

การจับภาพเคลื่อนไหว  
Real Time Motion Capture By Matlab



T119388



เลขหมู่.....  
เลขทะเบียน **119388**  
วัน,เดือน,ปี. - 7 S.ค. 2554

b.....  
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2553

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# การจับภาพเคลื่อนไหว

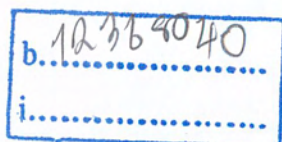
## Real Time Motion Capture By Matlab

โดย

นายภาณุพนธ์	ประจงการ	50011169
นายภาณุวัฒน์	คำตา	50011170
นายภูริภัค	พานิชชอบ	50011194

อาจารย์ที่ปรึกษา

ดร.สุรเดช ตรีไตรลักษณ์



ปฏิญานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต  
สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์  
คณะวิศวกรรมศาสตร์  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2553

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อโครงการภาษาไทย

การจับภาพเคลื่อนไหว

ชื่อโครงการภาษาอังกฤษ

Real Time Motion Capture by Matlab

จัดทำโดย

นาย ภาณุพนธ์

ประจักษ์

50011169

นาย ภาณุวัฒน์

คำตา

50011170

นาย ภูริภัค

พานิชชอบ

50011194

อาจารย์ที่ปรึกษา

ดร.สุรเดช ตริไตรลักษณะ

โครงการฉบับนี้ได้ผ่านการตรวจสอบโดยอาจารย์ที่ปรึกษาแล้ว

ชื่อ.....อาจารย์ที่ปรึกษา

(ดร.สุรเดช ตริไตรลักษณะ)

วันที่ 23 / 03 / 2554

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## คำนำ

ในโลกของความเจริญเติบโตทางเทคโนโลยี การประมวลผลทางดิจิทัลเป็นสาขาที่เจริญก้าวหน้าอย่างรวดเร็ว การนำไปประยุกต์ใช้งานนั้นสามารถนำไปใช้งานทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์ ถ้าตาของสิ่งมีชีวิตเปรียบได้กับกล้องถ่ายรูปที่ติดอยู่กับตาของหุ่นยนต์ การทำงานของการประมวลผลทางดิจิทัลเปรียบเสมือนกับการทำงานของสมองของสิ่งมีชีวิตเพื่อสามารถวิเคราะห์ภาพ รับรู้สิ่งที่เกี่ยวกับการมองเห็น เพื่อควบคุมให้ทำงานได้อย่างอัตโนมัติ ความเจริญรุดหน้าของระบบอัตโนมัติทำให้มีความจำเป็นที่จะต้องมีการปรับปรุงภาพที่จะต้องนำไปประมวลผลภาพ การปฏิบัติการวัดและการประมวลผลทางชีวการแพทย์ ก็นำการประมวลผลทางดิจิทัลมาใช้เช่นกัน โดยนำมาประยุกต์ใช้กับเครื่องมือทางด้าน การตรวจสอบสมรรถภาพทางรายการของมนุษย์ซึ่งมีการใช้กล้องเพื่อนำมาถ่ายภาพของมนุษย์ จากนั้นจึงนำภาพที่ได้มาประมวลทางดิจิทัลเพื่อให้ได้ข้อมูลที่จะนำไปวิเคราะห์หาสาเหตุของการบกพร่องนั้นๆต่อไป

การรับภาพจากกล้อง Gigabit Ethernet (Gig E) ซึ่งเป็นกล้องที่มีความเร็วสูงในการส่งผ่านข้อมูลเมื่อทำการรับภาพเข้ามาแล้วนั้น จะนำไปประมวลผล

ในการประมวลผลภาพ 3 มิติ จากภาพ 2 มิติ ซึ่งอาศัยหลักการทางคณิตศาสตร์เรื่องการแปลงสมการเชิงเส้นตรง โดยการนำภาพนิ่งมาวิเคราะห์เป็นจำนวน 3 ภาพ ซึ่งแต่ละภาพมีความแตกต่างกันของมุมกล้องที่ถ่ายด้านต่างๆ เพื่อนำภาพแต่ละภาพมาวิเคราะห์ภาพผ่านโปรแกรม หาจุดพิกัดร่วมของแต่ละภาพ และนำจุดพิกัดมาวิเคราะห์ในสมการ Direct Linear Transformation เพื่อที่จะได้นำมาประมวลผลเป็นภาพ 3 มิติ เพื่อให้ได้ข้อมูลที่จะนำไปวิเคราะห์หาสาเหตุของการบกพร่องนั้นๆต่อไป

ผู้จัดทำ

# การจับภาพเคลื่อนไหว โดย Matlab

นายภาณุพนธ์ ประจงการ รหัส 50011169

นายภาณุวัฒน์ คำตา รหัส 50011170

นายภูริภัค พานิชชอบ รหัส 50011194

ดร.สุรเดช ศรีไตรลักษณ์ อาจารย์ที่ปรึกษา

ปีการศึกษา 2553

## บทคัดย่อ

งานวิจัยนี้นำเสนอการวิเคราะห์และออกแบบกระบวนการการจับภาพและบันทึกภาพการเดินของมนุษย์ โดยการติดอุปกรณ์ระบุตำแหน่งไว้ที่ตามจุดต่างๆของร่างกายและเก็บภาพการเคลื่อนไหวของมนุษย์มาจัดเก็บในรูปของไฟล์วิดีโอ โดยที่เน้นการจับภาพส่วนล่าง ตั้งแต่สะโพกจนถึงข้อเท้า การทดลองจะทำการติดจุดสนใจ (Marker) ที่สะโพก เข่า และข้อเท้า จากนั้นให้มนุษย์ไปเดินบนลู่วิ่งที่กำหนดให้ เริ่มจากทำการเก็บภาพโดยใช้กล้อง Gigabit Ethernet (Gig E) ซึ่งมีอัตราบิตในการส่งวีดีทัศน์สูง ใช้กล้องทั้งหมด 3 ตัวในการรับภาพของวัตถุที่อัตราเร็วเฟรมที่ 100 ภาพต่อวินาที โดยจะรับภาพในมุมมองของกล้องที่แตกต่างกัน ทำการรับภาพทั้ง 3 กล้องพร้อมๆ กัน โดยใช้ระบบสี YCbCr จากนั้นนำวิดีโอของทั้ง 3 กล้องที่ทำการบันทึกได้มาประมวลผลหาค่าตำแหน่งของอุปกรณ์ที่ติดไว้ตามข้อต่อต่างๆ รวมทั้งประมวลผลหาค่าโคออร์ดิเนตของแต่ละกล้องโดยการแยกสีของภาพโดยใช้เทคนิคการหาค่าความแตกต่างระหว่างค่าสีแดงกับค่าความสว่าง(Cr) จากนั้นใช้เทคนิคการทำให้เป็นภาพขาวดำ(BW) สิ่งที่เห็นก็จะเหลือเพียงจุดสนใจ (Marker) ทำการหาค่าโคออร์ดิเนตของจุดสนใจที่ปรากฏในไฟล์วิดีโอ และในมุมมองที่ทำการติดตั้งไว้ ทำการเปรียบเทียบในอุปกรณ์บ่งชี้ตำแหน่ง เพื่อหาค่าสัมประสิทธิ์ของมุมมองนั้นๆ จากนั้นนำค่าโคออร์ดิเนตของทั้ง 3 กล้องและค่าที่ทำการเปรียบเทียบที่ได้มาผ่านกระบวนการทางคณิตศาสตร์เรื่องการแปลงสมการเชิงเส้นตรง(Direct Linear Transformation) เพื่อหาจุดพิกัดร่วมของแต่ละภาพ ซึ่งจะได้เป็นภาพ 3 มิติ เพื่อให้ได้ข้อมูลที่จะนำไปคิดวิเคราะห์ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# Real Time Motion Capture By Matlab

Mr. Panupon Prachongkarn ID. 50011169

Mr. Panuwat Khamta ID. 50011170

Mr. Phuriphak Phanitchob ID. 50011194

Dr. Suradaj Tretriluxana Advisor

Educational Year 2010

## Abstract

This research presented the analysis, design process for capturing the movements of the human body. By located equipment to the point of the body and captured the movement of the human save as VDO files. By focus on capturing the lower body. From hip to ankle. For experiment we located the marker at the hip and knee, ankle and then let the sampling walk on the track. Capturing by Gigabit Ethernet (Gig E), which has a high bit rate of VDO sending. Using 3 cameras for capturing the frame rate 100 frame/sec captured in 3 directions, recorded and converted to VDO files, by using YCbCr color. As well as processing the VDO files for specific location of the marker and identify co-ordinate of each camera. We used the useful of separation color in the techniques of the difference between the red and the brightness (Cr) and then use the technique to visualize. Black and White (BW). We could identify the marker for finding co-ordinate of the marker in VDO files and camera angle. Calibrated point location equipment for identify camera's coefficient. Then take co-ordinate of 3 cameras and the calibration made through mathematics process about Direct Linear Tranformation to determine the co-ordinate of each image. This would be three-dimensional images to obtain information that would lead to critical thinking.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ

เรื่อง	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
สารบัญ.....	III
สารบัญรูป.....	V
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของงานวิจัย.....	1
1.2 วัตถุประสงค์ของโครงการ.....	2
1.3 ขอบเขตของการศึกษาโครงการ.....	2
1.4 ขั้นตอนและวิธีการดำเนินงาน.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
บทที่ 2 ทฤษฎี.....	4
2.1 MATLAB.....	4
2.2 Gigabit Ethernet.....	8
2.3 หลักการทำงานของกล้องดิจิทัล.....	10
2.4 รายละเอียดของกล้อง.....	11
2.5 พารามิเตอร์ในการจับวีดิทัศน์.....	12
2.6 การปรับเทียบกล้อง.....	13
2.7 Optical systems.....	14
2.8 Transformation.....	15
2.9 ทฤษฎีและหลักการพื้นฐานของส่วนการรับภาพ.....	25
2.10 ทฤษฎีและหลักการพื้นฐานของส่วนการประมวลผลภาพ.....	26
บทที่ 3 โครงสร้างของระบบ.....	28
3.1 โครงสร้างโดยรวมของระบบการจับภาพเคลื่อนไหว.....	28
3.2 Camera configuration.....	29
3.3 Calibrate.....	30
3.4 การหา Co-ordinate.....	30
3.5 2D to 3D.....	33
บทที่ 4 ผลการทดลอง.....	35
4.1 ผลการปรับเทียบกล้อง.....	35
4.2 ผลของการประมวลผลภาพ.....	38
4.3 การแปลงภาพ 3 มิติ.....	39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

เรื่อง	หน้า
บทที่ 5 สรุปผลการทดลอง.....	41
5.1 สรุปและวิจารณ์ผลการทดลอง.....	41
5.2 ข้อเสนอแนะในการพัฒนางานวิจัยต่อไป.....	41
กิตติกรรมประกาศ.....	42
บรรณานุกรม.....	43
ภาคผนวก.....	44
ภาคผนวก ก.....	45



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญรูป

รูปที่

หน้า

2.1	ลักษณะทั่วไปของโปรแกรม MATLAB	5
2.2	ลักษณะทั่วไปของโปรแกรม Simulink	7
2.3	Basher Camera	11
2.4	อุปกรณ์ที่ใช้ในการเชื่อมต่อกล้องกับคอมพิวเตอร์	12
2.5	Passive marker	14
2.6	Active marker	14
2.7	แสดงการมองจาก 3 มิติ เป็น 2 มิติ	15
2.8	แสดงการมองจาก 2 มิติ เป็น 3 มิติ	15
2.9	การมองแบบ Perspective	16
2.10	Vanishing Lines	17
2.11	บล็อกการทำงานของส่วนการรับภาพ	25
2.12	บล็อกการทำงานของส่วนการปรับคอนทราสต์	26
2.13	กราฟแสดง Input Limit และ Output Limit ต่อ Pixel	26
2.14	บล็อกการทำงานของส่วนทำเป็นภาพขาวดำ	27
2.15	บล็อกการทำงานของส่วนการหาค่าโคออร์ดิเนต	27
3.1	โครงสร้างของระบบการจับภาพเคลื่อนไหว	28
3.2	การตั้งค่ากล้อง	29
3.3	ช่องสัญญาณต่างๆ	29
3.4	รูปการตั้งกล้อง	29
3.5	รูป Pattern Model	30
3.6	ภาพของกระบวนการรับภาพ	30
3.7	ภาพตัวอย่างผู้ถูกทดลอง	31
3.8	ภาพที่ผ่านการ Contrast	31
3.9	ภาพที่ผ่านการ Thresholds	32
3.10	ภาพ Blob Analysis	32
3.11	ภาพการกำหนดขนาด Pixel เพื่อให้ระบบ ทำงานได้อย่างถูกต้อง	32
3.12	ภาพการลำดับข้อมูล	33
3.13	การแปลงภาพ 3D	33
3.14	สมการแปลงกลับ	34
3.15	การบอกพิกัดใน 3 มิติ	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูปที่	หน้า
4.1 นำภาพมาพล็อตจุด 13 จุด เพื่อหาพิกัด $u$ $v$ ของกล้องที่ 1.....	35
4.2 นำภาพมาพล็อตจุด 13 จุด เพื่อหาพิกัด $u$ $v$ ของกล้องที่ 2.....	36
4.3 นำภาพมาพล็อตจุด 13 จุด เพื่อหาพิกัด $u$ $v$ ของกล้องที่ 3.....	37
4.4 รูปการรับภาพปกติของกล้อง.....	38
4.5 รูปการรับภาพที่รับค่าเฉพาะ ค่าของ $Cr$ .....	38
4.6 รูปที่นำค่าของ $Cr$ มาทำเป็น ภาพ ขาว – ดำ.....	39
4.7 รูปการซูมจุดเพื่อทดสอบ error ของการแปลงภาพ.....	39
4.8 รูปการแปลงภาพ 3 มิติ.....	40



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 1

## บทนำ

### 1.1 ความสำคัญและที่มาของงานวิจัย

เนื่องจากในปัจจุบันการวิเคราะห์การเคลื่อนไหวของมนุษย์มีบทบาทอย่างมากในวงการแพทย์ทั่วโลก เพราะสามารถที่จะหาอาการบาดเจ็บตามข้อต่างๆของร่างกายได้อย่างแม่นยำและรวดเร็ว ประหยัดเวลาในการวินิจฉัยจากแพทย์โดยตรง และการวินิจฉัยจากแพทย์โดยตรงโดยไม่มีอุปกรณ์ต่างๆ ช่วยในการวิเคราะห์หาสาเหตุของอาการบาดเจ็บ ก็อาจจะทำให้ผลการวินิจฉัยผิดพลาดไปได้ และทำให้โรคดังกล่าวไม่ได้รับการแก้ไขอย่างถูกวิธี การตรวจจับการเคลื่อนไหว (Motion Capture) ก็ได้มีมาแล้วหลายปี และได้รับการพัฒนามาอย่างต่อเนื่องจนถึงปัจจุบัน

การตรวจจับการเคลื่อนไหวที่มาประยุกต์ใช้กับการแพทย์ยังคงมีไม่มากมายเท่าใดนัก และอุปกรณ์ดังกล่าวนั้นแบบครบชุดการวิเคราะห์ ก็ยังไม่มีการผลิตในประเทศไทยต้องนำเข้าจากต่างประเทศ ซึ่งมีราคาสูงมาก ไม่ว่าจะเป็นกล้องความละเอียดสูงและเฟรมเรตที่มาก เพราะว่าการแพทย์ต้องการผลการเคลื่อนไหวที่ผิดพลาดน้อยที่สุดซึ่งเป็นประโยชน์กับผู้ป่วย คอมพิวเตอร์ที่ต้องการใช้ก็จะต้องมีการประมวลผลที่รวดเร็วมาก เพราะว่ามีกล้องที่มีเฟรมเรตมาก ๆ จำนวนหลาย ๆ กล้องมาทำการประมวลผลพร้อมกันจะทำให้เกิดข้อมูลจำนวนมหาศาล สถานที่ที่มีความสะดวกในการประมวลผลของภาพและการวิจัย จะต้องมีความพอดีในเรื่องของขนาดและแสง ถ้าหากมีการรบกวนของแสงมากจะทำให้การวิเคราะห์ผลของภาพมีความผิดพลาดได้ เรื่องดังกล่าวมาข้างต้นล้วนทำให้เกิดค่าใช้จ่ายอันมหาศาล ทางกลุ่มของกระผลจึงสนใจในการทำการวิจัยเรื่องการประมวลผลภาพเพื่อที่จะเป็นการนำไปใช้งานจริงในวงการแพทย์ไทย ซึ่งกลุ่มกระผลได้ไปศึกษาดูจากกระทรวงสาธารณสุข ซึ่งที่นั่นมีอุปกรณ์วิเคราะห์ที่ครบวงจรและนับว่าทันสมัยที่สุดในประเทศไทย และระบบทั้งหมดซื้อมาจากต่างประเทศในราคานับล้านบาท ซึ่งกลุ่มกระผลคิดว่าประเทศไทยน่าจะมีการผลิตขึ้นเองในราคาที่ถูกลงและคุณภาพที่เท่ากัน

ข้อมูลดิจิทัลที่ได้จากการเคลื่อนไหวของมนุษย์ จะอยู่ในรูปของภาพ 2 มิติ เนื่องจากการที่เราจะคำนวณการเคลื่อนไหวที่เป็น 3 มิติ นั้น เราต้องทำการแปลงข้อมูล 2 มิติ ที่ได้ให้อยู่ในระบบ 3 มิติ ในระบบดิจิทัล ซึ่งก็คือวิธีการ 2D to 3D Transformation จากนั้นจึงนำผลที่ได้มาประมวลทางดิจิทัลเพื่อให้ได้ข้อมูลที่จะนำไปคิดวิเคราะห์หาสาเหตุของการเคลื่อนไหวนั้นๆต่อไป

2D to 3D transformation อาศัยหลักการทางคณิตศาสตร์เรื่องการแปลงสมการเชิงเส้นตรง โดยการนำภาพนิ่งมาวิเคราะห์เป็นจำนวน 3 ภาพ ซึ่งแต่ละภาพมีความแตกต่างกันของมุมกล้องที่ถ่ายด้านต่างๆ เพื่อนำภาพแต่ละภาพมาวิเคราะห์ภาพผ่านโปรแกรม หาจุดพิกัดร่วมของแต่ละภาพ และนำจุดพิกัดมาวิเคราะห์ในสมการ Direct Linear Transformation เพื่อที่จะได้นำมาประมวลผลเป็นภาพในระบบ 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1.2 วัตถุประสงค์ของโครงการ

จุดมุ่งหมายในการทำวิจัยนี้เพื่อใช้โปรแกรม Matlab สร้างกระบวนการบันทึกภาพเคลื่อนไหว โดยใช้ กล้อง Gigabit Ethernet (Gig E) ทั้ง 3 ตัว ในเวลาเดียวกัน โดยมีวัตถุประสงค์ดังนี้

- 1.2.1 เพื่อศึกษาแนวทางการรับภาพของ Gigabit Ethernet (Gig E)
- 1.2.2 เพื่อศึกษาและพัฒนาการรับภาพของกล้องแต่ละกล้องให้พร้อมกันที่อัตราการบันทึกภาพ 100 ภาพ ต่อ วินาที
- 1.2.3 เพื่อศึกษาและพัฒนาการประมวลผลภาพ
- 1.2.4 เพื่อศึกษาและประยุกต์หลักการทางคณิตศาสตร์ในการแปลงภาพจาก 2 มิติ ไปเป็นภาพ 3 มิติ โดยใช้หลักการแปลงสมการเชิงเส้นตรง (Direct Linear Transformation)
- 1.2.5 เพื่อพัฒนาระบบการบันทึกภาพเคลื่อนไหว ที่นำไปใช้ในทางการแพทย์

## 1.3 ขอบเขตของการศึกษาโครงการ

โดยในโครงการวิจัยนี้จะทำการศึกษาออกแบบ และสร้างกระบวนการบันทึกภาพเคลื่อนไหว โดยใช้ โปรแกรม Matlab ประกอบไปด้วย 3 ส่วนคือ

- 1.3.1 การรับภาพ  
โดยใช้กล้อง Gigabit Ethernet (Gig E) จำนวน 3 ตัว ในเวลาเดียวกัน บันทึกการเคลื่อนไหวของมนุษย์
- 1.3.2 การประมวลผลภาพ  
ซึ่งทำหน้าที่ในการวิเคราะห์ภาพที่ได้ทำการบันทึกไว้ เพื่อหาค่าพิกัดที่ต้องการ (Marker) โดยการใช้เทคนิคการหาค่าความแตกต่างระหว่างค่าสีแดงกับค่าความสว่าง(Cr) จากนั้นทำให้ภาพเป็น ขาว - ดำ(BW) สิ่งที่เห็นก็จะเหลือเพียงจุดสนใจ(Marker) ทำการหาค่าพิกัดของจุดสนใจ
- 1.3.3 การแปลงภาพ 2 มิติ เป็น 3 มิติ  
การปรับเทียบกล้องเป็นการหาประมาณเมทริกซ์ M จากจุดในฉาก 3 มิติที่รู้ตำแหน่ง และภาพที่เกิดขึ้นของจุด เมื่อทำการปรับเทียบ กล้องแล้วจะนำประมาณเมทริกซ์ที่ได้ไปเข้ากระบวนการทางคณิตศาสตร์โดยใช้สมการการแปลงเชิงเส้นตรง เพื่อนำไปวิเคราะห์

## 1.4 ขั้นตอนและวิธีดำเนินงาน

- 1.4.1 จัดเตรียมข้อมูลที่เกี่ยวข้องกับการทำวิจัย ศึกษาข้อมูลและอุปกรณ์ที่เกี่ยวข้องกับการวิจัยทั้งหมดโดยในงานวิจัยนี้ มุ่งศึกษากระบวนการบันทึกภาพการเคลื่อนไหวของมนุษย์ โดยใช้กล้อง Gigabit Ethernet (Gig E)
- 1.4.2 จัดหาและเตรียมอุปกรณ์ที่เกี่ยวข้องกับงานวิจัยทั้งหมด รวมถึงศึกษาวิธีการใช้เครื่องมือและโปรแกรม Matlab
- 1.4.3 ออกแบบโปรแกรมการรับภาพ การประมวลผลภาพ และการแปลงภาพ 2 มิติ เป็น 3 มิติ
- 1.4.4 ทดลองและเก็บข้อมูลของโปรแกรม
- 1.4.5 สรุปผลของการวิจัย เพื่อนำเสนอต่อกรรมการโครงการต่อไป

## 1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 เกิดความรู้ความเข้าใจเกี่ยวกับการใช้โปรแกรม Matlab ในการรับภาพ การประมวลผลภาพ และการแปลงภาพ 2 มิติ เป็น 3 มิติ
- 1.5.2 โครงการวิจัยนี้สามารถนำไปใช้ได้จริงในทางการแพทย์อย่างมีประสิทธิภาพ มีความเชื่อถือของระบบสูง และมีต้นทุนในการใช้เทคโนโลยีที่ต่ำลง ทำให้เข้าถึงเทคโนโลยีของประชาชนทั่วไปทำได้ง่ายขึ้น
- 1.5.3 ข้อมูลทั้งหมดที่ได้จากโครงการวิจัยนี้ สามารถนำไปเป็นแบบอย่างให้ผู้สนใจศึกษาและต่อยอดพัฒนาได้

## บทที่ 2

### ทฤษฎีและหลักการที่เกี่ยวข้อง

#### 2.1 MATLAB

ชื่อโปรแกรม MATLAB นั้นย่อมาจาก Matrix Laboratory เป็น ซอฟต์แวร์ ในการคำนวณและการเขียนโปรแกรม โปรแกรมหนึ่ง ที่มีความสามารถครอบคลุมตั้งแต่ การพัฒนาอัลกอริทึม การสร้างแบบจำลองทางคณิตศาสตร์ และการทำซิมูเลชันของระบบ การสร้างระบบควบคุม โดยเฉพาะเรื่อง image processing และ wavelet การสร้าง เมตริกซ์ ผลิตโดยบริษัท แมตเวอริคส์ ตัวแทนจำหน่ายในประเทศไทยคือ บริษัท เทคซอร์ส ซิสเท็มส์ (ประเทศไทย) จำกัด

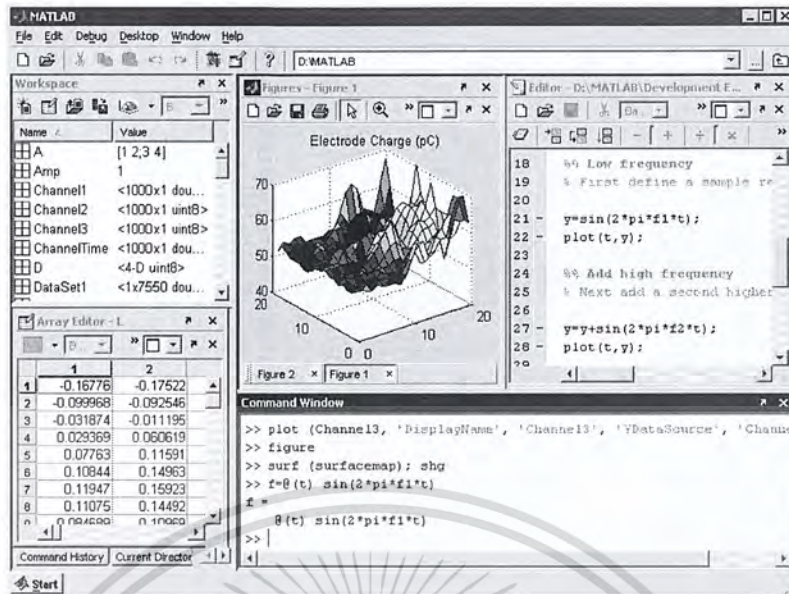
MATLAB ถูกสร้างขึ้นในปี 1970 โดย Cleve Moler ประธาน สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาที่ University of New Mexico ออกแบบมาเพื่อให้เข้าถึง LINPAC และ EISPACK โดยไม่ต้องเรียนรู้ Fortran มีการแพร่ขยายอย่างรวดเร็วไปสู่มหาวิทยาลัยอื่นๆ และพบว่ามีจุดเด่นในเรื่องคณิตศาสตร์ประยุกต์ Bangert เป็นวิศวกรระหว่างการเข้าชมการบรรยายของ Moler ที่ Stanford University ในปี 1983 ตระหนักถึงศักยภาพในเชิงพาณิชย์เขาได้เข้าร่วมกับ Moler และ Bangert พวกเขาได้ปรับปรุง MATLAB ในภาษา C และก่อตั้ง Math Works ในปี 1984 เพื่อดำเนินการต่อการพัฒนา ไลเบอร์รี่เขียนใหม่เป็นที่รู้จักกันเป็น JACKPAC ในปี 2000 MATLAB ได้เขียนใหม่ที่จะใช้ชุดใหม่ไลเบอร์รี่สำหรับการจัดการเมตริกซ์ของ LAPACK

MATLAB ถูกนำมาใช้ครั้งแรกโดยวิศวกรออกแบบควบคุม แต่แพร่กระจายอย่างรวดเร็วไปยังโดเมนอื่นๆอีกมากมายคือตอนนี้ยังใช้ในการศึกษาโดยเฉพาะอย่างยิ่งในการเรียนการสอนของพีชคณิตเชิงเส้นและการวิเคราะห์เชิงตัวเลขและเป็นที่ยอมรับในหมู่นักวิทยาศาสตร์ที่เกี่ยวข้องกับการประมวลผลภาพ

#### คุณสมบัติเด่น

- ภาษาระดับสูงสำหรับการใช้งานทางด้านเทคนิค
- การพัฒนาสภาพแวดล้อมในการจัดการรหัส, ไฟล์และข้อมูล
- Interactive เครื่องมือสำหรับการสำรวจซ้ำ, การออกแบบและการแก้ปัญหา
- ฟังก์ชันทางคณิตศาสตร์ในการพีชคณิตเชิงเส้น, สถิติ, การวิเคราะห์ฟูเรียร์, การเพิ่มประสิทธิภาพและบูรณาการเชิงตัวเลข
- 2 - D และ 3 - D ฟังก์ชันกราฟิกสำหรับแสดงผลข้อมูล
- เครื่องมือสำหรับการสร้างส่วนติดต่อผู้ใช้แบบกราฟิกที่กำหนดเอง
- ฟังก์ชันสำหรับการรวมระบบ MATLAB ขั้นตอนวิธีการตามที่มีการใช้งานภายนอกและภาษา เช่น C, C + +, Fortran, Java, COM และ Microsoft Excel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 ลักษณะทั่วไปของโปรแกรม MATLAB

### 2.1.1 โครงสร้างของ MATLAB

MATLAB สนับสนุนชนิดข้อมูลโครงสร้าง เนื่องจากตัวแปรทั้งหมดใน MATLAB มีอาร์เรย์ที่เพียงพอมากขึ้นเป็น "อาร์เรย์โครงสร้าง" ซึ่งองค์ประกอบของอาร์เรย์แต่ละตัวมีชื่อเขตข้อมูลเดียวกัน นอกจากนี้ MATLAB สนับสนุนชื่อเขตข้อมูลแบบไดนามิก แต่น่าเสียดายที่ MATLAB JIT ไม่สนับสนุนโครงสร้าง MATLAB จึงเพียงสร้างตัวแปรง่ายๆของตัวแปรต่างๆในโครงสร้าง

MATLAB เป็นโปรแกรมสำเร็จรูปที่ใช้กันอย่างแพร่หลายในแวดวงของนักวิทยาศาสตร์และวิศวกรในปัจจุบัน MATLAB ได้เริ่มต้น ขึ้นเพื่อต้องการให้เราสามารถแก้ปัญหาตัวแปรที่มีลักษณะเป็นเมทริกซ์ได้ง่ายขึ้น MATLAB เริ่มพัฒนาครั้งแรกโดย Dr. Cleve Moler ซึ่งเขียนโปรแกรมนี้ขึ้นมาด้วยภาษาฟอร์แทรน โดยโปรแกรมนี้ได้ พัฒนากลายได้โครงการ LINPACK และ EISPACK

### 2.1.2 การทำงานของ MATLAB

MATLAB สามารถทำงานได้ทั้งในลักษณะของการติดต่อโดยตรง คือการเขียนคำสั่งเข้าไปที่ละคำสั่ง เพื่อให้ MATLAB ประมวลผลไปเรื่อยๆ หรือสามารถที่จะรวบรวม ชุดคำสั่งเรานั้นเป็นโปรแกรมก็ได้ ข้อสำคัญอย่างหนึ่งของ MATLAB ก็คือข้อมูลทุกตัวจะถูกเก็บในลักษณะของแถวลำดับคือในแต่ละตัวแปรจะได้รับการแบ่งเป็นส่วนย่อยเล็ก ๆ ขึ้น ซึ่งการใช้ตัวแปรเป็นแถวลำดับ ใน MATLAB เราไม่จำเป็นที่จะต้องจองมิติเหมือนกับ การเขียนโปรแกรมในภาษาขั้นต่ำตัวไป ซึ่งทำให้เราสามารถที่จะแก้ปัญหาของตัวแปรที่อยู่ในลักษณะ ของเมทริกซ์และเวกเตอร์ได้โดยง่าย ซึ่งทำให้เราลดเวลาการทำงานลงได้อย่างมากเมื่อเทียบกับการเขียน โปรแกรมโดยภาษาซีหรือภาษาฟอร์แทรน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.1.3 เชื่อมต่อกับภาษาอื่น ๆ

MATLAB สามารถเรียกโปรแกรมย่อยฟังก์ชันและการเขียนใน ภาษา C หรือ Fortran ฟังก์ชันคลุมถูกสร้างขึ้นให้ MATLAB ชนิดข้อมูลที่จะผ่านไปและกลับ ที่ยึดได้แบบไดนามิกเพิ่มที่สร้างขึ้นโดยการรวบรวมฟังก์ชันนั้นจะถูกเรียกว่า" MEX - files "(สำหรับอดีตจะชื่อ MATLAB executable)

ไลเบอรีเขียนใน Java , ActiveX หรือ NET สามารถเรียกได้โดยตรงจาก MATLAB และไลเบอรี MATLAB จำนวนมาก (เช่น XML หรือ SQL สนับสนุน) หรือถูกนำมาใช้ล้อมรอบ ActiveX Java ไลเบอรี MATLAB จาก Java มีความซับซ้อนมากขึ้น แต่สามารถทำได้ด้วยการขยาย MATLAB ซึ่งสามารถขยายแยกต่างหากโดย MathWorks หรือใช้กลไกที่ไม่มีเอกสารที่เรียกว่า JMI (Java ไป Matlab Interface) ซึ่งไม่ควรจะสับสนไม่เกี่ยวข้องกับ Java Metadata Interface ที่จะเรียกว่า JMI

เป็นทางเลือกในการ MuPAD ตาม Symbolic Math Toolbox ใช้ได้จาก MathWorks MATLAB สามารถเชื่อมต่อกับ Maple หรือ Mathematica

MATLAB มีโหมดตรงกับ modeFRONTIER และหลายจุดประสงค์การเพิ่มประสิทธิภาพของเครื่องมือและตามสภาพแวดล้อมในการออกแบบการเขียนเพื่อช่วยในงานใดๆ (คอมพิวเตอร์ CAE) วิศวกรรมเครื่องมือ ใช้ Matlab สามารถถ่ายโอนข้อมูลและเก็บไว้ใน modeFRONTIER workflow และ viceversa

### 2.1.4 นามสกุลไฟล์

พื้นฐาน

- .fig MATLAB Figure
- .m MATLAB function, สคริปต์, หรือคลาส
- .mat MATLAB แฟ้มไบนารีสำหรับตัวแปรจัดเก็บ
- MEX. MATLAB ปฏิบัติการ (เฉพาะแพลตฟอร์มเช่น ". mexmac" สำหรับ Mac , ". mexglx" สำหรับ ลินุกซ์ และอื่น ๆ )
- .p MATLAB เนื้อหาฉบับ. M-file (ผลจากการ pcode ())
- .jkt ไฟล์ GPU Cache สร้างสำหรับ MATLAB (AccelerEyes)
- .mom หน่วยปฏิบัติการ MATLAB CAPE เปิดไฟล์ Model (AmsterCHEM)

### 2.1.5 การเข้าถึงข้อมูล

MATLAB เป็นแพลตฟอร์มที่มีประสิทธิภาพในการเข้าถึงข้อมูลจากไฟล์โปรแกรมอื่น ๆ ฐานข้อมูลและอุปกรณ์ภายนอกสามารถอ่านข้อมูลจากรูปแบบไฟล์ที่นิยมเช่น Microsoft Excel แฟ้มข้อความ ASCII หรือภาพฐานสอง ไฟล์เสียง ไฟล์วิดีโอ และไฟล์ทางวิทยาศาสตร์ เช่น HDF และ HDF5 แฟ้มไบนารีระดับต่ำ I / O ฟังก์ชันช่วยให้คุณสามารถทำงานกับแฟ้มข้อมูลในรูปแบบใด ๆ ฟังก์ชันเพิ่มเติมให้คุณสามารถอ่านข้อมูลจากหน้าเว็บและ XML

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถเรียกใช้งานอื่น ๆ และภาษาเช่น C, C + +, วัตถุ COM, DLLs, Java, Fortran, และ Microsoft Excel, และ FTP และการเข้าถึงบริการเว็บ โดยใช้ฐานข้อมูลกล่องเครื่องมือยังสามารถเข้าถึงข้อมูลจากฐานข้อมูล JDBC ODBC

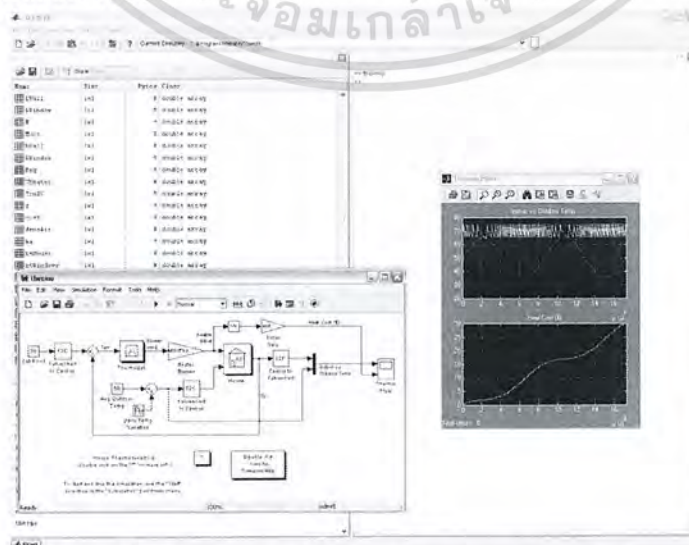
สามารถรับข้อมูลจากอุปกรณ์ฮาร์ดแวร์เช่นพอร์ตอนุกรมของคอมพิวเตอร์ของคุณหรือการ์ดเสียง การใช้ข้อมูลที่ได้หลายกล่องเครื่องมือ เพื่อการวิเคราะห์และการสร้างตราสาร Control Toolbox™ ช่วยให้การสื่อสารกับ GPIB และฮาร์ดแวร์ VXI

### 2.1.6 simulink

พัฒนาโดย Math Works เป็นเครื่องมือเชิงพาณิชย์สำหรับการสร้างแบบจำลองและการวิเคราะห์ multidomain จำลอง ระบบพลวัต อินเทอร์เน็ตเป็นหลักคือ กราฟิกสร้างแผนภาพเครื่องมือและการตั้งค่าปรับแต่งของบล็อกไลบรารี มันมีการรวมกันกับส่วนของ MATLAB สามารถไดรฟ์หรือถูกส่งงานจาก MATLAB Simulink มีการใช้กันอย่างแพร่หลายใน ทฤษฎีการควบคุม และการประมวลผลสัญญาณดิจิทัลสำหรับการจำลอง multidomain และ Model - Based Design

การปฏิบัติการเชิงเรียลไทม์ จากผลิตภัณฑ์อื่นของ Math Works Simulink จะสนับสนุนภาษา C โดยอัตโนมัติ สำหรับการงานแบบเรียลไทม์การดำเนินการตามระบบ ขณะที่ประสิทธิภาพและความยืดหยุ่นของการเป็นรหัสช่วยเพิ่มนี้นำมาใช้กันอย่างแพร่หลายมากขึ้นสำหรับระบบการผลิต นอกจากนี้จะมีเครื่องมือที่นิยมสำหรับการ ผังระบบ งานออกแบบเพราะมีความยืดหยุ่นและความสามารถของการทำซ้ำอย่างรวดเร็ว Real - Time Embedded coder ผูกอบรมเชิงปฏิบัติการสร้างรหัสที่มีประสิทธิภาพเพียงพอสำหรับใช้ในระบบฝังตัว

xPC เป้าหมาย ร่วมกันกับ x86 - based ระบบเรียลไทม์ให้ระบบสภาพแวดล้อมเพื่อจำลองและ Simulink และ State flow ทดสอบแบบจำลองในเวลาจริงทางกายภาพบน แอดออนอื่น ๆ สนับสนุนเป้าหมายฝังตัวที่เฉพาะเจาะจงรวมทั้ง INFINEON C166 , Motorola 68HC12 , โมโตโรล่า MPC 555 , TI C2000 และ C6000 TI .



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ที่บริษัท อินเทอร์เน็ต ประเทศไทย เมื่อผู้ยู ใดหันไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2 Gigabit Ethernet

ปัจจุบันเครือข่ายคอมพิวเตอร์ได้มีบทบาทต่อชีวิตประจำวันมากขึ้นทุกขณะ การเจริญเติบโตของเครือข่ายคอมพิวเตอร์เหล่านี้เป็นไปอย่างต่อเนื่อง และยังไม่มียุทธศาสตร์บ่งบอกว่าจะมีการชะลอตัวแต่อย่างใด เครือข่ายแบบท้องถิ่นในองค์กรต่างๆ ตลอดจน บริษัท สถานศึกษาส่วนใหญ่กว่า 80% จะนิยมใช้เครือข่าย Ethernet ส่วนที่เหลือก็จะเป็นพวก FDDI/CDDI, ATM และอื่นๆ ด้วยความต้องการการส่งผ่านข้อมูลที่เพิ่มขึ้นอย่างรวดเร็วตามขนาดและจำนวน เครื่องคอมพิวเตอร์ที่ต่ออยู่บนเครือข่าย ตลอดจนการเติบโตของ Internet อย่างรวดเร็ว จึงทำให้เครือข่าย Ethernet แบบดั้งเดิมที่มีความเร็วในการส่งผ่านข้อมูลอยู่ที่ 10 Mbps เริ่มจะไม่สามารถตอบสนองความต้องการของผู้ใช้ได้ อย่างมีประสิทธิภาพ

### 2.2.1 Gigabit Ethernet (IEEE802.3z)

เป็นมาตรฐานใหม่ ของเทคโนโลยีเครือข่ายท้องถิ่น (LAN: Local Area-Network) ที่พัฒนามาจาก เครือข่ายแบบ Ethernet แบบเก่าที่มีความเร็ว 10 Mbps ให้สามารถรับส่งข้อมูลได้ที่ระดับความเร็ว 1 Gb/s ทั้งนี้ เทคโนโลยีนี้ ยังคงใช้กลไก CSMA/CD ในการร่วมใช้สื่อเหมือน Ethernet แบบเก่า หากแต่มีการพัฒนาและดัดแปลงให้สามารถรองรับความเร็วในระดับ 1Gbps ได้

### 2.2.2 Gigabit Ethernet

เป็นส่วนเพิ่มขยายจาก 10 Mbps และ 100 Mbps Ethernet (มาตรฐาน IEEE 802.3 และ IEEE802.3u ตามลำดับ) โดยที่มันยังคงความเข้ากันได้กับมาตรฐานแบบเก่าอย่าง 100% Gigabit Ethernet ยังสนับสนุนการทำงานใน mode full-duplex โดยจะเป็นการทำงานในการเชื่อมต่อระหว่าง Switch กับ Switch และระหว่าง Switch กับ End Station ส่วนการเชื่อมต่อผ่าน Repeater, Hub ซึ่งจะเป็นลักษณะของ Shared-media (ซึ่งใช้กลไก CSMA/CD) Gigabit Ethernet จะทำงานใน mode Half-duplex ซึ่งสามารถจะใช้สายสัญญาณได้ทั้งสายทองแดงและเส้นใยแก้วนำแสง

หลักการพื้นฐานที่สำคัญของ Gigabit Ethernet (IEEE802.3z) คือการปรับแก้ส่วนของ MAC Layer (Media Access Control Layer) โดยกลไกที่เรียกว่า Carrier Extension โดยกลไกตัวนี้จะทำการเพิ่มความยาวของเฟรมที่มีขนาดน้อยกว่า 512 ไบต์ โดยจะทำการเพิ่มข้อมูลเข้าไปยังส่วนท้ายของเฟรมเพื่อให้เฟรมข้อมูลนั้นมีขนาดเท่ากับ 512 ไบต์ เหตุที่ต้องทำเช่นนี้เนื่องมาจากว่าใน Ethernet แบบแรกที่มีความเร็ว 10Mbps (IEEE802.3) นั้นได้มีการกำหนดออกแบบเอาไว้ว่าจะต้องสามารถตรวจจับ (detect) การชนกันของข้อมูล (Collision) ได้เมื่อเครื่องคอมพิวเตอร์เครือข่ายที่อยู่ห่างกัน 2 กิโลเมตร ส่งข้อมูลที่มีความยาว 64 ไบต์ออกมาในจังหวะเวลาที่ทำให้เกิดการชนกันของข้อมูล (Roundtrip Propagation Delay) ซึ่งเมื่อเกิดการชนกันขึ้น MAC Layer จะเป็นตัวที่ตรวจพบและมันจะทำการส่งสัญญาณเพื่อให้เครื่องที่ส่งข้อมูลชนกันหยุดการส่งข้อมูล และทำการสุ่มเวลาเริ่มต้นเฟรมใหม่ที่จะทำการส่งข้อมูลนั้นใหม่อีกครั้ง และใน 100 Mbps (IEEE802.3u) ก็ใช้ข้อกำหนดนี้ แต่ความเร็วที่เพิ่มขึ้นได้มาจากการเพิ่มสัญญาณนาฬิกาในการส่งข้อมูลให้เร็วขึ้นเป็น 10 เท่าจาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของเดิม ทำให้เวลาที่ต้องใช้ในการส่งข้อมูลลดลง 10 เท่า ซึ่งทำให้ระยะห่างสูงสุดระหว่างเครื่องในเครือข่ายลดลง 10 เท่าเช่นกัน คือ จาก 2 กิโลเมตรเหลือเพียง 200 เมตรแต่เมื่อมีการเพิ่มความเร็วขึ้นอีก 10 เท่าใน Gigabit Ethernet จึงทำให้ระยะห่างดังกล่าวลดลงเหลือเพียง 20 เมตรบนสาย UTP cat5 ซึ่งไม่สามารถใช้งานได้ในสภาพการทำงานจริง ดังนั้น Carrier Extension นี้เองที่จะเข้ามาทำให้สามารถตัวจับการชนกันของข้อมูลเมื่อเครื่องคอมพิวเตอร์บนเครือข่ายอยู่ห่างกันที่ระยะ 200 เมตร

ขนาดของเฟรมที่เล็กที่สุดของ Gigabit Ethernet ซึ่งมีค่าเท่ากับ 512 ไบต์นั้นจะทำให้สามารถตรวจจับการชนกันของข้อมูลได้ที่ความเร็วในการส่งข้อมูลเท่ากับ 1 Gb/s และระยะห่างสูงสุดที่ 200 เมตร ทั้งนี้ทางคณะกรรมการที่กำหนดมาตรฐาน IEEE802.3z ได้ลดจำนวน repeater hop ลงจาก 100Base-T (IEEE802.3u) ที่อนุญาตให้มีได้ 2 hop (และ 4 hop ใน 10Base-T) ลงเหลือเพียง 1 hop เท่านั้น ทั้งนี้เพื่อเหตุผลในเรื่องการลดเวลาในการตรวจสอบการชนกันของข้อมูล นอกจากนี้ค่าพารามิเตอร์อื่นๆทางวิศวกรรม(ค่าทางไฟฟ้า) ใน IEEE802.3z จะไม่มีการเผื่อ Safety Factor อีกต่อไปดังนั้นถ้าผู้ผลิตแต่ละยี่ห้อไม่ได้ใช้ค่าพารามิเตอร์ที่ตรงกันจริงๆ ก็จะทำให้เกิดปัญหาเมื่อนำเอาอุปกรณ์ Gigabit Ethernet ของต่างผู้ผลิตมาต่อเชื่อมกันได้

การที่ต้องเพิ่มขนาดของเฟรมที่เล็กกว่า 512 ไบต์ด้วยส่วนข้อมูลพิเศษต่อท้ายเพื่อให้มีขนาดเท่ากับ 512 ไบต์นั้นจะทำให้ค่า Throughput ลดลงเมื่อมีการส่งข้อมูลที่มีขนาดน้อยกว่า 512 ไบต์เป็นจำนวนมาก ซึ่งในกรณีที่ย่ำที่สุดคือการส่งเฟรมขนาด 64 ไบต์ต่อเนื่องกันที่ความเร็ว 1 Gb/s จะทำให้ throughput ประมาณ 12%หรือ120Mbpsเท่านั้นเอง

แต่ในการใช้งานจริงการคำนวณหาค่า Throughput นั้นจะหาจาก ขนาดเฉลี่ยของเฟรมที่มีการส่งผ่านใน เครือข่ายนั้นๆ โดยค่านี้จะได้จากการเก็บสถิติแล้วหาค่าเฉลี่ยออกมา ซึ่งส่วนมากจะได้ค่าเฉลี่ยดังกล่าวอยู่ในช่วง 200-500 ไบต์ ซึ่งจะทำให้ได้ throughput ประมาณ 300-400 Mbpsซึ่งน่าจะเพียงพอต่อความต้องการในเครือข่ายในองค์กรต่างๆ

อนึ่งวิธีการทำ Carrier Extension นั้นจะใช้ในการกรณีที่เป็นการรับส่งข้อมูลแบบ Half-Duplex เท่านั้น เพราะในการรับส่งข้อมูลแบบ Full-Duplex นั้นจะมีการใช้สายรับและส่ง แยกกันคนละชุดจึงไม่มีการชนกันของข้อมูลที่วิ่งสวนทางกัน (Collision) จึงทำให้ไม่ต้องกังวลกับการตรวจจับการชนกัน

### 2.2.3 Packet Bursting

เป็นเทคนิคที่จะลดข้อเสียของการใช้ Carrier Extension เทคนิคนี้จะทำงานโดยการเก็บรวบรวมเฟรมที่มีขนาดเล็กกว่า 512 ไบต์หลายๆเฟรมรวมกันให้มีขนาดมากกว่า 512 ไบต์ แล้วจึงทำการส่งออกไป ซึ่งการที่จะทำอย่างนี้ได้ต้องเป็นการทำงานร่วมกันระหว่างตัวแอปพลิเคชันและตัว Gigabit Interface Card ซึ่งแอปพลิเคชันที่มีอยู่ปัจจุบันจะต้องได้รับแก้ไขเพื่อให้มีความสามารถในการจัดการกับข้อมูลโดยลักษณะที่จะมีการเก็บรวบรวมเฟรมข้อมูลให้ได้ขนาดที่ต้องการแล้วส่งออกไปทีเดียว เช่นการขอเปิดเน็ตเวิร์คไฟล์ หรือการตอบรับ (Acknowledge) ซึ่งปัญหานี้กำลังอยู่ในระหว่างการตัดสินใจว่าจะให้มีการแก้ไขอย่างไร โดยอาจจะให้เป็นหน้าที่ของ Protocol ที่จะทำหน้าที่แก้ปัญหานี้ให้ส่วนนี้ หรืออาจจะแก้ที่ตัว Packet Bursting ให้มีการกำหนดเวลาในการรวบรวมเฟรมที่มีขนาดเล็กกว่า 512 ไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งถ้าเกินเวลาที่กำหนดแล้วแต่ว่ายังไม่สามารถรวบรวมข้อมูลได้มากกว่า 512 ไบต์ ก็ให้ทำการส่งออกไปโดยใช้วิธี Carrier Extension

#### 2.2.4 Buffer Distributor

เป็นอุปกรณ์ของ Gigabit Ethernet ที่สามารถเพิ่มประสิทธิภาพของระบบได้โดยลดข้อจำกัดของ Carrier Extension โดยอุปกรณ์นี้จะมีการทำงานที่รวมคุณลักษณะของ Repeater และ Switch เข้าด้วยกัน อุปกรณ์นี้จะใช้การเชื่อมต่อแบบ Full-Duplex และ Flow Control (IEEE802.3x) มันสามารถทำงานเหมือนกับ Repeater คือส่งข้อมูลทุก packet ไปยังทุกๆ Port ที่มีการเชื่อมต่ออยู่ และสามารถทำงานในลักษณะของ Switch คือการรับข้อมูลจากหลาย Port ได้พร้อมกันแล้วนำข้อมูลนั้นไปเก็บไว้ในหน่วยความจำ (Buffer) และเมื่อมีการเขียนลงจนเต็มทางอุปกรณ์นี้ก็จะใช้ Flow Control ส่งสัญญาณให้โหนดที่ส่งข้อมูลนั้นหยุดคอยจนกว่า Buffer นั้นจะว่างลงอีกครั้ง (หลังจากอุปกรณ์ได้ทำการส่งข้อมูลใน Buffer เหล่านั้นไปยังปลายทางเรียบร้อยแล้ว) วิธีนี้ก็จะสามารถให้ Throughput ได้เกือบ 100% แต่ข้อจำกัดของวิธีนี้คือทุกโหนดที่เชื่อมต่อกับอุปกรณ์นี้จะต้องเป็นแบบ Full-Duplex และสนับสนุนมาตรฐาน IEEE802.3x ด้วย

### 2.3 หลักการทำงานของกล้องดิจิตอล

หลักการทำงานของกล้องดิจิตอล คือ เมื่อทำการถ่ายภาพที่ถ่ายจะไปตกบน อุปกรณ์รับภาพที่ประกอบด้วยเซลล์ไวแสงจำนวนมาก และเซลล์เหล่านี้จะทำหน้าที่แปลงภาพซึ่งเป็นสัญญาณ Analog ไปเป็นสัญญาณ Digital โดยอุปกรณ์ที่เรียกว่า ATD ( Analog to digital converter) จากนั้นจึงเข้าสู่กระบวนการตรวจสอบความคมชัดของภาพและทำการบีบอัดเพื่อลดพื้นที่ในการจัดเก็บภาพก่อนจะส่งไปเก็บที่หน่วยความจำของกล้อง ในกล้องทุกตัว แนนอนหัวใจสำคัญที่สุดอันหนึ่งที่จะทำให้กล้องตัวนั้นถ่ายทอดรูปออกมาได้คือ Sensor รับภาพ ซึ่งมีหน้าที่รับแสงที่เข้ามาแล้วเปลี่ยนค่าแสงนั้นๆ เป็นสัญญาณดิจิตอล ซึ่งในปัจจุบันก็ยังมี Sensor รับภาพอยู่เพียง 2 แบบใหญ่ๆ เท่านั้น ซึ่งก็คือ CCD (ซีซีดี) และ CMOS (ซีมอส) ที่เป็นคู่แข่งที่สำคัญในท้องตลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.4 รายละเอียดและเทคโนโลยีของกล้องที่ใช้งาน



รูปที่ 2.3 Basher Camera

Specification	acA640-100gm/gc
Sensor Size (H x V pixels)	640 x 480
Sensor Type	Sony ICX618
Sensor Readout	Progressive Scan CCD
Optical Size	1/4"
Pixel Size (in $\mu\text{m}$ )	5.4 x 5.4 $\mu\text{m}$
Max. Frame Rate (at full resolution)	100 frames per second
Mono/Color	Mono / Color
Interface	Fast Ethernet (100 Mb/s) or Gigabit Ethernet (1000 Mb/s)
Video Output Format	Mono 8, Mono 16, Mono 12 packed, YUV 4:2:2, Raw 8, Bayer BG8, Bayer BG16, Bayer BG12 packed
Synchronization	Via external trigger, via the Ethernet connection, or free-run
Exposure Control	Freely programmable
Housing Size (L x W x H in mm)	42 x 29 x 29
Housing Temperature	Up to 50°C
Lens Mount	C-mount

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Digital I/O	1 opto-isolated input, 1 opto-isolated output
Power Requirements	Via Power over Ethernet (802.3af) or +12VDC (+/- 10%) via 6-pin Hirose connector
Power Consumption (typical)	2.3 W with PoE, 2.0 W with AUX
Weight (typical)	< 90g
Conformity	CE, FCC, GigE Vision, IP 30, RoHS, PoE (802.3af), UL (in preparation), GenICam
Software Driver	Basler pylon SDK incl. filter and performance driver
Operating System	Windows, Linux - 32 and 64 bit



รูปที่ 2.4 อุปกรณ์ที่ใช้ในการเชื่อมต่องานกับคอมพิวเตอร์

## 2.5 พารามิเตอร์ในการจับวิดีโอ (Video Capture Parameters)

ในการจับวิดีโอจากกล้องวิดีโอ หรืออุปกรณ์จับวิดีโอ พารามิเตอร์ที่จะต้องกำหนดได้แก่ อัตราเฟรม ขนาดภาพ และระบบสี ซึ่งรายละเอียดของพารามิเตอร์แต่ละตัวมีดังนี้

### 2.5.1 อัตราเฟรม (Frame Rate)

วิดีโอก็คือการนำภาพหลายๆภาพมาแสดงต่อกันนั่นเอง อัตราเฟรมเป็นค่าที่บ่งบอกถึงความเร็วในการเปลี่ยนภาพดังกล่าว มีหน่วยเป็นเฟรมต่อวินาที (fps) วิดีโอที่มีอัตราเฟรมสูงจะมีความราบรื่นในการรับชมมากกว่าวิดีโอที่มีอัตราเฟรมต่ำ แต่ก็ทำให้อัตราบิตในการส่งวิดีโอสูงขึ้นเช่นกัน โดยปกติแล้ววิดีโอของการฉายภาพยนตร์จะมีอัตราเฟรมอยู่ที่ 24 เฟรมต่อวินาที

## 2.5.2 ขนาดภาพ / ความละเอียด (Frame Size / Resolution)

ขนาดภาพของวิดีโอเป็นค่าที่ระบุถึงความกว้างและความสูงของภาพมีจำนวนกี่จุดภาพ (Pixel) วิดีโอที่มีขนาดภาพใหญ่ย่อมเป็นที่พึงพอใจแก่ผู้ใช้ แต่ก็ต้องใช้อัตราบิตในการส่งวิดีโอสูงกว่าวิดีโอที่มีขนาดภาพเล็กกว่า

## 2.5.3 ระบบสี (Color Space)

ระบบสีเป็นรูปแบบในการกำหนดค่าสำหรับแทนสี ระบบสีที่ใช้กันอย่างแพร่หลายในการสื่อสารวิดีโอได้แก่ ระบบ RGB และระบบ YCbCr โดยในระบบ RGB จุดจุดภาพแต่ละจุดจะถูกนำเสนอด้วยค่า 3 ค่าคือ ค่าสีแดง (R) ค่าสีเขียว (G) และค่าสีน้ำเงิน (B) ส่วนระบบสี YCbCr จะแบ่งข้อมูลของจุดภาพตามการรับรู้ของประสาทตาของคน โดยจะแยกข้อมูลเกี่ยวกับความสว่างออกจากข้อมูลเกี่ยวกับระดับสี นั่นคือใช้องค์ประกอบ Y เก็บค่าความสว่างของจุดภาพ ส่วน Cb คือค่าความแตกต่างระหว่างค่าสีน้ำเงินกับค่าความสว่าง (B-Y) องค์ประกอบ Cr คือค่าความแตกต่างระหว่างค่าสีแดงกับค่าความสว่าง (R-Y) ส่วนค่าความแตกต่างระหว่างค่าสีเขียวกับค่าความสว่าง (Cg) นั้นไม่จำเป็นต้องเก็บ เนื่องจากว่าค่าของ  $Cb + Cr + Cg$  มีค่าคงที่ ดังนั้นถ้าหากทราบเพียงค่า Cb และ Cr ก็สามารถหาค่า Cg ได้ อย่างไรก็ตามไม่ว่าจะเก็บข้อมูลของจุดภาพในระบบ RGB หรือ YCbCr ก็สามารถแปลงเป็นอีกระบบหนึ่งได้เสมอ โดยหากต้องการแปลงจากระบบ RGB เป็นระบบ YCbCr ก็สามารถแปลงได้ดังนี้

$$Y = 0.299R + 0.587G + 0.114B$$

$$Cb = 0.564(B - Y)$$

$$Cr = 0.713(R - Y)$$

หรือในทางกลับกัน การแปลงจากระบบ YCbCr เป็นระบบ RGB ก็สามารถทำได้โดยใช้สมการต่อไปนี้

$$R = Y + 1.402Cr$$

$$G = Y - 0.344Cb - 0.714Cr$$

$$B = Y + 1.772Cb$$

## 2.6 การปรับเทียบกล้อง

เป็นการหาประมาณเมทริกซ์  $M$  จากจุดในฉาก 3 มิติ ที่รู้ตำแหน่งและภาพที่เกิดขึ้น จากนั้นทำการหาพารามิเตอร์ภายในและพารามิเตอร์ภายนอก พิจารณาการปรับเทียบ ในการหาเมทริกซ์  $M$  เราจะทำการหาจุดในฉากที่เราทราบพิกัด  $X = [x, y, z]^T$  และจุดในภาพ 2 มิติ ที่สอดคล้องกัน  $[u, v]^T$  เราจะเห็นว่า จะได้ตัวแปรที่ไม่ทราบค่า 11 ตัวแปร แทนที่จะเป็น 12 ตัวแปร เนื่องจากเฟคเตอร์การสเกลที่ไม่ทราบค่าได้ ในการแก้สมการโฮโมจีเนียสเราต้องใช้อย่างน้อย 6 จุด ถ้ามีมากกว่า 6 จุด จะได้สมการ over-determined ซึ่งสามารถแก้ได้โดยใช้วิธี Least Square หรือใช้การหา Singular Value Decomposition (SVD) ของ  $A$  เมทริกซ์  $M$  ได้จากแถวสุดท้ายของ  $V$  เมื่อได้เมทริกซ์  $M$  แล้ว ขั้นตอนต่อไปทำการแยกพารามิเตอร์ภายใน (Intrinsic Parameter) ดังนั้นเมทริกซ์ย่อย  $3 \times 3$  แทนด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A เวกเตอร์คอลัมน์ขวามือสุดแทนด้วย b เวกเตอร์ที่มีพารามิเตอร์ระยะเคลื่อนที่สามารถแยกออกมาได้ง่ายโดย  $t = -A^{-1} b$  เวกเตอร์ t ให้ข้อมูลที่บอกถึงตำแหน่งกึ่งกลางของระนาบภาพ (Image Plane) จากนั้นพิจารณา  $A = KR$  โดยที่ K เป็นเมทริกซ์สามเหลี่ยมบน (Upper Triangle) และเมทริกซ์ R เป็นเมทริกซ์ออร์โธกอนอล การแยกเมทริกซ์ K และ R สามารถทำได้โดยใช้เทคนิค QR Decomposition สำหรับเมทริกซ์ A

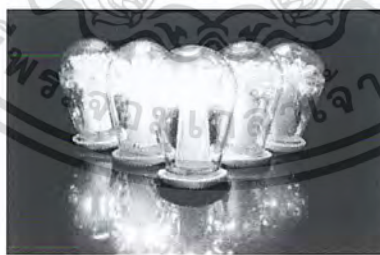
## 2.7 Optical systems



รูปที่ 2.5 Passive marker

### 2.7.1 Passive markers

เป็นระบบที่ใช้มาร์คเกอร์ที่เป็นวัสดุในการสะท้อนแสงกลับมายังกล้อง จากนั้นจะใช้กล้องทำการปรับค่า bright หรือทำการฟิลเตอร์ส่วนที่ไม่สนใจออกไปเหลือเฉพาะส่วนที่เราต้องการใช้ในการวิเคราะห์เท่านั้น มีลักษณะเบาและจะมีขนาดต่างๆที่ต่างกันออกไปแล้วแต่ระยะทางที่ทำการวิเคราะห์ อย่างเช่นขนาดความกว้างประมาณ 2.5 mm จะสามารถใช้ในการตรวจจับของริมฝีปากได้



รูปที่ 2.6 Active marker

### 2.7.2 Active markers

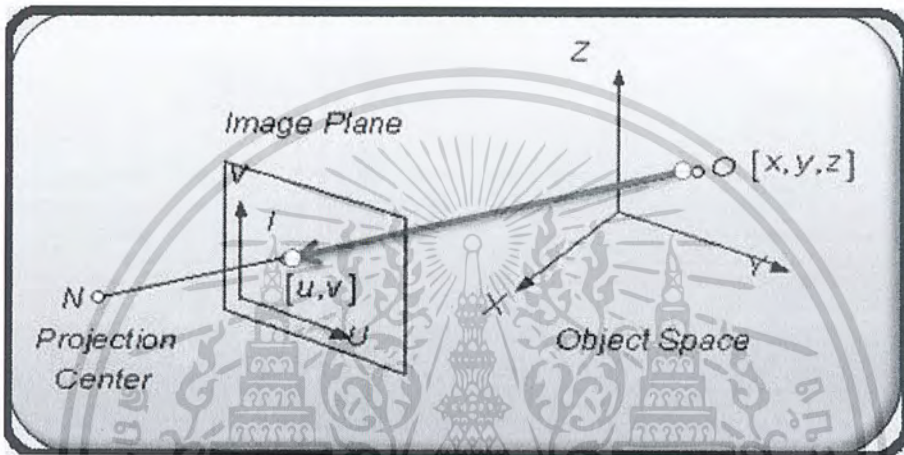
เป็นระบบที่ใช้มาร์คเกอร์ที่มีการเปล่งแสงออกมา เช่น หลอด LED เป็นต้น และส่งลำแสงไปยังกล้องเพื่อทำการบอกกล้องว่าตำแหน่งต่างๆของมาร์คเกอร์นั้นอยู่ที่ใด และใช้กล้องทำการปรับแต่งเพื่อหาจุดมาร์คเกอร์ที่แท้จริง มาร์คเกอร์ชนิดนี้ เหมาะกับการวัดแบบ real-time สำหรับหาตำแหน่งหรือการหมุน เหมาะสำหรับใช้ใน Biomedical เหมาะกับการวิเคราะห์ในทันทีที่มีแสงอื่นเข้ามาเกี่ยวข้อง หรือในสภาพแวดล้อมที่มีการ reflex ของแสงมีผลต่อการวิเคราะห์ผลภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.8 Transformation

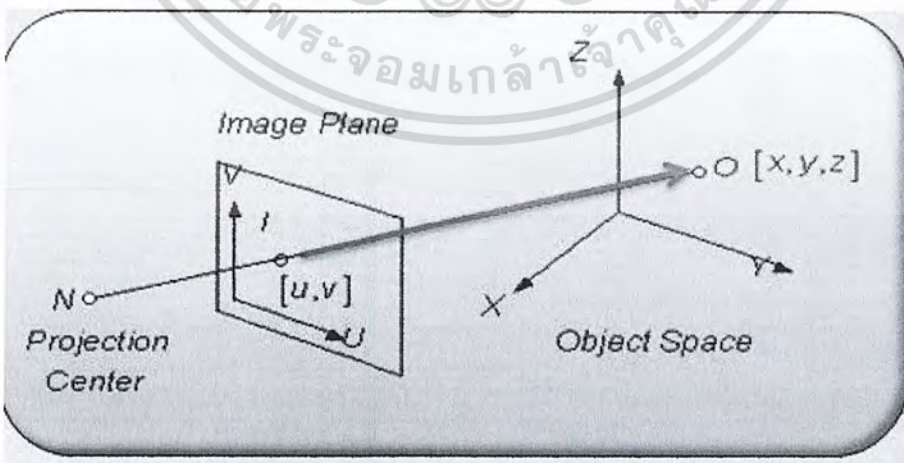
เป็นการแปลงข้อมูลภาพซึ่งในบทนี้จะกล่าวถึงสำหรับการแปลงภาพในสองมิติไปเป็นภาพในสามมิติ การแปลงภาพเป็นสิ่งจำเป็นมากสำหรับ Digital Image Processing เนื่องจากเป็นกระบวนการที่จะนำไปสู่การวิเคราะห์ภาพ (Digital Image Analysis) ที่ต้องการ

โดยการเห็นใน 3 มิติ (หรือเรียกว่า พิกัดโลก) จะมีแกนทางคณิตศาสตร์ อยู่ 3 แกน (X, Y, Z) เป็นพื้นฐานในการเห็นใน 3 มิติ เมื่อผ่านการถ่ายภาพผ่านเลนส์ผลที่ได้จะเป็นภาพ 2 มิติ (หรือเรียกว่า พิกัดภาพ) จะมีแกนทางคณิตศาสตร์ อยู่ 2 แกน (u,v) ดังแสดงได้ดังรูป



รูปที่ 2.7 แสดงการมองจาก 3 มิติ เป็น 2 มิติ

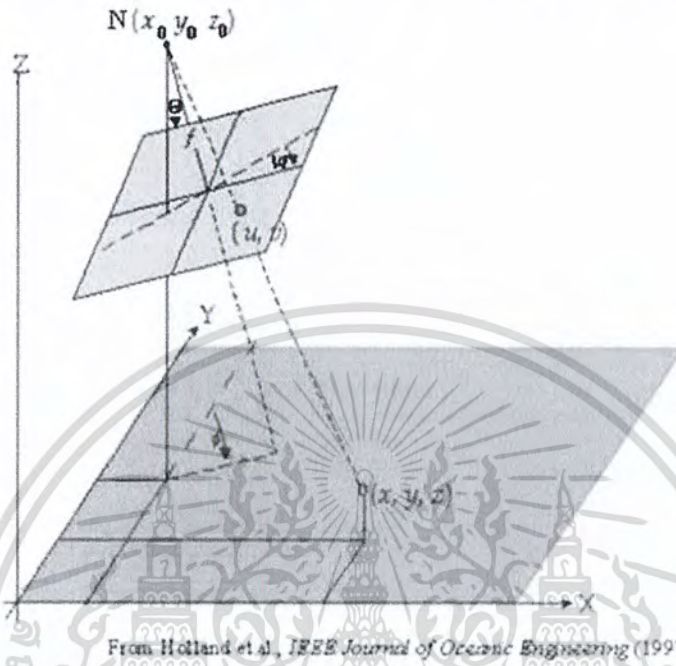
ในการแปลงกลับ การเห็นใน 2 มิติ (หรือเรียกว่า พิกัดภาพ) จะมีแกนทางคณิตศาสตร์ อยู่ 2 แกน (u, v) ซึ่งเมื่อผ่านการวิเคราะห์ประมวลผลผ่านโปรแกรม ผลที่ได้จะเป็นภาพ 3 มิติ (หรือเรียกว่า พิกัดภาพ) (X,Y,Z) จะมีแกนทางคณิตศาสตร์ อยู่ 3 แกน ดังแสดงได้ดังรูป



รูปที่ 2.8 แสดงการมองจาก 2 มิติ เป็น 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นการดึงข้อมูลออกจากภาพ 2 มิติ มีเหตุผลหลายประการที่ทำให้เห็นภาพ 3 มิติ ซึ่งในซึ่งมีความยากอย่างเช่น ระบบภาพของกล้องและตาของมนุษย์มีการมองแบบ เพอสเปคทีฟ ซึ่งทำให้มีการสูญเสียอย่างมากในการประมวลผลภาพจะมีผลทำให้ภาพที่ได้มีมุมมองที่คลาดเคลื่อน



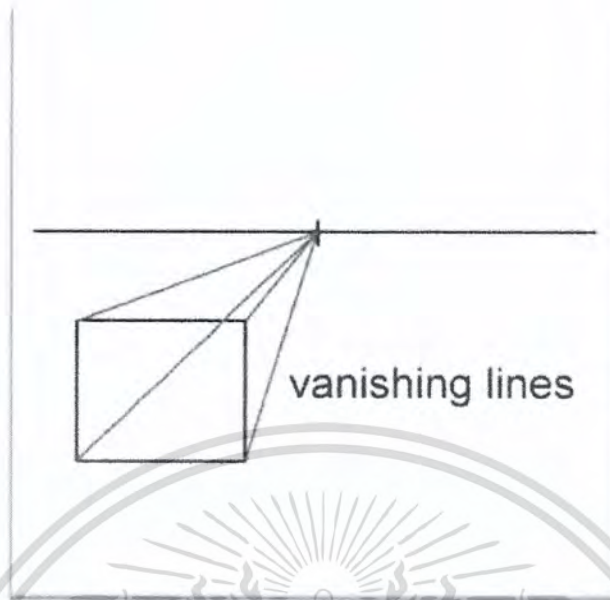
From Holland et al., *IEEE Journal of Oceanic Engineering* (1997)

รูปที่ 2.9 การมองแบบ Perspective

### 2.8.1 การแปลงแบบเพอสเปคทีฟ

ภาพ 2 มิติ ส่วนใหญ่ได้มาจากการมองวัตถุ 3 มิติ ผ่านเลนส์ของตาหรือเลนส์ของกล้อง ขบวนการเกิดภาพสามารถถูกจำลองได้โดยการใช้การแปลงแบบ เพอสเปคทีฟ (Perspective Transformation) ทุกจุดที่อยู่ในแนวเส้นตรงที่ลากจากจุดกึ่งกลางของการโปรเจกชันถึงจุดในฉากถูกโปรเจกชันลงบนจุดเดียวกัน การเห็นใน 3 มิติ เป็นกระบวนการย้อนกลับซึ่งเป็นการหาพิกัด 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.10 Vanishing Lines

การแปลงเพอสเปคทีฟ (Perspective Transformations) เป็นกระบวนการที่เกิดภาพในลักษณะที่แสงของวัตถุเข้าสู่จุดกึ่งกลางของเส้นที่ไประนาบของภาพ นั่นคือลำแสงของแต่ละจุดไม่ขนานกัน จะเห็นในลักษณะวัตถุที่อยู่ไกลจะมีลักษณะเล็กกว่าวัตถุที่อยู่ใกล้

ภาพ 2 มิติส่วนใหญ่ได้มาจากการมองวัตถุ 3 มิติ ผ่านเลนส์ของตาหรือเลนส์ของกล้อง อย่างที่สองคือ ความสัมพันธ์ใน 2 มิติ และเรขาคณิต ใน 3 มิติ ของจุดที่สอดคล้องกันในฉากมีความซับซ้อนมาก ทั้งนี้เนื่องจากค่าความเข้มของพิกเซล จะขึ้นอยู่กับ การวางตัวของผิว ชนิดที่ตำแหน่งของแสง ที่ตกกระทบตำแหน่งของผู้สังเกต

การแปลงจาก ภาพ 2 มิติ เป็นภาพ 3 มิติ คือการหาพิกัดของจุดที่ถูกโปรเจกชัน  $(x,y,z)$  ในหน่วยพิกัดของพิกเซลมีการกำหนดให้จุดพิกัดของภาพที่มีจุดกำเนิดที่มุมบนขวาของภาพ การแปลงจากความยาวเป็นพิกเซลเราจำเป็นต้องรู้ค่าการปรับสเกล (Aspect Ratio) ของแต่ละแกนของระบบพิกัดภาพ ค่าการปรับสเกลเป็นค่าที่ระบุจำนวนพิกเซลต่อหนึ่งหน่วยความยาว ในระบบภาพบางระบบพิกเซลของแถวถัดไปอาจมีการเอียงกัน เราจำเป็นต้องรู้ค่าที่เรียกว่าพารามิเตอร์ของการเอียง

### 2.8.2 Direct linear transformation

การแปลงเชิงเส้นตรง (DLT) เป็นวิธีการหาความสัมพันธ์ขั้นต้นที่จะแก้ชุดของตัวแปรความคล้ายคลึงกันจากการตั้ง

$$\mathbf{x}_k \propto \mathbf{A} \mathbf{y}_k \quad \text{โดย } k = 1, \dots, N$$

เวกเตอร์ที่รู้จักกันที่  $\mathbf{x}_k$  และ  $\mathbf{y}_k$ ,  $\propto$  หมายถึงการคูณสเกลามี unknown และ  $\mathbf{A}$  เป็นเมทริกซ์ (หรือการแปลงเชิงเส้น) ซึ่งมี unknown จะแก้สมการได้

ประเภทของความสัมพันธ์นี้จะปรากฏขึ้นบ่อยครั้งใน เรขาคณิตฉาย ตัวอย่างที่เป็นประโยชน์ รวมถึงความสัมพันธ์ระหว่าง 3 มิติในฉากของพวกเขาและการฉายภาพลงบนระนาบของ กล้อง pinhole และ hymnographies

### สมการเชิงเส้นสามัญ

$$\mathbf{x}_k = \mathbf{A} \mathbf{y}_k \quad \text{โดย } k = 1, \dots, N$$

จะสามารถแก้ไขได้เช่นโดยการเขียนเป็นสมการเมทริกซ์  $\mathbf{X} = \mathbf{A} \mathbf{Y}$  ที่เมทริกซ์  $\mathbf{X}$  และการ  $\mathbf{Y}$  มี พาทะ  $\mathbf{x}_k$  และการ  $\mathbf{y}_k$  ในคอลัมน์ของตน ระบุว่าไม่มีทางออกที่ไม่ซ้ำกันอยู่ก็จะได้รับโดย

$$\mathbf{A} = \mathbf{X} \mathbf{Y}^T (\mathbf{Y} \mathbf{Y}^T)^{-1}$$

Solutions ยังสามารถอธิบายไว้ในกรณีที่มีสมการมีเกินหรือต่ำกว่าที่กำหนดในสิ่งที่ทำให้ การแปลงเชิงเส้นตรงปัญหาที่แตกต่างเกินมาตรฐานจากกรณีขึ้นอยู่กับความเป็นจริงคือซ้ายและขวา ที่สองฝ่ายสามารถกำหนดสมการไม่รู้จักความแตกต่างซึ่งเป็นปัจจัยคูณ  $k$ . เป็นผลให้  $\mathbf{A}$  ไม่สามารถ คำนวณเป็นมาตรฐานในกรณีที่ แต่ความสัมพันธ์เหมือนกันถูกเขียนใหม่เป็นเนื้อเดียวกันสมการเชิงเส้น ที่เหมาะสมแล้วจะสามารถแก้ไขได้โดยวิธีการมาตรฐาน วิธีการรวมกันของความคล้ายคลึงกันเขียน สมการเชิงเส้นเป็นสมการที่เป็นเนื้อเดียวกันและการแก้พวกเขาโดยเป็นมาตรฐานที่เรียกว่า การเปลี่ยนแปลงขั้นตอนวิธีเชิงเส้นตรง หรือ ขั้นตอนวิธี DLT.

### ตัวอย่าง

ให้  $\mathbf{x}_k \in \mathbb{R}^2$  และการ  $\mathbf{y}_k \in \mathbb{R}^3$  เป็นสองชุดของเวกเตอร์ที่รู้จักและปัญหาคือการหา  $2 \times 3$  เมทริกซ์  $\mathbf{A}$  ดังกล่าว

$$\alpha_k \mathbf{x}_k = \mathbf{A} \mathbf{y}_k \quad \text{สำหรับการ } k = 1, \dots, N$$

ที่  $\alpha_k \neq 0$  เป็นปัจจัยที่สเกลไม่ทราบเกี่ยวกับสมการ  $k$ .

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อที่จะกำจัดของสเกลาร์ที่ไม่รู้จักและได้รับสมการที่เป็นเนื้อเดียวกันให้กำหนดป้องกันเมทริกซ์สมมาตร

$$\mathbf{H} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

และคูณทั้งสองด้านของสมการด้วย  $\mathbf{x}_k^T \mathbf{H}$  จากด้านซ้าย

$$\alpha_k \mathbf{x}_k^T \mathbf{H} \mathbf{x}_k = \mathbf{x}_k^T \mathbf{H} \mathbf{A} \mathbf{y}_k \quad \text{โดย } k = 1, \dots, N.$$

ตั้งแต่  $\mathbf{x}_k^T \mathbf{H} \mathbf{x}_k = 0$ , เป็นพจน์เดียวกันต่อไปซึ่งไม่ได้มีสเกลาร์ที่ไม่รู้จักที่อยู่ที่มี

$$0 = \mathbf{x}_k^T \mathbf{H} \mathbf{A} \mathbf{y}_k \quad \text{โดย } k = 1, \dots, N.$$

เพื่อที่จะแก้  $\mathbf{A}$  จากชุดของสมการนี้ให้พิจารณาองค์ประกอบของเวกเตอร์  $\mathbf{x}_k$  และการ  $\mathbf{y}_k$  และเมทริกซ์  $\mathbf{A}$ :

$$\mathbf{x}_k = \begin{pmatrix} x_{1k} \\ x_{2k} \end{pmatrix}, \quad \mathbf{y}_k = \begin{pmatrix} y_{1k} \\ y_{2k} \\ y_{3k} \end{pmatrix}, \quad \text{และ} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

และสมการข้างต้นจะกลายเป็นพจน์เดียวกัน

$$0 = a_{11} x_{2k} y_{1k} - a_{21} x_{1k} y_{1k} + a_{12} x_{2k} y_{2k} - a_{22} x_{1k} y_{2k} + a_{13} x_{2k} y_{3k} - a_{23} x_{1k} y_{3k} \\ \text{โดย } k = 1, \dots, N.$$

ยังสามารถเขียน

$$0 = \mathbf{b}_k^T \mathbf{a} \quad \text{โดย } k = 1, \dots, N$$

ที่  $\mathbf{b}_k$  และการ  $\mathbf{a}$  ทั้งคู่เป็นพหุ 6 มิติตามที่ระบุไว้

$$\mathbf{b}_k = \begin{pmatrix} x_{2k} y_{1k} \\ -x_{1k} y_{1k} \\ x_{2k} y_{2k} \\ -x_{1k} y_{2k} \\ x_{2k} y_{3k} \\ -x_{1k} y_{3k} \end{pmatrix} \quad \text{และการ} \quad \mathbf{a} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \\ a_{13} \\ a_{23} \end{pmatrix}.$$

ชุดของสมการที่เป็นเนื้อเดียวกันนี้ยังสามารถเขียนในรูปเมทริกซ์

$$\mathbf{0} = \mathbf{B} \mathbf{a}$$

ที่  $\mathbf{B}$  คือ  $N \times 6$  เมทริกซ์ซึ่งถือเป็นพหุ  $\mathbf{b}_k$  ในแถวของ ซึ่งหมายความว่า  $\mathbf{a}$  ตั้งอยู่ในพื้นที่ว่างของ  $\mathbf{B}$  และสามารถระบุได้ด้วยตัวอย่างเช่นโดยการ สลายตัวของมูลค่าเอกพจน์ ของ  $\mathbf{B}$ ;  $\mathbf{a}$  เป็นเส้นเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เดียวทางด้านขวาของ  $B$  ที่สอดคล้องกับค่าเอกพจน์มีค่าเท่ากับศูนย์ ครั้งหนึ่ง  $a$  ได้รับการกำหนดองค์ประกอบของ  $A$  สามารถพบได้โดยง่ายจากการจัดแจงใหม่เวกเตอร์ 6 มิติเพื่อ  $2 \times 3$  เมทริกซ์ที่ประกาศการปรับขนาดของ  $a$  หรือ  $A$  ไม่สำคัญ (ยกเว้นว่าจะต้องไม่เป็นศูนย์) ตั้งแต่สมการกำหนดแล้วอนุญาตให้มีการปรับที่ไม่รู้จัก

เวกเตอร์ในทางปฏิบัติ  $x^k$  และการ  $y^k$  อาจมีการรบกวนซึ่งหมายความว่าสมการความคล้ายคลึงกันเป็นเพียงที่ถูกต้องประมาณ ผลที่ตามมาอาจจะไม่เชิงเส้น  $a$  ซึ่งจะแก้สมการที่เป็นพจน์เดียวกัน  $0 = Ba$  เอง ในกรณีเหล่านี้เป็นโซลูชันที่รวมกำลังสองอย่างน้อยสามารถนำมาใช้โดยเลือก  $a$  เป็นเวกเตอร์เอกพจน์ขวาที่สอดคล้องกับค่าที่เล็กที่สุดของเอกพจน์  $B$ .

### 2.8.3 Matrix

Matrix เป็นเทคนิคทางคณิตศาสตร์ที่ใช้ในการหาคำตอบหรือหาค่าตัวแปรต้นในสมการเชิงเส้นที่มีมากกว่า 1 สมการ โดยจะได้คำตอบหรือค่าตัวแปรที่หาพร้อมกัน ขั้นตอนการคำนวณจะเริ่มจากการจัดรูป Matrices ให้อยู่ในรูปทางคณิตศาสตร์

$$[Y] = [A][X]$$

โดย  $[Y]$ ,  $[A]$  และ  $[X]$  คือ Matrix ที่เราต้องการหาคือ  $[X]$  จึงต้องจัดรูปสมการใหม่เป็นดังนี้

$$[X] = [A]^{-1}[Y]$$

เมื่อ  $[A]^{-1}$  คือ Inverse matrix ของ  $[A]$  เนื่องจากในการคำนวณทางคณิตศาสตร์ Matrix จะไม่มีการหาร จึงต้องใช้วิธีหา Inverse matrix ของ  $[A]$  แล้วนำไปคูณกับ  $[Y]$  เพื่อให้ได้  $[X]$  ซึ่งทำให้การหาค่าตัวแปรในสมการเชิงเส้น โดยใช้ Matrix มีไม่ได้ทำงานง่ายเหมือนการแก้สมการแบบธรรมดาทั่วไป แต่ถ้ามีสมการเชิงเส้นหลายๆสมการ การใช้ Matrix ถือว่าง่ายกว่าการแก้สมการแบบธรรมหลายเท่าตัว มีขั้นตอนการคำนวณตามแผนภาพต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



#### 2.8.4 พีชคณิตเชิงเส้น

เป็นสาขาหนึ่งของคณิตศาสตร์ที่ศึกษาเวกเตอร์ ปริภูมิเวกเตอร์ (หรืออีกชื่อหนึ่งคือ *ปริภูมิเชิงเส้น*) การแปลงเชิงเส้น และระบบสมการเชิงเส้น ปริภูมิเวกเตอร์เป็นเรื่องที่ได้ได้รับความสนใจอย่างมากในคณิตศาสตร์สมัยใหม่ เนื่องจากพีชคณิตเชิงเส้นถูกนำไปใช้อย่างกว้างขวางในคณิตศาสตร์สองสายหลักคือ พีชคณิตนามธรรมและการวิเคราะห์เชิงฟังก์ชัน พีชคณิตเชิงเส้นนั้นมีรูปแบบที่ชัดเจนในเรขาคณิตวิเคราะห์ และถูกขยายให้กว้างขึ้นในทฤษฎีตัวดำเนินการ และมีการประยุกต์ใช้อย่างแพร่หลายในวิทยาศาสตร์และสังคมศาสตร์ เนื่องจากแบบจำลองไม่เชิงเส้น (nonlinear model) ส่วนมากสามารถประมาณการณได้ด้วยแบบจำลองเชิงเส้น (linear model) การประยุกต์ใช้อย่างหนึ่งของพีชคณิตเชิงเส้นคือการแก้ระบบสมการเชิงเส้นหลายตัวแปร กรณีที่ง่ายที่สุดคือเมื่อมีจำนวนที่ไม่ทราบค่า (ตัวแปร) เท่ากับจำนวนของสมการ ดังนั้นเราสามารถแก้ปัญหา ระบบสมการเชิงเส้น  $n$  สมการ สำหรับจำนวนที่ไม่ทราบค่า  $n$  ตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยามให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Singular value decomposition

ใน พีชคณิตเชิงเส้น, ค่าการสลายตัวเอกพจน์ (SVD) ที่สำคัญคือ ตัวประกอบของรูปสี่เหลี่ยมจัตุรัส หรือ ขั้วชั้น เมทริกซ์กับการใช้งานจำนวนมากในการประมวลผลสัญญาณ และ สถิติ. โปรแกรมที่ใช้ SVD รวมถึงการใช้คอมพิวเตอร์ที่ pseudo inverse กำลังสองน้อยที่สุด เหมาะสมของข้อมูลการประมาณค่าเมทริกซ์และการกำหนด ตำแหน่ง, ช่วง และ พื้นที่ว่างของเมทริกซ์

### คำชี้แจงของทฤษฎีบท

สมมติว่า  $M$  คือ  $n \times m$  เมทริกซ์ ที่มีรายการที่มาจาก เขต  $K$  ซึ่งเป็นทั้งเขตของจำนวนจริง หรือของเขตข้อมูล ตัวเลขที่ขั้วชั้น. แล้วมีอยู่รูปตัวประกอบของ

$$M = U\Sigma V^*$$

ที่  $U$  เป็น  $m \times m$  รวมกันของเมทริกซ์ มากกว่า  $K$ ,  $\Sigma$  เมทริกซ์  $n \times m$  ในแนวทแยงของเมทริกซ์ ด้วยจำนวนจริงไม่เป็นค่าลบในแนวทแยงและ  $V, n \times n$  ของเมทริกซ์รวมกันมากกว่า  $K$ , หมายถึง ไขว้สังยุค ของ  $V$ . ตัวประกอบดังกล่าวเรียกว่า การสลายตัวมูลค่าเอกพจน์ ของ  $M$ .

การประจุมรวมกันคือการสังขั้วชั้นรายการในแนวทแยง  $\Sigma_{i,i}$  ในลำดับถัดลง ในกรณีนี้  $\Sigma$  เมทริกซ์ทแยงมุมจะถูกกำหนดโดยไม่ซ้ำกันโดย  $M$  (แต่เมทริกซ์  $U$  และ  $V$  ไม่ได้) รายการเส้นทแยงมุมของ  $\Sigma$  เป็นที่รู้จักกันเป็น ค่าเอกพจน์ ของ  $M$

### การประยุกต์ใช้ SVD

#### Pseudo inverse

ค่าการสลายตัวเอกพจน์สามารถนำมาใช้ในการคำนวณ pseudo inverse ของเมทริกซ์ แท้จริง pseudo inverse ของเมทริกซ์  $M$  กับการสลายตัวเอกพจน์ค่า  $M = U\Sigma V$  เป็น

$$M^+ = V\Sigma^+U^*$$

ที่  $\Sigma^+$  มันเป็น pseudo inverse ของ  $\Sigma$  ซึ่งเป็นที่เกิดขึ้นจากมาเลขทศรายการโดยการแทนที่ ซึ่งกันและกัน และ transposing ผลของเมทริกซ์ pseudo inverse เป็นหนึ่งในวิธีการแก้ปัญหา เชิงเส้นกำลังสองน้อยที่สุด

การแก้สมการเชิงเส้นที่เป็นเนื้อเดียวกันของชุด สมการเชิงเส้นที่เป็นเนื้อเดียวกัน สามารถเขียนเป็น  $Ax = 0$  สำหรับเมทริกซ์  $A$  เวกเตอร์และ  $x$  สถานการณ์โดยทั่วไปก็คือ  $A$  เป็นที่รู้จักกัน และไม่เป็นศูนย์  $x$  คือการได้รับการพิจารณาที่ตรงตามความสม ดังกล่าว  $x$  เป็นของ  $A$ 's พื้นที่ว่าง และบางครั้งเรียกว่า (ขวา) เวกเตอร์โมฆะ  $A \cdot x$  สามารถลักษณะเป็นเวกเตอร์เอกพจน์ขวาที่สอดคล้องกับค่าเอกพจน์ของ  $A$  ที่เป็นศูนย์ การสังเกตซึ่งหมายความว่าหาก  $A$  สแควร์มูลค่าเอกพจน์เมทริกซ์และไม่หายเป็นสมการได้ไม่ไม่ใช่ศูนย์  $x$  เป็นโซลูชัน และก็หมายความว่าถ้ามีหลายค่าเอกพจน์ใด ๆ การรวมกันเชิงเส้นของเวกเตอร์ที่สอดคล้องกันเอกพจน์ขวาเป็นโซลูชันที่ถูกต้อง

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตรงข้ามกับความหมายของเวกเตอร์นัล (ขวา) ไม่เป็นศูนย์  $\mathbf{x}$  ความพึงพอใจ  $\mathbf{x}^* \mathbf{A} = 0$ , ด้วย  $\mathbf{x}^*$  ไหว้แสดงถึงสังยุคของ  $\mathbf{x}$  เรียกว่าเวกเตอร์ว่างด้านซ้ายของ  $\mathbf{A}$

รวมกำลังสองน้อยที่สุดอย่างน้อย

รวมกำลังสองน้อยที่สุด ปัญหาหมายถึงการกำหนดเวกเตอร์  $\mathbf{x}$  ซึ่งช่วยลด 2 - Norm ของเวกเตอร์  $\mathbf{Ax}$  ภายใต้ข้อจำกัด  $\|\mathbf{x}\| = 1$ . ทางออกจะเปิดออกจะเวกเตอร์เอกพจน์ด้านขวาของ  $\mathbf{A}$  ที่สอดคล้องกับค่าเอกพจน์ที่เล็กที่สุด

Range, พื้นที่ว่างและจัดอันดับ

SVD อีกประยุกต์ใช้เป็นที่ให้การแสดงที่ชัดเจนในการ ช่วง และ พื้นที่ว่าง ของเมทริกซ์  $M$  เอกพจน์ขวาที่สอดคล้องกับค่าเอกพจน์ของ  $M$  span พื้นที่ว่างของ  $M$  ตัวอย่างเช่นพื้นที่ว่างสามารถถูกวัดโดยล่าสุดสองคอลัมน์ของ  $V$  ในตัวอย่างข้างต้น พาหะเอกพจน์ซ้ายที่ตรงกับที่ไม่ใช่ค่าศูนย์เอกพจน์ของ  $M$  ช่วงช่วง  $M$ . เป็นผลให้ อันดับ ของ  $M$  เท่ากับจำนวนของค่าที่ไม่ใช่ศูนย์ที่เป็นเอกพจน์เช่นเดียวกับจำนวนที่ไม่ใช่ศูนย์องค์ประกอบใน  $\Sigma$ .

ในพีชคณิตเชิงเส้นเชิงตัวเลขค่าเอกพจน์สามารถใช้ในการกำหนด ตำแหน่งที่มีประสิทธิภาพ ของเมทริกซ์เป็น ข้อผิดพลาดในการปัดเศษ อาจนำไปสู่การมีขนาดเล็ก แต่ไม่ใช่ศูนย์ค่าเอกพจน์ในอันดับที่บกพร่องของเมทริกซ์

### Low-rank matrix approximation

บางโปรแกรมในทางปฏิบัติจำเป็นต้องแก้ปัญหาของการประมาณค่าเมทริกซ์  $M$  กับเมทริกซ์อื่น  $\tilde{M}$  ซึ่งมีเฉพาะอันดับ  $r$ . ในกรณีที่การประมาณจะขึ้นอยู่กับการลด Frobenius Norm ของความแตกต่างระหว่าง  $M$  และ  $\tilde{M}$  ภายใต้ข้อจำกัด ที่  $\text{rank}(\tilde{M}) = r$  ปรากฏว่าการแก้ปัญหาคือให้โดยการ SVD  $M$  คือ

$$\tilde{M} = U\Sigma V^*$$

ที่  $\Sigma$  เป็นเช่นเดียวกับเมทริกซ์  $\Sigma$  ยกเว้นว่าจะมีเพียง  $r$  ใหญ่ที่สุดค่าเอกพจน์ (ค่าเอกพจน์อื่น ๆ จะถูกแทนที่ด้วยศูนย์) นี้เรียกว่า - Young ทฤษฎีบท Eckart เพราะมันเป็นเรื่องพิสูจน์ได้โดยทั้งสองผู้เขียนในปี 1936 (แต่ก็พบในภายหลังเพื่อได้รับทราบกันดีว่าก่อนหน้าผู้เขียน

หลักฐาน จะลด  $\|M - \tilde{M}\|_F$  ภายใต้  $\text{rank}(\tilde{M}) = r$ .

สมมติว่าของ SVD  $M = U\Sigma V^*$ . เมื่อ Norm Frobenius เป็น unitarily ไม่แปรเปลี่ยนเรามีคำสั่งเทียบเท่า :

$$\min_{\tilde{M}} \|\Sigma - U^* \tilde{M} V\|_F.$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทราบว่าตั้งแต่  $\Sigma$  เป็นเส้นทแยงมุม,  $U^* \tilde{M} V$  ควรเป็นแนวทแยงเพื่อลด Norm Frobenius. จำไว้ว่า **Laws Frobenius** เป็นรากที่สองของรายการผลรวมกำลังสองของโมดูลัสทั้งหมด ซึ่งสะท้อนว่า  $U$  และ  $V$  นอกจากนี้ยังมีเมทริกซ์เอกพจน์ของ  $\tilde{M}$ . ดังนั้นเราจึงคิดว่า  $\tilde{M}$  เพื่อลดข้อความข้างต้นมีรูปแบบ :

$$\tilde{M} = U S V^*,$$

ที่  $S$  เป็นเส้นทแยงมุม เส้นขวางรายการ **ของ**  $I$  ของ  $S$  ไม่จำเป็นต้องเป็นคำสั่งใน SVD

$$\min_M \|\Sigma - S\|_F \equiv \min_{s_i} \sqrt{\sum_{i=1}^n (\sigma_i - s_i)^2}.$$

จากข้อ จำกัด อันดับคือ  $S$  มี  $r$  ไม่เป็นศูนย์รายการในแนวทแยงคำสั่งที่กำหนดข้างต้นของการได้มาดังต่อไปนี้

$$\min_{s_i} \sqrt{\sum_{i=1}^r (\sigma_i - s_i)^2 + \sum_{i=r+1}^n \sigma_i^2} = \sqrt{\sum_{i=r+1}^n \sigma_i^2}.$$

ดังนั้นจึง  $\tilde{M}$  อันดับของ  $r$  คือการประมาณที่ดีที่สุดของ  $M$  ใน เกณฑ์ปกติ Frobenius ความรู้สึกเมื่อ  $\sigma_i = s_i$  ( $i = 1, \dots, r$ ) และเวกเตอร์ที่สอดคล้องกันเป็นเอกพจน์เช่นเดียวกับของ  $M$  SVD เมทริกซ์อีกตัวอย่างหนึ่งของการประมาณโดยสามารถแก้ ปัญหา Procrustes มุมฉากของการประมาณค่าใด ๆ ที่ให้เมทริกซ์ด้วยเมทริกซ์เชิงตั้งฉาก

### Separable models

SVD สามารถเป็นตัวย่อยสลายเมทริกซ์ลงในถ่วงน้ำหนัก, สิ่งข้อแยกผลรวมของเมทริกซ์ โดยแยกเราหมายถึงว่าเมทริกซ์  $A$  สามารถเขียนเป็น ผลลัพธ์ด้านนอก ของสองเวกเตอร์  $A = u \otimes v$  หรือในพิกัด  $A(i, j) = u(i)v(j)$ . โดยเฉพาะเอ็มเมทริกซ์สามารถแยกไปเป็น :

$$M = \sum_i A_i = \sum_i \sigma_i U_i \otimes V_i^\dagger.$$

นี่  $U_i$  และ  $V_i$  เป็น คำนวณ คอลัมน์ของเมทริกซ์ SVD ตรงกัน,  $\sigma_i$  / จะสั่งค่าเอกพจน์และแต่ละ  $A_i$  สามารถแยกได้ SVD สามารถใช้ในการหาตัวกรองการสลายตัวของการประมวลผลภาพตัวกรองแนวตั้งและแนวนอนแยกเป็น[1] โปรดทราบว่าศูนย์จำนวนไม่  $\sigma_i$  คือว่าเมทริกซ์อันดับของโมเดลแบบแยกมักจะเกิดขึ้นในระบบชีวภาพและการสลายตัว SVD จะมีประโยชน์ในการวิเคราะห์ระบบดังกล่าว ตัวอย่างเช่นบางเซลล์ภาพพื้นที่อ่อนไหวง่าย  $V_1$  เขตข้อมูลสามารถอธิบายได้ดี โดย Gabor Filter ในโดเมนพื้นที่คุณด้วยฟังก์ชันการมอดูเลตในโดเมนเวลา จึงได้รับการประเมินผ่านตัวกรองเชิงเส้น เช่น ความสัมพันธ์แบบย้อนกลับหนึ่งสามารถจัดเรียงสองมิติเชิงพื้นที่ในมิติหนึ่งจึงยอมสองมิติกรอง (เว้นวรรคเวลา) ที่สามารถย่อยสลายผ่าน SVD คอลัมน์แรกของ  $U$  ในการสลายตัวแล้ว Gabor SVD คือในขณะที่คอลัมน์แรกของ  $V$  แสดงถึงการปรับเวลา (หรือในทางกลับกัน) จากนั้นหนึ่งอาจ

$$\alpha = \sigma_1^2 / \sum_i \sigma_i^2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

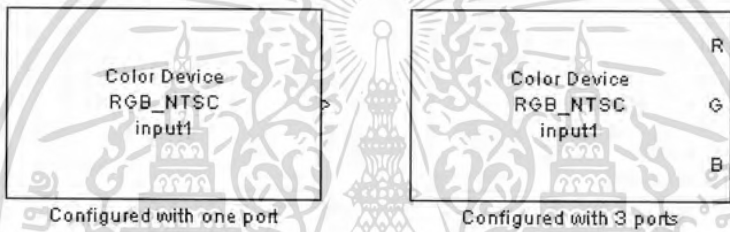
กำหนด separability ดัชนีความ ซึ่งเป็นส่วนของอำนาจในเอ็มเมทริกซ์ซึ่งเป็นสัดส่วนโดยแยกเมทริกซ์ครั้งแรกในการสลายตัว

### Nearest orthonormal matrix

ใน ปัญหา Procrustes ฉาก ที่เราต้องการเพื่อพิจารณา orthonormal เมทริกซ์ ใกล้เคียงกับการใช้ บรรทัดฐาน Frobenius. การแก้ปัญหาคือผลผลิต  $UV^*$  นี้ทำให้รู้สึกสังหรณ์ใจเพราะ orthonormal เมทริกซ์จะมีการสลายตัว  $UV^*$  ที่  $I$  เป็นเอกลักษณ์ของเมทริกซ์เพื่อที่ว่า ถ้า  $= U\Sigma V^*$  แล้วผลผลิต  $= UV^*$  จำนวนเงินที่จะเปลี่ยนค่าเอกพจน์กับคน

## 2.9 ทฤษฎีและหลักการพื้นฐานของส่วนการรับภาพ

### 2.9.1 การทำงานของส่วนการรับภาพ



รูปที่ 2.11 บล็อกการทำงานของส่วนการรับภาพ

จากบล็อกอุปกรณ์วิดีโอช่วยให้คุณสามารถรับข้อมูลภาพและวิดีโอสตรีมจากการอุปกรณ์ภาพเช่น กล้องและ grabbers กรอบเพื่อที่จะนำข้อมูลภาพในรูปแบบ Simulink บล็อกนี้ยังช่วยให้คุณกำหนดค่าและแสดงตัวอย่างการได้มาโดยตรงจาก Simulink จากบล็อกอุปกรณ์วิดีโอที่ค้นเปิดขึ้นให้ทำการเริ่มต้น, กำหนดค่าและการควบคุมอุปกรณ์การได้มาเปิดตั้งต้น และการกำหนดค่าเกิดขึ้นพร้อมกันที่จุดเริ่มต้นของการทำงานของแบบจำลอง ในช่วงเวลาการทำงานแบบจำลอง, บล็อกบัพเฟอร์ข้อมูลภาพส่งรูปสำหรับแต่ละขั้นตอนเวลาการจำลอง สามารถตั้งค่าบล็อกที่จะมีทั้งพอร์ตออกหนึ่งหรือทั้งสามพอร์ตส่งออกที่สอดคล้องกับการบีบอัดวงสีเช่นสีแดงเขียวและน้ำเงินหรือ Y, CB, Cr รูปด้านบนจะแสดงการกำหนดค่าทั้งสอง

### คุณสมบัติอื่นที่รองรับ ๑

จากบล็อกอุปกรณ์วิดีโอที่สนับสนุนการใช้โหมด Simulink Accelerator คุณลักษณะนี้จะเพิ่มความเร็วในการประมวลผลแบบจำลอง Simulink จากบล็อกอุปกรณ์วิดีโอที่สนับสนุนการใช้การแบบอ้างอิงคุณลักษณะนี้ช่วยให้โมเดลของคุณรวมถึงโมเดล Simulink อื่น ๆ เป็นส่วนประกอบแบบโมดูลสำหรับข้อมูลเพิ่มเติมเกี่ยวกับคุณลักษณะเหล่านี้ให้ดูที่เอกสาร Simulink

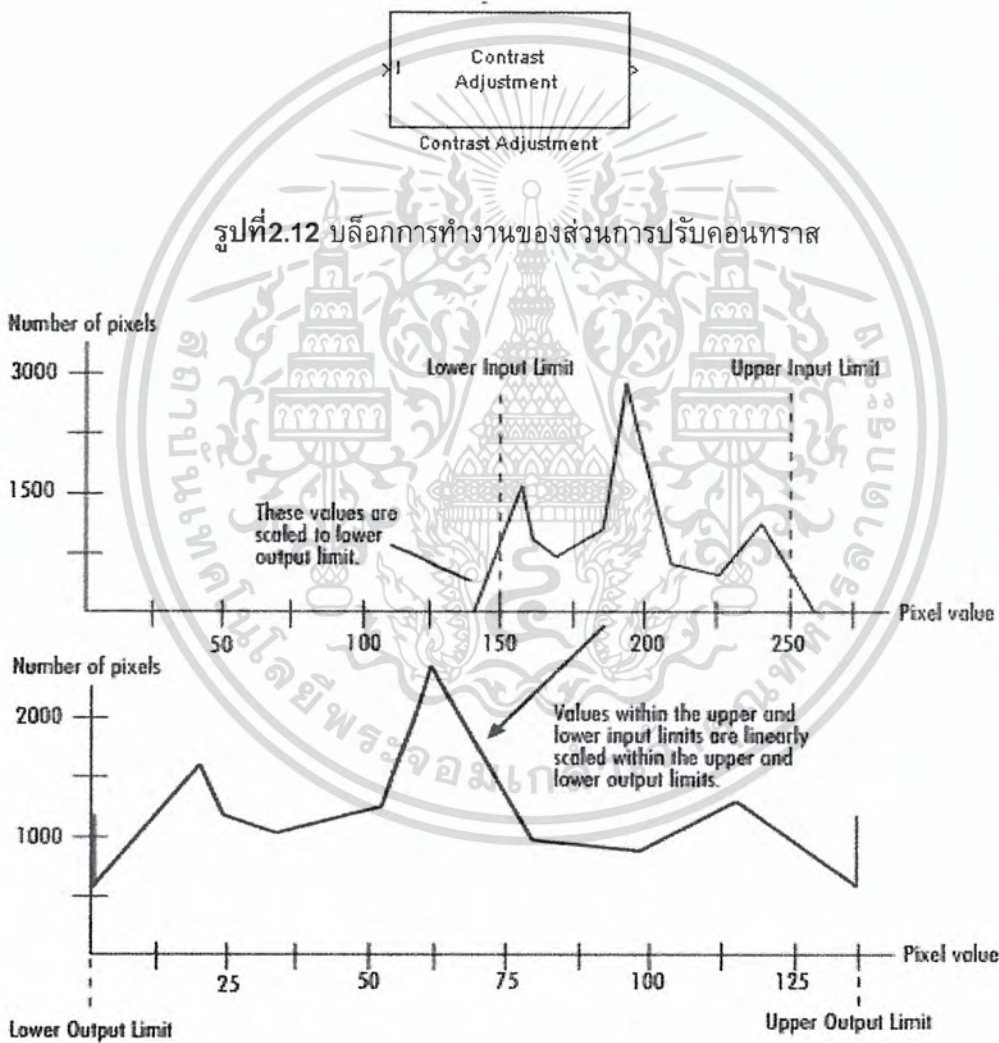
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากบล็อกสนับสนุนการใช้อุปกรณ์วีดิทัศน์ของคนรุ่นรหัสพร้อมด้วยฟังก์ชัน packNGo ที่มาของรหัสที่ต้องการและไลเบอรีที่ต้องการใช้ร่วมกันขึ้นอยู่กับส่วนถัดไป

## 2.10 ทฤษฎีและหลักการพื้นฐานของส่วนการประมวลผลภาพ

### 2.10.1 การปรับคอนทราส

การปรับตัวปรับความคมชัดของภาพโดยปรับค่าพิกเซลเป็นเส้นตรงระหว่างขีด จำกัด บนและล่าง ค่า Pixel ที่สูงกว่าหรือต่ำกว่าช่วงนี้สัมพันธ์กับค่าขีด จำกัดบนหรือลดลงตามลำดับ



รูปที่2.13 กราฟแสดง input limit และ output limit ต่อ pixel

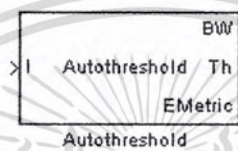
ในทางคณิตศาสตร์, การดำเนินการปรับความคมชัดได้รับการอธิบายโดยสมการดังต่อไปนี้ ซึ่งมีข้อจำกัดทางอินพุท [high\_in low\_in] และข้อจำกัดทางเอาต์พุท [high\_out low\_out]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Output = \begin{cases} low\_out, & Input \leq low\_in \\ low\_out + (Input - low\_in) \frac{high\_out - low\_out}{high\_in - low\_in}, & low\_in < Input < high\_in \\ high\_out, & Input \geq high\_in \end{cases}$$

### 2.10.2 การทำ threshold

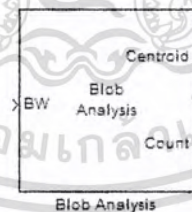
เป็นการแปลงความเข้มของภาพให้เป็นภาพไบนารี



รูปที่ 2.10 บล็อกการทำงานของส่วนทำเป็นภาพขาวดำ

บล็อก Auto threshold แปลงความเข้มของภาพไปยังภาพไบนารีโดยใช้ค่าเกณฑ์การคำนวณ โดยใช้วิธีของฮิสโตแกรม บล็อกนี้จะคำนวณค่าเกณฑ์นี้โดยการแยก Histogram ของภาพอินพุตดังกล่าว ที่ความแปรปรวนของกลุ่มแต่ละพิกเซลจะถูกย่อเล็กที่สุด

### 2.10.3 การหาค่าโคออร์ดิเนตของภาพจากบล็อก Blob Analysis



รูปที่ 2.10 บล็อกการทำงานของส่วนการหาค่าโคออร์ดิเนต

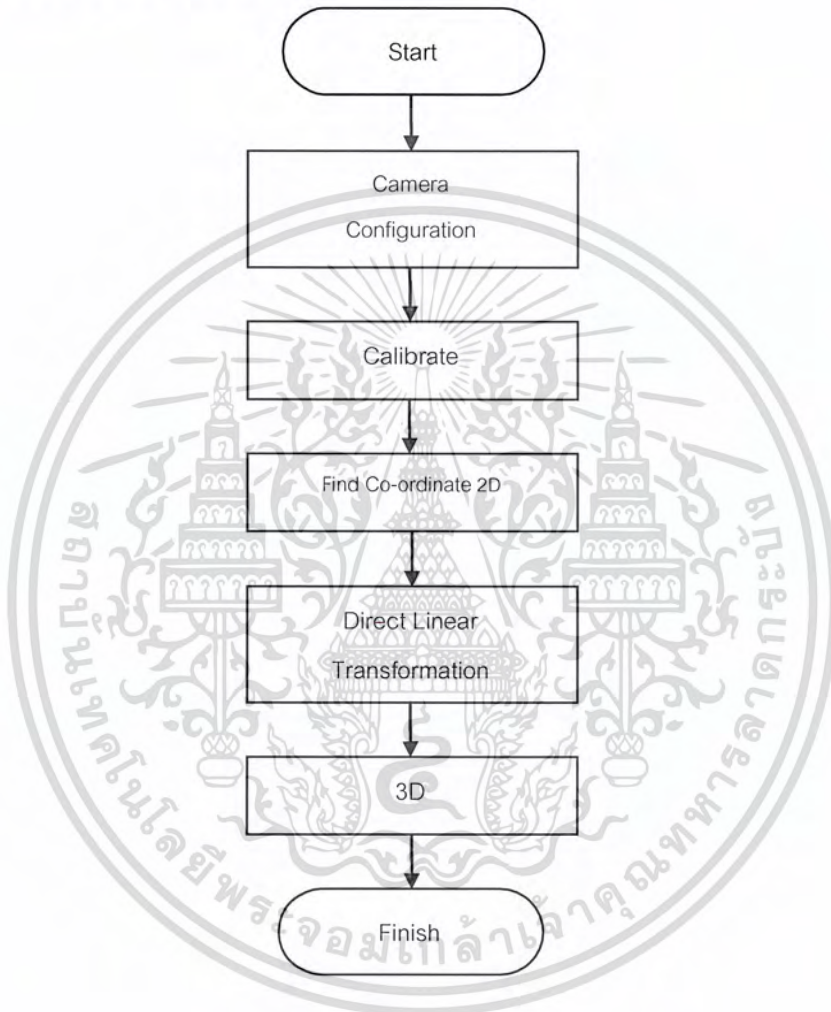
ใช้บล็อก Blob การวิเคราะห์ในการคำนวณทางสถิติสำหรับพื้นที่ที่มีข้อความในรูปแบบไบนารีใน ปริมาณ ของเมทริกซ์และการนับ บล็อก Blob สนับสนุนการวิเคราะห์สัญญาณขนาดตัวแปรที่เข้าและขา ออก สำหรับข้อมูลทาง pixel และความหมายเชิงพื้นที่ที่ประสานงานกับระบบ ที่ระบบการประสานงาน ในการประมวลผลภาพวิดีโอและ Block set ใช้บล็อกเลือกตัวแปรเพื่อเลือก blobs ตามสถิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 3

### การออกแบบการจับภาพเคลื่อนไหว

#### 3.1 โครงสร้างโดยรวมของระบบการจับภาพเคลื่อนไหว



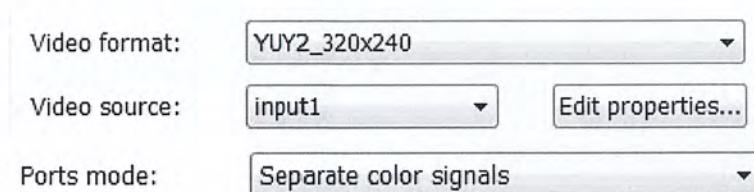
รูปที่ 3.1 โครงสร้างของระบบการจับภาพเคลื่อนไหว

จากรูปที่ 3.1 ผังวงจรการตรวจจับการเคลื่อนไหว เริ่มต้นด้วยการตั้งค่ากล้อง วางตำแหน่งอุปกรณ์ แล้วทำการ calibrate เพื่อหาสัมประสิทธิ์ในการแปลงภาพ จากนั้นทำการติด Active Marker ที่ตำแหน่งสำคัญทั้งหมด 3 ตำแหน่ง คือ สะโพก หัวเข่า และข้อเท้า ทำการจับภาพ และหาพิกัดของตำแหน่งสำคัญที่ได้กำหนดไว้แล้วจากกล้องทั้งสามตัว พร้อมกัน พร้อมกันนั้นนำค่าที่ได้มาประมวลผลโดยนำค่าสัมประสิทธิ์ที่และ พิกัดที่ได้ ไปเข้าสมการ DLT เพื่อแปลงออกมาเป็นตำแหน่งในระบบ 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

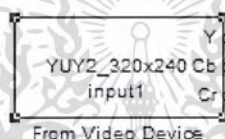
### 3.2 Camera configuration

Camera configuration คือ ตั้งหรือปรับค่ากล้องตามที่เราต้องการ



รูปที่ 3.2 การตั้งค่ากล้อง

โดยทำการตั้งวิดีโอในระบบ YUY2 และเปลี่ยน ขาออกของวิดีโอเป็นแบบแยกสัญญาณ จะได้สัญญาณทั้งหมด 3 แบบ คือ Y,Cb,Cr แล้วนำช่องสัญญาณ Cr ไปใช้งานในระบบ



รูปที่ 3.3 ช่องสัญญาณต่างๆ

จากนั้นนำกล้องมาวางในตำแหน่งต่างๆใน 3 มุม โดยกล้องแต่ละตัว ต้องมองเห็นการเดินทางของผู้ถูกทดลองตั้งแต่บริเวณ ข้อเท้าถึงสะโพก



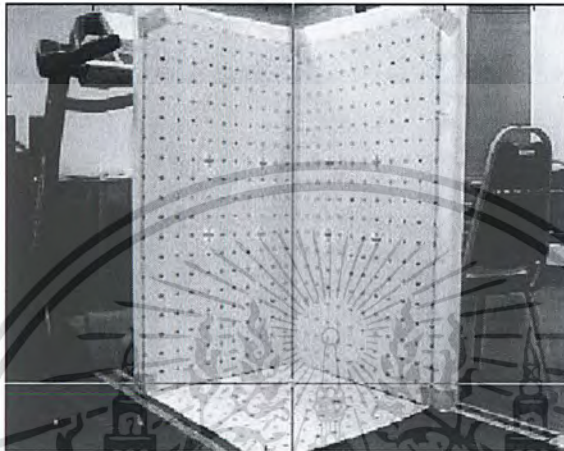
รูปที่ 3.4 การตั้งกล้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3 Calibrate

Calibrate คือการ เปรียบเทียบค่าของกล้องกับระบบ World space หรือการเปรียบเทียบพิกัดที่กล้องอ่านได้ กับ พิกัดที่แท้จริงบนโลก

การ calibrate อย่างง่าย สามารถทำได้โดยใช้ Pattern model กำหนดค่าที่แท้จริงบน World Space

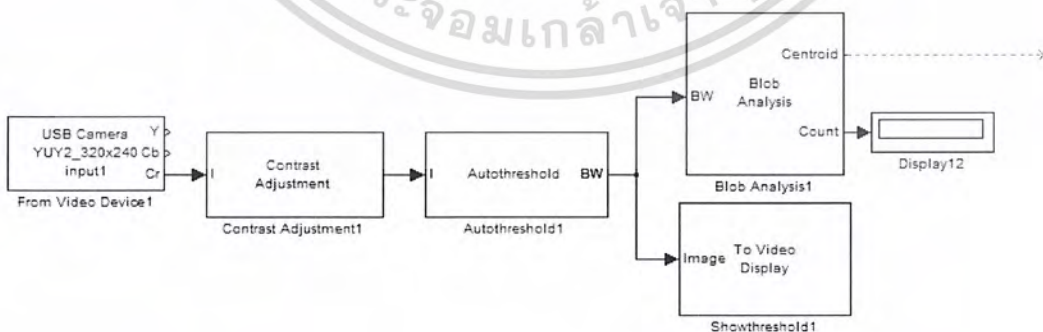


รูปที่ 3.5 Pattern Model

นำค่าที่กล้องอ่านจุดที่กำหนดค่าไว้แล้ว ทั้ง 3 กล้อง มาเข้าสมการ DLT (Direct Linear Transform) จะได้ค่าสัมประสิทธิ์การแปลงกลับในระนาบ 3 มิติ

### 3.4 การหา Co-ordinate

การหา Co-ordinate คือการหาพิกัดบนภาพที่กล้องอ่านได้ในแต่ละเฟรม

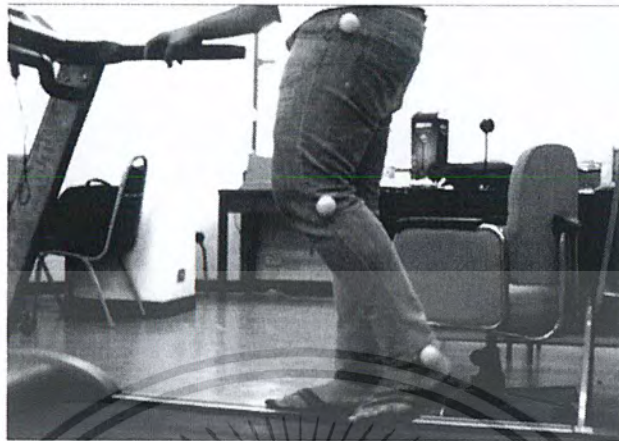


รูปที่ 3.6 กระบวนการรับภาพ

รูปที่ 3.3.4 แสดงถึงกระบวนการการรับภาพ โดย Input เป็น Video camera ในระบบสี YUY2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยแยก Out put ของวีดีโอเป็น Y,Cb,Cr แล้วทำการรับภาพ โดยให้ผู้ถูกทดลอง ติด Marker สีในโทนสีแดง ที่บริเวณข้อเท้า เข่า และ สะโพก แล้วทำการเดินบนลู่วิ่ง



รูปที่ 3.7 ตัวอย่างผู้ถูกทดลอง

ระบบ YCbCr จะแบ่งข้อมูลของจุดภาพตามการรับรู้ของประสาทตาของคน โดยจะแยกข้อมูลเกี่ยวกับความสว่างออกจากข้อมูลเกี่ยวกับระดับสี นั่นคือใช้องค์ประกอบ Y เก็บค่าความสว่างของจุดภาพ ส่วน Cb คือค่าความแตกต่างระหว่างค่าสีน้ำเงินกับค่าความสว่าง (B-Y) องค์ประกอบ Cr คือค่าความแตกต่างระหว่างค่าสีแดงกับค่าความสว่าง (R-Y)

ดังนั้นเราจึงใช้ช่องสัญญาณ Cr เท่านั้น เพื่อหาโทนสีแดงที่เด่นในภาพ เพื่อประโยชน์ในการหาค่าพิกัดของจุดที่เราต้องการ

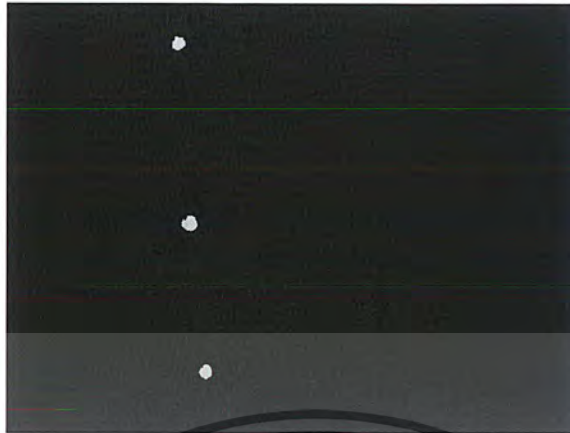
ภาพที่ได้ในแต่ละเฟรมในช่วงเวลาเดียวกัน จะเข้าสู่การ Contrast เพื่อให้ภาพมีความคมชัดยิ่งขึ้น



รูปที่ 3.8 ภาพที่ผ่านการ Contrast

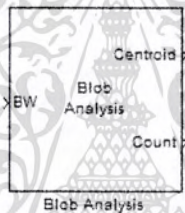
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากนั้นนำภาพที่ผ่านการ Contrast มาทำการ Thresholds



รูปที่ 3.9 ภาพที่ผ่านการ Thresholds

เมื่อได้ภาพที่ผ่านการ Thresholds แล้ว ที่ทำให้ภาพสามารถโชว์จุดที่เราต้องการได้ จากนั้นนำภาพที่ได้มาหาค่า Co-ordinate ของทั้ง 3 จุดโดยผ่านกระบวนการ Blob Analysis



รูปที่ 3.10 Blob Analysis

โดย Blob Analysis จะหาค่า Co-ordinate ของสีขาวต่างๆในภาพ โดยแสดงผลเป็นจุดศูนย์กลางของจุดนั้นๆ ดังนั้นเพื่อเป็นการป้องกันภาพที่มีสัญญาณรบกวน จำเป็นต้องกำหนดขนาดของ Pixel เพื่อให้ Blob Analysis เลือจุดที่สนใจได้อย่างถูกต้อง

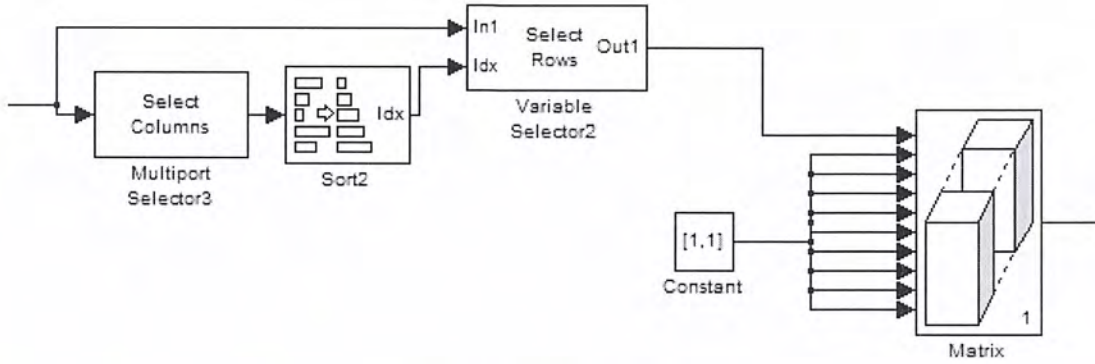
Specify minimum blob area in pixels: 10

Specify maximum blob area in pixels: 200

รูปที่ 3.11 การกำหนดขนาด Pixel เพื่อให้ระบบ ทำงานได้อย่างถูกต้อง

เมื่อผ่าน Blob Analysis แล้ว จะได้ Co - ordinate ของจุดทั้ง 3 จุด แต่ข้อมูลที่ได้จะไม่มีการเรียงลำดับหรือไม่ทราบได้ว่าพิกัดใด เป็นพิกัดบริเวณสะโพก เข้า หรือข้อเท้า ดังนั้นจึงต้องทำการเรียงข้อมูลเพื่อให้ทราบถึงพิกัดต่างๆ ที่บริเวณใดๆ

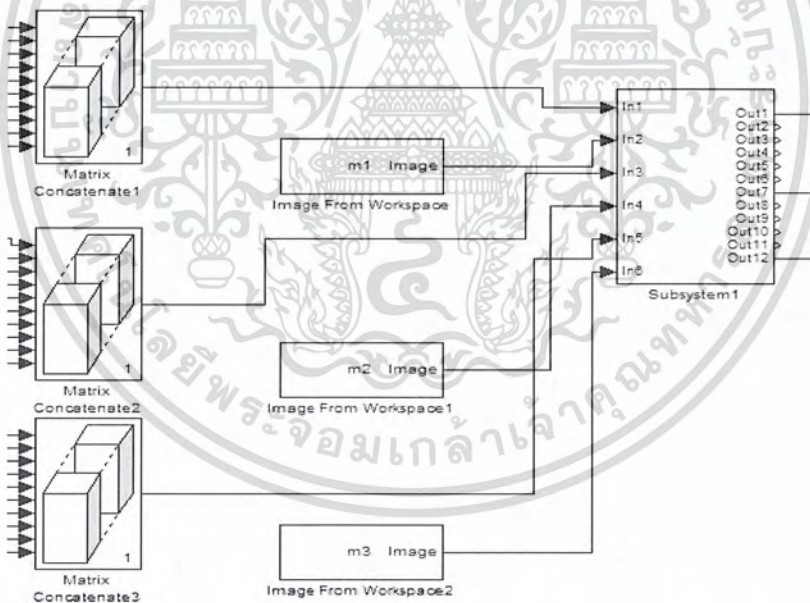
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.12 การลำดับข้อมูล

จากรูปที่ 3.8 ข้อมูลที่ผ่านการลำดับข้อมูลแล้ว จะถูกนำไปเข้าสู่กระบวนการแปลงกลับจากภาพ 2D เป็นภาพ 3D

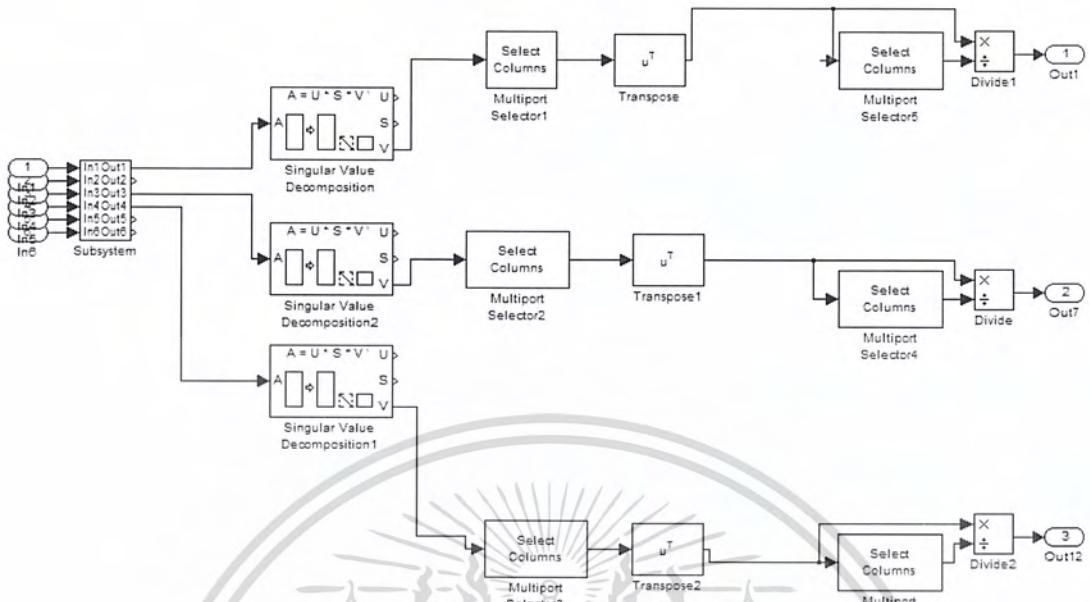
3.3 2D to 3D



รูปที่ 3.13 การแปลงภาพ 3D

ข้อมูลนี้ผ่านกระบวนการหา Co-ordinate และเรียงลำดับข้อมูลแล้ว จะถูกนำมาแปลงกลับเป็นพิกัดในระบบ 3D โดยทำการตั้งสัมประสิทธิ์การแปลงกลับที่ได้จากการ Calibrate มาเข้าสมการ DLT

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 สมการการแปลงกลับ

รูปที่ 3.9 แสดงถึงสมการการแปลงกลับ ที่ได้ออกแบบมาจากสมการ DLT เมื่อผ่านกระบวนการนี้ จะทำให้ทราบถึงพิกัดใน 3 มิติ (X,Y,Z) แล้วนำพิกัดที่ได้ ไปใช้งานต่อไป



รูปที่ 3.14 ภาพบอกพิกัดใน 3 มิติ

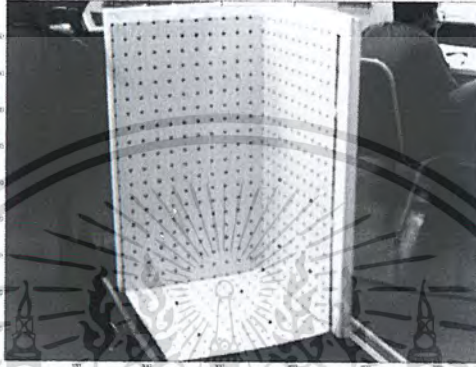
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

### ผลการทดลอง

#### 4.1 ผลการปรับเทียบกล้อง

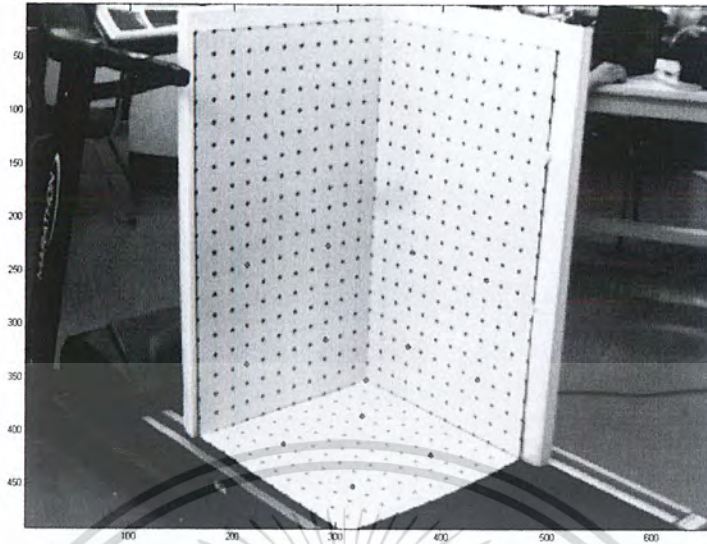
ในการปรับเทียบกล้อง จะทำการรับค่าจากอุปกรณ์ปรับเทียบ โดย มุมกล้อง 1 มุม จะได้ รูป 1 รูป และจะทำการรับค่า มา 13 จุด โดยจะทำทั้งหมด ทั้ง 3 กล้อง



รูปที่ 4.1 นำภาพมาพล็อตจุด 13 จุด เพื่อหา พิกัด  $u, v$  ของกล้องที่ 1

พิกัด world coordinate (X Y Z)			พิกัดในรูป (u v)	
X	Y	Z	u	v
15	0	15	308.6654	321.90935
15	0	40	306.7996	233.62523
40	0	15	219.7297	337.54299
40	0	40	214.7543	246.5
0	15	15	381.431	333.86449
0	15	40	380.8091	242.8215
0	40	15	423.7221	368.81028
0	40	40	424.344	273.16916
15	15	0	330.4329	396.39907
15	40	0	369.6144	437.78224
40	15	0	240.8752	412.95234
40	40	0	271.9716	456.17477
0	0	0	358.4197	365.13178

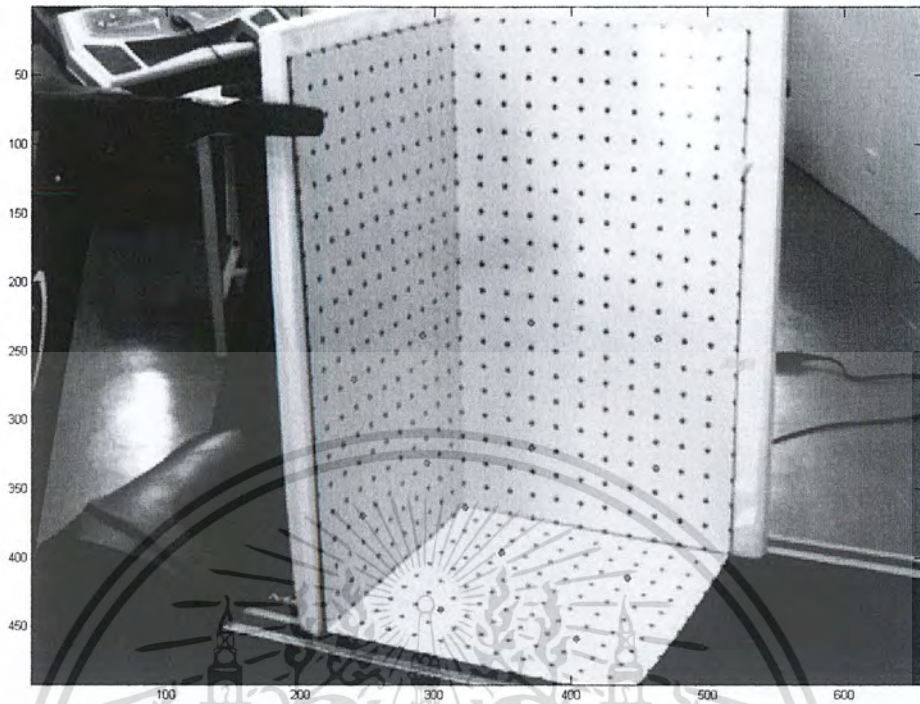
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 น้าภาพมาพล็อตจุด 13 จุด เพื่อหา พิกัด  $u v$  ของกล้องที่ 2

พิกัด world codinate (X Y Z)			พิกัดในรูป (u v)	
X	Y	Z	u	v
15	0	15	290.0076	314.5523
15	0	40	291.8733	227.1879
40	0	15	215.3762	338.4626
40	0	40	214.7543	244.6607
0	15	15	367.1267	320.9897
0	15	40	371.4802	232.7056
0	40	15	434.2949	352.257
0	40	40	441.758	257.5355
15	15	0	324.8355	387.2028
15	40	0	390.7599	423.0682
40	15	0	250.2042	412.9523
40	40	0	314.8847	452.4963
0	0	0	329.189	352.257

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



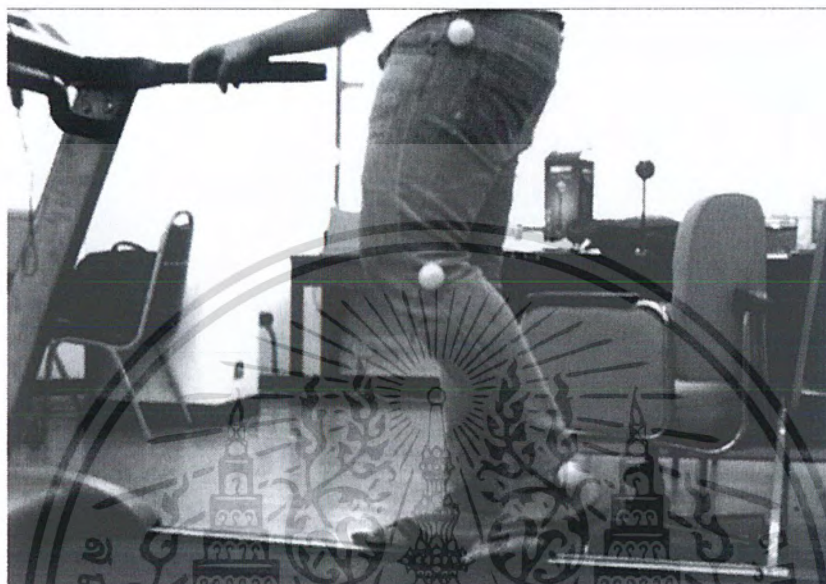
รูปที่ 4.3 นำภาพมาพล็อตจุด 13 จุด เพื่อหา พิกัด  $u v$  ของกล่องที่ 3

พิกัด world codinate (X Y Z)			พิกัดในรูป (u v)	
X	Y	Z	u	v
15	0	15	295.6049	331.1056
15	0	40	291.8733	238.2234
40	0	15	247.7164	370.6495
40	0	40	241.4972	270.4103
0	15	15	370.8582	320.0701
0	15	40	370.2363	228.1075
0	40	15	462.2817	334.7841
0	40	40	463.5255	240.9822
15	15	0	349.7127	396.3991
15	40	0	440.5142	413.872
40	15	0	304.9338	438.7019
40	40	0	403.1985	458.9336
0	0	0	322.3478	362.3729

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 4.2 ผลของการประมวลผลภาพ

เมื่อมีการรับภาพเข้ามาภาพจะมีลักษณะเป็นภาพสี โดยมีค่าของ YBbCr เป็นองค์ประกอบของ วีดีโอ จากนั้น จะทำการแยกค่า โดย จะรับเฉพาะ ค่า Cr ซึ่งเป็นค่าความแตกต่างระหว่างสีแดงกับ ค่าความสว่าง

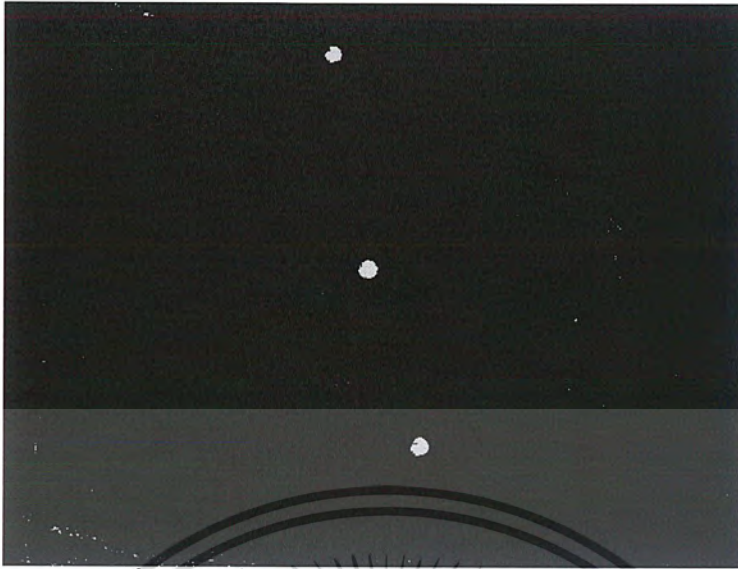


รูปที่ 4.4 รูปการรับภาพปกติของกล้อง



รูปที่ 4.5 รูปการรับภาพที่รับค่าเฉพาะ ค่าของ Cr

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

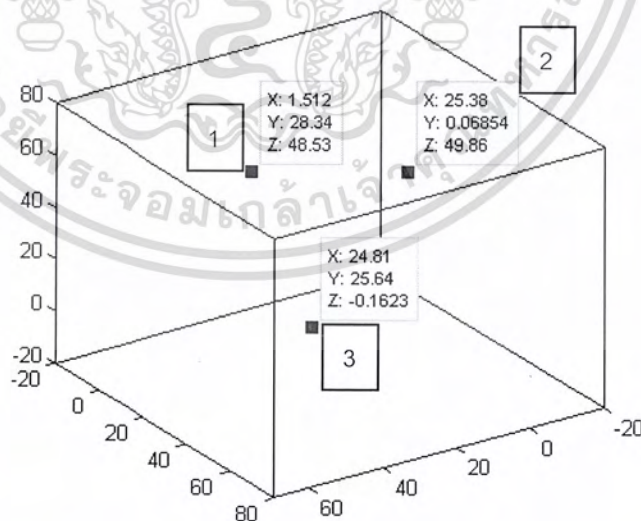


รูปที่ 4.6 รูปที่นำค่าของ Cr มาทำเป็นภาพ ขาว-ดำ

### 4.3 การแปลงภาพ 3 มิติ

การปรับเทียบกล้องเป็นการหาประมาณเมทริกซ์  $M$  จากจุดในฉาก 3 มิติที่รู้ตำแหน่ง และภาพที่เกิดขึ้นของจุด เมื่อทำการปรับเทียบกล้องแล้วจะนำประมาณเมทริกซ์ที่ได้ไปเข้ากระบวนการทางคณิตศาสตร์โดยใช้สมการการแปลงเชิงเส้นตรง

เมื่อเราทำการทดสอบโปรแกรม สุ่มพล็อตจุดค่าที่แท้จริงที่เราไม่ทราบค่า แล้วนำผลที่ได้มาสร้างเป็นภาพ 3 มิติ แล้วนำมาตรวจสอบกับค่าจริงที่วัดได้



รูปที่ 4.7 รูปการสุ่มจุดเพื่อทดสอบ error ของการแปลงภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

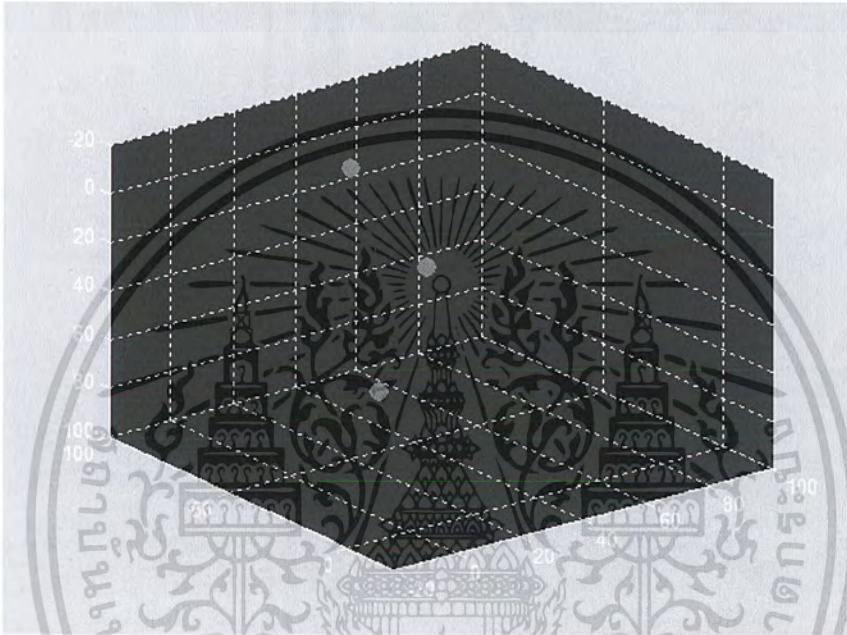
จากจุดที่ได้ เมื่อต้องการหาค่า error จะทำได้ดังนี้

$$\text{จุดแรก จุดที่แท้จริง คือ } 0 \ 30 \ 50 \text{ หาค่าโดย } \sqrt{(1.512-0)^2 + (28.34-25)^2 + (48.53-50)^2} = 3.95$$

$$\text{จุดสอง จุดที่แท้จริง คือ } 25 \ 0 \ 50 \text{ หาค่าโดย } \sqrt{(25.38-25)^2 + (0.068-0)^2 + (49.86-50)^2} = 0.41$$

$$\text{จุดสาม จุดที่แท้จริง คือ } 25 \ 25 \ 0 \text{ หาค่าโดย } \sqrt{(24.81-25)^2 + (25.64-25)^2 + (-0.16-0)^2} = 0.68$$

และเมื่อมีการรวมภาพของทั้ง 3 กล้องจากการทดลองการรับภาพมาแล้ว เมื่อนำมาแปลงเป็นภาพ 3 มิติ จะได้ซึ่งเป็นภาพเคลื่อนไหวต่อเนื่อง



รูปที่ 4.8 รูปการแปลงภาพ 3 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### สรุปผลการทดลอง

#### 5.1 สรุปและวิจารณ์ผลการทดลอง

โครงการนี้นำเสนอการออกแบบและกระบวนการจับภาพเคลื่อนไหว โดย โปรแกรม Matlab กระบวนการจับการเคลื่อนไหวนี้ ประกอบไปด้วยส่วนการรับภาพ ซึ่งต้องใช้ในการรับภาพโดยใช้ค่าของอัตราภาพที่สูงซึ่งเป็นการลดค่าความผิดพลาดได้อีกวิธีหนึ่ง โดยใช้กล้องทั้งหมด ทั้งหมด 3 ตัว ในส่วนของการปรับเทียบกล้องเพื่อหาค่า สัมประสิทธิ์ ของมุมกล้อง ซึ่งเป็นค่าที่จำเป็น ในการนำไปประมวลผล ส่วนการประมวลผลภาพเพื่อใช้หาค่าพิกัดของจุดสนใจ เป็นการหาพิกัดของอุปกรณ์บ่งชี้ตำแหน่ง โดยใช้เทคนิคการหาค่าความแตกต่างระหว่างสีแดงกับค่าแสงสว่าง(Cr) จากนั้นนำไปทำให้เป็นภาพขาว-ดำ(BW) ซึ่งจะได้จุดที่สนใจออกมา 3 จุด นำจุดที่ได้ เพื่อที่จะนำค่าเหล่านั้น มาประมวลผล เป็น 3 มิติ

จากการทำงานของชุดอุปกรณ์บันทึกภาพเคลื่อนไหว ใช้กล้อง 3 ตัว ในเวลาเดียวกันโดย โปรแกรม Matlab ดังกล่าว สามารถใช้กับมนุษย์ได้จริง ผลที่ได้นั้นเป็นไปตามการออกแบบซึ่งเป็นการเคลื่อนที่ของโมเดล 3 มิติ ที่ได้จากการรับค่าจากกล้อง

#### 5.2 ข้อเสนอแนะในการพัฒนางานวิจัยต่อไป

ในช่วงระหว่างการทำทดลองสร้างระบบการจับภาพเคลื่อนไหว โดยใช้โปรแกรม matlab นั้น ช่วงแรกประสบปัญหาของการรับภาพ ซึ่งภาพของกล้องแต่ละตัวทำการบันทึกภาพที่ไม่พร้อมกัน และอัตราภาพไม่ได้ตามที่ต้องการในขณะเก็บภาพ เป็นผลทำให้เป็น ผลทำให้ภาพที่ได้ไม่สามารถนำไปประมวลผลต่อได้ เพื่อการแก้ไขปัญหาดังกล่าว จึงควรเปลี่ยนและทดลองโปรแกรมอื่นๆ ที่มีฟังก์ชันที่สูงกว่าเพื่อเสถียรภาพของกระบวนการจับภาพเคลื่อนไหวที่ดีขึ้น

## กิตติกรรมประกาศ

การทำโครงการทาง Motion capture นี้สำเร็จได้ด้วยดี ทั้งนี้ได้รับความช่วยเหลือ ให้คำปรึกษา ทั้งทางด้านวิชาการและการปฏิบัติการตลอดจนช่วยเหลือสนับสนุนทางด้านต่าง ๆ จาก ดร.สุรเดช ตรีไตรลักษณ์

ขอกราบขอบพระคุณ ผศ.ดร.กิตติพล ชิตสกุล กรรมการสอบที่กรุณาให้ความช่วยเหลือ แนะนำสิ่งที่ดีที่สุดและตรวจแก้ไขโครงการฉบับนี้ให้ถูกต้องและสมบูรณ์ยิ่งขึ้น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บรรณานุกรม

- [1] รศ.ดร.ชูชาติ ปิณฑวิรุจน์ (2551) การประมวลผลภาพดิจิทัลขั้นสูงด้วย MATLAB.กรุงเทพฯ ห้างหุ้นส่วนจำกัด วี.เจ.พรีนติ้ง
- [2] ดร.ปริญญา สงวนสัตย์ (2553) คู่มือ matlab ฉบับสมบูรณ์.บริษัท ไอดีซี พรีเมียร์ จำกัด
- [3] International Archives of Photogrammetry, remote and Spatial Information Sciences, Vol,XXXVIII, Part 5, Commission V Symposium, Newcastle upon Tyne, UK. 2010
- [4] Daniel Bardsley / Bai Li Gabor Wavelet Based ,Correspondence Measure Estimation Methods, Aalborg University - Institute for Media Technology, Computer vision and graphics, Laurupvang 15, 2750 Ballerup, Denmark, www.aau.dk
- [5] <http://www.kwon3d.com/theory/dlt/dlt.html>
- [6] [http://en.wikipedia.org/wiki/Direct\\_linear\\_transformation](http://en.wikipedia.org/wiki/Direct_linear_transformation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก ก  
ทฤษฎีที่ใช้ในการใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

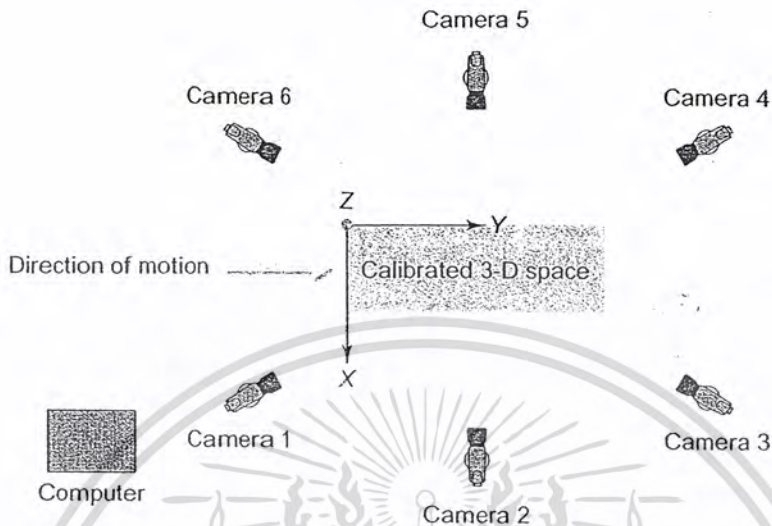


Figure 1.3 Typical multicamera experimental setup for a 3-D kinematic analysis of a walking stride.

of control points must contain at least six noncoplanar points ( $n \geq 6$ ), but usually many more than that are used. In fact, up to 20 control points are often used. The images of these control points are then captured by all cameras in a multicamera setup and digitized to produce 2-D coordinates of each point in each camera view. From the 2-D coordinates of the  $n$  noncoplanar control points, a set of  $2n$  equations is developed for each camera view. This set of equations is then solved for 11 DLT parameters ( $L_1$  through  $L_{11}$ ) relative to each camera. The DLT parameters reflect the relationship between 3-D space and the 2-D camera image. The equations for one camera are

$$\begin{aligned} x_i + L_1 X_i + L_2 Y_i + L_3 Z_i + L_4 \\ L_9 x_i X_i + L_{10} x_i Y_i + L_{11} x_i Z_i = 0 \end{aligned} \quad (2.14)$$

and

$$\begin{aligned} y_i + L_5 X_i + L_6 Y_i + L_7 Z_i + L_8 \\ L_9 y_i X_i + L_{10} y_i Y_i + L_{11} y_i Z_i = 0 \end{aligned} \quad (2.15)$$

where  $i$  is the number of the control point;  $x_i$  and  $y_i$  are the digitized 2-D coordinates for the  $i$ th control point;  $X_i$ ,  $Y_i$ , and  $Z_i$  are the known object space coordinates of the  $i$ th control point; and  $L_1$  to  $L_{11}$  are the DLT parameters.

The  $2n$  set of equations with 11 unknown parameters forms an overdetermined solution (i.e., more equations than unknowns) and is solved using a least-squares technique (Miller, Shapiro, and McLaughlin 1980). The DLT parameters contain implicit information relating to the camera orientations and locations. Once these DLT parameters are known, the control points can be removed from the laboratory. The markers that are placed on a subject can then be

captured by the same cameras as the subject moves through the calibrated object space. When these markers have been digitized in 2-D space in at least two cameras, the 2-D coordinates of the landmarks and the DLT parameters can be entered into the equations to solve for the 3-D coordinates using the same equation. In this case, however, we solve for  $X$ ,  $Y$ , and  $Z$ , the 3-D coordinates of the points.

The DLT method is not the only method used to determine 3-D coordinates from a set of 2-D coordinates. Other methods include nonlinear transformations. The techniques of all are similar in that known locations of markers in 3-D space are combined with their digitized images in 2-D space to determine a set of nonlinear equations from which the locations of unknown markers can be determined.

## COORDINATE SYSTEMS

Two coordinate systems must be defined when conducting a 3-D analysis. These are the *global* or *laboratory coordinate system* (GCS) and the *local coordinate system* (LCS).

### GLOBAL OR LABORATORY COORDINATE SYSTEM

The GCS is also referred to as the *inertial reference system*, and it is determined when the object space is defined during the 3-D data capture. This coordinate system—generally a right-handed, orthogonal system with an arbitrary origin—defines the fixed coordinate system in the laboratory from which all positions are ultimately derived. In this chapter, the GCS is designated using uppercase letters with the

ever, they are presented only briefly in appendix D of this book. Readers are urged to refer to more detailed presentations of the concepts in other sources.

## MATRICES

Many of the computations done in 3-D kinematics can be accomplished relatively efficiently using *matrix algebra*. A *matrix* is any rectangular array of numbers. Each number in the array is an *element*. If the array has  $m$  rows and  $n$  columns, the matrix is referred to as an  $m \times n$  matrix. Each element of the matrix can then be identified by its row and column positions. The number of rows and columns in a matrix denotes the order of the matrix. For example, the  $3 \times 3$  matrix  $[A]$  can be written as

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (2.8)$$

The element  $a_{23}$  is the element in the second row and the third column, or the element in the (2,3) place. Generally, each element in matrix  $[A]$  can be referred to as  $a_{ij}$ , where  $i$  is the row number and  $j$  is the column number. A matrix having only one row with several columns is a *row matrix*. For example,

$$[A] = [a_{11} \quad a_{12} \quad a_{13}] \quad (2.9)$$

Conversely, a matrix can have several rows but only one column. This type of matrix is referred to as a *column matrix*. For example,

$$[A] = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \quad (2.10)$$

We will encounter a number of special matrices in this chapter. A *square matrix* has the same number of rows and columns. A *diagonal matrix* is a square matrix in which the elements on the diagonal ( $a_{11}$ ,  $a_{22}$ ,  $\dots$ ,  $a_{nn}$ ) are nonzero and the others are zero. This matrix  $[A]$  is a square, diagonal matrix:

$$[A] = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \quad (2.11)$$

An *identity matrix* is a square matrix whose diagonal elements equal 1. Designated as  $[I]$  matrices, they are written like this:

$$[I] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.12)$$

The *transpose* of matrix  $[A]$  is designated as  $[A]^T$  and is one in which the rows and columns are interchanged.

$$\text{If } [A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad (2.13)$$

$$\text{then } [A]^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

## MATRIX OPERATIONS

A 3-D analysis involves many operations concerning matrices with which you must become familiar. These operations are presented in appendix E of this book. As with vector operations, readers are urged to refer to more detailed presentations of the concepts in other sources.

## COLLECTION OF THREE-DIMENSIONAL DATA

All current multicamera motion-analysis systems can collect 3-D data. The orientation and arrangement of the cameras are not as critical in a 3-D setup as in a 2-D setup (figure 2.3).

Each of the cameras only has to provide a set of 2-D coordinates, with each digitized landmark appearing in at least two cameras. From these sets of 2-D coordinates, 3-D spatial coordinates are generated. In most cases, the *direct linear transformation* (DLT) method (Abdel-Aziz and Karara 1971) is used to calculate the 3-D coordinates. The DLT method establishes a linear relationship between the 2-D camera coordinates of each body landmark and the representation of this landmark in 3-D space.

The technique to do this is relatively straightforward and, in this chapter, we only outline the methodology. More detailed explanations of DLT methods are presented in papers by Miller, Shapiro, and McLaughlin (1980); Shapiro (1978); and Marzan and Karara (1975). A set of  $n$  points, called *control points*, defines an *object space* or a coordinate system in the laboratory. The control points have known locations in real units in 3-D space. The set

## DIRECT LINEAR TRANSFORMATION BASED PHOTOGRAMMETRY SOFTWARE ON THE WEB

B. Molnár<sup>a</sup>

<sup>a</sup> Department of Photogrammetry and Geoinformatics, BME, Műegyetem rkp 3., Budapest, H-1111 – molnar.bence@fmt.bme.hu

Commission V, SS-3

**KEY WORDS:** Photogrammetry, Direct Linear Transformation, Huber-method, gross errors, weight matrix, web application

### ABSTRACT:

Internet offers a new direction for software development and also for photogrammetry. The affordable commercial digital cameras are widely available and their geometrical resolution is good enough to perform photogrammetric measurements. Seeing this, a widely accessible, web based photogrammetry application could be a new development direction for photogrammetry. For photogrammetric calculations Direct Linear Transformation is a well useable method as it does not need calibrated camera. The potential users are in the lack of experience, therefore the calculation method should be prepared to help them to achieve high accurate results. For example, gross errors in measurements cannot be avoided but the impacts of gross errors should be avoided. This study shows a technique how to reduce effects of gross errors with Huber-method and discusses the investigation for results.

### 1. INTRODUCTION

Availability of classical photogrammetry was limited due to expensive devices for example special cameras and stereocomparators. Nowadays everybody can afford a commercial camera with relatively high geometric resolution. Photogrammetric post-processing could be performed on a simple PC. Remaining limiting factors for users are the lack of experience and affordable software. This paper presents a solution to avoid these difficulties. A web based photogrammetry application was developed that allows to access the software by a single web browser. For calculation Direct Linear Transformation (DLT) is used, which enables the use of non calibrated digital cameras. The lack of user's expertise has to be considered during the calculation method. For example, gross errors have bad effects for classical least squares method, and inexperienced users can easily cause gross errors during digitizing. The effect of gross errors is minimized by Huber-method. The developed photogrammetry application can be a spatial measurement system not only for professionals, but for the wide public. To achieve an inexpensive, reliable, and widely available application, only open source and free software were used. Another goal was to apply all web standards to make the web application available in all graphical web browsers. In addition to the mentioned goals, accuracy issues should be assessed as well.

#### 1.1 Practical uses

There are two major reasons that motivated the software development. First, if somebody wants to buy or rent a real estate, he takes photo and the decision is made at home. These photos can be used for spatial measurements, why aren't they used? With photogrammetry the 3D model and a base plan can be created.

The other application area is education. For students it is an impressive application, they can try it at home, and also the calculation algorithm can be presented. Photogrammetry can be

presented in practice without the "black-box" calculation effect that usually can be experiences in case of the commercial software.

The web based photogrammetry is not a new idea, Grussenmeyer developed a web based photogrammetry application in Java programming language and discussed the possibilities of an internet based photogrammetry software (Grussenmeyer et al, 2001). However, this software still have limitations on software environment.

### 2. WEB BASED APPLICATION

Even more software is available in a web based version. For example many e-mail, office, route planning applications working on the internet. These applications usually are free to use, don't need software installations and available everywhere; that's why they are getting even more popular. This inspired the development of a web based photogrammetry application. The advantages and disadvantages of a web based application are shown in the Table 1.

Advantages	Disadvantages
No limit of operating system (multiplatform)	Broadband internet connection required
No need of software installation	Large data amount causes high network traffic
No specific rights are needed	Browser limitations
No need to upgrade, always the newest version is available	Private data handling (images)
Location and time independent	No 3D engine available
Regular user interface	
Inexpensive	

Table 1. Advantages and shortcomings of web based applications

On the development side there are some difficulties. In digital photogrammetry images have large size that must be transmitted over the internet in case of the web based approach. The transfer has to be well planned without data and resolution loss. For example, at the beginning of the work, only minimal part of

the images should be transferred, the other ones should be downloaded in the background while digitizing is made.

Another difficulty is the variety of browser types (i.e. Firefox, Internet Explorer, Opera, Google Chrome, etc.). They all work in a different way, it's not easy to create a browser independent web application. However, it would be beneficial, if one can work in the same user interface, even if he has to use Internet Explorer in the office, but Firefox is installed at home. This difficulty exist also in the classical desktop applications, it's difficult to develop a program that works on Windows, MacOS or Linux.

For development, the main issue was to use free and open source programs to reduce costs, and to ensure a dynamically developing programming background. This is true for the frontend and calculation part of the application too. The web page parsing is made by PHP, database backend is MySQL. The user interface is done with HTML, Javascript, SVG and VML and all sites apply the W3C web standards. The processing is done with Octave mathematical programming language that is similar to Matlab. The processing is done on the server, but in order to reduce server load, as much calculation as possible is processed on the client side. For example, 3D visualisation is done in the browser with SVG/VML and Javascript.

### 3. CALCULATION

#### 3.1 Direct Linear Transformation (DLT)

The commercial digital cameras are not calibrated. Neither the internal calibration data, nor the optical distortions are available that has negative effect on the final accuracy. The DLT enables performing calculations with high accuracy without calibration data (Abdel-Aziz and Karara, 1971). DLT describes a direct connection between 3D coordinates and image coordinates. Another advantage of DLT is the application of linear equations, which make the calculation algorithms more efficient. DLT defines 2 equations in each point, and has 11 parameters that means minimum 6 control points are needed for successful calculation. If all optical distortions are omitted, the equations are as follows (Karara, 1989):

$$\begin{aligned} L_1 X + L_2 Y + L_3 Z + L_4 - x L_9 X - x L_{10} Y - x L_{11} Z - x &= 0 \\ L_5 X + L_6 Y + L_7 Z + L_8 - y L_9 X - y L_{10} Y - y L_{11} Z - y &= 0 \end{aligned} \quad (1)$$

where  $L_i$  = DLT parameters  
 $x, y$  = image coordinates  
 $X, Y, Z$  = 3D coordinates

The calculation is processed in two steps, first the DLT parameters are estimated by control points for each image (orientation), then the unknown points can be calculated if they appear on more than 2 oriented image. Least Squares Method (LSM) is used (Detrekői, 1999) for adjustment with some modifications.

#### 3.2 Huber-method

If gross errors are present in measurements (image coordinates), LSM is not a good choice for adjustment. Gross errors cannot be avoided in photogrammetry, especially by inexperienced users. The main reason for gross errors is the false point identification on images. Estimations, which reduce the effect of gross errors are named robust estimations. A way for reducing

effects is to modify weight matrices, gross errors weight should be reduced. The algorithm is follows: first an initial adjustment is done with a unit matrix. A new weight matrix is defined based on the residual vector of previous adjustment, and this step is iterated.

The question is, how new weight matrix is defined. Several object functions were developed for an efficient filtering. The developed application uses Huber-method, since its object function is convex (Huber, 1981), what ensures, the estimation will be convergent. The object function is:

$$\rho(v) = \begin{cases} \frac{v^2}{2} & , |v| \leq a \\ a(|v| - \frac{a}{2}) & , |v| > a \end{cases} \quad (2)$$

where  $\rho(v)$  = object function  
 $v$  = residual  
 $a$  = threshold

The weight function is the derivation of object function and divided by  $x$ . As shown in Figure 2., the residuals smaller than the threshold get 1 for weight (this means, normal distribution is supposed), others get smaller values. During iterations gross error's weight decrease, and will have minimal effect on final result.

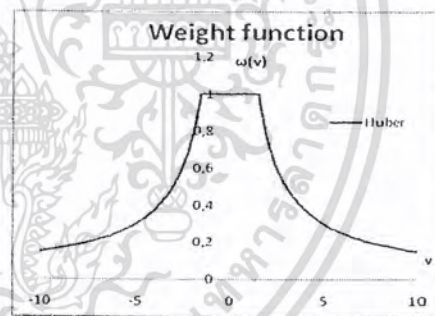


Figure 2. Huber-method weight function

The threshold value is  $1.5\sigma$  according to Huber (Huber, 1981), but it can be changed for a suitable value, for example  $3\sigma$ , as in geodesy usual. The photogrammetry web application uses the value 3 pixels and the number of iterations is 6. Experience proved that this is appropriate.

$$\omega(v) = \begin{cases} 1 & , |v| \leq a \\ \frac{a}{|v|} & , |v| > a \end{cases} \quad (3)$$

where  $\omega(v)$  = weight function  
 $v$  = residual  
 $a$  = threshold

Huber-method makes LSM a robust and widely applicable estimation. Using the Huber-method, the previously used algorithms can be kept, and calculation time doesn't rise exceptionally.

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การสงวนไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DLT describes one equation for x and one for y image coordinates. The coordinate pairs are not independent, thus if one of them has gross error, the other coordinate is also wrong. The weight must be the same for coordinate pairs.

**3.3 Huber-method in practice**

The performance of the calculation was tested in practice. A critical point of estimation is the first adjustment, if gross errors are dominant, the residual vector will be false. In this case the users have to be warned. It can be done if the RMS values of the first adjustment are high.

If RMSs are low, the calculation can be done, and Huber-method can be tested. In the particular test 4 images were used. The DLT parameters were determined by 20 control points and one unknown point were determined based on the DLT parameters. This process was performed several times, with different level of gross errors (test 1-3) on one point. The results were compared with that of without gross errors. A comparison was made also for extremely gross errors with and without Huber-method (test 4). Table 2 presents the image coordinates of point 8, in different tests, first column is without artificial values (clean measurements), others with gross errors (errors are in the last two rows).

[pixel]	clean	test 1	test 2	test 3	test 4
x	455.0419	475.0419	475.0419	575.0419	575.0419
y	588.4174	588.4174	548.4174	948.4174	948.4174
Δx	-	20	20	120	120
Δy	-	0	-40	360	360

Table 2. Test coordinates and differences [pixel]

Table 2 shows the final weights of measurements, the point 8 is with gross error, it has in test 1,2,3 very low value. The test 4 is without Huber-method. The weights of clean measurements are unchanged, the highest change is 0.141. Here are x and y coordinates are handled as independent values in order to see the real effect of Huber-method.

	clean		test 1		test 2		test 3		test 4	
1	0.207	0.865	0.217	0.942	0.207	0.724	0.198	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1
6	0.584	1	0.583	1	0.606	1	0.516	1	1	1
7	1	0.587	1	0.569	1	0.613	1	0.519	1	1
8	1	0.152	1	0.15	1	0.079	0.026	0.009	1	1
9	1	0.802	1	0.821	1	0.863	1	0.756	1	1
10	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1
12	0.761	1	0.739	1	0.686	1	0.818	1	1	1
13	0.367	0.548	0.363	0.528	0.363	0.584	0.405	0.529	1	1
14	0.506	1	0.480	1	0.470	1	0.462	1	1	1
15	0.494	0.602	0.482	0.575	0.453	0.545	0.464	0.545	1	1
16	1	0.484	1	0.471	1	0.515	1	0.465	1	1
17	1	1	1	1	1	1	1	1	1	1
18	0.56	0.461	0.539	0.455	0.606	0.459	0.579	0.415	1	1
19	0.184	1	0.182	1	0.187	1	0.179	1	1	1
20	1	0.252	1	0.249	1	0.245	1	0.265	1	1

Table 3. Weights of image coordinates

The maximum change of DLT parameters with Huber-method is 7% for L<sub>4</sub>, but all other values remain under 1%. Without Huber-method changes can be 158% (294% for L<sub>4</sub>)! The L<sub>4</sub> and L<sub>8</sub> are special parameters in the DLT equations, they represent a constant image coordinate offset.

	clean	test 1	test 2	test 3	test 4
L <sub>1</sub>	7.130589	7.121748	7.128731	7.079486	4.995442
L <sub>2</sub>	0.996978	0.998059	0.995439	0.97812	0.261669
L <sub>3</sub>	-6.81928	-6.82449	-6.83241	-6.83772	-7.51298
L <sub>4</sub>	-101.799	-100.813	-100.529	-94.4567	191.6648
L <sub>5</sub>	-0.45804	-0.46	-0.45406	-0.46558	-0.80934
L <sub>6</sub>	-7.66184	-7.65652	-7.66018	-7.62773	-6.26042
L <sub>7</sub>	-4.71138	-4.71009	-4.70971	-4.70847	-4.90712
L <sub>8</sub>	2568.476	2567.238	2567.051	2559.208	2213.25
L <sub>9</sub>	-0.00047	-0.00047	-0.00047	-0.00048	-0.0007
L <sub>10</sub>	0.000312	0.000312	0.000312	0.0003	-0.00018
L <sub>11</sub>	-0.00312	-0.00313	-0.00313	-0.00315	-0.00442

Table 4. Changes of DLT parameters

While DLT parameter changes are around 1%, the unknown point determination with Huber-method is very accurate; changes are under 0.025 mm. That means, the method can be well applied for photogrammetry, and the effect of gross errors can be reduced. (Test images had 3888x2592 pixels, object distance was ~3m.)

[mm]	clean	test 1	test 2	test 3	test 4
X	158.117	158.117	158.117	158.124	158.455
Y	204.591	204.582	204.568	204.585	205.446
Z	-0.12	-0.11	-0.095	-0.128	-2.316
ΔX	-	0	0	-0.007	-0.338
ΔY	-	0.009	0.023	0.006	-0.855
ΔZ	-	-0.01	-0.025	0.008	2.196

Table 5. Changes of unknown point coordinates [mm]

**3.4 Accuracy**

The final accuracy of measurements depends on multiple factors. An important one is the accuracy of control points' spatial coordinates. The critical factor is the digitizing accuracy. In digital photogrammetry subpixel digitizing is used to increase accuracy. It means that all pixels are divided to more pixels, and intensity is calculated for all subpixels depending on the neighbour pixels' intensity values. Modern browsers (e.g. Firefox) use this method if a photo is zoomed: the developed application uses the same method if available in browser. The digitizing gross errors are filtered as described above. As mentioned before, the user's experience is also an important factor. It also covers the photogrammetric network planning, i.e. the determination of camera positions and control point positions. The accuracy also depends on the quality of photos, the important factors are: geometric and radiometric resolution, optical distortion, and image sharpness.

**4. CONCLUSION**

With a web based photogrammetry application the number of potential photogrammetric users can be increased. Users only need to have a simple digital camera and a computer with a web browser and broadband internet connection. Measurements can be performed anytime all over the world, since user data are stored on server. Users can work fast and be effective due to well planned data transfer and calculation methods. Accuracy depends on users' experience, but the available accuracy is comparable to that of other photogrammetric methods. DLT offers the opportunity to use non calibrated cameras, and an effective linear equation system for calculation. The effects of gross errors can be reduced with Huber-method, thus the estimation will be convergent. The test shows that the gross errors' effect will be minimized even if false point identification is done. Least Squares Method combined with Huber-method can be reasonably applied for photogrammetric calculation.

## 5. REFERENCES

Abdel-Aziz Y. I, Karara H.M., 1971. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates. *ASP Symposium on Close-Range Photogrammetry*, Falls Church, VA, pp. 1-18.

Detrekői Á., 1999. *Adjustment calculation* (in Hungarian). Tankönyvkiadó, Budapest, pp. 1-685.

Fekete K., 2006. Close Range Photogrammetry and some Medical Practice (in Hungarian). PhD Thesis, Budapest.

Grussenmeyer P., Drap P., 2001. Possibilities and limits of Web photogrammetry - Experiences with the ARPENTEUR web based tool. *Photogrammetric Week 01*, Wichmann Verlag, Heidelberg

Huber P.J., 1981. *Robust Statistics*. John Wiley & Sons, New York.



# Efficiency of non-linear lens distortion models in biomechanical analysis of human movement

Serdar Aritan\*

Biomechanics Research Group, School of Sports Science and Technology, Hacettepe University, Ankara 06800, Turkey

## ARTICLE INFO

### Article history:

Received 2 August 2007

Received in revised form 2 August 2009

Accepted 27 January 2010

Available online xxx

### Keywords:

Image processing

Lens distortion

Distortion correction

Distortion mapping

DLT

Movement analysis

## ABSTRACT

Accurate camera calibration is required in applications, which involve quantitative measurements such as biomechanical analysis of the human movement. Some information concerning the measuring of human movement can be obtained by using uncalibrated cameras, however lens calibration is necessary when metric information is needed. When large field of view is required, wide-angle lenses are useful but they show significant non-linear distortion in the image. The aim of this study was to compare the accuracy of distortion models, which are used in biomechanical analysis of the human movement. A new calibration pattern was designed and the image of the calibration pattern was captured by using a wide-angle lens. The distorted image of the pattern was automatically processed to determine the dot centroids in the image. Lens distortion was modelled accounting for three sources of distortion: radial, decentring and thin prism distortions. These three effects result in seven distortion coefficients ( $k_1, k_2, k_3, p_1, p_2, s_1$ , and  $s_2$ ) in vertical and horizontal directions. Marzan and Karara (1975) [9] suggested 12, 14 and 16 parameters along with the Direct Linear Transformations (DLT) methods that comprise radial and decentring distortion models. In this study thin prism distortion [6] was also included to observe its effect. These non-linear distortion functions were minimised by using three different numerical methods, which were Least Squares, Levenberg-Marquand (LM) and Gauss-Newton (GM). In conclusion, the lens distortion models typically improved the accuracy. Applying the Non-linear Least Squares Optimisation Method (LM) showed the highest accuracy among the distortion models. While the average and the maximum error in the distorted image were 3.394% and 10.994%, respectively, they decreased to 1.591% and 3.524% in the application of DLT 18 method.

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

The importance of estimating pose and anthropometric measurements from an image sequence demands a relationship between the three-dimensional (3D) object coordinate and the image coordinate. Therefore, an accurate camera calibration is required in numerous applications such as: performance measurement for human factors engineering; posture and gait analysis for training athletes

and physically challenged people; animation of the human body and parts; automatic annotation of human activities in video databases; control in video games and virtual reality; and teleoperation of anthropometric robots.

Wide-angle lenses are often utilised in environments where there is little room to locate the cameras. Although, wide-angle lenses are useful where large field of view is required, they show significant non-linear distortion in the image. Generally, methods that correct "barrel" distortion must calculate a distortion centre and correct radial and tangential distortions [1]. Several researchers have introduced a variety of mathematical models of the image distortion and techniques to find the model parameters [2].

\* Tel.: +90 312 297 6893; fax: +90 312 299 2167.

E-mail addresses: [serdar.aritan@hacettepe.edu.tr](mailto:serdar.aritan@hacettepe.edu.tr), [serdar.aritan@ieec.org](mailto:serdar.aritan@ieec.org).

Smith et al. [3] found distortion parameters using a manual technique where dot locations were marked on the screen by hand. Tsai [4] suggested a second-order radial lens distortion correction for the camera calibration using off-the-shelf TV cameras and lenses. Nomura et al. [5] proposed a calibration technique for high-distortion TV camera lenses.

The precision used by researchers is sufficient for the most applications. However, in some cases, the camera lens needs to be accurately modelled. For these situations Weng et al. [6] modified the model proposed by Faugeras-Toscani [7] including three types of lens distortions, which were radial, decentering and thin prism types of distortions.

A method commonly used for the spatial reconstruction of a motion from recorded in two-dimensional (2D) images is the Direct Linear Transformation (DLT) [8]. The conventional 11-parameter DLT (DLT 11) does not include the lens distortion parameters, and therefore, distortion effects cannot be corrected. The lens distortion parameters can be integrated with DLT by increasing the number of additional parameters, which are related to the lens distortion [9-11]. Consequently, Marzan and Karara [9] suggested a DLT method with additional parameters (1-5) to account for lens distortions. These suggested 12, 14 and 16 parameter DLT methods will be called as DLT 12, DLT 14 and DLT 16 in the manuscript. DLT 12 and DLT 14 comprise only radial distortion, however DLT 16 also includes decentering distortion model. In this study thin prism distortion [6] was also included to observe its effects, which could be considered as an 18-parameter DLT (DLT 18). The aim of this study was to compare the efficiency of distortion models, which are used in biomechanical analysis of the human movement.

## 2. Methods

### 2.1. Data collection

In order to make accurate measurements, the calibration pattern must be composed of visible features, normally dots, lines, or crossing lines, placed at known locations with respect to each other in a calibration plane. Dots are usually preferred because their positions are easy to detect. Therefore a calibration pattern was constructed by using circular self-adhesive circular yellow labels (diameter of 20 mm) that were placed 20 mm apart on a black background. The total number of 384 dots was allocated in 24 pieces at horizontal and 16 pieces at vertical directions. The dimension of the effective part of the test pattern was 940 mm in width and 620 mm in height. The overall accuracy of the test pattern is better than 0.4 mm and local accuracy between dots is about 20  $\mu$ m. The calibration pattern was mounted on thin wooden plate and attached to a wall. A vertical line and a horizontal line, which intersect at the centre point of the test pattern, were marked for aligning with the camera. The picture of the calibration pattern can be seen in Fig. 1.

A video camera (Panasonic-MSVC) was placed on a professional tripod (Bogen 3029/3211), so that the captured image would fill the field of view (FOV) and not to be

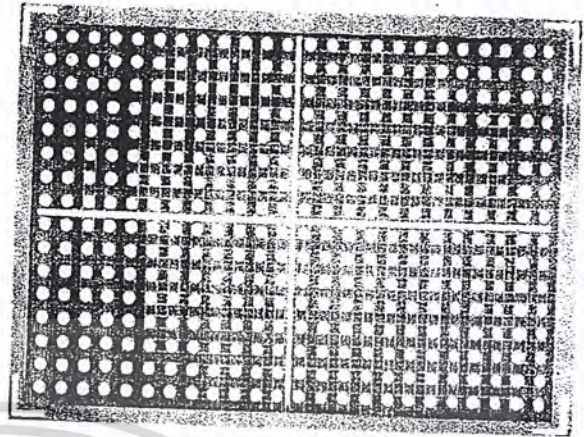


Fig. 1. The picture of the calibration pattern, which is consist of 384 pieces of circular dots.

blurred. The centre of lens was aligned with the centre of the test pattern. The image of the calibration pattern was recorded by a X0.5 wide-angle conversion unit. The distortion effect of the wide-angle unit can be seen in Fig. 2.

Finally, the image was captured and stored in a P4 IBM-Compatible computer by using Iomega video frame-grabber board with a resolution of 768 columns by 576 lines in 24-bit colour scale.

### 2.2. Image processing

Image segmentation and region identification were the most important steps leading to the analysis of processed image data. It was aimed for complete segmentation, which results in a set of disjoint regions corresponding uniquely with dots in the input image. Then all calibration dots were enumerated by using region identification method. All image processing were achieved by using custom written scripts in MATLAB (v.7.1, The Mathworks Inc., Natick, MA, USA). A simplified workflow of the image processing, which was applied to the image of circular dot pattern, is given in Fig. 3.

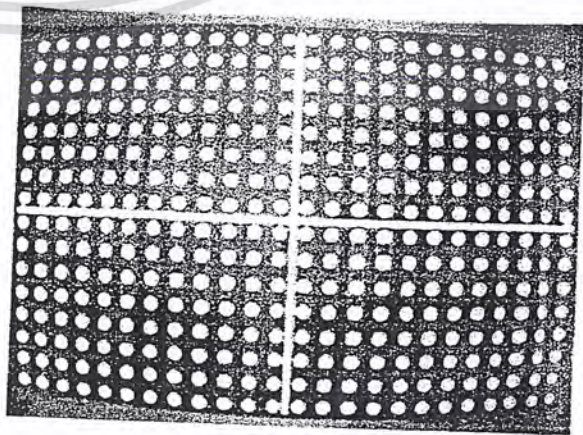


Fig. 2. The distorted picture of the calibration pattern that captured by using 0.5x wide-angle conversion lens.

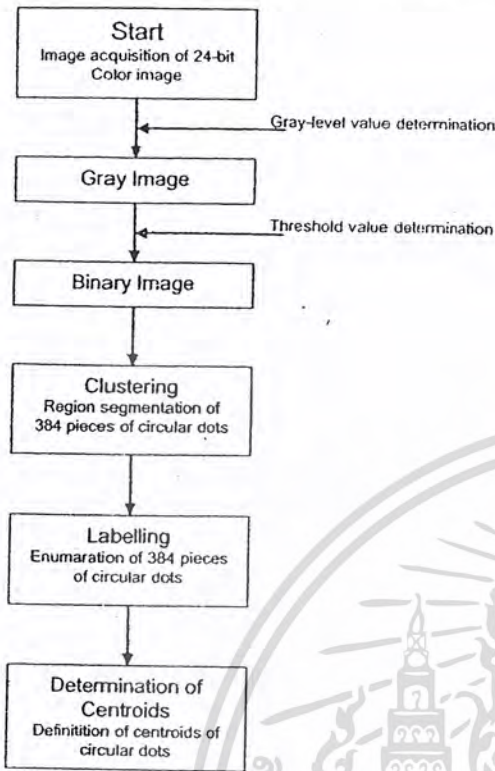


Fig. 3. Simplified work flow of the image processing.

2.2.1. Binary image creation

The calibration image needed to be converted to a binary image for the image segmentation. First, the 24-bit colour image of the calibration pattern was converted into the greyscale format for pixel processing. Since the dots did not touch each other and their grey-levels were clearly distinct from background grey-levels, thresholding method was suitable for segmentation. In this case, Otsu's method [12] was used to compute a global image threshold value. Then the grayscale image was converted into the binary image, by using the threshold value. The output binary image had values of 0 (black) for all pixels in the input image with luminance greater than threshold value and 1 (white) for all other pixels, which defined the background. Finally, isolated single pixels were detected by running a 3 x 3 median filter through the image and removed from the binary image (Fig. 4).

2.2.2. Dot-region identification

Region identification was necessary for dot enumeration. Labelling method was used to identify each dot boundary with a unique integer number. The largest integer label number gave the number of regions in the calibration image. The output of labelling method was a label matrix that was in the same size of the binary image. In the algorithm, the binary image was searched row by row and assigned a non-zero value to label matrix for each zero pixel in the binary image. The label value was chosen according to labels of the pixel's neighbours, where if all the neighbours were background pixels, a new unused label value was assigned in the label matrix. If there is just

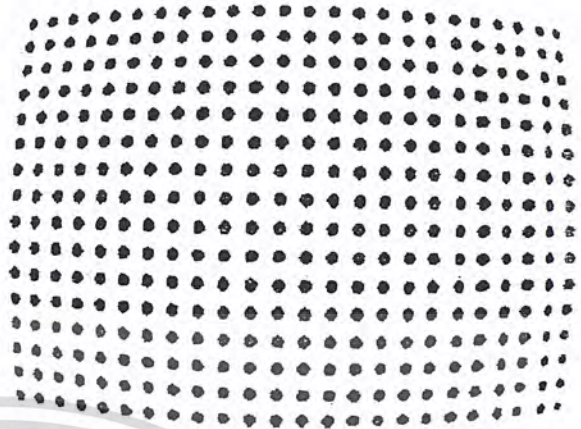


Fig. 4. Binary image of the calibration pattern.

one neighbouring pixel with a non-zero label value in the label matrix, this label value was assigned to the corresponding pixel. The sample of label matrix (nine calibration dots from the top-left corner) is shown in Fig. 5.

2.2.3. Dot-centroid estimation

In order to define centre of dot-regions, the first moment calculation was applied into the label matrix. As a result of the first moment calculation, each dot-region could be defined by its centroid. The centroids of the dot-regions are shown in Fig. 6.

2.3. Estimation of distortion centre

The commonly used lenses are spherical, so the transmitted light is symmetric to the distortion centre. When the distortion centre is computed, it was assumed that the distortion is radial, so that the model is symmetric to the distortion centre. Therefore, it is necessary to estimate the distortion centre for effective determination of the coef-

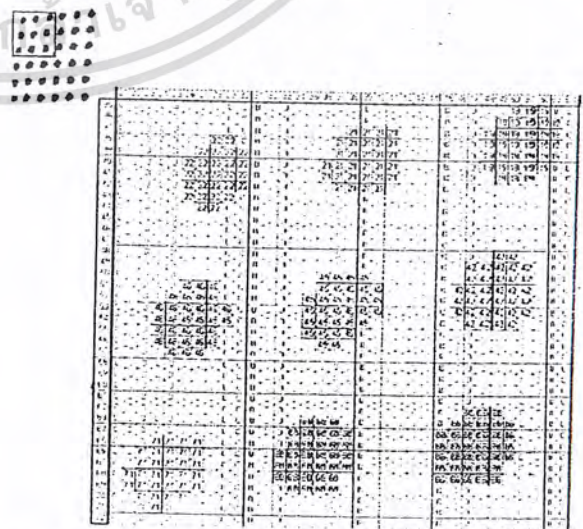


Fig. 5. Sample of label matrix from the top-left corner of the distorted image.

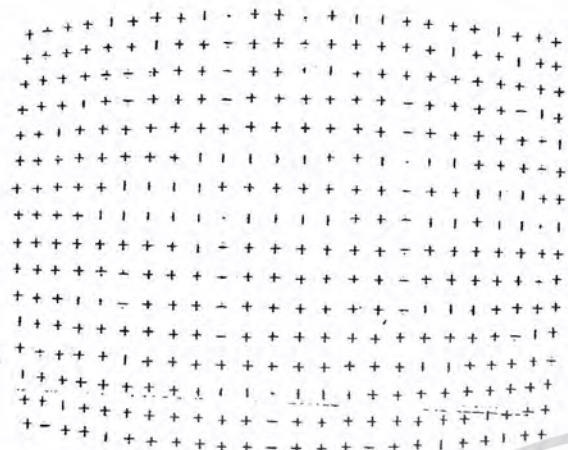


Fig. 6. The image of the dot-centroid positions.

coefficients of the distortion model. Distortion centre is a fixed point for an individual camera and once determined can be used for all the images obtained from that specific camera. Distortion centre  $(c_x, c_y)$  was calculated by using the same method as Asari et al. [13] applied. The distortion centre was a point where the curvature is zero, which was based on calculations of curvature from column-wise and row-wise distorted dots. It can be seen from the Fig. 6 that the curvature of columns decreases from the first to the 12th column and then increases from the 13th to the 24th, but the sign of the curvature changes. The same is also true for the rows.

2.4. Distortion model

In this research, distortion model which were proposed by Weng et al. [6] are used. This model includes three types of lens distortions, which are radial, decentering and thin prism types of distortions. The first one is caused by imperfect lens shape and it only reveals itself by radial positional error, which is called radial distortion  $(k)$ . The second and the third types of distortions are usually caused by improper lens and camera construction and generate both radial and tangential errors, which are called decentering  $(p)$  and thin prism distortions  $(s)$ , respectively.

Marzan and Karara [9] used non-linear distortion models to enhance DLT 11 method. In their suggestion DLT 11 method was expanded by adding one radial distortion coefficient  $(k_1)$  to DLT 12, three radial distortion coefficients  $(k_1, k_2, k_3)$  to DLT 14 and three radial and two decentering distortion coefficients  $(k_1, k_2, k_3, p_1, p_2)$  to DLT 16. In addition to Marzan and Karara's distortion methods, the thin prism distortion [6] coefficients  $(s_1, s_2)$  were also included into the distortion model. Accordingly, the model considers three types of distortions: radial distortion, decentering distortion and thin prism distortion. The total distortion will be the sum of the three distortions. These are;

radial,

$$\Delta_{ur} = u(k_1(u^2 + v^2) + k_2(u^2 + v^2)^2 + k_3(u^2 + v^2)^3) \quad (1)$$

$$\Delta_{vr} = v(k_1(u^2 + v^2) + k_2(u^2 + v^2)^2 + k_3(u^2 + v^2)^3) \quad (2)$$

decentering,

$$\Delta_{ud} = p_1(3u^2 + v^2) + 2p_2uv \quad (3)$$

$$\Delta_{vd} = 2p_1uv + p_2(u^2 + 3v^2) \quad (4)$$

and thin prism distortions.

$$\Delta_{up} = s_1(u^2 + v^2) \quad (5)$$

$$\Delta_{vp} = s_2(u^2 + v^2) \quad (6)$$

The total distortion in horizontal

$$\Delta_{ut} = \Delta_{ur} + \Delta_{ud} + \Delta_{up} \quad (7)$$

and the total distortion in vertical can be summarised.

$$\Delta_{vt} = \Delta_{vr} + \Delta_{vd} + \Delta_{vp} \quad (8)$$

where  $u$  and  $v$  represent the image coordinate of the centroids of the dots in horizontal and vertical directions respectively. These three effects result in seven distortion coefficients, which are  $k_1, k_2, k_3, p_1, p_2, s_1,$  and  $s_2$ .

2.5. Parameter estimation

The task of finding the lens distortion parameters was achieved by using built-in Matlab functions. First, the coordinate values of distorted dot centroids  $(\bar{u}, \bar{v})$  were recalculated with respect to the distortion centre, defined as  $(c_x, c_y)$  above.

$$\bar{u} = u - c_x \quad (9)$$

$$\bar{v} = v - c_y \quad (10)$$

Then, the radial distance of dot centroids from the distortion centre were computed.

$$r^2 = \bar{u}^2 + \bar{v}^2 \quad (11)$$

In order to utilise built-in Matlab functions, experimental data were organised in the matrix form for each DLT methods. These forms are presented below:

For DLT12;

$$\begin{bmatrix} \Delta_u \\ \Delta_v \end{bmatrix} = \begin{bmatrix} \bar{u}r^2 \\ \bar{v}r^2 \end{bmatrix} [k_1] \quad (12)$$

For DLT14;

$$\begin{bmatrix} \Delta_u \\ \Delta_v \end{bmatrix} = \begin{bmatrix} \bar{u}r^2 & \bar{u}(r^2)^2 & \bar{u}(r^2)^3 \\ \bar{v}(r^2) & \bar{v}(r^2)^2 & \bar{v}(r^2)^3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} \quad (13)$$

For DLT16;

$$\begin{bmatrix} \Delta_u \\ \Delta_v \end{bmatrix} = \begin{bmatrix} \bar{u}(r^2) & \bar{u}(r^2)^2 & \bar{u}(r^2)^3 & (3\bar{u}^2 + \bar{v}^2) & 2\bar{u}\bar{v} \\ \bar{v}(r^2) & \bar{v}(r^2)^2 & \bar{v}(r^2)^3 & 2\bar{u}\bar{v} & (\bar{u}^2 + 3\bar{v}^2) \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ p_1 \\ p_2 \end{bmatrix} \quad (14)$$

For DLT18;

$$\begin{bmatrix} \Delta_u \\ \Delta_v \end{bmatrix} = \begin{bmatrix} \bar{u}(r^2) & \bar{u}(r^2)^2 & \bar{u}(r^2)^3 & (3\bar{u}^2 + \bar{v}^2) & 2\bar{u}\bar{v} & r^2 & 0 \\ \bar{v}(r^2) & \bar{v}(r^2)^2 & \bar{v}(r^2)^3 & 2\bar{u}\bar{v} & (\bar{u}^2 + 3\bar{v}^2) & 0 & r^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ p_1 \\ p_2 \\ s_1 \\ s_2 \end{bmatrix} \quad (15)$$

The non-linear distortion equations can be minimised by using two different approaches [14,15]. The first one is Linear Least Squares Estimation (LSE) and the second approach is Non-linear Least Squares Optimisation (NLSO), which is an iterative method. Firstly, LSE was applied to estimate the coefficients (Eq. (12)–(15)) then by using these coefficients the sum of squared residuals (SSR) were calculated for all models. Secondly, in order to reduce SSR, NLSO method was applied to the model that showed the lowest squared residual. The distortion coefficients, which had been first computed by LSE, were used an initial estimate for NLSO. Levenberg–Marquardt (LM) and Gauss–Newton (GN) algorithms were chosen to observe the convergence of distinctive algorithms.

### 3. Results

The parameter estimation of Section 2.5 was applied to distortion model functions (Eqs. (7) and (8)). The estimates of the distortion coefficients of selected methods are shown in Table 1. This table also shows the sum of squared residuals of the selected methods.

NLSO method was only applied to 18-parameter DLT method. Although LM and GN were different algorithms,

they converged with the same coefficients. The distortion coefficients and the sum of squared residual of the method can be seen in Table 2.

In order to avoid digitisation effects, total distortion functions (Eqs. (7) and (8)) were used in an inverse manner to transfer the distorted image into the corrected image. For every pixel in the corrected image was determined by its original subpixel-exact position in the distorted image. Then, new gray value was calculated by utilising bilinear interpolation. Fig. 7 shows the image of Fig. 2 after compensation for lens distortion with the distortion model parameters of Table 2.

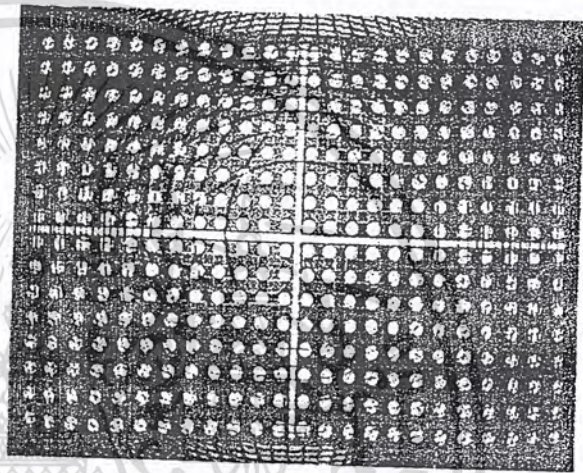


Fig. 7. The result of distortion correction procedure of 7-parameter (DLT 18) distortion model.

Table 1  
Distortion coefficients of selected methods, which were calculated by LSE, and their sum of squared residuals.

Model parameters	Used by DLT 12 Marzan and Karara [9]	Used by DLT 14 Marzan and Karara [9]	Used by DLT 16 Marzan and Karara [9]	Used by DLT 18 Weng et al. (1992)
$k_1$	$5.63853 \times 10^{-06}$	$2.61407 \times 10^{-06}$	$2.82759 \times 10^{-06}$	$2.85931 \times 10^{-06}$
$k_2$	–	$2.62541 \times 10^{-10}$	$2.37705 \times 10^{-10}$	$2.34014 \times 10^{-10}$
$k_3$	–	$-5.34995 \times 10^{-15}$	$-4.68691 \times 10^{-15}$	$-4.58837 \times 10^{-15}$
$p_1$	–	–	$-3.75890 \times 10^{-06}$	$-1.15545 \times 10^{-05}$
$p_2$	–	–	$-4.92624 \times 10^{-05}$	$-7.51179 \times 10^{-05}$
$s_1$	–	–	–	$2.13527 \times 10^{-05}$
$s_2$	–	–	–	$5.57397 \times 10^{-05}$
$\Sigma$ Residual <sup>2</sup> (pixels)	$1.21671 \times 10^4$	$1.19990 \times 10^4$	$1.13808 \times 10^4$	$1.13236 \times 10^4$

Table 2  
Distortion coefficients of 7-parameter (DLT 18) distortion model, which were calculated by NLSO, and their sum of squared residual.

Model parameters	Least Squares Estimation		
	Least squares method (used for initial values for NLSO)	Non-linear Least Squares Optimisation	
		Levenberg–Marquardt method	Gauss–Newton method
$k_1$	$2.85931 \times 10^{-06}$	$2.79717 \times 10^{-06}$	$2.79717 \times 10^{-06}$
$k_2$	$2.34014 \times 10^{-10}$	$2.39380 \times 10^{-10}$	$2.39380 \times 10^{-10}$
$k_3$	$-4.58837 \times 10^{-15}$	$-4.69794 \times 10^{-15}$	$-4.69794 \times 10^{-15}$
$p_1$	$-1.15545 \times 10^{-05}$	$-1.15586 \times 10^{-05}$	$-1.15586 \times 10^{-05}$
$p_2$	$-7.51179 \times 10^{-05}$	$-7.50935 \times 10^{-05}$	$-7.50936 \times 10^{-05}$
$s_1$	$2.13527 \times 10^{-05}$	$2.13395 \times 10^{-05}$	$2.13395 \times 10^{-05}$
$s_2$	$5.57397 \times 10^{-05}$	$5.57148 \times 10^{-05}$	$5.57148 \times 10^{-05}$
$\Sigma$ Residual <sup>2</sup> (pixels)	$1.13236 \times 10^4$	$1.13232 \times 10^4$	$1.13232 \times 10^4$

Please cite this article in press as: S. Aritan, Efficiency of non-linear lens distortion models in biomechanical analysis of human movement, Measurement (2010), doi:10.1016/j.measurement.2010.01.018



Fig. 8. Contour graphics of error distribution of distorted image (topmost) and performance of DLT 12 (middle-left), DLT 14 (middle-right), DLT 16 (bottom-left) and DLT 18 (bottom-right) distortion models.

The effects of the lens distortion error on an image can be viewed by a contour graphics due to the fact that it represents better elucidation as a quantitative measuring technique. That is why, the contour graphics of lens distortion and outcomes of the selected distortion models are presented in Fig. 8. Error distribution of the lens distortion and the models were calculated by using the

space between the distorted and the calibration locations. These error values in pixels were than normalised as percentage errors. In Fig. 8, the topmost and the largest of contour graphs shows the error distribution of wide-angle lens before applying any distortion models. There are four separate contour graphs below, which show performance of DLT 12 (middle-left), DLT 14 (mid-

ขอสงวนลิขสิทธิ์ในเอกสารนี้ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่อนุญาตให้เผยแพร่ไปโดยไม่ขออนุญาต  
 Please cite this article in press as: S. Aritan, Efficiency of non-linear lens distortion models in biomechanical analysis of human movement, Measurement (2010), doi:10.1016/j.measurement.2010.01.018  
 ขอสงวนลิขสิทธิ์ในเอกสารนี้ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่อนุญาตให้เผยแพร่ไปโดยไม่ขออนุญาต

**Table 3**  
Average and maximum errors of selected distortion models.

		Average error (%)	Maximum error (%)
Distorted image		3.394	10.994
Corrected image parameter used by DLT 12	$(k_1)$	1.646	3.909
Corrected image parameters used by DLT 14	$(k_1, k_2, k_3)$	1.621	4.078
Corrected image parameters used by DLT 16	$(k_1, k_2, k_3, p_1, p_2)$	1.592	3.609
Corrected image parameters used by DLT 18	$(k_1, k_2, k_3, p_1, p_2, s_1, s_2)$	1.592	3.526
Corrected image parameters used by DLT 18 with NLSO	$(k_1, k_2, k_3, p_1, p_2, s_1, s_2)$	1.591	3.524

dle-right), DLT 16 (bottom-left) and DLT 18 (bottom-right) distortion models respectively. Due to the similarity between DLT 18 with LSE and DLT 18 with NLSO, the contour graph of DLT 18 with NLSO is not shown in the Fig. 8.

It is clear from the topmost contour graph in Fig. 8 that the wide-angle lens caused small errors near the centre of the image and became progressively greater further away from the centre. The contour graphs of the distortion models in Fig. 8 illustrate that the correction procedure decreases the maximum error from 11% to 3%. The average and the maximum values of the contour graphs are summarised in Table 3.

As Table 3 shows, utilising DLT 18 with the NLSO method reduces the average lens distortion error by half. Further this method was able to reduce the maximum error by one third.

#### 4. Discussion and conclusion

Although the human movement can easily be obtained by using uncalibrated cameras, lens calibration plays a critical role when metric information is required. Therefore the aim of this study was to compare the efficiency of distortion models deliberately under the condition where the distortion was significant. In order to observe the lens distortion a new calibration pattern was designed. Regarding the test pattern, the dots on the pattern have to be large enough and adequately distributed to allow successful extraction. Certainly, substantial variation is also acceptable in this pattern as long as there is satisfactory information (known dot centroids) that covers the FOV of the camera to run the optimisation methods. The image of the distorted calibration pattern was automatically processed by custom written MATLAB scripts to find the dot centroids, which prevented combination of the lens distortion error with the manual digitising error.

Experimental results were able to demonstrate the performance of the non-linear distortion models that have been widely used in biomechanics. The result has also showed that radial lens distortion model (DLT 12 and DLT 14) results in significant improvement on correcting the lens distortion. Corrections of decentering (DLT 16) distortion also lead to considerable improvements over just correcting radial distortion. However the lowest maximum error value was achieved by addition of thin prism distortion (DLT 18) into the model. Finally, the NLSO method was applied to reduce the residual of DLT 18. The coefficients that obtained from the LSE were used as an initial estimate for the NLSO. Utilising this method slightly decreased the average and the maximum errors.

Although numerous commercial video-based movement analysis software use DLT 12 DLT 14 and DLT 16 algorithms for non-linear lens distortion, the thin prism distortion parameter does not integrate to their calibration procedures. This study shows that integration of thin prism distortion into the DLT possibly will increase the accuracy of the non-linear lens distortion corrections. To demonstrate the effect of this reconstruction of accuracy on kinematic data is presented in a video as a Supplementary material.

#### Acknowledgements

This work is funded by Hacettepe University Research Grant # 01.02.407.001. The helpful suggestions of Dr. Fatih Yaşar are gratefully acknowledged.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.measurement.2010.01.018.

#### References

- [1] P. Allard, J. Blanche, R. Aissaoui, Bases of three-dimensional reconstruction, in: P. Allard, I.A.F. Stokes, J.P. Blanche (Eds.), *Three-Dimensional Analysis of Human Movement, Human Kinetics, Champaign*, 1995, pp. 19–40.
- [2] H.J. Woltring, R. Huiskes, Stereophotogrammetry, in: N. Berme, A. Cappozzo (Eds.), *Biomechanics of Human Movement: Applications in Rehabilitation, Sports and Ergonomics*, Bertec Corporation, Worthington, OH, 1990, pp. 108–127.
- [3] W.E. Smith, N. Vakil, S.A. Maislin, Correction of distortion in endoscope images, *IEEE Transactions on Medical Imaging* 11 (1992) 117–122.
- [4] R.Y. Tsai, An effective and accurate camera calibration technique for 3D machine vision, in: *Proceedings of the IEEE Computer Vision Pattern Recognition*, Miami, FL, 1986, pp. 364–374.
- [5] Y. Nomura, M. Sagara, H. Naruse, A. Ide, Simple calibration algorithm for high-distortion-lens camera, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 14 (1992) 1095–1099.
- [6] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 14 (1992) 965–980.
- [7] O.D. Faugeras, G. Toscani, The calibration problem for stereo, in: *Proceedings of the IEEE Computer Vision and Pattern Recognition*, Miami FL, 1986, pp. 15–20.
- [8] Y.I. Abdel-Aziz, H.M. Karara, Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry, in: *Proceedings of the ASP Symposium on Close-Range Photogrammetry*, Urbana, IL, 1971, pp. 1–18.
- [9] G.T. Marzan, H.M. Karara, A computer program for direct linear transformation solution of the colinearity condition, and some applications of it, in: *Proceedings of the ASP Symposium on Close-Range Photogram*, Falls Church, VA, 1975, pp. 420–476.

- [10] N.R. Miller, R. Shapiro, T.M. Laughin, A Technique for obtaining spatial kinematics parameters of segments of biomechanical systems from cinematographic data, *Journal of Biomechanics* 13 (1980) 535-547.
- [11] H. Hatze, High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach, *Journal of Biomechanics* 21 (1988) 533-538.
- [12] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transaction on Systems, Man, and Cybernetics* 9 (1979) 62-66.
- ✓ [13] K.V. Asari, S. Kumar, D. Radhakrishnan, A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation, *IEEE Transactions on Medical Imaging* 18 (1999) 345-354.
- [14] W.H. Press, S.T. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, second ed., Cambridge University Press, Cambridge, 1992, pp. 656-706.
- [15] C.F. Van Loan, *Introduction to Scientific Computing: A Matrix-Vector Approach Using MATLAB*, second ed., Prentice-Hall, New Jersey, 2000, pp. 275-326.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

Please cite this article in press as: S. Aritan, Efficiency of non-linear lens distortion models in biomechanical analysis of human movement, *Measurement* (2010), doi:10.1016/j.measurement.2010.01.018