

โครงการวิจัยย่อยลำดับที่ 4
เรื่อง การรู้จำใบหน้า (Image Recognition)

1. ผู้รับผิดชอบโครงการ รศ.ดร. สมชาย จิตะพันธ์กุล

2. วัตถุประสงค์ของโครงการ รู้จำใบหน้า

เพื่อพัฒนาวิธีการแยกแยะใบหน้าของบุคคลต่างกันอย่างมีประสิทธิภาพได้ และนำวิธีการที่ได้ไปประยุกต์ใช้กับระบบรู้จำใบหน้า

1. ขอบเขตหรือเป้าหมายของโครงการ ศึกษาวิจัย และประยุกต์ใช้ ระบบการรู้จำใบหน้า

- 1.1. สร้างระบบรู้จำใบหน้าด้วยเทคนิคที่ปรับปรุงขึ้นใหม่
- 1.2. รูปใบหน้าคนที่ใช้เป็นภาพ 2 มิติ ที่มีส่วนของใบหน้าครบ
- 1.3. รูปใบหน้าคนที่ใช้มีการตรวจจับบริเวณ ใบหน้าที่ต้องการตรวจสอบมาแล้ว
- 1.4. รูปใบหน้าคนที่ใช้เป็นภาพหน้าตรงมีการหัน ก้มและเงยหน้าไม่เกิน 20 องศาจากภาพหน้าตรง
- 1.5. รูปใบหน้าคนที่ใช้สามารถแสดงอารมณ์แตกต่างกันได้
- 1.6. รูปใบหน้าคนที่ใช้ต้องมีขนาดมากกว่า 80 x 80 พิกเซล และคมชัดเมื่อมองด้วยตาเปล่า
- 1.7. รูปใบหน้าคนที่ใช้ต้องไม่มีสิ่งของที่บังบริเวณใบหน้าเกิน 10% ของส่วนที่เป็นใบหน้าทั้งหมด ยกเว้นในกรณีที่บุคคลนั้นสวมใส่สิ่งของดังกล่าวเป็นประจำ
- 1.8. รูปใบหน้าคนที่ใช้ทั้งหมดที่ใช้ฝึกฝนและทดสอบของบุคคลแต่ละคนจะต้องเก็บข้อมูลภายในช่วงเวลาเดียวกัน
- 1.9. อัตราความถูกต้องเฉลี่ยของระบบรู้จำไม่ต่ำกว่าร้อยละ 90 บนฐานข้อมูลที่มีบุคคลต่างกัน ไม่เกิน 40 คน และมีจำนวนภาพที่ใช้ฝึกฝนน้อยกว่าจำนวนภาพที่ใช้ทดสอบไม่เกินร้อยละ 50

2. ส่วนงานที่ได้ดำเนินการไปแล้ว

2.1. ศึกษาวิธีการรู้จำใบหน้าแบบต่างๆ และงานที่เกี่ยวข้อง โดยรวมสรุปได้ดังนี้

การรู้จำใบหน้าเป็นส่วนหนึ่งของแนวคิดทาง Biometrics ซึ่งใช้การสังเกตจากการวัดสิ่งต่างๆ ในร่างกาย ซึ่งแบ่งออกเป็น 2 ประเภท ดังนี้

1. Physiological Biometrics จะใช้ลักษณะเด่นที่เป็นอวัยวะต่างๆ ของร่างกาย เช่นที่ใช้กันในปัจจุบัน คือ ลายนิ้วมือ, ใบหน้า, เรขาคณิตของมือ, ม่านตา, เรตินา

2. Psychological Biometrics จะใช้สิ่งที่แสดงออกมา เช่น ที่ใช้ในปัจจุบัน คือ ลายมือ , ลายเซ็น , เสียง เป็นต้น

ซึ่งระบบ Biometrics นั้นจะใช้ระบบรู้จำรูปแบบเป็นหัวใจของระบบ ซึ่งทำให้แยกแยะบุคคลได้จาก 2 ประเภทที่กล่าวมาข้างต้น ในความจริงแล้วนั้น ระบบที่จะแยกแยะบุคคลได้นั้น ถ้าใช้เพียงวิธีเดียวในการทำงานนั้น จะให้ผลที่เชื่อถือได้ไม่มากนัก ซึ่งการรู้จำใบหน้า นั้น เป็นธรรมชาติ และไม่มี การรบกวนบุคคล แต่ว่าผลการทำงานได้เป็นที่ไม่น่าพอใจมากนัก แต่ทั้งที่การใช้เรตินาในการแยกแยะ นั้น ให้ผลที่เชื่อถือได้มากกว่าวิธีอื่นๆ แต่ ว่า มัน เป็นวิธีที่เกิดการรบกวนต่อบุคคลขึ้น ซึ่งหลายวิธี เช่น การรู้จำใบหน้าจึงได้ถูกคิดค้นขึ้นเพื่งแก้ปัญหาดังกล่าว

การรู้จำใบหน้าแบ่งเป็น 3 ระดับที่แตกต่างกัน คือ

1. Face Representation ประกอบด้วยรูปแบบของใบหน้าที่จะนำไปเข้ารหัส ซึ่งนั่นคือกฎเกณฑ์สำคัญๆ ในการแทนในฐานข้อมูลใบหน้าที่มี แต่ ว่าวิธีนี้ไม่สามารถทำได้ง่าย และมีแนวคิดที่ซับซ้อนที่จะให้ผลการแทนออกมาที่ดี ซึ่งวิธีนี้นั้นนำไปสู่แนวคิดที่ประสบความสำเร็จ เช่น การตรวจจับ (Detection) และการตรวจรู้ (Identification) สามารถแบ่งออกเป็น 3 ประเภท ดังนี้

1.1. Template-Based ซึ่งมีอยู่ 2 แบบ คือ แทนใบหน้าโดยค่าเฉลี่ยของ Template 2 มิติ กับค่าที่แทนวงรี ซึ่งเป็น โครงหน้า และ อวัยวะทั้งหมดบนใบหน้า

1.2. จะสมบูรณ์กว่าแบบแรกซึ่งจะมีการมองหลายๆ มุม ของใบหน้า และ ใช้ใบหน้าในหลายๆ ความรู้สึก จะดูตามส่วน ของ ตา, จมูก, ปากซึ่งเป็นส่วนที่สนใจ จุดเด่นของวิธีนี้ คือง่าย ส่วนจุดด้อย คือ ใช้หน่วยความจำมาก และให้สมรรถภาพไม่ดีในขั้นตอนการเปรียบเทียบ Feature-Based จะสนใจที่ตั้งและ ขนาดของอวัยวะ ขั้นตอนการเปรียบเทียบนั้นจะใช้เรขาคณิตโดยดูจากแต่ละรูปแบบของใบหน้า (ขนาด และที่ตั้งของอวัยวะต่างๆ) ในแต่ละรูปนั้นทุกระยะจะคำนวณจากอวัยวะในใบหน้า ส่วนผลที่ได้จะเป็นแบบ 1: 1 ในคุณลักษณะของใบหน้าที่ใช้ทดสอบ และ ส่วนของใบหน้าที่ใช้ในการฝึกสอน จะใช้การค้นหา Gradient ทางแนวตั้งในการหาส่วนบนสุดของศีรษะ, ตา, ฐานจมูก ส่วน Gradient ในแนวนอนนั้นใช้ในการตรวจหาด้านข้างของใบหน้า, จมูก ซึ่งในแต่ละใบหน้าจะคำนวณเวกเตอร์ขึ้น แล การรู้จำจะปฏิบัติกับตัวแยกแยะที่อยู่ใกล้เคียง Appearance-Based จะใช้การฉายใบหน้าลงไป ใบหน้าที่ลดมิติลง ซึ่ง Eigenface ก็จัดเป็นประเภทนี้ ซึ่งมันจะสร้างบนแกนฐานหลัก (PCA) ซึ่ง PCA จะเป็นเทคนิคที่นำมาศึกษาสำหรับรายงานนี้

