

โครงการวิจัยย่อยลำดับที่ 4
เรื่อง การรู้จำใบหน้า ปีที่ 2 (Face Recognition)

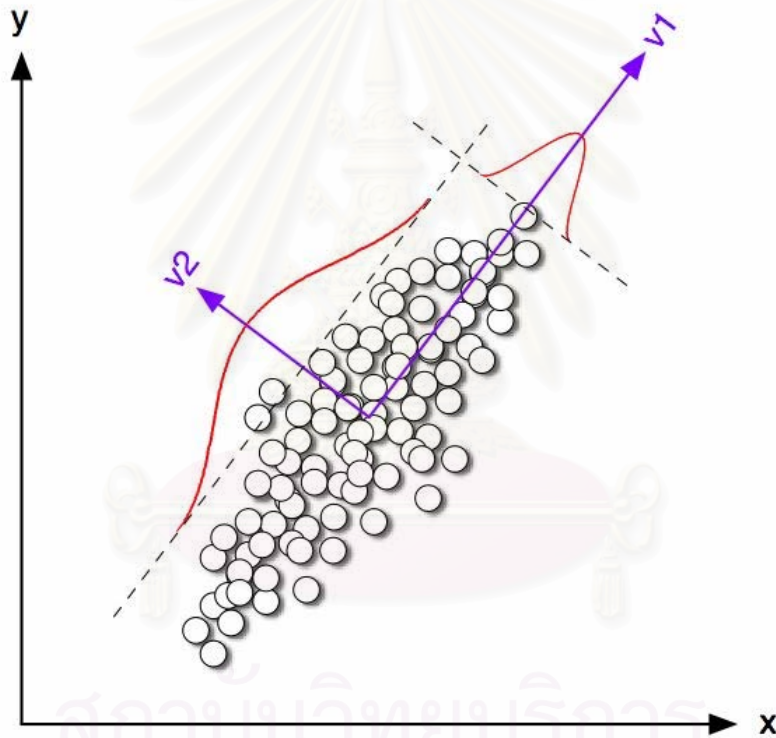
1. ผู้รับผิดชอบโครงการ รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล
2. วัตถุประสงค์ของโครงการ รู้จำใบหน้า

เพื่อพัฒนาวิธีการแยกแยะใบหน้าของบุคคลต่างกันออกจากกันได้ และนำวิธีการที่ได้ไปประยุกต์ใช้กับระบบรู้จำใบหน้า

3. ขอบเขตหรือเป้าหมายของโครงการ ศึกษาวิจัย และประยุกต์ใช้ ระบบการรู้จำใบหน้า
 - 3.1. สร้างระบบรู้จำใบหน้าด้วยเทคนิคที่ปรับปรุงขึ้นใหม่
 - 3.2. รูปใบหน้าคนที่ใช้เป็นภาพ 2 มิติ ที่มีส่วนของใบหน้าครบ
 - 3.3. รูปใบหน้าคนที่ใช้มีการตรวจจับบริเวณใบหน้าที่ต้องการตรวจสอบมาแล้ว
 - 3.4. รูปใบหน้าคนที่ใช้เป็นภาพหน้าตรงมีการหัน ก้มและเงยหน้าไม่เกิน 20 องศาจากภาพหน้าตรง
 - 3.5. รูปใบหน้าคนที่ใช้สามารถแสดงอารมณ์แตกต่างกันได้
 - 3.6. รูปใบหน้าคนที่ใช้ต้องมีขนาดมากกว่า 80 x 80 พิกเซล และคมชัดเมื่อมองด้วยตาเปล่า
 - 3.7. รูปใบหน้าคนที่ใช้ต้องไม่มีสิ่งของที่บดบังบริเวณใบหน้าเกิน 10% ของส่วนที่เป็นใบหน้าทั้งหมด ยกเว้นในกรณีที่บุคคลนั้นสวมใส่สิ่งของดังกล่าวเป็นประจำ
 - 3.8. รูปใบหน้าคนที่ใช้ทั้งหมดที่ใช้ฝึกฝนและทดสอบของบุคคลแต่ละคนจะต้องเก็บข้อมูลภายในช่วงเวลาเดียวกัน
 - 3.9. อัตราความถูกต้องเฉลี่ยของระบบรู้จำไม่ต่ำกว่าร้อยละ 90 บนฐานข้อมูลที่มีบุคคลต่างกันไม่เกิน 40 คน และมีจำนวนภาพที่ใช้ฝึกฝนน้อยกว่าจำนวนภาพที่ใช้ทดสอบไม่เกินร้อยละ 50
4. ส่วนงานที่ได้ดำเนินการไปแล้ว **Equation Section 4**
 - 4.1. ศึกษาวิธีการรู้จำใบหน้าแบบต่างๆ และงานที่เกี่ยวข้องโดยรวมสรุปได้ดังนี้
 - วิธี Subspace ได้แก่ PCA, LDA, ICA, 2DPCA, B2DPCA
 - วิธี Kernel ได้แก่ KPCA, KLDA (KFDA), K2DPCA
 - วิธี Multiresolution ได้แก่ Wavelet face
 - 4.2. ศึกษาการรู้จำใบหน้าด้วยเทคนิคต่างๆ ดังนี้

Principal Component Analysis (PCA)

วิธีการรู้จำใบหน้าที่ใช้ PCA เป็นหลักหรือเรียกว่า Eigenface ซึ่งมีแนวคิดคือ ฉายค่าความเข้มแสงของแต่ละพิกเซล (Pixel) ในภาพใบหน้าของกลุ่มตัวอย่างไปบนปริภูมิใหม่ซึ่งทำให้ความแปรปรวนร่วม (Covariance) มากที่สุด ในขณะที่แกนแต่ละแกนตั้งฉากกันและผ่านจุดศูนย์กลาง (Centroid) ด้วยการแปลงเชิงเส้น (Linear transformation) แกนดังกล่าวเรียกว่าแกน Principal Component (PC) โดยที่แกนแต่ละแกนแทนได้ด้วย เวกเตอร์เจาะจง (Eigen vector) จำนวนหนึ่งของเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) ที่สอดคล้องกับค่าเจาะจง (Eigen value) ที่มากที่สุดจำนวนหนึ่งดังรูปที่ 1 [1]



รูปที่ 1 ตัวอย่างการแปลง PCA แบบ 2 มิติ โดยแปลงจากปริภูมิคาร์ทีเซียน (x,y) ไปบนปริภูมิใหม่ (v_1,v_2)

การรู้จำใบหน้าภายในปริภูมิใหม่ที่เกิดจากการแปลงเชิงเส้นที่กล่าวมาทำได้โดยการวัดระยะยูคลิด (Euclidean distance) ระหว่างใบหน้าที่ต้องการทดสอบ กับใบหน้าที่ใช้ฝึกฝนทั้งหมด หากระยะทางที่สั้นที่สุดนั้นเกิดกับใบหน้าที่ใช้ฝึกฝนใด เราจะตัดสินว่าใบหน้าทดสอบดังกล่าวจะอยู่ในกลุ่ม (Class) เดียวกันกับใบหน้าฝึกฝนนั้น

ระเบียบวิธี PCA มีขั้นตอนดังนี้

กำหนดให้ Γ คือเวกเตอร์ของความเข้มแสงของพิกเซลของภาพที่ทำ Vectorization ซึ่งหลัก จากรูปที่นำเข้าไปเป็นกลุ่มฝึกฝน (Training set) มีขนาด $m \times n$ ให้เป็นเวกเตอร์ที่มีขนาด $mn \times 1$ และ Ψ คือเวกเตอร์เฉลี่ยของทุกรูปในกลุ่มฝึกสอนดังสมการที่ (4.1)

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (4.1)$$

โดย M คือจำนวนภาพทั้งหมดที่ใช้ในการฝึกฝน จากนั้นนำไปคำนวณหาเวกเตอร์ Φ ซึ่งคือ การทำให้ค่าเฉลี่ยของชุดฝึกฝนเป็นศูนย์ดังสมการที่ (4.2)

$$\Phi_i = \Gamma_i - \Psi, \quad (i=1, \dots, M) \quad (4.2)$$

กำหนดให้เมตริกซ์ \mathbf{A} คือเมตริกซ์ที่แต่ละหลักคือเวกเตอร์ Φ_i จำนวน M เวกเตอร์ดังสมการที่ (4.3)

$$\mathbf{A} = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M] \quad (4.3)$$

จากนั้นคำนวณเมตริกซ์ความแปรปรวนร่วม \mathbf{C} ของเมตริกซ์ \mathbf{A} ดังสมการที่ (4.4)

$$\mathbf{C} = \mathbf{A}\mathbf{A}^T \quad (4.4)$$

โดยเมตริกซ์ความแปรปรวนร่วม \mathbf{C} มีมิติขนาด $mn \times mn$ ซึ่งเป็นมิติขนาดใหญ่มาก ซึ่งไม่เหมาะสมที่จะนำไปหาค่าเจาะจง เนื่องจากต้องใช้เวลาและหน่วยความจำจำนวนมากในการคำนวณ แต่เนื่องจากจำนวนค่าเจาะจงที่ไม่เป็นศูนย์มีค่าเท่ากับจำนวนชั้น และจากคุณสมบัติดังสมการที่ (4.5)

$$\text{rank}(\mathbf{C}) = \text{rank}(\mathbf{A}\mathbf{A}^T) = \text{rank}(\mathbf{A}^T\mathbf{A}) = \text{rank}(\mathbf{L}) \quad (4.5)$$

โดย

$$\mathbf{L} = \mathbf{A}^T \mathbf{A} \quad (4.6)$$

ซึ่งเมตริกซ์ \mathbf{L} มีมิติขนาดเพียง $M \times M$ ดังนั้นการคำนวณค่าเฉพาะของเมตริกซ์ \mathbf{L} ใช้เวลาและหน่วยความจำน้อยกว่าการคำนวณเมตริกซ์ \mathbf{C} โดยตรง ดังนั้นจึงนำเมตริกซ์ \mathbf{L} ไปคำนวณค่าเฉพาะดังสมการที่ (4.7)

$$\mathbf{L}\mathbf{V} = \lambda\mathbf{V} \quad (4.7)$$

โดย λ คือเมตริกซ์เชิงที่มีค่าเฉพาะ λ ของเมตริกซ์ \mathbf{L} เป็นแกนทแยงมุมหลัก และ \mathbf{V} เป็นเมตริกซ์ขนาด $M \times M$ ที่แต่ละหลักคือเวกเตอร์เฉพาะของเมตริกซ์ \mathbf{L}

จากนั้นทำการหาเวกเตอร์เฉพาะของเมตริกซ์ \mathbf{C} ดังสมการที่ (4.8)

$$\mathbf{U} = \mathbf{A}\mathbf{V} \quad (4.8)$$

โดยเมตริกซ์ \mathbf{U} คือเมตริกซ์ที่แต่ละหลักคือเวกเตอร์เฉพาะของเมตริกซ์ \mathbf{C} ขนาด $mn \times M$ หรือคือเวกเตอร์เฉพาะของเมตริกซ์ \mathbf{C} จำนวน M ตัว

เนื่องจากเวกเตอร์เฉพาะที่สอดคล้องกับค่าเฉพาะที่มากที่สุดจะมีนัยสำคัญทางข้อมูลมากที่สุด ดังนั้นเราสามารถเลือกเวกเตอร์เฉพาะที่มากที่สุดเพียง d ตัว ดังนั้นเราสามารถลดขนาดมิติของเวกเตอร์คุณลักษณะได้อีกดังสมการที่ (4.9)

$$\Omega_i = \mathbf{W}^T \Phi_i \quad (4.9)$$

โดย Ω_i คือเวกเตอร์คุณลักษณะของภาพที่ i ในชุดฝึกฝน มีขนาด $d \times 1$ และเมตริกซ์ \mathbf{W} คือเมตริกซ์ที่แต่ละหลักคือเวกเตอร์เฉพาะของเมตริกซ์ \mathbf{C} จำนวน d ตัว

กำหนดให้ \mathbf{Y} คือเวกเตอร์ของภาพที่ต้องการทดสอบ เราสามารถฉายเวกเตอร์ \mathbf{Y} ไปบนปริภูมิคุณลักษณะได้ดังสมการที่

$$\omega = \mathbf{W}^T (\mathbf{Y} - \Psi) \quad (4.10)$$

โดย ω คือเวกเตอร์คุณลักษณะของภาพที่ต้องการทดสอบซึ่งอยู่ในปริภูมิเดียวกันกับเวกเตอร์ Ω ดังนั้นเราสามารถทราบได้ว่าเวกเตอร์ ω อยู่ใกล้กับเวกเตอร์ Ω ได้ได้ด้วยการหา ระยะยุคลิดที่สั้นที่สุดดังสมการที่ (4.11)

$$\mathcal{E} = \min_i (\|\Omega - \Omega_i\|^2) \quad (4.11)$$

โดย \mathcal{E} คือระยะยุคลิดระหว่างเวกเตอร์ ω และเวกเตอร์ Ω ที่สั้นที่สุด ดังนั้นเราสามารถกล่าวได้ว่าเวกเตอร์คุณลักษณะของภาพทดสอบอยู่ใกล้เวกเตอร์คุณลักษณะของภาพในชุดฝึกฝนในกลุ่มใด ภาพทดสอบดังกล่าวควรเป็นภาพของคนเดียวกันในกลุ่มนั้น



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Two-Dimensional Principal Component Analysis (2DPCA)

ในวิธีการ PCA ภาพ 2 มิติจะต้องถูกทำ Vectorization เพื่อให้เป็นเวกเตอร์ 1 มิติ เพื่อใช้คำนวณหาเมตริกซ์ความแปรปรวนร่วม ส่งผลทำให้ขนาดของเมตริกซ์ความแปรปรวนร่วมดังกล่าวมีขนาดใหญ่ทำให้การคำนวณหาค่าเจาะจงทำได้ยากและต้องใช้เวลาและหน่วยความจำจำนวนมาก และยังทำให้เกิดปัญหา Small Simple Size (SSS) หรือ Singularity ซึ่งเกิดขึ้นเป็นจำนวนตัวอย่างที่ใช้ฝึกฝนน้อยกว่าขนาดของเมตริกซ์ความแปรปรวนร่วมมาก ๆ เพื่อแก้ไขปัญหาต่าง ๆ เหล่านี้ 2DPCA ถูกเสนอขึ้นโดยนิยามเมตริกซ์ความแปรปรวนร่วมใหม่ เรียกว่า เมตริกซ์ความแปรปรวนร่วมเชิงภาพ (Image covariance matrix) ซึ่งคำนวณได้โดยตรงจากภาพแบบ 2 มิติ โดยที่ไม่ต้องทำ Vectorization ในขั้นตอนแรก ด้วยหลักการนี้เมตริกซ์ความแปรปรวนร่วมเชิงภาพจึงมีขนาดเล็กกว่าเดิมมาก คือ มีขนาดเท่ากับควมกว้างของรูปเดี่ยวเท่านั้น

ระเบียบวิธี 2DPCA มีขั้นตอนดังนี้

กำหนดให้เมตริกซ์ \mathbf{A} แทนภาพขนาด $m \times n$ และ \mathbf{x} เป็นเวกเตอร์การฉายขนาด $n \times 1$ เราสามารถทำการฉายเชิงเส้นได้ดังนี้

$$\mathbf{y} = \mathbf{Ax} \quad (4.12)$$

โดย \mathbf{y} คือ เวกเตอร์คุณสมบัติฉาย (Projected feature vector) ซึ่งมีมิติขนาด $m \times 1$ หรือเรียกอีกอย่างหนึ่งว่า Principal Component Vector (PCV) ซึ่งต่างกับในวิธี PCA เพื่อต้องการให้การกระจายของเวกเตอร์คุณสมบัติฉายมีค่ามากที่สุด เราจึงใช้บรรทัดฐานในการทำให้ฟังก์ชันผลบวกบนเส้นทแยงมุมหลัก (Trace) ของเมตริกซ์ความแปรปรวนร่วมมากที่สุดเช่นเดียวกันใน PCA ดังนี้

$$J(\mathbf{X}) = \text{tr}(\mathbf{S}_x) \quad (4.13)$$

โดย \mathbf{S}_x คือ เมตริกซ์ความแปรปรวนร่วม

$$\mathbf{S}_x = E[(\mathbf{y} - E\mathbf{y})(\mathbf{y} - E\mathbf{y})^T] \quad (4.14)$$

