การบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึกเชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง

นางสาวสมหญิง พันธุ์วิไล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2555

3D ENVIRONMENT RECONSTRUCTION USING ARTIFICIAL NEURAL NETWORK BASED DEPTH MAPS AND OPTICAL FLOW

Somying Phanwilai

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Telecommunication Engineering

Suranaree University of Technology

Academic Year 2012

การบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึก
เชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

____
(ผศ. ดร.อนันท์ อุ่นศิวิไลย์)
ประธานกรรมการ

____
(รศ. ดร.อาทิตย์ ศรีแก้ว)
กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)

____
(ผศ. ร.อ. ดร.ประโยชน์ คำสวัสดิ์)
กรรมการ

____
(รศ. ดร.กิตติ อัตถกิจมงคล)
กรรมการ

____
(ศ. ดร.ชูกิจ ลิมปิจำนงค์)
รองอธิการบดีฝ่ายวิชาการ

____
(รศ. ร.อ. ดร.กนต์ธร ชำนิประศาสน์)
คณบดีสำนักวิชาวิศวกรรมศาสตร์

สมหญิง พันธุ์วิไล : การบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึกเชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง (3D ENVIRONMENT RECONSTRUCTION USING ARTIFICIAL NEURAL NETWORK BASED DEPTH MAPS AND OPTICAL FLOW) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว, 117 หน้า.

การบูรณะสภาวะแวดล้อม 3 มิติโดยใช้กล้องแบบสเตอริโอ เป็นเทคโนโลยีที่นำมาใช้กับหุ่นยนต์เคลื่อนที่แบบอัตโนมัติ ระบบการมองเห็นของคอมพิวเตอร์ทำหน้าที่เป็นระบบการมองเห็นของหุ่นยนต์ โดยมีงานวิจัยมากมายพัฒนาระบบการมองเห็นของหุ่นยนต์ให้มีประสิทธิภาพที่ดีขึ้น ระบบการมองเห็นแบบสเตอริโอ ตัวสามารถหาระยะห่างระหว่างวัตถุและตัวกล้องได้โดยไม่ต้องอาศัยตัวตรวจรู้ใด ๆ มีระบบการทำงานพื้นฐานคือ การหาแผนที่ความไม่เท่ากัน (disparity maps) หรือแผนที่บ่งบอกความลึก (depth maps) เพื่อหาระยะห่างระหว่างวัตถุกับกล้องและการบูรณะสภาวะ 3 มิติ (3D reconstruction) หรือการแปลงพิกัดตำแหน่งสภาวะแวดล้อมในภาพให้เป็นพิกัดตำแหน่งของสภาวะแวดล้อมจริง ปัญหาของการหาแผนที่ความไม่เท่ากันคือการคำนวณหาความสอดคล้องภายในจุดภาพจากกล้องทั้ง 2 ตัว ในงานวิจัยนี้ใช้เครือข่ายฮอพฟิลด์ (Hopfield Network) ร่วมกับการไหลเชิงแสงเพื่อหาความสอดคล้องกันของจุดภาพ แล้วนำค่าที่ได้ไปบูรณะสภาวะ 3 มิติโดยการคำนวณเชิงเรขาคณิตเพื่อหาพิกัดจุดจริงของภาพที่ได้จาก 2 มิติเปลี่ยนแปลงเป็น 3 มิติ โดยระบบสามารถคำนวณหาระยะทางระหว่างกล้องและวัตถุได้อย่างมีประสิทธิภาพ เมื่อระบบทำการประมวลผลจะทำให้ทราบพิกัดของวัตถุที่เคลื่อนที่ภายในภาพ โดยอาศัยจุดกึ่งกลางและระยะห่างของกล้องและวัตถุเป็นจุดอ้างอิงพิกัดของวัตถุที่เกิดขึ้น เมื่อเทียบกับระยะจริงมีค่าความถูกต้องประมาณ 90 เปอร์เซนต์ทั้งภายในและภายนอกอาคาร

สาขาวิชา วิศวกรรมไฟฟ้า	ลายมือชื่อนักศึกษา_____________________
ปีการศึกษา 2555	ลายมือชื่ออาจารย์ที่ปรึกษา_______________

SOMYING PHANWILAI : 3D ENVIRONMENT RECONSTRUCTION USING ARTIFICIAL NEURAL NETWORK BASED DEPTH MAPS AND OPTICAL FLOW. THESIS ADVISOR : ASSOC. PROF. ARTHIT SRIKAEW, Ph.D., 117 PP.

3D ENVIRONMENT RECONSTRUCTION/HOPFIELD NEURAL NETWORK/STEREO VISION/STEREO CORRESPONDENCE/OPTICAL FLOW

3D environment reconstruction using stereo camera is one of the most crucial part in mobile robot development. The recent technology of computer vision has played an important role for robot vision system. Stereo vision system provides the ability to calculate a distance between an object in front of two cameras mounted in parallel which is essentially not required any other sensor. Stereo correspondence problem has then become one of the most challengeable tasks in stereo vision system. In order to calculate this distance, a corresponding portion of interesting object in each image has to be detected and matched. This work presents a hybrid approach of using two-dimensional Hopfield network to perform a stereo correspondence calculation from optical flow features in stereo images. The optical flow features from both left and right image are considered as the input for the network which has an advantage of a capability to detect object motion. Using such optical flow and a 2D network with ability to learn and recognize patterns, optical flow feature in each image can be matched efficiently. The proposed system demonstrates desirable results which can be mainly and effectively used in 3D

reconstruction at accuracy rate of object between stereo cameras distance about 90 percent both indoor and outdoor.

School of Electrical Engineering Student's Signature__________________

Academic Year 2012 Advisor's Signature__________________

กิตติกรรมประกาศ

วิทยานิพนธ์นี้ดำเนินการสำเร็จลุล่วงด้วยดี ผู้วิจัยขอกราบขอบพระคุณ บุคคลและกลุ่มบุคคลต่าง ๆ ที่ได้กรุณาให้คำปรึกษา แนะนำ ช่วยเหลือ อย่างดียิ่ง ทั้งในด้านวิชาการและการดำเนินงานวิจัย รวมถึงหน่วยงานต่าง ๆ ที่ช่วยอำนวยความสะดวกในการทำงานวิจัย อาทิ

รองศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษาแนะนำและแนะแนวทางอันเป็นประโยชน์ยิ่งต่องานวิจัย รวมถึงได้ช่วยตรวจทานและแก้ไขรายงานวิทยานิพนธ์เล่มนี้จนทำให้มีความสมบูรณ์ยิ่งขึ้น รวมทั้งเป็นกำลังใจให้กับผู้วิจัยเสมอมา

ผู้ช่วยศาสตราจารย์ ดร.อนันท์ อุ่นศิวิไลย์ รองศาสตราจารย์ ดร.กิตติ อัตถกิจมงคล ผู้ช่วยศาสตร์ตราจารย์ เรืออากาศเอก ดร.ประโยชน์ คำสวัสดิ์ และคณาจารย์ประจำสาขาวิชาวิศวกรรมไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่กรุณาให้คำปรึกษาด้านวิชาการเป็นอย่างดียิ่งมาโดยตลอด

ขอขอบคุณคุณบุคลากรสาขาวิชาวิศวกรรมไฟฟ้า และศูนย์เครื่องมือวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีสุรนารีทุกท่านที่อำนวยความสะดวกในการทำงานของข้าพเจ้า

ขอบคุณพี่ ๆ เพื่อน ๆ น้อง ๆ บัณฑิตศึกษาทุกท่าน รวมถึงมิตรสหายที่คอยถามไถ่และให้กำลังใจในการทำวิจัยมาโดยตลอด

ท้ายนี้ผู้วิจัยขอพระขอบคุณอาจารย์ผู้สอนทุกท่าน ที่ประสิทธิ์ประสาทวิชาความรู้ทางด้านต่าง ๆ ทั้งในอดีตและปัจจุบัน ขอกราบขอบพระคุณ คุณพ่อคุณแม่ รวมถึงญาติพี่น้องของผู้วิจัยทุกท่านที่ได้ให้ความรัก กำลังใจความอบอุ่น ความห่วงใย การอบรมเลี้ยงดูและให้การสนับสนุนทางด้านการศึกษาอย่างดียิ่งมาโดยตลอด

สมหญิง พันธุ์วิไล

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย) .... ก
บทคัดย่อ (ภาษาอังกฤษ).. ข
กิตติกรรมประกาศ.. ง
สารบัญ .... จ
สารบัญตาราง .... ฌ
สารบัญรูป .... ญ

บทที่

1 บทนำ 1
1.1 ความสำคัญของปัญหา 1
1.2 วัตถุประสงค์การวิจัย.. 2
1.3 ข้อตกลงเบื้องต้น 2
1.4 ขอบเขตของการวิจัย 2
1.5 วิธีการดำเนินงาน 2
1.6 ระเบียบวิธีวิจัย 3
1.7 เครื่องมือที่ใช้ในงานวิจัย.. 3
1.8 การเก็บรวบรวมข้อมูล 4
1.9 การวิเคราะห์ข้อมูล.. 4
1.10 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย.. 4

2 ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง 4
2.1 กล่าวนำ.. 5
2.2 การออกแบบระบบการมองเห็นของกล้อง 2 ตัว 5
2.2.1 ระบบการมองเห็นของกล้อง 2 ตัว.. 5
2.2.2 การทำงานของกล้องดิจิตอล.. 8
2.2.3 ความยาวโฟกัสของเลนส์ 9
2.3 การประมวลผลภาพเบื้องต้น.. 11

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญ(ต่อ)

หน้า

2.3.1 แบบจำลองภาพระดับเทา (Gray scale model) ... 12
2.3.2 แบบจำลองสี RGB ... 13
2.3.3 การหามุมด้วยวิธีของฮาร์ริส ... 14
2.4 การไหลเชิงแสง ... 15
2.4.1 Gradient Base Method ... 15
2.4.2 Region Base Method ... 18
2.5 เครือข่ายประสาทเทียม ... 19
2.5.1 ประวัติความเป็นมาของโครงข่ายประสาทเทียม ... 20
2.5.2 ความหมายและหลักการของโครงข่ายประสาทเทียม ... 21
2.5.3 ลักษณะของโครงข่ายประสาทเทียม ... 23
2.5.4 ประเภทของการเรียนรู้ของโครงข่ายประสาทเทียม ... 25
2.5.5 การประยุกต์ใช้งานโครงข่ายประสาทเทียม ... 25
2.6 เครือยข่ายฮอพฟิลด์ ... 26
2.6.1 โครงสร้างของเครือข่ายฮอพฟิลด์ ... 26
2.6.2 การเรียนรู้ของเครือข่ายฮอพฟิลด์ ... 27
2.6.3 การปรับค่าน้ำหนักประสาทของเครือข่ายฮอพฟิลด์ ... 29
2.7 ทฤษฎีเสถียรภาพเลียปูนอฟ ... 32
2.7.1 ทฤษฎีเลียปูนอฟจากการประมาณแบบเชิงเส้น ... 33
2.7.2 ทฤษฎีเลียปูนอฟจากการวิเคราะห์โดยตรง ... 34
2.8 งานวิจัยที่เกี่ยวข้อง ... 35
2.9 สรุป ... 37

3 การแก้ไขปัญหาความสดอคล้องกันของจุดภาพโดยใช้เครือข่ายประสาทฮอพฟิลด์ 2 มิติ ... 38

3.1 กล่าวนำ ... 38
3.2 โครงสร้างของเครือข่ายฮอพฟิลด์ 2 มิติ ... 38
3.3 การเรียนรู้ของเครือข่ายฮอพฟิลด์ 2 มิติ ... 39
3.4 การแก้ปัญหาความสอดคล้องกันของจุดภาพ ... 40

สารบัญ(ต่อ)

หน้า

3.5 ผลการทดสอบการแก้ปัญหาความสอดคล้องกันของจุดภาพ
โดยใช้เครือข่ายประสาทฮอพฟิลด์ 2 มิติ 44
3.6 สรุป 47
4 การบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธีแผนที่บ่งบอกความลึก
เชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง 48
4.1 กล่าวนำ 48
4.2 การคัดแยกวัตถุด้วยการไหลเชิงแสง 49
4.3 การแก้ปัญหาความสอดคล้องกันของจุดภาพ 53
4.4 การหาระยะห่างระหว่างกล้องกับวัตถุและการบูรณะสภาวะแวดล้อม 3 มิติ 58
4.5 ผลการทดสอบการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ 63
4.5.1 วัตถุเคลื่อนที่ภายในอาคาร 64
4.5.2 วัตถุเคลื่อนที่ภายนอกอาคาร 64
4.6 สรุป 70
5 วิเคราะห์และอภิปรายผลการทดลอง 71
5.1 กล่าวนำ 71
5.2 โครงสร้างของระบบการบูรณะสภาวะแวดล้อม 3 มิติ 71
5.3 ผลการทดสอบการคัดกรองการไหลเชิงแสง 72
5.4 ผลการทดสอบการจับคู่ด้วยเครือข่ายฮอพฟิล์ด 2 มิติ 77
5.5 การบูรณะสภาวะแวดล้อม 3 มิติ 80
5.6 ผลการหาระยะทางในสภาวะต่างๆ 84
5.7 สรุป 88
6 บทสรุปและข้อเสนอแนะ 90
6.1 สรุป 90
6.2 ข้อเสนอแนะ 91
รายการอ้างอิง 92

สารบัญ(ต่อ)

หน้า

ภาคผนวก

ภาคผนวก ก. ภาพต่างๆที่ใช้ในการทดสอบ 122

ภาคผนวก ข. ภาพผลการทดสอบการจับคู่ตำแหน่งในสภาวะต่างๆ 122

ภาคผนวก ค. บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในขณะศึกษา 122

ประวัติผู้เขียน 120

สารบัญตาราง

ตารางที่ **หน้า**

2.1 ค่ามาตรฐานของ Common Values ของภาพดิจิตอล 2 มิติ 12

4.1 ค่าความผิดพลาดของการจับคู่ภาพของกล้องซ้ายและขวา ในชุดตัวอย่างที่ 1-10 58

4.2 ข้อมูลที่นำมาใช้ในการคำนวณหาระยะทางจากผลการทดลอง ในตัวอย่างที่ 1 60

4.3 เปอร์เซนต์ความถูกต้องของระยะห่างระหว่างวัตถุกับกล้องที่คำนวณได้ กับค่าระยะห่างจริง 61

5.1 ความสามารถในการจับคู่ภาพของระบบและงานวิจัยที่เกี่ยวข้องในอดีต 80

5.2 ตารางแสดงระยะทางที่วัดได้สูงสุดในการติดตั้งกล้องและ ขนาดภาพต่าง ๆที่ป้อนให้กับระบบ (เมตร) 86

5.3 ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในระยะต่าง ๆ 86

5.4 ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่าง ๆ 88

สารบัญรูปภาพ

รูปที่ หน้า

2.1 การติดตั้งกล้องสองตัวแบบขนาน 5

2.2 ภาพจากกล้อง 2 ตัวติดตั้งแบบขนาน 6

2.3 องค์ประกอบของระบบการมองแบบกล้อง 2 ตัว 7

2.4 ลักษณะทางเรขาคณิตของกล้องสองตัว 8

2.5 กล้องดิจิตอลพร้อมเลนส์ 9

2.6 เลนส์ชนิดต่างๆ 10

2.7 จุดโฟกัสในกล้องดิจิตอลตามแกน x, y 10

2.8 ภาพในเชิงดิจิตอลเมื่อแบ่งค่าพิกเซลขนาด [m = 11, n = 4] 12

2.9 ค่าระดับเทา 13

2.10 ลูกบาศก์ของแบบจำลองสี RGB 14

2.11 การหามุมด้วยวิธีของฮาร์ริส 14

2.12 การไหลเชิงแสง 19

2.13 โครงสร้างระบบประสาท 22

2.14 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว 24

2.15 โครงข่ายประสาทเทียมแบบหลายชั้น 24

2.16 โครงสร้างของเครือข่ายฮอพฟิลด์ 26

2.17 แผนภาพค่าน้ำหนักประสาท 28

2.18 การไหลเวียนของพลังงานและการเปลี่ยนสถานะของเครือข่าย 31

2.19 พื้นผิวพลังงานเครือข่ายฮอพฟิล์ด 32

3.1 โครงสร้างเครือข่ายประสาทและค่าน้ำหนักประสาทของเครือข่ายฮอพฟิลด์ 39

3.2 C_{ikjl} function (non-linefunction) 42

สารบัญรูปภาพ (ต่อ)

รูปที่ **หน้า**

3.3 C_{ikjl} function โดยกำหนดให้ $\lambda = 1$ และ $\theta = 10$ 42

3.4 ระบบการทำงานของเครือข่ายฮอพฟิลด์ 43

3.5 ตัวอย่างการแก้ปัญหาด้วยเครือข่ายฮอพฟิลด์ 2 มิติ 45

3.6 ตัวอย่างการแก้ปัญหาด้วยเครือข่ายฮอพฟิลด์ 2 มิติ 46

4.1 โครงสร้างระบบการบูรณะสภาวะแวดล้อม 3 มิติ 48

4.2 การปรับขนาดและระดับสีของภาพต้นแบบ 49

4.3 การไหลเชิงแสง 50

4.4 ภาพขยายตำแหน่งการไหลเชิงแสงของภาพ 51

4.5 ความถี่ของตำแหน่งการเกิดการไหลเชิงแสงของภาพต้นแบบ 4.2 (ก 52

4.6 ตำแหน่งการเกิดการไหลเชิงแสงที่ทำการกรองแล้ว 52

4.7 ตัวอย่างการหาตำแหน่งการเกิดการไหลเชิงแสง 53

4.8 การพิจารณาค่า ΔD และ Δd 55

4.9 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 1 56

4.10 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 2 56

4.11 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 3 56

4.12 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 4 57

4.13 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 5 57

4.14 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 6 57

4.15 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 7 58

4.16 ตำแหน่งวัตถุที่ระยะต่างๆในระนาบ 3 มิติ 62

4.17 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ XY 62

สารบัญรูปภาพ (ต่อ)

รูปที่ **หน้า**

4.18 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ ZX 63
4.19 ผลการทดสอบการหาระยะวัตถุภายในอาคาร 65
4.20 ผลการทดสอบการหาระยะวัตถุภายในอาคาร 66
4.21 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร 67
4.22 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร 68
4.23 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร 69
5.1 โครงสร้างของระบบโดยรวม 71
5.2 การไหลเชิงแสงของลูคัส-คานาเด 72
5.3 การไหลเชิงแสงของลูคัส-คานาเด 73
5.4 เปรียบเทียบขนาดการไหลเชิงแสงในสภาวะต่างๆเมื่อกล้องถูกติดตั้งอยู่กับที่ 74
5.5 เปรียบเทียบขนาดการไหลเชิงแสงในสภาวะต่างๆเมื่อกล้องเคลื่อนที่ 75
5.6 ผลการคัดกรองการไหลเชิงแสง 76
5.7 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร 77
5.8 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร 78
5.9 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร 78
5.10 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายนอกอาคาร 79
5.11 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายนอกอาคาร 79
5.12 การบูรณะสภาวะแวดล้อม 3 มิติ 81
5.13 ผลการทดลองการคำนวณหาตำแหน่งในระนาบ 3 มิติ 82
5.14 ภาพที่เกิดจากการติดตั้งกล้องระยะห่างต่างๆ 84
ก.1 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 5 เซนติเมตร 94

สารบัญรูปภาพ (ต่อ)

รูปที่ **หน้า**

ก.2 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 15 เซนติเมตร 97
ก.3 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร 99
ก.4 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 25 เซนติเมตร 101
ก.5 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร 103
ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง........106

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

เทคโนโลยีการมองเห็นของคอมพิวเตอร์ (computer vision) เป็นเทคโนโลยีอย่างหนึ่งที่นำมาใช้ในงานหุ่นยนต์เคลื่อนที่อย่างแพร่หลาย ซึ่งส่วนใหญ่ใช้ในการพัฒนาเกี่ยวกับการหาตำแหน่งของวัตถุและสร้างเส้นทางเดินให้กับหุ่นยนต์ โดยการติดตั้งกล้องเข้ากับตัวหุ่นยนต์เพื่อแทนระบบการมองเห็นหรือแทนตาของหุ่นยนต์ ดังนั้นระบบการมองเห็นแบบกล้อง 2 ตัว (stereo vision) จึงเป็นเทคโนโลยีที่สำคัญอย่างหนึ่งในการพัฒนาเทคโนโลยีทางด้านหุ่นยนต์ ระบบการมองเห็นแบบกล้อง 2 ตัวของหุ่นยนต์เคลื่อนที่ ได้มีการค้นคว้าและพัฒนาวิธีการแบบใหม่ขึ้นตลอดเวลาเพื่อทำให้ระบบมีประสิทธิภาพที่ดีขึ้นและลดข้อจำกัดด้านต่าง ๆ ลง ซึ่งได้รับความสนใจอย่างมาก สำหรับเหตุผลที่ทำให้งานนี้น่าสนใจ คือ ข้อดีที่สำคัญของระบบการมองเห็นกล้อง 2 ตัว อยู่ที่การหาตำแหน่งและระยะห่างของวัตถุเพื่อใช้ในการหลบหลีกและนำทางของหุ่นยนต์โดยปราศจากการทำงานร่วมกับอุปกรณ์เครื่องตรวจวัดต่าง ๆ เช่น เครื่องรับรู้ (sensor) เลเซอร์ เป็นต้น

ระบบการมองเห็นแบบกล้อง 2 ตัวมีระบบการทำงานพื้นฐานคือ การหาแผนที่ความไม่เท่ากัน (disparity maps) หรือแผนที่บ่งบอกความลึก (depth maps) เพื่อหาระยะห่างระหว่างวัตถุกับกล้องและการบูรณะสภาวะ 3 มิติ (3D reconstruction) หรือการแปลงพิกัดตำแหน่งสภาวะแวดล้อมในภาพให้เป็นพิกัดตำแหน่งของสภาวะแวดล้อมจริง หลักการสำคัญในการหาแผนที่บ่งบอกความลึก อยู่ที่การคำนวณหาความสอดคล้องภายในภาพของกล้องทั้งสองตัว(stereo correspondence) การหาความสอดคล้องกันของจุดภาพต้องใช้จุดเด่นของทั้งสองภาพเป็นจุดอ้างอิงซึ่งจุดที่ได้จะกระจายอยู่ทั่วทั้งภาพ ซึ่งบ้างจุดก็ไม่จำเป็นในการนำมาคำนวณ และยังเหมาะกับภาพนิ่ง ซึ่งงานวิจัยนี้ได้นำมาประยุกต์ใช้กับภาพเคลื่อนไหว ดังนั้นจึงใช้ขนาดของการไหลเชิงแสงเป็นตัวคัดกรองจุดเด่นของภาพแล้วนำไปคำนวณหาความสอดคล้องภายในภาพของกล้องทั้งสองตัวด้วยวิธีการเครือข่ายฮอพฟิลด์สองมิติ (2D Hopfield Network)

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อออกแบบและพัฒนาอัลกอลิทึมในการหาความลึกของวัตถุในภาพด้วยระบบการ มองเห็นแบบสเตอริโอ

1.2.2 เพื่อออกแบบและพัฒนาอัลกอลิทึมบูรณะสภาวะแวดล้อม 3 มิติได้

1.3 ข้อตกลงเบื้องต้น

1.3.1 ระบบต้องอยู่ภายใต้เงื่อนไขของการกำหนดลักษณะการตั้งกล้องของการตั้งกล้องทั้ง 2 ตัว แบบขนาน

1.3.2 ระบบต้องอยู่ภายใต้เงื่อนไขของการกำหนดระยะห่างระหว่างกล้อง 2 ตัวคงที่

1.4 ขอบเขตของงานวิจัย

1.4.1 สร้างแผนที่บ่งบอกความลึกของสภวะแวดล้อมภายนอกอาคารได้

1.4.2 บูรณะสภาวะแวดล้อม 3 มิติภายนอกอาคารได้

1.4.3 กำหนดมาตรวัดความถูกต้องของแผนที่ได้อย่างชัดเจน

1.5 วิธีการดำเนินงาน

1.5.1 การหาแผนที่บ่งบอกความลึก (depth maps)

-ดำเนินการสืบค้นผลงานวิจัยที่เกี่ยวข้องกับระบบการมองเห็นแบบกล้อง 2 ตัวที่ปรากฏมา ก่อนแล้ว จากฐานข้อมูลต่าง ๆ เช่น IEEE และ Science Direct เป็นต้น

-ดำเนินการเก็บข้อมูลภาพแบบกล้อง 2 ตัวจากฐานข้อมูลต่าง ๆ เช่น อินเตอร์เน็ทเป็นต้น

-ดำเนินการออกแบบสร้างแบบจำลองการหาแผนที่บ่งบอกความลึกด้วยวิธีเครือข่ายฮอพฟิลด์

-ดำเนินการศึกษาการหาแผนที่บ่งบอกความลึกด้วยวิธีเครือข่ายฮอพฟิลด์

-ดำเนินการสร้างแบบจำลองการหาแผนที่บ่งบอกความลึกด้วยวิธีเครือข่ายฮอพฟิลด์

-ดำเนินการจำลองภาพต้นแบบขนาด 640x480 จุดภาพ (pixel) ในปริภูมิ RGB แล้วเปลี่ยนเป็นภาพระดับเทา

-ดำเนินการสร้างแบบจำลองการหาแผนที่บ่งบอกความลึก โดยเริ่มจากการคำนวณหาลักษณะเด่นของจุดภาพ โดยใช้วิธีหาขอบของแฮรีส (Harris detector)

-จะได้ลักษณะเด่นของจุดภาพจากกล้องทั้ง 2 ตัว หลังจากนั้นนำไปคำนวณในแบบจำลองเครือข่ายประสาทฮอพฟิลด์ 2 มิติ

-ดำเนินการทดสอบกับภาพวัตถุต่าง ๆ จากฐานข้อมูลและภาพสภาวะจริงกับแบบจำลองการหาแผนที่บ่งบอกความลึกที่จำลองขึ้น

-ดำเนินการสืบค้นผลงานวิจัยที่เกี่ยวข้องกับการไหลเชิงแสงที่ปรากฏมา ก่อนแล้วจากฐานข้อมูลต่าง ๆ เช่น IEEE และ Science Direct เป็นต้น

-ดำเนินการเก็บข้อมูลภาพเคลื่อนไหวโดยใช้กล้อง 2 สองตัวที่ติดตั้งกันแบบขนาน

-ดำเนินการศึกษานำระบบการจับกลุ่มการไหลเชิงแสงด้วยแผนผังคุณลักษณะจัดการตัวเองแบบดัดแปรที่มีผู้วิจัยไว้ก่อนหน้านี้

-ดำเนินการสร้างแบบจำลองการหาแผนที่บ่งบอกความลึกด้วยวิธีการไหลเชิงแสงร่วมกับเครือข่ายฮอพฟิลด์

-ดำเนินการทดสอบกับภาพวัตถุต่าง ๆ จากฐานข้อมูลและภาพสภาวะจริงกับแบบจำลองการหาแผนที่บ่งบอกความลึกที่จำลองขึ้น

1.5.2 การบูรณะสภาวะ 3 มิติ (3D reconstruction)

-ดำเนินการดำเนินการสืบค้นผลงานวิจัยที่เกี่ยวข้องกับการบูรณะสภาวะ 3 มิติที่ปรากฏมาก่อนแล้ว จากฐานข้อมูลต่าง ๆ เช่น IEEE และ Science Direct เป็นต้น

-ดำเนินการสร้างแบบจำลองการบูรณะสภาวะ 3 มิติด้วยวิธีการเชิงเรขาคณิตเพื่อหาพิกัดจริงของตำแหน่งวัตถุ

-ดำเนินการทดสอบกับภาพวัตถุต่าง ๆ จากฐานข้อมูลและภาพสภาวะจริงในอาคารกับแบบจำลองการการบูรณะสภาวะ 3 มิติที่จำลองขึ้น

1.6 ระเบียบวิธีวิจัย

- ใช้เครื่องคอมพิเตอร์แบบพกพา (Notebook) โปรแกรม MATLAB ในการสร้างแบบจำลอง
- ดำเนินงานโดยใช้ภาพเคลื่อนไหวจากกล้อง CCD เพื่อจำลองระบบ

1.7 เครื่องมือที่ใช้ในงานวิจัย

- โปรแกรม MATLAB

- คอมพิวเตอร์แบบพกพา

- กล้อง CCD

1.8 การเก็บรวบรวมข้อมูล

- เก็บข้อมูลภาพเคลื่อนไหวจากกล้อง โดยทำการเขียนโปรแกรม MATLAB เพื่อสร้างการบูรณะสภาวะ 3 มิติ

1.9 การวิเคราะห์ข้อมูล

- วิเคราะห์ข้อมูลด้วยเทคนิคเฉพาะทางวิศวกรรม

1.10 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

- สามารถเป็นแนวทางในการพัฒนาระบบการมองเห็นของคอมพิวเตอร์ ซึ่งเป็นเครื่องมือพื้นฐานที่จำเป็นในการวิจัยหลาย ๆ ด้าน เช่น การพัฒนาหุ่นยนต์ การพัฒนาระบบอุตสาหกรรมอัตโนมัติ ฯลฯ

- สามารถปรับปรุงวิธีการหาแผนที่บ่งบอกความลึกโดยใช้วิธีเครือข่ายฮอพฟิลด์ (Hopfield Neuron Network) ร่วมกับการไหลเชิงแสง (Optical flow)

- ได้ขั้นตอนวิธีสำหรับบูรณะสภาวะ 3 มิติด้วยวิธีแผนที่บ่งบอกความลึกมีประสิทธิภาพมากยิ่งขึ้นและมีความเร็วมากขึ้น

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 2

ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 กล่าวนำ

การหาแผนที่ 3 มิติ (3D Reconstruction) ต้องอาศัยกระบวนการที่ซับซ้อนโดยเริ่มจากการเก็บข้อมูลภาพจากกล้อง แล้วนำไปคัดกรองภาพเพื่อให้ได้ตำแหน่งจุดเด่นในภาพก่อนที่จะนำไปคำนวณในเครือข่ายประสาทเทียม ซึ่งขั้นตอนเหล่านี้ต้องการการประมวลผลที่รวดเร็วและแม่นยำ รวมทั้งต้องมีการวิเคราะห์เชิงปัญญาประดิษฐ์เพื่อที่จะคำนวณผลลัพท์ได้ถูกต้องยิ่งขึ้น จากการค้นคว้าจะพบได้ว่าในงานวิจัยนี้มีความรู้ด้านต่างๆ ทั้งการออกแบบระบบการมองเห็นของกล้อง การประมวลผลภาพ และการคำนวณเชิงเครือข่ายประสาท

2.2 การออกแบบระบบการมองเห็นของกล้องสองตัว

ระบบการใช้กล้องที่ใช้เทคนิคของการมองเห็นของเครื่องจักร (machine vision) ในการประมวลผลภาพแบบอัตโนมัตินั้นมักมีความยืดหยุ่นในการใช้งานสูง สามารถใช้ได้ในการตรวจจับช่องทางเดินรถ เส้นแบ่งทาง สิ่งกีดขวางบนถนน ยานพาหนะอื่น ๆ หรือลักษณะเด่นที่กำหนดจุดหมายปลายทางของการเคลื่อนที่ ภาพที่ใช้อาจเป็นภาพสีหรือภาพขาวดำก็ได้ขึ้นอยู่กับวิธีการ และข้อจำกัดทางด้านการประมวลผลข้อมูล

2.2.1 ระบบการมองเห็นของกล้อง 2 ตัว (stereo vision)

รูปที่ 2.1 การติดตั้งกล้องสองตัวแบบขนาน

ในปัจจุบันการวิจัยเกี่ยวกับการหาระยะห่างระหว่างกล้องและวัตถุนั้นนิยมใช้ระบบการมองเห็นของกล้องสองตัว (stereo vision) ซึ่งมีคุณสมบัติในการมองเห็นใกล้เคียงกับกับดวงตาของมนุษย์มาก เพราะใช้กล้องสองตัวจับภาพวัตถุ โดยทำการติดตั้งแบบขนานกันดังแสดงในรูปที่ 2.1

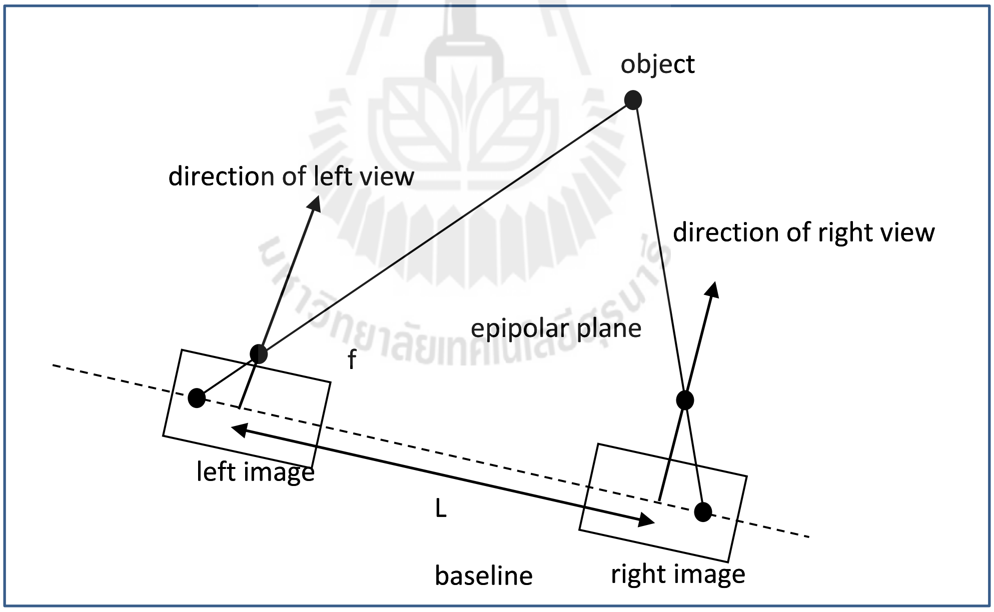
(ก)

(ข)

(ค)

รูปที่ 2.2 (ก) ภาพจากกล้องซ้าย (ข) ภาพจากกล้องขวา
(ค) ภาพจากกล้องซ้ายและกล้องขวามาซ้อนกัน

จากรูปที่ 2.2 (ค) จะเห็นได้ว่าภาพที่ได้จากกล้องสองตัวที่นำมาซ้อนกันวัตถุที่เกิดในแต่ละภาพมีตำแหน่งที่เกิดบนภาพต่างกันในแนวนอน จะเห็นได้ว่าแบบจำลองพื้นฐานของกล้องทั้งสองนี้จะแยกกันเพียงทิศทางแกน x เท่านั้น โดยจะมีระยะห่างเท่ากับระยะสัญฐาน (base line : b) และระนาบของภาพ (image plane) ของกล้องทั้งสองอยู่ในระนาบเดียวกัน ดังนั้นทำให้จุดที่เราสนใจของภาพจากภายนอกปรากฏลงบนระนาบของกล้องทั้งสองที่ตำปหน่งต่างกัน ซึ่งระยะระหว่างตำแหน่งบนระนาบนั้นเรียกว่าความต่างกัน (disparity) และระนาบที่ผ่านจุดศูนย์กลางของกล้องทั้งสองจุดที่เราสนใจนั้นเรียกว่าระนาบเอพิโพลาร์ (epipolar plane) ส่วนเส้นที่ตัดระนาบอิพิโพลาร์กับระนาบของภาพนั้นเรียกว่า อีพิโพลาร์ ไลน์ (Epipolar line) ซึ่งจากแบบจำลองของรูปที่ 2.3 แสดงนั้น ลักษณะจุดของภาพจากระนาบหนึ่งจะอยู่ในแถวเดียวกันกับอีกระนาบหนึ่งหมายความว่าเส้นระนาบเอพิโพลาจะอยู่ที่ตำแหน่งเดียวกันของระนาบภาพของกล้องทั้งสองแต่ในการกำหนดให้เป็นแบบนี้ได้นั้นแสดงว่าต้องกำหนดให้ความแตกต่างในแนวตั้ง (vertical disparity) นั้นมีค่าเท่ากับศูนย์



รูปที่ 2.3 องค์ประกอบของระบบการมองแบบกล้อง 2 ตัว

จากรูปที่ 2.4 จุด P จะปรากฏเป็นจุด p1 และ p2 บนระนาบภาพด้านซ้ายและระนาบภาพด้านขวาตามลำดับ และกำหนดให้จุดกำเนิดของพิกัดโลกอยู่ตำแหน่งเดียวกันกับตำแหน่งศูนย์กลางเลนส์ของกล้องด้านซ้าย ดังนั้นเราเปรียบเทียบสามเหลี่ยมคล้าย PMC_l แล p_lLC_r จะได้

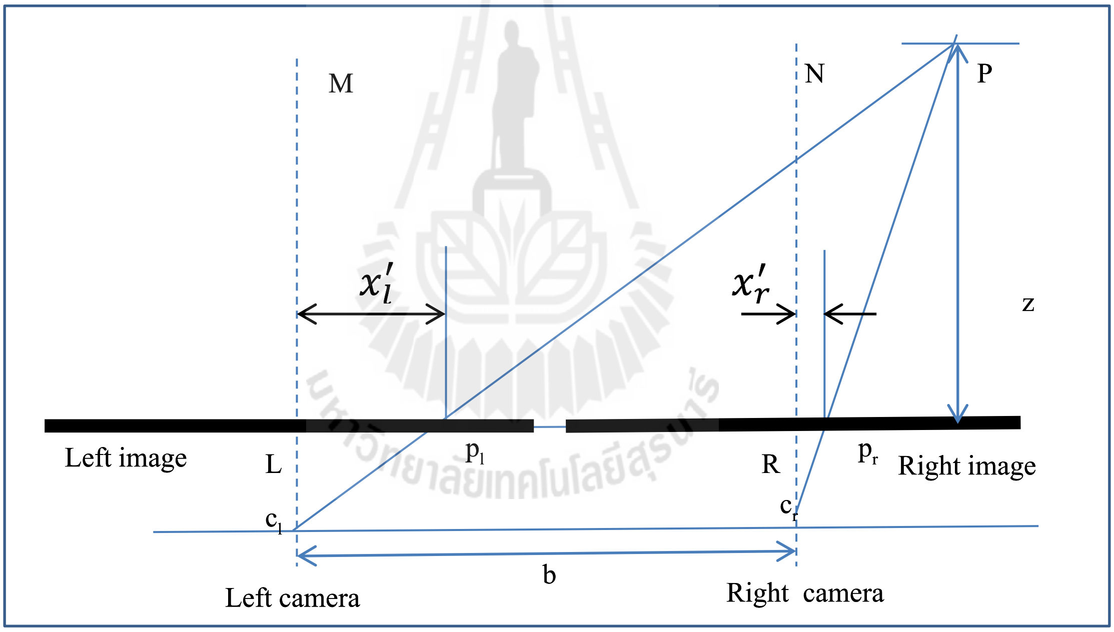
$$\frac{x}{z} = \frac{x_l'}{f} \tag{2.1}$$

ในทำนองเดียวกัน จากสามเหลี่ยมคล้าย PNC_l และ p_rRC_r จะได้

$$\frac{x-b}{z} = \frac{x_r'}{f} \tag{2.2}$$

รวมสมการ 2.1 และ 2.2 เข้าด้วยกันจะได้

$$z = \frac{bf}{(x_l' - x_r')} \tag{2.3}$$



รูปที่ 2.4 ลักษณะทางเรขาคณิตของกล้องสองตัว

ดังนั้นระยะ Z หรือระยะห่างระหว่างวัตถุกับจุดศูนย์กลางของภาพหาได้จากค่าความแตกต่างของภาพนั้นเอง หรือ การหาแผนที่ความไม่เท่ากัน (disparity maps)

2.2.2 การทำงานของกล้องดิจิตอล

การถ่ายภาพคือการวาดภาพด้วยแสง โดยอาศัยกล้องเป็นตัวกลางถ่ายทอดแสงลงบนแผ่นฟิล์มหรือตัวรับภาพในกล้องดิจิตอล หากไม่มีแสงก็ไม่มีภาพ แต่ในปัจจุบันกล้องที่นิยมใช้

เป็นกล้องดิจิตอล ซึ่งหลักการการทำงานเหมือนกลับกล้องฟิล์มเพียงแต่กล้องดิจิตอลใช้เซนเซอร์ในการรับภาพ แบ่งออกเป็น 2 ประเภท คือ ซีซีดี (CCD) และ ซีมอส (CMOS) กล้องดิจิตอลนั้นมีข้อดีกว่า คือเมื่อถ่ายภาพเสร็จแล้วสามารถดูภาพผ่านจอ LCD ได้ทันที นั่นเกิดจากการประมวลผลที่รวดเร็วทำให้สามารถนาภาพที่ได้จากเซ็นเซอร์ พร้อมกับการปรับแต่งสี ความสว่าง และความคมชัดที่เราปรับตั้งก่อนถ่าย

รูปที่ 2.5 กล้องดิจิตอลพร้อมเลนส์

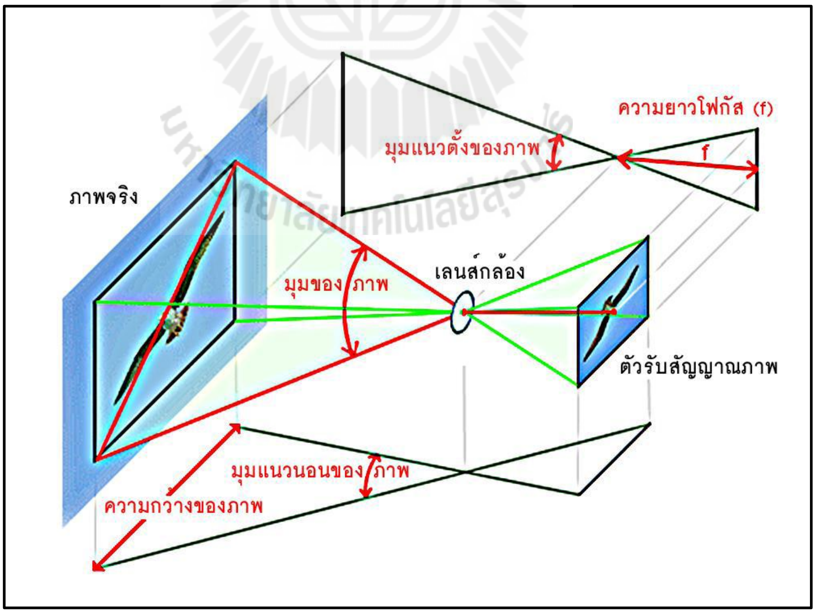
2.2.3 ความยาวโฟกัสของเลนส์ (Lens Focal Length)

เลนส์เป็นส่วนประกอบที่สาคัญของกล้องถ่ายภาพ ในกระบอกเลนส์จะประกอบไปด้วยชิ้นเลนส์หลายชิ้น จัดเรียงกันเป็นกลุ่มเพื่อทาหน้าที่ในการโฟกัสแสงให้ไปตกยังตัวรับภาพ (ฟิล์มหรือตัวรับภาพในกล้องดิจิตอล) เป็นตัวกลางในการถ่ายทอดแสงที่สะท้อนจากวัตถุไปยังตาแหน่งของตัวรับภาพ ดังนั้นคุณภาพของเลนส์จึงนับว่ามีส่วนสาคัญที่สุดส่วนหนึ่งในการบันทึกภาพ การถ่ายทอดความคมชัดทั้งในแง่ของรายละเอียดและสีสัน เลนส์ที่ใช้ในการบันทึกภาพจะมีทางยาวโฟกัสที่หลากหลาย เพื่อถ่ายทอดมุมมอง หรือองศาการรับภาพที่แตกต่างกัน ตามวัตถุประสงค์ของผู้ใช้งาน เลนส์ในปัจจุบันที่นิยมใช้กันส่วนใหญ่จะเป็นเลนส์ซูม (มีทางยาวโฟกัสปรับเปลี่ยนได้) ทาให้สามารถบันทึกภาพได้ในหลายมุมมองโดยไม่ต้องเปลี่ยนตาแหน่งในการบันทึกภาพ

ซึ่งเลนส์ด้านหน้าจะเป็นเลนส์นูนเมื่อแสงผ่านเข้ามา แสงก็จะเอียงเข้าหากันมารวมกันเป็นรูปกรวยจากนั้นก็จะกระจายออกมา ซึ่งจุดรวมแสงนี้เรียกว่าจุดโฟกัส (Focal Point) ซึ่งการกระจายแสงหลังจากผ่านจุดโฟกัส จะกระจายแสงออกจนได้แสงคุณภาพครอบคลุมพื้นที่ตัวรับ

สัญญาณภาพ ภาพที่ได้จะเป็นภาพจริงหัวกลับ ซึ่งระยะห่างจากจุดโฟกัสไปยังตัวรับสัญญาณภาพจะเรียกว่าความยาวโฟกัส (Focal Length)

รูปที่ 2.6 เลนส์ชนิดต่างๆ



รูป 2.7 จุดโฟกัสในกล้องดิจิตอลตามแกน x, y

2.3 การประมวลผลภาพเบื้องต้น

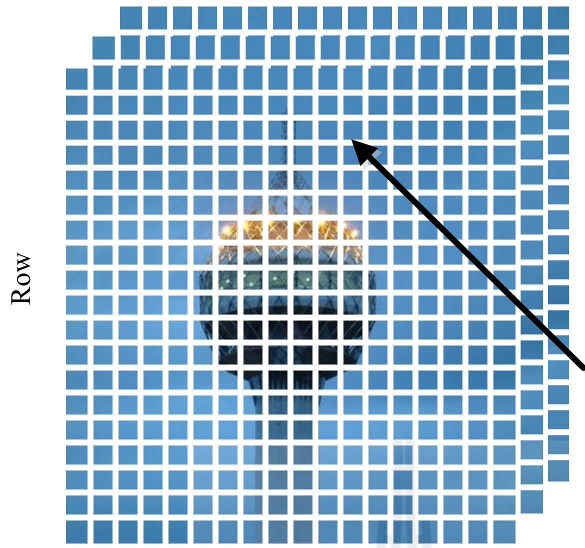
การแปลงภาพเชิงต่อเนื่อง (Continuous Image) แบบ 3 มิติ (ความกว้าง, ความสูงและความลึก) ให้เป็นภาพเชิงต่อเนื่อง 2 มิติ โดยใช้อุปกรณ์เชิงแสง (Optical Device) เช่นกล้องถ่ายภาพ การถ่ายภาพด้วยกล้องจะได้ภาพแบบ 2 มิติ (ความกว้างและความสูง) เพราะเราไม่อาจถ่ายความลึกของสถานที่มาด้วยได้ ความลึกที่กล่าวถึงคือค่าขอบของภาพ (Edge) หลักการประมวลผลภาพ ภาพถูกพิจารณาเป็นฟังก์ชัน 2 ตัวแปล เช่น A(x,y) เมื่อ A คือ Amplitude ของค่าความสว่างของภาพและค่า x และ y คือตำแหน่งของจุดในระบบพิกัดฉาก

ภาพเชิงดิจิตอล (Digital Image Definitions) แบบ 2 มิติแบบไม่ต่อเนื่องพิจารณาเป็นฟังก์ชัน 2 ตัวแปร a[m,n] จะถูกบรรยายในลักษณะเป็นช่องที่ไม่ต่อเนื่องซึ่งได้รับสืบทอดมาจากภาพแบบอนาล็อกเชิงต่อเนื่อง 2 มิติ ของฟังก์ชั่น A(x,y) โดยใช้กระบวนการ Sampling ทำให้ได้ข้อมูลแบบดิจิตอลออกมา การทำ Sampling คือการวัดสัญญาณที่เข้ามาโดยทำเป็นช่วงเวลาที่เท่าๆ กัน ลักษณะของภาพแบบดิจิตอล จะเกิดจากการแบ่งความกว้างและความยาวของภาพออกเป็นส่วนๆ เป็นจำนวนแถว (N) และจำนวนหลัก (M) จะได้ส่วนที่ตรงกันระหว่างแถวและหลักเป็นภาพชิ้นเล็กๆ เรียกว่าพิกเซล (Pixel) ค่าพิกัด [m,n] นิยามการอ้างอิงได้โดย (m = 0,1,2,…,m-1) และ (n = 0,1,2,…,m-1) โดยที่แต่ละพิกเซลมีค่าเป็นฟังก์ชั่นของความลึก (z) และฟังก์ชั่นของสี (λ)

ขั้นตอนการแปลงภาพให้เป็นภาพในเชิงดิจิตอลเราเรียกว่า มีขั้นตอนคือ การบันทึกภาพ (Image Acquisition) การสุ่มเลือกจุดตำแหน่ง (Image Sampling) และการประมาณค่าความเข้มของแสง (Image Quantization) แต่ละจุดภาพถูกแทนด้วยสีของภาพในโทนสีขาวดำ (Grayscale) และแทนสีดำด้วยค่าตัวเลข 0 แทนสีขาวด้วยตัวเลข 255 หรือเท่ากับ 2^8 เมื่อ 8 คือจำนวนบิตของหน่วยความจำที่ใช้ในการเก็บค่า สีดำถูกแทนด้วยรหัสในเลขฐานสองคือ 00000000_B และสีขาวถูกแทนด้วยรหัส 11111111_B และสีที่อยู่ระหว่างสีดำกับสีขาวก็จะไล่ไปตามลำดับการนับของบิตในเลขฐานสองดังแสดงในรูปที่ 2.8

เมื่อพิจารณาจากภาพแบบแบบไม่ต่อเนื่อง ค่าของฟังก์ชัน a[m,n] มีค่ามาตรฐานสำหรับใช้ในกระบวนการ Digital Image Processing มีที่มาจาก Video Standard และค่า Algorithmic Requirements เรียกว่า Common Values ดังตารางที่ 2.1

กรณีของ M = N = 2 K เมื่อ *K* = 8,9,… และค่าของระดับของโทนสีขาวดำ (Gray Level) มีค่าเท่ากับ 2B เมื่อ B เป็นจำนวนบิตในการแทนระดับความสว่างและถ้า B มากกว่า 1 เรียกว่า ภาพขาวดำ (Gray Image) ถ้าค่าของ B มีค่าเท่ากับ 1 เรียกว่าภาพไบนารี่ (Binary Image) จะมีค่าของระดับสีหรือค่าของแสงสว่าง 2 ระดับค่าระดับของโทนสีขาวดำที่จะใช้อ้างอิงคือสีดำ "0" และสีขาว "1" ชนิดของภาพดิจิตอลแบ่งออกเป็นหลายประเภท



รูปที่ 2.8 ภาพในเชิงดิจิตอลเมื่อแบ่งค่าพิกเซลขนาด [m = 11, n = 4]

ตารางที่ 2.1 ค่ามาตรฐานของ Common Values ของภาพดิจิตอล 2 มิติ

ตัวแปร	สัญลักษณ์	ค่าตัวอย่างพิกเซล
Rows	N	180,360,720,1024
Columns	M	120,240,480,960
Gray	L	0,128,255

การประมวลผลภาพ มีจุดประสงค์เพื่อพัฒนาและปรับปรุงข้อมูลที่เป็นภาพเพื่อให้สามารถแปลความหมายได้ดีขึ้น และเพื่อเตรียมข้อมูลฉาก (scene data) สำหรับให้คอมพิวเตอร์เข้าใจได้ ขบวนการประมวลผลภาพแบ่งออกได้เป็น 3 ขบวนการ คือ ขบวนการขั้นต้น เป็นการได้มาซึ่งภาพและการปรับปรุงภาพให้ดีขึ้น ได้แก่ การแปลงสัญญาณภาพ การเพิ่มความคมชัดของภาพ การกู้คืนสัญญาณภาพ เป็นต้น ขบวนการขั้นกลาง คือ ส่วนของการแยกองค์ประกอบและการดึงเอาคุณลักษณะที่สำคัญออกมา และขบวนการขั้นสูง คือ การตีความหมายและการจดจำภาพ

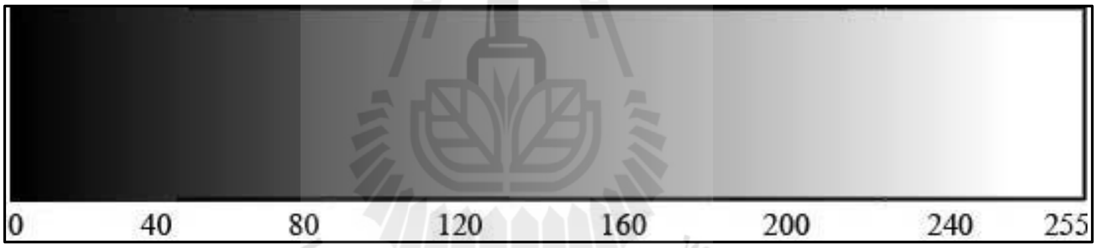
2.3.1 แบบจำลองระดับเทา (gray scale model)

ภาพระดับเทาสามารถเขียนแทนด้วย $f(x, y)$ โดยที่ขนาดของ f ที่ตำแหน่ง (x, y) คือค่าความเข้มของภาพที่จุดนั้น ๆ โดยที่ค่า $f(x, y)$ จะต้องไม่เป็นศูนย์และมีค่าจำกัด ซึ่งสามารถ

แยกองค์ประกอบได้เป็น 2 ส่วน คือ องค์ประกอบความสว่าง (illumination component : $i(x, y)$) และ องค์ประกอบการสะท้อนของแสง (reflectance component : $r(x, y)$) โดยที่ $f(x, y)$ เกิดจากการคูณกันขององค์ประกอบทั้งสองดังสมการที่ (2.4)

$$f(x, y) = i(x, y).r(x, y) \tag{2.4}$$

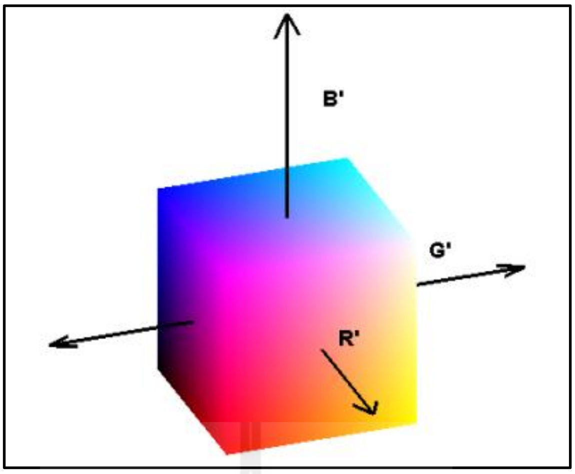
โดยที่ $i(x, y)$ จะต้องมีค่ามากกว่าศูนย์และมีค่าจำกัด ส่วน $r(x, y)$ จะต้องมีค่ามากกว่าศูนย์และน้อยกว่าหนึ่ง ภาพ f เป็นภาพที่มีเฉดสีเดียว ดังนั้นความเข้มของ f ที่ตำแหน่ง(x, y) เรียกว่าระดับเทา (gray level : l) ของภาพในตำแหน่งนั้น โดยที่ระดับเทาจะอยู่ในช่วง $L_{min} \leq l \leq L_{max}$ เรียกช่วง $[L_{min}, L_{max}]$ ว่าสเกลระดับเทา (gray scale) โดยปกติจะเลื่อนช่วงนี้ให้เป็น $[0, L]$ โดยที่ l เป็นศูนย์ คือจุดภาพที่มีสีดำ และที่ l เท่ากับ L คือจุดภาพที่มีสีขาว ส่วนจุดภาพที่มีค่า l ระหว่างนี้จะแสดงระดับสีของสีเทาแสดงดังรูปที่ 2.9



รูปที่ 2.9 ค่าระดับเทา

2.3.2 แบบจำลองสี RGB

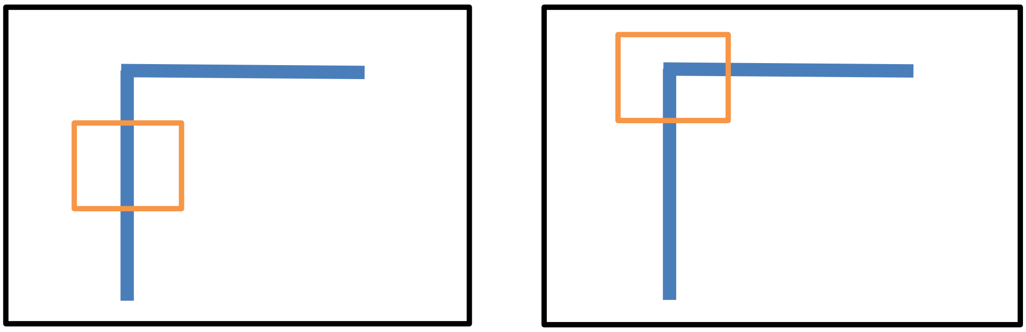
แบบจำลองสี RGB เป็นแบบจำลองที่ประกอบไปด้วยองค์ประกอบสเปกตรัมของ 3 แม่สีปฐมภูมิ (primary spectral components) ได้แก่ สีแดง (red : R) สีเขียว (green : G) และสีน้ำเงิน (blue : B) ซึ่งแต่ละสีอยู่บนระนาบภาพ (image planes) ที่เป็นอิสระต่อกันบนพื้นฐานของระบบพิกัดคาร์ทีเซียนดังแสดงในรูปที่ 2.10 กล้องสีส่วนมากที่ใช้ในการรับภาพดิจิทัล นิยมใช้แบบจำลองสี RGB เป็นรูปแบบในการรับภาพ ดังนั้นแบบจำลองสี RGB จึงเป็นแบบจำลองที่สำคัญในการประมวลผลภาพ



รูปที่ 2.10 ลูกบาศก์ของแบบจำลองสี RGB

2.3.3 การหามุมด้วยวิธีของฮาร์ริส (The Harris Corner Detector)

การหามุมด้วยวิธีของฮาร์ริสเป็นเทคนิคการหาจุดที่ได้รับความนิยม เนื่องจากมีจุดเด่นในเรื่องของ ความแม่นยำของการหมุน ขนาด ความแปรปรวนของแสงที่ทำให้ภาพเกิดความผิดเพี้ยนและภาพมีจุดรบกวน หลักการของฮาร์ริสมาจากการค้นหาจุดของภาพโดยมองผ่านหน้าต่างเล็กๆที่กำหนดขึ้น โดยภายในจะเป็นเส้นของรูปจากนั้นให้เลื่อนหน้าต่างเล็กๆ ที่กำหนดขึ้นไปในทิศทางต่างๆ ตามเส้นภายในภาพ เมื่อพบการเปลี่ยนแปลงมากๆ กล่าวคือเกิดการหักเหของเส้นมากๆ จะถือว่าจุดนั้นเป็นจุดมุม ดังแสดงในรูป 2.11



รูปที่ 2.11 การหามุมด้วยวิธีของฮาร์ริส

2.4 การไหลเชิงแสง

การไหลของแสงเป็นวิธีการหนึ่งที่มีประสิทธิภาพในการตรวจจับวัตถุเคลื่อนที่ภายในภาพได้ ซึ่งการสร้างการไหลของแสงนี้จะใช้การวิเคราะห์จากการเปลี่ยนแปลงของความเข้มแสงภายในภาพแต่ละลำดับภาพ (image sequence) โดยที่การไหลของแสงจะประกอบไปด้วยขนาดและทิศทางของความเข้มแสงที่เปลี่ยนตำแหน่งไปดังรูปที่ 2.8 แสดงตัวอย่างของการไหลของแสง การตรวจจับวัตถุโดยใช้การไหลของแสงจะใช้การวิเคราะห์กลุ่มของความเข้มแสงของวัตถุภายในภาพ ซึ่งความเข้มแสงของวัตถุที่เปลี่ยนตำแหน่ง จะคงความเข้มแสงเท่าเดิม และให้ผลของการไหลของแสงของกลุ่มแสงของวัตถุที่เปลี่ยนแปลงตำแหน่งไป เปลี่ยนทิศทางและขนาดที่ใกล้เคียงกัน

2.4.1 Gradient Base Method

กรรมวิธีนี้ใช้วิธีการสร้างการไหลของแสงโดยแก้ปัญหาอนุพันธ์เป็นหลัก โดยอาศัยหลักการของความคงที่ของความเข้มแสงจากวัตถุภายในภาพที่เปลี่ยนแปลงไป กำหนดให้ I คือความเข้มแสง สามารถเขียนเป็นสมการได้ดังนี้

$$I(x+dx, y+dy, t+dt) = I(x,y,t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + \cdots \qquad (2.5)$$

จะได้ว่า

$$\nabla I(x,y,t) + I_t = 0 \qquad (2.6)$$

สมการที่ 2.6 ได้มาจากการแก้สมการ อนุกรมเทย์เลอร์ (Horn B.K.P and Schunck B.G., 1981) จากสมการที่ 2.5 โดย $v = [v_x, v_y] = [\frac{dx}{dt}, \frac{dy}{dt}]$ และ $\nabla I = [\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}]$ เรียกสมการที่ 2.6 ว่า optical flow constraint equation (Barron J.L. and Thacker N.A., 1990) ในการพิจารณาระบบสมการเส้นตรงที่มี 2 ตัวแปร 1 สมการ ซึ่งมีแนวทางการแก้ดังนี้

2.4.1.1 วิธีของฮอร์นสชังก์ Horn and Schunck

Horn and Schunck (1993) ได้รวมเงื่อนไขของความลาดชัน (gradient) จากสมการ ที่ 2.5 ซึ่งเป็นเทอมของ ความราบเรียบรวม (global smoothness) เข้ากับเงื่อนไขของการประมาณความเร็ว $v(x, y,t) = (u,v)$ เพื่อพิจารณาหาค่าน้อยที่สุดของ

$$\int_D (\nabla I.v + I_t)^2 + \lambda^2(\|\nabla u\|_2^2 + \|\nabla v\|_2^2)dx \qquad (2.7)$$

โดยที่ขนาด (magnitude) ของ λ เป็นการแสดงให้เห็นถึงผลที่มีต่อเทอมของความราบเรียบ (smoothness) ให้λ = 0.5 (Horn B.K.P and Schunck B.G: 1993) จะทำให้ผลลัพธ์ของกรณีที่ทำการทดสอบได้ดี พิจารณาใช้สมการที่ 2.7 ในการวนซ้ำเพื่อทำการพิจารณาหาค่าน้อยที่สุดและได้ความเร็วของแสงของแต่ละจุดภาพ โดยการวนซ้ำจะเป็นไปตามสมการที่ 2.8 และ 2.9

$$u^{k+1} = u^k - \frac{I_x[I_x u^{-k} + I_y v^{-k} + I_t]}{\propto^2 + I_x^2 + I_y^2} \tag{2.8}$$

$$v^{k+1} = v^k - \frac{I_y[I_x u^{-k} + I_y v^{-k} + I_t]}{\propto^2 + I_x^2 + I_y^2} \tag{2.9}$$

โดยที่

k คือ จำนวนครั้งในการวนซ้ำ

u^0, v^0 คือ ค่าเริ่มต้นในการประมาณความเร็ว

$\bar{u}^k, \bar{v}^k$ คือ ค่าของตำแหน่งใกล้เคียงของ u^k, v^k

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

2.4.1.2 วิธีของลูคาสคาเนด Lucas & Kanade

อัลกอริทึมลูคาสคาเนด (Lucas-Kanade (LK)) ซึ่งถูกเสนอขึ้นในปี 1981 เป็นความพยายามหนึ่งที่จะใช้ในการติดตามการเคลื่อนไหวทั้งภาพ แต่ในปัจจุบันนี้ ด้วยเหตุที่วิธีนี้ง่ายที่จะประยุกต์ใช้กับเซตย่อย (Subset) ของจุดในรูปถูกอินพุทเข้ามา มันได้กลายมาเป็นเทคนิคในเรื่องการติดตามการเคลื่อนไหวเฉพาะที่ ที่สาคัญ อัลกอริทึม LK สามารถนาไปประยุกต์ใช้ในสภาพแวดล้อมที่เป็นการติดตามเฉพาะที่เพราะว่ามันจะน่าเชื่อถือเฉพาะกับข้อมูล ณ ตำแหน่งใดตำแหน่งหนึ่ง ซึ่งได้มาจากมาจากการสร้างกรอบเล็กล้อมรอบแต่ละจุดที่สนใจ ซึ่งตรงข้ามกันกับอัลกอริทึมของฮอร์นสชังก์ ข้อเสียของการใช้กรอบ ณ ตาแหน่งใดใด ที่เล็กในลูคาสคาเนดคือเมื่อมีการเคลื่อนที่ไปมากๆ จะทาให้จุดนั้นเลื่อนออกไปจากกรอบ ณ ตำแหน่งนั้นๆ และจะกลายเป็นอัลกอริทึมที่หาผลลัพธ์ไม่ได้ ปัญหานี้นาไปสู่การพัฒนาไปเป็นอัลกอริทึม LK แบบ "พีระมิด" ซึ่งการติดตามนั้นจะเริ่มต้นจากชั้นที่สูงที่สุดของพีระมิดของภาพ (กรอบเล็กที่รายละเอียดของภาพต่า) และทางานลงมาเรื่อยๆ จนถึงชั้นที่อยู่ต่าสุด (กรอบใหญ่ที่รายละเอียดของภาพสูง) การติดตามการ

เคลื่อนที่ไปตามพีระมิดของภาพจะทาให้กรอบ ณ ตาแหน่งนั้นๆ สามารถจับการเคลื่อนที่ที่มีการเคลื่อนที่ไปมากๆ ได้ แนวคิดพื้นฐานของอัลกอริทึม LK มาจากสมมติฐานสามอย่างคือ

1) *ความคงที่ของแสงสว่าง* พิกเซลจากรูปภาพของวัตถุในฉากไม่เปลี่ยนแปลงและไม่เคลื่อนที่จากเฟรมต่อเฟรม สาหรับภาพเกรย์สเกล (LK สามารถทาได้ในภาพสี) นั่นหมายความว่าความสว่างของพิกเซลไม่เปลี่ยนแปลงขณะที่มันถูกติดตามจากเฟรมต่อเฟรม

2) *ความนิ่งชั่วขณะ* หรือ “การเคลื่อนไหวที่เล็ก” การเคลื่อนไหวของพื้นผิววัตถุของแต่ละส่วนของภาพเปลี่ยนแปลงไปอย่างช้าๆ เมื่อเวลาผ่านไป ในเชิงปฏิบัติจะหมายถึงว่าการอินครีเมนต์ในขณะนั้นเร็วเพียงพอที่จะสัมพันธ์กับสเกลของการเคลื่อนไหวในรูปภาพซึ่งวัตถุไม่ได้เคลื่อนไหวอะไรมากมายจากเฟรมต่อเฟรม

3) *ความเชื่อมติดกันอย่างไม่มีช่องขาด* จุดที่อยู่ใกล้เคียงกันในฉากเป็นของพื้นผิววัตถุเดียวกัน มีการเคลื่อนไหวที่เหมือนกัน และฉายไปยังจุดใกล้เคียงกันบนแนวระนาบของภาพ

จากวิธีการของ Lucas & Kanade (1981) ได้นำเสนอวิธีการในการหาค่าน้ำหนักเหมาะสมของค่ากำลังสองน้อยสุด (LS) ของเงื่อนไขเฉพาะถิ่นอันดับหนึ่ง (local first-order) ของสมการที่ 2.5 เพื่อพิจารณาหาค่าเหมาะสมซึ่งเป็นค่าคงที่ของแบบจำลองสำหรับ v ในแต่ละจุดภาพโดยพิจารณาหาค่าน้อยที่สุดของสมการที่ 2.10

$$\sum_{x\in\Omega} W^2(x)[\nabla I.v + I_t]^2 \tag{2.10}$$

โดยที่ $W(x)$ คือ ฟังก์ชันหน้าต่าง (windows function) ซึ่งวิธีการแก้ปัญหาของสมการที่ 2.10 โดย n คือจุดที่เวลา t มีดังนี้

$$A^T W^2 A v = A^T W^2 b \tag{2.11}$$

$$A = [\nabla I(x_1), \ldots, \nabla I(x_n)]^T \tag{2.12}$$

$$W = diag[W(x_1), \ldots, W(x_n)] \tag{2.13}$$

$$b = -[I_t(x1), \ldots, I_t(x_n)]^T \tag{2.14}$$

$$v = [A^T W^2 A]^{-1} A^T W^2 b \tag{2.15}$$

ดังนั้นจะได้ผลลัพธ์ของสมการที่ 2.12 คือสมการที่ 2.15 ซึ่งรูปแบบของผลเฉลยจะเข้าใกล้รูปแบบของสมการ $A_T W_2 A$ ซึ่งไม่เป็นเอกพจน์ (nonsingular) ที่มีเมทริกซ์ขนาด 2 x 2

$$A^T W^2 A = \begin{bmatrix} \sum W^2(x) I_x^2(x) & \sum W^2(x) I_x(x) I_y(x) \\ \sum W^2(x) I_y(x) I_x(x) & \sum W^2(x) I_y^2(x) \end{bmatrix} \tag{2.16}$$

2.4.2 Region - Based Matching

วิธีการพิจารณาขอบเขตโดยความเหมือนของบริเวณหนึ่ง ๆ 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 $d = [d_x, d_y]$ เพื่อพิจารณาหาค่าเหมาะสมที่ดีที่สุดของขอบเขตของภาพที่เวลาต่างกันจากการค้นหาความเหมือนที่ดีที่สุดโดยทำการพิจารณาที่ค่ามากที่สุดของความเหมือน (บน d) ซึ่งในแต่ละวิธี จะปรับปรุงค่าความเร็วให้แม่นยำที่แตกต่างกัน แต่จะใช้หลักการที่มุ่งไปในทางเดียวกัน เช่น วิธีการบรรทัดฐานสหสัมพันธ์ไขว้ (normalized cross - correlation) หรือพิจารณาค่าน้อยที่สุดของระยะทาง (distance) อาทิเช่น วิธีการผลรวมรากที่สองของผลต่าง (Sum – Squared Difference : SSD) ซึ่งมีรูปแบบสมการที่ (2.17)

$$SSD_{1,2}(x,d) = \sum_{j=-n}^{n} \sum_{i=-n}^{n} W(i,j) \left[I_1\big(x + (i,j)\big) - I_2\big(x + d + (i,j)\big) \right]^2 \tag{2.17}$$

$$SSD_{1,2}(x,d) = W(x) * [I_1(x) - I_2(x+d)]^2 \tag{2.18}$$

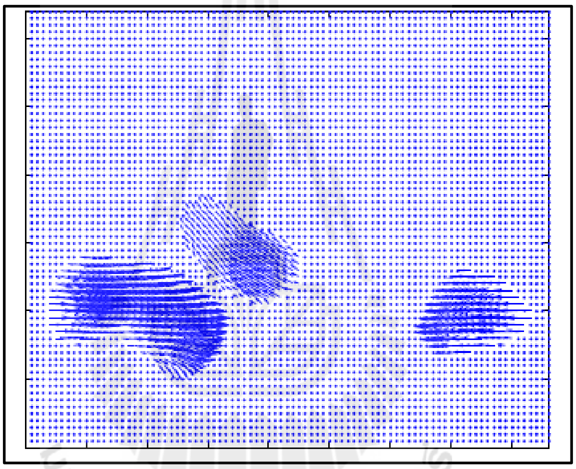
โดยที่ W หมายถึง ฟังก์ชันไม่ต่อเนื่องของหน้าต่าง 2 มิติ

d หมายถึง ค่าตัวเลขจำนวนเต็มของการเลื่อนตำแหน่ง

การพิจารณาค่าระยะทาง SSD การพิจารณาค่าความเหมือนแบบสหสัมพันธ์ไขว้ (cross - correlation) จะมีลักษณะวิธีการที่ใกล้เคียงกัน ในการพิจารณาค่าน้อยที่สุดของระยะทางการไหลของแสงในแต่ละวิธีข้างต้น ผลลัพธ์ที่ได้จะอยู่ในรูปของความเร็วในแต่ละจุดภาพ ดังแสดงในรูปที่ 2.12 ค) ผลการสร้างการไหลของแสงที่ได้จาก Lucas & Kanade โดยใช้ลูกศรแทนขนาดและทิศทางของความเร็วการไหลของแสง

(ก)

(ข)



(ค)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 2.12 การไหลของแสง ก) ภาพที่เวลา *t* ข) ภาพที่เวลา *t* +1
ค) ผลที่ได้จากวิธี Lucas & Kanade

2.5 เครือข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Neural network) เป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่เน้นด้านการคำนวณ และได้รับแรงบันดาลใจจากโมเดลทางชีวภาพ โดยมีความพยายามที่จะลอกเลียนความสามารถของสมองมนุษย์ ยกตัวอย่างเช่นพัฒนาการด้านสมองของเด็กตั้งแต่แรกเกิดจนถึงสามปีที่ในระยะเริ่มต้นของวัยนี้ เด็กจะเรียนรู้ว่าผู้ใดคือพ่อแม่ หรือคนแปลกหน้าในระยะต่อมาจะเริ่มตอบสนองเมื่อถูกเรียกชื่อ รู้จักแยกแยะสี และสิ่งของต่าง ๆ เริ่มเปล่งเสียงพูดเป็นคำ และเริ่มแสดงความคิดเห็น เป็นต้น จากลักษณะดังกล่าว จะเห็นว่า

ความสามารถในการเรียนรู้ของสมองมนุษย์มีการพัฒนาขึ้นตามลำดับ และทำงานได้อย่างมีประสิทธิภาพ แต่เมื่อเปรียบเทียบกับการทำงานของเครื่องคอมพิวเตอร์จะพบว่ามีการทำงานตามคำสั่งอย่างตรงไปตรงมาเท่านั้น ยกตัวอย่างเช่น เปิดแฟ้มข้อมูล การใช้อินเตอร์เน็ต และสืบค้นข้อมูล เป็นต้น การทำงานดังกล่าวไม่มีประสิทธิภาพพอที่จะให้เครื่องแสดงความคิดเห็น หรือวิพากย์วิจารณ์เกี่ยวกับข้อมูลต่าง ๆ ได้ ดังนั้น จะเห็นได้ว่าการทำงานของสมองมนุษย์มีความซับซ้อนเป็นอย่างมาก ซึ่งนักประสาทวิทยาตั้งแต่อดีตถึงปัจจุบันได้พยายามศึกษาค้นคว้ามาอย่างต่อเนื่องที่จะสร้างโปรแกรมคอมพิวเตอร์ที่มีลักษณะการประมวลผล คล้ายการทำงานของสมองมนุษย์และผลการศึกษาค้นคว้าที่ผ่านมา ส่วนหนึ่งก็สามารถนำไปประยุกต์ใช้ในสาขาวิชาต่าง ๆ ได้อย่างเป็นประโยชน์อย่าง

มากมายรูปแบบของโครงข่ายประสาทเทียมเป็นความก้าวหน้าอย่างหนึ่งของความพยายามที่จะพัฒนาโปรแกรมคอมพิวเตอร์ในปัจจุบันให้สามารถทำนายสิ่งต่าง ๆ ได้อย่างกว้างขวาง เนื่องจากโครงข่ายประสาทเทียมสามารถหาความสัมพันธ์ของข้อมูลนำเข้า (input) และข้อมูลส่งออก (output)ได้ โดยไม่จำเป็นต้องรู้ความสัมพันธ์ทางคณิตศาสตร์ของข้อมูลที่เป็นข้อมูลนำเข้าและข้อมูลส่งออกมาก่อน

2.5.1 ประวัติความเป็นมาของโครงข่ายประสาทเทียม

ปี พ.ศ. 2486 (Garson. 1998) ถือได้ว่าเป็นปีแห่งการกำเนิดของสาขาโครงข่ายประสาทเทียมในวงการวิทยาศาสตร์ โดยแม็คคัลลอช (McCulloch) และพิทส์ (Pitts) ได้เสนอแบบจำลองของเซลล์ประสาท และได้แสดงให้เห็นว่าในทางทฤษฎีนั้น โครงข่ายของแบบจำลองเซลล์ประสาทสามารถทำงานร่วมกับโปรแกรมคอมพิวเตอร์ได้ต่อมาปี พ.ศ. 2492 โดนัลด์ เฮบบ์ (DonaldHebb) ได้เสนอผลงานวิจัยว่า การเรียนรู้ของสมองสามารถอธิบายได้ด้วยรูปแบบของการประกอบเซลล์ประสาทเข้าด้วยกันเป็นโครงข่าย และได้เสนอกฎการเรียนรู้ของเฮบบ์ (Hebb's rule) ที่ทำให้โครงข่ายของเซลล์ประสาทเทียมที่แม็คคัลลอชและพิทส์เสนอไว้ สามารถเรียนรู้ปัญหาง่าย ๆได้สำเร็จ การเรียนรู้ในรูปแบบของเฮบบ์บนเซลล์ประสาทเทียมของแม็คคัลลอชและพิทส์นั้นเป็นการเรียนรู้แบบ “ไม่มีผู้สอน” ซึ่งในทางปฏิบัติแล้วโครงข่ายประสาทเทียมที่เรียนรู้ จะพยายามจัดกลุ่มข้อมูลที่โครงข่ายมองว่าคล้ายคลึงกันไปไว้ในกลุ่มเดียวกัน ซึ่งไม่เหมาะสมกับปัญหาประเภทที่ต้องมีการควบคุมกระบวนการเรียนรู้

ขณะที่ในช่วงปี พ.ศ. 2490 คอมพิวเตอร์ที่ทำงานเลียนแบบสมองเครื่องแรกของโลกถูกสร้างและทดสอบโดยมินสกี้ (Minsk) ซึ่งได้เสนอผลงานดังกล่าวในปี พ.ศ. 2511 ว่าเมื่อคอมพิวเตอร์ดังกล่าวได้รับการป้อนตัวอย่างสำหรับการเรียนรู้เข้าไป ก็จะสามารถปรับอัตราการ

ขยายสัญญาณในการเชื่อมโยงหรือ "ความแข็งแรงของการเชื่อมโยง" ระหว่างเซลล์ประสาทเทียมได้เองโดยอัตโนมัติ ซึ่งเป็นการแสดงการเรียนรู้ตัวอย่างที่ถูกป้อนเข้าไป

ในปี พ.ศ. 2501 (Garson. 1998)แฟรงค์ โรเซ็นแบลทท์ (Frank Rosenblatt) ได้พัฒนาลักษณะโครงข่ายประสาทเทียมขึ้นโดยใช้แบบจำลองของแม็คคัลลอชและพิทส์เป็นแนวทาง รวมทั้งเสนอวิธีการเรียนรูแบบใหม่สำหรับลักษณะโครงข่ายประสาทเทียมด้วย โครงข่ายประสาทเทียมดังกล่าวเรียกว่า เพอร์เซพตรอน (Perceptron)ซึ่งมีการเรียนรู้แบบ "มีผู้สอน" (supervisedlearning) โดยใช้การปรับความแข็งแรงของการเชื่อมโยง ซึ่งจะพิจารณาได้จากการเปรียบเทียบความรู้ของโครงข่ายประสาทเทียมกับความรู้ของ"ผู้สอน" (teacher) เพอร์เซพตรอนมีความเหมาะสมกับงานประเภท "การระบุชนิด" ซึ่งในระหว่างการเรียนรู้นั้น เพอร์เซพตรอนจะถูกสอนว่าข้อมูลตัวอย่างที่สอนเข้าไปแต่ละแบบนั้นจัดเป็นชนิดใดบ้างหากปัญหาและข้อมูลตัวอย่างมีความเหมาะสมเพอร์เซพตรอนจะสามารถระบุชนิดของข้อมูลที่ไม่เคยเห็นมาก่อนได้ถูกต้อง

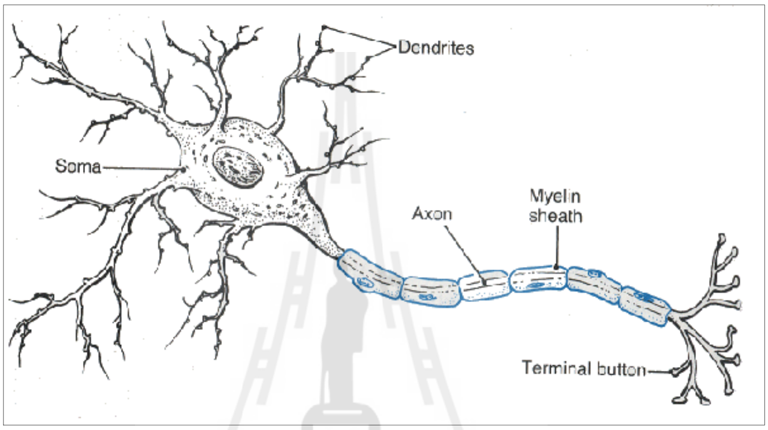
ในช่วงปี พ.ศ. 2500 เบอร์นาร์ด วิโดรว(Bernard Widrow) และมาร์เชียน ฮอฟฟ์ (MarcianHoff) ได้พัฒนาอุปกรณ์ที่เรียกว่า อดาไลน์(ADALINE; Adaptive Linear combiner) และกฎการเรียนรู้แบบใหม่ที่มีประสิทธิภาพสูง เรียกว่ากฎการเรียนรู้ของวินโดรว-ฮอฟฟ์ (Windrow-Hofflearning rule) ที่เป็นการเรียนรู้แบบ "มีผู้สอน"ซึ่งในเวลาต่อมา อุปกรณ์ดังกล่าวได้รับการขยายแนวคิดไปเป็นมาดาไลน์ (MADALINE ; ManyADALINEs) และได้ถูกนำไปประยุกต์ใช้ในการรู้จำรูปแบบ (pattern recognition) การพยากรณ์อากาศ และระบบควบคุมที่จำเป็นต้องมีการปรับเปลี่ยนระบบไปตามสภาพแวดล้อมต่าง ๆ

2.5.2 ความหมายและหลักการของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม คือ แบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์ โดยที่โครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาท(Neuron) และจุดประสานประสาทหรือไซแนปส์(Synapses) โดยโครงสร้างของการส่งสัญญาณประสาทประกอบจากการเชื่อมต่อระหว่างเซลล์ประสาทหลายพันล้านเซลล์ เซลล์ประสาทแต่ละเซลล์ประกอบด้วยแขนงรับสัญญาณประสาทซึ่งเป็นเสมือนหน่วยรับข้อมูลป้อนเข้าเรียกว่า "เดนไดรท์"(dendrites) และส่วนปลายของเซลล์ประสาทในการส่งสัญญาณประสาทซึ่งเป็นเสมือนหน่วยส่งข้อมูลออกของเซลล์ เรียกว่า "แอคซอน" (axon)โดยการส่งสัญญาณประสาทดังกล่าว อาจทำให้เกิดได้ทั้งการกระตุ้นและยับยั้ง ทั้งนี้นอกจากลักษณะดังกล่าวแล้ว วิธีการประมวลผลภายในเซลล์ประสาทแต่ละเซลล์ยังมีการขยายหรือลดขนาดของสัญญาณอีกด้วย โดยสัญญาณจากเดนไดรท์ต่าง ๆ จะรวมกันเข้าสู่เซลล์ประสาท และหากสัญญาณรวมมีความแรงเกิน

ค่าระดับ (threshold)ของเซลล์ประสาทนั้น ๆ เซลล์ประสาทก็จะส่งสัญญาณออกทางแอคซอนต่อไป

กระบวนการเรียนรู้ในสิ่งมีชีวิตจะมีผลให้เกิดการสร้างไซแนปส์ระหว่างเซลล์ประสาทขึ้นมาใหม่ หรือทำให้เกิดการเปลี่ยนสภาพของไซแนปส์ต่าง ๆ ในโครงข่ายของเซลล์ประสาท โครงข่ายประสาทของสิ่งมีชีวิตจึงไม่ได้ทำงานแบบเป็นลำดับขั้นตอน (sequential) แต่เพียงอย่างเดียว



รูปที่ 2.13 โครงสร้างระบบประสาท (http://www.mindcreators.com/NeuronBasics.htm)

โครงข่ายประสาทเทียมมีคุณลักษณะคล้ายกับการส่งผ่านสัญญาณประสาทในสมองของมนุษย์กล่าวคือ มีความสามารถในการรวบรวมความรู้(knowledge) โดยผ่านกระบวนการเรียนรู้ (learningprocess) และความรู้เหล่านี้จะจัดเก็บอยู่ในโครงข่ายในรูปแบบค่าน้ำหนัก (weight) ซึ่งสามารถปรับเปลี่ยนค่าได้เมื่อมีการเรียนรู้สิ่งใหม่ ๆ เข้าไปค่าน้ำหนักทำหน้าที่เปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ปัญหาเฉพาะอย่างของมนุษย์การประมวลผลต่าง ๆ เกิดขึ้นในหน่วยประมวลผลย่อย เรียกว่า โหนด (node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ (signal) ระหว่างโหนดที่เชื่อมต่อกัน (connection) จำลองมาจากการเชื่อมต่อของเดนไดร์ทและแอคซอนในระบบประสาทของมนุษย์ ภายในโหนดจะมีฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่า ฟังก์ชันกระตุ้น (activation function) หรือฟังก์ชันการแปลง (transfer function)ซึ่งทำหน้าที่เปรียบเสมือนกระบวนการทำงานในเซลล์ ดังรูปที่ 2.13โครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบ ดังนี้

1. ข้อมูลป้อนเข้า (input) เป็นข้อมูลที่เป็นตัวเลข หากเป็นข้อมูลเชิงคุณภาพ ต้องแปลงให้อยู่ในรูปเชิงปริมาณที่โครงข่ายประสาทเทียมยอมรับได้

2. ข้อมูลส่งออก (output) คือ ผลลัพธ์ที่เกิดขึ้นจริง (actual output) จากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม

3. ค่าน้ำหนัก (weights) คือ สิ่งที่ได้จากการเรียนรู้ของโครงข่ายประสาทเทียม หรือเรียกอีกอย่างหนึ่งว่า ค่าความรู้ (knowledge) ค่านี้จะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูลอื่นๆ ที่อยู่ในรูปแบบเดียวกัน

4. ฟังก์ชันผลรวม (Summation function: S) เป็นผลรวมของข้อมูลป้อนเข้า ( ) และค่าน้ำหนัก (wi)

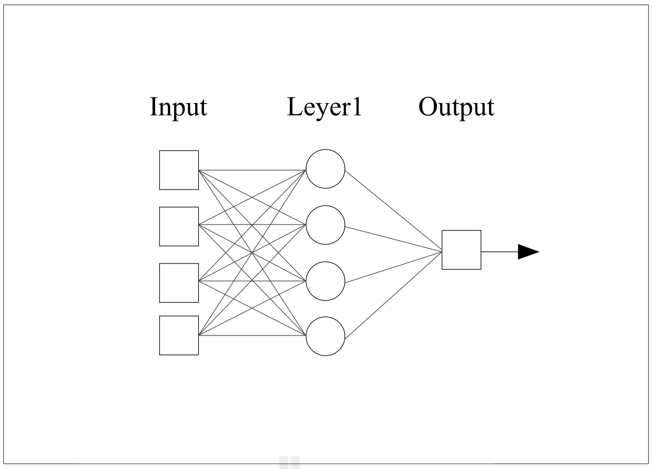
5. ฟังก์ชันการแปลง (transfer function) เป็นการคำนวณการจำลองการทำงานของโครงข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน (sigmoid function) ฟังก์ชันไฮเปอร์โบลิกแทนเจนต์ (hyperbolic tangent function) เป็นต้น

2.5.3 ลักษณะของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมประกอบด้วยเซลล์ประสาทเทียม หรือโหนดจำนวนมากเชื่อมต่อกันซึ่งการเชื่อมต่อแบ่งออกเป็นกลุ่มย่อย เรียกว่า ชั้น(layer) ชั้นแรก เป็นชั้นนำข้อมูลเข้าเรียกว่า ชั้นรับข้อมูลป้อนเข้า (input layer) ส่วนชั้นสุดท้ายเรียกว่าชั้นส่งข้อมูลออก (output layer) และชั้นที่อยู่ระหว่างชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเรียกว่า ชั้นแอบแฝง (hidden layer) ซึ่งโดยทั่วไปชั้นแอบแฝงอาจมีมากกว่า 1 ชั้นก็ได้ ด้วยเหตุนี้ จึงสามารถแบ่งประเภทของโครงข่ายประสาทเทียมตามจำนวนชั้นของโครงข่ายแบบกว้าง ๆ ได้ 2 แบบได้แก่ โครงข่ายแบบชั้นเดียว (single layer) และ โครงข่ายแบบหลายชั้น (multi-layer)

2.5.3.1 โครงข่ายแบบชั้นเดียว

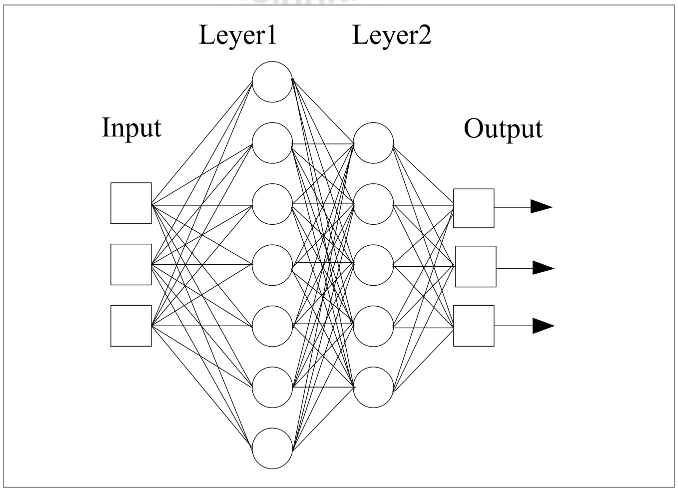
โครงข่ายแบบชั้นเดียว เป็นโครงข่ายประสาทเทียมอย่างง่ายที่มีเพียงชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเท่านั้น โหนดในชั้นรับข้อมูลป้อนเข้าทำหน้าที่รับข้อมูลเข้า (input value) แล้วส่งข้อมูลผ่านเส้นเชื่อมโยงต่าง ๆ ไปให้โหนดในชั้นส่งข้อมูลออก ความเข้มของสัญญาณหรือปริมาณข้อมูลที่นำเข้าสู่โหนดในชั้นส่งข้อมูลออกจะขึ้นอยู่กับค่าน้ำหนักที่อยู่บนเส้นเชื่อมโยง โหนดในชั้นส่งข้อมูลออกจะนำข้อมูลที่ได้รับมาคำนวณโดยใช้ฟังก์ชันทางคณิตศาสตร์ที่เรียกว่าฟังก์ชันการแปลง (transfer function)ที่เหมาะสมกับปัญหา แล้วส่งผลลัพธ์ที่ได้ออกมาเป็นข้อมูลส่งออก เช่น โครงข่ายแบบชั้นเดียวแบบเพอเซบตรอนอย่างง่าย (simple perceptron) และโครงข่ายโฮบฟิลด์ (Hopfield networks)ลักษณะโครงข่ายแบบชั้นเดียวแสดงดังรูปที่ 2.14



รูปที่ 2.14 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว

2.5.3.2. โครงข่ายแบบหลายชั้น

โครงข่ายแบบหลายชั้น เป็นโครงข่ายที่มีชั้นแอบแฝงตั้งแต่ 1 ชั้นขึ้นไป โครงข่ายแบบหลายชั้นจะใช้ในกรณีที่ปัญหามีความซับซ้อน ซึ่งโครงข่ายแบบชั้นเดียวไม่สามารถแก้ปัญหาได้ จึงเพิ่มจำนวนโหนดที่มีการคำนวณ หรือชั้นแอบแฝง ให้กับโครงข่าย ตัวอย่างของโครงข่ายแบบหลายชั้นเช่น การแพร่ย้อนกลับ (back propagation) เซลฟ์ออร์แกนไนซิ่งแมปซ์ (self-organizing maps)และเคาน์เตอร์พรอพะเกชัน (counter propagation) เป็นต้น ลักษณะโครงสร้างโครงข่ายแบบหลายชั้นแสดงดังรูปที่ 2.15



รูปที่ 2.15 โครงข่ายประสาทเทียมแบบหลายชั้น

2.5.4 ประเภทของการเรียนรู้ของโครงข่ายประสาทเทียม

2.5.4.1 การเรียนรู้แบบมีผู้สอน (Supervised learning)

ข้อมูลจะประกอบด้วยตัวอย่างข้อมูลที่ต้องการสอน และผลลัพธ์ที่ต้องการให้โครงข่ายสร้าง เมื่อมีการนำข้อมูลในลักษณะเดียวกันมาเป็นข้อมูลป้อนเข้า โครงข่ายจะกำหนดค่าผลลัพธ์ที่เป็นเป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัวโครงข่ายจะนำค่าผิดพลาดระหว่างค่าเป้าหมายกับค่าผลลัพธ์ที่ได้ มาใช้ในการปรับค่าน้ำหนัก เพื่อให้ค่าผลลัพธ์ที่ใกล้เคียงกับเป้าหมายมากที่สุด ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการสอนนักเรียนโดยมีครูผู้สอนคอยให้คำแนะนำ ตัวอย่างแบบจำลองนี้ได้แก่ การแพร่ย้อนกลับ และเพอเซบตรอน (perceptron) เป็นต้น

2.5.4.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)

การเรียนรู้แบบนี้จะสอนโครงข่ายโดยการนำข้อมูลป้อนเข้าอย่างต่อเนื่องเพียงอย่างเดียว ไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัว การปรับน้ำหนักจะใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่า โดยค่าน้ำหนักจะปรับตามกลุ่มที่ข้อมูลป้อนเข้าที่มีรูปแบบคล้ายคลึงกันถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการที่เราสามารถแยกแยะพันธุ์พืช พันธุ์สัตว์ ตามลักษณะรูปร่างของมันได้ด้วยตนเอง ตัวอย่างแบบจำลองนี้ได้แก่ เคาน์เตอร์พรอพะเกชัน (counter propagation: CPN) แบบจำลองอะแดปทีฟรีโซแนนซ์เทียรี่ (Adaptive Resonance Theory neural networks: ART) เป็นต้น

2.5.5 การประยุกต์ใช้งานโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมเป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ที่ประยุกต์ใช้ความรู้จากหลากหลายสาขามารวมเข้าด้วยกัน นับว่าเป็นศาสตร์ที่กำลังมีบทบาทอย่างยิ่งในปัจจุบัน กล่าวโดยสรุปโครงข่ายประสาทเทียมสามารถนำมาประยุกต์ใช้กับงานด้านต่างๆ ได้ดังนี้

1. การจำแนกรูปแบบ (pattern recognition) เช่น การมองเห็นวัตถุ หรือการวิเคราะห์เสียงพูดเพื่อแปลความหมาย

2. การทำนาย (prediction) หรือการพยากรณ์ (forecasting) เช่น การทำนายราคาหุ้นของตลาดหลักทรัพย์ การพยากรณ์อัตราการไหลของน้ำ การพยากรณ์ราคาสินค้า เป็นต้น

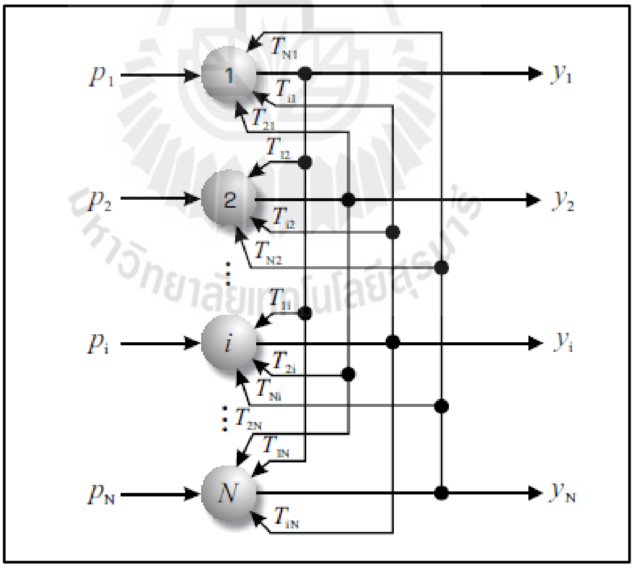
3. การควบคุม (control) เช่น การควบคุมระบบของเครื่องปรับอากาศ การควบคุมระบบเครื่องยนต์ และการควบคุมหุ่นยนต์ เป็นต้น

4. การหาความเหมาะสม (optimization) เช่น การเลือกระยะทางที่ใกล้ หรือสั้นที่สุดในการเดินทาง (shortest path)

5. การจัดกลุ่ม (clustering) และการจัดหมู่ (categorization) เช่น การวิเคราะห์ข้อมูลจากภาพถ่ายดาวเทียม หรือภาพถ่ายทางอากาศ

2.6 เครือข่ายฮอพฟิลด์

เครือข่ายฮอพฟิลด์เป็นเครือข่ายอัตสัมพันธ์ (autoassociative network) อย่างหนึ่งซึ่งประกอบไปด้วยนิวรอนชั้นเดียวที่มีการเชื่อมต่อกันอย่างทั่วถึง เครือข่ายฮอพฟิลด์สามารถทำงานได้ในลักษณะเดียวกันกับหน่วยความจำสัมพันธ์ (associative memory) งานของฮอพฟิลด์เป็นการนำเสนอแนวความคิดในรูปแบบผสมผสานระบบความจำในเครือข่ายในรูปของฟังก์ชันพลังงาน (จากสมการ Lyapunov) รวมไปถึงการประมวลผลแต่ละนิวรอนแบบไม่พร้อมเพรียงกัน (asynchronous processing) ข้อแตกต่างนี้ทำให้เครือข่ายฮอพฟิลดามารถลดปัญหาการใช้งานกับเครือข่ายที่มีขนาดใหญ่ได้ ข้อดีอีกอย่างหนึ่งก็คือองค์ประกอบในการประมวลผลไม่จำเป็นจะต้องใช้ข้อมูลความรู้จากเครือข่ายทั้งระบบ



รูปที่ 2.16 โครงสร้างของเครือข่ายฮอพฟิลด์

2.6.1 โครงสร้างของเครือข่ายฮอพฟิลด์

รูปที่ 2.12 แสดงโครงสร้างของเครือข่ายฮอพฟิลด์ สำหรับเครือข่ายแบบไม่ต่อเนื่อง (discrete) จะใช้เพียงค่าไบนารี (0 และ 1) หรือค่าสองขั้ว (1 และ -1) โครงสร้างดังกล่าว

ประกอบไปด้วย N นิวรอนซึ่งสามารถใช้กับข้อมูลแบบอินพุต p ที่มีขนาด N แต่ละนิวรอนจะถูกพิจารณาในรูปของสถานะ S (state) นิรอนแต่ละตัวจะเชื่อมโยงกับนิวรอนอื่นๆ ทั้งหมดในระบบยกเว้นตัวมันเอง ค่าน้ำหนักประสาทระหว่างนิวรอน i และ j คือ T_{ij} มีคุณสมบัติดังนี้

$$T_{ij} = T_{ji} \tag{2.17}$$

$$T_{ii} = 0 \tag{2.18}$$

สมการที่ 2.17 แสดงความสมมาตรของค่าน้ำหนักประสาทระหว่างนิวรอน i และ j ในขณะที่สมการที่ 2.18 แสดงว่าไม่มีการเชื่อมโยงค่าน้ำหนักประสาทของนิวรอนกับตัวมันเอง

2.6.2 การเรียนรู้ของเครือข่ายฮอพฟิลด์

สถานะของเครือข่ายฮอพฟิลด์จะขึ้นอยู่กับค่าพลังงานของนิวรอนในเครือข่าย การเปลี่ยนสถานะของนิวรอนขึ้นอยู่กับการเปลี่ยนแปลงของพลังงานดังกล่าวด้วย สถานะของเครือข่ายมีความสัมพันธ์กับพลังงานภายในเครือข่ายดังต่อไปนี้

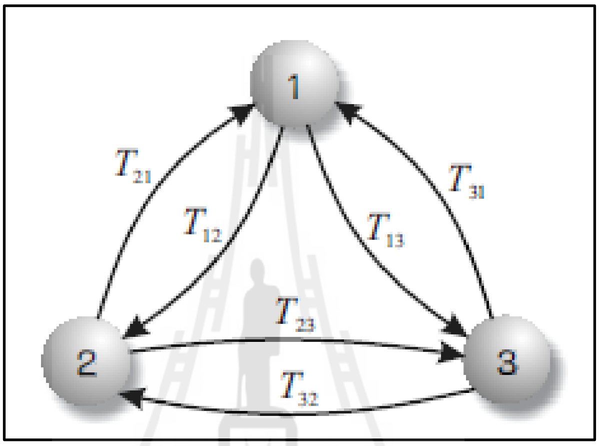
$$E = -\frac{1}{2}\sum_j \sum_{i,i\neq j} T_{ij} y_i y_j \tag{2.19}$$

โดยที่ T_{ij} คือค่าน้ำหนักประสาทระหว่างนิวรอนที่ i ไปยังนิวรอนที่ j และ y_i เป็นเอาต์พุตของนิวรอนที่ i ภายในเครือข่าย รูปที่ 2.13 แสดงการไหลเวียนของพลังงานในเครือข่ายพร้อมกับการเปลี่ยนสถานะที่เป็นไปได้ของเครือข่ายแบบ 3 และ 4 นิวรอน

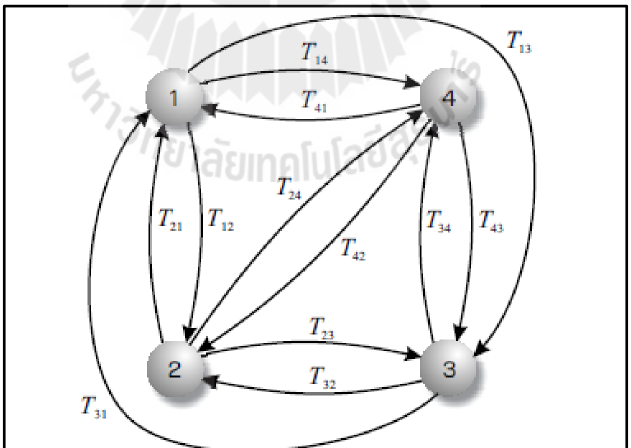
สถานะของเครือข่ายแทนด้วย $S^i(y_1, y_2, ..., y_N)$ โดยที่ $i = 0, 1, 2, \ldots, 2^N$ (จำนวนสถานะที่เป็นไปได้ทั้งหมดของเครือข่ายเท่ากับ 2^N) สังเกตว่าการเปลี่ยนสถานะของเครือข่ายจะเปลี่ยนทีละนิวรอน กล่าวคือในการเปลี่ยนสถานะแต่ละครั้ง องค์ประกอบ yi เพียงหนึ่งตัวเท่านั้นที่มีการเปลี่ยนสถานะ

เช่นเดียวกันกับเครือข่ายอื่นๆ เครือข่ายฮอพฟิลด์จะมีการวนรอบปรับค่าน้ำหนักประสาท โดยการปรับจะทำการปรับเพียงหนึ่งนิวรอนต่อครั้งจนกระทั่งค่าน้ำหนักประสาทลู่เข้าสู่คำตอบที่ต้องการ การลู่เข้าของเครือข่ายจะเกิดขึ้นเมื่อฟังก์ชั่นพลังงานภายในเครือข่ายให้ค่าน้อยที่สุดและไม่มีนิวรอนใดๆเกิดการเปลี่ยนสถานะ เครือข่ายฮอพฟิลด์ได้รับการพิสูจน์แล้วว่าสามารถลู่เข้าสถานะที่มีค่าพลังงานของเครือข่ายน้อยที่สุดได้ เมื่อมีการปรับค่าน้ำหนักประสาทอย่างเหมาะสม การเรียนรู้ของเครือข่ายฮอพฟิลด์เป็นไปในรูปแบบของฟังก์ชั่นพลังงาน เมื่อมีการ

นำเสนอรูปแบบอินพุตให้กับเครือข่าย เครือข่ายจะทำการปรับค่าน้ำหนักประสาทจนกระทั่งได้สถานะในการจดจำรูปแบบนั้น ๆ ที่ซึ่งฟังก์ชันพลังงานให้ค่าต่ำที่สุด ในระหว่างการเรียนรู้ของเครือข่ายฮอพฟิลด์ สามารถพิจารณาสถานะของเครือข่าย y_i ว่าเป็นหน่วยความจำระยะสั้น (short-term memory) ที่ซึ่งสามารถบันทึกและเรียกใช้รูปแบบอินพุตได้โดยการปรับเปลี่ยน y_i ในทำนองเดียวกัน น้ำหนักประสาท T_{ij} สามารถพิจารณาเป็นหน่วยความจำระยะยาว (long-term memory) ซึ่งถูกเก็บค่าไว้ภายในตัวของเครือข่ายเอง



(ก)



(ข)

รูปที่ 2.17 แผนภาพค่าน้ำหนักประสาท
(ก) เครือข่ายฮอพฟิลด์ที่มี 3 นิวรอน
(ข) เครือข่ายฮอพฟิลด์ที่มี 4 นิวรอน

2.6.3 การปรับค่าน้ำหนักประสาทของเครือข่ายฮอพฟิลด์

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

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

เอาต์พุตของนิวรอนตัวที่ j จะมีค่าดังนี้

$$y_i = \begin{cases} 1, s_j \geq 0 \\ 0, s_j < 0 \end{cases} \tag{2.21}$$

เนื่องจากในแต่ละรอบ เครือข่ายฮอพฟิลด์มีการปรับค่าน้ำหนักประสาทเพียงนิวรอนเดียว พิจารณานิวรอน j จะได้พลังงานของนิวรอน j นี้คือ

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

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

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

$$\Delta E_j = \Delta E_{\mathrm{j}}^{new} - \Delta E_{\mathrm{j}}^{old} \tag{2.24}$$

จากสมการที่ 2.23 จะได้

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

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

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

$$= -\frac{1}{2}\Delta y_j s_j \tag{2.28}$$

จากสมการข้างต้น สามารถแบ่งพิจาณาการเปลี่ยนแปลงของพลังงาน ΔE_j ของนิวรอนตัวที่ j ได้เป็น 3 กรณีดังนี้

1. ถ้านิวรอนตัวที่ j ไม่เปลี่ยนแปลงสถานะ นั่นคือ $\Delta y_j = 0$ ดังนั้นไม่มีการเปลี่ยนแปลง

ของพลังงาน

2. ถ้านิวรอนตัวที่ j เริ่มต้นที่สถานะ *1* และเปลี่ยนสถานะไปเป็น *0* จะได้ว่า

$$y_j^{old} = 1 \tag{2.29}$$

$$y_j^{new} = 0 \tag{2.30}$$

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

$$\Delta E_j < 0 \tag{2.31}$$

3. ถ้านิวรอนตัวที่ j เริ่มต้นที่สถานะ *0* และเปลี่ยนไปเป็น *1* จะได้ว่า

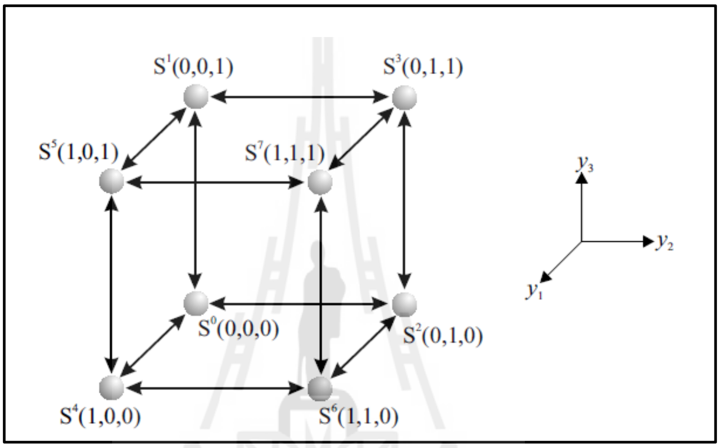
$$y_j^{old} = 0 \tag{2.32}$$

$$y_j^{new} = 1 \tag{2.33}$$

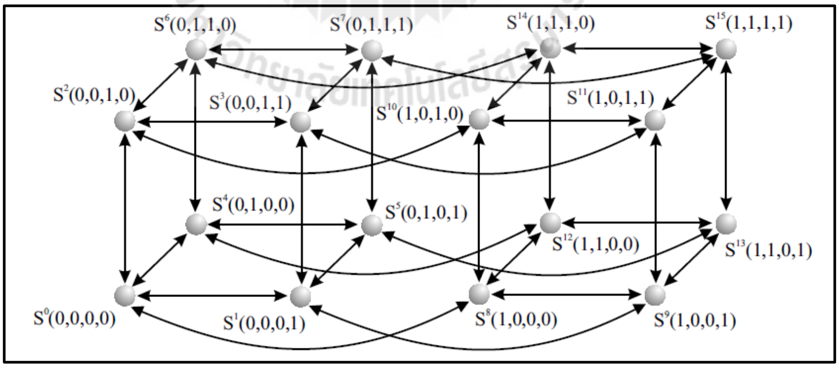
ดังนั้น $\Delta y_j = 1$ ซึ่งมีค่ามากกว่า *0* เมื่อพิจารณาจากสมการที่ *(2.21)* จะได้ว่า y_iจะเปลี่ยนสถานะจาก *0* ไปเป็น *1* เมื่อ $sj \geq 0$ ดังนั้นผลคูณ $\Delta y_j s_j$ จึงมีค่ามากกว่าหรือเท่ากับ *0* ซึ่งจะสรุปได้ว่า

$$\Delta E_j < 0 \tag{2.34}$$

จากการวิเคราะห์ข้างต้น ไม่ว่าการเปลี่ยนแปลงสถานะจะเป็นไปในทิศทางใดก็ตาม เครือข่ายฮอพฟิลด์จะให้ผลของต่าพลังงานที่ลดลงหรือคงที่เสมอ เมื่อทุกนิวรอนไม่มีการเปลี่ยนแปลงสถานะ การเปลี่ยนแปลงของพลังงานจะเท่ากับศูนย์หรือพลังงานมีค่าคงที่ ซึ่งสถานะที่เครือข่ายมีความเสถียรภาพ ณ ค่าพลังงานที่ต่ำที่สุดการประเมินเครือข่ายจะเริ่มจากการสุ่มนิวรอนแล้วทำการปรับค่าน้ำหนักประสาทของเครือข่าย เมื่อทุกนิวรอนได้รับการประเมิน และไม่มีนิรอนใดเกิดการเปลี่ยนแปลงสถานะแสดงว่าเครือข่ายได้ลู่เข้าแล้ว



(ก)

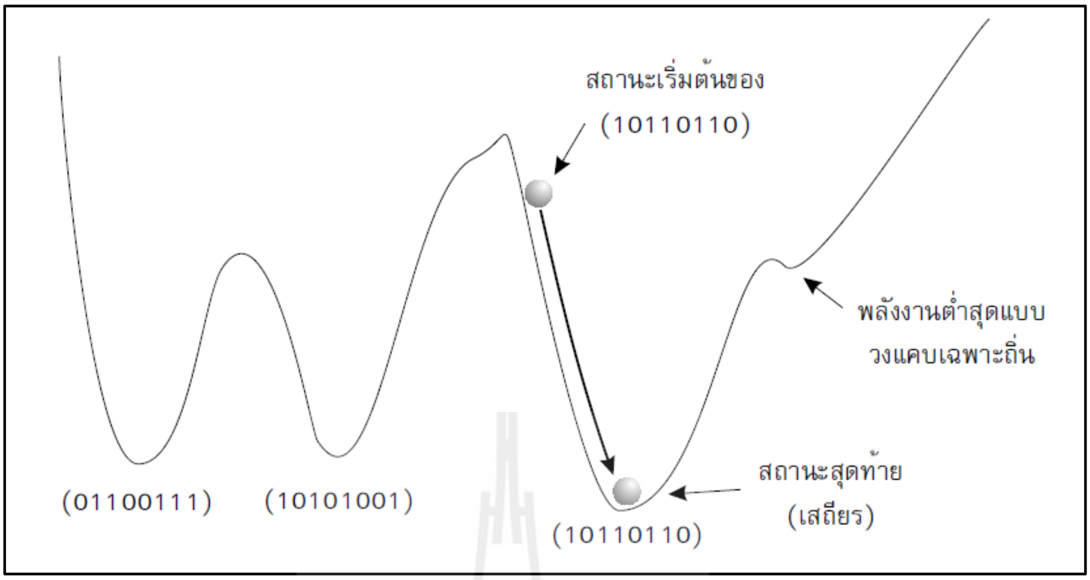


(ข)

รูปที่ 2.18 การไหลเวียนของพลังงานและการเปลี่ยนสถานะของเครือขาย

(ก) 3 นิวรอน

(ข) 4 นิวรอน



รูปที่ 2.19 พื้นผิวพลังงานเครือข่ายฮอพฟิล์ด

2.7 ทฤษฎีเสถียรภาพเลียปูนอฟ (Lyapunov stability theory)

ทฤษฎีเสถียรภาพเลียปูนอฟถูกนำเสนอครั้งแรกในช่วงปลายปี พ.ศ. 2435 โดยนักคณิตศาสตร์ชาวรัสเซียชื่อ A.M. Lyapunov โดยทฤษฎีหลักที่เกี่ยวข้องกับการวิเคราะห์เสถยีรภาพของระบบที่รู้จักกันทั่วไปมีอยู่ 2 ทฤษฎี คือ ทฤษฎีประมาณเชิงเส้น (linearization method) และทฤษฎีจากการวิเคราะห์โยตรง (direct method) โดยทษฤฎีจากการประมาณแบบเชิงเส้นเป็นการวิเคราะห์เสถียรภาพเฉพาะที่ (local stability) รอบจุดสมดุล ซึ่งมีข้อจำกัดของการใช้งานค่อนข้างมาก ในขณะที่ทฤษฎีจากการวิเคราะห์โยตรงสามารถจะใช้วิเคราะห์เสถียรภาพในวงกว้าง (global stability) ได้ ทำให้ทฤษฎีพื้นฐานหลักทางคณิตศาตร์ที่ใช้ทั่วไปสำหรับการวิเคราะห์เสถียรภาพระบบไม่เชิงเส้น

ความหมายหลักของเสถียรภาพตามแนวคิดของเลียปูนอฟคือจุดสมดุลถือว่าเป็นจุดสมดุลที่มีเสถียรภาพ ถ้าวิธีการเคลื่อนที่ของระบบภายหลังถูกรบกวนหรือกระทำให้เคลื่อนที่ออกจาสมดุลนั้นยังคงอยู่รอบๆจุดสมดุลนั้นตลอดไปและในกรณีที่เมื่อเวลาผ่านไปวิถีการเคลื่อนที่ลู่เข้าหาจุดสมดุลนั้นในที่สุด จะเรียกจุดสมดุลนี้ว่า เป็นจุดสมดุลที่มีเสถียรภาพแบบลู่เข้า (asymptotically stable) ส่วนจุดสมดุลที่ไม่มีคุณสมบัติตามความหมายของเสถียรภาพข้างต้นถือว่าเป็นจุดสมดุลที่ไม่เสถียรภาพ

ระบบเชิงเส้นที่มีจุดสมดุลที่มีเสถียรภาพจะเป็นระบบมีเสถียรภาพในวงกว้าง (globally stable) เสมอ แต่ข้อสรุปนี้ไม่เป็นจริงเสมอไปสำหรับระบบไม่เชิงเส้นด้วยเหตุผลหลายประการ เช่น ระบบไม่เชิงเส้นอาจจะมีจุดสมดุลหลายจุด ระบบไม่เชิงเส้นอาจจะมีพฤติกรรมที่ไม่ต่อเนื่อง เป็นต้น ส่งผลให้เสถียรภาพอาจจะเกิดขึ้นเฉพาะรอบๆจุดสมดุลนั้นด้วยขอบเขตจำกัดเท่านั้น ซึ่งสามารถสรุปได้เพียงว่าระบบมีเสถียรภาพเฉพาะที่ (locally stable)

2.7.1 ทฤษฎีลัยปูนอฟจากการประมาณแบบเชิงเส้น

ทฤษฎีลัยปูนอฟจากการประมาณแบบเชิงเส้นเป็นการวิเคราะห์เสถียรภาพรอบจุดสมดุลเฉพาะที่ของระบบไม่เชิงเส้นแบบไม่มีอินพุต ซึ่งมีขั้นตอนการวิเคราะห์ดังนี้

พิจารณาระบบแบบไม่มีอินพุตเขียนในรูปของสมการสถานะดังนี้

$$x = f(x) \tag{2.35}$$

โดยที่ x คือตัวแปรสถานะ และ f(x) คือ ฟังก์ชันเวกเตอร์ที่มีมิติเท่ากับตัวแปรสถานะ ในการวิเคราะห์จะเริ่มต้นด้วยการทำการประมาณระบบด้วยสมการระบบเชิงเส้นรอบจุดสมดุลที่สนใจ ในกรณีที่จุดสมดุลไม่เท่ากับศูนย์ ให้ทำการเปลี่ยนตัวแปร เพื่อให้จุดสมดุลของระบบตัวแปรใหม่เป็นศูนย์เสียก่อน จากนั้นทำการวิเคราะห์เสถียรภาพของสมการประมาณการที่เป็นระบบเชิงเส้น โดยใช้ทฤษฎีระบบเชิงเส้น เช่น การพิจารณาจากค่าโพล (pole) หรือค่าไอเกน (Eigenvalue) ของระบบ เป็นต้น สุดท้ายทฤษฎีเลียปูนอฟจากการประมาณแบบเชิงเส้นสรุปผลการวิเคราะห์ไว้ดังนี้

1. ถ้าสมการประมาณการด้วยระบบเชิงเส้นมีเสถียรภาพโดยแท้ (strictly stable) นั้นคือโพลทุกตัวของระบบเชิงเส้นมีค่าอยู่ทางซ้ายมือของแกนจินตภาพ จะสรุปได้ว่า จุดสมดุลของระบบไม่เชิงเส้นมีเสถียรภาพแบบลู่เข้า ดังนั้นระบบมีเสถียรภาพเฉพาะที่รอบจุดสมดุลที่สนใจ

2. ถ้าสมการประมาณการด้วยระบบเชิงเส้นไม่มีเสถียรภาพ (unstable) นั้นคือ มีโพลบางตัวของระบบเชิงเส้นมีค่าอยู่ทางขวามือของแกนจินตภาพ จะสรุปได้ว่า จุดสมดุลของระบบไม่เชิงเส้นไม่มีเสถียรภาพ

3. ถ้าสมการประมาณการด้วยระบบเชิงเส้นมีเสถียรภาพขอบ (marginally stable) นั้นคือ มีโพลงบางตัวของระบบเชิงเส้นมีค่าอยู่บนแกนจินตภาพ ในขณะที่เหลือมีค่าอยู่ทางซ้ายมือของแกนจอตภาพ จะไม่สามารถสรุปเสถียรภาพของระบบได้

จากรายละเอียดข้างต้น จะเห็นชัดเจนว่า จุดด้อยของทฤษฎีเลียปูยอฟจากการประมาณแบบเชิงเส้นประกอบด้วย การให้ผลลัพธ์ของการวิเคราะห์เสถีรภาพเป็นเพียงเสถียรภาพเฉพาะรอบจุดที่สมดุลสมดุลเท่านั้น แต่ทั้งนี้ไม่สามารถจะบอกได้ว่าระบบสมารถถูกรบกวนให้ห่างออกจากสมดุลไปได้ไกลเพียงไหน โดยที่ระบบยังคงมีเสถียรภาพอยู่ นอกจากนี้แล้ว ในกรณีของข้อที่ 3 ซึ่งไม่สามารถจะสรุปเสถียรภาพได้เลย ถ้าสมการประมาณการด้วยระบบเชิงเส้นมีเพียงเสถียรภาพขอบ อย่างไรข้อดีของทฤษฎีเลียปูนอฟจากการประมาณแบบเชิงเส้นคือ ความง่ายของการนำไปใช้งาน และเมื่อนำมารวมกับการใช้เทคนิคการจำลองการทำงานในคอมพิวเตอร์ สามารถที่จะบอกถึงเสถัยรภาพของระบบได้ระดับหนึ่งเช่นกัน

2.7.2 ทฤษฎีเลียปูนอฟจากการวิเคราะห์โดยตรง

ทฤษฎีเลียปูนอฟจากการวิเคราะห์โดยตรงเป็นทฤษฎีพื้นฐานหลักทางคณิตศาสตร์ที่ใช้ทั่วไปสำหรับการวิเคราะห์เสถียรภาพของจุดสมดุลของระบบไม่เชิงเส้น และสามารถจะใช้วิเคราะห์เสถียรภาพในวงกว้างได้ แนวคิดของทฤษฎีมาจากการสังเกตความสัมพันธ์ระหว่างระดับพลังงานและเสถียรภาพของระบบ เช่น ถ้าผลรวมของพลังงานจลน์และพลังงานศักย์ของระบบพลวัตลดลงอย่างต่อเนื่องสุดท้ายระบบจะหยุดนิ่ง ณ จุดสมดุล ซึ่งเป็นการบ่งบอกว่าถึงเสถียรภาพของจุดสมดุลของระบบนั้นเอง โดยพบว่าข้อสังเกตนี้เป็นจริงทั้งระบบเชิงเส้นและระบบไม่เชิงเส้น ทฤษฎีจากการวิเคราะห์โดยตรงของเลียปูนอฟสมามารถสรุปได้ดังต่อไปนี้

พิจารณาระบบแบบไม่มีอินพุตเขียนในรูปของสมการสถานะดังสมการที่ 2.35 โดยที่ x คือตัวแปรสถานะ และ $f(x)$ คือฟังก์ชันเวกเตอร์ที่มีมิติเท่ากับตัวแปรสถานะ และสมมุติให้จุดศูนย์เป็นจุดสมดุลที่ต้องการวิเคราะห์เสถียรภาพ นั้นคือ $x^* = 0$ และ $f(0) = 0$ ซึ่งในกรณีที่จุดสมดุลไม่ใช่จุดศูนย์ ก็ให้ทำการเปลี่ยนตัวแปรให้ระบบมีจุดศูนย์เป็นจุดสมดุลก่อนในลักษณะเช่นเดียวกับทฤษฎีเลียปูนอฟจากการประมาณแบบเชิงเส้นที่ผ่านมา ชั้นตอนในการวิเคราะห์ จะต้องกำหนดฟังก์ชันสเกลลาร์ ที่เป็นฟังก์ชั่นบวกแน่นอน (positive definite) ขึ้นมา ซึ่งฟังก์ชั่นอาจจะมีความหมายทางกายภาพ เช่น แสดงถึงระดับพลังงานของระบบ หรืออาจจะไม่มีความหมายทางกายภาพก็ได้

ทฤษฎีเสถียรภาพเลียปูนอฟนี้เป็นเพียงเงื่อนไขแบบเพียงพอ (sufficient condition) ไม่ใช่เงื่อนไขจำเป็น (necessary condition) ดังนั้นถ้าไม่สามารถหาฟังก์ชันเลียปูนอฟของระบบได้ ก็ยังไม่สามารถสรุปอะไรเกี่ยวกับเสถียรภาพได้ ซึ่งนี้ถือว่าเป็นข้อด้อยของทฤษฎีนี้

2.8 งานวิจัยที่เกี่ยวข้อง

วิธีการหาแผนที่บ่งบอกความลึกจากงานวิจัยต่าง ๆ ที่ผ่านมา (Barnard, Fischler, 1982 และ Dhond, Aggarwal, 1989) ได้ศึกษาเกี่ยวกับการเชื่อมต่อของระบบการมองเห็นของคอมพิวเตอร์ ได้พบว่าปัญหาหลักของระบบการมองเห็นแบบกล้อง 2 ตัวอยู่ที่การหาความสอดคล้องภายในภาพจากทั้ง 2 กล้อง เพื่อให้ได้ความลึกของสภาวะแวดล้อมภายในภาพเรียกว่าแผนที่บ่งบอกความลึก ต่อมามีงานวิจัยมากมายได้พัฒนาวิธีการหาแผนที่บ่งบอกความลึก 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) เป็นการสร้างกล่องจุดภาพเพื่อใช้ในการเปรียบเทียบหาความสัมพันธ์ของจุดภาพจากกล้องทั้ง 2 ตัวและลดข้อผิดพลาดจากการใช้วิธีคำนวณทุกจุด ซึ่งจากงานวิจัยนี้สามารถหาแผนที่บ่งบอกความลึกที่มีความชัดเจนได้ดีขึ้น จนในปี 1998 (S. Birchfield, C. Tomasi, 1998) ได้ศึกษาวิจัยแผนที่บ่งบอกความลึกด้วยวิธีจุดภาพต่อจุดภาพ (pixel-to-pixel) โดยการวัดค่าความเข้มของแต่ละจุดภาพเปรียบเทียบกัน ซึ่งผลที่ได้คือสามารถหาแผนที่บ่งบอกความลึกที่มีความชัดเจนขึ้นและหาบริเวณที่ซับซ้อนได้โดยใช้เวลาน้อยลง แต่ปัญหาที่เกิดขึ้นกับวิธีการนี้คืออัลกอริทึมในการคำนวณยุ่งยากและเหมาะสมกับงานเพียงบางประเภท ต่อมา (Y. Boykov, O. Veksler, and R. Zabih, 2001) ทำการวิจัยหาแผนที่บ่งบอกความลึกด้วยวิธีกราฟคัท (graph cut) โดยอัลกอริทึมวิธีกราฟคัทคำนวณหาความสัมพันธ์ของจุดภาพ ซึ่งแผนที่บ่งบอกความลึกที่หาขึ้นมีความชัดเจนสามารถแสดงรายละเอียดของวัตถุได้ แต่ปัญหาที่เกิดคืออัลกอลิทึมในการคำนวณยุ่งยากและใช้เวลานานเหมาะกับงานบางประเภท

จากผลงานวิจัยที่ผ่านมาจะเห็นได้ว่าการคำนวณหาค่าความสอดคล้องกันของจุดภาพนั้นมีอัลกอลิทึมที่ซับซ้อน (N. Nasrabadi, C. Choo, 1992) ได้นำหลักการของเครือข่ายฮอพฟิลด์มาใช้ศึกษาหาความสอดคล้องกันภายในภาพโดยนำจุดเด่นของทั้ง 2 ภาพ เข้าไปคำนวณในเครือข่ายฮอพฟิลด์ 2 มิติ (2D Hopfield Neuron Network) เพื่อลดการคำนวณทุกจุดภาพลง ซึ่งงานวิจัยนี้

สามารถหาแผนที่บ่งบอกความลึกได้ดีขึ้น มีความผิดพลาดเล็กน้อยและใช้เวลาในการคำนวณน้อยลง ลดข้อจำกัดของการติดตั้งกล้อง

หลังจากทำการหาแผนที่บ่งบอกความลึกแล้ว ขั้นตอนต่อไปคือการบูรณะสภาวะ 3 มิติจากแผนที่บ่งบอกความลึก การบูรณะสภาวะ 3 มิติคือการแปลงข้อมูลจากแผนที่บ่งบอกความลึกด้วยการคำนวณเชิงเลขาคณิต (triangular) เปลี่ยนเป็นข้อมูล 3 มิติหรือพิกัดจริง โดยที่แต่ละจุดภาพในแผนที่บ่งบอกความลึกแสดงถึงข้อมูลค่าความไม่เท่ากันของจุดภาพ (d) ซึ่งเป็น ตัวแปรที่สำคัญที่ใช้บอกความลึกของภาพที่ใช้ในการบูรณะสภาวะ 3 มิติ นอกจากนั้นระยะความยาวโฟกัสของกล้อง (f) และระยะห่างของการตั้งกล้องแบบขนานกันระหว่างกล้อง 2 ตัว (L) ก็เป็นตัวแปรสำคัญสำหรับการคำนวณหาสภาวะ 3 มิติหรือพิกัดจริงจากแผนที่บ่งบอกความลึก

สำหรับผลการบูรณะสภาวะ 3 มิติด้วยวิธีการนี้ (H. Kim, S. Yang, K. Sohn, 2003) กล่าวว่า การบูรณะสภาวะ 3 มิติด้วยวิธีการแปลงเชิงเรขาคณิต สามารถบูรณะสภาวะ 3 มิติจากแผนที่บ่งบอกความลึกได้ ซึ่งสามารถแยกระยะห่างของตำแหน่งวัตถุออกจากพื้นหลังได้ แต่ไม่สามารถบอกรายละเอียดที่ซับซ้อนของวัตถุหรือระยะห่างของวัตถุที่วางใกล้กันมาก ๆ ส่วนถ้าต้องการสภาวะ 3 มิติที่ต้องการรายละเอียดที่ซับซ้อน เช่น การสร้างหน้าคนหรือการค้นหาวัตถุ (H. Kim, S. Yang, K. Sohn, 2003) ได้เสนอให้ใช้วิธีการติดตั้งกล้องแบบหลายตัวแทนการติดตั้งกล้องเพียง 2 ตัว ส่วนการบูรณะสภาวะ 3 มิติด้วยวิธีการอื่น ๆ นอกจากวิธีการนี้เริ่มจาก (M. Pollefeys, L. Gool, 1997) ศึกษาการบูรณะสภาวะ 3 มิติด้วยการเปลี่ยนแนวการวางตัวของกล้องทั้ง 2 ตัวเพื่อให้เกิดมุมมองอื่น ๆ นั่นคือเป็นการเปลี่ยนมุมการวางกล้องกับวัตถุด้วยการขยับครั้งละ 5 องศารอบวัตถุเพื่อให้เห็นทุกด้านของวัตถุ ผลการวิจัยที่ได้มีความชัดเจนสามารถสร้างบริเวณที่ซับซ้อนได้ แต่ไม่เหมาะสมกับงานวิจัยที่ต้องการยึดตัวกล้องอยู่กับที่ ต่อมา (C. Strecha, L. Gool, 2002) สามารถบูรณะสภาวะ 3 มิติ โดยใช้แผนที่บ่งบอกความลึกเพียง 2 ชุดจากกล้อง 3 ตัวเพื่อลดข้อจำกัดเรื่องการขยับกล้อง ผลที่ได้นั้นสามารถหารายละเอียดที่ซับซ้อนได้มากขึ้น ส่วน (H. Kim, S. Yang, K. Sohn, 2003) ใช้กล้อง 2 ตัวบูรณะสภาวะ 3 มิติด้วยวิธีการแปลงเชิงเรขาคณิต ซึ่งสามารถบูรณะสภาวะ 3 มิติได้รายละเอียดเพียงแยกตำแหน่งของวัตถุออกจากพื้นหลัง แต่ไม่สามารถบ่งบอกรายละเอียดของวัตถุได้ ดังนั้นได้มีการแก้ไขสภาวะในการมองเห็นที่จำกัดนี้โดย (S. Kagami, Y. Takaoka, Y. Kida, K. Nishiwaki, and T. Kanade, 2005) ทำการวิจัยสร้างเส้นทางเดินให้กับหุ่นยนต์โดยการสร้างแนวทิศทางการเคลื่อนที่ของหุ่นยนต์เพื่อใช้เป็นแนวทางการเคลื่อนที่ด้วยการหาความสัมพันธ์จากภาพที่มองเห็นแบบลำดับต่อเนื่อง (image sequence) ด้วยการสร้างสภาวะ 3 มิติของแต่ละภาพมาเชื่อมต่อกัน ซึ่งผลการวิจัยที่ได้สามารถแก้ไขสภาวะการมองเห็นของหุ่นยนต์ที่มีจำกัดได้

2.9 สรุป

บทที่ 2 นี้ได้นำเสนอรายงานผลการสืบค้นวรรณกรรมย้อนหลังที่เกี่ยวข้องกับงานวิจัยที่จะดำเนินการจากฐานข้อมูล IEEE, IEE, ScienceDirect และอื่นๆ ซึ่งทำให้ทราบถึงแนวทางการวิจัยที่เกี่ยวข้อง ระเบียบวิธีการวิจัยที่ผู้วิจัยอื่นๆได้นำมาใช้ ผลการดำเนินงานข้อเสนอแนะต่าง ๆ ของนักวิจัยตั้งแต่อดีตจนถึงปัจจุบัน ซึ่งประกอบไปด้วยการหาแผนที่บ่งบอกความลึก และการหาระยะห่างระหว่างวัตถุกับกล้องโดยใช้กล้องสองตัว และพื้นฐานการวิเคราะห์ประมวลผลภาพ

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 3
การแก้ปัญหาความสอดคล้องกันของ
จุดภาพโดยใช้เครือข่ายประสาทฮอพฟิลด์ 2 มิติ

3.1 กล่าวนำ

ปัญหาความสอดคล้องกันของจุดภาพนั้น คือปัญหาสำคัญของการบูรณะแวดล้อม 3 มิติ จากงานวิจัยพบว่าเครือข่ายฮอพฟิลด์ 2 มิติ เป็นอีกวิธีการหนึ่งที่ใช้ในการแก้ปัญหานี้ โดยมีเครือข่ายฮอพฟิลด์เป็นต้นแบบ เครือข่ายฮอพฟิลด์ 2 มิติ มีชุดข้อมูลตำแหน่งที่น่าสนใจของกล้องซ้ายและขวาเป็นอินพุตให้กับระบบการแก้ปัญญาหาความสอดคล้องกันของจุดภาพหรือเรียกได้ว่าการจับคู่กันของจุดภาพ (matching) หลังจากนั้นเครือข่ายฮอพฟิลด์จะทำการวนรอบเพื่อหาคำตอบต่อไป ซึ่งในบทนี้ชุดข้อมูลที่นำมาวิจัย เป็นภาพนิ่งที่เกิดจากกล้องสองตัว เพื่อตรวจสอบความแม่นยำของระบบเครือข่ายฮอพฟิลด์ 2 มิติ

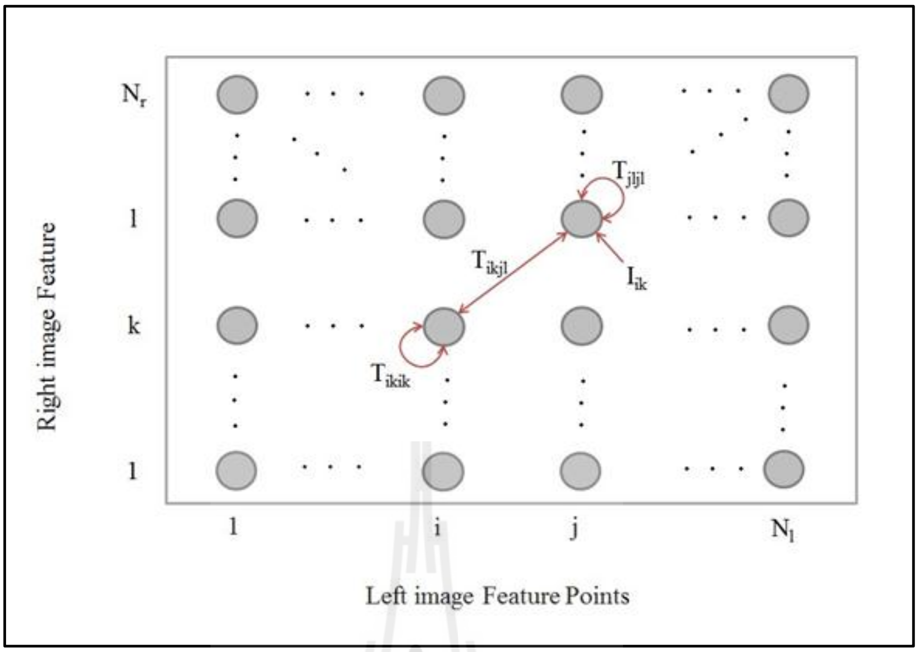
3.2 โครงสร้างของเครือข่ายฮอพฟิลด์ 2 มิติ

โครงสร้างของเครือข่ายฮอพฟิลด์ 2 มิติ ประกอบไปด้วย $N_r \times N_l$ นิวรอน โดยที่ N_r และ N_l คือจำนวนจุดที่สนใจของภาพทางด้านซ้ายและด้านขวาตามลำดับ สถานะของนิวรอน (state) ในเครือข่ายจะแสดงถึงความเป็นไปได้ของความสอดคล้องกันภายในจุดภาพของภาพทั้ง 2 กล้อง ภายในเครือข่ายนิวรอนจะเชื่อมโยงกับนิวรอนอื่น ๆ ในเครือข่ายทั้งหมดยกเว้นตัวมันเอง โดยมีค่านำหนักประสาทระหว่างนิวรอนคือ

$$T_{ikjl} = T_{jlik} \tag{3.1}$$

$$T_{ikik} = 0 \tag{3.2}$$

สมการที่ 3.1 แสดงถึงความสมมาตรของค่าน้ำหนักประสาทระหว่างนิวรอน ik และ jl ในขณะที่สมการที่ 3.2 แสดงว่าไม่มีการเชื่อมโยงค่าน้ำหนักประสาทของนิวรอนกับตัวมันเอง



รูปที่ 3.1 โครงสร้างเครือข่ายประสาทและค่าน้ำหนักประสาทของเครือข่ายฮอพฟิลด์

3.3 การเรียนรู้ของเครือข่ายฮอพฟิลด์ 2 มิติ

การเปลี่ยนแปลงสถานะของเครือข่ายฮอพฟิลด์ 2 มิติจะขึ้นอยู่กับค่าพลังงานของนิวรอนในเครือข่าย สถานะของเครือข่ายมีความสัมพันธ์กับพลังงานภายในเครือข่ายดังต่อไปนี้

$$E = -\left(\frac{1}{2}\right)\sum_{i=1}^{N_l}\sum_{k=1}^{N_r}\sum_{j=1}^{N_l}\sum_{l=1}^{N_r} T_{ikjl}V_{ik}V_{jl} - \sum_{i=1}^{N_l}\sum_{k=1}^{N_r} I_{ik}V_{ik} \qquad (3.3)$$

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} คือค่าอินพุตเริ่มต้นจะได้ว่า เมื่อนิวรอนที่ ik เปลี่ยนสถานะจะทำให้มีการเปลี่ยนแปลงค่าพลังงานด้วย ดังสมการ

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

การอัพเดทสถานะของนิวรอนคำนวณได้จาก

$$V_{ik} = 0\ if \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl} + I_{ik}\right] > 0 \tag{3.5}$$

$$V_{ik} = 1\ if \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl} + I_{ik}\right] < 0 \tag{3.6}$$

$$no\ chang\ if \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl} + I_{ik}\right] = 0 \tag{3.7}$$

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

3.4 การแก้ปัญหาความสอดคล้องกันของจุดภาพ

การแก้ปัญหาความสอดคล้องกันของจุดภาพจะใช้ค่าพลังงานของเครือข่ายประสาทเทียมดังสมการที่ 3.3 เมื่อค่าพลังงานต่ำสุดหรือค่าพลังงานคงที่แสดงถึงความเป็นไปได้ของความสอดคล้องกันของจุดภาพจากทั้ง 2 กล้อง ซึ่งมีส่วนประกอบของฟังก์ชั่นดังนี้

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

โดยเทอมแรกจะเป็นตัวบ่งชี้ความสอดคล้องกันของจุดภาพทั้งสองเทอมที่สองและสามคือเงื่อนไข (uniqueness constrain) ของแต่ละนิวรอน สามารถเขียนใหม่ได้เป็น

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

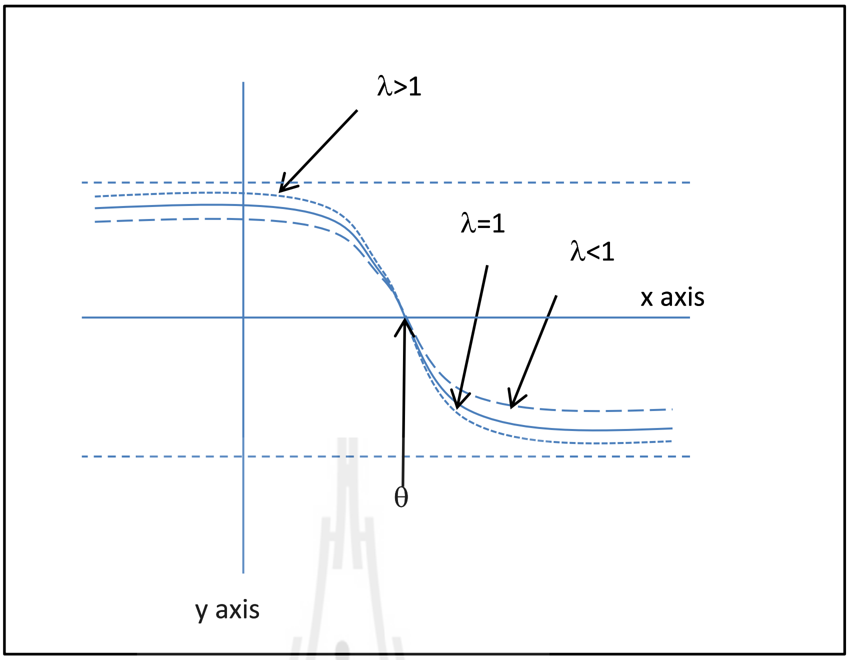
เมื่อนำสมการที่ 3.8 เปรียบเทียบกับสมการค่าพลังงาน จะได้ว่าสถานะของนิวรอน $V_{ik}=P_{ik}$ $V_{jl}=P_{jl}$ และ $I_{ik}=2$ และค่าน้ำหนักประสาท $T_{ikjl}=C_{ikjl}-\delta_{ij}-\delta_{lk}$ โดยที่ $\delta_{ij} = 1$ เมื่อ $i=j$ และ $\delta_{kl} =1$ เมื่อ $k=l$

ค่าน้ำหนักประสาท T_{ikjl} ได้มาจาก

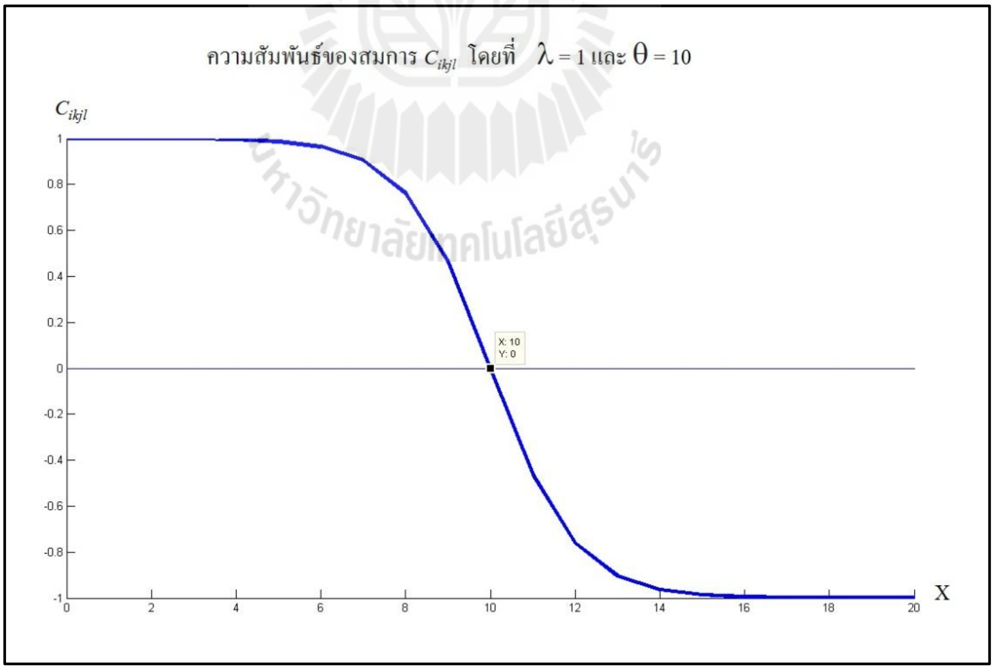
$$C_{ikjl} = (2/[1+e^{\lambda(X-\theta)}])-1 \quad (3.10)$$

$$X = [W_1|\Delta d|+W_2|\Delta D|] \quad (3.11)$$

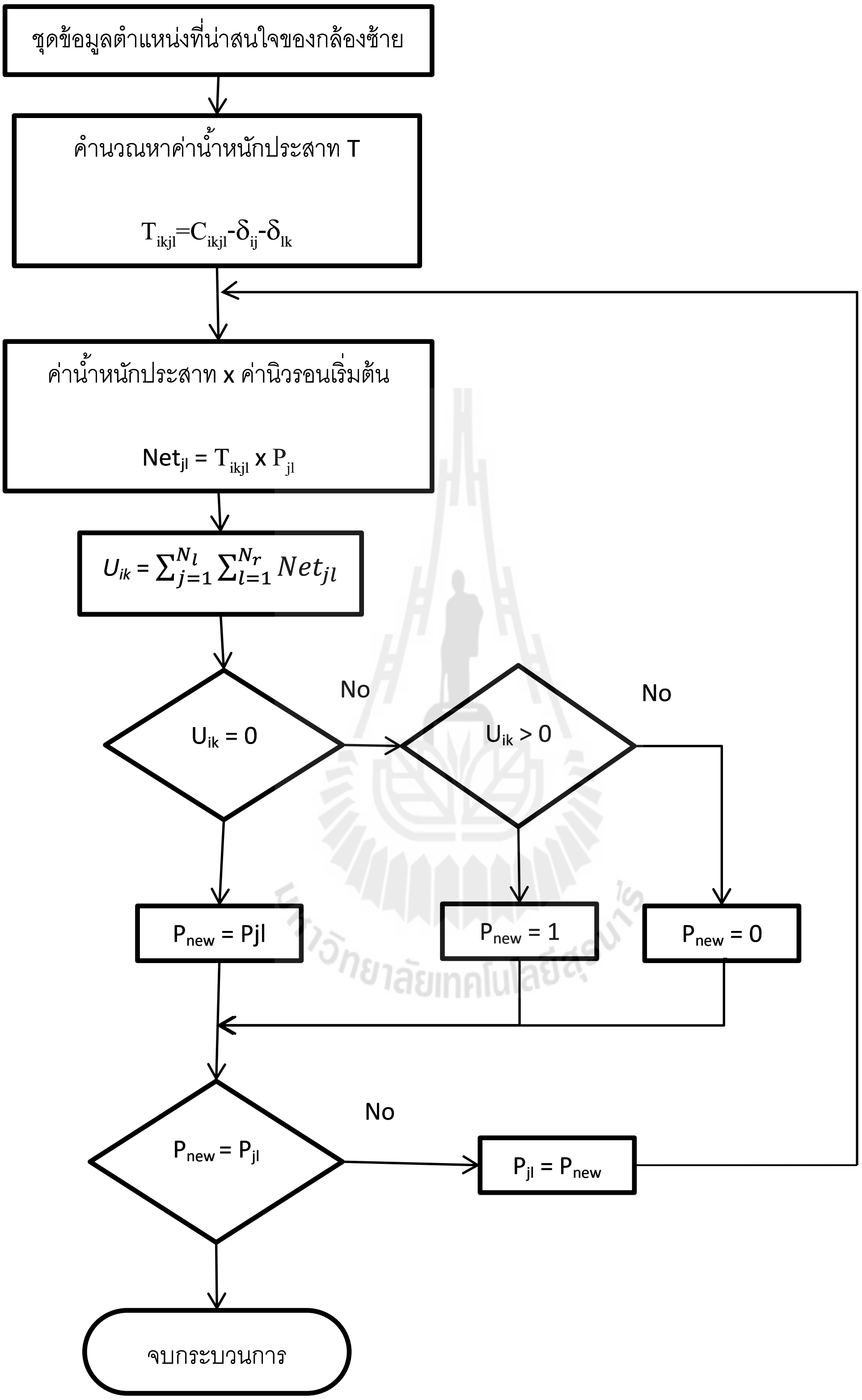
ซึ่งสมการที่ 3.10 เป็นฟังก์ชั่นไม่เป็นเชิงเส้น (non-linear function) เพื่อให้ค่าอยู่ในช่วง +1 และ -1 λ คือค่าความชันของฟังก์ชั่น θ คือจุดตัดแกน x ของฟังก์ชั่น จากผลการวิจัย (N. Nasrabadi, C. Choo, 1992) พบว่าค่าที่ดีที่สุดคือ $\lambda = 1$ และ $\theta = 10$ ค่า X หาได้จากสมการที่ (3.11) ซึ่ง Δd คือค่าระยะห่างระหว่างจุด (i,k) และ (j,l) ΔD คือค่าผลต่างระหว่างระยะห่างระหว่างจุด i และ j กับระยะห่างระหว่างจุด k และ l และค่าตัวแปร W_1 และ W_2 คือค่าคงที่โดย W1 + W_2=1 ซึ่งความสัมพันธ์ที่กล่าวมานั้นแสดงดังกราฟที่ 3.2 และ 3.3 ซึ่งสังเกตได้ว่า ค่า C_{ikjl} จะเริ่มลดลงเรื่อยๆ เมื่อค่า X มากขึ้น เมื่อพิจารณาค่า X จากสมการ 3.11 คือค่าความต่างกันของตำแหน่งที่พิจารณา ดังนั้นจะได้ว่า เมื่อค่า X = 0 คือค่าความต่างที่มีค่าน้อยของตำแหน่งที่สนใจจากทั้งสองภาพ แต่ยังสรุปไม่ได้ว่าเป็นจุดเดียวกัน แต่แสดงให้เห็นถึงความเป็นไปได้ว่าตำแหน่งนั้นคือตำแหน่งเดียวกัน แต่ถ้าค่า X = 10 หรือ X > 10 มากๆแสดงได้ว่าสองตำแหน่งนั้นมีความเป็นไปได้น้อยที่จะเป็นตำแหน่งเดียวกันเพราะกำหนดให้ค่าติดลบ



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



รูปที่ 3.3 C_{ikjl} function โดยกำหนดให้ λ = 1 และ θ = 10



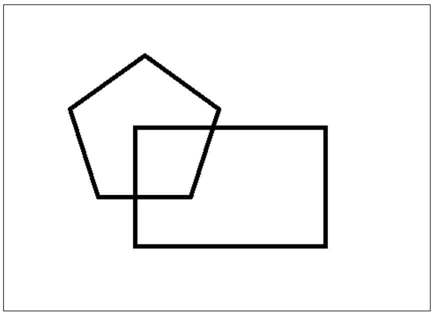
รูปที่ 3.4 ระบบการทำงานของเครือข่ายฮอพฟิลด์

เมื่อได้ค่าพารามิเตอร์ต่าง ๆ นำชุดข้อมูลที่ได้ป้อนเข้าสู่ระบบเครือข่ายฮอพฟิลด์เพื่อทำการวนรอบเพื่อหาความสอดคล้องของจุดภาพดังรูปที่ 3.3 เมื่อป้อนอินพุตเข้าสู่ระบบ เครือข่ายฮอพฟิลด์ จะทำการคำนวณหาเมตริกค่าน้ำหนักประสาทที่ได้จากสมการที่ 3.10 และสมการที่ 3.11 โดยมีเมตริกเท่ากับ $N_r \times N_l$ นิวรอน โดยที่ N_r และ N_l คือจำนวนจุดเด่นของภาพทางด้านซ้ายและด้านขวาตามลำดับ เมื่อได้ค่าเมตริกน้ำหนักประสาทแล้ว จึงนำไปคูณกับค่าเมตริกนิวรอนเริ่มต้น โดยค่านิวรอนเริ่มต้นเท่ากับ 0 แล้วหาผลรวมเพื่อหาสถานะของแต่ละนิวรอนในแต่ละรอบ แล้วตรวจสอบสถานะนิวรอนใหม่กับสถานะนิวรอนเริ่มต้น ว่ามีการเปลี่ยนแปลงหรือไม่ ถ้ามีการเปลี่ยนแปลง ให้เมตริกนิวรอนเริ่มต้นเท่ากับเมตริกนิวรอนใหม่ แล้วน้ำไปคำนวณกับค่าน้ำหนักประสาทจนกว่าเมตริกสถานะนิวรอนใหม่เท่ากับนิวรอนเดิมจึงสิ้นสุดกระบวนการ สถานะนิวรอนใหม่ที่สิ้นสุดกระบวนการแล้วนั้น ถ้าเท่ากับ 1 ในเมตริกตำแหน่งใดแสดงว่าตำแหน่งจากกล้องซ้ายและกล้องขวาในเมตริกตำแหน่งนั้นสอดคล้องกัน

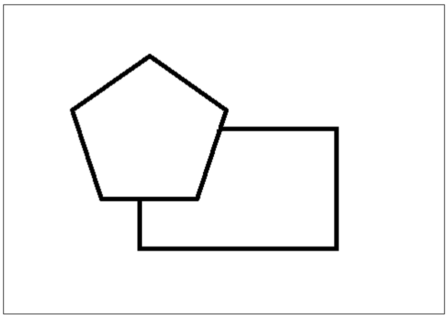
3.5 ผลการทดสอบการแก้ปัญหาความสอดคล้องกันของจุดภาพโดยใช้เครือข่ายประสาทฮอพฟิลด์ 2 มิติ

ตัวอย่างข้อมูลชุดแรกมาจากการจำลองต้นแบบเพื่อตรวจสอบการทำงานของระบบที่ใช้ในการแก้ปัญหาความสอดคล้องนี้ โดยใช้การตรวจจับแบบแฮริส (Harris's detector) แล้วป้อนชุดข้อมูลที่ได้เข้าเครือข่ายฮอพฟิลด์

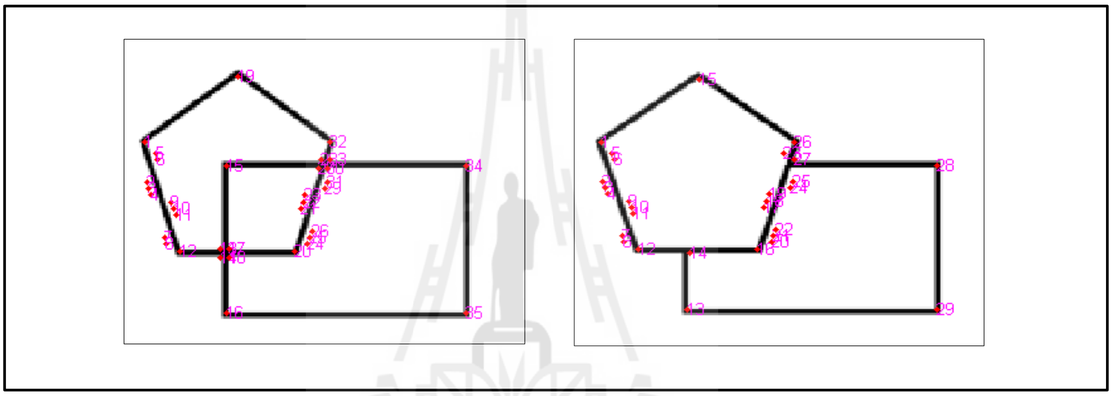
จากรูป 3.5(ก) และ (ข) ภาพต้นแบบซ้ายและขวามีจุดแตกต่างกัน เมื่อทำการหาตำแหน่งที่น่าสนใจด้วยวิธีแฮริสแล้วจะได้ ดังรูป 3.5(ค) ซึ่งภาพทางด้านซ้ายจะมีตำแหน่งที่น่าสนใจมากกว่าภาพทางด้านขวา เมื่อผ่านกระบวนการของเครือข่ายฮอพฟิลด์แล้วจะได้ดังรูปที่ 3.5 (ง) ตำแหน่งที่กำกับด้วยตัวเลขแสดงถึงจุดที่สอดคล้องกัน ซึ่งเห็นได้ว่า ตำแหน่งที่ตัวเลขกำกับไว้คือตำแหน่งเดียวกัน จึงทำการทดลองกับรูปจริงที่ถ่ายด้วยกล้อง 2 ตัว จากรูปที่ 3.6 (ง) จะเห็นได้ว่าตำแหน่งที่กำกับด้วยตัวเลข 1 ถึง 10 เป็นตำแหน่งที่สอดคล้องกันของภาพจากกล้องซ้ายและกล้องขวา จากการผลการทดลองตำแหน่งของวัตถุที่ต้องการคือกล่องที่อยู่ในภาพ แต่วิธีการคัดกรองของแฮริสตำแหน่งที่น่าสนใจได้กระจายไปทั่วภาพ ซึ่งเป็นตำแหน่งที่ไม่ต้องการทั้งยังเป็นภาระของระบบเครือข่ายฮอพฟิลด์ 2 มิติ ยิ่งอินพุตมากการทำงานของระบบก็ช้าลงด้วย ดังนั้นการใช้วิธีการนี้ในการคัดกรองวัตถุในภาพเคลื่อนไหวจึงไม่เหมาะสมนัก



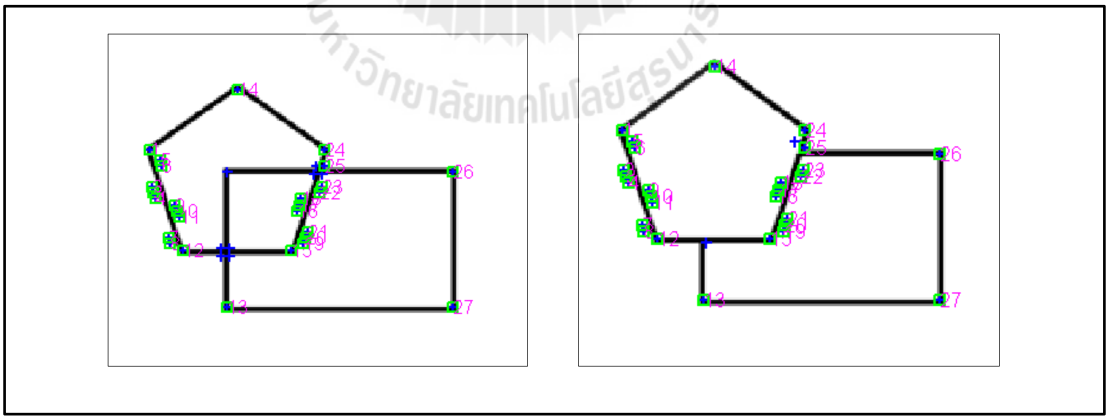
(ก)



(ข)



(ค)



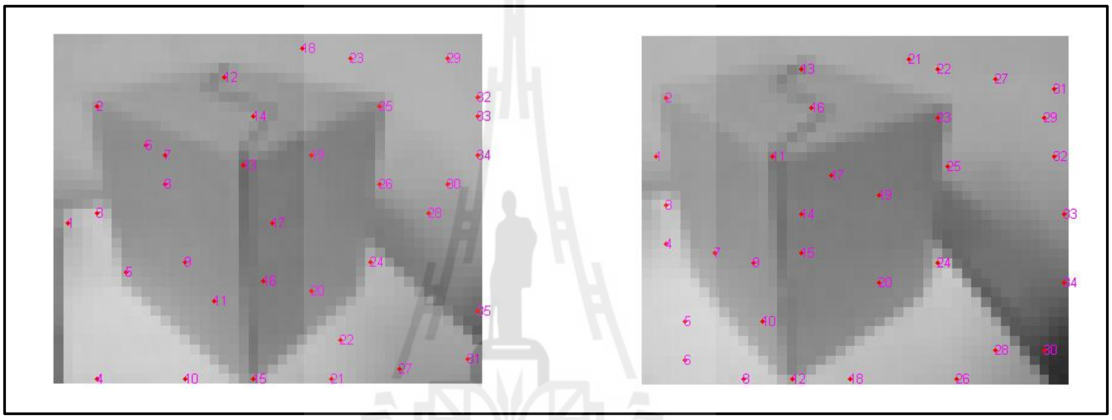
(ง)

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

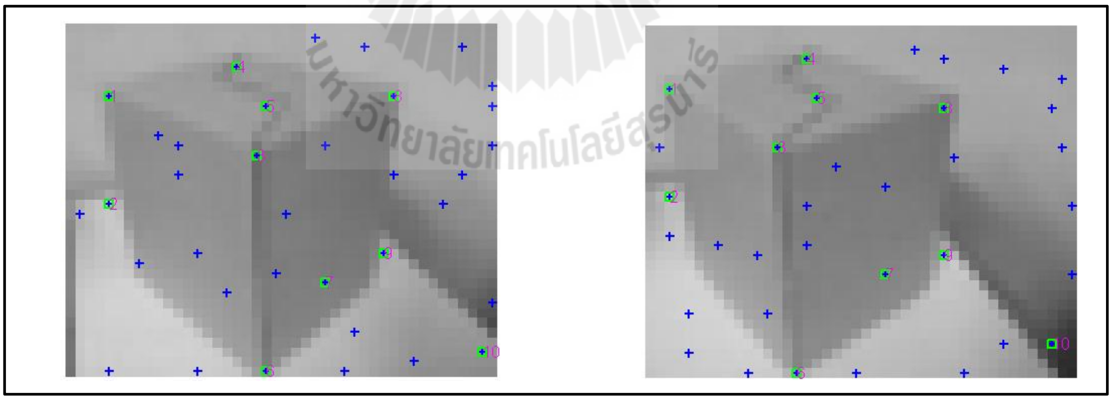
มหาวิทยาลัยเทคโนโลยีสุรนารี

(ก)

(ข)



(ค)



(ง)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 3.6 (ก) ภาพจากกล้องซ้าย (ข) ภาพจากกล้องขวา
(ค) ตำแหน่งที่เกิดจากการตรวจจับของแฮริส
(ง) ผลการแก้ปัญหาด้วยเครือข่ายฮอพฟิลด์ 2 มิติ

3.6 สรุป

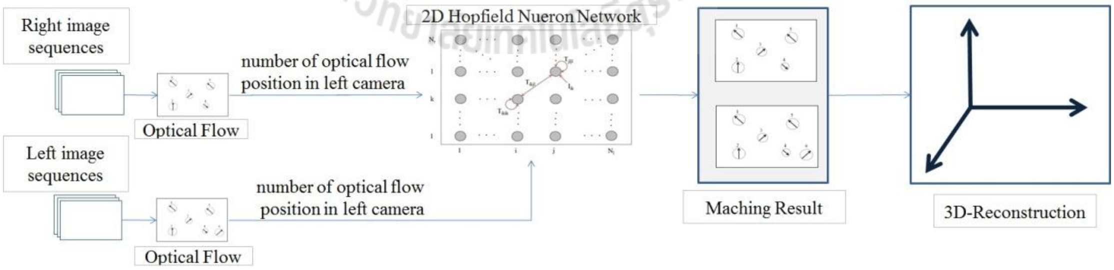
ระบบเครือข่ายฮอพฟิลด์สามารถแก้ปัญหาความสอดคล้องของจุดภาพจากกล้อง 2 ตัวได้ ซึ่งมีอินพุตที่เกิดจากตำแหน่งที่น่าสนใจเป็นตัวเริ่มกระบวรการ หลังจากนั้นเครือข่ายฮอพฟิลด์จะทำการวนรอบเพื่อหาคำตอบ การหาคำตอบของเครือข่ายประสาทเทียมนั้นต้องอาศัยอินพุตที่เหมาะสมกับวัตถุประสงค์งานนั้น ๆ ซึ่งในบทนี้ได้ใช้การหาขอบของแฮริสเพื่อหาจุดที่น่าสนใจของภาพทั้งสองซึ่งเหมาะกับงานที่ไม่มีความซับซ้อน โดยภาพที่ใช้มีมุมหรือขอบชัดเจนทำให้ง่ายต่อการตัดสินว่าจุดใดเป็นจุดเดียวกัน แต่เมื่อนำมาใช้ในรูปที่มีความซับซ้อนมากขึ้นการหาขอบของแฮริสไม่สามารถใช้ได้ จึงต้องใช้วิธีการอื่นการหาจุดที่น่าสนใจในบทต่อไป

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 4
การบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธี แผนที่บ่งบอกความลึกเชิงเครือข่ายประสาทเทียมและการไหลเชิงแสง

4.1 กล่าวนำ

ระบบการทำงานของการบูรณะแวดล้อม 3 มิติในงานวิจัยนี้ ได้ใช้การไหลเชิงแสงเป็นตัวคัดแยกวัตถุที่เคลื่อนที่ออกจากพื้นหลัง โดยวัตถุใด ๆ ที่มีการเคลื่อนที่ภายในภาพ เมื่อผ่านกระบวนการการไหลเชิงแสงแล้ว บริเวณที่วัตถุเคลื่อนที่ในภาพ จะเกิดการไหลของแสงขึ้นทำให้ได้ข้อมูลทั้งขนาดและมุมพร้อมกับตำแหน่งที่เกิดการไหลเชิงแสง ซึ่งข้อมูลเหล่านี้เป็นอินพุตให้กับระบบเครือข่ายฮอพฟิลด์ 2 มิติ ดังที่กล่าวไปแล้วในบทที่ 3 โครงสร้างของระบบแสดงดังรูป ที่ 4.1 โดยขั้นตอนของระบบรวมนี้จะเริ่มต้นด้วยการนำลำดับภาพจากกล้อง 2 ตัว เข้าสู่การเตรียมการประมวลผลด้วยการปรับขนาด การแปลงภาพจากภาพสีให้เป็นภาพระดับเทา จากนั้นนำภาพที่ผ่านการประมวลผลแล้วเข้ารับการประมวลผลการไหลเชิงแสง จะได้อินพุตป้อนให้กับการประมวลผลเครือข่ายฮอพฟิลด์ 2 มิติ เพื่อแก้ปัญหาความสอดคล้องกันของจุดภาพ แล้วจึงนำผลของการประมวลมาคำนวณหาระยะห่างระหว่างวัตถุกับกล้องและการบูรณะสภาวะแวดล้อม 3 มิติด้วยวิธีการทางคณิตศาสตร์ดังที่กล่าวไว้ในบทที่ 2



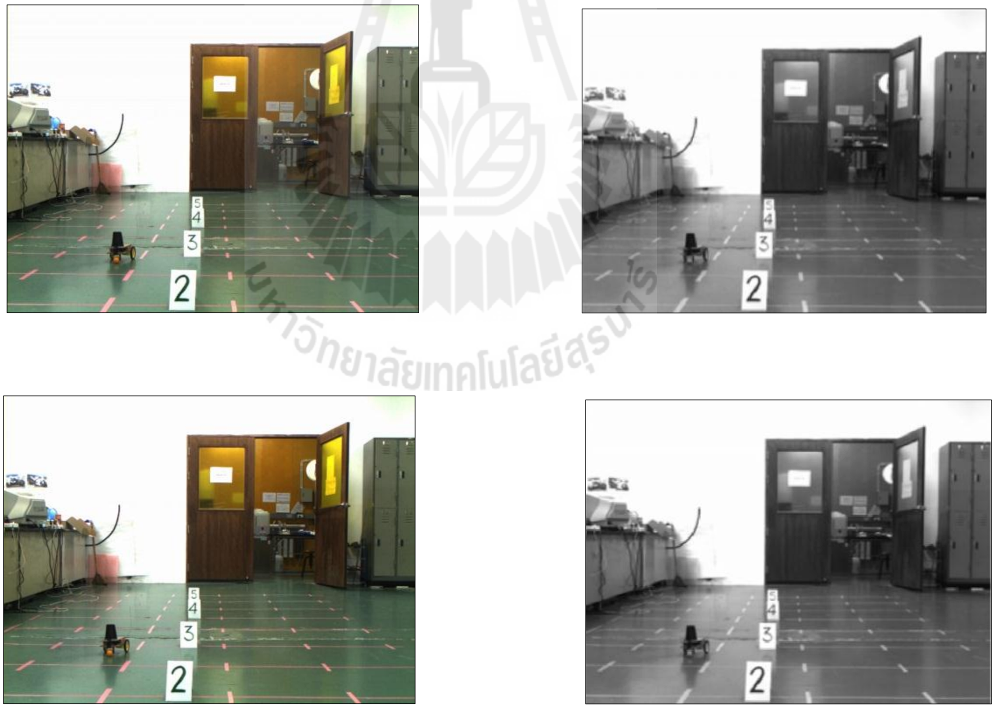
รูปที่ 4.1 โครงสร้างระบบการบูรณะสภาวะแวดล้อม 3 มิติ

4.2 กระบวนการคัดแยกวัตถุที่เคลื่อนที่

กระบวนการคัดแยกวัตถุสำหรับงานวิจัยนี้ใช้การไหลเชิงแสงในการคัดแยกและระบุตำแหน่งของวัตถุในภาพ เมื่อผ่านกระบวนการนี้แล้วจะได้ทั้งตำแหน่ง ขนาดและมุมของการไหลเชิงแสงซึ่งเป็นข้อมูลที่ถูกป้อนให้กระบวนการของเครือข่ายประสาทเทียมต่อไป การหาการไหลเชิงแสงของภาพจำเป็นจะต้องใช้ภาพที่มีเฟรมต่อเนื่องกัน 2 ภาพ โดยมีขั้นตอนดังนี้

ขั้นตอนที่ 1 การปรับขนาดและระดับสีของภาพต้นแบบ

การปรับระดับความเหมาะสมของขนาดภาพถือได้ว่าเป็นขั้นตอนที่สำคัญอีกขั้นตอนหนึ่งของกระบวนการ เมื่อทำการปรับขนาดภาพนอกจากข้อมูลบางส่วนจะหายไปตำแหน่งของวัตถุที่เกิดในแต่ละจุดภาพยังมีผลในการคำนวณในขั้นตอนการหาระยะห่างระหว่างกล้องอีกด้วย ในงานวิจัยนี้จะทำการปรับขนาดภาพให้มีขนาดเท่ากับ 640x480 จุดภาพ ซึ่งเป็นขนาดที่มีข้อมูลเพียงพอในการประมวลผลของระบบ จากนั้นทำการแปลงภาพจากภาพสี RGB ให้เป็นภาพระดับเทา ดังรูปที่ 4.2



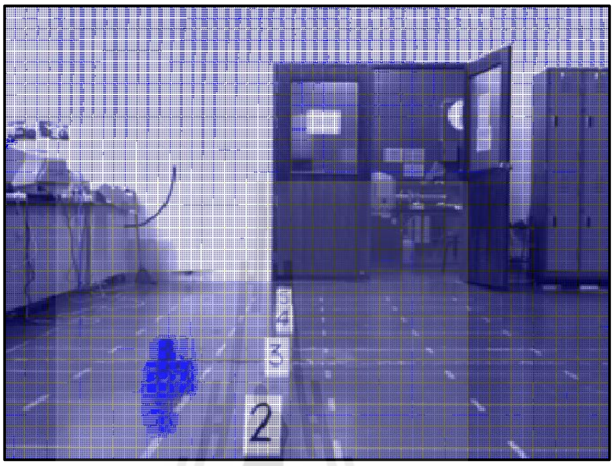
มหาวิทยาลัยเทคโนโลยีสุรนารี

(ก)ภาพต้นแบบ (ข) ภาพที่ผ่านการปรับขนาดและสีแล้ว

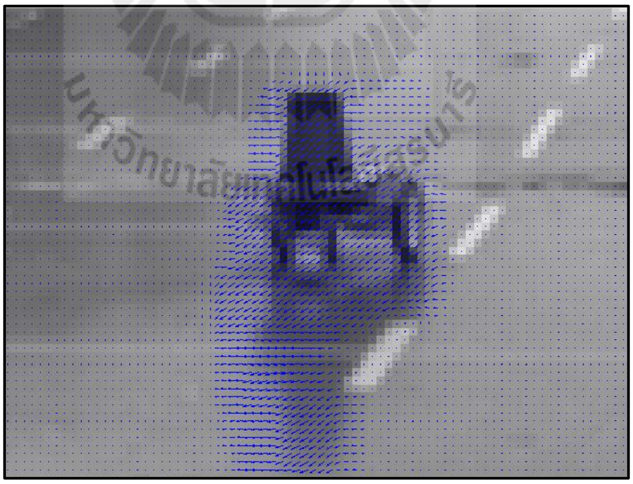
รูปที่ 4.2 การปรับขนาดและระดับสีของภาพต้นแบบ

ขั้นตอนที่ 2 การหาตำแหน่งการไหลเชิงแสงของวัตถุในภาพ

การไหลเชิงแสงสามารถตรวจจับวัตถุที่เคลื่อนที่ภายในภาพได้อย่างแม่นยำ โดยงานวิจัยนี้ได้ใช้การไหลเชิงแสงด้วยวิธีของลูคัสและคานาเด (Lucas & Kanade) โดยจะประกอบไปด้วยขนาดและทิศทาง แสดงได้ดังรูป 4.3



(ก) การไหลเชิงแสงของภาพที่ได้จากกล้องซ้ายที่ยังไม่ได้กรองขนาด



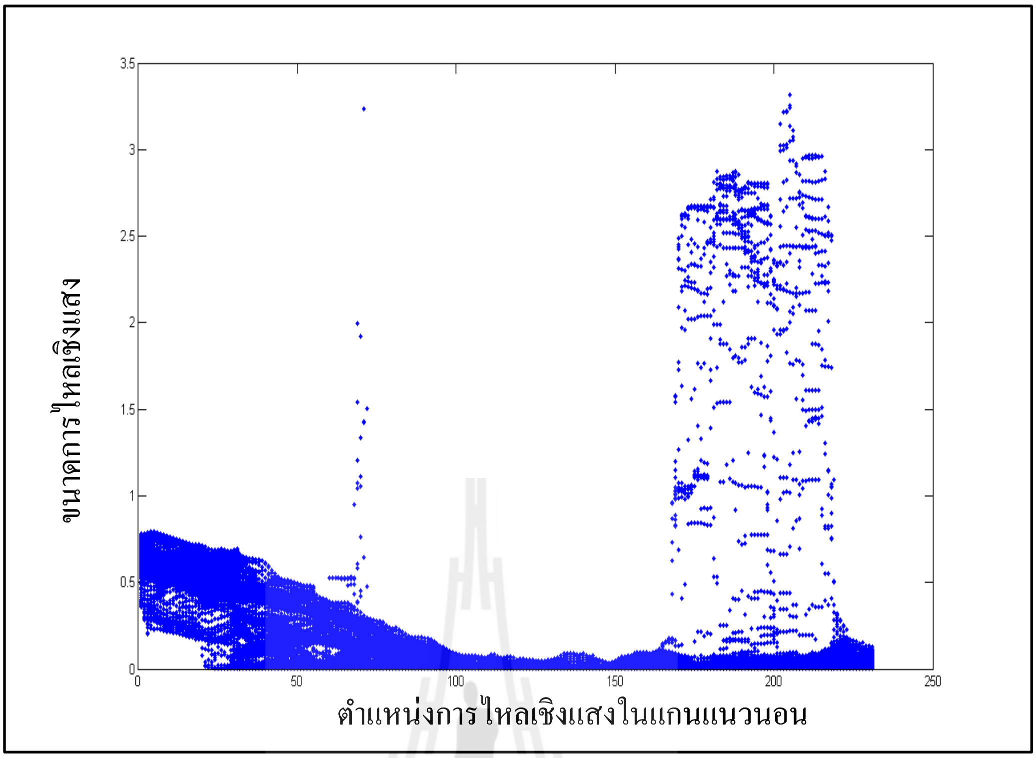
(ข) ภาพขยายตำแหน่งการไหลเชิงแสงของภาพ

รูปที่ 4.3 การไหลเชิงแสง

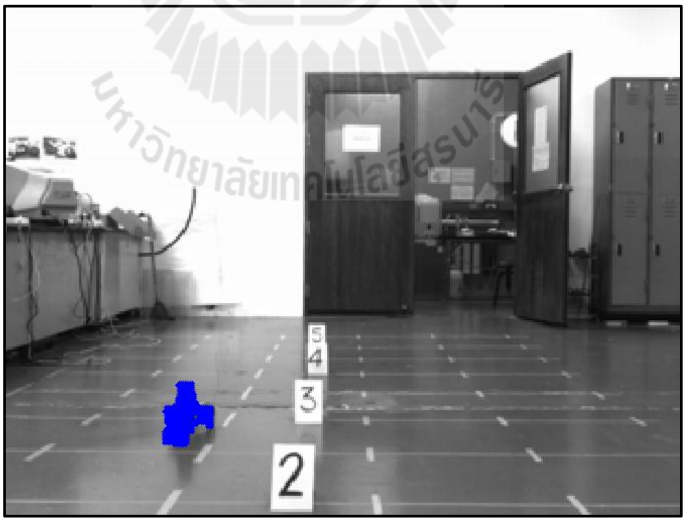
จากรูปที่ 4.3 จะเห็นได้ว่าการไหลเชิงแสงถึงแม้จะคัดแยกวัตถุที่เคลื่อนที่ได้ดี แต่ก็ยังมีตำแหน่งที่ไม่พิจารณาเกิดขึ้นด้วยดังแสดงไว้ที่รูปที่ 4.4 ซึ่งเป็นบริเวณที่ไม่พิจารณา จึงต้องทำการกรองโดยใช้ขนาดของการไหลเชิงแสง

รูปที่ 4.4 ภาพขยายตำแหน่งการไหลเชิงแสงของภาพ

จากรูปที่ 4.5 จะเห็นได้ว่าตำแหน่งที่เกิดการไหลเชิงแสงมีทั่วทั้งภาพ แต่ตำแหน่งที่วัตถุเคลื่อนที่นั้นจะมีขนาดการไหลเชิงแสงมากกว่าพื้นหลัง ซึ่งจะอยู่ในช่วง 2 - 3.5 เมื่อทำการกรองเรียบร้อยแล้วจะได้ผลการทดสอบดังรูปที่ 4.6 จะเห็นได้ว่าตำแหน่งของการไหลเชิงแสงจะเป็นตำแหน่งที่วัตถุเคลื่อนที่

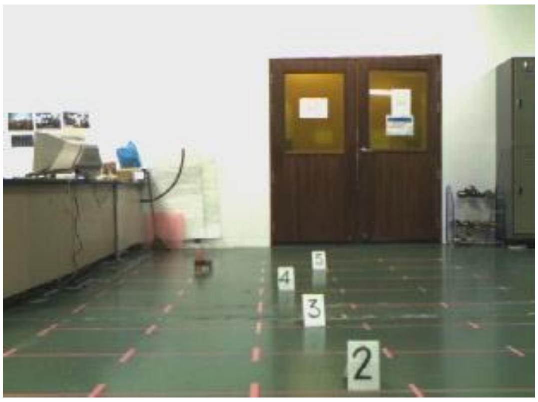


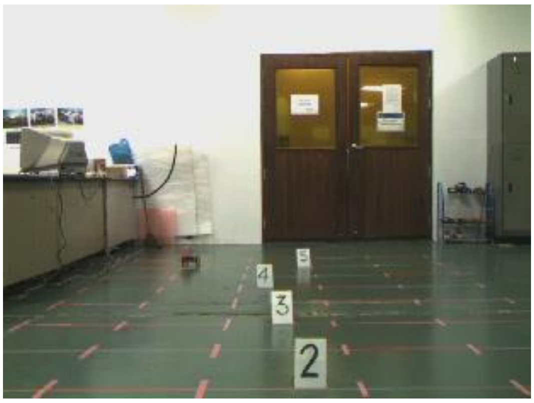
รูปที่ 4.5 ความถี่ของตำแหน่งการเกิดการไหลเชิงแสงของภาพต้นแบบ 4.2 (ก)



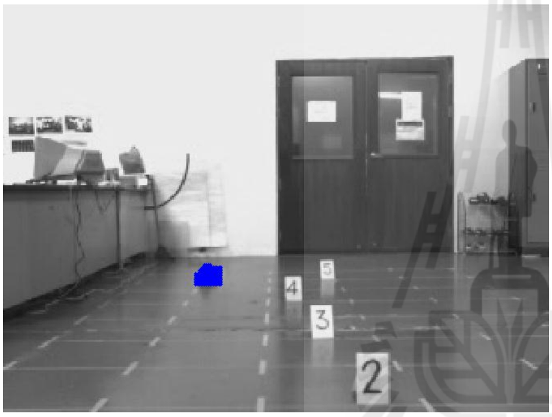
รูปที่ 4.6 ตำแหน่งการเกิดการไหลเชิงแสงที่ทำการกรองแล้ว

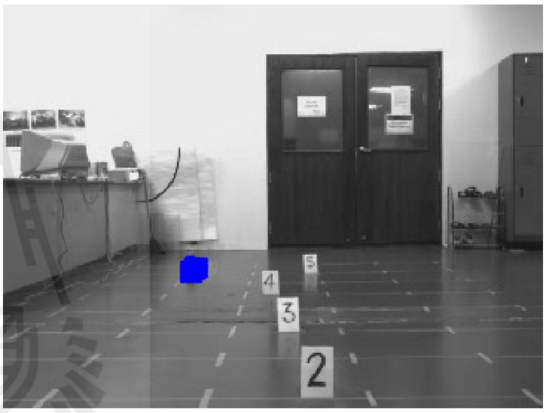
มหาวิทยาลัยเทคโนโลยีสุรนารี





(ก) (ข)





(ค) (ง)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 4.7 ตัวอย่างการหาตำแหน่งการเกิดการไหลเชิงแสง

(ก)ภาพต้นแบบจากล้องซ้าย

(ข)ภาพต้นแบบจากกล้องขวา

(ค) ภาพตำแหน่งการเกิดการไหลเชิงแสงจากกล้องซ้าย

(ง) ภาพตำแหน่งการเกิดการไหลเชิงแสง

4.3 การแก้ปัญหาความสอดคล้องกันของจุดภาพ

เมื่อผ่านการคัดแยกวัตถุที่เราสนใจด้วยการไหลเชิงแสงแล้ว จะได้ชุดข้อมูลออกมาเป็นตำแหน่งการเกิดการไหลเชิงแสงในแนวนอนและแนวตั้ง (x,y) ขนาดการไหลเชิงแสงและมุมของการไหลเชิงแสง ชุดข้อมูลเหล่านี้จะถูกนำไปประมวลผลในเครือข่ายฮอพฟิลด์ 2 มิติ ดังขั้นตอนที่

กล่าวไว้ในบทที่ 3 โดยเริ่มจาก คำนวณหาค่าน้ำหนักประสาท $T_{ikjl} = C_{ikjl}-\delta_{ij}-\delta_{lk}$ โดยการคำนวณหาค่า C_{ikjl} จะใช้สมการไม่เชิงเส้นมาช่วยพิจารณาเงื่อนไขดังต่อไปนี้

1. ตำแหน่งของจุดภาพจากกล้องซ้ายและขวาจะสัมพันธ์กันก็ต่อเมื่อมีมุมของการไหลเชิงแสงเท่ากัน

2. ตำแหน่งของจุดภาพจากล้องซ้ายและขวาจะสัมพันธ์กันก็ต่อเมื่อค่าผลต่างของขนาดจากจุดที่พิจารณากับตำแหน่งของจุดภาพอื่นๆภายในภาพเท่ากัน

จากเงื่อนไขทั้ง 2 ข้อสามารถเขียนเป็นสมการได้ดังสมการที่ 4.1

$$X = |\Delta A|+|\Delta L| \tag{4.1}$$

ซึ่ง X คือ ค่าที่ต้องการพิจารณา

ΔA คือ ค่าผลต่างของมุมการไหลเชิงแสง

ΔL คือ ค่าผลต่างของขนาดจุดที่พิจารณาโดยแยกพิจารณาเป็น 2 กรณี

จากความสัมพันธ์ของสมการที่ 4.1 เมื่อแยกพิจารณา ΔLจะได้ดังสมการที่ 4.2

$$X = |\Delta D|+ |\Delta d|+ |\Delta A| \tag{4.2}$$

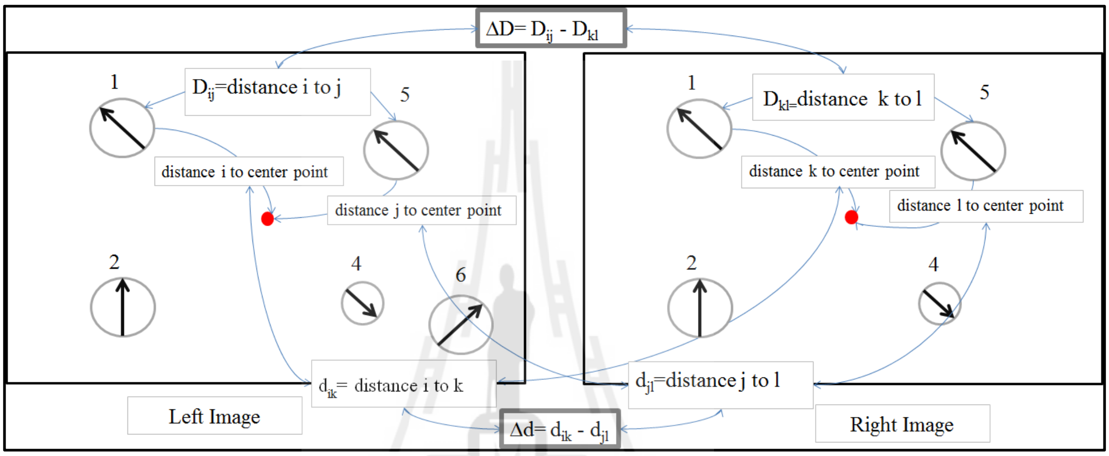
ซึ่ง ΔD คือ ค่าผลต่างของผลต่างของระยะจากตำแหน่งที่ i ถึงจุดอ้างอิง กับ ระยะจากตำแหน่งที่ k ถึงจุดอ้างอิง และ ผลต่างของระยะจากตำแหน่งที่ j ถึงจุดอ้างอิง กับ ระยะจากตำแหน่งที่ l ถึงจุดอ้างอิง

Δd คือ ค่าผลต่างระหว่างระยะจากตำแหน่งที่ i ถึงตำแหน่งที่ j กับระยะจากตำแหน่งที่ k ถึงตำแหน่งที่ l

ΔA คือ ค่าผลต่างของมุมจาการไหลเชิงแสงที่ตำแหน่ง i และ k

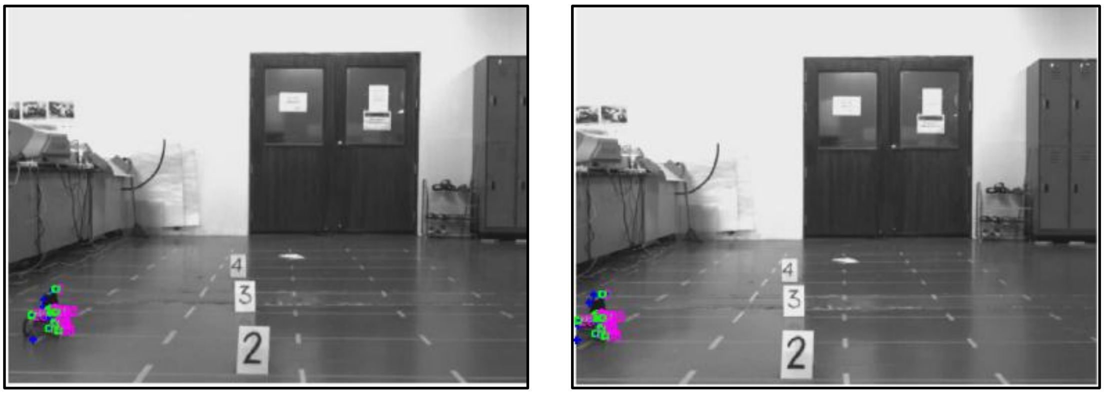
เมื่อกำหนดให้ภาพจากกล้องซ้ายคือภาพอ้างอิงจะพิจารณาค่าต่างๆได้ดังนี้ โดย ΔD เป็นตัวเปรียบเทียบตำแหน่งของตัวมันเองกับตำแหน่งรอบข้าง ก่อนที่จะนำผลที่ได้ไปเปรียบเทียบกับภาพจากกล้องขวา กล่าวคือ เมื่อพิจารณาตำแหน่งที่เกิดการไหลเชิงแสงในภาพซ้าย ระบบจะทำการเปรียบเทียบขนาดที่ตำแหน่งตัวมันเองกับตำแหน่งของสมาชิกในภาพซ้ายก่อนที่จะไปเปรียบเทียบตำแหน่งทุกตำแหน่งของสมาชิกที่เกิดการไหลเชิงแสงทางด้านขวา เมื่อพิจารณาแล้วจะเห็นได้ว่าถ้าตำแหน่งของภาพทางด้านซ้ายและขวาสัมพันธ์กันหรือเป็นจุดเดียวกันค่าที่ได้จะเท่ากับ 0

เช่นเดียวกันกับค่า Δd แต่ใช้ตำแหน่งกึ่งกลางภาพเป็นจุดอ้างอิง โดยพิจารณาจากการวัดระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพซ้ายเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพขวา แล้วจึงนำไปเปรียบเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพซ้ายเทียบกับระยะจุดกึ่งกลางถึงตำแหน่งที่พิจารณาที่ภาพขวาในจุดต่อไปทุกๆตำแหน่ง ในกรณีที่เป็นตำแหน่งเดียวกันควรมีค่าน้อยๆ หรือเข้าใกล้ 0 และ ΔA คือการเปรียบเทียบมุมระหว่างตำแหน่งที่พิจารณาที่ภาพซ้ายและขวา ถ้าเป็นมุมเดียวกันค่าที่ได้ต้องมีค่าน้อยๆ หรือเท่ากับ 0

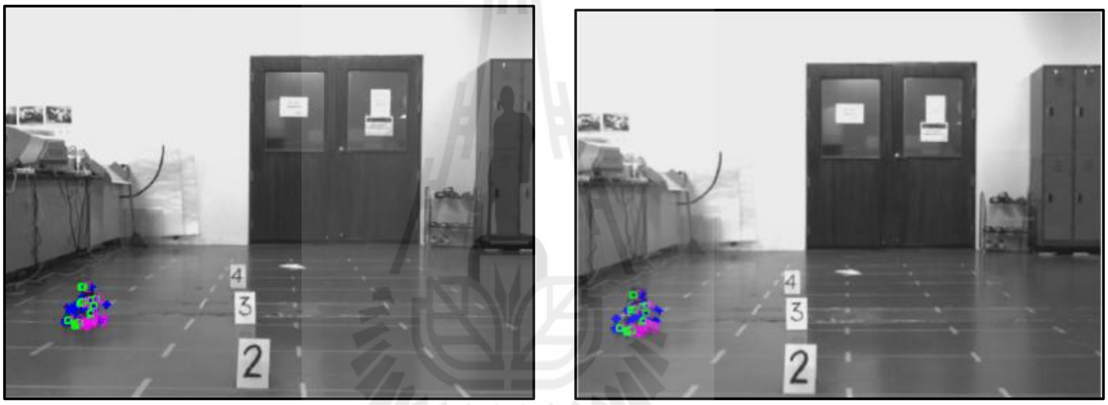


รูปที่ 4.8 การพิจารณาค่า ΔD และ Δd

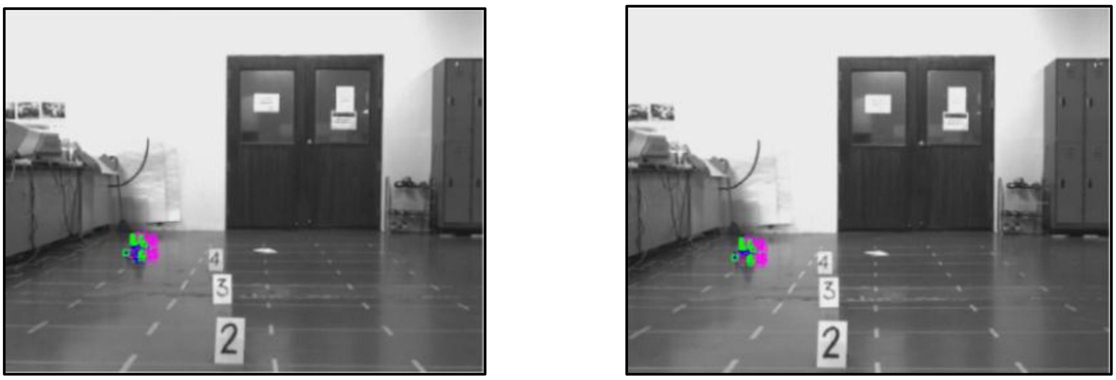
หลังจากนำชุดข้อมูลที่ได้ผ่านกระบวนการประมวลผลเครือข่ายฮอพฟิลด์ 2 มิติแล้ว ได้ผลการวิจัยดังรูปที่ 4.9 – 4.15 โดยชุดตัวอย่างที่ทำการทดสอบเป็นชุดตัวอย่างที่ถูกติดตั้งภายในอาคารซึ่งมีการเคลื่อนที่ของวัตถุอย่างอิสระและกล้องถูกติดตั้งอยู่กับที่ จากรูปที่ 4.9-4.15 จะเห็นได้ว่าวัตถุเคลื่อนที่เป็นแนวเส้นตรง เมื่อผ่านขั้นตอนการคัดแยกวัตถุโดยการไหลเชิงแสงแล้วจะสังเกตได้ว่าตำแหน่งการเกิดการไหลเชิงแสงหรือตำแหน่งที่ถูกจับคู่จะเคลื่อนที่ไปตามการเคลื่อนที่ของวัตถุ



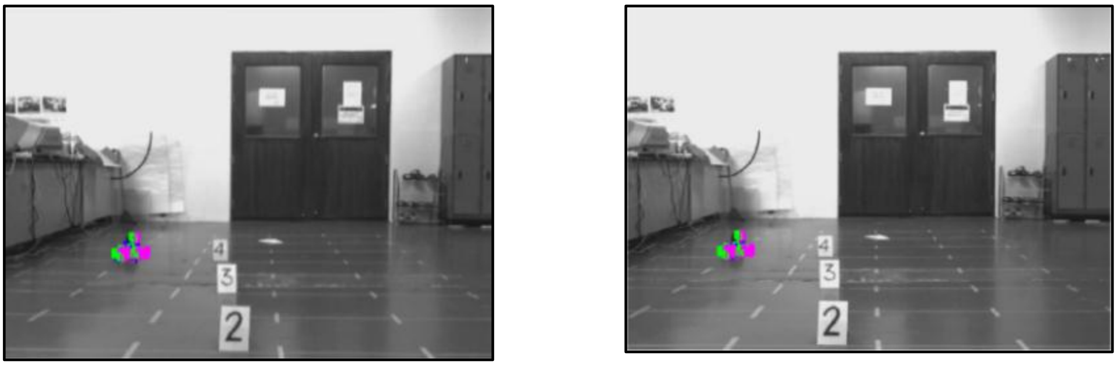
รูปที่ 4.9 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 1



รูปที่ 4.10 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 2

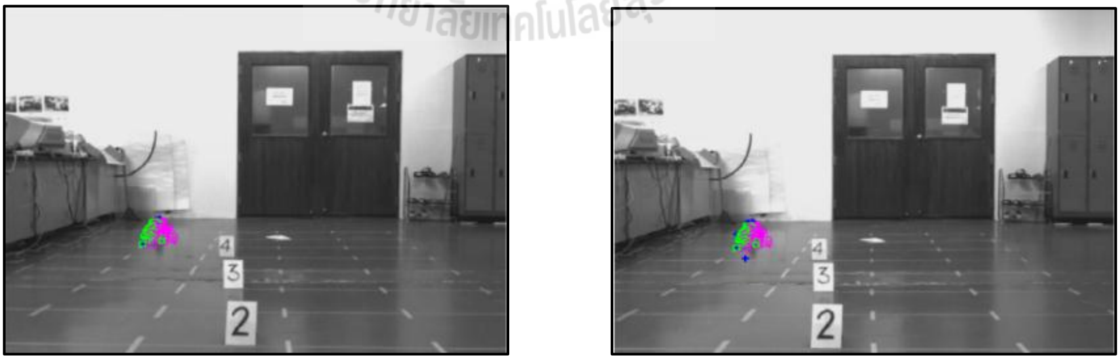


รูปที่ 4.11 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 3

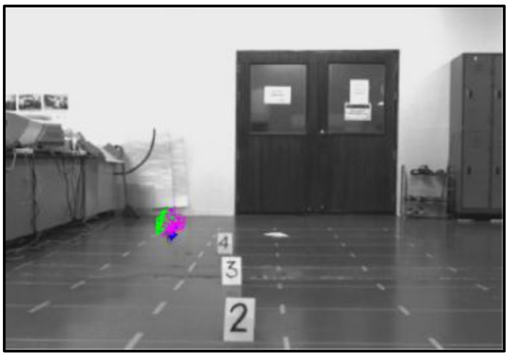


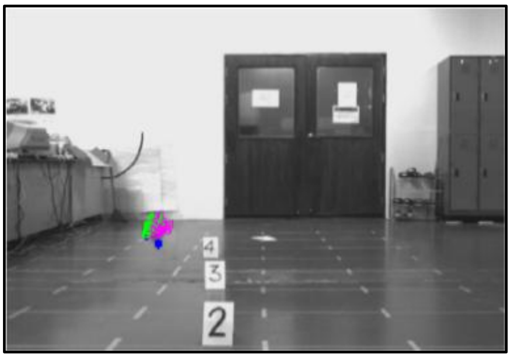
รูปที่ 4.12 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 4

รูปที่ 4.13 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 5



รูปที่ 4.14 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 6





รูปที่ 4.15 ผลการจับคู่ภาพของกล้องซ้ายและขวาของตัวอย่างที่ 7

รูปที่ 4.9 เป็นผลของการจับคู่ของตัวอย่างที่ 1 ผลการจับคู่ทั้งหมด 37 ตำแหน่ง มีผิดพลาด 2 ตำแหน่ง คิดเป็นความผิดพลาดร้อยละ 5.41 โดยความผิดพลาดของการจับคู่คิดจากจำนวนที่ถูกจับคู่กับจำนวนการจับคู่ที่ถูกต้อง โดยวัดจากการมองเห็นของมนุษย์ เพื่อตรวจสอบความแม่นยำของระบบจึงทำการทดสอบกับชุดตัวอย่างที่ 2-7 ซึ่งแสดงในรูปที่ 4.10 - 4.15 ผลปรากฏดังตารางที่ 4.1 ซึ่งค่าความผิดพลาดของข้อมูลทั้งหมดน้อยกว่าร้อยละ 10

ตารางที่ 4.1 ค่าความผิดพลาดของการจับคู่ภาพของกล้องซ้ายและขวาในตัวอย่างที่ 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

4.4 การหาระยะห่างระหว่างกล้องกับวัตถุและการบูรณะสภาวะแวดล้อม 3 มิติ

การหาระยะห่างระหว่างกล้องกับวัตถุสามารถคำนวณได้จาก

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

$$X = \frac{Z \times x'_l}{f} \tag{4.3}$$

$$Y = \frac{Z \times y'_l}{f} \tag{4.4}$$

โดยที่ Z คือ ระยะห่างระหว่างกล้องและวัตถุ

X,Y คือ ตำแหน่งในแนวนอนและแนวตั้งตามลำดับ โดยมีจุดอ้างอิงเป็นจุดกึ่งกลางของรูปภาพ

b คือ ระยะห่างระหว่างกล้องซ้ายและกล้องขวา

f คือ ความยาวโฟกัสของกล้อง

x'_l, x'_r คือ ระยะห่างระหว่างตำแหน่งในแนวแกนนอนกับจุดกึ่งกลางรูปของกล้องซ้ายและขวาตามลำดับ

y'_l คือ ระยะห่างระหว่างตำแหน่งในแนวแกนตั้งกับจุดกึ่งกลางรูป

จากตัวอย่างที่ 1 เมื่อผ่านขั้นตอนการจับคู่แล้วคำนวณจะได้ค่า Z ดังตารางที่ 4.2 เมื่อกำหนดให้ ระยะห่างระหว่างกล้องซ้ายและขวา (b) เท่ากับ 5 เซนติเมตร และความยาวโฟกัส (f) เท่ากับ 8 มิลลิเมตร เมื่อทำการคำนวณเสร็จสิ้นจะเห็นได้ว่าค่าที่ได้เท่ากับ 1.904 เมตรซึ่งค่าจริงของตัวอย่างคือ 2.5 เมตร จึงได้ทำการทดลองโดยใช้ตัวอย่างที่ 2 ถึง 7 ปรากฏว่ามีทุกตำแหน่งคลาดเคลื่อน จึงสังเกตได้ว่า ค่าที่ทำให้ผิดพลาดนั้นอยู่ที่ค่าความต่างของตำแหน่งแกน x ซึ่งเกิดจากการปรับขนาดรูปที่ได้จากกล้องให้มีขนาดเล็กลงจากเดิมซึ่งค่าที่ส่งผลให้การคำนวณคลาดเคลื่อนนั้นเป็นค่าคงที่ซึ่งเท่ากับ 5 จึงนำค่าคงที่นี้ไปลบออกจากค่าความต่างเดิมจะได้ผลการทดลองที่ปรับปรุงแล้วดังตารางที่ 4.3

เมื่อทำการคำนวณหาระยะห่างระหว่างวัตถุและกล้องได้แล้วจึงทำการคำนวณหาระยะในแนวแกนนอนและแกนตั้งของระบบจริง ซึ่งมีจุดอ้างอิงคือตรงจุดกึ่งกลางของภาพ ดังสมการที่ 4.3 และ 4.4 จะได้ผลการทดลองดังรูปที่ 4.13 – 4.15

ตารางที่ 4.2 ข้อมูลที่นำมาใช้ในการคำนวณหาระยะทางจากผลการทดลองในตัวอย่างที่ 1

ภาพซ้าย				ภาพขวา				$(x_l' - x_r')$	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

มหาวิทยาลัยเทคโนโลยีสุรนารี

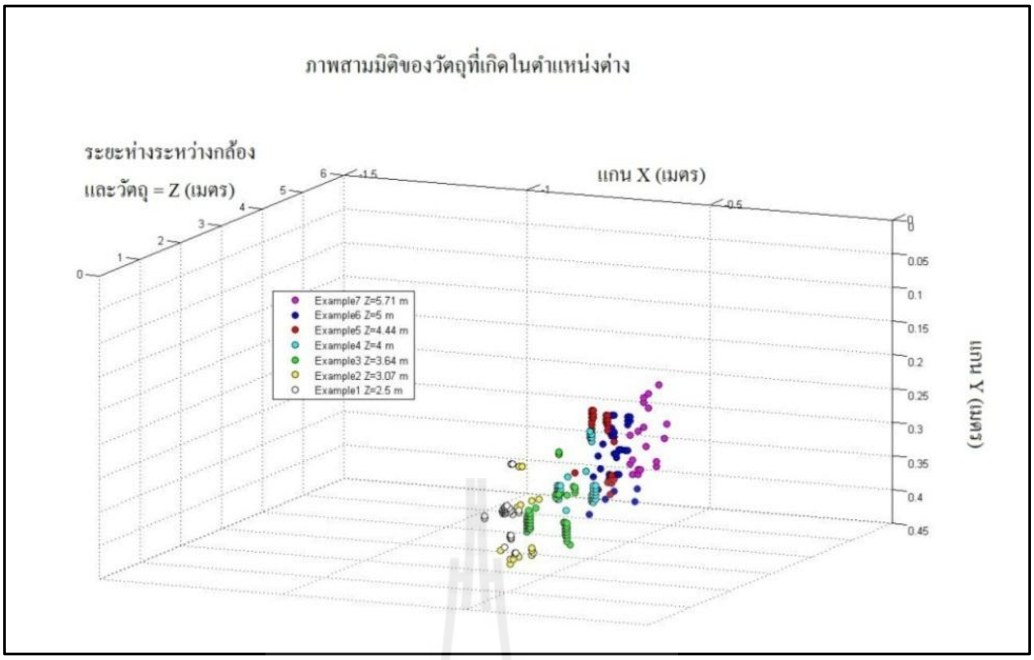
40	24	315	1.38839	49	3	319	1.421149	21	1.904

ตารางที่ 4.2 ข้อมูลที่นำมาใช้ในการคำนวณหาระยะทางจากผลการทดลองในตัวอย่างที่ 1 (ต่อ)

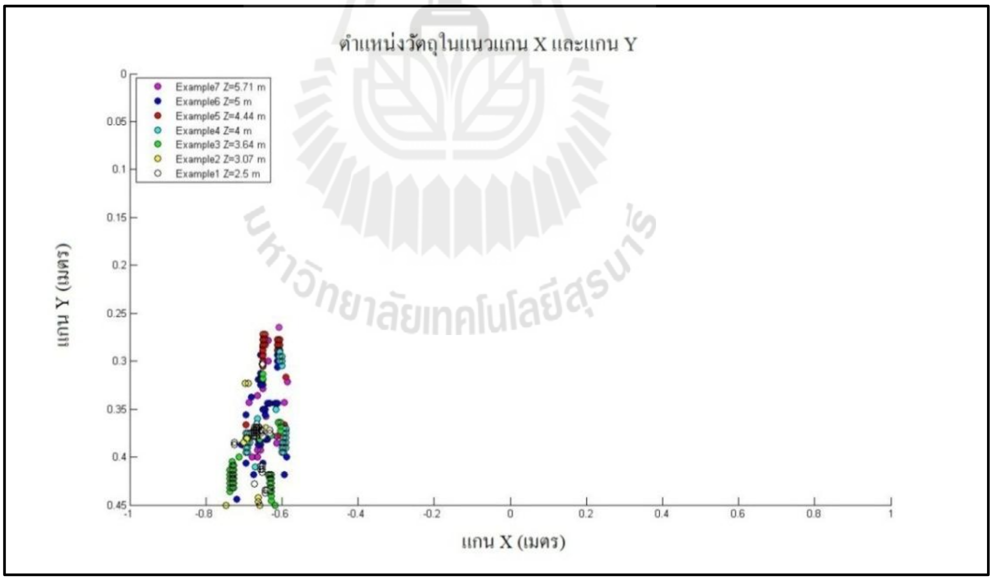
ภาพซ้าย				ภาพขวา				$(x_l' - x_r')$	Z
ลำดับ	x	y	มุม	ลำดับ	x	y	มุม		
41	24	316	1.404309	50	3	320	1.360304	21	1.904
44	46	323	1.428906	54	25	327	1.565879	21	1.904
45	47	323	1.303429	55	26	327	1.272486	21	1.904
46	46	324	1.362406	57	25	328	1.425486	21	1.904
47	47	324	1.29221	58	26	328	1.270589	21	1.904
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

ตารางที่ 4.3 เปอร์เซนต์ความถูกต้องของระยะห่างระหว่างวัตถุกับกล้องที่คำนวณได้กับค่าระยะห่างจริง

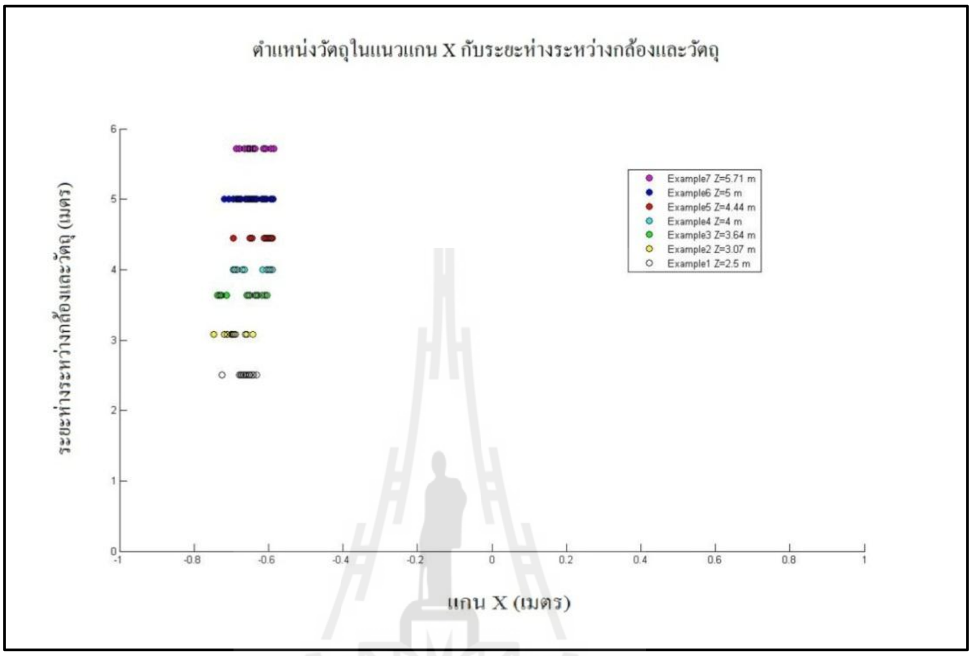
ชุดตัวอย่างที่	Z (เมตร)	ค่าจริง(เมตร)	ค่าความคลาดเคลื่อน(%)
1	2.5	2.5	0
2	3.08	3	2.67
3	3.64	3.5	4
4	4	4	0
5	4.44	4.5	6
6	5	5	0
7	5.714	5.5	3.89



รูปที่ 4.16 ตำแหน่งวัตถุที่ระยะต่างๆในระนาบ 3 มิติ



รูปที่ 4.17 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ XY



รูปที่ 4.18 ตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ ZX

รูปที่ 4.16 แสดงตำแหน่งวัตถุที่ระยะต่างๆ ในระนาบ 3 มิติของวัตถุตั้งแต่ตัวอย่างที่ 1-7 ซึ่งสังเกตได้ว่าพิกัดของวัตถุในภาพกับวัตถุที่เกิดในภาพจริงมีตำแหน่งที่ใกล้เคียงกัน สังเกตได้จากรูปที่ 4.17 เมื่อกำหนดให้แนวแกน Y เปรียบเสมือนพื้นของภาพจริงจะเห็นได้ว่าพิกัดของวัตถุที่คำนวณได้ของตัวอย่างที่ 1-7 อยู่ในแนวระนาบเดียวกัน จึงกำหนดได้ว่า ที่ระยะห่างจากจุดกึ่งกลางของการติดตั้งกล้องลงมา 0.45 เมตรคือพื้นของโลก และแนวการเคลื่อนที่ของวัตถุมีการเคลื่อนที่ไปทางด้ายซ้ายของรูปดังแสดงไว้รูปที่ 4.15 เมื่อเปรียบเทียบกับการเคลื่อนที่จริงของวัตถุจะสังเกตได้ว่ามีการเคลื่อนที่ไปในแนวเดียวกัน จากสูตรความสัมพันธ์ของสมการที่ 4.2 - 4.4 จะเห็นได้ว่า ถ้าสมการที่ 4.2 ซึ่งเป็นสมการหาระยะทางระหว่างวัตถุและกล้อง (Z) หาระยะผิดพลาดจะทำให้ระนาบตำแหน่งในแกน X และ Y ผิดพลาดไปด้วย ซึ่งตำแหน่งที่เกิดจากการพลอตกราฟในระนาบ X และ Y เป็นจุดเดียวกับการเกิดการไหลเชิงแสงที่ถูกจับคู่

4.5 ผลการทดสอบการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ

การทดสอบระบบการหาระยะระหว่างกล้องและวัตถุจะทำการเก็บข้อมูล ในช่วงสภาวะต่าง ๆ จากสภาพจริงที่ทั้งภายในอาคารและภายนอกอาคารทั้งเคลื่อนที่และไม่เคลื่อนที่ ผลที่ได้ระบบมีความคงทนต่อการเปลี่ยนแปลงสภาพแวดล้อมได้ดีและสามารถวัดระยะวัตถุที่เคลื่อนที่ได้อย่างมีประสิทธิภาพและได้ทำการทดสอบจากแฟ้มลำดับภาพทั้งหมด 2 แฟ้มลำดับภาพ ประกอบด้วยลำดับภาพทั้งหมด 1200 ลำดับภาพ โดยแบ่งออกเป็นประเภทได้ดังนี้

4.5.1 วัตถุเคลื่อนที่ภายในอาคาร

ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายในอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่แบ่งออกเป็น

-วัตถุเคลื่อนที่หนึ่งตำแหน่งภายในอาคาร

ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายในอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 1 ตำแหน่ง ผลการทดสอบเป็นไปตามรูปที่ 4.17

- วัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร

ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายในอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 2 ตำแหน่ง ผลการทดสอบเป็นไปตามรูปที่ 4.17

4.5.2 วัตถุเคลื่อนที่ภายนอกอาคาร

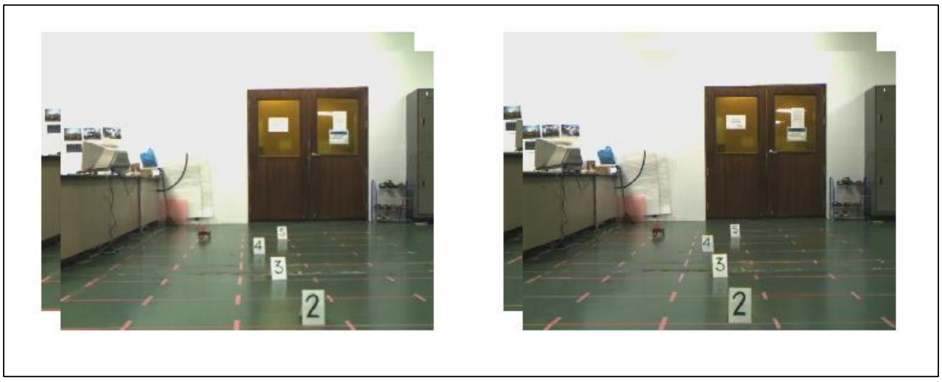
ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายนอกอาคารโดยกำหนดให้กล้องติดตั้งแบบขนานให้สามารถเคลื่อนที่ได้ ซึ่งการทดลองนี้แบ่งออกเป็น

-วัตถุเคลื่อนที่ภายนอกอาคาร โดยกล้องติดตั้งอยู่กับที่

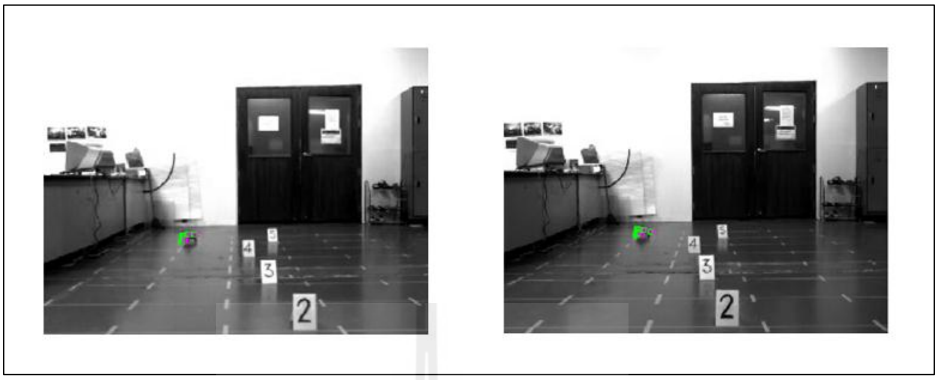
ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายนอกอาคารโดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ ผลการทดสอบเป็นไปตามรูปที่ 4.17

-วัตถุเคลื่อนที่ภายนอกอาคาร โดยกล้องเคลื่อนที่ไปพร้อมกับวัตถุ

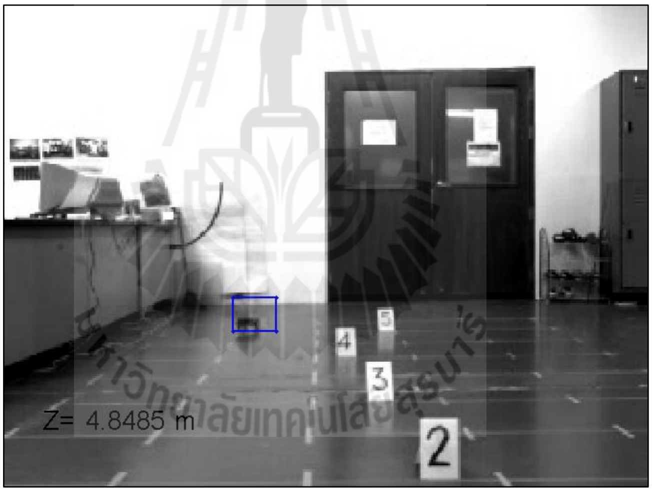
ในชุดตัวอย่างนี้เลือกนำสภาพแวดล้อมภายนอกอาคารโดยกำหนดให้กล้องถูกติดตั้งแบบเคลื่อนที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ ผลการทดสอบเป็นไปตามรูปที่ 4.17



(ก) ลำดับภาพจากกล้องซ้ายและขวา

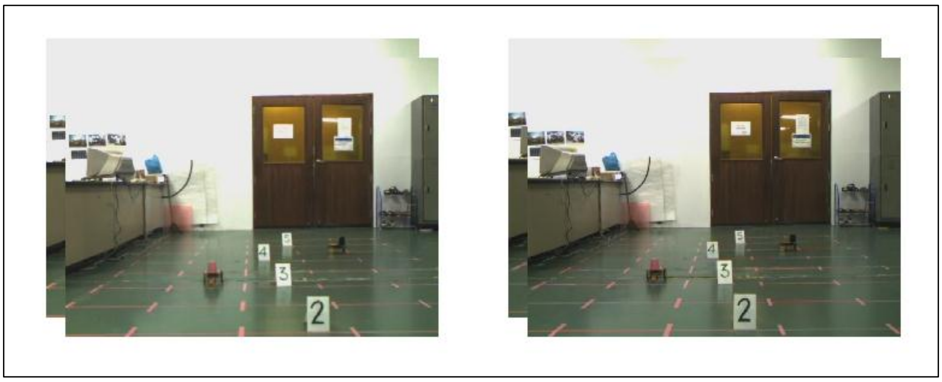


(ข) ผลการจับคู่ของภาพซ้ายและขวา



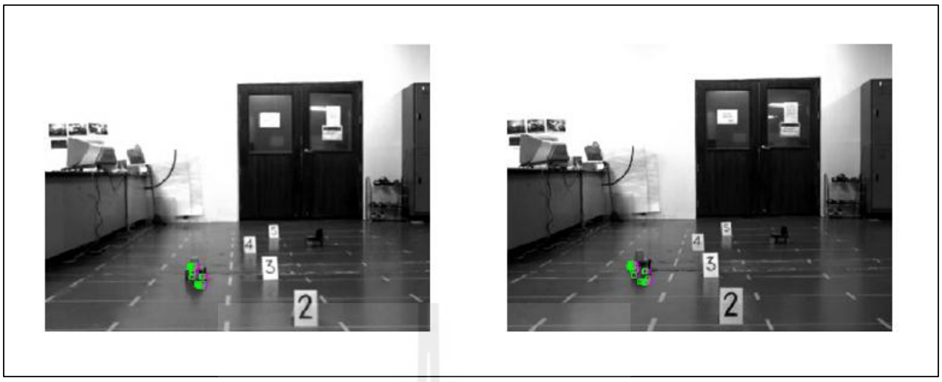
(ค) ผลการหาระยะทางจากกล้องและวัตถุ

รูปที่ 20 ผลการทดสอบการหาระยะวัตถุภายในอาคาร

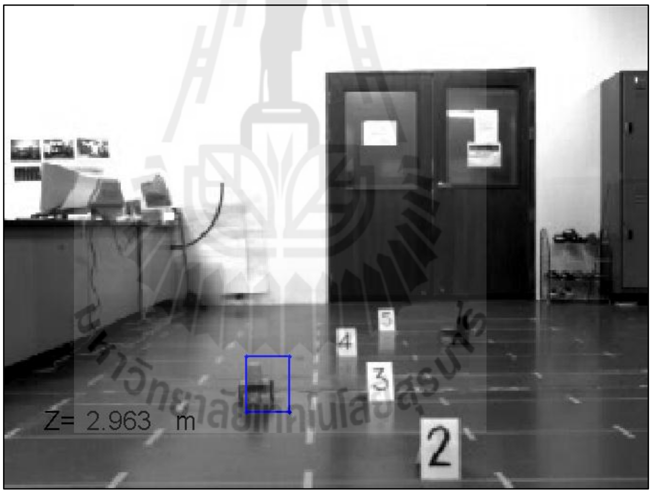


มหาวิทยาลัยเทคโนโลยีสุรนารี

(ก) ลำดับภาพจากกล้องซ้ายและขวา

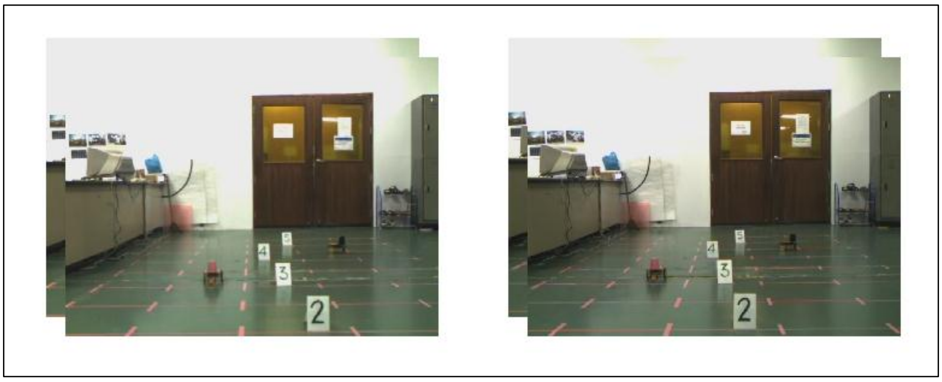


(ข) ผลการจับคู่ของภาพซ้ายและขวา

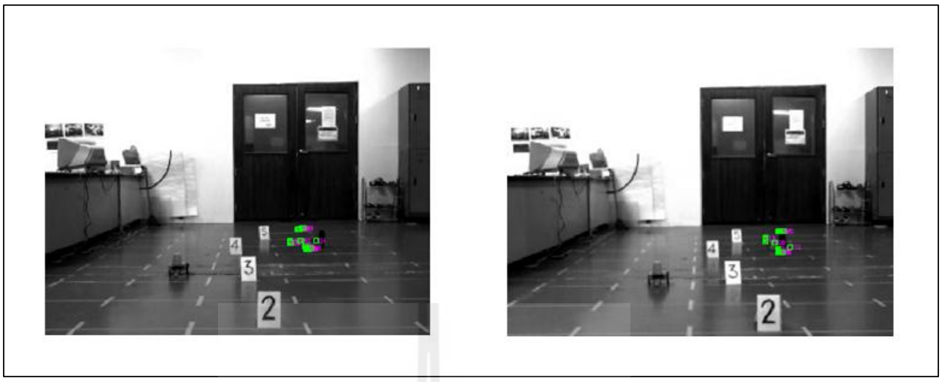


(ค) ผลการหาระยะทางจากกล้องและวัตถุ

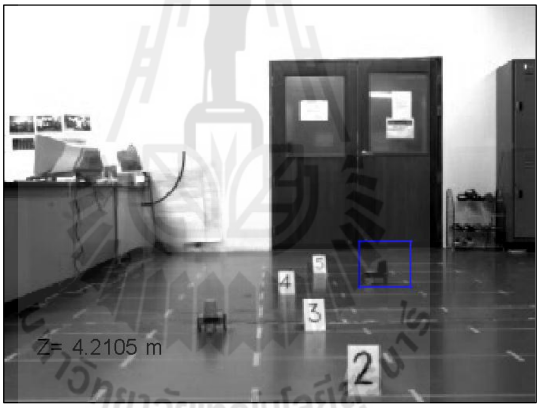
รูปที่ 4.16 ผลการทดสอบการหาระยะวัตถุภายในอาคาร



(ก) ลำดับภาพจากกล้องซ้ายและขวา



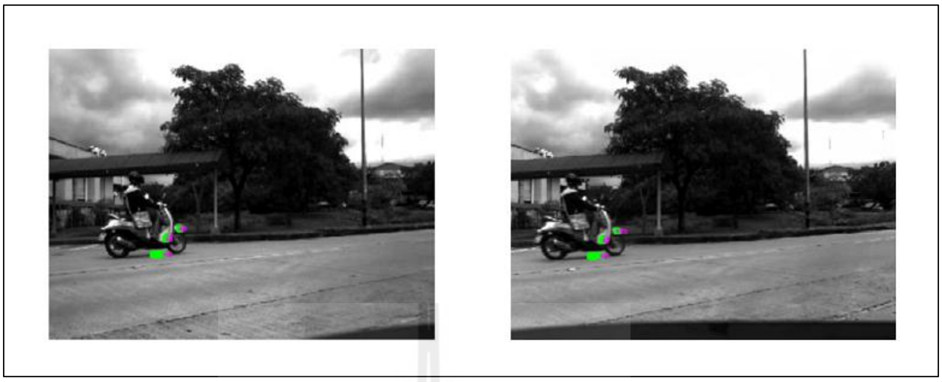
(ข) ผลการจับคู่ของภาพซ้ายและขวา



(ค) ผลการหาระยะทางจากกล้องและวัตถุ

รูปที่ 4.17 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร

(ก) ลำดับภาพจากกล้องซ้ายและขวา



(ข) ผลการจับคู่ของภาพซ้ายและขวา

(ค) ผลการหาระยะทางจากกล้องและวัตถุ

รูปที่ 4.18 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร

(ก) ลำดับภาพจากกล้องซ้ายและขวา

(ข) ผลการหาระยะทางจากกล้องและวัตถุ

รูปที่ 4.19 ผลการทดสอบการหาระยะวัตถุเคลื่อนที่ภายนอกอาคาร

4.5 สรุป

ระบบการหาระยะระหว่างกล้องและวัตถุโดยใช้เทคนิคเครือข่ายประสาทเทียมร่วมกับการไหลเชิงแสง สามารถหาจุดที่สอดคล้องกันของภาพซ้ายและขวาได้โดยอาศัยตำแหน่ง มุมและขนาดของการไหลเชิงแสงเป็นอินพุตให้กับระบบ เมื่อระบบทำการประมวลผลแล้วจะทำให้ทราบพิกัดของวัตถุที่เคลื่อนที่ภายในภาพ โดยอาศัยจุดกึ่งกลางและระยะห่างของกล้องและวัตถุเป็นจุดอ้างอิง พิกัดของวัตถุที่เกิดขึ้นเมื่อเทียบกับระยะจริงแล้วมีความแม่นยำอยู่ที่ประมาณ 90 เปอร์เซนต์

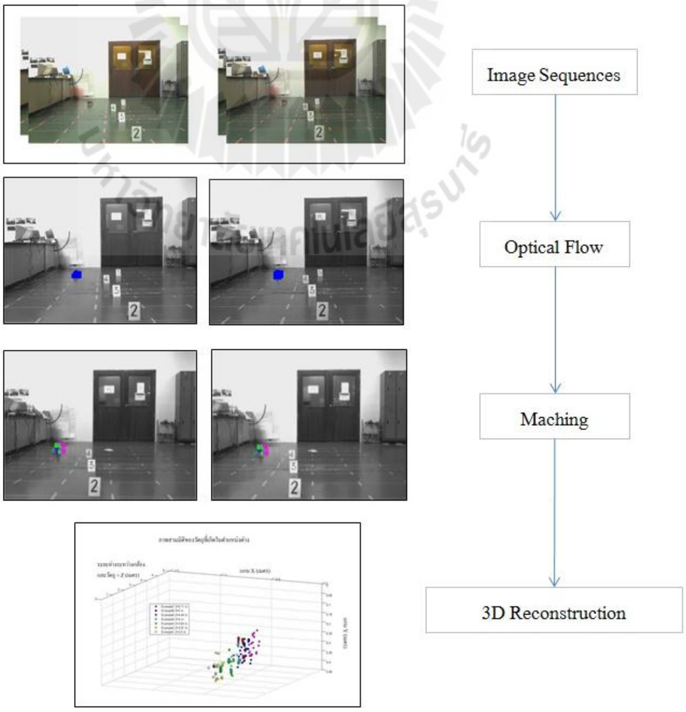
มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 5
วิเคราะห์และอภิปรายผลการทดลอง

5.1 บทนำ

ในบทนี้ได้นำเสนอการวิเคราะห์และอภิปรายผลการทดลองด้วยวิธีการดังที่กล่าวไปแล้วในบทที่ 4 โดยจะทำการวิเคราะห์ปัจจัยต่างๆ ที่เกี่ยวข้องกับระบบการบูรณะสภาวะแวดล้อม 3 มิติ ซึ่งได้แก่ การทดสอบอิทธิพลของการไหลเชิงแสงเมื่ออยู่ภายในและภายนอกอาคาร การจับคู่กันของจุดภาพด้วยเครือข่ายฮอพฟิลด์ 2 มิติ การวิเคราะห์ระยะห่างของการติดตั้งกล้องกับการหาระยะทางระหว่างกล้องกับวัตถุ และการหาระยะทางในสภาวะต่างๆ ดังรายละเอียดต่อไปนี้

5.2 โครงสร้างของระบบการบูรณะสภาวะแวดล้อม 3 มิติ



รูปที่ 5.1 โครงสร้างของระบบโดยรวม

ระบบการบูรณะสภาวะแวดล้อม 3 มิติสามารถสรุปการทำงานหลักของระบบออกเป็น 3 กระบวนการ คือ กระบวนการที่หนึ่ง เป็นกระบวนการคัดแยกวัตถุเคลื่อนที่โดยใช้การไหลเชิงแสง กระบวนการที่สอง เป็นกระบวนการการจับคู่ตำแหน่งที่สอดคล้องกันของจุดภาพรูปซ้ายและขวา และกระบวนการที่สาม เป็นกระบวนการบูรณะสภาวะแวดล้อม 3 มิติและการหาระยะทางระหว่างกล้องและวัตถุ ดังแสดงในรูปที่ 5.1

5.3 ผลการทดสอบการคัดกรองการไหลเชิงแสง

การไหลเชิงแสงใช้การวิเคราะห์จากการเปลี่ยนแปลงของความเข้มแสงภายในภาพแต่ละลำดับภาพ ดังนั้นในที่นี้จึงใช้วิธีการของลูคัส-คานาเดมาใช้ในภาพที่มีสภาวะแวดล้อมต่างๆกัน ได้แก่ ภาพที่เกิดภายในอาคารและภาพที่เกิดภายนอกอาคารซึ่งความเข้มแสงมีการเปลี่ยนแปลงอยู่ตลอดเวลา

(ก) ภาพภายในอาคาร

(ข) การไหลเชิงแสงที่ยังไม่ผ่านการกรอง

รูปที่ 5.2 การไหลเชิงแสงของลูคัส-คานาเด

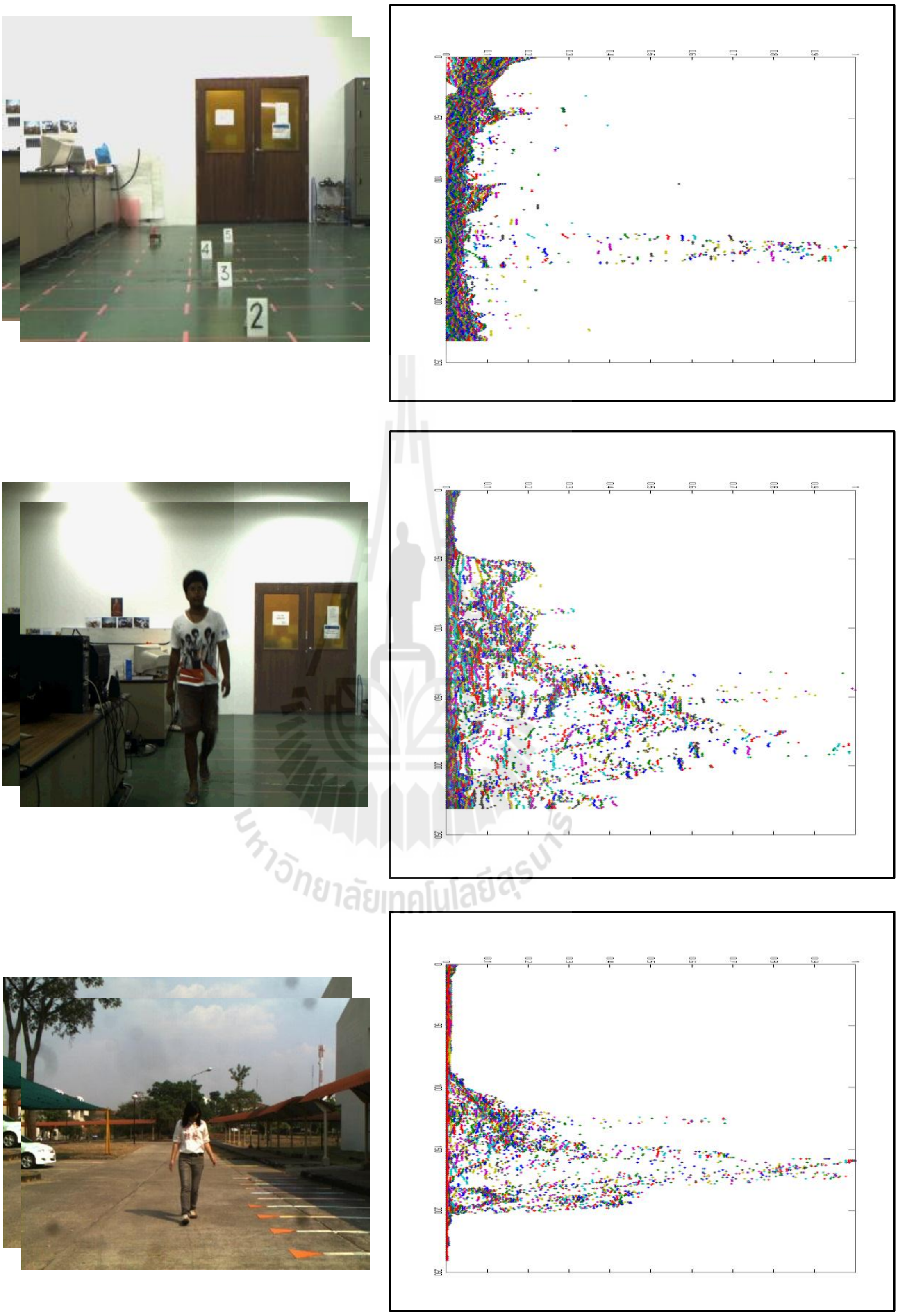
(ก) ภาพภายนอกอาคาร

(ข) การไหลเชิงแสงที่ยังไม่ผ่านการกรอง

รูปที่ 5.3 การไหลเชิงแสงของลูคัส-คานาเด

จากรูปที่ 5.2 และ 5.3 จะสังเกตได้ว่าเมื่อนำการไหลเชิงแสงมาใช้สามารถคัดแยกวัตถุออกจากพื้นหลังได้ดีทั้งภายในและภายนอกอาคาร แต่ยังมีการไหลเชิงแสงที่พื้นหลังอยู่จึงไม่สามารถนำมาใช้ได้เลยทันที นอกจากมีการไหลเชิงแสงที่พื้นหลังแล้วยังมีตำแหน่งที่ไม่ต้องการพิจารณาอีกด้วยจึงต้องคัดกรองสมาชิกการไหลเชิงแสงออก ด้วยการกรองโดยใช้ขนาดของการไหลเชิงแสงเอง ในแต่ละสภาวะแวดล้อมหรือแต่ละลำดับภาพการไหลเชิงแสงมีค่าไม่เท่ากัน ดังแสดงในรูปที่ 5.4

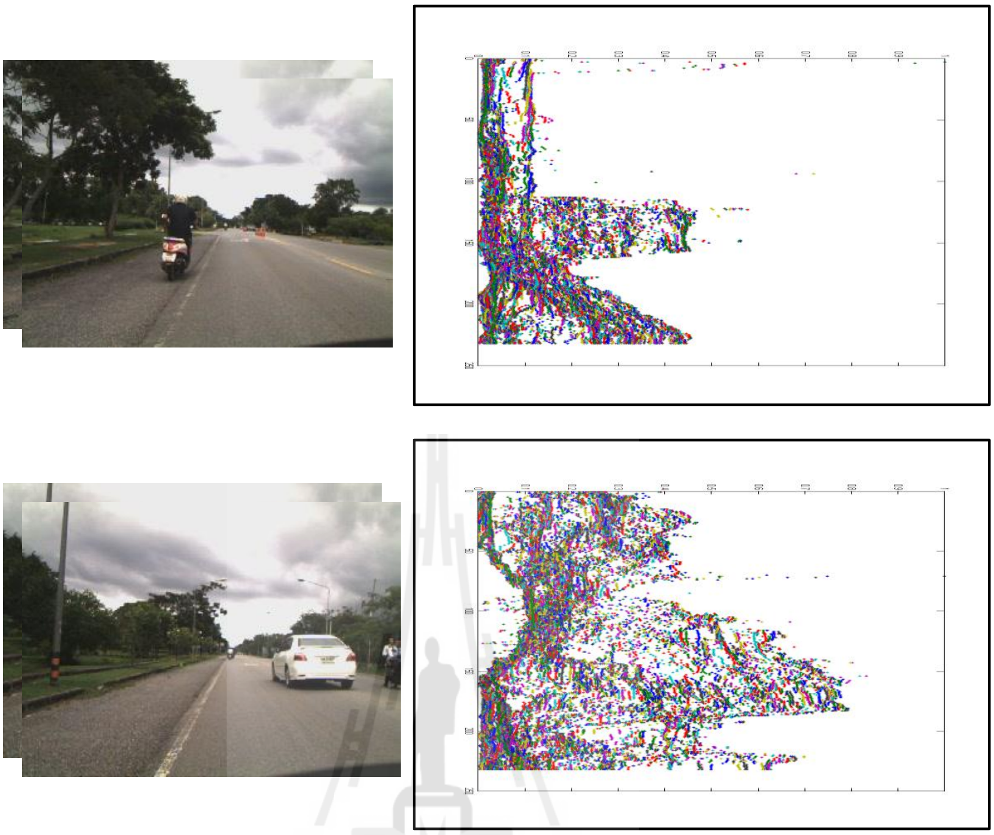
เมื่อกำหนดให้ขนาดของการไหลเชิงแสงสูงสุดในแต่ละภาพมีค่าเท่ากับ 1 ขนาดของวัตถุที่เคลื่อนที่จะมีค่ามากกว่าการไหลเชิงแสงของพื้นหลัง จะเห็นได้จากกราฟความสัมพันธ์ระหว่างขนาดและตำแหน่งการเกิดการไหลเชิงแสงในรูปที่ 5.4 ตำแหน่งที่วัตถุเคลื่อนที่จะมีขนาดการเกิดการไหลเชิงแสงมากกว่าบริเวณพื้นหลังจะอยู่ในช่วง 0.5 – 1 ทั้งในสภาวะที่อยู่ภายในและภายนอกอาคาร จึงเลือกขนาดการไหลเชิงแสงในช่วง 0.5-1 นี้ทำการคัดกรองภาพทั้งหมด แต่เมื่อนำไปใช้กับสถาวะที่กล้องเคลื่อนที่ไปพร้อมกับการตรวจจับภาพปรากฏว่าไม่สามารถคัดกรองได้เหมือนกับที่กล่าวไว้แล้วข้างต้น



(ก) ภาพต้นแบบในสภาวะต่าง ๆ (ข) ความถี่การเกิดการไหลเชิงแสง

รูปที่ 5.4 เปรียบเทียบขนาดการไหลเชิงแสงในสภาวะต่างๆเมื่อกล้องถูกติดตั้งอยู่กับที่

มหาวิทยาลัยเทคโนโลยีสุรนารี

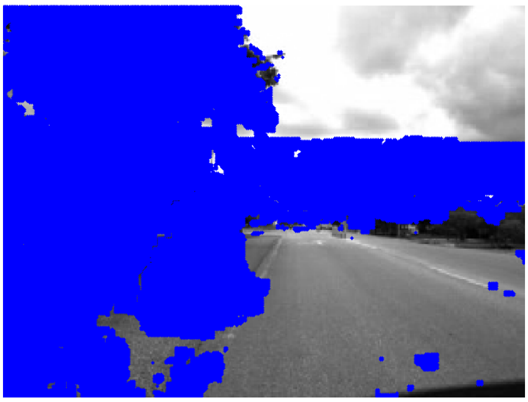


(ก) ภาพต้นแบบในสภาวะต่าง ๆ (ข) ความถี่การเกิดการไหลเชิงแสง

รูปที่ 5.5 เปรียบเทียบขนาดการไหลเชิงแสงในสภาวะต่างๆเมื่อกล้องเคลื่อนที่

จากการทดสอบในสภาวะที่กล้องได้ถูกติดตั้งแบบไม่เคลื่อนที่นั้นได้กำหนดช่วงการคัดกรองไว้ที่ 0.5 – 1 ซึ่งเมื่อทำการทดลองกับการติดตั้งกล้องแบบเคลื่อนที่ทำให้ไม่สามารถคัดกรองวัตถุได้ ซึ่งรูปที่ 5.5 แสดงกราฟความสัมพันธ์ขนาดการไหลเชิงแสงกับตำแหน่งในแกนแนวนอนของภาพได้ว่า บริเวณด้านซ้ายของกราฟซึ่งเป็นบริเวณขอบล่างของภาพมีการไหลเชิงแสงที่มีขนาดใกล้เคียงกับวัตถุที่เคลื่อนที่ ดังนั้นนอกจากการกรองด้วยขนาดของการไหลเชิงแสงแล้ว จึงต้องทำการตัดขอบออกด้วย เมื่อพิจารณารูปที่ขนาด 640x480 จะทำการตัดขอบภาพออก 30 แถว จะได้ผลการทดลองดังรูป 5.6 ซึ่งสามารถนำการคัดกรองนี้ไปใช้กับรูปที่กล้องถูกติดตั้งอยู่กับที่ได้เพราะภาพที่ได้เมื่อกล้องถูกติดตั้งอยู่กับที่นั้นไม่มีการเคลื่อนไหวบริเวณขอบภาพ

(ก) ภาพต้นแบบ



(ข) ภาพที่ยังไม่ได้ทำการกรองการไหลเชิงแสง

(ค) ภาพที่ทำการกรองแล้ว

รูปที่ 5.6 ผลการคัดกรองการไหลเชิงแสง

5.4 ผลการทดสอบการจับคู่ด้วยเครือข่ายฮอพฟิลด์ 2 มิติ

กระบวนการทำงานของเครือข่ายฮอพฟิลด์ 2 มิติมีอยู่ 3 ขั้นตอน ได้แก่ ขั้นตอนแรกนำอินพุตที่ได้มาคำนวณหาค่าน้ำหนักประสาท ขั้นตอนที่ 2 คือการประมวลผลของเครือข่าย และขั้นตอนสุดท้ายคือการปรับสถานะของแต่ละนิวรอน ขั้นตอนที่สำคัญคือขั้นตอนการคำนวณหาค่าน้ำหนักประสาทเพราะเป็นการรวมเอาองค์ประกอบและเงื่อนไขของการทำงานของเครือข่ายเอาไว้

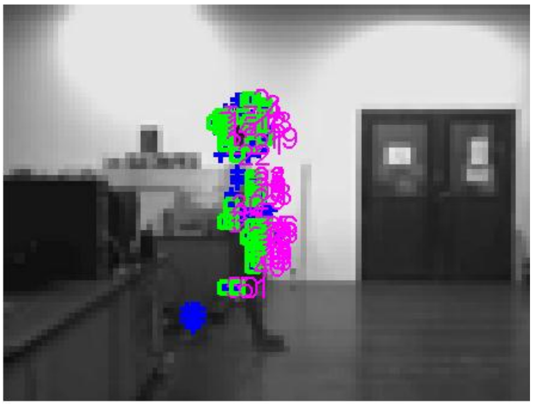
การคำนวณหาค่าน้ำหนักประสาทนั้นเริ่มจากการนำตำแหน่งและมุมของการไหลเชิงแสงของภาพซ้ายและภาพขวามาคำนวณโดยใช้ฟังก์ชั่นไม่เชิงเส้นตามสมการที่ 5.1

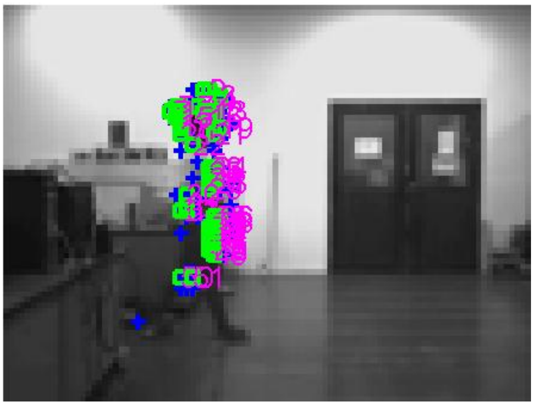
$$C_{ikjl} = (2/[1+e^{\lambda_{(X-\theta)}}])-1 \qquad (5.1)$$

โดยที่การทดลองครั้งนี้ใช้ $\theta = 0.2$ และ $\lambda = 1$ ซึ่งจะได้ผลการทดลองดังรูปที่ 5.7 – 5.11

(ก) ภาพซ้าย (ข) ภาพขวา

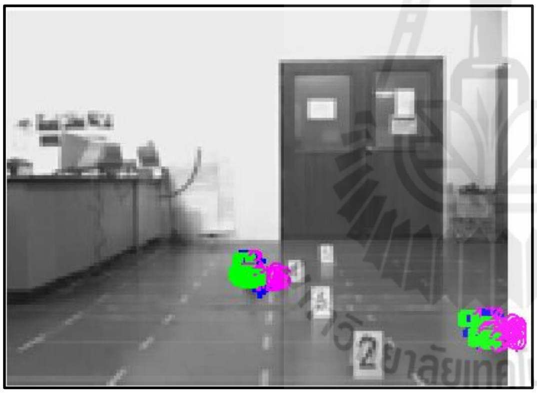
รูปที่ 5.7 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร

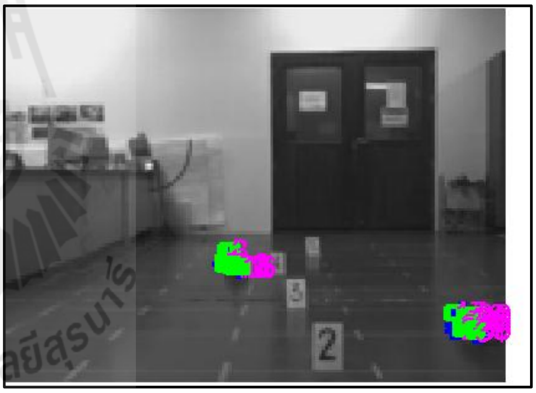




(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ 5.8 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร





(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ 5.9 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ 5.10 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายนอกอาคาร

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ 5.11 ผลการจับคู่ตำแหน่งภาพซ้ายและขวาของภาพที่เกิดภายในอาคาร

ประสิทธิภาพของการจับคู่ในกระบวนการนี้ขึ้นอยู่กับอินพุตและการเลือกใช้สมการที่นำมาใช้ในการคำนวณหาค่าน้ำหนักประสาทเทียม โดยพิจารณาโครงสร้างของเครือข่ายฮอพฟีลด์ 2 มิติ แล้วจะเป็น N_l x N_r นิวรอน ซึ่งเป็นโครงสร้างเมตริก 2 มิติ แต่เมื่อนำมาคำนวณหาค่าน้ำหนักประสาท T_{ijkl} จะทำให้ได้เมตริก 4 มิติขึ้นมา ด้วยเหตุนี้จึงทำให้ต้องใช้พื้นที่หน่วยความจำของคอมพิวเตอร์มากขึ้น อินพุตของเครือข่ายหรือนิวรอนของเครือข่าย N_l รวมกับ N_r ไม่ควรเกิน 200

นิวรอน แต่จำนวนนิวรอนก็มีอิทธิพลต่อความถูกต้องของระบบนั้นคือ จำนวนตำแหน่งอินพุตมากจะได้ผลการทดสอบที่แม่นยำกว่าอินพุตน้อยๆ

เมื่อดำเนินการจับคู่ตำแหน่งภายในภาพในสภาวะแวดล้อมต่างๆ ได้แก่ภาพภายในอาคารและภายนอกอาคาร พบว่าระบบสามารถทำการค้นหาและจับคู่ภาพได้อย่างมีประสิทธิภาพแต่ยังคงมีในกรณีที่ เมื่อกล้องเคลื่อนที่ทำให้การตรวจจับหาตำแหน่งการเกิดการไหลเชิงแสงในนั้นมีประสิทธิภาพน้อยลงทำให้ความถูกต้องในการจับคู่ของภาพซ้ายและขวาน้อยลงด้วย

และเมื่อพิจารณาความสามารถในการค้นหาและจับคู่ตำแหน่งภาพของระบบเปรียบเทียบกับงานวิจัยในอดีต ดังเสนอในตาราง 5.1 พบว่าการตรวจจับมีความถูกต้อง 96.84 เปอร์เซนต์ ซึ่งงานวิจัยในอดีตส่วนใหญ่จะเป็นภาพนิ่งและมีการควบคุมปัจจัยต่างๆ ได้แก่ แสงสว่าง การติดตั้งกล้อง ขนาดของภาพ ความซับซ้อนของพื้นหลังและวิธีการการหาตำแหน่งที่หน้าสนใจ ในขณะที่งายวิจัยนี้ได้นำเสนอการจับคู่ภาพเคลื่อนไหวและยังครอบคลุมถึงสภาพปัจจัยต่างๆ ทำให้ระบบสามารถค้นหาวัตถุและหาตำแหน่งได้อย่างแม่นยำ

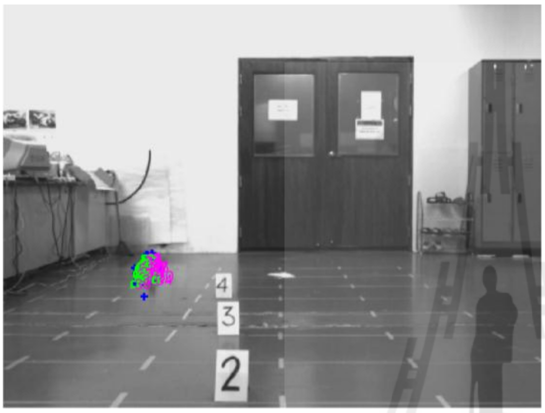
ตารางที่ 5.1 ความสามารถในการจับคู่ภาพของระบบและงานวิจัยที่เกี่ยวข้องในอดีต

ผู้แต่ง	Nasser, et al (1992)	Karim, et al (2002)	Somying (2012)
ประเภทการหาตำแหน่งที่น่าสนใจ	Moravec	Canny และMoravec	Optical Flow
ประสิทธิภาพของการจับคู่	94.34 (%)	93.94 %	96.84 %
จำนวนที่ถูกจับคู่	33	25	95
ประเภทของภาพ	ภาพนิ่ง	ภาพนิ่ง	ภาพต่อเนื่อง
ลายละเอียดภายในภาพ	ไม่ซับซ้อน	ซับซ้อน	ซับซ้อน
สภาพแวดล้อม	ภายในอาคาร	ภายในอาคาร	ภายในและภายนอกอาคาร

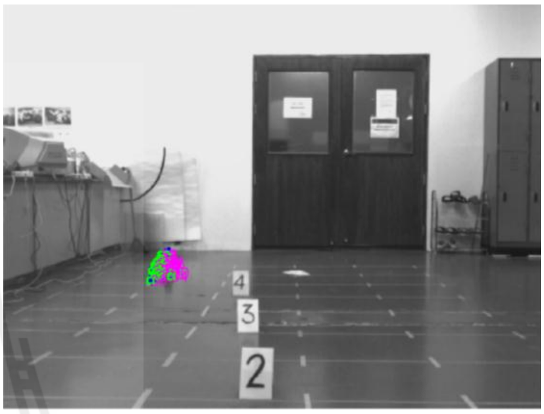
5.5 การบูรณะสภาวะแวดล้อม 3 มิติ

การคำนวณหาตำแหน่งพิกัด 3 มิติของวัตถุที่เกิดขึ้นในภาพ โดยใช้วิธีการทางเรขาคณิตนั้นมีปัจจัยที่ส่งผลต่อความแม่นยำได้แก่ ขนาดของภาพ และตำแหน่งการติดตั้งกล้อง

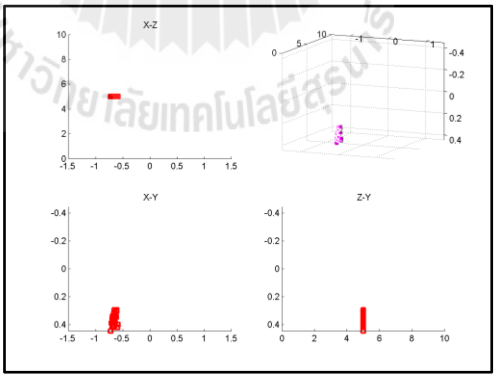
จากรูปที่ 5.12 จะเห็นได้ว่า เมื่อตำแหน่งที่เราสนใจจากภาพซ้ายและภาพขวา จากขั้นตอนการจับคู่ ถ้ากรณีที่การจับคู่ถูกต้องการคำนวณพิกัดจะได้ดังรูปที่ 5.12(ค) จะเห็นได้ว่าภาพฉายที่เกิดขึ้น จะมีตำแหน่งตรงตามวัตถุที่เกิดขึ้นในภาพ โดยใช้ภาพทางด้านซ้ายเป็นภาพอ้างอิง ตำแหน่งที่เกิดวัตถุในภาพคือบริเวณทางด้านซ้ายของรูป ตำแหน่งที่เกิดจากการคำนวณกราฟก็จะอยู่บริเวณทางด้านซ้าย เมื่อกำหนดให้จุกกึ่งกลางของภาพคือ (0,0)



(ก) ภาพซ้าย



(ข) ภาพขวา

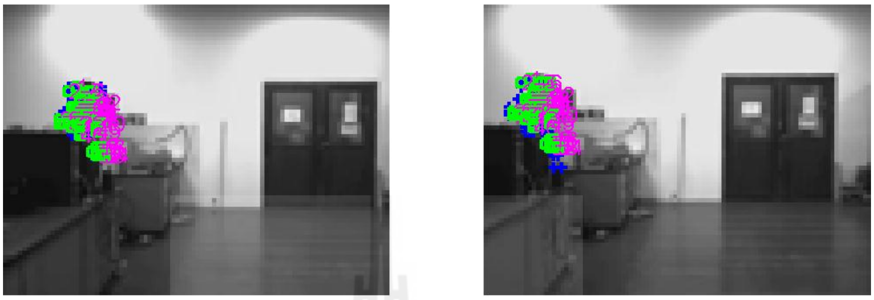


(ค) ภาพสามมิติและภาพฉายในระนาบต่างๆ

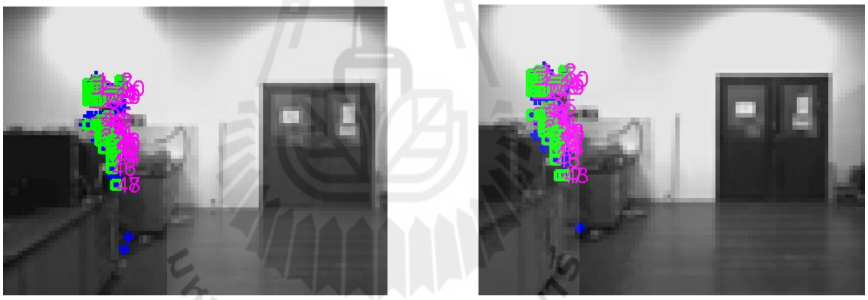
รูปที่ 5.12 การบูรณะสภาวะแวดล้อม 3 มิติ

มหาวิทยาลัยเทคโนโลยีสุรนารี

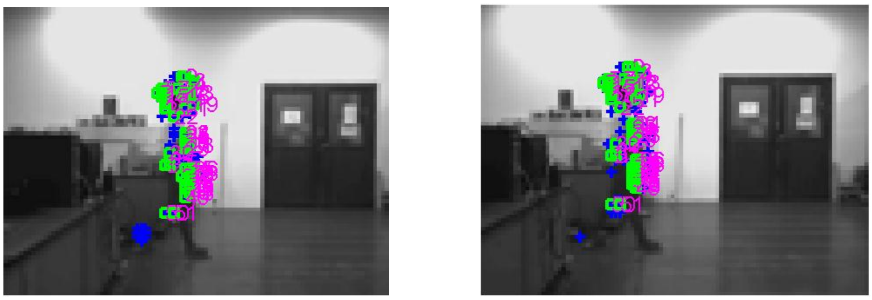
เมื่อทำการพิจารณาตำแหน่งที่เกิดวัตถุในระนาบ 3 มิติ ตำแหน่งที่เกิดขึ้นในภาพกับตำแหน่งที่เกิดขึ้นในการคำนวณจะสอดคล้องกัน สังเกตได้จากรูปที่ 5.13 วัตถุมีการเคลื่อนที่ จากด้านซ้ายไปด้านขวา ตำแหน่งที่ได้ก็มีการเคลื่อนที่จากซ้ายไปขวาเช่นกัน



(ก) การจับคู่กันของภาพซ้ายขวาเฟรมที่ 1

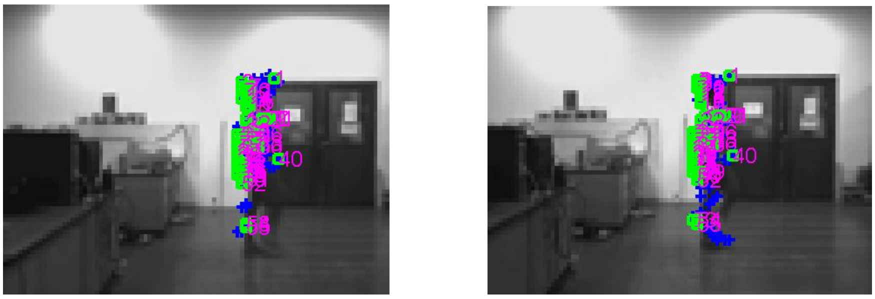


(ข) การจับคู่กันของภาพซ้ายขวาเฟรมที่ 2

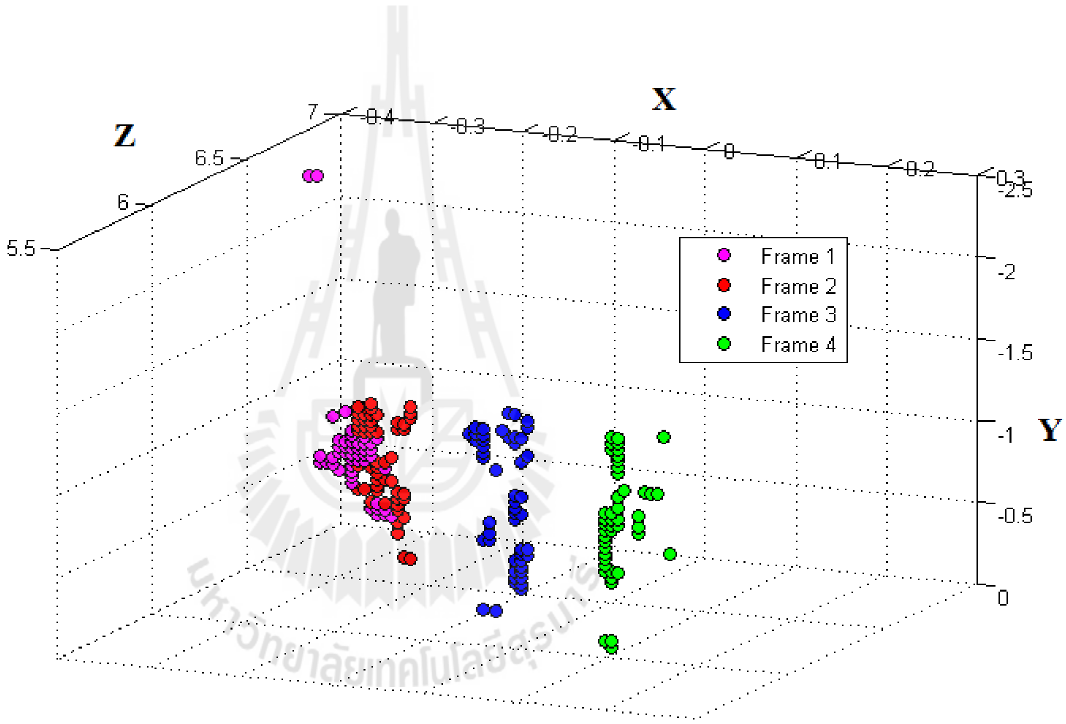


(ค) การจับคู่กันของภาพซ้ายขวาเฟรมที่ 3

รูปที่ 5.13 ผลการทดลองการคำนวณหาตำแหน่งในระนาบ 3 มิติ



(ง) การจับคู่กันของภาพซ้ายขวาเฟรมที่ 4



(จ) ภาพระนาบ 3 มิติ

รูปที่ 5.13 ผลการทดลองการคำนวณหาตำแหน่งในระนาบ 3 มิติ (ต่อ)

5.6 ผลการหาระยะทางในสภาวะต่างๆ

จากการผลการทดลองที่ผ่านมาจะเห็นได้ว่าความแม่นยำในการคำนวณหาตำแหน่งของวัตถุขึ้นอยู่กับการติดตั้งระยะห่างระหว่างกล้อง 2 ตัว ขนาดโฟกัสของเลนส์และขนาดของภาพที่ใช้ป้อนอินพุตในระบบ การติดตั้งกล้องยิ่งห่างกันมากช่วงการวัดระยะของวัตถุยิ่งละเอียดดังตารางที่ 5.2 แต่ข้อเสียของการติดตั้งกล้องห่างกันมากๆนั้น จะทำให้วัตถุที่ปรากฏด้านซ้ายหรือด้านขวาข้างใดข้างหนึ่งมากๆ จะทำให้วัตถุไม่ปรากฏในภาพและไม่สามารถคำนวณหาวัตถุได้ ความแตกต่างของภาพในแต่ละการติดตั้งระยะห่างระหว่างกล้องแสดงได้ดังรูปที่ 5.14

(ก) ภาพซ้ายและภาพขวาจากการติดตั้งกล้องห่างกัน 5 เซนติเมตร

(ข) ภาพซ้ายและภาพขวาจากการติดตั้งกล้องห่างกัน 10 เซนติเมตร

รูปที่ 5.14 ภาพที่เกิดจากการติดตั้งกล้องระยะห่างต่างๆ

(ค) ภาพซ้ายและภาพขวาจากการติดตั้งกล้องห่างกัน 15 เซนติเมตร

(ง) ภาพซ้ายและภาพขวาจากการติดตั้งกล้องห่างกัน 20 เซนติเมตร

รูปที่ 5.14 ภาพที่เกิดจากการติดตั้งกล้องระยะห่างต่างๆ (ต่อ)

เมื่อทำการพิจารณาหาค่าระยะทางของวัตถุกับกล้องตามลักษณะแวดล้อมต่างๆ โดยกำนหดให้

ตัวอย่างที่ 1 คือ วัตถุเคลื่อนที่หนึ่งตำแหน่งภายในอาคาร โดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 1 ตำแหน่ง

ตัวอย่างที่ 2 คือ วัตถุเคลื่อนที่มากกว่าหนึ่งตำแหน่งภายในอาคาร โดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ 2 ตำแหน่ง

ตัวอย่างที่ 3 คือ วัตถุเคลื่อนที่ภายนอกอาคาร โดยกล้องติดตั้งอยู่กับที่ โดยกำหนดให้กล้องถูกติดตั้งอยู่กับที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ

ตัวอย่างที่ 4 คือ วัตถุเคลื่อนที่ภายนอกอาคาร โดยกล้องเคลื่อนที่ไปพร้อมกับวัตถุ โดยกำหนดให้กล้องถูกติดตั้งแบบเคลื่อนที่และให้วัตถุเคลื่อนที่ได้อย่างอิสระ

ตารางที่ 5.2 ตารางแสดงระยะทางที่วัดได้สูงสุดในการติดตั้งกล้องและขนาดภาพต่าง ๆที่ป้อนให้กับระบบ (เมตร)

ขนาดภาพ	ระยะห่างในการติดตั้งกล้องซ้ายและกล้องขวา(เซนติเมตร)			
(pixel)	5	10	15	20
640x480	5	6	7	8

จากตารางที่ 5.2 จะเห็นได้ว่าเมื่อทำการติดตั้งกล้องซ้ายและขวาต่างกันจะทำให้การวัดหาระยะของวัตถุต่างกันด้วย ซึ่งในการทดสอบเมื่อทำการติดตั้งกล้องห่างกัน 5 เซนติเมตรระยะที่วัดได้สูงสุด 5 เมตร เมื่อทำการติดตั้งกล้องห่างกัน 10 เซนติเมตรระยะที่วัดได้สูงสุดที่ 6 เมตร เมื่อทำการติดตั้งกล้องห่างกัน 15 เซนติเมตรระยะที่วัดได้สูงสุดที่ 7 เมตร และเมื่อทำการติดตั้งกล้องห่างกัน 20 เซนติเมตรระยะที่วัดได้สูงสุดที่ 8 เมตร จะสังเกตได้ว่าระยะห่างในการติดตั้งกล้องแปรผันตรงกับระยะระหว่างกล้องและวัตถุ ซึ่งการติดตั้งนี้ส่งผลต่อค่าความต่างกันของภาพ (depth map) จึงทำให้วัดระยะได้มากขึ้นเมื่อติดตั้งกล้องห่างกันมากขึ้น แต่อย่างไรก็ตามเมื่อทำการติดตั้งกล้องห่างกันมากๆ จะทำให้วัตถุที่ปรากฏภายในภาพเมื่อถึงระยะหนึ่ง จะทำให้ข้อมูลภาพหายไป หรือกล้องไม่สามารถจับภาพได้ จึงต้องจำกัดขอบเขตของการวัดเพื่อให้ได้ค่าเหมาะสมที่สุด

ตารางที่ 5.3 ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในระยะต่างๆ

ระยะจริง	ระยะห่างในการติดตั้งกล้องซ้ายและกล้องขวา(เซนติเมตร)	ระยะที่วัดได้เฉลี่ย (เมตร)	เปอร์เซนต์ความถูกต้อง
2	5	2.10	95
	10	2.05	97.5
	15	2.03	98.5
	20	2.02	99
3	5	3.20	90
	10	3.13	93.5
	15	3.03	98.5
	20	3.04	98

ตารางที่ 5.3 ประสิทธิภาพในการหาระยะระหว่างกล้องและวัตถุในระยะต่างๆ (ต่อ)

ระยะจริง	ระยะห่างในการติดตั้งกล้องซ้ายและกล้องขวา(เซนติเมตร)	ระยะที่วัดได้เฉลี่ย	เปอร์เซนต์ความถูกต้อง
4	5	4.22	94.5
	10	4.10	97.5
	15	4.06	98.5
	20	4.05	98.75
5	5	4.72	94.4
	10	5.16	96.8
	15	5.10	98
	20	5.08	98.4
6	5	-	-
	10	5.93	98.6
	15	6.05	99
	20	6.03	99.4
7	5	-	-
	10	-	-
	15	7.05	99
	20	6.95	99
8	5	-	-
	10	-	-
	15	-	-
	20	8	100

หมายเหตุ สัญลักษณ์ (-) หมายถึง ไม่สามารถวัดค่าได้

จากตารางที่ 5.3 แสดงประสิทธิภาพในการวัดระยะในการติดตั้งกล้องในระยะต่างกันในสภาพแวดล้อมทั้งภายในและภายนอกอาคาร จะเห็นได้ว่า เมื่อต้องการวัดระยะระหว่างกล้องและวัตถุจาก 2 – 8 เมตร การติดตั้งระยะห่างกล้องที่ดีที่สุดคือ ระยะ 20 เซนติเมตรเพราะสามารถ

ครอบคลุมในการวัดช่วงที่ต้องการได้ แต่ต้องอยู่ในกรณีที่วัตถุไม่เคลื่อนที่ชิดขอบด้านใดด้านหนึ่งมากเกินไป

จากการทดลองที่ผ่านมาพบว่าระบบสามารถทำการค้นหาและวัดระยะทางได้อย่างมีประสิทธิภาพ แต่ยังคงมีในกรณีที่ภาพมีความซับซ้อนทำให้เกิดการตรวจจับวัตถุคลาดเคลื่อน ทำให้การหาระยะมีประสิทธิภาพน้อยลง โดยเฉพาะภาพที่เกิดจากตัวอย่างที่ 4 การเคลื่อนที่ของกล้องทำให้การตรวจจับโดยการไหลเชิงแสงต้องมีภาระมากขึ้นจึงส่งผลต่อระบบโดยรวมดังตารางที่ 5.4

ตารางที่ 5.4 ประสิทธิภาพของการหาระยะระหว่างกล้องและวัตถุในสภาพแวดล้อมต่างๆ

ตัวอย่างสภาพแวดล้อม	จำนวนเฟรมทั้งหมด	จำนวนเฟรมที่ผิดพลาด	เปอร์เซนต์ความถูกต้อง
1	750	69	90.8
2	250	27	89.2
3	500	46	90.8
4	250	35	86

5.7 สรุป

จากผลการทดสอบระบบเมื่อพิจารณาตามขั้นตอนต่างๆแล้วสามารถสรุปผลได้ดังนี้

1. การคัดกรองด้วยการไหลเชิงแสง สามารถตรวจจับวัตถุที่เคลื่อนที่ได้ดีทั้งภายในและภายนอกอาคาร โดยใช้ลักษณะเด่นของการไหลเชิงแสงที่ใช้ความเข้มแสงในการวิเคราะห์กระบวนการนี้ความต่างกันของแสงภายในและภายนอกอาคารจึงไม่ส่งผลกระทบ

2. การทดสอบการจับคู่ด้วยเครือข่ายฮอพฟิลด์ 2 มิติ ถึงแม้วิธีการนี้จะเป็นวิธีการเก่าแต่ก็ยังมีประสิทธิภาพในการจับคู่ได้อย่างแม่นยำโดยอาศัยพารามิเตอร์ต่างๆ อันได้แก่ ตำแหน่งและมุมการไหลเชิงแสง

3. การหาระยะทางและการบูรณะสภาวะแวดล้อม 3 มิติ สามารถหาระยะทางได้มีประสิทธิภาพถึงแม้ในบางตัวอย่างจะเกิดความผิดพลาดซึ่งสามารถสรุปถึงสาเหตุความผิดพลาดของระบบทั้งหมดได้ดังนี้

- การป้อนอินพุตให้กับระบบเครือข่ายฮอพฟิลด์จำกัด เนื่องจากหน่วยความจำของคอมพิวเตอร์ที่ใช้ในการทดสอบครั้งนี้มีจำกัด เมื่อทำการประมวลผลในขั้นตอนของเครือข่ายประสาทเทียมสามารถป้อนอินพุตได้เพียง 100 x 100 นิวรอนหรือจำนวนนิวรอนรวมกันแล้วไม่

เกิน 200 นิรอน ซึ่งการไหลเชิงแสงของวัตถุในภาพมีมากกว่านั้น ซึ่งเมื่อทำการกรองขนาดออกไประดับหนึ่งอาจจะทำให้ข้อมูลหายไปบ้างส่วนจึงทำให้ระบบประมวลผลแม่นยำน้อยลง

- ระยะห่างระหว่างวัตถุและกล้องที่เหมาะสมที่ได้จากการคำนวณเป็นไปตามตารางที่ 5.2 ดังนั้นเมื่อวัตถุเคลื่อนที่ออกห่างจากค่าที่เหมาะสมมากๆ ถึงแม้ในขั้นตอนของการจับคู่จะประมวลผลได้แต่เมื่อนำมาคำนวณหาระยะทางแล้วจะทำให้เกิดค่าผิดพลาดมากขึ้น

- การเคลื่อนที่ของวัตถุที่อยู่ชิดขอบทางด้านซ้ายหรือขวามากเกินไป หากวัตถุเคลื่อนที่อยู่ในตำแหน่งทางด้านซ้ายของกล้องหรือขวาของกล้องมากเกินไปจะทำให้ไม่สามารถตรวจจับวัตถุได้ถึงแม้จะเข้ามาในระยะใกล้ๆก็ตาม

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 6

บทสรุปและข้อเสนอแนะ

6.1 สรุป

งานวิจัยวิทยานิพนธ์นี้ได้ดำเนินการศึกษาและพัฒนาการบูรณะสภาวะแวดล้อมสามมิติ เพื่อสามารถนำไปประยุกต์ใช้ในการตรวจจับวัดระยะของวัตถุที่เคลื่อนที่แบบอัตโนมัติ ซึ่งระบบที่พัฒนาขึ้นประกอบไปด้วยระบบการตรวจจับการเคลื่อนที่ของวัตถุโดยใช้การไหลเชิงแสง การแก้ปัญหาความสอดคล้องกันของจุดภาพของกล้องสองตัวโดยใช้เครือข่ายฮอพฟิลด์ซึ่งเป็นเครือข่ายประสาทเทียมชนิดหนึ่ง และการจำลองพิกัดสามมิติของวัตถุที่เคลื่อนที่ภายในภาพ โดยการดำเนินงานวิจัยวิทยานิพนธ์ดังกล่าวสำเร็จลุล่วงตามวัตถุประสงค์ โดยสามารถสรุปผลการศึกษาวิจัยและพัฒนาทางวิศวกรรมเป็นข้อสรุปได้ดังต่อไปนี้

การปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง ได้รายงานไว้ในบทที่ 2 ผลการสืบค้นวรรณกรรมย้อนหลังที่เกี่ยวข้องกับงานวิจัยที่จะดำเนินการจากฐานข้อมูล IEEE, IEE, ScienceDirect และอื่นๆ ซึ่งทำให้ทราบถึงแนวทางการวิจัยที่เกี่ยวข้อง ระเบียบวิธีการวิจัยที่ผู้วิจัยอื่นๆได้นำมาใช้ ผลการดำเนินงานข้อเสนอแนะต่าง ๆ ของนักวิจัยตั้งแต่อดีตจนถึงปัจจุบัน ซึ่งประกอบไปด้วยการหาแผนที่บ่งบอกความลึก และการหาระยะห่างระหว่างวัตถุกับกล้องโดยใช้กล้องสองตัว และพื้นฐานการวิเคราะห์ประมวลผลภาพ

ในบทที่ 3 แสดงรายละเอียดของการแก้ปัญหาความสอดคล้องกันของจุดภาพโดยใช้เครือข่ายประสาทฮอพฟิลด์ 2 มิติ โดยแสดงรายละเอียดโครงสร้าง การเรียนรู้และวิธีการแก้ปัญหาของเครือข่าย ซึ่งระบบเครือข่ายฮอพฟิลด์สามารถแก้ปัญหาความสอดคล้องของจุดภาพจากกล้อง 2 ตัวได้ โดยมีอินพุตที่เกิดจากตำแหน่งที่น่าสนใจเป็นตัวเริ่มกระบวรการ หลังจากนั้นเครือข่ายฮอพฟิลด์จะทำการวนรอบเพื่อหาคำตอบจนได้จุดที่สอดคล้องกันของภาพจากล้องซ้ายและกล้องขวา โดยการทดสอบครั้งนี้ใช้วิธีการของแฮริสในการหาจุดที่น่าสนใจ ซึ่งไม่เหมาะกับการตรวจจับภาพที่มีการเคลื่อนที่ของวัตถุ อินพุตของระบบที่จะใช้ในขั้นตอนต่อไปจึงเปลี่ยนเป็นการไหลเชิงแสงแทน

ในบทที่ 4 แสดงรายละเอียดโครงสร้างของระบบการบูรณะแวดล้อม 3 มิติโดยเครือข่ายประสาทเทียมร่วมกับการไหลเชิงแสง ซึ่งแสดงวิธีและขั้นตอนในการแก้ปัญหาของระบบ เริ่ม

ตั้งแต่การป้อนลำดับภาพ การประมวลผลภาพ การหาการไหลเชิงแสง การแก้ปัญหาความสอดคล้องกันของจุดภาพ จนกระทั่งการคำนวณหาพิกัด 3 มิติ ระบบการหาระยะระหว่างกล้องและวัตถุโดยใช้เทคนิคเครือข่ายประสาทเทียมร่วมกับการไหลเชิงแสง สามารถหาจุดที่สอดคล้องกันของภาพซ้ายและขวาได้โดยอาศัยตำแหน่ง มุมและขนาดของการไหลเชิงแสงเป็นอินพุตให้กับระบบ เมื่อระบบทำการประมวลผลแล้วจะทำให้ทราบพิกัดของวัตถุที่เคลื่อนที่ภายในภาพ โดยอาศัยจุดกึ่งกลางและระยะห่างของกล้องและวัตถุเป็นจุดอ้างอิงพิกัดของวัตถุที่เกิดขึ้นเมื่อเทียบกับระยะจริงแล้วมีความแม่นยำเฉลี่ยที่ 90 เปอร์เซนต์ทั้งภายในและภายนอกอาคาร ซึ่งค่าความผิดพลาดเกิดจากการป้อนอินพุตให้กับระบบเครือข่ายฮอพฟิลด์จำกัด ระยะห่างระหว่างวัตถุและกล้องที่เหมาะสมอยู่ที่ 5 - 7 เมตรตามการติดตั้งระยะห่างระหว่างกล้องสองตัว และการเคลื่อนที่ของวัตถุที่อยู่ชิดขอบทางด้านซ้ายหรือขวามากเกินไป

6.2 ข้อเสนอแนะ

จากการดำเนินงานวิจัยวิทยานิพนธ์ที่ผ่านมาและผลที่ได้ทำให้เกิดแนวคิดและข้อเสนอแนะในการดำเนินงานวิจัยต่อไปในอนาคตดังต่อไปนี้

1.เพื่อให้ระบบสามารถนำไปประยุกต์ใช้งานได้หลากหลายควรทดสอบกับสภาพแวดล้อมต่างๆ เพื่อนำไปประยุกต์ใช้ได้เหมาะสม

2.เพื่อให้ระบบมีประสิทธิภาพที่ดีขึ้นควรปรับปรุงแก้ไขขั้นตอนที่ทำให้เกิดค่าความคลาดเคลื่อน โดยปรับปรุงพารามิเตอร์ต่างๆเพื่อให้สอดคล้องกับสภาพแวดล้อมที่ทำการทดสอบและทำให้ระบบทำงานได้เร็วขึ้น เพื่อให้ระบบสามารถทำงานได้ทันกับเวลาจริง (real time)

รายการอ้างอิง

อาทิตย์ ศรีแก้ว. (2546). **ปัญญาเชิงคำนวณ COMPUTERTATIONAL INTELLIGENCE.** สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี

เดวิด บรรเจิดพงศ์ชัย. "ระบบควบคุมพลวัต: การวิเคราะห์ การออกแบบ และการประยุกต์" พิมพ์ครั้งที่ 1 กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2551.

L. MaTThies, T. Kanade, R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequence", **International Journal of Computer Vision**, vol.3 1989,209- 236

Barnard, S.T. and Fischer, M.A. "Computational Stereo". **ACM Computing Surveys** , 14(4):553- 572, 1982.

B. D. Lucas and T. Kanade (1981), "An iterative image registration technique with an application to stereo vision". **Proceedings of Imaging Understanding Workshop**, pages 121-130

C. Leung, B. Appleton, C. Sun, "Fast stereo matching by iterated dynamic programming and quadtree subregioning", **Proceedings of the British Machine Vision Conference**, 2004, pp. 97–106.

Dhond, U.R. and Aggarwal, J.K. "structure from stereo : a review". **IEEE Trans. On Systems, Man, and Cybernetics**, 19, 1489-1510, 1989

Dimitrios Tzovaras, Nikos Grammalidis and Michael G. Strintzis, "Disparity field and depth map coding for multiview 3D image generation", **Signal Processing: Image Communication II, ELSEVIER**, 1998, page 205-203

D. Scharstein, R. Szeliski, "Taxonomy and evaluation of dense two-frame stereo correspondence algorithms", **International Journal of Computer Vision** 47(2002) 7–42.

D.V. Papadimitriou, T.J. Dennis, "Epipolar line estimation and rectification for stereo image pairs", **IEEE Trans. Image Process**. 5(4): 672-676, 1996

H. Kim, S. Yang, K. Sohn, "3D Reconstruction of stereo Image for Interaction Between Real and Virtual Worlds", **Proc. 2nd IEEE and ACM International Symposium On Mixed and Augmented Reality (ISMAR'03)**, IEEE, 2003

Karim Achour and Lyes Mahiddine. "Hopfield Neural Network Based Stereo Matching Algorithm", **Journal of Mathematical Imaging and Vision** 16: 17–29, 2002

M. Montemerlo and S. Thrum. "Large-Scale Robotic 3D Mapping of Urban Structures". **In Proc. of 9th International Symposium on Exprimental Robotics**, 2004.

M. Pollefeys and L. Van Gool, "Self-Calibration and Matric Reconstruction in spite of Varying and Unknow Internal Camera Paramater", **TR9707,ESAT-MI2, K.U.Leuyen**, 1997.

Nasser M. Nasrabadi, "Hopfield Network for Stereo Vision Correspondence". **IEEE Trans. On neural networks**, Vol.3, 1992

O. Faugeras, "Three-Dimention Computer Vision : A Geometric Viewpiont", **MIT Press**., London, 2001

R. Hartley and A. Zisserman. "Multiple view geometry in computer vision", **Cambridge**, 2000

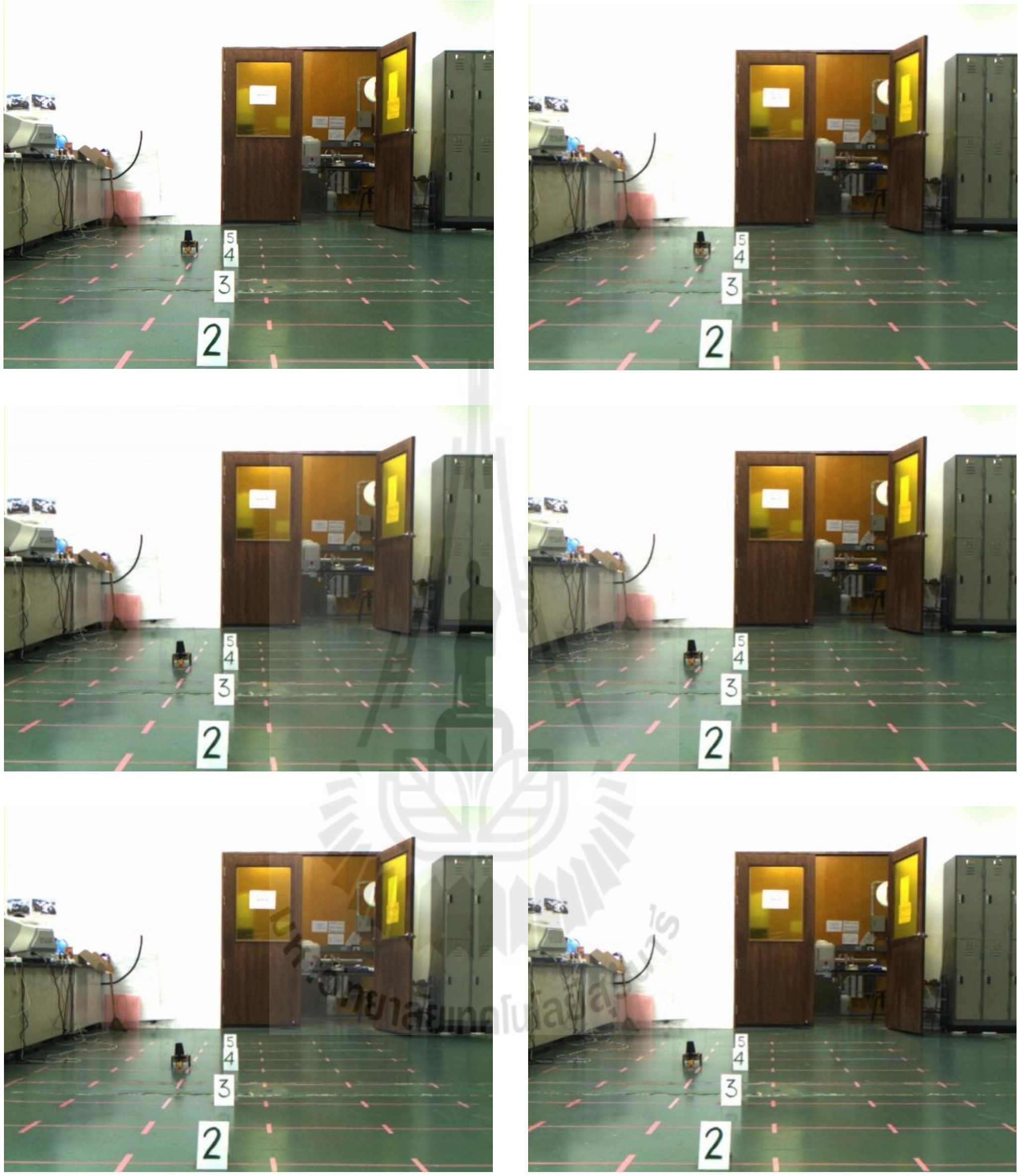
ภาคผนวก ก

ภาพต่าง ๆ ที่ใช้ในการทดสอบ

มหาวิทยาลัยเทคโนโลยีสุรนารี

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.1 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 5 เซนติเมตร



(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.1 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 5 เซนติเมตร (ต่อ)

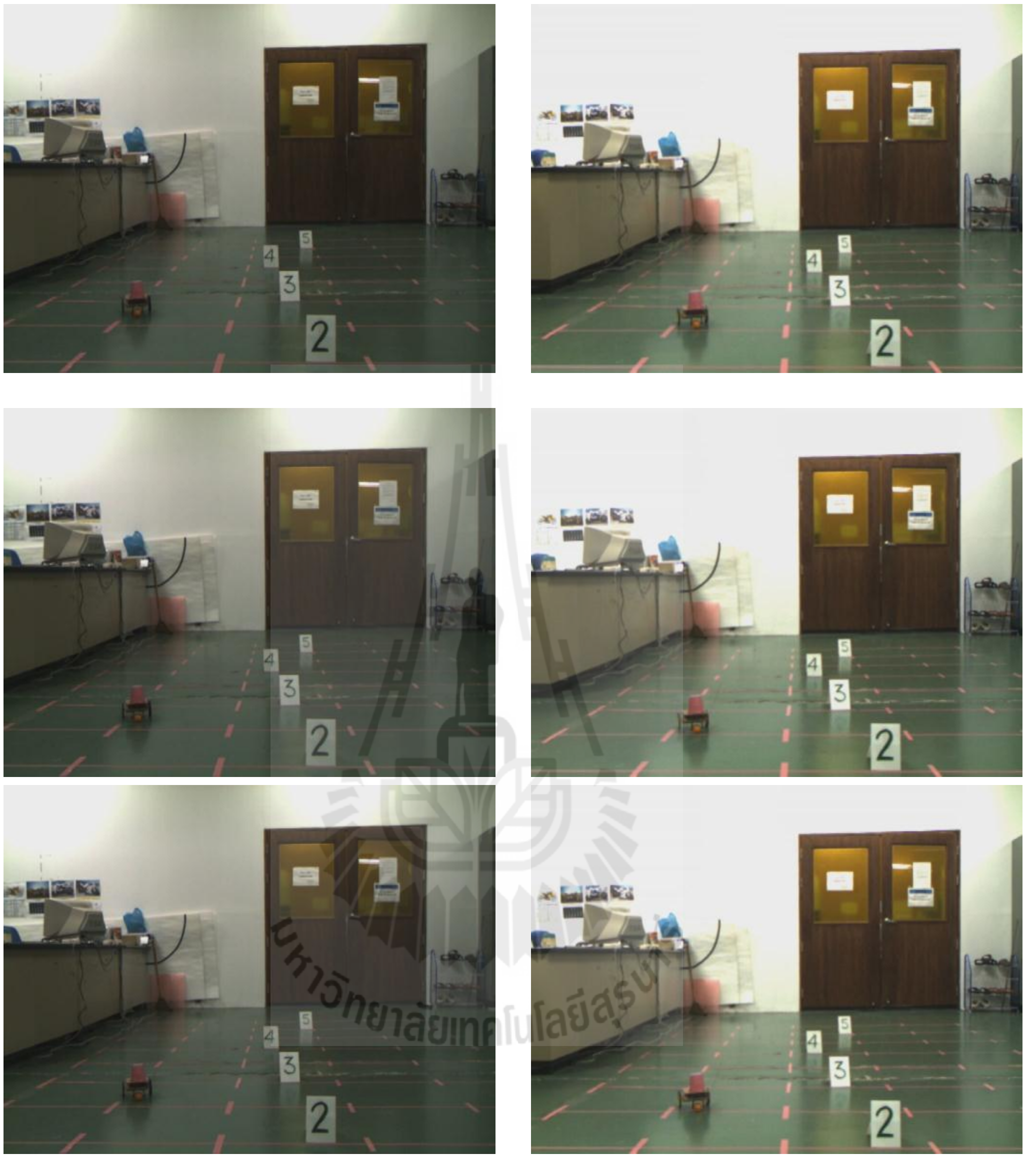
(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.2 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 10 เซนติเมตร

มหาวิทยาลัยเทคโนโลยีสุรนารี

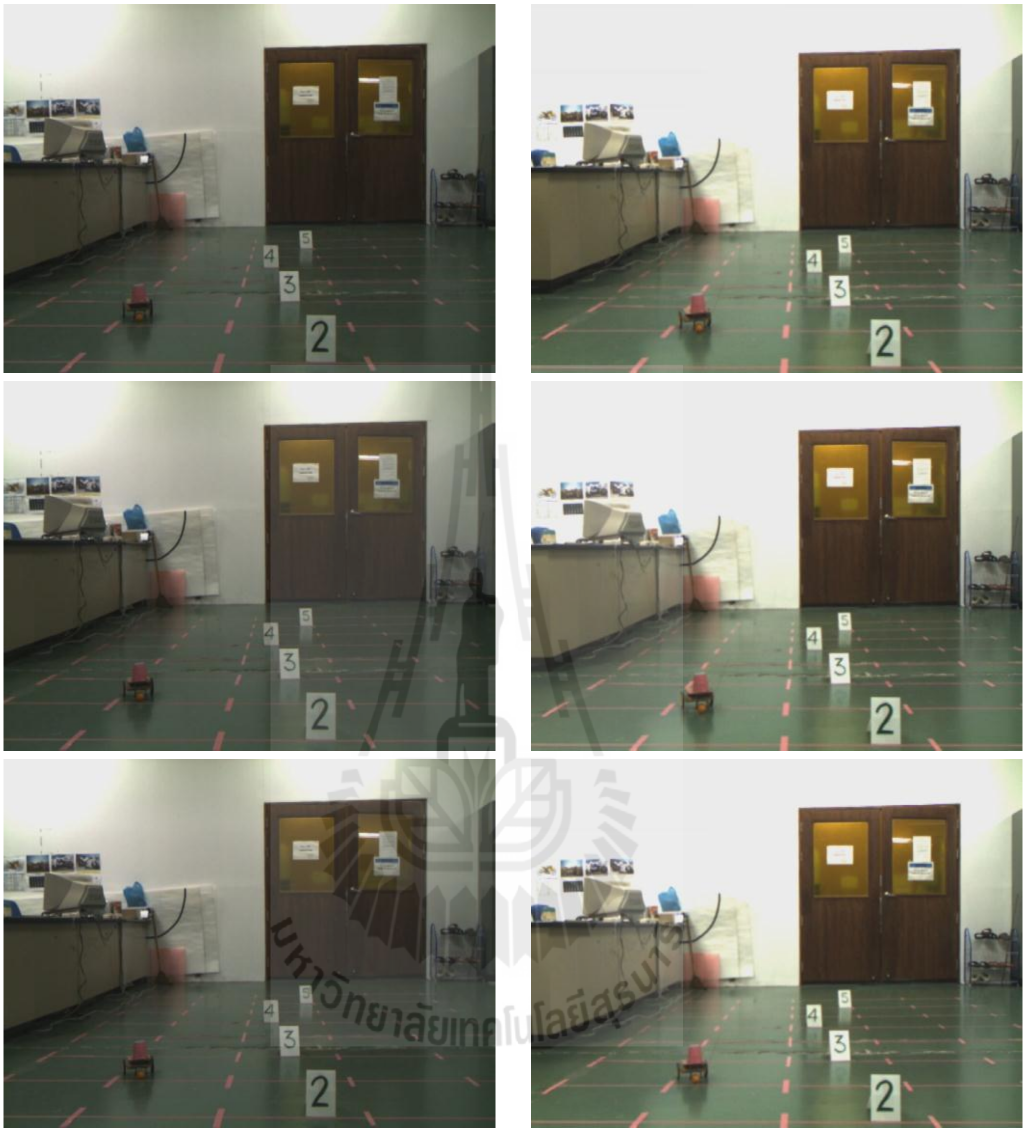
(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.2 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 10 เซนติเมตร (ต่อ)



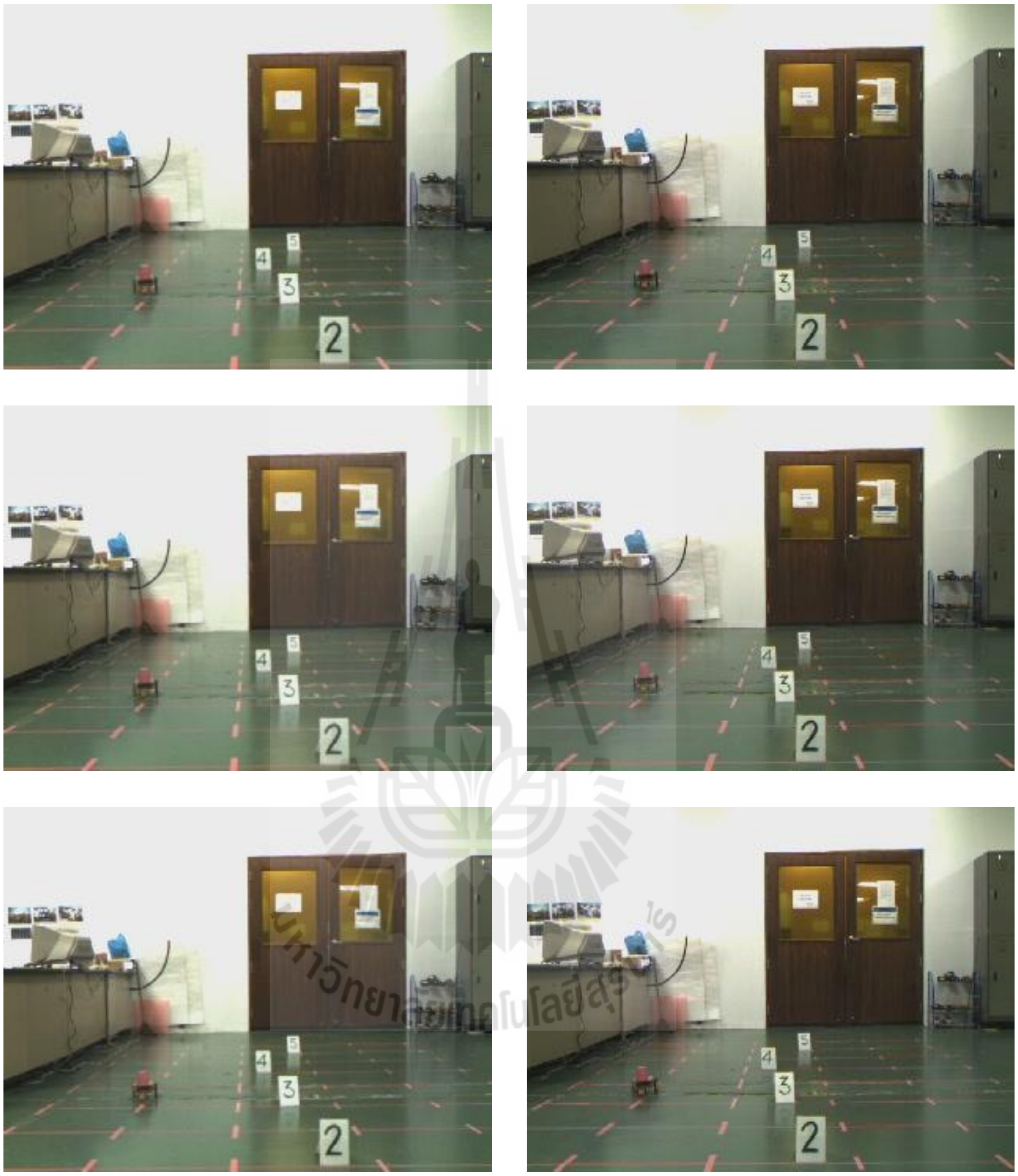
(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.3 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 15 เซนติเมตร



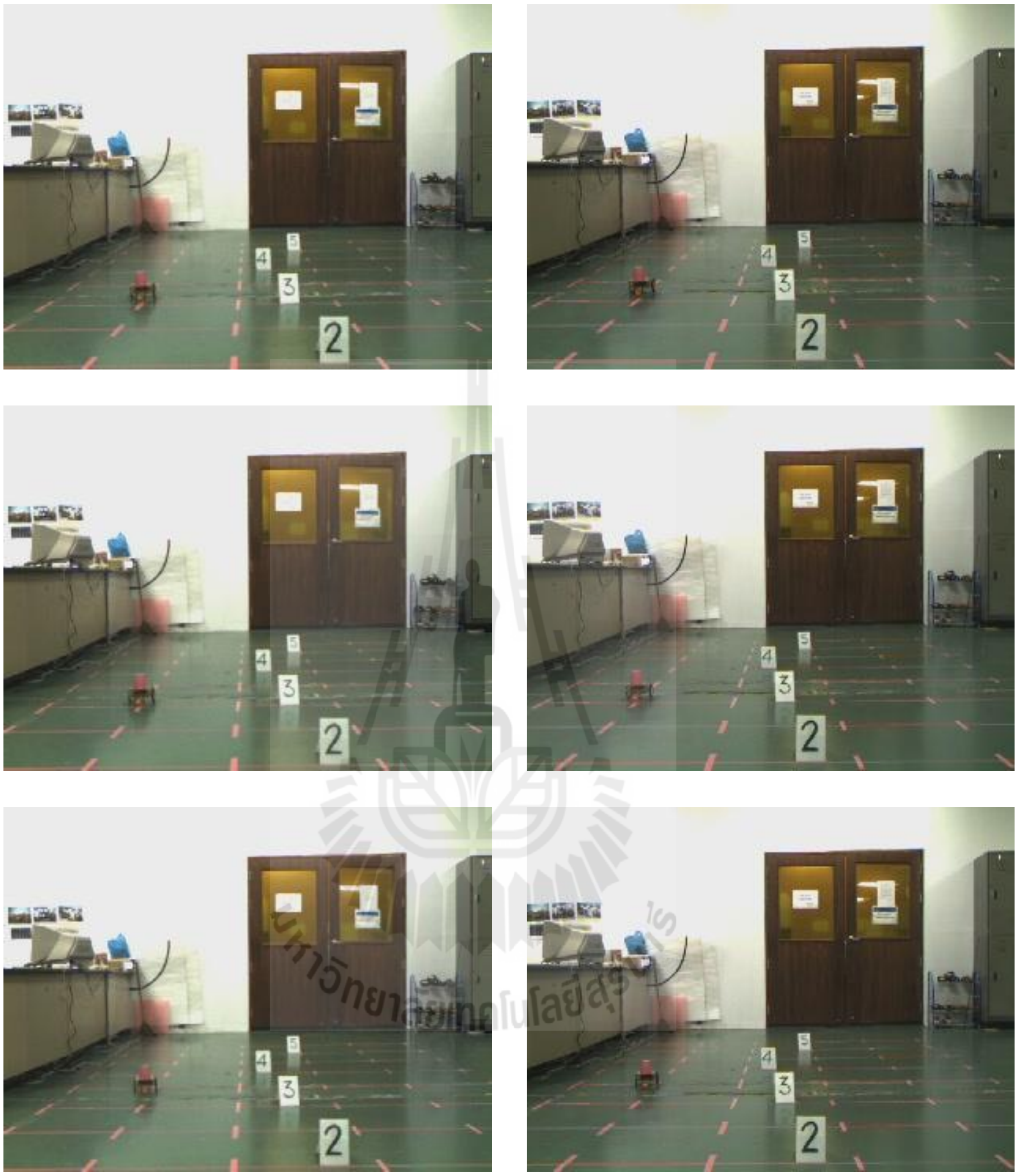
(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.3 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 15 เซนติเมตร (ต่อ)



(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.4 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร



(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.4 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร (ต่อ)

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ก.5 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร

(ก) ภาพซ้าย (ข) ภาพขวา

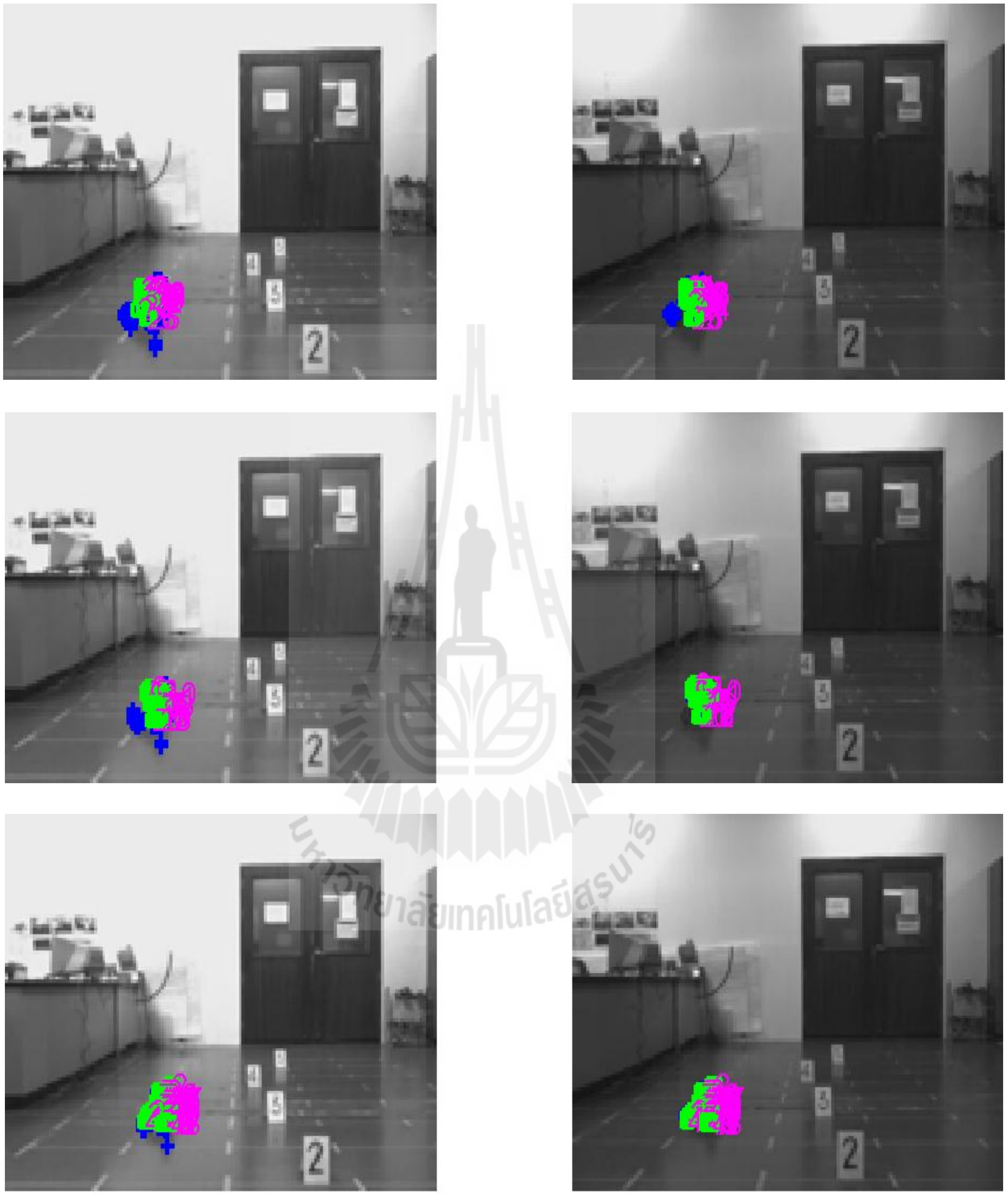
รูปที่ ก.5 ภาพต้นแบบที่ใช้ในการทดลองโดยการติดตั้งกล้องห่างกัน 20 เซนติเมตร (ต่อ)

ภาคผนวก ข

ภาพผลการทดสอบการจับคู่

ตำแหน่งของภาพจากกล้องซ้ายและขวา

มหาวิทยาลัยเทคโนโลยีสุรนารี

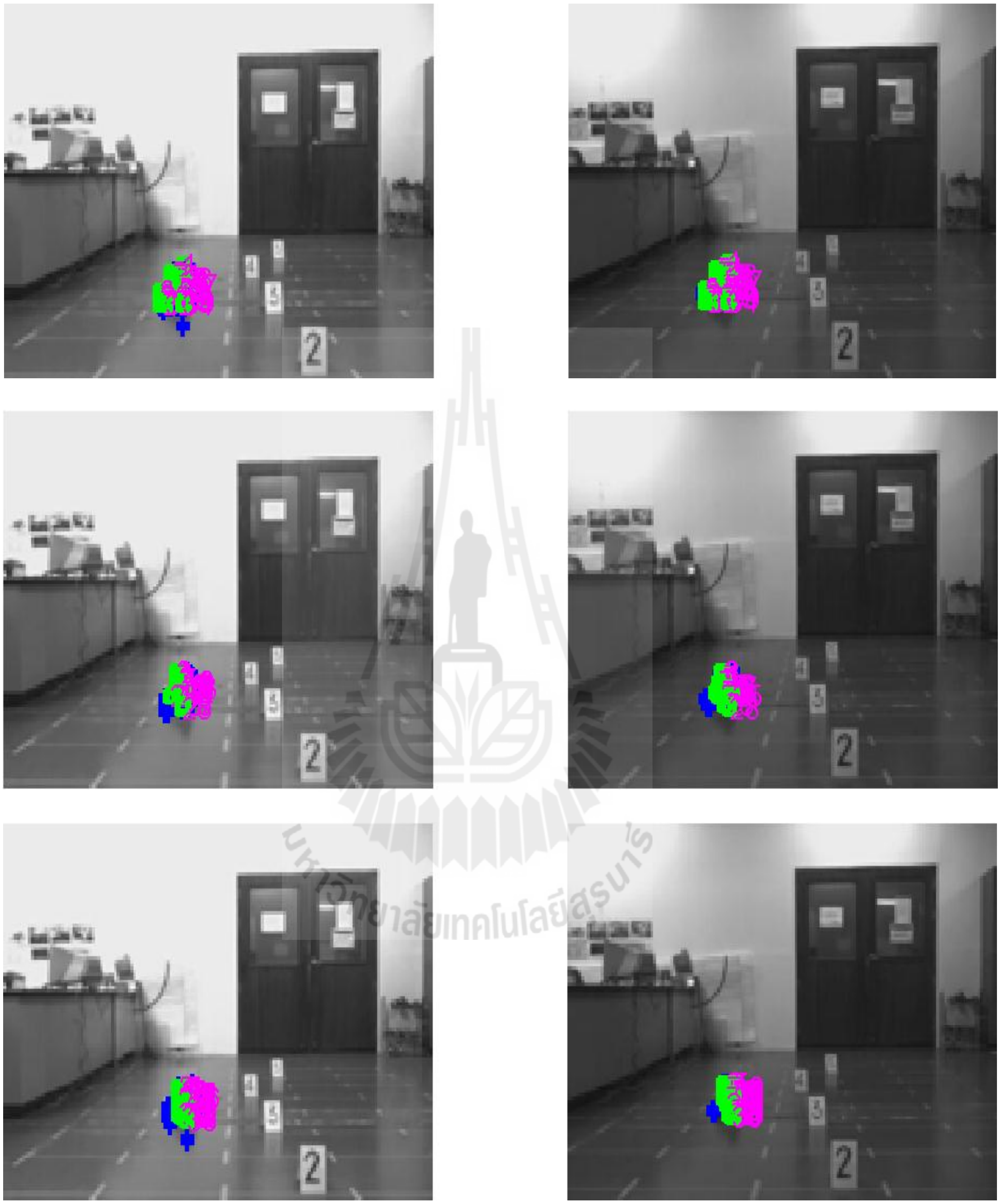


(ก) ภาพซ้าย　　　　　　　　(ข) ภาพขวา

รูปที่ ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง(ต่อ)



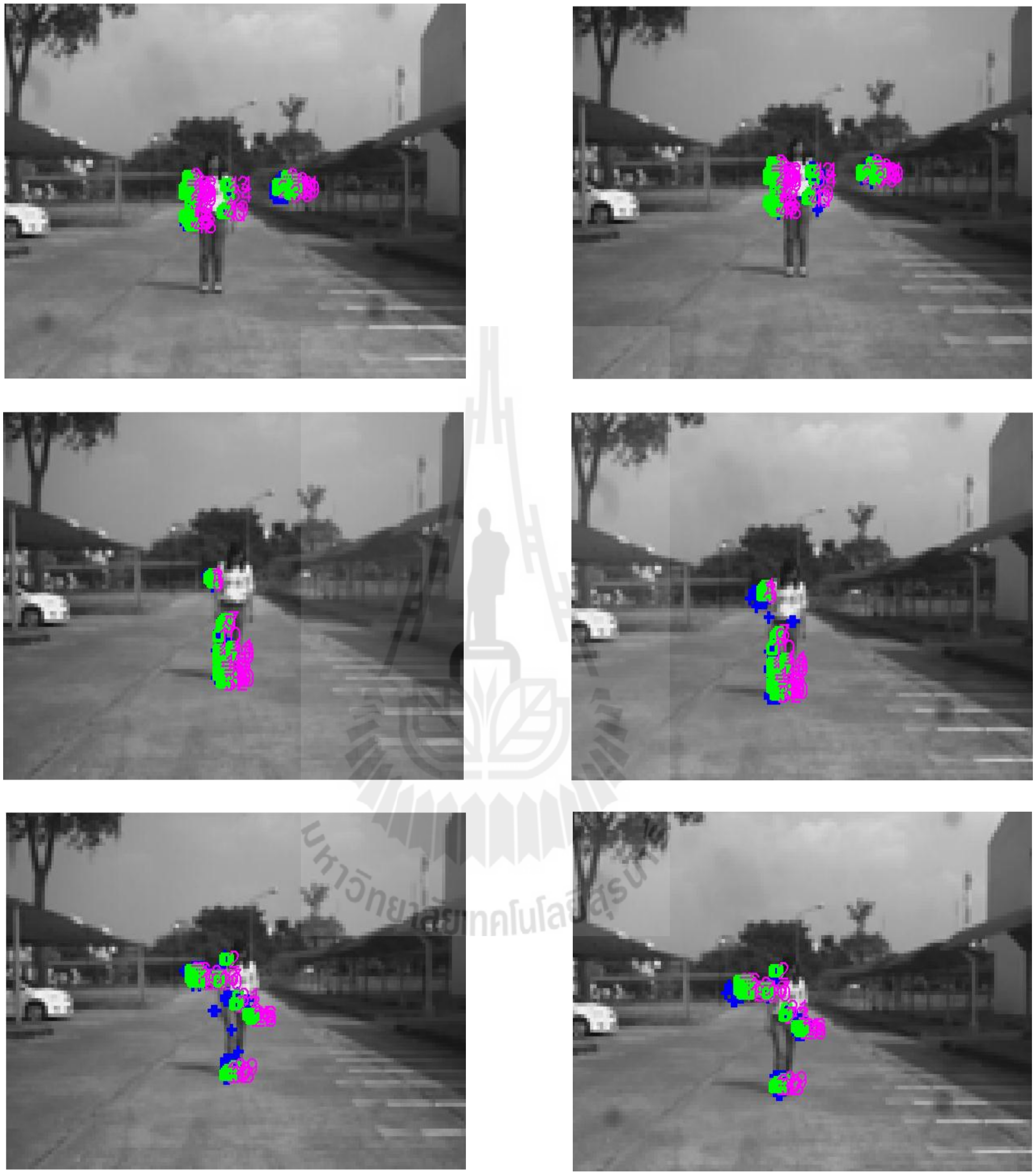
(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง(ต่อ)

มหาวิทยาลัยเทคโนโลยีสุรนารี

(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง(ต่อ)



(ก) ภาพซ้าย (ข) ภาพขวา

รูปที่ ข.1 ผลการทดลองการจับคู่ตำแหน่งด้วยเครือข่ายฮอพฟิลด์ 2 มิติร่วมกับการไหลเชิงแสง(ต่อ)

ภาคผนวก ค.

บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในขณะศึกษา

มหาวิทยาลัยเทคโนโลยีสุรนารี

รายชื่อบทความที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา

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, PP. 201-204

International Conference on Engineering, Applied Sciences, and Technology
November 21 – 24, 2012, Swissôtel Le Concorde, Bangkok, Thailand

Stereo Correspondence using 2-Dimensional Hopfield Neural Network and Optical Flow

Arthit Srikaew[1], Prayoth Kumsawat[2], Kitti Attakitmongcol[1] and Somying Phanwilai[1]
[1]School of Electrical Engineering
[2]School of Telecommunication Engineering
Institute of Engineering
Suranaree University of Technology
111 Suranaree, Muang District, Nakhonratchasima, Thailand 30000
ra@sut.ac.th, prayoth@sut.ac.th, kitti@sut.ac.th, B4810964@hotmail.com

***Abstract*— Stereo correspondence problem is one of the most important task in stereo vision system. One advantage of using such stereo vision system is the ability to calculate a distance between an object in front of two cameras mounted in parallel. In order to calculate this distance, a corresponding portion of interesting object in each image has to be detected and matched. This work presents a hybrid approach of using two-dimensional Hopfield network to perform a stereo correspondence calculation from optical flow features in stereo images. The optical flow features from both left and right image are considered as the input for the network which has an advantage of a capability to detect object motion. Using such optical flow and a 2D network with ability to learn and recognize patterns, optical flow feature in each image can be matched efficiently. The proposed system demonstrates desirable results which can be mainly and effectively used in 3D reconstruction.**

Keywords-2D Hopfield neural network, stereo vision, stereo correspondence, optical flow

I. Introduction

Computer vision becomes one of the most challenge and important part of automatic system. Stereo vision system has an advantage of ability to calculate distance between desired points and the stereo cameras. This can be considered as a distance sensor without any contact with objects of interest. Typically, there are two main parts involved with this distance calculation: disparity map and 3D reconstruction. There are many approaches proposed for stereo correspondence and disparity map calculation. In [1], image correspondence is computed at every position of left image and right image, it can match two positions but this is computationally expensive and inappropriate for complex image with high resolution of details. Previous work from [5] presents techniques to examine the coding of the depth map and disparity fields for stereo image communication applications. It is assumed that both left and right channels of the multi-view image sequence are coded using block- or object-based methods. A dynamic programming algorithm is used to estimate a disparity field between each stereo image pair. Taxonomy of dense and two-frame stereo methods were presented in [6]. The taxonomy is designed to assess the different components and design decisions made in individual stereo algorithms. Using this taxonomy, the existing stereo methods and present experiments are compared to evaluate the performance in many different variants. In [7], a fast minimization scheme was proposed in which the algorithm produces high quality stereo reconstructions for significantly reduced running time, requiring only a few seconds of computation. The minimization scheme is carried out using the iterated dynamic programming algorithm, which iterates over entire rows and columns for fast stereo matching. A quadtree subregioning process is also used for efficient computation of a matching cost volume where iterated dynamic programming operates on. From the previous work, the most challenge work is clearly to solve a stereo correspondence of the images from left and right camera. In [11], the 2D Hopfield network is developed and adapted to solve such problem. Features from both images are extracted and used as input for the network. In this paper, optical flows are extracted to use as a feature for network input. Using optical flow provides motion information and also reduces feature data. Details of optical flow and 2D Hopfield network used in this work are presented in the following section.

II. Two-Dimensional Hopfield Network

Hopfield network [11] is known as an auto-associative network consisting of fully connected neurons on a single layer. The network can also perform as an associative memory. During the network adaptation, its energy function decreases toward a local minimum which indicates a steady state convergence. In this work, a two-dimensional Hopfield network is used for finding stereo correspondence between left and right image. The network is composed of N_r x N_l neurons, where N_r and N_l are the number of features of interest from left and right image, respectively. The network architecture is shown in Fig. 1 with weights of all neuron connection.

Each neuron is fully connected to all other neurons in the network except itself. The weight of the connection between two neurons can be calculated using Equation (1) and (2) where $T_{\alpha\beta}$ represents the weight of the connection between n_α and n_β neurons.

$$T_{\alpha\beta} = T_{\beta\alpha} \quad (1)$$

$$T_{\alpha\alpha} = 0 \quad (2)$$

International Conference on Engineering, Applied Sciences, and Technology
November 21 – 24, 2012, Swissôtel Le Concorde, Bangkok, Thailand

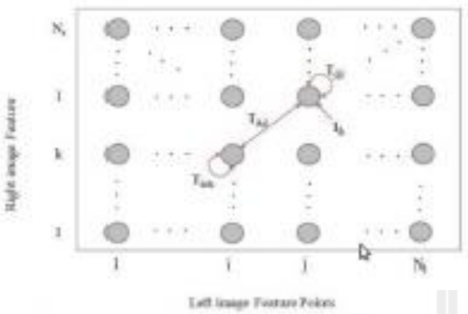


Figure 1. Two-dimensional Hopfield network architecture.

In (1) presents a weight symmetry between n_{ik} and n_{jl} neuron. There is no connection between neuron and itself as shown in (2) where the weight connection is equal zero. The Lyapounov energy function of 2D Hopfield network is given in (3) where V_{ik} and V_{jl} are the state of neuron *ik* and *jl*, respectively, and the I_{ik} is an external input value of every neuron. Each of every neuron of the network can only be either active or inactive (0 or 1) state.

$$E=-\frac{1}{2}\sum_{i=1}^{N_l}\sum_{k=1}^{N_r}\sum_{j=1}^{N_l}\sum_{l=1}^{N_r} T_{ikjl}V_{ik}V_{jl}-\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} I_{ik}V_{ik} \quad (3)$$

Previous work from [11] suggested that in order to minimize the network energy function, the state of each neuron can be updated upon the following (4) – (6) constraints. These conditions allow the network to change its state and the optimal solution is achieved when the network state does not change indicating the lowest level of its energy.

$$V_{ik}=0 \text{ if } \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl}+I_{ik}\right]>0 \quad (4)$$

$$V_{ik}=1 \text{ if } \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl}+I_{ik}\right]<0 \quad (5)$$

$$\text{no chang if } \left[\sum_{i=1}^{N_l}\sum_{k=1}^{N_r} T_{ikjl}V_{jl}+I_{ik}\right]=0 \quad (6)$$

In (4) – (6), state of neuron *ik* is changed when sum of weight of neuron *jl* and its I_{ik} changes. If the summation is greater than zero, neuron *ik* then changes its state to zero. If the summation is less than zero, neuron *ik* then changes its state to 1. If the summation is equal zero, the state of neuron *ik* then remains unchanged. For stereo correspondence problem, each feature position uses energy value from (7) in which the lowest or unchanged energy indicates the possible matching of the feature from both images. From [11], the weight $T_{ikjl} = C_{ikjl} - \delta_{ij} - \delta_{kl}$ where $\delta_{ij} = 1$ if $i = j$ and $\delta_{kl} = 1$ if $k = l$. C_{ikjl} is given in the following equation where $X = [W_1|\Delta d|+W_2|\Delta D|]$ with $\lambda = 1$ and $\theta = 10$ [11].

$$E=-\left(\frac{1}{2}\right)\sum_{i=1}^{N_l}\sum_{k=1}^{N_r}\sum_{j=1}^{N_l}\sum_{l=1}^{N_r}(C_{ikjl}-\delta_{ij}-\delta_{lk})P_{ik}P_{jl}-\sum_{i=1}^{N_l}\sum_{k=1}^{N_r}2P_{ik} \quad (7)$$

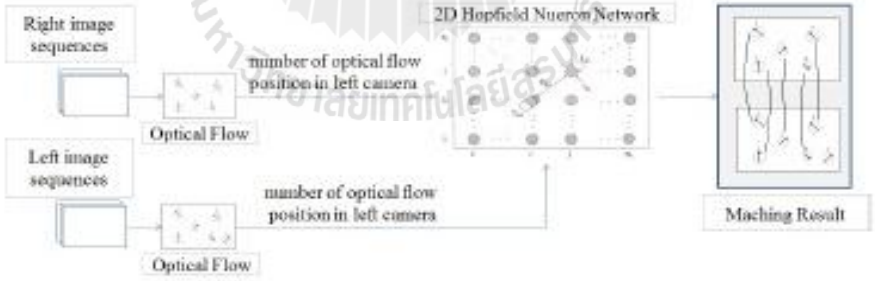


Figure 2. Stereo correspondence calculation diagram using 2D Hopfield network and optical flow features.

III. Stereo Correspondence Calculation using 2D Hopfield Network and Optical Flow

In this paper, optical flow is chosen to be a feature from left and right image. Lucas-Kanade method is used to compute optical flow [14]. The object of interest, hence, must be moving during the process of calculation. The stereo correspondence computation starts from computing optical flow of frame *n* and *n*-1 from each left and right camera.

Figure 3. example of typical optical flow

Fig. 3 shows an example of typical optical flow which clearly has too many optical flow occurrences within each image sequence. In order for the network to be able to process efficiently, the optical flow is filtered to have a suitable range of amplitude with less number of optical flows left in the image.

Figure 4. Computation of ΔD and Δd.

The next step is to calculate the weight of the network by computing ΔD and Δd of every optical flow position. ΔD is a difference of a distance between position *i* and *j* and a distance between position *k* and *l* while Δd is a difference of a distance between position *i* and *k* and a distance of position *j* and *l* as can be seen in Fig.4. This computation clearly shows that the matching optical flow position yields a very small value. The results are used in (7) with $W_1 = 0.6$ and $W_2 = 0.4$ where W_1 and W_2 are constants called difference measure from [11].

IV. Experimental Results and Discussions

The proposed stereo correspondence computation using 2D Hopfield network with optical flow feature provides desirable results. Figure 5 shows example results of optical flow feature matching. The number indicates each feature correspondence from both images. From this example, there are total of 40 matches found including 4 incorrect matches at number 23, 25, 27, and 30 (approximately 10 percent of error). The feature matching is corrected by human inspection in order to measure the performance of the proposed algorithm.

(a)

(b)

Figure 5. (a) Optical flows are computed and filtered to reduce number of optical flows (b) Example results of optical flow feature matching by using 2D Hopfield network.

Table 1 displays the overall performance of the stereo correspondence computation and the error from 6 stereo image samples using the proposed algorithm. The average error is acceptable with percentage of 7.33. This error mainly occurs at the optical flow positions which are very close to each other within the image hence very difficult to match.

TABLE I. EXAMPLE OF OPTICAL FLOW FEATURE MATCHING AND ITS ERROR

Example	Number of optical flow position		Number of match	Correct	Error (%)
	Left camera	*Right camera*			
1	13	12	3	3	0
2	18	31	7	7	0
3	30	45	15	13	13
4	68	87	13	12	7
5	82	113	40	36	10
6	96	85	41	35	14

Fig. 6 and 7 demonstrate example results of stereo correspondence computation using images from stereo camera having distance of 5 and 10 cm between cameras, respectively.

Figure 6. Example results of optical flow feature matching from stereo camera with 5-cm distance between both camera.

Figure 7. Example results of optical flow feature matching from stereo camera with 10-cm distance between both camera.

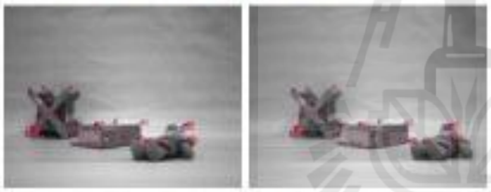


Figure 8. Example results of optical flow feature matching from stereo camera setup in indoor.

Figure 9. Example results of optical flow feature matching from stereo camera setup in indoor.

Fig. 8 and 9 present other example results of optical flow feature matching from indoor stereo camera system. The matching results rely heavily on the optical flow calculation. Higher resolution of the image can be used to achieve more details of optical flow which can be deployed for a 3D reconstruction effectively.

V. Conclusion

Stereo correspondence problem can be efficiently solved by using two-dimensional Hopfield network. In this work, the network is adapted to receive input features from optical flow which provide useful information of the image in many aspects. Results are desirable at the matching correction of 90 percent which can be effectively used in 3D reconstruction from stereo images. 2D Hopfield network greatly perform the matching and can be improved by optimizing the network parameters in the future work.

References

[1] L. MaTThies, T. Kanade, R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequence", International Journal of Computer Vision, vol.3 1989,209- 236

[2] Barnard, S.T. and Fischer, M.A. "Computational Stereo". ACM Computing Surveys 14(4):553- 572, 1982.

[3] D.V. Papadimitriou, T.J. Dennis, "Epipolar line estimation and rectification for stereo image pairs", IEEE Trans. Image Process. 5(4): 672-676, 1996

[4] Dhond, U.R. and Aggarwal, J.K. "structure from stereo : a review". IEEE Trans. On Systems, Man, and Cybernetics, 19, 1489-1510, 1989

[5] Dimitrios Tzovaras, Nikos Grammalidis and Michael G. Strintzis, "Disparity field and depth map coding for multiview 3D image generation", Signal Processing: Image Communication II, ELSEVIER, 1998, page 205-203

[6] D. Scharstein, R. Szeliski, Taxonomy and evaluation of dense two-frame stereo correspondence algorithms, International Journal of Computer Vision 47(2002) 7–42.

[7] C. Leung, B. Appleton, C. Sun, Fast stereo matching by iterated dynamic programming and quadtree subregioning, in: Proceedings of the British Machine Vision Conference, 2004, pp. 97–106.

[8] H. Kim, S. Yang, K. Sohn, "3D Reconstruction of stereo Image for Interaction Between Real and Virtual Worlds", Proc. 2nd IEEE and ACM International Symposium On Mixed and Augmented Reality (ISMAR'03), IEEE, 2003

[9] M. Montemerlo and S. Thrun. "Large-Scale Robotic 3D Mapping of Urban Structures". In Proc. of 9th International Symposium on Exprimental Robotics, 2004.

[10] M. Pollefeys and L. Van Gool, "Self-Calibration and Matric Reconstruction in spite of Varying and Unknow Internal Camera Parameter", TR9707,ESAT-MI2, K.U.Leuven, 1997.

[11] Nasser M. Nasrabadi, "Hopfield Network for Stereo Vision Correspondence". IEEE Trans. On neural networks, Vol.3, 1992

[12] O. Faugeras, "Three-Dimention Computer Vision : A Geometric Viewpiont", MIT Press., London, 2001

[13] R. Hartley and A. Zisserman. "Multiple view geometry in computer vision", Cambridge, 2000

[14] B. D. Lucas and T. Kanade (1981), "An iterative image registration technique with an application to stereo vision". Proceedings of Imaging Understanding Workshop, pages 121-130

ประวัติผู้เขียน

นางสาวสมหญิง พันธุ์วิไล เกิดวันที่28 กันยายน พ.ศ. 2529 ที่อำเภอเทิง จังหวัดเชียงราย เริ่มการศึกษาในระดับประถมศึกษาที่โรงเรียนวัดพระเกิดคงคาราม จังหวัดเชียงราย ระดับมัธยมศึกษาที่โรงเรียนเทิงวิทยาคม จังหวัดเชียงราย และสำเร็จการศึกษาระดับปริญญาตรีวิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้า) สำนักวิชาวิศวกรรมศาสตร์ จากมหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา เมื่อปีการศึกษา พ.ศ. 2551 หลังสำเร็จการศึกษา ได้ศึกษาต่อในระดับปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีสุรนารี

ในขณะศึกษาได้ทำหน้าที่เป็นผู้ช่วยสอนและอบรมในค่ายอิเล็กทรอนิกส์รุ่นเยาว์ (e-camp) และมีผลงานทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในขณะศึกษา ดังรายชื่อที่ปรากฏในภาคผนวก ค. ทั้งนี้ผู้เขียนมีความสนใจงานทางด้านการประมวลผลสัญญาณภาพ การมองเห็นของคอมพิวเตอร์และระบบเคลื่อนที่อัตโนมัติ

มหาวิทยาลัยเทคโนโลยีสุรนารี