2. Face Detection จะใช้ในส่วนของใบหน้าที่มีมุมต่างกัน ซึ่งซับซ้อน และต้องตรวจจับให้ได้ว่าส่วนไหนคือใบหน้า ซึ่งบางเทคนิคจะใช้รูปร่างในการตรวจ , ค้นหาจากสีผิว , ค้นหาอวัยวะบนใบหน้า หรือ จะเป็นการใช้ Artificial Neural Network

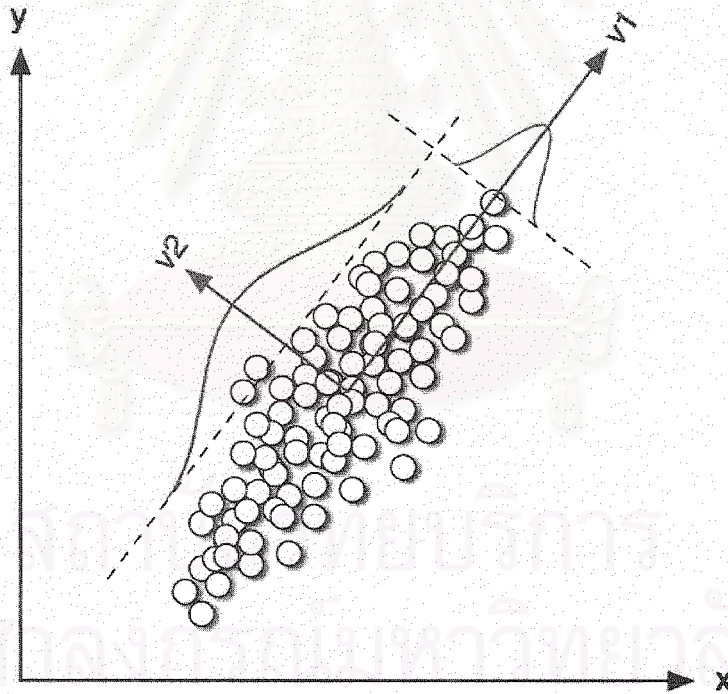
2.1. Automatic Face Recognition จะใช้การแยกแยะจากการเปรียบเทียบใบหน้าแบบ Facial Verification 1:1 หรือแบบ Face Identification 1:many ซึ่งจะเปรียบเทียบใบหน้าที่จะทดสอบกับ

ฐานข้อมูลที่มีอยู่ฐานข้อมูลที่ใช้ในการทดลองนั้น จากที่ศึกษามาพบว่ามีอยู่หลายประเภท เช่น ฐานข้อมูลที่น่าสนใจทางด้านแสง, ด้านความรู้สึกที่แสดงออกมา, ด้านมุมที่ใบหน้าหัน แต่การทดลองในรายงานนี้นั้น จะใช้ฐานข้อมูลที่แสดงความรู้สึกต่างๆ และมีใบหน้าที่หันได้ไม่เกิน ± 20 องศา นั่นคือฐานข้อมูล ORL และ จากที่อาจารย์ได้ให้มานั่นเอง

2.2. 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) โดยที่แกนแต่ละแกนแทนได้ด้วยเวกเตอร์เจาะจง (Eigen vector) จำนวนหนึ่งของเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) ที่สอดคล้องกับค่าเจาะจง (Eigen value) ที่มากที่สุดจำนวนหนึ่งดังรูปที่ 1 [1]



รูปที่ 1 ตัวอย่างการแปลง PCA แบบ 2 มิติ โดยแปลงจากปริภูมิคาร์ทีเซียน (x,y) ไปบนปริภูมิใหม่ (v1,v2)

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

ระเบียบวิธี PCA มีขั้นตอนดังนี้

กำหนดให้ Γ คือเวกเตอร์ของความเข้มแสงของพิกเซลของภาพที่ทำ Vectorization เชิงหลัก จากรูปที่นำเข้าไปเป็นกลุ่มฝึกฝน (Training set) มีขนาด $m \times n$ ให้เป็นเวกเตอร์ที่มีขนาด $mn \times 1$ และ Ψ คือเวกเตอร์เฉลี่ยของทุกๆ รูปในกลุ่มฝึกสอนดังสมการที่ (1)

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

โดย M คือจำนวนภาพทั้งหมดที่ใช้ในการฝึกฝน จากนั้นนำไปคำนวณหาเวกเตอร์ Φ ซึ่งคือ การทำให้ค่าเฉลี่ยของชุดฝึกฝนเป็นศูนย์ดังสมการที่ (2)

$$\Phi_i = \Gamma_i - \Psi, \quad (i=1, \dots, M) \quad (2)$$

กำหนดให้เมตริกซ์ A คือเมตริกซ์ที่แต่ละหลักคือเวกเตอร์ Φ_i จำนวน M เวกเตอร์ดังสมการที่ (3)

$$A = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M] \quad (3)$$

จากนั้นคำนวณเมตริกซ์ความแปรปรวนร่วม C ของเมตริกซ์ A ดังสมการที่ (4)

$$C = AA^T \quad (4)$$

โดยเมตริกซ์ความแปรปรวนร่วม C มีมิติขนาด $mn \times mn$ ซึ่งเป็นมิติขนาดใหญ่มาก ซึ่งไม่เหมาะสมที่จะนำไปหาค่าเจาะจง เนื่องจากต้องใช้เวลาและหน่วยความจำจำนวนมากในการคำนวณ แต่เนื่องจากจำนวนค่าเจาะจงที่ไม่เป็นศูนย์มีค่าเท่ากับจำนวนชั้น และจากคุณสมบัติดังสมการที่ (5)

$$\text{rank}(C) = \text{rank}(AA^T) = \text{rank}(A^T A) = \text{rank}(L) \quad (5)$$

โดย

$$L = A^T A \quad (6)$$

ซึ่งเมตริกซ์ L มีมิติขนาดเพียง $M \times M$ ดังนั้นการคำนวณค่าเฉพาะบนเมตริกซ์ L ใช้เวลาและหน่วยความจำน้อยกว่าการคำนวณบนเมตริกซ์ C โดยตรง ดังนั้นจึงนำเมตริกซ์ L ไปคำนวณค่าเฉพาะดังสมการที่ (7)

$$LV = \lambda V \quad (7)$$

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

จากนั้นทำการหาเวกเตอร์เฉพาะของเมตริกซ์ C ดังสมการที่ (8)

$$U = AV \quad (8)$$

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

เนื่องจากเวกเตอร์เฉพาะที่สอดคล้องกับค่าเฉพาะที่มากที่สุดจะมีนัยสำคัญทางข้อมูลมากที่สุด ดังนั้นเราสามารถเลือกเวกเตอร์เฉพาะที่มากที่สุดเพียง d ตัว ดังนั้นเราสามารถลดขนาดมิติของเวกเตอร์คุณลักษณะลงได้อีกดังสมการที่ (9)

$$\Omega_i = W^T \Phi_i \quad (9)$$

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

กำหนดให้ Y คือเวกเตอร์ของภาพที่ต้องการทดสอบ เราสามารถฉายเวกเตอร์ Y ไปบนปริภูมิคุณลักษณะได้ดังสมการที่

