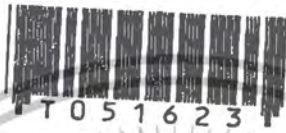


สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การรู้จำรูปภาพไบนารีโดยใช้ฮาดอร์ฟอาร์ทแมพ

A BINARY IMAGE RECOGNITION USING HAUSDORFF ARTMAP

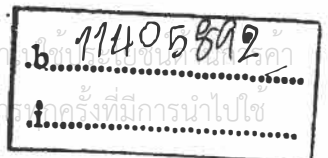


วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เลขหมู่.....
เลขทะเบียน **51623**
วัน,เดือน,ปี **6 ก.ค. 2547**

พ.ศ.2547

ISBN 974-9709-21-7



ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A BINARY IMAGE RECOGNITION USING HAUSDORFF ARTMAP



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

2004

ISBN 974-9709-21-7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2004

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การรู้จำรูปภาพใบนารีโดยใช้ฮาลูสตอร์ฟอาร์ทแมพ
นักศึกษา	นางสาวจงกลณี รุ่งเรือง
รหัสประจำตัว	44067042
ปริญญา	วิทยาศาสตร์มหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
พ.ศ.	2547
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ผศ.ดร.อาริต ธรรมโน

บทคัดย่อ

ในการรู้จำรูปแบบของมนุษย์นั้นมักอาศัยการจดจำโครงร่างแบบคร่าวๆ หรือภาพรวมของรูปแบบนั้นๆ งานวิจัยนี้จึงพยายามที่จะเลียนแบบการรู้จำของมนุษย์โดยใช้พีเจอาร์แบบรวมของรูปแบบที่เป็นภาพขาวดำ เช่น เส้น โครงร่าง เป็นต้น มาใช้แทนการหาพีเจอาร์แบบอื่นที่ซับซ้อนกว่า โดยได้นำเอาฮาลูสตอร์ฟดิสแทนซ์มาใช้ในการวัดความเหมือนของรูปแบบ เนื่องจากวิธีการนี้เป็นการวัดความเหมือนที่ใกล้เคียงกับวิธีการของมนุษย์มากกว่าการวัดความเหมือนแบบอื่นๆ และนอกจากนี้แล้วในการรู้จำรูปแบบได้นำเอาแนวทางของอาร์ทมาประยุกต์ใช้ร่วมกับฮาลูสตอร์ฟดิสแทนซ์ด้วยเพื่อช่วยแก้ไขปัญหาการรู้จำรูปแบบชนิดเดียวกันแต่มีหลายลักษณะจากการดำเนินการวิจัยและทดลองปรากฏว่าแนวทางใหม่นี้ให้ผลลัพธ์ที่ดีในการรู้จำรูปแบบตัวอย่างเช่น ลายนิ้วมือ, โครงร่างมือ, และใบหน้าของมนุษย์

Thesis Title	A Binary Image Recognition Using Hausdorff ARTMAP
Student	Miss Chongkolnee Rungruang
Student ID.	44067042
Degree	Master of Science
Year	2004
Thesis Advisor	Asst.Prof.Dr.Arit Thammano

ABSTRACT

Humans can categorize a pattern at a glance, because they recognize its global view and structure. In this research, the main idea is based on this human behavior; therefore, the global features of binary images such as edges and contours are used instead of complicated features. In addition, Hausdorff distance was chosen as a similarity measure, for it has been accepted more closely to human performance than other similarity metrics. For this pattern recognition, the Hausdorff distance is adapted to ART approach to solve the problem that there are different appearances in a class of pattern. From the experimental results of fingerprint, hand geometry, and human face recognition show that this new approach results in a good recognition rate.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้คงจะสำเร็จลุล่วงไปไม่ได้เลยหากไม่ได้รับคำปรึกษาแนะนำและความอนุเคราะห์จาก ผศ.ดร.อาริต ธรรมโน ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ที่ท่านได้ให้ความสนใจห่วงใยและติดตามความก้าวหน้าของการทำวิทยานิพนธ์อย่างใกล้ชิดเสมอมา ตลอดจนให้คำปรึกษาและข้อเสนอแนะที่เป็นประโยชน์อย่างยิ่งในการทำวิทยานิพนธ์แก่ข้าพเจ้า ข้าพเจ้ารู้สึกซาบซึ้งและขอขอบพระคุณเป็นอย่างสูง

ขอขอบคุณครอบครัวอันเป็นที่รักของข้าพเจ้าประกอบไปด้วยคุณพ่อ คุณแม่ และน้องชายที่เป็นกำลังใจอันดีให้แก่ข้าพเจ้าตลอดมา ข้าพเจ้าคงจะไม่สามารถผ่านพ้นอุปสรรคและความท้อแท้ใจไปได้เลยหากขาดกำลังใจและการสนับสนุนจากครอบครัว ขอขอบคุณญาติๆ ทุกคนที่เป็นกำลังใจให้ข้าพเจ้าด้วยเช่นกัน

นอกจากนี้แล้วข้าพเจ้าขอกราบขอบพระคุณครูอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทความรู้ให้แก่ข้าพเจ้าตั้งแต่อดีตจนถึงปัจจุบัน

สุดท้ายนี้ข้าพเจ้าขอขอบคุณเพื่อนๆ และพี่ๆ นักศึกษาทุกคนที่เป็นกำลังใจ ช่วยเหลือและเต็มใจให้คำปรึกษาในการทำวิทยานิพนธ์แก่ข้าพเจ้ามาโดยตลอด

จงกลณี รุ่งเรือง

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญรูป	VIII
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	1
1.3 สมมติฐานของการศึกษา	2
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย	2
1.5 ขอบเขตของการวิจัย	3
1.6 ขั้นตอนของการศึกษา	3
บทที่ 2 หลักการและทฤษฎีเบื้องต้นที่นำไปสู่นำแนวคิดของการวิจัย	4
2.1 แนวคิดในการมองเห็นและรู้จำวัตถุของมนุษย์	4
2.2 นีโอค็อกนิตรอนนิวรอลเน็ตเวิร์ค	5
2.3 ฮาอูสเตอร์ฟิดิสแทนซ์	7
2.4 ซิมพลิไฟด์ฟิชชี่อาร์ทแมพ	14
2.4.1 ความรู้เบื้องต้นเกี่ยวกับอาร์ทและซิมพลิไฟด์ฟิชชี่อาร์ทแมพ	14
2.4.2 กระบวนการทำงานของซิมพลิไฟด์ฟิชชี่อาร์ทแมพ	16
บทที่ 3 การรู้จำรูปภาพใบนารีโดยใช้ฮาอูสเตอร์ฟอาร์ทแมพ	19
3.1 บทนำ	19
3.2 กระบวนการเรียนรู้	20
3.3 กระบวนการทดสอบและการนำไปใช้งาน	26
3.4 การวิเคราะห์เปรียบเทียบความซับซ้อนในการคำนวณและประสิทธิภาพระหว่าง วิธีการฮาอูสเตอร์ฟอาร์ทแมพและวิธีการซิมพลิไฟด์ฟิชชี่อาร์ทแมพ	27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
3.4.1 ขนาดอินพุท	27
3.4.2 การคำนวณชอยซ์ฟังก์ชันและแมทซ์ฟังก์ชัน	27
3.4.3 ความซับซ้อนในการคำนวณหาค่าฮาลูสตอร์ฟติดแทนซ์กับการคำนวณ โดยใช้พีซี and โอเปอร์เรชัน	27
3.4.4 เปรียบเทียบประสิทธิภาพของตัววัดค่าความเหมือนของฮาลูสตอร์ฟ อาร์ทแมพกับซิมพลิไฟด์พีซีอาร์ทแมพ	30
บทที่ 4 การทดลองและผลการทดลอง	31
4.1 การรู้จำลายนิ้วมือด้วยฮาลูสตอร์ฟอาร์ทแมพ	32
4.1.1 กระบวนการเตรียมข้อมูล	32
4.1.1.1 การหาจุดอ้างอิง	32
4.1.1.2 การขจัดจุดซิงกูลาร์ที่ไม่ถูกต้องบริเวณพื้นหลัง	39
4.1.1.3 การแปลงภาพลายนิ้วมือแบบระดับสีเทาไปเป็นภาพแบบไบนารี ...	39
4.1.1.4 การเลือกรูปลายนิ้วมือที่เหมาะสมและปรับขนาดของรูป	41
4.1.1.5 การทำ thinning	43
4.1.2 กระบวนการเรียนรู้ภาพลายนิ้วมือ	44
4.1.3 กระบวนการทดสอบและผลลัพธ์ในการรู้จำลายนิ้วมือ	45
4.2 การรู้จำรูปร่างของมือ	46
4.2.1 การเตรียมข้อมูลภาพมือ	48
4.2.1.1 การแปลงจากภาพมือที่เป็นระดับสีเทาไปเป็นภาพแบบไบนารี	48
4.2.1.2 การหาเส้นขอบของฝ่ามือ	49
4.2.2 กระบวนการเรียนรู้และการทดสอบการรู้จำ	49
4.3 การรู้จำภาพใบหน้า	50
4.3.1 การเตรียมข้อมูลภาพใบหน้า	51
4.3.1.1 การหาจุดอ้างอิง	51
4.3.1.2 การหาเส้นขอบและการแปลงภาพไปเป็นภาพแบบไบนารี	51
4.3.1.3 การหมุนภาพ	52
4.3.1.4 การปรับขนาดภาพ	53

สารบัญ (ต่อ)

	หน้า
4.3.1.5 การเลื่อนตำแหน่งอ้างอิงของภาพและเลือกบริเวณของภาพ ในส่วนที่ต้องการ	53
4.3.2 กระบวนการเรียนรู้และการทดสอบผลการรู้จำ	53
4.3.3 การรู้จำภาพใบหน้าจากฐานข้อมูล ORL	54
4.3.4 การรู้จำภาพใบหน้าจากฐานข้อมูล MIT	58
4.3.5 การรู้จำภาพใบหน้าจากฐานข้อมูล Yale	61
4.4 การวิเคราะห์ถึงผลกระทบของค่า S ต่อประสิทธิภาพในการรู้จำของฮาลูสตอร์ฟ อาร์ทแมพ	64
4.5 การวิเคราะห์เปรียบเทียบการรู้จำใบหน้าด้วยวิธีฮาลูสตอร์ฟอาร์ทแมพกับวิธี สเปเชียลลีไอเกนเวทเต็ดดูบเบิลฮาลูสตอร์ฟดิสแทนซ์	67
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	70
เอกสารอ้างอิง	74
ภาคผนวก	77
ประวัติผู้เขียน	84

สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงค่าความเหมือนที่ได้จากตัววัดค่าความเหมือนชนิดต่างๆ เมื่อเปรียบเทียบ โดยใช้เลข 7 แบบต่างๆ ในรูปที่ 2.3	13
4.1 คุณสมบัติของวิธีการระบุบุคคลด้วยใบหน้า ลายนิ้วมือ และมือ	31
4.2 แสดงภาพลายนิ้วมือจากฐานข้อมูล FINGDB ที่ใช้ในการเรียนรู้ และภาพที่ใช้ทดสอบ	42
4.3 แสดงจำนวนรูปภาพในฐานข้อมูล FINGDB และจำนวนรูปภาพที่นำมาใช้ในการ ทดลอง	43
4.4 แสดงผลลัพธ์ในการรู้จำลายนิ้วมือโดย $p = 0.80$, ค่า $S = 2$, การเปลี่ยนตำแหน่งจะ ครอบคลุมบริเวณขยายออกจากจุดศูนย์กลางเดิมออกไปข้างละ 6 พิกเซล โดยทำ การย้ายจุดศูนย์กลางของรูปแบบอินพุตไปที่ละ 3 พิกเซล	46
4.5 แสดงจำนวนภาพมือที่ใช้ในการทดลอง	47
4.6 ผลการรู้จำภาพมือ โดยกำหนดให้ $p = 1.00$, $S = 4$	50
4.7 แสดงภาพที่ใช้ในการเรียนรู้ภาพใบหน้าจากฐานข้อมูล ORL	55
4.8 ผลลัพธ์การรู้จำภาพใบหน้า ORL (โดย $p = 0.80$, $S = 2$, และให้แคทากอรีไหนด จำนวน 153 ไหนด)	57
4.9 แสดงจำนวนภาพใบหน้าจากฐานข้อมูล MIT ที่ใช้ในการทดลอง	60
4.10 ผลลัพธ์การรู้จำภาพใบหน้า MIT (โดย $p = 0.60$, $S = 2$, และให้แคทากอรีไหนดจำนวน 223 ไหนด)	61
4.11 ผลลัพธ์ในการรู้จำภาพใบหน้าจากฐานข้อมูล Yale (โดย $p = 0.80$, $S = 2$, และให้ แคทากอรีไหนดจำนวน 39 ไหนด)	63
4.12 ภาพที่ใช้ในการเรียนรู้ของระบบที่ได้จากการสุ่มและให้ผลลัพธ์ในตารางที่ 4.11	64
4.13 ผลลัพธ์ในการรู้จำภาพใบหน้าจากฐานข้อมูล ORL และ MIT ด้วยวิธีการสเปเชียลลี ไอเกนเวทเต็ดดูบลิฮาสตอร์ฟิดิสแทนซ์และวิธีการฮาาสตอร์ฟอาร์ทแมพ	68

สารบัญรูป

รูปที่	หน้า
2.1 แผนภาพการทำงานของระบบการมองเห็นของมนุษย์	5
2.2 สถาปัตยกรรมของนีโอค็อกนิตรอนนิวรอลเน็ตเวิร์ค	6
2.3 เลข 7 ในรูปแบบต่างๆ ที่ใช้ในการเปรียบเทียบประสิทธิภาพของตัววัดค่าความเหมือน	9
2.4 สถาปัตยกรรมของซิมพลิไฟด์พีชชีอาร์ทแมพ	15
3.1 สถาปัตยกรรมของฮาสตูสเตอร์ฟอาร์ทแมพ	20
3.2 ตัวอย่างการปรับค่าน้ำหนักจากค่าน้ำหนักเดิม w^{old} ด้วยรูปแบบอินพุท I^m ไปเป็นค่า น้ำหนักใหม่ w^{new} เมื่อ $S = 1$	23
3.3 แสดงผลลัพธ์จากการดำเนินการด้วยวิธียุคลิเดียนดิสแทนซ์ทรานส์ฟอร์ม.....	29
4.1 ภาพลายนิ้วมือแบบโค้ง	33
4.2 ภาพลายนิ้วมือแบบโค้งกระโຈມ	33
4.3 ภาพลายนิ้วมือแบบมัดหวายมัดซ้าย	33
4.4 ภาพลายนิ้วมือแบบมัดหวายมัดขวา	34
4.5 ภาพลายนิ้วมือแบบก้นหอย	34
4.6 แสดงลักษณะสำคัญพิเศษบนลายนิ้วมือ	34
4.7 ตัวอย่างรูปภาพลายนิ้วมือจากฐานข้อมูล FINGDB	35
4.8 (a) ภาพลายนิ้วมือและตำแหน่งของจุดซิงกูลาร์ที่ยังคงมีจุดซิงกูลาร์ที่ผิดปกติอยู่บริเวณ พื้นหลัง (b) ภาพลายนิ้วมือที่ขจัดจุดซิงกูลาร์ที่ผิดปกติออกไปแล้ว	39
4.9 (a) ภาพลายนิ้วมือแบบไบนารีที่ได้จากวิธีการ RAT (b) ภาพลายนิ้วมือแบบไบนารีที่ได้จาก วิธีการ GT	40
4.10 ตัวอย่างภาพลายนิ้วมือที่คัดออกเนื่องจากไม่สามารถหาจุดซิงกูลาร์ที่แน่นอนได้	41
4.11 ตัวอย่างภาพลายนิ้วมือที่คัดออกเนื่องจากพื้นที่รอบจุดอ้างอิงมีขนาดเล็กกว่า 100x100 พิกเซล	41
4.12 ภาพลายนิ้วมือแบบไบนารีที่ผ่านกระบวนการ thinning แล้ว	44
4.13 แสดงตำแหน่งของจุดศูนย์กลางของภาพ (จุดสีดำ) และจุดศูนย์กลางใหม่ที่เปลี่ยน ตำแหน่งไปในการคำนวณแต่ละครั้ง (จุดสีเทา) เมื่อขยายจุดที่จะพิจารณาออกไปข้าง ละ 6 พิกเซลและทำการย้ายตำแหน่งไปที่ละ 3 พิกเซล โดยกรอบเส้นประสี่เหลี่ยมขนาด 80x80 พิกเซล คือ ส่วนของรูปแบบอินพุทที่จะเลื่อนตำแหน่งของจุดศูนย์กลางด้วย	

สารบัญญรูป (ต่อ)

รูปที่	หน้า
ตำแหน่งต่างๆ เพื่อที่จะทำการพิจารณาเปรียบเทียบกับค่าน้ำหนัก w'	44
4.14 แสดงการวัดค่าพีเจอร์ต่างๆ ที่ใช้ในการรู้จำภาพมือโดยทั่วไป	47
4.15 ตัวอย่างภาพมือที่ใช้ในการทดลอง	48
4.16 ภาพมือแบบไบนารีที่ได้จากการใช้วิธี GT	48
4.17 ภาพมือที่ได้รับการสกัดเส้นขอบ	49
4.18 แสดงตำแหน่งอ้างอิงของภาพใบหน้าโดยใช้ตำแหน่งของตาและระบุตำแหน่งด้วยมือ ...	51
4.19 (a) ภาพใบหน้าที่เป็นภาพแบบสีเทา (b) ภาพใบหน้าที่เป็นภาพหลายเส้นแบบระดับสีเทา โดยการใช้วิธี morphological gradient (c) ภาพแบบไบนารีโดยการใช้วิธีการ adaptive thresholding	52
4.20 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL, MIT, และ Yale แถวบนแสดงรูปภาพเดิม และแถวล่างแสดงรูปภาพแบบไบนารีที่ได้จากกระบวนการเตรียมข้อมูล	53
4.21 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL	54
4.22 ตัวอย่างภาพใบหน้าจากฐานข้อมูล MIT	59
4.23 ตัวอย่างภาพใบหน้าจากฐานข้อมูล Yale	62
4.24 แสดงผลความถูกต้องในการรู้จำลายนิ้วมือที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.80$ และ กำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟอาร์ทแมพเป็นลำดับที่ เหมือนกัน	65
4.25 แสดงผลความถูกต้องในการรู้จำภาพมือที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 1.00$ และ กำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟอาร์ทแมพเป็นลำดับที่ เหมือนกัน	65
4.26 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล ORL ที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.80$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟ อาร์ทแมพเป็นลำดับที่เหมือนกัน	66
4.27 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล MIT ที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.60$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟ อาร์ทแมพเป็นลำดับที่เหมือนกัน	66

สารบัญญรูป (ต่อ)

รูปที่

หน้า

4.28 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล Yale ที่สัมพันธ์กับค่า S ต่างๆ

เมื่อ $p = 0.80$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟ

อาร์ทแมพเป็นลำดับที่เหมือนกัน67



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การรู้จำรูปแบบที่เป็นรูปภาพโดยใช้โครงข่ายประสาทเทียมส่วนใหญ่ในปัจจุบันมักจะต้องทำการดึงค่าของลักษณะสำคัญของภาพออกมาก่อนซึ่งค่านี้มักจะเป็นค่าตัวเลขที่ได้มาจากการบวนการคำนวณที่ซับซ้อนและเข้าใจได้ยาก จึงน่าจะเป็นการดีกว่าถ้าสามารถที่จะดำเนินการรู้จำรูปแบบได้ด้วยวิธีการที่ง่ายและรวดเร็ว ซึ่งจากวิธีการในการมองเห็นและรู้จำวัตถุของมนุษย์นั้นสามารถที่จะจำแนกความแตกต่างของวัตถุได้จากการมองเพียงภาพรวมและโครงสร้างคร่าวๆ ของวัตถุนั้น เช่น เส้นขอบ รูปร่าง เป็นต้น ดังนั้นจึงจะนำเอาแนวคิดนี้มาใช้โดยการให้ข้อมูลสำหรับการเรียนรู้ของโครงข่ายประสาทเทียมมีลักษณะเป็นข้อมูลโครงร่างและเส้นขอบของรูปภาพที่เป็นสองมิติ เพราะข้อมูลโครงร่างและรูปร่างเหล่านี้จะเป็นข้อมูลที่สำคัญที่สุดในการรู้จำ เช่นเดียวกับการรู้จำของมนุษย์ แต่อย่างไรก็ดีในการสร้างโครงข่ายประสาทเทียมนั้นจำเป็นจะต้องเลือกใช้ตัววัดค่าความเหมือนของรูปแบบที่มีประสิทธิภาพใกล้เคียงกับมนุษย์ด้วยจึงจะทำให้ได้ประสิทธิภาพในการรู้จำที่ดี และวิธีการของโครงข่ายประสาทเทียมควรจะมีความสามารถเหมือนมนุษย์ในอันที่จะรู้จำสิ่งใหม่ได้โดยยังคงเก็บรักษาความรู้เก่าที่เคยเรียนรู้มาก่อนแล้วไว้ได้พร้อมกันด้วย อันจะทำให้สามารถแก้ไขปัญหาการรู้จำรูปแบบที่มีชนิดเดียวกันแต่มีหลายลักษณะได้ ดังนั้นการพัฒนาระบบการรู้จำที่รวมเอาคุณสมบัติและแนวคิดต่างๆ ดังกล่าวข้างต้นมาใช้ก็น่าจะทำให้ได้ระบบการรู้จำแบบใหม่ที่ง่ายและมีประสิทธิภาพใกล้เคียงกับมนุษย์มากยิ่งขึ้น

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1. เพื่อพัฒนาระบบการรู้จำรูปแบบซึ่งใช้โครงข่ายประสาทเทียมที่อาศัยแนวคิดจากวิธีการมองเห็นและรู้จำรูปแบบของมนุษย์เป็นพื้นฐาน โดยใช้ข้อมูลในการรู้จำเป็นอาร์เรย์ของรูปภาพ 2 มิติ
2. เพื่อศึกษาหาตัววัดความเหมือนของรูปแบบที่มีกระบวนการและประสิทธิภาพที่คล้ายคลึงกับมนุษย์มากที่สุด และหาวิธีการปรับใช้กับระบบการรู้จำ
3. เพื่อพัฒนาให้ได้ระบบการรู้จำที่ง่ายแก่การเข้าใจ ดำเนินการได้อย่างรวดเร็ว และให้ประสิทธิภาพในการรู้จำใกล้เคียงกับมนุษย์

1.3 สมมติฐานของการศึกษา

จากการศึกษาวิธีการมองเห็นและรู้จำรูปแบบของมนุษย์จะเห็นได้ว่าเมื่อมนุษย์รับภาพเข้ามาจาก 3 มิติจะแปลงมาอยู่ในรูปของข้อมูลที่เป็น 2 มิติ จากนั้นกระบวนการที่สำคัญในอันที่จะทำให้มนุษย์สามารถแยกแยะและรู้จำวัตถุได้อย่างรวดเร็วก็คือการสกัดเอาข้อมูลที่เป็นลักษณะโดยรวมแบบคร่าวๆ เช่น โครงร่าง รูปร่าง เส้นขอบ และลายเส้นออกมาจากภาพภาพนั้น ดังนั้นในกระบวนการรู้จำถ้านำเอาข้อมูลที่เป็นอาร์เรย์ 2 มิติของรูปร่างที่มีลักษณะเหล่านี้มาใช้ก็น่าจะทำให้การรู้จำรูปแบบได้ดีคล้ายคลึงกับมนุษย์และมีความรวดเร็วเข้าใจได้ง่าย เมื่อได้รูปแบบที่ใช้เป็นข้อมูลในการรู้จำแล้วอีกสิ่งหนึ่งก็คือตัววัดค่าความเหมือนซึ่งถ้าเลือกตัววัดค่าความเหมือนที่ได้รับการยอมรับว่ามีวิธีการและประสิทธิภาพใกล้เคียงกับมนุษย์ก็น่าจะทำให้กระบวนการรู้จำมีประสิทธิภาพมากยิ่งขึ้นด้วย อีกทั้งโครงสร้างของกระบวนการรู้จำรูปแบบควรต้องมีคุณสมบัติในการรู้จำ คือ สามารถเรียนรู้สิ่งใหม่ได้พร้อมทั้งยังคงความรู้เก่าไว้ได้อันเป็นคุณสมบัติที่มีเหมือนกับมนุษย์ก็ยิ่งน่าจะทำให้กระบวนการรู้จำนี้มีประสิทธิภาพดีคล้ายมนุษย์มากขึ้นไปอีก และท้ายที่สุดการปรับเอาแนวคิดทั้ง 3 ส่วนมาใช้ร่วมกันในการสร้างกระบวนการรู้จำก็น่าจะทำให้ได้กระบวนการรู้จำแบบใหม่ที่ดีและมีประสิทธิภาพคล้ายกับมนุษย์ตามที่ต้องการ

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

ในการวิจัยนี้ได้นำเอาแนวคิดในการมองเห็นและรู้จำของมนุษย์มาเป็นหลักการเบื้องต้นของแนวคิดในการพัฒนากระบวนการเรียนรู้ทั้งหมด และพยายามที่จะสร้างกระบวนการรู้จำใหม่ที่เลียนแบบมนุษย์และมีประสิทธิภาพใกล้เคียงกับมนุษย์ โดยกำหนดให้ข้อมูลที่ใช้ในการรู้จำของระบบเป็นข้อมูลภาพ 2 มิติที่ดึงเอาลักษณะเด่นของรูปร่างของวัตถุออกมาเพราะจากการศึกษาวิธีการมองเห็นและรู้จำวัตถุของมนุษย์พบว่ากระบวนการรู้จำที่สำคัญในอันที่จะทำให้มนุษย์ทำการรู้จำได้รวดเร็วคือ การสกัดเส้น ขอบ และรูปร่าง ออกจากวัตถุ แต่อย่างไรก็ตามในการรู้จำรูปภาพลายเส้นหรือภาพโครงร่างนี้มีความจำเป็นอย่างยิ่งที่จะต้องใช้วิธีการวัดค่าความเหมือนที่มีประสิทธิภาพสูงด้วยเช่นกัน ในงานวิจัยนี้จึงนำเอาวิธีการวัดค่าความเหมือนแบบฮาอูสดอร์ฟดิสแทนซ์ (Hausdorff distance) ที่มีประสิทธิภาพและได้รับการยอมรับว่ามีความใกล้เคียงกับวิธีการรู้จำของมนุษย์มากที่สุดมาเป็นตัววัดค่าความเหมือนของรูปแบบ และนำวิธีการวัดค่าความเหมือนด้วยฮาอูสดอร์ฟดิสแทนซ์มาปรับใช้ร่วมกับแนวทาง Adaptive Resonance Theory (ART) หรือที่เรียกว่าอาร์ทซึ่งมีคุณสมบัติในการรู้จำ คือ สามารถเรียนรู้สิ่งใหม่ได้พร้อมกับยังคงความรู้เก่าไว้ได้อันเป็นคุณสมบัติในการรู้จำของมนุษย์และแก้ปัญหาการจดจำรูปแบบชนิดเดียวแต่มีหลายลักษณะได้ ดังนั้นโดยสรุปงานวิจัยนี้จึงมุ่งที่จะศึกษาการใช้แนวทางอาร์ทเป็นพื้นฐานของ

กระบวนการเรียนรู้และพยายามหาวิธีการปรับเอาฮาอูสเตอร์ฟิตสแทนซ์มาใช้ร่วมกับแนวทางอาร์ท ให้ได้เพื่อรู้จำรูปภาพแบบ 2 มิติ

1.5 ขอบเขตของการวิจัย

งานวิจัยนี้เน้นการพัฒนากระบวนการรู้จำสำหรับรูปแบบที่เป็นรูปภาพซึ่งมีโครงสร้างที่ชัดเจนเท่านั้น เช่น ภาพลายนิ้วมือ ใบหน้า เป็นต้น โดยจะไม่ทำการรู้จำรูปแบบที่มีลักษณะอื่นๆ เช่น รูปภาพที่มีความซับซ้อน (เช่น ภาพทิวทัศน์) เป็นต้น และกระบวนการรู้จำที่พัฒนาขึ้นมาจะทำการรู้จำรูปแบบที่เป็นรูปภาพแบบไบนารีเท่านั้น

1.6 ขั้นตอนของการศึกษา

1. ศึกษากระบวนการมองเห็นและรู้จำของมนุษย์
2. ศึกษากระบวนการสกัดลายเส้น และโครงสร้างที่สำคัญออกจากรูปภาพ รวมถึงการปรับปรุงรูปภาพ
3. ศึกษาวิธีการวัดความเหมือนแบบฮาอูสเตอร์ฟิตสแทนซ์และการปรับใช้งาน
4. ศึกษาวิธีการของ Adaptive Resonance Theory แบบต่างๆ และเลือกวิธีการที่เหมาะสม
5. ศึกษาและทดลองปรับแนวคิดทั้งหมดมาใช้ร่วมกันด้วยวิธีการต่างๆ
6. วิเคราะห์และสรุปผลเพื่อเลือกหาแนวทางในการสร้างรูปแบบการเรียนรู้ที่ประยุกต์เอาแนวคิดทั้งหมดมาใช้ร่วมกัน
7. สร้างโครงสร้างและกระบวนการรู้จำแบบใหม่ตามแนวทางที่ได้
8. ทดลองและทดสอบกระบวนการรู้จำแบบใหม่กับข้อมูลรูปแบบต่างๆ ได้แก่ ลายนิ้วมือ ภาพของมือ และภาพใบหน้านามมนุษย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการและทฤษฎีเบื้องต้นที่นำไปสู่แนวคิดของการวิจัย

2.1 แนวคิดในการมองเห็นและรู้จำวัตถุของมนุษย์

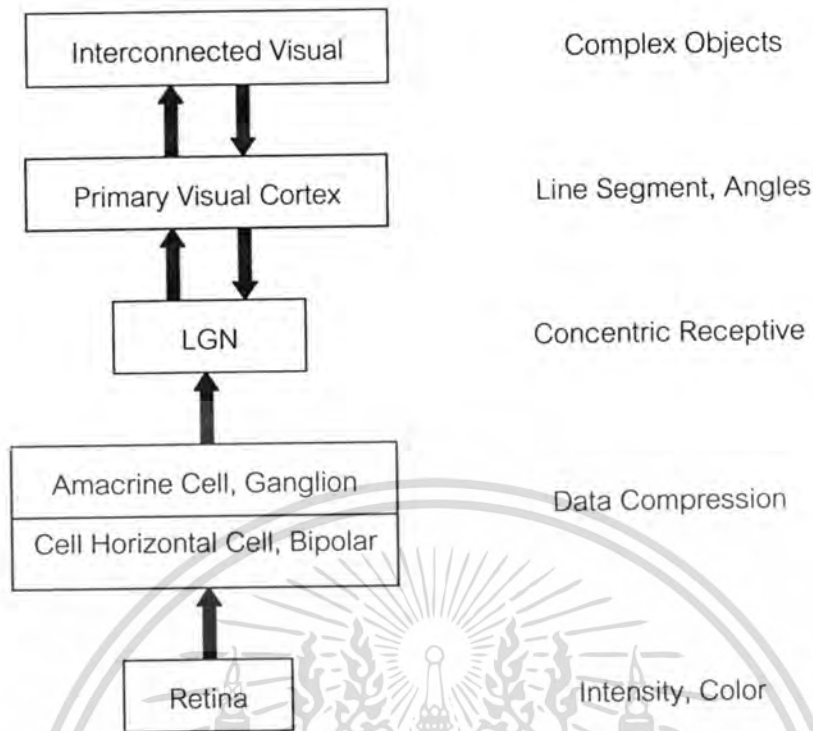
ในระบบการมองเห็นของมนุษย์นั้นจะดำเนินการเป็นลำดับขั้นตอน คือ เรตินา (retina) จะทำหน้าที่เป็นตัวรับภาพโดยมีอวัยวะของตัวรับภาพคือเซลล์รับภาพที่ประกอบไปด้วยเซลล์ 2 ชนิดคือ โคนเซลล์ (cone cell) และรีดเซลล์ (rod cell) โคนเซลล์จะทำหน้าที่ตอบสนองต่อสีโดยจะมีเซลล์ที่รับสี 3 ชนิดคือ เซลล์ที่ไวต่อแสงสีแดง สีเขียว และสีน้ำเงิน ส่วนรีดเซลล์จะเป็นเซลล์ที่ทำหน้าที่รับความเข้มของแสง [1]

จากนั้นเมื่อโคนเซลล์และรีดเซลล์รับข้อมูลการเห็นภาพเข้ามาแล้วจะยังไม่ส่งข้อมูลนี้ไปยังสมองโดยตรง แต่สัญญาณนี้จะถูกส่งจากเซลล์รับข้อมูลไปยังโครงข่ายซึ่งประกอบไปด้วยเซลล์ 4 ชนิด คือ ฮอริซันเทิลเซลล์ (horizontal cell) ไบโพลาร์เซลล์ (bipolar cell) อามาครินเซลล์ (amacrine cell) และแกงเกลียเซลล์ (ganglion cell) เซลล์เหล่านี้มีการเรียงตัวที่เชื่อมต่อกันเป็นลักษณะโครงข่าย (มีลำดับชั้นโยงใยถึงกัน) เพื่อที่จะทำหน้าที่รวมกลุ่มของสัญญาณที่รับมาจากเซลล์รับภาพของเรตินาแล้วส่งต่อไปยังสมองอีกทีหนึ่ง โดยแกงเกลียเซลล์จะทำหน้าที่ส่งสัญญาณที่ได้รับการบีบอัดในโครงข่ายนี้ไปยังประสาทรับภาพ จากขั้นตอนนี้จะเห็นได้ว่าสัญญาณภาพจะได้รับการบีบอัดให้มีขนาดเล็กลงเรื่อยๆ ในโครงข่ายของเซลล์ 4 ชนิดนี้ จากข้อมูลของเซลล์รับภาพที่เรตินาจำนวน 127 ล้านเซลล์จะได้รับการส่งออกไปยังสมองโดยแกงเกลียเซลล์เพียง 1 ล้านเซลล์เท่านั้น

ในส่วนต่อมาเมื่อเอาที่พุดจากแกงเกลียเซลล์มายังประสาทรับภาพจะมีลาเทอร์จีเจนิคูลาเทนิวคลี (lateral geniculate nuclei : LGN) ของทาลามัส (thalamus) ซึ่งทำหน้าที่เป็นตัวเชื่อมติดต่อกันระหว่างแกงเกลียเซลล์และวิซวลคอเทซ (visual cortex) ส่วนแรกของสมอง ที่วิซวลคอเทซ ส่วนแรกของสมองจะถูกจัดเรียงแบ่งออกเป็นชั้น แต่ละชั้นประกอบไปด้วยซิมเพิลเซลล์ (simple cell) และคอมเพลกซ์เซลล์ (complex cell) ซิมเพิลเซลล์จะตอบสนองต่อพีเจอร์แบบจีโอเมทริก (geometric feature) ได้แก่ ขอบ เส้น และมุมในตำแหน่งและทิศทางเชิงมุมที่เฉพาะเจาะจงในส่วนของ LGN ที่เป็นเหมือนส่วนการมองเห็นที่ละเอียดขึ้น คอมเพลกซ์เซลล์ในชั้นเดียวกันนี้จะตอบสนองกับพีเจอร์เช่นเดียวกับซิมเพิลเซลล์แต่ได้ในตำแหน่งที่หลากหลาย ส่วนต่อมาของวิซวลคอเทซเป็นส่วนที่มีการติดต่อเชื่อมโยงกันเพื่อสกัดพีเจอร์ที่ซับซ้อนขึ้นอันได้แก่ สี ทิศทางเชิงมุม การเคลื่อนไหว และขนาด

ขั้นตอนในการมองเห็นของมนุษย์สามารถที่จะสรุปได้ดังรูปต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 แผนภาพการทำงานของระบบการมองเห็นของมนุษย์

ในการรู้จำวัตถุของมนุษย์และวิธีการที่มนุษย์สามารถแยกความแตกต่างของวัตถุนั้นๆ ได้ก็เนื่องมาจากเมื่อมนุษย์มองภาพวัตถุที่เป็นภาพ 3 มิติ อาร์เรย์ 2 มิติของความยาวคลื่นหรือก็คือสีแต่ละสีและความเข้มของแสงจะถูกโปรเจกต์ลงบนเรตินา ซึ่งจากวิธีการและขั้นตอนการมองเห็นของมนุษย์ที่กล่าวมาแล้วข้างต้น จะเห็นได้ว่าขั้นตอนที่สำคัญในการดำเนินการรู้จำรูปแบบในเบื้องต้นนี้ก็คือการสกัดเส้นขอบจากอาร์เรย์ของเรตินานั้นเอง โดยเส้นขอบก็คือส่วนที่มีการเปลี่ยนแปลงของสีหรือความเข้มแสงในภาพ นอกจากนี้แล้วมนุษย์ยังมีความสามารถที่จะจดจำวัตถุที่อยู่ในหลายรูปแบบได้เช่น รูปแบบที่มีขนาดแตกต่างกัน วางตัวในทิศทางต่างๆ และอยู่ในตำแหน่งที่ไม่เหมือนกันได้ จากแนวคิดเหล่านี้เองได้มีการนำมาปรับใช้เพื่อพัฒนาวิธีการรู้จำวัตถุอย่างกว้างขวาง

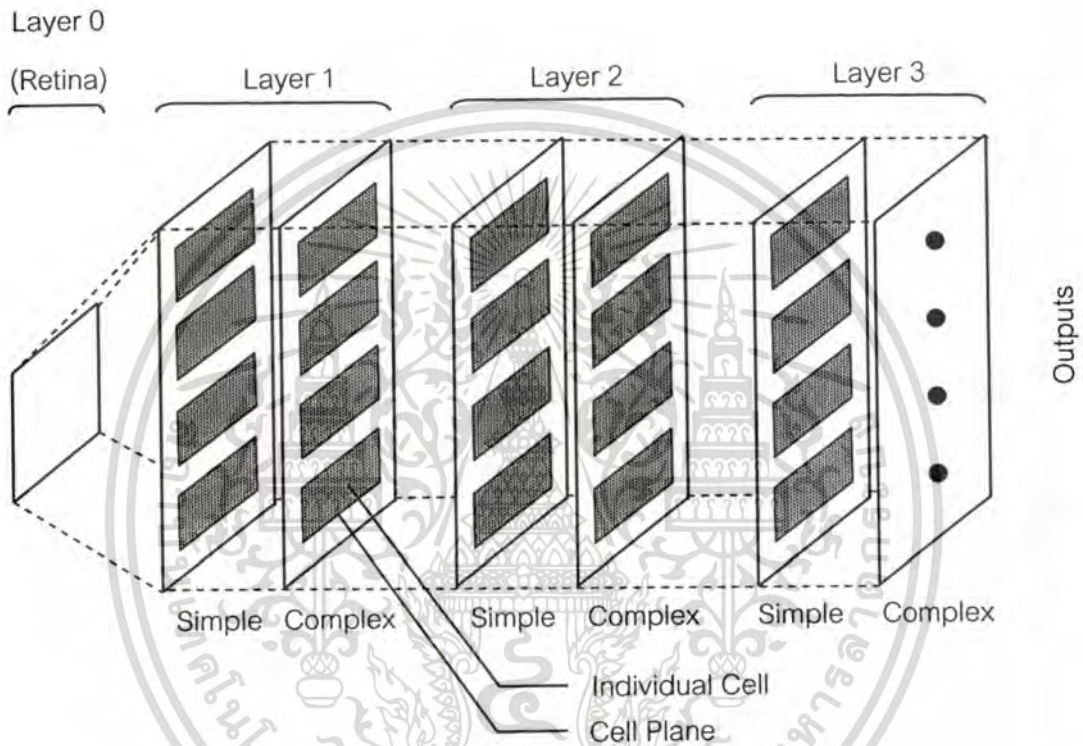
2.2 นีโอค็อกนิตรอนนิวโรลเน็ตเวิร์ค

นีโอค็อกนิตรอนนิวโรลเน็ตเวิร์ค (neocognitron neural network) เป็นการผสมผสานระหว่างโครงข่ายประสาท (neural network) และรูปแบบการทำงานของระบบการมองเห็นของมนุษย์ นีโอค็อกนิตรอนได้รับการพัฒนาจากพื้นฐานของรูปแบบของระบบการมองเห็นและวิธีการรู้จำวัตถุของมนุษย์ดังได้กล่าวมาแล้วในหัวข้อ 2.1 อินพุทของการรู้จำด้วยวิธีการนีโอค็อก-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นิตรอนนี้จะเป็นรูปภาพแบบสองมิติ (two-dimensional images) ซึ่งเลียนแบบมาจากอาร์เรย์ของเรตินาในการรับภาพ

โครงสร้างสถาปัตยกรรมของนีโอค็อกนิตรอนจะเป็นแบบไฮราคี (hierarchical organization) คือในระดับชั้นที่อยู่ต่ำกว่าจะดำเนินการที่มีความซับซ้อนน้อยกว่า และในระดับชั้นที่สูงขึ้นไปจะมีการทำงานที่ซับซ้อนมากขึ้นและเข้าใจได้ยากยิ่งขึ้น สถาปัตยกรรมของนีโอค็อกนิตรอนสามารถแสดงได้ดังรูปที่ 2.2



รูปที่ 2.2 สถาปัตยกรรมของนีโอค็อกนิตรอนนิวรอลเน็ตเวิร์ค

แต่ละชั้นในโครงสร้างของนีโอค็อกนิตรอนจะประกอบไปด้วยกลุ่มเซลล์ (cell) ซึ่งแบ่งออกเป็น 2 ประเภท คือ ซิมเพิลเซลล์ (simple cell) และคอมเพล็กซ์เซลล์ (complex cell) โดยซิมเพิลเซลล์จะทำหน้าที่รับอินพุตจากชั้นที่อยู่ก่อนหน้าและซิมเพิลเซลล์ในระดับชั้นแรกสุดจะทำหน้าที่รับอินพุตจากชั้นอินพุตหรือเรียกว่าเรตินา แต่ละซิมเพิลเซลล์จะตอบสนองต่อพีเจอร์ (feature) ที่แน่นอนในตำแหน่งที่ได้รับการระบุไว้ นอกจากนี้แล้วซิมเพิลเซลล์ยังถูกแบ่งออกเป็นเพลน (plane) ซึ่งเซลล์ในแต่ละเพลนนี้จะตอบสนองต่อพีเจอร์แบบเดียวกันแต่อยู่คนละตำแหน่งของเรตินา และทุกซิมเพิลเซลล์ในเพลนเดียวกันจะมีน้ำหนักอินพุตเหมือนกัน

คอมเพล็กซ์เซลล์ในชั้นหนึ่งๆ จะทำหน้าที่รับอินพุตจากซิมเพิลเซลล์ของระดับชั้นนั้น จากนั้นคอมเพล็กซ์เซลล์จะทำการรวมกลุ่มพีเจอร์ที่ได้รับการรู้จำจากซิมเพิลเซลล์เข้าด้วยกัน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นจะเห็นว่าคอมเพลกซ์เซลล์นั้นจะทำการตอบสนองต่อรูปแบบที่ซับซ้อนยิ่งขึ้น การที่คอมเพลกซ์เซลล์รับกลุ่มรูปแบบอินพุทจากซิมเพิลเซลล์ในชั้นเดียวกันนั้นทำให้ตำแหน่งที่ได้รับการตอบสนองของคอมเพลกซ์เซลล์น้อยลงกว่าซิมเพิลเซลล์ คอมเพลกซ์เซลล์ก็ได้รับการแบ่งออกเป็นเพลนเช่นเดียวกันกับซิมเพิลเซลล์ และในแต่ละเพลนก็จะประกอบไปด้วยเซลล์ซึ่งจะตอบสนองต่อวัตถุที่มีความคล้ายคลึงกัน ค่าน้ำหนักของคอมเพลกซ์เซลล์จะเป็นค่าที่กำหนดตายตัว (ไม่มีการเปลี่ยนแปลงในช่วงการเรียนรู้) และยังในระดับชั้นที่อยู่ห่างออกไปจากเรตินา การตอบสนองของคอมเพลกซ์เซลล์ก็จะมีค่าซับซ้อนมากยิ่งขึ้น

นี่คือคอนิตรอนให้ผลลัพธ์ในการจำที่ดี และข้อดีของนีโอค็อกนิตรอนคือความสามารถในการจำรูปแบบซึ่งอาศัยพีเจอร์ที่ถูกสกัดออกมาจึงไม่ขึ้นกับขนาดและตำแหน่งของรูปแบบ อีกทั้งยังสามารถที่จะเรียนรู้ได้ทั้งแบบ supervised และแบบ unsupervised แต่ก็มีข้อเสียคือมีการคำนวณที่ซับซ้อน และระบบการจำที่ใช้นีโอค็อกนิตรอนก็ยังมีขนาดใหญ่อีกด้วยเนื่องจากต้องใช้จำนวนนิวรอนมากและในแต่ละนิวรอนก็มีค่าน้ำหนักอีกเป็นจำนวนมากเช่นกัน ซึ่งทำให้การประยุกต์ไปใช้งานจริงทำได้ยาก

2.3 ฮาอูสตอร์ฟดิสแทนซ์

ฮาอูสตอร์ฟดิสแทนซ์ (Hausdorff distance : HD) ได้รับการยอมรับว่าเป็นตัววัดค่าความเหมือน (similarity metric) ที่มีประสิทธิภาพและมีหลักการทำงานใกล้เคียงกับวิธีการของมนุษย์มากกว่าตัววัดค่าความเหมือนอื่นๆ ที่ใช้กับโครงข่ายประสาทเทียมในปัจจุบันนี้ [1] ซึ่ง Rosandish [1] และ Xuejing Wu [2] ได้ทำการเปรียบเทียบให้เห็นความสามารถดังกล่าวของฮาอูสตอร์ฟดิสแทนซ์เมื่อเทียบกับตัววัดค่าความเหมือนอื่นๆ ฮาอูสตอร์ฟดิสแทนซ์เป็นการวัดระยะทางของแต่ละจุดในเซตของจุดอินพุท (input point set) ว่าอยู่ใกล้กับจุดบางจุดในเซตของจุดที่เป็นโมเดล (model point set) เป็นระยะทางเท่าใด กำหนดให้ $A = \{a_1, a_2, \dots, a_p\}$ และ $B = \{b_1, b_2, \dots, b_q\}$ เป็นเซต 2 เซตของจุดที่มีจำนวนสมาชิกจำกัด (finite point sets) ฮาอูสตอร์ฟดิสแทนซ์สามารถนิยามได้ดังนี้

$$H(A, B) = \max \{h(A, B), h(B, A)\} \quad (2.1)$$

เมื่อฟังก์ชัน $h(A, B)$ คำนวณมาจากไดเรกเตดฮาอูสตอร์ฟดิสแทนซ์ (directed Hausdorff distance) จาก A ไปยัง B ดังนี้

$$h(A, B) = \max_{a \in A} \{ \min_{b \in B} \{ \|a - b\| \} \} \quad (2.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $\|a - b\|$ ในที่นี้คือยูคลิดีเนียนดิสแทนซ์ (Euclidean distance) ระหว่างจุด a และ b ฮาอูสตอร์ฟดิสแทนซ์นิยามได้ดังนี้ คือ การหาจุดใน A ที่อยู่ห่างจากจุดใน B ที่ใกล้จุดนั้นที่สุดเป็นระยะทางมากที่สุด แล้ววัดระยะทางนั้น ในการนำฮาอูสตอร์ฟดิสแทนซ์มาใช้กับโครงข่ายประสาทเทียมจำเป็นต้องมีการปรับวิธีการคำนวณหาโคเรคเต็ดฮาอูสตอร์ฟดิสแทนซ์ในสมการ (2.2) เสียใหม่เพื่อเพิ่มประสิทธิภาพของการวัดค่าความเหมือนและลดผลกระทบจาก noise โดยจะต้องทำการคำนวณค่าโคเรคเต็ดฮาอูสตอร์ฟดิสแทนซ์ของทุกๆ จุด ซึ่งการคำนวณทำได้ดังนี้

$$d(a, B) = \min_{b \in B} \{\|a - b\|\} \quad (2.3)$$

เมื่อ $d(a, B)$ คือ ฮาอูสตอร์ฟดิสแทนซ์แบบจุด (pointwise Hausdorff distance) ของจุดใดๆ ใน A และโคเรคเต็ดฮาอูสตอร์ฟดิสแทนซ์จาก A ไปยัง B คำนวณได้จากสมการต่อไปนี้

$$h(A, B) = \frac{\sum_{a \in A} d(a, B)}{|A|} \quad (2.4)$$

โดย $|A|$ คือ จำนวนจุดในเซต A

เพื่อแสดงให้เห็นความเหมาะสมในการเป็นตัววัดค่าความเหมือนของรูปภาพแบบไบนารี และประสิทธิภาพของฮาอูสตอร์ฟดิสแทนซ์ ดังนั้นจึงขอแสดงการเปรียบเทียบประสิทธิภาพของฮาอูสตอร์ฟดิสแทนซ์กับตัววัดค่าความเหมือนอีกสามชนิด คือ แฮมมิงดิสแทนซ์ (Hamming distance), ดอทโปรดักท์ (dot product), และยูคลิดีเนียนดิสแทนซ์ (Euclidean distance) โดยวิธีการคำนวณเพื่อวัดค่าความเหมือนของเวกเตอร์ที่เป็นแบบไบนารี $\{0, 1\}$ ของตัววัดค่าความเหมือนทั้งสามประเภทนี้สามารถนิยามได้ดังต่อไปนี้ คือ

- ตัววัดค่าความเหมือนแบบแฮมมิง

$$1 - \frac{d}{n} \quad (2.5)$$

- ตัววัดค่าความเหมือนแบบดอทโปรดักท์

$$\frac{A \cdot W}{|A|} \quad (2.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ตัววัดค่าความเหมือนแบบยูคลิดีเนียน

$$1 - \sqrt{\frac{d}{n}} \quad (2.7)$$

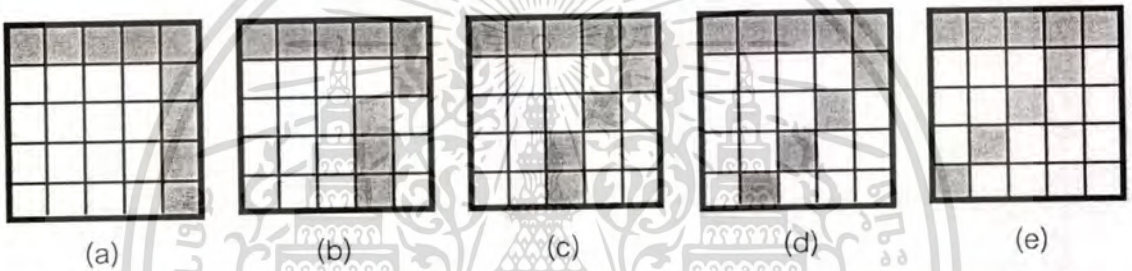
เมื่อ A คือ อินพุทเวกเตอร์

W คือ ค่าเวกเตอร์น้ำหนักของตัววัดค่าความเหมือนแต่ละตัว

d คือ จำนวนของอิลิเมนต์ (element) ที่มีค่าแตกต่างกันระหว่าง A และ W

n คือ จำนวนอิลิเมนต์ทั้งหมดใน A

$|A|$ คือ จำนวนของอิลิเมนต์ที่มีค่าเท่ากับ 1 ในเวกเตอร์ A



รูปที่ 2.3 เลข 7 ในรูปแบบต่างๆ ที่ใช้ในการเปรียบเทียบประสิทธิภาพของตัววัดค่าความเหมือน

ตัวอย่าง การคำนวณหาค่าฮามมิงคอร์ดที่ติดกันระหว่างเลข 7 แบบ (a) กับแบบ (b)

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
				a_{25}
				a_{35}
				a_{45}
				a_{55}

A

b_{11}	b_{12}	b_{13}	b_{14}	b_{15}
				b_{25}
			b_{34}	
			b_{44}	
			b_{54}	

B

ให้ เลข 7 แบบ (a) แทนด้วยเซตของจุด A

เลข 7 แบบ (b) แทนด้วยเซตของจุด B

$$A = \{a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{25}, a_{35}, a_{45}, a_{55}\}, |A| = 9$$

$$B = \{b_{11}, b_{12}, b_{13}, b_{14}, b_{15}, b_{25}, b_{34}, b_{44}, b_{54}\}, |B| = 9$$

ขั้นตอนการคำนวณมีดังนี้ คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) คำนวณหาค่าฮาอูลดอร์ฟติสแทนซ์แบบจุดของทุกๆ จุดใน A

$$d(a_{11}, B) = \min(\|a_{11} - b_{11}\|, \|a_{11} - b_{12}\|, \|a_{11} - b_{13}\|, \|a_{11} - b_{14}\|, \|a_{11} - b_{15}\|, \|a_{11} - b_{25}\|, \\ \|a_{11} - b_{34}\|, \|a_{11} - b_{44}\|, \|a_{11} - b_{54}\|)$$

$$= \min(0, 1, 2, 3, 4, 4.12, 3.61, 4.24, 5)$$

$$= 0$$

$$d(a_{12}, B) = \min(\|a_{12} - b_{11}\|, \|a_{12} - b_{12}\|, \|a_{12} - b_{13}\|, \|a_{12} - b_{14}\|, \|a_{12} - b_{15}\|, \|a_{12} - b_{25}\|, \\ \|a_{12} - b_{34}\|, \|a_{12} - b_{44}\|, \|a_{12} - b_{54}\|)$$

$$= \min(1, 0, 1, 2, 3, 3.16, 2.83, 3.61, 4.47)$$

$$= 0$$

$$d(a_{13}, B) = \min(\|a_{13} - b_{11}\|, \|a_{13} - b_{12}\|, \|a_{13} - b_{13}\|, \|a_{13} - b_{14}\|, \|a_{13} - b_{15}\|, \|a_{13} - b_{25}\|, \\ \|a_{13} - b_{34}\|, \|a_{13} - b_{44}\|, \|a_{13} - b_{54}\|)$$

$$= \min(2, 1, 0, 1, 2, 2.24, 2.24, 3.16, 4.12)$$

$$= 0$$

$$d(a_{14}, B) = \min(\|a_{14} - b_{11}\|, \|a_{14} - b_{12}\|, \|a_{14} - b_{13}\|, \|a_{14} - b_{14}\|, \|a_{14} - b_{15}\|, \|a_{14} - b_{25}\|, \\ \|a_{14} - b_{34}\|, \|a_{14} - b_{44}\|, \|a_{14} - b_{54}\|)$$

$$= \min(3, 2, 1, 0, 1, 1.41, 2, 3, 4)$$

$$= 0$$

$$d(a_{15}, B) = \min(\|a_{15} - b_{11}\|, \|a_{15} - b_{12}\|, \|a_{15} - b_{13}\|, \|a_{15} - b_{14}\|, \|a_{15} - b_{15}\|, \|a_{15} - b_{25}\|, \\ \|a_{15} - b_{34}\|, \|a_{15} - b_{44}\|, \|a_{15} - b_{54}\|)$$

$$= \min(4, 3, 2, 1, 0, 1, 2.24, 3.16, 4.12)$$

$$= 0$$

$$d(a_{25}, B) = \min(\|a_{25} - b_{11}\|, \|a_{25} - b_{12}\|, \|a_{25} - b_{13}\|, \|a_{25} - b_{14}\|, \|a_{25} - b_{15}\|, \|a_{25} - b_{25}\|, \\ \|a_{25} - b_{34}\|, \|a_{25} - b_{44}\|, \|a_{25} - b_{54}\|)$$

$$= \min(4.12, 3.16, 2.24, 1.41, 1, 0, 1.41, 2.24, 3.16)$$

$$= 0$$

$$d(a_{35}, B) = \min(\|a_{35} - b_{11}\|, \|a_{35} - b_{12}\|, \|a_{35} - b_{13}\|, \|a_{35} - b_{14}\|, \|a_{35} - b_{15}\|, \|a_{35} - b_{25}\|, \\ \|a_{35} - b_{34}\|, \|a_{35} - b_{44}\|, \|a_{35} - b_{54}\|)$$

$$= \min(4.47, 3.61, 2.83, 2.24, 2, 1, 1, 1.41, 2.24)$$

$$= 1$$

$$d(a_{45}, B) = \min(\|a_{45} - b_{11}\|, \|a_{45} - b_{12}\|, \|a_{45} - b_{13}\|, \|a_{45} - b_{14}\|, \|a_{45} - b_{15}\|, \|a_{45} - b_{25}\|, \\ \|a_{45} - b_{34}\|, \|a_{45} - b_{44}\|, \|a_{45} - b_{54}\|)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$= \min(5, 4.24, 3.61, 3.16, 3, 2, 1.44, 1, 1.44)$$

$$= 1$$

$$\begin{aligned} d(a_{55}, B) &= \min(\|a_{55} - b_{11}\|, \|a_{55} - b_{12}\|, \|a_{55} - b_{13}\|, \|a_{55} - b_{14}\|, \|a_{55} - b_{15}\|, \|a_{55} - b_{25}\|, \\ &\quad \|a_{55} - b_{34}\|, \|a_{55} - b_{44}\|, \|a_{55} - b_{54}\|) \\ &= \min(5.66, 5, 4.47, 4.12, 4, 3, 2.24, 1.44, 1) \\ &= 1 \end{aligned}$$

2) คำนวณหาค่า $h(A, B)$

$$\begin{aligned} h(A, B) &= (d(a_{11}, B) + d(a_{12}, B) + d(a_{13}, B) + d(a_{14}, B) + d(a_{15}, B) + d(a_{25}, B) + \\ &\quad d(a_{35}, B) + d(a_{45}, B) + d(a_{55}, B)) / |A| \\ &= (0 + 0 + 0 + 0 + 0 + 0 + 1 + 1 + 1) / 9 \\ &= 0.33 \end{aligned}$$

3) คำนวณหาค่าฮาอูสเตอร์ฟดิสแทนซ์แบบจุดของทุกๆ จุดใน B

$$\begin{aligned} d(b_{11}, A) &= \min(\|b_{11} - a_{11}\|, \|b_{11} - a_{12}\|, \|b_{11} - a_{13}\|, \|b_{11} - a_{14}\|, \|b_{11} - a_{15}\|, \|b_{11} - a_{25}\|, \\ &\quad \|b_{11} - a_{35}\|, \|b_{11} - b_{45}\|, \|b_{11} - a_{55}\|) \\ &= \min(0, 1, 2, 3, 4, 4.12, 4.47, 5, 5.66) \end{aligned}$$

$$= 0$$

$$\begin{aligned} d(b_{12}, A) &= \min(\|b_{12} - a_{11}\|, \|b_{12} - a_{12}\|, \|b_{12} - a_{13}\|, \|b_{12} - a_{14}\|, \|b_{12} - a_{15}\|, \|b_{12} - a_{25}\|, \\ &\quad \|b_{12} - a_{35}\|, \|b_{12} - b_{45}\|, \|b_{12} - a_{55}\|) \\ &= \min(1, 0, 1, 2, 3, 3.16, 3.61, 4.24, 5) \end{aligned}$$

$$= 0$$

$$\begin{aligned} d(b_{13}, A) &= \min(\|b_{13} - a_{11}\|, \|b_{13} - a_{12}\|, \|b_{13} - a_{13}\|, \|b_{13} - a_{14}\|, \|b_{13} - a_{15}\|, \|b_{13} - a_{25}\|, \\ &\quad \|b_{13} - a_{35}\|, \|b_{13} - b_{45}\|, \|b_{13} - a_{55}\|) \end{aligned}$$

$$= \min(2, 1, 0, 1, 2, 2.24, 2.83, 3.61, 4.47)$$

$$= 0$$

$$\begin{aligned} d(b_{14}, A) &= \min(\|b_{14} - a_{11}\|, \|b_{14} - a_{12}\|, \|b_{14} - a_{13}\|, \|b_{14} - a_{14}\|, \|b_{14} - a_{15}\|, \|b_{14} - a_{25}\|, \\ &\quad \|b_{14} - a_{35}\|, \|b_{14} - b_{45}\|, \|b_{14} - a_{55}\|) \end{aligned}$$

$$= \min(3, 2, 1, 0, 1, 1.41, 2.24, 3.16, 4.12)$$

$$= 0$$

$$\begin{aligned} d(b_{15}, A) &= \min(\|b_{15} - a_{11}\|, \|b_{15} - a_{12}\|, \|b_{15} - a_{13}\|, \|b_{15} - a_{14}\|, \|b_{15} - a_{15}\|, \|b_{15} - a_{25}\|, \\ &\quad \|b_{15} - a_{35}\|, \|b_{15} - b_{45}\|, \|b_{15} - a_{55}\|) \end{aligned}$$

$$= \min(4, 3, 2, 1, 0, 1, 2, 3, 4)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
&= 0 \\
d(b_{25}, A) &= \min(\|b_{25} - a_{11}\|, \|b_{25} - a_{12}\|, \|b_{25} - a_{13}\|, \|b_{25} - a_{14}\|, \|b_{25} - a_{15}\|, \|b_{25} - a_{25}\|, \\
&\quad \|b_{25} - a_{35}\|, \|b_{25} - a_{45}\|, \|b_{25} - a_{55}\|) \\
&= \min(4.12, 3.16, 2.24, 1.41, 1, 0, 1, 2, 3) \\
&= 0 \\
d(b_{34}, A) &= \min(\|b_{34} - a_{11}\|, \|b_{34} - a_{12}\|, \|b_{34} - a_{13}\|, \|b_{34} - a_{14}\|, \|b_{34} - a_{15}\|, \|b_{34} - a_{25}\|, \\
&\quad \|b_{34} - a_{35}\|, \|b_{34} - b_{45}\|, \|b_{34} - a_{55}\|) \\
&= \min(3.61, 2.83, 2.24, 2, 2.24, 1.41, 1, 1.41, 2.24) \\
&= 1 \\
d(b_{44}, A) &= \min(\|b_{44} - a_{11}\|, \|b_{44} - a_{12}\|, \|b_{44} - a_{13}\|, \|b_{44} - a_{14}\|, \|b_{44} - a_{15}\|, \|b_{44} - a_{25}\|, \\
&\quad \|b_{44} - a_{35}\|, \|b_{44} - b_{45}\|, \|b_{44} - a_{55}\|) \\
&= \min(4.24, 3.61, 3.16, 3, 3.16, 2.24, 1.41, 1, 1.14) \\
&= 1 \\
d(b_{54}, A) &= \min(\|b_{54} - a_{11}\|, \|b_{54} - a_{12}\|, \|b_{54} - a_{13}\|, \|b_{54} - a_{14}\|, \|b_{54} - a_{15}\|, \|b_{54} - a_{25}\|, \\
&\quad \|b_{54} - a_{35}\|, \|b_{54} - b_{45}\|, \|b_{54} - a_{55}\|) \\
&= \min(5, 4.47, 4.12, 4, 4.12, 3.16, 2.24, 1.41, 1) \\
&= 1
\end{aligned}$$

4) คำนวณหาค่า $h(B, A)$

$$\begin{aligned}
h(B, A) &= (d(b_{11}, A) + d(b_{12}, A) + d(b_{13}, A) + d(b_{14}, A) + d(b_{15}, A) + d(b_{25}, A) + \\
&\quad d(a_{34}, A) + d(b_{44}, A) + d(b_{54}, A)) / |B| \\
&= (0 + 0 + 0 + 0 + 0 + 0 + 1 + 1 + 1) / 9 \\
&= 0.33
\end{aligned}$$

5) คำนวณหาค่า $H(A, B)$

$$\begin{aligned}
H(A, B) &= \min(h(A, B), h(B, A)) \\
&= \min(0.33, 0.33) \\
&= 0.33
\end{aligned}$$

หมายเหตุ เนื่องจากค่าฮาอูสเตอร์ฟติดิสแทนซ์ที่ได้นี้เป็นค่าความแตกต่างและมีค่าอยู่ในช่วง 0 ถึง ความยาวเส้นทแยงมุมของรูปแบบ (มีค่าเท่ากับ 5.66) ดังนั้นเมื่อจะนำไปเปรียบเทียบกับค่าความเหมือนที่มีค่าอยู่ในช่วง 0 ถึง 1 จึงจำเป็นต้องดำเนินการดังนี้ คือ ปรับค่าฮาอูสเตอร์ฟติดิสแทนซ์ให้มีค่าอยู่ในช่วง 0 ถึง 1 โดยนำเอาค่าฮาอูสเตอร์ฟติดิสแทนซ์หารด้วยความยาวเส้นทแยง-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มุม (0.33/5.66) มีค่าเท่ากับ 0.058 จากนั้นปรับจากค่าความแตกต่างไปเป็นค่าความเหมือน (1 - 0.058) จะได้ค่าเท่ากับ 0.94

ตารางที่ 2.1 แสดงค่าความเหมือนที่ได้จากตัววัดค่าความเหมือนชนิดต่างๆ เมื่อเปรียบเทียบโดยใช้เลข 7 แบบต่างๆ ในรูปที่ 2.3

Similarity metric	Pattern				
	a	b	c	d	e
Hamming	1.00	0.76	0.76	0.76	0.68
Dot product	1.00	0.67	0.67	0.67	0.56
Euclidean	1.00	0.51	0.51	0.51	0.43
Hausdorff	1.00	0.94	0.90	0.88	0.80

จากรูปที่ 2.3 กำหนดให้รูปแบบ (a) เป็นรูปแบบอ้างอิง โดยจะทำการเปรียบเทียบความเหมือนของรูปแบบอื่นๆ กับรูปแบบ (a) จากรูปเลข 7 ทั้งหมดในรูปที่ 2.3 และจากผลการคำนวณค่าความเหมือนด้วยตัววัดค่าความเหมือนแบบต่างๆ ในตารางที่ 2.1 จะเห็นได้ว่ามีเพียงฮามมิงดิสแทนซ์ที่สามารถวัดค่าความเหมือนออกมาได้ตามความรู้สึกของมนุษย์ คือ เลข 7 รูปแบบ (b) จะมีความเหมือนกันกับรูปแบบ (a) มากที่สุด รูปแบบ (c), (d), และ (e) จะมีความเหมือนกับรูปแบบ (a) น้อยลงตามลำดับ เมื่อพิจารณาค่าความเหมือนที่ได้จากวิธีแฮมมิง ดอทโปรดักต์ และยูคลิเดียนจะเห็นได้ว่าตัววัดค่าความเหมือนทั้งสามไม่สามารถแยกความแตกต่างระหว่างรูปแบบของเลข 7 แบบ (b), (c), และ (d) ได้เลย ดังจะเห็นได้จากตารางที่ 2.1 ว่าตัววัดค่าความเหมือนทั้งสามนั้นให้ผลของค่าความเหมือนของรูปแบบ (a) เมื่อเทียบกับรูปแบบ (b), (c), และ (d) เป็นค่าที่เท่ากัน

จากที่กล่าวมาข้างต้นนอกจากฮามมิงดิสแทนซ์จะเป็นตัววัดค่าความเหมือนที่มีประสิทธิภาพใกล้เคียงกับมนุษย์แล้วจะเห็นได้ว่าฮามมิงดิสแทนซ์ยังมีคุณสมบัติอื่นๆ อันเหมาะสมที่จะนำไปใช้ในการรู้จำรูปแบบอีกด้วย ประการแรกคือเป็นการดำเนินการบนเซตที่เป็นเซตแบบปิดและเป็นเซตจำกัด (มีขอบเขตแน่นอน) ประการต่อมาคือให้ค่าที่ไม่เป็นค่าลบเสมอ นอกจากนี้แล้วยังมีคุณสมบัติ identity คือ ความสามารถที่รูปแบบแบบหนึ่งจะแสดงถึงรูปแบบชนิดนั้นได้เพียงอย่างเดียว โดยไม่ซ้ำกับรูปแบบชนิดอื่น มีคุณสมบัติ symmetry ที่แสดงว่าลำดับของการเปรียบเทียบกันระหว่างรูปแบบ 2 รูปแบบไม่มีผลต่อค่าความเหมือน และฮามมิงดิสแทนซ์ยังมีคุณสมบัติที่สำคัญอีกประการหนึ่งในการรู้จำรูปแบบคือ คุณสมบัติ triangle inequality ซึ่งหมายถึงถ้ารูปแบบสองรูปแบบมีความเหมือนกันน้อยมากก็เป็นไปไม่ได้ที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปแบบทั้งสองนี้ทั้งคู่จะมีความเหมือนกันกับรูปแบบที่สามใดๆ ก็ตาม อีกทั้งจากการทดลองและสังเกตพบว่าฮาอูสเตอร์ฟิดิสแทนซ์ยังสามารถที่จะทนต่อการหมุนและการย้ายตำแหน่งของรูปแบบได้เล็กน้อยอีกด้วย

2.4 ซิมพลิไฟด์ฟัซซีอาร์ทแมพ

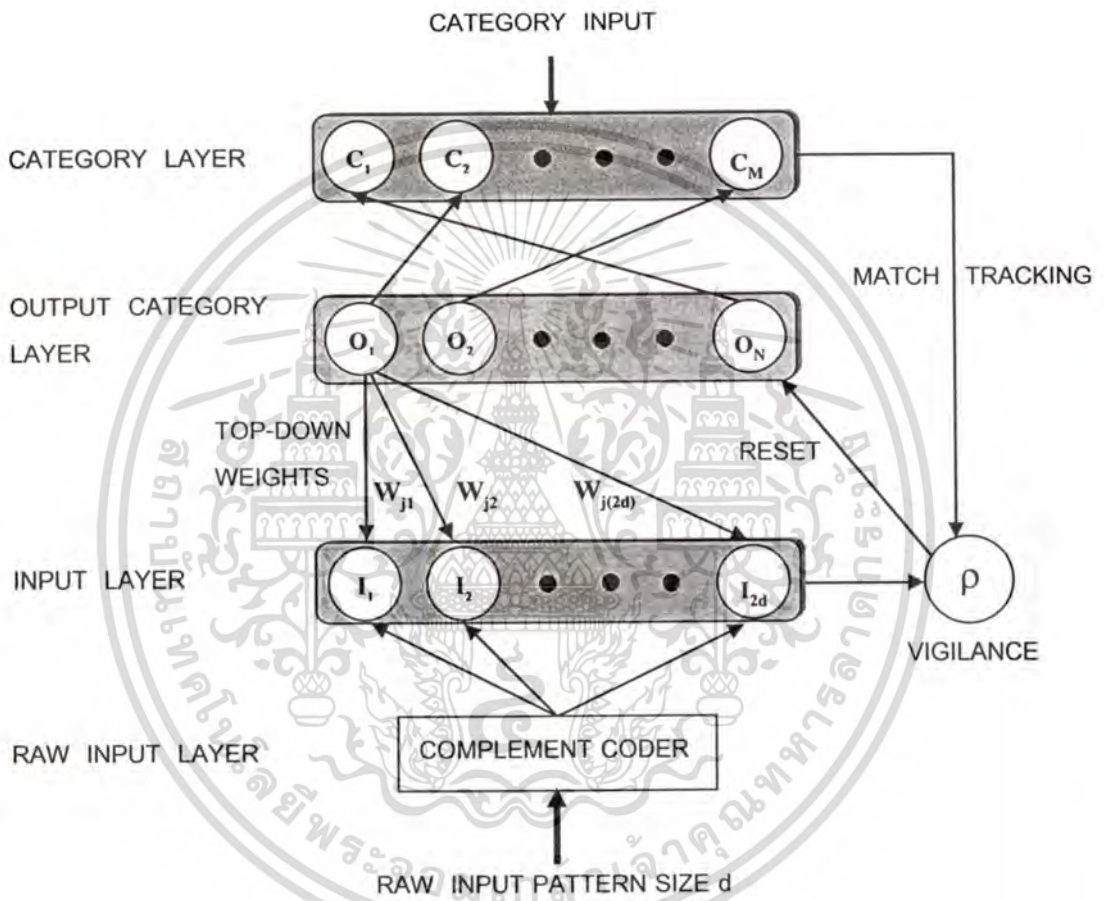
2.4.1 ความรู้เบื้องต้นเกี่ยวกับอาร์ทและซิมพลิไฟด์ฟัซซีอาร์ทแมพ

มนุษย์มีความสามารถที่จะจดจำสิ่งที่เคยได้รับการเรียนรู้มาแล้วและในขณะเดียวกันก็สามารถที่จะเรียนรู้ความรู้ใหม่ๆ เพิ่มได้ ซึ่งโครงข่ายประสาทเทียมส่วนใหญ่จะทำการปรับสิ่งที่เคยเรียนรู้ใหม่ทั้งหมดหรือลืมความรู้เก่าที่เคยเรียนรู้มาแล้วเมื่อต้องการจะเพิ่มความรู้ใหม่เข้าไปในระบบ ปัญหานี้เรียกว่า stability-plasticity dilemma [3] ซึ่ง Adaptive Resonance Theory (ART) ที่ได้รับการพัฒนาขึ้นโดย Carpenter และ Grossberg เพื่อใช้ในการรู้จำรูปแบบความสามารถที่จะแก้ไขปัญหาดังกล่าวข้างต้นได้ เนื่องจากอาร์ทมีคุณสมบัติสำคัญ 2 ประการ คือ สามารถคงความรู้ที่เคยเรียนรู้รูปแบบหรือชนิดของรูปแบบมาแล้วได้ (stability) และสามารถที่จะเรียนรู้รูปแบบใหม่ๆ ได้ไปพร้อมกัน (plasticity) แนวทางอาร์ทนี้ประกอบไปด้วยหลากหลายประเภท เช่น ART1, ART2, fuzzy ART, fuzzy ARTMAP เป็นต้น แต่อย่างไรก็ตามอาร์ททุกประเภทก็มีคุณสมบัติทั้งสองที่กล่าวมาแล้วทั้งสิ้น [4] โดยอาร์ทจะทำการปรับน้ำหนัก (weight) ก็ต่อเมื่อรูปแบบอินพุตนั้นมีความเหมือนกันกับรูปแบบเดิมที่เคยเรียนรู้ไปแล้วมากเพียงพอซึ่งเรียกว่าเกิดเรโซแนนซ์ (resonance) เท่านั้น และสำหรับรูปแบบอินพุตใหม่ที่ยังไม่เคยเรียนรู้มาก่อนอาร์ทจะสร้างเป็นนิวรอน (neuron) หรือโหนด (node) ใหม่โดยไม่กระทบกับนิวรอนเดิม ทั้งนี้อาร์ทยังสามารถเรียนรู้ชนิดข้อมูลที่มีรูปแบบอยู่ในหลากหลายลักษณะได้อีกด้วย

ซิมพลิไฟด์ฟัซซีอาร์ทแมพ (simplified fuzzy ARTMAP : SFAM) เป็นอาร์ทประเภทหนึ่งที่ได้ทำการปรับจากวิธีฟัซซีอาร์ทแมพให้มีความง่ายมากขึ้น [5] คำว่า “ฟัซซี” ในซิมพลิไฟด์ฟัซซีอาร์ทแมพมีความหมายว่าโครงข่ายนี้ได้นำเอาฟัซซีลอจิกมาใช้ ส่วนคำว่า “อาร์ท” แสดงให้เห็นว่าซิมพลิไฟด์ฟัซซีอาร์ทแมพใช้แนวทางของอาร์ท และส่วนคำว่า “แมพ” หมายความว่าวิธีการนี้ได้ดำเนินการเรียนรู้เพื่อที่จะจับคู่หรือแมพจากอินพุตไปเป็นเอาต์พุตที่ถูกต้อง

โครงสร้างสถาปัตยกรรมของซิมพลิไฟด์ฟัซซีอาร์ทแมพในรูปที่ 2.4 อินพุตที่เข้าไปในโครงข่ายจะต้องผ่านตัวคอมพลิเมนต์โคเดอร์ (complement coder) ซึ่งทำการปรับอินพุตให้มีขนาดเพิ่มขึ้นเป็นสองเท่าเพื่อช่วยให้โครงข่ายสามารถพิจารณาขอบเขตของประเภทรูปแบบนั้นๆ ได้ อินพุต (I) ที่ได้รับการเพิ่มขนาดจะถูกส่งต่อไปยังชั้นอินพุต (input layer) ค่าน้ำหนัก (weight, w) จากแต่ละโหนดของชั้นเอาต์พุตแคทากอรีจะได้รับการเชื่อมต่อกับชั้นอินพุต โดยจะ

เรียกชื่อของค่าน้ำหนักนี้ว่า ค่าน้ำหนักแบบบนลงล่าง (top-down weights) ในชั้นแคทากอรี (category layer) จะเป็นชั้นที่บ่งบอกถึงประเภทของกลุ่มข้อมูลซึ่งโครงข่ายจะต้องทำการเรียนรู้จำนวนทั้งหมด M ชนิด เอาร์ทพุทโหนดหนึ่งโหนดจัดอยู่ในประเภทได้เพียงประเภทเดียว ดังนั้นจึงสามารถที่จะชี้ไปยังตำแหน่งในชั้นแคทากอรีได้เพียงตำแหน่งเดียวเท่านั้น แคทากอรีอินพุทสามารถที่จะเพิ่มเข้าไปในชั้นแคทากอรีได้เฉพาะในระหว่างการสอนโครงข่ายเท่านั้น ซึ่งการสอนโครงข่ายนี้เป็นการสอนแบบ supervised



รูปที่ 2.4 สถาปัตยกรรมของซิมพลิไฟด์พีชชีอาร์ทแมพ

จากสถาปัตยกรรมของซิมพลิไฟด์พีชชีอาร์ทแมพสัญลักษณ์ ρ จะใช้แทนถึงค่าวิจิลแลนซ์ (vigilance) ซึ่งมีค่าได้ตั้งแต่ 0 ถึง 1 ค่าวิจิลแลนซ์จะเป็นค่าพารามิเตอร์ที่ผู้ใช้สามารถกำหนดค่าเริ่มต้นได้โดยจะกำหนดให้เป็นค่าพื้นฐาน (baseline) ค่าวิจิลแลนซ์จะเป็นตัวควบคุมการจัดกลุ่มของเอาร์ทพุทโหนดว่ารูปแบบอินพุทจะต้องมีความคล้ายคลึงกันมากน้อยเพียงใดจึงจะสามารถจัดอยู่ในกลุ่มเอาร์ทพุทเดียวกันได้ ยิ่งค่าวิจิลแลนซ์มีค่าสูงก็ยิ่งจะต้องมีความเหมือนกันมากจึงจะจัดอยู่ในกลุ่มเอาร์ทพุทเดียวกันได้ และเมื่อค่าวิจิลแลนซ์มีค่าต่ำอินพุทที่มีความเหมือนกันน้อยก็

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถจัดอยู่ในกลุ่มเดียวกันได้ ทั้งนี้ค่าวิจิลแลนซ์ยิ่งสูงก็จะยิ่งทำให้จำนวนโหนดของเอาร์ทพุท แคทากอรียิ่งมากขึ้นตามไปด้วย

ส่วนของการทำแมทซ์แทรคกิง (match tracking) เป็นส่วนที่โครงข่ายจะทำการปรับค่า วิจิลแลนซ์เองระหว่างการเรียนรู้ โดยจะทำการเพิ่มค่าขึ้นจากค่าพื้นฐานที่ผู้ใช้เป็นคนกำหนด การเพิ่มค่าจะเกิดขึ้นเมื่อมีความผิดพลาดในการจำแนกชนิดข้อมูลในระหว่างช่วงของการสอน โครงข่าย ในการทำแมทซ์แทรคกิงนั้นก็เพื่อให้มีการพิจารณาเพิ่มเอาร์ทพุทโหนดใหม่ขึ้นมาหรือก็คือการลดขนาดของขอบเขตของกลุ่มลง (ขอบเขตในการพิจารณาว่ารูปแบบนั้นๆ อยู่ภายในกลุ่มเดียวกันหรือไม่จะมีขนาดเล็กลง)

2.4.2 กระบวนการทำงานของซิมพลิไฟด์พีชชีอาร์ทแมพ

อินพุทที่เข้ามาจะต้องถูกทำการปรับโดยการทำคอมพลิเมนต์โคดดิ้ง (complement coding) ถ้ารูปแบบอินพุทเวกเตอร์ a มีจำนวนพีเจอร์ทั้งหมด d พีเจอร์ เวกเตอร์ที่ได้รับการทำคอมพลิเมนต์ \bar{a} จะสามารถกำหนดอยู่ในรูปสมการต่อไปนี้คือ

$$\bar{a} = 1 - a \quad (2.8)$$

เนื่องจากโครงข่ายนี้เป็นการดำเนินการที่ใช้พีชชีลอจิก ดังนั้นอินพุททั้งหมดในซิมพลิไฟด์พีชชีอาร์ทแมพจะมีค่าอยู่ระหว่าง 0 ถึง 1 อินพุทเวกเตอร์ที่ได้รับการดำเนินการทำคอมพลิเมนต์แล้วจะมีขนาดความยาวเป็นสองเท่าคือ $2d$ และสามารถแทนค่าของอินพุทได้ดังนี้

$$I = (a, \bar{a}) = (a_1, \dots, a_d, \bar{a}_1, \dots, \bar{a}_d) \quad (2.9)$$

$$|I| = |a, \bar{a}| = \sum_{i=1}^d a_i + (d - \sum_{i=1}^d a_i) = d \quad (2.10)$$

เมื่อค่า $|I|$ สามารถคำนวณได้ดังนี้

$$|p| = \sum_{i=1}^d p_i \quad (2.11)$$

กระบวนการทำงานของโครงข่ายนี้จำเป็นต้องมีวิธีการในการที่จะกระตุ้นการตอบสนองของชั้นเอาร์ทพุทที่มีต่ออินพุทที่เข้ามาในโครงข่าย เมื่อ I คืออินพุทที่อยู่ในรูปที่ได้รับการทำคอมพลิเมนต์แล้วเป็นเอกสารที่ส่งวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พลิเมนต์โคตติงแล้ว ค่าการกระตุ้นของเอาท์พุทสามารถเขียนแทนด้วย T_j สำหรับเอาท์พุทโหนดลำดับที่ j และมีค่าน้ำหนัก w_j ฟังก์ชันในการคำนวณค่านี้คือ

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|} \quad (2.12)$$

ค่า α จะได้รับการกำหนดให้เป็นค่าที่น้อยมากๆ ที่มีค่าเข้าใกล้ 0 ปกติจะให้มีค่าประมาณ 0.0000001 เอาท์พุทโหนดที่จะเป็นโหนดชนะคือโหนดที่มีค่า T_j สูงที่สุด

$$\text{Winner} = \max(T_j) \quad (2.13)$$

ถ้ามี T_j ที่มีค่ามากที่สุดมากกว่า 1 ค่า จะเลือกเอาท์พุทโหนดของโหนดที่ j ซึ่งเป็นโหนดแรกที่พบ (j มีค่าน้อยที่สุด) ชนิดของรูปแบบที่สัมพันธ์กับเอาท์พุทโหนดที่ชนะก็คือวิธีการจำแนกประเภทของรูปแบบอินพุทที่โครงข่ายนี้ดำเนินการนั่นเอง

แมตช์ฟังก์ชัน (match function) เป็นการนำเอาอินพุทที่เจอมาเปรียบเทียบกับค่าน้ำหนักโดยจะนำมาพิจารณาร่วมกับค่าวิจิลแลนซ์ (ρ) เพื่อดูว่าอินพุทมีความเหมือนกันกับเอาท์พุทโหนดที่เชื่อมโยงกันเพียงพอหรือไม่ หรือควรเพิ่มเอาท์พุทโหนดโหนดใหม่เพื่อเป็นตัวแทนของรูปแบบอินพุทนี้ ในการพิจารณามีเงื่อนไขคือถ้าค่าแมตช์ฟังก์ชันมีค่ามากกว่าค่าวิจิลแลนซ์แล้วแสดงว่าเกิดสถานะที่เรียกว่าเรโซแนนซ์ขึ้น ซึ่งเงื่อนไขนี้สามารถแสดงได้ดังต่อไปนี้

$$\frac{|I \wedge w_j|}{d} \geq \rho \quad (2.14)$$

การเกิดเรโซแนนซ์หมายถึงเอาท์พุทโหนดโหนดที่ j นั้นดีเพียงพอที่จะเป็นตัวแทนของอินพุท I ได้ (เอาท์พุทที่ j และอินพุทสามารถจัดอยู่ในกลุ่มเดียวกันได้) และได้ทำการพิจารณาต่อไปว่า ถ้าเอาท์พุทโหนด j นี้ซึ่งเป็นเอาท์พุทโหนดชนิด C และอินพุท I ก็มีชนิด C เช่นเดียวกันด้วยแล้วค่าน้ำหนักของเอาท์พุทโหนด j จะได้รับการปรับค่าเพื่อทำการเรียนรู้โดยใช้อินพุท I นี้ แต่ถ้าค่าของชนิดของเอาท์พุทโหนดกับอินพุทไม่เหมือนกันแสดงว่าเกิดสถานะที่ชนิดคำตอบไม่ถูกต้องหรือแคตทากอรีมิสแมตช์ (category mismatch) ค่าน้ำหนักจะไม่มีเปลี่ยนแปลงใดๆ โดยจะสังเกตได้ว่าเมื่อมีอินพุทตัวหนึ่งเข้ามาเพื่อให้ระบบทำการเรียนรู้จะมีเพียงเอาท์พุทโหนดโหนดเดียวเท่านั้นที่สามารถปรับค่าน้ำหนักได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงข่ายจะอยู่ในสถานะที่เรียกว่ามิสแมทช์รีเซต (mismatch reset) ถ้าค่าของแมทซ์ฟังก์ชันน้อยกว่าค่าของวิจแลนซ์ ซึ่งเป็นการบ่งบอกว่าเอาท์พุทโหนดที่กำลังพิจารณาอยู่นี้มีความคล้ายกันไม่มากพอที่จะเป็นตัวแทนของรูปแบบอินพุทนี้ ดังนั้นจะไม่มีค่าปรับน้ำหนักแม้ว่าคำตอบที่ได้จากเอาท์พุทโหนดนี้กับคำตอบที่แท้จริงของอินพุทจะเป็นคำตอบเดียวกัน

เมื่อโหนด j ซึ่งเป็นโหนดที่ชนะและผ่านเงื่อนไข (2.14) รวมถึงมีชนิดเดียวกันกับอินพุท I ด้วยแล้วค่าน้ำหนักจะได้รับการปรับค่าตามสมการต่อไปนี้คือ

$$w_j^{new} = \beta(I \wedge w_j^{old}) + (1 - \beta)w_j^{old} \quad (2.15)$$

ค่า β คือค่าอัตราการเรียนรู้ซึ่งจะมีค่ามากกว่า 0 แต่น้อยกว่าหรือเท่ากับ 1 ถ้า β ยิ่งมีค่ามากการเรียนรู้ก็จะรวดเร็วมากยิ่งขึ้น ในการปรับค่าน้ำหนักให้ง่ายขึ้นจะให้ค่าอัตราการเรียนรู้เป็น 1 ดังนั้นการปรับค่าน้ำหนักจะสามารถเขียนอยู่ในรูปสมการได้ดังต่อไปนี้คือ

$$w_j^{new} = (I \wedge w_j^{old}) \quad (2.16)$$

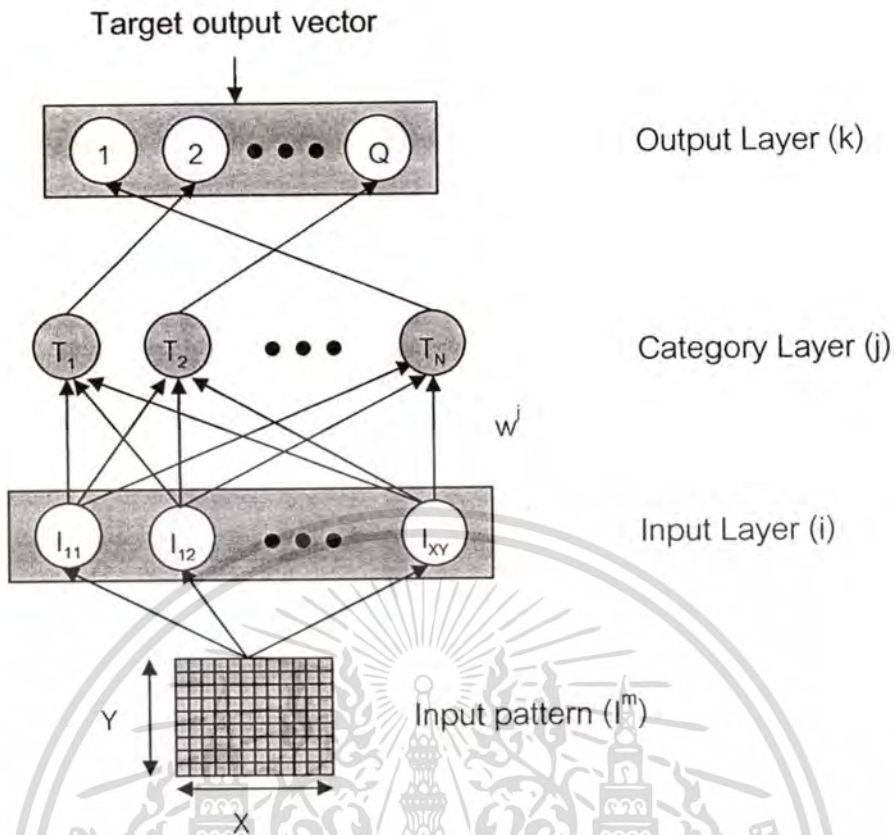
ในกรณีที่เงื่อนไข (2.14) เป็นเท็จ จะทำการเพิ่มค่าวิจแลนซ์ขึ้นให้มากกว่าค่าแมทซ์ฟังก์ชันของโหนดที่ชนะเล็กน้อย และพิจารณาโหนดที่มีค่า T มากที่สุดลำดับถัดไปเพื่อเป็นโหนดที่ชนะโหนดใหม่แทนโหนดเดิม ถ้าพิจารณาจนครบทุกโหนดแล้วไม่มีโหนดใดๆ ที่ผ่านเงื่อนไขนี้เลย หรือกรณีที่ประเภทของอินพุท I กับประเภทที่ได้จากเอาท์พุทโหนด J ที่ชนะเป็นคนละชนิดกันทุกครั้งก็จะทำการเพิ่มเอาท์พุทโหนดใหม่อีกหนึ่งโหนดโดยให้มีชนิดเป็นชนิดเดียวกันกับอินพุท I นั้น และมีค่าน้ำหนักเท่ากับค่าของอินพุท I ซึ่งได้รับการทำคอมพลิเมนต์แล้ว และกรณีที่อินพุทแรกหรืออินพุทนั้นเป็นชนิดที่ยังไม่ได้รับการเรียนรู้มาก่อนก็จะทำการสร้างเอาท์พุทโหนดโหนดใหม่ที่มีชนิดเป็นชนิดเดียวกับอินพุทนั้นและมีค่าน้ำหนักเท่ากับค่าของอินพุทนั้นที่ได้รับการทำคอมพลิเมนต์แล้ว

การรู้จำรูปภาพใบนารีโดยใช้ฮาอูสเตอร์ฟาร์ทแมพ

3.1 บทนำ

กระบวนการใหม่นี้ได้ปรับเอาฮาอูสเตอร์ฟิตสแทนซ์เข้ามาใช้ร่วมกับซิมพลิไฟด์พีชชีอาร์ทแมพ โดยนำเอาแนวคิดของฮาอูสเตอร์ฟิตสแทนซ์มาใช้แทนการดำเนินการแบบพีชชี และมีการรับอินพุทแบบเดียวกันกับโครงข่ายประสาทเทียมแบบนีโอค็อกนิตรอน ค่าฮาอูสเตอร์ฟิตสแทนซ์จะมีค่าได้ตั้งแต่ 0 ถึงระยะทางตามแนวเส้นทแยงมุมของภาพอินพุทซึ่งมีลักษณะเป็นอาร์เรย์ 2 มิติ เมื่อฮาอูสเตอร์ฟิตสแทนซ์มีค่าเท่ากับ 0 แสดงว่ารูปแบบที่นำมาเปรียบเทียบกับโมเดลสามารถแมทช์กันได้พอดี คือ มีความเหมือนกันทุกประการ ค่าน้ำหนักในกระบวนการนี้มีลักษณะเหมือนเป็นโมเดลอ้างอิง (reference model) ที่เป็นแบบใบนารีในวิธีการจับคู่รูปแบบ (pattern matching) โดยทั่วไป แต่แนวคิดนี้ต้องการสร้างโมเดลให้ครอบคลุมและเป็นตัวแทนที่ดีของกลุ่มข้อมูลได้ อีกทั้งสามารถปรับปรุงให้โมเดลนั้นมีคุณภาพที่ดียิ่งขึ้นได้อีกด้วยเนื่องจากแนวทางของอาร์ทสามารถสร้างโหนดเพิ่มเติมได้และค่าน้ำหนักก็สามารถที่จะปรับเปลี่ยนได้ในกระบวนการเรียนรู้ นอกจากนี้แนวทางใหม่นี้ยังอาศัยพื้นฐานแนวคิดมาจากการรู้จำและการมองเห็นของมนุษย์ดังกล่าวมาแล้วในบทที่ 2 โดยอาศัยส่วนของการรู้จำวัตถุที่สำคัญคือการรู้จำโครงร่าง เช่น เส้น ขอบ เป็นต้น สถาปัตยกรรมของกระบวนการรู้จำแบบใหม่นี้สามารถแสดงได้ดังรูปที่ 3.1

ฮาอูสเตอร์ฟาร์ทแมพเป็นโครงข่ายประสาทเทียมที่ประกอบไปด้วย 3 ระดับชั้น ชั้นแรกคือชั้นอินพุท (input layer) ซึ่งในชั้นนี้จะประกอบไปด้วยโหนดจำนวน $X \times Y$ โหนด แต่ละโหนดในชั้นอินพุทจะแทนพิกเซล (pixel) แต่ละพิกเซลของรูปแบบอินพุท ชั้นที่สองคือชั้นแคทากอรี (category layer) โหนดในชั้นที่สองจะได้รับการสร้างขึ้นในระหว่างกระบวนการเรียนรู้ แต่ละโหนดในชั้น แคทากอรีนี้จะเชื่อมต่อแบบทั่วถึงกับโหนดในชั้นอินพุทโดยจะเชื่อมกันผ่านทางน้ำหนัก w^j ซึ่งน้ำหนัก w^j จะมีขนาดมิติเท่ากับรูปแบบอินพุท และค่าน้ำหนักนี้จะเป็นตัวแทนของรูปแบบอ้างอิงของโหนดที่ j ในชั้นแคทากอรี ชั้นสุดท้ายคือชั้นเอาต์พุท (output layer) แต่ละโหนดในชั้นนี้จะเป็ตัวแทนแสดงถึงแต่ละประเภทหรือชนิด (class) ของรูปแบบซึ่งฮาอูสเตอร์ฟาร์ทแมพจะต้องทำการเรียนรู้ ในกระบวนการเรียนรู้แบบ supervised รูปแบบอินพุทแบบใบนารีจะถูกส่งเข้าไปยังโครงข่ายนี้พร้อมกับคำตอบ (target output vector) ซึ่งก็คือชนิดของรูปแบบอินพุทนั่นเอง



รูปที่ 3.1 สถาปัตยกรรมของฮาอูสดอร์ฟอาร์ทแมพ

3.2 กระบวนการเรียนรู้

ในกระบวนการเรียนรู้ที่รูปแบบอินพุตแบบไบนารี (I^m) ซึ่งเป็นรูปภาพของลักษณะโครงสร้างโดยรวม เช่น เส้น ขอบ และรูปร่าง ที่ส่งเข้ามาในระบบสามารถแสดงอยู่ในรูปแบบดังต่อไปนี้

$$I_{x,y}^m = \{1, 0\} \quad : x = 1, 2, \dots, X; y = 1, 2, \dots, Y \quad (3.1)$$

เมื่อ X และ Y คือมิติของรูปแบบอินพุต และ m แสดงลำดับของรูปแบบอินพุตที่ส่งไปยังโครงข่าย

เมื่อมีรูปแบบอินพุต (I^m) ใดๆ เข้ามายังระบบ ระบบก็จะดำเนินการคำนวณค่าชอยซ์ฟังก์ชัน (choice function) ของแต่ละโหนดลำดับที่ j ในชั้นแคตทอรี โดยค่าชอยซ์ฟังก์ชันสามารถคำนวณได้ดังนี้

$$T_j(I^m) = H(I^m, w^j) \quad (3.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $H(I^m, w^j)$ คือ ค่าฮาสเตอร์ฟดิสแทนซ์ระหว่างรูปแบบอินพุท I^m กับรูปแบบอ้างอิงหรือก็คือค่าน้ำหนักของโหนดลำดับที่ j (w^j) แล้วทำการเลือกโหนดที่ชนะ (โหนด J) เพื่อเป็นคำตอบของระบบ โดยพิจารณาเลือกจากโหนดในชั้นแคทากอรีที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดเมื่อเทียบกับโหนดอื่นๆ ในชั้นนี้ เพราะในแนวคิดนี้ค่าชอยซ์ฟังก์ชันยังมีค่าน้อยแสดงว่ายิ่งมีความเหมือนกันระหว่างรูปแบบอินพุทและรูปแบบอ้างอิงมากขึ้นเท่านั้น และถ้าค่าชอยซ์ฟังก์ชันมีค่ามากแสดงว่ารูปแบบทั้งสองก็มีความเหมือนกันน้อย เมื่อโหนดที่ชนะแทนด้วยโหนดที่ J ดังนั้น

$$T_J(I^m) = \min \{T_j(I^m)\} : j=1, 2, \dots, N \quad (3.3)$$

เมื่อ N คือ จำนวนโหนดทั้งหมดในชั้นแคทากอรี กรณีที่มีโหนดซึ่งมีค่าชอยซ์ฟังก์ชันน้อยที่สุดมากกว่าหนึ่งโหนดให้เลือกโหนดที่มีลำดับน้อยที่สุดเป็นโหนดที่ชนะ จากนั้นจะต้องดำเนินการตรวจสอบระดับความแตกต่างระหว่างรูปแบบอินพุทกับรูปแบบอ้างอิงของชนิดข้อมูลที่ถูกละเลือกเป็นคำตอบของระบบว่าอยู่ในระดับที่ยอมรับได้หรือไม่ โดยใช้การพิจารณาค่าวิจิแลนซ์พารามิเตอร์ (ρ) ตามเงื่อนไขต่อไปนี้

$$T_J(I^m) \leq \rho \quad (3.4)$$

ค่า ρ จะมีค่าอยู่ระหว่าง 0 ถึงความยาวของเส้นทแยงมุมของรูปแบบอินพุท ซึ่งในกระบวนการฮาสเตอร์ฟอาร์ทแมพนี้ยิ่งค่า ρ มีค่าน้อยจะยิ่งทำให้การจัดกลุ่มของข้อมูลมีความกระชับมากขึ้น กรณีที่เงื่อนไขที่ (3.4) ไม่เป็นจริงจะไม่ต้องทำการเพิ่มค่า ρ และหาค่าชอยซ์ฟังก์ชันตัวที่น้อยที่สุดถัดมาพิจารณาอีกเหมือนในซิมพลิไฟด์พีชชีอาร์ทแมพ เพราะในแนวทางนี้วิธีการคำนวณค่าชอยซ์ฟังก์ชันกับการตรวจสอบความเหมาะสมในการจัดกลุ่มข้อมูลหรือระดับความต่างกันของรูปแบบอินพุทกับรูปแบบอ้างอิงจะอาศัยค่าที่ได้จากค่าฮาสเตอร์ฟดิสแทนซ์เหมือนกัน ซึ่งทำให้โหนดที่ชนะเป็นโหนดที่มีการจัดกลุ่มข้อมูลที่ดีที่สุดด้วย กล่าวคือถ้าโหนดที่ชนะนี้ไม่ผ่านเงื่อนไขที่ (3.4) โหนดอื่นๆ ก็จะไม่สามารถผ่านเงื่อนไขนี้ได้เช่นกัน ดังนั้นกรณีที่เงื่อนไขที่ (3.4) ไม่เป็นจริงจึงทำการเพิ่มโหนด J ในชั้นแคทากอรีโหนดใหม่เข้าไปในโครงข่ายนี้ได้เลยเพื่อเป็นตัวแทนอ้างอิงของรูปแบบอินพุทนั้น และกำหนดค่าเริ่มต้นให้กับค่าน้ำหนักของโหนด J โหนดใหม่นี้ตามสมการดังต่อไปนี้

$$w^J = I^m \quad (3.5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โหนดใหม่ที่สร้างขึ้นมานี้จะทำให้เงื่อนไขที่ (3.4) เป็นจริงโดยอัตโนมัติ จากนั้นระบบจะเชื่อมโยงโหนด J ซึ่งเป็นโหนดที่ชนะเข้ากับชนิดคำตอบที่แท้จริงในชั้นเอาร์ทพุท

