



การทำฮาโลซิเนสภาพใบหน้าที่มีหลายมุมมองโดยการฝังตัวเชิงเส้นแบบเฉพาะส่วนที่
ปรับตัวได้

MULTIVIEW FACE HALLUCINATION BY ADAPTIVE LOCALLY LINEAR
EMBEDDING

โดย

เมธี ทองดี

คุณฉันทิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต

ปีการศึกษา 2556



**MULTIVIEW FACE HALLUCINATION BY ADAPTIVE LOCALLY LINEAR
EMBEDDING**

**BY
METEE THONGDEE**

**A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY IN INFORMATION
TECHNOLOGY**

FACULTY OF INFORMATION TECHNOLOGY

GRADUATE SCHOOL, RANGSIT UNIVERSITY

2013



คุษฎีนิพนธ์เรื่อง

การทำฮาอูซิเนสภาพใบหน้าที่มีหลายมุมมองโดยการฝังตัวเชิงเส้น
แบบเฉพาะส่วนที่ปรับตัวได้

โดย

เมธิ ทองดี

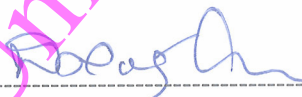
ได้รับการพิจารณาให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาปรัชญาคุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

มหาวิทยาลัยรังสิต

ปีการศึกษา 2556



รศ.ดร.บุญวัฒน์ อัดชู
ประธานกรรมการสอบ



ผศ.ดร.พิศิษฎ์ โภการัตน์กุล
กรรมการ



รศ.ดร.ชม กิมปาน
กรรมการ



ดร.เกียรติศักดิ์ ใหม่เจริญกุล
กรรมการ



ผศ.ดร.สิริพร สุภราทิตย์
กรรมการและอาจารย์ที่ปรึกษา



ผศ.ดร.ปริญญา สงวนสัตย์
กรรมการและอาจารย์ที่ปรึกษาร่วม

บัณฑิตวิทยาลัยรับรองแล้ว



(ผศ.ร.ต.หญิง ดร.วรรณิ์ สุขสาตร)

คณบดีบัณฑิตวิทยาลัย

5 พฤษภาคม 2557



Dissertation entitled

**MULTIVIEW FACE HALLUCINATION BY ADAPTIVE LOCALLY
LINEAR EMBEDDING**

by

METEE THONGDEE

was submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Information Technology

Rangsit University
Academic Year 2013

Boonwat Attachoo

Assoc.Prof.Boonwat Attachoo, D.Eng.
Examination Committee Chairperson

Phisit Phokharatkul

Asst.Prof.Phisit Phokharatkul, D.Eng.
Member

Chom Kimpan

Assoc.Prof.Chom Kimpan, D.Eng.
Member

Kiattisak Maichalernnukul

Kiattisak Maichalernnukul, D.Eng.
Member

Siriporn Supratid

Asst.Prof.Siriporn Supratid, D.Tech. Sci./CS.
Member and Advisor

Parinya Sanguansat

Asst.Prof.Parinya Sanguansat, D.Eng.
Member and Co-Advisor

Approved by Graduate School

Vannee Sooksatra

(Asst.Prof.Pl't.Off. Vannee Sooksatra, D.Eng.)

Dean of Graduate School

May 5, 2014

กิตติกรรมประกาศ

ในการที่คุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงลงได้ด้วยดี ผู้วิจัยใคร่ขอกราบขอบพระคุณท่าน ผู้ช่วยศาสตราจารย์ ดร.สิริพร ศุภราทิตย์ และผู้ช่วยศาสตราจารย์ ดร.ปริญญา สงวนสัตย์ อาจารย์ที่ปรึกษาคุษฎีนิพนธ์ที่กรุณาให้คำปรึกษาทั้งทางด้านการวิเคราะห์ เอกสารอ้างอิงและตรวจสอบต้นฉบับคุษฎีนิพนธ์ อันเป็นผลให้งานวิจัยนี้มีความสมบูรณ์ยิ่งขึ้นและยิ่งไปกว่านั้น ความเอาใจใส่ดูแลตลอดเวลาในการดำเนินงานเป็นสิ่งที่ผู้ทำคุษฎีนิพนธ์มีความซาบซึ้งใจเป็นอย่างสูงยิ่ง และใคร่ขอกราบขอบพระคุณรองศาสตราจารย์ ดร.บุญวัฒน์ อัครชู ประธานกรรมการสอบคุษฎีนิพนธ์ ที่ให้แนวความคิด ตลอดจนชี้แนะแนวทางที่เป็นประโยชน์ต่อการวิจัยนี้ อีกทั้งใคร่ขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ชม กิมปาน ผู้ช่วยศาสตราจารย์ ดร.พิศิษฎ์ โภคารัตน์กุล และอาจารย์ ดร.เกียรติศักดิ์ ไหม้เจริญกุล คณะกรรมการสอบคุษฎีนิพนธ์ ที่กรุณาให้ข้อเสนอแนะที่มีค่า ทำให้เสริมความสมบูรณ์ของการวิจัยได้ดียิ่งขึ้น

ท้ายที่สุดนี้ ผู้วิจัยสำนึกคุณค่าของตำรา เอกสารอ้างอิง และความรู้จากคณาจารย์ทุกท่านที่ได้นำมากล่าวอ้างไว้ด้วยความเคารพอย่างสูง ความอนุเคราะห์ทุนสนับสนุนจากมหาวิทยาลัยราชภัฏนครราชสีมา และกำลังใจจาก พ่อ แม่ ดร.เมธาวิ ทองดี และดร.ณัฐฉิณี ทองดี ตลอดจนผู้มีส่วนสนับสนุนทุกท่านที่ได้กล่าวไว้หรือที่ไม่ได้กล่าวไว้ในที่นี้ จะถูกจารึกในจิตใจของผู้วิจัยด้วยความระลึกถึงตลอดไป

เมธี ทองดี

ผู้วิจัย

5106732 : สาขาวิชาเอก: เทคโนโลยีสารสนเทศ; ปร.ด. (เทคโนโลยีสารสนเทศ)

คำสำคัญ : ภาพใบหน้าที่มีหลายมุมมอง, การทำฮาโลกราฟีในสภาพใบหน้า, ภาพความละเอียดสูง
 เมธี ทองดี: การทำฮาโลกราฟีในสภาพใบหน้าที่มีหลายมุมมองโดยการฝังตัวเชิงเส้นแบบเฉพาะ
 ส่วนที่ปรับตัวได้ (MULTIVIEW FACE HALLUCINATION BY ADAPTIVE LOCALLY
 LINEAR EMBEDDING) อาจารย์ที่ปรึกษา: ผศ.ดร.สิริพร สุภราทิพย์, อาจารย์ที่ปรึกษาร่วม:
 ผศ.ดร.ปริญญา สงวนสัตย์, 148 หน้า.

คุณฉันทิพนธ์นี้ มีจุดประสงค์เพื่อเสนอวิธีการใหม่ในการเพิ่มคุณภาพของการสร้างคืนภาพ
 ใบหน้าให้มีความละเอียดสูงจากภาพความละเอียดต่ำ ซึ่งเนื้อหาประกอบด้วยเทคนิคใหม่ คือ การ
 นำเสนอขั้นตอนวิธี Adaptive Locally Linear Embedding (ALLE) สำหรับการคำนวณหาค่าน้ำหนัก
 (W) โดยปรับปรุงจากขั้นตอนวิธี LLE เดิมในการทดลองสร้างคืนภาพใบหน้าที่มีความละเอียดสูงใน
 มุมมองหน้าตรง (Frontal view) โดยการพิจารณาจำนวน Neighbor ที่เกี่ยวข้องในแต่ละจุดตัวอย่าง
 จากค่าขีดเริ่มเปลี่ยนความคล้ายของภาพ (Threshold of similarity) จากภาพที่นำเข้า โดยไม่ได้
 คำนวณจากภาพฝึกฝนทั้งหมด แต่ตามความจริงเราไม่สามารถที่จะได้ภาพใบหน้าที่เป็นหน้าตรงได้
 ดังนั้นในคุณฉันทิพนธ์นี้ ได้เสนอวิธีการใหม่ในการสร้างคืนภาพที่มีหลายมุมมอง โดยมีวัตถุประสงค์
 เพื่อสร้างคืนภาพใบหน้าที่ตรง และใบหน้าในมุมมองอื่น ๆ โดยมีขั้นตอนการประมวลผลตามวิธีที่
 เสนอดังต่อไปนี้ เริ่มจากนำเข้าภาพใบหน้าที่ความละเอียดต่ำ 1 ภาพในมุมมอง หน้าตรง หน้าแยง หน้า
 ก้ม ใบหน้าด้านซ้าย และใบหน้าที่ด้านขวา จากนั้นภาพความละเอียดต่ำในมุมมองอื่น ๆ จะถูกสร้าง
 ขึ้นโดยวิธีการ ALLE ซึ่งใช้ ค่าขีดเริ่มเปลี่ยนความคล้ายของภาพ (Threshold of similarity) มาทำการ
 เลือกจุดที่ใกล้จุดที่สนใจ จากนั้นจะใช้ภาพค่าขีดเริ่มเปลี่ยนความคล้ายของภาพ (Threshold of
 similarity) เดิมในการสร้างภาพใบหน้าที่ความละเอียดสูงในทุกมุมมอง

ผลการทดลองบนฐานข้อมูลใบหน้าที่มาตรฐานพบว่า วิธีการที่นำเสนอขึ้นสามารถ
 เพิ่มคุณภาพของการสร้างคืนภาพใบหน้าที่ทั้งแบบหน้าตรงและใบหน้าที่หลายมุมมองได้ดีกว่าวิธีการ
 แบบดั้งเดิม

Student's Signature M.T

Thesis Advisor's Signature

Thesis Co-Advisor's Signature

5106732 : MAJOR: INFORMATION TECHNOLOGY; Ph.D.


KEYWORDS : MULTIVIEW, FACE HALLUCINATION, SUPER RESOLUTION


**METEE THONGDEE: MULTIVIEW FACE HALLUCINATION BY ADAPTIVE
 LOCALLY LINEAR EMBEDDING. DISSERTATION ADVISOR: ASST.PROF.
 SIRIPORN SUPRATID, D.Tech.Sc./CS, DISSERTATION CO-ADVISOR: ASST.PROF.
 PARINYA SANGUANSAT, D.ENG., 148 p.**

The objective of this dissertation is to propose the novel of frameworks in order to improve the quality of facial image reconstruction which providing the High-Resolution (HR) from the Low-Resolution (LR) image. The contents contain the newest technique, called Adaptive locally linear embedding (ALLE), referred to a modified version of LLE was proposed to apply with frontal view face hallucination; it uses a threshold of similarity for selecting the neighbors of each point. However, frontal face is barely captured in the real world. Therefore, this dissertation proposes a novel ALLE for multiview face hallucination. The main objective is to generate high quality of frontal and non-frontal face images. The processing steps, according to the proposed method are operated as follows; first, LR face in one of front, up, down, left or right views is fed as an input; then, the other views of such an LR image are generated by ALLE, which applies a threshold of similarity for selecting the neighbors of each point; and HR face images in all views of the same input object are achieved afterwards.

The experimental results show that the proposed method yields the better image quality of the reconstructed frontal and non-frontal face images over the baseline methods.

Student's Signature **M.T**

Thesis Advisor's Signature 

Thesis Co-Advisor's Signature 

สารบัญ

	หน้า
กิตติกรรมประกาศ	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
สารบัญ	ง
สารบัญตาราง	ช
สารบัญรูป	ซ
สัญลักษณ์และคำย่อ	ฐ
บทที่ 1	
บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 กรอบแนวคิดของการวิจัย	3
1.4 ขอบเขตการศึกษา	4
1.5 นิยามศัพท์	4
บทที่ 2	
แนวคิดและทฤษฎีที่เกี่ยวข้องกับงานวิจัย	6
2.1 Super Resolution Reconstruction (SRR)	6
2.1.1 Interpolation Based	6
2.1.2 Reconstruction Based	11
2.1.3 Learning Based Super-Resolution Method	11
2.2 Patch based	14
2.3 Linear Combination	17
2.4 Locally Linear Embedding(LLE)	18
2.5 งานวิจัยที่เกี่ยวข้อง	21

สารบัญ(ต่อ)

บทที่ 3	วิธีดำเนินงานวิจัย	23
	3.1 การสร้างฐานข้อมูลภาพใบหน้าของคนไทย	23
	3.2 กระบวนการในการทำวิจัย (Research Methodology)	25
	3.3 Adaptive Locally Linear Embedding (ALLE)	31
	3.3.1 ขั้นตอนการฮาหลูซิเนสภาพใบหน้า (face hallucination algorithm)	33
	3.4 Multiview Face Hallucination by Adaptive Locally Linear Embedding	35
	3.5 การประเมินคุณภาพของภาพ	37
บทที่ 4	ผลการวิจัย	38
	4.1 ฐานข้อมูลภาพใบหน้า	38
	4.1.1 The CAS-PEAL-R1	38
	4.1.2 ฐานข้อมูลภาพใบหน้าของคนไทย	40
	4.2 กระบวนการก่อนหน้า (Preprocessing)	41
	4.2.1 การจัดวางตำแหน่งของภาพใบหน้า (Face alignment)	41
	4.3 ผลการทดลอง (Experimental results)	46
	4.3.1 การฮาหลูซิเนสภาพใบหน้า (Face Hallucination)	46
	4.3.1.1 การฮาหลูซิเนสภาพใบหน้าจากฐานข้อมูลภาพ CAS-PEAL-R1	46
	4.3.1.2 การฮาหลูซิเนสภาพใบหน้าจากฐานข้อมูลภาพใบหน้าของคนไทย	55
	4.3.2 การฮาหลูซิเนสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination)	64
	4.3.2.1 การฮาหลูซิเนสภาพใบหน้าที่มีหลายมุมมองจากฐานข้อมูล CAS-PEAL-R1	64
	4.3.2.2 การฮาหลูซิเนสภาพใบหน้าที่มีหลายมุมมองจากฐานข้อมูลใบหน้าคนไทย	85

สารบัญ(ต่อ)

บทที่ 5	สรุปผลและข้อเสนอแนะ	106
	5.1 สรุปผลงานวิจัย	106
	5.2 ข้อเสนอแนะ	107
บรรณานุกรม		108
ภาคผนวก		112
ประวัติผู้วิจัย		148

มหาวิทยาลัยรังสิต
Rangsit University

สารบัญตาราง

ตารางที่	หน้า
4.1 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูงของใบหน้าตรง (K=100, $\theta = 0.1$)	54
4.2 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูงของใบหน้าตรง (K=200, $\theta = 0.2$)	54
4.3 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูงของใบหน้าตรง (K=100, $\theta = 0.1$)	63
4.4 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูงของใบหน้าตรง (K=200, $\theta = 0.2$)	63
4.5 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=100, $\theta = 0.1$)	83
4.6 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	83
4.7 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	84
4.8 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	84
4.9 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	84
4.10 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=100, $\theta = 0.1$)	104
4.11 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=100, $\theta = 0.1$)	104
4.12 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	105
4.13 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=100, $\theta = 0.1$)	105
4.14 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง (K=200, $\theta = 0.2$)	105

สารบัญรูป

รูปที่	หน้า
1.1 ตัวอย่างภาพใบหน้าคนร้ายที่ได้จากกล้องวงจรปิด	2
1.2 กรอบแนวคิดในการทำวิจัย	3
2.1 Bicubic-interpolation neighborhood	7
2.2 ตัวอย่างภาพค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x,y)	8
2.3 ตัวอย่างภาพใบหน้าที่สร้างโดย Bicubic Interpolation	10
2.4 ตัวอย่างรูปที่ผ่านวิธีการ Reconstruction Based	11
2.5 ตัวอย่างรูปที่ผ่านวิธีการ Reconstruction Based	11
2.6 ภาพความสัมพันธ์ระหว่างภาพความละเอียดสูงและภาพความละเอียดต่ำโดยใช้ตำแหน่ง (i, j)	14
2.7 ภาพใบหน้า patch matrix	15
2.8 Patch matrix	16
2.9 Patch ขนาด $n \times n$ จุดภาพ	16
2.10 Overlap patch	16
2.11 Linear Combination for Multiview Face	17
2.12 Locally Linear Embedding (LLE)	19
3.1 รูปแบบการตั้งกล้องถ่ายภาพ	24
3.2 ตัวอย่างการถ่ายภาพ	24
3.3 ตัวอย่างภาพใบหน้าในมุมมองต่างๆ	25
3.4 ตัวอย่างภาพจากฐานข้อมูล The CAS-PEAL-R1 ขนาด 360 x 480 จุดภาพ	26
3.5 การจัดวางตำแหน่งของภาพใบหน้า	27
3.6 ตัวอย่างกระบวนการจัดวางตำแหน่งภาพใบหน้า	28
3.7 ตัวอย่างภาพใบหน้าที่ผ่านกระบวนการจัดวางตำแหน่ง	28
3.8 ตัวอย่างภาพใบหน้าที่ผ่านกระบวนการตัดภาพ	28
3.9 ตัวอย่างภาพนำเข้าความละเอียดสูงขนาด 128 x 96 จุดภาพ	30

สารบัญรูป(ต่อ)

รูปที่	หน้า
3.10 ตัวอย่างภาพปรับลดความละเอียดขนาด 32 x 24 จุดภาพ	30
3.11 ตัวอย่างภาพนำเข้าความละเอียดต่ำขนาด 32 x 24 จุดภาพ	30
3.12 ตัวอย่างภาพความละเอียดต่ำมุมมองต่างๆที่สร้างขึ้นจากภาพนำเข้าขนาด 32 x 24 จุดภาพ	31
3.13 ตัวอย่างภาพความละเอียดสูงมุมมองต่างๆที่สร้างขึ้นขนาด 128 x 96 จุดภาพ	31
3.14 ภาพความสัมพันธ์ระหว่างภาพความละเอียดสูงและภาพความละเอียดต่ำ โดยใช้ตำแหน่ง (i, j)	32
3.15 Model position-patch	34
3.16 Multiview Face Hallucination by Adaptive Locally Linear Embedding framework	36
4.1 ตัวอย่างภาพจากฐานข้อมูล CAS-PEAL-R1	39
4.2 ตัวอย่างภาพจากฐานข้อมูลภาพใบหน้าคนไทย	40
4.3 แผนภาพแสดงกระบวนการก่อนหน้า	41
4.4 ตัวอย่างการจัดวางตำแหน่งของภาพใบหน้า	42
4.5 ตัวอย่างภาพใบหน้าความละเอียด 128 x 96 จุดภาพ (CAS-PEAL-R1 database)	42
4.6 ตัวอย่างภาพใบหน้าความละเอียด 128 x 96 จุดภาพ (Thai face database)	43
4.7 ตัวอย่างภาพใบหน้าความละเอียด 32 x 24 จุดภาพ (CAS-PEAL-R1 database)	44
4.8 ตัวอย่างภาพใบหน้าความละเอียด 32 x 24 จุดภาพ (Thai face database)	45
4.9 รูปภาพนำเข้าใบหน้าตรง ขนาด 32 x 24 จุดภาพ	47
4.10 รูปฮาโลจีนเนสภาพใบหน้าตรง (กำหนด $K = 100$, $\theta = 0.1$)	47
4.11 รูปฮาโลจีนเนสภาพใบหน้าตรง (กำหนด $K = 200$, $\theta = 0.2$)	47
4.12 รูปภาพนำเข้าใบหน้าเงย ขนาด 32 x 24 จุดภาพ	48
4.13 รูปฮาโลจีนเนสภาพใบหน้าเงย (กำหนด $K = 100$, $\theta = 0.1$)	48

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.14 รูปฮาโลจีนเนสภาพไบหน้าแงย (กำหนด $K = 200$, $\theta = 0.2$)	48
4.15 รูปภาพนำเข้าไบหน้าก้ม ขนาด 32 x 24 จุดภาพ	49
4.16 รูปฮาโลจีนเนสภาพไบหน้าก้ม (กำหนด $K = 100$, $\theta = 0.1$)	49
4.17 รูปฮาโลจีนเนสภาพไบหน้าก้ม (กำหนด $K = 200$, $\theta = 0.2$)	49
4.18 รูปภาพนำเข้าไบหน้าด้านซ้าย ขนาด 32 x 24 จุดภาพ	50
4.19 รูปฮาโลจีนเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 100$, $\theta = 0.1$)	50
4.20 รูปฮาโลจีนเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)	50
4.21 รูปภาพนำเข้าไบหน้าด้านขวา ขนาด 32 x 24 จุดภาพ	51
4.22 รูปฮาโลจีนเนสภาพไบหน้าด้านขวา (กำหนด $K = 100$, $\theta = 0.1$)	51
4.23 รูปฮาโลจีนเนสภาพไบหน้าด้านขวา (กำหนด $K = 200$, $\theta = 0.2$)	51
4.24 แสดงการเปรียบเทียบผลของการฮาโลจีนเนส โดยกำหนดค่า K ที่ต่างกัน (Single View)	52
4.25 แสดงการเปรียบเทียบผลของการฮาโลจีนเนส โดยกำหนดค่า θ ที่ต่างกัน (Single View)	53
4.26 รูปภาพนำเข้าหน้าตรง ขนาด 32 x 24 จุดภาพ	55
4.27 รูปฮาโลจีนเนสภาพไบหน้าตรง (กำหนด $K = 100$, $\theta = 0.1$)	55
4.28 รูปฮาโลจีนเนสภาพไบหน้าตรง (กำหนด $K = 200$, $\theta = 0.2$)	56
4.29 รูปภาพนำเข้าไบหน้าแงย ขนาด 32 x 24 จุดภาพ	56
4.30 รูปฮาโลจีนเนสภาพไบหน้าแงย (กำหนด $K = 100$, $\theta = 0.1$)	56
4.31 รูปฮาโลจีนเนสภาพไบหน้าแงย (กำหนด $K = 200$, $\theta = 0.2$)	57
4.32 รูปภาพนำเข้าไบหน้าก้ม ขนาด 32 x 24 จุดภาพ	57
4.33 รูปฮาโลจีนเนสภาพไบหน้าก้ม (กำหนด $K = 100$, $\theta = 0.1$)	57
4.34 รูปฮาโลจีนเนสภาพไบหน้าก้ม (กำหนด $K = 200$, $\theta = 0.2$)	58
4.35 รูปภาพนำเข้าไบหน้าด้านซ้าย ขนาด 32 x 24 จุดภาพ	58
4.36 รูปฮาโลจีนเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 100$, $\theta = 0.1$)	58
4.37 รูปฮาโลจีนเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)	59

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.38	59
4.39	59
4.40	60
4.41	61
4.42	62
4.43	64
4.44	64
4.45	65
4.46	65
4.47	65
4.48	66
4.49	67
4.50	68
4.51	69
4.52	70
4.53	71
4.54	72

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.55 รูปฮาโลแกรมภาพใบหลายมุมมองจากภาพนำเข้าไปทางด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)	73
4.56 รูปฮาโลแกรมภาพใบหลายมุมมองจากภาพนำเข้าไปทางด้านขวา (กำหนด $K = 100$, $\theta = 0.1$)	74
4.57 รูปฮาโลแกรมภาพใบหลายมุมมองจากภาพนำเข้าไปทางด้านขวา (กำหนด $K = 200$, $\theta = 0.2$)	75
4.58 แสดงการเปรียบเทียบผลของการฮาโลแกรมโดยกำหนดค่า K ที่ต่างกัน (Multiview)	76
4.59 แสดงการเปรียบเทียบผลของการฮาโลแกรมโดยกำหนดค่า θ ที่ต่างกัน (Multiview)	77
4.60 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Frontal View)	78
4.61 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Frontal View)	78
4.62 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Up View)	79
4.63 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Up View)	79
4.64 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Down View)	80
4.65 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Down View)	80
4.66 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Left View)	81
4.67 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Left View)	81

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.68 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Right View)	82
4.69 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Right View)	82
4.70 รูปภาพนำเข้าไปหน้าตรง ขนาด 32 x 24 จุดภาพ	85
4.71 รูปภาพนำเข้าไปหน้าแยง ขนาด 32 x 24 จุดภาพ	85
4.72 รูปภาพนำเข้าไปหน้าก้ม ขนาด 32 x 24 จุดภาพ	86
4.73 รูปภาพนำเข้าไปหน้าด้านซ้าย ขนาด 32 x 24 จุดภาพ	86
4.74 รูปภาพนำเข้าไปหน้าด้านขวา ขนาด 32 x 24 จุดภาพ	86
4.75 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าตรง (กำหนด $K = 100$, $\theta = 0.1$)	87
4.76 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าตรง (กำหนด $K = 200$, $\theta = 0.2$)	88
4.77 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าแยง (กำหนด $K = 100$, $\theta = 0.1$)	89
4.78 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าแยง (กำหนด $K = 200$, $\theta = 0.2$)	90
4.79 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าก้ม (กำหนด $K = 100$, $\theta = 0.1$)	91
4.80 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าก้ม (กำหนด $K = 200$, $\theta = 0.2$)	92
4.81 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าด้านซ้าย (กำหนด $K = 100$, $\theta = 0.1$)	93
4.82 รูปฮาโลจีนเนสภาพไบลายมุมมองจากภาพนำเข้าไปหน้าด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)	94

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.83 รูปฮาโลแกรมสีหลายมุมมองจากภาพนำเข้าใบหน้าด้านขวา (กำหนด $K = 100$, $\theta = 0.1$)	95
4.84 รูปฮาโลแกรมสีหลายมุมมองจากภาพนำเข้าใบหน้าด้านขวา (กำหนด $K = 200$, $\theta = 0.2$)	96
4.85 แสดงการเปรียบเทียบผลของการฮาโลแกรมสีโดยกำหนดค่า K ที่ต่างกัน (Multiview)	97
4.86 แสดงการเปรียบเทียบผลของการฮาโลแกรมสีโดยกำหนดค่า θ ที่ต่างกัน (Multiview)	98
4.87 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Frontal View)	99
4.88 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Frontal View)	99
4.89 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Up View)	100
4.90 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Up View)	100
4.91 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Down View)	101
4.92 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Down View)	101
4.93 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Left View)	102
4.94 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Left View)	102

สารบัญรูป(ต่อ)

รูปที่		หน้า
4.95	กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Right View)	103
4.96	กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Right View)	103

มหาวิทยาลัยรังสิต
Rangsit University

สัญลักษณ์และคำย่อ

สัญลักษณ์	ความหมาย
SR	Super-Resolution
SRR	Super-Resolution Reconstruction
HR	High-Resolution
LR	Low-Resolution
LLE	Locally Linear Embedding
ALLE	Adaptive Locally Linear Embedding
CCTV	Closed-circuit television
PSNR	Peak Signal-to-Noise Ratio

มหาวิทยาลัยรังสิต
Rangsit University

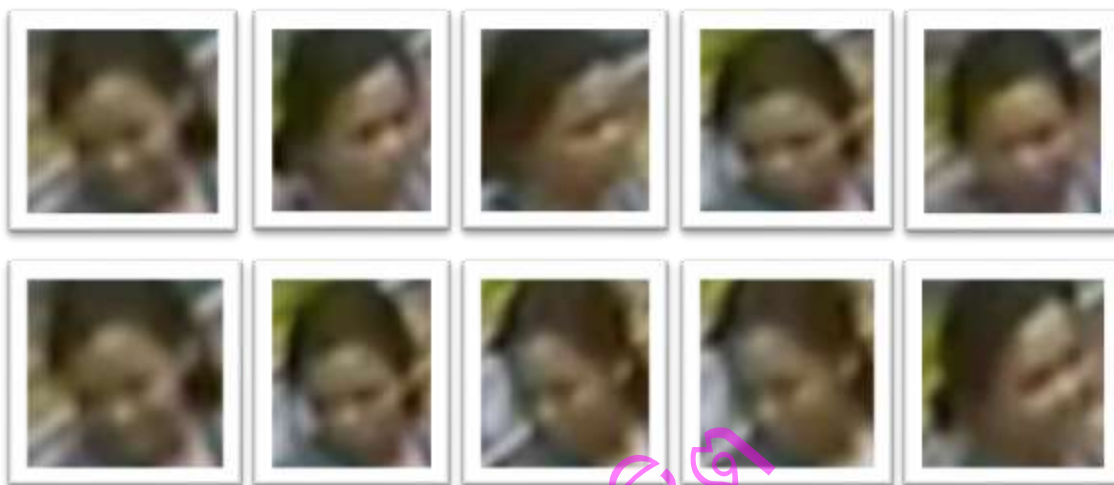
บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในสภาพเศรษฐกิจปัจจุบันอาชญากรรมเกิดขึ้นได้ตลอดเวลาการสูญหายของทรัพย์สินภายในบ้านที่พักอาศัยสถานที่ทำงานสถานที่ราชการสถานที่สาธารณะเช่นห้างสรรพสินค้าบางครั้งเกิดเหตุไม่คาดฝันมีโจรผู้ร้ายปล้นร้านทองกลางห้างขโมยเพชรและลอยนวลไปได้หรือในเหตุการณ์ความไม่สงบที่เกิดขึ้นใน 3 จังหวัดชายแดนภาคใต้ของประเทศไทยก็เกิดเหตุการณ์ที่ไม่คาดคิดขึ้นหลายต่อหลายครั้งเช่นการลอบวางระเบิดในสถานที่สาธารณะต่างๆตามถนนหน้าร้านค้าหน้าโรงแรมหรือเหตุการณ์ระเบิดที่เกิดขึ้นในกรุงเทพมหานครซึ่งสร้างความเสียหายต่อชีวิตและทรัพย์สินความเสียหายต่อการท่องเที่ยวทั้งยังเป็นการลดความเชื่อมั่นความน่าเชื่อถือของต่างประเทศในการที่จะเข้ามาทำธุรกิจหรือเข้ามาท่องเที่ยวที่ประเทศไทย

ระบบกล้องวงจรปิดหรือระบบโทรทัศน์วงจรปิด (Close Circuit Television (CCTV) System หรือ Video Surveillance System) เป็นระบบการบันทึกภาพเคลื่อนไหวที่ถูกจับภาพโดยกล้องวงจรปิด (CCTV Camera) ซึ่งเป็นระบบสำหรับการใช้เพื่อการรักษาความปลอดภัยหรือใช้เพื่อการสอดส่องดูแลเหตุการณ์หรือสถานการณ์ต่างๆที่นอกเหนือจากการรักษาความปลอดภัยซึ่งระบบ CCTV นี้ได้ถูกนำมาติดตั้งเพื่อใช้งานในการบันทึกภาพเหตุการณ์ที่เกิดขึ้นรวมทั้งจะสามารถเห็นใบหน้าของคนร้ายทั้งนี้เพื่อแก้ปัญหาในการติดตามและการจับกุมคนร้ายเพื่อมาดำเนินคดีปัญหาของการตรวจสอบการบันทึกข้อมูลโดยเฉพาะภาพใบหน้าของคนร้ายซึ่งภาพที่ได้จากกล้องวงจรปิด (CCTV Camera) เป็นภาพที่มีความละเอียดต่ำ (Low Resolution) ซึ่งจะมีปัญหาในการที่จะระบุตัวของคนร้ายได้ว่าคนร้ายเป็นใคร



รูปที่ 1.1 ตัวอย่างภาพใบหน้าคนร้ายที่ได้จากกล้องวงจรปิด

การแก้ปัญหาดังกล่าวต้องใช้กระบวนการ Super-Resolution (SR) เพื่อช่วยในการฟื้นฟูจากภาพที่มีความละเอียดต่ำ (Low Resolution) ให้เป็นภาพที่มีความละเอียดสูง (High Resolution) ซึ่งการสร้างคืนภาพโดยใช้กระบวนการ Super-Resolution (SR) นั้นต้องใช้ภาพจำนวนหลายๆภาพ เพื่อสร้างภาพใหม่นั้นคือข้อจำกัดของ Super-Resolution (SR) โดยซึ่งงานวิจัยที่ผ่านมาได้มีการพัฒนากระบวนการที่เรียกว่า Face Hallucination โดยให้ความสนใจเฉพาะภาพใบหน้าคนซึ่งภาพใบหน้าคนจะเป็นภาพที่มีลักษณะร่วมกันซึ่งเราสามารถฝึกฝนให้กับระบบได้โดยกระบวนการ Face Hallucination เป็นวิธีการสร้างคืนภาพจากภาพความละเอียดต่ำ (Low Resolution) ให้เป็นภาพที่มีความละเอียดสูง (High Resolution) โดยใช้ภาพเพียงภาพเดียวและเป็นภาพหน้าตรงเป็นหลักแต่ภาพที่ได้จากกล้องวงจรปิดไม่สามารถควบคุมเพื่อให้ได้ภาพที่เป็นหน้าตรง (Frontal Face) เท่านั้นและบางครั้งคนร้ายเองก็อาจจะรู้ว่ามิกกล้องวงจรปิด (CCTV Camera) ที่บันทึกภาพอยู่ที่พยายามหันหน้าไปทางอื่นเพื่อไม่ให้เห็นหน้าชัดเจนแบบตรงๆ โดยให้เห็นหน้าในส่วนด้านข้างซึ่งอาจจะเป็นส่วนซ้ายหรือด้านขวาอาจจะก้มหน้าลงก็ได้ทำให้การสร้างคืนภาพไม่มีประสิทธิภาพเท่าที่ควร

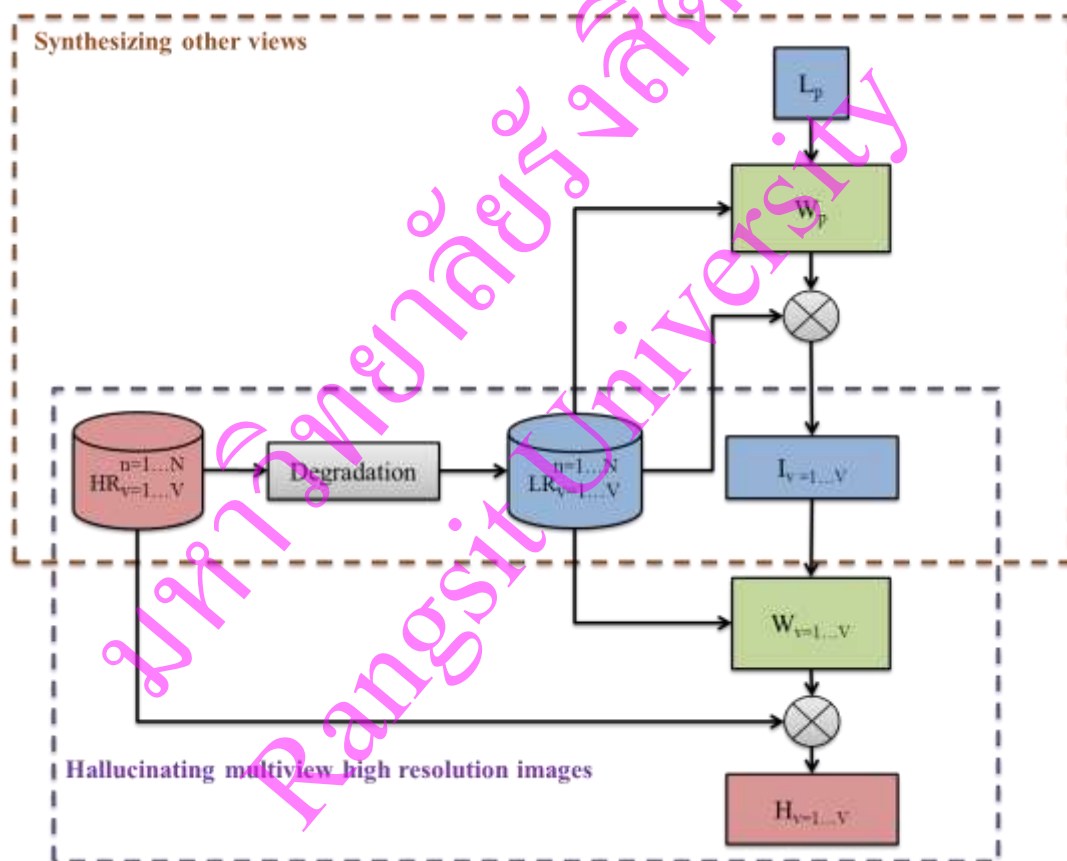
ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงใช้เทคนิคการสร้างคืนภาพกับภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination) เช่นภาพหน้าตรงภาพใบหน้าที่ด้านซ้ายภาพใบหน้าที่ด้านขวาภาพก้มหน้าหรือภาพเงาหน้าซึ่งต้องนำภาพที่บันทึกได้ผ่านกระบวนการการสร้างคืนภาพ (Reconstruction) ให้ได้ภาพที่มีความละเอียดสูงซึ่งจะสามารถนำไปใช้งานเพื่อสืบหาและติดตามคนร้ายมาดำเนินคดีต่อไปได้

1.2 วัตถุประสงค์

1.2.1 เพื่อเสนอวิธีการในการสร้างภาพใบหน้ามนุษย์ที่มีความละเอียดสูงจากภาพใบหน้าที่มีความละเอียดต่ำโดยตรงรับภาพใบหน้าในหลายมุมมอง

1.2.2 เพื่อสร้างฐานข้อมูลภาพใบหน้าของคนไทย

1.3 กรอบแนวคิด



รูปที่ 1.2 กรอบแนวคิดในการทำวิจัย

1.3.1 นำเข้าภาพใบหน้าที่มีหลายมุมมองซึ่งมีความละเอียดสูง $HR_{v=1...V}^{n=1...N}$

1.3.2 นำภาพที่นำเข้ามาผ่านกระบวนการ (Degradation) ทำให้เป็นภาพที่มีความละเอียดต่ำ $LR_{v=1...V}^{n=1...N}$

1.3.3 นำเข้าภาพใบหน้าที่ความละเอียดต่ำ 1 มุมมอง L_p

- 1.3.4 คำนวณหาค่าเฉลี่ยน้ำหนักของภาพ 1 ที่นำเข้า p
- 1.3.5 สร้างคืนภาพความละเอียดต่ำในมุมมองต่างๆขึ้นมา $I_{v=1\dots V}$
- 1.3.6 คำนวณหาค่าเฉลี่ยน้ำหนักของภาพในมุมมองต่างๆ $W_{v=1\dots V}$
- 1.3.7 สร้างคืนภาพความละเอียดสูงของทุกๆมุมมองจากภาพนำเข้า 1 มุมมอง $H_{v=1\dots V}$

1.4 ขอบเขตการศึกษา

- 1.4.1 ภาพที่ใช้ในการทำวิจัยนี้เป็นภาพใบหน้าของมนุษย์
- 1.4.2 จำนวนภาพที่ใช้ฝึกฝนระบบไม่ต่ำกว่า 500 ภาพจากจำนวนคนอย่างน้อย 100 คน โดยแต่ละคนจะมีภาพในมุมมองที่แตกต่างกันจำนวน 5 ภาพ ได้แก่ ภาพหน้าตรง ภาพหันซ้าย ภาพหันขวา ภาพก้มหน้า และ ภาพเงยหน้า
- 1.4.3 ขนาดของภาพที่เล็กที่สุดมีขนาดเท่ากับ 32×24 จุดภาพ ขนาดของภาพที่ขยายใหญ่ที่สุดมีขนาดเท่ากับ 4 เท่า หรือ 128×96 จุดภาพ
- 1.4.4 พัฒนาวิธีการในการสร้างภาพที่มีความละเอียดสูงจากภาพที่มีความละเอียดต่ำโดยใช้ภาพใบหน้าหลายๆมุมมองจากฐานข้อมูลภาพมาตรฐาน
- 1.4.5 สร้างฐานข้อมูลภาพใบหน้าของคนไทยไม่ต่ำกว่า 1,000 คนคนละ 15 มุมมอง

1.5 นิยามศัพท์เฉพาะ

- 1.5.1 Super Resolution Reconstruction (SRR) หมายถึงการสร้างคืนภาพความละเอียดสูง (Super Resolution) โดยการนำภาพที่มีความละเอียดต่ำ (Low Resolution) มาผ่านกระบวนการทางการประมวลผลรูปภาพหรือ Image Processing จะได้ภาพที่มีความละเอียดสูง (High Resolution)
- 1.5.2 Face Hallucination หมายถึง ภาพใบหน้าคนจากภาพความละเอียดต่ำภาพเดียวโดยอาศัยการเรียนรู้การฝึกฝนของระบบจากภาพตัวอย่างจำนวนหนึ่ง
- 1.5.3 Multiview หมายถึง ภาพใบหน้าที่มีหลายมุมมอง
- 1.5.4 Locally Linear Embedding (LLE) หมายถึง วิธีการลดมิติแบบไม่เชิงเส้นชนิดหนึ่ง
- 1.5.5 Patch หมายถึง ภาพย่อยซึ่งเป็นส่วนหนึ่งของภาพสมบูรณ์
- 1.5.6 Frontal Face หรือ Canonical Face หมายถึง ภาพใบหน้าที่ตรงของมนุษย์
- 1.5.7 Interpolation Based หมายถึง วิธีการที่ใช้ภาพความละเอียดต่ำภาพเดียว เพื่อสร้างคืน

ภาพความละเอียดสูง โดยใช้วิธีการคำนวณและสร้างจุดภาพ ขึ้นมาใหม่ตามจำนวนจุดภาพที่มีการเปลี่ยนแปลง จากการลดหรือขยายขนาดของภาพ

1.5.8 Learning Based หมายถึง วิธีการสร้างคืนภาพความละเอียดสูงจากการนำเข้าภาพความละเอียดต่ำภาพเดียว โดยจะใช้ภาพที่เก็บอยู่ในฐานข้อมูลมาสร้างคืนภาพ

1.5.8 Similarity หมายถึง ความคล้ายคลึงกัน

1.5.9 Nearest Neighbors หมายถึง จุดที่ใกล้ที่สุด

1.5.10 PSNR (Peak Signal-to-Noise Ratio) หมายถึง ค่าที่ใช้วัดคุณภาพของภาพเป็นค่ามาตรฐานที่นำมาใช้ในการเปรียบเทียบคุณภาพของรูปภาพดิจิทัล

มหาวิทยาลัยรังสิต
Rangsit University

บทที่ 2

แนวคิดและทฤษฎีที่เกี่ยวข้องกับงานวิจัย

การวิจัยครั้งนี้ผู้วิจัยได้ศึกษาทฤษฎี และ งานวิจัยต่างๆที่เกี่ยวกับ Face Hallucination โดยสรุปได้ดังต่อไปนี้

2.1 Super Resolution Reconstruction (SRR)

Super Resolution Reconstruction (SRR) หมายถึง การสร้างคืนภาพความละเอียดสูง (Super Resolution) โดยการนำภาพที่มีความละเอียดต่ำ (Low Resolution) มาผ่านกระบวนการทางการประมวลผลรูปภาพโดยสรุปวิธีการต่างๆ ได้ดังต่อไปนี้

2.1.1 Interpolation - Based

Interpolation Based (Kimmel R., 1999; Yang Li and Xueyin Lin, 2004; Jian Sun and others, 2008; Hua Shen and Shutao Li, 2009) เป็นวิธีการที่ใช้ภาพความละเอียดต่ำภาพเดียวเพื่อสร้างคืนภาพความละเอียดสูง โดยวิธีการ Interpolation คือ การคำนวณและสร้างจุดภาพ (Pixel) ขึ้นมาใหม่ตามจำนวนของจุดภาพที่มีการเปลี่ยนแปลงจากการลดหรือขยายขนาดของภาพ

2.1.1.1 Bicubic Interpolation

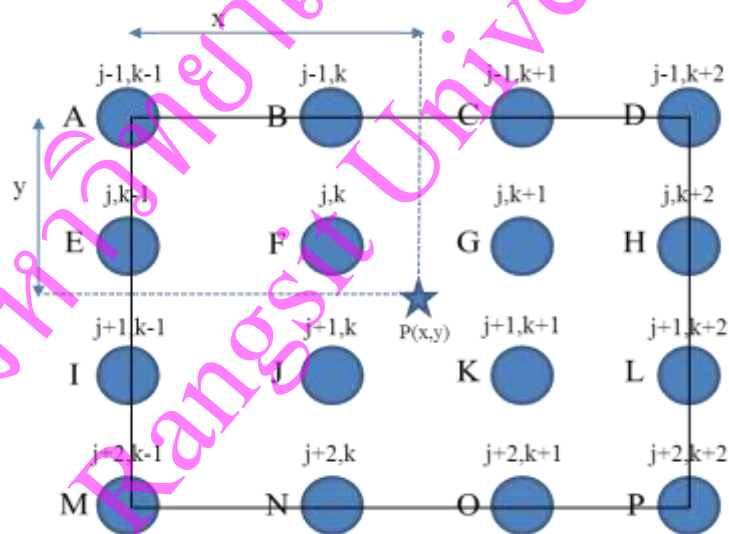
Bicubic Interpolation เป็นการคำนวณและสร้างจุดภาพ (Pixel) ขึ้นมาใหม่ตามจำนวนของจุดภาพที่มีการเปลี่ยนแปลง ซึ่งจะพิจารณาจำนวนจุดภาพที่ล้อมรอบทั้งหมด 16 จุดภาพ (4x4 จุดภาพ) ในขณะที่การคำนวณค่าเฉลี่ยเพื่อสร้างจุดภาพใหม่ โดยจุดภาพที่สร้างขึ้นใหม่จะให้ความนุ่มนวล ราบเรียบขึ้น (Wikipedia, <http://www.wikipedia.co.th>, 4 กุมภาพันธ์ 2555)

ในภาพต้นฉบับ จุดภาพใหม่ของ Bicubic ใช้ 16 จุดภาพ ในพื้นที่ใกล้เคียงกันของจุดภาพขนาด 4x4 จุดภาพ ซึ่งเป็นวิธีการที่ใช้มากที่สุดในโปรแกรมการแก้ไขภาพ และ กล้องดิจิทัลจำนวน

มากสำหรับการการคำนวณและสร้างจุดภาพ (Pixel) ขึ้นมาใหม่ตามจำนวนของจุดภาพที่มีการเปลี่ยนแปลง ตัวอย่าง โปรแกรม Adobe Photoshop CS มี 2 วิธีในการ Bicubic Interpolation คือ Bicubic Smoother เหนี่ยค่าสีจากจุดภาพข้างเคียงโดยรอบ แล้วคำนวณหาค่าสี ซึ่งจะควบคุมภาพให้มีความนุ่มนวลมากขึ้นกว่าเดิม และ Bicubic Sharper เหนี่ยค่าสีจากจุดภาพข้างเคียงโดยรอบ แล้วคำนวณหาค่าสี ซึ่งจะพยายามรักษาความคมชัดของภาพไว้ เหมาะสำหรับการย่อขนาดภาพให้เล็กลง (Zhang Xiang-guang, 2008)

Bicubic Interpolation ช่วยเพิ่มรูปแบบของฟังก์ชันความสว่างโดยการประมาณค่า วิธีการแบบดั้งเดิมจะใช้ 16 จุดภาพสำหรับวิธีการ Interpolation (Nuno-Maganda and Arias-Estrada, 2005) จากรูปที่ 2.1 ได้แสดงบริเวณใกล้เคียงของ Bicubic Interpolation

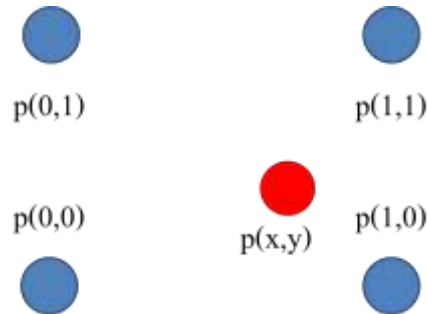
A, B, C, D, E, F, G, H, I, J, K, L, N, M, O, P แทนจุดภาพซึ่งมีขนาด 4×4 จุดภาพ (x,y) เป็นตัวแปรที่แสดงพิกัดของจุดภาพที่จะสร้างจุดภาพ (Pixel) ขึ้นมาใหม่



รูปที่ 2.1 Bicubic-Interpolation Neighborhood

ที่มา: Nuno-Maganda and Arias-Estrada, 2005

ตัวอย่าง การคำนวณค่าความเข้มแสงของจุดภาพโดยวิธี Bicubic Interpolation



รูปที่ 2.2 ตัวอย่างภาพค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x,y)

$p(x,y)$ คือ ค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x,y)
 (x,y) เป็นตัวแปรที่แสดงพิกัดของจุดภาพที่จะสร้างจุดภาพ (Pixel) ขึ้นมาใหม่
 การประมาณหา $p(x,y)$ ทำโดยวิธี Bicubic Interpolation ซึ่งมีนิยามดังนี้

$$p(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (2.1)$$

จาก (2.1) ค่าสัมประสิทธิ์ (a_{ij}) มีทั้งหมด 16 ค่าโดยสามารถคำนวณได้จาก 16 สมการดังนี้

$$f(0,0) = p(0,0) = a_{00} \quad (2.2)$$

$$f(1,0) = p(1,0) = a_{00} + a_{10} + a_{20} + a_{30} \quad (2.3)$$

$$f(0,1) = p(0,1) = a_{00} + a_{01} + a_{02} + a_{03} \quad (2.4)$$

$$f(1,1) = p(1,1) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} \quad (2.5)$$

$$f_x(0,0) = p_x(0,0) = a_{10} \quad (2.6)$$

$$f_x(1,0) = p_x(1,0) = a_{10} + 2a_{20} + 3a_{30} \quad (2.7)$$

$$f_x(0,1) = p_x(0,1) = a_{10} + a_{11} + a_{12} + a_{13} \quad (2.8)$$

$$f_x(1,1) = p_x(1,1) = \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} i \quad (2.9)$$

$$f_y(0,0) = p_y(0,0) = a_{01} \quad (2.10)$$

$$f_y(1,0) = p_y(1,0) = a_{01} + a_{11} + a_{21} + a_{31} \quad (2.11)$$

$$f_y(0,1) = p_y(0,1) = a_{01} + 2a_{02} + 3a_{03} \quad (2.12)$$

$$f_y(1,1) = p_y(1,1) = \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} j \quad (2.13)$$

$$f_{xy}(0,0) = p_{xy}(0,0) = a_{11} \quad (2.14)$$

$$f_{xy}(1,0) = p_{xy}(1,0) = a_{11} + 2a_{21} + 3a_{31} \quad (2.15)$$

$$f_{xy}(0,1) = p_{xy}(0,1) = a_{11} + 2a_{12} + 3a_{13} \quad (2.16)$$

$$f_{xy}(1,1) = p_{xy}(1,1) = \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} ij \quad (2.17)$$

โดยค่าของ p_x , p_y และ p_{xy} คำนวณได้จากสมการ

$$p_x(x, y) = \sum_{i=1}^3 \sum_{j=0}^3 a_{ij} i x^{i-1} y^j \quad (2.18)$$

$$p_y(x, y) = \sum_{i=0}^3 \sum_{j=1}^3 a_{ij} x^i j y^{j-1} \quad (2.19)$$

$$p_{xy}(x, y) = \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} ix^{i-1} jy^{j-1} \quad (2.20)$$

จากสมการที่ (2.2) ถึง (2.17) ค่าสัมประสิทธิ์ที่คำนวณได้จากระบบสมการต่อไปนี้

$$\begin{bmatrix} a_{00} \\ a_{10} \\ a_{20} \\ a_{30} \\ a_{01} \\ a_{11} \\ a_{21} \\ a_{31} \\ a_{02} \\ a_{12} \\ a_{22} \\ a_{32} \\ a_{03} \\ a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 \\ -3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 & 0 & 3 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 & 0 \\ 9 & -9 & -9 & 9 & 6 & 3 & -6 & -3 & 6 & -6 & 3 & -3 & 4 & 2 & 2 & 1 \\ -6 & 6 & 6 & -6 & -3 & -3 & 3 & 3 & -4 & 4 & -2 & 2 & -2 & -2 & -1 & -1 \\ 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ -6 & 6 & 6 & -6 & -4 & -2 & 4 & 2 & -3 & 3 & -3 & 3 & -2 & -1 & -2 & -1 \\ 4 & -4 & -4 & 4 & 2 & 2 & -2 & -2 & 2 & -2 & 2 & -2 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} f(0,0) \\ f(1,0) \\ f(0,1) \\ f(1,1) \\ f_x(0,0) \\ f_x(1,0) \\ f_x(0,1) \\ f_x(1,1) \\ f_y(0,0) \\ f_y(1,0) \\ f_y(0,1) \\ f_y(1,1) \\ f_{xy}(0,0) \\ f_{xy}(1,0) \\ f_{xy}(0,1) \\ f_{xy}(1,1) \end{bmatrix} \quad (2.21)$$



รูปที่ 2.3 ตัวอย่างภาพใบหน้าที่สร้างโดยวิธี Bicubic Interpolation

2.1.2 Reconstruction Based

เป็นวิธีการสร้างคืนภาพความละเอียดสูงโดยใช้ภาพเดียวกันของวัตถุหลายๆภาพ มาสร้างคืนภาพความละเอียดสูง โดยวิธีการนี้มีความเหมาะสมกับลักษณะการทำงานที่มีอุปกรณ์สำหรับการถ่ายภาพ เช่น ภาพถ่ายจากดาวเทียม เป็นต้น



รูปที่ 2.4 ตัวอย่างรูปที่ผ่านวิธีการ Reconstruction Based



รูปที่ 2.5 ตัวอย่างรูปที่ผ่านวิธีการ Reconstruction Based

2.1.3 Learning Based Super-Resolution Method

เป็นวิธีการที่ใช้ภาพฝึกฝนความละเอียดต่ำจำนวนหลายภาพเพื่อสร้างคืนภาพความละเอียดสูง ซึ่งภาพที่ได้ไม่จำเป็นต้องอยู่ในภาพชุดเดียวกัน (S. Baker, T. Kanada, 2000; CeLui and others, 2001; S. Baker, T. Kanada, 2002; Wang, 2005)

2.1.2.1 A Two-Step Approach

ขั้นตอนวิธี Two-Step เป็นการพัฒนาเพื่อการสร้างคืนภาพความละเอียดสูง (Super Resolution) จากภาพใบหน้าคนจากภาพความละเอียดต่ำ ภาพเดียว โดยอาศัยการเรียนรู้ การฝึกฝนของระบบ จากภาพตัวอย่างจำนวนหนึ่ง

ขั้นตอนที่ 1 ประยุกต์ PCA ใช้ในแบบจำลอง ภาพความละเอียดสูง และ ใช้วิธีการ Maximum a Posteriori (MAP) ประมาณค่าเพื่อสร้างคืนภาพโดยรวม จากนั้นสร้าง Nonparametric Markov Network ระหว่างภาพที่เหลืออยู่ และภาพโดยรวม เพื่อประมาณค่าภาพที่เหลืออยู่กับตัวอย่างภาพย่อย (CeLiu and others 2001)

โดย \mathbf{I} หมายถึง ภาพความละเอียดสูง \mathbf{h} หมายถึง ภาพความละเอียดต่ำ

l จำนวนจากสมการ

$$\mathbf{I}(a, b) = \frac{1}{s^2} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} \mathbf{h}(sa + i, sb + j) \quad (2.22)$$

โดย s หมายถึง จำนวนเต็ม

n หมายถึง จุดที่ไม่พึงประสงค์ที่เกิดขึ้นในภาพ

โดยกำหนดเงื่อนไข

\mathbf{h} คือ $\mathbf{n} - \mathbf{d}$

l คือ $\mathbf{m} - \mathbf{d}$

n คือ $\mathbf{m} - \mathbf{d}$ เวกเตอร์ ($\mathbf{m} = \mathbf{n} / s^2$)

สมการที่ (2.4) สามารถเขียนได้ดังนี้

$$\mathbf{I} = \mathbf{Dh} + \boldsymbol{\eta} \quad (2.23)$$

ค่าของ \mathbf{I} ตามเกณฑ์ของ MAP ค่าของ \mathbf{h} ที่เหมาะสมที่สุดสามารถหาได้โดย ค่าสูงสุดของความน่าจะเป็นจาก $P(\mathbf{h} / \mathbf{I})$

$$\mathbf{h} = \arg \max_{\mathbf{h}} P(\mathbf{I} / \mathbf{h}) P(\mathbf{h}) \quad (2.24)$$

Down-Sampling Model

$$\mathbf{h} = \mathbf{h}^g + \mathbf{h}^l \quad (2.25)$$

ภาพใบหน้า

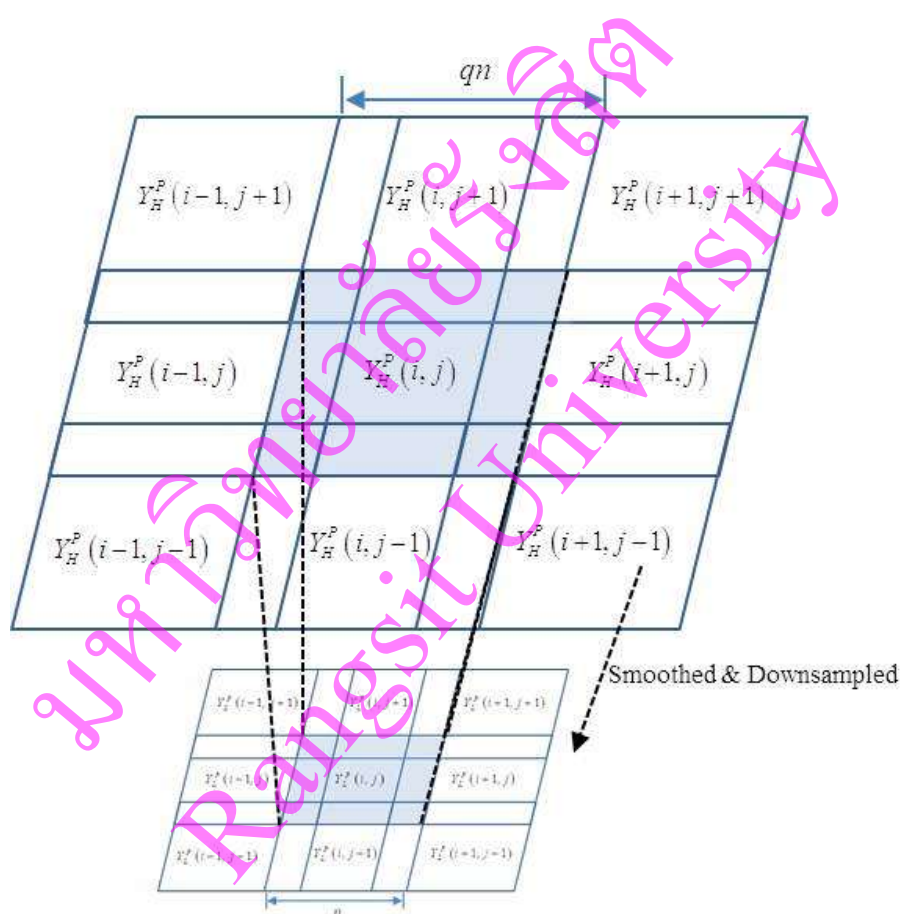
\mathbf{h}^g คือ ภาพโดยรวม (Global Image)

\mathbf{h}^l คือ ภาพที่เหลือที่มีการซูมตัวอย่างจากภาพย่อย

มหาวิทยาลัยรังสิต
Rangsit University

2.2 Patch Based

การตัดภาพภาพย่อยซึ่งเป็นส่วนหนึ่งของภาพสมบูรณ์ (X. Ma and others, 2009) เป็นวิธีการตัดภาพใบหน้าเป็นแผ่นปะ เพื่อที่จะใช้ในการเปรียบเทียบกับแผ่นปะที่มีความคล้ายกันในฐานข้อมูล แล้วจึงนำแผ่นปะแต่ละแผ่นมาต่อกันเป็นภาพใบหน้า แต่การวิธีการนำแผ่นปะมาต่อกันนั้นจะเกิดปัญหามีรอยต่อระหว่างแผ่นปะ ดังนั้นจึงใช้วิธีการทำ Overlap Patch เพื่อแก้ปัญหาดังกล่าว



รูปที่ 2.6 ภาพความสัมพันธ์ระหว่างภาพความละเอียดสูงและภาพความละเอียดต่ำ
โดยใช้ตำแหน่ง (i, j)

ที่มา: Xiang Ma and others, 2010 : 2226



รูปที่ 2.7 ภาพใบหน้า Patch Matrix

จากรูปที่ 2.3 กำหนดภาพใบหน้า เป็นแบบ Patch Matrix $\{Y^p(i, j)\}_{p=1}^N$ โดย N หมายถึง จำนวนของ Patch ในภาพ Y

จากรูปที่ 2.4 $Y^p(i, j)$ หมายถึง Patch Matrix โดยตำแหน่งของ Row มีค่าเท่ากับ i ตำแหน่งของ Column มีค่าเท่ากับ j

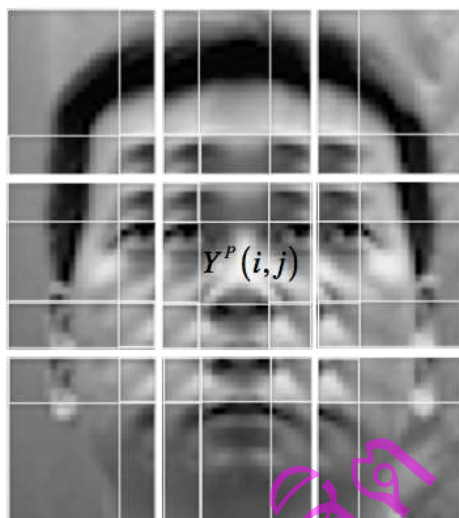
จากรูปที่ 2.5 ลักษณะของ Patch จะเป็นสี่เหลี่ยมมีขนาดเท่ากับ $n \times n$ จุดภาพ

กรณีที่เป็นภาพความละเอียดต่ำ (Low-Resolution Image) Patch Matrix มีค่าเท่ากับ $\{Y_L^p(i, j)\}_{p=1}^N$ โดย ถ้า n มีค่าเป็นจำนวนคู่ Patch ตำแหน่งที่ $Y_L^p(i, j)$ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $n \times [(n-1)/2]$

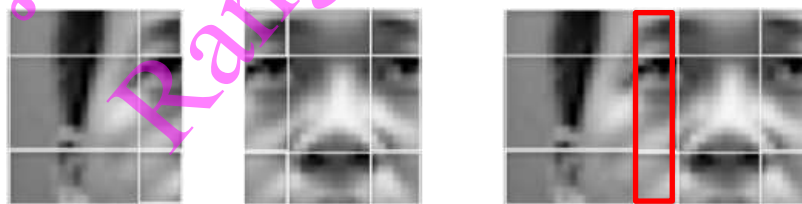
กรณีที่เป็นภาพความละเอียดสูง $Y_H^p(i, j)$ ลักษณะของ Patch จะเป็นสี่เหลี่ยมมีขนาดเท่ากับ $qn \times qn$ จุดภาพ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $(qn) \times [q(n-1)/2]$

โดย ถ้า n มีค่าเป็นจำนวนคู่ Patch ตำแหน่งที่ $Y_L^p(i, j)$ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $n \times (n/2)$ ถ้าเป็นภาพความละเอียดสูง $Y_H^p(i, j)$ จะมีค่าขนาดเท่ากับ $qn \times qn$ จุดภาพ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $(qn) \times (qn/2)$ จากรูปที่ 2.6

ตัวอย่าง กำหนดขนาดของ Patch เท่ากับ 25×25 จุดภาพ กรณีที่เป็นภาพความละเอียดต่ำ สามารถคำนวณการซ้อนทับบางส่วนกับแผ่นปะ (Overlap patch) จากสมการ $n \times [(n-1)/2]$ ซึ่งมีค่าเท่ากับ $25 \times (25-1)/2$ โดยขนาดของแผ่นปะที่ซ้อนทับบางส่วนจะมีขนาด 25×12 จุดภาพ กรณีเป็นภาพความละเอียดสูง กำหนดให้ q เท่ากับ 4 สามารถคำนวณจากสมการ $(qn) \times [q(n-1)/2]$ ซึ่งมีค่าเท่ากับ $(4 \times 25) \times [4 \times (25-1)/2]$ โดยขนาดของแผ่นปะที่ซ้อนทับบางส่วนจะมีขนาด 100×48 จุดภาพ



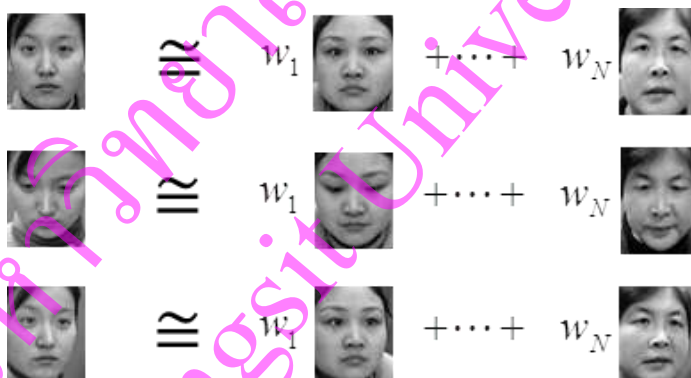
รูปที่ 2.8 Patch Matrix

รูปที่ 2.9 Patch ขนาด $n \times n$ จุดภาพ

รูปที่ 2.10 Overlap Patch

2.3 Linear Combination

Linear Combination เป็นวิธีการสร้างคืนภาพความละเอียดสูงจากการนำเข้าภาพความละเอียดต่ำภาพเดียว โดยที่จะใช้ภาพที่เก็บอยู่ในฐานข้อมูลมาสร้างคืนภาพ วิธีการนี้จะเหมาะสมกับการทำงานกับภาพใบหน้าเพราะ ภาพใบหน้าของเราจะมีส่วนที่มีความคล้ายกัน เช่น คิ้ว ปาก หรือ ตา เป็นต้น แต่การที่เราจะมีภาพใบหน้าทั้งหน้าที่มีความคล้ายกันเป็นเรื่องยาก เนื่องจากมีความคล้ายในแต่ละส่วนไม่เท่ากัน เช่น คิ้วอาจจะมีความคล้ายกับคิ้วของคนอื่น 5 คนซึ่งสัดส่วนความคล้ายมีความแตกต่างกัน คนที่ 1 อาจจะมีคล้าย 10% คนที่ 2 อาจจะมีคล้าย 20% คนที่ 3 อาจจะมีคล้าย 25% คนที่ 4 อาจจะมีคล้าย 30% และ คนที่ 5 อาจจะมีคล้าย 15% ซึ่งจะนำมาสร้างเป็นภาพคิ้วภาพใหม่ หรือ ตัวอย่าง วิธีการสกัดใบหน้าคนร้าย หรือโปรแกรมที่ใช้ในการสกัดภาพก็จะมีกระบวนการแยกส่วนต่างๆของใบหน้าในการเปรียบเทียบความคล้าย เช่น จมูก ปาก คาง หรือ ตา เป็นต้น แล้วจึงนำแต่ละส่วนที่มีความคล้ายมาสร้างเป็นภาพใหม่



รูปที่ 2.11 Linear Combination for Multiview Face

ภาพใบหน้าจะแสดงเป็นคอลัมน์เวกเตอร์ของจำนวนจุดภาพทั้งหมด ขึ้นอยู่กับความคล้ายคลึงกันของโครงสร้าง โดยรูปใบหน้าสามารถสังเคราะห์โดยใช้การรวมกันแบบเชิงเส้นของภาพที่ใช้ในการฝึกฝนให้กับระบบ ภาพใบหน้าในมุมมองต่างๆ ยังสามารถสร้างขึ้นมาใหม่โดยใช้การรวมกันแบบเชิงเส้นของภาพใบหน้าอื่น ๆ ในมุมมองเดียวกัน จากสมการ

$$\mathbf{I}_p = \mathbf{W}_p \mathbf{L}_p \quad (2.26)$$

โดย

- I_p คือ ภาพใบหน้าความละเอียดต่ำที่มุมมอง p
- W_p คือ ค่าความสำคัญของจุดภาพของมุมมอง p
- L_p คือ ภาพที่ใช้ในการฝึกฝนให้ระบบที่มุมมอง p

ภาพใบหน้าในมุมมองอื่นๆ สามารถคำนวณได้จากสมการ

$$I_o \approx W_p L_o \quad (2.27)$$

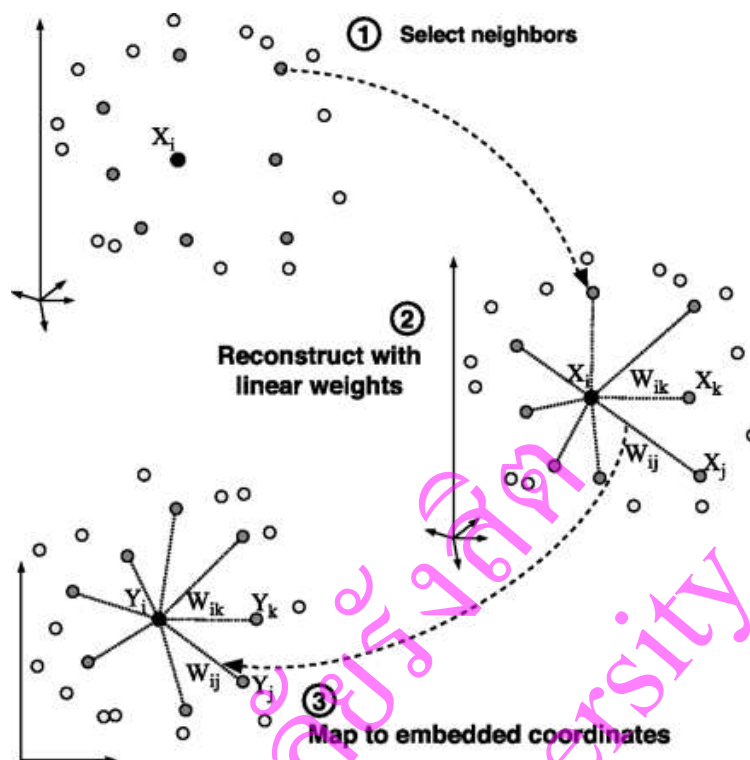
โดย

- I_o คือ ภาพใบหน้าความละเอียดต่ำที่มุมมองอื่นๆ
- L_o คือ ภาพที่ใช้ในการฝึกฝนให้ระบบที่มุมมองอื่นๆ

2.4 Locally Linear Embedding (LLE)

Locally Linear Embedding (LLE) (S. T. Roweis and L. K. Saul, 2000) วิธีการลดมิติแบบไม่เชิงเส้น โดยวิธีการนี้มีความคิดพื้นฐานดังนี้

- 1) เรียนรู้ลักษณะเชิงเส้นตรงเฉพาะที่ที่มีความสัมพันธ์ โดยข้อมูลเฉพาะที่ (Local Data)
- 2) ข้อมูลเฉพาะที่ (Local Data) มีลักษณะเป็นเชิงเส้นตรง เพราะ แผ่นปะ (Patch) ที่ไม่เป็นเชิงเส้นตรง (Nonlinear) จะมีลักษณะเป็นระนาบ (Plane) เมื่อมีขนาดเล็ก
- 3) ข้อมูลทั่วไป (Global Data) ไม่เป็นเชิงเส้นตรง เพราะเป็นรูปโค้ง (Curve)
- 4) ข้อมูลจะมีความสัมพันธ์กันแม้ว่าจะอยู่คนละปริภูมิ (Space)



รูปที่ 2.12 Locally Linear Embedding (LLE)

ที่มา: Sam T. Roweis and Lawrence K. Saul, 2000:2324

2.4.1 ขั้นตอนของวิธีการ Locally Linear Embedding (LLE)

2.4.1.1 คำนวณหาระยะห่างที่ใกล้ที่สุดจากจุด Data Point (x_i) ถึงจุดที่สนใจ (x_k) โดยใช้วิธีการ Euclidean Distance

2.4.1.2 คำนวณหาค่าน้ำหนัก (Weight) ที่ดีที่สุดของแต่ละ Data Point จากจุดที่สนใจ

เลือก w_{ij} ที่ทำให้ f จากสมการ (2.28) มีค่าน้อยที่สุด

$$f = \left\| \mathbf{x}_i - \sum_{j=1}^N w_{ij} \mathbf{x}_j \right\|^2 = \left\| \sum_{j=1}^N w_{ij} (\mathbf{x}_i - \mathbf{x}_j) \right\|^2 = \sum_{j=1}^N \sum_{k=1}^N w_{ij} w_{ik} C_{jk} \quad (2.28)$$

โดย N คือ จำนวนตัวอย่างทั้งหมด

\mathcal{N} คือ เซตของดัชนีของตัวอย่างที่อยู่ใกล้จุดที่สนใจ

K คือ จำนวนตัวอย่างที่อยู่ใกล้จุดที่สนใจ ($K = |\mathcal{N}|$) โดย $K < N$

คำนวณค่าความแปรปรวนร่วมจากสมการ

$$C_{jk} = (\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{x}_i - \mathbf{x}_k) \quad (2.29)$$

โดยมีเงื่อนไขคือ ผลรวมของค่าน้ำหนักของตัวอย่างที่อยู่ใกล้จุดที่สนใจจะมีค่าเท่ากับ 1 นอกจากนี้ค่าน้ำหนักของตัวอย่างอื่นๆ จะมีค่าเป็น 0

$$\sum_{j \in \mathcal{N}} W_{ij} = 1 \quad (2.30)$$

และ

$$W_{ij} = 0, j \notin \mathcal{N} \quad (2.31)$$

โดยการ Normalize ค่าน้ำหนักด้วย $Q = \sum_{j=1}^N \sum_{k=1}^N C_{jk}^{-1}$ จากสมการ

$$W_{ij} = \frac{\sum_{k \in \mathcal{N}} C_{jk}^{-1}}{Q} \quad (2.32)$$

นำค่าความแปรปรวนร่วมที่คำนวณได้มาใช้แก้ระบบเชิงเส้น (Solve Linear System) จากสมการ

$$\mathbf{C}\mathbf{W} = \mathbf{1} \quad (2.33)$$

คำนวณหาค่า \mathbf{W} จากสมการ

$$\mathbf{W} = \mathbf{C}^{-1}\mathbf{1} \quad (2.34)$$

2.4.1.3 คำนวณ Low Dimensional Vector ที่ดีที่สุดในการสร้างคืนภาพจาก
 ค่าน้ำหนัก w_{ij} ที่คำนวณได้มาใช้ใน Subspace y โดย y_i คือ จุดที่สนใจ (Data Point) ในอีกปริภูมิ
 (Space)

$$g = \sum_{i=1}^K \left\| \mathbf{y}_i - \sum_{j=1}^N w_{ij} \mathbf{y}_j \right\|^2 \quad (2.35)$$

2.5 งานวิจัยที่เกี่ยวข้อง

ปัจจุบันนี้ กล้องวิดีโอวงจรปิด ได้รับความนิยมนำมาใช้กันอย่างแพร่หลายในหลายสถานที่
 เช่น ธนาคาร ร้านค้า และลานจอดรถ ความปลอดภัยเป็นสิ่งที่สำคัญมาก รายละเอียดของภาพใบหน้า
 ที่ได้รับจากกล้องวิดีโอ มีความสำคัญสำหรับการระบุตัวบุคคล แต่ในหลายกรณี ภาพที่ได้จาก กล้อง
 วงจรปิดไม่สามารถระบุได้ เนื่องจาก ความละเอียดของภาพใบหน้าต่ำ ทำให้ใบหน้าบางส่วนขาด
 ความคมชัด ดังนั้น เพื่อให้ได้ใบหน้าที่มีรายละเอียด เพื่อการยืนยัน ระบุตัวตนได้ โดยอาศัยภาพ
 ความละเอียดต่ำ (LR) ของภาพใบหน้าเพื่อใช้ในการสร้างคืนภาพความละเอียดสูง (HR) โดยใช้
 เทคนิคที่เรียกว่า “Face Hallucination” หรือ “Face Super-Resolution” (X.G. Wang and X.O. Tang,
 2005) โดยเทคนิคดังกล่าวได้นำมาใช้ในเรื่องที่สำคัญ เช่น การถ่ายภาพทางการแพทย์ การถ่ายภาพ
 ดาวเทียม ระบบกล้องวงจร และ การฟื้นฟูความชัดขึ้นของภาพประวัติศาสตร์เก่า วิธีการ Super-
 Resolution เป็นเทคนิคในการสร้างคืนภาพความละเอียดสูง (HR) ภาพเดียวหรือจำนวนหลายภาพ
 จากภาพความละเอียดต่ำ (LR) หลายๆภาพ หรือภาพความละเอียดต่ำ (LR) ภาพเดียว

งานวิจัยทางด้าน Super-Resolution Reconstruction (SRR) ได้มีการนำเสนอ 2 วิธี คือ

1. วิธีการ Reconstruction Based 2. วิธีการ Learning Based วิธีการแรกนำภาพความละเอียดต่ำ (LR)
 จำนวนหลายๆภาพซึ่งเป็นภาพวัตถุเดียวกันเป็น ภาพนำเข้าเพื่อสร้างคืนภาพความละเอียดสูง (HR)
 ในขณะที่ วิธีการ Learning Based จะเป็นวิธีการสร้างคืนภาพจากการเรียนรู้ของระบบ โดยใช้
 จำนวนของข้อมูลตัวอย่างเพื่อสร้างคืนภาพ ข้อดีของวิธีการ Learning Based คือ การสร้างภาพบุคคล
 จากภาพความละเอียดต่ำภาพเดียว ซึ่งเป็นเทคนิคที่เรียกว่า “Face Hallucination” (Elad, M. and A.
 Feuer, 1999; S. Baker and T. Kanade, 2000;2003; Y. Hu and others, 2011; Gao, G. and J. Yang, 2014)
 จะเน้นสำหรับการใช้กับภาพใบหน้าของมนุษย์ จำนวนและวิธีการที่เกี่ยวข้องกับงานวิจัยทางด้านนี้
 ได้รับความสนใจมากขึ้น และ ให้ผลการทดลองที่มีประสิทธิภาพ เมื่อเทียบกับวิธีการอื่นๆ

ในปี 2009 นักวิจัย Xiang Ma ได้เสนอวิธีการสร้างคืนภาพความละเอียดสูง โดยใช้ตำแหน่งของแผ่นปะ (Position Patch)(X. Ma and others, 2009; 2010) เพื่อปรับปรุงคุณภาพของภาพ มีงานวิจัยที่นำเอาวิธีการนี้ไปใช้ในการสร้างคืนภาพใบหน้าความละเอียดสูง (K. Jia, S. G. Gong, 2008; Zhuang, Y. and others, 2007; Wei, L. and others, 2005; Park Sung Won and M. Savvides, 2007; Fan Wei and Yeung Dit-Yan, 2005; 2007) โดยในกระบวนการในการสร้างคืนภาพใบหน้านั้น สร้างจากชุดข้อมูลภาพความละเอียดสูง และชุดข้อมูลภาพความละเอียดต่ำ ภาพความละเอียดสูง ที่ได้จากวิธีที่เสนอ ดังกล่าว ถูกสร้างขึ้นจากตำแหน่งเดียวกันของแผ่นปะ (Patch) ของแต่ละภาพที่ใช้ในการฝึกฝนให้ระบบ

ผลจากการทดลองแสดงถึงคุณภาพของการสร้างคืนภาพที่ได้สูงกว่าวิธีการอื่น ได้แก่ Cubic B-spline Wangs Eigen-transformation method (X.G. Wang and X.O. Tang, 2005), Changs Neighbor Embedding (H. Chang and others, 2004) และ Zhuangs Locality Preserving method (Zhuang, Y. and others, 2007) และ มีการพัฒนางานวิจัยโดยได้มีจากนำ bilateral patches (Zhou and others, 2012) มาใช้ในงานวิจัย และ k nearest neighbors (KNN) ได้พัฒนามาใช้ในการในงานวิจัยในด้าน การสร้างคืนภาพใบหน้าด้วย (Yu, H. and others, 2011)

ในปี 2010 Xiang Ma ได้เสนอวิธีการในการสร้างคืนภาพใบหน้าหลายมุมมอง (X. Ma and others, 2010) โดยใช้ตำแหน่งของแผ่นปะ (X. Ma and others, 2009; 2010) ร่วมกับวิธีการ Locally Linear Embedding (LLE) ซึ่งวิธีการนี้ยังพบปัญหาในการสร้างคืนภาพใบหน้า นั่นคือ การกำหนดน้ำหนักที่เหมาะสมที่สุด ซึ่งในทางปฏิบัติสำหรับข้อมูลจริงน้ำหนักดังกล่าวจะกำหนดโดยใช้จำนวนคงที่ของจุดที่ใกล้กับจุดที่สนใจสำหรับทุกจุดไม่ได้เพราะจำนวนของจุดที่ใกล้กับจุดที่สนใจในแต่ละคนไม่เท่ากับคนอื่น ๆ

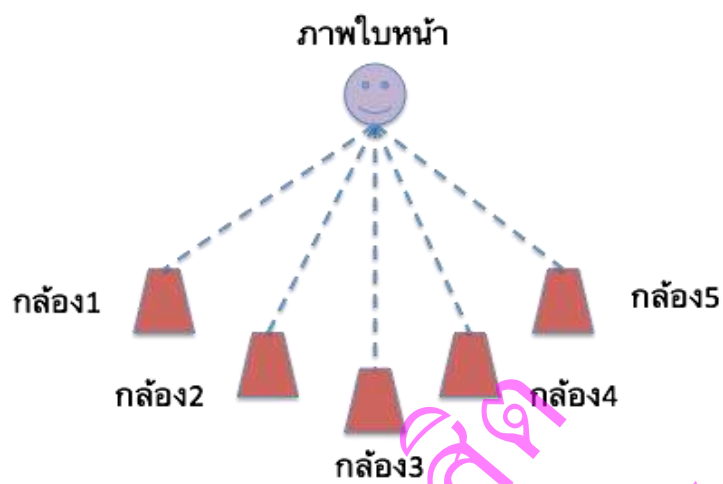
บทที่ 3

วิธีดำเนินงานวิจัย

ในบทนี้จะนำเสนอวิธีดำเนินงานวิจัย โดยกระบวนการในการทำวิจัย ประกอบด้วย ส่วนแรกหัวข้อ ที่ 3.1 จะอธิบายถึง การสร้างฐานข้อมูลภาพใบหน้าของคนไทย ส่วนที่ 2 หัวข้อที่ 3.2 จะเสนอกระบวนการในการทำวิจัย (Research Methodology) ส่วนที่ 3 หัวข้อที่ 3.3 จะอธิบายถึง วิธีการที่นำเสนอในการทำวิจัย (Adaptive Locally Linear Embedding) ส่วนที่ 4 อธิบายขั้นตอนการทำฮาโลซิเนสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination) ในหัวข้อที่ 3.4 และหัวข้อที่ 3.5 อธิบายถึงวิธีการประเมินคุณภาพของภาพ ตามลำดับ

3.1 การสร้างฐานข้อมูลภาพใบหน้าของคนไทย

การสร้างฐานข้อมูลภาพใบหน้าซึ่งเป็นภาพระดับเทา ในมุมมองที่แตกต่างกัน เพื่อใช้เผยแพร่ในงานวิจัยทางวิชาการระดับนานาชาติ ได้แก่ ภาพหน้าตรง ภาพหันซ้าย ภาพหันขวา ภาพก้มหน้า และ ภาพเงยหน้า โดยมีระยะห่างระหว่างกล้องถึงใบหน้า 0.8 เมตร ระยะความสูงของกล้อง 1.1 เมตร โดยแต่ละภาพจะมีองศาที่ต่างกัน 22.5 องศา (W. Gao and others, 2008:150)



รูปที่ 3.1 รูปแบบการตั้งกล้องถ่ายภาพ



รูปที่ 3.2 ตัวอย่างการถ่ายภาพ

ภาพที่ได้จะเป็นภาพใบหน้าของคนไทยโดยแต่ละคนจะมีภาพหน้าตรงในมุมมองต่างๆ 5 ภาพ ภาพก้มหน้าในมุมมองต่างๆ 5 ภาพและภาพเงยหน้าในมุมมองต่างๆ 5 ภาพ โดยแต่ละภาพจะมีขนาด 360x480 จุดภาพ



รูปที่ 3.3 ตัวอย่างภาพใบหน้าในมุมมองต่างๆ

3.2 กระบวนการในการทำวิจัย (Research Methodology)

3.2.1 ข้อมูลภาพที่ใช้ในงานวิจัย

3.2.1.1 ภาพที่ใช้ในงานวิจัยนี้เป็นภาพจากฐานข้อมูลภาพมาตรฐาน (The CAS-PEAL-R1) (W. Gao and others, 2008) โดยจะเป็นภาพระดับเทามีขนาด 360 x 480 จุดภาพ ซึ่งภาพที่ใช้จะเป็นภาพใบหน้าของคนที่มีมุมมองที่ต่างกันออกไป จำนวน 5 ภาพ ได้แก่ ภาพหน้าตรง ภาพหันซ้าย ภาพหันขวา ภาพก้มหน้า และ ภาพเงยหน้า จำนวนทั้งสิ้น 1,011 คน



รูปที่ 3.4 ตัวอย่างภาพจากฐานข้อมูล The CAS-PEAL-R1 ขนาด 360 x 480 จุดภาพ

3.2.2 วิธีการในการสร้างภาพที่มีความละเอียดสูงจากภาพที่มีความละเอียดต่ำ โดยใช้ภาพใบหน้า หลาย มุมมอง จากฐานข้อมูลภาพมาตรฐาน

กรอบแนวคิด (Conceptual Framework) การสร้างคืนภาพความละเอียดสูงที่มีหลายมุมมอง จากรูปที่ 1.1 สามารถอธิบายขั้นตอนการทำงานตามโครงร่างของงาน ได้ดังนี้

3.2.2.1 กระบวนการก่อนหน้า (Preprocessing)

1) การจัดวางตำแหน่งของภาพใบหน้า (Face alignment) เพื่อจะทำให้ตำแหน่งของภาพใบหน้าทุกภาพมีรูปแบบที่เหมาะสมในการประมวลผลต่อไป โดยในงานวิจัยนี้กำหนดตำแหน่งทั้ง 3 ตำแหน่งของภาพใบหน้าให้อยู่ในแนวเดียวกันโดยทำด้วยมือ (Manual) คือ ตำแหน่งกลางตาข้างซ้าย ตำแหน่งกลางตาข้างขวา และ ตำแหน่งกลางปาก (X. a. Q. L. Ma and others, 2012)

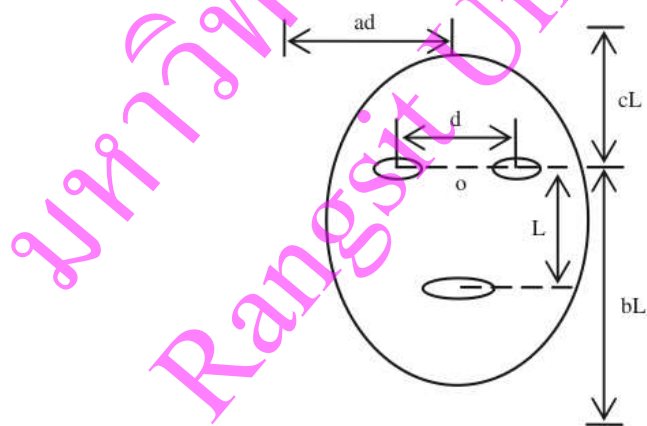
d คือ ระยะห่างระหว่างตำแหน่งกลางดวงตาของตาข้างซ้ายและตาข้างขวา

O คือ จุดกึ่งกลางของระยะห่างระหว่างตำแหน่งกลางดวงตาของตาข้างซ้ายและตาข้างขวา

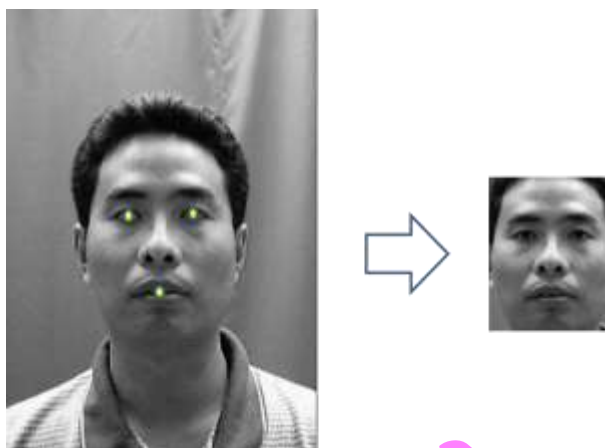
L คือ ระยะห่างจากจุดกึ่งกลางของระยะห่างระหว่างตำแหน่งกลางดวงตาของตาข้างซ้ายและตาข้างขวา ถึงจุดกลางของปาก

a , b และ c เป็นพารามิเตอร์ที่ปรับขนาดภาพใบหน้า

ระยะจาก O ไปทางซ้าย (หรือขวา) ด้านล่างและด้านบนของบริเวณที่น่าสนใจตามลำดับจะแสดงเป็น ad , BL และ CL



รูปที่ 3.5 การจัดวางตำแหน่งของภาพใบหน้า



รูปที่ 3.6 ตัวอย่างกระบวนการจัดวางตำแหน่งภาพใบหน้า



รูปที่ 3.7 ตัวอย่างภาพใบหน้าที่ผ่านกระบวนการจัดวางตำแหน่ง

2) ตัดภาพให้เหลือเฉพาะภาพใบหน้าส่วนที่ให้ความสนใจ และลดขนาดความละเอียดของภาพให้มีค่าเท่ากับ 128x96 จุดภาพ



รูปที่ 3.8 ตัวอย่างภาพใบหน้าที่ผ่านกระบวนการตัดภาพ

3.2.3 ขั้นตอนการสร้างคืนภาพความละเอียดต่ำ (Synthesizing Other View)

3.2.3.1 นำเข้าภาพใบหน้าที่มีหลายมุมมองที่มีความละเอียดขนาด 128x96

จุดภาพ

$$HR_{v=1\dots V}^{n=1\dots N}$$

3.2.3.2 ปรับลดขนาดความละเอียดของภาพให้มีค่าเท่ากับ 32x24 จุดภาพ

$$LR_{v=1\dots V}^{n=1\dots N}$$

จากสมการ

$$LR_v^n = DHR_v^n \quad (3.1)$$

3.2.3.3 นำเข้าภาพใบหน้าที่ความละเอียดต่ำ 1 มุมมอง L_p

3.2.3.4 คำนวณหาค่าน้ำหนักของภาพ (ในข้อ 3.2.3.3) ที่นำเข้า w_p

3.2.3.5 สร้างคืนภาพความละเอียดต่ำในมุมมองต่างๆ ขึ้นมา $I_{v=1\dots V}$

จากสมการ

$$I_v = W_p LR_v^n \quad (3.2)$$

3.2.4 ขั้นตอนการสร้างคืนภาพความละเอียดสูง (Hallucinating Multiview High Resolution Images)

3.2.4.1 คำนวณหาค่าน้ำหนักของภาพในมุมมองต่างๆ $w_{v=1\dots V}$ จากภาพความละเอียดต่ำในมุมมองต่างๆ $I_{v=1\dots V}$

3.2.4.2 สร้างคืนภาพความละเอียดสูงของทุกๆ มุมมองจากภาพนำเข้า 1 มุมมอง

$H_{v=1\dots V}$ จากสมการ

$$I = W_v HR_v^n \quad (3.3)$$

โดย

$HR_{v=1\dots V}^{n=1\dots N}$ คือ ภาพใบหน้าที่มีหลายมุมมองที่มีความละเอียดเท่ากับ 128 x 96 จุดภาพ

$LR_{v=1\dots V}^{n=1\dots N}$ คือ ภาพใบหน้าที่มีหลายมุมมองที่มีความละเอียดเท่ากับ 32 x 24 จุดภาพ

L_p คือ ภาพใบหน้าที่ความละเอียดต่ำ 1 มุมมอง

W_p คือ ค่าน้ำหนักของภาพ L_p

$I_{v=1\dots V}$ คือ ภาพความละเอียดต่ำในมุมมองต่างๆ ที่สร้างขึ้นมา

$W_{v=1\dots V}$ คือ ค่าน้ำหนักของภาพในมุมมองต่างๆ

$H_{v=1\dots V}$ คือ ภาพความละเอียดสูงในมุมมองต่างๆ ที่สร้างขึ้นมา



รูปที่ 3.9 ตัวอย่างภาพนำเข้าความละเอียดสูงขนาด 128 x 96 จุดภาพ



รูปที่ 3.10 ตัวอย่างภาพปรับลดความละเอียดขนาด 32 x 24 จุดภาพ



รูปที่ 3.11 ตัวอย่างภาพนำเข้าความละเอียดต่ำขนาด 32x24 จุดภาพ



รูปที่ 3.12 ตัวอย่างภาพความละเอียดต่ำมุมมองต่างๆที่สร้างขึ้น
จากภาพนำเข้าขนาด 32x24 จุดภาพ



รูปที่ 3.13 ตัวอย่างภาพความละเอียดสูงมุมมองต่างๆที่สร้างขึ้น
ขนาด 128x96 จุดภาพ

3.3 Adaptive Locally Linear Embedding (ALLE)

เนื่องด้วย วิธีการคำนวณหาค่าน้ำหนัก (W) ในงานวิจัยเรื่อง “Hallucinating face by position-patch” (X. Ma and others, 2010) เป็นงานวิจัยที่แสดงกระบวนการในการทำฮาโลซิเนสภาพใบหน้าของมนุษย์โดยการระบุตำแหน่งแผ่นปะของภาพ และคำนวณหาค่าน้ำหนัก (W) ตามหัวข้อ 3.2.3 ด้วยวิธีการ Locally Linear Embedding (LLE) ในการคำนวณหาค่าน้ำหนัก (W) โดยสรุปขั้นตอนและสมการ ดังนี้

กำหนดภาพใบหน้าเป็นแบบ Patch Matrix $\{Y^P(i, j)\}_{p=1}^N$

โดย

N หมายถึง จำนวนของ Patch ในภาพ Y

$Y^P(i, j)$ หมายถึง Patch Matrix โดยตำแหน่งของ Row มีค่าเท่ากับ i ตำแหน่งของ Column มีค่าเท่ากับ j

ลักษณะของ Patch จะเป็นสี่เหลี่ยมมีขนาดเท่ากับ $n \times n$ จุดภาพ

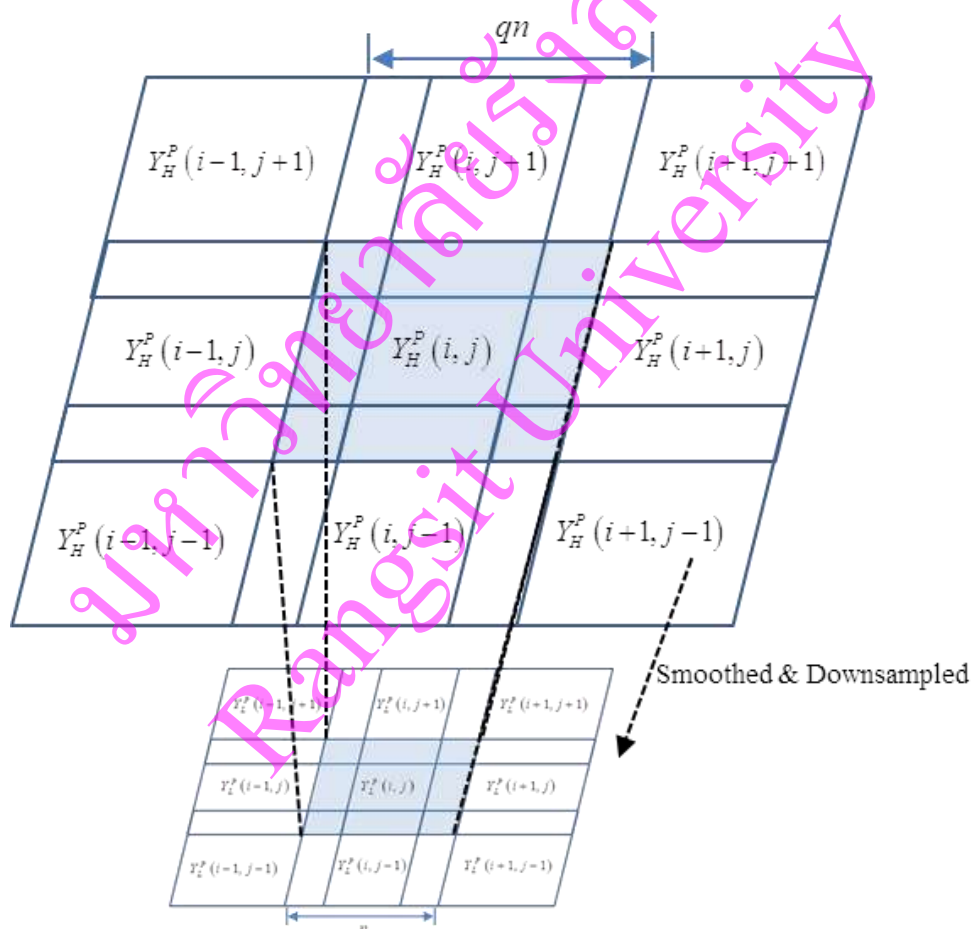
ภาพความละเอียดต่ำ (Low-Resolution Image) Patch Matrix มีค่าเท่ากับ $\{Y_L^P(i, j)\}_{p=1}^N$

โดย

ถ้า n มีค่าเป็นเลขคู่ Patch ตำแหน่งที่ $Y_L^P(i, j)$ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $n \times [(n-1)/2]$

ถ้าเป็นภาพความละเอียดสูง $Y_H^P(i, j)$ จะมีค่าขนาดเท่ากับ $qn \times qn$ จุดภาพ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $(qn) \times [q(n-1)/2]$

ถ้า n มีค่าเป็นเลขคู่ Patch ตำแหน่งที่ $Y_L^P(i, j)$ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $n \times (n/2)$ ถ้าเป็นภาพความละเอียดสูง $Y_H^P(i, j)$ จะมีค่าขนาดเท่ากับ $qn \times qn$ จุดภาพ จะซ้อนทับบางส่วนกับแผ่นปะ โดยมีขนาด $(qn) \times (qn/2)$ ตามรูปที่ 2.6



รูปที่ 3.14 ภาพความสัมพันธ์ระหว่างภาพความละเอียดสูงและ

ภาพความละเอียดต่ำโดยใช้ตำแหน่ง (i, j)

ที่มา: Xiang Ma and others, 2010 : 2226

3.3.1 ขั้นตอนการฮาหลูซิเนสภาพใบหน้า (Face Hallucination Algorithm)

3.3.1.1 นำเข้าภาพใบหน้าความละเอียดต่ำ (Low Resolution Image) ภาพความละเอียดต่ำที่ใช้ในการฝึกฝนระบบ (Low Resolution Training Image) และภาพความละเอียดสูงที่ใช้ในการฝึกฝนระบบ (High Resolution Training Image) ซึ่งเป็นแผ่นปะแบบซ้อนทับบางส่วน (Overlapping Patches) กำหนดโดย

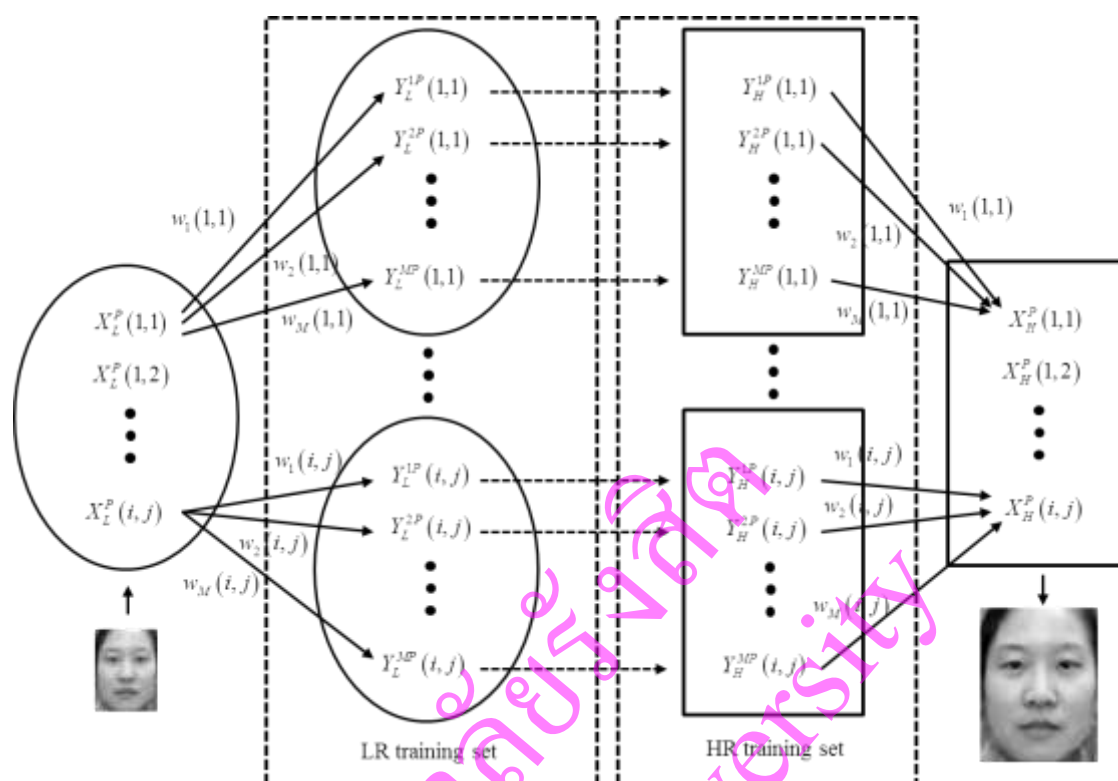
ภาพใบหน้าความละเอียดต่ำ แทนค่า โดย $\{X_L^p(i, j)\}_{p=1}^N$, X แทน ภาพใบหน้าที่ใช้ทดสอบ โดยกำหนดภาพใบหน้าแบบ patch matrix เท่ากับ $\{X^p(i, j)\}_{p=1}^N$ ภาพความละเอียดต่ำที่ใช้ในการฝึกฝนระบบ แทนค่าโดย $\{Y_L^{mp}(i, j)\}_{p=1}^N$ ภาพความละเอียดสูงที่ใช้ในการฝึกฝนระบบ แทนค่าโดย $\{Y_H^{mp}(i, j)\}_{p=1}^N$ $m = 1, 2, \dots, M$ โดย M คือ จำนวนภาพที่ใช้ในการฝึกฝนระบบ

3.3.1.2 ในแต่ละแผ่นปะ $X_L^p(i, j)$

- 1) คำนวณหาค่าน้ำหนัก $w(i, j)$
- 2) รวมแผ่นปะภาพความละเอียดสูง $X_H^p(i, j)$

3.3.1.3 รวมแผ่นปะภาพความละเอียดสูงเป็นภาพใบหน้าความละเอียดสูง

$$\{X_H^p(i, j)\}_{p=1}^N$$



รูปที่ 3.15 Model Position-Patch

ปัญหาที่พบคือ ด้วยวิธีการ Locally Linear Embedding (LLE) ในการคำนวณหาค่า น้ำหนัก (w) ต้องกำหนดจำนวนของ Neighbor ของทุกแผ่นปะ (Patch) ซึ่งเป็นวิธีการที่ไม่เหมาะสม ในการใช้กับข้อมูลจริง เพราะจำนวนของ Neighbor ในแต่ละแผ่นปะ (Patch) มีจำนวนที่ไม่เท่ากัน

งานวิจัยนี้นำเสนอขั้นตอนวิธี Adaptive Locally Linear Embedding (ALLE) สำหรับการ คำนวณหาค่าน้ำหนัก (w) โดยปรับปรุงจากขั้นตอนวิธี LLE เดิมให้สามารถพิจารณาจำนวน Neighbor ที่เกี่ยวข้องในแต่ละจุดตัวอย่างจากค่าขีดเริ่มเปลี่ยนความคล้ายของภาพ จากภาพที่นำเข้า โดยไม่ได้คำนวณจากภาพฝึกฝนทั้งหมด แทนการคำนวณหาค่าน้ำหนักเดิมโดยจะใช้ค่าขีดเริ่ม เปลี่ยนที่เหมือนกันเพื่อเลือก Neighbor ในแต่ละแผ่นปะ (Patch) ซึ่งวิธีการนี้จะบังคับให้ ค่าน้ำหนัก ของ Neighbor เป็น 0 ถ้ามากกว่าค่าขีดเริ่มเปลี่ยนที่กำหนด

Threshold of Similarity (θ) อธิบายโดย

$$\frac{\sum_{i=1}^K d_i}{n} \leq \theta \quad (3.4)$$

กำหนดให้

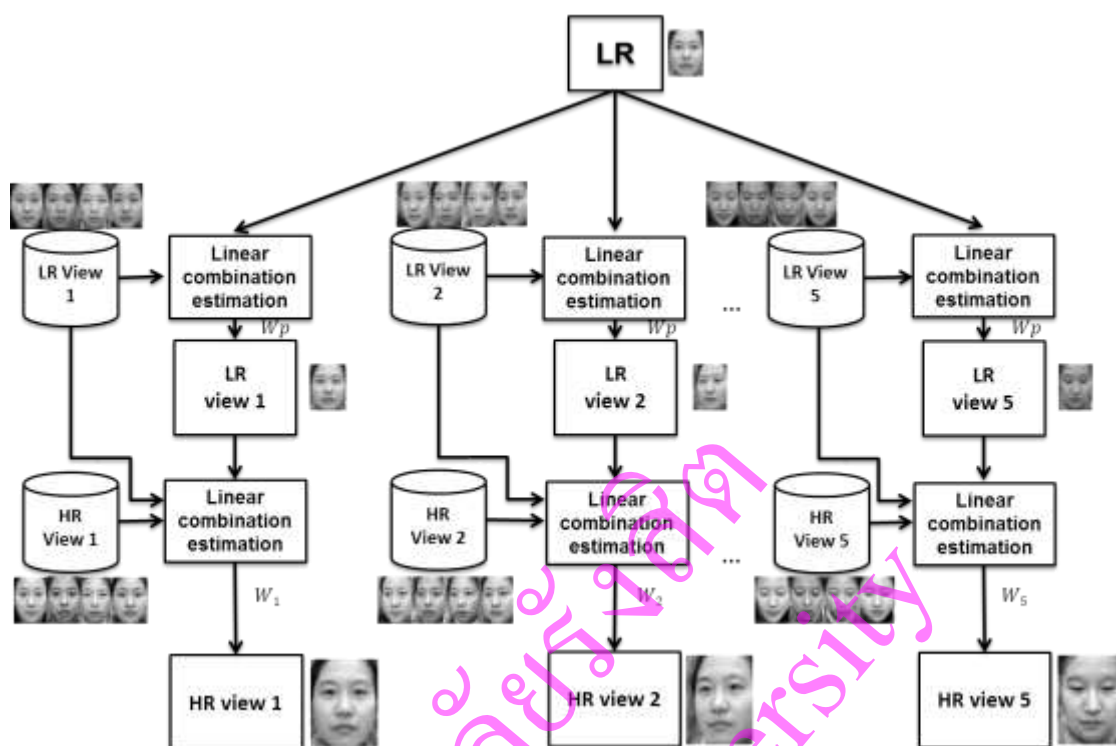
d_i ระยะห่างระหว่าง จากภาพนำเข้าถึงภาพที่ 1 ที่ใช้ในการฝึกฝน

$d_1 \leq d_2 \leq d_3 \leq \dots \leq d_n$ จำนวนของ Nearest Neighbors (K) ไม่ได้เป็นค่าเดียวกันกับ
ภาพนำเข้าตัวอย่างทั้งหมด

3.4 Multiview Face Hallucination by Adaptive Locally Linear Embedding

จากที่ผ่านมา กระบวนการ Face Hallucination เป็นวิธีการสร้างคืนภาพจากภาพความละเอียดต่ำ (Low Resolution) ให้เป็นภาพที่มีความละเอียดสูง (High Resolution) โดยใช้ภาพเพียงภาพเดียวและเป็นภาพหน้าตรงเป็นหลัก แต่ในความเป็นจริงภาพที่ได้จากกล้องวงจรปิดไม่สามารถควบคุมเพื่อให้ได้ภาพที่เป็นหน้าตรง (Frontal Face) เท่านั้น และบางครั้งผู้ต้องสงสัยเองก็อาจจะรู้ว่า มีกล้องวงจรปิด (CCTV Camera) ที่บันทึกภาพอยู่ ก็พยายามหันหน้าไปทางอื่น เพื่อไม่ให้เห็นใบหน้าชัดเจนแบบตรงๆ โดยให้เห็นหน้าในส่วนด้านข้างซึ่งอาจจะเป็นส่วนข้าง หรือ ด้านขวา อาจจะก้มหน้า หรือเงยหน้าไปทางอื่นก็ได้ ทำให้การสร้างคืนภาพไม่มีประสิทธิภาพเท่าที่ควร

ดังนั้นในงานวิจัยนี้ จึงทำการศึกษา การทำฮาโลซิเนสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination) โดยประยุกต์ใช้วิธีการ Adaptive Locally Linear Embedding (ALLE) ในการสร้างคืนภาพความละเอียดสูง



รูปที่ 3.16 Multiview Face Hallucination by Adaptive Locally Linear Embedding framework

3.4.1 ขั้นตอนที่น่าสนใจ (Proposed Face Hallucination Method)

3.4.1.1 นำเข้าภาพใบหน้าความละเอียดต่ำ (Low Resolution Image) ภาพความละเอียดต่ำที่ใช้ในการฝึกฝนระบบแต่ละมุมมอง (Low Resolution Training Image) และภาพความละเอียดสูงที่ใช้ในการฝึกฝนระบบแต่ละมุมมอง (High Resolution Training Image) ซึ่งเป็นแผ่นปะแบบซ้อนทับบางส่วน (Overlapping Patches)

3.4.1.2 ในแต่ละแผ่นปะ (Each Patch)

- 1) คำนวณค่าน้ำหนัก w ของภาพใบหน้าความละเอียดต่ำที่นำเข้า จากภาพความละเอียดต่ำที่ใช้ในการฝึกฝนระบบ ด้วยวิธีการ ALLE
- 2) รวมแผ่นปะภาพความละเอียดสูงของภาพความละเอียดต่ำที่นำเข้า โดยรวมแผ่นปะภาพความละเอียดสูงที่เหมือนกันของภาพความละเอียดต่ำที่ใช้ในการฝึกฝนระบบ โดยใช้ค่าน้ำหนักที่คำนวณได้
- 3) รวมแผ่นปะภาพความละเอียดสูงเป็นภาพใบหน้าความละเอียดสูงของแต่ละมุมมอง

3.5 การประเมินคุณภาพของภาพ

3.5.1 ประเมินคุณภาพของภาพโดยการวัดจากการมองเห็นของมนุษย์ (Subject Based)

3.5.2 ค่า PSNR (Peak Signal-to-Noise Ratio) ค่าที่ใช้วัดคุณภาพของภาพ เป็นค่ามาตรฐานที่นำมาใช้ในการเปรียบเทียบคุณภาพของรูปภาพดิจิทัล โดยการเปรียบเทียบกับรูปภาพต้นฉบับ ถ้าค่า PSNR สูงจะแสดงให้เห็นถึงคุณภาพของรูปที่ใกล้เคียงกับรูปภาพต้นฉบับ

สูตรคำนวณหาค่า PSNR

$$PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}} \quad (3.5)$$

255 คือ ค่าสูงสุดที่ข้อมูลขนาด 8 บิตสามารถแสดงได้

ค่า Mean Square Error (MSE) สามารถคำนวณได้จากสมการดังต่อไปนี้

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (3.6)$$

โดย mn คือ จำนวนจุดภาพแนวตั้งและแนวนอน

I คือ ภาพต้นฉบับ

K คือ ภาพที่ได้จากการประมวลผล

โดยผลที่ได้จากการทำฮาโลซิเนสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination) โดยประยุกต์ใช้วิธีการ Adaptive Locally Linear Embedding (ALLE) ในการสร้างคืนภาพความละเอียดสูง จะนำเสนอในบทที่ 4 ต่อไป

บทที่ 4

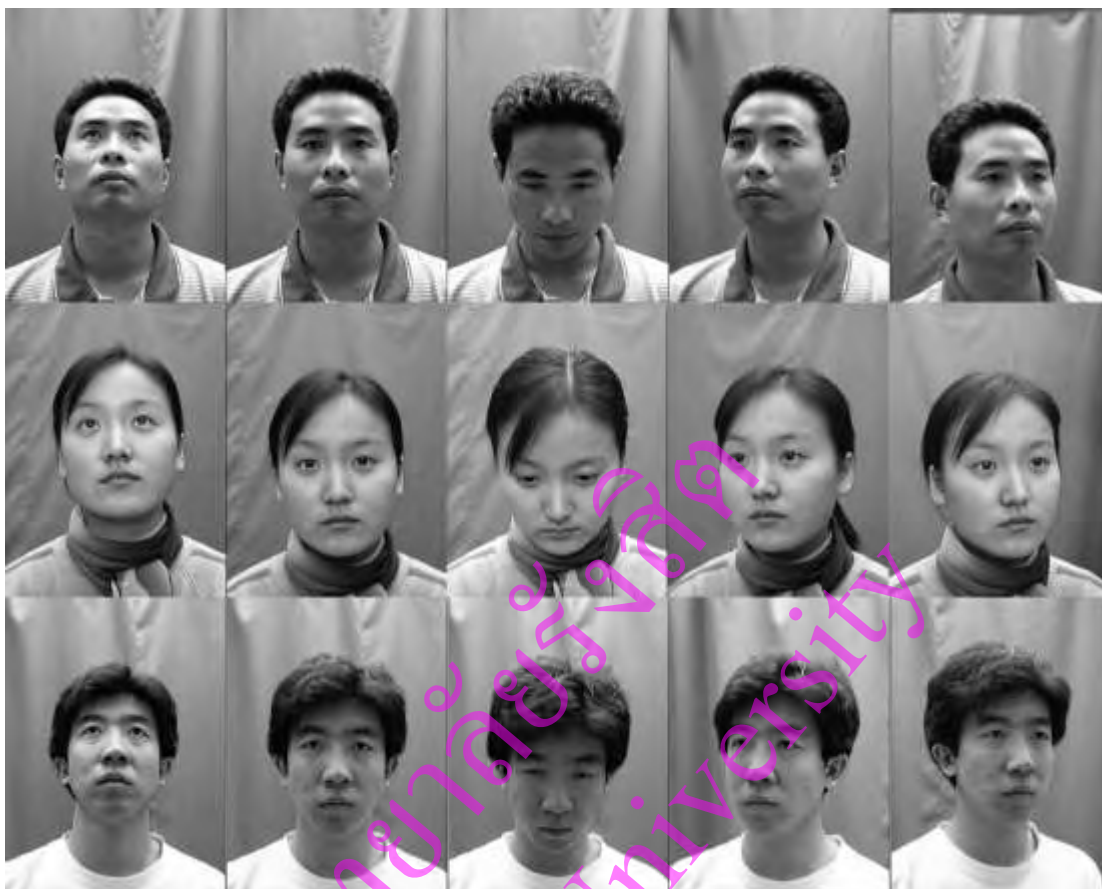
ผลการวิจัย

ผลของการวิจัยจะประกอบด้วยผลการสร้างฐานข้อมูลภาพใบหน้าของคนไทย แสดงในหัวข้อที่ 4.1 ฐานข้อมูลภาพใบหน้าที่ใช้ในการทำวิจัยนี้ประกอบด้วย ฐานข้อมูล The CAS-PEAL-R1 และ ฐานข้อมูลภาพใบหน้าของคนไทย แสดงในหัวข้อ 4.2 แสดงถึงกระบวนการก่อนหน้าผลการทดลอง แสดงในหัวข้อที่ 4.3 ผลการสร้างคืนภาพความละเอียดสูงจากมุมมองเดียว แสดงในหัวข้อที่ 4.3.1 หัวข้อที่ 4.3.2 แสดงผลการฮาหลูซิเนชันสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination) ตามลำดับ

4.1 ฐานข้อมูลภาพใบหน้า

4.1.1 The CAS-PEAL-R1

ฐานข้อมูล The CAS-PEAL-R1 (W. Gao and others, 2008) ที่ใช้ในงานวิจัยนี้เป็นฐานข้อมูลภาพใบหน้าที่มาตรฐาน ประกอบด้วยภาพใบหน้าจำนวนทั้งสิ้น 1,011 คน แบ่งเป็นภาพผู้ชายจำนวน 566 คน ผู้หญิง 445 คน โดยที่แต่ละคนจะประกอบไปด้วยภาพใบหน้า 5 มุมมอง ได้แก่ ภาพหน้าตรง ภาพใบหน้าที่ก้ม ภาพใบหน้าที่เงย ภาพใบหน้าที่ด้านซ้าย และภาพใบหน้าที่ด้านขวา โดยมีจำนวนภาพรวมทั้งสิ้น 5,055 ภาพ



รูปที่ 4.1 ตัวอย่างภาพจากฐานข้อมูล CAS-PEAL-R1

4.1.2 ฐานข้อมูลภาพใบหน้าของคนไทย

ฐานข้อมูลภาพใบหน้าของคนไทยที่ใช้ในการงานวิจัยนี้ ประกอบด้วยภาพใบหน้าจำนวนทั้งสิ้น 1,020 คน โดยภาพใบหน้าของแต่ละคนจะประกอบด้วยภาพหน้าตรง ภาพหน้าแยบ ภาพหน้าก้ม ภาพใบหน้าที่ด้านซ้าย และภาพใบหน้าที่ด้านขวา รวมจำนวนทั้งสิ้น 5,100 ภาพ โดยแต่ละภาพจะมีขนาด 360x480 จุดภาพ



รูปที่ 4.2 ตัวอย่างภาพจากฐานข้อมูลภาพใบหน้าคนไทย

4.2 กระบวนการก่อนหน้า (Preprocessing)

เนื่องจากภาพที่จะนำมาใช้ในงานวิจัย จากฐานข้อมูลภาพใบหน้าจาก หัวข้อที่ 4.1 มีข้อจำกัดที่เป็นปัญหาในการนำภาพใบหน้ามาประมวลผลกับระบบคอมพิวเตอร์ ได้แก่ ความคมชัดของภาพ ขนาดของภาพ ตำแหน่งของรูปภาพ สีพื้นหลังของภาพ ซึ่งปัญหาดังกล่าวสามารถแก้ไขได้ด้วยการนำภาพมาผ่านกระบวนการทำให้ภาพเป็นภาพที่มีมาตรฐานเดียวกันทั้งหมด ตามกระบวนการที่แสดงในรูปที่ 4.3



รูปที่ 4.3 แผนภาพแสดงกระบวนการก่อนหน้า

4.2.1 การจัดวางตำแหน่งของภาพใบหน้า (Face Alignment)

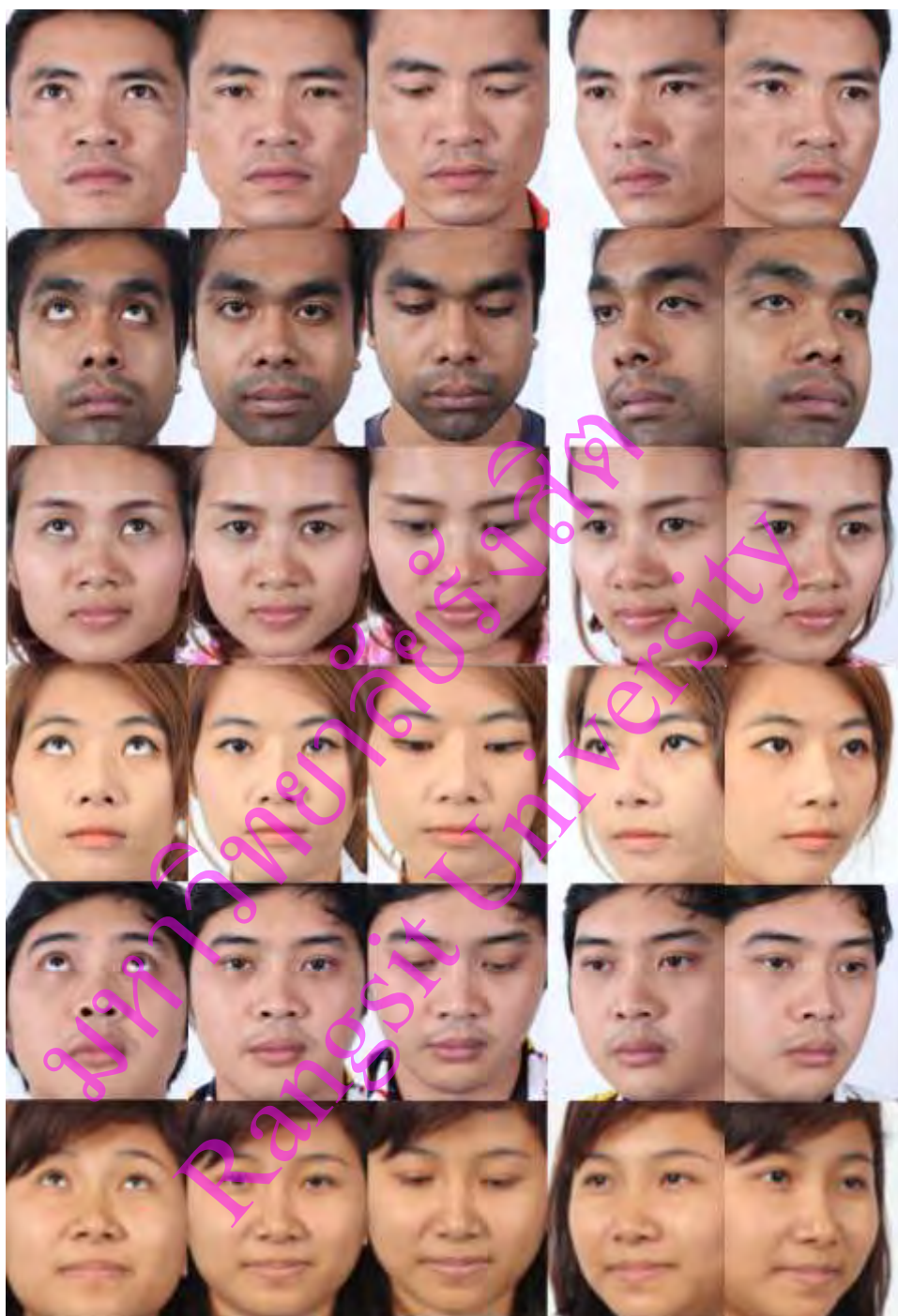
การจัดวางตำแหน่งของภาพใบหน้า (Face Alignment) เพื่อจะทำให้ตำแหน่งของภาพใบหน้าทุกภาพมีรูปแบบที่เหมาะสมในการประมวลผลต่อไป โดยในงานวิจัยนี้กำหนดตำแหน่งทั้ง 3 ตำแหน่งของภาพใบหน้าให้อยู่ในแนวเดียวกันโดยใช้วิธีการกำหนดตำแหน่งแบบ Manual คือ ตำแหน่งกลางตาข้างซ้าย ตำแหน่งกลางตาข้างขวา และ ตำแหน่งกลางปาก จากนั้นตัดภาพให้เหลือเฉพาะภาพใบหน้าส่วนที่ให้ความสนใจ และลดขนาดความละเอียดของภาพให้มีค่าเท่ากับ 128x96 จุดภาพ



รูปที่ 4.4 ตัวอย่างการจัดวางตำแหน่งของภาพใบหน้า



รูปที่ 4.5 ตัวอย่างภาพใบหน้าความละเอียด 128x96 จุดภาพ (CAS-PEAL-R1 Database)



รูปที่ 4.6 ตัวอย่างภาพใบหน้าความละเอียด 128x96 จุดภาพ (Thai Face Database)



รูปที่ 4.7 ตัวอย่างภาพใบหน้าความละเอียด 32x24 จุดภาพ (CAS-PEAL-R1 Database)



รูปที่ 4.8 ตัวอย่างภาพใบหน้าความละเอียด 32x24 จุดภาพ (Thai Face Database)

4.3 ผลการทดลอง (Experimental Results)

ในการวิจัยนี้ เราใช้ข้อมูลภาพใบหน้าจากฐานข้อมูลมาตรฐาน CAS-PEAL-R1 โดยการสุ่มเลือกภาพใบหน้าจำนวน 1,011 คน และฐานข้อมูลภาพใบหน้าของคนไทยที่ได้สร้างขึ้น โดยการสุ่มเลือกภาพใบหน้าทั้งสิ้น 5,100 ภาพ จากจำนวนคนทั้งหมด 1,020 คน โดยแต่ละคนจะมีภาพคนละ 5 มุมมองได้แก่ ภาพหน้าตรง ภาพก้มหน้า ภาพเงยหน้า ภาพใบหน้าที่ด้านซ้าย และภาพใบหน้าที่ด้านขวา ซึ่งภาพแต่ละภาพผ่านกระบวนการจัดวางตำแหน่งของภาพใบหน้า Face Alignment ทั้งสิ้น 3 ตำแหน่งคือ ตำแหน่งกลางตาซ้าย ตำแหน่งกลางตาขวา และ ตำแหน่งกลางปาก จากนั้นนำภาพที่ผ่านกระบวนการมาตัดภาพ (Crop) ให้มีความละเอียดขนาด 128x96 จุดภาพจากรูปที่ 4.5 เป็นภาพจากฐานข้อมูล CAS-PEAL-R1 และจากรูปที่ 4.6 เป็นภาพจากฐานข้อมูลภาพใบหน้าของคนไทย และ ภาพความละเอียดต่ำขนาด 32x24 จุดภาพจากรูปที่ 4.7 และจากรูปที่ 4.8 เป็นภาพจากฐานข้อมูลภาพใบหน้าของคนไทย

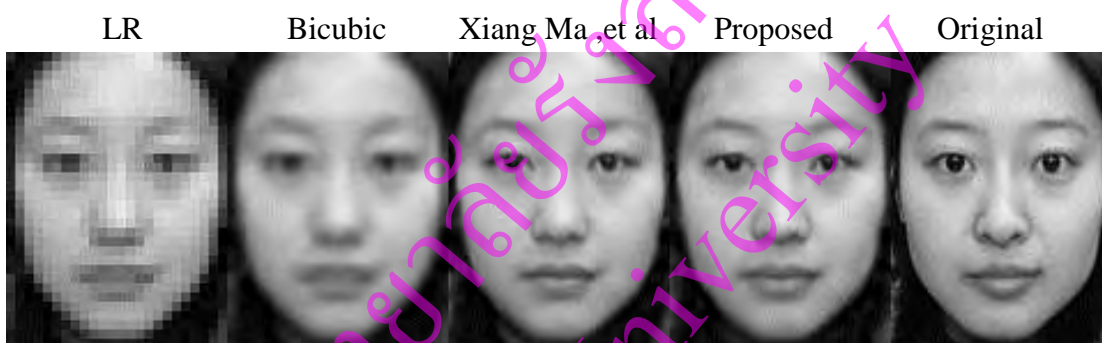
4.3.1 การฮาหลูซิเนสภาพใบหน้า (Face Hallucination)

4.3.1.1 การฮาหลูซิเนสภาพใบหน้าจากฐานข้อมูลภาพ CAS-PEAL-R1

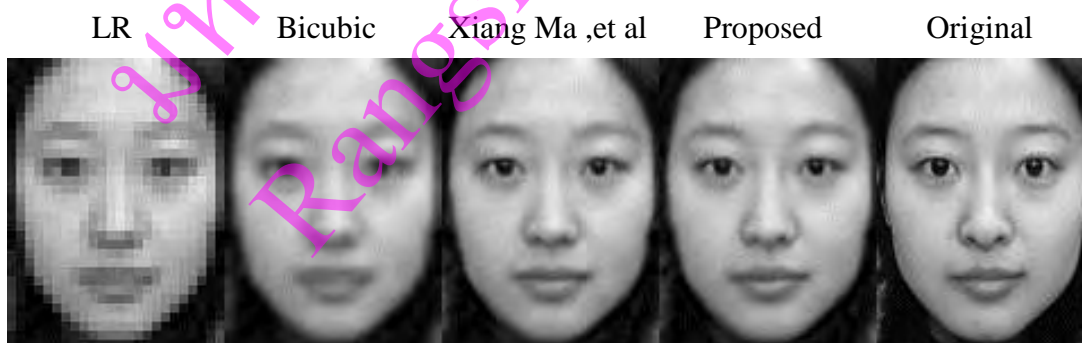
การทดลองโดยใช้ภาพจากฐานข้อมูลภาพ CAS-PEAL-R1 นำเข้าใบหน้ามุมมองใดมุมมองหนึ่งซึ่งเป็นภาพความละเอียดต่ำโดยมีขนาด 32x24 จุดภาพ ผลลัพธ์ที่ได้คือ การสร้างคืนภาพใบหน้าที่มีความละเอียดสูงของมุมมองเดียวกันโดยมีขนาด 128x96 จุดภาพ ซึ่งเป็นภาพที่ขยายจากภาพความละเอียดต่ำ 4 เท่า



รูปที่ 4.9 รูปภาพนำเข้าไปหน้าตรง ขนาด 32x24 จุดภาพ



รูปที่ 4.10 รูปฮาตุชินสภาพไปหน้าตรง (กำหนด $K = 100$, $\theta = 0.1$)



รูปที่ 4.11 รูปฮาตุชินสภาพไปหน้าตรง (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.12 รูปภาพนำเข้าไบหน้าเย ขนาด 32x24 จุดภาพ



รูปที่ 4.13 รูปฮาตุชินสภาพไบหน้าเย (กำหนด $K = 100$, $\theta = 0.1$)



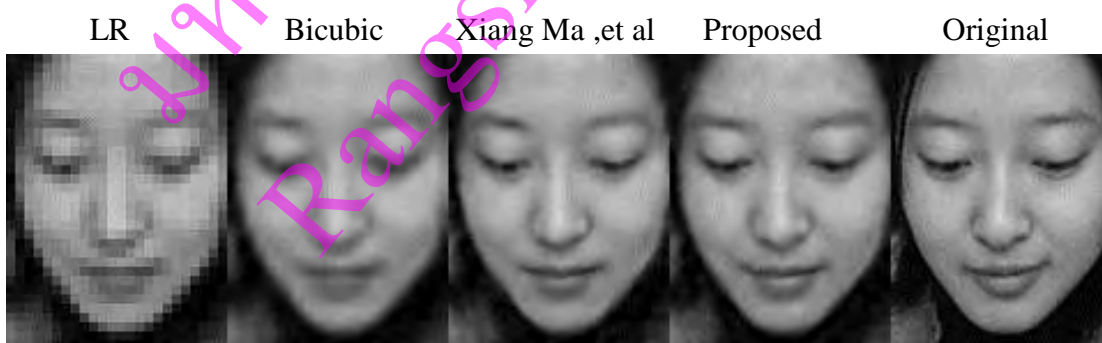
รูปที่ 4.14 รูปฮาตุชินสภาพไบหน้าเย (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.15 รูปภavnนำเข้าไบหน้ากัม ขนาด 32x24 จุดภาพ



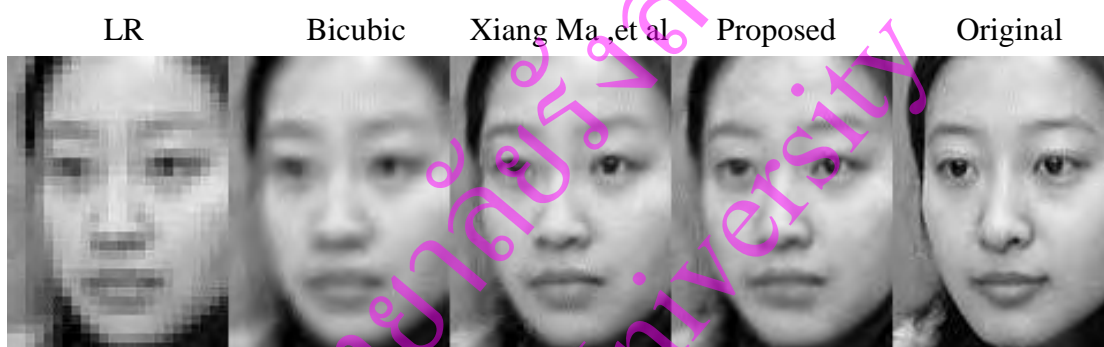
รูปที่ 4.16 รูปฮาลูชันสภาพไบหน้ากัม (กำหนด $K = 100$, $\theta = 0.1$)



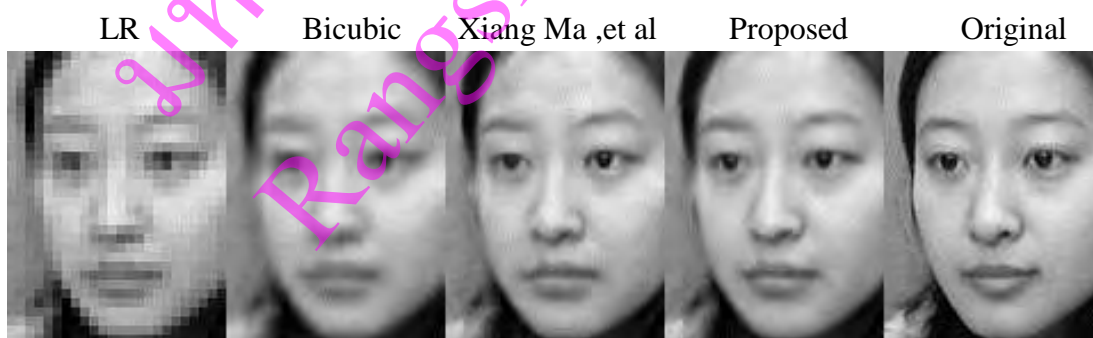
รูปที่ 4.17 รูปฮาลูชันสภาพไบหน้ากัม (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.18 รูปภาพนำเข้าไบหน้าด้านซ้าย ขนาด 32x24 จุดภาพ



รูปที่ 4.19 รูปฮาตุชินเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 100$, $\theta = 0.1$)



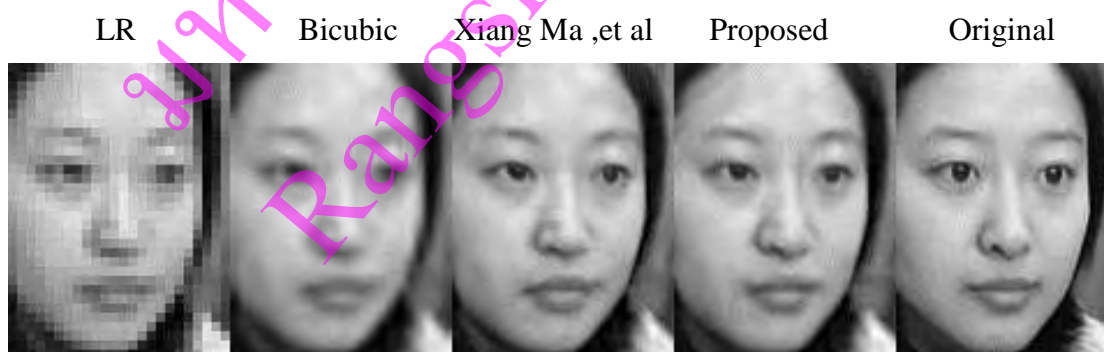
รูปที่ 4.20 รูปฮาตุชินเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)



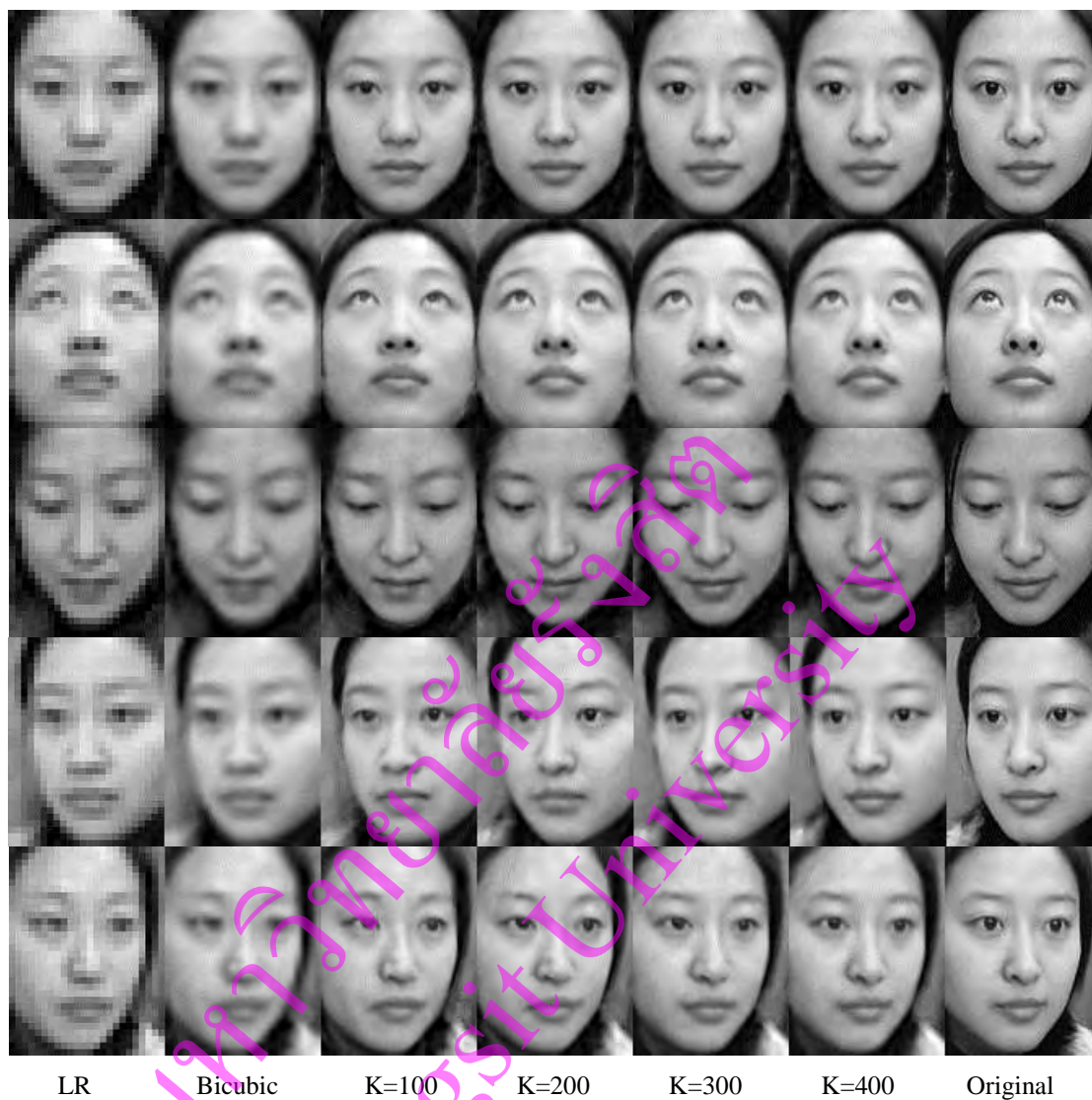
รูปที่ 4.21 รูปภาพนำเข้าไปบนหน้าด้านขวา ขนาด 32x24 จุดภาพ



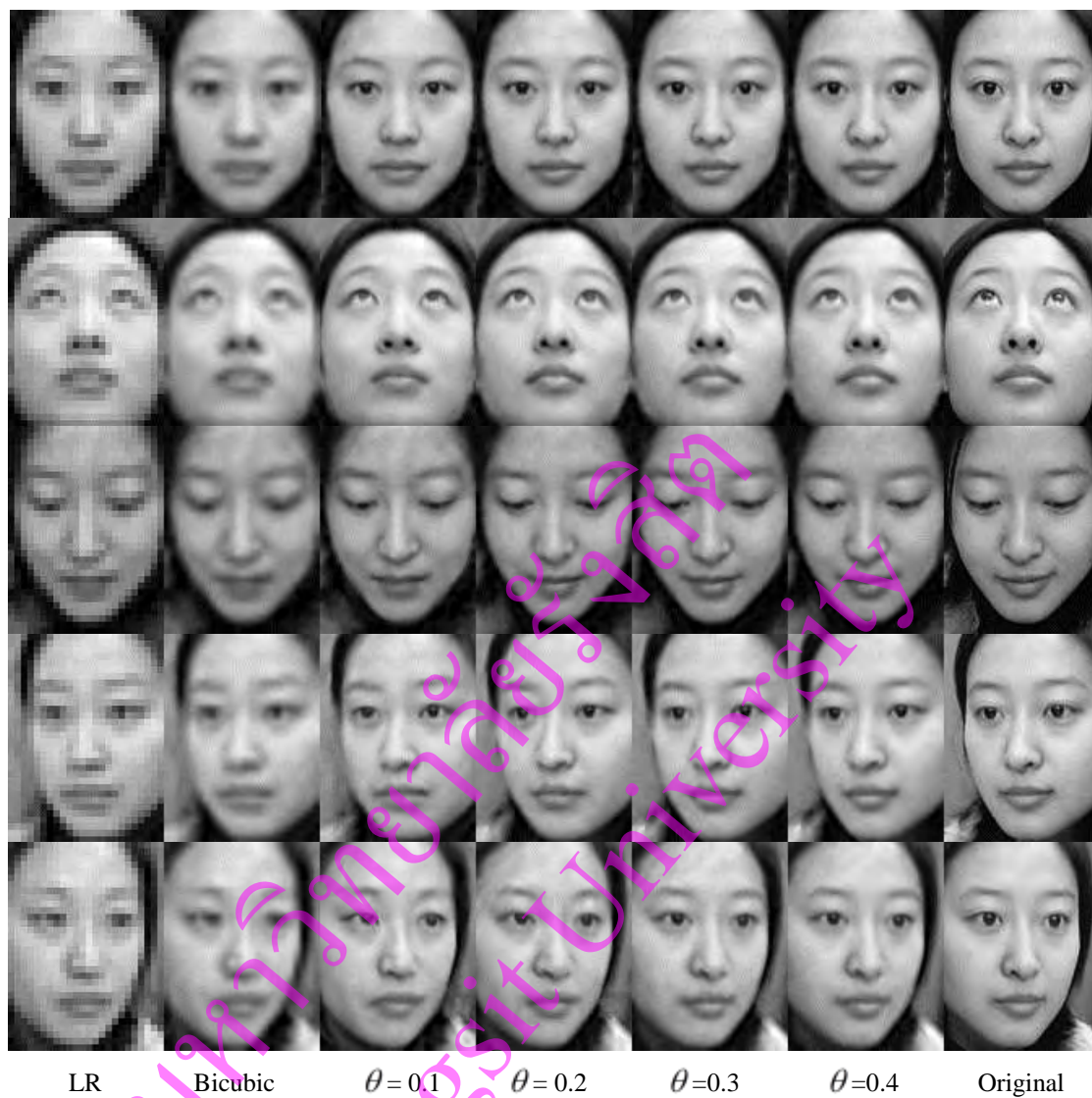
รูปที่ 4.22 รูปฮาตุชินเนสภาพใบหน้าด้านขวา (กำหนด $K = 100$, $\theta = 0.1$)



รูปที่ 4.23 รูปฮาตุชินเนสภาพใบหน้าด้านขวา (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.24 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า K ที่ต่างกัน (Single View)



รูปที่ 4.25 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า θ ที่ต่างกัน (Single View)

ผลการทดลองที่ 1 ซึ่งกำหนดค่า K (The Number of Nearest Neighbors) ให้มีค่าเท่ากับ 100 และ 200 กำหนดค่า θ (Threshold of Similarity) มีค่า 0.1 หรือ 10% และ 0.2 หรือ 20% ผลการทดลองแสดงให้เห็นว่าผลการเปรียบเทียบระหว่างวิธีการที่นำเสนอกับวิธีการ Bicubic และวิธีการของ Xiang Ma พบว่า วิธีการที่นำเสนอ สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง

ตารางที่ 4.1 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100, \theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	32.49	32.89
Up view	31.79	32.47
Down view	31.49	32.41
Left view	32.22	32.61
Right view	31.58	32.13

ตารางที่ 4.2 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	33.55	34.04
Up view	33.64	34.18
Down view	33.20	34.17
Left view	33.44	34.29
Right view	33.12	33.75

4.3.1.2 การฮาตุชินสภาพใบหน้าจากฐานข้อมูลภาพใบหน้าของคนไทย

การทดลองโดยใช้ภาพจากฐานข้อมูลภาพใบหน้าของคนไทย นำเข้าภาพใบหน้ามมมองใดมมมองหนึ่งซึ่งเป็นภาพความละเอียดต่ำโดยมีขนาด 32x24 จุดภาพ ผลลัพธ์ที่ได้คือ การสร้างคืนภาพใบหน้าที่มีความละเอียดสูงของมมมองเดียวกันโดยมีขนาด 128x96 จุดภาพ ซึ่งเป็นภาพที่ขยายจากภาพความละเอียดต่ำ 4 เท่า



รูปที่ 4.26 รูปภavnนำเข้าหน้าตรง ขนาด 32x24 จุดภาพ



รูปที่ 4.27 รูปฮาตุชินสภาพใบหน้าตรง (กำหนด $K = 100$, $\theta = 0.1$)



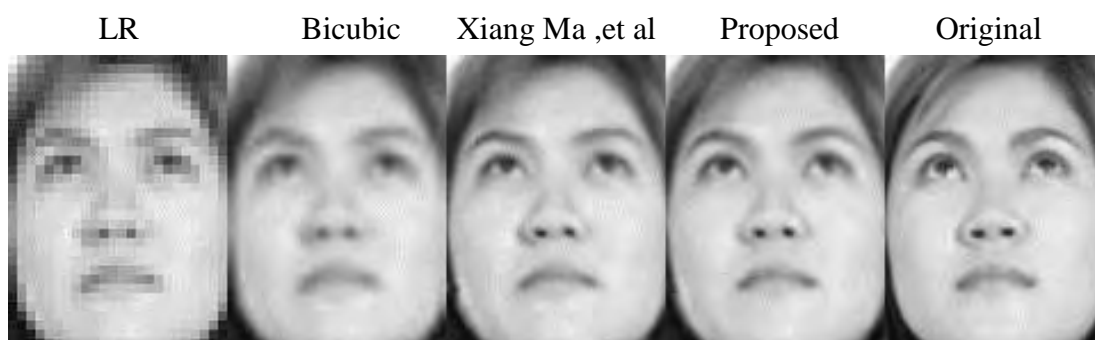
รูปที่ 4.28 รูปฮาตุชินเนสภาพใบหน้าตรง (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.29 รูปภาพนำเข้าใบหน้าเงย ขนาด 32x24 จุดภาพ



รูปที่ 4.30 รูปฮาตุชินเนสภาพใบหน้าเงย (กำหนด $K = 100$, $\theta = 0.1$)



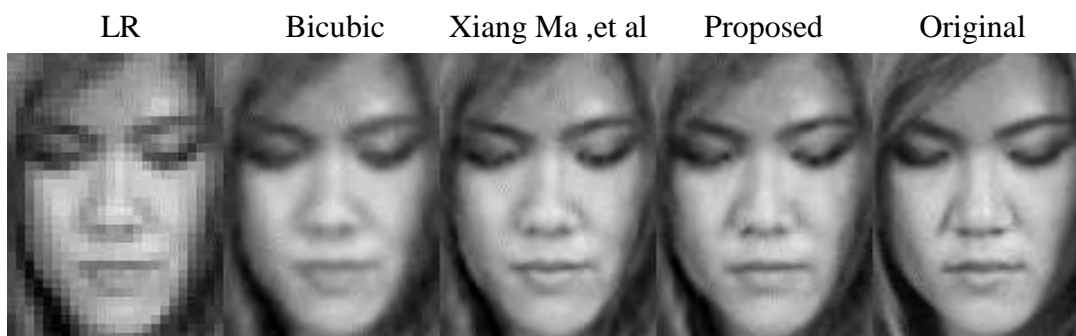
รูปที่ 4.31 รูปฮาตุชินเนสภาพโบน้าเงย (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.32 รูปภานำเข้าโบน้าก้ม ขนาด 32x24 จุดภาพ



รูปที่ 4.33 รูปฮาตุชินเนสภาพโบน้าก้ม (กำหนด $K = 100$, $\theta = 0.1$)



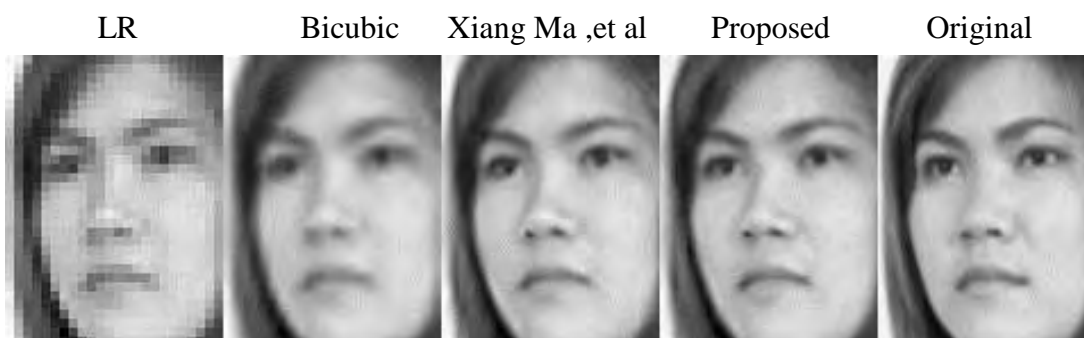
รูปที่ 4.34 รูปฮาตุชินนสภาพโบน้าก้ม (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.35 รูปภพนำเข้าโบน้าด้านซ้าย ขนาด 32×24 จดภพ



รูปที่ 4.36 รูปฮาตุชินนสภาพโบน้าด้านซ้าย (กำหนด $K = 100$, $\theta = 0.1$)



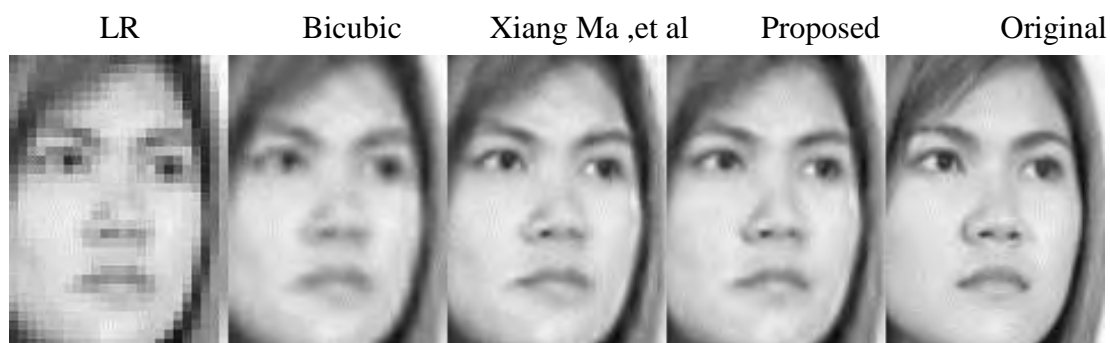
รูปที่ 4.37 รูปฮาตุชินเนสภาพไบหน้าด้านซ้าย (กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.38 รูปภาพนำเข้าไบหน้าด้านขวา ขนาด 32x24 จุดภาพ

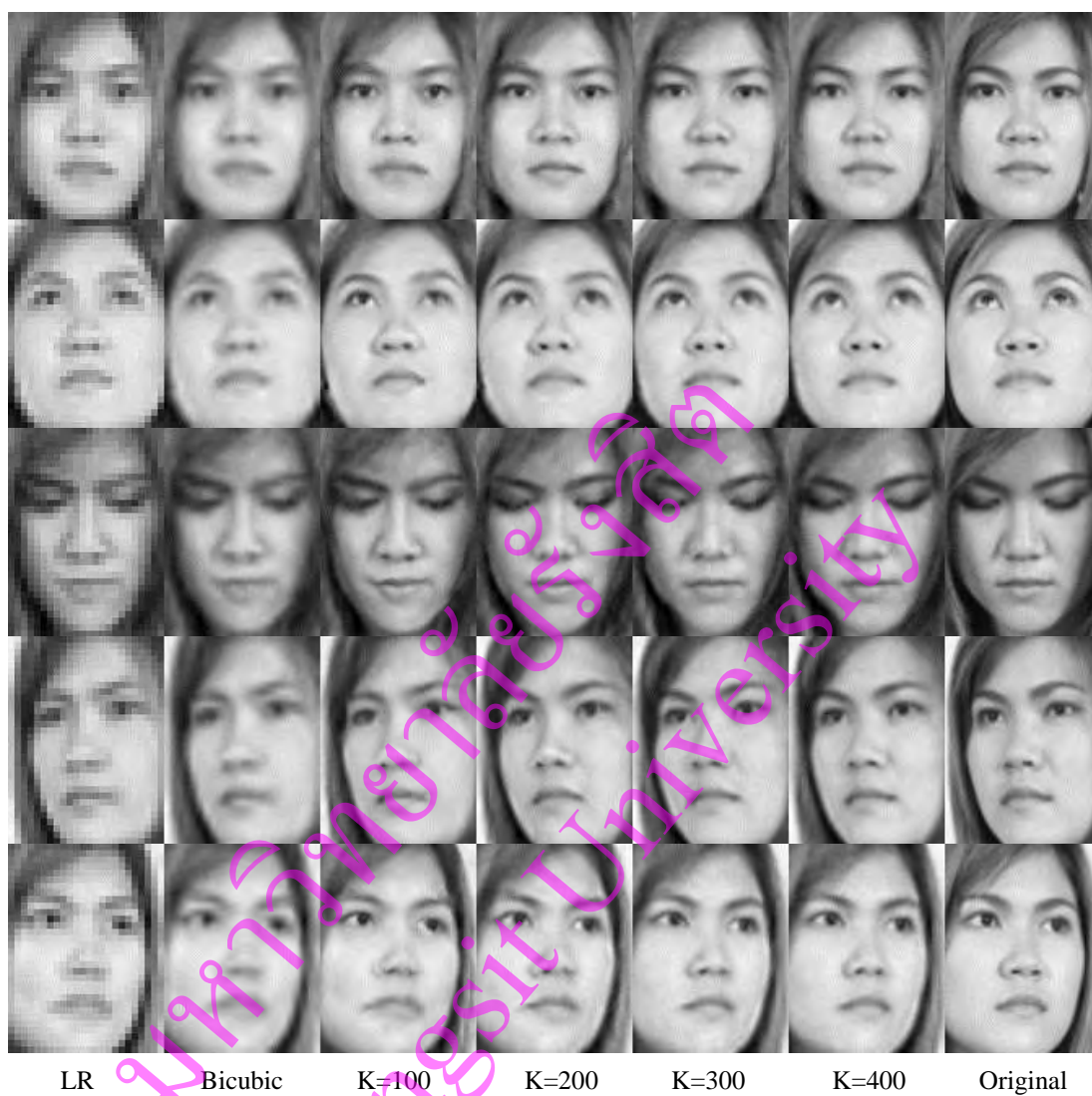


รูปที่ 4.39 รูปฮาตุชินเนสภาพไบหน้าด้านขวา (กำหนด $K = 100$, $\theta = 0.1$)



รูปที่ 4.40 รูปฮาอูซิเนสภาพใบหน้าด้านขวา (กำหนด $K = 200$, $\theta = 0.2$)

มหาวิทยาลัยรังสิต
Rangsit University



รูปที่ 4.41 แสดงการเปรียบเทียบผลของการฮาธุซิเนส
โดยกำหนดค่า K ที่ต่างกัน (Single View)



รูปที่ 4.42 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า θ ที่ต่างกัน (Single View)

ผลการทดลองที่ 2 เป็นการทดลองโดยใช้ภาพใบหน้าของคนไทย ซึ่งกำหนดค่า K (The Number of Nearest Neighbors) ให้มีค่าเท่ากับ 100 และ 200 กำหนดค่า θ (Threshold of Similarity) มีค่า 0.1 หรือ 10% และ 0.2 หรือ 20% จากรูปที่แสดงให้เห็นว่าผลการเปรียบเทียบระหว่างวิธีการที่นำเสนอกับวิธีการ Bicubic และ วิธีการของ Xiang Ma พบว่า วิธีการที่นำเสนอ สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง

ตารางที่ 4.3 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100, \theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	31.80	31.85
Up view	32.97	33.76
Down view	31.17	32.21
Left view	31.88	32.82
Right view	31.91	32.34

ตารางที่ 4.4 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	33.02	33.22
Up view	34.44	35.39
Down view	32.98	34.35
Left view	33.70	34.54
Right view	33.38	34.00

4.3.2 การฮาหลูซิเนสภาพใบหน้าที่มีหลายมุมมอง (Multiview Face Hallucination)

4.3.2.1 การฮาหลูซิเนสภาพใบหน้าที่มีหลายมุมมองฐานข้อมูล CAS-PEAL-R1

การทดลองโดยใช้ภาพจากฐานข้อมูลภาพ CAS-PEAL-R1 นำเข้าภาพใบหน้า มุมมองใดมุมมองหนึ่งซึ่งเป็นภาพความละเอียดต่ำโดยมีขนาด 32x24 จุดภาพ ผลลัพธ์ที่ได้คือ การสร้างคืนภาพใบหน้าที่มีความละเอียดสูงของ 5 มุมมอง ได้แก่ มุมมองใบหน้าที่ตรง มุมมองใบหน้าที่เอียง มุมมองใบหน้าที่ก้ม มุมมองใบหน้าที่ด้านซ้ายและมุมมองใบหน้าที่ด้านขวา โดยมีขนาด 128x96 จุดภาพ ซึ่งเป็นภาพที่ขยายจากภาพความละเอียดต่ำ 4 เท่า



รูปที่ 4.43 รูปภาพนำเข้าใบหน้าที่ตรง ขนาด 32x24 จุดภาพ



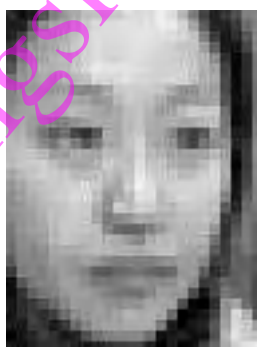
รูปที่ 4.44 รูปภาพนำเข้าใบหน้าที่เอียง ขนาด 32x24 จุดภาพ



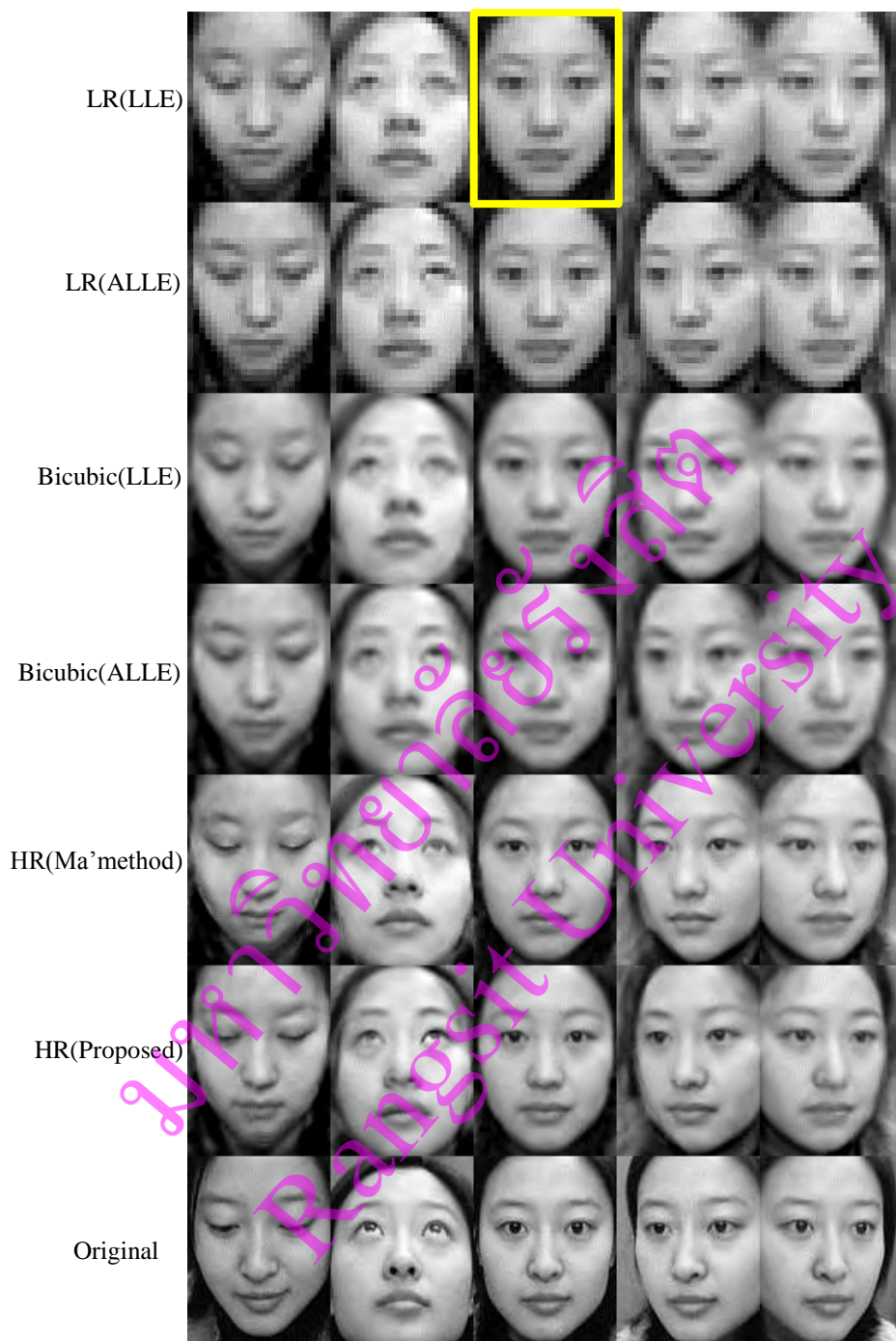
รูปที่ 4.45 รูปภาพนำเข้าไปบนหน้ากัม ขนาด 32x24 จุดภาพ



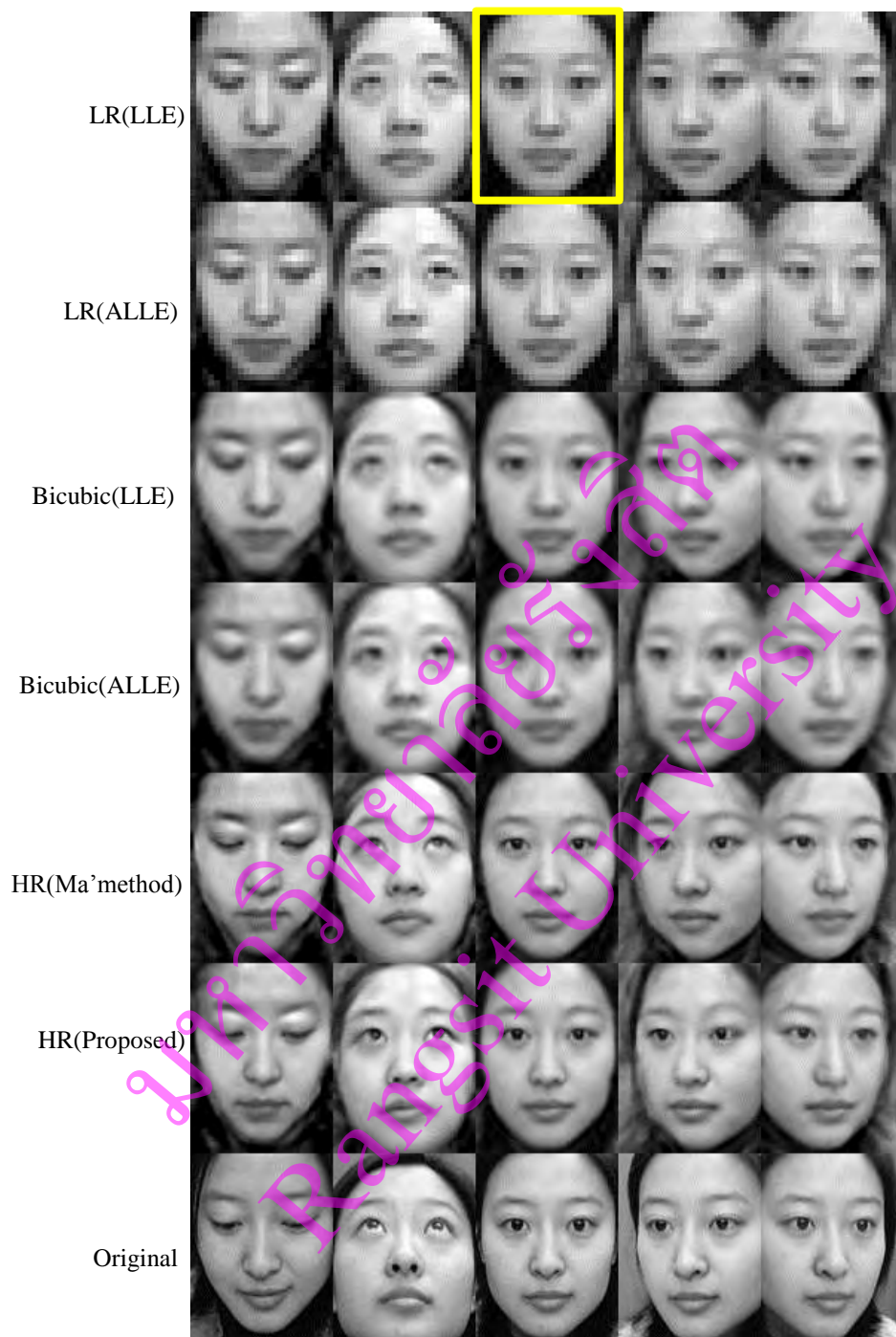
รูปที่ 4.46 รูปภาพนำเข้าไปบนหน้าด้านซ้าย ขนาด 32x24 จุดภาพ



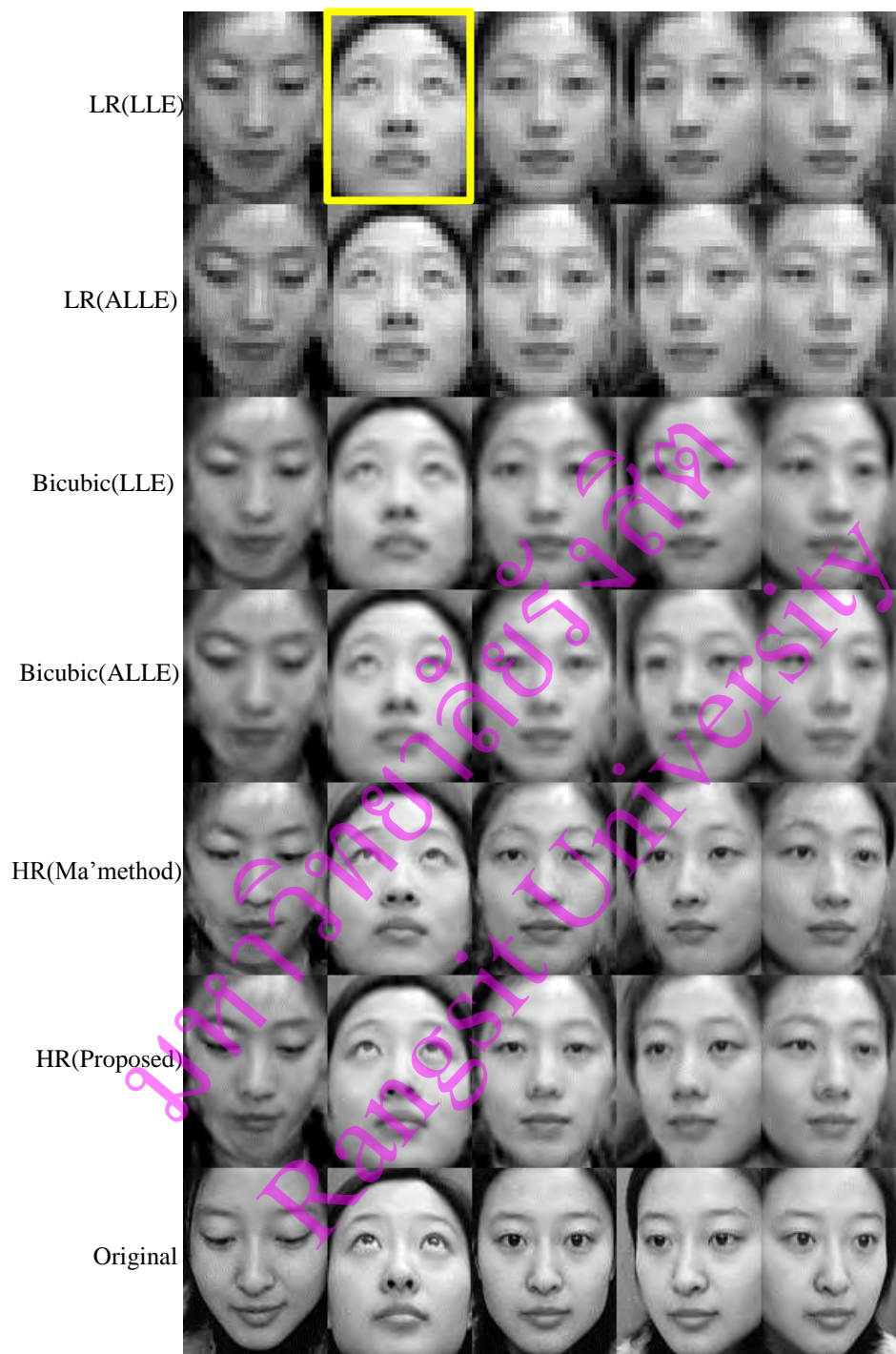
รูปที่ 4.47 รูปภาพนำเข้าไปบนหน้าด้านขวา ขนาด 32x24 จุดภาพ



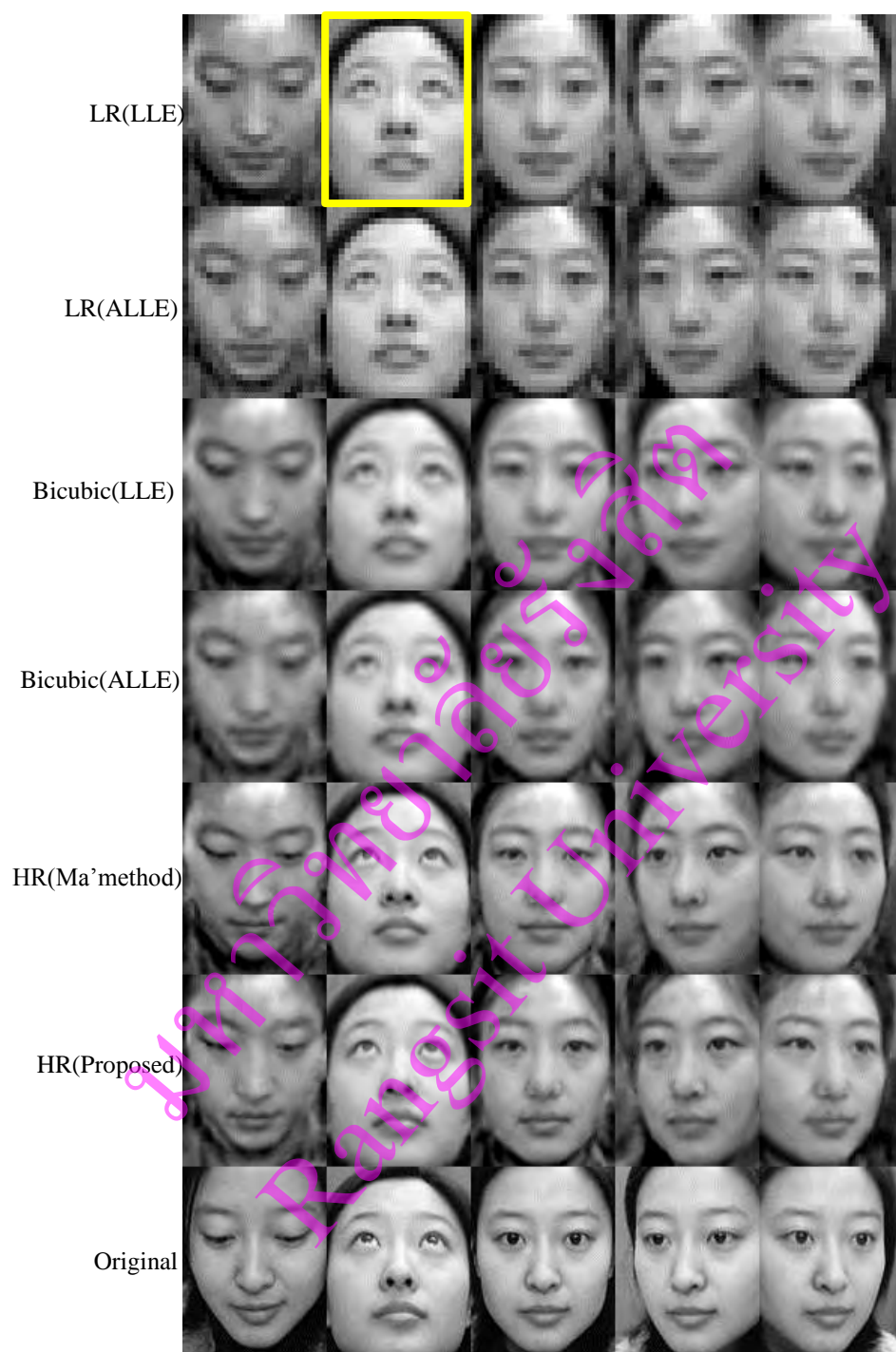
รูปที่ 4.48 รูปฮาโลจีนสภาพใบหลายมุมมองจากภาพนำเข้าไปหน้าตรง
(กำหนด $K = 100$, $\theta = 0.1$)



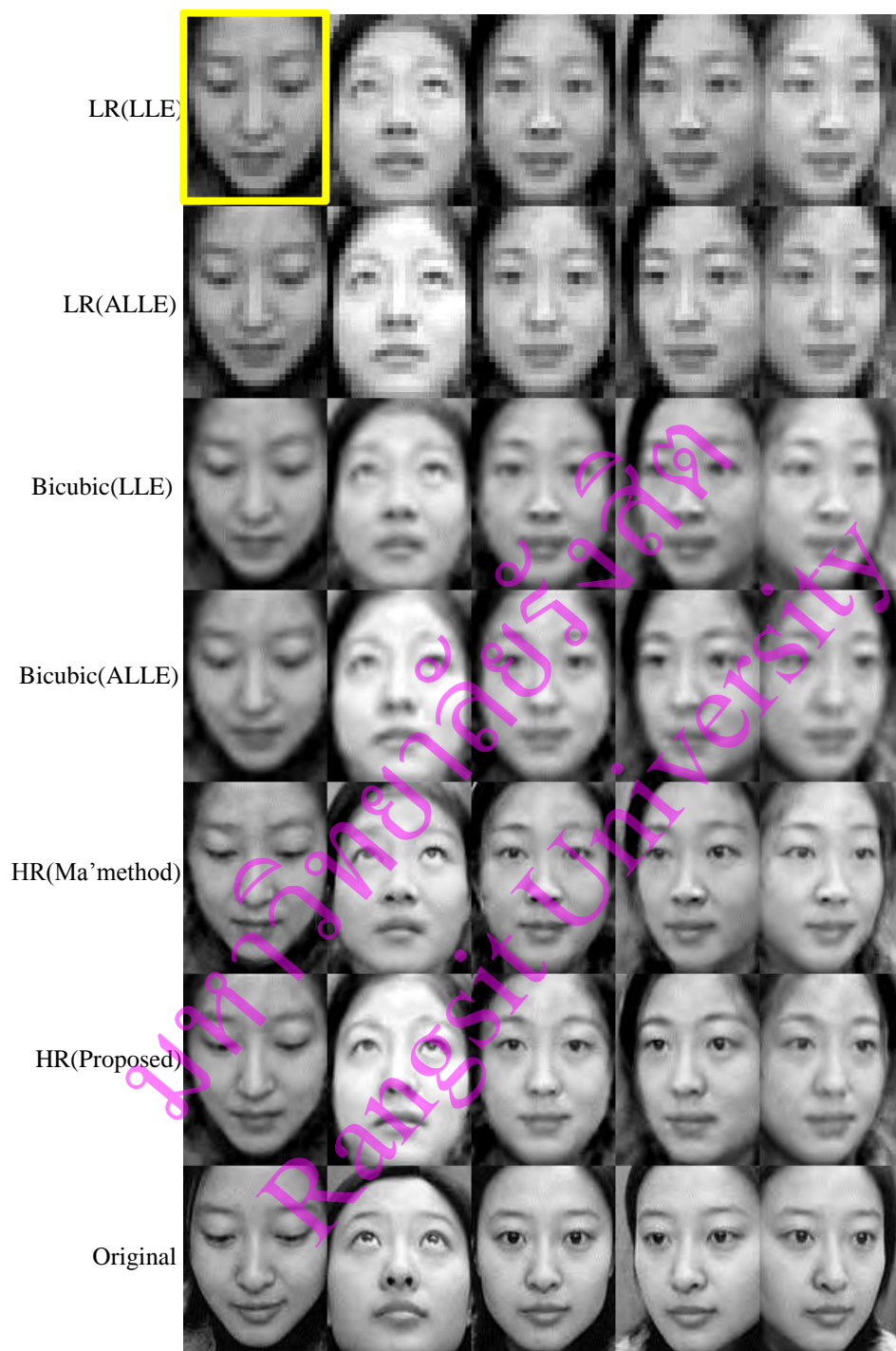
รูปที่ 4.49 รูปฮาตุชินเนสภาพโอบหลายมุมมองจากภาพนำเข้าใบหน้าตรง
(กำหนด $K = 200$, $\theta = 0.2$)



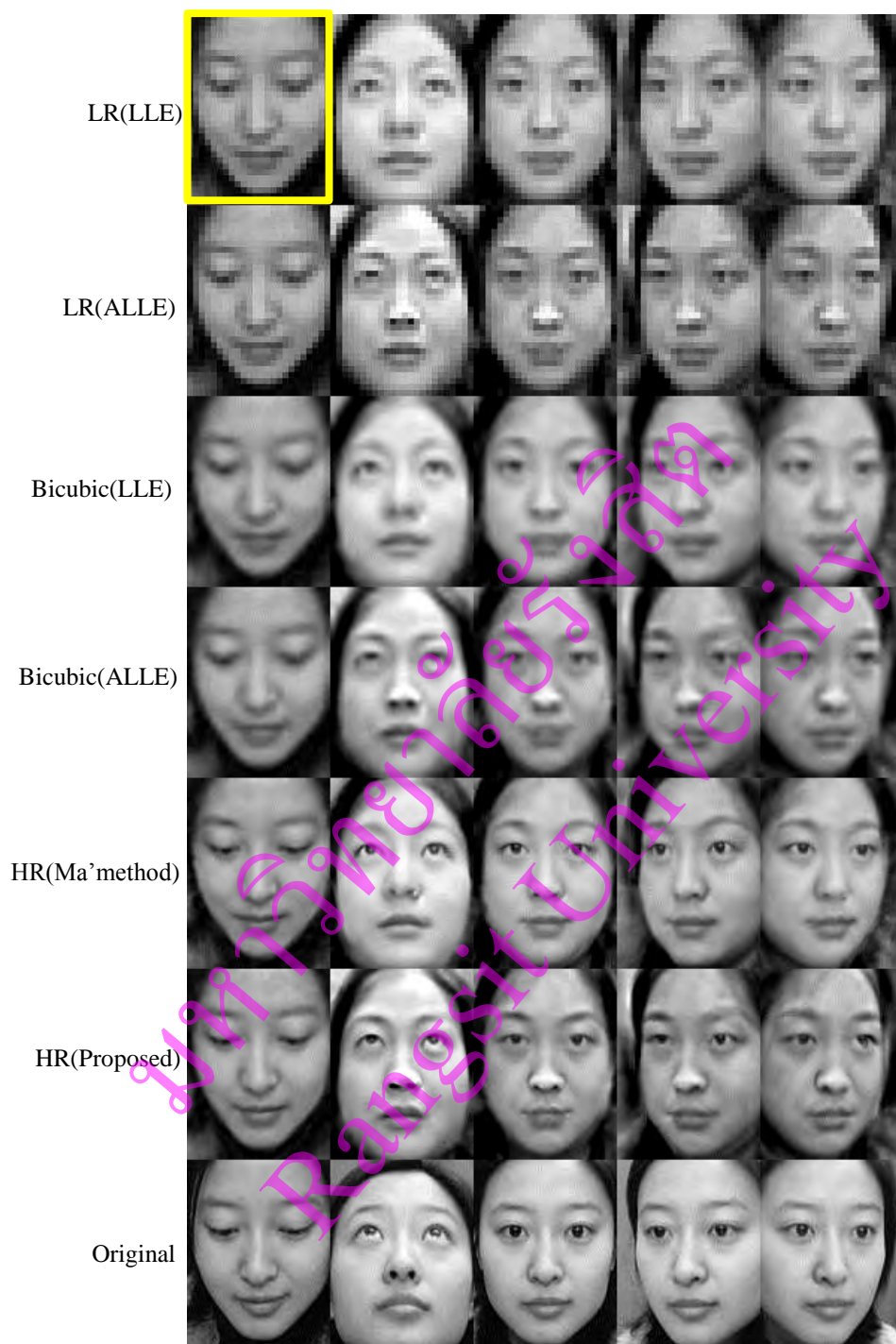
รูปที่ 4.50 รูปฮาตุชินสภาพใบหลายมุมมองจากภาพนำเข้าใบหน้าแยก
(กำหนด $K = 100$, $\theta = 0.1$)



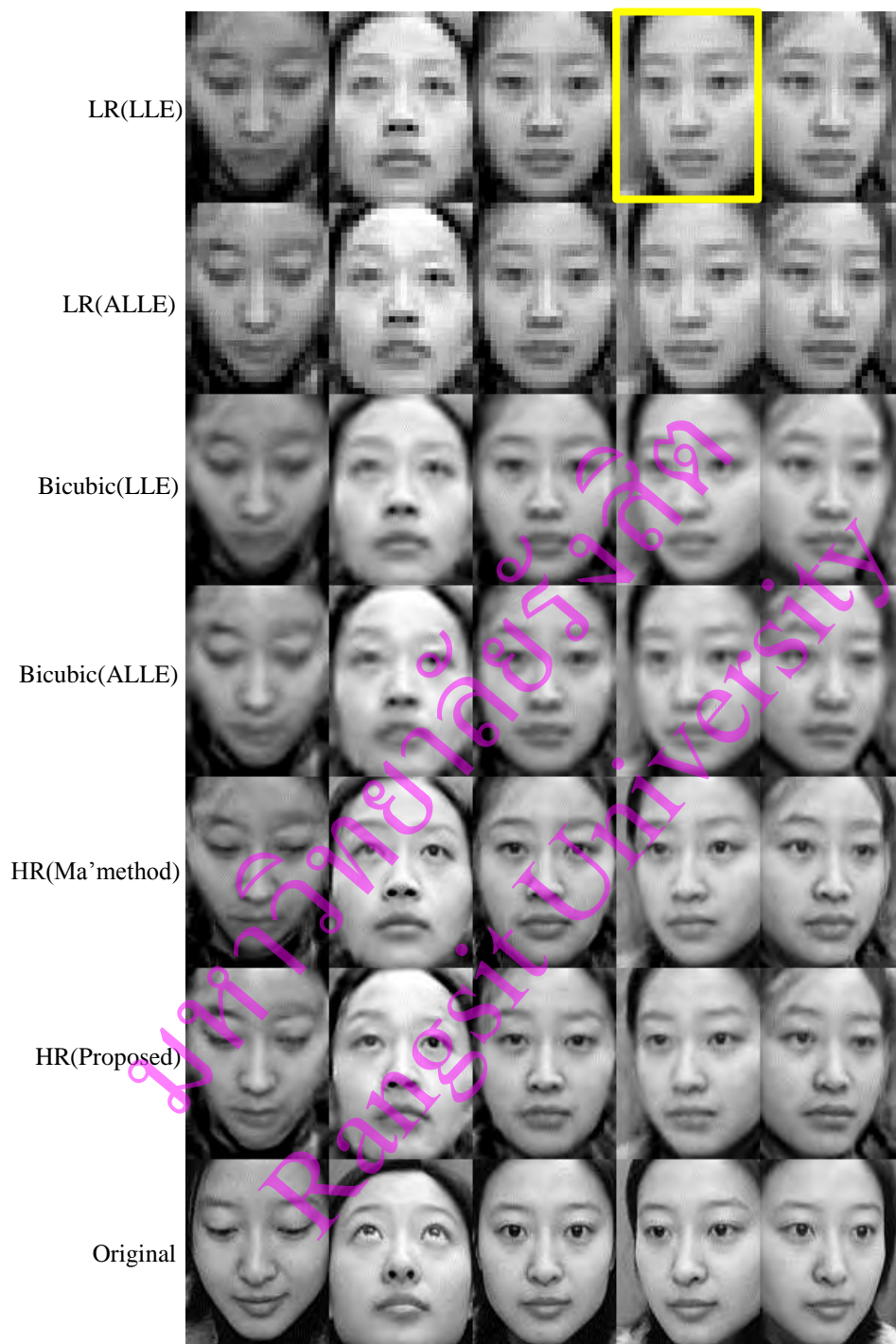
รูปที่ 4.51 รูปฮาตุจินสภาพใบหลายมุมมองจากภาพนำเข้าใบหน้าแยก
(กำหนด $K = 200, \theta = 0.2$)



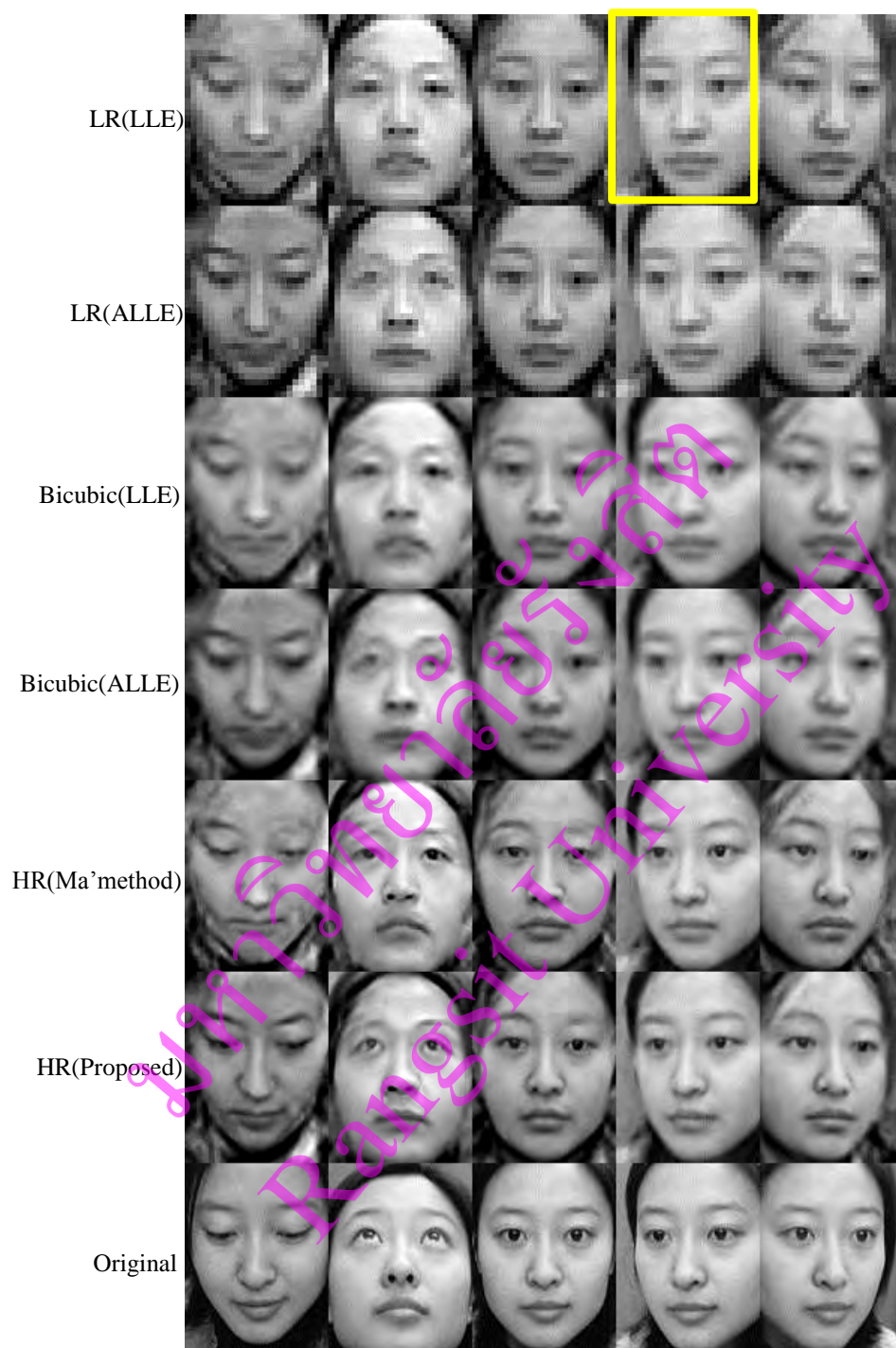
รูปที่ 4.52 รูปฮาตุจินเนสภาพใบหลายมุมมองจากภาพนำเข้าใบหน้าก้ม
(กำหนด $K = 100$, $\theta = 0.1$)



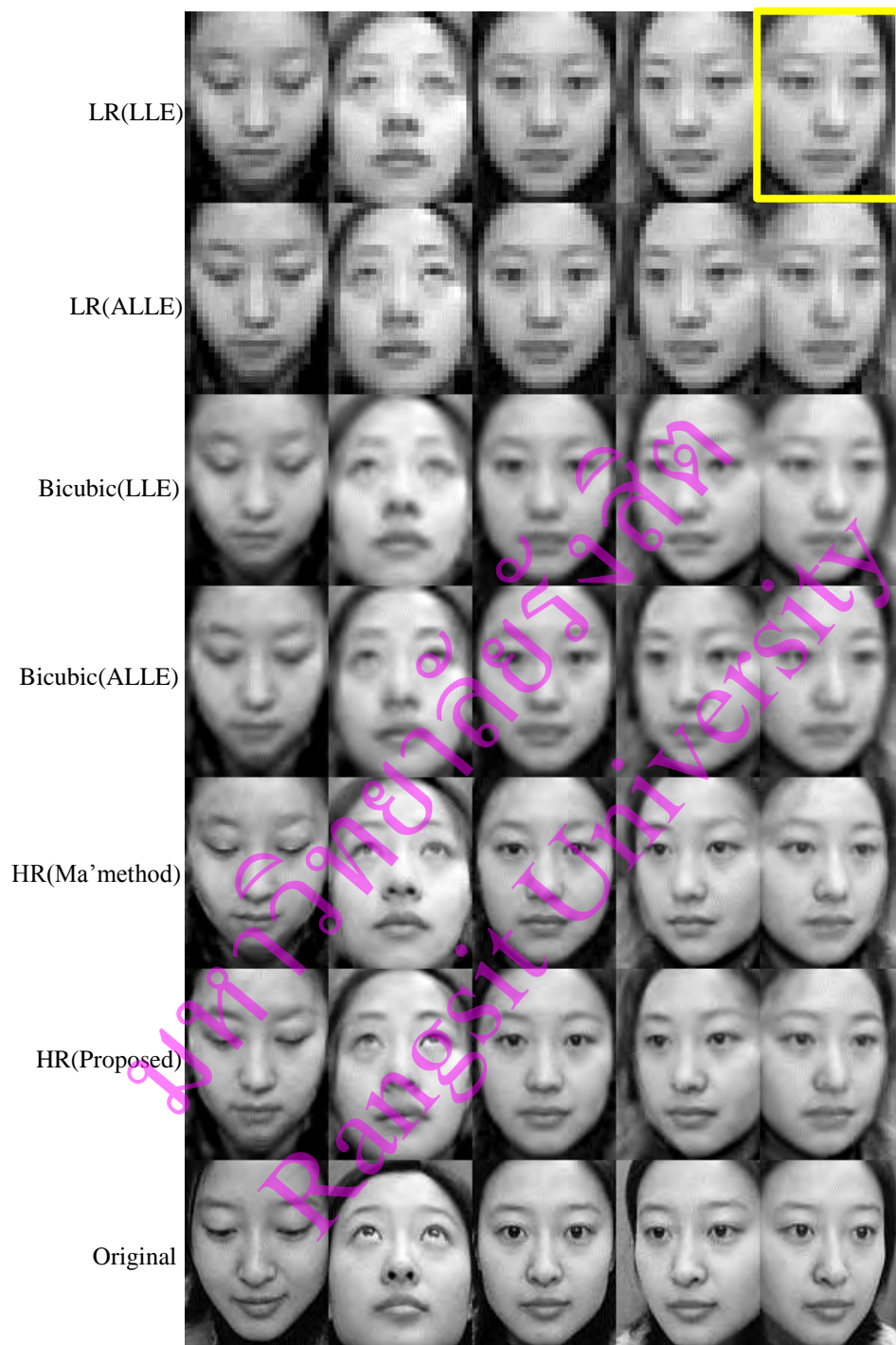
รูปที่ 4.53 รูปฮาตุจินเนสภาพโบหลายมุมมองจากภาพนำเข้าโบหน้าก้ม
(กำหนด $K = 200$, $\theta = 0.2$)



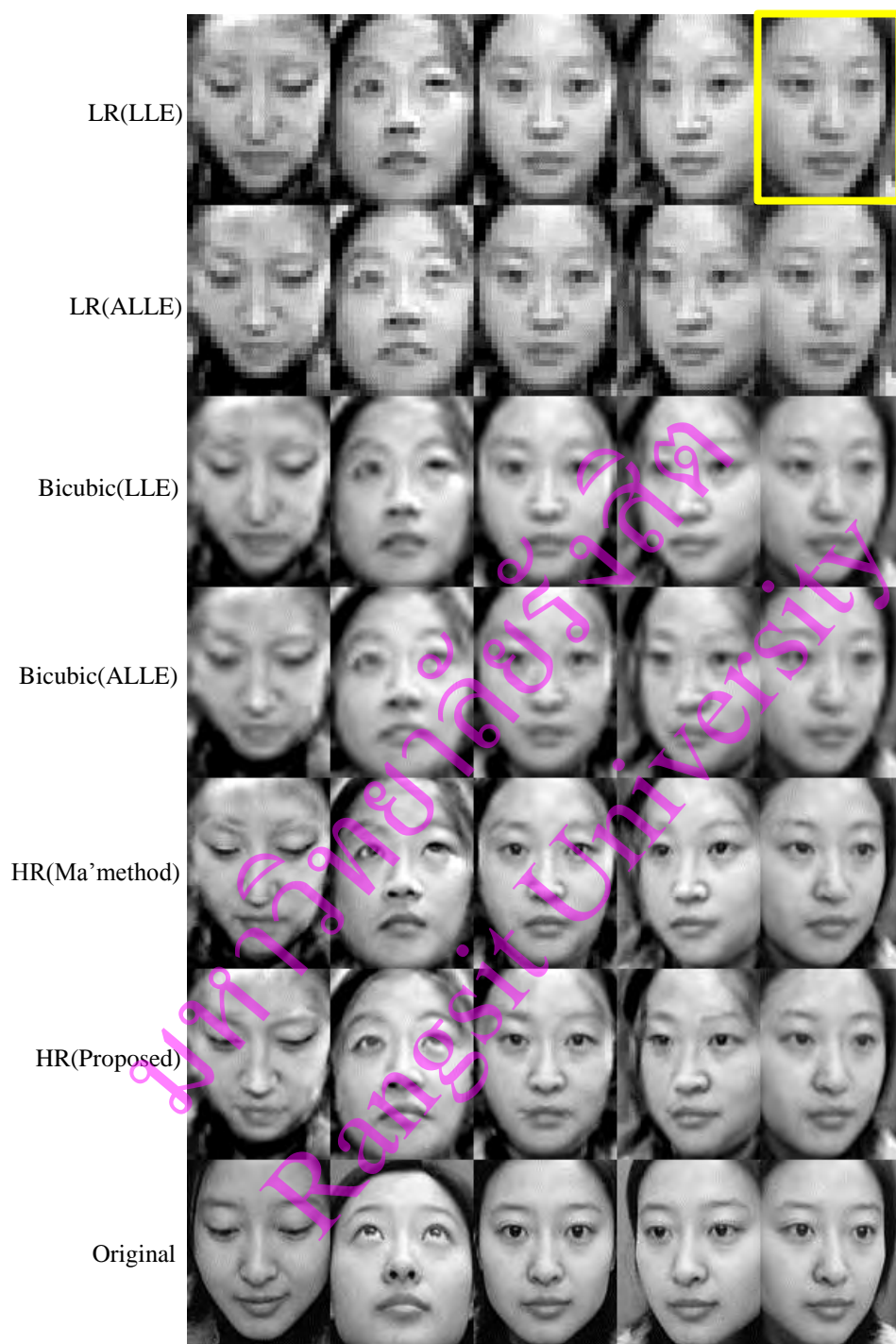
รูปที่ 4.54 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปบนหน้าด้านซ้าย
(กำหนด $K = 100$, $\theta = 0.1$)



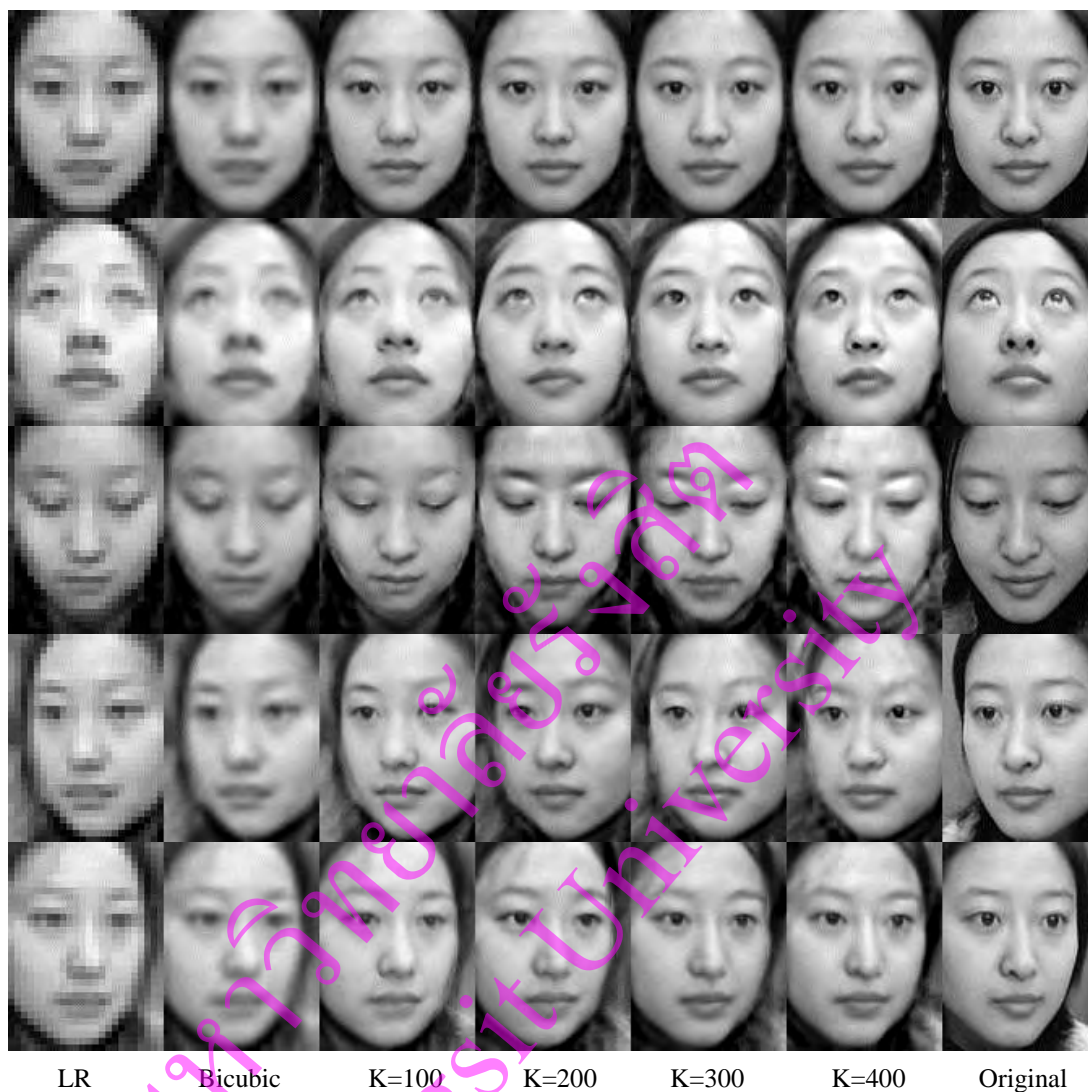
รูปที่ 4.55 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปบนหน้าด้านซ้าย
(กำหนด $K = 200, \theta = 0.2$)



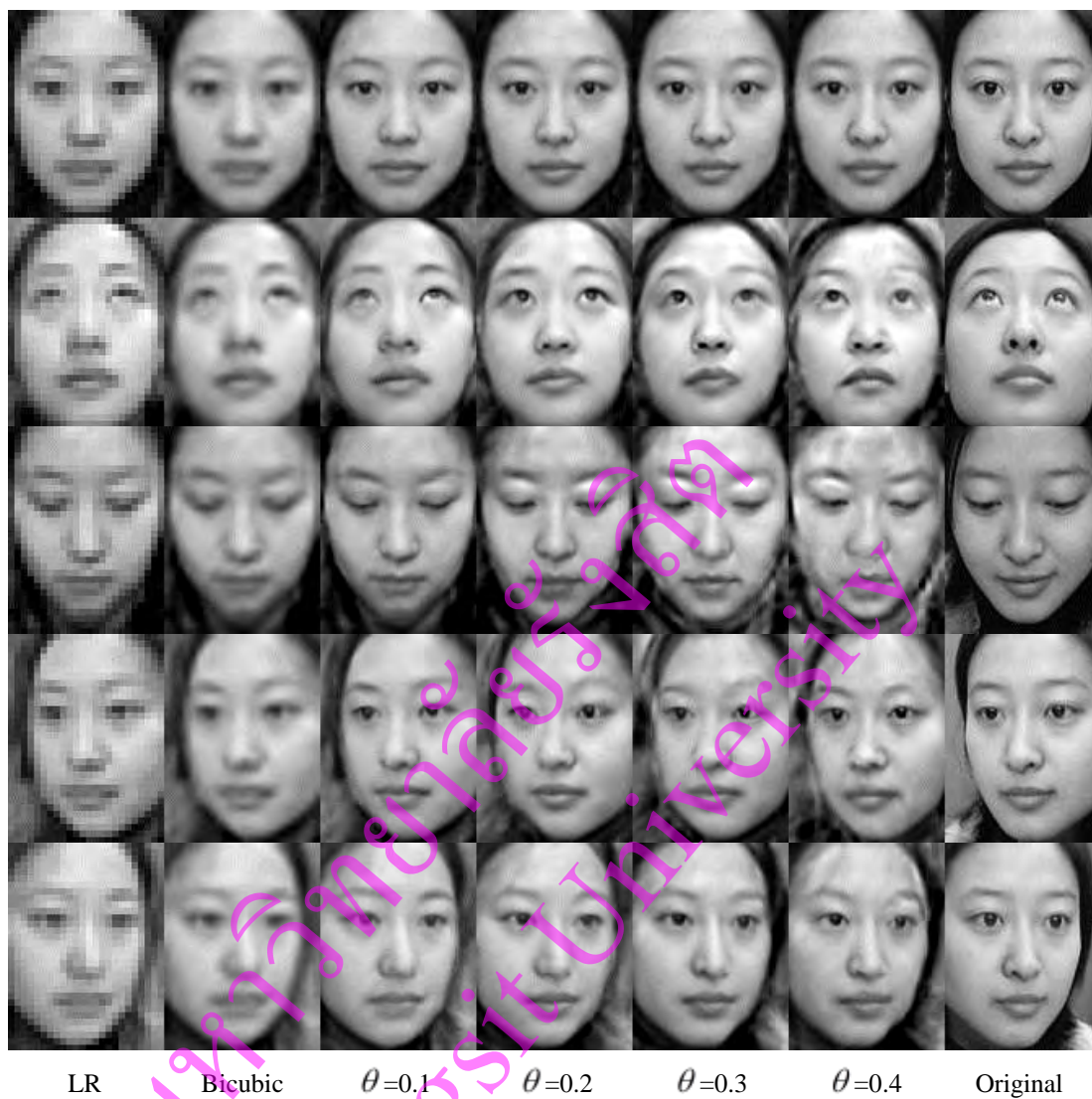
รูปที่ 4.56 รูปฮาตุชินสภาพใบหลายมุมมองจากภาพนำเข้าไปบนหน้าด้านขวา
(กำหนด $K = 100$, $\theta = 0.1$)



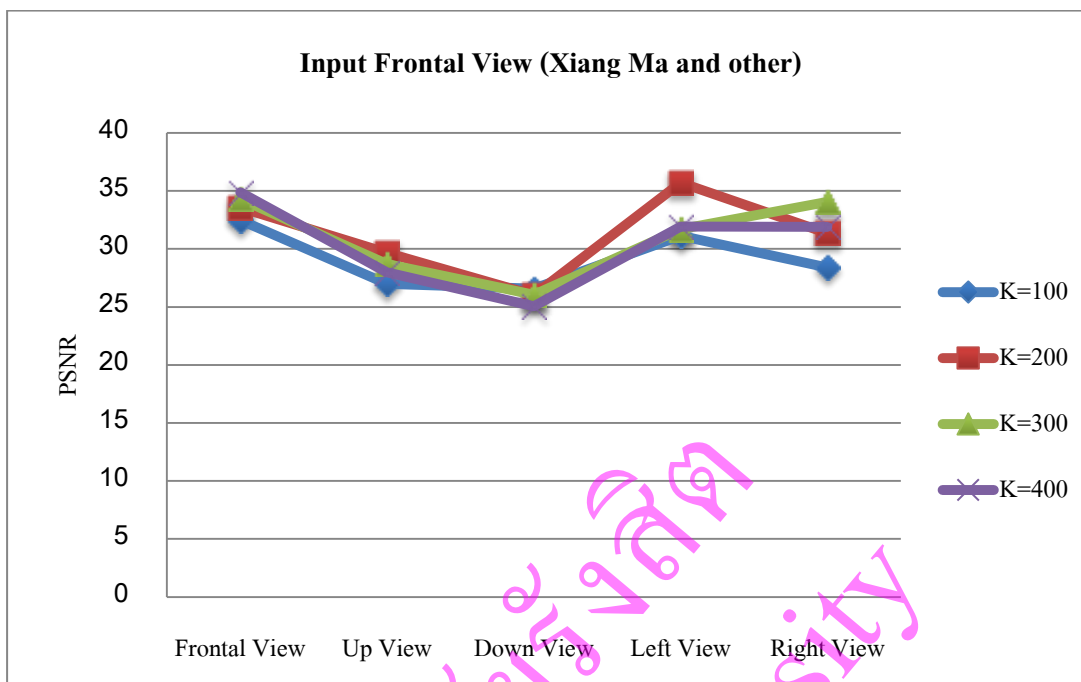
รูปที่ 4.57 รูปฮาโลจีนสภาพใบหลายมุมมองจากภาพนำเข้าไปบนหน้าด้านขวา
(กำหนด $K = 200$, $\theta = 0.2$)



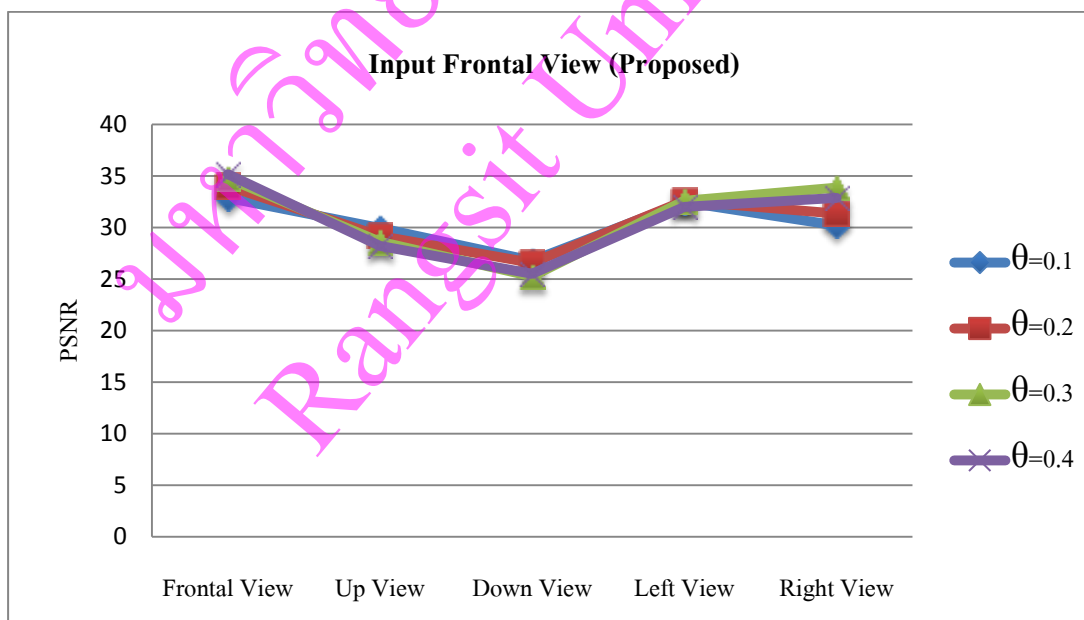
รูปที่ 4.58 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า K ที่ต่างกัน (Multiview)



รูปที่ 4.59 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า θ ที่ต่างกัน (Multiview)



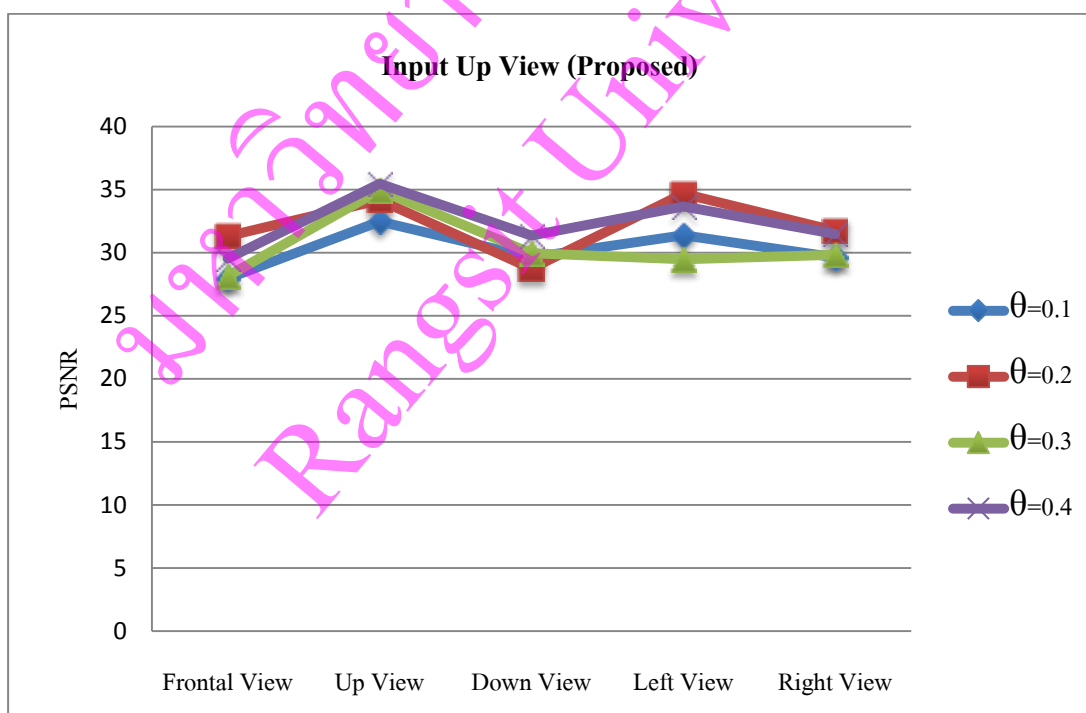
รูปที่ 4.60 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Frontal View)



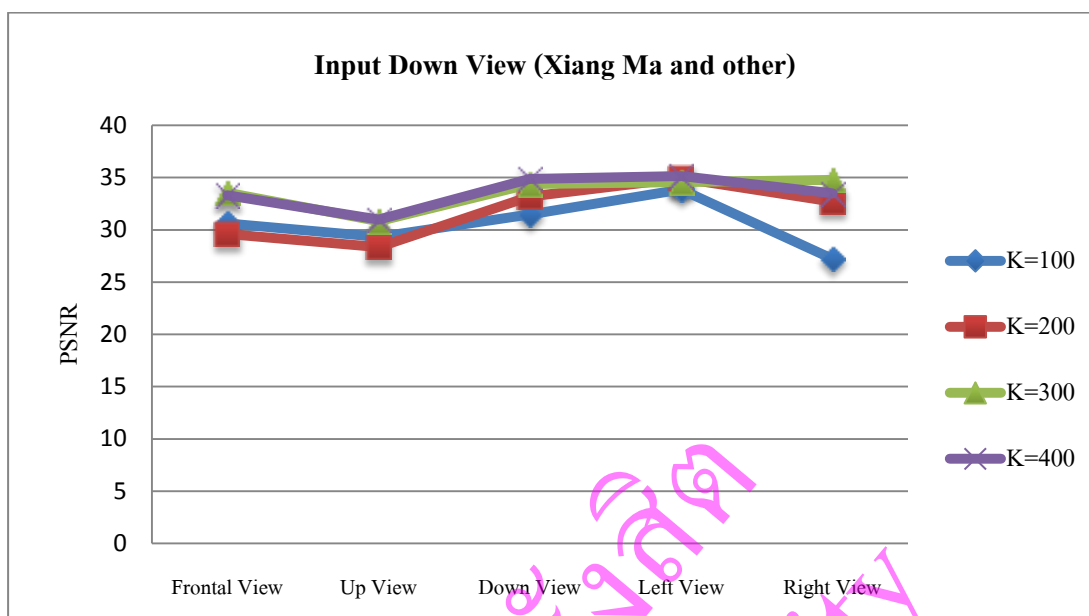
รูปที่ 4.61 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Frontal View)



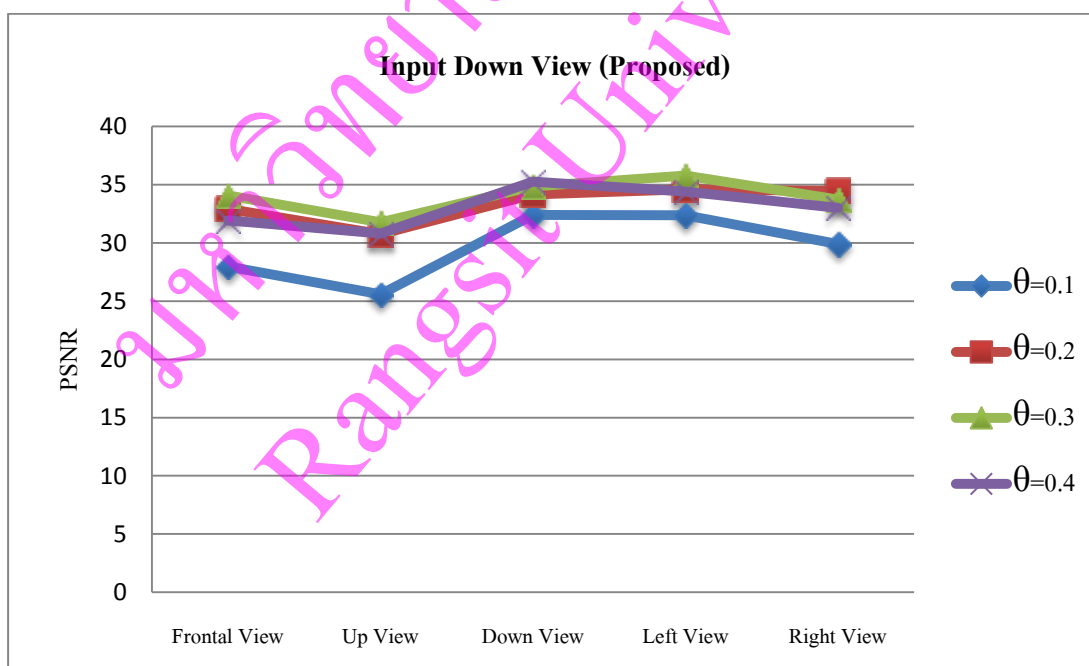
รูปที่ 4.62 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Up View)



รูปที่ 4.63 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Up View)



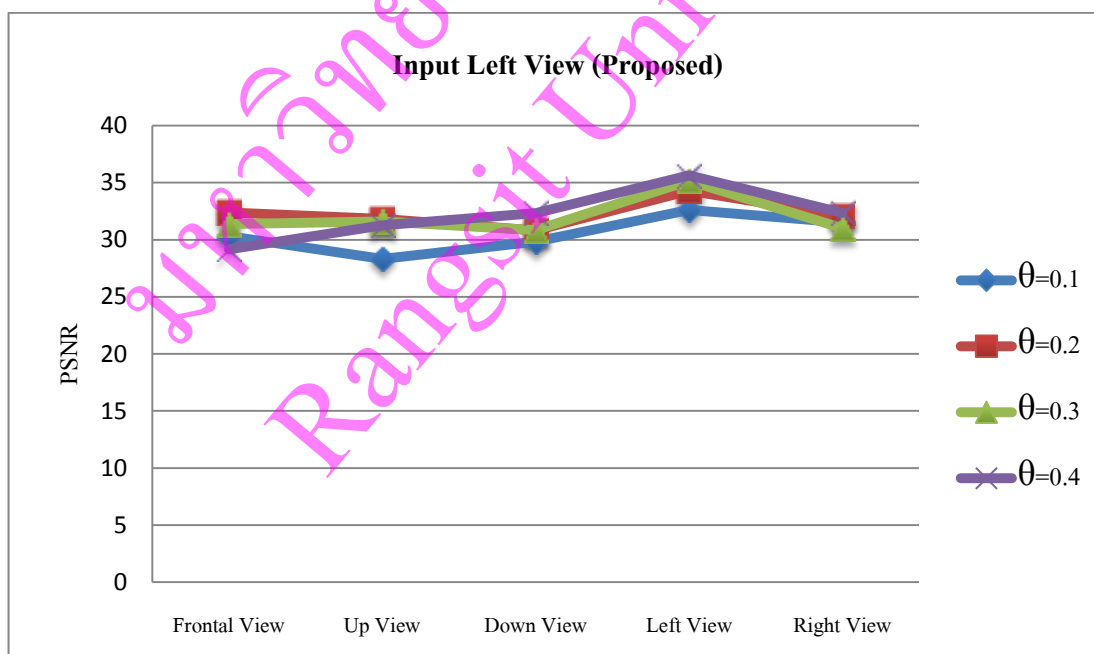
รูปที่ 4.64 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Down View)



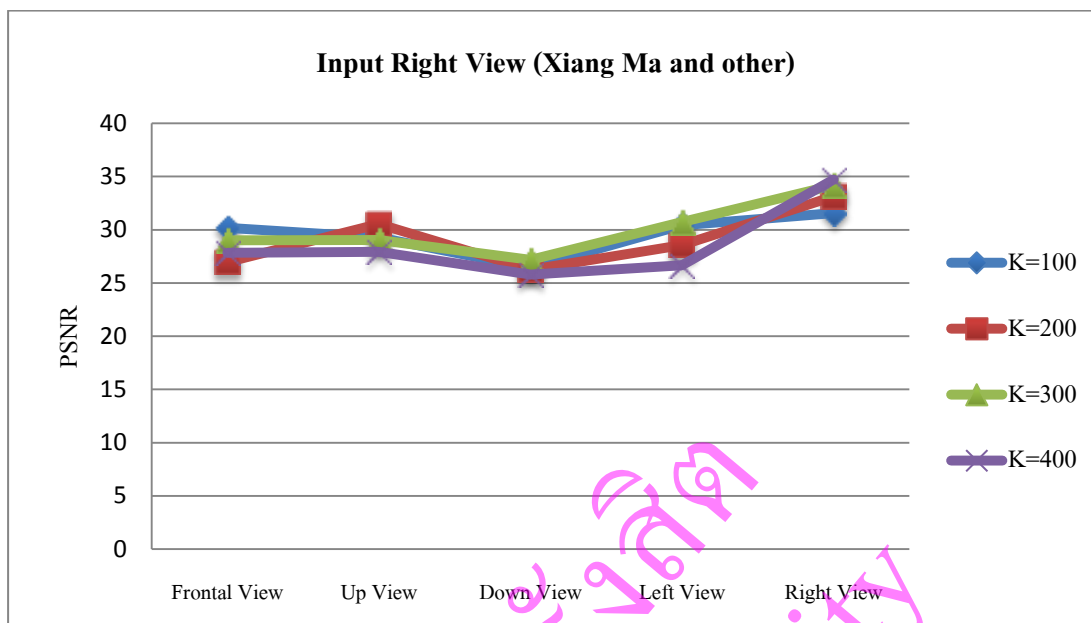
รูปที่ 4.65 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Down View)



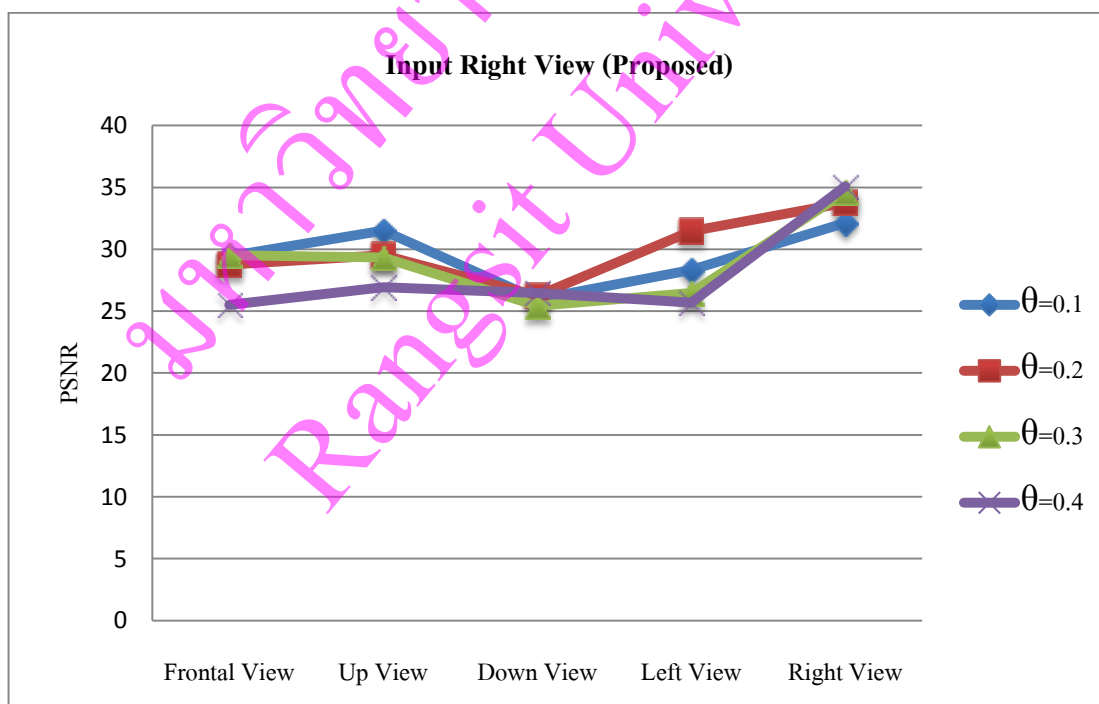
รูปที่ 4.66 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Left View)



รูปที่ 4.67 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Left View)



รูปที่ 4.68 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Right View)



รูปที่ 4.69 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Right View)

ผลการทดลองที่ 3 เป็นการทดลองโดยใช้ภาพใบหน้าของฐานข้อมูล CAS-PEAL-R1 โดยการสุ่มภาพนำเข้าจำนวนทั้งหมด 5,055 ภาพ จากซึ่งกำหนดค่า K (The Number of Nearest Neighbors) ให้มีค่าเท่ากับ 100 และ 200 กำหนดค่า θ (Threshold of Similarity) มีค่า 0.1 หรือ 10% และ 0.2 หรือ 20% จากผลการทดลองแสดงให้เห็นว่าผลการเปรียบเทียบระหว่างวิธีการที่นำเสนอ กับวิธีการ Bicubic และ วิธีการของ Xiang Ma พบว่า วิธีการที่นำเสนอ สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง จากตารางที่ 4.5 นำเข้าภาพหน้าตรง สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง โดยกำหนดค่า $K = 100$ และ $\theta = 0.1$ และจากตารางที่ 4.6 – 4.9 โดยการนำเข้าภาพใบหน้าในมุมมองหน้าแยง หน้าก้ม ใบหน้าด้านซ้าย และใบหน้าที่ด้านขวา ตามลำดับ วิธีการที่นำเสนอสามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง โดยกำหนดค่า $K = 200$ และ $\theta = 0.2$ โดยรูปที่ 4.60 – 4.69 แสดงการเปรียบเทียบค่า PSNR ของค่า K และ θ โดยค่า K มีค่า 100, 200, 300 และ 400 ค่า θ มีค่า 0.1, 0.2, 0.3 และ 0.4 ผลการเปรียบเทียบแสดงให้เห็นว่า ภาพใบหน้าที่สร้างคืนได้ในมุมมองเดียวกับภาพที่นำเข้าระบบ จะให้ประสิทธิภาพและผลดีกว่ามุมมองอื่นๆ ที่สร้างขึ้นมาได้

ตารางที่ 4.5 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100$, $\theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	32.49	32.89
Up view	26.98	29.94
Down view	26.59	26.80
Left view	31.13	32.47
Right view	28.40	30.19

ตารางที่ 4.6 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200$, $\theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	30.04	31.30
Up view	33.64	34.18
Down view	28.74	28.88
Left view	32.92	34.68
Right view	31.36	31.75

ตารางที่ 4.7 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	29.61	32.93
Up view	28.34	30.75
Down view	33.20	34.17
Left view	34.61	34.68
Right view	32.70	34.47

ตารางที่ 4.8 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	31.67	32.39
Up view	30.55	31.84
Down view	28.19	30.82
Left view	33.44	34.29
Right view	32.28	32.30

ตารางที่ 4.9 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	26.98	28.76
Up view	30.54	30.56
Down view	26.23	26.24
Left view	28.56	31.46
Right view	33.12	33.75

4.3.2.2 การฮาตุชินเนสภาพใบหน้าที่มีหลายมุมมองจากฐานข้อมูลใบหน้าคนไทย

การทดลองโดยใช้ภาพจากฐานข้อมูลภาพใบหน้าคนไทย นำเข้าภาพใบหน้า มุมมองใดมุมมองหนึ่งซึ่งเป็นภาพความละเอียดต่ำโดยมีขนาด 32x24 จุดภาพ ผลลัพธ์ที่ได้คือ การ สร้างคืนภาพใบหน้าที่มีความละเอียดสูงของ 5 มุมมอง ได้แก่ มุมมองใบหน้าตรง มุมมองใบหน้าเงย มุมมองใบหน้าที่้ม มุมมองใบหน้าที่้นซ้ายและมุมมองใบหน้าที่้นขวา โดยมีขนาด 128x96 จุดภาพ ซึ่งเป็นภาพที่ขยายจากภาพความละเอียดต่ำ 4 เท่า



รูปที่ 4.70 รูปภavnนำเข้าใบหน้าที่ตรง ขนาด 32x24 จุดภาพ



รูปที่ 4.71 รูปภavnนำเข้าใบหน้าที่เงย ขนาด 32x24 จุดภาพ



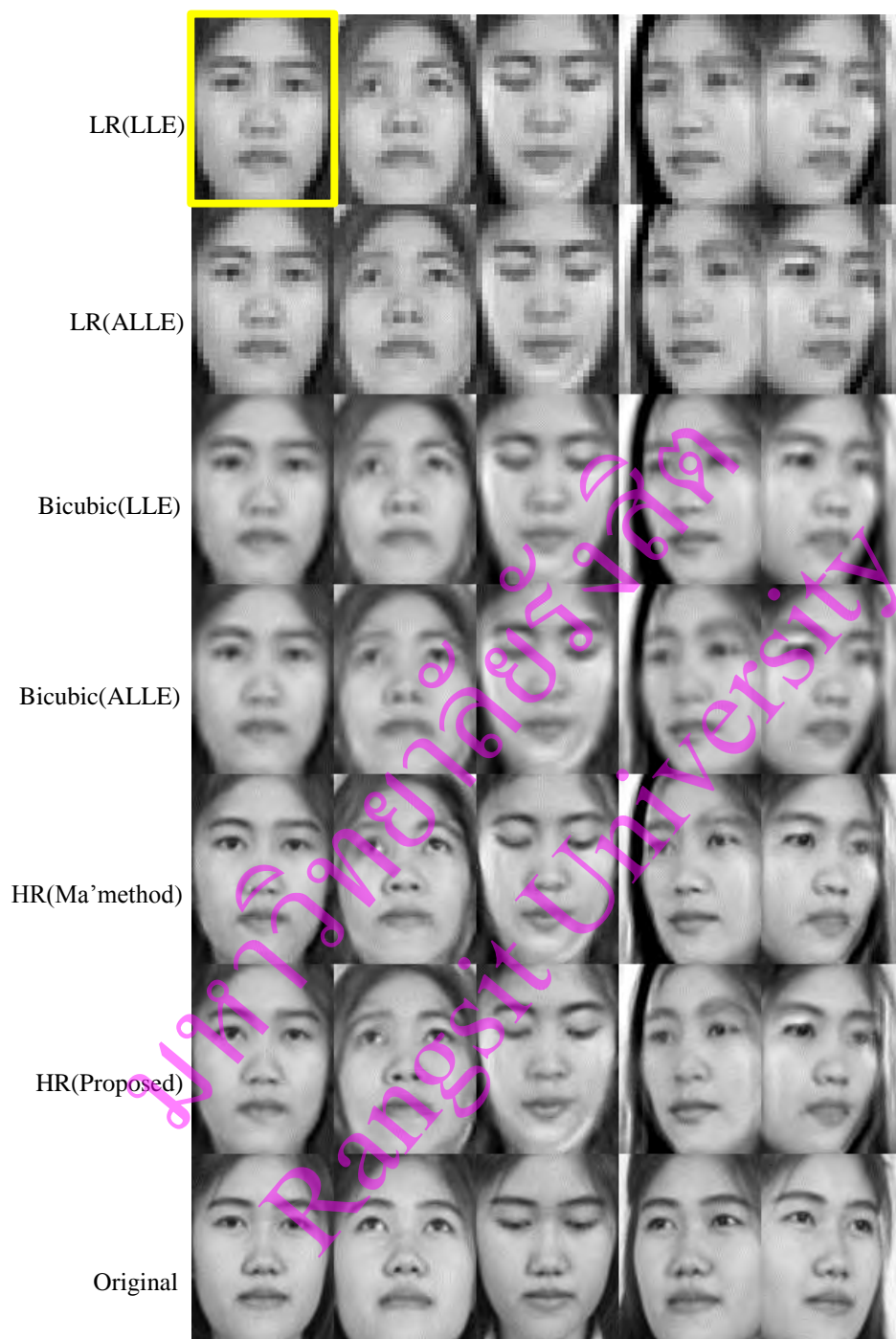
รูปที่ 4.72 รูปภาพนำเข้าไปบนหน้ากัม ขนาด 32x24 จุดภาพ



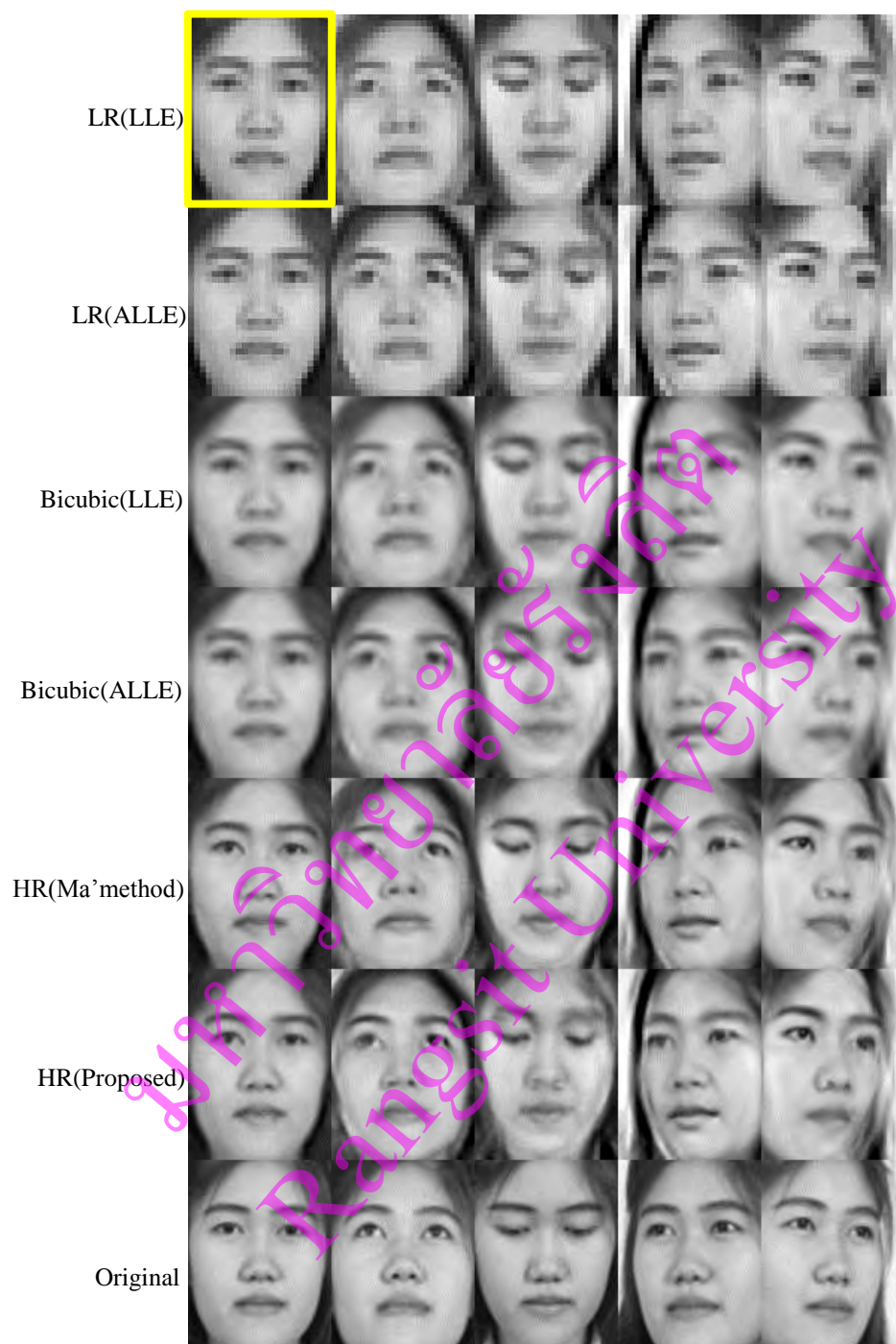
รูปที่ 4.73 รูปภาพนำเข้าไปบนหน้าด้านซ้าย ขนาด 32x24 จุดภาพ



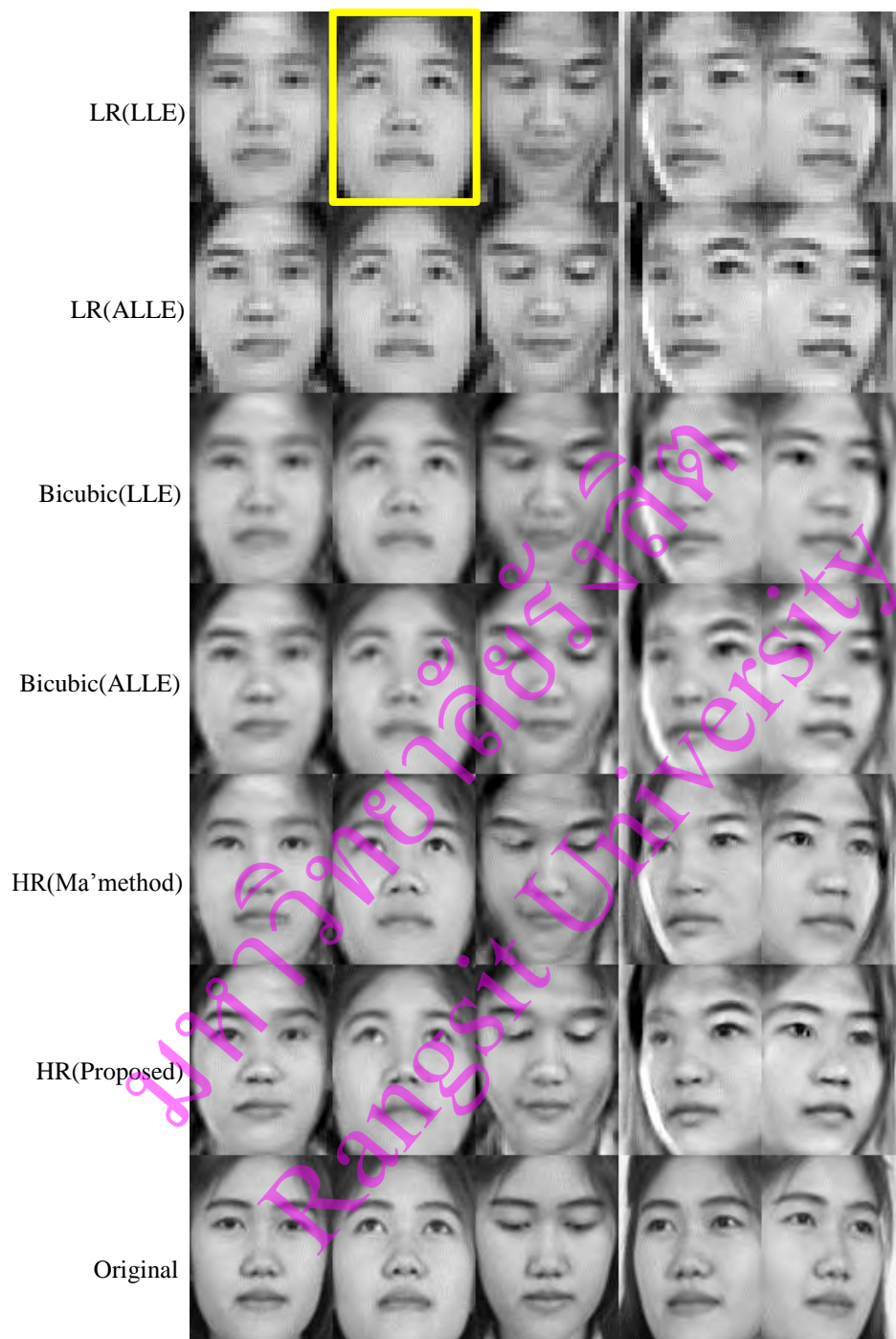
รูปที่ 4.74 รูปภาพนำเข้าไปบนหน้าด้านขวา ขนาด 32x24 จุดภาพ



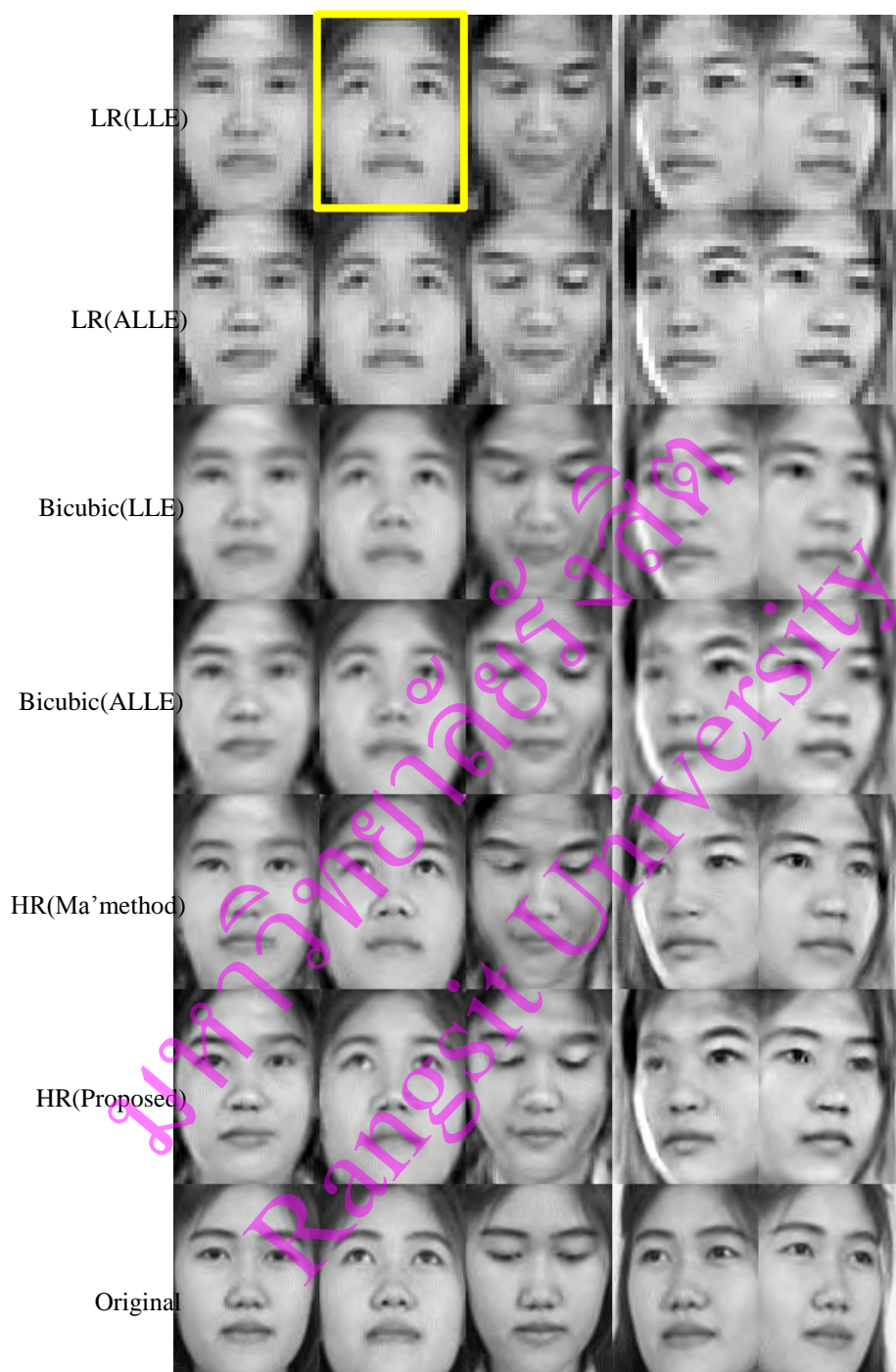
รูปที่ 4.75 รูปฮาซุชิเนสภาพใบหลายมุมมองจากภาพหน้าเข้าใบหน้าตรง
(กำหนด $K = 100$, $\theta = 0.1$)



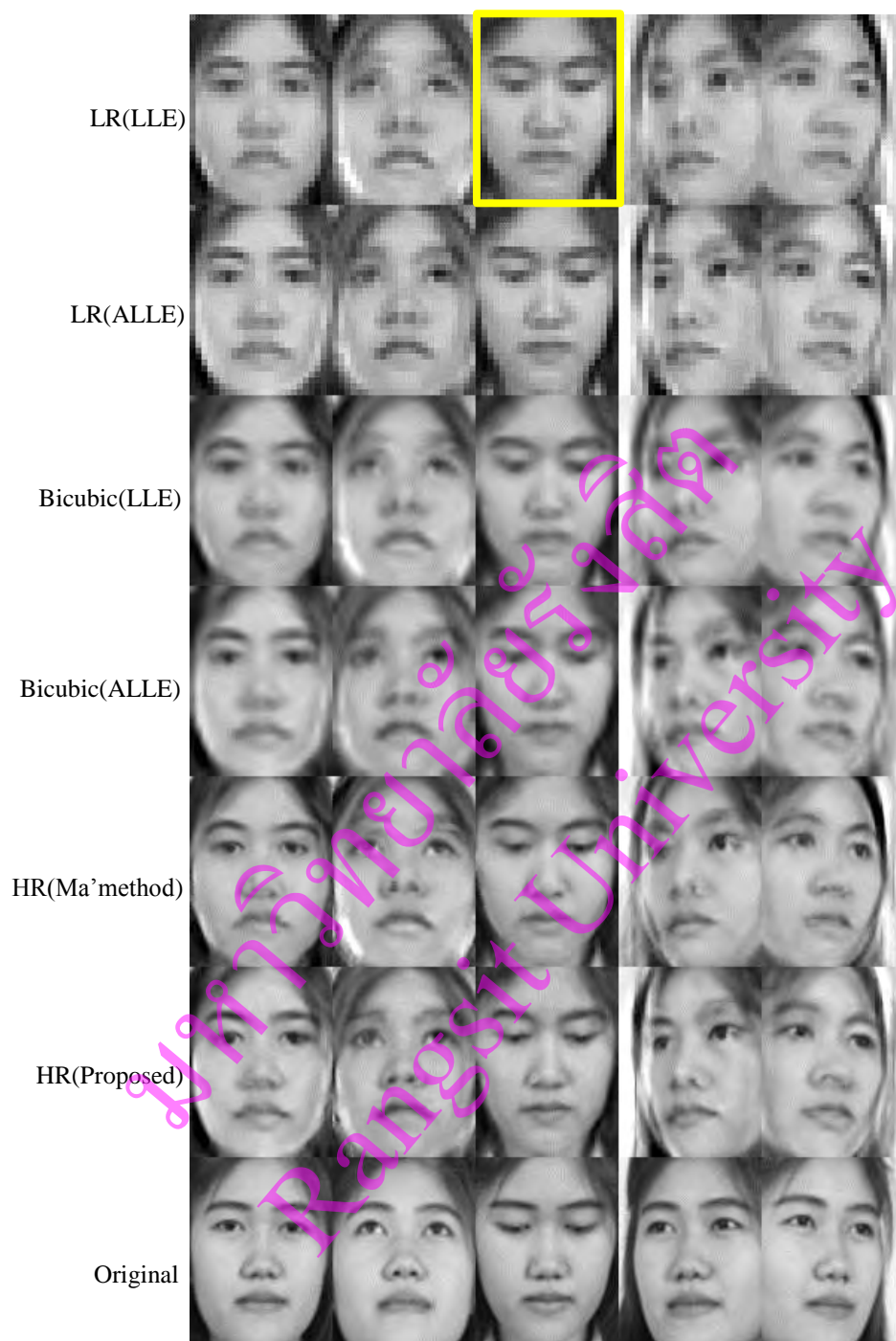
รูปที่ 4.76 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปบนหน้าตรง
(กำหนด $K = 200$, $\theta = 0.2$)



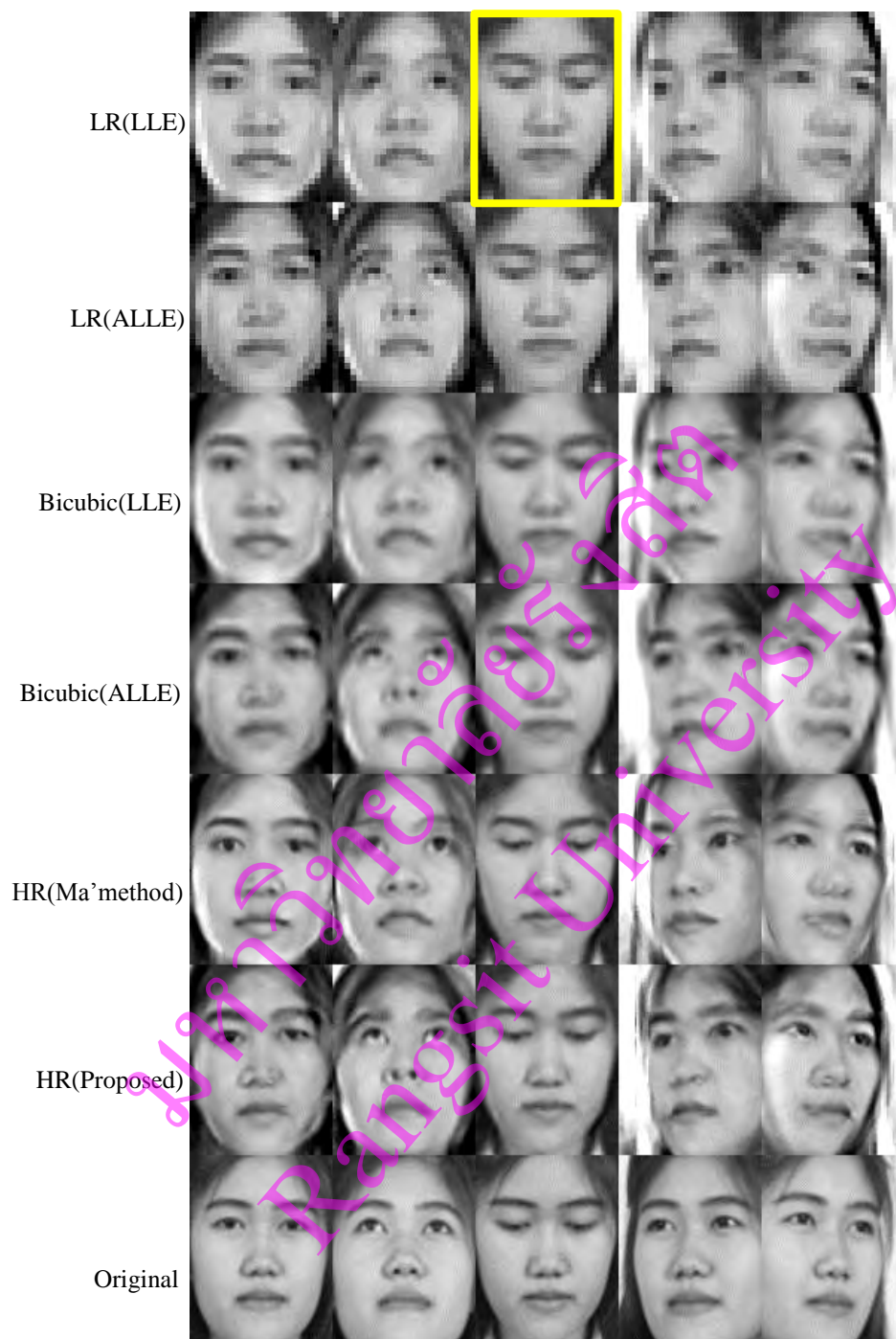
รูปที่ 4.77 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปหน้าเงย
(กำหนด $K = 100$, $\theta = 0.1$)



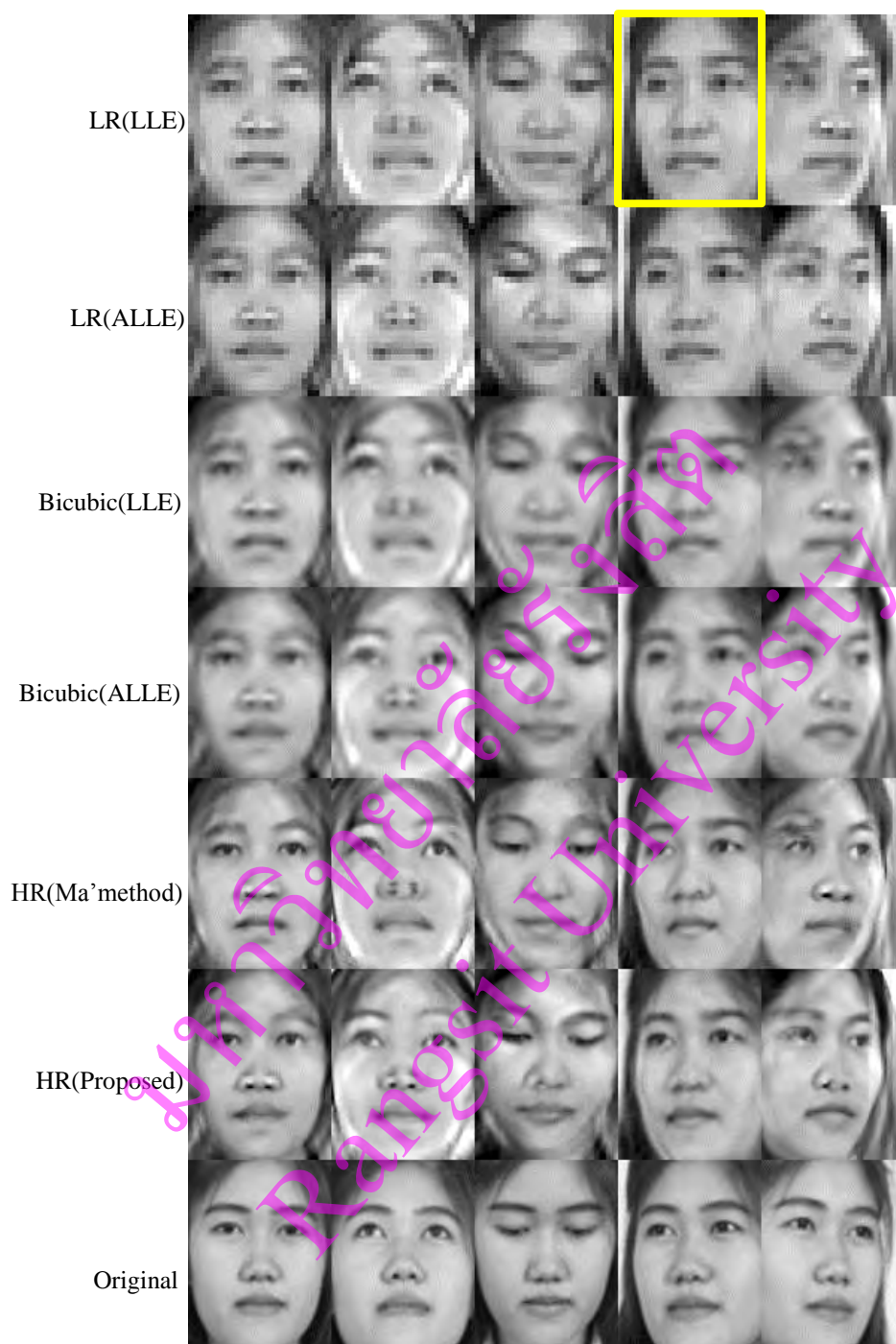
รูปที่ 4.78 รูปฮาตุชินเนสภาพโบหลายมุมมองจากภาพนำเข้าไปหน้าเงย
(กำหนด $K = 200$, $\theta = 0.2$)



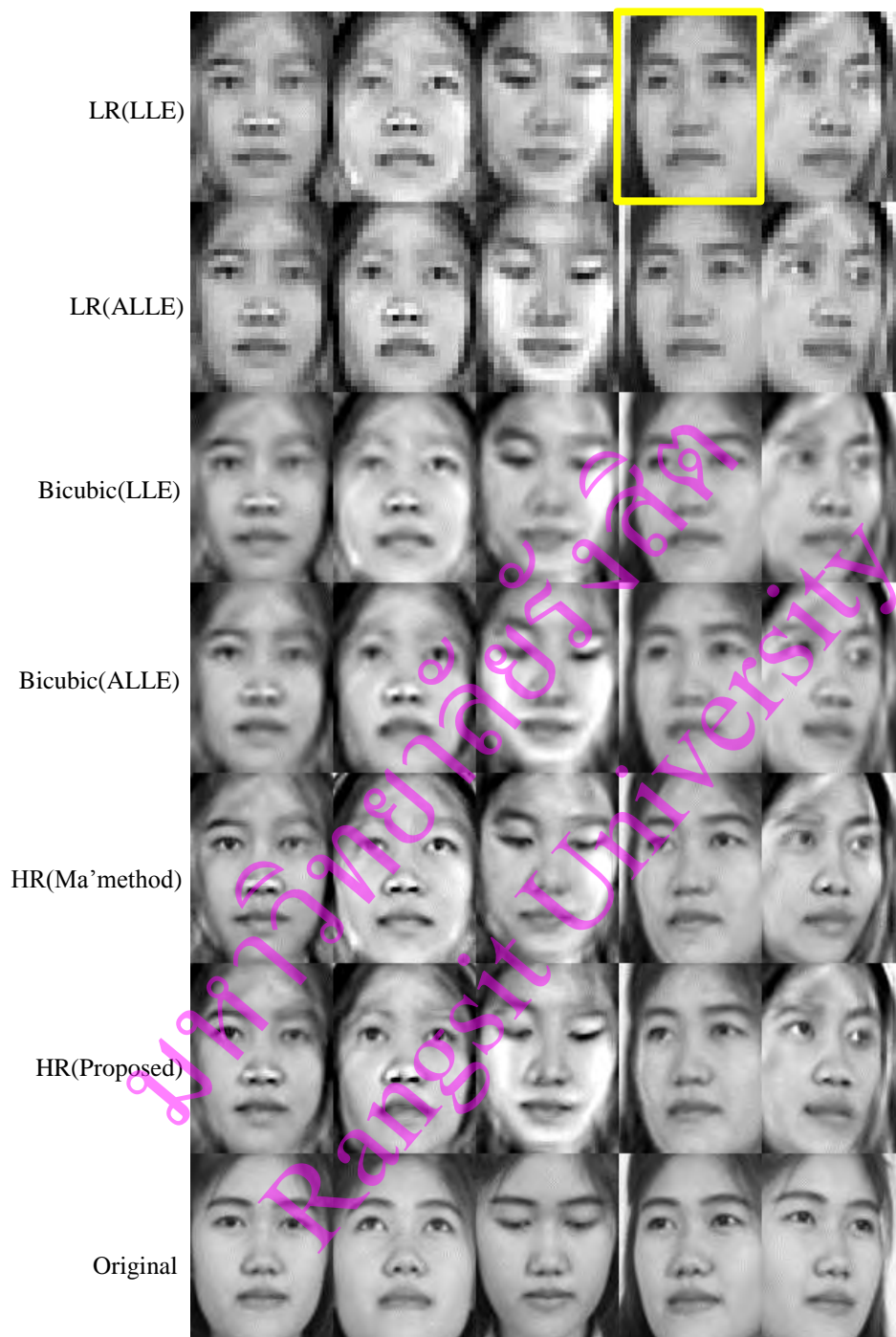
รูปที่ 4.79 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปหน้ากัม
(กำหนด $K = 100$, $\theta = 0.1$)



รูปที่ 4.80 รูปฮาตุชินเนสภาพใบหลายมุมมองจากภาพนำเข้าไปบน้ำกัม
(กำหนด $K = 200$, $\theta = 0.2$)



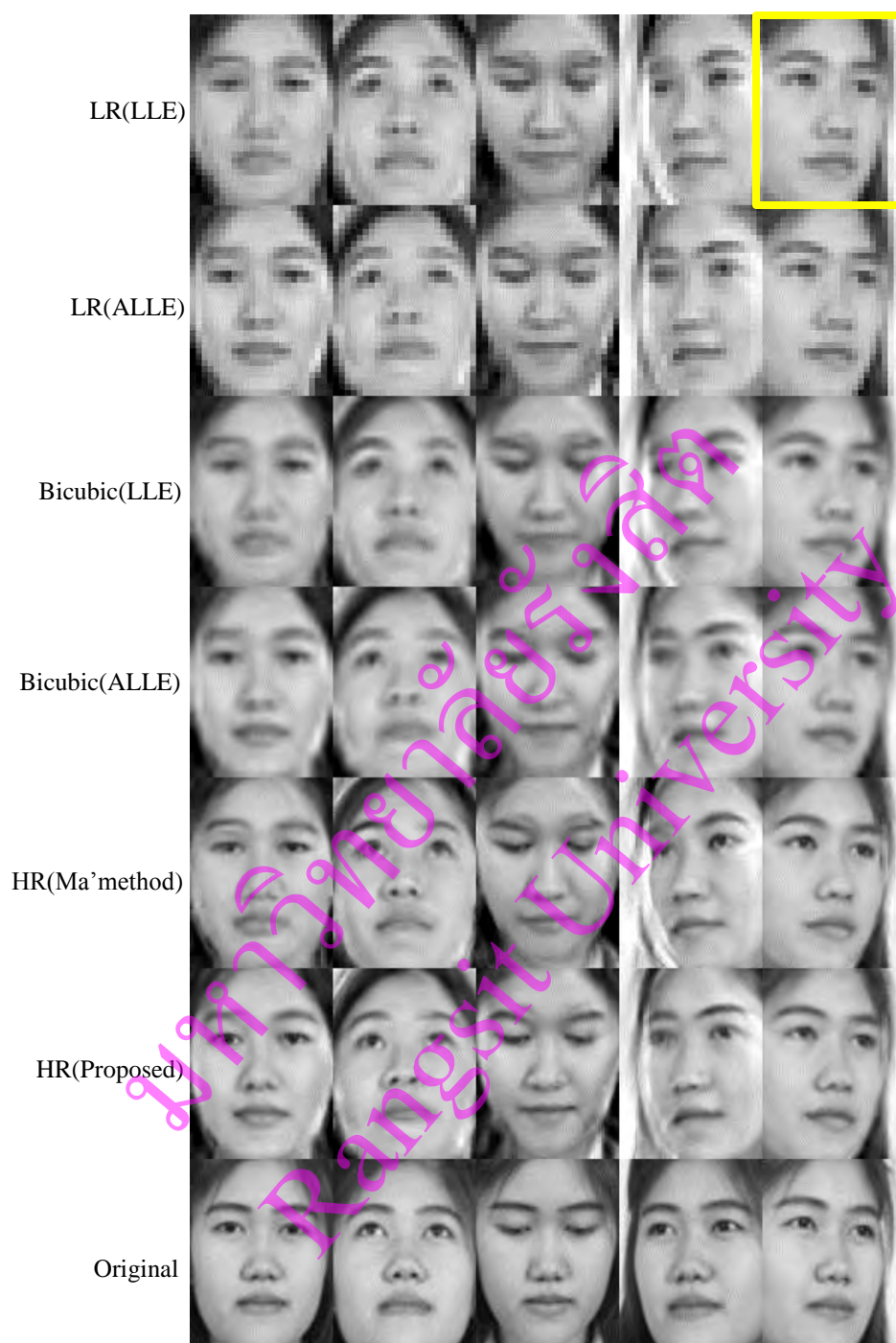
รูปที่ 4.81 รูปฮาตุชินเนสภาพใบหลายมุมมองจากภาพนำเข้าไปหน้าด้านซ้าย
(กำหนด $K = 100$, $\theta = 0.1$)



รูปที่ 4.82 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าไปหน้าด้านซ้าย
(กำหนด $K = 200$, $\theta = 0.2$)



รูปที่ 4.83 รูปฮาตุชินเนสภาพโอบหลายมุมมองจากภาพนำเข้าไบหน้าด้านขวา
(กำหนด $K = 100$, $\theta = 0.1$)



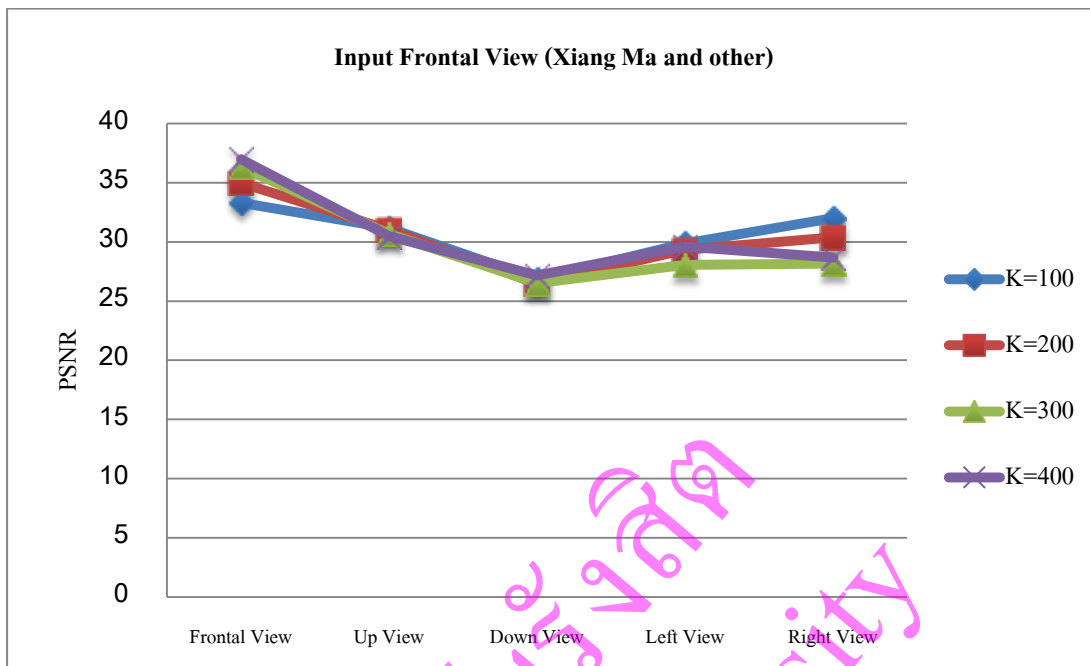
รูปที่ 4.84 รูปฮาตุชินเนสภาพโบลหลายมุมมองจากภาพนำเข้าใบหน้าด้านขวา
(กำหนด $K = 200$, $\theta = 0.2$)



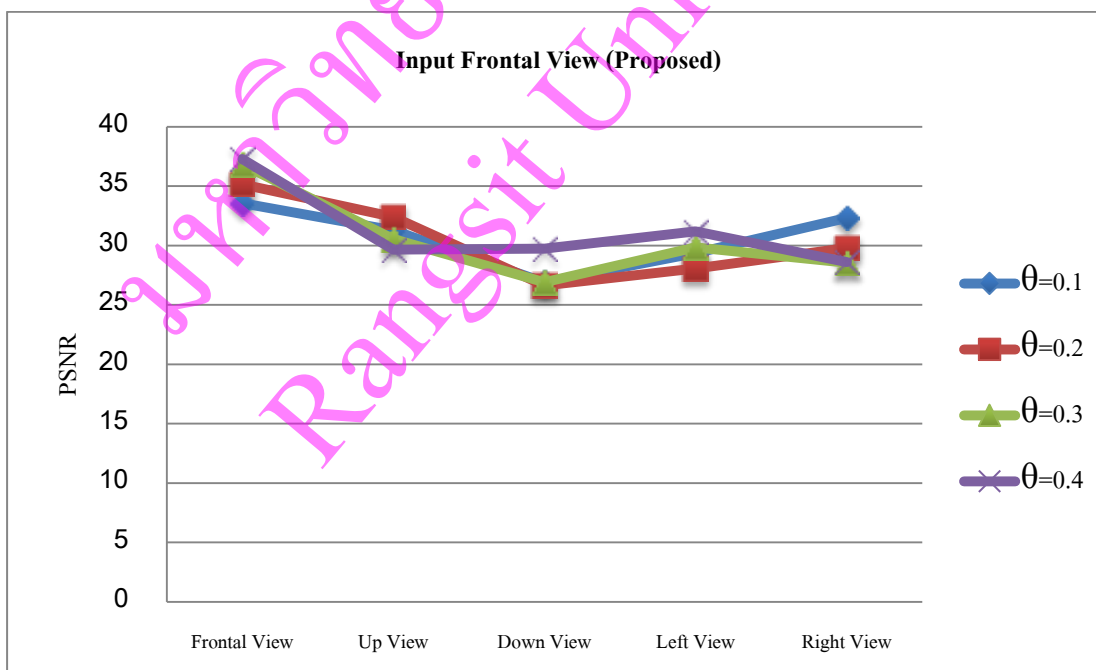
รูปที่ 4.85 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า K ที่ต่างกัน (Multiview)



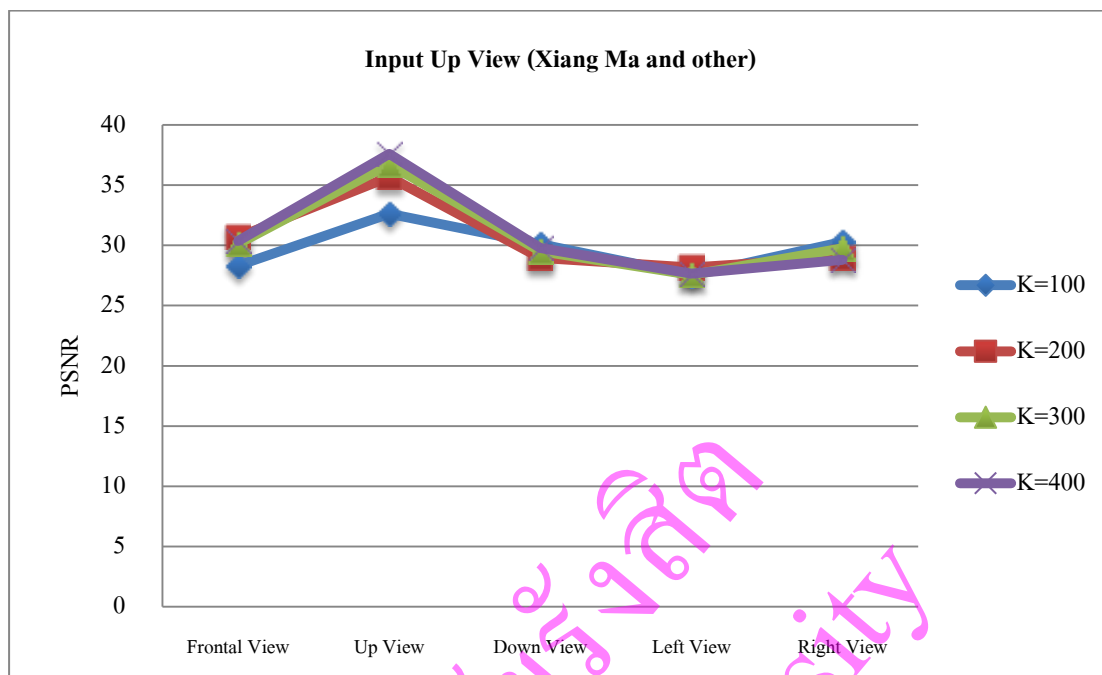
รูปที่ 4.86 แสดงการเปรียบเทียบผลของการฮาตุชินเนส
โดยกำหนดค่า θ ที่ต่างกัน (Multiview)



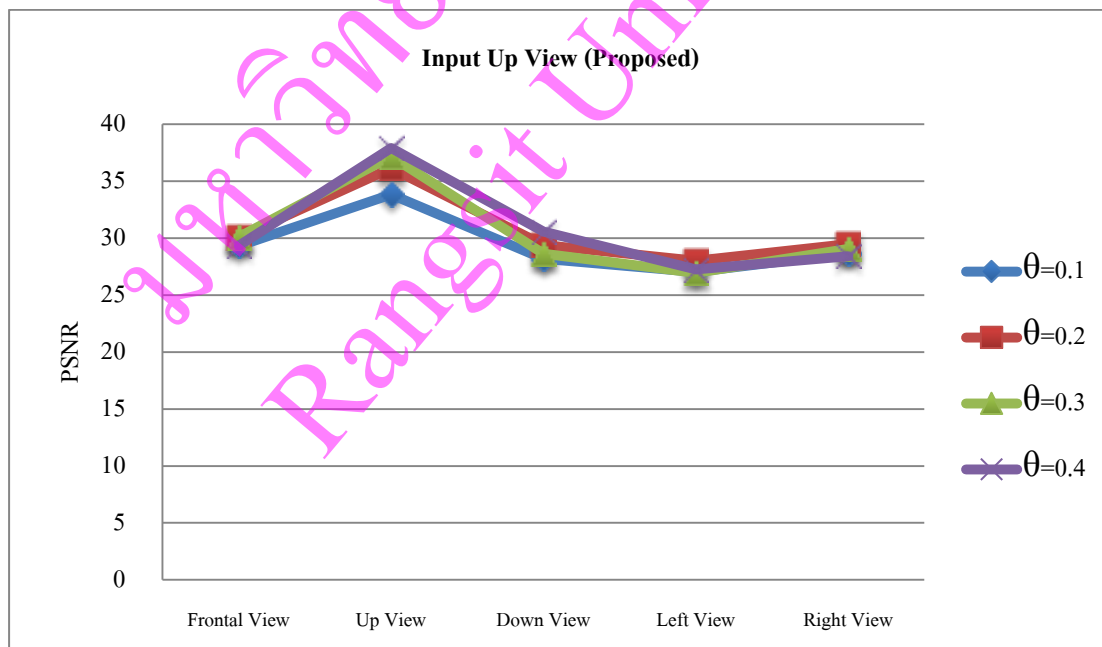
รูปที่ 4.87 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Frontal View)



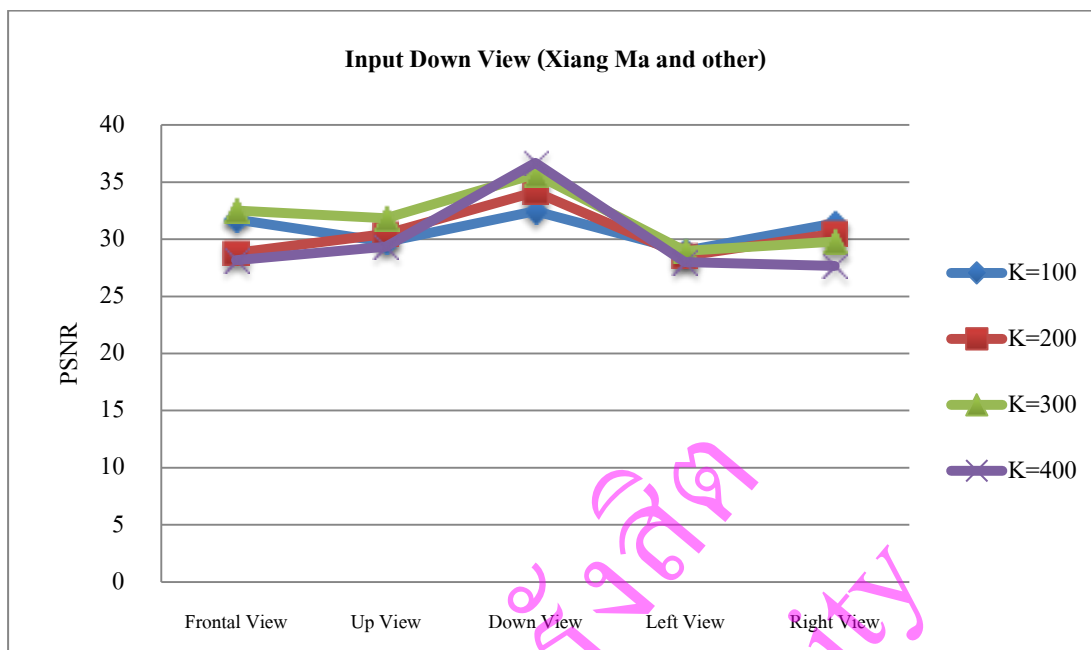
รูปที่ 4.88 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Frontal View)



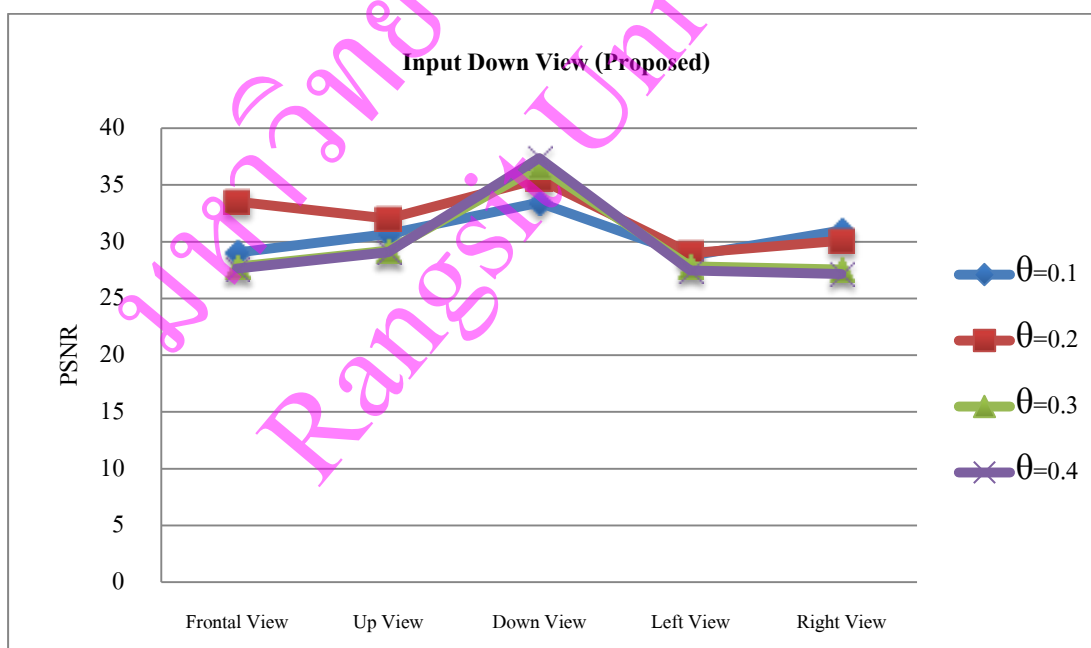
รูปที่ 4.89 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Up View)



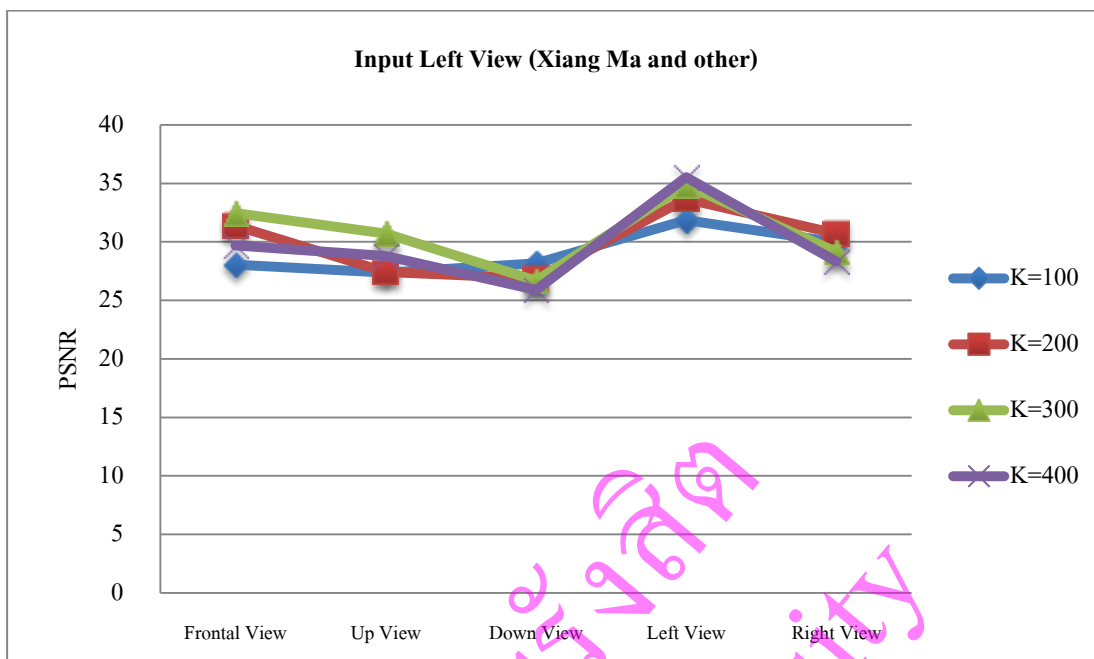
รูปที่ 4.90 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Up View)



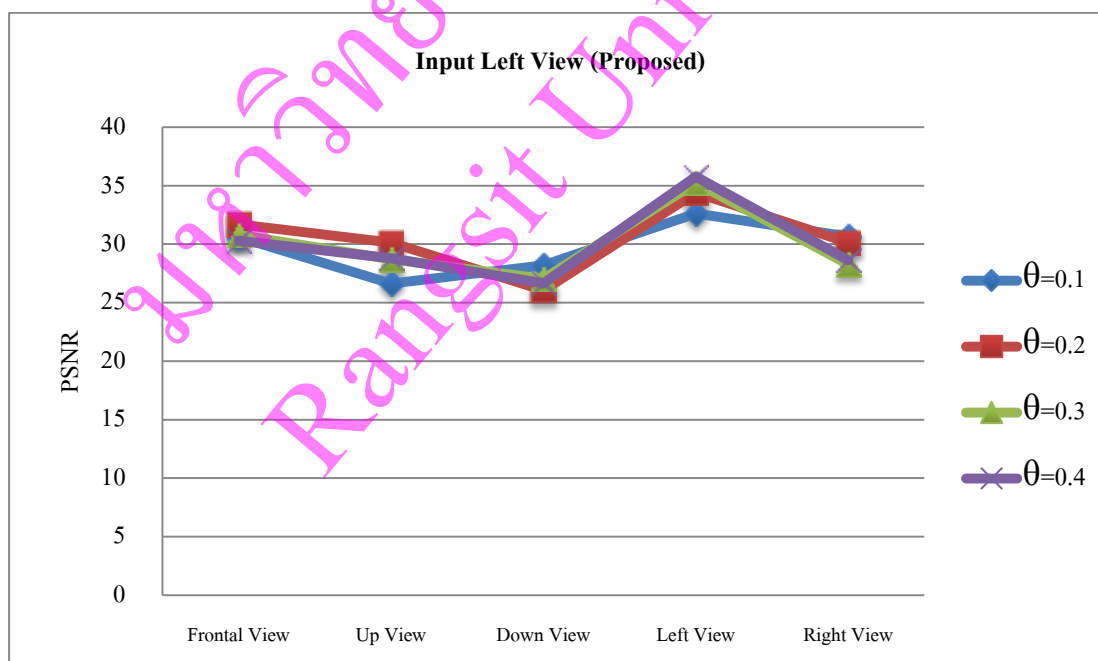
รูปที่ 4.91 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Down View)



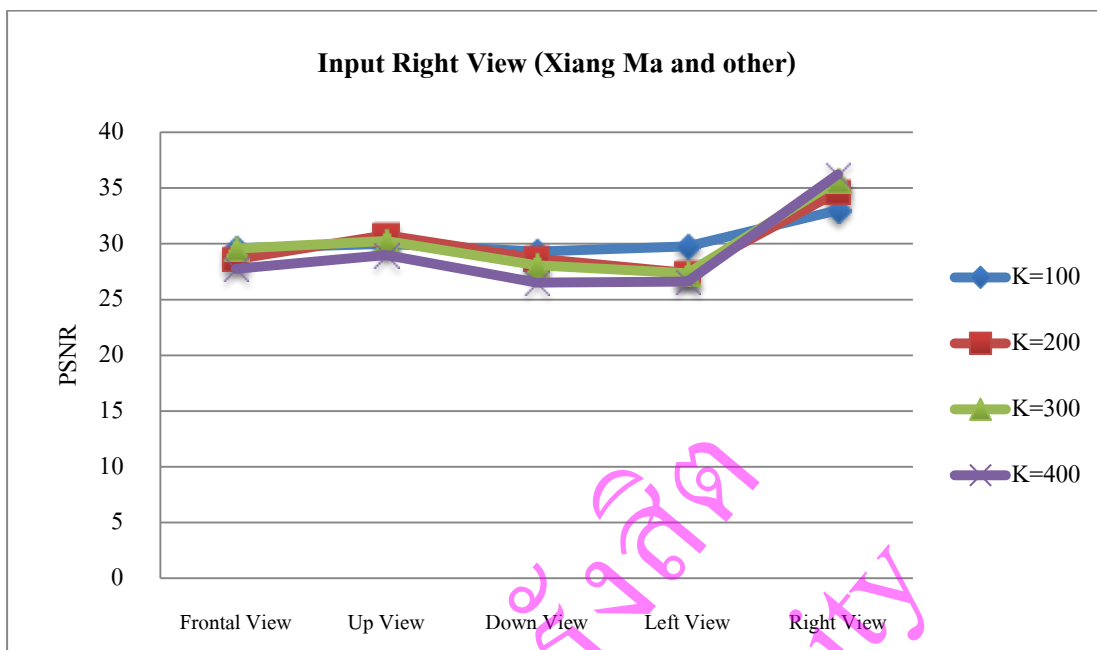
รูปที่ 4.92 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Down View)



รูปที่ 4.93 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Left View)



รูปที่ 4.94 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Left View)



รูปที่ 4.95 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า K ที่ต่างกัน (Input Right View)



รูปที่ 4.96 กราฟแสดงการเปรียบเทียบค่า PSNR โดยกำหนดค่า θ ที่ต่างกัน (Input Right View)

ผลการทดลองที่ 4 เป็นการทดลองโดยใช้ภาพใบหน้าของฐานข้อมูลใบหน้าคนไทย โดยสุ่มภาพใบหน้าจากจำนวนทั้งหมด 5,100 ภาพ ซึ่งกำหนดค่า K (The Number of Nearest Neighbors) ให้มีค่าเท่ากับ 100 และ 200 กำหนดค่า θ (Threshold of Similarity) มีค่า 0.1 หรือ 10% และ 0.2 หรือ 20% จากผลการทดลองแสดงให้เห็นว่าผลการเปรียบเทียบระหว่างวิธีการที่นำเสนอกับวิธีการ Bicubic และ วิธีการของ Xiang Ma พบว่า วิธีการที่นำเสนอ สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง จากตารางที่ 4.10, 4.11 และ 4.13 นำเข้าภาพหน้าตรง หน้าเงย และใบหน้าด้านซ้าย สามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง โดยกำหนดค่า $K = 100$ และ $\theta = 0.1$ และจากตารางที่ 4.12 และ 4.14 โดยการนำเข้าภาพใบหน้าในมุมมองหน้าก้ม และใบหน้าด้านขวา วิธีการที่นำเสนอสามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าทุกมุมมอง โดยกำหนดค่า $K = 200$ และ $\theta = 0.2$ โดยรูปที่ 4.87 – 4.96 แสดงการเปรียบเทียบค่า PSNR ของค่า K และ θ โดยค่า K มีค่า 100, 200, 300 และ 400 ค่า θ มีค่า 0.1, 0.2, 0.3 และ 0.4 ผลการเปรียบเทียบแสดงให้เห็นว่า ภาพใบหน้าที่สร้างคืนได้ในมุมมองเดียวกับภาพที่นำเข้าระบบ จะให้ประสิทธิภาพและผลดีกว่ามุมมองอื่นๆ ที่สร้างขึ้นมาได้

ตารางที่ 4.10 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100, \theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	33.31	33.57
Up view	31.16	31.29
Down view	26.85	26.92
Left view	29.88	29.90.
Right view	32.02	32.32

ตารางที่ 4.11 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100, \theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	28.32	29.30
Up view	32.62	33.82
Down view	28.18	30.09
Left view	27.53	28.03
Right view	28.99	29.47

ตารางที่ 4.12 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	28.82	33.54
Up view	30.47	32.03
Down view	34.12	35.54
Left view	28.53	28.99
Right view	30.48	31.00

ตารางที่ 4.13 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=100, \theta = 0.1$)

View	Xiang Ma ,et al	Proposed
Frontal view	28.06	30.61
Up view	27.42	30.14
Down view	28.17	28.20
Left view	33.66	34.41
Right view	29.98	30.67

ตารางที่ 4.14 ตารางแสดงค่า PSNR ในการสร้างคืนภาพความละเอียดสูง ($K=200, \theta = 0.2$)

View	Xiang Ma ,et al	Proposed
Frontal view	28.63	29.31
Up view	29.64.	29.99
Down view	28.68	29.26
Left view	26.73	26.76
Right view	34.64	35.17

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ผลของงานวิจัยนี้ได้ศึกษาถึงวิธีการสร้างภาพใบหน้าความละเอียดสูงขึ้นมาใหม่ จากภาพใบหน้าความละเอียดต่ำ ซึ่งเรียกว่า ฮาลูซิเนชันภาพใบหน้า (Face Hallucination) โดยนำเอาวิธีการ Learning Based มาใช้ในงานวิจัยนี้ โดยได้เลือกใช้ฐานข้อมูลภาพมาตรฐาน CAS-PEAL-R1 ซึ่งงานวิจัยนี้ได้เสนอวิธีการใหม่ ในการสร้างคืนภาพ โดยนำเอาวิธีการในการคำนวณหาหน้าหนัก (W) โดยวิธี LLE (Locally Linear Embedding) มาประยุกต์ใช้เพื่อที่จะพัฒนาให้มีประสิทธิภาพในการสร้างคืนภาพให้สูงขึ้น ซึ่งวิธีการที่เสนอเรียกว่า Adaptive Locally Linear Embedding (ALLE) โดยได้ทำการทดลองกับภาพใบหน้าในมุมมองเดียว และทำการทดลองกับภาพใบหน้าที่มีหลายมุมมอง และในงานวิจัยนี้ได้สร้างฐานข้อมูลภาพใบหน้าคนไทยในมุมมองต่างๆ จำนวนคนละ 15 ภาพ จำนวนคนทั้งสิ้น 1,020 คน รวมได้จำนวนฐานข้อมูลภาพใบหน้าคนไทยที่จะได้นำไปใช้ในงานวิจัยในอนาคต 15,300 ภาพ

ผลการทดลองของงานวิจัยนี้ โดยการทดลองกับฐานข้อมูลภาพ CAS-PEAL-R1 ทดลองใช้กับฐานข้อมูลภาพใบหน้าของคนไทย พบว่า กรณีที่นำเข้าภาพใบหน้าในมุมมองใด มุมมองหนึ่ง แล้ววิธีการที่นำเสนอสามารถสร้างคืนภาพใบหน้าในมุมมองนั้น ได้ผลของการทดลองมีประสิทธิภาพสูงกว่า วิธีการ Bicubic Interpolation และวิธีการของ Xiang Ma ซึ่งสามารถวัดผลและประเมินผลได้ด้วยสายตา และค่าของ PSNR

ผลการทดลองของงานวิจัย ในกรณีการสร้างคืนภาพความละเอียดสูงที่มีหลายมุมมอง โดยการทดลองกับฐานข้อมูลภาพ CAS-PEAL-R1 และ ทดลองใช้กับฐานข้อมูลภาพใบหน้าของคนไทย พบว่าการสร้างคืนภาพใบหน้าที่มีหลายมุมมองจากภาพนำเข้าในมุมมองใดมุมมองหนึ่ง โดยวิธีการที่นำเสนอสามารถสร้างคืนภาพใบหน้าในมุมมองต่างๆ ได้อย่างมีประสิทธิภาพ ทำให้ความชัดของรูปสูงขึ้น

5.2 ข้อเสนอแนะ

ในงานวิจัยนี้ เป็นงานวิจัยที่ต้องใช้ฐานข้อมูลภาพใบหน้าในการทดลอง ซึ่งกระบวนการก่อนหน้าการทดลอง (Preprocessing) มีความสำคัญอย่างมากต่อข้อมูลภาพ เนื่องจากว่า การจัดวางตำแหน่งของภาพใบหน้า (Face Alignment) ในงานวิจัยนี้ยังเป็นวิธีการที่ผู้วิจัยต้องทำเองด้วยระบบมือ อาจจะก่อให้เกิดปัญหาตำแหน่งของภาพที่ผ่านการระบุตำแหน่งและผ่านการตัดภาพอาจจะคลาดเคลื่อน ไม่ตรงตามตำแหน่งที่ต้องการ ซึ่งจะมีผลต่อการนำเอาภาพที่ได้ไปประมวลผลภาพต่อไป

มหาวิทยาลัยรังสิต
Rangsit University

บรรณานุกรม

- Baker Simon and Kanade Takeo. "Hallucinating Faces." *IEEE Transactions on Image Processing*. (2000) : 83-88.
- Baker Simon and Kanade Takeo. "Limits on Super-Resolution and How to Break Them." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24 (2002) : 1167-1183.
- Bing-Yu, Sun, Zhang Xiao-Ming, Li Jiuyong and Mao Xue-Min. "Feature Fusion Using Locally Linear Embedding for Classification." *IEEE Transactions on Neural Networks*. 21 (2010) : 163-168.
- CeLiu ,Heung-YeungShum and, Zhang Chang-Shui. "A Two-Step Approach to Hallucinating Faces:Global Parametric Model and Local Nonparametric Model." *IEEE Transactions on Image Processing*. 1 (2001) : 192-198.
- C. Jung, L. Jiao, B. Liu, M. Gong. "Position-patch based face hallucination using convex optimization." *IEEE Signal Processing Letters*. 18 (2011) : 367-370.
- Elad Michael and Feuer Arie. "Super-resolution reconstruction of image sequences." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 21 (1999) : 817 – 834.
- Freeman W. T. and Pasztor E. C. "Learning Low Level Vision." *Proceedings of the 7th IEEE International Conference on Computer Vision*. (1999) : 1182-1189.
- Gao, Guangwei, and Jian Yang. "A novel sparse representation based framework for face image super-resolution." *Neurocomputing*. 134 (2014) : 92-99.
- Hardie, R. C., K. J. Barnard and E. E. Armstrong. "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images." *IEEE Transactions on Image Processing*. 6 (1997) : 1621-1633.
- Hua Shen and Shutao Li. "Hallucinating Faces by Interpolation and Principal Component Analysis." *Proceedings of the Second International Symposium on Computational Intelligence and Design*. 1 (December 2009) : 295-298.

บรรณานุกรม(ต่อ)

- H. Chang, D.Y. Yeung and Y.M. Xiong. "Super-resolution through Neighbor Embedding." *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1 (July 2004).
- Jia Kui and Gong Shaogang. "Generalized Face Super-Resolution." *IEEE Transactions on Image Processing*. 17 (2008) : 873-886.
- Jian Sun ,Zongben Xu and, Heung-Yeung Shum. "Image super-resolution using gradient profile prior." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (June 2008) : 1-8.
- Kimmel R. "Demosaiicing: image reconstruction from color CCD samples." *IEEE Transactions on Image Processing*. 8 (1999) : 1221-1228.
- Lin Zhouchen and Shum Heung-Yeung. "Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 26 (2004) : 83-97.
- Liu Wei, Lin Dahua, and Tang Xiaoou. "Neighbor Combination and Transformation for Hallucinating Faces." *Proceedings of the IEEE International Conference on Multimedia and Expo*. (July 2005).
- Moon Gi, K. and S. Chaudhuri. "Super-resolution image reconstruction." *Signal Processing Magazine*. 20 (2003) : 19-20.
- Nuno-Maganda M. A. and Arias-Estrada M. O. "Real-time FPGA-based architecture for bicubic interpolation: an application for digital image scaling." *Proceedings of the International Conference on Reconfigurable Computing and FPGAs*. (September 2005).
- Park Sung Won, M. Savvides. "Robust Super-Resolution of Face Images by Iterative Compensating Neighborhood Relationships." *Proceedings of the Biometrics Symposium*. (September 2007) : 1-5.

บรรณานุกรม(ต่อ)

- Roweis, Sam T., and Lawrence K. Saul. "Nonlinear dimensionality reduction by locally linear embedding." *Science*. 22 (2000) : 2323-2326.
- Wang Xiaogang and Tang Xiaoou. "Hallucinating Face by Eigentransformation." *IEEE Transactions on System, Man and Cybernetics*. 35 (2005) : 425-434.
- Wei, Fan and Yeung Dit-Yan. "Image Hallucination Using Neighbor Embedding over Visual Primitive Manifolds." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2007) : 1-7.
- W. Gao, B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H.Zhang, D.B. Zhao. "The CAS-PEAL large-scale Chinese face database and baseline evaluations." *IEEE Transactions on System Man and Cybernetics (Part A)*. 38 (2008) : 149-161.
- Xiang Ma, Junping Zhang, and Chun Qi. "Position-based face hallucination method." *Proceedings of the IEEE Conference Multimedia and Expo*. (July 2009) : 290-293.
- Xiang Ma, Junping Zhang, and Chun Qi, "Hallucinating Face by Position-Patch." *Pattern Recognition*. 43 (2010) : 2224-2236.
- Xiang Ma, H. H., Shaopeng Wang, and Chun Qi. "A Simple Approach to Multiview Face Hallucination." *IEEE Signal Processing*. 17 (2010) : 579-582.
- X. a. Q. L. Ma, Hiệp and Philips, Wilfried and Song, Huansheng and Cui, Hua. "Sparse representation and position prior based face hallucination upon classified over- complete dictionaries." *Signal Processing*. 92 (2012) : 2066-2074.
- Yang Li and Xueyin Lin. "An improved two-step approach to hallucinating faces." *Proceedings of the Third International Conference on Image and Graphics*. (December 2004) : 298-301.

บรรณานุกรม(ต่อ)

- Yu, Hu, Lam Kin-Man, Qiu Guoping, and Shen Tingzhi. "From Local Pixel Structure to Global Image Super-Resolution: A New Face Hallucination Framework." *IEEE Transactions on Image Processing*. 20 (2011) : 433-445.
- Yueting Zhuang, Jian Zhang, and Fei Wu. "Hallucinating Faces: Lph Super-Resolution and Neighbor Reconstruction for Residue Compensation." *Pattern Recognition*. 40 (2007) : 3178-3194.
- Y. Hu, K.-Man Lam, T. Shen, W. Wang. "A novel kernel-based framework for facial-image Hallucination." *Image and Vision Computing*. 29 (2011) : 219-229.
- Zhang Xiang-guang. "A New Kind of Super-Resolution Reconstruction Algorithm Based on the ICM and the Constrained Cubic Spline Interpolation." *Proceedings of the Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*. (August 2008) : 530-534.
- Zhou, Fei, Biao Wang, and Qingmin Liao, Super-resolution for face image by bilateral patches, *IET Electron*. 48 (2012) : 1125-1126.

มหาวิทยาลัยรังสิต
Rangsit University

ภาคผนวก

ภาคผนวก ก.

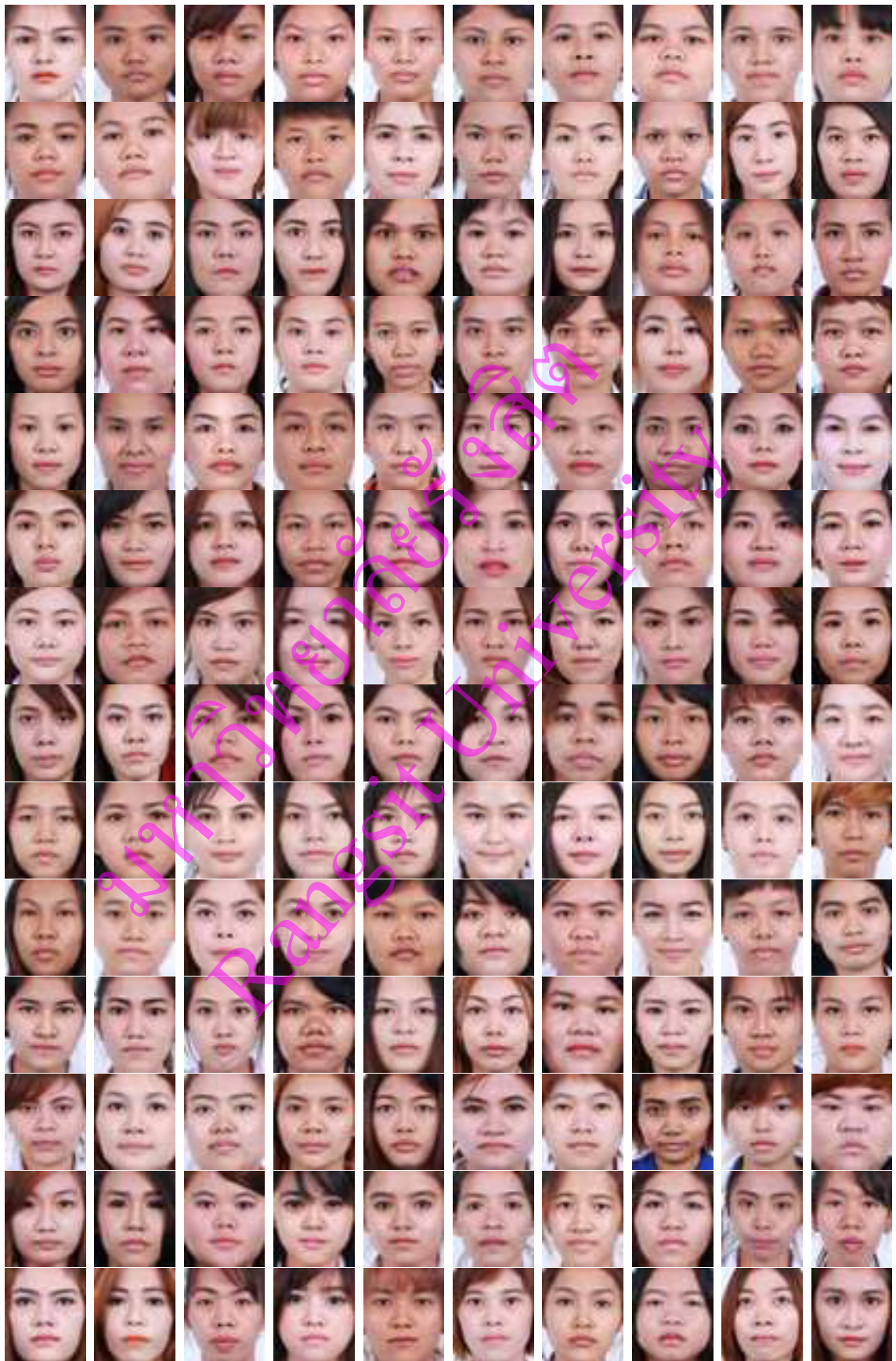
ตัวอย่างฐานข้อมูลภาพใบหน้าของคนไทย

มหาวิทยาลัยรังสิต
Rangsit University

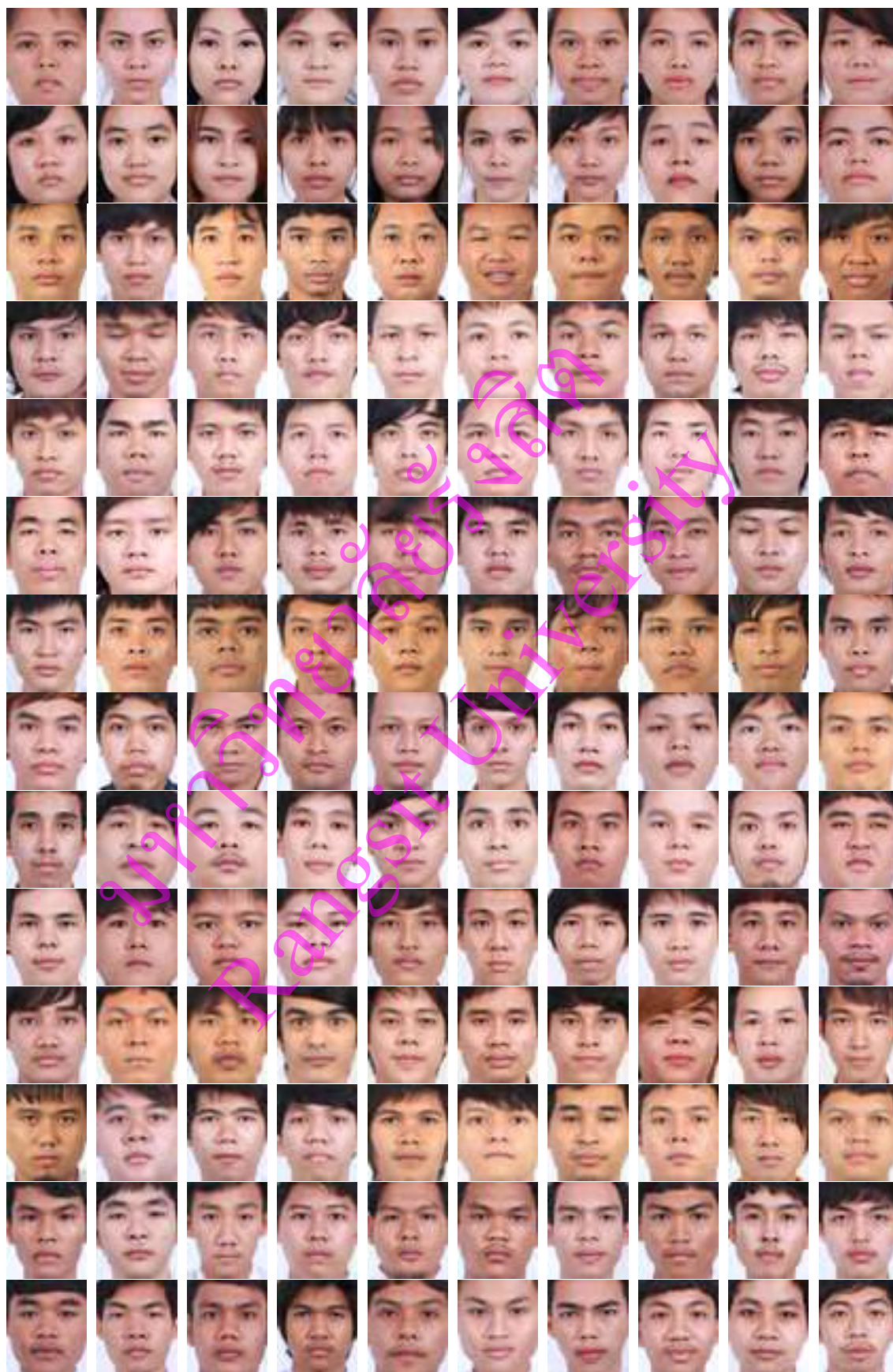


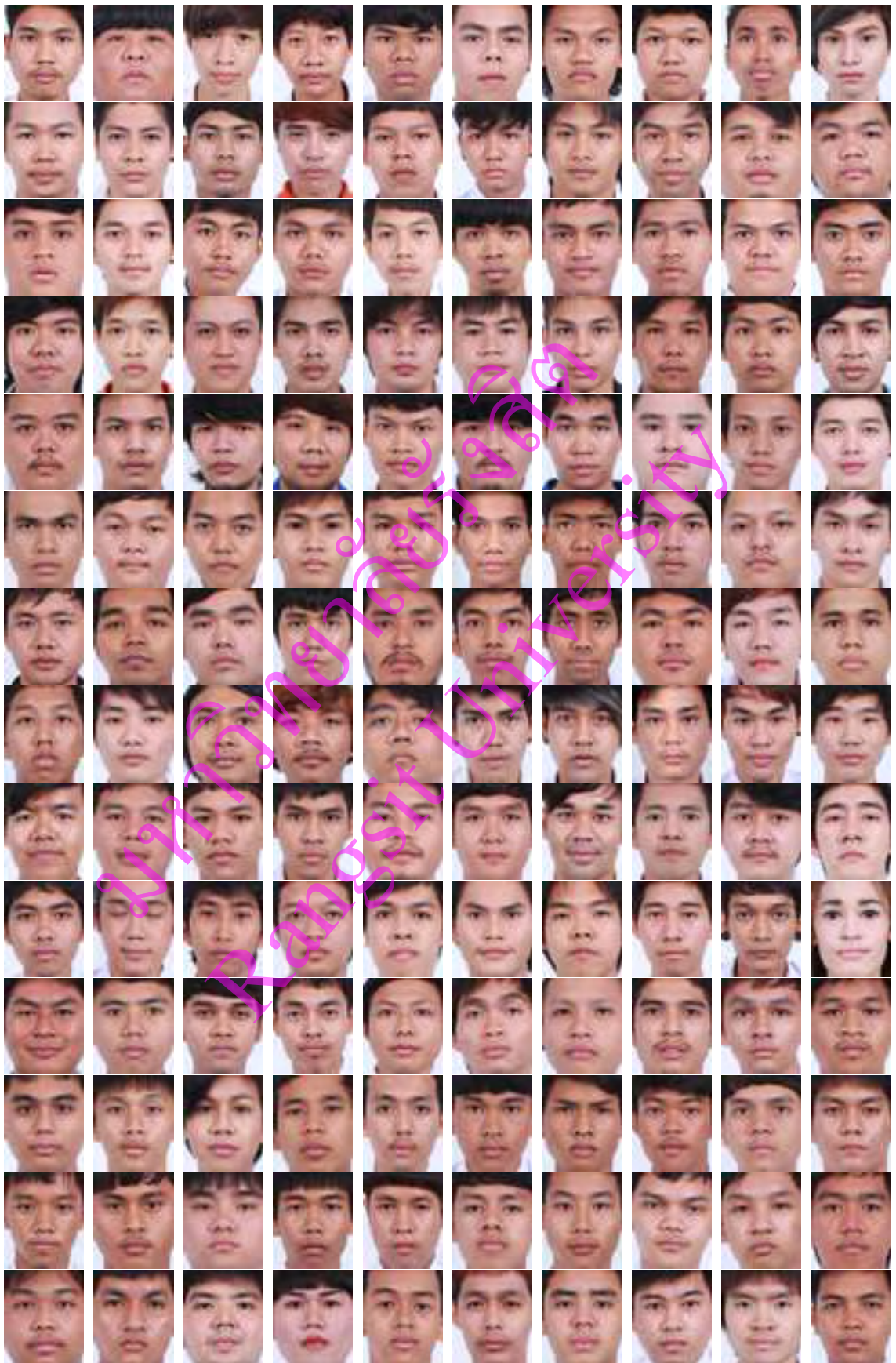














มหาวิทยาลัยรังสิต
Rangsit University

ภาคผนวก ข.

บทความการเข้าร่วมประชุมทางวิชาการระดับนานาชาติ

มหาวิทยาลัยรังสิต
Rangsit University

Adaptive Locally Linear Embedding For Face Hallucination

Metee Thongdee, Chom Kimpan (Rangsit University, Thailand)
Parinya Sanguansat (Panyapiwat Institute of Management, Thailand)

Abstract

In this paper, we proposed a novel subspace technique called Adaptive Locally Linear Embedding (ALLE). This technique can apply for face hallucination which is the reconstruction of a high-resolution face image from a single low-resolution image. Normally, the traditional Locally Linear Embedding (LLE) determines the optimal weights by the fixed number of neighbors for every point. This method is not suitable for real data because the number of neighbors in each point is not equal to others. The proposed ALLE uses a threshold of similarity for selecting the neighbors of each point. In this method, the weights are forced to zero for the neighbors that exceed the threshold. In our face hallucination framework, the learning based methods hallucinate high-resolution patches using the same position patches of each training image. The optimal weights of ALLE are estimated the relationship between these patches. The final high-resolution face image is formed by integrating the hallucinated patches. Experimental results show that the quality of reconstructed image is better than original Locally Linear Embedding (LLE) based face hallucination techniques.

Key words: Face hallucination, Super-resolution, Locally Linear Embedding, Adaptive Locally Linear Embedding

1. Introduction

Video surveillance cameras has been widely used in official places such as banks, stores and parking lots. The details of facial features obtained from surveillance video are important for identifying an individual but in many cases, the quality of images are not good because the resolution of a face image is normally low, causing some losses of facial features. Therefore, in order to obtain detailed facial features for the purpose of recognition, it is necessary to infer a high-resolution face image from the low-resolution one by the technique called face hallucination or face super-resolution⁽¹⁾. Super-resolution addresses various applications in a variety of important sectors, as diverse as medical imaging, satellite imaging, surveillance system, image enlarging in web pages, and restoration of old historic photographs. Super resolution is a technique to achieve a high-resolution (HR) image or multiple HR images from low-resolution (LR) image sequences or a single LR one. Under some circumstances, it is difficult or impossible to obtain image sequences.

Due to the limited information of image identification, reconstruction and expression analysis is a challenge to both humans and computer. For this reason, many methods have been proposed in Super-Resolution Reconstruction (SRR) and face hallucination⁽¹⁾⁻⁽⁹⁾⁽¹³⁾⁻⁽¹⁶⁾. In face hallucination, it can be performed in two different approaches: reconstruction-based and learning-based. Reconstruction-based approach uses multiple LR images of same object as the input for reconstructing a HR image, whereas learning-based approach uses a lot of training samples of same domain but different object for the reconstructing the HR image. The advantage of learning-based is it can reconstruct HR image from a single LR image.

One of the learning-based approach is the patch position which uses face image as well as image features to synthesize a high-resolution face image from a low resolu-

tion. The image position-patches are used to hallucinate the high-resolution image, and position-patches are not required to search the data from the training set. Compared with neighbor patches is widely used in face hallucination⁽⁹⁾⁻⁽¹²⁾ where neighborhood preservation for low-resolution and high-resolution patches rarely holds, position-patches hold the same position preservation between low-resolution and high-resolution patches, leading to more satisfactory results than neighbor patches. A number of related face hallucination methods have been proposed in recent years. Among them, learning based methods have received much attention because they can achieve high magnification factor and produce good super-resolved results compared with other methods.

Recently, Xiang Ma et al. proposed a face hallucination method based on position-patch⁽¹⁰⁾⁽¹⁷⁾. This method tried to improve image quality by using a one-step face hallucination based on position-patch instead of on neighbor-patch. In their work, the reconstruction of a high-resolution face image is created based on a set of high- and low-resolution training image pairs. The high resolution image obtained from their proposed method was generated by using the same position image patches of each training images. The result of this work provided the better image quality than other methods including Cubic B-spline, Wangs eigentransformation method⁽¹³⁾, Changs Neighbor Embedding⁽¹⁴⁾ and Zhuangs Locality Preserving method⁽¹⁵⁾. However, the traditional Locally Linear Embedding (LLE) which was used in their method still has problem causing from the determination of the optimal weights by using the fixed number of neighbors for every point. This method is not suitable for real data because the number of neighbors in each point is not equal to others.

In this paper, we resolved the problem of LLE by proposing a novel subspace technique, called Adaptive Locally Linear Embedding (ALLE). This technique can be applied for

face hallucination which is the reconstruction of a high-resolution face image from a single low-resolution image. In our method, we used a threshold of similarity for selecting the neighbors of each point. The optimal weights of ALLE were estimated the relationship between these patches. Then, the hallucinated patches were reconstructed based on the optimal weights to form a final high resolution face image. The result from our proposed method showed the image quality for reconstruction of the high-resolution face image better than the other methods.

Our paper is organized as follows: First, the description of locally linear embedding (LLE) concept is shown in section 2. Next, our Adaptive Locally Linear Embedding (ALLE) is proposed in section 3, the face hallucination method is presented in section 4 followed by the experimental results in section 5. Finally, conclusions are presented in section 6.

2. Locally Linear Embedding

The LLE⁽¹⁷⁾ algorithm attempts to discover nonlinear structure in the data. The algorithm is established under the assumption that any object is nearly flat on small scales. In this section, introduces the basics of LLE⁽¹⁸⁾. The LLE maps a data set $X = \{x_1, x_2, \dots, x_n\} \in R^{d \times n}$ globally to a data set $Y = \{y_1, y_2, \dots, y_n\} \in R^{m \times n}$. The data set (X, Y) is used to model the subsequently relationship. The basic principle of LLE is to minimize the reconstruction error of the set of all local neighborhoods in the data set. This takes two step 1) locally fitting hyper-planes around each samples x based on its K nearest neighbors and calculating reconstruction weights and 2) finding low-dimensional coordinates y_i for each x_i by minimizing a mapping function based on these weights.

(1) the cost function to be minimized is

$$\epsilon_w = \sum_{i=1}^n \left| x_i - \sum_{j=1}^K w_{i,j} x_{N(j)} \right| \dots \dots \dots (1)$$

This step is to make X_i be linearly reconstructed in terms of its neighbors $X_{N(1)}, X_{N(2)}, \dots, X_{N(K)}$. If X_i and X_j are not in the same neighborhood, $w_{i,j} = 0$. Weights $w_{i,j}$ are computed according to the least square principle and they sum up to 1. Weights $w_{i,j}$ are stored in an $n \times n$ sparse matrix W .

(2) the weight matrix W is fixed and new m -dimensional vectors y_i are sought by minimizing the criterion

$$\epsilon_y = \sum_{i=1}^n \left| y_i - \sum_{j=1}^K w_{i,j} y_{N(j)} \right|^2 \dots \dots \dots (2)$$

Rewriting (2) then gives

$$\epsilon_y = \sum_{i=1}^n \sum_{j=1}^n M_{i,j} y_i^T y_j = \text{tr} (YMY^T) \dots \dots \dots (3)$$

where M is an $n \times n$ matrix found as

$M = (I - W)^T (I - W)$ and Y constrains the y_i as its columns.(2) is reduced to an optimization problem as follows:

$$\min \epsilon (Y) = \text{tr} (YMY^T) \dots \dots \dots (4)$$

$$\text{s.t. } \frac{1}{n} Y Y^T = I$$

This is an eigenvalue problem. All eigenvectors of M are solutions, but the eigenvectors corresponding to the smallest eigenvalues minimize the objective function of (4).

3. Adaptive Locally Linear Embedding

In this section, the adaptive locally linear embedding (ALLE) is proposed. Normally, the weight matrix W is computed from the neighborhood in the same space of the training samples. Similar to LLE, ALLE computes the weights from the space of training samples which is adaptively built from only the neighborhood of each input, not all training samples. Since some information that contains in all training samples can make the optimizer misleading to another optimal values. The threshold of similarity θ is defined for building each subspace by

$$\frac{\sum_{i=1}^K d_i}{\sum_{i=1}^n d_i} \leq \theta \dots \dots \dots (5)$$

where d_i is the distance between input sample and the i^{th} training sample in which $d_1 \leq d_2 \leq d_3 \leq \dots \leq d_n$. That means the number of the nearest neighbors (K) is not the same value for all input samples.

4. Proposed Face Hallucination Method

The proposed face hallucination algorithm is summarized as follows:

1. Input low resolution image, low resolution training image and high resolution training image in overlapping patches.
2. For each patch,
 - (a) compute the optimal weights W of LR input from LR training set by ALLE.
 - (b) synthesize the HR patch of LR input by combining the corresponding HR patches of the LR training set with the optimal weights.
3. Concatenate and integrate the hallucinated high-resolution patches to form a facial image, which is the target high-resolution face image.

5. Experimental Results

Our face hallucination algorithm was performed on the CAS-PEAL face database⁽⁴⁾. We randomly selected 270 normal expression images of different people under the same light condition in CAS-PEAL face database. These face images were aligned manually using the locations of three points : centers of left and right eyeballs and center of the mouth⁽⁶⁾ and cropped to 128×96 pixel for high resolution face images and 32×24 pixel for low resolution face images. We compared our approach with baseline methods based on the same training set. These methods are Cubic-B-Spline and Mas position patch method⁽⁶⁾. For ALLE, we empirically found that the image quality is satisfied when threshold of similarity θ is set to 0.1 or 10 percent . The image quality is measured by the peak signal-to-noise ratio (PSNR) as

shown in Table 1. Our proposed method has the better results than Cubic-B-Spine and Ma's position patch method⁽⁶⁾. Some representative hallucinated results are shown in Fig. 1. The results of comparison of Difference due HR and image is reconstructed by our propose are shown in Fig. 2.



Fig. 1. Hallucinated results: (a) Low resolution image; (b) Bicubic; (c) Xiang Ma,et al.; (d) Proposed method; (e) Original image

6. Conclusion

In this study, we designed the adaptive locally linear embedding method for face hallucination. Experiments show that the proposed method generates results with the best image quality compared our approach with baseline methods based on the same training set. The image quality is measured by the peak signal-to-noise ratio (PSNR). Our future studies will focus on non frontal view or multi-view face hallucination.

Table 1. List of PSNR of hallucinated image

Bicubic	Xiang Ma,et al	Proposed
34.46 dB	36.24 dB	36.63 dB
33.86 dB	34.93 dB	35.31 dB
34.82 dB	35.55 dB	35.94 dB
34.17 dB	36.08 dB	36.66 dB
33.02 dB	34.27 dB	34.54 dB
34.20 dB	35.94 dB	36.38 dB



Fig. 2. Comparison of Difference due HR and image is reconstructed by our propose: (a) Difference HR image by Bicubic; (b) Difference HR image by Xiang Ma,et al.; (c) Difference HR image by Proposed method

References

- (1) S. Baker, T. Kanade: Hallucinating faces, in Proceedings of the International Conference on Automatic Face and Gesture Recognition, Grenoble, France, pp.83-88(2000).
- (2) S. Baker, T. Kanade: Limits on super-resolution and how to break them, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (9), pp.1167-1183(2002).
- (3) M. Elad, A. Feuer: Super-resolution reconstruction of image sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (9), pp.817-834(1999).
- (4) W. Gao, B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H.Zhang, D.B. Zhao: The CAS-PEAL large-scale Chinese face database and baseline evaluations, IEEE Transactions on System Man, and Cybernetics(Part A) (38), pp.149-161(2008).
- (5) X.G. Wang, X.O. Tang: Hallucinating face by eigen-transformation, IEEE Transactions on Systems Man and Cybernetics (Part C) 35 (3), pp.425-434(2005).
- (6) X. Ma, J. Zhang, and C. Qi: Hallucinating face by position-patch, Pattern Recognit., vol.43, pp.2224-2236(2010).
- (7) X. Ma, J. Zhang, and C. Qi: Position-based face hallucination method, in Proc. IEEE Conf. Multimedia and Expo, pp.290-293(2009).
- (8) Y.T. Zhuang, J. Zhang, F. Wu: Hallucinating faces:LPH super-resolution and neighbor reconstruction for residue compensation, Pattern Recognition, pp.3178-3194(2007).
- (9) W. Liu, D. Lin, X.O. Tang: Neighbor combination and

- transformation for hallucinating faces, in Proceedings of the IEEE International Conference on Multimedia and Expo , Netherlands , pp.478-484(2005).
- (10) S.W. Park, M. Savvides: Robust super-resolution of face images by iterative compensating neighborhood relationships, in Proceedings of the Biometrics Symposium, United States , (2008).
 - (11) W. Fan, D.Y. Yeung: Image hallucination using Neighbor Embedding over visual primitive manifolds, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition , United States , (2007).
 - (12) W. Fan, D.Y. Yeung: Image Hallucination using Neighbor Embedding over visual primitive manifolds, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition , pp.17(2007).
 - (13) W. T. Freeman and E. C. Pasztor: Learning Low Level Vision, in 7th IEEE International Conference on Computer Vision , pp.1182-1189(1999).
 - (14) R. C. Hardie , et al.: Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images, IEEE Transactions on Image Processing , vol.6 , pp.1621-1633(1997).
 - (15) Z. Lin and H-Y. Shum: Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation, IEEE Transactions on Pattern Analysis and Machine Intelligence , pp.83-97(2004).
 - (16) S. C. Park, et al.: Super-Resolution Image Reconstruction, IEEE Signal Processing Magazine , pp.21-36(2003).
 - (17) L. K. Saul and S. T. Roweis: Nonlinear dimensionality reduction by locally linear embedding, Science , pp.2323-2326(2000).
 - (18) Bing-Yu, S., Z. Xiao-Ming, et al.: Feature Fusion Using Locally Linear Embedding for Classification, Neural Networks, IEEE Transactions on , 21(1) , pp.163-168.
 - (19) H. Chang, D.Y. Yeung, Y.M. Xiong: Super-resolution through Neighbor Embedding, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition , Washington , pp.1275-1282(2004).

ICEAS 1403
Multiview Face Hallucination by
Adaptive Locally Linear Embedding

Metee Thongdee

Rangsit University, Thailand

metee.td@gmail.com

Chom Kimpan

Rangsit University, Thailand

chom@rsu.ac.th

Parinya Sanguansat

Panyapiwat Institute of Management, Thailand

sanguansat@yahoo.com

Abstract

In this paper, we propose Multiview Face Hallucination method by Adaptive Locally Linear Embedding (ALLE). This method can enhance a single view low resolution (LR) image up to multiview high resolution (HR) images. The input face image can be either frontal or nonfrontal face image. In our previous work, ALLE-based face hallucination was successfully applied to improve facial details from a single frontal view. Subsequently in this paper, the ALLE is applied to manipulate the non-frontal face details. By feeding a LR face in any view (Up, Down, Left, or Right) to the proposed Multiview Face Hallucination, the HR of the input will be generated in all views. Experimental results show that the quality of reconstructed image of our proposed framework is better than enhancing with the baseline in both interpolation and learning methods.

Keywords: Face hallucination, Super-resolution, Multiview, Locally Linear Embedding, Adaptive Locally Linear Embedding

1. Introduction

Generally, face Hallucination researches focus on only frontal face image because of the limitation of training samples. However, only frontal face cannot satisfy the real applications because the real face images, which are captured by video surveillance camera, are in many view angles. Video surveillance cameras have been widely used in official places such as banks, stores and parking lots. The details of facial features obtained from surveillance video are important for identifying an individual but in many cases, the quality of images are not good because the resolution of a face image is normally low, causing some losses of facial features. Therefore, in order to obtain detailed facial features for the purpose of recognition, it is necessary to infer a high-resolution face image from the low-resolution one by the technique called face

hallucination or face super-resolution [5]. Super-resolution addresses various applications in a variety of important sectors, as diverse as medical imaging, satellite imaging, surveillance system, image enlarging in web pages, and restoration of old historic photographs. Super resolution is a technique to achieve a high-resolution (HR) image or multiple HR images from low-resolution (LR) image sequences or a single LR one. Under some circumstances, it is difficult or impossible to obtain image sequences

Due to the limited information of image, identification, reconstruction and expression analysis is a challenge to both humans and computer. For this reason, many methods have been proposed in Super-Resolution Reconstruction (SRR) and face hallucination [1-3,13-16]. In face hallucination, it can be performed in two different approaches: reconstruction-based and learning-based. Reconstruction-based approach uses multiple LR images of same object as the input for reconstructing a HR image, whereas learning-based approach uses a lot of training samples of same domain but different object for the reconstructing the HR image. The advantage of learning-based is it can reconstruct HR image from a single LR image.

One of the learning-based approach is the patch position which uses face image as well as image features to synthesize a high-resolution face image from a low resolution. The image position-patches are used to hallucinate the high-resolution image, and position-patches are not required to search the data from the training set. Compared with neighbor patches is widely used in face hallucination [8-12] where neighborhood preservation for low-resolution and high-resolution patches rarely holds, position-patches hold the same position preservation between low-resolution and high-resolution patches, leading to more satisfactory results than neighbor patches. A number of related face hallucination methods have been proposed in recent years. Among them, learning based methods have received much attention because they can achieve high magnification factor and produce good super-resolved results compared with other methods.

Recently, Xiang Ma et al. proposed a face hallucination method based on position-patch [6-7]. This method tried to improve image quality by using a one-step face hallucination based on position-patch instead of on neighbor-patch. In their work, the reconstruction of a high-resolution face image could be created based on a set of high- and low- resolution training image pairs. The high resolution image obtained from their proposed method was generated by using the same position image patches of each training images. The result of this work provided the better image quality than other methods including Cubic B-spline, Wangs eigen-transformation method [5], Changs Neighbor Embedding [19] and Zhuangs Locality Preserving method [8].

The problem on uncontrollable of face images which is captured by camera are usually LR and nonfrontal face images. However, face hallucination method limited to

frontal face. Therefore, it is of practical significance to study how to resolve the HR multiview faces from a LR nonfrontal one.

Therefore, Xiang Ma et al. proposed a multiview face hallucination method[20] based on position-patch[6-7]. In their work, this method tried to resolve problem multiview face hallucination. A simple face transformation method that converts face transformation to a globally, predicting LR multiple views of a given LR one. Based on the LR faces synthesized, the local position-patch based face hallucination method [6] is used to incorporate face details. The method for mainly incorporating frontal face details was reported in previous work [6]. But, the traditional Locally Linear Embedding (LLE) which was used in their method still has problem causing from the determination of the optimal weights by using the fixed number of neighbors for every point. This method is not suitable for real data because the number of neighbors in each point is not equal to others.

In this paper, we resolved the problem of LLE by proposing a novel subspace technique, called Adaptive Locally Linear Embedding (ALLE). This technique can be applied for face hallucination which is the reconstruction of a high-resolution face image from a single low-resolution image. In our method, we used a threshold of similarity for selecting the neighbors of each point. The optimal weights of ALLE were estimated the relationship between these patches. Then, the hallucinated patches were reconstructed based on the optimal weights to form a final high resolution face image. The result from our proposed method showed the image quality for reconstruction of the high-resolution face image better than the other methods.

Our paper is organized as follows: First, the description of locally linear embedding (LLE) concept is shown in section 2. Next, our Adaptive Locally Linear Embedding (ALLE) is proposed in section 3, the multiview face hallucination method is presented in section 4 followed by the experimental results in section 5. Finally, conclusions are presented in section 6.

2. Locally Linear Embedding

The LLE [17] algorithm try to discover nonlinear structure in the data. The algorithm is existed under the assumption that any object is “nearly” flat on small scales. In this part, introduces the basics of LLE [18]. The LLE maps a data set $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in R^{d \times n}$ globally to a data set $Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\} \in R^{m \times n}$. The data set (X, Y) is used to model the subsequently relationship. The basic concept of LLE is to minimize the reconstruction error of the set of all local neighborhoods in the data set. This can categorized into two step 1) locally fitting hyper-planes around each samples \mathbf{x} based on its K nearest neighbors and calculating reconstruction weights and 2) finding low-dimensional coordinates \mathbf{y}_i for each \mathbf{x}_i by minimizing a mapping function based on these weights. (1) the cost function to be minimized is

$$\varepsilon_W = \sum_{i=1}^n \left| \mathbf{x}_i - \sum_{j=1}^K w_{i,j} \mathbf{x}_{N(j)} \right|^2 \quad (1)$$

This step is to make \mathbf{x}_i linearly reconstructed in terms of its neighbors

$\mathbf{x}_{N(1)}, \mathbf{x}_{N(2)}, \dots, \mathbf{x}_{N(K)}$. If \mathbf{x}_i and \mathbf{x}_j are not in the same neighborhood, $w_{i,j} = 0$

Weights $w_{i,j}$ are computed according to the least square principle and they sum up to 1. Weights $w_{i,j}$ are stored in an $n \times n$ sparse matrix W

(2) the weight matrix W is fixed and new m -dimensional vectors \mathbf{y}_i are sought by minimizing the criterion

$$\varepsilon_Y = \sum_{i=1}^n \left| \mathbf{y}_i - \sum_{j=1}^K w_{i,j} \mathbf{y}_{N(j)} \right|^2 \quad (2)$$

Rewriting (2) then gives

$$\varepsilon_Y = \sum_{i=1}^n \sum_{j=1}^n M_{i,j} \mathbf{y}_i^T \mathbf{y}_j = \text{tr}(YMY^T) \quad (3)$$

where M is an $n \times n$ matrix found as $M = (I - W)^T (I - W)$ and Y constrains the \mathbf{y}_i as its columns. (2) is reduced to an optimization problem as follows:

$$\begin{aligned} \min \varepsilon(Y) &= \text{tr}(YMY^T) \\ \text{s.t. } \frac{1}{n} YY^T &= I \end{aligned} \quad (4)$$

This is an eigenvalue problem. All eigenvectors of M are solutions, but the eigenvectors corresponding to the smallest eigenvalues minimize the objective function of (4).

3. Adaptive Locally Linear Embedding

In this section, the adaptive locally linear embedding (ALLE) is proposed. Normally, the weight matrix W is computed from the neighborhood in the same space of the training samples. Similar to LLE, ALLE computes the weights from the space of training samples which is adaptively built from only the neighborhood of each input, not all training samples. Since some information that contains in all training samples can make the optimizer misleading to another optimal values. The threshold of similarity θ is defined for building each subspace by

$$\frac{\sum_{i=1}^K d_i}{\sum_{i=1}^n d_i} \leq \theta \quad (5)$$

where d_i is the distance between input sample and the i^{th} training sample in which $d_1 \leq d_2 \leq d_3 \leq \dots \leq d_n$. That means the number of the nearest neighbors (K) is not the same value for all input samples.

4. Multiview Face Hallucination

This framework called multiview face hallucination or MFH algorithm[20] is based line method. We used ALLE method computes the optimal weights (W). The algorithm is summarized as follows:

1. Input LR face image at view p , LR training set at view p , multiview LR training set face images and multiview high resolution training set face images.
2. Compute the optimal weight W_p of LR input from LR training set at view p by ALLE.
3. For each view
 - (a) Synthesize multiview LR face images of the LR training set with the optimal weights W_p .
 - (b) Divide multiview LR face images, multiview LR training face images and multiview HR training face images respectively into position patches.
 - (c) Concatenate and integrate the hallucinate high resolution patches to form a facial image, which is the target high resolution face image.

5. Experimental Results

Our face hallucination algorithm was performed on the CAS-PEAL-R1 face database [4]. We randomly selected 2000 face images of 400 different individuals were randomly selected. Each individual has five different views left, right, up, down and frontal under the same light condition in CAS-PEAL face database. These face images were aligned manually using the locations of three points : centers of left and right eyeballs and center of the mouth [6]. Some aligned face images are shown in Fig. 1 and cropped to 128×96 pixel for high resolution face images and 32×24 pixel for low resolution face images. Some cropped face images are shown in Fig. 2. The optimal weights (K) is set to 150. We compared our approach with baseline methods based on the same training set. These methods are Cubic-B-Spline and MFH method [6]. For ALLE, we empirically found that the image quality is satisfied when threshold of similarity θ is set to 0.3. The image quality is measured by PSNR are shown in Table1. Our proposed method has the better results than Cubic-B-Spline and MFH method [20]. Some representative hallucinated MFH method results are shown in Fig. 3 and some representative our proposed method results are shown in Fig. 4.



Fig. 1 Aligned face images



Fig. 2 Cropped face images

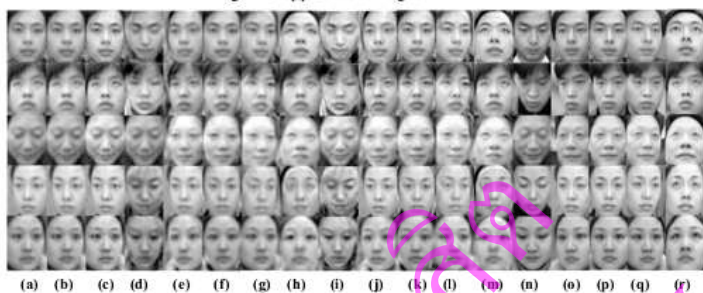


Fig. 3 Multiview face hallucination[20]. (a) LR(32×24) input faces at single view. (b) Bicubic. (c) Ma's position patch method[6]. (d)-(h) Face transformation[20]. (i)-(m) HR(128×96) MFH[20]. (n)-(r) Original images.

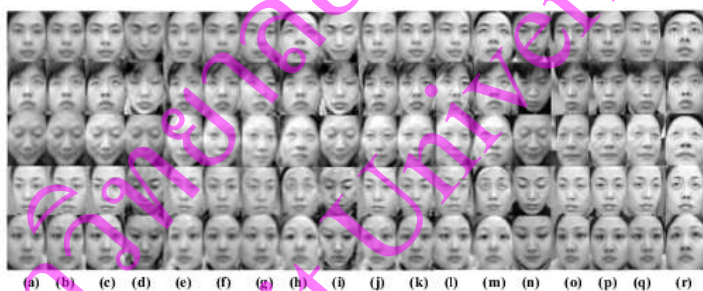


Fig. 4 Multiview face hallucination[20]. (a) LR(32×24) input faces at single view. (b) Bicubic. (c) Proposed method. (d)-(h) Face transformation[20]. (i)-(m) HR(128×96) Proposed method. (n)-(r) Original images.

Table 1. List of PSNR of Hallucinated image

Input	Down view (i)		Left view (j)		Frontal view (k)		Right view (l)		Up view (m)	
	MFH	ALLE	MFH	ALLE	MFH	ALLE	MFH	ALLE	MFH	ALLE
Frontal	27.24	27.73	26.54	26.65	31.40	31.47	31.73	32.76	28.80	29.29
Up view	24.98	25.07	26.87	26.96	26.45	26.62	26.11	26.18	31.36	31.40

Down view	31.34	31.38	26.23	26.89	28.51	28.65	27.43	28.62	29.97	30.49
Left view	29.19	30.39	32.71	32.88	32.62	32.82	31.92	32.58	35.34	35.66
Right view	32.60	32.61	27.71	27.72	32.29	32.29	31.60	31.60	29.72	29.75

6. Conclusion

In this paper, we propose Multiview Face Hallucination method by Adaptive Locally Linear Embedding (ALLE). Experimental results show that the quality of reconstructed image of our proposed framework is better than enhancing with the baseline in both interpolation and learning methods on the same training set. The image quality is measured by the peak signal-to-noise ratios (PSNR).

7. Acknowledgment and Legal Responsibility

This work has been supported by Research Grant for New Scholar(MRG5480162) from TRF (Thai Research Fund), CHE (Commission on Higher Education), and Panyapiwat Institute of Management, Thailand.

8. References

- [1] S. Baker, T. Kanade, "Hallucinating faces," in: *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, 2000, pp. 83–88.
- [2] S. Baker, T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (9) (2002)
- [3] M. Elad, A. Feuer, "Super-resolution reconstruction of image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (9) (1999)
- [4] W. Gao, B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H.Zhang, D.B. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations," *IEEE Transactions on System Man, and Cybernetics (Part A)* (38) (2008) 149–161.
- [5] X.G. Wang, X.O. Tang, "Hallucinating face by eigentransformation," *IEEE Transactions on Systems Man and Cybernetics (Part C)* 35 (3) (2005) 425–434.
- [6] X. Ma, J. Zhang, and C. Qi, "Hallucinating face by position-patch," *Pattern Recognit.*, vol. 43, pp.2224–2236, 2010.
- [7] X. Ma, J. Zhang, and C. Qi, "Position-based face hallucination method," in *Proc. IEEE Conf.Multimedia and Expo*, 2009, pp. 290–293.
- [8] Y.T. Zhuang, J. Zhang, F. Wu, "Hallucinating faces:LPH super-resolution and neighbor reconstruction for residue compensation," *Pattern Recognition* 40 (2007)
- [9] W. Liu, D. Lin, X.O. Tang, "Neighbor combination and transformation for hallucinating faces," in:*Proceedings of the IEEE International Conference on*

- Multimedia and Expo*, Netherlands, 2005, pp. 478–484.
- [10] S.W. Park, M. Savvides, “Robust super-resolution of face images by iterative compensating neighborhood relationships,” in: *Proceedings of the Biometrics Symposium*, United States, 2008, pp. 4430531.
- [11] W. Fan, D.Y. Yeung, “Image hallucination using Neighbor Embedding over visual primitive manifolds,” in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, United States, 2007.
- [12] W. Fan, D.Y. Yeung, “Image Hallucination using Neighbor Embedding over visual primitive manifolds,” in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2007, 17–22 June 2007, pp. 1–7.
- [13] W. T. Freeman and E. C. Pasztor, “Learning Low Level Vision,” in *7th IEEE International Conference on Computer Vision*, 1999, pp. 1182 - 1189.
- [14] R. C. Hardie, *et al.*, “Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images,” *IEEE Transactions on Image Processing*, vol. 6, pp. 1621-1633, 12 December 1997.
- [15] Z. Lin and H.-Y. Shum, “Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 83-97, 2004.
- [16] S. C. Park, *et al.* (2003) Super-Resolution Image Reconstruction, *IEEE Signal Processing Magazine*. 21-36.
- [17] L. K. Saul and S. P. Roweis, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, pp. 2323–2326, 2000.
- [18] Bing-Yu, S., Z. Xiao-Ming, *et al.* “Feature Fusion Using Locally Linear Embedding for Classification.” *Neural Networks*, *IEEE Transactions on* 21(1):
- [19] H. Chang, D.Y. Yeung, Y.M. Xiong. “Super-resolution through Neighbor Embedding,” in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, 2004, pp.1275–1282.
- [20] Xiang Ma, H. H., Shaopeng Wang, and Chun Qi (2010). “A Simple Approach to Multiview Face Hallucination.” *IEEE SIGNAL PROCESSING* 17: 579-582.

ภาคผนวก ค.

บทความทางวิชาการในวารสารระดับนานาชาติ

มหาวิทยาลัยรังสิต
Rangsit University

ACCEPTANCE LETTER

Dear Parinya Sanguansat , Metee Thongdee , Siriporn Supratid

Congratulations! I am pleased to announce that your paper, Adaptive Locally Linear Embedding For Multiview Face Hallucination, has been accepted by the Journal of Convergence Information Technology (ISSN: 1975-9320). You may check your paper's status by clicking "My Page" button in JCIT. <http://www.ajcit.org/jcit>

Aims and Focus of JCIT:

Journal of Convergence Information Technology is a peer reviewed international journal published by **AJCIT** (Advanced Institute of Convergence Information Technology). The aim of the journal is to provide fast publication of refereed, high quality original research papers in all branches of the convergence technologies and its applications.

JCIT focuses on the theories/technologies/architecture and its applications on the various aspects of advances in convergence and hybrid Information Technology. All research results on the traditional and convergence areas of information technologies and next generation information technologies shall be always welcome by JCIT.

JCIT is a refereed, multidisciplinary journal for bridging the latest advances in the convergence and hybrid information technology research. The journal only publishes well-written and academically validated manuscripts in both theoretical development and application research.

Date : 2014-03-27

Dr. Franz I.S. Ko (Advanced Institute of Convergence Information Technology, Korea / IBC, UK)

Franzko



Address: AJCIT, Myoungbo Bldg, 3F, Bunkin-dong 1-ga, Seo-gu, Busan, 602-816, South Korea
 Registration Number: 314-04-02154, TEL: International: +82-70-7713-9173, Domestic: 070-7713-9173

มหาวิทยาลัยรังสิต
 Rangsit University

Adaptive Locally Linear Embedding For Multiview Face Hallucination

¹Metee Thongdee, ¹Siriporn Supratid, ²Parinya Sanguansat

¹Faculty of Information Technology, Rangsit University, Pathum-Thani 12000, Thailand,
metee.td@gmail.com, siri_sup1@hotmail.com

²Faculty of Engineering and Technology, Panyapiwat Institute of Management, Nonthaburi,
Thailand, parinyasan@pim.ac.th

Abstract

The traditional Locally Linear Embedding (LLE) technique was applied for face hallucination. This technique determines the optimal weights by the fixed number of neighbors for every point. Our previous work, named an adaptive locally linear embedding (ALLE), referred to a modified version of LLE was proposed to apply with frontal view face hallucination; it uses a threshold of similarity for selecting the neighbors of each point. However, frontal face is barely captured in the real world. Therefore, this paper proposes a novel ALLE for multiview face hallucination. The main objective is to generate high quality of frontal and non-frontal face images. The processing steps, according to the proposed method are operated as follows: first, a low resolution (LR) face in one of front, up, down, left or right views is fed as an input; then, the other views of such an LR image are generated by ALLE, which applies a threshold of similarity for selecting the neighbors of each point; and high resolution (HR) face images in all views of the same input object are achieved afterwards. The experimental results show that the proposed method yields the better image quality of the reconstructed frontal and non-frontal face images over the baseline methods.

Keywords: Face hallucination, Super-resolution, Multiview, Locally Linear Embedding, Adaptive Locally Linear Embedding

1. Introduction

In these days, video surveillance cameras have been widely used in many places such as banks, stores and parking lots, where intensive security is critical. Details of facial features obtained from the surveillance video are important for identifying personal identity. However, in many cases, the images obtained from the surveillance cameras cannot be well identified due to low resolution of facial images that cause some losses of the facial features. Thus, in order to obtain detailed facial features for a purpose of personal recognition, it is necessary to infer a Low-Resolution (LR) facial image to a High-Resolution (HR) one by a technique called face hallucination or face super-resolution [4]. Such techniques are applied in a variety of important sectors, e.g., medical imaging, satellite imaging, surveillance system, image enlarging in web pages, and restoration of old historic photographs. A super resolution is a technique to produce a single or multiple high-resolution (HR) images from the low-resolution (LR) image sequences or a single LR image. Under some circumstances, it is impossible to obtain image sequences. Due to limited information of image identification, reconstruction and expression analysis is a challenge to both human and computer. Several super-resolution reconstruction (SRR) researches have been proposed, relying on two approaches: (1) reconstruction-based and (2) learning-based approach. Reconstruction-based approach employs multiple LR images of the same object as an input for reconstructing an HR image, whereas learning-based approach uses numbers of training samples from same domain with different objects to reconstruct the HR image. An advantage of the learning-based approach is its ability to reconstruct the HR image from the single LR image. In this work, learning-based super-resolution, which is also known as "face hallucination" [1-3, 13-16, 27-28] is focused for applying to human facial images. A number of related facial hallucination methods have been proposed in recent years. Indeed, learning based methods have acquired greater attention as they can achieve high magnification factor and produce positive super-resolved results compared to other methods.

Xiang Ma et al. proposed a facial hallucination method based on position-patch [6-7] to improve the image quality. Such Ma's method performs one-step facial hallucination based on position-patch instead of neighbor-patch. A patch position is one of the learning-based approaches that utilizes facial

image as well as image features to synthesize high-resolution facial images from the low-resolution images. In comparison, neighbor patches is widely used in face hallucination [5, 8-12, 23-25]. In their work, the reconstruction of a high-resolution facial image could be created based on a set of high and low resolution training image pairs. The high resolution image obtained from their proposed method was generated based on the same position image patches of each training images. The result of this work shows better image quality than many other methods including Cubic B-spline, Wangs Eigen-transformation method [4], Changs Neighbor Embedding [19] and Zhuangs Locality Preserving method [8]. In [23], F. Zhou, B. Wang and Q. Liao have extended the work in [7] by using bilateral patches. In [26], local pixel structure is learnt from k nearest neighbors (KNN) faces. However, there are some uncontrollable problems regarding this method, i.e., the facial images captured by the camera are LR and non-frontal facial images where such a facial hallucination method is limited to the frontal faces. Therefore, it is practically significant to study how to create the HR multi-viewed faces from the LR non-frontal images.

As a sequence, Xiang Ma et al. proposed a multi-viewed facial hallucination method [20] based on the position-patch [6-7]. It is a simple face transformation method that converts an LR face image to a globally, predicting LR multiple views of that given LR one. Based on the synthesized LR faces, facial details are incorporated by using the local position-patch [6]. Nevertheless, the traditional Locally Linear Embedding (LLE) technique, applied in such a hallucination method still confronts a problem, relevant to the determination of the optimal weights. Such weights are defined using fixed number of neighbors for every points. This is not practical for real world data because the number of neighbors in each point is not equal to others.

In this paper, a multiview face hallucination (MFH) using adaptive locally linear embedding (ALLE) technique is proposed for efficiently reconstructing high-resolution face images from a single low-resolution one. The optimal weights determination problem, addressed in LLE is resolved by using the previously proposed technique, ALLE [21]. The ALLE is applied to manipulate the non-frontal facial details. By feeding a single LR face in one of up, down, left, or right views to the proposed method, the HR images will then be generated in all views. The paper is organized as follows. In section 2, the concept of LLE for Face Hallucination is recalled. ALLE method, our previous work is presented in section 3. The novel proposed multiview face hallucination is presented in section 4. The experimental results are discussed and conclusion is made in section 5 and 6 respectively.

2. Locally Linear Embedding

The LLE [17] algorithm is employed to discover nonlinear structure in the data. The algorithm is developed under the assumption that any object is "nearly" flat on small scales. The basic of LLE is described in [18]. The original LLE try to map a data from one space globally to another space by embedding the locations of the neighbors for each point. For face hallucination, it is about LR and HR face image spaces. The main idea of LLE is to minimize the reconstruction error of the set of all local neighborhoods in the data set. The LLE process can be categorized into two steps as follows:

Step 1: locally fitting hyper-planes around each object \mathbf{x} based on its K -nearest neighbors

Step 2: Calculating the reconstruction weights by minimizing reconstruction error by the cost function below

$$E_{\mathbf{x}_i} = \sum_{j=1}^n \left| \mathbf{x}_i - \sum_{j=1}^K w_{i,j} \mathbf{x}_{N(j)} \right| \quad (1)$$

In this step, \mathbf{x}_i is linearly reconstructed in terms of its neighbors $\mathbf{x}_{N(1)}, \mathbf{x}_{N(2)}, \dots, \mathbf{x}_{N(K)}$. If \mathbf{x}_i and \mathbf{x}_j are not in the same neighborhood, then $w_{i,j} = 0$. Weights $w_{i,j}$ are computed according to the least square principle and they sum up to 1. Weights $w_{i,j}$ are stored in an $n \times n$ sparse matrix \mathbf{W}

For face hallucination, such reconstruction weights are computed in LR face image space and then applied to HR space to reconstruct the HR face image that corresponding to the LR input.

3. Adaptive Locally Linear Embedding

In this section, the adaptive locally linear embedding (ALLE) is proposed. Normally, the weight matrix \mathbf{W} is computed from the neighborhood in the same space of the training samples. Similar to LLE, ALLE computes the weights from the space of training samples which is adaptively built from only the neighborhood of each input, not all training samples. Since some information that contains in all training samples can make the optimizer misleading to another optimal values. The threshold of similarity θ is defined for building each subspace by

$$\frac{\sum_{j=1}^K d_j}{\sum_{i=1}^n d_i} \leq \theta \quad (2)$$

Where d_i is the distance between input sample and the i^{th} training sample in which $d_1 \leq d_2 \leq d_3 \leq \dots \leq d_n$. That means the number of the nearest neighbors (K) is not the same value for all input samples.

4. The Proposed Novel Multiview Face Hallucination

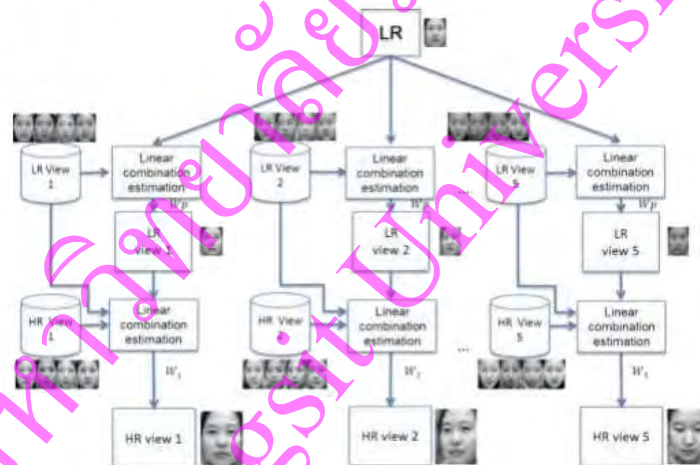


Figure 1. Multiview face hallucination Framework

The position-patch based multiview face hallucination has been proposed in MFH [20]. The framework is summarized in Figure 1; and the detail of processing steps is described as follows:

Step 1: Input LR face image at view p , LR training set and HR training set in all views.

Step 2: Estimate the linear combination to obtain the optimal weights W_p of LR input from LR training set at view p by LLE.

Step 3: for each view

- Synthesize other views of LR face images by the LR training set with the optimal weights W_p by LLE.
- Divide multiview LR face images, multiview LR training face images and multiview HR training face images respectively into position patches.
- Concatenate and integrate the hallucinated high resolution patches to form a facial image, which is the target high resolution face image.

Linear Combination for Multiview Face



Figure 2. Linear combination for multiview face

Face image is represented as a column vector of all pixel values. Based on structural similarity, face image can be synthesized using the linear combination of training objects [4]. In other words, a face image at unfixed view can also be reconstructed using linear combination of other objects in the same view aligned, as shown in Figure 2.

$$\mathbf{I}_p = \mathbf{W}_p \mathbf{L}_p \quad (3)$$

Where \mathbf{I}_p is the given LR face at view p , \mathbf{W}_p are the construction coefficients at view p that can be determined by (1), \mathbf{L}_p are the training faces at view p . Since the LR face images for other views \mathbf{L}_o are required. From the assumption in [20], the LR face images for other views \mathbf{I}_o can be approximated by

$$\mathbf{I}_o \approx \mathbf{W}_p \mathbf{L}_o \quad (4)$$

Since there is no information for determining the construction coefficients at other views. After obtaining LR in all views, these LR images are hallucinated to HR face images of all views by position-patch approach.

The proposed framework for multiview face hallucination is to replace the linear combinations in all processes by using ALLE instead of LLE in order to improve the performance of hallucinating face image in multiple views. That means the number of the neighbors can be adapted for each input patches. This would relieve the deviation from optimal values.

5. Experimental Results

Our face hallucination algorithm was performed on the CAS-PEAL-R1 face database [22]. 5,055 face images of 1011 different individuals are randomly selected. Each individual has five different views (left, right, up, down and frontal views) under the same light condition. These face images were aligned manually using the locations of three points: centers of left and right eyeballs and center of the mouth [6]. Figure. 3 shows some aligned face images, which are cropped to 32×24 pixel for low resolution face images and to 128×96 pixel for high resolution face images. Some cropped face images are shown in Figure. 4 and Figure. 5. Based on the same training sets, the proposed method is compared to baseline method, Bicubic and Ma's method [20]. Input face images are shown in Figure. 6 including (a) frontal view, (b) up view, (c) down view, (d) left view and (e) right view. Each individual input face image was generated into five different outputs of LR, synthesized LR face images and HR face images according to our framework. Face image outputs, generated by the proposed method were compared with Ma's method [20] as shown in Figure. 7 to 11. The experimental results show the difference between the results of high-resolution facial image by varying the number of neighbors, K as 100, 200 and 300 regarding LLE along with 0.1, 0.2 and 0.3 similarity threshold (θ) according to ALLE, respectively. With regard to ALLE, Table 1 indicates that the best quality of frontal view image can be found at 0.1 value of similarity threshold θ ; while Table 2 to 5 points that the best quality of other views is found at 0.2. In all cases of inputs, final HR images generated by the proposed method show superior image quality over Ma's method in every view.



Figure 3. Aligned face images



Figure 4. Low resolution cropped face images



Figure 5. High resolution cropped face images



Figure 6. Example input a face image on five different single views



Figure 7. Input a face image of frontal view Figure. 6 (a) Synthesized low-resolution images by Ma's method (b) Synthesized low-resolution images by the proposed method (c) Bicubic interpolation of (a) (d) Bicubic interpolation of (b) (e) Final HR results by Ma's method (f) Final HR results by the proposed method (g) Original images

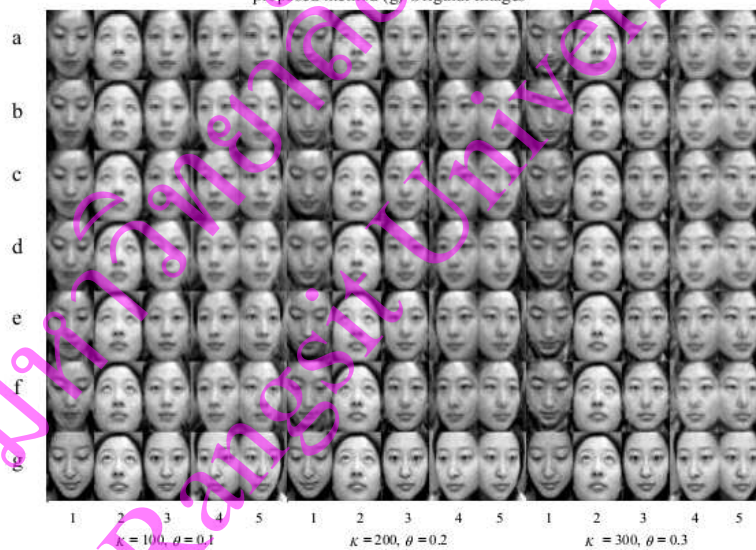


Figure 8. Input a face image of up view Figure.6(b) (a) Synthesized low-resolution images by Ma's method (b) Synthesized low-resolution images by the proposed method (c) Bicubic interpolation of (a) (d) Bicubic interpolation of (b) (e) Final HR results by Ma's method (f) Final HR results by the proposed method (g) Original images

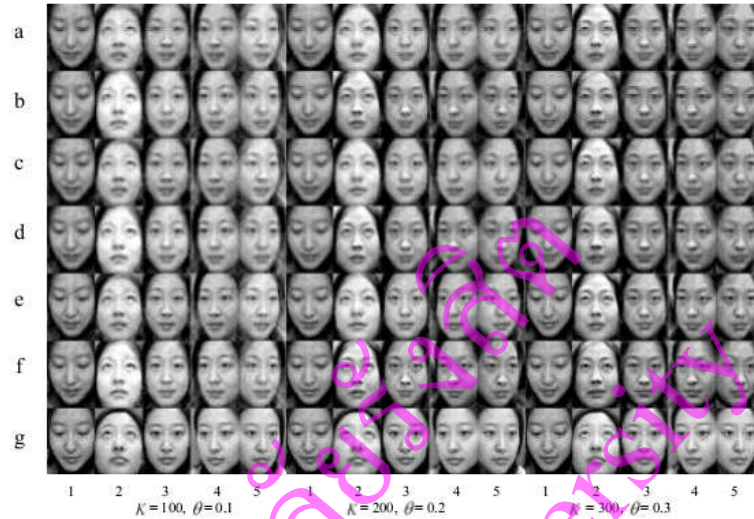


Figure 9. Input a face image of down view Figure.6(c) (a) Synthesized low-resolution images by Ma's method (b) Synthesized low-resolution images by the proposed method (c) Bicubic interpolation of (a) (d) Bicubic interpolation of (b) (e) Final HR results by Ma's method (f) Final HR results by the proposed method (g) Original images

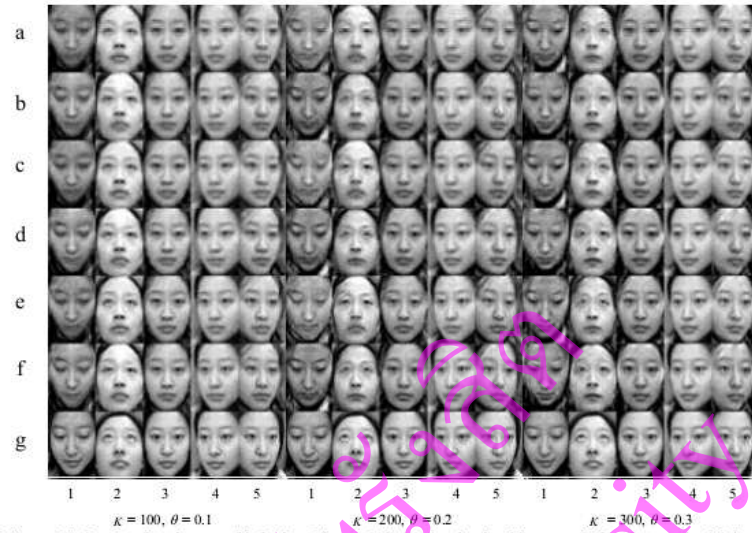


Figure 10. Input a face image of left view Figure.6(d) (a) Synthesized low-resolution images by Ma's method (b) Synthesized low-resolution images by the proposed method (c) Bicubic interpolation of (a) (d) Final HR results by Ma's method (e) Final HR results by the proposed method (f) Final HR results by the proposed method (g) Original images

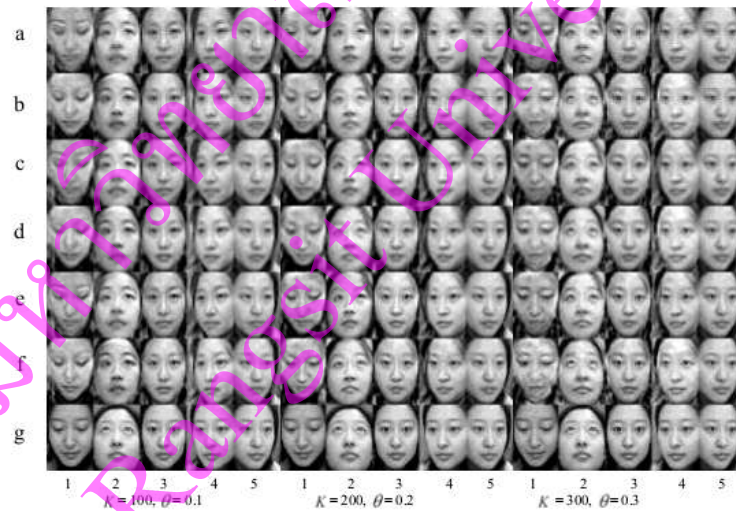


Figure 11. Input a face image of right view Figure.6(e) (a) Synthesized low-resolution images by Ma's method (b) Synthesized low-resolution images by the proposed method (c) Bicubic interpolation of (a) (d) Bicubic interpolation of (b) (e) Final HR results by Ma's method (f) Final HR results by the proposed method (g) Original images

Table 1. List of PSNR of Hallucinated image (input a face image on frontal view) $K=100$, $\theta = 0.1$

View	Ma's method	Our method
frontal	32.49	32.89
Up	26.98	29.94
Down	26.59	26.80
Left	31.13	32.47
Right	28.40	30.19

Table 2. List of PSNR of Hallucinated image (input a face image on up view) $K=200$, $\theta = 0.2$

View	Ma's method	Our method
frontal	30.04	31.30
Up	33.64	34.18
Down	28.74	28.88
Left	32.92	34.68
Right	31.36	31.75

Table 3. List of PSNR of Hallucinated image (input a face image on down view) $K=200$, $\theta = 0.2$

View	Ma's method	Our method
frontal	29.61	32.93
Up	28.34	30.75
Down	33.20	34.17
Left	34.61	34.68
Right	32.70	34.47

Table 4. List of PSNR of Hallucinated image (input a face image on left view) $K=200$, $\theta = 0.2$

View	Ma's method	Our method
frontal	31.67	32.39
Up	30.55	31.84
Down	28.19	30.82
Left	33.44	34.29
Right	32.28	32.30

Table 5. List of PSNR of Hallucinated image (input a face image on right view) $K=200$, $\theta = 0.2$

View	Ma's method	Our method
frontal	26.98	28.76
Up	30.54	30.56
Down	26.23	26.24
Left	28.56	31.46
Right	33.12	33.75

6. Conclusion

In this paper, a novel multiview face hallucination method, using Adaptive Locally Linear Embedding (ALLE) technique. Experimental results show the higher quality of reconstructed image of our proposed framework over those enhanced with the baseline in both interpolation and learning methods based on the same training set. High-resolution face images of five different views are generated from a single low-resolution face image. According to the experimental results, the reconstructed image will be more accurate if the view of the input is same as that of the output. Especially, the frontal, up and down views achieve better estimation than others. The result of the proposed method show superior reconstruction quality of the HR face image over other related methods in both visualization and PSNR. However, the alignment on low-resolution face images may not be accurate. An automatic face alignment algorithm will be development in our future work.

7. Acknowledgment

This work has been supported by Research Grant for New Scholar (MRG5480162) from TRF(Thai Research Fund), CHE (Commission on Higher Education), and Panyapiwat Institute of Management, Thailand.

8. Reference

- [1] S. Baker, T. Kanade, "Hallucinating faces", In Proceedings of the International Conference on Automatic Face and Gesture Recognition, pp.83-88, 2000.
- [2] Baker, S., T. Kanade, "Limits on Super-Resolution and How to Break Them", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp.1167-83, 2002.
- [3] Elad, M., A. Feuer, "Super-resolution reconstruction of image sequences", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 9, pp. 817-834, 1999.
- [4] X.G. Wang, X.O. Tang, "Hallucinating Face by eigentransformation", IEEE Transactions on Systems Man and Cybernetics(Part C), vol. 35, no. 3, pp.425-434, 2005.
- [5] K. Jia, S. G. Gong, "Generalized face super-resolution", IEEE Transactions on Image Processing, vol. 17, no. 6, pp.873-886, 2008.
- [6] X. Ma, J. Zhang, C. Qi, "Position-based face hallucination method", In Proceedings of the IEEE Conference Multimedia and Expo, pp.290-293, 2009.
- [7] X. Ma, J. Zhang, C. Qi, "Hallucinating face by position-patch", Pattern Recognition, vol. 43, no. 6, pp.2224-2236, 2010.
- [8] Zhuang, Y., J. Zhang, F. Wu, "Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation", Pattern Recognition, vol. 40, no. 11, pp. 3178-3194, 2007.
- [9] Wei, L., L. Dahua, T. Xiaoou, "Neighbor combination and transformation for hallucinating faces", In Proceedings of the Multimedia and Expo, 2005. ICME 2005, IEEE International Conference on, 2005.
- [10] Park Sung Won, M. Savvides, "Robust Super-Resolution of Face Images by Iterative Compensating Neighborhood Relationships", In Proceedings of the Biometrics Symposium, pp. 1-5, 2007.
- [11] Fan Wei, Yeung Dit-Yan, "Image Hallucination Using Neighbor Embedding over Visual Primitive Manifolds", In Proceedings of the Computer Vision and Pattern Recognition, CVPR '07, IEEE Conference on (2007), pp. 1-7, 2007.
- [12] W. Fan, D.Y. Yeung, "Image Hallucination using Neighbor Embedding over visual primitive manifolds", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.1-7, 2007.
- [13] W. T. Freeman, E. C. Pasztor, "Learning Low Level Vision", In Proceedings of the 7th IEEE International Conference on Computer Vision, pp.1182-1189, 1999.
- [14] Hardie, R. C., K. J. Barnard, E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images", IEEE Transactions on Image Processing vol. 6, no. 12, pp.1621-1633, 1997.

- [15] Z. Lin, H.-Y. Shum, "Fundamental Limits of Reconstruction-Based Super resolution Algorithms under Local Translation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp.83-97, 2004.
- [16] Moon Gi, K. and S. Chaudhuri, "Super-resolution image reconstruction", *Signal Processing Magazine, IEEE* 20(3), 2003.
- [17] S. T. Roweis, L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science*, vol. 22, no. 12, pp.2323-2326, 2000.
- [18] Bing-Yu, Sun, Zhang Xiao-Ming, Li Jiuyong, Mao Xue-Min. "Feature Fusion Using Locally Linear Embedding for Classification", *IEEE Transactions on Neural Networks*, vol. 21, pp. 163-168, 2010.
- [19] H. Chang, D.Y. Yeung, Y.M. Xiong. "Super-resolution through Neighbor Embedding", In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, pp.1275-1282, 2004.
- [20] Xiang Ma, H. H., Shaopeng Wang, and Chun Qi (2010). "A Simple Approach to Multiview Face Hallucination" *IEEE SIGNAL PROCESSING*, vol. 17, no. 6, pp.579-582, 2010.
- [21] Thongdee M., Kimpan C., Sanguansat P., "Adaptive Locally Linear Embedding For Face Hallucination", In *Proceedings of the Asian Conference on Information Systems*, Siem Reap, Cambodia, 2012.
- [22] W. Gao, B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H.Zhang, D.B. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations", *IEEE Transactions on System Man, and Cybernetics (Part A)*, vol. 38, pp.149-161, 2008.
- [23] F. Zhou, B. Wang, Q. Liao, Super-resolution for face image by bilateral patches, *IET Electron. Letter* 48(18), pp.1125-1126, 2012.
- [24] Zhou, Fei, Biao Wang, Qingmin Liao. "Super-Resolution for Facial Image Using Multilateral Affinity Function.", *Neurocomputing*, vol. 133, pp. 194-208, 2014.
- [25] C. Jung, L. Jiao, B. Liu, M. Gong. "Position-patch based face hallucination using convex optimization", *IEEE Signal Processing Letters*, vol. 18, no. 6, pp.367-370, 2011.
- [26] Yu, H., L. Kin-Man, Q. Guoping and S. Tingzhi, "From Local Pixel Structure to Global Image Super-Resolution: A New Face Hallucination Framework." *IEEE Transactions on Image Processing*, vol. 20, no.2, pp.433-445, 2011.
- [27] Y. Hu, K.-Man Lam, T. Shen, W. Wang, A novel kernel-based framework for facial-image hallucination, *Image and Vision Computing*, vol. 29, pp.219-229, 2011.
- [28] Gao, G., J. Yang. "A novel sparse representation based framework for face image super-resolution." *Neurocomputing*, vol.134, pp. 92-99, 2014.

ประวัติ ผู้วิจัย

ชื่อ	เมธี ทองดี
วัน เดือน ปี เกิด	15 กุมภาพันธ์ พ.ศ. 2515
สถานที่เกิด	นครปฐม ประเทศไทย
ประวัติการศึกษา	สถาบันราชภัฏนครปฐม วิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์, 2540 มหาวิทยาลัยสยาม บริหารธุรกิจมหาบัณฑิต สาขาวิชา การจัดการระบบสารสนเทศ, 2547 มหาวิทยาลัยรังสิต ปรัชญาคุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ, 2557
ที่อยู่ปัจจุบัน	381 หมู่บ้านัญธานี ต.จอหอ อ.เมือง จ.นครราชสีมา
โทรศัพท์	081-7187332
อีเมล	metee.nrru@gmail.com
สถานที่ทำงาน	คณะวิทยาการจัดการ โปรแกรมวิชาคอมพิวเตอร์ธุรกิจ มหาวิทยาลัยราชภัฏนครราชสีมา