ดังนั้นจากสมการที่ (4.12) และ (4.14) Trace ของเมตริกซ์ความแปรปรวนร่วมคือ

$$\begin{aligned}
tr(\mathbf{S}_x) &= tr\{E[(\mathbf{y} - E\mathbf{y})(\mathbf{y} - E\mathbf{y})^T]\} \\
&= tr\{E[(\mathbf{A} - E\mathbf{A})\mathbf{x}\mathbf{x}^T(\mathbf{A} - E\mathbf{A})^T]\} \\
&= tr\{E[\mathbf{x}^T(\mathbf{A} - E\mathbf{A})^T(\mathbf{A} - E\mathbf{A})\mathbf{x}]\} \\
&= tr\{\mathbf{x}^T E[(\mathbf{A} - E\mathbf{A})^T(\mathbf{A} - E\mathbf{A})]\mathbf{x}\} \\
&= tr\{\mathbf{x}^T \mathbf{G}\mathbf{x}\}
\end{aligned} \tag{4.15}$$

กำหนดให้

$$\mathbf{G} = E[(\mathbf{A} - E\mathbf{A})^T(\mathbf{A} - E\mathbf{A})] \tag{4.16}$$

โดย \mathbf{G} คือ เมตริกซ์ความแปรปรวนร่วมเชิงภาพ (Image covariance (scatter) matrix) ซึ่ง \mathbf{G} เป็นเมตริกซ์ Nonnegative definite ขนาด $n \times n$ ซึ่งเท่ากับขนาดความกว้างของภาพเดี่ยวเท่านั้น ดังนั้นเราสามารถคำนวณค่าเจาะจงของ \mathbf{G} ได้ง่ายและเร็วกว่า กำหนดให้มีจำนวนตัวอย่างภาพที่ใช้ฝึกฝน M ภาพ ดังนั้นเราสามารถคำนวณหาเมตริกซ์ความแปรปรวนร่วมเชิงภาพได้ดังนี้

$$\mathbf{G} = \frac{1}{M} \sum_{k=1}^M (\mathbf{A}_k - \bar{\mathbf{A}})^T (\mathbf{A}_k - \bar{\mathbf{A}}) \tag{4.17}$$

ซึ่งเวกเตอร์เจาะจงของ \mathbf{G} ก็คือ \mathbf{x} เช่นเดียวกับ PCA ค่าเจาะจงที่มากที่สุดค่าเดียวอาจไม่เพียงพอในการรู้จำ ดังนั้นเราจะเลือกค่าเจาะจงที่มากที่สุด d ลำดับ ดังนั้นการสกัดคุณลักษณะ (Feature Extraction) ทำได้ดังนี้

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \tag{4.18}$$

โดย $\mathbf{Y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_d]$ คือ เมตริกซ์คุณลักษณะมิติขนาด $m \times d$ ของ \mathbf{A} และ $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_d]$ เป็นเมตริกซ์การฉายมิติขนาด $n \times d$

กำหนดให้ \mathbf{A}_i เป็นภาพที่ต้องการทดสอบ ดังนั้นเราสามารถหาเมตริกซ์คุณลักษณะได้โดยการฉายดังสมการที่ (4.19)

$$\mathbf{B} = \mathbf{A}_i \mathbf{X} \tag{4.19}$$

เราสามารถวัดระยะห่างระหว่างเมตริกซ์ของภาพทดสอบและภาพที่ใช้ฝึกฝนได้โดยใช้สมการที่ (4.20)

$$d(\mathbf{B}, \mathbf{Y}_i) = \sum_{k=1}^d \|\mathbf{B}^{(k)} - \mathbf{Y}_i^{(k)}\|_2 \quad (4.20)$$

ดังนั้นหาก

$$i = \operatorname{argmin}(d(\mathbf{B}, \mathbf{Y}_i)) \quad (4.21)$$

เราจะกล่าวได้ว่า \mathbf{A}_i ควรอยู่ในกลุ่มที่ i



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Bilateral-projection based 2DPCA (B2DPCA)

B2DPCA เป็น 2DPCA ที่ใช้การฉายทั้งสองข้าง คือ ตามแนวแถวและหลักโดยคำนวณเมตริกซ์การฉายสองเมตริกซ์ดังนี้

$$\mathbf{B} = \mathbf{Z}^T \mathbf{A} \mathbf{X} \quad (4.22)$$

โดย \mathbf{X} คือเมตริกซ์การฉายเดิมใน 2DPCA ดังสมการที่ (4.18) และ \mathbf{Z} คือเมตริกซ์การฉายทางซ้ายซึ่งต้องกำหนดเมตริกซ์ความแปรปรวนร่วมเชิงภาพขึ้นใหม่ตามแนวหลักดังนี้

$$\mathbf{H} = \frac{1}{M} \sum_{k=1}^M (\mathbf{A}_k - \bar{\mathbf{A}})(\mathbf{A}_k - \bar{\mathbf{A}})^T \quad (4.23)$$

และหาด้วยวิธีการเดียวกันกับ \mathbf{X} หลังจากการฉายภาพ \mathbf{A} จะได้เมตริกซ์คุณลักษณะที่มีขนาดเล็กลงจากเดิม $m \times d$ เหลือเพียง $q \times d$ ซึ่ง q คือจำนวนค่าเจาะจงที่เลือกใช้ภายใน \mathbf{Z} ดังนั้น $q \leq m$ จากการทดลองใน [17] และ [18] พบว่าใช้เวลาในการรู้จำลดลงและอัตราการเรียนรู้จำสูงขึ้น สำหรับวิธีการวัดระยะห่างนั้นทำได้โดยสมการที่ (4.24)

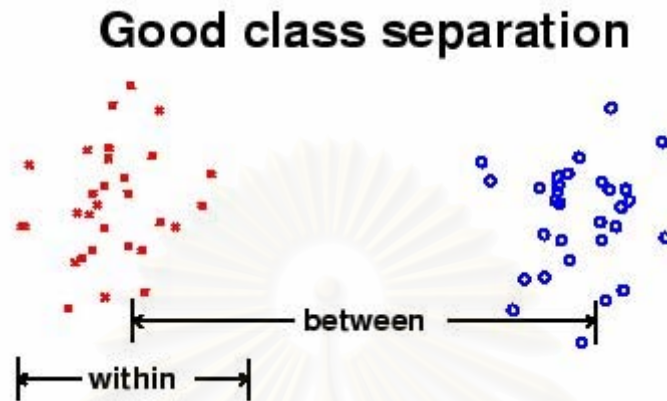
$$d(\mathbf{B}_i, \mathbf{B}_j) = \sqrt{\sum_{m=1}^q \sum_{n=1}^d (\mathbf{b}_{(m,n)_i} - \mathbf{b}_{(m,n)_j})^2} \quad (4.24)$$

ซึ่งในการทำงานเกี่ยวกับการรู้จำใน 2DPCA และ PCA ระยะห่างของตัวอย่างทดสอบใกล้เคียงกับตัวอย่างฝึกฝนได้ที่สุดเราจะจัดตัวอย่างทดสอบนั้นอยู่ในกลุ่มเดียวกับตัวอย่างฝึกฝนนั้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

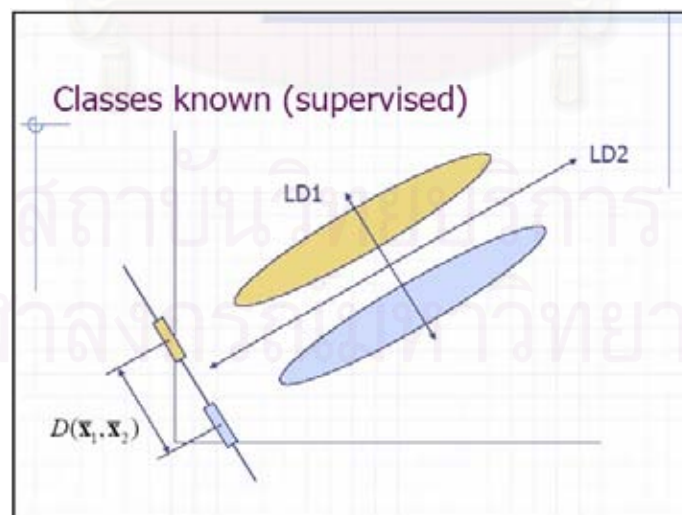
Linear Discriminant Analysis (LDA)

พิจารณาการกระจายตัวของข้อมูลระหว่างข้อมูล 2 กลุ่ม ดังรูปที่ 2 [2]



รูปที่ 2 ความสัมพันธ์ระหว่าง (between) และภายใน (within) กลุ่ม

การกระจายตัวของข้อมูลที่ดี (Good class separation) จะต้องมีความสัมพันธ์ระหว่างกลุ่มมาก ในขณะที่มีความสัมพันธ์ภายในกลุ่มน้อย หลักการของ LDA คือกำหนดบรรทัดฐานโดยให้การกระจายตัวภายในกลุ่มน้อยที่สุด แต่การกระจายตัวระหว่างกลุ่มมากที่สุดในเวลาเดียวกันดังรูปที่ 3 [2]



รูปที่ 3 ตัวอย่างการแปลง LDA แบบ 2 มิติ โดยแปลงจากปริภูมิคาร์ทีเซียน (x,y) ไปบนปริภูมิใหม่ $(LD1,LD2)$

กำหนดให้ \mathbf{S}_w เป็นเมตริกซ์การกระจายภายในกลุ่ม และ \mathbf{S}_b เป็นเมตริกซ์การกระจายระหว่างกลุ่มโดยบรรทัดฐาน LDA ดังนี้

$$J(\mathbf{w}) = \max \left(\frac{\mathbf{S}_b}{\mathbf{S}_w} \right) \quad (4.25)$$

โดยเวกเตอร์ \mathbf{w} เป็นเวกเตอร์เจาะจงของเมตริกซ์ $\mathbf{S}_w^{-1}\mathbf{S}_b$ เนื่องจากเวกเตอร์เจาะจงเพียงเวกเตอร์เดียวไม่เพียงพอในการรู้จำดังนั้นกำหนดให้

$$\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{w}_q] \quad (4.26)$$

โดย \mathbf{W} คือเมตริกซ์การฉายที่มีแต่ละหลักเป็นเวกเตอร์การฉาย \mathbf{w} จำนวน q เวกเตอร์ที่สอดคล้องกับค่าเจาะจงของเมตริกซ์ $\mathbf{S}_w^{-1}\mathbf{S}_b$ ที่มากที่สุด q ค่า

หลังการจากฝึกฝนแล้วเราจะได้เมตริกซ์การฉาย \mathbf{W} ในลักษณะเดียวกับวิธี PCA การฉายหรือการสกัดคุณลักษณะสามารถทำได้ดังสมการที่ (4.10) เช่นเดียวกัน และการวัดระยะห่างระหว่างคู่ของเวกเตอร์คุณลักษณะใด ๆ สามารถทำได้เช่นเดียวกับ PCA ดังสมการที่ (4.11) และใช้หลักเกณฑ์เดียวกันกับในวิธี PCA

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Two-Dimensional Linear Discriminant Analysis (2DLDA)

เราได้เสนอ 2DLDA ขึ้นด้วยหลักการเดียวกับ 2DPCA คือ กำหนดเมตริกซ์ความแปรปรวนร่วมขึ้นใหม่โดยใช้เมตริกซ์ของภาพตัวอย่างแบบ 2 มิติแทนเวกเตอร์แบบ 1 มิติ กำหนดให้ \mathbf{v} เป็นเวกเตอร์คุณลักษณะ \mathbf{A} เป็นเมตริกซ์ของภาพตัวอย่าง และ \mathbf{z} คือเวกเตอร์การฉายเชิงเส้น ดังนั้นเราสามารถเขียนการแปลงเชิงเส้นได้ดังนี้

$$\mathbf{v} = \mathbf{A}\mathbf{z} \quad (4.27)$$

โดยเราสามารถหาเวกเตอร์ \mathbf{z} ที่สอดคล้องกับบรรทัดฐานที่ทำให้เมตริกซ์การกระจายตัวระหว่างกลุ่ม (Between-class scatter matrix) \mathbf{S}_b มากที่สุด ในขณะที่เดียวกันก็ต้องทำให้เมตริกซ์การกระจายตัวภายในกลุ่ม (Within-class scatter matrix) \mathbf{S}_w มีค่าน้อยที่สุด เช่นเดียวกับบรรทัดฐานเดิมของ LDA

$$J(\mathbf{z}) = \frac{\text{tr}(\mathbf{S}_b)}{\text{tr}(\mathbf{S}_w)} \quad (4.28)$$

โดย \mathbf{S}_b และ \mathbf{S}_w สามารถหาได้จากสมการที่ (4.29) และ (4.30) ตามลำดับ

$$\mathbf{S}_b = \sum_{i=1}^K Pr(\omega_i) E[(E[\mathbf{v} | \omega = \omega_i] - E\mathbf{v})(E[\mathbf{v} | \omega = \omega_i] - E\mathbf{v})^T] \quad (4.29)$$

$$\mathbf{S}_w = \sum_{i=1}^K Pr(\omega_i) E[(\mathbf{v} - E\mathbf{v})(\mathbf{v} - E\mathbf{v})^T | \omega = \omega_i] \quad (4.30)$$

โดย $Pr(\omega_i)$ คือ ความน่าจะเป็นของกลุ่ม ω_i และ K คือ จำนวนกลุ่มทั้งหมดที่ใช้ฝึกฝน จากสมการที่ (4.28) (4.29) และ (4.30) จะได้

$$\begin{aligned}
J(\mathbf{z}) &= \frac{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) E \left[\left(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A} \right) \mathbf{z} \mathbf{z}^T \left(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A} \right)^T \right] \right)}{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) E \left[\left(\mathbf{A} - E\mathbf{A} \right) \mathbf{z} \mathbf{z}^T \left(\mathbf{A} - E\mathbf{A} \right)^T | \omega = \omega_i \right] \right)} \\
&= \frac{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) \mathbf{z}^T E \left[\left(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A} \right)^T \left(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A} \right) \right] \mathbf{z} \right)}{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) \mathbf{z}^T E \left[\left(\mathbf{A} - E\mathbf{A} \right)^T \left(\mathbf{A} - E\mathbf{A} \right) | \omega = \omega_i \right] \mathbf{z} \right)} \quad (4.31) \\
&= \frac{\text{tr} \left(\mathbf{z}^T \tilde{\mathbf{S}}_b \mathbf{z} \right)}{\text{tr} \left(\mathbf{z}^T \tilde{\mathbf{S}}_w \mathbf{z} \right)}
\end{aligned}$$

โดย

$$\tilde{\mathbf{S}}_b = \sum_{i=1}^K \frac{n_i}{K} (\bar{\mathbf{A}}_i - \bar{\mathbf{A}})^T (\bar{\mathbf{A}}_i - \bar{\mathbf{A}}) \quad (4.32)$$

$$\tilde{\mathbf{S}}_w = \sum_{i=1}^K \frac{n_i}{K} \sum_{\mathbf{A}_k \in \omega_i} (\mathbf{A}_k - \bar{\mathbf{A}}_i)^T (\mathbf{A}_k - \bar{\mathbf{A}}_i) \quad (4.33)$$

ซึ่ง \mathbf{z} สามารถหาได้จากเวกเตอร์เจาะจงของเมตริกซ์ $\tilde{\mathbf{S}}_w^{-1} \tilde{\mathbf{S}}_b$ ดังสมการที่ (4.34)

$$\tilde{\mathbf{S}}_b \mathbf{z} = \lambda \tilde{\mathbf{S}}_w \mathbf{z} \quad (4.34)$$

โดย λ คือค่าเจาะจงของเมตริกซ์ $\tilde{\mathbf{S}}_w^{-1} \tilde{\mathbf{S}}_b$ เช่นเดียวกับวิธี PCA เวกเตอร์เจาะจงเพียงเวกเตอร์เดียวไม่เพียงพอในการรู้จำ จึงต้องใช้เวกเตอร์เจาะจงจำนวนหนึ่งที่สอดคล้องกับค่าเจาะจงที่มากที่สุด q ค่า โดยกำหนดให้เมตริกซ์ $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_q]$ เป็นเมตริกซ์คุณลักษณะของภาพตัวอย่าง และ $\mathbf{Z} = [\mathbf{z}_1 \ \mathbf{z}_2 \ \dots \ \mathbf{z}_q]$ ดังการแปลงเชิงเส้น

$$\mathbf{V} = \mathbf{AZ} \quad (4.35)$$

ซึ่งการวัดระยะห่างทำได้แบบเดียวกับ 2DPCA

Two-Dimensional Linear Discriminant Analysis of Principal Component Vectors

เนื่องจาก 2DPCA เป็นการฉายภาพตัวอย่างไปบนปริภูมิคุณลักษณะ โดยการทำให้เมตริกซ์ความแปรปรวนร่วมของข้อมูลทั้งหมดมากที่สุดเท่าที่นั่น ซึ่งเป็นวิธีแบบ Unsupervised จึงไม่มีข้อมูลเกี่ยวกับกลุ่มอยู่ อาจไม่เพียงพอในการใช้แยกแยะกลุ่ม ในขณะที่เดียวกัน 2DLDA เป็นวิธีแบบ Supervised ซึ่งสามารถทำให้การกระจายตัวของข้อมูลระหว่างกลุ่มมีค่ามาก แต่ภายในกลุ่มมีค่าน้อย อย่างไรก็ตาม 2DLDA เพียงวิธีเดียวจะใช้ข้อมูลทั้งหมดเพื่อการแยกแยะกลุ่ม ในบางครั้งข้อมูลดังกล่าวไม่เหมาะสมที่จะใช้ในการแยกแยะข้อมูล เช่น บริเวณพื้นหลังของภาพ เป็นต้น แต่ในระบบที่เราเสนอขึ้นนี้ใช้ 2DPCA เพื่อลดมิติของข้อมูลที่ไม่สำคัญลงก่อน จากนั้นนำไปเป็นเมตริกซ์คุณลักษณะป้อนเข้าให้กับ 2DLDA ทำให้ประสิทธิภาพการรู้จำเพิ่มขึ้น จากสมการที่ (4.32) และ (4.33) จะเปลี่ยนเป็นสมการที่ (4.36) และ (4.37) ตามลำดับ

$$\tilde{\mathbf{S}}_b^{\mathbf{Y}} = \sum_{i=1}^K \frac{n_i}{K} (\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}})^T (\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}}) \quad (4.36)$$

$$\tilde{\mathbf{S}}_w^{\mathbf{Y}} = \sum_{i=1}^K \frac{n_i}{K} \sum_{\mathbf{Y}_k \in \omega_i} (\mathbf{Y}_k - \bar{\mathbf{Y}}_i)^T (\mathbf{Y}_k - \bar{\mathbf{Y}}_i) \quad (4.37)$$

โดย

$$\bar{\mathbf{Y}} = \frac{1}{M} \sum_{k=1}^M \mathbf{Y}_k \quad (4.38)$$

ดังนั้นการแปลงเชิงเส้นใหม่ได้โดยเมตริกซ์การฉายเชิงเส้น $\mathbf{L} = \mathbf{XZ}$ โดย \mathbf{X} และ \mathbf{Z} เป็นเมตริกซ์การฉายจากสมการที่ (4.18) และ (4.35)

$$\mathbf{D} = \mathbf{AL} \quad (4.39)$$

โดย \mathbf{D} คือ เมตริกซ์คุณลักษณะในปริภูมิคุณลักษณะ และสามารถใช้ในการหาระยะห่างเดียวกันกับวิธี 2DPCA

Face-Specific Subspace (FSS)

ใน PCA 2DPCA หรือ BDPCA นั้นเป็นวิธีแบบ Unsupervised คือไม่มีข้อมูลเกี่ยวกับกลุ่ม และใช้ตัวอย่างทั้งหมดในการฝึกฝน เพื่อให้ได้เมตริกซ์การฉาย และปริภูมิคุณลักษณะเพียงหนึ่งเดียว ตัวอย่างทั้งหมดทั้งตัวอย่างฝึกฝน และตัวอย่างทดสอบจะถูกฉายมาบนปริภูมินี้ จากนั้นจะวัดระยะห่างกันของตัวอย่างภายในปริภูมินี้ เพื่อเพิ่มประสิทธิภาพในกับ PCA ระบบรู้จักด้วยปริภูมีย่อยหน้าเจาะจง (Face-Specific Subspace, FSS) จึงถูกเสนอขึ้นใน [19] เพื่อปรับปรุงประสิทธิภาพของ PCA

ดังนั้นเมื่อปริภูมีย่อยหน้าเจาะจงสามารถเพิ่มประสิทธิภาพของ PCA ได้ดังนั้นควรเพิ่มประสิทธิภาพให้กับ 2DPCA และ B2DPCA ได้

ในกรณี 2DPCA กำหนดให้ \mathbf{G}_k คือเมตริกซ์ความแปรปรวนร่วมเชิงภาพของกลุ่มที่ k โดยใช้ตัวอย่างภาพในกลุ่มที่ k มาคำนวณดังนี้

$$\mathbf{G}_k = \frac{1}{M} \sum_{\mathbf{A}_c \in \omega_k} (\mathbf{A}_c - \bar{\mathbf{A}}_k)^T (\mathbf{A}_c - \bar{\mathbf{A}}_k) \quad (4.40)$$

โดย $\bar{\mathbf{A}}_k$ เป็นเมตริกซ์เฉลี่ยของตัวอย่างภาพในกลุ่มฝึกฝนที่ k โดยปริภูมีย่อยหน้าเจาะจงที่ k สามารถแทนได้ 3 ทูเปิลดังนี้

$$\mathfrak{R}_k^{2DPCA} = \{\mathbf{X}_k, \bar{\mathbf{A}}_k, d_k\} \quad (4.41)$$

ในทำนองเดียวกัน กรณี B2DPCA เมตริกซ์ความแปรปรวนร่วมเชิงภาพตามแนวหลักเพื่อคำนวณเมตริกซ์การฉายของกลุ่มที่ k \mathbf{Z}_k ได้ดังนี้

$$\mathbf{H}_k = \frac{1}{M} \sum_{\mathbf{A}_c \in \omega_k} (\mathbf{A}_c - \bar{\mathbf{A}}_k)(\mathbf{A}_c - \bar{\mathbf{A}}_k)^T \quad (4.42)$$

ดังนั้นจำนวนทูเปิลของ B2DPCA จึงเพิ่มขึ้นเป็น 5 ทูเปิลดังนี้

$$\mathfrak{R}_k^{B2DPCA} = \{\mathbf{Z}_k, \mathbf{X}_k, \bar{\mathbf{A}}_k, d_k, q_k\} \quad (4.43)$$

กำหนดให้ \mathbf{A}_i เป็นตัวอย่างป้อนเข้า เราสามารถหาภาพฉายของ \mathbf{A}_i บนปริภูมิย่อยหน้าเจาะจงของกลุ่มต่าง ๆ ทั้ง k ปริภูมิได้ดังนี้

$$\mathbf{U}_k = \begin{cases} \mathbf{W}_k \mathbf{X}_k; 2DPCA \\ \mathbf{Z}_k^T \mathbf{W}_k \mathbf{X}_k; B2DPCA \end{cases} \quad (4.44)$$

โดย \mathbf{U}_k คือภาพฉายของ \mathbf{A}_i บนปริภูมิย่อยที่ k และ $\mathbf{W}_k = (\mathbf{A}_i - \bar{\mathbf{A}}_k)$ เนื่องจากการวัดระยะห่างของภาพฉายที่อยู่ต่างปริภูมิกันทำไม่สามารรถได้เราจึงต้องทำการสร้างคืนภาพกลับมาที่ปริภูมิเดิมดังนี้

$$\mathbf{W}_k^r = \begin{cases} \mathbf{U}_k \mathbf{X}_k^T; 2DPCA \\ \mathbf{Z}_k \mathbf{U}_k \mathbf{X}_k^T; B2DPCA \end{cases} \quad (4.45)$$

โดย \mathbf{W}_k^r คือภาพที่ได้จากการสร้างคืนของ \mathbf{U}_k ดังนั้นเราสามารถหาค่าความผิดพลาดในการสร้างคืนจากปริภูมิย่อยใด ๆ ได้ดังนี้

$$\varepsilon_k(\mathbf{W}_k^r, \mathbf{A}_i) = \sum_{m=1}^{m_{row}} \sum_{n=1}^{n_{col}} |\mathbf{w}_{(m,n)_k}^r - \mathbf{a}_{(m,n)_i}^r| \quad (4.46)$$

โดย ε_k คือ ค่าความผิดพลาดในการสร้างคืนจากปริภูมิย่อย k เราสามารถกล่าวได้ว่าค่าความผิดพลาดจากปริภูมิย่อยใด ให้ค่าความผิดพลาดน้อยที่สุด

$$\varepsilon_i = \min_{1 \leq k \leq K} (\varepsilon_k) \quad (4.47)$$

ตัวอย่างทดสอบ \mathbf{A}_i ดังกล่าวจะเป็นกลุ่มที่ i เราจึงสามารถใช้ค่าความผิดพลาดนี้เป็นบรรทัดฐานในการรู้จำได้ อย่างไรก็ตาม ความซับซ้อนในการคำนวณจะมากขึ้น แต่ในขณะเดียวกันเราไม่จำเป็นต้องใช้หน่วยความจำที่ใช้เก็บเมตริกซ์คุณลักษณะของตัวอย่างฝึกฝนอีกต่อไป และจำนวนครั้งในการคำนวณระยะห่างจากเดิมเท่ากับจำนวนตัวอย่างที่ใช้ฝึกฝนทั้งหมด M ครั้ง เหลือเพียงเท่ากับจำนวนกลุ่ม K ครั้ง โดยสรุปเป็นตารางเปรียบเทียบได้ดังตารางที่ 1

ตารางที่ 1 การเปรียบเทียบหน่วยความจำที่ใช้ และความซับซ้อนในการคำนวณระหว่างวิธีการ
2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS

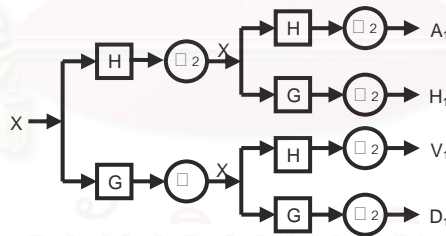
วิธีการ	หน่วยความจำที่ใช้		จำนวนการฉาย		จำนวนการคำนวณ ระยะห่าง
	เมตริกซ์ ฉาย	เมตริกซ์ คุณลักษณะ	ฝึกฝน	ทดสอบ	
2DPCA	nd	mdM	M	1	M
B2DPCA	$mq+nd$	dqM	$2M$	2	M
2DPCA+FSS	ndK	0	M	$2K$	K
B2DPCA+FSS	$K(nd + mq)$	0	$2M$	$4K$	K

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ความละเอียดหลายระดับ (Multiresolution Analysis)

วิธีความละเอียดหลายระดับนั้น เป็นวิธีการที่กำเนิดเครื่องมือในการวิเคราะห์สัญญาณที่มีความสามารถสูง ซึ่งนิยมใช้อย่างกว้างขวางในการสกัดคุณลักษณะ การบีบอัด และการลดสัญญาณรบกวน การแยกเวฟเลต (Wavelet Decomposition) ก็เป็นวิธีหนึ่งที่นิยมใช้ในเทคนิคความละเอียดหลายระดับทางการประมวลผลสัญญาณภาพ ซึ่งภาพนั้นจะถูกแปลงเป็นค่าทางสถิติ จากผลการผสมผสานที่ต่างกันของคุณลักษณะที่เปลี่ยนแปลงในทันทีทันใด ตัวอย่างเช่น ขอบ, บริเวณของตัวหนังสือ หรือ บริเวณที่มีความชัดเจนต่ำ โดยในที่นี่ได้เลือกให้การแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform หรือ DWT) เนื่องจากในการแปลงเวฟเลตแบบต่อเนื่อง (Continuous Wavelet Transform) นั้นจะให้ค่าข้อมูลซ้ำซ้อนมาก และต้องใช้การคำนวณที่ซับซ้อนมากกว่า

การแปลงเวฟเลตแบบ 2 มิติมีแนวคิดจากการประยุกต์อย่างต่อเนื่องของการแปลงเวฟเลตแบบ 1 มิติ ดังรูปที่ 2 ซึ่ง G คือ ตัวกรองความถี่สูงผ่าน (High-pass Filtering) และ H คือ ตัวกรองความถี่ต่ำผ่าน (Low-pass Filtering) โดย $\downarrow 2$ แสดงถึงการซัดตัวอย่างต่ำลง (Down sampling) 2 และในส่วนสุดท้าย จะได้ภาพแถบย่อยที่มีความละเอียด ขนาด $N/2 \times N/2$ ซึ่ง A_1 จะเป็นส่วนที่เก็บรายละเอียดความถี่ต่ำส่วนใหญ่ H_1 จะเป็นรายละเอียดทางด้านแนวนอน V_1 จะเป็นรายละเอียดทางด้านแนวตั้ง และสุดท้าย D_1 จะเป็นรายละเอียดทางด้านแนวทแยง



รูปที่ 4 แผนผังต้นไม้แทน 1 ระดับ ของ การแยกเวฟเลต 2 มิติ

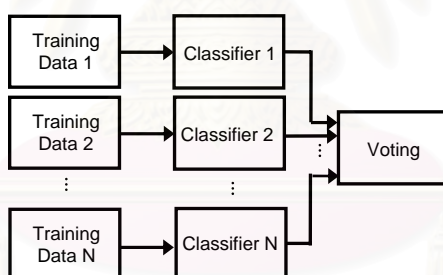
ในการจำแนกประเภทนั้น มีการตั้งสมมติฐานว่าความแตกต่างภายในกลุ่มจะอยู่ในส่วนความถี่สูง และลักษณะที่พอเพียงในการจำแนกนั้นจะอยู่ในส่วนความถี่ต่ำ [20]

วิธีชุดข้อมูลทั้งหมด (Ensemble Methods)

“ในการแก้ปัญหาที่ยากนั้น เราอาจจะต้องใช้การแทนที่ที่แตกต่างกัน เพื่อหยุดการโต้แย้งว่าเทคนิคใดนั้นดีที่สุดในการจำแนกแบบรูป... แทนที่เราจะทำงานที่ระดับการจัดการที่สูงขึ้น แต่หากสนใจที่จะค้นหาว่าจะสร้างระบบการจัดการอย่างไรเพื่อใช้ประโยชน์จากหลาย ๆ คุณสมบัติ และหลีกเลี่ยงข้อจำกัดที่แตกต่างในการเปรียบเทียบแต่ละสิ่งได้” [Minsky, 1991] ซึ่งอาจกล่าวได้ว่าชุดของจำแนก (Ensemble of Classifiers) คือ เซ็ตของตัวจำแนกที่มีการตัดสินใจแยกกัน โดยแต่ละตัวจำแนกเป็นอิสระต่อกัน และนำมาสรุปรวมกันโดยวิธีลงคะแนนเสียง ในชุดของตัวจำแนกนั้นแบ่งเป็น 2 ชนิด คือ

1. ชุดข้อมูลที่มีตัวจำแนกเหมือนกัน (Homogeneous Classifiers)
2. ชุดข้อมูลที่มีตัวจำแนกต่างกัน (Heterogeneous Classifiers)

การแบ่งชุดข้อมูลมีลักษณะเป็นดังรูปที่ 2 ซึ่งวิธีชุดข้อมูลทั้งสองประเภทนั้น ได้พิจารณาจากชุดข้อมูลการฝึกสอนที่ไม่เหมือนกัน ซึ่งแตกต่างกันในส่วนของชนิดตัวจำแนก สำหรับวิธีชุดข้อมูลที่มีตัวจำแนกเหมือนกันนั้น จะใช้ตัวจำแนกแต่ละตัวที่เหมือนกัน และให้ผลจากการจำแนกที่เป็นอิสระต่อกัน แต่ทว่าในกรณีของชุดข้อมูลที่มีตัวจำแนกต่างกัมนั้น จะใช้ตัวจำแนกที่แตกต่างกัน [21]



รูปที่ 5 ชุดข้อมูลตัวจำแนก

การสร้างชุดข้อมูลมีวิธีต่าง ๆ ดังนี้

1 การจัดการตัวอย่างที่ใช้ในการฝึกสอน (Manipulating the Training Examples)