กรณีที่เงื่อนไขที่ (3.4) เป็นจริงจะนำผลลัพธ์ซึ่งเป็นคำตอบที่สัมพันธ์กับโหนดที่ชนะไปเปรียบเทียบกับคำตอบที่แท้จริงของอินพุทนั้น ถ้าผลลัพธ์ที่ได้จากโครงข่ายไม่ถูกต้องตรงกับชนิดของอินพุทให้กลับไปเลือกโหนดที่เหลือในชั้นแคทากอรีที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดลำดับถัดไปมาเป็นโหนดที่ชนะโหนดใหม่ และดำเนินการพิจารณาเงื่อนไข (3.4) ดังกระบวนการที่กล่าวมาข้างต้น แต่ถ้าคำตอบของระบบ (คำตอบในชั้นเอาร์ทพุทที่เชื่อมโยงกันกับโหนดที่ชนะในชั้นแคทากอรี) ตรงกันกับคำตอบที่ถูกต้อง แสดงว่าเกิดเรโซแนนซ์ขึ้นก็จะทำการปรับค่าน้ำหนักของโหนดที่ J (w^J) ซึ่งเป็นโหนดที่ชนะตามขั้นตอนดังต่อไปนี้

$$b_{x,y}^J = \sum_{u=-S}^S \sum_{v=-S}^S I_{x+u,y+v}^m \quad (3.6)$$

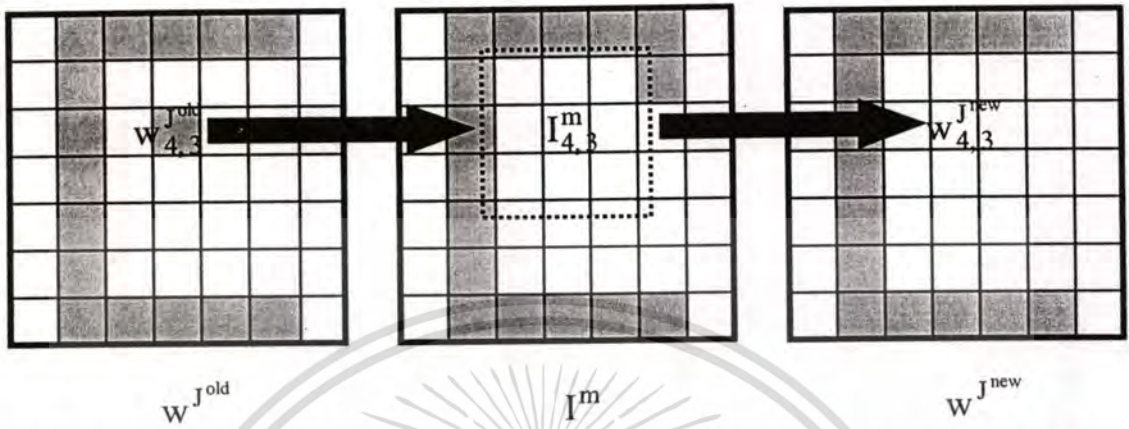
$$w_{x,y}^{J^{new}} = \begin{cases} w_{x,y}^{J^{old}}, & \text{if } b_{x,y}^J > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.7)$$

เมื่อ S คือค่าที่เป็นเลขจำนวนเต็มบวกซึ่งจะเป็นค่าที่บ่งบอกว่าจะพิจารณาพิกเซลในพื้นที่รอบตำแหน่ง x, y ออกไปเป็นจำนวนเท่าไร โดยในกรณีที่ค่านี้นี้มีค่าเท่ากับ S แสดงว่าจะต้องพิจารณาจุดใน I^m ทั้งหมด $(2S+1)^2$ พิกเซล $w_{x,y}^{J^{old}}$ คือ ค่าน้ำหนักในตำแหน่งพิกเซล x, y ของค่าน้ำหนักเดิมที่โหนด J และจากสมการ (3.6) และ (3.7) มีความหมายคือ ต้องการพิจารณาจุดภายในพื้นที่ขนาด $(2S+1) \times (2S+1)$ รอบพิกเซลในตำแหน่ง x, y ของรูปแบบอินพุท I^m ว่ามีพิกเซลใดๆ บ้างหรือไม่ที่มีค่าเท่ากับ 1 ถ้ามีพิกเซลที่มีค่าเท่ากับ 1 อยู่ในบริเวณพื้นที่นี้และ $w_{x,y}^{J^{old}}$ ก็มีค่าเท่ากับ 1 แล้ว $w_{x,y}^{J^{new}}$ จะมีค่าเท่ากับ 1 ด้วย ส่วนกรณีอื่นๆ ให้ $w_{x,y}^{J^{new}}$ มีค่าเท่ากับ 0 จุดประสงค์เพื่อพิจารณาว่ามีพิกเซลซึ่งเป็นเพื่อนบ้านกับ $I_{x,y}^m$ ภายในระยะ S ที่มีค่าเท่ากับ 1 หรือไม่ ถ้าไม่มีก็แสดงว่า $w_{x,y}^{J^{old}}$ ที่มีค่าเท่ากับ 1 นั้นน่าจะเป็นพิกเซลที่เป็น noise เพราะเมื่อพิจารณาพื้นที่รอบๆ พิกเซลนั้นใน I^m ซึ่งเป็นรูปแบบที่มีความเหมือนกันมากก็ไม่มีพิกเซลใดๆ ในบริเวณนี้เลยที่มีค่าเท่ากับ 1 จึงควรกำจัดค่าพิกเซลนี้ออกไปโดยให้ $w_{x,y}^{J^{new}}$ มีค่าเท่ากับ 0

นอกจากนี้จะสังเกตได้ว่าค่า S ในแนวทางนี้จะทำหน้าที่คล้ายตัวกำหนดอัตราการเรียนรู้ด้วย ทั้งนี้จะเห็นได้ว่ายิ่งค่า S มีค่าน้อยโอกาสที่จะทำการปรับค่าน้ำหนักก็มีมากขึ้น เช่น ถ้าค่า S มีค่าเท่ากับ 0 แสดงว่าให้ปรับน้ำหนักของโหนดนั้นโดยดำเนินการเหมือนการดำเนินการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบจุดต่อจุด (หรือก็คือการดำเนินการด้วยการ and) กับรูปแบบอินพุตที่ทำการเรียนรู้เข้ามา และถ้า $2S+1$ มีค่าเท่ากับความยาวของรูปแบบอินพุตก็แสดงว่าจะไม่มีการปรับค่าน้ำหนักใดๆ เลย



รูปที่ 3.2 ตัวอย่างการปรับค่าน้ำหนักจากค่าน้ำหนักเดิม $W^{J^{old}}$ ด้วยรูปแบบอินพุต I^m ไปเป็นค่าน้ำหนักใหม่ $W^{J^{new}}$ เมื่อ $S = 1$

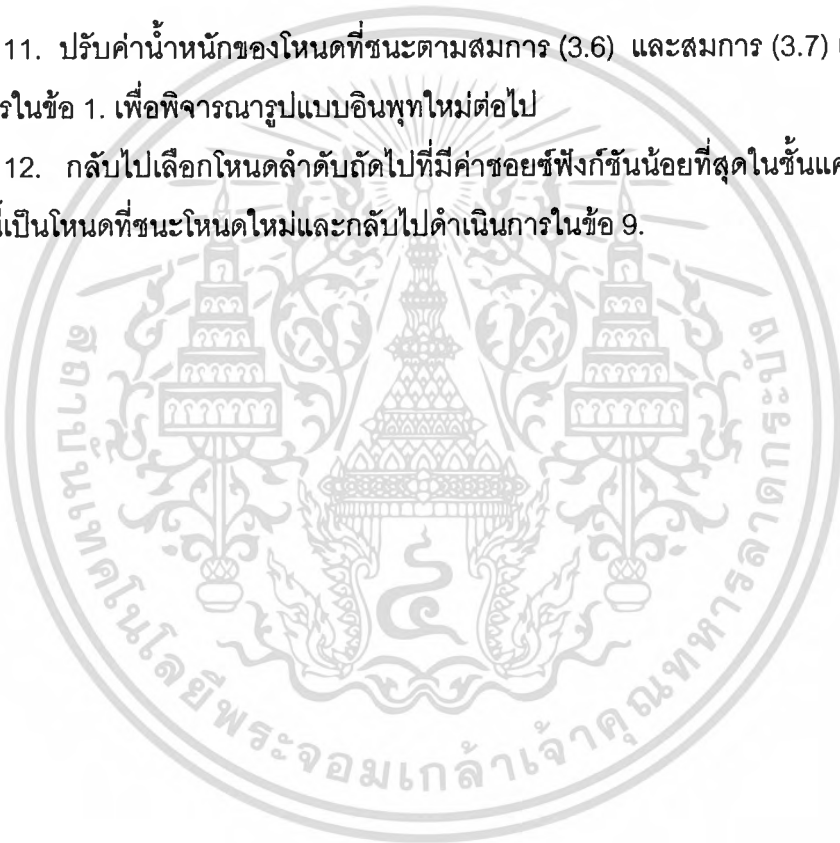
จากรูปที่ 3.2 จุดสีเทาจะแทนตำแหน่งพิกเซลที่มีค่าเท่ากับ 1 และจุดสีขาวแทนตำแหน่งที่มีค่าเท่ากับ 0 พิจารณาที่ตำแหน่ง (4, 3) ของค่าน้ำหนักเก่าเทียบกับบริเวณรอบๆ ในตำแหน่งเดียวกันของรูปแบบอินพุตออกไปจากตำแหน่งนั้นเป็นระยะทาง S ในที่นี้ S กำหนดให้มีค่าเท่ากับ 1 จะเห็นได้ว่าไม่มีค่าพิกเซลใดในบริเวณนี้ (พื้นที่ในกรอบเส้นประ) ที่มีค่าเท่ากับ 1 เลย ดังนั้นค่าน้ำหนักใหม่ที่ได้รับการปรับปรุงแล้วที่ตำแหน่ง (4, 3) จะมีค่าเท่ากับ 0

กระบวนการเรียนรู้ของฮาอูสเตอร์ฟอาร์ทแมพสามารถสรุปเป็นขั้นตอนได้ดังนี้ คือสำหรับทุกรูปแบบอินพุตที่ต้องการทำการเรียนรู้ให้ดำเนินการดังนี้

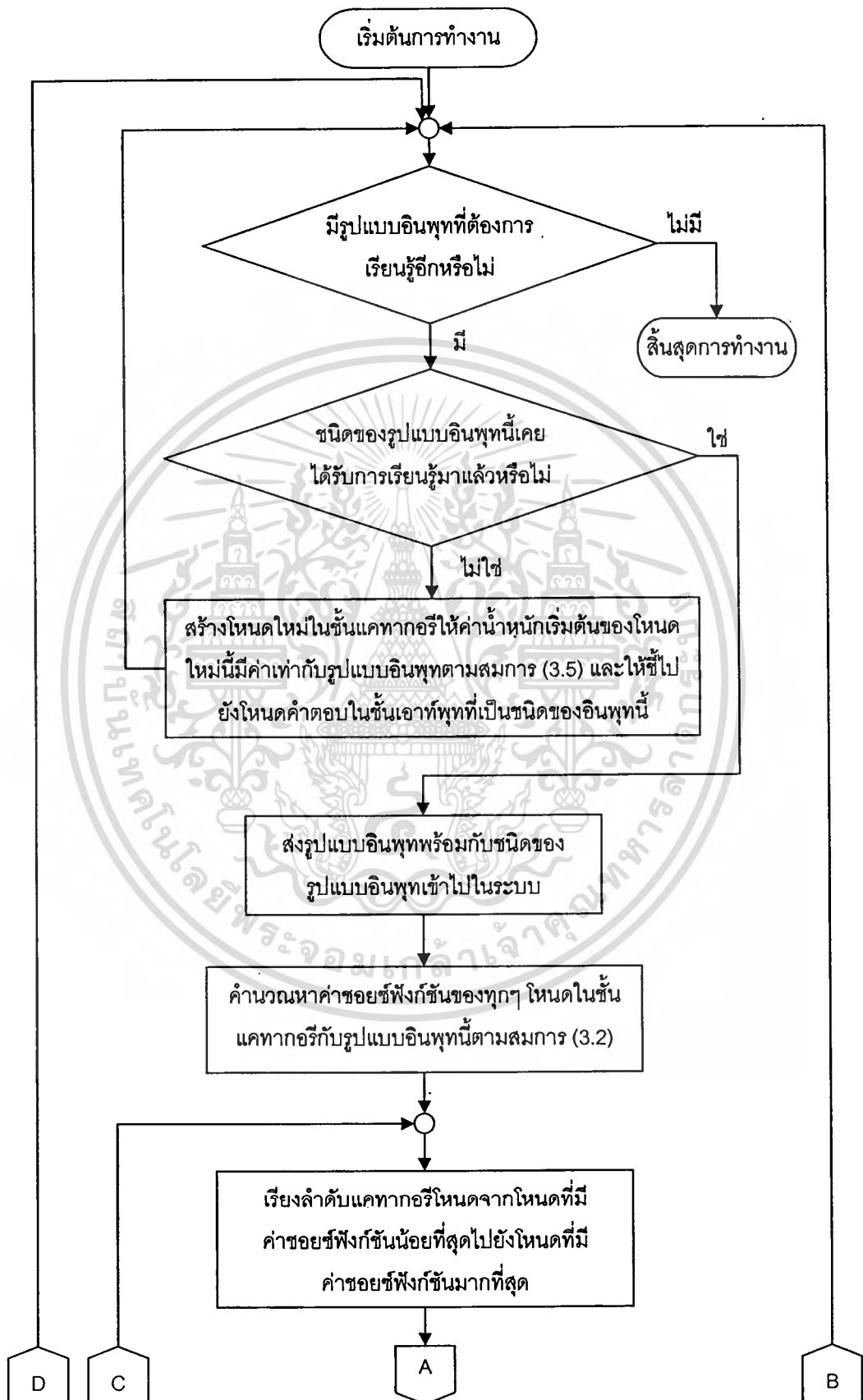
1. ถ้าประเภทของข้อมูลนั้นยังไม่เคยได้รับการเรียนรู้มาก่อนให้ดำเนินการในข้อ 2 ถ้าเป็นข้อมูลที่มีชนิดเป็นชนิดที่ระบบเคยเรียนรู้มาแล้วให้ข้ามไปดำเนินการในข้อ 5
2. สร้างโหนดใหม่ในชั้นแคทากอรีเพื่อเป็นรูปแบบอ้างอิงของอินพุตนี้
3. ให้ค่าน้ำหนักเริ่มต้นของโหนดใหม่นี้มีค่าเท่ากับรูปแบบอินพุตตามสมการ (3.5)
4. ให้โหนดใหม่นี้ชี้ไปยังคำตอบในชั้นเอาต์พุตที่เป็นชนิดเดียวกันกับอินพุตนี้ กลับไปดำเนินการในข้อ 1 เพื่อพิจารณารูปแบบอินพุตใหม่ต่อไป
5. ส่งรูปแบบอินพุตและชนิดคำตอบของรูปแบบอินพุตไปยังระบบ
6. คำนวณหาค่าชอยซ์ฟังก์ชันของทุกๆ โหนดในชั้นแคทากอรีกับรูปแบบอินพุตนี้ตามสมการ (3.2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

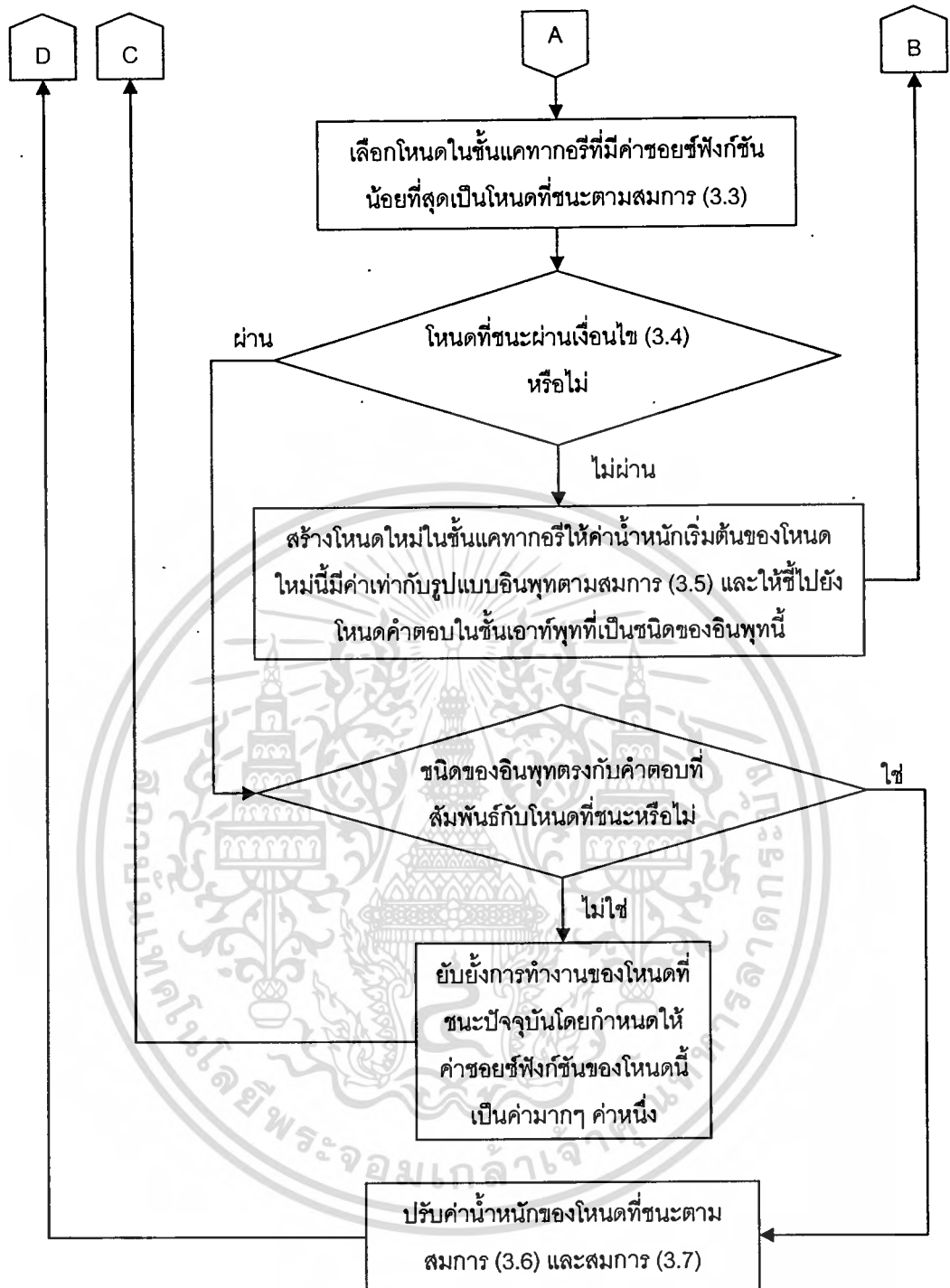
7. เรียงลำดับแคทากอรีไหนดจากไหนดที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดไปยังไหนดที่มีค่าชอยซ์ฟังก์ชันมาก
8. เลือกไหนดในชั้นแคทากอรีที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดเป็นไหนดที่ชนะซึ่งจะได้ไหนดที่ชนะตามสมการ (3.3)
9. จากไหนดที่ชนะนำไปตรวจสอบเงื่อนไข (3.4) ถ้าเงื่อนไขเป็นจริงให้ดำเนินการในข้อ 10
10. ถ้าเงื่อนไข (3.4) เป็นเท็จให้ไปดำเนินการในข้อ 2.
10. ถ้าไหนดที่ชนะให้คำตอบเดียวกันกับคำตอบของรูปแบบอินพุทนี้ให้ดำเนินการในข้อ 11.
11. ถ้าคำตอบที่ได้จากระบบกับคำตอบที่แท้จริงของรูปแบบอินพุทไม่ตรงกันให้ไปดำเนินการในข้อ 12.
11. ปรับค่าน้ำหนักของไหนดที่ชนะตามสมการ (3.6) และสมการ (3.7) แล้วจึงกลับไปดำเนินการในข้อ 1. เพื่อพิจารณารูปแบบอินพุทใหม่ต่อไป
12. กลับไปเลือกไหนดลำดับถัดไปที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดในชั้นแคทากอรี แล้วให้ไหนดนี้เป็นไหนดที่ชนะไหนดใหม่และกลับไปดำเนินการในข้อ 9.



ขั้นตอนการทำงานของฮาลูสตอร์ฟาร์ทแมพสามารถเขียนเป็นแผนผังลำดับงานได้ดังนี้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



3.3 กระบวนการทดสอบและการนำไปใช้งาน

ในส่วนของการทดสอบนี้เป็นการตรวจสอบว่าระบบที่ผ่านการเรียนรู้มาแล้วนั้นสามารถที่จะให้ผลความถูกต้องในการรู้จำมากน้อยเพียงใด ในการหาคำตอบของระบบทำได้โดยการคำนวณหาค่าชอยซ์ฟังก์ชันจากสมการ (3.2) และเลือกโหนดในชั้นแคตาคอรีที่ให้ค่าชอยซ์ฟังก์ชันน้อยที่สุดดังสมการ (3.3) เป็นโหนดที่ชนะ ชนิดของรูปแบบในชั้นอินพุทที่สัมพันธ์กับโหนดที่ชนะจะถือเป็นคำตอบของระบบ แล้วเปรียบเทียบกับคำตอบที่แท้จริงกับผลลัพธ์ที่ได้จากระบบเพื่อวัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัตราความถูกต้องในการรู้จำ และต้องทำการทดลองปรับค่า p และ S เป็นค่าต่างๆ เพื่อหาค่า p และ S ที่เหมาะสมเพื่อให้ได้ผลความถูกต้องในการรู้จำที่ดีที่สุด

การนำไปใช้งานจริงผลในการทำนาย คือ คำตอบที่ได้จากระบบซึ่งเป็นคำตอบของโหนดในชั้นเอาต์พุตที่สัมพันธ์กับโหนดที่ชนะในชั้นแคทากอรี

3.4 การวิเคราะห์เปรียบเทียบความซับซ้อนในการคำนวณและประสิทธิภาพระหว่างวิธีการฮาลอสเตอร์ฟาร์ทแมพและวิธีการซิมพลิไฟด์พีชชีอาร์ทแมพ

3.4.1 ขนาดอินพุต

เมื่อพิจารณาขนาดของรูปแบบอินพุตถ้าให้ขนาดอินพุตมีค่าเท่ากับ XY สำหรับวิธีการซิมพลิไฟด์พีชชีอาร์ทแมพจะต้องทำคอมพลิเมนต์โคดดิ้งดังได้กล่าวไว้ในหัวข้อ 2.4.2 ซึ่งจะทำให้ขนาดอินพุตมีค่าเพิ่มขึ้นเป็น 2 เท่า คือจะมีค่าเท่ากับ $2XY$ แต่สำหรับวิธีการฮาลอสเตอร์ฟาร์ทแมพขนาดอินพุตจะมีค่าเท่าเดิมคือ XY เพราะไม่มีการทำคอมพลิเมนต์โคดดิ้ง

3.4.2 การคำนวณชอยซ์ฟังก์ชันและแมทซ์ฟังก์ชัน

ในกระบวนการเรียนรู้ฮาลอสเตอร์ฟาร์ทแมพจะใช้สมการเดียวกันในการคำนวณหาค่าของชอยซ์ฟังก์ชันและแมทซ์ฟังก์ชันซึ่งจะคำนวณได้จากค่าฮาลอสเตอร์ฟิดิสแทนซ์ ทั้งนี้เป็นผลมาจากฮาลอสเตอร์ฟิดิสแทนซ์มีคุณสมบัติ symmetry คือค่าฮาลอสเตอร์ฟิดิสแทนซ์ของรูปแบบอินพุตเทียบกับรูปแบบอ้างอิง (ค่าน้ำหนัก) จะมีค่าเท่ากับค่าฮาลอสเตอร์ฟิดิสแทนซ์ของรูปแบบอ้างอิงเทียบกับรูปแบบอินพุต และด้วยเหตุผลที่ค่าชอยซ์ฟังก์ชันและแมทซ์ฟังก์ชันมีค่าเท่ากันทำให้กรณีซึ่งโหนดที่ชนะไม่ผ่านเงื่อนไขของวิจิลแลนซ์ระบบสามารถที่จะตัดสินใจเพิ่มโหนดใหม่เข้าไปในชั้นแคทากอรีได้ทันทีโดยไม่ต้องกลับไปพิจารณาหาโหนดใหม่ในชั้นแคทากอรีที่มีค่าชอยซ์ฟังก์ชันน้อยที่สุดลำดับถัดไปมาเป็นโหนดที่ชนะโหนดใหม่อีก ด้วยเหตุผลที่ว่าโหนดที่ชนะซึ่งมีค่าชอยซ์ฟังก์ชันน้อยที่สุดก็จะเป็นโหนดที่มีค่าแมทซ์ฟังก์ชันน้อยที่สุดด้วย หากโหนดนี้ไม่ผ่านเงื่อนไขของวิจิลแลนซ์แล้วโหนดอื่นๆ ก็จะไม่สามารถที่จะผ่านเงื่อนไขนี้ได้ด้วยเช่นกัน สำหรับกระบวนการซิมพลิไฟด์พีชชีอาร์ทแมพจะต้องกลับไปพิจารณาหาโหนดใหม่ในชั้นแคทากอรีที่มีค่าชอยซ์ฟังก์ชันมากที่สุดลำดับถัดไปมาเป็นโหนดที่ชนะแล้วนำไปดำเนินการพิจารณาเงื่อนไขวิจิลแลนซ์ต่อไป

3.4.3 ความซับซ้อนในการคำนวณหาค่าฮาลอสเตอร์ฟิดิสแทนซ์กับการคำนวณโดยใช้พีชชี and โอเปอร์เรชัน

เมื่อ ${}^X C_i$ คือ จำนวนรูปแบบที่แตกต่างกันซึ่งสามารถเกิดขึ้นได้เมื่อจำนวนจุดที่เป็น 1 ในรูปแบบนั้นๆ มีจำนวน i จุด การคำนวณค่าฮาลอสเตอร์ฟิดิสแทนซ์จะคำนวณเฉพาะจุดที่มีค่าเป็น 1 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เท่านั้น จำนวนจุดที่เป็น 1 เฉลี่ยจะมีค่าเท่ากับ $\frac{\sum_{i=1}^{XY} C_i}{\sum_{i=1}^{XY} C_i * i}$ จุด ซึ่งการคำนวณค่า ฮาอูสดอร์ฟดิสแทนซ์ในกระบวนการฮาอูสดอร์ฟอาร์แมพจะใช้เวลาในการคำนวณยูคลิเดียนดิสแทนซ์และเลือกค่าดิสแทนซ์ที่น้อยที่สุดเท่ากับจำนวนจุดที่มีค่าเป็น 1 ในรูปแบบอินพุทบวกกับจำนวนจุดที่มีค่าเป็น 1 ในรูปแบบอ้างอิง ทำให้ใช้เวลาเฉลี่ยในการคำนวณเท่ากับ

$$2 * (1 / \sum_{i=1}^{XY} C_i) * (\sum_{i=1}^{XY} C_i * i)$$

การคำนวณหาค่า $H(I, w)$ เมื่อ I แทนรูปแบบอินพุทใดๆ และ w แทนรูปแบบอ้างอิงใดๆ สามารถแสดงได้ดังนี้คือ

For all pixel (i, j) in I that $I(i, j) = 1$

For all pixel (u, v) in w that $w(u, v) = 1$

Find Euclidean distance from $I(i, j)$ to $w(u, v)$:

$$\| I(x, y) - w(u, v) \| = ((x-u)^2 + (y-v)^2)^{1/2}$$

End

$d(I(i, j), w)$ is the minimum of such distances

End

$$h(I, w) = d / |I|$$

For all pixel (u, v) in w that $w(u, v) = 1$

For all pixel (i, j) in I that $I(i, j) = 1$

Find Euclidean distance from $w(u, v)$ to $I(i, j)$:

$$\| I(x, y) - w(u, v) \| = ((x-u)^2 + (y-v)^2)^{1/2}$$

End

$d(w(u, v), I)$ is the minimum of such distances

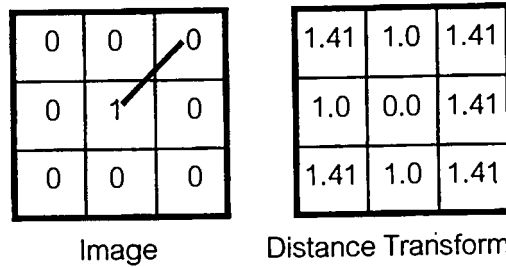
End

$$h(w, I) = d / |w|$$

$H(I, w)$ is the maximum between $h(I, w)$ and $h(w, I)$

เพื่อลดการคำนวณที่ซับซ้อนลงสามารถที่จะอาศัยหลักการของยูคลิเดียนดิสแทนซ์ทรานส์ฟอร์ม (Euclidean distance transform) โดยจะทำการหายูคลิเดียนดิสแทนซ์ทรานส์ฟอร์มของรูปแบบอินพุทก่อนแล้วจึงนำเอารูปแบบอินพุทนี้ไปเปรียบเทียบกับรูปแบบอ้างอิงในระบบ การแปลงรูปแบบอินพุทให้อยู่ในรูปยูคลิเดียนดิสแทนซ์ทรานส์ฟอร์มมีหลักการดังนี้คือทุกๆจุดในรูปแบบจะทำการหาระยะทางจากจุดนั้นไปยังจุดที่มีค่าเป็น 1 ที่ใกล้ที่สุดดังรูปที่ 3.3 ซึ่งถ้าใช้วิธีการดำเนินการใน [6] จะใช้เวลาในการคำนวณยูคลิเดียนดิสแทนซ์ทรานส์ฟอร์มเท่ากับ $O(X*Y)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แสดงผลลัพธ์จากการดำเนินการด้วยวิธียูคลิเดียนดิสแทนซ์ทรานส์ฟอร์ม

การคำนวณหาค่าฮิสโตแกรมฟิดิสแทนซ์ระหว่างรูปแบบอินพุตหนึ่งรูปแบบกับทุกๆ รูปแบบอ้างอิงจำนวน N รูปแบบรวมกับการใช้ยูคลิเดียนทรานส์ฟอร์มจะสามารถดำเนินการได้ดังนี้

Find Euclidean distance transform (E) of I

For $n = 1$ to N

$$d(w_n(x,y), I) = E(x,y)$$

$$h(w_n, I) = d(w_n, I) / |I|$$

For all pixel (i,j) in I that $I(i, j) = 1$

For all pixel (u, v) in w_n that $w_n(u, v) = 1$

Find Euclidean distance from $I(i, j)$ to $w_n(u, v)$

$$\| I(x, y) - w_n(u, v) \| = ((x-u)^2 + (y-v)^2)^{1/2}$$

End

$d(I(i, j), w_n)$ is the minimum of such distances

End

$$h(I, w_n) = d / |I|$$

$H(I, w_n)$ is the maximum between $h(I, w_n)$ and $h(w_n, I)$

End

เมื่ออาศัยวิธีการยูคลิเดียนทรานส์ฟอร์มเพื่อสร้าง look up table ในการหาค่าไดเรกต์ดีดฮิสโตแกรมฟิดิสแทนซ์แบบจุด ($d(w(x,y), I)$) ของทุกจุด x, y ใน w เมื่อเทียบกับ I จะทำให้เวลาที่ในแต่ละรอบในการคำนวณรูปแบบอินพุต I เทียบกับทุกรูปแบบอ้างอิงลดลงจากเดิมครึ่งหนึ่งแต่จะเพิ่มเวลาในส่วนของ การคำนวณหา ยูคลิเดียนทรานส์ฟอร์มเข้ามาด้วย

ดังนั้นจะเห็นได้ว่าวิธีการคำนวณในซิมพลิไฟด์พีซีอาร์ทแมพเป็นการคำนวณที่มีความซับซ้อนน้อยกว่าวิธีฮิสโตแกรมฟิดิสแทนซ์เพราะในแต่ละรอบของการคำนวณในซิมพลิไฟด์พีซีอาร์ทแมพเป็นเพียงการดำเนินการ and กันเท่านั้น ส่วนในวิธีฮิสโตแกรมฟิดิสแทนซ์แต่ละ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รอบของการคำนวณจะต้องทำการหาค่ายูคลิเดียนดิสแทนซ์และหาระยะทางที่น้อยที่สุดจากยูคลิเดียนดิสแทนซ์เหล่านั้นด้วย แต่อย่างไรก็ดีจากวิธีการคำนวณหาค่าฮาลูสเตอร์ฟิดิสแทนซ์ข้างต้นก็สามารถที่จะลดเวลาในการคำนวณในส่วนนี้ลงได้เกือบครึ่งหนึ่ง

3.4.4 เปรียบเทียบประสิทธิภาพของตัววัดค่าความเหมือนของฮาลูสเตอร์ฟาร์ทแมพกับซิมพลิไฟด์พีชชีอาร์ทแมพ

ในการวัดค่าความเหมือนของรูปแบบที่เป็นแบบไบนารีเมื่อใช้วิธีการ and กันตามวิธีการในซิมพลิไฟด์พีชชีอาร์ทแมพรูปแบบที่พิจารณาเปรียบเทียบกันจะต้องมีจุดที่มีค่าเป็น 1 ตรงกันเท่านั้นวิธีการนี้จึงจะให้ค่าคำตอบว่ามีความเหมือนกันที่จุดนั้น แต่ในรูปแบบใดๆ สองรูปแบบที่มีความคล้ายกันนั้นไม่จำเป็นที่จะต้องมีความเหมือนกันหมดทั้งภาพก็ได้ โดยในกระบวนการฮาลูสเตอร์ฟาร์ทแมพนำเอาฮาลูสเตอร์ฟิดิสแทนซ์มาใช้ซึ่งเป็นวิธีการที่ไม่ได้วัดค่าความเหมือนแบบจุดต่อจุดจึงมีประสิทธิภาพในการหาค่าความเหมือนของรูปภาพไบนารีที่ดีกว่าวิธีการซิมพลิไฟด์พีชชีอาร์ทแมพมาก แม้ว่าความซับซ้อนที่เกิดขึ้นจากการคำนวณฮาลูสเตอร์ฟิดิสแทนซ์จะมีมากกว่าก็ตาม แต่เหตุผลในแง่ของประสิทธิภาพในการวัดค่าความเหมือนที่ดีกว่ามากเมื่อเทียบกับการ and กันนั้นจึงสมควรที่จะใช้ฮาลูสเตอร์ฟิดิสแทนซ์เป็นตัววัดค่าความเหมือนแทนการ and กันแบบพีชชี นอกจากนี้แล้วในการปรับค่าน้ำหนักตามวิธีการเดิมซึ่งใช้หลักการของพีชชี and เช่นกันก็ไม่เป็นการเหมาะสมเท่ากับการปรับค่าน้ำหนักในกระบวนการฮาลูสเตอร์ฟาร์ทแมพซึ่งสามารถที่จะสร้างรูปแบบอ้างอิงได้ดีกว่าเพราะเป็นการพิจารณาปรับค่าน้ำหนักโดยดูจากค่าพิกเซลใกล้เคียงประกอบกันไม่ได้พิจารณาเฉพาะตำแหน่งที่ตรงกันเท่านั้น

บทที่ 4

การทดลองและผลการทดลอง

งานวิจัยชิ้นนี้ได้นำเอาวิธีการรู้จำรูปแบบแนวทางใหม่นี้มาทดลองกับข้อมูล 3 ชนิดด้วยกัน ประกอบด้วยข้อมูลลายนิ้วมือ ใบหน้า และมือ ซึ่งข้อมูลทั้ง 3 ประเภทนี้เป็นส่วนหนึ่งของเทคนิคการระบุบุคคลโดยใช้ลักษณะทางกายภาพหรือพฤติกรรมของมนุษย์ ซึ่งเทคนิคนี้เรียกว่า biometrics [7] ใน [7] ได้แสดงการเปรียบเทียบถึงคุณสมบัติของแต่ละเทคนิคในการระบุบุคคลทั้งสามประเภทนี้ไว้ดังตารางที่ 4.1 โดยเปรียบเทียบจากคุณสมบัติดังต่อไปนี้ คือ

- 1) universality คือ ลักษณะนี้เป็นลักษณะที่ควรจะมีอยู่ในทุกคน
- 2) uniqueness คือ สามารถใช้ลักษณะนี้ระบุบุคคลได้โดยไม่ซ้ำกัน
- 3) permanence คือ ลักษณะนี้จะไม่เปลี่ยนแปลงไปตามเวลา
- 4) collectability คือ ลักษณะที่สามารถถูกวัดได้ในเชิงปริมาณ
- 5) performance คือ ผลลัพธ์ความถูกต้องในการระบุบุคคลที่ได้รับ, ทรัพยากรที่ต้องใช้ และปัจจัยการทำงานที่มีผลต่อความถูกต้องในการระบุบุคคล
- 6) acceptability คือ ความรู้สึกของประชาชนในการยอมรับระบบ biometrics นั้นๆ
- 7) circumvention คือ ระดับที่บ่งบอกว่ายากแค่ไหนในการที่จะหลอกระบบ ด้วยวิธีการปลอมแปลงแบบต่างๆ

ตารางที่ 4.1 คุณสมบัติของวิธีการระบุบุคคลด้วยใบหน้า ลายนิ้วมือ และมือ

Biometric	คุณสมบัติข้อที่						
	1)	2)	3)	4)	5)	6)	7)
ใบหน้า	สูง	ต่ำ	กลาง	สูง	ต่ำ	สูง	ต่ำ
ลายนิ้วมือ	กลาง	สูง	สูง	กลาง	สูง	กลาง	สูง
มือ	กลาง	กลาง	กลาง	สูง	กลาง	กลาง	กลาง

จากตารางที่ 4.1 จะเห็นได้ว่าวิธีการระบุบุคคลด้วยใบหน้า ลายนิ้วมือ และมือมีคุณสมบัติที่เหมาะสมในการนำไปใช้ในงานระบุบุคคล ดังนั้นการที่งานวิจัยนี้ได้นำเอาข้อมูลทั้งสามประเภทข้างต้นมาทำการรู้จำจึงน่าจะทำให้เกิดประโยชน์ในการนำไปใช้งานด้านการระบุบุคคลต่อไป แต่อย่างไรก็ตามในแต่ละประเภทของข้อมูลจะมีวิธีการในการเตรียมข้อมูลที่แตกต่างกันออกไปก่อนที่จะนำข้อมูลนั้นๆ ไปดำเนินการรู้จำด้วยวิธีฮาดิสเตอร์ฟาร์ทแมพที่คิดขึ้น

นี้ เนื่องจากกระบวนการฮาอูสเตอร์ฟอาร์ทแมพจะรับข้อมูลอินพุตที่เป็นแบบไบนารีเท่านั้นจึงจำเป็นที่จะต้องแปลงข้อมูลเบื้องต้นให้เป็นข้อมูลแบบไบนารีเสียก่อนจึงจะนำไปดำเนินการรู้จำในส่วนของฮาอูสเตอร์ฟอาร์ทแมพได้ นอกจากนี้วิธีการเตรียมข้อมูลแล้วในบทนี้ได้กล่าวถึงการทดลองและผลการทดลองจากการดำเนินการรู้จำด้วยวิธีฮาอูสเตอร์ฟอาร์ทแมพด้วย รายละเอียดของกระบวนการทดลองในบทนี้ส่วนแรกมีเนื้อหาเกี่ยวกับการใช้ฮาอูสเตอร์ฟอาร์ทแมพในการรู้จำลายนิ้วมือ หัวข้อถัดมาอธิบายถึงการรู้จำรูปมือ และการรู้จำใบหน้ามนุษย์ นอกจากนี้ในส่วนสุดท้ายได้ทำการเปรียบเทียบผลการทดลองและวิเคราะห์การทำงานกับวิธีการอื่นอีกด้วย

4.1 การรู้จำลายนิ้วมือด้วยฮาอูสเตอร์ฟอาร์ทแมพ

ลายนิ้วมือได้รับการยอมรับและนำมาใช้ในการระบุบุคคลอย่างกว้างขวาง เช่น การใช้ลายนิ้วมือในงานสืบสวนสอบสวน เป็นต้น ทั้งนี้เนื่องจากลายนิ้วมือมีคุณสมบัติที่สำคัญคือ ลายนิ้วมือจะมีลักษณะไม่ซ้ำกันในบุคคลแต่ละคน และจะไม่มีเปลี่ยนแปลงตลอดชีวิตของบุคคลนั้น [7] ซึ่งลายนิ้วมือจะประกอบไปด้วยเส้นร่อง (furrow) และเส้นนูน (ridge) สลับกันไป นอกจากนี้แล้วลายนิ้วมือยังสามารถจำแนกเป็นหมวดหมู่หลักๆ ได้อีก 5 ประเภทคือ ลายนิ้วมือแบบก้นหอย (whorl) แบบมัดหวายบิดซ้าย (left loop) แบบมัดหวายบิดขวา (right loop) แบบโค้งแบนราบ (plain arch) และแบบโค้งกระโจม (tented arch) [8] ดังแสดงไว้ในรูปที่ 4.1 – 4.5 ระบบการรู้จำลายนิ้วมือส่วนใหญ่จะอาศัยลักษณะพิเศษ (minutiae) ที่อยู่ในลายนิ้วมือ 2 ประเภทคือ จุดปลาย (ending) และจุดแยก (bifurcation) ดังรูปที่ 4.6 โดยการสกัดเอาลักษณะพิเศษทั้งสองนี้ออกมาซึ่งจะใช้เวลาในการคำนวณมากและมีความซับซ้อน แต่วิธีการรู้จำด้วยฮาอูสเตอร์ฟอาร์ทแมพจะเป็นการรู้จำภาพลายนิ้วมือทั้งภาพโดยไม่จำเป็นต้องสกัดเอาลักษณะพิเศษแต่ละตำแหน่งออกมา

4.1.1 กระบวนการเตรียมข้อมูล

การเตรียมข้อมูลภาพลายนิ้วมือก่อนที่จะนำไปทำการรู้จำในกระบวนการรู้จำนี้จำเป็นต้องดำเนินการ 2 ขั้นตอนหลัก คือ การหาตำแหน่งอ้างอิงซึ่งจะใช้เป็นจุดศูนย์กลางของภาพ และแปลงภาพจากภาพระดับสีเทาไปเป็นภาพแบบขาวดำหรือภาพแบบไบนารี นอกจากนี้แล้วยังมีขั้นตอนย่อยๆ อีกหลายขั้นตอน โดยขั้นตอนทั้งหมดมีดังต่อไปนี้ คือ

4.1.1.1 การหาจุดอ้างอิง

เนื่องจากรูปลายนิ้วมือแต่ละรูปมีการจัดวางอยู่ในตำแหน่งที่ไม่เหมือนกัน จึงจำเป็นต้องหาตำแหน่งอ้างอิงของรูปลายนิ้วมือแต่ละรูป เพื่อปรับให้รูปลายนิ้วมืออยู่ในตำแหน่งที่เหมาะสมเหมือนกันทุกรูป (นำไปใช้เป็นจุดศูนย์กลางของรูป) โดยในการหาตำแหน่งอ้างอิงนี้กระทำได้โดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใช้พีเจอร์หรือลักษณะพิเศษที่มีอยู่ในลายนิ้วมือ คือ จุดซิงกูลาร์ (singular points : SPs) ซึ่งประกอบด้วยจุด 2 ประเภท ได้แก่ จุดคอร์ (core) และจุดเดลตา (delta) โดยจุดคอร์ คือ จุดที่เป็นส่วนบนสุดของเส้นโค้งที่อยู่ในชุด จุดเดลตา คือ จุดศูนย์กลางของพื้นที่สามเหลี่ยมซึ่งมีทิศทางของเส้นในสามทิศทางมาบรรจบกัน [9].



รูปที่ 4.1 ภาพลายนิ้วมือแบบโค้ง



รูปที่ 4.2 ภาพลายนิ้วมือแบบโค้งกระจัด



รูปที่ 4.3 ภาพลายนิ้วมือแบบมัดหวายปิดซ้าย

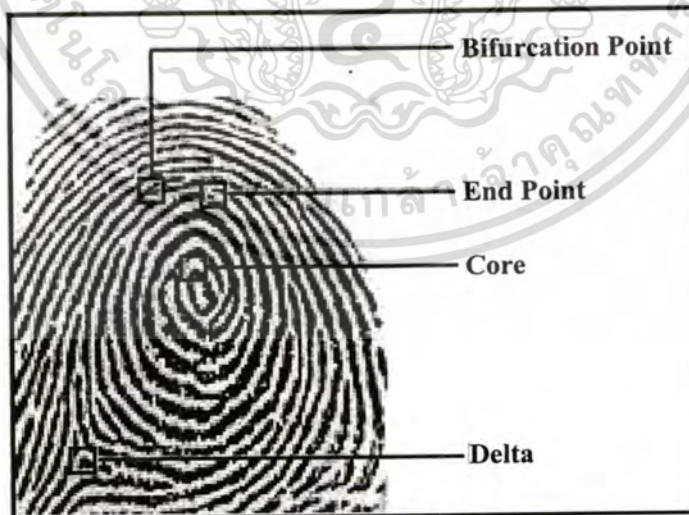
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 ภาพลายนิ้วมือแบบมัดหวยปิดขวา



รูปที่ 4.5 ภาพลายนิ้วมือแบบก้นหอย



□ Fingerprint feature

รูปที่ 4.6 แสดงลักษณะสำคัญพิเศษบนลายนิ้วมือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 ตัวอย่างรูปภาพลายนิ้วมือจากฐานข้อมูล FINGDB

งานวิจัยนี้ใช้วิธีการหาจุดซิงกูลาร์ตามแนวทางใน [10] ซึ่งหาตำแหน่งของจุดซิงกูลาร์โดยอาศัยการเปลี่ยนแปลงของเครื่องหมายในรูปภาพเชิงทิศทาง (directional image) วิธีการใน [10] สามารถหาจุดซิงกูลาร์ได้อย่างถูกต้องและใช้เวลาในการคำนวณน้อย ซึ่งใน [10] ได้นำเอากระบวนการทำงานนี้ไปทดสอบกับฐานข้อมูลลายนิ้วมือ FINGDB [11] แล้วให้ผลลัพธ์ที่ดี และในงานวิจัยนี้ก็ได้นำเอาฐานข้อมูล FINGDB มาใช้ในการทดสอบเช่นกันซึ่งฐานข้อมูล FINGDB นี้ประกอบไปด้วยภาพลายนิ้วมือจากนิ้วหัวแม่มือของคนจำนวน 21 คน คนละ 8 รูป จำนวนรูปทั้งหมดมี 168 รูปซึ่งเก็บภาพโดยใช้เครื่องสแกนที่มีคุณภาพดี แต่ละรูปเป็นรูป 8 บิต ขนาด 256 x 256 พิกเซล (pixel)

รายละเอียดของวิธีการใน [10] ประกอบไปด้วยขั้นตอนดังจะได้กล่าวต่อไปนี้ ขั้นตอนแรกคือการหาทิศทางของเส้นนูนเฉลี่ย (average ridge orientation) ของภาพลายนิ้วมือแบบระดับสีเทา (gray scale) โดยการคำนวณหาเกรเดียนท์ (gradient) ของภาพ ซึ่งสามารถหาค่าประมาณได้จากสูตรต่อไปนี้

$$DX_{ij} = A_{i,j+1} - A_{i,j-1} \quad (4.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$DY_{ij} = A_{i+1,j} - A_{i-1,j} \quad (4.2)$$

เมื่อ DX คือ เมทริกซ์ซึ่งเก็บค่าอนุพันธ์ตามแนวนอน (horizontal partial derivative) ของรูปในตำแหน่ง (i,j) DY คือ เมทริกซ์ซึ่งเก็บค่าอนุพันธ์ตามแนวตั้ง (vertical partial derivative) ของรูปในตำแหน่ง (i,j) และ A คือ ภาพลายนิ้วมือเริ่มต้นที่เป็นภาพระดับสีเทา

ขั้นตอนต่อมาคือการคำนวณหาเมทริกซ์ทิศทาง (orientation matrices) ของรูปภาพซึ่งในแนวทางนี้จะพิจารณาเพียงการเปลี่ยนแปลงเครื่องหมายของเมทริกซ์ทิศทางทั้งสองนี้เท่านั้น ดังนั้นจึงสามารถสรุปสูตรการคำนวณให้อยู่ในรูปอย่างง่ายได้ดังนี้ คือ

$$PX_{ij} = |DY_{ij}| - |DX_{ij}| \quad (4.3)$$

$$PY_{ij} = DY_{ij}DX_{ij} \quad (4.4)$$

จากสูตรข้างต้นค่า PX_{ij} และ PY_{ij} จะมีค่าเป็นศูนย์ทั้งคู่ก็ต่อเมื่อทั้ง DX_{ij} และ DY_{ij} มีค่าเป็นศูนย์ ซึ่งแสดงให้เห็นว่ามุมเชิงทิศทาง (orientation angle) ณ จุดที่ PX_{ij} และ PY_{ij} มีค่าเป็นศูนย์ ทั้งคู่ นั้นมีความคลุมเครือ (ไม่สามารถระบุทิศทางของจุดนั้นได้) จากคุณสมบัติของจุดซิงกูลาร์ทำให้สรุปได้ว่าจุดที่ทั้ง PX_{ij} และ PY_{ij} มีค่าเป็นศูนย์ก็คือจุดซิงกูลาร์นั่นเอง

หลังจากคำนวณหาเมทริกซ์ทิศทางทั้งสองได้แล้วขั้นตอนต่อมา คือ การ blur เมทริกซ์ทั้งสองนี้ โดยใช้ Gaussian low-pass filter เพื่อลดความไม่สม่ำเสมอของเส้นนูน (ridge irregularities) และจุดรบกวน (noise) ที่มีอยู่ในภาพ ซึ่งขนาดของ filter box จะต้องมีความเหมาะสมโดยสามารถหาขนาดที่เหมาะสมได้จากการทดลอง และนอกจากนี้แล้วค่าซิกมา (sigma) ของ Gaussian filter ก็ควรจะมีความเล็กที่สุดเท่าที่จะเป็นไปได้ แต่อย่างไรก็ตามการทำ blurring มากเกินไปก็จะทำให้ประสิทธิภาพในการหาจุดซิงกูลาร์ลดลงได้ด้วยเช่นกัน

ในการหาจุดซิงกูลาร์ขั้นตอนต่อมาคือการหาจุดตัดซึ่งเป็นจุดที่ทั้ง PX_{ij} และ PY_{ij} มีค่าเป็นศูนย์ แต่การเปลี่ยนแปลงเครื่องหมายในรูปภาพนั้นอาจมีความไม่สม่ำเสมอ ดังนั้นเพื่อแก้ปัญหาดังกล่าวนี้ในการคำนวณจะคำนวณการเปลี่ยนแปลงของเครื่องหมายเฉพาะจุดกลางของกรอบพื้นที่ซึ่งมีขนาด 3×3 เท่านั้น โดยในการคำนวณจะคำนวณทุกๆ 2 จุดของทั้งสองเมทริกซ์ พิจารณาว่าจุดนั้นมีการเปลี่ยนแปลงของเครื่องหมายหรือไม่โดยดูจากค่าที่น้อยที่สุดและค่าที่มากที่สุดของบล็อกขนาด 3×3 ดังได้กล่าวข้างต้น ถ้าค่าทั้งสองมีเครื่องหมายไม่เหมือนกันแสดงว่าจุดนั้นมีการเปลี่ยนแปลงของเครื่องหมาย และถ้าทั้งสองเมทริกซ์มีการเปลี่ยนแปลงของเครื่องหมาย ณ จุดนั้นทั้งคู่ก็แสดงว่าจุดนั้นเป็นจุดซิงกูลาร์ ในวิธีการนี้อาจให้จุดซิงกูลาร์หลายจุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในบล็อกที่อยู่ติดกันซึ่งสามารถกำจัดจุดเหล่านี้ทิ้งได้โดยถ้าพบจุดซิงกูลาร์ที่อยู่ใกล้ๆ กันก่อนแล้วก็ให้กำจัดจุดที่พบใหม่ทิ้งเสีย

ชนิดของจุดซิงกูลาร์ที่จุด (i,j) สามารถพิจารณาได้จากทิศทางของการเปลี่ยนแปลงเครื่องหมาย หรือก็คือเกรเดียนท์ของ PX และเกรเดียนท์ของ PY ที่จุด (i,j) นั้นเอง ถ้ามุมระหว่างเกรเดียนท์ทั้งสองนี้มีค่า 90 องศา แสดงว่าจุดซิงกูลาร์นั้นคือจุดเดลตา และถ้ามีค่า 270 องศา แสดงว่าจุดซิงกูลาร์นั้นเป็นจุดคอร์ มุมนี้กำหนดให้ทำมุมที่ทวนเข็มนาฬิกากับเกรเดียนท์ของ PX_{ij} ดังนั้นจึงไม่จำเป็นต้องคำนวณหาค่าของมุมที่แน่นอน ใช้แค่เพียงการพิจารณาว่ามุมนั้นมากกว่าหรือน้อยกว่า 180 องศา ก็เพียงพอโดยการคำนวณค่าจากค่าดีเทอร์มิแนนท์ (determinant) ของเมทริกซ์ที่เกิดจากเกรเดียนท์เวคเตอร์ (gradient vector) ทั้งสองซึ่งทำได้ดังนี้

$$\begin{vmatrix} \frac{\partial PX_{ij}}{\partial x} & \frac{\partial PY_{ij}}{\partial x} \\ \frac{\partial PX_{ij}}{\partial y} & \frac{\partial PY_{ij}}{\partial y} \end{vmatrix} = \frac{\partial PX_{ij}}{\partial x} \frac{\partial PY_{ij}}{\partial y} - \frac{\partial PY_{ij}}{\partial x} \frac{\partial PX_{ij}}{\partial y} \quad (4.5)$$

ถ้าดีเทอร์มิแนนท์มากกว่าศูนย์แสดงว่าจุดซิงกูลาร์นั้นเป็นจุดเดลตาและถ้าไม่ใช่แสดงว่าเป็นจุดคอร์

จุดซิงกูลาร์ที่ได้จากกระบวนการนี้อาจมีความผิดพลาดเนื่องจากบริเวณขอบของภาพลายนิ้วมือที่ได้รับการสแกนเข้ามาจะมีจุดรบกวนอยู่เป็นจำนวนมากซึ่งจุดรบกวนเหล่านี้เกิดจากนิ้วมือที่ได้รับการสแกนมีขนาดไม่ครอบคลุมทั้งบริเวณของการสแกน ปัญหานี้สามารถแก้ไขได้โดยการหารูปแบบของเวคเตอร์ทิศทาง (orientation vectors) ของบริเวณที่เป็นขอบของภาพซึ่งบริเวณนี้จะมีค่าเวคเตอร์ทิศทางน้อยกว่าบริเวณที่เป็นลายนิ้วมือจริงๆ ดังนั้นถ้าค่าสัมบูรณ์ของเวคเตอร์ทิศทางนั้นมีค่าน้อยกว่าค่า threshold ที่กำหนดแสดงว่าจุดนั้นไม่ใช่จุดซิงกูลาร์ที่แท้จริง แต่อย่างไรก็ตามอย่าลืมว่าที่จุดใกล้กับจุดซิงกูลาร์ค่าของเวคเตอร์ทิศทางจะมีค่าน้อยลงด้วยเช่นกันดังนั้นในการเปรียบเทียบค่าสัมบูรณ์ของเวคเตอร์ทิศทางกับค่า threshold นั้นจะต้องพิจารณาในจุดที่ห่างออกไปจากจุดซิงกูลาร์เล็กน้อย (ควรใช้จุดที่ห่างออกไปจากจุดซิงกูลาร์มากกว่า 2 จุด, $a > 2$) และในการพิจารณาเวคเตอร์ทิศทางจะพิจารณาเฉพาะ PY เท่านั้น นอกจากนี้การกำหนดค่า threshold ควรทำการทดลองเพื่อพิจารณาค่าที่เหมาะสม

Pseudocode ของกระบวนการนี้มีดังต่อไปนี้

```

n      = image size (n x n), assumed that the image is square
image  = the original grayscale fingerprint image as a (n x n) matrix
a      = threshold calculation point
w      = window size (w x w) for the filterbox
 $\sigma$   = sigma parameter of Gaussian filterbox
T      = threshold value for the PY term
PT     = Proximity threshold

for i, j := 2 to n-1 step 2
    dx := image(i,j+1) - image(i,j-1)
    dy := image(i+1,j) - image(i-1,j)
    PY(i/2,j/2) := dy*dx
    PX(i/2,j/2) := |dy| - |dx|
end for
PY := filter(PY, w,  $\sigma$ )
PX := filter(PX, w,  $\sigma$ )
For i, j := 2 to n/2-a step 2
    If |PY(i+a,j+a)| > T
        A := PX(i-1:i+1,j-1:j+1)
        B := PY(i-1:i+1,j-1:j+1)
        if sign(min(A)*max(A)) = -1 AND sign(min(B)*max(B)) = -1
            if "distance to previous found singular point" > PT
                if [A(2,1)-A(2,3)]*[B(1,2)-B(3,2)] - [B(2,1)-B(2,3)]*[A(1,2)-A(3,2)] > 0
                    "delta found at (2i, 2j)"
                else
                    "core found at (2i, 2j)"
                end if
            end if
        end if
    end if
end for
end for

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าคงที่ต่างๆ สำหรับวิธีการใน [10] ที่เหมาะสมกับฐานข้อมูล FINGDB [11] ได้รับการกำหนดไว้ดังนี้ คือ $n = 256$, $a = 5$, $w = 20$, $\sigma = 2.5$, $T = 150$, $PT = 15$

4.1.1.2 การขจัดจุดซิงกูลาร์ที่ไม่ถูกต้องบริเวณพื้นหลัง

จากการดำเนินการตามกระบวนการใน [10] ดังกล่าวข้างต้นยังมีจุดซิงกูลาร์ที่ไม่ถูกต้องเกิดขึ้นซึ่งจะเกิดบริเวณพื้นหลังของภาพที่ไม่ใช่ส่วนของลายนิ้วมืองดแสดงไว้ในรูปที่ 4.8 (a) จึงจำเป็นต้องทำการกำจัดจุดซิงกูลาร์ที่ผิดพลาดนี้โดยการคำนวณหาค่าเบี่ยงเบนมาตรฐาน (SD) ของจุดที่อยู่ใกล้กับจุดนั้นตามแนวแกน x และแนวแกน y โดยอาศัยสมมุติฐานที่ว่าค่าเบี่ยงเบนมาตรฐานบริเวณที่เป็นลายนิ้วมืองดมีค่ามากกว่าบริเวณที่เป็นพื้นหลังซึ่งมีลวดลายน้อยกว่า การดำเนินการในงานวิจัยนี้ได้พิจารณาจุดตามแนวแกน x จำนวน 20 จุดและแนวแกน y จำนวน 20 จุด โดยให้จุดที่ต้องการตรวจสอบอยู่ตรงกลาง แล้วคำนวณหาค่าเบี่ยงเบนมาตรฐานของจุดตามแกน x และจุดตามแกน y โดยค่าเบี่ยงเบนมาตรฐานทั้งสองจะต้องมีค่ามากกว่าค่า threshold ที่กำหนด ในที่นี้กำหนดให้มีค่าเท่ากับ 30 ถ้ามีค่าน้อยกว่าค่านี้แสดงว่าเป็นจุดที่อยู่ในบริเวณพื้นหลังไม่ใช่จุดซิงกูลาร์ที่แท้จริง



(a)

(b)

รูปที่ 4.8 (a) ภาพลายนิ้วมือและตำแหน่งของจุดซิงกูลาร์ซึ่งยังคงมีจุดซิงกูลาร์ที่ผิดพลาดอยู่ในบริเวณพื้นหลัง (b) ภาพลายนิ้วมือที่ขจัดจุดซิงกูลาร์ที่ผิดพลาดออกไปแล้ว

4.1.1.3 การแปลงภาพลายนิ้วมือแบบระดับสีเทาไปเป็นภาพแบบไบนารี

จากภาพลายนิ้วมือที่เป็นภาพระดับสีเทาในกระบวนการนี้จะทำการเปลี่ยนภาพดังกล่าวไปเป็นรูปภาพแบบไบนารีซึ่งจะมีค่าของแต่ละพิกเซลเป็น 0 หรือ 1 เท่านั้น ในงานวิจัยนี้ใช้วิธีการหาค่าเฉลี่ยแต่ละบริเวณมาเป็นค่า threshold หรือเรียกว่าวิธี Regional Average Thresholding (RAT) ที่นำเสนอใน [12] เพื่อแปลงภาพระดับสีเทาเป็นภาพแบบไบนารีโดยมีแนวคิดจากปัญหาเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการแปลงภาพซึ่งใช้ค่า threshold เพียงค่าเดียวทั้งภาพอาจจะทำให้พีเจอร์บางอย่างสูญหายไป ซึ่งแนวทางที่ใช้ค่า threshold ค่าเดียวกันทั้งภาพหรือเรียกว่าวิธี General Thresholding (GT) สามารถแสดงในรูปสมการได้ดังนี้

$$P(i, j) = \begin{cases} 1, & \text{if } I(i, j) > T \\ 0, & \text{if } I(i, j) \leq T \end{cases} \quad (4.6)$$

เมื่อ $I(i, j)$ คือ รูปภาพเริ่มต้นที่เป็นภาพระดับสีเทา, $P(i, j)$ คือ ภาพผลลัพธ์ที่เป็นภาพแบบไบนารี, T คือระดับของ threshold, และ $i = 0, \dots, N$ $j = 0, \dots, M$ โดยภาพมีขนาด $N \times M$



รูปที่ 4.9 (a) ภาพลายนิ้วมือแบบไบนารีที่ได้จากวิธีการ RAT (b) ภาพลายนิ้วมือแบบไบนารีที่ได้จากวิธีการ GT

วิธีการ RAT จะแบ่งภาพเริ่มต้นออกเป็นส่วนย่อยๆ ซึ่งในที่นี้แบ่งออกเป็นหน้าต่างขนาด 16×16 พิกเซล และหาค่า threshold ของแต่ละหน้าต่างโดยใช้ค่าเฉลี่ยของระดับค่าของสีเทาในแต่ละหน้าต่าง ซึ่งค่าเฉลี่ยของระดับสีเทาสามารถคำนวณได้ดังสมการดังต่อไปนี้

$$T = \frac{1}{N^2} \sum_{i=0}^N \sum_{j=0}^N I(i, j) \quad (4.7)$$

วิธีการ RAT เป็นวิธีที่สามารถแก้ไขปัญหาที่เกิดกับวิธี GT ได้ และทำให้คงรายละเอียดของภาพลายนิ้วมือไว้ได้ดีกว่าวิธี GT ดังแสดงไว้ในรูปที่ 4.9 เนื่องจากภาพลายนิ้วมือในแต่ละบริเวณเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะมีความเข้มของแสงไม่เท่ากันการใช้ค่า threshold ค่าเดียวตามแบบวิธี GT จะทำให้ข้อมูลลายเส้นของภาพลายนิ้วมือบางส่วนสูญหายไป

4.1.1.4 การเลือกรูปลายนิ้วมือที่เหมาะสมและปรับขนาดของรูป

ขั้นตอนนี้จะใช้จุดซิงกูลาร์ที่ได้จากระบวนการก่อนหน้านี้เป็นจุดอ้างอิง กรณีที่มีจุดซิงกูลาร์ทั้งจุดคอร์และเดลตาจะใช้จุดคอร์เป็นจุดอ้างอิง ถ้าพบจุดคอร์หรือเดลตาเพียงอย่างเดียวหนึ่งก็จะใช้จุดนั้นเป็นจุดอ้างอิง และเมื่อใช้จุดซิงกูลาร์ที่ได้รับการเลือกมาเป็นจุดศูนย์กลางของภาพจะทำการเลือกพื้นที่รอบจุดศูนย์กลางขนาด 100×100 พิกเซลเท่านั้น เนื่องจากบริเวณรอบจุดซิงกูลาร์จะเป็นบริเวณที่มีความสำคัญ (significance) มากที่สุด ภาพที่ไม่สามารถหาจุดซิงกูลาร์ได้หรือมีพื้นที่รอบจุดอ้างอิงน้อยกว่าที่กำหนดจะไม่ถูกนำมาพิจารณาตัวอย่างเช่นรูปที่ 4.10 และ 4.11 เมื่อสิ้นสุดขั้นตอนนี้ภาพในฐานข้อมูล FINGDB [11] จำนวน 168 รูปของ 21 คน จะได้ภาพที่เหมาะสมจำนวน 138 รูปจากบุคคล 20 คน มีขนาดของภาพ 100×100 พิกเซล โดยในการทดลองช่วงการเรียนรู้จะให้รูปภาพประมาณ 80% แรกของรูปภาพลายนิ้วมือของแต่ละคนในการรู้จำ และรูปในส่วนที่เหลือก็จะใช้ในการทดสอบเพื่อหาเปอร์เซ็นต์ความถูกต้องในการรู้จำ



รูปที่ 4.10 ตัวอย่างภาพลายนิ้วมือที่คัดออกเนื่องจากไม่สามารถหาจุดซิงกูลาร์ที่แน่นอนได้



รูปที่ 4.11 ตัวอย่างภาพลายนิ้วมือที่คัดออกเนื่องจากพื้นที่รอบจุดอ้างอิงมีขนาดน้อยกว่า 100×100 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 แสดงภาพลายนิ้วมือจากฐานข้อมูล FINGDB ที่ใช้ในการเรียนรู้ และภาพที่ใช้

ทดสอบ

บุคคลที่	ภาพที่ใช้ในการเรียนรู้	ภาพที่ใช้ในการทดสอบ
1	1_1, 1_3, 1_4, 1_5, 1_6	1_7, 1_8
2	2_1, 2_2, 2_3, 2_4, 2_5	2_6, 2_8
3	3_1, 3_2, 3_3, 3_4, 3_5, 3_6	3_7, 3_8
4	4_1, 4_2, 4_3, 4_4, 4_5	4_6, 4_8
5	5_1, 5_2, 5_3, 5_4, 5_5, 5_6	5_7, 5_8
6	6_1, 6_2, 6_3, 6_4, 6_5, 6_6	6_7, 6_8
7	7_1, 7_4, 7_5, 7_7	7_8
9	9_1, 9_2, 9_3, 9_4	9_5, 9_7
10	10_1, 10_2, 10_3, 10_4, 10_5, 10_6	10_7, 10_8
11	11_1, 11_2, 11_3, 11_4, 11_5, 11_6	11_7, 11_8
12	12_1, 12_2, 12_3, 12_4	12_6, 12_8
13	13_1, 13_2, 13_3, 13_4, 13_5, 13_6	13_7, 13_8
14	14_1, 14_2, 14_3, 14_4, 14_6	14_7, 14_8
15	15_2, 15_3, 15_4, 15_5, 15_6	15_7, 15_8
16	16_1, 16_2, 16_3, 16_4, 16_5, 16_6	16_7, 16_8
17	17_1, 17_2, 17_3, 17_5	17_7, 17_8
18	18_1, 18_2, 18_3, 18_4, 18_5, 18_6	18_7, 18_8
19	19_1, 19_4, 19_5, 19_6	19_7, 19_8
20	20_1, 20_2, 20_3, 20_4	20_7, 20_8
21	21_1, 21_4, 21_7	21_8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

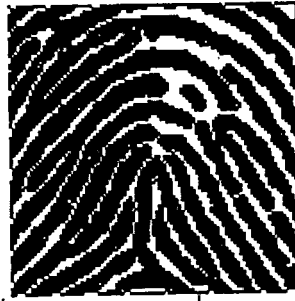
ตารางที่ 4.3 แสดงจำนวนรูปภาพในฐานข้อมูล FINGDB และจำนวนรูปภาพที่นำมาใช้ในการทดลอง

บุคคลที่	จำนวนภาพต่อคน	จำนวนภาพที่ใช้ในการเรียนรู้	จำนวนภาพที่ใช้ในการทดสอบ
1	7	5	2
2	7	5	2
3	8	6	2
4	7	5	2
5	8	6	2
6	8	6	2
7	5	4	1
9	6	4	2
10	8	6	2
11	8	6	2
12	6	4	2
13	8	6	2
14	7	5	2
15	7	5	2
16	8	6	2
17	6	4	2
18	8	6	2
19	6	4	2
20	6	4	2
21	4	3	1
รวม	138	100	38

4.1.1.5 การทำ thinning

การทำ thinning มีจุดประสงค์เพื่อให้ลายเส้นมีความเรียบมากขึ้นจัดส่วนที่ยื่นเกินออกมาจากตัวเส้น โดยมีการดำเนินการตามหลักการใน [13] ตัวอย่างภาพลายนิ้วมือที่ได้รับการทำ thinning และผ่านกระบวนการเตรียมข้อมูลก่อนหน้านี้อันได้แสดงไว้ในรูปที่ 4.12

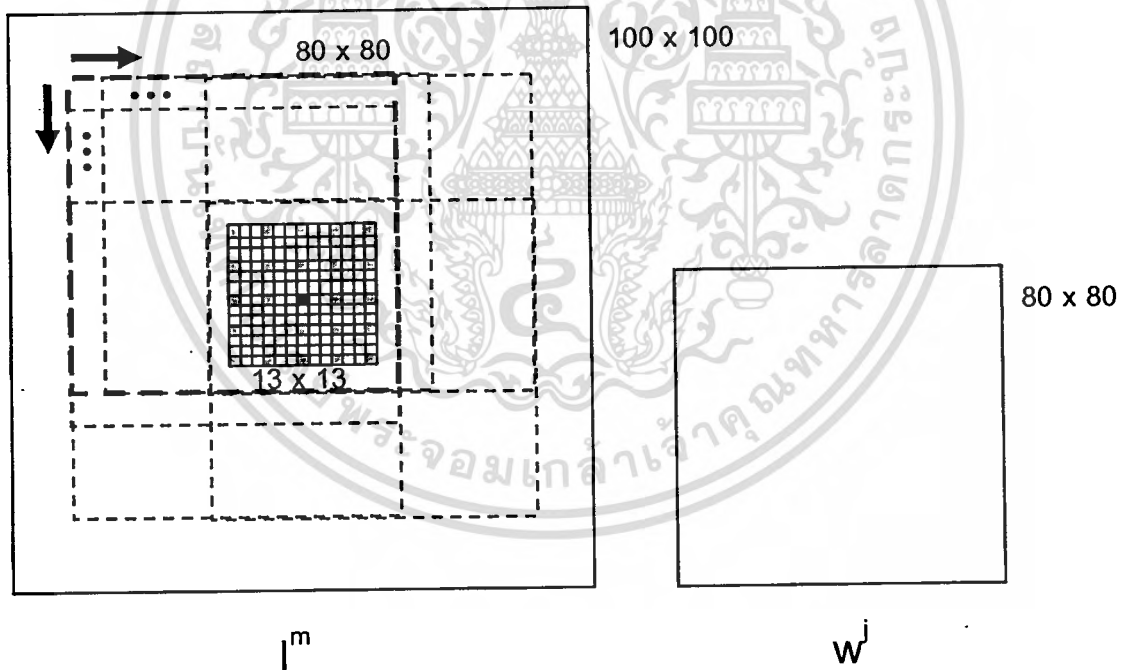
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 ภาพลายนิ้วมือแบบไบนารีที่ผ่านกระบวนการ thinning แล้ว

4.1.2 กระบวนการเรียนรู้ภาพลายนิ้วมือ

ในกระบวนการเรียนรู้จะทำการเรียนรู้แบบ supervised ตามวิธีการของฮาสตอร์ฟ อาร์ทแมพดังได้กล่าวข้างต้น แต่เนื่องจากภาพลายนิ้วมือที่ได้จากกระบวนการเตรียมข้อมูลซึ่งใช้วิธีหาซิงกูลาร์ตามวิธีการใน [10] แล้วนำจุดซิงกูลาร์ที่ได้มาเป็นจุดศูนย์กลางนั้น จุดซิงกูลาร์ที่ได้ อาจมีตำแหน่งคลาดเคลื่อนไปบ้างเล็กน้อย ดังนั้นในการคำนวณหาค่าฮาสตอร์ฟดิสแทนซ์หรือ ขอยซ์ฟังก์ชันในฮาสตอร์ฟอาร์ทแมพจึงจำเป็นต้องคำนวณค่าหลายครั้ง



รูปที่ 4.13 แสดงตำแหน่งของจุดศูนย์กลางของภาพ (จุดสีดำ) และจุดศูนย์กลางใหม่ที่เปลี่ยนตำแหน่งไปในการคำนวณแต่ละครั้ง (จุดสีเทา) เมื่อขยายจุดที่จะพิจารณาออกไปข้างละ 6 พิกเซล และทำการย้ายตำแหน่งไปที่ละ 3 พิกเซล โดยกรอบเส้นประสีเข้มขนาด 80x80 พิกเซล คือ ส่วนของรูปแบบอินพุทซึ่งจะเลื่อนตำแหน่งของจุดศูนย์กลางด้วยตำแหน่งต่างๆ เพื่อที่จะทำการพิจารณาเปรียบเทียบกับค่าน้ำหนัก w^j

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การคำนวณค่าฮาลูสตอร์ฟดิสแทนซ์แต่ละครั้งจะดำเนินการโดยใช้วิธีการเปลี่ยนบริเวณที่จะพิจารณาของรูปแบบอินพุทด้วยตำแหน่งจุดศูนย์กลางต่างๆ แล้วเปรียบเทียบกับรูปแบบอ้างอิง (reference pattern) หรือคือค่าน้ำหนักของแต่ละโหนด j การขยับตำแหน่งของจุดศูนย์กลางของรูปแบบอินพุทจะกินบริเวณขยายออกจากจุดศูนย์กลางของรูปแบบอินพุทเดิมออกไปข้างละ 6 พิกเซล โดยทำการย้ายจุดศูนย์กลางของรูปแบบอินพุทไปที่ละ 3 พิกเซลแล้วทำการเปรียบเทียบรูปแบบอินพุทที่ใช้จุดศูนย์กลางใหม่นี้กับรูปแบบอ้างอิง (ใช้จุดศูนย์กลางเดิม) ซึ่งการเปรียบเทียบจะใช้ส่วนของรูปแบบอินพุทและรูปแบบอ้างอิงเพียง 80×80 พิกเซลรอบจุดศูนย์กลางเท่านั้นเพื่อใช้ในการหาค่าฮาลูสตอร์ฟดิสแทนซ์ดังแสดงไว้ในรูปที่ 4.13 และจะทำการเปรียบเทียบซ้ำโดยเปลี่ยนจุดศูนย์กลางในรูปแบบอินพุทไปจนครบบริเวณที่กำหนด ซึ่งทำให้ต้องคำนวณค่าฮาลูสตอร์ฟดิสแทนซ์จำนวนทั้งหมด 25 ครั้ง จากนั้นเลือกค่าฮาลูสตอร์ฟดิสแทนซ์ที่น้อยที่สุดมาเป็นค่าชอยซ์ฟังก์ชันหรือค่าระดับความแตกต่างระหว่างรูปแบบอินพุทนี้กับค่าน้ำหนักของโหนด j และการปรับค่าน้ำหนักในกระบวนการเรียนรู้จะพิจารณาเฉพาะส่วนของรูปแบบอินพุทตรงบริเวณซึ่งทำให้รูปแบบอินพุทและรูปแบบอ้างอิงมีความเหมือนกันมากที่สุดเท่านั้น (ฮาลูสตอร์ฟดิสแทนซ์มีค่าน้อยที่สุด) โดยจะทำการตัดรูปแบบอินพุทมาเพียง 80×80 พิกเซลรอบจุดศูนย์กลางที่ทำให้มีความเหมือนกันมากที่สุด แล้วนำรูปแบบอินพุทส่วนนี้มาทำการเรียนรู้กับรูปแบบอ้างอิง (ปรับค่าน้ำหนัก) ต่อไป

4.1.3 กระบวนการทดสอบและผลลัพธ์ในการรู้จำลายนิ้วมือ

ในการหาค่าฮาลูสตอร์ฟดิสแทนซ์จะอาศัยการขยับตำแหน่งจุดศูนย์กลางเพื่อหาบริเวณซึ่งทำให้รูปแบบอ้างอิงและรูปแบบอินพุทแมทช์กันมากที่สุดเช่นเดียวกันกับกระบวนการเรียนรู้ในหัวข้อ 4.1.2 แล้วเลือกค่าฮาลูสตอร์ฟดิสแทนซ์ที่น้อยที่สุดเป็นค่าชอยซ์ฟังก์ชันของแต่ละโหนดเพื่อพิจารณาหาโหนดที่ชนะ และหาค่าคำตอบของระบบได้จากคำตอบในชั้นเอาต์พุทที่เชื่อมโยงกับโหนดที่ชนะในชั้นแคทาคอรี จากการทดลองได้ทำการปรับค่า S ในการเรียนรู้เป็นค่าต่างๆ และปรับค่า p ให้มีค่าต่างๆ กัน เพื่อหาค่า S และ p ซึ่งทำให้ได้ผลความถูกต้องในการรู้จำที่ดีที่สุด ซึ่งผลการทดลองเมื่อ $p = 0.80$, ค่า $S = 2$, การเปลี่ยนตำแหน่งครอบคลุมบริเวณขยายออกจากจุดศูนย์กลางออกไปข้างละ 6 พิกเซล โดยทำการย้ายจุดศูนย์กลางของรูปแบบอินพุทไปที่ละ 3 พิกเซล ปรากฏว่าได้ผลลัพธ์ในการรู้จำ 100% ดังตารางที่ 4.4

ตารางที่ 4.4 แสดงผลลัพธ์ในการรู้จำลายนิ้วมือโดย $p = 0.80$, ค่า $S = 2$, การเปลี่ยนตำแหน่งจะครอบคลุมบริเวณขยายออกจากจุดศูนย์กลางเดิมออกไปข้างละ 6 พิกเซล โดยทำการย้ายจุดศูนย์กลางของรูปแบบอินพุทไปที่ละ 3 พิกเซล

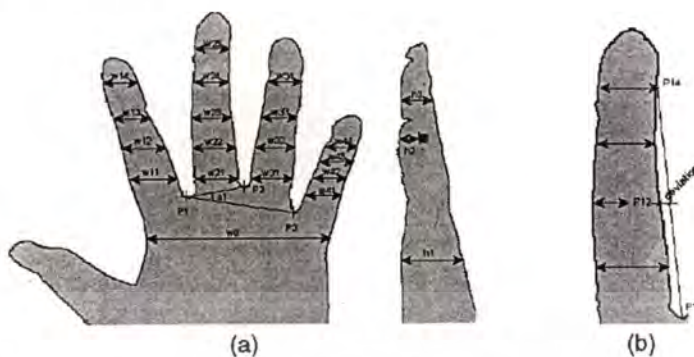
ลายนิ้วมือของบุคคลที่	จำนวน	จำนวนที่ตอบถูกต้อง
1	2	2
2	2	2
3	2	2
4	2	2
5	2	2
6	2	2
7	1	1
9	2	2
10	2	2
11	2	2
12	2	2
13	2	2
14	2	2
15	2	2
16	2	2
17	2	2
18	2	2
19	2	2
20	2	2
21	1	1
รวม	38	38
เปอร์เซ็นต์ความถูกต้อง	-	100

4.2 การรู้จำรูปร่างของมือ

การรู้จำภาพมือโดยส่วนใหญ่จะอาศัยการสกัดพีเจอร์จากภาพมือ [14, 15] ซึ่งมีความยุ่งยากและใช้เวลาในการคำนวณมากดังแสดงในรูปที่ 4.14 ในงานวิจัยนี้ได้ทำการรู้จำภาพมือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยอาศัยเฉพาะเส้นขอบของภาพมือเท่านั้นซึ่งเป็นการช่วยลดระยะเวลาในการคำนวณไปได้มาก และสามารถให้ผลลัพธ์ในการรู้จำที่ดี



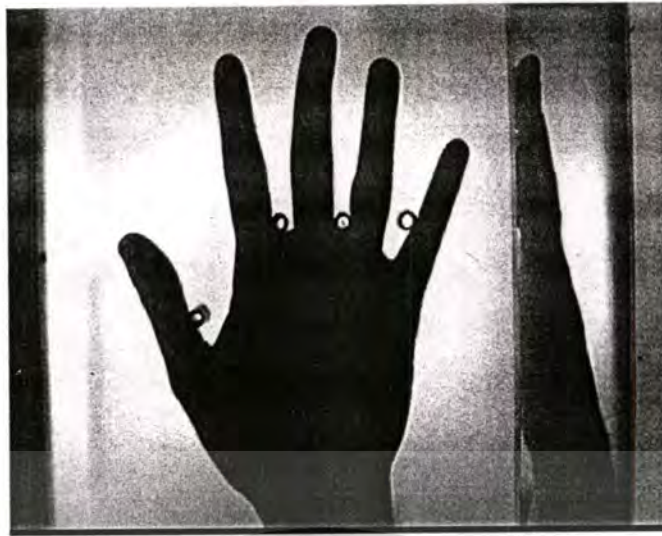
รูปที่ 4.14 แสดงการวัดค่าพีเจอร์ต่างๆ ที่ใช้ในการรู้จำภาพมือโดยทั่วไป

ภาพมือที่ใช้ในงานวิจัยนี้เป็นภาพมือที่นำมาจากฐานข้อมูลของ Biometrics Systems Lab., University of Bologna, Cesena-Italy [11] ประกอบไปด้วยภาพมือของบุคคล 8 คน มีทั้งหมด 66 ภาพ คือคนละประมาณ 8 ภาพ ซึ่งในการทดลองนี้ได้นำภาพทั้งหมดมาใช้ ดังตารางที่ 4.5

ตารางที่ 4.5 แสดงจำนวนภาพมือที่ใช้ในการทดลอง

ภาพมือของบุคคลที่	จำนวนภาพต่อคน	จำนวนภาพที่ใช้ในการเรียนรู้	จำนวนภาพที่ใช้ในการทดสอบ
1	9	7	2
3	8	6	2
4	8	6	2
5	8	6	2
6	9	7	2
7	7	5	2
8	9	7	2
9	8	6	2
รวม	66	50	16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.15 ตัวอย่างภาพมือที่ใช้ในการทดลอง

จากภาพมือในรูปที่ 4.15 นั้นในการทดลองนี้จะใช้เพียงภาพด้านบนของฝ่ามือขนาด 450x400 พิกเซลเท่านั้นในการรู้จำ โดยจะไม่นำเอาภาพฝ่ามือด้านข้างมาพิจารณาด้วย

4.2.1 การเตรียมข้อมูลภาพมือ

4.2.1.1 การแปลงจากภาพมือที่เป็นระดับสีเทาไปเป็นภาพแบบไบนารี

จากภาพมือที่เป็นแบบระดับสีเทาจะทำการใช้วิธี General Thresholding (GT) ดังได้กล่าวไว้ในกระบวนการเตรียมข้อมูลภาพลายนิ้วมือ เนื่องจากความแตกต่างระหว่างสีของพื้นหลัง และตัวรูปภาพมีความแตกต่างกันมากจึงสามารถใช้ค่า threshold ค่าเดียวดำเนินการกับทั้งภาพได้



รูปที่ 4.16 ภาพมือแบบไบนารีที่ได้จากการใช้วิธี GT

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1.2 การหาเส้นขอบของฝ่ามือ

เมื่อได้ภาพแบบไบนารีของฝ่ามือจากนั้นจะทำการหาเส้นขอบของฝ่ามือ โดยทำการพิจารณาพิกเซลรอบๆ ใน 8 ทิศทาง ซึ่งวิธีการดำเนินการมีดังต่อไปนี้ คือ

- 1) สแกนหาพิกเซลเริ่มต้นในรูปภาพ R กำหนดให้พิกเซลเริ่มต้นนี้แทนด้วย s
- 2) ให้พิกเซลที่กำลังพิจารณาแทนด้วย c จากนั้นให้ $c = s$ และพิกเซลที่ติดกับ s ทางทิศตะวันตกแทนด้วย b
- 3) จากพิกเซลที่ติดกับ c ทั้ง 8 ทิศทาง ให้เริ่มต้นด้วยพิกเซล b โดยเรียงลำดับตามเข็มนาฬิกา และแทนค่าด้วย n_1, n_2, \dots, n_8 หลังจากนั้นหา n_i ที่เป็นลำดับที่ i แรกใน R
- 4) ให้ $c = n_i$ และ $b = n_{i-1}$
- 5) ทำซ้ำในข้อ 3) และ 4) จนกระทั่ง $c = s$



รูปที่ 4.17 ภาพมือที่ได้รับการสกัดเส้นขอบ

4.2.2 กระบวนการเรียนรู้และการทดสอบการรู้จำ

เมื่อทำการรู้จำภาพมือด้วยวิธีฮาอูสเตอร์ฟอรัทแมพด้วยการใช้ภาพมือจำนวน 50 ภาพในการเรียนรู้ และทดสอบด้วยภาพมือจำนวน 16 ภาพ ซึ่งจากภาพมือของแต่ละบุคคลภาพในส่วนแรกจะใช้ในการเรียนรู้ และ 2 ภาพสุดท้ายจะใช้ในการรู้จำ โดยกำหนดให้ $p = 1.00$ และ $S = 4$ ปรากฏว่าให้ผลความถูกต้องในการรู้จำถึง 100% ดังแสดงรายละเอียดในตารางที่ 4.6

ตารางที่ 4.6 ผลการรู้จำภาพมือ โดยกำหนดให้ $p = 1.00$, $S = 4$

ภาพมือของบุคคลที่	จำนวนภาพที่ทดสอบ	จำนวนที่ตอบได้ถูกต้อง
1	2	2
3	2	2
4	2	2
5	2	2
6	2	2
7	2	2
8	2	2
9	2	2
รวม	16	16
คิดเป็นเปอร์เซ็นต์	-	100

4.3 การรู้จำภาพใบหน้า

การระบุบุคคลด้วยใบหน้าเป็นวิธีการที่มีความน่าสนใจเพราะเป็นวิธีการที่ไม่ยุ่งยากกับผู้ใช้ และลักษณะของใบหน้าก็เป็นคุณสมบัติเฉพาะตัวของบุคคลที่โดดเด่นซึ่งโดยทั่วไปในชีวิตประจำวันคนเราก็อาศัยการจดจำใบหน้าในการระบุตัวบุคคลอยู่แล้ว เทคนิควิธีในการรู้จำใบหน้านั้นมีอยู่หลากหลายวิธีด้วยกัน [16, 17] ซึ่งงานวิจัยส่วนใหญ่ที่ผ่านมาจะใช้ลักษณะโครงสร้างของใบหน้ามาเป็นพีเจอร์ในการรู้จำ เช่น ระยะระหว่างตาทั้งสองข้าง ความกว้างของศีรษะ ระยะทางจากดวงตาถึงปาก เป็นต้น นอกจากนี้แล้วยังมีอีกหลายงานวิจัยที่ได้นำเอาวิธีการทางสถิติมาใช้ และวิธีการรู้จำใบหน้าโดยใช้โครงข่ายประสาทเทียม [18] จากผลการศึกษาทางจิตวิทยากลับพบว่ามนุษย์สามารถที่จะแยกแยะใบหน้าจากการมองผ่านเพียงครั้งเดียวและสามารถจดจำภาพวาดลายเส้นของวัตถุได้อย่างรวดเร็วและถูกต้องเช่นเดียวกันกับการถ่ายภาพ [19] ดังนั้นจึงเป็นส่วนที่แสดงให้เห็นว่าวิธีการที่ใช้เพียงภาพลายเส้นของใบหน้าในการรู้จำจึงน่าจะเป็นวิธีการที่รวดเร็วกว่า และให้ผลความถูกต้องที่ดี ซึ่งในงานวิจัยนี้ได้ทดลองนำเอาโครงข่ายประสาทเทียมที่คิดขึ้นใหม่คือฮาลูสตอร์ฟอาร์ทแมพมาใช้ในการรู้จำใบหน้า โดยวิธีการรู้จำใบหน้าในงานวิจัยนี้อาศัยโครงสร้างโดยรวมของใบหน้าซึ่งก็คือภาพลายเส้นของใบหน้าที่เป็นภาพแบบใบนารีมาใช้เพื่อเป็นรูปแบบอินพุทโดยตรง แต่วิธีการรู้จำใบหน้าแบบอื่นๆ โดยส่วนใหญ่จะใช้ภาพใบหน้าที่เป็นระดับสีเทาและใช้พีเจอร์ที่ซับซ้อนดังได้กล่าวข้างต้นซึ่งทำให้ต้องใช้เวลาในการคำนวณมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1 การเตรียมข้อมูลภาพใบหน้า

รูปภาพใบหน้าที่เป็นภาพแบบระดับสีเทาจำเป็นจะต้องเปลี่ยนให้เป็นภาพลายเส้นที่เป็นแบบไบนารีก่อนจึงจะสามารถนำไปดำเนินการรู้จำด้วยฮาลูสตอร์ฟาร์ทแมพได้ นอกจากนี้แล้ววิธีการในการหาเส้นขอบของภาพใบหน้ายังเป็นกระบวนการที่สำคัญและมีผลโดยตรงต่อประสิทธิภาพของกระบวนการรู้จำในงานวิจัยนี้ กระบวนการต่างๆ ในเตรียมข้อมูลมีดังต่อไปนี้

4.3.1.1 การหาจุดอ้างอิง

ใช้ตำแหน่งของตาเป็นจุดอ้างอิงโดยในการหาจุดอ้างอิงใช้วิธีการแบบ manual ซึ่งระบุตำแหน่งโดยใช้คนเป็นผู้ค้นหา จะทำให้ได้จุดอ้างอิง 2 จุดคือตำแหน่งของตาซ้ายและตาดขวาดังรูปที่ 4.18



รูปที่ 4.18 แสดงตำแหน่งอ้างอิงของภาพใบหน้าโดยใช้ตำแหน่งของตาและระบุตำแหน่งด้วยมือ

4.3.1.2 การหาเส้นขอบและการแปลงภาพไปเป็นภาพแบบไบนารี

การใช้ภาพลายเส้นแทนภาพใบหน้าทั้งหมดเพื่อใช้ในการรู้จำใบหน้านั้นเป็นการแก้ปัญหาการเปลี่ยนแปลงของความเร็วที่ตกกระทบ (illumination) และให้ผลดีกับวิธีการที่ใช้ฮาลูสตอร์ฟิตสแทนซ์เป็นตัววัดค่าความเหมือน ขั้นตอนแรกจะต้องแปลงภาพระดับสีเทาไปเป็นภาพลายเส้นแบบสีเทา (gray edge image) $E(x,y)$ ก่อนโดยใช้ morphological operation [20] กับภาพใบหน้า $f(x,y)$ ดังแสดงได้ดังต่อไปนี้คือ

$$(f \oplus b) - (f \ominus b) = E \quad (4.8)$$

โดยในการดำเนินการวิจัยนี้ได้ทำการ dilation ภาพใบหน้าด้วย structure element ที่มีลักษณะเป็นสี่เหลี่ยมจัตุรัสขนาด 3x3 ก่อน แล้วทำการ erosion ภาพใบหน้าด้วย structure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

element แบบเดียวกัน แล้วนำรูปที่ได้จากการ dilation ตั้งแล้วลบ (subtraction) ด้วยภาพที่ได้จากการ erosion ภาพที่ได้หลังจากการ subtraction นี้คือ ภาพลายเส้นระดับสีเทา (gray edge image) ที่ต้องการ จากนั้นใช้วิธี adaptive thresholding ใน [21] เพื่อแปลงภาพลายเส้นที่ได้ในขั้นตอนก่อนหน้านี้เป็นภาพแบบไบนารี โดยวิธีการนี้ต้องการเน้นพีเจอร์สำคัญของใบหน้าและโครงหน้าซึ่งมักจะเป็นส่วนที่มีความลึกและมีค่าความเข้มของระดับสีเทา (gray level intensities) ต่ำกว่าของส่วนอื่นของใบหน้า ดังนั้นใน [22] จึงเสนอฟังก์ชันเพื่อนำมาใช้ในการพิจารณา คือ

$$n(x, y) = \frac{E(x, y)}{f(x, y)} \quad (4.9)$$

นำค่า $n(x, y)$ ที่ได้จากสมการ(4.9) มาจัดเรียงตามลำดับจากมากไปหาน้อย และกำหนดค่า threshold เพื่อให้จุดที่มีค่า $n(x, y)$ มากที่สุดตามลำดับได้รับการเลือกขึ้นมาตามเปอร์เซ็นต์ที่ต้องการ ในงานวิจัยนี้จะเลือกจุดที่มีค่า $n(x, y)$ มากที่สุด 30 % แรกให้เป็นจุดที่มีค่าเท่ากับ 1 และจุดที่เหลือจะให้ค่าเป็น 0 ซึ่งจะทำให้ได้ภาพแบบไบนารีที่ต้องการ



รูปที่ 4.19 (a) ภาพใบหน้าที่เป็นภาพแบบสีเทา (b) ภาพใบหน้าที่เป็นภาพลายเส้นแบบระดับสีเทาโดยการใช้วิธี morphological gradient (c) ภาพแบบไบนารีโดยใช้วิธีการ adaptive thresholding

4.3.1.3 การหมุนภาพ

จากการใช้ตาทั้งสองเป็นจุดอ้างอิงจะทำให้ได้จุดอ้างอิง 2 จุด ซึ่งสามารถทำการคำนวณหามุมระหว่างจุด 2 จุดนี้ได้ โดยในการหมุนภาพตั้งอยู่บนสมมุติฐานที่ว่าตาของคนเราทั้งสองข้างควรจะวางตัวอยู่ในแนวเดียวกัน เมื่อได้มุมระหว่างตาทั้งสองแล้วทำการหมุนภาพเพื่อให้จุดอ้างอิงทั้งสองอยู่ในตำแหน่งที่ขนานกันแนวระนาบ โดยการหมุนภาพจะใช้ทฤษฎี nearest neighbor interpolation จากนั้นทำการหาจุดอ้างอิงใหม่ของภาพที่ได้รับการหมุน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

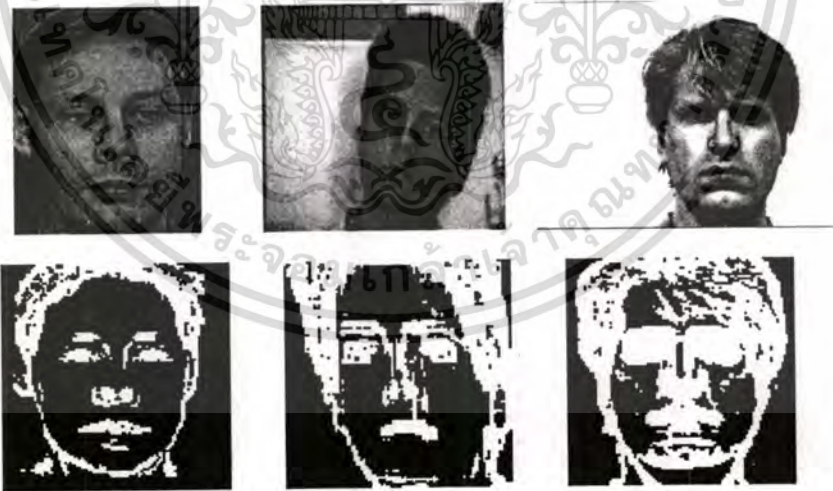
4.3.1.4 การปรับขนาดภาพ

จากจุดอ้างอิงของภาพที่ได้รับการหมุนแล้ว คำนวณหาระยะห่างระหว่างจุดอ้างอิงทั้งสอง เทียบกับระยะห่างที่ต้องการซึ่งในที่นี้ต้องการให้ทุกภาพมีระยะระหว่างตาทั้งสองข้างเท่ากับ 30 พิกเซล จากนั้นคำนวณหาอัตราส่วนของระยะระหว่างตาของภาพที่หมุนแล้วกับระยะระหว่างตาที่ต้องการ (เอา 30 ตั้งแล้วหารด้วยระยะระหว่างตาของรูปนั้น) อัตราส่วนนี้คือจำนวนเท่าที่ต้องการปรับขนาด ซึ่งในงานวิจัยนี้ในการปรับขนาดของรูปภาพได้ใช้ทฤษฎี nearest neighbor interpolation ในการดำเนินการ จากนั้นทำการหาจุดอ้างอิงใหม่ที่เกิดจากการปรับขนาดรูปภาพใบหน้า

4.3.1.5 การเลื่อนตำแหน่งอ้างอิงของภาพและเลือกบริเวณของภาพในส่วนที่ต้องการ

ในขั้นตอนนี้จะเลือกให้จุดอ้างอิงทั้งสองอยู่บริเวณที่เหมาะสม และอยู่ในตำแหน่งเดียวกันทุกรูป (ให้ตาซ้ายอยู่ในตำแหน่ง (45,60)) จากนั้นปรับให้ทุกภาพมีขนาด 100 x 100 พิกเซล โดยตัดส่วนของภาพที่เกินออกหรือภาพเมื่อปรับตำแหน่งแล้วมีขนาดน้อยกว่าที่ต้องการก็จะเติมส่วนที่ขาดไปให้เป็นที่ว่าง (มีค่าเป็นศูนย์)

ภาพใบหน้าที่ผ่านมากระบวนการเตรียมข้อมูลและจะส่งไปทำการเรียนรู้ในฮาลอสตอร์ฟ อาร์ทแมพสามารถแสดงตัวอย่างได้ดังรูปที่ 4.20



รูปที่ 4.20 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL, MIT, และ Yale แถวบนแสดงรูปภาพเดิม และแถวล่างแสดงรูปภาพแบบไบนารีที่ได้จากกระบวนการเตรียมข้อมูล

4.3.2 กระบวนการเรียนรู้และการทดสอบผลการรู้จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิจัยนี้นำเอากระบวนการหาอูสเตอร์ฟาร์ทแมพไปใช้ในการรู้จำภาพใบหน้าที่เป็นมาตรฐานจาก 3 ฐานข้อมูลด้วยกัน คือ ฐานข้อมูลภาพใบหน้า ORL [22], ฐานข้อมูลภาพใบหน้า MIT [23], และฐานข้อมูลภาพใบหน้า Yale [24]

4.3.3 การรู้จำภาพใบหน้าจากฐานข้อมูล ORL

ภาพใบหน้าในฐานข้อมูล ORL [22] ซึ่งประกอบไปด้วยภาพใบหน้าของบุคคลจำนวน 40 คน คนละ 10 ภาพ รวมทั้งสิ้น 400 ภาพ เป็นภาพแบบระดับสีเทา 8 บิต ขนาด 92x112 พิกเซล ภาพมีการให้แสงในตำแหน่งที่ต่างกัน และมีลักษณะการแสดงออกทางใบหน้าที่หลายหลาย ได้แก่ ภาพที่ลืมตา/หลับตา ยิ้ม/ไม่ยิ้ม นอกจากนี้ยังมีรายละเอียดของภาพที่แตกต่างกัน ได้แก่ ใส่แว่น/ไม่ใส่แว่น โดยในงานวิจัยนี้นำภาพทั้งหมดมาแบ่งเป็นภาพที่ใช้ในการเรียนรู้จำนวนคนละ 4 ภาพและใช้ในการทดสอบหาอัตราความถูกต้องในการรู้จำจำนวนคนละ 6 ภาพ รวมจำนวนภาพที่ใช้ในการเรียนรู้ทั้งหมด 160 ภาพและภาพที่ใช้ทดสอบทั้งหมด 240 ภาพ ดังแสดงรายละเอียดไว้ในตารางที่ 4.8 ภาพที่ใช้ในการเรียนรู้ที่เหมาะสมได้รับการเลือกขึ้นมาตามตารางที่ 4.7 เพื่อเป็นตัวแทนของกลุ่มข้อมูลใบหน้าของแต่ละบุคคล



รูปที่ 4.21 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL

ตารางที่ 4.7 แสดงภาพที่ใช้ในการเรียนรู้ภาพใบหน้าจากฐานข้อมูล ORL

บุคคล	ภาพใบหน้าที่ใช้ในการเรียนรู้
S1	2, 3, 6, 8
S2	12, 14, 15, 16
S3	21, 22, 25, 28
S4	31, 32, 34, 40
S5	42, 43, 46, 50
S6	51, 56, 57, 59
S7	61, 63, 64, 70
S8	73, 76, 79, 80
S9	83, 87, 89, 90
S10	94, 98, 99, 100
S11	102, 104, 105, 110
S12	113, 116, 117, 118
S13	124, 126, 128, 130
S14	131, 133, 139, 140
S15	142, 144, 145, 148
S16	157, 158, 159, 160
S17	164, 165, 167, 170
S18	172, 175, 176, 177
S19	183, 184, 186, 188
S20	191, 194, 198, 200
S21	204, 205, 207, 210
S22	212, 214, 216, 218
S23	225, 227, 229, 230
S24	231, 233, 237, 238
S25	241, 242, 248, 249
S26	252, 253, 254, 257
S27	261, 262, 267, 270
S28	272, 274, 276, 280

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 (ต่อ)

บุคคล	ภาพใบหน้าที่ใช้ในการเรียนรู้
S29	282, 285, 288, 289
S30	293, 294, 295, 296
S31	302, 307, 308, 310
S32	311, 314, 317, 318
S33	321, 325, 326, 328
S34	331, 334, 337, 338
S35	342, 343, 346, 349
S36	352, 355, 357, 360
S37	361, 366, 367, 370
S38	376, 377, 378, 379
S39	382, 384, 385, 390
S40	392, 393, 397, 400

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 ผลลัพธ์การรู้จำภาพใบหน้า ORL (โดย $p = 0.80$, $S = 2$, และให้แคทากอรีโหนด
จำนวน 153 โหนด)

บุคคล	จำนวนภาพที่ ใช้ในการ ทดสอบ	จำนวนภาพที่ ตอบได้ถูกต้อง	จำนวนภาพที่ ตอบผิด	คำตอบที่ตอบผิด
S1	6	5	1	S38
S2	6	6	0	-
S3	6	4	2	S40, S38
S4	6	5	1	S40
S5	6	6	0	-
S6	6	6	0	-
S7	6	6	0	-
S8	6	6	0	-
S9	6	5	1	S30
S10	6	5	1	S30
S11	6	6	0	-
S12	6	6	0	-
S13	6	5	1	S5
S14	6	6	0	-
S15	6	6	0	-
S16	6	6	0	-
S17	6	6	0	-
S18	6	6	0	-
S19	6	6	0	-
S20	6	6	0	-
S21	6	6	0	-
S22	6	6	0	-
S23	6	5	1	S38
S24	6	6	0	-
S25	6	6	0	-
S26	6	6	0	-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 (ต่อ)

บุคคล	จำนวนภาพที่ใช้ในการทดสอบ	จำนวนภาพที่ตอบได้ถูกต้อง	จำนวนภาพที่ตอบผิด	คำตอบที่ตอบผิด
S27	6	6	0	-
S28	6	6	0	-
S29	6	6	0	-
S30	6	6	0	-
S31	6	6	0	-
S32	6	4	2	S19, S8
S33	6	6	0	-
S34	6	6	0	-
S35	6	6	0	-
S36	6	6	0	-
S37	6	6	0	-
S38	6	6	0	-
S39	6	4	2	S29, S8
S40	6	6	0	-
รวม	240	228	12	-
คิดเป็นเปอร์เซ็นต์		95	5	-

4.3.4 การรู้จำภาพใบหน้าจากฐานข้อมูล MIT

ฐานข้อมูล MIT [23] ประกอบไปด้วยภาพใบหน้าของบุคคลจำนวน 16 คน คนละ 27 รูป รวมจำนวนทั้งหมด 432 รูป เป็นภาพระดับสีเทา 8 บิต ขนาด 128x120 พิกเซล นอกจากนี้ยังเป็นรูปซึ่งมีความแตกต่างในทิศทางการหันศีรษะ ความแตกต่างของการให้แสง และระยะความใกล้ไกลของภาพ ภาพที่มีทิศทางของศีรษะแตกต่างกันแบ่งได้เป็น 3 ประเภท คือ ภาพหน้าตรง ภาพเอียงศีรษะไปทางซ้ายและภาพเอียงศีรษะไปทางขวาโดยจะเอียงทำมุมประมาณ 22.5 องศา การให้แสงประกอบด้วย 3 ประเภท คือ การให้แสงอยู่เหนือศีรษะ ให้แสงทำมุม 45 องศา และให้แสงทำมุม 90 องศา ระยะใกล้ไกลของรูปประกอบด้วยภาพขนาดใหญ่ ขนาดกลาง และขนาดเล็ก ซึ่งภาพที่นำมาใช้ในการรู้จำและการทดสอบจะได้รับการสุ่มจากภาพทั้งหมด โดยสุ่มเพื่อนำไปใช้ในการเรียนรู้จำนวน 18 ภาพต่อคน รวม 288 ภาพ (เนื่องจากมี 2 ภาพที่ถูกคัดออกเพราะเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไฟล์ข้อมูลเสีย) และสุ่มเพื่อนำไปใช้ในการทดสอบทั้งหมด 144 ภาพ คือ 9 ภาพต่อคน ดังแสดงในตารางที่ 4.9



รูปที่ 4.22 ตัวอย่างภาพใบหน้าจากฐานข้อมูล MIT



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 แสดงจำนวนภาพใบหน้าจากฐานข้อมูล MIT ที่ใช้ในการทดลอง

บุคคล	จำนวนภาพ	จำนวนภาพที่ใช้ในการเรียนรู้	จำนวนในการทดสอบ
Billj	27	18	9
Brad	27	18	9
David	25	16	9
Foof	27	18	9
Irfan	27	18	9
Joel	27	18	9
Mike	27	18	9
Ming	27	18	9
Pascal	27	18	9
Robert	27	18	9
Stan	27	18	9
Stephen	27	18	9
Temp	27	18	9
Thad	27	18	9
Trevor	27	18	9
Vmb	27	18	9
Wave	27	18	9
รวม	430	286	144

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.10 ผลลัพธ์การรู้จำภาพใบหน้า MIT (โดย $p = 0.60$, $S = 2$, และให้แคทากอรีโหนด จำนวน 223 โหนด)

บุคคล	จำนวนภาพ	จำนวนภาพที่ตอบ ถูกต้อง	จำนวนภาพที่ตอบผิด
Billj	9	9	0
Brad	9	9	0
David	9	9	0
Foof	9	9	0
Irfan	9	9	0
Joel	9	9	0
Mike	9	9	0
Ming	9	9	0
Pascal	9	9	0
Robert	9	9	0
Stan	9	9	0
Stephen	9	9	0
Temp	9	9	0
Thad	9	9	0
Trevor	9	9	0
Vmb	9	9	0
Wave	9	9	0
รวม	144	144	0
คิดเป็นเปอร์เซ็นต์		100	0

4.3.5 การรู้จำภาพใบหน้าจากฐานข้อมูล Yale

ภาพใบหน้าจากฐานข้อมูล Yale [24] ประกอบไปด้วยภาพแบบ gif จำนวน 165 รูป ขนาด 320x243 พิกเซลของบุคคลจำนวน 15 คนโดยมีจำนวนภาพใบหน้าคนละ 11 รูป ใบหน้า ทั้ง 11 รูปจะประกอบไปด้วยใบหน้าที่แสดงอารมณ์ที่หลากหลายและมีลักษณะที่แตกต่างกัน ได้แก่ ภาพที่มีแสงเข้าจากตรงกลาง ภาพสวมแว่นตา ภาพใบหน้าที่มีความสุข ภาพที่แสงเข้าจากทางด้านซ้าย ภาพใบหน้าที่บุคคลไม่สวมแว่นตา ภาพใบหน้าในอารมณ์ปกติ ภาพที่แสงเข้าทางด้านขวา ภาพใบหน้าที่แสดงอารมณ์เศร้า ภาพใบหน้าที่แสดงความประหลาดใจ และภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบหน้าที่กำลังขยับตา ภาพใบหน้าที่ใช้ในการเรียนรู้จำนวน 6 ภาพต่อคนและภาพที่ใช้ทดสอบจำนวน 5 ภาพต่อคน รวมภาพที่ใช้ในการเรียนรู้จำนวน 90 รูป ภาพที่ใช้ในการทดสอบจำนวน 75 ภาพ โดยภาพที่นำมาใช้ในการรู้จำและการทดสอบได้มาจากการสุ่ม ซึ่งผลลัพธ์ในการรู้จำเมื่อกำหนดให้ $p = 0.80$ และ $s = 2$ ให้ผลความถูกต้องในการรู้จำ 96% ดังแสดงรายละเอียดไว้ในตารางที่ 4.11 ภาพใบหน้าที่เกิดจากการสุ่มขึ้นมาเพื่อใช้ในการเรียนรู้และให้ผลลัพธ์ความถูกต้องตามตารางที่ 4.11 มีรายละเอียดของภาพดังแสดงในตารางที่ 4.12



รูปที่ 4.23 ตัวอย่างภาพใบหน้าจากฐานข้อมูล Yale

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.11 ผลลัพธ์ในการรู้จำภาพใบหน้าจากฐานข้อมูล Yale (โดย $p = 0.80$, $S = 2$, และให้ แดทาทกร์โหนดจำนวน 39 โหนด)

บุคคลที่	จำนวนภาพที่ใช้ในการทดสอบ	จำนวนภาพที่ตอบถูกต้อง	จำนวนภาพที่ตอบผิด	คำตอบที่ตอบผิด
1	5	5	0	-
2	5	5	0	-
3	5	5	0	-
4	5	4	1	1
5	5	5	0	-
6	5	5	0	-
7	5	4	1	13
8	5	5	0	-
9	5	5	0	-
10	5	5	0	-
11	5	5	0	-
12	5	5	0	-
13	5	5	0	-
14	5	5	0	-
15	5	4	1	9
รวม	75	72	3	-
คิดเป็นเปอร์เซ็นต์		96	4	-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.12 ภาพที่ใช้ในการเรียนรู้ของระบบที่ได้จากการสุ่มและให้ผลลัพธ์ในตารางที่ 4.11

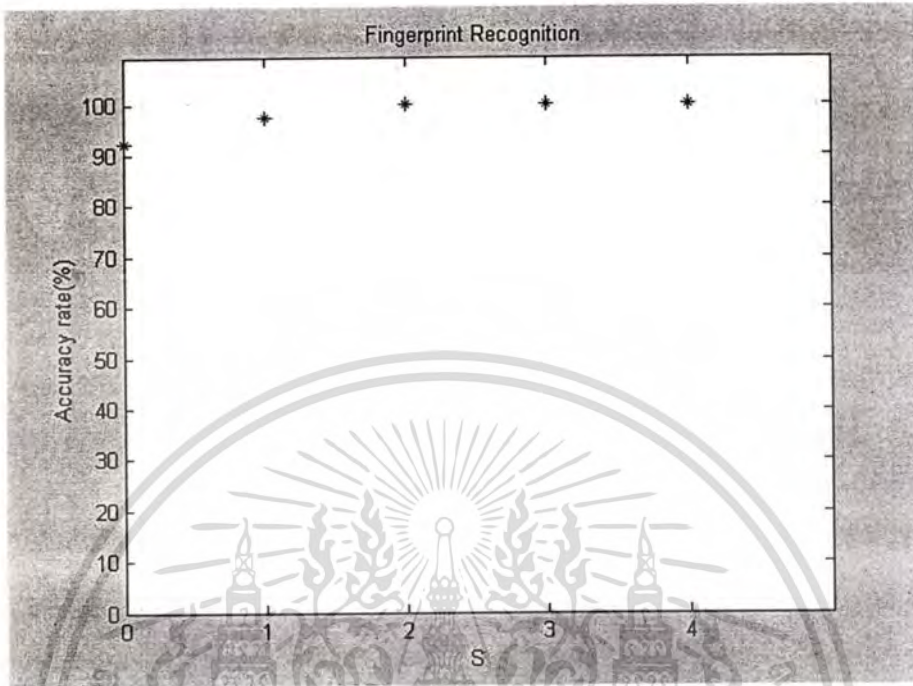
บุคคลที่	ภาพที่ใช้ในการเรียนรู้
1	center-light, happy, w/no glasses, sad, surprised, wink
2	center-light, left-light, w/no glasses, normal, surprised, wink
3	center-light, w/glasses, left-light, right-light, surprised, wink
4	happy, left-light, w/no glasses, normal, right-light, sleepy
5	w/glasses, happy, left-light, w/no glasses, right-light, sleepy
6	center-light, w/glasses, left-light, w/no glasses, surprised, wink
7	center-light, w/glasses, happy, left-light, right-light, sleepy
8	center-light, w/glasses, left-light, normal, right-light, sad
9	center-light, w/glasses, left-light, normal, right-light, sad
10	center-light, w/glasses, normal, right-light, sad, wink
11	center-light, left-light, w/no glasses, right-light, sleepy, wink
12	center-light, w/glasses, happy, left-light, right-light, surprised
13	center-light, w/glasses, w/no glasses, normal, sad, sleepy
14	w/glasses, happy, left-light, w/no glasses, sad, sleepy
15	center-light, left-light, w/no glasses, normal, right-light, surprised

4.4 การวิเคราะห์ถึงผลกระทบของค่า S ต่อประสิทธิภาพในการรู้จำของฮาอูสตอร์ฟาร์ทแมพ

เพื่อศึกษาและวิเคราะห์ถึงผลกระทบของค่า S ต่อประสิทธิภาพในการรู้จำของกระบวนการฮาอูสตอร์ฟาร์ทแมพจึงได้ทำการทดลองเปลี่ยนค่า S เป็นค่าต่างๆ โดยให้ค่า p เป็นค่าคงที่และกำหนดให้ลำดับในการป้อนรูปแบบอินพุตเพื่อทำการเรียนรู้เป็นลำดับที่เหมือนกันในทุกครั้งที่เปลี่ยนค่า S ผลการทดสอบเพื่อทำการรู้จำลายนิ้วมือ ภาพมือ ภาพใบหน้าจากฐานข้อมูล ORL, MIT, และ Yale ให้ผลดังกราฟในรูปที่ 4.24 – 4.27 ซึ่งจากกราฟดังกล่าวจะเห็นได้ว่ามีเพียงการรู้จำลายนิ้วมือเท่านั้นที่ค่า S จะมีผลกระทบต่อการรู้จำ โดยเมื่อค่า S มีค่าน้อยกว่า 2 จะทำให้ความสามารถในการรู้จำภาพลายนิ้วมือลดลงเล็กน้อย แต่ค่า S จะไม่มีผลกระทบต่อการรู้จำภาพมือ ภาพใบหน้าจากฐานข้อมูล ORL, MIT, และ Yale เลย ดังนั้นสรุปได้ว่าค่า S มีผลต่อกระบวนการรู้จำด้วยวิธีการฮาอูสตอร์ฟาร์ทแมพน้อยมาก แต่อย่างไรก็ตามไม่ควรที่จะกำหนดค่า S ให้มีขนาดเล็กเกินไปเพราะจุดบางจุดที่มีความสำคัญในรูปแบบอ้างอิงซึ่งไม่ใช่ noise

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

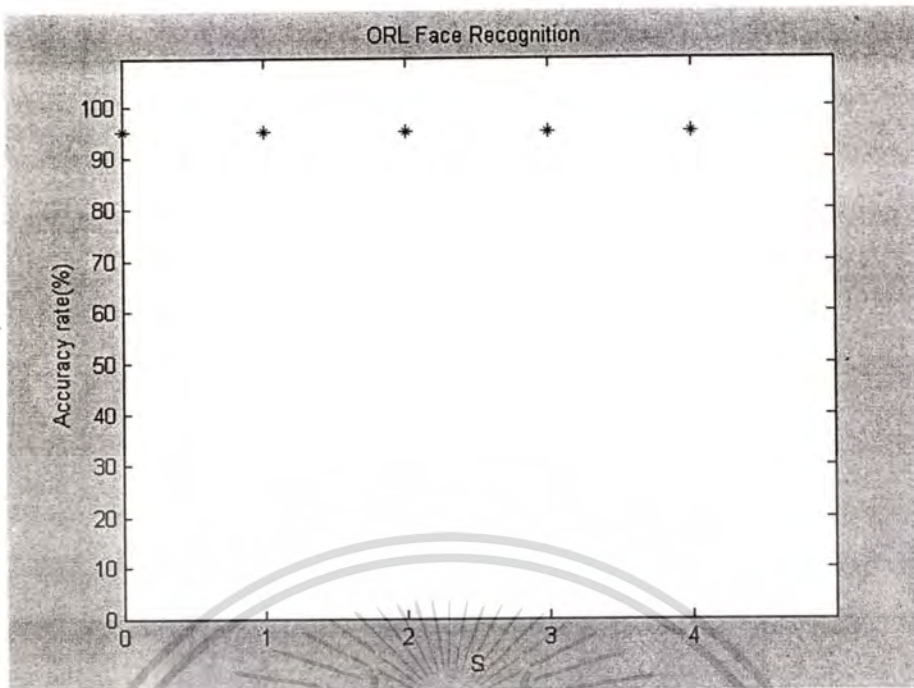
อาจถูกขจัดออกไปด้วยโดยไม่ตั้งใจ เนื่องจากเมื่อพิจารณาจุดนั้นเป็นออกเป็นระยะทาง S จะไม่มีจุดใดๆ อยู่เลยในรูปแบบอินพุทที่มีค่าเป็น 1 ทั้งนี้เป็นเพราะพิจารณาในพื้นที่ที่น้อยเกินไป



รูปที่ 4.24 แสดงผลความถูกต้องในการรู้จำลายนิ้วมือที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.80$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลอสตอร์ฟอาร์ทแมพเป็นลำดับที่เหมือนกัน



รูปที่ 4.25 แสดงผลความถูกต้องในการรู้จำภาพมือที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 1.00$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลอสตอร์ฟอาร์ทแมพเป็นลำดับที่เหมือนกัน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

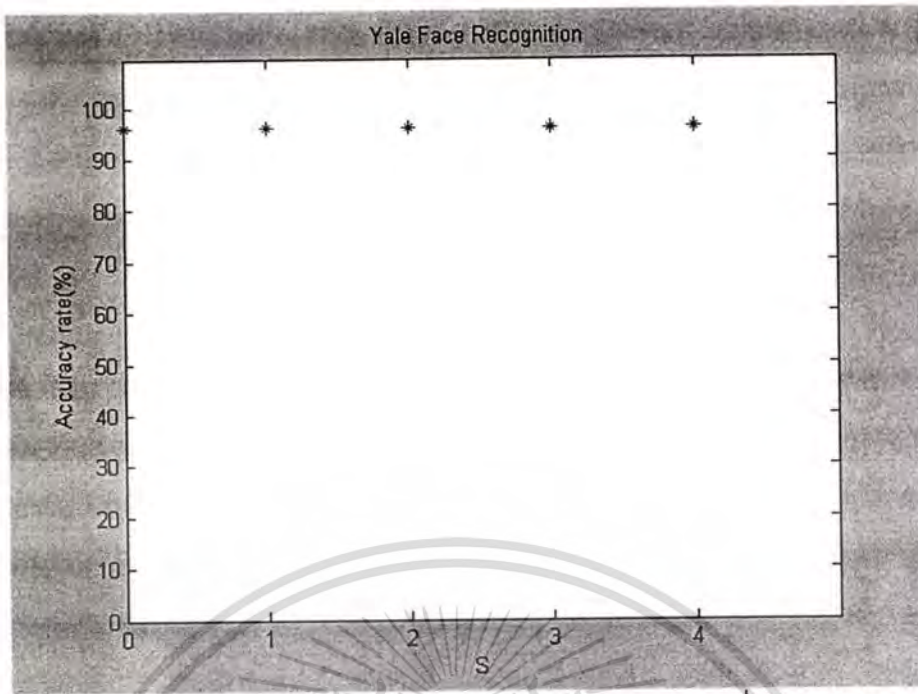


รูปที่ 4.26 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล ORL ที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.80$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟอาร์ทแมพเป็นลำดับที่เหมือนกัน



รูปที่ 4.27 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล MIT ที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $p = 0.60$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮาลูสตอร์ฟอาร์ทแมพเป็นลำดับที่เหมือนกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.28 แสดงผลความถูกต้องในการรู้จำใบหน้าจากฐานข้อมูล Yale ที่สัมพันธ์กับค่า S ต่างๆ เมื่อ $\rho = 0.80$ และกำหนดให้ลำดับในการส่งรูปแบบเข้าไปเรียนรู้ในฮิสโตแกรมพาร์ทแมพเป็นลำดับที่เหมือนกัน

4.5 การวิเคราะห์เปรียบเทียบการรู้จำใบหน้าด้วยวิธีฮิสโตแกรมพาร์ทแมพกับวิธีสเปเชียลลีไอเกนเวทเต็ดดับเบิลฮิสโตแกรมพาร์ทแมพ

เพื่อทำการวิเคราะห์เปรียบเทียบถึงประสิทธิภาพและการทำงานของกระบวนการรู้จำใบหน้าด้วยวิธีสเปเชียลลีไอเกนเวทเต็ดดับเบิลฮิสโตแกรมพาร์ทแมพ (spacially eigen weighted doubly Hausdorff distance : SEW2HD) ซึ่งได้รับการเสนอใน [21] กับวิธีการฮิสโตแกรมพาร์ทแมพซึ่งได้พัฒนาขึ้นในงานวิจัยนี้จึงได้ทำการรู้จำภาพใบหน้าจากฐานข้อมูล ORL และ MIT ด้วยวิธีการทั้งสองเปรียบเทียบกันซึ่งผลลัพธ์ความถูกต้องในการรู้จำได้แสดงไว้ดังตารางที่ 4.13 จากตารางที่ 4.13 จะเห็นได้ว่าการรู้จำภาพใบหน้าด้วยวิธีฮิสโตแกรมพาร์ทแมพให้ผลในการรู้จำที่ดีกว่าและมีจำนวนรูปแบบอ้างอิงน้อยกว่า

วิธีการ SEW2HD นั้นจะต้องทำการวัดค่าฮิสโตแกรมพาร์ทแมพดังสมการต่อไปนี้

$$H(A, B) = \max \{h_{sw}(A, B), h_{sw}(B, A)\} \quad (4.1)$$

เมื่อกำหนดให้ $A = \{a_1, a_2, \dots, a_p\}$ และ $B = \{b_1, b_2, \dots, b_q\}$ เป็นเซต 2 เซตของจุดที่มีจำนวนสมาชิกจำกัด และ $h_{sw}(A, B)$ คือ ค่าสเปเชียลลีไอเกนเวทเต็ดโดเร็คเตดฮิสโตแกรมพาร์ทแมพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แทนซ์ (spacially eigen weighted directed Hausdorff distance) จาก A ไป B ซึ่งคำนวณได้จากสมการต่อไปนี้คือ

$$h_{sw}(A, B) = \frac{1}{N_a} \cdot \sum_{N_a} w_e(b) \min \| a - b \| \quad (4.2)$$

จากสมการที่ (4.2) การหาค่าสเปเชียลไฮเอนเวทเต็ดไดเร็กเต็ดฮาอูสดอร์ฟดิสแทนซ์ จำเป็นจะต้องนำค่าไปคูณเข้ากับค่าน้ำหนักที่ได้จากไฮเอนเวทเต็ดฟังก์ชัน ($w_e(b)$) ด้วย ซึ่งในการคำนวณค่าไดเร็กเต็ดฮาอูสดอร์ฟดิสแทนซ์ด้วยสมการที่ (2.3) และ (2.4) จะไม่มีการพิจารณา ค่าน้ำหนักนี้ แต่อย่างไรก็ตามถ้าอาศัยหลักการสร้าง look up table ดังเสนอไว้ในหัวข้อที่ 3.4.3 โดยคำนวณผลคูณของค่าน้ำหนักแต่ละจุดกับค่าระยะทางที่น้อยที่สุดไว้ก่อนก็จะทำให้ความซับซ้อนในการคำนวณของทั้งสองวิธีนี้มีค่าเท่ากัน

ตารางที่ 4.13 ผลลัพธ์ในการรู้จำภาพใบหน้าจากฐานข้อมูล ORL และ MIT ด้วยวิธีการสเปเชียลไฮเอนเวทเต็ดคูบลิฮาอูสดอร์ฟดิสแทนซ์และวิธีการฮาอูสดอร์ฟอาร์ทแมพ

ฐานข้อมูล	จำนวนบุคคล	จำนวนภาพที่ใช้ train	จำนวนภาพที่ใช้ test	จำนวนรูปแบบอ้างอิง		ความถูกต้องในการรู้จำ	
				SEW2HD	HD ARTMAP	SEW2HD	HD ARTMAP
ORL	40	160	240	160	153	91%	95%
MIT	16	286	144	286	223	92%	100%

จำนวนครั้งในการคำนวณค่า SEW2HD ต่อหนึ่งภาพใบหน้าที่ต้องการทดสอบเพื่อหาคำตอบว่าเป็นใบหน้าของบุคคลใดจะต้องคำนวณค่า SEW2HD ระหว่างภาพใบหน้าที่จะทดสอบกับรูปแบบอ้างอิงทั้งหมดที่มีในระบบซึ่งมีค่าเท่ากับ 160 ภาพในการรู้จำใบหน้าจากฐานข้อมูล MIT และ 286 ภาพสำหรับการรู้จำใบหน้าในฐานข้อมูล ORL ดังตารางที่ 4.13 จากนั้นจึงทำการเลือกชนิดของรูปแบบอ้างอิงที่ให้ค่า SEW2HD น้อยที่สุดเป็นคำตอบของระบบ คือต้องคำนวณค่า SEW2HD จำนวน 160 ครั้งและ 286 ครั้งตามลำดับ ในกระบวนการฮาอูสดอร์ฟอาร์ทแมพจะคำนวณค่าฮาอูสดอร์ฟดิสแทนซ์ระหว่างภาพใบหน้าที่ต้องการทดสอบกับรูปแบบอ้างอิงซึ่งในที่นี้ก็คือค่าน้ำหนักของแต่ละโหนดในชั้นแคทากอรี แล้วเลือกโหนดที่จะชนะจากโหนดที่มีค่าฮาอูสดอร์ฟดิสแทนซ์น้อยที่สุดมาพิจารณาหาชนิดของใบหน้าสัมพันธ์กับโหนดนี้เพื่อเป็นคำตอบ ดังนั้นจำนวนครั้งในการคำนวณค่าฮาอูสดอร์ฟดิสแทนซ์สำหรับวิธีการฮาอูสดอร์ฟอาร์ทแมพจึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เท่ากับจำนวนโหนดในชั้นแคทากอรีซึ่งในการรู้จำใบหน้าจากฐานข้อมูล MIT และ ORL มีค่าเท่ากับ 153 และ 223 ตามลำดับ ดังจะเห็นได้ว่าเป็นค่าน้อยกว่าวิธีการ SEW2HD จึงสรุปได้ว่าวิธีการฮาอูสเตอร์ฟอาร์ทแมพมีจำนวนครั้งในการคำนวณที่น้อยกว่าวิธีการ SEW2HD ในส่วนของการทดสอบและนำไปใช้งานจริง เพราะวิธีการฮาอูสเตอร์ฟอาร์ทแมพสามารถสร้างจำนวนรูปแบบอ้างอิงที่น้อยกว่าได้ แต่อย่างไรก็ตามฮาอูสเตอร์ฟอาร์ทแมพจะใช้เวลาในการเรียนรู้มากกว่าวิธีการสเปเชียลลีไอเกนเวทเด็คโคเร็คเด็คฮาอูสเตอร์ฟอาร์ทแมพ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้อาศัยแนวคิดในการรู้จำโดยใช้โครงร่างของวัตถุ เช่น ลายเส้น เส้นขอบ เป็นต้น ในการรู้จำแทนการใช้ค่าพิกเซลที่มีความซับซ้อน เนื่องจากการศึกษาพบว่ามนุษย์สามารถรู้จำรูปภาพที่เป็นลายเส้นได้อย่างรวดเร็วและถูกต้อง ในงานวิจัยนี้จึงได้คิดค้นและเสนอกระบวนการรู้จำรูปภาพแบบไบนารีโดยใช้ฮิสโตแกรมฟอาร์ทแมนซ์ ซึ่งกระบวนการฮิสโตแกรมฟอาร์ทแมนซ์เป็นโครงข่ายประสาทเทียมที่พัฒนามาจากซิมพลิไฟด์ฟิซซีอาร์ทแมพ โดยนำเอาตัววัดค่าความเหมือนคือฮิสโตแกรมฟิตเนสที่ได้รับการยอมรับว่ามีการทำงานและประสิทธิภาพคล้ายมนุษย์มาใช้ในส่วนของการคำนวณชอยซ์ฟังก์ชันและการพิจารณาเงื่อนไขของวิเจแลนซ์แทนวิธีการดำเนินการแบบฟิซซี และฮิสโตแกรมฟอาร์ทแมนซ์ยังมีคุณสมบัติเช่นเดียวกับกับแนวทางของอาร์ทโดยทั่วไป คือสามารถรู้จำรูปแบบใหม่ๆ ได้โดยไม่จำเป็นต้องทึ่งความรู้เก่าที่เคยเรียนรู้มาแล้ว ทำให้แก้ไขปัญหาการรู้จำรูปแบบชนิดเดียวกันแต่มีหลายลักษณะได้ นอกจากนี้แล้วกระบวนการนี้ยังมีการปรับค่าน้ำหนักซึ่งเปรียบเสมือนตัวแบบอ้างอิงในการรู้จำ เพื่อให้ได้รูปแบบอ้างอิงที่มีประสิทธิภาพดียิ่งขึ้น โดยการปรับค่าน้ำหนักในฮิสโตแกรมฟอาร์ทแมนซ์ได้ใช้แนวคิดของการพิจารณาเพื่อกำจัดส่วนเกินหรือ noise ออกจากรูปแบบอ้างอิงในขณะที่มีการปรับค่าน้ำหนัก

ในการทดลองเพื่อทดสอบประสิทธิภาพในการทำงานของกระบวนการฮิสโตแกรมฟอาร์ทแมนซ์ได้นำเอาข้อมูลในการระบุตัวบุคคล 3 ประเภทมาใช้ ได้แก่ ลายนิ้วมือ มือและใบหน้า ซึ่งก็ให้ผลความถูกต้องในการรู้จำที่ดี คือ

- การรู้จำลายนิ้วมือจากฐานข้อมูลของ Biometrics Systems Lab., University of Bologna, Cesena-Italy โดยทำการรู้จำภาพลายนิ้วมือของคนจำนวน 20 คน ภาพลายนิ้วมือที่ใช้ในการเรียนรู้จำนวน 100 ภาพและภาพที่ใช้ในการทดสอบจำนวน 38 ภาพ ให้ผลความถูกต้องในการรู้จำ 100%
- การรู้จำโครงร่างมือจากฐานข้อมูลแหล่งเดียวกันกับลายนิ้วมือ โดยเป็นภาพมือของคนจำนวน 8 คน ภาพที่ใช้ในการเรียนรู้จำนวน 50 ภาพ ภาพที่ใช้ในการทดสอบจำนวน 16 ภาพ ให้ผลความถูกต้องในการรู้จำ 100%
- ฐานข้อมูลใบหน้า ORL ซึ่งประกอบไปด้วยภาพใบหน้าของบุคคลจำนวน 40 คน จำนวนภาพที่ใช้ในการเรียนรู้จำนวน 160 ภาพ และจำนวนภาพที่ใช้ในการทดสอบจำนวน 240 ภาพ เปอร์เซนต์ความถูกต้องในการรู้จำที่ได้ คือ 95%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ฐานข้อมูลใบหน้า MIT ประกอบไปด้วยภาพใบหน้าของคนจำนวน 16 คน ทำการ เรียนรู้ด้วยภาพใบหน้า 286 ภาพ และทดสอบผลความถูกต้องในการรู้จำด้วยภาพใบหน้าจำนวน 144 ภาพ ผลความถูกต้องที่ได้ คือ 100%

- ฐานข้อมูลใบหน้า Yale ทำการรู้จำภาพใบหน้าของบุคคลจำนวน 15 คน จำนวน ภาพใบหน้าที่ใช้ในกระบวนการเรียนรู้มีจำนวน 90 ภาพ และทดสอบประสิทธิภาพในการรู้จำด้วย ภาพจำนวน 75 ภาพ ผลความถูกต้องในการรู้จำที่ได้ คือ 96%

แต่อย่างไรก็ดีจะเห็นได้ว่าการรู้จำข้อมูลแต่ละแบบก็จำเป็นที่จะต้องทำการเตรียมข้อมูล ซึ่งมีวิธีการที่แตกต่างกันไปเพื่อแปลงข้อมูลให้อยู่ในรูปของภาพลายเส้นแบบไบนารี และ กระบวนการวิธีในการเตรียมข้อมูลมีผลอย่างยิ่งต่อประสิทธิภาพในการรู้จำของฮาคูสเตอร์ฟาร์ท-แมพ เพราะถ้าข้อมูลภาพแบบไบนารีที่ส่งเข้าไปทำการเรียนรู้ในระบบนั้นมีคุณภาพไม่ดีก็จะทำ ให้ผลความถูกต้องในการรู้จำของระบบลดลงไปด้วย

นอกจากนี้แล้วในกระบวนการเตรียมข้อมูลยังจำเป็นต้องทำการหาจุดอ้างอิงของภาพให้ ถูกต้อง เพื่อใช้ปรับขนาด ทิศทางการหมุน และการเลื่อนตำแหน่งของภาพระหว่างรูปแบบอินพุท กับรูปแบบอ้างอิงให้ตรงกัน เนื่องจากถ้าทำการเปรียบเทียบระหว่างรูปแบบอินพุทกับรูปแบบ อ้างอิงที่มีขนาดและการวางตัวที่ต่างกันมากก็จะยิ่งทำให้ความถูกต้องในการเปรียบเทียบลดลง โดยเฉพาะอย่างยิ่งรูปแบบอินพุทที่เป็นภาพลายเส้นที่มีรายละเอียดสูงเช่นในที่นี้คือภาพลายนิ้วมือ ตำแหน่งอ้างอิงที่ใช้จะผิดพลาดมากไม่ได้เพราะจะทำให้ตำแหน่งในการเปรียบเทียบคลาดเคลื่อน ไปส่งผลให้ผลการเปรียบเทียบหรือการวัดค่าความเหมือนมีความผิดพลาด ดังจะเห็นได้จากการ ทดลองรู้จำลายนิ้วมือในงานวิจัยนี้ต้องพยายามย้ายตำแหน่งอ้างอิงเป็นตำแหน่งต่างๆ เพื่อหา ตำแหน่งที่เหมาะสมที่สุด ดังนั้นวิธีการฮาคูสเตอร์ฟาร์ทแมพนี้จึงไม่เหมาะสมเท่าใดนักในการ นำไปใช้รู้จำภาพที่มีรายละเอียดสูง แต่จะให้ผลที่ติดว่าเมื่อนำมาใช้กับภาพที่เป็นลายเส้นที่ไม่ ละเอียดซับซ้อนมากนัก เช่น ภาพภาพโครงร่างต่างๆ เป็นต้น

การนำเอาฮาคูสเตอร์ฟิดิสแทนซ์มาเป็นตัววัดค่าความเหมือนมีข้อดีคือการดำเนินการ ของฮาคูสเตอร์ฟิดิสแทนซ์ไม่ได้เป็นการดำเนินการแบบจุดต่อจุด ทำให้รูปแบบที่นำมา เปรียบเทียบกันไม่จำเป็นต้องมีความเหมือนกันพอดีจึงจะสามารถระบุได้ว่าเป็นรูปแบบชนิด เดียวกัน ซึ่งทำให้ฮาคูสเตอร์ฟิดิสแทนซ์มีความเหมาะสมมากในการนำมาใช้วัดค่าความเหมือน ของวิธีการที่ไม่ใช่พีเจหรือใช้พีเจรวม (global feature) การที่ไม่ต้องทำการสกัดค่า พีเจจำนวนมากนักช่วยลดเวลาในการคำนวณไปได้มากก็จริงแต่ผลความผิดพลาดก็ย่อมมีได้ มากด้วยเช่นกัน เพราะเป็นการดำเนินการที่เปรียบเสมือนกับการพิจารณาอย่างคร่าวๆ ซึ่งถ้า ต้องการความละเอียดอ่อนและความถูกต้องในการพิจารณาสูงหรือรูปแบบที่ทำการรู้จำมีความ ซับซ้อนมากก็อาจจะทำให้เกิดความผิดพลาดได้ง่าย ดังนั้นวิธีการรู้จำที่ต้องใช้ความแม่นยำสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่น การรู้จำลายนิ้วมือของแต่ละบุคคล การรู้จำลายเซ็นต์ เป็นต้น กระบวนการรู้จำโดยใช้ ฮาอูสเตอร์ฟาร์ทแมพจึงอาจจะนำไปใช้ร่วมกับวิธีการรู้จำที่ต้องใช้ค่าพีเจอร์ โดยนำเอามาใช้ในส่วนของการกรองและคัดแยกชนิดของข้อมูลเบื้องต้น (classification) เพื่อให้ลดจำนวนในการคำนวณข้อมูลที่อยู่ในฐานข้อมูลขนาดใหญ่ลง หรือช่วยลดจำนวนครั้งในการคำนวณพีเจอร์ของรูปแบบลง เพราะได้แยกรูปแบบที่มีความเหมือนกันกับรูปแบบที่พิจารณาอยู่น้อยออกไปแล้วบางส่วน ต่อจากนั้นก็ทำการคำนวณหาค่าพีเจอร์ตามปกติ เพื่อเป็นการพิจารณาแบบละเอียดขึ้นต่อไป

การปรับค่าน้ำหนักของฮาอูสเตอร์ฟาร์ทแมพนั้น ทำได้เพียงแค่จัดส่วนที่เป็นส่วนเกินที่ผิดพลาดหรือ noise ออกไปได้เท่านั้น แต่ยังไม่สามารถที่จะเพิ่มส่วนที่ขาดหายไปของรูปแบบอ้างอิงให้สมบูรณ์ได้ ในการพัฒนาเพื่อปรับปรุงประสิทธิภาพของกระบวนการฮาอูสเตอร์ฟาร์ทแมพให้ดียิ่งขึ้นจึงน่าจะทำการพัฒนาในส่วนของการปรับค่าน้ำหนักในกระบวนการเรียนรู้ของระบบให้สามารถสร้างรูปแบบอ้างอิงให้มีประสิทธิภาพมากกว่านี้

กระบวนการฮาอูสเตอร์ฟาร์ทแมพนั้นลำดับของรูปแบบที่ทำการเรียนรู้มีผลต่อความถูกต้องในการรู้จำ ซึ่งจากการทดลองทำให้สังเกตได้ว่าเมื่อใช้รูปแบบชุดเดียวกันในการรู้จำแต่ลำดับในการรู้จำต่างกันทำให้ผลความถูกต้องในการรู้จำไม่เท่ากันได้ และจำนวนโหนดของระบบก็ไม่เท่ากัน เนื่องจากระบบจะกำหนดค่าเริ่มต้นของน้ำหนักที่เป็นตัวแบบอ้างอิงของแต่ละโหนดในชั้นแคทาคอริตามรูปแบบอินพุตแรกที่โหนดนั้นทำการเรียนรู้หรือจัดอยู่ในโหนดนั้น เมื่อมีรูปแบบอินพุตอื่นๆ เข้าในระบบและจัดอยู่ในโหนดเดียวกันนี้ รูปแบบอินพุตอื่นๆ จะทำได้เพียงแค่ปรับปรุงให้รูปแบบอ้างอิงเดิมนั้นมีคุณภาพที่ดีขึ้นเท่านั้นแต่จะมีผลน้อย ดังจะเห็นได้ว่ารูปแบบอินพุตตั้งต้นของแต่ละโหนดจะมีผลต่อลักษณะของรูปแบบอ้างอิงที่มีอยู่ในระบบมากที่สุด

การรู้จำโดยใช้ฮาอูสเตอร์ฟาร์ทแมพนั้นมีลักษณะที่คล้ายคลึงกับการรู้จำโดยใช้โครงข่ายประสาทเทียมโดยทั่วไปในแง่ที่ว่าระบบจะต้องได้รับการเรียนรู้ที่ครอบคลุมหรือเป็นตัวแทนของรูปแบบทั้งหมดจึงจะให้ผลลัพธ์ในการรู้จำที่ดี เพราะถ้ารูปแบบลักษณะใดที่มีความแตกต่างจากรูปแบบที่ระบบเคยเรียนรู้มาแล้วมาก ระบบก็ไม่สามารถที่จะจดจำหรือระบุได้ถูกต้อง ดังนั้นการเลือกข้อมูลอินพุตเพื่อให้ระบบทำการเรียนรู้จึงเป็นอีกปัจจัยหนึ่งที่สำคัญและมีผลต่อประสิทธิภาพในการรู้จำของฮาอูสเตอร์ฟาร์ทแมพอย่างยิ่ง

ในการเลือกค่า p และค่า s ที่เหมาะสมนั้นขึ้นอยู่กับชนิดของข้อมูล คือ ถ้าข้อมูลมีความละเอียดและซับซ้อนของลายเส้นน้อยค่า s ควรจะมีค่ามากและค่า p ควรจะมีค่ามาก เพราะถ้าให้ s มีค่าน้อยเกินไปอาจทำให้จุดที่ไม่ใช่ noise ในรูปแบบอ้างอิงถูกขจัดออกไปด้วย และในการวัดค่าความแตกต่างด้วยฮาอูสเตอร์ฟิตสแทนท์ผลลัพธ์ที่ได้จากรูปแบบที่มีลายเส้นน้อย เช่น ภาพเส้นขอบของฝ่ามือ เป็นต้น จะให้ค่าฮาอูสเตอร์ฟิตสแทนท์ที่มีค่ามากกว่ากรณีภาพที่มีลายเส้นถี่

กว่า คืออาจจะให้รูปแบบที่มีความเหมือนกันน้อยสามารถจัดอยู่ในกลุ่มเดียวกันได้ เพราะข้อมูลแบบนี้ส่วนใหญ่อาจอยู่ในชนิดเดียวกันแต่มีความเหมือนกันน้อย กรณีที่ภาพมีข้อมูลและรายละเอียดของลายเส้นมากค่า S ควรจะมีค่าน้อยเพราะการปรับค่าน้ำหนักจะได้ไม่กินบริเวณกว้างเกินไปจนไม่สามารถกำจัดจุดที่ไม่ต้องการหรือจุดที่เป็น noise ออกไปได้ และค่า p ควรจะมีค่าน้อยเพราะรูปแบบต่างชนิดกันจะมีความเหมือนกันมาก เช่น ภาพลายนิ้วมือ ของแต่ละบุคคลจะมีความแตกต่างกันไม่มาก การจัดกลุ่มของข้อมูลก็ควรจะทำให้มีความกระชับสูง คือ รูปแบบที่เหมือนกันมากจริงๆ จึงจะจัดให้อยู่ในกลุ่มเดียวกันได้ แต่อย่างไรก็ตามค่า p และค่า S ที่เหมาะสมอย่างแท้จริงสำหรับการรู้จำรูปแบบแต่ละประเภทสามารถทำการหาได้จากการทดลองข้อเสนอแนะเบื้องต้นนี้เป็นเพียงแนวทางในการปรับค่า p และค่า S เท่านั้น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] Rosandish R.G. **Intelligent Visual Inspection using Artificial Neural Networks.**
London : Chapman & Hall. 1997.
- [2] Xuejing Wu, Pengfei Shi. "Unconstrained Handwritten Numeral Recognition Using Hausdorff Distance and Multi-Layer Neural Network Classifier." Document Analysis and Recognition, 1999. ICDAR '99. Proceedings of the Fifth International Conference on, 20-22 Sept. 1999. pp. 249 -252.
- [3] Carpenter G.A., Grossberg S. "A massively parallel architecture for a self-organization neural pattern recognition machine." Computer Vision, Graphics and Image Processing. vol. 37, 1987. pp. 54-115.
- [4] Tay Y.H., et al. "Hand-written Postcode Recognition by Fuzzy ARTMAP Neural Network." Proc. of Malaysian Science and Technology Congress '97 (MSTC'97), Genting Highland, Malaysia.
- [5] Kasuba T. "Simplified Fuzzy ARTMAP." AI Expert, vol. 8, November 1993. pp. 18-25.
- [6] Heinz Breu, et al. "Linear Time Euclidean Distance Transform Algorithms." IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 17, no. 5, 1995. pp. 529-533.
- [7] A.K. Jain, et al. "An identity-authentication system using fingerprints." Proceedings of the IEEE. vol. 85, no. 9, 1997. pp. 1365 –1388.
- [8] Leong Chung Ern, G. Sulong. "Fingerprint classification approaches: an overview." Signal Processing and its Applications, Sixth International, Symposium on. vol. 1, 2001. pp. 347 –350.
- [9] Asker M. Bazen, Sabin H. Gerez. "Systematic methods for the Computation of Directional Fields and Singular Points of Fingerprints." IEEE transactions on pattern analysis and machine intelligence. vol. 24, no.7, July 2002.
- [10] Pualie Rämö, Marius Tico et al. "Optimized Singular point detection Algorithm for Fingerprint Images." Image Processing, 2001. Proceedings. 2001 International Conference on. vol. 3, 7-10 Oct. 2001. pp. 242 - 245.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [11] Biometrics Systems Lab., University of Bologna, Cesena-Italy, "Databases and Software" [Online]. Available : http://www.csr.unibo.it/research/biolab/bio_tree.html. 2003.
- [12] Mohamed S.M., Nyongesa H. O. "Automatic Fingerprint Classification System Using Fuzzy Neural Technique." Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on. vol. 1, 12-17 May 2002. pp. 358 – 362.
- [13] Haralick, Robert M., and Linda G. Shapiro. **Computer and Robot Vision**. Volume I. MA : Addison-Wesley. 1992.
- [14] R. Sanchez-Reillo, et al. "Biometric identification through hand geometry measurements." Pattern Analysis and Machine Intelligence, IEEE Transactions on. vol. 22, no. 10, 2000. pp. 1168 –1171.
- [15] R. Sanchez-Reillo. "Hand geometry pattern recognition through Gaussian mixture modeling." Pattern Recognition, 2000. Proceedings. 15th International Conference on. vol. 2, 2000. pp. 937 -940.
- [16] Chellappa R., Wilson C.L., and Sirciohey S. "Human and machine recognition of faces: a survey." Proc. IEEE. vol. 83, no. 5, 1991. pp. 705-740.
- [17] Grudin M.A. "On internal representations in face recognition systems." Pattern Recognition. vol. 33, no. 7, 2000. pp. 1161-1177.
- [18] Lin S., Kung S., and Lin L. "Face recognition/detection by probabilistic decision-based neural network." IEEE Trans. Neural Networks. vol. 81, no. 1, 1997. pp. 114-132.
- [19] Takács B. "Comparing face images using the modified Hausdorff distance." Pattern Recognition. vol. 31, no. 12, 1998. pp. 1873-1880.
- [20] Gonzalez R.C., Woods R.E. **Digital Image Processing**. Englewood Cliffs, NJ : Prentice-Hall. 2002.
- [21] Kwan-Ho Lin, Kin-Man Lam and Wan-Chi Siu / K.H. Lin et al. "Spatially eigen-weighted Hausdorff distances for human face recognition." Pattern Recognition, vol. 36, no. 8, August 2003. pp. 1827-1834.
- [22] AT&T Laboratories Cambridge University, UK. "ORL Face Database." [Online]. Available : <http://www.uk.research.att.com/facedatabase.html>. 2003.

- [23] MIT Media Laboratory Vision and Modeling Group. "MIT Face Database." [Online]. Available : <ftp://whitechapel.media.mit.edu/pub/images/>. 2003.
- [24] Yale University. "Yale University Face Database." [Online]. Available : <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>. 2004.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Proceedings of the 3rd WSEAS International Conference on ARTIFICIAL,
INTELLIGENCE, KNOWLEDGE ENGINEERING, DATA BASES



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HAUSDORFF ARTMAP FOR HUMAN FACE RECOGNITION

ARIT THAMMANO* AND CHONGKOLNEE RUNGRUANG**

Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang,
Bangkok, 10520 Thailand

E-mail: arit@it.kmitl.ac.th* and chongkolnee_r@yahoo.com**

Abstract: - Neural network's ability to automatically learn from examples makes it attractive and exciting. This paper proposes a new neural network approach specifically designed for solving two dimensional binary pattern recognition problems. This new neural network employs the concept of the Hausdorff distance to measure the likeness or similarity between the incoming input pattern and the reference patterns of each subject. The performance of this proposed model has been compared with the other different structures studied in the previous researches. The experimental results on two benchmark databases, the ORL and MIT face databases, show that the proposed network has the best performance among all.

Key-Words: - Neural Network, Face Recognition, Biometric Recognition System, Person Identification

1 Introduction

Person identification has received increasing attention in recent years. In general, there are three ways to identify an individual: the person knows something (e.g., a PIN, a password); the person possesses something (e.g., an ID card, a passport); or by measuring something about the person's body [1]. The later encompasses the biometric identification. Among all of the biometric identification methods, face recognition is the most natural, non-intrusive, and user-friendly biometric measure because it requires no disturbance to the person being identified. While more intrusive biometric recognition systems (e.g., palm, fingerprint) are presently more accurate, face recognition still has a critical role in certain domains since the person being identified may be at a distance from the sensor, the person does not have to be compliant, and recognition can be performed continuously.

A variety of techniques have been applied to deal with the face recognition problems. The reader should pay attention to Chellappa et al. [2] and the references therein for a more complete survey of previous research works on face recognition. In the early years, many researchers used the structure parameters of faces as the features in the facial image recognition. Kelly [3] used various kinds of facial features, including width of the head and distances between eyes, top of head to eyes, between eyes and nose and the distance from eyes to mouth. During the past decade, researchers have paid much attention to the statistical approaches. -- such as eigenfaces [4], KL transform [5], and SVD [6] --

and the neural network approaches [7, 8]. Neural network is very suitable for face recognition systems. It has the ability to automatically learn the rules from the given collection of representative examples, instead of following a set of human-designed rules [9]. Moreover, it is well-known that the neural network is more robust to noise than other methods. The Hausdorff ARTMAP neural network described in this paper employs the concept of the Hausdorff distance to measure the likeness or similarity between the incoming input pattern and the reference patterns of each subject. The ORL and MIT face databases are used in this study to evaluate the performance of the proposed neural network.

Following this introduction, section 2 briefly describes the fuzzy ARTMAP learning algorithm. The basic concept of the Hausdorff distance is introduced in section 3. The new Hausdorff ARTMAP is proposed in section 4. The preprocessing of the face images is described in section 5. In section 6, the experimental results are demonstrated and discussed.

2 Fuzzy ARTMAP

Fuzzy ARTMAP [10, 11, 12, 13, 14] is a supervised neural network composed of two fuzzy ART modules (ART_a and ART_b). The two fuzzy ARTs are linked together via the map field F^{ab} . Each fuzzy ART module has two layers: F_1^a and F_1^b are the input layers of each module, while F_2^a and F_2^b are the category layers. During the training phase, the

input vector and its desired output vector are presented to ART_a and ART_b respectively. The ART_a and ART_b modules classify the input and desired output vector into categories, then the map field uses a vigilance criterion to evaluate whether ART_a categories corresponds to ART_b categories. The criterion is as follows:

$$\frac{|y^b \wedge w_j^{ab}|}{|y^b|} \geq \rho_{ab} \quad (1)$$

where y^b is the output vector of ART_b , J is the index of the winning node in F_2^a , w_j^{ab} is the weights of the connections from the J^{th} node in F_2^a to the map field F^{ab} , and ρ_{ab} is the vigilance parameter of the map field. If the criterion is not respected, the vigilance parameter of ART_a , ρ_a , will be increased by a minimum value (just enough to force ART_a module to search for another category). However, when the vigilance criterion is respected, the weight vector is updated according to the equation:

$$w_j^{ab} = \beta_{ab} x^{ab} + (1 - \beta_{ab}) w_j^{ab} \quad (2)$$

where β_{ab} is the learning rate, which is between 0 and 1. x^{ab} is the output vector of F^{ab} .

3 Hausdorff Distance

The Hausdorff distance, when used as a measure of similarity between two two-dimensional binary patterns, has shown to agree closely with human performance [15]. The Hausdorff distance measures the extent to which each point of an input pattern lies near some point of a reference pattern. Given two finite sets $A = \{a_1, a_2, \dots, a_p\}$ and $B = \{b_1, b_2, \dots, b_q\}$, the Hausdorff distance between sets A and B is defined as:

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (3)$$

where the function $h(A, B)$ is called the directed Hausdorff distance from set A to set B , which can be computed as follows:

$$h(A, B) = \max_{a \in A} \{ \min_{b \in B} (\|a - b\|) \} \quad (4)$$

where $\|a - b\|$ is the Euclidean distance between point a and point b . The Hausdorff distance exhibits many desirable properties for pattern recognition. However, some modifications of the directed Hausdorff distance are made in this study in order to increase the noise immunity of the measurement.

$$h(A, B) = \frac{\sum_{a \in A} h(a, B)}{|A|} \quad (5)$$

where $|A|$ is the number of points in set A . $h(a, B)$ is the pointwise Hausdorff distance for point a . The pointwise Hausdorff distance is computed as follows:

$$h(a, B) = \min_{b \in B} (\|a - b\|) \quad (6)$$

4 Hausdorff ARTMAP

The architecture of the Hausdorff ARTMAP is a three-layer neural network as shown in Figure 1. The first layer is the input layer, which consists of $X \times Y$ nodes. Each node represents a pixel in the input pattern. The second layer is the category layer. The nodes in this second layer are constructed during the training phase. The third layer is the output layer. Each node in the output layer represents a class that the Hausdorff ARTMAP has to learn to recognize. During the supervised learning, the binary input pattern I^m is presented to the model, together with its respective target output vector. The input pattern is denoted by

$$I_{x,y}^m = \{1, 0\} : x=1, 2, \dots, X; y=1, 2, \dots, Y \quad (7)$$

where m is the m^{th} input pattern. X and Y are the dimensions of the input pattern.

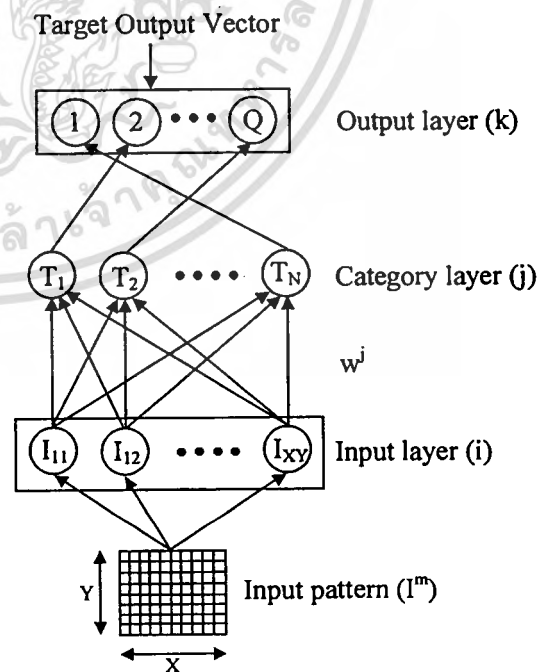


Figure 1 Architecture of the Hausdorff ARTMAP

Each node in the category layer is fully connected to the nodes in the input layer via the connections w^j . The weight w^j , which has the same dimension as the input pattern, represents the reference pattern of the j^{th} node in the category layer. Once the input pattern is transmitted to the category layer, the choice function of each j^{th} node in the category layer is evaluated as follows:

$$T_j(I^m) = H(I^m, w^j) \quad (8)$$

where $H(I^m, w^j)$ is the Hausdorff distance between the input pattern I^m and the reference pattern of the j^{th} node w^j . The system then makes a category choice by selecting the winning node J with minimum choice function value, among all the nodes j in the cluster layer. The category choice is indexed by J , where

$$T_j(I^m) = \min\{T_j(I^m)\} : j=1, 2, \dots, N \quad (9)$$

where N is the number of nodes in the category layer. In case of a tie, the node with the smallest index is chosen. Next, the vigilance criterion is evaluated to check whether the degree of mismatch between the input pattern and the reference pattern of the chosen category is within an acceptable level.

$$T_j(I^m) \leq \rho \quad (10)$$

where ρ is the vigilance parameter, which has the value between 0 and the length of the diagonal line. Resonance will occur if the chosen category meets the above criterion. However, if the condition in (10) is not satisfied, a new category node J is recruited to code the input pattern. The weight of this new node is initialized to be equal to the input pattern and this new node will automatically satisfy (10).

$$w^j = I^m \quad (11)$$

Next, the system associates the winning node J in the category layer with the target output vector. If the winning node J does not belong to the correct class defined by the target output vector, a new category node J is recruited and its weight is initialized using equation (11). Then the connection between a new category node and the target output is created. However, if the winning node represents the class to which I^m belongs, the weight vector w^j is then updated according to

$$b_{x,y}^j = \sum_{u=-S}^S \sum_{v=-S}^S I_{x+u, y+v}^m \quad (12)$$

$$w_{x,y}^{j \text{ new}} = \begin{cases} w_{x,y}^{j \text{ old}} & \text{if } b_{x,y}^j > 0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where S is a positive integer value; it tells the model how many pixels surrounding the current x, y position should be considered during the weight adjustment.

During testing, each test pattern is applied in turn and its class is predicted. The class whose category node returns the minimum output value is the result of the prediction.

5 Preprocessing

In this study, the performance of the Hausdorff ARTMAP is compared to the previous research work conducted by Lin et al. [16]. Therefore, in order to be comparable, the preprocessing step of this study replicated that of Lin et al.'s study. First, the gray-level edge image $E(x, y)$ is obtained by applying morphological operations [17] on the original face image $f(x, y)$. Then, the gray-level edge image is converted to a binary edge image using the adaptive threshold method.

$$n(x, y) = \frac{E(x, y)}{f(x, y)}$$

The values of the function $n(x, y)$ are then sorted in descending order, and the threshold is set so that 30% of the points with the largest magnitudes in $n(x, y)$ are selected.



Figure 2 Examples of the ORL and MIT Face Databases (a) the Original Facial Images; (b) the Binary Edge Images.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6 Experimental Results

To test the performance of the Hausdorff ARTMAP for face recognition, the experiments have been conducted on 2 databases: the ORL face database [18][19] (from the Olivetti Research Laboratory in Cambridge, UK.) and the MIT face database [20][21] (from the Vision and Modeling Group at the MIT Media Laboratory). Previously, Lin et al. [16] carried out a comparison study of many different techniques – the principal component analysis (PCA), the conventional Hausdorff distance (HD), the doubly modified Hausdorff distance (M2HD), the spatially weighted Hausdorff distance (SWHD), the spatially weighted doubly Hausdorff distance (SW2HD), the spatially eigen-weighted Hausdorff distance (SEWHD), and the spatially eigen-weighted doubly Hausdorff distance (SEW2HD) – on both ORL database and MIT database. The reported recognition rates based on the ORL database varied from 46 – 91%. The best recognition rate (91%) was achieved from the SEW2HD. For the MIT database, the best performance reported in their study was 92% which was also obtained from the SEW2HD.

For the ORL database, there are 10 different images of each of 40 distinct subjects; therefore, the total of 400 different face images are used in this experiment. The images of each subject were taken at different times with different lighting, facial expressions (open/closed eyes, smiling/non-smiling), and facial details (glasses/no-glasses). Four images of each subject are randomly chosen for training and the remaining six images are used for testing. The images used for training are shown in Table 1. Table 2 tabulates the recognition results of the Hausdorff ARTMAP on each subject of the ORL database. This best recognition performance (95%) is obtained when the vigilance parameter (ρ) is set at 0.8 and S is 2.

The MIT database contains 432 different face images corresponding to 16 distinct subjects. The subjects are all male. Each subject was digitized 27 times, varying the head orientation, the lighting, and the scale. In this experiment, randomly picked 144 images (9 images per subject) are included in the testing set. The Hausdorff ARTMAP with $\rho = 0.6$ and $S = 2$ gives a perfect accuracy (100%) on this database.

References:

[1] A. M. Martínez, M. Yang, and D. J. Kriegman, Special Issue on Face Recognition, *Computer*

Table 1 The Training Images of Each Subject in the ORL Database

Subject	Pattern Number	Training Pattern
S1	1 – 10	2, 3, 6, 8
S2	11 – 20	12, 14, 15, 16
S3	21 – 30	21, 22, 25, 28
S4	31 – 40	31, 32, 34, 40
S5	41 – 50	42, 43, 46, 50
S6	51 – 60	51, 56, 57, 59
S7	61 – 70	61, 63, 64, 70
S8	71 – 80	73, 76, 79, 80
S9	81 – 90	83, 87, 89, 90
S10	91 – 100	94, 98, 99, 100
S11	101 – 110	102, 104, 105, 110
S12	111 – 120	113, 116, 117, 118
S13	121 – 130	124, 126, 128, 130
S14	131 – 140	131, 133, 139, 140
S15	141 – 150	142, 144, 145, 148
S16	151 – 160	157, 158, 159, 160
S17	161 – 170	164, 165, 167, 170
S18	171 – 180	172, 175, 176, 177
S19	181 – 190	183, 184, 186, 188
S20	191 – 200	191, 194, 198, 200
S21	201 – 210	204, 205, 207, 210
S22	211 – 220	212, 214, 216, 218
S23	221 – 230	225, 227, 229, 230
S24	231 – 240	231, 233, 237, 238
S25	241 – 250	241, 242, 248, 249
S26	251 – 260	252, 253, 254, 257
S27	261 – 270	261, 262, 267, 270
S28	271 – 280	272, 274, 276, 280
S29	281 – 290	282, 285, 288, 289
S30	291 – 300	293, 294, 295, 296
S31	301 – 310	302, 307, 308, 310
S32	311 – 320	311, 314, 317, 318
S33	321 – 330	321, 325, 326, 328
S34	331 – 340	331, 334, 337, 338
S35	341 – 350	342, 343, 346, 349
S36	351 – 360	352, 355, 357, 360
S37	361 – 370	361, 366, 367, 370
S38	371 – 380	376, 377, 378, 379
S39	381 – 390	382, 384, 385, 390
S40	391 – 400	392, 393, 397, 400

Vision and Image Understanding, Vol. 91, Issue 1-2, Academic Press, 2003, pp. 1-5.

- [2] R. Chellappa, C. L. Wilson, and S. Sirohey, Human and Machine Recognition of Faces: A Survey, *Proceedings of the IEEE*, Vol. 83, No. 5, May 1995, pp. 705-740.
- [3] M. D. Kelly, Visual Identification of People by Computer, *Technical Report AI-130*, Stanford AI Project, Stanford, CA, 1970.

Table 2 The Recognition Results for the ORL Face Database

Subject Name	Number of Correct Pattern	Number of Incorrect Pattern	Incorrectly Assigned Subject
S1	5	1	S38
S2	6	0	-
S3	4	2	S40, S38
S4	5	1	S40
S5	6	0	-
S6	6	0	-
S7	6	0	-
S8	6	0	-
S9	5	1	S30
S10	5	1	S30
S11	6	0	-
S12	6	0	-
S13	5	1	S5
S14	6	0	-
S15	6	0	-
S16	6	0	-
S17	6	0	-
S18	6	0	-
S19	6	0	-
S20	6	0	-
S21	6	0	-
S22	6	0	-
S23	5	1	S38
S24	6	0	-
S25	6	0	-
S26	6	0	-
S27	6	0	-
S28	6	0	-
S29	6	0	-
S30	6	0	-
S31	6	0	-
S32	4	2	S19, S8
S33	6	0	-
S34	6	0	-
S35	6	0	-
S36	6	0	-
S37	6	0	-
S38	6	0	-
S39	4	2	S29, S8
S40	6	0	-
Total	228	12	

[4] M. A. Turk and A. P. Pentland, Face Recognition Using Eigenfaces, *Proceedings of the International Conference on Pattern Recognition*, 1991, pp. 586-591.

[5] S. Akamatsu, T. Sasaki, H. Fukamachi, and Y. Suenaga, A Robust Face Identification Scheme – KL Expansion of an Invariant Feature Space, *SPIE Proc.: Intelligent Robots and Computer*

Vision X: Algorithms and Techniques, Vol. 1607, 1991, pp. 71-84.

[6] Y. Cheng, K. Liu, J. Yang, and H. Wang, A Robust Algebraic Method for Human Face Recognition, *Proceedings of 11th International Conference on Pattern Recognition*, 1992, pp. 221-224.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [7] S. Y. Kung, S. H. Lin, and M. Fang, A Neural Network Approach to Face/Palm Recognition, *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing*, MA, USA, 1995, pp. 323-332.
- [8] H. M. El-Bakry and M. A. Abo Elsoud, Human Face Recognition Using Neural Networks, *Proceedings of 16th National Radio Science Conference (NRSC'99)*, Cairo, Egypt, February 1999.
- [9] S. Lin, S. Kung, and L. Lin, Face Recognition/Detection by Probabilistic Decision-Based Neural Network, *IEEE Transactions on Neural Networks*, Vol. 8, No. 1, January 1997, pp. 114-132.
- [10] G. A. Carpenter and S. Grossberg, A Self-Organizing Neural Network for Supervised Learning, Recognition, and Prediction, *IEEE Communications Magazine*, September 1992, pp. 38-49.
- [11] G. Carpenter, S. Grossberg, and K. Iizuka, Comparative Performance Measures of Fuzzy ARTMAP, Learned Vector Quantization, and Back Propagation for Handwritten Character Recognition, *Proceedings of the International Joint Conference on Neural Networks*, Vol. 1, June 1992, pp. 794-799.
- [12] N. Srinivasa, Learning and Generalization of Noisy Mappings Using a Modified PROBART Neural Network, *IEEE Transactions on Signal Processing*, Vol. 45, No. 10, October 1997.
- [13] C. P. Lim and R. F. Harrison, Probabilistic Fuzzy ARTMAP: An Autonomous Neural Network Architecture for Bayesian Probability Estimation, *Proceedings of Fourth International Conference on Artificial Neural Networks*, June 1995, pp. 148-153.
- [14] Y. H. Tay and M. Khalid, Comparison of Fuzzy ARTMAP and MLP Neural Networks for Handwritten Character Recognition, *Proceedings of International Federation of Automatic Control (IFAC) Symposium on Artificial Intelligent in Real-Time Control*, Kuala Lumpur, September 1997, pp. 363-371.
- [15] R. G. Rosandich, HAVNET: A New Neural Network Architecture for Pattern Recognition, *Neural Networks*, Vol. 10, No. 1, Pergamon Press, 1997, pp. 139-151.
- [16] K. Lin, K. Lam, and W. Siu, Spatially Eigen-weighted Hausdorff Distances for Human Face Recognition, *Pattern Recognition*, Vol. 36, Issue 8, Pergamon Press, 2003, pp. 1827-1834.
- [17] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 2002.
- [18] F. Samaria and A. Harter, Parameterisation of a Stochastic Model for Human Face Identification, *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, December 1994.
- [19] ORL Face Database, Retrieved from <http://www.uk.research.att.com/facedatabase.html>
- [20] M. Turk and A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991.
- [21] MIT Face Database, Retrieved from <ftp://whitechapel.media.mit.edu/pub/images/>

ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาวจงกลณี รุ่งเรือง .
วัน เดือน ปีเกิด	5 พฤษภาคม 2522 ที่จังหวัดตรัง
ที่อยู่	46/3 หมู่ 10 ต.ละมอ อ.นาโยง จ.ตรัง 92170 โทร. 0-7524-2059
ประวัติการศึกษา	วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยสงขลานครินทร์ 2544



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้