

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น

USING THE DOMINANT POINTS TO RECOGNIZE AND LOCATE

PARTIALLY OCCLUDED OBJECTS



นายพารา ลิ้มมะณีประเสริฐ
MR.PARA LIMMANEEPRASERTH

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เลขหมู่.....

เลขทะเบียน.....28906

วัน, เดือน, ปี.....1 พ.ย. 2540

พ.ศ. 2540

ISBN 974-621-850-6

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของสำนักหอสมุดกลางพระจอมเกล้าลาดกระบัง การนำเอกสารนี้ไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

USING THE DOMINANT POINTS TO RECOGNIZE AND LOCATE
PARTIALLY OCCLUDED OBJECTS



A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE
MASTER OF ENGINEERING (ELECTRICAL ENGINEERING)
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

1997

ISBN 974-621-850-6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วน โดยการใช้จุดเด่น
นักศึกษา	นายพารา ลิ้มมะณีประเสริฐ
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ผศ. เกษตร์ ศิริสันติสัมฤทธิ์
ระดับการศึกษา	วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า
ภาควิชา	วิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยี พระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.	2540

บทคัดย่อ

การจดจำภาพวัตถุเป็นส่วนสำคัญส่วนหนึ่งในระบบการมองเห็นของหุ่นยนต์ ซึ่งในวิทยานิพนธ์ฉบับนี้ จะได้นำเสนอหลักการ การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น ในการประมวลผลการจดจำภาพวัตถุสามารถแบ่งการทำงานได้เป็นสองขั้นตอน คือ ขั้นตอนหนึ่งจะเป็นส่วนที่ทำหน้าที่คำนวณหาลักษณะเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุออกมา ซึ่งลักษณะเด่นเหล่านี้มักจะเกิดขึ้นกับจุดที่มีค่าความโค้งสูงและเราเรียกจุดเหล่านี้ว่า "จุดเด่น" เนื่องจากจุดเด่นเหล่านี้เป็นข้อมูลที่มีความสำคัญและมีคุณสมบัติเพียงพอที่จะเป็นตัวบ่งบอกถึงคุณลักษณะของวัตถุได้ ดังนั้น ในขั้นตอนที่สองได้นำเสนอวิธีการคำนวณหาจุดเด่นของวัตถุออกมา โดยการใช้วิธีการตรวจสอบหาจุดเด่นด้วยเทคนิคการเปลี่ยนค่า สเกล-สเปซของตัวกรองเกาส์เซียน และตรึงจุดเด่นที่ได้นี้ให้อยู่ในตำแหน่งที่ถูกต้องตลอดเวลาในการเปลี่ยนค่า สเกล-สเปซ แต่ละค่า จากวิธีการดังกล่าวจะทำให้จุดเด่นที่คำนวณได้เป็นจุดเด่นที่เสถียรภาพ แม้จะมีการเปลี่ยนแปลงค่า สเกล-สเปซ ของตัวกรองเกาส์เซียนไปก็ตาม ส่วนในขั้นตอนที่สองนั้น จะเป็นส่วนที่ทำหน้าที่ประมวลผลการจับคู่ของจุดเด่น ซึ่งจะเป็นการจับคู่ระหว่างจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุที่ถูกบดบังบางส่วนกับจุดเด่นของภาพวัตถุต้นแบบที่เก็บเอาไว้ในฐานข้อมูล ในส่วนของการประมวลผลการจับคู่นี้ ได้นำเอาทฤษฎีโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์มาประยุกต์ใช้งาน เพื่อให้ได้ประสิทธิภาพสูงสุดของการประมวลผลการจับคู่ ซึ่งจะทำให้ภาพวัตถุที่ถูกบดบังบางส่วนสามารถที่จะจับคู่กับภาพวัตถุที่ถูกเก็บเอาไว้ในฐานข้อมูลได้อย่างแม่นยำ ถึงแม้ว่าภาพวัตถุนั้นจะถูกบดบังค่อนข้างมากและมีความซับซ้อนก็ตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title Using the Dominant Points to Recognize and Locate Partially Occluded Objects

Student Mr. Para Limmaneeprasert

Thesis Advisor Assistant Professor Kaset Sirisantisamrid

Level of Study Master of Engineering (Electrical Engineering)

Department Computer Engineering, Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang

Year 1997

Abstract

Object recognition is an essential part of computer vision system. In this thesis, we present the principle of object recognition by using the dominant points to recognize and locate partially occluded objects. Generally, object recognition processing can be subdivided into two major tasks: The first task is to extract the features of the object contour in a scene containing the high curvature points. These high curvature points are called the "dominant points". The dominant points along an object contour are rich in information content and are sufficient to characterize the shape of the object. Therefore, in this task we propose the extraction of the features by using the detection of dominant points with various scale-space of Gaussian filter. They, then, are traced back into the right location. From this method the dominant points do not change. They are stable when the scale-space of Gaussian filter is varied. In the second task of the object recognition. The dominant points are matched between the dominant points of a scene, consisting of partially occluded objects, and the set of models which are predefined in the model database. The matching task is implemented by Hopfield neural network. By this principle, the partially occluded objects can be exactly matched with the set of models in the database, though, the objects in a scene are heavily obstructed and complicated.

กิตติกรรมประกาศ

การจัดทำวิทยานิพนธ์ฉบับนี้ได้สำเร็จลุล่วงได้ด้วยดี เพราะได้รับคำปรึกษาและแนวทางในการดำเนินการศึกษาและวิจัยจาก ผู้ช่วยศาสตราจารย์ เกษตร์ ศิริสันติสัมฤทธิ์ ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ผู้ดำเนินการจัดทำวิทยานิพนธ์ จึงขอกราบขอบพระคุณท่านอาจารย์เป็นอย่างสูง

ขอกราบขอบพระคุณบิดา มารดา ที่เป็นผู้สร้างกำลังใจในการศึกษาเล่าเรียน และเป็นผู้ให้ทุกสิ่งทุกอย่าง ขอขอบพระคุณ พี่ ๆ ร่วมสายโลหิตที่ให้การสนับสนุนการศึกษา ตลอดเงินทุนในการศึกษาโดยตลอด

ขอขอบคุณ เพื่อน ๆ ที่มีส่วนร่วมในการให้ข้อมูลและแนะนำในการศึกษาและวิจัย

ท้ายที่สุดนี้ ขอขอบคุณผู้ให้ทุนอุดหนุนการทำวิทยานิพนธ์ บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พารา ลิมมะณีประเสริฐ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญภาพ	VIII
บทที่	
1. บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
จุดมุ่งหมายและวัตถุประสงค์	3
สมมุติฐานของการศึกษา	3
ขอบเขตของการศึกษา	4
รายละเอียดของวิทยานิพนธ์	4
2. การคำนวณหาจุดเด่นของเส้นขอบภาพ	5
บทนำ	5
คณิตศาสตร์ของตัวกรองเกาส์เซียน	5
1. คุณสมบัติของตัวกรองเกาส์เซียน	5
2. สมการสำหรับหาค่าความโค้ง	9
การประมวลผลเพื่อหาจุดเด่นของเส้นขอบภาพ	10
1. การประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน	10
2. การคำนวณหาค่าความโค้งด้วยการเปลี่ยนค่า สเกล-สเปซ ของตัวกรองเกาส์เซียน	14
3. การคำนวณหาค่าความโค้งด้วยตัวกรองเกาส์เซียนและคิวบิก บี-สไปร์น	18
4. สเกล-สเปซ อิมเมจ	30
5. การตรึงจุดเด่นให้คงที่โดยใช้ซอฟต์แวร์	31
สรุปและเปรียบเทียบผลการคำนวณหาจุดเด่นของแต่ละวิธี	35

สารบัญ (ต่อ)

บทที่	หน้า
3. โครงข่ายเซลล์ประสาทเทียม	37
บทนำ	37
แบบจำลองของเซลล์ประสาท	39
โครงข่ายเซลล์ประสาทเทียมแบบต่าง ๆ	43
1. โครงข่ายเซลล์ประสาทเทียมแบบเพอร์เซพตรอน (Perceptron)	43
2. โครงข่ายเซลล์ประสาทเทียมแบบแพร่กลับ (Back Propagation)	47
3. โครงข่ายเซลล์ประสาทเทียมแบบ Self-Organizing Map	50
4. โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ (Hopfield Neural Network)	52
4. การจดจำภาพวัตถุ	56
บทนำ	56
การประมวลผลการจดจำภาพวัตถุ	57
1. การคำนวณหาจุดเด่นของเส้นขอบภาพ	58
2. การประมวลผลการจับคู่ของจุดเด่นเพื่อจดจำภาพวัตถุ	63
การคำนวณตำแหน่งของวัตถุ	73
1. การคำนวณหาจุดศูนย์กลางของรูปหลายเหลี่ยม	73
2. การแปลงเส้นขอบภาพในระบบพิกัดโฮโมจีเนียสสำหรับภาพ 2 มิติ	74
การคำนวณหาค่าความผิดพลาดในการจับคู่ของจุดเด่น	79
สรุปปัจจัยที่ต้องพิจารณาเลือกพารามิเตอร์ในการประมวลผลการจดจำภาพวัตถุ	82
5. ผลการทดลอง	83
อุปกรณ์ในการทดลอง	83
ขั้นตอนการทดลอง	83
ตัวอย่างผลการทดลอง	86
ปัญหาที่เกิดขึ้นเมื่อเลือกใช้พารามิเตอร์ไม่เหมาะสม	113
สรุปผลการทดลอง	115
6. บทสรุป	116
สรุปผลการศึกษาและวิจัย	116
สรุปปัญหาที่เกิดขึ้นในการศึกษาและวิจัย	117
บรรณานุกรม	118
ภาคผนวก	120

สารบัญ (ต่อ)

	หน้า
ภาคผนวก ก.	121
ภาคผนวก ข.	126
ภาคผนวก ค.	136
ประวัติผู้เขียน	158



สารบัญตาราง

ตารางที่	หน้า
1. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น	80
2. ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย โครงข่ายเซลล์ประสาทเทียม	86
3. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 1)	89
4. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 2)	93
5. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 3)	96
6. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 4)	100
7. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 5)	104
8. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 6)	108
9. ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 7)	112
10. ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย โครงข่ายเซลล์ประสาทเทียม	113
11. ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย โครงข่ายเซลล์ประสาทเทียม	114

สารบัญภาพ

	หน้า
1. ภาพแสดงเส้นกราฟของตัวกรองเกาส์เซียน	6
2. ภาพแสดงคุณสมบัติของตัวกรองเกาส์เซียนในการกรองเส้นขอบภาพ ที่ค่า สเกล-สเปซ ต่าง ๆ กัน	8
3. ภาพแสดงการเก็บข้อมูลของเส้นขอบภาพ	11
4. ภาพแสดงขนาดหน้าต่างสำหรับเก็บค่าสัมประสิทธิ์ของตัวกรองเกาส์เซียน ที่ค่า สเกล-สเปซ ต่าง ๆ กัน	12
5. ภาพแสดงวิธีการประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน	13
6. ภาพต้นแบบรูปหลายเหลี่ยม	15
7. ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 3$ โดยใช้อัลกอริทึมที่หนึ่ง	16
8. ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 19$ โดยใช้อัลกอริทึมที่หนึ่ง	17
9. ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 3$ โดยใช้อัลกอริทึมที่สอง	21
10. ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 19$ โดยใช้อัลกอริทึมที่สอง	22
11. ภาพแสดงภาพต้นแบบภาพคิม	23
12. ภาพแสดงเส้นกราฟของค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง	24
13. ภาพแสดง สเกล-สเปซ อิมแมจ	30
14. ภาพ สเกล-สเปซ อิมแมจ แสดงการเลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่หนึ่ง	31
15. ภาพ สเกล-สเปซ อิมแมจ แสดงการเลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่สอง	31
16. ภาพแสดงวิธีการตรึงจุดเด่นให้คงที่โดยอาศัย สเกล-สเปซ อิมแมจ	33

สารบัญภาพ (ต่อ)

หน้า

17. ภาพ สเกล-สเปซ อิมเมจ แสดงการเลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่สาม	35
18. ภาพแสดงการแบ่งประเภทโครงข่ายเซลล์ประสาทเทียมตามกฎการเรียนรู้และเรขาคณิต	38
19. ภาพแสดงรูปแบบการเรียนรู้โดยมีครู	38
20. ภาพแสดงรูปแบบการเรียนรู้โดยไม่มีครู	39
21. ภาพแสดงแบบจำลองของเซลล์ประสาท	40
22. ภาพแสดงฟังก์ชันถ่ายโอนแบบต่าง ๆ	42
23. ภาพแสดงหลักการทำงานของเพอร์เซพตรอน	44
24. ภาพแสดงการจำแนกประเภทข้อมูลเพื่ออธิบายการลู่เข้าของเพอร์เซพตรอน	46
25. ภาพแสดงโครงข่ายแบบมีชั้นซ่อนสองชั้น	47
26. ภาพแสดงโครงสร้างของ Self-Organizing Map	50
27. ภาพแสดงโครงสร้างของโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์	52
28. ภาพแสดงขั้นตอนการประมวลผลการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วน โดยการใช้จุดเด่น	57
29. ภาพแสดงภาพต้นแบบระดับสีเทาที่ได้จากกล้อง CCD	59
30. ภาพแสดงเส้นขอบภาพที่ได้จากภาพที่ 29	60
31. ภาพแสดงจุดเด่นของเส้นขอบภาพ	61
32. ภาพแสดงภาพวัตถุที่วางทับซ้อนกันบางส่วน	62
33. ภาพแสดงความสัมพันธ์ระหว่างรูปสามเหลี่ยมสองรูป	64
34. ภาพแสดงค่าความโค้งและทิศทางความโค้งของจุดเด่นรูปหลายเหลี่ยม	66
35. ภาพแสดงการคำนวณหามุมของจุดเด่นด้วยเวกเตอร์	66
36. ภาพแสดงลำดับขั้นตอนการลู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม	72
37. ภาพแสดงรูปหลายเหลี่ยมที่เกิดจากการเชื่อมต่อกันของจุดเด่น	73
38. ภาพแสดงการแปลงและการแปลงผกผันสำหรับการเลื่อนเส้นขอบภาพ	75
39. ภาพแสดงการแปลงและการแปลงผกผันสำหรับการหมุนเส้นขอบภาพ	76
40. ภาพแสดงการแปลงสำหรับการย่อ/ขยายเส้นขอบภาพ	77
41. ภาพแสดงผลของการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น	80

สารบัญภาพ (ต่อ)

หน้า

42. ภาพแสดงพฤติกรรมของสมการหาค่าความผิดพลาดในการจับคู่ของจุดเด่น (สมการที่ (4-30))	81
43. ภาพแสดงการติดตั้งอุปกรณ์ในระบบการจดจำภาพวัตถุ	83
44. ภาพแสดงจุดเด่นของเส้นขอบภาพที่ถูกสร้างเป็นฐานข้อมูล	84
45. ภาพแสดงโครงสร้างการเก็บข้อมูลของภาพฐานข้อมูล	85
46. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 1)	87
47. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 1)	87
48. ภาพแสดงการกำหนดสถานะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 1)	88
49. ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 1)	88
50. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 1)	89
51. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 2)	90
52. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 2)	90
53. ภาพแสดงการกำหนดสถานะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 2)	91
54. ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 2)	92
55. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 2)	93
56. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 3)	94
57. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 3)	94
58. ภาพแสดงการกำหนดสถานะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3)	95
59. ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3)	95

สารบัญภาพ (ต่อ)

หน้า

60. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 3)	96
61. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 4)	97
62. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 4)	97
63. ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 4)	98
64. ภาพแสดงผลของการรู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 4)	99
65. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 4)	100
66. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 5)	101
67. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 5)	101
68. ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 5)	102
69. ภาพแสดงผลของการรู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 5)	103
70. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 5)	104
71. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 6)	105
72. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 6)	105
73. ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 6)	106
74. ภาพแสดงผลของการรู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 6)	107
75. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 6)	108
76. ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 7)	109
77. ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 7)	109

สารบัญภาพ (ต่อ)

	หน้า
78. ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 7)	110
79. ภาพแสดงผลของการลู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 7)	111
80. ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 7)	112
81. ภาพแสดงผลของการลู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3) (การประมวลผลการจับคู่โดยใช้ ค่าความโค้งและทิศทางความโค้งของจุดเด่น แต่เพียงอย่างเดียว)	114
82. ภาพแสดงผลของการลู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3) (การประมวลผลการจับคู่โดยใช้ มุมของจุดเด่น แต่เพียงอย่างเดียว)	115
83. ภาพแสดงการทับซ้อนกันของวัตถุที่เหมือนกันทุกประการ	117
84. ภาพแสดงหน้ากากสำหรับปรับปรุงเส้นขอบภาพ	122
85. ภาพแสดงหน้ากากสำหรับตรวจหาเส้นขอบภาพ	122
86. ภาพแสดงภาพต้นแบบสำหรับการประมวลผลหาเส้นขอบภาพ	123
87. ภาพแสดงผลที่ได้จากการปรับปรุงเส้นขอบภาพให้คมชัดขึ้นด้วยหน้ากากในภาพที่ 84	123
88. ภาพแสดงการแยกส่วนของข้อมูลออกจากส่วนที่เป็น Background ด้วย เทคนิค Thresholding	124
89. ภาพแสดงผลของการตรวจหาเส้นขอบภาพด้วยหน้ากากในภาพที่ 85	124
90. ภาพแสดงเส้นทางเดินของเส้นขอบภาพ	125
91. ภาพแสดงการเลื่อนของเส้นขอบภาพจากตำแหน่ง (x,y) ไปยังตำแหน่ง $(x+m,y+n)$	129
92. ภาพแสดงการหมุนของเส้นขอบภาพ	131
93. ภาพแสดงการย่อหรือขยายของเส้นขอบภาพ	133
94. ภาพแสดงค่าความโค้งและทิศทางความโค้งที่เสถียรภาพต่อการหมุนและย่อ/ขยาย ของเส้นขอบภาพ	135

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การจดจำภาพวัตถุเป็นส่วนหนึ่งในระบบการมองเห็นของหุ่นยนต์หรือแขนกล ซึ่งถูกพัฒนาขึ้นมาเพื่อให้ระบบการมองเห็นของหุ่นยนต์สามารถวิเคราะห์และจดจำภาพวัตถุได้ โดยมีการพัฒนาตั้งแต่การวิเคราะห์ภาพข้อมูลที่ไม่ซับซ้อน เช่น วัตถุแต่ละชิ้นวางอยู่ในตำแหน่งที่เป็นอิสระต่อกัน จนถึงกรณีภาพวัตถุมีความซับซ้อนมากขึ้น เช่น วัตถุแต่ละชิ้นอาจวางอยู่ในตำแหน่งที่สัมผัสกัน หรือแม้แต่การวางทับซ้อนกัน ซึ่งเป็นสาเหตุสำคัญที่ทำให้ข้อมูลภาพไม่สมบูรณ์ ไม่ชัดเจน และมีการสูญหายของข้อมูลไปบางส่วน ปัญหาที่เกิดขึ้นเป็นแรงผลักดันให้ผู้ที่ทำการวิจัยในสาขานี้ พยายามคิดค้นหาวิธีและเทคนิคต่าง ๆ เพื่อใช้แก้ไขปัญหาดังกล่าว ซึ่งวิทยานิพนธ์ฉบับนี้ ได้นำเสนอวิธีการแก้ไขปัญหาดังกล่าวด้วยหลักการ “การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้อัจฉริยะ”

ในการประมวลผลผลการจดจำภาพวัตถุจำเป็นจะต้องใช้คณิตศาสตร์ช่วยในการประมวลผล โดยคณิตศาสตร์ที่เลือกนำมาประยุกต์ใช้นั้น จะต้องสามารถตอบสนองต่อเงื่อนไขที่สำคัญ 3 ประการของการจดจำภาพวัตถุ คือ

- (1) จะต้องมีความเสถียรภาพต่อการเลื่อน หรือย้ายตำแหน่งของวัตถุ
- (2) จะต้องมีความเสถียรภาพต่อการหมุน หรือทิศทาง การวางของวัตถุ
- (3) จะต้องมีความเสถียรภาพต่อการย่อหรือขยายของภาพวัตถุ อันเนื่องมาจากการซูมของกล้อง (ไม่ควรเกิน $\pm 30\%$)

1. ความสำคัญของปัญหา :

ปัญหาที่เกิดขึ้นในการศึกษาและวิจัยเรื่อง “การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้อัจฉริยะ” พอที่จะกล่าวโดยสรุปได้ดังนี้

ในการจดจำภาพวัตถุที่ถูกบดบังบางส่วน สามารถที่จะแบ่งขั้นตอนการทำงานได้เป็น 2 ขั้นตอน คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- (1) ขั้นตอนที่หนึ่ง เป็นส่วนที่ทำหน้าที่พยายามดึงเอาจุดเด่นที่เกิดขึ้นบนเส้นขอบภาพของวัตถุออกมา
- (2) ขั้นตอนที่สอง เป็นส่วนที่ทำหน้าที่ประมวลผลการจับคู่ ระหว่างจุดเด่นของภาพวัตถุที่ถูกบดบัง

บางส่วนและจุดเด่นของภาพวัตถุที่สมบูรณ์ ซึ่งเก็บเอาไว้ในฐานข้อมูล

ปัญหาที่เกิดขึ้นในส่วนของขั้นตอนที่หนึ่ง พอที่จะกล่าวโดยสังเขป คือ วิธีการที่ใช้ในการคำนวณจุดเด่นที่เกิดขึ้นบนเส้นขอบภาพนั้น มักนิยมใช้ทฤษฎีของตัวกรองเกาส์เซียน ซึ่งเป็นวิธีการที่ดีมีประสิทธิภาพ แต่มีข้อควรระวัง คือ เมื่อมีการเปลี่ยนแปลงค่า สเกล-สเปซ ของตัวกรองเกาส์เซียนไป จะทำให้จุดเด่นมีการเลื่อนตำแหน่งไปตามการเปลี่ยนแปลงของค่า สเกล-สเปซ ด้วย ซึ่งจะทำให้จุดเด่นที่คำนวณได้นั้น เป็นจุดเด่นที่ไม่มีเสถียรภาพ

ส่วนปัญหาที่เกิดขึ้นในขั้นตอนที่สอง คือ เมื่อภาพวัตถุที่ถูกบดบังบางส่วนมีความซับซ้อนมากขึ้น จะทำให้ไม่สามารถทำการประมวลผลการจับคู่ ระหว่างจุดเด่นของภาพวัตถุที่ถูกบดบังบางส่วนกับจุดเด่นของภาพวัตถุที่ถูกเก็บไว้ในฐานข้อมูลได้อย่างมีประสิทธิภาพ

2. ทำไมปัญหานี้ถึงน่าสนใจ :

เนื่องจากปัญหาทั้ง 2 ประเด็นที่กล่าวมานั้น เป็นหัวใจสำคัญในการศึกษาและวิจัยการจดจำภาพวัตถุ ถ้าสามารถแก้ไขปัญหาทั้ง 2 ประเด็นนี้ได้ จะทำให้การประมวลผลการจดจำและบอกตำแหน่งของวัตถุมีประสิทธิภาพมากขึ้น โดยมีแนวทางที่จะแก้ไขปัญหาดังกล่าว คือ

- (1) ใช้วิธีการคำนวณหาจุดเด่นที่มีประสิทธิภาพเพื่อให้ได้จุดเด่นที่มีเสถียรภาพ แม้จะมีการเปลี่ยนแปลง ค่า สเกล-สเปซ ของตัวกรองเกาส์เซียนไปก็ตาม
- (2) ทาวิธีการที่มีประสิทธิภาพในการประมวลผลการจับคู่ของจุดเด่น เมื่อภาพของวัตถุที่ถูกบดบังบางส่วนมีความซับซ้อนมากขึ้น

3. หลักการที่นำเสนอ :

หลักการที่จะนำเสนอในการแก้ไขปัญหาดังกล่าวมาแล้วนั้น คือ

(1) ในประเด็นของการคำนวณหาจุดเด่น จะนำเสนอวิธีการคำนวณหาจุดเด่นด้วยเทคนิค การคำนวณหาจุดเด่นด้วยตัวกรองเกาส์เซียนและคิวบิต บี-สไฟน์ พร้อมกับใช้วิธี การตรึงจุดเด่นให้คงที่โดยการ ใช้ซอฟต์แวร์ ซึ่งจากหลักการดังกล่าวจะทำให้จุดเด่นที่คำนวณได้นั้น เป็นจุดเด่นที่มีเสถียรภาพ เหตุผลที่เลือกใช้เทคนิคต่าง ๆ ที่กล่าวมาแล้วนี้ เนื่องจากเทคนิคเหล่านี้สอดคล้องกับปัญหาที่ต้องการแก้ไข และสามารถตอบสนองต่อเงื่อนไขของการประมวลผลการจดจำภาพวัตถุได้อย่างสมบูรณ์

(2) ใช้ทฤษฎีโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ (Hopfield neural network) เข้ามาแก้ไข การประมวลผลการจับคู่ของจุดเด่น ทั้งนี้เนื่องจากโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ถูกพัฒนาขึ้นมา เพื่อแก้ไขปัญหา ปัญหาการเลือกคำตอบที่ดีที่สุด (Optimization Problem) ได้อย่างมีประสิทธิภาพ ซึ่งเงื่อนไข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไขของการประมวลผลการจับคู่ของจุดเด่นก็จัดอยู่ในกลุ่มของ **ปัญหาการเลือกคำตอบที่ดีที่สุด** ด้วยเช่นกัน จึงเป็นเหตุผลในการเลือกใช้วิธีดังกล่าวในการแก้ปัญหาการประมวลผลการจับคู่ของจุดเด่น ในกรณีที่ภาพวัตถุที่ถูกดบังบางส่วนมีความซับซ้อนมากขึ้น

จุดมุ่งหมายและวัตถุประสงค์

(1) เพื่อศึกษาและวิจัย วิธีการคำนวณหาจุดเด่นบนเส้นขอบภาพที่เสถียรภาพต่อการหมุน การย่อ/ขยาย และการเลื่อนของเส้นขอบภาพ และสิ่งที่สำคัญที่สุด คือ มีความเสถียรภาพต่อการเปลี่ยนแปลงค่าสเกล-สเปซ ของตัวกรองเกาส์เซียน

(2) เพื่อศึกษาและวิจัย กระบวนการการประมวลผลของจับคู่ของจุดเด่นให้มีประสิทธิภาพ โดยการนำเอาโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์เข้ามาประยุกต์ใช้งาน

สมมุติฐานของการศึกษา

จากหลักการและแนวทางการแก้ไขปัญหาก็กล่าวมา ทำให้คาดหวังได้ว่าจะสามารถคำนวณหาจุดเด่นที่เกิดขึ้นบนเส้นขอบภาพของวัตถุที่มีเสถียรภาพได้ ด้วยเทคนิคผสมผสานระหว่างการคำนวณหาจุดเด่นด้วยตัวกรองเกาส์เซียนและคิวบิค บี-สไปรน์ โดยใช้ร่วมกับเทคนิคการตรึงจุดเด่นให้คงที่โดยการใช้ซอฟต์แวร์ จากวิธีการดังกล่าวจะทำให้ได้ผลลัพธ์ที่ดีกว่าวิธีการหาจุดเด่นโดยทั่ว ๆ ไป นอกจากนี้ ในส่วนของการประมวลผลการจับคู่ของจุดเด่นยังได้นำทฤษฎีโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์มาประยุกต์ใช้งาน เพื่อใช้แก้ไขปัญหาในกรณีที่ภาพวัตถุมีความซับซ้อนมากขึ้น ทั้งนี้เนื่องจากโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ ถูกพัฒนาขึ้นเพื่อแก้ไขปัญห **ปัญหาการเลือกคำตอบที่ดีที่สุด** และในขณะเดียวกันปัญหาของการประมวลผลการจับคู่ก็เป็นปัญหาที่จัดอยู่ในกลุ่มของ **ปัญหาการเลือกคำตอบที่ดีที่สุด** เช่นกัน จึงนับได้ว่าเป็นวิธีการแก้ไขปัญหาที่ตรงประเด็นและคาดว่าจะได้ผลลัพธ์ที่ดีในการประมวลผลการจับคู่ของจุดเด่น

ขอบเขตของการศึกษา

ขอบเขตของการศึกษาและวิจัยจะมุ่งเน้นการจดจำภาพวัตถุในแบบสองมิติเท่านั้น โดยจะเน้นหนักอยู่ 2 กระบวนการ คือ กระบวนการคำนวณหาจุดเด่นของเส้นขอบภาพที่มีเสถียรภาพ และ กระบวนการจับคู่ของจุดเด่นด้วยโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ ส่วนการประมวลผลหาเส้นขอบภาพเป็นกระบวนการที่ทำหน้าที่แปลงภาพระดับสีเทาให้เป็นเส้นขอบภาพ เพื่อเป็นข้อมูลที่จะป้อนเข้าสู่กระบวนการคำนวณหาจุดเด่น ซึ่งกระบวนการดังกล่าวไม่ได้เป็นประเด็นหลักในการดำเนินการศึกษาและวิจัยในครั้งนี้ ดังนั้น จึงได้จัดกระบวนการประมวลผลหาเส้นขอบภาพเอาไว้เป็นตัวแปรควบคุมในการศึกษาและวิจัย

รายละเอียดของวิทยานิพนธ์

ในวิทยานิพนธ์ฉบับนี้ ได้แบ่งเนื้อหาที่ใช้ในการการดำเนินการวิจัยไว้เป็นบทต่าง ๆ ด้วยกันทั้งหมด 6 บท ซึ่งพอที่จะกล่าวโดยสังเขปดังนี้

1. บทที่ 1 บทนำ จะกล่าวถึงความเป็นมาของงานวิจัย และปัญหาต่าง ๆ ที่น่าสนใจซึ่งเกิดขึ้นกับงานวิจัยที่ผ่านมา ตลอดจนเสนอแนวทางในการแก้ไขปัญหาดังกล่าว
2. บทที่ 2 การคำนวณหาจุดเด่นของเส้นขอบภาพ สำคัญสำคัญในบทนี้ จะนำเสนอเทคนิคศาสตร์ที่ใช้ในการคำนวณหาจุดเด่น พร้อมทั้งอัลกอริทึมต่าง ๆ ที่ใช้ในการคำนวณ
3. บทที่ 3 โครงข่ายเซลล์ประสาทเทียม เป็นการนำเสนอทฤษฎีโครงข่ายเซลล์ประสาทเทียม โดยเริ่มตั้งแต่คณิตศาสตร์ของเซลล์ประสาทจนไปถึงรูปแบบจำลองของโครงข่ายเซลล์ประสาทแบบต่าง ๆ
4. บทที่ 4 การจดจำภาพวัตถุ ในบทนี้จะกล่าวถึงหลักการทำงานของ การจดจำภาพวัตถุ และการประมวลผลการจดจำภาพวัตถุ โดยนำทฤษฎีและหลักการที่ได้กล่าวไว้ในบทที่ 2 และบทที่ 3 มาประยุกต์ใช้งาน
5. บทที่ 5 ผลการทดลอง เป็นการแสดงให้เห็นถึงผลของการประมวลผลการจดจำภาพวัตถุ ในกรณีต่าง ๆ
6. บทที่ 6 บทสรุป จะกล่าวถึงปัญหาต่าง ๆ ที่เกิดขึ้นในการทำวิจัย และเสนอแนะแนวทางในการพัฒนาในขั้นต่อไป

บทที่ 2

การคำนวณหาจุดเด่นของเส้นขอบภาพ

บทนำ

ในกระบวนการประมวลผลการจดจำภาพวัตถุ จำเป็นอย่างยิ่งที่จะต้องลดปริมาณข้อมูลที่ไม่สำคัญของภาพลง เพื่อให้สามารถประมวลผลได้เร็วขึ้น ใช้พื้นที่หน่วยความจำน้อยลง และที่สำคัญอีกประการหนึ่งก็คือ ให้ง่ายต่อการสร้างแบบจำลองคณิตศาสตร์ เพื่อใช้ในการวิเคราะห์ข้อมูลและการสร้างโปรแกรม

การประมวลผลหาเส้นขอบภาพ เป็นอีกวิธีหนึ่งที่ยิมนำมาใช้ในการลดปริมาณข้อมูลที่ไม่สำคัญออกไป เพื่อให้เหลือแต่ข้อมูลสำคัญที่จะใช้ในการวิเคราะห์ เส้นขอบภาพที่ได้นั้น จะเป็นคุณลักษณะเฉพาะของภาพวัตถุนั้น ๆ และภายในเส้นขอบภาพก็จะประกอบไปด้วยส่วนหัวและส่วนโค้งต่าง ๆ ซึ่งก็เป็นลักษณะเฉพาะตัวของเส้นขอบภาพนั้น ๆ อีกเช่นกัน ลักษณะเฉพาะตัวเหล่านี้ที่เกิดขึ้นบนเส้นขอบภาพ จะเรียกว่า “ลักษณะเด่น” หรือ “จุดเด่น”

ดังนั้น ความหมายของคำว่า “จุดเด่น” จึงหมายถึง จุดที่มีค่าความโค้งสูงที่ปรากฏบนเส้นขอบภาพของวัตถุ ซึ่งจะปรากฏในตำแหน่งของเส้นขอบภาพที่เป็นส่วนหัวและส่วนโค้งเท่านั้น จุดเด่นของเส้นขอบภาพจึงเป็นคุณลักษณะเฉพาะตัวของภาพวัตถุนั้น ๆ ได้ และถูกนำมาใช้เป็นคุณสมบัติในการประมวลผลการจดจำภาพวัตถุ ส่วนวิธีการคำนวณหาจุดเด่นจะได้กล่าวโดยละเอียดในหัวข้อถัดไป

คณิตศาสตร์ของตัวกรองเกาส์เซียน

1. คุณสมบัติของตัวกรองเกาส์เซียน

สมการของตัวกรองเกาส์เซียนถูกนิยามไว้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

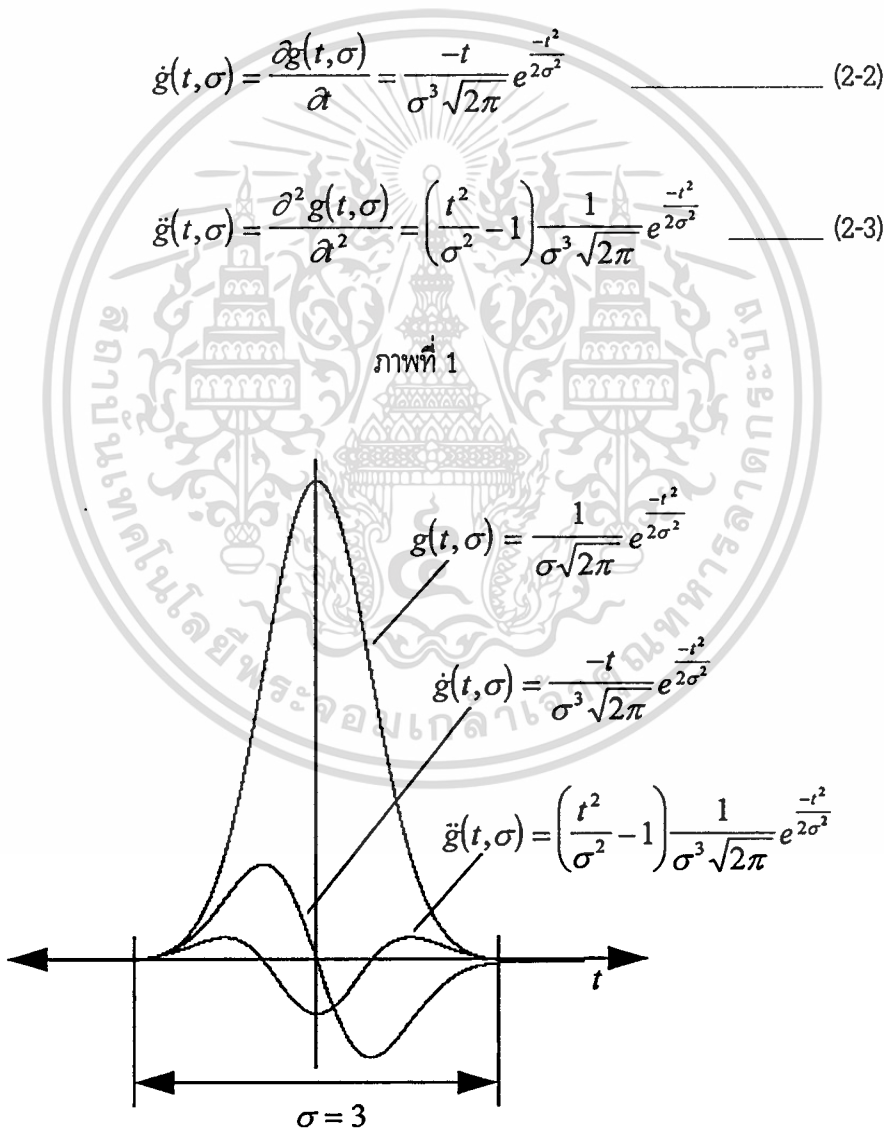
$$g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (2-1)$$

เมื่อ σ คือ ค่า สเกล-สเปซ ของตัวกรองเกาส์เซียน หรือค่าเบี่ยงเบนมาตรฐาน
 t คือ ตัวแปรอิสระ

ส่วนอนุพันธ์ย่อยอันดับที่หนึ่งและอันดับที่สอง สามารถหาได้จากสมการที่ (2-2) และ (2-3) ตามลำดับ และเส้นกราฟสมการที่ (2-1), (2-2) และ (2-3) ได้แสดงไว้ในภาพที่ 1

$$\dot{g}(t, \sigma) = \frac{\partial g(t, \sigma)}{\partial t} = \frac{-t}{\sigma^3\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (2-2)$$

$$\ddot{g}(t, \sigma) = \frac{\partial^2 g(t, \sigma)}{\partial t^2} = \left(\frac{t^2}{\sigma^2} - 1\right) \frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (2-3)$$



ภาพแสดงเส้นกราฟของสมการตัวกรองเกาส์เซียน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อผู้จัดทำให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

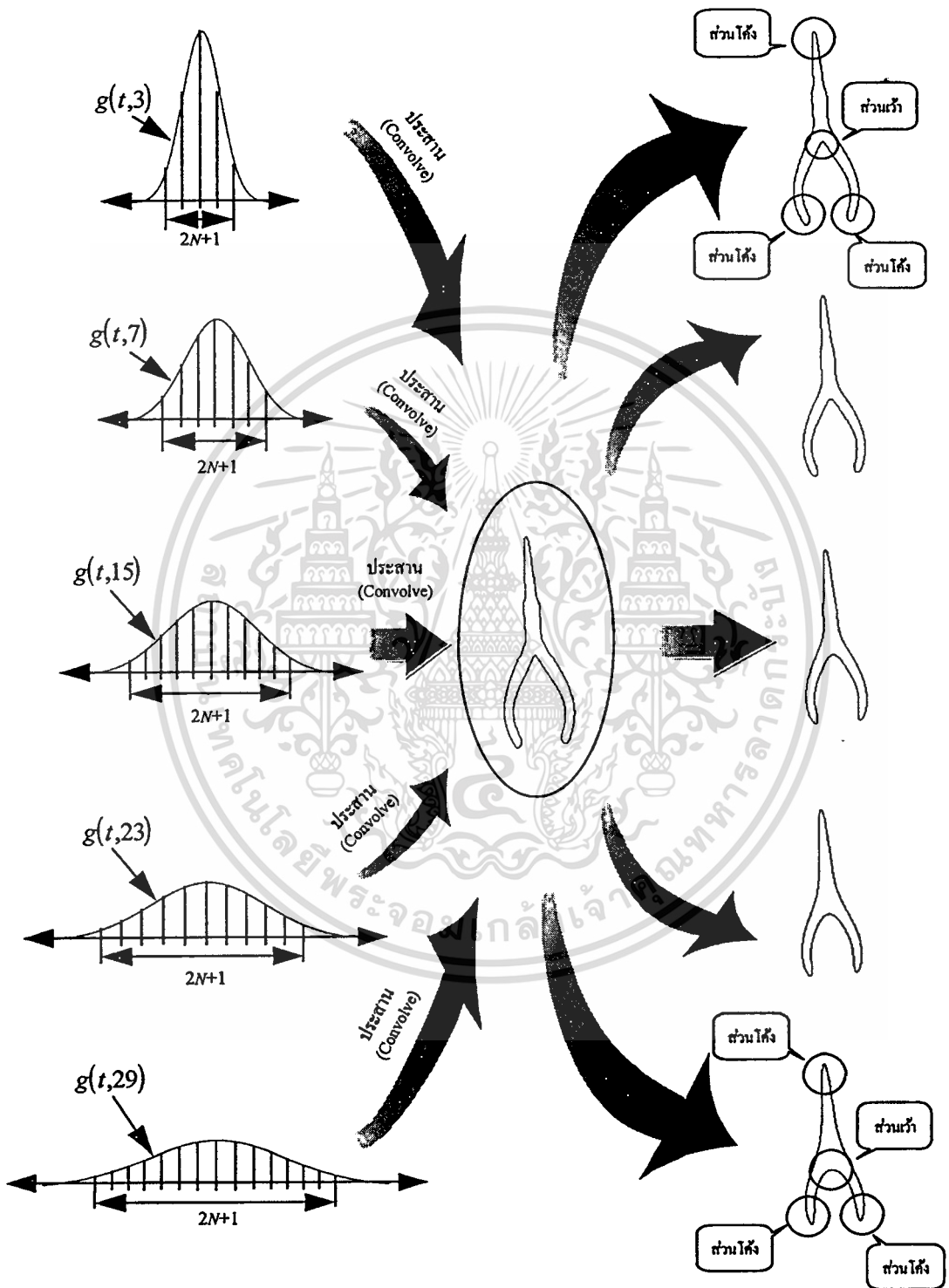
สมการเกาส์เซียนเป็นสมการที่มีคุณสมบัติเป็นตัวกรอง โดยมีค่า สเกล-สเปซ เป็นตัวควบคุมคุณสมบัติของการกรอง ถ้าค่า สเกล-สเปซ มีค่าน้อย ตัวกรองจะมีคุณสมบัติยอมให้ความถี่สูงผ่านไป แต่ในทางตรงกันข้าม ค่า สเกล-สเปซ มีค่ามาก ตัวกรองก็จะมีคุณสมบัติเป็นตัวกีดความถี่สูง ดังที่ได้พิสูจน์ให้เห็นในภาพที่ 2 ซึ่งเป็นการประสาน (Convolution) เส้นขอบภาพ (ภาพคิม) กับตัวกรองเกาส์เซียน โดยใช้สมการที่ (2-4) ซึ่งรายละเอียดของวิธีการประมวลผลจะได้กล่าวถึงในหัวข้อ การประมวลผลเพื่อหาจุดเด่นของเส้นขอบภาพ ต่อไป

$$\left. \begin{aligned} X(t, \sigma) &= x(t) * g(t, \sigma) \\ Y(t, \sigma) &= y(t) * g(t, \sigma) \end{aligned} \right\} \quad (2-4)$$

เมื่อ $x(t)$ และ $y(t)$ คือ ฟังก์ชันพารามตริกของเส้นขอบภาพ และสัญลักษณ์ $*$ คือ การประสาน

จากผลการกรองเส้นขอบภาพ (ภาพที่ 2) ได้แสดงให้เห็นคุณสมบัติตัวกรองเกาส์เซียนที่มีผลต่อการเปลี่ยนค่า สเกล-สเปซ ที่ค่าต่าง ๆ กัน จะสังเกตเห็นได้ว่าส่วนหัวและส่วนโค้งที่แท้จริงจะปรากฏอย่างเด่นชัด แม้ว่าจะทำการกรองเส้นขอบภาพด้วยตัวกรองแล้วก็ตาม ซึ่งเป็นเหตุผลที่สำคัญที่ถูกนำมาเป็นเงื่อนไขในการคำนวณหาจุดเด่นด้วยเทคนิคการเปลี่ยนค่า สเกล-สเปซ เพื่อคำนวณหาจุดเด่นที่มีเสถียรภาพ โดยอาศัยหลักการที่ว่า ส่วนหัวและส่วนโค้งของเส้นขอบภาพเป็นคุณลักษณะเฉพาะตัว ที่สามารถคำนวณเป็นค่าตัวเลขได้ด้วยสมการหาค่าความโค้ง และเป็นสิ่งกำหนดคุณลักษณะของ "จุดเด่น" ซึ่งส่วนหัวและส่วนโค้งของเส้นขอบภาพที่สามารถนิยามให้เป็นจุดเด่นได้นั้น จะต้องเป็นส่วนหัวและส่วนโค้งที่มีค่าความโค้งสูงเมื่อเทียบกับบริเวณข้างเคียง ซึ่งมีวิธีการคำนวณดังจะได้นำเสนอในหัวข้อถัดไป

ภาพที่ 2



ภาพแสดงคุณสมบัติของตัวกรองเกาส์เซียนในการกรองเส้นขอบภาพที่ค่า สเกล-สเปซ ต่าง ๆ กัน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. สมการสำหรับหาค่าความโค้ง

สมการสำหรับการคำนวณหาค่าความโค้งที่มีฟังก์ชันเส้นโค้ง $y=f(x)$ ถูกนิยามไว้ดังนี้

$$k = \frac{\left| \frac{d^2 y}{dx^2} \right|}{\left(1 + \left(\frac{dy}{dx} \right)^2 \right)^{3/2}} \quad (2-5)$$

ในทำนองเดียวกัน ฟังก์ชันเส้นโค้งในรูปแบบสมการพาราเมตริกก็สามารถคำนวณได้เช่นกัน โดยมี t เป็นตัวแปรอิสระของสมการพาราเมตริก เมื่อกำหนดให้

$$x = x(t) \quad (2-6a)$$

$$y = y(t) \quad (2-6b)$$

ถ้า $P(t)$ เป็นเวกเตอร์ของจุดใด ๆ ที่ปรากฏบนเส้นโค้งในรูปแบบสมการพาราเมตริก จะได้ว่า

$$P(t) = [x(t) \quad y(t)] \quad (2-7)$$

และอนุพันธ์ของ $P(t)$ สามารถหาได้จาก

$$\dot{P}(t) = [\dot{x}(t) \quad \dot{y}(t)] \quad (2-8a)$$

ค่าความชันของเส้นโค้ง หาได้จากสมการ (2-8b)

$$y' = \frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{\dot{y}(t)}{\dot{x}(t)} \quad (2-8b)$$

และในทำนองเดียวกัน อนุพันธ์อันดับที่สองของเส้นโค้ง $P(t)$ ก็สามารถคำนวณได้จาก

$$\ddot{P}(t) = [\ddot{x}(t) \quad \ddot{y}(t)] \quad (2-9a)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือ

$$\frac{d^2 y}{dx^2} = \frac{dy'}{dt} \bigg/ \frac{dx}{dt} \quad 1$$

$$\frac{d^2 y}{dx^2} = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t))^3} \quad \text{----- (2-9b)}$$

เพราะฉะนั้น สมการสำหรับหาค่าความโค้งของเส้นโค้งที่อยู่ในรูปแบบสมการพาราเมตริก จึงสามารถคำนวณได้จาก

$$k = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t))^3} = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\left(1 + \left(\frac{\dot{y}(t)}{\dot{x}(t)}\right)^2\right)^{3/2}} = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\left((\dot{x}(t))^2 + (\dot{y}(t))^2\right)^{3/2}} \quad \text{----- (2-10)}$$

สมการสำหรับหาค่าความโค้งที่แสดงในสมการที่ (2-10) เป็นสมการหาค่าความโค้งในรูปแบบพาราเมตริก ซึ่งเป็นสมการพื้นฐานที่จะถูกนำไปประยุกต์ใช้งานต่อไป

การประมวลผลเพื่อหาจุดเด่นของเส้นขอบภาพ

1. การประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน

การประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียนเป็นกระบวนการที่สำคัญ ที่จะถูกนำไปเป็นเครื่องมือในการประมวลผลเพื่อหาจุดเด่นของเส้นขอบภาพ ดังนั้น ในหัวข้อนี้จึงได้เกริ่นนำเพื่อให้ทราบถึง วิธีการเก็บข้อมูลของเส้นขอบภาพ วิธีการกำหนดขนาดหน้าต่างของตัวกรองเกาส์เซียนสำหรับการประสาน ตลอดจนวิธีการประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน ซึ่งที่กล่าวมาทั้งหมดนั้น จะทำการประมวลผลภายใต้หลักเกณฑ์การคำนวณคณิตศาสตร์แบบไม่ต่อเนื่อง (Discrete Mathematics)

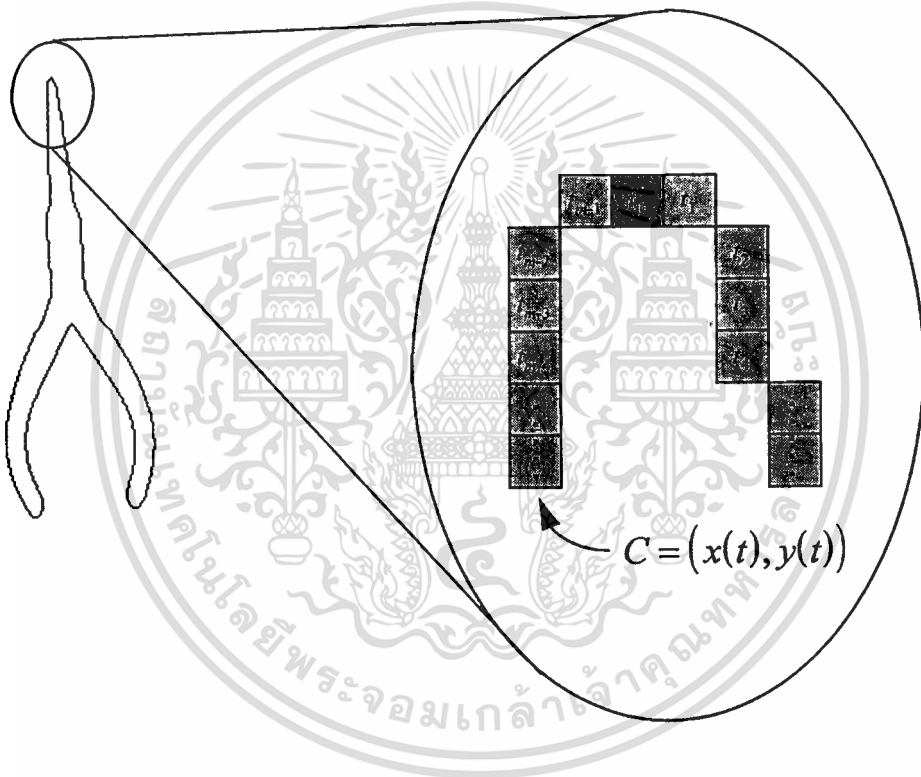
1.1 วิธีการเก็บข้อมูลของเส้นขอบภาพ

เส้นขอบภาพที่ปรากฏบนหน้าจอคอมพิวเตอร์ จะถูกแทนด้วยจุดภาพหลาย ๆ จุดวางเรียงต่อกัน ซึ่งจุดภาพเหล่านี้จะถูกเก็บตำแหน่งด้วยเลขจำนวนเต็ม ดังนั้น การบอกตำแหน่งของจุดภาพบนเส้นขอบ

¹ George B. Thomas, Jr. and Ross L. Finney, *Calculus and Analytic Geometry* (U.S.A. : Addison-Wesley, 1996), P. 745. เอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพจึงสามารถบอกแทนได้ด้วยฟังก์ชันแบบไม่ต่อเนื่อง $C=f(x,y)$ เมื่อ C คือเส้นขอบภาพ และ x, y เป็นพิกัดในแนวแกน x และ y ของตำแหน่งจุดภาพบนเส้นขอบภาพตามลำดับ เพื่อให้สอดคล้องต่อหลักเกณฑ์การคำนวณทางคณิตศาสตร์และเหมาะสมต่อการใช้งาน จึงเขียนฟังก์ชัน $f(x,y)$ ให้อยู่ในรูปของฟังก์ชันพารามetriคได้ คือ $C=(x(t),y(t))$ เมื่อ t แทนเส้นทางเดินของจุดภาพบนเส้นขอบภาพ และมีวิธีการเก็บข้อมูลของเส้นขอบภาพดังแสดงในภาพที่ 3

ภาพที่ 3



ภาพแสดงการเก็บข้อมูลของเส้นขอบภาพ

1.2 วิธีการกำหนดขนาดหน้าต่างของตัวกรองเกาส์เซียนสำหรับการประสาน

ตัวกรองเกาส์เซียนที่แสดงในสมการที่ (2-1) เป็นฟังก์ชันต่อเนื่อง แต่ข้อมูลที่จะใช้ในการประมวลผลร่วมกับตัวกรองเกาส์เซียน คือ เส้นขอบภาพ ซึ่งเป็นข้อมูลแบบไม่ต่อเนื่อง ดังนั้น ตัวกรองเกาส์เซียนจึงต้องถูกแซมพลิง (Sampling) ให้เป็นฟังก์ชันแบบไม่ต่อเนื่องด้วย ดังแสดงในภาพที่ 4 เพื่อให้สอดคล้องกับข้อมูลที่ใช้ในการประมวลผล ซึ่งโดยทั่วไปแล้ว ขนาดของหน้าต่างสำหรับเก็บค่าสัมประสิทธิ์ของตัวกรองเกาส์เซียนจะถูกกำหนดด้วยจำนวนแซมพลิง $2N+1$ โดยที่ N สามารถคำนวณได้จาก

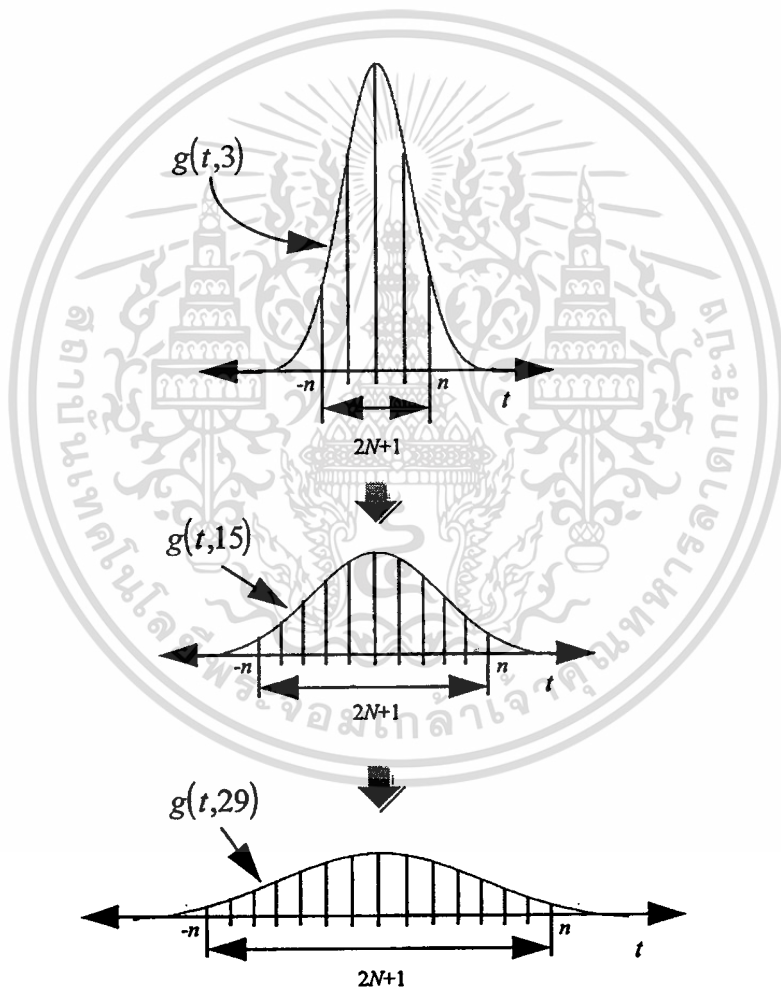
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$N = \min\{N_1, N_2\} \quad (2-11)$$

เมื่อ
$$N_1 = \left\{ n: \text{Max} \sum_{t=-n}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \leq 0.99 \right\} \quad (2-11a)$$

และ
$$N_2 = \lfloor 5\sigma \rfloor \quad (2-11b)$$

ภาพที่ 4

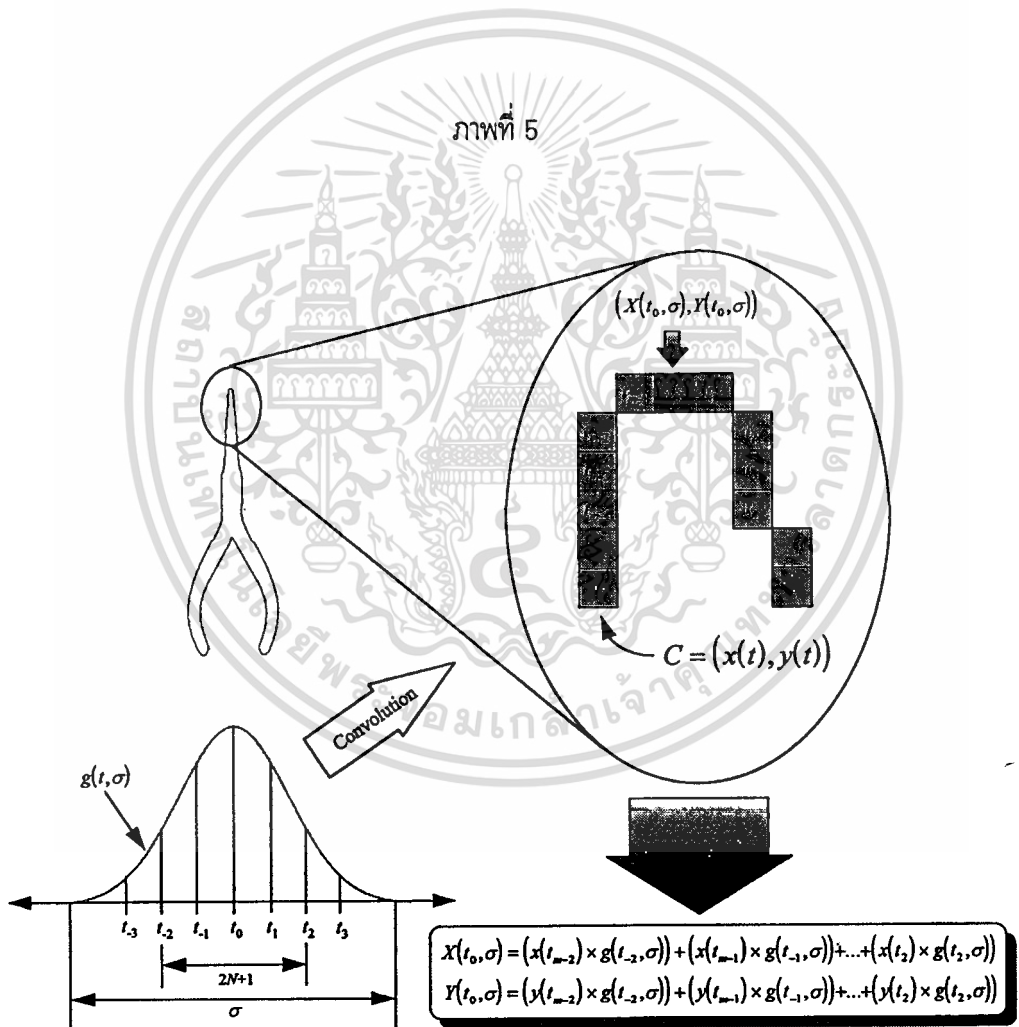


ภาพแสดงขนาดหน้าต่างต่างสำหรับเก็บค่าสัมประสิทธิ์ของตัวกรองเกาส์เซียนที่ค่า สเกล-สเปซ ต่าง ๆ กัน

* $\lfloor a \rfloor$ หมายถึง เลขจำนวนเต็มที่มีค่ามากที่สุด แต่น้อยกว่าหรือเท่ากับ a เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 วิธีการประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน

การประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน มีวิธีการประสานดังแสดงในภาพที่ 5 โดยข้อมูลที่นำมาประสานนั้น เป็นเส้นขอบภาพรูปคิม ซึ่งถูกเก็บตำแหน่งของจุดภาพไว้ในรูปแบบฟังก์ชันพาราเมตริก $C=(x(t),y(t))$ โดยมีตัวแปร t แทนตำแหน่งเส้นทางเดินของจุดภาพ ส่วนตัวกรองเกาส์เซียนจะถูกแซมพลิง โดยมีขนาดหน้าต่างสำหรับเก็บค่าสัมประสิทธิ์ของตัวกรองเกาส์เซียนเท่ากับ $2N+1$ และการแซมพลิงจะทำการแซมพลิงด้วยเลขจำนวนเต็ม เพื่อให้สอดคล้องกับข้อมูลของเส้นขอบภาพ ผลที่ได้จากการประสานก็คือ ค่า $X(t,\sigma)$ และ $Y(t,\sigma)$ ซึ่งจะถูกนำไปประมวลผลหาค่าความโค้งในขั้นตอนต่อไป และในภาพที่ 5 ก็ได้แสดงวิธีการประสานเส้นขอบภาพที่ตำแหน่ง t_0 กับตัวกรองเกาส์เซียนประกอบเป็นตัวอย่างด้วย



ภาพแสดงวิธีการประสานเส้นขอบภาพด้วยตัวกรองเกาส์เซียน

2. การคำนวณหาค่าความโค้งด้วยการเปลี่ยนค่า สเกล-สเปซ ของตัวกรองเกาส์เซียน

ดังที่ได้กล่าวมาแล้วว่า ค่าความโค้งเป็นสิ่งกำหนด “จุดเด่น” ที่ปรากฏบนเส้นขอบภาพ โดยจุดที่มีค่าความโค้งสูงเท่านั้นจึงจะสามารถกำหนดให้เป็นจุดเด่นได้ ดังนั้น ในลำดับขั้นตอนต่อไปนี้จะเป็นการนำเสนอวิธีการคำนวณหาค่าความโค้ง เพื่อกำหนดเป็นจุดเด่นด้วยเทคนิคการเปลี่ยนค่า สเกล-สเปซ ของตัวกรองเกาส์เซียน

เมื่อพิจารณาสมการที่ (2-10) จะพบว่าสมการดังกล่าวเป็นสมการหาค่าความโค้งโดยปราศจากตัวแปร สเกล-สเปซ ของตัวกรองเกาส์เซียน ซึ่งยังไม่ตรงตามวัตถุประสงค์ของการคำนวณหาค่าความโค้งด้วยการเปลี่ยนค่า สเกล-สเปซ ของตัวกรองเกาส์เซียน ดังนั้น เพื่อให้บรรลุวัตถุประสงค์ดังกล่าว ฟังก์ชันพาราเมตริก ฟังก์ชัน $x(t)$ และ $y(t)$ จึงถูกประสานกับสมการเกาส์เซียน และสามารถเขียนสมการหาค่าความโค้งได้ดังนี้

$$\begin{aligned}\dot{X}(t, \sigma) &= \frac{\partial X(t, \sigma)}{\partial t} \\ &= x(t) * \frac{\partial g(t, \sigma)}{\partial t} \\ &= x(t) * \dot{g}(t, \sigma)\end{aligned}\quad (2-12a)$$

$$\begin{aligned}\ddot{X}(t, \sigma) &= \frac{\partial \dot{X}(t, \sigma)}{\partial t} \\ &= x(t) * \frac{\partial \dot{g}(t, \sigma)}{\partial t} \\ &= x(t) * \ddot{g}(t, \sigma)\end{aligned}\quad (2-12b)$$

ส่วน $\dot{Y}(t, \sigma)$ และ $\ddot{Y}(t, \sigma)$ ก็สามารถคำนวณได้ในลักษณะเดียวกันกับสมการที่ (2-12a) และ (2-12b) ตามลำดับ ดังนั้น ค่าความโค้งจึงสามารถคำนวณได้จากสมการที่ (2-13)

$$K(t, \sigma) = \frac{\dot{X}(t, \sigma)\dot{Y}(t, \sigma) - Y(t, \sigma)\dot{X}(t, \sigma)}{\left(\dot{X}(t, \sigma)^2 + \dot{Y}(t, \sigma)^2\right)^{3/2}} \quad (2-13)$$

ภาพที่ 6 เป็นเส้นขอบภาพวัตถุต้นแบบรูปหลายเหลี่ยม ซึ่งถูกเก็บข้อมูลในรูปของฟังก์ชัน $C=f(x, y)$ เมื่อ C คือเส้นขอบภาพ และ x, y เป็นพิกัดในแนวแกน x และ y ของตำแหน่งจุดภาพของเส้นขอบภาพตามลำดับ เพื่อให้สะดวกต่อการคำนวณ จึงเขียนฟังก์ชัน $f(x, y)$ ให้อยู่ในรูปของฟังก์ชันพาราเมตริกได้ คือ $C=(x(t), y(t))$ เมื่อ t เป็นเส้นทางเดินของจุดภาพบนเส้นขอบภาพ ดังนั้น จุดเด่นจึงสามารถคำนวณได้จาก

สมการ (2-13) ซึ่งมีลำดับขั้นตอนการคำนวณดังแสดงไว้ในอัลกอริทึมที่หนึ่ง และได้ผลลัพธ์ดังแสดงในภาพที่ 7 และ 8

ภาพที่ 6



ภาพต้นแบบรูปหลายเหลี่ยม

อัลกอริทึมที่หนึ่ง

- (1) กำหนดสภาวะเริ่มต้นด้วยการกำหนดค่า $\sigma = 3$
- (2) คำนวณขนาดหน้าตาของตัวกรองเกาส์เซียนสำหรับเก็บค่าสัมประสิทธิ์ ซึ่งสามารถคำนวณได้จาก $2N+1$ ในสมการที่ (2-11)
- (3) คำนวณหาค่า $\dot{X}(t, \sigma)$, $\ddot{X}(t, \sigma)$, $\dot{Y}(t, \sigma)$ และ $\ddot{Y}(t, \sigma)$

$$\dot{X}(t, \sigma) = x(t) * \dot{g}(t, \sigma)$$

$$\ddot{X}(t, \sigma) = x(t) * \ddot{g}(t, \sigma)$$

$$\dot{Y}(t, \sigma) = y(t) * \dot{g}(t, \sigma)$$

$$\ddot{Y}(t, \sigma) = y(t) * \ddot{g}(t, \sigma)$$

- (4) นำค่า $\dot{X}(t, \sigma)$, $\ddot{X}(t, \sigma)$, $\dot{Y}(t, \sigma)$ และ $\ddot{Y}(t, \sigma)$ ที่คำนวณได้จาก (3) แทนลงในสมการที่ (2-13)

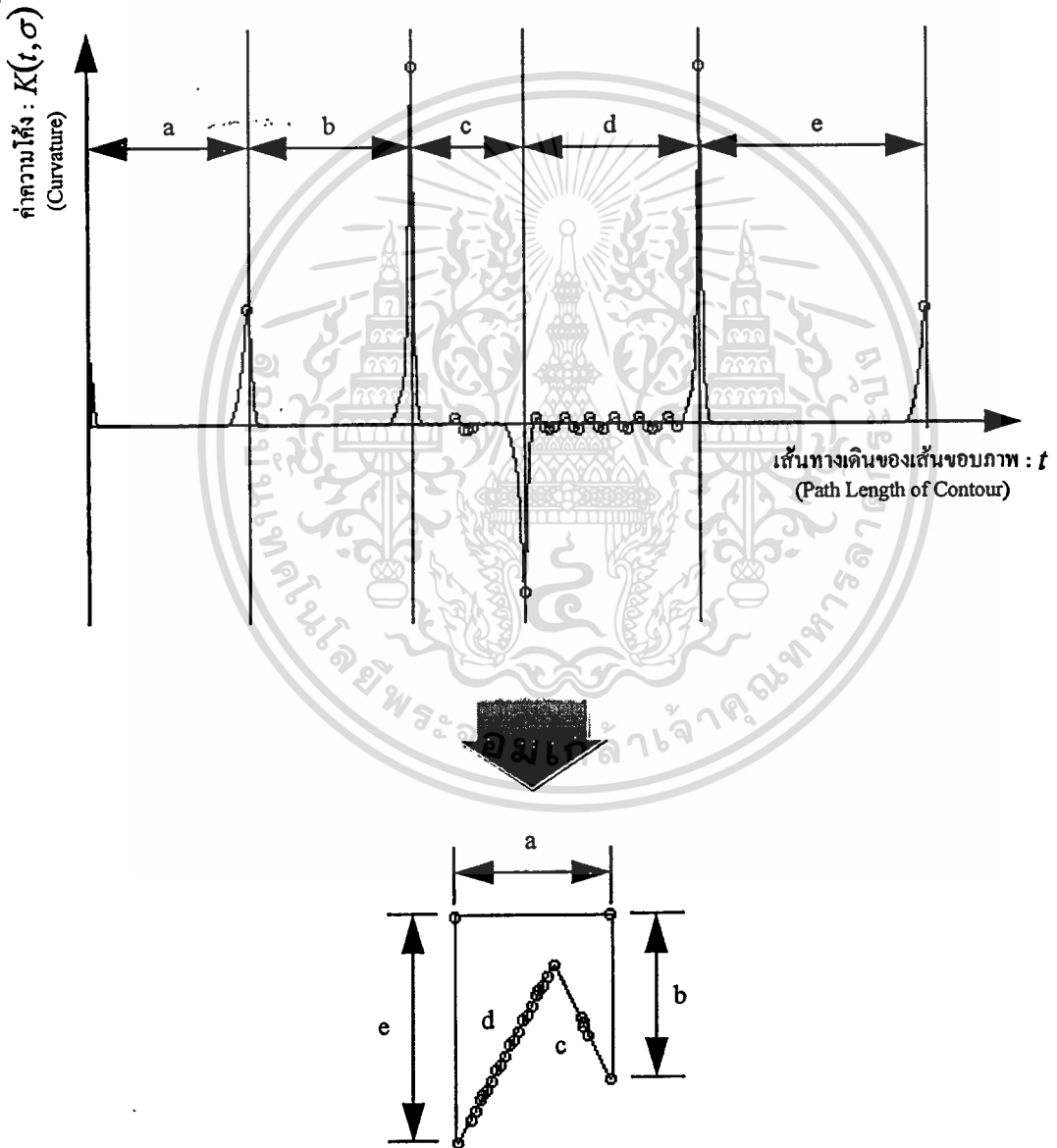
- (5) คำนวณหาจุดเด่นที่ปรากฏบนฟังก์ชัน $K(t, \sigma)$ โดยการพิจารณาเอาจุดที่มีการเปลี่ยนค่าความโค้งดังนี้

- (5.1) จุดที่มีค่าความโค้งของฟังก์ชัน $K(t, \sigma)$ เปลี่ยนจากค่าบวกไปเป็นลบ และมีค่ามากกว่า x (เมื่อ x เป็นค่า Threshold สำหรับกำหนดค่าความโค้งของจุดเด่น แนะนำให้ใช้ค่า 0.0035)

- (5.2) จุดที่มีค่าความโค้งของฟังก์ชัน $K(t, \sigma)$ เปลี่ยนจากค่าลบไปเป็นบวก และมีค่าน้อยกว่า $-x$

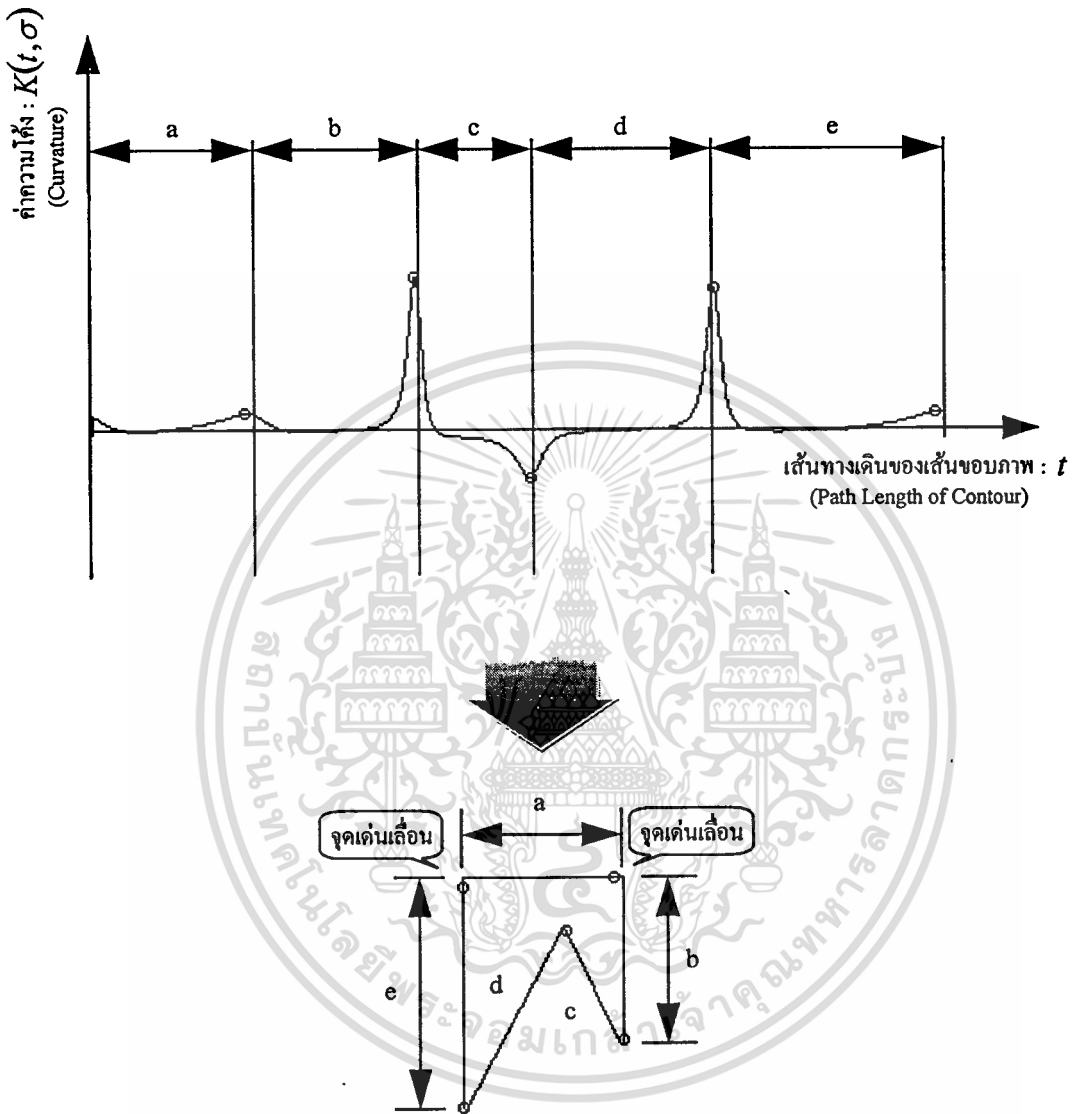
- (6) พิจารณาจำนวนของจุดเด่นที่คำนวณได้ก่อนหน้านี้ ถ้ามีค่าน้อยกว่าหรือเท่ากับจุดเด่นที่คำนวณได้ขณะนี้ ให้ข้ามไปทำขั้นตอนที่ (7) แต่ถ้ามากกว่า ให้เพิ่มค่า $\sigma = \sigma + 2$ แล้วย้อนกลับไปทำขั้นตอนที่ (2) ใหม่
- (7) หยุดการคำนวณ

ภาพที่ 7



ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 3$ โดยใช้อัลกอริทึมที่หนึ่ง เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ภายในเพื่อการศึกษาเท่านั้น เมื่อผู้ผู้ใดเห็นใบโฆษณาการคัดค้านว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 8



ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 19$ โดยใช้อัลกอริทึมที่หนึ่ง

จะเห็นได้ว่า เมื่อมีการเปลี่ยนค่า สเกล-สเปซ ให้เพิ่มขึ้นเรื่อย ๆ โดยเพิ่มทีละ 2 จะทำให้ตำแหน่งของจุดเด่นมีการเลื่อนไปจากตำแหน่งเดิม ดังที่ได้พิสูจน์ให้เห็นในภาพที่ 8 ซึ่งจะมีผลทำให้จุดเด่นที่คำนวณได้นั้น ไม่มีเสถียรภาพ ทั้งนี้เนื่องมาจากสมการที่ (2-12) และ (2-13) มีความไวสูงต่อการเปลี่ยนค่า สเกล-สเปซ

3. การคำนวณหาค่าความโค้งด้วยตัวกรองเกาส์เซียนและคิวบิก บี-สไปรน์

จากการเลื่อนของจุดเด่นอันเนื่องมาจากการเปลี่ยนค่า สเกล-สเปซ ของอัลกอริทึมที่หนึ่ง ดังแสดงในภาพที่ 8 ทำให้จุดเด่นที่คำนวณได้นั้น ไม่มีเสถียรภาพ จึงจำเป็นที่จะต้องหาแนวทางในการแก้ไขปัญหาดังกล่าวให้ดีขึ้น โดยได้นำเสนอการแก้ไขปัญหาด้วยเทคนิคการใช้ตัวกรองเกาส์เซียนร่วมกับ คิวบิก บี-สไปรน์ ดังจะกล่าวต่อไปนี้

ถ้ากำหนดให้ x และ y เป็นฟังก์ชันแบบพาราเมตริกของ บี-สไปรน์ ซึ่งเขียนแทนด้วยสมการดังนี้

$$x = f(u) = a_1u^3 + a_2u^2 + a_3u + a_4 \quad (2-14a)$$

$$y = g(u) = b_1u^3 + b_2u^2 + b_3u + b_4 \quad (2-14b)$$

เมื่อ $u \in [0,1]$

และถ้าเส้นโค้งที่กำลังพิจารณาถูกประมาณด้วยฟังก์ชันของ คิวบิก บี-สไปรน์ ดังนั้น x และ y สามารถเขียนเป็นสูตรสำเร็จได้ดังนี้

$$x(u) = TM_b G_x^T \quad (2-15a)$$

$$y(u) = TM_b G_y^T \quad (2-15b)$$

เมื่อ $M_b = \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix}; T = [u^3 \quad u^2 \quad u \quad 1];$

$$G_x^T = [x_{i-1} \quad x_i \quad x_{i+1} \quad x_{i+2}] \quad \text{และ} \quad G_y^T = [y_{i-1} \quad y_i \quad y_{i+1} \quad y_{i+2}]$$

โดยมีจุดควบคุมฟังก์ชัน คิวบิก บี-สไปรน์ ทั้งหมด 4 จุด ดังนี้

$$P_{i-1} = (x_{i-1}, y_{i-1}); \quad P_i = (x_i, y_i); \quad P_{i+1} = (x_{i+1}, y_{i+1}) \quad \text{และ} \\ P_{i+2} = (x_{i+2}, y_{i+2})$$

เมื่อแทนค่า G_x^T และ G_y^T ในสมการ (2-15) จะได้ว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$a_1 = (-x_{i-1} + 3x_i - 3x_{i+1} + x_{i+2}) / 6 \quad \text{-----} \quad (2-16a)$$

$$a_2 = (x_{i-1} - 2x_i + x_{i+1}) / 2 \quad \text{-----} \quad (2-16b)$$

$$a_3 = (-x_{i-1} + x_{i+1}) / 2 \quad \text{-----} \quad (2-16c)$$

$$a_4 = (x_{i-1} + 4x_i + x_{i+1}) / 6 \quad \text{-----} \quad (2-16d)$$

$$b_1 = (-y_{i-1} + 3y_i - 3y_{i+1} + y_{i+2}) / 6 \quad \text{-----} \quad (2-17a)$$

$$b_2 = (y_{i-1} - 2y_i + y_{i+1}) / 2 \quad \text{-----} \quad (2-17b)$$

$$b_3 = (-y_{i-1} + y_{i+1}) / 2 \quad \text{-----} \quad (2-17c)$$

$$b_4 = (y_{i-1} + 4y_i + y_{i+1}) / 6 \quad \text{-----} \quad (2-17d)$$

ค่าความโค้งที่ประมาณด้วยฟังก์ชัน คิวบิก บี-สไปไรน์ สามารถคำนวณได้ดังนี้

$$f'(u) = 3a_1u^2 + 2a_2u + a_3 \quad \text{-----} \quad (2-18a)$$

$$f''(u) = 6a_1u + 2a_2 \quad \text{-----} \quad (2-18b)$$

$$g'(u) = 3b_1u^2 + 2b_2u + b_3 \quad \text{-----} \quad (2-19a)$$

$$g''(u) = 6b_1u + 2b_2 \quad \text{-----} \quad (2-19b)$$

ที่ $u=0$ จะได้ว่า

$$f'(0) = a_3 ; f''(0) = 2a_2$$

$$g'(0) = b_3 ; g''(0) = 2b_2$$

เพราะฉะนั้น ค่าความโค้งที่ $u=0$ สามารถคำนวณได้จากการแทนค่า $f'(0), f''(0), g'(0)$ และ $g''(0)$ ลงในสมการที่ (2-10) จะได้ว่า

$$k_b = \frac{2(a_3b_2 - b_3a_2)}{(a_3^2 + b_3^2)^{3/2}} \quad \text{-----} \quad (2-20)$$

แทนค่าสมการที่ (2-16) และ (2-17) ลงในสมการที่ (2-20) ได้ว่า

$$k_b = \frac{4[(x_{i+1} - x_{i-1})(y_{i-1} - 2y_i + y_{i+1}) - (y_{i+1} - y_{i-1})(x_{i+1} - 2x_i + x_{i+1})]}{[(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2]^{3/2}} \quad \text{-----} \quad (2-21)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการหาค่าความโค้งที่ (2-21) จะพบว่า สมการดังกล่าวไม่ปรากฏตัวแปร สเกล-สเปซ ของตัวกรองเกาส์เซียนอีกเช่นกัน ดังนั้น จึงต้องประสานตัวกรองเกาส์เซียนกับฟังก์ชันพหุนามตรีโก $x(t)$ และ $y(t)$ ของเส้นขอบภาพตามลำดับ ดังแสดงในอัลกอริทึมที่สอง และได้ภาพที่ 6 มาทำการประมวลผลเพื่อหาจุดเด่นของเส้นขอบภาพด้วยอัลกอริทึมที่สอง ซึ่งได้ผลดังแสดงในภาพที่ 9 และ 10

อัลกอริทึมที่สอง

- (1) การกำหนดสภาวะเริ่มต้นด้วยการกำหนดค่า $\sigma = 3$
- (2) คำนวณขนาดหน้าต่างของตัวกรองเกาส์เซียนสำหรับเก็บค่าสัมประสิทธิ์ ซึ่งสามารถคำนวณได้จาก $2N+1$ ในสมการที่ (2-11)
- (3) คำนวณหาค่า a_2, a_3, b_2 และ b_3

$$a_3 = [X(t+1, \sigma) - X(t-1, \sigma)] / 2$$

$$a_2 = [X(t+1, \sigma) + X(t-1, \sigma) - 2X(t, \sigma)] / 2$$

$$b_3 = [Y(t+1, \sigma) - Y(t-1, \sigma)] / 2$$

$$b_2 = [Y(t+1, \sigma) + Y(t-1, \sigma) - 2Y(t, \sigma)] / 2$$

เมื่อ

$$X(t, \sigma) = x(t) * g(t, \sigma)$$

$$Y(t, \sigma) = y(t) * g(t, \sigma)$$

- (4) นำค่า a_2, a_3, b_2 และ b_3 ที่คำนวณได้จาก (3) แทนลงในสมการที่ (2-20)
- (5) ทำการประสานค่าความโค้งด้วยตัวกรองเกาส์เซียนดังนี้

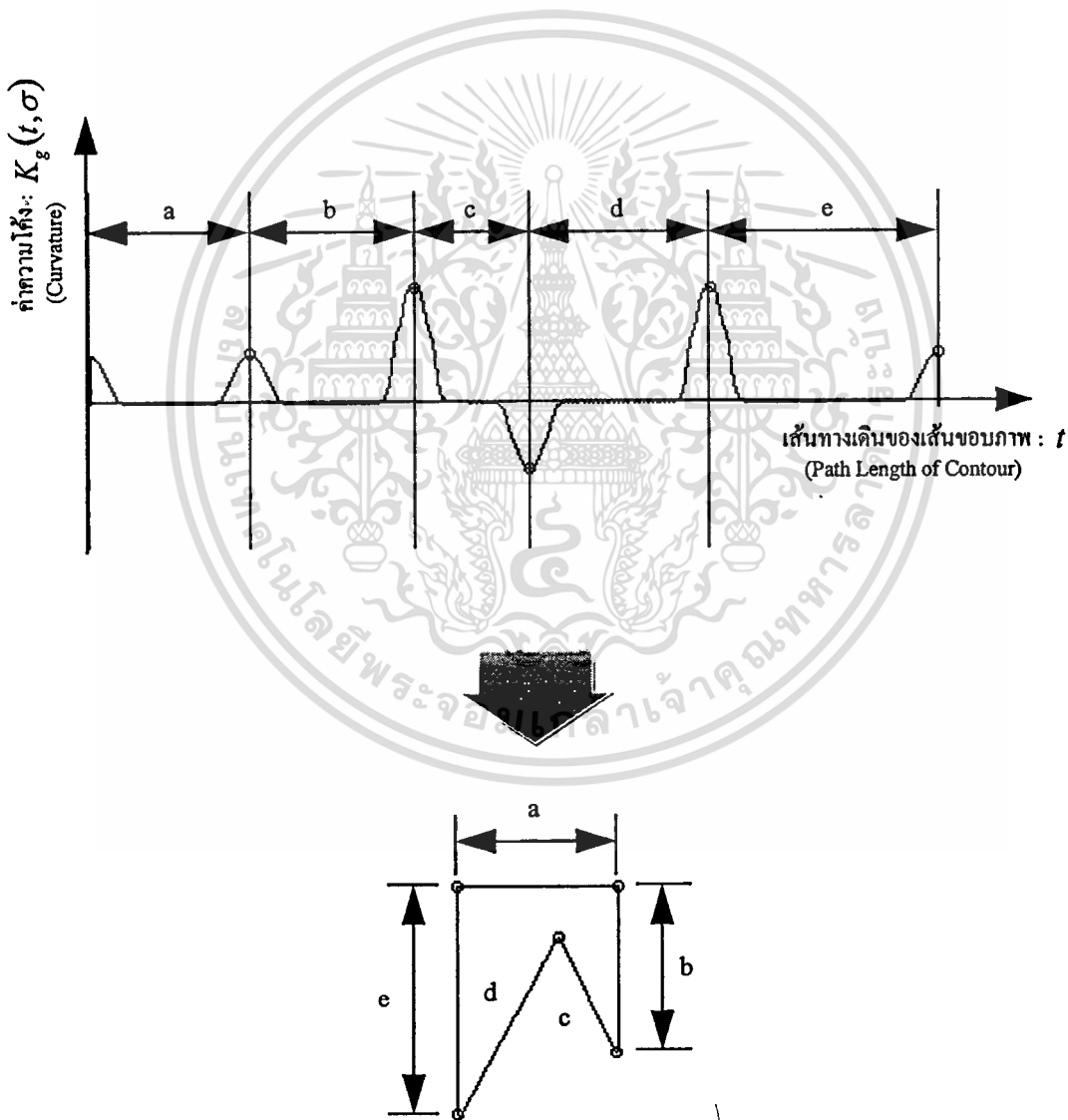
$$K_g(t, \sigma) = K_b(t, \sigma) * g(t, \sigma)$$

- (6) คำนวณหาจุดเด่นที่ปรากฏบนฟังก์ชัน $K_g(t, \sigma)$ โดยการพิจารณาเอาจุดที่มีการเปลี่ยนค่าความโค้งดังนี้

- (6.1) จุดที่มีค่าความโค้งของฟังก์ชัน $K_g(t, \sigma)$ เปลี่ยนจากค่าบวกไปเป็นลบ และมีค่ามากกว่า x

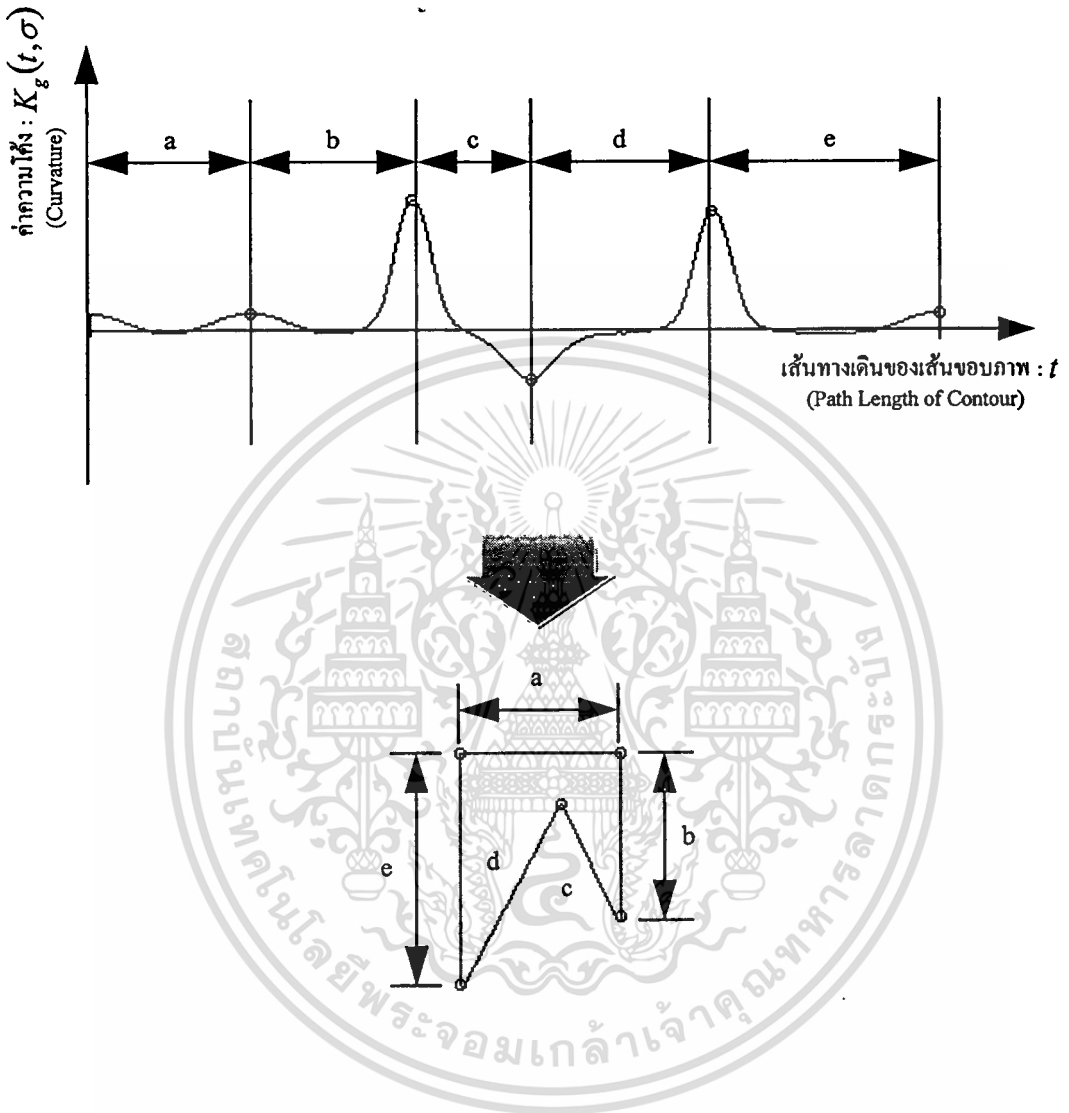
- (6.2) จุดที่มีค่าความโค้งของฟังก์ชัน $K_g(t, \sigma)$ เปลี่ยนจากค่าลบไปเป็นบวก และมีค่าน้อยกว่า $-x$
- (7) พิจารณาจำนวนของจุดเด่นที่คำนวณได้ก่อนหน้านี้ ถ้ามีค่าน้อยกว่าหรือเท่ากับจุดเด่นที่คำนวณได้ขณะนี้ ให้ข้ามไปทำขั้นตอนที่ (8) แต่ถ้ามากกว่า ให้เพิ่มค่า $\sigma = \sigma + 2$ แล้วย้อนกลับไปทำขั้นตอนที่ (2) ใหม่
- (8) หยุดการคำนวณ

ภาพที่ 9



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการใช้งานเฉพาะเท่านั้น เมื่อผู้ใดเห็นประโยชน์ในการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 10



ภาพแสดงค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นที่ $\sigma = 19$ โดยใช้อัลกอริทึมที่สอง

เมื่อเปรียบเทียบตำแหน่งของจุดเด่นที่แสดงในภาพที่ 8 และ 10 จะเห็นได้อย่างชัดเจนว่า ตำแหน่งของจุดเด่นที่คำนวณด้วยอัลกอริทึมที่สอง มีเสถียรภาพดีกว่าอัลกอริทึมที่หนึ่ง

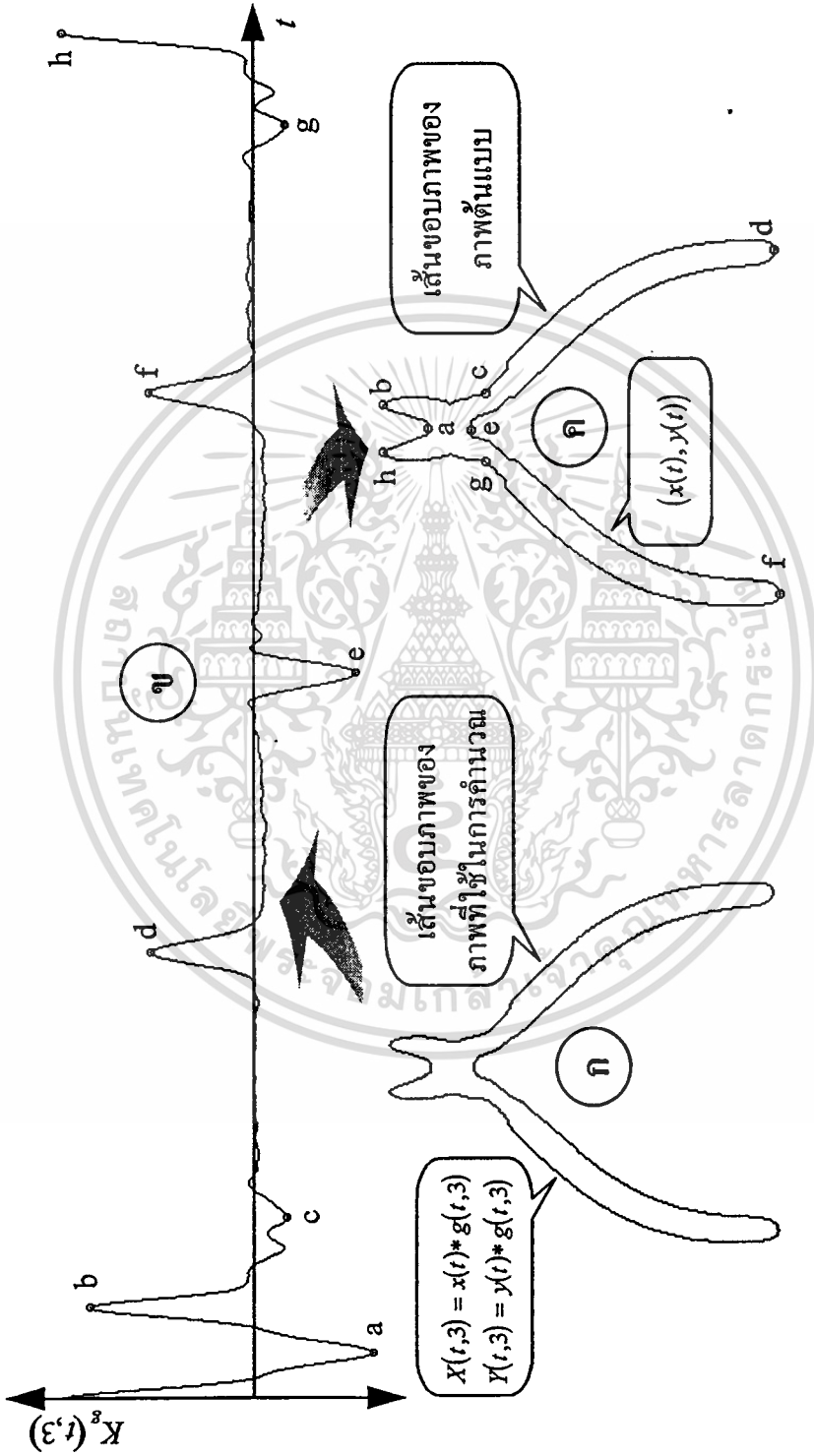
นอกจากนี้ เรายังได้ทดลองกับเส้นขอบภาพของภาพวัตถุจริงดังแสดงในภาพที่ 11 และได้ผลสรุปดังแสดงในภาพที่ 12 โดยได้ทดลองปรับค่า σ ตั้งแต่ 3-13 ซึ่งจำนวนจุดเด่นที่คำนวณได้จะมีค่าคงที่ ที่ค่า $\sigma = 7$ ถึง $\sigma = 13$ อนึ่ง ในการคำนวณหาค่าความโค้งของเส้นขอบภาพในอัลกอริทึมที่สอง เส้นขอบภาพจะถูกประสานกับสมการตัวกรองเกาส์เซียนเสียก่อน (ภาพที่ 12 (ก)) แล้วจึงนำผลลัพธ์ดังกล่าวไปคำนวณหา

ค่าความโค้ง $K_g(t, \sigma)$ (ภาพที่ 12 (ข)) จากนั้น ค่อยนำเอาตำแหน่งที่มีค่าความโค้งสูงที่ถูกกำหนดให้เป็นจุดเด่น นำมากำหนดเป็นจุดเด่นในเส้นขอบภาพต้นแบบ ดังแสดงในภาพที่ 12 (ค)

ภาพที่ 11



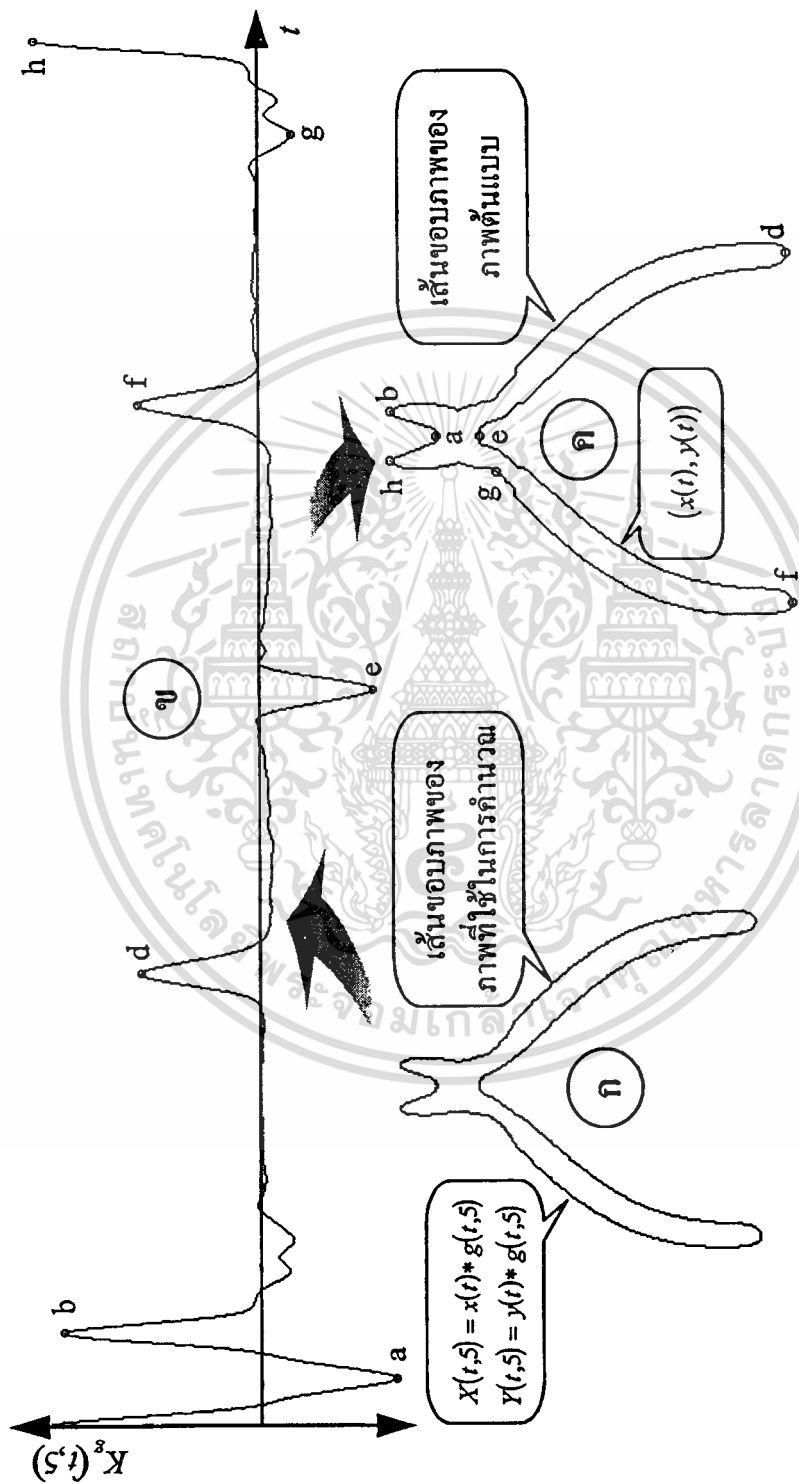
ภาพที่ 12



ภาพแสดงเส้นกราฟของค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นของภาพที่ 11 โดยใช้โอัลทอริทึมที่แสดง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

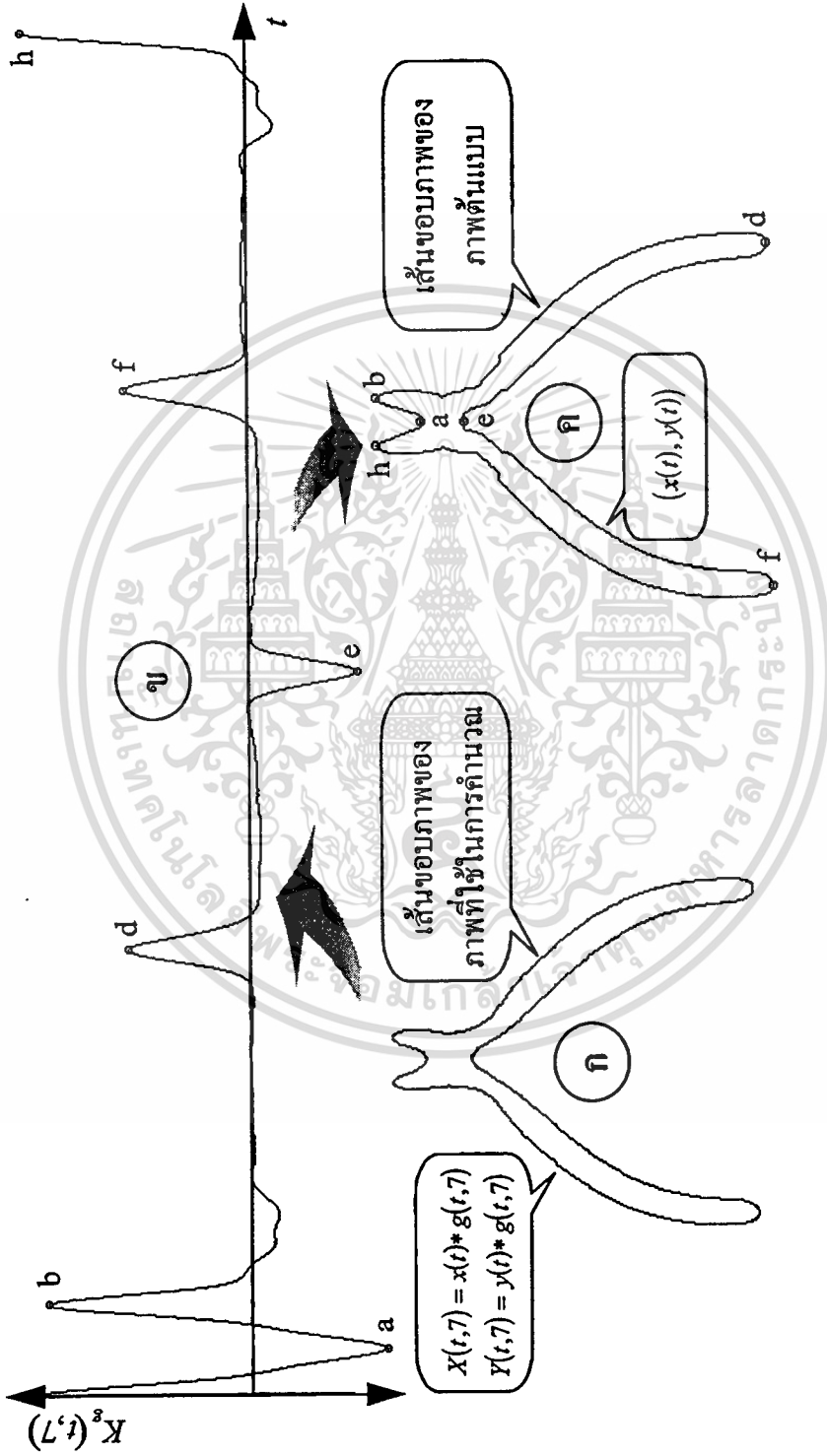
ภาพที่ 12 (ต่อ)



ภาพแสดงเส้นกราฟของค่าความถี่ที่ใช้กำหนดตำแหน่งของจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

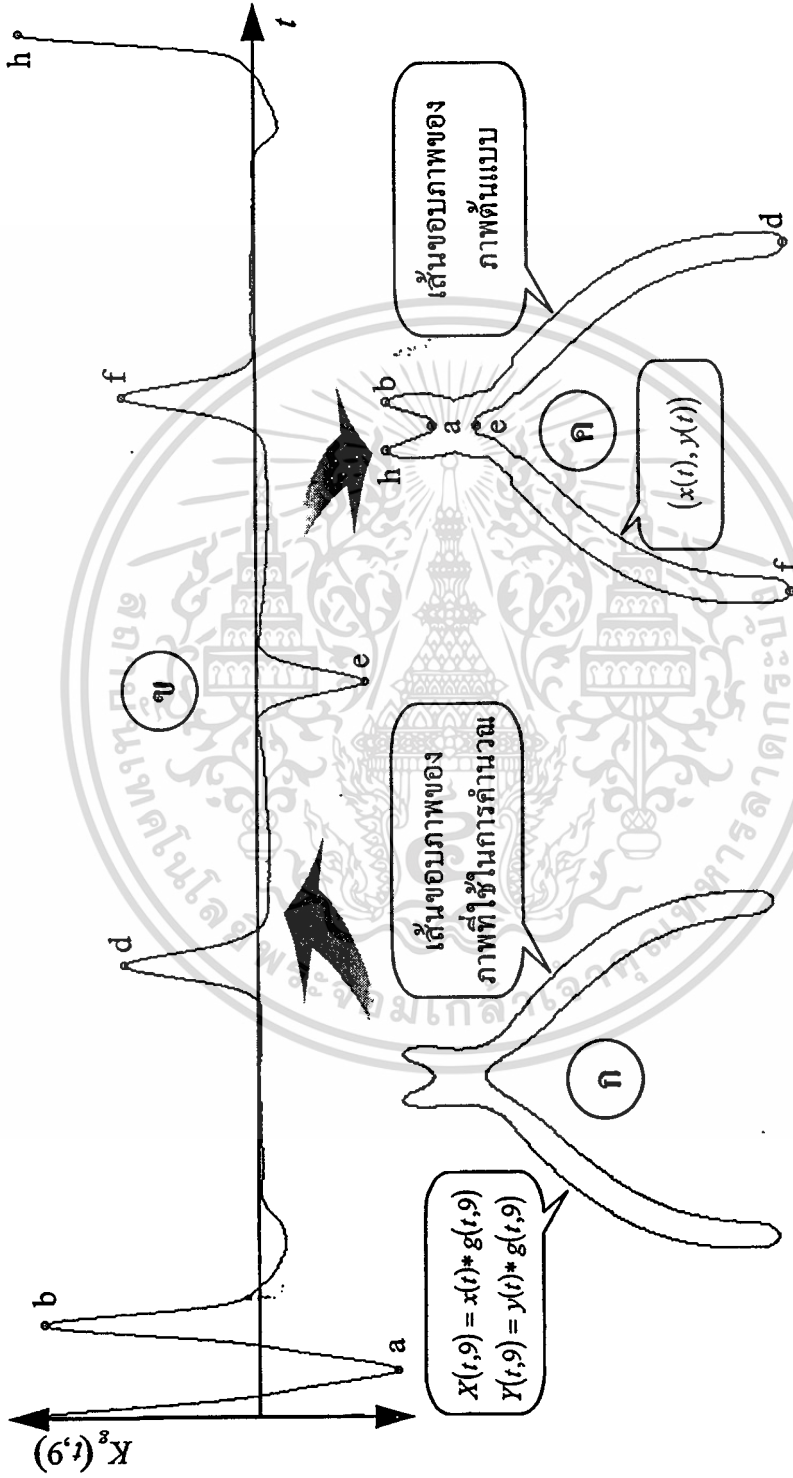
ภาพที่ 12 (ต่อ)



ภาพแสดงเส้นกราฟของค่าความโค้งที่ใช้กำหนดตำแหน่งจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง

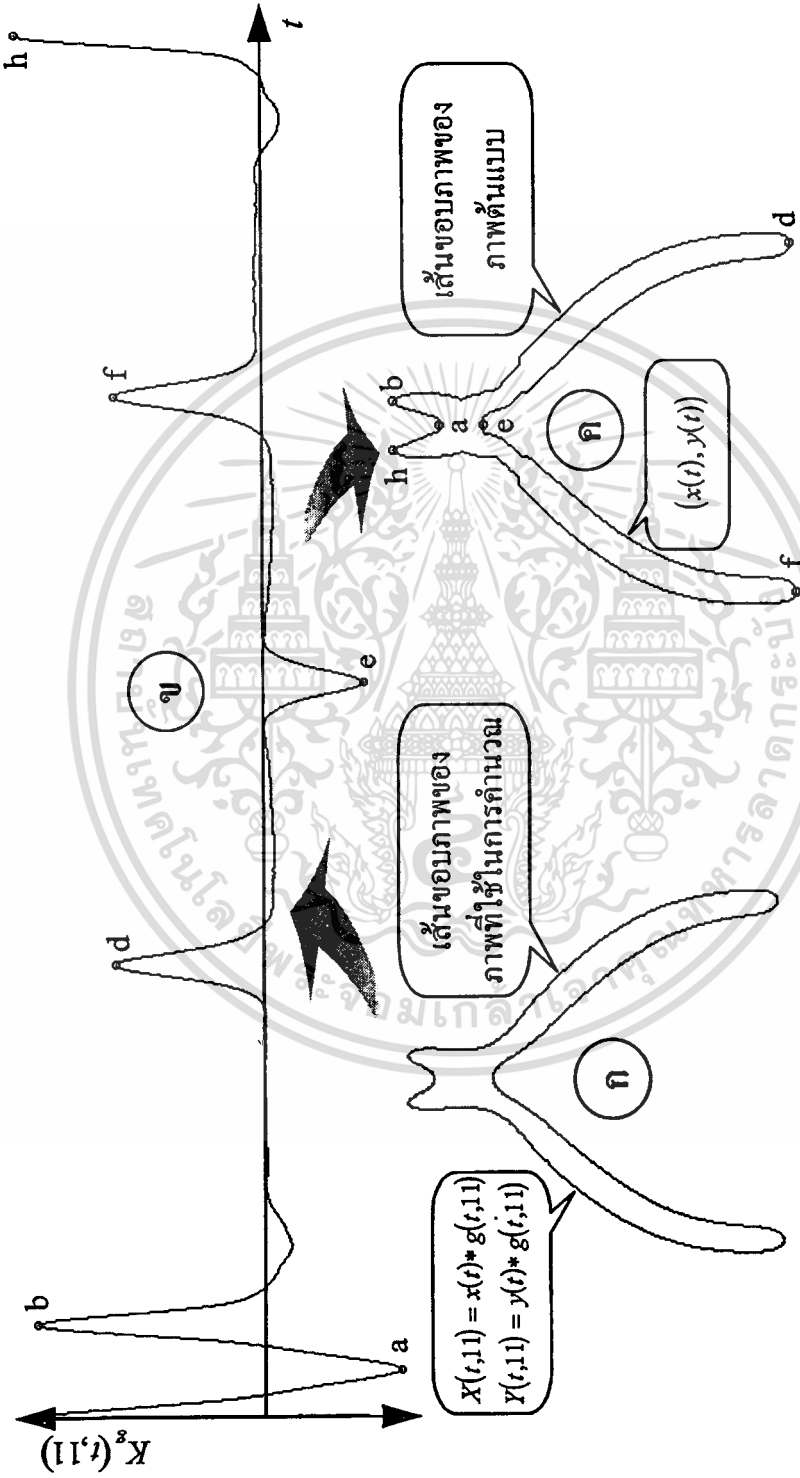
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 12 (ต่อ)



ภาพแสดงเส้นกราฟของค่าความถี่ที่ใช้กำหนดตำแหน่งของจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง

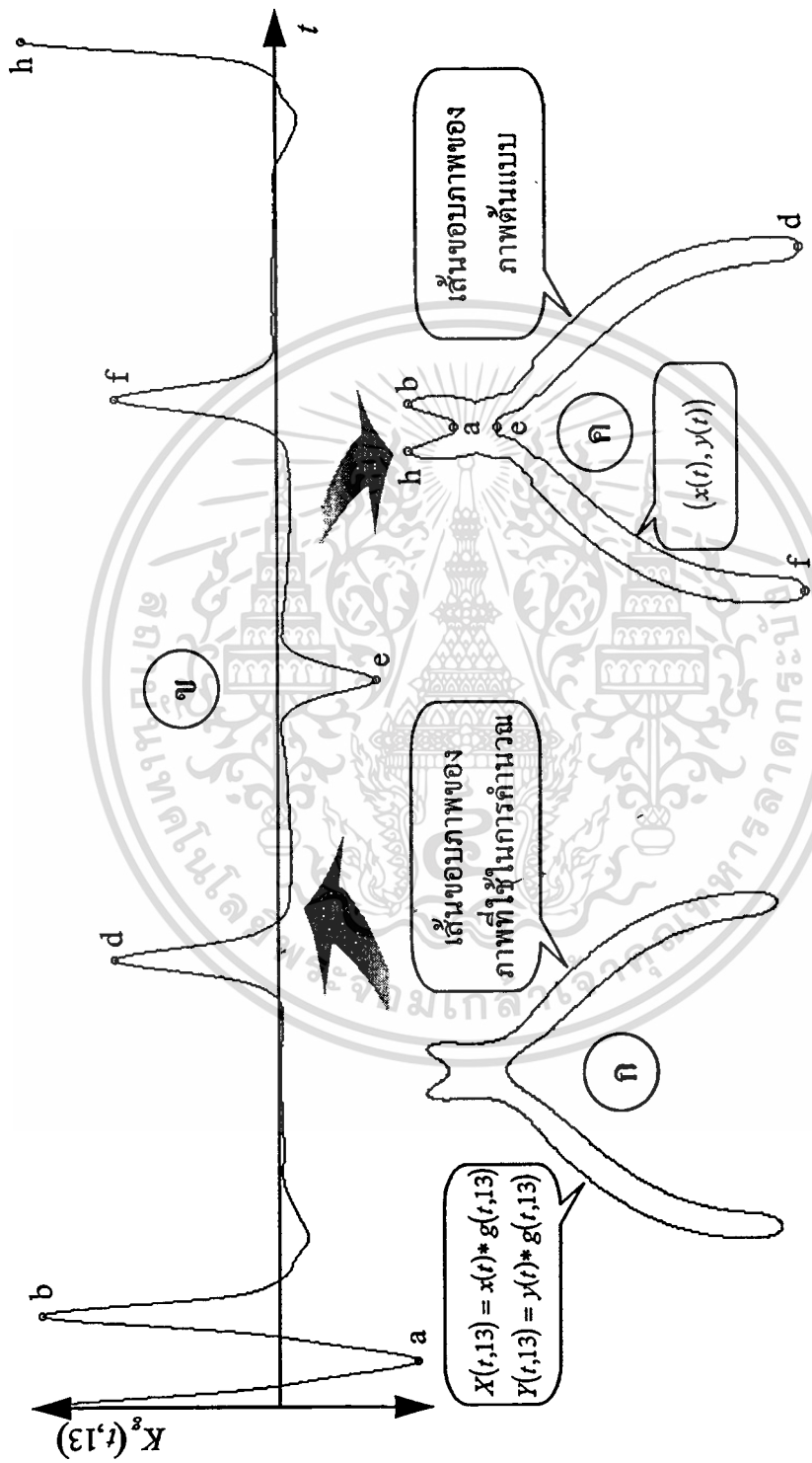
ภาพที่ 12 (ต่อ)



ภาพแสดงเส้นกราฟของค่าความโค้งที่ใช้กำหนดตำแหน่งของจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

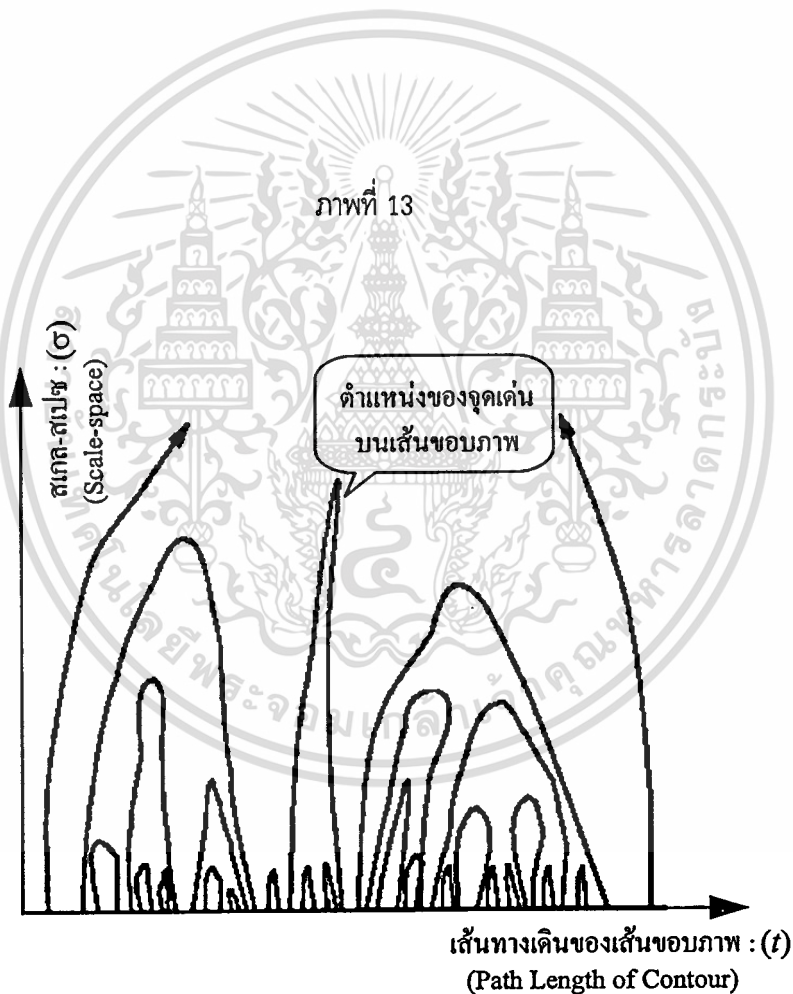
ภาพที่ 12 (ต่อ)



ภาพแสดงเส้นกราฟของค่าความโค้งที่ใช้กำหนดตำแหน่งจุดเด่นของภาพที่ 11 โดยใช้อัลกอริทึมที่สอง

4. สเกล-สเปซ อิมเมจ

สเกล-สเปซ อิมเมจ คือ การแสดงตำแหน่งของจุดเด่นบนเส้นขอบภาพในระนาบ (t, σ) เมื่อ t คือ เส้นทางเดินของเส้นขอบภาพ และ σ คือ สเกล-สเปซ ดังแสดงในภาพที่ 13 สเกล-สเปซ อิมเมจ สามารถนำมาเป็นเครื่องมือสำหรับวิเคราะห์ เพื่อดูพฤติกรรมเคลื่อนที่ของจุดเด่นที่ปรากฏบนเส้นขอบภาพของภาพวัตถุนั้น ๆ ว่ามีการเคลื่อนที่ของจุดเด่นเป็นเช่นไร ซึ่งโดยพฤติกรรมของ สเกล-สเปซ อิมเมจ เมื่อมีการเพิ่มของสเกล-สเปซ ให้มีค่าเพิ่มมากขึ้น จะทำให้จำนวนของจุดเด่นที่ปรากฏบนเส้นขอบภาพมีจำนวนลดลง จากอัลกอริทึมที่หนึ่งและอัลกอริทึมที่สองที่ใช้ในการคำนวณหาจุดเด่นที่ผ่านมา สามารถนำผลที่แสดงถึงตำแหน่งของจุดเด่นมาวิเคราะห์ด้วย สเกล-สเปซ อิมเมจ ได้เช่นกัน ดังแสดงในภาพที่ 14 และ 15



ภาพแสดง สเกล-สเปซ อิมเมจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 14



ภาพ สเกล-สเปซ อิมเมจ แสดงการเคลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่หนึ่ง

ภาพที่ 15



ภาพ สเกล-สเปซ อิมเมจ แสดงการเคลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่สอง

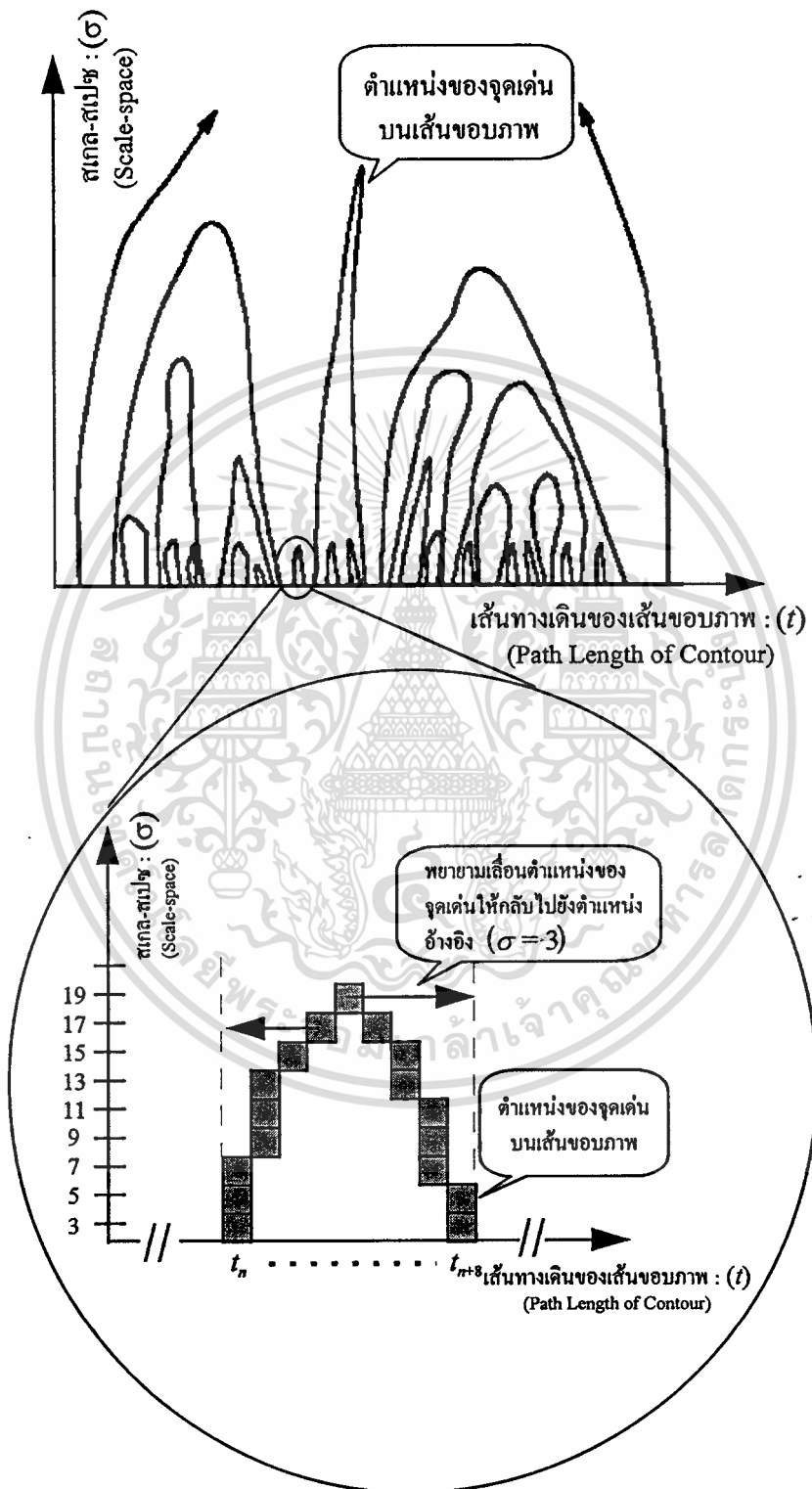
5. การตรึงจุดเด่นให้คงที่โดยใช้ซอฟต์แวร์

ถึงแม้ว่าจุดเด่นที่คำนวณได้โดยใช้อัลกอริทึมที่สอง จะมีการเคลื่อนของจุดเด่นน้อยกว่าอัลกอริทึมที่หนึ่งก็ตาม แต่เมื่อนำตำแหน่งของจุดเด่นที่คำนวณได้จากอัลกอริทึมที่สองมาพล็อตในระนาบ (t, σ) เพื่อดูพฤติกรรมเคลื่อนของจุดเด่นด้วย สเกล-สเปซ อิมเมจ ดังแสดงในภาพที่ 15 จะเห็นได้อย่างชัดเจนว่า การคำนวณหาจุดเด่นด้วยอัลกอริทึมดังกล่าว ยังคงมีการเคลื่อนของจุดเด่นอยู่ ดังนั้น การตรึงจุดเด่นให้คงที่โดยใช้ซอฟต์แวร์ จึงเป็นอีกแนวทางหนึ่งที่น่าสนใจแก้ไขปัญหาคือการเคลื่อนของจุดเด่นดังที่ได้กล่าวมา โดยอาศัยเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทฤษฎีของ สเกล-สเปซ อิมเมจ มาควบคุมการทำงานในแต่ละครั้งของการเปลี่ยนค่า 'สเกล-สเปซ' ซึ่งหลักการ
 ทำงานของอัลกอริทึมการตรึงจุดเด่นให้คงที่โดยการใช้ซอฟต์แวร์ คือ ความพยายามที่จะเลื่อนตำแหน่งของจุด
 เด่นที่มีค่า สเกล-สเปซ ที่สูงกว่า ไปยังตำแหน่งของจุดเด่นที่มีค่า สเกล-สเปซ ที่ต่ำกว่า ดังแสดงในภาพที่ 16
 และยังคงใช้อัลกอริทึมที่หนึ่งหรืออัลกอริทึมที่สอง ในการคำนวณหาจุดเด่นในแต่ละค่าของ สเกล-สเปซ โดย
 ในสภาวะเริ่มต้นของการทำงาน จะกำหนดค่า สเกล-สเปซ ต่ำ ๆ เพื่อเป็นตำแหน่งอ้างอิงสำหรับค่า สเกล-สเปซ
 ที่สูงกว่า ทั้งนี้เนื่องจากตำแหน่งของจุดเด่นที่เกิดขึ้นบนเส้นขอบภาพนั้น จะมีความเที่ยงตรงสูงในกรณีที่มีค่า
 สเกล-สเปซ มีค่าต่ำ ๆ ดังนั้น ในการกำหนดค่าเริ่มต้นสำหรับ สเกล-สเปซ จึงเลือกใช้ค่า $\sigma=3$ เป็นค่าเริ่ม
 ต้น แล้วค่อย ๆ เพิ่มค่า สเกล-สเปซ ขึ้นเป็น $\sigma=3,5,7,9,\dots$



ภาพที่ 16



ภาพแสดงวิธีการตรึงจุดเด่นให้คงที่โดยอาศัย สเกล-สเปซ อิมแมจ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์โดยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่ให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการประมวลผลของอัลกอริทึมสำหรับการเรียงจุดเด่นให้คงที่ จะต้องจัดเตรียมอาเรย์ไว้สำหรับเก็บค่าตำแหน่งของจุดเด่นทั้งหมด 3 ชุด คือ $P_1(i)$, $P_2(i)$ และ $P_3(i)$ โดยที่ $P_3(i)$ จะทำหน้าที่เก็บตำแหน่งของจุดเด่นที่ค่า $\sigma = 3$ และ $P_2(i)$ จะเก็บตำแหน่งของจุดเด่นที่เกิดขึ้น ณ ค่า สเกล-สเปซ ปัจจุบัน ส่วน $P_1(i)$ จะเก็บตำแหน่งของจุดเด่นที่เกิดขึ้นก่อนหน้านี ($\sigma - 2$) ซึ่งวิธีการแก้ปัญหาการเลื่อนของจุดเด่น ดังที่เกิดขึ้นในอัลกอริทึมที่หนึ่ง และอัลกอริทึมที่สอง สามารถสรุปเป็นอัลกอริทึมได้ดังนี้

อัลกอริทึมที่สาม

(1) คำนวณหาจุดเด่นด้วยอัลกอริทึมที่หนึ่งหรืออัลกอริทึมที่สอง โดยการกำหนดค่าเริ่มต้นด้วย $\sigma = 3$ แล้วทำการคำนวณหาจุดเด่นที่เกิดขึ้นบนเส้นโค้ง $K_g(t, \sigma)$ สำหรับอัลกอริทึมที่สองหรือ $K(t, \sigma)$ สำหรับอัลกอริทึมที่หนึ่ง ซึ่งจุดเด่นจะอยู่ในตำแหน่งของเส้นขอบภาพที่ $t = n_1, n_2, n_3, \dots, n_{M_1}$ จากนั้นเก็บตำแหน่งของจุดเด่นต่าง ๆ เอาไว้ในอาเรย์ $P_3(i)$ และ $P_1(i)$ เมื่อ $i = 1, 2, 3, \dots, M_1$

(2) เพิ่มค่าของ สเกล-สเปซ ขึ้นอีก 2 ($\sigma = \sigma + 2$) แล้วทำการคำนวณหาจุดเด่นบนเส้นโค้ง $K_g(t, \sigma)$ หรือ $K(t, \sigma)$ ซึ่งตำแหน่งของจุดเด่นจะอยู่บนเส้นขอบภาพที่ $t = m_1, m_2, m_3, \dots, m_{M_2}$ จากนั้นเก็บตำแหน่งของจุดเด่นต่าง ๆ เอาไว้ในอาเรย์ $P_2(i)$ เมื่อ $i = 1, 2, 3, \dots, M_2$ โดยทฤษฎีของสเกล-สเปซ จำนวนของจุดเด่นจะลดลง เมื่อค่า สเกล-สเปซ เพิ่มขึ้น นั่นคือ $M_2 \leq M_1$

(3) พยายามเลื่อนจุดเด่นที่ได้จากข้อ (2) ให้กลับมาอยู่ในตำแหน่งเดิมที่ค่า สเกล-สเปซ เท่ากับ 3 ดังแสดงในภาพที่ 16 โดยเขียนเป็น อัลกอริทึมได้ดังนี้

```

For i = 1 to M2
  Begin
    For j = 1 to M1
      Begin
        Search P1(k) So that
          Min
            k |P1(j) - P2(i)|
      End
      P3(i) = P3(k)
    End
  End.

```

(4) นำตำแหน่งของจุดเด่นที่ได้ใน $P_2(i)$ ย้ายไปเก็บไว้ใน $P_1(i)$ เพื่อเป็นตำแหน่งอ้างอิงในการเปรียบเทียบกับตำแหน่งของจุดเด่นชุดใหม่ ที่ได้จากการเพิ่มค่า $\sigma = \sigma + 2$ ซึ่งสามารถเขียนเป็นอัลกอริทึมได้ดังนี้

```

For i = 1 to M2
  Begin
    P1(i) = P2(i)
  End.

```

(5) ถ้าจำนวนของจุดเด่นที่คำนวณได้ก่อนหน้านี้ มีค่ามากกว่าหรือเท่ากับจุดเด่นที่คำนวณได้ขณะนี้ ให้หยุดการคำนวณ แต่ถ้าน้อยกว่า ให้เพิ่มค่า $\sigma = \sigma + 2$ แล้วย้อนกลับไปทำขั้นตอนที่ (2) ใหม่

เมื่อทำการพล็อตตำแหน่งของจุดเด่นที่ผ่านการประมวลผลด้วยอัลกอริทึมที่สามบนระนาบ (t, σ) ซึ่งเรียกว่า “สเกล-สเปซ อิมเมจ” ได้ผลดังแสดงในภาพที่ 17 ซึ่งจะเห็นได้อย่างชัดเจนว่า ตำแหน่งของจุดเด่นที่คำนวณหาด้วยอัลกอริทึมดังกล่าวมีเสถียรภาพดีที่สุดในกลุ่มของอัลกอริทึมทั้งหมดที่ได้นำเสนอมา



ภาพ สเกล-สเปซ อิมเมจ แสดงการเคลื่อนของจุดเด่นของภาพที่ 6 โดยใช้อัลกอริทึมที่สาม

สรุปและเปรียบเทียบผลการคำนวณหาจุดเด่นของแต่ละวิธี

จากอัลกอริทึมทั้งสามที่ได้กล่าวมาแล้วนั้น ได้พิสูจน์ให้เห็นอย่างชัดเจนในการคำนวณหาจุดเด่นและปัญหาที่เกิดขึ้น เนื่องจากการเปลี่ยนค่า สเกล-สเปซ ซึ่งพอที่จะสรุปข้อได้เปรียบและเสียเปรียบของแต่ละวิธีได้ ดังนี้ อัลกอริทึมที่หนึ่งเป็นอัลกอริทึมที่มีการเคลื่อนของจุดเด่นค่อนข้างชัดเจน เมื่อมีการเพิ่มค่า สเกล-สเปซ ให้ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำมาเผยแพร่ไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สูงขึ้น ดังแสดงในภาพที่ 8 และภาพที่ 14 ทำให้จุดเด่นที่คำนวณได้นั้นไม่มีเสถียรภาพ อีกทั้งยังใช้เวลาในการประมวลผลสูง เพราะจะต้องทำการประสานเส้นขอบภาพกับอนุพันธ์ย่อยอันดับที่หนึ่งและอันดับที่สองของการเกาส์เซียนใหม่ทุกครั้ง ที่มีการเปลี่ยนค่า สเกล-สเปซ จากข้อด้อยของอัลกอริทึมหนึ่ง จึงเปลี่ยนเทคนิคมา ใช้การประมาณค่าฟังก์ชันของเส้นโค้งด้วยคิวบิก บี-สไปรน์ จากเทคนิคดังกล่าว ทำให้สามารถลดเวลาที่ต้องใช้ในการประมวลผล และจุดเด่นที่คำนวณได้ก็มีเสถียรภาพมากขึ้น ดังแสดงในภาพที่ 10 และ 15 ส่วนวิธีสุดท้ายคือ อัลกอริทึมที่สาม เป็นการตรึงจุดเด่นให้คงที่โดยการใช้ซอฟต์แวร์ ซึ่งมีข้อดี คือ สามารถตรึงจุดเด่นให้คงที่ตลอดเวลา เมื่อมีการเปลี่ยนค่า สเกล-สเปซ ทำให้จุดเด่นที่คำนวณได้มีเสถียรภาพ อัลกอริทึมที่สามมักจะถูกใช้งานร่วมกับอัลกอริทึมที่หนึ่ง หรืออัลกอริทึมที่สองเสมอ นั่นคือ ประสิทธิภาพของอัลกอริทึมในการคำนวณหาจุดเด่น เป็นการนำข้อดีของแต่ละอัลกอริทึมมาผสมผสานร่วมกันในการประมวลผล จึงจะได้ประสิทธิภาพสูงสุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

โครงข่ายเซลล์ประสาทเทียม

บทนำ

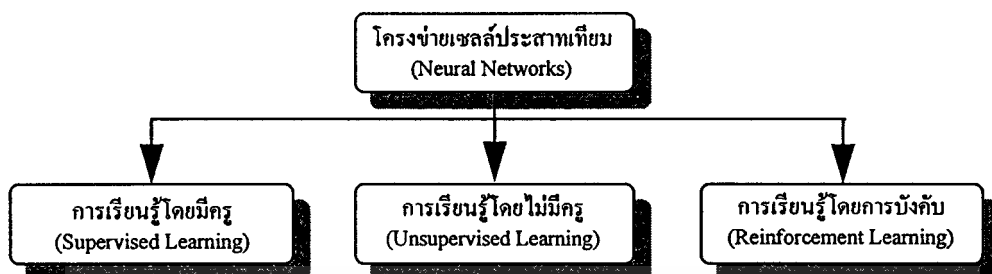
ในบทนี้จะได้กล่าวถึงรูปแบบของเซลล์ประสาท (Neural) และการวางตัวของเซลล์ประสาทจนกลายเป็นโครงข่าย ตลอดจนคณิตศาสตร์ที่ใช้ในการคำนวณ และการตัดสินใจของโครงข่ายเหล่านี้ จากอดีตจนถึงปัจจุบัน ได้มีผู้คิดค้นและพัฒนา รูปแบบของเซลล์ประสาทที่ขึ้นมามากมาย (มากกว่า 10 รูปแบบ) บางรูปแบบมีการพัฒนาขึ้นมาเพื่อจำลองการทำงานของสมองมนุษย์ ให้สามารถอธิบายการคิดและการเรียนรู้ของมนุษย์ว่าเกิดขึ้นได้อย่างไร และก็ยังมียังอีกหลายรูปแบบเช่นเดียวกันที่ถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหาทางวิศวกรรม โดยไม่มีพื้นฐานทางชีววิทยาเข้ามาเกี่ยวข้องแม้แต่น้อย รูปแบบเหล่านี้ถูกเรียกว่า “โครงข่ายเซลล์ประสาทเทียม” (Artificial Neural Networks) ในบทนี้จะได้ยกตัวอย่างรูปแบบโครงข่ายเซลล์ประสาทเทียมที่นิยมใช้กัน โดยเฉพาะอย่างยิ่งในการแก้ไขปัญหาทางวิศวกรรมและวิทยาการคอมพิวเตอร์ โดยในเนื้อหาหลักจะได้กล่าวถึงวิธีการคำนวณด้วยวิธีพิเศษ โดยอาศัยการคำนวณอย่างง่าย ๆ ของแต่ละเซลล์ประสาทที่มีการเชื่อมโยงถึงกันทางเรขาคณิต ซึ่งโดยทั่วไปแล้ว เราจะไม่พบการทำงานในลักษณะนี้กับระบบการทำงานของเซลล์ประสาทหรือสมองของมนุษย์และสัตว์แต่อย่างใด วิธีการเรียนรู้ (Learning) หรือค่าพารามิเตอร์ที่ใช้ในการปรับค่าของเซลล์ประสาท ก็จะไม่มีส่วนเกี่ยวข้องกับเชิงชีววิทยาทั้งสิ้น

ประเภทของโครงข่ายเซลล์ประสาทเทียมสามารถแบ่งออกมาได้หลายประเภท โดยขึ้นอยู่กับวิธีและกฎเกณฑ์ที่ใช้แบ่ง เช่นการแบ่งตาม *กฎการเรียนรู้และเรขาคณิต (Learning Rules and Geometry)* หรืออาจจะแบ่งตาม *การใช้งาน (Application)* ในกรณีที่ใช้แบ่งตาม *กฎการเรียนรู้และเรขาคณิต* สามารถแบ่งได้สามรูปแบบ คือ

- (ก) การเรียนรู้โดยมีครู (Supervised Learning)
- (ข) การเรียนรู้โดยไม่มีครู (Unsupervised Learning)
- (ค) การเรียนรู้โดยการบังคับ (Reinforcement Learning)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 18

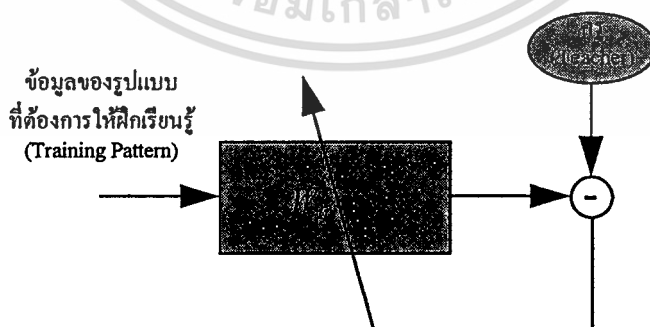


ภาพแสดงการแบ่งประเภทโครงข่ายเซลล์ประสาทเทียมตามกฎการเรียนรู้และเรขาคณิต

1. การเรียนรู้โดยมีครู (Supervised Learning)

หลักการทำงานของ การเรียนรู้โดยมีครู คือ ข้อมูลที่ต้องการให้ฝึกเรียนรู้จะเป็นคู่อินพุต/เอาต์พุต ดังแสดงไว้ในภาพที่ 19 ซึ่งจะถูกป้อนเข้าสู่โครงข่ายอย่างซ้ำ ๆ และระบบจะมีกฎเกณฑ์สำหรับปรับค่าพารามิเตอร์จนกระทั่งโครงข่ายให้เอาต์พุตออกมามีค่า "ใกล้เคียง" (Closed Enough) กับเอาต์พุตของข้อมูลที่ต้องการให้ฝึกเรียนรู้

ภาพที่ 19



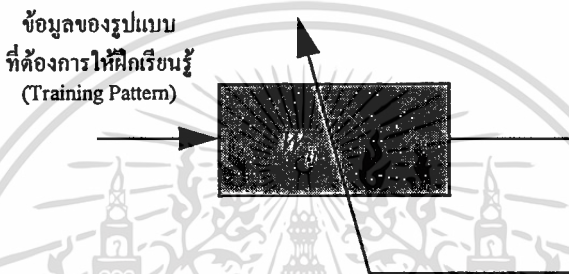
ภาพแสดงรูปแบบของการเรียนรู้โดยมีครู

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การเรียนรู้โดยไม่มีครุ (Unsupervised Learning)

หลักการงานของการเรียนรู้โดยไม่มีครุ คือ ข้อมูลที่ต้องการให้ฝึกเรียนรู้ จะมีเพียงข้อมูลส่วนที่เป็นอินพุตแต่เพียงอย่างเดียว ดังแสดงไว้ในภาพที่ 20 และโครงข่ายก็จะปรับค่าพารามิเตอร์ของมันเอง จนกระทั่งเมื่อให้อินพุตที่เหมือนกัน จะได้ค่าเอาต์พุตที่เหมือนกันด้วย

ภาพที่ 20



ภาพแสดงรูปแบบของการเรียนรู้โดยไม่มีครุ

3. การเรียนรู้โดยการบังคับ (Reinforcement Learning)

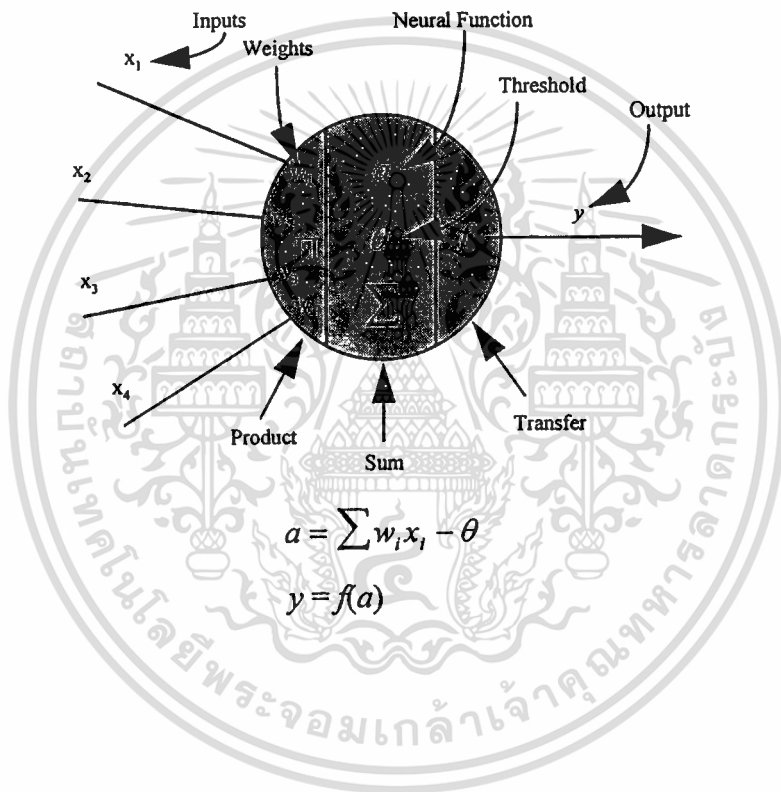
หลักการงานของการเรียนรู้โดยการบังคับ คือ ข้อมูลจะถูกป้อนให้โครงข่าย โดยบางครั้งอาจจะป้อนสัญญาณเพื่อบอกให้ทราบล่วงหน้าว่า อินพุตที่ป้อนให้ นั้น มันเคยล้มเหลว (Fail) หรือ สำเร็จ (Success) มาก่อน ซึ่งโครงข่ายจะปรับค่าพารามิเตอร์ เพื่อให้หลีกเลี่ยงความล้มเหลว หรือเพิ่มความสำเร็จ ทั้งนี้เพราะการตอบสนองที่ถูกต้องในแต่ละครั้งไม่อาจรู้ได้ ดังนั้น ความน่าจะเป็น (Probability) ของความล้มเหลว หรือ สำเร็จ จะถูกประมาณการไว้ก่อน และค่าพารามิเตอร์ก็จะถูกปรับตามนั้น

แบบจำลองของเซลล์ประสาท

เซลล์ประสาทจะเป็นส่วนที่เล็กที่สุด ซึ่งทำหน้าที่รับเอาข้อมูลเข้ามาพร้อม ๆ กันได้มากกว่าหนึ่งอินพุต แต่จะให้ข้อมูลออกเพียงหนึ่งเอาต์พุตเท่านั้น และเซลล์ประสาทเหล่านี้ เมื่อนำมาต่อเชื่อมโยงถึงกันก็จะกลายเป็นโครงข่ายประสาทเทียม ซึ่งสามารถเรียนรู้และปรับตัวได้เหมือนมนุษย์ โดยไม่ต้องอาศัยการป้อนข้อมูลล่วงหน้า การคำนวณจะกระทำโดยอัตโนมัติ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นโครงข่ายเซลล์ประสาทขึ้น โดยที่เซลล์ประสาทแต่ละเซลล์อาจจะถูกแทนด้วยสมการคณิตศาสตร์อย่างง่าย ๆ หรืออาจเป็นอุปกรณ์ธรรมดา ๆ (เมื่อนำไปประยุกต์ใช้งานฮาร์ดแวร์) ด้วยความสามารถในการเชื่อมโยงถึงกันและความเป็นไปได้ในการคำนวณเชิงขนาน ทำให้โครงข่ายเซลล์ประสาทเทียมเป็นเครื่องมือที่ดีในการแก้ไขปัญหทางวิศวกรรม ซึ่งได้แสดงแบบจำลองของเซลล์ประสาทให้เห็นดังในภาพที่ 21

ภาพที่ 21



ภาพแสดงแบบจำลองของเซลล์ประสาท

เมื่อ	X_i	คือ ข้อมูลของอินพุต
	W_i	คือ ค่าน้ำหนัก
	θ	คือ ค่า Threshold
	$a = \sum w_i x_i - \theta$	คือ ฟังก์ชันเซลล์ประสาท (Neural Function)
	f	คือ ค่า Scaling หรือ ฟังก์ชันถ่ายโอน (Transfer Function)
	y	คือ ค่าเอาต์พุตที่ได้จากเซลล์ประสาท
		หรือ $y = f(\sum w_i x_i - \theta)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพที่ 21 ถ้า $W > 0$ ค่าของข้อมูลอินพุตจะยังคงอยู่ เป็นเหตุให้เซลล์ประสาทให้ค่าเอาต์พุตออกมาที่ y มีค่าเป็นบวก ในทำนองเดียวกัน ถ้า $W_i < 0$ ค่าของข้อมูลอินพุตจะหายไป เป็นเหตุให้เซลล์ประสาทให้ค่าเอาต์พุตออกมามีค่าเป็น "0" หรือเป็นค่าติดลบ

ส่วนค่า Scaling หรือฟังก์ชันถ่ายโอน (f) โดยทั่วไปแล้วจะไม่มีรูปแบบ ซึ่งได้ยกตัวอย่างให้ดู 6 รูปแบบ ดังที่ได้แสดงไว้ในภาพที่ 22 ได้แก่

ฟังก์ชันที่ให้ค่าเป็นได้แค่ "0" หรือ "1" เท่านั้น ซึ่งเราจะเรียกฟังก์ชันนี้ว่า "Unit Step Function" และในบางครั้งอาจจะเขียนแทนด้วยสัญลักษณ์ $U(a)$

$$f(a) = \begin{cases} 1 & \text{if } a > 0 \\ 0 & \text{if } a < 0 \end{cases}$$

ฟังก์ชันที่ให้ค่าเป็นได้แค่ "-1" หรือ "1" เราจะเรียกฟังก์ชันนี้ว่า "Signum Function"

$$f(a) = \begin{cases} +1 & \text{if } a > 0 \\ -1 & \text{if } a < 0 \end{cases}$$

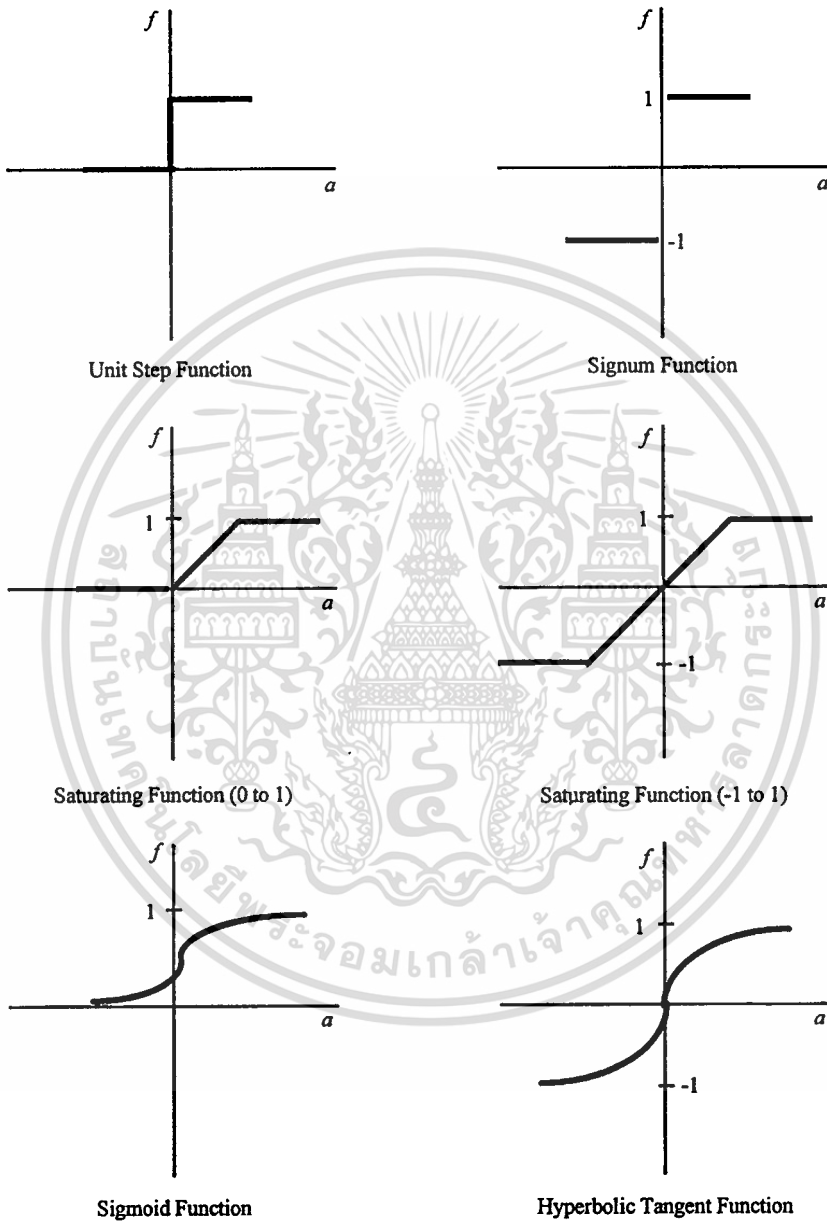
ฟังก์ชันที่อิ่มตัวระหว่างค่า 0 ถึง 1 เรียกฟังก์ชันนี้ว่า "Saturating Function (0 to 1)"

$$f(a) = \begin{cases} 1 & \text{if } a > 0 \\ a & \text{if } 0 < a < 1 \\ 0 & \text{if } a < 0 \end{cases}$$

ฟังก์ชันที่อิ่มตัวระหว่างค่า -1 ถึง 1 เรียกฟังก์ชันนี้ว่า "Saturating Function (-1 to 1)"

$$f(a) = \begin{cases} 1 & \text{if } a > 1 \\ a & \text{if } -1 < a < 1 \\ -1 & \text{if } a < -1 \end{cases}$$

ภาพที่ 22



ภาพแสดงฟังก์ชันถ่ายโอนแบบต่าง ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชัน Sigmoid

$$f(a) = \frac{1}{1 + \exp(-a)}$$

ฟังก์ชันข้างบนนี้เป็นที่นิยมนำไปประยุกต์ใช้งานกันเป็นที่แพร่หลาย ทั้งนี้เนื่องจากฟังก์ชันดังกล่าวสามารถหาค่าอนุพันธ์ได้ทุกค่า และมีขอบเขตของค่าอยู่ที่ 0 ถึง 1

$$\frac{df(a)}{da} = f(a)[1 - f(a)]$$

ด้วยคุณสมบัติที่โดดเด่นดังที่ได้กล่าวมาแล้วนั้น ฟังก์ชัน Sigmoid จึงถูกนำไปสร้างเป็นอัลกอริทึมสำหรับการเรียนรู้ โดยใช้หลักการของ Gradient เพื่อให้ง่ายต่อการประยุกต์ใช้งาน ส่วนการแปรเปลี่ยนค่าของฟังก์ชัน Sigmoid จะใช้ฟังก์ชันถ้ายไฮเพอร์โบลิก Hyperbolic Tangent โดยที่ค่าที่ได้ จะมีขอบเขตอยู่ระหว่าง -1 ถึง 1 และยังคงหาค่าอนุพันธ์ได้

$$f(a) = \frac{1 - \exp(2a)}{1 + \exp(-2a)}$$

$$\frac{df(a)}{da} = [1 + f(a)][1 - f(a)]$$

โครงข่ายเซลล์ประสาทเทียมแบบต่าง ๆ

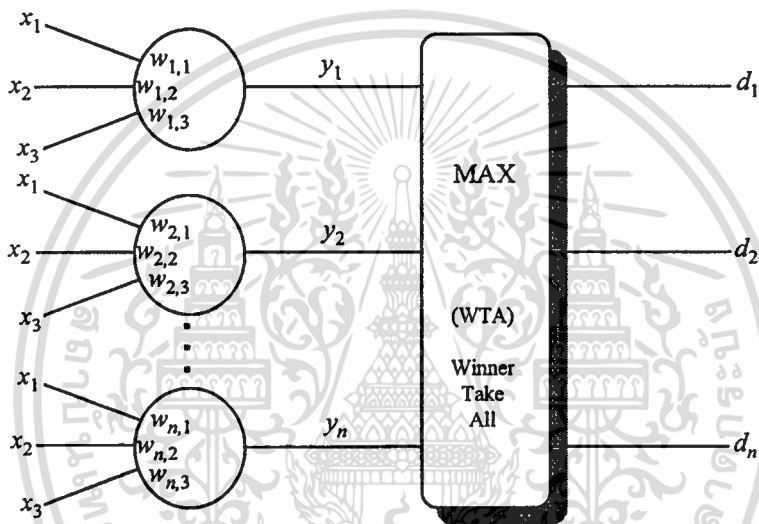
1. โครงข่ายเซลล์ประสาทเทียมแบบเพอร์เซพตรอน (Perceptron)

เพอร์เซพตรอนเป็นโครงข่ายเซลล์ประสาทเทียมที่มีจำนวนชั้นเพียงชั้นเดียว (Single Layer) และถูกออกแบบมาเพื่อใช้งานประเภทจดจำรูปแบบ (Pattern Recognition) โดยถูกจัดอยู่ในกลุ่มของโครงข่ายเซลล์ประสาทเทียมประเภทการเรียนรู้โดยมีครู ซึ่งมีลักษณะการทำงานดังที่จะกล่าวต่อไปนี้

ถ้ามีกลุ่มข้อมูลที่สามารถถูกแทนด้วย k ประเภทข้อมูล (Category) แล้วละก็ โครงข่ายเซลล์ประสาทเทียมแบบเพอร์เซพตรอน ก็จะต้องมี k เซลล์ประสาทภายในหนึ่งชั้น (Layer) เช่นกัน นั่นคือ จะแทนหนึ่งเซลล์ประสาทต่อหนึ่งประเภทข้อมูล และข้อมูลที่ป้อนเข้ามาจะถูกป้อนต่อไปให้เซลล์ประสาททั้งหมด k เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เซลล์ และมี k เอาท์พุทที่เป็นผลลัพธ์ซึ่งได้จาก k เซลล์ประสาท จากนั้น ผลลัพธ์ที่ได้นี้ก็ถูกป้อนไปยังขั้นตอนการตัดสินใจ (Decision) และเลือกเอาค่าที่มากที่สุดเป็นค่าที่ต้องการ โดยที่ขั้นตอนการตัดสินใจจะทำหน้าที่ตัดสินใจในการจัดประเภทข้อมูลที่รับเข้ามาจากอินพุทว่า สอดคล้องกับประเภทข้อมูลใด ดังแสดงไว้ในภาพที่ 23

ภาพที่ 23



การตัดสินใจ :

เวกเตอร์ : $X = (x_1, x_2, x_3, \dots, x_n)$

คำนวณเวกเตอร์ : $Y = (y_1, y_2, y_3, \dots, y_n)$



$$y_m = \text{MAX} \{ y_k \}$$

เมื่อ m เป็นดัชนีของค่า y ที่มากที่สุด

เอาท์พุท : $d_m = 1; d_i = 0$ เมื่อ $i \neq m$

อัลกอริทึมการฝึกเรียนรู้ของโครงข่ายเซลล์ประสาทเทียมแบบเพอร์เซพตรอน

(1) สร้างกลุ่มตัวอย่างข้อมูลของการฝึกเรียนรู้ขึ้นมา โดยสร้างมาจาก เวกเตอร์ X และ D เมื่อเวกเตอร์ X เป็นกลุ่มของข้อมูลอินพุต (x_i) ถูกสร้างไว้เป็นรูปแบบ (Pattern) เพื่อใช้ในการตัดสินใจ และเวกเตอร์ D เป็นกลุ่มของข้อมูลที่ส่งเข้ามา แต่ละคู่ของข้อมูลใน (X, D) กลุ่มจะถูกป้อนให้กับเพอร์เซพตรอน โดยกลุ่มข้อมูลที่ถูกป้อนไปให้กับเพอร์เซพตรอนในแต่ละรอบ จะเรียกว่า "Epoch" ซึ่งกระบวนการดังกล่าวจะกระทำซ้ำทุก ๆ Epoch

(2) กำหนดสภาวะเริ่มต้นให้กับเพอร์เซพตรอน โดยวิธีการสุ่มให้ค่า W_{ij} ทุกตัว ซึ่งมักจะกำหนดให้ค่าอัตราการเรียนรู้ L (Training Rate) มีค่าน้อย ๆ ประมาณ "1"

(3) กระบวนการประมวลผลหลัก (Main Loop) : ให้ดำเนินการตามขั้นตอนของ Epoch สำหรับกลุ่มของข้อมูลที่ต้องการให้ฝึกเรียนรู้ เป็นคู่ $(X, D_{desired})$ เมื่อ $D_{desired}$ คือ ค่า D ที่ต้องการ

ขั้นตอนที่ 1 ป้อนข้อมูล X แล้วให้คำนวณ $D_{computed}$

ขั้นตอนที่ 2 เปรียบเทียบค่า $D_{desired}$ กับ $D_{computed}$ ที่ได้จากขั้นตอนที่ 1

ขั้นตอนที่ 3 ถ้าค่า $D_{desired} = D_{computed}$ ให้กลับไปทำขั้นตอนที่ 1 ใหม่ เมื่อ Epoch ยังทำงานไม่เสร็จสิ้น หรือไปที่ขั้นตอนที่ 4 ถ้า Epoch ถูกทำงานจนเสร็จสิ้นหมดแล้ว

ถ้าค่า $D_{desired} \neq D_{computed}$ ให้ปรับค่าน้ำหนัก (w) เสียใหม่ดังสมการข้างล่างนี้

ถ้า $D_{desired}$ คือ $d_k=1$ และ $D_{computed}$ คือ $d_j=1$ (ซึ่งเป็นการตัดสินใจที่ผิด : เมื่อ k คือ ประเภทข้อมูลที่ต้องการ และ j เป็นการตัดสินใจผิด)

$$W_{ki}(new) = W_{ki}(old) + L_{xi} \quad \text{สำหรับ } i \text{ ทุกค่า}$$

$$W_{ji}(new) = W_{ji}(old) - L_{xi} \quad \text{สำหรับ } i \text{ ทุกค่า}$$

ย้อนกลับไปดำเนินการใน ขั้นตอนที่ 1 ใหม่

ขั้นตอนที่ 4 ถ้าไม่สามารถหาค่าที่ต้องการได้ในขั้นตอนของ Epoch ให้หยุด แต่ถ้าเป็นกรณีอื่นให้ย้อนกลับไปยัง กระบวนการประมวลผลหลัก (3) กระทำ Epoch ซ้ำ

อัลกอริทึมข้างบนนี้จะทำการประมวลผลจนกระทั่งสำเร็จ แต่สิ่งหนึ่งที่พึงควรระวัง คือ การเรียนรู้ของเพอร์เซพตรอนที่ไม่ลู่เข้า ดังนั้นจึงควรถือว่าการใช้เพอร์เซพตรอนจาก ทฤษฎีการลู่เข้าของเพอร์เซพตรอน (Perceptron Convergence Theorem) ซึ่งจะได้อธิบายต่อไปนี้

ทฤษฎีการลู่อิงของเพอร์เซพตรอน

ถ้ากลุ่มของข้อมูลที่ต้องการนำมาฝึกให้เรียนรู้ คือ เวกเตอร์ X ซึ่งเป็นข้อมูลประเภท จำแนกแบบเชิงเส้น (Linear Separable) อัลกอริทึมก็จะลู่อิงเข้าตามจำนวนของ Epoch ซึ่งมีจำนวนจำกัด

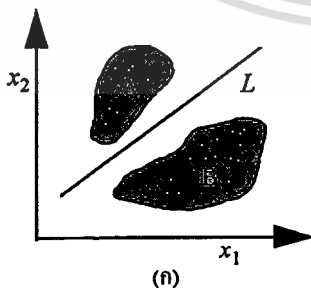
จากทฤษฎีบทที่กล่าวมานั้น ความหมายของคำว่า "จำแนกแบบเชิงเส้น" นั้น หมายถึง ถ้าทำการพล็อตตำแหน่งของสมาชิกในเวกเตอร์ X แต่ละตัว โดยใช้ค่า X เป็นค่าคอดอดิเนตของแกน แล้วประเภทของข้อมูล k สามารถจำแนกจากกันได้ด้วยเส้นตรง $k-1$ เส้น หรืออาจเป็นระนาบก็ได้ คงจะเห็นภาพได้ชัดเจนยิ่งขึ้นจากตัวอย่างต่อไปนี้ สมมุติว่ามีข้อมูลอินพุต 2 ค่า คือ x_1 และ x_2 ซึ่งเป็นข้อมูลที่สุ่มเข้ามา และเป็นข้อมูลเพียงสองประเภทเท่านั้นของการตัดสินใจ คือข้อมูลประเภท A และข้อมูลประเภท B เมื่อทำการพล็อตค่าของ x_1 และ x_2 แล้วจะได้ดังภาพที่ 24(ก) เมื่อทุกจุดที่อยู่ในข้อมูลประเภท A จะมีเวกเตอร์ D ดังนี้

$$D = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

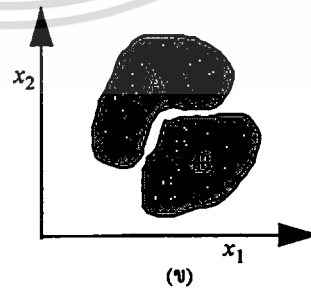
และขณะเดียวกัน ข้อมูลประเภท B ก็จะมีเวกเตอร์ D ดังนี้

$$D = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

ภาพที่ 24



(ก)
ข้อมูลที่สามารถจำแนก
แบบเชิงเส้น



(ข)
ข้อมูลที่ไม่สามารถจำแนกกันได้

ภาพแสดงการจำแนกประเภทข้อมูลเพื่ออธิบายการลู่อิงของเพอร์เซพตรอน

ข้อมูลที่อยู่กันเป็นกลุ่มดังแสดงในภาพที่ 24(ก) ซึ่งมีเส้นตรง L เป็นเส้นแบ่งกลุ่มของข้อมูลออกเป็นสองส่วน คือ ข้อมูลประเภท A และข้อมูลประเภท B ซึ่งกลุ่มข้อมูลที่วางตัวในลักษณะอย่างนี้ เราเรียกว่า “การจำแนกแบบเชิงเส้น” ดังนั้น อัลกอริทึมของเพอร์เซพตรอนจะลู่เข้าและให้คำตอบที่ถูกต้องเสมอ แต่ในโลกแห่งความเป็นจริงนั้น กลุ่มข้อมูลที่ใช้ในการฝึกเรียนรู้ จะมีการวางตัวดังแสดงในภาพที่ 24(ข) ซึ่งไม่มีเส้นตรงเส้นไหนเลยที่สามารถแบ่งข้อมูลประเภท A และข้อมูลประเภท B ออกจากกันได้ ฉะนั้น ข้อมูลที่ใช้ในการฝึกเรียนรู้ในลักษณะแบบนี้ จะให้ผลที่ไม่ยอมลู่เข้า เมื่อนำไปใช้กับอัลกอริทึมของเพอร์เซพตรอน โดยที่อัลกอริทึมจะทำงานวนซ้ำไปเรื่อย ๆ ถ้าปล่อยให้อัลกอริทึมทำงานเป็นเวลานานเกินไป ความสามารถในการจำแนกประเภทข้อมูลอาจจะให้คำตอบที่ไม่พึงประสงค์ได้ ดังนั้น ความสามารถในการจำแนกประเภทข้อมูลจึงเป็นสิ่งที่ต้องคำนึงถึงเป็นอย่างมาก ซึ่งโดยทั่วไปแล้วสำหรับอัลกอริทึมการแบ่งประเภทข้อมูลแบบอื่น ๆ จะยอมลดขีดความสามารถของการแบ่งประเภทข้อมูลลง โดยจะยอมให้ทำงานอยู่ประมาณ 90% ของเวลาที่ยังทำงานได้อย่างถูกต้อง เมื่อมันมีแนวโน้มที่จะไม่ลู่เข้า

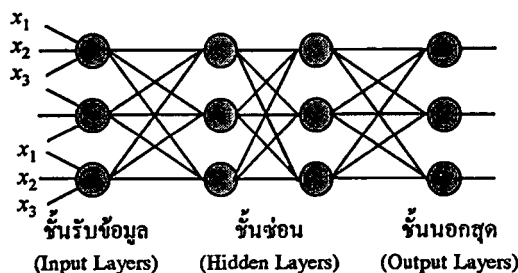
2. โครงข่ายเซลล์ประสาทเทียมแบบแพร่กลับ (Back Propagation)

โครงข่ายเซลล์ประสาทเทียมแบบแพร่กลับเป็นโครงข่ายแบบหลายชั้น (Multi Layer) โดยพัฒนาขึ้นมาเพื่อแก้ไขปัญหาในสิ่งที่โครงข่ายประสาทเทียมแบบชั้นเดียวไม่สามารถแก้ไขได้ เช่น ข้อจำกัดของการจำแนกแบบเชิงเส้น (Linear Separation) ของเพอร์เซพตรอน ซึ่งเป็นโครงข่ายเซลล์ประสาทแบบชั้นเดียว

โครงสร้างของเซลล์ประสาทแบบแพร่กลับ จะถูกวางเรียงกันเป็นชั้น ๆ ซึ่งมีด้วยกันทั้งหมด 3 ชั้น คือ

- (1) ชั้นรับข้อมูล (Input Layer)
- (2) ชั้นซ่อน (Hidden Layer)
- (3) ชั้นนอกสุด (Output Layer)

ภาพที่ 25



ในส่วนของชั้นรับข้อมูลจะประกอบไปด้วยเซลล์ประสาท ซึ่งทำหน้าที่รับข้อมูลจากภายนอกเข้ามา ถัดจากชั้นรับข้อมูลเข้ามา จะเป็นชั้นซ่อน ซึ่งประกอบไปด้วยเซลล์ประสาทที่ทำหน้าที่รับข้อมูลจากเซลล์ประสาทเซลล์อื่น ๆ เข้ามาทำการประมวลผล และจะให้ผลลัพธ์เพียงค่าเดียวและจะป้อนให้กับเซลล์ประสาทอื่น ๆ ต่อไป ส่วนชั้นนอกสุดนั้น จะประกอบไปด้วยเซลล์ประสาทที่จะให้ผลลัพธ์สุดท้ายออกมา

ถ้าในการทำงานของโครงข่าย มีการส่งผ่านของข้อมูลจากชั้นหนึ่งไปยังอีกชั้นหนึ่งถัดไป โดยไม่มีการป้อนข้อมูลให้เซลล์อื่นอีก เราจะเรียกการทำงานในลักษณะอย่างนี้ว่า "Feedforward" นั้นหมายความว่า ผลลัพธ์ที่ได้ของแต่ละเซลล์ประสาทในชั้นใด ๆ จะถูกป้อนสู่ชั้นถัดไปดังแสดงไว้ในภาพที่ 25 ซึ่งมีชั้นซ่อนอยู่สองชั้น

ในการรับข้อมูลของชั้นรับข้อมูล จะรับเพียง x_1 , x_2 และ x_3 เท่านั้น และขณะเดียวกัน ผลลัพธ์ที่ได้จากชั้นรับข้อมูล จะกลายเป็นข้อมูลที่จะป้อนให้กับชั้นซ่อนชั้นที่ 1 ผลลัพธ์ที่ได้จากชั้นซ่อนชั้นที่ 1 จะป้อนให้ชั้นซ่อนชั้นที่ 2 ต่อไป เซลล์ประสาทชั้นนอกสุดจะรับข้อมูลจากชั้นซ่อนชั้นที่ 2 เซลล์ประสาทที่ทำงานแบบ Feedforward ดังที่ได้กล่าวมาแล้วนั้น จำเป็นที่จะต้องมีฟังก์ชันถ่ายโอนที่สามารถหาค่าอนุพันธ์ได้ทุกค่า ดังนั้น ในการเลือกใช้งานฟังก์ชันจึงควรเลือกฟังก์ชันแบบ Sigmoid หรือ Hiperbolic Tangent และไม่ควรรใช้ฟังก์ชันแบบ Step หรือ ฟังก์ชันอิมพัลส์แบบเชิงเส้น ส่วนฟังก์ชัน Sigmoid เป็นฟังก์ชัน ที่นิยมนำมาใช้งานกันเป็นอย่างมาก ซึ่งสามารถหาค่าอนุพันธ์ได้ทุกค่า ดังแสดงในสมการข้างล่างนี้

$$f(a) = \frac{1}{1 + \exp(-a)}$$

$$\frac{df(a)}{da} = f(a)[1 - f(a)]$$

ส่วนในขั้นตอนการนำโครงข่ายไปใช้งาน ภายหลังจากที่คำนวณน้ำหนักของแต่ละเซลล์ประสาทถูกคำนวณจนเสร็จสิ้นแล้วโดยการฝึกเรียนรู้ ข้อมูลจากภายนอกจะถูกป้อนให้กับอินพุต x เซลล์ประสาทในชั้นรับข้อมูลจะตอบสนองต่อข้อมูลที่ป้อนเข้ามา และจะได้เอาท์พุทออกมาที่ชั้นรับข้อมูล ซึ่งเอาท์พุทที่ได้นี้จะกลายเป็นอินพุตของชั้นซ่อนต่อไป โดยลักษณะของการส่งผ่านข้อมูลจะส่งผ่านจากชั้นหนึ่งไปยังอีกชั้นหนึ่ง จนกระทั่งถึงชั้นนอกสุด ซึ่งจะเป็นเอาท์พุทที่ผ่านการตอบสนองของทุกเซลล์ประสาทแล้ว

อัลกอริทึมการฝึกเรียนรู้ของโครงข่ายเซลล์ประสาทเทียมแบบแพร่กลับ

อัลกอริทึมที่ใช้ฝึกเรียนรู้ของโครงข่ายเซลล์ประสาทเทียมแบบแพร่กลับ จะมีการทำงานแบบวนซ้ำ (Iteration) โดยใช้ Gradient Descent เป็นตัวควบคุมทำให้ผลรวมของความผิดพลาดกำลังสอง (sum of square error) มีค่าน้อยที่สุด เวกเตอร์ $\{C_i, B_i\}$ เป็นเวกเตอร์ที่ใช้ในการฝึกเรียนรู้ให้กับโครงข่ายเซลล์ประสาทที่มีการทำงานแบบ Feedforward โดยที่ เวกเตอร์ C จะใช้เป็นอินพุต ถ้ากำหนดให้

$$C = (c_1, c_2, c_3, \dots, c_n)$$

แล้วจะได้ว่า

$$x_1 = c_1, x_2 = c_2, x_3 = c_3, \dots, x_n = c_n$$

และมีเวกเตอร์ B เป็นเอาต์พุตที่ต้องการ ส่วนเอาต์พุตที่แท้จริง คือ

$$Y = (y_1, y_2, y_3, \dots, y_k)$$

ซึ่งจะถูกนำมาเปรียบเทียบกับเอาต์พุตที่ต้องการ

$$B = (b_1, b_2, b_3, \dots, b_k)$$

ส่วนเวกเตอร์ความผิดพลาด D จะถูกคำนวณโดย

$$D = (b_1 - y_1, b_2 - y_2, b_3 - y_3, \dots, b_k - y_k)$$

ค่าความผิดพลาดนี้จะถูกแพร่กลับไปยังโครงข่าย โดยเริ่มจากชั้นนอกสุดของโครงข่าย และต่อไปยังชั้นสุดท้ายของชั้นซ่อน และแพร่ไปยังทุก ๆ ชั้น จนถึงชั้นรับข้อมูล นี่เป็นการคำนวณเพื่อหาความผิดพลาดของแต่ละเซลล์ประสาทว่ามีค่ามากน้อยเพียงใด ค่าความผิดพลาดทั้งหมด d_j จะเป็นคุณสมบัติของเซลล์ประสาท j และท้ายสุด น้ำหนักของแต่ละเซลล์ประสาทจะถูกปรับเพื่อลดค่าความผิดพลาด ซึ่งลำดับขั้นตอนการทำงานสามารถอธิบายได้ดังนี้

- (1) กำหนดสภาวะเริ่มต้น : กำหนดค่าน้ำหนักของเซลล์ประสาทด้วยการสุ่ม และกำหนดอัตราการเรียนรู้ R

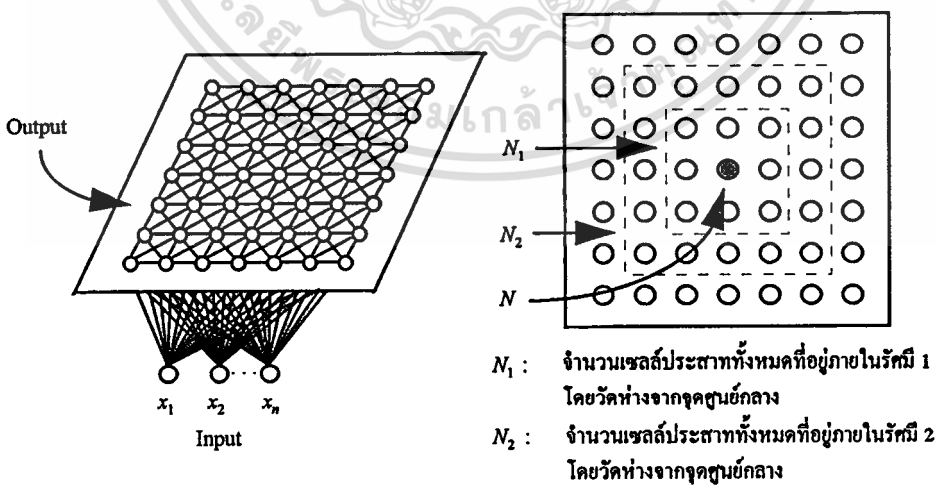
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- (2) กระบวนการประมวลผลหลัก (*Main Loop*) : ในแต่ละคู่ของเวกเตอร์ C และ B ให้ดำเนินการดังนี้
- (2.1) ให้อินพุตด้วยเวกเตอร์ C โดยป้อนแบบ Feedforward เพื่อคำนวณหาค่า Y
 - (2.2) คำนวณหาค่าความผิดพลาด $B-Y$
 - (2.3) ทำการแพร่กลับค่าความผิดพลาด เพื่อให้ได้ค่า d ในแต่ละเซลล์ประสาท
 - (2.4) ทำการปรับค่าน้ำหนักของแต่ละเซลล์ประสาท
- (3) ตรวจสอบ : หลังจากที่ผ่านมาจากกระบวนการประมวลผลหลักแล้ว ให้ตรวจสอบโครงข่ายเซลล์ประสาทว่าให้คำตอบที่ดีพอที่จะหยุดการทำงานได้หรือยัง ถ้ายังไม่ดีพอ ให้ทำการลดค่าอัตราการเรียนรู้ R แล้วย้อนกลับไปทำงานซ้ำในขั้นตอนที่ 2 ใหม่

3. โครงข่ายเซลล์ประสาทเทียมแบบ Self-Organizing Map

โครงข่ายเซลล์ประสาทเทียมแบบ Self-Organizing Map ซึ่งถูกจัดอยู่ในประเภท การเรียนรู้โดยไม่มีครู โดยโครงข่ายจะทำการปรับค่าน้ำหนักของเซลล์ประสาท เพื่อที่จะจัดกลุ่มของข้อมูลที่ได้รับเข้ามาทางอินพุตของโครงข่าย

ภาพที่ 26



ภาพแสดงโครงสร้างของ Self-Organizing Map

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงสร้างของโครงข่ายเซลล์ประสาทเทียมแบบนี้ จะประกอบด้วยชั้นของเซลล์ประสาทเพียงชั้นเดียว แต่จะมีการเรียงของเซลล์ประสาทแต่ละเซลล์ในลักษณะตาข่ายสี่เหลี่ยม ดังแสดงในภาพที่ 26 ซึ่งมีขนาด 7 แถว 7 หลัก (เหมือนเมตริกซ์) โดยเซลล์ประสาทแต่ละเซลล์จะถูกกำหนดหรือบอกตำแหน่งไว้ด้วยตัวเลขแถว (row number) และตัวเลขหลัก (column number) และในการกำหนดระยะทางความห่างระหว่างสองเซลล์ประสาทใด ๆ จะใช้ค่าตัวเลขสูงสุดของผลต่างในแถวหรือหลักในการแบ่งแยกสองเซลล์ประสาท ตัวอย่างเช่น ถ้าเซลล์ประสาทเซลล์ที่หนึ่ง อยู่ในตำแหน่งแถวที่ 3 และหลักที่ 5 และเซลล์ประสาทเซลล์ที่ 2 อยู่ในตำแหน่งแถวที่ 4 และหลักที่ 7 ผลต่างของแถวของเซลล์ประสาททั้งสอง คือ $4-3=1$ และผลต่างของหลักของเซลล์ประสาททั้งสองคือ $7-5=2$ เพราะฉะนั้น ค่าตัวเลขสูงสุดของผลต่างในแถวหรือหลัก คือ 2 เราจะนำหลักการดังกล่าวมาเป็นเครื่องกำหนดเซลล์ประสาทข้างเคียงและรัศมีของเซลล์ประสาทข้างเคียง เช่น รัศมีของเซลล์ประสาทข้างเคียง K สำหรับเซลล์ประสาท N คือ กลุ่มของเซลล์ประสาททั้งหมดที่อยู่ล้อมรอบเซลล์ประสาท N โดยมีระยะห่างจาก N เป็นระยะทาง K หรือน้อยกว่า ดังแสดงในภาพที่ 26

ส่วนน้ำหนักของเซลล์ประสาทจะถูกทำการ Normalized ดังนั้น ผลรวมของน้ำหนักยกกำลังสองจะมีค่าเท่ากับ "1" เอาท์พุทของเซลล์ประสาทจะได้จากการใช้เทคนิคเลือกผู้ชนะ (Winner Take All) แต่ละเซลล์จะรับข้อมูลอินพุท x_i มาจากแหล่งเดียวกัน โดยข้อมูลอินพุทจะถูกป้อนให้เรียงตามลำดับ และน้ำหนักของเซลล์ประสาทจะถูกปรับ โดยมีอัลกอริทึมสำหรับการเรียนรู้ดังนี้

อัลกอริทึมการเรียนรู้สำหรับโครงข่ายเซลล์ประสาทเทียมแบบ Self-Organizing Map

(1) กำหนดสถานะเริ่มต้น

กำหนดค่า R (อัตราการเรียนรู้ < 1)

กำหนดค่า K ซึ่งเป็นรัศมีของเซลล์ประสาท

กำหนดค่าน้ำหนักของเซลล์ประสาททั้งหมดโดยการสุ่มด้วยตัวเลขน้อย ๆ พร้อมกับทำการ Normalize ค่าน้ำหนัก

(2) กระบวนการประมวลผลหลัก (Main Loop)

(2.1) ป้อนข้อมูล X

(2.2) คำนวณค่า $a_i = \sum w_{ij} x_j$ ของทุกเซลล์ประสาท

(2.3) เลือกเซลล์ประสาทที่ให้ค่ามากที่สุด (N)

(2.4) ทำการปรับค่าของทุกเซลล์ประสาท i ภายในรัศมี K ของเซลล์ประสาทข้างเคียง N

(2.5) ทำการ Normalize น้ำหนัก

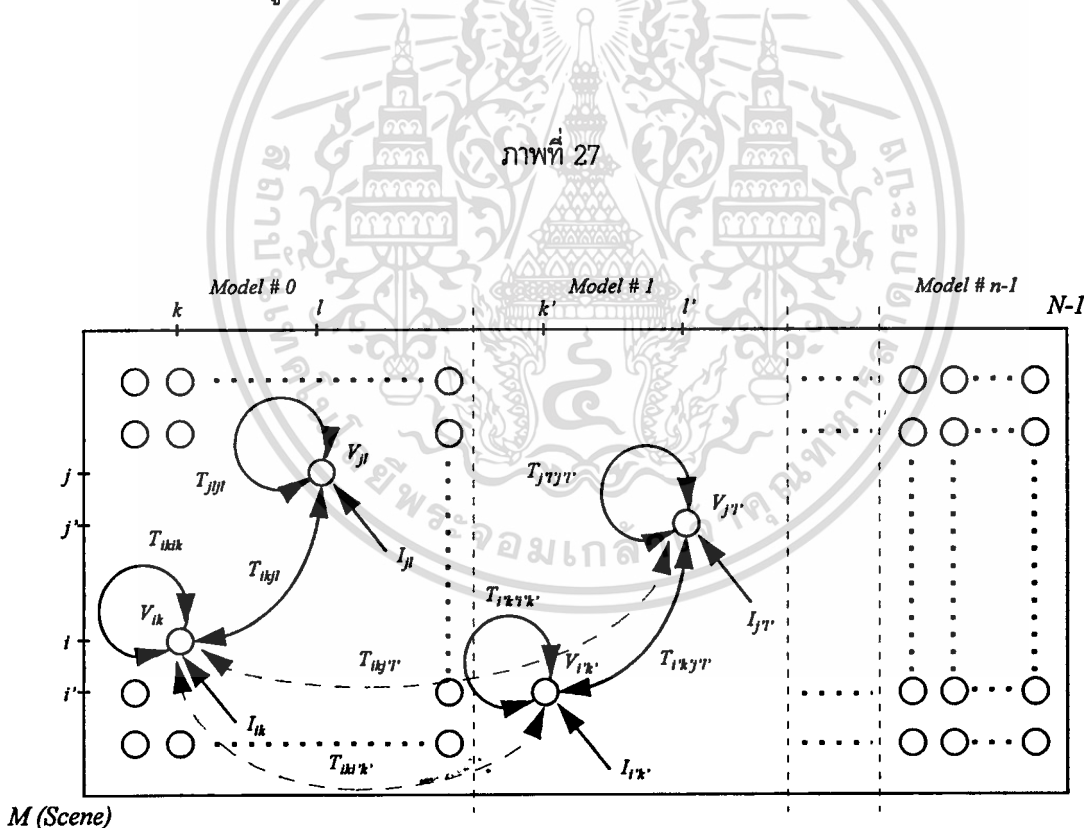
(3) เมื่อสิ้นสุดการทำงานจากกระบวนการประมวลผลหลัก ให้ลดค่า K และค่า R

ถ้า $K \neq 0$ ให้ย้อนกลับไปดำเนินการในขั้นตอนกระบวนการประมวลผลหลัก (2)

4. โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ (Hopfield Neural Network)

โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ เป็นโครงข่ายที่ถูกจัดอยู่ในประเภท การเรียนรู้โดยการ บังคับ (Reinforcement Learning) โดยถูกพัฒนาขึ้นมาเพื่อแก้ไข ปัญหา ปัญหาการเลือกคำตอบที่ดีที่สุด (Optimization Problem) ซึ่งมีหลักการทำงาน คือ การสร้างวิธีการตัดสินใจร่วมกัน โดยยืนอยู่บนพื้นฐานของข้อมูลที่ได้รับเข้ามาในแต่ละเซลล์ประสาทพร้อม ๆ กัน และมีการส่งผ่านข้อมูลระหว่างเซลล์ประสาท พร้อมกับบังคับให้โครงข่ายเซลล์ประสาท ลู่เข้าสู่สภาวะคงที่หรือเสถียรภาพ (Stable State) เพื่อที่จะได้ตัดสินใจกับข้อมูลที่ได้รับเข้ามานั้น

โครงสร้างของเซลล์ประสาทเทียมแบบโฮปฟิลด์ ดังแสดงในภาพที่ 27 จะเห็นว่าโครงข่ายเซลล์ประสาทเทียม มีการวางตัวของแต่ละเซลล์ประสาทอยู่ในรูปเมตริกซ์ (Matrix) ซึ่งมีขนาดเท่ากับ $M \times N$ แต่ละเซลล์ประสาทสามารถมีความสัมพันธ์กับเซลล์ประสาทอื่น ๆ ได้ ภายใต้ความเชื่อมโยงของ T_{ijkl} ขณะเดียวกันแต่ละเซลล์ประสาทก็จะรับเอาข้อมูลจากภายนอกเข้ามา (I_{ik}) ได้ด้วย



ภาพแสดงโครงสร้างของโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

ส่วนสมการต้นแบบของโฮปฟิลด์ E ถูกนิยามไว้ดังนี้

$$E = -\frac{1}{2} \sum_{x=1}^n \sum_{y=1}^n T_{xy} V_x V_y - \sum_{x=1}^n I_x V_x \quad (3-1)^1$$

เมื่อกำหนดให้ $x \Rightarrow i, k$ และ $y \Rightarrow j, l$ เพื่อให้สอดคล้องกับลักษณะของงาน จึงใช้ตัวย่อ 2 ตัว ดังแสดงไว้ในภาพที่ 27 ทำให้สามารถเขียนสมการที่ (3-1) ใหม่ได้เป็นสมการที่ (3-2)

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

เขียนสมการที่ (3-2) เลียนใหม่ ให้สมบูรณ์ได้เป็น

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

เมื่อ V_{ik} และ V_{jl} คือ สถานะของเซลล์ประสาท ณ ตำแหน่ง ik และ jl ตามลำดับ ซึ่งจะมีค่าเป็น "0" หรือ "1" เท่านั้น
 T_{ijkl} คือ ความสัมพันธ์ที่เชื่อมโยงถึงกันของเซลล์ประสาท ik กับ jl
 I_{ik} คือ ข้อมูลภายนอกที่ป้อนให้กับเซลล์ประสาท

คำตอบของสมการที่ (3-3) จะปรากฏ ก็ต่อเมื่อ สมการได้รับตัวเองเข้าสู่สภาวะพลังงานต่ำสุดแล้ว คำตอบที่ต้องการก็จะปรากฏอยู่ในแต่ละเซลล์ประสาทในโครงข่ายนั่นเอง แต่การที่สมการจะปรับตัวเองเพื่อเข้าสู่สภาวะพลังงานต่ำสุดได้นั้น เซลล์ประสาท V_{ik} จะต้องมีการเปลี่ยนแปลงสถานะตามสมการต่อไปนี้

$$V_{ik} \rightarrow 0 \quad \text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) < 0 \quad (3-4)$$

$$V_{ik} \rightarrow 1 \quad \text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) > 0 \quad (3-5)$$

$$V_{ik} \rightarrow \text{no change} \quad \text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) = 0 \quad (3-6)$$

¹ S.Y. Kung. Digital Neural Network (U.S.A. : Prentice Hall International Editions, 1993), PP 39,52
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งการเปลี่ยนแปลงสถานะของ V_{ik} ใด ๆ นั้น เป็นการปรับค่าพลังงานเฉพาะเซลล์ประสาทรุ่นเก่า นั้น นั้น สมการย่อยของการเปลี่ยนแปลงสถานะของ V_{ik} จึงเป็นการกำหนดให้ i กับ k เป็นค่าคงที่ คงเหลือแต่ j กับ l ที่เป็นตัวแปร จึงได้สมการย่อยของการลดค่าพลังงาน ดังสมการที่ (3-4), (3-5) และ (3-6) ตามลำดับ

อัลกอริทึมการทำงานของโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

- (1) กำหนดสถานะเริ่มต้นให้กับโครงข่าย โดยการกำหนดเซลล์ประสาท V_{ik} ให้มีค่าเป็น '0' หรือ '1' ตามเงื่อนไขสมการต่อไปนี้

$$V_{ik} = \begin{cases} 1 & \text{if } f(i,k) > \mu \\ 0 & \text{Otherwise} \end{cases}$$

เมื่อ $f(i,k)$ คือ ฟังก์ชันกำหนดค่าความสามารถเข้ากันได้ (Compatibility) ของเซลล์ประสาท ที่ i และ k

และ μ คือ ค่าคงที่สำหรับตั้งเป็นค่า Threshold ของฟังก์ชัน $f(i,k)$

- (2) กระบวนการประมวลผลหลัก :

(2.1) ทำการสุ่มเอาเซลล์ประสาท (i,k) ที่มีค่าเป็น "1" แล้วทำการคำนวณหาค่า U_{ik}

$$U_{ik} = \sum_{j=1}^M \sum_{l=1}^N (T_{ijl} - \delta_{ij} - \delta_{kl}) V_{jl} + I_{ik}$$

เมื่อ δ_{ij} และ δ_{kl} คือ Kronecker delta function

$$\text{โดยที่ } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{Otherwise} \end{cases} \quad \text{และ} \quad \delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{Otherwise} \end{cases}$$

(2.2) ทำการตัดสินใจเปลี่ยนแปลงสถานะของ V_{ik} ตามสมการต่อไปนี้

$$\begin{array}{ll} V_{ik} \rightarrow 0 & \text{ถ้า } U_{ik} < 0 \\ V_{ik} \rightarrow 1 & \text{ถ้า } U_{ik} > 0 \\ V_{ik} \rightarrow \text{no change} & \text{ถ้า } U_{ik} = 0 \end{array}$$

- (3) ตรวจสอบ : ตรวจสอบการเปลี่ยนแปลงสถานะของ V_{ik} ว่ายังมีการเปลี่ยนแปลงอยู่หรือไม่ ถ้ายังมีการเปลี่ยนแปลงให้กลับไปดำเนินการในกระบวนการประมวลผลหลัก (2) แต่ถ้าไม่มีการเปลี่ยนแปลงสถานะของ V_{ik} แสดงว่า U_{ik} ได้เข้าสู่สภาวะพลังงานต่ำสุดแล้ว ให้ไปดำเนินการต่อในขั้นตอนที่ 4
- (4) คำตอบสุดท้ายจะปรากฏที่ตำแหน่ง V_{ik} มีค่าเป็น "1" ในโครงข่าย



บทที่ 4

การจดจำภาพวัตถุ

บทนำ

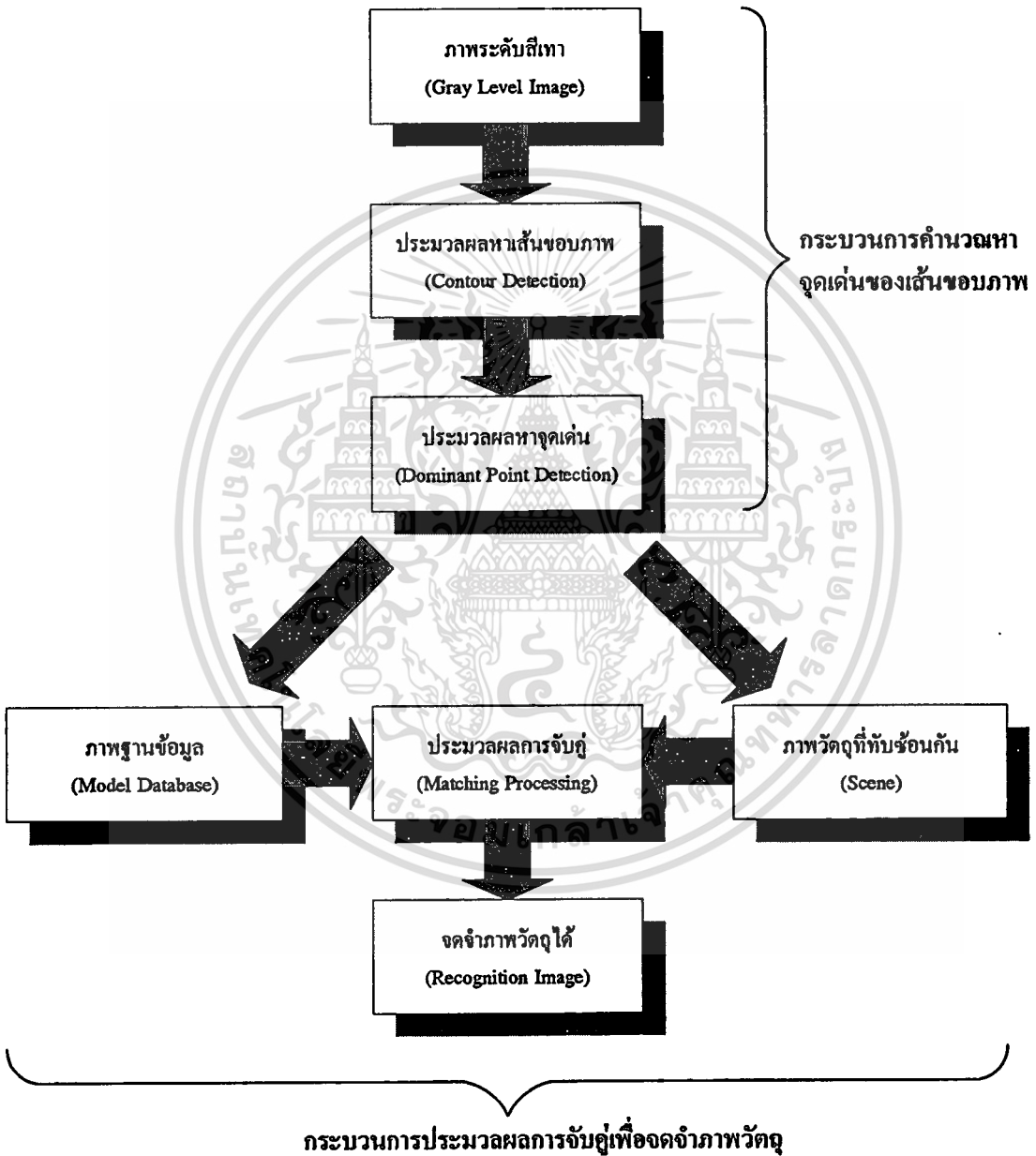
ปัจจุบัน ความต้องการใช้หุ่นยนต์และเซนกลในอุตสาหกรรมประกอบชิ้นส่วน และการตรวจสอบเริ่มมีมากขึ้น ซึ่งการทำงานของหุ่นยนต์เหล่านี้มักจะประกอบไปด้วยกระบวนการเลือกวัตถุ การหาตำแหน่งและทิศทางการวางของวัตถุ ขนาดของวัตถุ ขั้นตอนเหล่านี้ล้วนแต่เป็นปัจจัยในการพัฒนาระบบการมองเห็นของหุ่นยนต์มาโดยตลอด เพื่อให้สามารถหยิบจับชิ้นส่วนและวางชิ้นส่วนวัตถุได้อย่างถูกต้องและแม่นยำ แต่การพัฒนาไม่ได้หยุดอยู่เพียงที่หุ่นยนต์สามารถจับวางได้อย่างถูกต้องเท่านั้น แต่ต้องการให้ระบบการมองเห็นของหุ่นยนต์สามารถวิเคราะห์ภาพข้อมูลในสถานะการณ์ที่ข้อมูลของภาพวัตถุนั้นมีความไม่ชัดเจน มีการขาดหายไปบางส่วนหรือข้อมูลอื่นเนื่องมาจากการทับซ้อนกันหรือสัมผัสกันของวัตถุที่มากกว่าหนึ่งชิ้น ดังนั้น ในแนวทางการพัฒนาเพื่อแก้ไขปัญหาดังกล่าวจึงได้มีการพัฒนาการจดจำภาพวัตถุให้มีประสิทธิภาพ และมีความยืดหยุ่นต่อการใช้งานมากขึ้น

การประมวลผลการจดจำภาพวัตถุ คือ ความพยายามที่จะค้นหาส่วนของข้อมูลที่สูญหายไปบางส่วนหรือภาพวัตถุอื่นเนื่องมาจากการทับซ้อนกันบางส่วนหรือวัตถุ หรือการสัมผัสกันของวัตถุที่มีมากกว่าหนึ่งชิ้น เพื่อให้ระบบการมองเห็นของหุ่นยนต์ หรือเซนกลสามารถวิเคราะห์และจดจำวัตถุได้ รู้ตำแหน่งการวางและขนาดของวัตถุได้อย่างถูกต้อง เพื่อเป็นข้อมูลที่จะใช้ควบคุมการทำงานของเซนกลให้ดำเนินการไปได้อย่างถูกต้องและแม่นยำ

การจดจำภาพวัตถุมีเงื่อนไขที่สำคัญที่จะต้องคำนึงถึง คือ การจดจำภาพจะต้องสามารถตอบสนองและมีเสถียรภาพต่อการเลื่อนหรือย้ายตำแหน่งของภาพวัตถุ การหมุนของภาพวัตถุ ตลอดจนการย่อหรือขยายของภาพวัตถุ ฉะนั้น คณิตศาสตร์ที่นำมาทำการประมวลผลจะต้องสามารถตอบสนองต่อเงื่อนไขเหล่านี้ได้อย่างสมบูรณ์

การประมวลผลการจดจำภาพวัตถุ

ภาพที่ 28



ภาพแสดงขั้นตอนการประมวลผลการจดจำและบอกตำแหน่งของวัตถุ

ที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการงานของการจดจำภาพวัตถุได้แสดงไว้ในภาพที่ 28 ซึ่งมีกระบวนการที่สำคัญอยู่สองกระบวนการ คือ กระบวนการประมวลผลหาจุดเด่นของเส้นขอบภาพ และ การประมวลผลการจับคู่ของจุดเด่นเพื่อจดจำภาพวัตถุ

1. การคำนวณหาจุดเด่นของเส้นขอบภาพ

ในการคำนวณหาจุดเด่นของเส้นขอบภาพ มีลำดับขั้นตอนการทำงานโดยเริ่มจาก การรับภาพระดับสีเทา (Gray Level Image) เข้าสู่ระบบการประมวลผลโดยผ่านทางกล้อง CCD แล้วทำการแปลงให้เป็นเส้นขอบภาพ* หลังจากนั้น จึงนำเอาเส้นขอบภาพที่ได้มาทำการประมวลผลหาจุดเด่นตามวิธีการที่ได้กล่าวโดยละเอียดในบทที่ 2 ซึ่งในการคำนวณหาจุดเด่นของเส้นขอบภาพนั้น จะแบ่งออกเป็น 2 ส่วน คือ

1.1 การคำนวณหาจุดเด่นของเส้นขอบภาพที่สมบูรณ์เพื่อเก็บไว้ในฐานข้อมูล

ในสภาวะเริ่มต้นของการประมวลผลการจดจำภาพวัตถุ จำเป็นอย่างยิ่งที่จะต้องเตรียมข้อมูลที่ต้องเก็บไว้ในฐานข้อมูลเสียก่อน โดยการนำเอาภาพวัตถุที่สมบูรณ์แต่ละชิ้นมาทำการคำนวณหาจุดเด่นของภาพวัตถุ แล้วนำเอาจุดเด่นเหล่านั้นไปเก็บไว้ในฐานข้อมูล ซึ่งในขั้นตอนนี้จะทำการประมวลผลเพียงครั้งเดียวเท่านั้นในการเตรียมข้อมูลที่ต้องการให้กับฐานข้อมูล และฐานข้อมูลนี้สามารถที่จะขยายเพิ่มเติมหรือแก้ไขข้อมูลภายหลังได้ ผลของการสร้างฐานข้อมูลได้แสดงไว้ในภาพที่ 29 ถึง 31

1.2 การคำนวณหาจุดเด่นของเส้นขอบภาพที่มีวัตถุวางทับซ้อนกัน

การคำนวณหาจุดเด่นของเส้นขอบภาพที่มีวัตถุวางทับซ้อนกัน มีวิธีการและขั้นตอนเหมือนกับหัวข้อ 1.1 ซึ่งได้ผลดังที่ได้แสดงในภาพที่ 32

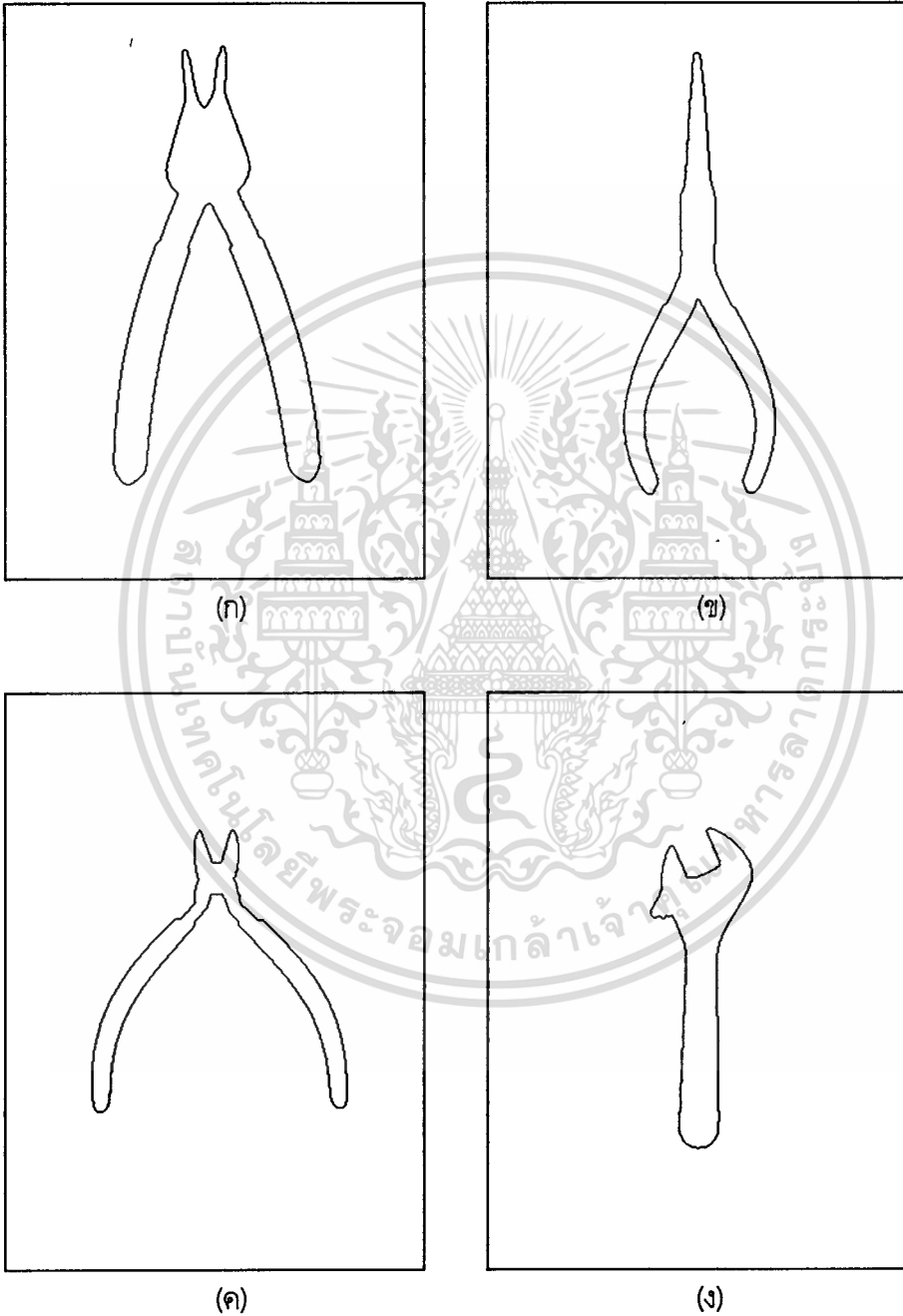
ภาพที่ 29



ภาพแสดงภาพต้นแบบระดับสีเทาที่ได้จากกล้อง CCD

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

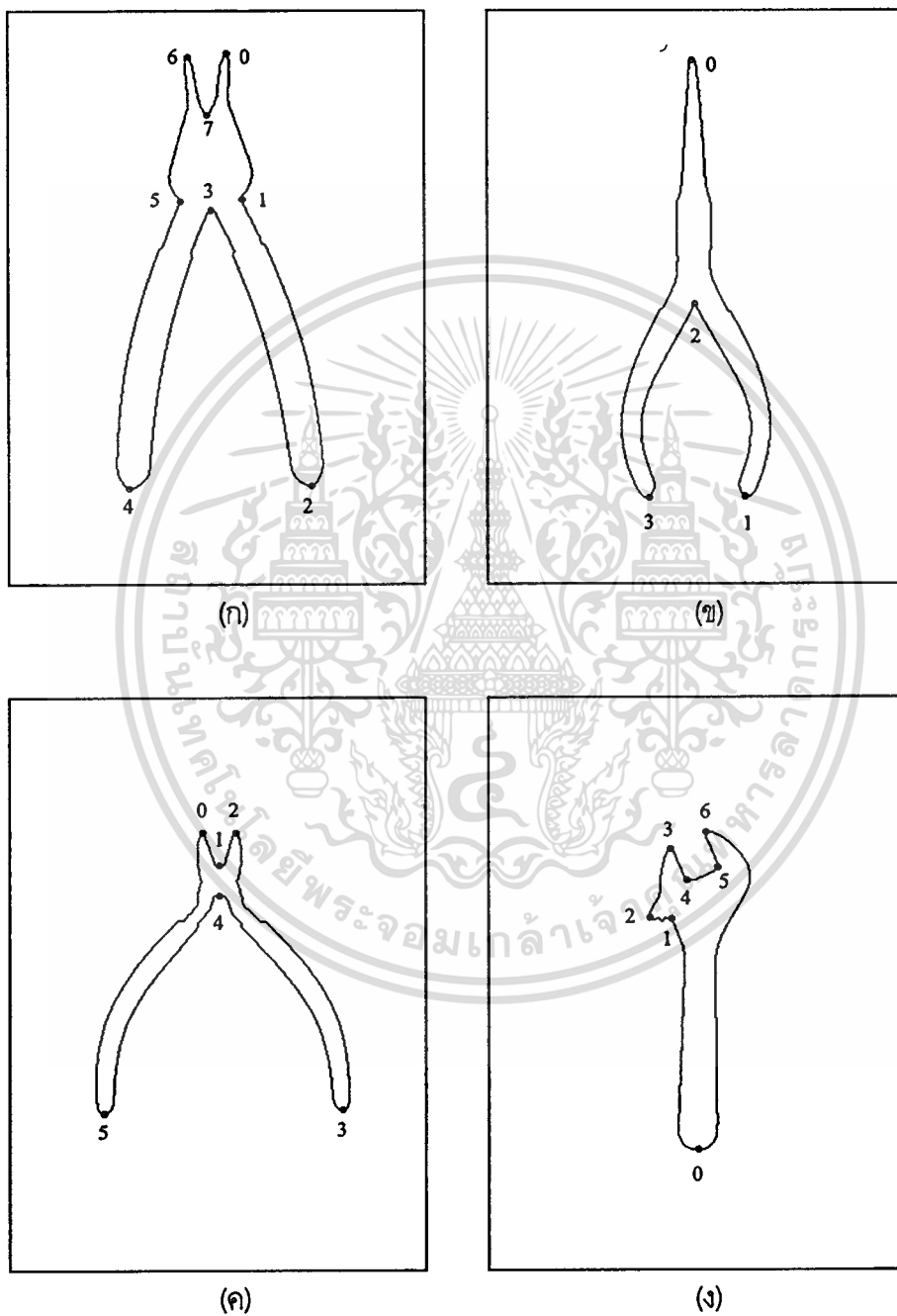
ภาพที่ 30



ภาพแสดงเส้นขอบภาพที่ได้จากภาพที่ 29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

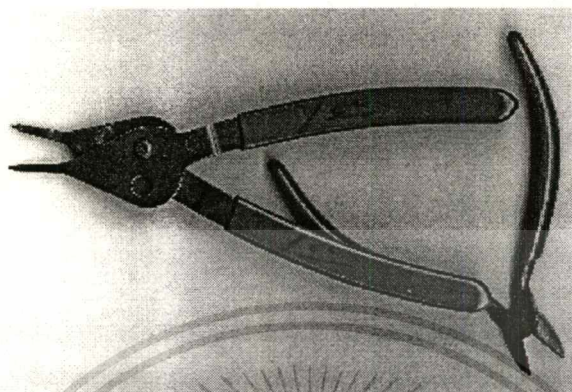
ภาพที่ 31



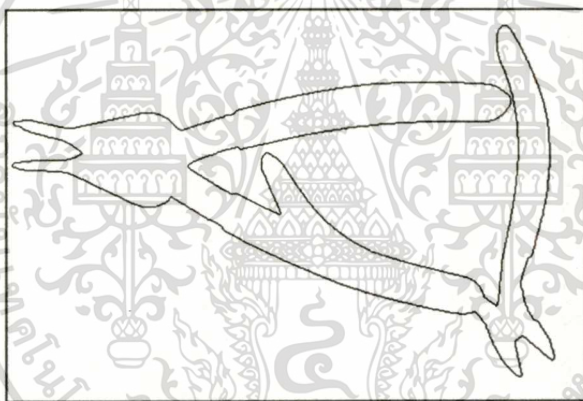
ภาพแสดงจุดเด่นของเส้นขอบภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

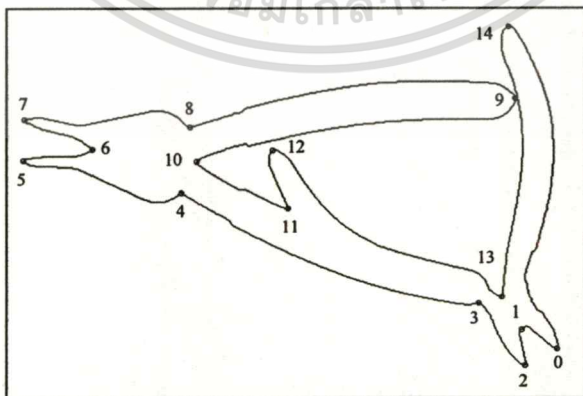
ภาพที่ 32



(ก)



(ข)



(ค)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับภาพแสดงภาพวัตถุที่วางทับซ้อนกันบางส่วน กรุณาอย่าให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การประมวลผลการจับคู่ของจุดเด่นเพื่อจดจำภาพวัตถุ

การประมวลผลการจับคู่ของจุดเด่น คือ ความพยายามที่จะจับคู่จุดเด่นระหว่างจุดเด่นของภาพวัตถุที่สมบูรณ์ในฐานข้อมูลกับจุดเด่นของภาพวัตถุที่ถูกทับซ้อนกันบางส่วน โดยมีเงื่อนไขในการจับคู่ที่สำคัญ 2 ประการ คือ

- (1) จุดเด่นของภาพวัตถุที่ถูกทับซ้อนกันบางส่วน จะต้องจับคู่กับจุดเด่นในฐานข้อมูลได้เพียงจุดเดียวเท่านั้น
- (2) จุดเด่นในฐานข้อมูล ก็จะต้องจับคู่กับจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนได้เพียงจุดเดียวเช่นกัน

จากเงื่อนไขทั้งสองประการที่ได้กล่าวมาแล้วนั้น เป็นการจับคู่แบบหนึ่งต่อหนึ่ง ซึ่งวิธีการที่มีประสิทธิภาพในการประมวลผลเพื่อการจับคู่ในลักษณะนี้ให้ถูกต้องและเหมาะสมที่สุด คือ การใช้โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ เนื่องด้วยลักษณะเด่นของการทำงานที่มีการนำเอาข้อมูลข้างเคียงเข้ามาพิจารณา ร่วมในการตัดสินใจ โดยใช้ค่าพารามิเตอร์ที่เป็นตัวกำหนดคุณลักษณะของจุดเด่นมาเป็นองค์ประกอบในการตัดสินใจ และมีสมการโฮปฟิลด์เป็นตัวควบคุม เพื่อให้ค่าพลังงานของโครงข่ายเซลล์ประสาทเทียมลู่เข้าสู่สภาวะพลังงานต่ำสุด และได้การจับคู่ที่เหมาะสมที่สุด ค่าพารามิเตอร์เหล่านี้ได้แก่ ค่าสเฟียริซิตี ค่าของมุม ค่าความโค้ง และทิศทางความโค้งของจุดเด่น อนึ่ง ปัจจัยสำคัญในการเลือกใช้ค่าพารามิเตอร์เหล่านี้ คือ ความสามารถในการตอบสนองต่อเงื่อนไขทั้ง 3 ประการของการจดจำภาพวัตถุ

2.1 ค่าพารามิเตอร์ที่ใช้ในการประมวลผลการจับคู่ของจุดเด่น

2.1.1 ค่าสเฟียริซิตี (Sphericity)

สเฟียริซิตี นิยามได้จากอัตราส่วนของค่าเฉลี่ยเรขาคณิต (Geometric Mean) ต่อค่าเฉลี่ยเลขคณิต (Arithmetic Mean) ของความยาวแกนหลักของวงรีภายในรูปสามเหลี่ยม

$$S = \frac{2\sqrt{a_1 a_2}}{a_1 + a_2} \quad (4-1)$$

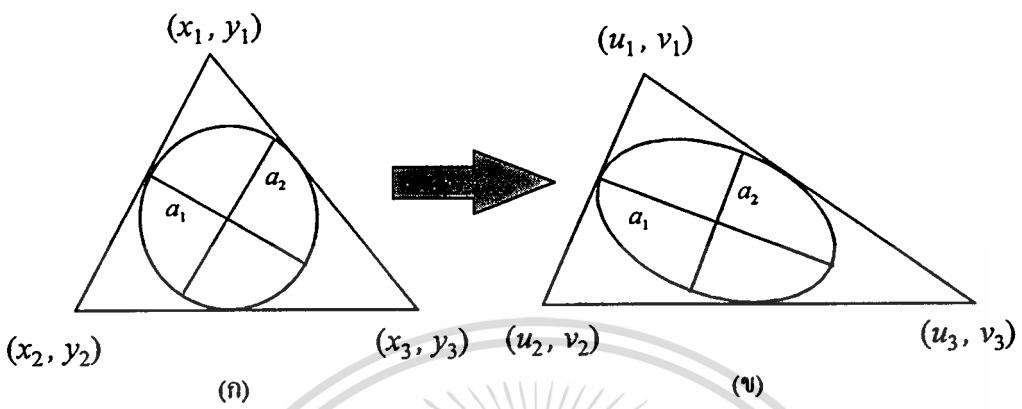
เมื่อ S คือ ค่าสเฟียริซิตี

a_1 และ a_2 คือ ความยาวแกนหลักของวงรี

จากนิยามที่กล่าวมานั้น สามารถนำเอาค่าของสเฟียริซิตี มาเป็นตัววัดความคล้ายหรือแตกต่างระหว่างรูปสามเหลี่ยมสองรูปใด ๆ ได้ ดังที่ได้แสดงไว้ในภาพที่ 33 โดยสรุปไว้ดังนี้

“ถ้าสามเหลี่ยมสองรูปใด ๆ มีความคล้ายกันค่า สเฟียริซิตี จะเท่ากับ ‘1’ ”

ภาพที่ 33



ภาพแสดงความสัมพันธ์ระหว่างรูปสามเหลี่ยมสองรูป

จากภาพที่ 33 สามารถพิจารณาความสัมพันธ์ของรูปสามเหลี่ยมรูป (ก) และ (ข) โดยอธิบายได้ ด้วย ความสัมพันธ์ในการแปลงรูป (Affine Transform) จาก x ไปยัง u ซึ่งกำหนดให้

$$u = Ax + t \tag{4-2}$$

เมื่อ $x = \begin{bmatrix} x \\ y \end{bmatrix}; u = \begin{bmatrix} u \\ v \end{bmatrix}; t = \begin{bmatrix} e \\ f \end{bmatrix}; A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ และ $\det(A) \neq 0$

ค่า ส.ป.ส. a, b, c, d, e และ f สามารถคำนวณได้จากสมการที่ (4-3)

$$\begin{bmatrix} a \\ b \\ e \end{bmatrix} = B^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ และ } \begin{bmatrix} c \\ d \\ f \end{bmatrix} = B^{-1} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{4-3}$$

เมื่อ $B = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$ และ $\det(B) \neq 0$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า (x_1, y_1) , (x_2, y_2) และ (x_3, y_3) คือจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุที่ทับซ้อนกัน เช่นเดียวกัน (u_1, v_1) , (u_2, v_2) และ (u_3, v_3) ก็คือจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุที่สมบูรณ์ซึ่งถูกเก็บไว้ในฐานข้อมูล

ถ้า $|A| > 0$ ในสมการที่ (4-2) เป็นจริง สามารถคำนวณค่าสเฟียริซิตี้ได้ดังนี้

$$S = \frac{t_1^2 + t_3^2 - (t_2^2 + t_4^2)}{(t_1^2 + t_2^2 + t_3^2 + t_4^2)} \quad (4-4)$$

เมื่อ $t_1 = a+d$; $t_2 = a-d$; $t_3 = b-c$; $t_4 = b+c$

2.1.2 ค่าความโค้ง (Curvature) และทิศทางความโค้งของจุดเด่น (Direction of Curvature)

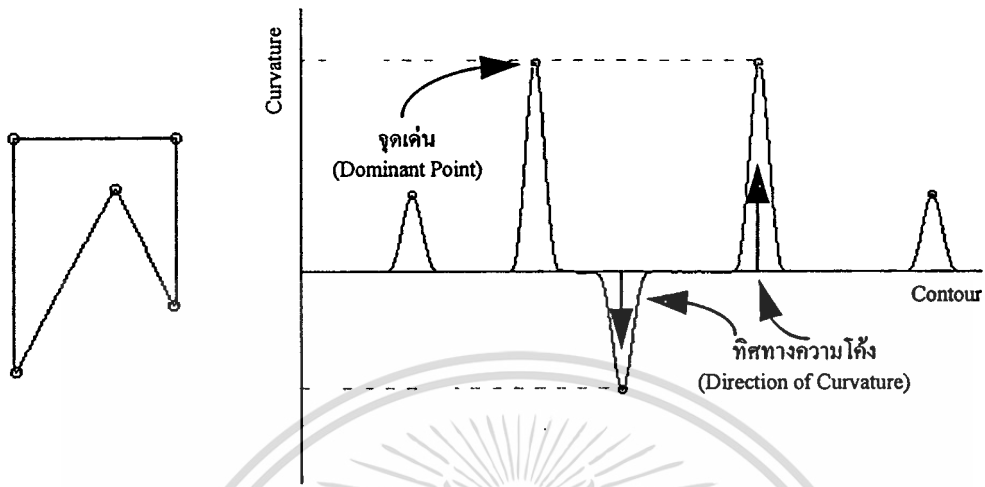
ค่าความโค้งและทิศทางความโค้งของจุดเด่น เป็นค่าพารามิเตอร์ที่สามารถบ่งบอกถึงความเหมือนหรือแตกต่างกันระหว่างจุดเด่นได้ โดยเฉพาะในการจับคู่ของจุดเด่น จุดเด่นที่ค่าความโค้งที่ใกล้เคียงกันและมีทิศทางความโค้งที่เหมือนกันเท่านั้นจึงจะสามารถจับคู่กันได้

$$F_1 = \begin{cases} 1 & \text{if } \left| \frac{Cu_m - Cu_s}{|Cu_m| + |Cu_s|} \right| < \lambda \text{ and} \\ & \text{if } (h_m \text{ and } h_s) \geq 0 \text{ or } (h_m \text{ and } h_s) < 0 \\ -1 & \text{Otherwise} \end{cases} \quad (4-5)$$

เมื่อ Cu_m คือ ค่าความโค้งของจุดเด่นที่ปรากฏในฐานข้อมูล
 Cu_s คือ ค่าความโค้งของจุดเด่นที่ปรากฏในภาพวัตถุที่ทับซ้อนกัน
 λ คือ ค่า Threshold ของค่าความโค้ง
 h_m และ h_s คือ ทิศทางของค่าความโค้งของจุดเด่นในฐานข้อมูล และภาพวัตถุที่ทับซ้อนกันตามลำดับ

ค่า Cu_m และ Cu_s คำนวณได้จากอัลกอริทึมที่สองและอัลกอริทึมที่สามในบทที่ 2

ภาพที่ 34

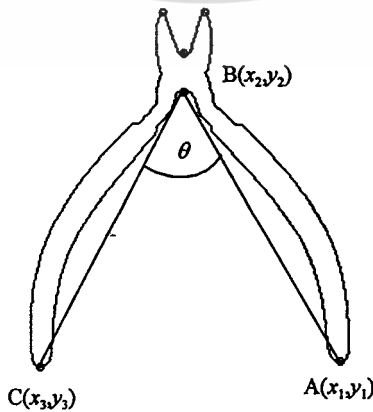


ภาพแสดงค่าความโค้งและทิศทางความโค้งของจุดเด่นรูปหลายเหลี่ยม

2.1.3 มุมของจุดเด่น (Corner)

มุมของจุดเด่น หมายถึง มุมที่เกิดจากการลากเส้นตรงเชื่อมต่อกันของจุดเด่นบนเส้นขอบภาพ ดังแสดงไว้ในภาพที่ 35 ดังนั้น มุมของจุดเด่นจึงเป็นค่าพารามิเตอร์อีกค่าหนึ่ง ที่สามารถนำมาเป็นตัวบ่งบอกคุณลักษณะของจุดเด่นได้ ซึ่งสามารถคำนวณหาได้ดังนี้

ภาพที่ 35



ภาพแสดงการคำนวณหามุมของจุดเด่นด้วยเวกเตอร์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อคุ้มครองสิทธิเท่านั้น เมื่อผู้ผู้ใดให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพที่ 35 เมื่อ $A(x_1, y_1)$, $B(x_2, y_2)$ และ $C(x_3, y_3)$ คือตำแหน่งของจุดเด่นที่สามารถคำนวณได้จากเส้นขอบภาพ จะได้ว่า

$$\overline{BA} \cdot \overline{BC} = |\overline{BA}| |\overline{BC}| \cos \theta \quad (4-6a)$$

$$\cos \theta = \frac{\overline{BA} \cdot \overline{BC}}{|\overline{BA}| |\overline{BC}|} \quad (4-6b)$$

เวกเตอร์ \overline{BA} และ \overline{BC} คำนวณได้จาก

$$\overline{BA} = (x_1 - x_2)\hat{i} + (y_1 - y_2)\hat{j} \quad (4-7)$$

และ

$$\overline{BC} = (x_3 - x_2)\hat{i} + (y_3 - y_2)\hat{j} \quad (4-8)$$

เมื่อ i และ j เป็นเวกเตอร์หนึ่งหน่วย

ดังนั้น

$$\overline{BA} \cdot \overline{BC} = (x_1 - x_2)(x_3 - x_2) + (y_1 - y_2)(y_3 - y_2) \quad (4-9)$$

และ

$$\theta = \cos^{-1} \left(\frac{(x_1 - x_2)(x_3 - x_2) + (y_1 - y_2)(y_3 - y_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}} \right) \quad (4-10)$$

$$F_2 = \begin{cases} 1 & \text{if } |Co_m - Co_s| < \beta \\ -1 & \text{Otherwise} \end{cases} \quad (4-11)$$

- เมื่อ Co_m คือ มุมของจุดเด่นที่ปรากฏในฐานข้อมูล
 Co_s คือ มุมของจุดเด่นที่ปรากฏในภาพวัตถุที่ทับซ้อนกัน
 β คือ ค่า Threshold ของผลต่างของมุมที่เกิดขึ้น

ค่า Co_m และ Co_s คำนวณได้จากสมการ (4-10)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 ฟังก์ชันพลังงานและสมการของโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

สมการที่ใช้ในการจับคู่ หรือบางครั้งมักจะถูกเรียกว่า “ฟังก์ชันพลังงาน” (Energy Function) ถูกนิยามไว้ดังนี้

$$e = -\sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N C_{ijkl} V_{ik} V_{jl} + \sum_{i=1}^M \left(1 - \sum_{k=1}^N V_{ik}\right)^2 + \sum_{k=1}^N \left(1 - \sum_{i=1}^M V_{ik}\right)^2 \quad (4-12)$$

เมื่อนิพจน์แรกของสมการ (4-12) ทำหน้าที่ในการจับคู่ระหว่าง i กับ k โดยมี C_{ijkl} เป็นตัวควบคุมของการจับคู่ และ V_{jl} เป็นข้อมูลข้างเคียงประกอบการพิจารณา ส่วนนิพจน์ที่ 2 และ 3 จะทำหน้าที่บังคับให้คำตอบที่ได้นั้น จะต้องมียกเว้นคำตอบเดียวและถูกต้องด้วย ซึ่งเป็นเงื่อนไขในการจับคู่แบบหนึ่งต่อหนึ่งนั่นเอง

จากสมการที่ (4-12) สามารถลดรูป เพื่อให้สามารถเทียบเคียงกับสมการของโฮปฟิลด์ สมการที่ (3-3) ในบทที่ 3 ได้ดังนี้

$$\begin{aligned} \sum_{i=1}^M \left(1 - \sum_{k=1}^N V_{ik}\right)^2 &= \sum_{i=1}^M \left(1 - 2 \sum_{k=1}^N V_{ik} + \left(\sum_{k=1}^N V_{ik}\right)^2\right) \quad (4-12a) \\ &= \sum_{i=1}^M (1) - 2 \sum_{i=1}^M \sum_{k=1}^N V_{ik} + \sum_{i=1}^M \left(\sum_{k=1}^N V_{ik}\right)^2 \\ &= \sum_{i=1}^M (1) - 2 \sum_{i=1}^M \sum_{k=1}^N V_{ik} + \sum_{i=1}^M \left(\sum_{k=1}^N \sum_{l=1}^N V_{ik} V_{il}\right) \\ &= M - 2 \sum_{i=1}^M \sum_{k=1}^N V_{ik} + \sum_{i=1}^M \sum_{k=1}^N \sum_{l=1}^N V_{ik} V_{il} \end{aligned}$$

ถ้ากำหนดให้ $V_{ik} V_{il} = V_{ik} V_{jl} \delta_{ij}$ จะได้ว่า

$$= M + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{ij} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \quad (4-12b)$$

และในทำนองเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\sum_{k=1}^N \left(1 - \sum_{i=1}^M V_{ik}\right)^2 = N + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{kl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \quad (4-12c)$$

แทนค่าสมการที่ (4-12b) และ (4-12c) กลับไปยังสมการที่ (4-12) จะได้ว่า

$$\begin{aligned} e &= M + N - \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N C_{ijkl} V_{ik} V_{jl} \\ &+ \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{ij} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \quad (4-12d) \\ &+ \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{kl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \end{aligned}$$

$$e = M + N - \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N (C_{ijkl} - \delta_{ij} - \delta_{kl}) V_{ik} V_{jl} - \sum_{i=1}^M \sum_{k=1}^N 4V_{ik} \quad (4-12e)$$

$$\frac{e - (M + N)}{2} = -\frac{1}{2} \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N (C_{ijkl} - \delta_{ij} - \delta_{kl}) V_{ik} V_{jl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \quad (4-13)$$

ซึ่งจะเห็นได้ว่า สมการที่ (4-13) สมมูลกับสมการที่ (3-3) ในบทที่ 3 ทุกประการ เมื่อ $T'_{ijkl} = C_{ijkl} - \delta_{ij} - \delta_{kl}$ และ $I_{ik} = 2$

2.3 อัลกอริทึมของการประมวลผลการจับคู่ด้วยโครงข่ายเซลล์ประสาทเทียมแบบไฮโปฟิลด์

ลำดับขั้นที่ 1 ให้กำหนดสถานะเริ่มต้นให้กับ V_{ik} (เซลล์ประสาท) ในโครงข่ายเซลล์ประสาทเทียมแบบไฮโปฟิลด์ ซึ่งมีขนาด $M \times N$ ภายใต้เงื่อนไขดังสมการ (4-14) เมื่อ M คือ จำนวนของจุดเด่นที่คำนวณได้จากภาพวัตถุที่ถูกทับซ้อน และ N คือ จำนวนจุดเด่นทั้งหมดของภาพที่เก็บไว้ในฐานข้อมูล

$$V_{ik} = \begin{cases} 1 & \text{if } Sphericity \geq \alpha \\ 0 & \text{Otherwise} \end{cases} \quad (4-14)$$

เมื่อ α คือ ค่า Threshold ของสเฟียริซิตี และมีค่าอยู่ระหว่าง $0 \leq \alpha \leq 1$

ในภาพที่ 36(ก) ได้แสดงให้เห็นการจับคู่ของจุดเด่นในภาพที่ 31(ก, ข, ค และ ง) ซึ่งเก็บไว้เป็นฐานข้อมูล กับจุดเด่นของภาพที่ 32(ข) ซึ่งเป็นภาพที่ถูกบดบังบางส่วน และมีจุดเด่นดังแสดงไว้ในภาพที่ 32(ค) โดยจะกำหนดให้ V_{ik} มีค่าเท่ากับ "1" หรือ "0" ตามเงื่อนไขของสมการที่ (4-14) ส่วนค่าของสเฟียริซิตีคำนวณ

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้จากสมการที่ (4-4) ในหัวข้อที่ 2.1.1 $V_{ik}=1$ หมายความว่าจุดเด่น ณ ตำแหน่ง i ใน ภาพที่ 32(ค) สามารถจับคู่กับจุดเด่นในตำแหน่ง k ใน ภาพฐานข้อมูล (ภาพที่ 31(ก, ข, ค และ ง)) ได้ ซึ่งจะเห็นว่า ในสภาวะเริ่มต้นของการกำหนดด้วยสมการที่ (4-14) ยังมีการจับคู่ที่ไม่เหมาะสม (Mismatch) อยู่ ดังแสดงให้เห็นในภาพที่ 36(ก) ซึ่งการจับคู่ที่ไม่เหมาะสมเหล่านี้ จะถูกกำจัดทิ้งไปด้วยสมการของไฮบิลด์ในลำดับขั้นตอนที่ 2 ถึง 4

แต่ถ้าพิจารณาพฤติกรรมในการจับคู่ของจุดเด่นให้ดีแล้วจะพบว่า จุดเด่นที่ควรจะจับคู่กันได้นั้น มักจะอยู่เป็นกลุ่มในแนว -45 องศา ของเมตริกซ์ $M \times N$ ในแต่ละภาพในฐานข้อมูล ดังนั้น จากพฤติกรรมดังกล่าว จึงสามารถกำจัดการจับคู่ที่ไม่เหมาะสมออกไปได้บางส่วน ซึ่งได้แสดงไว้ในภาพที่ 36(ข) และเรียกขั้นตอนนี้ว่า "การกรองเบื้องต้น" (Pre-filtering)

การกรองเบื้องต้นเป็นการกำหนดสภาวะเริ่มต้น ให้กับโครงข่ายเซลล์ประสาทเทียม เพื่อเป็นการกำจัดการจับคู่ที่ไม่เหมาะสมที่ได้จากสมการที่ (4-14) อีกครั้งหนึ่ง ซึ่งจะให้ผลดีสองประการ คือ เพิ่มความแม่นยำให้กับการทำงานของโครงข่ายเซลล์ประสาท และช่วยให้จำนวนครั้งของการทำงานซ้ำ (Iteration) ของโครงข่ายลดลง

ลำดับขั้นที่ 2 ให้เลือกตำแหน่งเซลล์ประสาท (i, k) ที่มีค่าเป็น "1" ใด ๆ ขึ้นมา แล้วคำนวณหาค่า U_{ik} ตามสมการที่ (4-15)

$$U_{ik} = \sum_{j=1}^M \sum_{l=1}^N (C_{ikjl} - \delta_{ij} - \delta_{kl}) V_{jl} + 2 \quad (4-15)$$

เมื่อ δ_{ij} และ δ_{kl} เป็น Kronecker delta function

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{Otherwise} \end{cases} \quad (4-15a)$$

$$\delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{Otherwise} \end{cases} \quad (4-15b)$$

และ

$$C_{ikjl} = \begin{cases} W_1 F_1(i, k) + W_2 F_1(j, l) + \\ W_3 F_2(i, k) + W_4 F_2(j, l) & \text{for } k \text{ and } l \in \text{the same model} \\ 0 & \text{for } k \text{ and } l \notin \text{the same model} \end{cases} \quad (4-15c)$$

เมื่อ $W_1 + W_2 + W_3 + W_4 = 1$ เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งรายละเอียดของ F_1 และ F_2 ได้อธิบายไว้ในหัวข้อที่ 2.1 ส่วนค่า W_1, W_2, W_3 และ W_4 จะเป็นค่าน้ำหนัก (Weight) ที่ใช้ถ่วงความสำคัญของ F_1 , และ F_2 ตามลำดับ

ลำดับขั้นที่ 3 ให้เปลี่ยนสถานะของ V_{ik} ตามสมการข้างล่างนี้

$$\begin{array}{lll} V_{ik} \rightarrow 0 & \text{ถ้า} & U_{ik} > 0.25 \\ V_{ik} \rightarrow 1 & \text{ถ้า} & U_{ik} < -0.25 \\ V_{ik} \rightarrow \text{no change} & \text{ถ้า} & -0.25 \leq U_{ik} \leq 0.25 \end{array}$$

เมื่อค่าตัวเลข "0.25" คือค่า Threshold ที่ได้จากการทดลอง

ลำดับขั้นที่ 4 ให้ตรวจสอบทุก ๆ V_{ik} ว่ามีการเปลี่ยนแปลงสถานะอยู่หรือไม่ ถ้ายังมีการเปลี่ยนแปลงอยู่ ให้กลับไปดำเนินการตาม ลำดับขั้นที่ 2 ใหม่ แต่ถ้าไม่มีการเปลี่ยนแปลง แสดงว่า V_{ik} ได้เข้าสู่สภาวะพลังงานต่ำสุดแล้ว ให้ดำเนินการต่อไป ลำดับขั้นที่ 5

ลำดับขั้นที่ 5 ลำดับขั้นนี้จะเป็นคำตอบสุดท้ายของ V_{ik} ที่มีการจับคู่ระหว่างจุดเด่นในฐานข้อมูลกับจุดเด่นของภาพ ในภาพที่ 32(ค) ซึ่งจะได้ผลดังภาพที่ 36(ค)

จากภาพที่ 36(ค) เมื่อพิจารณาเมตริกซ์ที่ได้ จะพบว่า ในแต่ละแถวจะปรากฏ "1" ไม่เกินหนึ่งค่า และในทำนองเดียวกัน ในแต่ละหลักก็ปรากฏ "1" ไม่เกินหนึ่งค่าเช่นกัน ซึ่งเมตริกซ์ที่มีสมาชิกจัดเรียงในลักษณะดังกล่าว จะเรียกว่า "Permutation Matrix" และจะให้ผลการจับคู่เป็นแบบหนึ่งต่อหนึ่งตามเงื่อนไขการจับคู่ที่ได้กล่าวมาแล้วในตอนต้น

จากผลการจับคู่ของจุดเด่นที่แสดงในภาพที่ 36(ค) เป็นเครื่องยืนยันให้เห็นว่า ภาพที่ 32(ค) มีวัตถุทับซ้อนกันอยู่สองชิ้น และภาพที่ 36(ค) ยังเป็นข้อมูลสำคัญในการตัดสินใจขั้นสุดท้ายในการประมวลผลการจดจำและบอกตำแหน่งของวัตถุที่ทับซ้อนกันบางส่วน ดังจะได้กล่าวรายละเอียดในหัวข้อถัดไป

ภาพที่ 36

	(ก)	(ข)	(ค)	(ง)		(ก)	(ข)	(ค)	(ง)	
	01234567	0123	012345	0123456	N	01234567	0123	012345	0123456	N
0	00000000	0000	100101	0000110		0	00000000	0000	100100	0000100
1	10001000	0000	010000	0000010		1	00000000	0000	010000	0000010
2	00000000	0000	101101	0000010		2	00000000	0000	001000	0000000
3	10001000	0000	000000	0000010		3	00001000	0000	000000	0000000
4	00010100	0000	000000	0000000		4	00000100	0000	000000	0000000
5	00100010	0100	000000	0000000		5	00000010	0000	000000	0000000
6	00000001	0000	000000	0000000		6	00000001	0000	000000	0000000
7	11011000	0001	000010	0000000		7	10000000	0000	000000	0000000
8	01010001	0000	000000	0000000		8	01000000	0000	000000	0000000
9	00100000	0000	000000	0000000		9	00100000	0000	000000	0000000
10	01010100	0000	100000	0000001		10	00010000	0000	000000	0000000
11	00000000	0000	001111	0000000		11	00000000	0000	001011	0000000
12	00000010	0000	101101	0000000		12	00000000	0000	100101	0000000
13	10001000	0000	000010	0000010		13	00001000	0000	000010	0000000
14	00110110	0100	101101	0000000		14	00000100	0000	001001	0000000

	(ก)	(ข)	(ค)	(ง)	
	01234567	0123	012345	0123456	N
0	00000000	0000	100000	0000000	
1	00000000	0000	010000	0000000	
2	00000000	0000	001000	0000000	
3	00000000	0000	000000	0000000	
4	00000000	0000	000000	0000000	
5	00000010	0000	000000	0000000	
6	00000001	0000	000000	0000000	
7	10000000	0000	000000	0000000	
8	01000000	0000	000000	0000000	
9	00100000	0000	000000	0000000	
10	00010000	0000	000000	0000000	
11	00000000	0000	000000	0000000	
12	00000000	0000	000100	0000000	
13	00000000	0000	000000	0000000	
14	00000000	0000	000001	0000000	

(ค)

(ก) แสดงการกำหนดสถานะเริ่มต้นของการจับคู่ให้กับเซลล์ประสาทของโครงข่าย

(ข) แสดงผลที่ได้จากการประมวลผล "การกรองเบื้องต้น"

(ค) แสดงผลของการเข้าสู่สถานะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม

การคำนวณหาตำแหน่งของวัตถุ

ในการประมวลผลการจดจำภาพวัตถุ จำเป็นอย่างยิ่งที่จะต้องชี้ตำแหน่งของวัตถุ บอกขนาดของวัตถุ ตลอดจนทิศทางในการทับซ้อนกันของวัตถุได้ ภายหลังจากประมวลผลการจับคู่ของจุดเด่นแล้ว ซึ่งมีวิธีการคำนวณหาตำแหน่งของวัตถุดังจะได้กล่าวต่อไปนี้

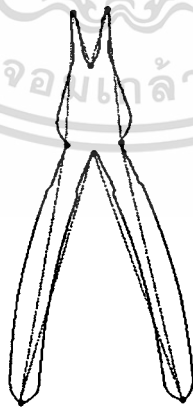
1. การคำนวณหาจุดศูนย์กลางของรูปหลายเหลี่ยม

จุดศูนย์กลางของรูปหลายเหลี่ยม ที่เกิดจากการเชื่อมต่อกันของจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุ ดังแสดงไว้ในภาพที่ 37 สามารถคำนวณได้ดังนี้

$$X_g = \frac{1}{n} \sum_{i=1}^n x_i ; \quad Y_g = \frac{1}{n} \sum_{i=1}^n y_i \quad (4-16)$$

เมื่อ x_i และ y_i คือ พิกัดในแนวแกน x และ y ของจุดเด่นตามลำดับ
 n คือ จำนวนจุดเด่นทั้งหมดที่ปรากฏในภาพ

ภาพที่ 37



ภาพแสดงรูปหลายเหลี่ยมที่เกิดจากการเชื่อมต่อกันของจุดเด่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การแปลงเส้นขอบภาพในระบบพิกัดโฮโมจีเนียสสำหรับภาพ 2 มิติ

2.1 การแปลงสำหรับการเลื่อนตำแหน่งเส้นขอบภาพ

การแปลงสำหรับการเลื่อนตำแหน่งเส้นขอบภาพ มีวิธีการดังที่ได้แสดงไว้ในภาพที่ 38 ซึ่งสามารถคำนวณได้จากเมตริกซ์การเลื่อนตำแหน่ง ดังนี้

$$P' = PT_1 \quad (4-17)$$

$$P' = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ M & N & 1 \end{bmatrix} \quad (4-18)$$

เมื่อ P' คือ เมตริกซ์ที่แสดงถึงตำแหน่งใหม่ของจุดภาพบนเส้นขอบภาพตามลำดับ

P คือ เมตริกซ์ของตำแหน่งจุดภาพบนเส้นขอบภาพเดิมก่อนที่จะทำการเลื่อนไป

T_1 คือ เมตริกซ์การแปลงสำหรับการเลื่อนเส้นขอบภาพ

M, N คือ พิกัดในแนวแกน x และ y ตามลำดับ ที่ต้องการย้ายไปยังตำแหน่งใหม่

2.2 การแปลงผกผันสำหรับการเลื่อนเส้นขอบภาพ

การแปลงผกผันสำหรับการเลื่อนเส้นขอบภาพ สามารถคำนวณได้จากเมตริกซ์ดังนี้

$$P' = PT_{-1} \quad (4-19)$$

$$P' = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -M & -N & 1 \end{bmatrix} \quad (4-20)$$

เมื่อ P' คือ เมตริกซ์ที่แสดงถึงตำแหน่งใหม่ของจุดภาพบนเส้นขอบภาพตามลำดับ

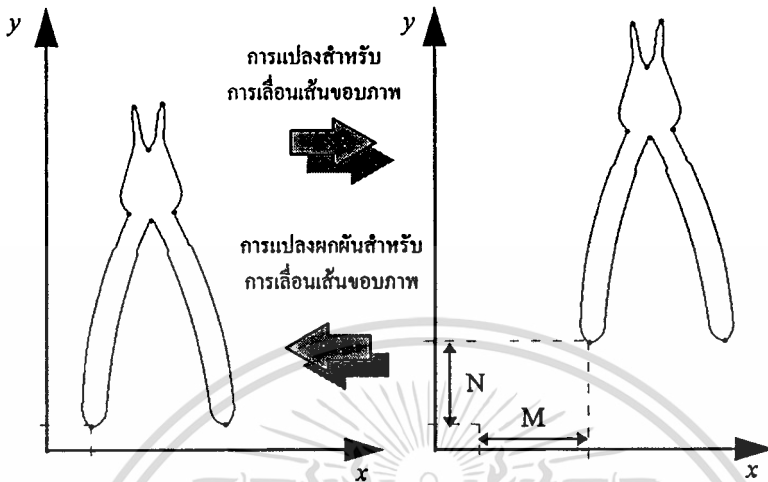
P คือ เมตริกซ์ของตำแหน่งจุดภาพบนเส้นขอบภาพเดิมก่อนที่จะทำการเลื่อนไป

T_{-1} คือ เมตริกซ์การแปลงผกผันสำหรับการเลื่อนเส้นขอบภาพ

$-M, -N$ คือ พิกัดในแนวแกน x และ y ตามลำดับ ที่ต้องการย้ายไปยังตำแหน่งใหม่

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง การนำเอกสารนี้ไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย

ภาพที่ 38



ภาพแสดงการแปลงและการแปลงพิกัดสำหรับการเลื่อนเส้นขอบภาพ

2.3 การแปลงสำหรับการหมุนเส้นขอบภาพ

การแปลงสำหรับการหมุนเส้นขอบภาพ ดังที่ได้แสดงไว้ในภาพที่ 39 สามารถคำนวณได้จากเมตริกซ์สำหรับการหมุนดังนี้

$$P' = PT_r \quad (4-21)$$

$$P' = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4-22)$$

- เมื่อ P' คือ เมตริกซ์ที่แสดงถึงตำแหน่งใหม่ของจุดภาพบนเส้นขอบภาพตามลำดับ
- P คือ เมตริกซ์ของตำแหน่งจุดภาพบนเส้นขอบภาพเดิม ก่อนที่จะทำการเลื่อนไป
- T_r คือ เมตริกซ์การแปลงสำหรับการหมุนเส้นขอบภาพ
- θ คือ มุมที่ต้องการหมุน

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์เพื่อการเรียนการสอนเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 การแปลงผกผันสำหรับการหมุนเส้นขอบภาพ

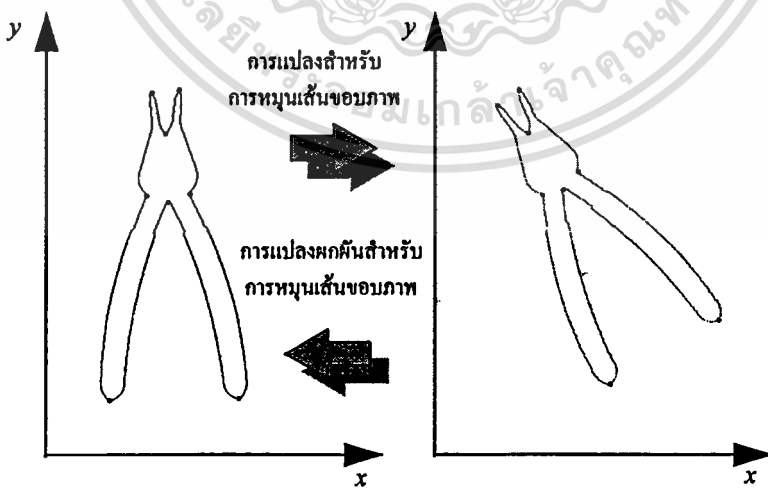
ในการคำนวณการแปลงผกผันสำหรับการหมุน สามารถทำได้โดยการแทนค่ามุม θ ด้วยค่าที่เป็นลบ ซึ่งคำนวณได้จากเมตริกซ์สำหรับการหมุนดังนี้

$$P' = PT_{r-1} \quad (4-23)$$

$$P' = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4-24)$$

- เมื่อ P' คือ เมตริกซ์ที่แสดงถึงตำแหน่งใหม่ของจุดภาพบนเส้นขอบภาพตามลำดับ
 P คือ เมตริกซ์ของตำแหน่งจุดภาพบนเส้นขอบภาพเดิม ก่อนที่จะทำการเลื่อนไป
 T_{r-1} คือ เมตริกซ์การแปลงผกผันสำหรับการหมุนเส้นขอบภาพ
 θ คือ มุมที่ต้องการหมุน

ภาพที่ 39



ภาพแสดงการแปลงและการแปลงผกผันสำหรับการหมุนเส้นขอบภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 การแปลงสำหรับการย่อ/ขยายเส้นขอบภาพ

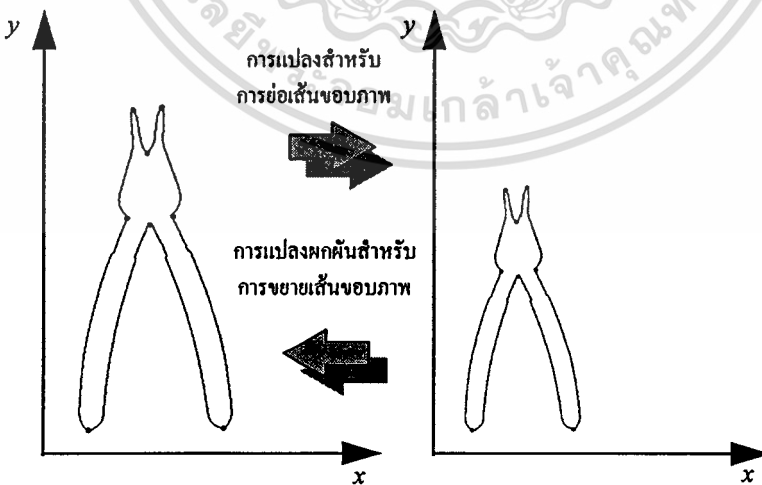
การแปลงสำหรับการย่อ/ขยายเส้นขอบภาพ คำนวณได้จากเมตริกซ์สำหรับการย่อ/ขยายดังนี้

$$P' = PT_S \quad (4-25)$$

$$P' = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4-26)$$

- เมื่อ P' คือ เมตริกซ์ที่แสดงถึงตำแหน่งใหม่ของจุดภาพบนเส้นขอบภาพตามลำดับ
 P คือ เมตริกซ์ของตำแหน่งจุดภาพบนเส้นขอบภาพเดิมก่อนที่จะทำการเลื่อนไป
 T_S คือ เมตริกซ์การแปลงสำหรับการเลื่อนเส้นขอบภาพ
 S_x, S_y คือ อัตราย่อ/ขยายในแนวแกน x และ y ตามลำดับ ที่ต้องย่อหรือขยาย
 (ย่อ : x และ $y < 1$, ขยาย : x และ $y > 1$)

ภาพที่ 40



ภาพแสดงการแปลงสำหรับการย่อ/ขยายเส้นขอบภาพ

2.6 การแปลงเส้นขอบภาพในฐานข้อมูลเพื่อหาตำแหน่งของภาพวัตถุที่ทับซ้อนกัน

การแปลงเส้นขอบภาพในฐานข้อมูลที่สามารถจดจำได้ เพื่อหาตำแหน่งของวัตถุที่ถูกทับซ้อนกัน สามารถคำนวณได้จากเมตริกซ์สำหรับการแปลงดังนี้

$$P_s = P_m T \quad \text{และ} \quad T = T_t T_r T_s \quad (4-27)$$

เมื่อ P_m คือ เมตริกซ์แสดงตำแหน่งของจุดภาพบนเส้นขอบภาพของวัตถุที่ทับซ้อนกัน

P_s คือ เมตริกซ์แสดงตำแหน่งของจุดภาพบนเส้นขอบภาพของวัตถุในฐานข้อมูล

โดย
$$T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad (4-27a)$$

เมื่อ
$$t_x = \left(\frac{1}{m} \sum_{i=1}^m x_{s_i} \right) - \left(\frac{1}{m} \sum_{i=1}^m x_{m_i} \right); \quad t_y = \left(\frac{1}{m} \sum_{i=1}^m y_{s_i} \right) - \left(\frac{1}{m} \sum_{i=1}^m y_{m_i} \right)$$

และ m คือ จำนวนจุดเด่นที่สามารถจับคู่ได้

$$T_r = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4-27b)$$

เมื่อ θ คือ มุมที่ต้องการจะหมุนไปยังตำแหน่งของวัตถุที่ถูกทับซ้อนกัน

และ
$$T_s = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4-27c)$$

เมื่อ S_x, S_y คือ อัตราการย่อ/ขยายในพิกัดแกน x และ y ตามลำดับ

ดังนั้น
$$T = \begin{bmatrix} S_x \cos \theta & -S_y \sin \theta & 0 \\ S_x \sin \theta & S_y \cos \theta & 0 \\ S_x(t_x \cos \theta + t_y \sin \theta) & S_y(-t_x \sin \theta + t_y \cos \theta) & 1 \end{bmatrix} \quad (4-27d)$$

การคำนวณหาค่าความผิดพลาดในการจับคู่ของจุดเด่น

ค่าความผิดพลาดอันเนื่องมาจาก การเลื่อน/ย้ายตำแหน่ง การหมุน และการย่อ/ขยายของเส้นขอบภาพในฐานข้อมูล เพื่อบอกตำแหน่งของภาพวัตถุที่ทับซ้อนกัน สามารถคำนวณได้จาก ค่าความผิดพลาดกำลังสอง (Square Error) ได้ดังนี้

$$\varepsilon = \sum_{i=0}^{m-1} \left((x_{si} - x_{ci})^2 + (y_{si} - y_{ci})^2 \right) \quad (4-28)$$

- เมื่อ ε คือ ค่าความผิดพลาดกำลังสอง
 m คือ จำนวนของจุดเด่นที่สามารถจับคู่ได้
 x_{si}, y_{si} คือ ตำแหน่งของจุดเด่นในแนวแกน x และ y ของภาพวัตถุที่ทับซ้อนกัน
 x_{ci}, y_{ci} คือ ค่าที่คำนวณได้จากการแปลงคอคออดิเนทของจุดเด่นในฐานข้อมูล ไปยังตำแหน่ง x_{si} และ y_{si} ตามลำดับ

ถึงแม้ว่า ในกระบวนการการจับคู่ของจุดเด่น โดยวิธีใช้โครงข่ายเซลล์ประสาทเทียมแบบไฮโปฟิลด์ ซึ่งเป็นวิธีการแก้ไขปัญหามีประสิทธิภาพ แต่เพื่อความแน่นอนในกระบวนการจับคู่ ว่ามีความถูกต้องเพียงใด จึงจำเป็นต้องมีการวัดความผิดพลาดในการจับคู่ของจุดเด่นโดยใช้สมการต่อไปนี้

$$\varepsilon' = \begin{cases} \left(1 + \left(\frac{n-2}{m-2} \right) \log_2 \left(\frac{n-2}{m-2} \right) \right) \bar{\varepsilon} & \text{for } m \geq 3 \\ \infty & \text{for } m = 0, 1, 2 \end{cases} \quad (4-29)^1$$

- เมื่อ ε' คือ ค่าความผิดพลาดในการจับคู่
 n คือ จำนวนจุดเด่นทั้งหมด ในแต่ละรูปในฐานข้อมูล
 m คือ จำนวนจุดเด่นที่สามารถจับคู่ได้
 $\bar{\varepsilon}$ คือ Normalized Square Error

โดยที่ $\bar{\varepsilon} = \frac{\varepsilon}{k}$; เมื่อ k คือ อัตราย่อ/ขยายของภาพ

¹ Nirwan Ansari and Edward J. Delp, "Partial Shape Recognition : A Landmark-Based Approach," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No. 5 (May 1990) : P. 479.

ซึ่งผลที่คำนวณได้จากสมการที่ (4-29) แสดงไว้ในตารางที่ 1 และภาพแสดงการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น ได้แสดงในภาพที่ 41

ตารางที่ 1

รูปในฐานข้อมูล (ภาพที่ 31)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ε'
ก	8	6	65.8880
ข	4	0	∞
ค	6	5	65.1900
ง	7	0	∞

ตารางสรุปค่าความผิดพลาดในการจับคู่ของจุดเด่น

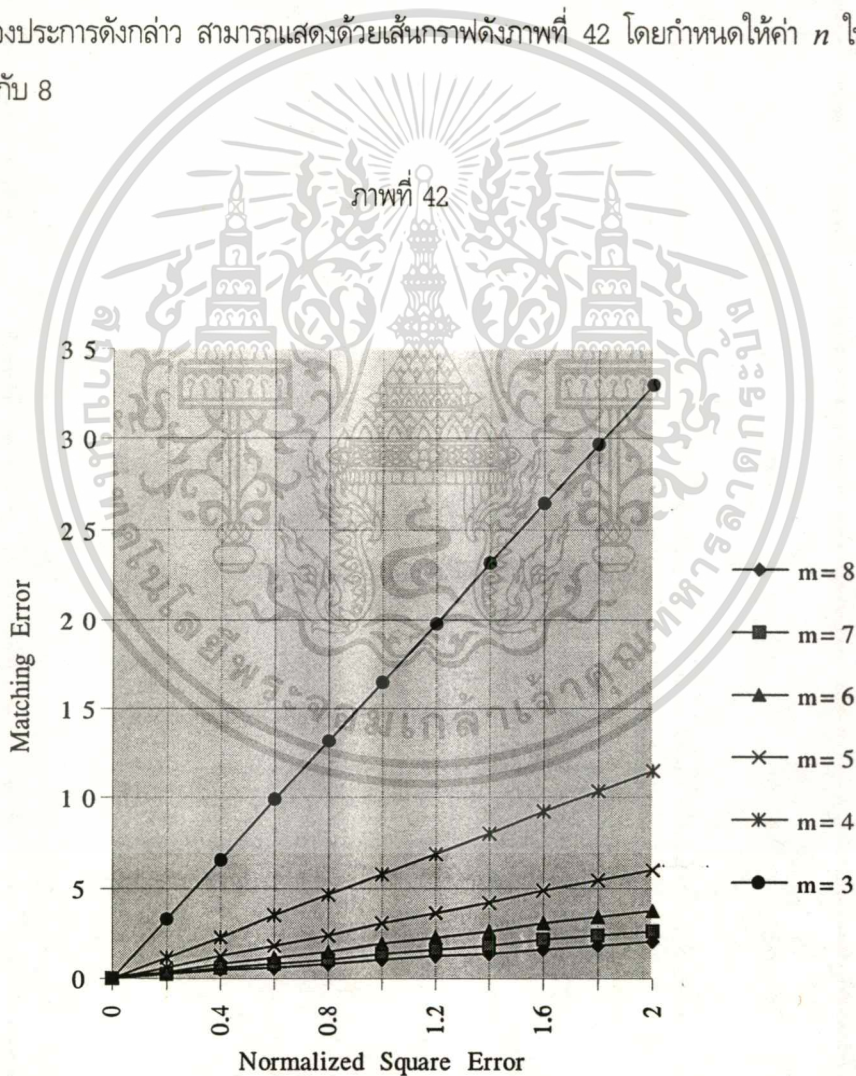
ภาพที่ 41



ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น

จากผลการคำนวณของสมการที่ (4-29) ซึ่งแสดงไว้ในตารางที่ 1 เพื่อเป็นการวัดผลขั้นสุดท้ายในการตัดสินใจ เพื่อบ่งบอกถึงความถูกต้องในการจดจำภาพวัตถุ อันเนื่องมาจากการประมวลผลการจับคู่ของจุดเด่นด้วยโครงข่ายเซลล์ประสาทเทียม เราพบว่า ถ้า ε' มีค่ามาก แสดงถึงผลของการประมวลผลการจับคู่ของจุดเด่นที่ผิดพลาดสูงเกินไป ซึ่งอาจเกิดจากการที่จำนวนจุดเด่นที่เลือกไม่เหมาะสม หรือการที่ค่า ε' ที่เลือกไม่เหมาะสม อย่างไรก็ตาม การปรับปรุงประสิทธิภาพของโครงข่ายเซลล์ประสาทเทียม อาจทำได้โดยการปรับค่า ε' ให้เหมาะสม หรือการเพิ่มจำนวนจุดเด่นที่เลือกให้มากขึ้น นอกจากนี้ การปรับปรุงประสิทธิภาพของโครงข่ายเซลล์ประสาทเทียม ยังสามารถทำได้โดยการปรับค่า ε' ให้เหมาะสม หรือการเพิ่มจำนวนจุดเด่นที่เลือกให้มากขึ้น

เด่น มีความน่าเชื่อถือน้อย แต่ในทางตรงกันข้าม ถ้า ε' มีค่าต่ำ แสดงถึงผลของการประมวลผลการจับคู่ของจุดเด่น มีความน่าเชื่อถือสูง แต่เมื่อพิจารณาสมการที่ (4-29) ให้ดีแล้วจะพบอีกว่า ปัจจัยที่จะส่งผลให้ ε' มีค่ามากหรือน้อยนั้น ขึ้นอยู่กับปัจจัยสองประการต่อไปนี้ : ประการแรก คือ m ซึ่งเป็นค่าแสดงถึงจำนวนจุดเด่นที่สามารถจับคู่ได้ ถ้า m มีค่ามาก จะส่งผลให้ ε' มีค่าต่ำ นั่นหมายความว่า มีความน่าเชื่อถือสูงในการประมวลผลการจับคู่ของจุดเด่น ประการที่สอง คือ ค่า $\bar{\varepsilon}$ ซึ่งเป็นค่าวัดความผิดพลาดกำลังสอง โดยเป็นการวัดความผิดพลาดอันเนื่องมาจากการเลื่อน/ย้ายตำแหน่ง การหมุน และการย่อ/ขยายของเส้นขอบภาพในฐานข้อมูลเมื่อเทียบกับเส้นขอบภาพของภาพวัตถุที่ทับซ้อนกัน ถ้าค่า $\bar{\varepsilon}$ มีค่าต่ำ แสดงว่า มีความผิดพลาดน้อยในการประมวลผลการจับคู่ของจุดเด่น นั่นหมายความว่า มีความน่าเชื่อถือสูงในการตัดสินใจการจดจำภาพวัตถุจากปัจจัยทั้งสองประการดังกล่าว สามารถแสดงด้วยเส้นกราฟดังภาพที่ 42 โดยกำหนดให้ค่า n ในสมการที่ (4-29) มีค่าเท่ากับ 8



เอกสารนี้ ภาพแสดงพฤติกรรมของสมการหาค่าความผิดพลาดในการจับคู่ของจุดเด่น (สมการที่ (4-29)) บนขั้นตอนการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปปัจจัยที่ต้องพิจารณาเลือกพารามิเตอร์ในการประมวลผลการจดจำภาพวัตถุ

ปัจจัยที่ต้องพิจารณาเลือกพารามิเตอร์ในการประมวลผลการจดจำภาพวัตถุ มีสิ่งที่จะต้องคำนึงถึง คือ พารามิเตอร์ที่เลือกมาใช้นั้น จะต้องสามารถตอบสนองต่อเงื่อนไขในการประมวลผลการจดจำภาพวัตถุได้ เงื่อนไขเหล่านั้น ได้แก่ ความเสถียรภาพต่อการเลื่อนตำแหน่ง การหมุน และการย่อหรือขยายภาพวัตถุ และโดยเฉพาะอย่างยิ่ง ในการเลือกใช้พารามิเตอร์ในการประมวลผลการจับคู่ของจุดเด่น โดยใช้โครงข่ายเซลล์ประสาทเทียมแบบไฮปฟิลด์นั้น สามารถที่จะนำพารามิเตอร์ต่าง ๆ มากกว่าหนึ่งค่ามาพิจารณาร่วมกันในการตัดสินใจ ด้วยสมการของไฮปฟิลด์ได้ ซึ่งในการดำเนินการศึกษาและวิจัย ได้ใช้พารามิเตอร์ทั้งหมด 4 ค่าด้วยกัน ได้แก่ ค่าสเฟียริซิตี ค่าของมุม ค่าความโค้งและทิศทางความโค้งของจุดเด่น เพื่อใช้วัดคุณสมบัติของจุดเด่นที่สามารถจับคู่กันได้ โดยมีโครงข่ายเซลล์ประสาทเทียมทำหน้าที่ในการประมวลผล ปัจจัยหนึ่งที่ต้องเลือกใช้ค่าพารามิเตอร์หลายค่า ก็เพื่อที่จะได้มาซึ่งความถูกต้องและแม่นยำในการประมวลผลการจับคู่ ถ้าเลือกใช้ค่าพารามิเตอร์เพียงค่าใดค่าหนึ่งแต่เพียงลำพัง จะทำให้ประสิทธิภาพของการประมวลผลการจับคู่ไม่อาจบรรลุผลตามวัตถุประสงค์ที่ต้องการได้ ซึ่งมีตัวอย่างในการประมวลผลดังกล่าวแสดงในบทผลการทดลอง



บทที่ 5

ผลการทดลอง

อุปกรณ์ในการทดลอง

- (1) เครื่องคอมพิวเตอร์ส่วนบุคคล พร้อมติดตั้ง Image Grabber Card
- (2) กล้องถ่ายภาพ วิดีโอ หรือกล้อง CCD
- (3) ชุดส่องแสงสว่าง
- (4) เครื่องรับโทรทัศน์

ขั้นตอนการทดลอง

ภาพที่ 43



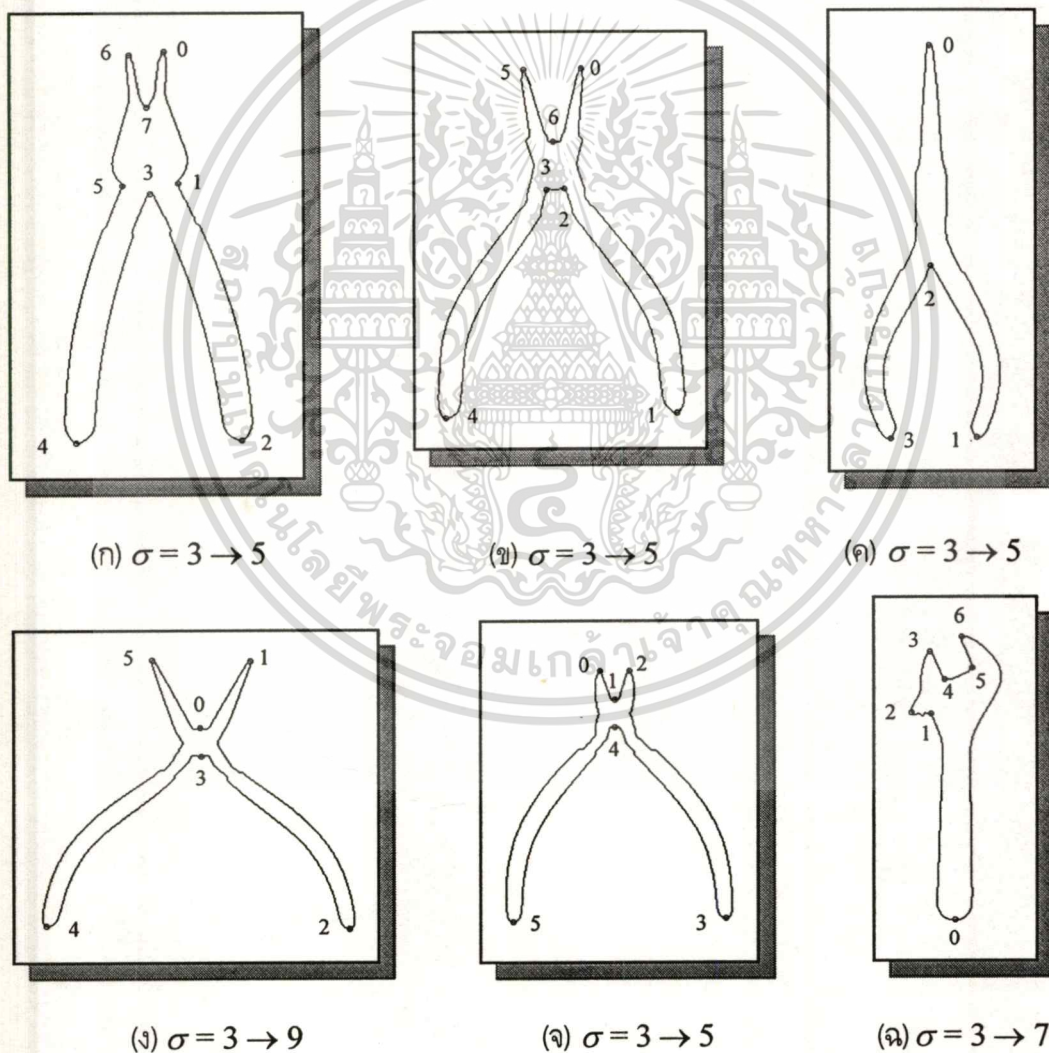
ภาพแสดงการติดตั้งอุปกรณ์ในระบบการจดจำภาพวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(1) ติดตั้งระบบดังแสดงในภาพที่ 43

(2) ทำการเก็บภาพวัตถุสำหรับสร้างฐานข้อมูล โดยทำการเก็บภาพวัตถุแต่ละชิ้นที่จะนำมาสร้างเป็นฐานข้อมูลผ่านเข้ามาทางกล้องวิดีโอ หรือ กล้อง CCD ต่อจากนั้น ดำเนินการประมวลผลหาเส้นขอบภาพของภาพวัตถุแต่ละชิ้น เมื่อได้เส้นขอบภาพของภาพวัตถุแต่ละชิ้นแล้ว ให้ดำเนินการประมวลผลหาจุดเด่นของเส้นขอบภาพเหล่านั้น โดยใช้อัลกอริทึมที่สองและอัลกอริทึมที่สาม ซึ่งได้กล่าวไว้โดยละเอียดในบทที่ 2 และได้ผลการประมวลผลการหาจุดเด่นเพื่อสร้างฐานข้อมูล ดังแสดงไว้ในภาพที่ 44 โดยมีโครงสร้างข้อมูลดังแสดงในภาพที่ 45

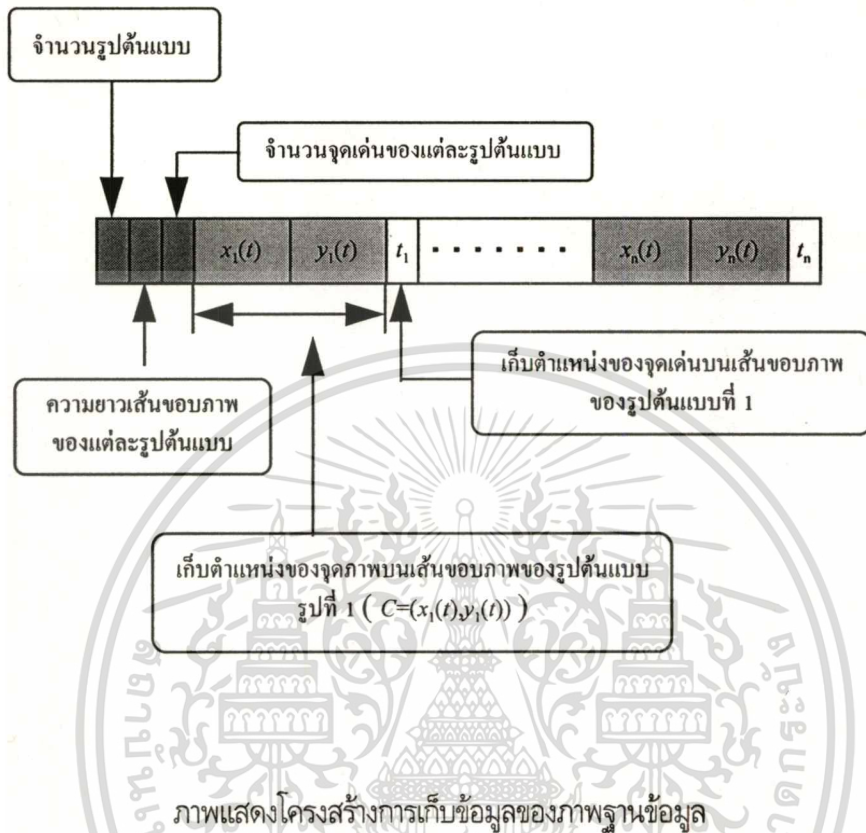
ภาพที่ 44



ภาพแสดงจุดเด่นของเส้นขอบภาพที่ถูกสร้างเป็นฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 45



(3) ทำการเก็บภาพตัวอย่างวัตถุที่มีการวางทับซ้อนกันบางส่วนหรือสัมผัสกัน เพื่อนำมาวิเคราะห์และประมวลผลการจดจำภาพวัตถุ โดยมีลำดับขั้นตอนการดำเนินการเช่นเดียวกับขั้นตอนที่ (2) ซึ่งจะได้ผลดังแสดงไว้ในภาพที่ 47, 52, 57, 62, 67, 72 และ 77

(4) ดำเนินการประมวลผลการจดจำภาพวัตถุที่มีการทับซ้อนกันบางส่วนหรือสัมผัสกัน ตามกรรมวิธีที่ได้กล่าวโดยละเอียดในบทที่ 4 โดยกำหนดค่าพารามิเตอร์ที่ใช้ในการประมวลผลด้วยโครงข่ายเซลล์ประสาทเทียมดังแสดงในตารางที่ 2 ส่วนค่า ε' ให้กำหนดไว้ที่ 200 โดยจะยอมรับภาพวัตถุที่สามารถจดจำได้จากการประมวลผล ถ้าค่า ε' มีค่าน้อยกว่า 200

ตารางที่ 2

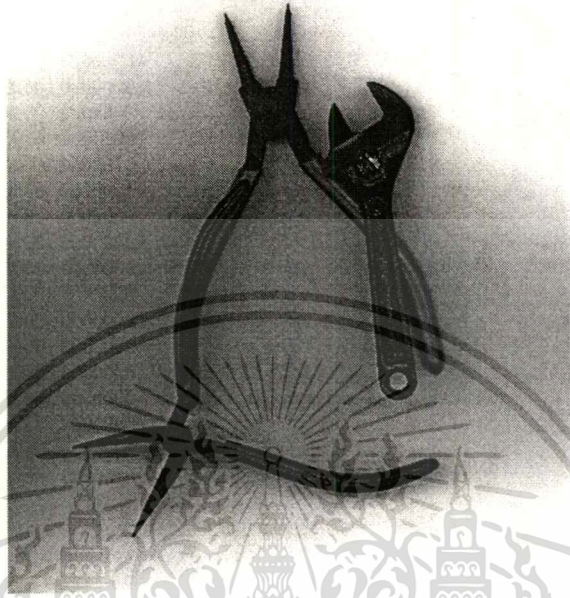
ชื่อพารามิเตอร์	ค่าพารามิเตอร์	สมการที่ใช้คำนวณ
ค่าสเปียร์ริซิติ	0.98-0.99	สมการที่ (4-4)
ค่าความโค้งของจุดเด่น	0.165	สมการที่ (4-5)
ค่ามุมของจุดเด่น	3-7 องศา	สมการที่ (4-10)
$W_1 = W_2$	0.25	สมการที่ (4-16c)
$W_3 = W_4$	0.25	สมการที่ (4-16c)

ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย
โครงข่ายเซลล์ประสาทเทียมแบบไฮปฟิลด์

ตัวอย่างผลการทดลอง

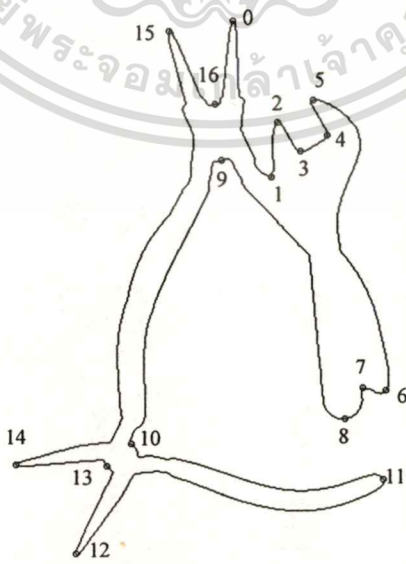
ในการทดลองครั้งนี้ได้สร้างฐานข้อมูลของภาพต้นแบบที่ผ่านการคำนวณหาจุดเด่นแล้ว ทั้งหมด 6 ตัวอย่าง ดังแสดงในภาพที่ 44 และได้ทดลองนำวัตถุทั้ง 6 ตัวอย่างมาวางให้เกิดการทับซ้อนกันหรือสัมผัสกันในลักษณะต่าง ๆ กัน ดังแสดงในภาพที่ 47, 52, 57, 62, 67, 72 และ 77 ตามลำดับ แล้วทำการประมวลผลการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น ซึ่งได้ผลดังแสดงในภาพที่ 50, 55, 60, 65, 70, 75, และ 80 ตามลำดับ

ภาพที่ 46



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 1)

ภาพที่ 47



ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 1)

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของกรมส่งเสริมการค้าระหว่างประเทศ กระทรวงพาณิชย์ โดยผู้ดูแลเว็บไซต์ระบบใบอนุญาตการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 48

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	1000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0001000	
3	00000000	0000000	0000	000000	000000	0000100	
4	00000000	0000000	0000	000000	000000	0000010	
5	00000000	0000000	0000	000000	000000	0000001	
6	00000000	0000000	0000	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000010	0000	000000	000000	0000000	
9	00000000	0000001	0000	000000	000000	0000000	
10	00000000	1000000	0000	000100	000000	0000000	
11	00000000	0000000	0000	000010	000000	0000000	
12	00000000	0000000	0000	000001	000000	0000000	
13	00000000	0000000	0000	100000	000000	0000000	
14	00000000	0000000	0000	010000	000000	0000000	
15	00000000	0000010	0000	000000	000000	0000000	
16	00000000	0000001	0000	000000	000000	0000000	

ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 1)

ภาพที่ 49

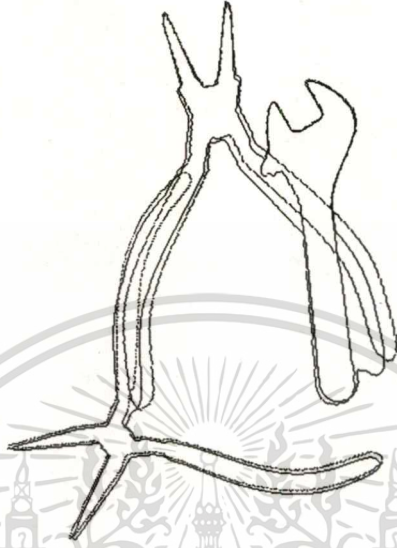
	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	1000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0001000	
3	00000000	0000000	0000	000000	000000	0000100	
4	00000000	0000000	0000	000000	000000	0000010	
5	00000000	0000000	0000	000000	000000	0000001	
6	00000000	0000000	0000	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000000	0000	000000	000000	0000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000100	000000	0000000	
11	00000000	0000000	0000	000010	000000	0000000	
12	00000000	0000000	0000	000001	000000	0000000	
13	00000000	0000000	0000	100000	000000	0000000	
14	00000000	0000000	0000	010000	000000	0000000	
15	00000000	0000010	0000	000000	000000	0000000	
16	00000000	0000001	0000	000000	000000	0000000	

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 50



ภาพแสดงผลของการจัดจำแนกออกตำแหน่งของวัตถุที่ถูกคบบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 1)

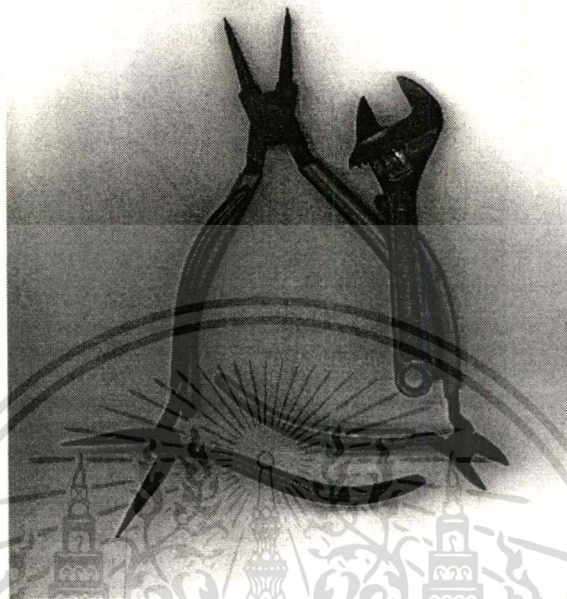
ตารางที่ 3

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	0	∞
ข	7	3	17.9886
ค	4	0	∞
ง	6	5	45.1040
จ	6	0	∞
ฉ	7	4	22.0167

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 1)

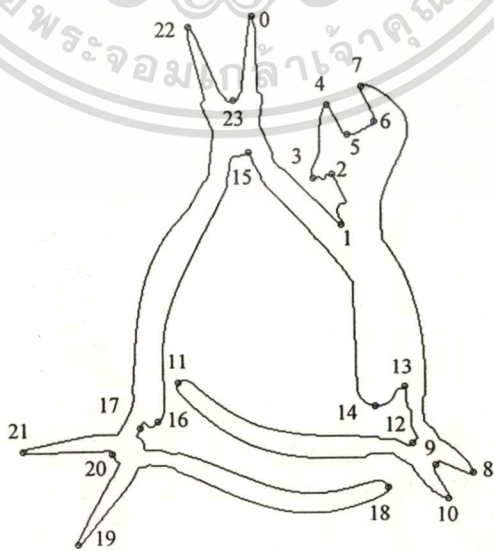
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 51



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 2)

ภาพที่ 52



เอกสารนี้เป็นเอกสารภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 2) ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 53

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	1000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0010000	
4	00000000	0000000	0000	000000	000000	0001000	
5	00000000	0000000	0000	000000	000000	0000100	
6	00000000	0000000	0000	000000	000000	0000010	
7	00000000	0000000	0000	000000	000000	0000001	
8	00000000	0000000	0000	000000	100100	0000000	
9	00000000	0000000	0000	000000	010010	0000000	
10	00000000	0000000	0000	000000	001001	0000000	
11	00000000	0000000	0000	000000	000100	0000000	
12	00000000	0000000	0000	000000	000010	0000000	
13	00000000	0000000	0000	000000	000000	0000000	
14	00000000	0000010	0000	000000	000000	0000000	
15	00000000	0000001	0000	000000	000000	0000000	
16	00000000	1000000	0000	000000	000000	0000000	
17	00000000	0000000	0000	000100	000000	0000000	
18	00000000	0000000	0000	000010	000000	0000000	
19	00000000	0000000	0000	000001	000000	0000000	
20	00000000	0000000	0000	100000	000000	0000000	
21	00000000	0000000	0000	010000	000000	0000000	
22	00000000	0000010	0000	000000	000000	0000000	
23	00000000	0000001	0000	000000	000000	0000000	

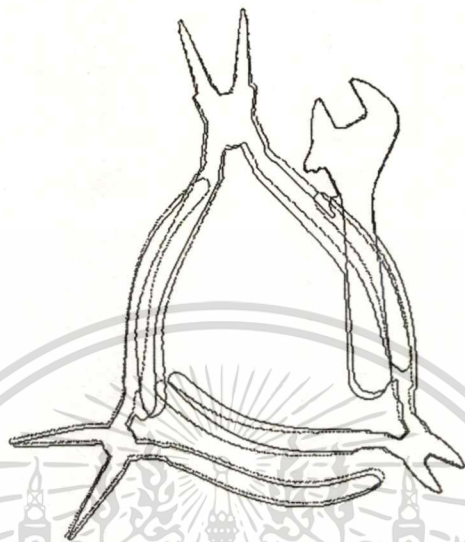
ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 2)

ภาพที่ 54

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	1000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0010000	
4	00000000	0000000	0000	000000	000000	0001000	
5	00000000	0000000	0000	000000	000000	0000100	
6	00000000	0000000	0000	000000	000000	0000010	
7	00000000	0000000	0000	000000	000000	0000001	
8	00000000	0000000	0000	000000	000000	1000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000000	001000	0000000	
11	00000000	0000000	0000	000000	000000	0000000	
12	00000000	0000000	0000	000000	000010	0000000	
13	00000000	0000000	0000	000000	000000	0000000	
14	00000000	0000000	0000	000000	000000	0000000	
15	00000000	0000000	0000	000000	000000	0000000	
16	00000000	0000000	0000	000000	000000	0000000	
17	00000000	0000000	0000	000000	000000	0000000	
18	00000000	0000000	0000	000010	000000	0000000	
19	00000000	0000000	0000	000001	000000	0000000	
20	00000000	0000000	0000	100000	000000	0000000	
21	00000000	0000000	0000	010000	000000	0000000	
22	00000000	0000010	0000	000000	000000	0000000	
23	00000000	0000001	0000	000000	000000	0000000	

ภาพแสดงผลของการลู่เข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 2)

ภาพที่ 55



ภาพแสดงผลของการตัดจำและบอกตำแหน่งของวัตุที่ถูกตัดบ้างส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 2)

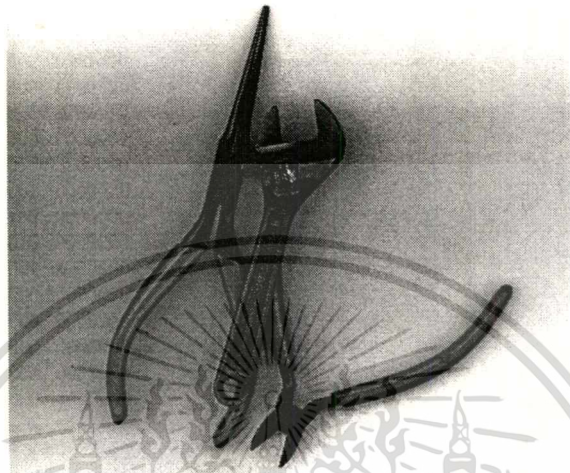
ตารางที่ 4

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	0	∞
ข	7	3	8.3930
ค	4	0	∞
ง	6	4	147.7869
จ	6	3	11.5709
ฉ	7	5	20.3604

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 2)

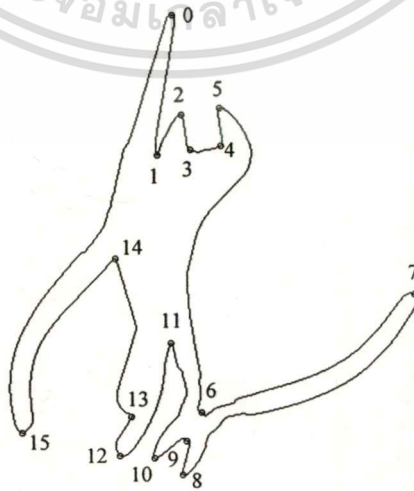
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 56



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 3)

ภาพที่ 57



ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 3)

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของกรมส่งเสริมการค้าระหว่างประเทศ กระทรวงพาณิชย์
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 58

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)
	01234567	0123456	0123	012345	012345	0123456
0	00000000	0000000	1000	000000	000000	0000000
1	00000000	0000000	0000	000000	000000	0000000
2	00000000	0000000	0000	000000	000000	0001000
3	00000000	0000000	0000	000000	000000	0000100
4	00000000	0000000	0000	000000	000000	0000010
5	00000000	0000000	0000	000000	000000	0000001
6	00000000	0000000	0000	000000	000010	0000000
7	00000000	0000000	0000	000000	000001	0000000
8	00000000	0000000	0000	000000	100100	0000000
9	00000000	0000000	0000	100000	010010	0000000
10	00000000	0000000	0000	010000	001001	0000000
11	00000000	0000000	0000	001000	000000	0000000
12	00000000	0000000	0000	000000	000000	0000000
13	00000000	0000000	0000	000000	000000	0000000
14	00000000	0000000	0010	000000	000000	0000000
15	00000000	0000000	0001	000000	000000	0000000

ภาพแสดงการกำหนดสถานะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 3)

ภาพที่ 59

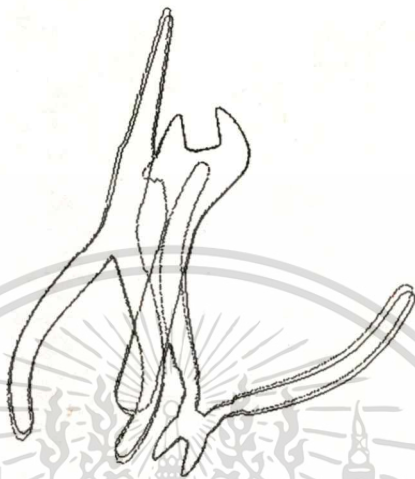
	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)
	01234567	0123456	0123	012345	012345	0123456
0	00000000	0000000	1000	000000	000000	0000000
1	00000000	0000000	0000	000000	000000	0000000
2	00000000	0000000	0000	000000	000000	0001000
3	00000000	0000000	0000	000000	000000	0000100
4	00000000	0000000	0000	000000	000000	0000010
5	00000000	0000000	0000	000000	000000	0000001
6	00000000	0000000	0000	000000	000010	0000000
7	00000000	0000000	0000	000000	000001	0000000
8	00000000	0000000	0000	000000	100000	0000000
9	00000000	0000000	0000	000000	010000	0000000
10	00000000	0000000	0000	000000	001000	0000000
11	00000000	0000000	0000	000000	000000	0000000
12	00000000	0000000	0000	000000	000000	0000000
13	00000000	0000000	0000	000000	000000	0000000
14	00000000	0000000	0010	000000	000000	0000000
15	00000000	0000000	0001	000000	000000	0000000

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 60



ภาพแสดงผลของการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 3)

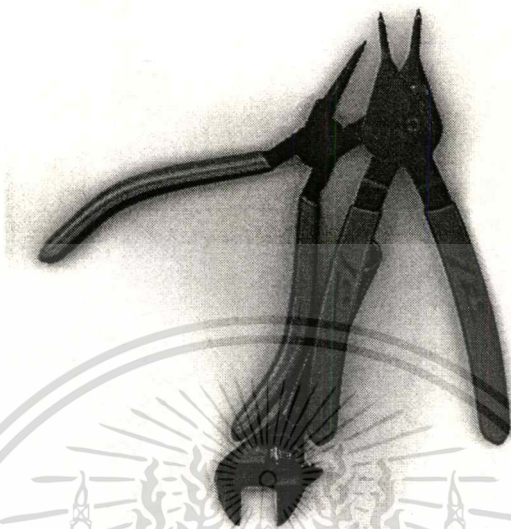
ตารางที่ 5

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	0	∞
ข	7	0	∞
ค	4	3	187.5002
ง	6	0	∞
จ	6	5	44.3560
ฉ	7	4	11.3005

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 3)

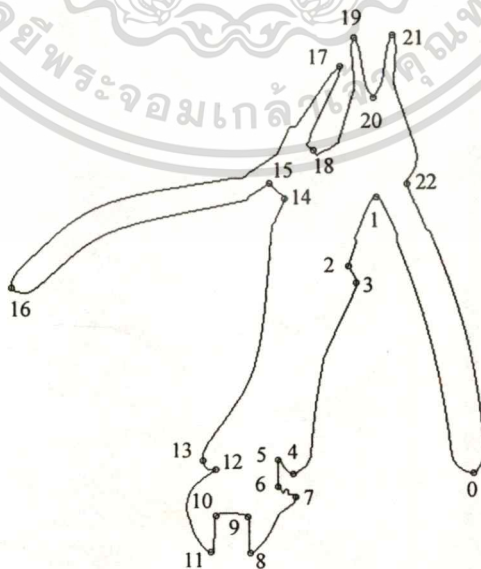
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 61



ภาพแสดงการทับซ้อนกันบางส่วนหรือล้มปลั๊กของวัตตุ (ตัวอย่างที่ 4)

ภาพที่ 62



เอกสารนี้เป็นเอกสารภาพแสดงจุดเด่นของภาพวัตตุที่ทับซ้อนกันบางส่วนหรือล้มปลั๊ก (ตัวอย่างที่ 4) ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 63

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00100000	0000000	0000	000000	000000	0000000	
1	00010000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0000000	
4	00000000	0000000	0000	000000	000000	0000000	
5	00000000	0000000	0000	000000	000000	0000000	
6	00000000	0000000	0000	000000	000000	0100000	
7	00000000	0000000	0000	000000	000000	0010000	
8	00000000	0000000	0000	000000	000000	0001000	
9	00000000	0000000	0000	000000	000000	0000100	
10	00000000	0000000	0000	000000	000000	0000010	
11	00000000	0000000	0000	000000	000000	0000001	
12	00000000	0000000	0000	000000	000000	0000000	
13	00000000	0100000	0000	000000	000000	0000000	
14	00000000	0010000	0000	000000	000000	0000000	
15	00000000	0001000	0000	000000	000000	0000000	
16	00000000	0000100	0000	000000	000000	0000000	
17	00000000	0000010	0000	000000	000000	0000000	
18	00000000	0000001	0000	000000	000000	0000000	
19	00100010	0000000	0000	000000	000000	0000000	
20	00010001	0000000	0000	000000	000000	0000000	
21	10001000	0000000	0000	000000	000000	0000000	
22	01000000	0000000	0000	000000	000000	0000000	

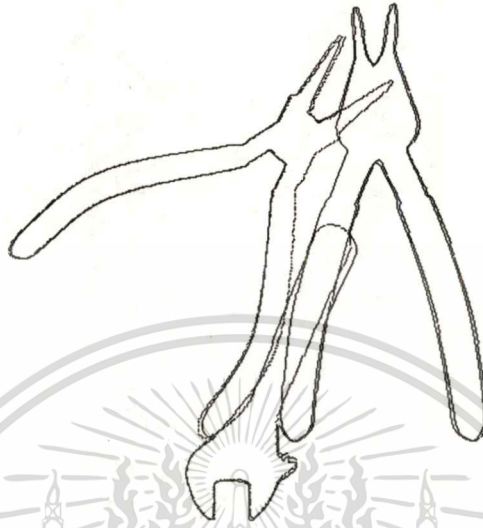
ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 4)

ภาพที่ 64

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)
	01234567	0123456	0123	012345	012345	0123456
0	00100000	0000000	0000	000000	000000	0000000
1	00010000	0000000	0000	000000	000000	0000000
2	00000000	0000000	0000	000000	000000	0000000
3	00000000	0000000	0000	000000	000000	0000000
4	00000000	0000000	0000	000000	000000	0000000
5	00000000	0000000	0000	000000	000000	0000000
6	00000000	0000000	0000	000000	000000	0100000
7	00000000	0000000	0000	000000	000000	0010000
8	00000000	0000000	0000	000000	000000	0001000
9	00000000	0000000	0000	000000	000000	0000100
10	00000000	0000000	0000	000000	000000	0000010
11	00000000	0000000	0000	000000	000000	0000001
12	00000000	0000000	0000	000000	000000	0000000
13	00000000	0000000	0000	000000	000000	0000000
14	00000000	0010000	0000	000000	000000	0000000
15	00000000	0000000	0000	000000	000000	0000000
16	00000000	0000100	0000	000000	000000	0000000
17	00000000	0000010	0000	000000	000000	0000000
18	00000000	0000001	0000	000000	000000	0000000
19	00000010	0000000	0000	000000	000000	0000000
20	00000001	0000000	0000	000000	000000	0000000
21	10000000	0000000	0000	000000	000000	0000000
22	01000000	0000000	0000	000000	000000	0000000

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 4)

ภาพที่ 65



ภาพแสดงผลของการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 4)

ตารางที่ 6

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	6	16.5756
ข	7	4	71.8551
ค	4	0	∞
ง	6	0	∞
จ	6	0	∞
ฉ	7	6	22.3658

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 4)

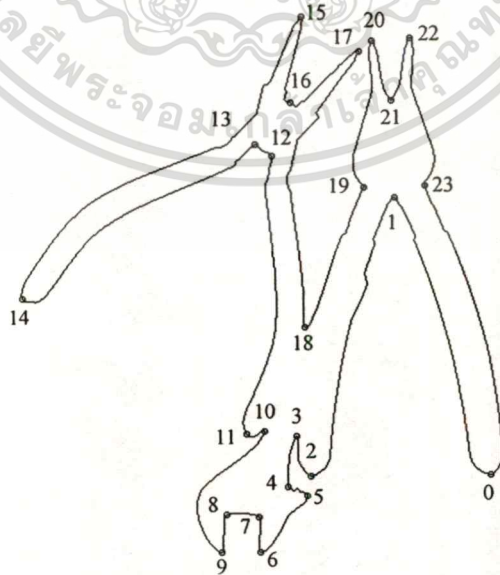
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 66



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 5)

ภาพที่ 67



ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 5)

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการแจ้งในเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้เนาเบเซประยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 68

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)
	01234567	0123456	0123	012345	012345	0123456
0	00101010	0000000	0000	000000	000000	0000000
1	00010101	0000000	0000	000000	000000	0000000
2	10001000	0000000	0000	000000	000000	0000000
3	00000000	0000000	0000	000000	000000	0000000
4	00000000	0000000	0000	000000	000000	0100000
5	00000000	0000000	0000	000000	000000	0010000
6	00000000	0000000	0000	000000	000000	0001000
7	00000000	0000000	0000	000000	000000	0000100
8	00000000	0000000	0000	000000	000000	0000010
9	00000000	0000000	0000	000000	000000	0000001
10	00000000	0000000	0000	000000	000000	0000000
11	00000000	0100000	0000	000000	000000	0000000
12	00000000	0010000	0000	000000	000000	0000000
13	00000000	0001000	0000	000000	000000	0000000
14	00000000	0000100	0000	000000	000000	0000000
15	00000000	0000010	0000	000000	000000	0000000
16	00000000	0000001	0000	000000	000000	0000000
17	00000000	1000000	0000	000000	000000	0000000
18	00000000	0000000	0000	000000	000000	0000000
19	00000100	0000000	0000	000000	000000	0000000
20	00100010	0000000	0000	000000	000000	0000000
21	00010001	0000000	0000	000000	000000	0000000
22	10001000	0000000	0000	000000	000000	0000000
23	01010000	0000000	0000	000000	000000	0000000

ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 5)

ภาพที่ 69

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	0000000	0000	000000	000000	0000000	
1	00010000	0000000	0000	000000	000000	0000000	
2	00001000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0000000	
4	00000000	0000000	0000	000000	000000	0100000	
5	00000000	0000000	0000	000000	000000	0010000	
6	00000000	0000000	0000	000000	000000	0001000	
7	00000000	0000000	0000	000000	000000	0000100	
8	00000000	0000000	0000	000000	000000	0000010	
9	00000000	0000000	0000	000000	000000	0000001	
10	00000000	0000000	0000	000000	000000	0000000	
11	00000000	0100000	0000	000000	000000	0000000	
12	00000000	0010000	0000	000000	000000	0000000	
13	00000000	0001000	0000	000000	000000	0000000	
14	00000000	0000100	0000	000000	000000	0000000	
15	00000000	0000010	0000	000000	000000	0000000	
16	00000000	0000001	0000	000000	000000	0000000	
17	00000000	1000000	0000	000000	000000	0000000	
18	00000000	0000000	0000	000000	000000	0000000	
19	00000000	0000000	0000	000000	000000	0000000	
20	00000010	0000000	0000	000000	000000	0000000	
21	00000001	0000000	0000	000000	000000	0000000	
22	10000000	0000000	0000	000000	000000	0000000	
23	00000000	0000000	0000	000000	000000	0000000	

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 5)

ภาพที่ 70



ภาพแสดงผลของการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 5)

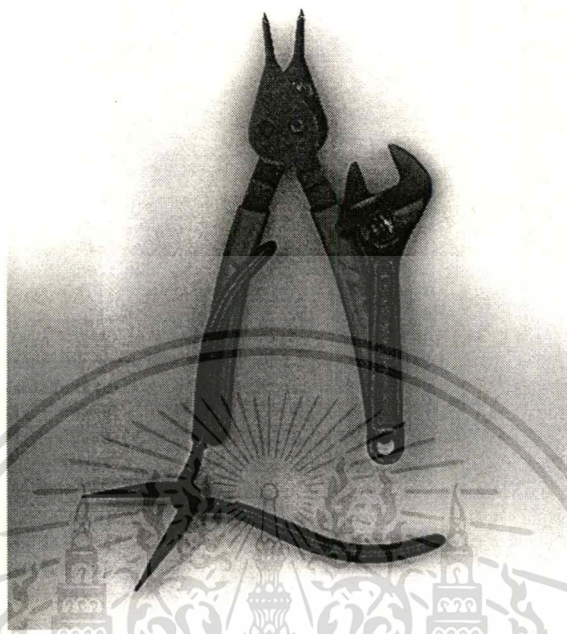
ตารางที่ 7

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	5	22.0240
ข	7	7	10.7778
ค	4	0	∞
ง	6	0	∞
จ	6	0	∞
ฉ	7	6	24.7290

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 5)

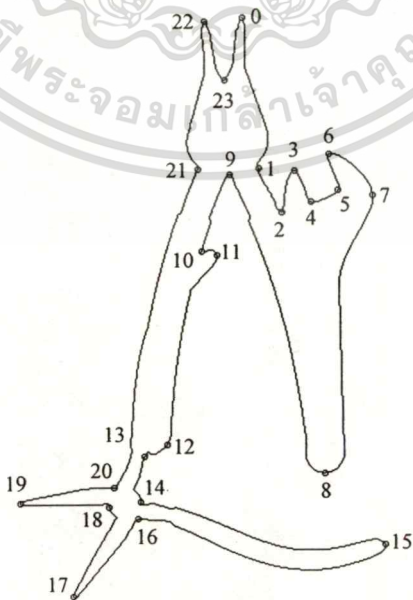
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 71



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 6)

ภาพที่ 72



ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 6)

เอกสารนี้เป็นเอกสารที่เผยแพร่หรือใช้เพื่อการเรียนการสอนเท่านั้น ไม่อนุญาตให้นำไปใช้โดยไม่ขออนุญาตจากเจ้าของลิขสิทธิ์
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 73

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	10001000	0000000	0000	000000	000000	0000000	
1	01000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0001000	
4	00000000	0000000	0000	000000	000000	0000100	
5	00000000	0000000	0000	000000	000000	0000010	
6	00000000	0000000	0000	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000000	0000	000000	000000	0000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000000	000000	0000000	
11	00000000	0000000	0000	000000	000000	0000000	
12	00000000	0000000	0000	000000	000000	0000000	
13	00000000	0000000	0000	000000	000000	0000000	
14	00000000	0000000	0000	000000	000000	0000000	
15	00000000	0000000	0000	000000	000000	0000000	
16	00000000	0000000	0000	000000	000000	0000000	
17	00000000	0000000	0000	000001	000000	0000000	
18	00000000	0000000	0000	100000	000000	0000000	
19	00000000	0000000	0000	010000	000000	0000000	
20	00000000	0000000	0000	000000	000000	0000000	
21	00000100	0000000	0000	000000	000000	0000000	
22	00100010	0000000	0000	000000	000000	0000000	
23	00010001	0000000	0000	000000	000000	0000000	

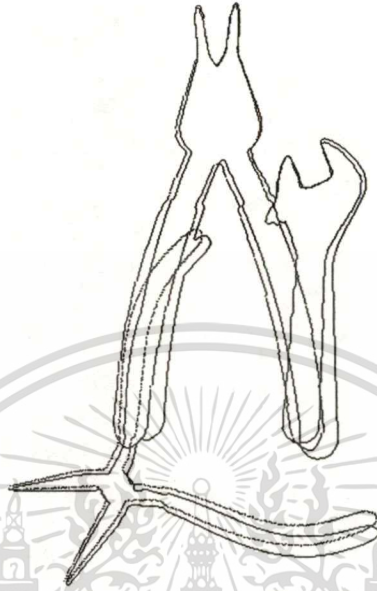
ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 6)

ภาพที่ 74

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	10000000	0000000	0000	000000	000000	0000000	
1	01000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	000000	0000000	
3	00000000	0000000	0000	000000	000000	0001000	
4	00000000	0000000	0000	000000	000000	0000100	
5	00000000	0000000	0000	000000	000000	0000010	
6	00000000	0000000	0000	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000000	0000	000000	000000	0000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000000	000000	0000000	
11	00000000	0000000	0000	000000	000000	0000000	
12	00000000	0000000	0000	000000	000000	0000000	
13	00000000	0000000	0000	000000	000000	0000000	
14	00000000	0000000	0000	000000	000000	0000000	
15	00000000	0000000	0000	000000	000000	0000000	
16	00000000	0000000	0000	000000	000000	0000000	
17	00000000	0000000	0000	000001	000000	0000000	
18	00000000	0000000	0000	100000	000000	0000000	
19	00000000	0000000	0000	010000	000000	0000000	
20	00000000	0000000	0000	000000	000000	0000000	
21	00000100	0000000	0000	000000	000000	0000000	
22	00000010	0000000	0000	000000	000000	0000000	
23	00000001	0000000	0000	000000	000000	0000000	

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 6)

ภาพที่ 75



ภาพแสดงผลของการจัดจำและบอกตำแหน่งของวัตถุที่ถูกบังคับบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 6)

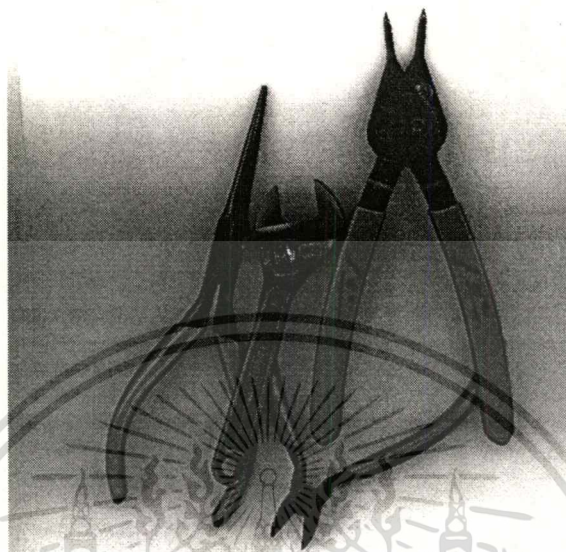
ตารางที่ 8

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ε'
ก	8	5	73.3312
ข	7	0	∞
ค	4	0	∞
ง	6	3	86.4200
จ	6	0	∞
ฉ	7	3	57.0192

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 6)

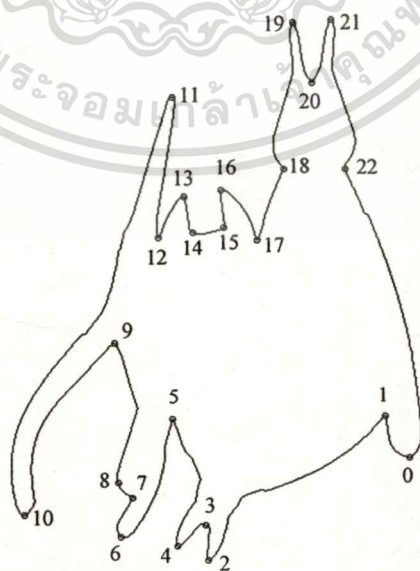
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 76



ภาพแสดงการทับซ้อนกันบางส่วนหรือสัมผัสกันของวัตถุ (ตัวอย่างที่ 7)

ภาพที่ 77



ภาพแสดงจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนหรือสัมผัสกัน (ตัวอย่างที่ 7)

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของสำนักงานทรัพย์สินส่วนพระมหากษัตริย์ โดยอยู่ภายใต้เงื่อนไขการใช้งานภายใต้เงื่อนไขการใช้งาน
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 78

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00100000	0000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	100100	0000000	
3	00000000	0000000	0000	100000	010010	0000000	
4	00000000	0000000	0000	010000	001001	0000000	
5	00000000	0000000	0000	001000	000000	0000000	
6	00000000	0000000	0000	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000000	0000	000000	000000	0000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000000	000000	0000000	
11	00000000	0000000	0000	000000	000000	0000000	
12	00000000	0000000	0000	000000	000000	0000000	
13	00000000	0000000	0000	000000	000000	0001000	
14	00000000	0000000	0000	000000	000000	0000100	
15	00000000	0000000	0000	000000	000000	0000010	
16	00000000	0000000	0000	000000	000000	0000001	
17	00000000	0000000	0000	000000	000000	0000000	
18	00000100	0000000	0000	000000	000000	0000000	
19	00100010	0000000	0000	000000	000000	0000000	
20	00010001	0000000	0000	000000	000000	0000000	
21	10001000	0000000	0000	000000	000000	0000000	
22	01000000	0000000	0000	000000	000000	0000000	

ภาพแสดงการกำหนดสภาวะเริ่มต้นในการจับคู่ของโครงข่ายของเซลล์ประสาทเทียม (ตัวอย่างที่ 7)

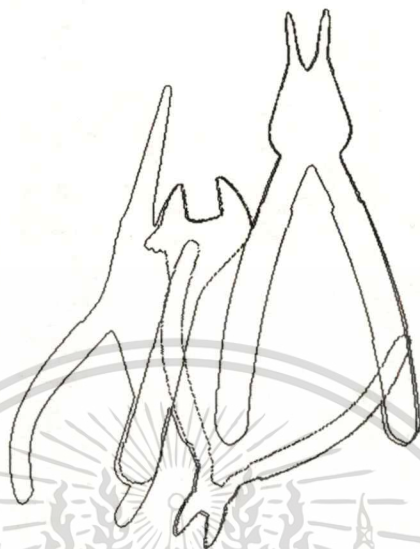
ภาพที่ 79

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)	N
	01234567	0123456	0123	012345	012345	0123456	
0	00000000	0000000	0000	000000	000000	0000000	
1	00000000	0000000	0000	000000	000000	0000000	
2	00000000	0000000	0000	000000	100000	0000000	
3	00000000	0000000	0000	000000	010000	0000000	
4	00000000	0000000	0000	000000	001000	0000000	
5	00000000	0000000	0000	000000	000000	0000000	
6	00000000	0000000	0001	000000	000000	0000000	
7	00000000	0000000	0000	000000	000000	0000000	
8	00000000	0000000	0000	000000	000000	0000000	
9	00000000	0000000	0000	000000	000000	0000000	
10	00000000	0000000	0000	000000	000000	0000000	
11	00000000	0000000	0000	000000	000000	0000000	
12	00000000	0000000	0000	000000	000000	0000000	
13	00000000	0000000	0000	000000	000000	0001000	
14	00000000	0000000	0000	000000	000000	0000100	
15	00000000	0000000	0000	000000	000000	0000010	
16	00000000	0000000	0000	000000	000000	0000001	
17	00000000	0000000	0000	000000	000000	0000000	
18	00000100	0000000	0000	000000	000000	0000000	
19	00000010	0000000	0000	000000	000000	0000000	
20	00000001	0000000	0000	000000	000000	0000000	
21	10000000	0000000	0000	000000	000000	0000000	
22	01000000	0000000	0000	000000	000000	0000000	

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 7)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 80



ภาพแสดงผลของการจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น (ตัวอย่างที่ 7)

ตารางที่ 9

รูปในฐานข้อมูล (ภาพที่ 44)	จำนวนจุดเด่นรวมของ แต่ละรูปในฐานข้อมูล	จำนวนจุดเด่นที่ สามารถจับคู่ได้	ϵ'
ก	8	5	9.9150
ข	7	0	∞
ค	4	1	∞
ง	6	0	∞
จ	6	3	36.4810
ฉ	7	4	26.4246

ตารางสรุปผลค่าความผิดพลาดในการจับคู่ของจุดเด่น (ตัวอย่างที่ 7)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างที่ 7 เป็นกรณีที่ไม่สามารถที่จะจดจำและบอกตำแหน่งของภาพวัตถุได้ทั้งหมดทุกชิ้น ซึ่งปัญหาที่เกิดขึ้นนี้ ไม่ได้เกิดจากข้อบกพร่องของการประมวลผลการจับคู่ด้วยโครงข่ายเซลล์ประสาทเทียมแต่อย่างใด แต่ได้เกิดจากการกำหนดสถานะเริ่มต้นที่ไม่เหมาะสมด้วยการกำหนดค่า *สเฟียริซิตี* ที่สูงเกินไปสำหรับการนี้ดังกล่าว ดังแสดงให้เห็นในภาพที่ 78 ทำให้เซลล์ประสาทที่ควรจะถูกจับคู่กันได้ถูกกำจัดทิ้งไปเสียก่อน ก่อนที่จะถูกป้อนเข้าสู่โครงข่ายเซลล์ประสาทเทียม จึงทำให้ได้ผลการทดลองดังแสดงในภาพที่ 80

ปัญหาที่เกิดขึ้นเมื่อเลือกใช้พารามิเตอร์ไม่เหมาะสม

การเลือกใช้พารามิเตอร์ก็เป็นปัจจัยสำคัญในการประมวลผลการจับคู่ของจุดเด่น ถ้าหากมีการเลือกใช้ค่าพารามิเตอร์อย่างไม่เหมาะสม เช่น การใช้จำนวนพารามิเตอร์ที่น้อยเกินไปในขั้นตอนของการประมวลผลการจับคู่ด้วยโครงข่ายเซลล์ประสาทแบบไฮบริด จะส่งผลทำให้ประสิทธิภาพในการประมวลผลลดลง ซึ่งได้ทดลองกับภาพที่ 58 โดยการกำหนดค่าพารามิเตอร์ดังแสดงในตารางที่ 10 จะได้ผลดังแสดงในภาพที่ 81 นั่นคือ ไม่สามารถกำจัดเซลล์ที่จับคู่ไม่เหมาะสมออกไปได้ และยังทำให้มีข้อมูลบางส่วนสูญหายไปด้วย ในทำนองเดียวกัน ถ้าใช้ค่าพารามิเตอร์ในตารางที่ 11 ก็จะได้ผลดังแสดงในภาพที่ 82 ซึ่งจะเกิดเหตุการณ์คล้ายกับภาพที่ 81 คือ ไม่สามารถกำจัดเซลล์ที่จับคู่ไม่เหมาะสมออกไปได้ แต่จะโชคดีกว่าที่ข้อมูลไม่สูญหาย

ตารางที่ 10

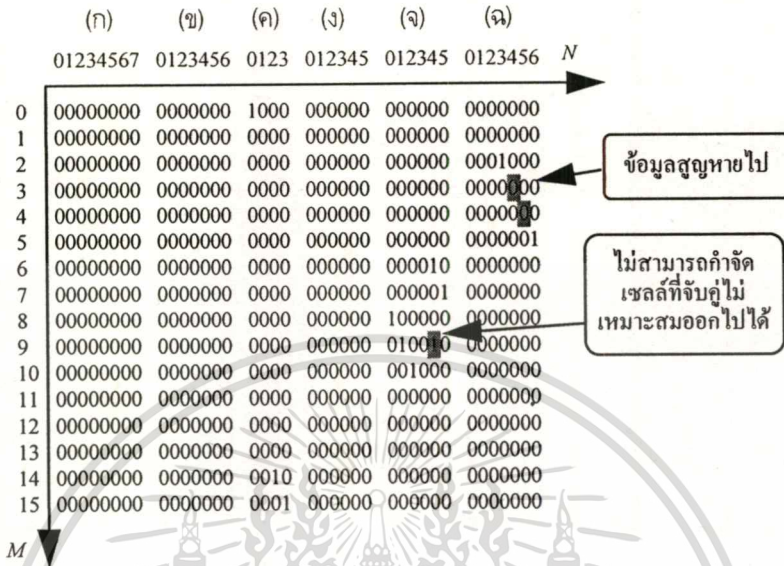
ชื่อพารามิเตอร์	ค่าพารามิเตอร์	สมการที่ใช้คำนวณ
ค่าสเฟียริซิตี	0.98-0.99	สมการที่ (4-4)
ค่าความโค้งของจุดเด่น	0.165	สมการที่ (4-5)
ค่ามุมของจุดเด่น	3-7 องศา	สมการที่ (4-10)
$W_1 = W_2$	0.5	สมการที่ (4-16c)
$W_3 = W_4$	0.0	สมการที่ (4-16c)

ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย

โครงข่ายเซลล์ประสาทเทียมแบบไฮบริด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 81



ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3) (การประมวลผลการจับคู่โดยใช้ ค่าความโค้งและทิศทางความโค้งของจุดเด่น แต่เพียงอย่างเดียว)

ตารางที่ 11

ชื่อพารามิเตอร์	ค่าพารามิเตอร์	สมการที่ใช้คำนวณ
ค่าสเฟียริซิตี	0.98-0.99	สมการที่ (4-4)
ค่าความโค้งของจุดเด่น	0.165	สมการที่ (4-5)
ค่ามุมของจุดเด่น	3-7 องศา	สมการที่ (4-10)
$W_1 = W_2$	0.0	สมการที่ (4-16c)
$W_3 = W_4$	0.5	สมการที่ (4-16c)

ตารางแสดงการกำหนดค่าพารามิเตอร์ที่ใช้ในการทดลองในขั้นตอนการประมวลผลด้วย

โครงข่ายเซลล์ประสาทเทียมแบบไฮปิลด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 82

	(ก)	(ข)	(ค)	(ง)	(จ)	(ฉ)
	01234567	0123456	0123	012345	012345	0123456
0	00000000	0000000	1000	000000	000000	0000000
1	00000000	0000000	0000	000000	000000	0000000
2	00000000	0000000	0000	000000	000000	0001000
3	00000000	0000000	0000	000000	000000	0000100
4	00000000	0000000	0000	000000	000000	0000010
5	00000000	0000000	0000	000000	000000	0000001
6	00000000	0000000	0000	000000	000010	0000000
7	00000000	0000000	0000	000000	000001	0000000
8	00000000	0000000	0000	000000	100000	0000000
9	00000000	0000000	0000	000000	010000	0000000
10	00000000	0000000	0000	000000	001000	0000000
11	00000000	0000000	0000	000000	000000	0000000
12	00000000	0000000	0000	000000	000000	0000000
13	00000000	0000000	0000	000000	000000	0000000
14	00000000	0000000	0010	000000	000000	0000000
15	00000000	0000000	0001	000000	000000	0000000

ไม่สามารถกำจัดเซลล์ที่จับคู่ไม่เหมาะสมออกไปได้

ภาพแสดงผลของการเข้าสู่สภาวะพลังงานต่ำสุดของโครงข่ายเซลล์ประสาทเทียม (ตัวอย่างที่ 3) (การประมวลผลการจับคู่โดยใช้ มุมของจุดเด่น แต่เพียงอย่างเดียว)

สรุปผลการทดลอง

จากตัวอย่างผลการทดลองที่ผ่านมา จะเห็นได้ว่า การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น สามารถประมวลผลได้อย่างมีประสิทธิภาพ โดยมี ϵ' เป็นค่าวัดความผิดพลาดอันเนื่องมาจากการประมวลผลการจับคู่ของจุดเด่นด้วยโครงข่ายเซลล์ประสาทเทียมแบบไฮบิลด์ และใช้ค่า ϵ' เป็นค่าสำหรับการตัดสินใจขั้นสุดท้ายว่าเป็นวัตถุที่คาดหมายไว้หรือไม่ ถ้าค่า ϵ' มีค่าต่ำ แสดงว่ามีความน่าเชื่อถือสูงต่อสิ่งที่คาดหมายไว้ด้วยการประมวลผลของโครงข่ายเซลล์ประสาทเทียมแบบไฮบิลด์ ซึ่งในการทดลองครั้งนี้ได้กำหนดค่า ϵ' ไว้ที่ 200 ถ้าค่า ϵ' มีค่ามากกว่า 200 จะปฏิเสธภาพในฐานะข้อมูลที่ประมวลผลได้

บทที่ 6

บทสรุป

สรุปผลการศึกษาและวิจัย

จากผลที่ได้ดำเนินการศึกษาและวิจัย เรื่อง "การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วน โดยการใช้จุดเด่น" โดยมีจุดมุ่งหมายเพื่อการศึกษาและวิจัยในสองประเด็นหลักของกระบวนการประมวลผลการจดจำภาพวัตถุ คือ ประเด็นที่หนึ่ง ต้องการแสวงหาวิธีการคำนวณหาจุดเด่นที่มีเสถียรภาพ และประเด็นที่สอง ต้องการหาวิธีที่มีประสิทธิภาพในการจับคู่ของจุดเด่นในฐานข้อมูลกับจุดเด่นของภาพวัตถุที่มีการวางทับซ้อนกันบางส่วน จากสาระสำคัญทั้งสองประเด็นเป็นสาเหตุให้ต้องดำเนินการศึกษาและวิจัย โดยประเด็นแรกนั้น ทำให้สามารถหาวิธีการที่จะคำนวณหาจุดเด่นที่มีเสถียรภาพ ด้วยเทคนิคการผสมผสานระหว่าง เทคนิคการคำนวณหาจุดเด่นด้วยตัวกรองเกาส์เซียนและคิวบิก ปี-สไฟร์น ร่วมกับ เทคนิคการตรึงจุดเด่นให้คงที่โดยใช้ซอฟต์แวร์ จากการใช้เทคนิคดังกล่าว ทำให้จุดเด่นที่คำนวณได้มีเสถียรภาพต่อการเลื่อนตำแหน่ง การหมุน และการย่อ/ขยายของภาพวัตถุ ซึ่งเป็นเงื่อนไขที่สำคัญของกระบวนการประมวลผลการจดจำภาพวัตถุ และโดยเฉพาะอย่างยิ่ง จุดเด่นที่คำนวณได้จากวิธีดังกล่าวยังเสถียรภาพต่อการเปลี่ยนค่า สเกล-สเปซ ของตัวกรองเกาส์เซียนอีกด้วย ดังแสดงให้เห็นในบทที่ 2 และ 5 ส่วนประเด็นที่สองของการศึกษาและวิจัยได้ประยุกต์ใช้โครงข่ายเซลล์ประสาทเทียมในการประมวลผลการจับคู่ของจุดเด่นที่มีประสิทธิภาพ โดยใช้คุณสมบัติของจุดเด่นที่คำนวณได้ด้วยเทคนิคผสมผสานที่กล่าวมาแล้วในขั้นต้น และด้วยคุณสมบัติของโครงข่ายเซลล์ประสาทเทียมที่สามารถวิเคราะห์และตัดสินใจในการประมวลผลการจับคู่ที่ดีที่สุดของจุดเด่นของภาพวัตถุในฐานข้อมูลที่สมบูรณ์กับภาพวัตถุที่ถูกบดบังบางส่วน ทำให้ผลของการประมวลผล การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น สามารถประมวลผลได้อย่างมีประสิทธิภาพ ดังแสดงให้เห็นในบทที่ 4 และบทผลการทดลอง (บทที่ 5)

สรุปปัญหาที่เกิดขึ้นในการศึกษาและวิจัย

ถึงแม้ว่าผลของการศึกษาและวิจัยเรื่อง "การจดจำและบอกตำแหน่งของวัตถุที่ถูกบดบังบางส่วนโดยการใช้จุดเด่น" จะสัมฤทธิ์ผลดังสมมุติฐานที่ได้ตั้งเอาไว้ แต่ก็ยังมีบางกรณีที่ไม่สามารถแก้ไขปัญหาการทับซ้อนกันของวัตถุด้วยวิธีการดังที่ได้กล่าวมาแล้วได้ เช่น การทับซ้อนกันของวัตถุที่เหมือนกันทุกประการ ดังแสดงในภาพที่ 83 ซึ่งกรณีดังกล่าวจะเกิดข้อขัดแย้งกับเงื่อนไขของการประมวลผลการจับคู่ ซึ่งได้กล่าวไว้ว่า

(1) จุดเด่นของภาพวัตถุที่ถูกทับซ้อนกันบางส่วน จะต้องจับคู่กับจุดเด่นในฐานข้อมูลได้เพียงจุดเดียวเท่านั้น

(2) จุดเด่นในฐานข้อมูล ก็จะต้องจับคู่กับจุดเด่นของภาพวัตถุที่ทับซ้อนกันบางส่วนได้เพียงจุดเดียวเช่นกัน

ดังนั้น ในส่วนของการประมวลผลการจับคู่ในกรณีดังกล่าว จึงควรหลีกเลี่ยงที่จะทำการประมวลผลโดยตรงจากวิธีที่ได้ดำเนินการศึกษาและวิจัยมา และควรได้รับการพัฒนาต่อไป



ภาพแสดงการทับซ้อนกันของวัตถุที่เหมือนกันทุกประการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- พารา ลิมมะณีประเสริฐ และ เกษตร์ ศิริสันติสัมฤทธิ์, "การตรวจหาจุดเด่นบนเส้นโค้งของขอบภาพโดยการเปลี่ยนค่าสเกล-สเปซของตัวกลอง," การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 17 1-2 ธันวาคม 2537 : 289-295
- พารา ลิมมะณีประเสริฐ และ เกษตร์ ศิริสันติสัมฤทธิ์, "โครงข่ายเซลล์ประสาทเทียมแบบไฮโปฟิลด์สำหรับการจดจำภาพวัตถุที่มีการทับซ้อนกันบางส่วน," วิศวกรรมลาดกระบัง ปีที่ 13, ฉบับที่ 1 (กรกฎาคม 2539) : 119-130
- Daniel M. Wuescher and Kim L. Boyer, "Robust Contour Decomposition Using a Constant Curvature Criterion," IEEE Trans. Pattern Anal. Machine Intell., Vol, 13, No. 1, January 1991.
- Farzin Mokhtarian and Alan Mackworth, "Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-8, No. 1, Jan 1986 : pp. 34-43
- George N. Bebis and George M. Papadourakis, "Object Recognition Using Invariant Object Boundary Representation and Neural Network Models," Pattern Recognition, Vol. 25, No. 1, 1992 : pp. 25-44
- Jian-Shing Wu and Jin-Jang Leou, "New Polygonal Approximation Schemes for Object Shape Representation," Pattern Recognition, Vol. 26, No. 4, 1993 : pp. 471-484
- Mark W. Koch and Rangasami L. Kashyap, "Using Polygons to Recognition and Locate Partially Occluded Objects," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-9, No. 4, July 1987 : pp. 483-494
- Nasser M. Nasrabadi and Chang Y. Choo, "Hopfield Network for Stereo Vision Correspondence," IEEE Trans. Neural Networks, Vol. 3 No. 1, Jan 1992 : pp. 5-13
- Nasser M. Nasrabadi and Wei Li, "Object Recognition by a Hopfield Neural Network," IEEE Trans. Syst., Man, Cybern., Vol. 21 No. 6, Nov/Dec 1991 : pp. 1523-1535

- Nicholas Ayache and Olivier D. Faugeras, "Hyper : A New Approach for the Recognition and Positioning of Two-Dimensional Objects," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-8, No. 1, Jan 1986 : pp. 44-56
- Nikhil R. Pal, Pratik Pal, and Anupam K. Basu, "A New Shape Representation Scheme and Its Application to Shape Discrimination Using a Neural Network," Pattern Recognition, Vol. 26, No. 4, 1993 : pp. 543-551
- Nirwan Ansari and Edward J. Delp, "Partial Shape Recognition : A Landmark-Based Approach," IEEE Trans. Pattern Anal. Machine Intell., Vol. 12, No.5, May 1990 : pp. 470-483
- NIRWAN ANSARI and EDWARD J. DELP, "On Detecting dominant points," Pattern Recognition, Vol. 24 No. 5, 1991 : pp. 441-451
- Nirwan Ansari and Kuowei Li, "Landmark-Based Shape Recognition by a Modified Hopfield Neural Network," Pattern Recognition, Vol. 26, No. 4, 1993 : pp. 531-542
- Peter W. Protzel and Daniel L. Palumbo, "Performance and Fault-Tolerance of Neural Networks for Optimization," IEEE Trans. Neural Networks, Vol. 4 No. 4, July 1993 : pp. 600-614
- Soo-Chang Pei and Chao-Nan Lin, "The Detection of Dominant Points on Digital Curves by Scale-Space Filtering," Pattern Recognition, Vol. 25, No. 11, 1992 : pp. 1307-1314
- Wei Wen and Andrei Lozzi, "Recognition and Inspection of Two-Dimensional Industrial Parts Using Subpolygons," Pattern Recognition, Vol. 25, No. 12, 1992 : pp. 1427-1434
- S. Y. Kung. Digital Neural Networks. USA : Prentice Hall International Editions, 1993.
- David F. Rogers and J. Alan Adams. Mathematical Elements for Computer Graphics. Singapore : McGraw-Hill, 1989.
- Milan Sonka, Vaclav Hlavac, and Roger Boyle. Image Processing, Analysis and Machine Vision. Cambridge : Chapman & Hall Computing, 1994.



ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประมวลผลหาเส้นขอบภาพ

ในการประมวลผลหาเส้นขอบภาพ จะเริ่มต้นจากการปรับปรุงเส้นขอบภาพที่ได้จากกล้องวีดีโอให้มีความคมชัดมากขึ้น โดยใช้เทคนิคของการปรับปรุงเส้นขอบภาพ (Edge Enhancement) ซึ่งมีหน้ากาก (Mask) ขนาด 3×3 และมีสัมประสิทธิ์ภายในหน้ากาดังที่แสดงในภาพที่ 84 นำไปประสานกับภาพที่ต้องการจะปรับปรุงเส้นขอบภาพ (ภาพที่ 86) ซึ่งจะได้ผลดังแสดงในภาพที่ 87

เมื่อได้เส้นขอบภาพที่คมชัดแล้ว จึงใช้เทคนิค Thresholding มาทำการแยกข้อมูลออกจากส่วนที่เป็น Background ซึ่งในการประมวลผลของรูปที่ 3 จะเลือกใช้ค่า Threshold ประมาณ 15 ซึ่งจะได้ผลดังแสดงในภาพที่ 88

ผลที่ได้จากการทำ Thresholding จะถูกนำมาผ่านกระบวนการตรวจหาเส้นขอบภาพ (Edge Detection) โดยใช้หน้ากากที่แสดงในภาพที่ 85 ซึ่งจะได้ผลดังแสดงในภาพที่ 89

ภาพที่ 84

-1	-1	-1
-1	9	-1
-1	-1	-1

ภาพแสดงหน้ากากสำหรับปรับปรุงเส้นขอบภาพ

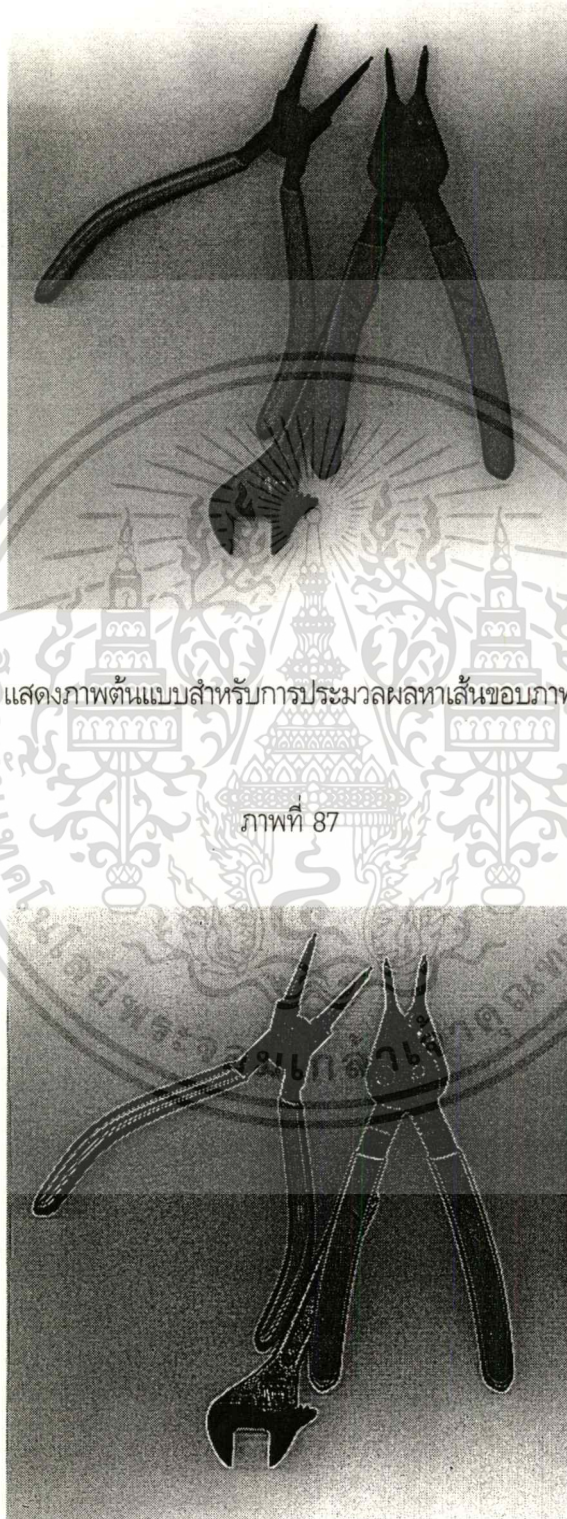
ภาพที่ 85

-1	-1	-1
-1	8	-1
-1	-1	-1

ภาพแสดงหน้ากากสำหรับตรวจหาเส้นขอบภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 86



แสดงภาพต้นแบบสำหรับการประมวลผลหาเส้นขอบภาพ

ภาพที่ 87

ภาพแสดงผลที่ได้จากการปรับปรุงเส้นขอบภาพให้คมชัดขึ้นด้วยหน้ากากในภาพที่ 84

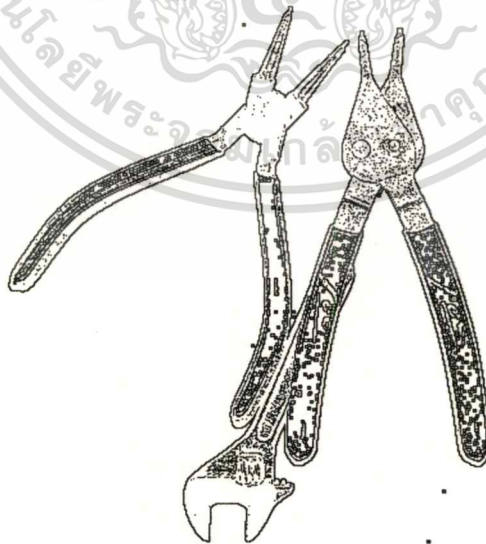
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 88



ภาพแสดงการแยกส่วนของข้อมูลออกจากส่วนที่เป็น Background ด้วยเทคนิค Thresholding

ภาพที่ 89



ภาพแสดงผลของการตรวจหาเส้นขอบภาพด้วยหน้ากากในภาพที่ 85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อได้เส้นขอบภาพดังแสดงในภาพที่ 89 แล้ว จึงทำการประมวลผลหาเส้นรอบขอบภาพ (Contour) ที่ยาวที่สุด ซึ่งจะได้ผลดังแสดงไว้ในรูปที่ 7

ภาพที่ 90



ภาพแสดงเส้นทางเดินของเส้นขอบภาพ

ในการประมวลผลหาเส้นขอบภาพนั้น คุณภาพของเส้นขอบภาพที่ประมวลผลได้ จะมีความคมชัดหรือไม่ อยู่ที่ปัจจัยดังต่อไปนี้ ปัจจัยแรก คือ จะต้องให้แสงสว่างกับวัตถุที่จะทำการเก็บภาพวัตถุนั้นอย่างพอเหมาะ โดยจะต้องให้แสงหลาย ๆ ทิศทาง เพื่อลบเงาของวัตถุที่จะเกิดขึ้น ปัจจัยที่สอง คือ คุณภาพของกล้องวิดีโอ กล้องที่มีคุณภาพสูงก็จะให้ภาพที่มีความคมชัดดี และปัจจัยสุดท้าย คือ ความละเอียดของ Image Grabber Card ถ้ามีความละเอียดสูงก็จะให้คุณภาพของภาพที่ได้มีความคมชัดดี



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก ข.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เงื่อนไขสำหรับการประมวลผลการจดจำภาพวัตถุ

ในการประมวลผลการจดจำภาพวัตถุ มีสิ่งที่ต้องพิจารณาเป็นสำคัญ คือ คณิตศาสตร์ที่นำมาประยุกต์ใช้งานการประมวลผล จะต้องสามารถตอบสนองต่อเงื่อนไขสำคัญ 3 ประการของการจดจำภาพวัตถุ คือ

1. จะต้องเสถียรภาพต่อการเลื่อนหรือย้ายตำแหน่งของภาพวัตถุ
2. จะต้องเสถียรภาพต่อการหมุนของภาพวัตถุ
3. จะต้องเสถียรภาพต่อการย่อหรือขยายของภาพวัตถุ

บทพิสูจน์ความมีเสถียรภาพของ สเฟียริซิตี ต่อการเลื่อน การหมุน และการย่อ/ขยาย¹

ดังที่ได้กล่าวมาแล้วในบทที่ 4 หัวข้อ 2.1.1 สำหรับ สเฟียริซิตี ของการแปลงรูปสามเหลี่ยม (Sphericity of a Triangular Transformation) ส่วนในภาคผนวกนี้ จะกล่าวถึงนิยามทั่วไปของ สเฟียริซิตี จนถึง Diffeomorphism² ตลอดจนคุณสมบัติที่ต้องการ เพื่อพิสูจน์ความเสถียรภาพของ สเฟียริซิตี ต่อการเลื่อน การหมุน และการย่อ/ขยาย

นิยามที่ 1 : สเฟียริซิตี ของ Diffeomorphism, $g: \Omega \rightarrow \bar{\Omega}$ ($\Omega, \bar{\Omega} \subset R^n$) สำหรับ $x \in \Omega$ ถูกกำหนดไว้ดังนี้

$$Y_g(x) = \frac{(\det(g''(x)g'(x)))^{\frac{1}{n}}}{\left(\frac{1}{n} \text{tr}(g''(x)g'(x))\right)}$$

เมื่อ $g'(x)$ คือ อนุพันธ์ของ $g(x)$

$g''(x)$ คือ ทรานส์โพสของ $g'(x)$

$\det(\)$ และ $\text{tr}(\)$ คือ ดีเทอร์มิแนนท์ และ เทรซ (trace) ของเมตริกซ์

¹ Nirwan Ansar and Edward J. Delp, "Partial Shape Recognition : A Landmark-Based Approach," IEEE Trans. Pattern Anal. Machine Intell., Vol. 12, No. 5, (May 1990) : pp. 473-474.

² Diffeomorphism คือ ความต่อเนื่องหนึ่งต่อหนึ่งของการ Mapping และการ Mapping กลับก็ยังคงต่อเนื่องด้วย นอกจากนี้ การ Mapping ทั้งสองแบบที่กล่าวมา ก็ยังคงมีความต่อเนื่องสำหรับอนุพันธ์ย่อยของมันอีกด้วย โดยใช้ประโยชน์ด้านการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เงื่อนไขสำหรับการประมวลผลการจดจำภาพวัตถุ

ในการประมวลผลการจดจำภาพวัตถุ มีสิ่งที่ต้องพิจารณาเป็นสำคัญ คือ คณิตศาสตร์ที่นำมาประยุกต์ใช้งานการประมวลผล จะต้องสามารถตอบสนองต่อเงื่อนไขสำคัญ 3 ประการของการจดจำภาพวัตถุ คือ

1. จะต้องเสถียรภาพต่อการเลื่อนหรือย้ายตำแหน่งของภาพวัตถุ
2. จะต้องเสถียรภาพต่อการหมุนของภาพวัตถุ
3. จะต้องเสถียรภาพต่อการย่อหรือขยายของภาพวัตถุ

บทพิสูจน์ความมีเสถียรภาพของ สเฟียร์ซิตี ต่อการเลื่อน การหมุน และการย่อ/ขยาย¹

ดังที่ได้กล่าวมาแล้วในบทที่ 4 หัวข้อ 2.1.1 สำหรับ สเฟียร์ซิตี ของการแปลงรูปสามเหลี่ยม (Sphericity of a Triangular Transformation) ส่วนในภาคผนวกนี้ จะกล่าวถึงนิยามทั่วไปของ สเฟียร์ซิตี จนถึง Diffeomorphism² ตลอดจนคุณสมบัติที่ต้องการ เพื่อพิสูจน์ความเสถียรภาพของ สเฟียร์ซิตี ต่อการเลื่อน การหมุน และการย่อ/ขยาย

นิยามที่ 1 : สเฟียร์ซิตี ของ Diffeomorphism, $g: \Omega \rightarrow \bar{\Omega}$ ($\Omega, \bar{\Omega} \subset R^n$) สำหรับ $x \in \Omega$ ถูกกำหนดไว้ดังนี้

$$Y_g(x) = \frac{(\det(g''(x)g'(x)))^{\frac{1}{n}}}{\left(\frac{1}{n} \text{tr}(g''(x)g'(x))\right)}$$

เมื่อ $g'(x)$ คือ อนุพันธ์ของ $g(x)$

$g''(x)$ คือ ทรานส์โพสของ $g'(x)$

$\det(\)$ และ $\text{tr}(\)$ คือ ดีเทอร์มิแนนท์ และ เทรซ (trace) ของเมตริกซ์

¹ Nirwan Ansar and Edward J. Delp, "Partial Shape Recognition : A Landmark-Based Approach," IEEE Trans. Pattern Anal. Machine Intell., Vol. 12, No. 5, (May 1990) : pp. 473-474.

² Diffeomorphism คือ ความต่อเนื่องหนึ่งต่อหนึ่งของการ Mapping และการ Mapping กลับก็ยังคงต่อเนื่องด้วย นอกเหนือจากนี้ การ Mapping ทั้งสองแบบที่กล่าวมา ก็ยังคงมีความต่อเนื่องสำหรับอนุพันธ์ย่อยของมันอีกด้วย โดยใช้ประโยชน์ด้านการคำนวณ ไม่ว่าจะเป็นกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้สะดวกต่อการคำนวณ จะทำการลดรูป $Y_g(x)$, $g'(x)$ และ $g''(x)$ ให้คงเหลือ Y_g , g' และ g'' ตามลำดับ g' เป็นการ Maps วงกลมรัศมีหนึ่งหน่วยใน Ω ไปเป็นวงรีใน $\bar{\Omega}$ โดยที่ $g''g'$ เป็นค่าบวก

ทฤษฎีบทแทรกที่ 1 : สเปกตรัมค่า Eigenvalue ของ $g''g'$:

$$Y_g = \frac{(\lambda_1 \lambda_2 \lambda_3 \dots \lambda_n)^{\frac{1}{n}}}{\left(\frac{1}{n} (\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_n) \right)}$$

เมื่อ $\lambda_i, i = 1, 2, 3, \dots, n$ คือ Eigenvalue ของ $g''g'$

พิสูจน์ :

$$\begin{aligned} \det(g''g') &= \det(\Phi' D \Phi) \\ &= \det(D \Phi' \Phi) \\ &= \det(D) = \lambda_1 \lambda_2 \lambda_3 \dots \lambda_n \end{aligned}$$

และ

$$\begin{aligned} \text{tr}(g''g') &= \text{tr}(\Phi' D \Phi) \\ &= \text{tr}(D \Phi' \Phi) \\ &= \text{tr}(D) = \lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_n \end{aligned}$$

ข้อสังเกต : ค่า Eigenvalue ของ $g''g'$ มีความสอดคล้องกับความยาวครึ่งหนึ่งของแกนวงรี เพราะว่า ค่าเฉลี่ยเรขาคณิตของลำดับเลขจำนวนจริงบวก จะมีค่าน้อยกว่าหรือเท่ากับ ค่าเฉลี่ยเลขคณิตของลำดับเลขจำนวนจริงบวกชุดเดียวกันเสมอ ดังนั้น ค่า สเปกตรัมค่า Eigenvalue จะมีความอยู่ในช่วง $[0, 1]$ ถ้า $Y_g(x) = 1$ สำหรับ $x \in \Omega$, $g(x)$ คือ conformal mapping

ทฤษฎีบทแทรกที่ 2 : สเปกตรัมค่า Eigenvalue ของ Diffeomorphism ที่ได้กล่าวไว้ในนิยามที่ 1 ยังคงมีเสถียรภาพ เมื่อ Diffeomorphism อยู่ภายใต้การแปลง $G = \{ \text{Translation, Rotation, Scaling} \}$

พิสูจน์ : ให้ g เป็น Diffeomorphism และกำหนดให้ g ถูกแปลงโดย G และได้ผลลัพธ์ \bar{g} นั่นคือ

$$\bar{g} = \alpha H g + h$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เมื่อ α คือ องค์ประกอบตัวคูณ (Scaling factor)
 H คือ เมทริกซ์สำหรับการหมุน (Rotation matrix)
 h คือ เวกเตอร์สำหรับการเลื่อนตำแหน่ง (Translation vector)

$$\begin{aligned} \text{ดังนั้น } \bar{g}'\bar{g}' &= (\alpha Hg')'(\alpha Hg') \\ &= \alpha^2 g''(H' H)g' \\ &= \alpha^2 g''g' \end{aligned}$$

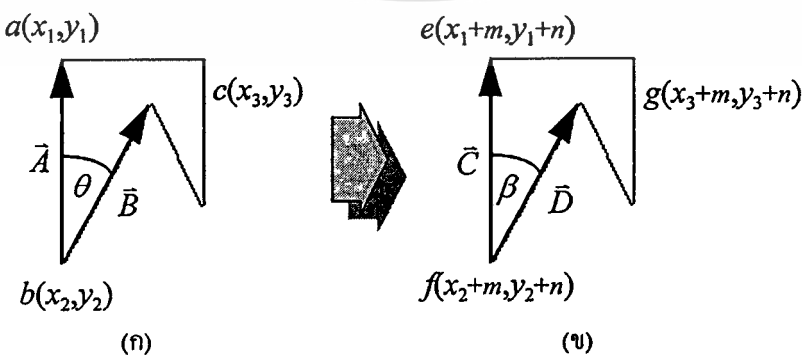
ใช้นิยามที่ 1 สรุปได้ว่า $Y_{\bar{g}} = Y_g$

จากบทพิสูจน์ของคณิตศาสตร์ที่กล่าวมา สรุปได้ว่า สเฟียริซิตี มีความเสถียรภาพต่อการเลื่อน การหมุน และการย่อ/ขยายของภาพวัตถุ

บทพิสูจน์ความเสถียรภาพของมุมของจุดเด่นต่อการเลื่อน, การหมุน และการย่อ/ขยาย

1. มุมของจุดเด่นที่เสถียรภาพต่อการเลื่อนของเส้นขอบภาพ จากภาพที่ 91(ก) มุมของจุดเด่นสามารถคำนวณได้ดังนี้

ภาพที่ 91



$$\theta = \text{Cos}^{-1} \left(\frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} \right) \quad \text{---} \quad \text{---} \quad \text{(ข-1)}$$

เมื่อ $\vec{A} = (x_1 - x_2)\vec{i} + (y_1 - y_2)\vec{j}$ _____ (ข-1a)

$\vec{B} = (x_3 - x_2)\vec{i} + (y_3 - y_2)\vec{j}$ _____ (ข-1b)

$|\vec{A}| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ _____ (ข-1c)

$|\vec{B}| = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}$ _____ (ข-1d)

ดังนั้น $\theta = \text{Cos}^{-1} \left(\frac{(x_1 - x_2)(x_3 - x_2) + (y_1 - y_2)(y_3 - y_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}} \right)$ (ข-2)

ในทำนองเดียวกัน มุมของจุดเด่นในภาพที่ 91(ข) สามารถคำนวณได้จาก

$$\beta = \text{Cos}^{-1} \left(\frac{\vec{C} \cdot \vec{D}}{|\vec{C}| |\vec{D}|} \right) \quad \text{---} \quad \text{(ข-3)}$$

เมื่อ $\vec{C} = (x_1 + m - x_2 - m)\vec{i} + (y_1 + n - y_2 - n)\vec{j}$ _____ (ข-3a)

$\vec{D} = (x_3 + m - x_2 - m)\vec{i} + (y_3 + n - y_2 + n)\vec{j}$ _____ (ข-3b)

$|\vec{C}| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ _____ (ข-3c)

$|\vec{D}| = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}$ _____ (ข-3d)

ดังนั้น $\beta = \text{Cos}^{-1} \left(\frac{(x_1 - x_2)(x_3 - x_2) + (y_1 - y_2)(y_3 - y_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}} \right)$ (ข-4)

จะเห็นได้ว่า สมการที่ (ข-2) เท่ากับสมการที่ (ข-4) ดังนั้น จึงสรุปได้ว่า แม้จะมีการเลื่อนเส้นขอบภาพไป ก็ไม่มีผลกระทบต่อมุมของจุดเด่น

2. มุมของจุดเด่นที่เสถียรภาพต่อการหมุนของเส้นขอบภาพ จากภาพที่ 92(ก) มุมของจุดเด่น (θ) สามารถคำนวณได้ดังนี้

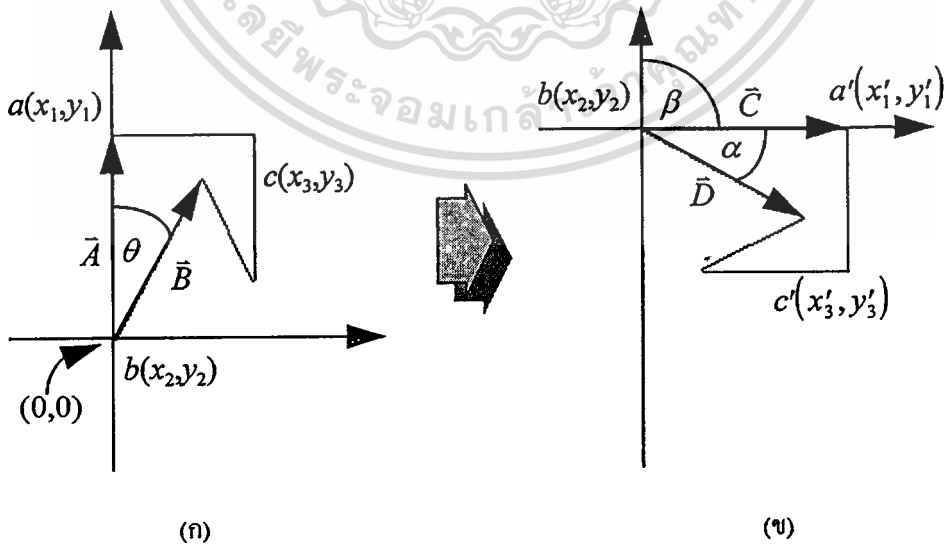
$$\theta = \text{Cos}^{-1} \left(\frac{(x_1)(x_3) + (y_1)(y_3)}{\sqrt{(x_1)^2 + (y_1)^2} \sqrt{(x_3)^2 + (y_3)^2}} \right) \quad \text{--- (ข-5)}$$

ทำการหมุนเวกเตอร์ \vec{A} และ \vec{B} ไปเป็นมุม β โดยมีจุดหมุนอยู่ที่จุด b จะได้ว่า

$$\begin{aligned} a'(x'_1, y'_1) &= [x_1 \quad y_1] \begin{bmatrix} \text{Cos}\beta & -\text{Sin}\beta \\ \text{Sin}\beta & \text{Cos}\beta \end{bmatrix} \\ &= [x_1 \text{Cos}\beta + y_1 \text{Sin}\beta \quad -x_1 \text{Sin}\beta + y_1 \text{Cos}\beta] \end{aligned} \quad \text{--- (ข-6a)}$$

$$\begin{aligned} c'(x'_3, y'_3) &= [x_3 \quad y_3] \begin{bmatrix} \text{Cos}\beta & -\text{Sin}\beta \\ \text{Sin}\beta & \text{Cos}\beta \end{bmatrix} \\ &= [x_3 \text{Cos}\beta + y_3 \text{Sin}\beta \quad -x_3 \text{Sin}\beta + y_3 \text{Cos}\beta] \end{aligned} \quad \text{--- (ข-6b)}$$

ภาพที่ 92



ในทำนองเดียวกัน เวกเตอร์ \vec{C} และ \vec{D} ก็สามารถคำนวณได้ดังนี้

$$\vec{C} = (x_1 \cos \beta + y_1 \sin \beta) \vec{i} + (-x_1 \sin \beta + y_1 \cos \beta) \vec{j} \quad \text{--- (ข-7a)}$$

$$\vec{D} = (x_3 \cos \beta + y_3 \sin \beta) \vec{i} + (-x_3 \sin \beta + y_3 \cos \beta) \vec{j} \quad \text{--- (ข-7b)}$$

$$|\vec{C}| = \sqrt{(x_1 \cos \beta + y_1 \sin \beta)^2 + (-x_1 \sin \beta + y_1 \cos \beta)^2} \quad \text{--- (ข-7c)}$$

$$|\vec{D}| = \sqrt{(x_3 \cos \beta + y_3 \sin \beta)^2 + (-x_3 \sin \beta + y_3 \cos \beta)^2} \quad \text{--- (ข-7d)}$$

$$\alpha = \cos^{-1} \left(\frac{(x_1 \cos \beta + y_1 \sin \beta)(x_3 \cos \beta + y_3 \sin \beta) + (-x_1 \sin \beta + y_1 \cos \beta)(-x_3 \sin \beta + y_3 \cos \beta)}{\sqrt{(x_1 \cos \beta + y_1 \sin \beta)^2 + (-x_1 \sin \beta + y_1 \cos \beta)^2} \times \sqrt{(x_3 \cos \beta + y_3 \sin \beta)^2 + (-x_3 \sin \beta + y_3 \cos \beta)^2}} \right) \quad \text{--- (ข-7e)}$$

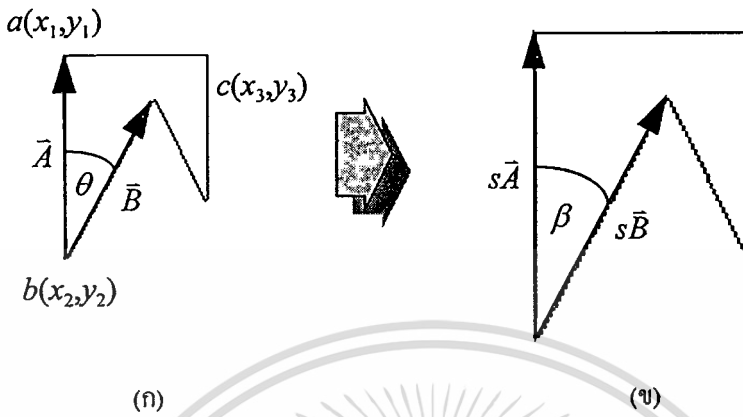
$$\theta = \cos^{-1} \left(\frac{(x_1)(x_3) + (y_1)(y_3)}{\sqrt{(x_1)^2 + (y_1)^2} \sqrt{(x_3)^2 + (y_3)^2}} \right) \quad \text{--- (ข-8)}$$

จะสังเกตเห็นได้ว่า สมการที่ (ข-5) เท่ากับสมการที่ (ข-8) ดังนั้น จึงสรุปได้ว่า การหมุนของเส้นขอบภาพไม่มีผลกระทบต่อมุมของจุดเด่น

3. มุมของจุดเด่นที่เสถียรภาพต่อการย่อหรือขยายเส้นขอบภาพ จากภาพที่ 93(ก) มุมของจุดเด่นสามารถคำนวณได้ดังนี้

$$\theta = \cos^{-1} \left(\frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} \right) \quad \text{--- (ข-9)}$$

ภาพที่ 93



ภาพแสดงการย่อหรือขยายของเส้นขอบภาพ

ถ้าเวกเตอร์ \vec{A} และ \vec{B} ถูกคูณด้วยอัตราย่อหรือขยาย s มุมระหว่างเวกเตอร์ $s\vec{A}$ และ $s\vec{B}$ สามารถคำนวณได้จาก

$$\begin{aligned}
 \beta &= \text{Cos}^{-1} \left(\frac{s\vec{A} \cdot s\vec{B}}{|s\vec{A}| |s\vec{B}|} \right) \\
 &= \text{Cos}^{-1} \left(\frac{s^2 (\vec{A} \cdot \vec{B})}{s^2 |\vec{A}| |\vec{B}|} \right) \quad \text{(ข-10)} \\
 &= \text{Cos}^{-1} \left(\frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} \right)
 \end{aligned}$$

เมื่อพิจารณาสมการที่ (ข-9) และ (ข-10) จะพบว่า สมการทั้งสองนั้นเท่ากันทุกประการ จึงสรุปได้ว่า การย่อ/ขยายเส้นขอบภาพ จะไม่มีผลกระทบต่อมุมของจุดเด่นแต่ประการใด

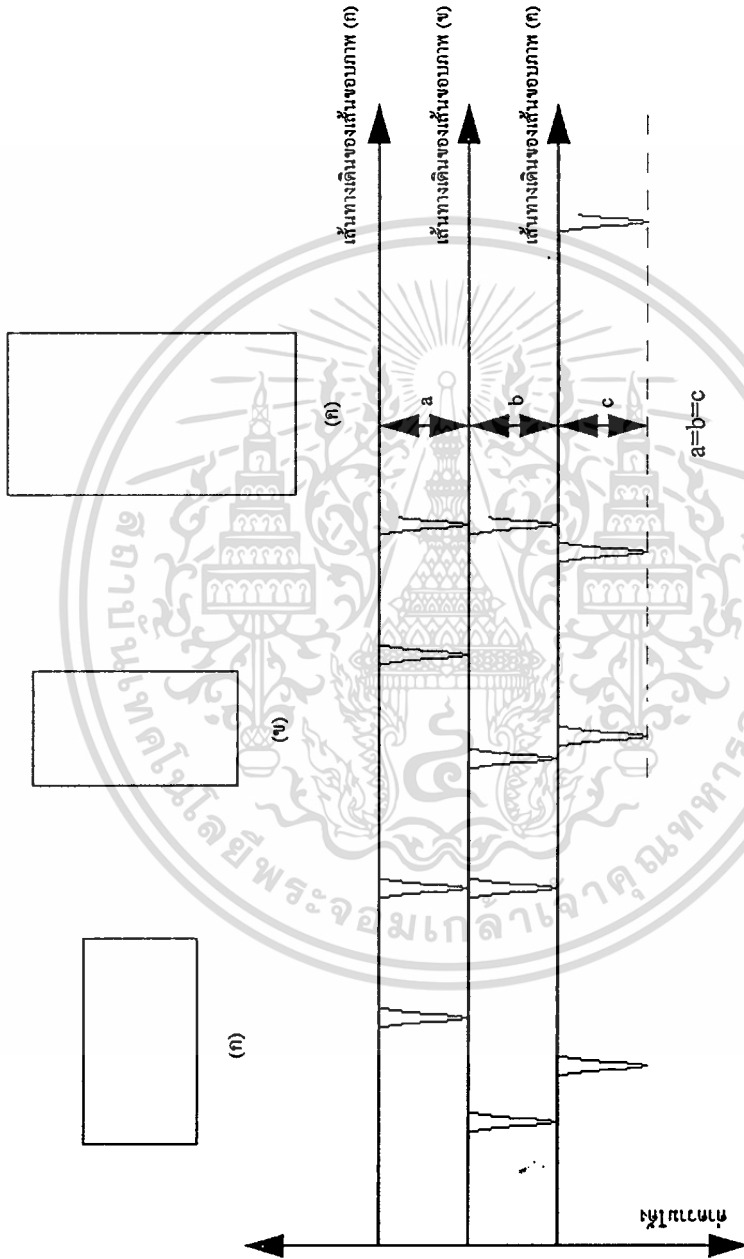
บทพิสูจน์ความมีเสถียรภาพของค่าความโค้งและทิศทางความโค้งของจุดเด่นต่อการเลื่อน การหมุน และการย่อ/ขยาย

ค่าความโค้งและทิศทางความโค้งของจุดเด่นจะเสถียรภาพต่อการเลื่อน การหมุน และการย่อ/ขยายของเส้นขอบภาพเสมอ ดังพิสูจน์ในภาพที่ 94 ซึ่งจะพบว่า แม้จะมีการหมุนภาพ (ภาพที่ 94(ข)) หรือการย่อ/ขยายภาพ (ภาพที่ 94 (ค)) ก็ไม่มีผลต่อค่าความโค้งและทิศทางความโค้งแต่อย่างใด นั่นคือ $a=b=c$ เสมอ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 94

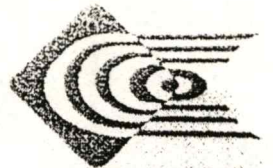


ภาพแสดงค่าความถี่และทิศทางความถี่ที่เสถียรภาพต่อการหมุนและเอียงของเส้นขอบภาพ

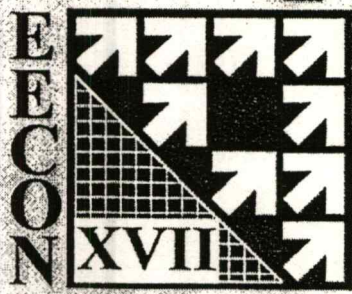
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



50th ANNIVERSARY
KMITNB
for Technology-Excellence



การประชุมวิชาการของวิศวกรรมไฟฟ้า
ครั้งที่ ๑๗ - ๑๘ ธันวาคม ๒๕๓๗
ณ คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจหาจุดเด่นบนเส้นโค้งของขอบภาพโดยการเปลี่ยนค่าสเกล-สเปซของตัวกรอง
Using variable scale-space filtering of Gaussian to detect dominant points

พาวา ลิ้มมะณีประเสริฐ * เกษตร์ ทิริสันติสัมพันธ์ **

* นักศึกษาปริญญาโท คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

** อาจารย์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ การตรวจหาจุดเด่นของขอบภาพเป็นขั้นตอนที่สำคัญอันหนึ่งในการรู้จำรูปร่างของวัตถุ วิธีการหนึ่งที่มีประสิทธิภาพในการตรวจหาจุดเด่น คือการใช้ตัวกรองเกาส์เซียนที่มีการเปลี่ยนค่าสเกล-สเปซ ในบทความนี้ได้นำเสนอวิธีการหาจุดเด่นพร้อมกันแสดงให้เห็นจุดเด่นที่ไม่เสถียรภาพเมื่อมีการเปลี่ยนค่าสเกล-สเปซ เราได้แนะนำวิธีการแก้ไขปัญหาดังกล่าว โดยการตรึงจุดเด่นให้คงที่ตลอดเวลาเมื่อมีการเปลี่ยนค่าสเกล-สเปซ ซึ่งจากวิธีการดังกล่าวจะทำให้ได้จุดเด่นที่ได้ไม่แปรตามการเคลื่อนย้ายตำแหน่งของวัตถุ การหมุนและการย่อขยาย

Abstract. The detection of dominant points is an important step for shape recognition. The effective method for finding the dominant points is to use variable scale-space filtering of Gaussian to detect them on digital curve. In this paper is presented how to detect the dominant points by variable scale-space filtering and show that the resultant is unstable of the dominant points, as the scale-space is changed. we suggest the method to solve this problem by tracing back the dominant point contours in the scale-space image. The dominant points do not change under translation, rotation and scaling condition from this method.

1. บทนำ

ในการจดจำรูปร่างของวัตถุ (shape recognition) เราจะใช้จุดเด่น (dominant points) เป็นตัวจดจำคุณลักษณะของวัตถุนั้นๆ ซึ่งจุดเด่นเหล่านี้ เราสามารถที่จะหาได้จากค่าความโค้งของเส้นรอบรูปของวัตถุนั้นเอง ในปัจจุบันเราสามารถแบ่งวิธีการหาจุดเด่นได้ 2 วิธีใหญ่ๆ คือ วิธีที่หาได้โดยตรงจากมุมที่เกิดขึ้นบนเส้นรอบรูป (contour) เช่น อัลกอริธึมของรัมเมอร์ (Ramer algorithm) ที่ได้เสนอไว้ใน [1] และวิธีที่หาได้จากภาพประมาณการรูปหลายเหลี่ยมแบบจริงเส้น (Linear Polygonal Approximation) ซึ่งเป็นวิธีการที่ใช้ในการผ่านสเกลในบทความนี้

ในการหาจุดเด่นของเส้นรอบรูปของวัตถุ จำเป็นจะต้องใช้ตัวกรองเกาส์เซียน (Gaussian filter) เข้าไปประมวล (convolution) กับเส้นรอบรูปของวัตถุเพื่อหาจุดเด่น ซึ่งในตัวกรองเกาส์เซียนเรา

สามารถที่จะเปลี่ยนค่าสเกล-สเปซ (scale-space) ของตัวกรองเกาส์เซียนได้ เพื่อหาจุดเด่นที่เหมาะสม แต่เมื่อเราเปลี่ยนค่าสเกล-สเปซไปแล้ว จะมีผลทำให้จุดเด่นที่ได้จากค่าสเกล-สเปซ ที่แตกต่างกันมีตำแหน่งที่ต่างกันด้วย ซึ่งในบทความนี้จะแสดงวิธีการแก้ไขปัญหาดังกล่าว โดยเมื่อมีการเปลี่ยนค่าสเกล-สเปซไปแล้ว จะยังคงตรึงจุดเด่นให้อยู่ในตำแหน่งเดิม

ในบทความนี้จะประกอบไปด้วยหัวข้อใหญ่ๆทั้งหมด 6 หัวข้อ ในหัวข้อแรกจะเป็นบทนำ ในหัวข้อที่ 2 จะกล่าวถึงคณิตศาสตร์ของตัวกรองเกาส์เซียนที่ใช้ในการหาจุดเด่น หัวข้อที่ 3 จะเป็นกรณีศึกษาคณิตศาสตร์ของตัวกรองเกาส์เซียนไม่ใช่วงรีกับเส้นรอบรูป ส่วนในหัวข้อที่ 4 นั้น จะเป็นการแก้ปัญหาการเคลื่อนของจุดเด่น เมื่อมีการเปลี่ยนค่า สเกล-สเปซ และในหัวข้อที่ 5 จะเป็นผลจากการทดลองพร้อมกับสรุปผลที่ได้ในหัวข้อที่ 6

2. คณิตศาสตร์ของตัวกรองเกาส์เซียน

จุดที่มีความโค้งสูงที่อยู่บนเส้นรอบรูป (sharp contour) จะเป็นตัวบ่งบอกถึงคุณลักษณะของเส้นรอบรูปของวัตถุนั้น (object contour) เราจะเรียกจุดเหล่านี้ว่า "จุดเด่น" ("dominant points") ซึ่งรายละเอียดในการหาจุดเด่นจะได้กล่าวต่อไป

2.1 การหาค่าความโค้งบนเส้นโค้ง

การคำนวณหาค่าความโค้งของเส้นรอบรูป ที่มีฟังก์ชัน เป็นฟังก์ชันที่ไม่ต่อเนื่อง (discrete function) จะไม่ง่ายเหมือนกับการคำนวณในกรณีที่เป็นฟังก์ชันแบบต่อเนื่อง (continuous function) ซึ่ง Mokhtarian และ Mackworth [2] ซึ่งได้ทำการพัฒนาสูตรในการหาค่าความโค้งแบบไม่ต่อเนื่องเอาไว้แล้ว และจะได้นำเสนอขอเป็นสังเขปดังต่อไปนี้

ฟังก์ชันของความโค้ง (curvature) $y(x)$ ถูกนิยามไว้ว่า

$$k(x) = \frac{|y''(x)|}{(1+(y'(x))^2)^{3/2}} \tag{1}$$

เมื่อ $y'(x) = \frac{dy}{dx}$

และ $y''(x) = \frac{d^2y}{dx^2}$ (2)

แต่เนื่องจากว่า เส้นรอบรูป (contour) เราได้มาจากภาพจริง (real image) ซึ่งภาพเหล่านี้ จะไม่มีคุณสมบัติของฟังก์ชัน $y = y(x)$ ด้วยเหตุผลดังกล่าว เราจึงต้องทำการแยกตัวแปร x และตัวแปร y ออกมาอยู่ในรูปแบบของสมการพารามตริก (parametric form) คือ $x(t)$ และ $y(t)$ เมื่อ $0 < t < L$ และ L เป็นความยาวส่วนโค้งของเส้นรอบรูป ดังนั้น

$$y'(x) = \frac{\dot{y}(t)}{\dot{x}(t)} \tag{3}$$

และ $y''(x) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t))^3}$ (4)

เมื่อ $\dot{x}(t) = \frac{dx}{dt}; \dot{y}(t) = \frac{dy}{dt}$; (5)

$$\ddot{x}(t) = \frac{d^2x}{dt^2}; \ddot{y}(t) = \frac{d^2y}{dt^2}$$

แทนค่าสมการ (5) ใน (1) เราจะได้

$$k(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \tag{6}$$

แต่ในการใช้งานจริง เราจะไม่นิยมหาค่าความโค้ง $k(t)$ ตามสมการที่ (5) และ (6) โดยตรง เราจะใช้วิธีการประสาน (convolution) ฟังก์ชัน $x(t)$ และ $y(t)$ ด้วยตัวกรองเกาส์เซียนแทน เพราะว่าตัวกรองเกาส์เซียนจะมีคุณสมบัติในการกำจัดสัญญาณรบกวน ซึ่งพหามิเตอร์ที่ทำการหาค่าความโค้งในการกำจัดสัญญาณรบกวนก็คือ σ (scale-space) นั่นคือ (ดูที่มาของสมการได้ในภาคผนวก)

$$\left. \begin{aligned} \dot{X}(t, \sigma) &= x(t) * \dot{g}(t, \sigma); \\ \dot{Y}(t, \sigma) &= y(t) * \dot{g}(t, \sigma); \\ \ddot{X}(t, \sigma) &= x(t) * \ddot{g}(t, \sigma); \\ \ddot{Y}(t, \sigma) &= y(t) * \ddot{g}(t, \sigma); \end{aligned} \right\} \tag{7}$$

เมื่อ $\dot{g}(t, \sigma) = \frac{\partial g(t, \sigma)}{\partial t};$
 $\ddot{g}(t, \sigma) = \frac{\partial^2 g(t, \sigma)}{\partial t^2}$ (8)

และ $g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$ (9)

เราสามารถเขียนสมการ (8) ใหม่ได้เป็น

$$k(t, \sigma) = \frac{\dot{X}(t, \sigma)\ddot{Y}(t, \sigma) - \dot{Y}(t, \sigma)\ddot{X}(t, \sigma)}{(X^2(t, \sigma) + Y^2(t, \sigma))^{3/2}} \tag{10}$$

2.2 การกรองเส้นขอบภาพเพื่อให้ราบเรียบ (smooth)

ถ้าเส้นขอบภาพไม่ราบเรียบ อันเนื่องมาจากมีสัญญาณรบกวน (quantization noise) ซึ่งจะส่งผลกระทบต่อจุดเด่น แต่เราสามารถทำให้เส้นขอบภาพนั้นให้ดีขึ้นได้ ด้วยวิธีการใช้ตัวกรองเกาส์เซียนเป็นตัวกำจัดสัญญาณรบกวนที่จะเข้ามารบกวนเส้นขอบภาพได้

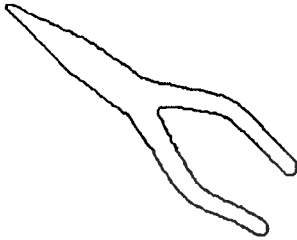
การกรองเส้นขอบภาพสามารถทำได้โดย นำเอาตัวกรองเกาส์เซียนไปทำการประสาน (convolution) กับเส้นขอบภาพ ซึ่งผลจากการกระทำดังกล่าว จะทำให้เส้นขอบภาพที่ถูกรบกวนจะราบเรียบขึ้น และเรายังสามารถปรับค่าสเกล-สเปซของตัวกรองเกาส์เซียน เพื่อให้ผลที่ได้มีความราบเรียบตามแต่เราต้องการได้อีกด้วย

สมการตัวกรองเกาส์เซียนได้แสดงไว้ใน สมการที่ (9) แล้ว ส่วนการประสาน (convolution) กับเส้นขอบภาพสามารถทำได้โดยสมการข้างล่างนี้

$$X(t, \sigma) = x(t) * g(t, \sigma)$$

$$Y(t, \sigma) = y(t) * g(t, \sigma)$$

เมื่อ $X(t, \sigma)$ และ $Y(t, \sigma)$ คือเส้นขอบภาพที่ผ่านการกรองด้วยตัวกรองเกาส์เซียนแล้ว ซึ่งผลจากการกระทำดังกล่าวกับรูปภาพ จะได้ผลดังแสดงไว้ในรูปที่ 2



รูปที่ 1 แสดงรูปต้นแบบที่ยังไม่ผ่านการขบวนการทำให้ราบเรียบ (smooth)



รูปที่ 2 เป็นรูปที่ผ่านการขบวนการทำให้ราบเรียบ (smooth) แล้ว ด้วยค่า $\sigma = 15$

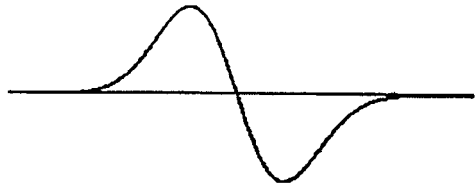
3. การนำคณิตศาสตร์ของตัวกรองเกาส์เซียนไปใช้งานจริงกับเส้นรอบรูป

ในหัวข้อที่ 2 สมการที่ (9) คือสมการเกาส์เซียนนั่นเอง สามารถเขียนได้เป็น

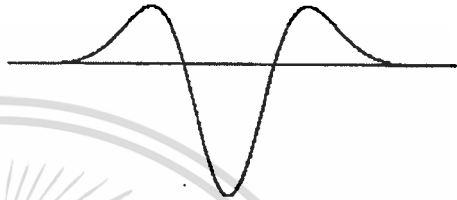
$$g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

$$g(t, \sigma) = \frac{-t}{\sigma^3\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (11)$$

$$g(t, \sigma) = \left(\frac{t^2}{\sigma^2} - 1\right) \frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (12)$$



รูปที่ 3 แสดงรูปตัวกรองเกาส์เซียนของอนุพันธ์ย่อยอันดับที่หนึ่งที่ $\sigma = 3$



รูปที่ 4 แสดงรูปตัวกรองเกาส์เซียนของอนุพันธ์ย่อยอันดับที่สองที่ $\sigma = 3$

ในการประสาน (convolution) ด้วยตัวกรองเกาส์เซียน เราจำเป็นต้องสร้างหน้าต่าง (window) เพื่อเก็บค่าสัมประสิทธิ์ของอนุพันธ์ย่อยอันดับที่หนึ่งและอนุพันธ์ย่อยอันดับที่สองของตัวกรองเกาส์เซียนที่จะนำไปประสาน (convolution) กับ $x(t)$ และ $y(t)$ ซึ่งขนาดของหน้าต่าง เราจะใช้ขนาดเท่ากับ $2N+1$ เมื่อค่า N สามารถหาได้จากสมการข้างล่างนี้

$$N = \min\{N_1, N_2\} \quad (13)$$

เมื่อ

$$N_1 = \{x: \text{Max} \sum_{t=-n}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \leq 0.99\} \quad (14)$$

และ $N_2 = \lfloor 5\sigma \rfloor$; เมื่อเครื่องหมาย $\lfloor a \rfloor$ หมายถึง เลขจำนวนเต็มที่มีมากที่สุดซึ่งน้อยกว่าหรือเท่ากับ a

เมื่อเราใส่สมการ (11) และ (12) ไปทำการประสาน (convolution) กับ $x(t)$ และ $y(t)$ โดยที่ $x(t)$ และ $y(t)$ คือค่าแทนของเส้นรอบรูปที่อยู่บน x และ y คอออกคือค่า ตามลำดับ ซึ่งเส้นรอบรูปนี้จะต้องเป็นเส้นรอบรูปที่ปิดเท่านั้น นั่นคือ

$$X(t, \sigma) = x(t) * g(t, \sigma); \quad (15)$$

$$Y(t, \sigma) = y(t) * g(t, \sigma);$$

$x(t)$ เป็นตำแหน่งของเส้นรอบรูปปิดสำหรับ x คอออกคือ

$$\begin{aligned} \dot{Y}(t, \sigma) &= y(t) * \dot{g}(t, \sigma); \\ Y(t, \sigma) &= y(t) * \dot{g}(t, \sigma); \end{aligned} \quad (16)$$

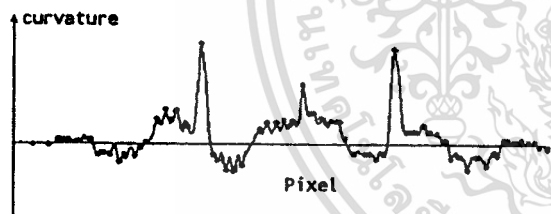
$y(t)$ เป็นตำแหน่งของเส้นรอบรูปที่ปิดสำหรับ y คอออกอินเท หลังจากนี้ให้นำค่า $\dot{X}(t, \sigma), X(t, \sigma), \dot{Y}(t, \sigma)$ และ $Y(t, \sigma)$ ไปแทนในสมการ (8) ซึ่งจะได้ค่าความโค้ง ดังนี้

$$k(t, \sigma) = \frac{\dot{X}(t, \sigma)\dot{Y}(t, \sigma) - X(t, \sigma)\ddot{Y}(t, \sigma)}{(\dot{X}^2(t, \sigma) + \dot{Y}^2(t, \sigma))^{3/2}}$$

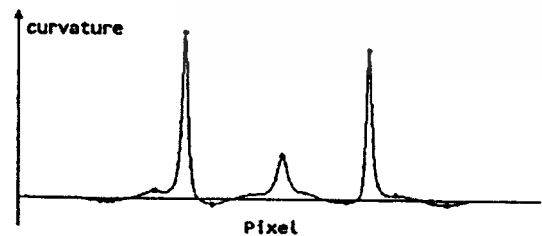
ซึ่งค่าความโค้งที่ต่ำที่สุดและสูงที่สุดจะเป็นจุดเด่นที่เกิดขึ้น ณ ตำแหน่ง t นั้นๆ บนเส้นรอบรูปปิด

รูปที่ 1 เป็นรูปต้นแบบที่จะนำมาทดลองหาจุดเด่น เมื่อเรานำรูปต้นแบบซึ่งถูกเก็บค่าตำแหน่ง x และ y ไว้ในรูปแบบของสมการพหุนามเทริก ไปทำการประสาน (convolution) กับสมการ $\dot{g}(t, \sigma)$ และ $\dot{g}(t, \sigma)$ ด้วยค่า $\sigma = 3$ และ $\sigma = 15$ ซึ่งจะได้ค่าความโค้งดังแสดงในรูปที่ 5 และ 6

ถ้าเราพิจารณารูปที่ 5 เราจะเห็นจุดสูงสุดและจุดต่ำสุดของค่าความโค้งที่เกิดขึ้นบนเส้นรอบรูปของวัตถุรูปที่ 1 ที่แสดงด้วยเครื่องหมาย \circ ซึ่งก็คือจุดเด่นของเส้นรอบรูปของวัตถุรูปที่ 1 นั้นเอง



รูปที่ 5 แสดงค่าความโค้งที่ได้จากรูปต้นแบบ (รูปคิม) โดยมีค่า $\sigma = 3$



รูปที่ 6 แสดงค่าความโค้งที่ได้จากรูปต้นแบบ (รูปคิม) โดยมีค่า $\sigma = 15$

เราสามารถที่จะหาจุดเด่นในแต่ละ สเกล-สเปซ โดยการเพิ่มค่า σ ไปทีละ $\sigma = \sigma + 2$ ซึ่งที่ค่า σ ต่างๆ เราสามารถหาจุดเด่นได้

ดังแสดงในรูปที่ 10 ซึ่งเราจะเห็นได้ว่าเมื่อมีการเปลี่ยนค่า σ (scale-space) ไปจะทำให้จุดเด่นมีการเคลื่อนไปตามค่า σ ที่เปลี่ยนไปด้วยการแก้ปัญหาที่กล่าวถึงในหัวข้อที่ 4



รูปที่ 7 แสดงถึงจุดเด่นของขอบภาพที่ $\sigma = 3$

4. การแก้ปัญหาการเลื่อนของจุดเด่น เมื่อมีการเปลี่ยนค่า σ (scale-space)

เนื่องจากค่าตำแหน่งของจุดเด่นที่เกิดขึ้นบนเส้นรอบรูปปิดนั้น จะมีความเที่ยงตรงในกรณีที่ค่า σ (สเกล-สเปซ) มีค่าน้อยๆ ดังนั้นในบทความนี้ เราจะเลือกค่า $\sigma = 3$ เป็นค่าเริ่มต้น แล้วค่อยๆ เพิ่มค่า σ ขึ้นเป็น $\sigma = 3, 5, 7, 9, \dots, 27$ และต้องจัดเตรียมอาร์เรย์เอาไว้เก็บค่าตำแหน่งของจุดเด่นอีก 3 ชุด คือ $P1(i), P2(i)$ และ $P3(i)$ โดยที่ $P3(i)$ จะทำหน้าที่เก็บตำแหน่งของจุดเด่นที่ค่า $\sigma = 3, P2(i)$ จะเก็บตำแหน่งของจุดเด่นที่เกิดขึ้น ณ ค่า σ ปัจจุบัน ส่วน $P1(i)$ จะเก็บตำแหน่งของจุดเด่นที่เกิดขึ้นก่อนหน้านี้ ($\sigma - 2$) ซึ่งวิธีการแก้ปัญหาการเลื่อนของจุดเด่น ดังที่เกิดขึ้นในหัวข้อที่ 3 สามารถสรุปเป็นขั้นตอนได้ดังนี้

(1) เริ่มต้นด้วยค่า $\sigma = 3$ (scale-space) แล้วทำการหาจุดเด่นที่เกิดขึ้นบนเส้นโค้งซึ่งจุดเด่นจะอยู่ในตำแหน่งของเส้นรอบรูปที่ $i = n_1, n_2, n_3, \dots, n_{M1}$ จากนั้นเก็บตำแหน่งของจุดเด่นต่างๆ เอาไว้ในอาร์เรย์ $P3(i)$ และ $P1(i)$ สำหรับ $i = 1, 2, 3, \dots, M1$

(2) เพิ่มค่าของ σ (scale-space) ขึ้นอีก 2, $\sigma = \sigma + 2$ แล้วหาจุดเด่นบนเส้นโค้ง $k(t, \sigma)$ ซึ่งตำแหน่งของจุดเด่นจะอยู่บนเส้นรอบรูปที่ $i = m_1, m_2, m_3, \dots, m_{M2}$ จากนั้นเก็บตำแหน่งของจุดเด่นต่างๆ เอาไว้ในอาร์เรย์ $P2(i)$ สำหรับ $i = 1, 2, 3, \dots, M2$ โดยทฤษฎีของสเกล-สเปซ [2] จำนวนของจุดเด่นจะลดลง เมื่อค่า σ เพิ่มขึ้น นั่นคือ $M2 \leq M1$

(3) พยายามเลื่อนจุดเด่นที่ได้จากข้อ 2 ให้อีกกลับมามีตำแหน่งเดิมที่ค่า σ เท่ากับ 3 โดยเขียนเป็น อัลกอริทึมได้ดังนี้

For $i = 1$ to $M2$

Begin

For $j = 1$ to $M1$

Begin

Search $P1(k)$ So that

$$\text{Min}_k |P1(j) - P2(i)|$$

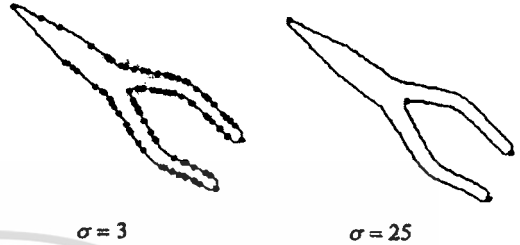
End

$$P3(i) = P3(k)$$

End.

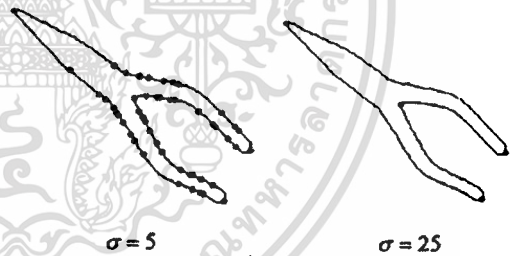
5. ผลการทดลอง

จากการทดลองเราจะใช้รูปคีมแบบที่เป็นรูปคีม เริ่มต้นด้วยค่า $\sigma = 3$ และเพิ่มค่า $\sigma = \sigma + 2$ ซึ่งจะได้จุดเด่นที่เกิดขึ้นบนเส้นขอบภาพดังแสดงไว้ในรูปที่ 7 ที่คำนวณได้จากสมการที่ (10) โดยตรง



รูปที่ 10 แสดงให้เห็นถึงจุดเด่นที่คำนวณได้ที่ค่า σ ต่างๆ

เราลองพิจารณารูปที่ 8 และ 10 ให้ดี จะเห็นได้ว่าจุดเด่นที่เกิดขึ้นที่ค่า σ ต่างๆ นั้นจะไม่อยู่ในตำแหน่งเดิมโดยตลอด จะมีการเลื่อนของจุดเด่นเกิดขึ้นซึ่งสามารถเช่นนี้เราเรียกว่า 'จุดเด่นไม่เสถียร' สังเกตได้จากกรณีที่ค่า $\sigma = 3$ กับ $\sigma = 25$ ซึ่งเมื่อเราแก้ปัญหาดังกล่าวดังที่ได้อธิบายไว้ในหัวข้อที่ 4 แล้วจะได้ผลดังแสดงไว้ในรูปที่ 9 และ 11



รูปที่ 11 แสดงให้เห็นถึงจุดเด่นที่ได้ทำการแก้ไขโดยการจูนจุดเด่นให้คงที่

(4) นำตำแหน่งของจุดเด่นที่ได้ใน $P2(i)$ ย้ายไปเก็บไว้ใน $P1(i)$ เพื่อเป็นตำแหน่งอ้างอิงในการเปรียบเทียบค่าตำแหน่งของจุดเด่นชุดใหม่ ที่ได้จากการเพิ่มค่า $\sigma = \sigma + 2$ ซึ่งสามารถเขียนเป็นอัลกอริทึมได้ดังนี้

For $i = 1$ to $M2$

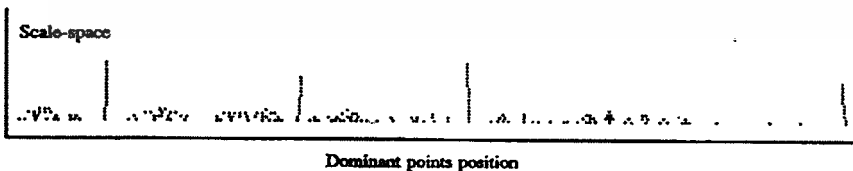
Begin

$$P1(i) = P2(i)$$

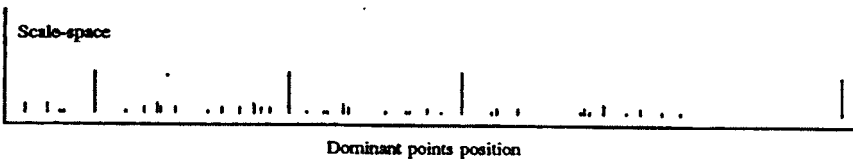
End.

(5) ย้อนกลับไปทำขั้นตอนที่ (2) ใหม่

เมื่อเราทำการพลอตตำแหน่งของจุดเด่นเทียบกับค่า สเกล-สเปซ จะได้ผลดังแสดงในรูปที่ 8 ซึ่งจะเห็นได้ว่า เมื่อมีการเปลี่ยนค่าของสเกล-สเปซไป จะทำให้ตำแหน่งของจุดเด่นเปลี่ยนแปลงไป ส่วนในรูปที่ 9 จะเป็นผลที่ได้จากการรวมการจูนจุดเด่นให้คงที่เมื่อเราเปลี่ยนค่าสเกล-สเปซไป



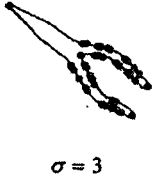
รูปที่ 8 เป็นรูปภาพแสดงความสัมพันธ์ระหว่างค่าสเกล-สเปซและตำแหน่งของจุดเด่น (ยังไม่ได้จูนจุดเด่น)



รูปที่ 9 เป็นรูปภาพแสดงความสัมพันธ์ระหว่างค่าสเกล-สเปซและตำแหน่งของจุดเด่น (หลังจากจูนจุดเด่นแล้ว)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนั้น เรายังทดลองย่อภาพรูปต้นแบบให้เล็กลงเหลือเพียง 0.5 เท่าของรูปต้นแบบ แล้วคำนวณหาจุดเด่น ซึ่งจะได้ผลดังแสดงไว้ในรูปที่ 12

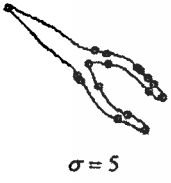


$\sigma = 3$



$\sigma = 9 - 15$

รูปที่ 12 แสดงจุดเด่นที่ได้จากการย่อขนาดของภาพทดลองครั้งแรกนี้



$\sigma = 5$



$\sigma = 9 - 15$

รูปที่ 13 แสดงถึงจุดเด่นที่ได้จากรูปที่ 12 แต่ได้ตรึงจุดเด่นเอาไว้

ทดลองหมุนภาพต้นแบบ (รูปคิม) ไปเป็นมุม 180° แล้วทำการคำนวณหาจุดเด่น ซึ่งได้ผลการทดลองดังแสดงไว้ในรูปที่ 14

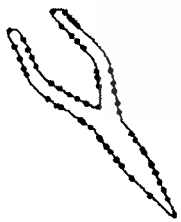


$\sigma = 3$



$\sigma = 25$

รูปที่ 14 แสดงจุดเด่นที่ได้จากการหมุนรูปต้นแบบ (รูปคิม) ไปเป็นมุม 180°



$\sigma = 5$

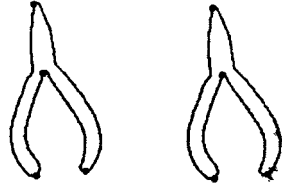


$\sigma = 25$

รูปที่ 15 แสดงให้เห็นรูปที่ได้ทำการตรึงจุดเอาไว้แล้ว

จากรูปที่ 14 ถ้าเราสังเกต จะเห็นจุดเด่นซึ่งเกิดขึ้นที่ปลายของคิมมีการเลื่อนไปจากตำแหน่งเดิม เมื่อเปรียบเทียบกับรูปที่ 15 แล้ว

จะเห็นได้ค่อนข้างชัดเจนว่า จุดเด่นที่แสดงไว้ในรูปที่ 15 จะเสถียรกว่ารูปที่ 14



รูปที่ 16 แสดงให้เห็นถึงจุดเด่นที่ไม่ได้ตรึงเอาไว้ (รูปซ้าย $\sigma = 23$) และจุดเด่นที่ผ่านการตรึงเอาไว้แล้ว (รูปขวา $\sigma = 23$)



(a) $\sigma = 3$



(b) $\sigma = 25$



(c) $\sigma = 25$

รูปที่ 11 (a) แสดงจุดเด่นที่เกิดขึ้นที่ค่า $\sigma = 3$ (b) เป็นจุดเด่นที่ไม่ได้ผ่านการตรึงเอาไว้ (c) เป็นจุดเด่นที่ผ่านการตรึงเอาไว้แล้ว

6. สรุป

จากที่ได้ทดลองทั้งหมด จะเห็นได้ว่า วิธีการในการตรึงจุดเด่นให้คงที่ได้นั้น ไม่ได้เป็นเรื่องที่ซับซ้อนแต่ประการใด แต่ด้วยวิธีการดังที่กล่าวมาแล้วนั้น ได้ให้ผลที่มีประสิทธิภาพที่ดีพอสมควร ซึ่งจากผลการทดลองที่ผ่านมาทั้งหมด จะเห็นได้ว่าจุดเด่นที่ได้หลังจากที่ได้ทำการตรึงจุดเด่นไว้แล้วนั้น จะไม่เปลี่ยนแปลงหรือเลื่อนแต่อย่างใด เมื่อเรามีการเคลื่อนย้ายภาพ การหมุนภาพ และการย่อหรือขยายภาพ จุดเด่นก็ยังคงตั้งอยู่ที่ตำแหน่งเดิม

ภาคผนวก

พิสูจน์สมการคณิตศาสตร์
สมการเกาส์เซียนถูกนิยามไว้ดังนี้

$$g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad \text{_____ (A)}$$

ผลที่ได้จากการประสาน (convolution) ของตำแหน่ง คอออกติงท
กับตัวกรองเกาส์เซียน จะได้ว่า

$$X(t, \sigma) = x(t) * g(t, \sigma); \quad \text{_____ (B)}$$

$$Y(t, \sigma) = y(t) * g(t, \sigma); \quad \text{_____ (C)}$$

และอนุพันธ์ย่อยอันดับที่หนึ่งของสมการ (B) ก็คือ

$$\begin{aligned} \dot{X}(t, \sigma) &= \frac{\partial X(t, \sigma)}{\partial t} \\ &= \frac{\partial (x(t) * g(t, \sigma))}{\partial t} \\ &= x(t) * \frac{\partial g(t, \sigma)}{\partial t} \\ &= x(t) * \dot{g}(t, \sigma) \end{aligned}$$

และอนุพันธ์ย่อยอันดับที่สองของสมการ (B) คือ

$$\begin{aligned} \ddot{X}(t, \sigma) &= \frac{\partial^2 X(t, \sigma)}{\partial t^2} \\ &= x(t) * \frac{\partial^2 (g(t, \sigma))}{\partial t^2} \\ &= x(t) * \ddot{g}(t, \sigma) \end{aligned}$$

ส่วนสมการ (C) ก็สามารถพิสูจน์ได้ในทำนองเดียวกัน

เอกสารอ้างอิง

- [1] NIRWAN ANSARI and EDWARD J.DELP, On Detecting dominant points, Pattern Recognition, Vol. 24 No. 5 pp. 441-451 1991.
- [2] FARZIN MOKHTARIAN and ALAN MACKWORTH, Scale-Based Description and Recognition of Planar Curve and Two-Dimensional Shapes, IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-8, No. 1, January 1986.
- [3] Daniel M. Wuescher and Kim L. Boyer, Robust Contour Decomposition Using a Constant Curvature Criterion, IEEE Trans. Pattern Anal. Machine Intell., Vol. 13, No. 1, January 1991.
- [4] Soo-Chang Pei and Chao-Nan Lin, The Detection of Dominant Points on Digital Curves by Scale-Space Filtering, Pattern Recognition, Vol. 25, No. 11, pp. 1307-1314, 1992

กิตติกรรมประกาศ

บทความนี้สามารถสำเร็จลงได้ด้วยดี เพราะได้รับความช่วยเหลือจากบุคคลต่อไปนี้ ขอขอบคุณ อาจารย์เกษตร์ ศิริสันติสัมฤทธิ์ ที่ได้ให้คำปรึกษาในการทำวิจัยและข้อมูลในการทำวิจัย ขอขอบคุณคุณดวงเดือน พงษ์เลาหพันธุ์ ที่ช่วยในการจัดทำเอกสาร อำนวยความสะดวกในเรื่องของซอฟต์แวร์และให้กำลังใจ และขอขอบคุณเพื่อนๆ ทุกคนที่ได้ให้ความช่วยเหลือจนงานสำเร็จลงได้ด้วยดี



ISSN 0125-1724

วิศวกรรม

ลาดกระบัง

ถนนวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

LADKRABANG ENGINEERING JOURNAL

Volume 13 Number 1

July, 1996

- | | | |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 11. | การออกแบบและการสร้างวงจรรองสัญญาณเชิงเลขแบบบิตเตอร์เวอร์ชันดับที่ 6
วินัย ทองตัน สมยศ จุณณะวิยะ กอบชัย เศษหาญ | 78 |
| 12. | วงจรรีความถี่เป็นสองเท่าโดยใช้มอสเฟต
ทวิช รุเมือง กอบชัย เศษหาญ | 91 |
| 13. | การออกแบบวงจรรูขนาออกแบบ 4 ควอดแรนท์โดยใช้ซีมอส
ฐิติพงศ์ บุคกงกรทศ ไทบุลย์ คู่ประกาย กอบชัย เศษหาญ | 95 |
| 14. | การประยุกต์คำนวณ HNSOFLC ในขบวนการ MIMO
สุรกิจ เด็คคุณสุข อิศาทร ทิทธิภัทรพันธุ์ งามกล งามวิวิทย์ | 101 |
| 15. | การจดจำอักษรพิมพ์ภาษาไทยโดยโครงข่ายประสาทเทียม แบบจำลอง Self Organizing Map
อัญญา วานิชทวีวัฒน์ บุญฉวี เครือตราฐ | 107 |
| 16. | โครงข่ายเซลล์ประสาทประสาทเทียมแบบไฮปไฟลด์สำหรับการจดจำภาพวัตถุที่มีการทับซ้อนกันบางส่วน
พารา ลิมนะณิประเสริฐ เกษตร์ ศิริสันติสัมพันธ์ อมรฤกษ์ | 119 |
| 17. | การออกแบบตัวควบคุม Invert pandulum สำหรับระบบที่เป็นเชิงเส้น
เกียรติอนันท์ สุรชู วิทยา ทิพย์สุวรรณพร | 131 |
| 18. | การใช้ไมโครคอมพิวเตอร์ในระบบวัดสัญญาณไฟฟ้าหัวใจแบบรวมศูนย์
วัฒนา โพธิ์เจริญ มนต์ สัจวงศิลป์ กิตติพล ธิศสกุล ทิรัช คู่ศิริวานิชกร | 137 |
| 19. | การออกแบบเครื่องกั้นน้ำส้มแบบต่อเนื่อง
พิชณู ศิริวานิชย์วงศ์ ราชัน พูลกิจ ปานมนต์ ศิริสมบุรณ์ ทิฉิต กิตติพันธ์ | 143 |

ฉบับพิเศษ

พระจอมเกล้าลาดกระบังเฉลิมพระเกียรติ

เทคโนโลยีพอเพียง

4-8 กรกฎาคม 2539

โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

สำหรับการจดจำภาพวัตถุที่มีการทับซ้อนกันบางส่วน

Hopfield Neural Network for Overlapping Object Recognition

พารา ลิมมะณีประเสริฐ ผศ. เกษตร์ ศิริสันติสัมฤทธิ์
คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

ในบทความนี้ได้นำเสนอหลักการการจดจำภาพวัตถุ ซึ่งเป็นกระบวนการที่สำคัญในระบบการมองเห็นของคอมพิวเตอร์ กระบวนการการจดจำภาพวัตถุสามารถแบ่งการทำงานได้เป็นสองขั้นตอน คือ ขั้นตอนการคำนวณหาจุดเด่นของภาพวัตถุ [1] และขั้นตอนการจับคู่ระหว่างจุดเด่นของภาพวัตถุที่ถูกบดบังบางส่วนกับภาพวัตถุที่สมบูรณ์ ซึ่งถูกเก็บไว้ในฐานข้อมูล ซึ่งเป็นขั้นตอนที่สำคัญที่จะกล่าวถึงในที่นี้ โดยหลักการทั่วไป การประมวลผลการจับคู่ มักจะเป็นการทำงานแบบลำดับในการควบคุมค่าพารามิเตอร์ต่างๆ และไม่พิจารณาข้อมูลข้างเคียงในการตัดสินใจ ซึ่งเป็นข้อด้อยของวิธีการเหล่านี้ ดังนั้น เพื่อให้ได้มาซึ่งความแม่นยำในกระบวนการจับคู่ จึงนำโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์เข้ามาประยุกต์ใช้งานด้วยลักษณะเด่นหลายประการ เช่น มีความสามารถในการคำนวณเชิงขนานได้ สามารถควบคุมค่าพารามิเตอร์ของจุดเด่นพร้อมๆกันได้มากกว่าหนึ่งค่า เช่น ค่าสเฟิริซิตี ค่าความโค้ง ทิศทางของความโค้ง และมุมของจุดเด่น ซึ่งค่าพารามิเตอร์เหล่านี้สามารถที่จะนำมาพิจารณาพร้อมๆกันและควบคุมด้วยสมการของโฮปฟิลด์ได้ นอกจากนี้ ในกระบวนการการจับคู่ นั้น การเลื่อนตำแหน่ง การหมุน และการย่อหรือขยายของภาพ จะไม่มีผลกระทบต่อในกระบวนการการประมวลผลแต่ประการใด

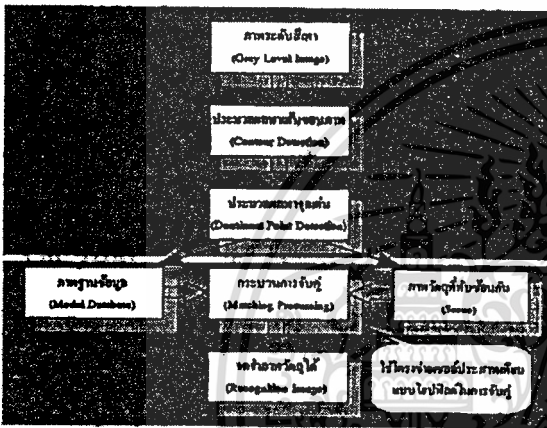
Abstract

In this paper, we present the principle of object recognition which is an important part of computer vision system. Object recognition processing can be classified into two tasks. The first task is to extract the dominant points from the object contour [1] and the second is to match the dominant points between scene, may be consisted of the overlapping object contour, and models. In particular, the second task that is the main idea of proposition in this paper. In several techniques, the matching process controls the parameters in sequence step and does not consider the neighbor information. These are the disadvantage points. In order to find the exact matching dominant points, so we implemented the matching task by a Hopfield neural network. The advantage of this technique is a mate can be performed in parallel computing, one scene can be matched all the models in the database at the same time, and we can also control the parameters of dominant point in one step. Such as, the sphericity, the curvature, the direction of curvature, and the angle of dominant point. These parameters can be combined together and controlled by Hopfield equation. Furthermore, during matching, the translation, rotation, and scaling of the object contour do not effect in processing, as using Hopfield neural network.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. บทนำ

ปัจจุบันการจดจำภาพวัตถุ (Object Recognition) เริ่มเข้ามามีบทบาทในระบบอัตโนมัติมากขึ้น เช่น ระบบอัตโนมัติที่ทำงานด้วยหุ่นยนต์ หรือแขนกล เป็นต้น ระบบการมองเห็นของหุ่นยนต์ หรือแขนกลที่ถูกควบคุมด้วยคอมพิวเตอร์ จะถูกเรียกว่า "ระบบการมองเห็นของคอมพิวเตอร์" (Computer Vision System) ซึ่งการจดจำภาพวัตถุ ก็เป็นส่วนหนึ่งของระบบนี้ด้วย และได้มีการพัฒนาอย่างต่อเนื่องมาโดยตลอด



รูปที่ 1 แสดงถึงขั้นตอนกระบวนการการจดจำภาพวัตถุ

หลักการทำงานของการจดจำภาพวัตถุ ได้แสดงไว้ในรูปที่ 1 โดยข้อมูลที่ป้อนเข้าสู่ระบบนั้น จะได้จากกล้อง CCD ซึ่งให้ภาพระดับสีเทา (256 ระดับ) และภาพนี้จะถูกแปลงให้เป็นเส้นขอบภาพด้วยเทคนิควิธีการประมวลผลหาเส้นขอบภาพ จากนั้น จะประมวลผลหาจุดเด่นจากเส้นขอบภาพนี้ด้วยวิธีการประมาณการรูปหลายเหลี่ยมแบบเชิงเส้น (Linear Polygonal Approximation) [1] และจุดเด่นเหล่านี้จะถูกนำไปเก็บไว้ในฐานข้อมูล เพื่อใช้เป็นข้อมูลในการจับคู่กับจุดเด่นของภาพวัตถุที่ทับซ้อนกันในขั้นตอนสุดท้าย นั่นคือ กระบวนการจับคู่ ซึ่งจะทำการประมวลผลจนสามารถจดจำภาพวัตถุได้

สิ่งที่ต้องคำนึงถึงในการประมวลผล การจดจำภาพวัตถุ คือ การเสถียรภาพต่อการเลื่อน/ย้ายตำแหน่ง, การหมุน และการย่อ/ขยายของภาพ นอกจากนี้ ในขั้นตอนของการจับคู่ ยังมีเงื่อนไขที่สำคัญ 2 ประการ คือ

- (1) จุดเด่นที่เกิดขึ้นบนเส้นขอบภาพ ที่มีวัตถุทับซ้อนกัน จะต้องจับคู่กับจุดเด่นในฐานข้อมูลได้เพียงจุดเดียวเท่านั้น

(2) จุดเด่นในฐานข้อมูล ก็จะต้องจับคู่กับจุดเด่นที่เกิดขึ้นบนเส้นขอบภาพที่มีวัตถุทับซ้อนกัน ได้เพียงจุดเดียวเช่นกัน

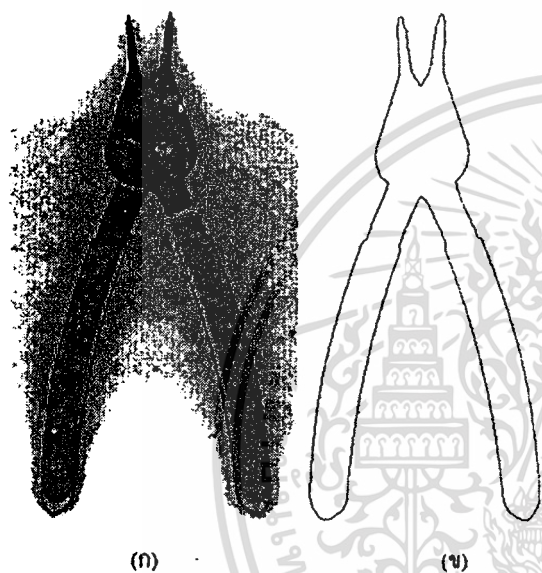
จากเงื่อนไขของการจับคู่ทั้งสองประการ ก่อให้เกิดปัญหาของการจดจำภาพวัตถุขึ้น เมื่อวัตถุมีการทับซ้อนหรือสัมผัสกัน ทั้งนี้ เนื่องจากการทับซ้อนกันของภาพวัตถุ ทำให้เกิดจุดเด่นจุดใหม่เพิ่มเติมขึ้นมา และในขณะที่เดียวกันจุดเด่นบางจุดก็จะถูกบดบังและสูญหายไป ทำให้ยากต่อการจับคู่ระหว่างจุดเด่นในฐานข้อมูลกับจุดเด่นของภาพวัตถุที่ทับซ้อนกัน

แต่ปัญหาดังกล่าวสามารถที่จะแก้ไขได้ด้วย หลักการของโครงข่ายเซลล์ประสาทเทียมแบบไฮโปลิเนอริ่ง ทั้งนี้ก็เพราะว่า ปัญหาที่เกิดขึ้นในกระบวนการจับคู่ ภายใต้เงื่อนไขการจับคู่แบบหนึ่งต่อหนึ่ง ซึ่งจะต้องคำนวณหาคู่ของจุดเด่นที่เหมาะสมที่สุดให้ได้ โดยอาศัยคุณสมบัติในการทำงานของไฮโปลิเนอริ่งที่มีลักษณะโคเคน คือ การพิจารณาข้อมูลก่อนการตัดสินใจจะนำข้อมูลข้างเคียงเข้ามาพิจารณาด้วย เพื่อให้ได้คำตอบที่ดีที่สุด โดยการบังคับให้สมการไฮโปลิเนอริ่งเข้าสู่ภาวะพลังงานต่ำสุด ด้วยการนำค่าพารามิเตอร์ที่บ่งบอกถึงคุณลักษณะของจุดเด่น มากกว่าหนึ่งค่าเข้ามาเป็นเงื่อนไขในการควบคุมไปพร้อมๆกัน ในเวลาเดียวกัน ซึ่งบทความนี้ได้ใช้พารามิเตอร์ทั้งหมด 4 ค่าด้วยกันในการควบคุมการจับคู่ของจุดเด่น ได้แก่ ค่าสเฟิซริจิตี [2], [5] ค่าความโค้ง, ทิศทางของความโค้ง [1] และค่ามุมของจุดเด่น [4] นอกจากนี้ การทำงานของโครงข่ายเซลล์ประสาทเทียมแบบไฮโปลิเนอริ่ง ยังตอบสนองต่อเงื่อนไขทั้ง 3 ประการของการจดจำภาพวัตถุ คือ เสถียรภาพต่อการเลื่อน/ย้าย, การหมุนและการย่อ/ขยายของภาพวัตถุอีกด้วย

สาระสำคัญของบทความนี้ จะประกอบไปด้วยหัวข้อต่างๆทั้งหมด 7 หัวข้อ ซึ่งจะขอกล่าวโดยสังเขปดังนี้ ในหัวข้อแรกนั้นจะเป็นบทนำ หัวข้อที่ 2 จะกล่าวถึงวิธีการคำนวณหาจุดเด่นของภาพวัตถุออกมา หัวข้อที่ 3 จะเป็นการนำเสนอ นิยามของ "Sphericity" ซึ่งเป็นทฤษฎีที่ถูกนำมาใช้ร่วมกับคณิตศาสตร์โครงข่ายเซลล์ประสาทเทียมแบบไฮโปลิเนอริ่งในหัวข้อที่ 4 ถัดจากนั้น ในหัวข้อที่ 5 จะอธิบายการประยุกต์ใช้งานโครงข่ายเซลล์ประสาทเทียมแบบไฮโปลิเนอริ่ง ในการประมวลผล การจดจำภาพวัตถุ ส่วนในหัวข้อที่ 6 และ 7 จะเป็นผลการทดลองและสรุปผลตามลำดับ

2. การประมวลผลหาจุดเด่น

ในการประมวลผลหาจุดเด่น จะเริ่มจากการแปลงข้อมูลของภาพที่เป็นภาพระดับสีเทา ดังแสดงในรูปที่ 2(ก) ให้เป็นเส้นขอบภาพเสียก่อน ด้วยเทคนิคการหาเส้นขอบภาพ ดังแสดงในรูปที่ 2 (ข) จากนั้น เส้นขอบภาพจึงจะถูกประมวลผลเพื่อหาจุดเด่นด้วยวิธีการ การประมาณการรูปหลายเหลี่ยมแบบเชิงเส้น [1] ซึ่งจะได้อัลกอริทึมของซิงเขปในหัวข้อที่ 2.1



รูปที่ 2 แสดงการแปลงภาพระดับสีเทาไปเป็นเส้นขอบภาพ

2.1 การหาค่าความโค้งบนเส้นโค้ง

ในการคำนวณหาค่าความโค้งบนเส้นโค้ง โดยวิธีต่อไปนี้จะต้องอยู่ภายใต้เงื่อนไขที่ว่า เส้นโค้งนั้นจะต้องเป็นเส้นโค้งปิด (Closed Contour) เท่านั้น และค่าความโค้งของเส้นโค้งใดๆสามารถคำนวณหาได้จากนิยามต่อไปนี้

$$k = \frac{|y''(x)|}{(1+(y'(x))^2)^{3/2}} \text{ เมื่อ } y'(x) = \frac{dy}{dx}$$

$$\text{และ } y''(x) = \frac{d^2y}{dx^2} \quad (2-1)$$

จากนิยามที่ได้กล่าวไว้ข้างบนนั้น สามารถที่จะแยกตัวแปร x และ y ออกมาอยู่ในรูปของสมการพารามตริก (Parametric form) ได้ คือ $x(t)$ และ $y(t)$ ตามลำดับ เมื่อ $0 < t < L$ และ L เป็นความยาวของเส้นโค้งปิด ดังนั้น จึงสามารถเขียนสมการ (2-1) ให้อยู่ในรูปสมการพารามตริกได้ดังนี้

$$y'(x) = \frac{\dot{y}(t)}{\dot{x}(t)} \text{ และ } y''(x) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t))^3} \quad (2-2)$$

$$\text{เมื่อ } \dot{x}(t) = \frac{dx}{dt} ; \dot{y}(t) = \frac{dy}{dt} \\ \ddot{x}(t) = \frac{d^2x}{dt^2} \text{ และ } \ddot{y}(t) = \frac{d^2y}{dt^2} \quad (2-3)$$

เมื่อแทนค่าสมการที่ (2-3) กลับไปยังสมการที่ (2-1) จะได้สมการหาค่าความโค้งในรูปของพารามตริกดังนี้

$$k(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \quad (2-4)$$

แต่ในการใช้งานจริง มักไม่นิยมที่จะหาค่าความโค้งจากสมการที่ (2-4) โดยตรง แต่จะใช้วิธีการประสาน (Convolution) ฟังก์ชัน $x(t)$ และ $y(t)$ ด้วยตัวกรองเกาส์เซียนแทน เพื่อเป็นการกำจัดสัญญาณรบกวนไปด้วยในตัว

$$\left. \begin{aligned} \dot{X}(t, \sigma) &= x(t) * \dot{g}(t, \sigma); \\ \dot{Y}(t, \sigma) &= y(t) * \dot{g}(t, \sigma); \\ \ddot{X}(t, \sigma) &= x(t) * \ddot{g}(t, \sigma); \\ \ddot{Y}(t, \sigma) &= y(t) * \ddot{g}(t, \sigma); \end{aligned} \right\} \quad (2-5)$$

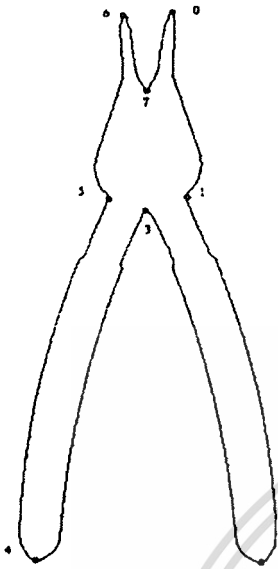
$$\left. \begin{aligned} \dot{g}(t, \sigma) &= \frac{\partial g(t, \sigma)}{\partial t}; \\ \ddot{g}(t, \sigma) &= \frac{\partial^2 g(t, \sigma)}{\partial t^2}; \end{aligned} \right\} \quad (2-6)$$

$$\text{และ } g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (2-7)$$

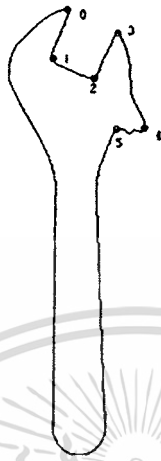
ดังนั้น จึงสามารถเขียนสมการที่ (2-4) ใหม่ได้เป็น

$$k(t, \sigma) = \frac{\dot{X}(t, \sigma)\ddot{Y}(t, \sigma) - \ddot{X}(t, \sigma)\dot{Y}(t, \sigma)}{(\dot{X}^2(t, \sigma) + \dot{Y}^2(t, \sigma))^{3/2}} \quad (2-8)$$

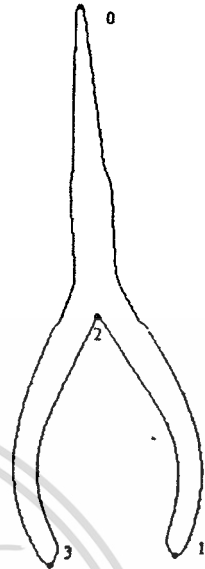
จุดเด่นที่ปรากฏในรูปที่ 3 จะถูกแทนด้วยเครื่องหมาย 'o' ซึ่งจุดเด่นเหล่านี้คำนวณได้จากสมการที่ (2-8) ด้วยวิธีการเปลี่ยนค่า สเกล-สเปซ โดยเริ่มจากค่าน้อยๆ (ประมาณ 3) แล้วเพิ่มขึ้นทีละ 2 จนกระทั่งถึงจุดอิ่มตัวและจุดเด่นยังคงเสถียรภาพอยู่



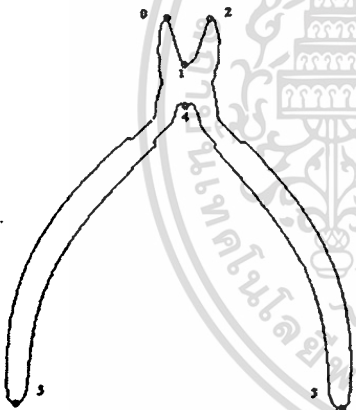
(ก) Model # 1



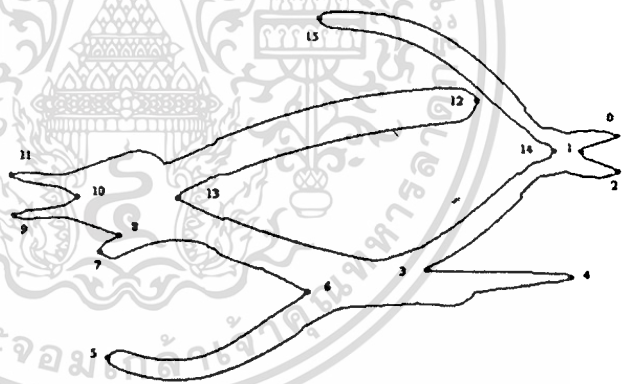
(ข) Model # 2



(ค) Model # 3

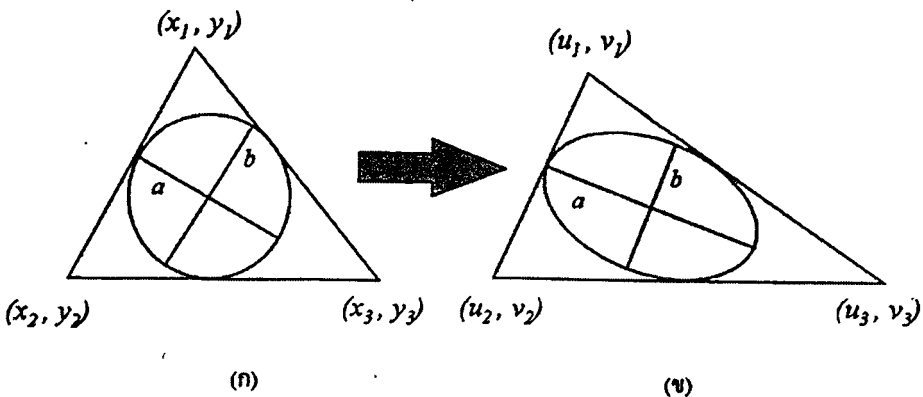


(ง) Model # 4



(จ)

รูปที่ 3 เป็นรูปเส้นโค้งปิดที่ผ่านการหาจุดเด่นแล้ว ซึ่งแสดงด้วยเครื่องหมาย '๐' และใช้ค่า $\sigma = 5 - 9$



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. นิยามของค่าสเฟียริซิตี (Sphericity)

สเฟียริซิตี ถูกนิยามไว้ว่า

$$S = \frac{2\sqrt{ab}}{a+b} \tag{3-1}$$

เมื่อ S คือ ค่าสเฟียริซิตี

a และ b คือ ความยาวแกนหลักของวงรี

จากนิยามที่กล่าวมานั้น สามารถนำเอาค่าของสเฟียริซิตี มาเป็นตัววัดความคล้ายหรือแตกต่างระหว่างรูปสามเหลี่ยมสองรูปใดๆ ได้ ดังที่ได้แสดงไว้ในรูปที่ 4 โดยสรุปได้ว่า ถ้าสามเหลี่ยมสองรูปใดๆ มีความคล้ายกันค่า สเฟียริซิตี จะเท่ากับ "1"

จากรูปที่ 4 สามารถพิจารณาความสัมพันธ์ของรูปสามเหลี่ยมรูป (ก) และ (ข) โดยอธิบายได้ด้วย ความสัมพันธ์ในการเปลี่ยนรูป (Affine Transform) [2], [5] จาก x ไปยัง u ซึ่งกำหนดให้

$$u = Ax + t \tag{3-2}$$

เมื่อ $x = \begin{bmatrix} x \\ y \end{bmatrix}$ $u = \begin{bmatrix} u \\ v \end{bmatrix}$ $t = \begin{bmatrix} e \\ f \end{bmatrix}$ $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

และ $\det(A) \neq 0$

ค่า ส.ป.ส. a, b, c, d, e และ f สามารถคำนวณได้จากสมการที่ (3-3)

$$\begin{bmatrix} a \\ b \\ e \end{bmatrix} = B^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ และ } \begin{bmatrix} c \\ d \\ f \end{bmatrix} = B^{-1} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{3-3}$$

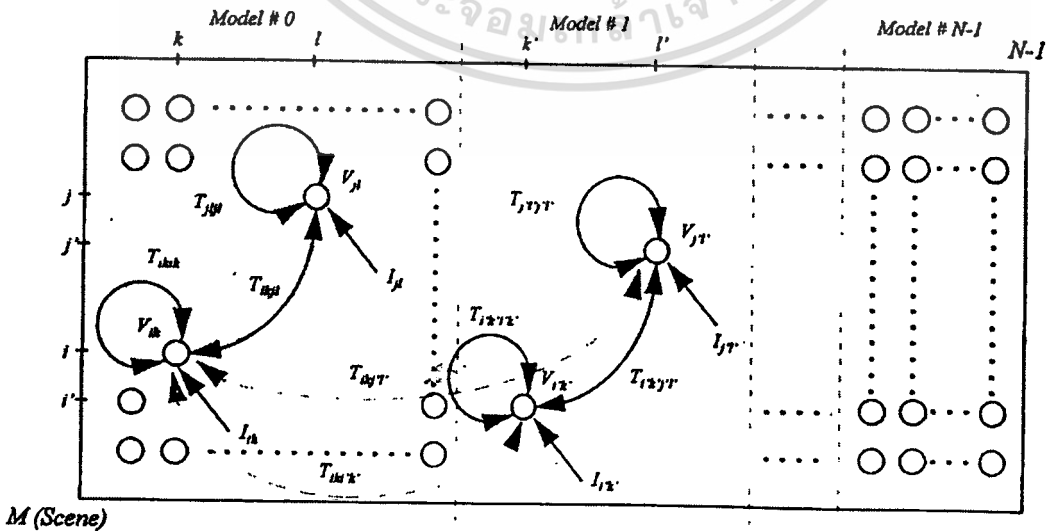
เมื่อ $B = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$ และ $\det(B) \neq 0$

ค่า (x_1, y_1) , (x_2, y_2) และ (x_3, y_3) คือจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุที่ทับซ้อนกัน เช่นเดียวกัน (u_1, v_1) , (u_2, v_2) และ (u_3, v_3) ก็คือจุดเด่นที่ปรากฏบนเส้นขอบภาพของวัตถุที่สมบูรณ์ซึ่งถูกเก็บไว้ในฐานข้อมูล

ถ้า $|A| > 0$ ในสมการที่ (3-2) เป็นจริง ก็สามารถคำนวณหาค่าสเฟียริซิตีได้ดังนี้

$$S = \frac{t_1^2 + t_3^2 - (t_2^2 + t_4^2)}{(t_1^2 + t_2^2 + t_3^2 + t_4^2)} \tag{3-4}$$

เมื่อ $t_1 = a+d$; $t_2 = a-d$; $t_3 = b-c$; $t_4 = b+c$



รูปที่ 5 โครงสร้างของโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ ในรูปแบบสองมิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. คณิตศาสตร์โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

หลักการทํางานของโครงข่ายเซลล์ประสาทเทียม คือ การสร้างวิธีการตัดสินใจร่วมกัน โดยขึ้นอยู่กับพื้นฐานของ ข้อมูลที่ได้รับเข้ามาในแต่ละเซลล์ประสาทหรือๆกัน และมีการส่งผ่านข้อมูลระหว่างเซลล์ประสาท พร้อมกับบังคับให้ โครงข่ายเซลล์ประสาท อยู่เข้าสู่สภาวะคงที่หรือเสถียรภาพ (Stable State) เพื่อที่จะได้ตัดสินใจกับข้อมูลที่ได้รับเข้ามานั้น

จากรูปที่ 5 จะเห็นว่าโครงข่ายเซลล์ประสาทเทียม มี การวางตัวของแต่ละเซลล์ประสาทอยู่ในรูปเมตริก (Matrix) ซึ่งมีขนาดเท่ากับ $M \times N$ แต่ละเซลล์ประสาทสามารถมีความ สัมพันธ์กับเซลล์ประสาทอื่นๆได้ ภายใต้ความเชื่อมโยงของ T_{ijkl} ดังรูปที่ 5 ขณะเดียวกันแต่ละเซลล์ประสาทก็จะรับเอา ข้อมูลจากภายนอกเข้ามา (I_{ik}) ได้ด้วย

ส่วนสมการค้นแบบของโฮปฟิลด์ ถูกนิยามไว้ดังนี้

$$E = -\frac{1}{2} \sum_{x=1}^n \sum_{y=1}^n T_{xy} V_x V_y - \sum_{x=1}^n I_x V_x \quad (4-1)$$

เมื่อกำหนดให้ $x \Rightarrow i, k$ และ $y \Rightarrow j, l$ เพื่อให้สอดคล้องกับลักษณะของงาน ซึ่งใช้ซับซ้อน 2 ตัว ดัง แสดงไว้ในรูปที่ 5 ทำให้สามารถเขียนสมการที่ (4-1) ใหม่ได้ เป็นสมการที่ (4-2)

$$E = -\frac{1}{2} \sum_{i,k} \sum_{j,l} T_{ijkl} V_{ik} V_{jl} - \sum_{i,k} I_{ik} V_{ik} \quad (4-2)$$

เขียนสมการที่ (4-2) เลียบใหม่ ให้สมบูรณ์ได้เป็น

$$E = -\frac{1}{2} \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{ik} V_{jl} - \sum_{i=1}^M \sum_{k=1}^N I_{ik} V_{ik} \quad (4-3)$$

เมื่อ V_{ik} และ V_{jl} คือ สถานะของเซลล์ประสาท ณ ตำแหน่ง ik และ jl ตามลำดับ ซึ่งจะมีค่าเป็น "0" หรือ "1" เท่านั้น

T_{ijkl} คือ ความสัมพันธ์ที่เชื่อมโยงถึงกันของเซลล์ประสาท ik กับ jl

I_{ik} และ I_{jl} คือ ข้อมูลที่ต้องป้อนให้กับเซลล์ประสาทในสภาวะเริ่มต้น

ค่าตอบของสมการที่ (4-3) จะปรากฏ ก็คือเมื่อ สมการได้ปรับตัวเองเข้าสู่สภาวะพลังงานต่ำสุดแล้ว ค่าตอบที่ ต้องการนี้ก็จะปรากฏอยู่ในแต่ละเซลล์ประสาทในโครงข่ายนั้น

เอง แต่การที่สมการจะปรับตัวเองเพื่อเข้าสู่สภาวะพลังงานต่ำสุดได้นั้น เซลล์ประสาท V_{ik} จะต้องมีการเปลี่ยนแปลงสถานะตามสมการต่อไปนี้

$$V_{ik} \rightarrow 0 \quad \text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) < 0 \quad (4-4)$$

$$V_{ik} \rightarrow 1 \quad \text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) > 0 \quad (4-5)$$

$V_{ik} \rightarrow \text{no change}$

$$\text{ถ้า} \quad \left(\sum_{j=1}^M \sum_{l=1}^N T_{ijkl} V_{jl} + I_{ik} \right) = 0 \quad (4-6)$$

ซึ่งการเปลี่ยนแปลงสถานะของ V_{ik} ใดๆนั้น เป็นการปรับค่าพลังงานเฉพาะเซลล์ประสาทนั้นเท่านั้น ดังนั้น สมการย่อยของการเปลี่ยนแปลงสถานะของ V_{ik} จึงเป็นการกำหนดให้ i กับ k เป็นค่าคงที่ คงเหลือแค่ j กับ l ที่เป็นตัวแปร จึงได้สมการย่อยของการลดค่าพลังงาน ดังสมการที่ (4-4), (4-5) และ (4-6) ตามลำดับ

ส่วนขั้นตอนของการนำไปใช้งาน จะได้อธิบายรายละเอียดในหัวข้อถัดไป

5. การประมวลผลการจดจำภาพวัตถุ

5.1 การประมวลผลด้วยโครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์

การประมวลผลการจดจำภาพวัตถุ คือ การวิเคราะห์ เพื่อจดจำภาพและคืนส่วนของข้อมูลที่สูญหายไป ดังที่ได้แสดงไว้ในรูปที่ 7 และ 8

ดังที่ได้กล่าวไว้ในตอนต้นแล้วว่า ในการประมวลผลการจดจำภาพวัตถุนั้น มีอยู่ด้วยกัน 2 ขั้นตอนหลักๆ คือ การคำนวณหาจุดเด่นของเส้นขอบภาพวัตถุออกมา และการจับคู่กันระหว่าง จุดเด่นที่เกิดขึ้นบนเส้นขอบภาพวัตถุที่ทับซ้อนและบดบังบางส่วน กับภาพวัตถุที่สมบูรณ์ซึ่งเก็บเอาไว้ในฐานข้อมูล ในขั้นตอนของการคำนวณหาจุดเด่นของเส้นขอบภาพออกมา ได้แสดงไว้ในรูปที่ 3 เรียบร้อยแล้ว ด้วยสมการคณิตศาสตร์ที่ได้อธิบายไว้ในหัวข้อ 2.1

ส่วนขั้นตอนที่สอง ซึ่งเป็นกระบวนการจับคู่ (Matching) นั้น จะได้อธิบายเป็นลำดับขั้นตอนดังนี้

สมการที่ใช้ในการจับคู่ หรือบางครั้งมักจะถูกเรียกว่า "ฟังก์ชันพลังงาน" (Energy Function) ถูกนิยามไว้ดังนี้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$e = -\sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N C_{ijkl} V_{ik} V_{jl} + \sum_{i=1}^M \left(1 - \sum_{k=1}^N V_{ik}\right)^2 + \sum_{k=1}^N \left(1 - \sum_{i=1}^M V_{ik}\right)^2 \tag{5-1}$$

เมื่อนิพจน์แรกของสมการ (5-1) ทำหน้าที่ในการจับคู่ระหว่าง i กับ k โดยมี C_{ijkl} เป็นตัวควบคุมของการจับคู่ และ V_{jl} เป็นข้อมูลข้างเคียงประกอบการพิจารณา ส่วนนิพจน์ที่ 2 และ 3 จะทำหน้าที่บังคับให้ค่าตอบที่ได้นั้น จะต้องไม่มีเพียงค่าตอบเดียวและถูกต้องด้วย (เป็นเงื่อนไขในการจับคู่แบบหนึ่งต่อหนึ่งนั่นเอง)

จากสมการที่ (5-1) สามารถลดรูป เพื่อให้สามารถเทียบเคียงกับสมการของไฮปอฟิลด์ (4-3) ได้ดังนี้

$$\sum_{i=1}^M \left(1 - \sum_{k=1}^N V_{ik}\right)^2 = \sum_{i=1}^M \left(1 - 2 \sum_{k=1}^N V_{ik} + \left(\sum_{k=1}^N V_{ik}\right)^2\right) \tag{5-2a}$$

$$= \sum_{i=1}^M (1) - 2 \sum_{i=1}^M \sum_{k=1}^N V_{ik} + \sum_{i=1}^M \left(\sum_{k=1}^N \sum_{l=1}^N V_{ik} V_{il}\right)$$

$$= M - 2 \sum_{i=1}^M \sum_{k=1}^N V_{ik} + \sum_{i=1}^M \sum_{k=1}^N \sum_{l=1}^N V_{ik} V_{il}$$

ถ้ากำหนดให้ $V_{ik} V_{il} = V_{ik} V_{jl} \delta_{ij}$ จะได้ว่า

$$= M + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{ij} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \tag{5-2b}$$

และในทำนองเดียวกัน

$$\sum_{k=1}^N \left(1 - \sum_{i=1}^M V_{ik}\right)^2 = N + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{kl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \tag{5-2c}$$

แทนค่าสมการที่ (5-2b) และ (5-2c) กลับไปยังสมการที่ (5-1) จะได้ว่า

$$e = M + N - \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N C_{ijkl} V_{ik} V_{jl} + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{ij} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} + \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N V_{ik} V_{jl} \delta_{kl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \tag{5-2d}$$

$$e = M + N - \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N (C_{ijkl} - \delta_{ij} - \delta_{kl}) V_{ik} V_{jl} - \sum_{i=1}^M \sum_{k=1}^N 4V_{ik} \tag{5-2c}$$

$$\frac{e - (M + N)}{2} = -\frac{1}{2} \sum_{i=1}^M \sum_{k=1}^N \sum_{j=1}^M \sum_{l=1}^N (C_{ijkl} - \delta_{ij} - \delta_{kl}) V_{ik} V_{jl} - \sum_{i=1}^M \sum_{k=1}^N 2V_{ik} \tag{5-3}$$

ซึ่งจะเห็นได้ว่า สมการที่ (5-3) สมมูลกับสมการที่ (4-3) ทุกประการ เมื่อ $T_{ijkl} = C_{ijkl} - \delta_{ij} - \delta_{kl}$ และ $I_{ik} = 2$

ถึงขั้นตอนนี้ ก็สามารถดำเนินการตามกรรมวิธีของไฮปอฟิลด์ได้แล้ว โดยเริ่มจากลำดับขั้นตอนที่ 1 ถึง 5

ลำดับขั้นที่ 1 ให้กำหนดสถานะเริ่มต้นให้กับ V_{ik} (เซลล์ประสาท) ในโครงข่ายเซลล์ประสาทเทียมแบบไฮปอฟิลด์ ซึ่งมีขนาด $M \times N$ ภายใต้เงื่อนไขของสมการ (5-4) (เมื่อ M คือ จำนวนของจุดเด่นที่คำนวณได้จากภาพวัตถุที่ถูกทับซ้อน และ N คือ จำนวนจุดเด่นทั้งหมดของภาพที่เก็บไว้ในฐานข้อมูล)

$$V_{ik} = \begin{cases} 1 & \text{if Sphericity} \geq \alpha \\ 0 & \text{Otherwise} \end{cases} \tag{5-4}$$

เมื่อ α คือ ค่า Threshold ของสเฟิริซิตี และมีค่าอยู่ระหว่าง $0 \leq \alpha \leq 1$

ในรูปที่ 6 ได้แสดงให้เห็นการจับคู่ของจุดเด่นในรูปที่ 3 (ก), (ข), (ค) และ (ง) ซึ่งเก็บไว้เป็นฐานข้อมูล กับจุดเด่นของรูปที่ 7 ซึ่งเป็นภาพที่ถูกบดบังบางส่วน และมีจุดเด่นดังแสดงไว้ในรูปที่ 3 (จ) โดยจะกำหนดให้ V_{ik} มีค่าเท่ากับ "1" หรือ "0" ตามเงื่อนไขของสมการที่ (5-4) ส่วนค่าของสเฟิริซิตีที่คำนวณได้จากสมการที่ (3-4) ในหัวข้อที่ 3 $V_{ik} = 1$ หมายความว่าจุดเด่น ณ ตำแหน่ง i ใน Scene (รูปที่ 3 (จ)) สามารถจับคู่กับจุดเด่นในตำแหน่ง k ใน Model (รูปที่ 3 (ก), (ข), (ค) และ (ง)) ได้ ซึ่งจะเห็นว่า ในสภาวะเริ่มต้นของการกำหนดด้วยสมการที่ (5-4) ยังมีการจับคู่ที่ไม่เหมาะสม (Mismatch) อยู่ ดังรูปที่ 6 (ก) ซึ่งการจับคู่ที่ไม่เหมาะสมเหล่านี้ จะถูกกำจัดทิ้งไปด้วยสมการของไฮปอฟิลด์ในลำดับขั้นตอนที่ 2 ถึง 4

แต่ถ้าพิจารณาพฤติกรรมในการจับคู่ของจุดเด่นให้ดีแล้วจะพบว่า จุดเด่นที่ควรจับคู่กันได้นั้น มักจะอยู่เป็นกลุ่ม

ในแนว -45 องศา ของเมตริก $M \times N$ ในแต่ละ Model ดังนั้น จากพฤติกรรมดังกล่าว จึงสามารถจัดการจับคู่ที่ไม่เหมาะสม ออกไปได้บางส่วน ซึ่งได้แสดงไว้ในรูปที่ 6 (ข) และเรียกขั้นตอนนี้ว่า "การกรองเบื้องต้น" (Pre-filtering)

การกรองเบื้องต้นเป็นการกำหนดสถานะเริ่มต้น ให้กับโครงข่ายเซลล์ประสาทเทียม เพื่อเป็นการจัดการจับคู่ที่ไม่เหมาะสมที่ได้จากสมการที่ (5-4) อีกครั้งหนึ่ง ซึ่งจะให้ผลดี สองประการ คือ เพิ่มความแม่นยำให้กับการทำงานของโครงข่ายเซลล์ประสาท และช่วยให้จำนวนครั้งของการทำงานซ้ำ (Iteration) ของโครงข่ายลดลง

ลำดับขั้นที่ 2 ให้เลือกตำแหน่งเซลล์ประสาท (i,k) ใดๆ ขึ้นมา แล้วคำนวณหาค่า U_{ik} ตามสมการที่ (5-5)

$$U_{ik} = \sum_{j=1}^M \sum_{l=1}^N (C_{ijkl} - \delta_{ij} - \delta_{kl}) V_{jl} + 2 \quad (5-5)$$

ซึ่งสมการที่ (5-5) ได้มาจากสมการที่ (4-4),(4-5)

และ (4-6) นั่นเอง โดยกำหนดให้

$$T_{ijkl} = C_{ijkl} - \delta_{ij} - \delta_{kl} \text{ และ } I_{ik} = 2$$

เมื่อ δ_{ij} และ δ_{kl} เป็น Kronecker delta function

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{Otherwise} \end{cases} \quad (5-5a)$$

$$\delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{Otherwise} \end{cases} \quad (5-5b)$$

และ

$$C_{ijkl} = \begin{cases} W_1 F_1 + W_2 F_2 + W_3 F_3 & \text{for } k \text{ and } l \\ \in \text{ the same model} \\ 0 & \text{for } k \text{ and } l \notin \text{ the same model} \end{cases} \quad (5-5c)$$

$$F_1 = \begin{cases} 1 & \text{if } |Co_m - Co_s| < \beta \\ -1 & \text{Otherwise} \end{cases} \quad (5-5d)$$

เมื่อ Co_m, Co_s และ β อธิบายไว้ในภาคผนวก (ก)

$$F_2 = \begin{cases} 1 & \text{if } \left| \frac{Cu_m - Cu_s}{|Cu_m| + |Cu_s|} \right| < \lambda \\ -1 & \text{Otherwise} \end{cases} \quad (5-5e)$$

เมื่อ Cu_m, Cu_s และ λ อธิบายไว้ในภาคผนวก (ข)

$$F_3 = \begin{cases} 1 & \text{if } (h_m \text{ and } h_s) \geq 0 \\ -1 & \text{or } (h_m \text{ and } h_s) < 0 \\ & \text{Otherwise} \end{cases} \quad (5-5f)$$

เมื่อ h_m และ h_s อธิบายไว้ในภาคผนวก (ค)

$$W_1 + W_2 + W_3 = 1$$

ซึ่งรายละเอียดของ F_1, F_2 และ F_3 ได้อธิบายไว้ในส่วนภาคผนวก ส่วนค่า W_1, W_2 , และ W_3 จะเป็นน้ำหนัก (Weigh) ที่ใช้วัดความสำคัญของ F_1, F_2 และ F_3 ตามลำดับ

ลำดับขั้นที่ 3 ให้เปลี่ยนสถานะของ V_{ik} ตามสมการข้างล่างนี้

$$\begin{aligned} V_{ik} &\rightarrow 0 && \text{ถ้า } U_{ik} > 2 \\ V_{ik} &\rightarrow 1 && \text{ถ้า } U_{ik} < -2 \\ V_{ik} &\rightarrow \text{no change} && \text{ถ้า } -2 \leq U_{ik} \leq 2 \end{aligned}$$

เมื่อค่าตัวเลข "2" คือค่า Threshold ที่ได้จากการทดลอง

ลำดับขั้นที่ 4 ให้ตรวจสอบทุกค่า V_{ik} ว่ามีการเปลี่ยนแปลงสถานะอยู่หรือไม่ ถ้ายังมีการเปลี่ยนแปลงอยู่ ให้กลับไปดำเนินการตาม ลำดับขั้นที่ 2 ใหม่ แต่ถ้าไม่มีการเปลี่ยนแปลง แสดงว่า V_{ik} ได้เข้าสู่สถานะพลังงานต่ำสุดแล้ว ให้ดำเนินการต่อไป ลำดับขั้นที่ 5

ลำดับขั้นที่ 5 ลำดับขั้นนี้จะเป็นคำตอบสุดท้ายของ V_{ik} ที่มีการจับคู่ระหว่างจุดเด่นในฐานข้อมูลกับจุดเด่นของภาพ ในรูปที่ 3 (จ) ซึ่งจะได้ผลดังรูปที่ 6 (ค)

5.2 การคำนวณค่าผิดพลาดในการจับคู่ของจุดเด่น

ค่าความผิดพลาดอันเนื่องมาจาก การเลื่อน/ย้ายตำแหน่ง การหมุน และการย่อ/ขยายของเส้นขอบภาพในฐานข้อมูล (Model) เพื่อบอกตำแหน่งของภาพวัตถุที่ทับซ้อนกัน สามารถคำนวณได้จาก ค่าความผิดพลาดกำลังสองที่น้อยที่สุด (Least Square Error) ได้ดังนี้

$$\varepsilon = \sum_{i=0}^{m-1} \left((x_s - x_o)^2 + (y_s - y_o)^2 \right) \quad (5-6)$$

เมื่อ ε คือ ค่าความผิดพลาดกำลังสองที่น้อยที่สุด

m คือ จำนวนของจุดเด่นที่สามารถจับคู่ได้

x_{sl}, y_{sl} คือ ตำแหน่งของจุดเด่นในแนวแกน x และ y ของภาพวัตถุที่ทับซ้อนกัน (Scene)

x_{ci} y_{ci} คือ ค่าที่คำนวณได้จากการแปลงคอดโคออดิเนทของจุด
เด่นในฐานข้อมูล ไปยังตำแหน่ง x_{si} และ y_{si} ตามลำดับ

โดยค่า ε ที่คำนวณได้ มีค่าขึ้นอยู่กับ

ถึงแม้ว่า ในกระบวนการการจับคู่ของจุดเด่น โดยวิธี
ใช้โครงข่ายเซลล์ประสาทเทียมแบบโฮปฟิลด์ จะเป็นวิธีการ
แก้ไขปัญหามีประสิทธิภาพ แต่เพื่อความแน่นอนในกระบวนการ
การจับคู่ ว่ามีความถูกต้องเพียงใด จึงจำเป็นที่จะต้องมีการวัด
ความผิดพลาดในการจับคู่ของจุดเด่น โดยใช้สมการต่อไปนี้

$$\varepsilon' = \begin{cases} \left(1 + \frac{(n-2)}{(m-2)} \log_2 \left(\frac{n-2}{m-2}\right)\right) \bar{\varepsilon} & \text{for } m \geq 3 \\ \infty & \text{for } m = 0, 1, 2 \end{cases} \quad (5-7)$$

เมื่อ ε' คือ ค่าความผิดพลาดในการจับคู่

n คือ จำนวนจุดเด่นทั้งหมด ในแต่ละ Model

m คือ จำนวนจุดเด่นที่สามารถจับคู่ได้

$\bar{\varepsilon}$ คือ Normalized Least Square Error

โดยที่ $\bar{\varepsilon} = \frac{\varepsilon}{k}$

เมื่อ k คือ อัตราข้อ/ขยายของภาพ

ซึ่งผลที่คำนวณได้จากสมการที่ (5-7) ได้แสดงไว้ใน

ตารางที่ 1 และ 2

N (Model)	Model = 1	Model = 2	Model = 3	Model = 4
0	00000000	000000	0000	100100
1	01000000	001000	0000	010000
2	00000000	000000	0000	001001
3	10001000	000000	0000	000000
4	00000000	000000	1000	000000
5	00000000	100000	0100	000001
6	00000000	000000	0010	000101
7	00000000	000000	0001	000000
8	00001100	000001	0000	010000
9	00100010	000001	0000	000000
10	00000001	000000	0000	000000
11	10001000	000001	0000	000000
12	00100000	000000	0000	000000
13	00010000	001101	0000	000000
14	01000010	000000	0000	000010
15	00001000	011000	0000	000011

M (Scene)

รูปที่ 6 (ก) สถานะเริ่มต้นของ V_{ik} ในโครงข่าย
ของรูปที่ 3 (จ)

N (Model)	Model = 1	Model = 2	Model = 3	Model = 4
0	00000000	000000	0000	100000
1	00000000	000000	0000	010000
2	00000000	000000	0000	001000
3	00000000	000000	0000	000000
4	00000000	000000	1000	000000
5	00000000	000000	0100	000000
6	00000000	000000	0010	000000
7	00000000	000000	0001	000000
8	00000100	000000	0000	000000
9	00000010	000000	0000	000000
10	00000001	000000	0000	000000
11	10000000	000000	0000	000000
12	00100000	000000	0000	000000
13	00010000	000000	0000	000000
14	00000000	000000	0000	000010
15	00000000	000000	0000	000001

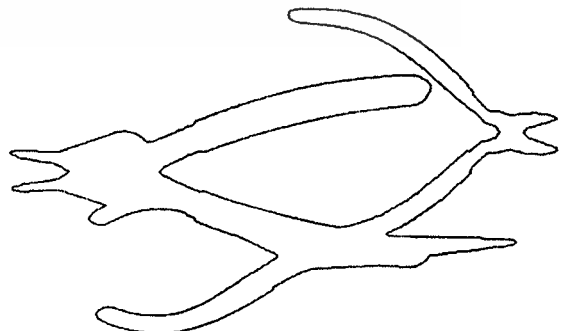
M (Scene)

รูปที่ 6 (ข) ผลที่ได้จาก “การกรองเบื้องต้น” ของรูปที่ 3 (จ)

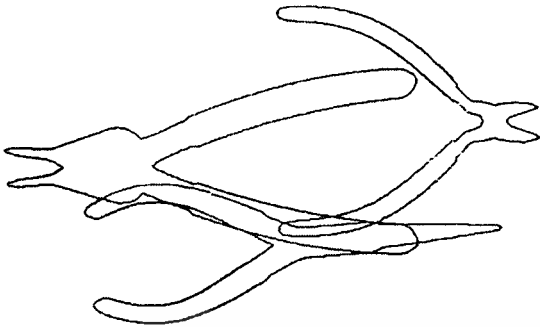
N (Model)	Model = 1	Model = 2	Model = 3	Model = 4
0	00000000	000000	0000	100000
1	00000000	000000	0000	010000
2	00000000	000000	0000	001000
3	00000000	000000	0000	000000
4	00000000	000000	1000	000000
5	00000000	000000	0100	000000
6	00000000	000000	0010	000000
7	00000000	000000	0001	000000
8	00000000	000000	0000	000000
9	00000010	000000	0000	000000
10	00000001	000000	0000	000000
11	10000000	000000	0000	000000
12	00100000	000000	0000	000000
13	00010000	000000	0000	000000
14	00000000	000000	0000	000010
15	00000000	000000	0000	000001

M (Scene)

รูปที่ 6 (ค) ค่าคอบซุคท้ายของการจับคู่ของรูปที่ 3 (จ)



รูปที่ 7 แสดงถึงภาพวัตถุที่มีการทับซ้อน
และถูกบดบังบางส่วน



รูปที่ 8 แสดงให้เห็นถึงการจดจำภาพวัตถุและคืนภาพส่วนที่ถูกบดบังบางส่วน

ตารางที่ 1 แสดงให้เห็นค่าความผิดพลาดในการจับคู่ระหว่างรูปในฐานะข้อมูลกับรูปที่ 8 (๑)

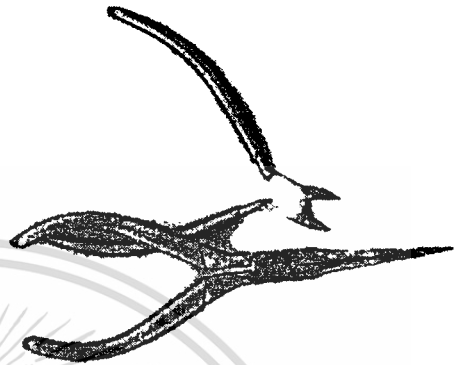
รูปในฐานะข้อมูล	จำนวนจุดเด่นรวมของแต่ละรูปในฐานะข้อมูล	จำนวนจุดเด่นที่สามารถจับคู่ได้	ϵ'
Model#1*	8	5	0.0000
Model#2	6	0	∞
Model#3*	4	4	0.8161
Model#4*	6	5	4.3442

หมายเหตุ : * เป็นรูปในฐานะข้อมูลที่สามารถจดจำได้ (ในตารางที่ 2 ก็มีความหมายเช่นเดียวกัน)

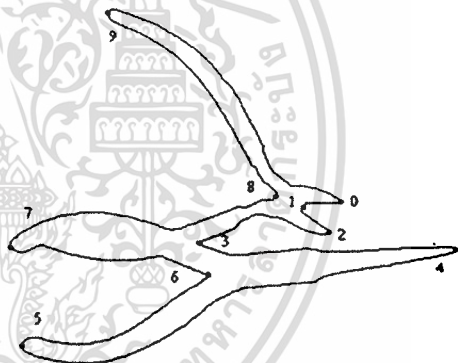
6. ผลการทดลอง

ผลการทดลอง การประมวลผลการจดจำภาพวัตถุด้วยวิธีของโครงข่ายเซลล์ประสาทเทียมแบบไฮปัสลิด ได้ผลดังแสดงไว้ในรูปที่ 8 ซึ่งเป็นภาพที่ได้จากการจดจำของรูปวัตถุ 3 ชิ้นที่วางทับซ้อนกัน จากจุดเด่นที่ได้จากรูปที่ 3 (๑) (เป็นภาพสังเคราะห์ขึ้น) โดยในการทดลอง ได้กำหนดค่า σ เพื่อปรับระดับประมาณ 0.97-0.99 ในการจับคู่และกำหนดเป็นสภาวะเริ่มต้นให้กับโครงข่ายเซลล์ประสาท ส่วนค่าพารามิเตอร์อีก 2 ค่า ได้กำหนดให้ β ใช้ค่าประมาณ 5-7 องศา และ λ ใช้ค่าประมาณ 0.015-0.020 พร้อมกับให้ค่าน้ำหนัก $W_1=0.6$ $W_2=0.2$ และ $W_3=0.2$ ซึ่งได้ผลการจับคู่ที่เป็นค่าขอบสุดท้ายดังในรูปที่ 6 (ค) นอกจากนี้ ยังได้คำนวณหาค่าความผิดพลาดในการจับคู่ (ϵ') เพื่อพิสูจน์ผลของการจับคู่อีกครั้งหนึ่ง ว่ามีความถูกต้องเพียงใด โดยค่า ϵ' ที่คำนวณได้มี ถ้ามีค่าน้อยๆ แสดงว่า มีความถูกต้องในการจับคู่สูง ถ้ามีค่ามาก แสดงว่า ในการจับคู่อาจจะผิดพลาดได้ และถ้าเป็นค่า อนันต์ นั้นหมายถึง การจับคู่ไม่ได้เลย หรือจับคู่ได้น้อยกว่า 2 จุด

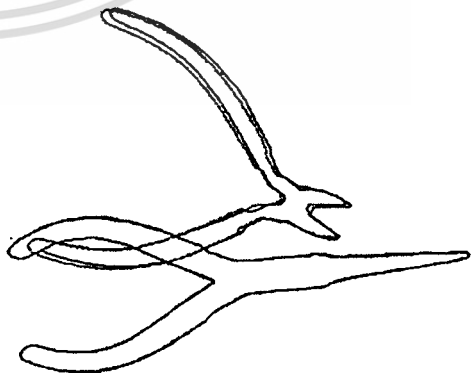
นอกจากนี้ ยังได้ทดลองกับภาพจริงที่ได้จากกล้อง CCD ซึ่งได้ผลการทดลองดังที่ได้แสดงไว้ในรูปที่ 11 โดยมีค่า ϵ' ดังที่ได้สรุปไว้ในตารางที่ 2



รูปที่ 9 เป็นรูปทึบสองอันที่วางทับซ้อนกัน ซึ่งถ่ายภาพผ่านกล้อง CCD



รูปที่ 10 แสดงถึงตำแหน่งของจุดเด่นบนเส้นขอบภาพที่วัตถุวางทับซ้อนกัน (ใช้ค่า $\sigma = 5-7$)



รูปที่ 11 แสดงให้เห็นถึงการจดจำภาพวัตถุและคืนส่วนที่ถูกทับซ้อนของวัตถุ

N (Model) Model # 1 Model # 2 Model # 3 Model # 4

M \ N	01234567	012345	0123	012345
0	00000000	000000	0000	100100
1	00000000	000000	0000	010000
2	00000000	000000	0000	001001
3	00000000	000000	0000	000000
4	00000000	000000	1000	000010
5	00000000	000000	0100	000000
6	00000000	000000	0010	000000
7	00000000	000000	0001	000000
8	00000000	000000	0000	000010
9	00000000	000000	0000	000101

M (Scene)

รูปที่ 12 แสดงให้เห็นค่าสถานะเริ่มต้นของ V_{ik} ภายในโครงข่ายเซลล์ประสาทแบบไฮโปฟิลด์ของรูปที่ 10

N (Model) Model # 1 Model # 2 Model # 3 Model # 4

M \ N	01234567	012345	0123	012345
0	00000000	000000	0000	100000
1	00000000	000000	0000	010000
2	00000000	000000	0000	001000
3	00000000	000000	0000	000000
4	00000000	000000	1000	000000
5	00000000	000000	0100	000000
6	00000000	000000	0010	000000
7	00000000	000000	0001	000000
8	00000000	000000	0000	000010
9	00000000	000000	0000	000001

M (Scene)

รูปที่ 13 สถานะสุดท้ายของโครงข่ายเซลล์ประสาทแบบไฮโปฟิลด์เมื่ออยู่ในสถานะพลังงานต่ำสุดของรูปที่ 10

ตารางที่ 2 แสดงให้เห็นค่าความผิดพลาดในการจับคู่ระหว่างรูปในฐานะข้อมูลกับรูปที่ 10

รูปในฐานะข้อมูล	จำนวนจุดเด่นรวมของแต่ละรูปในฐานะข้อมูล	จำนวนจุดเด่นที่สามารถจับคู่ได้	ϵ'
Model#1	8	0	∞
Model#2	6	0	∞
Model#3*	4	4	3.9044
Model#4*	6	5	25.8680

7. สรุปผล

โครงข่ายเซลล์ประสาทเทียมแบบไฮโปฟิลด์ เป็นวิธีการหนึ่งที่มีประสิทธิภาพ ในการนำมาประยุกต์ใช้งานกับการประมวลผลการจดจำภาพวัตถุ ทั้งนี้ก็สืบเนื่องมาจากความสามารถในการตอบสนองต่อเงื่อนไขหลัก ของการประมวลผล

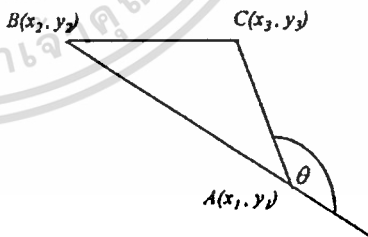
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษา เมื่อ i และ j เป็น unit vector ดังนั้น

การจดจำภาพวัตถุ ซึ่งได้แก่ การเสถียรภาพต่อการเลื่อนหรือย้ายตำแหน่ง การหมุน และการย่อ/ขยายภาพ นอกจากนี้ การใช้ค่า 4 พารามิเตอร์ (สเฟียร์ริซิตี, ความโค้ง, ทิศทางของความโค้ง และมุมของจุดเด่น) ในการบ่งบอกคุณลักษณะของจุดเด่นภายใต้ตัวแปร C_{ijkl} ของสมการไฮโปฟิลด์ ก็เพื่อให้ผลของการจับคู่มีความเหมาะสมมากที่สุด และสิ่งสำคัญที่ต้องคำนึงถึงคือ พารามิเตอร์ที่นำมาใช้นั้น จะต้องตอบสนองต่อเงื่อนไขทุกประการของการประมวลผลการจดจำภาพวัตถุได้ด้วย แต่ทั้งนี้ทั้งนั้น การที่จะทำให้โครงข่ายเซลล์ประสาทแบบไฮโปฟิลด์สามารถทำงานและแก้ไขปัญหาได้อย่างมีประสิทธิภาพได้นั้น จำเป็นอย่างยิ่งที่จะต้องคำนึงถึง ปัจจัยที่จะนำมาเป็นข้อมูลให้กับสมการของไฮโปฟิลด์ด้วย ปัจจัยเหล่านี้ได้แก่ การจัดแสงให้กับกล้องถ่ายภาพแบบ CCD เพื่อให้ได้ภาพที่สมบูรณ์ จะได้เป็นข้อมูลที่ถูกต้องในการประมวลผล นอกจากนั้น การบวนการคำนวณหาดีเอ็นเอจุดเด่นออกมา ก็เป็นปัจจัยที่สำคัญเช่นกัน ตำแหน่งของจุดเด่นเหล่านั้น จะต้องเที่ยงตรงถูกต้องและเสถียรภาพ ถ้าสามารถควบคุมปัจจัยหลักเหล่านี้ได้แล้ว จะได้ผลลัพธ์ที่ถูกต้องและแม่นยำ

8. ภาคผนวก

(ก) มุมของจุดเด่น

จากรูป มุม θ สามารถคำนวณได้จาก



$$\overline{AB} \cdot \overline{BC} = |\overline{AB}| |\overline{BC}| \cos \theta \quad (ก-1)$$

$$\cos \theta = \frac{\overline{AB} \cdot \overline{BC}}{|\overline{AB}| |\overline{BC}|} \quad (ก-2)$$

เวกเตอร์ \overline{AB} และ \overline{BC} คำนวณได้จาก

$$\overline{AB} = (x_2 - x_1)i + (y_2 - y_1)j \quad (ก-3)$$

$$\text{และ } \overline{BC} = (x_3 - x_2)i + (y_3 - y_2)j \quad (ก-4)$$

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\overline{AB} \cdot \overline{BC} = \frac{(x_3 - x_2)(x_2 - x_1) + (y_3 - y_2)(y_2 - y_1)}{\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \quad (n-5)$$

และ

$$\cos \theta = \frac{(x_3 - x_2)(x_2 - x_1) + (y_3 - y_2)(y_2 - y_1)}{\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \quad (n-6)$$

$$F_1 = \begin{cases} 1 & \text{if } |Co_m - Co_s| < \beta \\ -1 & \text{Otherwise} \end{cases}$$

เมื่อ Co_m คือ มุมของจุดเด่นที่ปรากฏในฐานข้อมูล (Model) รูปที่ 3 (ก), (ข), (ค) และ (ง)

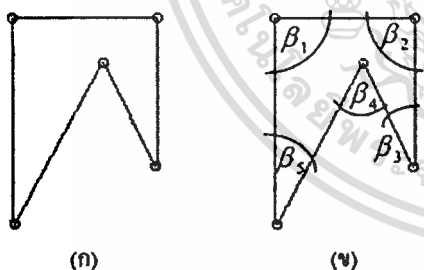
Co_s คือ มุมของจุดเด่นที่ปรากฏในภาพวัตถุที่รับ

เข้ามาพิจารณา (Scene) ในรูปที่ 3 (จ)

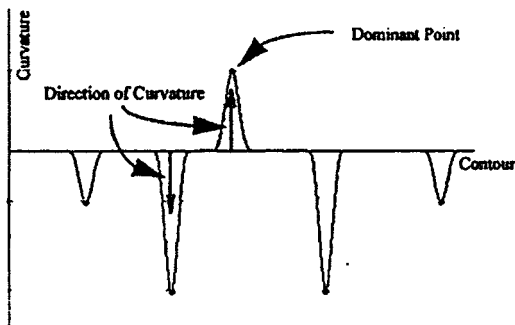
และ β คือ ค่า Threshold ของผลต่างของมุมที่เกิดขึ้น (ซึ่งโดยทั่วไปใช้ค่าประมาณ 3-7 องศา)

ค่า Co_m และ Co_s คำนวณได้จากสมการ (n-6)

(ข) ค่าความโค้งของจุดเด่น



(ก) จุดเด่นของรูปหลายเหลี่ยม (ข) มุมของจุดเด่น



รูปแสดงถึงค่าความโค้งและทิศทางของความโค้งของรูปข้างบน

$$F_2 = \begin{cases} 1 & \text{if } \left| \frac{Cu_m - Cu_s}{|Cu_m| + |Cu_s|} \right| < \lambda \\ -1 & \text{Otherwise} \end{cases} \quad (v-1)$$

เมื่อ Cu_m คือ ค่าความโค้ง (Curvature) ของจุดเด่นที่ปรากฏใน Model รูปที่ 3 (ก), (ข), (ค) และ (ง)

Cu_s คือ ค่าความโค้ง (Curvature) ของจุดเด่นที่ปรากฏใน Scene รูปที่ 3 (จ)

และ λ คือ ค่า Threshold ของค่าความโค้ง

ค่า Cu_m และ Cu_s คำนวณได้จากสมการ (2-8)

(ค) ทิศทางความโค้งของจุดเด่น

$$F_3 = \begin{cases} 1 & \text{if } (h_m \text{ and } h_s) \geq 0 \\ -1 & \text{or } (h_m \text{ and } h_s) < 0 \\ \text{Otherwise} \end{cases} \quad (k-1)$$

เมื่อ h_m และ h_s คือ ทิศทางของค่าความโค้งของจุดเด่นใน Model และ Scene ตามลำดับ

9. เอกสารอ้างอิง

- [1] พารา ลิมนะพีระเสวีรุ, เกษตร์ ศิริกันตีสัตถุทธิ, "การตรวจหาจุดเด่นบนเส้นโค้งของขอบภาพโดยการเปลี่ยนค่าสเกล-สเปซของตัวกรอง," การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 17 1-2 ธันวาคม 2537
- [2] Nirwan Ansari, Kuowei Li, "Landmark-Based Shape Recognition by a Modified Hopfield Neural Network," Pattern Recognition, Vol. 26, No. 4, pp. 531-542, 1993
- [3] Nasser M. Nasrabadi, and Wei Li "Object Recognition by a Hopfield Neural Network," IBBE Trans. Syst., Man, Cybern., Vol. 21 No. 6, pp. 1523-1535, Nov/Dec 1991
- [4] George N. Bebis and George M. Papadourakis, "Object Recognition Using Invariant Object Boundary Representation and Neural Network Models," Pattern Recognition, Vol. 25, No. 1, pp. 25-44, 1992
- [5] Nirwan Ansari, and Edward J. Delp, "Patial Shape Recognition : A Landmark-Based Approach," IBBE Trans. Pattern Anal. Machine Intell., Vol. 12, No.5, pp.

ประวัติผู้เขียน

นายพารา ลิมมะณีประเสริฐ เกิดเมื่อวันที่ 2 ตุลาคม พ.ศ. 2509 ที่จังหวัดหนองคาย สำเร็จการศึกษา
 ชั้นปริญญาโทสหกรรมศาสตรบัณฑิต (สาขาเทคโนโลยีคอมพิวเตอร์อุตสาหกรรม) เกียรตินิยมอันดับสอง จาก
 สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อปีการศึกษา 2534 เริ่มเข้าทำงานครั้งแรกเป็น
 พนักงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย ในตำแหน่งวิศวกรระดับ 4 เป็นระยะเวลา 1 ปี และเข้ารับราชการ
 ในสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในตำแหน่งวิศวกร 3 เป็นระยะเวลา 3 ปี ปัจจุบันเป็น
 อาจารย์พิเศษในสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

