

การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA
FACE RECOGNITION USING WPS-CSLDA ALGORITHM

วีรรัตน์ ลีลาธรรม
WEERARAT LEEATHAM

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาคณะหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคม

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ศ.ศ. 2550

KMITL-2007-EN-M-010-061

การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA

FACE RECOGNITION USING WPS-CSLDA ALGORITHM



วีรรัตน์ ลีลาธรรม

WEERARAT LEELATHAM

เลขหมู่.....
เลขทะเบียน..... **76686**
วัน,เดือน,ปี..... - 6 S.ก. 2550

b. 118 50231
i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคม

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2550

KMITL-2007-EN-M-010-081

FACE RECOGNITION USING WPS-CSLDA ALGORITHM

WEERARAT LEELATHAM

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

2007

KMITL-2007-EN-M-010-081

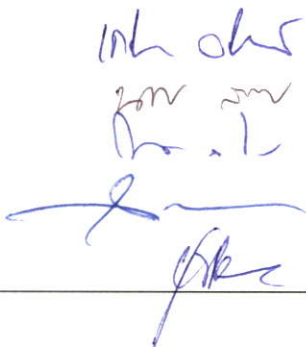
COPYRIGHT 2007

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA
Face Recognition Using WPS-CSLDA Algorithm
นักศึกษา นายวีรรัตน์ ลีสาธรรม
รหัสประจำตัว 48060973
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา วิศวกรรมโทรคมนาคม
อาจารย์ที่ปรึกษาวิทยานิพนธ์ รศ.ดร.ยุทธพงษ์ รังสรรค์เสรี

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ผศ.เกรียงไกร	วงศ์โรจนภรณ์	
ผศ.นภัทร	สระเอี่ยม	
รศ.ดร.โกสินทร์	จ่านงไทย	
ผศ.ดร.สุทธิชัย	นพนาถิพงษ์	
รศ.ดร.ยุทธพงษ์	รังสรรค์เสรี	

วัน / เดือน / ปี ที่สอบ 4 ตุลาคม 2550 เวลา 11.30-13.30 น.

สถานที่สอบ ณ อาคาร A ห้องประชุม 2 ชั้น 3

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.จารุวัตร เจริญสุข)
คณบดีบัณฑิตวิทยาลัย

วันที่.....๕.....เดือน.....๓๑๑๒.....พ.ศ.....๒๕๕๐.....

หัวข้อวิทยานิพนธ์	การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA
นักศึกษา	นายวีรวัฒน์ สีลาธรรม
รหัสนักศึกษา	48060973
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมโทรคมนาคม
พ. ศ.	2550
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ. ดร. ยุทธพงษ์ รังสรรค์เสรี

บทคัดย่อ

วิทยานิพนธ์นี้ได้นำเสนอการพัฒนาอัลกอริทึมใหม่บนพื้นฐานของ LDA มีชื่อว่า Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) โดยที่ WPS-CSLDA สามารถแก้ปัญหาการปนกันของข้อมูล โดยได้นำ Weighted Pairwise Scatter มาคำนวณการแบ่งแยกข้อมูลระหว่างกลุ่มที่มีการปนกันเพื่อให้ได้ลักษณะการแบ่งแยกที่มีประสิทธิภาพสูงกว่าวิธีที่ผ่านมา ความแตกต่างระหว่าง วิธีการที่นำเสนอ กับ วิธีอื่น คือ WPS-CSLDA จะนำ Weighted Pairwise Scatter มาใช้ในการคำนวณ อัลกอริทึมที่พัฒนาแล้วได้นำมาประยุกต์กับงานด้านการรู้จำใบหน้า จากการทดลองจะแสดงให้เห็นว่า WPS-CSLDA มีประสิทธิภาพในการคำนวณและการรู้จำที่ดี

Thesis Title	Face Recognition Using WPS-CSLDA Algorithm
Student	Mr. Weerarat Leelatham
Student ID.	48060973
Degree	Master of Engineering
Program	Telecommunications Engineering
Year	2007
Thesis Advisor	Assc. Prof. Dr.Yuttapong Rangsanseri

ABSTRACT

In this thesis, we propose a novel algorithm based on LDA, Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) for solving the confusable data in face recognition application. WPS-CSLDA is used to compute between class scatter discrimination for higher efficiency. The key difference between WPS-CSLDA and classical LDA is using Weighted Pairwise Scatter for the computation. This developed algorithm is applied to the face recognition works. The experimental results indicated that the recognition rate of our algorithm is better than that of the conventional method.

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จได้อย่างดีจากการให้โอกาส คำแนะนำ และคำปรึกษาจาก รศ. ดร. ยุทธพงษ์ ริงสรรค์เสรี ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ข้าพเจ้ารู้สึกซาบซึ้งในความอนุเคราะห์ของท่านอาจารย์ และขอขอบพระคุณเป็นอย่างสูงที่มอบประสบการณ์หลายๆ อย่าง ให้แก่ข้าพเจ้า

ขอกราบขอบพระคุณคณาจารย์ภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกๆ ท่านที่ได้ประสิทธิ์ ประสาทวิชาแก่ข้าพเจ้า

ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่มีส่วนช่วยให้คำแนะนำปรึกษาต่างๆ และคอยให้กำลังใจเสมอมา

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และคอยสนับสนุนในทุกๆ เรื่องมาโดยตลอด จนทำให้ข้าพเจ้าสามารถทำวิทยานิพนธ์เล่มนี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์อันพึงมาจากวิทยานิพนธ์เล่มนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

วีรรัตน์ ติลาธรรม

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VIII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของวิทยานิพนธ์.....	1
1.2 จุดมุ่งหมายและวัตถุประสงค์ของวิทยานิพนธ์.....	1
1.3 สมมุติฐานของการศึกษาและทฤษฎีแนวความคิดที่ใช้ในการวิจัย.....	1
1.4 ขอบเขตของวิทยานิพนธ์.....	2
1.5 รายละเอียดของวิทยานิพนธ์.....	3
บทที่ 2 การประมวลผลภาพและการรู้จำใบหน้า.....	4
2.1 ความซับซ้อนและความยุ่งยากในการรู้จำใบหน้าด้วยคอมพิวเตอร์.....	4
2.2 การนำระบบรู้จำใบหน้าไปใช้จริงในปัจจุบัน.....	5
2.2.1 ฝ้าคลุมคคผลแปลกหน้า.....	5
2.2.2 หนังสือเดินทางอิเล็กทรอนิกส์ (E-Passport).....	5
2.2.3 ล็อกหน้าจอ (Screen Lock).....	5
2.2.4 การค้นหาในฐานข้อมูล.....	5
2.3 สิ่งที่เปลี่ยนแปลงได้บนภาพใบหน้าที่มีผลกระทบในการรู้จำใบหน้า.....	6
2.3.1 สภาพแวดล้อม.....	6
2.3.2 การแสดงออกทางอารมณ์บนใบหน้า.....	6
2.3.3 อื่น ๆ.....	6
2.4 วิธีในการรู้จำภาพใบหน้าที่มีการวิจัยกันในปัจจุบัน.....	7
2.4.1 ลักษณะทางเรขาคณิต (Geometrical Features).....	7
2.4.2 การเปรียบเทียบกราฟ (Graph Matching).....	8
2.4.3 วิธีการทางสถิติ (Statistical).....	8

สารบัญ (ต่อ)

	หน้า
2.4.3.1 การวิเคราะห์องค์ประกอบหลัก.....	9
2.4.3.2 การวิเคราะห์การแบ่งแยกเชิงเส้น.....	10
2.4.4 โครงข่ายประสาทเทียม (Neural network).....	11
2.4.5 วิธีอื่น ๆ (Other Methods).....	12
บทที่ 3 ทฤษฎีพื้นฐานที่เกี่ยวข้อง.....	13
3.1 การวิเคราะห์องค์ประกอบหลัก (PCA).....	13
3.1.1 แนวความคิดของการวิเคราะห์องค์ประกอบหลัก.....	13
3.1.2 หลักการคำนวณของการวิเคราะห์องค์ประกอบหลัก.....	15
3.1.3 การหาพีเจอร์ของภาพ.....	17
3.1.4 การสร้างภาพใหม่.....	17
3.2 การวิเคราะห์การแบ่งแยกเชิงเส้น (LDA).....	18
3.2.1 แนวความคิดของการวิเคราะห์การแบ่งแยกเชิงเส้น.....	18
3.2.2 หลักการคำนวณของการวิเคราะห์การแบ่งแยกเชิงเส้น.....	25
3.2.3 การหาพีเจอร์ของภาพ.....	26
3.2.4 การสร้างภาพใหม่.....	26
3.3 Two – Dimensional Principal Component Analysis (2DPCA).....	27
3.3.1 แนวความคิดของวิธี 2DPCA.....	27
3.3.2 การหาพีเจอร์ของภาพ.....	29
3.3.3 การสร้างภาพใหม่.....	29
3.3.4 หลักการคำนวณของวิธี 2DPCA.....	30
3.4 Two – Dimensional Linear Discriminant Analysis (2DLDA).....	31
3.4.1 แนวความคิดของวิธี 2DLDA.....	31
3.4.2 การหาพีเจอร์ของภาพ.....	33
3.4.3 การสร้างภาพใหม่.....	33
3.4.4 หลักการคำนวณของวิธี 2DLDA.....	34
3.5 Eigenimages.....	35
3.5.1 แนวความคิดของวิธี Eigenimages.....	36

สารบัญ (ต่อ)

	หน้า
3.5.2 การหาพีเจอร์ของภาพ.....	39
3.5.3 การสร้างภาพใหม่.....	39
3.5.4 หลักการคำนวณของวิธี Eigenimages	39
3.6 Coupled Subspace Linear Discriminant Analysis (CSLDA).....	41
3.6.1 แนวความคิดของวิธีการ CSLDA	41
3.6.2 การหาพีเจอร์ของภาพ.....	44
3.6.3 การสร้างภาพใหม่.....	45
3.6.4 หลักการคำนวณของวิธี CSLDA	45
บทที่ 4 วิธีการที่นำเสนอ (WPS-CSLDA).....	47
4.1 Weighted Pairwise Scatter	47
4.2 แนวคิดของ WPS-CSLDA.....	50
4.3 การหาพีเจอร์ของภาพ	54
4.4 การสร้างภาพใหม่.....	54
4.5 วิธีคำนวณ WPS-CSLDA.....	54
4.6 วิธีการจำแนกข้อมูลภาพใบหน้า (Images Classification)	56
บทที่ 5 การทดลองและผลการทดลอง	57
5.1 การทดลองบนฐานข้อมูลภาพใบหน้า ORL	57
5.1.1 การทดสอบการสร้างภาพใหม่ด้วยวิธี WPS-CSLDA บนฐานข้อมูล ORL.....	58
5.1.2 การทดลองเปรียบเทียบความถูกต้องในการรู้จำใบหน้าระหว่าง วิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages.....	58
5.1.3 การทดลองเปรียบเทียบความถูกต้องในการรู้จำใบหน้าระหว่าง วิธี WPS-CSLDA กับวิธีอื่น ๆ	63
5.2 การทดลองบนฐานข้อมูลภาพใบหน้า AR.....	66
5.2.1 การทดลองกับเงื่อนไขการถ่ายภาพต่างเวลา.....	68
5.2.2 การทดลองกับเงื่อนไขการแสดงออกทางอารมณ์บนใบหน้า.....	70
5.2.3 การทดลองภายใต้เงื่อนไขของความสว่างบนใบหน้า.....	71

สารบัญ (ต่อ)

	หน้า
5.3 การทดลองบนฐานข้อมูลภาพใบหน้า Yale.....	72
5.3.1 การทดลองภายใต้เงื่อนไขการเปลี่ยนแปลงของการแสดงออกบนใบหน้าและ ความสว่างบนใบหน้า.....	73
5.3.2 การทดลองความถูกต้องในการรู้จำวิธี WPS-CSLDA ภายใต้เงื่อนไขของ จำนวนตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอน	74
บทที่ 6 สรุปผลการวิจัยและแนวทางการพัฒนา.....	77
6.1 สรุปผลการวิจัย.....	77
6.2 แนวทางในการพัฒนา	78
เอกสารอ้างอิง	79
ภาคผนวก.....	82
ภาคผนวก ก.	83
ภาคผนวก ข.	84
ประวัติผู้เขียน.....	96

สารบัญตาราง

ตารางที่	หน้า
5.1 การเปรียบเทียบการรู้จำสูงสุดของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้ในการฝึกสอน.....	62
5.2 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าสูงสุดของวิธี 2DLDA, วิธี LDA, วิธี 2DPCA และวิธี PCA กับวิธี WPS-CSLDA.....	66
5.3 เปรียบเทียบวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้ตัวแปรของการถ่าย ภาพต่างเวลา เงื่อนไขการแสดงออกทางอารมณ์บนใบหน้า เงื่อนไขแสงบนใบหน้า.....	69
5.4 เปรียบเทียบร้อยละความถูกต้องในการรู้จำสูงสุดของวิธีทั้ง 3 วิธีบนฐานข้อมูล Yale.....	73
5.5 การเปรียบเทียบการรู้จำสูงสุดของวิธี WPS-CSLDA ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้ ในการฝึกสอน.....	75

สารบัญรูป

รูปที่	ที่
3.1 ลำดับขั้นการพัฒนาทฤษฎีการรู้จำใบหน้า PCA และ LDA.....	13
3.2 ตัวอย่างภาพที่เปลี่ยนจากเมตริกภาพเป็นเวกเตอร์ภาพ.....	15
3.3 ตัวอย่างภาพใบหน้าบุคคลหนึ่งทีสร้างใหม่ด้วยวิธี PCA จากฐานข้อมูล ORL.....	17
3.4 ผลของการเลือกค่าของ w ที่ใช้ในการฉายภาพ.....	19
3.5 การเลือกแกนการฉายจากระยะทางระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉาย 2 กลุ่ม.....	20
3.6 ตัวอย่างการฉายภาพที่มีการแบ่งแยกได้ดีของตัวอย่าง 2 กลุ่ม.....	21
3.7 ตัวอย่างภาพใบหน้าบุคคลหนึ่งทีสร้างใหม่ด้วยวิธี LDA จากฐานข้อมูล ORL.....	27
3.8 ตัวอย่างภาพใบหน้าทีสร้างใหม่ด้วยวิธี 2DPCA จากฐานข้อมูล ORL.....	30
3.9 ตัวอย่างภาพใบหน้าทีสร้างใหม่ด้วยวิธี 2LDAจากฐานข้อมูล ORL.....	34
4.1 ข้อมูลทีมีความแปรปรวนเท่ากันทีตำแหน่งต่างกัน.....	48
5.1 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล ORL.....	57
5.2 ภาพใบหน้าทีสร้างขึ้นใหม่ด้วยวิธี WPS-CSLDA.....	58
5.3 ความถูกต้องในการรู้จำใบหน้าของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages.....	60
5.4 ความถูกต้องในการรู้จำใบหน้าของวิธี WPS-CSLDA ภายใต้เงื่อนไขของตัวอย่างภาพ ใบหน้าทีใช้ในการฝึกสอน.....	61
5.5 ความถูกต้องในการรู้จำใบหน้าของวิธี CSLDA ภายใต้เงื่อนไขของตัวอย่างภาพใบหน้าที ใช้ในการฝึกสอน.....	61
5.6 ความถูกต้องในการรู้จำใบหน้าของวิธี Eigenimages ภายใต้เงื่อนไขของตัวอย่างภาพ ใบหน้าทีใช้ในการฝึกสอน.....	62
5.7 เปรียบเทียบความถูกต้องในการรู้จำใบหน้าในแนวทแยงของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages.....	63
5.8 ความถูกต้องในการรู้จำของวิธี 2DLDA กับวิธี PCA วิธี LDA และวิธี 2DPCA.....	65
5.9 ตัวอย่างภาพใบหน้า AR ของบุคคลคนหนึ่งทีถูกแบ่งเป็นสองกลุ่ม.....	67
5.10 ตัวอย่างภาพใบหน้าของ 1 บุคคลทีถูกเลือกมาใช้ในการทดลอง.....	68
5.11 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าในแนวทแยงของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้ตัวแปรของการถ่ายภาพต่างเวลา.....	69
5.12 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าแนวทแยงภายใต้เงื่อนไขการแสดงออก ทางอารมณ์บนใบหน้าของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages.....	70

สารบัญรูป (ต่อ)

รูปที่	ที่
5.13 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าแนวทแยงภายใต้เงื่อนไขของความสว่าง บนใบหน้าของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages	72
5.14 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล Yale.....	73
5.15 เปรียบเทียบความถูกต้องในการรู้จำใบหน้า Yale ในแนวทแยงของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages	74
5.16 เปรียบเทียบความถูกต้องในการรู้จำใบหน้า Yale ของวิธี WPS-CSLDA ภายใต้เงื่อนไข ของตัวอย่างภาพที่ใช้ในการฝึกสอน	76
ก. รูปแบบและขั้นตอนการคำนวณที่ใช้ในการทดลอง	83

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของวิทยานิพนธ์

ปัจจุบันนี้งานด้านการรู้จำใบหน้าเป็นเรื่องที่กำลังได้รับความสนใจอย่างมาก มีวิธีการรู้จำใบหน้าหลายวิธี ซึ่งหนึ่งในนั้นคือวิธีการทางสถิติ ซึ่งวิธีการนี้ได้ถูกนำมาประยุกต์ใช้งานด้านการรู้จำใบหน้าอย่างกว้างขวาง แต่วิธีการที่ใช้อยู่ในปัจจุบันยังให้ความถูกต้องในการรู้จำไม่ดีเท่าที่ควร จากการศึกษาพบว่า วิธีการคำนวณที่ใช้อยู่ในปัจจุบันยังสามารถที่จะพัฒนาให้มีประสิทธิภาพในการรู้จำใบหน้าเพิ่มขึ้นได้

1.2 จุดมุ่งหมายและวัตถุประสงค์ของวิทยานิพนธ์

1.2.1 เพื่อศึกษาทฤษฎีการรู้จำใบหน้า โดยเฉพาะหลักการการหาพีเจอร์์ของภาพบนพื้นฐานของการแปลงเชิงเส้น (LDA)

1.2.2 ศึกษาการนำ Weighted Pairwise Scatter มาใช้ในการคำนวณหาพีเจอร์์ของภาพบนพื้นฐานของการแปลงเชิงเส้น

1.2.3 วิเคราะห์และเปรียบเทียบความถูกต้องในการรู้จำใบหน้าวิธีการใหม่ที่วิทยานิพนธ์นี้นำเสนอ คือ Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) กับวิธีการที่มีมาก่อนหน้านี้

1.3 สมมุติฐานของการศึกษาและทฤษฎีแนวความคิดที่ใช้ในการวิจัย

วิธีการทางสถิติเป็นเครื่องมือที่ถูกนำมาใช้สำหรับการหาพีเจอร์์ (feature extraction) ของภาพ และมีหลายวิธีที่มีประสิทธิภาพสูงได้ถูกนำมาใช้ การวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ซึ่งเป็นวิธีที่สำคัญมากในการพัฒนาทางสถิติ แต่ว่าวิธี PCA และวิธี LDA จะใช้การคำนวณเฉพาะเวกเตอร์ 1 มิติ ดังนั้นเมื่อต้องการใช้วิธีนี้กับภาพจึงจำเป็นต้องแปลงเมตริกภาพให้เป็นเวกเตอร์ภาพ หลังจากนั้นจึงใช้พื้นฐานของเวกเตอร์ในการคำนวณหาเมตริกความแปรปรวนร่วม (covariance matrix) เพื่อคำนวณหาเวกเตอร์ที่เหมาะสมที่สุดที่ใช้ในการฉายภาพ วิธีการนี้ถูกเรียกว่า PCA [1] และ LDA [2] เมื่อนำวิธีการดังกล่าวมาใช้ในการงานด้านการรู้จำใบหน้าเป็นผลให้ได้กระบวนการรู้จำที่ดี แต่ว่าภาพใบหน้าที่ใช้ในกระบวนการรู้จำจะมีมิติมาก เช่น ภาพใบหน้าที่มีขนาดความละเอียด 112 x 92 จุดภาพ เมื่อเปลี่ยนเป็นเวกเตอร์จะมีขนาด 10,304 มิติ จากเหตุผลดังกล่าวจึงทำให้การคำนวณด้วยวิธี PCA และวิธี LDA ทำได้ยากและ

ใช้เวลามาก ต่อมานักวิจัยได้พัฒนารูปแบบการคำนวณหาพีเจอร์์ของภาพใบหน้า ให้คำนวณง่ายยิ่งขึ้นและใช้เวลาในการคำนวณน้อยกว่าเดิม แนวทางการวิจัยและพัฒนาจึงดังต่อไปนี้

Yang [3] ได้เสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติมาคำนวณโดยตรง วิธีการนี้มีชื่อว่า การวิเคราะห์องค์ประกอบหลักของภาพ (Image Principal Component Analysis: IMPCA) โดยที่วิธีการนี้สามารถลดมิติได้เฉพาะแนวหลัก (column) ของภาพ แต่ไม่สามารถลดมิติของแถว (row) จากภาพได้

Kongsontana [4] ได้นำเสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติ มาคำนวณโดยตรง และได้นำข้อมูลของกลุ่มมาพิจารณา มีชื่อว่า 2DLDA แต่วิธีนี้สามารถลดมิติได้เฉพาะหลัก (column) ของภาพเช่นกัน ไม่สามารถลดมิติได้พร้อมกันทั้งแถวและหลักของภาพ

Zhang [5] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลักโดยใช้เมตริกการฉายสองตัว วิธีการนี้มีชื่อว่า Eigenimages แต่วิธีการนี้ไม่ได้นำข้อมูลของกลุ่มมาพิจารณา

สิทธิพันธ์ [6] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลักโดยใช้เมตริกการฉายสองตัวและได้นำข้อมูลของกลุ่มมาพิจารณา วิธีการนี้มีชื่อว่า CSLDA แต่ยังคงพบว่ายังมีการปนกันของข้อมูลที่มีการกระจายระหว่างกลุ่ม ซึ่งมีผลต่อความถูกต้องในการรู้จำใบหน้า

วิธีการแก้ปัญหาการปนกันของข้อมูลระหว่างกลุ่ม ซึ่งมีผลต่อการรู้จำภาพใบหน้า นั้น ต้องหาทฤษฎีมาแก้ปัญหาการปนกันของข้อมูลระหว่างกลุ่ม จากการศึกษาพบว่า การแก้ปัญหาการปนกันของข้อมูลระหว่างกลุ่มนี้เสนอโดย

Li [7] แก้ปัญหาการปนกันของข้อมูลระหว่างกลุ่มบนพื้นฐาน LDA โดยนำ Weighted Pairwise Scatter มาใช้แก้ปัญหาการปนกันระหว่างกลุ่มของข้อมูลเสียง

ดังนั้นวิทยานิพนธ์นี้จึงได้นำเสนอวิธีการหาพีเจอร์์ของภาพวิธีใหม่ มีชื่อว่า Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) สำหรับการหาพีเจอร์์ของภาพ วิธีการที่นำเสนอนี้เป็นการประยุกต์โดยนำ Weighted Pairwise Scatter [7] ที่ใช้สำหรับการแก้ปัญหาการปนกันของข้อมูลเสียงมาประยุกต์ใช้ในการคำนวณเมตริกการกระจายระหว่างกลุ่มของการหาพีเจอร์์ CSLDA [6] ของภาพซึ่งเป็นอัลกอริทึมบนพื้นฐานของ LDA เพื่อทำให้ข้อมูลระหว่างกลุ่มที่มีการปนกันแยกออกจากกัน เมื่อข้อมูลที่ปนกันแยกห่างจากกันแล้วทำให้ผลการรู้จำภาพใบหน้าดีขึ้นกว่าเดิม

1.4 ขอบเขตของวิทยานิพนธ์

วิทยานิพนธ์นี้ทำการศึกษาอัลกอริทึมใหม่ของการรู้จำใบหน้า มีชื่อว่า WPS-CSLDA ซึ่งวิธีการนี้เป็นการนำ Weighted Pairwise Scatter มาใช้แก้ปัญหาการปนกันของข้อมูลที่เกิดจากการคำนวณหาพีเจอร์์ CSLDA แล้วทำการเปรียบเทียบความถูกต้องในการรู้จำใบหน้ากับวิธี Eigenimages และ วิธี CSLDA กับฐานข้อมูล 3 แห่ง ได้แก่ ฐานข้อมูลภาพใบหน้า ORL ฐาน

ข้อมูลภาพใบหน้า AR และฐานข้อมูลภาพใบหน้า Yale โดยที่ฐานข้อมูลภาพใบหน้า ORL และฐานข้อมูลภาพใบหน้า Yale นั้นภาพใบหน้าทั้งหมดเป็นภาพระดับสีเทา ส่วนฐานข้อมูลภาพใบหน้า AR เป็นภาพใบหน้าทั้งหมดเป็นภาพสีจึงทำการเปลี่ยนภาพทั้งหมดเป็นภาพระดับสีเทา และถ้าฐานข้อมูลใดมีภาพใบหน้าขนาดใหญ่ก็จะทำการลดขนาดภาพโดยการตัดภาพด้วยมือจากโปรแกรม Photoshop และในการคำนวณหาความถูกต้องในการรู้จำได้ใช้โปรแกรม Matlab 7.0

1.5 รายละเอียดของวิทยานิพนธ์

เนื้อหาของวิทยานิพนธ์นี้ได้แบ่งเป็น 6 บท เพื่อแสดงถึงวิธีการที่เคยใช้ในการรู้จำใบหน้า ทฤษฎีที่นำมาใช้ในวิทยานิพนธ์ วิธีการรู้จำใบหน้าที่น่าสนใจ ผลการทดลอง สรุป และแนวทางในการพัฒนาต่อไป ดังมีรายละเอียดในบทต่างๆ ดังต่อไปนี้

บทที่ 1 บทนำ

กล่าวถึง ความเป็นมาและความสำคัญของวิทยานิพนธ์ จุดมุ่งหมายและวัตถุประสงค์ของวิทยานิพนธ์ สมมติฐานของการศึกษาและทฤษฎีแนวความคิดที่ใช้ในการวิจัย ขอบเขตของวิทยานิพนธ์ และรายละเอียดของวิทยานิพนธ์

บทที่ 2 การประมวลผลภาพและการรู้จำใบหน้า

กล่าวถึง ความยุ่งยากในการรู้จำด้วยคอมพิวเตอร์ ประโยชน์ในการนำไปประยุกต์ใช้งาน และวิธีการในการรู้จำใบหน้าแบบต่างๆ ที่มีการวิจัยออกมาจนถึงปัจจุบัน

บทที่ 3 ทฤษฎีพื้นฐานที่เกี่ยวข้อง

บทนี้จะเป็นการนำเสนอวิธีการรู้จำใบหน้าโดยใช้วิธีการทางสถิติแบบต่างๆ ที่เป็นพื้นฐานของการวิจัยในวิทยานิพนธ์เล่มนี้ ซึ่งประกอบด้วย วิธีการวิเคราะห์องค์ประกอบหลัก (PCA), วิธีการแบ่งแยกเชิงเส้น (LDA), วิธี 2DPCA, วิธี 2DLDA, วิธี Eigenimages และวิธี CSLDA

บทที่ 4 WPS-CSLDA

บทนี้เป็นการนำเสนอการรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA ซึ่งเป็นวิธีการใหม่ในการรู้จำใบหน้าที่วิทยานิพนธ์เล่มนี้นำเสนอ ว่ามีแนวคิดและหลักการคำนวณหาฟิเจอร์อย่างไร

บทที่ 5 ผลการทดลองหรือการวิเคราะห์

บทนี้กล่าวถึง การวิเคราะห์และผลการทดลองเพื่อทดสอบแนวคิดที่น่าสนใจว่า การทดลองนั้นได้ผลเป็นเช่นไร

บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ

บทนี้จะเป็นการสรุปผลจากการวิจัยที่ผ่านมารวมถึงข้อเสนอแนะสำหรับแนวทางในการพัฒนาต่อไป

บทที่ 2

การประมวลผลภาพและการรู้จำใบหน้า

ในการประมวลผลและการรู้จำภาพใบหน้านั้นมีเงื่อนไขและข้อจำกัดในการรู้จำมากมาย ดังนั้น หัวข้อต่อจากนี้ไปจะเป็นการแสดงให้เห็นว่าการประมวลผลและการรู้จำภาพใบหน้านั้นมีเงื่อนไขและข้อจำกัดอะไรบ้าง รวมถึงวิธีการที่มีการนำไปใช้จริงในปัจจุบัน

2.1 ความซับซ้อนและความยุ่งยากในการรู้จำใบหน้าด้วยคอมพิวเตอร์

การรู้จำใบหน้านั้นมนุษย์ดูเหมือนจะทำได้อย่างรวดเร็วและง่ายดายเหมือนกับไม่ต้องใช้ความพยายามมากนัก แต่ก็ไม่ได้หมายความว่า การใช้คอมพิวเตอร์รู้จำได้นั้นจะเป็นเรื่องง่าย ๆ ไปด้วย

ความยากอันดับแรก คือ ในการประยุกต์ใช้งานจริง ๆ ระบบจำเป็นจะต้องมีความถูกต้องสูงมาก เช่น ระบบรักษาความปลอดภัย, E-Passport และระบบ ATM Cash Machine ที่ใช้การรู้จำใบหน้าในการระบุบุคคลนั้น จำเป็นจะต้องมีความเร็วสูงและต้องไม่ผิดพลาดเลย การประยุกต์ทางด้านการตรวจจับคนที่เจ้าหน้าที่ทางกฎหมายต้องการจับตัวและผู้ที่เกี่ยวข้องนั้น ระบบต้องมีการติดตั้งที่สถานที่สาธารณะต่าง ๆ ตัวอย่างเช่น สนามบิน ระบบจะต้องสามารถตรวจจับบุคคลผิดกฎหมายที่ต้องการออกมาให้ได้ในขณะที่ทุกคนต่างกำลังเดินหรือเคลื่อนไหวกันอยู่ ซึ่งการทำงานลักษณะนี้เป็นเรื่องยากมาก ดังนั้นระบบส่วนใหญ่ที่ทำงาน หรือบทความที่วิจัยกันอยู่ต่างมีขอบเขตความสามารถและขีดจำกัดในการประยุกต์ด้วยกันทั้งหมด

ความยากอันดับที่สอง คือ ในการรู้จำใบหน้าที่มีประสิทธิภาพจริง ๆ นั้นไม่สามารถใช้การประมวลผลภาพแบบเก่า ๆ เช่น การเทียบภาพ (Image Matching) ได้ ซึ่งส่วนที่ท้าทายนั้นแบ่งเป็นสองส่วน คือ ส่วนแรกเป็นส่วนการตรวจจับใบหน้า ที่จะต้องทำการตัดสินใจว่าในภาพนั้นมีภาพใบหน้าอยู่หรือไม่ และอยู่ตำแหน่งใดของภาพ และส่วนที่สองการรู้จำใบหน้าที่จะสามารถระบุได้ว่าเป็นบุคคลใดจากฐานข้อมูลที่มีอยู่ ซึ่งใบหน้านั้นยากที่จะตรวจจับและรู้จำทั้งนี้เพราะข้อเท็จจริงที่ว่าภาพใบหน้านั้นไม่มีรูปแบบตายตัว ซึ่งใบหน้าอาจมีการเปลี่ยนแปลงลักษณะต่าง ๆ ได้ทั้งการส่องสว่าง การแสดงออกทางอารมณ์ การวางตัวของศีรษะ ทรงผม คิ้ว หนวด ไล่แว่น สวมหมวก ทำให้สิ่งเหล่านี้เป็นอุปสรรคทั้งสิ้น

2.2 การนำระบบรู้จำใบหน้าไปใช้จริงในปัจจุบัน

2.2.1 ฝ่าคูบุคคลแปลกหน้า

เป็นการทำงานอัตโนมัติ ในการฝ่าคูบุคคลที่ไม่มีอยู่ในฐานข้อมูล มันสามารถส่งเสียงเตือนเมื่อเห็นคนแปลกหน้า และสามารถเก็บบันทึกภาพใบหน้าได้ด้วย ซึ่งทำให้สามารถกลับมาดูภาพใบหน้าของคนบุกลูกได้เมื่อต้องการหรือภายหลังถึงสิ้นสุดวัน อีกทั้งยังมีการบันทึกเวลาที่บุคคลนั้นเข้ามาให้เห็นด้วย ซึ่งจะสามารถดูได้ง่ายกว่าการดูวิดีโอทั้งหมดเพราะการทำงานด้วยวิธีนี้จะเก็บเฉพาะภาพใบหน้าของผู้บุกลูกเท่านั้น

2.2.2 หนังสือเดินทางอิเล็กทรอนิกส์ (E-Passport)

เป็นการตรวจสอบหนังสือเดินทางของบุคคลที่เดินทางเข้า-ออกระหว่างประเทศกับฐานข้อมูลอาชญากร ตัวอย่างเช่น ตำรวจต้องการฝ่าคูบุคคลต่าง ๆ ที่เดินทางเข้า-ออกประเทศการทำงานในลักษณะเช่นนี้ สามารถบันทึกข้อมูลวันเวลาและจำนวนครั้งที่บุคคลนั้นเข้ามาและออกไปให้เห็นด้วย

2.2.3 ล็อกหน้าจอ (Screen Lock)

เป็นการทำงานที่เหมือนกับการล็อกหน้าจอแบบปกติที่ใช้การป้อนรหัสเพื่อผ่านเพื่อเข้าสู่ระบบคอมพิวเตอร์หรือระบบเครือข่ายนั่นเอง เว้นแต่ว่ามันจะมีการปลดล็อกอัตโนมัติเมื่อมีผู้ใช้กลับมาและมองไปที่กล้องจับภาพ หมายความว่าจะไม่ยอมให้บุคคลอื่น ๆ เข้ามาใช้เครื่องหรือเข้าสู่ระบบได้ ระบบนี้สามารถออกแบบให้ปลดล็อกเพียงคนเดียวที่ยินยอมนั้นกลับมาอีกครั้ง หรือสามารถออกแบบให้ปลดล็อกเมื่อกลุ่มใบหน้าที่อยู่ในบานข้อมูลคนใดคนหนึ่งกลับมาได้

2.2.4 การค้นหาในฐานข้อมูล

การทำงานนี้สามารถค้นหาภาพใบหน้าในฐานข้อมูลภาพขนาดใหญ่ หรือเป็นฐานข้อมูลภาพเคลื่อนไหวที่เป็นวิดีโอขนาดใหญ่ก็ได้ โดยจะมีการตรวจหาอย่างอัตโนมัติในทุก ๆ ภาพหรือในทุก ๆ เฟรมของวิดีโอ ที่มีภาพบุคคลที่เราต้องการหาซึ่งอาจจะต้องการหาหลายคนในคราวเดียว ตัวอย่างเช่น การค้นหาภาพจากแถบบันทึกวิดีโอ ซึ่งเป็นเรื่องที่น่าเบื่อและเป็นกระบวนการที่ต้องใช้เวลามากคือการค้นหาภาพบุคคลที่ต้องการ ซึ่งถ้าหากวิดีโอมีขนาดใหญ่หรือต้องทำบ่อย ๆ กลับไปกลับมาก็จะเป็นเรื่องที่เสียเวลามาก แต่ถ้าหากใช้โปรแกรมการรู้จำใบหน้าทำการค้นหาแทนซึ่งสามารถทำงานได้อย่างไม่รู้จักเหน็ดเหนื่อยในการค้นหาต่อเนื่องกันเป็นชั่วโมง ๆ โดยไม่มีการหยุดพัก หรือถ้าเป็นการค้นหาในฐานข้อมูลขนาดใหญ่เช่น ในฐานข้อมูลของกรมตำรวจเป็นต้น มันจะเป็นการหาบุคคลที่มีใบหน้าคล้ายกับภาพที่เราต้องการค้นหา โดยจะแสดงรายชื่อ ประวัติ และภาพออกมาตามลำดับความคล้าย อาจจะเป็น 5 หรือ 10 ภาพเป็นต้น

สามารถสรุปเป็นสองแนวทางใหญ่ ๆ คือ

1) กรณีต้องการค้นหาบุคคลในฐานข้อมูลใบหน้าขนาดใหญ่ (เช่น ฐานข้อมูลของกรมตำรวจ) ระบบเช่นนี้จะต้องแสดงภาพและประวัติของบุคคลที่มีความคล้ายออกมา โดยที่ระบบไม่จำเป็นต้องมีความเร็วแบบเวลาจริงในการทำงาน (Non Real Time)

2) กรณีต้องการชี้ระบุตัวบุคคลแบบเวลาจริง (เช่น ในระบบเฝ้าดูเพื่อรักษาความปลอดภัยและระบบติดตามตำแหน่งเป็นต้น) ระบบนี้จำเป็นต้องใช้เวลาในการทำงานที่รวดเร็วและมีความถูกต้องแม่นยำที่สูงมาก

2.3 สิ่งที่เปลี่ยนแปลงได้บนภาพใบหน้าที่มีผลกระทบในการรู้จำใบหน้า

2.3.1 สภาพแวดล้อม

- การส่องสว่าง
- พื้นหลัง

2.3.2 การแสดงออกทางอารมณ์บนใบหน้า

- การขมวดคิ้ว
- ลักษณะของปาก (ยิ้ม หัวเราะ)
- กระพริบตา

2.3.3 อื่น ๆ

- การใส่แว่น
- สวมหมวก
- ทรงผม
- หนวด
- ขนาดของศีรษะเนื่องจากระยะใกล้ไกลจากกล้อง(Scaling)
- ตำแหน่งและการวางตัวของศีรษะ

2.4 วิธีในการรู้จำภาพใบหน้าที่มีการวิจัยกันในปัจจุบัน

จากการสำรวจในบทความต่าง ๆ มีวิธีการมากมายที่ใช้ในการรู้จำใบหน้าซึ่งสามารถแบ่งเป็นกลุ่มได้ดังนี้

- 2.4.1 ลักษณะทางเรขาคณิต (Geometrical Features)
- 2.4.2 การเปรียบเทียบกราฟ (Graph Matching)
- 2.4.3 วิธีการทางสถิติ (Statistical)
- 2.4.4 โครงข่ายประสาทเทียม (Neural network)
- 2.4.5 วิธีอื่น ๆ (Other Methods)

2.4.1 ลักษณะทางเรขาคณิต (Geometrical Features)

มีนักวิจัยจำนวนมากที่เลือกใช้การรู้จำใบหน้าในลักษณะทางเรขาคณิต คืออาศัยอัตราส่วนหรือสัดส่วนของระยะทางจากตาซ้ายไปตาขวา จากตาไปจมูก จากจมูกไปปาก หรือใช้รูปร่างของปาก รูปร่างของตา และรูปร่างของคาง ซึ่งเป็นลักษณะทางกายภาพของใบหน้า ดังมีรายละเอียดของงานวิจัยดังนี้

Kanade [8] ได้นำเสนอวิธีการแยกลักษณะเด่น (Automatic Feature Extraction) โดยใช้อัตราส่วนของระยะทาง (Ratio of Distance) โดยมีผลการรู้จำอยู่ระหว่าง 45% -75% ของฐานข้อมูลภาพทั้งหมด 20 บุคคล

Brunelli และ Poggio [9] ได้คำนวณเซตของลักษณะทางเรขาคณิตเช่น ความกว้างของจมูก ตำแหน่งของปากบนภาพใบหน้า และรูปร่างของคาง โดยมีอัตราการรู้จำ 90% ของฐานข้อมูล 47 คน อย่างไรก็ตามในงานวิจัยนี้ได้แสดงให้เห็นว่าเมื่อถ้าใช้วิธีเปรียบเทียบเทมเพลต (templet) อย่างง่ายนั้น จะให้ผลการรู้จำ 100% ที่ฐานข้อมูลเดียวกัน

Cox และคณะ [10] ได้เสนอเทคนิคระยะทางผสม (Mixture Distance) ที่ทำให้ได้อัตราการรู้จำสูงถึง 95% โดยใช้ฐานข้อมูลที่สุ่มมาทีละ 96 ภาพ จากทั้งหมด 685 ภาพ ซึ่งใบหน้าของแต่ละภาพนั้นจะถูกแทนด้วยระยะทางต่าง ๆ ถึง 30 อย่าง

ระบบการรู้จำใบหน้าแบบที่มีการพิจารณาลักษณะทางเรขาคณิตนี้ เป็นระบบที่มีประโยชน์และสามารถนำไปใช้ได้ผ่านทางปฏิบัติในการค้นหาในฐานข้อมูลขนาดใหญ่ได้ ถ้ามีการวัดระยะทางจากจุดที่เป็นลักษณะเด่นได้อย่างถูกต้องทั้งนี้เพราะข้อมูลที่ได้อาจมีจำนวนน้อยมากเมื่อเทียบกับการคำนวณแบบอื่น ทำให้เวลาในการค้นหาน้อยกว่าวิธีอื่น ๆ ดังนั้นจึงมีการคิดสร้างระบบฐานข้อมูลขึ้นมาสนับสนุนวิธีการนี้ ซึ่งฐานข้อมูลดังกล่าวถูกเรียกว่าฐานข้อมูลภาพถ่ายด้าน (Mug shot Database) ทั้งนี้เพราะถ้ามีภาพด้านข้างด้วยจะช่วยให้การค้นหาลักษณะเด่นได้ถูกต้องมากขึ้น แต่ความถูกต้องก็ยังขึ้นอยู่กับอัลกอริทึมที่ใช้ด้วย

2.4.2 การเปรียบเทียบกราฟ (Graph Matching)

เป็นการมองภาพในลักษณะที่เป็นเวกเตอร์ของกราฟซึ่งมีจุดและเส้นในการเชื่อมต่อเป็นโครงร่างที่แนบไปบนใบหน้าตามลักษณะเด่นที่พิจารณา เช่น รูปร่างกราฟของตา ปาก คาง หู และคิ้ว เป็นต้น

Lades และคณะ [11] เสนอสถาปัตยกรรมการเชื่อมต่อแบบยืดหยุ่น (Dynamic Link Architecture) สำหรับการรู้จำใบหน้าแบบไม่ขึ้นอยู่กับความเพี้ยน โดยใช้การเทียบกราฟแบบยืดหดได้ (Elastic Graph Matching) เพื่อหากราฟในฐานข้อมูลที่ใกล้เคียงที่สุด เมื่อวัตถุเป้าหมายจะถูกแทนด้วยกราฟห่าง ๆ กันที่มีจำนวนไม่มากนัก (Sparse Graph) ซึ่งข้อมูลจะถูกติดป้าย (Label) แทนด้วยฟังก์ชันความโค้งและขอบต่าง ๆ ก็ถูกติดป้ายซึ่งแทนด้วยระยะทางเรขาคณิต เมื่อทำการเปรียบเทียบเข้ามาหนึ่งภาพ จะทำการค้นหาในฐานข้อมูลทั้งหมด 87 คน โดยภาพที่นำมาทดสอบอาจมีหลายลักษณะที่แตกต่างกัน เช่น มีการหมุนของศีรษะ ดังนั้นเวลาที่ใช้ในการค้นต่อครั้งเป็น 25 วินาที แต่บทความนี้ไม่ได้แสดงผลการรู้จำได้เลย

เมื่อ Wiskott และคณะ [12] ได้ปรับปรุงกระบวนการของ Lades โดยทำการทดสอบฐานข้อมูล 300 คน จากฐานข้อมูล FERET ปรากฏว่ามีผลการรู้จำสูงถึง 97.30% แต่ว่าการทดลองนี้ไม่ได้บอกเวลาที่ใช้ในการรู้จำ และตัวประมวลผลที่ใช้ ซึ่งอาจเป็นไปได้ว่าระบบดังกล่าวจำเป็นต้องใช้ตัวประมวลผลที่มีประสิทธิภาพและความเร็วสูงมาก เพราะวิธีการนี้คงใช้เวลาในการประมวลผลนาน

2.4.3 วิธีการทางสถิติ

วิธีการทางสถิติประกอบด้วยวิธีการเปรียบเทียบเทมเพลต ซึ่งเป็นพื้นฐานของระบบ โดยที่ภาพที่ใช้สอนและภาพที่ใช้ทดสอบจะถูกเปรียบเทียบกัน (Match) โดยการวัดความสัมพันธ์ระหว่างภาพที่ใช้สอนกับภาพที่ใช้ทดสอบ อย่างไรก็ตามวิธีการทางสถิติยังประกอบด้วยวิธีที่ใช้การแปลงเชิงเส้น (Linear Transform) อันได้แก่ การวิเคราะห์องค์ประกอบหลัก (PCA) การวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) เป็นต้น โดยระบบที่ใช้การแปลงเชิงเส้นเป็นวิธีที่เข้ามาแก้ไขวิธีการเปรียบเทียบเทมเพลต เพราะว่าวิธีวิธีการเปรียบเทียบเทมเพลตจะทำการแยกแยะภาพที่ใช้ในการสอนกับภาพที่ใช้ทดสอบในปริภูมิที่มีขนาดใหญ่

วิธีการเปรียบเทียบเทมเพลต [9] นี้ทำงานโดยการหาความสัมพันธ์ของภาพ 2 ภาพโดยตรงเมื่อภาพมีขนาดเดียวกัน มีองค์ประกอบของใบหน้าวางอยู่ตรงกัน และมีการส่องสว่างของแสงในทิศทางเดียวกัน จะให้ผลการรู้จำอย่างมีประสิทธิภาพที่ดีถึง 100% แต่ถ้าภาพที่เข้ามาทดสอบมีการเปลี่ยนแปลงไปจากเดิม ต้องมีการประมวลผลก่อน (preprocessing) เพื่อให้ภาพเข้าผู้ตำแหน่งหรือการส่องสว่างใกล้เคียงกัน

เนื่องจากวิธีการเปรียบเทียบเทมเพลตนี้ ทำงานโดยการหาสหสัมพันธ์ของภาพ 2 ภาพจึงเกิดข้อจำกัดมาก ดังนั้นวิธีการที่สามารถหาพีเจอร์ที่แท้จริงของภาพและลดขนาดของภาพได้คือวิธีที่ใช้การระบบการแปลงเชิงเส้น วิธีการนี้สามารถตัดข้อมูลที่ เป็นผลกระทบต่อการรู้จำใบหน้าอันได้แก่ การส่องสว่าง การแสดงทางอารมณ์บนใบหน้าได้

2.4.3.1 การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PCA)

การวิเคราะห์องค์ประกอบหลักเป็นที่ทราบกันดีว่าเหมือนวิธีการยูเนนเลิฟ (Karhunen Loeve Expansion) โดยวิธีการนี้สามารถหาพีเจอร์ของภาพและการสร้างภาพใหม่ได้ ซึ่งวิธีการนี้ถูกนำมาใช้ในงานต่าง ๆ เช่น การจดจำรูปแบบและการมองเห็นของคอมพิวเตอร์

Sirowich และ Kirby [13] ซึ่งพวกเขาเป็นกลุ่มแรกที่ได้ประยุกต์การวิเคราะห์องค์ประกอบหลักมาใช้ในการรู้จำใบหน้ามนุษย์

Turk และ Pentland [1] ได้เสนอวิธีการรู้จำใบหน้าที่มีชื่อว่า ใบหน้าไอเกน (Eigenface) วิธีการนี้สามารถทำได้โดยการฉายภาพใบหน้าลงในปริภูมิใบหน้าไอเกน (Eigenface space) โดยใช้ไอเกนเวกเตอร์เป็นแกนการฉาย ซึ่งไอเกนเวกเตอร์คำนวณมาจากเมตริกความแปรปรวนของภาพใบหน้า หลังจากได้ใบหน้าไอเกนแล้วนำไปค้นหาในฐานข้อมูลใบหน้าไอเกน ในการทดลองนี้ได้ใช้ฐานข้อมูลใบหน้า 2,500 ภาพ ของ 16 บุคคล ที่มีการเคลื่อนไหว การวางตำแหน่งศีรษะในหลาย ๆ ลักษณะ มีการขยับขยายและมีการเปลี่ยนแปลงของแสง ผลการทดลองเมื่อมีการเปลี่ยนแปลงของแสงมีความสามารถในการรู้จำ 96% เมื่อการวางตำแหน่งศีรษะมีการเปลี่ยนแปลงจะมีความสามารถในการรู้จำถึง 85% และในการรู้จำภาพที่มีการขยับขยายมีความถูกต้องสูงถึง 64% โดยตัวแปรของการขยับขยายของภาพนั้น ได้ใช้อัลกอริทึมทำการปรับขนาดศีรษะให้มีขนาดเดียวกับฐานข้อมูลใบหน้าไอเกน หลังจากนั้นทำการตัดภาพให้เหลือเฉพาะส่วนที่เป็นใบหน้าเพื่อลดผลจากการเปลี่ยนแปลงของมุมและฉากหลัง

จากแนวความคิดของ Turk และ Pentland เป็นผลให้มีการนำไปประยุกต์กับงานด้านต่าง ๆ อีกมากอันได้แก่

Pentland และคณะ [14] มีผลการรู้จำที่ดีในฐานข้อมูลขนาดใหญ่ โดยมีความถูกต้องในการรู้จำ 95% ในฐานข้อมูล 200 คน ที่สุ่มมาจากฐานข้อมูลขนาด 3,000 คน

Moghaddame และ Pentland [15] ได้ใช้วิธีการหาใบหน้าไอเกนในการเข้ารหัสภาพใบหน้าสำหรับงานด้านการโทรภาพ (video telephone)

Lee และคณะ [16] ได้เสนอวิธีการใช้ PCA ในการตรวจหาภาพส่วนหัวในภาพพื้นหลังที่ซับซ้อนของแต่ละบุคคลและทำการรู้จำภาพ โดยทำการเปรียบเทียบกับฐานข้อมูลแต่ละบุคคลที่รู้จักอยู่แล้ว

Lee และคณะ [17] ได้ทำการศึกษาวิธี PCA เมื่อใช้กับฐานข้อมูลภาพใบหน้าขนาดใหญ่

Crowly และ Schwerdt [18] ได้เสนอวิธีการ PCA ในการเข้ารหัสและบีบอัดข้อมูลวิดีโอ

ในการรู้จำใบหน้าด้วยวิธี PCA ภาพใบหน้าที่เป็น 2 มิติ ซึ่งเก็บในรูปแบบเมตริก ก่อนการคำนวณจำเป็นต้องเปลี่ยนจากเมตริกภาพเป็นเวกเตอร์ภาพ จึงเป็นผลให้ภาพมีมิติที่มากในปริภูมิ (space) ดังนั้นจึงเป็นการยากที่จะคำนวณหาเมตริกการกระจายร่วม เนื่องจากเมตริกการกระจายร่วมจะมีขนาดใหญ่เมื่อเปรียบเทียบกับจำนวนภาพที่ใช้สอน ทำให้เมื่อคำนวณหาไอเกนเวกเตอร์และค่าไอเกนแล้วคำนวณได้ยากและใช้เวลามาก Yang และคณะ [3] ได้พัฒนาหลักการใหม่ที่น่าสนใจมาจากรูปแบบ PCA เพื่อแก้ไขจุดบกพร่องนี้ จึงได้เสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติมาคำนวณโดยตรง วิธีการนี้มีชื่อว่า การวิเคราะห์องค์ประกอบหลักของภาพ (Image Principal Component Analysis: IMPCA หรือ 2DPCA) จากผลการทดลองของ Yang แสดงให้เห็นถึงประโยชน์ของเวลาที่ใช้ในการคำนวณและความถูกต้องในการรู้จำที่ดีกว่าวิธี PCA

ต่อมา Zhang [5] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลักโดยใช้แกนการฉายสองแกน วิธีการนี้มีชื่อว่า Eigenimage วิธีการนี้สามารถคำนวณได้เร็วกว่า PCA และ 2DPCA เนื่องจากมีแกนการฉาย 2 แกน ทำให้ได้ความถูกต้องในการรู้จำที่ดีกว่าทั้ง 2 วิธีที่ผ่านมา

2.4.3.2 การวิเคราะห์การแบ่งแยกเชิงเส้น (Linear Discriminant Analysis : LDA)

R.A. Fisher [19] ได้พัฒนาวิธีการแบ่งแยกเชิงเส้น มาใช้ในงานด้านการจดจำวัตถุ โดยให้ชื่อว่าวิธีการแบ่งแยกเชิงเส้นของฟิชเชอร์ (Fisher Linear Discriminant: FLD)

Etemad และ Chellappa [20] ได้เสนอวิธีการแบ่งแยกเชิงเส้นของฟิชเชอร์ (Fisher Linear Discriminant: FLD) สำหรับกระบวนการรู้จำใบหน้า จุดมุ่งหมายของวิธีการนี้ต้องการหาการฉายภาพของข้อมูลภาพใบหน้าที่มีการกระจายระหว่างกลุ่มมาก ๆ และการกระจายภายในกลุ่มน้อย ๆ วิธีการนี้สามารถทำได้โดยคำนวณหาเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มของภาพที่ใช้สอน แล้วนำเมตริกทั้งสองมาคำนวณหาไอเกนเวกเตอร์และค่าไอเกนที่สมนัยกันจากมากไปน้อย เพื่อใช้ในการฉายภาพ ดังนั้นเราจะได้ฟีเจอร์ของภาพใหม่โดยการฉายภาพที่ใช้สอนลงบนไอเกนเวกเตอร์

ปริภูมีย่อยของ LDA [2], [21], [22], [23] วิธีการนี้ได้รวมเอาวิธี PCA และ LDA เข้าด้วยกัน ดังนั้นวิธีนี้จะประกอบด้วย 2 ขั้นตอน ขั้นตอนแรก ภาพจะถูกฉายลงบนปริภูมิของใบหน้าไอเกน โดยวิธี PCA และขั้นตอนที่สอง นำใบหน้าไอเกนมาฉายด้วยวิธี LDA ในวิธีการนี้จะสามารถลดผลกระทบของตัวแปรที่มีผลกระทบในการรู้จำใบหน้า และสามารถพิจารณาข้อมูลการแบ่งแยกของกลุ่มได้ด้วย

Kongsontana [4] ได้นำเสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติ มาคำนวณโดยตรง และได้นำข้อมูลของกลุ่มมาพิจารณามีชื่อว่า 2DLDA แต่วิธีนี้สามารถลดมิติได้เฉพาะหลัก (column) ของภาพไม่สามารถลดมิติได้พร้อมกันทั้งแถวและหลักของภาพ

สิทธิพันธ์ [6] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแนวและหลัก โดยใช้แกนการฉายสองแกนและได้นำข้อมูลของกลุ่มมาพิจารณา วิธีการนี้มีชื่อว่า CSLDA แต่ยังคงพบว่ายังมีการปนกันของข้อมูลที่มีการกระจายระหว่างกลุ่ม ซึ่งมีผลต่อความถูกต้องในการรู้จำใบหน้า

2.4.4 โครงข่ายประสาทเทียม (Neural network)

Temdee [24] และคณะ ได้ทำการหาเส้นขอบของภาพใบหน้าและนำมาตัดเป็นส่วน ๆ ขององค์ประกอบภาพใบหน้าเช่น ตา จมูก ริมฝีปาก แล้วนำมาเข้ารหัสภาพแต่ละส่วนเพื่อใช้เป็นอินพุตให้กับโครงข่ายประสาทเทียมแบบแพร่กระจายกลับ (Backpropagation Neural Network) สำหรับใช้ในการรู้จำใบหน้า ซึ่งได้ทำการทดลองกับฐานข้อมูล ORL ซึ่งได้ผลความถูกต้องในการรู้จำ 85%

Er และคณะ [25] ใช้หลักการวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) มาทำการลดขนาดของข้อมูลเพื่อเป็นอินพุตให้กับ Radial Basis Function (RBF) ซึ่งวิธีการนี้ให้ผลความถูกต้องในการรู้จำ 94%

Demer [26] ใช้หลักการวิเคราะห์องค์ประกอบหลัก (PCA) มาคำนวณหาองค์ประกอบหลัก 50 ค่าแรกแล้วนำมาเป็นอินพุตของโครงข่ายประสาทเทียมแบบหลายชั้น (Standard Multilayer Perceptron) เพื่อใช้ในการรู้จำใบหน้า ซึ่งผลการรู้จำใบหน้าด้วยวิธีนี้ให้ผลการรู้จำที่ดี แต่ฐานข้อมูลที่น่ามาทดลองนั้นค่อนข้างง่ายเกินไป โดยที่ภาพที่ใช้ทดลองเป็นภาพด้านหน้า ไม่มีการเปลี่ยนแปลงของแสง ไม่มีการเปลี่ยนแปลงจากการหมุน และฐานข้อมูลมีเพียง 20 คนเท่านั้น

Lawrence และคณะ [27] ได้ใช้โครงข่ายประสาทเทียมแบบผสม (Hybrid Neural Network) โดยใช้การผสมของโครงข่ายประสาทเทียมแบบ Self-Organizing Map Neural Network (SOM) เพื่อลดข้อมูลภาพใบหน้า และใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) ในการรู้จำใบหน้า ซึ่ง Lawrence และคณะได้เปรียบเทียบผลการรู้จำกับฐานข้อมูล ORL ได้ดังนี้

1. ใช้ SOM ในการลดมิติข้อมูล และใช้ CNN ในการรู้จำ ให้ความถูกต้องในการรู้จำ 96.20%
2. ใช้การแปลงแบบคาร์สูงเนนลิฟ (KL Transform) ในการลดข้อมูลภาพใบหน้า และใช้ CNN ในการรู้จำ ให้ความถูกต้องในการรู้จำ 94.70%
3. ใช้ SOM ในการลดข้อมูล และใช้โครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) ในการรู้จำ ซึ่งให้ผลความถูกต้องในการรู้จำ 60.00%
4. ใช้การแปลงแบบคาร์สูงเนนลิฟให้ผลความถูกต้อง 89.50%

2.4.5 วิธีอื่น ๆ (Other Methods)

ข้อมูลเชิงระยะทาง (Range Data)

เป็นวิธีการรู้จำใบหน้าที่แตกต่างกันจากวิธีอื่น โดยการใช้ข้อมูลเชิงระยะทาง ซึ่งข้อมูลภาพของแต่ละบุคคลได้มาจากการสแกนด้วยระบบเลเซอร์สแกน ที่มีข้อมูลเชิงลึกประกอบด้วย แล้วทำการชี้เฉพาะบุคคลในรูป 3 มิติ [28]

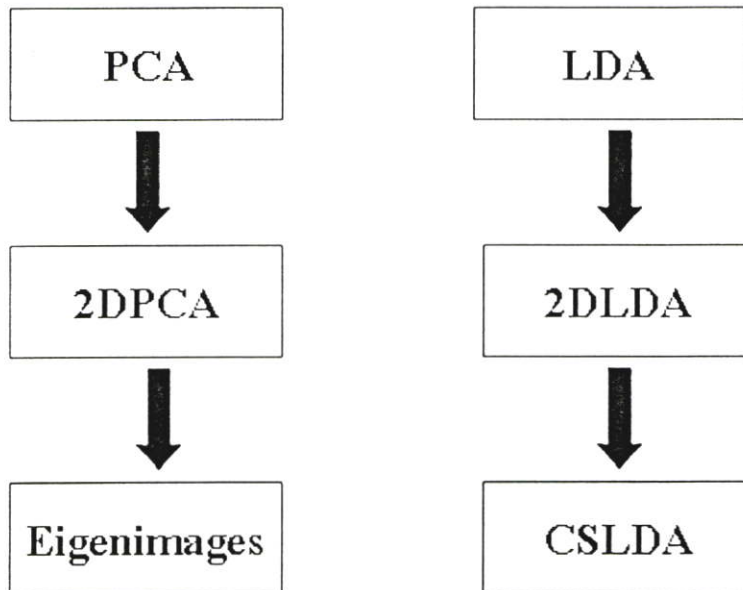
การสแกนด้วยอินฟราเรด (Infrared Scanning)

Yoshitomi [29] วิธีการนี้ได้ใช้แสงอินฟราเรดตรวจจับใบหน้าเพื่อหาการกระจายของอุณหภูมิบนใบหน้าและโครงร่างใบหน้า เมื่อนำการกระจายของอุณหภูมิมาหาค่าเฉลี่ยของอุณหภูมิเพื่อใช้เป็นอินพุตให้กับโครงข่ายประสาทเทียม และข้อมูลโครงร่างใบหน้านำมาปรับปรุงขนาดและตำแหน่งเพื่อใช้กับวิธีการแยกแยะโดยการสอน (Supervised Classification) ผลที่ได้จากทั้งสองวิธีนำมา รวมกันเพื่อชี้เฉพาะบุคคล

บทที่ 3

ทฤษฎีพื้นฐานที่เกี่ยวข้อง

บทที่ผ่านมาทำให้เราทราบถึงวิธีการรู้จำใบหน้าชนิดต่าง ๆ ซึ่งมีหลากหลายวิธีที่ใช้ในปัจจุบัน แต่วิธีการรู้จำภาพใบหน้าที่วิทยานิพนธ์นี้สนใจ คือ วิธีการรู้จำใบหน้าโดยใช้วิธีการทางสถิติ การรู้จำด้วยวิธีการทางสถิติมีทฤษฎีในการรู้จำใบหน้าที่สำคัญ 2 ทฤษฎี ได้แก่ การวิเคราะห์องค์ประกอบหลัก (PCA) และ การวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ทั้ง 2 ทฤษฎีนี้เป็นทฤษฎีพื้นฐานที่สำคัญในการพัฒนาวิธีการรู้จำภาพใบหน้าด้วยวิธีทางสถิติมาจนถึงปัจจุบัน ขั้นตอนการพัฒนาวิธีการรู้จำภาพใบหน้าของทั้ง 2 ทฤษฎีนั้นแสดงในรูปแบบที่ 3.1 โดยที่เนื้อหาทฤษฎี และรายละเอียดจะกล่าวในหัวข้อต่อไป



รูปที่ 3.1 ลำดับขั้นการพัฒนาทฤษฎีการรู้จำใบหน้า PCA และ LDA

3.1 การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

3.1.1 แนวความคิดของการวิเคราะห์องค์ประกอบหลัก

การวิเคราะห์องค์ประกอบหลักเป็นการแทนเซตของข้อมูลด้วยเซตข้อมูลใหม่ที่มีขนาดเล็กกว่าข้อมูลเดิม โดยข้อมูลใหม่นี้จะไม่มีความสัมพันธ์กันระหว่างข้อมูลภายในเซตเลย ซึ่งข้อมูลใหม่นี้ถูกเรียกว่า องค์ประกอบหลัก (Principal Components) ถ้ามีเซตข้อมูลที่มีความสัมพันธ์กัน

อยู่ภายในข้อมูล ซึ่งมีขนาด D มิติ สามารถใช้การรวมกันเชิงเส้นในการลดข้อมูลลงเป็น d มิติ โดยที่ d มีค่าน้อยกว่า D

Hotelling [30] ได้นำเสนอ PCA ซึ่งนำแนวคิดมาจากการการแปลงเชิงเส้นของตัวแปรที่มีความสัมพันธ์กัน (correlate) ไปเป็นตัวแปรใหม่ที่ไม่สัมพันธ์กันตามลำดับการลดลงของความแปรปรวน

กำหนดให้มีเวกเตอร์ t ใด ๆ ที่มีขนาด D มิติ มีค่าเฉลี่ย คือ \bar{t} และค่าเมตริกความแปรปรวนร่วม Σ ขนาด $D \times D$ ต้องการคำนวณหาเวกเตอร์ใหม่ y ที่ไม่มีความสัมพันธ์กันและมีค่าความแปรปรวนลดลงตามลำดับ จากหลักการรวมกันเชิงเส้นสามารถแสดงเป็นสมการได้ดังนี้

$$y_k = u_{1k}t_1 + u_{2k}t_2 \dots + u_{Dk}t_D = u_k^T t \quad (3.1)$$

จากสมการสามารถเขียนให้อยู่ในรูปของเมตริกได้ โดยที่ Y เวกเตอร์มีขนาด $d \times 1$ t มีขนาด $D \times 1$ และ U มีขนาด $D \times d$

$$Y = U^T t \quad (3.2)$$

จากสมการที่ (3.1) จะเห็นว่า u เป็นเพียงแกนของทิศทางเท่านั้นจึงกำหนดให้ u เป็นเวกเตอร์หนึ่งหน่วยที่มีคุณสมบัติคือ

$$u_k^T u_k = \sum_{k=1}^D u_{kj}^2 = 1 \quad (3.3)$$

จากแนวคิดการการแปลงเชิงเส้นของตัวแปรที่มีความสัมพันธ์กัน (Correlate) ไปเป็นตัวแปรใหม่ที่ไม่สัมพันธ์กัน ตามลำดับ การลดลงของความแปรปรวน ดังนั้นการคำนวณหาเมตริกความแปรปรวนร่วมของ Y ทำได้โดย

k^{TH} องค์ประกอบหลัก = การรวมกันเชิงเส้นของ $u_k^T t$ ที่มีค่าความแปรปรวนมากที่สุด $\max(u_k^T U)$

$$\begin{aligned} y_k &= u_k^T t \\ &= E[(U^T(t - \bar{t}))^2] \\ &= E[U^T(t - \bar{t})(U^T(t - \bar{t}))^T] \\ &= E[U^T(t - \bar{t})(t - \bar{t})^T U] \\ &= U^T E[(t - \bar{t})(t - \bar{t})^T] U \\ &= U^T \Sigma U \end{aligned} \quad (3.4)$$

จากสมการที่ (3.4) สามารถแก้ปัญหาโดยใช้วิธีการ Lagrange multipliers ดังนั้นจึงสามารถสรุปได้ว่าแกนการฉายก็คือไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกน λ ที่เรียงจากมากไปน้อยตามลำดับ ซึ่งคำนวณมาจากเมตริกการกระจาย

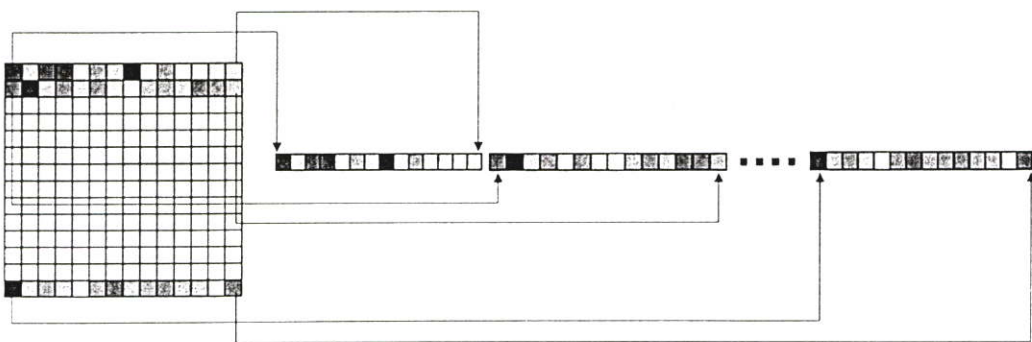
3.1.2 หลักการคำนวณของการวิเคราะห์องค์ประกอบหลัก

วิธีหาปริภูมิใบหน้าไอเกน (Eigenface) ที่นำเสนอโดย Turk และ Pentland [1] เป็นวิธีที่ประยุกต์มาจากวิธี PCA ซึ่งนำมาใช้ในการรู้จำใบหน้า โดยการนำภาพใบหน้าแต่ละภาพมากำหนดเป็นจุดในปริภูมิภาพใบหน้า ปริภูมิภาพใบหน้าจะมีมิติเท่ากับจำนวนพิกเซลของภาพ เช่น ภาพระดับสีเทามีขนาดเท่ากับ $m \times n$ จะได้มิติของปริภูมิภาพคือ D ซึ่ง D จะเท่ากับ $m \times n$ แล้วทำการลดมิติไปเป็นปริภูมิที่มีมิติน้อยลงเพื่อหาพีเจอร์ของภาพ ซึ่งถูกเรียกว่าปริภูมิใบหน้าไอเกน (Eigenface Space)

วิธีการหาปริภูมิใบหน้าไอเกน ทำโดยคำนวณหาไอเกนเวกเตอร์ของเมตริกความแปรปรวนของภาพที่ใช้สอน ที่สมนัยกับค่าไอเกนที่มีค่าความแปรปรวนจากมากไปน้อย เมื่อนำภาพที่ใช้สอนไปฉายด้วยไอเกนเวกเตอร์ที่ได้ จะได้ปริภูมิใบหน้าไอเกน

ในการทดสอบ จะนำภาพที่ใช้สอนทั้งหมดฉายด้วยวิธี Eigenface (PCA) จะได้ปริภูมิใบหน้าไอเกนที่ใช้สอน แล้วนำภาพที่ใช้ในการทดสอบมาฉายด้วยไอเกนเวกเตอร์ของภาพที่ใช้ในการสอน จะได้ปริภูมิใบหน้าไอเกนที่ใช้ในการทดสอบ ในการจำแนกภาพใบหน้าทำโดยหาระยะทางของปริภูมิที่ใช้สอนกับปริภูมิที่ใช้ทดสอบ

วิธีการคำนวณ เมื่อกำหนดให้ภาพที่ใช้สอน M ภาพ คือ A_1, A_2, \dots, A_M แต่ละภาพมีขนาด $m \times n$ จุดภาพ โดยที่แต่ละภาพจะถูกเปลี่ยนให้เป็นเวกเตอร์ r ตามรูปที่ 3.2 จากรูปที่ 3.2 นำแต่ละแถวของภาพมาเรียงต่อกันจากบนลงล่างให้เป็นเวกเตอร์ ดังนั้นจะได้แต่ละเวกเตอร์จะมีขนาด $1 \times D$ ($D = m \times n$)



รูปที่ 3.2 ตัวอย่างภาพที่เปลี่ยนจากเมตริกภาพเป็นเวกเตอร์ภาพ

หลังจากภาพที่ใช้สอนทั้งหมดถูกเปลี่ยนให้เป็นเวกเตอร์ทั้งหมดแล้ว นำเวกเตอร์ทั้งหมดมาเรียงตามตาม ลำดับนี้ ได้เป็นเมตริก องค์ประกอบตัวอย่างที่ใช้ในการสอนตามสมการ

$$t = [t_1, t_2, \dots, t_M] \quad (3.5)$$

คำนวณค่าเฉลี่ยของภาพใบหน้า (Mean Face) เพื่อหาจุดศูนย์กลางของข้อมูลใหม่

$$\bar{t} = \frac{1}{M} \sum_{i=1}^M t_i \quad (3.6)$$

คำนวณหาภาพที่ถูกลบด้วยค่าเฉลี่ยของภาพใบหน้า (Mean subtracted image)

$$T_i = t_i - \bar{t} \quad (3.7)$$

สร้างเมตริกที่ใช้ในการคำนวณหาเมตริกความแปรปรวนร่วม

$$T = [T_1, T_2, \dots, T_M] \quad (3.8)$$

คำนวณหาเมตริกความแปรปรวนร่วม (Covariance Matrix) เพื่อวัดการกระจายของข้อมูล

$$S = T.T^T = \frac{1}{M} \sum_{i=1}^M T_i T_i^T \quad (3.9)$$

คำนวณหาจำนวนค่าไอเกนเวกเตอร์ ที่สมนัยกับค่าไอเกน จากเมตริกความแปรปรวนร่วมด้วยสมการ

$$Su = \lambda u \quad (3.10)$$

เรียงลำดับค่าไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปน้อย แล้วเลือกเฉพาะไอเกนเวกเตอร์ที่สัมพันธ์กับค่าไอเกนที่ไม่เท่ากับศูนย์ ดังนั้นจะได้เมตริกใหม่ของไอเกนเวกเตอร์คือ

$$U_{PCA} = [u_1 u_2 u_3 \dots u_{M-1}] \quad (3.11)$$

3.1.3 การหาพีเจอร์ของภาพ (Feature Extraction)

เมื่อได้เวกเตอร์ที่เหมาะสมที่ใช้ในการฉายภาพของ PCA คือ $U_{PCA} = [u_1, u_2, \dots, u_{M-1}]$ โดยที่ d เป็นค่าที่เลือกระหว่าง 1 ถึง $M-1$ และถ้ามีภาพที่ใช้สอนทั้งหมด M ภาพ จะสามารถหาพีเจอร์ของภาพที่ใช้สอนทั้งหมดจากสมการ

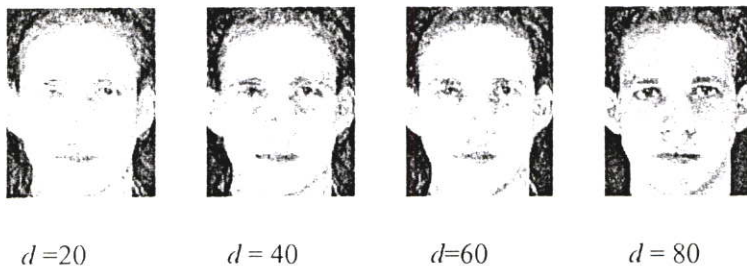
$$Y_k = U_d^T (t_k - \bar{t}); k = 1, 2, \dots, M \quad (3.12)$$

3.1.4 การสร้างภาพใหม่ (Reconstruction)

วิธี PCA สามารถนำเอาพีเจอร์ของภาพและแกนการฉายที่เหมาะสมที่สุดของ PCA มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

โดยที่ $Y = [y_1, \dots, y_d]$ เป็นพีเจอร์ของภาพ และ $U_d = [u_1, u_2, \dots, u_d]$ เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นจาก $y = U_d^T (t - \bar{t})$ และ U_d เป็นออร์โธโนมอล ($U^{-1} = U^T$) จะได้สมการการสร้างภาพใหม่คือ

$$\tilde{t} = U_d y + \bar{t} \quad (3.13)$$



รูปที่ 3.3 ตัวอย่างภาพใบหน้าบุคคลหนึ่งที่สร้างใหม่ด้วยวิธี PCA จากฐานข้อมูล ORL

3.2 การวิเคราะห์การแบ่งแยกเชิงเส้น (LDA)

3.2.1 แนวความคิดของการวิเคราะห์การแบ่งแยกเชิงเส้น

เมื่อวิธี PCA ถูกใช้สำหรับการแทนข้อมูลใหม่ (Representation data) โดยไม่คำนึงถึงการแบ่งแยกข้อมูลระหว่างกลุ่ม แนวคิดของวิธี LDA มาจากความต้องการที่จะฉายข้อมูลจาก D มิติลงบนเส้น เพื่อให้ข้อมูลนั้นมีการแบ่งแยกของกลุ่มที่ดี ถ้าตัวอย่างข้อมูลมีขนาด D มิติ มีการรวมตัวกันแน่น เมื่อทำการฉายลงบนแกนใด ๆ ทำให้เกิดการปนกันของตัวอย่างทุก ๆ กลุ่ม ดังนั้นจึงทำให้กระบวนการรู้จำแย่งลงไปด้วย อย่างไรก็ตามการแก้ปัญหาสามารถทำได้การทดลองเคลื่อนเส้นไปรอบ ๆ เพื่อดูตำแหน่งของเส้นที่ทำให้การฉายตัวอย่างข้อมูลมีการแบ่งแยกที่ดี

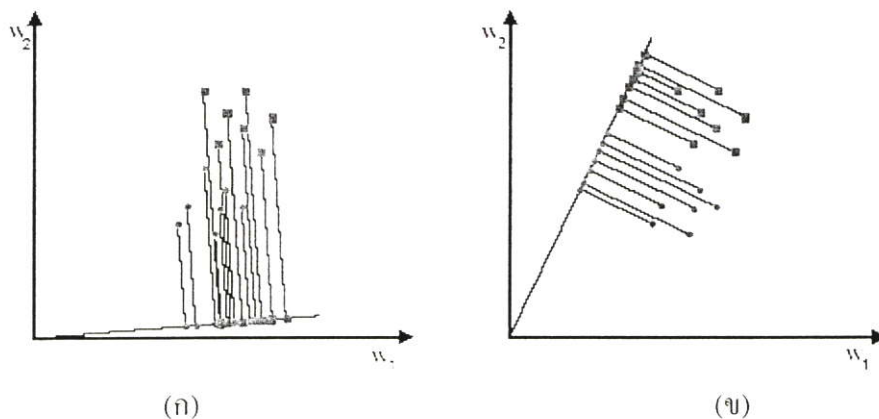
การหาแกนการฉายที่ทำให้ตัวอย่างแบ่งแยกได้ดี ขั้นแรกได้ทำการทดลองกับตัวอย่างเพียง 2 กลุ่มเพื่อดูความสัมพันธ์ แล้วหลังจากนั้นจึงขยายเป็นจำนวนกลุ่มที่มากขึ้น

เมื่อกำหนดให้จำนวนตัวอย่างทั้งหมดคือ t_1, t_2, \dots, t_M แต่ละตัวอย่างมีขนาด D มิติ โดยถูกแบ่งออกเป็นสองกลุ่ม T_1 และ T_2 กลุ่มที่ 1 มีจำนวนตัวอย่าง N_1 และกลุ่มที่ 2 มีจำนวนตัวอย่าง N_2 จากหลักการรวมกันเชิงเส้น (Linear Combination) จะได้สมการ

$$Y = W^T t \quad (3.14)$$

จะได้ y_1, y_2, \dots, y_M เป็นตัวอย่างที่ถูกฉายซึ่งถูกแบ่งเป็นสองกลุ่มคือ \tilde{T}_1 และ \tilde{T}_2 โดยที่แต่ละค่าของ y_M มาจากการฉายค่าของ t_M ลงบนเส้นในทิศทางของ w

จากแนวความคิดของวิธีการนี้คือต้องการให้ตัวอย่างในกลุ่ม T_1 ที่ถูกฉายตกลงบนเส้นให้อยู่ในกลุ่มเดียวกัน ส่วนตัวอย่างที่เหลือก็ให้ตกลงบนเส้นแล้วมาอยู่ในกลุ่มที่เหลือจึงจะทำให้การแบ่งแยกมีประสิทธิภาพ ดังนั้นสิ่งที่สำคัญที่สุดคือการเลือกแกนการฉาย w รูปที่ 3.4 แสดงผลของการเลือกค่าของ w ที่ใช้ในการฉายของตัวอย่างภาพ 2 มิติ จากรูปจะเห็นว่าเมื่อเลือกแกนการฉาย w เหมือนรูปที่ 3.4 (ก) จะทำให้ตัวอย่างที่ถูกฉายของสองกลุ่มมีการปนกัน แต่ถ้าเลือกแกนการฉาย w เหมือนรูปที่ 3.4 (ข) จะทำให้กลุ่มตัวอย่างมีการแบ่งแยกกันอย่างชัดเจน



รูปที่ 3.4 ผลของการเลือกค่าของ w ที่ใช้ในการฉายภาพ

(ก) การเลือกค่าของ w ที่ทำให้ตัวอย่างที่ถูกฉายของสองกลุ่มมีการปนกัน

(ข) การเลือกค่าของ w ที่ทำให้ตัวอย่างที่ถูกฉายมีการแบ่งแยกกันอย่างชัดเจน

วิธีการที่จะหาทิศทางการฉายที่ดีที่สุด โดยทดลองวัดการแบ่งแยกของกลุ่มจากผลต่างของค่าเฉลี่ยตัวอย่างที่ถูกฉายในแต่ละกลุ่ม โดยที่ผลต่างของค่าเฉลี่ยตัวอย่างที่ถูกฉายมีค่ามากจะแสดงให้เห็นว่าข้อมูลของแต่ละกลุ่มตัวอย่างที่ถูกฉายห่างกันมาก ถ้ากำหนดให้ \bar{t}_i คือ ค่าเฉลี่ยแต่ละกลุ่มตัวอย่าง และ \tilde{m}_i คือค่าเฉลี่ยแต่ละกลุ่มของตัวอย่างที่ถูกฉาย

$$\bar{t}_i = \frac{1}{N_i} \sum_{t \in T_i} t \quad (3.15)$$

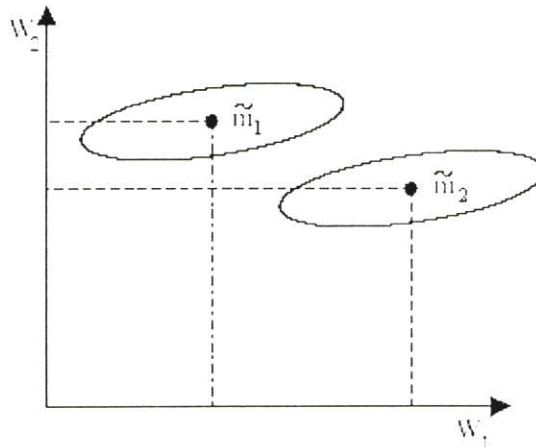
และ

$$\tilde{m}_i = \frac{1}{N_i} \sum_{y \in \tilde{T}_i} y \quad (3.16)$$

$$= \frac{1}{N_i} \sum_{t \in T_i} w^T t \quad (3.17)$$

จากการทดลองเลือกระยะทางระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉายเป็นสมการวัตถุประสงค์ (Objective function) จะได้สมการวัตถุประสงค์เป็น

$$J(w) = |\tilde{m}_1 - \tilde{m}_2| = |w^T (\bar{t}_1 - \bar{t}_2)| \quad (3.18)$$



รูปที่ 3.5 การเลือกแกนการฉายจากระยะทางระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉาย 2 กลุ่ม

จากรูปที่ 3.5 เมื่อเลือกแกนการฉายเป็น w_2 ตัวอย่างที่ถูกฉายลงบนแกนนี้จะมีการปนกันน้อยกว่าเมื่อเลือกแกนการฉายเป็น w_1 ทั้งที่ระยะห่างระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉายของแกนการฉาย w_2 มีค่าน้อยกว่าแกนการฉาย w_1 ซึ่งแสดงให้เห็นว่าการใช้ระยะห่างระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉายเพียงอย่างเดียวยังไม่เพียงพอที่จะทำให้การฉายตัวอย่างมีการแบ่งแยกที่ดี

Fisher [2] ได้นำเสนอการแก้ปัญหานี้ โดยการหาค่ามากที่สุดของฟังก์ชันอัตราส่วนของระยะทางระหว่างค่าเฉลี่ยของตัวอย่างที่ถูกฉายกับผลรวมการกระจายภายในกลุ่มของตัวอย่างที่ถูกฉาย เมื่อการกระจายภายในกลุ่มจะมีค่าเท่ากับค่าความแปรปรวนของตัวอย่างภายในกลุ่ม โดยที่ค่าการกระจายของแต่ละกลุ่มตัวอย่างที่ถูกฉายคือ

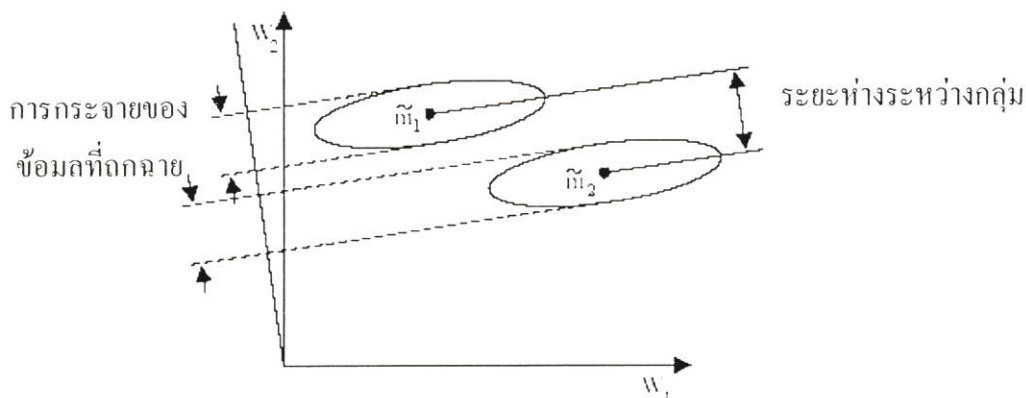
$$\tilde{S}_i^2 = \sum_{y \in T_i} (y - \tilde{m}_i)^2 \quad (3.19)$$

และ $\tilde{S}_1^2 + \tilde{S}_2^2$ ถูกเรียกว่า การกระจายภายในกลุ่มของตัวอย่างที่ถูกฉาย

ดังนั้นสมการการแบ่งแยกเชิงเส้นที่ได้จากการแปลงเชิงเส้น $w^T t$ ซึ่งมีค่ามากที่สุดของข้อมูล 2 กลุ่มจะได้สมการวัตถุประสงค์เป็น

$$J(w) = \frac{|\tilde{m}_1^2 + \tilde{m}_2^2|}{\tilde{S}_1^2 + \tilde{S}_2^2} \quad (3.20)$$

เมื่อสมการที่ (3.20) มีค่ามากที่สุดจะทำให้การแบ่งแยกระหว่างข้อมูล 2 กลุ่มทำได้ดังรูปที่ 3.7



รูปที่ 3.6 แสดงตัวอย่างการฉายภาพที่มีการแบ่งแยกได้ดีของตัวอย่าง 2 กลุ่ม

จากรูปที่ 3.6 จะเห็นว่าเมื่อใช้ค่าการกระจายภายในกลุ่มมาคิดคำนวณด้วย ทำให้การฉายข้อมูลของ 2 กลุ่มแบ่งแยกกันอย่างชัดเจน โดยตัวอย่างที่ถูกฉายในกลุ่มเดียวกันจะรวมตัวอยู่ใกล้ ๆ กัน ขณะที่ยังคงไว้ซึ่งระยะห่างมาก ๆ ของค่าเฉลี่ยแต่ละกลุ่ม

เมื่อได้แนวคิดในการแบ่งแยกข้อมูลที่ดีแล้ว ลำดับต่อมาที่หาวิธีในการหาแกนการฉายแกนของตัวอย่างข้อมูล โดยสามารถหามาจากเมตริกการกระจายภายในกลุ่มและระหว่างกลุ่มของตัวอย่าง กำหนดให้เมตริกการกระจายภายในกลุ่มคือ

$$S_i = \sum_{t \in T_i} (t - \bar{t}_i)(t - \bar{t}_i)^T \quad (3.21)$$

แต่ในกรณีนี้พิจารณาแค่ 2 กลุ่มจะได้เมตริกการกระจายภายในกลุ่มเป็น

$$S_w = S_1 + S_2 \quad (3.22)$$

ความสัมพันธ์ของเมตริกการกระจายภายในกลุ่มของตัวอย่างที่ถูกฉายกับเมตริกการกระจายภายในกลุ่มของตัวอย่างคือ

$$\begin{aligned} \bar{S}_i^2 &= \sum_{y \in T_i} (y - \bar{m}_i)^2 \\ &= \sum_{t \in T_i} (w^T t - w^T \bar{t}_i)^2 \\ &= \sum_{t_k \in T_i} w^T (t_k - \bar{t}_i)(t_k - \bar{t}_i)^T w \\ &= w^T S_i w \end{aligned} \quad (3.23)$$

ดังนั้นผลรวมของเมตริกการกระจายของข้อมูลที่ถูกลาย

$$\tilde{S}_1^2 + \tilde{S}_2^2 = w^T S_W w \quad (3.24)$$

ในวิธีเดียวกันผลต่างของค่าเฉลี่ยที่ถูกลายสามารถแสดงในรูปผลต่างของค่าเฉลี่ยตัวอย่างได้คือ

$$\begin{aligned} |\tilde{m}_1 - \tilde{m}_2|^2 &= |w^T \bar{I}_1 - w^T \bar{I}_2|^2 \\ &= w^T (\bar{I}_1 - \bar{I}_2)(\bar{I}_1 - \bar{I}_2)^T w \\ &= w^T S_B w \end{aligned} \quad (3.25)$$

ซึ่ง S_W ถูกเรียกว่าเมตริกการกระจายในกลุ่มและ S_B เรียกว่าเมตริกการกระจายระหว่างกลุ่ม โดยที่เมตริกทั้งสองจะมีขนาดสมมาตร

ดังนั้นจากสมการที่ 3.20 สามารถแสดงสมการในตัวแปรของ S_W และ S_B เป็น

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (3.26)$$

ในการแก้สมการ (3.26) เพื่อหาเกณฑ์การฉาย โดยการใช้องค์การหา Differentiate เทียบกับ w และกำหนดให้มีค่าเท่ากับศูนย์จะได้

$$S_B w = \lambda S_W w \quad (3.27)$$

เมื่อ S_W ไม่เป็นเอกฐาน (nonsingular) เราสามารถเปลี่ยนสมการใหม่เป็น

$$S_W^{-1} S_B w = \lambda w \quad (3.28)$$

แต่ในกรณีนี้พิจารณาจำนวนกลุ่มเพียง 2 กลุ่ม จึงไม่จำเป็นจะต้องแก้สมการที่ (3.28) โดยการหาค่าไอเกนเวกเตอร์และค่าไอของ $S_W^{-1} S_B$ เพราะในความเป็นจริง $S_B w$ จะมีทิศทางเดียวกับ $(\bar{I}_1 - \bar{I}_2)$ ดังนั้นการหาค่า w ที่เหมาะสมที่สุดในกรณีนี้คือ

$$w = S_W^{-1} (\bar{I}_1 - \bar{I}_2) \quad (3.29)$$

ดังนั้นจะได้แผนการฉาย w ที่มาจากการแบ่งแยกเชิงเส้นสามารถคำนวณได้โดยการหาค่าอัตราส่วนที่มากที่สุดของการกระจายภายในกลุ่มและการกระจายระหว่างกลุ่ม จากการทดลองใช้ข้อมูล 2 กลุ่มมาทำการฉายจะพบว่าข้อมูลของตัวอย่าง D มิติจะสามารถมิติลงเป็น 1 มิติได้

หลังจากทดลองใช้ข้อมูล 2 กลุ่มมาทดลองหาแผนการฉายที่เหมาะสมที่สุด ซึ่งจะแสดงให้เห็นความสัมพันธ์และวิธีในการหาแผนการฉาย ลำดับต่อมาจึงขยายจำนวนกลุ่มของตัวอย่างข้อมูลเป็นจำนวน L กลุ่ม เป็นผลทำให้สมการการแบ่งแยกเชิงเส้นจะประกอบไปด้วยฟังก์ชันการแบ่งแยกทั้งหมด L

จากการทดลองใช้ตัวอย่าง 2 กลุ่มในการลดมิติจะมีเมตริกที่ใช้ในการหาแผนการฉายอยู่สองเมตริกคือ เมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม ดังนั้นในกรณีของ L กลุ่ม ก็สามารถหาเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มเป็นเมตริกการกระจายภายในกลุ่มคือ

$$S_w = \sum_{i=1}^L S_i \quad (3.30)$$

เมื่อการกระจายของแต่ละกลุ่มและค่าเฉลี่ยของแต่ละกลุ่มเป็น

$$S_i = \sum_{t \in T_i} (t - \bar{t}_i)(t - \bar{t}_i)^T \quad \text{และ} \quad \bar{t}_i = \frac{1}{N_i} \sum_{t \in T_i} t \quad (3.31)$$

และสามารถหาเมตริกการกระจายระหว่างกลุ่มได้คือ

$$S_B = \sum_{i=1}^L N_i (\bar{t}_i - \bar{t})(\bar{t}_i - \bar{t})^T \quad (3.32)$$

โดยที่ \bar{t} คือ ค่าเฉลี่ยของตัวอย่างทั้งหมด

$$\bar{t} = \frac{1}{m} \sum_{i=1}^M t_i \quad (3.33)$$

การฉายข้อมูลจากปริภูมิ D มิติ ไปเป็นปริภูมิ $L-1$ มิติด้วยวิธีการแบ่งแยกเชิงเส้นจะประกอบไปด้วยฟังก์ชันการแบ่งแยกทั้งหมด $L-1$

$$y_i = w_i^T t \quad i = 1, \dots, L-1 \quad (3.34)$$

เปลี่ยนสมการให้อยู่ในรูปของเมตริกจะได้

$$Y = W^T t \quad (3.35)$$

เมื่อตัวอย่างในการฉายมีความสัมพันธ์กันกับตัวอย่างที่ถูกฉาย ดังนั้นจึงสามารถหาเวกเตอร์ค่าเฉลี่ยและเมตริกการกระจายได้ดังนี้

$$\tilde{m}_i = \frac{1}{N_i} \sum_{y \in T_i} y \quad (3.36)$$

$$\tilde{m} = \frac{1}{m} \sum_{i=1}^L N_i \tilde{m}_i \quad (3.37)$$

$$\tilde{S}_W = \sum_{i=1}^L \sum_{y \in T_i} (y - \tilde{m}_i)(y - \tilde{m}_i)^T \quad (3.38)$$

$$\tilde{S}_B = \sum_{i=1}^L N_i (\tilde{m}_i - \tilde{m})(\tilde{m}_i - \tilde{m})^T \quad (3.39)$$

จากแนวความคิดของตัวอย่างที่ถูกแบ่งเป็น 2 กลุ่มจะได้ความสัมพันธ์ สามารถหาเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มของตัวอย่างที่ถูกฉายได้คือ

$$\tilde{S}_W = w^T S_W w \quad (3.40)$$

$$\tilde{S}_B = w^T S_B w \quad (3.41)$$

ดังนั้นความสัมพันธ์ระหว่างเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มที่ใช้ในการฉายข้อมูลให้มีปริภูมิมีขนาดลดลงสามารถทำได้โดย หาค่าส่วนที่มากที่สุดของเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม แต่เนื่องจากการวัดการกระจายต้องมีขนาดเป็นสเกลจึงจำเป็นต้องใช้การหาดีเทอร์มิแนนซ์ (determinant) เข้ามาช่วย ดังนั้นสมการวัดประสิทธิภาพคือ

$$J(w) = \frac{|\tilde{S}_B|}{|\tilde{S}_W|} = \frac{|w^T S_B w|}{|w^T S_W w|} \quad (3.42)$$

การแก้สมการ (3.42) จะใช้การหาโอเกนเวกเตอร์ที่สมนัยกับค่าโอเกนที่เรียงจากมากไปน้อยตามลำดับของสมการที่ (3.43)

$$S_B w_i = \lambda_i S_W w_i \quad (3.43)$$

เมื่อเรียงลำดับค่าไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปน้อย แล้วเลือกเฉพาะไอเกนเวกเตอร์ที่สัมพันธ์กับค่าไอเกนที่ไม่เท่ากับศูนย์จะได้เมตริกใหม่ของไอเกนเวกเตอร์คือ

$$W_{LDA} = [w_1 w_2 w_3 \dots w_L] \quad (3.44)$$

ดังนั้นจะได้การฉายภาพจากปริภูมิ D มิติ ไปเป็นปริภูมิ L มิติ ที่ทำให้อัตราส่วนระหว่างเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มมีค่ามากที่สุด

3.2.2 หลักการคำนวณของการวิเคราะห์การแบ่งแยกเชิงเส้น

วิธี LDA สามารถนำมาใช้ในการรู้จำใบหน้า โดยการนำภาพใบหน้าแต่ละภาพมากำหนดเป็นจุดในปริภูมิภาพใบหน้า ปริภูมิภาพใบหน้าจะมีมิติเท่ากับจำนวนพิกเซลของภาพ เช่น ภาพระดับสีเทามีขนาดเท่ากับ $m \times n$ จะได้มิติของปริภูมิภาพคือ D ซึ่ง D จะเท่ากับ $m \times n$ แล้วทำการลดมิติไปเป็นปริภูมิที่มีมิติน้อยลงเพื่อหาฟีเจอร์ของภาพ

วิธี LDA ทำโดยคำนวณหาไอเกนเวกเตอร์จากเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่มของภาพที่ใช้สอน ที่สมนัยกับค่าไอเกนที่มีค่าเรียงลำดับจากมากไปน้อย เมื่อนำภาพที่ใช้สอนไปฉายด้วยไอเกนเวกเตอร์ที่ได้ จะได้ปริภูมิที่มีมิติน้อยลง

ในการทดสอบ จะนำภาพที่ใช้สอนทั้งหมดฉายด้วยวิธี LDA จะได้ปริภูมิใบหน้าที่ใช้สอน แล้วนำภาพที่ใช้ในการทดสอบมาฉายด้วยไอเกนเวกเตอร์ของภาพที่ใช้ในการสอน จะได้ปริภูมิใบหน้าที่ใช้ในการทดสอบ ในการแยกแยะภาพใบหน้าทำโดยหาระยะทางของปริภูมิที่ใช้สอนกับปริภูมิที่ใช้ทดสอบ

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ ดังนั้น $A_j (j = 1, 2, \dots, M)$ แต่ละภาพมีขนาด $m \times n$ จุดภาพ โดยที่แต่ละภาพถูกเปลี่ยนให้เป็นเวกเตอร์ t เมื่อแต่ละเวกเตอร์ t มีขนาด $1 \times D$ ($D = m \times n$) จัดแบ่งเป็นภาพทั้งหมดออกเป็น $T_i (i = 1, 2, \dots, L)$ โดยที่ L คือจำนวนกลุ่ม \bar{t}_i คือค่าเฉลี่ยที่คำนวณจากทุกภาพ \bar{t}_i คือค่าเฉลี่ยที่คำนวณเฉพาะภาพภายในกลุ่ม T_i ซึ่งแต่ละกลุ่มมี N_i ภาพ การคำนวณกระทำได้ตามขั้นตอนดังนี้

คำนวณเมตริกการกระจายภายในกลุ่ม

เมตริกการกระจายภายในกลุ่มจะเป็นการวัดค่าการกระจายระหว่างภาพภายในแต่ละกลุ่ม เมื่อมีภาพทั้งหมด L กลุ่ม ดังนั้นการคำนวณจะใช้การหาผลรวมของเมตริกความแปรปรวน (Covariance matrix) ในแต่ละกลุ่มมาคำนวณหา

$$S_w = \sum_{i=1}^L \sum_{t \in T_i} (t - \bar{t}_i)(t - \bar{t}_i)^T \quad (3.45)$$

คำนวณเมตริกการกระจายระหว่างกลุ่ม

เมตริกการกระจายระหว่างกลุ่มจะเป็นการวัดค่าการกระจายระหว่างกลุ่ม สามารถคำนวณหาโดยการหาผลรวมเมตริกความแปรปรวนของความแตกต่างระหว่างค่าเฉลี่ยรวมกับค่าเฉลี่ยของแต่ละกลุ่ม

$$S_B = \sum_{i=1}^L N_i (\bar{t}_i - \bar{t})(\bar{t}_i - \bar{t})^T \quad (3.46)$$

คำนวณค่าไอเกนเวกเตอร์และค่าไอเกน

คำนวณหาค่าไอเกนเวกเตอร์และค่าไอเกนของเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม

$$S_B w = \lambda S_{Bw} \quad (3.47)$$

เลือกค่าไอเกนเวกเตอร์

เรียงไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปน้อยและเลือกไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่ไม่เท่ากับศูนย์ ดังนั้นไอเกนเวกเตอร์จะมีทั้งหมด L

$$W_{LDA} = [w_1, w_2, \dots, w_L] \quad (3.48)$$

3.2.3 การหาพีเจอร์ของภาพ

เมื่อได้เวกเตอร์ที่เหมาะสมที่ใช้ในการฉายภาพของ LDA คือ $W_{LDA} = [w_1, w_2, \dots, w_d]$ โดยที่ d เป็นค่าที่เลือกระหว่าง 1 ถึง $L-1$ และมีภาพที่ใช้สอนทั้งหมด M ภาพ จะสามารถหาพีเจอร์ของภาพที่ใช้สอนทั้งหมดจากสมการ

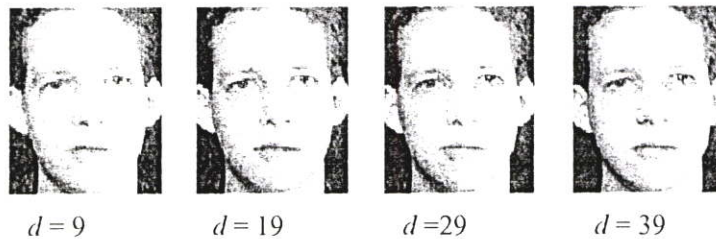
$$Y_k = W_d^T (t_k - \bar{t}); k = 1, 2, \dots, M \quad (3.49)$$

3.2.4 การสร้างภาพใหม่

วิธีของ LDA สามารถนำเอาพีเจอร์ของภาพและแกนการฉายที่เหมาะสมที่สุดของวิธีการ LDA มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

โดยที่ $Y = [y_1, \dots, y_d]$ เป็นพีเจอร์ของภาพ และ $U_d = [u_1, u_2, \dots, u_d]$ เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นจาก $Y = W_d^T (t - \bar{t})$ และ W_d เป็นอโรโธมอด ($W^{-1} = W^T$) จะได้สมการการสร้างภาพใหม่คือ

$$\tilde{t} = W_d y + \bar{t} \quad (3.50)$$



รูปที่ 3.7 ตัวอย่างภาพใบหน้าบุคคลหนึ่งทีสร้างขึ้นด้วยวิธี LDA จากฐานข้อมูล ORL

3.3 Two – Dimensional Principal Component Analysis (2DPCA)

เมื่อนำหลักการของวิธี PCA มาใช้ในการรู้จำใบหน้าที่เป็นภาพสองมิติ จะเห็นว่า ก่อนการคำนวณจำเป็นต้องเปลี่ยนจากเมตริกภาพให้เป็นเวกเตอร์ภาพทำให้มิติมีค่ามาก เมื่อนำมาหาค่าเมตริกการกระจายรวมเป็นผลให้เมตริกมีขนาดใหญ่ ทำให้การคำนวณหาค่าไอเกนเวกเตอร์ที่ใช้ในการฉายจึงทำได้ยากและใช้เวลาในการคำนวณมาก

เทคนิคใหม่ที่ถูกนำเสนอโดย Yang สามารถนำภาพที่ใช้สอนมาคำนวณได้โดยตรง วิธีการนี้มีชื่อว่า Two – Dimensional Principal Component Analysis (2DPCA) โดยที่ความแตกต่างระหว่าง PCA และ 2DPCA คือ สมการการแปลงเชิงเส้น ซึ่งวิธี 2DPCA จะใช้สมการ $Y = AX$ โดยที่ X คือ เมตริกการฉาย และ A เป็นเมตริกภาพ ขณะที่ PCA จะใช้สมการ $Y = U^T t$ โดยที่ U คือ แกนการฉาย และ เป็นเวกเตอร์ภาพ เป็นผลให้ภาพที่ใช้ในการสอนไม่จำเป็นต้องเปลี่ยนจากเมตริกภาพเป็นเวกเตอร์ภาพและในการคำนวณหาค่าเมตริกการกระจายรวมของภาพสามารถนำภาพที่ใช้สอนมาคำนวณหาได้ตรง ซึ่งเป็นผลทำให้ขนาดของเมตริกการกระจายรวมของภาพที่ใช้วิธี 2DPCA มีขนาดน้อยกว่า วิธี PCA จึงทำให้เวลาในการคำนวณหาค่าของไอเกนเวกเตอร์มีค่าน้อยมาก หลักการคำนวณของวิธี 2DPCA คือการคำนวณหาเมตริกความแปรปรวนร่วมจากภาพที่ใช้ในการสอนทั้งหมด แล้วคำนวณหาค่าไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่เรียงลำดับจากมากไปน้อย แล้วจึงนำมาฉายภาพใบหน้าให้มีขนาดลดลง เช่น ภาพใบหน้ามีขนาด 112×92 จุดภาพ เมื่อเลือกไอเกนเวกเตอร์ 12 มิติ พีเจอร์ของภาพใบหน้าใหม่จะมีขนาดเป็น 112×12 จุดภาพ ซึ่งจะเห็นว่าจำนวนมิติของคอลัมน์มีค่าลดลง

3.3.1 แนวความคิดของวิธี 2DPCA

กำหนดให้ X ประกอบด้วยเวกเตอร์ขนาด n มิติ และ A คือภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนแกน X ตามหลักการแปลงเชิงเส้น จะได้

$$Y = AX \quad (3.51)$$

ดังนั้นจะได้ Y ประกอบด้วยเวกเตอร์ขนาด m มิติ โดยแต่ละเวกเตอร์ของ Y คือพีเจอร์
เวกเตอร์ของภาพ A

วิธีในการหาเวกเตอร์ X ที่ใช้ในการฉายที่คี่นั้น สามารถคำนวณจากการหาผลรวมค่าไอเกน
(Trace) ของเมตริกความแปรปรวนร่วมของพีเจอร์เวกเตอร์ Y จากแนวความคิดนี้ สามารถสร้างเป็น
หลักการได้ตามสมการ

$$J(x) = tr(S_x) \quad (3.52)$$

เมื่อ S_x คือเมตริกความแปรปรวนร่วมของพีเจอร์เวกเตอร์ Y และ $tr(S_x)$ คือการหาผลรวม
ค่าไอเกนของเมตริก S_x

ลักษณะที่สำคัญของการหาค่ามากที่สุด (Maximize) ในสมการที่ (3.52) ก็เพื่อหาทิศทางการ
ฉาย X ที่มีความแปรปรวนมากที่สุด โดยที่เมตริกความแปรปรวนร่วม (S_x) สามารถหาจาก

$$\begin{aligned} S_x &= E(Y - E(Y))(Y - E(Y))^T \\ &= [AX - E(AX)] [AX - E(AX)]^T \\ &= E[(A - E(A))X] [(A - E(A))X]^T \end{aligned} \quad (3.53)$$

ถ้า

$$tr(S_x) = X^T [E(A - E(A))^T (A - E(A))] X \quad (3.54)$$

เมื่อกำหนดให้

$$G = E[(A - EA)^T (A - EA)] \quad (3.55)$$

เมตริก G คือ เมตริกความแปรปรวนร่วมของภาพ (เมตริกการกระจายของภาพ) ที่มีขนาด $n \times n$

เมื่อกำหนดให้มีภาพตัวอย่างที่ใช้สอนทั้งหมด M ภาพ เป็นเมตริกที่มีขนาด $m \times n$ ดังนั้น A_j
($j = 1, 2, \dots, M$) และค่าเฉลี่ยภาพของภาพทั้งหมดคือ \bar{A} การคำนวณหาเมตริกความแปรปรวนร่วมของ
ภาพ สามารถนำภาพมาคำนวณได้โดยตรงจาก

$$G = \frac{1}{M} \sum_{j=1}^M [(A_j - \bar{A})^T (A_j - \bar{A})] \quad (3.56)$$

จากสมการที่ (3.52) สามารถแทนค่าสมการที่ (3.54) และสมการที่ (3.56) เป็น

$$J(X) = X^T G X \quad (3.57)$$

สมการที่ (3.57) มีชื่อว่า สมการทั่วไปของผลรวมการกระจาย (generalized total scatter criterion) แกน X ที่ได้จากสมการนี้ที่มีค่ามากที่สุดมีชื่อว่า แกนที่เหมาะสมที่สุดในการฉาย (optimal projection axis) ในการหาค่ามากที่สุดนั้นก็เพื่อต้องการให้การฉายภาพลงบนแกน X แล้วเมตริกใหม่ที่ถูกฉายมีการกระจายตัวมากที่สุด

แกนที่เหมาะสมที่สุดที่ใช้ในการฉาย X_{opt} สามารถหาได้จากการหาค่าไอเกนเวกเตอร์ ของ G ที่สมนัยกับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m ซึ่งแต่ละไอเกนเวกเตอร์เป็นออร์โธโนมอล (Orthonormal)

$$X_{2DPCA} = [X_1 \ X_2 \ \dots \ X_m] \quad (3.58)$$

3.3.2 การหาพีเจอร์ของภาพ

เวกเตอร์ที่เหมาะสมที่ใช้ในการฉายภาพของ 2DPCA คือ X_1, X_2, \dots, X_m ถูกนำมาใช้หาพีเจอร์ของภาพ แต่เนื่องต้องการหาพีเจอร์ของภาพฉะนั้นจึงเลือกแกนในการฉายเพียง 1 ถึง d เท่านั้น เมื่อกำหนดให้ A เป็นภาพตัวอย่างดังนั้นการหาพีเจอร์ของภาพ A คือ

$$Y_k = A X_k, \quad k = 1, 2, \dots, d \quad (3.59)$$

เวกเตอร์ใหม่ที่ได้จากการฉาย Y_1, Y_2, \dots, Y_d มีชื่อว่า เวกเตอร์องค์ประกอบหลัก (principal component vector) ของภาพตัวอย่าง A เมื่อนำเวกเตอร์ขององค์ประกอบหลักรวมเป็นเมตริก $Y = [y_1, y_2, \dots, y_d]$ ที่มีขนาดเป็น $m \times d$ เมตริก Y มีชื่อว่า พีเจอร์เมตริกภาพ A

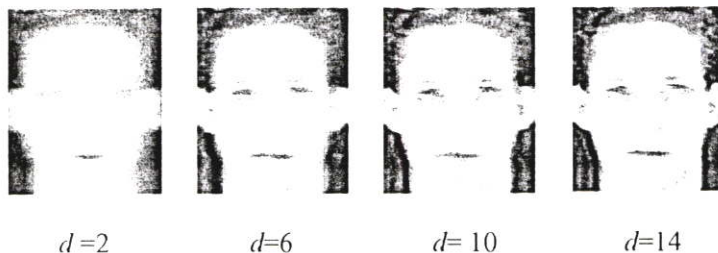
3.3.3 การสร้างภาพใหม่

ในวิธีของ 2DPCA สามารถนำเอาพีเจอร์เมตริกและแกนการฉายที่เหมาะสมที่สุดมาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ $Y = [y_1, \dots, y_d]$ เป็นพีเจอร์เมตริกของภาพ A และ $X = [x_1, \dots, x_d]$ เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นจาก $Y_k = A X_k$ และ X เป็นออร์โธโนมอล ($X^{-1} = X^T$) จะได้สมการการสร้างภาพใหม่คือ

$$\tilde{A} = Y X^T = \sum_{k=1}^d y_k x_k^T \quad (3.60)$$

ในการสร้างภาพใบหน้าใหม่นั้นถ้าเลือกจำนวนแกนการฉาย d เท่ากับ n จะได้ภาพที่สร้างใหม่เหมือนกับภาพ A แต่ถ้าเลือก d น้อยกว่า n ภาพ \tilde{A} จะความเหมือนกับภาพ A น้อยลงตามลำดับ



รูปที่ 3.8 ตัวอย่างภาพใบหน้าที่สร้างใหม่ด้วยวิธี 2DPCA จากฐานข้อมูล ORL

3.3.4 หลักการคำนวณของวิธี 2DPCA

เมื่อกำหนดให้มีภาพที่ใช้สอน M ภาพ คือ A_1, A_2, \dots, A_M แต่ละภาพมีขนาด $m \times n$ จุดภาพ คำนวณหาค่าเฉลี่ยของภาพใบหน้า (Mean Face)

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i$$

คำนวณหาเมตริกความแปรปรวนร่วม (Covariance Matrix)

$$G = \frac{1}{M} \sum_{j=1}^M [(A_j - \bar{A})^T (A_j - \bar{A})]$$

คำนวณหาค่าไอเกนเวกเตอร์ ที่สมนัยกับค่าไอเกน จากเมตริกความแปรปรวนร่วมด้วยสมการ

$$GX = \lambda X$$

เรียงลำดับค่าไอเกนเวกเตอร์ ที่สมนัยกับค่าไอเกนจากมากไปน้อย แล้วเลือกแกนการฉายมา d มิติ ดังนั้นจะได้เมตริกใหม่ของไอเกนเวกเตอร์ คือ

$$X = [x_1, x_2, x_3, \dots, x_d]$$

การหาพีเจอร์ของภาพ

$$Y_k = AX_k, k = 1, 2, \dots, d$$

การสร้างภาพใหม่

$$\tilde{A} = YX^T_A = \sum_{k=1}^d y_k x_k^T$$

3.4 Two – Dimensional Linear Discriminant Analysis (2DLDA)

วิธีการ 2DLDA เป็นวิธีการใหม่ที่ประยุกต์มาจากวิธี LDA และ 2DPCA โดยสามารถนำเมตริกภาพมาคำนวณหาเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม แล้วนำมาคำนวณหาไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปน้อยตามลำดับ จากนั้นนำภาพใบหน้ามาฉายลงบนไอเกนเวกเตอร์ จะได้ฟีเจอร์เมตริกที่มีจำนวนมิติของคอลัมน์ภาพน้อยลง

3.4.1 แนวความคิดของวิธี 2DLDA

กำหนดให้ X ประกอบด้วยเวกเตอร์ขนาด n มิติ และ A เป็นภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนแกน X ตามหลักการแปลงเชิงเส้นจะได้

$$Y = AX \quad (3.61)$$

ดังนั้นจะได้ Y ประกอบด้วยเวกเตอร์ขนาด m มิติ โดยแต่ละเวกเตอร์ของ Y คือฟีเจอร์เวกเตอร์ของภาพ A

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ A_j ($j = 1, 2, \dots, M$) จัดแบ่งเป็นกลุ่ม T_i ($i = 1, 2, \dots, L$) โดยที่ L คือจำนวนกลุ่ม \bar{A} คือค่าเฉลี่ยที่คำนวณจากทุกภาพ \bar{A}_i คือค่าเฉลี่ยที่คำนวณเฉพาะภาพภายในกลุ่ม T_i ซึ่งแต่ละกลุ่มมี N_i ภาพ และ P_i คือกลุ่ม T_i ที่ถูกฉายลงบนแกน X

หลังจากภาพที่ใช้ฝึกสอนถูกฉายตามสมการที่ (3.61) ลงบนแกน X จะได้ฟีเจอร์เมตริกเป็น

$$Y_j = A_j X, \quad j = 1, 2, \dots, M \quad (3.62)$$

วิธีในการหาแกนการฉาย X สามารถคำนวณได้โดยการหาผลรวมค่าไอเกน (Trace) ของเมตริกความแปรปรวนร่วมของภาพตัวอย่างที่ถูกฉาย จากแนวความคิดนี้สามารถสร้างเป็นสมการได้คือ

$$J(X) = \frac{P_B}{P_W} \quad (3.63)$$

เมื่อ P_B และ P_W คำนวณจาก

$$P_B = n(\tilde{S}_B)$$

$$P_W = n(\tilde{S}_W)$$

โดยที่

\tilde{S}_B คือ เมตริกการกระจายระหว่างกลุ่มของภาพตัวอย่างที่ถูกฉาย

\tilde{S}_W คือ เมตริกการกระจายภายในกลุ่มของภาพตัวอย่างที่ถูกฉาย

$$\begin{aligned}\tilde{S}_B &= \sum_{i=1}^L N_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T \\ &= \sum_{i=1}^L N_i [(\bar{A}_i - \bar{A})X][(\bar{A}_i - \bar{A})X]^T\end{aligned}\quad (3.64)$$

$$\begin{aligned}\tilde{S}_W &= \sum_{i=1}^L \sum_{Y_k \in P_i} (Y_k - \bar{Y}_i)(Y_k - \bar{Y}_i)^T \\ &= \sum_{i=1}^L \sum_{Y_k \in P_i} [(A_k - \bar{A}_i)X][(A_k - \bar{A}_i)X]^T\end{aligned}\quad (3.65)$$

ถ้า

$$\begin{aligned}n(\tilde{S}_B) &= X^T \left[\sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A}) \right] X \\ &= X^T S_B X\end{aligned}\quad (3.66)$$

$$\begin{aligned}n(\tilde{S}_W) &= X^T \left[\sum_{i=1}^L \sum_{Y_k \in P_i} (A_k - \bar{A}_i)^T (A_k - \bar{A}_i) \right] X \\ &= X^T S_W X\end{aligned}\quad (3.67)$$

ในการคำนวณหา S_B และ S_W สามารถนำภาพที่ใช้ฝึกสอนมาคำนวณหาได้โดยตรงและจากสมการที่ (3.63) สามารถเปลี่ยนเป็นสมการใหม่ได้คือ

$$J(X) = \frac{X^T S_B X}{X^T S_W X}\quad (3.68)$$

สมการที่ (3.68) มีชื่อว่าสมการหลักการฉายเชิงเส้นที่ยึดถือประเภทเป็นเกณฑ์ (class specific linear projection criterion) โดยเวกเตอร์ X ที่มาจากการหาค่ามากที่สุดของ $J(X)$ มีชื่อว่าเกณฑ์การฉายที่เหมาะสมที่สุด

$$X_{\text{opt}} = \arg \max_X J(X) \quad (3.69)$$

ถ้า S_B^{-1} ไม่เป็นเมตริกเอกฐาน (non singular) สามารถหาเมตริก X โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B X_{\text{opt}} = \lambda S_W X_{\text{opt}} \quad (3.70)$$

โดยปกติเกณฑ์การฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_B^{-1} S_W$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง n ซึ่งแต่ละไอเกนเวกเตอร์เป็นออร์โธโนมอล (Orthonormal)

$$X_{\text{2DLDA}} = [x_1 x_2 \dots x_n] \quad (3.71)$$

3.4.2 การหาพีเจอร์ของภาพ

เวกเตอร์ที่เหมาะสมที่สุดที่ใช้ในการฉายภาพของ 2DLDA คือ x_1, x_2, \dots, x_n ถูกนำมาใช้หาพีเจอร์ของภาพ แต่เนื่องต้องการหาพีเจอร์ของภาพจะนั้นจึงเลือกเกณฑ์ในการฉายเพียง 1 ถึง d เท่านั้น

$$Y_k = AX_k, k = 1, 2, \dots, d \quad (3.72)$$

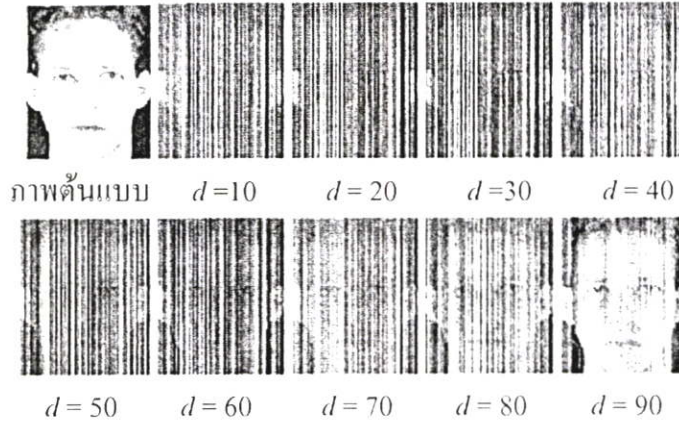
เวกเตอร์ใหม่ที่ได้จากการฉาย y_1, \dots, y_d มีชื่อว่าพีเจอร์เวกเตอร์ เมื่อนำพีเจอร์เวกเตอร์รวมเป็นเมตริก $Y = [y_1, \dots, y_d]$ ที่มีขนาดเป็น $m \times d$ เมตริก Y ที่ได้มีชื่อว่าพีเจอร์เมตริกของภาพ A

3.4.3 การสร้างภาพใหม่

ในวิธีของ 2DLDA สามารถนำเอาพีเจอร์เมตริกและเกณฑ์การฉายที่เหมาะสมที่สุดมาสร้างเป็นภาพใบหน้าใหม่ได้โดย กำหนดให้ A เป็นภาพ โดยที่ $Y = [y_1, \dots, y_d]$ เป็นพีเจอร์เมตริกของภาพ A และ $X_{\text{2DLDA}} = [x_1, \dots, x_d]$ เป็นเกณฑ์การฉายที่เหมาะสมที่สุด จากสมการ $Y = AX$ และ X เป็นออร์โธโนมอล ($X^{-1} = X^T$) ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = YX^T = \sum_{k=1}^d y_k x_k^T \quad (3.73)$$

ในการสร้างภาพใบหน้าใหม่ขึ้นถ้าเลือกจำนวนแกนการฉาย d เท่ากับ n จะได้ภาพที่สร้างใหม่เหมือนกับภาพ A แต่ถ้าเลือก d น้อยกว่า n ภาพ \tilde{A} จะความเหมือนกับภาพ A น้อยลงตามลำดับ



