**สารบัญ**

[บทที่ 1 บทนำ 3](#_Toc283977933)

[1.1 แนวคิดของโครงงาน 3](#_Toc283977934)

[1.2 ความสำคัญของโครงงาน 3](#_Toc283977935)

[1.3 ความเป็นมาของโครงงาน 3](#_Toc283977936)

[บทที่ 2 วัตถุประสงค์และเป้าหมาย 5](#_Toc283977937)

[บทที่ 3 ทฤษฎีและหลักการที่เกี่ยวข้อง 6](#_Toc283977938)

[3.1 การปรับปรุงภาพ (Image Enhancement) 6](#_Toc283977939)

[3.2 ปริภูมิสี RGB และปริภูมิสี HSV (HSV Color space and HSV Color space) 7](#_Toc283977940)

[3.3 Morphology 8](#_Toc283977941)

[3.4 Segmentation 9](#_Toc283977942)

[3.5 การทำ Camera calibration ด้วย MATLAB 10](#_Toc283977943)

[3.6 หลักการและการคำนวณที่เกี่ยวข้องกับการทำ Stereo Vision 11](#_Toc283977944)

[3.7 Physic Engine 12](#_Toc283977945)

[3.8 OpenCV 14](#_Toc283977946)

[3.9 OpenGL 14](#_Toc283977947)

[3.10 การปรับเปลี่ยนของผลลัพธ์เนื่องจากการเปลี่ยนระยะห่างระหว่างกล้อง 15](#_Toc283977948)

[บทที่ 4 รายละเอียดของการพัฒนา 16](#_Toc283977949)

[4.1 ภาพรวมของการพัฒนาระบบ 16](#_Toc283977950)

[4.2 Input / Output Specification 16](#_Toc283977951)

[4.3 Functional Specification 17](#_Toc283977952)

[4.4 โครงสร้างของซอฟต์แวร์ 17](#_Toc283977953)

[4.5 ตัวอย่างโปรแกรม 18](#_Toc283977954)

[4.6 ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา 19](#_Toc283977955)

[4.7 เครื่องมือที่ใช้ในการพัฒนา 19](#_Toc283977956)

[4.8 คุณลักษณะของอุปกรณ์ที่ใช้กับโปรแกรม 20](#_Toc283977957)

[บทที่ 5 กลุ่มผู้ใช้โปรแกรม 22](#_Toc283977958)

[บทที่ 6 ผลการทดสอบโปรแกรม 23](#_Toc283977959)

[บทที่ 7 ปัญหาและอุปสรรค 24](#_Toc283977960)

[บทที่ 8 แนวทางการพัฒนาต่อในอนาคต 25](#_Toc283977961)

[บทที่ 9 ข้อสรุปและข้อเสนอแนะ 26](#_Toc283977962)

[บทที่ 10 เอกสารอ้างอิง 27](#_Toc283977963)

บทที่ 1 บทนำ

1. แนวคิดของโครงงาน

ในปัจจุบันเทคโนโลยี Augmented reality กำลังเข้ามามีบทบาทมากขึ้นต่อการดำรงชีวิตของเรา ทั้งจากทาง Smart phone ที่เริ่มมีการนำเทคโนโลยีนี้เข้ามาแก้ไขปัญหาต่างๆ เช่น การผสานข้อมูลสถานที่บนแผนที่และตำแหน่งจาก GPS เข้ากับภาพถ่ายจากกล้องวีดีโอเพื่อบอกทิศทาง, การรู้จำและให้ข้อมูลสถานที่สำคัญโดยการวิเคราะห์จากภาพวิดีโอและตำแหน่งจาก GPS เป็นต้น นอกจากนั้นยังมีการใช้เทคโนโลยีนี้บนอุปกรณ์อื่นๆ เช่น การเติมลูกเล่นบนภาพถ่ายให้สอดคล้องกับใบหน้าคน, การเสริม effect บนเกมส์กระดานเพื่อให้เกมมีความสมจริงและสนุกยิ่งขึ้น เป็นต้น ยิ่งไปกว่านั้นยังมีการผสานวัตถุเสมือนเข้ากับโลกจริงด้วยการวิเคราะห์ภาพจากโลกจริงก่อนแล้วแสดงภาพวัตถุเสมือนซ้อนลงไปเหมือนกับว่ามีวัตถุจริงอยู่ตรงนั้นเช่น เกม Kweekie, Parallel Tracking and Mapping for Small AR Workspaces เป็นต้น

1. ความสำคัญของโครงงาน

โครงงานวิศวกรรมคอมพิวเตอร์ “ระบบวัตถุเสมือน” นี้ถูกจัดทำขึ้น เพื่อทำการศึกษาการควบคุมวัตถุเสมือนด้วยมือในสามมิติผ่านทางกล้องเว็บแคมโดยเน้นการจำลองการปฏิสัมพันธ์ระหว่างมือและวัตถุเสมือนในเชิงกายภาพ คือทำให้เราสามารถจับและเคลื่อนย้ายวัตถุเสมือนในสามมิติได้ ทั้งนี้เพื่อเป็นการเสริมความสมจริงของวัตถุดังกล่าวอันจะทำให้เกิดประโยชน์กับงานด้าน Augmented reality โดยเฉพาะในด้านการนำเทคโนโลยีดังกล่าวมาใช้ในการสร้างระบบวัตถุเสมือนจริงที่ต้องการให้มีปฏิสัมพันธ์กับคนผ่านการจับด้วยมือ เช่น ระบบวัตถุเสมือนสำหรับทำการทดลองทางวิทยาศาสตร์, ระบบวัตถุเสมือนของเล่นตัวต่อเสริมทักษะ เป็นต้น โดยผลสำเร็จจะเป็นโปรแกรมระบบวัตถุเสมือนจริงที่สามารถปรับให้แสดงระบบวัตถุ ใดๆก็ได้ตามกับชุดโมเดลที่โหลดให้กับระบบ

1. ความเป็นมาของโครงงาน

ถึงแม้ว่าจะมีผลงานหลายๆ ชิ้นที่พยายามแสดงวัตถุเสมือนให้สมจริงด้วยการแสดงเป็นภาพสามมิติ และพยายามทำส่วนติดต่อผู้ใช้ให้สามารถใช้ท่ามือในการติดต่อกับระบบผ่านทางกล้องเว็บแคม เช่น Sixth Sense, Hand Gesture Recognition for interactive interfaces, HandVu เป็นต้น แต่อย่างไรก็ตาม เราไม่สามารถใช้มือจับวัตถุที่แสดงนั้นได้ ซึ่งผลงานเหล่านั้นเน้นการวิเคราะห์แยกแยะท่าทางมือเพื่อทำการควบคุมระบบ เน้นการควบคุมระบบในสองมิติ ทำให้ไม่สามารถควบคุมวัตถุที่เป็นสามมิติได้อย่างสมจริงและเป็นธรรมชาติ การพัฒนาการจำลองปฏิสัมพันธ์เชิงกายภาพระหว่างมือและวัตถุเสมือนจึงเป็นสิ่งสำคัญที่จะช่วยให้การควบคุมวัตถุเสมือนเหล่านั้นง่าย และเป็นธรรมชาติมากยิ่งขึ้น อันจะทำให้เกิดการพัฒนา Application ที่หลากหลายยิ่งขึ้นในรูปแบบของ Augmented reality ที่สมจริงมายิ่งขึ้น นอกจากนั้นยังสามารถนำไปประยุกต์ใช้ในการทำการควบคุมคอมพิวเตอร์ และ Virtual reality ได้อีกด้วย

บทที่ 2 วัตถุประสงค์และเป้าหมาย

โครงงานระบบวัตถุเสมือน เป็นระบบจำลองการควบคุมวัตถุในสามมิติ ซึ่งมีวัตถุประสงค์และเป้าหมายของโครงงานดังต่อไปนี้

1. เพื่อทำให้สามารถ จับ, ผลัก และ บิด วัตถุเสมือนใน Augmented reality ด้วยมือได้
2. เพื่อสร้างชุดโมเดลวัตถุเสมือนจริงตัวอย่างทดสอบ 1 ชุด คือชุดของเล่นตัวต่อเสริมทักษะ

บทที่ 3 ทฤษฎีและหลักการที่เกี่ยวข้อง

1. การปรับปรุงภาพ (Image Enhancement)

กระบวนการปรับปรุงคุณภาพของภาพเป็นการปรับปรุงภาพให้เหมาะสมกับการประมวลผลต่างๆ เนื่องจากภาพที่รับเข้ามานั้นอาจมีสัญญาณรบกวน จึงต้องปรับปรุงคุณภาพของภาพก่อนที่จะนำภาพนั้นไปประมวลผลทางคอมพิวเตอร์วิชันต่อไป ซึ่งจะทำให้ประมวลผลได้ดีขึ้น โดยทั่วไปการปรับปรุงภาพทำได้ทั้งใน spatial domain และ frequency domain แต่ในโครงงานนี้ใช้การปรับปรุงภาพใน spatial domain ซึ่งมีจุดประสงค์เพื่อ smooth ภาพ

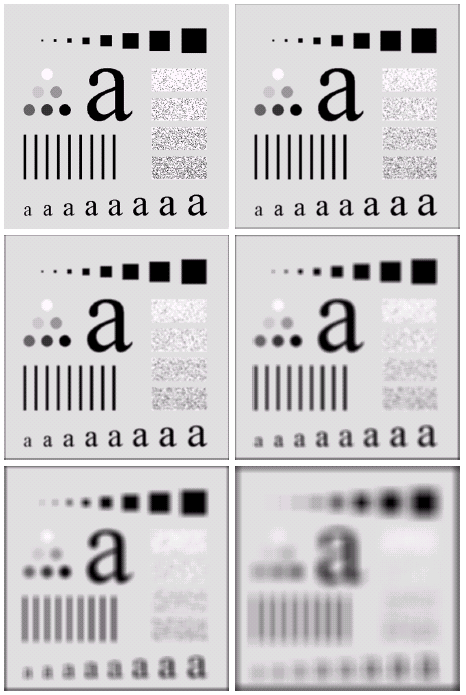
การ smoothing นั้นทำได้โดยใช้การ convolution ซึ่งเป็นการนำ mask กวาดไปบนภาพที่ละจุดภาพจากซ้ายไปขวา และจากบนลงล่าง ตามสมการ

โดยที่ คือ mask ขนาด mxn

คือ input

รูปที่ 1 ตัวอย่าง mask ขนาด 3x3

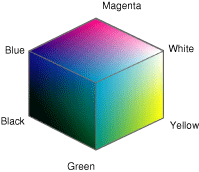
พิจารณารูปข้างล่าง จะพบว่าขนาดของ mask ส่งผลต่อภาพ output เป็นอย่างมาก



รูปที่ 2 การ Smoothing โดยใช้ mean filtering โดยใช้ mask ที่มีขนาดต่างๆ กัน

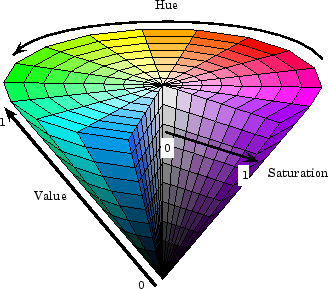
1. ปริภูมิสี RGB และปริภูมิสี HSV (HSV Color space and HSV Color space)

ปริภูมิสี RGB เป็นปริภูมิสีที่ใช้กันอย่างแพร่หลายในระบบภาพดิจิตอลคอมพิวเตอร์ โดยแทนค่าสีของจุดภาพแต่ละจุดภาพด้วยเวกเตอร์สามมิติ ซึ่งแทนค่าสีปฐมภูมิได้แก่ สีแดง, สีเขียว และ สีน้ำเงิน โดยมีช่วงของค่าสีระหว่าง 0-255 ซึ่งปริภูมิสี RGB จะมีลักษณะเป็นรูปลูกบาศก์ ดังรูป



รูปที่ 3 ปริภูมิสี RGB

ปริภูมิสี HSV ใช้หลักการแยกความสว่างออกจากเนื้อสีของจุดภาพ โดยปริภูมิสี HSV นั้นแทนค่าด้วยเวกเตอร์สามมิติ ซึ่งประกออบด้วย H แทนค่าเนื้อสี (Hue), S แทนค่าความอิ่มตัวของสี (Saturation) และ V แทนค่าความสว่างของแสง (Value) ซึ่งปริภูมิสี HSV จะมีลักษณะเป็นรูปกรวย ดังรูป



รูปที่ 4 ปริภูมิสี HSV

การแปลงปริภูมิสี RGB มาเป็นปริภูมิสี HSV ทำได้โดย

*โดยที่*

1. Morphology

Morphology เป็นตัวดำเนินการที่ก่อให้เกิดการเปลี่ยนแปลงรูปร่างของวัตถุในภาพ ซึ่งในที่นี้จะกล่าวถึงการดำเนินการ morphology บนภาพ binary ซึ่งประกอบด้วยตัวดำเนินการ 2 ตัว ได้แก่ ภาพอินพุตที่เป็นไบนารี่ และ mask

การดำเนินการ morphology พื้นฐานประกอบด้วย Dilation, Erosion, Closing, Opening ซึ่งในโครงงานนี้จะใช้การดำเนินการ Opening เพื่อลดสัญญาณรบกวน

* Dilation เป็นการขยายรูปร่างของวัตถุในภาพอินพุต ตามสมการ
* Erosion เป็นการหดรูปร่างของวัตถุในภาพอินพุต ตามสมการ
* Opening ใช้ลบจุดเล็กๆ หรือขอบที่ยื่นออกมาบริเวณขอบของวัตถุในภาพ

1. Segmentation

กระบวนการ Segmentation เป็นกระบวนการแยกวัตถุ หรือองค์ประกอบต่างๆ ออกจากภาพอินพุท โดยในโครงงานนี้ จะใช้สีของ marker ซึ่งมีความเด่นชัดจากสีของพื้นหลังในการ segmentation ซึ่งมีขั้นตอนดังนี้

1. ปรับปรุงคุณภาพของภาพ โดยใช้เทคนิคการเบลอภาพ
2. แปลงปริภูมิสีจากปริภูมิ RGB เป็นปริภูมิ HSV
3. ตั้งค่า threshold ของค่า H, S และ V ให้สอดคล้องกับสีของ marker
4. กำหนดให้ค่าของจุดสีที่เป็นองค์ประกอบที่ต้องการพิจารณามีค่าเป็นสีขาว (255) และกำหนดให้ค่าของจุดสีที่ไม่ใช่องค์ประกอบที่ต้องการพิจารณามีค่าเป็นสีดำ (0)



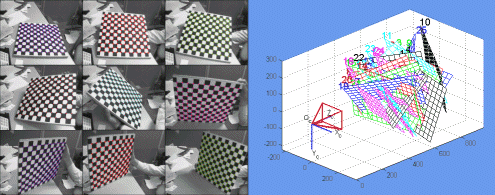


รูปที่ 5 กระบวนการ segmentation เพื่อหาวัตถุด้านซ้าย

1. การทำ Camera calibration ด้วย MATLAB

เป็นการหาคุณสมบัติของระบบกล้อง 2 อย่างได้แก่ สมบัติของกล้องแต่ละตัว คือ Intrinsic matrix ซึ่งเก็บ focal lengths และ center of image และ Distortion matrix สำหรับปรับแก้ความบิดเบี้ยวของภาพ และสมบัติของระบบกล้องคือ การเอียงตัวกันของกล้องทั้งสอง และระยะห่างระหว่างกล้อง

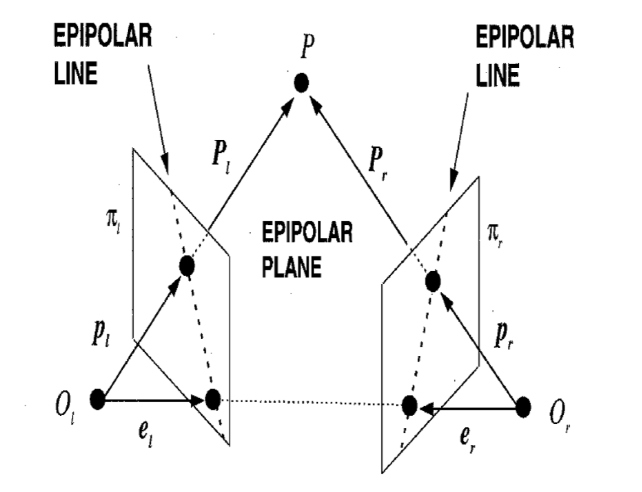
เลือกใช้ Camera calibration toolbox บน MATLAB ของ Jean-Yves Bouguet เพราะเป็น toolbox ที่สามารถปรับหาค่าที่ต้องการได้อย่างละเอียด รู้ค่าคลาดเคลื่อนและมีผู้ใช้กันอย่างแพร่หลาย การทำงานจะใช้ตารางหมากรุกสำหรับการทำ calibration คือทราบขนาดของช่อง และจำนวนช่องที่แน่นอนเป็นตัวอ้างอิง โดยถ่ายภาพในหลายๆ มุมมอง แล้วนำไปหาจุดตัดของช่องตารางหมากรุก แล้วคำนวณขึ้นเป็นคุณสมบัติของระบบกล้องที่ต้องการ



รูปที่ 6 ภาพแสดงการทำ Camera calibration ด้วย MATLAB

1. หลักการและการคำนวณที่เกี่ยวข้องกับการทำ Stereo Vision

* Epipolar geometry



รูปที่ 7 ภาพแสดง epipolar geometry

เป็นเรขาคณิตของระบบ Stereo vision ประกอบไปด้วยส่วนสำคัญต่างๆดังนี้

1. Center of projection (Ol ,Or) จุดรวม Projection line ของแต่ละกล้อง
2. Projection plane () ระนาบของ Image ของแต่ละกล้อง
3. Object point (P) จุดวัตถุ
4. Image points () จุด image ของ P บนแต่ละ Image plane
5. Epipoles () จุด image ของ center of projection ของอีกกล้องบน Image plane
6. Epipolar plane ระนาบระหว่างจุด
7. Epipolar lines เส้นตรงที่ Epipolar plane ตัดกับ Projection planes ของแต่ละกล้อง

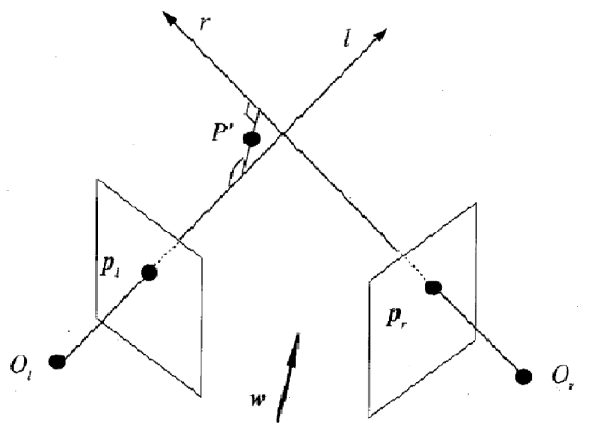
ประโยชน์ของ Epipolar geometry คือช่วยในการแก้ปัญหา correspondence problem และเป็นโมเดลในการคำนวณทางเรขาคณิตของระบบ เช่น การทำ rectification, reconstruction

* Correspondence problem

เนื่องจากโครงการนี้ใช้ marker ติดที่ปลายนิ้วแยกสีเพื่อหานิ้วมือ ทำให้ไม่ต้องแก้ปัญหา correspondence โดยสามารถจับคู่ปลายนิ้วในสองภาพด้วยสีได้เลย

* Reconstruction

ใช้การทำ Triangulation



รูปที่ 8 ภาพแสดงการทำ Triangulation

เป็นการหาตำแหน่งสามมิติของวัตถุโดยตรงจากข้อมูลทางเรขาคณิตของระบบที่มี ทั้งนี้ความถูกต้องของข้อมูลที่ได้จึงขึ้นอยู่กับระดับความถูกต้องของข้อมูลทางกายภาพของระบบที่หาได้

1. Physic Engine

เป็นส่วนที่ต้องตั้งสมมติฐานให้วัตถุในระบบเพิ่มเติม เพราะเนื่องจากวัตถุเสมือนไม่สามารถมีการตอบสนองกลับ (Feedback) มาที่มือได้ดังนั้นเมื่อถูกควบคุมด้วยมือวัตถุจะอยู่ในสภาพคล้าย ไม่มีมวลเพราะไม่สามารถฝืนแรงของมือได้และเกาะติดไปกับนิ้ว โดยแยกการคำนวณออกเป็นกรณีต่างๆ ตามจำนวนนิ้วที่มาชนวัตถุ

ให้จุด A, B, C แทนนิ้วที่ 1, 2, 3 ที่มาชนกับวัตถุ ตามลำดับ โดย A’, B’, C’ จะแทนตำแหน่งถัดไปของนิ้วดังกล่าว, cubePoints แทนจุดของกล่องวัตถุเสมือน และ normal แทนเส้นตั้งฉากระนาบพื้น

* พิจารณาการสัมผัส 1 นิ้ว

A

A’

normal

รูปที่ 9 ภาพแสดงเรขาคณิตของการเคลื่อนวัตถุด้วยหนึ่งนิ้ว

* พิจารณาการสัมผัส 2 นิ้ว

A

B

A’

B’

A

B

A’

B’

B

รูปที่ 10 ภาพแสดงเรขาคณิตของการเคลื่อนวัตถุด้วยสองนิ้ว

1. ย้ายกล่องไปที่ A’
2. ทำการหมุนกล่องรอบจุด A’ตามการกวาดของเวกเตอร์ AB ไป A’B’

* พิจารณาการสัมผัส 3 นิ้ว

A

B

A’

B’

A

B

A’

B’

C

C’

C

C’

รูปที่ 11 ภาพแสดงเรขาคณิตของการเคลื่อนวัตถุด้วยสามนิ้ว

ทำการเคลื่อนที่ในแบบ 2 นิ้วก่อน แล้วเพิ่มการหมุนที่เกิดจาก AC และ A’C’ เฉพาะส่วนองค์ประกอบตามแกน A’B’ รอบจุด A’

ทั้งนี้เลือกเก็บข้อมูลของกล่องด้วย Homogeneous transformation matrix สำหรับแปลงจากคู่อันดับของวัตถุเป็นคู่อันดับของโลกจริง และมิติของกล่องได้แก่ กว้าง, ยาว, ลึก เพื่อความสะดวกในการคำนวณและการทำงานร่วมกับ OpenGL

1. OpenCV

OpenCV เป็น library ที่ใช้สำหรับการประมวลผลภาพ เพื่อนำไปพัฒนาต่อในงานด้านคอมพิวเตอร์วิชัน โดยสามารถประมวลผลภาพได้ทั้งภาพนิ่งและภาพเคลื่อนไหว โดยมีฟังก์ชันสำเร็จรูปสำหรับจัดการข้อมูลภาพ และการประมวลผลภาพพื้นฐานเช่น การหาขอบภาพ การกรองข้อมูลภาพ เป็นต้น

1. OpenGL

OpenGL เป็น library สำหรับแสดง graphic 3D ที่เลือกใช้ในงานนี้ ใช้ในการแสดงภาพสองมุมมอง ในมุมมองแรกคือมุมมองกล้อง ทำการตั้งค่าการ projection ของ OpenGL ให้สอดคล้องกับค่าของกล้องจริง ทำให้ ปลายนิ้วและวัตถุเสมือนที่วาดซ้อนลงบนมุมมองนี้มีความสอดคล้องกับภาพจากกล้อง

มุมมองที่สองคือมุมมองบน มีขึ้นเพื่อให้สามารถทราบระยะลึกของวัตถุได้เพราะไม่สามารถทราบได้จากภาพในมุมมองกล้อง

1. การปรับเปลี่ยนของผลลัพธ์เนื่องจากการเปลี่ยนระยะห่างระหว่างกล้อง

การจัดกล้องในท่ามาตรฐานของระบบคือการจัดให้ภาพที่ได้อยู่ร่วมระนาบกัน ทำให้สามารถประมาณระยะลึกของวัตถุได้จากสมการอย่างง่าย (แม้ว่าจะไม่มีการทำ rectify)

ถ้าจัดระยะห่างของกล้องใหม่ ให้ต่างจากค่าที่กำหนดไว้เดิม T พบว่าระบบจะยังทำงานได้เหมือนเดิมคือยังสามารถเคลื่อนตำแหน่งของวัตถุตามตำแหน่งของ marker เพียงแต่ระบบจะปรับเปลี่ยนขนาดวัตถุเสมือนไปกล่าวคือ ถ้าจัดกล้องให้ห่างกันน้อยกว่า T จะทำให้ความแตกต่างของภาพจากกล้องทั้งสองลดลง กล่าวคือ ค่า d สำหรับวัตถุที่ระยะเดิมลดลง เป็นผลให้ได้ค่า Z มากขึ้น แต่เนื่องจากขนาดของมือในแต่ละภาพยังคงเดิม ระบบจึงเข้าใจว่ามือใหญ่ขึ้น หรือในอีกมุมมองหนึ่งคือขนาดของวัตถุเสมือนเล็กลง และเลื่อนเข้าไปอยู่ใกล้กล้องมากขึ้น ในทางกลับกัน การจัดกล้องให้ห่างกันมากกว่า T จะทำให้เกิดผลในทางตรงกันข้าม คือวัตถุเสมือนใหญ่ขึ้น และอยู่ใกล้กล้องมากยิ่งขึ้น

กล่าวโดยสรุปคือถ้าเป็นการทำงานที่ไม่ต้องการให้ขนาดวัตถุเสมือนที่ตั้งไว้มีค่าเท่าขนาดในโลกจริง สามารถตั้งกล้องที่ความห่างใดๆก็ได้โดยจะยังสามารถควบคุมวัตถุเสมือนด้วย marker ได้ตามเดิม

บทที่ 4 รายละเอียดของการพัฒนา

* 1. ภาพรวมของการพัฒนาระบบ

โครงการวิจัย “ระบบวัตถุเสมือนที่สามารถจับด้วยมือได้” คือการพัฒนาระบบที่สามารถจำลองปฏิสัมพันธ์เชิงกายภาพระหว่างมือกับวัตถุเสมือนใน Augmented reality ได้ โดยสามารถสรุปภาพรวมของการพัฒนาได้ดังต่อไปนี้

* Hand feature extraction

คือการแยกแยะมือออกจากภาพ และหาจุดสำคัญของมือคือ ปลายนิ้ว เพื่อทำการคำนวณพิกัดสามมิติในขั้นต่อไป โดยใช้ปลอกนิ้วเป็น Marker ติดที่ปลายนิ้ว แยกสีตามนิ้ว เพื่อช่วยในการแยกแยะมือออกจากภาพ

* Stereopsis

เป็นวิธีที่ใช้ในการหาตำแหน่งสามมิติของปลายนิ้วที่หาได้จากขั้นก่อนหน้าโดยอาศัยภาพที่ได้จากกล้องสองตัวที่ต่างมุมมองกันในการคำนวณหาความลึก มีประเด็นที่ต้องพิจารณาดังนี้

* + Correspondence

ใช้สีของ marker ในการระบุจุดที่แทนปลายนิ้วเดียวกันในแต่ละภาพที่คู่กัน แทนการsearch corresponding point แบบปกติ

* + Reconstruction

ใช้การทำ Triangulation เพื่อหาพิกัดสามมิติของปลายนิ้ว

* Physic Engine

เป็นส่วนคำนวณความสัมพันธ์ระหว่างพิกัดสามมิติของมือกับวัตถุเสมือน

* Final Application

รวมส่วนประกอบทั้งหมดเข้าด้วยกัน

* 1. Input / Output Specification
* Input: ภาพวีดีโอจากกล้องเว็บแคมจาก 2 มุมมองที่ใกล้เคียงกัน ซึ่งแสดงการเคลื่อนไหวของมือผู้ใช้
* Output: การตอบสนองของระบบวัตถุเสมือนที่สอดคล้องกับมือ
  1. Functional Specification

ข้อกำหนดเชิงหน้าที่ของโครงงานระบบวัตถุเสมือน สามารถอธิบายแยกเป็น 2 ส่วน คือ

* Hand recognition unit:

เป้าหมาย: หาพิกัดได้ถูกต้องและต่อเนื่อง 80% แบบ real-time

* Physic engine unit:

เป้าหมาย: ทำการจำลองความสัมพันธ์ทางกายภาพได้ใกล้เคียงความจริงแบบ real-time

* Graphic display unit:

เป้าหมาย: แสดงภาพได้แบบ real-time

* 1. โครงสร้างของซอฟต์แวร์

โครงงานระบบวัตถุเสมือน ได้แบ่งออกเป็นหน่วยย่อยๆ โดยแต่ละหน่วยมีโครงสร้างการทำงานที่สัมพันธ์กันเป็นลำดับขั้นตอนดังแผนผังลำดับงานด้านล่างนี้

Hand recognition unit

Physic engine unit

Graphic display unit

รูปที่ 12 โครงสร้างของซอฟต์แวร์

โดยแต่ละส่วนมีรายละเอียด ดังนี้

* Hand recognition unit:

หาพิกัดสามมิติของจุดสำคัญบนมือได้จากข้อมูลภาพของกล้องเว็บแคม

* Physic engine unit:

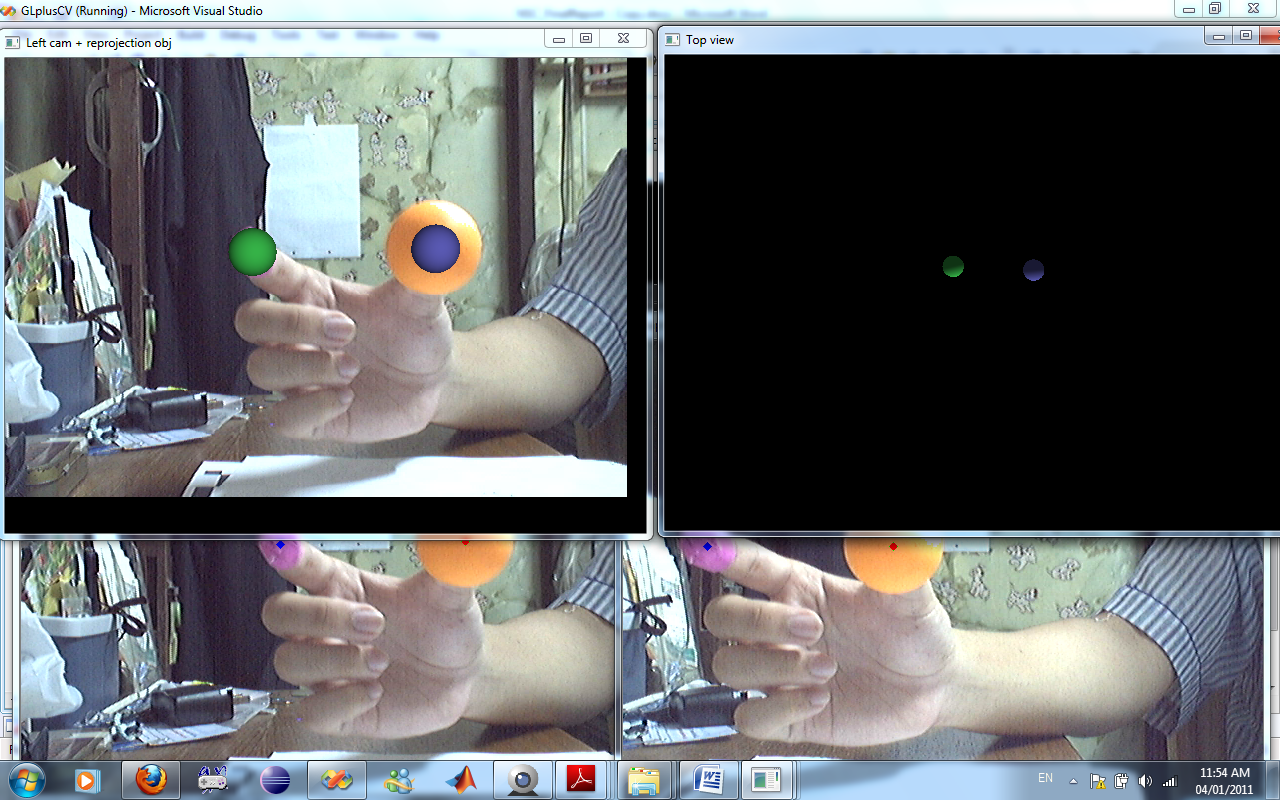
คำนวณปฏิสัมพันธ์เชิงกายภาพระหว่างมือและวัตถุเสมือนได้ใน 3 รูปแบบคือ จับ, ผลัก, บิด

* Graphic display unit:

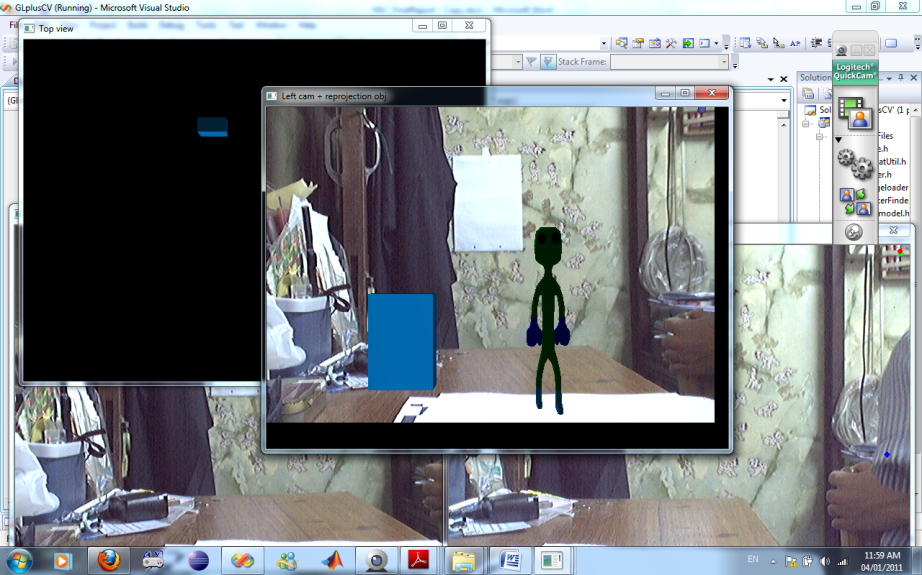
แสดงวัตถุซ้อนลงบนภาพจากเว็บแคมตามผลการคำนวณของ Physic engine ได้ โดยจะแสดงวัตถุเสมือนในกรอบอ้างอิงของกล้องเสมอ (โครงการจึงจะไม่มีการใช้ตัวอ้างอิงตำแหน่งในโลกจริง (marker))

* 1. ตัวอย่างโปรแกรม

โปรแกรมที่พัฒนาขึ้นใช้วัตถุสามมิติแทนวัตถุในโลกจริงและในโลกเสมือนดังนี้

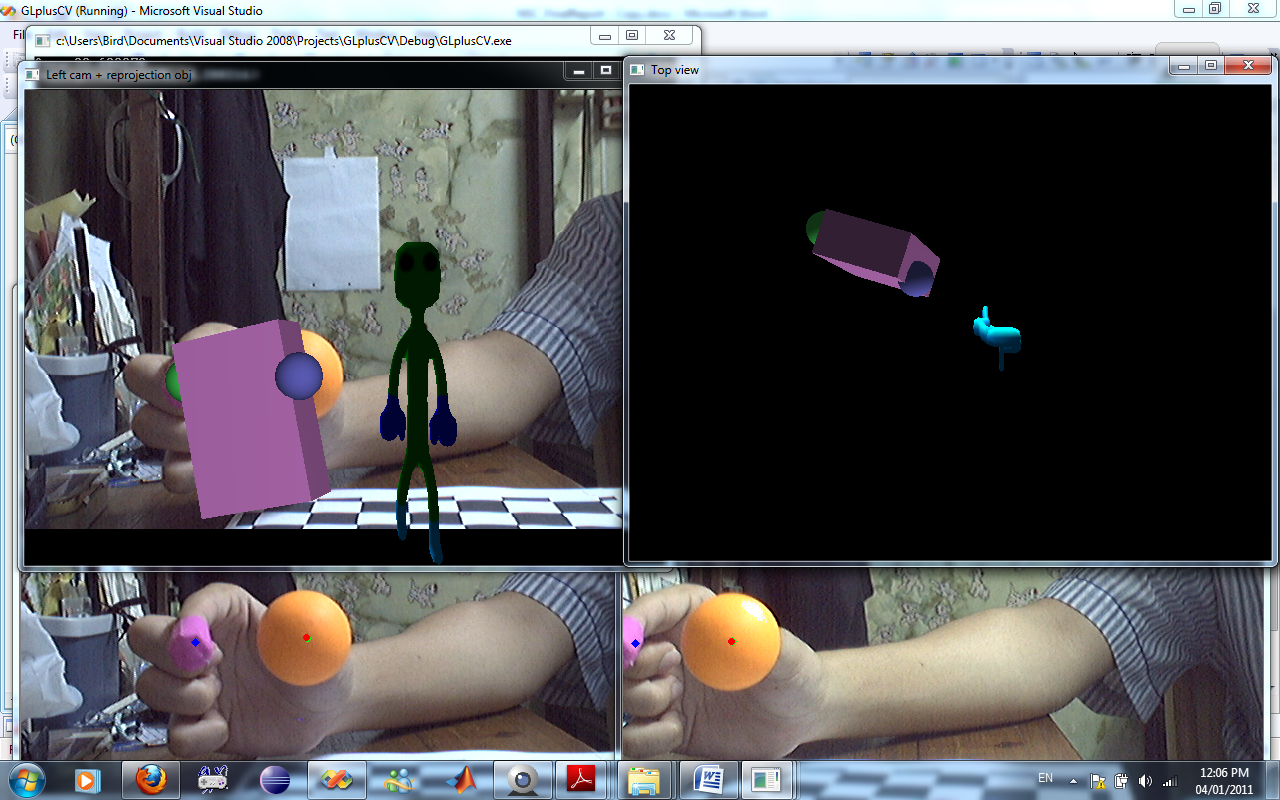


รูปที่ 13 ทรงกลมแทนปลอกนิ้ว (Marker) ที่เจอในระบบ



รูปที่ 14 วัตถุเสมือนรูปสี่เหลี่ยมและ model แทนวัตถุเสมือน

โดยจะมีสองมุมมองให้ใช้สำหรับการควบคุมคือมุมมองกล้องและมุมมองบน



รูปที่ 15 มุมมองกล้องและมุมมอง Top-view

* 1. ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา
* เพื่อพัฒนาระบบวัตถุเสมือนจริงที่สามารถทำการปฏิสัมพันธ์กับมือในเชิงกายภาพในลักษณะของ Augmented reality (คือรับ input ผ่านกล้องและแสดงภาพวัตถุซ้อนลงบนภาพจากกล้อง)
* สร้างชุดโมเดลวัตถุเสมือนจริงตัวอย่างทดสอบ 1 ชุด คือชุดของเล่นตัวต่อเสริมทักษะ
* เนื่องจากใช้เว็บแคมในการพัฒนาจึงต้องการสภาวะแวดล้อมในการใช้งานที่มีแสงเพียงพอ
* Physic engine ที่ทำการพัฒนาขึ้นจะมีข้อจำกัดดังนี้
  + มีการคำนวณความสัมพันธ์เฉพาะระหว่างมือกับวัตถุ จะไม่พิจารณาการชนของกันของวัตถุเสมือน
  + วัตถุเสมือนเป็นรูปลูกบาศก์ (สามารถใช้วัตถุรูปทรงอื่นในลูกบาศก์ใสได้ในลักษณะของ bounding box)

ทั้งนี้เพื่อให้งานมีประเด็นอยู่ในเรื่อง Computer vision เป็นหลัก

* 1. เครื่องมือที่ใช้ในการพัฒนา
* Microsoft Visual Studio 2008

เป็น IDE ที่ใช้ในการพัฒนาโครงงาน โดยใช้ภาษา C++ เป็นหลัก

* OpenCV

เป็น library ใช้สำหรับการประมวลผลด้าน Image Processing และ Computer Vision

* OpenGL

เป็น library ใช้สำหรับแสดงกราฟิกสามมิติ

* GLUT (OpenGL Utility Toolkit)

ใช้ในการควบคุมการทำงานของระบบวินโดวส์ เพื่อให้โปรแกรมที่เราสร้างขึ้นมาสามารถถูกนำไปใช้บน platform อื่นๆ ได้

* GLUI (GLUT-Based User Interface Library)

ไลบรารีตัวเชื่อมสำหรับภาษา C++ บนพื้นฐานของ GLUT ซึ่งจะประกอบด้วยตัวคอนโทรลในลักษณะเดียวกันกับวินโดวส์แบบต่างๆ สำหรับใช้กับ OpenGL

* กล้องเว็บแคมจำนวน 2 ตัว

ใช้สำหรับรับภาพอินพุตจาก 2 มุมมอง

* 1. คุณลักษณะของอุปกรณ์ที่ใช้กับโปรแกรม

ชุดกล้อง Webcam Logitech รุ่น QuickCam Pro 4000(V-UT16) 2 ตัว วางตัวขนานกัน, ห่างจากกัน 3.5” ยึดฐานมั่นคง



รูปที่ 16 ชุดกล้อง Webcam Logitech รุ่น QuickCam Pro 4000(V-UT16)

บทที่ 5 กลุ่มผู้ใช้โปรแกรม

โครงงานวัตถุเสมือน มีกลุ่มเป้าหมายผู้ใช้งานดังนี้

* ผู้ใช้ที่ต้องการจัดการกับวัตถุสามมิติด้วยการเคลื่อนย้ายด้วยมือ ทั้งใน Augmented reality
* ผู้ใช้ที่ต้องการควบคุมคอมพิวเตอร์ด้วยมือ ซึ่งระบบทำให้สามารถควบคุมได้ทั้งสองมิติแบบปกติ และมิติความลึกด้วย
* กลุ่มผู้ใช้ที่สนใจในเทคโนโลยีประเภท Augmented Reality หรือสนใจในผลงานด้าน Computer Vision

บทที่ 6 ผลการทดสอบโปรแกรม

บทที่ 7 ปัญหาและอุปสรรค

ในระหว่างการพัฒนาโครงงาน ผู้พัฒนาพบปัญหาในการดำเนินโครงงาน ดังนี้

1. ปัญหาเรื่องความเชี่ยวชาญในการเขียนโปรแกรมด้วยภาษา C++

เนื่องจากความไม่ชำนาญในภาษาที่ใช้เขียนโปรแกรมในช่วงแรก ๆ จึงทำให้ในแต่ละขั้นตอนการทำงาน จำเป็นต้องตรวจสอบความถูกต้องก่อนเสมอ ทำให้การทำงานเป็นไปอย่างล่าช้า และยากต่อการตรวจสอบ ซึ่งปัญหานี้มีแนวทางการแก้ไขคือ ศึกษาการทำงานของคำสั่งที่ใช้เป็นประจำอย่างละเอียด จะทำให้การทำงานเป็นไปอย่างมีประสิทธิภาพ และรวดเร็วมากยิ่งขึ้น

1. ปัญหาความไม่ชำนาญในการใช้เครื่องมือ OpenCV

เนื่องจากเครื่องมือ OpenCV นั้นเป็นอุปกรณ์ที่มีความซับซ้อนในตัวเองเป็นอย่างมาก ทำให้ต้องใช้เวลาในการค้นคว้าหาคำตอบเพื่อไขปัญหาข้อสงสัยมาก ทั้งจากทางหนังสือและทางอินเตอร์เน็ต เช่นการกลับกันของ RGB เป็น BGR และ การไม่ Release Frame ในแต่ละ loop เป็นต้น ตัวอย่างที่กล่าวข้างต้นทำให้ผลลัพธ์ของระบบไม่เป็นไปตามที่ต้องการ และต้องใช้เวลามากในการค้นคว้าหาสาเหตุ ปัญหานี้มีแนวทางการแก้ไขได้โดย ศึกษาค้นคว้าจากอินเตอร์เน็ต และ ได้รับแนะนำจากอาจารย์ที่ปรึกษาโครงการและรุ่นพี่ในห้องทดลอง

บทที่ 8 แนวทางการพัฒนาต่อในอนาคต

ในแต่ละส่วนที่ได้ทำการพัฒนาขึ้น ผู้พัฒนาได้เห็นถึงแนวทางในการต่อยอดดังนี้

* ส่วนการหาปลายนิ้ว

สามารถใช้ Algorithm ที่ซับซ้อนขึ้นเพื่อที่จะหาปลายนิ้วโดยไม่ต้องอาศัย marker เลยก็ได้

* ส่วนของกล้อง

สามารถเพิ่มความถูกต้องของตำแหน่งปลายนิ้วจากสองกล้อง ให้ละเอียดมากยิ่งขึ้น, สอดคล้องกันเชิงเวลามากยิ่งขึ้น (Synchronized) และมีอัตราการรับภาพที่เร็วยิ่งขึ้น (frame per sec) ได้ด้วยการใช้กล้องที่ดียิ่งขึ้น

* ส่วนของ Physic engine

สามารถเพิ่มเติมความสมจริงด้วยการนำ Physic engine ขนาดใหญ่เช่น Open dynamic engine (ODE) มาใส่การคำนวณเพิ่มเติมเช่น แรงต่างๆ, และการชนกันระหว่างวัตถุเสมือน

* ส่วนแสดงผล
  + สามารถพัฒนา User interface แบบสามมิติเพื่อให้สามารถใช้งานระบบได้หลากหลายมากยิ่งขึ้น
  + สามารถแสดงผลด้วยจอภาพแบบ Head mounted display ซึ่งจะทำให้ระบบมีความสมจริงมากยิ่งขึ้น
  + สามารถพัฒนาให้แสดงผลใน Virtual reality หรือใช้ในงานควบคุมคอมพิวเตอร์แทนได้

ทั้งนี้ทางผู้พัฒนาได้ทำ การพัฒนา API ให้นักพัฒนาคนอื่นๆ สามารถเรียกใช้ และนำไปประยุกต์ใช้กับงานประเภทอื่นๆ ได้ตามที่ต้องการโดยแยกออกเป็น Class ต่างๆอย่างละเอียด แต่อย่างไรก็ตาม การนำ API ไปใช้ จะต้องอาศัยทักษะการเขียนโปรแกรม และการเรียนรู้การใช้งานอยู่บ้างซึ่งสามารถทำได้จากการศึกษา Source code และโปรแกรมตัวอย่าง

บทที่ 9 ข้อสรุปและข้อเสนอแนะ

โดยสรุปแล้ว ระบบสามารถจำลองปฏิสัมพันธ์เชิงกายภาพได้ผลลัพธ์และความถูกต้องเป็นที่น่าพอใจ อย่างไรก็ตาม การใช้ปลอกนิ้วสี เป็น marker ทำให้มีข้อจำกัดเรื่องความเข้มของแสง และสีรบกวนในฉากหลังทำให้เกิดความผิดพลาดขึ้นเป็นบางครั้ง

ทั้งนี้ จึงยังสามารถนำไปพัฒนาต่อยอด เพื่อทำให้ระบบมีความสมบูรณ์ และสามารถประยุกต์ใช้ในงานที่หลากหลายมากยิ่งขึ้น นอกจากนี้ โครงการนี้ยังได้นำความรู้ทางด้าน Digital Image Processing และ Computer Vision มาประยุกต์ใช้เข้าด้วยกัน โดยใช้เทคนิคหลายๆอย่าง จึงหวังว่าจะเป็นประโยชน์ต่อผู้ที่ศึกษาการทำงานของโครงการนี้

บทที่ 10 เอกสารอ้างอิง

[1] Emanuele Trucco and Alessandro Verri, “Introductory Techniques for 3-D Computer Vision”, Prentice Hall, 1998

[2] Gary Bradski and Adrian Kaehler, “Learning OpenCV”, 1st edition, United State of America: O’Reilly, 2008

[3] Wikitude, <http://www.wikitude.org/>

[4] Austrian startup Mobilizy unveils augmented reality driving directions, <http://social.venturebeat.com/2009/08/28/austrian-startup-mobilizy-unveils-augmented-reality-driving-directions/>

[5] Augmented Reality - Layar Reality Browser, [http://www.layar.com](http://www.layar.com/)

[6] Layar, worlds first mobile Augmented Reality browser, <http://www.youtube.com/watch?v=b64_16K2e08>

[7] New YouCam 3 Webcam Software, <http://www.cyberlink.com/products/youcam/overview_en_US.html>

[8] PIT Strategy, <http://www.augmented-reality-games.com/page.asp?pageid=215>

[9] Kweekies, <http://www.int13.net/kweekies-augmented-reality/en/>

[10] Georg Klein and David Murray, “Parallel Tracking and Mapping for Small AR Workspaces”, Active Vision Laboratory, University of Oxford, 2007, <http://www.robots.ox.ac.uk/~gk/publications/KleinMurray2007ISMAR.pdf>

[11] Parallel Tracking and Mapping (PTAM) news, <http://ewokrampage.wordpress.com/>

[12] Pranav Mistry and Pattie Maes and Liyan Chang, “WUW - Wear Ur World - A Wearable Gestural Interface”, MIT Media Lab, <http://fluid.media.mit.edu:4000/assets/_pubs/wip177-mistry.pdf>

[13] C. Keskin, A. Erkan, L. Akarun, “REAL TIME HAND TRACKING AND 3D GESTURE RECOGNITION FOR INTERACTIVE INTERFACES USING HMM”, Boğaziçi University, 2009, <http://cs.nyu.edu/~naz/docs/icann.pdf>

[14] Handvu Hand Gesture Recognition, <http://www.movesinstitute.org/~kolsch/HandVu/HandVu.html>

[15] HandVu for virtual and augmented environments, 2004, <http://www.movesinstitute.org/~kolsch/HandVu/hv_ARVR.wmv>

[16] Le Dung and Makoto Mizukawa, “Fast Hand Feature Extraction Based on Connected Component Labeling, Distance Transform and Hough Transform”, Shibaura Institute of Technology, 2009

[17] Camera Calibration Toolbox for Matlab, <http://www.vision.caltech.edu/bouguetj/calib_doc>

[18] ผศ.ดร.ธนารัตน์ ชลิดาพงศ์, เอกสารประกอบการสอนวิชาคอมพิวเตอร์วิชัน 2/2553