บทที่ 3

Modified Hausdorff ARTMAP for Binary Image Recognition

3.1 บทนำ

ลักษณะทางกายภาพ และพฤติกรรมของมนุษย์ทางกายภาพของมนุษย์ ถูกนำมาประยุกต์ใช้อย่างแพร่หลายในปัจจุบัน ทั้งในส่วนของระบบสมาชิก การเข้า-ออกเวลาทำงาน รวมทั้งระบบรักษาความปลอดภัยในสนามบิน หรือธนาคาร โดยเทคนิคการระบุบุคคลโดยใช้คุณสมบัติทางกายภาพ หรือพฤติกรรมของมนุษย์เรียกว่า Biometrics [6] คุณสมบัติทางกายภาพ หรือพฤติกรรมของมนุษย์ที่จัดอยู่ใน Biometrics สามารถแบ่งออกเป็น 2 กลุ่ม [7] คือ กลุ่มคุณสมบัติทางกายภาพ (physiological) ประกอบด้วย รูม่านตา (retinal/iris) ใบหน้า (face) ลายนิ้วมือ (fingerprint) และกลุ่มคุณสมบัติทางพฤติกรรม ประกอบด้วย ลายเซ็น (signature) และเสียง (voice)

งานวิจัยชิ้นนี้ใช้ข้อมูลใบหน้า ลายนิ้วมือ และลายเซ็น โดยทั้ง 3 ชุดข้อมูลนี้ต้องเข้าสู่กระบวนการเตรียมข้อมูลก่อนการใช้งานจริง (pre-processing) ซึ่งข้อมูลแต่ละชุดมีวิธีการเตรียมข้อมูลที่แตกต่างกัน และได้นำเสนอในหัวข้อถัดไป

3.2 การเตรียมข้อมูล

การเตรียมข้อมูล (pre-processing) คือ กระบวนการจัดการกับข้อมูลดิบ (raw input) ให้มีความเหมาะสมกับอัลกอริทึมที่ต้องการนำเสนอ โดยอาจประกอบด้วยกระบวนการต่างๆ เช่น การกำจัดจุดรบกวน (noise) การปรับค่าของข้อมูล หรือปรับขนาดของข้อมูล (normalization) การหาโครงร่างรูปภาพ (skeleton)

3.2.1 การเตรียมข้อมูลภาพใบหน้า

รูปภาพใบหน้าในชุดข้อมูลเดิมเป็นภาพระดับสีเทา (gray scale image) มีความจำเป็นที่ต้องเปลี่ยนมาเป็นภาพขาวดำ (binary image) เพื่อใช้ในงานวิจัยนี้ และการหาเส้นขอบของภาพใบหน้าก็มีความสำคัญเช่นกัน กระบวนการเตรียมข้อมูลภาพใบหน้ามีดังต่อไปนี้

3.2.1.1 การหาตำแหน่งอ้างอิง

ตาทั้งสองข้างถูกใช้เป็นตำแหน่งอ้างอิง วิธีการหาตำแหน่งเป็นแบบ manual คือการหาตำแหน่งของดวงตาทั้งสองข้าง โดยคนเป็นผู้หา ตำแหน่งอ้างอิงที่ได้แสดงในรูปที่ 3.1

รูปที่ 3.1 แสดงตำแหน่งอ้างอิงของภาพใบหน้า

3.2.1.2 การหาเส้นขอบของภาพ และการแปลงภาพระดับสีเทาเป็นภาพขาวดำ

การใช้เส้นขอบของภาพแทนภาพใบหน้าทั้งหมด เพื่อแก้ปัญหาการเปลี่ยนแปลงความเข้มแสงที่ตกกระทบ (illumination) การแปลงภาพจากภาพระดับสีเทา เป็นภาพลายเส้นระดับสีเทา (gray edge image) E(x,y) โดยใช้ morphological operation [8] กับภาพใบหน้า f(x,y) ขั้นตอนการแปลงภาพแสดงได้ดังนี้

$$(f \oplus b) - (f \ominus b) = E \tag{3.1}$$

ในการดำเนินการวิจัยนี้ได้นำเอาภาพใบหน้าต้นฉบับทำ dilation ภาพใบหน้าด้วย structure element ที่มีลักษณะที่เป็นสี่เหลี่ยมจัตุรัสขนาด 3×3 จากนั้นเอาภาพใบหน้าต้นฉบับทำ erosion ภาพใบหน้าด้วย structure element แบบเดียวกัน นำภาพที่ได้จากทำ dilation ตั้งแล้วลบ (subtraction) ด้วยภาพที่ได้จากการทำ erosion โดยภาพที่ได้จากการทำ subtraction คือภาพลายเส้นระดับสีเทาที่ต้องการ จากนั้นใช้วิธี adaptive thresholding ใน [9] เพื่อแปลงภาพลายเส้นระดับสีเทาเป็นภาพขาวดำที่ต้องการ โดยวิธีการใน [9] ให้ความสำคัญกับใบหน้า และโครงหน้าซึ่งมักจะเป็นส่วนที่มีความลึก และค่าความเข้มของระดับสีเทา (gray level intensities) ต่ำกว่าส่วนอื่นๆ ของใบหน้า ฟังก์ชันที่ [9] เลือกใช้คือ

$$n(x,y) = \frac{E(x,y)}{f(x,y)} \tag{3.2}$$

นำค่า $n(x,y)$ ที่ได้จากสมการ (3.2) มาเรียงลำดับจากมากไปหาน้อย จากนั้นกำหนดค่า threshold เพื่อเลือกจุดที่มีค่า $n(x,y)$ มากที่สุดตามลำดับที่ได้เรียงไว้ โดยเลือกจำนวนจุดตามเปอร์เซ็นที่ต้องการ ในงานวิจัยนี้เลือกจุดที่มีค่า $n(x,y)$ มากที่สุด 30% แรกตามลำดับที่เรียงไว้ โดยให้จุดนั้นมีค่าเท่ากับ 1 ส่วนจุดที่เหลือให้มีค่าเท่ากับ 0 ซึ่งจะได้ภาพขาวดำตามต้องการ รูปที่ 3.2

แสดงตัวอย่างในการแปลงภาพระดับสีเทา เป็นภาพขาวดำ รูป (a) แสดงภาพใบหน้าต้นฉบับเป็นภาพระดับสีเทา รูป (b) แสดงภาพใบหน้าแบบลายเส้นระดับสีเทา และรูป (c) แสดงภาพใบหน้าขาวดำ

(a) (b) (c)

รูปที่ 3.2 (a) ภาพใบหน้าต้นฉบับ (b) ภาพใบหน้าแบบลายเส้นระดับสีเทา (c) ภาพใบหน้าขาวดำ

3.2.1.3 การหมุนภาพ

การกำหนดให้ตาทั้งสองข้างเป็นจุดอ้างอิง ทำให้สามารถที่คำนวณมุมระหว่างจุด 2 จุดนี้ได้ สมมติฐานในการหมุนภาพคือ ตาของเราทั้งสองข้างควรวางอยู่ในแนวเดียวกัน ดังนั้นเมื่อได้มุมระหว่างตาทั้งสองแล้ว ทำการหมุนภาพเพื่อให้จุดอ้างอิงทั้งสองอยู่ในตำแหน่งที่ขนานกับแนวระนาบ ทฤษฎีสำหรับการหมุนภาพที่ใช้คือ ทฤษฎี nearest neighbor interpolation จากนั้นหาจุดอ้างอิงใหม่จากภาพที่หมุนแล้ว

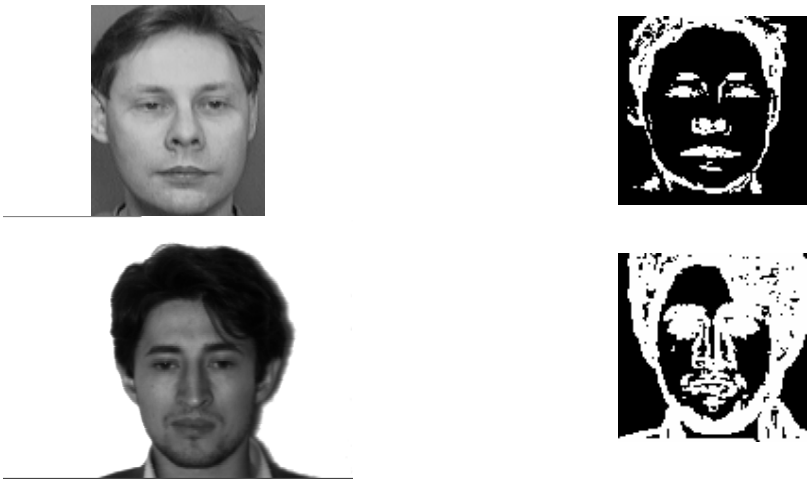
3.2.1.4 การปรับขนาดภาพ

จาก 3.2.1.3 ได้จุดอ้างอิงใหม่จากภาพที่หมุนแล้ว ขั้นตอนต่อไปคือการคำนวณระยะห่างระหว่างจุดอ้างอิงทั้งสอง เทียบกับระยะห่างที่ต้องการ ซึ่งในงานวิจัยนี้กำหนดให้ทุกภาพมีระยะระหว่างดวงตาทั้งสองข้างเท่ากับ 30 พิกเซล จากนั้นคำนวณหาอัตราส่วนของระยะระหว่างดวงตาของภาพที่หมุนแล้ว กับระยะระหว่างดวงตาที่ต้องการ อัตราส่วนนี้คือจำนวนเท่าที่ต้องการปรับขนาด งานวิจัยนี้เลือกใช้ทฤษฎี nearest neighbor interpolation ในการปรับขนาดภาพ จากนั้นทำการหาจุดอ้างอิงใหม่จากภาพใบหน้าที่ปรับขนาดแล้ว

3.2.1.5 การเลื่อนตำแหน่งอ้างอิงของภาพใบหน้า และการเลือกบริเวณของภาพในส่วนที่ต้องการ

ขั้นตอนนี้ทำการเลื่อนจุดอ้างอิงทั้งสองให้อยู่ในบริเวณที่เหมาะสม และอยู่ในตำแหน่งเดียวกันทุกรูป โดยให้ตาซ้ายอยู่ที่ตำแหน่ง (45,60) จากนั้นปรับขนาดภาพให้มีขนาด 100×100 พิกเซล เท่ากันทุกภาพ โดยตัดส่วนเกินของภาพออกหากมีขนาดเกินจากขนาดที่กำหนด และหากภาพที่ได้มีขนาดน้อยกว่าที่กำหนดไว้จะทำการเติมส่วนที่ขาดไปให้เป็นช่องว่าง (มีค่าเป็นศูนย์)

ภาพใบหน้าที่ผ่านกระบวนการเตรียมข้อมูลแสดงตัวอย่างในรูปที่ 3.3 ในคอลัมน์ (a) คือภาพใบหน้าต้นฉบับ และในคอลัมน์ (b) คือภาพใบหน้าหลังจากกระบวนการเตรียมข้อมูล



(a) (b)

รูปที่ 3.3 (a) ภาพใบหน้าต้นฉบับ (b) ภาพใบหน้าหลังจากกระบวนการเตรียมข้อมูล

3.2.2 การเตรียมข้อมูลภาพลายนิ้วมือ

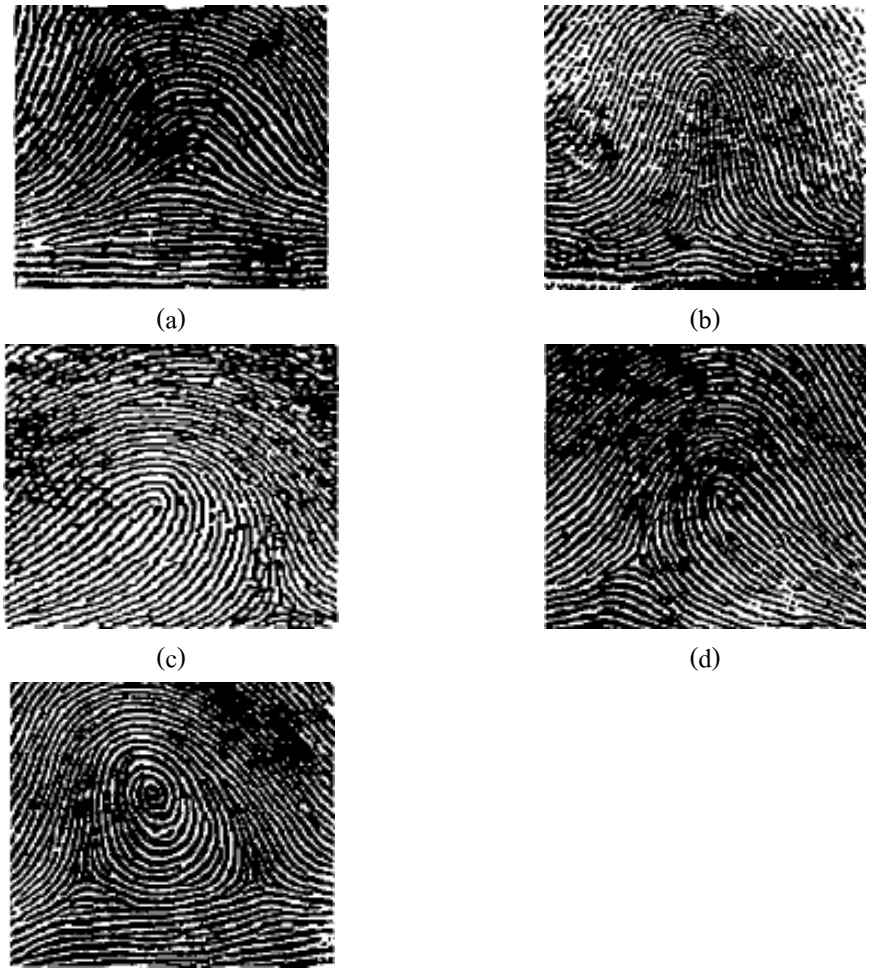
ขั้นตอนสำคัญหลัก 2 ขั้นตอนในการเตรียมข้อมูลภาพลายนิ้วมือ คือ ขั้นตอนการหาตำแหน่งอ้างอิง ซึ่งจะนำมาใช้เป็นจุดศูนย์กลางของภาพ และขั้นตอนการแปลงภาพจากภาพระดับสีเทา เป็นภาพแบบขาวดำ หรือภาพขาวดำ

ขั้นตอนการเตรียมข้อมูลภาพลายนิ้วมือมีขั้นตอนต่างๆ ดังนี้

3.2.2.1 การหาจุดอ้างอิง

รูปลายนิ้วมือแต่ละรูปมีการจัดวางตำแหน่งที่ไม่เหมือนกัน ดังนั้นจึงจำเป็นต้องหาตำแหน่งอ้างอิงของรูปลายนิ้วมือทุกๆ รูป และใช้ตำแหน่งอ้างอิงนี้เป็นจุดศูนย์กลางของภาพ

การหาตำแหน่งอ้างอิงทำได้โดยเลือกใช้ลักษณะพิเศษ (feature) ที่มีอยู่ในลายนิ้วมือ คือ จุดซิงกูลาร์ (singular points : SPs) ซึ่งประกอบด้วยจุด 2 ประเภท คือ จุดคอร์ (core) และจุดเดลตา (delta) จุดคอร์ คือ จุดที่เป็นส่วนบนสุดของเส้นโค้งที่อยู่ในสุด จุดเดลตา คือ จุดศูนย์กลางของพื้นที่สามเหลี่ยม ซึ่งมีทิศทางของเส้นในสามทิศทางมาบรรจบกัน [10] ลายนิ้วมือโดยทั่วไปมีอยู่ 5 แบบ [11] คือ ลายนิ้วมือแบบโค้ง ลายนิ้วมือแบบโค้งกระโจม ลายนิ้วมือแบบมัดหวายปัดซ้าย ลายนิ้วมือแบบมัดหวายปัดขวา ลายนิ้วมือแบบก้นหอย ดังแสดงในรูปที่ 3.4 และรูปที่ 3.5 แสดงลักษณะพิเศษบนลายนิ้วมือ



รูปที่ 3.4 (a) ลายนิ้วมือแบบโค้ง (b) ลายนิ้วมือแบบโค้งกระโจม (c) ลายนิ้วมือแบบมัดหวายปัดซ้าย (d) ลายนิ้วมือแบบมัดหวายปัดขวา (e) ลายนิ้วมือแบบก้นหอย

รูปที่ 3.5 แสดงลักษณะพิเศษบนลายนิ้วมือ

งานวิจัยนี้ใช้วิธีการหาจุดซิงกูลาร์ตามแนวทางใน [12] การหาจุดซิงกูลาร์ใน [12] อาศัยการเปลี่ยนแปลงของเครื่องหมายในรูปภาพเชิงทิศทาง (directional image) ซึ่งสามารถหาจุดซิงกูลาร์ได้อย่างถูกต้อง และใช้เวลาในการคำนวณน้อย วิธีการใน [12] ถูกนำไปทดสอบกับฐานข้อมูลลายนิ้วมือ FINGDB [13] แล้วให้ผลลัพธ์ดี และงานวิจัยนี้เลือกใช้ฐานข้อมูลลายนิ้วมือ FINGDB ในการทดสอบด้วยเช่นกัน

3.2.2.2 การแปลงภาพลายนิ้วมือระดับสีเทา เป็นภาพขาวดำ

งานวิจัยนี้ใช้วิธีการในการแปลงภาพระดับสีเทา เป็นภาพแบบขาวดำที่เรียกว่า Regional Average Thresholding (RAT) [14] เป็นการหาค่าเฉลี่ยในแต่ละบริเวณมาเป็นค่า threshold แนวคิดของ RAT มาจากการแก้ปัญหาการแปลงภาพระดับสีเทา เป็นภาพขาวดำ โดยการใช้ค่า threshold เพียงค่าเดียว ซึ่งทำให้เกิดการสูญเสียฟีเจอร์ (feature) บางอย่าง วิธีการที่ใช้ค่า threshold เพียงค่าเดียวคือ General Thresholding (GT) สามารถแสดงสมการได้ดังนี้

$$P(i, j) = \begin{cases} 1, & \text{if } I(i, j) > T \\ 0, & \text{if } I(i, j) \le T \end{cases} \tag{3.3}$$

เมื่อ $I(i, j)$ คือ รูปภาพเริ่มต้นที่เป็นภาพขาวเทา

$P(i, j)$ คือ ภาพขาวดำที่ได้จากการแปลง

T คือ ระดับของ threshold

$i = 0, 1, \ldots, N$

$j = 0, 1, \ldots, M$

ขนาดภาพ $= N \times M$

วิธีการ RAT จะแบ่งภาพเริ่มต้นออกเป็นส่วนหน้าต่างย่อยๆ โดยใน [14] แบ่งออกขนาด 16×16 พิกเซล จากนั้นหาค่าเฉลี่ยของระดับค่าสีเทาในแต่ละหน้าต่าง ค่าเฉลี่ยของระดับค่าสีเทาหาได้จากสมการที่ 3.4

$$T = \frac{1}{N^2}\sum_{i=0}^{N}\sum_{j=0}^{N} I(i, j) \tag{3.4}$$

วิธีการ RAT นี้สามารถแก้ปัญหาที่เกิดขึ้นกับวิธี GT ได้ และยังคงลายเส้นของลายนิ้วมือได้ดีกว่าวิธี GT ด้วย ในรูปที่ 3.6 แสดงภาพการแปลงลายนิ้วมือจากภาพระดับสีเทา เป็นภาพขาวดำ

(a) (b)

รูปที่ 3.6 (a) ภาพลายนิ้วมือขาวดำที่ได้จากวิธี RAT (b) ภาพลายนิ้วมือขาวดำที่ได้จากวิธีการ GT

3.2.2.3 การปรับขนาดภาพ และการเลือกภาพลายนิ้วมือ

การปรับขนาดภาพงานวิจัยนี้เลือกใช้จุดซิงกูลาร์ที่หาได้จากขั้นตอนที่ 3.2.2.1 โดยหากจุดซิงกูลาร์ที่หาได้มีทั้งจุดคอร์ และจุดเดลตา งานวิจัยนี้จะเลือกใช้จุดคอร์เป็นจุดอ้างอิง แต่หากจุดซิงกูลาร์ที่หาได้มีเพียงจุดคอร์ หรือจุดเดลตา อย่างใดอย่างหนึ่ง จะเลือกจุดนั้นเป็นจุดอ้างอิง

เมื่อได้จุดซิงกูลาร์ที่ใช้เป็นจุดอ้างอิงแล้ว ทำการเลือกภาพในพื้นที่รอบจุดนี้ ขนาด 100×100 พิกเซล เป็นภาพที่ใช้งานจริง เนื่องจากบริเวณรอบจุดซิงกูลาร์เป็นบริเวณที่มีความสำคัญมาก (significance) มากที่สุด รูปที่ 3.7 เป็นตัวอย่างภาพลายนิ้วมือที่ไม่สามารถนำมาใช้งานได้

(a) (b)

รูปที่ 3.7 (a) ภาพลายนิ้วมือที่ไม่สามารถหาจุดซิงกูลาร์ที่แน่นอนได้ (b) ภาพลายนิ้วมือที่จุดอ้างอิงน้อยกว่าขนาดที่กำหนด

3.2.2.4 การทำ thinning

การทำ thinning คือ การขจัดส่วนเกินที่ยื่นออกมาจากตัวเส้น ซึ่งทำให้ลายเส้นมีความเรียบมากขึ้น การทำ thinning ทำตามหลักการใน [15] ตัวอย่างภาพลายนิ้วมือที่ทำ thinning แล้วแสดงในรูปที่ 3.8

(a) (b)

รูปที่ 3.8 (a) ภาพลายนิ้วมือต้นแบบ (b) ภาพลายนิ้วมือแบบขาวดำที่ทำ thinning แล้ว

3.2.3 การเตรียมข้อมูลภาพลายเซ็น

การเตรียมข้อมูลภาพลายเซ็นจำเป็นที่จะต้องทำ skeleton การตัดพื้นที่รอบข้างลายเซ็นที่ไม่ได้ใช้งาน รวมทั้งทำการปรับขนาดภาพลายเซ็นให้มีขนาดที่เท่ากันทั้งหมดด้วย การเตรียมข้อมูลลายเซ็นมีขั้นตอนต่างๆ ดังนี้

3.2.3.1 การนำภาพลายเซ็นเข้าคอมพิวเตอร์

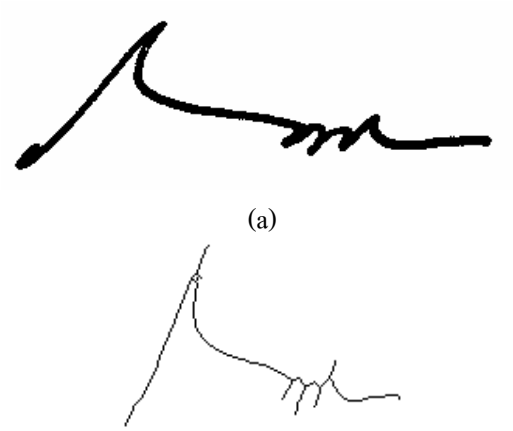
ลายเซ็นทั้งหมดถูกเซ็นบนกระดาษขนาด A4 สีขาว ในการนำภาพลายเซ็นเข้าคอมพิวเตอร์ที่งานวิจัยนี้เลือกใช้คือการใช้เครื่องสแกนเนอร์ การสแกนภาพใช้ความละเอียด 300 dpi และเลือกประเภทภาพเป็นภาพขาวดำ

3.2.3.2 การตัดขอบภาพ และการปรับขนาดภาพ

หลังจากที่เราทำการสแกนภาพลายเซ็นแล้ว ขั้นตอนถัดไปคือการตัดพื้นหลัง (background) รอบนอกในส่วนเกินจากลายเซ็น เพื่อให้ภาพมีขนาดที่พอดีกับลายเซ็น จากนั้นทำการปรับขนาดภาพให้มีขนาดเท่ากับ 100×150 พิกเซล เท่ากันทุกภาพ

3.2.3.3 การทำ skeleton

ขั้นตอนถัดมาคือการทำ skeleton เป็นการหาโครงสร้างในส่วนแกนกลางของลายเส้น ในรูปที่ 3.9 (a) แสดงภาพตัวอย่างลายเซ็นต้นฉบับ และภาพ (b) เป็นภาพลายเซ็นที่ทำ pre-processing แล้ว



(a)

(b)

รูปที่ 3.9 (a) ภาพลายเซ็นต้นฉบับ (b) ภาพลายเซ็นที่ผ่านการทำ pre-processing แล้ว

3.3 การย้ายจุดศูนย์กลางอ้างอิง

การย้ายจุดศูนย์กลางอ้างอิง หรือการสแกน คือการย้ายจุดศูนย์กลางจากตำแหน่งจุดศูนย์กลางของภาพเดิม ไปยังตำแหน่งรอบๆ จุดศูนย์กลางนั้นใช้ควบคู่กับการเปรียบเทียบภาพในอัลกอริทึมที่นำเสนอ การย้ายตำแหน่งเพื่อให้เกิดความถูกต้องมากที่สุด หากภาพนั้นมีการเคลื่อนออกจากตำแหน่งจุดศูนย์กลางภาพ

การสแกนจะเกิดจากการกำหนดค่าพารามิเตอร์จากผู้ใช้โดยพารามิเตอร์ที่ใช้มีดังนี้

- จำนวนพิกเซลที่จะขยายออกไปจากจุดศูนย์กลางรูปภาพในแนวแกน x (x_area)
- จำนวนพิกเซลที่จะขยายออกไปจากจุดศูนย์กลางรูปภาพในแนวแกน y (y_area)
- จำนวนพิกเซลที่ข้ามเพื่อกำหนดจุดศูนย์กลางใหม่จุดถัดไป (step)

ตัวอย่างการหาจุดศูนย์กลาง เรากำหนดให้ x_area = 6, y_area = 6 step = 2 เราจะทำการขยายพื้นที่การพิจารณาดังนี้ ขยายไปทางด้านซ้าย และด้านขวาของจุดศูนย์กลางภาพ (x_area) เป็นจำนวนด้านละ 6 พิกเซล ขยายไปทางด้านบน และด้านล่างของจุดศูนย์กลางภาพ (y_area) เป็นจำนวนด้านละ 6 พิกเซล เราจะได้พื้นที่การพิจารณาขนาด 13×13 พิกเซล^2 จากนั้นให้ พิกเซล ซ้ายบนเป็นจุดศูนย์กลางใหม่ที่จะพิจารณา และทำการเลื่อนไปทางขวาอีก 2 พิกเซล (step) เพื่อกำหนดจุดศูนย์กลางใหม่ ทำไปจนครบพื้นที่การพิจารณา ด้วยวิธีนี้เราจะได้จุดศูนย์กลางใหม่ทั้งหมด 25 จุด ดังรูปที่ 3.10 เส้นสีดำทึบ คือขนาดภาพ จุดตรงกลาง คือจุดศูนย์กลางภาพ และจุดทั้ง 25 จุด คือจุดศูนย์กลางที่ใช้เพื่อการพิจารณา

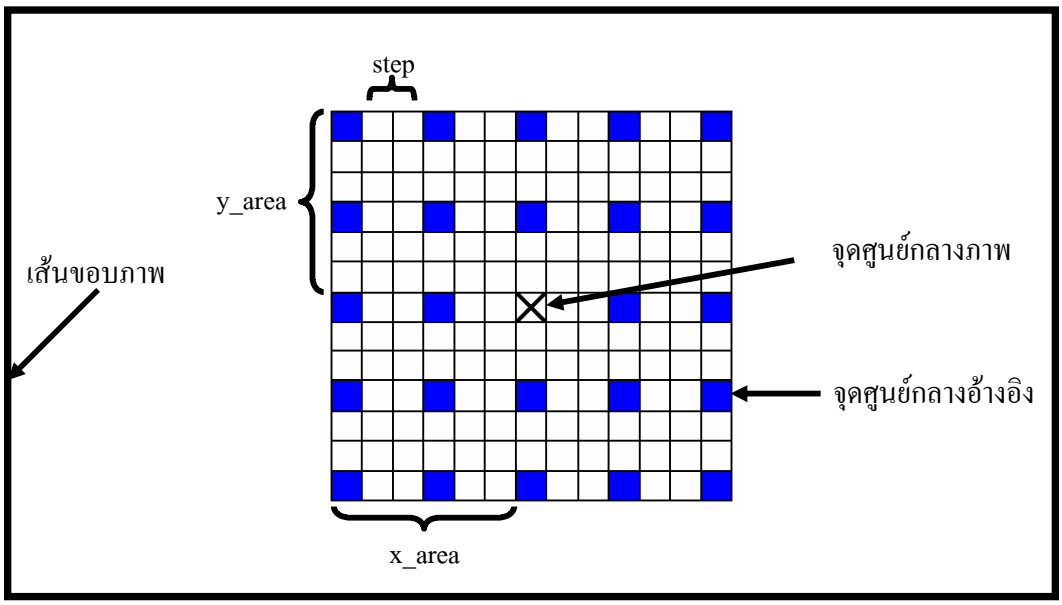
วิธีการคำนวณหาจำนวนจุดศูนย์กลาง สามารถหาได้ตามสมการนี้

$$e = \left[\left(2 \times \left(\frac{x_area}{(step+1)}\right)\right)+1\right] \times \left[\left(2 \times \left(\frac{y_area}{(step+1)}\right)\right)+1\right] \tag{3.5}$$

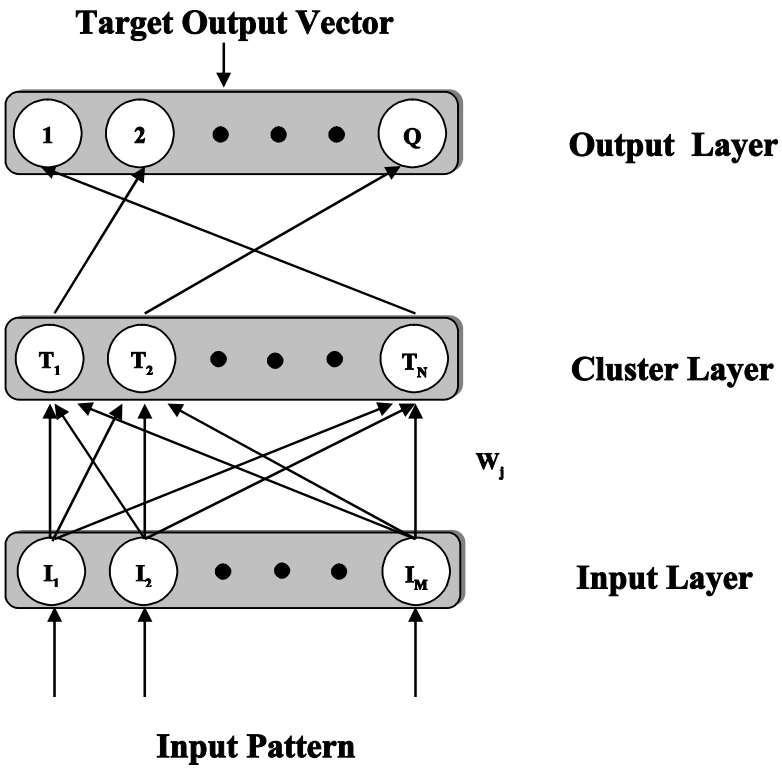
3.4 สถาปัตยกรรมของ Modified Hausdorff ARTMAP