การจัดการชุดข้อมูลวิธีนี้ทำขึ้นเพื่อสร้างสมมติฐานหลาย ๆ แบบ ซึ่งขั้นตอนวิธีการเรียนรู้จำเป็นต้องดำเนินการหลายครั้ง โดยในแต่ละครั้งที่ดำเนินการ จะมีการสุ่มเลือกเซตย่อย (Subset) ของตัวอย่างที่ใช้ในการฝึกสอนต่างกัน โดยเทคนิคนี้จะเป็นดีสำหรับกรณีของขั้นตอนวิธีการเรียนรู้ที่ไม่เสถียร (Unstable Learning Algorithm) ซึ่งคือ ขั้นตอนวิธีที่ตัวจำแนกข้อมูลขาออกมีการเปลี่ยนแปลงมาก เมื่อมีข้อมูลที่ใช้ฝึกสอนเปลี่ยนแปลงเพียงเล็กน้อย วิธีการจัดการตัวอย่างที่ใช้ในการฝึกสอนเบื้องต้นซึ่งเป็นที่รู้จักกันดีนั้นคือ การแบกกิง (Bagging), การบูสติง (Boosting)

2 การจัดการคุณสมบัติของข้อมูลขาเข้า (Manipulating the Input Features)

เทคนิคนี้จะเป็นการสร้างตัวจำแนกหลาย ๆ ตัว (Multiple Classifiers) จากการแทนคุณสมบัติแบบต่าง ๆ ของข้อมูลที่ใช้ฝึกสอนเริ่มแรก (โดยในงานวิจัยนี้ได้เลือกใช้การแปลงเวฟเล็ตในการแทนคุณสมบัติของข้อมูลฝึกสอนเริ่มแรก) เพื่อให้ได้ข้อมูลที่ใช้ฝึกสอนชุดใหม่ และทำการจัดข้อมูลที่ได้จากการแทนคุณสมบัติดังกล่าวเป็นเซตต่าง ๆ โดยนำข้อมูลที่จัดแล้วนั้นไปใช้ต่อไปในขั้นตอนการเรียนรู้ จะสังเกตได้ว่าเทคนิคนี้เหมาะกับงานที่คุณสมบัติของข้อมูลขาเข้ามีความซ้ำซาก (Redundant) สูง

3 การจัดการที่ข้อมูลออก (Manipulating the Output Targets)

เทคนิคนี้จะใช้ในการสร้างชุดข้อมูลที่ใช้ในการจำแนกที่เหมาะสม เพื่อจัดการกับข้อมูลขาออก โดยเทคนิคนี้อธิบายได้ว่าเป็น การเข้ารหัสขาออกที่ตรวจจับความผิดพลาด (Error-correcting Output Coding) สมมติว่ามีจำนวน K คลาสที่มีขนาดใหญ่ จากนั้นปัญหาการเรียนรู้จะสร้างได้โดยสุ่มส่วนต่าง ๆ ใน K คลาสให้เป็น 2 เซตย่อย A_i และ B_i ในข้อมูลขาเข้าจะทำการตั้งชื่อใหม่ (Re-label) โดยให้ในกลุ่ม A_i นั้นจะเปลี่ยนให้เป็น 0 และกลุ่ม B_i จะเปลี่ยนให้เป็น 1 ข้อมูลที่ตั้งชื่อใหม่นี้จะใช้ในขั้นตอนวิธีการเรียนรู้ ซึ่งสร้างตัวจำแนก h_i จากการทำงานซ้ำจำนวน L ครั้ง จะให้เซตย่อยที่แตกต่างกันใน A_i และ B_i ทำให้ได้ชุดข้อมูล L ตัวจำแนก คือ h_1, \dots, h_L

เมื่อได้ชุดข้อมูลเข้า x ใหม่แล้ว จะทำการจำแนกได้โดยนำตัวจำแนก h_i แต่ละตัวมาทำการจำแนก x ซึ่งถ้า $h_i(x) = 0$ จะลงคะแนนเสียงในแต่ละคลาสของ A_i และถ้า $h_i(x) = 1$ จะลงคะแนนเสียงให้แต่ละคลาสของ B_i หลังจากทีแต่ละตัวจำแนก L ได้ลงคะแนนเสียงแล้ว จะเลือกคลาสที่มีคะแนนเสียงมากที่สุดเป็นผลจากการทำนายที่ได้ของชุดข้อมูล

4 การใส่ข้อมูลแบบสุ่ม (Injecting Randomness)

เทคนิคนี้จะสร้างชุดข้อมูลตัวจำแนก โดยการใส่ข้อมูลแบบสุ่มลงในขั้นตอนวิธีการเรียนรู้ และไม่มีการเปลี่ยนแปลงข้อมูลที่ใช้ฝึกสอน เทคนิคนี้มีจุดประสงค์เพื่อให้ได้ผลการจำแนกที่แตกต่างกันเพิ่มขึ้นกว่ากรณีที่ไม่ได้ใส่ข้อมูลแบบสุ่มลงในขั้นตอนวิธีการเรียนรู้ ยกตัวอย่างเช่น ในข่ายงานระบบประสาท (Neural Network) ที่มีการเลือกใช้กระบวนการคิดย้อนกลับ (Backpropagation Algorithm) พบว่าในการกำหนดค่าเริ่มต้นของตัวถ่วงน้ำหนัก (Weight) ของข่ายงานนั้นเป็นแบบสุ่ม ซึ่งถ้ามีข้อมูลที่ใช้ในการฝึกสอนเดียวกัน แต่จะมีการกำหนดค่าตัวถ่วงน้ำหนักที่ต่างกัน สามารถให้ผลการจำแนกที่แตกต่างกันได้

ในงานวิจัยนี้ได้เลือกใช้ชุดข้อมูลตัวจำแนกแบบชุดข้อมูลที่มีตัวจำแนกเหมือนกัน ประกอบกับการจัดการคุณสมบัติของข้อมูลขาเข้า และวิธีการจัดการตัวอย่างที่ใช้ในการฝึกสอนนั้นก็ก็เป็นทางเลือก

หนึ่งซึ่งใช้ในการคัดเลือกคุณลักษณะ โดยแนวคิดดังกล่าวนี้เป็นวิธีการในแต่ละส่วนสำหรับ
ออกแบบการจัดการข้อมูลขาเข้าที่จะก่อนนำไปสกัดคุณลักษณะโดยวิธีวิเคราะห์ปริมาณน้อย เพื่อ
ปรับปรุงให้ได้อัตราการเรียนรู้ที่ดีขึ้น



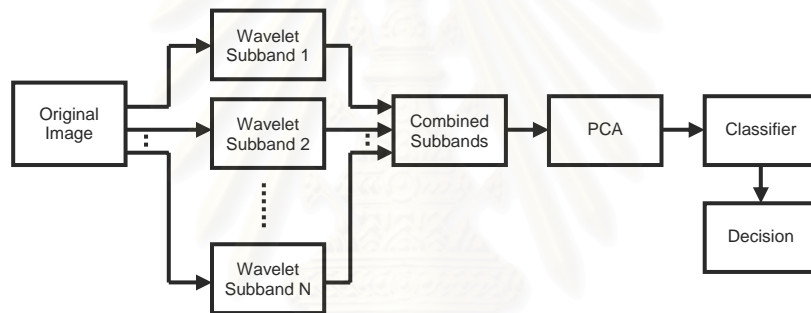
สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

การรวมแถบย่อยแบบต่าง ๆ (Fusion Subbands)

เนื่องจากได้มีการนำแนวคิดของวิธีความละเอียดหลายระดับมาประยุกต์ใช้กับวิธี PCA โดยนำเอากลุ่มแถบย่อยย่อยต่าง ๆ มารวมกัน [6] โดยมีวิธีการรวมอยู่ 3 แบบ คือ

1 การรวมแบบใช้ข้อมูล (Data Fusion)

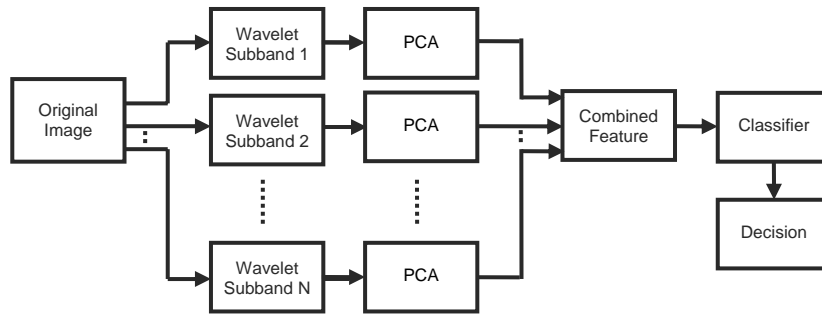
การรวมแบบใช้ข้อมูล กระทำโดยรวมข้อมูล ซึ่งคือแถบย่อยต่าง ๆ ที่ได้จากการแปลงเวฟเล็ตแบบไม่ต่อเนื่องโดยตรง หลังจากนั้นนำแถบย่อยย่อยมารวมกัน ทำให้ได้เป็นเวกเตอร์ใหม่ที่เกิดจากหลายแถบย่อยมาต่อกัน จากนั้นนำไปสกัดคุณลักษณะโดยการวิเคราะห์ปริภูมิย่อย ซึ่งในโครงร่างนี้เลือกใช้วิธี PCA ลำดับต่อไปนั้นได้ทำการพิจารณาระยะห่างโดยใช้ระยะทางแบบต่าง ๆ และสุดท้ายทำการตัดสินใจว่ารู้จำได้ว่าเป็นบุคคลใด ดังรูป



รูปที่ 6 แผนภาพบล็อกของการรวมแบบใช้ข้อมูล

2 การรวมแบบใช้คุณลักษณะ (Feature Fusion)

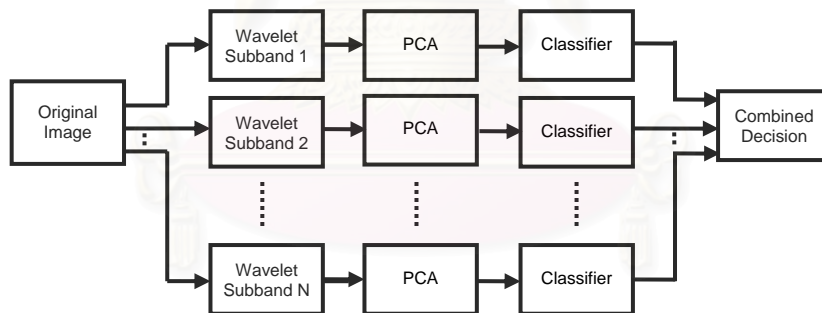
การรวมแบบใช้คุณลักษณะนั้น พิจารณาจากคุณลักษณะที่ได้จากการวิเคราะห์ปริภูมิย่อยของแต่ละแถบย่อยแยกจากกัน หลังจากนั้นนำคุณลักษณะที่ได้มารวมกัน และทำการจำแนกโดยใช้ตัวจำแนกเดียวกันกับข้อ 1 และทำการตัดสินใจว่าเป็นบุคคลใดเป็นขั้นตอนสุดท้าย ซึ่งการรวมวิธีนี้ต่างกับการรวมแบบใช้ข้อมูลที่มีการวิเคราะห์ปริภูมิย่อยแยกกัน หลังจากนั้นขั้นตอนการทำงานเช่นเดียวกันโดยการรวมแบบใช้คุณลักษณะได้แสดงในรูป



รูปที่ 7 แผนภาพบล็อกของการรวมแบบใช้คุณลักษณะ

3 การรวมแบบใช้ผลการตัดสินใจ (Decision Fusion)

การรวมแบบใช้ผลการตัดสินใจจะทำการพิจารณาทุกแถบย่อยแยกกันในทุก ๆ ขั้นตอน ซึ่งส่งผลให้มีผลการตัดสินใจออกมาหลายแบบ ดังนั้นในขั้นตอนสุดท้ายต้องมีการสรุปผล ดังนั้นจึงได้ทำการรวมโดยวิธีเสียงส่วนมาก โดยจะสรุปผลรวมจากคะแนนเสียงที่มีการลงคะแนนข้างมาก ซึ่งเป็นวิธีที่ง่าย และเหมาะสมในกรณีนี้ จากนั้นสรุปว่าเป็นบุคคลใด ดังแสดงในรูปซึ่งวิธีนี้เป็นหนึ่งในระบบหลายตัวจำแนกซึ่งไม่มีการถ่วงน้ำหนักการตัดสินใจนั่นเอง

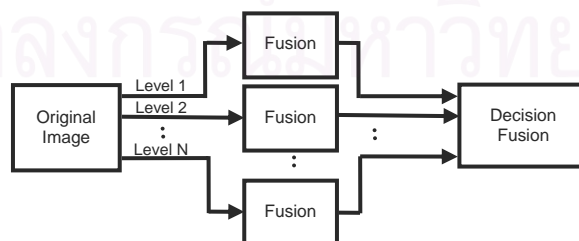


รูปที่ 8 แผนภาพบล็อกของการรวมแบบใช้ผลการตัดสินใจ

แนวคิดที่นำเสนอสำหรับการใช้วิธีความละเอียดหลายระดับ

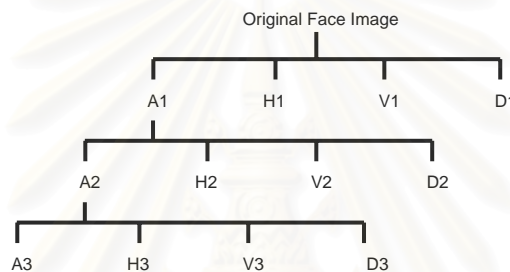
นำเสนอการใช้ระบบหลายตัวจำแนกสองระดับในการสร้างชุดโครงสร้างการรวมแถบย่อย ซึ่งสามารถพิจารณาได้ว่าเป็นการรวมชุดโครงสร้างการรวมแถบย่อยแต่ละโครงสร้างเข้าด้วยกัน โดยมีการพิจารณาว่าเป็นการรวมในสองระดับ ดังแสดงในรูป โดยแต่ละขั้นตอนนั้นจะเป็นการรวมแถบย่อยเฉพาะแบบ ซึ่งกล่าวได้ว่าเป็นการรวมแถบย่อยเพียงครั้งเดียว ในแต่ละโครงสร้างการรวมแถบย่อยนั้นจะเป็นอิสระต่อโครงสร้างการรวมแถบย่อยอื่น ๆ ซึ่งข้อมูลขาเข้าของแต่ละโครงสร้างการรวมแถบย่อยในงานวิจัยนี้จะถูกแปลงข้อมูล (Data Transform) จากเวกเตอร์ข้อมูลขาเริ่มต้นก่อน ซึ่งการแปลงข้อมูลในงานวิจัยนี้เลือกใช้การแยกเวฟเล็ตแบบไม่ต่อเนื่อง ดังนั้นในแต่ละขั้นตอนจะถูกฝึกสอนจากข้อมูลขาเข้าเดียวกันแต่มีการแทนสัญญาณที่แตกต่างกัน จากนั้นแต่ละโครงสร้างการรวมแถบย่อยจะมีการฝึกสอน, ทดสอบ และจำแนกของตัวโครงสร้างนั้น ๆ เอง เมื่อเสร็จสิ้นการทดสอบ และได้ผลการจำแนกของแต่ละโครงสร้างแล้ว จะนำผลการจำแนกของแต่ละโครงสร้างมาหาผลสรุปร่วมกัน ซึ่งในแต่ละคลาสนั้นจะมีการถ่วงน้ำหนักที่เท่ากัน เนื่องจากฐานข้อมูลใบหน้าที่ในการทดลองนั้นมีขนาดเล็กเกินไปที่จะมีการถ่วงน้ำหนักได้

สำหรับการแทนข้อมูลขาเข้านั้น ในงานวิจัยนี้ได้ประยุกต์ใช้การจัดการคุณสมบัติของข้อมูลขาเข้าซึ่งเป็นหนึ่งในวิธีชุดข้อมูลทั้งหมด โดยการใช้การแปลงเวฟเล็ตแบบไม่ต่อเนื่อง ที่เป็นการแปลงสัญญาณจากโดเมนเวลาไปยังโดเมนความถี่ จากการคำนวณเวฟเล็ตในหลายระดับ ทำให้ได้ความละเอียดทางเวลา และความถี่อย่างหลากหลาย ซึ่งจากข้อมูลต้นแบบเดียวกันจะมีแทนที่แตกต่างกัน โดยมีความละเอียดทางเวลา และความถี่ต่างกันนั่นเอง โดยมีขั้นตอนการทำงานเป็นดังรูป สำหรับโครงสร้างการรวมแถบย่อยที่ใช้ในในแต่ละชุดโครงสร้างนั้นเป็นโครงสร้างแถบย่อยประเภทเดียวกันทั้งหมด โดยเลือก 1 ใน 3 แบบนั่นเอง ซึ่งคือ การรวมโดยใช้ข้อมูล, การรวมโดยใช้ลักษณะเด่น และการรวมโดยใช้ผลการตัดสินใจ



รูปที่ 9 แผนภาพบล็อกแนวคิดที่นำเสนอ

ในงานวิจัยแบบต้องถูกกำหนดให้มีขนาด 128×128 จุดภาพก่อนการนำไปแปลงข้อมูลซึ่งใช้แยกเวฟเล็ตแบบไม่ต่อเนื่องดังที่กล่าวมาข้างต้น ซึ่งมีจุดประสงค์ของการกำหนดขนาดภาพต้นแบบเพื่อจะได้รูปที่มีขนาดที่เท่ากับขนาดต้นแบบโดยไม่มีการสูญเสียใด ๆ เมื่อทำการคืนกลับทางเวฟเล็ตซึ่งแผนภาพต้นไม้ของการแยกเวฟเล็ต 2 มิติที่ใช้ในงานวิจัยนี้แสดงอยู่ในรูปที่ 7 โดยในระดับแรกจะได้แถบย่อยซึ่งคือ A1, H1, V1 และ D1 จะมีความละเอียดที่ 64×64 จุดภาพ จากนั้นในระดับที่สองจะทำการแยกเวฟเล็ตเฉพาะกับแถบย่อย A1 เท่านั้น เนื่องจากในแถบย่อยอื่น ๆ มีประสิทธิภาพในการจำแนกต่ำ ซึ่งแถบย่อยที่ได้ในระดับที่สอง คือ A2, H2, V2 และ D2 จะมีขนาด 32×32 จุดภาพ และในระดับที่สามจะทำการแยกเวฟเล็ตเฉพาะกับแถบย่อย A2 เท่านั้นจากเหตุผลเดียวกันกับในระดับที่สอง โดยได้แถบย่อยที่อยู่ในระดับสุดท้าย คือ A3, H3, V3 และ D3 จะมีขนาดเป็น 16×16 จุดภาพ



รูปที่ 10 แผนภาพต้นไม้ของการแยกเวฟเล็ตที่ใช้ในโครงการวิจัย

กลุ่มแถบย่อยที่ทำงานวิจัยได้ทำการเลือกแบบต้นไม้ออกเทพ (Octave Tree) ซึ่งทำได้ 3 ระดับคือ

ระดับที่ 1 เลือก A1 H1 V1 D1 ซึ่งเป็นแถบย่อยที่ระดับที่ 1 ทั้งหมด

ระดับที่ 2 เลือก H1 V1 D1 A2 H2 V2 D2 ซึ่งเป็นแถบย่อยที่เกิดขึ้นระดับที่ 1 ยกเว้น A1 และแถบย่อยที่เกิดจาก A1 ในระดับที่ 2

ระดับที่ 3 เลือก H1 V1 D1 H2 V2 D2 A3 H3 V3 D3 ซึ่งเป็นแถบย่อยที่เกิดขึ้นระดับที่ 1 ยกเว้น A1 และ แถบย่อยที่เกิดจาก A1 ในระดับที่ 2 ยกเว้น A2 และ แถบย่อยในระดับที่ 3 ทั้งหมด

ฟังก์ชันฐานหลัก (Basis Function) ของเวฟเล็ตในงานวิจัยนี้ได้เลือกใช้เวฟเล็ตดาบาชี่ 4 (Daubechies 4 Wavelets) ในขั้นตอนของการรู้จำนั้นสามารถทำได้โดยนำตัวอย่างใบหน้าที่ใช้ทดสอบซึ่งได้ทำการฉายแล้ว ไปหาระยะห่างกับกลุ่มตัวอย่างใบหน้าที่ใช้ฝึกสอนซึ่งได้ทำการฉายแล้วทุก ๆ ภาพ ถ้าระยะห่างระหว่างตัวอย่างใบหน้าที่ใช้ทดสอบนั้น ๆ กับตัวอย่างใบหน้าที่ใช้ฝึกสอนใดมีค่าน้อยที่สุด ก็จะตัดสินใจว่าเป็นใบหน้าของบุคคลที่ใช้ฝึกสอนคนนั้น

Image Cross-Covariance Analysis (ICCA)

เมื่อสังเกตความสัมพันธ์ระหว่าง Image covariance matrix ของ 2DPCA ในสมการ (4.16) และ Covariance matrix ของ PCA ในสมการ (4.4) พบว่า

$$\mathbf{G}(i, j) = \sum_{k=1}^m \mathbf{C}(m(i-1)+k, m(j-1)+k) \quad (4.48)$$

โดย $\mathbf{G}(i, j)$ และ $\mathbf{C}(i, j)$ คือ element ที่แถว i^{th} และหลัก j^{th} ของ \mathbf{G} และ \mathbf{C} ตามลำดับ และ m คือความสูงของรูป

เพื่อให้เห็นภาพกำหนดให้ภาพที่ใช้ฝึกฝนทั้งหมดมีขนาด 3 คูณ 3 ดังนั้น Covariance matrix ของภาพเหล่านี้จึงมีขนาด 9 คูณ 9 และขนาดของ Image covariance matrix มีขนาดเพียง 3 คูณ 3 ดังแสดงในรูปที่ 11 พบว่าแต่ละ element ของ \mathbf{G} เกิดจากการบวกกันของ element ทุก element บน \mathbf{C} ที่มีป้ายชื่อเดียวกัน



รูปที่ 11 ความสัมพันธ์ระหว่าง \mathbf{G} และ \mathbf{C}

จากที่กล่าวมาเราสามารถสรุปได้ว่า Image covariance matrix นั้นรวบรวมข้อมูลเพียง $1/m$ ของข้อมูลทั้งหมดใน Covariance matrix เดิม เราพบว่าข้อมูลที่ถูกลบทิ้งไปนั้นไม่ใช่ข้อมูลที่เปล่าประโยชน์ดังจะแสดงในเห็นในหัวข้อผลการทดลองต่อไป

เพื่อจะพิจารณาข้อมูลที่ขาดหายไปนี้เรากำหนดวิธีขึ้นใหม่เรียกว่า Image Cross-Covariance Analysis (ICCA) โดยการนิยาม Image Cross-Covariance matrix ขึ้นมาใหม่ดังนี้

$$\mathbf{G}_L = E \left[(\mathbf{B}_L - E[\mathbf{B}_L])^T (\mathbf{A} - E[\mathbf{A}]) \right] \quad (4.49)$$

เมื่อ \mathbf{B}_L คือภาพที่เลื่อนไปจำนวน L ครั้งโดยใช้ระเบียบวิธีดังกล่าวอย่างฟังก์ชัน MATLAB ดังต่อไปนี้

```
function B=Shifting(A,L)
if(L==1), B=A; return;
index = [2:size(A,2),1];
for i=1:L-1
B=A(2:size(A,1),:);
B=[B;A(1,index)];
A=B;
end
```

โดยตัวอย่างภาพที่ได้การวิธีดังกล่าวแสดงในรูปที่ 12

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 12 ตัวอย่างภาพที่เกิดจากการเลื่อนบนฐานข้อมูล ORL

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการทดลอง

ฐานข้อมูลที่ใช้ในการทดลองประกอบด้วย ฐานข้อมูลรู้ใบหน้าจำนวน 3 ฐานข้อมูล ได้แก่ ฐานข้อมูล Yale AR และ ORL และฐานข้อมูล MSTAR

- ฐานข้อมูล Yale

ฐานข้อมูล Yale [15] ประกอบด้วยภาพสกุล Graphic Interchange Format (GIF) ซึ่งเป็นการบีบอัดภาพแบบไม่มีการสูญเสียข้อมูลแบบ 8 บิต จำนวน 165 ภาพ โดยแบ่งออกเป็น 15 กลุ่ม แต่ละกลุ่มประกอบด้วยภาพจากบุคคลเดียวกัน 11 ภาพ แต่มีการแสดงอารมณ์รวมทั้งทิศทางที่แสงตกกระทบต่าง ๆ กัน ดังนี้ แสงตกกระทบกลางใบหน้า (center-light) สวมแว่นตา (with glasses) อารมณ์สุข (happy) แสงตกกระทบทางซ้ายของภาพใบหน้า (left-light) ไม่สวมแว่น (without glasses) ปกติ (normal) แสงตกกระทบทางขวาของภาพใบหน้า (right-light) อารมณ์เศร้า (sad) หลับตา (sleepy) ตกใจ (surprised) และ ขยิบตา (wink) ดังแสดงในรูปที่ 13 แต่ละภาพถูกตัดบริเวณใบหน้าด้วยมือและกำหนดขนาดใหม่เป็น 100 x 80 ตารางพิกเซล



รูปที่ 13 ตัวอย่างภาพใบหน้าจากฐานข้อมูล Yale

- ฐานข้อมูล AR

ฐานข้อมูล AR ถูกสร้างขึ้นโดย Aleix Martinez และ Robert Benavente ภายใน Computer Vision Center (CVC) ที่ U.A.B. ประกอบด้วยภาพสีจำนวนมากกว่า 4,000 ภาพจาก 126 คน ซึ่งเป็นผู้ชาย 70 คน และผู้หญิง 56 คน ภาพทั้งหมดเป็นภาพใบหน้าตรงที่แสดงอารมณ์ ทิศทางที่แสงตกกระทบ และ วัตถุบังใบหน้า (แว่นตากันแดด หรือผ้าผืนคอ) แตกต่างกัน โดยการเก็บข้อมูลของแต่ละบุคคลจะแบ่งเป็น 2 ชุด โดยแต่ละชุดมีระยะเวลาห่างกัน 2 สัปดาห์ ในการทดลองของ

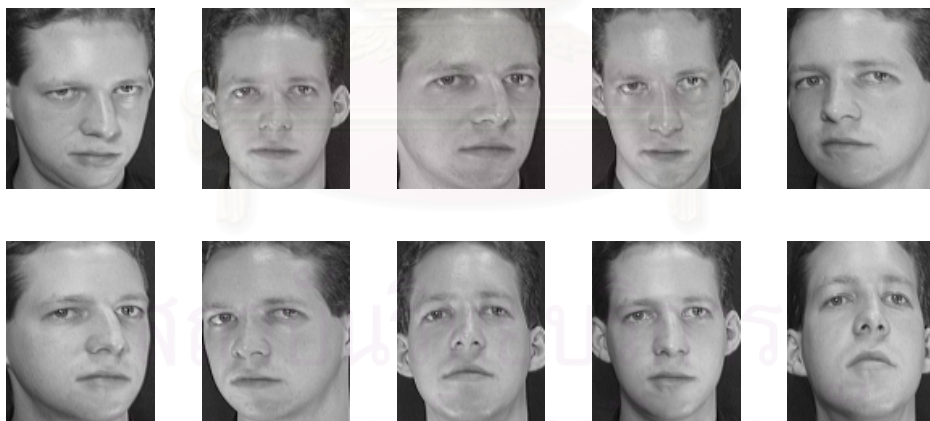
เราจะพิจารณาเพียงกลุ่มละ 14 ภาพโดยไม่นับรวมภาพที่มีวัตถุบดบังใบหน้า จากนั้นตัด และ กำหนดขนาดใหม่เป็น 100 x 80 จากนั้นเปลี่ยนเป็นภาพสีเทา 256 ระดับดังแสดงในรูปที่ 14



รูปที่ 14 ตัวอย่างภาพใบหน้าจากฐานข้อมูล AR

- ฐานข้อมูล ORL

ฐานข้อมูล ORL ประกอบด้วยภาพใบหน้าจาก 40 คนโดยแต่ละคนเก็บข้อมูลไว้เป็นจำนวน 10 ภาพในมุมมองที่ต่างกัน 20 องศา รวมถึงการแสดงอารมณ์ การลืมตาและหลับตาที่ต่าง ๆ กัน และ บางภาพของบางคนเก็บข้อมูลที่เวลาต่างกัน ภาพทั้งหมดมีขนาด 112 x 92 ดังแสดงในรูปที่ 15



รูปที่ 15 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL

- ฐานข้อมูล MSTAR

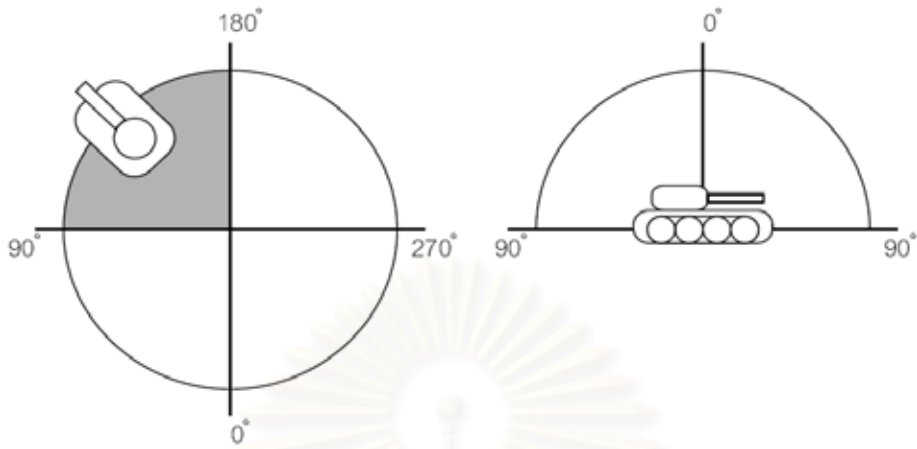
ฐานข้อมูล MSTAR ประกอบด้วยชุดข้อมูล Synthetic Aperture Radar (SAR) ความละเอียดสูงรวบรวมโดยห้องวิจัยโปรแกรม DARPA/Wright laboratory Moving and Stationary Target Acquisition and Recognition (MSTAR) ชุดข้อมูลประกอบด้วยภาพ SAR ขนาด 128 x 128 ของ

รถทางทหาร 3 แบบ ได้แก่ BMP2 armored personal carriers (APCs), BTR70 APCs และรถถัง T72 ซึ่งรูป SAR ดังกล่าวขึ้นกับการวางตัวของรถเป้าหมายที่ต่างกัน ดังนั้นการตรวจรู้ทิศทาง การวางตัวของรถก่อน (มุมมอง Aspect) ดังรูปที่ 16 สามารถเพิ่มประสิทธิภาพการรู้จำได้มากขึ้นซึ่งในการทดลองของเราจะทำด้วยมือ ตัวอย่างรูปภาพจากฐานข้อมูล MSTAR แสดงดังรูปที่ 17 โดยภาพทั้งหมดถูกตัดเอาเฉพาะบริเวณกลางภาพขนาด 80 x 80 ตารางพิกเซล

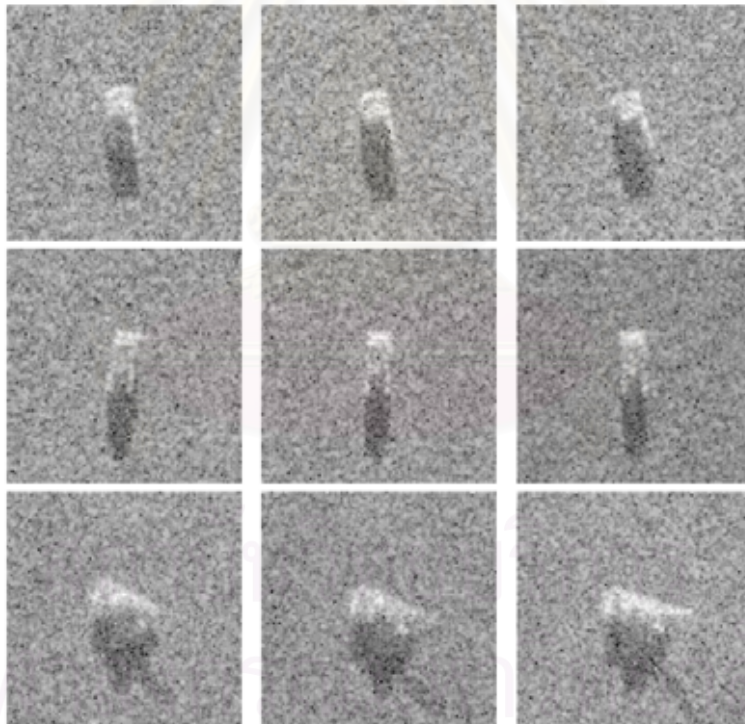


สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 แสดงชุดฝึกฝน และตารางที่ 3 แสดงชุดทดสอบ



รูปที่ 16 ภาพซ้ายคือมุม Aspect และภาพขวาคือมุม Depression



รูปที่ 17 ตัวอย่างภาพ SAR จากฐานข้อมูล MSTAR โดยแถวบนเป็นรถ BMP2 armored personal carriers (APCs) แถวกลางเป็นรถ BTR70 APCs และแถวล่างเป็นรถถัง T72

ตารางที่ 2 ชุดฝึกฝนของฐานข้อมูล MSTAR

ชนิดของรถ	หมายเลขรถ	หมายเลขซีเรียล	มุม Depression (องศา)	จำนวนภาพ
BMP2	1	9563	17	233
	2	9566		231
	3	C21		233
BTR70	1	C71	17	233
T72	1	132	17	232
	2	812		231
	3	S7		228
รวม	-	-	17	1,621

ตารางที่ 3 ชุดทดสอบของฐานข้อมูล MSTAR

ชนิดของรถ	หมายเลขรถ	หมายเลขซีเรียล	มุม Depression (องศา)	จำนวนภาพ
BMP2	1	9563	15	195
	2	9566		196
	3	C21		196
BTR70	1	C71	15	196
T72	1	132	15	196
	2	812		195
	3	S7		191
รวม	-	-	15	1,365

การทดลองที่ 1 ประสิทธิภาพของระบบรู้จำใบหน้าด้วยวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA

ตารางที่ 4 และตารางที่ 5 แสดงการเปรียบเทียบวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA ณ จุดที่มีประสิทธิภาพการรู้จำมากที่สุดบนฐานข้อมูล Yale และ ORL โดยพบว่าเมตริกซ์คุณลักษณะของ 2DPCA มีขนาด 100×23 เมตริกซ์คุณลักษณะของ 2DLDA มีขนาด 100×14 และเมตริกซ์คุณลักษณะของ 2DPCA+2DLDA มีขนาด 100×7 ซึ่งวิธีการ 2DPCA+2DLDA มีมิติของเมตริกซ์คุณลักษณะขนาดเล็กที่สุด และใช้เวลาทั้งฝึกฝนและทดสอบน้อยที่สุดบนฐานข้อมูล Yale และยังให้ความถูกต้องในการรู้จำมากที่สุด ในขณะที่เดียวกัน 2DPCA+2DLDA ที่ทดสอบบนฐานข้อมูล ORL ก็ให้ความถูกต้องในการรู้จำมากที่สุดเช่นกัน โดยมีมิติของเมตริกซ์คุณลักษณะเท่ากับของ 2DPCA แต่ใช้เวลามากกว่า 2DPCA เพียงไม่มาก อัตราการรู้จำของ 2DLDA บนฐานข้อมูล ORL มีค่าน้อยลงตามสมมติฐานที่ว่า 2DLDA จะนำข้อมูลที่ไม่มีนัยในการแยกแยะกลุ่มมาใช้ในการแยกแยะ แต่เมื่อนำ 2DPCA มาเป็นตัวสกัดคุณลักษณะก่อนป้อนเข้า 2DLDA พบว่าสามารถเพิ่มอัตราการรู้จำได้ดีขึ้นจริง

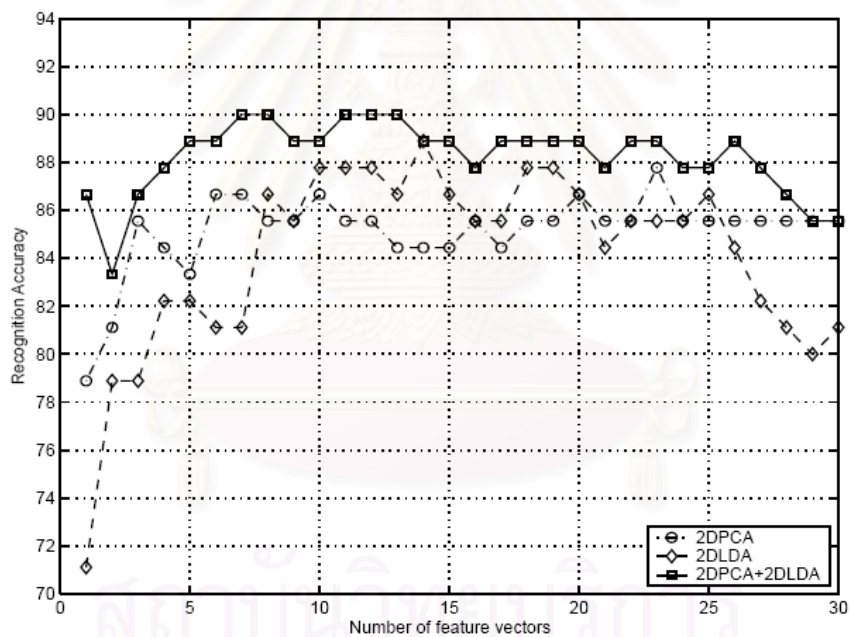
รูปที่ 18 และรูปที่ 19 แสดงความถูกต้องในการรู้จำเมื่อมีการเปลี่ยนแปลงจำนวนค่าเจาะจงที่ใช้ในการหาเวกเตอร์เจาะจงหรือ d แต่เนื่องจากอัตราการรู้จำของ 2DPCA+2DLDA ขึ้นกับจำนวนค่าเจาะจงที่ใช้ทั้ง 2DPCA และ 2DLDA ดังนั้นเส้นกราฟที่แสดงของ 2DPCA+2DLDA จึงเป็นเส้นกราฟที่แสดงจุดสูงสุดของกราฟในรูปที่ 20 และรูปที่ 21

ตารางที่ 4 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล Yale

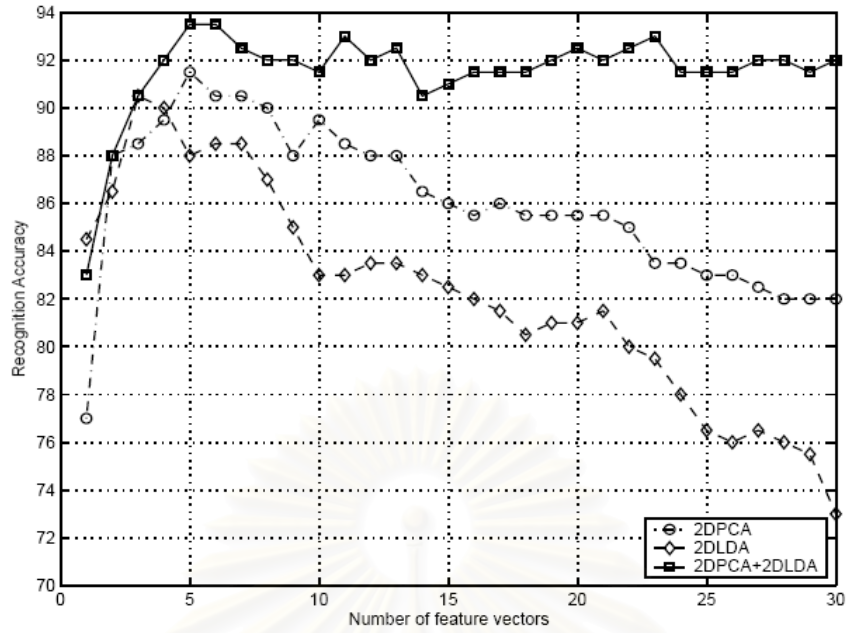
วิธีการ	ความถูกต้อง (%)	d	q	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	87.78	23	-	100×23	0.17 (1)	3.14 (1)	3.31 (1)
2DLDA	88.89	-	14	100×14	0.16 (0.94)	1.94 (0.62)	2.10 (0.63)
2DPCA+2DLDA	90.00	21	7	100×7	0.24 (1.14)	1.01 (0.32)	1.25 (0.38)

ตารางที่ 5 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล ORL

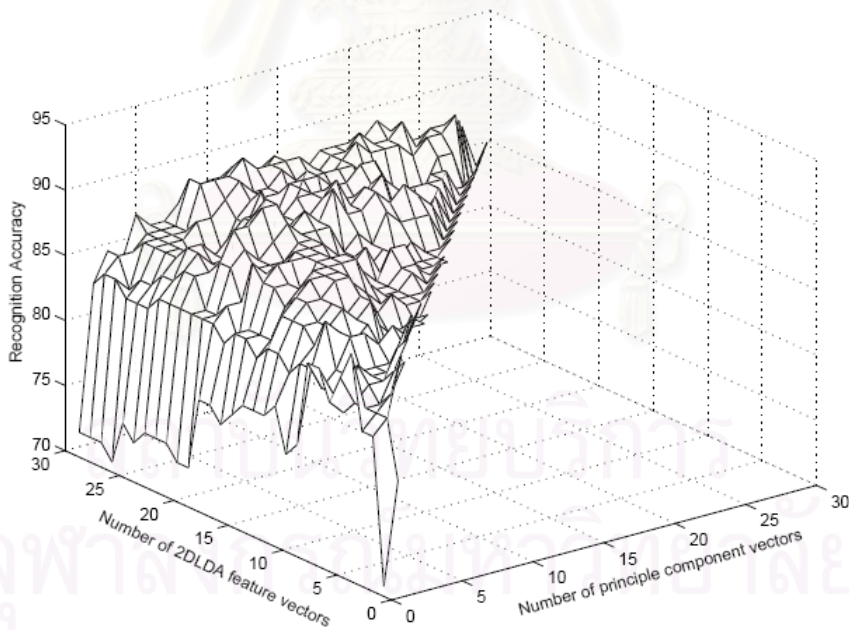
วิธีการ	ความถูกต้อง (%)	d	q	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	91.50	5	-	112 x 5	0.33 (1)	4.33 (1)	4.69 (1)
2DLDA	90.50	-	3	112 x 3	0.50 (1.51)	2.75 (0.64)	3.28 (0.7)
2DPCA+2DLDA	93.50	14	5	112 x 5	0.64 (1.94)	4.33 (1)	5.22 (1.07)



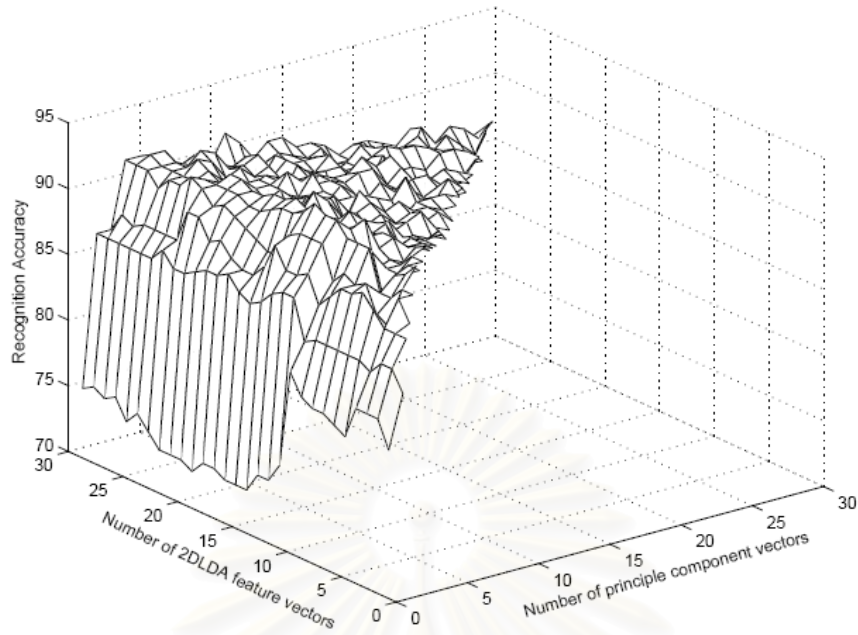
รูปที่ 18 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล Yale



รูปที่ 19 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล ORL



รูปที่ 20 ความถูกต้องในการรู้จำของวิธี 2DPCA+2DLDA เทียบกับจำนวนเวกเตอร์คุณลักษณะของ 2DLDA และ 2DPCA บนฐานข้อมูล Yale

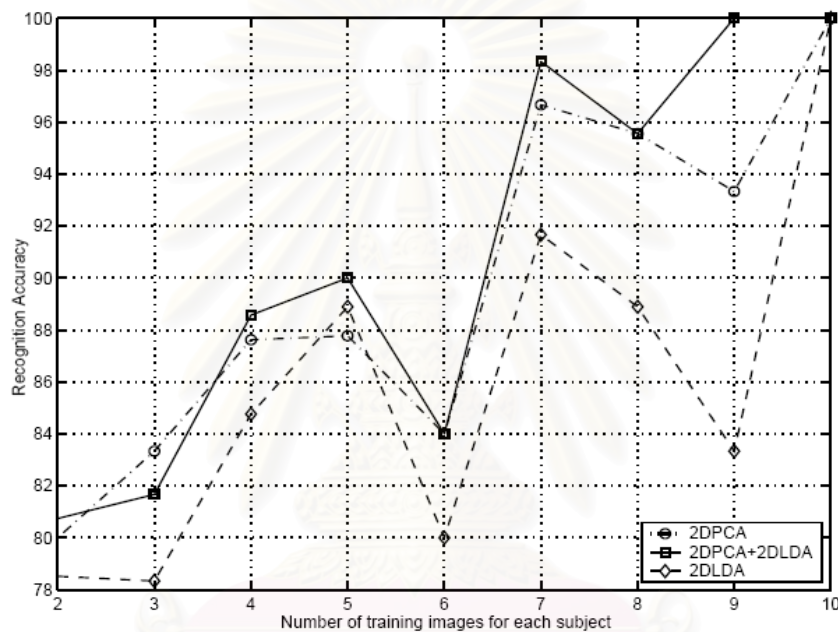


รูปที่ 21 ความถูกต้องในการรู้จำของวิธี 2DPCA+2DLDA เทียบกับจำนวนเวกเตอร์คุณลักษณะของ 2DLDA และ 2DPCA บนฐานข้อมูล ORL

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

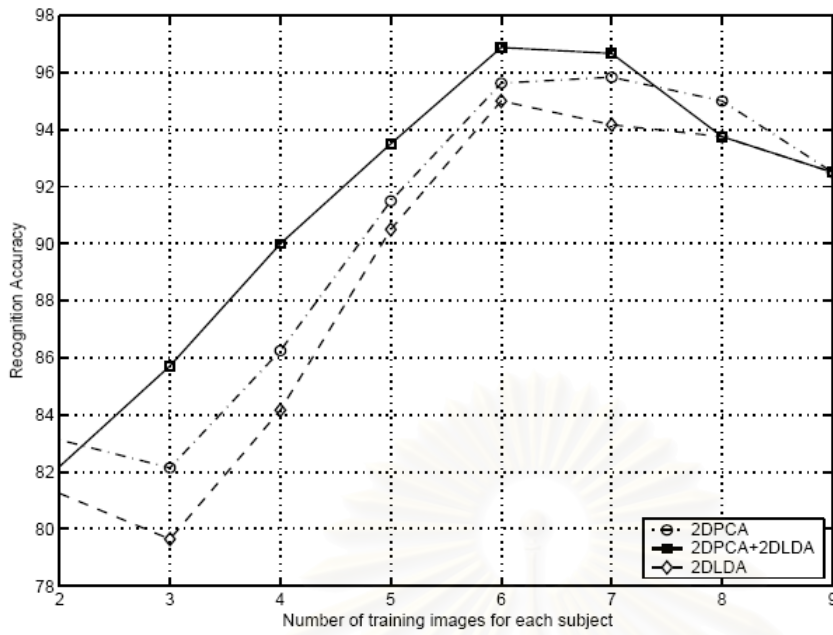
การทดลองที่ 2 ประสิทธิภาพของระบบรู้จำใบหน้าด้วยวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA เมื่อมีการเปลี่ยนแปลงจำนวนตัวอย่างในการฝึกฝน

บนฐานข้อมูล Yale ทำการเปลี่ยนแปลงจำนวนตัวอย่างฝึกฝนตั้งแต่ 2 ถึง 10 โดยภาพทดสอบที่ใช้คือภาพที่ไม่ได้ใช้ในการฝึกฝน ในขณะที่เดียวกันบนฐานข้อมูล ORL ทำการเปลี่ยนแปลงจำนวนตัวอย่างฝึกฝนตั้งแต่ 2 ถึง 9 โดยภาพทดสอบที่ใช้คือภาพที่ไม่ได้ใช้ในการฝึกฝนได้ดังรูปที่ 22 และรูปที่ 23 ตามลำดับ



รูปที่ 22 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA เมื่อมีการเปลี่ยนแปลงจำนวนใบหน้าที่ใช้ฝึกฝนในแต่ละกลุ่มบนฐานข้อมูล Yale

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 23 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA เมื่อมีการเปลี่ยนแปลงจำนวนใบหน้าที่ใช้ฝึกฝนในแต่ละกลุ่มบนฐานข้อมูล ORL

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การทดลองที่ 3 ประสิทธิภาพของระบบรู้จำใบหน้าด้วยวิธีการปริภูมิย่อยเจาะจงหน้าเปรียบเทียบกับ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS

ประยุกต์ใช้วิธีปริภูมิย่อยเจาะจงหน้ากับ 2DPCA และ B2DPCA บนฐานข้อมูล Yale AR และ MSTAR ได้ผลดังตารางที่ 6 ตารางที่ 7 และตารางที่ 8 โดย d และ q คือจำนวนค่าเจาะจงที่ใช้ในแนวแถวและหลัก ตามลำดับ โดยใช้ฐานข้อมูล MSTAR ใช้ข้อมูลเชิงขนาดมาสร้างเป็นภาพแล้วทำการตรวจรู้จำตำแหน่งการวางตัวของรถด้วยมือ โดยแบบเป็นช่วงช่วงละ 5 10 15 20 30 และ 40 องศา

ตารางที่ 6 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS บนฐานข้อมูล Yale

วิธีการ	ความถูกต้อง (%)	d	q	มิติ
2DPCA	87.78	5	-	100 x 5
B2DPCA	92.22	23	22	22 x 23
2DPCA+FSS	92.22	5	-	100 x 5
B2DPCA+FSS	94.44	1	1	1 x 1

ตารางที่ 7 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS บนฐานข้อมูล AR

วิธีการ	ความถูกต้อง (%)	d	q	มิติ
2DPCA	54.49	5	-	100 x 5
B2DPCA	55.45	23	22	22 x 23
2DPCA+FSS	56.41	5	-	100 x 5
B2DPCA+FSS	59.46	1	1	1 x 1

ตารางที่ 8 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS ที่ช่วงมุมต่าง ๆ บนฐานข้อมูล MSTAR

วิธีการ	ช่วงมุม (องศา)	ความถูกต้อง (%)	d	q
2DPCA	5	98.28	5	-
	10	98.35	5	-
	15	98.73	5	-
	20	98.57	5	-
	30	98.66	5	-
	40	98.24	5	-
B2DPCA	5	98.24	5	10
	10	98.14	5	10
	15	98.14	5	10
	20	98.15	5	10
	30	98.35	5	10
	40	97.62	5	10
2DPCA+FSS	5	93.10	5	-
	10	94.57	5	-
	15	94.22	5	-
	20	93.14	5	-
	30	90.22	5	-
	40	86.62	5	-
B2DPCA+FSS	5	94.54	5	10
	10	95.93	5	10
	15	95.60	5	10
	20	93.80	5	10
	30	91.82	5	10
	40	88.16	5	10

การทดลองที่ 4 ประสิทธิภาพของระบบรู้จำใบหน้าด้วยวิธีใช้วิธีความละเอียดหลายระดับ

1 ทำการทดลองโดยใช้วิธี PCA ในการสกัดคุณลักษณะ

ทำการทดลองบนฐานข้อมูลใบหน้า Yale โดยใช้ใบหน้าในการฝึกสอนจำนวน 5 ใบหน้า และใช้ทดสอบจำนวน 6 ใบหน้าซึ่งได้ผลการทดลองดังตาราง พบว่าการนำระบบหลายตัวจำแนกมาประยุกต์ใช้กับวิธีความละเอียดหลายระดับและ PCA นั้นให้ผลการทดลองที่ดีขึ้น แต่ว่ากลุ่มแถบย่อยที่ให้ผลดีที่สุดของการทดลองโดยใช้ฐานข้อมูล Yale นั้น ต่างกับกลุ่มที่ให้ผลดีที่สุดในการทดลองบนฐานข้อมูล ORL

เพื่อลดขั้นตอนการคำนวณที่ใช้ในการทดลองลง จึงทำการทดลองโดยใช้ค่าขีดแบ่งของพลังงานจากค่าลักษณะเฉพาะที่ 90% ดังสมการที่ 25 โดยขั้นตอนดังกล่าวทำเพื่อลดภาระการคำนวณลง ซึ่งจำนวนเวกเตอร์ลักษณะเฉพาะที่ได้จากการคำนวณค่าขีดแบ่งนั้นไม่สามารถรับรองได้ว่าให้ผลการรู้จำสูงสุด แต่จะให้ผลการรู้จำที่ต่างกันเพียงเล็กน้อย ในบางกรณีอาจค่าเท่ากับผลการรู้จำสูงสุดที่เป็นไปได้

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{j=1}^N \lambda_j} = 0.9 \quad (4.50)$$

ตารางที่ 9 ผลการทดลองโดยใช้วิธี PCA บนฐานข้อมูล ORL

	Data fusion			Feature Fusion			Decision Fusion		
	L1	L2	CC	L1	L2	CC	L1	L2	CC
0th Level (Original)	85.50	78.50	85.00	85.50	78.50	85.00	85.50	78.50	85.00
1st Level	82.00	77.50	89.50	83.00	78.50	84.50	57.00	64.00	90.00
2nd Level	76.50	82.00	93.00	82.50	80.00	89.00	77.50	77.50	90.50
3rd Level	71.00	75.00	91.50	77.50	80.00	85.50	80.50	82.00	93.00
1st-3rd Level (Data fusion)				82.00	81.00	93.00	75.00	82.00	93.00
1st-3rd Level (Feature fusion)							81.50	83.00	88.50
1st-3rd Level (Decision fusion)							79.00	82.00	92.50
0th-3rd Level (Data fusion)				82.00	81.00	93.00	84.00	78.50	91.50
0th-3rd Level (Feature fusion)							84.00	81.00	89.50
0th-3rd Level (Decision fusion)							84.50	84.00	94.00

ตารางที่ 10 ผลการทดลองโดยใช้วิธี PCA บนฐานข้อมูล Yale

	Data fusion			Feature Fusion			Decision Fusion		
	L1	L2	CC	L1	L2	CC	L1	L2	CC
0th Level (Original)	91.11	87.78	81.11	91.11	87.78	81.11	91.11	87.78	81.11
1st Level	88.89	87.78	86.67	88.89	85.56	85.56	77.78	83.33	88.89
2nd Level	82.22	84.44	90.00	84.44	85.56	87.78	86.67	85.56	90.00
3rd Level	80.00	86.67	88.89	84.44	88.89	88.89	90.00	91.11	91.11
1st-3rd Level (Data fusion)				87.88	92.42	93.94	83.33	85.56	90.00
1st-3rd Level (Feature fusion)							86.67	84.44	87.78
1st-3rd Level (Decision fusion)							86.67	86.67	87.78
0th-3rd Level (Data fusion)				87.88	92.42	93.94	88.89	87.78	88.89
0th-3rd Level (Feature fusion)							84.44	85.56	88.89
0th-3rd Level (Decision fusion)							91.11	90.00	90.00

2 ทำการทดลองโดยใช้วิธี 2DPCA ในการสกัดคุณลักษณะ

จากผลการทดลองที่ได้ดังตาราง พบว่าคุณลักษณะที่ได้จากการแปลงเวฟเลิตแบบไม่ต่อเนื่องนั้น ในแต่ละแถบย่อยซึ่งอยู่ในระดับที่ต่างกัน จะมีขนาดไม่เท่ากัน ทำให้ไม่สามารถทำการรวมแบบข้อมูลในระบบตัวจำแนกในระดับที่หนึ่งได้ และจากผลการทดลองพบว่า วิธีนี้ไม่สามารถปรับปรุงผลการรู้จำได้ สำหรับฐานข้อมูล ORL ทำได้ดีที่สุดได้เพียงเท่ากับการใช้เพียงวิธี 2DPCA เพียงวิธีเดียวในการคำนวณเท่านั้น ส่วนฐานข้อมูล Yale นั้น ไม่สามารถปรับปรุงผลการรู้จำได้

จากอัตราการรู้จำของวิธี 2DPCA นั้น มีอัตราการรู้จำสูงสุดเพียงบางค่า เนื่องจากกราฟของอัตราการรู้จำนั้นมีลักษณะเป็นภูเขา ซึ่งจะเพิ่ม และลดลงอย่างรวดเร็ว ซึ่งต่างกับกราฟอัตราการรู้จำของวิธี PCA ที่เมื่อได้อัตราการรู้จำสูงสุดแล้ว มักจะคงที่ที่ค่าสูงสุด หรือลดลงเพียงเล็กน้อยต่อไป จากสาเหตุดังกล่าวทำให้ไม่สามารถใช้ขีดแบ่งของพลังงานจากค่าลักษณะเฉพาะเพื่อลดขั้นตอนการคำนวณสำหรับวิธี 2DPCA ลงได้

ตารางที่ 11 ผลการทดลองโดยใช้วิธี 2DPCA บนฐานข้อมูล ORL

	Data fusion			Feature Fusion			Decision Fusion		
	sumL1	sumL2	sumCC	sumL1	sumL2	sumCC	sumL1	sumL2	sumCC
0th Level (Original)	94.00	92.50	92.50	94.00	91.50	91.50	94.00	92.50	92.50
1st Level				89.00	90.50	89.50	83.00	78.00	83.50
2nd Level				92.00	91.00	93.00	78.00	88.00	90.50
3rd Level				91.50	90.50	91.00	83.50	87.00	90.00
1st-3rd Level (Data fusion)							91.50	91.00	92.00
1st-3rd Level (Feature fusion)							90.00	88.00	87.00
1st-3rd Level (Decision fusion)							92.00	90.50	92.00
0th-3rd Level (Data fusion)							94.00	90.00	87.00

ตารางที่ 12 ผลการทดลองโดยใช้วิธี 2DPCA บนฐานข้อมูล Yale

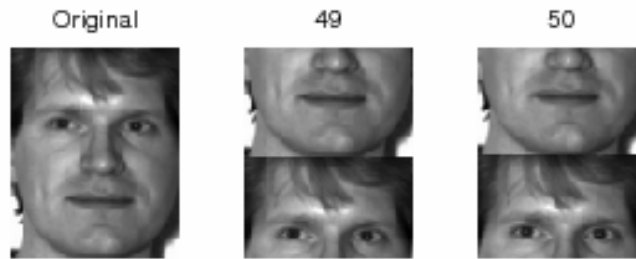
	Data fusion			Feature Fusion			Decision Fusion		
	sumL1	sumL2	sumCC	sumL1	sumL2	sumCC	sumL1	sumL2	sumCC
0th Level (Original)	94.444	94.444	97.778	96.667	92.222	91.111	94.444	94.444	97.778
1st Level				88.889	88.889	88.889	93.333	91.111	95.556
2nd Level				87.778	87.778	92.222	94.444	90.000	94.444
3rd Level				88.889	92.222	94.444	96.667	91.111	95.556
1st-3rd Level (Data fusion)							88.889	88.889	92.222
1st-3rd Level (Feature fusion)							94.444	91.111	94.444
1st-3rd Level (Decision fusion)							87.778	87.778	90.000
0th-3rd Level (Data fusion)							96.667	92.222	95.556

การทดลองที่ 5 ประสิทธิภาพของระบบรู้จำใบหน้าด้วยวิธีใช้วิธี ICCA

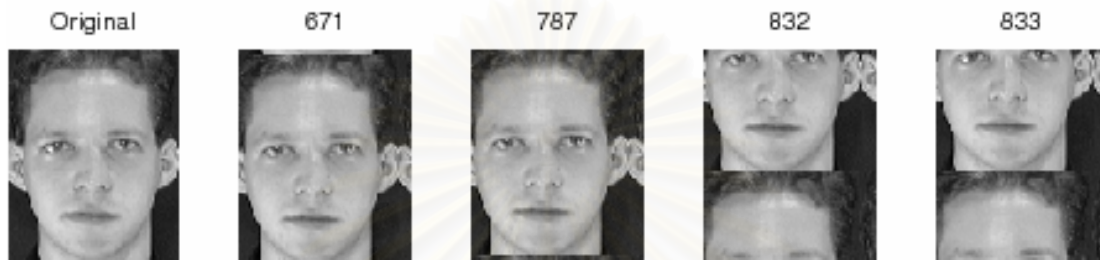
ประยุกต์ใช้วิธี ICCA และ 2DPCA บนฐานข้อมูล Yale AR และ ORL ได้ผลดังตารางที่ 13 โดย d คือจำนวนค่าเจาะจงสูงสุดที่ใช้ตามลำดับ และ L คือจำนวนการเลื่อนภาพ

ตารางที่ 13 ผลการทดลองโดยใช้วิธี ICCA บนฐานข้อมูล Yale AR และ ORL

Database	Method	Accuracy	d	L	Train time	Test time
Yale	2DPCA	95.56	20	1	0.03	0.34
	ICCA	97.78	20	49	0.03	0.34
ORL	2DPCA	92.50	5	1	0.08	0.51
	ICCA	95.00	5	671	0.11	0.51
AR	2DPCA	60.74	2	1	0.16	2.34
	ICCA	61.70	2	9	1.53	2.34



รูปที่ 24 ลักษณะการเลื่อนภาพของรูปที่ทำให้ได้ความถูกต้องสูงสุดบนฐานข้อมูล Yale

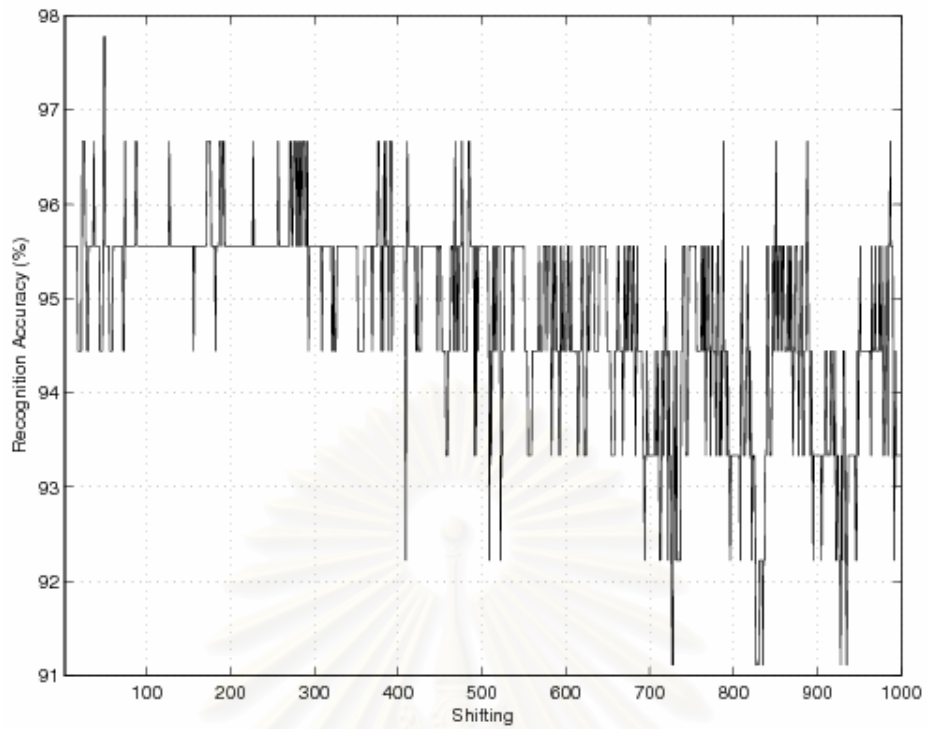


รูปที่ 25 ลักษณะการ Shift ของรูปที่ทำให้ได้ความถูกต้องสูงสุดบนฐานข้อมูล ORL

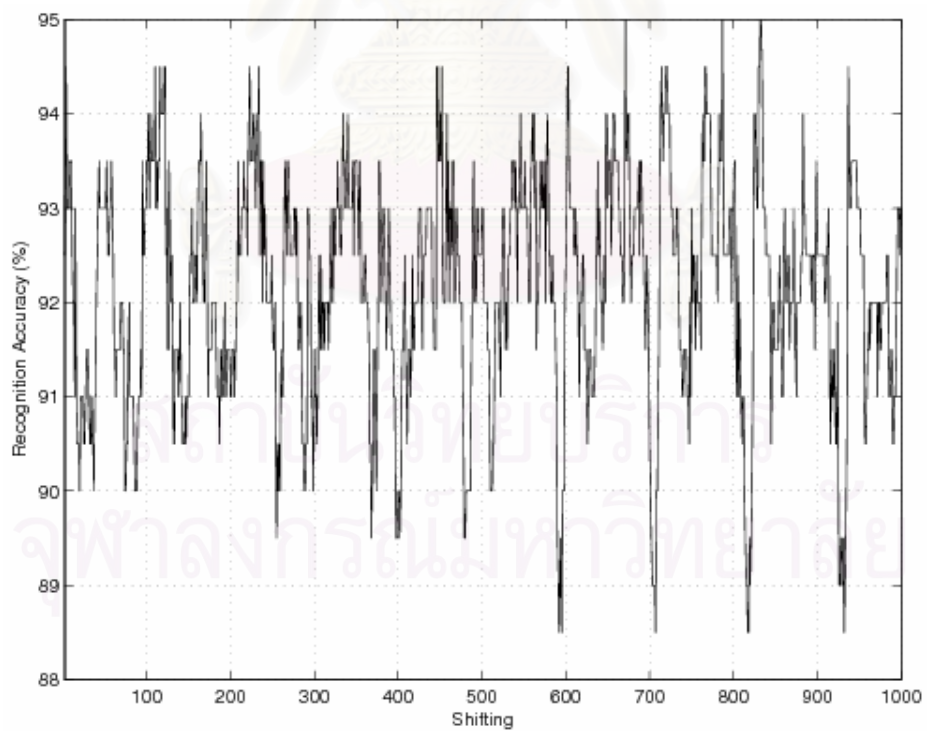


รูปที่ 26 ลักษณะการ Shift ของรูปที่ทำให้ได้ความถูกต้องสูงสุดบนฐานข้อมูล AR

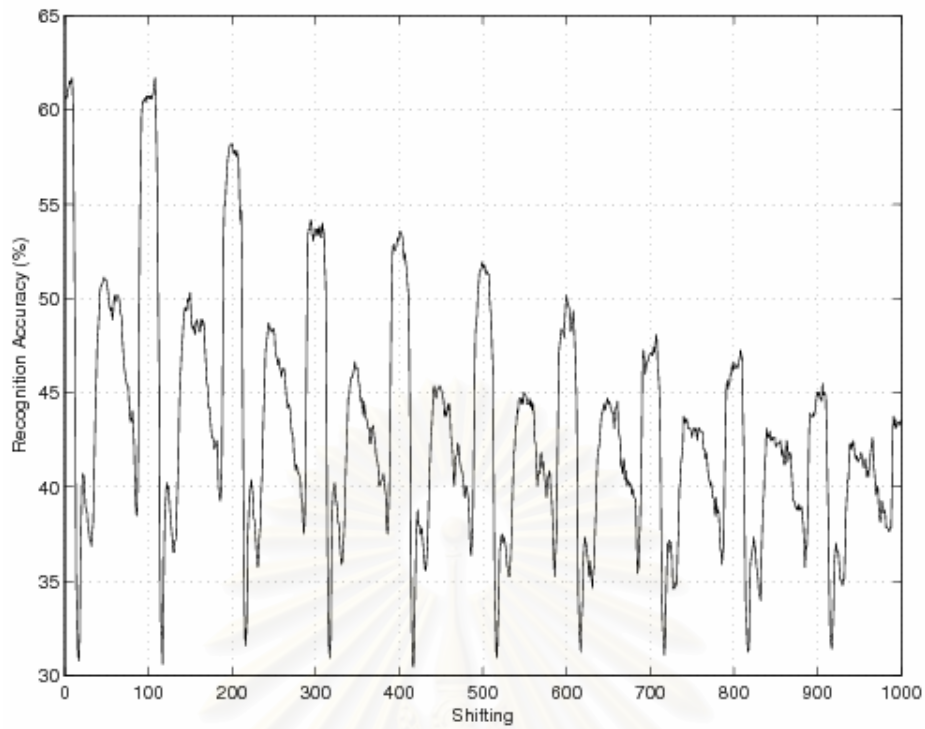
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 27 ความถูกต้องในการรู้จำของวิธี ICCA เมื่อมีการเปลี่ยนแปลงจำนวนการเลื่อนภาพบน
ฐานข้อมูล Yale



รูปที่ 28 ความถูกต้องในการรู้จำของวิธี ICCA เมื่อมีการเปลี่ยนแปลงจำนวนการเลื่อนภาพบน
ฐานข้อมูล ORL



รูปที่ 29 ความถูกต้องในการรู้จำของวิธี ICCA เมื่อมีการเปลี่ยนแปลงจำนวนการเลื่อนภาพบน
ฐานข้อมูล AR

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุปผลการทดลอง

เราได้เสนอวิธีการใหม่ขึ้นมาดังนี้

- Two-Dimensional Linear Discriminant Analysis (2DLDA)
- Two-Dimensional Linear Discriminant Analysis of Principal Component Vectors (2DPCA+2DLDA)
- Face-Specified Subspace based Two-Dimensional Principal Component Analysis (2DPCA+FSS)
- Face-Specified Subspace Bilateral-projection based Two-Dimensional Principal Component Analysis (B2DPCA+FSS)
- Decision Combination of Multiple Classifier Systems for Multiresolution Face Recognition
- Image Cross-Covariance Analysis (ICCA)

จากการทดลองพบว่า ระบบรู้จำใบหน้าแบบ 2DPCA+2DLDA มีประสิทธิภาพในการรู้จำเพิ่มขึ้น และใช้เวลาในการรู้จำเพิ่มขึ้นไม่มาก รวมถึงสามารถลดจำนวนขนาดมิติของเมตริกซ์คุณลักษณะลงได้อีก โดยการใช้ 2DLDA ที่เสนอขึ้นใหม่เป็นตัวแยกประเภทหลังจากสกัดคุณลักษณะด้วย 2DPCA มาก่อน

ระบบรู้จำที่ใช้ปริภูมิย่อยเจาะจงเชิงภาพที่ใช้ 2DPCA และ B2DPCA สามารถเพิ่มอัตราการรู้จำบนฐานข้อมูล Yale และ AR ได้ แต่กลับทำให้อัตราการรู้จำบนฐานข้อมูล MSTAR มีค่าน้อยลง ทั้งนี้เนื่องมาจาก จำนวนตัวอย่างที่ใช้ในการฝึกฝนในแต่ละกลุ่มมีมากอยู่แล้วทำให้การประมาณค่าของเมตริกซ์ความแปรปรวนร่วมของ 2DPCA และ B2DPCA ทำได้ดีกว่า แต่วิธีการที่ใช้ปริภูมิย่อยเจาะจงเชิงภาพเหมาะกับข้อมูลที่มีจำนวนตัวอย่างที่ใช้ฝึกฝนน้อย

ระบบรู้จำที่ใช้ความละเอียดภาพหลายระดับการใช้ระบบหลายตัวจำแนกสองระดับในการสร้างชุดโครงสร้างการรวมแถบย่อยนั้น ถ้าวิเคราะห์ในส่วนของการสกัดคุณลักษณะแล้ว จากการทดลองพบว่าแถบย่อยที่ได้จากการทำเวฟเล็ตแบบไม่ต่อเนื่องนั้นในแต่ละระดับ จะได้ขนาดของแถบย่อยที่ไม่เท่ากัน แต่เมื่อใช้ PCA ในการสกัดคุณลักษณะแล้วการมีขนาดที่ไม่เท่ากันไม่เป็นปัญหา เนื่องจาก PCA มีการคิดเริ่มต้นเวกเตอร์ แต่ในส่วนของการรวมแถบย่อยนั้น เริ่มต้นมีการคำนวณโดยใช้เมตริกซ์ของภาพนั้นๆ ทำให้เมื่อทำการรวมวิธีต่างๆ เกิดปัญหาขึ้น ซึ่งเมื่อสังเกตจากผลการทดลองโดยรวมแล้วพบว่า วิธีที่นำเสนอขึ้นเมื่อใช้ PCA ในการสกัดคุณลักษณะจะให้อัตรา

การรู้จำที่พัฒนาขึ้นจากการใช้ PCA เพียงอย่างเดียวมาก โดยจะสังเกตได้จากทั้งสองฐานข้อมูล แต่ผลการทดลองเมื่อใช้ 2DPCA ในการสกัดคุณลักษณะพบว่าอัตราการรู้จำไม่สามารถพัฒนาขึ้นได้เลย ซึ่งเป็นผลจากคุณลักษณะที่ใช้ไม่เหมาะสมกับวิธีที่นำเสนอนี้ วิธีที่นำเสนอพบว่าเป็นวิธีหนึ่งซึ่งสามารถปรับปรุงผลการรู้จำให้ได้ดีขึ้นมากเมื่อเทียบกับการใช้ PCA พื้นฐาน ให้อัตราการรู้จำที่สูงกว่างานวิจัยที่ได้ใช้เพียงแถบย่อยเพียงแถบเดียว และรวมถึงให้อัตราการรู้จำที่สูงกว่าการใช้ระบบหลายตัวจำแนกในระดับเดียวได้ ซึ่งมีข้อสังเกตว่าวิธีที่นำเสนอนี้มีการใช้จำนวนแถบย่อยให้การวิเคราะห์ที่มากกว่า และมีการคำนวณที่ซับซ้อนกว่ามาก ถ้าสนใจแต่ในแง่ของอัตราการรู้จำได้แล้ว วิธีนี้ถือว่าเป็นวิธีที่น่าสนใจมากวิธีหนึ่ง

ส่วนวิธีการ ICCA นั้นมีข้อดีคือสามารถได้ความถูกต้องในการรู้จำสูงกว่า 2DPCA โดยที่มีความซับซ้อนในการทดสอบภาพตัวอย่างไม่ต่างกัน แต่ข้อเสียคือยังไม่มีเงื่อนไขที่เหมาะสมในการหาจำนวนการเลื่อนของภาพ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5. ส่วนที่จะดำเนินการต่อไป

5.1. ศึกษาการรู้จำใบหน้าเพิ่มเติม

6. ผลผลิตและหรือความสัมฤทธิ์ผลของงานที่ได้ดำเนินการไปแล้ว

6.1. วารสารวิชาการนานาชาติเรื่อง Two-Dimensional Linear Discriminant Analysis of Principle Component Vectors for Face Recognition ซึ่งได้รับการตอบรับเพื่อตีพิมพ์จาก The Institute of Electronics, Information and Communication Engineers (IEICE) Transaction on Information and Systems

6.2. บทความวิชาการนานาชาติเรื่อง Two-Dimensional Linear Discriminant Analysis of Principle Component Vectors for Face Recognition ซึ่งได้รับการตอบรับให้ไปนำเสนอที่งานประชุม 2006 IEEE International Conference Acoustics, Speech, and Signal Processing (ICASSP 2006) ณ เมืองตูรูส ประเทศฝรั่งเศส วันที่ 14-16 พฤษภาคม พ.ศ. 2549

6.3. บทความทางวิชาการนานาชาติเรื่อง Class-Specific Subspace-Based Two-Dimensional Principal Component Analysis for Face Recognition ซึ่งได้รับการตอบรับให้ไปนำเสนอที่งานประชุม 2006 the 18th International on Pattern Recognition (ICPR 2006) ณ เมืองฮ่องกง ประเทศจีน วันที่ 20-24 สิงหาคม พ.ศ. 2549

6.4. Decision Combination of Multiple Classifier Systems for Multiresolution Face Recognition ซึ่งได้รับการตอบรับให้ไปนำเสนอที่งานประชุม 2007 the fourth annual international conference organized by Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI) Association ณ จังหวัดเชียงราย ประเทศไทย วันที่ 9-12 พฤษภาคม พ.ศ. 2550

รายการอ้างอิง

- [1] <http://web.media.mit.edu/~tristan/phd/dissertation/chapter5.html>
- [2] <http://www.dtreg.com/lda.htm>
- [3] L. Sirovich and M. Kirby, Low-dimensional procedure for characterization of human faces," J. Optical Soc. Am., vol.4, pp.519-524, 1987.
- [4] M. Turk and A. Pentland, Eigenfaces for recognition," J. of Cognitive Neuroscience, vol.3, no.1, pp.71-86, 1991.
- [5] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. and Mach. Intell., vol.19, pp.711-720, July 1997.
- [6] W. Zhao, R. Chellappa, and A. Krishnaswamy, Discriminant analysis of principle components for face recognition," Japan, IEEE 3rd Inter. Conf. on Automatic Face and Gesture Recognition, 1998.
- [7] W. Zhao, R. Chellappa, and N. Nandhakumar, Empirical performance analysis of linear discriminant classifiers," CVPR, pp.164-171, IEEE Computer Society, 1998.
- [8] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, 1990.
- [9] L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, A new LDA based face recognition system which can solve the small sample size problem," Pattern Recognition, vol.33, no.10, pp.1713-1726, 2000.
- [10] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, Regularized discriminant analysis for the small sample size problem in face recognition," Pattern Recogn. Lett., vol.24, no.16, pp.3079-3087, 2003.
- [11] R. Huang, Q. Liu, H. Lu, and S. Ma, Solving the small sample size problem of LDA," Pattern Recognition, vol.3, pp.29-32, 2002.
- [12] J. Yang, D. Zhang, A.F. Frangi, and J. Yu Yang, Two-dimensional PCA: A new approach to appearance-based face representation and recognition," IEEE Trans. Pattern Anal. and Mach. Intell., vol.26, pp.131-137, Jan. 2004.
- [13] J. Ye, R. Janardan, and Q. Li, Two-dimensional linear discriminant analysis," in Advances in Neural Information Processing Systems 17, ed. L.K. Saul, Y. Weiss, and L. Bottou, pp.1569-1576, MIT Press, Cambridge, MA, 2005.

- [14] J. Yang and J.Y. Yang, From image vector to matrix: A straightforward image projection technique IMPCA vs. PCA," Pattern Recognition, vol.35, no.9, pp.1997-1999, 2002.
- [15] Yale, The Yale face database," 1997. Available from <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
- [16] ORL, The ORL face database at the AT&T (Olivetti) research laboratory," 1992. Available from <http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html>.
- [17] D. Zhang and Z.H. Zhou, (2D)²PCA: 2-directional 2-dimensional PCA for efficient face representation and recognition," Neurocomputing, 2005.
- [18] H. Kong, X. Li, L. Wang, E.K. Teoh, J.G. Wang, and R. Venkateswarlu, Generalized 2D principal component analysis," IEEE International Joint Conference on Neural Networks (IJCNN), 2005.
- [19] S. Shan, W. Gao, and D. Zhao, Face recognition based on face-specific subspace," International Journal of Imaging Systems and Technology, vol.13, no.1, pp.23-32, 2003.
- [20] H. K. Ekenel and B. Sankur, "Multiresolution Face Recognition," *Image and Vision Computing*, vol. 3, no. 5, pp. 469-477, 2005.
- [21] T. G. Dietterich, "Machine Learning Research: Four Current Direction," *AI Magazine, Winter*, pp. 97-136, 1997.
- [22] G. C. Feng, P. C. Yuen, and D. Q. Dai, "Human Face Recognition Using PCA on Wavelet Subband," SPIE Journal of Electronic Imaging, vol. 9, no.2, pp. 226-233, 2000.
- [23] J. T. Chien and C. C. Wu, "Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 12, December 2002.
- [24] V. T. Brennen, "Principal Component Analysis with Multiresolution," Ph. D. thesis, University of Florida, May, 2001.
- [25] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma, and N. Otsu, "Face Recognition System Using Local Autocorrelations and Multiscale Integration," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 10, October 1996.

- [26] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision Combination in Multiple Classifier Systems," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 1, January 1994.
- [27] A. K. Jain, X. Lu, and Y. Wang, "Combining Classifiers for Face Recognition," in Proc. ICME, pp. 13-16, Baltimore, MD, July, 2003.
- [28] T. H.-Man, "Face Recognition Committee Machine: Methodology, Experiments and a System Application," M. thesis, Chinese University of Hong Kong, July, 2003.
- [29] J. A. Benediktsson and O. K. Ersoy, "Parallel Consensual Neural Networks," IEEE Trans. Neural Networks, vol. 8, no. 1, pp. 54-64, 1997.
- [30] Atalay, "Face Recognition Using Eigenface," M. Sc. thesis, Istanbul Technical University, 1996.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย