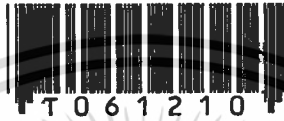


สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวีดีโอตัวเดียวสำหรับ  
ระบบแสดงภาพสเตอริโอแบบเปลี่ยนมุมมองได้

A LEG POSITION ESTIMATION ALGORITHM USING A SINGLE CAMERA  
FOR AN ADAPTABLE-VIEWPOINT STEREOSCOPIC DISPLAY SYSTEM



ณรงค์ศักดิ์ พุดเฟือก

NARONGSAK PUTPUEK

เลขหมู่.....  
เลขทะเบียน... 61210  
วัน,เดือน,ปี... 17 ก.ค. 2549

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

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาเทคโนโลยีสารสนเทศ  
บัณฑิตวิทยาลัย  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
พ.ศ. 2548  
ISBN 974-15-2044-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**A LEG POSITION ESTIMATION ALGORITHM USING A SINGLE CAMERA  
FOR AN ADAPTABLE-VIEWPOINT STEREOSCOPIC DISPLAY SYSTEM**



**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY  
SCHOOL OF GRADUATE STUDIES  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

**2005**

**ISBN 974-15-2044-1**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2005**

**SCHOOL OF GRADUATE STUDIES**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**บัณฑิตวิทยาลัย**  
**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**  
**ใบรับรองวิทยานิพนธ์**

---

**หัวข้อวิทยานิพนธ์** อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวีดีโอตัวเดียวสำหรับระบบ  
แสดงภาพสเตอริโอแบบเปลี่ยนมุมมองได้  
A LEG POSITION ESTIMATION ALGORITHM USING A SINGLE  
CAMERA FOR AN ADAPTABLE-VIEWPOINT STEREOSCOPIC  
DISPLAY SYSTEM

**ชื่อนักศึกษา** นายณรงค์ศักดิ์ พุคเพือก  
**รหัสประจำตัว** 43067144  
**ปริญญา** วิทยาศาสตรมหาบัณฑิต  
**สาขาวิชา** เทคโนโลยีสารสนเทศ  
**อาจารย์ผู้ควบคุมวิทยานิพนธ์** รศ.ดร.นพพร โชติกกำจร

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.ดร.นพพร	โชติกกำจร	
รศ.ดร.บุญวัฒน์	อัคร	
รศ.ดร.อาริต	ธรรมโน	
ดร.อรัญญา	วลัยรัชต์	
ดร.ธนารัตน์	ชวลิตาพงศ์	

วัน/เดือน/ปี ที่สอบ 10 ตุลาคม 2548 เวลา 8.00 น. เป็นต้นไป

สถานที่สอบ ณ ห้อง M04 (ชั้นลอย) คณะเทคโนโลยีสารสนเทศ



วันที่.....?.....เดือน.....๕๔๗๘.....พ.ศ.....๒๕๔๘.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวีดีโอตัวเดียว สำหรับระบบแสดงภาพสเตอริโอแบบเปลี่ยนมุมมองได้
นักศึกษา	นายณรงค์ศักดิ์ พุดเผือก
รหัสประจำตัว	43067144
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
พ.ศ.	2548
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.ดร.นพพร โชติคกำจร

### บทคัดย่อ

ระบบแสดงภาพสเตอริโอแสดงภาพที่มองผ่านตาซ้าย และตาขวาของผู้ใช้จากจุดมองที่กำหนดได้ ทำให้ผู้ใช้รู้สึกถึงความลึกของภาพได้เหมือนกับมองภาพวัตถุจริง แต่อย่างไรก็ตามระบบแสดงภาพสเตอริโอโดยปกติ ไม่สามารถแสดงภาพที่มีความสอดคล้องกับการเปลี่ยนตำแหน่งของผู้ใช้ได้ ระบบแสดงภาพสเตอริโอเชิงโต้ตอบสามารถสร้างภาพที่สอดคล้องกับตำแหน่งของผู้ใช้ ซึ่งทำให้ผู้ใช้สามารถมองเห็นภาพที่ถูกต้องตามตำแหน่งที่เปลี่ยนไป ดังนั้นข้อมูลตำแหน่งของผู้ใช้จะมีความจำเป็นสำหรับระบบการแสดงผลนี้ ในงานวิจัยนี้จึงนำเสนอวิธีการตรวจวัดตำแหน่งของผู้ใช้ด้วยอัลกอริทึมสำหรับตรวจวัดตำแหน่งขาของผู้ใช้ ร่วมกับกล้องวีดีโอ 1 ตัว โดยวิธีการนี้พัฒนาขึ้นสำหรับการประยุกต์ใช้งานกับระบบแสดงภาพสเตอริโอเชิงโต้ตอบในระดับราคาต่ำ เช่น เกมส์คอมพิวเตอร์ เป็นต้น การคำนวณตำแหน่งของผู้ใช้ที่ยืนอยู่กับพื้นราบสามารถคำนวณได้โดยการแยกโครงสร้างขาจากภาพวีดีโอแต่ละภาพ แล้วนำโครงสร้างขาที่ได้ไปคำนวณหาตำแหน่งสันเท้า ซึ่งแทนตำแหน่งของผู้ใช้ด้วยวิธีการแปลงแบบโฮโมกราฟฟี วิธีการนี้ผู้ใช้ไม่ต้องสวมใส่อุปกรณ์ตรวจวัดตำแหน่งหรือการทำเครื่องหมายบนร่างกาย จากการทดลองวิธีการตรวจวัดตำแหน่งขาสามารถคำนวณตำแหน่งของผู้ใช้ที่ยืนอยู่กับพื้นราบมีค่าผิดพลาดเฉลี่ย 10 เซนติเมตร

<b>Thesis</b>	A Leg Position Estimation Algorithm Using A Single Camera for An Adaptable-Viewpoint Stereoscopic Display System
<b>Student</b>	Mr. Narongsak Phutphuek
<b>Student ID</b>	43067144
<b>Degree</b>	Master of Science
<b>Programme</b>	Information Technology
<b>Year</b>	2005
<b>Thesis Advisor</b>	Assoc. Prof. Dr. Nopporn Chotikakamthorn

### ABSTARCT

A stereoscopic display system presents the left and right eyes of the viewer with images from different perspective viewpoints, to provide depth perception. However, a standard stereoscopic display system cannot deliver an image with correct perspective when user changes his/her viewing position. With an interactive stereoscopic display system, a pair of displayed images is synthesized based on the current viewer's position. Thus, a viewer always sees the image pair with correct perspective. For such system, however, information about current viewer's position is required. In this paper, we describe an interactive stereoscopic system that employs a vision-based leg tracking algorithm for the estimation of the viewer position. The method targets at a low-cost application such as computer game etc. It makes use of a cheap single video camera and a fast algorithm for viewer position estimation. A viewer's current planar position is calculated from a leg skeleton extracted from each video snapshot. The skeleton is then used to estimate a viewer leg end-point (heel) position using a homography transformation. The method dose not requires a viewer to wear any tracking sensor or marker. From the experiment, it was found that the proposed method can estimate a viewer planar position with the averaged estimation error of 10 cm.

# กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สามารถสำเร็จลุล่วงไปด้วยดี ก็เพราะได้รับการสนับสนุน และความช่วยเหลือ จากบุคคลเหล่านี้

ขอกราบขอบพระคุณ บิดา มารดา และ ญาติๆ ของนักวิจัยที่ให้การสนับสนุนและช่วยเหลือในด้านต่าง ๆ อีกทั้งเป็นกำลังใจให้กับผู้วิจัยเสมอมา

ขอขอบพระคุณ รศ.ดร.นพพร โชติกกำทร ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ซึ่งให้แนวคิดและคำปรึกษาต่างๆ เป็นอย่างดี อีกทั้งยังให้การสนับสนุนเอกสารและอุปกรณ์ที่เกี่ยวข้องกับงานวิจัยเป็นอย่างดี

ขอขอบคุณ พี่ๆ เพื่อนๆ น้องๆ ภาควิชาคอมพิวเตอร์ พี่ๆ คณะวิทยาศาสตร์ และพี่ๆ เพื่อนๆ มหาวิทยาลัยราชภัฏราชนครินทร์ ทุกคนที่ให้คำปรึกษาและให้กำลังใจเป็นอย่างดี

ขอขอบคุณน้องชาว MVLAB ทุกคน ไม่ว่าจะเป็นน้องกบ ที่ให้คำปรึกษาและที่พัก น้องเปิง น้องเต๋อ น้องนน น้องนะ รวมถึงน้องๆ ทุกคน ที่ให้คำปรึกษาด้วยดี

อนึ่ง งานวิจัยที่นำเสนอในวิทยานิพนธ์ฉบับนี้นั้น ส่วนหนึ่งได้งานปฏิบัติภายในห้องปฏิบัติการ Multimedia and Virtual Research (MVLAB) ของสำนักวิจัยการสื่อสารและเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

คุณค่าและประโยชน์ที่พึงได้รับจากวิทยานิพนธ์ฉบับนี้ ขอมอบแด่ผู้มีพระคุณทุกท่าน

ณรงค์ศักดิ์ พุดเผือก

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	1
1.3 สมมุติฐานของการศึกษา.....	1
1.4 ขอบเขตของการดำเนินงานวิจัย.....	2
1.5 ขั้นตอนการศึกษา.....	2
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	2
1.7 โครงสร้างวิทยานิพนธ์.....	2
บทที่ 2 งานวิจัยที่เกี่ยวข้อง.....	3
บทที่ 3 ทฤษฎีที่เกี่ยวข้องในงานวิจัย.....	9
3.1 คอมพิวเตอร์กราฟิก.....	9
3.1.1 มิติและระบบพิกัดในคอมพิวเตอร์กราฟิก.....	9
3.1.2 การแปลงเชิงเรขาคณิต.....	11
3.2 รูปภาพดิจิทัล.....	14
3.3 การประมวลผลภาพดิจิทัล.....	15
3.4 การแสดงภาพสเตอริโอ.....	17
3.4.1 การมองเห็นภาพ 3 มิติ.....	17
3.4.2 การแสดงภาพสเตอริโอด้วยคอมพิวเตอร์.....	18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
3.4.3 ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ.....	19
3.4.4 หลักการสร้างภาพสเตอริโอแบบเพอร์สเพกทีฟ [18].....	20
3.5 Homography [8].....	23
3.6 การตรวจวัดการเคลื่อนไหวของมนุษย์.....	26
3.6.1 ชนิดของการตรวจวัดตำแหน่ง [3].....	26
3.6.2 เปรียบเทียบข้อดีและข้อเสียของวิธีการตรวจวัดตำแหน่ง.....	28
<b>บทที่ 4 อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ.....</b>	<b>30</b>
4.1 แนวคิดเบื้องต้น.....	30
4.2 การแบ่งส่วนของร่างกายมนุษย์.....	30
4.3 ขนาดและระยะพิกัดของภาพวิดีโอ.....	31
4.4 อัลกอริทึมในการตรวจวัดตำแหน่งขา.....	32
4.4.1 กระบวนการลบภาพพื้นหลัง.....	32
4.4.2 กระบวนการแยกภาพขา.....	33
4.4.3 กระบวนการแยกโครงสร้างขา.....	34
4.4.4 กระบวนการคำนวณหาตำแหน่งสันเท้า.....	35
4.4.5 กระบวนการคำนวณตำแหน่งของผู้ใช้.....	36
4.5 อัลกอริทึมสำหรับแก้ไขข้อผิดพลาด.....	36
<b>บทที่ 5 ผลการทดลอง.....</b>	<b>37</b>
5.1 อุปกรณ์ที่ใช้ในการทดลองและระบบที่ใช้ในการทดลอง.....	37
5.1.1 อุปกรณ์ที่ใช้ในการทดลอง.....	37
5.1.2 ระบบที่ใช้ในการทดลอง.....	37
5.1.3 การกำหนดค่าเริ่มต้นของ Homography matrix.....	38
5.1.4 การกำหนดตำแหน่งอ้างอิงบนพื้น.....	38
5.2 ผู้ร่วมในการทดลอง.....	40

## สารบัญ (ต่อ)

	หน้า
5.3 การทดลอง .....	40
5.3.1 การวัดตำแหน่งโดยขึ้นอยู่กับที่.....	40
5.3.2 การวัดตำแหน่งโดยการเดินเป็นเส้นตรง .....	43
5.3.3 การวัดตำแหน่งโดยการเดินวนไปมา.....	45
5.3.3 การวัดคุณภาพจากการใช้งาน.....	49
บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ .....	52
6.1 สรุปผลการวิจัย.....	52
6.2 ข้อเสนอแนะสำหรับการพัฒนาในอนาคต.....	54
เอกสารอ้างอิง.....	55
ภาคผนวก.....	57
ภาคผนวก ก. ผู้ร่วมในการทดลอง.....	58
ภาคผนวก ข. บทความและผลงานวิจัยที่ได้รับการตีพิมพ์.....	59
ประวัติผู้เขียน .....	73

# สารบัญตาราง

ตารางที่	หน้า
3.1 ผลลัพธ์ที่ได้จากการคำนวณตำแหน่งในการสร้างภาพตาซ้ายและตาขวา.....	23
5.1 ตำแหน่งอ้างอิงที่ใช้สำหรับตรวจสอบความถูกต้องจากการทดลอง.....	39
5.1 (ต่อ).....	40
5.2 ลักษณะของการขึ้นที่ใช้ในการทดลอง .....	41
5.3 ค่าความผิดพลาดของค่าตำแหน่งที่คำนวณได้เปรียบเทียบกับตำแหน่งอ้างอิง .....	42
5.4 ทิศทางการเดินที่ใช้ในการทดลอง.....	44
5.5 ค่าผิดพลาดของการคำนวณตำแหน่งจากทิศทางการเดินที่ใช้ในการทดลอง .....	44
5.6 ผลการทดลองของรูปแบบการเดินที่ 1 โดยคิดค่าความผิดพลาดของการคำนวณค่าตำแหน่งที่น้อยที่สุด .....	47
5.7 ค่าผิดพลาดของการคำนวณตำแหน่งจากรูปแบบการเดินที่ใช้ในการทดลอง ที่ค่า Delay = 0... 48	
5.8 ค่าผิดพลาดของการคำนวณตำแหน่งจากรูปแบบการเดินที่ใช้ในการทดลอง เมื่อมีค่าความผิดพลาดน้อยที่สุด.....	49
5.9 เปรียบเทียบค่าความผิดพลาดที่น้อยที่สุดของการเดินแต่ละรูปแบบ .....	49
5.10 การวิเคราะห์ผลระยะห่างของวัตถุจำลองต่อคุณภาพของภาพที่ผู้ทดลองรับรู้.....	50

# สารบัญรูป

รูปที่	หน้า
2.1 Personal space station (PSS) .....	3
2.2 วิธีการแสดงภาพสเตอริโอเชิงโต้ตอบที่สามารถใช้ได้กับผู้ใช้ตั้งแต่ 3 คนขึ้นไป ของ Y. Kitamura, T. Konishi, S. Yamamoto and F. Kishino .....	4
2.3 วิธีการตรวจวัดท่าทางของขา ของ Chin-Chun Chang and Wen-Hsiang Tsai.....	5
2.4 ระบบตรวจวัดตำแหน่งศีรษะด้วยกล้องวิดีโอ ของ M. Ribo, A. Pinz, and A. Fuhrmann.....	6
2.5 วิธีการตรวจวัดตำแหน่งของตา ของ T. Ohno and N. Mukawa .....	7
2.6 วิธีการตรวจวัดตำแหน่งของตา ของ R. Ruddaraju, A. Haro, K. Nagel, Quan T. Tran, Irfan A. Essa, G. Abowd and E. D. Mynatt .....	7
2.7 วิธีการตรวจวัดตำแหน่งของศีรษะสำหรับ Fish Tank VR ของ J. Rekimoto .....	8
3.1 การประยุกต์ใช้งานคอมพิวเตอร์กราฟิก ด้านระบบการแสดงผลภาพเสมือนจริง [19].....	9
3.2 จุด เส้น รูป 2 มิติ และรูป 3 มิติ .....	10
3.3 ระบบพิกัดคาร์ทีเซียนแบบ 2 มิติ .....	10
3.4 ระบบพิกัดคาร์ทีเซียนแบบ 3 มิติ.....	11
3.5 กระบวนการแปลงภาพธรรมชาติ(Continuous image)เป็นภาพดิจิทัล(Digital image).....	14
3.6 ระบบพิกัดของภาพดิจิทัล .....	15
3.7 ส่วนประกอบของระบบการประมวลผลภาพดิจิทัลเบื้องต้น .....	16
3.8 การลบพื้นหลังจากภาพออก (ก)ภาพที่ต้องการลบพื้นหลัง, (ข)ภาพพื้นหลัง และ (ค)ภาพผลลัพธ์ .....	17
3.9 หลักการมองเห็นภาพ 3 มิติของมนุษย์ [20] .....	17
3.10 อุปกรณ์สำหรับระบบการแสดงผลภาพสเตอริโอด้วยคอมพิวเตอร์ [2] .....	18
3.11 ภาพ 3 มิติที่สร้างด้วยคอมพิวเตอร์ (ก)ภาพของตาซ้าย, (ข)ภาพของตาขวา และ(ค)ภาพเมื่อมองด้วยตาเปล่า [21] .....	19
3.12 ภาพที่มองผ่านแว่นตา (Shutter glasses) (ก)ภาพที่มองเห็นผ่านตาซ้าย และ(ข)ภาพที่มองเห็นผ่านตาขวา [21].....	19
3.13 ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ [22].....	20
3.14 หลักการสร้างภาพสเตอริโอแบบพอร์สเพคทีฟ .....	20
3.15 หลักการฉายภาพบนระนาบ.....	21
3.16 ตัวอย่างการสร้างภาพสเตอริโอ.....	22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป(ต่อ)

รูปที่	หน้า
3.17 แบบจำลองการสร้างภาพด้วยกล้อง .....	23
3.18 ตัวอย่างของวิธีการสำหรับการตรวจวัดการเคลื่อนไหวของมนุษย์ [3].....	26
3.19 เครื่องมือตรวจวัดตำแหน่งชื่อ FASTRAK ของบริษัท Polhemus [24].....	27
3.20 วิธีการตรวจวัดตำแหน่งแบบ Vision based with markers [25].....	28
3.21 วิธีการตรวจวัดตำแหน่งแบบ Vision based without markers .....	28
4.1 การแบ่งร่างกายมนุษย์ออกเป็น 5 ส่วน [23].....	31
4.2 ระยะพิกัดของภาพที่ใช้ในงานวิจัย.....	31
4.3 โครงสร้างและขั้นตอนของอัลกอริทึมในการตรวจวัดตำแหน่งขา.....	32
4.4 กระบวนการ Background subtraction .....	33
4.5 การคำนวณหา Center of Mass.....	33
5.1 ระบบที่ใช้ในการทดลอง (ก)ภาพด้านบน และ(ข)ภาพด้านข้าง .....	37
5.2 จุดอ้างอิงที่สอดคล้องกัน 4 จุดระหว่าง World plane กับ Image plane.....	38
5.3 การกำหนดตำแหน่งอ้างอิง .....	39
5.4 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งเปรียบเทียบกับตำแหน่งอ้างอิง .....	42
5.5 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจากรูปแบบการอื่นที่ใช้ในการทดลอง .....	43
5.6 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจากทิศทางการเดินที่ใช้ในการทดลอง.....	45
5.7 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจาก Smoothing algorithm .....	45
5.8 รูปแบบการเดินที่ใช้ในการทดลอง (ก) รูปแบบที่ 1 (ข) รูปแบบที่ 2 (ค) รูปแบบที่ 3 (ง) รูปแบบที่ 4 (จ) รูปแบบที่ 5 และ (ฉ) รูปแบบที่ 6.....	46
5.9 ผลการทดลองของการเดินในรูปแบบที่ 1 เมื่อคำนวณค่าตำแหน่งใหม่ด้วย Smoothing algorithm (ก) $N = 0$ (ข) $N = 10$ และ (ค) $N = 20$ .....	47
5.10 ภาพของวัตถุจำลองที่สร้างขึ้นจากข้อมูลตำแหน่งจากอุปกรณ์ตรวจวัดตำแหน่งและอัลกอริทึมในการตรวจวัดตำแหน่งขา .....	50
6.1 (ก) การเดินในทิศทางด้านซ้าย ไปด้านขวา (ข) การเดินในทิศทางด้านหลังมาด้านหน้า .....	53
ก.1 ผู้ร่วมในการทดลองทั้งหมด 10 คน.....	58

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

การมองเห็นของมนุษย์โดยปกติจะมีการมองเห็นภาพเป็นแบบ 3 มิติ ซึ่งเกิดจากการที่มนุษย์มองภาพด้วยตา 2 ข้าง ซึ่งภาพที่มองเห็นผ่านตาทั้ง 2 ข้างจะไม่เหมือนกัน เมื่อสมองนำภาพที่ดวงตาทั้ง 2 ข้างเห็นไปตีความก็จะได้เป็นภาพ 3 มิติ คือเป็นภาพที่มีความลึกของวัตถุในระยะต่างๆ ทำให้มนุษย์สามารถระบุระยะของวัตถุในภาพได้ หลักการดังกล่าวได้ถูกนำมาใช้ในการแสดงภาพบนจอภาพคอมพิวเตอร์ ทำให้สามารถมองเห็นภาพบนจอคอมพิวเตอร์เป็นภาพ 3 มิติได้ การแสดงภาพบนจอคอมพิวเตอร์ดังกล่าวเรียกว่า ระบบการแสดงผลภาพสเตอริโอ (Stereoscopic viewing system)

ระบบการแสดงผลภาพสเตอริโอที่มีการพัฒนาให้สามารถแสดงผลสอดคล้องกับตำแหน่งการเคลื่อนไหวของมนุษย์ได้ เรียกว่า ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ (Interactive stereoscopic viewing system) ระบบดังกล่าวจำเป็นต้องทราบตำแหน่งปัจจุบันของผู้ดูเพื่อการปรับเปลี่ยนมุมมองของภาพที่แสดงผลสอดคล้องกับตำแหน่งการมอง เครื่องมือที่สามารถนำมาใช้ในการตรวจวัดตำแหน่งของมนุษย์ ได้แก่ เครื่องตรวจวัดตำแหน่งด้วยสนามแม่เหล็กไฟฟ้า แต่เนื่องจากเครื่องมือดังกล่าวในปัจจุบันยังมีราคาแพง ไม่เหมาะสมที่จะนำมาใช้กับแอปพลิเคชันราคาถูกเช่น เกมสโตร์คอมพิวเตอร์ เป็นต้น ในงานวิจัยนี้ จึงนำเสนอวิธีการตรวจวัดตำแหน่งของมนุษย์ ด้วยวิธีการที่ง่าย รวดเร็วและมีต้นทุนของอุปกรณ์ที่นำมาใช้มีราคาถูก ซึ่งมีความเหมาะสมกับการประยุกต์ใช้ในงานที่มีข้อจำกัดในเรื่องงบประมาณ

### 1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

วัตถุประสงค์ของการศึกษาวิทยานิพนธ์ฉบับนี้ เพื่อศึกษาและพัฒนาวิธีการตรวจวัดตำแหน่งของผู้ใช้โดยใช้กล้องวิดีโอตัวเพียงเดียวสำหรับระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ โดยอาศัยอุปกรณ์ที่มีราคาถูก และสามารถทำงานได้กับเครื่องคอมพิวเตอร์ส่วนบุคคลโดยทั่วไป

### 1.3 สมมุติฐานของการศึกษา

1.3.1 ระยะห่างของผู้ใช้จากกล้องมีผลต่อความถูกต้องในการคำนวณหาตำแหน่ง

1.3.2 ที่ระดับของความผิดพลาดในการคำนวณหาตำแหน่งที่เท่ากัน ในระยะห่างที่ต่างกันมีผลต่อคุณภาพของการทำงานที่ผู้ใช้รับรู้

## 1.4 ขอบเขตของการดำเนินงานวิจัย

1.4.1 ใช้กล้องเว็บแคมจำนวน 1 ตัว ที่มีการเชื่อมต่อแบบ USB

1.4.2 ขนาดของพื้นที่ที่สามารถทำงานได้มีความกว้างไม่เกิน 2x2 เมตรและมีผู้ใช้เพียง 1 คนอยู่ในพื้นที่ที่กำหนด

1.4.3 พื้นที่ที่ใช้ทำการทดลองมีการควบคุมสี และการสะท้อนแสง

1.4.4 พัฒนาโปรแกรมประยุกต์ ที่นำวิธีการตรวจวัดตำแหน่งของผู้ใช้มาใช้งานและสามารถทำงานได้แบบเวลาจริง (Real time)

## 1.5 ขั้นตอนการศึกษา

1.5.1 กำหนดหัวข้อ เป้าหมาย วัตถุประสงค์ และ ขอบเขตของการทำวิทยานิพนธ์

1.5.2 ศึกษาทฤษฎี และ หลักการพื้นฐานที่เกี่ยวข้องกับงานวิจัย

1.5.3 ศึกษาวิธีการต่างๆ ที่มีอยู่ แนวความคิด หลักการ ข้อดี และข้อบกพร่องของแต่ละวิธีการ

1.5.4 พัฒนาอัลกอริทึมสำหรับตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ

1.5.5 นำอัลกอริทึมสำหรับตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอมาพัฒนาเป็นโปรแกรม

1.5.6 ทำการทดลอง ปรับปรุง และ สรุปผล

1.5.7 จัดทำเอกสารประกอบวิทยานิพนธ์

## 1.6 ประโยชน์ที่คาดว่าจะได้รับ

สามารถพัฒนาวิธีการและโปรแกรมตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ ที่มีความเหมาะสมกับแอปพลิเคชันราคาถูก เช่น การแสดงผลิตภัณฑ์ และเกมส์ เป็นต้น

## 1.7 โครงสร้างวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้แบ่งเป็นส่วนที่สำคัญดังนี้

บทที่ 1 บทนำ กล่าวถึง ความเป็นมาและความสำคัญของปัญหา ความมุ่งหมายและวัตถุประสงค์ของการศึกษา สมมุติฐานของการศึกษา ขอบเขตของการดำเนินงานวิจัย ขั้นตอนการดำเนินงานวิจัย และประโยชน์ที่คาดว่าจะได้รับ และโครงสร้างวิทยานิพนธ์

บทที่ 2 งานวิจัยที่เกี่ยวข้อง กล่าวถึงงานวิจัยที่ผ่านมาในอดีตถึงปัจจุบัน

บทที่ 3 ทฤษฎีที่เกี่ยวข้องในงานวิจัย

บทที่ 4 อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ กล่าวถึงวิธีการในการคำนวณหาตำแหน่งขาของผู้ใช้ด้วยกล้องวิดีโอ

บทที่ 5 ผลการทดลอง ในการตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ

บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ

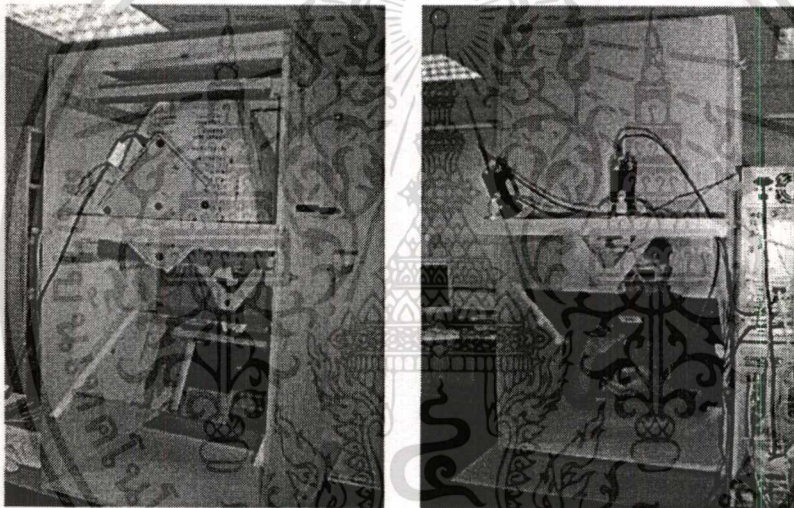
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

### งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวกับการตรวจวัดตำแหน่งสำหรับระบบแสดงผลภาพสเตอริโอเชิงโต้ตอบและการใช้งานในรูปแบบต่าง ๆ สรุปได้ดังนี้

J. D. Mulder and R. van Liere [5] ได้พัฒนา Personal Space Station (PSS) ขึ้นมาซึ่งเป็นระบบ VR/AR แบบตั้งโต๊ะ โดยมีเป้าหมายดังนี้คือ ค่าใช้จ่ายต่ำ ออกแบบง่าย และสามารถโต้ตอบเป็นแบบ 3 มิติได้ ระบบนี้พัฒนาโดยใช้เทคนิคการตรวจวัดตำแหน่งศีรษะโดยใช้กล้องวิดีโอ (Vision based) 2 ตัว ระบบ PSS นี้ ผู้ใช้งานจะนั่งอยู่ด้านหน้าของเครื่องที่มีกระจกสะท้อนภาพสเตอริโอมาจากจอภาพด้านบน ดังรูปที่ 2.1



รูปที่ 2.1 Personal Space Station (PSS)

ระบบที่แสดงในภาพอาศัยกล้องราคาไม่สูงนัก ใดๆก็ดี ระบบดังกล่าวมีความจำเป็นต้องใช้กล้องจำนวน 2 ตัว ซึ่งต้องมีการปรับค่าก่อนนำไปใช้งาน (calibration) นอกจากนี้องค์ประกอบอื่น ๆ เช่น กระจกสะท้อน ทำให้ไม่สะดวกในการติดตั้งใช้งาน อีกทั้งไม่สามารถใช้งานในกรณีที่ผู้ใช้มีการเคลื่อนย้ายตำแหน่ง (เดินไปมา)

Y. Kitamura, T. Konishi, S. Yamamoto and F. Kishino [9] ได้พัฒนาวิธีการแสดงผลภาพสเตอริโอเชิงโต้ตอบที่สามารถใช้ได้กับผู้ใช้ตั้งแต่ 3 คนขึ้นไปดังรูปที่ 3 งานวิจัยนี้จะใช้วิธีการนำแผ่นที่เป็นหน้ากากที่เจาะรูตรงกลางมาปิดทับอุปกรณ์แสดงผลและใช้ Polhemus Fastrack สำหรับตรวจวัดตำแหน่งศีรษะของผู้ใช้ ข้อมูลตำแหน่งของผู้ใช้ที่ได้นั้นจะนำมาคำนวณเพื่อสร้างเอกสารนี้เป็นเอกสารที่ส่งวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่สอดคล้องกับตำแหน่งของผู้ใช้ ซึ่งผู้ใช้แต่ละคนจะมองเห็นภาพผ่านตรงจมองที่ต่างกัน ดังรูปที่ 2.2

ระบบการแสดงผลนี้ถึงแม้จะสามารถใช้งานได้กับผู้ใช้งานหลายคนก็ตาม แต่มีความจำเป็นต้องใช้อุปกรณ์ตรวจวัดตำแหน่งแบบสนามแม่เหล็กไฟฟ้าติดไว้ที่ตัว ซึ่งมีราคาแพง และไม่สะดวกหากมีการเปลี่ยนตำแหน่งของผู้ใช้ เนื่องจากอุปกรณ์ตรวจวัดตำแหน่งจะมีสายต่อออกมาจากอุปกรณ์ตรวจวัด

Chin-Chun Chang and Wen-Hsiang Tsai [10] ได้ทำการพัฒนาวิธีการตรวจวัดท่าทางของขาเพื่อนำไปใช้ในแอปพลิเคชันด้าน Virtual Reality โดยมีเป้าหมายเพื่อให้ผู้ใช้สามารถใช้งานควบคุมการเคลื่อนที่ในโลกเสมือน งานวิจัยนี้จะใช้หลักการเก็บลักษณะการขึ้นในรูปแบบต่างๆ เป็นคำสั่งเพื่อใช้ควบคุมการทำงาน โดยรูปแบบนี้จะถูกเก็บไว้เป็นต้นแบบ ในขั้นตอนการตรวจวัดท่าทางจะติดมาร์คเกอร์ไว้ที่รองเท้าเพื่อเป็นจุดอ้างอิง แล้วใช้กล้องวิดีโอ 1 ตัวสำหรับจับภาพของเท้าเพื่อนำมาประมวลผลรูปแบบการเคลื่อนที่ของเท้าไปเปรียบเทียบกับรูปแบบที่กำหนดไว้ ดังรูปที่ 2.3

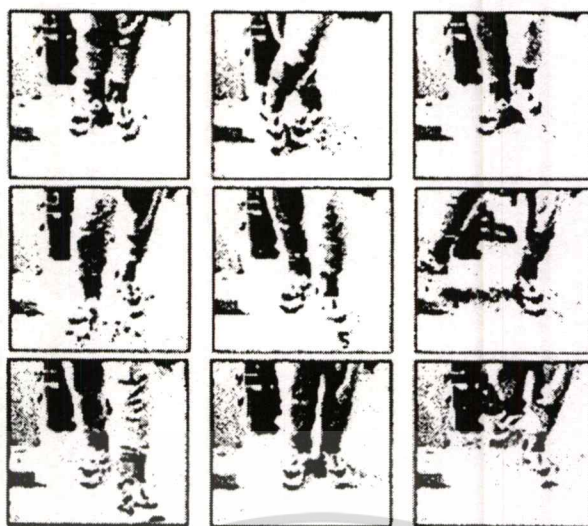
วิธีการนี้สามารถนำมาใช้ทำควบคุมการเคลื่อนที่ในโลกเสมือนได้ แต่ในการทำงานยังจำเป็นต้องมีมาร์คเกอร์ช่วยในการตรวจวัดท่าทาง และมีจุดประสงค์เพื่อการใช้งานในการทำท่าทางของขาในการสั่งงานคอมพิวเตอร์ วิธีดังกล่าวไม่สามารถใช้ในการคำนวณหาตำแหน่งของผู้ใช้ได้



รูปที่ 2.2 วิธีการแสดงผลสเตอริโอเชิงโต้ตอบที่สามารถใช้ได้กับผู้ใช้งานตั้งแต่ 3 คนขึ้นไป ของ

Y. Kitamura, T. Konishi, S. Yamamoto and F. Kishino

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 วิธีการตรวจวัดท่าทางของขา ของ Chin-Chun Chang and Wen-Hsiang Tsai

M. Ribo, A. Pinz, and A. Fuhrmann [12] ได้พัฒนาระบบตรวจวัดตำแหน่งศีรษะด้วยกล้องวิดีโอสำหรับ VR/AR แอปพลิเคชัน ซึ่งสามารถวัดตำแหน่งศีรษะแบบ 6 แกน ได้แก่ตำแหน่งแนวพิกัด  $x$ ,  $y$  และ  $z$  กับองศาการหมุนรอบแกน (yaw, pitch, roll) ในงานวิจัยนี้จะใช้กล้องวิดีโอ 2 ตัวที่เป็นแบบ Progressive scan CCD ความเร็ว 30 เฟรมต่อวินาที และมาร์คเกอร์ติดไว้ที่ Shutter Glass การคำนวณหาตำแหน่งศีรษะในงานวิจัยนี้จะทำการค้นหาตำแหน่งของมาร์คเกอร์ด้วยวิธีการ Blob detection เมื่อได้ตำแหน่งของมาร์คเกอร์ของภาพด้านซ้ายและขวา จึงนำตำแหน่งที่ได้มาคำนวณหา Projective matrix ของภาพด้านซ้ายและขวา เมื่อได้ Matrix ที่ต้องการจึงนำไปคำนวณหาตำแหน่ง 3 มิติโดย วิธีการ 3D Reconstruction

ระบบนี้สามารถที่จะคำนวณตำแหน่งออกมาเป็น 3 มิติได้ แต่มีความจำเป็นต้องติดตั้งมาร์คเกอร์ร่วมด้วย ซึ่งอาจจะทำให้มีความยุ่งยากในการใช้งาน เช่น การติดตั้งมาร์คเกอร์ และการที่ต้องใช้กล้องวิดีโอ 2 ตัวในการตรวจวัดจำเป็นต้องมีการปรับค่าก่อนการใช้งาน และกล้องที่ใช้มีราคาแพง

T. Ohno and N. Mukawa [14] ได้พัฒนาวิธีการโต้ตอบกับคอมพิวเตอร์ด้วยการตรวจวัดตำแหน่งการมองของดวงตา โดยวิธีการนี้จะสามารถเคลื่อนที่ศีรษะได้อย่างอิสระและการ Calibration นั้นทำได้ง่าย ในงานวิจัยนี้มี 2 ระบบคือ 1. ระบบวัดตำแหน่งการมองของดวงตา ใช้กล้องอินฟราเรด 1 ตัวติดตั้งไว้ด้านล่างและสามารถหมุนได้ ใช้สำหรับจับภาพแสดงอินฟราเรดที่ดวงตา เพื่อนำมาคำนวณหาตำแหน่งการมองของดวงตาบนจอภาพ 2. ระบบตรวจวัดตำแหน่งของตา จะประกอบด้วยกล้อง 2 ตัวติดตั้งอยู่ด้านบน ใช้สำหรับจับภาพหน้าของผู้ใช้เพื่อคำนวณหาตำแหน่งของตา การทำงานของระบบนี้จะมีการ Calibration ระบบด้วยการมองจุด 2 จุดบนจอภาพ เอกสารฉบับนี้เผยแพร่โดยศูนย์วิจัยเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี กรุงเทพมหานคร ไม่ว่ากรรมใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 ระบบตรวจวัดตำแหน่งศีรษะด้วยกล้องวิดีโอ ของ M. Ribo, A. Pinz, and A. Fuhrmann

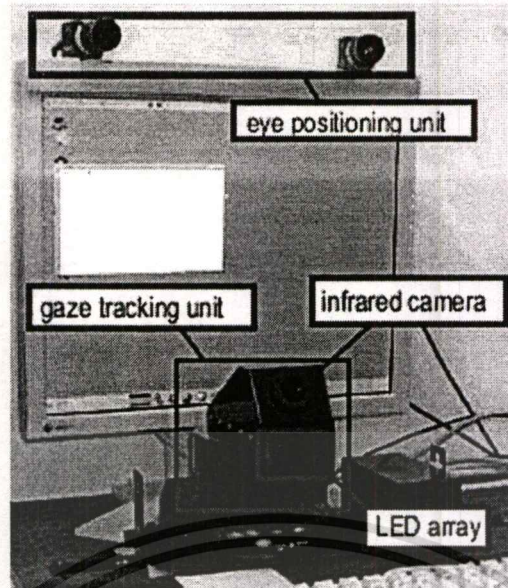
เพื่อปรับความถูกต้อง เมื่อเริ่มใช้งานระบบที่ 1 จะใช้สำหรับจับภาพของดวงตาและคำนวณหาตำแหน่งการมองของดวงตา เมื่อผู้ใช้มีการเปลี่ยนตำแหน่งของศีรษะทำให้ระบบที่ 1 ไม่สามารถคำนวณหาตำแหน่งการมองได้ จึงนำข้อมูลจากระบบที่ 2 มาใช้ในการเลื่อนกล้องไปยังตำแหน่งของตาที่ระบบที่ 2 คำนวณได้

เนื่องจากระบบการตรวจวัดนี้ เป็นระบบการตรวจวัดตำแหน่งที่ผู้ใช้งานจำเป็นต้องต้องนั่งอยู่กับที่ ไม่สามารถใช้งานได้ในกรณีที่ผู้ใช้งานมีการเปลี่ยนตำแหน่ง เช่น การเดิน และจำเป็นจะต้องมีอุปกรณ์ประกอบหลายอย่าง ซึ่งทำให้ระบบการตรวจวัดมีราคาแพง

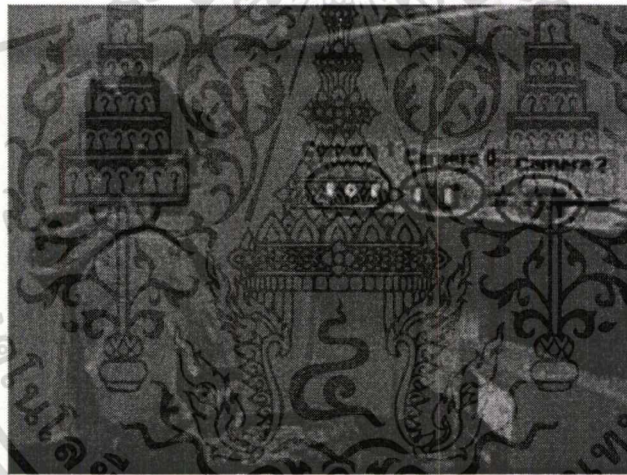
R. Ruddaraju, A. Haro, K. Nagel, Quan T. Tran, Irfan A. Essa, G. Abowd and E. D. Mynatt [15] ได้พัฒนาระบบโต้ตอบกับผู้ใช้อยู่ด้วยวิธีการตรวจวัดตำแหน่งของตา งานวิจัยนี้จะใช้กล้องวิดีโอแบบอินฟราเรด 3 ตัวในการจับภาพดวงตาเพื่อคำนวณหาตำแหน่งของศีรษะ และใช้วิธีการ Fisher's linear discriminant สำหรับค้นหาตำแหน่งของดวงตาซึ่งวิธีการนี้จะไม่ขึ้นอยู่กับ การเปลี่ยนแปลงของแสง โดยความถูกต้องโดยรวมอยู่ที่ 81.25-87.25 %

งานวิจัยดังกล่าวมีความจำเป็นต้องใช้กล้องวิดีโอแบบอินฟราเรด 3 ตัว ซึ่งมีราคาแพง และจำเป็นต้องมีการปรับค่าเริ่มต้นก่อนการใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5 วิธีการตรวจวัดตำแหน่งของตา ของ T. Ohno and N. Mukawa



รูปที่ 2.6 วิธีการตรวจวัดตำแหน่งของตา ของ R. Ruddaraju, A. Haro, K. Nagel, Quan T. Tran, Irfan A. Essa, G. Abowd and E. D. Mynatt

J. Rekimoto [16] ได้พัฒนาวิธีการตรวจวัดตำแหน่งของศีรษะสำหรับ Fish Tank VR โดยไม่มีอุปกรณ์เสริมบนศีรษะ การคำนวณหาตำแหน่งของศีรษะจะใช้กล้องวิดีโอ 1 ตัวสำหรับถ่ายภาพหน้าของผู้ใช้ โดยจะทำการลบภาพพื้นหลังออกด้วยวิธีการ Image Subtraction เมื่อได้ภาพหน้าของผู้ใช้แล้ว จึงนำภาพหน้ามาทำการค้นหาตำแหน่งของศีรษะผู้ใช้ด้วยวิธีการ template matching หลังจากที่ได้ตำแหน่งของศีรษะในภาพแล้วจึงนำตำแหน่งที่ได้มาคำนวณตำแหน่งของศีรษะใน World Coordinate

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการนี้จะใช้วิธีการ template matching ซึ่งเมื่อเปลี่ยนผู้ใช้จำเป็นจะต้องมีการปรับค่าเริ่มต้นก่อนการใช้งานใหม่ และในการใช้งานผู้ใช้ไม่สามารถเปลี่ยนตำแหน่งได้ เช่น การเดิน



รูปที่ 2.7 วิธีการตรวจวัดตำแหน่งของศีรษะสำหรับ Fish Tank VR ของ J. Rekimoto



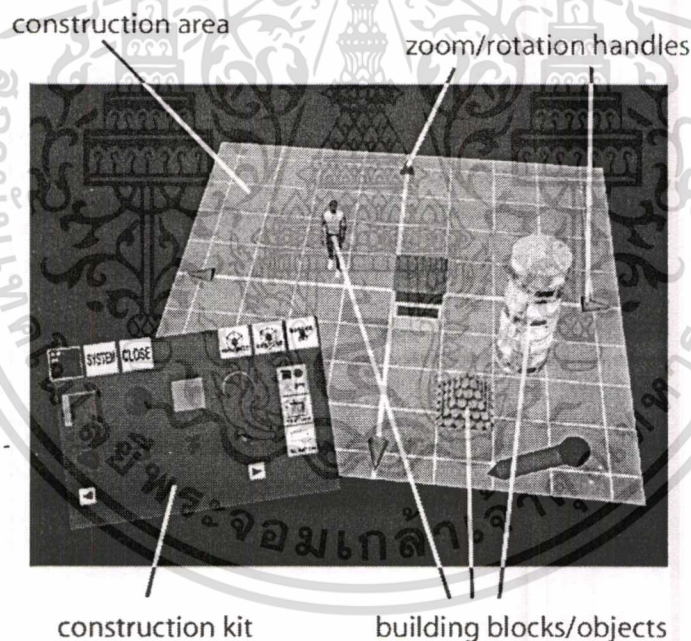
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 3

# ทฤษฎีที่เกี่ยวข้องในงานวิจัย

### 3.1 คอมพิวเตอร์กราฟิก

คอมพิวเตอร์กราฟิก(Computer graphics) หรือเรียกย่อๆ ว่า ซีจี (CG) เป็นการนำเครื่องคอมพิวเตอร์ประมวลผลข้อมูลซึ่งอยู่ในรูปของข้อมูลตัวเลข ตัวอักษร หรือสัญลักษณ์ต่างๆ แล้วแสดงผลลัพธ์ทางจอภาพเป็นภาพที่มนุษย์สามารถเข้าใจได้ เช่น รูปทรงเรขาคณิต สี สัน ลวดลายและแสงเงาต่างๆ เป็นต้น ในปัจจุบันเราสามารถพบเห็นการนำคอมพิวเตอร์กราฟิกไปประยุกต์ใช้งานในหลายสาขา เช่น วิทยาศาสตร์ งานศิลปะ วิศวกรรม ธุรกิจ อุตสาหกรรม การแพทย์ การทหาร ความบันเทิง งานโฆษณา การศึกษา การฝึกอบรม และอำนวยความสะดวกในบ้านเรือน รูปที่ 3.1 แสดงถึงการนำคอมพิวเตอร์กราฟิกไปใช้งานในด้านระบบการแสดงผลภาพเสมือนจริง

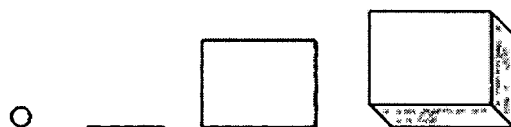


รูปที่ 3.1 การประยุกต์ใช้งานคอมพิวเตอร์กราฟิก ด้านระบบการแสดงผลภาพเสมือนจริง [19]

#### 3.1.1 มิติและระบบพิกัดในคอมพิวเตอร์กราฟิก

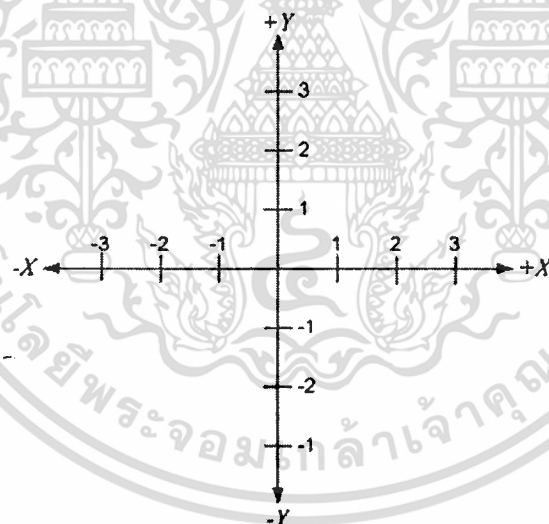
การสร้างภาพในคอมพิวเตอร์กราฟิก เป็นกระบวนการวาดภาพลงบนพื้นที่ว่างที่กำหนด ซึ่งองค์ประกอบเบื้องต้นของภาพพื้นฐานในคอมพิวเตอร์กราฟิกคือ จุด ซึ่งมีการแสดงออกมาเพียงตำแหน่ง แต่ไม่มีทิศทางและมิติของจุด ถ้ามีการกำหนดจุดขึ้นมา 2 จุด แล้วทำการสร้างจุดให้จุดสองจุดดูเหมือนจะมีการเชื่อมต่อกันจะเรียกว่าเส้น จากเส้นเมื่อมีการวาดเส้นเป็นรูปร่างต่างๆ เช่น สี่เหลี่ยมก็จะเรียกว่า รูปร่างสองมิติ หรือระนาบ และจากรูปทรงสี่เหลี่ยมถ้ามีการวาดรูปสี่เหลี่ยมให้เอียงเอกริมเป็นเอียงหรือทแยงมุมสำหรับใช้ในการแข่งขันเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้เอนกประสงค์ในการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นรูปในลักษณะรูปลูกบาศก์ก็จะเรียกว่ารูปทรงปริมาตร ซึ่งมีลักษณะเป็นวัตถุ 3 มิติ ดังนั้นการสร้างภาพในคอมพิวเตอร์กราฟิกจะสามารถแสดงภาพได้ใน 2 รูปแบบคือ ภาพแบบ 2 มิติ และภาพแบบ 3 มิติ ดังรูปที่ 3.2



รูปที่ 3.2 จุด เส้น รูป 2 มิติ และรูป 3 มิติ

รูปแบบการแสดงผลแบบ 2 มิติที่ใช้งานในระบบคอมพิวเตอร์กราฟิกนั้น ได้ใช้หลักการของนักคณิตศาสตร์และนักปราชญ์ที่ชื่อ Rene Descartes ซึ่งเรียกวิธีการของนักปราชญ์นี้ว่า ระบบพิกัดคาร์ทีเซียน (Cartesian coordinates) ในระบบพิกัดนี้จะใช้วิธีการแสดงระบบพิกัดโดยการลากเส้นตรง 2 เส้นตัดกันโดยเส้นตรงทั้งสองเส้นนั้นจะตั้งฉากกัน โดยกำหนดให้เส้นในแนวตั้งคือแกน  $Y$  และเส้นในแนวนอนคือแกน  $X$  ดังรูปที่ 3.3

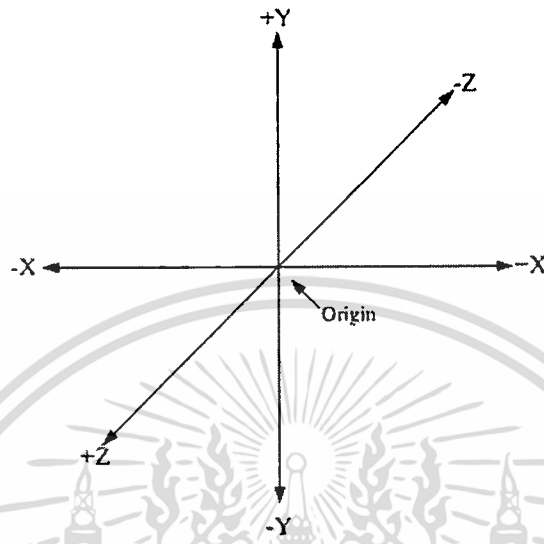


รูปที่ 3.3 ระบบพิกัดคาร์ทีเซียนแบบ 2 มิติ

ระบบพิกัดคาร์ทีเซียนจะประกอบด้วยพื้นที่ที่เกิดจากการตัดกันของเส้นตรงสองเส้น จุดที่เส้นตรงสองเส้นตัดกันเรียกว่าจุดกำเนิด (Origin) ในแต่ละแกนจะมีการแบ่งมาตรโดยมีค่าเริ่มต้นเป็นศูนย์จากจุดกำเนิดและเพิ่มค่าในระยะที่เท่ากันไปตามแกนทั้งสองด้านด้วยค่าบวกและลบ การที่จะอ่านตำแหน่งของจุดในระบบพิกัดคาร์ทีเซียน ทำได้โดยการอ่านค่าจากมาตรบนแกน  $X$  และแกน  $Y$  หรือถ้าต้องการกำหนดตำแหน่งของจุดทำได้โดยการกำหนดค่าแกน  $X$  และแกน  $Y$  ลงไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบพิกัดคาร์ทีเซียนแบบ 3 มิติจะมีการเพิ่มแกนขึ้นมาอีก 1 แกนคือแกน  $Z$  เป็นแกนที่ลากผ่านจุดกำเนิดและตั้งฉากกับแกน  $X$  และแกน  $Y$  ดังรูปที่ 3.4 การอ่านค่าตำแหน่งหรือการระบุตำแหน่งของจุดจะเหมือนกับระบบแบบ 2 มิติ แต่จะมีค่าเพิ่มขึ้นมาคือค่าของแกน  $Z$



รูปที่ 3.4 ระบบพิกัดคาร์ทีเซียนแบบ 3 มิติ

### 3.1.2 การแปลงเชิงเรขาคณิต (Geometric transformation)

การแปลงเชิงเรขาคณิตในคอมพิวเตอร์กราฟิกเป็นการเปลี่ยนตำแหน่งหรือขนาดของจุดหรือภาพที่วาด ซึ่งได้แก่ การเคลื่อนย้ายภาพ (Translation), การปรับขนาดภาพ (Scaling) และ การหมุนภาพ (Rotation)

#### 3.1.2.1 การแปลงเชิงเรขาคณิต 2 มิติ (Two dimensional geometric transformations)

การแปลงเชิงเรขาคณิต 2 มิติสำหรับคอมพิวเตอร์โดยทั่วไปจะใช้เมทริกซ์ ขนาด  $3 \times 3$  เนื่องจากสามารถนำมาใช้ในการเขียนโปรแกรมด้วยคอมพิวเตอร์ได้ง่าย และสามารถผสมผสานการแปลงภาพแบบต่างๆ ไว้ด้วยกันได้ และในการคำนวณด้วยเมทริกซ์นั้นจำเป็นจะต้องมีการเปลี่ยนพิกัด 2 มิติให้อยู่ในรูปของ พิกัด โฮโมจีเนียส (Homogeneous coordinates) เสียก่อน เช่น พิกัด  $(x, y)$  จะเขียนอยู่ในรูปของ  $[x \ y \ h]$  โดยทั่วไปจะกำหนดให้  $h$  มีค่าเท่ากับ 1

1. การเคลื่อนย้ายภาพ (translation) หมายถึง จุดใดๆ ที่ถูกวาดในระบบพิกัดคาร์ทีเซียนแบบ 2 มิติถูกเคลื่อนย้ายไปยังตำแหน่งต่างๆ โดยการเปลี่ยนค่าพิกัด  $(x, y)$  โดยสามารถเขียนได้ดังสมการ

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ  $x, y$  = ค่าพิกัดเดิม  
 $t_x, t_y$  = ระยะทางที่ต้องการเคลื่อนย้ายไปตามแกนอนและตั้ง  
 $x', y'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

2. การปรับขนาดภาพ (Scaling) คือการยืดหรือหดภาพ ซึ่งทำให้ภาพที่ได้มีขนาดเล็กหรือใหญ่ขึ้น โดยสามารถเขียนได้ดังสมการ

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.2)$$

เมื่อ  $x, y$  = ค่าพิกัดเดิม  
 $s_x, s_y$  = ขนาดของภาพที่ต้องการตามแกนอนและตั้ง  
 $x', y'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

3. การหมุนภาพ (Rotation) คือการหมุนภาพ โดยยึดเอาจุดกำเนิด (Origin) เป็นจุดหมุน สามารถเขียนได้ดังสมการ

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.3)$$

เมื่อ  $x, y$  = ค่าพิกัดเดิม  
 $\theta$  = ค่าองศาในการหมุนภาพ  
 $x', y'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

### 3.1.2.2 การแปลงเชิงเรขาคณิต 3 มิติ (Three dimensional geometric transformations)

การแปลงเชิงเรขาคณิต 3 มิติสำหรับคอมพิวเตอร์จะใช้เมทริกซ์ ขนาด 4x4 และในการคำนวณด้วยเมทริกซ์นั้นจำเป็นจะต้องมีการเปลี่ยนพิกัด 3 มิติให้อยู่ในรูปของ พิกัดโฮโมจีเนียส (Homogeneous coordinates) เสียก่อน เช่น พิกัด  $(x, y, z)$  จะเขียนอยู่ในรูปของ  $[x \ y \ z \ h]$  โดยทั่วไปจะกำหนดให้  $h$  มีค่าเท่ากับ 1

1. การเคลื่อนย้ายภาพ (Translation) คือการเคลื่อนย้ายภาพไปยังตำแหน่งต่างๆ ในพิกัด 3 มิติ โดยสามารถเขียนได้ดังสมการ

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.4)$$

เมื่อ  $x, y, z$  = ค่าพิกัดเดิม  
 $t_x, t_y, t_z$  = ระยะทางที่ต้องการเคลื่อนย้ายไปตามแกนอนและตั้ง  
 $x', y', z'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

2. การปรับขนาดภาพ (Scaling) คือการยืดหรือหดภาพ ซึ่งทำให้ภาพที่ได้มีขนาดเล็กหรือใหญ่ขึ้น โดยสามารถเขียนได้ดังสมการ

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.5)$$

เมื่อ  $x, y, z$  = ค่าพิกัดเดิม  
 $s_x, s_y, s_z$  = ขนาดของภาพที่ต้องการตามแกนอนและตั้ง  
 $x', y', z'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

3. การหมุนภาพ (Rotation) คือการหมุนภาพ โดยยึดแกน  $x, y, z$  เป็นจุดหมุน สามารถเขียนได้ดังสมการ

การหมุนโดยยึดแกน  $x$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.6)$$

การหมุนโดยยึดแกน  $y$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.7)$$

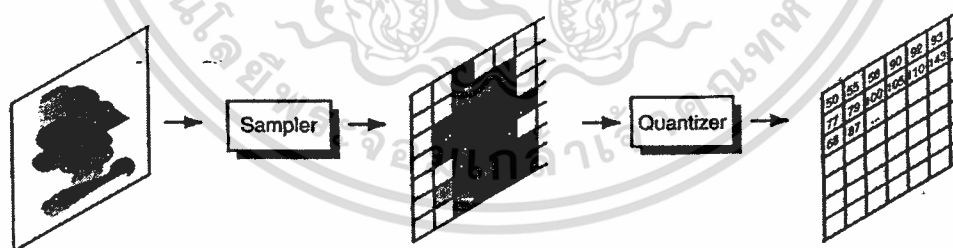
การหมุนโดยยึดแกน  $z$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.8)$$

เมื่อ  $x, y, z$  = ค่าพิกัดเดิม  
 $\theta$  = ค่าองศาในการหมุนภาพ  
 $x', y', z'$  = ค่าพิกัดใหม่ที่คำนวณได้จากสมการ

### 3.2 รูปภาพดิจิทัล (Digital image)

รูปภาพที่มนุษย์เรามองเห็นโดยทั่วไป คือการสะท้อนของแสงที่ออกมาจากวัตถุแล้วมาตกกระทบที่ดวงตาของมนุษย์ โดยแหล่งกำเนิดแสงอาจจะมาจากแสงของดวงอาทิตย์หรือแสงไฟ ซึ่งภาพที่มนุษย์เห็นนั้นจะเป็นภาพที่มีลักษณะต่อเนื่อง (Continuous) โดยในภาพจะเป็นค่าการเปลี่ยนแปลงของความสว่างและสี ถ้าต้องการนำภาพมาประมวลผลกับเครื่องคอมพิวเตอร์ก็ต้องมีการแปลงภาพให้อยู่ในรูปของข้อมูลตัวเลขที่คอมพิวเตอร์สามารถนำไปประมวลผลได้ ดังนั้นจำเป็นจะต้องนำรูปภาพที่มีลักษณะต่อเนื่อง (Continuous Image) มาผ่านกระบวนการแปลงภาพให้อยู่ในรูปของภาพดิจิทัล ซึ่งกระบวนการดังกล่าวจะประกอบไปด้วย 2 ขั้นตอน คือ การแซมปลิง (Sampling) และการปรับในเชิงปริมาณ (Quantization) [1] ดังภาพที่ 3.5



รูปที่ 3.5 กระบวนการแปลงภาพธรรมชาติ (Continuous image) เป็นภาพดิจิทัล (Digital image)

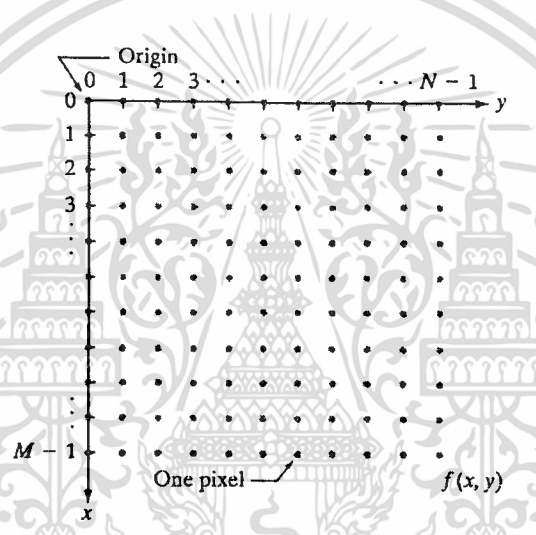
ขั้นตอนแรกคือการแซมปลิง โดยขั้นตอนนี้จะทำการเปลี่ยนความสว่างที่อยู่ในภาพต่อเนื่อง ออกเป็นค่าความสว่างที่ไม่ต่อเนื่องซึ่งก็ขึ้นอยู่กับว่าต้องการขนาดของภาพเท่าใด เช่น 200x200 หลังจากนั้นขั้นตอนต่อไปคือการปรับในเชิงปริมาณจะเปลี่ยนค่าความสว่างที่ได้ให้เป็นค่าตัวเลข ซึ่งจะอยู่ในขอบเขตของสี (Gray level หรือ Brightness intensity) ที่มีระดับตั้งแต่ค่าเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จนถึงขาว (0 ถึง 255) หลังจากนั้นข้อมูลที่ได้จะถูกเก็บอยู่ในรูปของอาเรย์ขนาด 2 มิติ ซึ่งแต่ละตำแหน่งของอาเรย์จะแทนค่าแต่ละจุดในรูปภาพดิจิทัล แต่ละตำแหน่งของอาเรย์จะเรียกว่าจุดภาพ (Pixel)

จากจุดภาพที่อยู่ภายในอาเรย์ 2 มิติ สามารถกำหนดให้อยู่ในรูปของระบบพิกัดเชิงเส้น (Cartesian coordinate) ได้โดยกำหนดให้  $(x, y)$  แทนตำแหน่งใดๆ ของจุดภายในอาเรย์ในแนวนอนและแนวตั้งตามลำดับของรูปภาพดิจิทัลซึ่งมีความสัมพันธ์กับรูปภาพต่อเนื่อง ซึ่งเมื่อนำตำแหน่งที่มีความสัมพันธ์กันนี้มาประกอบกันก็จะสามารถแสดงผลเป็นภาพได้

จากอาเรย์ของภาพดิจิทัล 2 มิติสามารถเขียนให้อยู่ในรูปของฟังก์ชัน  $f(x,y)$  มีขนาด  $M$  แถว และ  $N$  คอลัมน์ โดยที่  $M$  คือความสูงของภาพและ  $N$  คือความกว้างของภาพ ดังรูปที่ 3.6 และสามารถเขียนอยู่ในรูปของเมตริกขนาด  $M \times N$  ดังสมการที่ 3.9



รูปที่ 3.6 ระบบพิกัดของภาพดิจิทัล

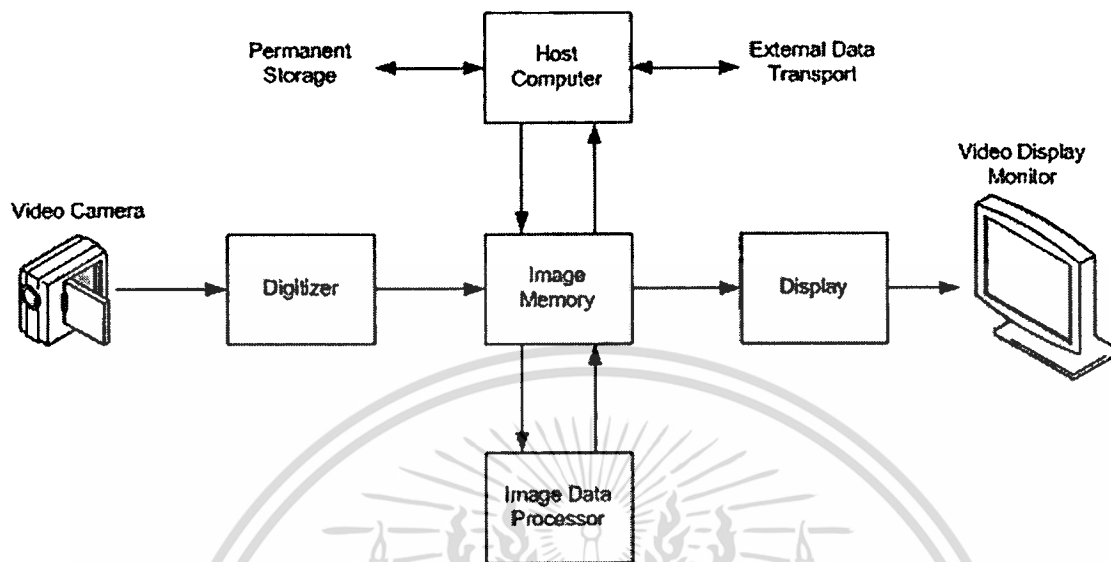
$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1) & \dots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix} \tag{3.9}$$

### 3.3 การประมวลผลภาพดิจิทัล (Digital image processing)

การประมวลผลภาพดิจิทัลหมายถึงกระบวนการหรือวิธีการใดๆ ที่กระทำกับภาพดิจิทัลเพื่อที่จะแก้ไข ปรับปรุง หรือวิเคราะห์ภาพดิจิทัล เพื่อทำการปรับปรุงภาพดิจิทัลให้ดีขึ้น หรือแยกส่วนของภาพที่ต้องการออกมา ซึ่งกระบวนการนี้จะใช้เครื่องคอมพิวเตอร์และซอฟต์แวร์ที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประกอบด้วยฟังก์ชันทางคณิตศาสตร์สำหรับดำเนินการกับรูปภาพในลักษณะต่างๆ รูปที่ 3.7 แสดงส่วนประกอบของระบบการประมวลผลภาพดิจิทัลเบื้องต้น



รูปที่ 3.7 ส่วนประกอบของระบบการประมวลผลภาพดิจิทัลเบื้องต้น

### 3.3.1 กระบวนการพีชคณิต ( Algebraic Operations )

การประมวลผลภาพแบบพีชคณิตคือการนำภาพดิจิทัล 2 ภาพมาผ่านกระบวนการ หาผลบวก หาคความแตกต่าง หาผลคูณ หรือ ผลหาร โดยกระบวนการนี้จะเป็นการประมวลผลแบบจุดภาพกับจุดภาพ โดยมีสมการคือ

$$C(x, y) = A(x, y) + B(x, y) \quad (3.10)$$

$$C(x, y) = A(x, y) - B(x, y) \quad (3.11)$$

$$C(x, y) = A(x, y) * B(x, y) \quad (3.12)$$

$$C(x, y) = A(x, y) / B(x, y) \quad (3.13)$$

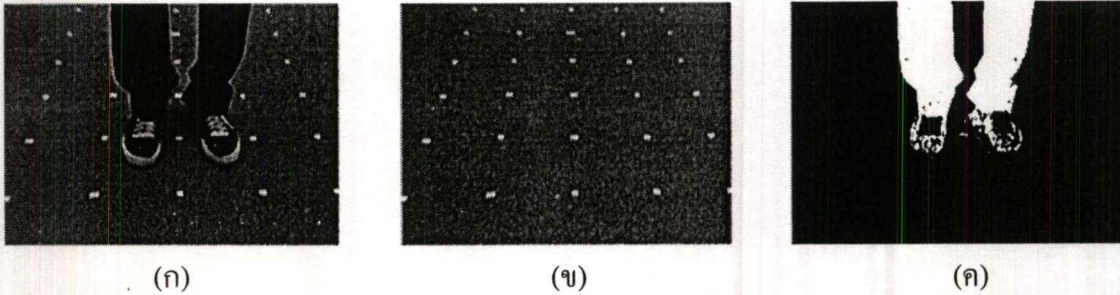
เมื่อ  $A(x, y)$  = ภาพอินพุตที่นำมาใช้กับกระบวนการพีชคณิต

$B(x, y)$  = ภาพอินพุตที่นำมาใช้กับกระบวนการพีชคณิต

$C(x, y)$  = ภาพผลลัพธ์ที่ได้จากกระบวนการพีชคณิต

1) การลบภาพ (Image subtraction) เป็นวิธีการที่สามารถใช้ลบสิ่งที่ไม่ต้องการในภาพออกไปได้ โดยสามารถนำไปประยุกต์ใช้งานได้หลายรูปแบบ เช่น การลบภาพพื้นหลังจากภาพออก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



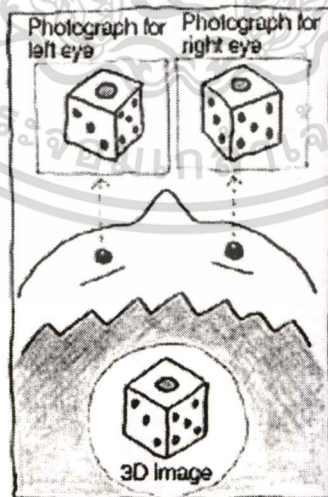
รูปที่ 3.8 การลบพื้นหลังจากภาพออก (ก) ภาพที่ต้องการลบพื้นหลัง, (ข) ภาพพื้นหลัง และ (ค) ภาพผลลัพธ์

จากรูปที่ 3.8 เป็นตัวอย่างการนำการลบภาพมาประยุกต์ใช้งาน วิธีการนี้จะเรียกว่า การลบภาพพื้นหลัง (Background Subtraction) โดยนำไปใช้สำหรับภาพที่ต้องการลบภาพพื้นหลังออก เพื่อให้เหลือเฉพาะสิ่งที่ต้องการเท่านั้น

### 3.4 การแสดงภาพสเตอริโอ (Stereoscopic display)

#### 3.4.1 การมองเห็นภาพ 3 มิติ

การมองเห็นภาพโดยปกติของมนุษย์จะเห็นภาพในลักษณะของภาพ 3 มิติ คือมีความลึกของภาพที่เห็น ทำให้มนุษย์สามารถระบุระยะของวัตถุที่มองเห็นจากภาพได้ การมองเห็นภาพ 3 มิติของมนุษย์ เกิดจากการที่สมองของมนุษย์นำภาพที่มองเห็นผ่านจากดวงตาข้างซ้ายและดวงตาข้างขวา มาสร้างเป็นภาพขึ้นใหม่ทำให้มนุษย์สามารถมองเห็นภาพเป็น 3 มิติได้ ดังรูปที่ 3.9



รูปที่ 3.9 หลักการมองเห็นภาพ 3 มิติของมนุษย์ [20]

### 3.4.2 การแสดงภาพสเตอริโอด้วยคอมพิวเตอร์

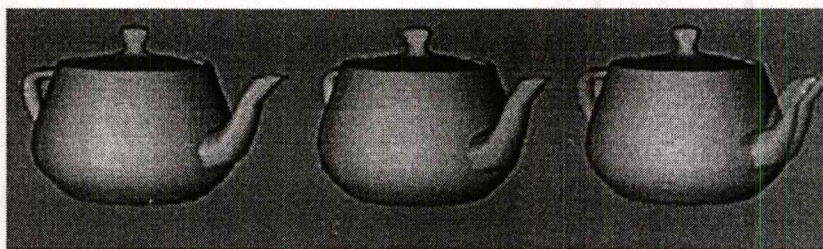
การแสดงภาพบนจอภาพคอมพิวเตอร์ที่มนุษย์มองเห็น โดยทั่วไปจะเป็นภาพแบบ 2 มิติ คือ มีความกว้างและความยาวของภาพเท่านั้น การที่จะทำให้มนุษย์สามารถเห็นภาพจากจอภาพคอมพิวเตอร์เป็นภาพ 3 มิติแบบที่เห็นภาพตามปกติของมนุษย์ จึงต้องใช้วิธีการแสดงภาพเลียนแบบการมองเห็นภาพ 3 มิติของมนุษย์ โดยจะสร้างภาพบนจอภาพซึ่งภาพที่สร้างขึ้นจะมีความสัมพันธ์กับภาพที่ตาข้างซ้ายและตาข้างขวาจะมองเห็น และใช้แว่นตา (Shutter glasses) เป็นเครื่องมือช่วยทำให้การมองเห็นภาพบนจอภาพมีความสัมพันธ์กับภาพที่จะมองเห็นผ่านตาซ้ายและตาขวา อุปกรณ์สำหรับระบบการแสดงผลสเตอริโอด้วยคอมพิวเตอร์ ดังรูปที่ 3.10



รูปที่ 3.10 อุปกรณ์สำหรับระบบการแสดงผลสเตอริโอด้วยคอมพิวเตอร์ [2]

หลักการมองเห็นภาพ 3 มิติจากจอภาพคอมพิวเตอร์นั้น จะอาศัยอุปกรณ์ 2 ชิ้นด้วยกันคือ จอภาพ และแว่นตา (Shutter glasses) ดังรูปที่ 3.10 เครื่องคอมพิวเตอร์จะทำหน้าที่สร้างภาพสำหรับดวงตาข้างซ้ายและดวงตาข้างขวา ซึ่งถ้ามองด้วยตาเปล่าภาพที่จอคอมพิวเตอร์จะเห็นเหมือนเป็นภาพซ้อน ดังรูปที่ 3.11 ดังนั้นถ้าจะทำให้มองเห็นภาพสำหรับดวงตาข้างซ้ายและดวงตาข้างขวาได้ จึงต้องใช้แว่นตาช่วย วัสดุที่ทำเป็นเลนส์ของแว่นตาจะทำด้วย LCD แว่นตาจะมีหน้าที่ปิดเปิดภาพที่เห็นให้มีความสัมพันธ์กับภาพที่สร้างขึ้นที่หน้าจอ โดยเมื่อภาพของตาซ้ายปรากฏขึ้นที่จอภาพ แว่นตาทางด้านขวาจะปิดและเปิดตาซ้ายทำให้มองเห็นภาพของตาซ้ายที่จอภาพ และเมื่อภาพของตาขวาปรากฏขึ้นที่จอภาพแว่นตาทางด้านซ้ายจะปิดและเปิดตาขวาทำให้มองเห็นภาพของตาขวาที่จอภาพ จากหลักการนี้จึงทำให้สามารถมองเห็นภาพที่หน้าจอคอมพิวเตอร์เป็นภาพ 3 มิติได้ ดังรูปที่ 3.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

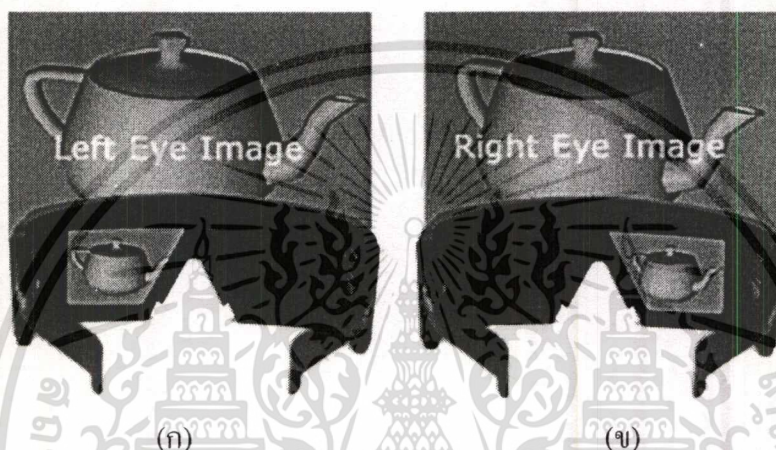


(ก)

(ข)

(ค)

รูปที่ 3.11 ภาพ 3 มิติที่สร้างด้วยคอมพิวเตอร์ (ก)ภาพของตาซ้าย, (ข)ภาพของตาขวา และ (ค)ภาพเมื่อมองด้วยตาเปล่า [21]



(ก)

(ข)

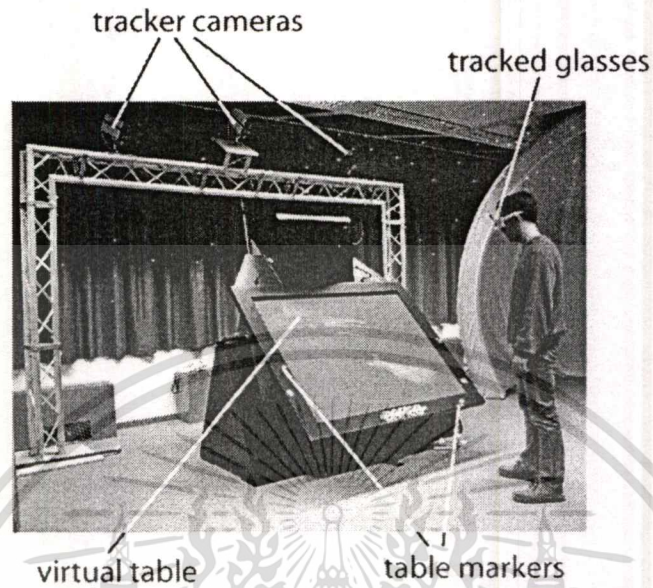
รูปที่ 3.12 ภาพที่มองผ่านแว่นตา (Shutter glasses) (ก) ภาพที่มองเห็นผ่านตาซ้าย และ(ข) ภาพที่มองเห็นผ่านตาขวา [21]

#### 3.4.3 ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ (Interactive stereoscopic viewing system)

จากการที่สามารถนำเครื่องคอมพิวเตอร์และแว่นตา (Shutter glasses) มาใช้งานร่วมกันทำให้สามารถแสดงผลภาพ 3 มิติเลียนแบบการมองเห็นภาพ 3 มิติของมนุษย์ได้ แต่การแสดงผลภาพสเตอริโอโดยปกติมีข้อจำกัด เช่น เมื่อผู้ใช้งานเปลี่ยนตำแหน่งในการมองภาพ ภาพบนจอภาพก็ไม่สามารถที่จะสร้างภาพให้มีความสอดคล้องกับตำแหน่งของผู้ใช้งานที่เปลี่ยนไปได้ ทำให้การมองเห็นภาพสเตอริโอที่ได้ยังไม่สมบูรณ์แบบเหมือนกับการมองเห็นภาพ 3 มิติตามธรรมชาติของมนุษย์

จากข้อจำกัดของการแสดงผลภาพสเตอริโอโดยปกติ จึงได้มีการพัฒนาระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ (Interactive stereoscopic viewing system) ขึ้นมาเพื่อแก้ไขข้อจำกัดดังกล่าว การแสดงผลภาพสเตอริโอเชิงโต้ตอบจะสามารถสร้างภาพที่มีความสัมพันธ์กับผู้ใช้งานที่เปลี่ยนตำแหน่งในการมองภาพได้ จึงทำให้ระบบการแสดงผลภาพสเตอริโอมีประสิทธิภาพมากขึ้น โดยการสร้างภาพที่มีความสัมพันธ์กับผู้ใช้งานที่เปลี่ยนตำแหน่งจะใช้วิธีการตรวจวัดตำแหน่งของผู้ใช้ ซึ่ง

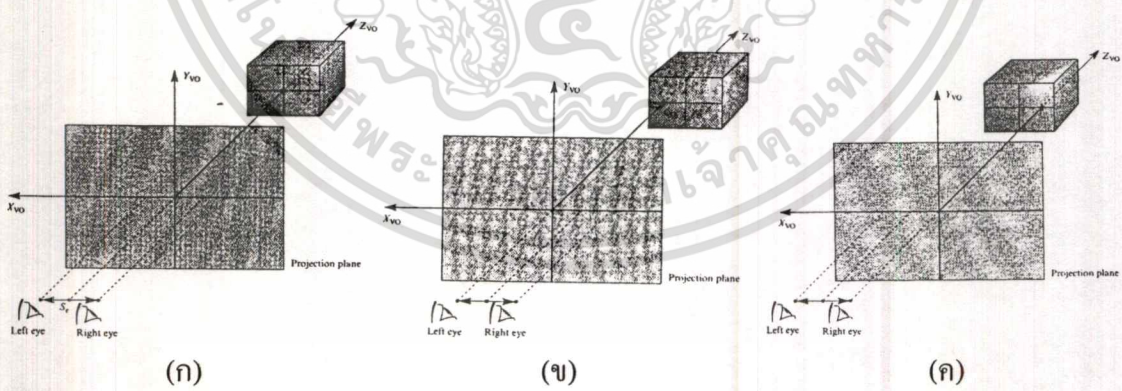
อาจจะใช้เครื่องมือและวิธีการต่างๆ เข้ามาช่วยในการตรวจวัดตำแหน่งของผู้ใช้ ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ ดังรูปที่ 3.13



รูปที่ 3.13 ระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ [22]

3.4.4 หลักการสร้างภาพสเตอริโอแบบเพอร์สเปกทีฟ [18]

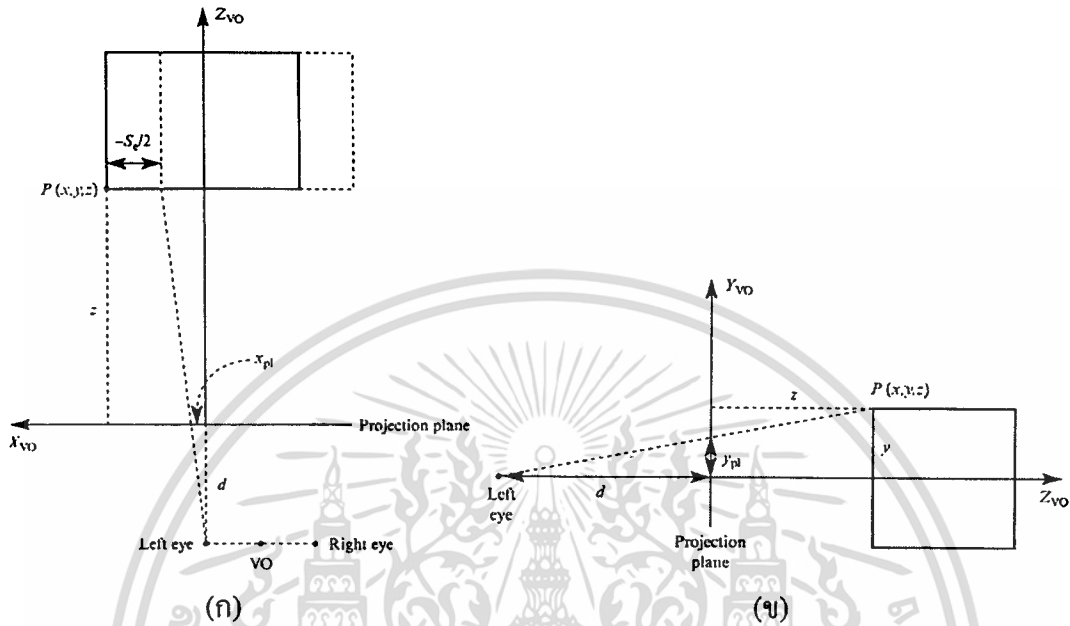
การสร้างภาพสเตอริโอด้วยคอมพิวเตอร์ สามารถนำหลักการจากการสร้างภาพวัตถุ 3 มิติแบบเพอร์สเปกทีฟมาใช้ในการสร้างภาพสำหรับตาข้างซ้ายและตาข้างขวาได้



รูปที่ 3.14 หลักการสร้างภาพสเตอริโอแบบเพอร์สเปกทีฟ

จากรูปที่ 3.14 (ก) จะเห็นว่าดวงตาทั้ง 2 ข้างเพ็งมองไปที่กล่องสี่เหลี่ยม ถ้าต้องการสร้างภาพสำหรับดวงตาข้างซ้าย จะกำหนดให้จุดมองของตาทั้งสองข้างอยู่ที่จุดเริ่มต้น แล้วทำการเลื่อนกล่องสี่เหลี่ยมไปทางด้านขวา ดังรูปที่ 3.14 (ข) และทำการสร้างภาพเพอร์สเปกทีฟของกล่อง ถ้าเอกสารนี้เป็นเอกสารทสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยญาติเห็นาไปเซประเยชนคานการค้ำไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต้องการสร้างภาพสำหรับดวงตาข้างขวาจะกำหนดให้จุดมองของตาทั้งสองข้างอยู่ที่จุดเริ่มต้น แล้วทำการเลื่อนกล้องสี่เหลี่ยมไปทางด้านซ้าย ดังรูปที่ 3.14 (ค) และทำการสร้างภาพเพอร์สเปกทีฟของกล้อง



รูปที่ 3.15 หลักการฉายภาพนรนาบ

จากรูปที่ 3.15 (ก) แสดงให้เห็นถึงจุดมองเสมือน (Virtual observer) ของดวงตาข้างซ้ายที่จุดเริ่มต้น โดยจุดมองภาพจะมีระยะห่างจากจุดเริ่มต้นเท่ากับ  $d$  และมีระยะห่างระหว่างดวงตาทั้งสองข้างเท่ากับ  $S_e$  ที่ตำแหน่งของวัตถุ  $P(x, y, z)$  จะมีการเลื่อนตำแหน่งไปเท่ากับ  $-S_e/2$  ดังนั้นจึงสามารถเขียนเป็นสมการได้ดังนี้

$$\frac{x_{pl}}{d} = \frac{x - \frac{S_e}{2}}{z + d} \tag{3.14}$$

ดังนั้น

$$x_{pl} = \frac{d \left[ x - \frac{S_e}{2} \right]}{z + d} \tag{3.15}$$

จากรูปที่ 3.15 (ข) แสดงให้เห็นถึงจุดมองเสมือนด้านข้างของดวงตาข้างซ้าย และสามารถเขียนเป็นสมการได้ดังนี้

$$\frac{y_{pl}}{d} = \frac{y}{z + d} \tag{3.16}$$

ดังนั้น

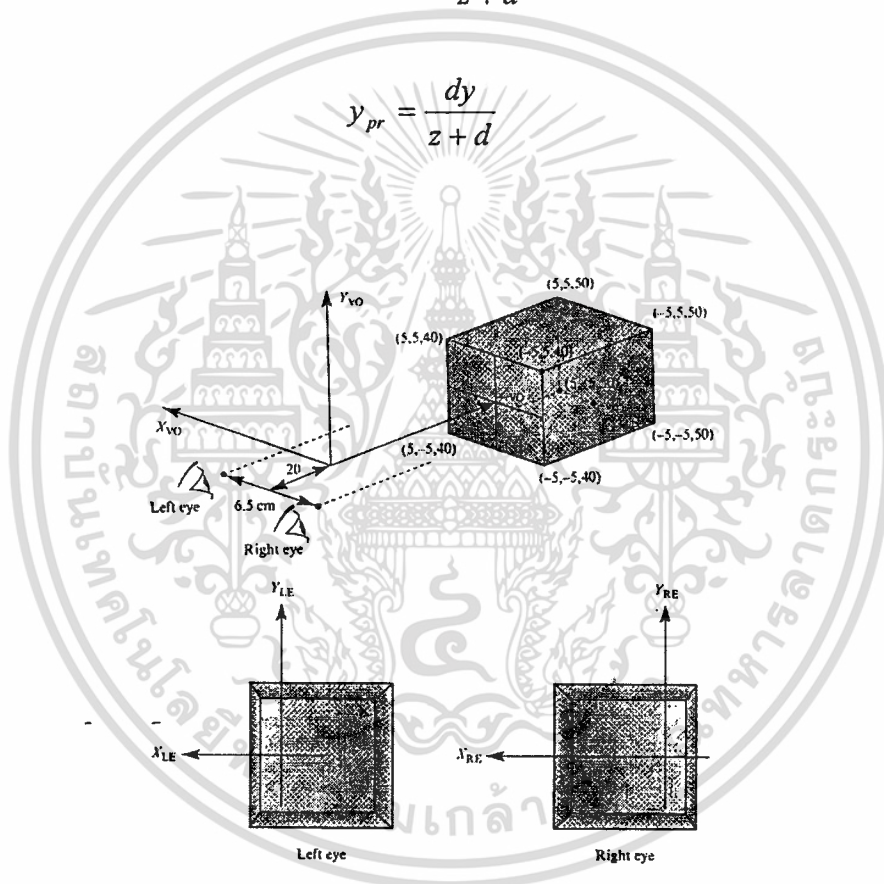
$$y_{pr} = \frac{dy}{z+d} \quad (3.17)$$

ดังนั้นที่จุด  $P(x,y,z)$  บนวัตถุที่มองจะอยู่จะมีตำแหน่งใน Projection plane เท่ากับ  $(x_{pr}, y_{pr}, 0)$  และในกรณีของภาพสำหรับตาข้างขวาสามารถคำนวณหาค่า  $x_{pr}$  และ  $y_{pr}$  ได้จากสมการดังต่อไปนี้

$$x_{pr} = \frac{d \left[ x + \frac{s_e}{2} \right]}{z+d} \quad (3.18)$$

ดังนั้น

$$y_{pr} = \frac{dy}{z+d} \quad (3.19)$$



รูปที่ 3.16 ตัวอย่างการสร้างภาพสเตอริโอ

จากรูปที่ 3.16 เป็นตัวอย่างของการคำนวณตำแหน่งของวัตถุบน Projection plane สำหรับการสร้างภาพสเตอริโอ โดยกำหนดให้มีระยะห่างระหว่างดวงตาข้างซ้ายและขวาเท่ากับ 6.5 เซนติเมตรและมองไปที่ลูกบาศก์ขนาด 10 เซนติเมตรที่ห่างจากจุดมอง 60 เซนติเมตร มีระยะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

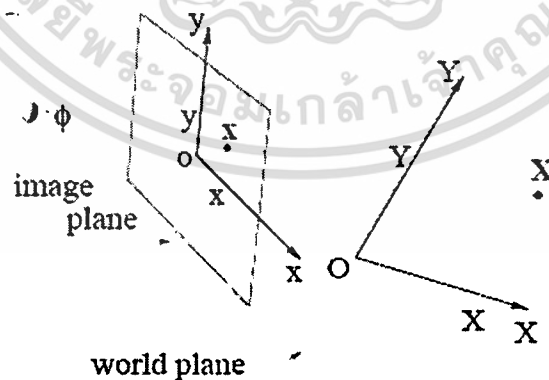
ระหว่างดวงตากับ Projection plane เท่ากับ 20 เซนติเมตร ดังนั้นสามารถคำนวณค่าของ  $x_{pl}$ ,  $y_{pl}$ ,  $x_{pr}$  และ  $y_{pr}$  ได้ดังตารางที่ 3.1

ตารางที่ 3.1 ผลลัพธ์ที่ได้จากการคำนวณตำแหน่งในการสร้างภาพตาซ้ายและตาขวา

$x$	$y$	$z$	$x_{pl}$	$y_{pl}$	$x_{pr}$	$y_{pr}$
5	5	40	0.58	1.67	2.75	1.67
5	-5	40	0.58	-1.67	2.75	-1.67
-5	5	40	-2.75	1.67	-0.58	1.67
-5	-5	40	-2.75	-1.67	-0.58	-1.67
5	5	50	0.5	1.43	2.36	1.43
5	-5	50	0.5	-1.43	2.36	-1.43
-5	5	50	-2.36	1.43	-0.5	1.43
-5	-5	50	-2.36	-1.43	-0.5	-1.43

### 3.5 Homography [8]

การฉายภาพจาก World plan ไปยัง Image plan จะใช้หลักการของ Plane projective transformation วิธีการนี้จะใช้วิธีการคำนวณหา Transformation matrix ซึ่งมีขนาด 3x3 โดยจะใช้สัญลักษณ์แทน Matrix นี้คือ  $H$  ดังนั้น Transformation matrix  $H$  นี้จะเรียกว่า Homography หรือ Projective transformation



รูปที่ 3.17 แบบจำลองการสร้างภาพด้วยกล้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.17 จุด  $X$  ใน World plane สามารถ Transform ไปเป็นจุด  $x$  ใน image plane โดยสามารถคำนวณได้จากสมการที่ 3.20

$$X = Hx \quad (3.20)$$

เมื่อ  $X$  คือ  $(X \ Y \ Z)^T$  เป็นตำแหน่งพิกัดในรูปแบบ Homogenous Vector บน World plane

$x$  คือ  $(x \ y \ z)^T$  เป็นตำแหน่งพิกัดในรูปแบบ Homogenous Vector บน Image plane

$H$  คือ Transformation matrix ขนาด 3x3

ดังนั้นจากสมการที่ 3.20 สามารถเขียนสมการใหม่ได้ดังนี้

$$\begin{pmatrix} XW \\ YW \\ W \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.21)$$

ทำการเปลี่ยนผลลัพธ์ที่ได้ให้อยู่ในพิกัด Homogenous จากสมการที่ 3.21 จะได้ค่า  $W = gx + hy + i$  ดังนั้นจึงสามารถเขียนสมการได้ใหม่ดังนี้

$$\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \frac{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}}{\begin{pmatrix} g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}} \quad (3.22)$$

จากสมการที่ 3.22 สามารถหาค่าของ  $X$  และ  $Y$  ได้จากสมการต่อไปนี้

$$X = \frac{ax + by + c}{gx + hy + 1} \quad (3.23)$$

$$Y = \frac{dx + ey + f}{gx + hy + 1} \quad (3.24)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น  $X$  และ  $Y$  จะมีค่าเท่ากับ

$$X = ax + by + c - gXx - hYy \quad (3.25)$$

$$Y = dx + ey + f - gYx - hYy \quad (3.26)$$

ทำการเพิ่มนิพจน์ที่เป็น 0 เข้าไปในสมการทั้ง 2 จะได้

$$X = ax + by + c + 0d + 0e + 0f - gXx - hYy \quad (3.27)$$

$$Y = 0a + 0b + 0c + dx + ey + f - gYx - hYy \quad (3.28)$$

ดังนั้นสามารถเขียนสมการใหม่ในรูปของ Matrix

$$\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & -Xx & -Yy \\ 0 & 0 & 0 & x & y & 1 & -Yx & -Yy \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix} = \begin{pmatrix} X \\ Y \end{pmatrix} \quad (3.29)$$

จากสมการที่ 3.29 จำนวนค่าของ  $(a \ b \ c \ d \ e \ f \ g \ h)^T$  จำเป็นจะต้องใช้สมการอย่างน้อย 8 สมการ หรือจุดจำนวน 4 จุดที่สอดคล้องกันระหว่าง World plane กับ Image plane ดังสมการ

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -X_3x_3 & -X_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -Y_3x_3 & -Y_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -X_4x_4 & -X_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -Y_4x_4 & -Y_4y_4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix} = \begin{pmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \\ X_4 \\ Y_4 \end{pmatrix} \quad (3.30)$$

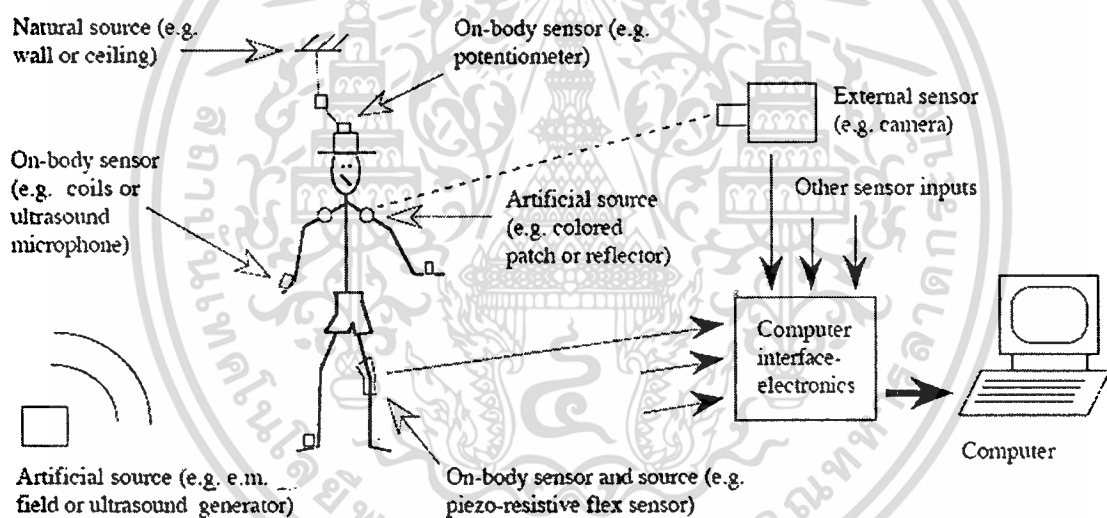
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการที่ 3.30 สามารถนำไปใช้คำนวณค่าของ Transformation matrix ได้

จากหลักการของ Homography หรือ Projective transformation สามารถนำไปประยุกต์ใช้ในการคำนวณหาตำแหน่งใน World plane จากภาพเพียงภาพเดียวได้ โดยถ้าต้องการให้มีความถูกต้องมากขึ้นจำเป็นจะต้องใช้จุดอ้างอิงระหว่าง World plane กับ Image plane มากกว่า 4 จุดขึ้นไป เพื่อนำมาคำนวณหา Transformation matrix

### 3.6 การตรวจวัดการเคลื่อนไหวของมนุษย์

ระบบการตรวจวัดการเคลื่อนไหวของมนุษย์ เป็นระบบที่สามารถตรวจวัดตำแหน่งการเคลื่อนไหวของมนุษย์ได้แบบเวลาจริง ในการตรวจวัดอาจจะมีการใช้อุปกรณ์ติดตั้งลงบนตัวของมนุษย์ เช่น อุปกรณ์ตัววัดด้วยสนามแม่เหล็กไฟฟ้า หรือการตรวจวัดด้วยกล้องวิดีโอ ซึ่งอาจจะเป็นการตรวจวัดทั้งหมดของร่างกาย หรือ ใช้เฉพาะบางส่วน ดังรูปที่ 3.18



รูปที่ 3.18 ตัวอย่างของวิธีการสำหรับการตรวจวัดการเคลื่อนไหวของมนุษย์ [3]

#### 3.6.1 ชนิดของการตรวจวัดตำแหน่ง [3]

การตรวจวัดการเคลื่อนไหวของมนุษย์นั้นมียู่ด้วยกันหลากหลายวิธีการ ดังรูปที่ 3.18 ในปัจจุบันสามารถแบ่งวิธีการที่ใช้ในการตรวจวัดตำแหน่งได้ 3 วิธีการด้วยกันคือ

3.6.1.1 Non-vision based วิธีการนี้จะใช้เซ็นเซอร์ติดไว้กับตัวมนุษย์เพื่อใช้สำหรับรวบรวมข้อมูลตำแหน่งการเคลื่อนไหวหรือการเปลี่ยนตำแหน่งของร่างกายมนุษย์ โดยสามารถจำแนกเทคโนโลยีที่ใช้ในการตรวจวัดคือ เครื่องกล, การตรวจจับความเฉื่อย, การตรวจจับคลื่นเสียง, คลื่นวิทยุหรือไมโครเวฟ และการตรวจวัดสนามแม่เหล็ก สำหรับเทคโนโลยีต่างๆ ที่นำมาใช้ในการ

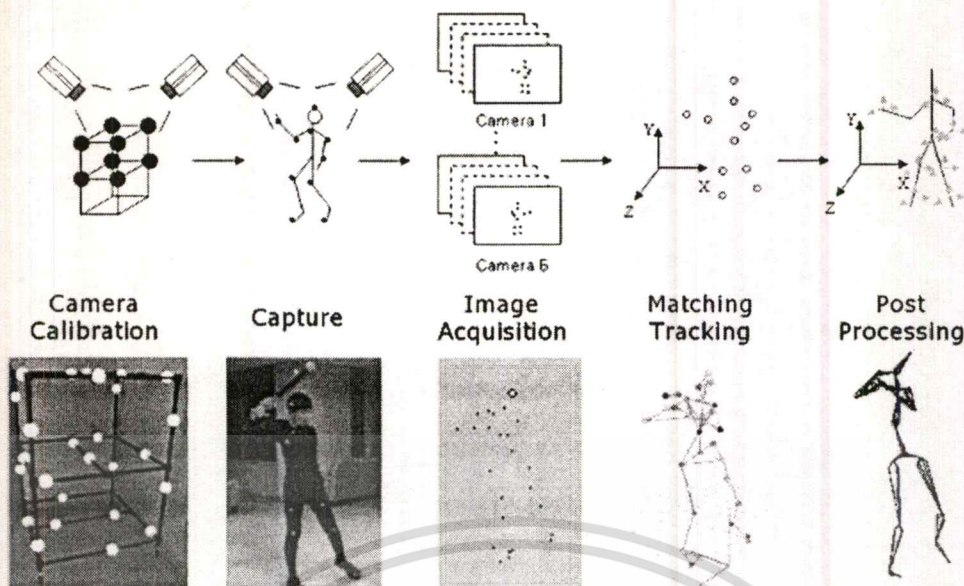
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตรวจวัดนี้มีทั้งข้อดีและข้อจำกัดแตกต่างกันไป ซึ่งข้อจำกัดนี้จะประกอบไปด้วย ข้อกำหนดของรูปแบบการใช้งาน ข้อกำหนดของการตรวจวัด และข้อกำหนดเกี่ยวกับสภาพแวดล้อม จึงทำให้การนำเทคโนโลยีนี้ไปใช้งานได้ ในสภาวะแวดล้อมที่แตกต่างกันไป และอุปกรณ์หรือเครื่องมือสำหรับวิธีการนี้จะมีราคาแพง ซึ่งไม่เหมาะกับการนำมาใช้ในเครื่องคอมพิวเตอร์ส่วนบุคคลทั่วไป ตัวอย่างของเครื่องมือตรวจวัดตำแหน่งแบบใช้สนามแม่เหล็กชื่อ FASTRAK ของบริษัท Polhemus ดังรูปที่ 3.19



รูปที่ 3.19 เครื่องมือตรวจวัดตำแหน่งชื่อ FASTRAK ของบริษัท Polhemus [24]

3.6.1.2 Vision based with markers วิธีการนี้จะใช้อุปกรณ์ทางแสงหรือกล้องวิดีโอ ร่วมกับการทำเครื่องหมายบนร่างกาย  $\otimes$  สำหรับเป็นจุดอ้างอิงในการตรวจจับการเคลื่อนไหวของร่างกายมนุษย์ วิธีการนี้สามารถนำไปใช้ในการตรวจวัดการเคลื่อนไหวของร่างกายมนุษย์ซึ่งมีโครงสร้างที่มีความซับซ้อนสูงได้เป็นอย่างดี และวิธีการตรวจวัดตำแหน่งแบบนี้จะนำไปใช้ในงานวิจัยทางด้านการแพทย์, วิทยาศาสตร์การกีฬา, ด้านวิศวกรรม และการสร้างภาพอนิเมชันที่ต้องการการเคลื่อนไหวที่มีความสมบูรณ์แบบ ตัวอย่างดังรูปที่ 3.20



รูปที่ 3.20 วิธีการตรวจวัดตำแหน่งแบบ Vision based with markers [25]

3.6.1.3 Vision based without markers วิธีการนี้จะใช้อุปกรณ์ทางแสงหรือกล้องวิดีโอเป็นเครื่องมือในการตรวจวัดตำแหน่ง โดยจะจับร่างกายของมนุษย์หรือส่วนใดส่วนหนึ่งที่ต้องการซึ่งจะขึ้นอยู่กับวิธีของผู้ที่นำวิธีการนี้ไปใช้ว่าจะตรวจวัดจากส่วนใดของร่างกาย เช่น หัว หน้า ดวงตา แขน มือ หรือขา เป็นต้น สำหรับวิธีการนี้ถ้าเป็นการเจาะจงไปยังส่วนใดส่วนหนึ่งของร่างกาย เช่น ดวงตา จะต้องใช้กล้องที่มีความละเอียดสูงและระยะเวลาใช้งานจะไม่สูงมากเนื่องจากข้อจำกัดดังกล่าว อีกทั้งประสิทธิภาพที่ได้ยังขึ้นอยู่กับเครื่องคอมพิวเตอร์ที่นำมาประมวลผลด้วย ตัวอย่างดังรูปที่ 3.21



รูปที่ 3.21 วิธีการตรวจวัดตำแหน่งแบบ Vision based without markers

### 3.6.2 เปรียบเทียบข้อดีและข้อเสียของวิธีการตรวจวัดตำแหน่ง

การตรวจวัดตำแหน่งโดยวิธีการที่กล่าวมามีทั้งข้อดีและข้อเสียแตกต่างกัน โดยสามารถเปรียบเทียบได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6.2.1 ราคา สำหรับวิธีการตรวจวัดตำแหน่งด้วย Non-vision based tracking จะใช้อุปกรณ์ที่มีราคาแพงกว่า วิธีการตรวจวัดตำแหน่งแบบ Vision based แต่จะได้คุณภาพของข้อมูลตำแหน่งที่สูงกว่า แต่เนื่องจากราคาที่แพงของการตรวจวัดตำแหน่งแบบ Non-vision based จึงไม่เหมาะสมกับการนำมาใช้งานในการพัฒนาแอปพลิเคชันราคาต่ำ

3.6.2.2 ความสะดวกสบาย สำหรับวิธีการตรวจวัดตำแหน่งด้วย Non-vision based tracking จะใช้อุปกรณ์มาติดตั้งไว้ที่ร่างกายและมีสายเชื่อมต่อกับอุปกรณ์ที่บนร่างกายไปยังเครื่องรับทำให้การเคลื่อนไหวของร่างกายและพื้นที่ใช้งานถูกจำกัด สำหรับวิธีการแบบ Vision based ร่วมกับมาร์คเกอร์ จำเป็นต้องมีมาร์คเกอร์ติดตั้งไว้กับร่างกายซึ่งทำให้มีขั้นตอนในการใช้งานที่ยุงยาก วิธีการแบบ Vision based แบบไม่ใช้มาร์คเกอร์ จะมีความสะดวกสบายในการติดตั้งการใช้งาน และไม่ถูกจำกัดพื้นที่ใช้งาน

3.6.2.3 สัญญาณรบกวน (Noise) เมื่อเปรียบเทียบระหว่างวิธีที่ไม่ใช้กล้องวิดีโอ กับวิธีที่ใช้กล้องวิดีโอ สัญญาณรบกวนของวิธีที่ไม่ใช้กล้องวิดีโอ เช่น วิธีที่ใช้สนามแม่เหล็กไฟฟ้าอาจจะเกิดสัญญาณรบกวนจากคลื่นวิทยุและอุปกรณ์ที่สามารถส่งสัญญาณสนามแม่เหล็กไฟฟ้าได้ สำหรับวิธีที่ใช้กล้องวิดีโอจะมีสัญญาณรบกวนที่เกิดจากบิจัยหลายอย่าง เช่น แสงสว่าง, ความแตกต่างของสี เป็นต้น ถ้ามีการติดตั้งระบบที่ดีก็จะสามารถทำให้สัญญาณรบกวนน้อยลง

3.6.2.4 ความละเอียด (Accuracy) ความละเอียดของข้อมูลสำหรับการตรวจวัดตำแหน่งด้วยกล้องวิดีโอแบบไม่ใช้มาร์คเกอร์อาจจะให้ความละเอียดของข้อมูลต่ำ เมื่อเทียบกับการตรวจวัดตำแหน่งแบบไม่ใช้กล้องวิดีโอ แต่ก็มีความเหมาะสมเพียงพอกับแอปพลิเคชันราคาต่ำ เช่น เกมส์ ที่ไม่ต้องการความละเอียดและความถูกต้องของข้อมูลสูงมาก

## บทที่ 4

# อัลกอริทึมในการตรวจวัดตำแหน่งขาด้วยกล้องวิดีโอ

### 4.1 แนวคิดเบื้องต้น

การแสดงภาพสเตอริโอเชิงโต้ตอบ เป็นวิธีการแสดงภาพสเตอริโอที่สามารถสร้างภาพที่มีความสัมพันธ์กับการเปลี่ยนตำแหน่งของผู้ใช้ได้ วิธีการตรวจวัดตำแหน่งของผู้ใช้ด้วยกล้องวิดีโอแบบไม่ใช้มาร์คเกอร์เป็นอีกวิธีหนึ่งที่น่าสนใจในการตรวจวัดตำแหน่งของผู้ใช้ ซึ่งการตรวจวัดตำแหน่งนี้จะเป็นการเจาะจงไปที่ส่วนใดส่วนหนึ่งของร่างกายมนุษย์ เช่น ศีรษะ ลำตัว ขา เท้า เป็นต้น โดยที่วิธีการดังกล่าวไม่ต้องอาศัยมาร์คเกอร์ ทำให้สะดวกในการติดตั้งและต่อผู้ใช้งาน

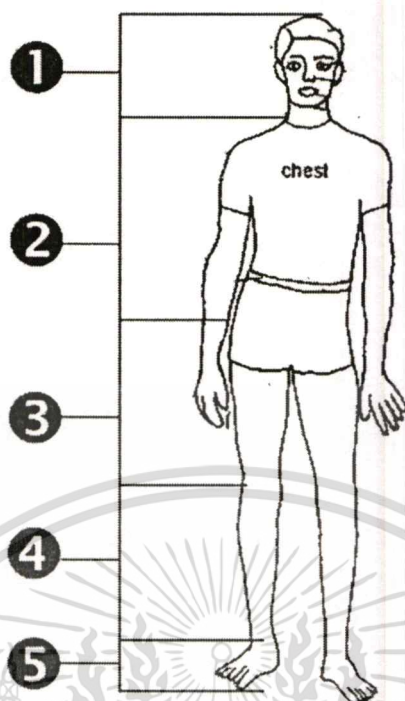
ในระบบการแสดงภาพสเตอริโอเชิงโต้ตอบ อุปกรณ์ที่ใช้แสดงผลภาพจะมีการติดตั้งอยู่กับที่ โดยผู้ใช้เท่านั้นที่จะมีการเปลี่ยนตำแหน่งในการมอง การก้ม/เงยหน้า ไม่มีผลกับการแสดงภาพสเตอริโอเชิงโต้ตอบ เพราะฉะนั้นข้อมูลที่มีความจำเป็นสำหรับระบบการแสดงผลนี้คือค่าตำแหน่งของผู้ใช้ในระบบพิกัด  $X$  และ  $Y$  เท่านั้น

ในการแสดงภาพสเตอริโอด้วยเครื่องคอมพิวเตอร์ส่วนบุคคล ซึ่งเป็นลักษณะการใช้งานที่เป็นเป้าหมายหลักของงานวิจัยนี้ ราคาและความซับซ้อนของระบบเป็นปัจจัยที่มีความสำคัญ ดังนั้นในงานวิจัยนี้ การคำนวณหาตำแหน่งการยืนของผู้ใช้ จะใช้วิธีการที่ง่าย รวดเร็วและอุปกรณ์มีราคาต่ำ โดยการคำนวณตำแหน่งดังกล่าว อาศัยกล้องราคาถูก 1 ตัว ทำการจับภาพขาและเท้าของผู้ใช้ ข้อมูลดังกล่าว จะถูกนำไปใช้หาตำแหน่งที่ยืนในระนาบของพื้น ต่อไป โดยการจับภาพร่างกายท่อนล่างของผู้ใช้ มีข้อดีหลายประการ ได้แก่

- สามารถควบคุมภาพพื้นหลังได้ง่าย เนื่องจากพื้นที่ที่ผู้ใช้งานยืนอยู่จะไม่มีการเปลี่ยนแปลงของพื้น เช่น การเดินผ่านของบุคคลอื่น เป็นต้น
- ง่ายต่อการคำนวณค่าตำแหน่งด้วย Homography
- ผู้ใช้งานรู้สึกมีความเป็นส่วนตัวมากกว่า การตรวจวัดตำแหน่งที่ใช้กล้องจับภาพไปที่หน้าหรือทั้งร่างกายของผู้ใช้งาน

### 4.2 การแบ่งส่วนของร่างกายมนุษย์

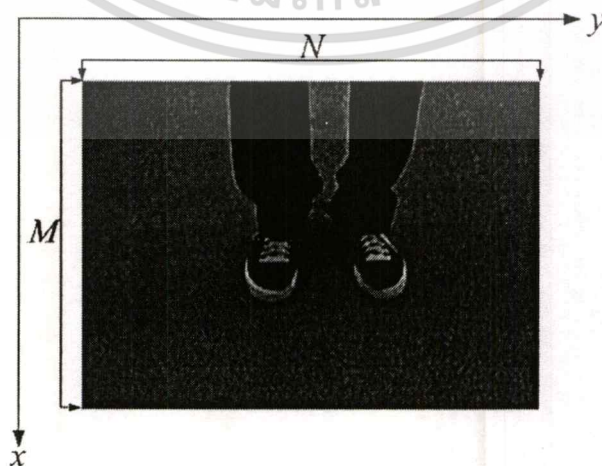
ในงานวิจัยนี้ทำการแบ่งส่วนของร่างกายมนุษย์ออกเป็น 5 ส่วนด้วยกันคือ ศีรษะ ลำตัว ขา และ เท้า ดังรูปที่ 4.1 โดยจะนำส่วนที่ 4 และ 5 มาใช้ในการตรวจวัดตำแหน่งของผู้ใช้ที่ยืนอยู่กับพื้นราบ



รูปที่ 4.1 การแบ่งร่างกายมนุษย์ออกเป็น 5 ส่วน[23]

#### 4.3 ขนาดและระยะพิกัดของภาพวิดีโอ

ในงานวิจัยนี้จะใช้อุปกรณ์กล้องวิดีโอที่มีราคาต่ำ และมีการเชื่อมต่อกับเครื่องคอมพิวเตอร์ที่ง่าย เช่น เชื่อมต่อผ่านพอร์ต USB เป็นต้น ขนาดของภาพวิดีโอที่ใช้ 320x240 จุดภาพ ตามคุณสมบัติของกล้องวิดีโอ โดยกำหนดให้  $N$  คือขนาดของภาพตามแนวแกน  $y$  และ  $M$  คือขนาดของภาพตามแนวแกน  $x$  และมีการกำหนดระยะพิกัดของภาพเพื่อนำมาใช้ในกระบวนการประมวลผลตรวจวัดตำแหน่ง ดังรูปที่ 4.2

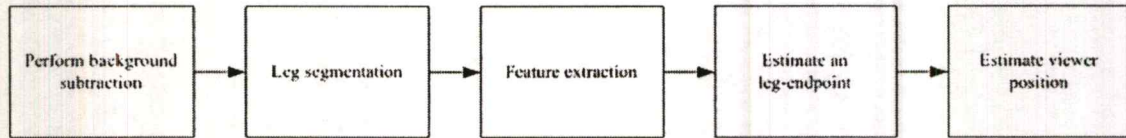


รูปที่ 4.2 ระยะพิกัดของภาพที่ใช้ในงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.4 อัลกอริทึมในการตรวจวัดตำแหน่งขา

อัลกอริทึมที่ใช้สำหรับตรวจวัดตำแหน่งขานั้น มีขั้นตอนการทำงานดังรูปที่ 4.3



รูปที่ 4.3 โครงสร้างและขั้นตอนของอัลกอริทึมในการตรวจวัดตำแหน่งขา

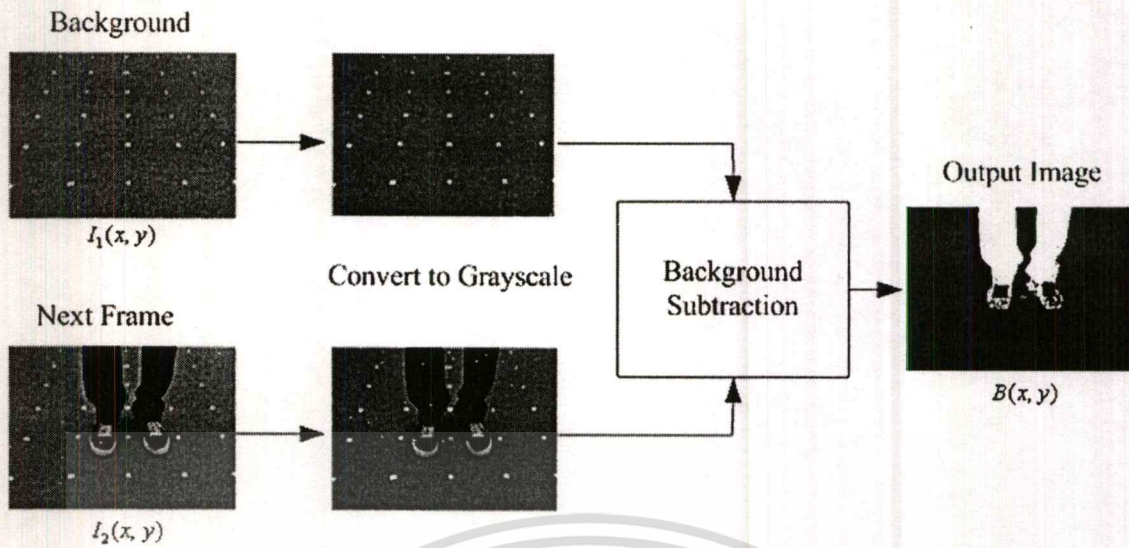
จากขั้นตอนการทำงานของอัลกอริทึมในรูปที่ 4.3 สามารถอธิบายการทำงานได้ดังนี้

##### 4.4.1 กระบวนการลบภาพพื้นหลัง (Perform background subtraction)

ภาพที่ได้จากกล้องวีดีโอเว็บแคมจะประกอบด้วยพื้นหลังและขาของผู้ใช้ ดังนั้นจำเป็นจะต้องลบภาพพื้นหลังออกเพื่อให้เหลือแต่ภาพที่มีเฉพาะขาของผู้ใช้เท่านั้น ในขั้นตอนนี้จึงใช้วิธีการลบภาพพื้นหลัง (Background subtraction) เพื่อลบภาพพื้นหลังที่ไม่ต้องการออกไป วิธีการนี้จะมีการเก็บภาพพื้นหลังไว้ แล้วนำพื้นหลังและภาพที่ต้องการลบภาพพื้นหลังมาเปลี่ยนเป็นภาพแบบ Grayscale จากนั้นนำภาพทั้งสองไปผ่านกระบวนการ Background subtraction ดังรูปที่ 4.4 และสามารถคำนวณได้จากสมการที่ 4.1

$$B(x, y) = f_{bw} |I_1(x, y) - I_2(x, y)| \quad (4.1)$$

เมื่อ  $I_1(x, y)$  คือ ภาพพื้นหลัง  
 $I_2(x, y)$  คือ ภาพที่ต้องการลบภาพพื้นหลัง  
 $B(x, y)$  คือ ภาพที่ได้จากวิธีการ Background subtraction



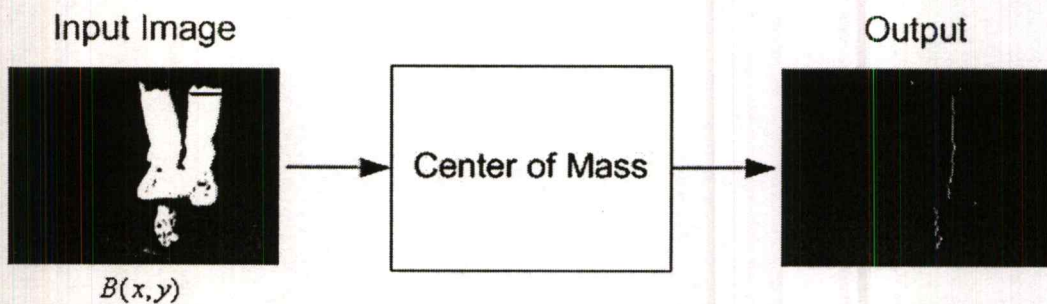
รูปที่ 4.4 กระบวนการ Background subtraction

4.4.2 กระบวนการแยกภาพขา (Leg segmentation)

เมื่อได้ภาพจากกระบวนการลบภาพพื้นหลัง (Background subtraction) ขึ้นต่อไปเป็นการค้นหาส่วนที่เป็นขาของผู้ใช้ โดยปกติเมื่อนุษย์ยืนตรงขาจะมีลักษณะเป็นเส้นตรง และเมื่อนำภาพขาของผู้ใช้มาคำนวณหาจุดกึ่งกลางของภาพในแต่ละแถว แล้วนำไปวาดเป็นกราฟก็จะทำให้ได้ตำแหน่งที่ผู้ใช้นอนอยู่ในภาพ ดังนั้นจากหลักการดังกล่าวการคำนวณหาจุดกึ่งกลางของภาพตามแนวนอนจะใช้วิธีการ Center of mass คำนวณได้จากสมการที่ 4.2 และกระบวนการแสดงดังรูปที่ 4.5

$$r_x = \frac{\sum_{y=1}^N B(x,y) \cdot y}{\sum_{y=1}^N B(x,y)} \tag{4.2}$$

เมื่อ  $r_x$  คือ จุดกึ่งกลางของภาพตามแนวนอนที่คำนวณได้



รูปที่ 4.5 การคำนวณหา Center of Mass

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.4.3 กระบวนการแยกโครงสร้างขา (Leg skeleton extraction)

เนื่องจากภาพที่ได้จากการลบภาพพื้นหลัง อาจจะมีบางส่วนในภาพที่ไม่ต้องการ คิ้ว เช่น เงา เป็นต้น ดังนั้นกระบวนการแยกโครงสร้างขาจะทำการตัดสิ่งที่ไม่ต้องการในภาพออก ให้เหลือเฉพาะโครงสร้างของขาเท่านั้น โดยแบ่งออกเป็น 2 ขั้นตอนดังนี้

4.4.3.1 การเลือกจุดที่ต้องการจากขั้นตอนกระบวนการแยกภาพขา (Leg segmentation) โดยจุดที่เลือกจะใช้ค่า Threshold ที่คำนวณได้จากสมการที่ 4.3 และ สมการที่ 4.4

$$S_x = \sum_{y=1}^N B(x, y) \quad (4.3)$$

เมื่อ  $S_x$  คือ ผลรวมของค่า Pixel ทั้งหมดตามแนวนอน ณ แถวที่  $x$

$$Th_1 = \frac{\sum_{x=1}^M S_x}{M} \quad (4.4)$$

เมื่อ  $Th_1$  คือ ค่า Threshold ที่นำมาใช้สำหรับเลือกจุดที่เหมาะสม

ค่า Threshold ที่คำนวณได้จากสมการที่ 4.4 จะใช้เป็นตัวนำเข้าไปเลือกจุดที่ต้องการจากกระบวนการแยกภาพขา (Leg segmentation) โดยการเลือกจุด  $r_x$  ณ แถวที่  $x$  ที่มีผลรวมของจำนวนจุดภาพตามแนวนอนมากกว่าค่า  $Th_1$  ดังสมการที่ 4.5

$$R_x = \{x | S_x > Th_1\} \quad (4.5)$$

เมื่อ  $R_x$  คือ เซตของแถวที่ถูกเลือก ซึ่งมีค่า  $S_x > Th_1$

4.4.3.2 จากจุด  $r_x, x \in R_x$  ซึ่งหาได้จากในขั้นตอนที่แล้ว นำจุดที่ได้มาคำนวณหาตัวแปร  $m$  และ  $b$  ของสมการเส้นตรง  $y = mx + b$  เพื่อใช้เป็นเส้นอ้างอิงที่จะใช้คำนวณหาตำแหน่งของจุดที่ผู้ใช้ยื่นในขั้นตอนต่อไป การคำนวณจะใช้วิธีการ Least squares approximation ดังสมการที่ 4.6 และ 4.7

$$m = \frac{P\left(\sum_{x \in R_x} x r_x\right) - \left(\sum_{x \in R_x} x\right)\left(\sum_{x \in R_x} r_x\right)}{P\left(\sum_{x \in R_x} x^2\right) - \left(\sum_{x \in R_x} x\right)^2} \quad (4.6)$$

$$b = \frac{\left(\sum_{x \in R_x} r_x\right)\left(\sum_{x \in R_x} x^2\right) - \left(\sum_{x \in R_x} x\right)\left(\sum_{x \in R_x} x r_x\right)}{P\left(\sum_{x \in R_x} x^2\right) - \left(\sum_{x \in R_x} x\right)^2} \quad (4.7)$$

- เมื่อ  $m$  คือ ค่าความชันของสมการเส้นตรง  
 $b$  คือ จุดที่เส้นตรงตัดกับแกน  $y$   
 $P$  คือ จำนวนของจุดที่นำมาใช้ในการคำนวณ (จำนวนสมาชิกของ  $R_x$ )

#### 4.4.4 กระบวนการคำนวณหาตำแหน่งสันเท้า (Leg-endpoint Estimation)

เมื่อได้ลักษณะ โครงสร้างของขาแล้ว ขั้นตอนต่อไปคือทำการคำนวณหาจุดที่เป็นตำแหน่งของสันเท้า สามารถแทนด้วยตัวแปร  $C_{xy}$  โดยจะนำจุดที่เป็นโครงสร้างของขามาเปรียบเทียบกับค่าจากสมการเส้นตรงที่คำนวณได้ ว่ามีระยะห่างจากสมการเส้นตรงเท่าใด ดังสมการที่ 4.8

$$e_{x'} = |r_{x'} - (mx' + b)|, x' \in R_x \quad (4.8)$$

- เมื่อ  $e_{x'}$  คือ ค่าระยะห่างของจุด โครงสร้างขาจากสมการเส้นตรง

จากค่าระยะห่าง  $e_{x'}$  ที่ได้จากสมการที่ 4.8 นำมาหาค่าเฉลี่ยของค่าระยะห่างดังสมการต่อไปนี้

$$A = \frac{\sum_{x'}^P e_{x'}}{P} \quad (4.9)$$

- เมื่อ  $A$  คือ ค่าเฉลี่ยของค่าระยะห่างของจุด โครงสร้างขาจากสมการเส้นตรง

จากนั้นทำการเลือกจุดที่เป็นโครงสร้างของขาจาก  $r_x$  โดยเลือกเฉพาะจุดที่มีค่าระยะห่างของจุดน้อยกว่าค่าเฉลี่ยที่คำนวณได้ โดยพิจารณาจากจุดที่เลือกที่มีค่า  $x$  มากที่สุดไปหาค่าน้อยที่สุด ถ้าจุดใดมีค่าระยะห่างจากสมการเส้นตรงน้อยกว่าค่าของ  $A$  เป็นตำแหน่งแรก ให้เลือกจุดนั้นเป็นตำแหน่งสั้นเท้า  $C_{xy}$  แล้วนำจุดที่เลือกมาคำนวณหาตำแหน่งของผู้ใช้ใน World plane ต่อไป

#### 4.4.5 กระบวนการคำนวณตำแหน่งของผู้ใช้ (Estimate viewer position)

เมื่อได้ค่าตำแหน่งสั้นเท้า  $C_{xy}$  จากขั้นตอนที่ผ่านมา จะนำค่าตำแหน่งที่คำนวณได้นี้ไปแทนในสมการที่ 3.23 และสมการที่ 3.24 ก็จะทำได้ค่าตำแหน่งของผู้ใช้ใน World plane ค่าตำแหน่งที่คำนวณได้แทนด้วยตัวแปร  $W_{xy}$  ระบบแสดงภาพสเตอริโอเชิงโต้ตอบจะนำค่าตำแหน่งนี้ไปสร้างภาพที่มีความสัมพันธ์กับตำแหน่งของผู้ใช้

#### 4.5 อัลกอริทึมสำหรับแก้ไขข้อผิดพลาด

ค่าตำแหน่งใน World plane ของผู้ใช้ที่คำนวณได้จากอัลกอริทึมในการตรวจวัดตำแหน่งขา อาจจะมีข้อผิดพลาดเกิดขึ้นได้ เช่น เกิดจากการเปลี่ยนแปลงของแสง หรือ เงาที่พื้น เป็นต้น ดังนั้นเพื่อให้ค่าตำแหน่งที่คำนวณได้มีความเหมาะสมกับการนำไปใช้งานกับระบบการแสดงผลภาพสเตอริโอเชิงโต้ตอบ จึงต้องมีอัลกอริทึมสำหรับแก้ไขข้อผิดพลาดของตำแหน่งที่คำนวณได้

- อัลกอริทึมปรับค่าตำแหน่งให้เหมาะสม (Smoothing algorithm)

อัลกอริทึมนี้ใช้สำหรับปรับค่าตำแหน่งของผู้ใช้ที่คำนวณได้ ให้มีความเหมาะสมที่จะนำไปใช้กับระบบแสดงภาพสเตอริโอเชิงโต้ตอบ โดยจะนำค่าตำแหน่ง  $W_{xy}$  มาคำนวณดังสมการที่ 4.10

$$W_{i,xy}(t) = \frac{\sum_{m=1}^{N-1} W_{xy}(t-m)}{N} \quad (4.10)$$

เมื่อ	$W_{i,xy}(t)$	คือ ค่าตำแหน่ง $W_{xy}$ ตัวที่ $i$ ในเวลา $t$ ใดๆ
	$N$	คือ จำนวนของข้อมูลตำแหน่งที่นำมาใช้ในการคำนวณ
	$t$	คือ ค่าดัชนีเวลาของค่าตำแหน่ง $W_{xy}$

## บทที่ 5

### ผลการทดลอง

#### 5.1 อุปกรณ์และสภาพแวดล้อมในการทดลอง

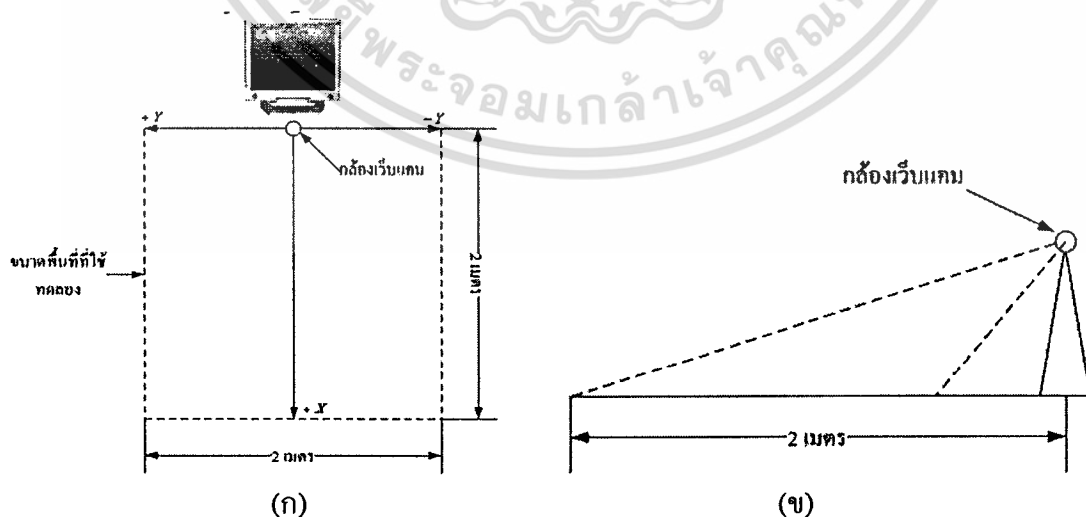
##### 5.1.1 อุปกรณ์ที่ใช้ในการทดลอง

อุปกรณ์ที่ใช้ในการเก็บข้อมูลประกอบด้วย

- เครื่องคอมพิวเตอร์ PC Intel Pentium III 900 MHz
- กล้องวีดีโอเว็บแคม Logitech รุ่น QuickCam Pro 4000, ความละเอียดของภาพวีดีโอสูงสุด 640x480 จุดภาพ, รองรับ USB 2.0 และ 1.1
- พัฒนาโปรแกรมตรวจวัดตำแหน่งจากโปรแกรม Microsoft® Visual C++ 6.0 และใช้ TLIB:Tracking Library [11] เพื่อประมวลผลภาพที่ได้จากกล้องวีดีโอเว็บแคม

##### 5.1.2 การจัดวางอุปกรณ์เพื่อการทดลอง

พื้นที่และการจัดวางอุปกรณ์ที่ใช้ในการทดลอง แสดงดังรูปที่ 5.1 จากรูป 5.1 (ก) ระบบจะประกอบด้วยเครื่องคอมพิวเตอร์สำหรับประมวลผลตำแหน่ง กล้องวีดีโอเว็บแคมตั้งอยู่ด้านหน้า และพื้นที่ที่ใช้ในการทดลองมีขนาด 2x2 เมตร กำหนดระบบพิกัดในแนวนอนเป็นแกน  $Y$  และในแนวตั้งเมื่อเทียบกับเครื่องคอมพิวเตอร์เป็นแกน  $X$  กล้องวีดีโอเว็บแคมตั้งอยู่ที่ตำแหน่ง  $X=0, Y=0$  โดยกล้องวีดีโอจะปรับตำแหน่งรับภาพไปที่พื้นเพื่อให้กล้องจับภาพพื้นภายในขอบเขตที่กำหนดดังรูปที่ 5.1 (ข)

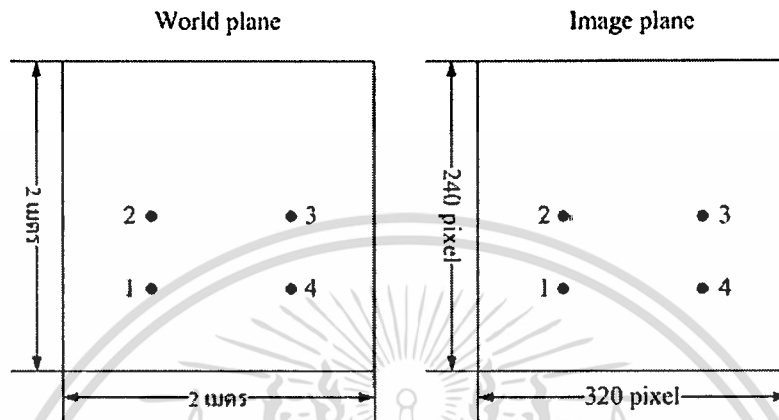


รูปที่ 5.1 ระบบที่ใช้ในการทดลอง (ก)ภาพด้านบน และ (ข)ภาพด้านข้าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.1.3 การคำนวณค่าของ Homography matrix

จากหลักการของ Homography ที่กล่าวไว้ในบทที่ 3 ก่อนการทดลองจำเป็นจะต้องทำการคำนวณค่าของ Homography matrix โดยจะใช้จุดอ้างอิงทั้งหมด 4 จุดที่สอดคล้องกันระหว่าง World plane กับ Image plane ดังรูปที่ 5.2



รูปที่ 5.2 จุดอ้างอิงที่สอดคล้องกัน 4 จุดระหว่าง World plane กับ Image plane

จากรูปที่ 5. 2 จุดอ้างอิงบน World plane กำหนดให้อยู่ ณ ตำแหน่งต่าง ๆ ดังนี้

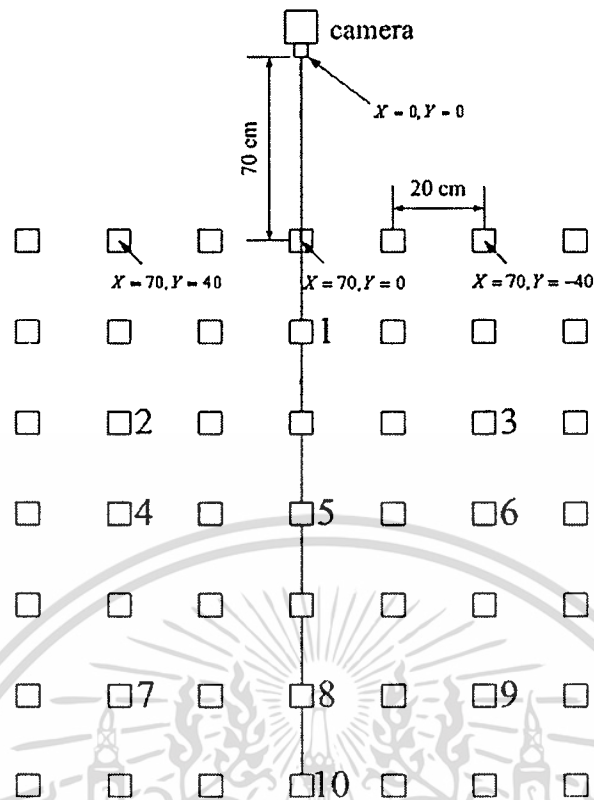
- ตำแหน่งที่ 1  $X_1 = 90$  ซม. ,  $Y_1 = -40$  ซม.
- ตำแหน่งที่ 2  $X_2 = 130$  ซม. ,  $Y_2 = -40$  ซม.
- ตำแหน่งที่ 3  $X_3 = 130$  ซม. ,  $Y_3 = 40$  ซม.
- ตำแหน่งที่ 4  $X_4 = 90$  ซม. ,  $Y_4 = 40$  ซม.

สำหรับจุดอ้างอิงบน Image plane ค่าแต่ละตำแหน่งอาจจะมีการเปลี่ยนแปลงได้ เนื่องมาจากการปรับระดับความสูงของกล้องและการปรับตำแหน่งการรับภาพของกล้องที่จับภาพไปที่พื้น เมื่อได้ค่าตำแหน่งของจุดอ้างอิงทั้ง 4 จุดใน World plane และ Image plane นำค่าตำแหน่งที่ได้แทนลงในสมการที่ 3.30 แล้วคำนวณหาค่าของ Homography matrix จากสมการ ค่าของ Homography matrix ที่ได้จะนำไปใช้ในการคำนวณหาตำแหน่งของผู้ใช้ใน World plane หลังจากที่สามารถคำนวณหาตำแหน่งของผู้ใช้จาก Image plane ได้แล้ว

### 5.1.4 การกำหนดตำแหน่งอ้างอิงบนพื้น

การกำหนดตำแหน่งอ้างอิงบนพื้น ทำขึ้นเพื่อใช้ในการตรวจสอบความถูกต้องของค่าตำแหน่งที่คำนวณได้ สำหรับการทดลองที่กำหนดให้ผู้ใช้นั่งอยู่บนตำแหน่งอ้างอิงที่กำหนดลงบนพื้น แล้วนำภาพที่ได้จากกล้องวิดีโอมาคำนวณหาตำแหน่งของผู้ใช้ ตำแหน่งต่าง ๆ ที่ใช้ในการทดลอง แสดงดังรูปที่ 5.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.3 การกำหนดตำแหน่งอ้างอิง

จากรูปที่ 5.3 กำหนดให้ตำแหน่งที่กล้องตั้งอยู่มีค่า  $X=0, Y=0$  โดยด้านขวาเป็น  $-Y$  และด้านซ้ายเป็น  $+Y$  และตำแหน่งที่ใช้ในการตรวจวัดความถูกต้องมีระยะห่างของแต่ละตำแหน่งเท่ากับ 20 เซนติเมตร ในการทดลองจะใช้ตำแหน่งทั้งหมด 10 ตำแหน่ง ดังตารางที่ 5.1

ตารางที่ 5.1 ตำแหน่งอ้างอิงที่ใช้สำหรับตรวจสอบความถูกต้องจากการทดลอง

ตำแหน่งที่	$X$ (เซนติเมตร)	$Y$ (เซนติเมตร)
1	90	0
2	110	40
3	110	-40
4	130	40
5	130	0
6	130	-40
7	170	40
8	170	0
9	170	-40

### ตารางที่ 5.1 (ต่อ)

ตำแหน่งที่	$X$ (เซนติเมตร)	$Y$ (เซนติเมตร)
10	190	0

### 5.2 ผู้ร่วมในการทดลอง

ในงานวิจัยนี้ ใช้ผู้ทดลองทั้งหมด 10 คน ซึ่งแต่ละคนจะมีลักษณะสรีระของร่างกายที่แตกต่างกัน และมีการแต่งกายที่แตกต่างกัน โดยผู้ทดลองแต่ละคนจะทำการทดลองคนละ 10 ครั้งในแต่ละการทดลอง โดยในการทดลองกำหนดให้ (ภาพของผู้ร่วมทดสอบแสดงในภาคผนวก ก)

$L$  คือ ระยะห่างระหว่างปลายเท้าทั้งสองด้าน มีหน่วยเป็นเซนติเมตร

$d$  คือ ค่าองศาในการขึ้นของผู้ทดลองที่ทำมุมกับกล้อง

### 5.3 การทดลอง

การทดสอบความถูกต้องของอัลกอริทึมในการวัดตำแหน่งขาได้แบ่งออกเป็น 4 การทดลอง คือ

- การวัดตำแหน่ง โดยยืนอยู่กับที่
- การวัดตำแหน่ง โดยการเดินเป็นเส้นตรง
- การวัดตำแหน่ง โดยการเดินวนไปมา
- การวัดคุณภาพจากการใช้งาน

#### 5.3.1 การวัดตำแหน่ง โดยยืนอยู่กับที่

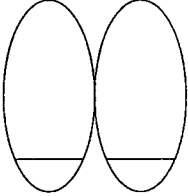
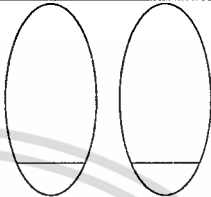
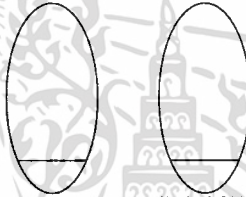
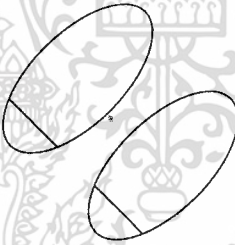
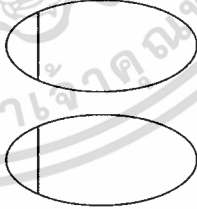
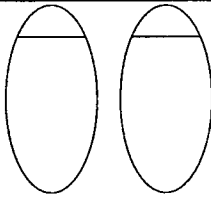
ในการทดลองนี้มีวัตถุประสงค์ดังต่อไปนี้

- เพื่อทดสอบความถูกต้องในการคำนวณค่าตำแหน่ง จากลักษณะการขึ้นของผู้ทดลองที่แตกต่างกัน 6 รูปแบบ

- เพื่อทดสอบความถูกต้องในการคำนวณค่าตำแหน่ง เมื่อผู้ทดลองยืนอยู่บนตำแหน่งอ้างอิงทั้งหมด 10 จุด

การทดลองนี้ กำหนดให้ผู้ทดลองยืนอยู่บนตำแหน่งอ้างอิงทั้งหมด 10 ตำแหน่งดังรูปที่ 5.3 โดยแต่ละตำแหน่งผู้ทดลองจะขึ้นในลักษณะที่แตกต่างกัน 6 รูปแบบ ดังตารางที่ 5.2 แล้วนำข้อมูลตำแหน่งที่คำนวณได้มาเปรียบเทียบความถูกต้องกับค่าตำแหน่งจริงของตำแหน่งอ้างอิงทั้งหมด 10 ตำแหน่ง

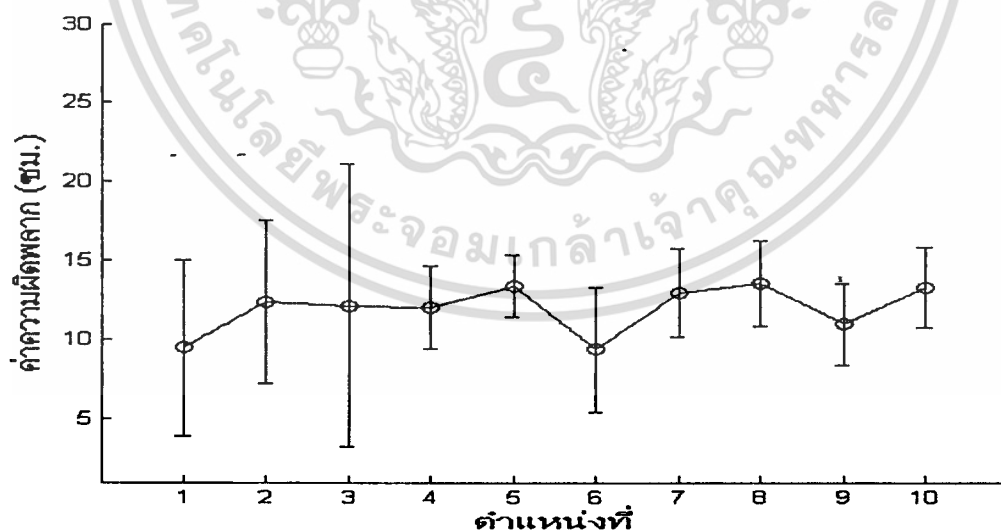
ตารางที่ 5.2 ลักษณะของการยื่นที่ใช้ในการทดลอง

รูปแบบที่	ลักษณะการยื่น	
1	$L = 0$ ซม.	
2	$L = 20$ ซม.	
3	$L = 40$ ซม.	
4	$d = 45$ องศา	
5	$d = 90$ องศา	
6	$d = 180$ องศา	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

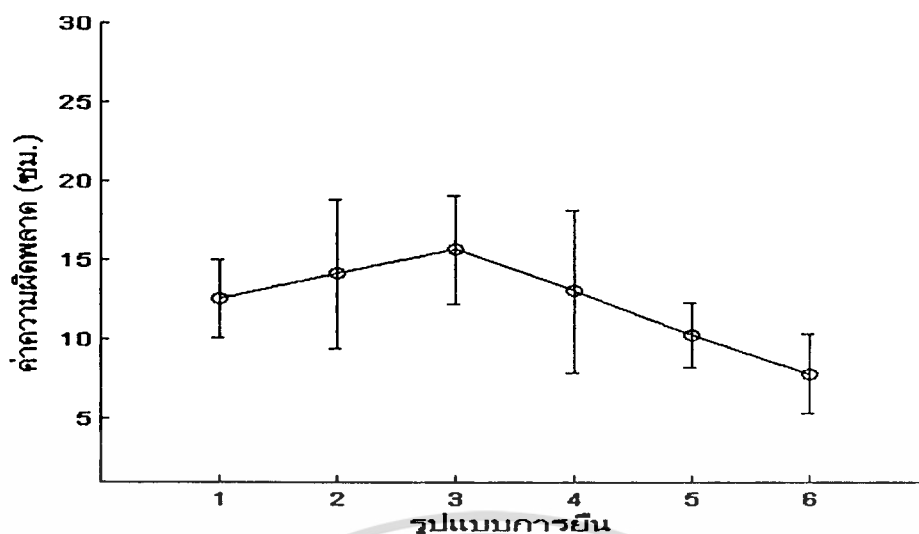
ตารางที่ 5.3 ค่าความผิดพลาดของค่าตำแหน่งที่คำนวณได้เปรียบเทียบกับตำแหน่งอ้างอิง

ตำแหน่งที่	X (ซม.)	Y (ซม.)	ค่าความผิดพลาดของการยึนรูปแบบที่ (ซม.)						ค่าเฉลี่ย
			1	2	3	4	5	6	
1	90	0	15.18	14.64	12.72	01.80	08.16	04.51	09.50
2	110	40	09.86	11.45	21.69	13.76	11.38	06.25	12.40
3	110	-40	12.03	25.33	20.52	04.78	06.86	03.55	12.18
4	130	40	09.80	12.91	13.52	15.06	13.11	07.99	12.06
5	130	0	16.01	14.78	14.25	12.76	10.63	12.25	13.45
6	130	-40	09.75	06.19	16.92	09.00	07.75	06.75	09.40
7	170	40	13.76	13.81	13.81	16.78	11.16	08.55	12.98
8	170	0	15.26	14.78	14.77	16.18	11.55	09.02	13.59
9	170	-40	12.00	14.02	10.76	13.15	08.87	07.29	11.01
10	190	0	14.50	14.09	14.81	16.02	11.54	09.09	13.34
ค่าเฉลี่ย			12.55	14.15	15.67	13.05	10.32	7.86	



รูปที่ 5.4 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งเปรียบเทียบกับตำแหน่งอ้างอิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.5 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจากรูปแบบการย่นที่ใช้ในการทดลอง

จากการทดลอง ผลของความถูกต้องของค่าตำแหน่งที่คำนวณได้ดังรูปที่ 5.4 ตำแหน่งที่มีความถูกต้องของค่าตำแหน่งมากที่สุดคือตำแหน่งที่ 1 มีค่าความผิดพลาดเท่ากับ 2.66 เซนติเมตร ซึ่งมีระยะ  $X = 90$  เซนติเมตร และ  $Y = 0$  เซนติเมตร ซึ่งเป็นตำแหน่งที่อยู่ใกล้กับกล้องวิดีโอมากที่สุด

ผลของความถูกต้องของค่าตำแหน่งที่คำนวณได้เมื่อย่นในรูปแบบที่กำหนดดังรูปที่ 5.5 รูปแบบการย่นที่มีค่าความถูกต้องของค่าตำแหน่งมากที่สุดคือการย่นในรูปแบบที่ 6 ซึ่งเป็นการย่นทำมุมกับกล้องวิดีโอ 180 องศา หรือการย่นหันหลังให้กล้องวิดีโอ

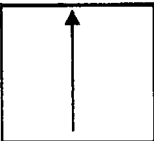
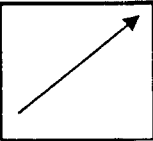
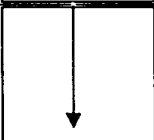
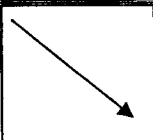




### 5.3.2 การวัดตำแหน่ง โดยการเดินเป็นเส้นตรง

ในการทดลองนี้มีวัตถุประสงค์ดังต่อไปนี้

- เพื่อทดสอบความถูกต้องในการคำนวณค่าตำแหน่ง เมื่อเดินในทิศทางที่แตกต่างกัน
- เพื่อหาจำนวนค่าตำแหน่งสำหรับ Smoothing algorithm ที่ทำให้มีความถูกต้องของค่าตำแหน่งที่คำนวณได้มีความถูกต้องสูงที่สุด

การทดลองนี้ กำหนดให้ผู้ทำการทดลองเดินภายในพื้นที่ทำงานขนาด  $2 \times 2$  เมตร ตามทิศทางที่แตกต่างกัน 8 รูปแบบ ดังตารางที่ 5.4 และนำค่าตำแหน่งที่คำนวณได้ มาคำนวณค่าตำแหน่งใหม่ด้วย Smoothing algorithm โดยกำหนด  $N = 4, 6, 8, 10, 12, 14, 16, 18$  และ  $20$  ซึ่งการเดินในแต่ละทิศทางจะใช้เวลาในการเดินเฉลี่ยเท่ากับ 3 วินาที

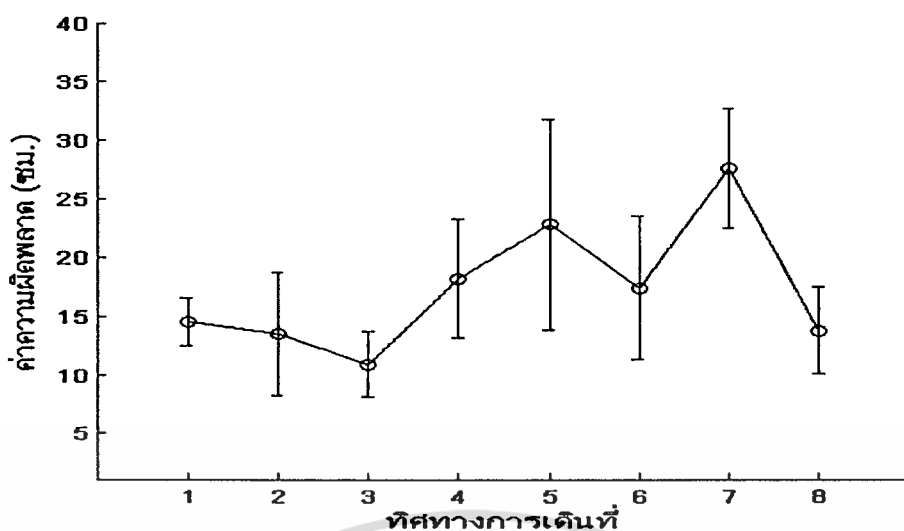
ตารางที่ 5.4 ทิศทางการเดินที่ใช้ในการทดลอง

รูปแบบที่	ทิศทางการเดิน	รูปแบบที่	ทิศทางการเดิน
1		5	
2		6	
3		7	
4		8	

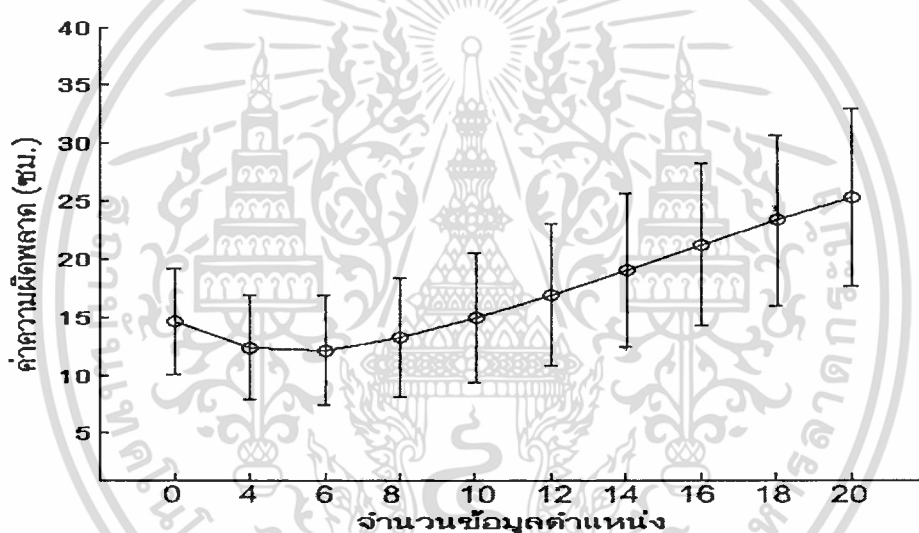
ตารางที่ 5.5 ค่าผิดพลาดของการคำนวณตำแหน่งจากทิศทางการเดินที่ใช้ในการทดลอง

รูปแบบที่	จำนวนข้อมูลตำแหน่ง										ค่าเฉลี่ย
	0	4	6	8	10	12	14	16	18	20	
1	16.07	13.73	12.42	12.46	12.65	13.11	14.07	15.45	16.93	18.35	14.52
2	10.40	07.32	07.48	09.16	11.29	13.67	15.90	18.00	19.98	21.89	13.51
3	10.17	09.04	08.22	08.00	08.83	09.90	11.22	12.78	14.49	16.27	10.89
4	13.19	12.35	13.12	14.56	16.37	18.41	20.59	22.65	24.50	26.14	18.19
5	11.61	11.40	14.14	17.70	21.25	24.59	27.80	30.72	33.31	35.67	22.82
6	15.87	11.53	09.87	11.25	13.92	16.78	19.67	22.51	25.16	27.61	17.42
7	24.16	22.41	22.25	23.35	24.98	27.04	29.41	31.86	34.28	36.66	27.64
8	15.38	11.23	09.66	09.62	10.62	12.10	14.04	16.21	18.32	20.34	13.75
ค่าเฉลี่ย	14.61	12.38	12.15	13.26	14.99	16.95	19.09	21.27	23.37	25.37	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.6 ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจากทิศทางการเดินทางที่ใช้ในการทดลอง



รูปที่ 5.7-ค่าความผิดพลาดในการคำนวณค่าตำแหน่งจาก Smoothing algorithm

จากผลการทดลองทิศทางการเดินทางดังรูปที่ 5.6 รูปแบบที่ 3 ซึ่งเป็นการเดินทางจากด้านซ้ายไปด้านขวาของพื้นที่ทำการทดลอง เป็นทิศทางที่มีความถูกต้องของค่าตำแหน่งที่คำนวณได้สูงที่สุดซึ่งมีค่าความผิดพลาดของค่าตำแหน่งที่คำนวณได้เท่ากับ 10.89 เซนติเมตร และผลของความถูกต้องของค่าตำแหน่งที่คำนวณด้วย Smoothing algorithm ดังรูปที่ 5.7 จำนวนข้อมูลตำแหน่งเท่ากับ 6 ตำแหน่ง จะทำให้การคำนวณค่าตำแหน่งมีความถูกต้องสูงที่สุดโดยมีค่าความผิดพลาดของค่าตำแหน่งที่คำนวณได้เท่ากับ 12.15 เซนติเมตร

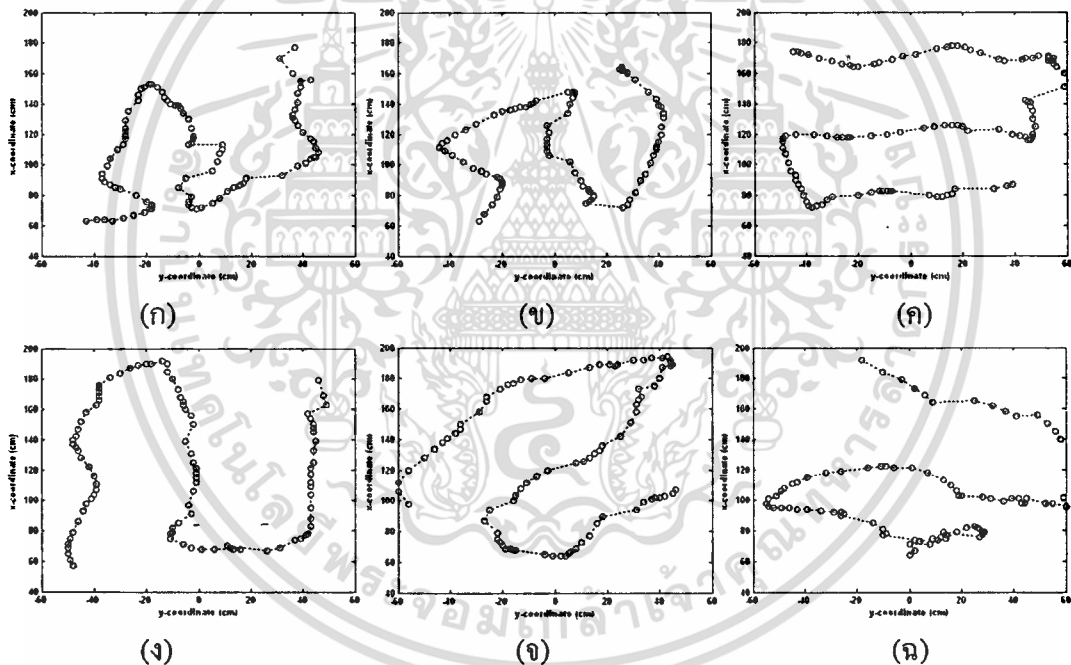
### 5.3.3 การวัดตำแหน่ง โดยการเดินวนไปมา

ในการทดลองนี้มีวัตถุประสงค์ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เพื่อทดสอบความถูกต้องในการคำนวณค่าตำแหน่ง จากรูปแบบการเดินที่กำหนด
- เพื่อคำนวณหาจำนวนค่าตำแหน่งสำหรับ Smoothing algorithm ที่ทำให้มีค่าความถูกต้องของค่าตำแหน่งที่คำนวณ ได้มีความถูกต้องสูงที่สุด
- เพื่อคำนวณหาจำนวนค่าตำแหน่งที่มีการ Delay จากการใช้ Smoothing algorithm ที่ทำให้มีค่าความถูกต้องของค่าตำแหน่งที่คำนวณ ได้มีความถูกต้องสูงที่สุด

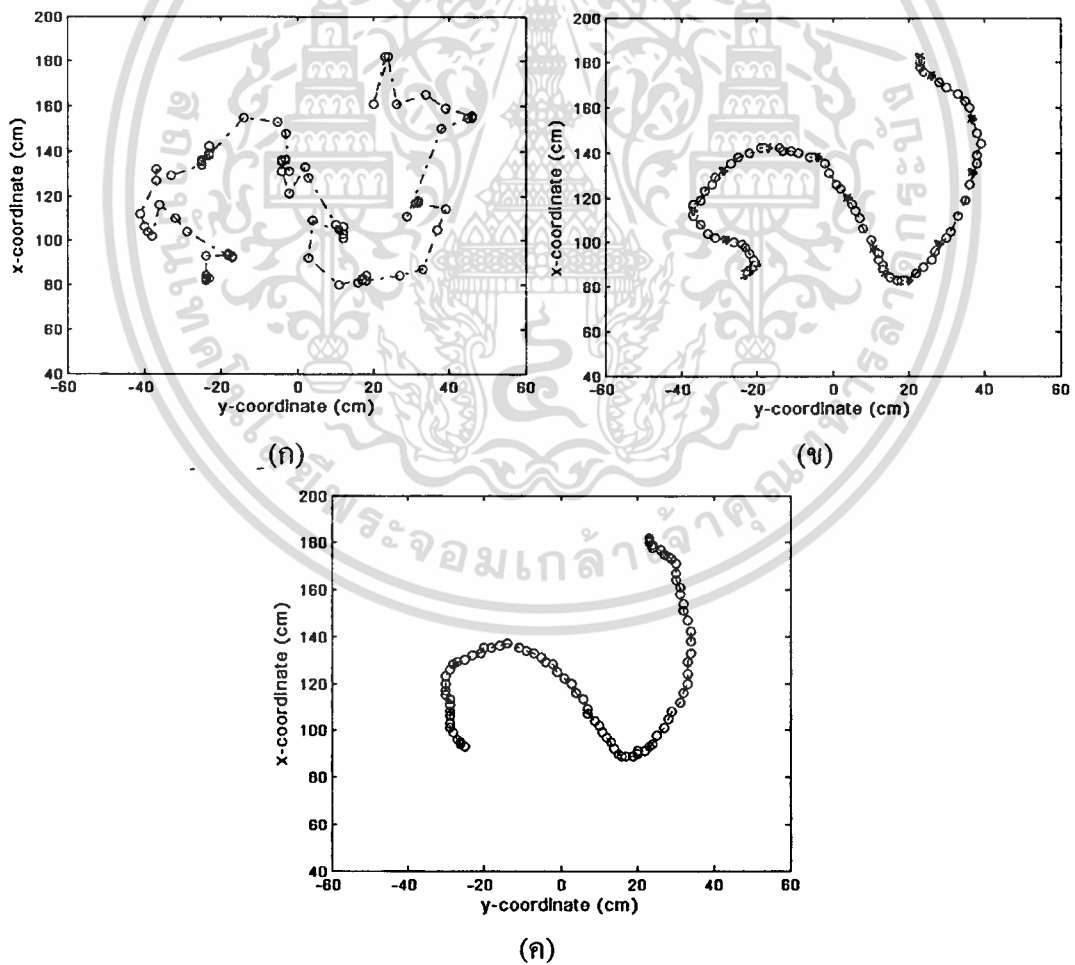
การทดลองนี้ กำหนดให้ผู้ทำการทดลองเดินภายในพื้นที่ทำงานขนาด 2x2 เมตร ตามรูปแบบการเดินที่กำหนด 6 รูปแบบ ดังรูปที่ 5.8 และนำค่าตำแหน่งที่คำนวณได้ มาคำนวณค่าตำแหน่งใหม่ด้วย Smoothing algorithm โดยกำหนด  $N = 4, 6, 8, 10, 12, 14, 16, 18$  และ 20 และค่าตำแหน่งที่คำนวณได้จาก Smoothing algorithm นำมาคำนวณหาตำแหน่งด้วยค่า Delay=1, 2, 3, 4, 5, 6, 7, 8, 9 และ 10 ตำแหน่ง ซึ่งการเดินในแต่ละรูปแบบจะใช้เวลาในการเดินเฉลี่ยเท่ากับ 12.33 วินาที



รูปที่ 5.8 รูปแบบการเดินที่ใช้ในการทดลอง (ก) รูปแบบที่ 1 (ข) รูปแบบที่ 2 (ค) รูปแบบที่ 3 (ง) รูปแบบที่ 4 (จ) รูปแบบที่ 5 และ (ฉ) รูปแบบที่ 6

ตารางที่ 5.6 ผลการทดลองของรูปแบบการเดินที่ 1 โดยคิดค่าความผิดพลาดของการคำนวณค่าตำแหน่งที่น้อยที่สุด

จำนวนค่าตำแหน่ง	Delay(ตำแหน่ง)	ค่าความผิดพลาด(ชม.)
0	2	8.67
4	0	11.13
6	0	9.46
8	1	8.92
10	2	8.58
12	3	8.34
14	4	8.25
16	4	8.40
18	5	8.57
20	6	8.89



รูปที่ 5.9 ผลการทดลองของการเดินในรูปแบบที่ 1 เมื่อคำนวณค่าตำแหน่งใหม่ด้วย Smoothing

algorithm (ก)  $N=0$  (ข)  $N=10$  และ (ค)  $N=20$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองของรูปแบบการเดินที่ 1 ดังตารางที่ 5.6 เมื่อกำหนดให้  $N = 0$  สำหรับ Smoothing algorithm ค่าความผิดพลาดที่ได้เท่ากับ 8.67 เซนติเมตร และมีค่า Delay = 2 ตำแหน่ง ซึ่งการ Delay ของค่าตำแหน่งที่เกิดขึ้นเนื่องมาจากประสิทธิภาพในการประมวลผลของเครื่องคอมพิวเตอร์ และเมื่อกำหนดให้  $N = 14$  ค่าความผิดพลาดที่ได้เท่ากับ 8.25 เซนติเมตร เป็นค่าความผิดพลาดที่น้อยที่สุด และมีค่า Delay = 4 ตำแหน่ง ดังนั้นเมื่อใช้จำนวนค่าตำแหน่งสำหรับ Smoothing algorithm สูงก็จะได้ค่าความถูกต้องของค่าตำแหน่งที่คำนวณได้สูงขึ้น แต่จะมีการ Delay ของค่าตำแหน่งสูงขึ้นด้วย และจากรูปที่ 5.9 เป็นการแสดงค่าตำแหน่งที่คำนวณได้ของการเดินในรูปแบบที่ 1 จากรูปที่ 5.9 (ก) กำหนดให้  $N = 0$  ค่าตำแหน่งที่คำนวณได้จะมีความต่อเนื่องของค่าตำแหน่งน้อย รูปที่ 5.9 (ค) กำหนดให้  $N = 20$  ค่าตำแหน่งที่คำนวณได้จะมีความต่อเนื่องมากกว่าในรูปที่ 5.9 (ก) ดังนั้นการใช้ค่า  $N$  สูงก็จะทำให้ค่าตำแหน่งที่คำนวณได้มีความต่อเนื่องมากขึ้น ซึ่งเมื่อนำค่าตำแหน่งที่คำนวณได้ไปใช้ในการสร้างภาพสำหรับแอปพลิเคชันที่แสดงผลภาพสเตอริโอ จะทำให้ภาพที่แสดงมีการสั่นไหวของภาพที่ได้น้อยกว่าเมื่อค่า  $N = 0$

ตารางที่ 5.7 ค่าผิดพลาดของการคำนวณตำแหน่งจากรูปแบบการเดินที่ใช้ในการทดลอง ที่ค่า

Delay = 0

รูปแบบที่	จำนวนข้อมูลตำแหน่ง										ค่าเฉลี่ย
	0	4	6	8	10	12	14	16	18	20	
1	14.07	11.13	9.46	9.04	9.59	10.67	12.32	14.14	15.89	17.56	12.39
2	13.94	10.47	9.41	9.71	11.16	13.12	15.42	17.89	20.46	22.96	14.46
3	14.76	11.98	10.97	10.84	11.44	12.79	14.54	16.56	18.71	20.84	14.34
4	15.68	11.28	9.60	10.29	12.04	14.58	17.79	21.01	24.14	27.16	16.36
5	16.55	13.13	12.69	13.93	15.99	18.48	21.17	23.85	26.46	28.92	19.12
6	15.29	12.67	12.47	13.48	14.82	16.40	18.22	20.04	21.83	23.56	16.88
ค่าเฉลี่ย	15.05	11.78	10.77	11.22	12.51	14.34	16.58	18.92	21.25	23.50	

ตารางที่ 5.8 ค่าผิดพลาดของการคำนวณตำแหน่งจากรูปแบบการเดินที่ใช้ในการทดลอง เมื่อมีค่าความผิดพลาดน้อยที่สุด

จำนวน ข้อมูล ตำแหน่ง	รูปแบบที่											
	1		2		3		4		5		6	
	error	delay	error	delay	error	delay	error	delay	error	delay	error	delay
0	8.67	2	9.18	2	10.30	2	9.59	2	16.55	0	15.29	0
4	11.13	0	10.47	0	11.98	0	11.28	0	13.13	0	12.67	0
6	9.46	0	9.41	0	10.97	0	9.60	0	12.15	1	12.43	1
8	8.92	1	9.18	1	10.66	1	9.41	1	11.41	2	12.27	1
10	8.58	2	9.18	2	10.38	2	9.43	2	10.98	3	12.21	2
12	8.34	3	9.30	3	10.18	3	9.46	3	10.87	4	12.23	3
14	8.25	4	9.56	4	10.09	4	9.55	4	11.05	5	12.35	4
16	8.40	4	9.94	5	10.16	5	9.88	5	11.34	6	12.61	5
18	8.57	5	10.36	6	10.29	6	10.51	6	11.72	7	13.01	6
20	8.89	6	10.83	7	10.51	6	11.30	7	12.15	8	13.63	6

ตารางที่ 5.9 เปรียบเทียบค่าความผิดพลาดที่น้อยที่สุดของการเดินแต่ละรูปแบบ

รูปแบบที่	ค่าความผิดพลาด (ชม.)	จำนวนข้อมูลตำแหน่ง	delay
1	8.25	14	4
2	9.18	8	1
3	10.09	14	4
4	9.41	8	1
5	10.87	12	4
6	12.21	10	2
ค่าเฉลี่ย	10	11	3

จากผลการทดลองของรูปแบบการเดินทั้ง 6 รูปแบบ ดังตารางที่ 5.9 ค่าเฉลี่ยความผิดพลาดจากการคำนวณค่าตำแหน่งเท่ากับ 10 เซนติเมตร และ ข้อมูลตำแหน่งที่ใช้สำหรับ Smoothing algorithm เฉลี่ยเท่ากับ 11 ตำแหน่ง และมีค่า Delay ของค่าตำแหน่งที่คำนวณได้เท่ากับ 3 ตำแหน่ง

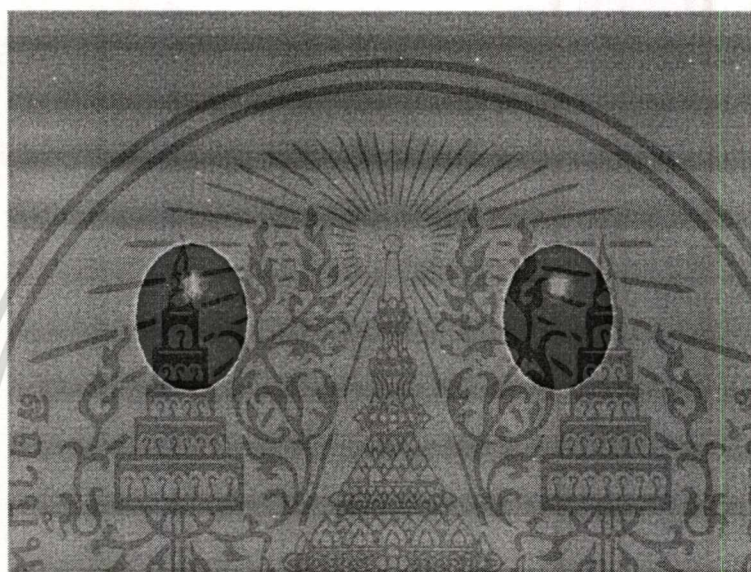
#### 5.3.4 การวัดคุณภาพจากการใช้งาน

ในการทดลองนี้มีวัตถุประสงค์ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เพื่อทดสอบระยะห่างของวัตถุจำลองที่สร้างบนจอภาพกับผู้ทดลองมีผลต่อคุณภาพของการใช้งาน

การทดลองนี้จะนำข้อมูลตำแหน่งที่ได้จากอุปกรณ์ตรวจวัดตำแหน่งแบบสนามแม่เหล็กไฟฟ้า และข้อมูลตำแหน่งที่คำนวณได้จากอัลกอริทึมในการตรวจวัดตำแหน่งมาสร้างเป็นภาพอนิเมชันของวัตถุจำลอง ดังรูปที่ 5.10 ภาพอนิเมชันของวัตถุจำลองจะสร้างโดยกำหนดระยะห่างที่ 90, 110, 130, 150, 170 และ 190 เซนติเมตร ผู้ทดลองจะดูภาพของวัตถุจำลองทั้ง 2 เปรียบเทียบกัน แล้วตอบแบบสอบถามถึงระดับของคุณภาพของภาพวัตถุจำลองที่มองเห็น



รูปที่ 5.10 ภาพของวัตถุจำลองที่สร้างขึ้นจากข้อมูลตำแหน่งจากอุปกรณ์ตรวจวัดตำแหน่งและอัลกอริทึมในการตรวจวัดตำแหน่ง

ตารางที่ 5.10 การวิเคราะห์ผลระยะห่างของวัตถุจำลองต่อคุณภาพของภาพที่ผู้ทดลองรับรู้

รายการ	คุณภาพของภาพ	
	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน
1. ระยะห่าง 90 เซนติเมตร	3.5	0.53
2. ระยะห่าง 110 เซนติเมตร	3.5	0.53
3. ระยะห่าง 130 เซนติเมตร	4.5	0.53
4. ระยะห่าง 150 เซนติเมตร	4.5	0.53
5. ระยะห่าง 170 เซนติเมตร	4.7	0.48
6. ระยะห่าง 190 เซนติเมตร	5.0	0.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### การแปลความหมายผลของการวิเคราะห์

ค่าเฉลี่ยตั้งแต่ 0.00 – 1.00 คือ คุณภาพของภาพน้อยที่สุด

ค่าเฉลี่ยตั้งแต่ 1.01 – 2.00 คือ คุณภาพของภาพน้อย

ค่าเฉลี่ยตั้งแต่ 2.01 – 3.00 คือ คุณภาพของภาพปานกลาง

ค่าเฉลี่ยตั้งแต่ 3.01 – 4.00 คือ คุณภาพของภาพสูงมาก

ค่าเฉลี่ยตั้งแต่ 4.01 – 5.00 คือ คุณภาพของภาพสูงมากที่สุด

จากตารางที่ 5.10 ที่ระยะห่าง 90 เซนติเมตรมีคุณภาพของภาพในระดับสูงมาก ที่ระยะห่าง 110 เซนติเมตรมีคุณภาพของภาพในระดับสูงมาก ที่ระยะห่าง 130 เซนติเมตรมีคุณภาพของภาพในระดับสูงมากที่สุด ที่ระยะห่าง 150 เซนติเมตรมีคุณภาพของภาพในระดับสูงมากที่สุด ที่ระยะห่าง 170 เซนติเมตรมีคุณภาพของภาพในระดับสูงมากที่สุด ที่ระยะห่าง 190 เซนติเมตรมีคุณภาพของภาพในระดับสูงมากที่สุด



## บทที่ 6

# สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอวิธีการตรวจวัดตำแหน่งขาด้วยกล้องวีดีโอตัวเดียวสำหรับระบบการแสดงผลสเตอริโอเชิงโต้ตอบ โดยใช้วิธีการที่ง่าย รวดเร็ว และต้นทุนของอุปกรณ์ที่นำมาใช้มีราคาต่ำ ซึ่งมีความเหมาะสมสำหรับคอมพิวเตอร์ส่วนบุคคลที่มีความสามารถในการแสดงผลสเตอริโอเชิงโต้ตอบ

### 6.1 สรุปผลการวิจัย

จากเทคนิคที่พัฒนาขึ้นเพื่อการตรวจวัดตำแหน่งขาตลอดจนวิธีในการปรับปรุงค่าที่คำนวณได้ในการวิจัย ผู้วิจัยได้ทำการทดสอบประสิทธิภาพของวิธีการดังกล่าว โดยคำนึงถึงปัจจัยต่าง ๆ ที่เกี่ยวข้อง สรุปผลการศึกษาได้ดังต่อไปนี้

1) จากการทดลองที่ 1 ซึ่งกำหนดให้ผู้ทำการทดลองยืนบนตำแหน่งอ้างอิง โดยแต่ละตำแหน่งมีการกำหนดรูปแบบการยืนทั้งหมด 6 รูปแบบ ผลลัพธ์ที่ได้จากการทดลองลักษณะการยืนหันหลังให้กล้องวีดีโอ 180 องศา จะมีค่าความถูกต้องของค่าตำแหน่งที่คำนวณได้สูงที่สุด และการยืนหันหน้าให้กับกล้องวีดีโอ 0 องศาและระยะปลายเท้าทั้ง 2 ด้านห่าง 40 เซนติเมตร จะมีค่าความถูกต้องของค่าตำแหน่งที่คำนวณได้ต่ำที่สุด เนื่องจากอัลกอริทึมในการตรวจวัดตำแหน่งขาเป็นการคำนวณหาจุดปลายของเท้าบนระนาบ การยืนหันหน้าให้กล้องวีดีโอจะมีส่วนของปลายเท้ายื่นออกมาทำให้การตรวจวัดตำแหน่งที่ได้เป็นตำแหน่งของปลายเท้า และการยืนหันหลังให้กล้องวีดีโอ 180 องศา จะไม่มีส่วนของเท้ายื่นออกมาทำให้การตรวจวัดตำแหน่งมีความถูกต้องสูง

ผลลัพธ์ที่ได้จากการทดลองเมื่อระยะห่างจากกล้องวีดีโอสูงขึ้น ความถูกต้องของค่าตำแหน่งที่คำนวณได้จะมีความถูกต้องน้อยลง เนื่องจากเมื่อระยะห่างสูงขึ้นขนาดของขาของผู้ทดลองในภาพวีดีโอก็จะมีขนาดเล็ก เมื่อเทียบกับขนาดของขาที่มีระยะห่างจากกล้องวีดีโอน้อยกว่า ดังนั้นเมื่อมีระยะห่างสูง ขนาดของขาในภาพวีดีโอก็จะมีขนาดเล็กจึงทำให้จำนวนของจุดภาพที่นำมาใช้ในการคำนวณน้อยทำให้มีค่าความผิดพลาดจากการคำนวณค่าตำแหน่งสูง

2) จากการทดลองที่ 2 ซึ่งกำหนดให้ผู้ทำการทดลองเดินในทิศทางที่กำหนดภายในพื้นที่ทำการทดลองทั้งหมด 8 รูปแบบ พบว่าทิศทางการเดินในแนวขวางกับทิศทางของกล้อง จะมีค่าความถูกต้องของค่าตำแหน่งที่คำนวณได้สูงที่สุด เนื่องจากอัลกอริทึมในการคำนวณค่าตำแหน่งจะสามารถคำนวณค่าตำแหน่งได้ถูกต้องสูง เท้าทั้งสองข้างควรอยู่ในลักษณะสมมาตร แต่การเดินจะมีการยกเท้าขึ้น ทำให้เสียความสมมาตรของเท้าทั้ง 2 ไป การเดินในทิศทางขวางกับทิศทางของกล้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำให้เท้าทั้ง 2 ข้างมีลักษณะใกล้เคียงกับความสมมาตรสูงกว่าท่าเดินในทิศทางอื่น ๆ (รูปที่ 6.1 ประกอบ)



(ก)



(ข)

รูปที่ 6.1 (ก) การเดินในทิศทางด้านซ้ายไปด้านขวา (ข) การเดินในทิศทางด้านหลังมาด้านหน้า

3) ในการทดลองที่ 3 กำหนดให้ผู้ทำการทดลองเดินภายในพื้นที่ทำงานขนาด 2x2 เมตร ตามรูปแบบการเดินที่กำหนด 6 รูปแบบและนำค่าตำแหน่งที่คำนวณได้ มาคำนวณค่าตำแหน่งใหม่ด้วย Smoothing algorithm ผลลัพธ์ที่ได้จากการทดลอง ค่าเฉลี่ยความผิดพลาดจากการคำนวณค่าตำแหน่งในการเดินทั้ง 6 รูปแบบเท่ากับ 10 เซนติเมตร และ ข้อมูลตำแหน่งที่ใช้สำหรับ Smoothing algorithm เฉลี่ยเท่ากับ 11 ตำแหน่ง และมีค่า Delay ของค่าตำแหน่งที่คำนวณได้เท่ากับ 3 ตำแหน่ง และจากการทดลองเมื่อใช้จำนวนข้อมูลตำแหน่งสำหรับ Smoothing algorithm น้อยความต่อเนื่องของค่าตำแหน่งจะน้อยกว่าการใช้จำนวนค่าตำแหน่งสำหรับ Smoothing algorithm ที่สูงกว่า แต่การใช้จำนวนค่าตำแหน่งสูงก็จะส่งผลให้มีการ Delay ของค่าตำแหน่งที่คำนวณได้สูงตามไปด้วย

4) ในการทดลองที่ 4 เป็นการศึกษาผลอันเนื่องมาจากระยะการมองต่อคุณภาพของภาพที่ได้ในเชิงคุณภาพ (subjective measure) โดยการนำข้อมูลตำแหน่งที่ได้จากอุปกรณ์ตรวจวัดตำแหน่งแบบสนามแม่เหล็กไฟฟ้า และข้อมูลตำแหน่งที่คำนวณได้จากอัลกอริทึมในการตรวจวัดตำแหน่งมาสร้างเป็นภาพอนิเมชันของวัตถุจำลอง โดยกำหนดระยะห่างที่ 90, 110, 130, 150, 170 และ 190 เซนติเมตร ผู้ทดลองจะดูภาพทั้ง 2 เปรียบเทียบกัน แล้วตอบแบบสอบถามถึงระดับของคุณภาพของภาพที่มองเห็น ผลลัพธ์ที่ได้จากการทดลองเมื่อระยะของวัตถุจำลองที่สร้างบนจอภาพมีระยะห่างสูงขึ้น จะทำให้คุณภาพของภาพที่ผู้ทดลองมองเห็นมีคุณภาพสูงขึ้น เนื่องจากเมื่อระยะห่างสูงขึ้น ภาพที่ผู้ทดลองมองเห็นบนจอภาพจะมีขนาดเล็กลง ทำให้ผู้ทดลองมองเห็นความผิดพลาดที่เกิดขึ้นจากการคำนวณค่าตำแหน่งได้น้อย จึงทำให้คุณภาพของภาพที่ผู้ทดลองมองเห็นมีคุณภาพสูงขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 6.2 ข้อเสนอแนะสำหรับการพัฒนาในอนาคต

ในงานวิจัยนี้ได้ใช้อัลกอริทึมในการลบภาพพื้นหลังแบบง่าย ทำให้เมื่อมีการเปลี่ยนแปลงของแสง เช่น มากขึ้นหรือน้อยลง ทำให้อัลกอริทึมลบภาพพื้นหลังไม่สามารถทำงานได้อย่างมีประสิทธิภาพ จึงจำเป็นต้องมีการพัฒนาอัลกอริทึมลบภาพพื้นหลังที่มีความคงทนต่อ การเปลี่ยนแปลงของแสงมาใช้งาน

อัลกอริทึมที่ใช้ในการคำนวณค่าตำแหน่งในงานวิจัย ยังไม่สามารถคำนวณหาได้ว่าผู้ทดลองยืนหันหน้าในทิศทางใด หากต้องการนำวิธีดังกล่าวไปประยุกต์ใช้ในงานด้านอื่น ๆ เช่น เกมสกีอาร์ จำเป็นต้องมีการพัฒนาอัลกอริทึมในการคำนวณหาทิศทางการยืนของผู้ทดลอง



## เอกสารอ้างอิง

- [1] Kenneth R. Castleman. **Digital Image Processing**, Prentice-Hall incorporation. 1996.
- [2] L. Lipton. "StereoGraphics Developers' Hand Book" **Stereographics Corporation**, 1997.
- [3] Huiyu Zhou and Housheng Hu. "A Survey – Human Movement Tracking and Stroke Rehabilitation" **Department of Computer Sciences University of Essex United Kingdom**, December 8, 2004.
- [4] Sigrid Brunner, Oliver Bimber and Steffen Mader. "Report on Tracking Technology" **VirtualShowcases**, September 10, 2003.
- [5] J. D. Mulder and R. van Liere. "The Personal Space Station: Bringing Interaction Within Reach" **Proceeding of the Virtual Reality International Conference 2002**, 2002. pp. 73 – 81.
- [6] J. D. Mulder, J.Jansen and A. van Rhijn." An Affordable Optical Head Tracking System for Desktop VR/AR Systems" **Eurographics Workshop on Virtual Environment (2003)**. 2003. pp.215-233.
- [7] R. van Liere, J. D. Mulder. "Optical Tracking Using Projective Invariant Marker Pattern Properties" **Proceedings of the IEEE Virtual Reality 2003**. 2003. pp.191-198.
- [8] A.Criminisi, I. Reid, and A. Zisserman. "A Plane Measuring Device" **Image and Vision Computing**. June, 1999. pp. 625 - 634.
- [9] Y. Kitamura, T. Konishi, S. Yamamoto, and F. Kishino. "Interactive stereoscopic display for three or more users" **Computer Graphics Annual Conference Series (Proc. of SIGGRAPH 2001)**. 2001. pp.231-239.
- [10] Chin-Chun Chang and Wen-Hsiang Tsai. "Vision-Based Tracking and Interpretation of Human Leg Movement for Virtual Reality Applications" **IEEE Transactions on Circuits and Systems for Video Technology**. January.,2001. pp.9-24.
- [11] Sebastien Grange, Terrence Fong, and Charles Baur. "TLIB:a Real-time Computer Vision Library for HIC" **IEEE Transactions on Circuits and Systems for Video Technology**. January, 2001. pp.9-24.
- [12] M. Ribo, A. Pinz, and A. Fuhrmann. "A new Optical Tracking System for Virtual and Augmented Reality Applications" **In Proceedings of the IEEE Instrumentation and Measurement Technical Conference**. May, 2001. pp.1935-1936.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [13] M. Ribo. "State of the art report on optical tracking" **Technical Report VRVis 2001-25**. 2001.
- [14] T. Ohno and N. Mukawa. "A free-head, simple calibration, gaze tracking system that enables gaze-based interaction" **Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications**. March, 2004. pp.115-122.
- [15] R. Ruddaraju, A. Haro, K. Nagel, Quan T. Tran, Irfan A. Essa, G. Abowd and E. D. Mynatt. "Perceptual User Interfaces using Vision-based Eye Tracking" **Proceedings of the 5th international conference on Multimodal interfaces**. 2003. pp.227-233.
- [16] J. Rekimoto. "A Vision-Based Head Tracker for Fish Tank Virtual Reality – VR without Head Gear" **Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)**. 1995. pp.94-100.
- [17] A. Mulder. "Human movement tracking technology" **Technical Report 94-1, Simon Fraser University**. 1994.
- [18] John Vince. **Virtual Reality Systems**, Addison-Wesley. 1995.
- [19] Michael Wögerbauer. "Virtual Construction Kit," [Online]. Available : <http://www.vrvis.at/vr/constructionkit/index.html>. 2005.
- [20] Nikon Corporation . "Principle of Stereographs," [Online]. Available : <http://www.nikon.co.jp/main/eng/feelnikon/kumon/12e.htm>. 2005.
- [21] Virtual Soldier Research. "Virtual Reality For Santos™," [Online]. Available : <http://www.digital-humans.org/Report2004>. 2004.
- [22] Rainer Splechtma and Anton Fuhrmann. "Extending the Working Volume of Projection-based Mixed Reality Systems," [Online]. Available : <http://www.digital-humans.org/Report2004>. 2004.
- [23] Spotlight. "Parts of the human body," [Online]. Available : <http://www.spotlightradio.net/wordlist.php>. 2005.
- [24] Polhemus. "Fastrak The Fast and Easy Digital Tracker," [Online]. Available : <http://www.polhemus.com/fastrak.htm>. 2004.
- [25] J. Ahn, C. Sul, E. Park and K. Wohn. "Development of LAN-based Optical Motion Capture System," [Online]. Available : <http://vr.kaist.ac.kr/~chocchoggi/>. 2002.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

ผู้ร่วมในการทดลอง

ในภาคผนวก ก. เป็นผู้ร่วมในการทดลองในงานวิจัยนี้ ซึ่งมีทั้งหมด 10 คน ซึ่งแต่ละคนจะมีลักษณะสรีระของร่างกาย และการแต่งกายที่แตกต่างกัน



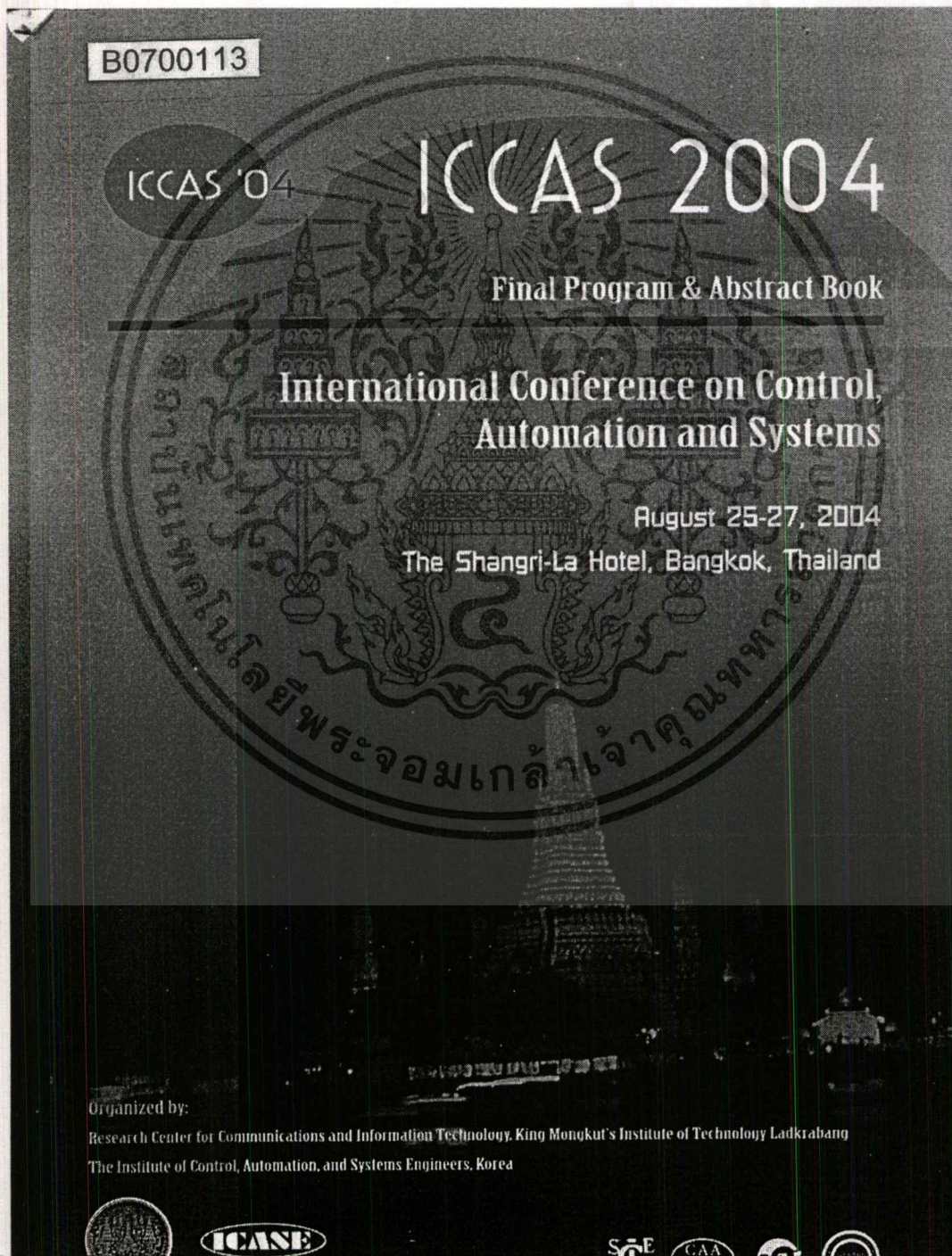
รูปที่ ก.1 ผู้ร่วมในการทดลองทั้งหมด 10 คน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข.

## บทความและผลงานวิจัยที่ได้รับการตีพิมพ์

1. Narongsak putpuek Chotikakamthorn. "A Fast Vision-based Head Tracking Method for Interactive Stereoscopic Viewing", International Conference on Control, Automation and Systems, August 25-27,2004. pp. 725 – 728



13.30 ~ 13.50 TP05-1

Digital Watermarking Technique for Images with Perspective Distortion

Nopporn Chotikakamthorn and Wiyada Yawar

In this paper, a problem of geometrically distorted images is considered. In particular, the paper discusses the detection of a watermark from a photographed image of the watermarked picture. The image is possibly obtained by using a digital camera. This watermark detection problem is made difficult by various geometric distortions added to the original picture through the printing and photographing processes. In particular, the paper focuses on the geometric distortion due to a projective transformation, as part of a camera 3D-to-2D imaging process. It is well-known that a cross ratio of collinear points is invariant under a perspective projection. By exploiting this fact, a projective-invariant digital watermarking technique is developed. By detecting the picture's corners, and the image center point at the intersection of two main diagonal lines, predefined cross ratios are used to compute the watermark embedded locations. From those identified embedding pixel locations, a watermark can be detected by performing a correlation between a watermark pattern and the image over those pixels. The proposed method does not require an inverse transformation on the distorted image, thus simplifying the detection process. Performance of the proposed method has been analyzed through computer experiments.

perception. However, because the algorithm and device used are designed for the computation, the estimation is typically not precise enough to provide a flicker-free interactive viewing. An error concealment method is thus proposed to alleviate the problem. This concealment method should be sufficient for applications that do not require a high degree of visual realism and interaction.

14.50 ~ 15.10 TP05-3

Hybrid HMM for Transitional Gesture Classification in Thai Sign Language Translation

Arune Jarwanawat, Nopporn Chotikakamthorn and Worawit Warapan

A human sign language is generally composed of both static and dynamic gestures. Each gesture is represented by a hand shape, its position, and hand movement (for a dynamic gesture). One of the problems found in automated sign language translation is on segmenting a hand movement that is part of a transitional movement from one hand gesture to another. This transitional gesture conveys no meaning, but serves as a connecting period between two consecutive gestures. Based on the observation that many dynamic gestures as appeared in Thai sign language dictionary are of quasi-periodic nature, a method was developed to differentiate between a (meaningful) dynamic gesture and a transitional movement. However, there are some meaningful dynamic gestures that are of non-periodic nature. These gestures cannot be distinguished from a transitional movement by using the signal quasi-periodicity. This paper proposes a hybrid method using a combination of the periodicity-based gesture segmentation method with a HMM-based gesture classifier. The HMM classifier is used here to detect dynamic signs of non-periodic nature. Combined with the periodic-based gesture segmentation method, this hybrid scheme can be used to identify segments of a transitional movement. In addition, due to the use of quasi-periodic nature of many dynamic sign gestures, dimensionality of the HMM part of the proposed method is significantly reduced, resulting in computational saving as compared with a standard HMM-based method. Through experiment with real measurement, the proposed method's recognition performance is reported.

13.50 ~ 14.10 TP05-2

Preliminary programming for librarization of Haptic Primitives based on constructive solid geometry and god-object

Do-Hyung Jin, Ki-Uk Kyung and Dang-Soa Kwon

In this paper, we propose "the haptic primitive" for haptic rendering without solving complicated parametric equations. In order to develop "the haptic primitive", we adopted "the God-Object Method" as a haptic rendering algorithm and applied "Constructive Solid Geometry (CSG)" to manage haptic objects. Beyond the "ghost library" of PHANTOM this research can be utilized as a basic component in developing tools and libraries that aim to simplify the haptic modeling. It also can be applied to tactile display modules and temporal display modules, and ultimately developed into a One-Stop Haptic Modeling Tool, enabling users to create tangible CAD systems or Tangible e-commerce systems more conveniently.

14.10 ~ 14.30 TP05-3

Multiresolution Wavelet-Based Disparity Estimation for Stereo Image Compression

Chompoonuch Tangcharoen and Rattikom Varakulsinpunth

The ordinary stereo image of an object consists of data of left and right views. Therefore, the left and right image pairs have to be transmitted simultaneously in order to display 3-dimensional video at the remote site. However, due to the twice data in comparing with a monoscopic image of the same object, it needs to be compressed for fast transmission and resource saving. Hence, it needs an effective coding algorithm for compressing stereo image. It was found previously that compressing left and right frames independently will achieve the compression ratio lower than compressing by utilizing the spatial redundancy between both frames. Therefore, in this paper, we study the stereo image compression technique based on the multiresolution wavelet transform using varied disparity-block size for estimation and compensation. The size of disparity-block in the stereo pair sub-bands are scaling on a coarse-to-fine wavelet coefficients strategy. Finally, the reference left image and residual right image after disparity estimation and compensation are coded by using SPIHT coding. The considered method demonstrates good performance in both PSNR measures and visual quality for stereo image.

14.30 ~ 14.50 TP05-4

A Fast Vision-based Head Tracking Method for Interactive Stereoscopic Viewing

Narungsak Putpuek and Nopporn Chotikakamthorn

In this paper, the problem of a viewer's head tracking in a desktop-based interactive stereoscopic display system is considered. A fast and low-cost approach to the problem is important for such a computing environment. The system under consideration utilizes a shuttle glass for stereoscopic display. The proposed method makes use of an image taken from a single low-cost video camera. By using a simple feature extraction algorithm, the obtained points corresponding to the image of the user-worn shuttle glass are used to estimate the glass center, its local 'yaw' angle, as measured with respect to the glass center, and its global 'yaw' angle as measured with respect to the camera location. With these estimations, the stereoscopic image synthetic program utilizes those values to interactively adjust the two-view stereoscopic image pair as displayed on a computer screen. The adjustment is carried out such that the so-obtained stereoscopic picture, when viewed from a current user position, provides a close-to-real perspective and depth

13.30 ~ 13.5

Design an

Water co existence of a that contain v be found at Venice of the us better unit presents the which was o open-frame areas in the integral- de results sho model of the will be impl

13.50 ~ 1

Auto Un

In the r vessel, pr rolled pla integrity c such as u under wa The insp with a big and han transport polar on Setup b several an unde experim our robr the res

14.10

A

The future cente: grow recen RCP( will be robot RCP, and r abilit that: algo syst this dev to ti It is The be: cor the fre res: int Pt sy

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรรมใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## A Fast Vision-based Head Tracking Method for Interactive Stereoscopic Viewing

Narongsak Puptuek, and Nopporn Chotikakamthorn

Faculty of Information Technology &  
Research Center for Communications and Information Technology  
King Mongkut's Institute of Technology Ladkrabang, Bangkok Thailand  
(E-mail: s3067144@kmitl.ac.th)

**Abstract:** In this paper a problem of a viewer's head tracking in a desktop-based interactive stereoscopic display system is considered. A fast and low-cost approach to the problem is important for such a computing environment. The system under consideration utilizes a shuttle glass for stereoscopic display. The proposed method makes use of an image taken from a single low-cost video camera. By using a simple feature extraction algorithm, the obtained points corresponding to the image of the user-worn shuttle glass are used to estimate the glass center, its local 'yaw' angle, as measured with respect to the glass center, and its global 'yaw' angle as measured with respect to the camera location. With these estimations, the stereoscopic image synthetic program utilizes those values to adaptively adjust the two-view stereoscopic image pair as displayed on a computer screen. The adjustment is carried out such that the so-obtained stereoscopic picture, when viewed from a current user position, provides a close-to-real perspective and depth perception. However, because the algorithm and device used are designed for fast computation, the estimation is typically not precise enough to provide a flicker-free interactive viewing. An error concealment method is thus proposed to alleviate the problem. This concealment method should be sufficient for applications that do not require a high degree of visual realism and interaction.

**Keywords:** Stereoscopic Display, Computer Vision, Interactive System

### 1. INTRODUCTION

A stereoscopic display system presents the left and right eyes of the viewer with images from different perspective viewpoints, just as the viewer sees the virtual world [1]. From these two slightly different views, the eye-brain synthesizes an image of the world with stereoscopic depth. A single, not double, image is seen since the two are fused by the mind into one. Thus, the system can provide an illusion of depth perception just like a human vision experience when viewing a real-world object. Stereoscopic display is used for a variety of applications from games to computer-aided design. However, a standard stereoscopic display system is limited to delivering an image pair perfectly viewed only from a single view point. Changing of viewer's position does not result in perspective change of the viewed scene. Interactive stereoscopic display system provides an answer to the problem. By adaptively adjusting the rendering view point of the synthetic stereoscopic image pair according to the current viewer's position, a high degree of virtual realism is obtained. To obtain the important information about current viewer's spatial position, a head tracking sensor/subsystem may be employed. Tracking of a viewer's head position is achieved by using one of these sensing technologies: electromagnetic sensor, optical sensor, as well as inertial, acoustical or mechanical sensing devices.

Use of these sensing devices is limited to high-end applications such as computer-aided design or VR-based simulation systems. For desktop applications, simpler technique is required. Examples of previous work on desktop-based virtual and augmented reality (VR/AR) systems include those of [2-3]. The Personal Space Station (PSS) is described in [4]. The system allows its user to perform 3D interactive. In this system, all interactive 3D tasks are performed directly with the hands or by using task specific input devices. The PSS is designed with three major goals: low costs, ergonomics and direct 3D interaction. The system is based on an optical-based head tracking [4] technique. Two standard FireWire cameras are used to provide tracking. In

PSS system user is seated in front of the mirror which reflects the stereoscopic images of the virtual world as displayed by the monitor.

In this paper, we present a fast and low-cost optical head tracking system that uses single low-cost camera and a simple shuttle glass objection segmentation method. By exploiting the color and shape of the shuttle glass device, additional markers are not required. An algorithm is developed to provide rough estimation of the viewer's position. Error concealment method is proposed to alleviate the inaccuracy in viewer's position estimation.

The paper is organized as follows. First, the proposed interactive stereoscopic viewing system is introduced along with general image processing notations. Next, detail of user's head position estimation is given in Section 3. Experiment results are provided with some discussion in Section 4. A conclusion is given in Section 5.

### 2. VISION-BASED INTERACTIVE STEREO SCOPIC VIEWING SYSTEM

The interactive stereoscopic viewing system is shown in Fig. 1. From the figure, a single video camera is placed on top of the CRT display unit. A stereoscopic shuttle glass is used to provide scene depth information to a viewer. A viewer is free to change his/her position relative to the camera. It is assumed, however, that local head rotation about the local z-axis is negligible. This assumption is made to simplify the heading tracking algorithm. The assumption should be approximately met for the usage scenario under study.

From Fig. 1 a user wearing a stereoscopic shuttle glass watches a stereoscopic computer-generated picture as shown in the display unit, semi-interactively. The term 'semi-interaction' here refers to the ability of the system to allow its user to change his/her physical viewpoint (e.g., by stepping to the right/left or back/front), and the rendered content is adjusted accordingly in a stepwise manner to provide certain degree of virtual realism. To create such virtual realism, the left and right images from which the

stereoscopic picture is constructed must be rendered based on a viewer position.

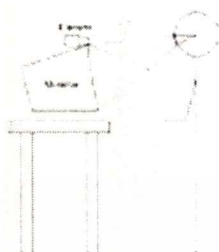


Fig. 1 The interactive stereoscopic viewing system



Fig. 2 Rendering of a stereoscopic image pair based on a viewer position



Fig. 3 Image plane coordinate

To realize such a system, estimates of a viewer's 'yaw' angles must be made available. In this paper, a simple vision-based method for estimating the global and local 'yaw' angles of a viewer's head is described. Estimations of both angular values ( $\alpha$  and  $\theta$ ) as shown in Fig. 2) are obtained by analyzing an image snapshot using a single low-cost video camera. Let's  $I(x, y)$  be the intensity function at the  $(x, y)$  pixel coordinate of the gray-scale image taken by a video camera at any time instance. Note that, here the  $(0, 0)$  pixel is located at the center of the image (see Fig. 3). From any snapshot  $I(x, y)$ , a user global nominal 'yaw' angle, as

measured with respect to the video camera coordinate is first estimated from the segments of  $I(x, y)$  corresponding to the left and right viewing areas of the shuttle glass. Then, estimation of the viewer's distance  $d$  is next obtained. The so-obtained estimated angle  $\theta$  and distance  $d$  are then used in making a decision on whether to change the rendering viewpoint of the stereoscopic image pair as displayed by the system's CRT.

In the next section, detail of an algorithm used to estimate  $\theta$  and  $d$  is described.

### 3. ALGORITHM FOR HEAD'S POSITIONAL ESTIMATION

#### 3.1 Image segmentation and feature extraction algorithm

The first step in the head's positional estimation is to perform necessary feature extraction from the camera image snapshot. The algorithm is described as follows.

##### Image Segmentation and feature extraction algorithm

1. Convert  $I(x, y)$  to a binary image. Segments of images areas of black pixels are then identified.
2. For each segmented black-pixel area in the binary version of  $I(x, y)$ , calculate the segment centroid position (as denoted by  $(x_i, y_i)$  in Fig. 4).
3. Let  $d_x$  and  $d_y$  represent the minimum width and height of the image area corresponding to either of the two glass's viewing (LCD) areas. From the two values, for each segmented black-pixel area, starting from the segment centroid position, find pixels at the four boundaries of the segment (see Fig. 4). These pixel positions are denoted by  $(x_j, y_j), j = 1, \dots, 8$ .
4. For each black-pixel image area, from the values of  $(x_j, y_j)$  as obtained from Step 3, determine  $m_{xy}$ .

$m_{x_{\text{center}}}, m_{y_{\text{center}}}$  and  $m_{x_{\text{edge}}}$  (see Fig. 4) from

$$m_{xy} = \frac{y_2 - y_1}{x_2 - x_1} \quad (1)$$

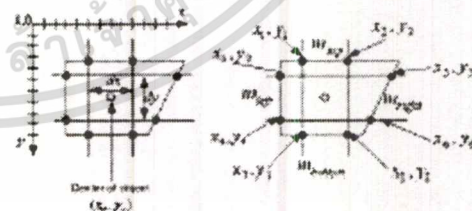


Fig. 4 Parameters corresponding to each segmented black-pixel area in  $I(x, y)$

$$m_{\text{left}} = \frac{X_0 - X_1}{Y_0 - Y_1} \quad (2)$$

$$m_{\text{right}} = \frac{X_2 - X_0}{Y_2 - Y_0} \quad (3)$$

$$m_{\text{bottom}} = \frac{Y_1 - Y_2}{X_1 - X_2} \quad (4)$$

5. For each black-pixel image area, calculate two standard deviation values as defined by the following equations:

$$\sigma_1 = \sqrt{\frac{(m_{\text{left}} - \bar{m}_{\text{left}})^2 + (m_{\text{top}} - \bar{m}_{\text{top}})^2 + (m_{\text{bottom}} - \bar{m}_{\text{bottom}})^2}{2}} \quad (5)$$

$$\sigma_2 = \sqrt{\frac{(m_{\text{right}} - \bar{m}_{\text{right}})^2 + (m_{\text{left}} - \bar{m}_{\text{left}})^2 + (m_{\text{bottom}} - \bar{m}_{\text{bottom}})^2}{2}} \quad (6)$$

where

$$\bar{m}_{\text{top}} = \frac{(m_{\text{left}} + m_{\text{right}} + m_{\text{bottom}})}{3} \quad (7)$$

$$\bar{m}_{\text{right}} = \frac{(m_{\text{right}} + m_{\text{left}} + m_{\text{bottom}})}{3} \quad (8)$$

6. Of the two values  $\sigma_1$  and  $\sigma_2$ , choose the smaller value and denote it by  $\sigma_{\text{min}}$ .
7. Among all segmented areas, select the two black-pixel segmented areas in the binary version of  $I(x, y)$  that have the smallest values of  $\sigma_{\text{min}}$ .

The two selected segments are then identified as corresponding to the two shuttle glass's viewing areas.

From the identified viewing areas as described, the next section details how to estimate  $\theta$  and  $d$ .

### 3.2 Estimation of the user's relative distance from a screen

It is first assumed that the camera focal length  $f$  is known. From the image  $I(x, y)$  obtained from the camera at any time instance, apply the feature extraction and segmentation algorithms as described above. From the so-obtained two shuttle glass's viewing areas as appeared in  $I(x, y)$ , the left and right borders, as well as the center point of the shuttle glass can be identified. It is assumed in the study that the user's head is oriented such that its z-axis is parallel to that of the camera, and the three projected points as mentioned pass through the image center. No assumption is made, however, regarding to the head's 'yaw' angle relative to the glass center (as denoted by  $\alpha$  in Fig. 2). Based on these assumptions, the distances between the three points, measured from the image center along the image's x-axis, are  $m_0$ ,  $m_1$ , and  $m_2$  as shown in Fig. 2. Therefore, the three angular values in Fig. 2 can be calculated from the following equations:

$$\theta = \tan^{-1}(m_0 / f) \quad (9)$$

$$\theta_1 = \tan^{-1}(m_1 / f) \quad (10)$$

$$\theta_2 = \tan^{-1}(m_2 / f) \quad (11)$$

From Fig. 2, it can be deduced that  $x_0$ ,  $x_1$ ,  $x_2$  are related to one another by the following equations:

$$x_0 = l \cos(\alpha) + l \sin(\alpha) \tan(\theta_1) - x_1 \quad (12)$$

$$x_1 = l \cos(\alpha) + l \sin(\alpha) \tan(\theta_2) + x_0 \quad (13)$$

where  $l$  is half the width of the shuttle glass. In addition, it can be verified that:

$$x_1 = d m_1 / f, \quad i = 0, 1, 2. \quad (14)$$

From Eqs. (9-14),  $\alpha$  can be obtained from:

$$\alpha = \tan^{-1}(A/B) \quad (15)$$

where

$$A = (m_0 - m_2) - (m_1 - m_0) \quad (16)$$

$$B = (m_1 - m_2) \tan(\theta_1) - (m_0 - m_2) \tan(\theta_2) \quad (17)$$

With  $\alpha$  as obtained from Eq. (15),  $x_1$ ,  $x_2$  can be obtained from Eqs. (12-13). As a result, by using Eq. (14),  $d$  is obtained from:

$$d = f(x_1 - x_2) / (m_1 - m_2) \quad (18)$$

Having obtained the estimates of both  $d$  and  $\theta$ , the graphic rendering subsystem can use these values to adjust the rendering such that the left and right images match the user viewing position.

In practice, by asking a user to perform initial calibration process, the rough estimate of  $f$  can be obtained.

### 3.3 Error concealment method

Because the estimation as detailed in Section 3.2 is not precise, due to limitation in terms of the image acquisition device's accuracy, as well as the need to reduce algorithm complexity as much as possible. Therefore, truly real-time interactive system is unachievable with sufficiently low flicker. The system developed here must therefore work in a semi-interactive mode. In particular, change of rendering viewpoint is made based on the following conditions:

- Change is made whenever the user's head position is sufficiently stationary. Head stationarity can be detected by calculating standard deviation of the two parameter estimates ( $d$  and  $\theta$ ) among those obtained from a few (say, 3-4) consecutive image snapshots.
- In addition, the whole viewing space is divided into discrete set of viewing areas. Only when the user's head position falls into a different viewing space that the system changes the rendering viewpoint.
- Switching between two image pairs of different viewpoints is performed gradually, by fading out the image pair corresponding to the current rendering viewpoint and fading in the image pair corresponding to the new rendering viewpoint.

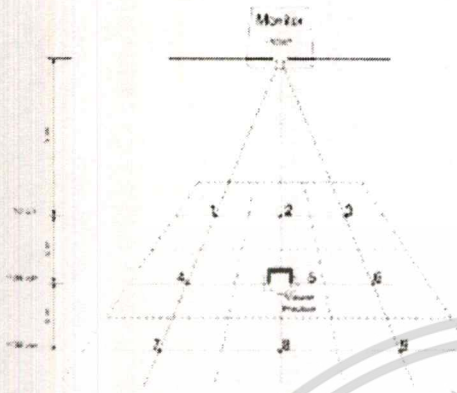


Fig. 5 Division of the whole viewing space into a few smaller viewing areas.

#### 4. EXPERIMENT

In the experiment, the viewing space is divided into 9 smaller viewing areas. The center position of each area is shown in Fig. 5.

First, the initial calibration step was performed using sample image snapshots. Each was taken from each of 9 different viewing areas. These sample images were used to obtain the estimate of the camera focal length  $f$  by using Eq. (13). The value of  $X_1 - X_2$  in Eq. (13) corresponds to the width of the shunt glass as measured from a real shunt glass.

Next, for each of 9 viewer positions, 10 image snapshots were taken, and used to estimate  $\theta$  and  $d$ . The result is shown in Table 1.

The estimates of  $\theta$  and  $d$  were then used to plot the viewer position over the segmented viewing areas, as shown in Fig. 6. The system chooses the rendering viewpoint based on the area where the estimated viewer position falls on. From Table 2, it was found that, based on the estimated viewer position, the averaged percentages of correct viewer position classification is 97.78%.

Table 1 Means and standard deviations of the estimates of  $\theta$  and  $d$  for each of 9 different view positions

Positions	$\theta$ (degree)		$d$ (cm.)	
	Mean	Std.	Mean	Std.
1	-17.42	0.41	79.25	0.90
2	0.60	1.41	82.94	1.05
3	14.76	0.63	74.27	1.12
4	-17.74	0.58	104.02	1.68
5	-0.64	0.53	113.95	1.27
6	14.64	0.64	99.70	0.82
7	-18.74	0.65	125.25	3.33
8	-2.27	0.36	134.39	1.99
9	15.73	0.67	128.10	7.16

Table 2 Positions classification

Position No.	Correct Classification (%)	Incorrect Classification (%)
1	100%	0%
2	100%	0%
3	100%	0%
4	100%	0%
5	90%	10%
6	100%	0%
7	90%	10%
8	100%	0%
9	100%	0%
Total	97.78%	2.22%

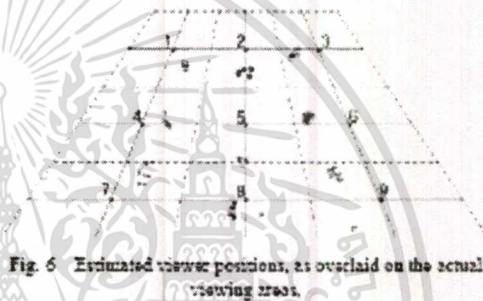


Fig. 6 Estimated viewer positions, as overlaid on the actual viewing areas.

#### 5. CONCLUSION

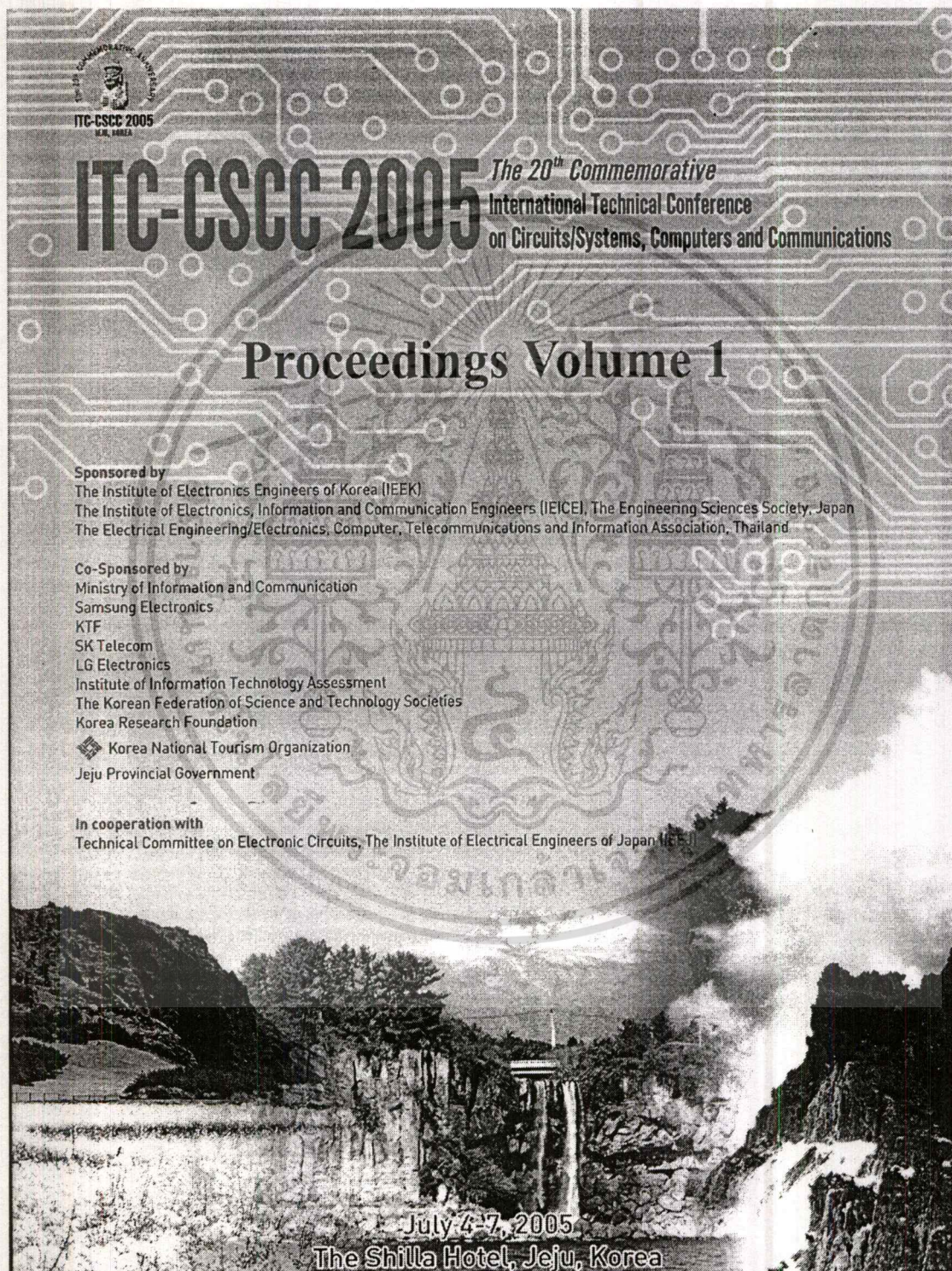
In this paper, a simple interactive stereoscopic viewing system has been described. A simple image segmentation and feature extraction method has been developed. From the so-obtained features, a method for estimation of the viewer's head position has been detailed. Error concealment strategy has been given to alleviate the problem due to imprecise head positional parameter estimation.

#### REFERENCES

- [1] L. Lipton, "StereoGraphics Developers' Handbook", Stereographics Corporation, 1997.
- [2] J. D. Mulder, J. Jensen, and A. van Rijn, "An Affordable Optical Head Tracking System for Desktop VR/AR Systems", *Eurographics Workshop on Virtual Environments (2003)*, pp. 215-233, 2003.
- [3] J. D. Mulder and R. van Liere, "The Personal Space Station: Bringing Interaction Within Reach", *Proceedings of the Virtual Reality International Conference 2002, VRIC 2002*, pp. 73-81, 2002.
- [4] R. van Liere, J. D. Mulder, "Optical Tracking Using Projective Invariant Marker Pattern Properties", *Proceedings of the IEEE Virtual Reality 2003*, pp. 191-198, 2003.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Narongsak putpuek, Natapon Pantuwong, Nopporn Chotikakamthorn and Olarn Wongwirat.  
 “A Low-cost Vision-based Leg Tracking Algorithm for an Interactive Stereoscopic System”  
 ITC-CSSC 2005, pp. 981 - 982 .



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## A Low-cost Vision-based Leg Tracking Algorithm for an Interactive Stereoscopic System

Narongsak Phurphuek, Natapon Panruwong, Nopporn Chotikakamthorn, and Olarn Wongwirat  
 Faculty of Information Technology & Research Center of Communications and Information Technology  
 King Mongkut's Institute of Technology Ladkrabang  
 Chalokkrung Road, Ladkrabang, Bangkok, Thailand  
 E-mail: S3067144@kmitl.ac.th

**Abstract:** A stereoscopic display system delivers images from different perspective viewpoints to the viewer's left and right eyes, providing a depth perception. A typical stereoscopic display system cannot deliver an image with correct perspective when a viewer changes his/her viewing position. With an interactive stereoscopic display system, a pair of displayed images is synthesized based on the current viewer's position. For such system, however, information about current viewer's position is required. In this paper, we present an interactive stereoscopic system that employs a vision-based leg tracking algorithm for the estimation of the viewer position. The method targets at a low-cost application such as a computer game or desktop-based products. It makes use of a single inexpensive video camera and a fast algorithm for viewer position estimation.

### 1. Introduction

An interactive stereoscopic display system can provide an illusion of spatial depth and deliver a correct-perspective image pair when a viewer changes his/her viewpoint [1]. For such a system, continually-updated information about current viewer position is needed. To obtain the viewer's position, various human movement tracking techniques may be employed. Tracking techniques can be classified into a non-vision based, a vision based with markers, and a vision based without markers methods [2]. Generally, a non-vision based method employs either a magnetic-based or acoustic-based tracking sensors. Those sensors, however, are generally expensive and require that a sensor device is first placed on a viewer body. Due to such limitations, this paper emphasizes a vision-based approach.

In [3], a desktop VR system is described. The system, called PSS, is a vision-based method with markers. Two cameras are used to track the space in which the interaction takes place. The system described in [4] is another vision-based method with markers. This system uses two low-cost standard FireWire cameras to track user head position and orientation in 3D. Other vision-based methods with markers include those of [5-6]. The system described in [5] uses a single camera for tracking and interpreting leg motion in image sequences to provide a gesture-based human-computer interaction. The method in [6] uses 2 camera and markers placed on a HMD to track viewer 3D head position. The working space of the method, however, appears to be limited.

In this paper, the main objective is to develop a viewer position estimation to be used in a low-cost, small-to-medium working area, interactive stereoscopic system. Thus, a single camera, instead of two, is used. In addition, we consider a shutter glass-based stereoscopic system,

instead of an HMD. Due to the nature of the shutter-glass system, only a planar viewer position is needed. Unlike a HMD-based system, a viewer head orientation is irrelevant here. This fact has been exploited to simplify a viewer tracking algorithm. The proposed method uses a video capture of a viewer's legs for the estimation of his/her position.

### 2. Problem Formulation

A model of a system setup are shown in Figure 1. From the figure, the desktop-based interactive stereoscopic system is equipped with a single low-cost video camera with a USB interface. A camera points toward the lower part of the viewer, with the floor in front covered most of the picture background. The setup is motivated by the following observations and requirements:



Figure 1. a) The interactive stereoscopic system. b) a snapshot of viewer legs.

- By pointing a camera against the floor, we obtain better control on the picture background as taken by the camera. The setting as proposed here allows efficient background subtraction without complicating system setting.
- It is less intrusive in pointing toward the lower part of the body.
- Such camera setting allows a simple, yet efficient, algorithm based on planar homography to be developed for the estimation of a viewer position.

### 3. Viewer Position Estimation Using Legs Images

With the system setup as shown in Figure 1 (a), video snapshots of the viewer's legs are taken (see Figure 1(b)). From each frame of images, in combination with a background-only snapshot taken during the system setup time, an algorithm for estimating the leg position with respect to the floor, is described below.

- Let  $I_1(x, y)$  be a grayscale video snapshot, and  $I_2(x, y)$  be a background image. Let  $M$  and  $N$  be the number of image's rows and columns respectively. Based on these two images, perform a background subtraction operation. Denote the background subtracted image as  $B(x, y)$ .
- Compute the horizontal centroid for each row according to the following equations

$$r_x = \frac{\sum_{y=1}^N B(x, y) \cdot y}{\sum_{y=1}^N B(x, y)} \quad (1)$$

$$r_{x,i} = \frac{\sum_{y=1}^N B(x, y)}{M} \quad (2)$$

- Based on Eqs (1-2), select rows where the corresponding vertical profiles are greater than  $\alpha T_{x,i}$  ( $\alpha = 0.75$  here). Let  $(x_i, r_{x,i}), i = 1, \dots, Q$ , be coordinates of the centroids corresponding to such  $Q$  selected rows.
- Based on  $Q$  points in Step 3, estimate the parameters ( $m$  and  $b$ ) of a linear equation  $y = mx + b$  by solving a least square equation.
- For each image row ( $x$ ), calculate the horizontal distance between the row's centroid  $r_x$  and the point on the line  $y = mx + b$  obtained from Step 4.
- Starting from the last row of the image upward, find the first image row where the corresponding error is lesser than a pre-defined threshold. Denote the threshold as  $T_{x,2}$ .
- The point so-obtained from Step 6 is used to compute a viewer position measured in the world-plane (see Figure 1 (a)), by using a planar homographic projection. In computing this world-plane position, a homography matrix must first be estimated based on at least four point correspondences [7].

Given a viewer position  $(X, Y)$ , along with a nominal viewer height, the viewer 3-D position is applied to the stereoscopic 3D computer generated module, to deliver a stereo image pair that matches the viewer position.

#### 4. Experiment

In the experiment, a low-cost off-the-shelf USB video camera is used. The camera has a resolution of 640 x 480 pixels, and a frame rate up to 30 frames per second. A computer with Pentium-III 1.13 MHz processor, and 512 MB memory with USB 1.0, was used for the computation.

Actual positions of the ten viewing points used for the experiment, measured in a world coordinate. Except for the four markers used in the initialization stage, these markers were used for the purpose of calculating estimation error for the experiment only.

#### 5. Conclusion

In this paper, a simple vision-based leg tracking algorithm for low-cost application has been described. The method uses a single camera and the simple viewer tracking algorithm. From Figure 2, The algorithm has an averaged error in estimating the viewer leg position of 15% along the X axis, and of 12.5 % along the Y axis. This level of

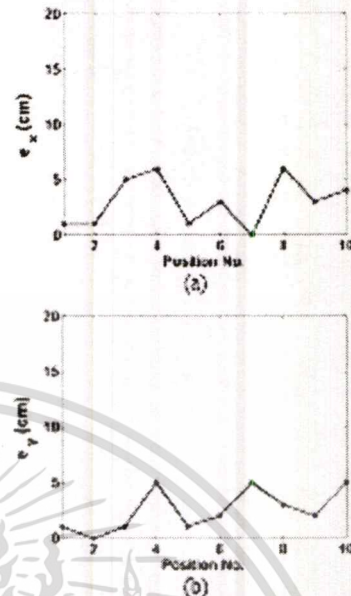


Figure 2. Averaged error of the estimated viewer position at different viewpoints (a) x coordinate (b) y coordinate

accuracy should be acceptable for an interactive stereoscopic viewing application that does not demands a highest degree of virtual realism and interactivity.

#### References

- [1] G. Burdea, and P. Coiffet, *Virtual Reality Technology* (Chapter 2), John Wiley & Sons, Inc., 1994.
- [2] H. Zhou, and E. Hu, "A Survey-Human Movement Tracking and Stroke Rehabilitation", Department of Computer Sciences University of Essex United Kingdom, Location, 2004.
- [3] J.D. Mulder, and R. van Lier, "The Personal Space Station: Bringing Interaction Within Reach", *Proceeding of the Virtual Reality International Conference 2002*, pp. 73-81, 2002.
- [4] J.D. Mulder, J. Jansen, and A. van Rhijn, "An Affordable Optical Head Tracking System for Desktop VR/AR Systems", *Eurographics Workshop on Virtual Environment (2003)*, pp. 215-233, 2003.
- [5] C.C. Chang and W.H. Tsai, "Vision-Based Tracking and Interpretation of Human Leg Movement for Virtual Reality Applications", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.11, No.1, pp.9-24, January 2001.
- [6] M. Ribo, A. Pinz, and A. Fuhrmann, "A new Optical Tracking System for Virtual and Augmented Reality Applications", *Proceedings of the IEEE Instrumentation and Measurement Technical Conference*, pp.1935-1936, May 2001.
- [7] A. Criminisi, I. Reid, and A. Zisserman, "A Plane Measuring Device", *Image and Vision Computing Vol 17* No.8, pp. 625-634, June 1999.

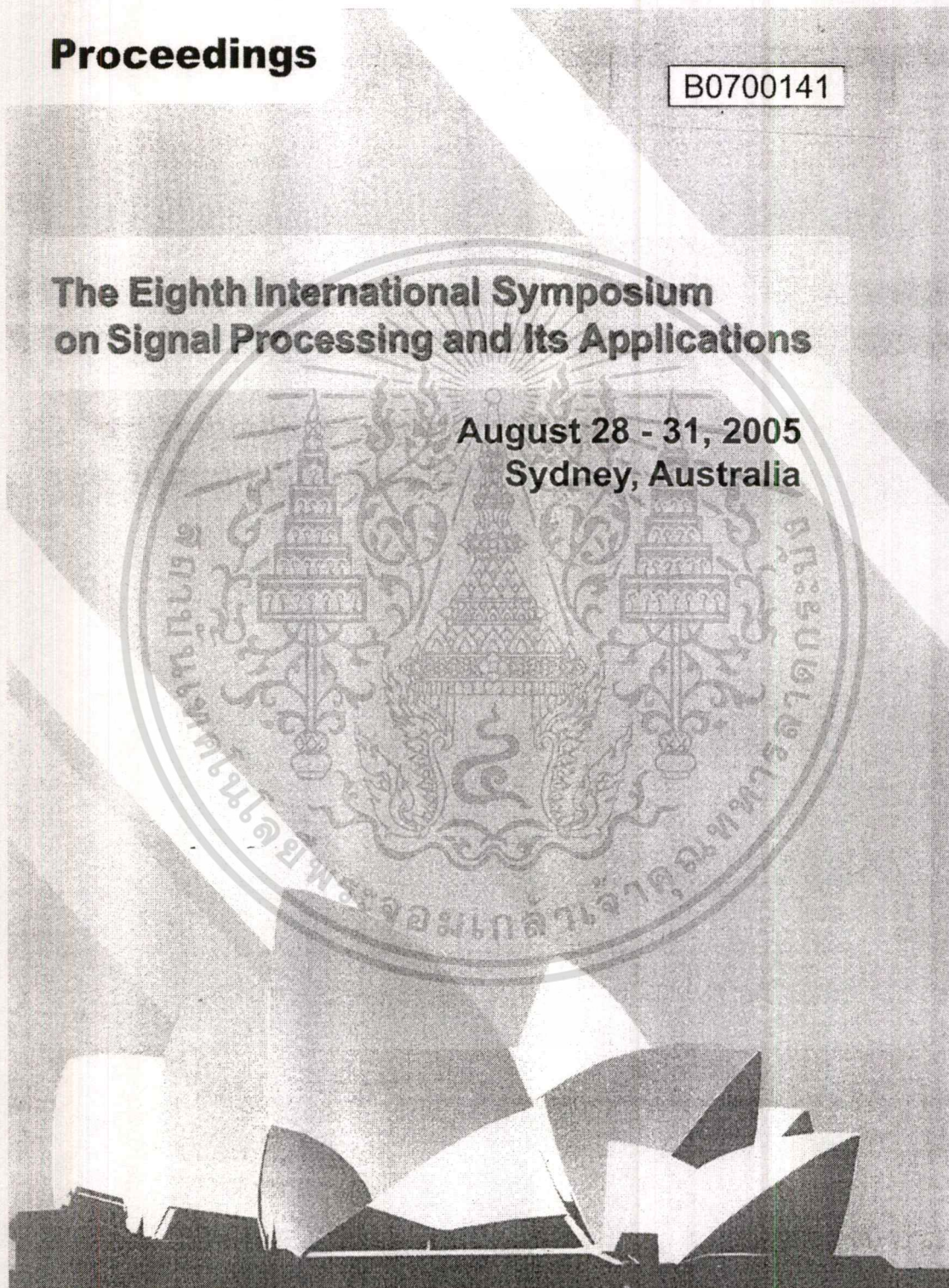
3. Narongsak putpuek and Nopporn Chotikakamthorn. "An Interactive Stereoscopic Viewing system Using Fast Vision-based Led End-point Position Estimation" ISSPA 2005, pp. 631 - 634

## Proceedings

B0700141

### The Eighth International Symposium on Signal Processing and Its Applications

August 28 - 31, 2005  
Sydney, Australia



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## AN INTERACTIVE STEREOSCOPIC VIEWING SYSTEM USING FAST VISION-BASED LEG END-POINT POSITION ESTIMATION

*Narongsak Phuaphuek and Nopporn Chotikakamthorn*

Faculty of Information Technology & Research Center of Communications and Information Technology  
King Mongkut's Institute of Technology Ladkrabang  
Chalongkrong Road, Bangkok, Thailand  
e-mail: s3067144@kmitl.ac.th

### ABSTRACT

In this paper, a problem of interactive stereoscopic display is considered. To provide a stereo image pair with correct viewing angle, a stereoscopic system must be able to identify a viewer spatial location. A fast and low-cost viewer location estimation method using a single video camera is described in this paper. The method tracks a viewer's standing position by estimating his/her leg end-point on the floor. The so-obtained position is then transformed to a 3D coordinate by using a planar homographic projection. Because the estimated viewer position is noisy, a smoothing and semi-interactive updating algorithm is developed to reduce a flicker caused by the imprecise estimate.

### 1. INTRODUCTION

A stereoscopic display system presents the left and right eyes of the viewer with images from different perspective viewpoints, just as the viewer sees the visual world. From these two slightly different views, a brain synthesizes an image of the world with stereoscopic depth. A single, not double, image is seen since the two are fused by the mind into one. Thus, the system can provide an illusion of depth perception just like a human vision experience when viewing a real-world object. Stereoscopic display is used for a variety of applications from games to computer-aided design. However, a standard stereoscopic display system is limited to delivering an image pair perfectly viewed only from a single view point. Changing of viewer's position does not result in perspective change of the viewed scene. Interactive stereoscopic display system provides an answer to the problem. By adaptively adjusting the rendering view point of the synthetic stereoscopic image pair according to the current viewer's position, a high degree of virtual realism is obtained. To obtain the important information about current viewer's spatial position, a head tracking sensor/subsystem may be employed. Tracking of a viewer's head position is achieved by using one of these sensing technologies: electromagnetic sensor, optical sensor, inertial/acoustical or mechanical sensing devices.

Use of these sensing devices is limited to high-end applications such as computer-aided design or VR-based simulation systems. In addition, those sensor-based

technologies places restriction on viewer movement to a certain extent. An alternative solution to those sensor-based methods is a vision-based viewer position estimation.

Vision-based methods use one or more video cameras to collect information required for the estimation. For years, methods for tracking various human body parts have been developed for VR and computer game applications [1]. For our problem, vision-based head tracking and pose estimation like those in [2-3] may be used to obtain a viewer direction with respect to a display unit (provided that the cameras used are placed sufficiently close to the unit). In addition to the methods using a stereo camera like those of [2-3], some other methods have been proposed for head tracking using a single video camera [4-5]. The latter approach is more economical as compared to the stereo camera approach.

In addition to head tracking, methods developed for tracking a human body [6-7] can also be used to obtain a viewer direction with respect to a display unit. However, viewer direction information as obtained by these methods, as well as those head tracking methods just mentioned, is not sufficient for the interactive stereoscopic viewing application. A distance from a display unit is also required so that a stereo image pair with appropriate perspective can be delivered.

In principle, a multi-view image information is required for a three-dimensional positional information to be extracted by a vision-based method. For our problem, however, due to its certain constraints and restrictive requirements, a 3-D viewer positional information can be derived from its projected 2-D positional information. By exploiting those application-specific conditions, this paper describes a method for estimating a viewer position, to be used as a main component of a low-cost interactive stereoscopic viewing system. By targeting at a low-cost desktop system, a single camera is used to reduce complexity. Legs of a viewer are pictured by the camera. Unlike a human leg tracking in [8], where the purpose is to classify the leg shape into one of the predefined classes, here the resulting image is used for the estimation of a viewer position.

Organization of the paper is as follows. First, a brief summary of the system under study is described. Next, the leg position estimation method is detailed. A smoothing and updating algorithm to reduce the problem

caused by estimation inaccuracy is described in Section 4. Experiment results are provided with some discussion in Section 5. A conclusion is given in Section 6.

## 2. PROBLEM FORMULATION

A model of a human and a system setup are shown in Figure 1. From Fig. 1a, a human body is modeled as composed of four sticks. Of the four, at this stage of research, only the bottom two (filled ones) are used for position estimation. The upper two parts will be taken into account in future study. The system under study is shown in Fig. 1b. Here, the desktop-based interactive stereoscopic system is equipped with a single low-cost video camera with a USB interface. A camera points toward the lower part of the viewer, with the floor in front covered most of the picture background (see Fig. 1c). The setup is motivated by the following observations and requirements:

- By pointing a camera against the floor, we obtain better control on the picture background as taken by the camera. In many outdoor as well as ad-hoc indoor settings, with a camera being tilted up in a typical eye-level angle, it is very difficult to control the background behind the viewer. As a result, use of a simple background subtraction method is less efficient under such setting. The setting as proposed here allows efficient background subtraction without complicating system setting.
- It appears to be less intrusive in pointing toward the lower part of the body.
- Such camera setting allows a simple, yet efficient, algorithm based on planar homography to be developed for the estimation of a viewer position ( $x$ - $y$  plane).

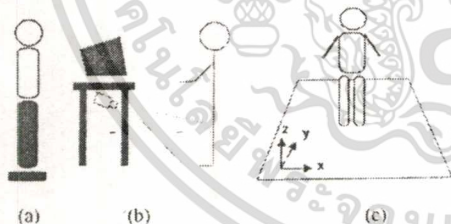


Fig. 1. a) A human body model. b) Camera setting. c) Most of the picture area is filled with anything inside the rectangular area shown in the figure.

## 3. VIEWER POSITION ESTIMATION USING LEGS IMAGES

With the system setup as shown in Figure 1, a sample of the pictures taken is shown in Figure 2. From such

image, combined with a background-only snapshot taken during the system setup time, an algorithm for estimating the viewer position using a legs' end point, is described below.

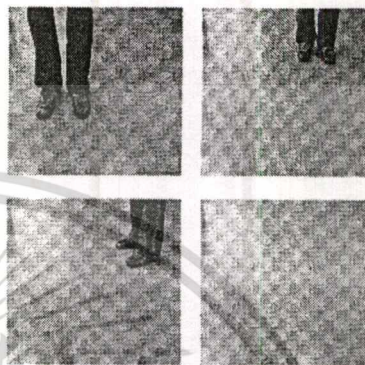


Fig. 2. Samples of pictures taken under the system setting as illustrated in Fig. 1.

### Estimation Algorithm

1. Let  $I(m, n)$  be any current greyscale image, and  $I_b(m, n)$  be a background image, perform a background subtraction operation. Denote the background subtracted image  $I_s(m, n)$ . In addition, let  $M$  and  $N$  be the number of rows and columns respectively.
2. Compute the image profile and moment according to the following equations

$$p_v(m) = \sum_{n=1}^N I_s(m, n) \quad (1)$$

$$p_{mo}(m) = \sum_{n=1}^N I_s(m, n)n / p_v(m) \quad (2)$$

3. Find an average value of  $p_v(m)$ , as denote by  $\bar{p}_v$ . Then, choose  $p_{mo}(m)$  which correspond to  $p_v(m) \geq \bar{p}_v$ . Let  $Q$  be the number of  $p_{mo}(m)$  chosen using this criterion. Then, form a set of points  $q_i, i = 1, \dots, Q$ , of which their coordinates are given by  $(m, p_{mo}(m))$ .
4. Find a straight line that approximates the leg skeleton. To do so, estimate line parameters ( $a.m + b.n = 1$ ) by solving the least squared equation constructed from  $q_i$ .
5. From the line as obtained by Step 4, compute the estimation error between each  $q_i$  and the point evaluated from the so-obtained linear

equation for the same row  $m$ .

6. To get a viewer's standing point, starting from the last image row ( $m = Af$ ), find the first  $m$  that the error as computed by Step 5 is under a pre-defined threshold. Denote that row number  $m_p$ .
7. Use  $m_p$  to compute a viewer position on the x-y plane (see Figure 1c), by using a planar homographic projection. In computing this x-y position, a homography matrix must first be estimated [9]. In practice, this can be obtained by using at least four pre-defined markers placed on the floor within the camera viewing angle. Denote the estimated x-y position as  $c_{xy}$ .

Given  $c_{xy}$ , along with a nominal viewer height, the viewer 3-D position is applied to the stereoscopic 3D computer generated module, to deliver a stereo image pair that matches the viewer position.

#### 4. A SMOOTHING AND UPDATING ALGORITHM FOR SEMI-INTERACTIVE STEREOSCOPIC VIEWING

Because the estimation as detailed in Section 3 is not precise, due to limitation in terms of the image acquisition device's accuracy, as well as the need to reduce algorithm complexity to make it suitable for desktop applications. Therefore, truly real-time interactive system is unachievable with sufficiently low flicker. The system developed here must therefore work in a semi-interactive mode. In particular, change of rendering viewpoint is made based on the following conditions

- A smoothed/interpolated version of  $c_{xy}$  is used for rendering a stereoscopic image pair. This can be obtained by averaging  $c_{xy}$  over a few consecutive frames.
- Updating an image pair when a viewpoint is changed may be performed gradually, by fading out the image pair corresponding to the current rendering viewpoint and fading in the image pair corresponding to the new rendering viewpoint.
- Fading is not applied if the viewer position is sufficiently stationary. Still standing is detected by calculating a standard deviation in the estimate  $c_{xy}$  over a few consecutive video frames.

#### 5. EXPERIMENT

Two experiments were conducted to test the method as described. The first one aimed at measuring the estimation accuracy for the case of still standing. The second one aimed at evaluating the method accuracy when a viewer moves around.

In this experiment, a viewer was asked to stand on 10 different points within a rectangular area of 1.5 m x 1.5 m, in front of a display unit. From a pre-taken background image and those obtained with the viewer present in the scene, the viewer's legs end point was estimated using the method described in Section 3. The result is compared with a correct value, obtained by manual registration. The resulting estimation error, measured in pixels and normalized by the image sizes, is  $1.1\% \pm 1.0\%$  along a horizontal direction, and is  $4.6\% \pm 4.2\%$  along a vertical direction. Pictures obtained from the experiment, overlaid by the estimated leg end points, are shown in Fig. 3. Estimation errors at different viewer positions are shown in Table. 1.

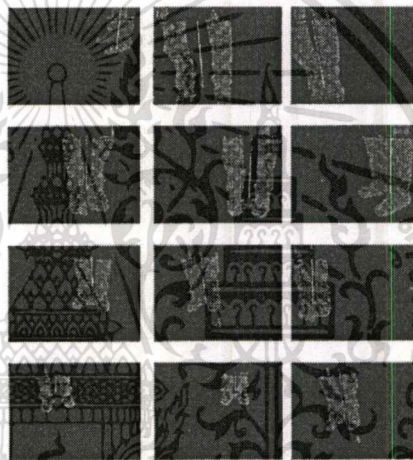


Fig. 3. Parts of the experimental results (the straight line overlaid on the image is that obtained by the algorithm. The line lower end point is the estimate of the legs end-point position).

Table. 1. Estimation errors for different viewer positions (still standing case).

Positions	$y(cm.)$	$x(cm.)$	$e_x(cm.)$	$e_y(cm.)$
1	90	20	1	1
2	90	0	1	0
3	90	-20	5	1
4	130	20	6	5
5	130	0	1	1
6	130	-20	3	2
7	150	0	0	5
8	170	20	6	3
9	170	0	3	2
10	170	-20	4	5

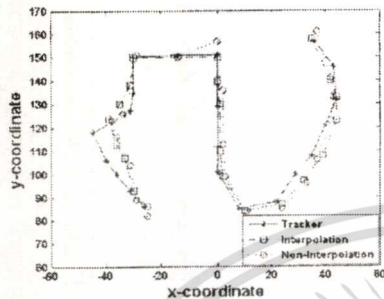


Fig.4. Actual and estimated (with and without interpolation) viewer movement contours.

In the second experiment, a 3-dimensional electromagnetic tracking sensor is placed on a viewer body. The position information as obtained from the device is used as a benchmark against which the position estimated by the proposed method is compared. In the experiment, a viewer was asked to move around within the pre-defined work area in a rather spontaneous manner. A sample of the actual and estimated movement contours is shown in Fig. 4.

## 6. CONCLUSION

In this paper, a method for estimating a viewing position as applied to a desktop-based stereoscopic viewing system has been described. The method is based on the use of a low-cost video camera, with vision-based image processing techniques. From pictures of human legs taken by a camera, an algorithm has been developed to estimate the position of the legs end-point. From this estimated viewing standing point, a planar homographic transform is applied to arrive at a 3D estimate of a viewer position. A smoothing and updating algorithm has also included as an error concealment strategy. Its purpose is to reduce image flickers resulting from a noisy and imprecise estimate.

## 7. REFERENCES

- [1] Chiu-Chun Chang, and Wen-Hsiang Tsai, "Vision-based Tracking and Interpretation of Human Leg Movement for Virtual Reality Applications," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, no. 1, pp. 9-24, Jan 2001.
- [2] Jurriann D. Mulder, Jack Jansen, and Arjen van Rhijn, "An Affordable Optical Head Tracking System for Desktop VR/AR Systems," *Proc. of the workshop on virtual environments*, pp. 215-223, 2003.
- [3] Louis-Philippe Morency, Ali Rahimi, Neal Checka, and Trevor Darrell, "Fast stereo-based head tracking for interactive environments," *Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition*, pp. 375 - 380, 20-21 May 2002.
- [4] Marco La Cascia, Stan Sclaroff, and Vassilis Athitsos, "Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Registration of Texture-Mapped 3D Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 322 - 336, April 2000.
- [5] Lisa M Brown, "3D head using motion adaptive texture-mapping," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 998-1003, Dec 2001.
- [6] Yu Huang, and Thomas S. Huang, "Model-based human body tracking," *Proc. Int. Conf. Pattern Recognition*, vol. 1, pp. 552-555, Aug. 2002.
- [7] Huazhong Ning, Liang Wang, Weining Hu, and Tieniu Tan, "Model-based tracking of human walking in monocular image sequences," *IEEE Region 10 Conf. Computers, Communications, Control and Power Engineering (TENCON)*, vol.1, pp. 537-540, Oct. 2002.
- [8] William Freeman, et al., "Computer Vision for Interactive Computer Graphics," *IEEE Computer Graphics and Applications*, pp. 42-53, May-June 1998.
- [9] Oliver Faugeras, and Quang-Tuan Luong, "The Geometry of Multiple Images," *The MIT Press*, 2001.

