

การตรวจจับเปลวเพลิงในภาพโดยใช้เทคนิคโครงข่ายประสาทเทียม

FLAME DETECTION IN IMAGE USING NEURAL NETWORK

ไกรสิทธิ์ ทองนพคุณ

KRAISEE THONGNOPAKUN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2550

การตรวจจับเปลวเพลิงในภาพโดยใช้เทคนิคโครงข่ายประสาทเทียม

FLAME DETECTION IN IMAGE USING NEURAL NETWORK



ไกรสิทธิ์ ทองนพคุณ

KRAISEE THONGNOPAKUN

เลขหมู่.....  
เลขทะเบียน..... 74529  
วัน,เดือน,ปี..... - 3 ต.ค. 2550

11802480  
b.....  
i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมการวัดคุม

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2550

**FLAME DETECTION IN IMAGE USING NEURAL NETWORK**

**KRAISEE THONGNOPAKUN**

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

**2007**

**COPYRIGHT 2007**

**SCHOOL OF GRADUATE STUDIES**

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

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

---

**หัวข้อวิทยานิพนธ์**      การตรวจจับเปลวเพลิงในภาพโดยใช้เทคนิคโครงข่ายประสาทเทียม  
Flame Detection in Image Using Neural Network

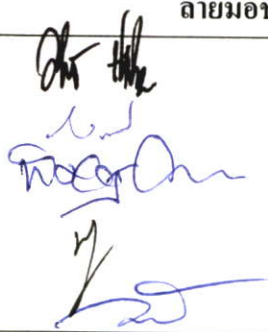
**นักศึกษา**                      นายไกรสิทธิ์      ทองนพคุณ

**รหัสประจำตัว**              45061228

**ปริญญา**                      วิศวกรรมศาสตรมหาบัณฑิต

**สาขาวิชา**                    วิศวกรรมการวัดคุม

**อาจารย์ที่ปรึกษาวิทยานิพนธ์**    รศ.วิทยา ทิพย์สุวรรณพร

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ผศ.อาจินต์	น่วมสำราญ	
ผศ.ทรงชัย	วีระทวีมาศ	
ผศ.ดร.พิศิษฐ์	โกการัตน์กุล	
รศ.ดร.ฟูศักดิ์	ชีวิสุวิทย์	
รศ.วิทยา	ทิพย์สุวรรณพร	

วัน / เดือน / ปี ที่สอบ 23 เมษายน 2550 เวลา 09.00-11.00 น.

สถานที่สอบ ณ อาคาร 12 ชั้น 4 (ห้อง E12-403)

  
บัณฑิตวิทยาลัยรับรองแล้ว  
(รศ.ดร.จารุวัตร เจริญสุข)  
คณบดีบัณฑิตวิทยาลัย

วันที่.....๒๓.....เดือน.....พฤษภาคม.....พ.ศ.....๒๕๕๐.....

หัวข้อวิทยานิพนธ์	การตรวจจับเปลวเพลิงในภาพ โดยใช้เทคนิคโครงข่ายประสาทเทียม
นักศึกษา	นายไกรสิทธิ์ ทองนพคุณ
รหัสนักศึกษา	45061228
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมการวัดคุม
พ.ศ.	2550
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.วิทยา ทิพย์สุวรรณพร

### บทคัดย่อ

วิทยานิพนธ์ฉบับนี้มีจุดมุ่งหมายที่จะศึกษาเกี่ยวกับ การค้นหาผลที่เกิดจากกรรมวิธีการเกิดภาพ (Image Processing) โดยใช้เทคนิคโครงข่ายประสาทเทียม (Artificial neural network, ANN) เพื่อตรวจจับเปลวเพลิงที่เกิดขึ้นภายในภาพ ซึ่งจะใช้ภาพเปลวเพลิงเป็นลักษณะสำคัญของอินพุท ในการพิจารณา ภาพต่างๆจะเอามาจากภาพของการเกิดเปลวเพลิง แล้วจะทำการแก้ไขและปรับปรุงภาพ และทำการหาลักษณะสำคัญของภาพ เพื่อประยุกต์ใช้กับ โครงข่ายประสาทเทียมทำการหาการเกิดเปลวเพลิงภายในภาพ

<b>Thesis Title</b>	Flame detection in image using neural network
<b>Student</b>	Mr. Kraisee Thongnopakun
<b>Student ID.</b>	45061228
<b>Degree</b>	Master of Engineering
<b>Program</b>	Instrumentation Engineering
<b>Year</b>	2007
<b>Thesis Advisor</b>	Assoc.Prof. Vittaya Tipsuwannaporn

### **ABTRACT**

In this thesis, the aim of study to investigate the effectiveness of image processing using artificial neural networks to detect flame images. This uses features of flame images in a simulation fire as input elements, will be considered. Flame images are taken features of the images. For the artificial neural network (ANN) applications are extracted from a rectified image. This considers the difference in distance from the important characters to detect flame images.

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ ผู้เขียนขอกราบขอบพระคุณ รศ. วิทยา ทิพย์สุวรรณพร ที่ได้กรุณาให้คำปรึกษา คำแนะนำ และปรับปรุงแก้ไขข้อบกพร่องต่างๆ แก่ผู้เขียนด้วยดีตลอดมา

พร้อมทั้งขอขอบพระคุณ รศ. ทวีพล ชื้อสัคย์ รศ.ดร. พุศศักดิ์ ชิวสุวิทย์ รศ.ดร.สาธิต อินทจักร์ และคณาจารย์ภาควิชาวิศวกรรมการวัดคุมทุกท่าน ที่กรุณาให้ความช่วยเหลือและคำแนะนำต่างๆ ให้วิทยานิพนธ์ฉบับนี้สมบูรณ์มากยิ่งขึ้น

ตลอดจนเพื่อนๆ นิสิตปริญญาโททุกท่าน ที่ให้ความช่วยเหลือและเป็นกำลังใจมาโดยตลอด

สุดท้ายนี้ผู้เขียนขอกราบขอบพระคุณบิดา มารดา และบุคคลในครอบครัวที่คอยช่วยเหลือให้กำลังใจและอยู่เบื้องหลังความสำเร็จของผู้เขียนตลอดมา หากวิทยานิพนธ์ฉบับนี้มีความบกพร่อง ผิดพลาดประการใด ผู้เขียนขออภัยไว้ด้วยความยินดี

ไกรสิทธิ์ ทองนพคุณ

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	VIII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	1
1.3 ขอบเขตของการวิจัย.....	2
1.4 ขั้นตอนการศึกษา.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
บทที่ 2 การประมวลผลภาพดิจิทัล.....	4
2.1 บทนำ.....	4
2.2 การแสดงภาพดิจิทัล (Digital image representation).....	6
2.3 รูปแบบจำลองของสี (Color models).....	7
2.3.1 รูปแบบจำลองสีแบบอาร์จีบี (RGB color model).....	7
2.3.2 การแสดงภาพสีบนเครื่องไมโครคอมพิวเตอร์ (Display of color image on microcomputer).....	8
2.4 การประมวลผลภาพ (Image processing).....	9
2.4.1 การประมวลผลแบบจุดภาพ (Pixel point processing).....	9
2.4.2 การประมวลผลแบบกลุ่มจุดภาพ (Pixel group processing).....	9
2.5 ฮิสโตแกรมของภาพ (Image histogram).....	11
2.5.1 การยืดคอนทราสต์แบบเชิงเส้น (Linear contrast stretch).....	14
2.5.2 การยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง (Piecewise linear contrast stretch).....	17
2.5.3 การปรับเท่าฮิสโตแกรม (Histogram equalization).....	18
2.5.4 ขีดเริ่มเปลี่ยนฮิสโตแกรม (Histogram thresholding).....	22

## สารบัญ(ต่อ)

	หน้า
2.6 การตรวจจับขอบในภาพ (Edge detection).....	23
2.6.1 การใช้เทมเพลตในการตรวจจับขอบ.....	24
2.7 การปรับเรียบให้ภาพ (Image smoothing).....	28
2.7.1 การปรับเรียบด้วยค่าเฉลี่ย.....	28
บทที่ 3 โครงข่ายประสาทเทียม.....	31
3.1 บทนำ.....	31
3.2 ความรู้พื้นฐานทางกายภาพของระบบเซลล์ประสาท.....	31
3.3 แบบจำลองโครงข่ายเซลล์ประสาท.....	36
3.4 การเรียนรู้ของโครงข่าย.....	40
3.4.1 การสอนแบบชี้แนะหรือดูแล (Supervised Learning).....	40
3.4.2 การสอนแบบไม่มีการชี้แนะหรือไม่มีการดูแล (Unsupervised Learning).....	40
3.5 แบบจำลองโครงข่ายประสาทเทียม.....	41
3.5.1 โครงข่ายที่ส่งสัญญาณไปข้างหน้า (Feedforward Networks).....	42
3.5.2 โครงข่ายที่มีการป้อนกลับ (Feedback Networks).....	42
3.6 การแพร่กระจายกลับ.....	43
3.7 กฎเซลล์ออดเนกประสงค์.....	46
3.8 สรุป.....	49
บทที่ 4 ผลการทดลองการตรวจจับเปลวเพลิงภายในภาพ.....	50
4.1 อุปกรณ์ที่ใช้ในการทดลอง.....	50
4.2 ผลการเตรียมข้อมูลภาพ.....	50
4.2.1 การออกแบบการหาค่าฮิสโตแกรมของภาพ.....	50
4.2.2 การปรับปรุงภาพ.....	51
4.2.3 การจำแนกภาพ.....	52
4.2.4 การหาขอบภาพ.....	53
4.2.5 การแยกคุณลักษณะเด่น.....	53
4.2.5.1 การหาทิศทางหลัก.....	53

## สารบัญ(ต่อ)

	หน้า
4.3 ผลการฝึกการรู้จักโดยใช้โครงข่ายประสาทเทียม.....	54
4.3.1 การรับค่าอินพุต.....	54
4.3.2 การฝึกการรู้จัก.....	55
4.3.2.1 ขั้นตอนการฝึกการรู้จัก.....	55
4.3.2.2 การกำหนดจำนวนเซลล์ในโครงข่ายประสาทเทียม.....	56
4.3.3 ผลการทดสอบการรู้จัก.....	56
บทที่ 5 บทสรุป และแนวทางในการพัฒนาต่อไป.....	78
5.1 สรุป.....	78
5.2 วิจัยณ์และข้อเสนอแนะ.....	78
5.3 แนวทางในการพัฒนาต่อไป.....	78
เอกสารอ้างอิง.....	80
ภาคผนวก ผลงานที่ได้รับการตีพิมพ์.....	82
ประวัติผู้เขียน.....	87

## สารบัญตาราง

ตารางที่	หน้า
2.1 ลักษณะการกระจายของค่าระดับสีเทาของภาพที่มีจำนวนระดับสีเทา 8 ระดับ.....	21
3.1 แสดงการปรับค่าถ่วงน้ำหนักสำหรับกฎการสอนแบบต่างๆ.....	41
4.1 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 1 พิกเซล.....	59
4.2 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 10 พิกเซล.....	60
4.3 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 50 พิกเซล.....	61
4.4 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 100 พิกเซล.....	62
4.5 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 500 พิกเซล.....	63
4.6 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 1,000 พิกเซล.....	64
4.7 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 3,000 พิกเซล.....	65
4.8 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 6,500 พิกเซล.....	66
4.9 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 13,000 พิกเซล.....	67
4.10 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 20,000 พิกเซล.....	68
4.11 แสดงความถูกต้องในการจำแนกแบบ 2 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 6,500 พิกเซล (ประมาณ 10 %).....	69
4.12 แสดงความถูกต้องในการจำแนกแบบ 2 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 13,000 พิกเซล (ประมาณ 20 %).....	72
4.13 แสดงความถูกต้องในการจำแนกแบบ 2 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 20,000 พิกเซล (ประมาณ 30 %).....	74

# สารบัญรูป

รูปที่	หน้า
2.1 ระบบเก็บภาพดิจิทัล.....	5
2.2 แสดงภาพดิจิทัลแบบขาวดำ.....	6
2.3 แสดงรูปแบบจำลองสีแบบอาร์จีบี.....	8
2.4 แสดงไดอะแกรมของฮีสโตแกรมสามมิติของการแสดงภาพสีโดยการใช้รูปแบบการจำลองสีแบบอาร์จีบีที่มีขนาด 24 บิต ต่อจุดภาพ.....	8
2.5 หลักการประมวลผลแบบจุดภาพ.....	9
2.6 เทมเพลตขนาด 3x3.....	11
2.7 ภาพถ่ายเปลวเพลิงกับฮีสโตแกรมที่สอดคล้อง.....	12
2.8 ฟังก์ชันความหนาแน่นความน่าจะเป็นและความน่าจะเป็นสะสมที่สอดคล้อง.....	12
2.9 ภาพที่มีลักษณะมืด เมื่อข้อมูลของระดับสีเทาของจุดภาพส่วนใหญ่ปรากฏอยู่ทางด้านระดับความสว่างค่าต่ำ ๆ.....	13
2.10 ภาพที่มีลักษณะสว่าง เมื่อข้อมูลของระดับสีเทาของจุดภาพส่วนใหญ่ปรากฏอยู่ทางด้านระดับความสว่างค่าสูง ๆ.....	14
2.11 ภาพที่ดี เมื่อจุดภาพมีค่าระดับความสว่างเกือบทุกระดับ.....	14
2.12 การแปลงฮีสโตแกรมโดยการยืดคอนทราสต์แบบเชิงเส้น.....	15
2.13 ฟังก์ชันที่ใช้ในการแปลงฮีสโตแกรมของการยืดคอนทราสต์แบบเชิงเส้น.....	16
2.14 ภาพที่ได้จากการยืดคอนทราสต์แบบเชิงเส้น.....	16
2.15 ค่าฟังก์ชันตัวอย่างของการยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง.....	17
2.16 ภาพจากการยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง.....	17
2.17 การปรับเท่าฮีสโตแกรมของตารางที่ 2.1.....	21
2.18 การปรับเท่าฮีสโตแกรมภาพ.....	22
2.19 การแปลงภาพไบนารีโดยใช้ขีดเริ่มเปลี่ยนฮีสโตแกรม.....	23
2.20 เทมเพลตเกรเดียนต์ของลาปลาเซียนแบบต่าง ๆ.....	24
2.21 เทมเพลต 8 ทิศทาง.....	25
2.22 การประยุกต์ใช้เทมเพลตทิศตะวันตกแบบสามระดับกับข้อมูลภาพ.....	26
2.23 เทมเพลตสองทิศทางที่นิยมใช้กัน.....	26
2.24 การสร้างภาพของขอบหรือเกรเดียนต์.....	28
2.25 เทมเพลต $W(i, j)$ ที่ใช้ในการปรับเรียบภาพด้วยค่าเฉลี่ย.....	29
2.26 การปรับเรียบด้วยค่าเฉลี่ย.....	30

## สารบัญรูป(ต่อ)

รูปที่	หน้า
3.1 การสื่อสารข้อมูลในระบบประสาท.....	32
3.2 โครงสร้างของเซลล์ประสาทและการส่งสัญญาณกระแสประสาท.....	33
3.3 แสดงลักษณะการเกิดศักย์ไฟฟ้าขณะทำงาน.....	35
3.4 แสดงภาวะต่างๆ ของเซลล์ประสาท.....	35
3.5 แสดงศักย์ไฟฟ้าขณะทำงานที่เกิดจากการรวมกันของซิแนปส์.....	36
3.6 แบบจำลองของเซลล์ประสาทของ McCulloch – Pitts และ โครงข่ายลอจิกพื้นฐาน.....	37
3.7 แบบจำลองโครงข่ายเซลล์ประสาทเทียม.....	38
3.8 Activation Function ชนิดต่างๆ.....	39
3.9 โครงข่าย Feed forward แบบชั้นเดียว.....	42
3.10 โครงข่ายป้อนกลับแบบเวลาไม่ต่อเนื่องชั้นเดียว.....	43
3.11 โครงข่าย Multilayer Perceptron ที่มี 3 ชั้น.....	43
3.12 แสดงความสัมพันธ์ระหว่างตัวแปร $X$ กับ $Y$ .....	44
3.13 แสดงความสัมพันธ์ระหว่างตัวแปร $X$ กับ $Y$ โดยมีชั้นฮิดเดนเป็นตัวเชื่อม.....	45
4.1 แสดงไม่มีเปลวเพลิงเกิดขึ้นภายในภาพและฮิสโตแกรมที่สอดคล้อง.....	50
4.2 แสดงเปลวเพลิงที่เกิดขึ้นภายในภาพโดยเปลวเพลิง และฮิสโตแกรมที่สอดคล้อง.....	50
4.3 ภาพเมื่อทำการปรับปรุงภาพ.....	51
4.4 แสดงผลการทำ Thresholding.....	52
4.5 แสดงผลการหาขอบภาพ เมื่อให้ค่า Threshold=180.....	53
4.6 แสดงขั้นตอนการหาค่าลักษณะเด่นโดยวิธีการหาทิศทางของลายเส้น.....	54
4.7 แสดงการหาค่าลักษณะเด่นของภาพ.....	55
4.8 อินพุทของโครงข่าย.....	55
4.9 แสดงขั้นตอนการฝึกการรู้จำ.....	56
4.10 แสดงจำนวนเซลล์ประสาทในโครงข่ายประสาทเทียม.....	57
4.11 แสดงขั้นตอนการทดสอบการรู้จำ.....	58
4.12 แสดงข้อมูลจากตารางที่ 4.1.....	59
4.13 แสดงข้อมูลจากตารางที่ 4.2.....	60
4.14 แสดงข้อมูลจากตารางที่ 4.3.....	61
4.15 แสดงข้อมูลจากตารางที่ 4.4.....	62
4.16 แสดงข้อมูลจากตารางที่ 4.5.....	63

## สารบัญรูป(ต่อ)

รูปที่	หน้า
4.17 แสดงข้อมูลจากตารางที่ 4.6.....	64
4.18 แสดงข้อมูลจากตารางที่ 4.7.....	65
4.19 แสดงข้อมูลจากตารางที่ 4.8.....	66
4.20 แสดงข้อมูลจากตารางที่ 4.9.....	67
4.21 แสดงข้อมูลจากตารางที่ 4.10.....	68

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

อัครกัฏยเป็นปรากฏการณ์ธรรมชาติที่เกิดขึ้น เมื่อสสารมีอุณหภูมิสูงขึ้นจนถึงอุณหภูมิวิกฤติจะเกิดปฏิกิริยาทางเคมี เช่น เมื่อถูกกักบอออกซิเจนจะเกิดความร้อน เปลวเพลิง แสงสว่าง ควัน ไอ น้ำ คาร์บอนมอนนอกไซด์ ซึ่งในการเกิดอัครกัฏยในแต่ละครั้งก็จะทำให้เกิดความอันตรายต่อชีวิต และทำให้สูญเสียทรัพย์สิน รวมถึงจะมีผลกระทบต่อความเสียหายทางเศรษฐกิจ และเป็นเวลาหลายศตวรรษมาแล้วที่มนุษย์ได้ประดิษฐ์และพัฒนาเครื่องจักรขึ้นมาใช้งาน ซึ่งเราจะเห็นได้ว่าเครื่องจักรเหล่านี้จะถูกใช้สำหรับงานที่ที่มีความยุ่งยากซับซ้อน ถ้าหากใช้คนทำก็อาจจะทำให้เกิดความผิดพลาดได้ง่าย จนกระทั่งได้มีการพัฒนาเครื่องคอมพิวเตอร์ขึ้นมา และในปัจจุบันเครื่องคอมพิวเตอร์ได้เข้ามามีบทบาทเพิ่มขึ้นเรื่อยๆ ซึ่งงานส่วนใหญ่ที่ใช้คอมพิวเตอร์คือ งานจัดเก็บข้อมูล งานประมวลผลข้อมูล ดังนั้นนักวิทยาศาสตร์ได้เล็งเห็นความสำคัญนี้ จึงได้พัฒนาคอมพิวเตอร์ให้มีความฉลาดเพิ่มมากขึ้น โดยพยายามให้เครื่องคอมพิวเตอร์และโปรแกรมที่ใช้ มีการเรียนรู้ จดจำ และสามารถตัดสินใจเองได้ ฉะนั้นทฤษฎีของโครงข่ายประสาทเทียม(Neural Network System) เป็นทฤษฎีหนึ่งที่น่าสนใจไปประยุกต์ใช้งานในการวิเคราะห์ข้อมูลและเพิ่มประสิทธิภาพให้กับคอมพิวเตอร์

ในวิทยานิพนธ์ฉบับนี้จึงได้นำเสนอเกี่ยวกับการตรวจจับเปลวเพลิงภายในภาพ โดยอาศัยหลักการประมวลผลภาพดิจิทัล (Digital Image Processing) ทำงานร่วมกับโครงข่ายประสาทเทียม เพื่อทำการตรวจจับภาพว่าภาพไหนเป็นภาพที่มีเปลวเพลิง และภาพไหนเป็นภาพที่ไม่มีเปลวเพลิง เพื่อที่จะได้นำการวิจัยนี้ไปประยุกต์ใช้กับการออกแบบเป็นตัวตรวจจับการเกิดเพลิงไหม้อัตโนมติต่อไป

### 1.2 วัตถุประสงค์ของการวิจัย

- 1.2.1 เพื่อหาคุณสมบัติและขั้นตอนการทำงานของการประมวลผลภาพ
- 1.2.2 เพื่อศึกษาและพัฒนาประสิทธิภาพในการแยกแยะกลุ่มของข้อมูลภาพ
- 1.2.3 เพื่อศึกษาคุณสมบัติและขั้นตอนการทำงานของโครงข่ายประสาทเทียม
- 1.2.4 เพื่อศึกษาและพัฒนาการนำโครงข่ายประสาทเทียมมาใช้ในการวิเคราะห์ภาพถ่าย
- 1.2.5 เพื่อหาจุดบกพร่องที่เกิดขึ้นจากการนำเอาทฤษฎีโครงข่ายประสาทเทียมมาใช้งาน

### 1.3 ขอบเขตของการวิจัย

เป็นการนำเสนอการหาเปลวเพลิงภายในภาพ โดยใช้หลักการประมวลผลภาพดิจิทัลทำงานร่วมกับโครงข่ายประสาทเทียม ในการวิเคราะห์ว่าภาพไหนเป็นภาพที่มีเปลวเพลิงและภาพไหนเป็นภาพที่ไม่มีเปลวเพลิง โดยจะทำการวิเคราะห์ภาพดิจิทัลหาค่าต่างๆ ของภาพ เช่นหาค่าฮิสโตแกรม ค่าเฉลี่ยของค่าความสว่างของภาพ ค่าความเบี่ยงเบนมาตรฐานของภาพ เป็นต้น เพื่อนำไปทำการกำหนดอินพุตให้กับโครงข่ายประสาทเทียม และทำการสอน(Training) ให้โครงข่ายประสาทเทียมสามารถวิเคราะห์ได้ว่าภาพไหนเป็นภาพของเปลวเพลิง และภาพไหนเป็นภาพที่ไม่มีเปลวเพลิง

### 1.4 ขั้นตอนการศึกษา

ในปัจจุบันได้มีการนำเอาการประมวลผลภาพดิจิทัล (Digital Image Processing) และโครงข่ายประสาทเทียม (Artificial neural network, ANN) มาใช้กันอย่างแพร่หลาย ในวิทยานิพนธ์ฉบับนี้ได้นำเสนอเกี่ยวกับ การตรวจจับเปลวเพลิงภายในภาพ โดยใช้เทคนิคโครงข่ายประสาทเทียม โดยจะแบ่งขั้นตอนการศึกษาออกเป็น 5 บท ดังนี้

#### บทที่ 1 บทนำ

กล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ของการวิจัย ขอบเขตของการวิจัย และขั้นตอนของการศึกษา ประโยชน์ที่คาดว่าจะได้รับ

#### บทที่ 2 การประมวลผลภาพดิจิทัล

กล่าวถึงบทนำ การแสดงภาพดิจิทัล รูปแบบการจำลองของสี การประมวลผลภาพ ฮิสโตแกรมของภาพ การตรวจจับขอบในภาพ และการปรับเรียบให้ภาพ

#### บทที่ 3 โครงข่ายประสาทเทียม

กล่าวถึงความรู้พื้นฐานทางกายภาพของระบบเซลล์ประสาท แบบจำลองโครงข่ายเซลล์ประสาท การเรียนรู้ของโครงข่าย แบบจำลองของโครงข่ายประสาทเทียม การแพร่กระจายกลับ กฎเดลต้าอเนกประสงค์ และสรุป

#### บทที่ 4 ผลการทดลองการตรวจจับเปลวเพลิงภายในภาพ

กล่าวถึงอุปกรณ์ที่ใช้ในการทดลอง ผลการเตรียมข้อมูลภาพ และผลการหาภาพเปลวเพลิงโดยใช้โครงข่ายประสาทเทียม

บทที่ 5 บทสรุป และแนวทางในการพัฒนาต่อไป

กล่าวถึงผลสรุป วิจัยและข้อเสนอแนะ พร้อมทั้งเสนอแนวทางในการพัฒนาต่อไป

### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 ได้รับความรู้เกี่ยวกับขั้นตอนการทำงานของกรมผลผลิต
- 1.5.2 ได้รับความรู้ในการนำโครงข่ายประสาทเทียมมาประยุกต์ใช้ในการใช้งานจริง
- 1.5.3 สามารถนำระบบการทำงานในวิทยานิพนธ์ฉบับนี้ไปปรับใช้ได้กับงานวิจัยอื่นที่เกี่ยวข้องกับภาพสองมิติ

## บทที่ 2

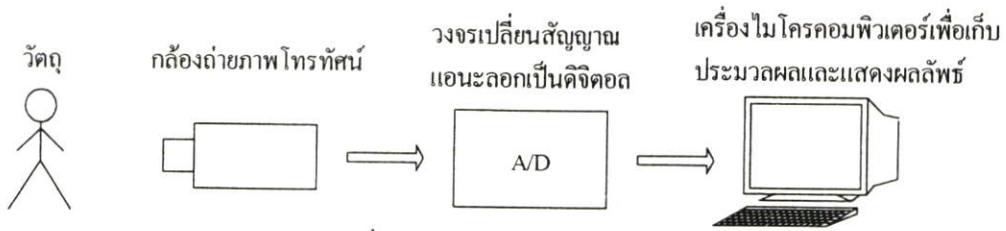
# การประมวลผลภาพดิจิทัล

### 2.1 บทนำ

ข้อมูลภาพได้จากแสงตกกระทบวัตถุแล้วสะท้อนผ่านเลนส์เข้าสู่ตัวบันทึกภาพ ซึ่งอาจอยู่ในรูปของฟิล์ม (Film) หรือตัวตรวจจับ (Sensor) ดังนั้นกระบวนการถ่ายภาพจึงเป็นวิธีการแปลงข้อมูลเชิงกายภาพ (Physical) ของวัตถุสามมิติ (Three-dimensional) มาเป็นข้อมูลภาพ (Image) ของวัตถุสองมิติ (Two-dimensional) สำหรับกรณีการบันทึกภาพของวัตถุด้วยกล้องถ่ายภาพโทรทัศน์ (Television camera) นั้น จะทำการบันทึกภาพอยู่ในรูปของสัญญาณไฟฟ้าแบบแอนะล็อก (Analog) ในขั้นตอนนี้ตัวตรวจจับจะทำการวัดค่าความเข้มของจุดเล็ก ๆ ทีละจุดไปเรื่อย ๆ ตามแนวเส้นทางที่กำหนดไว้ ซึ่งปกติจะไล่จากซ้ายไปขวาและบนลงล่าง ค่าที่ตัวตรวจจับวัดได้นี้เป็นค่าความเข้มแสงของจุดภาพนั้น ในการพิจารณาภาพหนึ่งภาพใดที่ได้จากกล้องถ่ายภาพโทรทัศน์จะเป็นภาพนิ่ง (Still picture) จากการนำภาพนิ่งที่เวลาต่าง ๆ มาจัดเรียงกันเป็นจำนวนมากแล้วให้แสดงผลในช่วงเวลาสั้น ๆ ภาพที่ได้จะมีลักษณะ เคลื่อนไหว (Moving picture) สำหรับในวิทยานิพนธ์ฉบับนี้จะกล่าวถึงเฉพาะภาพนิ่งเท่านั้น

สัญญาณภาพในแต่ละเส้นสแกน (Scan) ที่ได้จากกล้องถ่ายภาพโทรทัศน์ ซึ่งมีลักษณะเป็นแรงดันไฟฟ้าแบบแอนะล็อกจะสอดคล้องกับความสว่างบนวัตถุที่ถ่ายได้ กล่าวคือ บริเวณที่สว่างจะให้ค่าแรงดันไฟฟ้าสูง ส่วนบริเวณที่มืดจะให้ค่าแรงดันไฟฟ้าต่ำ แต่เนื่องจากสัญญาณภาพที่ได้จะต้องนำไปประมวลผลด้วยเครื่องไมโครคอมพิวเตอร์ (Micro computer) เพื่อจุดประสงค์หนึ่งจุดประสงค์ใด ดังนั้นแต่ละเส้นสแกนจะถูกชักตัวอย่าง (Sampling) การชักตัวอย่างจะต้องเป็นไปตามอัตราในควิสต์ (Nyquist rate) สำหรับบล็อก ไดอะแกรม (Block diagram) ของระบบเก็บข้อมูลภาพดิจิทัล (Digital image acquisition system) แสดงได้ดังรูปที่ 2.1

ในการชักตัวอย่างของแต่ละเส้นสแกนจะเป็นการบอกถึงรายละเอียด (Resolution) ของข้อมูลภาพเชิงตัวเลข ทั้งนี้จะต้องขึ้นกับขีดความสามารถทางฮาร์ดแวร์ (Hardware) ของตัวแปลงสัญญาณแอนะล็อกเป็นข้อมูลเชิงตัวเลข (A/D Converter) ว่าสามารถทำงานได้ที่ความถี่ของการชักตัวอย่าง (Sampling frequency) เท่าไร รวมทั้งช่วงเวลาเข้าถึง (Access time) ของแรม (RAM) ที่ใช้เก็บข้อมูลชั่วคราวก่อนจะถ่ายโอน (Transfer) ข้อมูลดิจิทัลเข้าสู่หน่วยความจำของเครื่องคอมพิวเตอร์ เพื่อเก็บข้อมูลไว้อย่างถาวรหรือเพื่อการประมวลผลภาพ โดยปกติแล้วตัวแปลง A/D มักนิยมใช้ขนาด 8 บิต เพื่อให้สอดคล้องกับระบบทางฮาร์ดแวร์ของเครื่องคอมพิวเตอร์



รูปที่ 2.1 ระบบเก็บภาพดิจิทัล

จากการเก็บข้อมูลภาพดิจิทัลในแต่ละเส้นสแกน ถ้าหากภาพ ๆ หนึ่งมีจำนวนหลายเส้นสแกน ดังนั้นข้อมูลภาพดิจิทัลที่ได้สามารถเขียนอยู่ในรูปฟังก์ชันสองมิติของความเข้มของแสง  $f(x, y)$  โดยค่าของฟังก์ชันจะแสดงถึงค่าความเข้มของแสงที่พิกัดสเปเชียล (Spatial) เป็น  $(x, y)$  ถ้าหากใช้ A/D ขนาด 8 บิต จะพบว่าค่าความเข้มของแสงจะแบ่งออกได้  $2^8$  หรือ 256 ระดับของสีเทา เมื่อค่าระดับสีเทาเป็น 0 จะหมายถึงจุดภาพนั้นมีค่าความเข้มของแสงต่ำจะให้จุดภาพเป็นสีดำ ในทางตรงกันข้ามถ้าหากค่าระดับสีเทาเป็น 255 แสดงว่าจุดภาพนั้นมีค่าความเข้มของแสงสูงจะให้จุดภาพเป็นสีขาว ส่วนค่าอื่น ๆ ที่อยู่ระหว่างค่าทั้งสองคือ 0 ถึง 255 ก็จะมีระดับสีที่ไล่จากดำไปหาขาวนั่นเอง ภาพดิจิทัลเมื่อนำมาประมวลผลในเครื่องคอมพิวเตอร์ที่อยู่ในรูปฟังก์ชัน  $f(x, y)$  สามารถเขียนอยู่ในรูปเมทริกซ์ (Matrix) สองมิติ โดยแต่ละมิติจะแทนพิกัดสเปเชียลและค่าสัมประสิทธิ์ของเมทริกซ์จะเป็นค่าระดับสีเทาของจุดภาพนั้น อย่างเช่น ภาพขนาด  $M \times N$  เขียนอยู่ในรูปเมทริกซ์ได้เป็น

$$f(x, y) = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ f(2,1) & f(2,2) & \cdots & f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ f(M,1) & f(M,2) & \cdots & f(M,N) \end{bmatrix} \quad (2.1)$$

และสามารถเขียนอยู่ในรูปเวกเตอร์ (Vector) ได้โดยการดึงข้อมูลภาพทีละแถวในแนวนอน (Row) หรือทีละแถวในแนวตั้ง (Column) การเขียนข้อมูลในรูปเวกเตอร์จะได้ รูปแบบที่กะทัดรัดกว่า และสามารถนำไปประยุกต์ใช้กับการประมวลสัญญาณดิจิทัล หนึ่งมิติที่มีอยู่

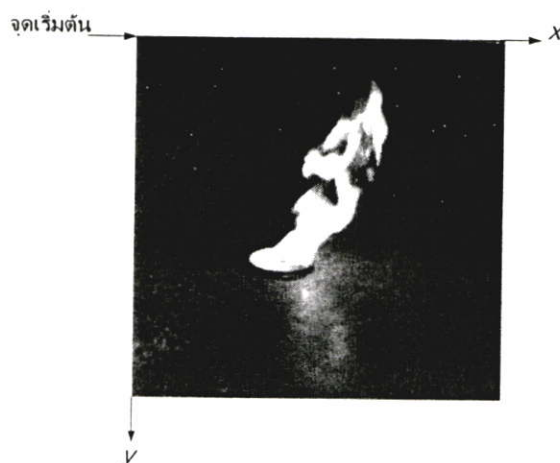
ในการนำข้อมูลภาพดิจิทัลไปใช้งานด้านต่าง ๆ จะต้องทำการประมวลผล ซึ่งสามารถแบ่งการประมวลผลออกเป็นกลุ่มใหญ่ ๆ ได้สองกลุ่มคือ กลุ่มแรกเป็นการประมวลผลเชิงจุด (Point operation) จะเป็นวิธีการประมวลผลอย่างง่าย โดยจะทำการแปลงค่าระดับสีเทาของจุดภาพต้นฉบับ (Original image) ให้เป็นค่าระดับสีเทาใหม่ในภาพผลลัพธ์ (Transformed image) หลังการแปลง เทคนิคที่จัดว่าเป็นการประมวลผลเชิงจุดคือ การปรับปรุงฮิสโตแกรม (Histogram modification) เป็นต้น กลุ่มที่สองเป็นการประมวลผลเชิงกลุ่มข้างเคียง (Neighborhood) เป็นวิธีการ

นำค่าระดับสีเทาของจุดภาพข้างเคียงที่อยู่รอบๆ มาประมวลผลด้วย ซึ่งเทคนิคที่จัดว่าเป็นการประมวลผลเชิงกลุ่มข้างเคียงคือ การกรอง สัญญาณภาพ (Filtering) เป็นต้น

## 2.2 การแสดงภาพดิจิทัล (Digital image representation) [2]

ข้อมูลภาพดิจิทัลจะมีลักษณะเป็นแถวและลำดับ(Array) แต่ละจุดภาพจะอยู่ในรูปสี่เหลี่ยมเล็กๆ ในทางกายภาพ แต่ละจุดภาพที่ถูกจัดวางอยู่ภาพจะมีตำแหน่งบอกพิกัดทางสเปเชียล (Spatial coordinate) ในรูปคาร์ทีเซียน (Cartesian) เป็น  $(x, y)$  เมื่อ  $x$  เป็นลำดับตำแหน่งในแนวนอน และ  $y$  เป็นลำดับตำแหน่งในแนวตั้ง ปกติแล้วตำแหน่งจุดภาพ  $(0,0)$  จะอยู่ทางด้านซ้ายมือสุดของภาพ การจัดลำดับตำแหน่งของจุดภาพจะเรียงจากซ้ายไปขวาในแต่ละเส้นภาพ และการจัดลำดับของเส้นภาพจะเรียงจากบนลงล่าง ถ้าหากว่าเขียนให้อยู่ในรูปของฟังก์ชัน  $f(x, y)$  จะทำให้ทราบว่าค่าของฟังก์ชันจะเป็นระดับความสว่างของจุดภาพดังกล่าว ซึ่งบางครั้งจะเรียกว่าค่าระดับสีเทา (Brightness or gray level) ในแต่ละภาพจะมีจำนวนระดับสีเทาที่ระดับขึ้นอยู่กับจำนวนบิตที่นำมาเข้ารหัสว่ามีกี่บิต (Bit) ในบางครั้งถ้ากำหนดให้ฟังก์ชันหรือค่าระดับสีเทาอยู่ในแกน  $z$  ก็จะสามารถพล็อต (Plot) ข้อมูลภาพดิจิทัลให้เป็นภาพสามมิติได้ แต่การมองภาพเป็นการมองในแนวตั้งฉาก ดังนั้นค่าความสูงของระดับสีเทาในแกน  $z$  จึงถูกปรับเปลี่ยนให้อยู่ในความดำของจุดภาพในสองมิติแทนดังแสดงในรูปที่ 2.2

หากจุดภาพที่อยู่ประชิดติดกันมีค่าความแตกต่างของระดับสีเทาที่จะเกิดเป็นขอบของวัตถุในภาพขึ้น การเปลี่ยนแปลงค่าระดับสีของจุดภาพต่างๆ นี้เองทำให้ผู้มองสามารถแยกแยะรายละเอียดของภาพ หรือสามารถจำแนก (Classify) วัตถุต่างๆ ในภาพได้ เนื่องจากตำแหน่งโคออดิเนต  $(x, y)$  จะเป็นเลขจำนวนเต็มหน่วย (Integer) ดังนั้นแต่ละตำแหน่งโคออดิเนตจึงเป็นจุดภาพที่เรียกว่า พิกเซล (Pixel หรือ Picture element) โดยขนาดของจุดภาพขึ้นอยู่กับค่า  $x$  และ  $y$  ที่จะแบ่งไป ตัวอย่างเช่น รูปที่ 2.2 ซึ่งเป็นภาพขนาด  $256 \times 256$  จุดภาพ และค่าความแตกต่างของความสว่างหรือระดับสีเทาเป็น 256 ระดับ



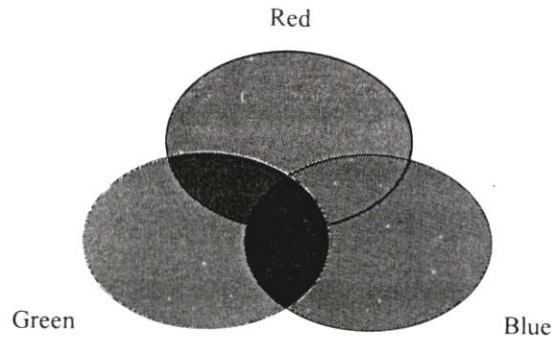
รูปที่ 2.2 แสดงภาพดิจิทัลแบบขาวดำ

## 2.3 รูปแบบจำลองของสี (Color models)

การแสดงผลภาพสีมีลักษณะเป็นรูประบบพิกัดแบบสามมิติ แต่ละสีเกิดจากการผสมระหว่างแกนทั้งสาม โดยรูปแบบจำลองของสีในปัจจุบันได้มีการเปลี่ยนไปตามฮาร์ดแวร์ เช่น จอภาพสี เครื่องพิมพ์ และการประยุกต์ใช้งานในด้านต่างๆ รูปแบบของการจำลองสีที่นิยมใช้กันมากสำหรับจอมอนิเตอร์สี และกล้องวิดีโอคือ RGB model (Red, Green, Blue) สำหรับรูปแบบจำลองของสีแบบ CMY model (Cyan, Magenta, Yellow) นิยมใช้กับเครื่องพิมพ์สี โดยที่ Y คือ Luminance และ I กับ Q คือส่วนของ Chromatic ที่เป็น In phase และ Quadrature ตามลำดับ ส่วนรูปแบบจำลองความถี่สีที่ใช้ในการเปลี่ยนไปเปลี่ยนมาระหว่างรูปแบบจำลองของสีแบบต่างๆ คือ ระบบรูปแบบจำลอง HIS (Hue, Saturation, Intensity) และรูปแบบจำลอง HSV (Hue, Saturation, Value) รูปแบบจำลองสีที่นิยมใช้กันมากสำหรับการประมวลผลภาพ (Image Processing) นั้นคือ RGB model, CMY model, และ HIS model สำหรับในบทนี้จะอธิบายเฉพาะรูปแบบจำลองของสีแบบอาร์จีบีเท่านั้น

### 2.3.1 รูปแบบจำลองสีแบบอาร์จีบี (RGB color mode)

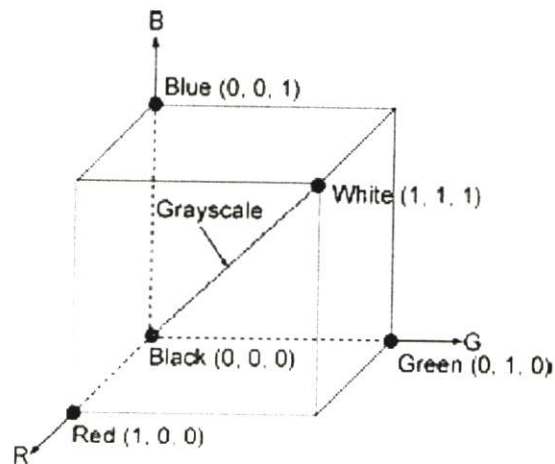
ลำของแสงเมื่อผ่านแท่งแก้วปริซึมแล้วแสงที่ออกมาไม่ได้เป็นสีขาว แต่จะประกอบไปด้วยสเปกตรัมที่ต่อเนื่องของสีที่มีย่านจากสีม่วง (Violet) ไปจนถึงสีแดง (Red) โคนพื้นฐานแล้วสีที่มนุษย์สามารถมองเห็นวัตถุนั้น เกิดจากการธรรมชาติของแสงในการสะท้อนของวัตถุ ซึ่งแสงทั้งหมดทุกแสงจะถูกมองเป็นสีขาว วัตถุที่สะท้อนบางส่วนของสเปกตรัม (Spectrum) จะใช้บางเฉดของสี อย่างเช่นวัตถุที่มีแสงสีเขียวจะสะท้อนช่วงคลื่นที่มีความยาว 500 ถึง 700 นาโนเมตร โดยการดูดซับเอาความยาวคลื่นที่เหลือออกนั้นไว้หมด ทำให้เราสามารถมองเห็นวัตถุชิ้นนั้นเป็นสีเขียวนั่นเอง ในย่านที่มนุษย์สามารถมองเห็นได้นั้นความยาวคลื่นของแสงจะเป็นตัวกำหนดสีที่อยู่ในหน่วยนาโนเมตร (nm) ในระบบรูปแบบจำลองสีแบบอาร์จีบี (RGB color model) นั้นจะมีสีหลักๆ อยู่ 3 สีคือ สีแดง (Red color) สีเขียว (Green color) สีน้ำเงิน (Blue color) จะถูกเรียกว่าสีปฐมภูมิ (Primary color) เมื่อมีการผสมของสีแดง สีเขียว และสีน้ำเงิน เข้าด้วยกันก็จะได้สีต่างๆ ออกมามากมายดังแสดงในรูปที่ 2.3 ซึ่งสังเกตเห็นว่า สีน้ำเงินผสมสีแดงจะได้สีม่วง สีน้ำเงินผสมสีเขียวจะได้สีฟ้าแกมเขียว สีแดงผสมสีเขียวได้สีเหลือง วิธีการนี้เป็นวิธีที่มีประสิทธิภาพดีเพียงพอในการแสดงผลภาพ และมีความนิยมใช้กันมาก เช่น การแสดงผลในมอนิเตอร์ (Monitor) สีสำหรับใช้ในคอมพิวเตอร์ทั่วไป เป็นต้น



รูปที่ 2.3 แสดงรูปแบบจำลองสีแบบอาร์จีบี

### 2.3.2 การแสดงภาพสีบนเครื่องไมโครคอมพิวเตอร์ (Display of color image on microcomputer)

ข้อมูลภาพสี เป็นข้อมูลขนาด 8 บิต ดังนั้นข้อมูลสีหนึ่งจุดจะมีขนาด 24 บิต หรือเรียกว่าเป็นข้อมูล 24 บิต ต่อหนึ่งจุดภาพ (Bit per Pixel) นั่นคือแต่ละจุดภาพจะมีความเป็นไปได้ที่จะมีข้อมูลภาพที่แตกต่างกันถึง  $2^{24}$  ระดับ คือเท่ากับ 16.7 ล้านระดับที่แตกต่างกัน ถ้ามีการแสดงผลภาพสีนำข้อมูลภาพของทั้งสามสีมาพล็อตเป็นฮิสโตแกรม (Histogram) สามมิติ ซึ่งแต่ละแกนคือค่าระดับสีเทาของข้อมูลภาพที่มีค่าได้ตั้งแต่ 0-255 ทำให้เกิดบล็อกสีเหลี่ยมเล็กๆภายในฮิสโตแกรมดังกล่าวจำนวน  $256 \times 256 \times 256$  บล็อก โคนแต่ละบล็อกคือข้อมูลของเฉดสีที่เกิดขึ้นในแต่ละจุดภาพนั่นเอง ฮิสโตแกรมของภาพดังแสดงในรูปที่ 2.4



รูปที่ 2.4 แสดงไดอะแกรมของฮิสโตแกรมสามมิติของการแสดงภาพสีโดยการใช้รูปแบบการจำลองสีแบบอาร์จีบีที่มีขนาด 24 บิต ต่อจุดภาพ (1 แทนค่า 255)

## 2.4 การประมวลผลภาพ (Image Processing)

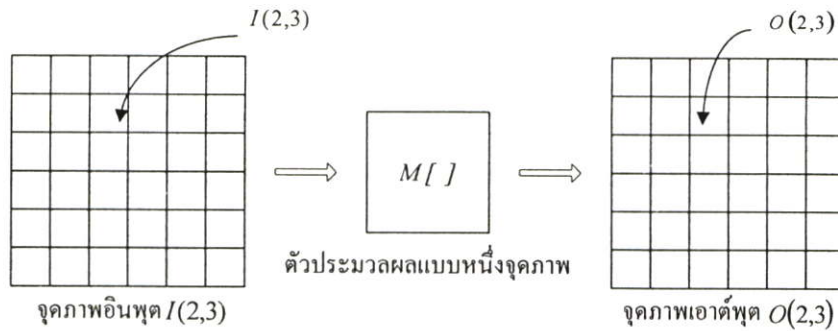
ในการประมวลผลข้อมูลภาพดิจิทัลจะทำได้ในสองลักษณะ โดยลักษณะแรกเป็นการประมวลผลแบบจุดภาพ (Pixel point processing) ส่วนลักษณะที่สองเป็นการประมวลผลแบบกลุ่มจุดภาพ (Pixel group processing)

### 2.4.1 การประมวลผลแบบจุดภาพ (Pixel point processing)

เป็นวิธีการประมวลผลที่ง่ายที่สุด จะทำการเปลี่ยนแปลงค่าระดับสีเทาของจุดภาพแบบจุดต่อจุดเป็นหลักการมูลฐาน ค่าระดับสีเทาของแต่ละจุดภาพที่เปลี่ยนไปโดยมากได้มาจากการคำนวณทางคณิตศาสตร์หรือทางตรรก [4] สมการที่ใช้ในการประมวลผลแบบจุดภาพเขียนได้เป็น

$$O(x, y) = M[I(x, y)] \quad (2.2)$$

เมื่อ  $I(x, y)$  เป็นภาพอินพุต  $O(x, y)$  เป็นภาพเอาต์พุต และ  $M[ ]$  เป็นฟังก์ชันในการแทน (Mapping) ข้อมูล โดยจุดภาพที่พิกัดหนึ่งจะถูกปรับปรุงเปลี่ยนแปลงค่าระดับสีเทาด้วยฟังก์ชันที่กำหนดไว้ ค่าระดับสีเทาใหม่ของจุดภาพดังกล่าวจะถูกนำมาแทนค่าลงในตำแหน่งพิกัดเดิมของภาพผลลัพธ์เอาต์พุต หลักการของการประมวลผลแบบจุดภาพดังแสดงในรูปที่ 2.5



รูปที่ 2.5 หลักการประมวลผลแบบจุดภาพ

### 2.4.2 การประมวลผลแบบกลุ่มจุดภาพ (Pixel group processing)

จากการประมวลผลแบบจุดภาพนั้นมักจะเป็นการเปลี่ยนแปลงค่าระดับสีเทาของ จุดภาพ อย่างไม่รู้ทิศทาง วิธีการที่วิธีนี้จะไม่สามารถเปลี่ยนแปลงรายละเอียดทางสเปเชียลในภาพ เนื่องจากเป็นขบวนการแทนค่าแบบจุดภาพต่อจุดภาพของจุดภาพอินพุตกับจุดภาพเอาต์พุต ซึ่งขบวนการดังกล่าวไม่ได้คำนึงถึงจุดภาพละแวกข้างเคียง (Neighbor) กับจุดภาพอินพุต ในขณะที่การประมวลผลแบบกลุ่มจุดภาพจะนำเอาค่าระดับสีเทาของจุดภาพละแวกข้างเคียงมาร่วมพิจารณาและคำนวณค่าใหม่ เพื่อนำมาแทนค่าลงไป ในจุดภาพกึ่งกลางของกลุ่มที่กำลังสนใจอยู่ ลักษณะนี้จะทำให้เกิดการเปลี่ยนแปลงรายละเอียดทางสเปเชียลในภาพ

ถ้าหากบริเวณใดของภาพดิจิทัลที่มีการเปลี่ยนแปลงค่าระดับสีเทาของจุดภาพ หรือความสว่างทันทีทันใด เช่น จากมืดเป็นสว่างหรือจากสว่างเป็นมืด บริเวณดังกล่าวจะให้ความถี่สูง ส่วนบริเวณใดที่ค่าระดับสีเทามีการเปลี่ยนแปลงน้อยหรือให้ค่าความสว่างเกือบเท่ากันตลอดบริเวณนั้น จะให้ความถี่ต่ำ โดยปกติแล้วบริเวณที่เป็นขอบวัตถุ (Edges) ในภาพจะเป็นบริเวณที่มีความถี่สูง ดังนั้นถ้าหากการนำค่าความสว่างของจุดภาพที่สนใจกับจุดภาพในละแวกข้างเคียงมาทำการประมวลผล ซึ่งส่งผลให้ค่าจุดภาพที่สนใจเปลี่ยนแปลงไปย่อมส่งผลให้เกิดการเปลี่ยนแปลงความถี่ของจุดภาพในละแวกนั้น ในการประมวลผลโดยการใช้กลุ่มจุดภาพบางครั้งอาจทำให้ขอบวัตถุคมชัดขึ้น หรือในทางตรงกันข้าม ความคมชัดอาจถูกทำลายไป ลักษณะเช่นนี้เรียกว่าเกิดการกรองทางสเปเชียล (Spatial filtering)

การประมวลผลแบบกลุ่มจุดภาพ จะใช้วิธีการถ่วงน้ำหนัก (Weighted) ให้กับจุดภาพต่าง ๆ ในละแวกข้างเคียง กลุ่มของจุดภาพที่จะถ่วงน้ำหนักนี้เรียกว่า เทมเพลต (Template) หรือเคอร์เนลส์ (Kernels) จะมีขนาด  $3 \times 3$  หรือ  $5 \times 5$  โดยปกติแล้วมักจะใช้เทมเพลตขนาด  $3 \times 3$  ทั้งนี้เพราะถ้าใช้เทมเพลตขนาดใหญ่จะใช้เวลาในการประมวลผลนาน ในการประมวลผลนั้นจะนำเทมเพลตที่เลือกไว้มาทาบบนภาพดิจิทัล โดยให้จุดกึ่งกลางของเทมเพลตทับกับจุดภาพที่สนใจ ซึ่งจุดภาพดังกล่าวจะถูกเปลี่ยนแปลงค่าระดับสีเทาไปตามการถ่วงน้ำหนัก (Weighting) ของจุดภาพในเทมเพลต จากการเลื่อนเทมเพลตไปที่ละจุดภาพจากซ้ายไปขวาและจากบนลงล่าง ลักษณะการทำงานดังกล่าวจะเรียกว่าการทำคอนโวลูชันในทางสเปเชียล (Spatial convolution)

ถ้าหากเทมเพลตขนาด  $3 \times 3$  ที่เลือกใช้มีค่าถ่วงน้ำหนักทั้ง 9 จุด หรือสัมประสิทธิ์ทั้ง 9 ของการทำคอนโวลูชันดังแสดงในรูปที่ 2.6 ในการประมวลผลเพื่อเปลี่ยนแปลงค่าระดับสีเทาที่ตำแหน่ง  $(x, y)$  ใด ๆ ซึ่งอยู่กึ่งกลางของเทมเพลตของการประมวลผลแต่ละครั้ง การทำคอนโวลูชันภาพอินพุต  $I(x, y)$  กับเทมเพลตของรูปที่ 2.6 จะได้ภาพเอาต์พุตที่จุด  $(x, y)$  เป็น  $O(x, y)$  โดย

$$\begin{aligned} O(x, y) = & W(1,1)I(x-1, y-1) + W(1,2)I(x-1, y) + W(1,3)I(x-1, y+1) \\ & W(2,1)I(x, y-1) + W(2,2)I(x, y) + W(2,3)I(x, y+1) \\ & W(3,1)I(x+1, y-1) + W(3,2)I(x+1, y) + W(3,3)I(x+1, y+1) \end{aligned} \quad (2.3)$$

$$O(x, y) = \sum_{i=1}^3 \sum_{j=1}^3 W(i, j)I(x+2-i, y+2-j) \quad (2.4)$$

หรือเขียนเป็นสูตร

W(1,1)	W(1,2)	W(1,3)
W(2,1)	W(2,2)	W(2,3)
W(3,1)	W(3,2)	W(3,3)

รูปที่ 2.6 เเทมเพลตขนาด 3x3

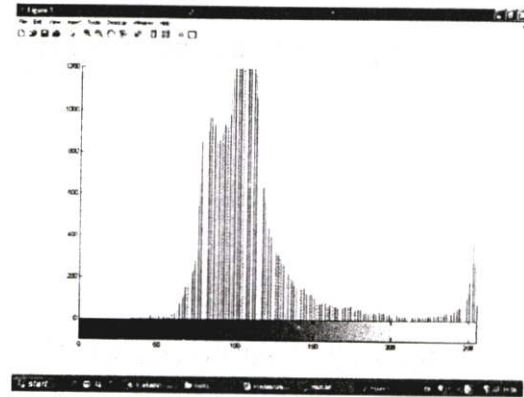
ในการประมวลผลแบบกลุ่มจุดภาพที่ใช้เทมเพลตขนาด 3x3 นั้น ปกติแล้วเส้นภาพเส้นแรกบนสุดและเส้นภาพล่างสุดกับเส้นภาพแนวตั้งแถวแรกกับแถวสุดท้ายจะไม่มีผลการประมวลผล ทั้งนี้เพราะเทมเพลตจะล้นขอบภาพทำให้ประมวลผลไม่ได้ ดังนั้นเพื่อให้ภาพเอาท์พุทขนาดเท่าเดิมจะมีการเติมค่าของจุดภาพให้ขอบทั้งสี่ด้านของภาพเอาท์พุท ในการเติมค่าให้จุดภาพดังกล่าว จะต้องมีการหลักการในการเติมเพื่อให้ได้ภาพผลลัพธ์ที่สมบูรณ์ตามวัตถุประสงค์ของการใช้เทมเพลตนั้น ๆ เช่น ถ้าเทมเพลตที่ใช้เป็นการหาค่าเกรเดียนต์ (Gradient) บริเวณขอบทั้งสี่ด้านของภาพที่ไม่มีผลการประมวลผลจะถูกแทนด้วยค่าศูนย์ เพราะถือว่าบริเวณขอบภาพมีอัตราการเปลี่ยนแปลงค่าระดับสีเทาต่ำ แต่ถ้าเป็นการกรองทางสปีเชียลก็ให้ใส่จุดภาพบริเวณขอบทั้งสี่ด้านของภาพผลลัพธ์ด้วยค่าจุดภาพเดิมของภาพอินพุท อย่างไรก็ตามค่าของจุดภาพที่จะเติมให้ขอบภาพเอาท์พุทก็ไม่มีกำหนดตายตัวว่าเป็นอย่างไรแน่นอน ทั้งนี้ขึ้นอยู่กับความเหมาะสมและความพอใจของผู้ประมวลผลภาพ

## 2.5 ฮิสโตแกรมของภาพ (Image histogram )

ภาพดิจิทัล แต่ละจุดภาพจะถูกเข้ารหัสตามค่าความสว่างของจุดภาพ โดยปกติแล้วแต่ละจุดภาพจะถูกเข้ารหัสไว้ 8 บิต ทำให้มีระดับความสว่างจากดำไปขาวจำนวน 256 ระดับ ความสว่างของจุดภาพค่าต่าง ๆ จะกระจายไปทั่วภาพดิจิทัล ในการตรวจนับจำนวนจุดภาพที่มีระดับความสว่างเท่ากันว่ามีกี่จุดภาพในแต่ละภาพ จะเป็นความถี่ (Frequency) ของระดับความสว่างนั้น ความถี่ของระดับความสว่างทั้ง 256 ระดับจะถูกนำมาเขียนเป็นกราฟแท่ง ซึ่งเรียกกันว่า ฮิสโตแกรม (Histogram) [6, 7] ตัวอย่างภาพถ่ายเปลวเพลิงรูปที่ 2.7 (ก) จะให้ฮิสโตแกรมที่สอดคล้องดังแสดงในรูปที่ 2.7 (ข)



(ก) ภาพถ่ายเปลวเพลิง

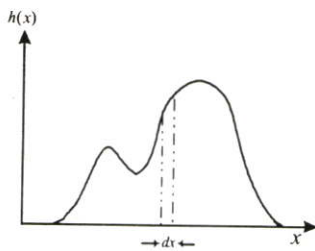


(ข) ฮิสโตแกรมของภาพ (ก)

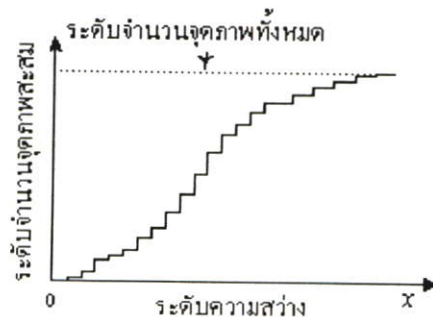
รูปที่ 2.7 ภาพถ่ายเปลวเพลิงกับฮิสโตแกรมที่สอดคล้อง

ฮิสโตแกรมของภาพจะให้ข้อมูลความสว่างของจุดภาพเท่านั้น โดยจะไม่มีข้อมูลสเปซเชียล (Spatial information) หรือตำแหน่งของจุดภาพ ดังนั้นภาพสองภาพที่มีรายละเอียดแตกต่างกัน อาจทำให้ฮิสโตแกรมที่เหมือนกันก็ได้ ประเด็นสำคัญของฮิสโตแกรมคือสามารถมองเป็นการกระจายความน่าจะเป็นแบบดิสครีต (Discrete probability distribution) ถ้าหากนำจำนวนจุดภาพทั้งหมดของภาพไปหารจำนวนจุดภาพของแต่ละระดับความสว่าง ดังนั้นความสูงของแต่ละแท่งของกราฟ จะเป็นตัวบ่งบอกถึงโอกาสในการเกิดค่าความสว่างระดับดังกล่าวของจุดภาพในบริเวณใด ๆ ก็ได้ในภาพ ถ้าหาก  $h(x)$  เป็นฟังก์ชันความน่าจะเป็นของระดับความสว่างและมี  $dx$  เป็นค่าความสว่างที่เพิ่มขึ้น โดยปกติแล้ว  $dx = 1$  หรือบางครั้งอาจจะเขียนได้ว่า  $dx = (L - 1) / L$  ถ้าหาก  $L$  เป็นจำนวนระดับความสว่าง ทั้งหมด

ในทางสถิติแล้ว  $h(x)$  จะเป็นฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function) หรือฮิสโตแกรม ซึ่งจะให้  $\int_0^x h(x) dx$  เป็นฟังก์ชันความน่าจะเป็นสะสม (Cumulative probability function) หรือฮิสโตแกรมรวบรวมดังแสดงในรูปที่ 2.8



(ก) ฮิสโตแกรม



(ข) ฮิสโตแกรมสะสม

รูปที่ 2.8 ฟังก์ชันความหนาแน่นความน่าจะเป็นและความน่าจะเป็นสะสมที่สอดคล้อง

ฮิสโตแกรมของภาพจะมีส่วนช่วยบอกคุณลักษณะของภาพนั้น ๆ อย่างเช่น ถ้าหากฮิสโตแกรมของภาพรวมตัวกระจุกกันทางด้านซ้าย กล่าวคือ จุดภาพส่วนใหญ่มีค่าระดับความสว่างต่ำภาพที่ได้จะมีค่างแสดงในรูปที่ 2.9 ในทางตรงกันข้าม ถ้าหากฮิสโตแกรมของภาพรวมตัวกระจุกกันทางด้านขวา ซึ่งจุดภาพส่วนใหญ่มีค่าระดับความสว่างสูง ภาพที่ได้จะสว่างจ้าค่างแสดงในรูปที่ 2.10 โดยปกติแล้วภาพที่ดีควรมีการกระจายของระดับสีเทาทุก ๆ ระดับตลอดย่านของระดับความสว่างที่เข้ารหัสไว้ค่างแสดงในรูปที่ 2.11 ซึ่งภาพดังกล่าวจะให้คอนทราสต์ในภาพสูง

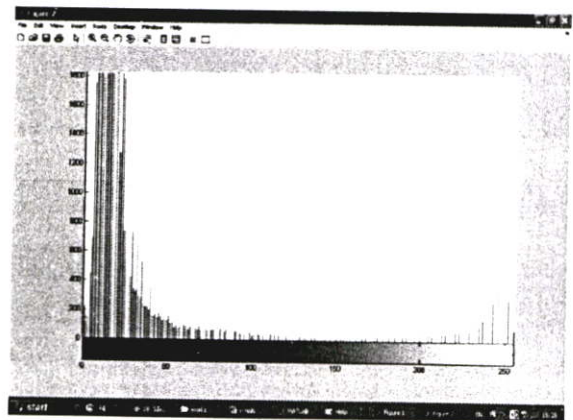
สำหรับค่าคอนทราสต์ของภาพอาจคำนวณได้จากอัตราส่วนในสมการข้างล่าง

$$\text{ค่าคอนทราสต์ในภาพ} = \frac{\text{ค่าระดับความสว่างสูงสุดในภาพ} - \text{ค่าระดับความสว่างต่ำสุดในภาพ}}{\text{จำนวนระดับความสว่างทั้งหมดจากการเข้ารหัส}} \quad (2.5)$$

เนื่องจากการเปลี่ยนแปลงฮิสโตแกรมของภาพมีผลทำให้จุดภาพเกิดการเปลี่ยนแปลงค่าระดับความสว่าง ซึ่งการเปลี่ยนแปลงเกิดขึ้นจากจุดภาพหนึ่งจุดภาพใดโดยไม่เกี่ยวข้องกับค่าระดับความสว่างของจุดภาพข้างเคียง ดังนั้นการปรับปรุงเปลี่ยนแปลงฮิสโตแกรมจะเป็นการประมวลผลภาพแบบจุด สำหรับภาพผลลัพธ์หลังการเปลี่ยนแปลงฮิสโตแกรมอาจจะช่วยให้มองเห็นรายละเอียดในภาพชัดเจนขึ้นหรืออาจจะแย่งก็ได้ ดังนั้นในการประมวลผลภาพวิธีการเบื้องต้นทั่ว ๆ ไปที่รู้จักกันคือ การเปลี่ยนแปลงฮิสโตแกรมของภาพนั่นเอง



(ก) ภาพดิจิทัล

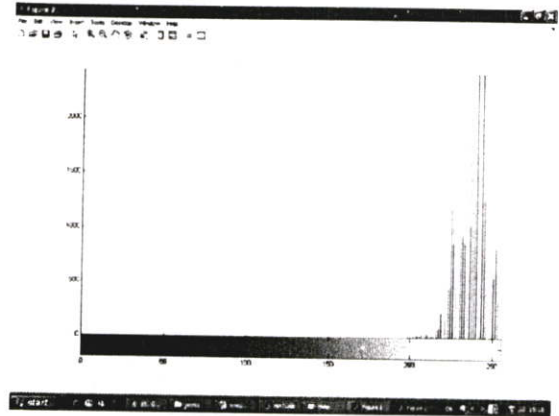


(ข) ฮิสโตแกรมที่สอดคล้อง

รูปที่ 2.9 ภาพที่มีลักษณะมืด เมื่อข้อมูลของระดับสีเทาของจุดภาพส่วนใหญ่ปรากฏอยู่ทางด้านระดับ ความสว่างค่าต่ำ ๆ



(ก) ภาพดิจิทัล

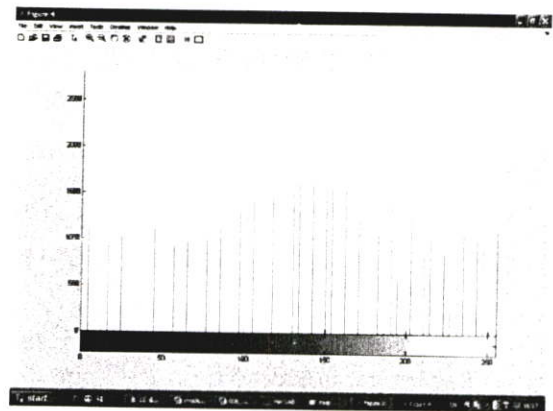


(ข) ฮิสโตแกรมที่สอดคล้อง

รูปที่ 2.10 ภาพที่มีลักษณะสว่าง เมื่อข้อมูลของระดับสีเทาของจุดภาพส่วนใหญ่ปรากฏอยู่ทางด้าน ระดับความสว่างค่าสูง ๆ



(ก) ภาพดิจิทัล



(ข) ฮิสโตแกรมที่สอดคล้อง

รูปที่ 2.11 ภาพที่ดี เมื่อจุดภาพมีค่าระดับความสว่างเกือบทุกระดับ

### 2.5.1 การยืดคอนทราสต์แบบเชิงเส้น (Linear contrast stretch)

จากการที่ค่าระดับสีเทาของจุดภาพต่าง ๆ มีการรวมตัวเป็นกระจุกอยู่ในช่วงแคบ ๆ ช่วงใดช่วงหนึ่งของฮิสโตแกรมดังแสดงในรูปที่ 2.9 หรือรูปที่ 2.10 ภาพเหล่านี้จะให้ความแตกต่างหรือคอนทราสต์ในภาพต่ำยากที่จะทำการจำแนกรายละเอียดในภาพ บางครั้งอาจจะกล่าวได้ว่าภาพดังกล่าวมีพิสัยพลวัต (Dynamic range) แคบ ในการเพิ่มพิสัยพลวัตให้กว้างขึ้น (หรือแคบลงในทางกลับกัน) สามารถทำได้โดยใช้สมการเชิงเส้นที่เรียกกันในเรื่องว่า การยืดคอนทราสต์แบบเชิงเส้น

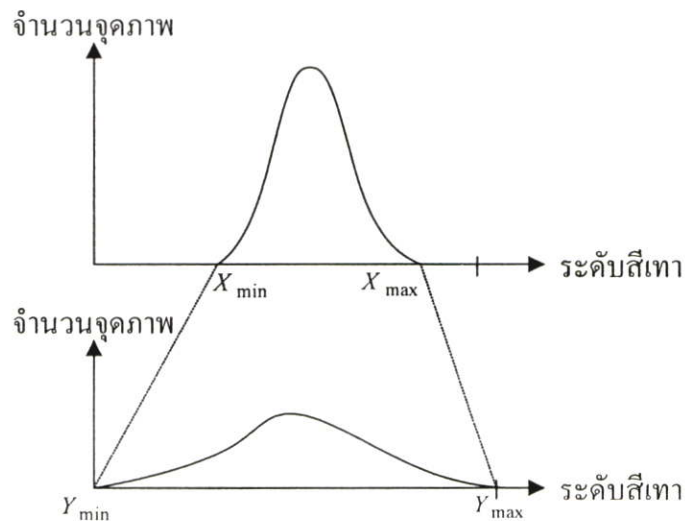
ในการยืดคอนทราสต์แบบเชิงเส้น จะต้องทำการกำหนดค่าสูงสุดและต่ำสุดของค่าระดับสีเทาจากฮิสโตแกรมของภาพต้นฉบับ (Original image) ที่จะยืดไปอยู่ที่ค่าสูงสุดและต่ำสุดของค่าระดับสีเทาในฮิสโตแกรมของภาพใหม่ที่ใด เช่น ถ้ากำหนดให้  $X_{min}$  และ  $X_{max}$  เป็นช่วงของฮิสโตแกรมภาพต้นฉบับจะถูกยืดไปเป็น  $Y_{min}$  และ  $Y_{max}$  ใน ฮิสโตแกรมภาพใหม่ดังแสดงในรูปที่ 2.12 ถ้าหากค่า  $X$  เป็นค่าระดับสีเทาใด ๆ ที่จะทำ การยืด  $[X_{min} \leq X \leq X_{max}]$  เพื่อไม่ให้เกิดปัญหาในการคำนวณ ปกติมักจะกำหนดให้ค่าระดับสีเทา  $X$  ใด ๆ ที่มีค่าต่ำกว่าค่า  $X_{min}$  จะถูก

กำหนดให้ค่าเป็น  $X_{\min}$  และค่าระดับ สีเทา  $X$  ใดๆ ที่มีค่าสูงกว่าค่า  $X_{\max}$  จะถูกกำหนดให้ มีค่าเป็น  $X_{\max}$  ค่าระดับสีเทาใหม่  $Y$  หลังการแปลงคำนวณได้ตามสมการที่ (2.6)

$$Y = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})} (Y_{\max} - Y_{\min}) + Y_{\min} \quad (2.6)$$

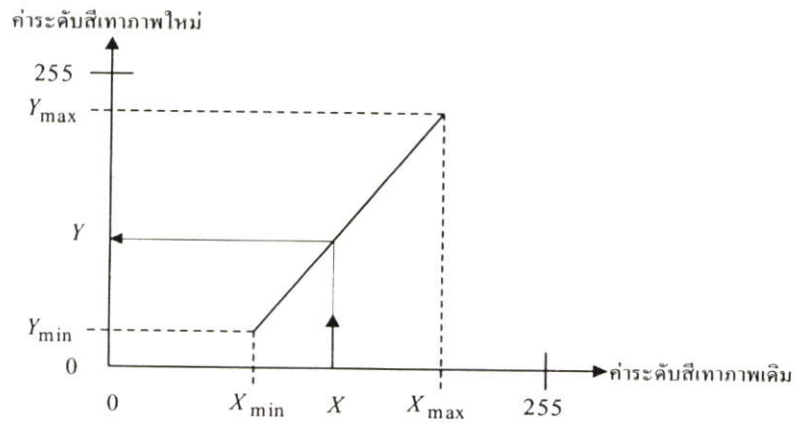
โดยค่า  $Y$  อยู่ระหว่าง  $Y_{\min}$  ถึง  $Y_{\max}$  ค่าระดับสีเทา  $Y$  หลังการแปลงส่วนใหญ่มักจะเป็น จำนวนอิงคระชนี (Floating point number) ถึงแม้ว่าค่าระดับสีเทาเดิม  $X$  จะเป็นจำนวนเต็ม (Integer) ก็ตาม แต่โดยปกติแล้วค่าระดับสีเทาที่ใช้งานกันอยู่ในภาพดิจิทัลมักจะเป็นเลขจำนวนเต็ม ดังนั้นค่าระดับสีเทา  $Y$  ที่คำนวณได้นั้นต้องทำการปัดเศษ (Round off) จึงมีผลทำให้ค่าระดับสีเทาบางค่าจะขาดหายไป ฮิสโตแกรมใหม่หลังการบีบคอนแทรสต์แบบ เชิงเส้นอาจจะขาดความต่อเนื่อง

เนื่องจากการแปลงค่าระดับสีเทาจากการบีบคอนแทรสต์แบบเชิงเส้นนี้เป็นการประมวลผลภาพแบบจุดภาพ เพื่อให้การแปลงทำได้อย่างรวดเร็วมักนิยมนำตารางเปิดดูมาใช้งาน ซึ่งถ้าหากค่าระดับสีเทาเดิมและค่าระดับสีเทาใหม่มีค่าระหว่าง 0 ถึง 255 สูตรที่ใช้ในการคำนวณจากสมการที่ (2.6) จะถูกนำมาใช้คำนวณไม่เกิน 256 ครั้ง [7] จึงทำให้เวลาในการคำนวณที่ต้องใช้ในการแปลงภาพขนาดใหญ่ลดลงได้อย่างมาก



รูปที่ 2.12 การแปลงฮิสโตแกรมโดยการบีบคอนแทรสต์แบบเชิงเส้น

สำหรับฟังก์ชันที่ใช้ในการแปลงฮิสโตแกรมจากการยืดคอนทราสต์แบบเชิงเส้น สามารถแสดงได้ดังรูปที่ 2.13



รูปที่ 2.13 ฟังก์ชันที่ใช้ในการแปลงฮิสโตแกรมของการยืดคอนทราสต์แบบเชิงเส้น

รูปที่ 2.14 เป็นตัวอย่างจากการแปลงค่าระดับสีเทาของภาพโดยการยืดคอนทราสต์แบบเชิงเส้น ซึ่งภาพเดิมในรูปที่ 2.14 (ก) จะมีพิสัยพลวัตของระดับสีเทาแคบ ๆ ทำให้ รายละเอียดในภาพยากต่อการจำแนก หลังการยืดคอนทราสต์จะพบว่าความแตกต่างภายในภาพเด่นชัดขึ้นดังรูปที่ 2.14 (ข)



(ก) ภาพต้นฉบับ

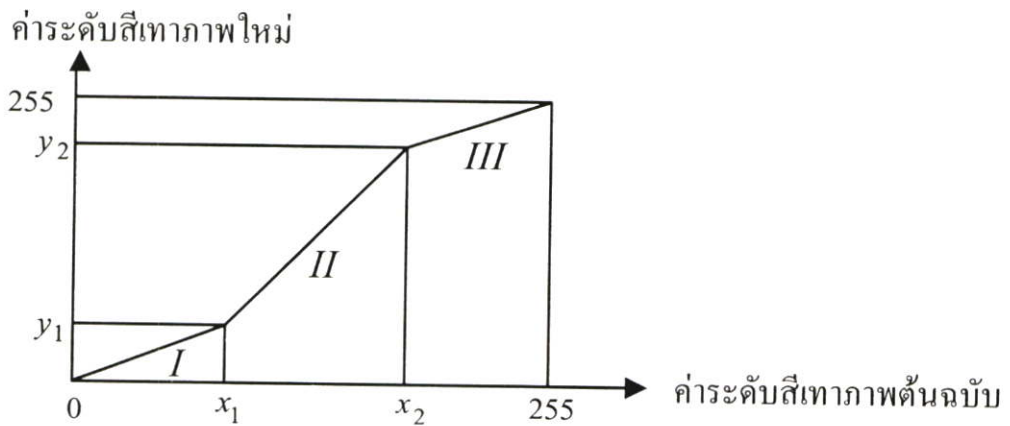


(ข) ภาพใหม่

รูปที่ 2.14 ภาพที่ได้จากการยืดคอนทราสต์แบบเชิงเส้น

2.5.2 การยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง (Piecewise linear contrast stretch)

บางครั้งในการปรับปรุงความแตกต่างภายในภาพโดยการยืดคอนทราสต์แบบเชิงเส้น อาจจะแบ่งช่วงการยืดคอนทราสต์แบบเชิงเส้นได้หลายช่วง ทั้งนี้เพราะค่าระดับสีเทาบางช่วงในภาพต้นฉบับอาจมีความจำเป็นต้องทำให้เกิดความแตกต่างสูง อย่างเช่นในช่วงที่ II ของฮิสโตแกรมรูปที่ 2.15 ในขณะที่ค่าระดับสีเทาในช่วงที่ III ต้องการให้เกิดความแตกต่างน้อยกว่า โดยในแต่ละช่วงสามารถใช้สมการที่ (2.6) ช่วยในการคำนวณค่าระดับสีเทาใหม่ที่ได้จากการแปลง



รูปที่ 2.15 ค่าฟังก์ชันตัวอย่างของการยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง

ตัวอย่างของภาพที่ได้จากการประมวลผลโดยการยืดคอนทราสต์แบบเชิงเส้น ช่วงแคบ ๆ หลาย ๆ ช่วง ดังแสดงในรูปที่ 2.16



(ก) ภาพต้นฉบับ



(ข) ภาพผลลัพธ์

รูปที่ 2.16 ภาพจากการยืดคอนทราสต์แบบเชิงเส้นช่วงแคบ ๆ หลาย ๆ ช่วง

### 2.5.3 การปรับเท่าฮิสโตแกรม (Histogram equalization) [2, 6]

ในการแปลงฮิสโตแกรมถ้าหากว่าค่าระดับสีเทาเดิมก่อนการแปลงคือ  $x$  ซึ่งจะถูกทำให้เป็นบรรทัดฐาน (Normalize) อยู่ในช่วง  $[0,1]$  โดยฟังก์ชันในการแปลงคือ  $T(x)$  หลังการแปลงจะได้ค่าระดับสีเทาใหม่เป็น  $y$  ซึ่งเขียนได้ดังสมการที่ (2.7)

$$y = T(x) \quad (2.7)$$

เงื่อนไขของฟังก์ชันในการแปลงมีอยู่ 2 ประการคือ

ก).  $T(x)$  เป็นผลการแปลงที่ให้ค่าเดียว (Single-valued) และมีการเพิ่มขึ้นในทิศทางเดียว (Monotonically increasing) ภายในช่วง  $0 \leq x \leq 1$

ข)  $0 \leq T(x) \leq 1$  สำหรับ  $0 \leq x \leq 1$

ส่วนการแปลงย้อนกลับ (Inverse transform) จากตัวแปร  $y$  เป็นตัวแปร  $x$  จะได้

$$x = T^{-1}(y) \quad \text{เมื่อ } 0 \leq y \leq 1 \quad (2.8)$$

ซึ่งตัวแปลงผกผันก็ต้องสอดคล้องกับเงื่อนไขทั้ง 2 ประการดังกล่าวเช่นกัน ถ้าให้  $p_x(x)$  และ  $p_y(y)$  เป็นฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function) ของค่าระดับสีเทาดั้งเดิมกับค่าระดับสีเทาหลังการแปลง ให้ความสัมพันธ์ดังสมการที่ (2.9)

$$p_y(y) = \left[ p_x(x) \frac{dx}{dy} \right]_{x=T^{-1}(y)} \quad (2.9)$$

ส่วนฟังก์ชันที่ใช้ในการปรับเท่าฮิสโตแกรมคือ

$$y = T(x) = \int_0^x p_x(w) dw \quad ; 0 \leq x \leq 1 \quad (2.10)$$

เมื่อตัวแปร  $w$  เป็นตัวแปรหุ่น (Dummy variable) เทอมทางด้านขวามือของสมการที่ (2.10) ซึ่งเป็นเทอมการอินทิเกรต (Integrate) นั้นเป็นฟังก์ชันกระจายแบบสะสม (Cumulative Distribution Function) ของตัวแปร  $x$  จากการทำอนุพันธ์สมการที่ (2.11) เทียบกับตัวแปร  $x$  จะได้

$$\frac{dy}{dx} = p_x(x) \quad (2.11)$$

ทำการแทนสมการที่ (2.11) ในสมการที่ (2.9) พบว่า

$$\begin{aligned}
 p_y(y) &= \left[ p_x(x) \cdot \frac{1}{p_x(x)} \right]_{x=T^{-1}(y)} \\
 &= [1]_{x=T^{-1}(y)} \\
 p_y(y) &= 1 \quad \text{เมื่อ} \quad 0 \leq y \leq 1
 \end{aligned} \tag{2.12}$$

ซึ่งพบว่าฟังก์ชันความหนาแน่นความน่าจะเป็นของค่าระดับสีเทา  $y$  ที่ได้หลังการแปลง จะมีลักษณะเป็นความหนาแน่นสม่ำเสมอ (Uniform) ภายในช่วงค่าของตัวแปร  $y$

ในการประมวลผลภาพดิจิทัลนั้น ฟังก์ชันต่าง ๆ จะต้องทำให้อยู่ในรูปดิคริต อย่างเช่น ความน่าจะเป็นของค่าระดับสีเทาที่ระดับ  $k$  จากภาพต้นฉบับเขียนได้เป็น

$$p_x(x_k) = \frac{N_k}{N} \tag{2.13}$$

เมื่อ  $0 \leq x_k \leq 1$  และ  $k = 0, 1, \dots, L-1$  โดย  $L$  เป็นจำนวนระดับสีเทาทั้งหมด  $N_k$  เป็นจำนวนจุดภาพทั้งหมดในภาพที่มีค่าระดับสีเทาเป็น  $k$  ในขณะที่  $N$  เป็นจำนวนจุดภาพ ทั้งหมดในภาพนั้น เมื่อนำค่า  $p_x(x_k)$  ไปทำการพล็อตร่วมกับ  $x_k$  จะได้กราฟ (Graph) ซึ่งเรียกกันว่าฮิสโตแกรมนั่นเอง ในการปรับแก้ฮิสโตแกรมของสมการที่ (2.10) เขียนอยู่ในลักษณะดิคริตได้เป็น

$$\begin{aligned}
 y_k &= T(x_k) = \sum_{i=0}^k \frac{N_i}{N} \\
 &= \sum_{i=0}^k p_x(x_i)
 \end{aligned} \tag{2.14}$$

เมื่อ  $0 \leq x_k \leq 1$  และ  $k = 0, 1, 2, \dots, L-1$

ในการคำนวณหาค่า  $y_k$  จากสมการที่ (2.14) จะได้

$$\begin{aligned}
 y_0 &= \sum_{i=0}^0 p_x(x_i) = p_x(x_0) \\
 &= 0.0225 \\
 y_1 &= \sum_{i=0}^1 p_x(x_i) = p_x(x_0) + p_x(x_1) \\
 &= 0.0225 + 0.2124 = 0.2349
 \end{aligned}$$

สำหรับค่าอื่นคู่ได้จากคอลัมน์ (Column) ที่สี่ของตารางที่ 2.1 เนื่องจากจำนวนของระดับ สีเทา ก่อนทำการแปลงและหลังทำการแปลงจะมีจำนวนเท่ากัน และค่าความห่างของระดับ สีเทา เท่า ๆ กันในกรณีนี้คือห่างกัน  $1/7$  ดังนั้นค่าระดับสีเทา  $y_k$  ที่ได้จะต้องถูกปรับค่าเข้าหาค่าระดับสีเทามาตรฐานที่ใกล้เคียงที่สุด ซึ่งในกรณีนี้ค่าระดับสีเทามาตรฐานคือ  $0, 1/7, 2/7, \dots, 6/7$  และ  $1$  ดังนั้นจากการแปลงฮิสโตแกรมด้วยวิธีการปรับเท่าจะได้ค่าระดับสีเทาใหม่ที่ใกล้เคียงกับระดับมาตรฐาน ดังแสดงในเทอมขวามือของคอลัมน์ที่สี่ของตารางที่ 2.1

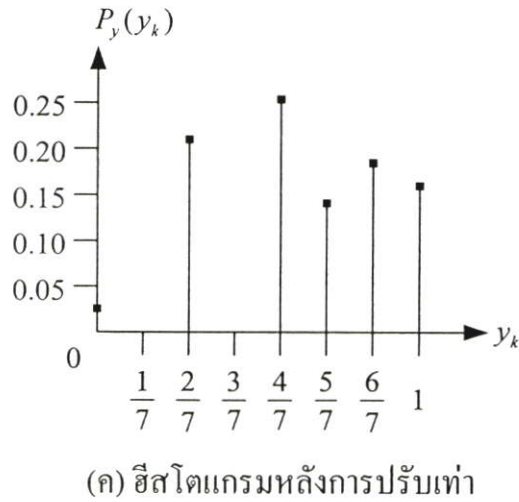
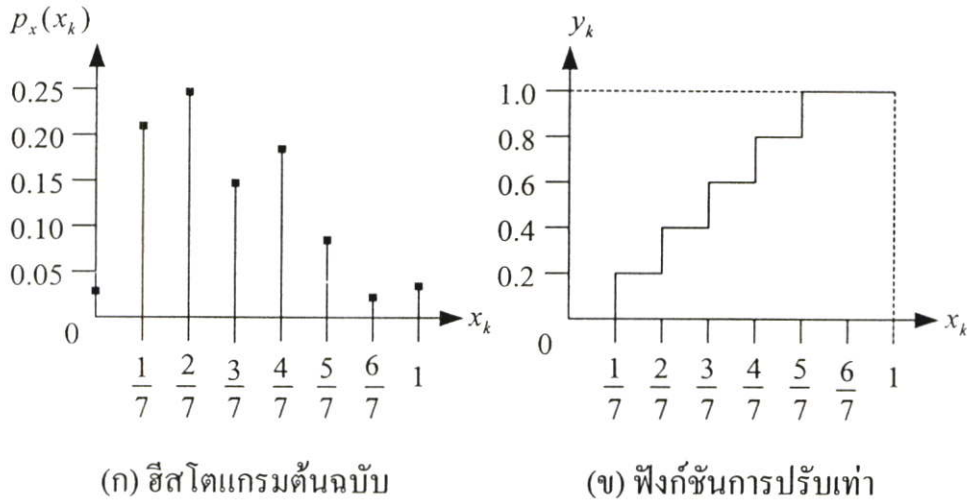
ตัวอย่าง ภาพขนาด  $64 \times 64$  จุดภาพ มีการกระจายของค่าระดับสีเทาดังแสดงใน ตารางที่ 2.1

ตารางที่ 2.1 ลักษณะการกระจายของค่าระดับสีเทาของภาพที่มีจำนวนระดับสีเทา 8 ระดับ

$x_k$	$N_k$	$p_x(x_k) = \frac{N_k}{N}$	$y_k = \sum_{i=0}^k p_x(x_i)$
$x_0 = 0$	92	0.0225	$0.0225 \approx 0$
$x_1 = \frac{1}{7} = 0.1429$	870	0.2124	$0.2349 \approx 2/7$
$x_2 = \frac{2}{7} = 0.2857$	1103	0.2693	$0.5042 \approx 4/7$
$x_3 = \frac{3}{7} = 0.4286$	603	0.1472	$0.6514 \approx 5/7$
$x_4 = \frac{4}{7} = 0.5714$	760	0.1855	$0.8369 \approx 6/7$
$x_5 = \frac{5}{7} = 0.7143$	408	0.0996	$0.9365 \approx 1$
$x_6 = \frac{6}{7} = 0.8571$	116	0.0283	$0.9648 \approx 1$
$x_7 = 1$	144	0.0352	$1.0000 \approx 1$

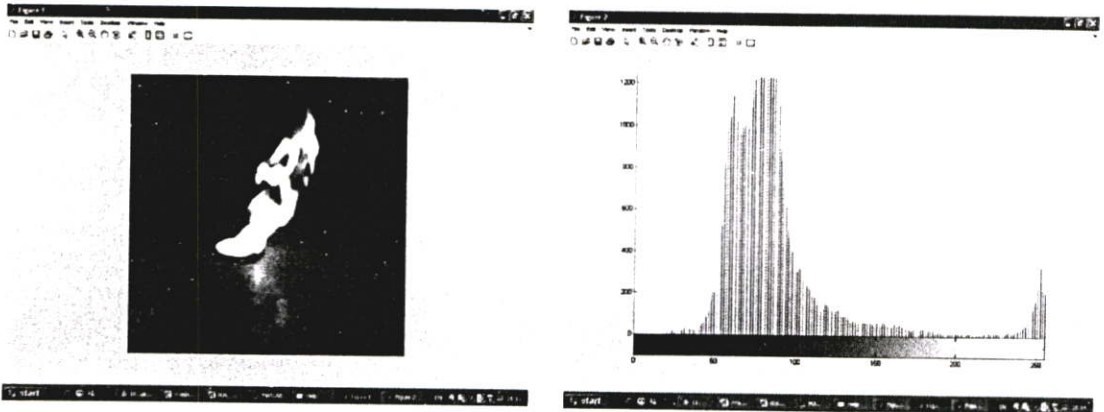
“  $\approx$  ” หมายถึงการประมาณ

จากการปรับเท่าฮิสโตแกรมพบว่า จุดภาพที่มีค่าระดับสีเทา 0 จะยังคงอยู่ที่ 0 ใน ฮิสโตแกรมใหม่ จุดภาพที่มีค่าระดับสีเทาเดิมเป็น  $1/7$  จะถูกแมป (Map) ไปอยู่ที่ค่าระดับ สีเทา  $2/7$  ของฮิสโตแกรมใหม่ จุดภาพที่มีค่าระดับสีเทาเดิมเป็น  $2/7, 3/7$  และ  $4/7$  จะถูกแมปไปอยู่ที่ค่าระดับสีเทา  $4/7, 5/7$  และ  $6/7$  ในฮิสโตแกรมใหม่ ส่วนค่าระดับสีเทาเดิมที่เป็น  $5/7, 6/7$  และ  $1$  จะถูกแมปรวมไปอยู่ที่ค่าระดับสีเทา 1 ในฮิสโตแกรมใหม่ ดังนั้นจุดภาพที่มีค่าระดับสีเทา 1 ในฮิสโตแกรมใหม่จะมีจำนวน  $408+116+144 = 668$  จุดภาพ หรือจะมีค่าความน่าจะเป็นรวมกันเท่ากับ  $0.1631$  สำหรับฮิสโตแกรมต้นฉบับและฮิสโตแกรมใหม่ที่ได้หลังการปรับเท่าฮิสโตแกรมดังแสดง ในรูปที่ 2.17

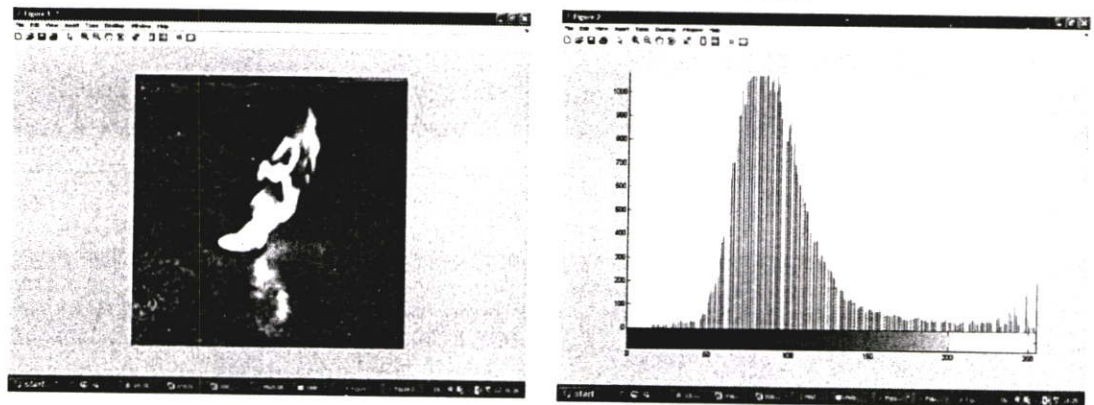


รูปที่ 2.17 การปรับเท่าฮิสโตแกรมของตารางที่ 2.1

เมื่อทำการประยุกต์การปรับเท่าฮิสโตแกรมให้กับภาพในรูปที่ 2.18 (ก) ก็จะได้ภาพดังรูปที่ 2.18 (ข) ซึ่งจะพบว่ารายละเอียดของภาพจะเด่นชัดขึ้น



(ก) ภาพต้นฉบับและฮิสโตแกรมที่สอดคล้อง



(ข) ภาพหลังการปรับเท่าฮิสโตแกรมและฮิสโตแกรมที่สอดคล้อง

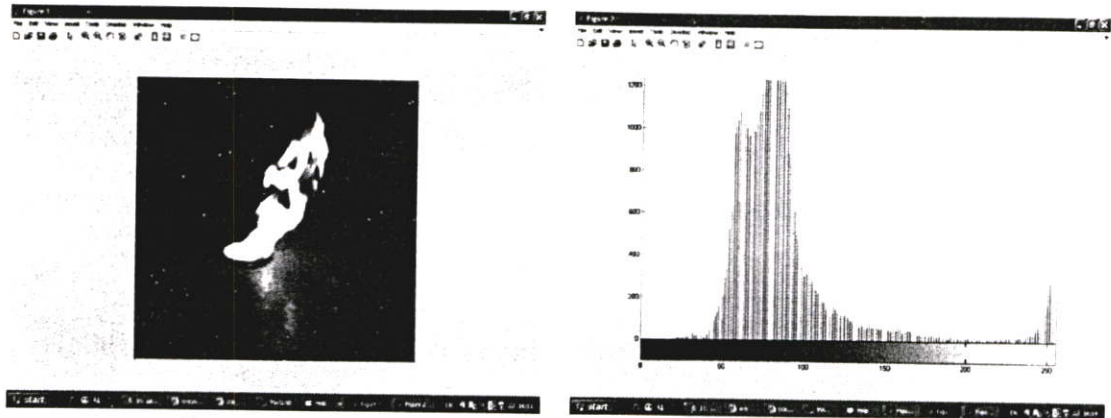
รูปที่ 2.18 การปรับเท่าฮิสโตแกรมภาพ

### 2.5.4 ขีดเริ่มเปลี่ยนฮิสโตแกรม (Histogram thresholding)

วิธีการปรับปรุงคุณภาพนี้จะเป็นการแปลงภาพระดับสีเทาให้เป็นภาพไบนารี (Binary) ซึ่งจะให้จุดภาพมีค่าระดับสีเทาเพียงสองระดับ กล่าวคือถ้าหากค่าระดับสีเทาของจุดภาพมีค่าที่ต่ำกว่าหรือเท่ากับค่าขีดเริ่มเปลี่ยน (Thresholding) จุดภาพนั้นจะถูกกำหนดค่าใหม่ให้เป็น 0 (จะเป็นจุดภาพที่มีสีดำ) ในทางตรงกันข้ามถ้าหากค่าระดับสีเทาของจุดภาพมีค่าสูงกว่าค่าของขีดเริ่มเปลี่ยน จุดภาพนั้นก็จะถูกกำหนดค่าใหม่ให้เป็น 255 (จะเป็นจุดภาพ ที่มีสีขาว) สมการที่ใช้ในการแปลงคือ

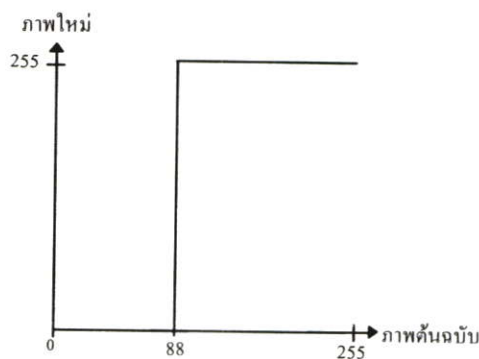
$$y = \begin{cases} 0 & ; x \leq T \\ 255 & ; x > T \end{cases} \quad (2.15)$$

เมื่อ  $x$  คือค่าระดับสีเทาของจุดภาพต้นฉบับและ  $y$  เป็นค่าระดับสีเทาที่ผ่านการแปลง โดย  $T$  เป็นค่าขีดเริ่มเปลี่ยน ตัวอย่างของภาพที่ได้จากการแปลงโดยใช้ขีดเริ่มเปลี่ยนฮิสโตแกรม ดังแสดงในรูปที่ 2.19

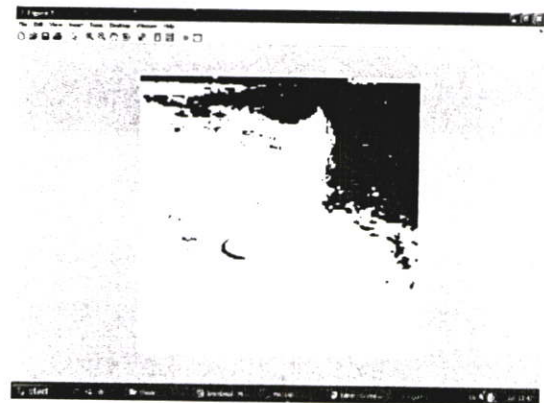


(ก) ภาพต้นฉบับ (Original Image)

(ข) ฮิสโตแกรมของภาพ (ก)



(ค) ฟังก์ชันในการแปลงเมื่อกำหนดขีดเริ่มเปลี่ยนฮิสโตแกรมไว้ที่ 88



(ง) ภาพไบนารีที่ได้หลังการแปลงด้วยขีดเริ่มเปลี่ยนฮิสโตแกรม

รูปที่ 2.19 การแปลงภาพไบนารีโดยใช้ขีดเริ่มเปลี่ยนฮิสโตแกรม

## 2.6 การตรวจจับขอบในภาพ (Edge detection)

ขอบต่าง ๆ ภายในภาพเกิดจากการเปลี่ยนแปลงของค่าระดับสีเทาแบบทันทีทันใดจากค่าระดับสีเทาดำ ๆ ไปเป็นค่าระดับสีเทาสูง ๆ หรือในทางตรงกันข้ามจะเปลี่ยนจากค่าระดับสีเทาสูง ๆ ไปเป็นค่าระดับสีเทาดำ ๆ ดังนั้นการตรวจจับขอบจึงเป็นการหาค่าความลาดชัน (Slope) ระหว่างจุดภาพ โดยทั่วไปมักเรียกว่า ค่าเกรเดียนท์ (Gradient) ขอบต่าง ๆ มักจะเป็นโครงร่างของวัตถุภายในภาพ โครงร่างต่าง ๆ ภายในภาพจะถูกนำไปใช้ประโยชน์ในการวิเคราะห์ตรวจสอบการเคลื่อนไหวของวัตถุจากเฟรม (Frame) หนึ่งไปยังอีกเฟรมหนึ่ง หรืออาจถูกนำไปใช้งานการรู้จำวัตถุ (Object recognition) เป็นต้น

### 2.6.1 การใช้เทมเพลตในการตรวจจับขอบ

เนื่องจากการตรวจจับขอบภายในภาพจะเป็นการหาค่าความแตกต่างของจุดภาพจากจุดภาพข้างเคียงว่ามีมากน้อยเพียงไร เพื่อไม่ให้เกิดการไบแอส (Bias) ขึ้นในพื้นที่เอกพันธ์ (Homogeneous) เทมเพลตที่นำมาใช้ตรวจจับขอบจึงมีเงื่อนไขว่าผลรวมของน้ำหนักถ่วงในแต่ละเทมเพลตจะต้องเป็นศูนย์ ดังนั้นถ้าหากนำไปประยุกต์กับข้อมูลภาพแล้วบริเวณใดที่มีค่าระดับสีเทาเท่ากันหมดจะให้ค่าผลลัพธ์หลังการประมวลผลเป็นศูนย์ (ตรวจสอบดูได้จากรูปที่ 2.20 และ 2.21) เทมเพลตที่ใช้ในการตรวจจับขอบบางครั้งเรียกว่า เทมเพลตเกรเดียนต์ (Gradient template) ค่าผลลัพธ์ที่ได้จากการคำนวณนี้จะเป็นค่าเกรเดียนต์ของจุดภาพนั้น

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

(ก) เทมเพลต

0	-1	0
-1	4	-1
0	-1	0

(ข) เทมเพลต

1	-2	1
-2	4	-2
1	-2	1

(ค) เทมเพลต

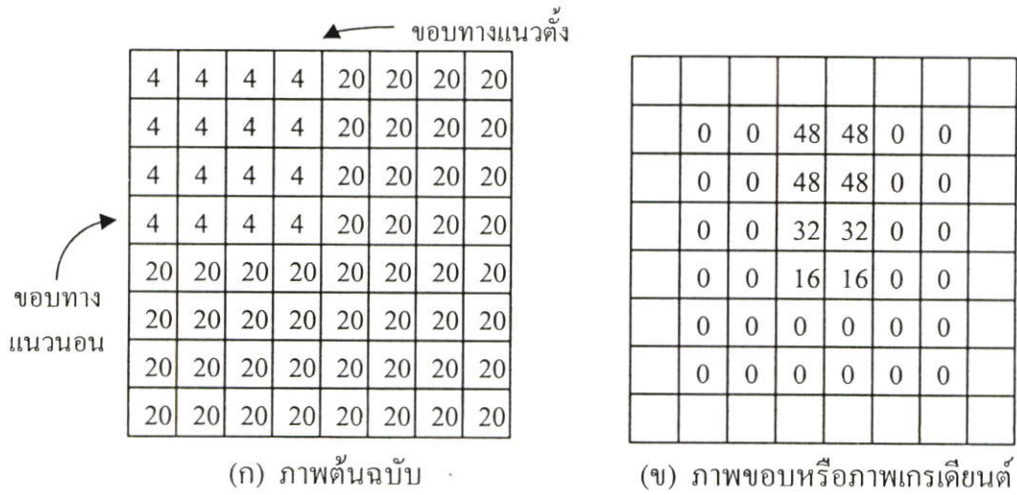
รูปที่ 2.20 เทมเพลตเกรเดียนต์ของลาปลาเซียนแบบต่าง ๆ

เทมเพลตเกรเดียนต์ที่นำมาใช้ในการตรวจจับขอบจะแบ่งเป็นสองกลุ่มคือ กลุ่มที่ไม่คำนึงถึงทิศทางของขอบในภาพ (Omni directional) อย่างเช่น เทมเพลตของ Laplacian ดังแสดงในรูปที่ 2.20 เทมเพลตนี้จะถูกนำไปทำคอนโวลูชันกับภาพโดยใช้สมการที่ (2.4) โดยเทมเพลตจะถูกแทนลงในเทอม  $W(i, j)$  ของสมการดังกล่าว เทมเพลตดังกล่าวจะถูกนำมาทาบบกับข้อมูลภาพ โดยให้จุดกึ่งกลางของเทมเพลตทับลงบนจุดที่ต้องการคำนวณหาค่า เกรเดียนต์

ทิศทางของขอบ	Prewitt	Kirsh	สามระดับ	ห้าระดับ
↓ เหนือ	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$
↖ ตะวันตกเฉียงเหนือ	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}$
← ตะวันตก	$\begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & -1 \end{bmatrix}$	$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$
↙ ตะวันตกเฉียงใต้	$\begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$
↓ ใต้	$\begin{bmatrix} -1 & -1 & -1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$
↘ ตะวันออกเฉียงใต้	$\begin{bmatrix} -1 & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$
↗ ตะวันออกเฉียง	$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$
↖ ตะวันออกเฉียงเหนือ	$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$

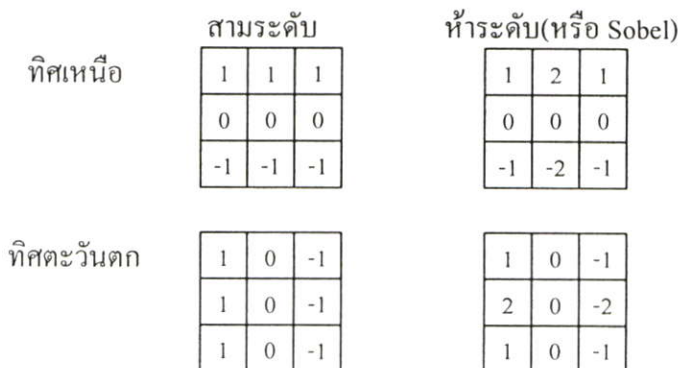
รูปที่ 2.21 เทมเพลต 8 ทิศทาง [14]

เทมเพลตเกรเดียนต์อีกกลุ่ม จะเป็นกลุ่มที่คำนึงถึงทิศทาง (Compass) [13] เช่น เทมเพลตของ Prewitt ของ Kirsch เป็นต้น ดังแสดงในรูปที่ 2.21 เทมเพลตแต่ละแบบจะมี 8 ทิศทาง เทมเพลตแต่ละทิศทางจะถูกนำมาคำนวณหาค่าเกรเดียนต์ว่าเทมเพลตทิศไหนให้ค่าเกรเดียนต์สูงสุดก็จะเลือกค่าเกรเดียนต์จากเทมเพลตนั้น ตัวอย่างเช่น การประยุกต์ใช้ เทมเพลตทิศตะวันตกแบบสามระดับ ในรูปที่ 2.21 กับข้อมูลภาพในรูปที่ 2.22 (ก) จะได้ค่าเกรเดียนต์ในรูปที่ 2.22 (ข) ซึ่งเทมเพลตนี้จะให้ค่าเกรเดียนต์ของขอบทางแนวตั้งมีค่าสูง ส่วนขอบทางแนวนอนจะพบว่าเทมเพลตนี้ให้ค่าเป็นศูนย์



รูปที่ 2.22 การประยุกต์ใช้เทมเพลตทิศตะวันตกแบบสามระดับกับข้อมูลภาพ

ในการใช้เทมเพลตทั้งหมด 8 ทิศทางจะต้องใช้เวลาในการคำนวณสูงมาก ดังนั้น เพื่อลดเวลาในการคำนวณจึงมีการนำเอาเทมเพลตเพียงสองทิศที่ตั้งฉากกัน (Orthogonal template) มาคำนวณหาค่าเกรเดียนต์ร่วมกัน ซึ่งปกติมักจะเลือกเทมเพลตทิศเหนือกับทิศตะวันตกดังแสดงในรูปที่ 2.23 เทมเพลตทั้งสองทิศจะให้ค่าเกรเดียนต์ที่ประมาณจากอนุพันธ์พาร์เชียล (Partial derivative) ตามทิศทางของขอบวัตถุ



รูปที่ 2.23 เทมเพลตสองทิศทางที่นิยมใช้กัน

ในภาพ โดยเทมเพลตทิศเหนือจะให้ค่าเกรเดียนต์สูงสำหรับบริเวณที่มีขอบทางแนวนอน และเทมเพลตทิศตะวันตกจะให้ค่าเกรเดียนต์สูงสำหรับบริเวณที่มีขอบทางแนวตั้ง ส่วนบริเวณที่ขอบเป็นเส้นเฉียงก็จะให้ค่าเกรเดียนต์จากการเฉลี่ยของเทมเพลตทั้งสอง สำหรับทิศทางของขอบหรือเกรเดียนต์นี้จะได้จากค่าอาร์คแทนเจน (Arctangent) ของค่าเกรเดียนต์จากเทมเพลตทั้งสอง

เทมเพลตห้าระดับทั้งสองเทมเพลตในรูปที่ 2.23 นั้น บางครั้งเรียกว่าเทมเพลตของ Sobel จะเป็นที่ยอมรับใช้ในการคำนวณหาค่าเกรเดียนต์กันมาก ทั้งนี้เพราะนำไปใช้สร้างเป็นฮาร์ดแวร์ได้ง่าย โดยค่าถ่วงน้ำหนักให้จุดภาพจะเป็น 1 และ 2 ถ้าตัวถ่วงน้ำหนักเป็น 1 ก็ใช้ค่าของข้อมูลภาพนั้นเลย แต่ถ้าตัวถ่วงน้ำหนักเป็น 2 ก็ให้ทำการเลื่อนบิตค่าของข้อมูลจุดภาพไปทางซ้ายมือหนึ่งบิต (Shift left)

ถ้าหากกำหนดให้เทมเพลตทิศเหนือในรูปที่ 2.23 เป็น  $W_h(i, j)$  และเทมเพลตทิศตะวันตกเป็น  $W_v(i, j)$  เทมเพลตทั้งคู่จะถูกนำมาทำคอนโวลูชันกับข้อมูลภาพต้นฉบับ  $I(x, y)$  โดยจะให้ค่าเกรเดียนต์ของขอบในแนวนอนและแนวตั้งเป็น  $G_h(x, y)$  กับ  $G_v(x, y)$  ดังสมการที่ (2.16) และ (2.17) ตามลำดับ

$$G_h(x, y) = \sum_{i=1}^3 \sum_{j=1}^3 W_h(i, j) I(x+2-i, y+2-j) \quad (2.16)$$

$$G_v(x, y) = \sum_{i=1}^3 \sum_{j=1}^3 W_v(i, j) I(x+2-i, y+2-j) \quad (2.17)$$

ค่าเกรเดียนต์เฉลี่ยของภาพเอทพุท  $O(x, y)$  ได้จากสมการที่ (2.18) คือ

$$O(x, y) = \sqrt{[G_h(x, y)]^2 + [G_v(x, y)]^2} \quad (2.18)$$

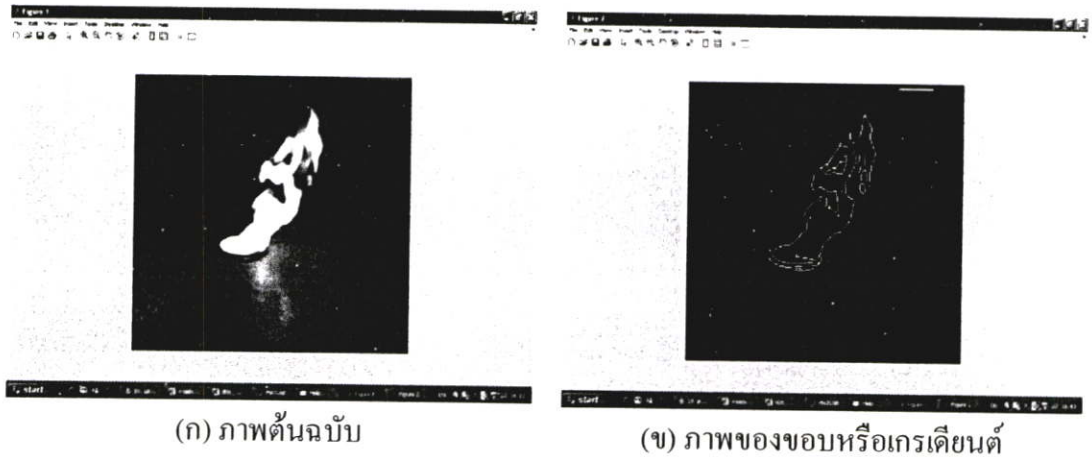
ในบางครั้งเพื่อลดเวลาการคำนวณลง ค่าเกรเดียนต์ผลลัพธ์  $O(x, y)$  จะคำนวณจากผลรวมของค่าสัมบูรณ์ (Absolute value) ของค่าเกรเดียนต์ทั้งสอง

$$O(x, y) = |G_h(x, y)| + |G_v(x, y)| \quad (2.19)$$

สำหรับทิศทางของขอบหรือเกรเดียนต์คือ  $\theta(x, y)$  นั้นคำนวณได้ดังนี้

$$\theta(x, y) = \tan^{-1} \left[ \frac{G_h(x, y)}{G_v(x, y)} \right] \quad (2.20)$$

ตัวอย่างของการประยุกต์ใช้เทมเพลตทั้งสองของ Sobel กับภาพรูปที่ 2.24(ก) จะได้ภาพที่เป็นค่าของขอบหรือภาพเกรเดียนต์ ดังแสดงในรูปที่ 2.24(ข)



รูปที่ 2.24 การสร้างภาพของขอบหรือเกรเดียนต์

## 2.7 การปรับเรียบให้ภาพ (Image smoothing)

เนื่องจากภาพที่จัดเก็บมาได้ อาจจะมีสัญญาณรบกวนแบบสุ่ม (Random noise) มาปรากฏซ้อนทับ (Superimpose) บนความสว่างหรือค่าระดับสีเทาของจุดภาพ ซึ่งสัญญาณรบกวนดังกล่าว บางครั้งเกิดจากตัวจัดเก็บข้อมูลภาพหรือเกิดขึ้นในระหว่างการส่งข้อมูลภาพผ่านระบบการสื่อสาร ขบวนการปรับเรียบให้ภาพจึงถูกนำมาประยุกต์ใช้เพื่อกำจัดสัญญาณรบกวนที่แปลกปลอมเข้ามาอยู่ในภาพ เนื่องจากสัญญาณรบกวนมักมีสเปกตรัมอยู่ที่ความถี่สูง ๆ ดังนั้นการกำจัดสัญญาณรบกวนโดยการปรับเรียบให้ภาพก็คือการกรองที่ให้ความถี่ต่ำผ่านนั่นเอง ในการกำจัดสัญญาณรบกวนทำได้ทั้งในโดเมนสเปเชียล (Spatial domain) และในโดเมนความถี่ (Frequency domain) ในที่นี้จะกล่าวถึงเฉพาะการกำจัดสัญญาณรบกวนด้วยการปรับเรียบให้ภาพทางด้านโดเมนสเปเชียลเท่านั้น

### 2.7.1 การปรับเรียบด้วยค่าเฉลี่ย

การกำจัดสัญญาณรบกวนจากการปรับเรียบด้วยค่าเฉลี่ยนี้ จะเป็นการนำเอาค่าระดับสีเทาของจุดภาพข้างเคียงของจุดภาพที่ต้องการประมวลผลมาหาค่าเฉลี่ย เพื่อจะนำไปแทนค่าระดับสีเทาของจุดภาพนั้น จากการนำค่าระดับสีเทาของจุดภาพข้างเคียงมาใช้ในการปรับเรียบให้ภาพบางครั้ง อาจจะเรียกขบวนการนี้ว่า การเฉลี่ยค่าในละแวกข้างเคียง (Neighborhood averaging) จำนวนจุดภาพในละแวกข้างเคียงที่ใช้โดยมากมักจะเป็น เทมเพลตจัตุรัสขนาด  $3 \times 3$  จุดภาพหรือ  $5 \times 5$  จุดภาพ ปกติแล้วนิยมใช้เทมเพลตขนาด  $3 \times 3$  จุดภาพเพื่อลดเวลาในการคำนวณ การปรับเรียบด้วยค่าเฉลี่ยจะเป็นการประมวลผลแบบกลุ่มจุดภาพ ซึ่งจะอยู่ในลักษณะการทำคอนโวลูชันระหว่างภาพกับเทมเพลตที่ใช้ เพื่อบริหารค่าระดับสีเทาใหม่ของจุดภาพตำแหน่งเดิม ในการปรับเรียบ

ค่าเฉลี่ยนี้จะนำเทมเพลตที่ใช้ไปทาบบนภาพอินพุทหรือภาพต้นฉบับ ซึ่งจุดกึ่งกลางของเทมเพลตจะถูกลงบนจุดภาพที่สนใจที่จะถูกทำการปรับแก้ค่าระดับสีเทา สมการที่ใช้สำหรับเทมเพลตขนาด 3x3 จุดภาพคือ

$$S(x, y) = \sum_{i=1}^3 \sum_{j=1}^3 W(i, j)I(x+2-i, y+2-j) \quad (2.21)$$

เมื่อ  $W(i, j)$  เป็นเทมเพลตที่ใช้ และ  $I(x, y)$  เป็นภาพอินพุท ส่วน  $S(x, y)$  เป็นภาพผลลัพธ์จากการปรับแก้ด้วยค่าเฉลี่ย

ลักษณะเทมเพลตที่ใช้อาจจะแบ่งเป็นสองลักษณะดังแสดงในรูปที่ 2.25 โดยในรูปที่ 2.25(ก) เป็นการคำนวณค่าเฉลี่ยที่ให้ความสำคัญของจุดภาพข้างเคียงเท่ากันหมดทุกจุด ส่วนเทมเพลตในรูปที่ 2.25 (ข) จะเป็นการให้ความสำคัญของจุดภาพที่อยู่ใกล้กับจุดกึ่งกลางที่มีความสำคัญสูงกว่าจุดภาพที่อยู่ห่างไกลออกไป สำหรับเลข 9 ที่เป็นตัวหารในเทมเพลต รูปที่ 2.25(ก) และเลข 16 ที่เป็นตัวหารในเทมเพลต รูปที่ 2.25(ข) จะเป็นค่าถ่วงน้ำหนักเพื่อรักษาคุณสมบัติเอกพันธ์ (Homogeneity) ของพื้นที่ในภาพ กล่าวคือ ถ้าจุดภาพที่ต้องการทำการประมวลผลมีค่าระดับสีเทาเท่ากับค่าระดับสีเทาของจุดภาพต่าง ๆ ในระนาบข้างเคียง การประมวลผลจะต้องไม่ไปปรับเปลี่ยนค่าของจุดภาพที่สนใจนั้น

$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

(ก) ความสำคัญเท่าเทียมกัน

$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$
$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$
$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

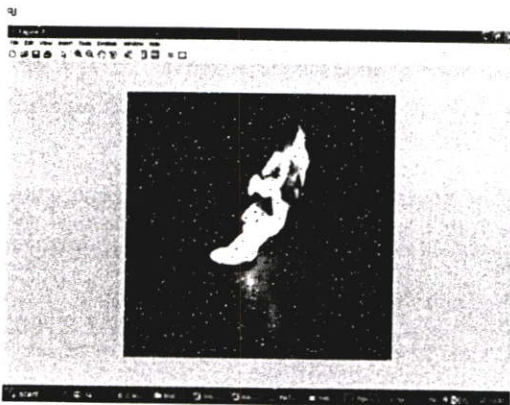
(ข) ความสำคัญขึ้นอยู่กับระยะห่าง

รูปที่ 2.25 เทมเพลต  $W(i, j)$  ที่ใช้ในการปรับแก้ภาพด้วยค่าเฉลี่ย

ภาพตัวอย่างในรูปที่ 2.26(ก) เป็นภาพอินพุทที่มีสัญญาณรบกวนปรากฏอยู่ในภาพ หลังการใช้เทมเพลต รูปที่ 2.25(ก) ในการปรับแก้ด้วยค่าเฉลี่ยจะให้ผลดังรูปที่ 2.26 (ข) ซึ่งพบว่าสัญญาณรบกวนจะถูกกำจัดหายไป แต่ขณะเดียวกันขอบต่าง ๆ ในภาพจะถูกทำลายไปด้วย ขอบต่าง ๆ เหล่านี้จะเป็นข้อมูลที่มีความถี่สูง ในการป้องกันหรือลดการ สูญเสียรายละเอียดที่มีความถี่สูงนั้นสามารถทำได้ โดยการประยุกต์ใช้ค่าขีดเริ่มเปลี่ยนกับ สมการที่ (2.22) กล่าวคือ

$$N(x, y) = \begin{cases} S(x, y) & ; |I(x, y) - S(x, y)| > T \\ I(x, y) & ; \text{อื่นๆ} \end{cases} \quad (2.22)$$

ขบวนการของสมการที่ (2.22) นี้จะแทนค่าระดับสีเทาที่ตำแหน่ง  $(x, y)$  ของภาพผลลัพธ์  $N(x, y)$  ด้วยค่า  $S(x, y)$  ถ้าหากค่าระดับสีเทาภาพเดิม  $I(x, y)$  กับค่าระดับสีเทาจากการปรับค่าด้วยการเฉลี่ย  $S(x, y)$  นั้นให้ผลต่างสัมบูรณ์มากกว่าค่าขีดเริ่มเปลี่ยน  $T$  แต่ถ้าหากไม่ใช่ก็ให้ค่าระดับสีเทาของภาพเดิมเหมือนเดิมคือ  $I(x, y)$  จากการใช้สมการที่ (2.22) กับภาพในรูปที่ 2.26 (ก) จะได้ภาพผลลัพธ์ในรูปที่ 2.26 (ข) ซึ่งพบว่าความคมชัดในภาพที่ได้จะดีกว่าภาพผลลัพธ์ในรูปที่ 2.26 (จ)



(ก) ภาพที่มีสัญญาณรบกวนเป็นจุดๆ



(ข) ภาพหลังการใช้เทมเพลตในรูปที่ 2.23(ก)



(ค) ภาพหลังการใช้เทมเพลตรวมกับขีดเริ่มเปลี่ยน เมื่อกำหนดให้  $T = 30$   
รูปที่ 2.26 การปรับเรียบด้วยค่าเฉลี่ย

## บทที่ 3

# โครงข่ายประสาทเทียม

### 3.1 บทนำ

หลายศตวรรษมาแล้วที่มนุษย์ได้ประดิษฐ์และพัฒนาเครื่องจักรขึ้นมา เราจะเห็นว่าเครื่องจักรเหล่านี้จะถูกใช้สำหรับงานที่มีความยุ่งยากซับซ้อนและเป็นงานที่ซ้ำๆ ซากๆ ซึ่งถ้าใช้คนทำอาจทำให้เกิดความผิดพลาดได้ง่าย ในสมัยแรกๆ อาจเริ่มจากเครื่องจักรที่มีการทำงานที่ไม่ยุ่งยากซับซ้อน เช่น กาน, รอก เป็นต้น และต่อมาได้ประดิษฐ์สิ่งที่มีความซับซ้อนขึ้นมาเรื่อยๆ เช่น เครื่องจักรกล, หุ่นยนต์, และคอมพิวเตอร์ เป็นต้น จนกระทั่งปัจจุบันวิศวกรและนักวิทยาศาสตร์แขนงต่างๆ ได้พยายามที่จะพัฒนาเครื่องคอมพิวเตอร์ให้มีความฉลาดยิ่งขึ้น โดยพยายามให้เครื่องคอมพิวเตอร์มีการเรียนรู้และสามารถตัดสินใจได้เองหรือเป็นส่วนช่วยในการตัดสินใจขั้นต้นได้ ระบบแบบจำลองโครงข่ายเซลล์ประสาทหรือโครงข่ายประสาทเทียม(Artificial Neural Network: ANN) เป็นตัวอย่างหนึ่งที่จะทำให้เครื่องคอมพิวเตอร์ต่างๆ มีศักยภาพเพิ่มมากขึ้น

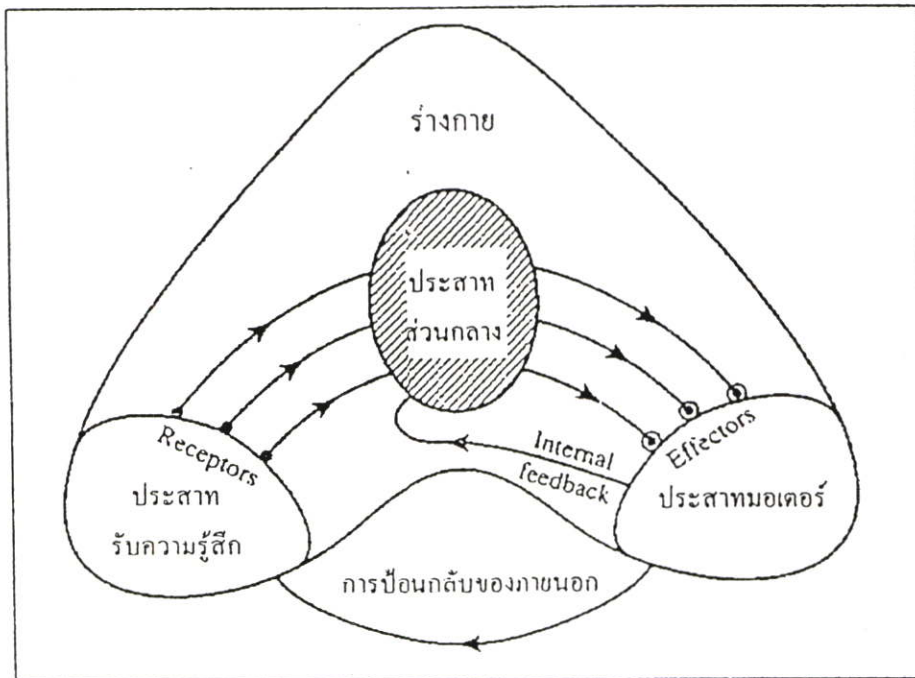
เมื่อเราพิจารณาถึงมนุษย์ และสัตว์ จะเห็นว่าสามารถรู้จักภาพและวิเคราะห์แยกแยะเสียงได้ดีกว่าคอมพิวเตอร์ที่ทันสมัยที่สุดแม้เป็นคอมพิวเตอร์ที่ถูกสร้างขึ้นมาทางด้านนี้โดยเฉพาะก็ตาม และนักวิจัยในสาขานี้ได้คาดหวังกันว่าโครงข่ายประสาทเทียม จะเป็นคลื่นลูกใหม่ที่ใช้ในการประมวลผลข่าวสารสำหรับใช้แทนหรือสนับสนุนการคำนวณแบบเดิมโครงข่ายประสาทเทียม จะช่วยเสริมความสามารถของคอมพิวเตอร์สำหรับการประมวลผลในระบบใหญ่ๆ ซึ่งระบบนี้สามารถเรียนรู้และตัดสินใจสำหรับกระบวนการของระบบนั้นๆ ได้

สำหรับการทำงานของโครงข่ายประสาทเทียมจะถอดแบบมาจากการทำงานของระบบสมองมนุษย์ คือจะมีการส่งผ่านข้อมูลกัน โดยการเชื่อมต่อของโหนดกันเป็นโครงข่ายร่างแหอย่างหนาแน่น และมีการทำงานในลักษณะขนาน ซึ่งในหัวข้อถัดไปจะอธิบายลักษณะการทำงานของระบบเซลล์ประสาทของมนุษย์เพื่อจะได้เทียบเคียงการทำงานและการทำให้สามารถเข้าใจการทำงานของโครงข่ายประสาทเทียมได้ดีขึ้น

### 3.2 ความรู้พื้นฐานทางกายภาพของระบบเซลล์ประสาท

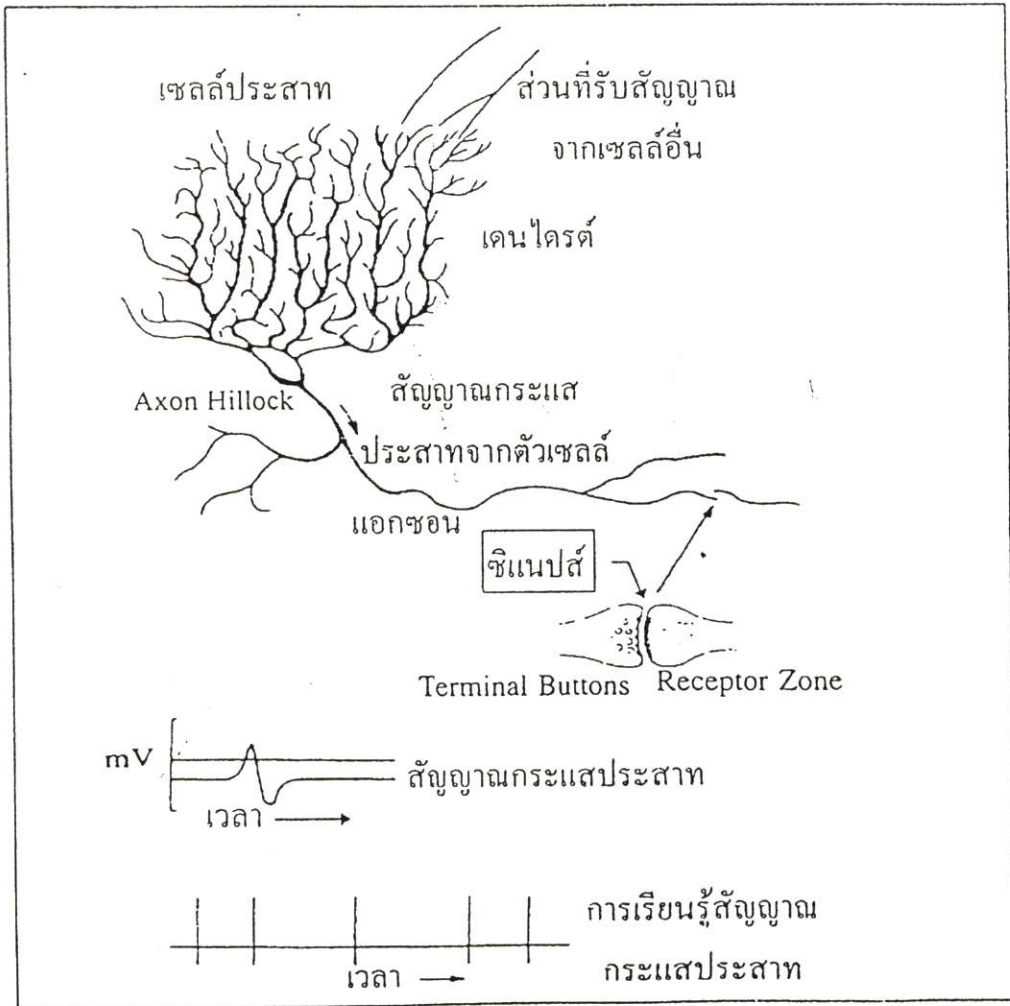
ระบบประสาทมีหน้าที่ควบคุมรักษาสมดุลภายในร่างกายให้คงที่ ระบบประสาทของมนุษย์เป็นระบบโครงสร้างที่ซับซ้อนซึ่งจะประกอบด้วยเซลล์ประสาท (Neurons) ประมาณ  $10^{11}$  เซลล์ และเซลล์ประสาทแต่ละเซลล์จะเชื่อมต่อกับเซลล์ประสาทตัวอื่นๆ อีกประมาณ  $10^4$  เซลล์ ระบบประสาทมีอยู่ 3 ส่วนด้วยกัน (รูปที่ 3.1) คือ

1. ระบบประสาทส่วนกลาง (Central Nervous System)
2. ระบบประสาทรับความรู้สึก (Sensory Systems)
3. ระบบประสาทมอเตอร์ (Motor Systems)



รูปที่ 3.1 การสื่อสารข้อมูลในระบบประสาท [8]

จากรูปที่ 3.1 ระบบประสาทส่วนกลางจะทำหน้าที่วิเคราะห์ข้อมูลต่างๆ ที่รับมาจากระบบประสาทรับความรู้สึกที่ประกอบด้วยเซลล์ประสาทรับความรู้สึก (Sensory Receptor) ซึ่งจะรับรู้เกี่ยวกับการเปลี่ยนแปลงสิ่งแวดล้อมรอบตัว และจะส่งข่าวสารไปยังระบบประสาทส่วนกลาง และเมื่อประสาทส่วนกลางประมวลผลแล้วก็จะส่งสัญญาณไปควบคุมสั่งการระบบประสาทมอเตอร์ของอวัยวะแสดงผล (Effectors) ให้ทำงาน ซึ่งจะทำให้เกิดการตอบสนองต่อการเปลี่ยนแปลงทั้งภายในและภายนอกทั่วร่างกาย



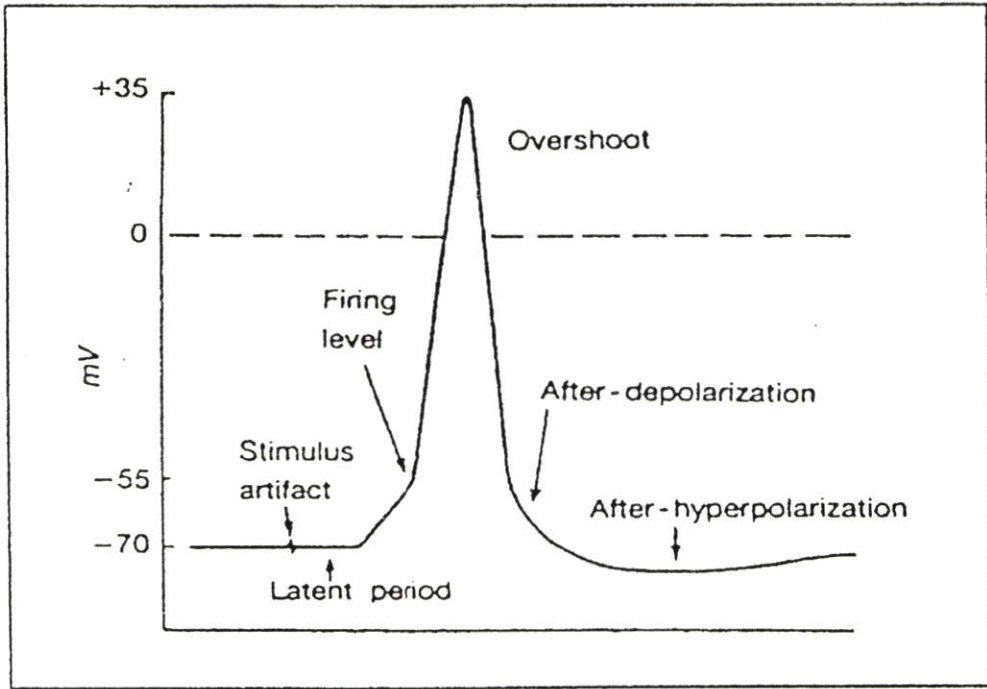
รูปที่ 3.2 โครงสร้างของเซลล์ประสาทและการส่งสัญญาณกระแสประสาท [8]

เซลล์ประสาทแต่ละเซลล์ประกอบด้วย ตัวเซลล์ (Cell Body) และแขนงที่แยกออกจากตัวเซลล์ ซึ่งได้แก่ แอกซอน (Axon) และ เดนไดรต์ (Dendrite) (รูปที่ 3.2) แอกซอนมีจุดเริ่มต้นจากบริเวณส่วนที่หนาที่สุดของเซลล์ประสาทที่เรียกว่า Axon Hillock เซลล์ประสาทแต่ละเซลล์จะมีแอกซอนเพียงหนึ่งอันปลายของแอกซอนจะแยกเป็นแขนงหลายเส้นและไปสิ้นสุดที่ Terminal Buttons และที่จุดสิ้นสุดนี้จะมีตัวรับกระแสประสาท (Nerve Impulse) ที่เรียกว่า Receptor Zone ซึ่งต่ออยู่กับ เดนไดรต์ของเซลล์ประสาทตัวอื่น กระแสประสาทหรือข้อมูลที่อยู่ในรูปของการเปลี่ยนแปลงทางไฟฟ้าเคมี จะไปกระตุ้นให้ตัวเซลล์ส่งสัญญาณไปตามแอกซอน ดังนั้นแอกซอนจะทำหน้าที่นำสัญญาณกระแสประสาทออกจากตัวเซลล์ไปสู่เซลล์ประสาทตัวอื่น และเดนไดรต์จะทำหน้าที่รับกระแสประสาทจากเซลล์อื่นเข้าสู่ตัวเซลล์

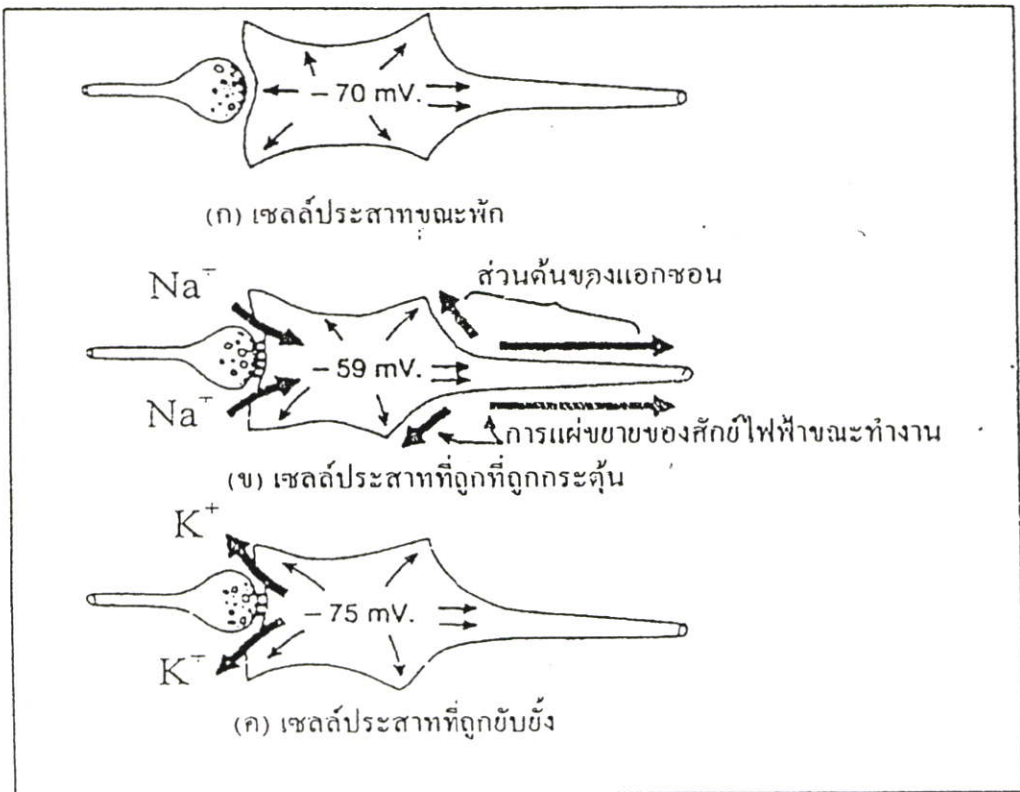
คุณสมบัติของเซลล์ประสาทโดยทั่วไปคือ จะมีความไวต่อการถูกกระตุ้น เมื่อเซลล์ประสาทถูกกระตุ้นจะทำให้เกิดการเปลี่ยนแปลงสมมูลของไอออนบริเวณ Receptor Zone ซึ่งจะทำให้ซึ่งจะทำให้เกิดความต่างศักย์ขึ้น และเมื่อความต่างศักย์ของจุดนี้สูงกว่าเทรชโฮลด์ก็จะทำให้เกิดการส่งสัญญาณกระแสประสาทไปยังเซลล์อื่นๆ

ความต่างศักย์ของเยื่อหุ้มเซลล์ประสาทขณะพักประมาณ  $-60$  ถึง  $-90$  mV ถ้ามีการกระตุ้นเซลล์ประสาทจะมีการแก้โพลาไรซ์ (Depolarization) ของเยื่อหุ้มเซลล์ และหลังจากแก้โพลาไรซ์จนความต่างศักย์เพิ่มขึ้นเป็น  $15$  mV แล้ว จะทำให้อัตราแก้โพลาไรซ์เพิ่มขึ้นอย่างรวดเร็ว (รูปที่ 3.3) เมื่อความต่างศักย์เพิ่มขึ้นจนถึงค่าสูงสุด ( $\approx 35$  mV) จะทำให้อัตราการแก้โพลาไรซ์ลดลงอย่างรวดเร็ว ( $\approx 70\%$ ) และค่อยๆ ลดลงกลับสู่สภาวะปกติ หลังจากเข้าสู่สภาวะปกติแล้วความต่างศักย์จะลดลงต่ำกว่าระดับปกติซึ่งจะเรียกช่วงนี้ว่า After Hyper polarization ถ้าดับการเปลี่ยนแปลงความต่างศักย์ทั้งหมดนี้เรียกว่า ศักย์ไฟฟ้าขณะทำงาน (Action Potential)

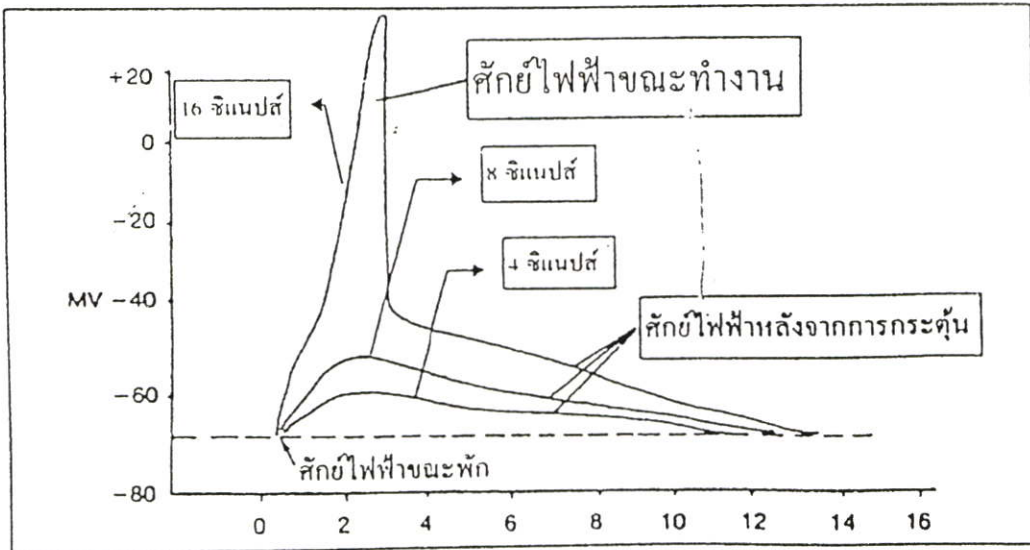
กระแสประสาทที่ส่งผ่านจากเซลล์หนึ่งไปยังอีเซลล์หนึ่ง จะต้องผ่านรอยต่อระหว่าง Terminal Buttons และ Receptor Zone ซึ่งเรียกรอยต่อนี้ว่า ซิแนปส์ (Synapse) การส่งผ่านกระแสประสาทกันระหว่างเซลล์ (Neurotransmission) ซิแนปส์จะปล่อยสารเคมีเพื่อไปกระตุ้น (Excitatory) สำหรับบางกรณี และในบางกรณีก็ปล่อยสารเคมีไปเพื่อยับยั้ง (Inhibitory) และในขณะใดขณะหนึ่งซิแนปส์จะยอมให้เกิดขึ้นเพียงกรณีเดียวเท่านั้น คือถ้าไม่กระตุ้นก็จะปล่อยสารเพื่อยับยั้งการส่งผ่านกระแสประสาท ในรูปที่ 3.4 แสดงสภาวะต่างๆ ของเซลล์ประสาท ในกรณีที่ซิแนปส์เกิดสภาวะการกระตุ้นความต่างศักย์ระหว่างผนังเซลล์จะลดลงเป็น  $-58$  mV จะทำให้เกิดการโพลาไรซ์ขึ้น ซึ่งถ้าสูงพอถึงภาวะ Firing Level ก็จะทำให้เกิดศักย์ไฟฟ้าขณะทำงานขึ้น และการที่จะเกิดสภาวะเช่นนี้ขึ้นได้ต้องอาศัยซิแนปส์ สิบถึงหลายร้อยตัวในเวลาเดียวกันโดยกระบวนการ Summation ซึ่งซิแนปส์เพียงตัวเดียวจะไม่สามารถกระตุ้นเซลล์ประสาทให้ถึงภาวะ Firing Level ได้ ในรูปที่ 3.5 แสดงศักย์ไฟฟ้าขณะทำงานที่เกิดขึ้นจากการรวมกันของซิแนปส์จำนวนต่างๆ กัน



รูปที่ 3.3 แสดงลักษณะการเกิดศักย์ไฟฟ้าขณะทำงาน [9]



รูปที่ 3.4 แสดงภาวะต่างๆ ของเซลล์ประสาท



รูปที่ 3.5 แสดงศักย์ไฟฟ้าขณะทำงานที่เกิดจากการรวมกันของซินแนปส์

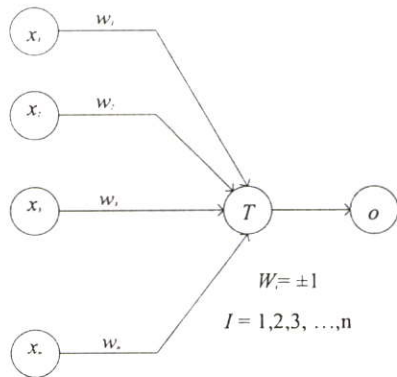
ในสภาวะกระตุ้นจะทำให้เกิดการส่งผ่านสัญญาณกระแสประสาท ไอออนของโซเดียม ( $\text{Na}^+$ ) โปรตัสเซียม ( $\text{K}^+$ ) และ คลอไรด์ ( $\text{Cl}^-$ ) จะซึมผ่านเยื่อหุ้มเซลล์ได้มากขึ้นส่วนในสภาวะยับยั้ง โปรตัสเซียม และ คลอไรด์ จะไหลออกมาจากเซลล์และโซเดียมไม่มีการไหลเข้าเซลล์ซึ่งจะเป็นเหตุให้ประจุบวกภายในเซลล์ลดจำนวนลงทำให้ความต่างศักย์ภายในเซลล์เป็นลบมากขึ้น (ประมาณ -75 mV) ซึ่งจะส่งผลให้เซลล์ประสาทถูกกระตุ้นยากกว่าปกติ คือต้องใช้ความต่างศักย์เพิ่มขึ้นเป็น 16 mV ซึ่งสามารถกระตุ้นเซลล์ประสาทให้ถึง Firing Level ได้ ซึ่งระยะนี้จะเรียกว่าระยะการดื้อ (Refractory Period)

### 3.3 แบบจำลองโครงข่ายเซลล์ประสาท

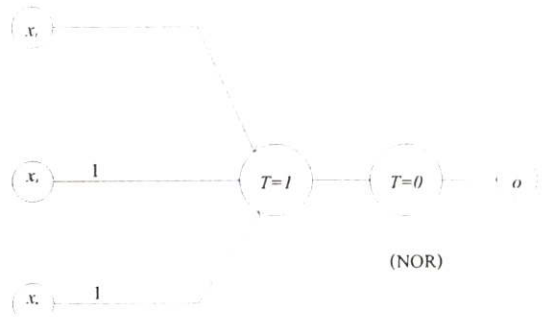
มาพิจารณากันถึงระบบการทำงานของแบบจำลองเซลล์ประสาทที่ใช้ในการประมวลผล โดยใช้คอมพิวเตอร์ที่นำไปใช้ควบคุมระบบต่างๆ แบบจำลองที่จะกล่าวมาเป็นแบบแรกในที่นี้ถูกเสนอโดย McCulloch และ Pitts ในปี ค.ศ. 1943 ซึ่งแบบจำลองของเซลล์ประสาทแสดงในรูปที่ 3.6 ก อินพุต  $x_i$  (สำหรับ  $i = 1, 2, 3, \dots, n$ ) จะมีค่า  $\{0, 1\}$  ซึ่งจะขึ้นอยู่กับสัญญาณอินพุตสัญญาณจากเซลล์อื่นในขณะนั้นว่าจะมีหรือไม่มีสัญญาณ ส่วนสัญญาณที่จะส่งต่อไปยังเซลล์ถัดไปซึ่งเป็นเซลล์แสดงผล(เซลล์ของผลลัพธ์ จะแทนด้วย  $o$ ) และ Firing level ของแบบจำลองนี้ถูกกำหนดโดย

$$o^{k+1} = \begin{cases} 1 & \text{ถ้า } \sum_{i=1}^n w_i x_i^k \geq T \\ 0 & \text{ถ้า } \sum_{i=1}^n w_i x_i^k < T \end{cases} \quad (3.1)$$

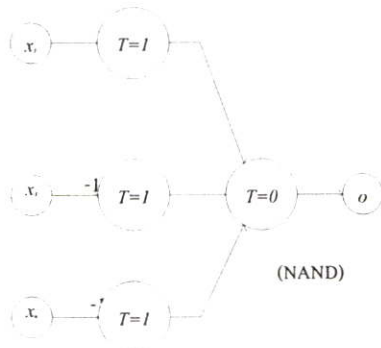
โดยที่  $k = 0, 1, 2, \dots$  เป็นช่วงเวลาแบบไม่ต่อเนื่อง  $w_i$  เป็นค่าถ่วงน้ำหนักที่เชื่อมต่อกับอินพุตที่  $i$  ซึ่งถ้า  $w_i = +1$  แสดงถึงสถานะการกระตุ้นของซินแนปส์ และถ้า  $w_i = -1$  ซินแนปส์จะมีการยับยั้งการส่งผ่านสัญญาณและ  $T$  เป็นค่าความต่างศักย์เทรชโฮลด์หรือขีดเริ่มเปลี่ยน ซึ่งถ้าค่าผลรวมของผลคูณระหว่างค่าถ่วงน้ำหนักกับสัญญาณอินพุตที่จะต้องมากกว่า  $T$  ซึ่งจะมีสัญญาณผ่านไปยังเซลล์อื่นได้



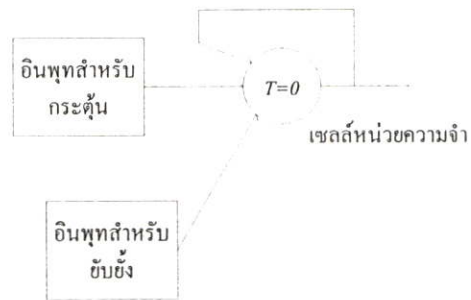
(ก) ไดอะแกรมแบบจำลอง



(ข) เกต NOR



(ค) เกต NAND



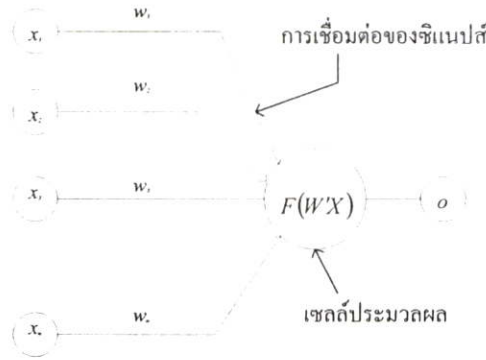
(ง) เซลล์หน่วยความจำ

**รูปที่ 3.6** แบบจำลองของเซลล์ประสาทของ McCulloch – Pitts และ โครงข่ายลอจิกพื้นฐาน

แบบจำลองของเซลล์ประสาทของ McCulloch – Pitts แม้จะเป็นโครงข่ายแบบง่าย ๆ แต่แบบจำลองนี้สามารถแทนตัวดำเนินการทางลอจิกได้ คือตัวดำเนินการ NOT, OR, และ AND เมื่อเรากำหนดค่าถ่วงน้ำหนักและค่าเทรชโฮลด์ที่เหมาะสมให้กับโครงข่าย ในรูปที่ 3.6 (ข) และ 3.6 (ค) เป็นแบบจำลองที่ใช้แทนตัวดำเนินการซึ่งเป็นเกตที่ใช้ในพีชคณิตบูลีนคือ NOR และ NAND ซึ่งมีอินพุตสามตัว และในรูปที่ 3.6 ง เป็นเซลล์ที่ใช้แทนหน่วยความจำซึ่งเป็นโครงข่ายที่มีการป้อนกลับจากเอาต์พุต

โครงข่ายประสาทเทียมที่จะกล่าวต่อไปนี้เป็นแบบจำลองที่ใช้เป็นต้นแบบในการศึกษาแบบจำลองและระบบการเรียนรู้ชนิดต่างๆ ที่จะกล่าวถึงต่อไป ซึ่งในโครงข่ายนี้จะคล้ายกับ

แบบจำลองของ McCulloch – Pitts ส่วนที่ต่างกันคือ ค่าของตัวแปรต่างๆ ที่ใช้ในแบบจำลอง เนื่องจากค่าของตัวแปรในแบบจำลองของโครงข่ายประสาทเทียมเป็นเลขจำนวนจริง และค่าถ่วงน้ำหนักที่ได้จากการเรียนรู้ ซึ่งแบบจำลองนี้จะแสดงในรูปที่ 3.7



รูปที่ 3.7 แบบจำลองโครงข่ายเซลล์ประสาทเทียม

จากรูปที่ 3.7 แสดงโครงข่ายการเชื่อมต่อของแบบจำลองเซลล์ประสาทที่สามารถสอนให้โครงข่ายตัดสินใจได้ โดยมี  $x_i$  เป็นสัญญาณอินพุต  $w_i$  เป็นค่าถ่วงน้ำหนักที่ได้จากการสอน โครงข่ายและในแต่ละโหนดในโครงข่ายจะใช้แทนเซลล์ประสาทในแต่ละเซลล์ ซึ่งบางครั้งจะเรียกว่าหน่วยประมวลผลพื้นฐาน (Process Element Unit) และมีซิแนปส์ซึ่งจะเชื่อมต่อโหนดเพื่อใช้ในการส่งผ่านสัญญาณ การกระตุ้นหรือยับยั้งสัญญาณจะขึ้นอยู่กับค่าถ่วงน้ำหนัก  $w_i$  และสำหรับสัญญาณเอาต์พุตสามารถคำนวณได้ดังนี้

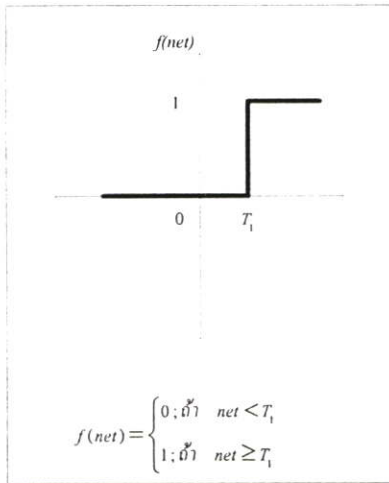
$$o = f(WX) \quad (3.2)$$

โดยที่  $W$  เป็นเวกเตอร์ของค่าถ่วงน้ำหนักซึ่งสามารถกำหนดได้ดังนี้

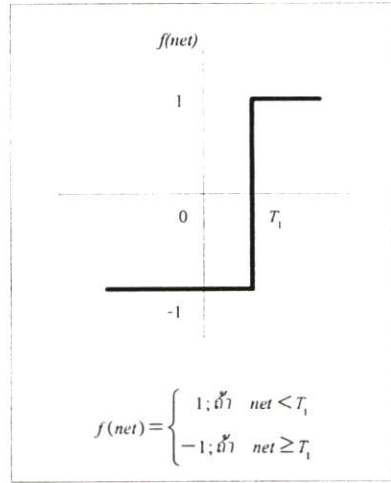
$$W \equiv [w_1, w_2, w_3, \dots, w_n]$$

และ  $X$  เป็นเวกเตอร์อินพุต  $X \equiv [x_1, x_2, x_3, \dots, x_n]$  เมื่อ  $i$  เป็นตัวดำเนินการทรานสโพสค์ของเมตริกซ์ ฟังก์ชันกำหนดสัญญาณเอาต์พุตในสมการที่ (3.2) ถูกเรียกว่าฟังก์ชันการเร่งเร้าหรือแอกติเวชันฟังก์ชัน (Activation Function) ซึ่งมีคุณสมบัติคล้ายกับกราฟของศักย์ไฟฟ้าขณะทำงานในรูปที่ 3.3 Activation Function [10-13] มีด้วยกันสองชนิดคือ ชนิดที่เป็นเชิงเส้นและชนิดที่ไม่เป็นเชิงเส้นดังรูปที่ 3.8

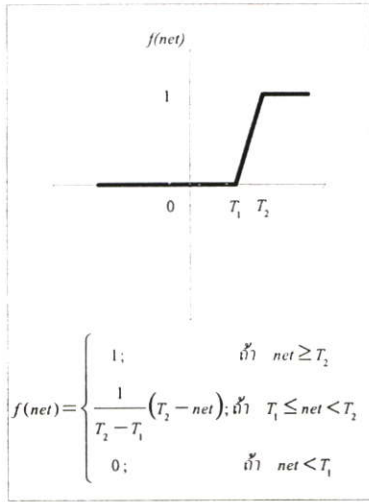
กำหนดให้  $net = W'X = \sum_{i=1}^n w_i x_i$  และ  $T_1, T_2$  เป็นค่าเทรชโฮลด์



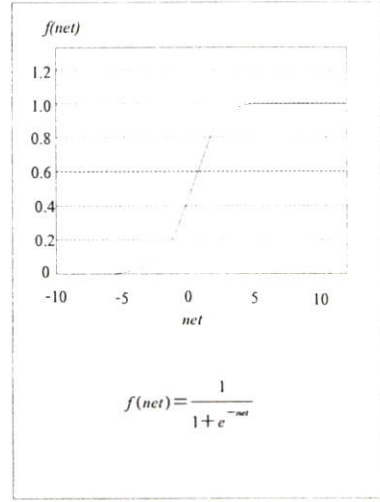
(ก) Threshold Logic



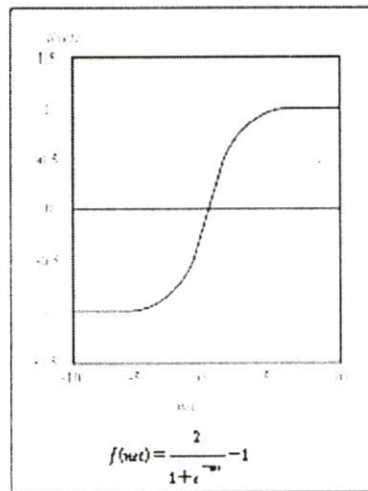
(ข) Bipolar Binary



(ค) Linear Threshold



(ง) Sigmoid



(จ) bipolar Continuous

รูปที่ 3.8 Activation Function ชนิดต่างๆ

สำหรับ Activation Function ในรูปที่ 3.8 ที่เป็นแบบไม่ต่อเนื่อง (รูปที่ 3.8 ก-ค) และแบบต่อเนื่อง (รูปที่ 3.8 ง-จ) ค่าที่ได้จากฟังก์ชันเหล่านี้มีสองลักษณะ [8] คือ สองสภาพขั้วหรือไบโพลาร์ (Bipolar) หรือโดเมนของ  $f(net)$  จะอยู่ในช่วงปิด  $[0,1]$  การที่จะกำหนด Activation Function ใดคุณสมบัติที่ดีที่สุดคงเป็นไปได้ เนื่องจากโครงข่ายประสาทเทียมไม่มีโครงข่ายที่แน่นอน นั่นคือโครงข่ายหนึ่งจะใช้กับปัญหาใดปัญหาหนึ่งเท่านั้น ดังนั้นการเลือก Activation Function ใดต้องพิจารณาให้เหมาะสมกับปัญหานั้นๆ

### 3.4 การเรียนรู้ของโครงข่าย

การเรียนรู้ของโครงข่ายประสาทเทียมจะมีประสิทธิภาพเพียงใดนั้นขึ้นอยู่กับค่าถ่วงน้ำหนักของโครงข่าย ซึ่งการสอน (Training) โครงข่ายก็คือค่าถ่วงน้ำหนักที่เหมาะสมให้แก่โครงข่ายนั้นๆ วิธีการสอนโครงข่ายประสาทเทียมมีอยู่ด้วยกันสองแบบ คือ

#### 3.4.1 การสอนแบบชี้แนะหรือดูแล (Supervised Learning)

การสอนโดยวิธีนี้จะกำหนดเซตของการสอนให้กับโครงข่าย ซึ่งเซตนี้ประกอบด้วยอินพุตและเอาต์พุตที่ต้องการ (Output Desired) เมื่อป้อนอินพุตให้กับโครงข่าย โครงข่ายก็จะทำการประมวลผลจนได้คำตอบและค่าถ่วงน้ำหนักออกมาชุดหนึ่ง สำหรับคำตอบที่ได้จากโครงข่ายจะถูกนำมาคำนวณค่าความผิดพลาด โดยวัดเป็นระยะทางว่ามีความห่างจากคำตอบที่ต้องการของอินพุตในชุดเดียวกันมากน้อยเท่าไร ถ้ายังมีความผิดพลาดสูงอยู่ก็จะมี การปรับค่าถ่วงน้ำหนัก และทำการสอนต่อไปจนกว่าค่าความผิดพลาดระหว่างคำตอบของโครงข่ายกับเอาต์พุตที่ต้องการมีค่าน้อยพอที่จะยอมรับได้จึงจะหยุดการสอน และค่าถ่วงน้ำหนักที่ได้ก็จะเป็นเหมือนฟังก์ชันที่ได้ในการแปลงข้อมูล

#### 3.4.2 การสอนแบบไม่มีการชี้แนะหรือไม่มีการดูแล (Unsupervised Learning)

การสอนโดยวิธีนี้จะป้อนอินพุตเข้าสู่โครงข่าย และภายในโครงข่ายจะมีเอาต์พุตโหนดอยู่หลายโหนดด้วยกัน โดยแต่ละโหนดจะแทนกลุ่มของข้อมูลที่มีคุณสมบัติเหมือนกัน เมื่อป้อนอินพุตเข้าสู่โครงข่ายๆ จะคำนวณค่าความสัมพันธ์ที่มีอยู่ภายในเซตของอินพุต โดยอาศัยค่าถ่วงน้ำหนักเป็นตัวแยกความแตกต่างของอินพุตไปเก็บไว้ในโหนดเอาต์พุตของโครงข่าย การสอนโดยวิธีนี้จะไม่สามารถระบุได้ว่าเอาต์พุตโหนดใดเป็นของข้อมูลกลุ่มไหน ผู้ใช้ต้องกำหนดเอง ซึ่งจากการสอนแบบชี้แนะที่โครงข่ายสามารถระบุกลุ่มของเอาต์พุตได้อย่างแน่นอน

การสอนโครงข่ายเซลล์ประสาทแบบจำลอง เป็นการหาฟังก์ชันการแปลงและฟังก์ชันการแปลงที่ได้มีคุณสมบัติไม่เป็นเชิงเส้น ซึ่งฟังก์ชันการแปลงของโครงข่ายประสาทเทียม ในที่นี้คือเซตของค่าถ่วงน้ำหนักของโครงข่าย ดังนั้นฟังก์ชันการแปลงจะมีศักยภาพภาพมากน้อยเพียงใดนั้นจะขึ้นอยู่กับค่าถ่วงน้ำหนักของโครงข่ายนั้นๆ ว่ามีเสถียรภาพมากน้อยเพียงใด และค่าถ่วงน้ำหนักคำนวณได้จากการสอนโครงข่าย ซึ่งการสอนโครงข่ายมีหลายแบบด้วยกันเช่น กฎการสอนของ

Hebb, กฎการสอนขอแบบ Perceptron ของ Rosenblatt, กฎการสอนแบบเบลล์, กฎการสอนของ Widrow-Hoff, กฎการสอนโดยใช้สหสัมพันธ์, กฎการสอนแบบ Winner-Take-All, และกฎการสอนแบบ Outstar ของ Grossberg และในตารางที่ 3.1 ได้สรุปการสอนโครงข่ายของวิธีต่างๆ ไว้

ตารางที่ 3.1 แสดงการปรับค่าถ่วงน้ำหนักสำหรับกฎการสอนแบบต่างๆ

กฎการสอน	การปรับค่าถ่วงน้ำหนัก	ค่าเริ่มต้นของค่าถ่วงน้ำหนัก	การ สอน	Activation Function
Hebb	$\Delta w_{ij} = cf(W_i^1 X) x_j$	0	U	Any
Preceptron	$\Delta w_{ij} = c(d_i - f(W_i^1 X)) x_j$	จำนวนจริงใดๆ	S	Binary Bipolar or Binary Unipolar
Delta	$\Delta w_i = c(d_i - o_i) f'(W_i^1 X) X$	จำนวนจริงใดๆ	S	ฟังก์ชันต่อเนื่อง
Widrow-Hoff	$\Delta w_{ij} = c(d_i - W_i^1 X) x_j$	จำนวนจริงใดๆ	S	Any
Correlation	$\Delta w_{ij} = cd_i x_j$	0	S	Any
Winner-Take-All	$\Delta w_{mj} = c(x_j - w_{mj})$ $m$ : โหนดที่เป็น Winner	Random Normalized	U	ฟังก์ชันต่อเนื่อง
Outstar	$\Delta w_{ij} = c(d_i - w_{ij})$	0	S	ฟังก์ชันต่อเนื่อง

C: เป็นค่าคงที่ของอัตราการเรียนรู้ [10-13]

S: การสอนแบบชี้หน้า, U: การสอนแบบไม่ชี้หน้า

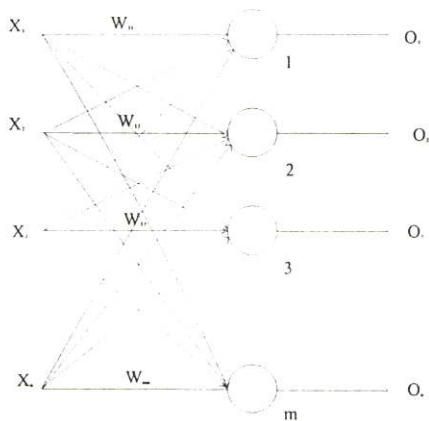
$\Delta w_{ij}$  : ค่าถ่วงน้ำหนักที่ถูกปรับค่าและมีการเชื่อมต่อระหว่างอินพุตโหนดที่  $i$  และเอาต์พุตโหนดที่  $j$

### 3.5 แบบจำลองโครงข่ายประสาทเทียม

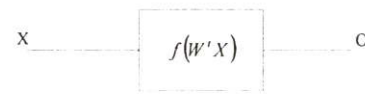
จากที่ได้กล่าวมาแล้วเกี่ยวกับโครงข่ายเซลล์ประสาท ทั้งที่เป็นแบบจำลองและเป็นโครงข่ายของเซลล์ประสาทมนุษย์จริง จะมีการเชื่อมต่อกันของโหนดในลักษณะของโครงข่ายอย่างหนาแน่น เพื่อให้โครงข่ายสามารถเรียนรู้และสามารถจดจำสิ่งที่เรียนรู้มาแล้วได้ ซึ่งการเชื่อมโยงของโครงข่ายจะมีสองลักษณะด้วยกันคือ

### 3.5.1 โครงข่ายที่ส่งสัญญาณไปข้างหน้า (Feed forward Networks)

โครงข่ายชนิดนี้ประกอบด้วยชั้นต่างๆ ของโครงข่าย โดยชั้นแรกจะเป็นอินพุตและชั้นสุดท้ายจะเป็นชั้นเอาต์พุต ส่วนระหว่างชั้นอินพุตกับชั้นเอาต์พุตอาจจะมี หรือไม่มีชั้นที่แทรกอยู่ภายในอีกก็ได้ ซึ่งจะขึ้นอยู่กับอัลกอริทึมที่ใช้ในการสอนโครงข่าย เช่นถ้าเป็นโครงข่าย Perceptron แบบหลายชั้น (Multilayer Perceptron) ก็จะมีชั้นที่อยู่ระหว่างอินพุตกับเอาต์พุตอีก ซึ่งอาจจะมีมากกว่าหนึ่งชั้นก็ได้ ส่วนโครงข่าย Self-Organizing Map ของ Kohonen [14] จะมีเพียงชั้นของอินพุตกับเอาต์พุตเท่านั้น การเชื่อมต่อระหว่างเครือข่ายแบบ Feed forward จะมีค่าถ่วงน้ำหนักเป็นตัวเชื่อม และสัญญาณอินพุตที่เข้ามาจะถูกส่งไปตามทิศทางของลูกศรจนถึงชั้นของเอาต์พุต โดยไม่มีการป้อนกลับดังรูปที่ 3.9



(ก) ลักษณะการเชื่อมโยง

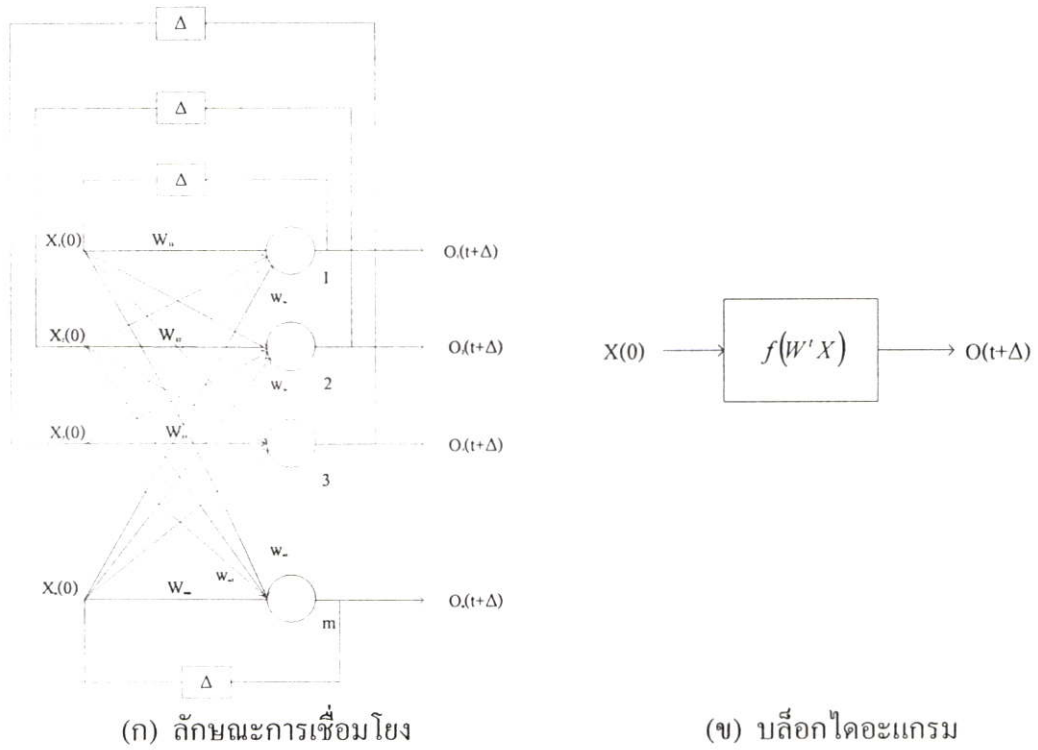


(ข) บล็อกไดอะแกรม

รูปที่ 3.9 โครงข่าย Feed forward แบบชั้นเดียว

### 3.5.2 โครงข่ายที่มีการป้อนกลับ (Feedback Networks)

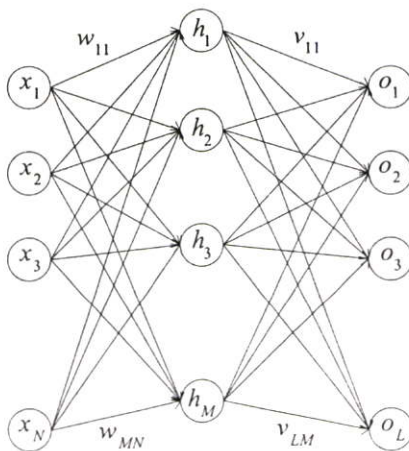
ในส่วนแรกของโครงข่ายนี้จะเป็นโครงข่าย Feed forward เหมือนกับแบบแรก และส่วนที่เพิ่มเข้ามาคือส่วนของการป้อนกลับดังแสดงในรูปที่ 3.10 และการป้อนกลับจะมีการหน่วงเวลาไปจากเดิมเท่ากับ  $\Delta$  ซึ่งโครงข่ายในรูปที่ 3.10 จะเรียกว่า โครงข่ายรีเคอร์เรนต์ (Recurrent Network) [8]



รูปที่ 3.10 โครงข่ายป้อนกลับแบบเวลาไม่ต่อเนื่องชั้นเดียว

### 3.6 การแพร่กระจายกลับ

การแพร่กระจายกลับหรือแบคโพรพาเกชัน (Back-Propagation) เป็นขั้นตอนที่ใช้สอนโครงข่ายแบบ Multilayer Perceptron ซึ่งเป็นแบบจำลองโครงข่ายเซลล์ประสาทที่มีการเชื่อมโยงกันเป็นโครงข่ายแบบชั้นๆ ดังในรูปที่ 3.11



รูปที่ 3.11 โครงข่าย Multilayer Perceptron ที่มี 3 ชั้น

$X = \{x_1, x_2, \dots, x_N\}$ ; เซตของอินพุต (Input Layer)

$H = \{h_1, h_2, \dots, h_M\}$ ; เซตของฮิดเดน (Hidden Layer)

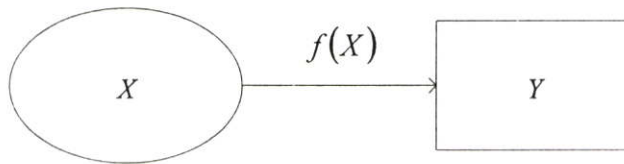
$O = \{o_1, o_2, \dots, o_L\}$ ; เซตของเอาต์พุต (Output Layer)

$W$  = เมตริกซ์ของค่าถ่วงน้ำหนักที่เชื่อมต่อระหว่างชั้นอินพุตกับฮิดเดน

$V$  = เมตริกซ์ของค่าถ่วงน้ำหนักที่เชื่อมต่อระหว่างชั้นฮิดเดนกับชั้นเอาต์พุต

ซึ่งโครงข่ายนี้จะมีการเชื่อมโยงกัน 3 ชั้น ซึ่งประกอบด้วยชั้นของอินพุตที่มีเซลล์ประสาทอยู่  $N$  โหนด ถัดมาเป็นชั้นของฮิดเดนหรือชั้นภายใน (Hidden Layer) ซึ่งประกอบด้วยโหนดต่างๆ จำนวน  $M$  โหนด และชั้นสุดท้ายคือชั้นของเอาต์พุตที่มีโหนดต่างๆ อยู่  $L$  โหนด โครงข่ายแบบ Multilayer Perceptron ในรูปที่ 3.11 แต่ละโหนดในชั้นเดียวกันจะไม่มี การเชื่อมต่อกัน การเชื่อมโยงกันจะมีเฉพาะระหว่างชั้นเท่านั้น และการเชื่อมต่อนี้จะต่อถึงกันทุกโหนด โครงข่ายแบบ Multilayer Perceptron ไม่จำเป็นต้องมีสามชั้นเหมือนในรูปที่ 3.11 อาจจะมีจำนวนชั้นมากกว่านี้ก็ได้ ซึ่งอาจมีสี่ชั้น โดยการเพิ่มชั้นฮิดเดนเข้าไปอีกหนึ่งชั้น หรือถ้าต้องการจำนวนชั้นมากกว่านี้ก็สามารถทำได้ โดยการเพิ่มชั้นของฮิดเดน

ชั้นฮิดเดน [8, 14] เป็นตัวเพิ่มความสามารถให้แก่โครงข่ายเซลล์ประสาท โครงข่ายแบบ Multilayer Perceptron ถ้าไม่มีชั้นฮิดเดนก็จะกลายเป็นโครงข่ายแบบ Perceptron อย่างเช่น โครงข่ายแบบ Perceptron ที่เสนอโดย Rosenblatt ซึ่ง Minsky และ Papert ได้แสดงให้เห็นว่าโครงข่ายเซลล์ประสาทจำลองนี้เป็นโครงข่ายที่มีความสัมพันธ์กันระหว่างอินพุตกับเอาต์พุตในลักษณะเชิงเส้นซึ่งจะไม่สามารถแก้ปัญหาเอ็กคลูซีฟ (Exclusive OR : XOR) ได้ [14] ในกรณีที่ไม่มีชั้นฮิดเดน ความสัมพันธ์ของฟังก์ชันการแปลงดังแสดงในรูปที่ 3.12

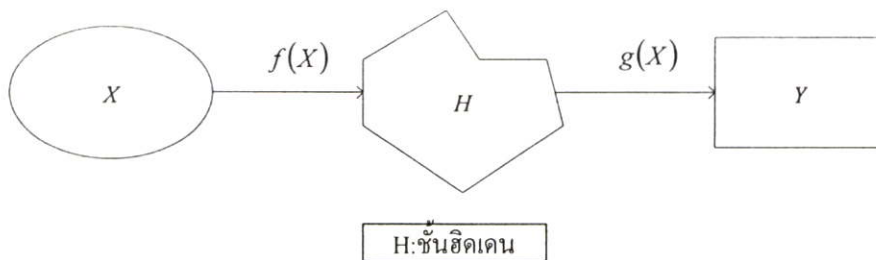


รูปที่ 3.12 แสดงความสัมพันธ์ระหว่างตัวแปร  $X$  กับ  $Y$

จากรูปที่ 3.12 เมื่อ  $X$  เป็นตัวแปรอิสระ  $Y$  เป็นตัวแปรตามและ  $f(X)$  เป็นฟังก์ชันการแปลงจาก  $X$  ไปยัง  $Y$  ในกรณีที่  $X$  เป็นอิสระต่อกันแบบเชิงเส้น (Linearly Independent) เราสามารถคำนวณฟังก์ชันการแปลง  $f(X)$  ได้ ดังตัวอย่างเช่น ถ้า  $\alpha_1 = \{1,0\}$  และ  $\alpha_2 = \{0,1\}$  เมื่อ  $X = \{\alpha_1, \alpha_2\}$  และ  $Y' = [1, 1]$  ดังนั้นเราสามารถคำนวณฟังก์ชันการแปลงได้ดังนี้

$$X f(X) = Y \rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

เมื่อฟังก์ชันการแปลง  $f(X) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  แต่ถ้าตัวแปร  $X$  ไม่เป็นอิสระต่อกันแบบเชิงเส้น (Linearly Independent) เราไม่สามารถจะกำหนดฟังก์ชันการแปลง  $f(X)$  ได้ เช่น ถ้า  $\alpha_1 = \{0,0\}$  และ  $\alpha_2 = \{1,1\}$  เมื่อ  $X = \{\alpha_1, \alpha_2\}$  และ  $Y' = [1, 1]$  จะได้ว่า  $X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$  ซึ่งในกรณีนี้เราไม่สามารถคำนวณฟังก์ชัน  $f(X)$  ที่สอดคล้องกับ  $Y$  ได้ และในกรณีนี้จะเหมือนกับปัญหา XOR แต่ถ้าเพิ่มชั้นฮิดเดนเข้าไปในรูปที่ 3.12 จะทำให้สามารถแก้ปัญหาคือเป็น Linearly Independent นี้ได้ [8,14] (ดูรูปที่ 3.13)



รูปที่ 3.13 แสดงความสัมพันธ์ระหว่างตัวแปร  $X$  กับ  $Y$  โดยมีชั้นฮิดเดนเป็นตัวเชื่อม

ในการหาความสัมพันธ์จาก  $X$  ไปยัง  $Y$  โดยชั้นของฮิดเดนในทาง โครงข่ายประสาทเทียม จะทำให้ฟังก์ชัน  $f(X)$  และ  $g(X)$  ไม่เป็นเชิงเส้นซึ่ง  $f(X)$  และ  $g(X)$  ในที่นี้คือค่าถ่วงน้ำหนักที่ได้จากการสอนโครงข่ายนั่นเอง การที่จะกำหนดชั้นของฮิดเดนว่าในโครงข่ายหนึ่งๆ ควรมีฮิดเดนกี่ชั้น และแต่ละชั้นประกอบด้วยโหนดกี่โหนดนั้น ไม่มีกฎเกณฑ์ที่แน่นอน ดังนั้นการกำหนดจำนวนชั้นและจำนวนโหนดของแต่ละชั้นของฮิดเดนสามารถทำได้โดยการทดลอง

### 3.7 กฎเดลต้าอเนกประสงค์

กฎเดลต้า (Delta Rule) ถูกพัฒนาขึ้นมาเพื่อใช้สอนโครงข่ายเซลล์ประสาทจำลอง [12] ซึ่งครั้งแรกใช้ในการสอนโครงข่ายเซลล์ประสาทจำลอง Perceptron ซึ่งจะเป็นกฎการสอน Perceptron แบบต่อเนื่อง และต่อมาได้พัฒนาการสอนเดลต้ารูต์ให้ใช้ได้กับโครงข่ายที่มีการเชื่อมต่อกันหลายชั้นอย่าง Multilayer Perceptron จึงเรียกกฎเดลต้าที่ถูกพัฒนาขึ้นใหม่นี้ว่า กฎเดลต้าอเนกประสงค์ (Generalized Delta Rule: GDR)

ขั้นตอนการสอนโครงข่าย Multilayer Perceptron จะเรียกว่า ขั้นตอนการสอนแบบค่าความผิดพลาดแพร่กระจายกลับหรือ Error Back-propagation ซึ่งเป็นการแพร่กระจายกลับของค่าความผิดพลาดที่เกิดขึ้นในชั้นเอาต์พุตที่ต้องการกับเอาต์พุตที่คำนวณได้ โดยคำนวณย้อนกลับจากชั้นเอาต์พุตผ่านชั้นฮิดเดนตลอดมาจนถึงชั้นอินพุต เพื่อทำการปรับค่าถ่วงน้ำหนัก จากรูปที่ 3.11 เมื่อทำการสอนโครงข่ายนี้ ซึ่งมีการสอนแบบชี้นำ ดังนั้นในการสอนจะต้องป้อนเซตของข้อมูลจะใช้สอนซึ่งประกอบด้วยเซตของอินพุต และเซตของเอาต์พุตที่ต้องการซึ่งจะสอดคล้องกับเซตของอินพุต กำหนดให้  $P$  เป็นจำนวนเซตทั้งหมดที่ใช้ในการสอนโครงข่าย ดังนั้นเซตของอินพุตและเอาต์พุตที่ต้องการจะมีอยู่  $P$  เซต, ถ้าให้  $D$  เป็นเซตของเอาต์พุตที่ต้องการจะได้ว่า  $D_p = \{d_1, d_2, d_3, \dots, d_L\}$ , เมื่อ  $P=1, 2, 3, \dots, P$

$W$  เป็นเมตริกซ์ของค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างชั้นของอินพุตกับชั้นของฮิดเดนซึ่ง  $W$  มีขนาด  $M$  แถว  $N$  คอลัมน์

$$W = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1N} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2N} \\ w_{31} & w_{32} & w_{33} & \dots & w_{3N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{M1} & w_{M2} & w_{M3} & \dots & w_{MN} \end{bmatrix}$$

$V$  เป็นเมตริกซ์ของค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างชั้นฮิดเดนไปยังชั้นเอาต์พุต ซึ่งมีขนาด  $L$  แถว  $M$  คอลัมน์

$$V = \begin{bmatrix} v_{11} & v_{12} & v_{13} & \dots & v_{1M} \\ v_{21} & v_{22} & v_{23} & \dots & v_{2M} \\ v_{31} & v_{32} & v_{33} & \dots & v_{3M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{L1} & v_{L2} & v_{L3} & \dots & v_{LM} \end{bmatrix}$$

และ  $f(net)$  เป็น Activation Function ใดๆ ดังที่แสดงในรูปที่ 3.8

เมื่อทำการป้อนเซตของข้อมูลที่ใส่สอนเข้าสู่โครงข่ายจะสามารถคำนวณ โหนดของฮิดเดนที่  $m$  ได้ดังนี้

$$\begin{aligned} net_m &= \sum_{n=1}^N w_{mn} x_n \\ h_m &= f(net_m) \end{aligned} \quad (3.3)$$

และสามารถคำนวณเอาต์พุต โหนดที่  $l$  ได้ดังนี้

$$\begin{aligned} net_l &= \sum_{m=1}^M v_{lm} x_m \\ o_l &= f(net_l) \end{aligned} \quad (3.4)$$

เมื่อคำนวณชั้นเอาต์พุตครบทุกโหนดแล้ว ขั้นต่อไปเป็นการปรับค่าถ่วงน้ำหนักของโครงข่าย โดยใช้ค่าผลรวมของค่าความคลาดเคลื่อนกำลังสอง ระหว่างค่าของผลลัพธ์ที่ได้จากโครงข่ายกับค่าเอาต์พุตที่ต้องการของแพทเทิร์นที่  $p$  ซึ่งสามารถคำนวณได้ดังนี้

$$E_p = \frac{1}{2} \sum_{l=1}^L (o_{pl} - d_{pl})^2 \quad (3.5)$$

ดังนั้นในการปรับค่าถ่วงน้ำหนักที่เชื่อมต่อระหว่างชั้นเอาต์พุตกับชั้นฮิดเดนสำหรับแพทเทิร์นที่  $p$  ( $H_p$ ) สามารถคำนวณได้ดังนี้

$$\begin{aligned} \Delta V &= -\eta \frac{\partial E_p}{\partial V} \\ &= \alpha V + \eta \delta_p^V H_p \end{aligned} \quad (3.6)$$

โดยที่  $\delta_p^V = (D_p - O_p) O_p (1 - O_p)$  ซึ่งเป็นความคลาดเคลื่อนภายในชั้นของเอาต์พุต

$\alpha$  เป็นค่าโมเมนตัม [8, 12]

$\eta$  เป็นค่าอัตราการเรียนรู้ (Learning Rate) ซึ่งเป็นค่าคงที่ [8, 12]

และการปรับค่าถ่วงน้ำหนักที่เชื่อมต่อระหว่างชั้นฮิดเดนกับชั้นอินพุตสามารถคำนวณได้ดังนี้

$$\Delta W' = \alpha W + \eta \delta_p^W X_p \quad (3.7)$$

โดยที่  $\delta_p^W = H_p (1 - H_p) (\delta_p^V V)$  ซึ่งเป็นค่าความคลาดเคลื่อนภายในชั้นฮิดเดน สำหรับ  $\alpha$  และ  $\eta$  จะมีคุณสมบัติเหมือนกับสมการที่ (3.6)

ในกรณีที่โครงข่ายมี  $I$  ชั้นสามารถปรับค่าถ่วงน้ำหนักโดยวิธีค่าความผิดพลาดกระจายกลับได้ดังนี้

$$\Delta W^i = \alpha W^i + \eta \delta_p^i X_p^{i-1} \quad (3.8)$$

เมื่อ  $W^i$  เป็นเมตริกซ์ค่าถ่วงน้ำหนักที่อยู่ระหว่างชั้นที่  $i$  กับ  $i-1$  ( $i=1, 2, 3, \dots, I$ ),  $X_p^{i-1}$  เป็นเซตของอินพุตแพทเทิร์นที่  $p$  จากชั้นที่  $i-1$  และ  $\delta_p^i$  เป็นเวกเตอร์ของความคลาดเคลื่อนสำหรับชั้นที่  $i$  สำหรับชั้นเอาต์พุตหรือชั้นที่  $I$  สามารถคำนวณได้ดังนี้

$$\delta_p^i = (D_p - O_p^i) O_p^i (1 - O_p^i) \quad (3.9)$$

โดยที่  $O_p^i$  คือชั้นของเอาต์พุต และสำหรับความคลาดเคลื่อนของชั้นที่  $i$  ใดๆ โดยที่  $i \neq I$  สามารถคำนวณได้ดังนี้

$$\delta_p^i = O_p^i (1 - O_p^i) (\delta_p^{i+1} W^{i+1}) \quad (3.10)$$

โดยที่  $O_p^i$  คือค่าเอาต์พุตชั้นที่  $i$  แพทเทิร์นที่  $p$  และสำหรับ  $\delta_p^{i+1}$  และ  $W^{i+1}$  เป็นค่าความคลาดเคลื่อนและเมตริกซ์ค่าถ่วงน้ำหนักของชั้นถัดไป

ปัจจัยที่ทำให้ประสิทธิภาพของการสอนโครงข่ายด้วยวิธี Back-propagation เพิ่มขึ้นมีดังนี้

- การกำหนดค่าเริ่มต้นให้กับเมตริกซ์ถ่วงน้ำหนัก ก่อนที่จะทำการสอนโครงข่าย Multilayer Perceptron โดยวิธี Error Back-propagation จำเป็นต้องกำหนดค่าเริ่มต้นให้กับเมตริกซ์ค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างชั้นทุกชั้น โดยค่านี้จะเป็นเลขจำนวนจริงที่มีค่าน้อยๆ ที่ได้มาจากการสุ่มค่าเริ่มต้นของเมตริกซ์เชื่อมโยงจะมีผลต่อเวลาที่ใช้ในการสอน และอาจจะส่งผลถึงค่าถ่วงน้ำหนักหลังจากที่ได้สอนโครงข่ายไปแล้วว่ามีศักยภาพมากน้อยเพียงใดในการแปลงข้อมูลอินพุตไปสู่เอาต์พุต

- อัตราเร็วในการเรียนรู้ (Learn Rate:  $\eta$ ) ซึ่งโดยทั่วไปแล้วค่านี้จะเป็นค่าคงที่ที่มีค่าอยู่ระหว่าง 0.05-0.25 [14] แต่มีงานวิจัยบางส่วนที่ค่าอัตราเร็วในการเรียนรู้มีการปรับค่าได้เช่น กำหนดให้อัตราเร็วในการเรียนรู้ผันแปรตามจำนวนรอบที่สอน [16], อัตราเร็วในการเรียนรู้ผันแปรตามค่าความผิดพลาดเฉลี่ยของระบบ [17] และอัตราเร็วในการเรียนรู้ผันแปรตามค่าความแปรปรวนของแต่ละแพทเทิร์น [15]

- โมเมนตัม (Momentum:  $\alpha$ ) สำหรับการสอนแบบเจอเนรอลไลซ์เซลล์ต่ำ โดยทั่วไปค่าโมเมนตัมอาจจะใช้หรือไม่ใช้ก็ได้ แต่ค่านี้จะมีคุณสมบัติป้องกันการแกว่ง (Oscillate) ของระบบ

โดยค่านี้อาจสัมพันธ์กับอัตราความเร็วในการเรียนรู้ คือถ้าอัตราเร็วในการเรียนรู้มีค่ามากแต่ค่าโมเมนต์ดัมมีค่าน้อยจะทำให้ระบบโครงข่ายเกิดการแกว่ง

-ไบแอส (Bias) โหนดไบแอสสำหรับโครงข่ายโดยทั่วไปอาจจะมีหรือไม่มีก็ได้ ในกรณีของโครงข่ายที่ใช้ในระบบที่เป็น Real-time แล้วจะไม่ค่อยมีการใช้โหนดไบแอส เนื่องจากจะเสียเวลาในการคำนวณ และถ้ามีการออกแบบโครงข่ายอย่างเหมาะสมแล้วเทอมไบแอสไม่จำเป็นต้องมี แต่ถ้าจำนวนโครงข่ายไม่เหมาะสมคือมีจำนวนโหนดน้อยเกินไป การเพิ่มโหนดไบแอสจะทำให้โครงข่ายสามารถจดจำได้ดีขึ้น โหนดไบแอสจะถูกเพิ่มเข้าไปเป็นโหนดๆ หนึ่งในแต่ละชั้น ยกเว้นในชั้นของผลลัพธ์ และโหนดไบแอสที่เพิ่มเข้าไปมักเป็นค่าคงที่ และผลของการเพิ่มโหนดไบแอสเข้าไปในโครงข่ายจะทำให้ประสิทธิภาพของการทำงานดีขึ้น [18]

### 3.8 สรุป

โครงข่ายประสาทเทียมเป็นเครื่องมือช่วยให้คอมพิวเตอร์มีความสามารถเพิ่มมากขึ้น (นอกเหนือจากการคำนวณด้วยหลักทางคณิตศาสตร์ทั่วไปและการเก็บข้อมูล) โดยเฉพาะการประมวลผลข่าวสารที่มีความยุ่งยากซับซ้อน เช่นการทำให้คอมพิวเตอร์สามารถรู้จักภาพและเข้าใจภาษาพูดของมนุษย์ ซึ่งในงานเช่นนี้จะต้องมีการตัดสินใจที่ต้องอาศัยความรู้และประสบการณ์ ซึ่งถ้าใช้เทคนิคทางคณิตศาสตร์ธรรมดาแก้ปัญหานี้ก็ทำให้มีระบบที่มีความซับซ้อนมาก แต่ถ้าใช้โครงข่ายประสาทเทียม ก็จะช่วยลดความยุ่งยากลงได้มากทีเดียว แต่อย่าลืมว่าในปัจจุบันโครงข่ายประสาทเทียมเหมือนฟังก์ชันการแปลงเท่านั้น ดังนั้นเมื่อมีข้อมูลที่เข้ามาใหม่และมีการแจกแจงที่ต่างจากเดิมมาก (ต่างจากข้อมูลที่เคยสอนไว้มาก) ระบบนี้ก็จะมีความคลาดเคลื่อนสูงเช่นกัน

แต่อย่างไรก็ตาม โครงข่ายประสาทเทียมยังมีประโยชน์อยู่มาก เพียงแต่เราต้องเลือกปัญหาให้เหมาะสมกับระบบก็จะช่วยลดความซับซ้อนของระบบลงได้มาก อย่างเช่นทำให้คอมพิวเตอร์จดจำรูปแบบของตัวอักษรแต่ละตัว ที่เป็นทั้งตัวพิมพ์และลายมือเขียน

## บทที่ 4

# ผลการทดลองการตรวจจับเปลวเพลิงภายในภาพ

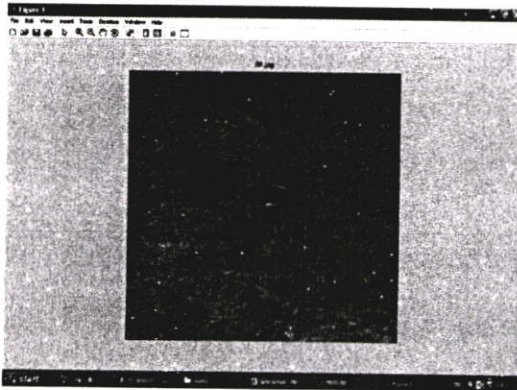
### 4.1 อุปกรณ์ที่ใช้ในการทดลอง

- คอมพิวเตอร์ Acer Intel Core Duo 1.66 MHz ที่มีหน่วยความจำแรม 512 MB
- กล้องที่สามารถรับภาพได้ขนาด 2288x1712 จุด (แต่จะทำการย่อรูปให้เหลือขนาด 256x256 จุด เพื่อที่จะทำการจำแนกได้รวดเร็วยิ่งขึ้น)
- ในวิทยานิพนธ์ฉบับนี้ จะทดลองการทำงานโดยใช้โปรแกรม MATLAB 7.0

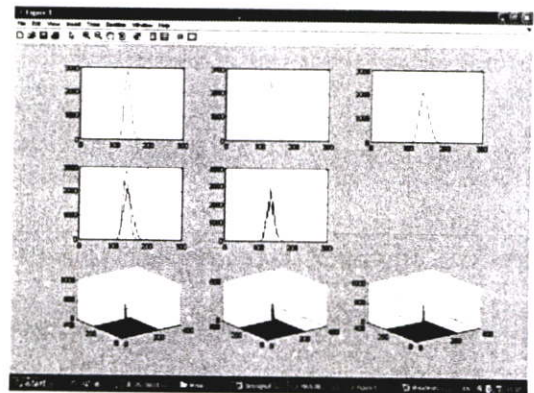
### 4.2 ผลการเตรียมข้อมูลภาพ

#### 4.2.1 การออกแบบการหาค่าฮิสโตแกรมของภาพ

การหาค่าฮิสโตแกรมของภาพโดยใช้การจำลองโปรแกรม MATLAB ในขณะที่ไม่มีการเกิดเปลวเพลิงภายในภาพ และมีเปลวเพลิงเกิดขึ้นภายในภาพ จากรูปที่ 4.1 และรูปที่ 4.2 จะเห็นว่าค่าฮิสโตแกรมต่างๆ ที่ได้จะมีความแตกต่างกัน

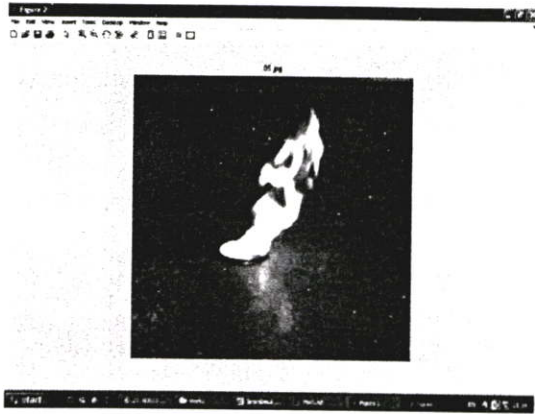


(ก) ภาพดิจิทัล

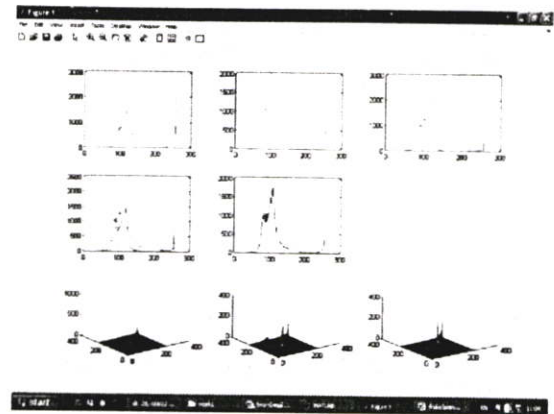


(ข) ฮิสโตแกรมที่สอดคล้อง

รูปที่ 4.1 แสดงไม่มีเปลวเพลิงเกิดขึ้นภายในภาพและฮิสโตแกรมที่สอดคล้อง



(ก) ภาพดิจิตอล



(ข) ฮิสโตแกรมที่สอดคล้อง

รูปที่ 4.2 แสดงเปลวเพลิงที่เกิดขึ้นภายในภาพโดยเปลวเพลิง และฮิสโตแกรมที่สอดคล้อง

#### 4.2.2 การปรับปรุงภาพ

จากการปรับปรุงภาพที่ได้กล่าวมาแล้วในบทที่ 2 ในวิทยานิพนธ์ฉบับนี้จะใช้การปรับปรุงฮิสโตแกรม ซึ่งการเปลี่ยนแปลงฮิสโตแกรมของภาพมีผลทำให้จุดภาพเกิดการเปลี่ยนแปลงค่าระดับความสว่าง การเปลี่ยนแปลงเกิดขึ้นจากจุดภาพหนึ่งจุดภาพใดโดยไม่เกี่ยวข้องกับค่าระดับความสว่างของจุดภาพข้างเคียง สำหรับภาพผลลัพธ์หลังการเปลี่ยนแปลงฮิสโตแกรมอาจจะช่วยให้มองเห็นรายละเอียดในภาพชัดเจนขึ้นหรืออาจจะแย่งก็ได้ แต่ในการเพิ่มความสว่างให้ภาพนั้นไม่สมควรที่จะเพิ่มความสว่างให้กับภาพมากเกินไป เพราะจะทำให้รายละเอียดหรือองค์ประกอบบางอย่างของภาพบางอย่างหายไป ผลที่ได้จากการปรับปรุงภาพจะแสดงดังรูปที่ 4.3



(ก) ภาพต้นฉบับ

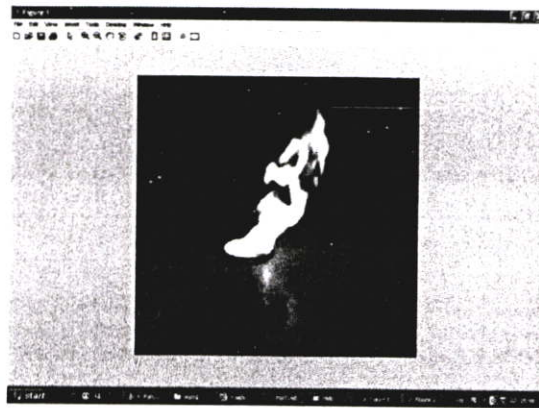


(ข) ภาพหลังการปรับปรุง

รูปที่ 4.3 ภาพเมื่อทำการปรับปรุงภาพ

### 4.2.3 การจำแนกภาพ

เนื่องจากกระบวนการการจำแนกภาพได้ใช้วิธีในการปรับปรุงภาพคือ Thresholding เพื่อทำการตัดภาพพื้นหลังของภาพออกไป ให้เหลือเพียงภาพของเปลวเพลิงอย่างเดียว ผลการทดสอบในส่วนของการปรับปรุงภาพด้วยวิธีการ Thresholding ทำให้ข้อมูลภาพเหลือเพียงข้อมูล 0 คือ พิกเซลที่เป็นสีดำ และ 1 คือข้อมูลภาพที่เป็นสีขาวเท่านั้น ซึ่งข้อมูลภาพที่ได้จะนำไปทำการหาขอบภาพให้ได้ภาพของเปลวเพลิงที่ต้องการเพื่อนำไปให้โครงข่ายประสาทเทียมเรียนรู้และทำการจำแนกต่อไป



(ก) ภาพต้นฉบับ



(ข) หลังการทำ Thresholding เมื่อค่า  $Th=150$



(ค) หลังการทำ Thresholding เมื่อค่า  $Th=180$

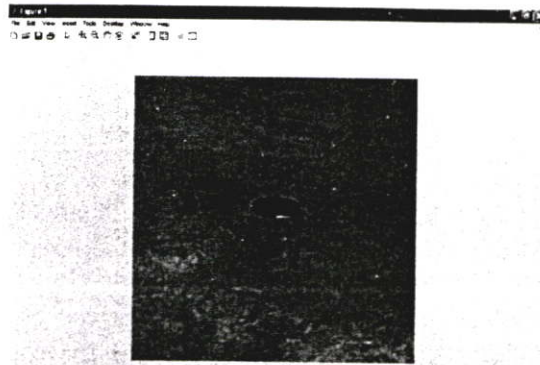


(ง) หลังการทำ Thresholding เมื่อค่า  $Th=220$

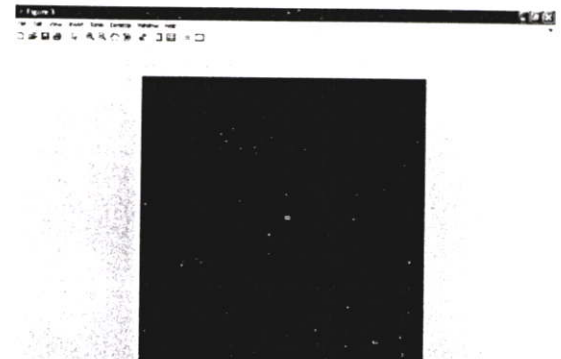
รูปที่ 4.4 แสดงผลการทำ Thresholding

#### 4.2.4 การหาขอบภาพ

การหาขอบภาพเป็นขั้นตอนหนึ่งในการเตรียมข้อมูลเพื่อฝึกการรู้จำของโครงข่ายประสาทเทียม



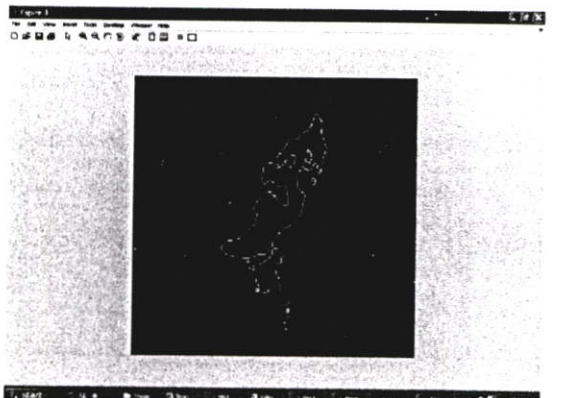
(ก) ภาพต้นฉบับเมื่อไม่มีแปลวเพลิง



(ข) ขอบของภาพ เมื่อค่า Th=180



(ค) ภาพต้นฉบับเมื่อมีแปลวเพลิง



(ง) ขอบของภาพ เมื่อค่า Th=180

รูปที่ 4.5 แสดงผลการหาขอบภาพ เมื่อให้ค่า Threshold=180

#### 4.2.5 การแยกคุณลักษณะเด่น

การแยกคุณลักษณะเด่นเป็นกระบวนการ การลดขนาดของข้อมูลให้เหมาะสมกับการฝึก การรู้จำโดยใช้โครงข่ายประสาทเทียม โดยการดึงเอาลักษณะเด่นของสัญญาณออกมาแต่ยังคง คุณลักษณะหลักๆ ของข้อมูลเอาไว้

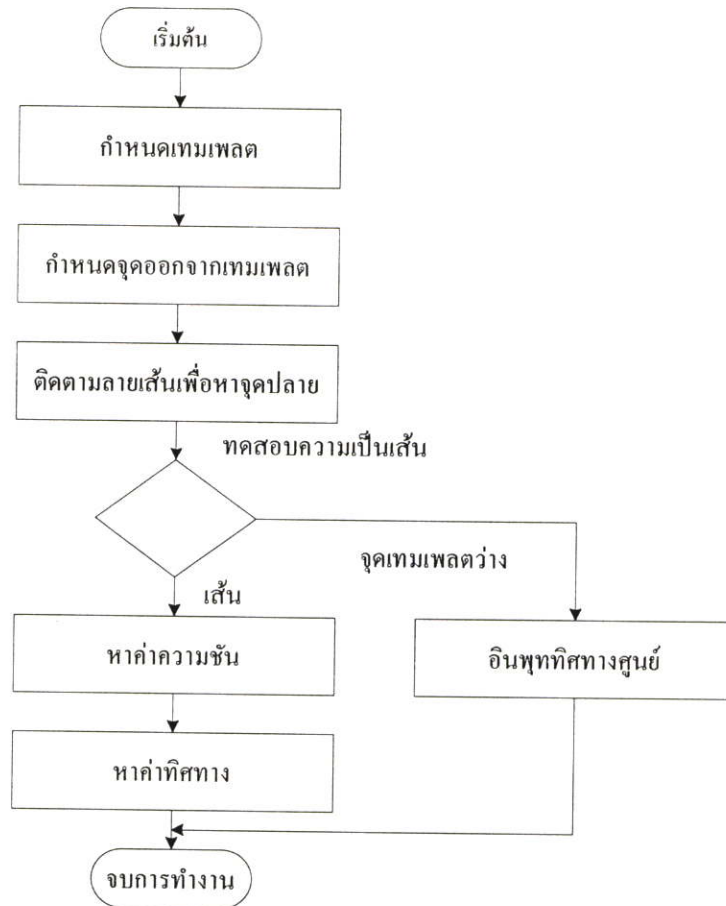
##### 4.2.5.1 การหาทิศทางหลัก

การหาทิศทางหลักของภาพแปลวเพลิง เป็นการพิจารณาค่าความชันของลายเส้นของแปลว เพลิง

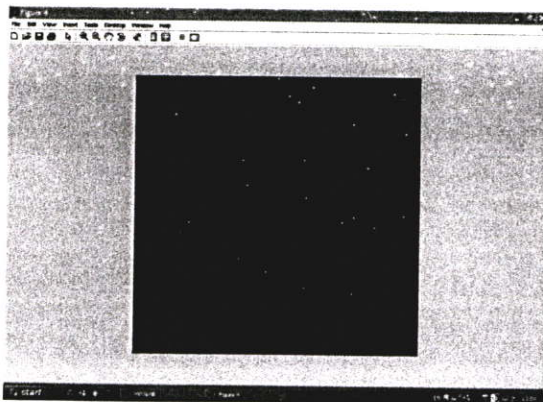
ขั้นตอนการหาทิศทาง

1. กำหนดเทมเพลต
2. กำจัดสัญญาณที่เป็นจุดออกไป
3. ติดตามลายเส้นเพื่อหาจุดเริ่มต้นและจุดสิ้นสุด โดยเริ่มจากจุดที่อยู่ตรงมุมบนซ้ายเป็น จุดเริ่มต้นแล้วเลื่อนจุดจากซ้ายไปขวา บนลงล่าง จนครบทุกจุดบนเทมเพลต

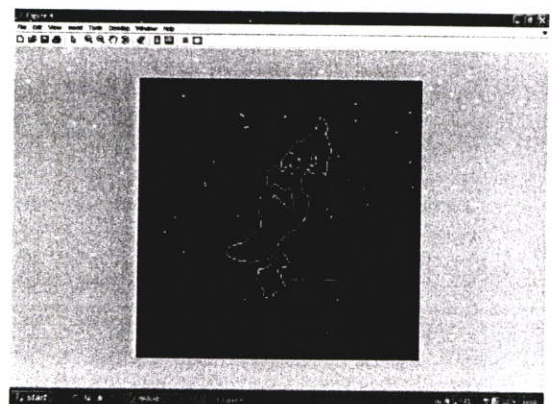
4. พิจารณาว่าได้ครบหรือไม่
5. คำนวณหาค่าความชัน
6. พิจารณาหาค่าความชันว่ามีทิศทางในลักษณะใด



รูปที่ 4.6 แสดงขั้นตอนการหาค่าลักษณะเด่นโดยวิธีการหาทิศทางของลายเส้น



(ก) ขอบภาพเมื่อไม่มีเปลวเพลิง



(ข) ขอบภาพเมื่อไม่มีเปลวเพลิง

1	1	1	1	1	1	1	1	1	1	1	1	1	4	4	4
1	1	1	1	1	1	1	1	1	1	1	1	1	4	4	4
1	1	1	1	1	1	1	1	1	1	4	1	1	4	1	4
1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	3
1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	3
1	1	1	1	1	1	1	1	2	2	1	2	1	4	2	2
1	1	1	1	1	1	1	1	2	2	2	3	4	4	3	2
1	1	1	1	1	1	1	1	3	3	3	4	4	4	3	1

(ค) ค่าลักษณะเด่นของภาพ (ก)

(ง) ค่าลักษณะเด่นของภาพ (ข)

รูปที่ 4.7 แสดงการหาค่าลักษณะเด่นของภาพ

### 4.3 ผลการฝึกการรู้จำโดยใช้โครงข่ายประสาทเทียม

การฝึกการรู้จำโดยใช้โครงข่ายประสาทเทียมแบบแพร่กลับ มีขั้นตอนการดำเนินการดังต่อไปนี้

#### 4.3.1 การรับค่าอินพุท

อินพุทสำหรับฝึกการรู้จำคือผลลัพธ์ที่ได้จากการหาค่าลักษณะเด่นในรูปที่ 4.7 และไปผ่านการแปลงข้อมูลให้มีลักษณะเป็นเวกเตอร์เพื่อที่จะส่งไปยังชั้นอินพุทของโครงข่ายดังรูปที่ 4.8

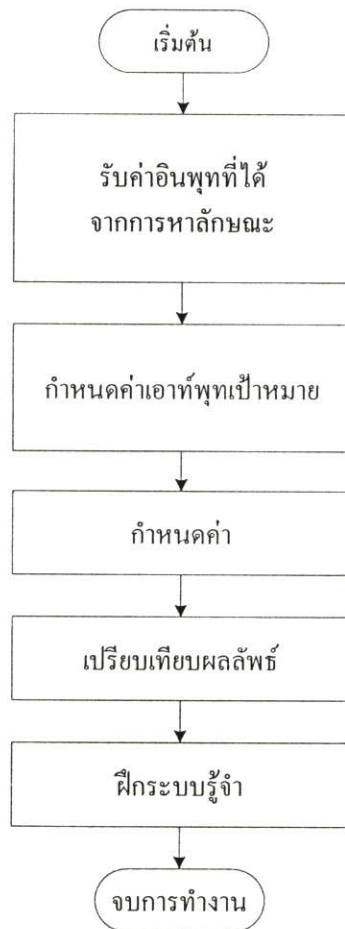
$$Xi = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1N} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2N} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{M1} & x_{M2} & x_{M3} & \cdots & x_{MN} \end{bmatrix} \Rightarrow \vec{V}_i = \{x_{11}, x_{12}, x_{13}, \dots, x_{1N}, x_{21}, x_{22}, \dots, x_{MN}\}$$

รูปที่ 4.8 อินพุทของโครงข่าย

### 4.3.2 การฝึกการรู้จำ

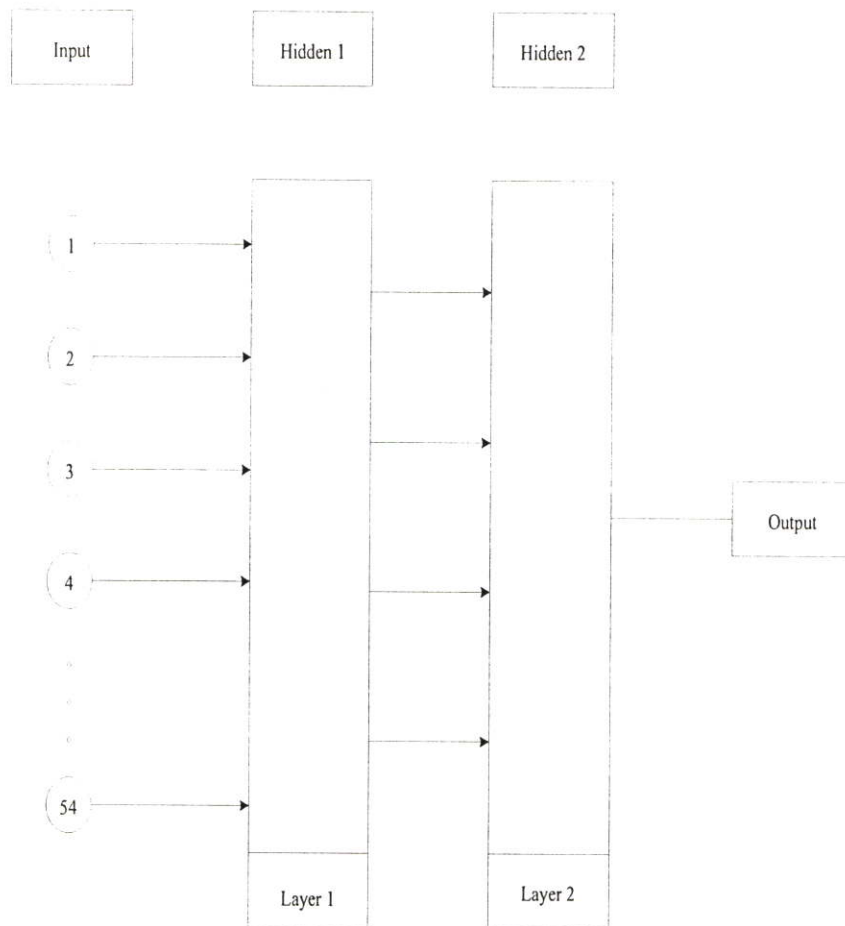
#### 4.3.2.1 ขั้นตอนการฝึกการรู้จำ

ประกอบด้วยขั้นตอนการกำหนดค่าโครงข่าย (Net) จำนวนเซลล์ประสาทในชั้นต่างๆ การกำหนดฟังก์ชันสำหรับการหาค่ากระตุ้น (Activate) การกำหนดฟังก์ชันสำหรับการฝึก (Training) การกำหนดรอบของการฝึกการรู้จำ (Epoch) การกำหนดค่าผิดพลาดที่ยอมรับได้ และการฝึกโครงข่ายประสาทเทียม



รูปที่ 4.9 แสดงขั้นตอนการฝึกการรู้จำ

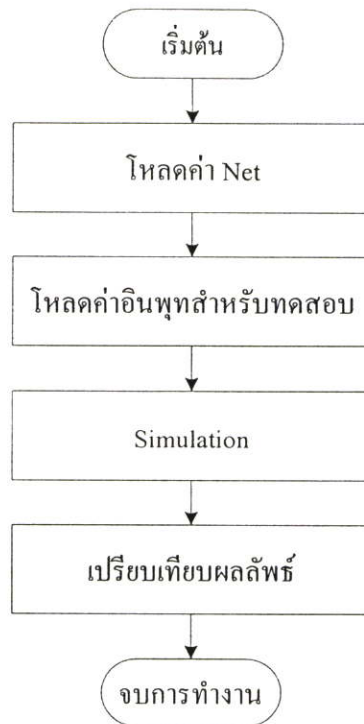
#### 4.3.2.2 การกำหนดจำนวนเซลล์ในโครงข่ายประสาทเทียม



รูปที่ 4.10 แสดงจำนวนเซลล์ประสาทในโครงข่ายประสาทเทียม

#### 4.3.3 ผลการทดสอบการรู้จำ

การทดสอบการรู้จำเป็นการวัดประสิทธิภาพของการฝึกโครงข่ายประสาทเทียม โดยนำค่าที่ได้จากการฝึก คือ Net และอินพุตสำหรับการทดสอบมา Simulate แล้วนำค่าที่ได้มาเปรียบเทียบกับเป็นภาพของเปลวเพลิงหรือไม่



รูปที่ 4.11 แสดงขั้นตอนการทดสอบการรู้จำ

ในการป้อนข้อมูลที่สำคัญบางอย่างเพื่อให้โครงข่ายประสาทเทียมนำไปใช้ในการประมวลผลนั้นจะต้องประกอบด้วย จำนวนชั้นซ่อน จำนวนเซลล์ในแต่ละชั้นซ่อน ค่าความผิดพลาดที่ต้องการ และอัตราการเรียนรู้

ในส่วนภาพเปลวเพลิงที่เป็นอินพุตให้กับโครงข่ายประสาทเทียมเรียนรู้ นั้นจะเป็นภาพขนาด 256x256 นั่นคือมีจำนวน 65,536 พิกเซลต่อภาพ

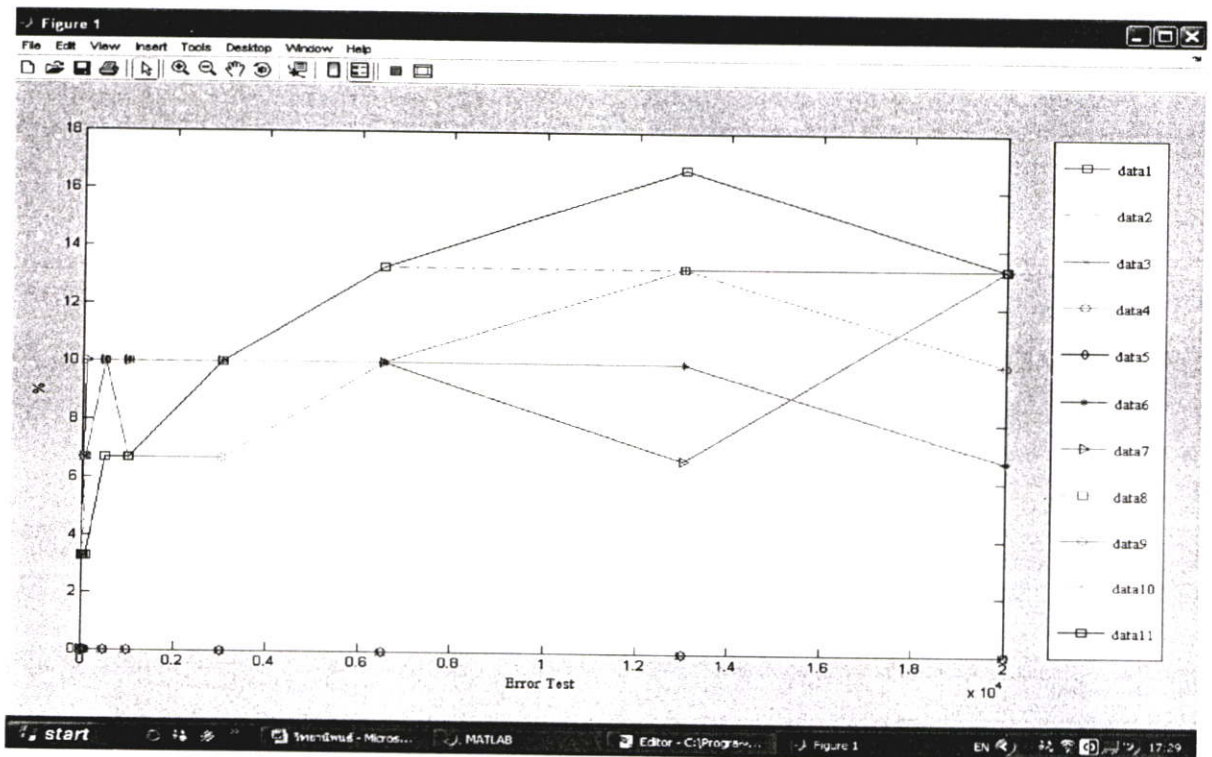
ในส่วนของการทดสอบการจำแนก จะเป็นการป้อนภาพที่ไม่ได้ให้โครงข่ายประสาทเทียมทำการเรียนรู้ จำนวน 30 ภาพ เพื่อทำการสังเกตผลการทำงานของโครงข่าย

- แบบ 1 ชั้นซ่อน จะใช้จำนวนจุดผิดพลาดในการทดลองให้โครงข่ายประสาทเทียมทำการจำแนกที่ 1, 10, 50, 100, 500, 1,000, 3,000, 6,500, 13,000 และ 20,000 จุด จำนวนเซลล์ในชั้นซ่อนจะใช้ที่ 1, 5, 10, 50, 100, 200, 500, 1,000, 2,000, 5,000 และ 10,000

- แบบ 2 ชั้นซ่อน จะใช้จำนวนจุดผิดพลาดในการทดลองให้โครงข่ายประสาทเทียมทำการจำแนกที่ 1, 10, 50, 100, 500, 1,000, 3,000, 6,500, 13,000 และ 20,000 จุด จำนวนเซลล์ในชั้นซ่อนที่ 1 จะใช้ที่ 50, 100, 200, 500, 1,000, 2,000 และ 5,000 ส่วนจำนวนเซลล์ในชั้นซ่อนที่ 2 จะใช้ที่ 50, 100, 200, 500, 1,000, 2,000 และ 5,000 เช่นกัน

ตารางที่ 4.1 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 1 พิกเซล

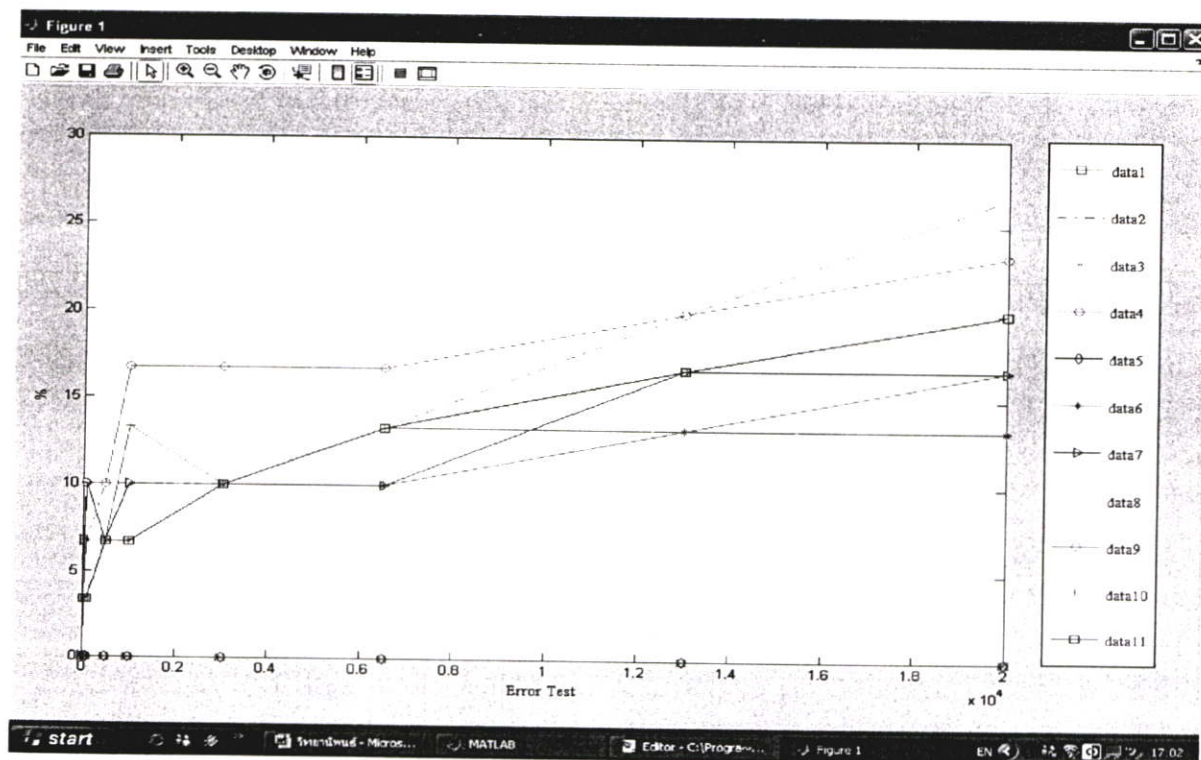
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0
200	0	3.3	6.7	6.7	10	10	10	10	10	6.7
500	0	3.3	6.7	10	10	10	10	10	6.7	13.3
1,000	0	3.3	3.3	6.7	10	10	10	13.3	13.3	13.3
2,000	0	3.3	6.7	6.7	10	6.7	6.7	10	13.3	10
5,000	0	3.3	6.7	6.7	10	6.7	10	10	13.3	13.3
10,000	0	3.3	3.3	3.3	6.7	6.7	10	13.3	16.7	13.3



รูปที่ 4.12 แสดงข้อมูลจากตารางที่ 4.1

ตารางที่ 4.2 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 10 พิกเซล

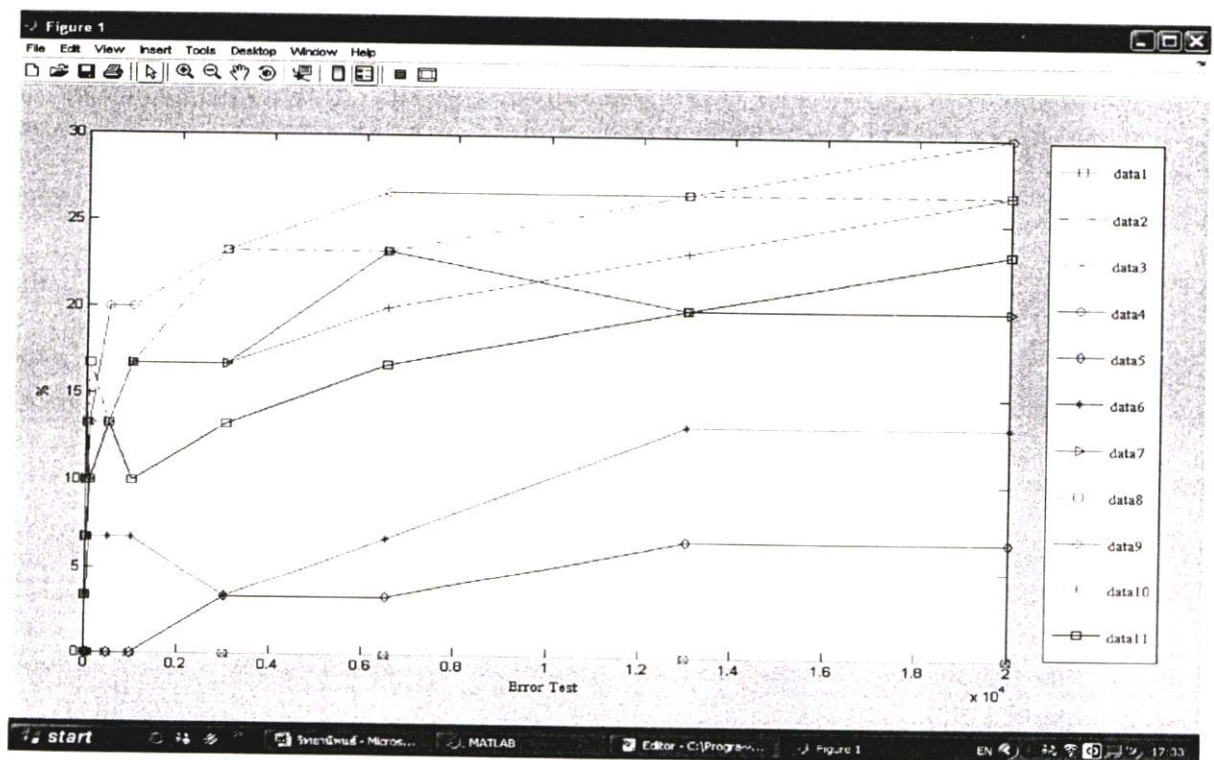
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0
200	0	3.3	3.3	6.7	6.7	6.7	10	10	13.3	13.3
500	0	0	6.7	10	6.7	10	10	10	16.7	16.7
1,000	0	0	6.7	6.7	10	10	10	13.3	20	26.7
2,000	0	3.3	3.3	10	10	16.7	16.7	16.7	20	23.3
5,000	0	3.3	3.3	6.7	6.7	13.3	10	13.3	13.3	16.7
10,000	0	0	3.3	3.3	6.7	6.7	10	13.3	16.7	20



รูปที่ 4.13 แสดงข้อมูลจากตารางที่ 4.2

ตารางที่ 4.3 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 50 พิกเซล

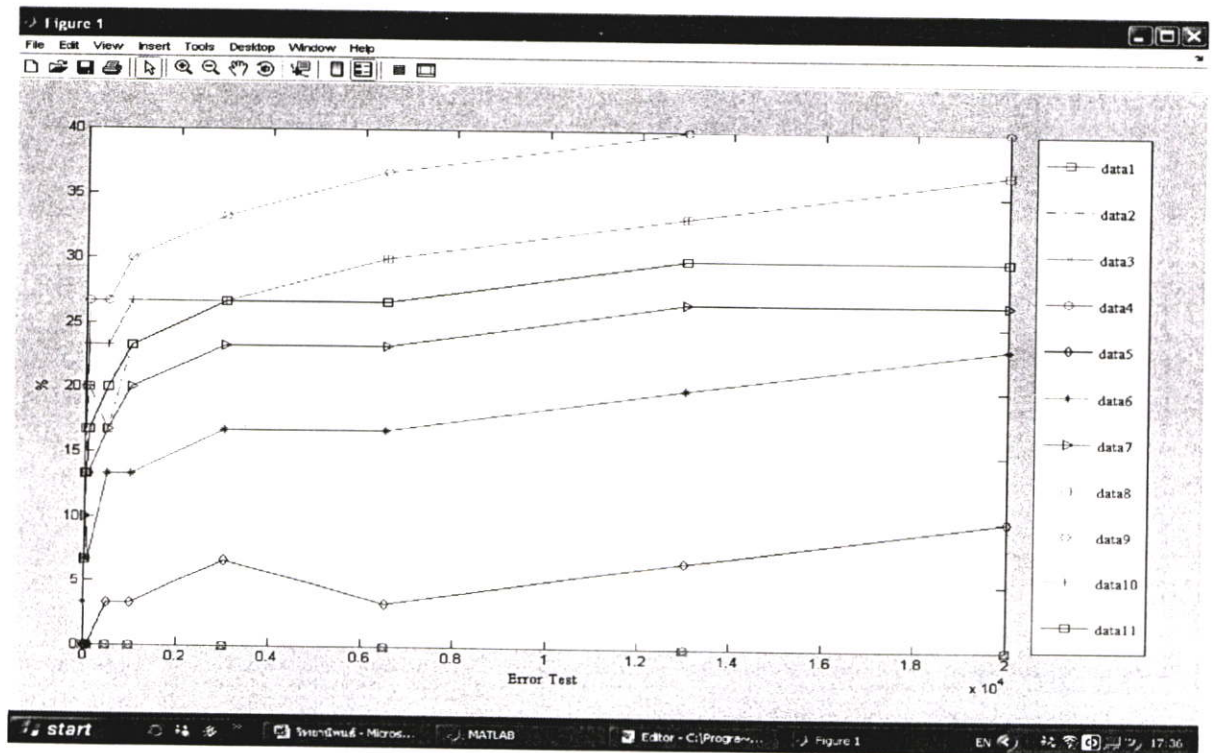
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	3.3	3.3	6.7	6.7
200	3.3	3.3	3.3	6.7	6.7	6.7	3.3	6.7	13.3	13.3
500	6.7	10	13.3	10	13.3	16.7	16.7	23.3	20	20
1,000	6.7	10	13.3	16.7	13.3	16.7	23.3	23.3	26.7	26.7
2,000	3.3	6.7	13.3	13.3	20	20	23.3	26.7	26.7	30
5,000	3.3	6.7	6.7	13.3	13.3	16.7	16.7	20	23.3	26.7
10,000	3.3	6.7	6.7	10	13.3	10	13.3	16.7	20	23.3



รูปที่ 4.14 แสดงข้อมูลจากตารางที่ 4.3

ตารางที่ 4.4 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ่อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 100 พิกเซล

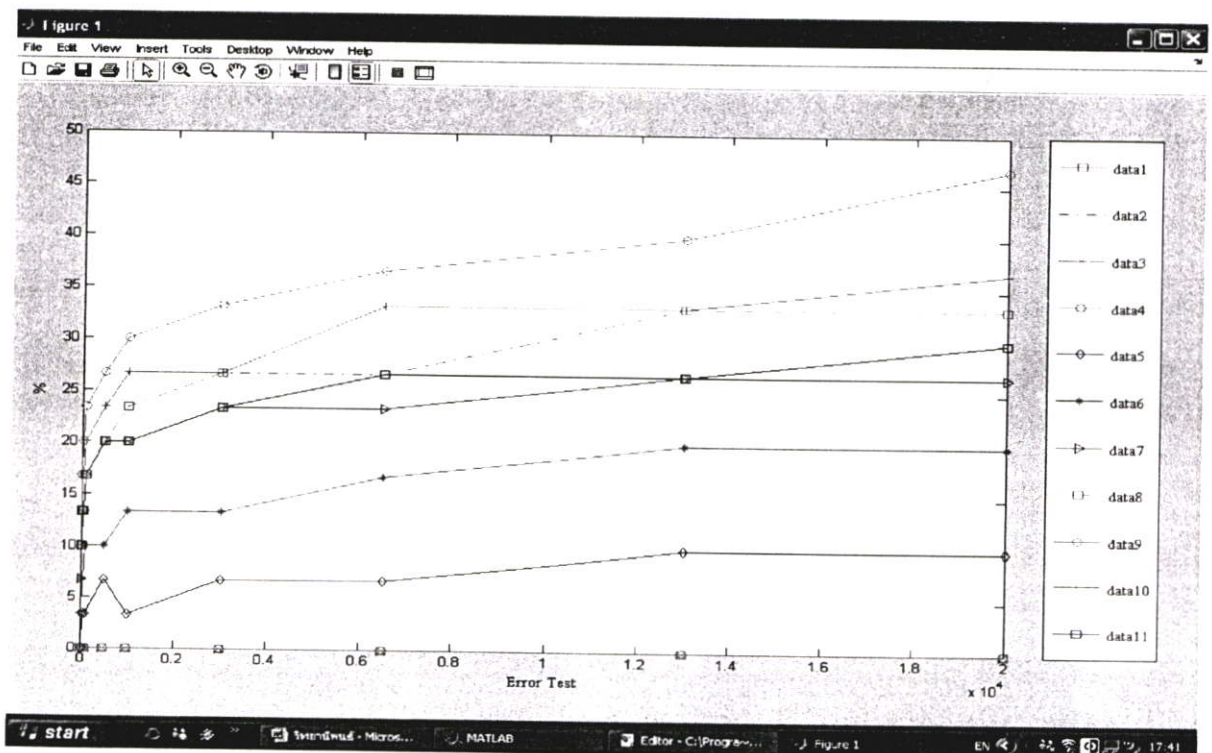
จำนวนเซลล์ในชั้นซ่อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	3.3	3.3	6.7	3.3	6.7	10
200	3.3	6.7	10	6.7	13.3	13.3	16.7	16.7	20	23.3
500	6.7	10	13.3	13.3	16.7	20	23.3	23.3	26.7	26.7
1,000	13.3	13.3	16.7	20	16.7	23.3	26.7	30	33.3	36.7
2,000	13.3	13.3	20	26.7	26.7	30	33.3	36.7	40	40
5,000	10	13.3	20	23.3	23.3	26.7	26.7	30	33.3	36.7
10,000	6.7	13.3	13.3	16.7	20	23.3	26.7	26.7	30	30



รูปที่ 4.15 แสดงข้อมูลจากตารางที่ 4.4

ตารางที่ 4.5 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 500 พิกเซล

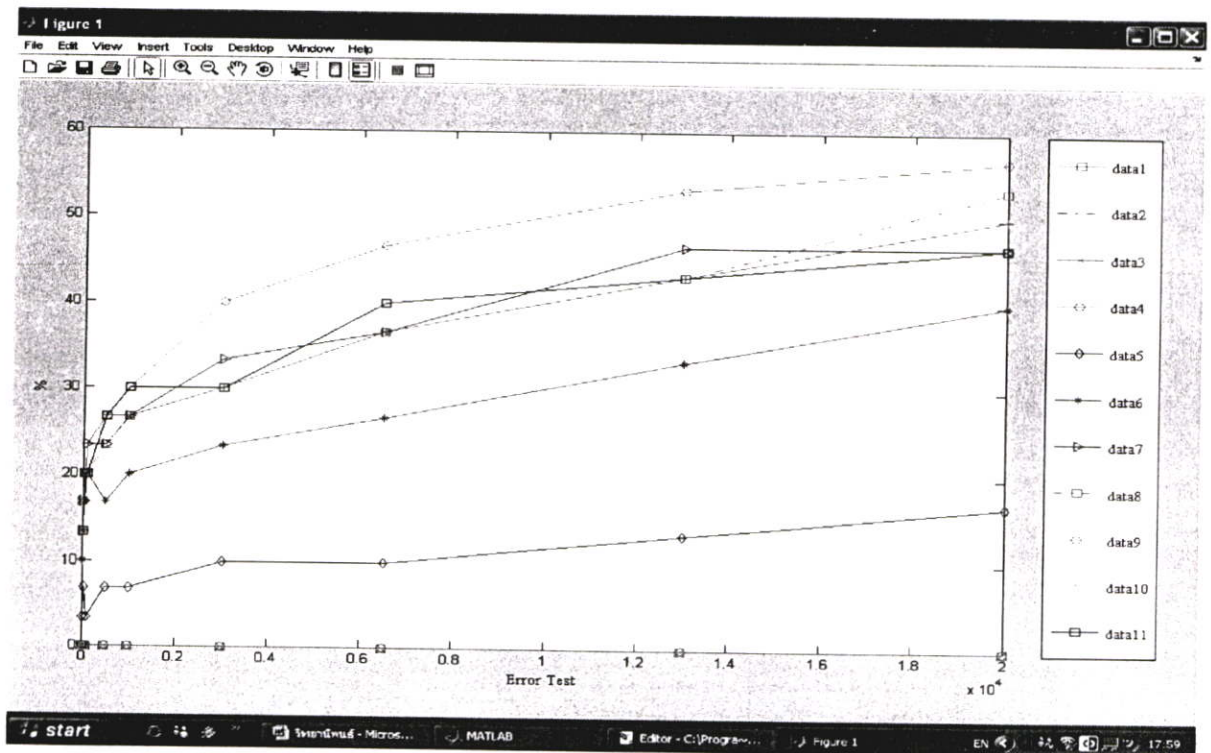
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	0	0	3.3	3.3	6.7	3.3	6.7	6.7	10	10
200	3.3	3.3	6.7	10	10	13.3	13.3	16.7	20	20
500	6.7	10	13.3	16.7	20	20	23.3	23.3	26.7	26.7
1,000	10	13.3	13.3	16.7	20	23.3	26.7	26.7	33.3	33.3
2,000	13.3	16.7	20	23.3	26.7	30	33.3	36.7	40	46.7
5,000	13.3	16.7	16.7	20	23.3	26.7	26.7	33.3	33.3	36.7
10,000	10	13.3	13.3	16.7	20	20	23.3	26.7	26.7	30



รูปที่ 4.16 แสดงข้อมูลจากตารางที่ 4.5

ตารางที่ 4.6 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 1,000 พิกเซล

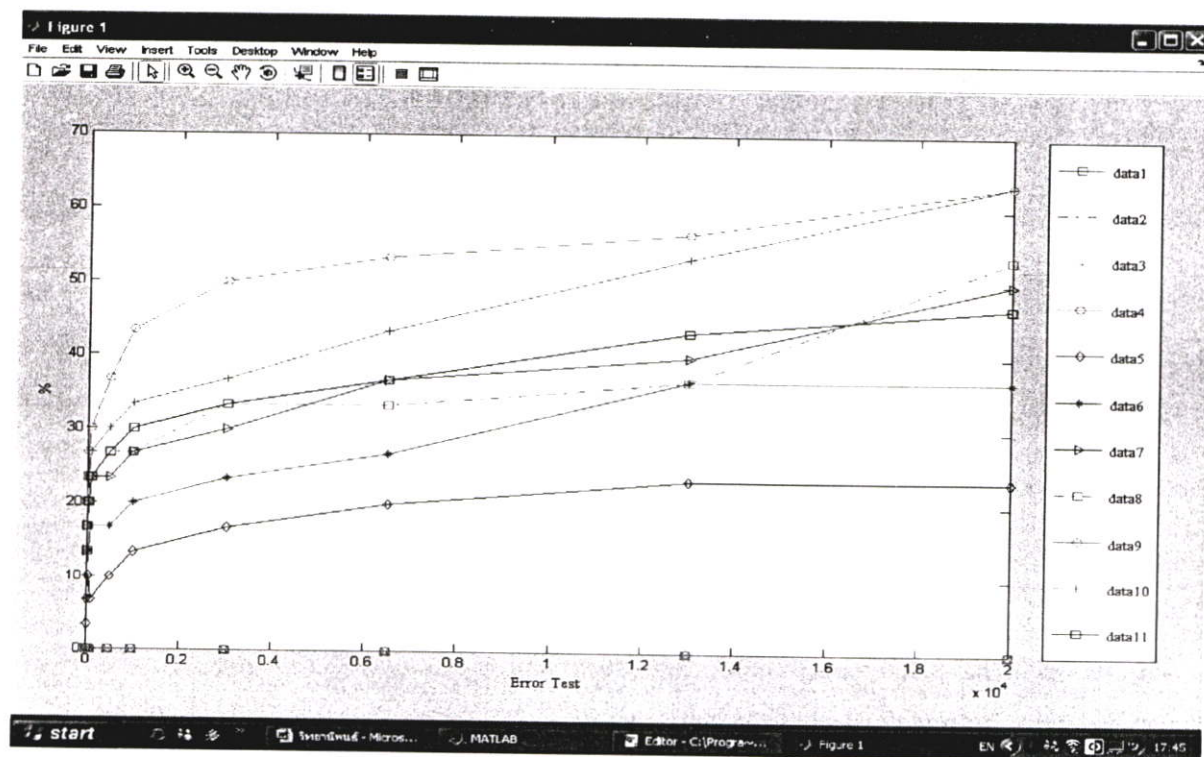
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	3.3	3.3	6.7	3.3	6.7	6.7	10	10	13.3	16.7
200	10	13.3	16.7	20	16.7	20	23.3	26.7	33.3	40
500	13.3	16.7	16.7	23.3	23.3	26.7	33.3	36.7	46.7	46.7
1,000	13.3	16.7	20	20	23.3	26.7	30	36.7	43.3	53.3
2,000	16.7	16.7	20	20	26.7	30	40	46.7	53.3	56.7
5,000	16.7	13.3	16.7	23.3	26.7	26.7	30	36.7	43.3	50
10,000	13.3	13.3	20	20	26.7	30	30	40	43.3	46.7



รูปที่ 4.17 แสดงข้อมูลจากตารางที่ 4.6

ตารางที่ 4.7 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 3,000 พิกเซล

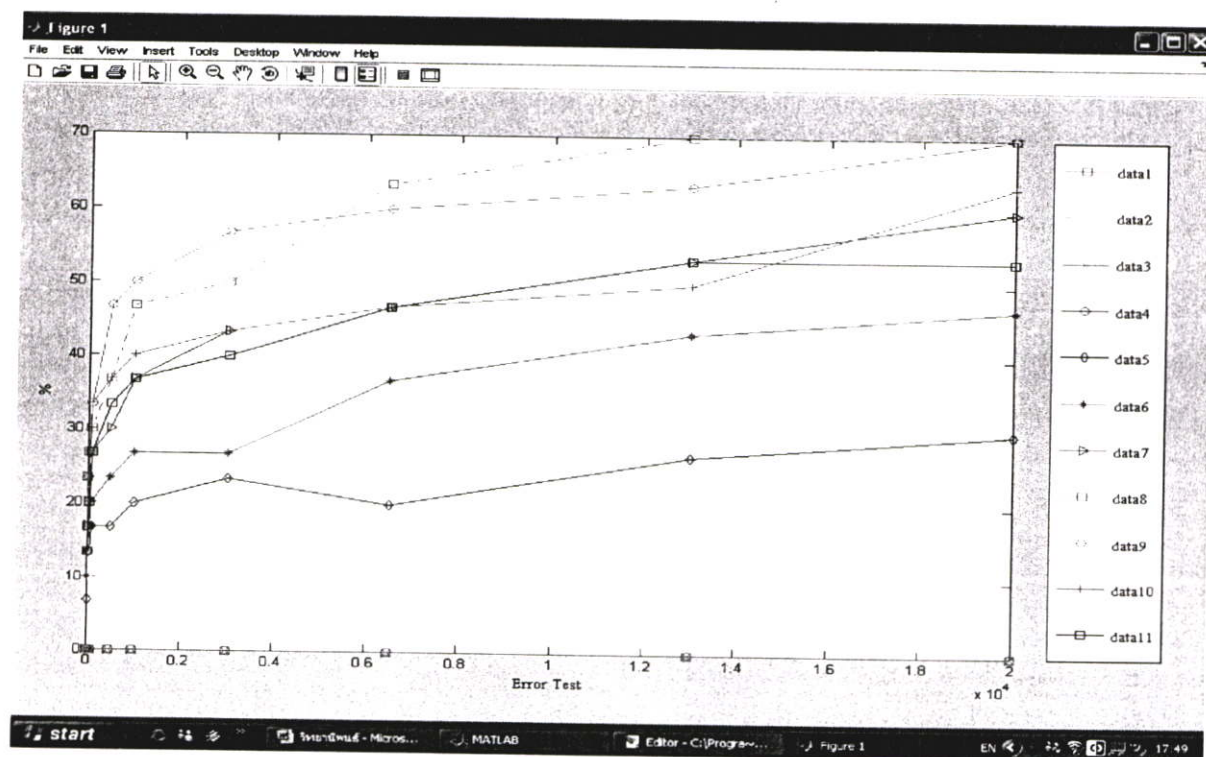
จำนวนเซลล์ในชั้นซ้อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	3.3	6.7	10	6.7	10	13.3	16.7	20	23.3	23.3
200	6.7	10	13.3	16.7	16.7	20	23.3	26.7	36.7	36.7
500	13.3	16.7	20	23.3	23.3	26.7	30	36.7	40	50
1,000	13.3	20	23.3	23.3	26.7	26.7	33.3	33.3	36.7	53.3
2,000	16.7	16.7	26.7	30	36.7	43.3	50	53.3	56.7	63.3
5,000	16.7	20	23.3	26.7	30	33.3	36.7	43.3	53.3	63.3
10,000	13.3	16.7	20	23.3	26.7	30	33.3	36.7	43.3	46.7



รูปที่ 4.18 แสดงข้อมูลจากตารางที่ 4.7

ตารางที่ 4.8 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ่อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 6,500 พิกเซล

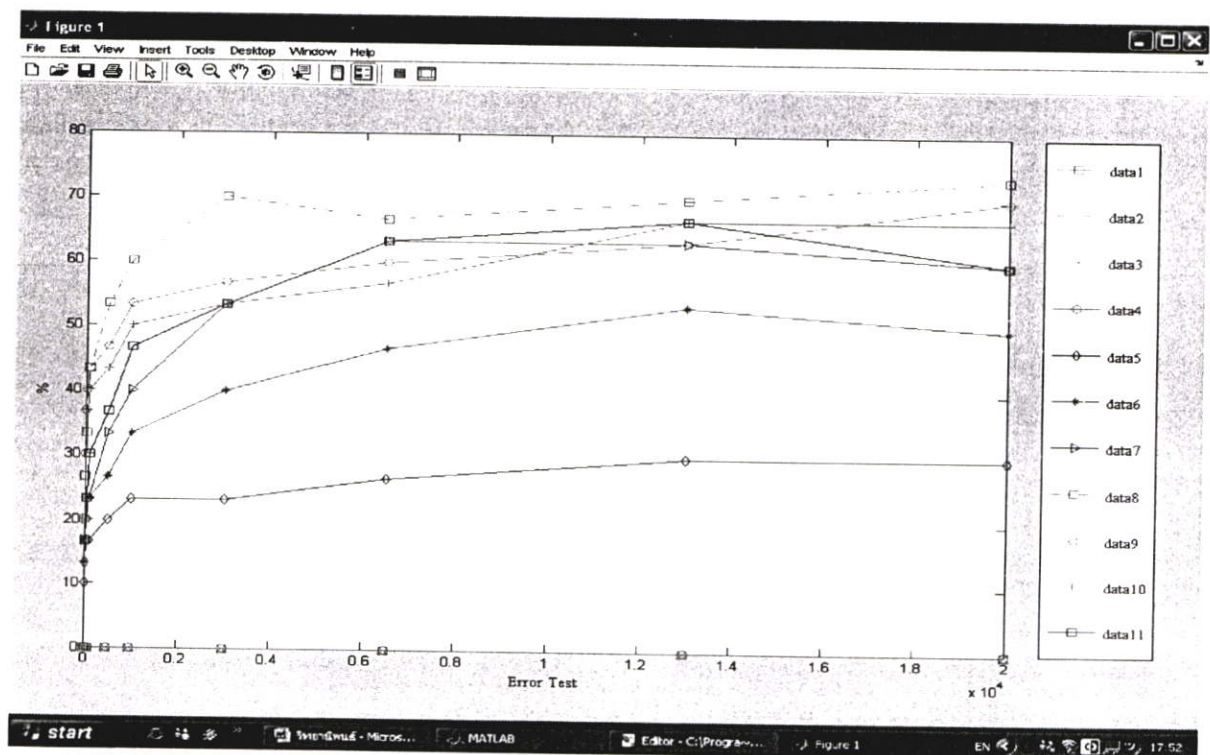
จำนวนเซลล์ในชั้นซ่อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	6.7	13.3	13.3	16.7	16.7	20	23.3	20	26.7	30
200	10	13.3	16.7	20	23.3	26.7	26.7	36.7	43.3	46.7
500	16.7	20	16.7	26.7	30	36.7	43.3	46.7	53.3	60
1,000	16.7	23.3	26.7	30	36.7	46.7	50	63.3	70	70
2,000	13.3	16.7	23.3	33.3	46.7	50	56.7	60	63.3	70
5,000	13.3	16.7	26.7	33.3	36.7	40	43.3	46.7	50	63.3
10,000	13.3	16.7	20	26.7	33.3	36.7	40	46.7	53.3	53.3



รูปที่ 4.19 แสดงข้อมูลจากตารางที่ 4.8

ตารางที่ 4.9 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ่อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 13,000 พิกเซล

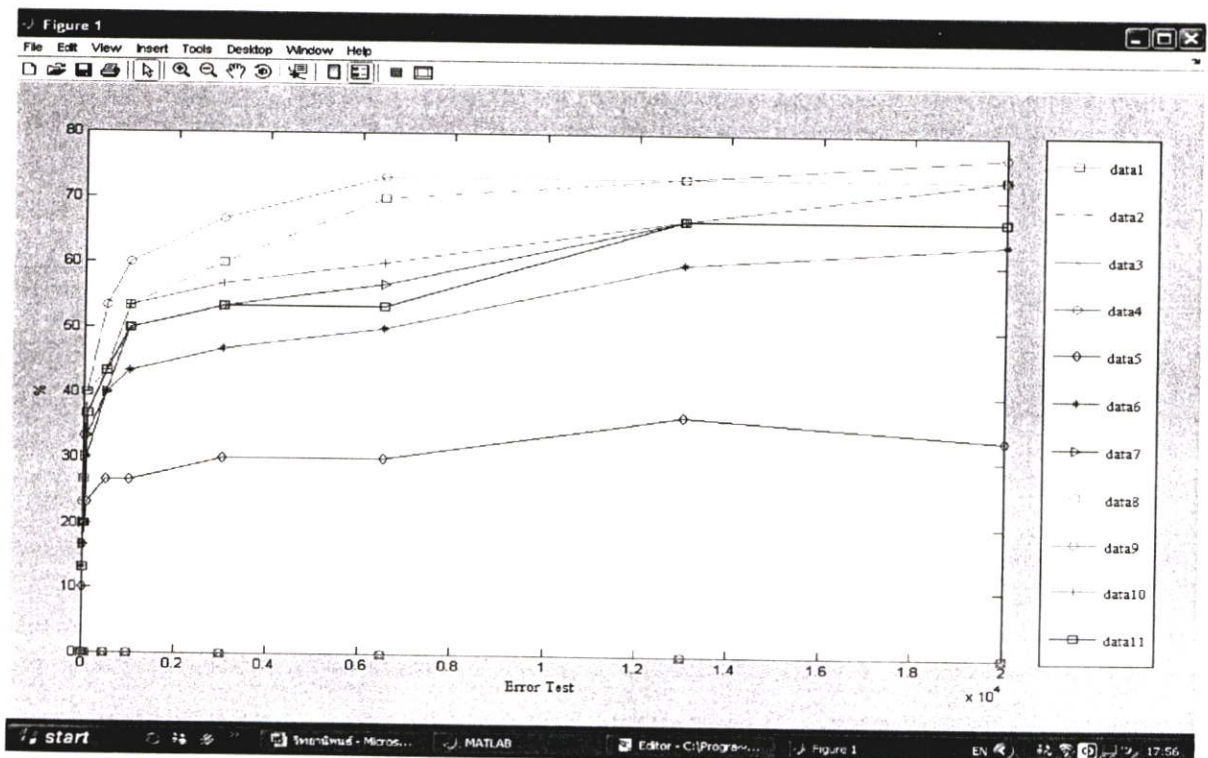
จำนวนเซลล์ในชั้นซ่อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	10	13.3	16.7	16.7	20	23.3	23.3	26.7	30	30
200	13.3	16.7	23.3	23.3	26.7	33.3	40	46.7	53.3	50
500	16.7	16.7	20	23.3	33.3	40	53.3	63.3	63.3	60
1,000	20	26.7	33.3	43.3	53.3	60	70	66.7	70	73.3
2,000	30	36.7	40	43.3	46.7	53.3	56.7	60	63.3	70
5,000	20	23.3	36.7	40	43.3	50	53.3	56.7	66.7	66.7
10,000	16.7	16.7	23.3	30	36.7	46.7	53.3	63.3	66.7	60



รูปที่ 4.20 แสดงข้อมูลจากตารางที่ 4.9

ตารางที่ 4.10 แสดงความถูกต้องในการจำแนกแบบ 1 ชั้นซ่อน เมื่อใช้จำนวนจุดผิดพลาดในการเรียนรู้ 20,000 พิกเซล

จำนวนเซลล์ในชั้นซ่อนที่ 1	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
	1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
100	10	16.7	20	23.3	26.7	26.7	30	30	36.7	33.3
200	13.3	16.7	26.7	30	40	43.3	46.7	50	60	63.3
500	16.7	20	30	33.3	40	50	53.3	56.7	66.7	73.3
1,000	20	20	26.7	40	43.3	53.3	60	70	73.3	73.3
2,000	20	23.3	33.3	40	53.3	60	66.7	73.3	73.3	76.7
5,000	16.7	23.3	26.7	36.7	43.3	53.3	56.7	60	66.7	73.3
10,000	13.3	20	20	36.7	43.3	50	53.3	53.3	66.7	66.7



รูปที่ 4.21 แสดงข้อมูลจากตารางที่ 4.10

จากกราฟแสดงข้อมูล

Data 1	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	1 เซลล์
Data 2	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	5 เซลล์
Data 3	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	10 เซลล์
Data 4	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	50 เซลล์
Data 5	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	100 เซลล์
Data 6	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	200 เซลล์
Data 7	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	500 เซลล์
Data 8	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	1,000 เซลล์
Data 9	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	2,000 เซลล์
Data 10	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	5,000 เซลล์
Data 11	คือ จำนวนเซลล์ในชั้นที่ 1 มีจำนวน	10,000 เซลล์

ตารางที่ 4.11 แสดงความถูกต้องในการจำแนกแบบ 2 ชั้นซ้อน เมื่อใช้จำนวนจุดผิดพลาดในการ  
เรียนรู้ 6,500 พิกเซล (ประมาณ 10%)

จำนวน เซลล์ในชั้น ซ้อนที่ 1	จำนวน เซลล์ในชั้น ซ้อนที่ 2	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ										
		1	10	50	100	500	1,000	3,000	6,500	13,000	20,000	
50	50	0	0	0	0	0	0	0	0	0	0	0
50	100	0	0	0	0	0	0	0	0	0	0	0
50	200	0	3.3	6.7	6.7	6.7	10	10.3	16.7	20	26.7	
50	500	0	3.3	6.7	6.7	10	13.3	13.3	16.7	20	23.3	
50	1,000	3.3	6.7	3.3	10	13.3	16.7	20	23.3	23.3	26.7	
50	2,000	3.3	6.7	6.7	10	13.3	16.7	16.7	20	20	23.3	
50	5,000	0	0	0	0	0	0	0	0	0	0	0
50	10,000	0	0	0	0	0	0	0	0	0	0	0
100	50	0	0	0	0	0	0	0	0	0	0	0
100	100	16.7	16.7	20	23.3	30	33.3	36.7	48.3	46.7	53.3	
100	200	10	16.7	20	26.7	26.7	30	33.3	43.3	43.3	46.7	
100	500	10	20	16.7	20	23.3	23.3	26.7	30	33.3	36.7	

ตารางที่ 4.11 (ต่อ)

จำนวน เซลล์ในชั้น ช่องที่ 1	จำนวน เซลล์ในชั้น ช่องที่ 2	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
		1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
100	1,000	6.7	10	16.7	20	23.3	26.7	26.7	30	30	33.3
100	2,000	6.7	13.3	13.3	16.7	16.7	16.7	20	23.3	26.7	30
100	5,000	0	0	0	0	0	0	0	0	0	0
100	10,000	0	0	0	0	0	0	0	0	0	0
200	50	10	16.7	20	23.3	23.3	30	33.3	43.3	43.3	50
200	100	16.7	16.7	23.3	23.3	26.7	30	33.3	36.7	43.3	46.7
200	200	10	13.3	16.7	20	23.3	33.3	36.7	43.3	53.3	60
200	500	6.7	6.7	10	13.3	20	26.7	33.3	40	43.3	50
200	1,000	3.3	6.7	6.7	10	16.7	20	20	23.3	30	36.7
200	2,000	3.3	3.3	6.7	13.3	16.7	16.7	20	23.3	26.7	33.3
200	5,000	0	0	0	0	0	0	0	0	0	0
200	10,000	0	0	0	0	0	0	0	0	0	0
500	50	0	3.3	6.7	13.3	16.7	16.7	20	20	23.3	26.7
500	100	3.3	10	16.7	20	20	23.3	26.7	33.3	36.7	40
500	200	3.3	6.7	13.3	16.7	20	23.3	26.7	40	46.7	56.7
500	500	6.7	10	16.7	20	26.7	33.3	40	56.7	66.7	73.3
500	1,000	6.7	6.7	10	16.7	23.3	30	33.3	43.3	60	70
500	2,000	3.3	3.3	6.7	10	16.7	20	30	36.7	46.7	53.3
500	5,000	3.3	6.7	10	13.3	16.7	20	26.7	33.3	43.3	46.7
500	10,000	0	0	0	0	0	0	0	0	0	0
1,000	50	3.3	6.7	10	13.3	16.7	16.7	20	23.3	26.7	30
1,000	100	6.7	10	16.7	20	23.3	26.7	30	33.3	36.7	43.3
1,000	200	6.7	10	16.7	20	23.3	26.7	30	33.3	40	46.7







ตารางที่ 4.12 (ต่อ)

จำนวน เซลล์ในชั้น ช่องที่ 1	จำนวน เซลล์ในชั้น ช่องที่ 2	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
		1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
5,000	50	0	0	0	0	0	0	0	0	0	0
5,000	100	0	0	0	0	0	0	0	0	0	0
5,000	200	0	0	0	0	0	0	0	0	0	0
5,000	500	3.3	6.7	10	16.7	23.3	26.7	30	33.3	40	46.7
5,000	1,000	0	0	0	0	0	0	0	0	0	0
5,000	2,000	0	0	0	0	0	0	0	0	0	0
5,000	5,000	0	0	0	0	0	0	0	0	0	0
5,000	10,000	0	0	0	0	0	0	0	0	0	0

ตารางที่ 4.13 แสดงความถูกต้องในการจำแนกแบบ 2 ชั้นช่อง เมื่อใช้จำนวนจุดผิดพลาดในการ  
เรียนรู้ 20,000 พิกเซล (ประมาณ 30 %)

จำนวน เซลล์ในชั้น ช่องที่ 1	จำนวน เซลล์ในชั้น ช่องที่ 2	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
		1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
50	50	0	0	0	0	0	0	0	0	0	0
50	100	0	0	0	0	0	0	0	0	0	0
50	200	6.7	10	10	13.3	13.3	16.7	16.7	20	33.3	36.7
50	500	6.7	10	6.7	10	13.3	16.7	20	26.7	30	33.3
50	1,000	6.7	6.7	10	13.3	16.7	20	20	23.3	26.7	30
50	2,000	6.7	10	10	13.3	16.7	16.7	20	23.3	23.3	30
50	5,000	0	0	0	0	0	0	0	0	0	0
50	10,000	0	0	0	0	0	0	0	0	0	0
100	50	0	0	0	0	0	0	0	0	0	0
100	100	10	16.7	20	26.7	30	33.3	36.7	40	46.7	56.7
100	200	10	13.3	20	26.7	30	33.3	40	43.3	46.7	53.3

ตารางที่ 4.13 (ต่อ)

จำนวน เซลล์ในชั้น ชั้นที่ 1	จำนวน เซลล์ในชั้น ชั้นที่ 2	จำนวนจุดผิดพลาดที่ใช้ในการทดสอบ									
		1	10	50	100	500	1,000	3,000	6,500	13,000	20,000
100	500	10	16.7	20	20	23.3	26.7	30	33.3	36.7	40
100	1,000	10	13.3	16.7	20	23.3	26.7	26.7	30	33.3	36.7
100	2,000	6.7	13.3	16.7	13.3	16.7	20	23.3	26.7	26.7	30
100	5,000	0	0	0	0	0	0	0	0	0	0
100	10,000	0	0	0	0	0	0	0	0	0	0
200	50	6.7	13.3	16.7	20	23.3	30	36.7	46.7	50	60
200	100	10	13.3	16.7	23.3	26.7	30	36.7	46.7	50	53.3
200	200	10	13.3	16.7	20	33.3	40	43.3	53.3	56.7	66.7
200	500	6.7	10	13.3	20	26.7	36.7	40	43.3	46.7	53.3
200	1,000	6.7	10	10	13.3	20	23.3	30	36.7	40	53.3
200	2,000	6.7	10	6.7	10	16.7	20	23.3	23.3	30	40
200	5,000	0	0	0	0	0	0	0	0	0	0
200	10,000	0	0	0	0	0	0	0	0	0	0
500	50	6.7	6.7	10	13.3	20	23.3	30	33.3	36.7	46.7
500	100	10	16.7	20	23.3	26.7	33.3	36.7	40	46.7	50
500	200	6.7	13.3	16.7	20	26.7	33.3	40	50	56.7	66.7
500	500	6.7	10	16.7	20	26.7	36.7	46.7	66.7	73.3	86.7
500	1,000	6.7	10	16.7	23.3	30	33.3	40	43.3	56.7	66.7
500	2,000	6.7	10	16.7	20	30	33.3	40	46.7	53.3	60
500	5,000	0	0	0	0	0	0	0	0	0	0
500	10,000	0	0	0	0	0	0	0	0	0	0
1,000	50	6.7	10	13.3	16.7	20	20	23.3	26.7	30	33.3
1,000	100	6.7	13.3	16.7	20	26.7	30	33.3	36.7	40	43.3
1,000	200	6.7	10	16.7	20	26.7	26.7	30	36.7	40	46.7



จากตารางที่ 4.1 ถึง 4.10 จะเป็นการแสดงถึงเปอร์เซ็นต์ความถูกต้องของการจำแนกเปลวเพลิงภายในภาพในโครงข่ายแบบ 1 ชั้นซ่อน จะสามารถได้ข้อสังเกตจากการจำแนกดังนี้

จำนวนเซลล์ในการจำแนกที่ได้ผลดีที่สุดจะอยู่ในช่วง 1,000 ถึง 5,000 เซลล์

เมื่อใช้จำนวนพิกเซลความผิดพลาด 1 ถึง 5 จุด ในการเรียนรู้และจำแนกจะมีความถูกต้องในการจำแนกต่ำมากจนถึงไม่สามารถจำแนกได้

เมื่อพิจารณาเวลาที่ใช้ในการฝึกโครงข่ายประสาทเทียมพบว่าที่เซลล์ประสาทที่ชั้นซ่อนเท่ากับ 100 เซลล์จะใช้เวลามากที่สุดซึ่งจะใช้เวลาประมาณ 15 วินาที หรือบางทีก็ไม่สามารถจำแนกได้ ต่อมาเมื่อเพิ่มเซลล์ประสาทในชั้นซ่อนขึ้นไปเรื่อยๆ พบว่าจะใช้เวลาใกล้เคียงกันคืออยู่ในช่วง 1-5 วินาที ซึ่งถือว่าไม่แตกต่างกันมากนัก

จากตารางที่ 4.11 ถึง 4.13 จะเป็นการแสดงถึงเปอร์เซ็นต์ความถูกต้องของการจำแนกเปลวเพลิงภายในภาพในโครงข่ายแบบ 2 ชั้นซ่อน จะสามารถได้ข้อสังเกตจากการจำแนกดังนี้

จากตารางที่ 4.11 จะสามารถจำแนกได้ถูกต้องสูงสุด 73.3 เปอร์เซ็นต์ โดยใช้จำนวนเซลล์ในชั้นที่ 1 และ ชั้นที่ 2 เป็น 500 และ 500 ตามลำดับ และความถูกต้องน้อยที่สุดที่จำแนกได้เป็นศูนย์

จากตารางที่ 4.11 เมื่อใช้จำนวนเซลล์ในชั้นซ่อน 100 และ 50 โครงข่ายจะไม่สามารถลู่เข้าได้ และ เมื่อใช้จำนวนเซลล์ในชั้นที่ 2 เป็น 10,000 เซลล์ โครงข่ายจะไม่สามารถลู่เข้าได้เช่นกัน

จากตารางที่ 4.12 จะสามารถจำแนกได้ถูกต้องสูงสุด 83.3 เปอร์เซ็นต์ โดยใช้จำนวนเซลล์ในชั้นที่ 1 และ ชั้นที่ 2 เป็น 500 และ 500 ตามลำดับ และความถูกต้องน้อยที่สุดที่จำแนกได้เป็นศูนย์

จากตารางที่ 4.12 เมื่อใช้จำนวนเซลล์ในชั้นซ่อน 100 และ 50 โครงข่ายจะไม่สามารถลู่เข้าได้ และ เมื่อใช้จำนวนเซลล์ในชั้นที่ 2 เป็น 10,000 เซลล์ โครงข่ายจะไม่สามารถลู่เข้าได้เช่นกัน

จากตารางที่ 4.13 จะสามารถจำแนกได้ถูกต้องสูงสุด 86.7 เปอร์เซ็นต์ โดยใช้จำนวนเซลล์ในชั้นที่ 1 และ ชั้นที่ 2 เป็น 500 และ 500 ตามลำดับ และความถูกต้องน้อยที่สุดที่จำแนกได้เป็นศูนย์

จากตารางที่ 4.13 เมื่อใช้จำนวนเซลล์ในชั้นซ่อน 100 และ 50 โครงข่ายจะไม่สามารถลู่เข้าได้ และ เมื่อใช้จำนวนเซลล์ในชั้นที่ 2 เป็น 10,000 เซลล์ โครงข่ายจะไม่สามารถลู่เข้าได้เช่นกัน

เมื่อพิจารณาเวลาที่ใช้ในการฝึกโครงข่ายประสาทเทียมของตารางที่ 4.11 ถึง 4.13 พบว่าจะใช้เวลาใกล้เคียงกันคืออยู่ในช่วง 2-7 วินาที ซึ่งถือว่าไม่แตกต่างกันมากนัก

## บทที่ 5

# บทสรุป และแนวทางในการพัฒนาต่อไป

จากการทดลองใช้โครงข่ายประสาทเทียมแพร่กระจายกลับแบบ 1 และ 2 ชั้นซ่อน ในการจำแนกเปลวเพลิงภายในภาพ ทำให้ทราบถึงวิธีการทำงานของโครงข่ายประสาทเทียม และประสิทธิภาพในการทำงานของการจำแนกเปลวเพลิงภายในภาพ

### 5.1 สรุป

โครงข่ายประสาทเทียมแบบ 2 ชั้นซ่อน จะให้ผลการจำแนกดีกว่าโครงข่ายประสาทเทียมแบบ 1 ชั้นซ่อน

เมื่อใช้จุดผิดพลาดในการเรียนรู้มากกว่า 13,000 จุด ขึ้นไป จะให้ค่าความถูกต้องในการจำแนกที่ดี

โครงข่ายประสาทเทียมแบบ 1 ชั้นซ่อน เมื่อใช้จำนวนเซลล์ 1,000 ถึง 5,000 เซลล์ จะให้ความถูกต้องในการจำแนกสูงที่สุด

โครงข่ายประสาทเทียมแบบ 2 ชั้นซ่อน เมื่อจำนวนเซลล์ในชั้นซ่อนที่ 1 และ ชั้นซ่อนที่ 2 คือ 500 และ 500 เซลล์ ตามลำดับ จะให้ความถูกต้องในการจำแนกสูงที่สุด

### 5.2 วิจารณ์และข้อเสนอแนะ

ในการรับภาพควรคำนึงถึงความสว่างของแสง ต้องให้ความสว่างเพียงพอและสม่ำเสมอทั้งภาพ เพื่อให้ข้อมูลภาพเปลวเพลิงในการป้อนเข้าโครงข่ายประสาทเทียมสมบูรณ์ที่สุด

ในการจัดเตรียมข้อมูลภาพ เช่น การทำ Thresholding และ Contrast มีความสำคัญมาก ซึ่งถ้าปรับได้ดีก็จะสามารถให้ผลในการจำแนกดีตามไปด้วย

ในการทดลองในแต่ละครั้งจะมีข้อมูลของค่าตัวแปรต่างๆหลายค่า ดังนั้นควรมีการจดบันทึก หรือมีวิธีการเก็บข้อมูลที่ดีเพื่อป้องกันความสับสนหรือการสูญหายที่อาจเกิดขึ้นได้

### 5.3 แนวทางในการพัฒนาต่อไป

ควรใช้คอมพิวเตอร์ที่มีทรัพยากรมากเพียงพอสำหรับการฝึกการเรียนรู้ จำ โดยให้คอมพิวเตอร์มีหน่วยความจำ(RAM) มากขึ้น และมีหน่วยประมวลผลกลาง (CPU) มีความเร็วเพิ่มขึ้น จะทำให้การฝึกใช้เวลาน้อยลง

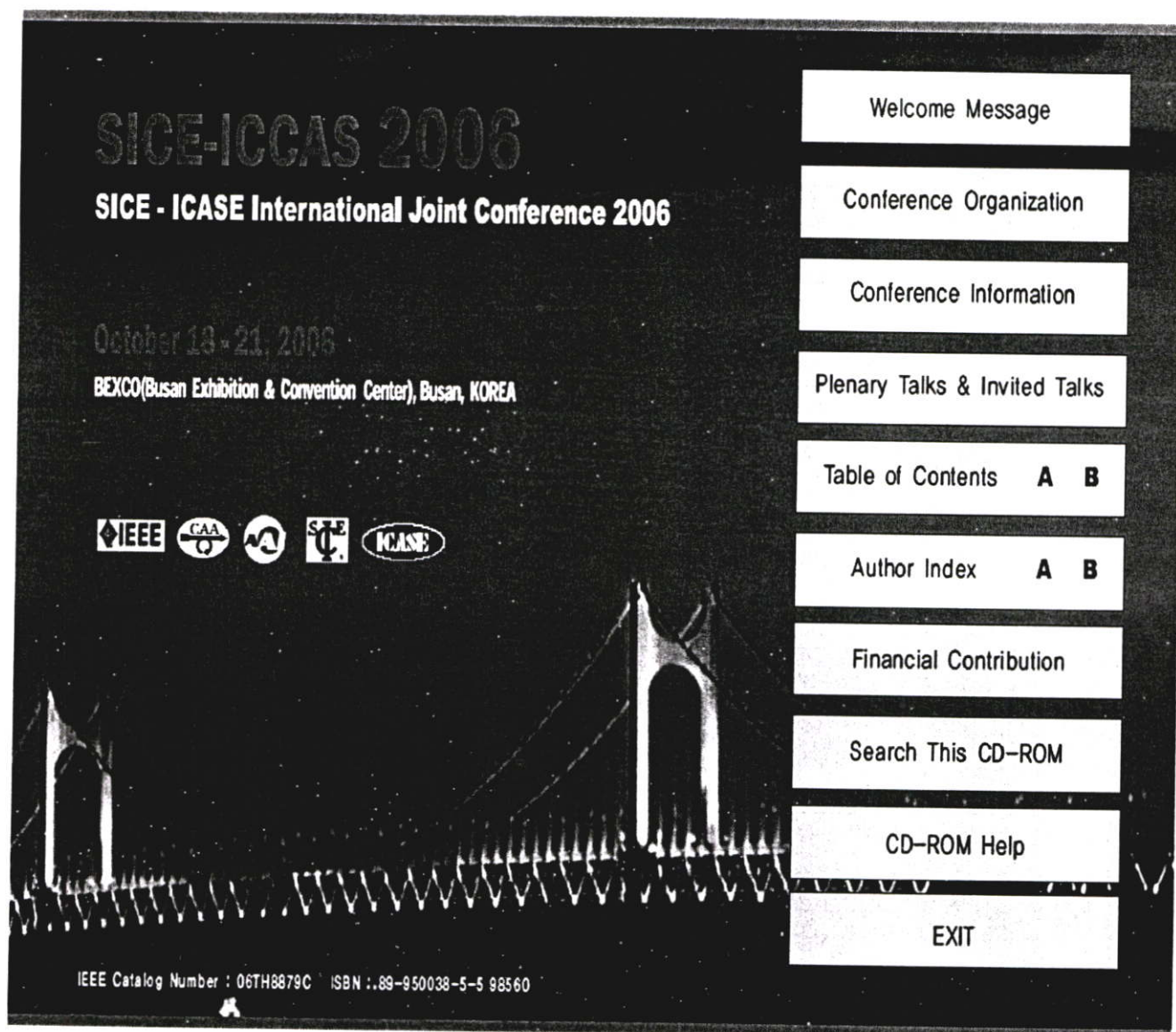
ในอนาคต สามารถประยุกต์ใช้การตรวจจับเปลวเพลิงภายในภาพนี้ เพื่อประยุกต์เป็นตัวตรวจจับเพลิงไหม้ภายในอาคาร, ในโรงงานอุตสาหกรรม , ไฟไหม้ป่า หรือไฟไหม้ที่อื่นๆ ได้

## เอกสารอ้างอิง

- [1] R.C.Gonzaleez and R.E. Woods, “Digital Image Processing”, 2<sup>nd</sup> Edition, United States of America:, Addison Wesley Publishing Company, 1992.
- [2] R.C. Gonzalez and P. Wintz, “Digital Image Processing”, 2<sup>nd</sup> Edition, Addison Wesley Publishing Company, Reading, Massachusetts, 1987.
- [3] J.G Proakis and D.G. Manolakis, “Digital singnal processing: principles, algorithms and applications,” 2<sup>nd</sup> edition, Maxwell Macmillan International Editions, 1992.
- [4] R.Michael Hord, “Digital Image Processing of Remotely Sensed Data”, Academic Press, New York, 1982.
- [5] Gregory A. Baxes, “Digital Image Processing : Principles and Applications”, John Wiley & Sons, Inc. New York, 1994
- [6] John A. Richards, “Remote Sensing Digital Image Analysis: An Introduction”, Springer-Verlag, Berlin, 1986.
- [7] H.A. Andrews, “Computer Techniques in Image Processing”, Academic Press, New York, 1970.
- [8] J.M. Zurada, Introduction to Artificial Neural System, West Publishing Company, Singapore, 1992
- [9] ราตรี สุดทรวง, ประสาทศรวิทยา, สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, พิมพ์ครั้งที่ 2 , 2535
- [10] A.A. Kempka, “Artificial Neural Networks: Part I”, AI Expert, June, 1994.
- [11] A.A. Kempka, “Artificial Neural Networks: Part I”, AII Expert, June, 1994.
- [12] D.E. Rumelhart, and J.L. McClelland, and the PDP Research Group, Parallel Distributed Processing; Exploration in the Microstructure of Cognition, Volume 1: Foundation, MIT Press Massachusetts, USA., 1986.
- [13] D.E. Rumelhart, and J.L. McClelland, Exploration in Parallel Distributed Processing; A Handbook of Model Programs; and Exercises, MIT Press Massachusetts, USA., 1988.
- [14] J.A. Freeman, and D.M. Skapara, Neural Networks, Applications, and Programming Techniques, Addison Wesley, USA., 1991.

- [15] กิตติ ไพฑูรย์วัฒนกิจ สาริต อินทจักร์ และ ไพธิน บุญเดช, “นิเวรอลเน็ตเวิร์คบนทรานสพิวเตอร์กับการหาขอบภาพ,” ประชุมใหญ่วิชาการทางวิศวกรรมประจำปี 2537, วิศวกรรมสถานแห่งประเทศไทยฯ, (2537) หน้า EE 75-90
- [16] T. Kohonen, Self-Organization and Associative Memory, Springer-Verlag, USA., 1989.
- [17] กิตติ ไพฑูรย์วัฒนกิจ และ เศรษฐพล ลิมปราชญา, “การใช้นิเวรอลเน็ตเวิร์คเพื่อจดจำแพทเทิร์น,” วารสารสำนักงานคณะกรรมการวิจัยแห่งชาติ, ปีที่ 25 เล่มที่ 1, (ม.ค. - มิ.ย. 2536) หน้า 45-67
- [18] J.E. Dayhoff, Neural Network Architectures an Introduction, Van Nostrand Reinhold, 1990.

ภาคผนวก  
ผลงานที่ได้รับการตีพิมพ์



SICE-ICASE International Joint Conference 2006  
Oct. 18-21, 2006 in Bexco, Busan, Korea

### Fire detection using neural network

V. Tipsuwanporn<sup>1</sup>, V. Krongratana<sup>2</sup>, S. Gulpanich<sup>3</sup> and K. Thongnopakun<sup>4</sup>  
<sup>1,2,3,4</sup> Department of Instrumentation Engineering, Faculty of Engineering,  
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand  
(Tel. +66 2326 7347 Ext. 102; E-mail: ktvittay@kmitl.ac.th)

**Abstract:** This study aims to investigate the effectiveness of image processing by using artificially neural network to detect fire in the room by using camera. This uses features of flame images in a simulation fire as input elements. Flame images are taken from the dynamic image, and features of the images for the ANN application are extracted from a rectified image which considers the difference in distance from the CCD color to the fire source.

**Keywords:** image processing, neural network.

## 1. INTRODUCTION

In this paper, studies an application of image processing by using artificially neural network to detect fire in the room by using camera. In our study, an image from CCD color is analyzed to construct a set of features to feed into ANN, which uses the features of image flame in a simulation fire occurring as input elements. Features of the image for the ANN are extracted from the rectified image which considers the difference in distance from the CCD to the fire source. From our study, the distinctive relation Red intensity and Green intensity in the case of fire flame is known. ANN is employed to identify fire that happen in the room, the model of ANN is hierarchical type in that rooming with back propagation mechanism. From the experimentation, the proposed model can detect the fire occurring.

## 2. EXAMPLE FIRE EXPERIMENT

Fig1. Show the example fire experiment of image processing to detect fire in the room. Determine is area  $4 \times 5 \times 4$  m. by size image flame approximately 30cm. and distance far away between CCD camera and fire source approximately 2m. Therefore example is the histogram to compare difference image flame and image non-flame.



Fig.1 Simulation fire experiment.

## 3. FLOW OF PROCESSING

The fire is detected by image processing and the ANN from the photographed dynamic image is group of 30 frames per second and one image is taken from ever fame.

An image in the state where non flames is used as the background image. Since flame in an image has the feature that its Red intensity is generally high. Red intensity differences between the background image and each frame are calculated

Finally, the fire is judged by the hierarchical type ANN with the feature parameters as input elements. The above processing method is repeated.

Fig. 2 [1-2] show the processing sequence for a one-frame image

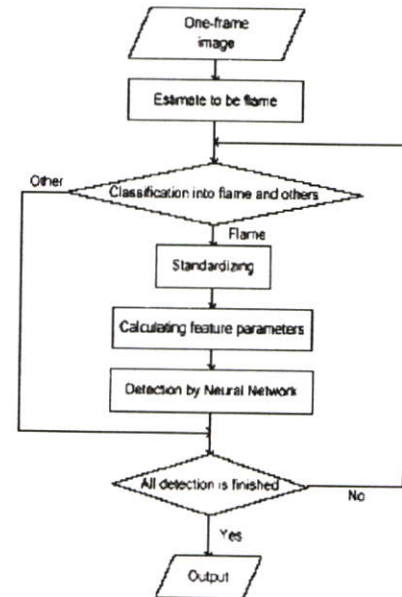


Fig.2 Flow of processing.

#### 4. EXTRACTION OF FLAME IMAGE

Each color corresponds to the basic spectrum factors of red R, green G, and blue B in the RGB model. The color model is based on the Cartesian coordinate system. The color subspace is a cube, as shown in Fig. 3 [6]. The RGB value consists of three vertexes, one black point, and one white point, where black and white are the original point and the farthest vertex point from the original point, respectively. The grayscale color is spread along the line that connects two white points from the black color in this model, and the color is on the cube space or the inside point that is defined as the vector that is extended from the original point.

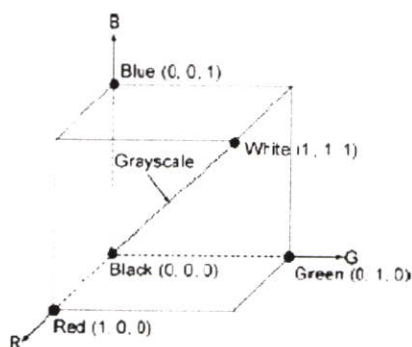


Fig. 3 RGB color cube.

In this paper, the original RGB images are applied as images converted to grayscale. The following conversion (1) is a conceptual expression as shown in Fig. 3 [6]. However, (2) is used in practical applications. In this paper, (2) is also applied to convert the RGB color to grayscale.

$$\text{Gray} = \frac{R + G + B}{3} \quad (1)$$

$$\text{Gray} = (0.30 \times R) + (0.59 \times G) + (0.11 \times B). \quad (2)$$

For comparison, a value of each color in a pixel  $(x, y)$  in the one-frame image  $f_N$  is defined as  $f_N(x, y)$  [3,4] (Equation (3)). A value in the background image is similarly defined as  $f_B(x, y)$  (Equation (4)), and these difference values  $f_{out}(x, y)$  are calculated by formula (5). This is the threshold defined by considering noise. This processing for the all pixels extracts pixels estimated to be flame.

$$f_N(x, y) = (R_N(x, y), G_N(x, y), B_N(x, y)) \quad (3)$$

$$f_B(x, y) = (R_B(x, y), G_B(x, y), B_B(x, y)) \quad (4)$$

$$f_{out}(x, y) = f_N(x, y) \text{ if } R_N(x, y) - R_B(x, y) > Th.$$

$$f_{out}(x, y) = (0, 0, 0) \text{ if } R_N(x, y) - R_B(x, y) \leq Th \quad (5)$$

Where

$f_N(x, y)$ : Pixel in one-frame image

$R_N(x, y), G_N(x, y), B_N(x, y)$ :

Red, Green, Blue intensity of  $f_N(x, y)$

$f_B(x, y)$ : Pixel in background image

$R_B(x, y), G_B(x, y), B_B(x, y)$ :

Red, Green, Blue intensity of  $f_B(x, y)$

$f_{out}$ : Pixel in output image

$Th$ : Threshold

#### 5. FIRE DETECTION BY ANN

The fire is judged by the hierarchical type ANN with the feature parameters as input element. Fig. 4 is the outline of the hierarchical type ANN used for this detection [5]. Input data is a total of Red, Green and Blue. The middle class and output layers two units of a fire output and a non-fire output.

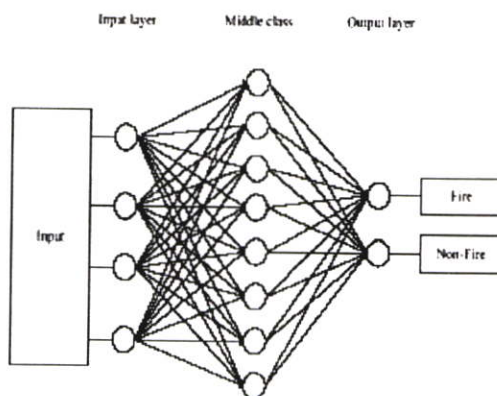


Fig. 4. Outline of a hierarchical type ANN.

If a fire output is 0.7 or more and a non-fire output is 0.3 or less, it is defined as the fire and its opposite is defined as a non-fire. Moreover, with other outputs, it is defined as judgment being impossible.

For the classification into flame and other, the value of each color in the pixels investigated. From our present study, the distinctive relation between Red intensity and Green intensity in the case of a fire flame is known. Fig.5 is the relation of the pixels. It shows the value is distribution in a specific zone.

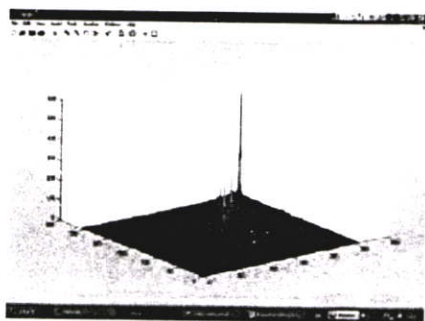


Fig.5. Color area of flame between Red intensity and Green intensity.

Fig.6, Fig.7, and Fig.8. It show extraction of flame by difference processing and histogram.

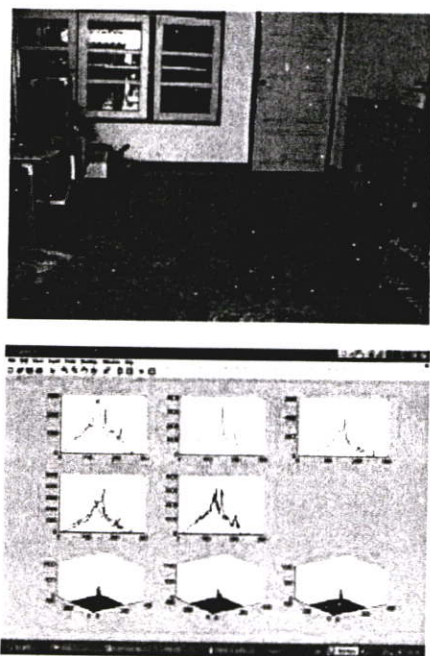


Fig.6. Show histogram of non-flame processing.

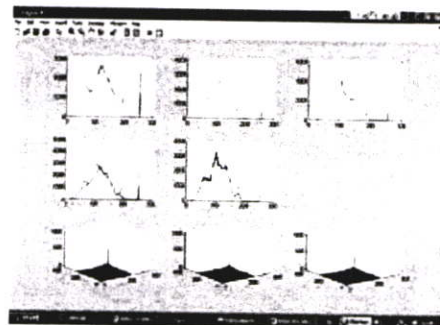


Fig.7. Show histogram of flame processing.

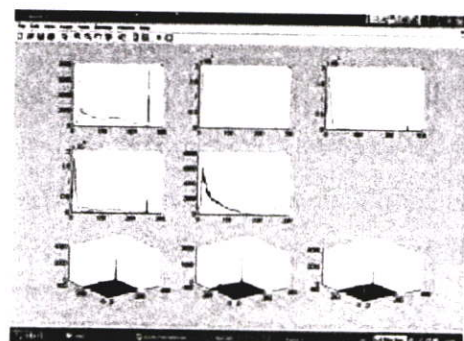
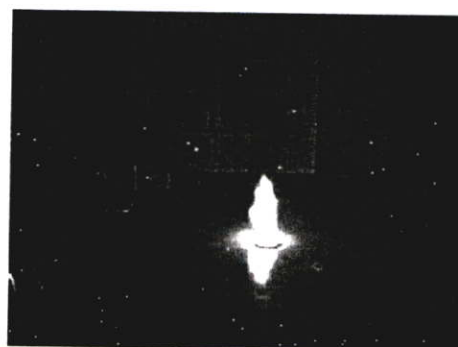


Fig.8. Show histogram of flame processing at night.

## 6. CONCLUSION

In this paper the ANN can detection of image flame information which considers the difference from the CCD color. From the histogram we have seen to difference between histogram of non-flame and histogram of flame are used as input elements for color information, training by the ANN.

The future can be applying use detection fire in the building, the factory and the forest etc.

## REFERENCES

- [1] T. Ono.,H. Ishii and K. Kawamura, "Application of neural network to analyses of CCD colour TV-camera image for the detection of car fires in expressway tunnels," *Fire Safety Journal*, 2005.
- [2] T. Ono.,H. Ishii and K. Kawamura, "Fire detection in road tunnel by neural network applied to itv-camera images. *JAFSE annual symposium 2003*.pp.142-3, 2003.
- [3] W.K. Pratt, *Digital Image Processing*, Second Edition, USA: Jon Wiley & Sons, Inc 1991.
- [4] R.C.Gonzalez and R.E. woods, *Digital Image Processing*, USA: Addison-Wiley, 1992.
- [5] S. Lu and A. Szeto, "Hierachical Artificial Neural Networks for Enhancement,"*Pattern Recognition*, Vol.26, No.8, pp.1149-1163, 1993
- [6] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. New York: McGraw-Hill, 1995.

## ประวัติผู้เขียน

ชื่อ-นามสกุล	นาย ไกรสิทธิ์ ทองนพคุณ
วัน เดือน ปีเกิด	เกิดเมื่อวันที่ 9 มีนาคม พ.ศ. 2520 ที่จังหวัดสงขลา
ที่อยู่	บ้านเลขที่ 45 ถนนเทศบาล 16 ตำบลพะตง อำเภอหาดใหญ่ จังหวัดสงขลา 90230
ประวัติการศึกษา	ปีการศึกษา 2542 สำเร็จการศึกษาหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมอิเล็กทรอนิกส์ จากมหาวิทยาลัยเอเชียอาคเนย์ ปีการศึกษา 2538 สำเร็จการศึกษาระดับประกาศนียบัตรวิชาชีพ สาขา อิเล็กทรอนิกส์ จากวิทยาลัยเทคนิคหาดใหญ่ ปีการศึกษา 2535 สำเร็จการศึกษาระดับชั้นมัธยมศึกษาปีที่ 3 จากโรงเรียน หาดใหญ่วิทยาลัย