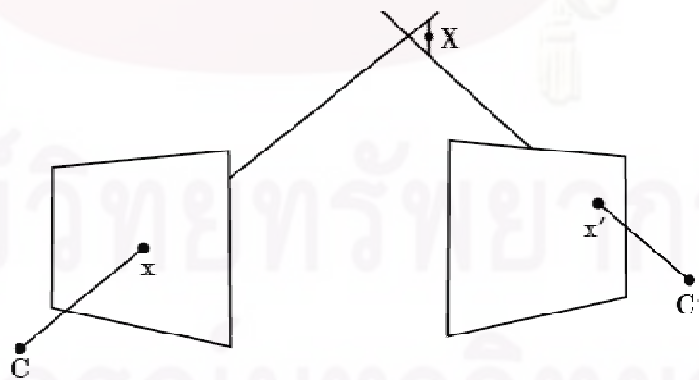
2.6 กาสร้างภาพ 3 มิติ

กระบวนการของการสร้างภาพ 3 มิติ หรือการสร้างเป็นรูปสามเหลี่ยม (Triangulation) ด้วยการสมมติฐานของการฉายกลับของรังสีของจุดศูนย์กลางกล้องตั้งนั้นจุดภาพและจุดตัดในระบบ 3 มิติเป็นดังแสดงในภาพ



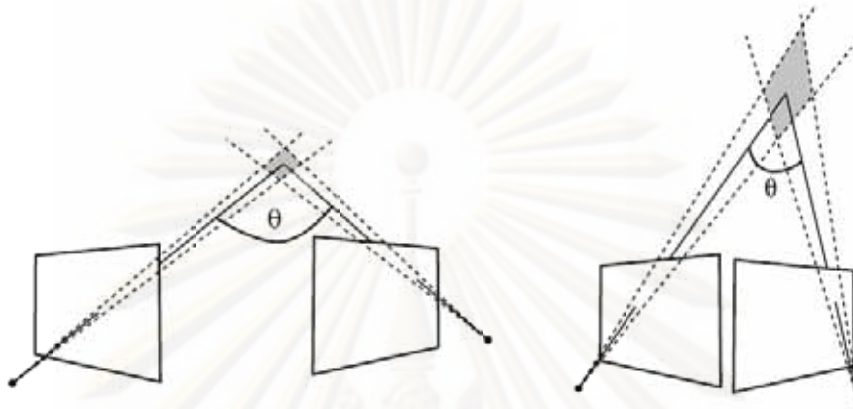
รูปที่ 2.14 การสร้างภาพในอุดมคติด้วยจุดตัดของรังสีการฉายกลับ

ถ้า $x_i = PX_i$ และ $x'_i = P'X_i$ ทุกคู่จุดของภาพ (x_i, x'_i) เป็นไปตามเงื่อนไขของอีพิโพลาร์ $x'_i F x_i = 0$ และ $x'_{n,i} T E x_{n,i} = 0$ ตามลำดับ สามารถคำนวณหาจุด X_i ในระบบ 3 มิติได้โดยตรง โดยทั่วไปในความเป็นจริงภาพที่ได้จะมีสัญญาณรบกวนและการฉายกลับของรังสีจะไม่อยู่บนจุดตัดดังแสดงในรูปที่ 2.1



รูปที่ 2.15 การสร้างภาพที่ไม่อยู่ในอุดมคติ

ความผิดพลาดของจุดในการสร้างภาพ 3 มิติโดยทั่วไปคือการฉายกลับของรังสีที่ไม่มี
ความแน่นอนการสร้างจุดในระบบ 3 มิติ ตัวบ่งชี้ความถูกต้องคือเส้นฐาน และถ้าเราต้องการ
ความถูกต้องในการสร้างจุด 3 มิติ เราจะต้องรู้มุมจุดตัด θ ดังแสดงในรูปที่ 2.16



รูปที่ 2.16 ความไม่แน่นอนของจุดในระบบการสร้างภาพ 3 มิติ

2.6.1 การสร้าง 3 มิติในภาพ 3 มุมมอง (3-View Methods)

ในส่วนนี้อธิบายถึงการสร้างภาพ 3 มิติในภาพ 3 มุมมองซึ่งก็คล้ายคลึงกันเมตริกซ์
พื้นฐาน ความสัมพันธ์กันของโครงสร้างทางคณิตศาสตร์และความสอดคล้องกันของข้อมูลใน 3
มุมมอง เรียกว่า ไตรโฟคัลเทนเซอร์ (Trifocal Tensor)

การคำนวณของไตรโฟคัลเทนเซอร์มีความคล้ายคลึงกับเมตริกซ์พื้นฐาน ถ้าเรารู้เมตริกซ์
กล้องเราสามารถประมาณค่าของจุดในระบบ 3 มิติได้ดังนี้

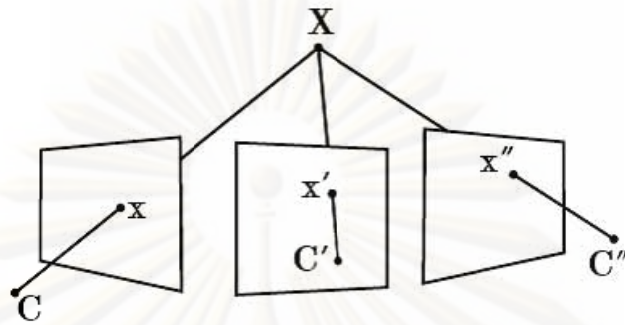
เมตริกซ์กล้อง 3 กล้อง $P=[I \ 0], P'=[A \ a_4]$ และ $P''=[B \ b_4]$ สามารถคำนวณไตรโฟ
คัลเทนเซอร์ได้โดยตรงดังนี้

$$T_i = a_i b_4^T - a_4 b_i^T \quad (2.48)$$

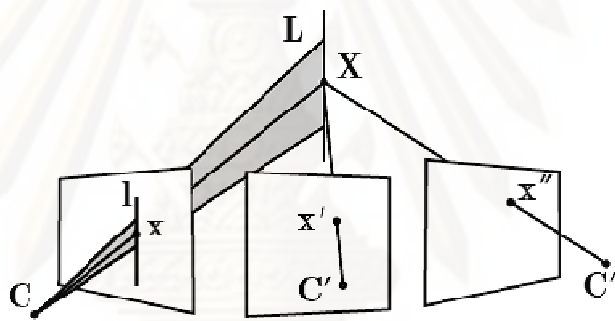
ซึ่ง a_i และ b_i จำนวนคอลัมน์ของเมตริกซ์กล้อง A และ B ที่ $i=1...3$

ความสัมพันธ์แบบต่างๆ ของไตรโฟคัลเทนเซอร์ แสดงดังรูปที่ 2.17 ซึ่งเป็นการแสดง
ความสัมพันธ์ในแบบต่างๆ ของภาพ 3 มุมมอง ซึ่งไม่เหมือนเมตริกซ์พื้นฐานที่ใช้ความสอดคล้อง

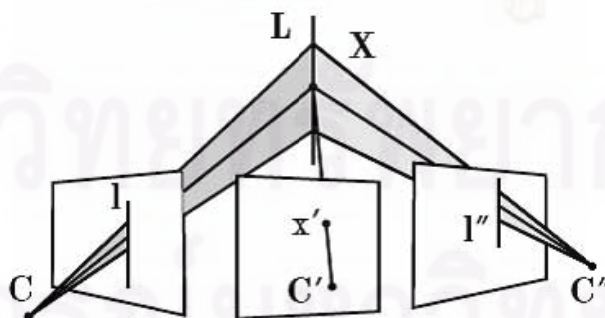
ของจุด ไตรโพล์แทนเซอร์ใช้ความสอดคล้องได้หลายแบบคือ แบบจุด-จุด-จุด แบบจุด-จุด-เส้น และแบบจุด-เส้น-เส้น



(ก) ความสัมพันธ์แบบจุด-จุด-จุด



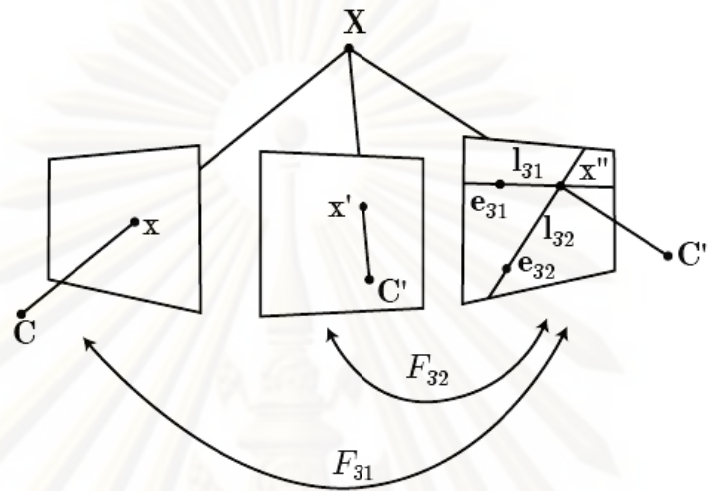
(ข) ความสัมพันธ์แบบจุด-จุด-เส้น



(ค) ความสัมพันธ์แบบจุด-เส้น-เส้น

รูปที่ 2.17 ความสัมพันธ์ของไตรโพล์

การแปลงจุดในไตรโพลีคัลเทนเซอร์ถ้ากล้อง (P, P', P'') และการฉายในภาพ 2 ภาพ (x, x') ความสัมพันธ์ของจุด x'' ในภาพที่ 3 สามารถหาได้โดยตรงดังแสดงในรูปที่ 2.18



รูปที่ 2.18 การแปลงจุดด้วย 2 เมตริกซ์พื้นฐาน

สมการของความสัมพันธ์แบบจุด-จุด-จุด แสดงได้ดังนี้

$$[x']_x \left(\sum_i x^i T_i \right) [x'']_x = 0_{3 \times 3} \quad (2.49)$$

ซึ่งคล้ายคลึงกับเงื่อนไขของอีพิโพลาร์ดังสมการที่ (2.41) ภาพ 2 มุมมองสามารถหาพินดาเมนทัลได้โดยตรงด้วยการคำนวณหาอีพิโพล (e, e') ให้ u_i และ v_i ทางด้านซ้ายและขวาของ T_i เป็นเวกเตอร์ศูนย์ (Null Vectors) ดังนั้นอีพิโพลเป็นเวกเตอร์ศูนย์

$$e'^T [u_1, u_2, u_3] = 0 \quad (2.50)$$

$$e''^T [v_1, v_2, v_3] = 0 \quad (2.51)$$

และเมตริกซ์พื้นฐานสามารถเขียนได้ดังนี้

$$F_{21} = [e']_x [T_1, T_2, T_3] e'' \quad (2.52)$$

$$F_{31} = [e'']_x [T_1^T, T_2^T, T_3^T] e' \quad (2.53)$$

เมตริกซ์พื้นฐานสามารถคำนวณหาความสัมพันธ์ของจุด x'' ใน 3 มุมมองได้จากการคำนวณหาความสัมพันธ์ของ $x \leftrightarrow x'$

$$x'' = (F_{31}x) \times (F_{32}x') \quad (2.54)$$

2.6.2 การสร้าง 3 มิติในภาพ N มุมมอง (N-View Methods)

วิธีการสร้างภาพ 3 มิติ ในภาพ N มุมมองยังได้ครอบคลุมถึงภาพวีดิทัศน์ (Video) ด้วยรูปแบบของเรขาคณิตอิมพีโพลาร์สำหรับ 2 มุมมอง (เมตริกซ์พื้นฐาน) และ 3 มุมมอง (ไตรโพลคัลเทนเซอร์) สามารถนำมาใช้กับ n มุมมองได้

สำหรับภาพที่เป็นแบบแข็งเกร็ง (Rigid) มีวิธีการสร้างภาพ 3 มิติที่น่าสนใจเรียกว่า บันเดิลแอดจัสต์เมนต์ (Bundle Adjustment) มีกระบวนการหาค่าที่เหมาะสมที่สุดในบริเวณ (Global Optimization) ปกติใช้วิธีการ Levenberg-Marquardt (LM) [Levenberg 1944, Marquardt 1963] บันเดิลแอดจัสต์เมนต์เป็นการหาค่าความผิดพลาดที่น้อยที่สุดของการฉาย

$$\min_{\hat{P}^i, \hat{X}_j} \sum_{i,j} d(\hat{P}^i \hat{X}_j, x_j)^2 \quad (2.55)$$

$d(\cdot, \cdot)$ เป็นระยะทางภาพเรขาคณิตระหว่างเวกเตอร์ 2 เวกเตอร์

ส่วนของการบันเดิลแอดจัสต์เมนต์ มาจากการปรับตัวของรังสีจากแต่ละจุดศูนย์กลางกล้องและจุดจุด 3 มิติ ซึ่งเมื่อต้องการสร้างจุดใน 3 มิติ จึงจำเป็นต้องมีการหาค่าที่เหมาะสมที่สุด

บทนี้ได้อธิบายถึงพื้นฐานของสมการทางคณิตศาสตร์ที่จำเป็นต่อการสร้างภาพ 3 มิติ โดยได้อธิบายถึงหลักการทั่วไปทางด้านคอมพิวเตอร์กราฟฟิก, เรขาคณิตในภาพ 1 มุมมอง อธิบายแบบจำลองของกล้องรูเข็มเกี่ยวกับค่าพารามิเตอร์ต่างๆ ของกล้องซึ่งแสดงถึงเรขาคณิตอิมพีโพลาร์ด้วยเมตริกซ์พื้นฐาน ซึ่งมีความสำคัญต่อความสัมพันธ์ของภาพ 2 มุมมอง การสร้างจุด 3 มิติจากจุด 2 มิติด้วยการฉายแบบรูปสามเหลี่ยมและสุดท้ายอธิบายเรขาคณิตใน 3 มุมมองและ N มุมมอง

บทที่ 3

วิธีการที่นำเสนอ

บทนี้กล่าวถึงวิธีการสร้างภาพใบหน้าคน 3 มิติจากหลายมุมมองโดยแบ่งวิธีการสร้างภาพใบหน้าคน 3 มิติออกเป็น 2 ส่วนคือ การแบ่งส่วนภาพ (Image Segmentation) ใบหน้าคน 2 มิติ เราใช้การแบ่งส่วนภาพใบหน้าคนด้วยวิธีการจำลองรูปร่างแบบแอคทีฟ (Active Shape Models) หรือ ASM และส่วนของการสร้างภาพ 3 มิติ เราสร้างภาพใบหน้าคน 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรง (Direct Linear Transformation) หรือ DLT

3.1 การแบ่งส่วนภาพ

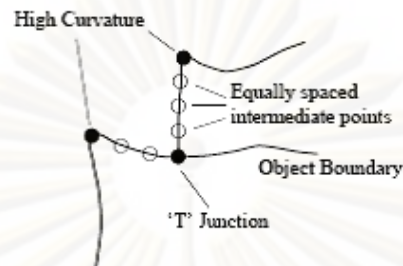
การแบ่งส่วนภาพเป็นขั้นตอนเบื้องต้นที่สำคัญสำหรับงานในการประมวลผลภาพ (Image Processing) ที่ต้องการตัดส่วนใดส่วนหนึ่งของภาพออกมา เพื่อนำไปประมวลผลในขั้นตอนถัดไป ขั้นตอนวิธีการที่ใช้ในการแบ่งส่วนภาพนั้นได้มีหลากหลายวิธีด้วยกัน เช่น วิธี Thresholding, วิธี Region Growing, วิธี K-Means, วิธี Watershed, วิธี Graph cuts, วิธี แอ็กทีฟคอนทัวร์ (Active Contour) และวิธีการจำลองรูปร่างแบบแอ็กทีฟ (Active Shape Models) เป็นต้น โดยในวิทยานิพนธ์นี้ ได้สนใจและเลือกที่จะพัฒนาวิธีการจำลองรูปร่างแบบแอ็กทีฟหรือ ASM [16]

วิธีการจำลองรูปร่างแบบแอ็กทีฟเป็นการใช้ข้อมูลทางสถิติเข้ามาช่วยในการจำลองรูปร่าง (Statistical Shape Models) เราจะใช้รูปร่าง (Shape) ของวัตถุในภาพอยู่ในรูปชุดข้อมูลจุดจำนวน n จุด โดยทั่วไปชุดข้อมูลเหล่านี้จะอยู่ในระบบ 2 มิติหรือ 3 มิติก็ได้ โดยปกติทั่วไปองค์ประกอบของรูปร่างในภาพจะไม่แปรผันกับการแปลง (Transformation) รูปร่างของวัตถุในภาพจะไม่เปลี่ยนแปลงเมื่อมีการเลื่อนขนาน (Translation) การหมุนวัตถุ (Rotation) และการสเกล (Scaling) การแปลงภาพในภาพ 3 มิติมีความคล้ายคลึงกับการแปลงภาพ 2 มิติ และการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอ็กทีฟเป็นการวิเคราะห์ (Analysis) รูปร่างของวัตถุภายในภาพ และเป็นการสังเคราะห์ (Synthesis) รูปร่างใหม่ของวัตถุที่มีความคล้ายคลึงกับรูปร่างในกลุ่มตัวอย่าง (Training) จากแบบจำลองและข้อมูลของกลุ่มตัวอย่างได้มาจากการสร้างขึ้นจากภาพของกลุ่มตัวอย่างซึ่งสร้างขึ้นเป็นระบบจุดสำคัญ (Landmarks)

3.1.1 การกำหนดจุดสำคัญของวัตถุบนภาพ

การกำหนดจุดสำคัญของวัตถุจะพิจารณาถึงบริเวณต่างๆ ของวัตถุภายในภาพ วิธีที่ง่าย ๆ สำหรับการสร้างกลุ่มตัวอย่างคือการสร้างด้วยคน การค้นหาบริเวณของวัตถุที่มีความสอดคล้อง

กันของจุดในแต่ละภาพ ในภาพ 2 มิติบริเวณที่สามารถจะกำหนดเป็นจุดสำคัญได้เช่น บริเวณขอบของวัตถุ เส้นโค้งและมุมต่างๆ และลักษณะที่สำคัญบนวัตถุ จากรูปที่ 3.1 แสดงถึงบริเวณต่างๆที่กำหนดจุดสำคัญของวัตถุภายในภาพ



รูปที่ 3.1 การกำหนดจุดสำคัญของวัตถุ Cootes, T. F. et al. [16]

รูปร่างของวัตถุมีขนาดเท่ากับ n จุดใน d มิติเราจะอธิบายรูปร่างของวัตถุได้ด้วยเวกเตอร์องค์ประกอบ nd ใน 2 มิติเราอธิบายด้วยจุด $\{(x_i, y_i)\}$ จำนวน n จุดมีเวกเตอร์องค์ประกอบ $2n$ โดยที่ x เขียนได้ดังสมการ

$$x = (x_1, \dots, x_n, y_1, \dots, y_n)^T \quad (3.1)$$

3.1.2 แบบจำลองการแจกแจงจุด

แบบจำลองการแจกแจงจุด (Point Distribution Models) หรือ PDM เป็นเทคนิคที่ใช้ในการหารูปร่าง (Shape) ของใบหน้าคน ประกอบด้วย 2 ส่วนคือ ส่วนของกระบวนการปรับแนว (Alignment) ของกลุ่มตัวอย่าง และส่วนของการหาความแปรปรวนแบบจำลองรูปร่าง การคำนวณแบบจำลองการแจกแจงจุดเป็นการประมาณค่าของกลุ่มข้อมูลตัวอย่าง (Training Data Set) โดยการหาค่าเฉลี่ยตำแหน่งของจุดสำคัญ (Landmarks) ที่ใช้แทนรูปร่างของใบหน้าคน ซึ่งจากการกำหนดจุดสำคัญของภาพใบหน้าคนในแต่ละภาพของกลุ่มตัวอย่างเราจะทำการปรับแนวภาพเนื่องมาจากจุดที่ได้จากภาพใบหน้าคนมีขนาด ตำแหน่ง และมุมของภาพไม่เท่ากัน



รูปที่ 3.2 ตัวอย่างของใบหน้าที่มีลักษณะต่างกัน : Cootes, T. F. et al. [16]

3.1.3 กระบวนการปรับแนวของกลุ่มตัวอย่าง

กระบวนการปรับแนวของข้อมูลตัวอย่าง (Aligning the Training Data) เป็นสิ่งที่จำเป็นในกระบวนการปรับแนวของรูปร่างตัวอย่าง (Training Shape) ในการประมาณค่าโดยการกำหนดค่าของแต่ละตัวอย่างของการเลื่อนขนาน (Translation) การสเกล (Scaling) และการหมุนวัตถุ (Rotation) เป็นตัวกำหนดในการหาค่าจุดที่เหมือนกันและแตกต่างกันของรูปร่าง กระบวนการปรับแนวใช้เพื่อหารูปร่างค่าเฉลี่ย (Mean Shape) ของกลุ่มข้อมูลตัวอย่าง วิธีที่ได้รับความนิยมมากในกระบวนการปรับแนวของรูปร่างในระบบพิกัด (Coordinate) คือการวิเคราะห์รูปร่างของกลุ่มข้อมูลตัวอย่างขั้นตอนวิธีกระบวนการปรับแนวแบบนี้เรียกว่าวิธีการโปรครัสเตอร์ (Procruster Method) [17] กระบวนการปรับแนวของรูปร่างเป็นการรวมตัวกันของจุดกลุ่มข้อมูลตัวอย่างเพื่อหาค่าของระยะทางในแต่ละจุดของรูปร่าง

ขั้นตอนวิธีกระบวนการปรับแนวได้โดยการปรับแนวของคู่รูปร่างโดยให้

$$p_i = [x_i \ y_i]^T \quad (3.2)$$

P_i เป็นรูปร่างของกลุ่มตัวอย่าง

ซึ่งกลุ่มจุดรูปร่างเหล่านี้ อาจมีการเปลี่ยนแปลงแบบหมุน θ การปรับมาตรา s และการเลื่อนขนาน (t_x, t_y) ได้ดังสมการ

$$p'_i = T(p_i) = s \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (3.3)$$

โดยที่ $p'_i = [x'_i \ y'_i]$ คือตำแหน่งที่เปลี่ยนไปของ p_i

ดังนั้นผลรวมถ่วงน้ำหนักของค่าความคลาดเคลื่อนของจุด p_i คือ

$$E = \sum_{i=1}^N w_i \left\| s \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} - \begin{bmatrix} x'_i \\ y'_i \end{bmatrix} \right\|^2 \quad (3.4)$$

ซึ่ง w_i เป็นตัวถ่วงน้ำหนักของจุดสำคัญ p_i

$\| \cdot \|^2$ คือ นอร์มกำลังสอง (L_2 norms)

เราสามารถหาค่าพารามิเตอร์ θ, s, t_x, t_y ที่ทำให้สมการ (3.4) มีค่าน้อยที่สุดได้โดยการหาค่าของพารามิเตอร์แต่ละตัวที่ทำให้อนุพันธ์ของ E เทียบกับพารามิเตอร์นั้นมีค่าเป็นศูนย์ ดังนี้

$$\frac{\partial E}{\partial t_x} = 0, \quad \frac{\partial E}{\partial t_y} = 0, \quad \frac{\partial E}{\partial s} = 0$$

พารามิเตอร์ s, t_x, t_y สามารถแก้สมการได้ด้วยสมการเชิงเส้น (Linear Equation) ตามลำดับดังนี้

$$s \sum_{i=1}^N w_i q(y_i, -x_i, \theta) - N t_x = - \sum_{i=1}^N w_i x'_i \quad (3.5)$$

$$s \sum_{i=1}^N w_i q(-x_i, -y_i, \theta) - N t_y = - \sum_{i=1}^N w_i y'_i \quad (3.6)$$

$$\begin{aligned} & s \sum_{i=1}^N w_i^2 (q^2(y_i, -x_i, \theta) + q^2(x_i, y_i, \theta)) - t_x \sum_{i=1}^N w_i q(y_i, -x_i, \theta) - t_y \sum_{i=1}^N w_i q(-x_i, y_i, \theta) \\ & = - \sum_{i=1}^N w_i x'_i q(y_i, -x_i, \theta) + \sum_{i=1}^N w_i y'_i q(x_i, -y_i, \theta) \end{aligned} \quad (3.7)$$

โดยที่ $q(a, b, \theta) = a \sin \theta + b \cos \theta$

N คือ จำนวนจุดของข้อมูลในรูปร่างของใบหน้า

สำหรับพารามิเตอร์ θ นั้นเราไม่สามารถแก้ด้วยสมการเชิงเส้นได้ แต่เราสามารถหาค่า θ ได้จากค่า s, t_x, t_y ของสมการ (3.5) ถึง (3.7) ดังสมการ (3.8)

$$\theta = \arg \min_{\theta} E(s, t_x, t_y) \quad (3.8)$$

การหาค่าเฉลี่ยของรูปร่างของวัตถุโดยกระบวนการปรับแนวนั้นเราจะได้ค่าเฉลี่ยของรูปร่างวัตถุเป็น \bar{P} แต่ละรูปร่างให้คู่อันดับ N คู่ ได้ค่าเฉลี่ยรูปร่างดังนี้

$$p_j = [x_1^j, y_1^j, x_2^j, y_2^j, \dots, x_N^j, y_N^j]^T \quad (3.9)$$

$$F = \{p_1, p_2, p_3, \dots, p_M\} \quad (3.10)$$

$$\bar{P} = \frac{1}{M} \sum_{j=1}^M p_j \quad (3.11)$$

โดยที่ M คือจำนวนชุดของข้อมูลตัวอย่าง

\bar{P} ค่าเฉลี่ยรูปร่างของกลุ่มตัวอย่าง

3.1.4 ความแปรปรวนรูปร่างแบบจำลอง (Modeling Shape Variation)

สมมติให้กลุ่มตัวอย่าง M มีชุดข้อมูลจำนวน p_i จุดโดยทั่วไปกระบวนการปรับแนวในระบบพิกัดเวกเตอร์ที่ใช้จะมีการแจกแจงใน nd มิติ เราสามารถที่จะแจกแจงแบบจำลองนี้ได้ และเราสามารถที่จะสร้างกลุ่มตัวอย่างใหม่ที่มีความคล้ายคลึงกับกลุ่มตัวอย่างเดิมได้และทดสอบรูปร่างใหม่ที่จะตัดสินใจในความเป็นไปได้ของรูปร่างใหม่ด้วย

เราหาพารามิเตอร์ของแบบจำลองได้จาก $p = V(b)$, b เป็นพารามิเตอร์ของแบบจำลอง ดังนั้นแบบจำลองสามารถใช้ในการสร้างเวกเตอร์ p ใหม่ได้ ถ้าเราสามารถแจกแจงพารามิเตอร์แบบจำลอง $p(b)$ เราจะกำหนดค่า p เพื่อสร้างแบบจำลองให้มีความคล้ายคลึงกับรูปร่างของกลุ่มตัวอย่าง

เราแก้ปัญหาในการประมาณค่าการสร้างแบบจำลองโดยภายในข้อจำกัดของข้อมูลจาก nd มิติโดยลดจำนวนมิติลงด้วยประมาณค่าของวิธีการวิเคราะห์องค์ประกอบแกนหลัก (Principal Component Analysis) หรือ PCA เป็นเทคนิคในการผสมลักษณะเด่นของเวกเตอร์ข้อมูลภาพ เพื่อสร้างเวกเตอร์ใหม่ที่อยู่ในปริภูมิที่มีมิติน้อยกว่าเวกเตอร์เดิม โดยการผสมที่เราใช้นั้นจะเป็นการผสมเชิงเส้น (Linear Combination) นั่นคือ การเอาเวกเตอร์ลักษณะเด่นมาคูณกับค่าคงที่แล้วนำมารวมกัน ซึ่งการวิเคราะห์องค์ประกอบแกนหลักมีข้อดี คือ เป็นการลดจำนวนพารามิเตอร์ของ

ระบบ เพราะจำนวนพหามิตอร์เหล่านี้ ขึ้นอยู่กับจำนวนมิติของเวกเตอร์ข้อมูลภาพ ซึ่งประมาณค่าดังนี้

1. คำนวณหาค่าเฉลี่ยของข้อมูลด้วย

$$\bar{p} = \frac{1}{M} \sum_{i=1}^m p_i \quad (3.12)$$

2. คำนวณค่าความแปรปรวนร่วม

$$S = \frac{1}{M-1} \sum_{i=1}^M (p_i - \bar{p})(p_i - \bar{p})^T \quad (3.13)$$

3. คำนวณเวกเตอร์ลักษณะเฉพาะ (Eigenvectors) ϕ_i และความสอดคล้องของค่าลักษณะเฉพาะ λ_i ของ S (ค่าของ $\lambda_i \geq \lambda_{i+1}$) การคำนวณเวกเตอร์ลักษณะเฉพาะดูในภาคผนวก ก

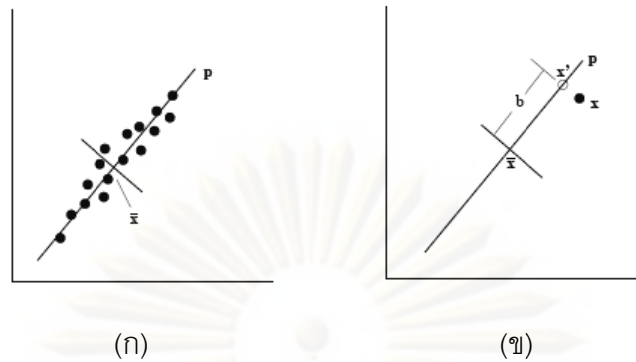
ถ้า ϕ มีเวกเตอร์ลักษณะเฉพาะที่สัมพันธ์กับค่าลักษณะเฉพาะมีขนาดใหญ่ เราสามารถประมาณค่าของกลุ่มตัวอย่างได้ด้วย

$$x \approx \bar{x} + \phi b \quad (3.14)$$

ที่ $\phi = (\phi_1 | \phi_2 | \dots | \phi_t)$ และ b เป็นเวกเตอร์ขนาด t มิติโดย

$$b = \phi^T (x - \bar{x}) \quad (3.15)$$

เวกเตอร์ b ถูกกำหนดด้วยพหามิตอร์ของการเปลี่ยนแบบจำลองโดยแปรเปลี่ยนขององค์ประกอบของ b ที่เราสามารถแปรเปลี่ยนรูปร่าง x ด้วยสมการที่ (3.14)



รูปที่ 3.3 PAC ในเวกเตอร์ 2 มิติ (ก) p เป็นแกนหลัก

(ข) ประมาณค่าจุด x ไปจุด x' ในบริเวณใกล้เคียงบนเส้นแกนหลัก Cootes, T. F. et al. [16]

รูปร่างในกลุ่มตัวอย่างสามารถประมาณค่า โดยใช้ค่าเฉลี่ยรูปร่าง ที่ได้จากรูปร่างกลุ่มตัวอย่าง ได้ดังสมการ

$$p_j = \bar{p} + \mathbf{V}\mathbf{b} \quad (3.16)$$

$\mathbf{V} = [v_1 \ v_2 \ \dots \ v_r]$ เป็นเมตริกซ์ของเวกเตอร์ลักษณะเฉพาะ (Eigenvector)

$\mathbf{b} = [b_1 \ b_2 \ \dots \ b_r]^T$ เป็นค่าถ่วงน้ำหนักของแต่ละเวกเตอร์ลักษณะเฉพาะ

ในการแยกเวกเตอร์ลักษณะเฉพาะ (Eigen Decomposition) นอกจากจะช่วยในการลดขนาดของข้อมูลภาพที่ต้องการคำนวณลงได้แล้ว เวกเตอร์การฉายที่หาได้จะมีคุณสมบัติในการดึงข้อมูลของภาพส่วนที่มีค่าความแปรปรวนร่วมของกลุ่มข้อมูลภาพสูงสุดออกมานั้น ซึ่งถือเป็นส่วนข้อมูลที่มีความสำคัญของภาพที่เหมาะสมแก่การนำมาใช้วิเคราะห์ โดยส่วนของข้อมูลภาพที่มีความสำคัญน้อยจะถูกจัดว่าเป็นสัญญาณที่ไม่ต้องการในการประมวลผลจึงถูกตัดทิ้งไป เมื่อทำการแปลงตามสมการจะได้เวกเตอร์คุณลักษณะซึ่งมีเฉพาะข้อมูลส่วนที่มีความสำคัญของภาพแต่ละภาพไปทำการวิเคราะห์ความแตกต่างของภาพต่อไป

3.1.5 การฟิตแบบจำลอง

การฟิตแบบจำลอง (Fitting Models) สามารถอธิบายได้ด้วยค่าพารามิเตอร์รูปร่าง b ร่วมกับการแปลงแบบจำลองในระบบพิกัดของภาพซึ่งมีความคล้ายคลึงกับการแปลงโดยการเลื่อนขนาน (x_t, y_t) การหมุนวัตถุ θ และการสเกล s ของแบบจำลองในภาพ

การเลื่อนขนานของจุดของแบบจำลองในภาพ x คือ

$$x = T_{x_t, y_t, s, \theta}(\bar{x} + \theta b) \quad (3.17)$$

ฟังก์ชันของ $T_{x_t, y_t, s, \theta}$ หมุนวัตถุด้วย θ สเกลด้วย s และย้ายตำแหน่งด้วย (x_t, y_t) เราสามารถหาจุด (xy) ได้ดังสมการที่ (3.18)

$$T_{x_t, y_t, s, \theta} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \begin{pmatrix} s \cos \theta & s \sin \theta \\ -s \sin \theta & s \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.18)$$

เราหาค่าพารามิเตอร์รูปร่างและพารามิเตอร์ระบุตำแหน่ง (Pose Parameters) ที่เข้าคู่กัน (Match) ของแบบจำลองที่แทนที่ x ด้วยจุดภาพ Y ค่าที่น้อยที่สุดของผลรวมระยะทางกำลังสองระหว่างความสอดคล้องแบบจำลองและจุดภาพจะมีค่าเท่ากัน เป็นดังสมการ

$$|Y - T_{x_t, y_t, s, \theta}(\bar{x} + \theta b)|^2 \quad (3.19)$$

ขั้นตอนการประมาณค่าของแบบจำลอง

1. กำหนดค่าเริ่มต้นของพารามิเตอร์รูปร่าง b เท่ากับศูนย์
2. สร้างแบบจำลอง $x = \bar{x} + \theta b$
3. หาพารามิเตอร์ระบุตำแหน่ง (X_t, Y_t, s, θ) เพื่อเมท x เข้าไปใน Y อธิบายในภาคผนวก ข
4. ผันกลับ (Invert) พารามิเตอร์ระบุตำแหน่งและให้ Y ฉายเข้าไปในระบบพิกัดในแบบจำลอง

$$y = T_{x_t, y_t, s, \theta}^{-1}(Y) \quad (3.20)$$

5. ค่า y เข้าไปในระนาบของ \bar{x}
6. ปรับค่าพารามิเตอร์แบบจำลองให้เข้าคู่กับ y

$$b = \phi^T (y - \bar{x}) \quad (3.21)$$

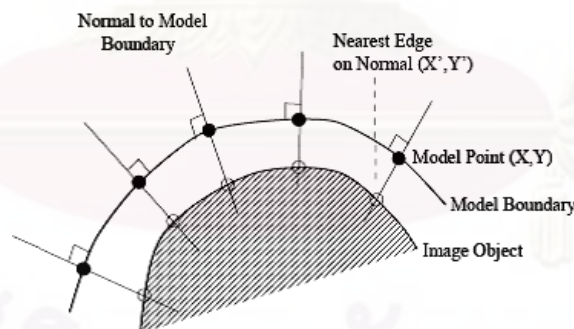
7. ถ้ายังไม่คู่เข้าหาค่าที่เหมาะสมก็กลับไปทำตามข้อที่ 2 อีกครั้ง

3.1.6 ฟังก์ชันการฟิต

เราเลือกฟังก์ชันการฟิตของแบบจำลองด้วยการสมมติให้รูปร่างแบบจำลองที่เราใช้เป็นขอบของวัตถุและเราหาขอบจริงของวัตถุที่มีความแข็งแรงด้วยการวัดระยะทางระหว่างจุดแบบจำลองและจุดภาพที่มีความแข็งแรงในบริเวณใกล้เคียงในภาพ ดังแสดงในรูปที่ 3.4

ถ้าตำแหน่งจุดแบบจำลองอยู่ในเวกเตอร์ X และจุดขอบในบริเวณใกล้เคียงในแต่ละจุดแบบจำลองคือ X' ค่าความผิดพลาดหาได้ดังสมการที่ (3.22)

$$F(b, X_i, Y_i, s, \theta) = |X' - X|^2 \quad (3.22)$$



รูปที่ 3.4 ขอบของวัตถุและจุดแบบจำลอง: Cootes, T. F. et al. [16]

เราสามารถที่จะสร้างจุด X ของแบบจำลองเข้าไปในภาพด้วยการกำหนดตำแหน่ง การหมุนและการสเกลโดยสมการที่ (3.17) การประมาณค่าของแบบจำลองในการฟิตข้อมูลได้ดังต่อไปนี้

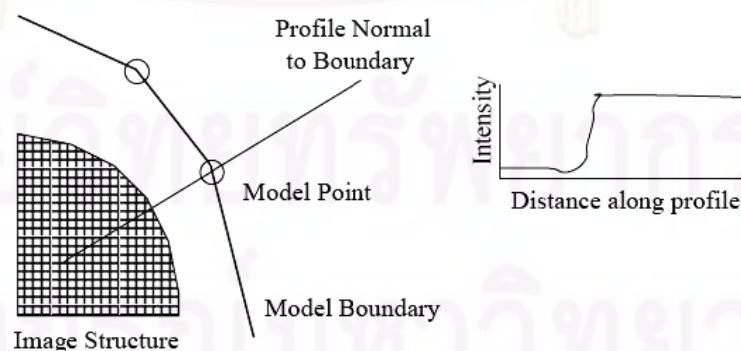
1. กำหนดจุดเริ่มต้นในภาพรอบแต่ละจุด X_i หาตำแหน่งที่ใกล้ที่สุดสำหรับการเข้าสู่คู่ของจุด X'_i
2. ปรับค่าพารามิเตอร์ (X_i, Y_i, s, θ, b) ในการพีตเพื่อหาตำแหน่งจุด X ใหม่
3. ทำซ้ำจนกว่าจะได้รูปร่างที่ดีที่สุด

มีการกำหนดเส้นค้นหา (Profiles) ที่ขอบของภาพในแต่ละจุดของแบบจำลองดังแสดงในรูปที่ 3.4 แสดงเส้นขอบของวัตถุและจุดแบบจำลองโดยแต่ละจุดแบบจำลองมีเส้นค้นหาขอบของวัตถุโดยที่เส้นค้นหาจะตั้งฉากกับจุดแบบจำลอง ถ้าเราให้ขอบแบบจำลองมีความสอดคล้องกับขอบของวัตถุเราสามารถตำแหน่งขอบที่ชัดเจนได้ด้วยเส้นค้นหา ซึ่งจะเป็นตำแหน่งใหม่ที่จุดของแบบจำลองจะย้ายไป ซึ่งเป็นตำแหน่งที่ขนาดเกรเดียนท์มากที่สุดบนเส้นค้นหา และมีการปรับค่าการหมุน การปรับมาตราและการเลื่อนขนาด รูปร่างของรูปแบบเพื่อให้เข้ากับวัตถุของภาพได้ดังสมการ

$$p = s \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} (\mathbf{V}\mathbf{b} + \bar{p}) + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (3.23)$$

เมื่อเราปรับข้อมูลให้เข้ากับวัตถุในภาพ โดยการปรับค่าการหมุน การปรับมาตรา การเลื่อนขนาดและรูปร่างของพารามิเตอร์ \mathbf{b}

$$p + dp \approx \bar{p}_i + \mathbf{V}(\mathbf{b} + d\mathbf{b}) \quad (3.24)$$



รูปที่ 3.5 ตัวอย่างเส้นค้นหาจุดแต่ละจุดแบบจำลอง: Cootes, T. F. et al. [16]

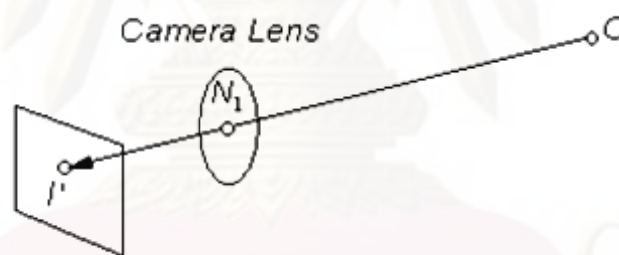
วิธีการแบ่งส่วนภาพใบหน้าคนด้วยวิธีการจำลองรูปร่างแบบเอกทิพสามารถแบ่งส่วนภาพของใบหน้าคนได้ดีเนื่องจากรูปร่างที่ได้หลังจากการแบ่งส่วนภาพใบหน้าคนแล้วยังคงรูปร่างของใบหน้าไว้ได้ด้วย

การสร้างภาพใบหน้าคน 3 มิติเราสร้างด้วยวิธีการแปลงเชิงเส้นโดยตรงซึ่งก็มีส่วนของการเทียบวัตถุกล้องเพื่อหาค่าพารามิเตอร์กล้องและการสร้างภาพ 3 มิติ

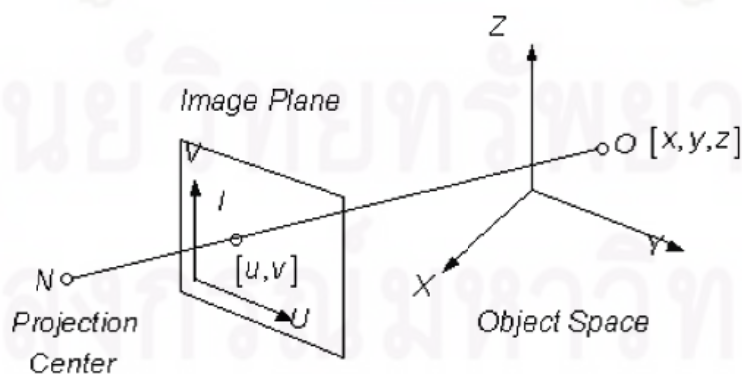
3.2 การแปลงเชิงเส้นโดยตรง

การแปลงเชิงเส้นโดยตรง (Direct linear Transformation) หรือ DLT เป็นวิธีการเทียบวัตถุกล้องภายในตัวเอง วิธีการ DLT เป็นการนำเสนอด้วยการฉายเมตริกซ์ผกผันเทียม (Pseudo Inverse) ซึ่งอธิบายด้วยความสัมพันธ์ระหว่างพิกัด 2 มิติและพิกัด 3 มิติ

การถ่ายภาพด้วยกล้องคือการเมท (Mapping) วัตถุในระบบจริงเข้าไปในระบบเสมือนในภาพที่ 3.6 แสดงการเมทของจุดบนวัตถุเข้าไปในจุดภาพ

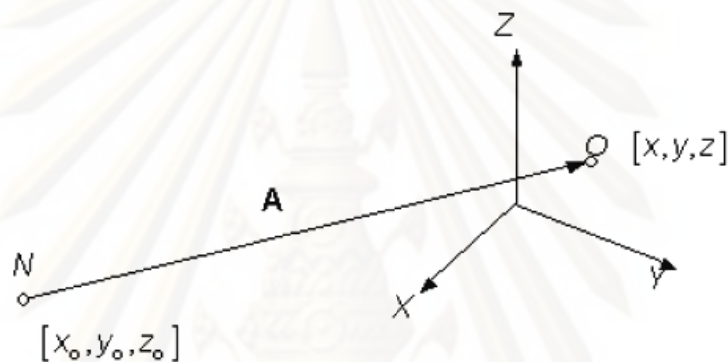


รูปที่ 3.6 การเมทของจุดบนวัตถุเข้าไปในจุดภาพ: Richard, H. and Andrew, Z.[26]



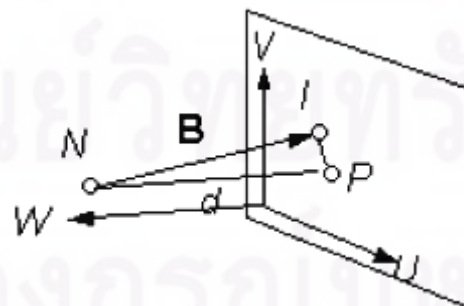
รูปที่ 3.7 ความสัมพันธ์การฉายภาพระบบจริงกับระบบภาพ: Richard, H. and Andrew, Z.[26]

รูปที่ 3.7 อธิบายถึงความสัมพันธ์ของการฉายภาพ โดยที่วัตถุ O เป็นวัตถุที่มีการฉายเข้าไปในภาพและระนาบที่ฉายเรียกว่า ระนาบภาพ (Image Plane) ที่จุด N เป็นศูนย์กลางการฉาย (Projection Center) ในระบบของจุดบนวัตถุและจุดบนภาพเรากำหนดโดย ปริภูมิวัตถุ (Object-Space Reference) ในพิกัด XYZ (XYZ System) และอ้างอิงระนาบภาพ (Image-Plane Reference) ในพิกัด uv (uv System) วัตถุ O มีพิกัดในปริภูมิวัตถุที่จุด $[x, y, z]$ และฉายจุดของวัตถุที่จุด $[x, y, z]$ ลงบนระนาบอ้างอิงภาพได้จุด $[u, v]$ เป็นพิกัดของจุดภาพ I จุด I, N และ O เรียกว่า เส้นตรงร่วม (Collinear) และเราเรียกว่า เงื่อนไขเส้นตรงร่วม (Collinearity Condition) เป็นพื้นฐานของวิธีของ DLT



รูปที่ 3.8 ความสัมพันธ์ตำแหน่งการฉายและปริภูมิวัตถุ: Richard, H. and Andrew, Z.[26]

รูปที่ 3.8 สมมติให้ตำแหน่งของจุดศูนย์กลางการฉาย N มีพิกัด $[x_0, y_0, z_0]$ เวกเตอร์ A เป็นเส้นตรงจากจุด N ถึง O เขียนได้เป็น $[x - x_0, y - y_0, z - z_0]$



$$I = [u, v, 0]$$

$$P = [u, v, 0]$$

$$N = [u_0, v_0, d]$$

รูปที่ 3.9 ระนาบภาพ 3 มิติ: Richard, H. and Andrew, Z. [26]

รูปที่ 3.8 ระนาบภาพ 3 มิติ โดยระบบ W ของจุดบนระนาบภาพ 3 มิติมีตำแหน่งของจุด I กลายเป็น $[u, v, 0]$ จุด P เป็นจุดหลัก (Principal Point) ถ้าลากเส้นจากจุดศูนย์กลางการฉาย N ถึงระนาบภาพที่ขนานกับแกน W และตั้งฉากกับระนาบภาพเรียกว่า แกนหลัก (Principal Axis) และจุดหลักเป็นจุดร่วมของแกนหลัก ระยะทาง d เป็นระยะทางหลัก (Principal Distance) ระหว่าง จุด P และจุด N เวกเตอร์ B จากจุด N และ I กลายเป็น $[u - u_0, v - v_0, 0]$

ดังนั้นเวกเตอร์ A ในรูปที่ 3.8 และเวกเตอร์ B ในรูปที่ 3.9 เขียนเวกเตอร์ด้วยเงื่อนไขเส้นตรงร่วมได้ดังสมการ

$$B = cA \quad (3.25)$$

ที่ c เป็นค่าสเกล และสามารถเขียนการแปลงเวกเตอร์ของภาพได้ด้วยสมการที่ (3.26)

$$T_{I/o} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \quad (3.26)$$

$$\begin{bmatrix} u - u_0 \\ v - v_0 \\ -d \end{bmatrix} = c \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \\ 1 \end{bmatrix} \quad (3.27)$$

$$\begin{aligned} u - u_0 &= c [m_{11}(x - x_0) + m_{12}(y - y_0) + m_{13}(z - z_0) + m_{14}] \\ v - v_0 &= c [m_{21}(x - x_0) + m_{22}(y - y_0) + m_{23}(z - z_0) + m_{24}] \\ -d &= c [m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}] \end{aligned} \quad (3.28)$$

จากสมการที่ (3.28) หาค่า c ได้ดังสมการที่ (3.29)

$$c = \frac{-d}{m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}} \quad (3.29)$$

แทนค่า c สมการที่ (3.29) เข้าไปในสมการที่ (3.28) ได้ดังสมการที่ (3.30)

$$\begin{aligned} u - u_0 &= -d \frac{m_{11}(x - x_0) + m_{12}(y - y_0) + m_{13}(z - z_0) + m_{14}}{m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}} \\ v - v_0 &= -d \frac{m_{21}(x - x_0) + m_{22}(y - y_0) + m_{23}(z - z_0) + m_{24}}{m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}} \end{aligned} \quad (3.30)$$

u, v, u_0 และ v_0 เป็นพิกัดในระนาบภาพความยาวมีหน่วยเป็น เซนติเมตร และในระบบ
ดิจิทัลความยาวมีหน่วยเป็นพิกเซล ดังนั้นในสมการที่ (3.29) จะกลายเป็นสมการที่ (3.31)

$$\begin{aligned} u - u_0 &\Rightarrow \lambda_u (u - u_0) \\ v - v_0 &\Rightarrow \lambda_v (v - v_0) \\ u - u_0 &= -\frac{d}{\lambda_u} \frac{m_{11}(x - x_0) + m_{12}(y - y_0) + m_{13}(z - z_0) + m_{14}}{m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}} \\ v - v_0 &= -\frac{d}{\lambda_v} \frac{m_{21}(x - x_0) + m_{22}(y - y_0) + m_{23}(z - z_0) + m_{24}}{m_{31}(x - x_0) + m_{32}(y - y_0) + m_{33}(z - z_0) + m_{34}} \end{aligned} \quad (3.31)$$

เราเขียนสมการให้อยู่ในรูปใหม่ได้เป็นดังสมการที่ (3.32)

$$\begin{aligned} u &= \frac{L_1 x + L_2 y + L_3 z + L_4}{L_9 x + L_{10} y + L_{11} z + 1} \\ v &= \frac{L_5 x + L_6 y + L_7 z + L_8}{L_9 x + L_{10} y + L_{11} z + 1} \end{aligned} \quad (3.32)$$

โดยที่

$$[d_u, d_v] \equiv \left[\frac{d}{\lambda_u}, \frac{d}{\lambda_v} \right]$$

$$D = -(x_0 m_{31} + y_0 m_{32} + z_0 m_{33} + m_{34})$$

$$\begin{aligned}
L_1 &= \frac{u_0 m_{31} - d_u m_{11}}{D} \\
L_2 &= \frac{u_0 m_{32} - d_u m_{12}}{D} \\
L_3 &= \frac{u_0 m_{33} - d_u m_{13}}{D} \\
L_4 &= \frac{(d_u m_{11} - u_0 m_{31}) x_0 + (d_u m_{12} - u_0 m_{32}) y_0 + (d_u m_{13} - u_0 m_{33}) z_0}{D} \\
L_5 &= \frac{v_0 m_{31} - d_v m_{11}}{D} \\
L_6 &= \frac{v_0 m_{32} - d_v m_{12}}{D} \\
L_7 &= \frac{v_0 m_{33} - d_v m_{13}}{D} \\
L_8 &= \frac{(d_v m_{11} - v_0 m_{31}) x_0 + (d_v m_{12} - v_0 m_{32}) y_0 + (d_v m_{13} - v_0 m_{33}) z_0}{D} \\
L_9 &= \frac{m_{31}}{D} \\
L_{10} &= \frac{m_{32}}{D} \\
L_{11} &= \frac{m_{33}}{D}
\end{aligned} \tag{3.33}$$

สัมประสิทธิ์ L_1 ถึง L_{11} ในสมการที่ (3.32) คือค่าพารามิเตอร์ DLT (DLT Parameters) ซึ่งตัวที่แสดงถึงความสัมพันธ์ระหว่างปริภูมิวัตถุและระนาบภาพ

3.2.1 การเทียบวัตถุกล้อง

วิธีการของ DLT ในส่วนแรกเราจะทำการเทียบวัตถุกล้อง (Camera Calibration) ก่อเพื่อหาค่า DLT พารามิเตอร์ จากสมการที่ (3.32) เขียนใหม่ได้เป็นสมการที่ (3.34)

$$\begin{aligned}
u &= L_1 x + L_2 y + L_3 z + L_4 - L_9 u x - L_{10} u y - L_{11} u z \\
v &= L_5 x + L_6 y + L_7 z + L_8 - L_9 v x - L_{10} v y - L_{11} v z
\end{aligned} \tag{3.34}$$

สมการที่ (3.34) เขียนใหม่ได้

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -ux & -uy & -uz \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -vx & -vy & -vz \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \cdot \\ L_{11} \end{bmatrix} \tag{3.35}$$

ถ้าจุดบนภาพมี n จุดเขียนใหม่ได้เป็นดังสมการที่ (3.36)

$$\begin{bmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 & -u_1z_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1x_1 & -v_1y_1 & -v_1z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & 1 & 0 & 0 & 0 & 0 & -u_nx_n & -u_ny_n & -u_nz_n \\ 0 & 0 & 0 & 0 & x_n & y_n & z_n & 1 & -v_nx_n & -v_ny_n & -v_nz_n \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_{11} \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ \vdots \\ u_n \\ v_n \end{bmatrix} \quad (3.36)$$

จากสมการที่ (3.36) เขียนให้อยู่ในรูปอย่างง่ายได้ดังนี้

$$X \cdot L = Y \quad (3.37)$$

ค่า DLT พารามิเตอร์หาได้ด้วยวิธีกำลังสองน้อยที่สุด (Least Square Method) ดังแสดงในสมการที่ (3.38)

$$\begin{aligned} X \cdot L &= Y \\ (X^t \cdot X) \cdot L &= X^t \cdot Y \\ (X^t \cdot X)^{-1} \cdot (X^t \cdot X) \cdot L &= (X^t \cdot X)^{-1} (X^t \cdot Y) \\ L &= (X^t \cdot X)^{-1} \cdot (X^t \cdot Y) \end{aligned} \quad (3.38)$$

3.3 การสร้างภาพ 3 มิติ (3D Reconstruction)

จากส่วนของการที่เวิรดกกล้องเราสามารถคำนวณค่าพารามิเตอร์กล้องได้ จากสมการที่ (3.34) จัดเรียงใหม่ได้เป็นดังสมการที่ (3.39)

$$\begin{aligned} (uL_9 - L_1)x + (uL_{10} - L_2)y + (uL_{11} - L_3)z &= L_4 - u \\ (vL_9 - L_5)x + (vL_{10} - L_6)y + (vL_{11} - L_7)z &= L_8 - v \end{aligned} \quad (3.39)$$

เขียนสมการที่ (3.39) เขียนใหม่ในรูปเมตริกซ์ได้ดังนี้

$$\begin{bmatrix} uL_9 - L_1 & uL_{10} - L_2 & uL_{11} - L_3 \\ vL_9 - L_5 & vL_{10} - L_6 & vL_{11} - L_7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L_4 - u \\ L_8 - v \end{bmatrix} \quad (3.40)$$

ถ้าจุดบนภาพมี n จุดเขียนใหม่ได้เป็นดังสมการที่ (3.41)

$$\begin{bmatrix} u^{(1)}L_9^{(1)} - L_1^{(1)} & u^{(1)}L_{10}^{(1)} - L_2^{(1)} & u^{(1)}L_{11}^{(1)} - L_3^{(1)} \\ v^{(1)}L_9^{(1)} - L_5^{(1)} & v^{(1)}L_{10}^{(1)} - L_6^{(1)} & v^{(1)}L_{11}^{(1)} - L_7^{(1)} \\ \vdots & \vdots & \vdots \\ u^{(m)}L_9^{(m)} - L_1^{(m)} & u^{(m)}L_{10}^{(m)} - L_2^{(m)} & u^{(m)}L_{11}^{(m)} - L_3^{(m)} \\ v^{(m)}L_9^{(m)} - L_5^{(m)} & v^{(m)}L_{10}^{(m)} - L_6^{(m)} & v^{(m)}L_{11}^{(m)} - L_7^{(m)} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L_4^{(1)} - u^{(1)} \\ L_8^{(1)} - v^{(1)} \\ \vdots \\ L_4^{(m)} - u^{(m)} \\ L_8^{(m)} - v^{(m)} \end{bmatrix} \quad (3.41)$$

จากสมการที่ (3.41) สามารถหาค่า xyz ในระบบ 3 มิติ ได้ด้วยวิธีกำลังสองน้อยที่สุดเหมือนดังในสมการที่ (3.38)

การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองด้วยวิธีการแปลงเชิงเส้นโดยตรงมีความง่ายไม่ยุ่งยากซับซ้อนในการเทียบวัดกล้องและยังใช้เวลาในการประมวลผลรวดเร็วอีกด้วย

ศูนย์วิทยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย

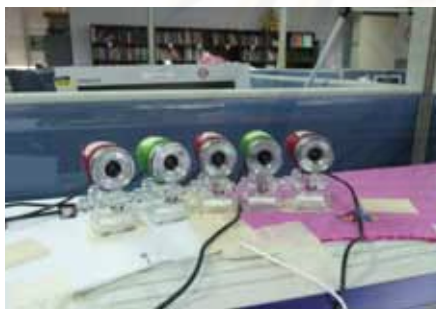
บทที่ 4

ผลการทดลอง

ในบทนี้จะแสดงผลการทดลองในการสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง ด้วยวิธีการแปลงโดยตรงหรือ DLT โดยแบ่งการทดลองออกเป็นสองส่วนในส่วนแรกเป็นการแบ่งส่วนภาพใบหน้าคนด้วยวิธีการจำลองรูปร่างแบบเอกทิพหรือ ASM และส่วนที่สองเป็นส่วนของการสร้างภาพ 3 มิติด้วยวิธีการแปลงโดยตรงหรือ DLT

4.1 อุปกรณ์ที่ใช้ในการทดลอง

การทดลองสร้างภาพ 3 มิติมีอุปกรณ์ที่ใช้ในการทดลองประกอบไปด้วย กล้องถ่ายภาพเว็บแคม ยี่ห้อ MD Tech ความละเอียดขนาด 5 ล้านพิกเซลจำนวน 5 ตัวซึ่งกล้องแต่ละตัววางห่างกัน 5 องศาและวางอยู่ในระนาบเดียวกัน กล้องสี่เหลี่ยมขนาด 10 เซนติเมตร 20 เซนติเมตร และ 30 เซนติเมตร กล้องโอบิกามีรูปดาวทรงกลม และใบหน้าคน ดังแสดงในรูปตามลำดับ



(ก)



(ข)



(ค)



(ง)



(จ)



(ฉ)

รูปที่ 4.1 อุปกรณ์ที่ใช้ในการทดลอง (ก) กล้องเว็บแคม (ข) กล้องขนาด 10 เซนติเมตร
(ค) กล้องขนาด 20 เซนติเมตร (ง) กล้องขนาด 30 เซนติเมตร
(จ) กล้องโอริกามีรูปดาวทรงกลม (ฉ) ใบหน้าคน

4.2 การแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟ

การแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟมีบานข้อมูลที่ใช้ในการทดลอง 2 ส่วนคือส่วนที่เป็นฐานข้อมูลภาพใบหน้าที่ใช้เป็นกลุ่มตัวอย่าง (Training Data) และส่วนที่เป็นภาพใบหน้าที่ใช้ในการแบ่งส่วนภาพ (Test Data)

4.2.1 ฐานข้อมูลภาพ

ฐานข้อมูลภาพใบหน้าที่ใช้เป็นกลุ่มตัวอย่างเป็นการเก็บภาพใบหน้าคนหลายๆ คนในการสร้างเป็นฐานข้อมูล และในการทดลองนี้เราใช้ฐานข้อมูลของ The Color Feret Database [29] ในการเป็นกลุ่มตัวอย่างของขั้นตอนการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟ ตัวอย่างของภาพใบหน้าที่แสดงในรูปที่ 4.2 ภาพใบหน้าที่ใช้ในการแบ่งส่วนภาพเป็นภาพใบหน้าที่เราใช้กล้องเว็บแคมถ่ายภาพขึ้นมาเองดังแสดงในรูปที่ 4.3

ข้อมูลภาพที่เราใช้ในการทดลองที่ถ่ายภาพใบหน้าคนจากกล้องเว็บแคมเราได้ใช้ข้อมูลภาพใบหน้าคน 2 มิติจำนวน 5 คนแต่ละคนมีภาพถ่ายใบหน้า 5 มุมมองดังนั้นภาพใบหน้าคน 2 มิติที่ใช้ในการทดลองมีทั้งหมด 25 มุมมอง



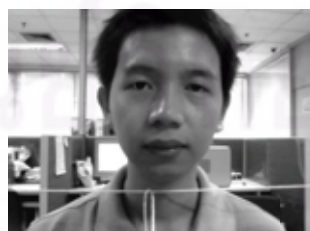
รูปที่ 4.2 ตัวอย่างรูปที่ใช้เป็นกลุ่มตัวอย่าง: The Color FERET Database [29]



(ก)



(ข)



(ค)



(ง)



(จ)

รูปที่ 4.3 ภาพถ่ายที่ได้จากกล้องเว็บแคม (ก), (ข), (ค), (ง) และ (จ)
เป็นภาพจากกล้องตัวที่ 1, 2, 3, 4 และ 5 ตามลำดับ

4.2.2 การสร้างจุดของกลุ่มตัวอย่าง

การสร้างจุดของกลุ่มตัวอย่างเป็นการกำหนดจุดที่มีคุณลักษณะเด่น (Feature Point) ของภาพใบหน้าเพื่อให้งานต่อการแบ่งส่วนภาพ โดยเราได้กำหนดคุณลักษณะเด่นบนภาพใบหน้าเป็นดังนี้ เรากำหนดจุดบนภาพใบหน้าทั้งหมด 300 จุด แบ่งออกเป็น ตาซ้าย 14 จุด ตาขวา 14 จุด จมูก 32 จุด ปาก 34 จุด โคนงหน้า 40 จุด และบริเวณใบหน้า 166 จุด ซึ่งแสดงดังรูปที่ 4.4

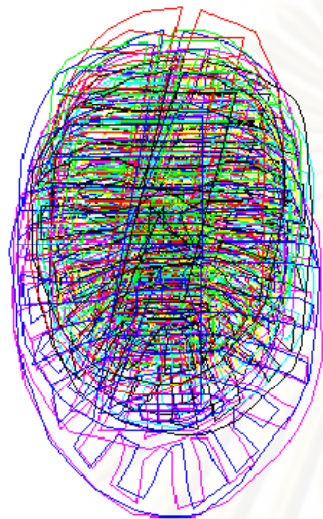


รูปที่ 4.4 จุดบนใบหน้า 300 จุด

4.2.3 กระบวนการปรับแนวรูปร่าง

กระบวนการปรับแนวภาพเป็นการปรับรูปร่างของกลุ่มตัวอย่างที่มีอยู่หลายๆ รูปร่างซึ่งมีการหมุนวัตถุ การเลื่อนขนานและการสเกล ที่มีขนาดไม่เท่ากันให้มีลักษณะที่ใกล้เคียงกันมากที่สุด ซึ่ง

เมื่อเราได้จุดรูปร่างที่เราต้องการในหลายๆ รูปร่างแล้วจะเข้าสู่กระบวนการปรับแนว แสดงดังรูปที่ 4.5 (ก) เป็นรูปที่แสดงการได้จุดรูปร่างของใบหน้า 20 รูปร่างโดยยังไม่ผ่านการปรับแนว (Alignment) และเมื่อนำมาชุดข้อมูลจุดรูปร่างมาผ่านกระบวนการปรับแนวจะได้ดังรูปที่ 4.5 (ข)



(ก)



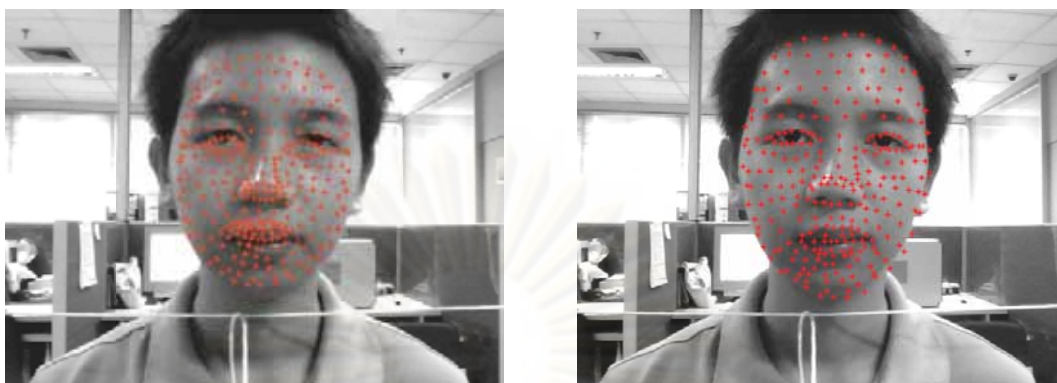
(ข)

รูปที่ 4.5 ชุดข้อมูลจุดที่ได้จากใบหน้า 20 หน้า (ก) จุดบนใบหน้าก่อนการปรับแนว
(ข) จุดบนใบหน้าหลังการปรับแนว

จากรูปที่ 4.5 จะเห็นความแตกต่างของรูปที่ 4.5 (ก) และรูปที่ 4.5(ข) ซึ่งรูปที่ 4.5(ก) จะเห็นว่ามีการหมุนวัตถุ การเลื่อนขนาน และการสเกลที่มีลักษณะต่างๆ ที่ไม่ใกล้เคียงกันใน 20 ชุดข้อมูล และรูปที่ 4.5 (ข) ภาพที่ได้จะมีการหมุนวัตถุ การเลื่อนขนาน และการสเกลที่มีขนาดใกล้เคียงกัน กล่าวคือเป็นการปรับแนวของรูปร่างของข้อมูลกลุ่มตัวอย่างให้มีขนาดใกล้เคียงกัน เพื่อหารูปร่างเฉลี่ย (Mean Shape) นั้นเอง

4.2.4 การฟิตรูปร่างของแบบจำลอง

การฟิตรูปร่างของแบบจำลองเป็นการปรับตัวของรูปร่างเฉลี่ยของชุดข้อมูลตัวอย่างที่ได้จากกระบวนการปรับแนวรูปร่างเพื่อให้รูปร่างเฉลี่ยปรับตัวเข้าหาขอบของภาพวิธีการปรับตัวของรูปร่างเฉลี่ยได้อธิบายในบทที่ 3 วิธีการฟิตรูปร่างของแบบจำลอง เพื่อแบ่งส่วนภาพของภาพใบหน้าคน ตัวอย่างการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบเอกทิวแสดงดังรูปที่ 4.6



(ก)

(ข)

รูปที่ 4.6 การแบ่งส่วนภาพด้วยแบบจำลองรูปร่างแบบแอกทิฟ (ก) การกำหนดจุดเริ่มต้นของรูปร่างแบบจำลอง (ข) ผลของการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟ

4.3 การสร้างภาพใบหน้าคน 3 มิติ

การสร้างภาพใบหน้าคน 3 มิติ ในวิทยานิพนธ์นี้เราใช้วิธีการแปลงโดยตรงหรือ DLT เราแบ่งการสร้างภาพออกเป็น 2 ส่วนคือ การเทียบวัดกล้องและการสร้างภาพ 3 มิติ ซึ่งได้ทดลองดังนี้

4.3.1 การเทียบวัดกล้อง

การเทียบวัดกล้องของวิธีการแปลงโดยตรงเป็นการเทียบวัดกล้องภายในตัวเองเพื่อหาค่าพารามิเตอร์กล้อง มีขั้นตอนการเทียบวัดกล้องดังนี้

1. ภาพถ่ายที่ใช้เป็นภาพอ้างอิงในแต่ละกล้องได้ชุดภาพทั้งหมด 5 ภาพซึ่งในการทดลองเราใช้กล่องสี่เหลี่ยมแสดงดังรูปที่ 4.7



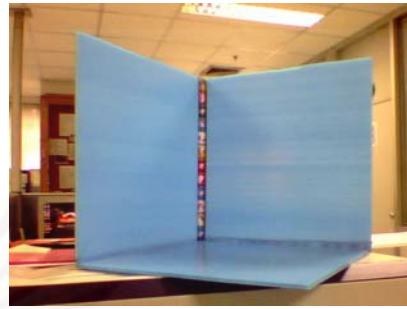
(ก)



(ข)



(ค)



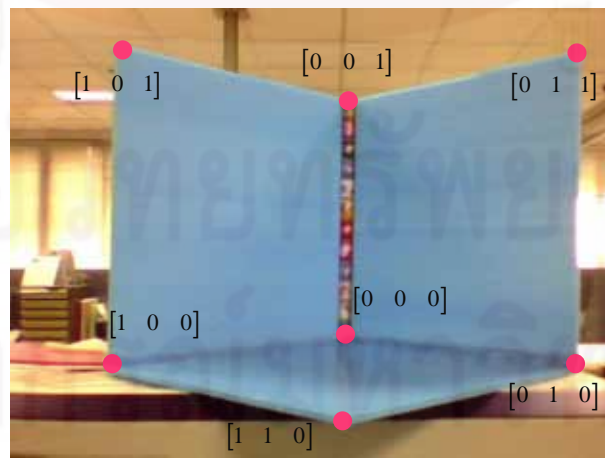
(ง)



(จ)

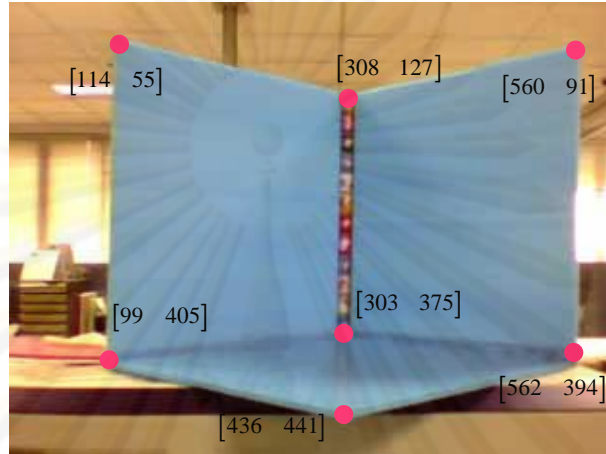
รูปที่ 4.7 ชุดภาพจากกล้องแต่ละมุมมอง (ก), (ข), (ค), (ง) และ (จ)
เป็นภาพจากกล้องที่ 1,2,3,4 และ 5 ตามลำดับ

2. สมมติพิกัดอ้างอิง xyz ในระบบ 3 มิติดังแสดงในรูปที่ 4.8 โดยกำหนดให้เป็นดังนี้
 $xyz = [0\ 0\ 0; 0\ 1\ 0; 0\ 0\ 1; 0\ 1\ 1; 1\ 0\ 0; 1\ 1\ 0; 1\ 0\ 1]$



รูปที่ 4.8 การอ้างอิงพิกัดในระบบ 3 มิติ

3. กำหนดพิกัดอ้างอิงในระบบ 2 มิติของภาพอ้างอิงโดยเราจะต้องกำหนดพิกัดอ้างอิงในระบบ 2 มิติทุกภาพของชุดภาพอ้างอิง กล่าวคือเราต้องกำหนดพิกัดระบบ 2 มิติในภาพทั้ง 5 ภาพ ตัวอย่างการกำหนดจุดดังแสดงในรูปที่ 4.9



รูปที่ 4.9 ตัวอย่างพิกัดอ้างอิงในระบบ 2 มิติ

4. จากขั้นตอนที่ 2 และ 3 ทำให้เรารู้พิกัดของภาพอ้างอิงในระบบ 2 มิติและ 3 มิติ และเราคำนวณค่า L_1 ถึง L_{11} ได้ดังสมการที่ (4.1)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -ux & -uy & -uz \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -vx & -vy & -vz \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \cdot \\ L_{11} \end{bmatrix} \quad (4.1)$$

การประมาณค่า L ซึ่งเป็นค่าพารามิเตอร์กล้อง

4.3.2 การสร้างภาพ 3 มิติ

เมื่อเรารู้ค่าพารามิเตอร์กล้องจากขั้นตอนการเทียบวัดกล้องแล้วเราสามารถสร้างภาพ 3 มิติจากหลายมุมมองด้วยสมการที่ (4.2)

$$\begin{bmatrix} uL_9 - L_1 & uL_{10} - L_2 & uL_{11} - L_3 \\ vL_9 - L_5 & vL_{10} - L_6 & vL_{11} - L_7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L_4 - u \\ L_8 - v \end{bmatrix} \quad (4.2)$$

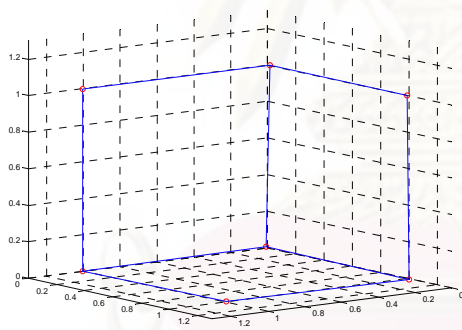
กรณีจุดบนภาพมี n จุดเขียนใหม่ได้เป็นดังสมการที่ (4.3)

$$\begin{bmatrix} u^{(1)}L_9^{(1)} - L_1^{(1)} & u^{(1)}L_{10}^{(1)} - L_2^{(1)} & u^{(1)}L_{11}^{(1)} - L_3^{(1)} \\ v^{(1)}L_9^{(1)} - L_5^{(1)} & v^{(1)}L_{10}^{(1)} - L_6^{(1)} & v^{(1)}L_{11}^{(1)} - L_7^{(1)} \\ \vdots & \vdots & \vdots \\ u^{(m)}L_9^{(m)} - L_1^{(m)} & u^{(m)}L_{10}^{(m)} - L_2^{(m)} & u^{(m)}L_{11}^{(m)} - L_3^{(m)} \\ v^{(m)}L_9^{(m)} - L_5^{(m)} & v^{(m)}L_{10}^{(m)} - L_6^{(m)} & v^{(m)}L_{11}^{(m)} - L_7^{(m)} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L_4^{(1)} - u^{(1)} \\ L_8^{(1)} - v^{(1)} \\ \vdots \\ L_4^{(m)} - u^{(m)} \\ L_8^{(m)} - v^{(m)} \end{bmatrix} \quad (4.3)$$

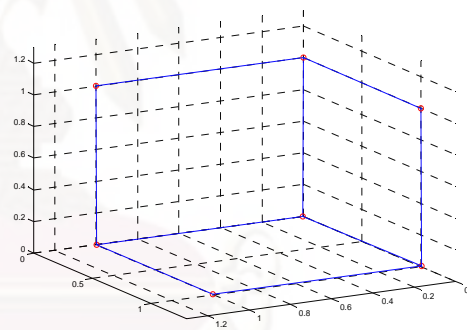
4.4 การทดลองสร้างภาพ 3 มิติ

4.4.1 การสร้างภาพ 3 มิติหลายมุมมองของกล่องสี่เหลี่ยม

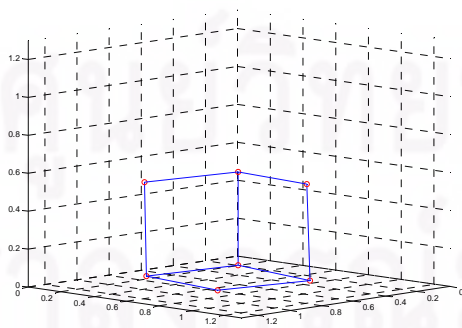
การสร้างภาพ 3 มิติหลายมุมมองของกล่องสี่เหลี่ยมนี้เป็นการทดสอบการสร้างภาพ 3 มิติของวิธีการแปลงโดยตรง ดังรูปที่ 4.10 สร้างกล่องสี่เหลี่ยมขนาด 10 เซนติเมตร, 20 เซนติเมตร และ 30 เซนติเมตรโดยอ้างอิงกับกล่องสี่เหลี่ยมขนาดตัวเอง



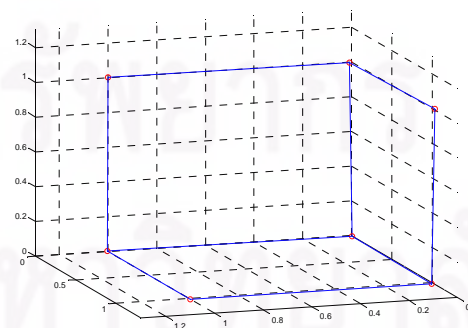
(ก)



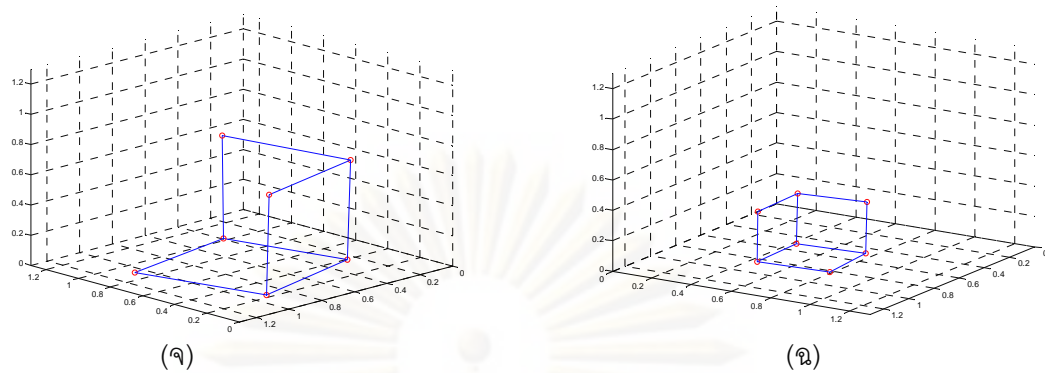
(ข)



(ค)



(ง)

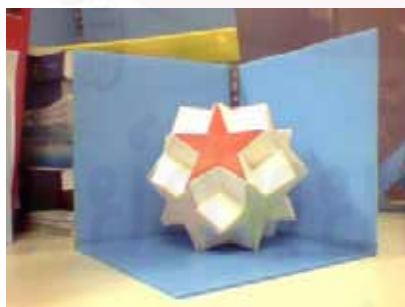


รูปที่ 4.10 การสร้างภาพ 3 มิติของกล่องสี่เหลี่ยม

- (ก) กล่องสี่เหลี่ยมอ้างอิงขนาด 10 เซนติเมตรสร้างกล่องขนาด 10 เซนติเมตร
- (ข) กล่องสี่เหลี่ยมอ้างอิงขนาด 20 เซนติเมตรสร้างกล่องขนาด 20 เซนติเมตร
- (ค) กล่องสี่เหลี่ยมอ้างอิงขนาด 20 เซนติเมตรสร้างกล่องขนาด 10 เซนติเมตร
- (ง) กล่องสี่เหลี่ยมอ้างอิงขนาด 30 เซนติเมตรสร้างกล่องขนาด 30 เซนติเมตร
- (จ) กล่องสี่เหลี่ยมอ้างอิงขนาด 30 เซนติเมตรสร้างกล่องขนาด 20 เซนติเมตร
- (ฉ) กล่องสี่เหลี่ยมอ้างอิงขนาด 30 เซนติเมตรสร้างกล่องขนาด 10 เซนติเมตร

4.4.2 การสร้างภาพ 3 มิติหลายมุมมองของโริกามีรูปร่างดาว

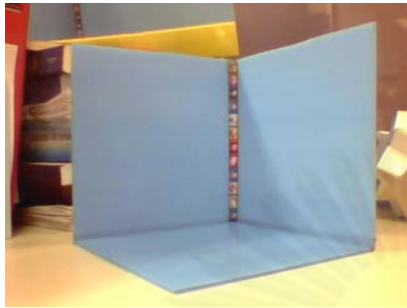
การสร้างภาพ 3 มิติของโริกามีรูปร่างดาวเป็นการสร้างจากภาพ 2 มุมมองโดยใช้กล่องสี่เหลี่ยมขนาด 20 เซนติเมตรเป็นตัวอ้างอิง ผลลัพธ์แสดงดังรูปที่ 4.11



(ก)



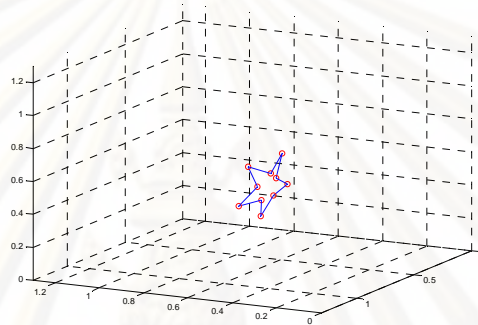
(ข)



(ค)



(ง)

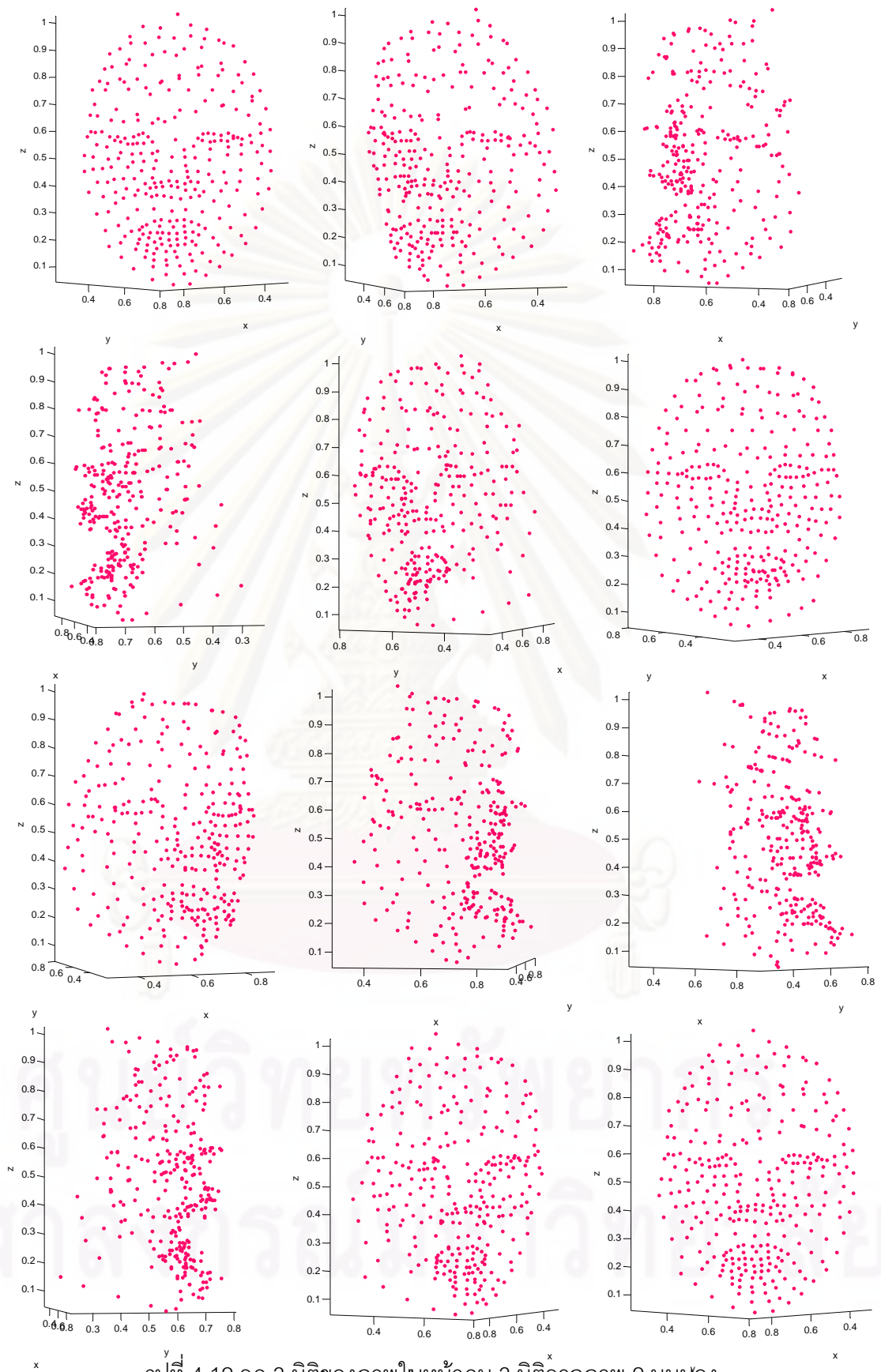


(จ)

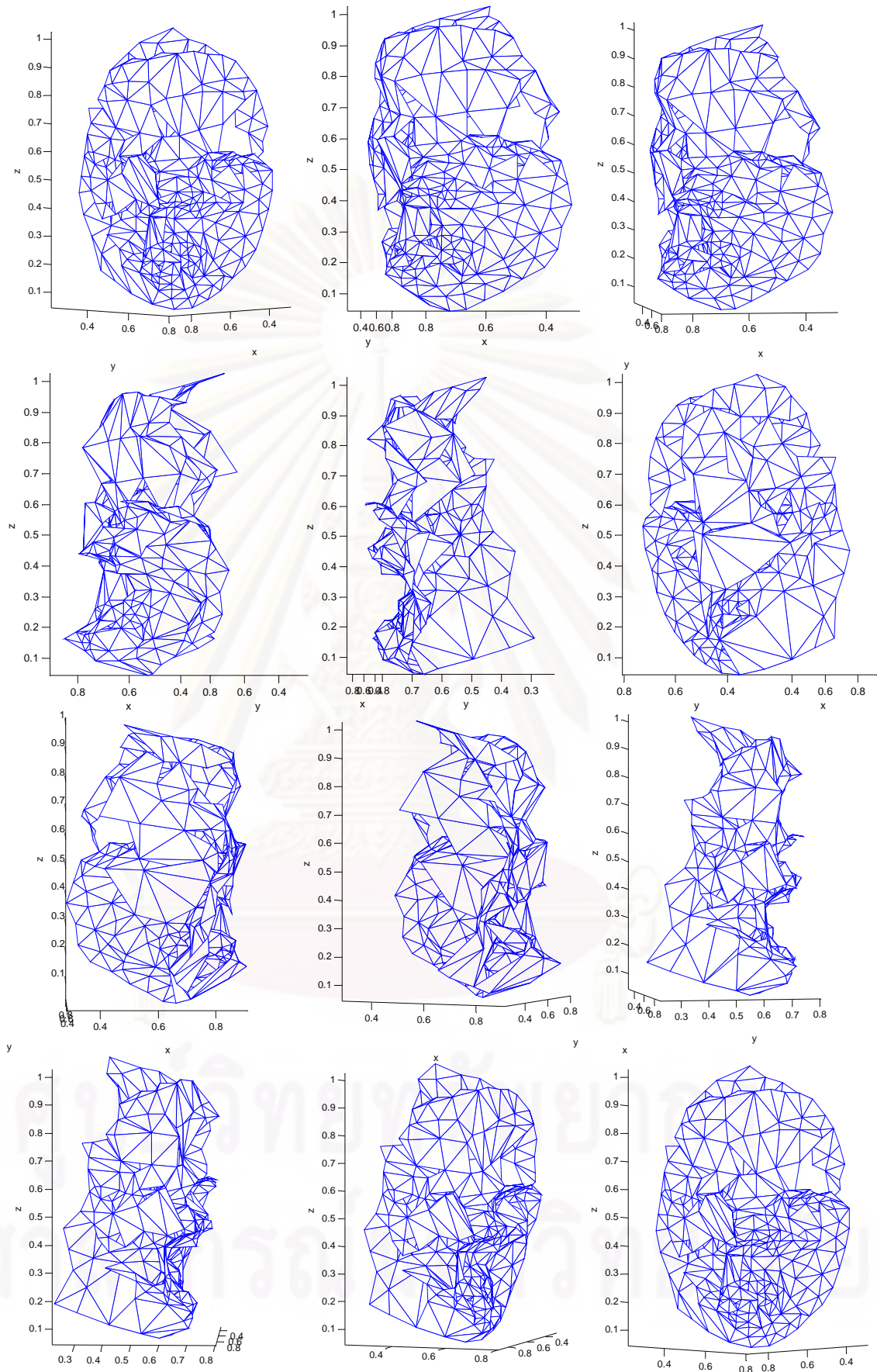
รูปที่ 4.11 สร้างภาพโอริกามีรูปดาวจากภาพ 2 มุมมอง (ก)และ(ข) ภาพโอริกามีรูปดาว 2 มุมมอง (ค) และ (ง) กล้องสี่เหลี่ยมข้างอิง 2 มุมมอง (จ) จุดของภาพดาวในระบบ 3 มิติ

4.4.3 การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง

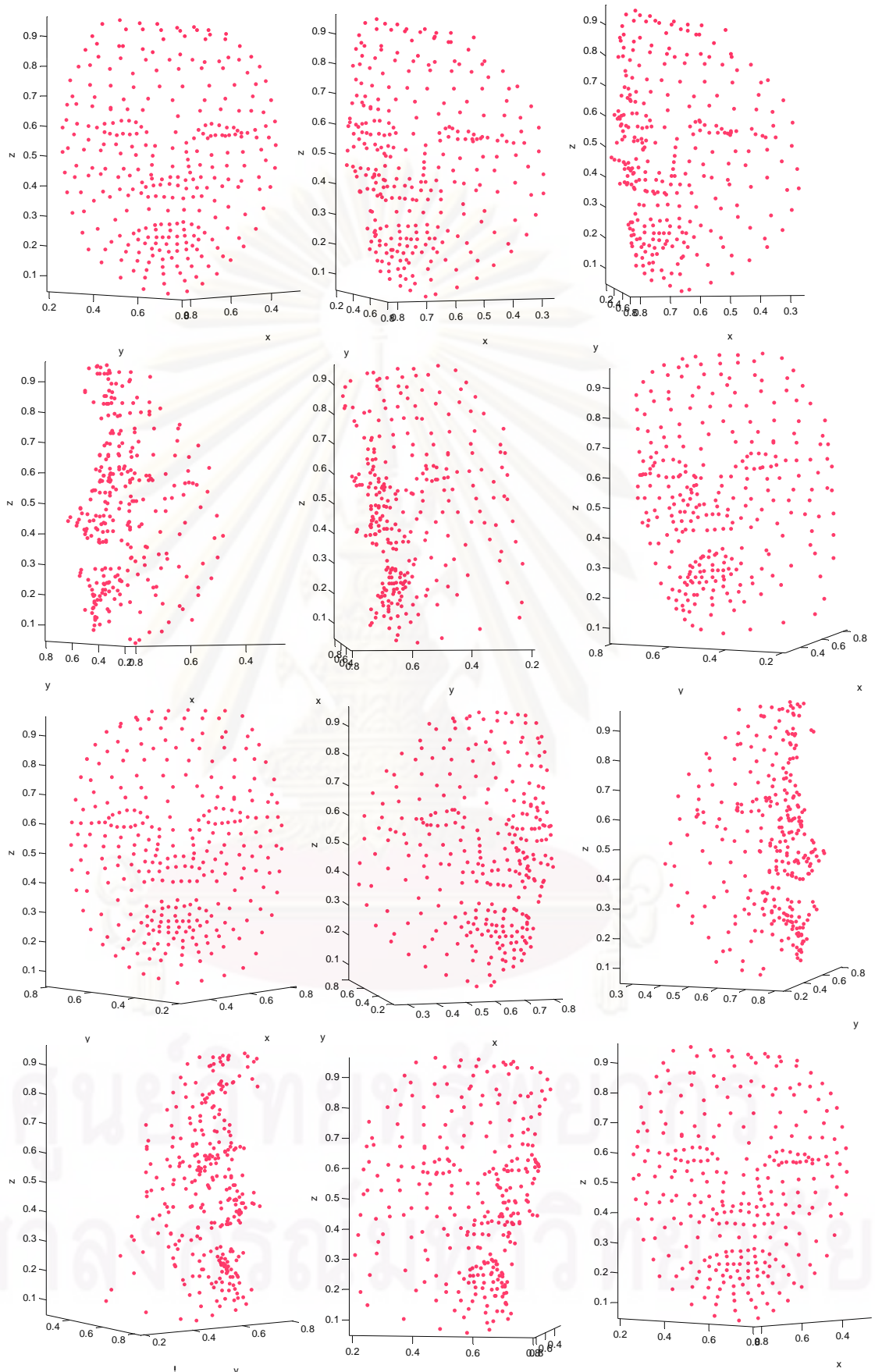
เราใช้ข้อมูลภาพใบหน้าคนตามรูปที่ 4.3 ทั้งหมด 5 ภาพและใช้กล้องสี่เหลี่ยมในรูปที่ 4.7 เป็นตัวอ้างอิง การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง เราได้กำหนดการสร้างภาพ 3 มิติคือเราสร้างภาพ 3 มิติด้วย 5 มุมมองและการกำหนดการถ่ายภาพเราได้วางกล้องแต่ละตัวห่างกัน 5 องศา โดยเราไม่ได้ทำการหมุนกล้องหรือหมุนวัตถุเลย ดังนั้นภาพที่ได้ในแต่ละมุมมองจึงเป็นภาพที่แตกต่างกัน 5 องศา โดยเราได้ทำการทดลองสร้างภาพใบหน้าคน 3 มิติจาก 2 มุมมองเราใช้ภาพที่ 1 และภาพที่ 3 สร้างภาพใบหน้าคน 3 มิติจาก 3 มุมมองเราใช้ภาพที่ 1, 3 และ 5 การสร้างภาพใบหน้าคน 3 มิติจาก 4 มุมมองเราใช้ภาพที่ 1, 2, 3 และ 4 และสร้างภาพใบหน้าคน 3 มิติจาก 5 มุมมองเราใช้ภาพที่ 1, 2, 3, 4 และ 5 ตัวอย่างของผลการทดลองแสดงดังรูปข้างล่าง โดยดูจากขวาไปซ้ายและบนลงล่าง



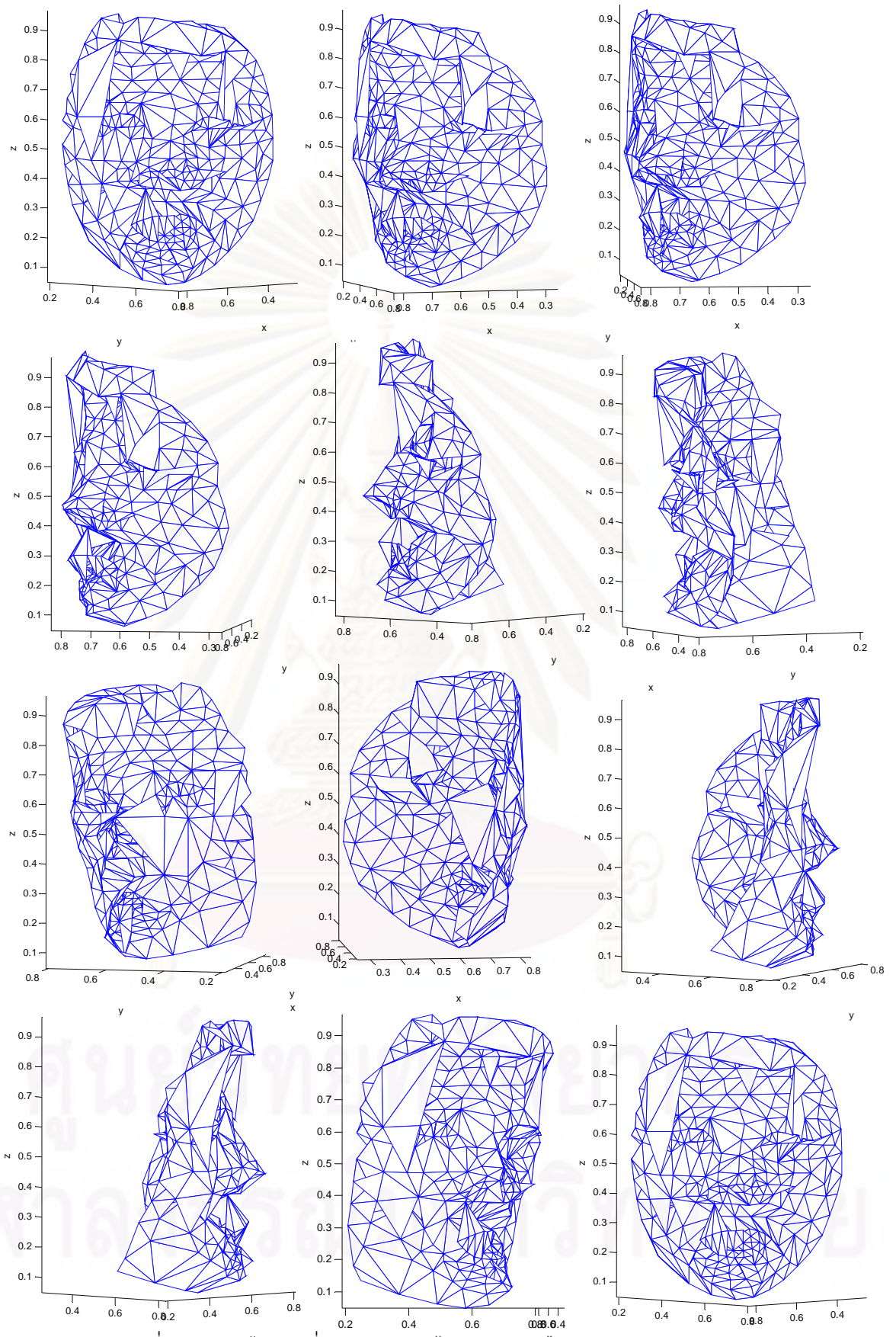
รูปที่ 4.12 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 2 มุมมอง



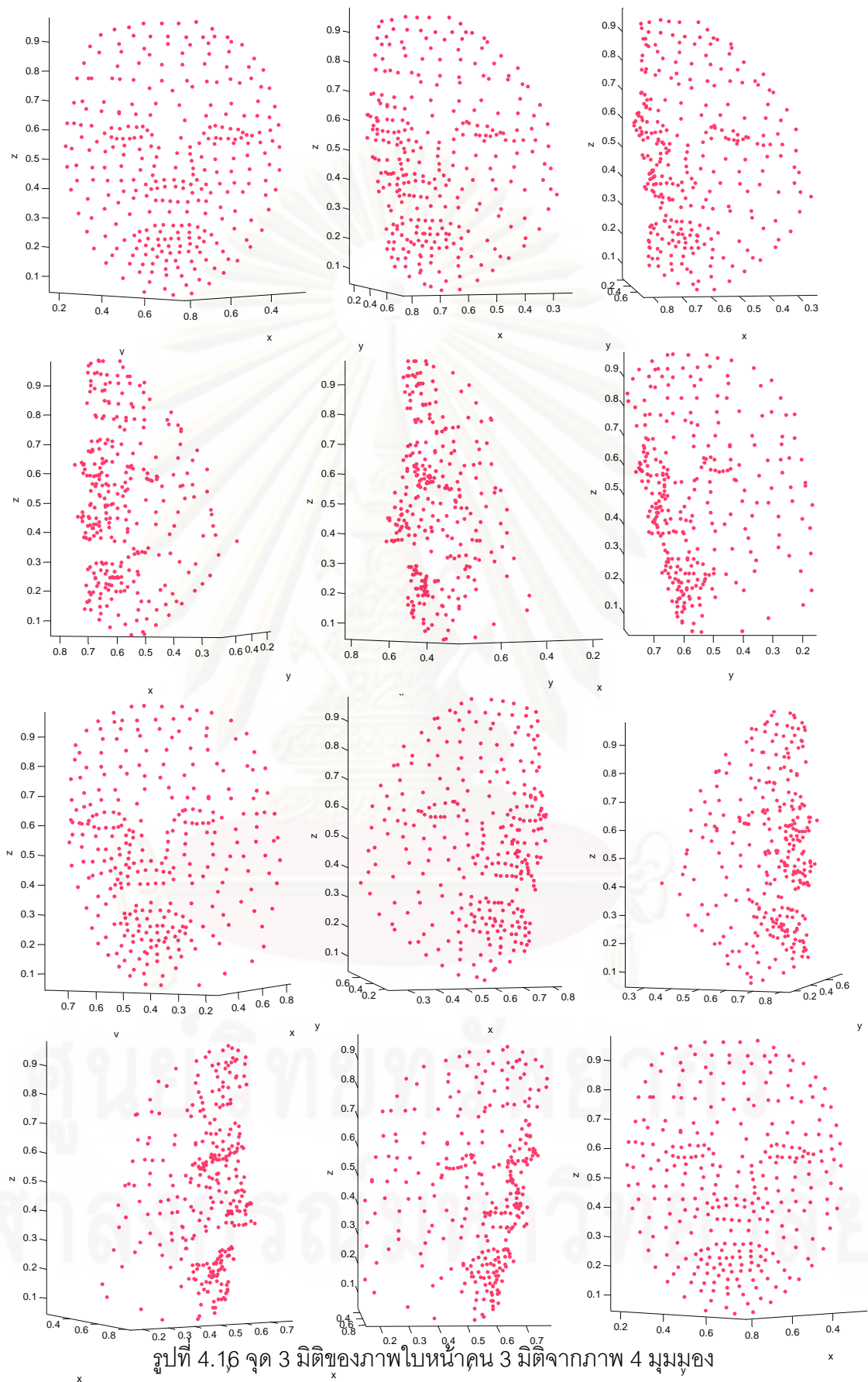
รูปที่ 4.13 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 2 มุมมอง

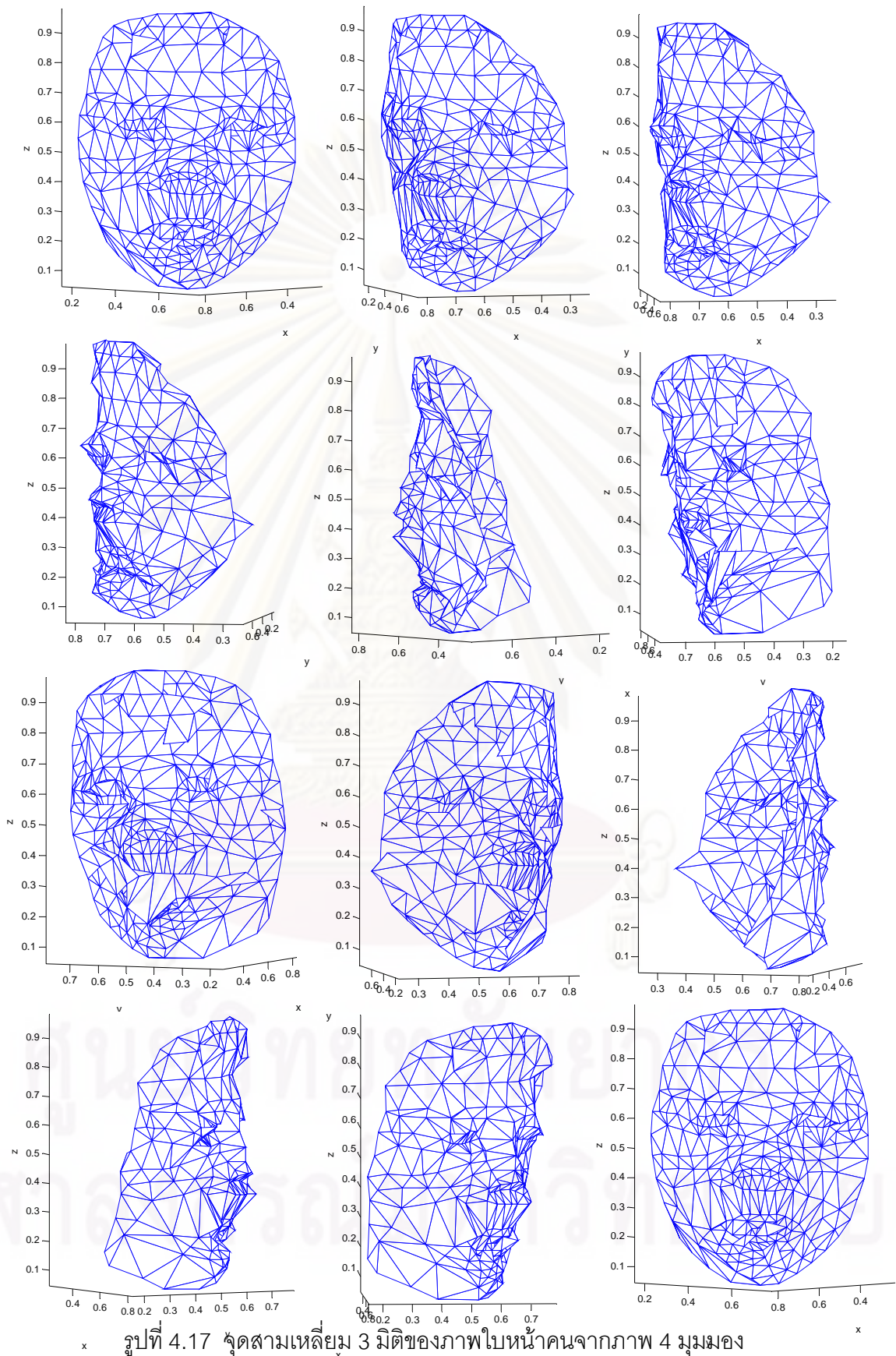


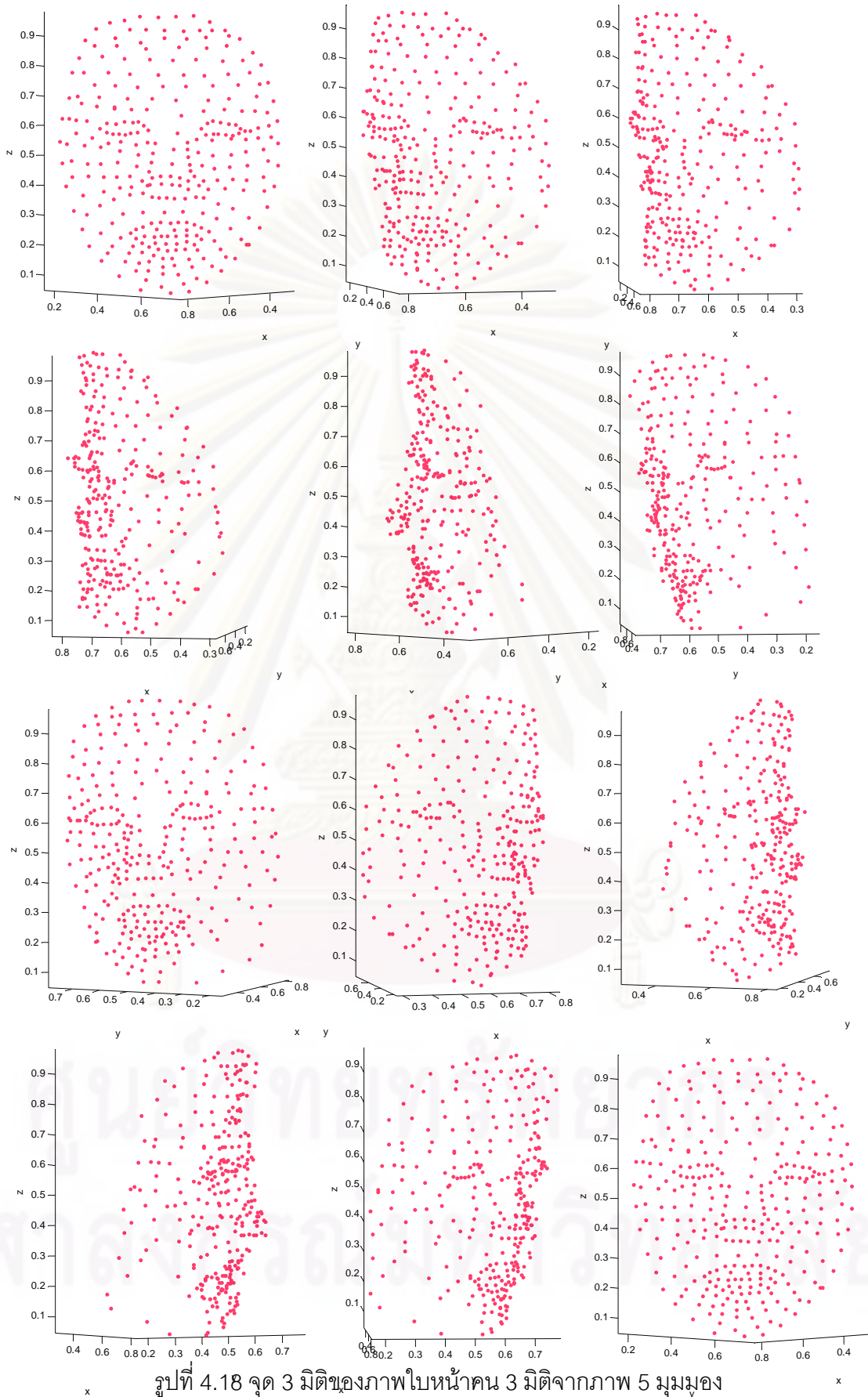
รูปที่ 4.14 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 3 มุมมั่ง



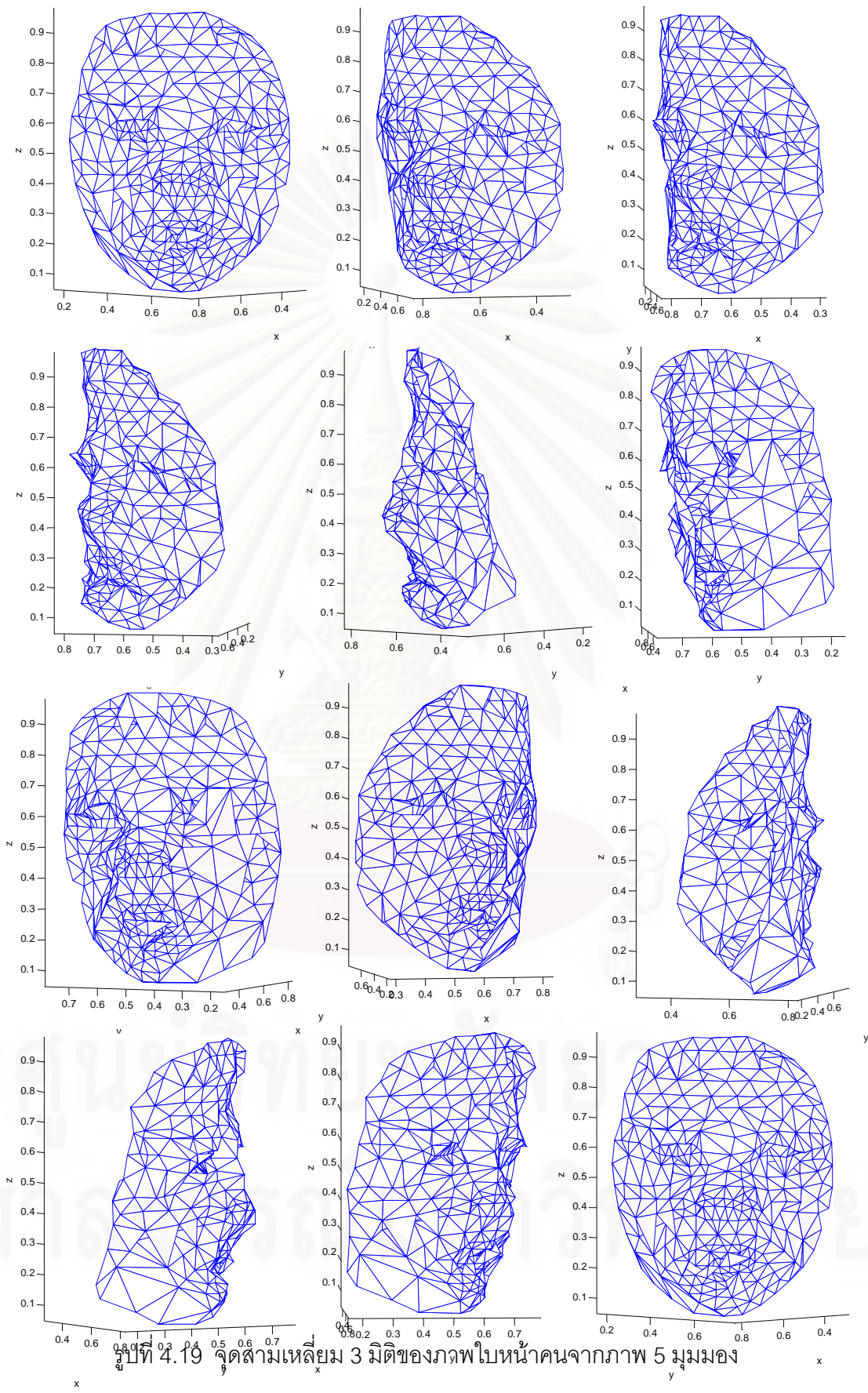
รูปที่ 4.15 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 3 มุมมอง



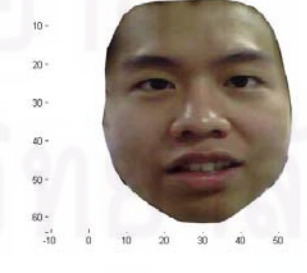
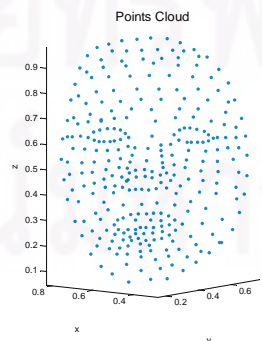
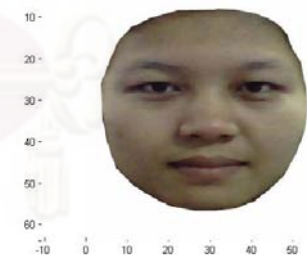
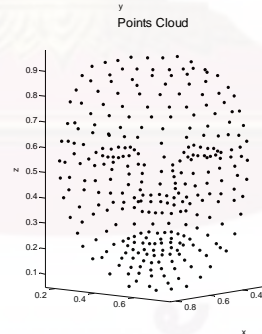
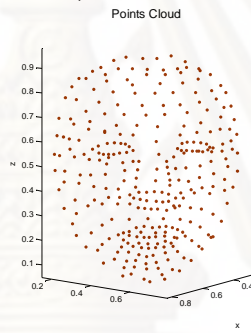
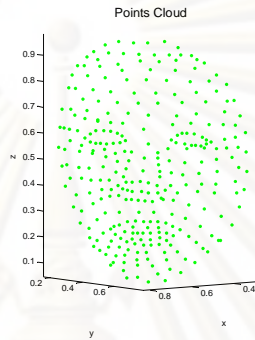
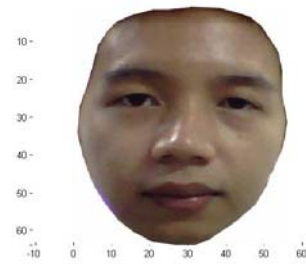
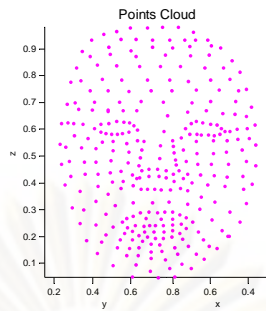




รูปที่ 4.18 จุด 3 มิติของภาพใบหน้าคน 3 มิติจากภาพ 5 มุมมอง



รูปที่ 4.19 จุดสามเหลี่ยม 3 มิติของภาพใบหน้าคนจากภาพ 5 มุมมอง



รูปที่ 4.20 ภาพใบหน้าคน 3 มิติ

จากรูปที่ 4.20 แสดงภาพใบหน้าคน 3 มิติจากหลาย ๆ ฐานข้อมูลโดยรูปซ้ายมือเป็นภาพถ่ายใบหน้าคน 2 มิติ รูปกลางเป็นจุดภาพใบหน้าคน 3 มิติและรูปด้านขวาเป็นรูปภาพใบหน้าคน 3 มิติ เราได้ทดลองสร้างภาพใบหน้าคน 3 มิติจำนวน 5 ฐานข้อมูล และได้หาประสิทธิภาพของการสร้างภาพ 3 มิติจากภาพหลายมุมมองโดยการแยกแยะความเหมือนของภาพใบหน้าคน 2 มิติและ 3 มิติด้วยมนุษย์ เราใช้ข้อมูลทั้งหมด 5 ภาพใบหน้าให้ผลค่าเฉลี่ยความเหมือนของภาพใบหน้าคน 2 มิติกับภาพใบหน้าคน 3 มิติ ดังนี้ ค่าเฉลี่ยความเหมือนของใบหน้าคนในฐานข้อมูลที่ 1 มีค่าความเหมือน 80.7% ใบหน้าคนในฐานข้อมูลที่ 2 มีค่าความเหมือน 86.75 % ใบหน้าคนในฐานข้อมูลที่ 3 มีค่าความเหมือน 79.75 % ใบหน้าคนในฐานข้อมูลที่ 4 มีค่าความเหมือน 76.5 % และใบหน้าคนในฐานข้อมูลที่ 5 มีค่าความเหมือน 85.15 % ตามลำดับ



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้นำเสนอการสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมอง โดยเสนอการสร้างภาพ 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรงหรือ DLT เพราะเป็นวิธีการที่ง่ายไม่ยุ่งยากและซับซ้อนน้อยกว่าวิธีการอื่น โดยแบ่งการทดลองออกเป็น 2 ส่วนคือการแบ่งส่วนภาพและการสร้างภาพ 3 มิติ และเริ่มการทดลองด้วยการถ่ายภาพใบหน้าคนด้วยกล้องแบบเว็บแคมจำนวน 5 ตัวซึ่งได้ภาพ 5 มุมมองด้วย แล้วนำภาพเหล่านั้นมาทำการแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟ

การแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟ ข้อดีของวิธีการนี้คือสามารถคงรูปร่างของวัตถุที่เราต้องการแบ่งส่วนภาพได้ ซึ่งเป็นผลดีต่อการแบ่งส่วนภาพในกรณีที่ขอบของวัตถุไม่ชัดเจน เนื่องใบหน้าคนมีขอบของภาพบริเวณคางกับคอไม่ชัดเจนดังนั้นการแบ่งส่วนภาพด้วยวิธีอื่นอาจจะติดปัญหาตรงส่วนนี้ได้ แต่การแบ่งส่วนภาพด้วยวิธีการจำลองรูปร่างแบบแอกทิฟยังคงรูปร่างของใบหน้าไว้ได้และในภาพใบหน้าคนจะมีจุดที่เป็นคุณลักษณะเด่นน้อยทำให้การค้นหาจุดบนใบหน้ายาก ตัวอย่างเช่นบริเวณแก้มและหน้าผากการค้นหาจะเป็นไปได้ยากมาก เนื่องจากมีบริเวณดังกล่าวไม่มีจุดที่เป็นคุณลักษณะเด่นเลย และวิธีการจำลองรูปร่างแบบแอกทิฟสามารถที่จะหาจุดบริเวณนี้ได้ด้วยการสร้างกลุ่มตัวอย่าง กล่าวคือวิธีการจำลองรูปร่างแบบแอกทิฟสามารถแบ่งส่วนภาพและหาจุดต่าง ๆ ในบริเวณใบหน้าได้ในครั้งเดียว โดยไม่ต้องแยกการแบ่งส่วนภาพและการค้นหาจุดออกจากกัน ส่วนข้อดีของวิธีการจำลองรูปร่างแบบแอกทิฟคือจะต้องมีการใช้ข้อมูลใบหน้าคนจำนวนมากเพื่อใช้ในการเป็นกลุ่มตัวอย่างของภาพเพื่อจะทำให้รูปร่างแบบจำลองที่ใช้ในการแบ่งส่วนภาพสามารถแบ่งส่วนภาพได้ดีขึ้นและจะใช้เวลาค่อนข้างนานในการสร้างกลุ่มตัวอย่างด้วย

สร้างภาพใบหน้าคน 3 มิติหลายมุมมองเราสร้างภาพ 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรง เป็นการสร้างภาพ 3 มิติด้วยการใช้อ้างอิงของวัตถุในระบบ 3 มิติเพื่อหาค่าพารามิเตอร์กล้อง โดยเริ่มจากการใช้ภาพถ่ายของกล้องสี่เหลี่ยมเป็นจุดอ้างอิงในระบบ 2 มิติและ 3 มิติเพื่อคำนวณหาพารามิเตอร์กล้อง เมื่อรู้ค่าพารามิเตอร์กล้องแล้วเราจะสร้างภาพ 3 มิติได้ด้วยจุดต่าง ๆ บนใบหน้าคนในแต่ละมุมมอง จากการแบ่งส่วนภาพเราจะได้จุดบนภาพใบหน้าในระบบ 2 มิติในหลายๆ มุมมอง เมื่อเรามีค่าพารามิเตอร์กล้องและจุดบนใบหน้าแล้วเราก็สามารถสร้างภาพ

ใบหน้าคน 3 มิติได้ การสร้างภาพ 3 มิติด้วยวิธีการแปลงเชิงเส้นโดยตรงนี้มีข้อดีคือเป็นวิธีการที่ไม่มีความยุ่งยากซับซ้อนและมีความรวดเร็วในการในการเทียบวัดกล้องและยังสามารถคำนวณจุด 3 มิติได้โดยตรงใช้เวลาในการคำนวณน้อย ข้อด้อยคือถ้าภาพที่ใช้ในการสร้างภาพ 3 มิติน้อยมุมมองจะทำให้จุดที่ได้ใน 3 มิติมีค่าความผิดพลาดมาก ในการทดลองจะเห็นว่าการสร้างภาพใบหน้าคน 3 มิติจากภาพ 2 มิติใน 2 มุมมองให้ค่าความผิดพลาดมากที่สุดเมื่อเทียบกับ 3 มุมมอง 4 มุมมอง และ 5 มุมมอง โดยที่จุด 3 มิติของภาพ 5 มุมมองให้ผลดีกว่ามุมมองอื่น ๆ

5.2 ข้อเสนอแนะ

การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองเป็นเทคโนโลยีที่ค่อนข้างใหม่มีการศึกษาวิจัยกันไม่มากนักเนื่องมาจากมีข้อจำกัดทางด้านอุปกรณ์ถ้าต้องการความถูกต้องแม่นยำมาก การสร้างภาพใบหน้าคน 3 มิติจากภาพหลายมุมมองสามารถนำไปประยุกต์ใช้กับวัตถุอื่นที่ไม่ใช่ใบหน้าคนหรือใช้กับกล้องวงจรปิดเพื่อใช้สร้างใบหน้าคนในระบุดัตว์ผู้ร้ายได้

รายการอ้างอิง

- [1] Parke, F.I. Computer Generated Animation of Face. In ACM National Conference, November 1972.
- [2] Yuankui, H., Ying, Z., and Zengfu, W. Reconstruction of 3D Face from a Single 2D Image for Face Recognition. IEEE International Workshop on VS-PETS, 2005.
- [3] Seitz, S.M., and Dyer, C.R. Photorealistic Scene Reconstruction by Voxel Coloring. IEEE International on CVPR, 1997.
- [4] Kolmogorov, V., and Zabih, R. Multi-Camera Scene Reconstruction via Graph Cuts. Proceedings of ECCV, 2002.
- [5] Faugeras, O., and Keriven, R. Complete Dense Stereovision Using Level Set Methods. European Conf. on Computer Vision, 1998.
- [6] Curless, B., and Levoy, M. A Volumetric Method for Building Complex Models From Range Images. Proceeding of SIGGRAPH, 1996.
- [7] Tang, C.K., and Medioni, G. Curvature-Augmented Tensor Voting for Shape Inference From Noisy 3D Data. IEEE International on TPAMI, 24(6):858–864, 2002.
- [8] Zhao, H.K., Osher, S., Merriman, B., and Kang, M. Implicit and Non-Parametric Shape Reconstruction From Unorganized Data Using a Variational Level Set Method. CVIU, 2000.
- [9] Whitaker, R. A Level-Set Approach to 3D Reconstruction From Range Data, IJCV, 29(3):203–231, 1998.
- [10] Lhuillier, M., and Quan, L. Surface Reconstruction by Integrating 3D and 2D Data of Multiple Views. In: IEEE Conf. on Computer Vision, 2003.
- [11] Zeng, G., Sylvain, P., Long, Q., and Lhuillier, M. Surface Reconstruction by Propagating 3D Stereo Data in Multiple 2D Images. Proceedings of ECCV, 2005.
- [12] Yuxiao, H., Dalong, J., Shuicheng, Y., Lei, Z., and Hongjiang, Z. Automatic 3D Reconstruction for Face Recognition. IEEE Conf on FGR, 2004.
- [13] Volker, B., and Thomas, V. A Morphable Model for The Synthesis of 3D Face. Proceeding of SIGGRAPH, 1999.

- [14] Mark, C., Chia-Yen, C., Gareth, B., Patrice, D., and Georgy, G. Evaluation of 3D Face Analysis and Synthesis Techniques. IEEE International Conference on Multimedia and Expo, 2004.
- [15] Tsai, R. A versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-The-Shelf TV Cameras and Lenses. IEEE Journal of Robotics and Automation, 1987.
- [16] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. Active Shape Models – Their Training and Application. Computer Vision and Image Understanding, 1995.
- [17] Cootes, T. F., and Taylor, C. J. Active Shape Models - Smart Snakes. In British Machine Vision Conference, 1992.
- [18] Goodall, C. Procrustes methods in the statistical analysis of shape. Journal of the Royal Statistical Society B, 53(2):285{339, 1991.
- [19] Frangi, A. F., Niessen, W. J., and Viergever, M. A. Three-Dimensional Modeling for Functional Analysis of Cardiac Images: A review. IEEE Transactions on Medical Imaging, January 2001.
- [20] Malladi, R., Sethian, J., and Vemuri, B. C. Shape Modeling With Front Propagation: A Level Set Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, February 1995.
- [21] Rafael, C. G., Richard, E. W., and Steven, L. E. Digital Image Processing. Pearson Education 2004.
- [22] Svoboda, T., Kybic, J., and Hlavac, V. Image Processing Analysis and Machine Vision. 2007.
- [23] Zhang, L., Ai, H., Xin, S., Huang, C., Tsukiji, C., and Lao, S. Robust Face Alignment Based on Local Texture Classifiers. In: Proceedings of ICIP. 2005.
- [24] Yan, S., Li, M., Zhang, H., and Cheng, Q. Ranking Prior Likelihood Distributions for Bayesian Shape Localization Framework. In: Proceedings of ICCV. 2003.
- [25] Cootes, T., Taylor, C. and Lanitis, A. Active Shape Models: Evaluation of a Multi-Resolution Method for Improving Image Search. Proceedings of the British

Machine Vision Conference, 1994.

- [26] Richard, H., and Andrew, Z. Multiple View Geometry in Computer Vision Second Edition. Cambridge University Press, March 2004.
- [27] Wei, Q., Zhou, G., Zhu, Q., and Zhang Q. On DLT Method for CCD Camera Calibration. 3rd International Conference on Signal Processing ICSIGP, 1996.
- [28] Zahalka ,F. Experiment of Accuracy of DLT Method. Experiment of the Accuracy of Space Points Reconstruction Using DLT Method. Symposium on the 3D Analysis, June 2006.
- [29] The Color FERET Database. National Institute of Standards and Technology.
- [30] Thomas, V. Synthesis of Novel Views From a Single Face Image. International Journal of Computer Vision, 1998.
- [31] Wallraven, C., Blanz, V., and Vetter, T. 3D-Reconstruction of Faces: Combining Stereo with Class-Based Knowledge. Proceedings of DAGM 1999,
- [32] Zhou, Y., Zhang, W., Tang, X., and Shum, H. A Bayesian Mixture Model for Multi-View Face Alignment. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

การวิเคราะห์องค์ประกอบแกนหลัก หรือ PCA

เราจะอธิบายการใช้ PCA ในเวกเตอร์ x_i เมตริกซ์ความแปรปรวนร่วม $n \times n$ ซึ่งอาจจะมีค่าที่ใหญ่มากแต่เราก็สามารถหาเวกเตอร์ลักษณะเฉพาะและค่าลักษณะเฉพาะได้จากเมตริกซ์ $s \times s$ ขนาดเล็กๆ ที่ได้จากข้อมูลที่มีการจำแนกเวกเตอร์ลักษณะเฉพาะที่มีเมตริกซ์เป็นจัตุรัส

เมตริกซ์ D เป็นข้อมูลของแต่ละเวกเตอร์

$$D = ((x_1 - \bar{x}) | \dots | (x_s - \bar{x})) \quad (\text{ก.1})$$

เมตริกซ์แปรปรวนร่วมเขียนได้ดังสมการ

$$S = \frac{1}{s} DD^T \quad (\text{ก.2})$$

กำหนดให้ T เป็นเมตริกซ์ $s \times s$

$$T = \frac{1}{s} D^T D \quad (\text{ก.3})$$

การปรับแนวรูปร่าง

การแก้ปัญหาของการหาค่าพารามิเตอร์ที่เหมาะสมของการปรับแนวด้วยการใช้ค่าถ่วงน้ำหนักของค่าความแตกต่างที่น้อยที่สุดของผลรวมกำลังสอง ดังนั้นเราสมมติข้อมูล n จุดใน 2 ชุดข้อมูล x_i และ x'_i กำหนดการแปลง $x' = T_t(x)$ ด้วยพารามิเตอร์ t ดังนั้นหาค่าที่น้อยที่สุดได้ด้วยสมการ (ก.4)

$$E = \sum_{i=1}^n (x'_i - T_t(x_i))^T W_i (x'_i - T_t(x_i)) \quad (\text{ก.4})$$

การย้ายตำแหน่ง

การย้ายตำแหน่งสามารถเขียนสมการได้ดังนี้

$$T_t(x) = x + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (\text{ก.5})$$

เมื่อ $t = (t_x, t_y)^T$ แก้ปัญหาของพารามิเตอร์ (t_x, t_y) ด้วยสมการเชิงเส้น (Linear Equation)

$$\left(\sum_{i=1}^n W_i \right) \begin{pmatrix} t_x \\ t_y \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n W_i (x'_i - x_i) \end{pmatrix} \quad (\text{ก.6})$$

กรณีที่มีตัวถ่วงน้ำหนักแบบไอโซโทรปิก (Isotropic Weights) $W_i = w_i I$ แก้ปัญหาโดย

$$t = \left(\sum_{i=1}^n W_i \right)^{-1} \left(\sum_{i=1}^n w_i (x'_i - x_i) \right) \quad (\text{ก.7})$$

กรณีที่ไม่มีตัวถ่วงน้ำหนักแก้ปัญหาด้วยความแตกต่างของค่าเฉลี่ย

$$t = \bar{x}' - \bar{x} \quad (\text{ก.8})$$

การแปลงและการสเกล

$$T_t(x) = sx + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (\text{ก.9})$$

และ $t = (s, t_x, t_y)^T$ แก้ปัญหาของพารามิเตอร์ (t_x, t_y) ด้วยสมการเชิงเส้น

$$\begin{pmatrix} \frac{S_x w_x}{S w_x} & \frac{S^T w_x}{S w} \end{pmatrix} \begin{pmatrix} s \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} \frac{S_x w'_x}{S w_x} \end{pmatrix} \quad (\text{ก.10})$$

ซึ่ง

$$\begin{aligned} S_x w_x &= \sum x_i^T W_i x_i & S w_x &= \sum W_i x_i & S w &= \sum W_i \\ S_x w_{x'} &= \sum x_i^T W_i x'_i & S w_{x'} &= \sum W_i x'_i \end{aligned} \quad (\text{ก.11})$$

กรณีที่ไม่มีตัวถ่วงน้ำหนักแก้ปัญหาง่ายๆ ได้ดังนี้

$$\begin{pmatrix} S_{xx} + S_{yy} & S_x & S_y \\ S_x & n & 0 \\ S_y & 0 & n \end{pmatrix} \begin{pmatrix} s \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} S_{xx'} + S_{yy'} \\ S_{x'} \\ S_{y'} \end{pmatrix} \quad (\text{ก.12})$$

ซึ่ง

$$\begin{aligned} S_{xx} &= \sum x_i^2 & S_{yy} &= \sum y_i^2 & S_x &= \sum x_i & S_y &= \sum y_i \\ S_{xx'} &= \sum x_i x'_i & S_{yy'} &= \sum y_i y'_i & S_{x'} &= \sum x'_i & S_{y'} &= \sum y'_i \end{aligned} \quad (\text{ก.13})$$

ภาคผนวก ข
ดัชนีคำศัพท์

3D Face Reconstruction	การสร้างภาพใบหน้าคน 3 มิติ
Active Shape Model	การจำลองรูปร่างแบบแอคทีฟ
Affine Transform	การแปลงสัมพรรค
Alignment	การปรับแนว
Analysis	การวิเคราะห์
Baseline	เส้นฐาน
Canonical Camera	แบบบัญญัติกล้อง
Calibration	การเทียบวัด
Central Projection	ศูนย์กลางการฉาย
Collinearity Condition	เงื่อนไขเชิงเส้นร่วม
Computer Vision	การรับรู้ภาพของคอมพิวเตอร์
Coordinate	พิกัด
Correlation Matrix	เมตริกซ์สหสัมพันธ์
Correspondence Point	จุดความสอดคล้อง
Covariance Matrix	เมตริกซ์ความแปรปรวนร่วม
Digital Image Processing	การประมวลผลภาพดิจิทัล
Distortion	การบิดเบือน
Direct Linear Transformation	การแปลงเชิงเส้นโดยตรง
Essential Matrix	เมตริกซ์จำเป็น
Eigen Decomposition	การแยกเวกเตอร์ลักษณะเฉพาะ
Eigenvalue	ค่าลักษณะเฉพาะ
Eigenvector	เวกเตอร์ลักษณะเฉพาะ
Element	องค์ประกอบ
Euclidean	ยูคลิด
Extrinsic Parameter	ค่าพารามิเตอร์ภายนอก
Feature Point	จุดคุณลักษณะเด่น
Fundamental Matrix	เมตริกซ์พื้นฐาน

Geometric	เรขาคณิต
Image Point	จุดภาพ
Image Segmentation	การแบ่งส่วนภาพ
Intrinsic Parameter	ค่าพารามิเตอร์ภายใน
Inverse Transformation	การแปลงกลับ
Landmark	จุดกำหนด
Least Square	กำลังสองน้อยที่สุด
Level Set	เซตระดับ
Linear Combination	การผสมเชิงเส้น
Linear Equation	สมการเชิงเส้น
Mean Shape	รูปร่างเฉลี่ย
Mapping	การเมท
Model	แบบจำลอง
Multiview	หลายมุมมอง
Non Linear	ไม่เป็นเชิงเส้น
Optimization	การหาค่าเหมาะสมที่สุด
Pinhole Camera	กล้องรูเข็ม
Point Distribution Models	แบบจำลองการแจกแจงจุด
Procraster Method	วิธีการโปรครัสเตอร์
Pose	การระบุตำแหน่ง
Principal Component Analysis	การวิเคราะห์องค์ประกอบแกนหลัก
Principal Point	จุดหลัก
Projection	การฉาย
Pseudo Inverse	การแปลงกลับเทียม
Reference Point	จุดอ้างอิง
Rotation	การหมุนวัตถุ
Scaling	การสเกล
Shape	รูปร่าง
Similarity	ความคล้าย
Surface	พื้นผิว

Synthesis

การสังเคราะห์

Training Image

ภาพกลุ่มตัวอย่าง

Transformation

การแปลง

Translation

การเลื่อนขนาน

World Point

จุดจริง



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค
บทความที่ได้รับการตีพิมพ์

Manassanan Srikham, Charnchai Pluempitwiriwaj and Thitiporn Chanwimaluang.
Comparison of Dense Matching Algorithms in Noisy Image. Proceedings of The 2nd
International Conference on Digital Image Processing (ICDIP 2010), Proc. of SPIE Vol.
7546, Singapore, 26-28 February 2010.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวมนัสนันท์ ศรีคำ เกิดวันที่ 23 พฤศจิกายน พ.ศ. 2524 ที่จังหวัดศรีสะเกษ สำเร็จการศึกษาระดับปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้ากำลัง คณะวิศวกรรมศาสตร์ มหาวิทยาลัยหอการค้าไทย ในปีการศึกษา 2548 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต ในสังกัดห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2550



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย