

รหัสโครงการ SUT7-711-51-12-49



รายงานการวิจัย

การบูรณะสถานะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึก
และการไหลเชิงแสง

**3D ENVIRONMENT RECONSTRUCTION BASED-ON DEPTH
MAPS AND OPTICAL FLOW**

ได้รับทุนอุดหนุนการวิจัยจาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว



รายงานการวิจัย

การบูรณะสถานะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึก

และการไหลเชิงแสง

3D ENVIRONMENT RECONSTRUCTION BASED-ON DEPTH MAPS AND OPTICAL FLOW

คณะผู้วิจัย

หัวหน้าโครงการ

รองศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว

สาขาวิชาวิศวกรรมไฟฟ้า

สำนักวิชาวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ผู้ร่วมวิจัย

นางสาวสมหญิง พันธุ์ไฉ

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณ พ.ศ.2551

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

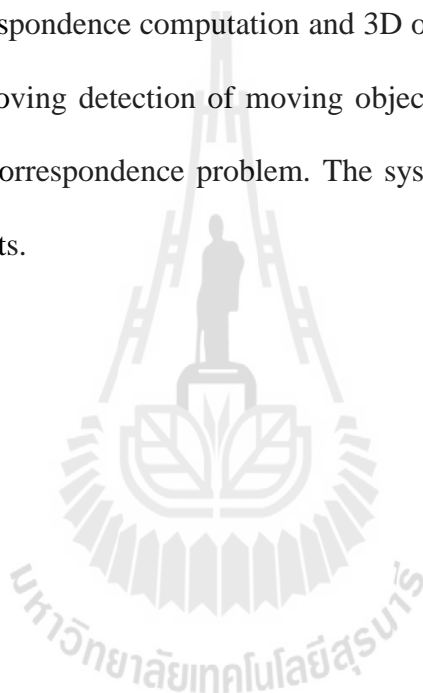
บทคัดย่อ

ในปัจจุบันเทคโนโลยีทางด้านหุ่นยนต์เคลื่อนที่ (mobile robot) เป็นเทคโนโลยีที่สำคัญและได้มีการนำมาประยุกต์ใช้ในชีวิตประจำวันของมนุษย์มากขึ้น ซึ่งได้มีการค้นคว้าและวิจัยอย่างต่อเนื่องเพื่อพัฒนาเทคโนโลยีทางด้านหุ่นยนต์เคลื่อนที่ให้เกิดประสิทธิภาพใกล้เคียงกับมนุษย์มากที่สุด เทคโนโลยีการมองเห็นของคอมพิวเตอร์ (computer vision) เป็นเทคโนโลยีอย่างหนึ่งที่นำมาใช้ในงานหุ่นยนต์เคลื่อนที่อย่างแพร่หลาย ซึ่งส่วนใหญ่ใช้ในการพัฒนาเกี่ยวกับการหาตำแหน่งของวัตถุและสร้างเส้นทางเดินให้กับหุ่นยนต์ โดยการติดกล้องเข้ากับตัวหุ่นยนต์เพื่อแทนระบบการมองเห็นหรือแทนตาของหุ่นยนต์ ดังนั้นระบบการมองเห็นแบบกล้อง 2 ตัว (stereo vision) จึงเป็นเทคโนโลยีที่สำคัญอย่างหนึ่งในการพัฒนาเทคโนโลยีทางด้านหุ่นยนต์ งานวิจัยนี้ได้ดำเนินการศึกษาและพัฒนาการบูรณะสถานะแวดล้อมสามมิติเพื่อสามารถนำไปประยุกต์ใช้ในการตรวจจับวัตรระยะของวัตถุที่เคลื่อนที่แบบอัตโนมัติ ซึ่งระบบที่พัฒนาขึ้นประกอบไปด้วยระบบการตรวจจับการเคลื่อนที่ของวัตถุโดยใช้การไหลเชิงแสง การแก้ปัญหาความสอดคล้องกันของจุดภาพของกล้องสองตัว ด้วยเครือข่ายประสาทเทียมแบบฮอปฟิลด์สองมิติและการจำลองพิกัดสามมิติของวัตถุที่เคลื่อนที่ภายในภาพ โดยระบบที่ได้มีความถูกต้องแม่นยำถึง 90 เปอร์เซ็นต์



Abstract

Now a day, mobile robot technology becomes important and closer to our daily life. Computer vision is one of the technologies which has been mainly developed to locate objects and build paths for a mobile robot using stereo vision. This research presents a 3D environment reconstruction based-on depth maps and optical flow. The system can be applied to automatically measure distance of moving object. The proposed system is composed of motion detection, stereo correspondence computation and 3D object position calculation. The optical flow is used for improving detection of moving object and the Hopfield network is deployed for solving stereo correspondence problem. The system provides desirable results with accuracy up to 90 percents.



กิตติกรรมประกาศ

โครงการวิจัยนี้ได้รับความร่วมมือช่วยเหลือในด้านต่างๆ จากหลายๆ ฝ่าย จนสำเร็จไปได้ด้วยดี คณะผู้วิจัยขอขอบพระคุณศูนย์เครื่องมือวิทยาศาสตร์และเทคโนโลยี ที่ให้ความเอื้อเฟื้อทั้งทางด้านสถานที่ เครื่องมือและบุคลากร ขอขอบพระคุณสาขาวิชาวิศวกรรมไฟฟ้าและสำนักวิศวกรรมศาสตร์สำหรับการสนับสนุนในทุกๆ ด้าน การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณพ.ศ. 2551

คณะผู้วิจัย



สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ฉ
สารบัญรูป.....	ช
บทที่	
1 บทนำ	1
1.1 ความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์การวิจัย.....	2
1.3 ประโยชน์ที่คาดว่าจะได้รับ.....	2
2 ปรัชญ่วรรณกรรมและงานวิจัยที่เกี่ยวข้อง	3
2.1 ระบบการมองเห็นแบบกล้องสองตัว.....	3
2.2 การไหลเชิงแสง.....	6
2.3 เครือข่ายฮอปฟิลด์.....	10
2.4 งานวิจัยที่เกี่ยวข้อง.....	15
3 การแก้ปัญหาคความสอดคล้องกันของจุดภาพโดยใช้เครือข่ายประสาทฮอปฟิลด์ 2 มิติ (2D-Hopfiled Neuron Network)	18
3.1 บทนำ.....	18
3.2 โครงสร้างของเครือข่ายฮอปฟิลด์ 2 มิติ.....	18
3.3 การเรียนรู้ของเครือข่ายฮอปฟิลด์ 2 มิติ.....	19
3.4 การแก้ปัญหาคความสอดคล้องกันของจุดภาพ.....	20
3.5 ผลการทดสอบการแก้ปัญหาคความสอดคล้องกันของจุดภาพ โดยใช้เครือข่ายประสาทเทียมแบบฮอปฟิลด์ 2 มิติ.....	23
4 การบูรณะสถานะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึกเชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง	27
4.1 บทนำ.....	27

สารบัญ (ต่อ)

หน้า

4.2	การคัดแยกวัตถุด้วยการไหลเชิงแสง.....	28
4.3	การแก้ปัญหาความสอดคล้องกันของจุดภาพ.....	30
4.4	การหาระยะห่างระหว่างกล้องกับวัตถุและการบูรณะสภาวะแวดล้อม 3 มิติ.....	35
4.5	ผลการทดสอบการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ.....	40
5	สรุปและข้อเสนอแนะ.....	46
5.1	สรุป.....	46
5.2	ข้อเสนอแนะ.....	47
	รายการอ้างอิง.....	48
	ภาคผนวก	
	ภาคผนวก ก. บทความที่ได้รับการตีพิมพ์เผยแพร่.....	49
	ภาคผนวก ข. ประวัติผู้วิจัย.....	51



สารบัญตาราง

ตารางที่	หน้า
4.1	ค่า w_1 , w_2 และ w_3 ที่เป็นไปได้ตามเงื่อนไข $\Delta A > \Delta D > \Delta a$ 81
4.2	ค่าความผิดพลาดของการจับคู่ภาพของกล้องซ้ายและขวาในชุดตัวอย่างที่ 1-10 82
4.3	ข้อมูลที่น่าสนใจในการคำนวณหาระยะทางจากผลการทดลองในตัวอย่างที่ 1 82
4.4	เปอร์เซ็นต์ความถูกต้องของระยะห่างระหว่างวัตถุกับกล้องที่คำนวณได้กับค่าระยะห่างจริง 82
4.5	ระยะทางที่วัดได้สูงสุดในการติดตั้งกล้องและขนาดภาพต่างๆที่ป้อนให้กับระบบ 82
4.6	ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ 82



สารบัญรูป

รูปที่	หน้า
2.1 การติดตั้งกล้องสองตัวแบบขนาน.....	4
2.2 องค์ประกอบของระบบการมองแบบกล้อง 2 ตัว	6
2.3 ลักษณะทางเรขาคณิตของกล้องสองตัว.....	7
2.4 การไหลของแสง ก) ภาพที่เวลา t ข) ภาพที่เวลา $t+1$ ค) ผลที่ได้จากวิธี Lucas & Kanade.....	8
2.5 โครงสร้างของเครือข่ายซอฟฟีลด์.....	9
2.6 ก) เครือข่ายซอฟฟีลด์ที่มี 3 นิวรอน ข) เครือข่ายซอฟฟีลด์ที่มี 4 นิวรอน.....	11
3.1 โครงสร้างเครือข่ายประสาทและค่าน้ำหนักประสาทของเครือข่ายซอฟฟีลด์ (N. Nasrabadi, C. Choo, 1992).....	12
3.2 C_{ijkl} function (non-linear function).....	12
3.3 แผนผังการทำงานของเครือข่ายซอฟฟีลด์.....	13
3.4 ก) ภาพต้นแบบทางด้านซ้าย ข) ภาพต้นแบบทางด้านขวา ค) ตำแหน่งที่เกิดจากการ ตรวจจับของแฮริส ง) ผลการแก้ปัญหาด้วยเครือข่ายซอฟฟีลด์ 2 มิติ.....	15
3.5 ก) ภาพจากกล้องซ้าย ข) ภาพจากกล้องขวา ค) ตำแหน่งที่เกิดจากการตรวจจับของ Harris ง) ผลการแก้ปัญหาด้วยเครือข่ายซอฟฟีลด์ 2 มิติ.....	16
4.1 โครงสร้างระบบการบูรณะสภาวะแวดล้อม 3 มิติ.....	25
4.2 ชุดข้อมูลลำดับภาพที่ได้จากกล้องซ้าย.....	26
4.3 ชุดข้อมูลลำดับภาพที่ได้จากกล้องขวา.....	27
4.4 ก) การไหลเชิงแสงของภาพที่ได้จากกล้องซ้ายที่ยังไม่ได้กรองขนาด ข) ภาพขยายการ ไหลเชิงแสงของภาพที่ได้จากกล้องซ้ายที่ยังไม่ได้กรองขนาด ค) ภาพขยายการไหลเชิง แสงของภาพที่ได้จากกล้องขวายที่ยังไม่ได้กรองขนาด	27
4.5 การคำนวณหาค่า ΔD และ Δd	27
4.6 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 1.....	27
4.7 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 2.....	27
4.8 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 3.....	27
4.9 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 4.....	27
4.10 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 5.....	27
4.11 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 6.....	27

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.12 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 7.....	27
4.13 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ 3 มิติ.....	27
4.14 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ XY.....	27
4.15 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ ZX.....	27
4.16 ผลการทดสอบการหาระยะวัตถุภายในอาคาร.....	27
4.17 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร.....	27
4.18 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร.....	27
4.19 ผลการทดสอบหาพิสัย 3 มิติของคนเคลื่อนที่ภายในอาคาร.....	27



บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

เทคโนโลยีการมองเห็นของคอมพิวเตอร์ (computer vision) เป็นเทคโนโลยีอย่างหนึ่งที่น่ามาใช้ในการงานหุ่นยนต์เคลื่อนที่อย่างแพร่หลาย ซึ่งส่วนใหญ่ใช้ในการพัฒนาเกี่ยวกับการหาตำแหน่งของวัตถุและสร้างเส้นทางเดินให้กับหุ่นยนต์ โดยการติดตั้งกล้องเข้ากับตัวหุ่นยนต์เพื่อแทนระบบการมองเห็นหรือแทนตาของหุ่นยนต์ ดังนั้นระบบการมองเห็นแบบกล้อง 2 ตัว (stereo vision) จึงเป็นเทคโนโลยีที่สำคัญอย่างหนึ่งในการพัฒนาเทคโนโลยีทางด้านหุ่นยนต์ ระบบการมองเห็นแบบกล้อง 2 ตัวของหุ่นยนต์เคลื่อนที่ ได้มีการค้นคว้าและพัฒนาวิธีการแบบใหม่ขึ้นตลอดเวลาเพื่อให้ระบบมีประสิทธิภาพที่ดีขึ้นและลดข้อจำกัดด้านต่าง ๆ ลง ซึ่งได้รับความสนใจอย่างมาก สำหรับเหตุผลที่ทำให้งานนี้น่าสนใจ คือ ข้อดีที่สำคัญของระบบการมองเห็นแบบกล้อง 2 ตัวอยู่ที่การหาตำแหน่งและระยะห่างของวัตถุเพื่อใช้ในการหลบหลีกและนำทางของหุ่นยนต์โดยปราศจากการทำงานร่วมกับอุปกรณ์เครื่องตรวจวัดต่าง ๆ เช่น เครื่องรับรู้ (sensor) เลเซอร์ เป็นต้น

ระบบการมองเห็นแบบกล้อง 2 ตัวมีระบบการทำงานพื้นฐานคือ การหาแผนที่ความไม่เท่ากัน (disparity maps) หรือแผนที่บ่งบอกความลึก (depth maps) เพื่อหาระยะห่างระหว่างวัตถุกับกล้องและการบูรณะสถานะ 3 มิติ (3D reconstruction) หรือการแปลงพิกัดตำแหน่งสถานะแวดล้อมในภาพให้เป็นพิกัดตำแหน่งของสถานะแวดล้อมจริง หลักการสำคัญในการหาแผนที่บ่งบอกความลึก อยู่ที่การคำนวณหาความสอดคล้องภายในภาพของกล้องทั้งสองตัว (stereo correspondence) การหาความสอดคล้องกันของจุดภาพต้องใช้จุดเด่นของทั้งสองภาพเป็นจุดอ้างอิงซึ่งจุดที่ได้จะกระจายอยู่ทั่วทั้งภาพ ซึ่งบางจุดก็ไม่จำเป็นในการนำมาคำนวณ และยังไม่เหมาะกับภาพนิ่ง ซึ่งงานวิจัยนี้ได้นำมาประยุกต์ใช้กับภาพเคลื่อนไหว ดังนั้นจึงใช้ขนาดของการไหลเชิงแสงเป็นตัวคัดกรองจุดเด่นของภาพแล้วนำไปคำนวณหาความสอดคล้องภายในภาพของกล้องทั้งสองตัวด้วยวิธีการเครือข่ายฮอปฟิลด์สองมิติ (2D Hopfield Network)

1.2 วัตถุประสงค์ของการวิจัย

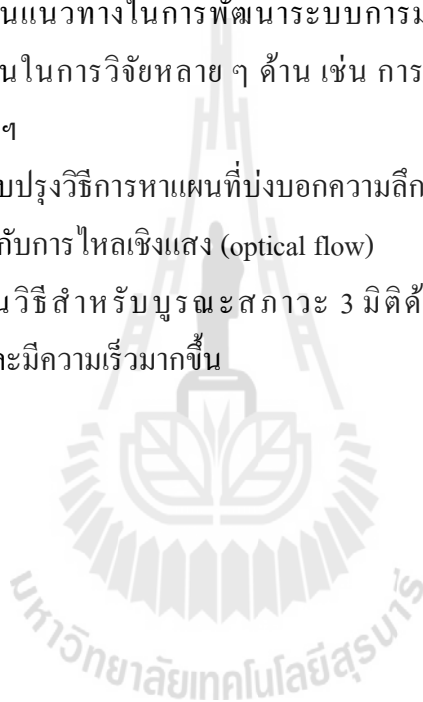
- 1.2.1 เพื่อออกแบบและพัฒนาอัลกอริทึมในการหาความลึกของวัตถุในภาพ
- 1.2.2 เพื่อออกแบบและพัฒนาอัลกอริทึมบูรณะสถานะแวดล้อม 3 มิติได้
- 1.2.3 เพื่อออกแบบและพัฒนากระบวนการสร้างแผนที่แบบ 3 มิติได้

1.3 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1.3.1 สามารถเป็นแนวทางในการพัฒนาระบบการมองเห็นของคอมพิวเตอร์ ซึ่งเป็นเครื่องมือพื้นฐานที่จำเป็นในการวิจัยหลาย ๆ ด้าน เช่น การพัฒนาหุ่นยนต์ การพัฒนาระบบอุตสาหกรรมอัตโนมัติ ฯลฯ

1.3.2 สามารถปรับปรุงวิธีการหาแผนที่บ่งบอกความลึกโดยใช้เครือข่ายประสาทเทียมแบบฮอปฟิลด์ (Hopfield) ร่วมกับการไหลเชิงแสง (optical flow)

1.3.3 ได้ขั้นตอนวิธีสำหรับบูรณะสถานะ 3 มิติด้วยวิธีแผนที่บ่งบอกความลึกมีประสิทธิภาพมากยิ่งขึ้นและมีความเร็วมากขึ้น



บทที่ 2

ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

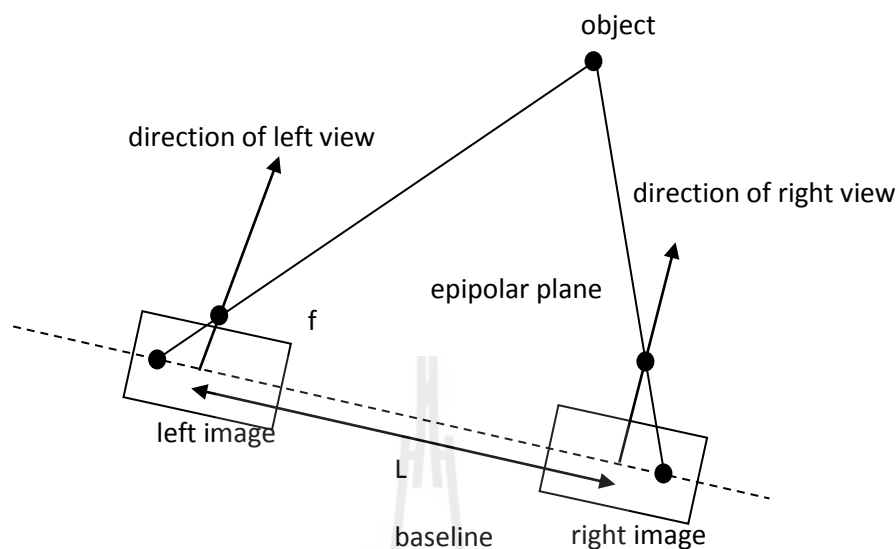
การหาแผนที่ 3 มิติ (3D reconstruction) ต้องอาศัยกระบวนการที่ซับซ้อน โดยเริ่มจากการเก็บข้อมูลภาพจากกล้อง แล้วนำไปคัดกรองภาพเพื่อให้ได้ตำแหน่งจุดเด่นในภาพก่อนที่จะนำไปคำนวณในเครือข่ายประสาทเทียม ซึ่งขั้นตอนเหล่านี้ต้องการการประมวลผลที่รวดเร็วและแม่นยำ รวมทั้งต้องมีการวิเคราะห์เชิงปัญญาประดิษฐ์เพื่อที่จะคำนวณผลลัพธ์ที่ได้ถูกต้องยิ่งขึ้น จากการศึกษาพบว่าในงานวิจัยด้านนี้ต้องมีองค์ความรู้ด้านต่างๆ ทั้งการออกแบบระบบการมองเห็นของกล้อง การประมวลผลภาพ และการคำนวณเชิงเครือข่ายประสาท

2.1 ระบบการมองเห็นแบบกล้องสองตัว

ในปัจจุบันการวิจัยเกี่ยวกับการหาระยะห่างระหว่างกล้องและวัตถุนั้นนิยมใช้ระบบการมองเห็นเห็นของกล้องสองตัว (stereo vision) ซึ่งมีคุณสมบัติในการมองเห็นใกล้เคียงกับกับดวงตาของมนุษย์มาก เพราะใช้กล้องสองตัวจับภาพวัตถุตั้งแสดงในรูปที่ 2.1

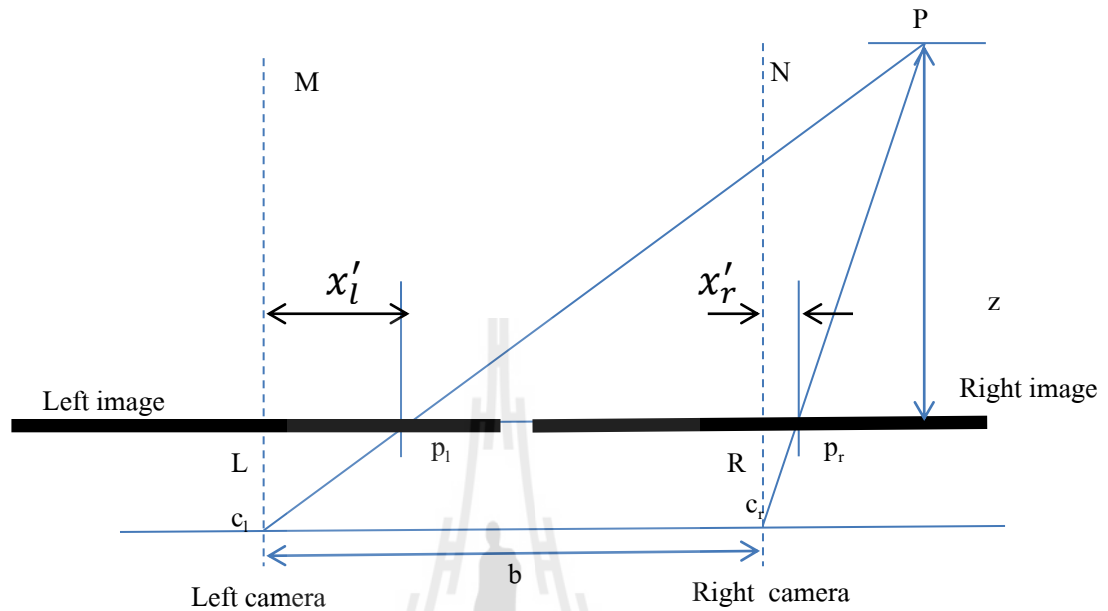


รูปที่ 2.1 การติดตั้งกล้องสองตัวแบบขนาน



รูปที่ 2.2 องค์ประกอบของระบบการมองเห็นแบบกล้อง 2 ตัว

จากรูปที่ 2.2 จะเห็นได้ว่าแบบจำลองพื้นฐานของกล้องทั้งสองนี้จะแยกกันเพียงทิศทางแกน x เท่านั้น โดยจะมีระยะห่างเท่ากับระยะสัญญาณ (base line : b) และระนาบของภาพ (image plane) ของกล้องทั้งสองอยู่ในระนาบเดียวกัน ดังนั้นทำให้จุดที่เราสนใจของภาพจากภายนอกปรากฏลงบนระนาบของกล้องทั้งสองที่ตำแหน่งต่างกัน ซึ่งระยะระหว่างตำแหน่งบนระนาบนั้นเรียกว่าความต่างกัน (disparity) และระนาบที่ผ่านจุดศูนย์กลางของกล้องทั้งสองจุดที่เราสนใจนั้นเรียกว่าระนาบเอพิโพลาร์ (epipolar plane) ส่วนเส้นที่ตัดระนาบเอพิโพลาร์กับระนาบของภาพนั้นเรียกว่าเส้นเอพิโพลาร์ (epipolar line) ซึ่งจากแบบจำลองของรูปที่ 2.2 แสดงนั้น ลักษณะจุดของภาพจากระนาบหนึ่งจะอยู่ในแถวเดียวกันกับอีกระนาบหนึ่งหมายความว่าเส้นระนาบเอพิโพลาร์จะอยู่ที่ตำแหน่งเดียวกันของระนาบภาพของกล้องทั้งสอง แต่ในการกำหนดให้เป็นแบบนี้ได้นั้นแสดงว่าต้องกำหนดให้ความแตกต่างในแนวตั้ง (vertical disparity) มีค่าเท่ากับศูนย์



รูปที่ 2.3 ลักษณะทางเรขาคณิตของกล้องสองตัว

จากรูปที่ 2.3 จุด P จะปรากฏเป็นจุด p_l และ p_r บนระนาบภาพด้านซ้ายและระนาบภาพด้านขวาตามลำดับ และกำหนดให้จุดกำเนิดของพิกัดโลกอยู่ตำแหน่งเดียวกันกับตำแหน่งศูนย์กลางเลนส์ของกล้องด้านซ้าย ดังนั้นเราเปรียบเทียบสามเหลี่ยมคล้าย PMc_l และ p_lLc_r จะได้

$$\frac{x}{z} = \frac{x'_l}{f} \quad (2.1)$$

ในทำนองเดียวกันจากสามเหลี่ยมคล้าย PNc_r และ p_rRc_l จะได้

$$\frac{x-b}{z} = \frac{x'_r}{f} \quad (2.2)$$

ทำการรวมสมการที่ 2.1 และ 2.2 เข้าด้วยกันจะได้

$$Z = \frac{bf}{(x'_l - x'_r)} \quad (2.3)$$

ดังนั้นระยะ Z หรือระยะห่างระหว่างวัตถุกับจุดศูนย์กลางของภาพหาได้จากค่าความแตกต่างของภาพนั่นเอง ซึ่งคือการหาแผนที่ความไม่เท่ากัน (disparity maps)

2.2 การไหลเชิงแสง

การไหลของแสงเป็นวิธีการหนึ่งที่มีประสิทธิภาพในการตรวจจับวัตถุเคลื่อนที่ภายในภาพได้ ซึ่งการคำนวณการไหลของแสงนี้จะใช้การวิเคราะห์จากการเปลี่ยนแปลงของความเข้มแสงภายในภาพแต่ละลำดับภาพ (image sequence) โดยที่การไหลของแสงจะประกอบไปด้วยขนาดและทิศทางของความเข้มแสงที่เปลี่ยนตำแหน่งไปดังรูปที่ 2.4 แสดงตัวอย่างของการไหลของแสงการตรวจจับวัตถุโดยใช้การไหลของแสงจะใช้การวิเคราะห์กลุ่มของความเข้มแสงของวัตถุภายในภาพซึ่งความเข้มแสงของวัตถุที่เปลี่ยนตำแหน่งจะคงความเข้มแสงเท่าเดิมและให้ผลของการไหลของแสงของกลุ่มแสงของวัตถุที่เปลี่ยนแปลงตำแหน่งไปเปลี่ยนทิศทางและขนาดที่ใกล้เคียงกัน

2.2.1 วิธีเชิงเกรเดียนต์ (Gradient Base Method)

กรรมวิธีนี้ใช้วิธีการสร้างการไหลของแสงโดยแก้ปัญหาอนุพันธ์เป็นหลักโดยอาศัยหลักการของความคงที่ของความเข้มแสงจากวัตถุภายในภาพที่เปลี่ยนแปลงไปกำหนดให้ I คือความเข้มแสงสามารถเขียนเป็นสมการได้ดังนี้

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \dots \quad (2.4)$$

จะได้ว่า

$$\nabla I(x, y, t) + I_t = 0 \quad (2.5)$$

สมการที่ 2.5 ได้มาจากการแก้สมการอนุกรมเทย์เลอร์ (Horn B.K.P and Schunck B.G., 1981) จากสมการที่ 2.4 โดย $v = [v_x, v_y] = \left[\frac{dx}{dt}, \frac{dy}{dt} \right]$ และ $\nabla I = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right]$ เรียกสมการที่ 2.5 ว่า optical flow constraint equation (Barron J.L. and Thacker N.A., 1990) ในการพิจารณา ระบบสมการเส้นตรงที่มี 2 ตัวแปร 1 สมการมีแนวทางการแก้ดังนี้

2.2.1.1 Horn & Schunck

Horn & Schunck (1993) ได้รวมเงื่อนไขของความลาดชัน (gradient) จากสมการที่ 2.5 ซึ่งเป็นเทอมของความราบเรียบรวม (global smoothness) เข้ากับเงื่อนไขของการประมาณความเร็ว $v(x, y, t) = (u, v)$ เพื่อพิจารณาหาค่าน้อยที่สุดของ

$$\int_D (\nabla I \cdot v + I_t)^2 + \lambda^2 (\|\nabla u\|_2^2 + \|\nabla v\|_2^2) dx \quad (2.6)$$

โดยที่ขนาด (magnitude) ของ λ เป็นการแสดงให้เห็นถึงผลที่มีต่อเทอมของความราบเรียบ (smoothness) ให้ $\lambda = 0.5$ (Horn B.K.P and Schunck B.G: 1993) จะทำให้ผลลัพธ์ของกรณีที่ทำ การทดสอบได้ดี พิจารณาใช้สมการที่ 2.6 ในการวนซ้ำเพื่อทำการพิจารณาหาค่าน้อยที่สุดและได้ ความเร็วของแสงของแต่ละจุดภาพโดยการวนซ้ำจะเป็นไปตามสมการที่ 2.7 และ 2.8

$$u^{k+1} = u^k - \frac{I_x[I_x u^{-k} + I_y v^{-k} + I_t]}{\alpha^2 + I_x^2 + I_y^2} \quad (2.7)$$

$$v^{k+1} = v^k - \frac{I_y[I_x u^{-k} + I_y v^{-k} + I_t]}{\alpha^2 + I_x^2 + I_y^2} \quad (2.8)$$

โดยที่

- k คือจำนวนครั้งในการวนซ้ำ
- u^0, v^0 คือค่าเริ่มต้นในการประมาณความเร็ว
- u^{-k}, v^{-k} คือค่าของตำแหน่งใกล้เคียงของ u^k, v^k

ในกระบวนการแบบดั้งเดิม (Horn B.K.P and Schunck B.G, 1993) ใช้สมการเชิงอนุพันธ์อันดับหนึ่งในการประมาณค่าความสว่าง ในที่นี้ได้พิจารณาการประมาณจากความสัมพันธ์ของสมการอนุพันธ์เชิงตัวเลขและสามารถพิจารณาจากค่าความคลาดเคลื่อน

2.2.1.2 Lucas & Kanade

จากวิธีการของ Lucas & Kanade (1981) ได้นำเสนอวิธีการในการหาค่า น้ำหนักเหมาะสมของค่ากำลังสองน้อยสุด (LS) ของเงื่อนไขเฉพาะถิ่นอันดับหนึ่ง (local first-order) ของสมการที่ 2.5 เพื่อพิจารณาหาค่าเหมาะสมซึ่งเป็นค่าคงที่ของแบบจำลองสำหรับ v ในแต่ละ จุดภาพโดยพิจารณาหาค่าน้อยที่สุดของสมการที่ 2.9

$$\sum_{x \in \Omega} W^2(x) [\nabla I \cdot v + I_t]^2 \quad (2.9)$$

โดยที่ $W(x)$ คือฟังก์ชันหน้าต่าง (window function) ซึ่งวิธีการแก้ปัญหาของ สมการที่ 2.9 โดย n คือจุดที่เวลา t มีดังนี้

$$A^T W^2 A v = A^T W^2 b \quad (2.10)$$

$$A = [\nabla I(x_1), \dots, \nabla I(x_n)]^T \quad (2.11)$$

$$W = \text{diag}[W(x_1), \dots, W(x_n)] \quad (2.12)$$

$$b = -[I_t(x_1), \dots, I_t(x_n)]^T \quad (2.13)$$

$$v = [A^T W^2 A]^{-1} A^T W^2 b \quad (2.14)$$

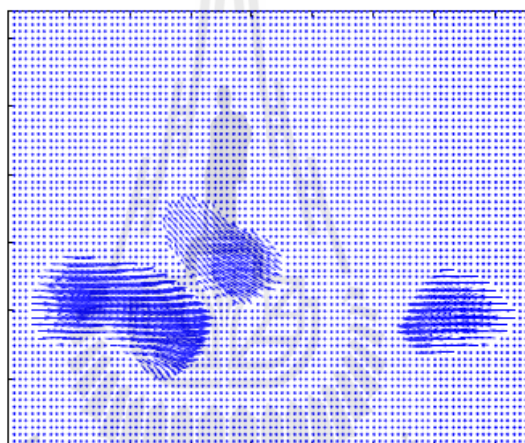
ดังนั้นจะได้ผลลัพธ์ของสมการที่ 2.11 คือสมการที่ 2.14 ซึ่งรูปแบบของผลเฉลยจะเข้าใกล้ รูปแบบของสมการ $A^T W^2 A$ ซึ่งไม่เป็นเอกพจน์ (nonsingular) ที่มีเมทริกซ์ขนาด 2×2

$$A^T W^2 A = \begin{bmatrix} \sum W^2(x) I_x^2(x) & \sum W^2(x) I_x(x) I_y(x) \\ \sum W^2(x) I_y(x) I_x(x) & \sum W^2(x) I_y^2(x) \end{bmatrix} \quad (2.15)$$



(ก)

(ข)



(ค)

รูปที่ 2.4 การไหลของแสง ก) ภาพที่เวลา t ข) ภาพที่เวลา $t+1$ ค) ผลที่ได้จากวิธี Lucas & Kanade

2.2.2 วิธีจับคู่เชิงพื้นที่ (Region - Based Matching)

วิธีการพิจารณาขอบเขตโดยความเหมือนของบริเวณหนึ่งๆ ซึ่งได้นำเสนอโดย Glazer F., Reynolds G., and Anandan P., (1983), Anandan (1989), Burt P. J., Yen C., and Xu X., (1983), Little J. J., Bultho H. H., and Poggio T. A., (1988), Little J. J. and Verri A., (1989) เป็นการอธิบายความเร็ว v ที่ได้ทำการเลื่อน $d = [d_x, d_y]$ เพื่อพิจารณาค่าเหมาะสมที่สุดของขอบเขตของภาพที่เวลาต่างกันจากการค้นหาความเหมือนที่ดีที่สุดโดยทำการพิจารณาที่ค่ามากที่สุดของความเหมือน (บน d) ซึ่งในแต่ละวิธีจะปรับปรุงค่าความเร็วให้แม่นยำที่แตกต่างกันแต่จะใช้หลักการที่มุ่งไปในทางเดียวกันเช่นวิธีการบรรทัดฐานสหสัมพันธ์ไขว้ (normalized cross - correlation) หรือ

พิจารณาค่าน้อยที่สุดของระยะทาง (distance) อาทิเช่นวิธีการผลรวมรากที่สองของผลต่าง (Sum – Squared Difference : SSD) ซึ่งมีรูปแบบสมการดังนี้

$$SSD_{1,2}(x, d) = \sum_{j=-n}^n \sum_{i=-n}^n W(i, j) [I_1(x + (i, j)) - I_2(x + d + (i, j))]^2$$

$$SSD_{1,2}(x, d) = W(x) * [I_1(x) - I_2(x + d)]^2$$

โดยที่ W หมายถึงฟังก์ชันไม่ต่อเนื่องของหน้าต่าง 2 มิติ

d หมายถึงค่าตัวเลขจำนวนเต็มของการเลื่อนตำแหน่ง

การพิจารณาค่าระยะทาง SSD การพิจารณาค่าความเหมือนแบบสหสัมพันธ์ไขว้ (cross - correlation) จะมีลักษณะวิธีการที่ใกล้เคียงกันในการพิจารณาค่าน้อยที่สุดของระยะทาง

การไหลของแสงในแต่ละวิธีข้างต้นผลลัพธ์ที่ได้จะอยู่ในรูปของความเร็วในแต่ละจุดภาพดังแสดงในรูปที่ 2.4 ค) ผลการสร้างการไหลของแสงที่ได้จาก Lucas & Kanade โดยใช้กฎศรแทนขนาดและทิศทางของความเร็วการไหลของแสง

2.3 เครือข่ายฮอปฟิลด์

เครือข่ายฮอปฟิลด์เป็นเครือข่ายอัตโนมัติ (autoassociative network) อย่างหนึ่งซึ่งประกอบไปด้วยนิวรอนชั้นเดียวที่มีการเชื่อมต่อกันอย่างทั่วถึง เครือข่ายฮอปฟิลด์สามารถทำงานได้ในลักษณะเดียวกันกับหน่วยความจำสัมพันธ์ (associative memory) งานของฮอปฟิลด์เป็นการนำเสนอแนวความคิดในรูปแบบผสมผสานระบบความจำในเครือข่ายในรูปของฟังก์ชันพลังงาน (จากสมการ Lyapunov) รวมไปถึงการประมวลผลแต่ละนิวรอนแบบไม่พร้อมเพรียงกัน (asynchronous processing) ข้อแตกต่างนี้ทำให้เครือข่ายฮอปฟิลด์สามารถลดปัญหาการใช้งานกับเครือข่ายที่มีขนาดใหญ่ได้ ข้อดีอีกอย่างหนึ่งก็คือองค์ประกอบในการประมวลผลไม่จำเป็นต้องใช้ข้อมูลความรู้จากเครือข่ายทั้งระบบ

2.3.1 โครงสร้างของเครือข่ายฮอปฟิลด์

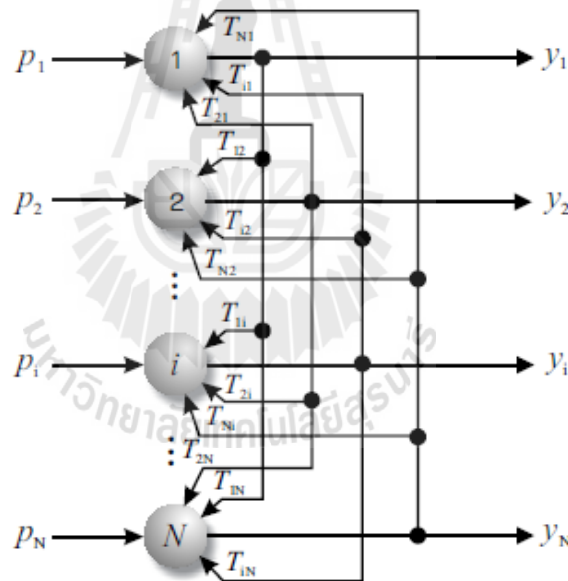
รูปที่ 2.5 แสดงโครงสร้างของเครือข่ายฮอปฟิลด์ สำหรับเครือข่ายแบบไม่ต่อเนื่อง (discrete) จะใช้เพียงค่าไบนารี (0 และ 1) หรือค่าสองขั้ว (1 และ -1) โครงสร้างดังกล่าวประกอบไปด้วย N นิวรอนซึ่งสามารถใช้กับข้อมูลแบบอินพุต p ที่มีขนาด N แต่ละนิวรอนจะถูกพิจารณาในรูป

ของสถานะ S (state) นิรeronแต่ละตัวจะเชื่อมโยงกับนิรeronอื่นๆ ทั้งหมดในระบบยกเว้นตัวมันเอง
 ค่าน้ำหนักประสาทรหว่างนิรeron i และ j คือ T_{ij} มีคุณสมบัติดังนี้

$$T_{ij} = T_{ji} \quad (2.17)$$

$$T_{ii} = 0 \quad (2.18)$$

สมการที่ 2.17 แสดงความสมมาตรของค่าน้ำหนักประสาทรหว่างนิรeron i และ j ในขณะที่
 สมการที่ 2.18 แสดงว่าไม่มีการเชื่อมโยงค่าน้ำหนักประสาทของนิรeronกับตัวมันเอง



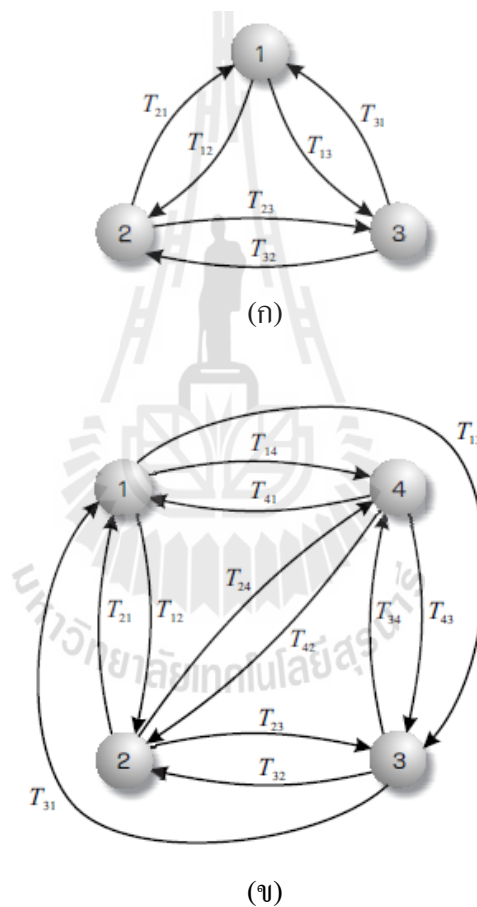
รูปที่ 2.5 โครงสร้างของเครือข่ายฮอปฟิลด์

2.3.2 การเรียนรู้ของเครือข่ายฮอปฟิลด์

สถานะของเครือข่ายฮอปฟิลด์จะขึ้นอยู่กับค่าพลังงานของนิรeronในเครือข่าย การเปลี่ยนแปลงของนิรeronขึ้นอยู่กับเปลี่ยนแปลงของพลังงานดังกล่าวด้วย สถานะของเครือข่ายมีความสัมพันธ์กับพลังงานภายในเครือข่ายดังต่อไปนี้

$$E = -\frac{1}{2} \sum_j \sum_{i, i \neq j} T_{ij} y_i y_j \quad (2.19)$$

โดยที่ T_{ij} คือค่าน้ำหนักประสาทรหว่างนิวรอนที่ i ไปยังนิวรอนที่ j และ y_i เป็นเอาต์พุตของนิวรอนที่ i ภายในเครือข่าย รูปที่ 2.6 แสดงการไหลเวียนของพลังงานในเครือข่ายพร้อมกับการเปลี่ยนสถานะที่เป็นไปได้ของเครือข่ายแบบ 3 และ 4 นิวรอน



รูปที่ 2.6 (ก) เครือข่ายฮอปฟิลด์ที่มี 3 นิวรอน (ข) เครือข่ายฮอปฟิลด์ที่มี 4 นิวรอน

สถานะของเครือข่ายแทนด้วย $S^i(y_1, y_2, \dots, y_N)$ โดยที่ $i = 0, 1, 2, \dots, 2^N$ (จำนวนสถานะที่เป็นไปได้ทั้งหมดของเครือข่ายเท่ากับ 2^N) สังเกตว่าการเปลี่ยนสถานะของเครือข่ายจะเปลี่ยนทีละนิวรอน กล่าวคือในการเปลี่ยนสถานะแต่ละครั้ง องค์ประกอบ y_i เพียงหนึ่งตัวเท่านั้นที่มีการเปลี่ยนสถานะ

เช่นเดียวกันกับเครือข่ายอื่นๆ เครือข่ายฮอปฟิลด์จะมีการวนรอบปรับค่าน้ำหนักประสาท โดยการปรับจะทำการปรับเพียงหนึ่งนิวรอนต่อครั้งจนกระทั่งค่าน้ำหนักประสาทเข้าสู่ค่าตอบที่ต้องการ การลู่เข้าของเครือข่ายจะเกิดขึ้นเมื่อฟังก์ชันพลังงานภายในเครือข่ายให้ค่าน้อยที่สุดและไม่มีนิวรอนใดๆ เกิดการเปลี่ยนสถานะ เครือข่ายฮอปฟิลด์ได้รับการพิสูจน์แล้วว่าสามารถลู่เข้าสถานะที่มีค่าพลังงานของเครือข่ายน้อยที่สุดได้ เมื่อมีการปรับค่าน้ำหนักประสาทอย่างเหมาะสม การเรียนรู้ของเครือข่ายฮอปฟิลด์เป็นไปในรูปแบบของฟังก์ชันพลังงาน เมื่อมีการนำเสนอรูปแบบอินพุตให้กับเครือข่าย เครือข่ายจะทำการปรับค่าน้ำหนักประสาทจนกระทั่งได้สถานะในการจดจำรูปแบบนั้น ๆ ที่ซึ่งฟังก์ชันพลังงานให้ค่าต่ำที่สุด ในระหว่างการเรียนรู้ของเครือข่ายฮอปฟิลด์สามารถพิจารณาสถานะของเครือข่าย y_i ว่าเป็นหน่วยความจำระยะสั้น (short-term memory) ที่ซึ่งสามารถบันทึกและเรียกใช้รูปแบบอินพุตได้โดยการปรับเปลี่ยน y_i ในทำนองเดียวกัน น้ำหนักประสาท T_{ij} สามารถพิจารณาเป็นหน่วยความจำระยะยาว (long-term memory) ซึ่งถูกเก็บค่าไว้ในตัวของเครือข่ายเอง

2.3.3 การปรับค่าน้ำหนักประสาทของเครือข่ายฮอปฟิลด์

พิจารณาอินพุตตัวกระตุ้น S_j (stimulus) ที่ป้อนให้กับนิวรอนตัวที่ j ซึ่งเป็นผลรวมจากเอาต์พุตของนิวรอนตัวอื่นภายในเครือข่ายที่ไม่ใช่ตัวที่ j เองผ่านน้ำหนักประสาท T_{ij} ดังนี้

$$s_j = \sum_{i=1, i \neq j}^N y_i T_{ij} \quad (2.20)$$

เอาต์พุตของนิวรอนตัวที่ j จะมีค่าดังนี้

$$y_i = \begin{cases} 1, & s_j \geq 0 \\ 0, & s_j < 0 \end{cases} \quad (2.21)$$

เนื่องจากในแต่ละรอบ เครือข่ายฮอปฟิลด์มีการปรับค่าน้ำหนักประสาทเพียงนิวรอนเดียว พิจารณานิวรอน j จะได้พลังงานของนิวรอน j นี้คือ

$$E_j = -\frac{1}{2} \sum_{i, i \neq j} T_{ij} y_i y_j \quad (2.22)$$

$$= y_j \left(-\frac{1}{2} \sum_{i, i \neq j} T_{ij} y_i \right) \quad (2.23)$$

จะได้นิรอน j เกิดการเปลี่ยนแปลงค่าจาก y_j^{old} ไปยังค่าใหม่ y_j^{new} ซึ่งทำให้เกิดการเปลี่ยนแปลงของพลังงานคือ ΔE_j ดังนี้

$$\Delta E_j = \Delta E_j^{new} - \Delta E_j^{old} \quad (2.24)$$

จากสมการที่ 2.23 จะได้

$$\Delta E = y_j^{new} \left(-\frac{1}{2} \sum_{i,i \neq j} T_{ij} y_i \right) - y_j^{old} \left(-\frac{1}{2} \sum_{i,i \neq j} T_{ij} y_i \right) \quad (2.25)$$

$$= (y_j^{new} - y_j^{old}) \left(-\frac{1}{2} \sum_{i,i \neq j} T_{ij} y_i \right) \quad (2.26)$$

$$= \Delta y_j \left(-\frac{1}{2} \sum_{i,i \neq j} T_{ij} y_i \right) \quad (2.27)$$

$$= -\frac{1}{2} \Delta y_j s_j \quad (2.28)$$

จากสมการข้างต้น สามารถแบ่งพิจารณาการเปลี่ยนแปลงของพลังงาน ΔE_j ของนิรอนตัวที่ j ได้เป็น 3 กรณีดังนี้

1. ถ้านิรอนตัวที่ j ไม่เปลี่ยนแปลงสถานะ นั่นคือ $\Delta y_j = 0$ ดังนั้นไม่มีการเปลี่ยนแปลงของพลังงาน

2. ถ้านิรอนตัวที่ j เริ่มต้นที่สถานะ 1 และเปลี่ยนสถานะไปเป็น 0 จะได้ว่า

$$y_j^{old} = 1 \quad (2.29)$$

$$y_j^{new} = 0 \quad (2.30)$$

ดังนั้น $\Delta y_j = -1$ ซึ่งมีค่าน้อยกว่า 0 เมื่อพิจารณาจากสมการที่ 2.21 จะได้ว่า y_j จะเปลี่ยนสถานะจาก 1 ไปเป็น 0 เมื่อ $s_j < 0$ ดังนั้นผลคูณ $\Delta y_j s_j$ จึงมีค่ามากกว่า 0 ซึ่งจะสรุปได้ว่า

$$\Delta E_j < 0 \quad (2.31)$$

3. ถ้านิวรอนตัวที่ j เริ่มต้นที่สถานะ 0 และเปลี่ยนไปเป็น 1 จะได้ว่า

$$y_j^{old} = 0 \quad (2.32)$$

$$y_j^{new} = 1 \quad (2.33)$$

ดังนั้น $\Delta y_j = 1$ ซึ่งมีความมากกว่า 0 เมื่อพิจารณาจากสมการที่ (2.21) จะได้ว่า y_j จะเปลี่ยนสถานะจาก 0 ไปเป็น 1 เมื่อ $s_j \geq 0$ ดังนั้นผลคูณ $\Delta y_j s_j$ จึงมีความมากกว่าหรือเท่ากับ 0 ซึ่งจะสรุปได้ว่า

$$\Delta E_j < 0 \quad (2.34)$$

จากการวิเคราะห์ข้างต้น ไม่ว่าจะการเปลี่ยนแปลงสถานะจะเป็นไปในทิศทางใดก็ตาม เครือข่ายฮอปฟิลด์จะให้ผลของค่าพลังงานที่ลดลงหรือคงที่เสมอ เมื่อทุกนิวรอนไม่มีการเปลี่ยนแปลงสถานะ การเปลี่ยนแปลงของพลังงานจะเท่ากับศูนย์หรือพลังงานมีค่าคงที่ ซึ่งเป็นสถานะที่เครือข่ายมีความเสถียรภาพ ณ ค่าพลังงานที่ต่ำที่สุด การประเมินเครือข่ายจะเริ่มจากการสุ่มนิวรอนแล้วทำการปรับค่าน้ำหนักประสาทของเครือข่าย เมื่อทุกนิวรอนได้รับการประเมิน และไม่มีนิวรอนใดเกิดการเปลี่ยนแปลงสถานะแสดงว่าเครือข่ายได้ลู่เข้าแล้ว

2.4 งานวิจัยที่เกี่ยวข้อง

วิธีการหาแผนที่บ่งบอกความลึกจากงานวิจัยต่าง ๆ ที่ผ่านมา (Barnard, Fischler, 1982 และ Dhond, Aggarwal, 1989) ได้ศึกษาเกี่ยวกับการเชื่อมต่อของระบบการมองเห็นของคอมพิวเตอร์ ได้พบว่าปัญหาหลักของระบบการมองเห็นแบบกล้อง 2 ตัวอยู่ที่การหาความสอดคล้องภายในภาพจากทั้ง 2 กล้อง เพื่อให้ได้ความลึกของสถานะแวดล้อมภายในภาพเรียกว่าแผนที่บ่งบอกความลึก ต่อมา มีงานวิจัยมากมายได้พัฒนาวิธีการหาแผนที่บ่งบอกความลึก เริ่มจาก (L. Matthies, T. Kanade, R. Szeliski, 1989) ศึกษาวิธีการหาแผนที่บ่งบอกความลึกด้วยการหาความสอดคล้องของภาพ จากการเปรียบเทียบหาความสัมพันธ์ทั้งหมดของจุดภาพจากกล้องทั้ง 2 ตัว ซึ่งผลการวิจัยสามารถหาแผนที่บ่งบอกความลึกได้แต่มีความผิดพลาดมาก ไม่สามารถหาบริเวณที่ซับซ้อนได้และใช้เวลานาน

เนื่องจากต้องการเปรียบเทียบความสัมพันธ์ทุกจุดภาพ หลังจากนั้น (T. Kanade, M. Okutomi, 1994) ได้พัฒนาประสิทธิภาพของวิธีการหาแผนภาพที่บ่งบอกความลึกด้วยวิธีการหาผลรวมของผลต่างกำลังสอง (sum square difference: SSD) เป็นการสร้างกล่องจุดภาพเพื่อใช้ในการเปรียบเทียบหาความสัมพันธ์ของจุดภาพจากกล่องทั้ง 2 ตัวและลดข้อผิดพลาดจากการใช้วิธีคำนวณทุกจุด ซึ่งจากงานวิจัยนี้สามารถหาแผนที่บ่งบอกความลึกที่มีความชัดเจนได้ดีขึ้น จนในปี 1998 (S. Birchfield, C. Tomasi, 1998) ได้ศึกษาวิจัยแผนที่บ่งบอกความลึกด้วยวิธีจุดภาพต่อจุดภาพ (pixel-to-pixel) โดยการวัดค่าความเข้มของแต่ละจุดภาพเปรียบเทียบกับกัน ซึ่งผลที่ได้คือสามารถหาแผนที่บ่งบอกความลึกที่มีความชัดเจนขึ้นและหาบริเวณที่ซับซ้อนได้โดยใช้เวลาน้อยลง แต่ปัญหาที่เกิดขึ้นกับวิธีการนี้คืออัลกอริทึมในการคำนวณยุ่งยากและเหมาะสมกับงานเพียงบางประเภท ต่อมา (Y. Boykov, O. Veksler, and R. Zabih, 2001) ทำการวิจัยหาแผนที่บ่งบอกความลึกด้วยวิธีกราฟคัท (graph cut) โดยอัลกอริทึมวิธีกราฟคัทคำนวณหาความสัมพันธ์ของจุดภาพ ซึ่งแผนที่บ่งบอกความลึกที่หาขึ้นมีความชัดเจนสามารถแสดงรายละเอียดของวัตถุได้ แต่ปัญหาที่เกิดขึ้นคืออัลกอริทึมในการคำนวณยุ่งยากและใช้เวลานานเหมาะสมกับงานบางประเภท

จากผลงานวิจัยที่ผ่านมาจะเห็นได้ว่าการคำนวณหาค่าความสอดคล้องกันของจุดภาพนั้นมีอัลกอริทึมที่ซับซ้อน (N. Nasrabadi, C. Choo, 1992) ได้นำหลักการของเครือข่ายฮอปฟิลด์มาใช้ศึกษาหาความสอดคล้องกันภายในภาพ โดยนำจุดเด่นของทั้ง 2 ภาพ เข้าไปคำนวณในเครือข่ายฮอปฟิลด์ 2 มิติ (2D Hopfield Neuron Network) เพื่อลดการคำนวณทุกจุดภาพลง ซึ่งงานวิจัยนี้สามารถหาแผนที่บ่งบอกความลึกได้ดีขึ้น มีความผิดพลาดเล็กน้อยและใช้เวลาในการคำนวณน้อยลง ลดข้อจำกัดของการติดตั้งกล้อง

หลังจากทำการหาแผนที่บ่งบอกความลึกแล้ว ขั้นตอนต่อไปคือการบูรณะสภาวะ 3 มิติจากแผนที่บ่งบอกความลึก การบูรณะสภาวะ 3 มิติคือการแปลงข้อมูลจากแผนที่บ่งบอกความลึกด้วยการคำนวณเชิงเลขาคณิต (triangular) เปลี่ยนเป็นข้อมูล 3 มิติหรือพิกัดจริง โดยที่แต่ละจุดภาพในแผนที่บ่งบอกความลึกแสดงถึงข้อมูลค่าความไม่เท่ากันของจุดภาพ (d) ซึ่งเป็นตัวแปรที่สำคัญที่ใช้บอกความลึกของภาพที่ใช้ในการบูรณะสภาวะ 3 มิตินอกจากนั้นระยะความยาวโฟกัสของกล้อง (f)

และระยะห่างของการตั้งกล้องแบบขนานกันระหว่างกล้อง 2 ตัว (L) ก็เป็นตัวแปรสำคัญสำหรับการคำนวณหาสถานะ 3 มิติหรือพิกัดจริงจากแผนที่บ่งบอกความลึก

สำหรับผลการบูรณะสถานะ 3 มิติด้วยวิธีการนี้ (H. Kim, S. Yang, K. Sohn, 2003) กล่าวว่า การบูรณะสถานะ 3 มิติด้วยวิธีการแปลงเชิงเรขาคณิต สามารถบูรณะสถานะ 3 มิติจากแผนที่บ่งบอกความลึกได้ ซึ่งสามารถแยกระยะห่างของตำแหน่งวัตถุออกจากพื้นหลังได้ แต่ไม่สามารถบอกรายละเอียดที่ซับซ้อนของวัตถุหรือระยะห่างของวัตถุที่วางใกล้กันมากๆ ส่วนถ้าต้องการสถานะ 3 มิติที่ต้องการรายละเอียดที่ซับซ้อน เช่น การสร้างหน้าคนหรือการค้นหาวัดดู (H. Kim, S. Yang, K. Sohn, 2003) ได้เสนอให้ใช้วิธีการติดตั้งกล้องแบบหลายตัวแทนการติดตั้งกล้องเพียง 2 ตัว ส่วนการบูรณะสถานะ 3 มิติด้วยวิธีการอื่น ๆ นอกจากวิธีการนี้เริ่มจาก (M. Pollefeys, L. Gool, 1997) ศึกษาการบูรณะสถานะ 3 มิติด้วยการเปลี่ยนแนวการวางตัวของกล้องทั้ง 2 ตัวเพื่อให้เกิดมุมมองอื่น ๆ นั่นคือการเปลี่ยนมุมการวางกล้องกับวัตถุด้วยการขยับครึ่งละ 5 องศารอบวัตถุเพื่อให้เห็นทุกด้านของวัตถุ ผลการวิจัยที่ได้มีความชัดเจนสามารถสร้างบริเวณที่ซับซ้อนได้ แต่ไม่เหมาะสมกับงานวิจัยที่ต้องการยึดตัวกล้องอยู่กับที่ต่อมา (C. Strecha, L. Gool, 2002) สามารถบูรณะสถานะ 3 มิติ โดยใช้แผนที่บ่งบอกความลึกเพียง 2 ชุดจากกล้อง 3 ตัวเพื่อลดข้อจำกัดเรื่องการขยับกล้อง ผลที่ได้นั้นสามารถหารายละเอียดที่ซับซ้อนได้มากขึ้น ส่วน (H. Kim, S. Yang, K. Sohn, 2003) ใช้กล้อง 2 ตัวบูรณะสถานะ 3 มิติด้วยวิธีการแปลงเชิงเรขาคณิต ซึ่งสามารถบูรณะสถานะ 3 มิติได้รายละเอียดเพียงแยกตำแหน่งของวัตถุออกจากพื้นหลัง แต่ไม่สามารถบ่งบอกรายละเอียดของวัตถุได้ ดังนั้นได้มีการแก้ไขสถานะในการมองเห็นที่จำกัดนี้โดย (S. Kagami, Y. Takaoka, Y. Kida, K. Nishiwaki, and T. Kanade, 2005) ทำการวิจัยสร้างเส้นทางเดินให้กับหุ่นยนต์โดยการสร้างแนวทิศทางการเคลื่อนที่ของหุ่นยนต์เพื่อใช้เป็นแนวทางการเคลื่อนที่ด้วยการหาความสัมพันธ์จากภาพที่มองเห็นแบบลำดับต่อเนื่อง (image sequence) ด้วยการสร้างสถานะ 3 มิติของแต่ละภาพมาเชื่อมต่อกัน ซึ่งผลการวิจัยที่ได้สามารถแก้ไขสถานะการมองเห็นของหุ่นยนต์ที่มีจำกัดได้

บทที่ 3

การแก้ปัญหาความสอดคล้องกันของจุดภาพโดยใช้

เครือข่ายประสาทฮอปฟิลด์ 2 มิติ (2D Hopfield Neuron Network)

3.1 กล่าวนำ

ปัญหาความสอดคล้องกันของจุดภาพนั้น คือปัญหาสำคัญของการบูรณะแวลด้อม 3 มิติ จากงานวิจัยพบว่าเครือข่ายฮอปฟิลด์ 2 มิติ เป็นอีกวิธีการหนึ่งที่ใช้ในการแก้ปัญหานี้ โดยมีเครือข่ายฮอปฟิลด์เป็นต้นแบบ เครือข่ายฮอปฟิลด์ 2 มิติ มีชุดข้อมูลตำแหน่งที่น่าสนใจของกล้องซ้ายและขวาเป็นอินพุตให้กับระบบ หลังจากนั้นเครือข่ายฮอปฟิลด์จะทำการวนรอบเพื่อหาคำตอบต่อไป ซึ่งในบทนี้ชุดข้อมูลที่นำมาวิจัย เป็นภาพนิ่งที่เกิดจากกล้องสองตัว เพื่อตรวจสอบความแม่นยำของระบบเครือข่ายฮอปฟิลด์ 2 มิติ

3.2 โครงสร้างของเครือข่ายฮอปฟิลด์ 2 มิติ

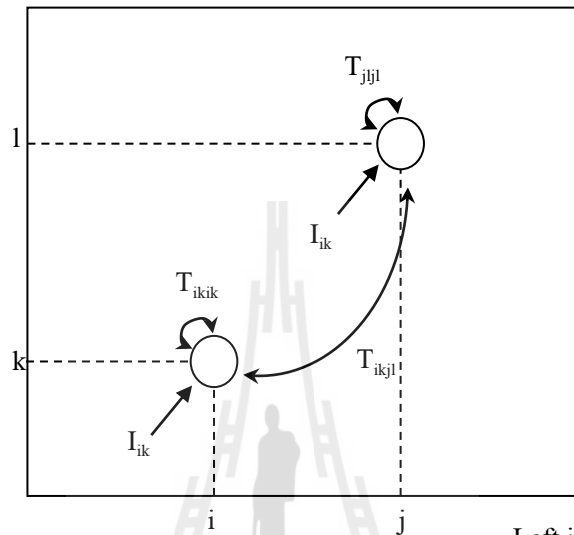
โครงสร้างของเครือข่ายฮอปฟิลด์ 2 มิติ ประกอบไปด้วย $N_r \times N_c$ นิวรอน โดยที่ N_r และ N_c คือจำนวนจุดเด่นของภาพทางด้านซ้ายและด้านขวาตามลำดับ สถานะของนิวรอน (state) ในเครือข่ายจะแสดงถึงความเป็นไปได้ของความสอดคล้องกันภายในจุดภาพของภาพทั้ง 2 กล้อง นิวรอนจะเชื่อมโยงกับนิวรอนอื่น ๆ ในเครือข่ายทั้งหมดยกเว้นตัวมันเอง โดยมีค่าน้ำหนักประสาทรหว่างนิวรอนคือ

$$T_{ikjl} = T_{jlik} \quad (3.1)$$

$$T_{ikik} = 0 \quad (3.2)$$

สมการที่ 3.1 แสดงถึงความสมมาตรของค่าน้ำหนักประสาทรหว่างนิวรอน ik และ jl ในขณะที่สมการที่ 3.2 แสดงว่าไม่มีการเชื่อมโยงค่าน้ำหนักประสาทของนิวรอนกับตัวมันเอง

Right image Feature Points



Left image Feature Points

รูปที่ 3.1 โครงสร้างเครือข่ายประสาทและค่าน้ำหนักประสาทของเครือข่ายฮอปฟิลด์ (N. Nasrabadi, C. Choo, 1992)

3.3 การเรียนรู้ของเครือข่ายฮอปฟิลด์ 2 มิติ

การเปลี่ยนแปลงสถานะของเครือข่ายฮอปฟิลด์ 2 มิติจะขึ้นอยู่กับค่าพลังงานของนิวรอนในเครือข่าย สถานะของเครือข่ายมีความสัมพันธ์กับพลังงานภายในเครือข่ายดังต่อไปนี้

$$E = -(1/2) \sum_{i=1}^{N_i} \sum_{k=1}^{N_r} \sum_{j=1}^{N_l} \sum_{l=1}^{N_r} T_{ikjl} V_{ik} V_{jl} - \sum_{i=1}^{N_i} \sum_{k=1}^{N_r} I_{ik} V_{ik} \quad (3.3)$$

โดยที่ V_{ik} และ V_{jl} คือสถานะของนิวรอน ik และ jl ตามลำดับ

T_{ikjl} คือ ค่าน้ำหนักประสาทของสองนิวรอน

I_{ik} คือ ค่าอินพุตเริ่มต้น (initial input) ของแต่ละนิวรอน

จากสมการที่ 3.3 จะเห็นได้ว่า เมื่อมีการป้อนกลับของนิวรอนที่ $T_{ikik} = 0$ และ I_{ik} คือค่าอินพุตเริ่มต้น จะได้ว่า เมื่อนิวรอนที่ ik เปลี่ยนสถานะจะทำให้มีการเปลี่ยนแปลงค่าพลังงานด้วย ดังสมการ

$$\Delta E_{ik} = [\sum_{i=1}^{N_l} \sum_{k=1}^{N_r} T_{ikjl} V_{jl} + I_{ik}] \Delta V_{ik} \quad (3.4)$$

การปรับสถานะของนิวรอนคำนวณได้จาก

$$V_{ik} = 0 \text{ if } [\sum_{i=1}^{N_l} \sum_{k=1}^{N_r} T_{ikjl} V_{jl} + I_{ik}] > 0 \quad (3.5)$$

$$V_{ik} = 1 \text{ if } [\sum_{i=1}^{N_l} \sum_{k=1}^{N_r} T_{ikjl} V_{jl} + I_{ik}] < 0 \quad (3.6)$$

$$\text{no chang if } [\sum_{i=1}^{N_l} \sum_{k=1}^{N_r} T_{ikjl} V_{jl} + I_{ik}] = 0 \quad (3.7)$$

จากสมการสถานะของนิวรอนที่ ik จะเปลี่ยนสถานะก็ต่อเมื่อผลรวมของค่าน้ำหนักประสาทสถานะในนิวรอน jl และค่าอินพุตเริ่มต้นเมื่อผลรวมมากกว่า 0 จะทำให้สถานะนิวรอนที่ ik เปลี่ยนเป็น 1 เมื่อผลรวมน้อยกว่า 0 จะเปลี่ยนเป็น 0 และสถานะที่ ik จะไม่เปลี่ยนแปลงเมื่อผลรวมเท่ากับ 0

3.4 การแก้ปัญหาคงความสอดคล้องกันของจุดภาพ

การแก้ปัญหาคงความสอดคล้องกันของจุดภาพจะใช้ค่าพลังงานของเครือข่ายประสาทเทียม ดังสมการที่ 3.3 เมื่อค่าพลังงานต่ำสุดหรือค่าพลังงานคงที่แสดงถึงความเป็นไปได้ของความสอดคล้องกันของจุดภาพจากทั้ง 2 กล้อง ซึ่งมีส่วนประกอบของฟังก์ชันดังนี้

$$E = -\left(\frac{1}{2}\right) \sum_{i=1}^{N_l} \sum_{k=1}^{N_r} \sum_{j=1}^{N_l} \sum_{l=1}^{N_r} C_{ikjl} P_{ik} P_{jl} + \sum_{i=1}^{N_l} (1 - \sum_{k=1}^{N_r} P_{ik})^2 + \sum_{k=1}^{N_r} (1 - \sum_{i=1}^{N_l} P_{ik})^2 \quad (3.8)$$

โดยเทอมแรกจะเป็นตัวบ่งชี้ความสอดคล้องกันของจุดภาพทั้งสอง ส่วนเทอมที่สองและสามคือเงื่อนไข (uniqueness constrain) ของแต่ละนิวรอน สามารถเขียนใหม่ได้เป็น

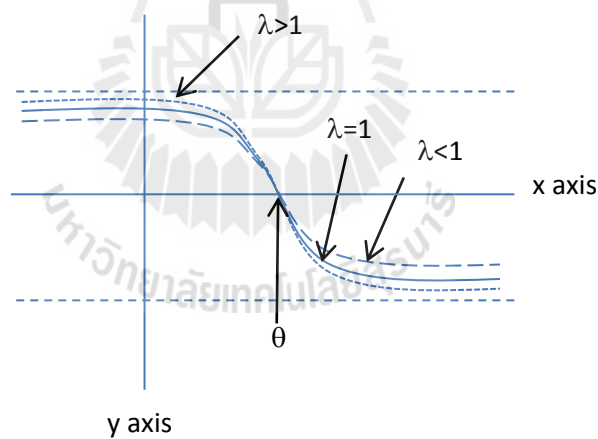
$$E = -\left(\frac{1}{2}\right) \sum_{i=1}^{N_l} \sum_{k=1}^{N_r} \sum_{j=1}^{N_l} \sum_{l=1}^{N_r} (C_{ikjl} - \delta_{ij} - \delta_{lk}) P_{ik} P_{jl} - \sum_{i=1}^{N_l} \sum_{k=1}^{N_r} 2P_{ik} \quad (3.9)$$

เมื่อนำสมการที่ 3.8 เปรียบเทียบกับสมการค่าพลังงานจะได้ว่าสถานะของนิวรอน $V_{ik}=P_{ik}$ $V_{jl}=P_{jl}$ และ $I_{ik}=2$ และค่าน้ำหนักประสาท $T_{ikjl}=C_{ikjl}-\delta_{ij}-\delta_{lk}$ โดยที่ $\delta_{ij}=1$ เมื่อ $i=j$ และ $\delta_{kl}=1$ เมื่อ $k=l$

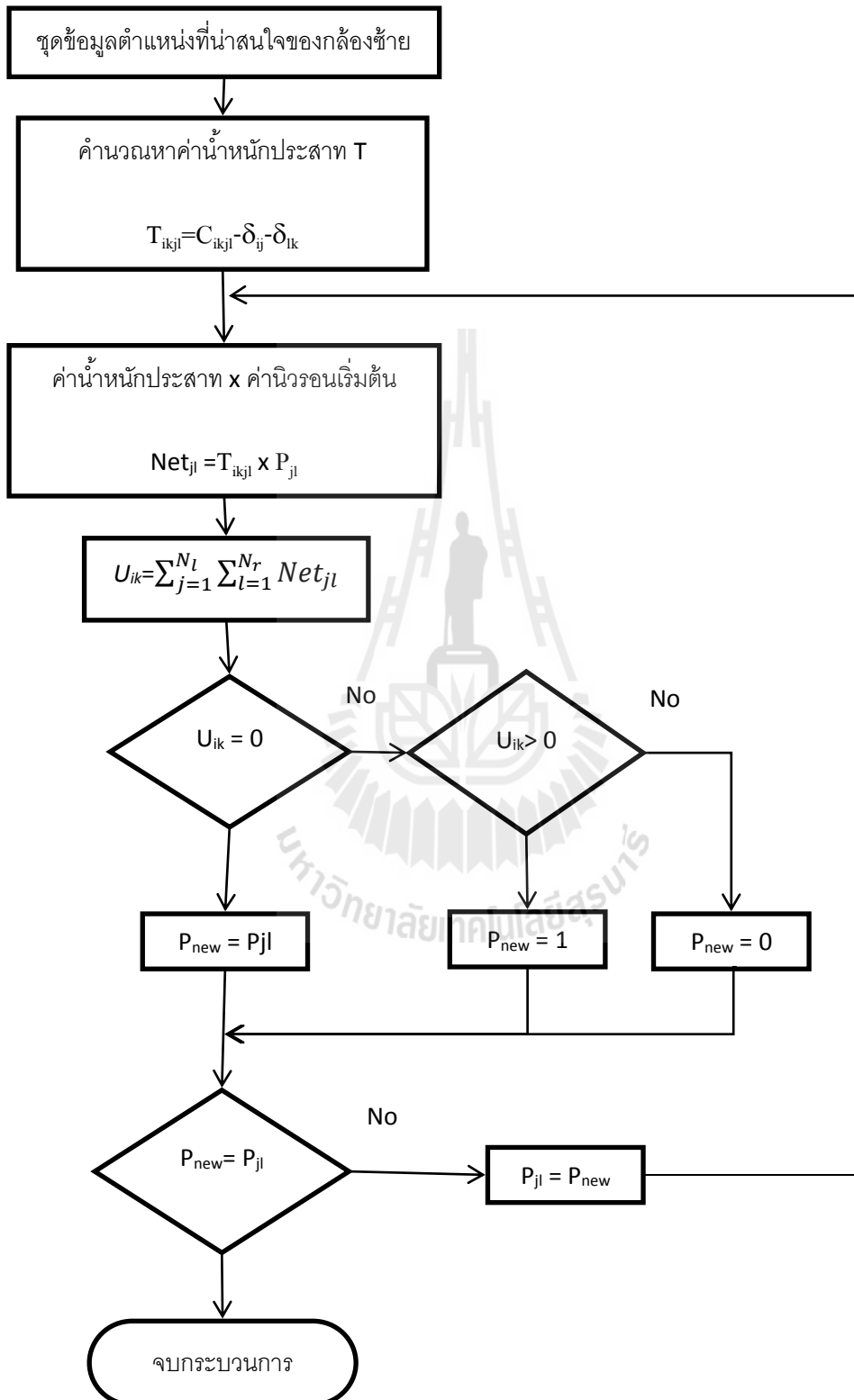
ค่าน้ำหนักประสาท T_{ikjl} ได้มาจาก

$$C_{ikjl} = (2 / [1 + e^{\lambda(x-\theta)}]) - 1 \quad (3.10)$$

$$X = [w_1 |\Delta d| + w_2 |\Delta D|] \quad (3.11)$$



รูปที่ 3.2 C_{ikjl} function (non-linear function)



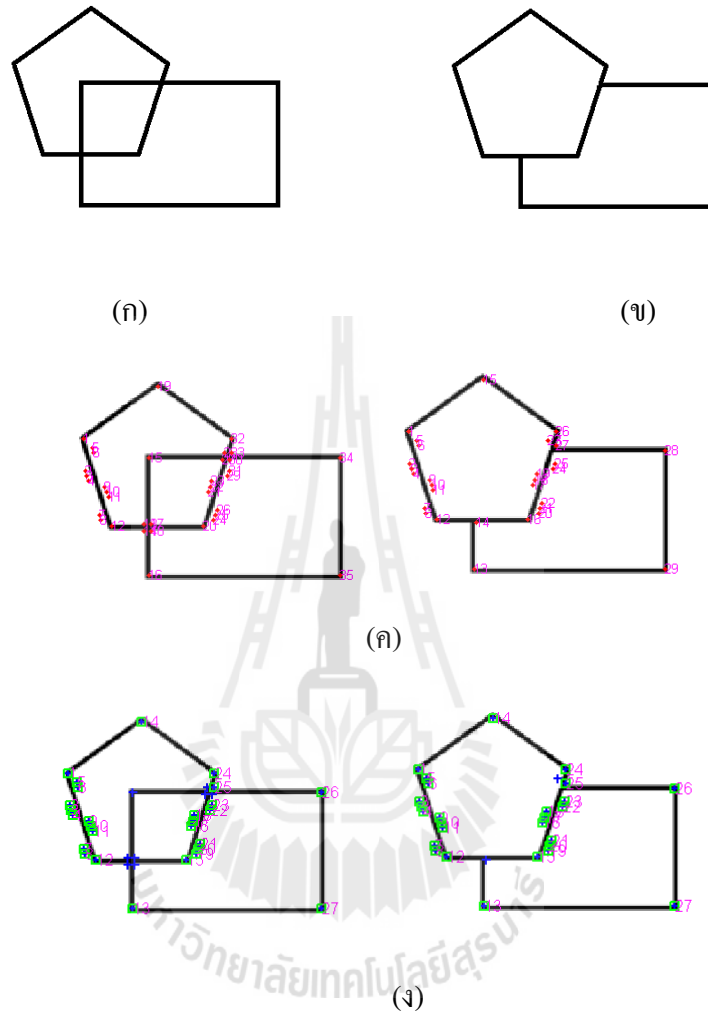
รูปที่ 3.3 แผนผังการทำงานของเครือข่ายฮอปฟิลด์

ซึ่งสมการที่ 3.10 เป็นฟังก์ชันไม่เป็นเชิงเส้น (non-linear function) เพื่อให้ค่าอยู่ในช่วง +1 และ -1 λ คือค่าความชันของฟังก์ชัน θ คือจุดตัดแกน x ของฟังก์ชัน จากผลการวิจัยของ N. Nasrabadi, C. Choo, 1992 พบว่าค่าที่ดีที่สุดคือ $\lambda = 1$ และ $\theta = 10$ ดังนั้นค่าน้ำหนักประสาทจึงขึ้นอยู่กับตัวแปร X จะได้ว่า เมื่อค่า $X=0$ คือค่าที่แสดงได้ว่าเป็นจุดที่ดีที่สุด แต่ถ้าค่า X มากๆ แสดงได้ว่าสองจุดนั้นไม่สอดคล้องกัน ค่า X หาได้จากสมการที่ (3.11) ซึ่ง Δd คือค่าระยะห่างระหว่างจุด (i,k) และ (j,l) ΔD คือค่าผลต่างระหว่างระยะห่างระหว่างจุด i และ j กับระยะห่างระหว่างจุด k และ l W_1 และ W_2 คือค่าคงที่โดย $W_1 + W_2 = 1$

เมื่อได้ค่าพารามิเตอร์ต่าง ๆ นำชุดข้อมูลที่ได้ป้อนเข้าสู่ระบบเครือข่ายฮอปฟิลด์เพื่อทำการวนรอบเพื่อหาความสอดคล้องของจุดภาพดังรูปที่ 3.3 เมื่อป้อนอินพุตเข้าสู่ระบบ เครือข่ายฮอปฟิลด์ จะทำการคำนวณหาเมตริกค่าน้ำหนักประสาทที่ได้จากสมการที่ 3.10 และสมการที่ 3.11 โดยมีเมตริกขนาดเท่ากับ $N_r \times N_i$ นิเวรอน โดยที่ N_r และ N_i คือจำนวนจุดเด่นของภาพทางด้านซ้ายและด้านขวาตามลำดับเมื่อได้ค่าเมตริกน้ำหนักประสาทแล้ว จึงนำไปคูณกับค่าเมตริกนิเวรอนเริ่มต้นโดยค่านิเวรอนเริ่มต้นเท่ากับ 0 แล้วหาผลรวมเพื่อหาสถานะของแต่ละนิเวรอนในแต่ละรอบ แล้วตรวจสอบสถานะนิเวรอนใหม่กับสถานะนิเวรอนเริ่มต้น ว่ามีการเปลี่ยนแปลงหรือไม่ ถ้ามีการเปลี่ยนแปลง ให้เมตริกนิเวรอนเริ่มต้นเท่ากับเมตริกนิเวรอนใหม่ แล้วนำไปคำนวณกับค่าน้ำหนักประสาทจนกว่าเมตริกสถานะนิเวรอนใหม่เท่ากับนิเวรอนเดิมจึงสิ้นสุดกระบวนการ สถานะนิเวรอนใหม่ที่สิ้นสุดกระบวนการแล้วนั้น ถ้าเท่ากับ 1 ในเมตริกตำแหน่งใดแสดงว่าตำแหน่งจากกล้องซ้ายและกล้องขวาในเมตริกตำแหน่งนั้นสอดคล้องกัน

3.5 ผลการทดสอบการแก้ปัญหาความสอดคล้องกันของจุดภาพโดยใช้เครือข่ายประสาทเทียมแบบฮอปฟิลด์ 2 มิติ

ตัวอย่างข้อมูลชุดแรกมาจากการจำลองต้นแบบเพื่อตรวจสอบการทำงานของระบบที่ใช้ในการแก้ปัญหาความสอดคล้องนี้ โดยใช้การตรวจจับแบบ Harris (Harris's detector) แล้วป้อนชุดข้อมูลที่ได้เข้าเครือข่ายฮอปฟิลด์



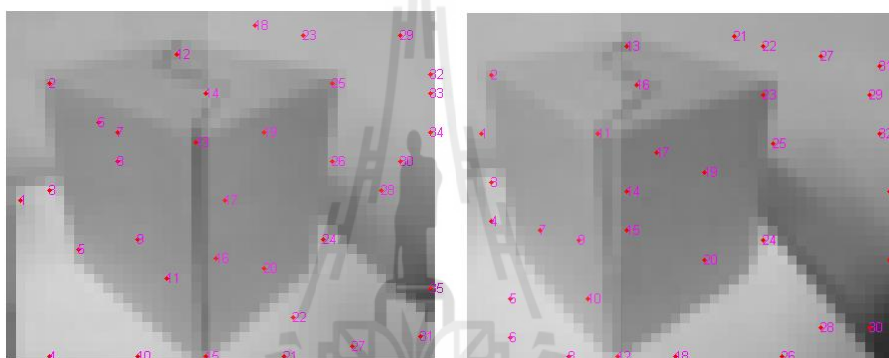
รูปที่ 3.4 (ก) ภาพต้นแบบทางด้านซ้าย (ข) ภาพต้นแบบทางด้านขวา (ค) ตำแหน่งที่เกิดจากการตรวจจับของแฮริส (ง) ผลการแก้ปัญหาด้วยเครือข่ายฮอฟฟิลด์ 2 มิติ

จากรูปที่ 3.4 (ก) และ (ข) ภาพต้นแบบซ้ายและขวามีจุดแตกต่างกัน เมื่อทำการหาตำแหน่งที่น่าสนใจด้วยวิธี Harris แล้วจะได้ ดังรูปที่ 3.4 (ค) ซึ่งภาพทางด้านซ้ายจะมีตำแหน่งที่น่าสนใจมากกว่าภาพทางด้านขวา เมื่อผ่านกระบวนการของเครือข่ายฮอฟฟิลด์แล้วจะได้ดังรูปที่ 3.4 (ง) ตำแหน่งที่กำกับด้วยตัวเลขแสดงถึงจุดที่สอดคล้องกัน ซึ่งเห็นได้ว่า ตำแหน่งที่ตัวเลขกำกับไว้คือตำแหน่งเดียวกันจึงทำการทดลองกับรูปจริงที่ถ่ายด้วยกล้อง 2 ตัว

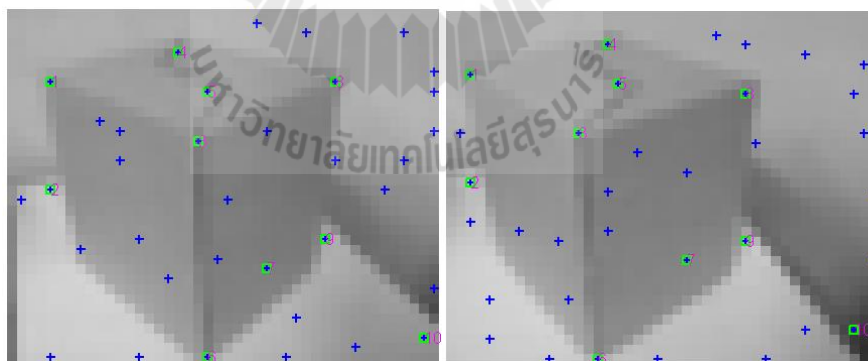


(ก)

(ข)



(ค)



(ง)

รูปที่ 3.5 (ก) ภาพจากกล้องซ้าย (ข) ภาพจากกล้องขวา (ค) ตำแหน่งที่เกิดจากการตรวจจับของ Harris (ง) ผลการแก้ปัญหาด้วยเครือข่ายฮอปฟิลด์ 2 มิติ

จากรูปที่ 3.5 (ง) จะเห็นว่าตำแหน่งที่กำกับด้วยตัวเลข 1 ถึง 10 เป็นตำแหน่งที่สอดคล้องกันของภาพจากกล้องซ้ายและกล้องขวา จากการผลการทดลองตำแหน่งของวัตถุที่ต้องการคือกล่อง

ที่อยู่ในภาพ แต่วิธีการคัดกรองของแฮริสตำแหน่งที่น่าสนใจได้กระจายไปทั่วภาพ ซึ่งเป็นตำแหน่งที่ไม่ต้องการ ทั้งยังเป็นภาระของระบบเครือข่ายฮอฟฟิลด์ 2 มิติ ยิ่งอินพุตมากการทำงานของระบบก็ช้าลงด้วย ดังนั้นการใช้วิธีการนี้ในการคัดกรองวัตถุในภาพเคลื่อนไหวจึงไม่เหมาะสมนัก



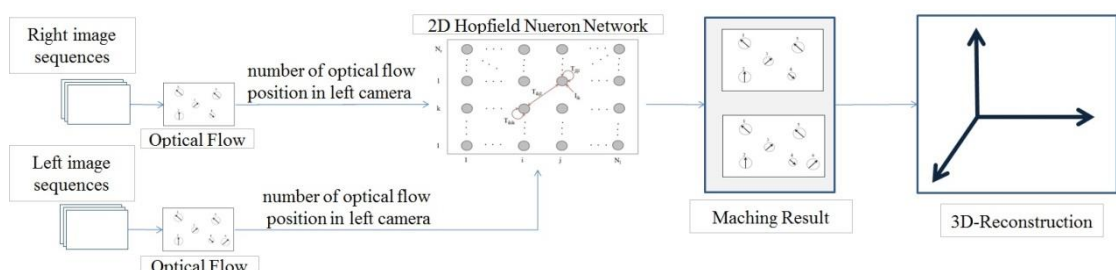
บทที่ 4

การบูรณะสถานะแวดล้อม 3 มิติด้วยวิธี

แผนที่บ่งบอกความลึกเชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง

4.1 กล่าวนำ

ระบบการทำงานของระบบบูรณะสถานะแวดล้อม 3 มิติในงานวิจัยนี้ ได้ใช้การไหลเชิงแสงเป็นตัวคัดแยกวัตถุที่เคลื่อนที่ออกจากพื้นหลัง โดยวัตถุใด ๆ ที่มีการเคลื่อนที่ภายในภาพ เมื่อผ่านกระบวนการการไหลเชิงแสงแล้ว บริเวณที่วัตถุเคลื่อนที่ในภาพ จะเกิดการไหลของแสงขึ้นทำให้ได้ข้อมูลทั้งขนาดและมุมพร้อมกับตำแหน่งที่เกิดการไหลเชิงแสง ซึ่งข้อมูลเหล่านี้เป็นอินพุตให้กับระบบเครือข่ายฮอปฟิลด์ดังที่กล่าวไปแล้วในบทที่ 3 โครงสร้างของระบบแสดงดังรูปที่ 4.1 โดยขั้นตอนของระบบรวมนี้จะเริ่มต้นด้วยการนำลำดับภาพจากกล้อง 2 ตัว เข้าสู่การเตรียมการประมวลผลด้วยการปรับขนาด การแปลงภาพจากภาพสีให้เป็นภาพระดับเทา จากนั้นนำภาพที่ผ่านการประมวลผลแล้วเข้ารับการประมวลผลการไหลเชิงแสง จะได้อินพุตป้อนให้กับการประมวลผลเครือข่ายฮอปฟิลด์เพื่อแก้ปัญหาค่าความสอดคล้อง แล้วจึงนำผลของการประมวลผลมาคำนวณหาระยะห่างระหว่างวัตถุกับกล้องและการบูรณะสถานะแวดล้อม 3 มิติด้วยวิธีการทางคณิตศาสตร์ดังที่กล่าวไว้ในบทที่ 2

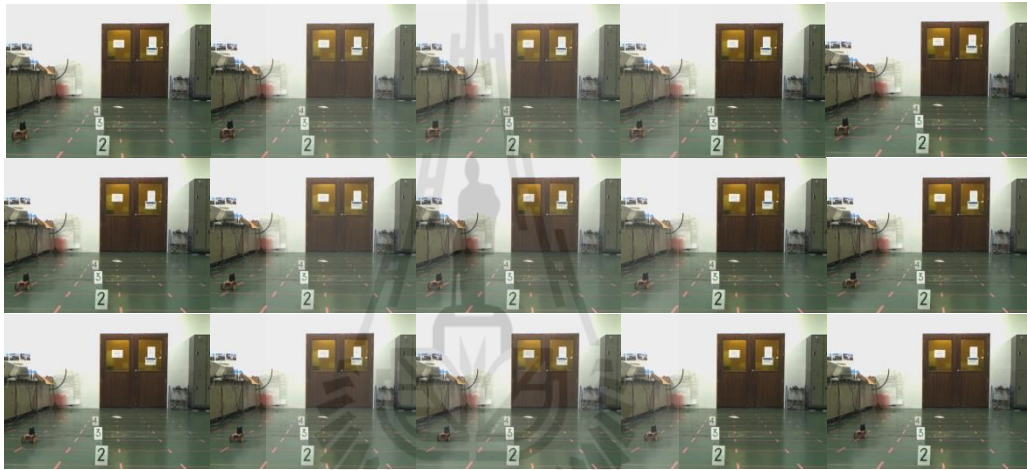


รูปที่ 4.1 โครงสร้างระบบการบูรณะสถานะแวดล้อม 3 มิติ

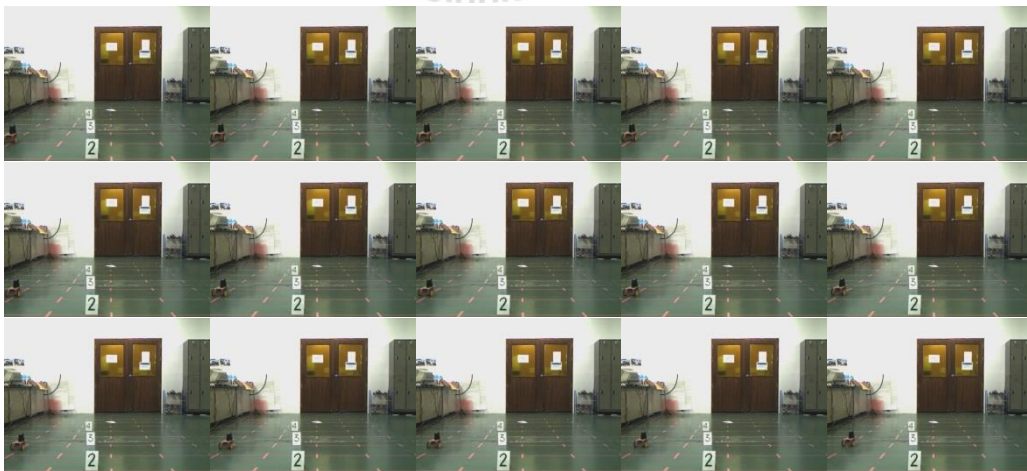
4.2 การคัดแยกวัตถุด้วยการไหลเชิงแสง

ในส่วนนี้จะกล่าวถึงการไหลเชิงแสงที่นอกจากจะนำมาตรวจจับวัตถุที่เคลื่อนไหวแล้วยังสามารถนำขนาดของการไหลเชิงแสงมาเป็นตัวคัดแยกวัตถุออกจากพื้นหลังอีกด้วย รวมทั้งยังเป็นตัวอินพุตให้กับระบบเครือข่ายประสาทเทียมเพื่อแก้ปัญหาความสอดคล้องกันของจุดภาพ แบ่งเป็นขั้นตอนได้ดังนี้

ขั้นตอนที่ 1 รับชุดข้อมูลภาพอินพุตจากกล้องดังรูปที่ 4.2 และรูปที่ 4.3



รูปที่ 4.2 ชุดข้อมูลลำดับภาพที่ได้จากกล้องซ้าย



รูปที่ 4.3 ชุดข้อมูลลำดับภาพที่ได้จากกล้องขวา

ขั้นตอนที่ 2 นำภาพที่ได้มาผ่านกระบวนการประมวลผลโดยแปลงจากภาพสีมาเป็นภาพระดับเทา

ขั้นตอนที่ 3 ประมวลผลการไหลเชิงแสงและกำหนดช่วงขนาดของการไหลเชิงแสงที่ได้เพื่อหาวัตถุที่เคลื่อนไหวในภาพ แล้วบันทึกข้อมูลตำแหน่ง ขนาดและมุมที่เกิดการไหลเชิงแสง



(ก)



(ข)

(ค)

- รูปที่ 4.4 (ก) การไหลเชิงแสงของภาพที่ได้จากกล้องซ้ายที่ยังไม่ได้กรองขนาด
 (ข) ภาพขยายการไหลเชิงแสงของภาพที่ได้จากกล้องซ้ายที่ยังไม่ได้กรองขนาด
 (ค) ภาพขยายการไหลเชิงแสงของภาพที่ได้จากกล้องขวาที่ยังไม่ได้กรองขนาด

จากวิธีการของ Lucas & Kanade ทำให้ได้ตำแหน่งของการเคลื่อนที่ดังรูปที่ 4.3(ก) ตำแหน่งที่เกิดการไหลเชิงแสงจะอยู่บริเวณรอบๆ วัตถุซึ่งรูปที่ 4.3(ข) และ (ค) ทำให้เห็นขนาดและมุมของการไหลเชิงแสงที่เกิดขึ้นได้ชัดเจนยิ่งขึ้น จะสังเกตได้ว่าการไหลเชิงแสงของภาพทั้งสองภาพมีขนาดใกล้เคียงกันทั้งตำแหน่ง มุมและขนาด แต่อย่างไรก็ตามเมื่อนำไปประมวลผลในขั้นตอนต่อไปจำนวนตำแหน่งที่เกิดการไหลเชิงแสงมีมากเกินไประบบไม่สามารถประมวลได้จึงต้องทำการกรองขนาดของการไหลเชิงแสงอีกครั้งโดยกำหนดให้ขนาดการไหลอยู่ในช่วง 0.7 - 0.5

4.3 การแก้ปัญหาคงความสอดคล้องกันของจุดภาพ

เมื่อผ่านการตัดแยกวัตถุที่เราสนใจด้วยการไหลเชิงแสงแล้ว จะได้ชุดข้อมูลออกมาเป็นตำแหน่งการเกิดการไหลเชิงแสงในแกนแนวนอนและแนวแกนตั้ง (x,y) ขนาดการไหลเชิงแสงและมุมของการไหลเชิงแสงชุดข้อมูลเหล่านี้จะถูกลำนำไปประมวลผลในเครือข่ายฮอปฟิลด์ โดยเริ่มจากคำนวณค่า $C_{ijkl} = (2 / [1 + e^{\lambda(x-\theta)}]) - 1$ โดย X หาได้จาก

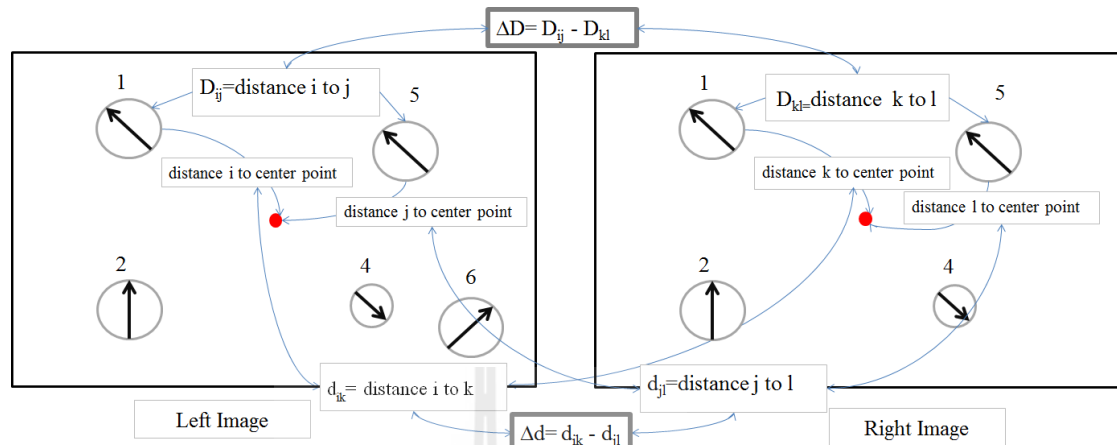
$$X = [w_1|\Delta D| + w_2|\Delta d| + w_3|\Delta A|] \quad (4.1)$$

ซึ่ง ΔD คือ ค่าผลต่างของผลต่างของระยะจากตำแหน่งที่ i ถึงจุดอ้างอิงกับระยะจากตำแหน่งที่ k ถึงจุดอ้างอิงและผลต่างของระยะจากตำแหน่งที่ j ถึงจุดอ้างอิงกับระยะจากตำแหน่งที่ 1 ถึงจุดอ้างอิง

Δd คือ ค่าผลต่างระหว่างระยะจากตำแหน่งที่ i ถึงตำแหน่งที่ j กับระยะจากตำแหน่งที่ k ถึงตำแหน่งที่ 1

ΔA คือ ค่าผลต่างของมุมจากการไหลเชิงแสงที่ตำแหน่ง i และ k

w_1, w_2 และ w_3 คือ ค่าตัวชี้วัดความสำคัญของเงื่อนไขโดยที่รวมกันแล้วเท่ากับ 1



รูปที่ 4.5 การคำนวณค่า ΔD และ Δd

สมการที่ 4.1 เป็นเงื่อนไขแรกของระบบการทำงานของเครือข่ายประสาทเทียมอธิบายได้จากรูปที่ 4.4 โดยมี ΔD เป็นตัวเปรียบเทียบตำแหน่งของตัวเองกับตำแหน่งรอบข้าง ก่อนที่จะนำผลที่ได้ไปเปรียบเทียบกับอีกภาพ กล่าวคือ เมื่อพิจารณาตำแหน่งที่เกิดการไหลเชิงแสงในภาพซ้ายระบบจะทำการเปรียบเทียบขนาดที่ตำแหน่งตัวเองกับตำแหน่งของสมาชิกในภาพ ก่อนที่จะไปเปรียบเทียบตำแหน่งทุกตำแหน่งของสมาชิกที่เกิดการไหลเชิงแสงทางด้านขวา เมื่อพิจารณาแล้วจะเห็นได้ว่า ถ้าตำแหน่งของภาพทางด้านซ้ายและขวาสัมพันธ์กันหรือเป็นจุดเดียวกันค่าที่ได้จะเท่ากับ 0 เช่นเดียวกับกับค่า Δd แต่ใช้ตำแหน่งกึ่งกลางภาพเป็นจุดอ้างอิง โดยพิจารณาจากการวัดระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพซ้ายเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพขวา แล้วจึงนำไปเปรียบเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพซ้ายเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพขวาในจุดต่อไปทุกๆตำแหน่ง ในกรณีที่เป็นตำแหน่งเดียวกันควรมีค่าน้อยๆ หรือเข้าใกล้ 0 และ Δd คือการเปรียบเทียบมุมระหว่างตำแหน่งที่พิจารณาที่ภาพซ้ายและขวา ถ้าเป็นมุมเดียวกันค่าที่ได้ต้องมีค่าน้อยๆ หรือเท่ากับ 0

เมื่อพิจารณาความสำคัญของเงื่อนไขกำหนดให้ $\Delta d > \Delta D > \Delta d$ จากคำอธิบายข้างต้นจะเห็นได้ว่า ถ้าตำแหน่งทางภาพซ้ายเป็นตำแหน่งเดียวกันกับภาพขวา $\Delta D = 0$ และ $\Delta d = 0$ จะเหลือ Δd เพื่อพิจารณาความสัมพันธ์ที่เป็นไปได้ จาก N. Nasrabadi และ C. Choo (1992) ได้นำเสนอไว้ว่าค่า X ที่ดีควรมีค่าน้อยๆ ดังนั้น Δd ควรมีค่าน้อยลงด้วย เมื่อกำหนดให้ w_1, w_2 และ w_3 คือค่า

ตัวชี้วัดความสำคัญของเงื่อนไขของ ΔD , Δd และ ΔA ตามลำดับ ดังนั้นการหาค่า W_1 , W_2 และ W_3 จากการสุ่มใช้ค่าที่เป็นไปได้ตามเงื่อนไขจึงมีอยู่ 4 กรณีดังนี้

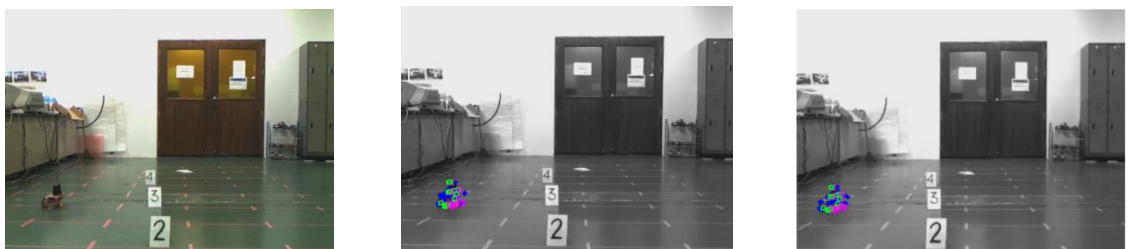
ตารางที่ 4.1 ค่า W_1 , W_2 และ W_3 ที่เป็นไปได้ตามเงื่อนไข $\Delta A > \Delta D > \Delta d$

กรณี	W_1	W_2	W_3
1	0.5	0.4	0.2
2	0.5	0.3	0.1
3	0.6	0.3	0.1
4	0.7	0.2	0.1

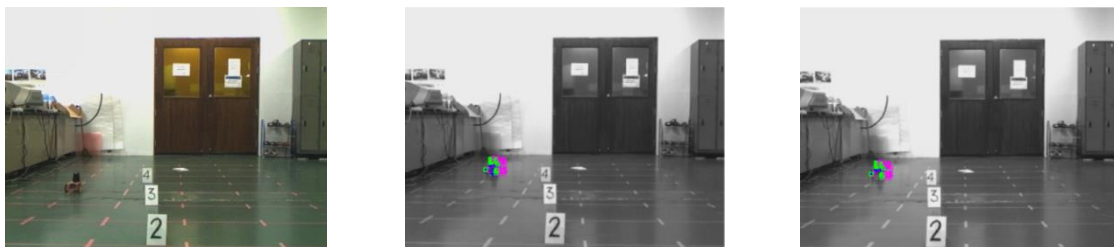
นอกจากค่าในตารางที่ 4.1 แล้วยังสามารถทำการสุ่มค่าอื่น ๆ เพื่อทำการทดลองผลปรากฏว่า เมื่อกำหนดให้ $W_1 = 0.6$, $W_2 = 0.3$ และ $W_3 = 0.1$ เป็นค่าที่เหมาะสมที่สุดโดยวัดจากความถูกต้องของการจับคู่ตำแหน่งจากภาพซ้ายและขวาหลังจากนำชุดข้อมูลที่ได้ผ่านกระบวนการประมวลผลเครือข่ายฮอปฟิลด์แล้ว ได้ผลลัพธ์ดังรูป



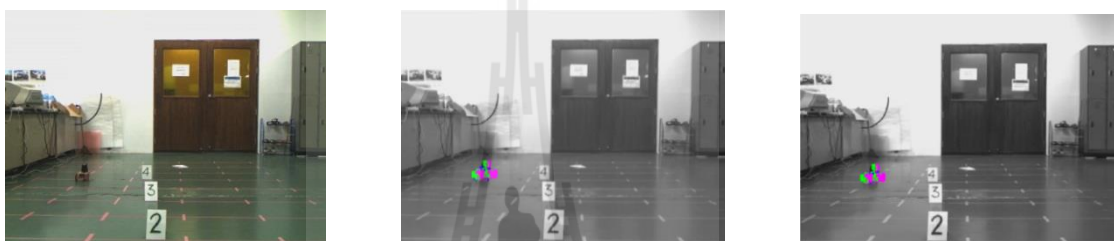
รูปที่ 4.6 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 1



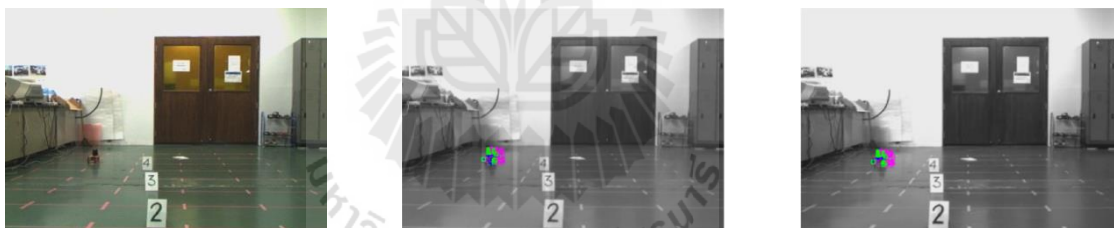
รูปที่ 4.7 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 2



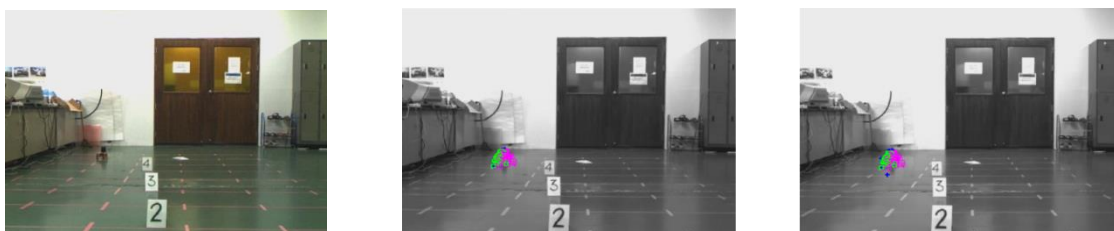
รูปที่ 4.8 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 3



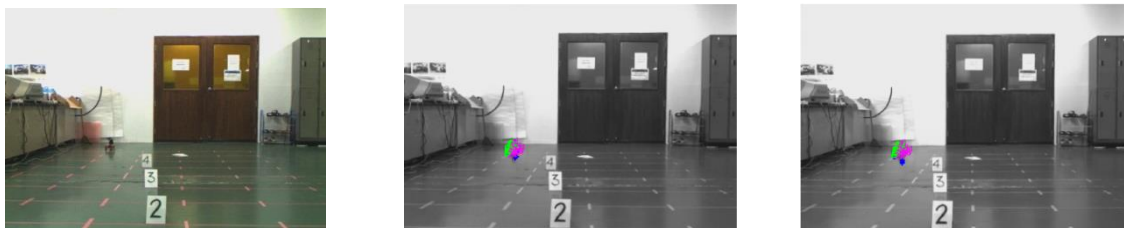
รูปที่ 4.9 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 4



รูปที่ 4.10 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 5



รูปที่ 4.11 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 6



รูปที่ 4.12 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 7

ตารางที่ 4.2 ค่าความผิดพลาดของการจับคู่ภาพของกล้องซ้ายและขวาในชุดตัวอย่างที่ 1-10

ตัวอย่าง	จำนวนตำแหน่งที่สนใจ		จำนวนที่ถูกจับคู่	จำนวนคู่ที่ถูก	ความผิดพลาด (%)
	ภาพซ้าย	ภาพขวา			
1	56	76	37	35	5.41
2	70	86	19	18	5.26
3	77	93	45	45	0
4	96	62	36	36	0
5	77	62	26	26	0
6	92	92	35	35	0
7	73	92	22	20	9.09

จากผลการทดสอบจากชุดตัวอย่างดังรูปที่ 4.5 - 4.14 สรุปได้ดังตารางที่ 4.2 ความถูกต้องของการจับคู่ตำแหน่งของภาพซ้ายและภาพขวาวัดจากการมองเห็นของมนุษย์ รูปที่ 4.5 เป็นผลของการจับคู่ของชุดตัวอย่างที่ 1 ผลการจับคู่ทั้งหมด 37 ตำแหน่ง มีผิดพลาด 2 ตำแหน่ง คิดเป็นความผิดพลาดร้อยละ 5.41 โดยความผิดพลาดของการจับคู่คิดจากจำนวนที่ถูกจับคู่กับจำนวนการจับคู่ที่ถูกต้อง เพื่อตรวจสอบความแม่นยำของระบบจึงทำการทดสอบกับชุดตัวอย่างที่ 2-7 ผลปรากฏดังตารางที่ 4.2 ซึ่งค่าความผิดพลาดของชุดข้อมูลทั้งหมดน้อยกว่าร้อยละ 10 ซึ่งเป็นค่าที่ยอมรับได้

4.4 การหาระยะห่างระหว่างกล้องกับวัตถุและการบูรณะสถานะแวดล้อม 3 มิติ

การหาระยะห่างระหว่างกล้องกับวัตถุสามารถคำนวณได้จาก

$$Z = \frac{bf}{(x'_l - x'_r)} \quad (4.2)$$

$$X = \frac{Z \times x'_l}{f} \quad (4.3)$$

$$Y = \frac{Z \times y'_l}{f} \quad (4.4)$$

- โดยที่ Z คือระยะห่างระหว่างกล้องและวัตถุ
- X, Y คือตำแหน่งในแนวนอนและแนวตั้งตามลำดับ โดยมีจุดอ้างอิงเป็นจุดกึ่งกลางของรูปภาพ
- b คือระยะห่างระหว่างกล้องซ้ายและกล้องขวา
- f คือความยาวโฟกัสของกล้อง
- x'_l, x'_r คือระยะห่างระหว่างตำแหน่งในแนวแกนขนานกับจุดกึ่งกลางรูปของกล้องซ้ายและขวาตามลำดับ
- y'_l คือระยะห่างระหว่างตำแหน่งในแนวแกนตั้งกับจุดกึ่งกลางรูป
- จากชุดตัวอย่างที่ 1 เมื่อผ่านขั้นตอนการจับคู่แล้วคำนวณจะได้ค่า Z ดังตารางเมื่อกำหนดให้ระยะห่างระหว่างกล้องซ้ายและขวา (ค่า b) เท่ากับ 5 เซนติเมตร และความยาวโฟกัส (ค่า f) เท่ากับ 8 มิลลิเมตร

ตารางที่ 4.3 ข้อมูลที่นำมาใช้ในการคำนวณหาระยะทางจากผลการทดลองในตัวอย่างที่ 1

ภาพซ้าย				ภาพขวา				$(x'_i - x'_r)$	z
ลำดับ	x	y	มุม	ลำดับ	x	y	มุม		
3	47	289	-3.08262	4	26	293	-2.95477	21	1.904
4	48	289	-3.02566	5	27	293	-2.91377	21	1.904
11	43	309	-3.0954	13	22	313	-3.11651	21	1.904
12	41	310	-3.03913	17	20	314	-3.09158	21	1.904
13	42	310	-3.02712	18	21	314	-3.05545	21	1.904
14	43	310	-3.07331	19	22	314	-3.0789	21	1.904
15	44	310	-3.11756	20	23	314	-3.11063	21	1.904
16	45	310	-3.13152	21	24	314	-3.12698	21	1.904
17	40	311	-3.03498	23	19	315	-3.05535	21	1.904
18	41	311	-2.96094	24	20	315	-2.9762	21	1.904
19	42	311	-3.0056	25	21	315	-2.96911	21	1.904
20	43	311	-3.10396	26	22	315	-3.06289	21	1.904
21	46	311	-3.10725	28	25	315	-3.13237	21	1.904
22	47	311	-3.10543	29	26	315	-3.06978	21	1.904
25	54	311	-2.84269	31	33	315	-2.93584	21	1.904
26	39	312	-2.97273	32	18	316	-3.03596	21	1.904
27	40	312	-2.84948	33	19	316	-2.93042	21	1.904
28	41	312	-2.89623	34	20	316	-2.88358	21	1.904
29	42	312	-3.00859	35	21	316	-2.95036	21	1.904
30	43	312	-3.10801	36	22	316	-3.06218	21	1.904
32	46	312	-3.10164	39	25	316	-3.13237	21	1.904
34	54	312	-2.88723	42	33	316	-2.98091	21	1.904
35	40	313	-2.90153	43	19	317	-2.9641	21	1.904
36	41	313	-2.87835	44	20	317	-2.83454	21	1.904
37	42	313	-2.99874	45	21	317	-2.90857	21	1.904
38	48	313	2.688392	47	27	317	2.757066	21	1.904
40	24	315	1.38839	49	3	319	1.421149	21	1.904
41	24	316	1.404309	50	3	320	1.360304	21	1.904
44	46	323	1.428906	54	25	327	1.565879	21	1.904
45	47	323	1.303429	55	26	327	1.272486	21	1.904
46	46	324	1.362406	57	25	328	1.425486	21	1.904
47	47	324	1.29221	58	26	328	1.270589	21	1.904
48	47	325	1.314975	59	26	329	1.314925	21	1.904
49	41	329	-3.12293	61	20	333	-3.1045	21	1.904
50	50	331	1.594185	68	29	335	1.58619	21	1.904
51	51	331	1.621852	69	30	335	1.579088	21	1.904
53	50	332	1.614758	70	29	336	1.597621	21	1.904

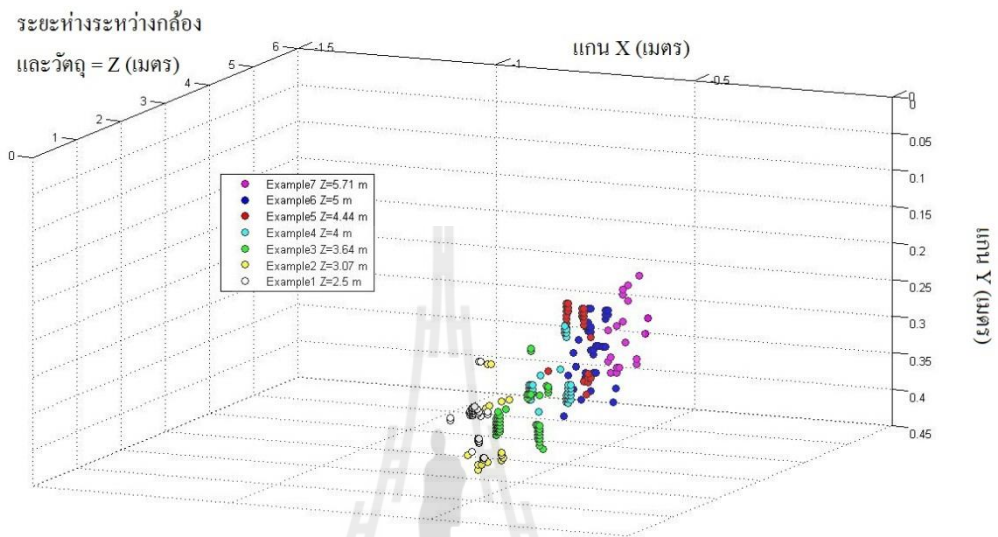
เมื่อทำการคำนวณเสร็จสิ้นจะเห็นได้ว่าค่าที่ได้เท่ากับ 1.904 เมตรซึ่งค่าจริงของชุดตัวอย่างคือ 2.5 เมตร จึงได้ทำการทดลองในตัวอย่างที่ 2 ถึง 7 ปรากฏว่ามีทุกตำแหน่งคลาดเคลื่อน จึงสังเกตได้ว่าค่าที่ทำให้ผิดพลาดนั้นอยู่ที่ค่าความต่างของตำแหน่งแกน x ซึ่งเกิดจากการปรับขนาดรูปที่ได้จากกล้องให้มีขนาดเล็กลงจากเดิมซึ่งค่าที่ส่งผลให้การคำนวณคลาดเคลื่อนนั้นเป็นค่าคงที่ซึ่งเท่ากับ 5 g เมื่อนำค่าคงที่ดังกล่าวนี้ไปลบออกจากค่าความต่างเดิมจะได้ผลการทดลองที่ปรับปรุงแล้วดังตารางที่ 4.4

ตารางที่ 4.4 เปอร์เซนต์ความถูกต้องของระยะห่างระหว่างวัตถุกับกล้องที่คำนวณได้กับค่าระยะห่างจริง

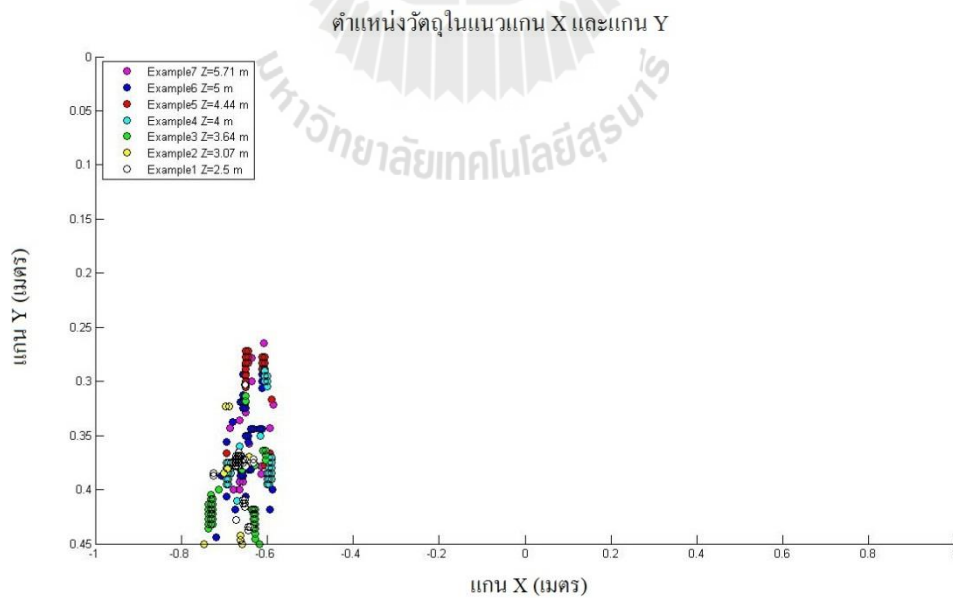
ชุดตัวอย่างที่	Z (เมตร)	ค่าจริง (เมตร)	ค่าความคลาดเคลื่อน (%)
1	2.5	2.5	0
2	3.08	3	2.67
3	3.64	3.5	4
4	4	4	0
5	4.44	4.5	6
6	5	5	0
7	5.714	5.5	3.89

เมื่อทำการคำนวณหาระยะห่างระหว่างวัตถุและกล้องได้แล้วจึงทำการคำนวณหาระยะในแนวแกนนอนและแกนตั้งของระบบจริง ซึ่งมีจุดอ้างอิงคือตรงจุดกึ่งกลางของภาพ ดังสมการที่ 4.3 และ 4.4 จะได้ผลการทดลองดังรูปที่ 4.13 – 4.15

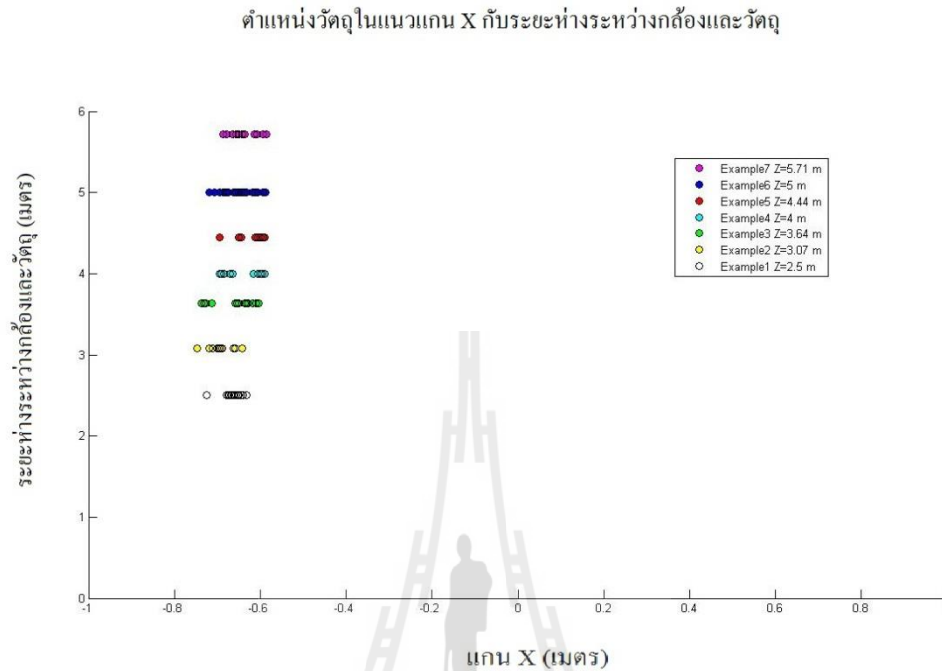
ภาพสามมิติของวัตถุที่เกิดขึ้นในตำแหน่งต่าง



รูปที่ 4.13 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ 3 มิติ



รูปที่ 4.14 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ XY



รูปที่ 4.15 ตำแหน่งวัตถุที่ระยะต่างๆในระนาบ ZX

รูปที่ 4.13 แสดงตำแหน่งวัตถุที่ระยะต่างๆในระนาบ 3 มิติของวัตถุตั้งแต่ตัวอย่างที่ 1-7 ซึ่งสังเกตได้ว่าพิกัดของวัตถุในภาพเสมือนกับวัตถุที่เกิดในภาพจริง สังเกตได้จากรูปที่ 4.14 เมื่อกำหนดให้แนวแกน Y เปรียบเสมือนพื้นของภาพจริงจะเห็นได้ว่าพิกัดของวัตถุที่คำนวณได้ของตัวอย่างที่ 1-7 อยู่ในแนวระนาบเดียวกัน จึงกำหนดได้ว่า ที่ระยะห่างจากจุดกึ่งกลางของการติดตั้งกล้องลงมา 0.45 เมตรคือพื้นของโลก และแนวการเคลื่อนที่ของวัตถุยังเหมือนกับภาพต้นแบบดังแสดงไว้รูปที่ 4.15

จากการผลการทดลองที่ผ่านมาจะเห็นได้ว่าความแม่นยำในการคำนวณหาตำแหน่งของวัตถุขึ้นอยู่กับการติดตั้งระยะห่างระหว่างกล้อง 2 ตัว ขนาดโฟกัสของเลนส์และขนาดของภาพที่ใช้ป้อนอินพุตในระบบ การติดตั้งกล้องยิ่งห่างกันมากช่วงการวัดระยะของวัตถุยิ่งละเอียดคงตาราง แต่ข้อเสียของการติดตั้งกล้องห่างกันมากก็นั้น จะทำให้วัตถุที่ปรากฏด้านซ้ายหรือด้านขวาข้างใดข้างหนึ่งมากๆ จะทำให้วัตถุไม่ปรากฏในภาพและไม่สามารถคำนวณหาวัตถุได้

ตารางที่ 4.5 ระยะทางที่วัดได้สูงสุดในการติดตั้งกล้องและขนาดภาพต่าง ๆ ที่ป้อนให้กับระบบ

ขนาดภาพ (pixel)	ระยะห่างในการติดตั้งกล้องซ้ายและกล้องขวา (เซนติเมตร)		
	5	10	15
512 x 384	7	8	10

ขนาดภาพในตารางเป็นการปรับจากรูปต้นแบบที่ขนาด 1024x768 พิกเซล

4.5 ผลการทดสอบการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ

การทดสอบระบบการหาระยะระหว่างกล้องและวัตถุจะทำการเก็บข้อมูลในช่วงสภาวะต่างๆจากสภาพจริงที่ทั้งภายในอาคารและภายนอกอาคารทั้งเคลื่อนที่และไม่เคลื่อนที่ ผลที่ได้ระบบมีความคงทนต่อการเปลี่ยนแปลงสภาพแวดล้อมได้ดีและสามารถวัดระยะวัตถุที่เคลื่อนที่ได้อย่างมีประสิทธิภาพและได้ทำการทดสอบจากเพิ่มลำดับภาพต่างๆ โดยแบ่งออกเป็นประเภทได้ดังนี้

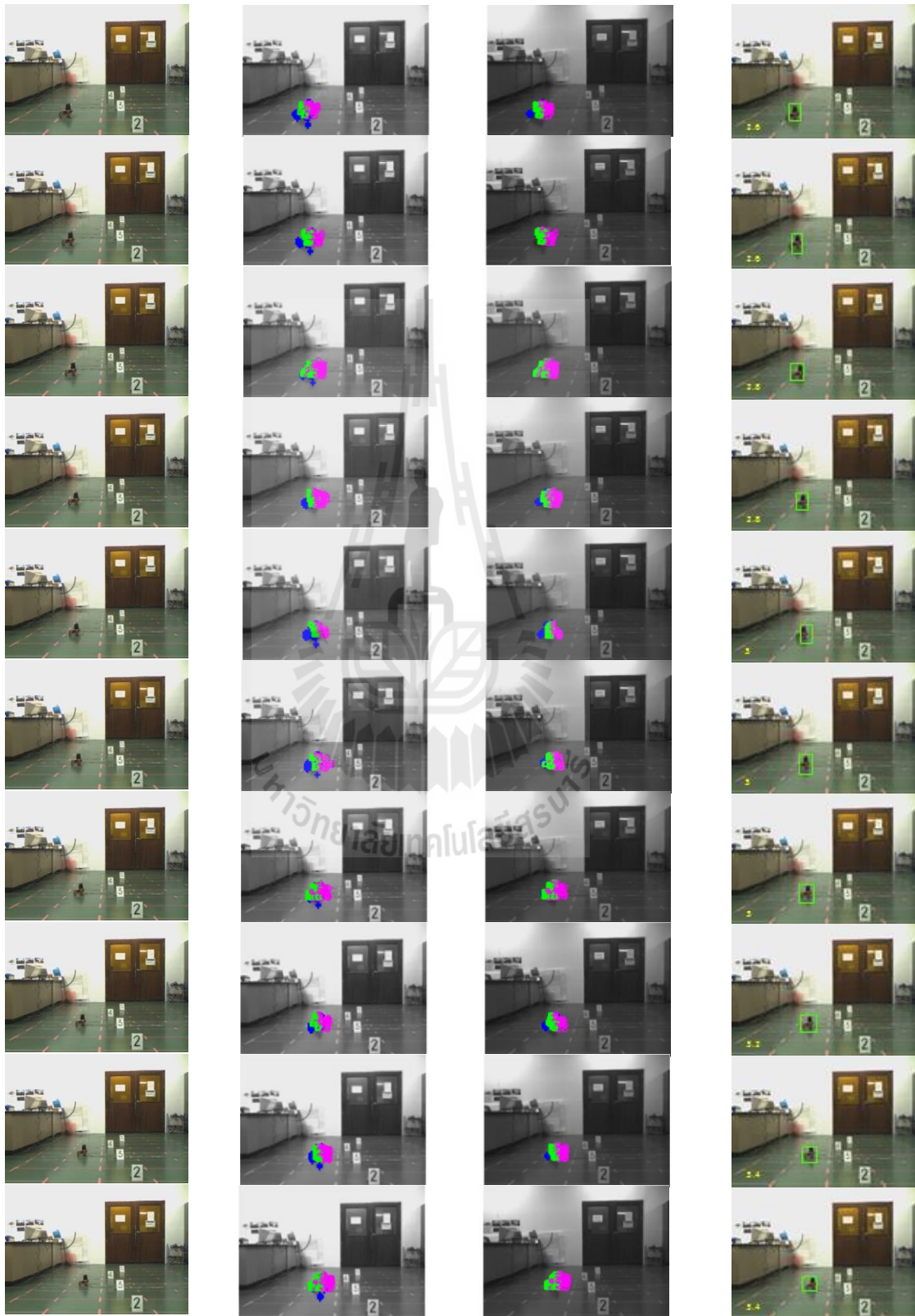
4.5.1 วัตถุเคลื่อนที่ภายในอาคาร ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายในอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 1 ตำแหน่งผลการทดสอบเป็นไปตามรูปที่ 4.16

4.5.2 วัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายในอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 2 ตำแหน่ง ผลการทดสอบเป็นไปตามรูปที่ 4.17

4.5.3 วัตถุเคลื่อนที่ภายนอกอาคาร ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายนอกอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระผลการทดสอบเป็นไปตามรูปที่ 4.18

ตารางที่ 4.6 ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ

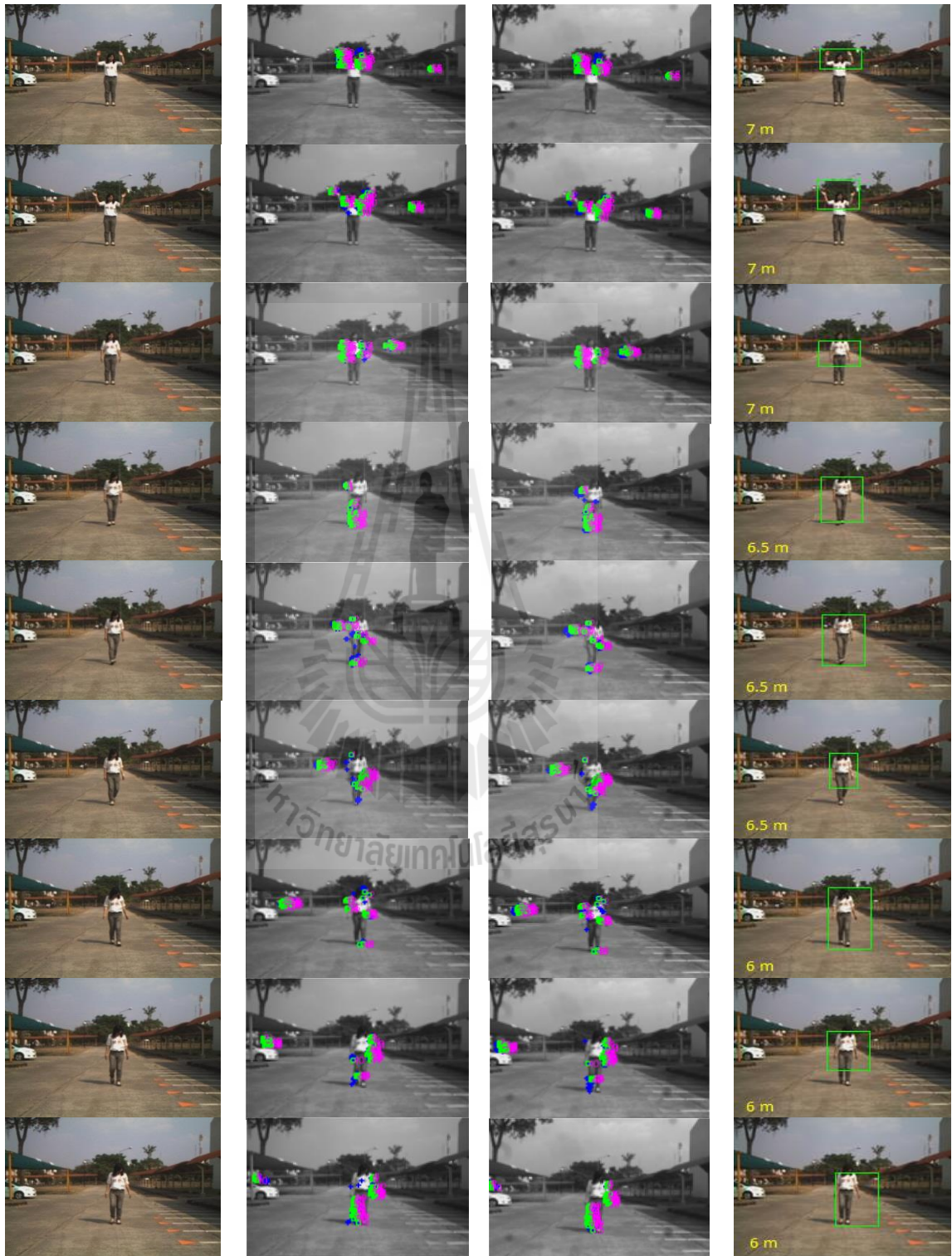
ตัวอย่างสภาพแวดล้อม	จำนวนเฟรมทั้งหมด	จำนวนเฟรมที่ผิดพลาด	เปอร์เซ็นต์ความผิดพลาด
1	1000	95	9.5
2	1000	98	9.8
3	500	40	8



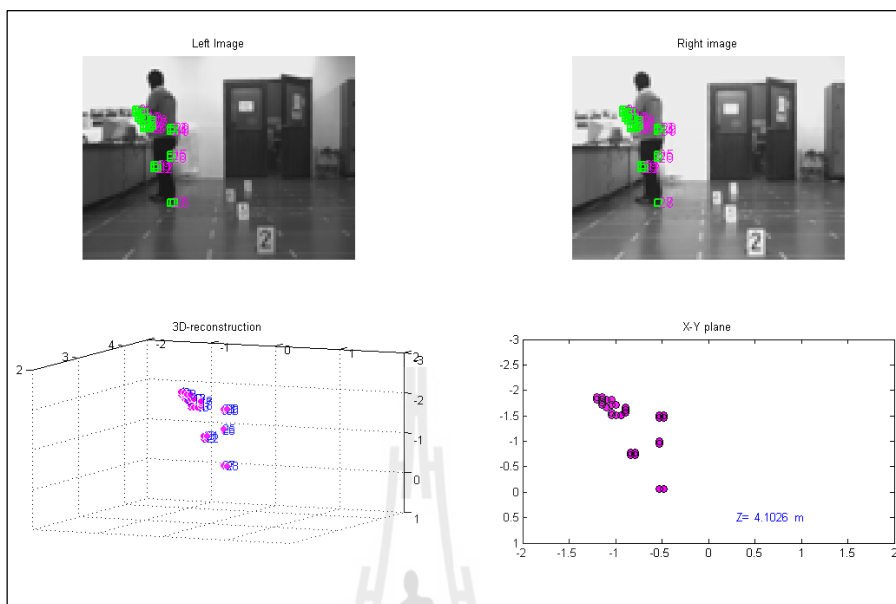
รูปที่ 4.16 ผลการทดสอบการหาระยะวัตถุภายในอาคาร



รูปที่ 4.17 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร



รูปที่ 4.18 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร



รูปที่ 4.19 ผลการทดสอบหาพิกัด 3 มิติของคนเคลื่อนที่ภายในอาคาร

จากรูปที่ 4.19 ระบบสามารถคำนวณหาพิกัดการเคลื่อนที่ของมนุษย์ได้ เมื่อผ่านระบบการบูรณะสถานะแวดล้อม 3 มิติแล้ว จะเห็นได้ว่า ภาพจริงที่เกิดขึ้น มีการเคลื่อนที่ไปทางซ้ายของรูป เมื่อทำการจับคู่การไหลเชิงแสงของภาพซ้ายและขวาแล้วจะได้พิกัด 3 มิติออกมาดังรูป จากสูตรความสัมพันธ์ของสมการที่ 4.2 - 4.4 จะเห็นได้ว่า ถ้าสมการที่ 4.2 ซึ่งเป็นสมการหาระยะทางระหว่างวัตถุและกล้อง (Z) หาระยะผิดพลาดจะทำให้ระนาบตำแหน่งในแกน X และ Y ผิดพลาดไปด้วย ซึ่งรูปที่ 4.19 ได้แสดงถึงความแม่นยำในการจับคู่การไหลเชิงแสง ซึ่งตำแหน่งที่เกิดจากการพลอตกราฟในระนาบ X และ Y เป็นจุดเดียวกับการเกิดการไหลเชิงแสงที่ถูกจับคู่

จากการทดสอบที่ผ่านมาจะเห็นได้ว่าระบบยังไม่สามารถตรวจสอบหาระยะทางของวัตถุได้ทั้งหมด สามารถสรุปถึงสาเหตุความผิดพลาดของระบบทั้งหมดได้ดังนี้

1. การป้อนอินพุตให้กับระบบเครือข่ายซอฟต์แวร์จำกัด เนื่องจากหน่วยความจำของคอมพิวเตอร์ที่ใช้ในการทดสอบครั้งนี้มีจำกัด เมื่อทำการประมวลผลในขั้นตอนของเครือข่ายประสาทเทียมสามารถป้อนอินพุตได้เพียง 100x 100 นิวรอนหรือจำนวนนิวรอนรวมกันแล้วไม่เกิน 200 นิวรอน ซึ่งการไหลเชิงแสงของวัตถุในภาพมีมากกว่านั้น ซึ่งเมื่อทำการกรองขนาดออกไประดับหนึ่งอาจจะทำให้ข้อมูลหายไปบางส่วนจึงทำให้ระบบประมวลผลแม่นยำน้อยลง

2. ระยะห่างระหว่างวัตถุและกล้องที่เหมาะสมที่ได้จากการคำนวณคือ 7 เมตร ดังนั้นเมื่อวัตถุเคลื่อนที่ออกห่างจากกล้องมากกว่า 7 เมตร ถึงแม้ในขั้นตอนของการจับคู่จะประมวลผลได้แต่เมื่อนำมาคำนวณหาระยะทางแล้วจะทำให้เกิดค่าผิดพลาดมากขึ้น

3. การเคลื่อนที่ของวัตถุที่อยู่ชิดขอบทางด้านซ้ายหรือขวามากเกินไป หากวัตถุเคลื่อนที่อยู่ในตำแหน่งทางด้านซ้ายของกล้องหรือขวาของกล้องมากเกินไปจะทำให้ไม่สามารถตรวจจับวัตถุได้ถึงแม้จะเข้ามาในระยะใกล้ๆก็ตาม



บทที่ 5

สรุปและข้อเสนอแนะ

5.1 สรุป

งานวิจัยนี้ได้ดำเนินการศึกษาและพัฒนาการบูรณะสภาวะแวดล้อมสามมิติเพื่อสามารถนำไปประยุกต์ใช้ในการตรวจจับวัฏระยะของวัตถุที่เคลื่อนที่แบบอัตโนมัติ ซึ่งระบบที่พัฒนาขึ้นประกอบไปด้วยระบบการตรวจจับการเคลื่อนที่ของวัตถุโดยใช้การไหลเชิงแสง การแก้ปัญหาความสอดคล้องกันของจุดภาพของกล้องสองตัว โดยใช้เครือข่ายซอฟต์แวร์ซึ่งเป็นเครือข่ายประสาทเทียมชนิดหนึ่ง และการจำลองพิกัดสามมิติของวัตถุที่เคลื่อนที่ภายในภาพ โดยการดำเนินงานวิจัยดังกล่าวสำเร็จคล่องตามวัตถุประสงค์ โดยสามารถสรุปผลการศึกษาวิจัยและพัฒนาทางวิศวกรรมเป็นข้อสรุปได้ดังต่อไปนี้

การปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง ได้รายงานไว้ในบทที่ 2 ผลการสืบค้นวรรณกรรมย้อนหลังที่เกี่ยวข้องกับงานวิจัยที่จะดำเนินการจากฐานข้อมูล IEEE, IEE, ScienceDirect และอื่นๆ ซึ่งทำให้ทราบถึงแนวทางการวิจัยที่เกี่ยวข้อง ระเบียบวิธีการวิจัยที่ผู้วิจัยอื่นๆ ได้นำมาใช้ ผลการดำเนินงานข้อเสนอแนะต่าง ๆ ของนักวิจัยตั้งแต่อดีตจนถึงปัจจุบัน ซึ่งประกอบไปด้วยการหาแผนที่บ่งบอกความลึก และการหาระยะห่างระหว่างวัตถุกับกล้อง โดยใช้กล้องสองตัว และพื้นฐานการวิเคราะห์ประมวลผลภาพ

ในบทที่ 3 แสดงรายละเอียดของการแก้ปัญหาความสอดคล้องกันของจุดภาพ โดยใช้เครือข่ายประสาทซอฟต์แวร์ 2 มิติ โดยแสดงรายละเอียดโครงสร้าง การเรียนรู้และวิธีการแก้ปัญหาของเครือข่าย ซึ่งระบบเครือข่ายซอฟต์แวร์สามารถแก้ปัญหาความสอดคล้องกันของจุดภาพจากกล้อง 2 ตัวได้ โดยมีอุปสรรคที่เกิดจากตำแหน่งที่น่าสนใจเป็นตัวเริ่มกระบวนการ หลังจากนั้นเครือข่ายซอฟต์แวร์จะทำการวนรอบเพื่อหาคำตอบจนได้จุดที่สอดคล้องกันของภาพจากกล้องซ้ายและกล้องขวา โดยการทดสอบครั้งนี้ใช้วิธีการของ Harris ในการหาจุดที่น่าสนใจ ซึ่งไม่เหมาะกับการตรวจจับ

ภาพที่มีการเคลื่อนที่ของวัตถุ อินพุตของระบบที่จะใช้ในขั้นตอนต่อไปจึงเปลี่ยนเป็นการไหลเชิงแสงแทน

ในบทที่ 4 แสดงรายละเอียดโครงสร้างของระบบการบูรณะแวล้อม 3 มิติ โดยเครือข่ายประสาทเทียมร่วมกับการไหลเชิงแสง ซึ่งแสดงวิธีและขั้นตอนในการแก้ปัญหาของระบบ เริ่มตั้งแต่การป้อนลำดับภาพ การประมวลผลภาพ การหาการไหลเชิงแสง การแก้ปัญหาความสอดคล้องกันของจุดภาพ จนกระทั่งการคำนวณหาพิกัด 3 มิติ ระบบการหาระยะระหว่างกล้องและวัตถุโดยใช้เทคนิคเครือข่ายประสาทเทียมร่วมกับการไหลเชิงแสง สามารถหาจุดที่สอดคล้องกันของภาพซ้ายและขวาได้โดยอาศัยตำแหน่ง มุมและขนาดของการไหลเชิงแสงเป็นอินพุตให้กับระบบ เมื่อระบบทำการประมวลผลแล้วจะทำให้ทราบพิกัดของวัตถุที่เคลื่อนที่ภายในภาพ โดยอาศัยจุดกึ่งกลางและระยะห่างของกล้องและวัตถุเป็นจุดอ้างอิงพิกัดของวัตถุที่เกิดขึ้นเมื่อเทียบกับระยะจริงแล้วมีความแม่นยำอยู่ที่ 90 เปอร์เซ็นต์ทั้งภายในและภายนอกอาคาร ซึ่งค่าความผิดพลาดเกิดจากการป้อนอินพุตให้กับระบบเครือข่ายซอฟท์แวร์จำกัด ระยะห่างระหว่างวัตถุและกล้องที่เหมาะสมอยู่ที่ 7 เมตร และการเคลื่อนที่ของวัตถุที่อยู่ชิดขอบทางด้านซ้ายหรือขวามากเกินไป

5.2 ข้อเสนอแนะ

จากการดำเนินงานวิจัยที่ผ่านมาและผลที่ได้ทำให้เกิดแนวคิดและข้อเสนอแนะในการดำเนินงานวิจัยต่อไปในอนาคตดังต่อไปนี้

1. เพื่อให้ระบบสามารถนำไปประยุกต์ใช้งานได้หลากหลายควรทดสอบกับสภาพแวดล้อมต่างๆ เพื่อนำไปประยุกต์ใช้ได้เหมาะสม
2. เพื่อให้ระบบมีประสิทธิภาพที่ดีขึ้นควรปรับปรุงแก้ไขขั้นตอนที่ทำให้เกิดค่าความคลาดเคลื่อน โดยปรับปรุงพารามิเตอร์ต่างๆ เพื่อให้สอดคล้องกับสภาพแวดล้อมที่ทำการทดสอบ และทำให้ระบบทำงานได้เร็วขึ้น เพื่อให้ระบบสามารถทำงานได้ทันกับเวลาจริง (real time) รวมถึงการปรับเทียบ (calibration) กล้องทั้ง 2 ตัว ให้ได้คุณลักษณะภายในที่เหมือนกันให้มากที่สุด
3. เพื่อให้ได้ผลการหาจุดสอดคล้องที่ดียิ่งขึ้น สามารถประยุกต์นำเอากระบวนการหาค่าที่เหมาะสมที่สุดมาใช้ โดยเฉพาะเทคนิคทางปัญญาประดิษฐ์ จะช่วยลดความผิดพลาดของการจับคู่ได้

รายการอ้างอิง

- S.T. Barnard and M.A. Fischer. (1982). Computational Stereo. **ACM Computing Surveys**. 14(4):553-572
- B. D. Lucas and T. Kanade. (1981). An iterative image registration technique with an application to stereo vision. **Proceedings of Imaging Understanding Workshop**, 121-130.
- U.R. Dhond and J.K. Aggarwal. (1989). Structure from stereo : a review. **IEEE Transaction on Systems, Man, and Cybernetics**, (19): 1489-1510.
- D.V. Papadimitriou and T.J. Dennis. (1996). Epipolar line estimation and rectification for stereo image pairs. **IEEE Transaction on Image Process**. 5(4): 672-676.
- H. Kim, S. Yang and K. Sohn. (2003). 3D Reconstruction of stereo image for interaction between real and virtual worlds. **Proceedings of 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)**. IEEE.
- L. MaTThies, T. Kanade and R. Szeliski. (1989). Kalman filter-based algorithms for estimating depth from image sequence. **International Journal of Computer Vision**. (3):209-236.
- M. Montemerlo and S. Thrum. (2004). Large-scale robotic 3D mapping of urban structures. **Proceedings of 9th International Symposium on Experimental Robotics**.
- M. Pollefeys and L. Van Gool. (1997). Self-calibration and matrix reconstruction in spite of varying and unknow internal camera parameter. **TR9707,ESAT-MI2, K.U. Leuven**.
- Nasser M. Nasrabadi. (1992). Hopfield network for stereo vision correspondence. **IEEE Transaction on Neural Networks**. (3).
- O. Faugeras. (2001). Three-dimension computer vision : a geometric viewpoint. **MIT Press**. London.
- R. Hartley and A. Zisserman. (2000). Multiple view geometry in computer vision. **Cambridge**.
- อาทิตย์ ศรีแก้ว. (2010). **ปัญญาเชิงคำนวณ**. มหาวิทยาลัยเทคโนโลยีสุรนารี.



ภาคผนวก ก

บทความที่ได้รับการตีพิมพ์เผยแพร่

รายชื่อบทความที่ได้รับการตีพิมพ์เผยแพร่

Arthit Srikaew, Prayoth Kumsawat, Kitti Attakitmongcol and Somying Phanwilai, “Stereo Correspondence using 2-Dimensional Hopfield Neural Network and Optical Flow”, *International Conference on Engineering, Applied Sciences, and Technology*, Bangkok Thailand, November 21-24, 2012, pages 201-204.



ภาคผนวก ข

ประวัติผู้วิจัย



ประวัติผู้วิจัย

ดร.อาทิตย์ ศรีแก้ว เกิดเมื่อวันที่ 19 พฤศจิกายน พ.ศ. 2515 สำเร็จการศึกษาระดับปริญญาตรีใน สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์จาก สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อ พ.ศ. 2537 สำเร็จการศึกษาระดับ ปริญญาโทและปริญญาเอกสาขาวิศวกรรมไฟฟ้าจาก Vanderbilt University ประเทศสหรัฐอเมริกา เมื่อ พ.ศ. 2540 และ 2543 ตามลำดับ ปัจจุบันดำรงตำแหน่งรองศาสตราจารย์ สาขาวิชาวิศวกรรม ไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี มีความสนใจงานวิจัยทางด้านการมองเห็นของคอมพิวเตอร์และหุ่นยนต์ การประมวลผลภาพ และระบบทางปัญญาประดิษฐ์

นางสาวสมหญิง พันธุ์ไธ เกิดเมื่อวันที่ 28 กันยายน พ.ศ. 2529 ที่อำเภอเทิง จังหวัดเชียงราย สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายจากโรงเรียนเทิงวิทยาคม จังหวัดเชียงราย สำเร็จการศึกษาระดับปริญญาวิทยาศาสตรบัณฑิต (วิศวกรรมไฟฟ้า) จากมหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา เมื่อปีการศึกษา 2551 จากนั้นได้เข้าศึกษาต่อในระดับปริญญาโท สาขาวิชาวิศวกรรมไฟฟ้า ที่มหาวิทยาลัยเทคโนโลยีสุรนารี ขณะศึกษาระดับปริญญาโทได้เป็นผู้สอนและร่วมจัดทำรายการค่ายอิเล็กทรอนิกส์รุ่นเยาว์ (e-Camp) ในปี 2010 2011 และ 2012