รูปที่ 3.9 ตัวอย่างภาพใบหน้าที่สร้างใหม่ด้วยวิธี 2DLDA จากฐานข้อมูล ORL

3.4.4 หลักการคำนวณของวิธี 2DLDA

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ A_j ($j = 1, 2, \dots, M$) จัดแบ่งเป็นกลุ่ม T_i ($i = 1, 2, \dots, L$) โดยที่ L คือจำนวนกลุ่ม

คำนวณเมตริกการกระจายภายในกลุ่ม

เมตริกการกระจายภายในกลุ่มจะเป็นการวัดค่าการกระจายระหว่างภาพภายในแต่ละกลุ่ม เมื่อมีภาพทั้งหมด L กลุ่ม ดังนั้นการคำนวณจะใช้การหาผลรวมของเมตริกความแปรปรวน (Covariance matrix) ในแต่ละกลุ่มมาคำนวณหา

$$S_W = \sum_{i=1}^L \sum_{A_k \in T_i} (A_k - \bar{A}_i)^T (A_k - \bar{A}_i)$$

คำนวณเมตริกการกระจายระหว่างกลุ่ม

เมตริกการกระจายระหว่างกลุ่มจะเป็นการวัดค่าการกระจายระหว่างกลุ่ม สามารถคำนวณหาโดยการหาผลรวมเมตริกความแปรปรวนของความแตกต่างระหว่างค่าเฉลี่ยรวมกับค่าเฉลี่ยของแต่ละกลุ่ม

$$S_B = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

คำนวณค่าไอเกนเวกเตอร์และค่าไอเกน

คำนวณหาค่าไอเกนเวกเตอร์และค่าไอเกนของเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม

$$S_B X_{opt} = \lambda S_W X_{opt}$$

เลือกค่าไอเกนเวกเตอร์

เรียงไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปน้อย แล้วเลือกแกนในการฉาย 1 ถึง d โดย $d \leq n$

$$X = [x_1, x_2, \dots, x_d]$$

การหาพีเจียร์ของภาพ

$$Y_k = AX_k, k = 1, 2, \dots, d$$

การสร้างภาพใหม่

$$\tilde{A} = YX^T = \sum_{k=1}^d y_k x_k^T$$

3.5 Eigenimages

วิธีการรู้จำใบหน้าที่ผ่านมาจะเห็นได้ว่าการพัฒนาวิธีการหาพีเจียร์ของภาพซึ่งการคำนวณหาพีเจียร์เริ่มจากการใช้เวกเตอร์มาคำนวณโดยได้นำภาพมาแปลงเป็นเวกเตอร์ภาพ แล้วนำมาลดมิติข้อมูล จากนั้นนำไปหาความค่าความถูกต้องในการรู้จำ ได้แก่ วิธี PCA และ วิธี LDA การพัฒนาขั้นต่อมาได้นำภาพมาคำนวณโดยตรง คือ วิธี 2DPCA และวิธี 2DLDA การคำนวณหาพีเจียร์ 2 วิธีนี้ไม่ต้องเปลี่ยนภาพให้เป็นเวกเตอร์ภาพ ใช้ภาพมาคำนวณโดยตรงซึ่งเป็นการลดขั้นตอนการคำนวณหาพีเจียร์ จึงทำให้เวลาที่ใช้ในการคำนวณน้อยลงและได้ความถูกต้องในการรู้จำที่ดีกว่าเดิม แต่ วิธี 2DPCA และ วิธี 2DLDA ยังสามารถพัฒนาการคำนวณหาพีเจียร์ เพื่อให้ได้ความถูกต้องในการรู้จำที่ดีกว่าเดิมได้ เนื่องจาก วิธี 2DPCA และ วิธี 2DLDA สามารถคำนวณหาพีเจียร์ ของภาพได้เพียงแค่ด้านเดียว

การพัฒนาขั้นต่อมา Zhang [5] ได้นำภาพมาคำนวณโดยตรง การคำนวณนี้สามารถทำการคำนวณได้ทั้ง 2 ด้าน คือ แถว และ หลัก ของภาพ วิธีการนี้เรียกว่า Eigenimages โดยที่ความแตกต่างระหว่างวิธี 2PCA และ Eigenimages คือ สมการการแปลงเชิงเส้น ซึ่ง วิธี Eigenimages จะใช้ $Y = U^T AV$ โดยที่ U และ V เป็นเมตริกการฉายในแนวหลัก (column) และเมตริกการฉายในแนวแถว (row) ตามลำดับ A คือ เมตริกภาพ ซึ่งจะเห็นได้ว่ามีเมตริกการฉาย 2 ตัว ต่างจากวิธี 2DPCA จะใช้ $Y = AX$ ซึ่ง X เป็นแกนการฉายแกนเดียว จะเห็นได้ว่าวิธี Eigenimages มีความละเอียดในการคำนวณมากกว่าวิธี 2DPCA

3.5.1 แนวความคิดของวิธี Eigenimages

กำหนดให้ $U \in R^{m \times r}$ และ $V \in R^{n \times c}$ และ A คือภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนเมตริก U และ V ตามหลักการแปลงจะได้

$$Y = U^T AV \quad (3.74)$$

ดังนั้นจะได้ Y คือ พีเจอร์เมตริกของภาพ A ที่มีขนาด $r < m$ และ $c < n$

วิธีการในการหาเมตริกการฉาย U และ V เราไม่สามารถ U และ V พร้อมกันได้ ดังนั้นจึงทำการคำนวณทีละตัว โดยกำหนดให้ $V = I$ (I คือ เมตริกเอกลักษณ์) เพื่อหา U และต่อมากำหนด $U = I$ เพื่อหาค่า V

การคำนวณหา U

$$\text{กำหนดให้ } V = I$$

จาก

$$J(U) = tr(S_U) \quad (3.75)$$

เมื่อ S_U คือ เมตริกความแปรปรวนร่วมของพีเจอร์เมตริก Y ในแนวหลัก และ $tr(S_U)$ คือ การหาผลรวมค่าไอเกนของเมตริก S_U

ลักษณะที่สำคัญของการหาค่ามากที่สุด (Maximize) ในสมการที่ (3.75) ก็เพื่อหาทิศทางการฉาย U ที่มีความแปรปรวนมากที่สุด โดยที่เมตริกความแปรปรวนร่วม (S_U) สามารถหาได้จาก

$$\begin{aligned} S_U &= E(Y - E(Y))(Y - E(Y))^T \\ &= E[U^T AV - E(U^T AV)][U^T AV - E(U^T AV)]^T \\ &= E[U^T (A - E(A))V][U^T (A - E(A))V]^T \end{aligned} \quad (3.76)$$

ถ้า $V = I$

$$\text{tr}(S_U) = U^T [E(A - E(A))^T ((A - E(A)))U] \quad (3.77)$$

เมื่อกำหนดให้

$$G_U = [E(A - E(A))^T ((A - E(A)))] \quad (3.78)$$

เมตริก G_U คือ เมตริกความแปรปรวนร่วมของภาพแนวหลักที่มีขนาด $m \times n$

เมื่อกำหนดให้มีภาพตัวอย่างที่ใช้สอนทั้งหมด M ภาพ เป็นเมตริกที่มีขนาด $m \times n$ ดังนั้น $A_j (j = 1, 2, \dots, M)$ และค่าเฉลี่ยภาพของภาพทั้งหมดคือ \bar{A} การคำนวณหาเมตริกความแปรปรวนร่วมของภาพ สามารถนำภาพมาคำนวณได้โดยตรงจาก

$$G_U = \frac{1}{M} \sum_{j=1}^M [(A_j - \bar{A})^T (A_j - \bar{A})] \quad (3.79)$$

จากสมการที่ (3.75) สามารถแทนค่าสมการที่ (3.77) และสมการที่ (3.79) เป็น

$$J(U) = U^T G_U U \quad (3.80)$$

สมการที่ (3.80) มีชื่อว่า สมการทั่วไปของผลรวมการกระจาย เมตริก U ที่ได้จากสมการนี้ มีค่ามากที่สุดมีชื่อว่า เมตริกที่เหมาะสมที่สุดในการฉาย ในการหาค่ามากที่สุดนั้นก็เพื่อต้องการให้การฉายภาพลงบนเมตริก U แล้วเมตริกใหม่ที่ถูกฉายมีการกระจายตัวมากที่สุด

เมตริกการฉายที่เหมาะสมที่สุดที่ใช้ในการฉาย U_{opt} สามารถหาได้จากการหาค่าไอเกนเวกเตอร์ ของ G ที่สมนัยกับค่าที่มากที่สุดของค่าไอเกนค่าที่มากที่สุดตั้งแต่ 1 ถึง m แต่ในความเป็นจริงเนื่องจากต้องการหาพีเจอร์ของภาพนั้นจึงเลือกแกนการฉายเพียง 1 ถึง r เท่านั้น ดังนั้น เมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$

$$U = [u_1 u_2 \dots u_r] \quad (3.81)$$

การคำนวณ V

กำหนดให้ $U = I$

จาก

$$J(V) = \text{tr}(S_V) \quad (3.82)$$

เมื่อ S_V คือ เมตริกความแปรปรวนร่วมของพีเจอร์เมตริก Y ในแนวแถว และ $tr(S_V)$ คือ การหาผลรวมค่าไอเกนของเมตริก S_V

ลักษณะที่สำคัญของการหาค่ามากที่สุด (Maximize) ในสมการที่ (3.82) ก็เพื่อหาทิศทางการฉาย V ที่มีความแปรปรวนมากที่สุด โดยที่เมตริกความแปรปรวนร่วม (S_V) สามารถหาได้จาก

$$\begin{aligned} S_V &= E(Y - E(Y))(Y - E(Y))^T \\ &= E[U^T AV - E(U^T AV)][U^T AV - E(U^T AV)]^T \\ &= E[U^T (A - E(A))V][U^T (A - E(A))V]^T \end{aligned} \quad (3.83)$$

ถ้า $V = I$

$$tr(S_V) = V^T [E(A - E(A))(A - E(A))^T] V \quad (3.84)$$

เมื่อกำหนดให้

$$G_V = [E(A - E(A))(A - E(A))^T] \quad (3.85)$$

เมตริก G_V คือ เมตริกความแปรปรวนร่วมของภาพแนวแถวที่มีขนาด $m \times m$

กำหนดให้ $A_j = [(A_j^{(1)})^T (A_j^{(2)})^T \dots (A_j^{(m)})^T]^T$ และ $\bar{A} = [(\bar{A}^{(1)})^T (\bar{A}^{(2)})^T \dots (\bar{A}^{(m)})^T]^T$ ขณะที่ $A_j^{(i)}$ และ $\bar{A}^{(i)}$ แสดงเวกเตอร์แถวที่ i ของ A_j และ \bar{A} ตามลำดับ สามารถนำภาพมาคำนวณได้โดยตรงจาก

$$G_V = \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^m (A_j^{(i)} - \bar{A}^{(i)})(A_j^{(i)} - \bar{A}^{(i)})^T \quad (3.86)$$

จากสมการที่ (3.82) สามารถแทนค่าสมการที่ (3.84) และสมการที่ (3.86) เป็น

$$J(V) = V^T G_V V \quad (3.87)$$

สมการที่ (3.87) มีชื่อว่า สมการทั่วไปของผลรวมการกระจาย เมตริก V ที่ได้จากสมการนี้ มีค่ามากที่สุดมีชื่อว่า เมตริกที่เหมาะสมที่สุดในการฉาย ในการหาค่ามากที่สุดนั้นก็เพื่อต้องการให้การฉายภาพลงบนเมตริก V แล้วเมตริกใหม่ที่ถูกฉายมีการกระจายตัวมากที่สุด

แกนการฉายที่เหมาะสมที่สุดที่ใช้ในการฉาย V_{opt} สามารถหาได้จากการหาค่าไอเกนเวกเตอร์ ของ G ที่สมนัยกับค่าไอเกนค่าที่มากที่สุดตั้งแต่ 1 ถึง n แต่ในความเป็นจริงเนื่องจาก

ต้องการหาพีเจียร์ของภาพนั้นจึงเลือกแกนการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$V = [v_1 v_2 \dots v_c] \quad (3.88)$$

3.5.2 การหาพีเจียร์ของภาพ

หลังจากที่ได้เมตริกที่เหมาะสมที่ใช้ในการฉายภาพของวิธี Eigenimages คือ $[u_1 u_2 \dots u_r]$ และ $[v_1 v_2 \dots v_c]$ ถูกนำมาใช้หาพีเจียร์ของภาพ เมื่อกำหนดให้ A เป็นภาพตัวอย่างดังนั้นการหาพีเจียร์ของภาพ A คือ

$$Y_j = U^T A_j V, j = 1, 2, \dots, M \quad (3.89)$$

เมตริกใหม่ที่ได้จากการฉาย Y_1, Y_2, \dots, Y_M มีชื่อว่าเมตริกองค์ประกอบหลักของภาพตัวอย่าง A ที่มีขนาด $r \times c$ โดยที่ $r \leq m$ และ $c \leq n$ เมตริก Y มีชื่อว่าพีเจียร์เมตริกภาพ A

3.5.3 การสร้างภาพใหม่

วิธีของ Eigenimages สามารถนำเอาพีเจียร์เมตริกและเมตริกการฉายที่เหมาะสมที่สุด U และ V มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ Y เป็นพีเจียร์เมตริกของภาพ A โดยที่ $U = [u_1 u_2 \dots u_r]$ และ $V = [v_1 v_2 \dots v_c]$ เป็นเมตริกการฉายที่เหมาะสมที่สุด ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = UYV^T \quad (3.90)$$

3.5.4 หลักการคำนวณของวิธี Eigenimages

เมื่อกำหนดให้มีภาพที่ใช้สอน M ภาพคือ A_1, A_2, \dots, A_M แต่ละภาพมีขนาด $m \times n$ จุดภาพคำนวณค่าเฉลี่ยของภาพใบหน้า (Mean Face)

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i$$

คำนวณหาเมตริกความแปรปรวนร่วมแนวหลัก

$$G_U = \frac{1}{M} \sum_{j=1}^M [(A_j - \bar{A})^T (A_j - \bar{A})]$$

คำนวณหาเมตริกความแปรปรวนร่วมแนวแถว

$$G_V = \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^m (A_j^{(i)} - \bar{A}^{(i)})(A_j^{(i)} - \bar{A}^{(i)})^T$$

คำนวณหาค่าไอเกนเวกเตอร์ ที่สมนัยกับค่าไอเกนจากเมตริกความแปรปรวนร่วมด้วยสมการ

$$G_U U_{opt} = \lambda U_{opt}$$

$$G_V V_{opt} = \lambda V_{opt}$$

เลือกค่า G_U ที่สมนัยกับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m แต่เนื่องจากต้องการหาพีเจอร์์ของภาพ ฉะนั้นจึงเลือกแกนในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$ ในทำนองเดียวกัน เลือกค่า G_V ที่สมนัยกับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง n แต่เนื่องจากต้องการหาพีเจอร์์ของภาพ ฉะนั้นจึงเลือกแกนในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$U = [u_1 u_2 \dots u_r]$$

$$V = [v_1 v_2 \dots v_c]$$

การหาพีเจอร์์ของภาพ

$$Y_j = U^T A_j V, \quad j = 1, 2, \dots, M$$

การสร้างภาพใหม่

$$\tilde{A} = U Y V^T$$

3.6 Coupled Subspace Linear Discriminant Analysis (CSLDA)

วิธีการ CSLDA เป็นวิธีการหาพีเจอร်ของภาพที่พัฒนามาจากวิธีการหาพีเจอร်บนพื้นฐานของ LDA ต่อจากวิธี 2DLDA ซึ่งวิธีการนี้เป็นการพัฒนาพีเจอร်ให้สามารถลดมิติข้อมูลได้ทั้ง แกว และ หลัก ของภาพ เช่นเดียวกับวิธี Eigenimages

สิทธิพนธ์ [6] ได้พัฒนาวิธีการ CSLDA นี้ขึ้นมาให้สามารถลดมิติข้อมูลได้ทั้ง แกว และ หลักของภาพ บนพื้นฐานการหาพีเจอร်ของ LDA ซึ่งได้นำข้อมูลของกลุ่มมาพิจารณาแตกต่างจากการหาพีเจอร်บนพื้นฐานของ PCA ที่ไม่ได้นำข้อมูลกลุ่มมาใช้ ในการคำนวณหาพีเจอร်

3.6.1 แนวความคิดของวิธีการ CSLDA

กำหนดให้เมตริกการฉาย $U \in R^{m \times r}$ และ $V \in R^{m \times c}$ A คือภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนเมตริก U และ V ตามหลักการแปลงเชิงเส้น

$$Y = U^T A V \quad (3.91)$$

ดังนั้นจะได้ Y คือ พีเจอร်เมตริกของภาพ A ที่มีขนาด $r \times c$ โดยที่ $r < m$ และ $c < n$

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวนภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม T_i และ P_i คือกลุ่มที่ถูกฉาย หลังจากภาพที่ใช้สอนถูกฉายลงบนเมตริก U และ V จะได้พีเจอร်เมตริกเป็น

$$Y_j = U^T A_j V, \quad j = 1, 2, \dots, M \quad (3.92)$$

วิธีในการหาเมตริกการฉาย U และ V สามารถคำนวณได้โดยการหาผลรวมค่าไอเกน (trace) ของเมตริกความแปรปรวนร่วมจากภาพตัวอย่างที่ถูกฉาย [1] จากแนวความคิดสามารถสร้างเป็นสมการได้คือ

$$J(U, V) = \frac{P_B}{P_W} \quad (3.93)$$

เมื่อ

$$P_B = \text{tr}(\tilde{S}_B)$$

$$P_W = \text{tr}(\tilde{S}_W)$$

โดยที่

\tilde{S}_B คือ เมตริกการกระจายระหว่างกลุ่มของภาพตัวอย่างที่ถูกฉาย

\tilde{S}_W คือ เมตริกการกระจายภายในกลุ่มของภาพตัวอย่างที่ถูกฉาย

$$\begin{aligned}\tilde{S}_B &= \sum_{i=1}^L N_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T \\ &= \sum_{i=1}^L N_i [U(\bar{A}_i - \bar{A})V][U(\bar{A}_i - \bar{A})V]^T\end{aligned}\quad (3.94)$$

$$\begin{aligned}\tilde{S}_W &= \sum_{i=1}^L \sum_{Y_k \in P_i} (Y_k - \bar{Y}_i)(Y_k - \bar{Y}_i)^T \\ &= \sum_{i=1}^L \sum_{A_k \in P_i} [U(A_k - \bar{A}_i)V][U(A_k - \bar{A}_i)V]^T\end{aligned}\quad (3.95)$$

ถ้า

$$tr(\tilde{S}_B) = tr\left(\sum_{i=1}^L N_i U^T (\bar{A}_i - \bar{A})VV^T (\bar{A}_i - \bar{A})^T U\right) \quad (3.96)$$

$$tr(\tilde{S}_W) = tr\left(\sum_{i=1}^L \sum_{A_k \in P_i} U^T (A_k - \bar{A}_i)V V^T (A_k - \bar{A}_i)^T U\right) \quad (3.97)$$

ในการคำนวณหา U และ V ไม่สามารถทำได้พร้อมกัน ดังนั้นจึงทำการคำนวณทีละค่า โดยกำหนดให้ $V = I$ (I คือ เมตริกเอกลักษณ์) เพื่อหาค่าของ U ลำดับต่อมากำหนด $U = I$ คำนวณหาค่าของ V

การคำนวณหา U

กำหนดให้ $V = I$

$$tr(\tilde{S}_B) = U^T S_B^U U, \quad tr(\tilde{S}_W) = U^T S_W^U U \quad (3.98)$$

โดยที่

$$\begin{aligned}S_B^U &= \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T \\ S_W^U &= \sum_{i=1}^L \sum_{A_k \in P_i} N_i (A_k - \bar{A}_i)(A_k - \bar{A}_i)^T\end{aligned}$$

จากสมการที่ (3.93) สามารถแทนค่าใหม่ได้คือ

$$J(U) = \frac{U^T S_B^U U}{U^T S_W^U U} \quad (3.99)$$

การแก้สมการหา U ที่เหมาะสมที่สุดในการฉายสามารถทำได้โดยการหาค่ามากที่สุดของ $J(U)$

$$U_{opt} = \arg \max_U J(U) \quad (3.100)$$

ถ้า S_W^U ไม่เป็นเมตริกเอกฐาน (nonsingular) สามารถหาเมตริก U_{opt} โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B^U U_{opt} = \lambda S_W^U U_{opt} \quad (3.101)$$

โดยปกติเมตริกการฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_B^{U^{-1}} S_B^U$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m แต่ในความเป็นจริงเนื่องจากการหาพีเจอร์ของภาพะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$

$$U = [u_1 u_2 \dots u_r] \quad (3.102)$$

การคำนวณหา V

กำหนดให้ $U = I$

$$tr(\tilde{S}_B) = V^T S_B^V V, \quad tr(\tilde{S}_W) = V^T S_W^V V \quad (3.103)$$

โดยที่

$$S_B^V = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

$$S_W^V = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i)^T (\bar{A}_k - \bar{A}_i)$$

ตามคุณสมบัติของ $tr(AB) = tr(BA)$ เมื่อ A และ B เป็นเมตริกจากสมการที่ (3.93) สามารถแทนค่าใหม่ได้คือ

$$J(V) = \frac{V^T S_B^V V}{V^T S_W^V V} \quad (3.104)$$

การแก้สมการหา V ที่เหมาะสมที่สุดสำหรับการฉายสามารถทำได้โดยการหาค่ามากที่สุดของ $J(V)$

$$V_{opt} = \arg \max_V J(V) \quad (3.105)$$

ถ้า S_B^T ไม่เป็นเมตริกเอกฐาน (nonsingular) สามารถหาเมตริก V_{opt} โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B^T V_{opt} = \lambda S_W^T V_{opt} \quad (3.106)$$

โดยปกติเมตริกการฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_B^{T-1} S_B^T$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง n แต่ในความเป็นจริงเนื่องจากต้องการหาพีเจอร်ของภาพฉะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$V = [v_1 v_2 \dots v_c] \quad (3.107)$$

3.6.2 การหาพีเจอร်ของภาพ

หลังจากภาพที่ใช้สอนถูกฉายลงบนเมตริก U และ V โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม จะได้พีเจอร်เมตริกเป็น

$$Y_j = U^T A_j V, j = 1, 2, \dots, M \quad (3.108)$$

เมตริกใหม่ที่ได้จากการฉาย Y_1, Y_2, \dots, Y_M มีชื่อว่าเมตริกองค์ประกอบหลักของภาพตัวอย่าง A ที่มีขนาด $r \times c$ โดยที่ $r \leq m$ และ $c \leq n$ เมตริก Y มีชื่อว่าพีเจอร်เมตริกภาพ A

3.6.3 การสร้างภาพใหม่

วิธี CSLDA สามารถนำเอาพีเจอร์เมตริกและเมตริกการฉายที่เหมาะสมที่สุด U และ V มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ Y เป็นพีเจอร์เมตริกของภาพ A โดยที่ $U = [u_1 u_2 \dots u_r]$ และ $V = [v_1 v_2 \dots v_c]$ เป็นเมตริกการฉายที่เหมาะสมที่สุด ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = UYV^T \quad (3.109)$$

3.6.4 หลักการคำนวณของวิธี CSLDA

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวนภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม

คำนวณเมตริกการกระจายภายในกลุ่ม

เมตริกการกระจายภายในกลุ่มจะเป็นการวัดค่าการกระจายระหว่างภาพภายในกลุ่มเมื่อมีภาพทั้งหมด L กลุ่ม ดังนั้นการคำนวณจะใช้การหาผลรวมของเมตริกความแปรปรวนในแต่ละกลุ่มมาคำนวณ

$$S_W^U = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i) (\bar{A}_k - \bar{A}_i)^T$$

$$S_W^V = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i)^T (\bar{A}_k - \bar{A}_i)$$

คำนวณเมตริกการกระจายระหว่างกลุ่ม

เมตริกการกระจายระหว่างกลุ่มจะเป็นการวัดค่าการกระจายระหว่างกลุ่ม สามารถคำนวณโดยการหาผลรวมเมตริกความแปรปรวนของค่าความแตกต่างระหว่างค่าเฉลี่ยรวมกับค่าเฉลี่ยแต่ละกลุ่ม

$$S_B^U = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A}) (\bar{A}_i - \bar{A})^T$$

$$S_B^V = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

คำนวณไอเจนเวกเตอร์และค่าไอเจน

คำนวณไอเจนเวกเตอร์และค่าไอเจนของเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม

$$S_B^U U_{opt} = \lambda S_W^U U_{opt}$$

$$S_B^V V_{opt} = \lambda S_W^V V_{opt}$$

เลือกค่าไอเจนเวกเตอร์

เลือกค่า S_B^U ที่สมนัยกับค่าไอเจนที่มากที่สุดตั้งแต่ 1 ถึง m แต่เนื่องจากต้องการหาพีเจอร์ของภาพ ฉะนั้นจึงเลือกแกนในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$ ในทำนองเดียวกัน เลือกค่า S_B^V ที่สมนัยกับค่าไอเจนที่มากที่สุดตั้งแต่ 1 ถึง n แต่เนื่องจากต้องการหาพีเจอร์ของภาพ ฉะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$U = [u_1 u_2 \dots u_r]$$

$$V = [v_1 v_2 \dots v_c]$$

การหาพีเจอร์ของภาพ

$$Y_j = U^T A_j V, \quad j = 1, 2, \dots, M$$

การสร้างภาพใหม่

$$\tilde{A} = U Y V^T$$

บทที่ 4

วิธีการที่นำเสนอ (WPS-CSLDA)

จากบทที่ 3 ทำให้เราทราบถึงแนวทางการพัฒนาวิธีการรู้จำภาพใบหน้าโดยใช้วิธีทางสถิติ และวิธีการคำนวณ ซึ่งพบว่าวิธีการรู้จำภาพใบหน้าทางสถิตินั้นมีทฤษฎีที่สำคัญ 2 ทฤษฎี ได้แก่ ทฤษฎีบนพื้นฐานของ PCA และทฤษฎีบนพื้นฐานของ LDA แต่สิ่งที่เราสนใจนั้นคือ ทฤษฎีบนพื้นฐานของ LDA เนื่องจากมีการนำข้อมูลของกลุ่มมาพิจารณา จากการศึกษาพบว่า ทฤษฎีบนพื้นฐานของ LDA ที่นำข้อมูลของกลุ่มข้อมูลมาพิจารณานั้น ให้ความถูกต้องในการรู้จำภาพใบหน้ามากกว่าทฤษฎีที่อยู่บนพื้นฐานของ PCA แต่ในทฤษฎีบนพื้นฐานของ LDA ยังพบว่าเมื่อได้นำข้อมูลของกลุ่มมาพิจารณาแล้วนั้นยังมีข้อมูลระหว่างกลุ่มบางส่วนมีการปนกัน ซึ่งทำให้ประสิทธิภาพในการคำนวณความถูกต้องในการรู้จำลดลง

ปัญหาดังกล่าวเราได้้นำ weighted pairwise scatter ที่นำเสนอโดย Li [7] ที่ใช้แก้ปัญหาการปนกันของข้อมูลเสียงมาประยุกต์ใช้กับแก้ปัญหาการปนกันของข้อมูลภาพ โดยได้นำ weighted pairwise scatter มาคำนวณหาเมตริกการกระจายระหว่างกลุ่มของการคำนวณพีเจอร์ CSLDA ทฤษฎีและวิธีการคำนวณจะนำเสนอในหัวข้อต่อไป

4.1 Weighted Pairwise Scatter

ตามทฤษฎีของ LDA [2] โดยทั่วไปเทอมแรก $\tilde{S}_W = w^T S_W w$ จะเรียกว่าเมตริกการกระจายภายในกลุ่ม เทอมที่สอง $\tilde{S}_B = w^T S_B w$ จะเรียกว่าเมตริกการกระจายระหว่างกลุ่ม ถ้า w เป็นการฉายเส้นตรง (linear projection) ซึ่งอยู่ในปริภูมิพีเจอร์ (feature space) ใหม่ การกระจายภายในกลุ่มและระหว่างกลุ่มจะเป็น $w^T S_W w$ และ $w^T S_B w$

เป็นที่รู้ทั่วกันว่า การกระจายระหว่างกลุ่มจะรองรับข้อมูลการแบ่งแยก (discriminant information) แนวคิดของ LDA คือ จะทำให้อัตราส่วนเมตริกการกระจายระหว่างกลุ่มและเมตริกการกระจายภายในกลุ่มหลังจากการแปลง (transform) มีค่ามากที่สุด ซึ่งจะทำได้โดยการแปลงที่ยังคงข้อมูลการแบ่งแยกมากที่สุด ในขณะที่ขนาดของมิติจะลดลง เราต้องการที่จะทำให้อัตราส่วนดังกล่าวมีค่ามากที่สุด

$$\max_w \frac{|w^T S_B w|}{|w^T S_W w|}$$

คำตอบในการแก้ปัญหานี้จะใช้ ไอเกนเวกเตอร์ (Eigenvector) ที่สมนัยกับค่าไอเกน (Eigenvalue) ที่เรียงจากมากไปน้อยจากสมการ

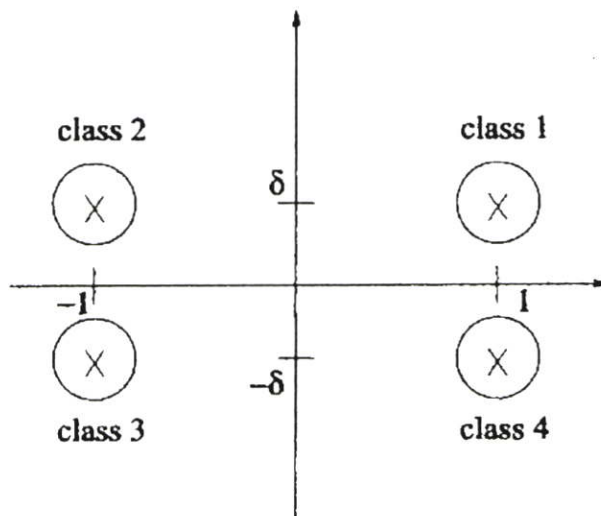
$$S_B w_i = \lambda S_W w_i$$

รูปแบบของเมตริกการกระจายระหว่างกลุ่มสามารถหาได้จาก

$$S_B = \sum_{i=1}^L N_i (\bar{t}_i - \bar{t})(\bar{t}_i - \bar{t})^T$$

โดยที่ \bar{t} คือ ค่าเฉลี่ยของตัวอย่างทั้งหมด

สมการนี้เป็นการวัดว่าจะแบ่งค่าเฉลี่ยของแต่ละกลุ่มอย่างไรจากจุดศูนย์กลางข้อมูล มันเป็นการดีที่จะได้ค่า S_B มากๆ เพราะจะแสดงว่ากลุ่มต่างๆ จะมีการแยกออกจากกันมากในปริภูมิการแปลง (transformed space) ดังนั้นเป็นการง่ายที่จะแบ่งแยก จากการแสดงนี้จะไม่ชัดเจนว่ากลุ่มต่างๆ จะถูกแยกออกอย่างไร ในตัวอย่างต่อไปนี้จะทำให้เข้าใจจุดที่เราพิจารณา



รูปที่ 4.1 ข้อมูลที่มีความแปรปรวนเท่ากันที่ตำแหน่งต่างกัน

ตัวอย่าง กำหนดให้มีข้อมูล 4 กลุ่มแต่ละกลุ่มมีจำนวนของพีเจอร์เวคเตอร์เหมือนกันและมีความแปรปรวนเท่ากัน ดังนี้

$$\begin{aligned}\bar{i}_1 &= (1, \delta) & \bar{i}_2 &= (-1, \delta) \\ \bar{i}_3 &= (-1, -\delta) & \bar{i}_4 &= (1, -\delta)\end{aligned}$$

ปัญหานี้จะแสดงในรูปที่ 4.1 ในกรณีนี้เมตริกการกระจายระหว่างกลุ่มเป็นดังนี้
กำหนดให้ $\delta = 0$

$$\frac{1}{4}S_B = \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

เมตริกการกระจายระหว่างกลุ่มจะไม่เกิดการแบ่งแยกในแนวตั้ง จะแยกได้แต่แนวนอนเพียงอย่างเดียว เราสามารถบอกได้ว่าความแปรปรวนของกลุ่มข้อมูลคู่อื่นๆ จะมีอิทธิพลต่อเมตริกการกระจายตัวระหว่างกลุ่มมากกว่า (1,4) และ (2,3) การพิจารณาถึงปัญหาในการแยกความแปรปรวนของกลุ่มข้อมูลคู่ (1,4) และ (2,3) มีความสำคัญมากกว่าอื่นๆ เพราะกลุ่มข้อมูลคู่นี้จะมีความยุ่งยากมากกว่า การกระจายระหว่างกลุ่มยังคงไม่ยึดจับ (capture) ข้อมูลนี้ ดังนั้นสรุปได้ว่า วิธีที่เป็นที่ยอมรับจะเป็นตัวอย่าง ข้อมูลการแยกแยะตามต้องการได้ไม่แม่นยำ

จากตัวอย่างข้างต้นจะเห็นว่าถ้ามีบางกลุ่มที่อยู่ใกล้ๆ กันมาเปรียบเทียบกับกลุ่มอื่นๆ เมตริกการกระจายระหว่างกลุ่มส่วนมากจะไม่สนใจข้อมูลการแยกแยะระหว่างกลุ่มที่อยู่ใกล้ๆ กัน

การพิจารณาในส่วนที่ผ่านมานำเราไปสู่การกำหนดเมตริกการกระจายระหว่างกลุ่มที่เท่ากับผลรวมของเมตริก Weight pairwise scatter [7] เนื่องจากการคงข้อมูลการแยกแยะที่เพียงพอ เราจะต้องปรับค่าถ่วงน้ำหนัก (weight) การปรับที่ได้รับการยอมรับคือ จะปรับค่าถ่วงน้ำหนักให้เท่ากับอินเวอร์ส (inverse) ของค่าเฉลี่ยระหว่างกลุ่มของระยะทางยูคลีเดียน (Euclidean distance)

$$w = \frac{1}{\|\bar{i}_i - \bar{i}\|} = \frac{1}{(\bar{i}_i - \bar{i})^T (\bar{i}_i - \bar{i})} \quad (4.1)$$

เราจะให้ความสำคัญกับกลุ่มที่มีค่าเฉลี่ยใกล้กับกลุ่มอื่นๆ มากกว่ากลุ่มที่มีค่าเฉลี่ยห่างกัน ในกรณีนี้กลุ่มที่มีการปนกันของข้อมูลมากกว่าจะได้รับการสนใจมากกว่า และกลุ่มที่มีความการปนกันของข้อมูลน้อยกว่าจะได้รับความสนใจน้อยกว่า เนื่องมาจากรูปแบบปกติของถ่วงน้ำหนัก (weight) เมตริกการกระจายระหว่างชั้นจะเป็น

$$S_{B_{norm}} = \frac{1}{N_i} \sum_{i=1}^L N_i \frac{(\bar{t}_i - \bar{t})(\bar{t}_i - \bar{t})^T}{(\bar{t}_i - \bar{t})^T (\bar{t}_i - \bar{t})} \quad (4.2)$$

จากปัญหาที่กำหนดในตัวอย่าง ความผกผันระหว่างกลุ่มที่ถูกคำนวณโดยการใส่ (B) จะเป็นดังนี้
กำหนดให้ $\delta = 0$

$$\frac{1}{4} S_{B_{norm}} = \begin{pmatrix} 1 + \frac{1}{1 + \delta^2} & 0 \\ 0 & 1 + \frac{\delta^2}{1 + \delta^2} \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

ดังนั้นการใส่เมตริกความแปรปรวนระหว่างกลุ่ม (B) จะไม่เกิดอะไรขึ้นกับการเข้าใกล้กันของคู่ (1,4) และ (2,3) เราจะทำให้การกระจายของมันอยู่ในการคำนวณ WPS-CSLDA นี่คือนี่ที่ต้องการ เนื่องจากปัญหาการแยกข้อมูล

เมื่อเราได้เมตริกการกระจายระหว่างกลุ่มใหม่ เราจะแก้สมการ การกระจายระหว่างกลุ่ม ด้วยเมตริกใหม่นี้ นอกเหนือจากนั้นยังคงเดิม

4.2 แนวคิดของ Weighted Pairwise Scatter Coupled Subspace Linear

Discriminant Analysis (WPS-CSLDA)

กำหนดให้แกนการฉาย $U \in \mathbb{R}^{m \times r}$ และ $V \in \mathbb{R}^{n \times c}$ A คือ ภาพที่มีขนาด $m \times n$ จากการฉาย ภาพ A ลงบนเมตริกการฉาย U และ V ตามหลักการแปลงเชิงเส้น

$$Y = U^T A V \quad (4.3)$$

ดังนั้นจะได้ Y คือ พีเจอร์เมตริกของภาพ A ที่มีขนาด $r \times c$ โดยที่ $r \leq m$ และ $c \leq n$

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวน ภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม T_i และ P_i คือกลุ่มที่ถูกฉาย หลังจากภาพที่ใช้สอนถูกฉายลงบนเมตริกการฉาย U และ V จะได้พีเจอร์เมตริกเป็น

$$Y_j = U^T A_j V, j = 1, 2, \dots, M \quad (4.4)$$

วิธีในการหาเมตริกการฉาย U และ V สามารถคำนวณได้โดยการหาผลรวมค่าไอเกน (trace) ของเมตริกความแปรปรวนร่วมจากภาพตัวอย่างที่ถูกฉาย [1] จากแนวความคิดสามารถสร้างเป็นสมการได้คือ

$$J(U, V) = \frac{P_B}{P_W} \quad (4.5)$$

เมื่อ

$$P_B = \text{tr}(\tilde{S}_B)$$

$$P_W = \text{tr}(\tilde{S}_W)$$

โดยที่

\tilde{S}_B คือ เมตริกการกระจายระหว่างกลุ่มของภาพตัวอย่างที่ถูกฉาย

\tilde{S}_W คือ เมตริกการกระจายภายในกลุ่มของภาพตัวอย่างที่ถูกฉาย

$$\begin{aligned} \tilde{S}_B &= \sum_{i=1}^L \varpi N_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T \\ &= \sum_{i=1}^L \varpi N_i [U(\bar{A}_i - \bar{A})V][U(\bar{A}_i - \bar{A})V]^T \end{aligned} \quad (4.6)$$

$$\begin{aligned} \tilde{S}_W &= \sum_{i=1}^L \sum_{Y_k \in P_i} (Y_k - \bar{Y}_i)(Y_k - \bar{Y}_i)^T \\ &= \sum_{i=1}^L \sum_{A_k \in P_i} [U(A_k - \bar{A}_i)V][U(A_k - \bar{A}_i)V]^T \end{aligned} \quad (4.7)$$

ถ้า

$$n(\tilde{S}_B) = n\left(\sum_{i=1}^L \varpi N_i U^T (\bar{A}_i - \bar{A})VV^T (\bar{A}_i - \bar{A})^T U\right) \quad (4.8)$$

$$n(\tilde{S}_W) = n\left(\sum_{i=1}^L \sum_{A_k \in P_i} U^T (A_k - \bar{A}_i)V V^T (A_k - \bar{A}_i)^T U\right) \quad (4.9)$$

เมื่อ ϖ คือ weighted pairwise scatter

การคำนวณหา U และ V นั้นไม่สามารถทำได้พร้อมกัน ดังนั้นจึงทำการคำนวณทีละค่า โดยกำหนดให้ $V = I$ (I คือ เมตริกเอกลักษณ์) เพื่อหาค่าของ U ลำดับต่อมากำหนด $U = I$ คำนวณหาค่าของ V

การคำนวณหา U

กำหนดให้ $V = I$

$$m(\tilde{S}_B) = U^T S_B^U U, \quad m(\tilde{S}_W) = U^T S_W^U U \quad (4.10)$$

โดยที่

$$S_B^U = \sum_{i=1}^L w_i N_i (\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T$$

$$S_W^U = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i)(\bar{A}_k - \bar{A}_i)^T$$

จากสมการที่ (4.3) สามารถแทนค่าใหม่ได้คือ

$$J(U) = \frac{U^T S_B^U U}{U^T S_W^U U} \quad (4.11)$$

การแก้สมการหา U ที่เหมาะสมที่สุดในการฉายสามารถทำได้โดยการหาค่ามากที่สุดของ $J(U)$

$$U_{opt} = \arg \max_U J(U) \quad (4.12)$$

ถ้า S_W^U ไม่เป็นเมตริกเอกฐาน (non singular) สามารถหาเมตริก U_{opt} โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B^U U_{opt} = \lambda S_W^U U_{opt} \quad (4.13)$$

โดยปกติเมตริกการฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_B^{U^{-1}} S_W^U$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m แต่ในความเป็นจริงเนื่องจากต้องการหาพีเจอร์ของภาพฉะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$

$$U = [u_1 u_2 \dots u_r] \quad (4.14)$$

การคำนวณหา V

กำหนดให้ $C = I$

$$tr(\tilde{S}_B) = V^T S_B^V V, \quad tr(\tilde{S}_W) = V^T S_W^V V \quad (4.15)$$

โดยที่

$$S_B^V = \sum_{i=1}^L \omega N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

$$S_W^V = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i)^T (\bar{A}_k - \bar{A}_i)$$

ตามคุณสมบัติของ $tr(AB) = tr(BA)$ เมื่อ A และ B เป็นเมตริกจากสมการที่ (4.3) สามารถแทนค่าใหม่ได้ คือ

$$J(V) = \frac{V^T S_B^V V}{V^T S_W^V V} \quad (4.16)$$

การแก้สมการหา V ที่เหมาะสมที่สุดในการฉายสามารถทำได้โดยการหาค่ามากที่สุดของ $J(V)$

$$V_{opt} = \arg \max_V J(V) \quad (4.17)$$

ถ้า S_W^V ไม่เป็นเมตริกเอกฐาน (nonsingular) สามารถหาเมตริก V_{opt} โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B^V V_{opt} = \lambda S_W^V V_{opt} \quad (4.18)$$

ในการทำงานเดียวกันเมตริกการฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_W^{V^{-1}} S_B^V$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง n แต่ในความเป็นจริงเนื่องจาก

ต้องการหาพีเจอรซ์ของภาพฉะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$V = [v_1 v_2 \dots v_c] \quad (4.19)$$

4.3 การหาพีเจอรซ์ของภาพ

หลังจากภาพที่ใช้สอนถูกฉายลงบนเมตริก U และ V โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม จะได้พีเจอรซ์เมตริกเป็น

$$Y_j = U^T A_j V, j = 1, 2, \dots, M \quad (4.20)$$

เมตริกใหม่ที่ได้จากการฉาย Y_1, Y_2, \dots, Y_M มีชื่อว่าเมตริกองค์ประกอบหลักของภาพตัวอย่าง A ที่มีขนาด $r \times c$ โดยที่ $r \leq m$ และ $c \leq n$ เมตริก Y มีชื่อว่าพีเจอรซ์เมตริกภาพ A

4.4 การสร้างภาพใหม่

วิธี CSLDA สามารถนำเอาพีเจอรซ์เมตริกและเมตริกการฉายที่เหมาะสมที่สุด U และ V มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ Y เป็นพีเจอรซ์เมตริกของภาพ A โดยที่ $U = [u_1 u_2 \dots u_r]$ และ $V = [v_1 v_2 \dots v_c]$ เป็นเมตริกการฉายที่เหมาะสมที่สุด ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = UYV^T \quad (4.21)$$

4.5 วิธีการคำนวณ WPS-CSLDA

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวนภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม

คำนวณเมตริกการกระจายภายในกลุ่ม

เมตริกการกระจายภายในกลุ่มจะเป็นการวัดค่าการกระจายระหว่างภาพภายในกลุ่มเมื่อมีภาพทั้งหมด L กลุ่ม ดังนั้นการคำนวณจะใช้การหาผลรวมของเมตริกความแปรปรวนในแต่ละกลุ่มมาคำนวณ

$$S_W^U = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i) (\bar{A}_k - \bar{A}_i)^T$$

$$S_W^V = \sum_{i=1}^L \sum_{A_k \in P_i} N_i (\bar{A}_k - \bar{A}_i)^T (\bar{A}_k - \bar{A}_i)$$

คำนวณ weighted pairwise scatter

weighted pairwise scatter เป็นค่าที่จะนำไปใช้ในเมตริกการกระจายระหว่างกลุ่มเพื่อทำการแบ่งแยกข้อมูลที่อยู่ใกล้กันให้แยกจากกัน

$$\varpi = \frac{1}{\|\bar{A}_i - \bar{A}\|} = \frac{1}{(\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})}$$

คำนวณเมตริกการกระจายระหว่างกลุ่ม

เมตริกการกระจายระหว่างกลุ่มจะเป็นการวัดค่ากระจายระหว่างกลุ่ม สามารถคำนวณโดยการหาผลรวมเมตริกความแปรปรวนของค่าความแตกต่างระหว่างค่าเฉลี่ยรวมกับค่าเฉลี่ยแต่ละกลุ่ม

$$S_B^U = \sum_{i=1}^L \varpi N_i (\bar{A}_i - \bar{A}) (\bar{A}_i - \bar{A})^T$$

$$S_B^V = \sum_{i=1}^L \varpi N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

คำนวณไอเกนเวกเตอร์และค่าไอเกน

คำนวณไอเกนเวกเตอร์และค่าไอเกนของเมตริกการกระจายภายในกลุ่มและเมตริกการกระจายระหว่างกลุ่ม

$$S_B^U U_{opt} = \lambda S_W^U U_{opt}$$

$$S_B^V V_{opt} = \lambda S_W^V V_{opt}$$

เลือกค่าไอเกนเวกเตอร์

เลือกค่า S_B^U ที่สมนัยกับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m แต่เนื่องจากต้องการหาพีเจอร์ของภาพ ฉะนั้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$ ในทำนองเดียวกัน เลือกค่า S_B^V ที่สมนัยกับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง n แต่

เนื่องจากต้องการหาพีเจอร์ของภาพจะเน้นจึงเลือกเมตริกในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมตริกใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $n \times c$

$$U = [u_1 u_2 \dots u_r]$$

$$V = [v_1 v_2 \dots v_c]$$

การหาพีเจอร์ของภาพ

$$Y_j = U^T A_j V, j = 1, 2, \dots, M$$

4.6 วิธีการจำแนกข้อมูลภาพใบหน้า (Images Classification)

เมื่อคำนวณหาพีเจอร์เมตริกของภาพด้วยวิธี WPS-CSLDA แล้ว จะใช้วิธีการจำแนกข้อมูลด้วยวิธี 1-Nearest-Neighbor (INN) ซึ่งเป็นการประยุกต์การคำนวณหาระยะทางยูคลีเดียนเพื่อใช้กับเมตริกโดย กำหนดให้ภาพใบหน้า A_1 และ A_2 ถูกนำมาหาพีเจอร์เมตริกเป็น $Y_1 = [y_1^1, \dots, y_1^d]$ และ $Y_2 = [y_2^1, \dots, y_2^d]$ ถ้าต้องการหาความเหมือนกันของภาพที่ถูกฉายทั้งสอง ทำได้โดยการหาระยะทาง

$$d(Y_1, Y_2) = \sum_{k=1}^d \|y_1^k - y_2^k\|_2 \quad (4.22)$$

เมื่อ $\|y_1^k - y_2^k\|_2$ คือ ระยะทางยูคลีเดียนระหว่างสองพีเจอร์เมตริก

ถ้าพีเจอร์เมตริกของภาพที่ใช้ฝึกสอนทั้งหมด คือ Y_1, Y_2, \dots, Y_M (M คือ จำนวนภาพที่ใช้ฝึกสอน) เมื่อแต่ละภาพเป็นส่วนหนึ่งของกลุ่ม T_i ในการทดสอบ จะนำภาพตัวอย่างหาระยะทางยูคลีเดียนเทียบกับภาพทุกภาพที่ใช้ในการฝึกสอน จากนั้นจะพิจารณาภาพตัวอย่างที่เทียบกับภาพฝึกสอนแล้ว ได้ระยะทางยูคลีเดียนที่สั้นที่สุด ว่าอยู่ในกลุ่มใดของภาพฝึกสอนจากนั้นก็กำหนดให้พีเจอร์ของภาพตัวอย่าง Y กลุ่มนั้นด้วย

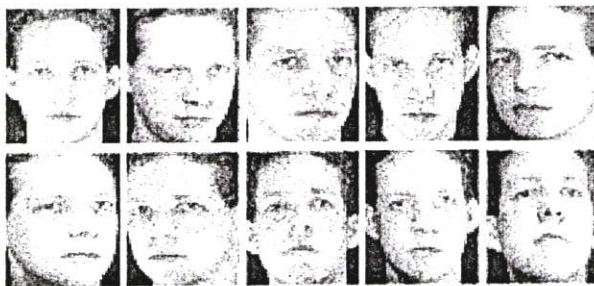
บทที่ 5

การทดลองและผลการทดลอง

บทนี้เป็นการทดลองเปรียบเทียบความถูกต้องในการรู้จำระหว่างวิธี Eigenimages วิธี CSLDA วิธี WPS-CSLDA และวิธีอื่นๆ ภายใต้งैเงื่อนไขต่าง ๆ ที่ทำให้การรู้จำใบหน้ามีความผิดพลาด ซึ่งในการรู้จำใบหน้านั้นได้ทดลองกับฐานข้อมูลใบหน้า 3 แห่ง ได้แก่ ฐานข้อมูลใบหน้า ORL [31] ฐานข้อมูลใบหน้า AR [32],[33] และฐานข้อมูลใบหน้า Yale [34] โดยที่ฐานข้อมูลใบหน้า ORL ใช้ทดสอบความถูกต้องในการรู้จำใบหน้าภายใต้งैเงื่อนไข การเปลี่ยนแปลงตำแหน่งและจำนวนตัวอย่างในการเรียนรู้ ฐานข้อมูลใบหน้า AR ใช้ทดสอบความถูกต้องในการรู้จำใบหน้าภายใต้งैเงื่อนไขของระยะเวลาในการถ่ายภาพ การแสดงออกทางอารมณ์บนใบหน้า ความสว่างบนใบหน้า และฐานข้อมูลใบหน้า Yale ใช้ทดสอบความถูกต้องในการรู้จำใบหน้าภายใต้งैเงื่อนไขการเปลี่ยนแปลงการแสดงออกบนใบหน้า ความสว่างบนใบหน้า ลักษณะและรูปแบบการคำนวณหาความถูกต้องในการรู้จำภาพใบหน้า แสดงในรูปของภาคผนวก ก.

5.1 การทดลองบนฐานข้อมูลภาพใบหน้า ORL

ฐานข้อมูลภาพใบหน้าของ ORL ถูกสร้างขึ้นในช่วงเดือนเมษายน พ.ศ. 2538 ถึง เดือนเมษายน พ.ศ. 2540 ที่ Olivetti Research Laboratory ของมหาวิทยาลัย Cambridge ประเทศสหรัฐอเมริกา ภาพใบหน้าที่ในฐานข้อมูลนี้ทั้งหมดจะเป็นภาพระดับสีเทาและมีพื้นหลังเป็นสีดำ ขนาด 112 x 92 จุดภาพ ถ่ายจากใบหน้าของ 40 บุคคล แบ่งเป็นผู้ชาย 36 คน และเป็นผู้หญิง 4 คน คนละ 10 ภาพที่แตกต่างกัน เช่น ภาพที่มีการแสดงออกทางอารมณ์บนใบหน้า (การปิดตาหรือเปิดตา ยิ้มหรือไม่ยิ้ม) ภาพที่มีรายละเอียดอื่น ๆ (การใส่แว่นหรือถอดแว่น) ภาพที่การก้มหน้าหรือเงยหน้า หันซ้ายหรือขวา ไม่เกินร้อยละ 20 โดยที่ รูปที่ 5.1 เป็นตัวอย่างภาพใบหน้าทั้ง 10 ภาพของบุคคลเดียว

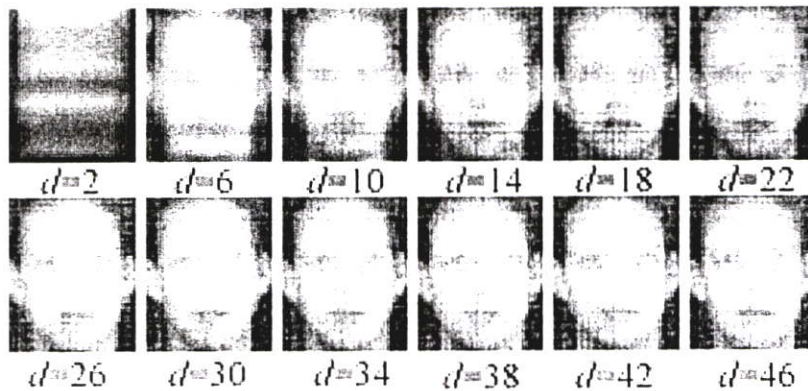


รูปที่ 5.1 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล ORL

5.1.1 การทดสอบการสร้างภาพใหม่ด้วยวิธี WPS-CSLDA บนฐานข้อมูล ORL

เนื่องจาก WPS-CSLDA เป็นทฤษฎีบนพื้นฐานของ LDA ดังนั้นจึงมีคุณสมบัติอย่างหนึ่งเหมือนกันคือ สามารถสร้างภาพใบหน้าใหม่ได้ การที่จะทดสอบว่าวิธี WPS-CSLDA ที่ได้เสนอในวิทยานิพนธ์นี้มีความถูกต้องหรือไม่จึงทำการทดลองสร้างภาพใบหน้าใหม่

การทดสอบคุณสมบัติการสร้างภาพใหม่ของวิธี WPS-CSLDA นั้น ได้ทำการสุ่มเลือกภาพที่ใช้ฝึกสอนจำนวน 5 ภาพต่อหนึ่งบุคคล เพื่อนำมาหาแกนการฉายที่เหมาะสมที่สุดด้วยวิธี WPS-CSLDA เมื่อนำภาพแรกในรูปที่ 5.1 มาทดลองฉายลงบนแกนการฉายที่เหมาะสมที่สุดจะได้พีเจอร์เมตริกที่ใช้ในการสร้างภาพใหม่ รูปที่ 5.2 แสดงให้เห็นว่าเมื่อจำนวน มิติ d (d คือ $r = c$) เพิ่มขึ้นภาพที่สร้างใหม่จะมีความเหมือนภาพต้นแบบมากขึ้น



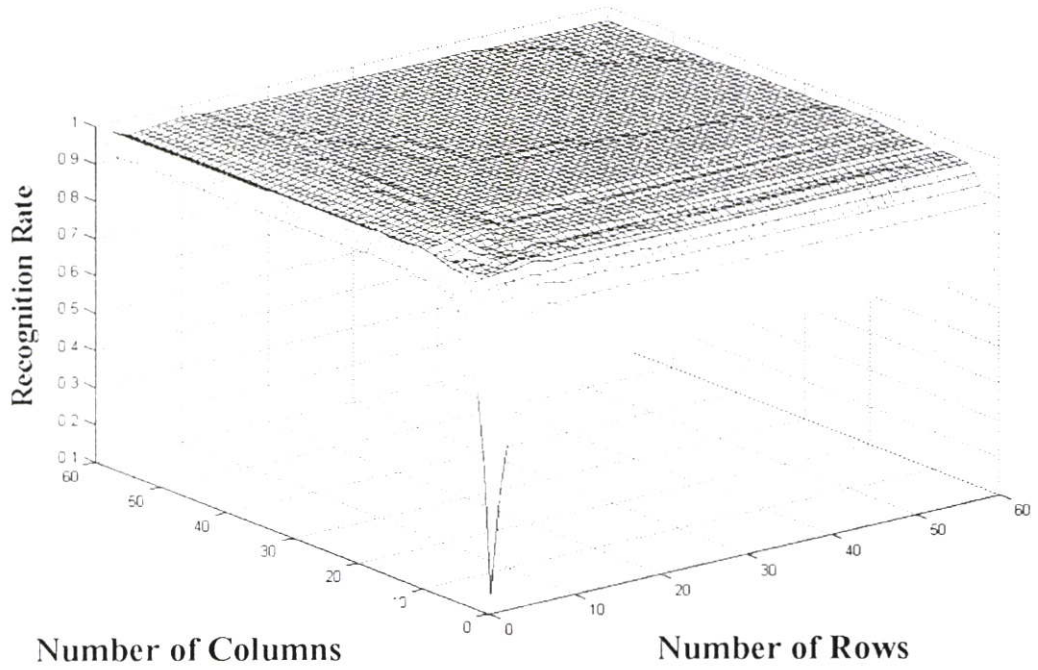
รูปที่ 5.2 ภาพใบหน้าที่สร้างขึ้นใหม่ด้วยวิธี WPS-CSLDA

จากรูปที่ 5.2 พบว่าวิธี WPS-CSLDA ก็มีคุณสมบัติการสร้างภาพใหม่เหมือนกับวิธีบนพื้นฐานของ LDA โดยที่ภาพที่สร้างใหม่จะเหมือนกับภาพต้นแบบขึ้นอยู่กับจำนวนมิติของแกนการฉายทั้ง 2 แกน ถ้าจำนวนมิติมากขึ้นภาพใบหน้าที่สร้างใหม่ก็จะมีเหมือนมากขึ้นตามไปด้วย

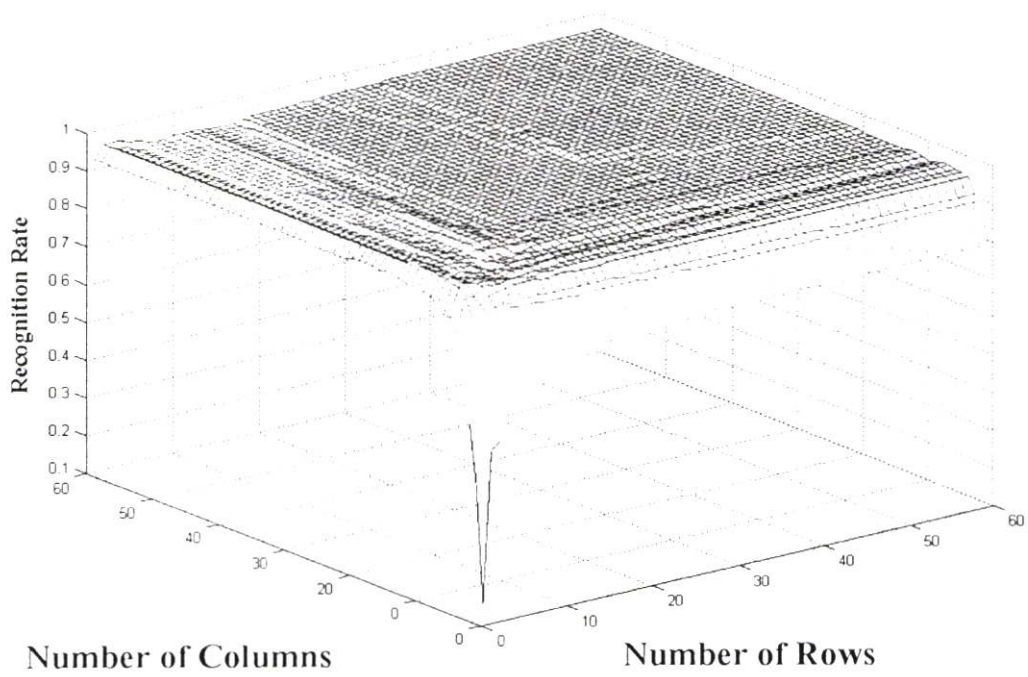
5.1.2 การทดลองเปรียบเทียบความถูกต้องในการรู้จำใบหน้าระหว่างวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages

การทดลองที่ 2 นี้จะทำการเปรียบเทียบความถูกต้องในการรู้จำระหว่างวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages ภายใต้เงื่อนไขของจำนวนตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอน โดยแบ่งการทดสอบเป็น 4 ครั้ง ครั้งแรกภาพใบหน้าที่ใช้ในการฝึกสอนจะถูกเลือกมาอย่างสุ่ม 2 ภาพต่อหนึ่งบุคคล และภาพที่เหลือถูกนำมาใช้ในการทดสอบความถูกต้อง การทดสอบครั้งต่อๆ มาภาพที่ใช้ในการฝึกสอนจะถูกเลือกมาอย่างสุ่มเพิ่มเป็น 3, 4 และ 5 ภาพ ตามลำดับ

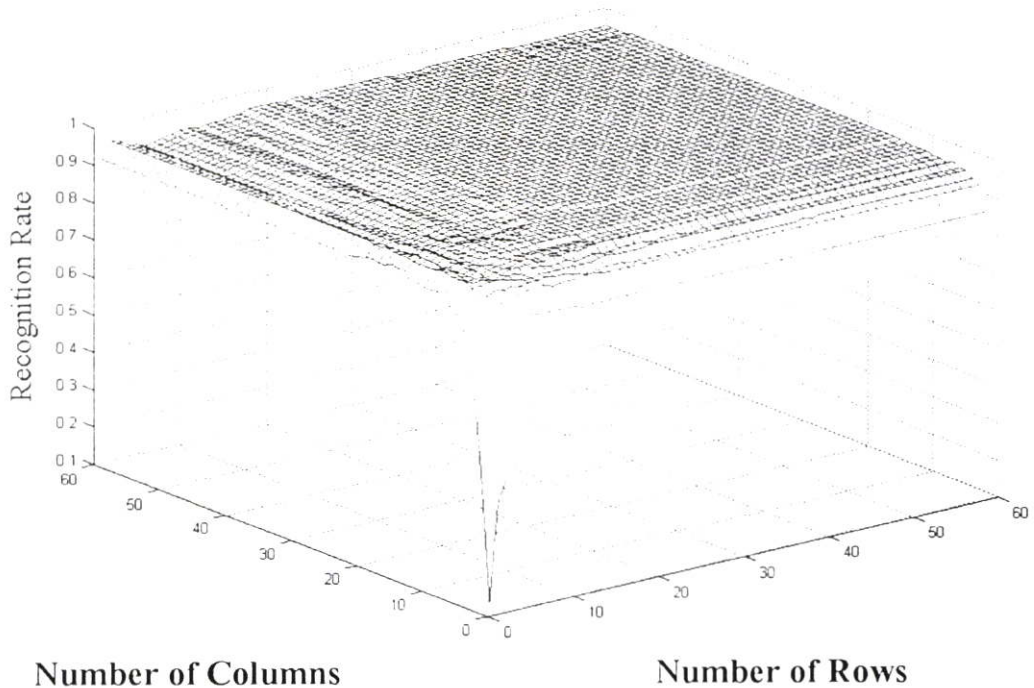
การทดลองของทั้ง 3 วิธีนั้นได้ทำการหาความถูกต้องในการรู้จำภาพใบหน้าที่ตั้งแต่พีเจอร์ที่ $1 \times 1, 1 \times 2, 1 \times 3$ ไปจนถึง พีเจอร์ที่ 60×60 ดังแสดงในรูปที่ 5.3



(n)



(v)

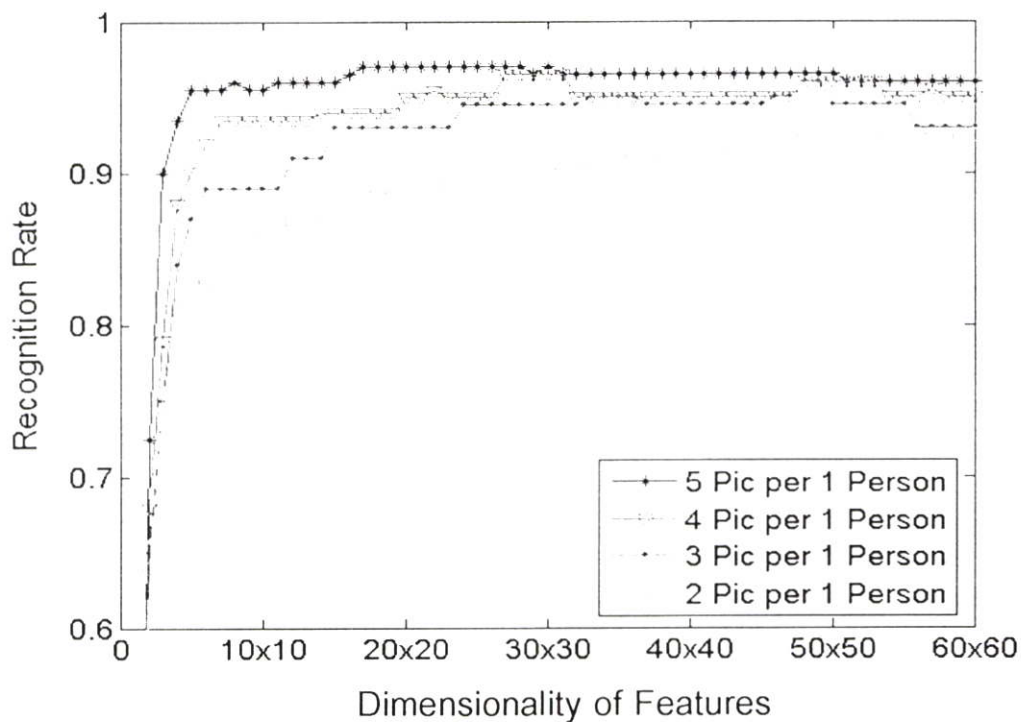


(ค)

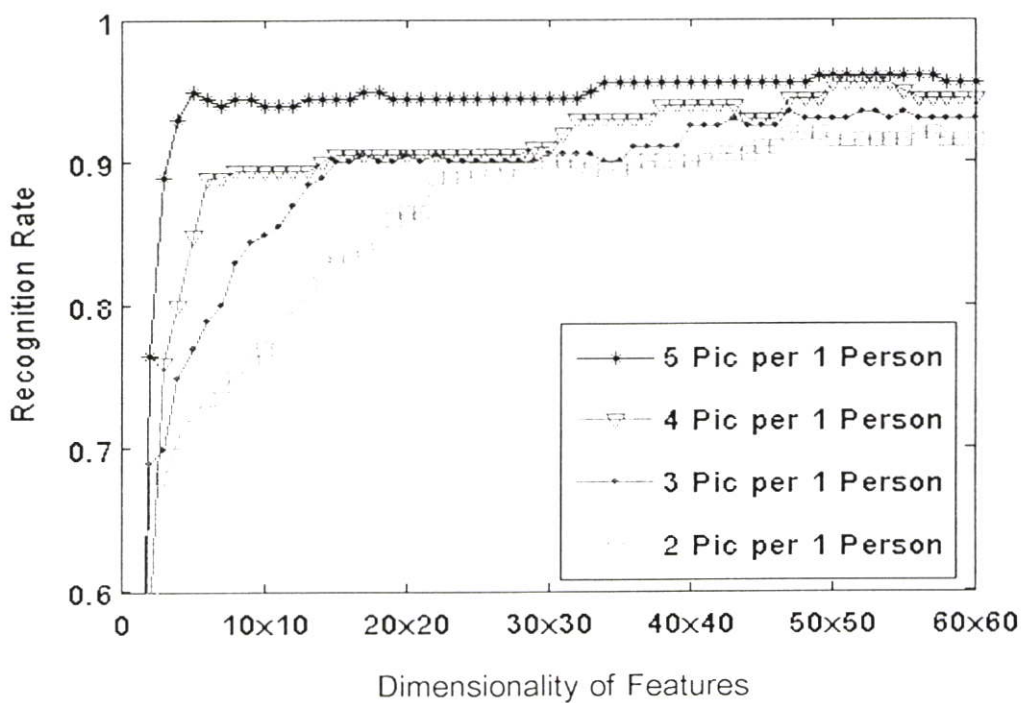
รูปที่ 5.3 ความถูกต้องในการรู้จำใบหน้าของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages

- (ก) ความถูกต้องในการรู้จำของวิธี WPS-CSLDA
- (ข) ความถูกต้องในการรู้จำของวิธี CSLDA
- (ค) ความถูกต้องในการรู้จำของวิธี Eigenimages

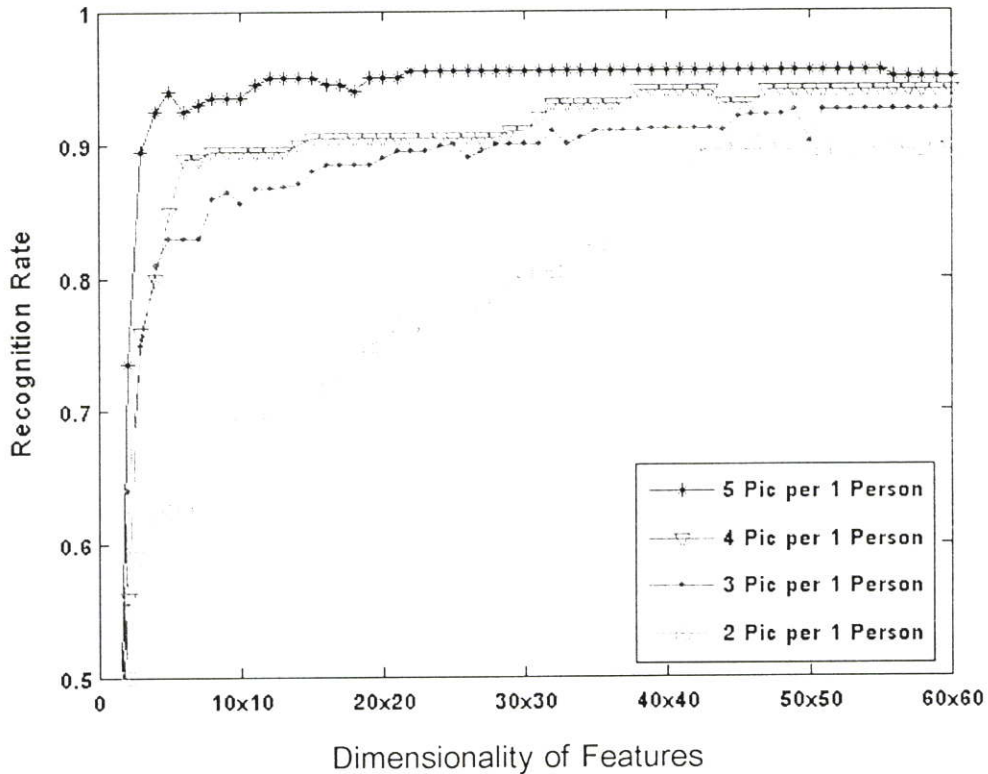
การเปรียบเทียบความถูกต้องในการรู้จำของทั้ง 3 วิธีนั้น ทำได้ยากเนื่องจากความถูกต้องในการรู้จำของทั้ง 3 วิธี แสดงเป็นกราฟ 3 มิติ ดังรูปที่ 5.3 ฉะนั้นการเปรียบเทียบจึงจำเป็นต้องเลือกความถูกต้องในการรู้จำมา 1 แนวเพื่อมาเปรียบเทียบ ในที่นี้จะนำความถูกต้องการรู้จำในแนวทแยงมาเปรียบเทียบ ผลการเปรียบเทียบภายใต้เงื่อนไขของภาพตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอนของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages แสดงในรูปที่ 5.4 รูปที่ 5.5 และ รูปที่ 5.6 ตามลำดับ



รูปที่ 5.4 ความถูกต้องในการรู้จำใบหน้าของวิธี WPS-CSLDA ภายใต้เงื่อนไขของตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอน



รูปที่ 5.5 ความถูกต้องในการรู้จำใบหน้าของวิธี CSLDA ภายใต้เงื่อนไขของตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอน



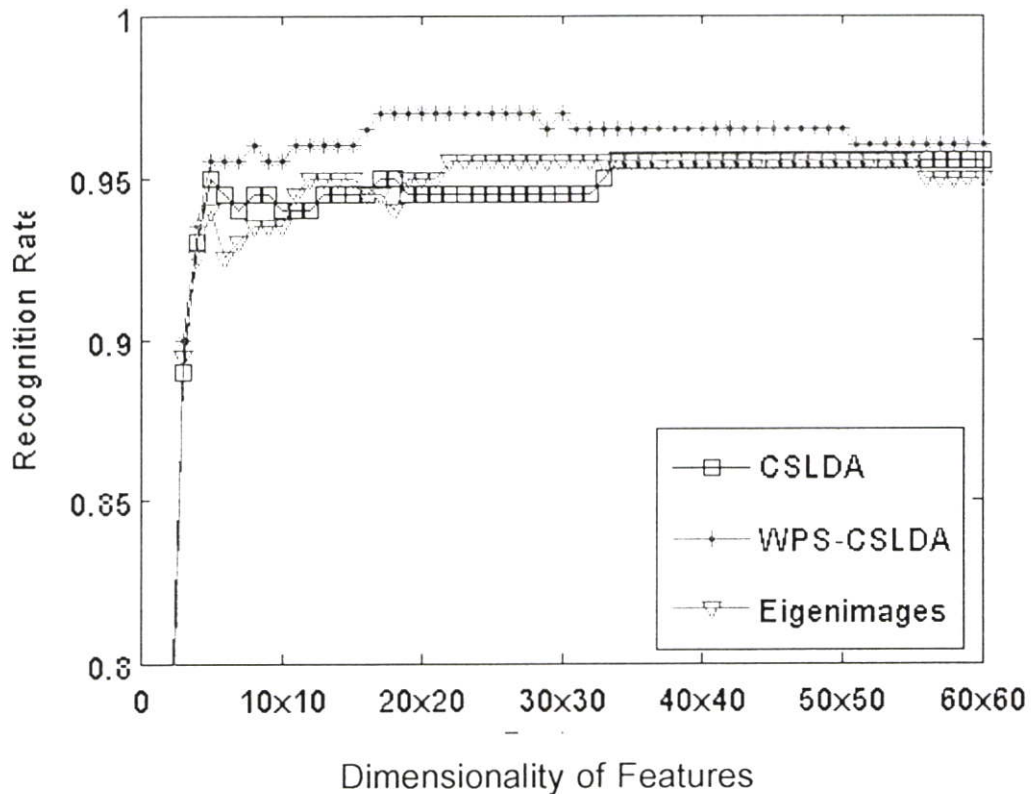
รูปที่ 5.6 ความถูกต้องในการรู้จำใบหน้าของวิธี Eigenimages ภายใต้เงื่อนไขของตัวอย่างภาพ ใบหน้าที่ใช้ในการฝึกสอน

จากรูปที่ 5.4 รูปที่ 5.5 และรูปที่ 5.6 สามารถหาเปอร์เซ็นต์ความถูกต้องในการรู้จำภาพ ใบหน้าสูงสุดดัง ตารางที่ 5.1 เป็นการเปรียบเทียบเปอร์เซ็นต์ความถูกต้องในการรู้จำใบหน้า ระหว่างวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้ในการฝึกสอน

ตารางที่ 5.1 การเปรียบเทียบการรู้จำสูงสุดของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้ในการฝึกสอน

จำนวนภาพต่อบุคคลที่ใช้ในการฝึกสอน	WPS-CSLDA		CSLDA		Eigenimages	
	ความถูกต้องในการรู้จำสูงสุด (%)	ขนาดฟีเจอร์	ความถูกต้องในการรู้จำสูงสุด (%)	ขนาดฟีเจอร์	ความถูกต้องในการรู้จำสูงสุด (%)	ขนาดฟีเจอร์
2	94.37	12x2	93.44	14x2	91.01	13x2
3	96.40	21x2	95.71	22x3	93.33	13x3
4	97.35	14x2	96.25	10x2	94.17	21x2
5	98.50	13x2	97.50	13x2	95.50	18x2

จากตารางที่ 5.1 แสดงให้เห็นว่าเปอร์เซ็นต์ความถูกต้องในการรู้จำใบหน้าสูงสุดของ วิธี WPS-CSLDA เมื่อเปรียบเทียบกับวิธี CSLDA และวิธี Eigenimages แล้วแสดงให้เห็นว่าวิธี WPS-CSLDA มีความถูกต้องในการรู้จำใบหน้าสูงสุด มีค่ามากที่สุด และเมื่อนำผลการรู้จำภาพใบหน้าใน แนวทางนี้ ทั้งหมดมาเปรียบเทียบกันทั้ง 3 วิธี โดยใช้ภาพของบุคคลที่ใช้ในการฝึกสอน 5 ภาพ พบว่าวิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าภายใต้เงื่อนไขของภาพตัวอย่างภาพใบหน้าที่ใช้ ในการฝึกสอนดีที่สุด ดังรูปที่ 5.7

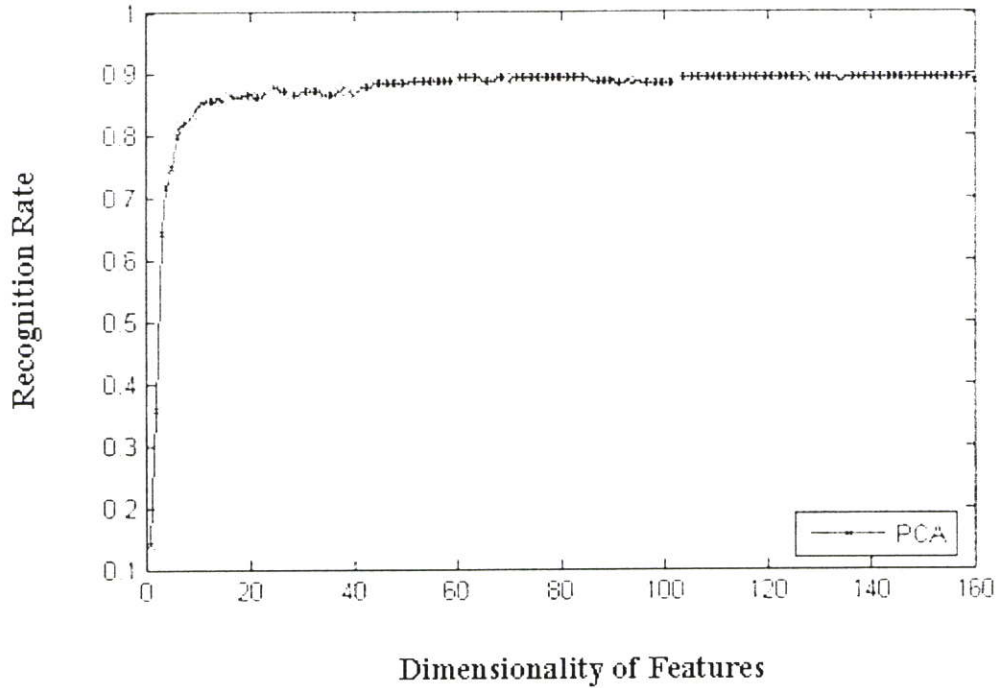


รูปที่ 5.7 เปรียบเทียบความถูกต้องในการรู้จำใบหน้าในแนวทางของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages

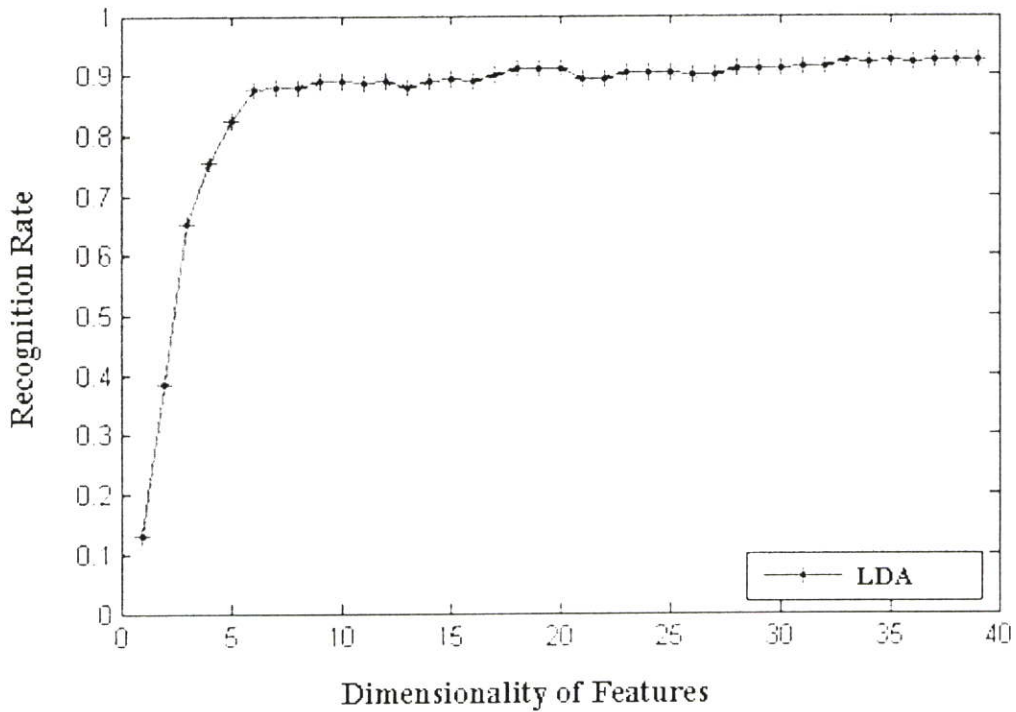
5.1.3 การทดลองเปรียบเทียบความถูกต้องในการรู้จำใบหน้าระหว่างวิธี WPS-CSLDA กับวิธีอื่น ๆ

การทดลองที่ 3 จะทำการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าวิธี WPS-CSLDA กับวิธีการอื่น ๆ คือ วิธี PCA, วิธี LDA, วิธี 2DPCA และวิธี 2DLDA ในการทดลองภาพที่ใช้ในการฝึกสอนถูกเลือกมาอย่างสุ่มจำนวน 5 ภาพต่อหนึ่งบุคคลและภาพที่เหลือจะถูกนำมาใช้ทดสอบ ดังนั้นภาพที่ใช้ฝึกสอนมีจำนวนทั้งหมด 200 ภาพ และภาพที่ใช้ในการทดสอบ 200 ภาพ ในการทดลอง วิธี PCA ได้ทำการหาค่าร้อยละความถูกต้องในการรู้จำตั้งแต่พีเจอร์ที่ 1 จนถึงพีเจอร์มากที่สุดที่ 160 (โดยที่ N คือจำนวนภาพทั้งหมด) และวิธี LDA ได้หาค่าร้อยละความถูกต้องในการรู้จำ

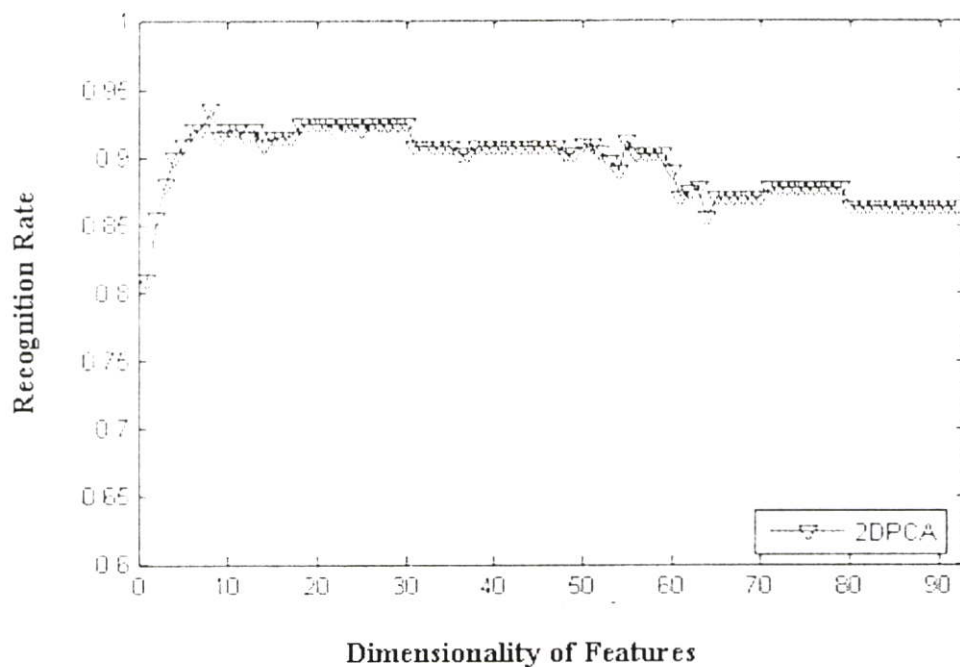
ตั้งแต่ฟีเจอร์ที่ 1 จนถึงฟีเจอร์มากที่สุดที่ 39 ส่วนวิธี 2DPCA และวิธี 2DLDA ได้ทำการหาค่าร้อยละความถูกต้องในการรู้จำตั้งแต่ฟีเจอร์ที่ 112 x 1 จนถึงฟีเจอร์มากที่สุดที่ 112 x 92 ซึ่งผลการทดลองแสดงใน รูปที่ 5.8



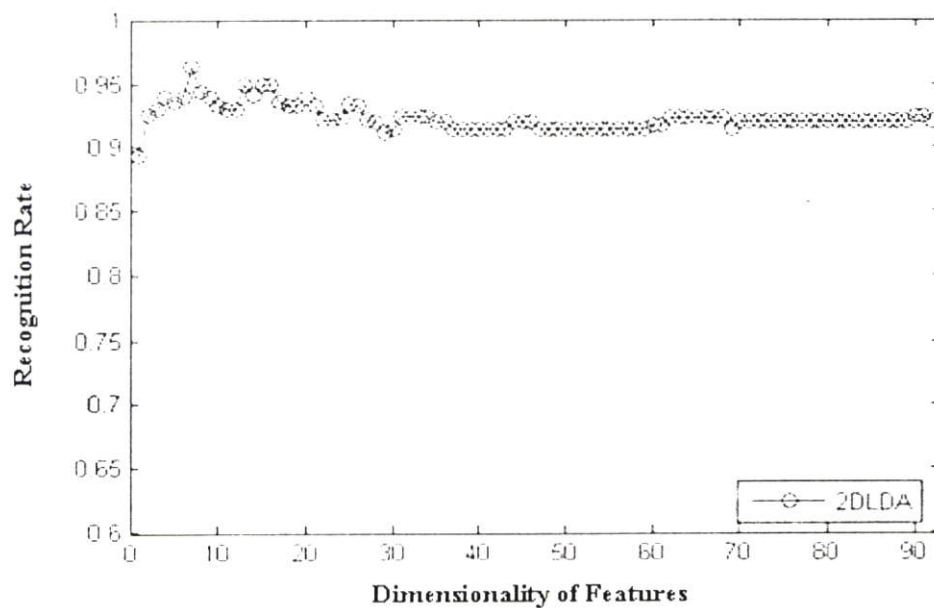
(ก)



(ข)



(ค)



(ง)

รูปที่ 5.8 ความถูกต้องในการรู้จำของวิธี 2DLDA กับวิธี PCA วิธี LDA และวิธี 2DPCA

- (ก) ความถูกต้องในการรู้จำของวิธี PCA
- (ข) ความถูกต้องในการรู้จำของวิธี LDA
- (ค) ความถูกต้องในการรู้จำของวิธี 2DPCA
- (ง) ความถูกต้องในการรู้จำของวิธี 2DLDA

จากรูป 5.8 แสดงให้เห็นว่าความถูกต้องในการรู้จำของวิธี 2DLDA, วิธี LDA, วิธี 2DPCA และวิธี PCA ไม่สามารถนำเส้นกราฟมาเปรียบเทียบกับความถูกต้องในการรู้จำของวิธี WPS-CSLDA ได้ เนื่องจาก ความถูกต้องในการรู้จำของวิธี WPS-CSLDA แสดงเป็นกราฟ 3 มิติ ดังรูปที่ 5.3 (ก) ซึ่งมีขนาดของพีเจอร์ต่างกัน ดังนั้นในการเปรียบเทียบของวิธี 2DLDA, วิธี LDA, วิธี 2DPCA และวิธี PCA กับวิธี WPS-CSLDA นั้น จะนำเฉพาะ ความถูกต้องในการรู้จำภาพใบหน้าสูงสุดมาเปรียบเทียบกัน ดังแสดงในตารางที่ 5.2

ตารางที่ 5.2 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าสูงสุดของวิธี 2DLDA,วิธี LDA, วิธี 2DPCA และวิธี PCA กับวิธี WPS-CSLDA

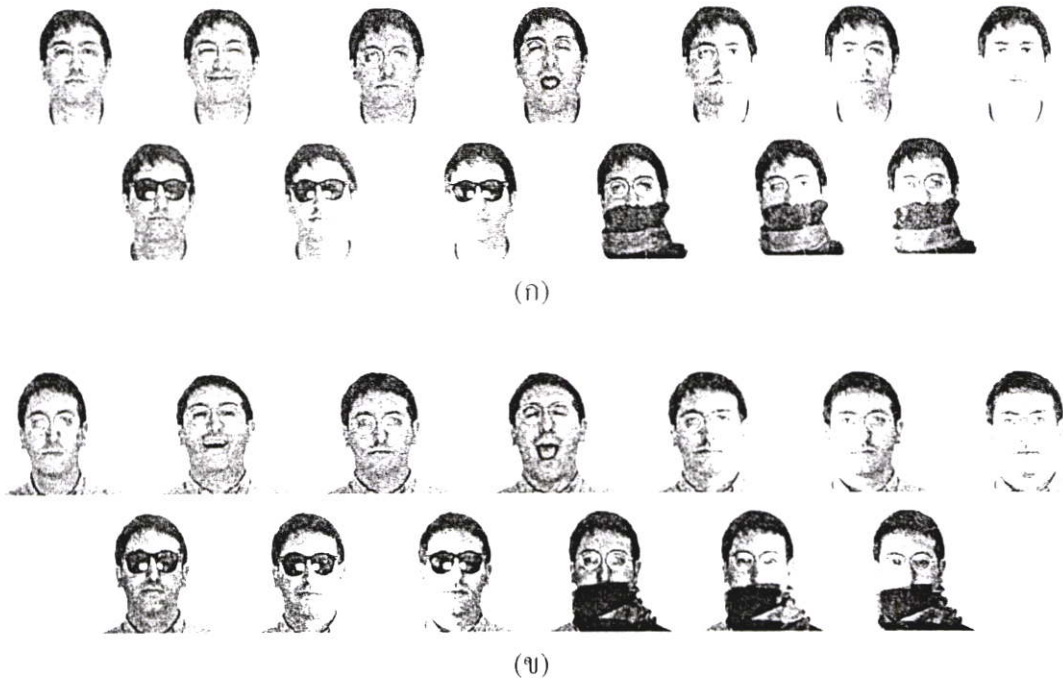
วิธี	เปอร์เซ็นต์ความถูกต้องในการรู้จำสูงสุด	ขนาดของพีเจอร์
WPS-CSLDA	98.50	13x2
PCA	89.57	60
LDA	92.53	33
2DPCA	93.50	112 x 8
2DLDA	96.00	112 x 7

จากตารางที่ 5.2 พบว่าวิธี WPS-CSLDA มีความถูกต้องในการรู้จำภาพใบหน้าสูงสุด ดีที่สุดใน 4 วิธี โดยผลการรู้จำภาพใบหน้าที่ดีที่สุดของวิธี WPS-CSLDA คือ 98.50 เปอร์เซ็นต์ ที่พีเจอร์มีขนาดเท่ากับ 13x2

5.2 การทดลองบนฐานข้อมูลภาพใบหน้า AR

ฐานข้อมูลภาพใบหน้าของ AR ถูกสร้างขึ้นโดย Aleix Martinez และ Robert Benavente ใน Computer Vision Center (CVC) [35] ประกอบด้วยภาพใบหน้าสีขนาด 768 x 576 จุดภาพ ที่ความละเอียดของสี 24 บิต มากกว่า 4,000 ภาพ ของ 126 บุคคล แบ่งเป็นผู้ชาย 70 คน และเป็นผู้หญิง 56 คน ซึ่งถ่ายเฉพาะด้านหน้าที่แตกต่างกันตามเงื่อนไขต่าง ๆ เช่น เงื่อนไขการแสดงออกทางอารมณ์บนใบหน้า เงื่อนไขของความสว่างบนใบหน้า เงื่อนไขการสวมแว่นตากันแดดและผ้าพันคอ

ภาพใบหน้าของฐานข้อมูล AR ที่สามารถให้นำมาใช้ได้มีเพียง 120 บุคคล แบ่งเป็นชาย 65 คน และหญิง 55 คน โดยภาพใบหน้าของแต่ละบุคคลจะถูกแบ่งออกเป็น 2 กลุ่ม กลุ่มแรกประกอบด้วยภาพใบหน้า 13 ภาพที่แตกต่างกัน ส่วนกลุ่มที่สองจะมีลักษณะภาพใบหน้าเหมือนกับกลุ่มแรกแต่ถ่ายภาพห่างกันเป็นระยะเวลา 2 สัปดาห์ ดังแสดงในรูปที่ 5.9 รูปที่ 5.9 เป็นตัวอย่างภาพใบหน้า AR ของบุคคลคนหนึ่งที่ถูกแบ่งเป็นสองกลุ่ม โดยที่รูปที่ 5.9 (ก) เป็นกลุ่มที่ 1 และรูปที่ 5.9 (ข) เป็นกลุ่มที่ 2



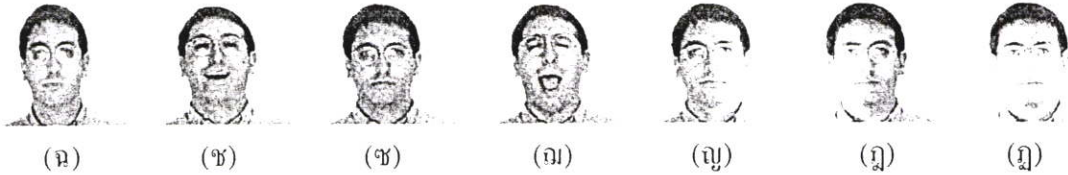
รูปที่ 5.9 ตัวอย่างภาพใบหน้า AR ของบุคคลคนหนึ่งที่ถูกแบ่งเป็นสองกลุ่ม
 (ก) ตัวอย่างภาพใบหน้าของบุคคลคนหนึ่งที่ถูกแบ่งเป็นกลุ่มที่ 1
 (ข) ตัวอย่างภาพใบหน้าของบุคคลคนหนึ่งที่ถูกแบ่งเป็นกลุ่มที่ 2

ก่อนการทดลองได้ทำการเลือกภาพใบหน้าของแต่ละบุคคล ทั้ง 120 คนมาตามรูปที่ 5.10 โดยที่รูปในกลุ่มที่ 1 ประกอบด้วยรูปที่ 5.10 (ก) เป็นภาพใบหน้าปกติ รูปที่ 5.10 (ข) เป็นภาพใบหน้ายิ้ม รูปที่ 5.10 (ค) เป็นภาพใบหน้าโกรธ รูปที่ 5.10 (ค) เป็นภาพใบหน้าแสดงความกลัว รูปที่ 5.10 (ฅ) เป็นภาพใบที่มีแสงส่องด้านขวา รูปที่ 5.10 (ง) เป็นภาพใบที่มีแสงส่องด้านซ้าย และรูปที่ 5.10 (จ) เป็นภาพใบที่มีแสงส่องทั้งสองด้าน ส่วนกลุ่มที่ 2 ทำเหมือนกลุ่มแรกจะได้ดังรูปที่ 5.10 (ฉ) , รูปที่ 5.10 (ซ) , รูปที่ 5.10 (ซ), รูปที่ 5.10 (ฌ) , รูปที่ 5.10 (ญ) , รูปที่ 5.10 (ฎ) , รูปที่ 5.10 (ฎ) ดังนั้นภาพที่ทำการตัดทิ้งจากรูปที่ 5.10 คือภาพที่มีการสวมแว่นตาและผ้าพันคอ หลังจากนั้นทำการแปลงภาพทั้งหมดให้เป็นภาพระดับสีเทาที่มีความละเอียดของสี 8 บิต และเนื่องจากภาพใบหน้าในข้อมูลนี้ขนาดใหญ่จึงปรับภาพให้มีขนาด 50 x 40 จุดภาพ

กลุ่มที่ 1



กลุ่มที่ 2



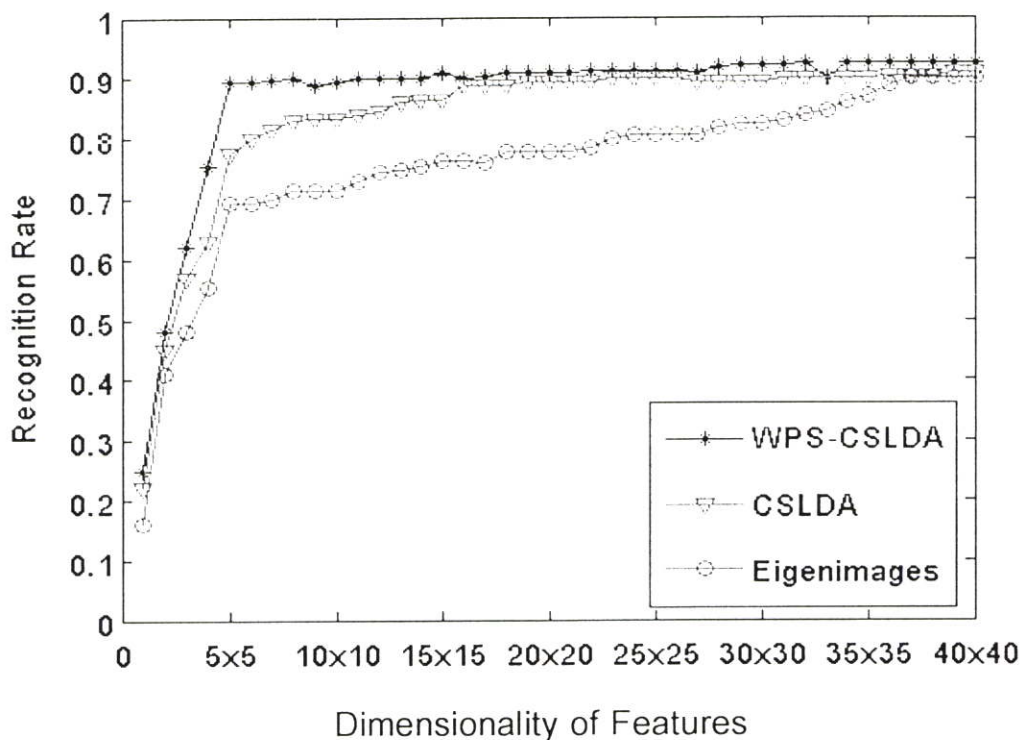
รูปที่ 5.10 ตัวอย่างภาพใบหน้าของ 1 บุคคลที่ถูกเลือกมาใช้ในการทดลอง

5.2.1 การทดลองกับเงื่อนไขการถ่ายภาพต่างเวลา

การทดลองนี้ได้ทำการเปรียบเทียบวิธี WPS-CSLDA กับ วิธี CSLDA และวิธี Eigenimages ภายใต้เงื่อนไขการถ่ายภาพต่างเวลา โดยการทดลองได้ใช้ภาพจากรูปที่ 5.10 กลุ่มที่ 1 ทั้งหมดมาใช้สอนและภาพจากรูปที่ 5.10 กลุ่มที่ 2 ถูกนำมาใช้ทดสอบ ดังนั้นภาพที่ใช้ในการสอนทั้งหมดจึงมี 840 ภาพ และภาพที่ใช้ในการทดสอบทั้งหมด 840 ภาพ

การหาพีเจอร์ของภาพของทั้ง 3 วิธีนั้นได้ทำการหาความถูกต้องในการรู้จำภาพใบหน้า ตั้งแต่พีเจอร์ที่ $1 \times 1, 1 \times 2, 1 \times 3, \dots, 2 \times 1, 2 \times 2$ ไปจนถึง พีเจอร์ที่ 40×40 จากนั้นจะใช้วิธี INN เป็นการจำแนกบุคคลโดยใช้ความใกล้เคียงกันมากที่สุดของภาพใบหน้าที่ถูกหาพีเจอร์ ผลการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าของทั้ง 3 วิธีเปรียบเทียบกันได้ยากเนื่องจากเป็นกราฟ 3 มิติคล้ายกับรูปที่ 5.4 ดังนั้นในการเปรียบเทียบ จะนำแนวทแยงมุมของกราฟความถูกต้องในการรู้จำทั้ง 3 วิธีมาเปรียบเทียบกัน ดังแสดงใน รูปที่ 5.11

การคำนวณหาเวลาที่ใช้ในการหาพีเจอร์ของภาพจางทั้ง 3 วิธีแสดงในตารางที่ 5.3 จากตารางที่ 5.3 จะเห็นได้ว่า ทั้ง 3 วิธี มีความเร็วในการคำนวณใกล้เคียงกันจะแตกต่างกันแค่ระดับทศนิยม แต่เมื่อดูที่ความถูกต้องในการรู้จำภาพใบหน้าจะพบว่า ความถูกต้องในการรู้จำภาพใบหน้าด้วยวิธี WPS-CSLDA มีค่าดีที่สุด



รูปที่ 5.11 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าในแนวทแยงของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้ตัวแปรของการถ่ายภาพต่างเวลา

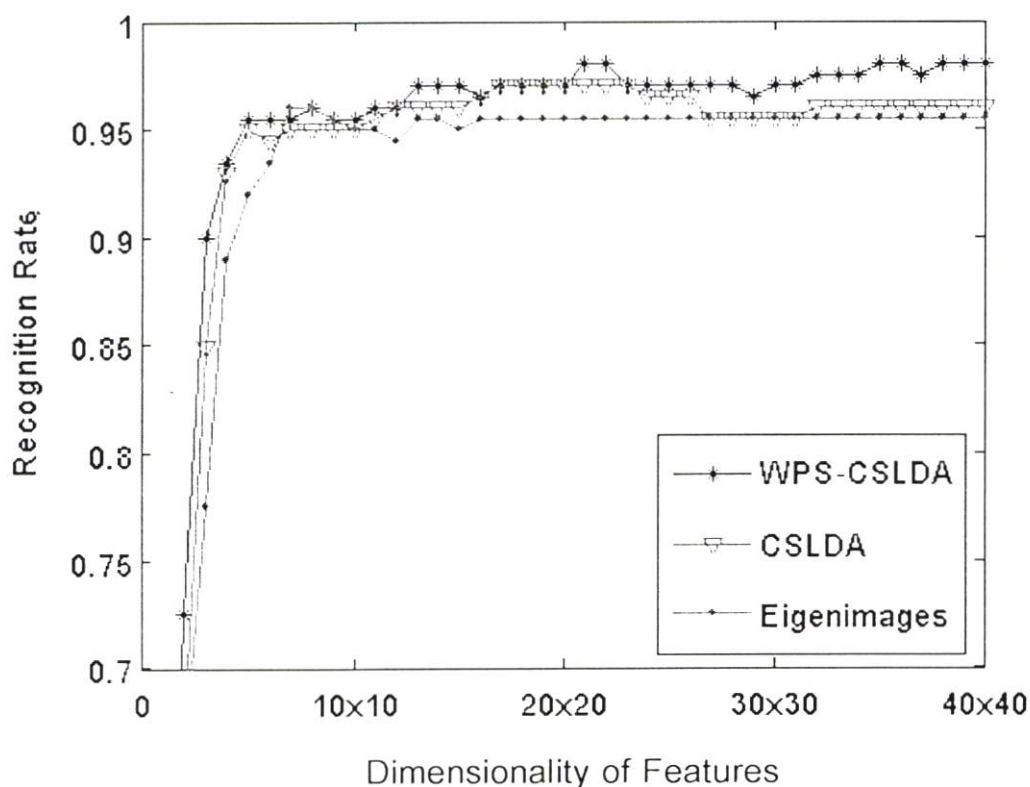
ตารางที่ 5.3 เปรียบเทียบวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้ตัวแปรของการถ่ายภาพต่างเวลา เงื่อนไขการแสดงออกทางอารมณ์บนใบหน้า เงื่อนไขแสงบนใบหน้า

การทดลองภายใต้เงื่อนไขของการเปลี่ยนแปลงบนใบหน้า		เวลาในการคำนวณหาพีเจอร์ของภาพ (วินาที)	ความถูกต้องในการรู้จำสูงสุด (เปอร์เซ็นต์)
เงื่อนไข	วิธี		
การถ่ายภาพต่างเวลา	WPS-CSLDA	0.553	93.75
	CSLDA	0.501	91.40
	Eigenimages	0.482	90.275
การแสดงออกทางอารมณ์บนใบหน้า	WPS-CSLDA	0.442	98.90
	CSLDA	0.410	98.12
	Eigenimages	0.378	97.47
ตัวแปรทางแสงบนภาพใบหน้า	WPS-CSLDA	0.402	94.65
	CSLDA	0.384	92.63
	Eigenimages	0.327	91.02

5.2.2 การทดลองกับเงื่อนไขการแสดงผลออกทางอารมณ์บนใบหน้า

การทดลองนี้ได้ทำการเปรียบเทียบวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้เงื่อนไขการแสดงผลออกทางอารมณ์บนใบหน้า ก่อนการทดลองได้ทำการเลือกภาพจากฐานข้อมูลทั้งหมดมาเป็นฐานข้อมูลใหม่ ดังนั้นภาพใบหน้าของแต่ละบุคคลที่ถูกเลือกจากกลุ่มที่ 1 เป็น 5.10(ก), 5.10(ข), 5.10(ค), 5.10(ค) และจากกลุ่มที่ 2 เป็น 5.10(ฉ), 5.10(ช), 5.10(ช), 5.10(ฉ) ซึ่งภาพที่เลือกทั้งหมดจะอยู่ภายใต้เงื่อนไขการแสดงผลออกทางอารมณ์บนใบหน้า หลังจากนั้นได้ทำการเลือกภาพที่ใช้ในการสอนจากแต่ละบุคคลมาสองภาพเป็น 5.10(ก) และ 5.10(ฉ) ส่วนภาพที่เหลือถูกนำมาทดสอบความถูกต้องในการรู้จำเป็น 5.10(ข), 5.10(ค), 5.10(ค) และ 5.10(ช), 5.10(ช), 5.10(ฉ) ดังนั้นภาพที่ใช้สอนทั้งหมดจะมี 240 ภาพและภาพที่ใช้ทดสอบ 720 ภาพ

การทดลองที่ 5.2.2 นี้ การหาพีเจอร์ของภาพของทั้ง 3 วิธีนั้นได้ทำการหาความถูกต้องในการรู้จำภาพใบหน้าตั้งแต่พีเจอร์ที่ $1 \times 1, 1 \times 2, 1 \times 3, \dots, 2 \times 1, 2 \times 2$ ไปจนถึง พีเจอร์ที่ 40×40 จากนั้นจะใช้วิธี INN เป็นการจำแนกบุคคลโดยใช้ความใกล้เคียงกันมากที่สุดของภาพใบหน้าที่ถูกหาพีเจอร์แล้ว ผลการรู้จำภาพใบหน้าสูงสุดของทั้ง 3 วิธีแสดงในตารางที่ 5.3 และผลการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าในแนวทแยงของทั้ง 3 วิธี แสดงในรูปที่ 5.12



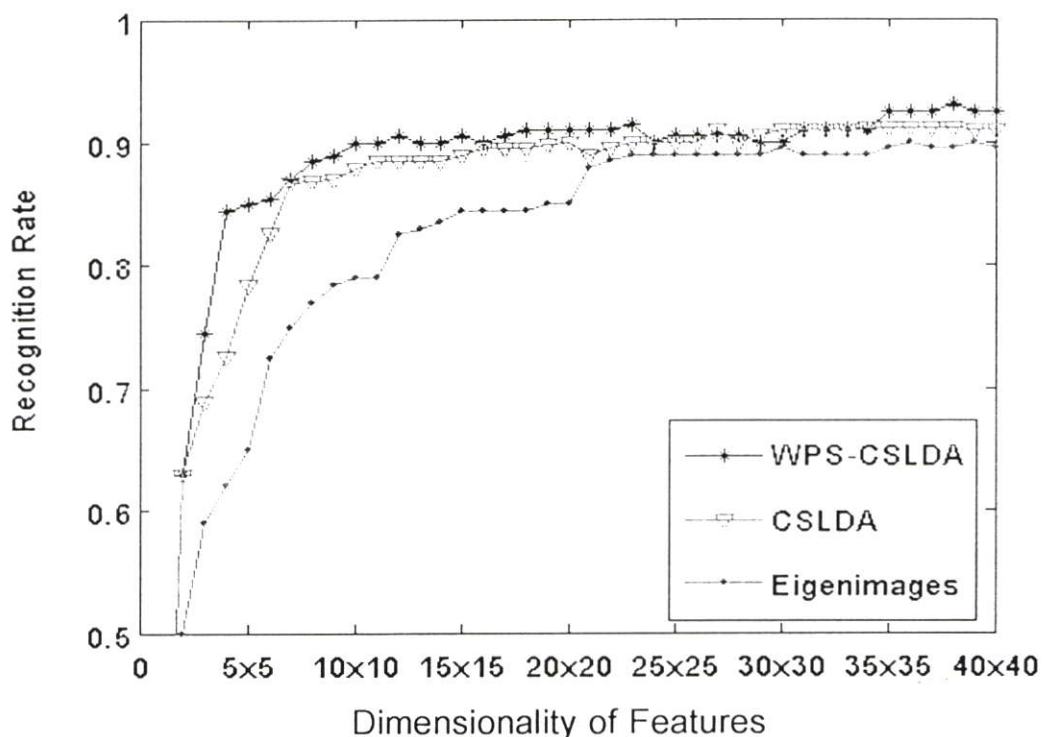
รูปที่ 5.12 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าแนวทแยงภายใต้เงื่อนไขการแสดงผลออกทางอารมณ์บนใบหน้าของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages

จากรูปที่ 5.12 จะเห็นได้ว่าเส้นกราฟของการรู้จำภาพใบหน้าวิธี WPS-CSLDA ค่าดีที่สุดเมื่อเปรียบเทียบกับวิธี CSLDA และวิธี Eigenimages

5.2.3 การทดลองภายใต้เงื่อนไขของความสว่างบนใบหน้า

การทดลองนี้ได้ทำการเปรียบเทียบวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages ภายใต้เงื่อนไขของความสว่างบนใบหน้า โดยทำการเลือกภาพจากฐานข้อมูลทั้งหมดมาเป็นฐานข้อมูลใหม่ซึ่งจะนำมาใช้ในการทดลอง ดังนั้นภาพใบหน้าของแต่ละบุคคลจากกลุ่มที่ 1 เป็นรูปที่ 5.10(ก), รูปที่ 5.10(ข), รูปที่ 5.10(ค), รูปที่ 5.10(ง) และจากกลุ่มที่ 2 เป็นรูปที่ 5.10(ฉ), รูปที่ 5.10(จ), รูปที่ 5.10(ฉ), รูปที่ 5.10(ฉ) หลังจากนั้นทำการเลือกภาพมาอย่างสุ่มคนละ 2 ภาพ จากกลุ่มที่ 1 หนึ่งภาพและจากกลุ่มที่ 2 อีกหนึ่งภาพเพื่อนำมาใช้ในการสอน ส่วนภาพที่เหลือจากการเลือกถูกนำมาใช้ในการทดสอบความถูกต้องในการรู้จำ ดังนั้นภาพที่ใช้ฝึกสอนมีจำนวนทั้งหมด 240 ภาพ และภาพที่ใช้ในการทดสอบ 720 ภาพ

การทดลองที่ 5.2.3 นี้ การหาพีเจอร์ของภาพของทั้ง 3 วิธีนั้นได้ทำการหาความถูกต้องในการรู้จำภาพใบหน้าที่ตั้งแต่พีเจอร์ที่ $1 \times 1, 1 \times 2, 1 \times 3, \dots, 2 \times 1, 2 \times 2$ ไปจนถึง พีเจอร์ที่ 40×40 จากนั้นจะใช้วิธี INN เป็นการจำแนกบุคคลโดยใช้ความใกล้เคียงกันมากที่สุดของภาพใบหน้าที่ถูกหาพีเจอร์แล้ว ผลการรู้จำภาพใบหน้าสูงสุดของทั้ง 3 วิธีแสดงในตารางที่ 5.3 และ ผลการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าในแนวทแยงของทั้ง 3 วิธี แสดงในรูปที่ 5.13

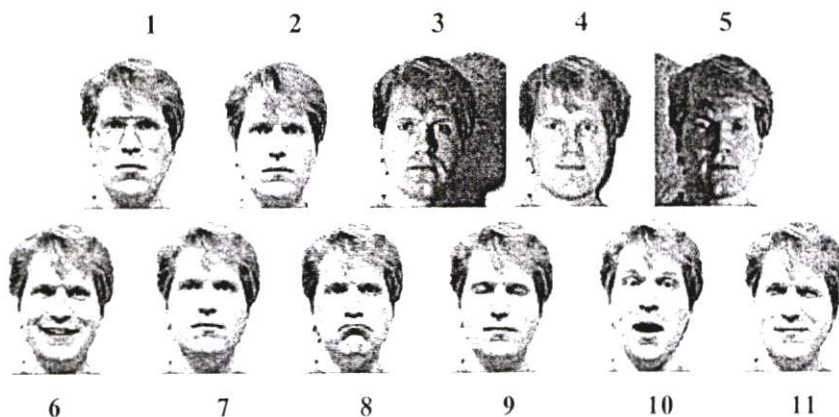


รูปที่ 5.13 เปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าแนวทแยงภายใต้เงื่อนไขของความสว่างบนใบหน้าของวิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimages

5.3 การทดลองบนฐานข้อมูลภาพใบหน้า Yale

ฐานข้อมูลภาพใบหน้าของ Yale ถูกสร้างขึ้นมาจาก Yale Center for Computational Vision and Control โดยภาพใบหน้าทั้งหมดเป็นภาพระดับสีเทาจากหลังมีสีขาวขนาด 320 x 243 จุดภาพ ถ่ายจาก 15 บุคคล คนละภาพ 11 ภาพที่แตกต่างกัน ซึ่งในฐานข้อมูลภาพใบหน้าจะมีทั้งผู้หญิงและผู้ชาย โดยรูปที่ 5.14 เป็นตัวอย่างภาพใบหน้าของบุคคลคนหนึ่ง ซึ่งภาพแรกจะเป็นภาพที่มีแสงล้อมรอบบนลักษณะใบหน้าที่เป็นปกติและบุคคลคนนั้นไม่ใส่แว่น ในภาพที่สองจะเป็นภาพที่มีแสงล้อมรอบบนลักษณะใบหน้าที่เป็นปกติและบุคคลคนนั้นไม่ใส่แว่น ถ้าปกติบุคคลนั้นใส่แว่น ภาพที่ 3 - 5 เป็นภาพที่ได้มาโดยการส่องสว่างของบนใบหน้าธรรมชาติโดยใช้ Luxolamp วางใน 3 จุดตามลำดับ ภาพที่หกลือได้มาจากภาพที่ล้อมรอบไปด้วยแสง แต่มีลักษณะที่แสดงออกทางบนใบหน้าที่แตกต่างกัน (เช่น มีความสุข, เศร้า, หลับ, กระพริบตา, และประหลาดใจ)

เนื่องจากภาพใบหน้าในข้อมูลนี้มีขนาดใหญ่ทำให้การคำนวณใช้เวลามาก ก่อนการทดลองภาพทั้งหมดจึงถูกตัดและปรับให้มีขนาด 100 x 80 จุดภาพ



รูปที่ 5.14 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล Yale

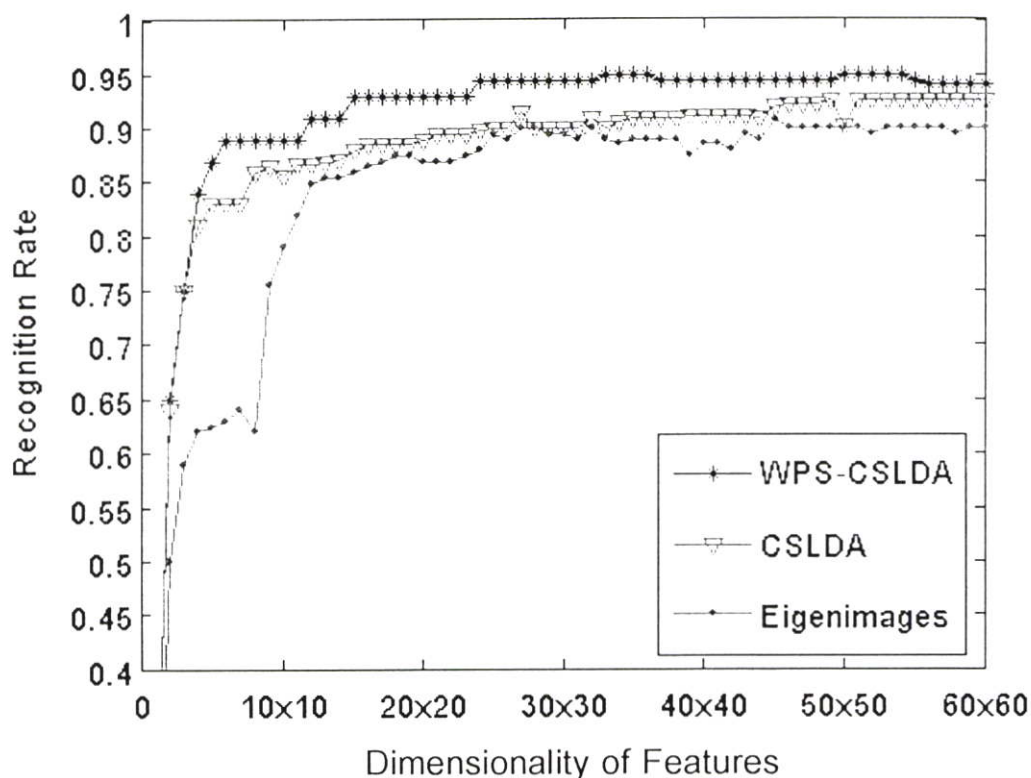
5.3.1 การทดลองภายใต้เงื่อนไขการเปลี่ยนแปลงของการแสดงออกบนใบหน้าและความสว่างบนใบหน้า

การทดลองกับฐานข้อมูล Yale นี้เป็นการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าของวิธี WPS-CSLDA กับวิธี CSLDA และวิธี Eigenimages โดยการเปรียบเทียบนั้นได้ทำการหาความถูกต้องในการรู้จำภาพใบหน้าที่ตั้งแต่พีเจอร์ที่ $1 \times 1, 1 \times 2, 1 \times 3, \dots, 2 \times 1, 2 \times 2$ ไปจนถึง พีเจอร์ที่ 60×60 จากนั้นจะใช้วิธี 1NN เป็นการจำแนกบุคคลโดยใช้ความใกล้เคียงกันมากที่สุดของภาพใบหน้าที่ถูกหาพีเจอร์ การทดสอบภาพใบหน้าที่ใช้ในการฝึกสอนจะถูกเลือกมาอย่างสุ่ม 5 ภาพต่อหนึ่งบุคคล และ 6 ภาพที่เหลือจะถูกนำมาในการทดสอบความถูกต้อง

ผลการเปรียบเทียบความถูกต้องในการรู้จำสูงสุดแสดงในตารางที่ 5.4 และเมื่อจะนำกราฟความถูกต้องในการรู้จำภาพใบหน้าที่ของทั้ง 3 วิธีมาเปรียบเทียบกันนั้น ทำได้ยากเนื่องจากกราฟแสดงผลการเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าที่ของทั้ง 3 วิธีเป็นกราฟ 3 มิติคล้ายกับรูปที่ 5.4 ดังนั้นในการเปรียบเทียบจะนำแนวทแยงมุมของกราฟความถูกต้องในการรู้จำภาพใบหน้าที่ของทั้ง 3 วิธีมาเปรียบเทียบกัน ดังแสดงใน รูปที่ 5.15

ตารางที่ 5.4 เปรียบเทียบความถูกต้องในการรู้จำสูงสุดของวิธีทั้ง 3 วิธีบนฐานข้อมูล Yale

วิธีการหาพีเจอร์	ความถูกต้องในการรู้จำสูงสุด (เปอร์เซ็นต์)	ขนาดพีเจอร์
WPS-CSLDA	95.70	12x2
CSLDA	93.85	13x2
Eigenimages	92.54	14x3



รูปที่ 5.15 เปรียบเทียบความถูกต้องในการรู้จำใบหน้า Yale ในแนวทแยงของวิธี WPS-CSLDA วิธี CSLDA และวิธี Eigenimages

ผลการเปรียบเทียบความถูกต้องในการรู้จำสูงสุดของวิธีทั้ง 3 วิธีบนฐานข้อมูล Yale จากตารางที่ 5.4 แสดงให้เห็นอย่างชัดเจนว่าความถูกต้องในการรู้จำสูงสุดของวิธี WPS-CSLDA มีผลการรู้จำภาพใบหน้าที่ดีที่สุด คือ 95.70% ที่ขนาดฟีเจอร์ 12x2 รองลงมาได้แก่วิธี CSLDA มีความถูกต้องในการรู้จำภาพใบหน้าสูงสุด 93.85% ที่ขนาดฟีเจอร์ 13x2 และสุดท้ายคือวิธี Eigenimages มีผลการรู้จำภาพใบหน้าที่ดีที่สุด 92.54% ที่ขนาดฟีเจอร์ 14x3 และเมื่อนำความถูกต้องในแนวทแยงมุมมาสร้างเป็นเส้นกราฟของทั้ง 3 วิธีดังรูปที่ 5.15 จะให้เห็นได้อย่างชัดเจนว่าความถูกต้องในการรู้จำของวิธี WPS-CSLDA นั้นมีเส้นกราฟความถูกต้องในแนวทแยงมุมสูงที่สุด

5.3.2 การทดสอบความถูกต้องในการรู้จำวิธี WPS-CSLDA ภายใต้เงื่อนไขของจำนวนตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอน

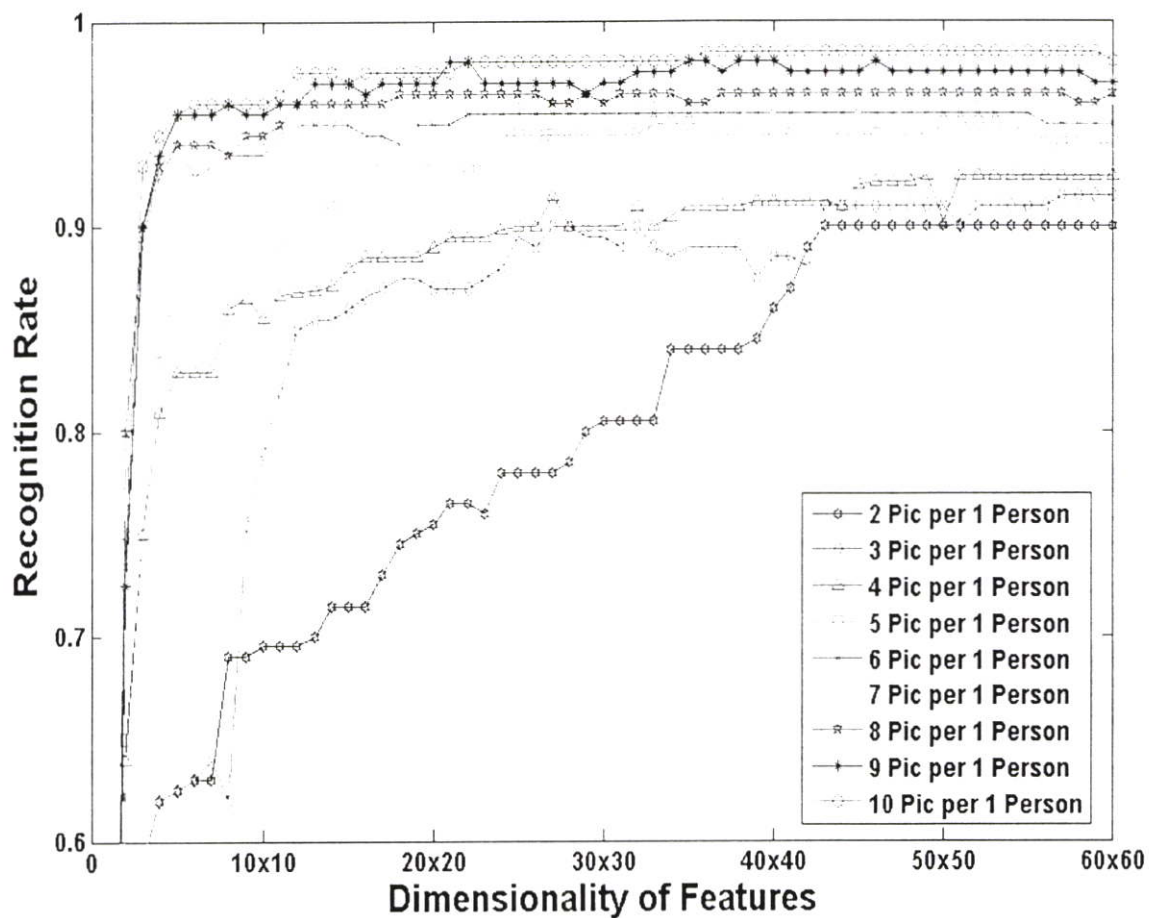
การทดลองภายใต้เงื่อนไขของจำนวนตัวอย่างภาพใบหน้าที่ใช้ในการฝึกสอนนี้แบ่งการทดสอบเป็น 8 ครั้ง ครั้งแรกภาพใบหน้าที่ใช้ในการฝึกสอนจะถูกเลือกมาอย่างสุ่ม 2 ภาพต่อหนึ่งบุคคล และภาพที่เหลือถูกนำมาใช้ในการทดสอบความถูกต้อง การทดสอบครั้งต่อๆ มา ภาพที่ใช้ในการฝึกสอนจะถูกเลือกมาอย่างสุ่มเพิ่มเป็น 3, 4, 5 จนถึง 10 ภาพ ตามลำดับ

ตารางที่ 5.5 การเปรียบเทียบการรู้จำสูงสุดของวิธี WPS-CSLDA ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้ในการฝึกสอน

จำนวนภาพต่อบุคคลที่ใช้ในการฝึกสอน (Training Set)	จำนวนภาพต่อบุคคลที่ใช้ในการทดสอบ (Test Set)	WPS-CSLDA	
		ความถูกต้องในการรู้จำสูงสุด (%)	ขนาดพีเจอร์
2	9	90.79	12x2
3	8	91.98	21x2
4	7	93.11	14x2
5	6	95.70	12x2
6	5	96.23	13x2
7	4	96.99	13x3
8	3	97.73	13x2
9	2	98.12	13x2
10	1	98.98	13x3

จากตารางที่ 5.5 แสดงให้เห็นว่าจำนวนภาพใบหน้าที่ใช้ในการฝึกสอนมีผลต่อความถูกต้องในการรู้จำภาพใบหน้า โดยความถูกต้องในการรู้จำภาพใบหน้าสูงสุดอยู่ที่ใช้จำนวนภาพในการฝึกสอน 10 ภาพ ได้ความถูกต้องในการรู้จำภาพใบหน้า 98.98% ที่ขนาดพีเจอร์ 13x3

ในการเปรียบเทียบความถูกต้องในการรู้จำทำได้ยากเนื่องจากความถูกต้องในการรู้จำภาพใบหน้า เป็นกราฟ 3 มิติ ดังนั้นจึงได้นำแนวทแยงมุมมาเปรียบเทียบ ผลการเปรียบเทียบดังแสดงในรูปที่ 5.16



รูปที่ 5.16 เปรียบเทียบความถูกต้องในการรู้จำใบหน้า Yale ของวิธี WPS-CSLDA ภายใต้เงื่อนไขของตัวอย่างภาพที่ใช้ในการฝึกสอน

บทที่ 6

สรุปผลการวิจัยและแนวทางในการพัฒนา

6.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้เป็นการศึกษาการรู้จำใบหน้าโดยใช้วิธี Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) ซึ่งเป็นวิธีการรู้จำใบหน้าบนพื้นฐานของการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ด้วยวิธีการทางสถิติ

วิธีการรู้จำภาพใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA นี้สามารถหาพีเจอร์ของภาพได้ทั้งแถว (row) และหลัก (column) ของภาพเช่นเดียวกับอัลกอริทึม CSLDA แต่อัลกอริทึมที่น่าเสนอนี้ยังได้นำ Weighted Pairwise Scatter ที่ใช้สำหรับแก้ปัญหาการปนกันของกลุ่มข้อมูลจากการรู้จำเสียงมาประยุกต์ใช้ในการคำนวณการรู้จำภาพใบหน้า

การคำนวณความถูกต้องในการรู้จำภาพหน้านั้นเราคำนวณ Weighted Pairwise Scatter เพื่อคัดแยกข้อมูลที่มีการปนกันระหว่างกลุ่มซึ่งมีผลต่อความถูกต้องในการรู้จำภาพใบหน้า ให้มีการแยกออกจากกัน จากผลการทดลองภายใต้สภาวะและเงื่อนไขภาพใบหน้าต่างๆ สามารถสรุปผลการทดลองได้ ดังนี้

ฐานข้อมูลใบหน้า ORL

การเปรียบเทียบความถูกต้องในการรู้จำภาพใบหน้าสูงสุดของวิธี WPS-CSLDA เมื่อเปรียบเทียบกับวิธี CSLDA, วิธี Eigenimages, วิธี 2DLDA, วิธี 2DPCA, วิธี LDA และวิธี PCA แล้วพบว่าความถูกต้องในการรู้จำภาพใบหน้าวิธี WPS-CSLDA มีผลการรู้จำภาพใบหน้าดีที่สุดโดยมีผลการรู้จำภาพใบหน้า 98.50%

ฐานข้อมูลใบหน้า AR

การเปรียบเทียบความถูกต้องของภาพใบหน้าในฐานข้อมูลใบหน้า AR นี้เป็นการเปรียบเทียบความถูกต้องในการรู้จำ วิธี WPS-CSLDA กับ วิธี CSLDA และวิธี Eigenimages โดยการเปรียบเทียบสามารถแบ่งได้เป็นการทดลองย่อย 3 การทดลองตามเงื่อนไขบนภาพใบหน้าที่นี้ได้แก่ 1. การถ่ายภาพต่างเวลา 2. การแสดงออกทางอารมณ์บนใบหน้า และ 3. ตัวแปรทางแสงบนภาพใบหน้า จากการทดลองหาความถูกต้องในการรู้จำภาพใบหน้าทั้ง 3 เงื่อนไขนี้ ปรากฏว่าผลการรู้จำภาพใบหน้าวิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าดีที่สุดทั้ง 3 เงื่อนไข โดยที่เงื่อนไขการถ่ายภาพต่างเวลาวิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าสูงสุดที่ 93.75% เงื่อนไขการแสดงออกทางอารมณ์บนใบหน้าวิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าสูงสุดที่ 98.90% และเงื่อนไขตัวแปรทางแสงบนภาพใบหน้าวิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าสูงสุดที่

94.65% และ เมื่อนำเวลาที่ใช้คำนวณมาเปรียบเทียบกันพบว่าเวลาที่ใช้ในการคำนวณเทียบไม่แตกต่างกันเลยแต่วิธี WPS-CSLDA ให้ผลการรู้จำภาพใบหน้าดีที่สุด

ฐานข้อมูลใบหน้า Yale

การเปรียบเทียบความถูกต้องของภาพใบหน้าในฐานข้อมูลใบหน้า Yale นี้ เป็นการเปรียบเทียบความถูกต้องในการรู้จำ วิธี WPS-CSLDA กับวิธี CSLDA และวิธี Eigenimages ภายใต้เงื่อนไขการเปลี่ยนแปลงของการแสดงออกบนใบหน้าและความสว่างใบหน้า ผลการทดลองแสดงให้เห็นว่าความถูกต้องในการรู้จำใบหน้าสูงสุดของวิธี WPS-CSLDA มีความถูกต้องในการรู้จำสูงสุด คือ 95.70%

จากการทดลองทั้ง 3 ฐานข้อมูลใบหน้าภายใต้เงื่อนไขต่างๆ ที่ทำให้การรู้จำใบหน้าเกิดความผิดพลาดได้แก่ การเปลี่ยนตำแหน่งและจำนวนตัวอย่างในการเรียนรู้ การแสดงออกบนใบหน้าและความสว่าง พบว่าความถูกต้องในการรู้จำภาพใบหน้าด้วยวิธี WPS-CSLDA ให้ผลความถูกต้องในการรู้จำดีที่สุด

6.2 แนวทางในการพัฒนา

งานวิจัยนี้ได้ทำการทดลองจากฐานข้อมูลใบหน้า 3 ฐานข้อมูล คือ ฐานข้อมูล ORL , ฐานข้อมูล AR, และฐานข้อมูล Yale แต่ในความเป็นจริงยังมีฐานข้อมูลใบหน้าอีกหลายฐานข้อมูลที่มีเงื่อนไขของฐานข้อมูลแตกต่างจากนี้ ดังนั้นจะต้องทำการพัฒนาโปรแกรมนี้ให้ครอบคลุมกับเงื่อนไขที่หลากหลายมากขึ้น

เอกสารอ้างอิง

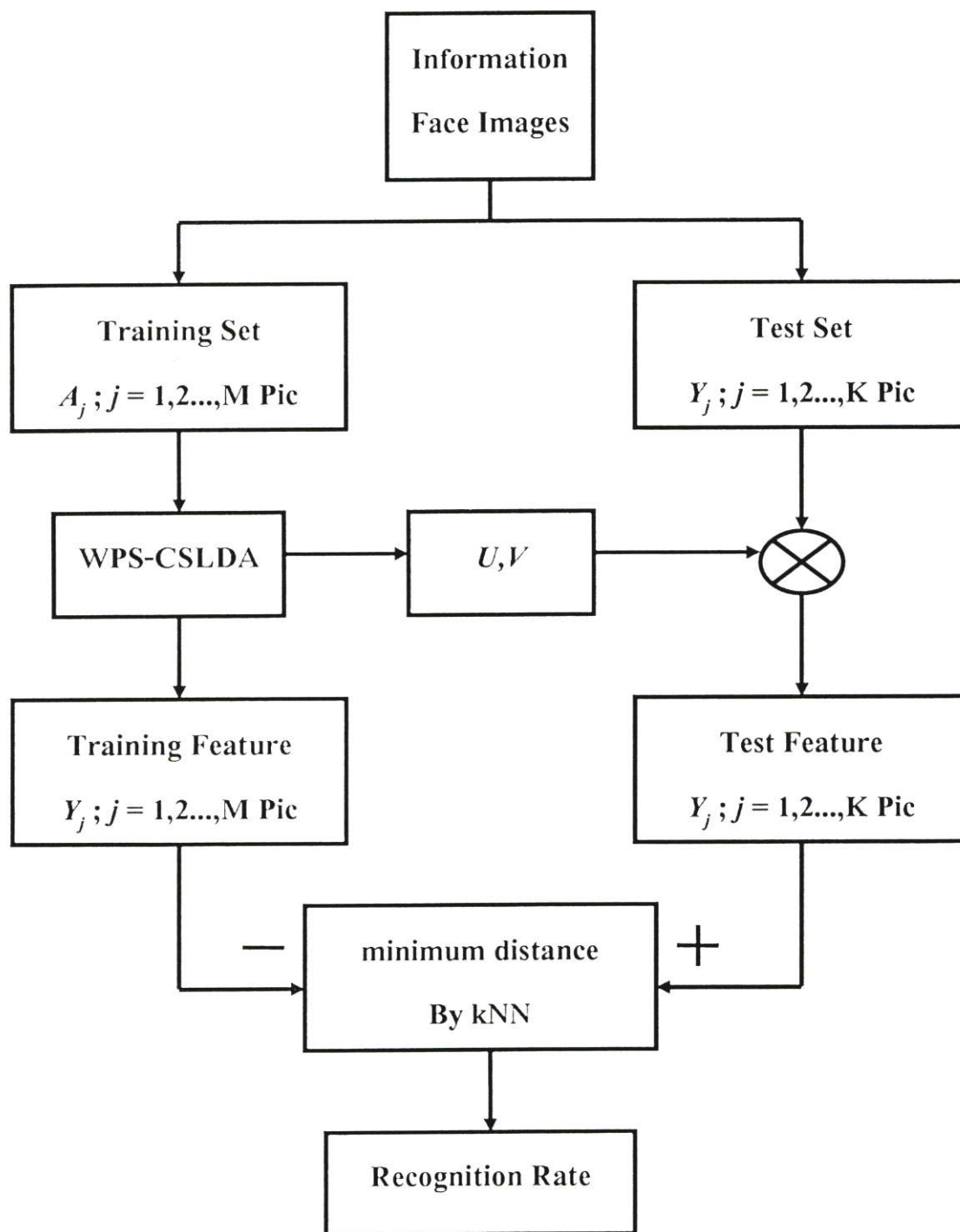
- [1] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, 1991.
- [2] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, Jul. 1997.
- [3] J. Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 1, Jan. 2004.
- [4] S. Kongsontana and Y. Rangsanteri, "Face Recognition Using 2DLDA Algorithm," Eighth International Symposium on Signal Processing and Its Applications, ISSPA 2005, August 2005.
- [5] D. Zhang, S. Chen, J. Liu, "Representing Image Matrices: Eigenimages vs. Eigenvectors," In International Symposium on Neural Networks (ISNN 2005), pp. 659-664, 2005.
- [6] สิทธิชนันท์ คงสนทนา, ยุทธพงษ์ รังสรรค์เสรี และวีรรัตน์ สีลาธรรม "การรู้จำใบหน้าโดยใช้หลักการของ CSLDA," เอกสารการประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 28 Vol.2 ตุลาคม 2548.
- [7] Y. Li, Y. Gao, H. Erdogan "Weighted Pairwise Scatter To Improve Linear Discriminant Analysis" in ICSP-2000. vol. 4, pp.608-611, 2000.
- [8] T. Kanade, "Picture Processing by Computer Complex and Recognition of Human Faces," Ph.D. dissertation, Kyoto Univ., Japan, 1973.
- [9] R. Brunelli and T. Poggio, "Face Recognition : Features versus Templates," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, Oct. 1993.
- [10] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based Face Recognition using mixture-distance," in Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 1996.
- [11] M. Lades, J. C. Vorbruggen, J. Buhann, J. Lange, C. Von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," IEEE Transactions on Computers, Vol. 42, March 1993.
- [12] L. Wiskott, J. M. Fellous, N. Kruger, and C. Von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," IEEE, 1997.

- [13] L. Sirovich and M. Kirby, "Low-Dimensional Procedure for Characterization of Human Faces," *J. Optical Soc. Am.*, vol. 4, 1987.
- [14] A. Pentland, T. Straner, N. Etcoff, A. Masoiu, O. Oliyide, and M. Turk, "Experiments with Eigenface," in *Proc. Looking at people Wkshp, Int. Joint Conf. Artificial Intell.*, Chamberry, France, 1993.
- [15] A. Pentland, B. Moghaddam, and T. Staner, "View-based and Modular Eigenspaces for Face recognition," in *Proc. IEEE conf. Computer Vision and Pattern Recognition 1994*.
- [16] S. J. Lee, S. B. Yung, J. W. Kwon, and S. H. Hong, "Face Detection and Recognition Using PCA," *IEEE TENCON*, 1999.
- [17] S. Z. Lee and J. Lu, "Generalizing Capacity of Face Database for Face Recognition," *IEEE*, 1999.
- [18] J. L. Crowley, and K. Schwerdt, "Robust Tracking and Compression for Video Communication," *IEEE*, 1999.
- [19] R. A. Fisher, "The use of multiple Measures in Taxonomic Problems, *Ann. Eugenics*, 1936.
- [20] K. Etemad and R. Chellappa, "Face Recognition Using Discriminant Eigenvectors," *IEEE*, 1996.
- [21] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, and J. Weng, "Discriminant Analysis of Principal Component for Face Recognition," *International Conference on Automatic Face and Gesture Recognition*, 1998.
- [22] W. Zhao, R. Chellappa, and N. Nandhakumar, "Empirical Performance Analysis of Linear Discriminant Classifiers," *IEEE*, 1998.
- [23] W. Zhao, "Subspace Methods in Object/Face Recognition," *IEEE*, 1999.
- [24] P. Temdec, D. Khawparisuth, and k. Chamnoghthai, "Face Recognition by Using Fractal Encoding and Backpropagation Neural Network," *Fifth International Symposium on Signal Processing and its Applications, ISSPA '99, Brisbane, Australia, August 1999*.
- [25] M. J. Er, S. Wu, and J. Lu, "Face Recognition Using Radial Basis Function (RBF) Neural Networks," *Proceedings of the 38th Conference on Decision & Control*, Phoenix, Arizona USA, *IEEE*, 1999.
- [26] D. Demers and G. W. Cottrell, "Nonlinear Dimensionality Reduction," in *Advances in Neural Information Processing System 5*, S. J. Hanson, J. D. Cowan, C. Lee Giles, Eds. San Mateo, Ca: Morgan Kaufmann, 1993.
- [27] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolution Neural-network Approach," *IEEE, Trans. Neural Networks*, vol.8, no. 1, January 1997.

- [28] S. A. Rizvi, P. J. Phillips, and H. Moon, "A verification protocol and statistical performance analysis for face recognition," IEEE Proc. Conf. Computer Vision and Pattern recognition (CVPR), Santa Barbara, June 1998.
- [29] Y. Yoshitomi, T. Miyaura, S. Tomito, and S. Kimura, "Face Identification Using Thermal Image Processing," IEEE International Workshop on Robot and Human Communication, IEEE, 1997.
- [30] Hotelling. A Primer of Multivariate Statistics. Lawrence Erlbaum Associates, 2001
- [31] <http://www.cam-orl.co.uk>
- [32] A.M. Martinez and R. Benavente, "The AR Face Database," CVC Technical Report, no. 24, June 1998.
- [33] A.M. Martinez and R. Benavente, "The AR Face Database," http://rvll.ecn.purdue.edu/~aleix/aleix_face_DB.html, 2003.
- [34] <ftp://plucky.cs.yale.edu/CVC/pub/images/yaleface>

ภาคผนวก

ภาคผนวก ก
รูปแบบการทดลอง



ก. รูปแบบและขั้นตอนการคำนวณที่ใช้ในการทดลอง

ภาคผนวก ข

ผลงานที่ได้รับการตีพิมพ์

- [1] สิทธิชนันท์ คงสนทนา, ยุทธพงษ์ รังสรรค์เสรี และ วีรรัตน์ ธีลาธรรม “การรู้จำใบหน้าโดยใช้หลักการของ CSLDA,” การประชุมทางวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 28, ตุลาคม 2548.
- [2] วีรรัตน์ ธีลาธรรม, ยุทธพงษ์ รังสรรค์เสรี, “การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA,” วิศวกรรมลาดกระบัง, ปีที่ 24, ฉบับที่ 2, มิถุนายน 2550.

การรู้จำใบหน้าโดยใช้หลักการของ CSLDA Face Recognition Using CSLDA Algorithm

สิงธินันท์ กอนสงทนา พุทธพรม ธีรสารศรี วิจิรัตน์ อิศรธรรม

ภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ถนนจตุรุมจรูญ เขตคลองจั่น กรุงเทพมหานคร 10520

โทร. 0-2461-0894 E-mail: s.kongsontana@yahoo.com, tangsawseri@yahoo.com, beung278@yahoo.com

บทคัดย่อ

บทความนี้ได้เสนออัลกอริทึมใหม่ของ LDA มีชื่อ Coupled Subspace Linear Discriminant Analysis (CSLDA) โดยที่ CSLDA สามารถแก้ปัญหาการเป็นเมตริกเอกฐานได้ ความแตกต่างระหว่าง CSLDA และ LDA คือ ลักษณะในการคำนวณ ซึ่ง LDA จะใช้เวกเตอร์ในการคำนวณขณะที่ CSLDA จะใช้เมตริกในการคำนวณ บทความนี้ได้ทำการศึกษารวมกันของวิธี CSLDA กับวิธี LDA ซึ่งมีชื่อว่า CSLDA+LDA อัลกอริทึมที่พัฒนาแล้วได้นำมาประยุกต์กับงานด้านการรู้จำใบหน้า จากการศึกษาแสดงให้เห็นว่า CSLDA และ CSLDA+LDA มีประสิทธิภาพในการคำนวณและการรู้จำที่ดี

คำสำคัญ: การวิเคราะห์การแบ่งแยกเชิงเส้น การลดขนาดมิติของข้อมูล การรู้จำใบหน้า

Abstract

In this paper, we are propose a novel LDA algorithm, namely CSLDA, which stands for Coupled Subspaces Linear Discriminant Analysis. CSLDA overcomes the singularity problem implicitly. The key difference between CSLDA and classical LDA lies in the model for data representation. Classical LDA works with vectorized representations of data, while the CSLDA algorithm works with data in matrix representation. To further reduce the dimension by CSLDA, the combination of CSLDA and classical LDA namely CSLDA+LDA, is studied, where LDA is preceded by CSLDA. The proposed algorithms are applied on face recognition. A. Experiments show that CSLDA and CSLDA+LDA achieve computing efficiency and competitive recognition accuracy

Keywords: Linear Discriminant Analysis, Dimensionality Reduction, Face Recognition.

1. บทนำ

วิธีการแบบสถิติเชิงเส้นที่ได้รับความนิยมมาใช้ในการระบุการระบุเฟส (feature extraction) ของภาพ และมีหลายวิธีที่มีประสิทธิภาพสูงได้ถูกการวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ซึ่งเป็นวิธีที่สำคัญมากในการพัฒนาแบบสถิติ แต่อย่างไรก็ตามวิธี PCA และวิธี LDA จะใช้การคำนวณเฉพาะเวกเตอร์ 1 มิติ ดังนั้นเมื่อต้องการใช้วิธีที่มีลักษณะที่จำเป็นจะต้องเปลี่ยนเมตริกภาพให้เป็นเวกเตอร์ภาพหลังจากนั้นจึงใช้พื้นฐานของเวกเตอร์ในการคำนวณหาเมตริกความแปรปรวนร่วม (covariance matrix) เพื่อคำนวณหาเวกเตอร์ที่เหมาะสมที่สุดที่ใช้ในการระบุภาพ วิธีที่เรานี้ถูกเรียกว่า Eigenface [1] และ Fisherface [2] เมื่อ นำวิธีการดังกล่าวมาใช้ในงานด้านการจดจำใบหน้าเป็นผลให้ได้กระบวนการจดจำที่ดี แต่ว่าภาพใบหน้าที่ใช้ในกระบวนการจดจำจะมีมิติมากเช่น ภาพใบหน้าที่มีขนาดความละเอียด 112 x 92 จุดภาพ เมื่อแปลงเป็นเวกเตอร์จะมีความยาว 10,304 มิติ ซึ่งเหตุผลดังกล่าวจึงทำให้การคำนวณด้วยวิธี Eigenface และ วิธี Fisherface นำได้ยากและใช้เวลานาน

Yang [3] ได้เสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้เทคนิคเป็น 2 มิติบนตำแหน่งพิกเซล วิธีการนี้มีชื่อว่า การวิเคราะห์องค์ประกอบหลักของภาพ (Image Principal Component Analysis: IMPCA) โดยที่วิธีการนี้สามารถลดมิติได้เฉพาะหลัก (column) ของภาพ แต่ไม่สามารถลดมิติแถว (row) ของภาพได้

Zhang [4] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลัก โดยใช้การลดขนาดของเมตริก วิธีการนี้มีชื่อว่า Eigenimages แต่วิธีการนี้ไม่ได้นำข้อมูลจากกลุ่มมาพิจารณา

บทความนี้เสนอวิธี Coupled Subspace Linear Discriminant Analysis (CSLDA) สำหรับการระบุเฟสของภาพ วิธีการนี้มีขั้นตอนที่แยกกันเป็น 2 ขั้นตอนวิธีแรกคือ LDA ที่นำข้อมูลจากกลุ่มมาพิจารณา ซึ่งความแตกต่างระหว่าง CSLDA และ LDA คือ ลักษณะในการคำนวณ โดยที่ LDA จะใช้เวกเตอร์ในการคำนวณ ขณะที่ CSLDA ใช้เมตริกในการคำนวณ เมื่อคำนวณด้วยวิธี CSLDA จะสามารถแก้ปัญหาการเป็นเมตริกเอกฐานและสามารถลดมิติของภาพได้ทั้งแถวและหลัก บทความนี้ยังได้ศึกษารวมกันของวิธี CSLDA กับวิธี LDA ซึ่งมีชื่อว่า CSLDA+LDA เพื่อทำการเปรียบเทียบกันกับวิธี PCA+LDA [5] และ Fisherface

DS020

2. Coupled Subspace Linear Discriminant Analysis

2.1 แนวความคิด และ หลักการคำนวณ

กำหนดให้แก่อนุกรการฉาย $U \in R^{m \times m}$ และ $V \in R^{m \times c}$ ภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนแกน U และ V ตามหลักการแปลงเชิงเส้น

$$Y = U^T A V \tag{1}$$

ดังนั้นจะได้ Y คือพีเพอร์เมตริกของภาพ A ที่มีขนาด $r \times c$ โดยที่ $r < m$ และ $c < n$

เมื่อกำหนดให้มีตัวอย่างภาพทั้งหมด M ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวนภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, $A_j (j = 1, 2, \dots, M)$ ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ $\bar{A}_i (i = 1, 2, \dots, L)$ คือค่าเฉลี่ยของภาพแต่ละกลุ่ม T_i และ P_i คือกลุ่มที่ถูกฉาย

หลังจากภาพที่ใช้สอนถูกฉายลงบนแกน U และ V จะใช้พีเพอร์เมตริกเป็น

$$Y_j = U^T A_j V, \quad j = 1, 2, \dots, M \tag{2}$$

วิธีในการหาแก่อนุกรการฉาย U และ V สามารถคำนวณได้โดยหาผลรวมค่าไอเกน (trace) ของเมตริกความแปรปรวนร่วมจากภาพตัวอย่างที่ถูกฉาย [1] จากแนวความคิดสามารถสร้างเป็นสมการได้คือ

$$J(U, V) = \frac{P_B}{P_W} \tag{3}$$

เมื่อ

$$P_B = \mu(\tilde{S}_B)$$

$$P_W = \mu(\tilde{S}_W)$$

โดยที่

\tilde{S}_B คือเมตริกการกระจายระหว่างกลุ่มของภาพตัวอย่างที่ถูกฉาย

\tilde{S}_W คือเมตริกการกระจายภายในกลุ่มของภาพตัวอย่างที่ถูกฉาย

$$\begin{aligned} \tilde{S}_B &= \sum_{i=1}^L N_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T \\ &= \sum_{i=1}^L N_i [U(\bar{A}_i - \bar{A})V][U(\bar{A}_i - \bar{A})V]^T \end{aligned} \tag{4}$$

$$\begin{aligned} \tilde{S}_W &= \sum_{i=1}^L \sum_{Y_k \in P_i} (Y_k - \bar{Y}_i)(Y_k - \bar{Y}_i)^T \\ &= \sum_{i=1}^L \sum_{k \in P_i} [U(A_k - \bar{A}_i)V][U(A_k - \bar{A}_i)V]^T \end{aligned} \tag{5}$$

ถ้า

$$\mu(\tilde{S}_B) = \mu\left(\sum_{i=1}^L N_i U^T (\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T U\right) \tag{6}$$

$$\mu(\tilde{S}_W) = \mu\left(\sum_{i=1}^L \sum_{k \in P_i} U^T (A_k - \bar{A}_i)(A_k - \bar{A}_i)^T U\right) \tag{7}$$

ในการคำนวณหา U และ V ให้สามารถนำไปพร้อมกัน ดังนั้นจึงทำการคำนวณเพื่อหาค่า U ก่อนโดยนำ $V = I$ (คือ เมตริกเอกลักษณ์เพื่อหาค่าของ V แล้วค่อยหาค่าของ $U = I$ คำนวณค่าของ V)

การคำนวณ U

กำหนดให้ $V = I$

$$\mu(\tilde{S}_B) = U^T S_B^U U, \quad \mu(\tilde{S}_W) = U^T S_W^U U \tag{8}$$

โดยที่

$$S_B^U = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T$$

$$S_W^U = \sum_{i=1}^L \sum_{k \in P_i} (A_k - \bar{A}_i)(A_k - \bar{A}_i)^T$$

จากสมการนี้เราสามารถหาค่า U ได้คือ

$$J(U) = \frac{U^T S_B^U U}{U^T S_W^U U} \tag{9}$$

ในการหาค่าของ U ที่เหมาะสมที่สุดในการฉายสามารถทำได้โดยการหาค่าของ $J(U)$

$$U_{opt} = \arg \max_U J(U) \tag{10}$$

ถ้า S_W^U ไม่เป็นเมตริกเอกฐาน (non singular) สามารถหาเมตริก U_{opt} โดยการคำนวณค่าไอเกนเบตอร์และค่าไอเกนที่ได้จากสมการ

$$S_B^U U_{opt} = \lambda S_W^U U_{opt} \tag{11}$$

โดยที่ไอเกนการกระจายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกน

เบตอร์ของ $S_B^U S_W^U^{-1}$ S_B^U ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดแล้ว 1 ถึง m แต่ในบางกรณีอาจมีกรณีที่การหาพีเพอร์เมตริกของ U นั้นจึงเลือกแกนในการฉายเป็น r แก่ r เท่านั้น ดังนั้นเมตริก U ที่มี r แก่ r ในการฉายจึงมีขนาดเป็น $m \times r$

$$U = [u_1, \dots, u_r] \tag{12}$$

การคำนวณ V

กำหนดให้ $V = I$

$$\mu(\tilde{S}_B) = V^T S_B^V V, \quad \mu(\tilde{S}_W) = V^T S_W^V V \tag{13}$$

โดยที่

$$S_B^F = \sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A})$$

$$S_W^F = \sum_{i=1}^L \sum_{k \in P_i} (A_k - \bar{A}_i)^T (A_k - \bar{A}_i)$$

ความคลุมสมมาตรของเทรซ $tr(AB) = tr(BA)$ เมื่อ A และ B เป็นเมทริกซ์
จากสมการที่ (3) สามารถแทนค่าใหม่ได้คือ

$$J(V) = \frac{V^T S_B^F V}{V^T S_W^F V} \tag{14}$$

ในการแก้สมการหา V ที่เหมาะสมที่สุดในกรณีที่สามารถนำได้โดย
การหาค่ามากที่สุดของ $J(V)$

$$V_{opt} = \arg \max_V J(V) \tag{15}$$

ถ้า S_W^F ไม่เป็นเมทริกซ์เอกฐาน (non singular) สามารถหาเมทริกซ์ V_{opt}
โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนที่ได้จากสมการ

$$S_B^F V_{opt} - \lambda S_W^F V_{opt} \tag{16}$$

โดยปกติเมทริกซ์ที่เหมาะสมที่สุดที่ใช้ในการจัดเรียงจากไอเกน
เวกเตอร์ของ $S_W^{F^{-1}} S_B^F$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง
 m แต่ในความเป็นจริงเนื่องจากต้องการหาฟิเจอร์จากภาพจะนับจึงเลือก
แทนในการฉายเพียง 1 ถึง c เท่านั้น ดังนั้นเมทริกซ์ใหม่ที่ใช้ในการจัดเรียง
ขนาดเป็น $n \times c$

$$V = [v_1 v_2 \dots v_c] \tag{17}$$

2.2 การสร้างภาพใหม่

ในวิธีของ CSLDA สามารถนำเอาฟิเจอร์เมทริกซ์และเมทริกซ์การจัดเรียง
ที่เหมาะสมที่สุด U และ V มาสร้างเป็นภาพใหม่ได้ด้วยขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ Y เป็นฟิเจอร์เมทริกซ์ของภาพ A โดยที่
 $U = [u_1, \dots, u_c]$ และ $V = [v_1, \dots, v_c]$ เป็นเมทริกซ์การจัดเรียงที่เหมาะสมที่สุด
ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = UV^T \tag{18}$$

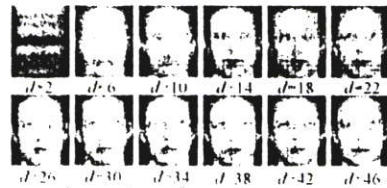
3. การทดลองและการวิเคราะห์

ฐานข้อมูลภาพใบหน้าของ ORI (<http://www.cam.ac.uk/ori>) มี
ทั้งหมดเป็นภาพระดับสีเทาและมีขนาด 112×92 จุดภาพ ถ่ายจากใบหน้า
ของ 40 บุคคล คนละ 10 ภาพที่แตกต่างกัน เช่น ภาพที่มีการแสดงออก
ทางใบหน้า (การปิดตาหรือเปิดตา คิ้วหรือไม่มี) ภาพที่มีรายละเอียด (การใส่แว่นหรือถอดแว่น) ภาพที่มีการหันหน้าหรือเบี่ยงหน้า หันซ้ายหรือ
ขวา ไม่เกินร้อยละ 20 โดยรูปที่ 1 เป็นตัวอย่างภาพใบหน้าทั้ง 10 ภาพของ
บุคคลเดียว ในการทดลองทั้งหมดจะใช้ 1-Nearest Neighbor (INN)
เป็นวิธีแบ่งแบบ (Classifier)



รูปที่ 1 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล ORI

การทดลองแรกเป็นการทดสอบคุณสมบัติการสร้างภาพใหม่ด้วยวิธี
CSLDA ในการทดลอง ภาพที่ใช้ฝึกสอนถูกเลือกมาอย่างสุ่มจำนวน 5
ภาพต่อหนึ่งบุคคล เมื่อนำมาแทนที่บนกระดาษที่เฉพาะสมที่ด้วยวิธี
CSLDA เมื่อนำเฉพาะรูปที่ 1 มาฉายลงบนกระดาษที่เฉพาะสม
ที่ดีที่สุดทั้งสควมจะ ได้ฟิเจอร์เมทริกซ์ที่ใช้ในการสร้างภาพใหม่ จากรูปที่ 2
เมื่อจำนวน d (d คือ $r - c$) เพิ่มขึ้นภาพที่สร้างใหม่จะมีความเหมือนภาพ
ต้นแบบมากขึ้น



รูปที่ 2 ภาพที่สร้างขึ้นใหม่ของบุคคลคนหนึ่ง

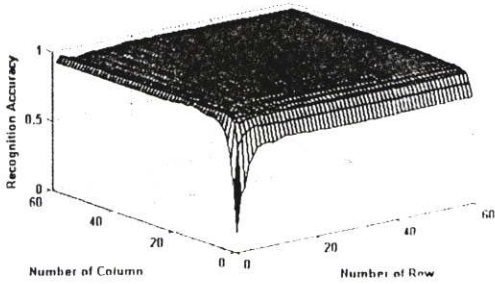
การทดลองที่สองได้เป็นการเปรียบเทียบความถูกต้องระหว่าง วิธี
CSLDA กับ วิธี Eigenimage ภายใต้เงื่อนไขของจำนวนตัวอย่างภาพใบ
หน้าที่ใช้สอน โดยแบ่งการทดลองออกเป็น 3 ครั้ง ครั้งแรกภาพใบหน้าที่ใช้
ใช้ฝึกสอนจะถูกแยกออกมาอย่างสุ่ม 3 ภาพต่อหนึ่งบุคคล และภาพที่
เหลือถูกนำมาใช้ทดสอบความถูกต้อง การทดลองครั้งถัด ๆ มา ภาพที่ใช้
ฝึกสอนจะถูกเลือกแบบสุ่ม 4, 5 ภาพต่อหนึ่งบุคคลตามลำดับ

ตารางที่ 1 เปรียบเทียบความถูกต้องในการรู้จำสุ่มระหว่าง Eigenimage
และ CSLDA ที่สัมพันธ์กับจำนวนตัวอย่างที่ใช้สอน

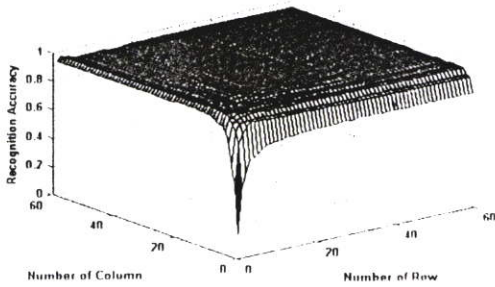
Training sample per class	Eigenimages	CSLDA
3	0.9333(13x3)	0.9571(22x3)
4	0.9417(21x2)	0.9625(10x2)
5	0.9550(18x2)	0.9750(13x2)

จกตารางที่ 1 เป็นได้ว่าวิธี CSLDA มีความถูกต้องในการรู้จำสูง
มากกว่าวิธี Eigenimages เมื่อเทียบกับ CSLDA ได้ประสิทธิภาพดีกว่าวิธี LDA
ที่นำข้อมูลจากกลุ่มมาพิจารณาด้วย รูปที่ 3 และ 4 เป็นตัวอย่างของการหา
ความถูกต้องในการรู้จำของวิธี Eigenimage และ CSLDA

DS020



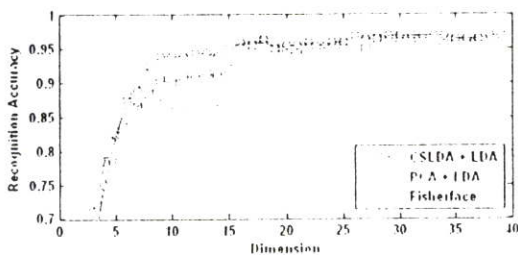
รูปที่ 3 ความถูกต้องในการรู้จำของ Eigenimages



รูปที่ 4 ความถูกต้องในการรู้จำของ CSLDA

การทดลองที่สามเปรียบเทียบกระบวนการของ CSLDA+LDA, PCA+LDA, Fisherface บนฐานข้อมูล ORL ภาพที่ใช้สอนถูกเลือกมาอย่างสุ่มจำนวน 5 ภาพต่อหนึ่งบุคคลและภาพที่เหลือถูกนำมาใช้สำหรับทดสอบ ดังนั้นภาพที่ใช้สอนมีจำนวนทั้งหมด 200 ภาพ และภาพที่ใช้ในการทดสอบ 200 ภาพ

ในการเปรียบเทียบ วิธีของ Fisherface กำหนดให้ PCA มีขนาดเมตริกเป็น $M \times L$ (160) มิติและกำหนดให้ LDA มีขนาดมิติไม่เกิน $L - 1$ (39) มิติ ส่วนวิธีของ PCA+LDA กำหนดให้ PCA มีมิติไม่เกิน $M - L$ และ LDA มีขนาดมิติไม่เกิน $L - 1$ วิธี CSLDA+LDA กำหนดให้ CSLDA มีขนาด 8×8 และใช้ LDA ลดมิติไม่เกิน $L - 1$



รูปที่ 5 เปรียบเทียบผลการทดลองจัดระบบที่ CSLDA+LDA, PCA+LDA และ Fisherface

จากรูปที่ 5 แสดงให้เห็นว่า CSLDA+LDA มีผลการรู้จำที่ดีกว่า PCA+LDA และ Fisherface

ตารางที่ 2 เปรียบเทียบเวลา (วินาที) ที่ใช้คำนวณหาไอเจนรัวภาพทั้ง 3 วิธี

CSLDA+LDA	PCA+LDA	Fisherface
7.1676	36.5478	40.9780

ตารางที่ 2 แสดงเวลาที่ใช้ในการสอนทั้งสามวิธี (CPU: Pentium IV 1.7 GHz, RAM: 256M) ทั้งสามวิธีถูกคำนวณด้วยโปรแกรม Matlab 6.5 จากตารางที่ 1 จะเห็นว่าวิธี CSLDA+LDA ใช้เวลาในการคำนวณน้อยกว่าวิธี PCA+LDA และ วิธี Fisherface มากเพราะว่า วิธี CSLDA+LDA จะจัดลดรูปเมตริกที่มีขนาดเมตริกเป็น 92×92 และ 112×112 แต่วิธี PCA+LDA และ Fisherface ต้องจัดลดรูปเมตริกที่มีขนาด 10304×10304 ดังนั้นการคำนวณจึงทำได้รวดเร็วกว่าวิธีอื่นๆ

4. สรุปผลการทดลอง

บทความนี้เสนอวิธีการใหม่สำหรับหารหาไอเจนรัวภาพที่มีชื่อว่า CSLDA โดยที่ความแตกต่างระหว่าง CSLDA และ LDA คือ CSLDA จะใช้การแทนที่เป็นเมตริกตามค่าขนาดโดยวิธีของ LDA ต้องเปลี่ยนเมตริกภาพเป็นเวกเตอร์ภาพก่อนการคำนวณ ในการคำนวณด้วยวิธี CSLDA สามารถนำไปแทนที่เป็นการเป็นเมตริกของของวิธี LDA บทความนี้ยังได้ทำการเปรียบเทียบความเร็วกันของ LDA และ CSLDA ซึ่งมีชื่อว่า CSLDA+LDA จากผลการทดลองพบว่ากระบวนการของ CSLDA+LDA ให้ผลการรู้จำสูงกว่าวิธีอื่นๆของ Eigenimages และวิธี CSLDA+LDA ก็พบว่าให้ผลการรู้จำดีกว่า PCA+LDA, Fisherface และใช้เวลาในการคำนวณน้อยกว่าวิธีอื่นๆ

5. เอกสารอ้างอิง

- [1] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991
- [2] P. N. Bellamune, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997
- [3] J. Yang, "Two Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 135-137, 2004
- [4] D. Zhang, S. Chen, J. Liu, "Representing Image Matrices: Eigenimages vs. Eigenvectors," In International Symposium on Neural Networks (ISSN: 2005), pp. 659-664 China, 2005
- [5] J. Ye, R. Janardan, C.H. Park, H. Park, "An optimization criterion for generalized discriminant analysis on undersampled problems," IEEE Transactions on Pattern Analysis and Machine Intelligence, No. 8, pp. 928-994, 2004



ลาดกระบัง

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

LADKRABANG ENGINEERING JOURNAL

ปีที่ 24 ฉบับที่ 2

มิถุนายน 2550

1.	การผลิตกระแสไฟฟ้าและพลังงานความร้อนจากระบบเซลล์เชื้อเพลิงแบบออกไซด์ของแข็งโดยใช้ ก๊าซธรรมชาติ และ ก๊าซสังเคราะห์ที่ได้จากชีวมวลเป็นแหล่งพลังงาน	1
	วาทิต วัฒนุญไพศาล รพีพงศ์ สุวรรณวางกูร นริศรา ทองบุญชู	
2.	การปรับปรุงตัวเร่งปฏิกิริยานิกเกิลบนตัวรองรับอะลูมินาเพื่อใช้ผลิตไฮโดรเจนจากปฏิกิริยารีฟอร์มมิ่ง เอทานอลด้วยไอน้ำ	7
	ดวงกมล ณ ระนอง วิศรุต ใหญ่ยอด พัฒนา เทียมวิลาวณิชย์ อภาภรณ์ จันทร์ปรีกษ์ กันยรัตน์ บัวเกตุ ณัฐนิชา ณัฐนรากร พรสวรรค์ กาญจนวนิชย์กุล	
3.	การเร่งปฏิกิริยาจัดไฮโดรเจนของเอทานอลเพื่อได้อะซีตัลดีไฮด์บนตัวเร่งปฏิกิริยานิกเกิล	13
	อาทิตย์ เนรมิตตพงษ์ วิชาดา อัดไพบูลย์ สุธาสินี เนรมิตตพงษ์	
4.	การสังเคราะห์เอทิลีนจากเอทานอลบนตัวเร่งปฏิกิริยาโลหะฟอสเฟต	19
	อาทิตย์ เนรมิตตพงษ์ วัชรภรณ์ ปราณี สุธาสินี เนรมิตตพงษ์	
5.	การปรับปรุงสมบัติป้องกันไอโซนของยางธรรมชาติโดยใช้ไมโครคริสตัลไลน์แวกซ์	25
	เศรษฐกร พรหมศิริ กาเหว่า เทพนอก สุรัตน์ อารีรัตน์	
6.	แผงเซลล์แสงอาทิตย์ทันเดมอะมอร์ฟัสซิลิคอนประสิทธิภาพสูงที่สร้างโดยกระบวนการเชิงปริมาณ	31
	ศรัณย์ สาธารมณีน วิสุทธิ์ รุติรุ่งเรือง เกรียงศักดิ์ สุวรรณโพธิ์ศรี นิรุช ปันเกตุ	
7.	ลักษณะสมบัติของซิลิคอนโฟโตไดโอดแบบ PIN ที่มีชั้นพอร์ซิลิคอนจากการกัดย้อมสีเป็นชั้นลด การสะท้อนแสง	37
	จิรวัดณ์ ปราบเขต สมเกียรติ ศุภเดช	
8.	การสังเคราะห์อนุภาคนาโนของเหล็กด้วยวิธี CVD จากละอองเฟอร์โรเซน เอทานอล	43
	ธนูทอง คำยา สุทธิชัย ทัยสิทธิ์ศักดิ์	
9.	เครื่องผลิตก๊าซไอโซนแบบใช้แหล่งจ่ายเฟสซิฟท์ที่ดับปลิวเอ็มฟูลบริดจ์คอนเวอร์เตอร์	49
	อนุรักษ มั่นสุข วิจิตร กิณเรศ	
10.	การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA	55
	วีรรัตน์ ลีลาธรรม ยุทธพงษ์ รังสรรค์เสรี	
11.	การจัดแบ่งภาระงานของเวปคลัสเตอร์ด้วยลำดับเสมือน	61
	กริช กองศรีมา มยุรี เลิศเวชกุล	
12.	ปรากฏการณ์การเคลื่อนที่วนของวัสดุเมล็ดเล็กภายใต้การสั่นแนวตั้ง	67
	สกันธ์ คล่องบุญจิต	
13.	Equations for Determining Characteristics of Resonance-Type Silencers	73
	M. Lokitsangtong A. Sindhuphak T. Komori	
14.	Design Techniques for Energy Efficient Multiplier	77
	W. Suntiamorntut C. Vongchumyen	

การรู้จำใบหน้าโดยใช้อัลกอริทึม WPS-CSLDA

Face Recognition Using WPS-CSLDA Algorithm

วีรรัตน์ ตีลาธรรม ยุทธพงษ์ รังสรรค์เสรี

ภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

บทความนี้ได้นำเสนอการพัฒนาอัลกอริทึมใหม่บนพื้นฐานของ LDA มีชื่อว่า Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) โดยที่ WPS-CSLDA สามารถแก้ปัญหาการปนกันของข้อมูลระหว่างกลุ่ม โดยได้นำ Weighted Pairwise Scatter มาคำนวณการแบ่งแยกข้อมูลระหว่างกลุ่มที่มีการปนกันเพื่อให้ได้ลักษณะการแบ่งแยกที่มีประสิทธิภาพสูงกว่าวิธีที่ผ่านมา ความแตกต่างระหว่าง WPS-CSLDA กับ วิธีอื่น คือ WPS-CSLDA จะนำ Weighted Pairwise Scatter มาใช้ในการคำนวณ อัลกอริทึมที่พัฒนาแล้วได้นำมาประยุกต์กับงานด้านการรู้จำใบหน้าจากการทดลองแสดงให้เห็นว่า WPS-CSLDA มีประสิทธิภาพในการคำนวณและการรู้จำที่ดีกว่าวิธีอื่น

Abstract

In this paper, we propose a novel algorithm based on LDA, Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) for solving the confusable data in face recognition application. WPS-CSLDA is used to compute between class scatter discrimination for higher efficiency. The key difference between WPS-CSLDA and classical LDA is using Weighted Pairwise Scatter for the computation. This algorithm is applied to the face recognition works. The experimental results indicated that the recognition rate of our algorithm is better than that of the conventional method.

1. บทนำ

วิธีการทางสถิติเป็นเครื่องมือที่ถูกนำมาใช้สำหรับการหาฟีเจอร์ (feature extraction) ของภาพ และมีหลายวิธีที่มีประสิทธิภาพสูงได้ถูกนำมาใช้ การวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ซึ่งเป็นวิธีที่สำคัญมากในการพัฒนาทางสถิติ แต่่ววิธี PCA และวิธี LDA จะใช้การคำนวณเฉพาะเวกเตอร์ 1 มิติ ดังนั้นเมื่อต้องการใช้วิธีนี้กับภาพจึงจำเป็นต้องเปลี่ยนเมตริกภาพให้เป็นเวกเตอร์ภาพ หลังจากนั้นจึงใช้พื้นฐานของเวกเตอร์ในการคำนวณหาเมตริกความแปรปรวนร่วม (covariance matrix) เพื่อคำนวณหาเวกเตอร์

ที่เหมาะสม ที่สุดที่ใช้ในการฉายภาพ วิธีการนี้ถูกเรียกว่า Eigenface [1] และ Fisherface [2] เมื่อ นำวิธีการดังกล่าวมาใช้ในงานด้านการรู้จำใบหน้าเป็นผลให้ได้กระบวนการรู้จำที่ดี แต่ภาพใบหน้าที่ใช้ในกระบวนการรู้จำจะมีมิติมาก เช่น ภาพใบหน้าที่มีขนาดความละเอียด 112 x 92 จุดภาพ เมื่อเปลี่ยนเป็นเวกเตอร์จะมีขนาด 10,304 มิติ จากเหตุผลดังกล่าวจึงทำให้การคำนวณด้วยวิธี Eigenface และวิธี Fisherface ทำได้ยากและใช้เวลานาน

Yang [3] ได้เสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติมาคำนวณโดยตรง วิธีการนี้มีชื่อว่า การวิเคราะห์องค์ประกอบหลักของภาพ

(Image Principal Component Analysis: IMPCA) โดยที่วิธีการนี้สามารถลดมิติได้เฉพาะหลัก (column) ของภาพ แต่ไม่สามารถลดมิติแถว (row) ของภาพได้

Kongsontana [4] ได้เสนอวิธีการคำนวณหาเมตริกความแปรปรวนร่วมโดยใช้ภาพที่เป็น 2 มิติ มาคำนวณโดยตรง และได้นำข้อมูลของกลุ่มมาพิจารณา มีชื่อว่า 2DLDA แต่วิธีการนี้สามารถลดมิติได้เฉพาะหลัก (column) ของภาพเช่นกัน ไม่สามารถลดมิติได้พร้อมกันทั้งแถวและหลักของภาพ

Zhang [5] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลักโดยใช้แกนการฉายสองแกน วิธีการนี้มีชื่อว่า Eigenimage แต่วิธีการนี้ไม่ได้นำข้อมูลของกลุ่มมาพิจารณา

สิทธิพันธ์ [6] ได้เสนอวิธีการที่สามารถลดมิติของภาพได้ทั้งแถวและหลักโดยใช้แกนการฉายสองแกนและได้นำข้อมูลของกลุ่มมาพิจารณา วิธีการนี้มีชื่อว่า CSLDA แต่ยังคงพบว่ามี การปนกันของข้อมูลที่มีการกระจายระหว่างกลุ่ม ซึ่งมีผลต่อความถูกต้องในการรู้จำใบหน้า

Li [7] ได้นำเสนอวิธีการแก้ปัญหาการปนกันของข้อมูลบนพื้นฐาน LDA โดยนำ Weighted Pairwise Scatter มาใช้แก้ปัญหการปนกันของข้อมูล

บทความนี้เสนอวิธี Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA) สำหรับการหาพีเจอร์ของภาพ วิธีการนี้เป็นการประยุกต์ [6] มาใช้กับ [7] เพื่อทำให้ข้อมูลระหว่างกลุ่มที่มีการปนกัน แยกออกจากกันมากขึ้น ซึ่งทำให้ผลการรู้จำใบหน้าดีขึ้นกว่าเดิม

2. Weighted Pairwise Scatter Coupled Subspace Linear Discriminant Analysis (WPS-CSLDA)

ทฤษฎีการวิเคราะห์การแบ่งแยกเชิงเส้นนั้นจะนำข้อมูลของกลุ่มมาพิจารณา ปัญหาที่พบมากในการพิจารณาข้อมูลกลุ่มนั้น คือ การปนกันของข้อมูลที่อยู่คนละกลุ่ม ในส่วนถัดไปจะเป็นการนำทฤษฎีเข้ามาแก้ปัญหการปนกันของข้อมูลนี้ โดยทฤษฎีนี้มีชื่อ "Weighted Pairwise Scatter"

2.1 แนวความคิด Weighted Pairwise Scatter

ทฤษฎี LDA ทั่วไปเทอมแรก $\tilde{S}_w = w^T S_w w$ จะเรียกว่าเมตริกการกระจายภายในกลุ่ม เทอมที่สอง $\tilde{S}_B = w^T S_B w$ จะเรียกว่าเมตริกการกระจายระหว่างกลุ่ม ถ้า w เป็นการฉายเส้นตรง (linear projection) ซึ่งอยู่ในปริภูมิพีเจอร์ (feature space) ใหม่ การกระจายภายในกลุ่มและระหว่างกลุ่มจะเป็น $w^T S_w w$ และ $w^T S_B w$

แนวคิดของ LDA คือ จะทำให้ข้อคร่าส่วนเมตริกการกระจายระหว่างกลุ่มและเมตริกการกระจายภายในกลุ่ม หลังจากการแปลง (transform) มีค่ามากที่สุด ซึ่งจะทำให้สามารถเลือกการแปลงที่ยังคงข้อมูลการแบ่งแยกมากที่สุด ในขณะที่ขนาดของมิติจะลดลง เราต้องการที่จะทำให้สมการเป้าหมาย (objective function) มีค่ามากที่สุด

$$\max_w \frac{|w^T S_B w|}{|w^T S_w w|}$$

คำตอบในการแก้ปัญหานี้จะใช้ ไอเกนเวกเตอร์ (Eigenvector) ที่สัมพันธ์กับค่าไอเกน (Eigenvalue) ที่เรียงจากมากไปน้อยจากสมการ

$$S_B w = \lambda S_w w$$

รูปแบบของเมตริกการกระจายระหว่างกลุ่มสามารถหาได้จาก

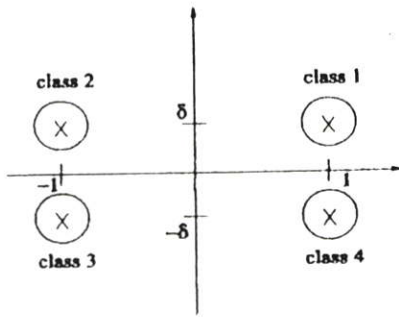
$$S_B = \sum_{i=1}^L N_i (\bar{i}_i - \bar{i})(\bar{i}_i - \bar{i})^T$$

โดยที่ S_B คือ เมตริกการกระจายระหว่างกลุ่ม

\bar{i} คือ ค่าเฉลี่ยของตัวอย่างทั้งหมด

\bar{i}_i คือ ค่าเฉลี่ยตัวอย่างของแต่ละกลุ่ม

สมการนี้เป็นการคำนวณว่าจะแบ่งค่าเฉลี่ยข้อมูลของแต่ละกลุ่มอย่างไรจากจุดศูนย์กลางข้อมูล มันเป็นการตีที่จะได้ค่า S_B มากๆ เพราะจะแสดงว่ากลุ่มต่างๆ จะมีการแบ่งแยกออกจากกันมากในปริภูมิการแปลง (transformed space) แต่จากสมการไม่สามารถบอกได้ชัดเจนว่ากลุ่มต่างๆ จะถูกแยกออกอย่างไร ตัวอย่างต่อไปนี้จะทำให้เข้าใจจุดที่พิจารณา



รูปที่ 1 ข้อมูลที่มีความแปรปรวนเท่ากันที่ตำแหน่งต่างกัน

ตัวอย่าง กำหนดให้มีข้อมูล 4 กลุ่มแต่ละกลุ่มมีจำนวนของพีเจอร์เวคเตอร์เหมือนกันและมีความแปรปรวนเท่ากัน ดังรูปที่ 1 คือ

$$\begin{aligned} \bar{t}_1 &= (1, \delta) & \bar{t}_2 &= (-1, \delta) \\ \bar{t}_3 &= (-1, -\delta) & \bar{t}_4 &= (1, -\delta) \end{aligned}$$

ปัญหานี้จะแสดงในรูปที่ 1 ในกรณีนี้เมตริกการกระจายระหว่างกลุ่มเป็นดังนี้

กำหนดให้ $\delta = 0$

$$\frac{1}{4} S_B = \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

เมตริกการกระจายระหว่างกลุ่มจะไม่เกิดการแบ่งแยกในแนวตั้ง จะแยกได้แค่แนวนอนเพียงอย่างเดียว เราสามารถบอกได้ว่าความแปรปรวนของกลุ่มข้อมูลคู่อื่นๆ จะมีอิทธิพลต่อเมตริกการกระจายตัวระหว่างกลุ่มมากกว่า (1,4) และ (2,3) การพิจารณาถึงปัญหาในการแยกความแปรปรวนของกลุ่มข้อมูลคู่ (1,4) และ (2,3) มีความสำคัญมากกว่าอื่นๆ เพราะกลุ่มข้อมูลคู่นี้มีความยุ่งยากมากกว่า การกระจายระหว่างกลุ่มยังคงไม่ยึดจับ (capture) ข้อมูลนี้

จากตัวอย่างข้างต้นจะเห็นว่าถ้ามีข้อมูลบางกลุ่มที่อยู่ใกล้ๆ กันมาเปรียบเทียบกับกลุ่มอื่นๆ เมตริกการกระจายระหว่างกลุ่มส่วนมากจะไม่สนใจข้อมูลการแบ่งแยกระหว่างกลุ่มที่อยู่ใกล้ๆ กัน

การพิจารณาในส่วนที่ผ่านมานำเราไปสู่การกำหนดเมตริกการกระจายระหว่างกลุ่มที่เท่ากับผลรวมของเมตริก Weighted pairwise scatter เนื่องจากการคงข้อมูลการแยกแยะที่เพียงพอเราจะต้องปรับค่าถ่วงน้ำหนัก (weight) การปรับที่ได้รับการยอมรับ คือ จะปรับค่าถ่วง

น้ำหนักให้เท่ากับ อินเวิร์ส (inverse) ระยะทางยูคลิดีียน (Euclidean distance) ของผลต่างค่าเฉลี่ยระหว่างกลุ่ม

$$w = \frac{1}{\|\bar{t}_i - \bar{t}_j\|} = \frac{1}{(\bar{t}_i - \bar{t}_j)^T (\bar{t}_i - \bar{t}_j)} \quad (A)$$

เราจะให้ความสำคัญกับกลุ่มข้อมูลที่มีค่าเฉลี่ยใกล้กับกลุ่มอื่นๆ มากกว่ากลุ่มที่มีค่าเฉลี่ยห่างกัน ในกรณีนี้กลุ่มที่มีการปนกันของข้อมูลมากกว่าจะได้รับการสนใจมากกว่า และกลุ่มที่มีความแปรปรวนของข้อมูลน้อยกว่าจะได้รับความสนใจน้อยกว่า เนื่องมาจากรูปแบบปกติของถ่วงน้ำหนัก เมตริกการกระจายระหว่างกลุ่มจะเป็น

$$S_{B_{inv}} = \frac{1}{N_i} \sum_{i=1}^L N_i \frac{(\bar{t}_i - \bar{t})(\bar{t}_i - \bar{t})^T}{(\bar{t}_i - \bar{t})^T (\bar{t}_i - \bar{t})} \quad (B)$$

จากปัญหาที่กำหนดในตัวอย่าง อินเวิร์สระหว่างกลุ่มที่ถูกคำนวณโดย (B) จะเป็นดังนี้

กำหนดให้ $\delta = 0$

$$\frac{1}{4} S_{B_{inv}} = \begin{pmatrix} 1 + \frac{1}{1 + \delta^2} & 0 \\ 0 & 1 + \frac{\delta^2}{1 + \delta^2} \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

ดังนั้นการใช้เมตริกความแปรปรวนระหว่างกลุ่ม (B) จะไม่เกิดอะไรขึ้นกับการเข้าใกล้กันของคู่ (1,4) และ (2,3) เราจะทำให้การกระจายของมันอยู่ในการคำนวณ WPS-CSLDA นี่คือนสิ่งที่ต้องการเนื่องจากปัญหาการแยกข้อมูล

เมื่อเราได้เมตริกการกระจายระหว่างกลุ่มใหม่ เราจะสามารถ การกระจายระหว่างกลุ่มด้วยเมตริกใหม่นี้ นอกเหนือจากนั้นยังคงเดิม

2.2 การคำนวณพีเจอร์ WPS-CSLDA

กำหนดให้แฉกการฉาย $U \in \mathbb{R}^{m \times m}$ และ $V \in \mathbb{R}^{m \times c}$ A คือภาพที่มีขนาด $m \times n$ จากการฉายภาพ A ลงบนแกน U และ V ตามหลักการแปลงเชิงเส้น

$$Y = U^T A V \quad (1)$$

ดังนั้นจะได้ Y คือ พีเจอร์เมตริกของภาพ A ที่มีขนาด $r \times c$ โดยที่ $r < m$ และ $c < n$

เมื่อกำหนดให้มิตัวอย่างภาพทั้งหมด L ภาพ และมีจำนวนกลุ่ม L กลุ่ม แต่ละกลุ่มมีจำนวนภาพในกลุ่ม N_i ภาพ โดยภาพแต่ละภาพที่ใช้สอนมีขนาดเป็น $m \times n$, A_j ($j = 1, 2, \dots, M$) ค่าเฉลี่ยของภาพทั้งหมดที่ใช้สอนคือ \bar{A} และ \bar{A}_i ($i = 1, 2, \dots, L$) คือค่าเฉลี่ยของภาพแต่ละกลุ่ม P_i คือกลุ่มที่ถูกฉาย หลังจากภาพที่ใช้สอนถูกฉายลงบนแกน U และ V จะได้พีเจอร์ทริกเป็น

$$Y_j = U^T A_j V, \quad j = 1, 2, \dots, M \quad (2)$$

วิธีในการหาแกนการฉาย U และ V สามารถคำนวณได้โดยการหาผลรวมค่าไอเกน (trace) ของเมทริกซ์ความแปรปรวนร่วมจากภาพตัวอย่างที่ถูกฉาย [1] จากแนวความคิดนี้สามารถสร้างเป็นสมการได้คือ

$$J(U, V) = \frac{P_B}{P_W} \quad (3)$$

เมื่อ

$$P_B = tr(\tilde{S}_B)$$

$$P_W = tr(\tilde{S}_W)$$

โดยที่

\tilde{S}_B คือ เมทริกซ์การกระจายระหว่างกลุ่มของภาพตัวอย่างที่ถูกฉาย

\tilde{S}_W คือ เมทริกซ์การกระจายภายในกลุ่มของภาพตัวอย่างที่ถูกฉาย

$$\begin{aligned} \tilde{S}_B &= \sum_{i=1}^L \omega N_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T \\ &= \sum_{i=1}^L \omega N_i [U(\bar{A}_i - \bar{A})V] [U(\bar{A}_i - \bar{A})V]^T \end{aligned} \quad (4)$$

$$\begin{aligned} \tilde{S}_W &= \sum_{i=1}^L \sum_{Y_k \in P_i} (Y_k - \bar{Y}_i)(Y_k - \bar{Y}_i)^T \\ &= \sum_{i=1}^L \sum_{Y_k \in P_i} [U(A_k - \bar{A}_i)V] [U(A_k - \bar{A}_i)V]^T \end{aligned} \quad (5)$$

ถ้า

$$tr(\tilde{S}_B) = tr\left(\sum_{i=1}^L \omega N_i U^T (\bar{A}_i - \bar{A}) V V^T (\bar{A}_i - \bar{A})^T U\right) \quad (6)$$

$$tr(\tilde{S}_W) = tr\left(\sum_{i=1}^L \sum_{Y_k \in P_i} U^T (A_k - \bar{A}_i) V V^T (A_k - \bar{A}_i)^T U\right) \quad (7)$$

เมื่อ ω คือ weighted pairwise scatter

ในการคำนวณหา U และ V ไม่สามารถทำได้พร้อมกัน ดังนั้นจึงทำการคำนวณทีละค่า โดยกำหนดให้ $V = I$ (I คือ เมทริกซ์เอกลักษณ์) เพื่อหาค่าของ U ลำดับต่อมา กำหนด $U \neq I$ คำนวณหาค่าของ V

การคำนวณหา U

กำหนดให้ $V = I$

$$tr(\tilde{S}_B) = U^T S_B^U U, \quad tr(\tilde{S}_W) = U^T S_W^U U \quad (8)$$

โดยที่

$$S_B^U = \sum_{i=1}^L \omega N_i (\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T$$

$$S_W^U = \sum_{i=1}^L \sum_{Y_k \in P_i} (A_k - \bar{A}_i)(A_k - \bar{A}_i)^T$$

จากสมการที่ (3) สามารถแทนค่าใหม่ได้คือ

$$J(U) = \frac{U^T S_B^U U}{U^T S_W^U U} \quad (9)$$

ในการแก้สมการหา U ที่เหมาะสมที่สุดในการฉาย สามารถทำได้โดยการหาค่ามากที่สุดของ $J(U)$

$$U_{opt} = \arg \max_U J(U) \quad (10)$$

ถ้า S_W^U ไม่เป็นเมทริกซ์เอกฐาน (non singular) สามารถหาเมทริกซ์ U_{opt} โดยการคำนวณไอเกนเวกเตอร์และค่าไอเกนได้จากสมการ

$$S_B^U U_{opt} = \lambda S_W^U U_{opt} \quad (11)$$

โดยปกติแกนการฉายที่เหมาะสมที่สุดที่ใช้ในการฉายมาจากไอเกนเวกเตอร์ของ $S_B^U^{-1} S_W^U$ ที่สัมพันธ์กับค่าไอเกนที่มากที่สุดตั้งแต่ 1 ถึง m แต่ในความเป็นจริงเนื่องจากต้องการหาพีเจอร์ทริกของภาพ ฉะนั้นจึงเลือกแกนในการฉายเพียง 1 ถึง r เท่านั้น ดังนั้นเมทริกซ์ใหม่ที่ใช้ในการฉายจึงมีขนาดเป็น $m \times r$

$$U = [u_1, u_2, \dots, u_r] \quad (12)$$

การคำนวณหา V

การคำนวณหา V กำหนดให้ $U=I$ (I คือ เมทริกซ์เอกลักษณ์) จากนั้นทำการคำนวณทั้งหมื่นเหมือนหา U ตามสมการที่ (8) ถึง (12) โดยเปลี่ยนตัวแปร U เป็น V

2.3 การสร้างภาพใหม่

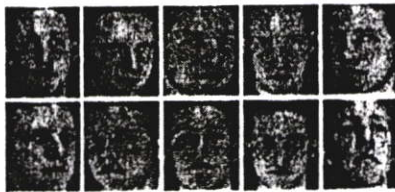
วิธี WPS-CSLDA สามารถนำเอาพีเจอร์เมตริกและแกนการฉายที่เหมาะสมที่สุด U และ V มาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังนี้

กำหนดให้ A เป็นภาพ โดยที่ V เป็นพีเจอร์เมตริกของภาพ A โดยที่ $U = [u_1, \dots, u_r]$ และ $V = [v_1, \dots, v_r]$ เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นสมการในการสร้างภาพใหม่คือ

$$\tilde{A} = UVV^T \tag{13}$$

3. การทดลองและการวิเคราะห์

ฐานข้อมูลภาพใบหน้าของ ORL (<http://www.cam.ac.uk>) ทั้งหมดเป็นภาพระดับสีเทาและมีขนาด 112 x 92 จุดภาพ ถ่ายจากใบหน้าของ 40 บุคคล คนละ 10 ภาพที่แตกต่างกัน เช่น ภาพที่มีการแสดงออกทางใบหน้า (การปิดตาหรือเปิดตา ยิ้มหรือไม่ยิ้ม) ภาพที่มีรายละเอียดอื่น ๆ (การใส่แว่นหรือถอดแว่น) ภาพที่มีการก้มหน้าหรือเงยหน้า หันซ้ายหรือขวา ไม่เกินร้อยละ 20 โดย รูปที่ 2 เป็นตัวอย่างภาพใบหน้าทั้ง 10 ภาพของบุคคลเดียว ในการทดลองทั้งหมดจะใช้ 1-Nearest-Neighbor (1NN) เป็นวิธีการแบ่งแยก (classifier)

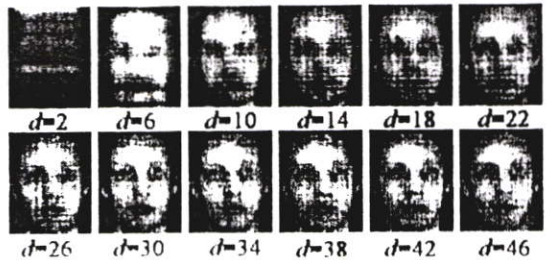


รูปที่ 2 ตัวอย่างภาพใบหน้า 1 บุคคลของฐานข้อมูล ORL.

การทดลองแรกเปรียบเทียบความถูกต้องในการรู้จำใบหน้าสูงสุดของ 3 วิธี คือ วิธี WPS-CSLDA, วิธี CSLDA และวิธี Eigenimage ภายใต้เงื่อนไขของจำนวนตัวอย่างภาพใบหน้าที่ใช้สอน โดยแบ่งการทดสอบออกเป็น 3 ครั้ง ครั้งแรกภาพใบหน้าที่ใช้ฝึกสอนจะถูกแบ่ง

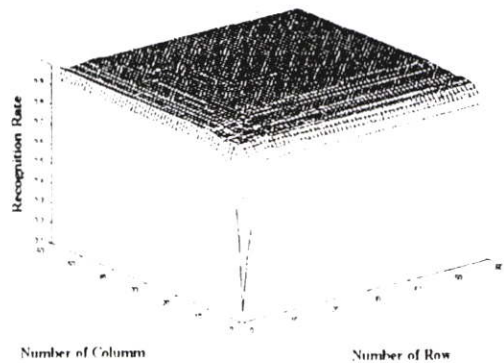
ออกเป็นกลุ่ม 3 ภาพต่อหน้าบุคคล และภาพที่เหลือจะถูกนำมาใช้ทดสอบความถูกต้อง การทดสอบครั้งต่อ ๆ มา ภาพที่ใช้ฝึกสอนจะถูกเลือกมาอย่างสุ่ม 4 และ 5 ภาพต่อหน้าบุคคล ตามลำดับ ดังแสดงในตารางที่ 1

การทดลองที่ 2 ทำการทดสอบคุณสมบัติการสร้างภาพใหม่ด้วยวิธี WPS-CSLDA ในการทดลองภาพที่ใช้ฝึกสอนถูกเลือกมาอย่างสุ่มจำนวน 5 ภาพต่อหน้าบุคคล เพื่อนำมาหาแกนการฉายที่เหมาะสมที่สุดด้วยวิธี WPS-CSLDA เมื่อนำภาพแรกของรูปที่ 2 มาฉายลงบนแกนการฉายที่เหมาะสมที่สุดทั้งสองแกนจะได้พีเจอร์เมตริกที่ใช้ในการสร้างภาพใหม่ รูปที่ 3 แสดงให้เห็นว่าเมื่อจำนวน d (d คือ $r - c$) เพิ่มขึ้นภาพที่สร้างใหม่จะมีความเหมือนภาพต้นแบบมากขึ้น



รูปที่ 3 ภาพที่สร้างขึ้นใหม่โดย WPS-CSLDA

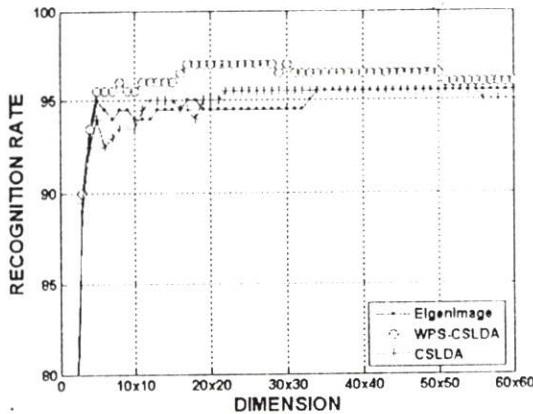
การเปรียบเทียบความถูกต้องในการรู้จำของทั้ง 3 วิธี นั้น ทำได้ยากเนื่องจากความถูกต้องในการรู้จำของทั้ง 3 วิธี แสดงเป็นกราฟ 3 มิติ ดังรูปที่ 4 ฉะนั้นในการเปรียบเทียบจึงจำเป็นต้องเลือกความถูกต้องในการรู้จำมา 1 แนวเพื่อนำมาเปรียบเทียบ ในที่นี้จะนำความถูกต้องการรู้จำในแนวทแยงมาเปรียบเทียบ ผลการเปรียบเทียบแสดงในรูปที่ 5



รูปที่ 4 ความถูกต้องในการรู้จำของ WPS-CSLDA

ตารางที่ 1 การเปรียบเทียบการรู้จำสูงสุดของทั้ง 3 วิธี

Training sample per class	WPS-CSLDA		CSLDA		Eigenimage	
	Accuracy (%)	Dimension	Accuracy (%)	Dimension	Accuracy (%)	Dimension
3	96.40	21x2	95.71	22x3	93.33	13x3
4	97.35	14x2	96.25	10x2	94.17	21x2
5	98.50	13x2	97.50	13x2	95.50	18x2



รูปที่ 5 เปรียบเทียบร้อยละความถูกต้องระหว่าง CSLDA, Eigenimage และ WPS-CSLDA

4. สรุปผลการทดลอง

บทความนี้เสนอหลักการ ใหม่สำหรับวิธีการหาพีเจอร์ของภาพที่มีชื่อว่า WPS-CSLDA วิธีการนี้เป็นการนำ Weighted Pairwise Scatter มาใช้ในการคำนวณเมตริกการกระจายระหว่างกลุ่ม ของการหาพีเจอร์ CSLDA เพื่อลดมิติของภาพทั้งแถวและหลัก นอกจากนี้ยังทำให้ข้อมูลระหว่างกลุ่มที่มีการปนกันแยกห่างออกจากกัน จากการทดลองแสดงให้เห็นว่าการรู้จำใบหน้าสูงสุดของวิธี WPS-CSLDA เมื่อเปรียบเทียบกับวิธี Eigenimage และวิธี CSLDA มีความถูกต้องในการรู้จำใบหน้าดีที่สุด และเมื่อนำการรู้จำทั้งหมดมาเปรียบเทียบกันทั้ง 3 วิธีพบว่า WPS-CSLDA ให้ผลการรู้จำดีที่สุด

5. เอกสารอ้างอิง

[1] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

- [2] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.
- [3] J. Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, No. 1, pp. 135-137, 2004.
- [4] S. Kongsontana and Y. Rangsanteri, "Face Recognition Using 2DLDA Algorithm," Eighth International Symposium on Signal Processing and Its Applications, ISSPA 2005, August 2005.
- [5] D. Zhang, S. Chen, J. Liu, "Representing Image Matrices: Eigenimages vs. Eigenvectors," In International Symposium on Neural Networks (ISNN 2005), pp. 659-664, 2005.
- [6] สิทธิพันธ์ คงสมทนา, พุทธพงษ์ รังสรรค์เสรี และ วีรวัฒน์ ลีลาธรรม "การรู้จำใบหน้าโดยใช้หลักการของ CSLDA," เอกสารการประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 28 Vol.2 ตุลาคม 2548.
- [7] Y. Li, Y. Gao, H. Erdogan "Weighted Pairwise Scatter To Improve Linear Discriminant Analysis" in ICSP-2000, vol. 4, pp.608-611, 2000.

ประวัติผู้เขียน

ผู้เขียน	นายวีรรัตน์ สีลาธรรม
วันเดือนปีเกิด	14 ตุลาคม 2524
สถานที่เกิด	กรุงเทพมหานคร
ที่อยู่	53 ซ.อักษรลักษณ์ ถ.สุขุมวิท ต.ปากน้ำ อ.เมืองฯ จ.สมุทรปราการ 10270 โทร. 08-689-305-71 e-mail : neung278@gmail.com, neung278@yahoo.com
ประวัติการศึกษา	ปริญญาวิทยาศาสตรบัณฑิต สาขาฟิสิกส์ประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง 2546