กระบวนการการเรียนรู้นี้จะเป็นการนำเอาซิมพลิไฟด์ฟัซซีอาร์ทแมพมาใช้ โดยนำเอาฮาอูสดอร์ฟดิสแทนซ์มาปรับใช้แทนการดำเนินการแบบฟัซซี (Fuzzy Operation) ในส่วนของอินพุทนั้นจะรับเข้าสู่โครงข่ายในรูปแบบของภาพสองมิติ (two-dimensional images) และภาพนั้นเป็นภาพขาวดำเท่านั้น (Binary Image)

วิธีการนี้ประกอบด้วย 3 ระดับชั้น ชั้นแรกคือชั้นอินพุท (Input Layer) ชั้นสองคือชั้นคลัสเตอร์ (Cluster Layer) ชั้นสุดท้ายคือชั้นเอาท์พุท (Output Layer)



รูปที่ 3.10 แสดงจุดศูนย์กลางใหม่ที่ถูกกำหนดขึ้นเพื่อใช้ในการเปรียบเทียบภาพ



รูปที่ 3.11 สถาปัตยกรรมของ Modified Hausdorff ARTMAP

ชั้นที่ 1 Input Layer มีขนาด M โหนด ซึ่งทำหน้าที่รับ input pattern ขนาด M มิติ

ชั้นที่ 2 Cluster Layer โหนดในชั้นนี้จะถูกสร้างขึ้นมาในระหว่างกระบวนการเรียนรู้ แต่ละโหนดใน Cluster Layer จะเชื่อมต่อแบบทั่วถึงกับทุกโหนดใน Input Layer โดยจะเชื่อมต่อผ่านทางค่าน้ำหนัก ซึ่งค่าน้ำหนักจะมีขนาดไม่แน่นอน แต่จะมีขนาดใหญ่ไม่เกินขนาดภาพของ input pattern และค่าน้ำหนักนี้จะเป็นตัวแทนของรูปแบบอ้างอิงใน Cluster Layer

ชั้นที่ 3 Output Layer แต่ละโหนดในชั้นนี้จะเป็นตัวแทนแสดงแต่ละประเภท หรือชนิด (class) เนื่องจากกระบวนการเรียนรู้นี้เป็นแบบ supervised ดังนั้นคำตอบจะถูกส่งเข้าไปในโครงข่ายพร้อมกับอินพุท นั่นก็คือชนิดของรูปแบบอินพุท (target output vector) นั่นเอง โดยคำตอบที่ถูกส่งเข้าไปในโครงข่ายจะถูกส่งเข้าไปเฉพาะช่วงของการเรียนรู้เท่านั้น

ค่าน้ำหนักในกระบวนการนี้จะใช้ในลักษณะโมเดลอ้างอิง (Reference Model) โดยต้องการสร้างโมเดลให้ครอบคลุม และเป็นตัวแทนกลุ่มที่ดีของกลุ่มข้อมูล โดยขนาดของโมเดลนั้นสามารถที่จะขยายได้จากขนาดที่กำหนดไว้ และจากแนวคิดที่ต้องการให้ลำดับในการรู้จำไม่มีผลต่อผลการทดลองจึงได้ เพิ่มส่วนของการหาตัวแทนกลุ่มด้วย

3.5 กระบวนการเรียนรู้

การเรียนรู้ของกระบวนการนี้ใช้ input pattern เป็นแบบขาวดำ (I^m) ซึ่งเป็นลักษณะของโครงสร้างโดยรวม ที่ถูกส่งเข้ามาในระบบสามารถแสดงได้ดังนี้

$$I^m_{x,y} = \{1,0\}, x = 1,2,..X, y = 1,2,..Y \tag{3.6}$$

เมื่อ I คือ รูปภาพของลักษณะโครงสร้างโดยรวม

m คือ ลำดับของ input pattern ที่ถูกส่งไปยังโครงข่าย

X และ Y คือจำนวนพิกเซล ตามแนวแกน x และแกน y ของรูปภาพที่นำเข้ามาเป็นอินพุทของระบบ

ในระหว่างกระบวนการเรียนรู้ input pattern จะถูกแบ่งเป็น 2 กลุ่ม คือ (1) กลุ่มของคลาสที่กำลังพิจารณา และ (2) กลุ่มของคลาสที่อยู่นอกการพิจารณา หมายถึง กลุ่มของคลาสที่เหลือที่ไม่ได้อยู่ในกลุ่มที่ 1 การแบ่งกลุ่มนี้เพื่อต้องการให้ลำดับของ input pattern ไม่มีผลกระทบต่อลักษณะของโครงข่าย

กระบวนการเรียนรู้ของวิธีการนี้มีขั้นตอน ดังนี้

1. สร้างเซตของ input pattern P และ Q

กำหนดให้ P เป็นเซตของ input pattern ทั้งหมดที่เป็นสมาชิกของคลาสที่กำลังพิจารณา

$$P = \{P_j\} = \{I^m \mid I^m \in A\} \tag{3.7}$$

เมื่อ j = 1,2,…,c คือ อินเด็กซ์ของ input pattern ในเซต P

A คือ คลาสที่กำลังพิจารณา

c คือ จำนวน input pattern ที่เป็นสมาชิกของคลาส A

กำหนดให้ Q เป็นเซตของ input pattern ทั้งหมดที่เป็นสมาชิกของคลาสที่อยู่นอกการพิจารณา

$$Q = \{Q_e\} = \{I^m \mid I^m \in B\} \tag{3.8}$$

เมื่อ e = 1,2,…,d คือ อินเด็กซ์ของ input pattern ในเซต Q

B คือ คลาสที่อยู่นอกการพิจารณา

d คือ จำนวน input pattern ที่เป็นสมาชิกของคลาส B

2. สร้าง reference pattern

ขั้นตอนการหา reference pattern มีกระบวนการการย้ายจุดศูนย์กลาง ดังที่กล่าวไว้แล้วในบทที่ 3

ขั้นตอนการสร้าง reference pattern นี้พยายามหาตัวแทนที่ดีที่สุดของรูปแบบที่กำลังพิจารณา ซึ่งจะทำการพิจารณา input pattern ทั้งหมดในเซต P และเซต Q โดยใจความสำคัญของขั้นตอนนี้คือ พยายามหา reference pattern ที่มีความเหมือนกันมากที่สุดในกลุ่มที่กำลังพิจารณา ในขณะเดียวกันต้องมีความแตกต่างกับกลุ่มนอกการพิจารณามากที่สุดด้วย ในการสร้าง reference pattern มีขั้นตอนย่อย 2 ขั้นตอนดังนี้

2.1 หา reference pattern ในเซต P

ขั้นตอนนี้จะเป็นการเปรียบเทียบรูปแบบกันภายในกลุ่ม เพื่อหาตัวแทนกลุ่มที่มีความเหมือนกันมากที่สุด โดยการเปรียบเทียบหาได้จาก

$$d_i = \min_{j=1}^{c}(\min_{n=1}^{e}(H(w_i, P_j))) \tag{3.9}$$

เมื่อ i คือ อินเด็กซ์ของ reference pattern โดยที่ i = 1,2,…,c

n = 1,2,…,e

e = จำนวนจุดศูนย์กลางที่ใช้เพื่อพิจารณา

j = 1,2,…,c

$i \neq j$

d คือ ค่าความเหมือนที่หาได้

$H(w_i, P_j)$ คือค่าฮาอูดอร์ฟดิสแทนซ์ระหว่าง reference pattern w_i กับ input pattern P_j

w คือ reference pattern ของ P โดยที่ในตอนแรกนั้นขนาดของ reference pattern จะมีขนาดเท่ากับที่ผู้ใช้ได้กำหนดไว้

จากนั้นนำค่า d_i มาพิจารณาความคล้ายคลึงตามเงื่อนไขนี้

$$d_i \leq \rho_{similar} \tag{3.10}$$

ค่า $\rho_{similar}$ จะมีค่าระหว่าง 0 ถึงความยาวของเส้นทแยงมุมของ reference pattern ในขนาดปัจจุบัน ซึ่งในกระบวนการนี้ค่า $\rho_{similar}$ หมายถึง ค่าความคล้ายคลึงกันของ reference pattern กับ P_j กรณี (3.10) เป็นจริง ให้ทำข้อ 2.2 ต่อ แต่หาก (3.10) เป็นเท็จจะทำข้อ 2.1 ซ้ำ โดยเพิ่มขนาดของ reference pattern ด้วย การเพิ่มขนาดนั้นจะเพิ่มขนาดให้ใหญ่ขึ้นตามที่กำหนดไว้ และขนาดที่ได้จะมีขนาดไม่ใหญ่เกินขนาดของ input pattern นั้น

2.2 ตรวจสอบความคล้ายคลึงที่น้อยพอของ reference pattern กับ Q

ขั้นตอนนี้จะนำเอา reference pattern ที่ได้จากข้อ 2.1 มาหาค่าความเหมือน โดยเปรียบเทียบกับทุกๆ input pattern ใน Q โดยหาได้จาก

$$dO_i = \min_{k=1}^{d}(\min_{n=1}^{e}(H(w_i, Q_k))) \tag{3.11}$$

เมื่อ i คือ อินเด็กซ์ของ reference pattern โดยที่ i = 1,2,…,c

k = 1,2,…,d

dO คือ ค่าความเหมือนที่หาได้

$H(w_i, Q_k)$ คือค่าฮาอูดอร์ฟดิสแทนซ์ระหว่าง reference pattern เปรียบเทียบกับทุกๆ input pattern ใน Q

จากนั้นนำค่า dO_i มาพิจารณาตามเงื่อนไขนี้

$$dO_i > \rho_{dissimilar} \tag{3.12}$$

ค่า $\rho_{dissimilar}$ จะมีค่าระหว่าง 0 ถึงความยาวของเส้นทแยงมุมของ reference pattern ซึ่งในกระบวนการนี้ค่า $\rho_{dissimilar}$ หมายถึง ค่าความแตกต่างของ reference pattern เมื่อเทียบกับทุกๆ input pattern ใน Q กรณี (3.12) เป็นเท็จให้ย้อนกลับไปทำข้อ 2.1 ใหม่ โดยเพิ่มขนาดของ reference pattern ที่กำลังพิจารณาด้วย หาก (3.12) เป็นจริง reference pattern ที่ได้จะเป็นตัวแทนของ input pattern ปัจจุบัน และทำการหา reference pattern ของ input pattern ถัดไปจนกว่าจะครบทุก input pattern ใน P

3. สร้างเวคเตอร์แสดงการเป็นสมาชิก

ขั้นตอนนี้จะเป็นการบอกว่า reference pattern ที่หาได้มีความสามารถมากพอที่จะเป็นตัวแทนของ reference pattern อื่นได้หรือไม่ โดยนำเอา reference pattern ที่ได้ทั้งหมดมาหาค่าฮาอูสดอร์ฟดิสแทนซ์เทียบกับทุก input pattern ใน P โดยหาได้จาก

$$dD_{ij} = \min(H(w_i, P_j)) \tag{3.13}$$

เมื่อ dD คือ ค่าความเหมือนที่หาได้

$i = 1,2,\ldots,c$

$j = 1,2,\ldots,c$

$H(w_i, P_j)$ คือ ค่าฮาอูดอร์ฟดิสแทนซ์ระหว่าง reference pattern w_i เปรียบเทียบกับทุกๆ input pattern ใน P

แล้วนำมาเปรียบเทียบกับค่า $\rho_{similar}$ โดยการเปรียบเทียบนี้จะบอกถึงความสามารถของ reference pattern ที่หาได้ในการเป็นตัวแทนแทน input pattern ในเซต P ซึ่งการเก็บค่าการแสดงเป็นสมาชิกจะเป็นไว้ในเวคเตอร์ M ที่มีขนาดเท่ากับจำนวนของ input pattern ในเซต P โดยเวคเตอร์ M เป็นไบนารีเวคเตอร์ที่มีค่าเฉพาะ 0 และ 1 เท่านั้น ซึ่ง 0 หมายถึงการไม่เป็นสมาชิก และ 1 หมายถึงการเป็นสมาชิก ตามเงื่อนไขดังนี้

$$M(i,g) = \begin{cases} 1 \ , \ \text{if } dD_{ij} \le \rho_{similar} \\ 0 \ , \ \text{otherwise} \end{cases} \quad (3.14)$$

เมื่อ $M(i,g)$ คือ เวคเตอร์แสดงการเป็นสมาชิกสำหรับ reference pattern w_i

4. เลือก reference pattern ที่เหมาะสม

จากข้อ 3. ทำให้เราทราบว่า reference pattern ตัวใดสามารถเป็นตัวแทนที่ดีได้ ซึ่งในขั้นตอนนี้จะทำการหา reference model ของโครงข่ายจาก reference pattern ที่หาได้ โดยใจความสำคัญของขั้นตอนนี้คือ หา reference model ที่สามารถเป็นตัวแทนของ input pattern ใน P ที่กำลังพิจารณาได้ โดยมีข้อจำกัดในการเลือกคือ เลือกจำนวน reference model ให้ได้จำนวนน้อยที่สุดที่สามารถครอบคลุมได้ทุก input pattern ใน P

reference model จะถูกเลือกทีละตัว โดยมีลำดับในการเลือกซึ่งพิจารณาตามเงื่อนไขต่อไปนี้

- เลือก reference pattern ที่ครอบคลุมจำนวนสมาชิกมากที่สุด
- เลือก reference pattern ที่ให้ค่าฮาอูดอร์ฟดิสแทนซ์น้อยที่สุด

โดยจะให้ความสำคัญกับเงื่อนไขแรกมากกว่า หากพบว่า reference pattern มากกว่า 1 ตัวที่เป็นไปตามเงื่อนไขแรก ก็จะพิจารณาเงื่อนไขถัดมา และขั้นตอนจะสิ้นสุดลงเมื่อสามารถหา reference model ครอบคลุ่มสมาชิกทั้งหมดในเซต P

วิธีการเลือก reference model ที่เหมาะสม สามารถเขียนได้ดังนี้

1. เลือก reference pattern ที่ครอบคลุมสมาชิกมากที่สุด เพื่อเป็น reference model ตัวแรก

2. ตรวจสอบ reference pattern ที่ถูกเลือก หากครอบคลุมสมาชิกทุกตัวแล้วให้สิ้นสุดการหา reference model แต่หากยังไม่ครอบคลุมสมาชิกทุกตัวให้หา reference pattern ตัวอื่นที่ทำให้การครอบคลุมจำนวนสมาชิกมากขึ้นเมื่อรวมกับจำนวนสมาชิกที่ได้จากข้อ 1 หากการหา reference pattern ในข้อนี้ได้ reference pattern 1 reference pattern ให้ reference pattern นั้นเป็น reference model ทันที แต่หากได้มากกว่า 1 ตัวให้ดำเนินการข้อ 3

3. ตรวจสอบค่าฮาอูดอร์ฟดิสแทนซ์ของ reference pattern ที่ได้จากกระบวนการเรียนรู้ข้อ 2.1 จากนั้นเลือก reference pattern ที่ให้ค่าฮาอูดอร์ฟดิสแทนซ์น้อยที่สุด เป็น reference model

4. ตรวจสอบความครอบคลุมสมาชิก หากครอบคลุมสมาชิกครบทุกตัวแล้วให้สิ้นสุดการหา reference model แต่หากยังไม่ครอบคลุมครบให้วนกลับไปทำข้อ 1 จนกว่าจะครอบคลุมสมาชิกทุกตัว

ตัวอย่างการเลือก reference pattern ที่เหมาะสม

จากข้อ 3. เราจะได้เวคเตอร์แสดงการเป็นสมาชิกดังตารางที่ 3.1 โดยในตัวอย่างมี input pattern จำนวน 3 pattern ในคลาสที่กำลังพิจารณาอยู่

ตารางที่ 3.1 ตัวอย่างเวคเตอร์แสดงการเป็นสมาชิก

	input pattern no. 1	input pattern no. 2	input pattern no. 3
ref. pattern no. 1	1	1	0
ref. pattern no. 2	0	1	0
ref. pattern no.3	0	0	1

จากตารางที่ 3.1 การพิจารณาในครั้งแรกจะพิจารณา reference pattern ที่ครอบคลุมจำนวนสมาชิกมากที่สุดก่อน ดังนั้นระบบจะเลือก reference pattern no. 1 เป็น reference model ตัวแรก ระบบจะทำการเลือก reference pattern ตัวถัดไป โดย reference pattern no. 3 จะถูกเลือกเป็น reference model ตัวที่สอง เนื่องจาก reference pattern no.3 นี้ทำให้ความครอบคลุมสมาชิกของ reference model เพิ่มมากขึ้น จากนั้นระบบจะทำการเลือก reference model ตัวถัดไป แต่เนื่องจากในตัวอย่างนี้ reference model ที่หาได้สามารถครอบคลุมจำนวนสมาชิกทั้งหมดแล้ว ดังนั้นขั้นตอนการเลือก reference pattern ที่เหมาะสมสำหรับตัวอย่างนี้จะได้ reference model จำนวน 2 ตัว

5. กำหนดให้ reference model เป็น model ของโครงข่าย

จากข้อ 4 เราจะได้ reference model ที่สามารถเป็นตัวแทน input pattern ทุกตัวในเซต P จากนั้นทำการกำหนดให้ reference model ที่ได้เป็น model ของโครงข่าย โดยมีชนิดของคลาสเป็นชนิดเดียวกับคลาสที่เรากำลังพิจารณาอยู่

6. วนทำซ้ำ

วนทำซ้ำข้อ 1 ถึงข้อ 5 กับคลาสอื่น จนกระทั้งครบทุกคลาส