$$\omega = W^T (Y - \Psi) \quad (10)$$

โดย ω คือเวกเตอร์คุณลักษณะของภาพที่ต้องการทดสอบซึ่งอยู่ในปริภูมิเดียวกับกับเวกเตอร์ Ω ดังนั้นเราสามารถทราบได้ว่าเวกเตอร์ ω อยู่ใกล้กับเวกเตอร์ Ω ได้ด้วยการหารระยะยุคลิดที่สั้นที่สุดดังสมการที่ (11)

$$\varepsilon = \min_i (\|\Omega - \Omega_i\|^2) \quad (11)$$

โดย ε 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 ส่งผลทำให้ขนาดของเมตริกซ์ความแปรปรวนร่วมดังกล่าวมีขนาดใหญ่ทำให้การคำนวณหาค่าจะทำได้ยากและต้องใช้เวลาและหน่วยความจำจำนวนมาก และยังทำให้เกิดปัญหา Small Simple Size (SSS) หรือ Singularity ซึ่งจะเกิดขึ้นเป็นจำนวนตัวอย่างที่ใช้ฝึกฝนน้อยกว่าขนาดของเมตริกซ์ความแปรปรวนร่วมมาก ๆ เพื่อแก้ไขปัญหาดังกล่าวเหล่านี้ 2DPCA ถูกเสนอขึ้นโดยนิยามเมตริกซ์ความแปรปรวนร่วมใหม่ เรียกว่า เมตริกซ์ความแปรปรวนร่วมเชิงภาพ (Image covariance matrix) ซึ่งคำนวณได้โดยตรงจากภาพแบบ 2 มิติ โดยที่ไม่ต้องทำ Vectorization ในขั้นตอนแรก ด้วยหลักการนี้ เมตริกซ์ความแปรปรวนร่วมเชิงภาพจึงมีขนาดเล็กกว่าเดิมมาก คือ มีขนาดเท่ากับความกว้างของรูปเดี่ยวเท่านั้น

ระเบียบวิธี 2DPCA มีขั้นตอนดังนี้

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

$$y = Ax \quad (12)$$

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

$$J(X) = tr(S_x) \quad (13)$$

โดย S_x คือ เมตริกซ์ความแปรปรวนร่วม

$$S_x = E[(y - E_y)(y - E_y)^T] \quad (14)$$

ดังนั้นจากสมการที่ (12) และ (14) Trace ของเมตริกซ์ความแปรปรวนร่วมคือ

$$\begin{aligned}
tr(S_x) &= tr\{E[(y - Ey)(y - Ey)^T]\} \\
&= tr\{E[(A - EA)xx^T(A - EA)^T]\} \\
&= tr\{E[x^T(A - EA)^T(A - EA)x]\} \\
&= tr\{x^T E[(A - EA)^T(A - EA)]x\} \\
&= tr\{x^T Gx\}
\end{aligned} \tag{15}$$

กำหนดให้

$$G = E[(A - EA)^T(A - EA)] \tag{16}$$

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

$$G = \frac{1}{M} \sum_{k=1}^M (A_k - \bar{A})^T (A_k - \bar{A}) \tag{17}$$

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

$$Y = AX \tag{18}$$

โดย $Y = [y_1 \ y_2 \ \dots \ y_d]$ คือ เมตริกซ์คุณลักษณะมิติขนาด $m \times d$ ของ A และ $X = [x_1 \ x_2 \ \dots \ x_d]$ เป็นเมตริกซ์การฉายมิติขนาด $n \times d$

กำหนดให้ A_i เป็นภาพที่ต้องการทดสอบ ดังนั้นเราสามารถหาเมตริกซ์คุณลักษณะได้โดยการฉายดังสมการที่ (19)

$$B = A_i X \tag{19}$$

เราสามารถวัดระยะห่างระหว่างเมตริกซ์ของภาพทดสอบและภาพที่ใช้ฝึกฝนได้โดยใช้สมการที่ (20)

$$d(\mathbf{B}, \mathbf{Y}_i) = \sum_{k=1}^d \|\mathbf{B}^{(k)} - \mathbf{Y}_i^{(k)}\|_2 \quad (20)$$

ดังนั้นหาก

$$i = \operatorname{argmin}(d(\mathbf{B}, \mathbf{Y}_i)) \quad (21)$$

เราจะกล่าวได้ว่า \mathbf{A}_i ควรอยู่ในกลุ่มที่ i



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Bilateral-projection based 2DPCA (B2DPCA)

B2DPCA เป็น 2DPCA ที่ใช้การฉายทั้งสองข้าง คือ ตามแนวแถวและหลัก โดยคำนวณเมตริกซ์การฉายสองเมตริกซ์ดังนี้

$$\mathbf{B} = \mathbf{Z}^T \mathbf{A} \mathbf{X} \quad (22)$$

โดย \mathbf{X} คือเมตริกซ์การฉายเดิมใน 2DPCA ดังสมการที่ (18) และ \mathbf{Z} คือเมตริกซ์การฉายทางซ้ายซึ่งต้องกำหนดเมตริกซ์ความแปรปรวนร่วมเชิงภาพขึ้นใหม่ตามแนวหลักดังนี้

$$\mathbf{H} = \frac{1}{M} \sum_{k=1}^M (\mathbf{A}_k - \bar{\mathbf{A}})(\mathbf{A}_k - \bar{\mathbf{A}})^T \quad (23)$$

และหาด้วยวิธีการเดียวกันกับ \mathbf{X} หลังจากการฉายภาพ \mathbf{A} จะได้เมตริกซ์คุณลักษณะที่มีขนาดเล็กลงจากเดิม $m \times d$ เหลือเพียง $q \times d$ ซึ่ง q คือจำนวนค่าเจาะจงที่เลือกใช้ภายใน \mathbf{Z} ดังนั้น $q \leq m$ จาก การทดลองใน [17] และ [18] พบว่าใช้เวลาในการรู้จำลดลงและอัตราการรู้จำสูงขึ้น สำหรับวิธีการวัดระยะห่างนั้นทำได้โดยสมการที่ (24)

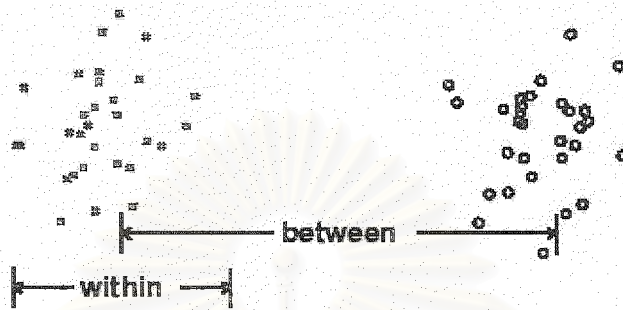
$$d(\mathbf{B}_i, \mathbf{B}_j) = \sqrt{\sum_{m=1}^q \sum_{n=1}^d (\mathbf{b}_{(m,n)_i} - \mathbf{b}_{(m,n)_j})^2} \quad (24)$$

ซึ่งในทำนองเดียวกับการรู้จำใน 2DPCA และ PCA ระยะห่างของตัวอย่างทดสอบใกล้กับตัวอย่างฝึกฝนได้ที่สุดเราจะจัดตัวอย่างทดสอบนั้นอยู่ในกลุ่มเดียวกับตัวอย่างฝึกฝนนั้น

Linear Discriminant Analysis (LDA)

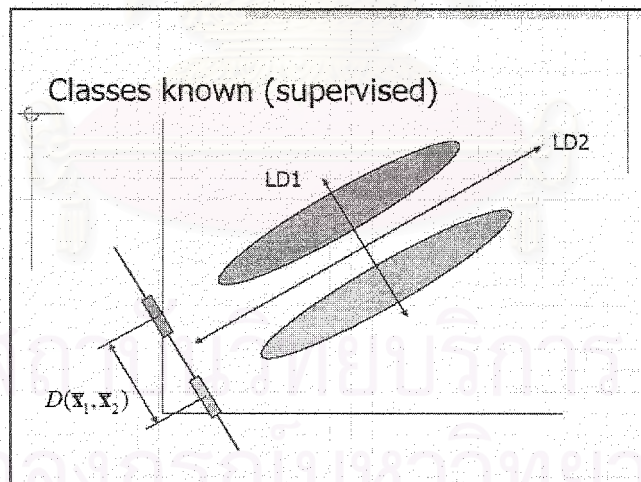
พิจารณาการกระจายตัวของข้อมูลระหว่างข้อมูล 2 กลุ่ม ดังรูปที่ 2 [2]

Good class separation



รูปที่ 2 ความสัมพันธ์ระหว่าง (between) และภายใน (within) กลุ่ม

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



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

กำหนดให้ S_w เป็นเมตริกซ์การกระจายภายในกลุ่ม และ S_b เป็นเมตริกซ์การกระจายระหว่างกลุ่ม โดยบรรทัดฐาน LDA ดังนี้

$$J(\mathbf{w}) = \max \left(\frac{S_b}{S_w} \right) \quad (25)$$

โดยเวกเตอร์ \mathbf{w} เป็นเวกเตอร์เจาะจงของเมตริกซ์ $S_w^{-1}S_b$ เนื่องจากเวกเตอร์เจาะจงเพียงเวกเตอร์เดียวไม่เพียงพอในการรู้จำดังนั้นกำหนดให้

$$\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{w}_q] \quad (26)$$

โดย \mathbf{W} คือเมตริกซ์การฉายที่มีแต่ละหลักเป็นเวกเตอร์การฉาย \mathbf{w} จำนวน q เวกเตอร์ที่สอดคล้องกับค่าเจาะจงของเมตริกซ์ $S_w^{-1}S_b$ ที่มากที่สุด q ค่า

หลังจากการฝึกฝนแล้วเราจะได้เมตริกซ์การฉาย \mathbf{W} ในลักษณะเดียวกับวิธี PCA การฉายหรือการสกัดคุณลักษณะสามารถทำได้ดังสมการที่ (10) เช่นเดียวกัน และการวัดระยะห่างระหว่างคู่ของเวกเตอร์คุณลักษณะใด ๆ สามารถทำได้เช่นเดียวกับ PCA ดังสมการที่ (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

$$\mathbf{v} = \mathbf{A}\mathbf{z} \quad (27)$$

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

$$J(\mathbf{z}) = \frac{\text{tr}(\mathbf{S}_b)}{\text{tr}(\mathbf{S}_w)} \quad (28)$$

โดย \mathbf{S}_b และ \mathbf{S}_w สามารถหาได้จากสมการที่ (29) และ (30) ตามลำดับ

$$\mathbf{S}_b = \sum_{i=1}^K \text{Pr}(\omega_i) E[(E[\mathbf{v} | \omega = \omega_i] - E\mathbf{v})(E[\mathbf{v} | \omega = \omega_i] - E\mathbf{v})^T] \quad (29)$$

$$\mathbf{S}_w = \sum_{i=1}^K \text{Pr}(\omega_i) E[(\mathbf{v} - E\mathbf{v})(\mathbf{v} - E\mathbf{v})^T | \omega = \omega_i] \quad (30)$$

โดย $\text{Pr}(\omega_i)$ คือ ความน่าจะเป็นของกลุ่ม ω_i และ K คือ จำนวนกลุ่มทั้งหมดที่ใช้ฝึกฝน จากสมการที่ (28), (29) และ (30) จะได้

$$\begin{aligned}
J(\mathbf{z}) &= \frac{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) E \left[(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A}) \mathbf{z} \mathbf{z}^T (E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A})^T \right) \right]}{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) E \left[(\mathbf{A} - E\mathbf{A}) \mathbf{z} \mathbf{z}^T (\mathbf{A} - E\mathbf{A})^T | \omega = \omega_i \right] \right)} \\
&= \frac{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) \mathbf{z}^T E \left[(E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A})^T (E[\mathbf{A} | \omega = \omega_i] - E\mathbf{A}) \right] \mathbf{z} \right)}{\text{tr} \left(\sum_{i=1}^K \text{Pr}(\omega_i) \mathbf{z}^T E \left[(\mathbf{A} - E\mathbf{A})^T (\mathbf{A} - E\mathbf{A}) | \omega = \omega_i \right] \mathbf{z} \right)} \\
&= \frac{\text{tr}(\mathbf{z}^T \tilde{\mathbf{S}}_b \mathbf{z})}{\text{tr}(\mathbf{z}^T \tilde{\mathbf{S}}_w \mathbf{z})} \tag{31}
\end{aligned}$$

โดย

$$\tilde{\mathbf{S}}_b = \sum_{i=1}^K \frac{n_i}{K} (\bar{\mathbf{A}}_i - \bar{\mathbf{A}})^T (\bar{\mathbf{A}}_i - \bar{\mathbf{A}}) \tag{32}$$

$$\tilde{\mathbf{S}}_w = \sum_{i=1}^K \frac{n_i}{K} \sum_{\mathbf{A}_k \in \omega_i} (\mathbf{A}_k - \bar{\mathbf{A}}_i)^T (\mathbf{A}_k - \bar{\mathbf{A}}_i) \tag{33}$$

ซึ่ง \mathbf{z} สามารถหาได้จากเวกเตอร์เจาะจงของเมตริกซ์ $\tilde{\mathbf{S}}_w^{-1} \tilde{\mathbf{S}}_b$ ดังสมการที่ (34)

$$\tilde{\mathbf{S}}_b \mathbf{z} = \lambda \tilde{\mathbf{S}}_w \mathbf{z} \tag{34}$$

โดย λ คือค่าเจาะจงของเมตริกซ์ $\tilde{\mathbf{S}}_w^{-1} \tilde{\mathbf{S}}_b$ เช่นเดียวกับวิธี PCA เวกเตอร์เจาะจงเพียงเวกเตอร์เดียวไม่เพียงพอในการรู้จำ จึงต้องใช้เวกเตอร์เจาะจงจำนวนหนึ่งที่สุดคดคล้องกับค่าเจาะจงที่มากที่สุด 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

$$\mathbf{V} = \mathbf{AZ} \tag{35}$$

ซึ่งการวัดระยะห่างทำได้แบบเดียวกับ 2DPCA

Two-Dimensional Linear Discriminant Analysis of Principal Component Vectors

เนื่องจาก 2DPCA เป็นการฉายภาพตัวอย่างไปบนปริภูมิคุณลักษณะ โดยการทำให้เมตริกซ์ความแปรปรวนร่วมของข้อมูลทั้งหมดมากที่สุดเท่านั้น ซึ่งเป็นวิธีแบบ Unsupervised จึงไม่มีข้อมูลเกี่ยวกับกลุ่มอยู่ อาจไม่เพียงพอในการใช้แยกแยะกลุ่ม ในขณะที่เดียวกัน 2DLDA เป็นวิธีแบบ Supervised ซึ่งสามารถทำให้การกระจายตัวของข้อมูลระหว่างกลุ่มมีค่ามาก แต่ภายในกลุ่มมีค่าน้อย อย่างไรก็ตาม 2DLDA เพียงวิธีเดียวจะใช้ข้อมูลทั้งหมดเพื่อการแยกแยะกลุ่ม ในบางครั้งข้อมูลดังกล่าวไม่เหมาะสมที่จะใช้ในการแยกแยะข้อมูล เช่น บริเวณพื้นหลังของภาพ เป็นต้น แต่ในระบบที่เราเสนอขึ้นนี้ใช้ 2DPCA เพื่อลดมิติของข้อมูลที่ไม่สำคัญลงก่อน จากนั้นนำไปเป็นเมตริกซ์คุณลักษณะป้อนเข้าให้กับ 2DLDA ทำให้ประสิทธิภาพการรู้จำเพิ่มขึ้น จากสมการที่ (32) และ (33) จะเปลี่ยนเป็นสมการที่ (36) และ (37) ตามลำดับ

$$\tilde{S}_b^Y = \sum_{i=1}^K \frac{n_i}{K} (\bar{Y}_i - \bar{Y})^T (\bar{Y}_i - \bar{Y}) \quad (36)$$

$$\tilde{S}_w^Y = \sum_{i=1}^K \frac{n_i}{K} \sum_{Y_k \in \omega_i} (Y_k - \bar{Y}_i)^T (Y_k - \bar{Y}_i) \quad (37)$$

โดย

$$\bar{Y} = \frac{1}{M} \sum_{k=1}^M Y_k \quad (38)$$

ดังนั้นการแปลงเชิงเส้นใหม่ได้โดยเมตริกซ์การฉายเชิงเส้น $L = XZ$ โดย X และ Z เป็นเมตริกซ์การฉายจากสมการที่ (18) และ (35)

$$D = AL \quad (39)$$

โดย 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 G_k คือเมตริกซ์ความแปรปรวนร่วมเชิงภาพของกลุ่มที่ k โดยใช้ตัวอย่างภาพในกลุ่มที่ k มาคำนวณดังนี้

$$G_k = \frac{1}{M} \sum_{A_c \in \omega_k} (A_c - \bar{A}_k)^T (A_c - \bar{A}_k) \quad (40)$$

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

$$\mathfrak{H}_k^{2DPCA} = \{X_k, \bar{A}_k, d_k\} \quad (41)$$

ในทำนองเดียวกัน กรณี B2DPCA เมตริกซ์ความแปรปรวนร่วมเชิงภาพตามแนวหลักเพื่อคำนวณเมตริกซ์การฉายของกลุ่มที่ k Z_k ได้ดังนี้

$$H_k = \frac{1}{M} \sum_{A_c \in \omega_k} (A_c - \bar{A}_k)(A_c - \bar{A}_k)^T \quad (42)$$

ดังนั้นจำนวนทูเปิลของ B2DPCA จึงเพิ่มขึ้นเป็น 5 ทูเปิลดังนี้

$$\mathfrak{H}_k^{B2DPCA} = \{Z_k, X_k, \bar{A}_k, d_k, q_k\} \quad (43)$$

กำหนดให้ A_c เป็นตัวอย่างป้อนเข้าเราสามารถหาภาพฉายของ A_c บนปริภูมีย่อยหน้าเจาะจงของกลุ่มต่าง ๆ ทั้ง k ปริภูมิได้ดังนี้

$$U_k = \begin{cases} W_k X_k; 2DPCA \\ Z_k^T W_k X_k; B2DPCA \end{cases} \quad (44)$$

โดย U_k คือภาพฉายของ A_i บนปริภูมิย่อยที่ k และ $W_k = (A_i - \bar{A}_k)$ เนื่องจากการวัดระยะห่างของภาพฉายที่อยู่ต่างปริภูมิกันทำให้เราจึงต้องทำการสร้างคืนภาพกลับมาที่ปริภูมิเดิมดังนี้

$$W_k^r = \begin{cases} U_k X_k^T; 2DPCA \\ Z_k U_k X_k^T; B2DPCA \end{cases} \quad (45)$$

โดย W_k^r คือภาพที่ได้จากการสร้างคืนของ U_k ดังนั้นเราสามารถหาค่าความผิดพลาดในการสร้างคืนจากปริภูมิย่อยใด ๆ ได้ดังนี้

$$\varepsilon_k(W_k^r, A_i) = \sum_{m=1}^{m_{row}} \sum_{n=1}^{n_{col}} |w_{(m,n)_k}^r - a_{(m,n)_i}^r| \quad (46)$$

โดย ε_k คือ ค่าความผิดพลาดในการสร้างคืนจากปริภูมิย่อย k เราสามารถกล่าวได้ว่าค่าความผิดพลาดจากปริภูมิย่อยใด ให้ค่าความผิดพลาดน้อยที่สุด

$$\varepsilon_i = \min_{1 \leq k \leq K} (\varepsilon_k) \quad (47)$$

ตัวอย่างทดสอบ A_i ดังกล่าวจะเป็นกลุ่มที่ i เราจึงสามารถใช้ค่าความผิดพลาดนี้เป็นบรรทัดฐานในการรู้จำได้ อย่างไรก็ตาม ความซับซ้อนในการคำนวณจะมากขึ้น แต่ในขณะเดียวกันเราไม่จำเป็นต้องใช้หน่วยความจำที่ใช้เก็บเมตริกซ์คุณลักษณะของตัวอย่างฝึกฝนอีกต่อไป และจำนวนครั้งในการคำนวณระยะห่างจากเดิมเท่ากับจำนวนตัวอย่างที่ใช้ฝึกฝนทั้งหมด M ครั้งเหลือเพียงเท่ากับจำนวนกลุ่ม K ครั้ง โดยสรุปเป็นตารางเปรียบเทียบได้ดังตารางที่ 1

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

วิธีการ	หน่วยความจำที่ใช้		จำนวนการฉาย		จำนวนการคำนวณระยะห่าง
	เมตริกซ์ฉาย	เมตริกซ์คุณลักษณะ	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

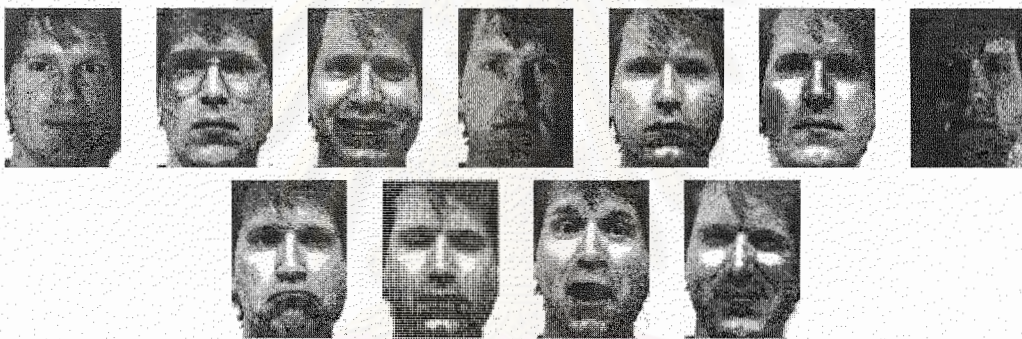
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

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 แต่มีการแสดงอารมณ์รวมทั้งทิศทางที่แสดงตกระทบต่าง ๆ กัน ดังนี้ แสดงตกระทบกลางใบหน้า (center-light) สวมแว่นตา (with glasses) อารมณ์สุข (happy) แสดงตกระทบทางซ้ายของภาพใบหน้า (left-light) ไม่สวมแว่น (without glasses) ปกติ (normal) แสดงตกระทบทางขวาของภาพใบหน้า (right-light) อารมณ์เศร้า (sad) หลับตา (sleepy) ตกใจ (surprised) และ ขยิบตา (wink) ดังแสดงในรูปที่ 4 แต่ละภาพถูกตัดบริเวณ ใบหน้าด้วยมือและกำหนดขนาดใหม่เป็น 100 x 80 ตารางพิกเซล



รูปที่ 4 ตัวอย่างภาพใบหน้าจากฐานข้อมูล Yale

● ฐานข้อมูล AR

ฐานข้อมูล AR ถูกสร้างขึ้นโดย Alex Martinez และ Robert Benavente ภายใน Computer Vision Center (CVC) ที่ U.A.B. ประกอบด้วยภาพสีจำนวนมากว่า 4,000 ภาพจาก 126 คนซึ่งเป็นผู้ชาย 70 คน และผู้หญิง 56 คน ภาพทั้งหมดเป็นภาพใบหน้าตรงที่แสดงอารมณ์ ทิศทางที่แสดงตกระทบ และ วัตถุบังบังใบหน้า (แว่นตากันแดด หรือผ้าผืนคอ) แตกต่างกัน โดยการเก็บข้อมูลของแต่ละบุคคลจะแบ่งเป็น 2 ชุด โดยแต่ละชุดมีระยะเวลาห่างกัน 2 สัปดาห์ ในการทดลองของเราจะพิจารณาเพียงกลุ่มละ 14 ภาพโดยไม่นับรวมภาพที่มีวัตถุบังบังใบหน้า จากนั้นตัด และกำหนดขนาดใหม่เป็น 100 x 80 จากนั้นเปลี่ยนเป็นภาพสีเทา 256 ระดับดังแสดงในรูปที่ 5



รูปที่ 5 ตัวอย่างภาพใบหน้าจากฐานข้อมูล AR

● ฐานข้อมูล ORL

ฐานข้อมูล ORL ประกอบด้วยภาพใบหน้าจาก 40 คน โดยแต่ละคนเก็บข้อมูลไว้เป็นจำนวน 10 ภาพในมุมมองที่ต่างกัน 20 องศา รวมถึงการแสดงอารมณ์ การลืมตาและหลับตาที่ต่าง ๆ กัน และบางภาพของบางคนเก็บข้อมูลที่เวลาต่างกัน ภาพทั้งหมดมีขนาด 112 x 92 ดังแสดงในรูปที่ 6

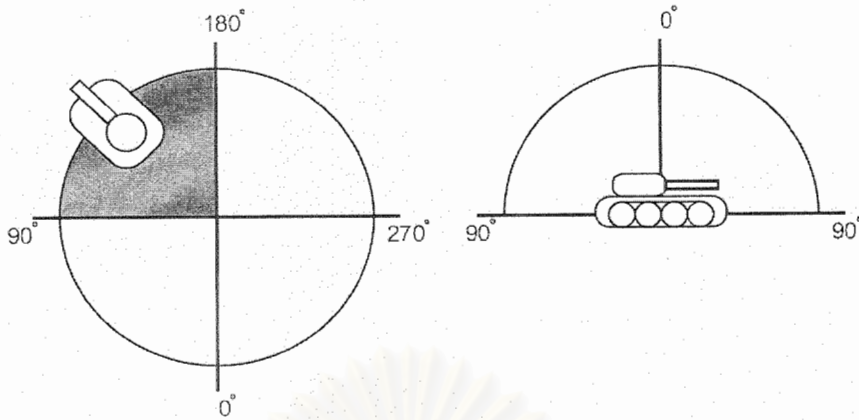


รูปที่ 6 ตัวอย่างภาพใบหน้าจากฐานข้อมูล ORL

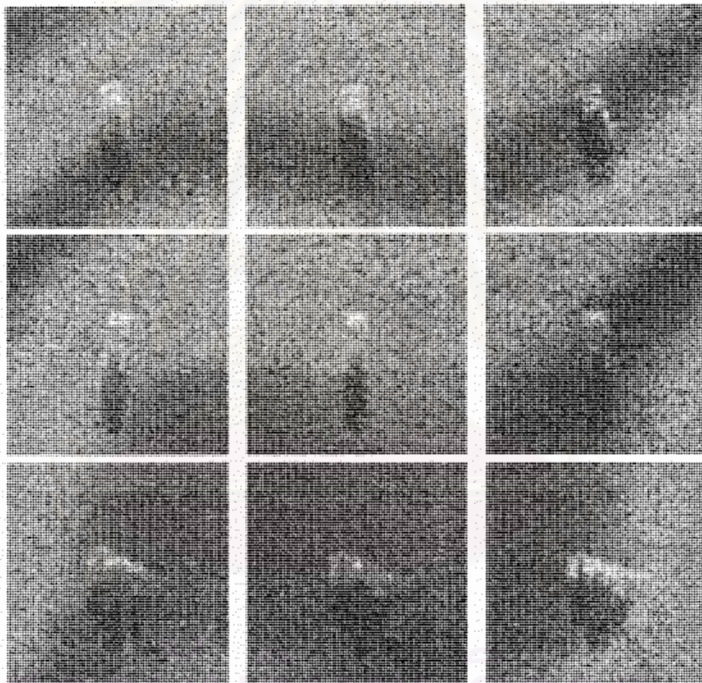
● ฐานข้อมูล MSTAR

ฐานข้อมูล MSTAR ประกอบด้วยชุดข้อมูล Synthetic Aperture Radar (SAR) ความละเอียดสูง รวบรวมโดยห้องวิจัยโปรแกรม DARPA/Wright laboratory Moving and Stationary Target Acquisition and Recognition (MSTAR) ชุดข้อมูลประกอบด้วยภาพ SAR ขนาด 128 x 128 ของรถทางทหาร 3 แบบ ได้แก่ BMP2 armored personal carriers (APCs), BTR70 APCs และ รถถัง T72 ซึ่งรูป SAR ดังกล่าวขึ้นกับการวางตัวของรถเป้าหมายที่ต่างกัน ดังนั้นการตรวจรู้ทิศทางการวางตัวของรถก่อน (มุม Aspect) ดังรูปที่ 7 สามารถเพิ่มประสิทธิภาพการรู้จำได้มากขึ้นซึ่งในการทดลองของเราจะทำได้ด้วยมือ ตัวอย่างรูปภาพจากฐานข้อมูล MSTAR แสดงดังรูปที่ 8 โดยภาพทั้งหมดถูกตัดเอาเฉพาะบริเวณกลางภาพขนาด 80 x 80 ตารางพิกเซล

ตารางที่ 2 แสดงชุดฝึกฝน และตารางที่ 3 แสดงชุดทดสอบ



รูปที่ 7 ภาพซ้ายคือมุมมอง Aspect และภาพขวาคือมุมมอง Depression



รูปที่ 8 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 ซึ่งวิธี 2DPCA+2DLDA มีมิติของเมตริกซ์คุณลักษณะขนาดเล็กที่สุด และใช้เวลาทั้งฝึกฝนและทดสอบน้อยที่สุดบนฐานข้อมูล Yale และยังให้ความถูกต้องในการรู้จำมากที่สุด ในขณะที่เดียวกัน 2DPCA+2DLDA ที่ทดสอบบนฐานข้อมูล ORL ก็ให้ความถูกต้องในการรู้จำที่สูงสุดเช่นกัน โดยมีมิติของเมตริกซ์คุณลักษณะเท่ากับของ 2DPCA แต่ใช้เวลามากกว่า 2DPCA เพียงไม่มาก อัตราการรู้จำของ 2DLDA บนฐานข้อมูล ORL มีค่าน้อยลงตามสมมติฐานที่ว่า 2DLDA จะนำข้อมูลที่ไม่มียัยในการแยกแยะกลุ่มมาใช้ในการแยกแยะ แต่เมื่อนำ 2DPCA มาเป็นตัวสกัดคุณลักษณะก่อนป้อนเข้า 2DLDA พบว่าสามารถเพิ่มอัตราการรู้จำได้ดียิ่งขึ้นจริง

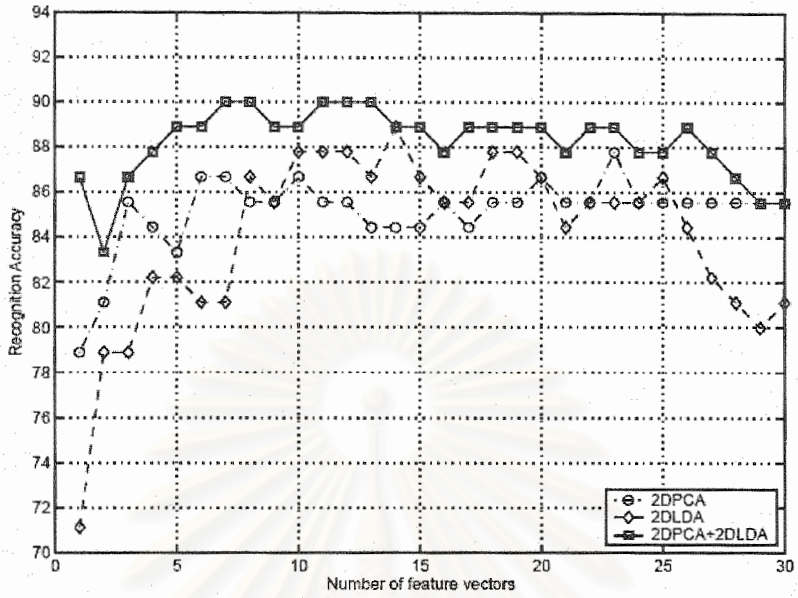
รูปที่ 9 และรูปที่ 10 แสดงความถูกต้องในการรู้จำเมื่อมีการเปลี่ยนแปลงจำนวนค่าเจาะจงที่ใช้ในการหาเวกเตอร์เจาะจงหรือ d แต่เนื่องจากอัตราการรู้จำของ 2DPCA+2DLDA ขึ้นกับจำนวนค่าเจาะจงที่ใช้ทั้ง 2DPCA และ 2DLDA ดังนั้นเส้นกราฟที่แสดงของ 2DPCA+2DLDA จึงเป็นเส้นกราฟที่แสดงจุดสูงสุดของกราฟในรูปที่ 11 และรูปที่ 12

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

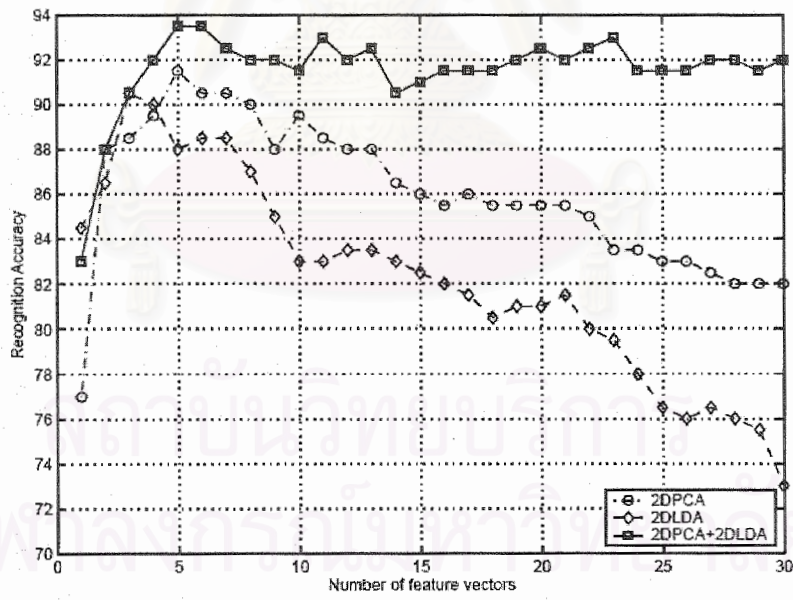
วิธีการ	ความถูกต้อง (%)	d	q	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	87.78	23	-	100×23	0.17 (1)	3.14 (1)	3.31 (1)
2DLDA	88.89	-	14	100×14	0.16 (0.94)	1.94 (0.62)	2.10 (0.63)
2DPCA+2DLDA	90.00	21	7	100×7	0.24 (1.14)	1.01 (0.32)	1.25 (0.38)

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

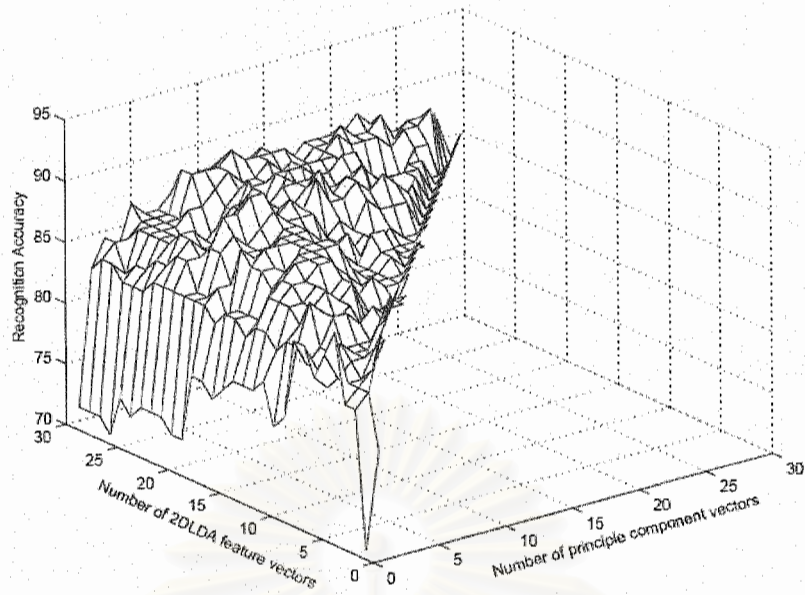
วิธีการ	ความถูกต้อง (%)	d	q	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	91.50	5	-	112×5	0.33 (1)	4.33 (1)	4.69 (1)
2DLDA	90.50	-	3	112×3	0.50 (1.51)	2.75 (0.64)	3.28 (0.7)
2DPCA+2DLDA	93.50	14	5	112×5	0.64 (1.94)	4.33 (1)	5.22 (1.07)



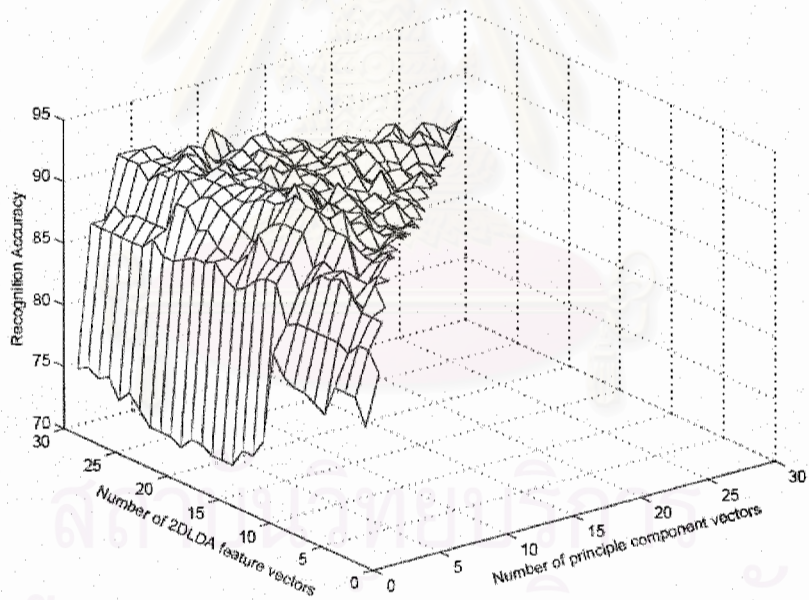
รูปที่ 9 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล Yale



รูปที่ 10 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล ORL



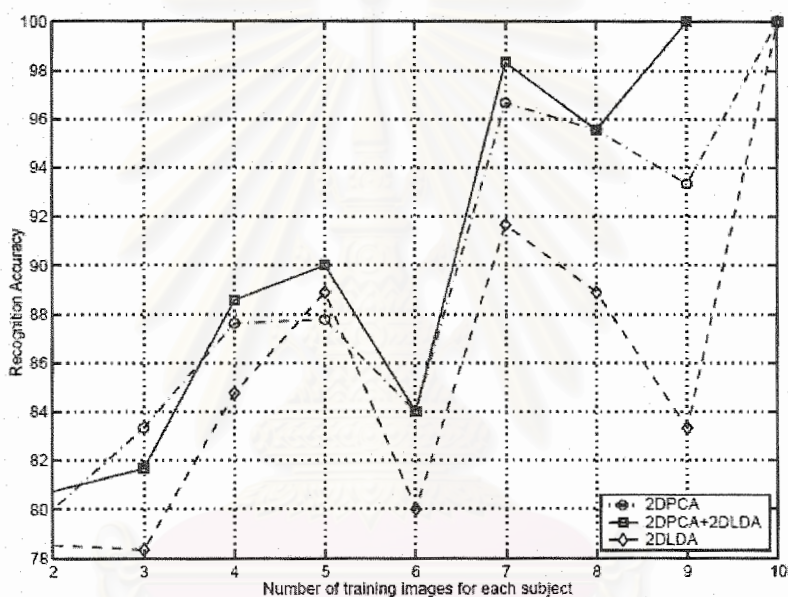
รูปที่ 11 ความถูกต้องในการรู้จำของวิธี 2DPCA+2DLDA เทียบกับจำนวนเวกเตอร์คุณลักษณะของ 2DLDA และ 2DPCA บนฐานข้อมูล Yale



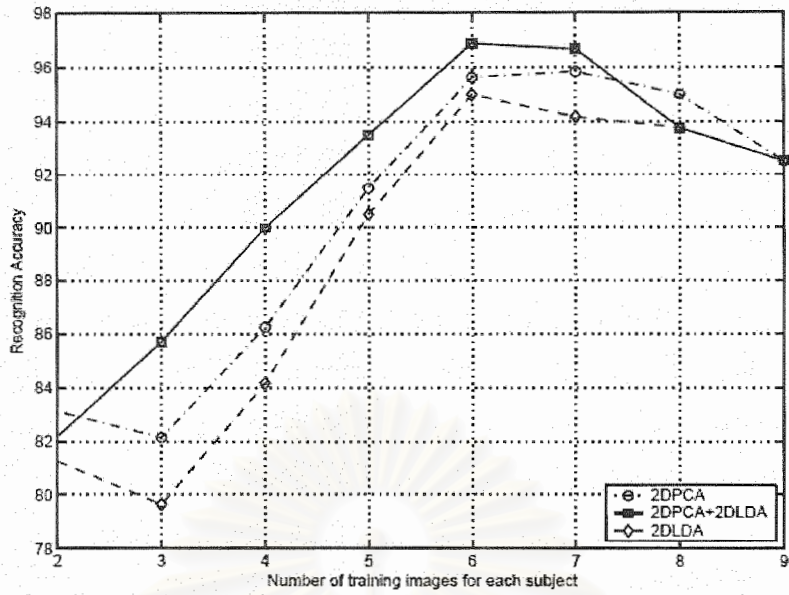
รูปที่ 12 ความถูกต้องในการรู้จำของวิธี 2DPCA+2DLDA เทียบกับจำนวนเวกเตอร์คุณลักษณะของ 2DLDA และ 2DPCA บนฐานข้อมูล ORL

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

บนฐานข้อมูล Yale ทำการเปลี่ยนแปลงจำนวนตัวอย่างฝึกฝนตั้งแต่ 2 ถึง 10 โดยภาพทดสอบที่ใช้คือภาพที่ไม่ได้ใช้ในการฝึกฝน ในขณะที่เดียวกันบนฐานข้อมูล ORL ทำการเปลี่ยนแปลงจำนวนตัวอย่างฝึกฝนตั้งแต่ 2 ถึง 9 โดยภาพทดสอบที่ใช้คือภาพที่ไม่ได้ใช้ในการฝึกฝน ได้ดังรูปที่ 13 และรูปที่ 14 ตามลำดับ



รูปที่ 13 ความถูกต้องในการรู้จำของวิธี 2DPCA 2DLDA และ 2DPCA+2DLDA เมื่อมีการเปลี่ยนแปลงจำนวนใบหน้าที่ใช้ฝึกฝนในแต่ละกลุ่มบนฐานข้อมูล Yale



รูปที่ 14 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 คือจำนวนค่าเจาะจงที่ใช้ในแนวแถว และหลัก ตามลำดับ โดยใช้ฐานข้อมูล MSTAR ใช้ข้อมูลเชิงขนาดมาสร้างเป็นภาพแล้วทำการตรวจรู้ตำแหน่งการวางตัวของรถด้วยมือ 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

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)

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

ระบบรู้จำที่ใช้ปริภูมิย่อยเจาะจงเชิงภาพที่ใช้ 2DPCA และ B2DPCA สามารถเพิ่มอัตราการเรียนรู้จากรู้อบบฐานข้อมูล Yale และ AR ได้ แต่กลับทำให้อัตราการเรียนรู้จบบนฐานข้อมูล MSTAR มีค่าน้อยลง ทั้งนี้เนื่องมาจากจำนวนตัวอย่างที่ใช้ในการฝึกฝนในแต่ละกลุ่มมีมากอยู่แล้วทำให้การประมาณค่าของเมตริกซ์ความแปรปรวนร่วมของ 2DPCA และ B2DPCA ทำได้ดีกว่า แต่วิธีการที่ใช้ปริภูมิย่อยเจาะจงเชิงภาพเหมาะกับข้อมูลที่มีจำนวนตัวอย่างที่ใช้ฝึกฝนน้อย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3. ส่วนที่จะดำเนินการต่อไป

3.1. ศึกษาการรู้จำใบหน้าเพิ่มเติม

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.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย