

บทที่ 3

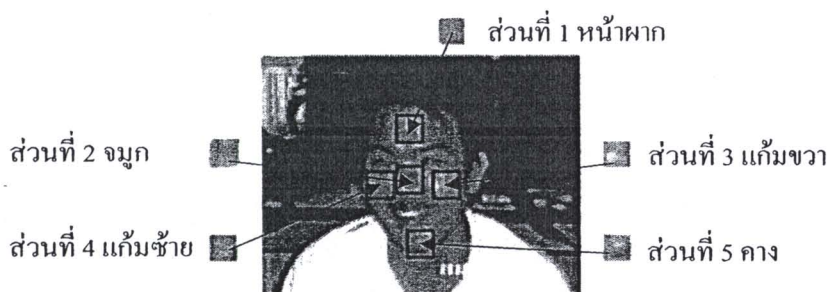
ระบบการตรวจจับใบหน้าคนด้วยโครงข่าย ART

ระบบการตรวจจับใบหน้าคนด้วยโครงข่าย ART นี้ ประกอบไปด้วยระบบย่อย 3 ส่วนได้แก่ระบบคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง ระบบแปลงข้อมูลภาพเป็นการวิเคราะห์องค์ประกอบหลักและระบบตรวจจับใบหน้าคนด้วยโครงข่าย ART รายละเอียดต่างๆ มีดังต่อไปนี้

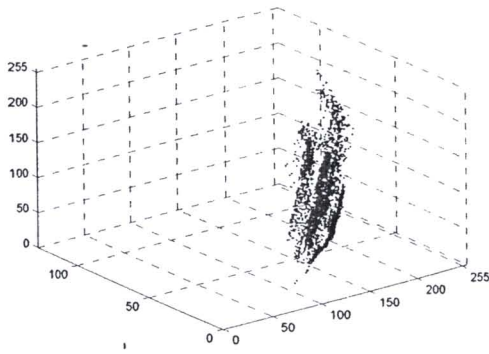
3.1 ระบบการคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง

สีผิวมนุษย์นั้น มีความหลากหลายแตกต่างกันไปขึ้นอยู่กับลักษณะพันธุกรรมและเชื้อชาติ แต่จะพบว่าความแตกต่างของสีผิวมนุษย์ที่เกิดขึ้นนั้น มาจากความแตกต่างทางด้านความเข้มแสง (intensity) มากกว่าทางด้านความเป็นสี (chrominance) ดังนั้นการแยกสีผิวมนุษย์ทำได้โดยการคัดส่วนที่เป็นความเข้มแสงออก พิจารณาเฉพาะส่วนของความเป็นสี ทำให้ได้ค่าสีผิวมนุษย์ที่แตกต่างจากค่าสีอื่น ๆ จากนั้นใช้ช่วงของค่าสีผิวมนุษย์นี้เป็นอินพุตให้โครงข่ายประสาทเทียมแบบแพร่กลับทำการคัดแยกสีผิวมนุษย์ออกจากสีอื่นๆ เพื่อกำหนดขอบเขตในการค้นหาหน้ามนุษย์ให้แคบลง (ชมพู่ ทรัพย์ปทุมสิน, 2548)

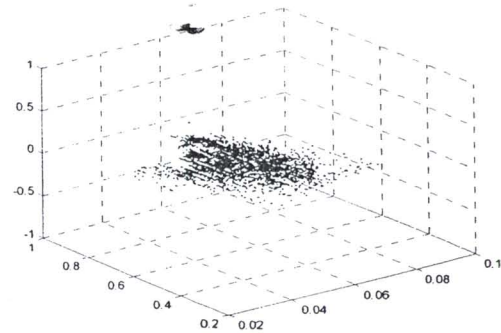
3.1.1 ค่าสีผิวมนุษย์ (Human Color) เป็นค่าที่สามารถคัดแยกออกจากสีของสิ่งแวดล้อมได้ โดยแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV เพื่อแยกส่วนความเข้มแสงออกจากส่วนความเป็นสี เนื่องจากปริภูมิสี HSV สามารถกำจัดผลของความแตกต่างของค่าความเข้มแสงของสีผิวมนุษย์ และลดข้อจำกัดทางด้านแสงสว่างได้จากการคัดส่วนของค่า V ซึ่งเป็นค่าที่ใช้บอกระดับความสว่างของภาพออก จากนั้นคัดส่วนของภาพที่เป็นสีผิวมนุษย์มาเป็นอินพุตให้โครงข่าย โดยการเก็บตัวอย่างอินพุตจากภาพนั้นจะทำการเก็บตัวอย่างบริเวณใบหน้า และภาพใบหน้า 1 ภาพนั้นจะแบ่งบริเวณการเก็บตัวอย่างสีผิวมนุษย์ออกเป็น 5 ส่วนด้วยกัน เนื่องจากทั้ง 5 ส่วนนี้เป็นส่วนของสีผิวมนุษย์ที่ไม่มีสีอื่นที่ไม่ใช่สีผิวมนุษย์ผสมอยู่ด้วย และเป็นจุดหลักบนใบหน้าทำให้ง่ายในการเก็บข้อมูล โดยข้อมูลที่เก็บได้นั้นมีขนาดใหญ่เหมาะสม และในการเก็บข้อมูลถึง 5 จุดนั้น เพื่อให้ได้ตัวอย่างที่ครอบคลุมส่วนต่างๆ ของใบหน้ามากที่สุด ดังแสดงในรูปที่ 3.1 และค่าสีผิวมนุษย์ในปริภูมิสี RGB และ HSV แสดงในรูปที่ 3.2



รูปที่ 3.1 ตัวอย่างอินพุตของโครงข่ายประสาทเทียม



(ก)



(ข)

รูปที่ 3.2 ค่าสีผิวมนุษย์ (ก) ปริภูมิ RGB (ข) ปริภูมิ HSV

3.1.2 โครงข่ายประสาทเทียม (อาทิตย์ ศรีแก้ว, 2552) เป็นโครงข่ายที่สามารถปรับตัวเองต่อการตอบสนองของอินพุต ตามกฎการเรียนรู้ (learning rule) และเมื่อโครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายจะสามารถทำงานที่กำหนดไว้ได้ โดยโครงข่ายประสาทเทียมถูกพัฒนาคิดค้นมาจากหลักการทำงานของสมองมนุษย์ ซึ่งประกอบไปด้วยหน่วยประมวลผลเรียกว่าเซลล์ประสาทเทียม (neuron) โดยจำนวนเซลล์ประสาทในสมองมนุษย์นั้นมีอยู่ประมาณ 10^{11} เซลล์ และมีการเชื่อมต่อกันอย่างมากมาย สมองมนุษย์จึงเสมือนเป็นคอมพิวเตอร์ที่สามารถปรับตัวเองได้ ไม่เป็นเชิงเส้น และทำงานแบบขนานในการดูแลจัดการการทำงานร่วมกันของเซลล์ประสาทในสมอง ดังนั้นโครงข่ายประสาทเทียมที่เลียนแบบมาจากการทำงานของสมองมนุษย์นี้ จึงมีความสามารถในการเรียนรู้จากตัวอย่าง และการทำให้เป็นกรณีทั่วไป (generalize) ซึ่งถือว่าเป็นคุณลักษณะสำคัญของโครงข่ายประสาทเทียม โดยโครงข่ายจะถูกฝึกสอนด้วยรูปแบบต่าง ๆ ที่ต้องการให้โครงข่ายเรียนรู้ผ่านกฎการเรียนรู้ ขบวนการเรียนรู้ได้ของโครงข่ายนั่นเองที่ทำให้โครงข่ายมีความแตกต่างจากการทำงานของโปรแกรมคอมพิวเตอร์อื่น ๆ และการที่โครงข่ายถูกทำให้เป็นกรณีทั่วไปทำให้โครงข่ายสามารถจำแนกแยกแยะรูปแบบของอินพุตแบบใหม่ ๆ ที่โครงข่ายไม่รู้จักรมาก่อนได้ โครงข่ายจะทำการเก็บข้อมูลความรู้ในระหว่างขั้นตอนของการเรียนรู้ โดยเก็บไว้ที่น้ำหนักประสาท (synaptic weights) โครงสร้างของตัวเซลล์ประสาทเทียมภายในโครงข่ายมีอยู่หลายชนิด ซึ่งเป็นองค์ประกอบสำคัญที่ทำให้คุณลักษณะต่าง ๆ ของโครงข่ายแตกต่างกันออกไป ไม่ว่าจะเป็นการจัดวางเรียงตัวของเซลล์ประสาทเทียม กฎการเรียนรู้ที่ทำให้เกิดการปรับเปลี่ยนค่าของจุดประสาท และเงื่อนไขในการฝึกฝนของ นอกจากนี้โครงข่ายประสาทเทียมยังมีข้อดีอีกมากมาย เช่น มีความทนทานต่อความบกพร่อง เพราะข้อมูลภายในโครงข่ายได้ถูกกระจายไปทั้งโครงข่ายตามเซลล์ประสาทเทียมต่าง ๆ การจะทำให้ทั้งระบบไม่สามารถทำงานได้นั้น จะต้องทำให้เกิดความเสียหายอย่างหนักเท่านั้น และโครงข่ายประสาทเทียมมีความสามารถในการโต้ตอบ และตอบสนองต่อสภาวะแวดล้อมได้ โดยเมื่อสภาวะแวดล้อม

เปลี่ยนไป ตัวโครงข่ายจะสามารถตอบสนองกับการเปลี่ยนแปลงนั้น ๆ แล้วทำการฝึกฝนให้เข้ากับสภาวะแวดล้อมใหม่ได้

3.1.3 การคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ โครงข่ายประสาทเทียมแบบแพร่กลับ เป็น โครงข่ายประสาทเทียมแบบหลายชั้น มีการเรียนรู้แบบมีผู้ฝึกสอน (supervised learning) คือ เรียนรู้ที่จะสร้างผลลัพธ์ที่ต้องการให้ได้ตามตัวอย่างที่ป้อนให้ และลดค่าความผิดพลาดของเอาต์พุตให้น้อยที่สุดโดยเทียบกับค่าน้ำหนักประสาท ซึ่งเป็นการวิเคราะห์หาค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean square error) ในอัลกอริทึมแบบแพร่กลับ มีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้แสดงในสมการที่ 2.9

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_Q, t_Q\} \quad (3.1)$$

ในการคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับจำเป็นต้องกำหนดคู่อินพุตและเป้าหมายที่ชัดเจน เพื่อให้โครงข่ายสามารถทำงานได้อย่างถูกต้องที่สุด โดยจะทำการเก็บตัวอย่างสีผิวมนุษย์จากที่ได้กล่าวไว้ในหัวข้อที่ 3.1.1 มาเป็นคู่อินพุตและเป้าหมายของโครงข่าย ซึ่งใช้ค่า H และค่า S ของภาพในช่วงที่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าเป็นสีผิวมนุษย์ และค่า H และค่า S ของภาพ ในช่วงที่ไม่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าไม่ใช่สีผิวมนุษย์มาฝึกสอนโครงข่าย ดังความสัมพันธ์ในสมการที่ 3.2

$$\left\{ P_{skin} = \begin{bmatrix} H_{skin} \\ S_{skin} \end{bmatrix}, t_{skin} = 1 \right\} \\ \left\{ P_{nonskin} = \begin{bmatrix} H_{nonskin} \\ S_{nonskin} \end{bmatrix}, t_{nonskin} = 0 \right\} \quad (3.2)$$

โดยกำหนดตัวอย่างสีผิวมนุษย์ที่ใช้ฝึกสอนโครงข่ายจำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ จากจำนวนคน 20 คนซึ่งมีสีผิวแตกต่างกันออกไป และตัวอย่างภาพที่ไม่ใช่สีผิวมนุษย์จำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ ต่อมาทำการกำหนดตัวชี้วัดประสิทธิภาพ (performance index) เมื่อระบบมีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้ ทำการป้อนแต่ละอินพุตให้กับโครงข่าย เอาต์พุตที่ได้จะถูกนำไปเปรียบเทียบกับเป้าหมาย เพื่อให้ค่าความผิดพลาดของเอาต์พุตและเป้าหมายมีค่าน้อยที่สุด ซึ่งค่าความผิดพลาดระหว่างเอาต์พุตและเป้าหมายนี้เองเป็นตัวชี้วัดประสิทธิภาพ โดยตัวชี้วัดประสิทธิภาพดังกล่าวนี้ก็คือค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean-square error) ของเอาต์พุตและเป้าหมาย

จากนั้นเพื่อให้ระบบมีประสิทธิภาพมากที่สุด จำเป็นจะต้องออกแบบพารามิเตอร์ของโครงข่ายประสาทเทียมแบบแพร่กลับซึ่งได้แก่ จำนวนชั้น จำนวนเซลล์ประสาทเทียม และถ่วงโอน ให้เหมาะสมกับงานที่ทำ ซึ่งในงานวิจัยเรื่องวิธีการใหม่แบบพันทางในการแยกส่วนมือจากภาพสี (ชมพู ทรัพย์ปทุมสิน, 2548) ได้ทำการทดลองหาพารามิเตอร์ที่เหมาะสมของโครงข่ายประสาทเทียมแบบแพร่กลับในการคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง โดยใช้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยเป็นตัวชี้วัดประสิทธิภาพ ได้ผลการทดลองดังตารางที่ 3.1

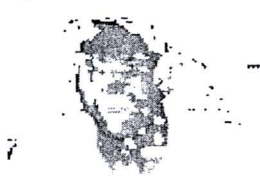
จากผลการทดลอง โครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยที่สุดคือ โครงข่าย 2-10-1 (ชั้นอินพุตมีเซลล์ประสาทเทียม 2 เซลล์ มีชั้นซ่อนเร้น 1 ชั้นมีเซลล์ประสาทเทียม 10 เซลล์ และชั้นเอาต์พุตมีเซลล์ประสาทเทียม 1 เซลล์) ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ (Hyperbolic Tangent Sigmoid) และแบบลอจ ซิกมอยด์ (Log-Sigmoid) ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยเท่ากับ 0.0408147 และ โครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยรองลงมา คือ โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ และแบบลอจ ซิกมอยด์ ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ย 0.0430907 ดังนั้นจากผลการทดลอง จึงเลือกใช้โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ เนื่องจากโครงข่าย 2-5-1 นั้นให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยกว่าไม่มากนักเมื่อเทียบกับโครงข่าย 2-10-1 แต่ใช้จำนวนเซลล์ประสาทเทียมน้อยกว่าโครงข่าย 2-10-1 ทำให้โครงข่ายมีขนาดเล็กกว่าโดยที่ประสิทธิภาพของโครงข่ายใกล้เคียงกัน โดยผลการคัดแยกสีผิวแสดงในรูปที่ 3.3 และสถาปัตยกรรมโครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์แสดงในรูปที่ 3.4

ตารางที่ 3.1 ผลการฝึกสอนโครงข่ายประสาทเทียมเพื่อเลือกพารามิเตอร์ของโครงข่าย

โครงข่าย	ฟังก์ชันถ่ายโอน	ค่าความผิดพลาดแบบกำลังสองเฉลี่ย
2-2-1	Tansig – Tansig	0.0823830
2-2-1	Tansig – Purelin	0.0781679
2-2-1	Purelin – Purelin	0.1501080
2-2-1	Tansig – Logsig	0.0498176
2-5-1	Tansig – Purelin	0.0746313
2-5-1	Tansig – Logsig	0.0430907
2-10-1	Tansig – Purelin	0.0689867
2-10-1	Tansig – Logsig	0.0408147
2-5-3-1	Tansig – Logsig – Tansig	0.5086721
2-5-3-1	Tansig – Logsig – Logsig	0.0682925
2-5-3-1	Tansig – Logsig – Purelin	0.0678671
2-5-3-2-1	Tansig – Logsig – Logsig - Purelin	0.2496315



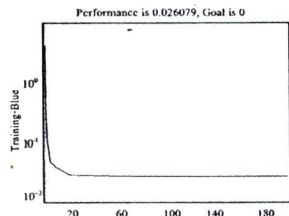
(ก)



(ข)



(ค)

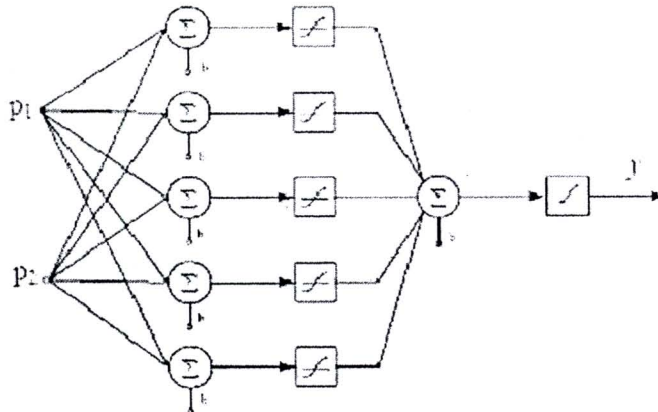


(ง)



(จ)

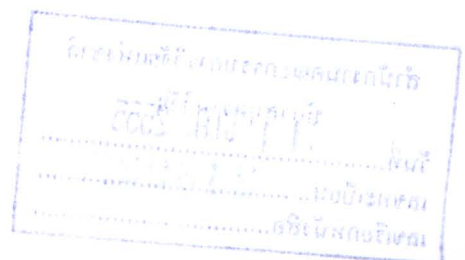
รูปที่ 3.3 ผลการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ (ก) ภาพต้นแบบ (ข) ผลการคัดแยกสีผิวมนุษย์ (ค) ขอบเขตของสีผิวมนุษย์ (ง) ผลการฝึกสอนโครงข่าย (จ) ขอบเขตการค้นหาหน้าคน



รูปที่ 3.4 โครงสร้างของโครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์

3.2 ระบบการแปลงข้อมูลภาพเป็นการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

วิธีการวิเคราะห์องค์ประกอบหลัก เป็นการแทนชุดข้อมูลใหม่ที่มีขนาดเล็กกว่าลงในชุดข้อมูลเดิม ซึ่งชุดข้อมูลใหม่นั้นจะมีความสัมพันธ์ระหว่างข้อมูลภายในชุดน้อยมาก โดยชุดข้อมูลเดิมที่มีความสัมพันธ์



ภายในข้อมูลขนาด D มิติ สามารถใช้การรวมเชิงเส้นในการลดข้อมูลเป็น d มิติ โดยที่ $d < D$ และเรียกชุดข้อมูลใหม่นี้ว่า องค์ประกอบหลัก (Principal Components)

วิธีการวิเคราะห์องค์ประกอบหลักสามารถทำได้ 2 วิธี คือ ทฤษฎีผลรวมของระยะทางที่น้อยที่สุด (Least Squared Distance) นำเสนอโดย Duda (2001) และ วิธีการเปลี่ยนตัวแปร นำเสนอโดย Hotelling (2001)

1. ทฤษฎีผลรวมของระยะทางที่น้อยที่สุด (Least Squared Distance) เป็นวิธีที่มีแนวคิดมาจากการหาจุดและเส้นที่มีความเหมาะสมที่สุดของชุดข้อมูลตัวอย่างขนาด M ค่า โดยแต่ละค่ามีขนาด D มิติ ทำให้ได้ข้อมูลตัวอย่างทั้งหมด คือ $t_1, t_2, \dots, t_M$ แล้วแทนข้อมูลตัวอย่างนั้นด้วยข้อมูลใหม่ ขนาด D มิติ และ 1 มิติ โดยที่ t_0 คือข้อมูลที่ถูกแทนเป็น 0 มิติ (0-dimensional representation) เพื่อหาความสัมพันธ์ของข้อมูลตัวอย่างกับข้อมูลใหม่ โดยใช้ t_0 มาคำนวณหาค่าผลรวมความผิดพลาดน้อยที่สุด แสดงในสมการที่ 3.3

$$J_0(t_0) = \sum_{k=1}^M \|t_0 - t_k\|^2 \quad (3.3)$$

โดยค่า t_0 ที่ทำให้ J_0 มีค่าน้อยที่สุดหาได้จาก

$$\bar{t} = \frac{1}{M} \sum_{k=1}^M t_k \quad (3.4)$$

t คือค่าเฉลี่ยตัวอย่าง

$$\begin{aligned} J_0(t_0) &= \sum_{k=1}^M \|(T_0 - \bar{t}) - (t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2 \sum_{k=1}^M (T_0 - \bar{t})^T (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2(T_0 - \bar{t})^T \sum_{k=1}^M (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \end{aligned} \quad (3.5)$$

ทำการฉายข้อมูลลงบนเส้นค่าเฉลี่ยตัวอย่าง ได้สมการความสัมพันธ์ของตัวแปรใหม่ดังสมการที่ 3.6 โดยที่ u คือเวกเตอร์ขนาด 1 หน่วย ที่มีทิศทางตามเส้นค่าเฉลี่ย และ y คือขนาดที่สัมพันธ์กันของระยะจาก t ถึง $\bar{t}$

$$t = \bar{t} + yu \quad (3.6)$$

พบว่าแกนฉายของข้อมูล 1 มิติ ก็คือไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่มากที่สุดตามลำดับ ดังนั้นในแนวทางเดียวกันเราสามารถขยายการแทนข้อมูลในรูปของ d มิติ โดยใช้การฉายข้อมูลลงเส้นข้อมูล d ดังแสดงในสมการที่ 3.7

$$t = \bar{t} + y_d u_d \quad (3.7)$$

และจะเรียกค่า y_d ว่าองค์ประกอบหลัก (Principal component) ของข้อมูลขนาด d มิติ



2. วิธีการเปลี่ยนตัวแปร มีแนวคิดมาจากการแปลงเชิงเส้นของตัวแปรที่มีความสัมพันธ์กันกับตัวแปรใหม่ที่มีความสัมพันธ์กันน้อยตามลำดับการลดลงของความแปรปรวน โดยกำหนดให้เวกเตอร์ t ใด ๆ ขนาด D มิติ มีค่าเฉลี่ย $\bar{t}$ และเมทริกซ์ความแปรปรวนร่วม ขนาด $D \times D$ มิติ สามารถคำนวณหาเวกเตอร์ใหม่ y ที่มีความสัมพันธ์กันน้อยตามลำดับการลดลงของค่าความแปรปรวน จากหลักการรวมกันเชิงเส้นแสดงความสัมพันธ์ดังสมการที่ 3.8

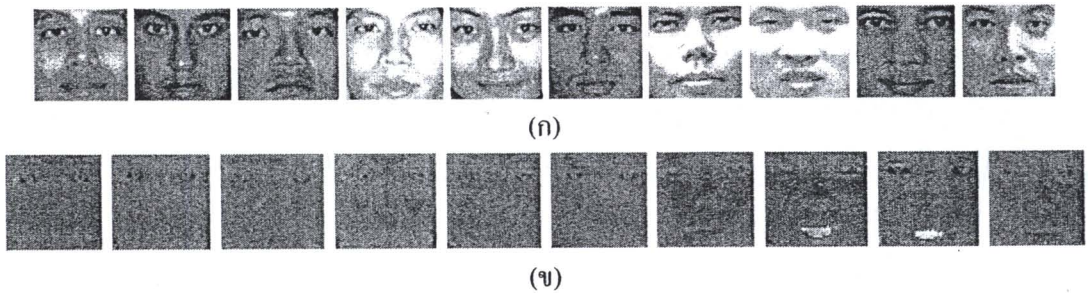
$$y_k = u_1 t_1 + u_2 t_2 + \dots + u_{Dk} t_D = \mathbf{u}_k^T \mathbf{t} \quad (3.8)$$

จากสมการที่ 3.8 สามารถเขียนสมการใหม่ได้โดยที่ y มีขนาด $d \times 1$ มิติ จากเวกเตอร์ t ขนาด $D \times 1$ มิติ และ $\mathbf{U}$ ขนาด $D \times d$ มิติ

$$y = \mathbf{U}^T \mathbf{t} \quad (3.9)$$

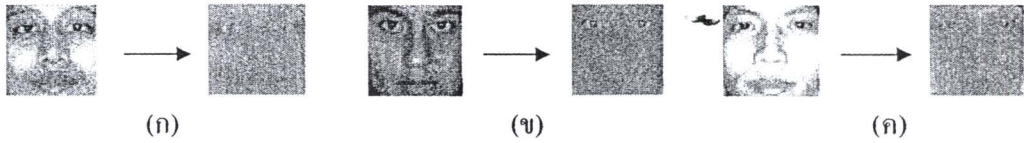
จากสมการที่ 3.8 พบว่า $\mathbf{u}$ เป็นเมทริกซ์ที่ระบุแกนทิศทางขนาด 1 หน่วย และในสมการที่ 3.9 เมทริกซ์ $\mathbf{U}$ คือเมทริกซ์ความแปรปรวนร่วม และสามารถหาไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่เรียงจากมากไปน้อยได้จากเมทริกซ์ความแปรปรวนร่วม $\mathbf{U}$

3.2.1 การหาค่าไอเกนและไอเกนเวกเตอร์ ด้วยหลักการวิเคราะห์องค์ประกอบ เป็นวิธีการหาหน้าไอเกน (Eigenface) ในปริภูมิหน้า (face Space) นำเสนอโดย Turk และ Pentland โดยทำการคำนวณหาไอเกนเวกเตอร์ของเมทริกซ์ความแปรปรวนของภาพหน้าตัวอย่าง ที่สมนัยกับค่าไอเกนที่มีค่าความแปรปรวนจากมากไปน้อย โดยนำภาพหน้าตัวอย่างดังกล่าวไปฉายด้วยไอเกนเวกเตอร์ จะได้ไอเกนเวกเตอร์ในปริภูมิหน้าไอเกน ซึ่งสามารถคำนวณได้โดยนำภาพหน้าตัวอย่าง จำนวน N ภาพ คือ $A_1, A_2, A_3, \dots, A_N$ ที่มีมิติขนาดเท่ากับคือ $m \times n$ มิติ



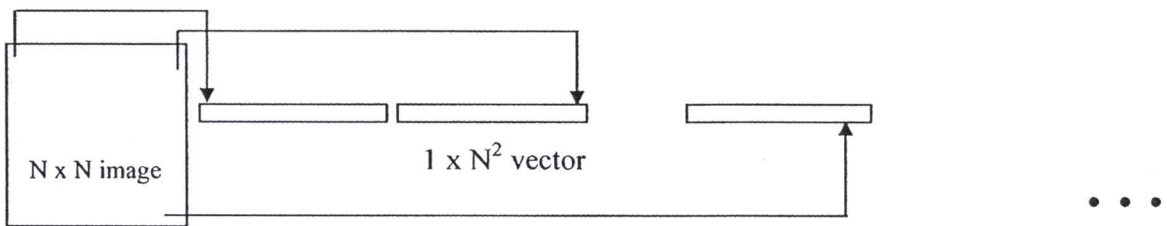
รูปที่ 3.5 ตัวอย่างภาพหน้าตัวอย่างในการหาหน้าไอเกน (ก) ตัวอย่างภาพในปริภูมิ RGB (ข) ตัวอย่างภาพในปริภูมิ HSV

โดยในรูปที่ 3.5 แสดงภาพตัวอย่างก่อนจะนำภาพมาทำการหาค่าไอเกนเวกเตอร์ซึ่ง จำเป็นต้องนำภาพในปริภูมิ RGB มาแปลงเป็นปริภูมิ HSV เพื่อลดผลของค่าความสว่าง ซึ่งมีผลต่อความถูกต้องของระบบการตรวจจับหน้าคน แล้วจึงแปลงภาพที่ทำการลดผลของความสว่างแล้วนั้นเป็นปริภูมิ RGB อีกครั้ง และแปลงให้อยู่ในระดับเทาเพื่อให้เป็นภาพระนาบเดียว ดังแสดงในรูปที่ 3.6



รูปที่ 3.6 ภาพระดับเทาที่ได้จากการลดผลของความสว่าง (ก) สีผิวกลาง (ข) สีผิวเข้ม (ค) สีผิวอ่อน

แปลงเมทริกซ์ภาพตัวอย่างเป็นเวกเตอร์ขนาด $1 \times D$ ($D=mxn$) มิติ ดังแสดงในรูปที่ 3.7



รูปที่ 3.7 ตัวอย่างการแปลงเมทริกซ์ภาพเป็นเวกเตอร์ภาพ

จากนั้นนำเวกเตอร์ที่ได้มามาจัดให้อยู่ในรูปของเมทริกซ์ ได้

$$\mathbf{I} = \{I_1, I_2, I_3, \dots, I_N\} \quad (3.10)$$

และนำเมทริกซ์ $\mathbf{I}$ ที่ได้มาหาค่ากลางของชุดข้อมูลจากสมการที่ 3.11

$$\Phi_i = I_i - \frac{1}{N} \sum_{i=1}^N I_i \quad (3.11)$$

จะได้เวกเตอร์ค่ากลางของชุดข้อมูลในสมการที่ 3.12

$$\mathbf{A} = \{\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_N\} \quad (3.12)$$

จากนั้นนำเมทริกซ์ $\mathbf{A}$ ไปหาค่าเมทริกซ์ความแปรปรวน (covariance matrix: $\mathbf{C}$) โดยเมทริกซ์ที่ได้มีขนาด $D \times D$ มิติ ดังแสดงในสมการที่ 3.13 และ 3.14

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^N \Phi_i \Phi_i^T = \mathbf{A} \mathbf{A}^T \quad (3.13)$$

$$\mathbf{A} \mathbf{v}_i = u_i \quad (3.14)$$

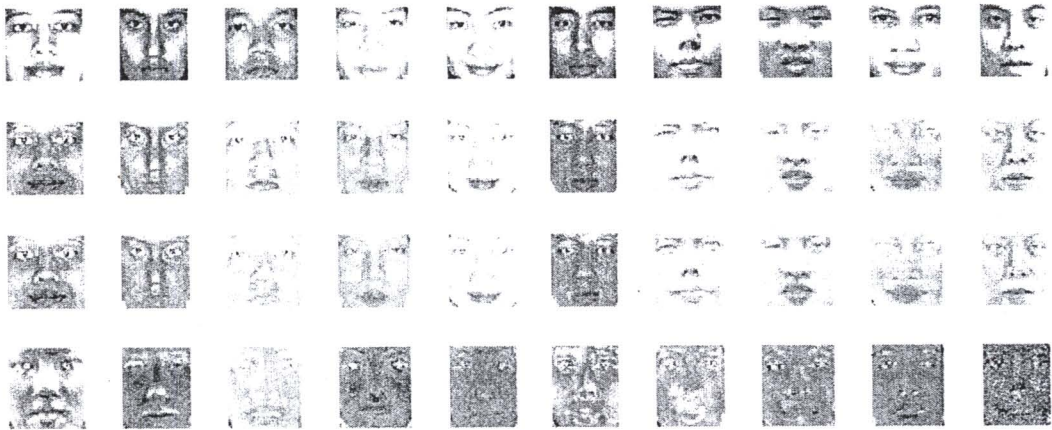
จะได้ค่าไอเกน u_i จำนวนเท่ากับจำนวนภาพต้นแบบคือ N ค่า และไอเกนเวกเตอร์ $\mathbf{v}_i$ ขนาด $D \times D$ จุดภาพ ซึ่งเป็นข้อมูลขนาดใหญ่ พิจารณาเห็นว่าสามารถลดขนาดของค่าทั้งสองลงได้ โดยค่าไอเกนและไอเกนเวกเตอร์ของ $\mathbf{A} \mathbf{A}^T$ และ $\mathbf{A}^T \mathbf{A}$ นั้นจะเป็นค่าที่มีความสอดคล้องกัน ดังนั้นเราจะหาค่าไอเกนและไอเกนเวกเตอร์จาก $\mathbf{L} = \mathbf{A}^T \mathbf{A}$ จะได้ค่าไอเกน N ค่าและไอเกนเวกเตอร์มีขนาด $1 \times N$ มิติ

จากค่าไอเกนเวกเตอร์ สามารถหาค่าหน้าไอเกน (Eigenface) ได้ โดยที่หน้าไอเกนค่า 1 นั้นจะแสดงลักษณะเฉพาะของหน้าได้ชัดเจนกว่าหน้าไอเกนค่า 2 ดังนั้น เราสามารถตัดหน้าไอเกนค่า 10 ออกได้โดยไม่

ทำให้เกิดความผิดพลาดกับระบบ เนื่องจากค่าไอเกนค่าที่ 10 นั้นแสดงคุณสมบัติของความเป็นหน้าไว้น้อยมากจนสามารถตัดค่านี้ทิ้งได้ โดยไม่ส่งผลต่อการแปลงเป็นหน้าไอเกน ดังนั้นเราจะสนใจหน้าไอเกน k ค่า โดยที่ $1 \leq k < N$ และแสดงการหาหน้าไอเกนในโดเมนหน้า (face space) จากสมการที่ 3.15

$$uu_k = (Av_k)^T \left(I_i - \frac{1}{N} \sum_{i=1}^N I_i \right) \quad (3.15)$$

ภาพต้นแบบจำนวน 10 ภาพสามารถหาภาพหน้าไอเกนได้ดังแสดงในรูปที่ 3.8



รูปที่ 3.8 ภาพหน้าไอเกน แถวที่ 1 ภาพต้นแบบในระดับเทา แถวที่ 2 ภาพต้นแบบในระดับเทาที่ทำการลดผลของความสว่างแล้ว แถวที่ 3 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 1 แถวที่ 4 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 10

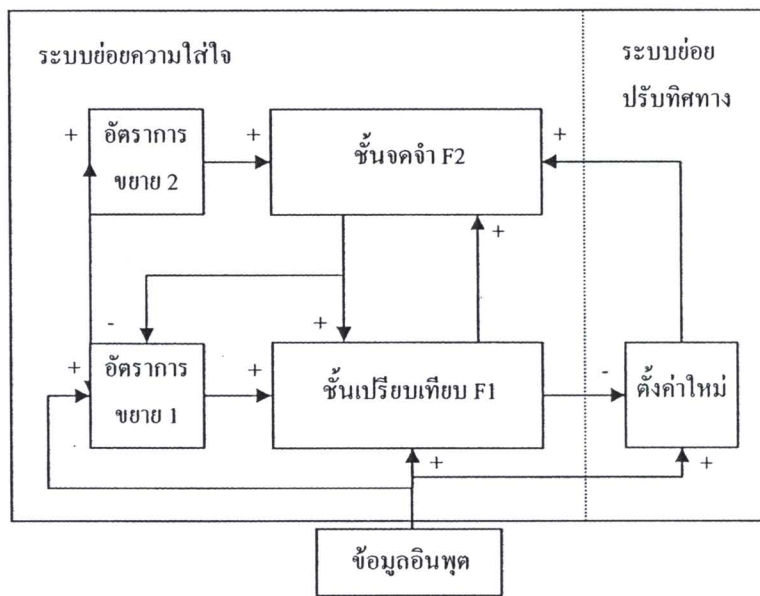
3.3 โครงข่าย ART (Adaptive Resonance Theory Network)

โครงข่าย ART เป็นโครงข่ายประสาทเทียมแบบไม่มีการฝึกสอน โดยเมื่อทำการป้อนค่าให้โครงข่าย โครงข่ายจะจดจำข้อมูลนั้นเป็นกลุ่มข้อมูลชุดที่ 1 เมื่อป้อนข้อมูลใหม่ที่มีค่าแตกต่างจากข้อมูลชุดแรก ระบบจะทำการเพิ่มกลุ่มข้อมูลใหม่เพื่อเก็บข้อมูลที่มีความแตกต่างกัน ซึ่งแตกต่างจากโครงข่ายประสาทเทียมแบบอื่นที่เมื่อทำการป้อนข้อมูลให้โครงข่ายด้วยข้อมูลชุดที่ 1 แล้ว ต้องการเพิ่มข้อมูลใหม่ให้โครงข่ายเรียนรู้เพิ่มเติม จำเป็นจะต้องทำการฝึกสอนโครงข่ายใหม่อีกครั้ง ด้วยข้อมูลชุดใหม่นั้น และทำให้โครงข่ายลืมความรู้เก่าที่ได้เรียนรู้มาจากข้อมูลชุดเดิมซึ่งปัญหานี้เรียกว่า ทวิตเสถียรภาพสภาพพลาสติก (stability-plasticity dilemma) (Guo, Li, and Chan, 1998) ดังนั้น Carpenter and Grossberg (1991) จึงได้พัฒนาโครงข่าย ART เพื่อแก้ปัญหาดังกล่าว กล่าวคือโครงข่าย ART มีคุณสมบัติสำคัญ 2 ประการ คือ สามารถคงความรู้ที่เคยเรียนรู้หรือรูปแบบที่เรียนรู้มาแล้วได้ (stability) และสามารถเรียนรู้ความรู้หรือรูปแบบใหม่เพิ่มเติมได้ไปพร้อม ๆ กัน (plasticity)

3.3.1 สถาปัตยกรรมของโครงข่าย ART โครงข่ายประสาทเทียม ART เป็นโครงข่ายประสาทเทียมแบบมีการฝึกสอนโดยใช้กฎการเรียนรู้แบบแข่งขัน แบ่งการทำงานออกเป็น 2 ระบบย่อย คือ

1. ระบบย่อยความใส่ใจ (attentional subsystem) มีโครงสร้างแบบ 2 ชั้น คือชั้นเปรียบเทียบ (comparison layer: F1) และชั้นจดจำ (recognition layer: F2)
2. ระบบย่อยปรับทิศทาง (orienting subsystem) ทำการปรับเทียบค่าน้ำหนักประสาทให้มีความสอดคล้องกับค่าอินพุตของระบบที่เข้ามา

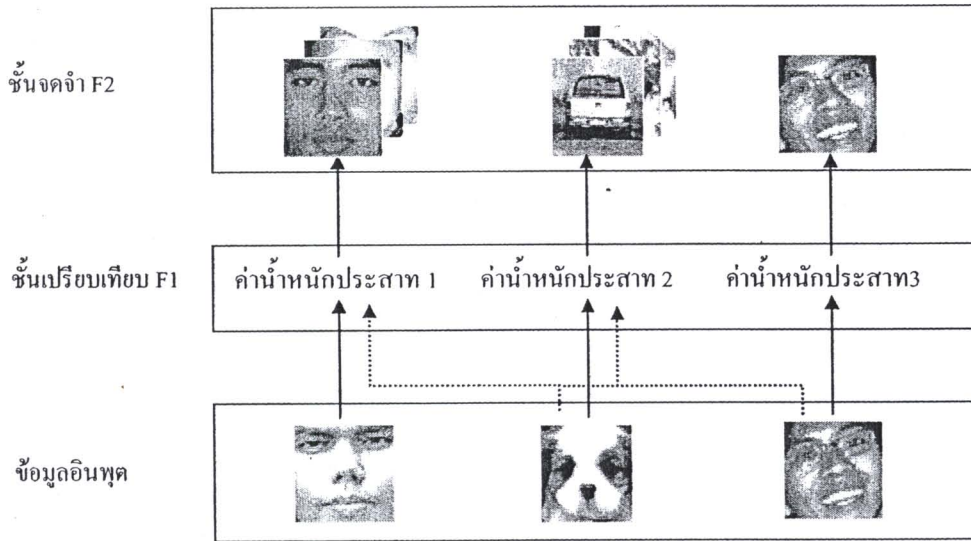
โครงสร้างของระบบทั้งสองแสดงในรูปที่ 3.9



รูปที่ 3.9 แผนผังการทำงานของโครงข่าย ART

โครงข่ายประสาทเทียม ART ที่ได้รับการฝึกสอนแล้วจะเก็บค่าข้อมูลอยู่ในรูปของค่าน้ำหนักประสาท (weight) โดยในชั้นเปรียบเทียบจะนำข้อมูลอินพุต (input data) ที่เข้ามานั้น มาพิจารณาเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่อยู่ในชั้นจดจำ เรียกค่านี้อีกว่าค่าเทียบเคียง ถ้าข้อมูลอินพุตนั้นมีค่าสอดคล้องกับค่าน้ำหนักประสาทของกลุ่มใด ค่าเทียบเคียงที่ได้จากการเปรียบเทียบกับค่าน้ำหนักประสาทกลุ่มนั้นจะมีค่าน้อย และทำการคัดแยกกลุ่มข้อมูลอินพุตโดยการเปรียบเทียบค่าสอดคล้อง (vigilance: ρ) กับค่าเทียบเคียง ถ้าข้อมูลที่ได้มีค่าเทียบเคียงมากกว่าค่าสอดคล้อง ระบบจะทำการปรับค่าน้ำหนักประสาทของกลุ่มข้อมูลนั้น เพื่อทำการเก็บค่าข้อมูลใหม่ในกลุ่มที่มีค่าสอดคล้องกัน ถ้าข้อมูลที่ได้ มีค่าเทียบเคียงน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตนั้น ไม่สอดคล้องกับค่าน้ำหนักประสาทในกลุ่มนั้น ระบบจะส่งค่าน้ำหนักประสาทของกลุ่มใหม่มาทำการเปรียบเทียบกับข้อมูลอินพุตต่อไป แต่ถ้าข้อมูลอินพุตนั้น ไม่สอดคล้องกับ

กลุ่มข้อมูลในชั้นจดจำกลุ่มใดเลระบบจะทำการสร้างกลุ่มใหม่และเก็บข้อมูลอินพุตนั้นไว้ในกลุ่มใหม่ ดังแสดงในรูปที่ 3.10



โดย — แสดงว่าค่าเทียบเคียงมากกว่าค่าสอดส่อง

..... แสดงว่าค่าเทียบเคียงน้อยกว่าค่าสอดส่อง

รูปที่ 3.10 โครงข่ายประสาทเทียม ART

3.3.2 ประเภทของโครงข่าย ART สามารถแบ่งออกได้เป็น 4 ประเภท โดยจะแตกต่างกันไปตามข้อมูลที่รับเข้ามาและรูปแบบของการเรียนรู้ ได้แก่

3.3.2.1 โครงข่าย ART1 (Carpenter and Grossberg, 1987) เป็นระบบที่มีความซับซ้อนน้อยที่สุด โดยจะรับข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น โดยข้อมูลอินพุตจะส่งให้ชั้นเปรียบเทียบ F1 ตัดสินใจว่าข้อมูลอินพุตนั้นควรเป็นข้อมูลกลุ่มใด ด้วยสมการที่ 3.16 และ 3.17

$$Y_j = \sum_{i=1}^n (b_{ij} I_i) \quad (3.16)$$

$$Y_j = \text{Max}(Y_j) \quad (3.17)$$

โดย I คือ อินพุตของระบบ

b คือ ไบอัส ของระบบ

Y_j คือ ค่ามากที่สุดของกลุ่มที่ J

3.3.2.2 โครงข่าย ART2 (Carpenter and Grossberg, 1987) เป็นระบบที่ทำการพัฒนาต่อจาก ART1 ให้สามารถรับข้อมูลเชิงอุปมาน (analog data) จึงมีโครงสร้างของระบบที่ซับซ้อนมากกว่า ART1 โดยที่ ART2 นั้นจะมีโครงสร้างเหมือนกับ ART1 ยกเว้นในชั้นเปรียบเทียบ F1 จะทำการเพิ่ม ชั้นย่อย

(sublayers) เพื่อให้สามารถรับข้อมูลแบบเชิงอุปมานได้ โดยข้อมูลอินพุตเชิงอุปมานนั้นมีขนาดใหญ่และมีความแตกต่างของข้อมูลมากกว่าข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น ที่ซ่อนให้โครงข่าย ART1 ดังนั้นชั้นย่อยที่เพิ่มมาในชั้นเปรียบเทียบกับ F1 จึงทำหน้าที่ย่อข้อมูลขนาดใหญ่ให้เล็กลงตามสมการที่ 3.18

$$f(x) = \begin{cases} x & \text{if } x \geq \theta \\ 0 & \text{if } x < \theta \end{cases} \quad (3.18)$$

โดยค่า θ เป็นค่าที่ผู้ใช้โครงข่ายกำหนดเลือกค่าตามความเหมาะสมของข้อมูลอินพุต

3.3.2.3 โครงข่าย ARTMAP (Carpenter, Grossberg, and Reynolds, 1991) เป็นระบบที่มีผู้ฝึกสอน (supervised) และเสมือนมีโครงข่าย ART 2 ชุด โดยชุดแรกจะทำหน้าที่รับข้อมูลอินพุต และชุดที่ 2 รับข้อมูลเอาต์พุตที่มีความสัมพันธ์กับข้อมูลอินพุต ระบบจะทำการเรียนรู้จากความสัมพันธ์ของข้อมูลอินพุตกับเอาต์พุตนั้น โดยคำว่า ARTMAP นั้นมีความหมายว่าระบบจะดำเนินการเรียนรู้เพื่อจับคู่หรือแมตช์ (match) ข้อมูลอินพุตและเอาต์พุตที่ถูกต้อง โดยอินพุตก่อนเข้าโครงข่ายต้องผ่านการทำตัวเข้ารหัสแบบเติมเต็ม (Complement Coder) เพื่อปรับข้อมูลอินพุตให้มีขนาดเป็นสองเท่า ทำให้โครงข่ายสามารถพิจารณาขอบเขตของรูปแบบข้อมูลได้ครอบคลุมขึ้น จากนั้นส่งข้อมูลไปทำการเปรียบเทียบเพื่อหากลุ่มข้อมูลที่เหมาะสมที่สุด โดยทำการเปรียบเทียบด้วยสมการที่ 3.4 และ 3.5

$$Y_j(I) = \frac{|I \cap w_j|}{\alpha + |w_j|} \quad (3.19)$$

$$Y(I)_J = \text{Max}(Y_j) \quad (3.20)$$

เมื่อ $\cap$ ในสมการที่ 3.19 คือตรรกศาสตร์ส่วนร่วมรวม (logical AND intersection) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทที่มีค่าเหมือนกัน โดยกำหนดให้ค่า α มีค่าน้อยมากๆ และ Y_j คือกลุ่มที่มีความสอดคล้องกับอินพุตมากที่สุด

3.3.2.4 โครงข่าย Fuzzy ART หรือ Fuzzy ARTMAP (Carpenter, Grossberg, and Rosen, 1991) เป็นระบบที่ได้รับการพัฒนามาจากโครงข่าย ART และโครงข่าย ARTMAP โดยนำตรรกศาสตร์คลุมเครือ (fuzzy logic) มารวมกับโครงข่าย ดังนั้นในการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทจึงใช้กฎตรรกศาสตร์มาพิจารณาด้วย ดังสมการที่ 3.21 และ 3.22

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|} \quad (3.21)$$

$$T(I)_J = \text{Max}(T_j) \quad (3.22)$$

เมื่อ $\wedge$ ในสมการที่ 3.21 คือตรรกศาสตร์คลุมเครือค่าต่ำสุดรวม (fuzzy AND minimum) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทโดยพิจารณาค่าที่ต่ำที่สุดในการรวมกัน และถ้าค่า T_j มีค่ามากที่สุดมากกว่า 1 ค่าระบบจะทำการเลือกเอาต์พุตของข้อมูลกลุ่มแรกเป็นกลุ่มที่เหมาะสมที่สุดของอินพุตนั้น

3.4 ระบบตรวจจับหน้าคนจากภาพด้วยโครงข่าย ART

ค่าไอเกนเวกเตอร์ที่คำนวณได้จากข้างต้น จะถูกนำมาเป็นอินพุตของโครงข่าย ART ซึ่งในที่นี้พิจารณาใช้โครงข่าย ART แบบ ARTMAP เนื่องจากเป็นโครงข่ายที่พัฒนามาจาก ART1 และ ART2 เพื่อให้สามารถรับข้อมูลอินพุตที่มีความซับซ้อนมากขึ้น ซึ่งหลักการในการคัดแยกกลุ่มข้อมูลของโครงข่ายทำให้สามารถแยกข้อมูลที่มีความซับซ้อนมากได้ชัดเจนมากกว่า ART1 และ ART2 ส่วนโครงข่ายแบบ Fuzzy ART หรือ Fuzzy ARTMAP นั้น ใช้หลักตรรกศาสตร์คลุมเครือมาช่วยในการคัดแยกกลุ่มข้อมูล ซึ่งมีโครงสร้างที่ซับซ้อนมากขึ้นไป ในการคัดแยกกลุ่มข้อมูลที่ต้องการแยกเพียงความเป็นหน้ามนุษย์และความไม่เป็นหน้ามนุษย์เท่านั้น

3.4.1 โครงสร้างของระบบ ป้อนชุดข้อมูลตัวอย่างเพื่อกำหนดกลุ่มให้โครงข่าย ART ด้วยค่าไอเกนเวกเตอร์ที่ได้จากใบหน้าตัวอย่าง โดยใช้ใบหน้าตัวอย่างจำนวน 20 หน้า ทำให้ได้ไอเกนเวกเตอร์ที่เป็นชุดข้อมูลตัวอย่างของหน้าคนเป็นกลุ่มข้อมูล กลุ่มที่ 1 ของโครงข่ายมีขนาด 20×20 มิติ และในทำนองเดียวกันทำการป้อนข้อมูลตัวอย่างที่ไม่ใช่ใบหน้าคนให้โครงข่าย ART ทำให้ชั้นเปรียบเทียบในโครงข่าย ART มีข้อมูลจดจำอยู่ 2 กลุ่ม คือ กลุ่มหน้าคนและกลุ่มที่ไม่ใช่หน้าคน โดยในการเปรียบเทียบขั้นแรกนี้เป็นการเปรียบเทียบเพื่อตรวจจับหน้าคน โดยเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่เก็บค่าในชั้นจดจำ และโครงข่ายจะทำการจำแนกกลุ่มข้อมูลที่เข้าโครงข่าย โดยเมื่อมีชุดข้อมูลใหม่เข้าสู่โครงข่ายโครงข่ายจะทำการตรวจสอบชุดข้อมูลนั้นโดยการเปรียบเทียบค่าเทียบเคียง (match: μ) กับค่าสอดคล้อง (ρ) เพื่อหาค่าตรวจสอบ (check: C) ในพิจารณาว่า โครงข่ายทำการคัดแยกกลุ่มข้อมูลได้ถูกต้องเพียงใด ถ้าข้อมูลตัวอย่างที่ป้อนให้โครงข่าย ART เป็นข้อมูลหน้าคน จะสามารถพิจารณาได้ 3 กรณีคือ

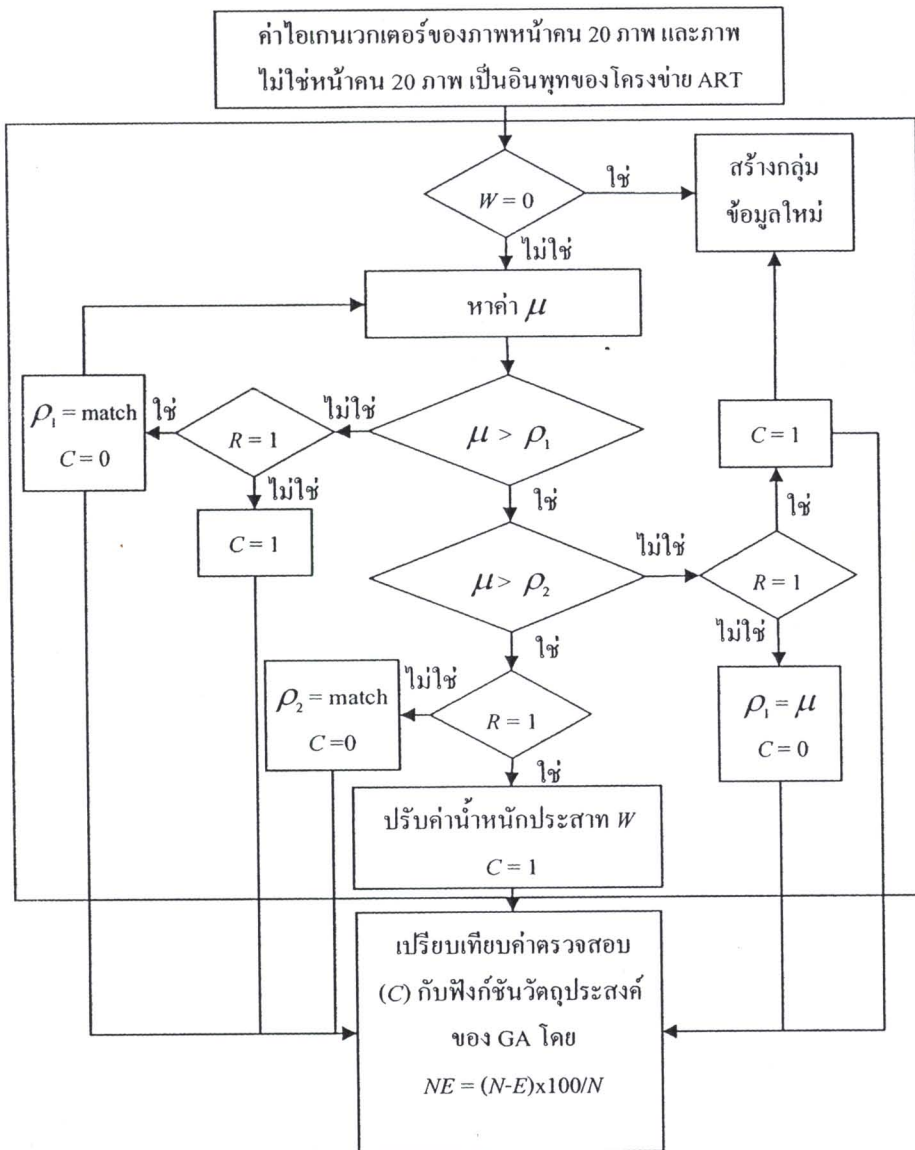
1. ข้อมูลตัวอย่างมีค่า μ มากกว่า ρ_1 และ ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 1 และโครงข่าย ART จะทำการปรับค่าน้ำหนักประสาทของกลุ่มข้อมูล
2. ข้อมูลตัวอย่างมีค่า μ มากกว่าค่า ρ_1 แต่น้อยกว่า ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 1 และโครงข่าย ART จะทำการสร้างกลุ่มข้อมูลใหม่
3. ข้อมูลตัวอย่างมีค่า μ น้อยกว่า ρ_1 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_1 = \mu$

และถ้าข้อมูลตัวอย่างที่ป้อนให้โครงข่าย ART เป็นข้อมูลที่ไม่ใช่หน้าคน จะสามารถพิจารณาได้ 3 กรณีเช่นกันคือ

1. ข้อมูลตัวอย่างมีค่า μ มากกว่า ρ_1 และ ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_2 = \mu$
2. ข้อมูลตัวอย่างมีค่า μ มากกว่าค่า ρ_1 แต่น้อยกว่า ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_1 = \mu$
3. ข้อมูลตัวอย่างมีค่า μ น้อยกว่า ρ_1 จะให้ค่าตรวจสอบมีค่าเป็น 1

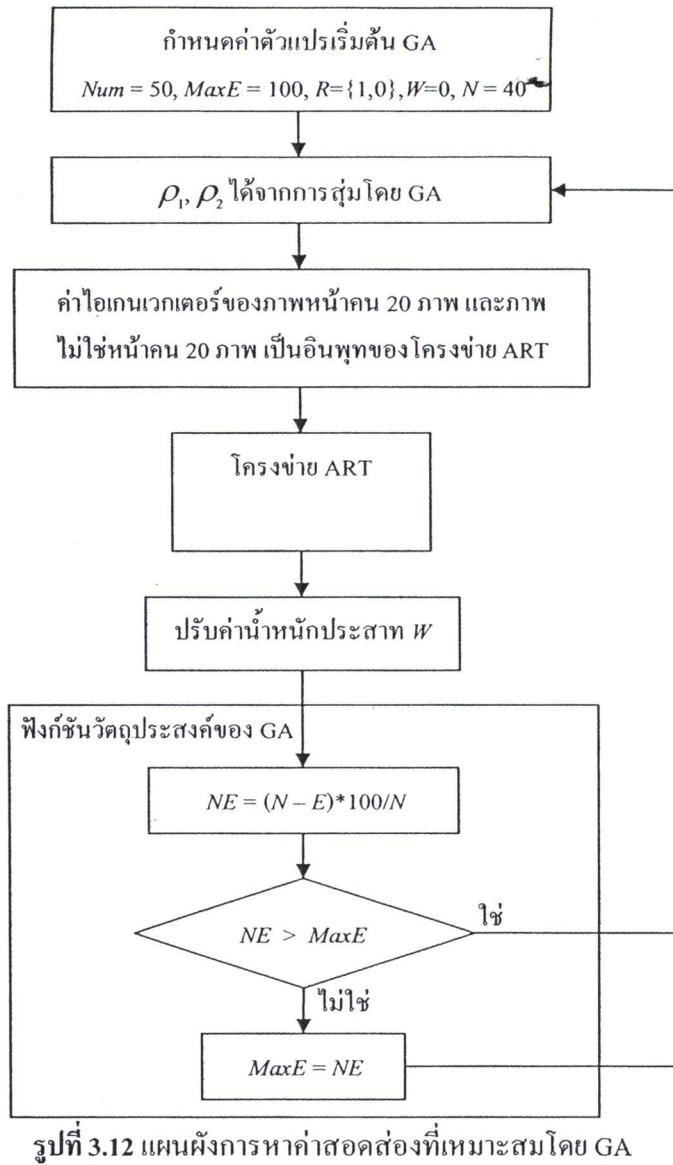
ในงานวิจัยนี้ ได้เพิ่มกลไกการกำหนดค่าพารามิเตอร์สอดส่องที่เหมาะสม โดยใช้จินเนติกอัลกอริทึม (Genetic Algorithm) ทำการค้นหาค่าที่เหมาะสมที่สุด (optimization) แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดส่องที่เหมาะสมด้วย GA มีแสดงในรูปที่ 3.11

