

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

โปรแกรม Calibration สำหรับการสร้างภาพ 3 มิติที่มีความแม่นยำสูง
Versatile Camera Calibration Technique for High-Accuracy 3D Machine
Vision

โดย

นาย ฉัตรชัย เฉลิมโชคเจริญกิจ
นาย วุฒิสันตะบุตร

เลขหมู่.....
เลขทะเบียน..... 73157
วัน,เดือน,ปี..... 6 ก.ค. 2550

b. 1178699
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมบัณฑิต
สาขาวิชาอิเล็กทรอนิกส์
คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2548

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม Calibration สำหรับการสร้างภาพ 3 มิติที่มีความแม่นยำสูง
Versatile Camera Calibration Technique for High-Accuracy 3D Machine
Vision

โดย

นาย ฉัตรชัย เฉลิมโชคเจริญกิจ รหัส 45010148
นาย วุฒ สันตะบุตร รหัส 45010684

อาจารย์ที่ปรึกษา
ดร. ศิริเดช บุญแสง

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมบัณฑิต
สาขาวิชาอิเล็กทรอนิกส์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2548

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายงานปีการศึกษา 2548

ภาควิชา อิเล็กทรอนิกส์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง โปรแกรม Calibration สำหรับการสร้างภาพ 3 มิติที่มีความแม่นยำสูง

ผู้จัดทำ

- | | | |
|----------------|-------------------|---------------|
| 1. นาย จักรชัย | เฉลิม โชคเจริญกิจ | รหัส 45010148 |
| 2. นาย วุฒ | สันตะบุตร | รหัส 45010684 |



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม calibration สำหรับการสร้างภาพ 3 มิติที่มีความแม่นยำสูง

นาย ฉัตรชัย เฉลิมโชคเจริญกิจ รหัส 45010148

นาย วุฒ สันตะบุตร รหัส 45010684

ดร. ศิริเดช บุญแสง อาจารย์ที่ปรึกษา

ภาคเรียนที่ 2 ปีการศึกษาที่ 2548

บทคัดย่อ

ระบบแสง 3 มิติแบบราคาถูก อาศัยเทคนิค camera calibration ที่ช่วยให้สามารถสร้างสร้างภาพ 3 มิติที่ระบุตำแหน่งได้อย่างแม่นยำและมีความละเอียดที่ระดับมิลลิเมตร ระบบประกอบด้วย เลเซอร์ชี้ตำแหน่ง, กล้อง webcam 2 ตัว และโปรแกรมประมวลผล (Labview และ Matlab) การทำงานของระบบ ใช้เลเซอร์กวาดไปบนวัตถุ โดยมีกล้อง webcam 2 ตัวซึ่งอยู่คนละตำแหน่งบันทึกภาพวิดีโอ (ไฟล์ .avi) พร้อมกัน ไฟล์ .avi ที่ได้จะถูกนำไปป้อนใส่โปรแกรมคอมพิวเตอร์ซึ่งจะประมวลผลออกมาได้เป็นจุดภาพ 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Camera Calibration Technique for Machine Vision

Mr. Chatchai Chalermchokejaroenkij

ID.45010148

Mr. Warit Santaputra ID.45010684

Dr. Siridech Boonsang advisor

2nd Semester, 2005

Abstract

This 3D scanning system, with the use of camera calibration technique, can create a highly accurate 3D image within millimeter resolution. The system consists of a laser pointer, 2 webcams and computer programs (Labview and Matlab). The scanning process can be done by sweeping laser on the target while 2 webcams record the videos (.avi file) synchronously. The recorded movie files then will be processed by a computer program in order to create 3D dots of image.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ขอขอบคุณ

อาจารย์ ศิริเดช บุญแสง (อาจารย์ที่ปรึกษา) และอาจารย์ภาคิเล็กทรอนิกส์ทุกท่านที่ให้การอุปการะในการให้คำปรึกษา และแนะนำเกี่ยวกับปฏิญานิพนธ์เรื่องนี้ และทางภาคิเล็กทรอนิกส์ที่ให้อิมเครื่องมือในการทดลอง ตลอดจนสั่งสอนให้ความรู้จนสามารถนำมาประยุกต์ใช้งานในการทำปฏิญานิพนธ์ครั้งนี้

คุณพ่อ คุณแม่ ที่คอยให้ความเป็นห่วง และให้การสนับสนุนทางการเงินตลอดมา จนทำให้ปฏิญานิพนธ์นี้สำเร็จ โดยสมบูรณ์ได้ด้วยดี

คุณค่าและประโยชน์อันพึงมีในปฏิญานิพนธ์นี้ ผู้จัดทำขอมอบแด่ผู้มีพระคุณทุกท่าน

.....
(นาย ฉัตรชัย เฉลิมโชคเจริญกิจ)

.....
(นาย วุฒ สันตะบุตร)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	ก
กิตติกรรมประกาศ	ค
บทที่1 บทนำ	1
บทที่2 ทฤษฎีและหลักการเบื้องต้น	3
2.1 Calibration	3
2.2 Digital Image Processing	13
2.3 Stereo Vision: Triangulation	17
บทที่3 ขั้นตอนการทดลอง	24
3.1 Camera Calibration	24
3.2 ขั้นตอนการสร้างภาพ 3 มิติ	29
3.3 ขั้นตอนการทดลองระบบเพื่อหาความแม่นยำของระบบ	30
บทที่4 ผลการทดลอง	31
4.1 Calibration	31
4.2 ผลการทดลองการสร้างภาพ 3 มิติ	33
บทที่5 สรุปผลการทดลอง	38
ภาคผนวก	39
Front Panel ของ โปรแกรม	40
Block Diagram ของ โปรแกรม	41
บรรณานุกรม	46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

	หน้า
รูปที่ 1.1 บล็อกไดอะแกรมของระบบ	1
รูปที่ 2.1 การ project ใน 3 มิติและการ distort ของเลนส์กล้อง	3
รูปที่ 2.2 ขั้นตอนทั้ง 4 ในการแปลงจากพิกัด 3 มิติโลก ไปเป็นพิกัดภาพในคอมพิวเตอร์	4
รูปที่ 2.3 การใช้ระนาบรูปตารางหมากรุกในการคาลิเบรทกล้อง	8
รูปที่ 2.4 แสดงการหา pattern matching รูป a. เป็น pattern image รูป b.-d. แสดงตำแหน่งที่ match ได้	14
รูปที่ 2.5 การหา correlation	16
รูปที่ 2.6 ตำแหน่งของกล้องและพิกัดต่างๆ	18
รูปที่ 2.7 พิกัดของภาพกับ frame	19
รูปที่ 2.8 รูปมาตรฐานของการทำ stereo triangulation	21
รูปที่ 2.9 stereo triangulation เมื่อกล้องสองตัวไม่ขนานกัน	22
รูปที่ 3.1 แสดงภาพระนาบรูปตารางหมากรุกที่ใช้ในการทดลอง	24
รูปที่ 3.2 แสดงภาพการถ่ายภาพระนาบรูปตารางหมากรุก โดยกล้อง web cam ทั้ง 2 ตัวซึ่งตั้งอยู่กับขาตั้งกล้อง	25
รูปที่ 3.3 แสดงภาพ plate ที่ใช้กำหนดมุมในการหั่นของระนาบรูปตารางหมากรุก	25
รูปที่ 3.4 ระนาบรูปตารางหมากรุกทิศทางที่ 4 จากกล้องด้านขวา	26
รูปที่ 3.5 ภาพถ่ายของระนาบรูปตารางหมากรุกจากกล้องด้านซ้ายทั้ง 10 ทิศทาง	26
รูปที่ 3.6 ภาพถ่ายของระนาบรูปตารางหมากรุกจากกล้องด้านขวาทั้ง 10 ทิศทาง	27
รูปที่ 3.7 แสดงการหามุมจากภาพของกล้องด้านขวาทิศทางที่ 4	27
รูปที่ 3.8 แสดงมุมที่ถูกรวม extract ออกมาได้	28
รูปที่ 3.9 อุปกรณ์และการทดลอง pattern matching	29
รูปที่ 4.1 แสดงภาพค่า pixel error ของจุดที่ใช้ calibrate ทุกจุด (error ไม่เกิน 0.8 pixel)	32

สารบัญภาพ

	หน้า
รูปที่ 4.2 แสดงภาพ extrinsic ของกล้องคั่นขวากับระนาบรูป ตารางหมากรุกในทิศทางต่างๆ	33
รูปที่ 4.3 กราฟแสดงความสัมพันธ์ระหว่างระยะวัตถุต่อ % error	34
รูปที่ 4.4 ภาพของวัตถุที่ใช้สแกน	35
รูปที่ 4.5 ภาพที่สแกนได้ของรูปที่ 4.3 หน่วยเป็น mm.	36
รูปที่ 4.6 ภาพของวัตถุที่ใช้สแกน	36
รูปที่ 4.7 ภาพที่สแกนได้ของรูปที่ 4.5 หน่วยเป็น mm.	37



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

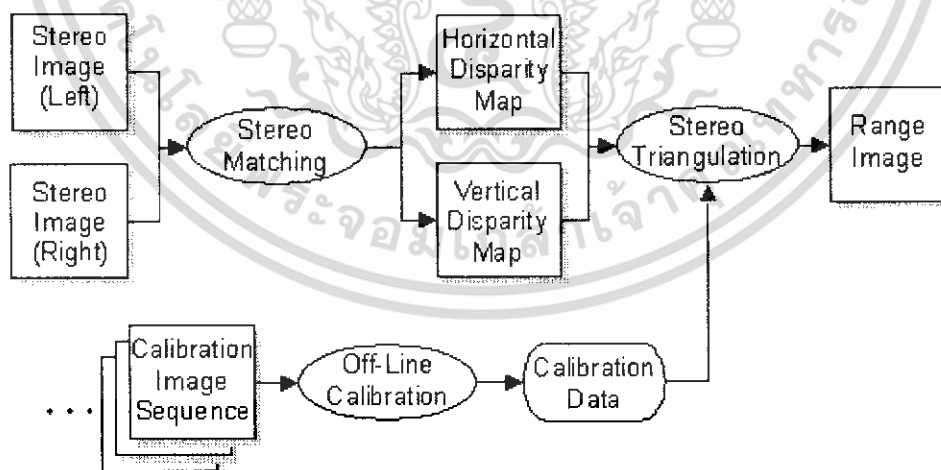
บทที่ 1

บทนำ

ปัจจุบันมีการใช้ประโยชน์จากเครื่องสแกน 3 มิติ อย่างกว้างขวาง เช่น การตรวจสอบความถูกต้อง, สัตยกรรมพลาสติก, การผลิตเครื่องแต่งกาย และอื่นๆอีกมากมาย ระบบสแกนนี้จะสร้างพื้นผิว 3 มิติ ในลักษณะของจุดในตำแหน่งพิกัด 3 มิติ

แม้ว่าเทคโนโลยีการใช้อัลตราซาวด์หรือเทคโนโลยีทางเครื่องกลต่างๆจะทำให้ได้รับข้อมูลได้เช่นกันแต่การสแกนทางแสงจะมีข้อดีตรงที่มันไม่ต้องสัมผัสกับผิววัตถุและให้ความละเอียดถูกต้องสูง ดังนั้นเทคโนโลยีทางแสงจึงถูกใช้อย่างกว้างขวาง ระบบสแกน 3 มิติที่มีความถูกต้องแม่นยำและยืดหยุ่นสูง สามารถเคลื่อนย้ายได้อย่างสะดวกง่ายดายและไม่จำกัดขนาดของวัตถุที่ถูกสแกน อีกทั้งอุปกรณ์ที่ใช้มีราคาถูกและสามารถหาซื้อได้ทั่วไปตามท้องตลาด

ลักษณะของระบบสแกน 3 มิติประกอบด้วย Laser pointer, กล้องวิดีโอ 1 คู่ และระบบประมวลผลภาพในคอมพิวเตอร์ การทำงานของระบบจะทำงานโดยใช้ Laser pointer สแกนไปตามพื้นผิวของวัตถุ (พิกัด 2 มิติ ที่จุดแสงซึ่งคำนวณที่พิกเซลที่ศูนย์กลางของจุดแสง) แล้วใช้กล้องวิดีโอถ่าย ให้ระบบดำเนินการแสดงตำแหน่ง 3 มิติของวัตถุที่ถูกสแกนด้วยวิธี stereo triangulation ระบบจะรู้จักจุดแสงของ Laser pointer ได้จากลักษณะรูปร่างโดยใช้ pattern matching กล้องวิดีโอทั้งสองตัวที่ใช้ในระบบสแกนนี้จะถูกนำไปปรับเทียบก่อนเพื่อหาค่า intrinsic และ extrinsic parameter ซึ่งจะทำให้ค่าที่สแกนได้มีความถูกต้องแม่นยำ



รูปที่ 1.1 บล็อกไดอะแกรมของระบบ

จากรูปที่ 1.1 จะเห็นว่าระบบประกอบไปด้วย 3 ส่วนหลัก คือ

1. Calibration คือการหาค่า parameter ต่างๆของกล้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Pattern matching คือการหาภาพต้นแบบซึ่งก็คือภาพจุดแสงเลเซอร์บนวัตถุที่ถูกสแกน
3. Stereo triangulation คือการคำนวณหาตำแหน่งใน 3 มิติโดยใช้ข้อมูลที่ได้จาก 2 ชั้น
ตอนแรก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

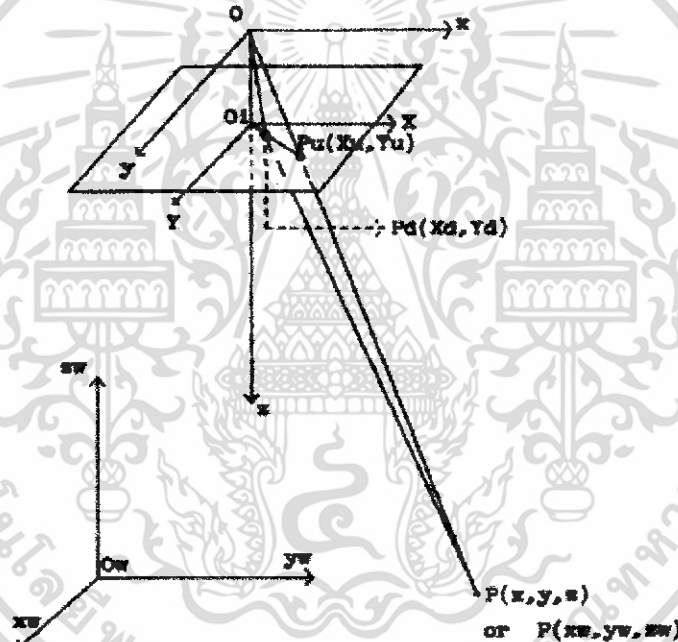
บทที่ 2

ทฤษฎีและหลักการเบื้องต้น

2.1 Calibration

การ calibrate กล้องนี้จะทำใน 3 มิติเพื่อหาค่าปัจจัยต่างๆของกล้องประกอบด้วยปัจจัยภายในของกล้องเช่น focal length (ระยะ โฟกัส) ,Lens distortion (การบิดเบือนของเลนส์) เป็นต้น เรียกว่าค่า intrinsic parameter และปัจจัยภายนอกของกล้องนั้นคือตำแหน่งใน 3 มิติของกล้องและความสัมพันธ์ระหว่างพิกัดของกล้องและพิกัด โลกเรียกว่าค่า extrinsic parameter

หลักในการ calibrate คือ การแปลงพิกัด 3 มิติของ โลก(3D world coordinate) เป็นพิกัดของกล้อง มีทั้งหมด 4 ขั้นตอน

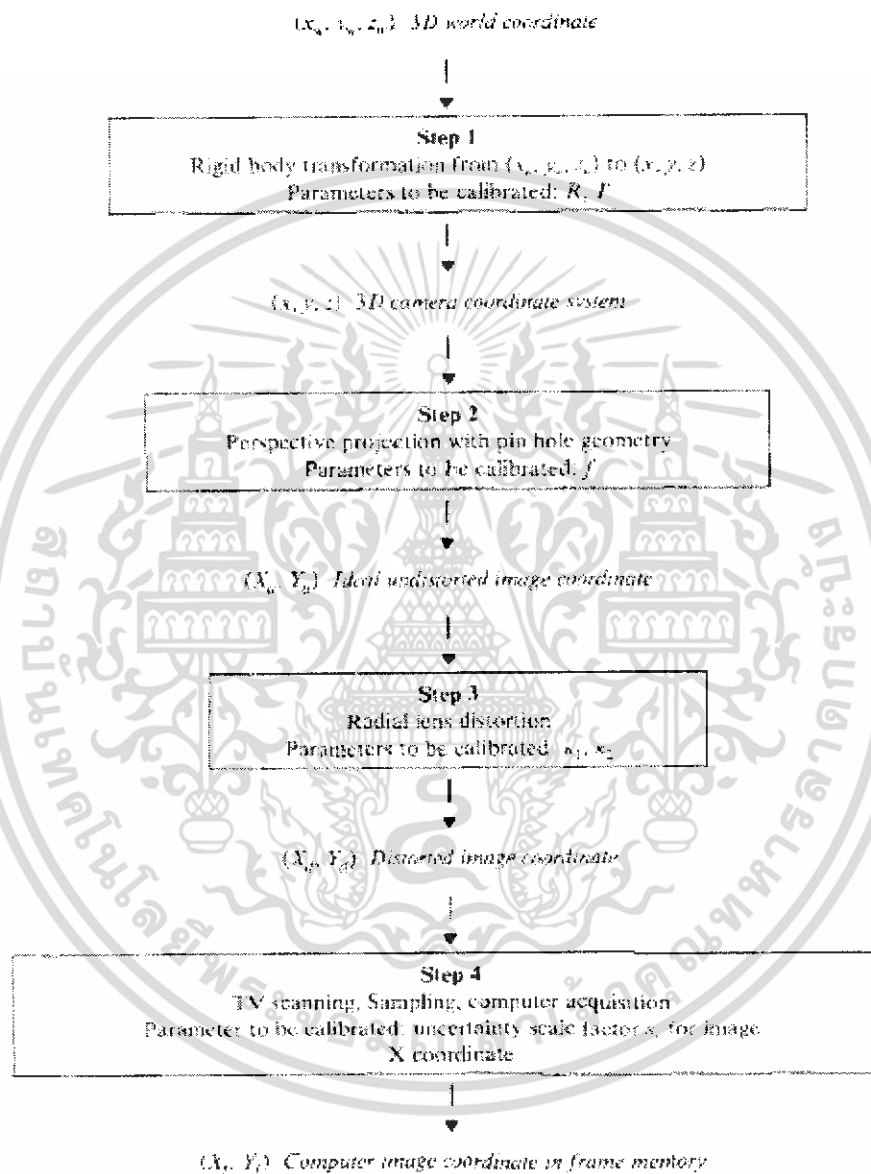


รูปที่ 2.1 การ project ใน 3 มิติและการ distort ของเลนส์กล้อง

จากรูปที่ 2.1 แสดงโมเดลของกล้อง (x_w, y_w, z_w) เป็นพิกัด 3 มิติ ของวัตถุ P ในระบบพิกัด 3 มิติโลก (x, y, z) เป็นพิกัด 3 มิติของวัตถุ P ในระบบพิกัด 3 มิติของกล้องซึ่งจุดศูนย์กลางอยู่ที่ O (X, Y) เป็นระบบพิกัดของภาพจุดศูนย์กลางอยู่ที่ O_1 และขนานกับแกน x และ y f คือระยะห่างระหว่าง image plane กับ O (X_c, Y_c) คือพิกัดภาพ(ในแกน x, y, z) (X_c, Y_c) คือพิกัดภาพที่แท้จริง ซึ่งต่างจาก (X_c, Y_c) เนื่องจากที่ (X_c, Y_c) มีการ distort ของเลนส์อยู่ สำหรับ (X_c, Y_c) คือพิกัดที่ใช้ในคอมพิวเตอร์ซึ่งเป็นจำนวนของพิกเซลในแกน x และ y ของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าพารามิเตอร์ต่างๆจะต้องถูก calibrate จากพิกัดของภาพใน image plane ไปสู่พิกัดของคอมพิวเตอร์ การแปลงจะต้องแปลงจากพิกัด (x_w, y_w, z_w) ไปเป็น (X_p, Y_p)



รูปที่ 2.2 ขั้นตอนทั้ง 4 ในการแปลงจากพิกัด 3 มิติโลกไปเป็นพิกัดภาพในคอมพิวเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 1 Rigid body transformation

เป็นการแปลงระบบพิกัด 3 มิติโลก พิกัด (x_w, y_w, z_w) ไปเป็นระบบพิกัด 3 มิติของกล้อง โดย
ใช้ค่า extrinsic parameter คือค่า R และ T ตามความสัมพันธ์

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T \quad (2.1)$$

ค่า R เป็นค่า 3x3 เมทริกซ์ของการหมุน(rotation) รอบจุด origin

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2.2)$$

และ T เป็นเวกเตอร์ของการเลื่อนตำแหน่ง (Translation)

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

ค่าพารามิเตอร์ที่ใช้ในการ Calibrate คือค่า R และ T

ขั้นตอนที่ 2 เป็นการแปลงจากพิกัด 3 มิติของกล้อง (x, y, z) ไปเป็นพิกัดภาพสมมติฐาน (ไม่มีการ
distort) (X_u, Y_u) โดยการใช้การ projection ใน 3 มิติของกล้องแบบรูเข็ม (pinhole camera) ตามสมการ

$$X_u = f(x/z) \quad (2.4)$$

$$Y_u = f(y/z) \quad (2.5)$$

ค่าพารามิเตอร์ที่ใช้ calibrate คือค่าระยะโฟกัส f

ขั้นตอนที่ 3 ค่าการ distort ของเลนส์คือ

$$X_d + D_x = X_u \quad (2.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Y_d + D_y = Y_u \quad (2.7)$$

ซึ่ง (X_d, Y_d) คือพิกัดภาพที่แท้จริง (มีการ distort) บน image plane

$$D_x = X_d(k_1 r^2 + k_2 r^4 + \dots) \quad (2.8)$$

$$D_y = Y_d(k_1 r^2 + k_2 r^4 + \dots) \quad (2.9)$$

$$r = (X_d^2 + Y_d^2)^{1/2} \quad (2.10)$$

ค่าพารามิเตอร์ที่ใช้ calibrate คือค่าสัมประสิทธิ์การบิดเบือน (Distortion coefficient) k_i การบิดเบือนมี 2 ชนิดคือ ทางรัศมีและทางเส้นสัมผัส แต่ในการใช้งานจริงนั้นจะพิจารณาเฉพาะการบิดเบือนทางรัศมีเท่านั้น เนื่องจากหากพิจารณาทั้ง 2 ตัวจะทำให้สมการยุ่งยากและไม่เสถียร ดังนั้นเราจึงพิจารณาเฉพาะการบิดเบือนทางรัศมีเท่านั้น

ขั้นตอนที่ 4 เป็นการแปลงจากพิกัดภาพที่แท้จริง (X_d, Y_d) ไปยังพิกัดภาพของคอมพิวเตอร์ (X_r, Y_r) โดยสมการ

$$X_r = s_x d'_x{}^{-1} X_d + C_x \quad (2.11)$$

$$Y_r = d'_y{}^{-1} Y_d + C_y \quad (2.12)$$

$$d'_x = d_x (N_x / N_x) \quad (2.13)$$

โดยที่ (X_r, Y_r) คือจำนวน row และ column ของจุดพิกเซลของภาพที่สนใจในคอมพิวเตอร์

(C_x, C_y) คือจำนวน row และ column ของจุดศูนย์กลางของภาพในคอมพิวเตอร์

d_x คือ ระยะระหว่างจุดศูนย์กลางของจุดที่ใกล้กันในแกน x

d_y คือ ระยะระหว่างจุดศูนย์กลางของจุดที่ใกล้กันในแกน y

N_x คือจำนวนจุดเซนเซอร์ในแกน x

N_x คือจำนวนพิกเซลในเส้นที่แสดง โดยคอมพิวเตอร์

ค่าพารามิเตอร์ที่ใช้คาไลเบรท คือค่าแฟกเตอร์ความคลาดเคลื่อนของภาพ s_x ในการแปลงระหว่างพิกัดภาพจากคอมพิวเตอร์กับพิกัดภาพจริงนั้นต้องใช้ในการเทียบระยะห่างระหว่าง 2 พิกเซลที่ใกล้กันทั้งทาง row และ column ภายในกรอบภาพทางคอมพิวเตอร์กับพิกัดภาพจริง

สมการที่เกี่ยวข้องในการแปลงพิกัด 3 มิติของโลกไปเป็นพิกัดภาพ 2 มิติในคอมพิวเตอร์ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยการรวม 3 ชั้นตอนสุดท้ายจะได้พิกัด(X,Y) ของคอมพิวเตอร์เกี่ยวข้องกับพิกัด(x,y,z) ของจุดวัตถุในระบบพิกัดกล้อง 3 มิติโดยสมการต่อไปนี้

$$s_x^{-1}d'_x X + s_x^{-1}d'_x X_{k_1} r^2 = f \frac{x}{z} \quad (2.14)$$

$$d'_y Y + d_y Y_{k_1} r^2 = f \frac{y}{z} \quad (2.15)$$

เมื่อ
$$r = \sqrt{(s_x^{-1}d'_x X)^2 + (d_y Y)^2}$$

แทนสมการ 2.1 ลงในสมการ 2.14 และ 2.15 จะได้

$$s_x^{-1}d'_x X + s_x^{-1}d'_x X_{k_1} r^2 = f \frac{r_1 X_w + r_2 Y_w + r_3 Z_w + T_x}{r_7 X_w + r_8 Y_w + r_9 Z_w + T_z} \quad (2.16)$$

และ

$$d'_y Y + d_y Y_{k_1} r^2 = f \frac{r_4 X_w + r_5 Y_w + r_6 Z_w + T_y}{r_7 X_w + r_8 Y_w + r_9 Z_w + T_z} \quad (2.17)$$

ค่าพารามิเตอร์ที่ใช้ในการแปลงพิกัดในรูปที่ 2.2 สามารถแบ่งได้เป็น 2 ชนิดคือ

2.1.1 extrinsic parameter

คือค่าพารามิเตอร์ในขั้นตอนที่ 1 สำหรับแปลงจากระบบพิกัด 3 มิติของโลกไปเป็นระบบพิกัด 3 มิติของกล้องที่มีจุดศูนย์กลางที่ optical center (O) ค่า extrinsic parameter มี 6 ตัวคือ ค่ามุมหักเห θ , ค่าความลาด ϕ และค่าความลาดเอียง ψ สำหรับการหมุนและอีก 3 ตัวแปรสำหรับการเลื่อนคือ T_x, T_y, T_z เมทริกซ์ของการหมุน R สามารถแสดงเป็นฟังก์ชันของ θ, ϕ และ ψ ได้ดังนี้

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

2.1.2 intrinsic parameter

คือค่าพารามิเตอร์ในขั้นตอนที่ 2 – 4 สำหรับการแปลงพิกัดวัตถุ 3 มิติไปเป็นพิกัดกล้องแล้วแปลงไปเป็นพิกัดภาพในคอมพิวเตอร์ มีทั้งหมด 6 ตัวคือ

f คือค่าระยะ โฟกัสหรือระยะห่างระหว่าง image plane กับจุดศูนย์กลาง

k_1 คือค่าสัมประสิทธิ์การบิดเบือนของเลนส์

s_x คือค่า factor ความคลาดเคลื่อนสำหรับ x เนื่องจาก error ของเวลาในกล้องทีวี

(C_x, C_y) คือพิกัดภาพในคอมพิวเตอร์



รูปที่ 2.3 การใช้ระนาบรูปตารางหมากรุกในการคาไลเบรทกล้อง

2.1.3 ขั้นตอนการคำนวณ

A) คำนวณพิกัดที่ถูก distort (X_d, Y_d)

i) ถ่ายภาพระนาบหมากรุกเข้าในคอมพิวเตอร์แล้วทำการตรวจหาจำนวน row และ column ที่แต่ละจุดที่ทำการคาไลเบรท เรียกว่า (X_n, Y_n)

ii) ได้ค่า N_x, N_y, d_x, d_y จากสมการ (2.13) โดยใช้ข้อมูลของกล้องและขนาดเฟรมจากผู้ผลิต

iii) ให้ (C_x, C_y) เป็นพิกเซลที่จุดศูนย์กลางของภาพในเฟรม

iv) คำนวณ (X_{di}, Y_{di}) โดยใช้สมการ (2.11) และ (2.12) ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$X_{di} = s_x^{-1} d'_x (X_{fi} - C_x) \tag{2.19}$$

$$Y_{di} = d_y (Y_{fi} - C_y) \tag{2.20}$$

สำหรับ i เท่ากับ $1, 2, \dots, N$ และ N คือจำนวนทั้งหมดของจุดที่ทำการคาถิเบรท

B) คำนวณตัวแปร $T_y^{-1} r_1, T_y^{-1} r_2, T_y^{-1} T_x, T_y^{-1} r_4, T_y^{-1} r_5$

สำหรับแต่ละจุด $i (x_{wi}, y_{wi}, z_{wi})$ และค่า (X_{di}, Y_{di}) จากการคำนวณในข้อ a) สามารถตั้งสมการเชิงเส้นได้ดังนี้

$$[Y_{di} X_{di} \quad Y_{di} X_{di} \quad Y_{di} - X_{di} X_{di} \quad X_{di} Y_{di}] \begin{bmatrix} T_y^{-1} r_1 \\ T_y^{-1} r_2 \\ T_y^{-1} T_x \\ T_y^{-1} r_4 \\ T_y^{-1} r_5 \end{bmatrix} = X_{di} \tag{2.21}$$

เมื่อ N (จำนวนจุดวัตถุ) ต้องมากกว่า 5 จุด จึงจะสามารถตั้งสมการเชิงเส้นนี้ได้

C) คำนวณ $(r_1, \dots, r_5, T_x, T_y)$ จาก $(T_y^{-1} r_1, T_y^{-1} r_2, T_y^{-1} T_x, T_y^{-1} r_4, T_y^{-1} r_5)$

ให้ C เป็น เมทริกซ์ย่อย 2×2 ของเมทริกซ์การหมุน R

$$C = \begin{bmatrix} r'_1 & r'_2 \\ r'_4 & r'_5 \end{bmatrix} = \begin{bmatrix} r_1/T_y & r_2/T_y \\ r_4/T_y & r_5/T_y \end{bmatrix} \tag{2.22}$$

$$r'_2 = \frac{S_r - |S_r^2 - 4(r'_1 r'_5 - r'_4 r'_2)|^{1/2}}{2(r'_1 r'_5 - r'_4 r'_2)} \tag{2.23}$$

เมื่อ $S_r = r_1'^2 + r_2'^2 + r_4'^2 + r_5'^2$

พิจารณาเครื่องหมายของ T_y

- i) ให้จุดวัตถุ i ในพิกัดภาพในคอม (X_{fi}, Y_{fi}) ห่างจากศูนย์กลางภาพ (C_x, C_y) พิกัดโลกของวัตถุจะเป็น (x_{wi}, y_{wi}, z_{wi})
- ii) ให้เครื่องหมายของ T_y เป็น +1
- iii) คำนวณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
 r_1 &= (T_y^{-1} r_1) \cdot T_y & r_2 &= (T_y^{-1} r_2) \cdot T_y \\
 r_3 &= (T_y^{-1} r_3) \cdot T_y \\
 r_5 &= (T_y^{-1} r_5) \cdot T_y & T_1 &= (T_y^{-1} T_1) \cdot T_y \\
 x &= r_1 x_w + r_2 y_w + T_x & y &= r_4 x_w + r_5 y_w + T_2
 \end{aligned}$$

iv) ถ้า x, X เครื่องหมายเหมือนกันและ y, Y มีเครื่องหมายเหมือนกันแล้ว T_y จะมีเครื่องหมายเป็น +1 ไม่เช่นนั้น T_y จะมีเครื่องหมายเป็น -1

คำนวณเมทริกซ์ 3 มิติของการหมุน R หรือ r_1, r_2, \dots, r_9

i) คำนวณ

$$\begin{aligned}
 r_1 &= (T_y^{-1} r_1) \cdot T_y & r_2 &= (T_y^{-1} r_2) \cdot T_y \\
 r_4 &= (T_y^{-1} r_4) \cdot T_y \\
 r_5 &= (T_y^{-1} r_5) \cdot T_y & T_1 &= (T_y^{-1} T_1) \cdot T_y
 \end{aligned}$$

ii) คำนวณ R ด้วยสมการต่อไปนี้

$$R = \begin{bmatrix} r_1 & r_2 & (1-r_1^2-r_2^2)^{1/2} \\ r_4 & r_5 & s(1-r_4^2-r_5^2)^{1/2} \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (2.24)$$

เมื่อ $s = -(r_1 r_4 + r_2 r_5)$ ส่วน r_7, r_8, r_9 หาได้จากผลของ 2 row แรกโดยใช้คุณสมบัติ Orthonormal และคุณสมบัติมือขวาของ R

iii) คำนวณระยะโฟกัส f ถ้า น้อย f กว่าศูนย์จะได้สมการ 2.25

$$R = \begin{bmatrix} r_1 & r_2 & -(1-r_1^2-r_2^2)^{1/2} \\ r_4 & r_5 & -s(1-r_4^2-r_5^2)^{1/2} \\ -r_7 & -r_8 & r_9 \end{bmatrix} \quad (2.25)$$

สมการ (2.24) และ (2.25) จะต้องถูกเลือกใช้อันใดอันหนึ่งเท่านั้น วิธีที่จะใช้เลือกสมการที่เหมาะสมนั้นคือการใช้สมการเชิงเส้นสมการ (2.26) ด้านล่างสำหรับคำนวณค่าประมาณของ f และ T_z โดยไม่สนใจค่าการบิดเบือนของเลนส์ เนื่องจากค่าประมาณของ f ไม่จำเป็นต้องสนใจ คำนวณเอาเฉพาะเครื่องหมายเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

D) คำนวณค่าประมาณของ f และ T_z โดยไม่สนใจการบิดเบือนของเลนส์

$$[y_i \quad -d_i Y_i] \begin{bmatrix} f \\ T_z \end{bmatrix} = w_i d_i Y_i \quad (2.26)$$

เมื่อ

$$y_i = r_4 x_{wi} + r_5 y_{wi} + T_y$$

$$w_i = r_7 x_{wi} + r_8 y_{wi}$$

E) คำนวณค่าที่แท้จริงของ f, T_z และ k_1

แก้สมการ (2.17) โดยให้ f, T_z และ k_1 เป็นค่าตัวแปร คำนวณโดยใช้ค่า f, T_z จากสมการที่ (2.26) เป็นค่าเริ่มต้นและใช้ค่า $k_1 = 0$ เป็นค่าเริ่มต้น

2.1.4 Multiple Viewing position calibration

เมื่อมุมมองของภาพมีมากกว่า 1 มุมมองโดยที่แต่ละมุมมองมีความแตกต่างทั้งในเรื่องตำแหน่งและความสัมพันธ์กับจุดที่ calibrate โดยการใช้กล้องตัวเดียว ค่า extrinsic parameter ของแต่ละมุมมองย่อมแตกต่างกันแต่ค่า intrinsic parameter เหมือนกันเนื่องจากใช้กล้องตัวเดียวกัน

การวัดความแม่นยำของการคาลิเบรทกล้องเราจะวัดในระบบพิกัด 3 มิติ ผ่านทางวิธี stereo triangulation โดยใช้ค่าพารามิเตอร์จากการคาลิเบรทกล้อง

1. หากคาลิเบรทกล้องแบบมุมมองเดียวจะต้องต่อกล้องอีกตัวอย่างแน่นหนากับตัวแรกแล้วทำการคาลิเบรทกล้องทั้ง 2 ตัว (จุดประสงค์ของกล้องตัวที่ 2 คือเพื่อให้สามารถใช้ stereo triangulation)
2. ได้พิกัดภาพ 2 มิติ สำหรับจุดที่ทดสอบของพิกัด 3 มิติ (ที่รู้ความสัมพันธ์กับพิกัด 3 มิติของโลก)
3. คำนวณพิกัด 3 มิติของจุดที่ทดสอบนั้นในระบบพิกัดโลกโดยใช้ stereo triangulation ถ้าใช้การคาลิเบรทแบบหลายมุมมองเพียง 2 มุมมองเท่านั้นก็เพียงพอสำหรับ stereo triangulation แล้ว ถ้าใช้การคาลิเบรทแบบมุมมองเดียวต้องใช้กล้อง 2 ตัวต่อเชื่อมเข้าด้วยกันอย่างแน่นหนาในการคาลิเบรทและจึงใช้ stereo triangulation ในการหาพิกัด
4. ความแม่นยำในการคาลิเบรทประเมินโดยการเปรียบเทียบความแตกต่างระหว่างพิกัด 3 มิติที่เรารู้ค่ากับพิกัดที่คำนวณได้จากข้อ 3 การเปรียบเทียบอาจเปรียบเทียบในพิกัด 3 มิติของโลกหรือในพิกัด 3 มิติของกล้องก็ได้ แต่เราจะใช้การเปรียบเทียบพิกัด 3 มิติของกล้องเพราะในระบบพิกัดนี้จะเปรียบเทียบได้ง่ายกว่า เช่น พิกัด z คือความลึก แกน x และ y จะขนานกับแกน XY ของ image plane

2.1.5 ทฤษฎี Upper Bound of error สำหรับการวัดใน 3 มิติ

ค่า error ของการวัดใน 3 มิติของพิกัด x,y,z โดยการใช้ stereo triangulation คือ

$$\text{error}_{\text{total}} \approx \left[\left(1 + \frac{1}{\sqrt{6N_0N_f}} \right) \frac{z}{f} + \frac{\|T_x\|}{L\sqrt{6N_0}} \left(1 + \frac{1}{N_f} \right) + \frac{1}{2\sqrt{6N_0}} + \frac{1}{2\sqrt{6N_0}f} \frac{z}{f} \right] \frac{z}{|T_x|} \cdot \delta + \Delta q \quad (2.27)$$

เมื่อ δ คือค่า error ของตำแหน่งภาพโดยประมาณ

N_0 คือจำนวนจุดทั้งหมดที่ใช้ในการคาลิเบรท

$\|T_x\|$ คือระยะห่างระหว่างจุด optical center ของทั้ง 2 สมองมอง

N_f คือจำนวนของมุมมองที่ใช้ในการคาลิเบรท

L คือขนาดของพื้นที่ใช้งานใน image plane ที่ถูกสแกนโดยกล้อง

Δq คือจุดที่ไม่ชัดเจนใน 3 มิติ

2.1.6 ผลของจำนวนจุดที่ใช้ในการคาลิเบรทต่อความแม่นยำในการคาลิเบรท

จากสมการที่ 22 สังเกตว่าสามารถแบ่งได้เป็น 2 กลุ่มคือกลุ่มของพจน์ที่มี N_0 และ N_f เป็นตัวหารเรียกว่า $\text{error}_{\text{calib}}$ กับกลุ่มที่ไม่โดน N_0 หรือ N_f หาร (ไม่ได้รับอิทธิพลของ N_0, N_f) เรียกว่า

$$\text{error}_{\text{noncalib}} = \left[\frac{1}{\sqrt{6N_0N_f}} \frac{z}{f} + \frac{\|T_x\|}{L\sqrt{6N_0}} \left(1 + \frac{1}{N_f} \right) + \frac{1}{2\sqrt{6N_0}} + \frac{1}{2\sqrt{6N_0}f} \frac{z}{f} \right] \frac{z}{|T_x|} \cdot \delta \quad (2.28)$$

$$\text{error}_{\text{noncalib}} = \frac{z}{f} \frac{z}{|T_x|} \delta + \Delta q \quad (2.29)$$

จะเห็นว่าค่า $\text{error}_{\text{calib}}$ จะขึ้นอยู่กับ N_0 และ N_f ถ้า N_0 มีค่ามากๆ จะทำให้ $\text{error}_{\text{calib}} \ll \text{error}_{\text{noncalib}}$ และ

$\text{error}_{\text{total}} \approx \text{error}_{\text{noncalib}}$ เพราะฉะนั้น ยิ่งใช้จำนวนจุดในการคาลิเบรทมากจะทำให้ได้ค่า $\text{error}_{\text{total}}$ น้อย

ตัวบ่งบอกว่าจำนวนจุดที่ใช้ในการคาลิเบรทมากเพียงพอหรือไม่ นั่นดูได้จากอัตราส่วน

$$\text{error}_{\text{calib}} / \text{error}_{\text{noncalib}}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 Digital Image Processing

ในการประมวลผลภาพทางดิจิทัลนั้นเป็นการนำข้อมูลรูปภาพเชิงดิจิทัลไปประมวลผล เพื่อให้ได้ข้อมูลที่ต้องการที่สามารถนำมาใช้งานหรือวิเคราะห์ได้ ซึ่งในการประมวลผลนั้นจะมีเทคนิคที่ใช้หลากหลายและแตกต่างกันออกไป

2.2.1 Object Recognition

เป็นเทคนิคอย่างหนึ่งในการประมวลผลภาพทางดิจิทัลเพื่อให้ระบบได้ทำความรู้จักกับ object หรือ pattern ที่เราต้องการนำไปใช้ ในการพิจารณา object recognition หรือ pattern recognition นั้น เราพิจารณาจาก 2 ประเด็นหลักๆคือ

1. ปริมาณ (decision-theoretic) เช่น ความยาว, พื้นที่ และความละเอียด
2. ลักษณะโครงสร้าง (structure) เช่น ความสัมพันธ์ของโครงสร้าง, ขอบเขต และช่วงที่พิจารณา

ในที่นี้จะขอพูดถึงเฉพาะ decision-theoretic recognition ซึ่งเป็นวิธีที่จะนำมาใช้ใน pattern matching ต่อไป

2.2.2 Pattern and Pattern Classes

Pattern ในที่นี้หมายถึงรูปแบบของภาพซึ่งถูกนำมาเรียบเรียง ซึ่งเราจะนำมาพิจารณาในหัวข้อต่อไป ระดับของ pattern เราสามารถแบ่งออกได้เป็น class เรากำหนดให้ pattern class เป็น $\omega_1, \omega_2, \dots, \omega_W$ เมื่อ W คือจำนวนของ class Pattern recognition โดย machine อาศัยเทคนิคการกำหนดลำดับ pattern ในแต่ละ class แบบอัตโนมัติ

การเรียบเรียงข้อมูลของ pattern ประกอบไปด้วยข้อมูลเชิงเวกเตอร์ (ใช้ในการวิเคราะห์เชิงปริมาณ), strings และ trees (ใช้ในการวิเคราะห์เชิงลักษณะโครงสร้าง) กำหนดให้ pattern vector เป็น x โดยจะมีรูปแบบดังสมการที่ (2.30)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ M \\ \vdots \\ x_n \end{bmatrix} \quad (2.30)$$

เมื่อ x บอกลักษณะของลำดับข้อมูล, n คือจำนวนของลักษณะข้อมูลของ pattern โดย pattern vector แสดงข้อมูลเป็น column (ในที่นี้เป็น $n \times 1$ metric)

2.2.3 Recognition based on decision theoretic methods

Recognition ด้วยการพิจารณาคุณสมบัติเชิงปริมาณนั้นอาศัยการวิเคราะห์ข้อมูลเป็นฟังก์ชัน กำหนดให้ $x=(x_1, x_2, \dots, x_n)^T$ แสดงเวกเตอร์ของ pattern ขนาด n มิติ สำหรับ W pattern class $\omega_1, \omega_2, \dots, \omega_W$ ให้ $d_1(x), d_2(x), \dots, d_W(x)$ เป็นฟังก์ชันปริมาณของ W เมื่อ pattern x เป็นส่วนหนึ่งของ class ω_1 จะได้

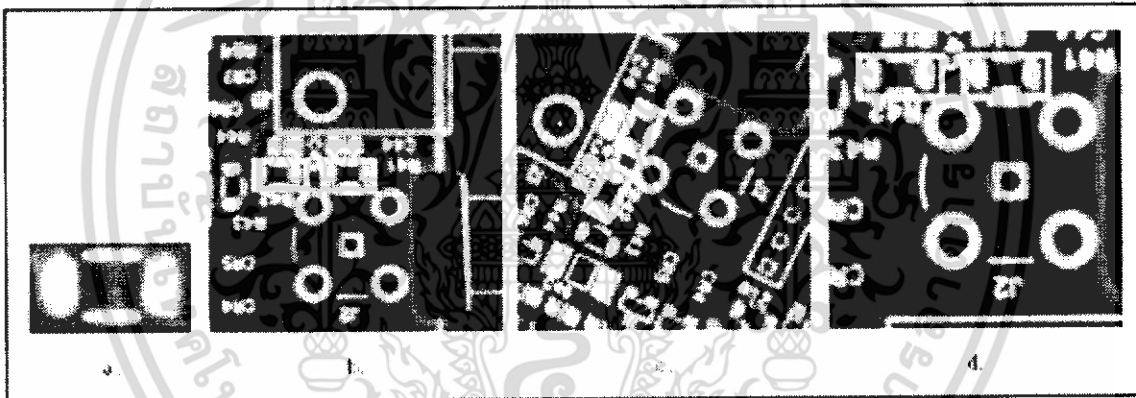
$$d_1(x) < d_j(x) \quad j=1, 2, \dots, W; j \neq 1 \quad (2.31)$$

2.2.4 Pattern Matching

Pattern Matching เป็นการจดจำภาพต้นแบบ (sub-image) แล้วนำไปค้นหาเพื่อจับคู่ในภาพแม่แบบ (image)

เทคนิค recognition ที่จะนำไปใช้ในการ match ในที่นี้จะขอเสนอวิธีทำ pattern matching 2 วิธีคือ

1. Minimum distance classifier
2. Correlation



รูปที่ (2.4) แสดงการค้นหา pattern matching รูป a. เป็น pattern image รูป b.-d. แสดงตำแหน่งที่ match ได้

2.2.4.1 Minimum distance classifier

เป็นการคำนวณด้วยปริมาณเชิง vector เราจะใช้การคำนวณแต่ละ class กับ prototype pattern vector โดยจะเป็นการหาระยะห่างระหว่าง unknown pattern กับแต่ละ prototype vector ระยะห่างที่สั้นที่สุดจะถูกนำมาคำนวณ เรากำหนดให้ prototype ของ pattern class เป็นค่าเฉลี่ย vector ของ pattern class นั้นๆ จะได้ว่า

$$m_j = \frac{1}{N_j} \sum_{x \in \omega_j} x, \quad j=1, 2, \dots, W \quad (2.32)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ N_j คือจำนวนของ pattern vector ของ class ω_j ในการระบุ class ให้กับ unknown vector x สามารถทำได้โดยการส่งไปอยู่ใน class ที่ใกล้เคียงกับ prototype ใช้กฎระยะทางของ Euclidean

$$D_j(x) = \|x - m_j\| \quad j=1, 2, \dots, W \quad (2.33)$$

เมื่อ $D_j(x)$ มีค่าเป็นระยะที่น้อยที่สุดจะบ่งบอกได้ถึง การ match ที่ดีที่สุด

2.2.4.2 Correlation

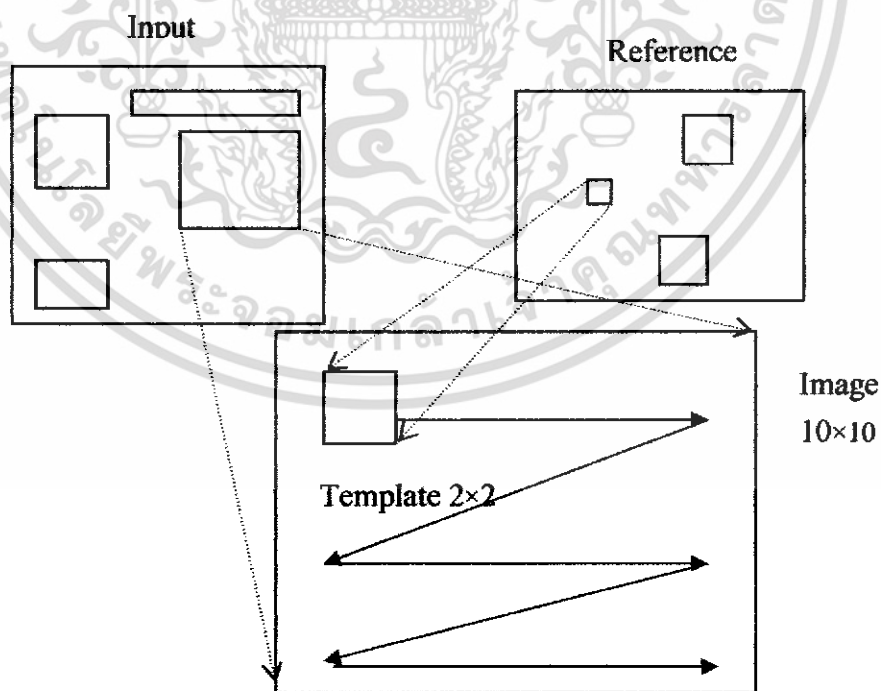
การ match ด้วยวิธี correlation จะเป็นการคำนวณในเทอมของ image โดยตรงโดย correlation เป็นความสัมพันธ์แบบหนึ่งใน Fourier transform

กำหนดให้ sub-image เป็น template $w(x,y)$ ขนาด $K \times L$ และ image $f(x,y)$ ขนาด $M \times N$ เมื่อ $K \leq M$ และ $L \leq N$ ความสัมพันธ์แบบ correlation ระหว่าง $f(x,y)$ และ $w(x,y)$ เขียนได้เป็น

$$c(i,j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x,y) f(x+i,y+j) \quad (2.34)$$

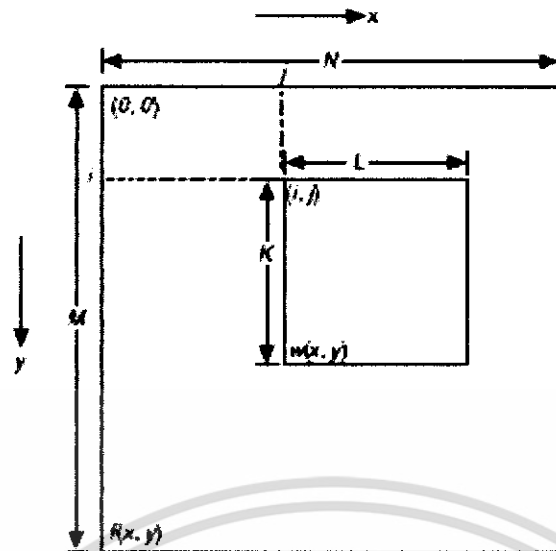
เมื่อ $x=0, 1, 2, \dots, M-1, y=0, 1, 2, \dots, N-1$

รูปที่ (2.5) แสดงขั้นตอนของ correlation จากสมการที่ (1) โดยมีจุด origin ของ f อยู่ที่จุดซ้ายบนสุด (0,0) และจุด origin ของ w อยู่ที่จุดกึ่งกลางของภาพ w สำหรับค่าหนึ่งๆของ (x,y) สมมติให้เป็น (x_0, y_0) ใน f แทนในสมการที่ (2.34) จะได้ค่า c มาค่าหนึ่ง เมื่อ x และ y มีค่าเปลี่ยนไป (w เคลื่อนไปรอบๆพื้นที่ใน f) ค่า $c(i,j)$ ที่มากที่สุดจะบ่งบอกถึงการจับคู่ที่ดีที่สุด



(a)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(b) ขั้นตอนของ correlation ของ f และ w
รูปที่ (2.5) การหา correlation

ความแม่นยำของ correlation จะลดลงเมื่อค่าค่าของ x และ y อยู่ในตำแหน่งใกล้ขอบของ image f นอกจากนี้ข้อเสียเปรียบของวิธีนี้คือเมื่อมีการเปลี่ยนแปลง amplitude ของ f และ w ค่าของ $c(i, j)$ จะเปลี่ยนไปด้วย ตัวอย่างเช่น เมื่อทุกค่าของ f ถูกเพิ่มเป็น 2 เท่า ค่า $c(i, j)$ จะเพิ่มเป็น 2 เท่าด้วย ซึ่งวิธีแก้ไขสามารถทำได้โดย ทำการ match โดยใช้ correlation coefficient ซึ่งนิยามได้ดังสมการที่ (2.35)

$$R(i, j) = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w})(f(x+i, y+j) - \bar{f}(i, j))}{\left[\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w})^2 \right]^{1/2} \left[\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (f(x+i, y+j) - \bar{f}(i, j))^2 \right]^{1/2}} \quad (2.35)$$

เมื่อ $x=0, 1, 2, \dots, M-1, y=0, 1, 2, \dots, N-1$

\bar{w} คือ ค่าเฉลี่ยของ pixel ใน w และ \bar{f} คือค่าเฉลี่ยของ f ในช่วงที่ทับกันสนิทกับตำแหน่งของ w ผลรวมที่ได้จะอยู่ในช่วงที่พิสัยรวมของทั้ง x และ y ค่า correlation coefficient $R(i, j)$ อยู่ในช่วง -1 ถึง 1 และเป็นอิสระต่อการเปลี่ยน scale ใน amplitude ของ f และ w

แม้ว่าการทำ normalized correlation โดย หา correlation coefficient จะช่วยแก้ปัญหาที่เกิดกับการเปลี่ยน amplitude ได้ แต่ก็ยังมีปัญหาในเรื่องของการเปลี่ยนขนาดและการหมุนอยู่ การ normalize สำหรับขนาดต้องใช้การ spatial scale ซึ่งจะทำให้การคำนวณช้าลงมาก โดยเราต้องทำการเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แม้ว่าการทำ normalized correlation โดย หา correlation coefficient จะช่วยแก้ปัญหาที่เกิดกับการเปลี่ยน amplitude ได้ แต่ก็ยังมีปัญหาในเรื่องของการเปลี่ยนขนาดและการหมุนอยู่ การ normalize สำหรับขนาดต้องใช้การ spatial scale ซึ่งจะทำให้การคำนวณช้าลงมาก โดยเราต้องทำการปรับขนาดของ template $w(x, y)$ ให้มีขนาดเท่ากับ pattern ใน image f ก่อนแล้วจึงนำไปหา correlation การหมุนนั้นเป็นเรื่องที่ยากยิ่งกว่าโดย เรา สามารถหมุน template $w(x, y)$ ให้ทำมุมเท่ากับมุมของ pattern ใน image f แล้วจึงนำไปหา correlation แต่โดยทั่วไปแล้วเราไม่สามารถทราบมุมที่แน่นอนของ pattern ที่ถูกหมุนได้ในการคำนวณจึงทำให้เสียเวลามาก

ข้อเสียของ correlation อีกอย่างคือใช้เวลาในการคำนวณค่อนข้างนานและความแม่นยำต่ำ วิธีที่จะแก้ไขคือ การใช้ข้อมูลของขอบ, มุม, และ โครงสร้างทางเรขาคณิตของ image เช่นจำนวนของเส้นตรงหรือวงกลมใน image มาใช้เป็นข้อมูลอ้างอิง จะทำให้การ match ถูกจำกัดจุดสนใจลง ซึ่งจะทำให้ความแม่นยำและรวดเร็วเพิ่มขึ้น

วิธีที่จะลดจุดสนใจนั้นในที่นี้ขอเสนอวิธี pyramidal matching โดยเป็นการลดขนาดของ image และ template โดย image และ template จะถูกนำไป sampling เพื่อลดความละเอียด (resolution) โดย image และ template สามารถลดเหลือขนาดเป็น 1 ใน 4 ของภาพเดิมทำให้การ match เร็วขึ้นได้

นอกจากนี้การหา correlation ใน frequency domain ก็สามารถทำได้เช่นกัน โดยการใช้ fast Fourier transform โดยถ้า image และ template มีขนาดเท่ากันแล้ววิธีนี้จะมีประสิทธิภาพมากกว่า การหา correlation ใน spatial domain ใน frequency domain นั้น correlation จะ ได้มาจากการคูณ fast Fourier transform ของ image ด้วย complex conjugate ของ fast Fourier transform ของ template แต่อย่างไรก็ตามในการนำมาใช้จริงยังทำได้ยาก

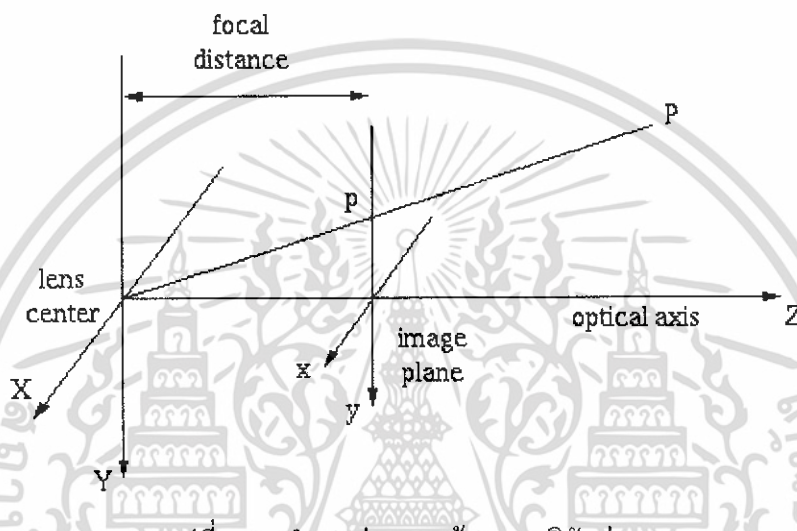
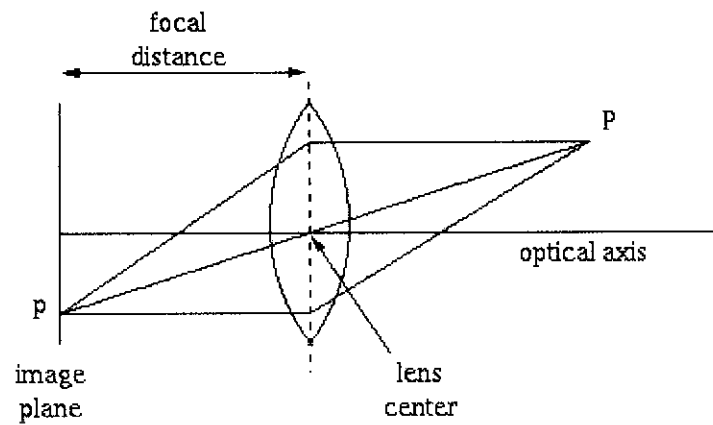
2.3 Stereo Vision: Triangulation

การวิเคราะห์ stereo triangulation เป็นการคำนวณตำแหน่งพิกัด 3-มิติ ของจุดที่พิจารณาในรูปภาพ โดยใช้ระบบ stereo vision ประกอบด้วยมุมมอง 2 ทิศทาง

2.3.1 Camera model

ตัวอย่างกล้องแบบรูเข็มที่ใช้สำหรับการสร้างภาพ 3-มิติ แสดงดังรูปที่ (2.6)

73157



รูปที่ (2.6) ตำแหน่งของกล้องและพิกัดต่างๆ

ความสัมพันธ์ระหว่างพิกัด โลก (world coordinate) ของจุด $P(X,Y,Z)$ กับพิกัดของ plane image (x,y) ในกล้องคือ

$$x = f\left(\frac{X}{Z}\right)$$

$$y = f\left(\frac{Y}{Z}\right)$$

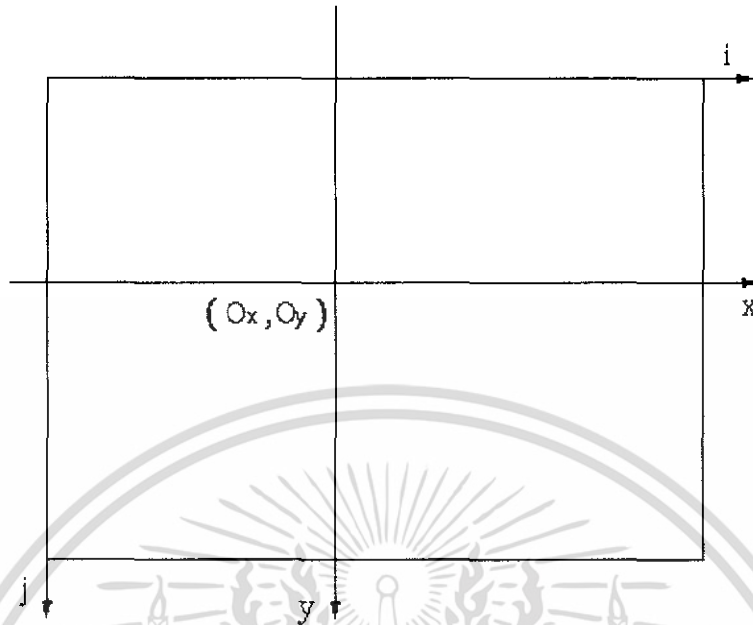
เมื่อ f คือ ระยะโฟกัสของเลนส์

2.3.2 Frame to image coordinate

โดยทั่วไปรูปภาพแบบดิจิทัลจะถูกเก็บไว้ใน frame buffer ซึ่งมีลักษณะเป็น matrix ของ pixels ซึ่งมีขนาด column W และ row H จากรูปที่ (2.7) กำหนดให้ (i,j) เป็นพิกัดของ frame ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยมีจุด origin อยู่ที่จุดมุมซ้ายบน ให้จุด (O_x, O_y) เป็นจุดโฟกัสของเลนส์ (จุดตัดระหว่าง optical axis กับ plane image) ในพิกัดเฟรม และให้ (x, y) เป็นพิกัดภาพ



รูปที่ (2.7) พิกัดของภาพกับ frame

ความสัมพันธ์ระหว่างพิกัดของภาพกับพิกัด frame แสดงได้ดังสมการ

$$x = (i - O_x) \times S_x \quad (2.36)$$

$$y = (j - O_y) \times S_y \quad (2.37)$$

เมื่อ S_x, S_y เป็นระยะแนวราบและแนวตั้งของ pixel ที่ติดกัน ใน frame buffer

2.3.3 Radial distortion

ความผิดเพี้ยนของเลนส์มีวิธีแก้ไขดังสมการนี้

$$x = x_d \times (1 + k_1 r_d^2) \quad (2.38)$$

$$y = y_d \times (1 + k_1 r_d^2) \quad (2.39)$$

เมื่อ (x_d, y_d) คือพิกัดภาพของภาพที่ผิดเพี้ยน โดย r_d เท่ากับ $\sqrt{x_d^2 + y_d^2}$ และ k_1 เป็นค่าคงที่โดยขึ้นอยู่กับความผิดเพี้ยนของเลนส์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์ระหว่างพิกัดภาพและเฟรมขณะเกิด lens distortion คือ

$$x = (i_d - O_x) \times S_x \times (1 + k_1 r_d^2) \quad (2.40)$$

$$y = (j_d - O_y) \times S_y \times (1 + k_1 r_d^2) \quad (2.41)$$

2.3.4 Scanning and sampling

ถ้าระยะของจุดที่อยู่ติดกันจากกล้องดิจิทัลเป็น (P_x, P_y) แล้วระยะของ pixels ในภาพเป็น

$$S_x = a_x * P_x \quad (2.42)$$

$$S_y = P_y \quad (2.43)$$

เมื่อ a_x คือ scale factor ขึ้นกับการระยะการเคลื่อนที่ค้นหาในแนวราบ และความถี่ sampling โดยในทางกลับกันจะไม่มี sampling ในแนวดิ่ง

2.3.5 Triangulation

ตำแหน่ง 3-มิติ (X, Y, Z) ที่จุด P สามารถนำมาสร้างบน image plane ของกล้องได้เมื่อเราทราบตำแหน่งและมุมที่แน่นอนของกล้องทั้งสอง

กำหนดให้พิกัดโลก 3-มิติ อ้างอิงเป็นระบบอ้างอิงของกล้องทางซ้าย กล้องทางขวาจะเคลื่อนที่และหมุน โดยขึ้นกับกล้องทางซ้าย การแปลงจะประกอบไปด้วยตัวแปร 6 ค่า

เงื่อนไขที่ง่ายที่สุดที่เราพิจารณาคือเมื่อ optical axis ของกล้องทั้ง 2 ขนานกันและกล้องทางขวามือเคลื่อนที่เฉพาะตามแกน X

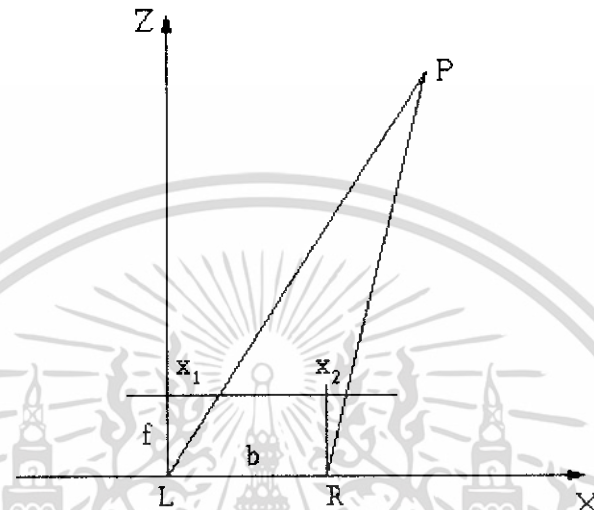
2.3.6 Geometry for parallel cameras

กำหนดให้ตำแหน่งของกล้องตามรูปที่ (3) เป็นตำแหน่งมาตรฐานที่เราจะใช้พิจารณา

1. L และ R เป็นกล้องซ้ายและขวาตามลำดับ โดยมีมุมมองที่ขนานกัน และมี f เป็นระยะโฟกัสของกล้องทั้งสอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. baseline ซึ่งเชื่อมต่อกล้องทั้งสองและตั้งฉากกับ optical axis กำหนดให้ b คือระยะห่างระหว่างจุดศูนย์กลางเลนส์ของกล้องทั้งสอง
3. Optical axis อยู่ใน plane XZ มีแกน XY ขนานกับ image plane ของกล้องทั้งสอง แกน X อยู่ที่ตำแหน่งของ baseline และที่จุด origin $O(X,Y,Z)$ ของจุดอ้างอิงโลก (world reference) เป็นจุดกึ่งกลางเลนส์ของกล้องซ้าย



รูปที่ (2.8) รูปมาตรฐานของการทำ stereo triangulation

จากรูปนำข้อกำหนดต่างๆมาเขียนเป็นสมการของ stereo triangulation ได้ดังนี้

$$Z = (b * f) / (x_1 - x_2) \quad (2.44)$$

$$X = x_1 * Z / f \quad (2.45)$$

$$Y = y_1 * Z / f \quad (2.46)$$

2.3.7 Non-parallel cameras

เมื่อกล้องทั้งสองไม่ขนานกัน กล่าวคือกล้องทางขวาถูกหมุน ในที่นี้พิจารณาที่มุมที่หมุนมีขนาดเล็ก

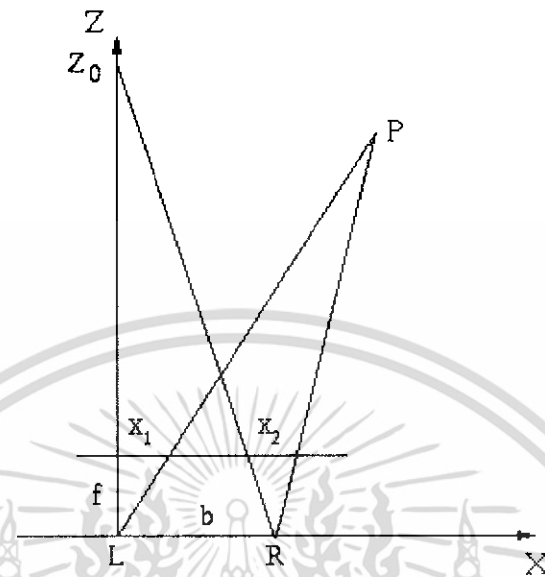
การหมุนรอบแกน Y (θ)

ในกรณีนี้ optical axis จะไม่ขนานกันแต่ก็ยังอยู่บนระนาบ XZ โดยมีจุดตัดที่จุด $(0,0,Z_0)$ เรียกว่า fixation point โดยตำแหน่งของจุดนี้จะอยู่ข้างหลังกล้องได้เมื่อ $(Z_0 < 0)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้ θ เป็นมุมที่เปลี่ยนไป จะได้ว่า

$$Z_0 = b / \tan(\theta)$$



รูปที่ (2.9) stereo triangulation เมื่อกล้องสองตัวไม่ขนานกัน
ถ้ามุมที่หมุนไปมีขนาดเล็ก จากรูปที่ (2.9) เราสามารถเขียนสมการได้ว่า

$$Z = (b * f) / (x_1 - x_2 + f * b / Z_0) \quad (2.47)$$

การหมุนรอบแกน X (ϕ)

การหมุนรอบแกน X มีผลกระทบต่อ การสร้างพิกัด Y เรา กำหนดให้ ϕ เป็นมุมที่หมุน จะ
ได้ stereo triangulation เป็น

$$Z = (b * f) / (x_1 - x_2)$$

$$X = x_1 * Z / f$$

$$Y = y_1 * Z / f + \tan(\phi) * Z \quad (2.48)$$

การหมุนรอบแกน Z (ψ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การหมุนรอบ optical axis โดยทั่วไปเราจะแก้ไขด้วยการหมุนภาพก่อน แล้วจึงทำการคำนวณ matching และ triangulation

ในที่นี้กำหนดให้มุมของการหมุนของกล้องขวรอบ optical axis เป็น ψ

2.3.8 เงื่อนไขทั่วไป

กำหนดให้ translation vector T และ rotation matrix R อธิบายการเปลี่ยนตำแหน่งจากกล้องซ้ายไปกล้องขวา คำตอบของ stereo triangulation หาได้จากสมการ

$$p' = RT(p - T) \quad (2.49)$$

เมื่อ p และ p' คือพิกัดของ P ในพิกัดของกล้องซ้ายและขวาตามลำดับ และ RT เป็น transpose หรือ inverse ของ matrix R



บทที่ 3

ขั้นตอนการทดลอง

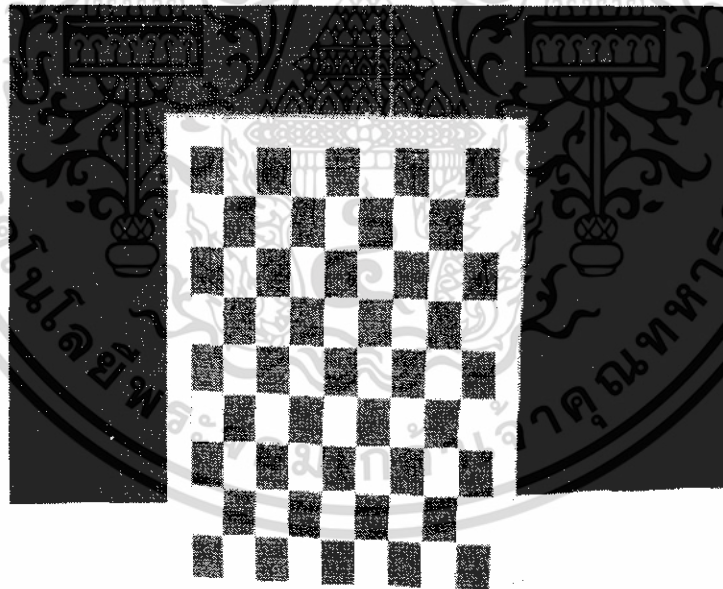
อุปกรณ์ที่ใช้ในการทดลองประกอบด้วย

1. ระนาบรูปดาวรางหมากรุก
2. กล้อง web cam จำนวน 2 ตัว
3. laser pointer 1 ตัว
4. โปรแกรมที่ใช้ในการทำ calibration, pattern matching และ triangulation

3.1 camera calibration

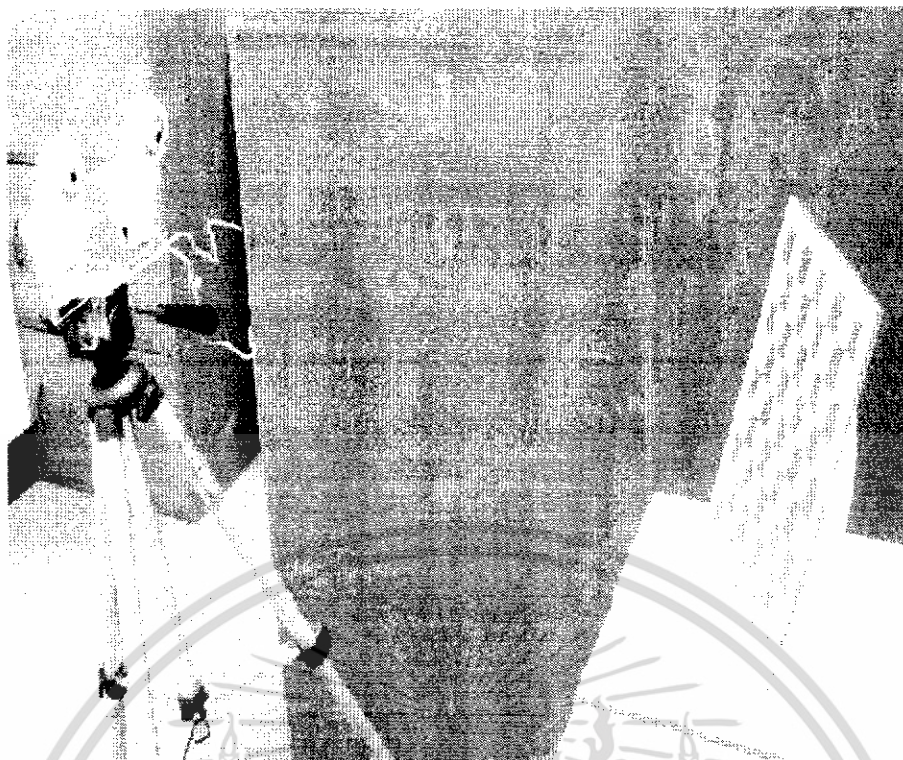
ขั้นตอนการทดลอง

1. ใช้กล้อง web cam ทั้ง 2 ตัวถ่ายภาพนิ่งของระนาบรูปดาวรางหมากรุกพร้อมกัน โดยหันระนาบในทิศทางต่างๆกันทั้งหมด 10 ทิศทาง (ค่า N , ของกล้องทั้ง 2 ตัวเท่ากับ 10) เพราะฉะนั้นจะได้ภาพของระนาบรูปดาวรางหมากรุกทั้งหมด 20 ภาพจากกล้องทั้ง 2 ตัว



รูป 3.1 แสดงภาพระนาบรูปดาวรางหมากรุกที่ใช้ในการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

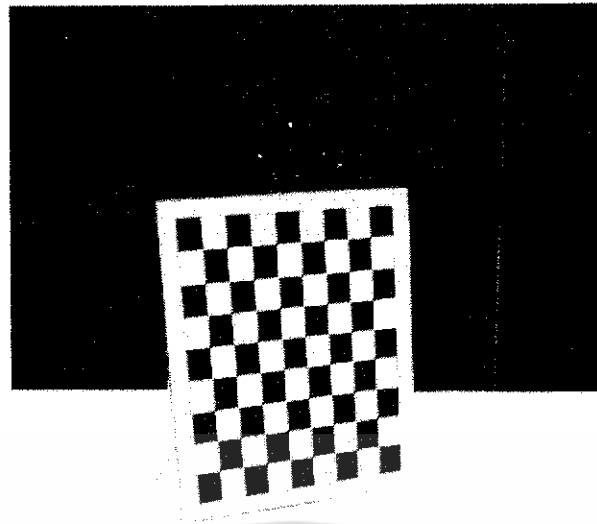


รูป 3.2 แสดงภาพการถ่ายภาพพระนารูปตรางหมากรุกโดยกล้อง web cam ทั้ง 2 ตัวซึ่งต่ออยู่กับขาตั้งกล้อง

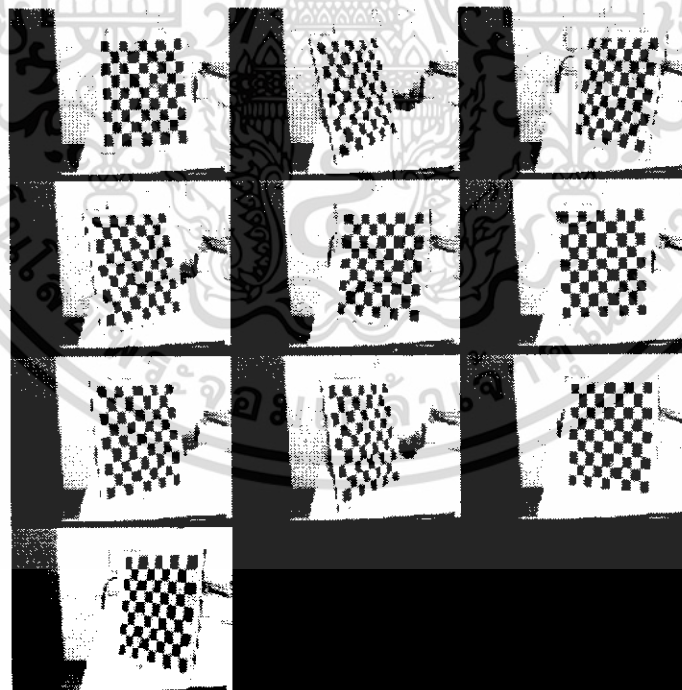


รูปที่ 3.3 แสดงภาพ plate ที่ใช้กำหนดมุมในการหันของพระนารูปตรางหมากรุก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

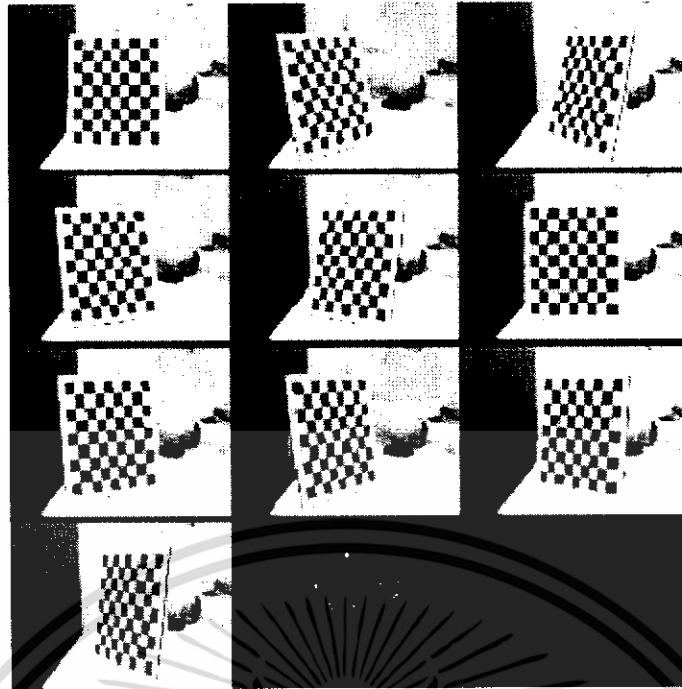


รูปที่ 3.4 ระนาบรูปตารางหมากรุกทิศทางที่ 4 จากกล้องด้านขวา



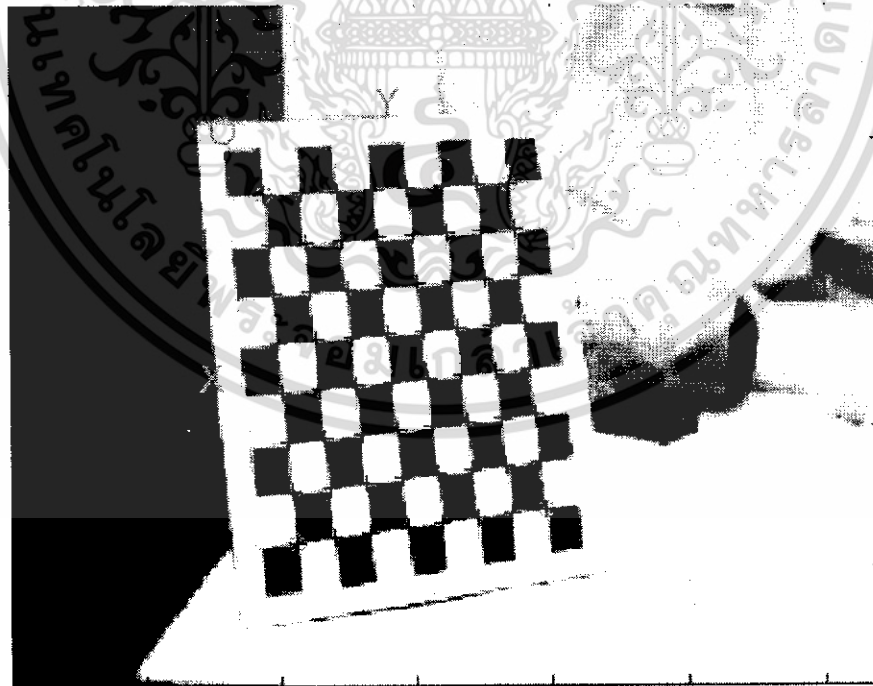
รูป 3.5 ภาพถ่ายของระนาบรูปตารางหมากรุกจากกล้องด้านซ้ายทั้ง 10 ทิศทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



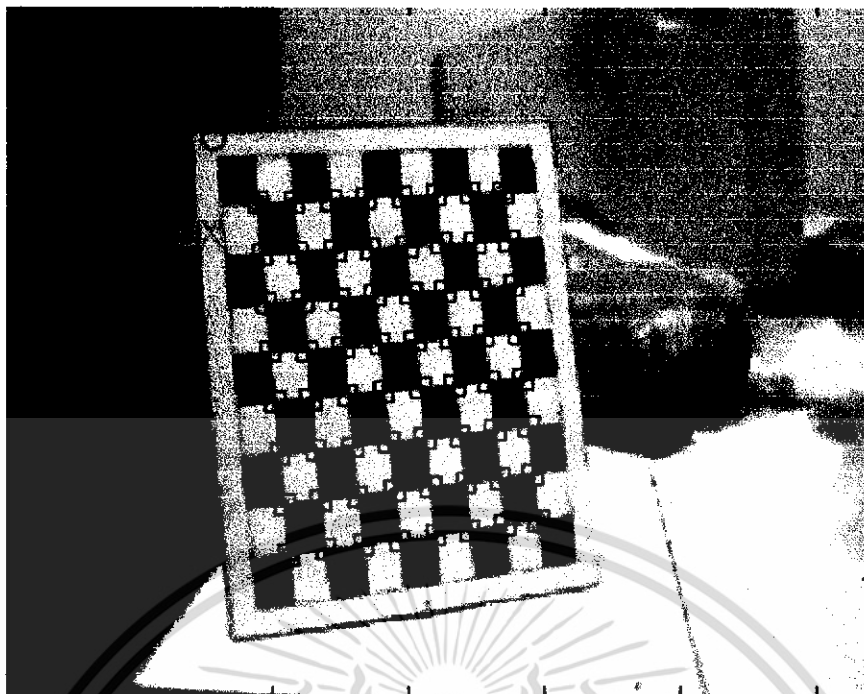
รูป 3.6 ภาพถ่ายของระนาบรูปดาวทรงกลมจากกล้องด้านขวาทั้ง 10 ทิศทาง

2. นำภาพทั้งหมดของกล้องด้านขวาไปทำการหามุมของรูปสี่เหลี่ยมในดาวทรงกลม
- รูปที่ 3.7 และ 3.8



รูปที่ 3.7 แสดงการหามุมจากภาพของกล้องด้านขวาทิศทางที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 แสดงมุมที่ถูก extract ออกมาได้

จากรูปที่ 3.8) จะ ได้ค่า input ในการ calibrate คือค่า dX , dY และค่า N_0 ซึ่งคือค่าจำนวนมุมของระนาบรูปตารางหมากรุกที่ extract ออกมาได้ นั้นเองจากภาพได้มุมทั้งหมด 64 มุม ภาพมีจำนวน 10 ภาพ เพราะฉะนั้นค่า N_0 เท่ากับ $64 \times 10 = 640$ ส่วนค่า dX และ dY นั้น เราจะกำหนดให้เป็น 19 mm. ทั้งคู่ เนื่องจากด้านของสี่เหลี่ยมในรูปตารางหมากรุกที่ใช้มีค่าเป็น 19 mm. ทั้ง 2 ด้าน เพราะฉะนั้น พิกัดที่วัดออกมาได้จะมีหน่วยเป็น mm.

3. นำค่าตัวแปรต่างๆที่ได้ไปทำการ calibrate โดยโปรแกรม matlab แล้วทำเช่นเดิมกับภาพที่ได้จากกล้องด้านซ้าย

4. จะ ได้ค่า intrinsic parameter ซึ่งคือค่า Focal Length (f_c) , Principal point (cc) , Skew (α_c) , Distortion (kc) และ Pixel error (err) และค่า extrinsic parameter ซึ่งคือค่า Rotation vector (om) และ Translation vector (T) ของกล้องทั้ง 2 ตัว

3.2 ขั้นตอนการสร้างภาพ 3 มิติ

ขั้นตอนนี้จะเป็นการประมวลผลของระบบทั้งหมดเพื่อนำไปสร้างภาพ 3 มิติโดยจะประกอบด้วยส่วนของ pattern matching และ triangulation ซึ่งในโครงการนี้เราใช้โปรแกรม NI LabVIEW ช่วยในการเขียนโปรแกรมขึ้นมา ซึ่งใน LabVIEW นี้ประกอบไปด้วยฟังก์ชันการทำงานแบบ machine vision ซึ่งทำให้ง่ายต่อการนำไปใช้งาน

3.2.1 Pattern matching

ขั้นตอนนี้เป็นการทดลองเพื่อหาพิกัด (x,y) ของจุดตำแหน่งที่พิจารณาเพื่อจะนำไปใช้ในการ triangulate ต่อไป โดยใช้อุปกรณ์และการทดลองดังรูปที่ (3.9)



รูปที่ (3.9) อุปกรณ์และการทดลอง pattern matching

ขั้นตอนในการหาพิกัด (x,y)

1. ใช้จุดเลเซอร์เป็นตัวระบุตำแหน่งที่เราจะใช้พิจารณา
2. ทำการบันทึกวิถีโอดีด้วยกล้อง 2 ตัวซึ่งอยู่ห่างกัน
3. นำไฟล์วิถีโอดีที่ได้มา run ใน โปรแกรมซึ่งใช้เทคนิค pattern matching

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตำแหน่งและพิกัด (x,y) ของ template ใน image input นั้นๆ และนำค่าไปใช้ในส่วน of stereo triangulation

3.2.2 Stereo triangulation

ขั้นตอนนี้ เป็นการหาพิกัด 3 มิติของวัตถุที่ถูกชี้โดย laser pointer โดยนำค่า intrinsic และ extrinsic parameter จากการ calibrate และค่าพิกัด x,y ที่ได้จาก pattern matching มาคำนวณหาตำแหน่งใน 3 มิติของวัตถุโดยใช้เทคนิค stereo triangulation เมื่อทราบตำแหน่งแล้วก็สามารถนำมาแสดงผลเป็นกราฟจุด 3 มิติได้

3.3 ขั้นตอนการทดลองระบบเพื่อหาความแม่นยำของระบบ

1. ทดลอง กำหนดระยะห่างระหว่างกล้อง 3 ระยะ คือที่ 5 cm., 10 cm., และ 15 cm. และวางตำแหน่งของวัตถุอ้างอิงซึ่งมีขนาด 30 mm. ห่างจากกล้อง 4 ระยะคือที่ 15 cm., 20 cm., 25 cm. และ 30 cm.
2. ทำการสแกนเพื่อหาขนาดของวัตถุแล้วนำมาเปรียบเทียบกับขนาดจริงของวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

4.1 calibration

ผลของค่า intrinsic parameter ประกอบไปด้วย

- Focal Length (fc) คือค่าระยะโฟกัส เป็นเวกเตอร์ 2×1 ประกอบด้วย fc(1) และ fc(2) โดยแสดงเป็นหน่วยของ horizontal และ vertical pixels ตามลำดับ โดยปกติค่าของ fc(1) และ fc(2) จะมีค่าใกล้เคียงกัน อัตราส่วน fc(1) / fc(2) เรียกว่า aspect ratio มีค่าห่างจาก 1 ถ้า pixel ไม่เป็นสี่เหลี่ยม ดังนั้น ธรรมชาติของ pixel ก็ต้องจะไม่เป็นรูปสี่เหลี่ยมสมบูรณ์
- Principal point (cc) แสดงเป็นพิกัด pixel ของแกน x และ y ตามลำดับ
- Skew (alpha_c) คือค่าสัมประสิทธิ์ของมุมระหว่างแกน x และ y ของ pixel ค่านี้จะถูกเซตเป็น 0 (มุม 90 องศา) ตลอดเพราะในปัจจุบันกล้องที่ผลิตออกมาจะถูกประมาณให้พิกเซลเป็นรูปสี่เหลี่ยม
- Distortion (kc) คือค่าสัมประสิทธิ์ของการบิดเบือนของเลนส์ แสดงค่าเป็นเวกเตอร์ 5×1 ประกอบด้วยค่า kc(1) , kc(2) , kc(3) , kc(4) และ kc(5) ค่า kc(1) ถึง kc(4) เป็นค่าการ distort ทางรัศมีส่วนค่า kc(5) เป็นค่าการ distort ทางเส้นสัมผัสซึ่งโดยปกติค่า kc(5) นี้จะถูกเซตให้เป็น 0 ตลอด เพราะโดยการประมาณว่าเลนส์ที่ผลิตได้ในปัจจุบันมีความสมบูรณ์ของตำแหน่งศูนย์กลาง

ผลของค่า extrinsic parameter ประกอบไปด้วย

- Rotation vector (om) เป็นเวกเตอร์ขนาด 3×1 แสดงค่า rotation ระหว่างกล้องทั้ง 2 ตัว ค่า om สัมพันธ์กับ rotation matrix (R) คือ $R = Rodrigues(om)$
- Translation vector (T) มีขนาด 3×1 ประกอบด้วย T_x, T_y และ T_z ตามลำดับ แสดงค่า translation ระหว่างกล้องทั้ง 2 ตัว

ที่ระยะห่างระหว่างกล้อง 7 cm.

ค่า intrinsic parameter ของกล้องด้านขวาเป็นดังนี้

$$\begin{aligned}
 fc_left &= [451.15038 \quad 450.18022] \pm [2.56557 \quad 2.51871] \\
 cc_left &= [147.72152 \quad 98.59365] \pm [3.10861 \quad 3.34596] \\
 alpha_c_left &= [0.00000] \pm [0.00000] \Rightarrow \text{angle of pixel axes} = 90.0 \pm 0 \text{ degrees} \\
 kc_left &= [-0.23567 \quad 0.09977 \quad -0.00362 \quad -0.00223 \quad 0.00000] \pm [0.03315 \quad 0.25685 \\
 &\quad 0.00153 \quad 0.00116 \quad 0.00000]
 \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า intrinsic parameter ของกล้องด้านซ้ายเป็นดังนี้

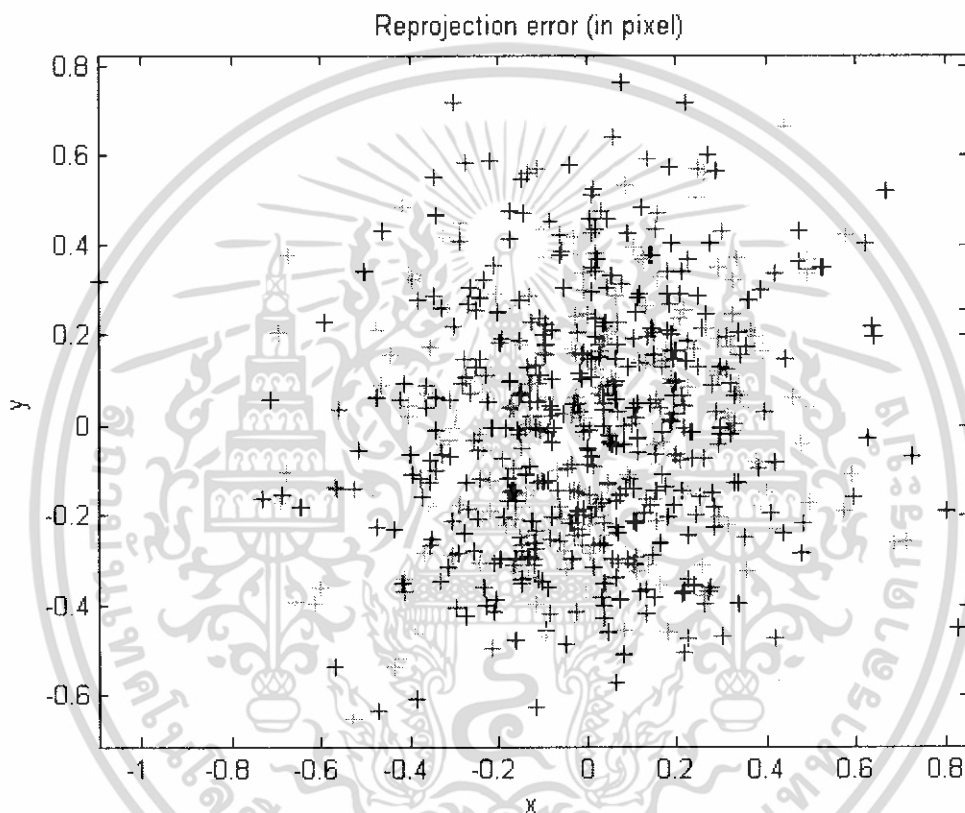
$$fc_right = [450.38868 \quad 449.52331] \pm [2.52869 \quad 2.48404]$$

$$cc_right = [168.65950 \quad 98.24767] \pm [3.16183 \quad 3.31703]$$

$$\alpha_c_right = [0.00000] \pm [0.00000] \Rightarrow \text{angle of pixel axes} = 90.0 \pm 0 \text{ degrees}$$

$$kc_right = [-0.22814 \quad 0.04698 \quad -0.00281 \quad -0.00050 \quad 0.00000] \pm [0.03691 \quad 0.32751 \\ 0.00153 \quad 0.00117 \quad 0.00000]$$

ค่าด้านหลังคือค่าของ error ความคลาดเคลื่อนของ intrinsic parameter



รูปที่ 4.1 แสดงภาพค่า pixel error ของจุดที่ใช้ calibrate ทุกจุด (error ไม่เกิน 0.8 pixel)

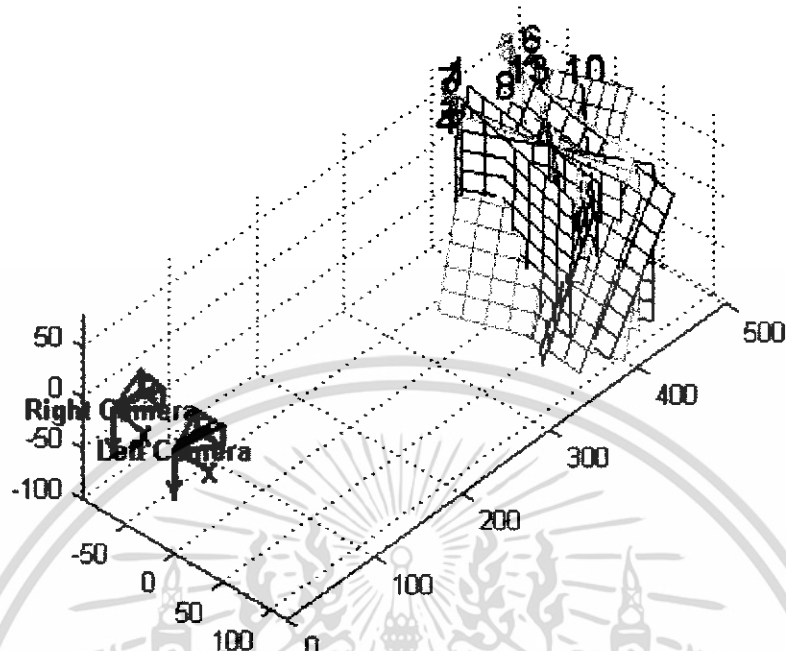
ค่า extrinsic parameter ของกล้องทั้ง 2 ตัวเป็นดังนี้

$$om = [0.00778 \quad -0.17091 \quad -0.00695] \pm [0.00963 \quad 0.00913 \quad 0.00100]$$

$$T = [68.40765 \quad -1.38162 \quad 6.01761] \pm [0.55284 \quad 0.43933 \quad 2.59171]$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Extrinsic parameters



รูปที่ 4.2 แสดงภาพ extrinsic ของกล้องทั้งสองกับระนาบรูปตารางหมากรุกในทิศทางต่างๆ

4.2 ผลการทดลองการสร้างภาพ 3 มิติ

4.2.1 ผลการทดลองหาความแม่นยำของระบบ

ตำแหน่งของวัตถุอ้างอิงซึ่งมีขนาด 30 mm.

ที่ระยะห่างระหว่างกล้องทั้งสอง 5 cm.

ระยะห่างของวัตถุ (cm.)	ขนาดของวัตถุที่วัดได้ (mm.)	% ความผิดพลาด
15	28.64	4.53
20	29.12	2.93
25	28.16	6.13
30	27.54	8.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตำแหน่งของวัตถุอ้างอิงซึ่งมีขนาด 30 mm.

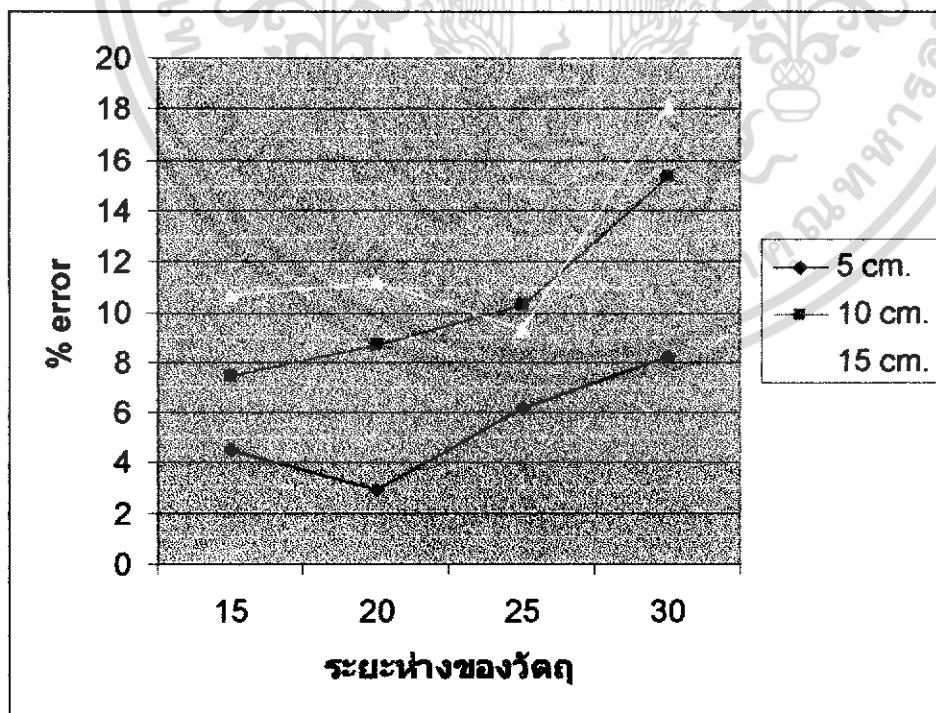
ที่ระยะห่างระหว่างกล้องทั้งสอง 10 cm.

ระยะห่างของวัตถุ (cm.)	ขนาดของวัตถุที่วัดได้ (mm.)	% ความผิดพลาด
15	27.77	7.43
20	27.38	8.73
25	26.91	10.3
30	25.41	15.3

ตำแหน่งของวัตถุอ้างอิงซึ่งมีขนาด 30 mm.

ที่ระยะห่างระหว่างกล้องทั้งสอง 15 cm.

ระยะห่างของวัตถุ (cm.)	ขนาดของวัตถุที่วัดได้ (mm.)	% ความผิดพลาด
15	26.78	10.7
20	26.63	11.23
25	27.19	9.36
30	24.56	18.13

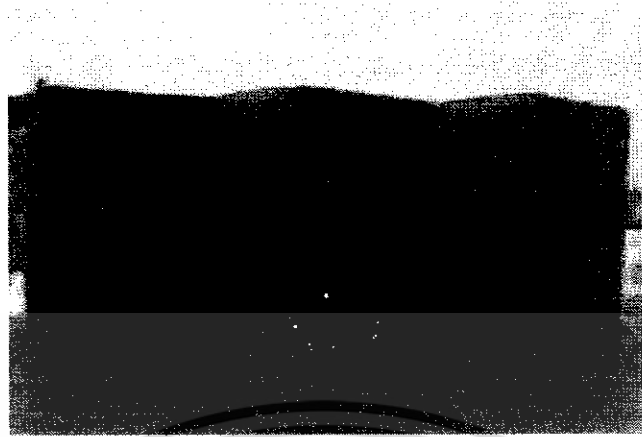


รูปที่ 4.3 กราฟแสดงความสัมพันธ์ระหว่างระยะวัตถุต่อ % error

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 ผลการทดลองสร้างภาพ 3 มิติ

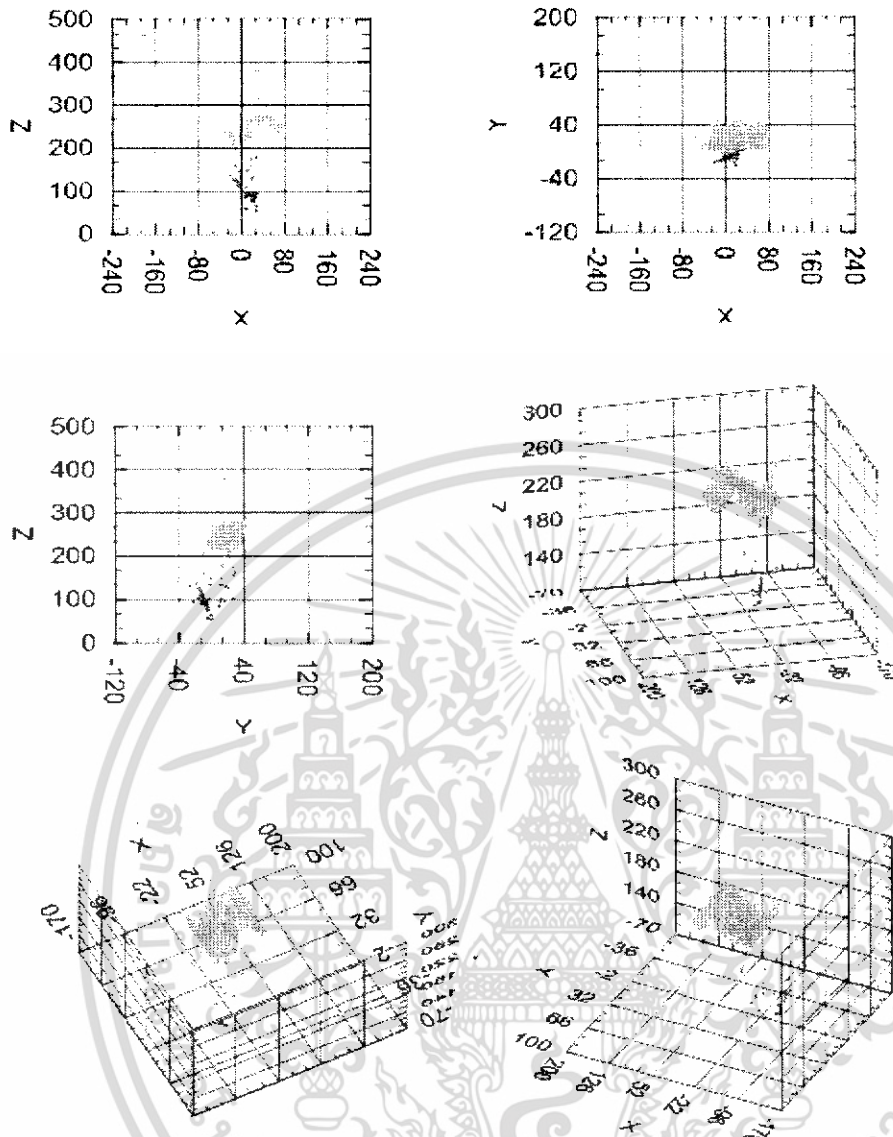
สแกนวัตถุจากรูปที่ 4.4 ที่ระยะห่างระหว่างกล้อง 7 cm. และวัตถุห่างจากกล้อง 20 cm.



รูปที่ 4.4 ภาพของวัตถุที่ใช้สแกน

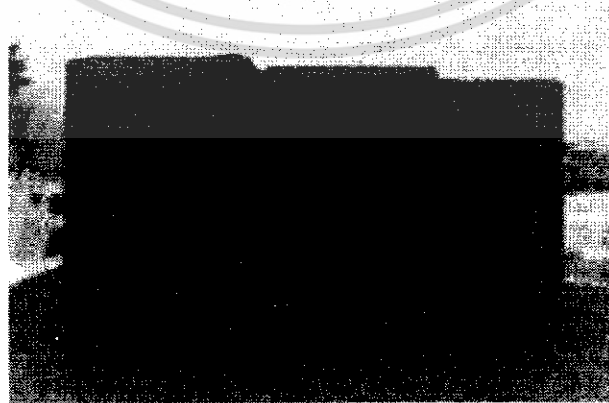


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



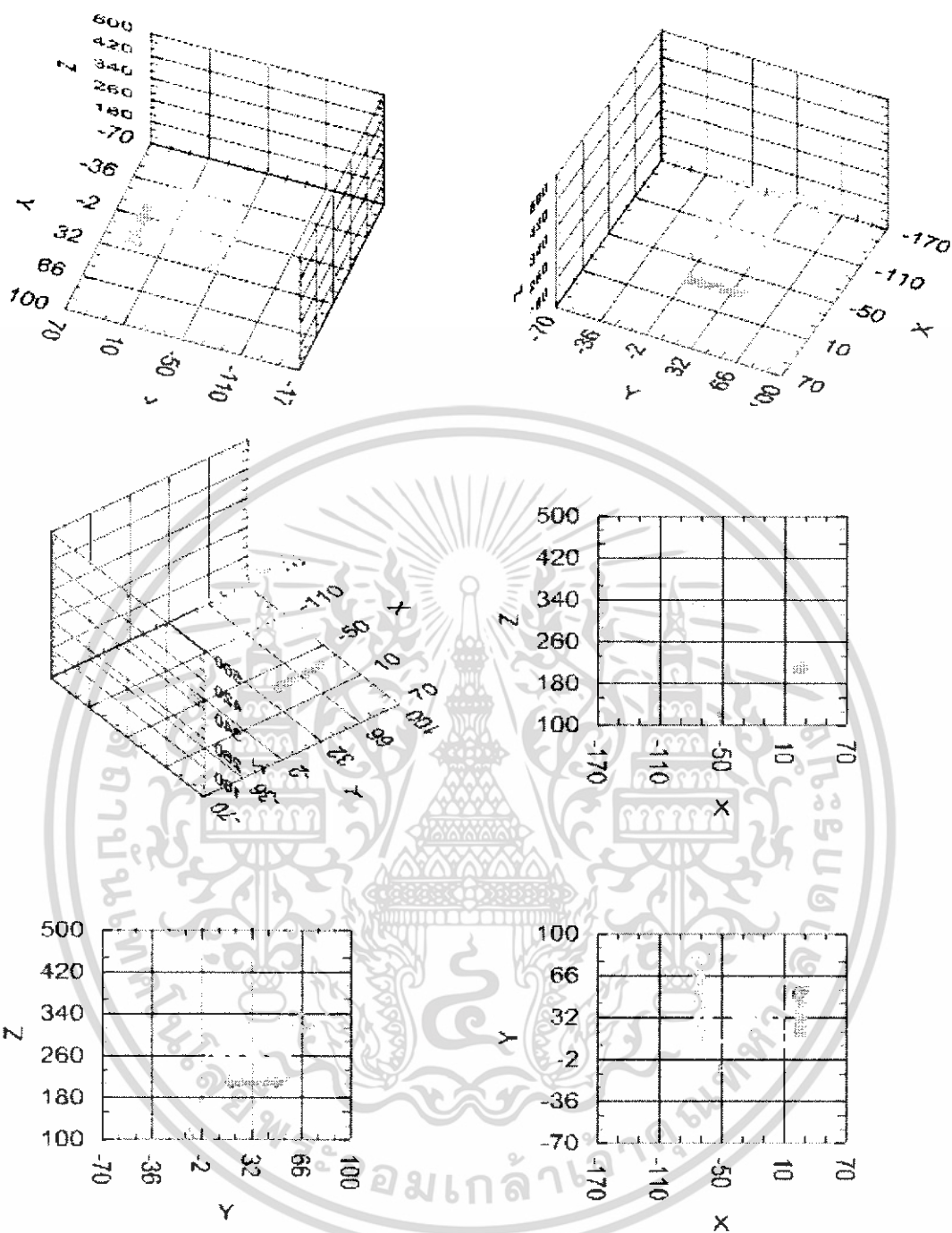
รูปที่ 4.5 ภาพที่สแกนได้ของรูปที่ 4.3 หน่วยเป็น mm.

สแกนวัตถุตั้งรูปที่ 4.6 ที่ระยะห่างระหว่างกล้อง 7 cm. และวัตถุห่างจากกล้อง 20 cm.



รูปที่ 4.6 ภาพของวัตถุที่ใช้สแกน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 ภาพที่สแกนได้ของรูปที่ 4.5 หน่วยเป็น mm.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการทดลอง

โครงการนี้เป็นารสร้างภาพ 3 มิติ โดยใช้เทคนิค camera calibration โดยระบบมีราคาถูก, สะดวก, ยืดหยุ่น และสร้างภาพ 3 มิติที่มีความแม่นยำได้ดี

ในการทดลองนี้เป็นการเป็นการทำเพื่อหาค่า parameter ต่างๆ และความสัมพันธ์ของกล้อง ทั้ง 2 ตัว โดยจะคำนวณค่า intrinsic parameter ของกล้องทั้ง 2 ที่ละตัว เพราะฉะนั้น ไม่มีข้อจำกัดว่า จะต้องใช้กล้องตัวเดียวกัน(ถ่ายทีละมุมมอง)หรือยี่ห้อเดียวกันในการทดลอง และมีการคำนวณค่า extrinsic parameter ทำให้การทดลองนี้มีความยืดหยุ่นสูงเพราะจะไม่จำกัดว่ากล้องจะต้องอยู่ห่างกันเท่าไรหรือทำมุมต่อกันแค่ไหน แต่อย่างไรก็ตามหากกล้องทำมุมต่อกันมากเกินไป(กล้องหันหัวคนละทิศกัน) จะเกิด error ขึ้นคือระบบจะไม่สามารถคำนวณค่า extrinsic ได้

จากการทดลองเพื่อหาความแม่นยำของระบบพบว่า ในระยะห่างของกล้องและวัตถุที่เหมาะสมคือ 5 cm. และ 20 cm. ตามลำดับ ที่ระยะห่างของวัตถุมากความคลาดเคลื่อนจะมากขึ้นด้วย เนื่องจากความคลาดเคลื่อนของแต่ละ pixel เมื่อวัตถุอยู่ไกลจะส่งผลให้เกิดการคำนวณผิดพลาดได้มาก

ปัญหาที่พบมากในการทดลองนี้คือ เนื่องจากโครงการนี้มีค่าใช้จ่ายต่ำกล้องที่ใช้มีความละเอียดต่ำ (320×240 pixel) จึงเป็นผลให้ความแม่นยำน้อยกว่าที่ควรจะเป็น โดยเมื่อวัตถุห่างเกิน 30 cm. ความผิดพลาดจะมากถึง 10% และถ้าไกลกว่านั้นจะคลาดเคลื่อนจนไม่สามารถยอมรับได้ เพื่อให้ได้ความแม่นยำสูง ควรใช้กล้องที่มีความละเอียดสูง และวางตำแหน่งให้เหมาะสม สภาพแวดล้อมก็เป็นปัญหาที่พบมากโดยในสถานะที่มีแสงรบกวน แสงจะทำให้ detect จุดเลเซอร์ได้ยากส่งผลให้เกิดการผิดพลาดขึ้น การแก้ไขจึงควรทดลองในห้องมืด

บรรณานุกรม

1. Roger Y. TSALIEEE Journal of Robotics and Automatio,Vol.RA-3,No.4,August 1987
2. T.Fang and L.A.Piegl, "Delaunay Triangulation in three dimension,"IEEE Graphics and Applications, Vol.15,No.5,1995,pp.62-69
3. Rafael C. Gonzalez and Richard E. Woods.Digital Image Processing.Prentice-Hall,Inc.,2002

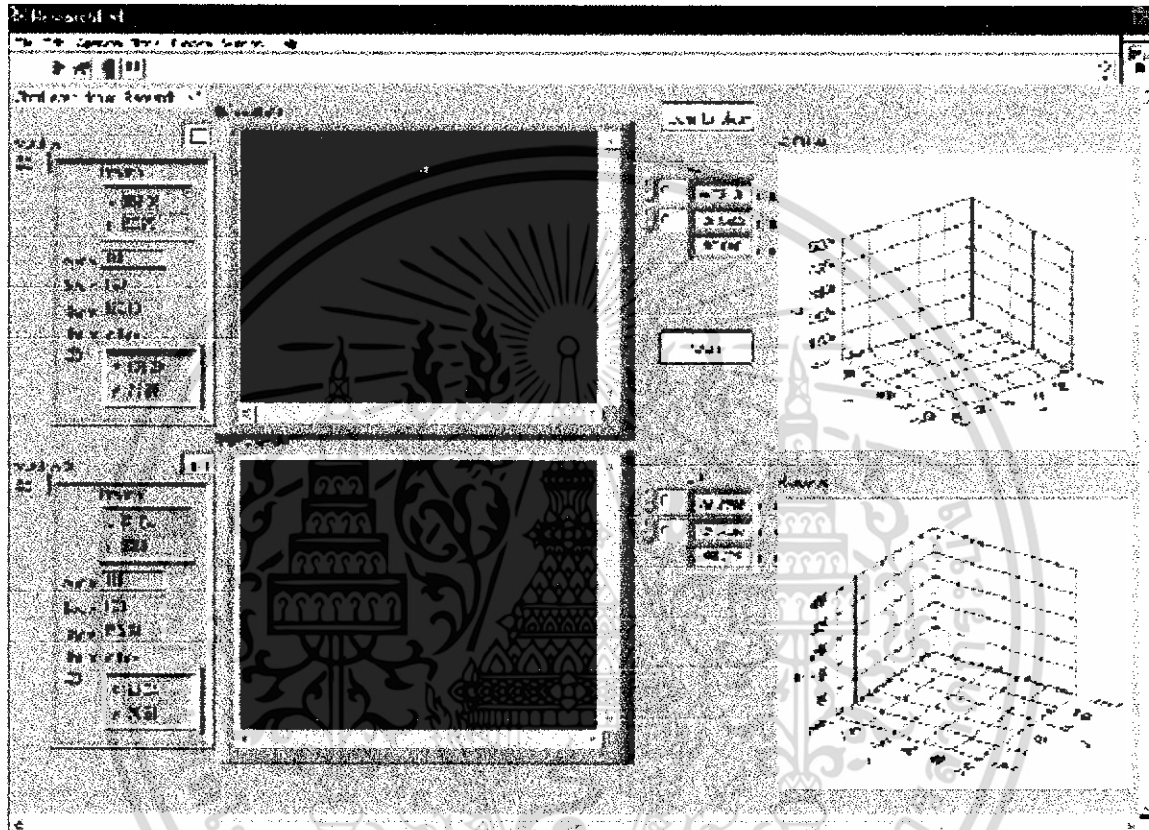


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

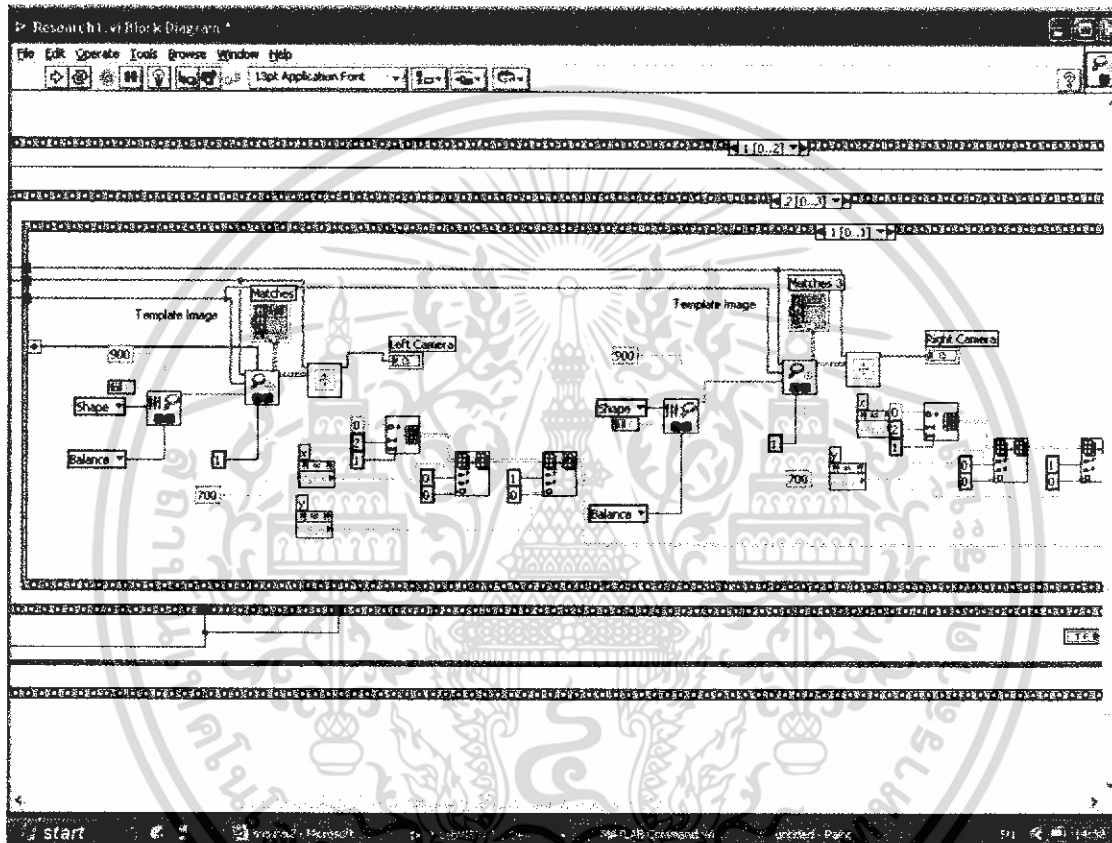
Front Panel ของ โปรแกรม



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

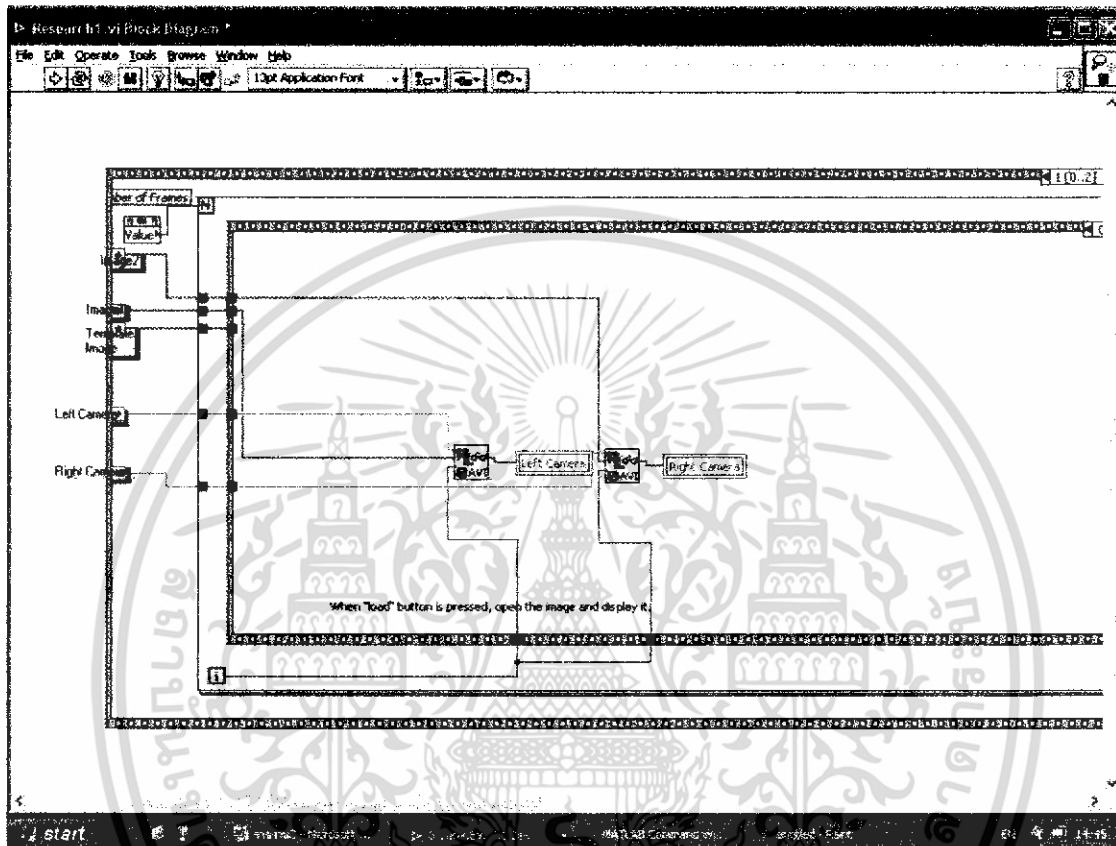
Block Diagram ของ โปรแกรม

- Pattern Matching



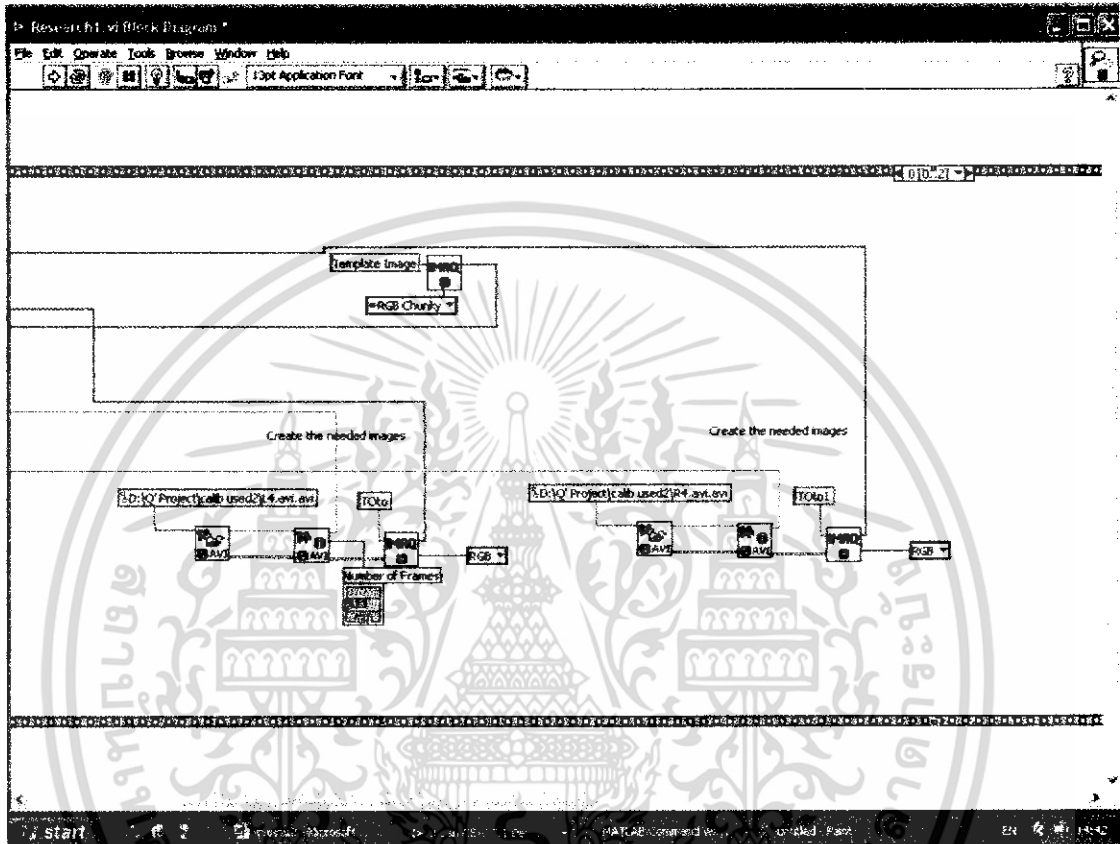
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Open .avi file



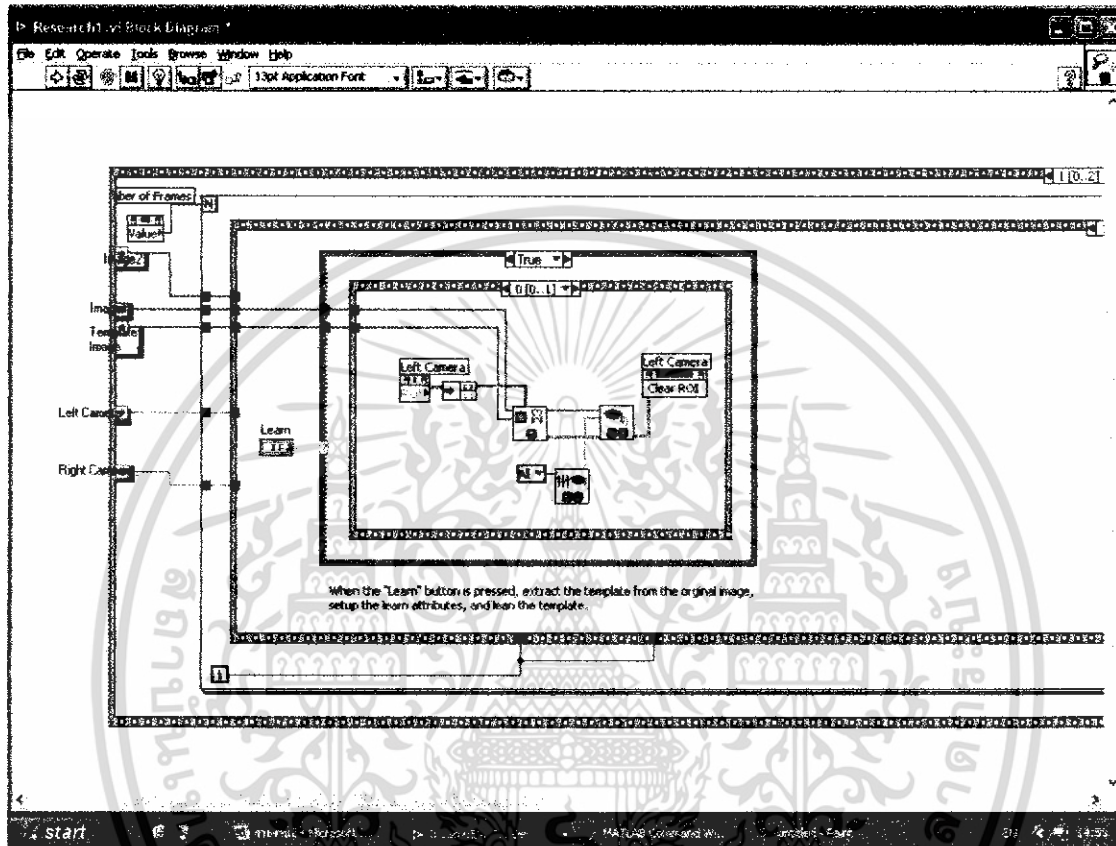
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Load .avi file



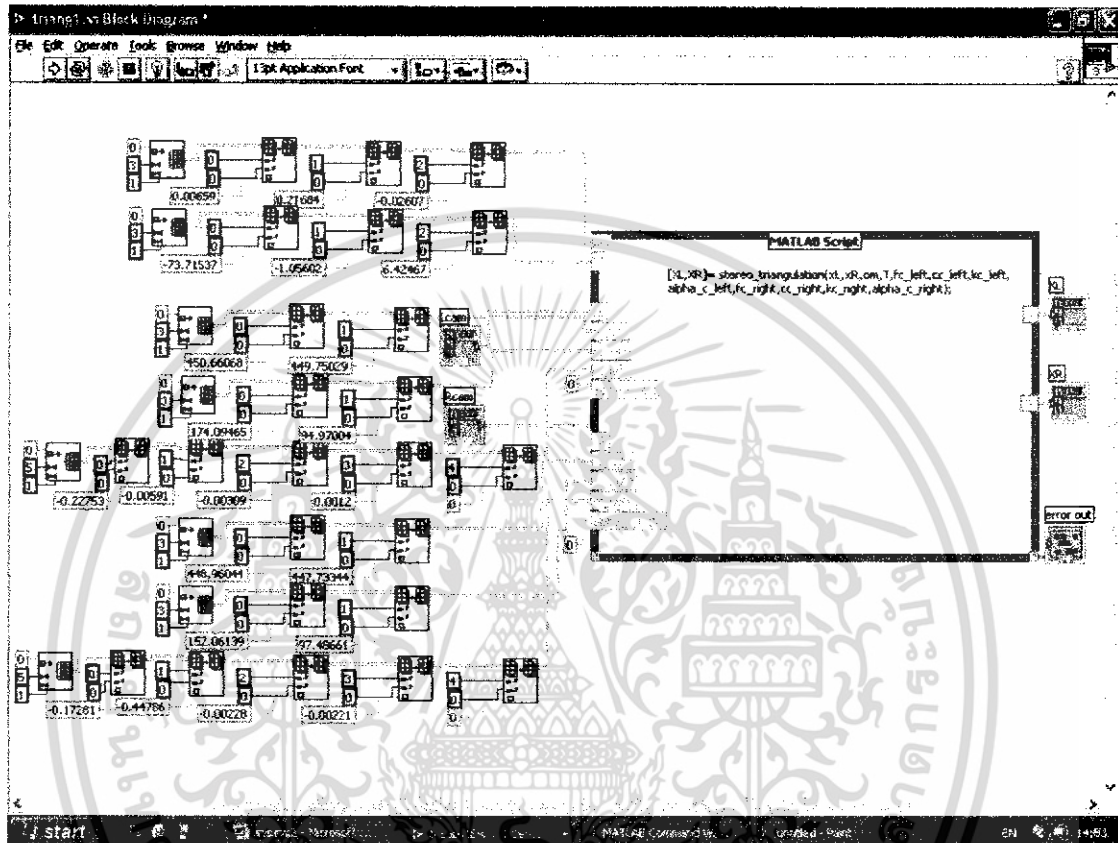
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Learn Template



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Stereo Triangulation



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้