เนื่องจากชุดข้อมูลที่เป็นหน้าคนนั้น เป็นข้อมูลที่เกี่ยวข้องกับองค์ประกอบของหน้าในภาพที่มีความหลากหลายของหน้าแตกต่างกันตาม ลักษณะการวางท่า อารมณ์ของหน้า ขนาดของหน้า เป็นต้น ทำให้การกำหนดค่าสอดส่องที่เหมาะสมทำได้ยาก ดังนั้นจึงนำ GA มาช่วยในการหาค่าสอดส่องที่เหมาะสมที่สุดในการแบ่งกลุ่มข้อมูลของโครงข่าย ART โดยทำการสร้างฟังก์ชันวัตถุประสงค์ให้ทำการแยกข้อมูลของหน้าคนจำนวน 20 หน้า และข้อมูลที่ไม่ใช่หน้าคนอีก 20 หน้าผ่านโครงข่าย ART โดยกำหนดค่าพารามิเตอร์เริ่มต้นของ GA ได้แก่จำนวนประชากร (population) จำนวน 50 ประชากร ($Num = 50$) ค่าความผิดพลาดเริ่มต้น เท่ากับ 100 ($MaxE = 100$) ค่าค้นแบบ (R) เป็นเซตแสดงสถานะของอินพุตว่าเป็นหน้าคนหรือไม่ใช่หน้าคน ในที่นี้ใช้ตัวอย่างที่เป็นหน้าคน 20 ตัวอย่าง และไม่ใช่หน้าคน 20 ตัวอย่าง ดังนั้นจะได้ค่า R เป็นเซตของตัวอย่างหน้าคน และตัวอย่างที่ไม่ใช่หน้าคน โดยที่ A มีค่าเป็น 1 และ B มีค่าเป็น 0 ค่าน้ำหนักประสาทเริ่มต้นเป็น 0 ($W = 0$) และจำนวนภาพตัวอย่างในการทดลอง 40 ตัวอย่าง ($N = 40$) จากนั้น ทำการกำหนดค่าตัวแปรใน GA ค่าสอดส่องเริ่มต้น ρ_1 และ ρ_2 โดยการสุ่มเลือกอิสระ และส่งค่าสอดส่องที่ได้จากการสุ่มให้โครงข่าย ART

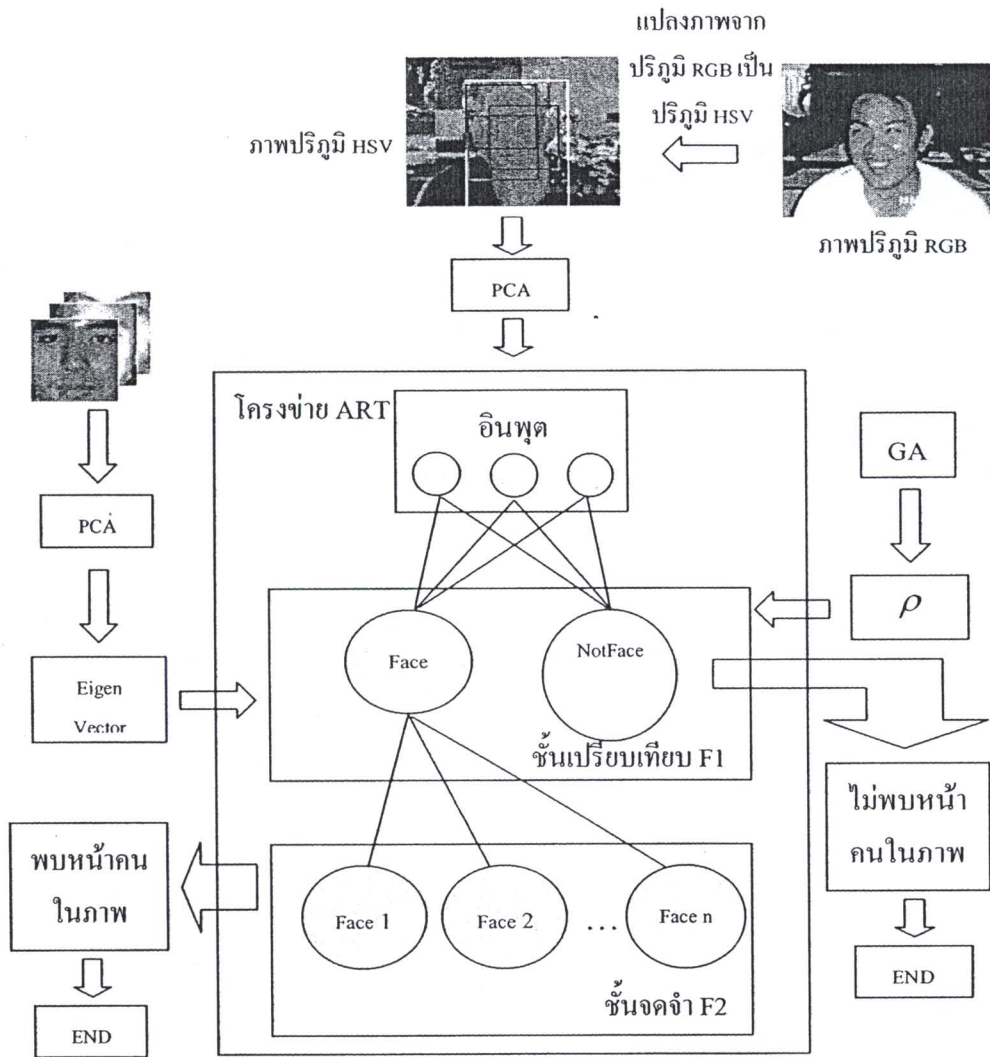


รูปที่ 3.11 แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA

เมื่อได้ค่าตรวจสอบแล้วจะส่งค่าที่ได้ไปยังวัตถุประสงค์ของ GA เพื่อหาค่าความผิดพลาด (Net error: NE) โดย GA จะทำการปรับปรุงพันธุกรรมของประชากรเพื่อให้ได้ตัวแปรที่มีค่าความผิดพลาดที่น้อยที่สุด และสิ้นสุดการทำงานเมื่อ GA ครบจำนวนรอบสูงสุดที่กำหนดไว้ แสดงแผนผังการหาค่าสอดคล้องที่เหมาะสมโดย GA ในรูปที่ 3.12



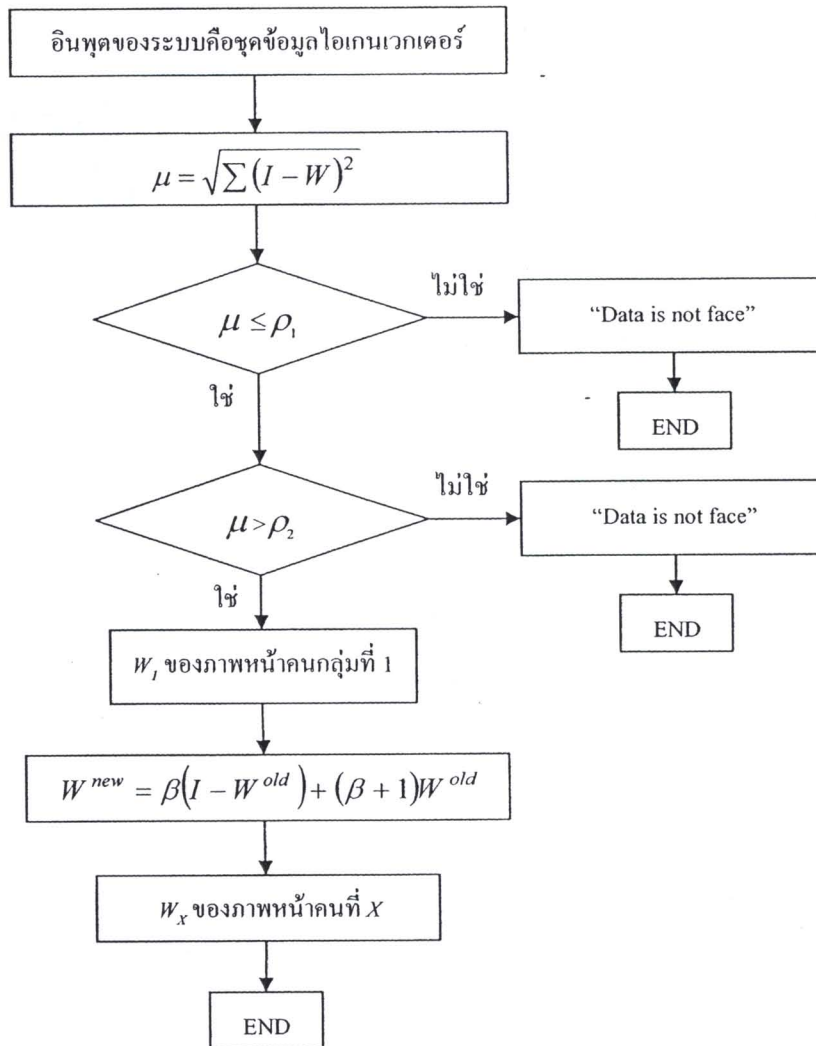
เมื่อได้ค่าสอดคล้องที่เหมาะสมกับระบบแล้ว นำค่าที่ได้นั้นมากำหนดเป็นค่าสอดคล้องของโครงข่าย ART โดยโครงสร้างของระบบการตรวจจับหน้าคนด้วยโครงข่าย ART มีโครงสร้างดังแสดงในรูปที่ 3.13



รูปที่ 3.13 โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART

จากนั้นนำค่าสอดคล้องที่ได้รับการกระบวนการ GA มาใช้ในการเปรียบเทียบกับค่าเทียบเคียงที่ได้จากการคำนวณน้ำหนักประสาทของกลุ่มข้อมูลที่เหมาะสมที่สุด และในชั้นที่ 2 จะทำการเปรียบเทียบเพื่อแยกประเภทของหน้าที่มีลักษณะต่างกัน โดยแสดงโครงสร้างของโครงข่าย ART ในรูปที่ 3.14

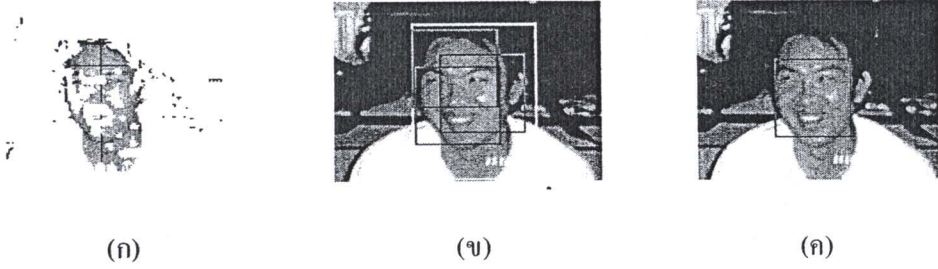




รูปที่ 3.14 แผนผังโครงข่าย ART ในการคัดแยกหน้าคน

นำภาพที่ต้องการตรวจสอบมาแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV จากนั้นมาทำการคัดแยกสีผิวมนุษย์ เพื่อกำหนดขอบเขตในการค้นหาให้แคบลง ได้ขอบเขตในการค้นหาที่คาดว่าจะหน้าจะเป็นหน้าคน เนื่องจากมีค่าอยู่ในช่วงของสีผิวมนุษย์นั้น ต่อมาทำการหาขนาดของกล่องข้อมูลที่จะเข้าโครงข่าย ART ในรูปที่ 3.15 (ข) ในกรอบสี่เหลี่ยมคือขอบเขตทั้งหมดที่มีค่าสีผิวมนุษย์ที่ได้จากการคัดแยกสีผิวมนุษย์อยู่ และกรอบสีแดงคือกล่องข้อมูลที่จะป้อนเข้าโครงข่าย ART โดยสามารถหาได้จาก การหาดำแหน่งของค่าจุดสีของค่าสีผิวมนุษย์ที่มีค่ามากที่สุดทั้งในแนวดิ่ง และแนวนอน เพื่อหาค่าที่น้อยที่สุดของทั้ง 2 แนวเป็นความยาวของกล่องข้อมูล แสดงในรูปที่ 3.15 (ก) เมื่อได้ขนาดของกล่องข้อมูลแล้ว ทำการหาดำแหน่งของกล่องภาพที่จะเป็นอินพุตป้อนให้โครงข่าย ART โดยทำการเลือกกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องข้อมูลไม่น้อยกว่า 2 ใน 3 ของกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องมากที่สุด เป็นข้อมูลอินพุต

ป้อนเข้าโครงข่าย ART โดยทำการเปรียบเทียบกล่องข้อมูลอินพุตที่จะเข้าโครงข่ายกับข้อมูลหน้าในชั้นเปรียบเทียบชั้นแรกเพื่อหาตำแหน่งที่เหมาะสมที่สุดที่มีหน้าคนปรากฏอยู่ดังแสดงในรูปที่ 3.15 (ก)



รูปที่ 3.15 (ก) เส้นกำหนดขนาดกล่องข้อมูล (ข) กล่องข้อมูลอินพุตที่ป้อนเข้าโครงข่าย ART
(ค) กล่องข้อมูลที่ดีที่สุดที่เป็นตำแหน่งของหน้าคนในภาพ

3.4.2 หลักการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART ทำการสร้างโครงข่ายประสาทเทียม แปลงข้อมูลอินพุตของค่าไอเกนเวกเตอร์ในสมการที่ 3.23 ให้อยู่ในรูปของข้อมูลส่วนเติมเต็ม (complement data) แสดงในสมการที่ 3.24

$$v_1, v_2, \dots, v_M \quad (3.23)$$

$$I = (v, v^c) = (v_1, v_2, \dots, v_M, 1 - v_1, 1 - v_2, \dots, 1 - v_M) \quad (3.24)$$

โดย I คือข้อมูลส่วนเติมเต็ม

V คือไอเกนเวกเตอร์

ฝึกสอนโครงข่าย ด้วยกฎการเรียนรู้แบบแข่งขันแสดงในสมการที่ 3.25

$$W_i^{new} = \begin{cases} W_i^{old} & \text{if } \mu \geq \rho \\ \beta(I - W_i^{old}) + (1 + \beta)W_i^{old} & \text{if } \mu < \rho \end{cases} \quad (3.25)$$

โดย W^{new} คือค่าน้ำหนักประสาทใหม่

W^{old} คือค่าน้ำหนักประสาทเดิม

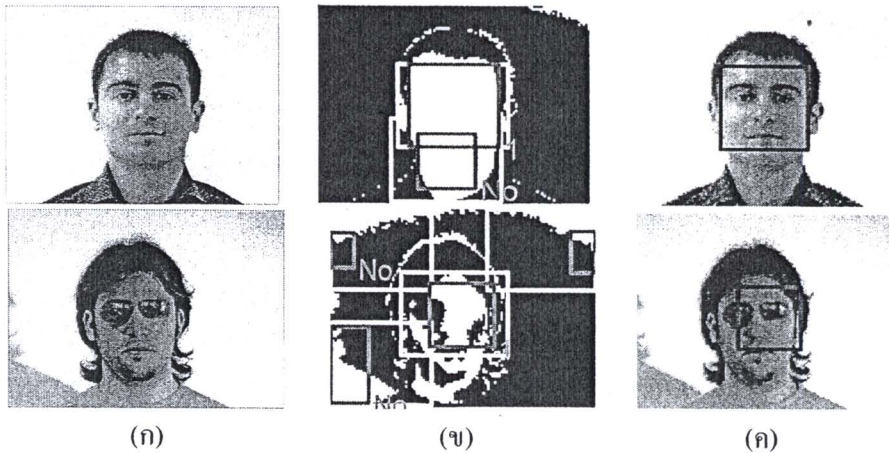
β คือค่าอัตราการเรียนรู้มีค่าอยู่ระหว่าง 0 ถึง 1

โครงข่าย ART จะทำการคัดแยกกลุ่มข้อมูลออกเป็น 2 กลุ่มข้อมูลคือ กลุ่มที่ 1 เป็นข้อมูลของภาพใบหน้าบุคคล และกลุ่มที่ 2 เป็นข้อมูลของภาพที่ไม่ใช่ใบหน้าบุคคล โดยพิจารณาคัดแยกด้วยค่าเทียบเคียง (μ) ดังแสดงในสมการที่ 3.26 เปรียบเทียบกับค่าสอดคล้อง (ρ)

$$\mu = \sqrt{(I - w_i^{old})^2} \quad (3.26)$$

ถ้าค่าเทียบเคียงที่ได้มีค่าน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตกับค่าน้ำหนักประสาทที่ใช้ไม่ได้ของกลุ่มเดียวกัน ระบบจะทำการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทของกลุ่มใหม่ จนกว่าจะได้กลุ่มที่ตรงกัน และถ้าข้อมูลอินพุตที่ได้นี้ ไม่ตรงกับกลุ่มข้อมูลของระบบกลุ่มใดเลย ระบบจะทำการจดจำข้อมูลอินพุตนี้ในกลุ่มที่สร้างขึ้นใหม่ต่อไป

3.4.3 ตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART เมื่อกำหนดขอบเขตการ ค้นหาโดยการพิจารณาบริเวณที่เป็นสีผิวมนุษย์เท่านั้น ระบบจะทำการสร้างกล่องข้อมูลอินพุต และนำกล่อง ข้อมูลนั้นไปแปลงเป็นไอเกนเวกเตอร์เพื่อเป็นอินพุตให้โครงข่าย ART เพื่อตัดแยกความเป็นหน้าคน และ ไม่ใช่หน้าคน โดยเปรียบเทียบค่าเทียบเคียงของข้อมูลกับค่าสอดคล้องของระบบ แสดงได้ในรูปที่ 3.16



รูปที่ 3.16 ภาพตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART

(ก) ตัวอย่างภาพต้นแบบ (ข) ตัวอย่างภาพผลการทำงานของโครงข่าย ART

(ค) ตัวอย่างภาพผลการตรวจจับหน้าคนด้วยโครงข่าย ART

จากรูปที่ 3.16 (ข) พบว่าระบบจะทำการสร้างกล่องข้อมูลอินพุตที่มีค่าสีผิวมนุษย์มาก 1 ใน 4 ของพื้นที่กล่องทั้งหมดเป็นอินพุตของระบบ แสดงในกรอบภาพสี่เหลี่ยม และจุดภาพสีขาวแสดงถึงค่าจุดภาพนั้นมีความน่าจะเป็นที่จะเป็นสีผิวมนุษย์ โดยกล่องข้อมูลนี้อาจมีได้หลายอันขึ้นอยู่กับจำนวนจุดภาพที่เป็นสีผิวมนุษย์ในภาพนั้นมีค่ามากน้อยเพียงใด ต่อจากนั้นระบบจะทำการสร้างกล่องข้อมูลเพื่อเป็นอินพุตให้โครงข่าย แสดงในกรอบภาพสีแดงเพื่อหาจุดที่มีความน่าจะเป็นว่าเป็นหน้าคนภายในกล่องภาพแต่ละกล่อง แล้วแสดงผลการคัดแยกหน้าคน โดยระบบจะแสดงตัวอักษร “No” บนกล่องภาพที่ไม่ใช่หน้าคน และแสดงตัวเลขบนกล่องภาพที่เป็นหน้าคน โดยถ้าในภาพนั้นระบบตรวจพบหน้าคนจำนวน 3 คน ระบบจะแสดงตัวเลขบนกล่องที่เป็นหน้าคนเรียงกันไปจาก 1 ถึง 3