

การประยุกต์ใช้เทคนิคการขยายรายละเอียดบริเวณที่สนใจสำหรับระบบวีดิทัศน์ใฝ่สำรวจ



นายโกศล ปัญญาโสภณเลิศ

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

APPLICATION OF REGION-OF-INTEREST SUPERRESOLUTION TECHNIQUE
FOR VIDEO SURVEILLANCE SYSTEM



Mr. Kosol Punyasoponlert

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2007
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การประยุกต์ใช้เทคนิคการขยายรายละเอียดบริเวณที่สนใจสำหรับ
ระบบวิถีทัศน์ใฝ่ระยะวัง

โดย

นายโกศล ปัญญาโสภณเลิศ

สาขาวิชา

วิศวกรรมไฟฟ้า

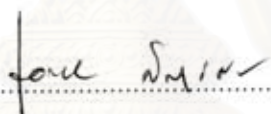
อาจารย์ที่ปรึกษา

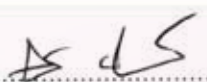
อาจารย์ ดร.สุรีย ทุมรินทร์

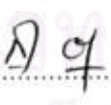
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต


..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.ติเรก ลาวัญย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.เอกชัย ลีลาวัศม์)


..... อาจารย์ที่ปรึกษา
(อาจารย์ ดร.สุรีย ทุมรินทร์)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุภาวดี อร่ามวิทย์)

สถาบันนิเทศบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โกศล ปัญญาโสภณเลิศ : การประยุกต์ใช้เทคนิคการขยายรายละเอียดบริเวณที่สนใจสำหรับระบบวีดิทัศน์เฝ้าระวัง (APPLICATION OF REGION-OF-INTEREST SUPERRESOLUTION TECHNIQUE FOR VIDEO SURVEILLANCE SYSTEM) อาจารย์ที่ปรึกษา : ดร.สุรีย พุ่มรินทร์, 68 หน้า

วิทยานิพนธ์ฉบับนี้นำเสนอการประยุกต์ใช้เทคนิคการขยายรายละเอียดภาพสำหรับเพิ่มภาพวีดิทัศน์จากระบบเฝ้าระวัง โดยออกแบบเป็นโปรแกรมในภาษา C++ สำหรับขยายรายละเอียดภาพในบริเวณที่สนใจ โดยใช้การซ้อนทับภาพโดยกำหนดตำแหน่งแบบอิงลักษณะเด่นร่วมกับการแปลงภาพแบบทศนิยม และสามารถซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติได้โดยใช้ระเบียบวิธีคัดกรองแรมแซค ในส่วนของการแก้สมการแบบจำลองของปัญหาได้เลือกใช้ระเบียบวิธีเชิงกำหนดแบบคุมค่า และคำนวณโดยระเบียบวิธีทำซ้ำแบบลาดลงขั้นที่สุด จากผลการทดลองพบว่าภาพที่ผ่านการขยายรายละเอียดโดยวิธีการที่นำเสนอมีอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด สูงกว่าภาพที่ผ่านการขยายรายละเอียดด้วยวิธีไบคิวบิกโดยเฉลี่ย 1.8 เดซิเบล และใช้เวลาประมวลผลประมาณ 10-12 วินาทีต่อ 1 ภาพที่ขนาด 320x240 จุดภาพ โดยภาพผลลัพธ์ที่ได้มีรายละเอียด ความคมชัด และความเปรียบต่างดีขึ้นอย่างเห็นได้ชัด ซึ่งจากผลการทดลองแสดงให้เห็นว่าโปรแกรมที่นำเสนอสามารถปรับปรุงรายละเอียดของภาพวีดิทัศน์ได้ใกล้เคียงกับภาพความละเอียดสูงต้นแบบมากกว่าวิธีที่ใช้ภาพความละเอียดต่ำเพียงภาพเดียว โดยไม่จำเป็นต้องปรับเปลี่ยนอุปกรณ์ในระบบบันทึกภาพ

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อนิสิต.....โกศล ปัญญาโสภณเลิศ.....
สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....AUS.....
ปีการศึกษา.....2550.....

4870232621 : MAJOR ELECTRICAL ENGINEERING

KEYWORDS : IMAGE SUPERRESOLUTION / AUTO-REGISTRATION / RANSAC / OPTICAL FLOW / REGULARIZED DETERMINISTIC METHOD / STEEPEST DESCENT ITERATIVE METHOD.

KOSOL PUNYASOPONLERT : APPLICATION OF REGION-OF-INTEREST SUPERRESOLUTION TECHNIQUE FOR VIDEO SURVEILLANCE SYSTEM. THESIS ADVISOR : SUREE PUMRIN, Ph.D., 68 pp.

This thesis proposes the technique of region-of-interest image Superresolution for image sequences recorded by surveillance video camera, presented in form of C++ software. Registration step is feature-based with image perspective transformation, processes automatically using RANSAC classification method. To solve the problem model equation, regularized deterministic method is chosen and Steepest Descent iterative method is used for calculation. The experiment results show the advantage of proposed method over conventional ones such as Bicubic methods with averagely 1.8 dB higher PSNR and processing time is 10-12 seconds approximately for a 320x240-pixel image. The result images show better details, contrast and higher sharpness obviously. According to experimental results, video image resolution improvement using proposed method is evident without any modification of system equipments.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department..... Electrical Engineering..... Student's signature..... *Kosol Punyasoponlert*.....
Field of Study..... Electrical Engineering..... Advisor's signature..... *Suree Pumrin*.....
Academic Year..... 2007.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีด้วยความช่วยเหลือจาก อ.ดร.สุรีย์ พุ่มรินทร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้คอยให้แนะนำที่เป็นประโยชน์อย่างยิ่งแก่ผู้เขียนตลอดช่วงเวลาที่ทำงานวิจัย อีกทั้งท่านยังได้ช่วยจัดหาเงินทุนสำหรับงานวิจัยแก่ผู้เขียนอีกด้วย

ขอขอบคุณโครงการการประเมินผลและวิเคราะห์ข่าวสารวิธีทัศน์สำหรับระบบตรวจการณ์ (VIPA) ที่ได้ให้เงินทุนสนับสนุนงานวิจัยในด้านที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้ และขอขอบคุณ คณาจารย์และเพื่อนๆ ในโครงการสำหรับคำแนะนำตลอดระยะเวลาของโครงการ

ขอขอบคุณพี่ๆ เพื่อนๆ และน้องๆ ทุกคนในห้องปฏิบัติการวิจัยออกแบบและประยุกต์วงจรรวมที่คอยให้ความช่วยเหลือ ให้ข้อคิดเห็นและคำแนะนำต่างๆ ตลอดมา

ขอขอบคุณเพื่อนๆ ในรุ่น 82 และพี่น้องวิศวฯ จุฬาฯ รุ่นอื่นๆ ทุกคนที่แวะเวียนมาเยี่ยมเยียน ผู้เขียนที่ห้องปฏิบัติการวิจัยเสมอมา

สุดท้ายนี้ ผู้เขียนใคร่ขอกราบขอบพระคุณบิดา มารดา ที่คอยให้การสนับสนุนทางด้านการศึกษา และให้อิสระในการเลือกเส้นทางชีวิตแก่ลูกๆ ตลอดมา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บทที่ 1 บทนำ.....	1
1.1 แนวคิดและเหตุผลในการทำวิทยานิพนธ์.....	1
1.2 วัตถุประสงค์ในการทำวิทยานิพนธ์.....	2
1.3 ขอบเขตของงานวิจัย.....	2
1.4 ขั้นตอนการดำเนินงานวิจัย.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
1.6 ลำดับขั้นตอนในการนำเสนอผลงานวิจัย.....	3
บทที่ 2 ปรัชญาศรัทธาธรรม.....	4
2.1 หลักการพื้นฐานการขยายรายละเอียดภาพ.....	4
2.1.1 เงื่อนไขการขยายรายละเอียดภาพ.....	5
2.2 การขยายรายละเอียดภาพจากแฟ้มภาพวีดิทัศน์.....	6
2.3 แบบจำลองพื้นฐานของปัญหา.....	7
2.4 งานวิจัยที่ผ่านมา.....	8
2.4.1 ระเบียบวิธีประมาณค่าในช่วงแบบไม่เอกรูป.....	8
2.4.2 ระเบียบวิธีเชิงโดเมนความถี่.....	9
2.4.3 ระเบียบวิธีขายกลับแบบทำซ้ำ.....	10
2.4.4 ระเบียบวิธีขายลงบนเซตฐาน.....	10
2.4.5 ระเบียบวิธีเชิงกำหนดแบบคุมค่า.....	11
2.4.6 ระเบียบวิธีเชิงสุ่มแบบคุมค่า.....	12
2.5 สรุปท้ายบท.....	14
บทที่ 3 ปัญหาการสร้างกลับภาพความละเอียดสูง.....	16
3.1 ปัจจัยลดทอนคุณภาพในแบบจำลองของปัญหา.....	16

3.1.1 ความพร่ามัว.....	16
3.1.2 สัญญาณรบกวน.....	18
3.2 ตัวแปรและตัวดำเนินการต่างๆ ในแบบจำลองของปัญหา.....	19
3.2.1 ภาพความละเอียดต่ำ.....	20
3.2.2 ภาพความละเอียดสูง.....	21
3.2.3 ตัวดำเนินการบิดโค้ง.....	21
3.2.4 ตัวดำเนินการพร่ามัว.....	23
3.2.5 ตัวดำเนินการชักสุมตัวอย่างขาด.....	24
3.2.6 สัญญาณรบกวน.....	25
3.3 ปัญหาผกผันและปัญหาลักษณะเลว.....	25
3.4 เงื่อนไขและลักษณะของเพิ่มภาพวีดิทัศน์ที่ใช้ในงานวิจัย.....	26
3.4.1 การบิดโค้งเชิงเส้น.....	26
3.4.2 ความเร็วเคลื่อนที่ของวัตถุต่ำ.....	27
3.4.3 ความละเอียดเชิงพื้นที่สูงเพียงพอ.....	28
3.5 สรุปท้ายบท.....	28
บทที่ 4 ระเบียบวิธีที่เลือกใช้ในการออกแบบโปรแกรม.....	29
4.1 การซ่อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ.....	29
4.1.1 กรณีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด.....	29
4.1.2 ระเบียบวิธีคัดกรองแรนแซค.....	30
4.1.3 กรณีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่.....	33
4.2 การรวมภาพ.....	34
4.2.1 การรวมภาพแบบเฉลี่ยไม่ถ่วงน้ำหนัก.....	34
4.2.2 การรวมภาพแบบเฉลี่ยถ่วงน้ำหนัก.....	35
4.3 ระเบียบวิธีที่ใช้ในการแก้สมการแบบจำลองของปัญหา.....	35
4.4 การขยายรายละเอียดภาพสี.....	37
4.5 สรุปท้ายบท.....	37
บทที่ 5 การออกแบบโปรแกรมขยายรายละเอียดภาพ.....	39
5.1 ไบรารีโอเพ่นซีวี.....	39
5.2 ฟังก์ชันหลักที่ใช้ในการเขียนโปรแกรม.....	39
5.2.1 ฟังก์ชันในไลบรารีโอเพ่นซีวี.....	39
5.2.2 ฟังก์ชันที่เขียนขึ้นมาใหม่.....	41

5.3 การออกแบบโปรแกรม.....	42
5.4 โปรแกรมขยายรายละเอียดภาพที่นำเสนอ	44
5.5 สรุปท้ายบท	46
บทที่ 6 ผลการทดสอบโปรแกรม	47
6.1 ดัชนีที่ใช้ในการวัดผลการทดสอบ	47
6.2 ผลการทดสอบกับเพิ่มภาพสังเคราะห์	47
6.2.1 กรณีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด	48
6.2.2 กรณีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่	52
6.3 ผลการทดสอบกับเพิ่มภาพทั่วไป	53
6.4 เวลาที่ใช้ในการประมวลผล.....	54
6.5 สรุปท้ายบท	55
บทที่ 7 ข้อสรุปและข้อเสนอแนะ.....	56
7.1 สรุป	56
7.2 ข้อเสนอแนะและข้อควรปรับปรุง	57
รายการอ้างอิง	59
บรรณานุกรม.....	60
ภาคผนวก	62
ภาคผนวก ก	
บทความที่ได้รับการตีพิมพ์ในการประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 30.....	63
ประวัติผู้เขียนวิทยานิพนธ์	68

สารบัญตาราง

	หน้า
ตารางที่ 2-1 งานวิจัยด้านการขยายรายละเอียดภาพที่มีมาก่อนหน้านี้.....	14
ตารางที่ 6-1 สรุปผลการทดลองการเพิ่มภาพที่มีการเคลื่อนไหวแบบเลื่อนขนานทั้งหมด	51



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

	หน้า
รูปที่ 2-1 การขยายรายละเอียดภาพด้วยภาพวิธีการประมาณค่าในช่วงด้วยระเบียบวิธีต่างๆ	4
รูปที่ 2-2 การสร้างกลับภาพความละเอียดสูง.....	5
รูปที่ 2-3 ระยะเลื่อนระหว่างภาพ 2 ภาพ	5
รูปที่ 2-4 การขยายรายละเอียดภาพที่มีตัวประกอบขยายเท่ากับ 2.....	6
รูปที่ 2-5 การขยายรายละเอียดภาพจากแฟ้มภาพวีดิทัศน์.....	7
รูปที่ 2-6 การซ้อนทับภาพโดยกำหนดตำแหน่ง	8
รูปที่ 2-7 การประมาณค่าในช่วงแบบไม่เอกรูป.....	9
รูปที่ 3-1 การเกิดความพรั่มัวเชิงแสง	17
รูปที่ 3-2 ผลของความพรั่มัวเชิงแสง	17
รูปที่ 3-3 ผลของความพรั่มัวเชิงเคลื่อนที่	17
รูปที่ 3-4 ภาพที่บันทึกจากตัวตรวจรู้ภาพที่มีความละเอียดแตกต่างกัน.....	18
รูปที่ 3-5 กราฟสัญญาณรบกวน 1 มิติ ทั้ง 3 ประเภท	19
รูปที่ 3-6 จุดควบคุมในแต่ละภาพ	22
รูปที่ 3-7 การจับคู่จุดควบคุมระหว่างภาพ 2 ภาพ.....	22
รูปที่ 3-8 การแปลงเรขาคณิตแบบต่างๆ	23
รูปที่ 3-9 ตัวดำเนินการชักสุมตัวอย่างแบบเลือกค่า	24
รูปที่ 3-10 ตัวดำเนินการชักสุมตัวอย่างแบบเฉลี่ยค่า	24
รูปที่ 3-11 แฟ้มภาพวีดิทัศน์ที่มีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด	26
รูปที่ 3-12 แฟ้มภาพวีดิทัศน์ที่มีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่.....	26
รูปที่ 3-13 ภาพใบหน้าคนซึ่งเป็นวัตถุแบบไม่คงรูป.....	27
รูปที่ 3-14 ภาพใบหน้าคนซึ่งเกิดการบิดบังตัวเอง.....	27
รูปที่ 4-1 การค้นหาตำแหน่งของลักษณะเด่นในภาพ.....	30
รูปที่ 4-2 การจับคู่ลักษณะเด่นที่สอดคล้อง	30
รูปที่ 4-3 การคำนวณเมตริกซ์แปลงภาพจากจุดควบคุมจำนวน 4 คู่	31
รูปที่ 4-4 การหาตำแหน่งทำนายของจุดควบคุมในเฟรมอ้างอิงด้วยเมตริกซ์แปลงภาพ.....	31
รูปที่ 4-5 ความคลาดเคลื่อนระหว่างตำแหน่งทำนายกับตำแหน่งจริงของจุดควบคุมในเฟรมข้างเคียง .	32
รูปที่ 4-6 ชุดของจุดควบคุมชุดใหม่สำหรับการคำนวณในรอบต่อไป	32
รูปที่ 4-7 การรวมภาพแบบเฉลี่ยไม่ถ่วงน้ำหนัก	34

รูปที่ 4-8 การรวมภาพแบบเฉลี่ยถ่วงน้ำหนัก.....	35
รูปที่ 4-9 การขยายรายละเอียดภาพสี	38
รูปที่ 5-1 แผนภาพการทำงานของโปรแกรมขยายรายละเอียดภาพที่นำเสนอ	42
รูปที่ 5-2 ผังงานการทำงานของโปรแกรม.....	43
รูปที่ 5-3 หน้าต่างแสดงผลของโปรแกรมขยายรายละเอียดภาพ.....	45
รูปที่ 6-1 ผลการทดลองที่ 1	48
รูปที่ 6-2 ผลการทดลองที่ 2.....	49
รูปที่ 6-3 ผลการทดลองที่ 3.....	50
รูปที่ 6-4 ผลการทดลองที่ 4.....	51
รูปที่ 6-5 ผลการทดลองที่ 5.....	52
รูปที่ 6-6 ผลการทดลองที่ 6.....	53
รูปที่ 6-7 ผลการทดลองที่ 7.....	54
รูปที่ 6-8 ค่า PSNR ที่รอบทำซ้ำต่างๆ ในการทดลองที่ 1.....	55

บทที่ 1

บทนำ

1.1 แนวคิดและเหตุผลในการทำวิทยานิพนธ์

ในปัจจุบันได้มีการนำระบบบันทึกภาพเคลื่อนไหวมาใช้ในระบบรักษาความปลอดภัยและระบบเฝ้าระวังตามสถานที่ต่างๆ มากขึ้น โดยใช้กล้องวีดิทัศน์วงจรปิดติดตั้งตามจุดต่างๆ เพื่อบันทึกภาพบุคคลและเหตุการณ์ที่น่าสงสัย และจัดเก็บเป็นข้อมูลตามช่วงเวลาในรูปแบบดิจิทัลเพื่อประโยชน์ในการตรวจสอบภายหลังเมื่อมีเหตุร้ายเช่น การโจรกรรม หรือเหตุวินาศกรรมเกิดขึ้น อย่างไรก็ตาม การเพิ่มภาพวีดิทัศน์ที่บันทึกนี้บางครั้งไม่สามารถให้ข้อมูลที่สำคัญเช่น ใบหน้าผู้ต้องสงสัย แผ่นป้ายทะเบียนพาหนะหรือลักษณะของวัตถุต้องสงสัยแก่ผู้วิเคราะห์ได้เนื่องจากมีรายละเอียดของวัตถุในภาพไม่เพียงพอ ซึ่งเป็นผลมาจากข้อจำกัดต่างๆ ที่เกิดขึ้นระหว่างการบันทึกภาพ เช่น ระยะเวลาบันทึกภาพไกลเกินไป ความพัวมัว (Blur) ซึ่งเกิดจากระบบทัศนูปกรณ์ในกล้อง หรือความพัวมัวที่เกิดจากการที่ตัวตรวจจับภาพ (Image Sensor) มีความละเอียดไม่เพียงพอที่จะเก็บรายละเอียดของวัตถุที่ต้องการได้ครบถ้วน

จากปัญหาข้างต้นทำให้การเพิ่มคุณภาพของภาพ (Image Enhancement) สำหรับเพิ่มภาพวีดิทัศน์เป็นหัวข้อที่ได้รับความสนใจจากนักวิจัยและนักวิชาการในวงกว้างและมีการค้นคว้าพัฒนาตั้งแต่ปี ค.ศ. 1970 มาโดยตลอด [1] โดยเริ่มจากการศึกษาและพัฒนาหลักการขยายรายละเอียดสำหรับชุดของภาพนิ่ง ภาพที่ผ่านการเพิ่มคุณภาพแล้วนั้นอาจมีความละเอียดเชิงพื้นที่ (Spatial Resolution) เพิ่มขึ้นทำให้สามารถเห็นรายละเอียดต่างๆ ได้ดีขึ้น หรืออาจมีความละเอียดเชิงพื้นที่เท่าเดิมแต่ความพัวมัวลดลง ความคมชัดและความเปรียบต่าง (Sharpness & Contrast) ดีขึ้นซึ่งจัดเป็นการบูรณะภาพ (Image Restoration) โดยจำนวนของภาพต้นแบบที่ต้องใช้ในการสร้างภาพผลลัพธ์นั้นอาจเป็นภาพเดียว ภาพหลายๆ ภาพของวัตถุเดียวกันซึ่งบันทึกจากหลายมุมกล้อง หรือเป็นชุดของภาพต่อเนื่องจากเพิ่มภาพเคลื่อนไหวก็ได้ โดยขึ้นอยู่กับขั้นตอนวิธี (Algorithm) ที่ใช้

การขยายรายละเอียดภาพ (Image Superresolution) ซึ่งเป็นระเบียบวิธีสำหรับเพิ่มความละเอียดเชิงพื้นที่ของภาพให้สูงขึ้น นับเป็นหนึ่งในวิธีการเพิ่มคุณภาพของภาพนิ่งและภาพเคลื่อนไหวซึ่งได้รับความสนใจจากนักวิจัยมาเป็นเวลานาน เนื่องจากการใช้ข้อมูลของภาพที่มีอยู่ได้อย่างเกิดประโยชน์สูงสุด และไม่จำเป็นต้องปรับเปลี่ยนแก้ไขอุปกรณ์ในระบบบันทึกภาพ รวมถึงไม่ต้องใช้พื้นที่จัดเก็บข้อมูลเพิ่มเติม ทำให้ประหยัดค่าใช้จ่ายและสามารถนำระเบียบวิธีนี้ไปประยุกต์ใช้กับระบบบันทึกภาพทั่วไป ซึ่งรวมถึงอุปกรณ์บันทึกภาพระดับผู้บริโภค เช่น กล้องคอมแพคต์ดิจิทัล หรือกล้องถ่ายรูปดิจิทัลในโทรศัพท์เคลื่อนที่ซึ่งมีขายทั่วไปได้อีกด้วย

1.2 วัตถุประสงค์ในการทำวิทยานิพนธ์

- 1) ศึกษาและพัฒนาขั้นตอนวิธีขยายรายละเอียดภาพบริเวณที่สนใจ
- 2) ออกแบบโปรแกรมคอมพิวเตอร์สำหรับขยายรายละเอียดภาพเพิ่มวีดิทัศน์ด้วยภาษา C/C++

1.3 ขอบเขตของงานวิจัย

ออกแบบโปรแกรมคอมพิวเตอร์ด้วยภาษา C/C++ เพื่อรับภาพจากแฟ้มภาพวีดิทัศน์ชนิด AVI มาผ่านการขยายรายละเอียดในบริเวณที่สนใจ โดยมีคุณสมบัติทางเทคนิคดังนี้

- 1) สามารถเลือกตัวประกอบการขยายได้ในช่วง 2 - 4 เท่า
- 2) เลือกบริเวณที่ต้องการขยายรายละเอียดภาพได้
- 3) ซ้อนทับภาพโดยกำหนดตำแหน่งโดยอัตโนมัติสำหรับกรณีที่วัตถุในภาพมีการเคลื่อนไหวแบบเลื่อนขนาน และไม่มีปัญหาการบิดบังได้
- 4) ให้ภาพผลลัพธ์ที่ถูกขยายรายละเอียดเป็นภาพนิ่งสีหรือขาวดำแบบ 8 บิต

1.4 ขั้นตอนการดำเนินงานวิจัย

- 1) ศึกษาขั้นตอนวิธีการขยายรายละเอียดภาพจากรายงานที่มีผู้วิจัยอยู่ก่อนหน้า
- 2) ทดสอบขั้นตอนวิธี และเขียนโปรแกรมขยายรายละเอียดภาพด้วยภาษา C++ ร่วมกับไลบรารี OpenCV เพื่อดูผลลัพธ์ได้แก่ระดับความละเอียดที่เพิ่มขึ้น และเวลาที่ใช้ในการประมวลผล
- 3) ทดสอบขั้นตอนวิธีและโปรแกรมกับแฟ้มภาพวีดิทัศน์
- 4) ปรับปรุงแก้ไขข้อผิดพลาดที่เกิดขึ้น
- 5) สรุปและเขียนวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้รับความรู้ความเข้าใจเกี่ยวกับการเขียนโปรแกรมภาษา C++ ควบคู่ไปการใช้ไลบรารี OpenCV ในการประมวลสัญญาณภาพแบบดิจิทัล
- 2) ได้รับความรู้ความเข้าใจในการแก้ระบบสมการขนาดใหญ่เพื่อหาค่าที่เหมาะสมที่สุดโดยระเบียบวิธีทำซ้ำ
- 3) เพิ่มประสิทธิภาพให้กับระบบเฝ้าระวังภัยอัตโนมัติ
- 4) สามารถนำไปประยุกต์ใช้ในงานประมวลผลภาพต่างๆ ที่ต้องการภาพถ่ายดิจิทัลความละเอียดสูง หรือเพิ่มความละเอียดของภาพ เช่น ระบบตรวจสอบป้ายทะเบียนยานพาหนะบนท้องถนน กล้องถ่ายภาพดิจิทัลความละเอียดต่ำ งานถ่ายภาพทางด้านดาราศาสตร์ เป็นต้น

1.6 ลำดับขั้นตอนในการนำเสนองานวิจัย

งานวิจัยชิ้นนี้แบ่งออกเป็น 7 บท โดยบทที่ 2 จะกล่าวถึงหลักการพื้นฐานของการขยายรายละเอียดภาพ แบบจำลองพื้นฐานของปัญหา และงานวิจัยที่เกี่ยวข้องในช่วงที่ผ่านมา บทที่ 3 กล่าวถึงปัญหาการสร้างกลับภาพความละเอียดสูง แบบจำลอง ตัวแปรและตัวดำเนินการที่มีผลต่อปัญหา รวมถึงรายละเอียดของระเบียบวิธีที่เลือกใช้ในงานวิจัยนี้ จากนั้นในบทที่ 4 จะกล่าวถึงขั้นตอนวิธีที่ใช้ในการออกแบบโปรแกรม ซึ่งได้แก่การซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ การรวมภาพ และการแก้สมการแบบจำลองของปัญหา รวมถึงการขยายรายละเอียดภาพสี่ สำหรับบทที่ 5 จะอธิบายถึงการออกแบบโปรแกรมขยายรายละเอียดภาพ ไคบรารีโอเฟนซีวี และฟังก์ชันหลักที่ใช้ในการประมวลผล จากนั้นจะนำเสนอผลการทดสอบโปรแกรมกับภาพวิดีโอที่สนใจความละเอียดต่ำภายใต้เงื่อนไขต่างๆ ในบทที่ 6 สุดท้ายในบทที่ 7 จะเป็นการสรุปผลงานวิจัยและรวบรวมข้อเสนอแนะสำหรับปรับปรุงและพัฒนา งานวิจัยต่อไป



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ปริทรรศน์วรรณกรรม

ในบทนี้จะกล่าวถึงหลักการพื้นฐานการขยายรายละเอียดภาพ แบบจำลองพื้นฐานของปัญหา และงานวิจัยที่เกี่ยวข้องในช่วงที่ผ่านมาโดยสังเขป ระเบียบวิธีขยายรายละเอียดภาพแบบต่างๆ และข้อดีข้อเสียของแต่ละระเบียบวิธีในงานวิจัยเหล่านี้

2.1 หลักการพื้นฐานการขยายรายละเอียดภาพ

การขยายรายละเอียดภาพเป็นการเพิ่มความละเอียดเชิงพื้นที่ของภาพ (Spatial Resolution) กล่าวคือจำนวนจุดภาพของภาพความละเอียดต่ำ (Low Resolution Image: LR) ให้เพิ่มมากขึ้นเป็นภาพความละเอียดสูง (High Resolution Image: HR) โดยมีตัวประกอบการขยาย (Enhancement Factor: r) เป็นตัวกำหนดขนาดของภาพผลลัพธ์ที่ได้ ตัวอย่างเช่น ภาพต้นแบบความละเอียดต่ำขนาด 60×80 จุดภาพ ถูกเพิ่มความละเอียดด้วยตัวประกอบการขยายเท่ากับ 4 จะทำให้ได้ภาพผลลัพธ์ความละเอียดสูงขนาด 240×320 จุดภาพ โดยปกติการขยายรายละเอียดภาพสามารถทำได้โดยอาศัยฟังก์ชันประมาณค่าในช่วง (Interpolation Function) ในซอฟต์แวร์ประมวลผลภาพทั่วไป เช่น การประมาณในช่วงแบบไบลิเนียร์ (Bilinear Interpolation) หรือแบบไบคิวบิก (Bicubic Interpolation) ดังแสดงในรูปที่ 2-1



60 x 84 Pixel
Low Resolution Image

240 x 336 Pixels
Using Bilinear Interpolation

240 x 336 Pixels
Using Bicubic Interpolation

รูปที่ 2-1 การขยายรายละเอียดภาพด้วยวิธีการประมาณค่าในช่วงด้วยระเบียบวิธีต่างๆ

อย่างไรก็ดี เนื่องจากระเบียบวิธีเหล่านี้อาศัยข้อมูลที่มีอยู่ในภาพความละเอียดต่ำเพียงภาพเดียว ทำให้ไม่มีรายละเอียดเพิ่มขึ้นในภาพผลลัพธ์แม้จะมีความละเอียดเชิงพื้นที่สูงขึ้น เพื่อให้ได้ภาพที่มีรายละเอียดของวัตถุในภาพมากขึ้นจึงได้มีแนวคิดในการนำภาพของวัตถุเดียวกันแต่ถูกบันทึกจากมุมมองหรือช่วงเวลาต่างกันหลายๆ ภาพมารวบรวมข้อมูลที่แตกต่างกันเพื่อนำมารวมอยู่ในภาพเพียง

ภาพเดี่ยวดังแสดงในรูปที่ 2-2 จึงสามารถสร้างเป็นภาพความละเอียดสูงที่มีข้อมูลเพิ่มขึ้น โดยภาพต้นแบบความละเอียดต่ำนี้อาจเป็นชุดภาพต่อเนื่องที่ได้จากภาพเคลื่อนไหว หรือภาพของวัตถุเดียวกัน แต่ถูกบันทึกด้วยมุมกล้องที่แตกต่างกันก็ได้



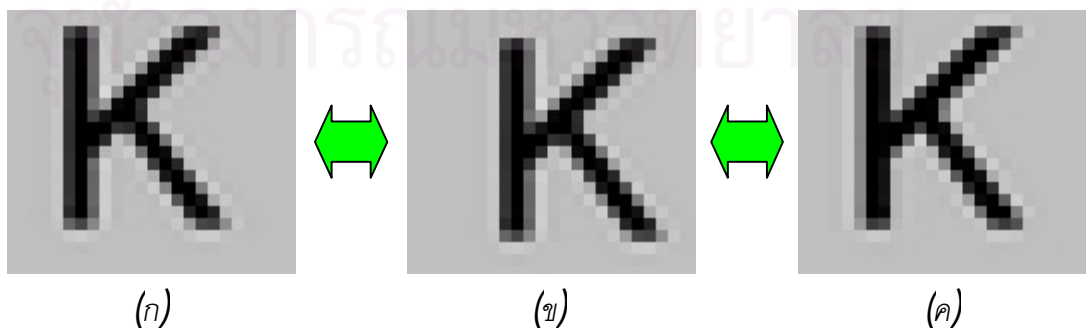
รูปที่ 2-2 การสร้างกลับภาพความละเอียดสูง

2.1.1 เงื่อนไขการขยายรายละเอียดภาพ

ปัจจัยสำคัญที่ทำให้กระบวนการขยายรายละเอียดภาพมีความเป็นไปได้ ได้แก่

1) ระยะเวลาไม่เต็มจุดภาพ (Sub-pixel Shift)

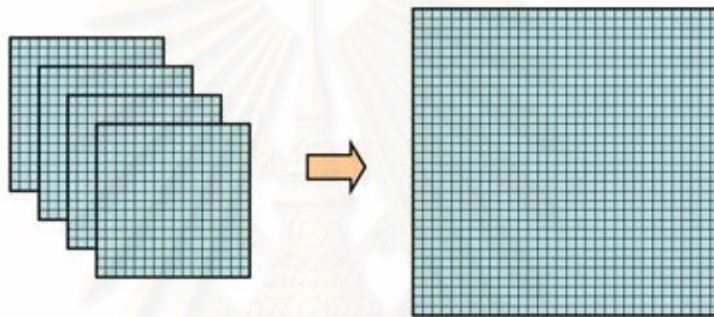
เกิดจากการที่วัตถุในภาพต้นแบบความละเอียดต่ำแต่ละภาพมีการเลื่อนตำแหน่งของวัตถุในภาพเป็นระยะใดๆ (Arbitrary Shift) ซึ่งหากระยะที่วัตถุเลื่อนที่ไประหว่างแต่ละภาพเป็นจำนวนเต็มจุดภาพพอดี เมื่อผ่านการซักสุ่มตัวอย่างอย่างช้าลงด้วยตัวตรวจรู้ภาพจะทำให้ไม่เกิดข้อมูลที่แตกต่างระหว่างแต่ละภาพ จึงไม่สามารถนำไปสร้างเป็นภาพความละเอียดสูงที่มีปริมาณข้อมูลมากขึ้นได้ โดยระยะดังกล่าวเรียกว่าระยะเลื่อนวิกฤต (Critical Shift) ดังแสดงในรูปที่ 2-3 (ก)-(ข) ซึ่งเห็นได้ว่าวัตถุในภาพทั้งสองมีลักษณะเหมือนกันทุกประการ ซึ่งต่างจากกรณีของภาพในรูปที่ 2-3 (ข)-(ค) ซึ่งมีระยะเลื่อนไม่เป็นค่าวิกฤต



รูปที่ 2-3 ระยะเวลาเลื่อนระหว่างภาพ 2 ภาพ (ก)-(ข) เป็นค่าวิกฤต (ข)-(ค) ไม่เป็นค่าวิกฤต

2) จำนวนภาพความละเอียดต่ำที่เพียงพอ

เนื่องจากการขยายรายละเอียดภาพอาศัยหลักการที่ว่า ยังมีข้อมูลที่แตกต่างกัน (ชุดภาพความละเอียดต่ำ) เป็นจำนวนมาก ก็จะสามารถนำข้อมูลที่จำเป็นมาบรรจุรวมกันในภาพความละเอียดสูงได้ โดยหากต้องการสร้างภาพความละเอียดสูงขนาด $rM \times rN$ จากชุดภาพความละเอียดต่ำขนาด $M \times N$ ต้องใช้ภาพความละเอียดต่ำเป็นจำนวนไม่ต่ำกว่า r^2 ภาพ ซึ่งในทางทฤษฎีถือว่า มีข้อมูลซึ่งได้แก่จุดภาพที่แตกต่างกันเป็นจำนวนมากเพียงพอที่จะนำไปสร้างเป็นภาพความละเอียดสูงดังกล่าวได้ ดังแสดงตัวอย่างในรูปที่ 2-4 ซึ่งเป็นการขยายรายละเอียดภาพด้วยตัวประกอบการขยายเท่ากับ 2 นั่นคือเป็นการสร้างภาพความละเอียดสูงขนาด 32×32 จุดภาพ จากภาพความละเอียดต่ำขนาด 16×16 จุดภาพ จำนวน 4 ภาพ



รูปที่ 2-4 การขยายรายละเอียดภาพที่มีตัวประกอบการขยายเท่ากับ 2

เงื่อนไขข้อที่ 2 นี้ นับเป็นปัญหาที่มักพบในกระบวนการขยายรายละเอียดภาพ เนื่องจากในทางปฏิบัติพบว่าภาพความละเอียดต่ำที่ได้จากระบบบันทึกภาพมักมีจำนวนน้อยกว่าจำนวนขั้นต่ำ ทำให้กลายเป็นปัญหาค่ากำหนดขาด (**Underdetermined System**) จึงจำเป็นต้องหาวิธีการประมาณค่าข้อมูลของจุดภาพที่ขาดหายไป ดังจะได้กล่าวถึงโดยละเอียดในบทที่ 3

2.2 การขยายรายละเอียดภาพจากแฟ้มภาพวิดีโอ

โดยหลักการที่กล่าวมาข้างต้น เมื่อพิจารณาชุดภาพต่อเนื่องที่ได้จากแฟ้มภาพวิดีโอว่าเป็นภาพของวัตถุเดียวกันที่ถูกบันทึกด้วยมุมกล้องต่างกันและที่เวลาต่างๆ กันดังรูปที่ 2-5 ทำให้สามารถนำหลักการขยายรายละเอียดภาพมาประยุกต์ใช้กับแฟ้มภาพวิดีโอได้ โดยมีเงื่อนไขว่าในชุดภาพต่อเนื่องควรจะมีวัตถุที่สนใจปรากฏอยู่ทุกภาพ และวัตถุดังกล่าวมีการเปลี่ยนแปลงไปเล็กน้อยเพื่อให้เกิดข้อมูลของภาพวัตถุที่แตกต่างกันในแต่ละภาพ แต่ไม่แตกต่างกันมากเกินไปเพื่อให้การระบุตำแหน่งจุดภาพของวัตถุนั้นเฟรมภาพความละเอียดสูงเป็นไปอย่างถูกต้อง โดยในงานวิจัยนี้ได้เลือกเฟรมภาพที่สนใจและเฟรมภาพข้างเคียงจำนวนรวม 8 ภาพสำหรับการขยายรายละเอียดด้วยตัวประกอบการขยายในช่วงตั้งแต่ 2 ถึง 4 เท่า



รูปที่ 2-5 การขยายรายละเอียดภาพจากแฟ้มภาพวิดีโอ

2.3 แบบจำลองพื้นฐานของปัญหา

การขยายรายละเอียดภาพโดยทั่วไปนิยมเริ่มจากการตั้งแบบจำลองที่อธิบายกระบวนการบันทึกจากเหตุการณ์ซึ่งถูกมองว่าเป็นภาพความละเอียดสูง ผ่านกระบวนการลดทอนคุณภาพเป็นภาพความละเอียดต่ำ [1] ดังแสดงในสมการ (2.1)

$$y = H * x + h \quad (2.1)$$

โดย x คือภาพต้นแบบ คุณประสาน (Convolute) กับตัวดำเนินการลดทอนคุณภาพ H (Degrading Operator) จากนั้นจึงบวกเข้ากับสัญญาณรบกวน h ทำให้ได้ภาพผลลัพธ์ที่ถูกลดทอนคุณภาพ y ตัวดำเนินการ H เป็นผลรวมของตัวดำเนินการลดทอนคุณภาพย่อย เช่น ความพร่ามัว ชนิดต่างๆ การซ้กซึมตัวอย่างขาลง เป็นต้น โดยจะได้กล่าวถึงรายละเอียดในบทที่ 3

จากสมการ (2.1) หากให้ H และ h มีค่าแตกต่างกันทั้งหมด L แบบ ก็จะสามารถสร้างภาพความละเอียดต่ำต่างๆ กันจำนวน L ภาพได้จากภาพความละเอียดสูงเพียงภาพเดียว ดังสมการ

$$y_k = H_k * x + h_k \quad k = 1, 2, 3, \dots, L \quad (2.2)$$

โดย y_k คือภาพความละเอียดต่ำลำดับที่ k ที่ได้จากการนำภาพความละเอียดสูง x ผ่านกระบวนการลดทอนคุณภาพลำดับที่ k จะเห็นได้ว่าการขยายรายละเอียดภาพก็คือกระบวนการสร้างกลับภาพความละเอียดสูงภาพหนึ่งจากภาพความละเอียดต่ำจำนวนหลายๆ ภาพนั่นเอง ปัญหาที่จัดเป็นปัญหาประเภทผกผัน (Inverse Problem) ซึ่งเป็นการคำนวณหาสัญญาณขาเข้าของระบบซึ่งในที่นี้คือภาพความละเอียดสูงต้นแบบ โดยการประมาณจากสัญญาณขาออกนั้นคือภาพความละเอียดต่ำ

ที่มีอยู่ ในการหาผลเฉลยของปัญหาจึงต้องทำการประมาณค่าตัวแปรและตัวดำเนินการ รวมถึงวิธีการแก้สมการที่เหมาะสม

2.4 งานวิจัยที่ผ่านมา

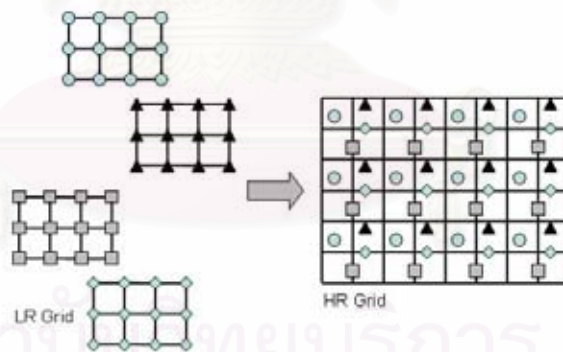
ในอดีตที่ผ่านมาได้มีกลุ่มผู้สนใจทำการวิจัยเรื่องการขยายรายละเอียดภาพเป็นจำนวนมาก ซึ่งในแต่ละกลุ่มก็ได้เสนอวิธีการที่แตกต่างกันไป โดยสามารถจำแนกชั้นตอนวิธีตามแนวทางหลักๆ ได้ดังนี้

2.4.1 ระเบียบวิธีประมาณค่าในช่วงแบบไม่เอกรูป (Non-uniform Interpolation Method)

หรืออีกชื่อหนึ่งว่าระเบียบวิธีเลื่อนและบวก (Shift & Add Method) [1, 2] จัดเป็นระเบียบวิธีแบบขั้นตอนเดียวซึ่งใช้เวลาคำนวณต่ำและเป็นระเบียบวิธีที่เรียบง่ายที่สุด แบ่งเป็น 3 ขั้นตอนหลักได้แก่

- การซ้อนทับภาพโดยกำหนดตำแหน่ง (Image Registration)

เป็นการระบุตำแหน่งของวัตถุในแต่ละภาพเมื่อเทียบกับภาพอ้างอิงภาพหนึ่งๆ เพื่อซ้อนทับภาพวัตถุลงบนกรอบอ้างอิงของภาพความละเอียดสูงได้อย่างถูกต้องดังรูปที่ 2-6 ซึ่งนับเป็นขั้นตอนที่สำคัญมากในกระบวนการขยายรายละเอียดภาพ เนื่องจากเป็นขั้นตอนที่ระบุว่าข้อมูลจากจุดภาพหนึ่งๆ ควรจะอยู่ ณ ตำแหน่งใดในกรอบภาพความละเอียดสูง ภาพผลลัพธ์ที่ได้จะมีรายละเอียดดีเพียงใดนั้นจึงขึ้นอยู่กับความถูกต้องแม่นยำของขั้นตอนนี้เป็นหลัก

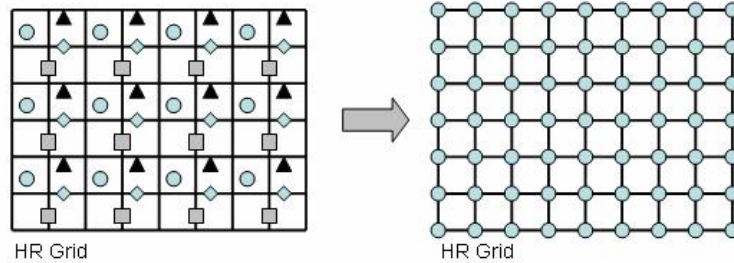


รูปที่ 2-6 การซ้อนทับภาพโดยกำหนดตำแหน่ง

หลังจากที่ทราบตำแหน่งของจุดภาพบนกริดภาพความละเอียดสูงแล้ว จึงนำข้อมูลจุดภาพทั้งหมดมาบรรจุรวมกันในเฟรมภาพความละเอียดสูง ซึ่งบางครั้งเรียกขั้นตอนนี้ร่วมกับการซ้อนทับภาพโดยกำหนดตำแหน่งว่า การรวมภาพ (Image Fusion)

- การประมาณค่าในช่วงแบบไม่เอกรูป (Non-uniform Interpolation)

เนื่องจากจุดภาพที่ถูกแปลงไปสู่พิกัดในภาพอ้างอิงนั้นอาจไม่ตรงกับค่าพิกัดของกริด (grid) ของภาพความละเอียดสูงซึ่งเป็นเลขจำนวนเต็มพอดี จึงต้องมีการประมาณค่าในช่วงเพื่อคำนวณค่าจุดภาพที่เหมาะสมที่สุดสำหรับแต่ละตำแหน่งพิกัดดังแสดงตัวอย่างในรูปที่ 2-7



รูปที่ 2-7 การประมาณค่าในช่วงแบบไม่เอกรูป

- การกำจัดความพร่ามัว (Deblurring)

เป็นการกำจัดหรือลดผลจากความพร่ามัว เพื่อเพิ่มความเปรียบต่าง (Contrast) ในภาพ ลัพธ์ให้มากขึ้น รวมถึงการเพิ่มความคมชัดของลักษณะเด่นในภาพที่เป็นมุมและขอบ (Corners & Edges) ซึ่งเป็นรายละเอียดที่สำคัญในภาพ โดยประมวลผลร่วมกับตัวเนกาทีฟที่ทราบค่าหรือโดยประมาณจากภาพต้นแบบ

ขั้นตอนการขยายรายละเอียดภาพทั้ง 3 ในระเบียบวิธีประมาณค่าในช่วงแบบไม่เอกรูปนี้นับเป็นขั้นตอนพื้นฐานที่จำเป็นสำหรับกระบวนการขยายรายละเอียดภาพทั่วไป โดยจะพบว่าในขั้นตอนย่อยของระเบียบวิธีอื่นๆ มักจะประกอบไปด้วยขั้นตอนหลักทั้ง 3 นี้เสมอ

2.4.2 ระเบียบวิธีเชิงโดเมนความถี่ (Frequency Domain Method)

นำเสนอครั้งแรกโดย Tsai และ Huang ในปี 1984 [1] โดยการแปลงสัญญาณภาพเชิงโดเมนพื้นที่ (Spatial Domain) ไปสู่โดเมนความถี่ด้วยการแปลงฟูเรียร์ (Fourier Transform) ระเบียบวิธีนี้มีข้อดีตรงที่ใช้การคำนวณที่เรียบง่ายไม่ซับซ้อนโดยอาศัยหลักการ 3 ข้อได้แก่

- (1) คุณสมบัติการเลื่อนของการแปลงฟูเรียร์
- (2) ความสัมพันธ์เคลือบแฝง (Aliasing) ระหว่างผลการแปลงฟูเรียร์แบบต่อเนื่อง (Continuous Fourier Transform: CFT) ของภาพความละเอียดสูง กับผลการแปลงฟูเรียร์แบบเต็มหน่วย (Discrete Fourier Transform: DFT) ของภาพความละเอียดต่ำ
- (3) สมมุติฐานที่ว่าภาพความละเอียดสูงนั้นเป็นสัญญาณข้อมูลที่มีความถี่แบบจำกัดแถบ (Band-limited Signal)

อย่างไรก็ตาม ระเบียบวิธีนี้ใช้ได้กับการเคลื่อนที่แบบเลื่อนขนานทั้งหมด (Global Translational Motion) เท่านั้น ทำให้วิธีนี้ถูกจำกัดอยู่เพียงในกรณีที่วัตถุในภาพอยู่นิ่งกับที่ และกล้องบันทึกภาพเป็นฝ่ายเคลื่อนไหวเท่านั้น เช่น การถ่ายภาพระยะไกล หรือการถ่ายภาพจากดาวเทียม เป็นต้น

2.4.3 ระเบียบวิธีฉายกลับแบบทำซ้ำ (Iterative Back-Projection Method: IBP)

นำเสนอโดย Irani และ Peleg ในปี 1991 [1] ซึ่งเป็นการจำลองภาพต้นแบบความละเอียดสูงและป้อนเข้าสู่ระบบบันทึกภาพแบบผกผันที่มีตัวดำเนินการและตัวแปรต่างๆ ที่ประมาณค่าขึ้นมา ทำให้ได้ภาพความละเอียดต่ำและนำไปเปรียบเทียบกับภาพความละเอียดต่ำที่บันทึกได้เพื่อคำนวณค่าความคลาดเคลื่อนและนำไปปรับปรุงภาพต้นแบบความละเอียดสูงที่ประมาณขึ้นมาในตอนแรก จากนั้นจึงทำซ้ำเพื่อลดค่าความผิดพลาดลงให้น้อยที่สุดตามสมการ

$$x^{n+1} = x^n + \sum (y_k - H_k * x^n) * h^{BP} \quad (2.4)$$

โดยที่ h^{BP} เป็นตัวดำเนินการซึ่งสามารถกำหนดได้อย่างอิสระ โดยอาจกำหนดให้เป็นตัวดำเนินการที่กำหนดเงื่อนไขหรือคุณสมบัติบางอย่างที่ต้องการให้มีในภาพผลลัพธ์ก็ได้ ระเบียบวิธีนี้มีข้อดีในด้านความเรียบง่ายแต่ก็ใช้เวลานานและอาจให้ผลเฉลยได้หลายคำตอบ นอกจากนี้การทำซ้ำอาจจะไม่ลู่เข้า (Converge) เสมอไป

2.4.4 ระเบียบวิธีฉายลงบนเซตนูน (Projection onto Convex Sets Method: POCS)

นำเสนอครั้งแรกโดย Stark และ Oskoui ในปี 1989 [1] จากนั้นจึงถูกพัฒนาโดย Tekalp, Ozkan และ Sezan [3] เพื่อให้สามารถใช้กับแบบจำลองที่มีสัญญาณรบกวนด้วยได้ ระเบียบวิธีนี้อาศัยสมมุติฐานว่ามีเซตนูน (Convex Set: C_i) ของภาพที่มีลักษณะสมบัติต่างๆ ที่ภาพผลลัพธ์ควรจะมี โดยภาพผลลัพธ์ที่เป็นผลเฉลยนั้นจะอยู่ในบริเวณจุดตัด (Intersection) ของเซตเหล่านี้ ($C_s = \bigcap_{i=1}^m C_k$) ซึ่งจุดตัดนี้เป็นเซตนูนเช่นกัน การหาผลเฉลยสามารถทำได้โดยการฉายภาพจุดเริ่มต้น (Initializing Point) ไปยังเซตลักษณะสมบัติต่างๆ ผ่านตัวดำเนินการฉาย P_i (Projection Operator) ลงบนเซตลักษณะสมบัติ C_i ทีละเซต ตามความสัมพันธ์เวียนบังเกิด

$$x^{n+1} = P_m P_{m-1} \dots P_2 P_1 x^n \quad (2.5)$$

โดยเซตลักษณะสมบัติ C_i นั้นนิยามโดย

$$C_i = \{x \mid |r^{(x)}| < d_i\} \quad (2.6)$$

$$r^{(x)} = y_k - \sum x * H_k \quad (2.7)$$

เมื่อ $r^{(x)}$ คือเศษเหลือ (Residue) ซึ่งเป็นความคลาดเคลื่อนในการประมาณค่า x และ d_i คือค่าความมั่นใจเชิงสถิติหรือค่าความคลาดเคลื่อนที่ยอมรับได้ ซึ่งกำหนดจากลักษณะสมบัติของสัญญาณ

รบกวน ขนาดของ $r^{(x)}$ นั้นอาจถูกกำหนดด้วยการวัดระยะห่างในปริภูมิ L_2 หรือปริภูมิ ∞ เงื่อนไขพิเศษอาจถูกเพิ่มเติมเข้าไปในระเบียบวิธีนี้ได้เพื่อให้ผลเฉลยที่ได้มีความถูกต้องมากยิ่งขึ้น เช่น เงื่อนไขด้านขนาดสูงสุดและต่ำสุดของสัญญาณ (Amplitude Constraint) ตัวอย่างของงานวิจัยที่ใช้ระเบียบวิธีนี้ เช่น [4] ซึ่งเป็นกรณีพิเศษโดยการฉายค่าตอบไปยังเซตฐาน 2 เซตสลับไปมาจนกระทั่งคำตอบลู่ออกเข้า

ระเบียบวิธีนี้มีข้อดีในด้านการคำนวณที่ไม่ซับซ้อน และสามารถใช้ประโยชน์จากข้อมูลรู้ก่อน (Prior) มาช่วยในการประมาณค่าผลเฉลยได้ แต่มีข้อเสียในด้านความหลากหลายของผลเฉลยที่ได้ นอกจากนี้การคำนวณยังลู่เข้าช้าและใช้เวลาคำนวณมาก

2.4.5 ระเบียบวิธีเชิงกำหนดแบบคุมค่า (Regularized Deterministic Method)

จุดมุ่งหมายในการแก้ปัญหการสร้างกลับภาพความละเอียดสูง หรือการขยายรายละเอียดภาพ นั่นคือการหาผลเฉลย x ที่สอดคล้องกับภาพที่ได้จากการสังเกต y_k ทุกภาพและตัวดำเนินการ H_k และ h_k หรืออีกนัยหนึ่งคือมีความคลาดเคลื่อนน้อยที่สุด [1, 2]

โดยการสมมติให้สัญญาณรบกวนในภาพเป็นสัญญาณรบกวนสีขาวแบบเกาส์ที่มีค่าเฉลี่ยเป็นศูนย์ จะสามารถเขียนพจน์ความคลาดเคลื่อนเพื่อหาผลเฉลย x ได้ดังสมการ (2.7)

$$Error = \sum_{k=1}^N \|y_k - H_k x\|_p^p \quad (2.7)$$

โดยพจน์ทางขวามือคือความคลาดเคลื่อนซึ่งคำนวณจากผลรวมของระยะห่างในปริภูมิ L_p เมื่อ p คืออันดับของนอร์ม เนื่องจากกรณีจำนวนภาพความละเอียดต่ำไม่เพียงพอสำหรับจำนวนขั้นต่ำ หรือที่เรียกว่ากรณีกำหนดขาด (Underdetermined Case) ซึ่งเป็นกรณีที่พบเป็นส่วนใหญ่ ทำให้ปัญหาการสร้างกลับภาพความละเอียดสูงมีแนวโน้มเป็นปัญหาลักษณะเลว (Ill-posed Problem: ดูรายละเอียดเพิ่มเติมในบทที่ 3) และจำเป็นต้องใส่พจน์คุมค่า (Regularization Term) เพื่อช่วยประมาณค่าจุดภาพที่ขาดหาย และปรับปรุงแบบจำลองของปัญหาให้กลายเป็นปัญหาลักษณะดี (Well-posed) และผลเฉลยมีเสถียรภาพมากขึ้น ดังเช่นสมการ (2.8) ซึ่งแสดงตัวอย่างหนึ่งของพจน์ความคลาดเคลื่อนที่ใส่พจน์คุมค่าในปริภูมิ L_2

$$Error = \sum_{k=1}^N \|y_k - H_k x\|_2^2 + a \|Cx\|_2^2 \quad (2.8)$$

วิธีนี้เรียกว่าวิธีกำลังสองน้อยสุดเงื่อนไขบังคับ (Constrained Least Square Method: CLS) โดย $\|\cdot\|$ คือระยะห่างในปริภูมิหรือนอร์ม (Norm) อันดับ 2 พจน์แรกทางขวามือของสมการคือพจน์ความคลาดเคลื่อน ส่วนพจน์ที่สองทางขวามือเป็นพจน์คุมค่า โดย a คือตัวแปรเสริมการคุมค่า

(Regularization Parameter) ซึ่งทำหน้าที่ถ่วงดุลระหว่างรายละเอียด (Fidelity) กับความราบเรียบ (Smoothness) ของภาพผลลัพธ์ความละเอียดสูง C เป็นตัวดำเนินการเพื่อคงคุณสมบัติบางประการของภาพผลลัพธ์ x ไว้ ซึ่งโดยทั่วไปนิยมเลือกใช้ C เป็นตัวกรองผ่านสูง (High Pass Filter) เมื่อต้องการจำกัดความราบเรียบ (Smoothness) ของภาพผลลัพธ์ หรืออาจเป็นเพียงเมตริกซ์เอกลักษณ์ (Identity Matrix) เพื่อจำกัดพลังงานรวมของภาพผลลัพธ์ก็ได้

ผลเฉลยของปัญหา x ซึ่งเป็นภาพความละเอียดสูงจะหาได้จากค่าประมาณ \hat{x} ที่ทำให้พจน์ CLS มีค่าน้อยที่สุด กล่าวคือ

$$x = \arg \min_{\hat{x}} \left[\sum_{k=1}^N \|y_k - H_k \hat{x}\|_2^2 + a \|C\hat{x}\|_2^2 \right] \quad (2.9)$$

โดยระเบียบวิธีหาผลเฉลยแบบทำซ้ำ เช่น วิธีลาดลงชันที่สุด (Steepest Descent Method หรือ Gradient Method) ทำให้สามารถเขียนผลเฉลยในรูปแบบปิดได้เป็น

$$x_{n+1} = x_n + b \left[\sum_{k=1}^N H_k^T (y_k - H_k x_n) - a C^T C x_n \right] \quad (2.10)$$

อย่างไรก็ดี วิธีการนี้มักให้ผลเฉลยที่ลู่เข้าช้า (Slow Convergence) และจะช้ายิ่งขึ้นเมื่อใช้กับภาพที่มีขนาดใหญ่ เมื่อนำไปประยุกต์ใช้ในงานที่ต้องการความรวดเร็วในการประมวลผล เช่น ระบบเวลาจริง (Real Time System) จึงจำเป็นต้องพิจารณาวิธีการทำซ้ำที่ลู่เข้าได้รวดเร็วกว่านี้ เช่น ระเบียบวิธีเกรเดียนท์สังยุค (Conjugated Gradient Method)

2.4.6 ระเบียบวิธีเชิงสุ่มแบบคุมค่า (Regularized Stochastic Method)

ระเบียบวิธีเชิงสุ่มแบบคุมค่า หรืออีกชื่อหนึ่งว่าระเบียบวิธีแบบเบย์ (Bayesian Method) [1, 5] วิธีนี้จะให้ผลลัพธ์ที่ดีเมื่อทราบฟังก์ชันความหนาแน่นความน่าจะเป็นภายหลัง (Posterior Probability Density Function) ของภาพต้นแบบ โดยการประมาณค่า x ด้วยข้อมูลรู้ภายหลังมากที่สุด (Maximum a Posteriori: MAP) ซึ่งทำให้ความน่าจะเป็นภายหลังมีค่ามากที่สุด กล่าวคือเมื่อกำหนดชุดภาพความละเอียดต่ำ y_1, y_2, \dots, y_N มาให้ ผลเฉลย x ซึ่งเป็นภาพความละเอียดสูง จะมีค่าเท่ากับค่าประมาณ \hat{x} ที่ทำให้ความน่าจะเป็นแบบมีเงื่อนไขมีค่าสูงสุด

$$x = \arg \max_{\hat{x}} P(\hat{x} | y_1, y_2, \dots, y_N) \quad (2.11)$$

โดยการใส่ฟังก์ชันลอการิทึมและอาศัยกฎของเบย์ ทำให้สามารถจัดรูปสมการในรูปของความน่าจะเป็นแบบลอการิทึม (Log-likelihood) ได้เป็น

$$x = \arg \max_{\hat{x}} \{ \ln P(y_1, y_2, \dots, y_N | \hat{x}) + \ln P(\hat{x}) \} \quad (2.12)$$

ซึ่ง $P(y_1, y_2, \dots, y_N | \hat{x})$ สามารถหาได้จากผลคูณของความน่าจะเป็นย่อยของภาพความละเอียดต่ำแต่ละภาพดังสมการ

$$P(y_k | \hat{x}) = \prod_{\forall(x,y)} \frac{1}{s_k \sqrt{2p}} \exp\left(-\frac{(\hat{y}_k(x,y) - y_k(x,y))^2}{2s_k^2}\right) \quad (2.13)$$

เมื่อ $\hat{y}_k = H_k \hat{x}$ ซึ่งเมื่อรวมความน่าจะเป็นจากภาพความละเอียดต่ำทุกภาพแล้ว จะสามารถเขียนความน่าจะเป็นแบบลอการิทึม ได้เป็น

$$\sum_{\forall k} \ell(y_k) = -\sum_{\forall k} \|H_k \hat{x} - y_k\|^2 \quad (2.14)$$

โดยที่ $P(\hat{x})$ สามารถคำนวณได้จากข้อมูลรู้ก่อนแบบกิบส์ (Gibbs Prior) ซึ่งนิยามความหนาแน่นความน่าจะเป็นตามสมการ

$$P(X = x) = \frac{1}{Z} \exp\{-U(x)\} = \frac{1}{Z} \exp\left\{-\sum_{c \in S} j_c(x)\right\} \quad (2.15)$$

โดย Z คือตัวปรับบรรทัดฐาน (Normalizing Constant) $U(x)$ คือฟังก์ชันพลังงาน และ $j_c(x)$ เป็นฟังก์ชันศักย์ซึ่งขึ้นกับค่าของจุดภาพในอาณาบริเวณ c และ S เป็นเซตของอาณาบริเวณดังกล่าว

โดยการนิยามให้ $j_c(x) = (D^{(n)} x)^2$ เมื่อ $D^{(n)}$ เป็นผลต่างสี่บเนื่องอันดับที่ n (n^{th} Order Difference) ของภาพเพื่อรักษาลักษณะที่เป็นขอบ (Edge) ไว้ โดยนิยามให้ n มีค่าเท่ากับ 1 และ 2 นั้นคือ $j_c(x)$ เป็นอนุพันธ์อันดับหนึ่งและสองของภาพในแนวตั้ง แนวนอน และแนวทแยงมุมนั่นเอง โดยการสมมุติว่าสัญญาณรบกวนในภาพเป็นแบบแจกแจงเอกลักษณ์อิสระ (Independent Identical Distributed: i.i.d.) แบบเกาส์และมีค่าเฉลี่ยเป็นศูนย์ จะทำให้สามารถคำนวณผลเฉลย x ซึ่งเป็นภาพความละเอียดสูงได้จาก

$$x = \arg \min_{\hat{x}} \left[\sum_{k=1}^N \|y_k - H_k \hat{x}\|_p^p + a \sum_{c \in S} j_c(\hat{x}) \right] \quad (2.16)$$

ระเบียบวิธีเชิงสุมแบบคงค่านี้มีข้อดีในด้านความยืดหยุ่นและเสถียรภาพต่อสัญญาณรบกวนและความผิดพลาดจากการประมาณค่าตัวแปรเสริมต่างๆ นอกจากนี้ยังใช้ประโยชน์จากข้อมูลรู้ก่อน (Prior) มาช่วยปรับปรุงผลเฉลย ทำให้มั่นใจว่าผลเฉลยที่คำนวณได้นั้นมีอยู่หนึ่งเดียว (Unique)

จากงานวิจัยเกี่ยวกับการขยายรายละเอียดภาพในแต่ละระเบียบวิธีที่ได้กล่าวมาข้างต้น สามารถสรุปเป็นตารางได้ดังนี้

ตารางที่ 1 : สรุปงานวิจัยด้านการขยายรายละเอียดภาพที่มีผู้วิจัยมาก่อนหน้านี้

Researchers	H. Ur & D. Gross [1]	R.Y. Tsai & T.S.Huang [1]	M. Irani & S. Peleg [1]	H. Stark & P. Oskoui [1]	B.C. Tom & Katsaggelos [1]
Method	Non-uniform Interpolation	Frequency Domain	Iterative Back Projection	Projection onto Convex Sets	Deterministic & Stochastic Regularized
Object Motion	Global & Local Translational	Global Translational Only	Global & Local Translational	Global & Local Translational	Global & Local Translational
Computational Cost	Low	Low	Low	High	High
Characteristic of Solution	Optimality not guaranteed	-	Slow Convergence	Slow Convergence	Moderate Convergence
Inclusion of Priors	No	No	No	Yes	Yes
Number of Solution	Multiple Solutions	Unique Solution	Multiple Solutions	Multiple Solutions	Unique Solution

2.5 สรุปท้ายบท

การขยายรายละเอียดภาพอาศัยแนวคิดการนำข้อมูลภาพของวัตถุเดียวกันจากหลายๆ มุมมองมารวมกันในภาพภาพเดียวที่มีความละเอียดเชิงพื้นที่สูงขึ้น โดยอาศัยการซ้กซึมตัวอย่างช้าลงที่ระยะเลื่อนไม่เต็มจุดภาพ และจำนวนภาพความละเอียดต่ำที่เพียงพอทำให้การสร้างกลับภาพความละเอียดสูงมีความเป็นไปได้ โดยการพิจารณาชุดภาพต่อเนื่องที่ได้จากแฟ้มภาพวิดีโอที่ค้นเสมือนเป็นชุดของภาพความละเอียดต่ำของวัตถุเดียวกันที่ถูกบันทึกที่เวลาและมุมกล้องต่างๆ กันทำให้สามารถนำหลักการขยายรายละเอียดภาพมาประยุกต์ใช้กับแฟ้มภาพวิดีโอได้ โดยการพิจารณาการขยายรายละเอียดภาพเป็นกระบวนการย้อนกลับของการลดทอนคุณภาพผ่านตัวแปรต่างๆ เช่นการบิดโค้ง ความพรั่มัว

และการชักสุมตัวอย่างซาหลง ทำให้สามารถตั้งแบบจำลองของปัญหา และศึกษาค้นคว้าเพื่อหาระเบียบวิธีที่เหมาะสมในการแก้ปัญหาได้ เช่น ระเบียบวิธีประมาณค่าในช่วงแบบไม่เอกรูป ระเบียบวิธีฉายภาพลงบนเซตนูน ระเบียบวิธีเชิงกำหนดแบบคุ่มค่า เป็นต้น ซึ่งแต่ละระเบียบวิธีก็มีข้อดีข้อเสียแตกต่างกันไปในเรื่องระยะเวลาที่ใช้ในการประมวลผล อัตราการลู่เข้าของคำตอบ เสถียรภาพและลักษณะสมบัติของคำตอบที่ได้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

ปัญหาการสร้างกลับภาพความละเอียดสูง

จากแบบจำลองของปัญหาที่ได้กล่าวไว้ในบทที่ 2 จะเห็นว่าการขยายรายละเอียดภาพสามารถพิจารณาเป็นปัญหาการสร้างกลับภาพความละเอียดสูงได้ ในบทนี้จะกล่าวถึงรายละเอียดเพิ่มเติมของแบบจำลองของปัญหา ปัจจัยที่มีผลต่อการบันทึกภาพ ตัวแปรหลัก ตัวแปรเสริมและตัวดำเนินการในแบบจำลอง รวมถึงเงื่อนไขที่ต้องพิจารณาและลักษณะของเพิ่มภาพวีดิทัศน์ที่ใช้ในงานวิจัยนี้

3.1 ปัจจัยลดทอนคุณภาพในแบบจำลองของปัญหา

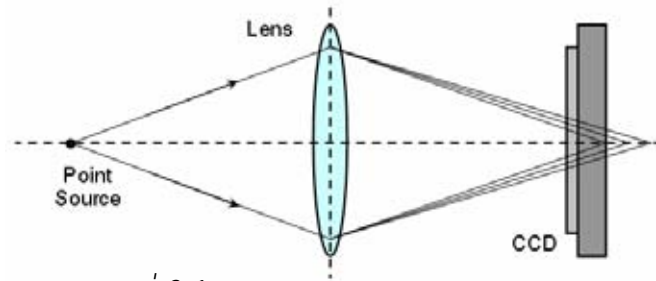
ในการตั้งแบบจำลองของปัญหาเพื่อหาคำตอบของสมการ จำเป็นต้องมีการศึกษาลักษณะสมบัติของปัจจัยต่างๆ ที่มีผลกระทบต่อตัวแปร เพื่อที่จะสามารถประมาณลักษณะสมบัติของตัวดำเนินการต่างๆ ในแบบจำลองซึ่งโดยทั่วไปมักไม่เป็นที่ทราบค่าแน่ชัดได้อย่างเหมาะสม ปัจจัยหลักที่มีผลต่อคุณภาพของภาพในกระบวนการบันทึกภาพได้แก่

3.1.1 ความพร่ามัว (Blur)

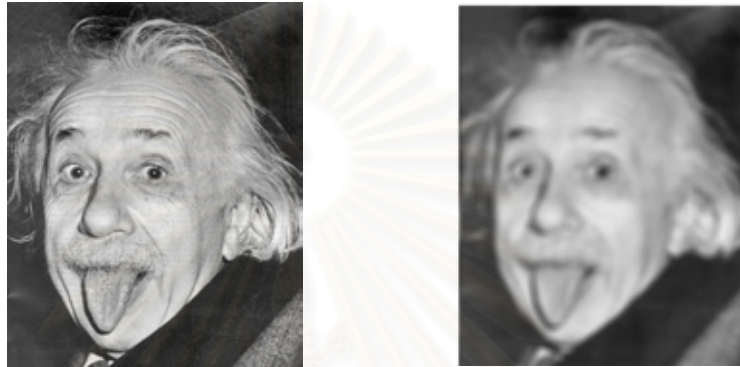
ความพร่ามัวเป็นปรากฏการณ์ของแสงซึ่งเมื่อสะท้อนจากวัตถุผ่านตัวกลางและระบบทัศนอุปกรณ์ในระบบบันทึกภาพไปตกยังตัวตรวจรู้ภาพโดยมีลักษณะไม่เป็นจุดสมบูรณ์ กล่าวคือแสงมีการกระจายตัวไปยังบริเวณรอบๆ โดยฟังก์ชันที่อธิบายลักษณะความพร่ามัวนี้เรียกว่าฟังก์ชันแผ่กระจายจุด (Point Spread Function: PSF) โดยปกติแล้วเป็นการยากที่จะทราบคุณสมบัติของฟังก์ชันแผ่กระจายจุดของระบบบันทึกภาพได้ ในทางปฏิบัติจึงใช้การประมาณฟังก์ชันในแบบจำลองให้ใกล้เคียงความจริงที่สุด ความพร่ามัวที่มักพบในกระบวนการบันทึกภาพนั้นแบ่งได้เป็น 3 ประเภทใหญ่ๆ ได้แก่

1) ความพร่ามัวเชิงแสง (Optical Blur)

เกิดจากการที่ระบบเลนส์ในกล้องจัดวางตัวในระยະที่ไม่เหมาะสม หรือเกิดจากเลนส์ที่ใช้มีคุณภาพต่ำทั้งในด้านวัสดุที่ใช้ทำเลนส์และความโค้งของผิวเลนส์กล่าวคือเกิดความคลาดทรงกลม (Spherical Aberration) ทำให้แสงจากวัตถุตามรูปที่ 3-1 (โดยพิจารณาว่าทุกๆ จุดบนผิววัตถุเป็นแหล่งกำเนิดแสงแบบจุด) เมื่อเดินทางผ่านระบบเลนส์แล้วไม่ถูกโฟกัสเป็นจุดเมื่อตกกระทบฉากรับภาพ ทำให้ภาพที่ได้มีลักษณะพร่ามัวไม่คมชัดดังรูปที่ 3-2 แม้จะเลือกใช้ตัวตรวจรู้ภาพที่มีความละเอียดเชิงพื้นที่สูง นอกจากนี้ยังเกิดจากการที่แสงซึ่งมีความยาวคลื่นต่างๆ กันเดินทางผ่านเลนส์จึงทำให้เกิดการหักเหของแสงในแต่ช่วงความถี่แตกต่างกันไป



รูปที่ 3-1 การเกิดความพรำมัวเชิงแสง



รูปที่ 3-2 ผลของความพรำมัวเชิงแสง

2) ความพรำมัวการเคลื่อนที่ (Motion Blur)

ในช่วงเวลาขณะทำการบันทึกภาพ วัตถุเป้าหมายหรือตัวกล้องบันทึกภาพเองอาจมีการเคลื่อนไหว และเนื่องจากช่วงเวลาเปิดหน้ากล้องมีค่าไม่เป็นศูนย์ (Non-zero Aperture Time) ทำให้แสงจากวัตถุซึ่งควรจะถูกโฟกัสเป็นจุดนิ่งบนตัวตรวจรู้ภาพถูกลากเป็นแนวตามทิศทางการเคลื่อนที่ตลอดช่วงเวลาที่เปิดหน้ากล้อง [6] ทำให้ภาพที่ได้ไม่คมชัด ดังแสดงในรูปที่ 3-3 ซึ่งแสดงผลของความพรำมัวจากการเคลื่อนที่ในแนวเฉียงขึ้นไปทางขวาเล็กน้อย โดยหากเวลาเปิดหน้ากล้องนานขึ้น จะยิ่งสังเกตเห็นผลได้ชัดเจนมากขึ้น ตัวอย่างของความพรำมัวชนิดนี้เช่นการถ่ายภาพในที่ที่มีแสงน้อย หรือการถ่ายภาพดวงดาวโดยการเปิดหน้ากล้องเป็นเวลานาน



รูปที่ 3-3 ผลของความพรำมัวการเคลื่อนที่

3) ความพร่ามัวตัวตรวจรู้ (Sensor Blur)

ความพร่ามัวชนิดนี้เกิดจากตัวตรวจรู้ภาพที่มีความละเอียดเชิงพื้นที่จำกัด บันทึกแสงแบบเฉลี่ยจากวัตถุหรือฉากเป้าหมายซึ่งมีความละเอียดเชิงพื้นที่แบบต่อเนื่องหรือไม่จำกัดลงสู่หน่วยรับแสงบนตัวตรวจรู้ภาพ [1, 7] ทำให้ไม่สามารถเก็บรายละเอียดจากวัตถุได้ทั้งหมด ความพร่ามัวชนิดนี้จึงเกิดกับกล้องถ่ายภาพดิจิทัลทุกประเภท โดยตัวตรวจรู้ภาพที่มีความละเอียดสูงกว่าจะสามารถเก็บรายละเอียดของวัตถุได้มากกว่า ดังแสดงในรูปที่ 3-4

ในทางปฏิบัติ ความพร่ามัวจากตัวตรวจรู้ภาพนี้อาจถูกพิจารณาเป็นการชักสุมตัวอย่างขาลงโดยตัวตรวจรู้ภาพ โดยข้อมูลที่บันทึกได้จะมีปริมาณน้อยกว่าข้อมูลจริงจากวัตถุเป้าหมาย



ภาพจากตัวตรวจรู้ภาพความละเอียดสูง

ภาพจากตัวตรวจรู้ภาพความละเอียดต่ำ

รูปที่ 3-4 ภาพที่บันทึกจากตัวตรวจรู้ภาพที่มีความละเอียดแตกต่างกัน

3.1.2 สัญญาณรบกวน (Noise)

สัญญาณรบกวนโดยทั่วไปแบ่งออกเป็น 2 ประเภทหลักๆ ได้แก่ สัญญาณรบกวนแบบบวก (Additive Noise) และสัญญาณรบกวนแบบคูณ (Multiplicative Noise) ในงานวิจัยนี้จะกล่าวถึงเฉพาะสัญญาณรบกวนแบบบวก ตัวอย่างแบบจำลองของสัญญาณรบกวนที่นิยมใช้ได้แก่

1) สัญญาณรบกวนแบบเกาส์ (Gaussian Noise)

แบบจำลองของสัญญาณรบกวนแบบเกาส์มีรูปแบบตามสมการ (3.1)

$$h(x) = \frac{1}{s\sqrt{p}} e^{-\frac{x^2}{2s^2}} \quad (3.1)$$

2) สัญญาณรบกวนแบบลาปลาซ (Laplacian Noise)

แบบจำลองของสัญญาณแบบลาปลาซมีรูปแบบตามสมการ (3.2)

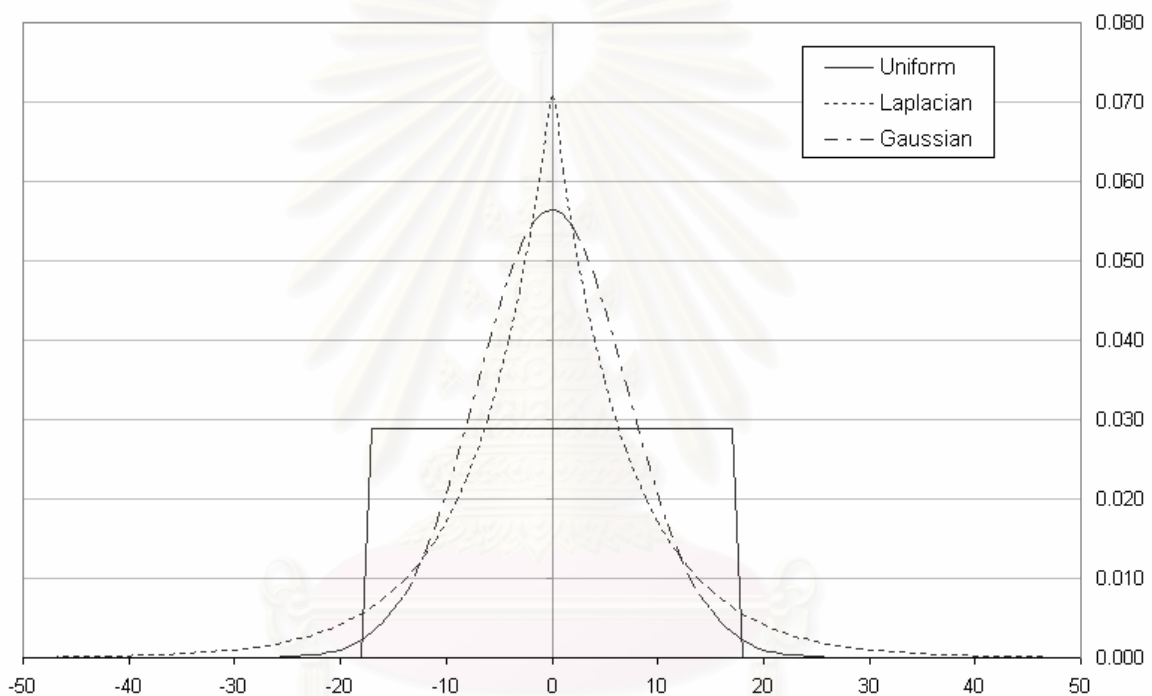
$$h(x) = \frac{1}{s\sqrt{2}} e^{-\frac{\sqrt{2}|x|}{s}} \quad (3.2)$$

3) สัญญาณรบกวนแบบเอกรูป (Uniform Noise)

แบบจำลองของสัญญาณแบบเอกรูปมีรูปแบบตามสมการ (3.3)

$$h(x) = \begin{cases} \frac{1}{s2\sqrt{3}} & \text{for } |x| \leq s\sqrt{3} \\ 0 & \text{else} \end{cases} \quad (3.3)$$

สัญญาณรบกวนข้างต้นเป็นแบบจำลองของสัญญาณ 1 มิติ ซึ่งมีกราฟแสดงตัวอย่างในรูปที่ 3-5 สำหรับกรณี $s = 10$



รูปที่ 3-5 กราฟสัญญาณรบกวน 1 มิติทั้ง 3 ประเภท

เมื่อนำไปใช้กับงานประมวลผลภาพซึ่งเป็นสัญญาณ 2 มิติ โดยการพิจารณาสัญญาณรบกวนเป็นแบบแบบแจกแจงเอกลักษณะอิสระ จะได้ว่า

$$h(x, y) = h(x) \cdot h(y) \quad (3.4)$$

3.2 ตัวแปรและตัวดำเนินการต่างๆ ในแบบจำลองของปัญหา

จากสมการแบบจำลองของปัญหา ในบทที่ 2

$$y_k = H_k * x + h_k \quad k = 1, 2, 3, \dots, L \quad (3.5)$$

เมื่อนำมาเขียนใหม่โดยพิจารณาว่า H_k เป็นผลรวมของตัวดำเนินการชนิดต่างๆ จะได้ว่า

$$y_k = D_k * B_k * W_k * x + h_k \quad k = 1, 2, 3, \dots, L \quad (3.6)$$

โดย W_k คือตัวดำเนินการบิดโค้ง (Warping Operator) B_k เป็นตัวดำเนินการพรางมัว (Blur Operator) และ D_k คือตัวดำเนินการชักสุมตัวอย่างขาลง (Down-sampling Operator) และ h_k เป็นสัญญาณรบกวน โดยที่ x เป็นภาพความละเอียดสูงและ y_k เป็นภาพความละเอียดต่ำลำดับที่ k จากจำนวนทั้งหมด L ภาพ

ในงานวิจัยที่ผ่านมาส่วนใหญ่ [1, 2, 4, 7] นิยมนำ y_k และ x ซึ่งเดิมเป็นตัวแปรแบบแถว-สดมภ์ มาจัดรูปใหม่แบบศัพท์พจนานุกรม (Lexicographic Order) เป็นเวกเตอร์แนวตั้งขนาด MN และเปลี่ยนตัวดำเนินการย่อยแต่ละตัวเป็นเมตริกซ์ขนาด $MN \times MN$ ซึ่งทำให้ตัวดำเนินการคูณประสาน (*) ถูกเปลี่ยนเป็นตัวดำเนินการคูณแบบธรรมดา (\cdot) ทำให้สามารถเขียนสมการ (3.6) ในรูปของเวกเตอร์-เมตริกซ์ได้เป็น

$$y_k = D_k B_k W_k x + h_k \quad k = 1, 2, 3, \dots, L \quad (3.7)$$

อย่างไรก็ตาม การจัดรูปตัวแปรในรูปเวกเตอร์-เมตริกซ์นี้ทำให้ตัวดำเนินการ W_k , B_k และ D_k ซึ่งเป็นเมตริกซ์มีขนาดใหญ่มาก เป็นผลให้ใช้เวลาในการคำนวณมากแม้ว่าเมตริกซ์เหล่านี้จะประกอบไปด้วยสมาชิกที่มีค่าเป็นศูนย์อยู่จำนวนมาก (Sparse Matrix) ก็ตาม เพื่อหลีกเลี่ยงปัญหาดังกล่าว งานวิจัยนี้จึงคงรูปแบบตัวแปรเป็นแบบแถว-สดมภ์ตามเดิมโดยพิจารณาว่าภาพความละเอียดสูง x ถูกลดทอนคุณภาพผ่านกระบวนการบิดโค้ง การทำให้ภาพพรางมัว การชักสุมตัวอย่างขาลง และการเจ็บบนด้วยสัญญาณรบกวน โดยใช้สัญลักษณ์เป็นตัวดำเนินการ W_k , B_k , D_k และ h_k ตามลำดับ และเพื่อความสะดวกในการทำความเข้าใจกับตัวดำเนินการเหล่านี้ จึงขอละสัญลักษณ์ตัวดำเนินการคูณประสาน * ระหว่างตัวแปรที่เป็นรูปภาพกับตัวดำเนินการต่างๆ ไว้ในฐานที่เข้าใจว่าเป็นการดำเนินการผ่านตัวลดทอนคุณภาพดังกล่าว รายละเอียดของตัวแปรและตัวดำเนินการต่างๆ ที่ใช้ในงานวิจัยมีดังต่อไปนี้

3.2.1 ภาพความละเอียดต่ำ

y_k หรือภาพที่ได้จากการสังเกตลำดับที่ k (k^{th} Measurement) ถูกพิจารณาเป็นเมตริกซ์ขนาด $M \times N$ จุดภาพ ซึ่งสมาชิกแต่ละตัวแสดงค่าระดับความเข้มเป็นจำนวนเต็มอยู่ในช่วง 0 ถึง 255 โดยค่า 0 แทนค่าระดับความเข้มต่ำสุด และค่า 255 แทนค่าระดับความเข้มสูงสุดสำหรับแต่ละช่องสัญญาณของภาพสี ซึ่งเป็นค่าจุดภาพที่ใช้ในมาตรฐานภาพสีแบบ 8 บิต

3.2.2 ภาพความละเอียดสูง

x ในทางทฤษฎีคือฉากเหตุการณ์จริงซึ่งเป็นภาพที่มีความละเอียดเชิงพื้นที่สูงเป็นอนันต์ ซึ่งหากใช้สมการแบบจำลองใน (3.5) จะพบว่าไม่สามารถสร้างกลับภาพที่มีปริมาณข้อมูลมากเป็นอนันต์ได้ ในทางปฏิบัติจึงสมมติให้ x เป็นภาพที่มีความละเอียดสูงที่มีข้อมูลเป็นจำนวนจำกัด โดยเป็นตัวอย่างแบบแถว-สดมภ์ซึ่งมีคุณสมบัติเช่นเดียวกันกับ y_k แต่มีขนาด $r_1 M \times r_2 N$ จุดภาพ เมื่อ r_1 และ r_2 คือตัวประกอบการขยายในแนวตั้งและแนวนอนตามลำดับ โดยทั่วไปนิยมกำหนดให้ $r_1 = r_2 = r$ เพื่อความสะดวกในการวิเคราะห์ปัญหาและรักษาอัตราส่วนลักษณะ (Aspect Ratio) ของภาพ

เมื่อพิจารณาจำนวนข้อมูล จะเห็นว่า y_k แต่ละภาพมีจำนวนจุดภาพเท่ากับ MN จุดภาพ และ x มีจำนวนจุดภาพเท่ากับ $r^2 MN$ จุดภาพ ในทางทฤษฎีหากสมมติให้จุดภาพจากภาพความละเอียดต่ำทั้งหมดเป็นข้อมูลชนิดไม่ซ้ำกัน (Non-redundant Information) จะได้ว่า จำนวนขั้นต่ำของภาพความละเอียดต่ำที่ต้องใช้ในกระบวนการสร้างกลับภาพความละเอียดสูงเท่ากับ r^2 ภาพ สำหรับงานวิจัยนี้ซึ่งใช้ตัวประกอบการขยายเท่ากับ 4 เป็นหลักในการทำขยายรายละเอียดเพิ่มภาพวิดีโอที่ต้นนั้นจึงต้องใช้จำนวนภาพความละเอียดต่ำขั้นต่ำถึง 16 ภาพ แต่ในทางปฏิบัติมักพบว่าไม่สามารถรวบรวมชุดภาพหนึ่งต่อหนึ่งของวัตถุหนึ่งๆ เช่น ภาพคนเป็นจำนวนมากเช่นนี้ได้ เนื่องจากลักษณะของวัตถุในชุดภาพหนึ่งต่อหนึ่งจะมีลักษณะเปลี่ยนไปมากเมื่อเปรียบเทียบระหว่างภาพแรกกับภาพสุดท้าย ทำให้ไม่สามารถทำขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งได้อย่างแม่นยำ

ในงานวิจัยนี้จึงเลือกจำนวนภาพความละเอียดต่ำอยู่ในช่วง 4-8 ภาพ สำหรับใช้กับตัวประกอบการขยายในช่วง 2-4 เท่า โดยเน้นที่ค่า 4 เท่าเป็นหลัก เนื่องจากสามารถสังเกตเห็นรายละเอียดที่เพิ่มขึ้นได้มากกว่ากรณี 2 เท่าอย่างเห็นได้ชัดเมื่อเปรียบเทียบกับ การขยายรายละเอียดภาพด้วยวิธีประมาณค่าในช่วงแบบไบคิวบิก

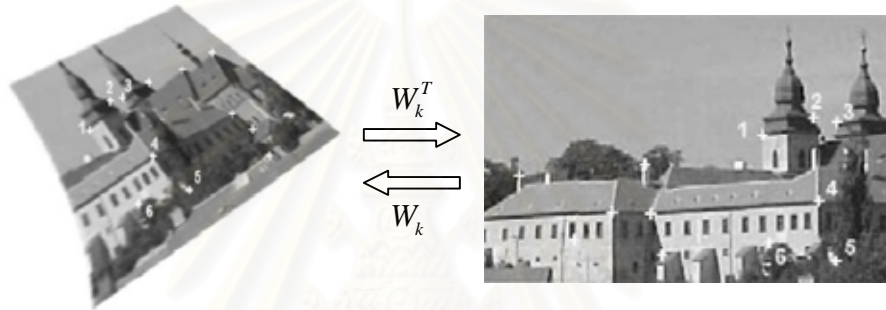
3.2.3 ตัวดำเนินการบิดโค้ง

w_k เป็นตัวดำเนินการหลักในขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่ง (Image Registration) ซึ่งทำหน้าที่แปลงค่าพิกัดของจุดภาพในภาพหนึ่งไปสู่ค่าพิกัดใหม่ในอีกภาพหนึ่ง ในงานวิจัยนี้เลือกใช้การซ้อนทับภาพโดยกำหนดตำแหน่งแบบอิงลักษณะเด่นในภาพ (Feature-based Registration) [8] กล่าวคือเป็นการซ้อนทับภาพโดยอาศัยจุดควบคุม (Control Points: CPs) ซึ่งเป็นตำแหน่งของลักษณะเด่นเช่นบริเวณเส้นขอบ จุดตัดของเส้นขอบ โดยจุดควบคุมที่ได้นั้นจะต้องปรากฏอยู่ในทุกภาพที่จะทำการซ้อนทับ และไม่แปรเปลี่ยนเมื่อภาพถูกบิดโค้ง ดังรูปที่ 3-6 ซึ่งแสดงจุดควบคุม (ตำแหน่งกากบาทสีขาว)



รูปที่ 3-6 จุดควบคุมในแต่ละภาพ

โดยการจับคู่จุดควบคุม (Corresponding CPs Matching) ระหว่างภาพ 2 ภาพที่ถูกต้อง ทำให้สามารถคำนวณเมตริกซ์แปลงภาพ (Transformation Matrix) สำหรับแปลงระบบพิกัดภาพดังแสดงในรูปที่ 3-7 ได้



รูปที่ 3-7 การจับคู่จุดควบคุมระหว่างภาพ 2 ภาพ

การแปลงพิกัดภาพนั้นจัดเป็นการแปลงแบบเรขาคณิต (Geometric Transform) ในงานวิจัยนี้จะพิจารณาเฉพาะการแปลงเชิงเส้น (Linear Transformation) ซึ่งโดยทั่วไปแบ่งออกเป็น 3 ประเภทใหญ่ๆ ได้แก่

- การแปลงคล้าย (Similarity Transform) ใช้กับภาพที่มีการเปลี่ยนแปลงตำแหน่งแบบเลื่อนที่ (Translation) แบบหมุน (Rotation) ซึ่งแกนหมุนตั้งฉากกับระนาบของภาพ และแบบเปลี่ยนมาตราส่วน (Scaling) โดยรักษาค่าอัตราส่วนลักษณะของวัตถุในภาพไว้ และใช้จุดควบคุมเพียง 2 จุด โดยมีสมการดังนี้

$$u = s[x \cdot \cos(f) - y \cdot \sin(f)] + t_x \quad (3.8 A)$$

$$v = s[x \cdot \sin(f) + y \cdot \cos(f)] + t_y \quad (3.8 B)$$

- การแปลงสัมพรรค (Affine Transform) เป็นการแปลงซึ่งมีสมบัติเหมือนกับการแปลงคล้าย แต่สามารถเปลี่ยนมาตราส่วนได้ทั้งแกนตั้งและแกนนอนแยกอิสระ โดยใช้จุดควบคุมจำนวน 3 จุดตามสมการ

$$u = a_{11} \cdot x + a_{12} \cdot y + a_{13} \quad (3.9 A)$$

$$v = a_{21} \cdot x + a_{22} \cdot y + a_{23} \quad (3.9 B)$$

- การแปลงทัศนมิติ (Perspective Transform) เป็นการแปลงที่มีการคำนวณระยะใกล้ไกลที่ไม่เท่ากันทุกจุดในภาพเข้ามาเกี่ยวข้องด้วย กล่าวคือครอบคลุมการบิดภาพรอบแกนที่อยู่บนระนาบเดียวกับภาพ (Tilt) โดยใช้จุดควบคุมจำนวน 4 จุดตามสมการ 3-10

$$u = \frac{a_{11} \cdot x + a_{12} \cdot y + a_{13}}{a_{31} \cdot x + a_{32} \cdot y + a_{33}} \quad (3.10 A)$$

$$v = \frac{a_{21} \cdot x + a_{22} \cdot y + a_{23}}{a_{31} \cdot x + a_{32} \cdot y + a_{33}} \quad (3.10 B)$$

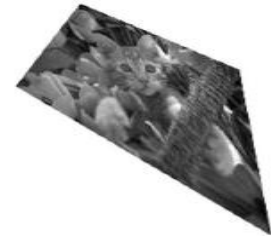
ตัวอย่างการแปลงภาพเชิงเส้นประเภทต่างๆ แสดงไว้ในรูปที่ 3-8



(ก) การแปลงคล้าย



(ข) การแปลงสัมพรรค



(ค) การแปลงทัศนมิติ

รูปที่ 3-8 การแปลงเรขาคณิตแบบต่างๆ

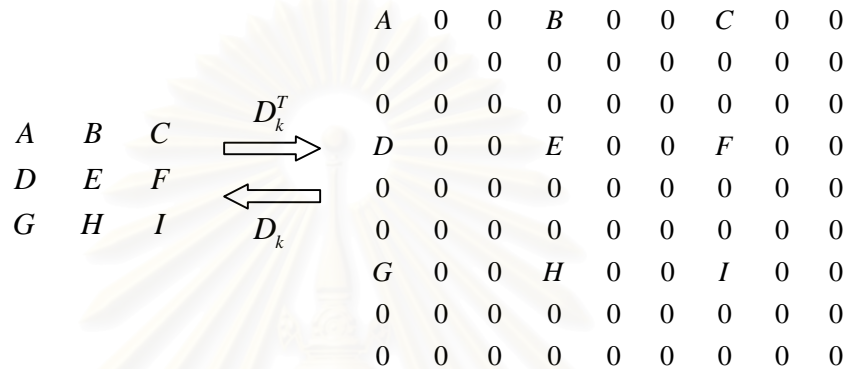
ในงานวิจัยนี้ได้เลือกใช้การแปลงทัศนมิติ ซึ่งครอบคลุมกรณีที่มีการเคลื่อนที่เป็นแบบพ้องรูปเชิงระนาบ (Planar Homography Transformation) [5] ซึ่งจะได้กล่าวถึงอย่างละเอียดในบทที่ 4 โดยการแปลงชนิดนี้จะใช้ได้กับแฟ้มภาพเคลื่อนไหวที่มีอัตราเฟรม (Frame Rate) สูง และกรณีที่วัตถุในภาพมีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่ (Local Translational Motion) นอกจากนี้การกระจัดของจุดภาพ (Pixel Displacement) อันเนื่องมาจากตัวดำเนินการบิดโค้งนี้จะถูกปิดเศษให้เป็นจำนวนเต็มจุดภาพที่ใกล้เคียงค่าเดิมมากที่สุดในรอบอ้างอิงของภาพความละเอียดสูง โดยไม่กระทบต่อคุณภาพของภาพผลลัพธ์เนื่องจากกระทำในกริดความละเอียดสูง

3.2.4 ตัวดำเนินการพรมัว

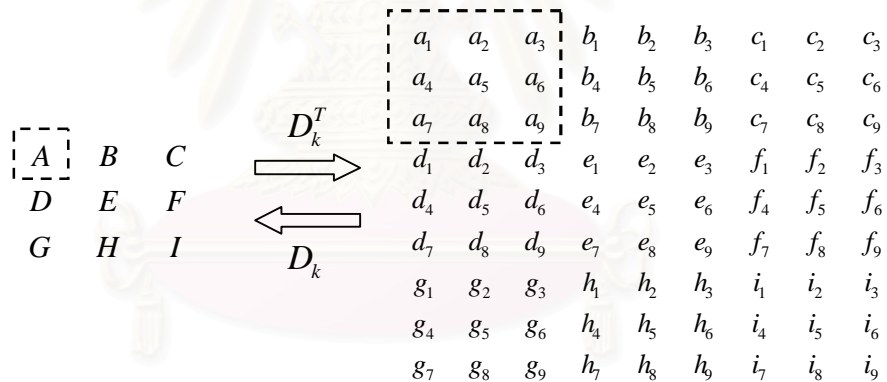
B_k ทำหน้าที่อธิบายลักษณะสมบัติของฟังก์ชันกระจายจุด ซึ่งในงานวิจัยนี้ได้เลือกพิจารณาเฉพาะความพรมัวเชิงแสง โดยประมาณค่าฟังก์ชันกระจายจุดด้วยเคอร์เนลความพรมัวแบบเกาส์ (Gaussian Blur Kernel) ที่มีขนาดอยู่ในช่วง 3×3 ถึง 7×7 และมีค่าเบี่ยงเบนมาตรฐาน (Standard Deviation: s) อยู่ในช่วง 1 ถึง 5 โดยจะอธิบายถึงวิธีการประมาณค่าที่เหมาะสมของขนาดเคอร์เนลและค่า s ในบทที่ 5

3.2.5 ตัวดำเนินการชักสุมตัวอย่างขาลง

ทำหน้าที่อธิบายพฤติกรรมกรรมกรบนที่ข้อมูลของตัวตรวจจู้ภาพ ซึ่งแบ่งออกเป็น 2 แบบหลักๆ ได้แก่ตัวดำเนินการสุมแบบเลือกค่า (Pixel-selection) [2] ซึ่งเป็นการชักสุมแบบเลือกค่าจุดภาพเฉพาะจุดที่ระยะและตำแหน่งคงที่ และแบบเฉลี่ยค่า (Spatial Average) [1, 7] ซึ่งเก็บข้อมูลจุดภาพโดยการเฉลี่ยค่าจุดภาพในบริเวณหนึ่งๆ ดังแสดงในรูปที่ 3-9 และ 3-10 ซึ่งเป็นตัวอย่างของการชักสุมตัวอย่างโดยมีตัวประกอบการขยายเท่ากับ 3



รูปที่ 3-9 ตัวดำเนินการชักสุมตัวอย่างแบบเลือกค่า



รูปที่ 3-10 ตัวดำเนินการชักสุมตัวอย่างแบบเฉลี่ยค่า

โดยในรูปที่ 3-10 แต่ละจุดภาพของภาพความละเอียดต่ำทางด้านซ้ายมือ นั้นได้มาจากการเฉลี่ยค่าจุดภาพของภาพความละเอียดสูงทางด้านขวามือขนาด 3x3 ดังสมการ (3.11)

$$X = \frac{\sum_{i=1}^9 x_i}{9} \quad X = A, B, C, \dots \quad x = a, b, c, \dots \quad (3.11)$$

ในงานวิจัยนี้ได้เลือกตัวดำเนินการชักสุมตัวอย่างกระจายจุดทั้งแบบขาขึ้นและขาลงดังในรูปที่ 3-9 เพื่อให้สอดคล้องกับแบบจำลองของปัญหา ซึ่งหากเลือกตัวดำเนินการแบบเฉลี่ยค่าจะเกิดปัญหาในขั้นตอนการชักสุมขาขึ้นเนื่องจากไม่สามารถประมาณค่าข้อมูลของจุดภาพที่เพิ่มขึ้นได้

3.2.6 สัญญาณรบกวน

สัญญาณรบกวนในแบบจำลองของงานวิจัยนี้ได้สมมุติสัญญาณรบกวนสีขาวแบบแฉกแฉง เอกลักษณะอิสระ 2 มิติของเกาส์ ซึ่งมีแบบจำลองดังสมการ (3.1) และ (3.4)

3.3 ปัญหาผกผันและปัญหาลักษณะเลว (Inverse Problem & Ill-posed Problem)

กระบวนการสร้างกลับภาพความละเอียดสูงเป็นการประมาณค่าตัวแปรเสริม (ตัวดำเนินการลดทอนคุณภาพต่าง ๆ) และสัญญาณเข้า (ภาพความละเอียดสูง) ของแบบจำลองจากสัญญาณออกของระบบหรือผลการวัด (ภาพความละเอียดต่ำ) ที่มีอยู่ ปัญหาลักษณะนี้จึงจัดเป็นปัญหาผกผัน ซึ่งโดยทั่วไปมักจะมีลักษณะเลว (Ill-posed) เนื่องมาจากการขาดเงื่อนไขที่จำเป็น 3 ประการที่จะทำให้เป็นปัญหาลักษณะดี (Well-posed) [9] ได้แก่

- 1) การมีอยู่ของผลเฉลย (Existence of Solution)
- 2) ผลเฉลยมีอยู่หนึ่งเดียว (Uniqueness of Solution)
- 3) ความต่อเนื่องของผลเฉลยซึ่งอิงข้อมูล (Continuity of Data-dependent Solution)

จากแบบจำลองของปัญหาการสร้างกลับภาพความละเอียดสูง ในกรณีกำหนดขาด (Underdetermined) [1, 2] ซึ่งจำนวนจุดภาพสุทธิจากภาพความละเอียดต่ำทั้งหมดมีจำนวนน้อยกว่าจำนวนจุดภาพในภาพความละเอียดสูงที่ต้องการ เมื่อพยายามแก้สมการของปัญหาโดยตรงจะพบว่ามีผลเฉลย คือภาพความละเอียดสูง เป็นจำนวนอนันต์ที่สอดคล้องกับสมการของปัญหาอันเนื่องมาจากมีจุดภาพที่ขาดหาย (Missing Pixels) ในภาพผลลัพธ์ความละเอียดสูง ทำให้เป็นปัญหาลักษณะเลว เนื่องจากขาดเงื่อนไขข้อ 2 นอกจากนี้ตัวดำเนินการพรัามัวใน H_x ในแบบจำลองของปัญหาซึ่งเป็นตัวดำเนินการที่มีสภาวะเลว (Ill-conditioned) ยังทำให้ผลเฉลยไม่มีเสถียรภาพสำหรับกรณีกำหนดเกิน (Overdetermined) อีกด้วย กล่าวคือความคลาดเคลื่อนเพียงเล็กน้อยระหว่างขั้นตอนการคำนวณเช่น การปิดเศษหรือสัญญาณรบกวน จะทำให้ผลเฉลยสุดท้ายเปลี่ยนแปลงไปอย่างมาก ทำให้เข้าข่ายปัญหา ลักษณะเลวเนื่องจากขาดเงื่อนไขข้อ 3

โดยทั่วไปมักพบว่าจำนวนภาพความละเอียดต่ำในปัญหาสร้างกลับภาพความละเอียดสูงมักมีจำนวนน้อยกว่า r^2 ทำให้เป็นปัญหาสภาวะเลวแบบกำหนดขาด เพื่อปรับปรุงแบบจำลองของปัญหาให้เป็นลักษณะดี จึงได้มีการใส่พจน์คุมค่า (Regularization Term) เพื่อควบคุมให้ผลเฉลยมีคุณสมบัติที่ต้องการ เช่น ข้อจำกัดด้านแอมพลิจูดหรือพลังงานรวมของผลเฉลย หรือความราบเรียบของผลเฉลย ดังจะได้กล่าวรายละเอียดในบทที่ 4

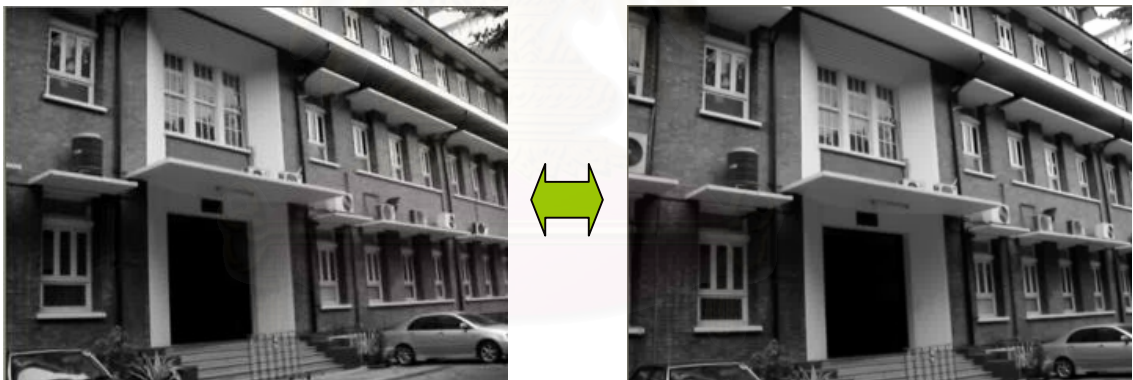
3.4 เงื่อนไขและลักษณะของแฟ้มภาพวิดีโอที่ใช้นในงานวิจัย

เพื่อให้ระเบียบวิธีที่เลือกใช้นในงานวิจัยนี้เกิดประสิทธิภาพสูงสุด จึงได้มีการกำหนดเงื่อนไขและลักษณะของแฟ้มภาพวิดีโอที่เข้ากับโปรแกรมที่นำเสนอไว้ดังนี้

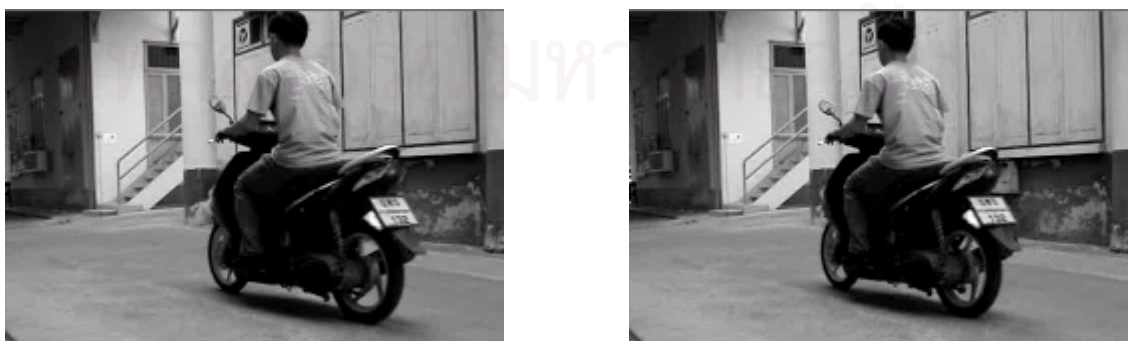
3.4.1 การบิดโค้งเชิงเส้น (Linear Warping)

เพื่อให้โปรแกรมขยายรายละเอียดภาพในงานวิจัยนี้สามารถใช้งานการแปลงภาพแบบเชิงเส้นกับวัตถุที่สนใจได้ จึงได้มีการกำหนดลักษณะการเคลื่อนที่และลักษณะทางเรขาคณิตที่เหมาะสมของวัตถุที่สนใจ โดยมีองค์ประกอบดังต่อไปนี้

- การเคลื่อนที่ของวัตถุเป็นการเคลื่อนที่แบบเลื่อนขนานทั้งหมด (Global Translational Motion) กล่าวคือวัตถุในฉากทั้งหมดมีการเคลื่อนที่แบบเลื่อนขนานเหมือนกัน เช่นการบันทึกภาพระยะไกล การหมุนกวาดกล้องไปรอบๆ ดังตัวอย่างในรูปที่ 3-11 หรือเป็นการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่ (Local Translational Motion) กล่าวคือแต่ละวัตถุในฉากมีการเคลื่อนที่แบบเลื่อนขนานแตกต่างกัน เช่น ภาพวิดีโอที่คนซึ่งกล้องบันทึกติดตั้งอยู่กับที่ โดยมีวัตถุเคลื่อนไหวและฉากหลังอยู่นิ่งดังตัวอย่างในรูปที่ 3-12



รูปที่ 3-11 แฟ้มภาพวิดีโอที่ที่มีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด



รูปที่ 3-12 แฟ้มภาพวิดีโอที่ที่มีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่

- วัตถุที่สนใจมีความคงรูป (Rigidity) ซึ่งในความหมายของการประมวลผลภาพ จะหมายถึง ภาพของวัตถุที่มีระยะระหว่างจุดสองจุดใดๆ บนผิววัตถุเป็นระยะคงที่ แต่ในกรณีภาพ 2 มิติที่ถูกฉายจากวัตถุ 3 มิติที่มีการหมุน จะทำให้ระยะดังกล่าวไม่คงที่โดยขึ้นกับแบบจำลองเรขาคณิตและรูปแบบการเคลื่อนที่ของวัตถุ เช่น ภาพใบหน้าคนๆ เดียวกันแต่มีสีหน้าหรืออารมณ์ต่างกันดังในรูปที่ 4-3 การที่วัตถุที่สนใจในภาพขาดความคงรูปทำให้ไม่สามารถใช้การแปลงเชิงเส้นกับภาพเหล่านี้ได้



รูปที่ 3-13 ภาพใบหน้าคนซึ่งเป็นวัตถุแบบไม่คงรูป

- วัตถุที่สนใจไม่มีปัญหาการบดบัง (Occlusion) ได้แก่การถูกวัตถุอื่นบดบัง เช่น วัตถุในฉากเคลื่อนมาซ้อนกัน เช่น คนเดินสวนกัน หรือหายไปใต้มุมอับ เช่น คนเดินเข้าไปหลังกำแพง และการบดบังตัวเองของวัตถุที่สนใจซึ่งเกิดจากการหมุนรอบตัวเอง เช่น การหันหน้าไปมา ทำให้ใบหน้าซีดจางหรือขวาทหายไป ปัญหาการบดบังนี้นับเป็นปัญหาที่แก้ไขยากที่สุด เนื่องจากไม่สามารถประมาณค่ารายละเอียดที่ถูกบดบังได้



รูปที่ 3-14 ภาพใบหน้าคนซึ่งเกิดการบดบังตัวเอง

3.4.2 ความเร็วเคลื่อนที่ของวัตถุต่ำ

เมื่อความเร็วของวัตถุเคลื่อนที่ในภาพมีค่าต่ำเมื่อเทียบกับอัตราเฟรมของการบันทึกภาพ ซึ่งโดยทั่วไปมีค่าเท่ากับ 25 เฟรมต่อวินาที จะสามารถประมาณได้ว่าระยะเวลาเปิดหน้ากล้องสำหรับบันทึกมีค่าเป็นศูนย์ (Zero Aperture Time) ทำให้สามารถละเลยความพัวมัวที่เกิดจากการเคลื่อนที่ของกล้องหรือวัตถุได้

3.4.3 ความละเอียดเชิงพื้นที่สูงเพียงพอ

สำหรับโปรแกรมในการหาจุดอ้างอิงเพื่อซ้อนทับภาพโดยกำหนดตำแหน่ง โดยจากผลการทดลองในงานวิจัยนี้พบว่าช่วงขนาดของวัตถุที่สนใจในภาพที่สามารถนำมาผ่านการขยายรายละเอียดด้วยโปรแกรมที่นำเสนอได้อยู่ที่ประมาณ 40x30 ถึง 80x60 จุดภาพ

3.5 สรุปท้ายบท

การขยายรายละเอียดภาพจัดเป็นปัญหาย้อนกลับของการลดทอนคุณภาพของภาพซึ่งจำเป็นต้องมีการประมาณค่าตัวดำเนินการและตัวแปรเสริมต่างๆ ในกระบวนการให้ใกล้เคียงกับความเป็นจริงเพื่อให้ได้คำตอบคือสัญญาณเข้าที่ถูกต้อง ซึ่งโดยส่วนใหญ่พบว่าภาพความละเอียดต่ำที่มีอยู่มีจำนวนต่ำกว่าจำนวนขั้นต่ำทำให้เป็นปัญหาแบบกำหนดขาด และทำให้เข้าข่ายปัญหาลักษณะเลว จึงจำเป็นต้องใส่พจน์คุมค่าลงในสมการต้นทุนเพื่อปรับปัญหาให้มีลักษณะดี นอกจากนี้เพื่อให้การประมวลผลเป็นไปอย่างมีประสิทธิภาพ จำเป็นต้องพิจารณาถึงลักษณะการบิดโค้งหรือการเคลื่อนที่ของวัตถุในภาพเพื่อให้สามารถแปลงภาพด้วยการแปลงเชิงเส้นได้ รวมถึงความเร็วเคลื่อนที่ของวัตถุต้องมีค่าน้อย และความละเอียดเชิงพื้นที่ของบริเวณที่สนใจต้องมีค่าสูงเพียงพอ เพื่อให้สามารถซ้อนทับภาพได้อย่างถูกต้องแม่นยำ

บทที่ 4

ระเบียบวิธีที่เลือกใช้ในการออกแบบโปรแกรม

ในบทนี้จะอธิบายถึงระเบียบวิธีที่ใช้ในขั้นตอนต่างๆ ของการขยายรายละเอียดภาพ ได้แก่ การซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ การรวมภาพ และการแก้สมการแบบจำลองของปัญหา รวมถึงการประยุกต์ระเบียบวิธีดังกล่าวเพื่อให้โปรแกรมสามารถประมวลผลภาพสีได้

4.1 การซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ

ขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งเป็นขั้นตอนที่ทำให้ข้อมูลจุดภาพจากภาพความละเอียดต่ำในแต่ละภาพมาปรากฏอยู่บนกรอบภาพความละเอียดสูงในตำแหน่งที่ถูกต้อง [1] ซึ่งนับเป็นขั้นตอนที่มีความสำคัญต่อการขยายรายละเอียดภาพอย่างมาก เนื่องจากหากซ้อนทับภาพผิดตำแหน่งแล้ว จะทำให้คุณภาพของภาพความละเอียดสูงกลับแย่ลงดังจะแสดงตัวอย่างในบทที่ 7 เพื่อให้ขั้นตอนนี้เป็นไปอย่างมีประสิทธิภาพจึงต้องเลือกแบบจำลองการแปลงภาพที่เหมาะสมกับการเคลื่อนที่ของวัตถุที่สนใจภายในภาพ โดยการซ้อนทับภาพต้องมีความแม่นยำในระดับต่ำกว่าจุดภาพของภาพความละเอียดต่ำ หรืออยู่ในระดับเดียวกับจุดภาพของภาพความละเอียดสูง

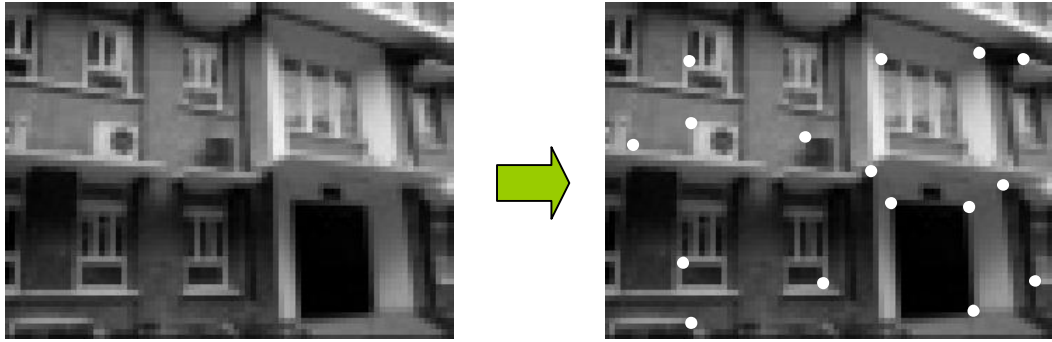
การออกแบบโปรแกรมให้สามารถซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติได้ ทำให้โปรแกรมมีความยืดหยุ่น สามารถนำไปใช้ขยายรายละเอียดภาพของวัตถุใดๆ ในแฟ้มภาพวีดิทัศน์ได้ โดยไม่จำเป็นต้องทราบรูปแบบการเคลื่อนที่ของวัตถุนั้นๆ มาก่อน อย่างไรก็ตาม ประสิทธิภาพของการซ้อนทับภาพยังคงมีข้อจำกัดเช่น แบบจำลองเรขาคณิตและการเคลื่อนที่ของวัตถุที่สนใจภายในภาพ ในหัวข้อนี้จะอธิบายถึงขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติสำหรับภาพวีดิทัศน์ที่วัตถุในภาพมีการเคลื่อนที่แบบเลื่อนขนาน โดยแบ่งเป็นการเลื่อนขนาดทั้งหมด และการเลื่อนขนาดเฉพาะที่

4.1.1 กรณีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด

แบ่งออกเป็น 3 ขั้นตอนหลักได้แก่

- การตรวจจับลักษณะเด่นในภาพ (Feature Detection)

ใช้ฟังก์ชันตรวจจับลักษณะเด่นในไลบรารีโอเพ่นซีวี ตรวจสอบหาลักษณะเด่นภายในภาพ ได้แก่ มุม และขอบจำนวน 16 จุดที่มีความชัดมากที่สุดดังแสดงในรูปที่ 4-1 (วัดจากค่าลักษณะเฉพาะหรือ Eigenvalue) สำหรับใช้เป็นจุดควบคุมโดยมีความแม่นยำในระดับต่ำกว่าจุดภาพ และมีระยะห่างระหว่างจุดไม่ต่ำกว่าค่าๆ หนึ่งเพื่อให้จุดควบคุมไม่กระจุกตัวอยู่ที่บริเวณเดียว ทำให้การคำนวณมีความถูกต้องมากยิ่งขึ้น



รูปที่ 4-1 การค้นหาตำแหน่งของลักษณะเด่นในภาพ

- การจับคู่ลักษณะเด่นที่สอดคล้องกัน (Corresponding Control Point Matching)

โดยการให้ฟังก์ชันการไหลเชิงแสงเพื่อติดตามการเคลื่อนที่ของลักษณะหรือจุดควบคุมที่ตรวจพบในขั้นตอนที่แล้ว เพื่อหาตำแหน่งของจุดควบคุมที่สอดคล้องกันในเฟรมข้างเคียง โดยคำนวณระหว่างภาพอ้างอิงกับภาพข้างเคียงคู่หนึ่งๆ จนครบทุกจุด ดังแสดงในรูปที่ 4-2



เฟรมอ้างอิง

เฟรมข้างเคียง

รูปที่ 4-2 การจับคู่ลักษณะเด่นที่สอดคล้อง

- การคำนวณเมตริกซ์แปลงภาพ (Image Transformation Matrix Calculation)

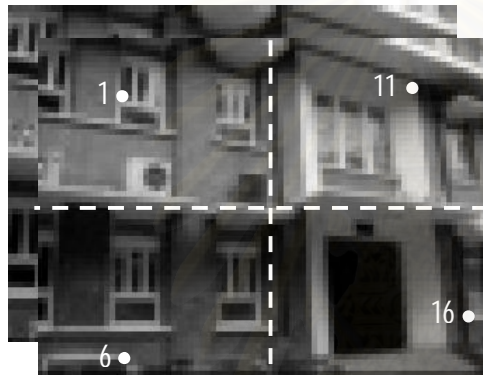
โดยทั่วไปในการคำนวณเมตริกซ์แปลงภาพแบบทศนิยมมิติจะใช้จุดควบคุมจำนวน 4 คู่ แต่เพื่อให้เมตริกซ์แปลงภาพมีความถูกต้องแม่นยำมากที่สุด เมื่อได้จุดควบคุมและคู่ทั้งหมดที่สอดคล้องกันทั้ง 16 คู่แล้วจึงนำมาคัดกรองด้วยระเบียบวิธีแรนแซคเพื่อหาคู่จุดควบคุมที่เหมาะสม และกำจัดคู่จุดควบคุมที่ไม่เข้าพวกทิ้งไป จากนั้นจึงนำไปคำนวณเมตริกซ์แปลงภาพต่อไปโดยใช้วิธีการที่เหมาะสมที่สุด (Optimization Method)

4.1.2 ระเบียบวิธีคัดกรองแรนแซค (RANSAC) [5]

ระเบียบวิธีแรนแซคเป็นระเบียบวิธีคัดกรองแบบทำซ้ำเพื่อคัดรวมวัตถุที่สนใจไว้ด้วยกัน ซึ่งในงานวิจัยนี้ได้นำมาใช้เพื่อคัดเลือกจุดควบคุมและคู่ของจุดควบคุมที่สอดคล้องที่เหมาะสมซึ่งทำให้โปรแกรมสามารถคำนวณเมตริกซ์แปลงภาพได้โดยมีความผิดพลาดน้อยที่สุด

ขั้นตอนของระเบียบวิธีเรขาคณิตสำหรับการเคลื่อนไหวแบบพ้องรูปเชิงระนาบ มีดังนี้

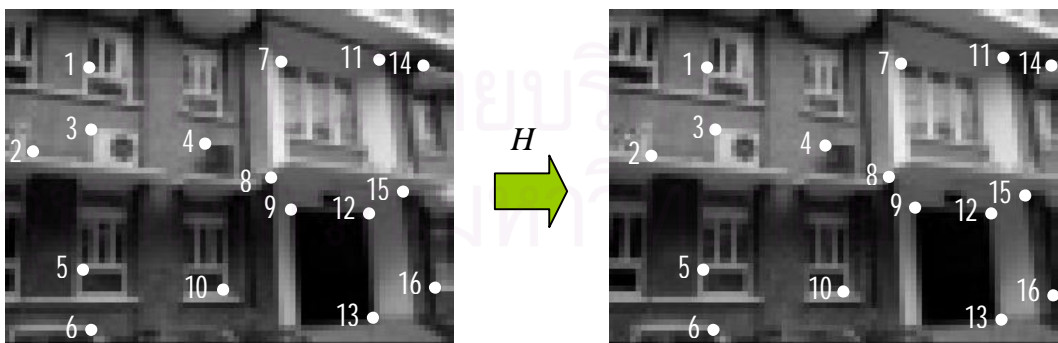
- 1) เลือกจุดควบคุมและคู่ที่สอดคล้องมาจำนวน 4 คู่ โดยการยึดหลักว่า ในการแปลงเชิงเส้นนั้น จุดควบคุมที่ดีควรจะกระจายครอบคลุมบริเวณที่จะทำการแปลงภาพ ซึ่งในที่นี้ได้แก่ภาพทั้งภาพ เพื่อให้พื้นที่ที่ถูกล้อมรอบด้วยจุดควบคุมมีความแม่นยำในการแปลง และหลีกเลี่ยงการที่ชุดของจุดควบคุมกระจุกตัวอยู่บริเวณใดบริเวณหนึ่งของภาพ ซึ่งทำให้บริเวณภาพนอกจุดควบคุมเกิดความคลาดเคลื่อนในการแปลงได้มาก จึงใช้วิธีการแบ่งภาพออกเป็น 4 ส่วน ซึ่งในแต่ละส่วนมีจำนวนจุดควบคุมเท่ากัน และสุ่มเลือกจุดควบคุมและคู่ที่สอดคล้อง 1 คู่จากแต่ละส่วน โดยในตัวอย่างนี้ได้เลือกจุดควบคุมที่ 1 11 6 และ 16 ดังแสดงตัวอย่างในรูปที่ 4-3 และนำมาคำนวณเมตริกซ์แปลงภาพทัศนมิติ H ตามสมการ (3.10)



$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

รูปที่ 4-3 การคำนวณเมตริกซ์แปลงภาพจากจุดควบคุมจำนวน 4 คู่

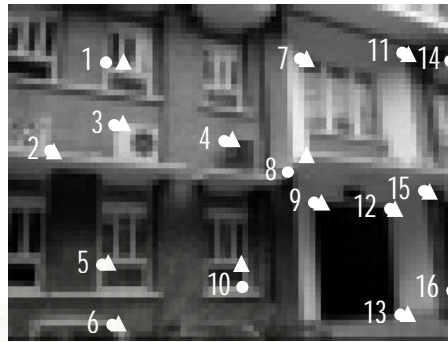
- 2) นำเมตริกซ์ H ที่ได้มาทำนายตำแหน่งของจุดควบคุมของภาพอ้างอิงที่เปลี่ยนไปทุกจุดในเฟรมข้างเคียงดังในรูปที่ 4-4 จากนั้นเปรียบเทียบตำแหน่งที่คำนวณได้กับตำแหน่งของจุดควบคุมของจริงในเฟรมข้างเคียง บันทึกค่าระยะคลาดเคลื่อนสำหรับทุกๆ จุดควบคุมไว้



รูปที่ 4-4 การหาตำแหน่งทำนายของจุดควบคุมในเฟรมอ้างอิงด้วยเมตริกซ์แปลงภาพ

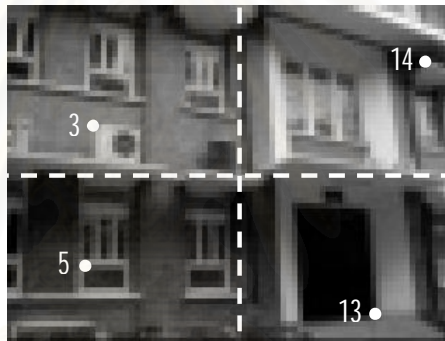
- 3) คำนวณหาจำนวนจุดในกลุ่ม (Inlier) โดยพิจารณาค่าผิดพลาดระหว่างตำแหน่งทำนายกับตำแหน่งจริงของจุดควบคุมในขั้นตอนที่แล้ว โดยหากมีค่าน้อยกว่าค่าขีดเริ่มเปลี่ยนที่กำหนดไว้จะถือว่าเป็นจุดในกลุ่ม แต่หากค่าผิดพลาดมีค่ามากกว่าค่าขีดเปลี่ยนจะถือว่าเป็นจุดนอกกลุ่ม

เป็นจุดนอกกลุ่ม (Outlier) บันทึกจำนวนจุดในกลุ่มที่นับได้สำหรับชุดของจุดควบคุมตั้งต้น
หนึ่งๆ ที่เลือกมาในขั้นตอน 1) โดยในรูปที่ 4-5 นี้จุดนอกกลุ่มได้แก่จุดควบคุมที่ 18 และ 10



รูปที่ 4-5 ความคลาดเคลื่อนระหว่างตำแหน่งทำนายกับตำแหน่งจริงของจุดควบคุมในเฟรมข้างเคียง

- 4) เลือกชุดของจุดควบคุม 4 จุดชุดใหม่ที่แตกต่างจากขั้นตอน 1) แล้วทำซ้ำจนถึงขั้นตอน 3)
ทำซ้ำเช่นนี้จนครบทุกชุดจุดควบคุมที่เป็นไปได้ทั้งหมด



รูปที่ 4-6 ชุดของจุดควบคุมชุดใหม่สำหรับการคำนวณในรอบถัดไป

- 5) เลือกชุดจุดควบคุมตั้งต้นที่มีจำนวนจุดในกลุ่มมากที่สุด และจุดควบคุมที่เป็นจุดในกลุ่ม
ทั้งหมด มาคำนวณหาเมตริกซ์แปลงภาพทัศนมิติอีกครั้งด้วยระเบียบวิธีกำลังสองน้อยสุด
โดยการแก้ระบบสมการในรูปแบบเมตริกซ์ดังสมการ (4.1) [12]

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -y_1u_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2u_2 & -y_2u_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3u_3 & -y_3u_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -x_nu_n & -y_nu_n \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1v_1 & -y_1v_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_2 & -y_2v_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3v_3 & -y_3v_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & x_n & y_n & 1 & -x_nv_n & -y_nv_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \\ v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix} \quad (4.1)$$

เมื่อ n คือจำนวนของจุดในกลุ่มทั้งหมดซึ่งมีค่าตั้งแต่ 4 ถึง 16 จุด (x, y) คือตำแหน่งพิกัดของจุดภาพในเฟรมอ้างอิง (u, v) คือตำแหน่งพิกัดของจุดภาพในเฟรมข้างเคียงใดๆ h_{ij} คือสมาชิกในแถวที่ i และสดมภ์ที่ j ของเมตริกซ์แปลงภาพ H โดยที่ $h_{33}=1$ เมื่อได้เมตริกซ์ H ซึ่งเป็นเมตริกซ์แปลงภาพไปข้างหน้า (Forward Transformation Matrix) แล้ว จึงนำไปหาเมตริกซ์ผกผันเพื่อหาเมตริกซ์แปลงภาพผกผัน (Inverse Transformation Matrix) สำหรับแปลงภาพจากตำแหน่งพิกัดของเฟรมข้างเคียงกลับไปสู่ตำแหน่งพิกัดของเฟรมอ้างอิง

จากขั้นตอนข้างต้น ทำให้ได้เมตริกซ์แปลงภาพซึ่งมีความผิดพลาดน้อยที่สุดระหว่างภาพอ้างอิงกับภาพข้างเคียงคู่หนึ่งๆ ระเบียบวิธีแรมแซคนี้ยังสามารถนำไปใช้คัดแยกการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่ ซึ่งมักพบในกรณีของระบบเฝ้าระวังโดยมีวัตถุในฉากกำลังเคลื่อนที่ในขณะที่ฉากหลังอยู่นิ่ง โดยสามารถคัดแยกการเคลื่อนที่ของวัตถุและฉากออกจากกันได้

4.1.3 กรณีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่

ในกรณีที่วัตถุที่สนใจมีการเคลื่อนที่แตกต่างจากวัตถุอื่นๆ ในภาพหรือฉากหลัง เช่นกรณีที่กำลังบันทึกภาพถูกติดตั้งอยู่กับที่ในระบบเฝ้าระวัง ทำให้ฉากหลังปรากฏอยู่นิ่ง หรือมีวัตถุเคลื่อนที่หลายชิ้นรวมถึงวัตถุที่สนใจ เช่นบุคคล หรือยานพาหนะเคลื่อนที่ในฉากท่ามกลางวัตถุอื่นๆ ในกรณีนี้ จำเป็นต้องใช้การตรวจจับวัตถุเคลื่อนที่เพื่อแยกการเคลื่อนที่ของวัตถุที่สนใจออกจากวัตถุอื่นๆ และฉากหลัง และซ้อนทับภาพโดยกำหนดตำแหน่งเฉพาะวัตถุที่สนใจได้อย่างถูกต้อง

ในงานวิจัยนี้ได้มุ่งเน้นไปที่กรณีที่มีการเคลื่อนที่หลักเพียง 2 กลุ่มในฉาก ได้แก่ กรณีที่ฉากหลังอยู่นิ่งและมีวัตถุเคลื่อนที่ที่วัตถุที่สนใจเคลื่อนที่เพียงวัตถุเดียว เนื่องจากโดยทั่วไปในการใช้งานโปรแกรมผู้ใช้จะกำหนดบริเวณที่จะขยายรายละเอียดภาพเอง ทำให้บริเวณดังกล่าวมีเพียงวัตถุที่สนใจและฉากหลังเท่านั้น จากนั้นจึงตรวจจับวัตถุที่สนใจโดยการใช้ระเบียบวิธีต่อไปนี้

- ระเบียบวิธีลบฉากหลัง (Background Subtraction Method)

วิธีนี้อาศัยหลักการพื้นฐานคือนำภาพฉากหลังเปล่าซึ่งไม่มีวัตถุเคลื่อนที่ในภาพมาลบออกจากภาพฉากหลังเดียวกันที่มีวัตถุเคลื่อนที่ปรากฏอยู่ ทำให้ได้ภาพที่มีแต่วัตถุเคลื่อนที่ ซึ่งสามารถนำไปใช้เป็นแผ่นบัง (Mask) เพื่อตรวจจับจุดควบคุมเคลื่อนที่ตามขั้นตอนซ้อนทับภาพ โดยกำหนดตำแหน่งตามปกติได้ วิธีนี้มีข้อดีในแง่ความเรียบง่าย และสามารถสร้างแผ่นบังที่มีรูปร่างตรงกับวัตถุอย่างแม่นยำได้ แต่จะทำได้ดีกับกรณีที่กล้องถูกติดตั้งอยู่กับที่และอยู่ในมุมมองโดยต้องมีการบันทึกฉากหลังเปล่าเก็บไว้ทุกครั้งที่มีการเปลี่ยนตำแหน่งกล้อง หากเป็นกรณีที่กล้องบันทึกถูกติดตั้งไว้กลางแจ้ง จะเกิดปัญหาความสว่างของแสงและเงาจากดวงอาทิตย์ที่เปลี่ยนไป ทำให้ผลการตรวจจับวัตถุผิดพลาดได้

- ระเบียบวิธีจำแนกวัตถุเคลื่อนที่ (Moving Object Classification Method)

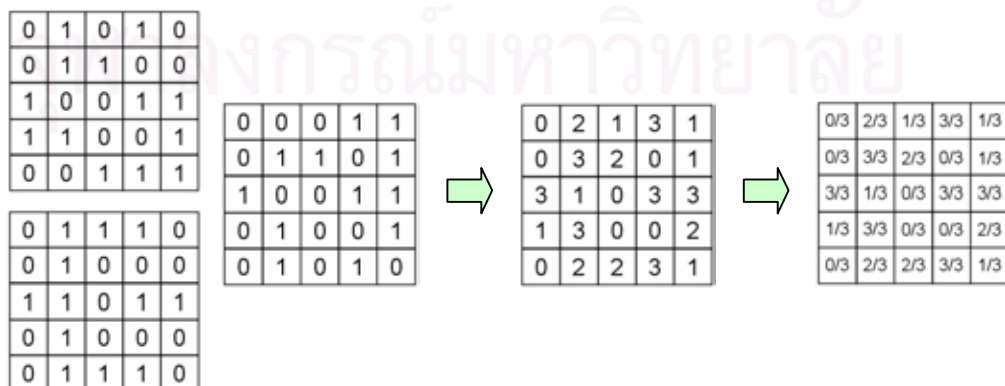
วิธีนี้อาศัยการตรวจจับจุดวัตถุเคลื่อนที่ทั้งหมดภายในภาพแล้วนำมาจำแนกเป็นวัตถุแต่ละชิ้นที่มีการเคลื่อนที่แตกต่างกัน โดยใช้การจำแนกเวกเตอร์เคลื่อนที่ของจุดควบคุมทั้งหมดในภาพ เพื่อคัดเวกเตอร์ที่มีลักษณะการเคลื่อนที่เหมือนกันให้อยู่ในกลุ่มเดียวกัน วิธีการนี้มีการคำนวณที่ซับซ้อนมากกว่า และเกิดข้อผิดพลาดได้ง่ายกว่าวิธีแรก แต่หากออกแบบการคำนวณได้ถูกต้องจะสามารถนำไปใช้งานได้ครอบคลุมมากกว่าวิธีแรกเนื่องจากไม่ขึ้นกับสภาพแสงเงาในฉาก และไม่จำเป็นต้องบันทึกฉากหลังเปล่าเก็บไว้ล่วงหน้า นอกจากนี้ยังสามารถนำวิธีนี้ไปใช้กับกล้องที่มีการเคลื่อนไหว เช่นการหมุนกวาดได้อีกด้วย

4.2 การรวมภาพ

หลังจากทราบตำแหน่งของจุดภาพทั้งหมดบนเฟรมอ้างอิงแล้ว จึงนำข้อมูลจุดภาพเหล่านี้มาบรรจุในภาพเดียวกัน ซึ่งขั้นตอนดังกล่าวเรียกว่าการรวมภาพ ในงานวิจัยนี้ได้ทดสอบระเบียบวิธีรวมภาพเฉลี่ย (Mean Method) ซึ่งเป็นการใช้จากข้อมูลจุดภาพที่มีอยู่ทุกจุดในการประมาณค่าจุดภาพในภาพความละเอียดสูง โดยแบ่งเป็น 2 แนวทางได้แก่

4.2.1 การรวมภาพเฉลี่ยแบบไม่ถ่วงน้ำหนัก

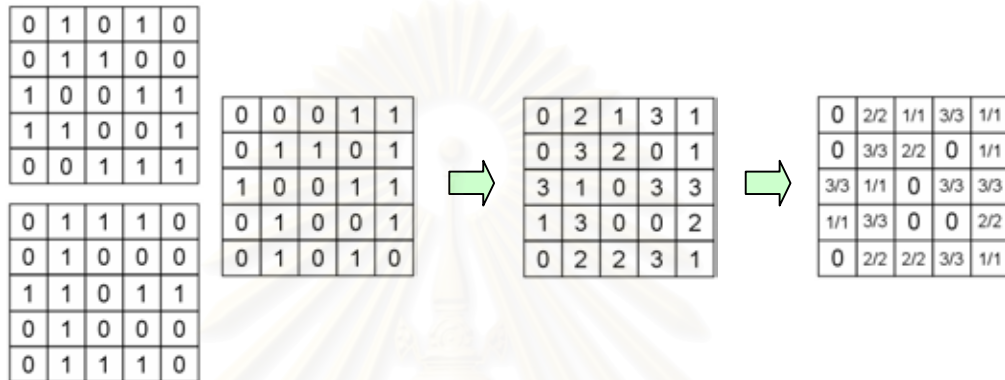
ด้วยวิธีนี้ จุดภาพบนเฟรมความละเอียดสูงจุดหนึ่งๆ จะได้รับข้อมูลจุดภาพจากภาพความละเอียดต่ำทุกภาพโดยมีความสำคัญเท่าๆ กันแม้ว่าจะมีจุดภาพจากภาพความละเอียดต่ำบางจุดที่ไม่มีข้อมูล ดังแสดงในรูปที่ 4-7 ซึ่งเป็นตัวอย่างการรวมภาพความละเอียดต่ำจำนวน 3 ภาพเป็นภาพความละเอียดสูง 1 ภาพ โดยตัวเลขในแต่ละจุดภาพแสดงจำนวนข้อมูลที่ได้จากภาพความละเอียดต่ำ จะเห็นได้ว่าทุกตำแหน่งในภาพจะถูกเฉลี่ยด้วยตัวคูณ $\frac{1}{3}$ เท่าๆ กันทั้งหมด ทำให้ตำแหน่งที่มีข้อมูลจุดภาพจากภาพความละเอียดต่ำเพียงภาพเดียวหรือ 2 ภาพถูกเฉลี่ยร่วมกับจุดภาพที่ไม่มีข้อมูล ทำให้ความเปรียบต่าง ณ จุดนั้นลดลง ภาพที่ได้จึงมีลักษณะค่อนข้างมืด



รูปที่ 4-7 การรวมภาพแบบเฉลี่ยไม่ถ่วงน้ำหนัก

4.2.2 การรวมภาพเฉลี่ยแบบถ่วงน้ำหนัก

ด้วยวิธีนี้ จุดภาพบนเฟรมความละเอียดสูงจุดหนึ่งๆ จะได้รับข้อมูลจุดภาพตามปริมาณจุดภาพจากภาพความละเอียดต่ำ ในกรณีนี้ จุดภาพใดที่มีจำนวนข้อมูลจุดภาพจากภาพความละเอียดต่ำเป็นจำนวนมากกว่า จะมีผลต่อภาพลัพธ์มากกว่าจุดภาพที่มีจำนวนข้อมูลจุดภาพจากภาพความละเอียดต่ำเป็นจำนวนน้อยกว่า หรืออีกนัยหนึ่งคือตัวคูณถ่วงน้ำหนักจะมีค่าน้อยเมื่อจำนวนข้อมูลจากภาพความละเอียดต่ำมีค่าน้อย ทำให้ภาพลัพธ์ที่ได้มีรายละเอียดชัดเจนมากขึ้น ดังตัวอย่างในรูปที่ 4-8



รูปที่ 4-8 การรวมภาพเฉลี่ยแบบถ่วงน้ำหนัก

4.3 ระเบียบวิธีที่ใช้ในการแก้สมการแบบจำลองของปัญหา

ในงานวิจัยนี้ได้เลือกระเบียบวิธีการสร้างกลับภาพความละเอียดสูงเชิงกำหนดแบบคุ่มค่าซึ่งได้กล่าวถึงในบทที่ 2 ดังแสดงในสมการ

$$x = \arg \min_{\hat{x}} \left[\sum_{k=1}^L \|y_k - H_k \hat{x}\|_p^p + a \|C\hat{x}\|_q^q \right] \quad (4.2)$$

เมื่อ p คืออันดับของระยะห่าง $\|\cdot\|$ หรือนอร์มซึ่งนิยมใช้ค่าในช่วง $1 \leq p \leq 2$ โดยใช้ควบคู่กับพจน์คุ่มค่า ซึ่งเรียกโดยรวมว่าฟังก์ชันต้นทุน (Cost Function) และ a คือตัวคูณลากวางซึ่งทำหน้าที่ถ่วงดุลระหว่างพจน์รายละเอียด (พจน์แรกทางขวามือ) และพจน์ความราบเรียบของคำตอบ (พจน์ที่สองทางขวามือ)

การเลือกอันดับ p นอร์มมีความสำคัญต่อกระบวนการหาผลเฉลย x เนื่องจากระเบียบวิธีที่ดีนั้นต้องมีความทนทานต่อข้อมูลนอกกลุ่ม (Outlier) และสัญญาณรบกวน จากผลการวิจัยใน [2] บ่งชี้ว่าการเลือกอันดับของนอร์มเท่ากับ 1 ซึ่งเทียบได้กับการหาค่ามัธยฐาน (Median) ของข้อมูลนั้น มีความทนทานต่อข้อมูลนอกกลุ่มมากกว่าการเลือกอันดับของนอร์มเท่ากับ 2 ซึ่งเทียบได้กับการหาค่าเฉลี่ย (Mean) เนื่องจากข้อมูลนอกกลุ่มโดยส่วนใหญ่จะมีค่าแตกต่างไปจากข้อมูลปกติอย่างมาก ซึ่งเมื่อ

เลือกใช้วิธีหาค่าเฉลี่ยแล้ว จะทำให้ค่ากลางที่หาได้คลาดเคลื่อนไปมาก แต่เมื่อใช้วิธีหาค่ามัธยฐานแล้ว ข้อมูลนอกกลุ่มเหล่านี้จะไม่มีผลกระทบต่อกลุ่มข้อมูลปกติ เนื่องจากวิธีการนี้จะคัดเลือกเพียงข้อมูลเดียวที่อยู่ตำแหน่งกึ่งกลางเท่านั้น

อย่างไรก็ตาม ระเบียบวิธีในนอร์มอันดับ 1 นั้นจะใช้ได้ดีในกรณีกำหนดเกิน ซึ่งมีภาพความละเอียดต่ำเป็นจำนวนมากกว่า r^2 เท่านั้น เมื่อนำไปใช้กับชุดภาพต่อเนื่องจากเพิ่มภาพวีดิทัศน์ ซึ่งมีจำนวนภาพความละเอียดต่ำเพียง 6-8 ภาพสำหรับตัวประกอบการขยายในช่วง 2-4 เท่า ทำให้เป็นปัญหากำหนดขนาด และจุดภาพหนึ่งๆ บนภาพความละเอียดสูงอาจมีข้อมูลจากภาพความละเอียดต่ำเพียงจุดเดียวหรืออาจไม่มีข้อมูลเลย ในกรณีนี้ระเบียบวิธีในนอร์มอันดับ 1 และ 2 จะให้ผลลัพธ์แบบเดียวกัน [2] ซึ่งในงานวิจัยนี้ได้เลือกระเบียบวิธีในนอร์มอันดับ 2 เนื่องจากสอดคล้องกับสัญญาณรบกวนแบบเกาส์ในแบบจำลองของปัญหา

สำหรับพจน์คุมค่า $\|Cx\|_q^q$ นั้นมีหน้าที่หลักคือประมาณค่าจุดภาพในตำแหน่งที่ไม่มีข้อมูลจากภาพความละเอียดต่ำ ทำให้ภาพผลลัพธ์ที่ได้มีความราบเรียบมากขึ้น ซึ่งมีประโยชน์ในกรณีที่ภาพความละเอียดต่ำมีจำนวนน้อย โดยพจน์คุมค่าที่ดีนั้น ควรจะรักษาคุณสมบัติที่ต้องการให้มีในภาพผลลัพธ์ความละเอียดสูงไว้ได้ เช่น ความคมชัดของลักษณะเด่นที่เป็นขอบ (Edge) โดยยังคงความสามารถในการกำจัดข้อมูลนอกกลุ่มและสัญญาณรบกวนไว้ได้อยู่ พจน์คุมค่าที่นิยมใช้ในกระบวนการขยายรายละเอียดภาพได้แก่พจน์คุมค่าแบบทิคอนอฟ (Tikhonov Regularization Term) ซึ่งเป็นตัวดำเนินการลาปลาซ (Laplacian Operator) ที่มีลักษณะเป็นตัวกรองเต็มหน่วยแบบผ่านสูง (Discrete High Pass Filter) และเป็นตัวตรวจจับลักษณะขอบ (Edge Detector) โดยมีลักษณะดังแสดงในสมการ (3.10) ซึ่งเป็นตัวดำเนินการที่เลือกใช้ในงานวิจัยนี้

$$C = \frac{1}{4} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.3)$$

ตัวดำเนินการดังกล่าวทำหน้าที่ประมาณค่าจุดภาพรอบข้างที่ขาดหาย โดยขนาด 3×3 ของตัวดำเนินการนี้ครอบคลุมพื้นที่บริเวณรอบๆ เพียงพอสำหรับการขยายรายละเอียดภาพด้วยตัวประกอบการขยายในช่วง 2 ถึง 4 เท่า จากการทดลองพบว่าหากตัวดำเนินการลาปลาซมีขนาดตั้งแต่ 5×5 ขึ้นไปจะทำให้ภาพผลลัพธ์ที่ได้เกิดการอ้อมตัวของจุดภาพและผิดเพี้ยนไปอย่างมากเมื่อเปรียบเทียบกับภาพความละเอียดสูงต้นแบบ

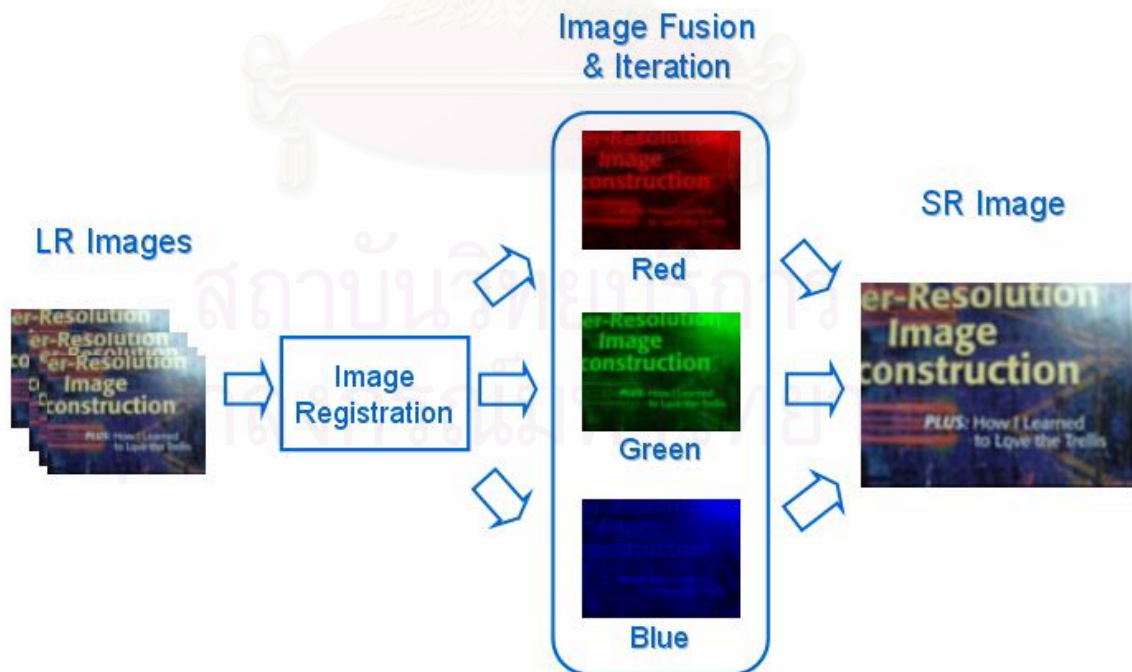
เมื่อแก้สมการด้วยระเบียบวิธีทำซ้ำแบบลาดลงชันที่สุด (Steepest Descent Iterative Method) จะสามารถเขียนผลเฉลยจากการวนรอบแต่ละรอบในรูปแบบปิด (Closed-form) ได้เป็น [1]

$$x_{n+1} = x_n + b \left[\sum_{k=1}^L H_k^T (y_k - H_k x_n) - a C^T C x_n \right] \quad (4.4)$$

โดยที่ภาพเริ่มต้น x_0 อาจเป็นภาพเปล่า (Null Image) หรือเป็นภาพที่ถูกปรับปรุงคุณภาพของภาพ (Image Enhancement) มาก่อน เช่น ภาพความละเอียดต่ำที่ถูกขยายขนาดด้วยวิธีไบคิวบิกก็ได้ ซึ่งจากการทดลองพบว่าในกรณีหลังคำตอบจะลู่เข้าได้เร็วกว่า b คือระยะก้าว (Step Size) ของคำตอบที่ลู่เข้าสู่ผลเฉลยในแต่ละรอบการทำซ้ำ ซึ่งอาจเลือกใช้แบบระยะก้าวคงที่ (Fixed) หรือแบบปรับตัวได้ (Adaptive) ก็ได้ โดยหากเลือกระยะก้าวขนาดใหญ่จะทำให้คำตอบลู่เข้าได้เร็ว แต่จะมีปัญหาเรื่องการอ้อมตัวของจุดทำยในช่องทำยๆ ของการวนรอบ

4.4 การขยายรายละเอียดภาพสี

ด้วยหลักการที่กล่าวข้างต้นซึ่งเป็นขั้นตอนขยายรายละเอียดภาพทั่วไปกล่าวคือ สำหรับภาพแบบช่องสัญญาณเดี่ยว (ภาพขาวดำ) จึงสามารถนำมาประยุกต์ใช้กับการขยายรายละเอียดภาพสีได้ โดยการพิจารณาภาพสีต้นแบบว่าเป็นภาพเดี่ยว 3 ช่องสัญญาณ โดยมีขั้นตอนซ้อนทับภาพโดยกำหนดตำแหน่งร่วมกัน แล้วทำการประมวลผลภาพแยกอิสระจากกันที่ละช่องสัญญาณสี จากนั้นจึงนำภาพผลลัพธ์ที่ได้มาซ้อนกลับใน 3 ช่องสัญญาณ ทำให้ได้ภาพสีตามต้องการดังแสดงในรูปที่ 4-9



รูปที่ 4-9 การขยายรายละเอียดภาพสี

ในงานวิจัยนี้ได้ทำการทดลองประมวลผลภาพสีในปริภูมิสีหลายแบบได้แก่ ปริภูมิสีแบบ RGB และปริภูมิสีแบบ YCbCr โดยอาศัยการแปลงเชิงเส้นจากปริภูมิ RGB ไปยังปริภูมิ YCbCr สำหรับภาพสีแบบ 8 บิตตามความสัมพันธ์ [10]

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ Cb &= 128 - 0.169R - 0.331G + 0.500B \\ Cr &= 128 + 0.500R - 0.418G - 0.081B \end{aligned} \quad (4.5)$$

เมื่อ R, G, B แทนค่าจุดภาพในช่องสัญญาณสีแดง สีเขียว สีน้ำเงิน และ Y, Cb, Cr แทนค่าจุดภาพในช่องสัญญาณสองสว่าง สัญญาณสีน้ำเงิน และสัญญาณสีแดงตามลำดับ

4.5 สรุปท้ายบท

ในบทนี้ได้กล่าวถึงระเบียบวิธีที่เลือกใช้ในการออกแบบโปรแกรมสำหรับแต่ละขั้นตอนของการขยายรายละเอียดภาพ ในขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งนั้นได้ถูกออกแบบให้สามารถซ้อนทับภาพโดยอัตโนมัติได้โดยใช้ระเบียบวิธีอิงลักษณะเด่นร่วมกับระเบียบวิธีคัดกรองแรนแซค ทำให้สามารถคำนวณเมตริกซ์แปลงภาพได้อย่างถูกต้องแม่นยำมากยิ่งขึ้น ในขั้นตอนการรวมภาพได้เลือกใช้ระเบียบวิธีรวมภาพแบบเฉลี่ยถ่วงน้ำหนักซึ่งให้ภาพผลลัพธ์ที่มีความเปรียบต่างดีกว่า ในส่วนของการแก้สมการแบบจำลองของปัญหาได้เลือกใช้ระเบียบวิธีเชิงกำหนดร่วมกับพจน์ค่าแบบทีโคเนอฟ และวิธีทำซ้ำแบบลาดลงขั้นที่สุด จากหลักการดังกล่าวจึงสามารถนำประยุกต์ใช้สำหรับการขยายรายละเอียดภาพสีได้โดยพิจารณาของสัญญาณสีแต่ละช่องว่าเป็นช่องสัญญาณเดี่ยวแยกอิสระกันซึ่งมีรูปแบบการแปลงภาพแบบเดียวกัน จากนั้นจึงนำผลลัพธ์ที่ได้จากแต่ละช่องสัญญาณมารวมกลับเป็นภาพสีได้

บทที่ 5

การออกแบบโปรแกรมขยายรายละเอียดภาพ

ในบทนี้จะกล่าวถึงการออกแบบโปรแกรมร่วมกับไลบรารีโอเพนซีวี ฟังก์ชันที่จำเป็นต่อการประมวลผล ขั้นตอนการประมวลผลโดยละเอียดของโปรแกรมขยายรายละเอียดภาพในงานวิจัย

5.1 ไลบรารีโอเพนซีวี

ในงานวิจัยนี้ได้เลือกใช้โปรแกรม **Microsoft Visual Studio 2005** สำหรับออกแบบโปรแกรมขยายรายละเอียดภาพในภาษา **C++** ร่วมกับฟังก์ชันประมวลผลภาพในไลบรารีโอเพนซีวี (**OpenCV**) ซึ่งย่อมาจาก **Open Source Computer Vision Library** เป็นไลบรารีภาษา **C++** ที่ถูกออกแบบและพัฒนาโดยบริษัทอินเทลคอร์ปอเรชันสำหรับใช้ในงานด้านคอมพิวเตอร์วิทัศน์เวลาจริง (**Real Time Computer Vision**) ซึ่งเปิดเผยมอบให้บุคคลภายนอกสามารถนำไลบรารีนี้ไปใช้ได้อย่างอิสระ โดยภายในบรรจุฟังก์ชันที่จำเป็นและสะดวกต่อการประมวลผลสัญญาณภาพ เช่น การดำเนินการทางคณิตศาสตร์และทางตรรก การระบุและตรวจจับวัตถุในภาพ การแปลงภาพ เป็นต้น โดยมีคู่มือการใช้งานอย่างละเอียดมาให้ ผู้สนใจสามารถดูรายละเอียดของไลบรารีนี้ได้ใน [11]

5.2 ฟังก์ชันหลักที่ใช้ในการเขียนโปรแกรม

เนื่องจากบางฟังก์ชันในไลบรารีโอเพนซีวีไม่สามารถตอบสนองความต้องการทั้งหมดของขั้นตอนวิธีที่เลือกใช้ได้ จึงจำเป็นต้องออกแบบบางฟังก์ชันขึ้นมาใหม่ โดยฟังก์ชันหลักที่จำเป็นในงานวิจัยนี้ได้แก่

5.2.1 ฟังก์ชันในไลบรารีโอเพนซีวี

a) cvGoodFeatureToTrack

ใช้สำหรับค้นหาตำแหน่งของลักษณะเด่นเช่น มุม จุดตัดของเส้น หรือขอบวัตถุในภาพ โดยใช้ระเบียบวิธีตรวจจับมุมแบบค่าลักษณะเฉพาะน้อยสุด (**Minimal Eigenvalue Corner Tracking**) หรือวิธีตรวจจับมุมแบบแฮร์ริส (**Harris Corner Tracking**) โดยสามารถกำหนดจำนวนจุดควบคุมที่ต้องการตรวจจับและระยะน้อยสุดระหว่างจุดควบคุมแต่ละจุดได้ ซึ่งโปรแกรมจะคัดเลือกจุดควบคุมที่มีค่าลักษณะเฉพาะมากที่สุดและอยู่ห่างกันไม่ต่ำกว่าค่าที่ระบุไว้ นอกจากนี้ยังสามารถกำหนดขนาดของหน้าต่างที่ใช้ในการเฉลี่ยค่าได้ ลักษณะเด่นที่ตรวจหาได้นี้จะถูกใช้เป็นจุดควบคุมสำหรับขั้นตอนถัดไปโดยกำหนดตำแหน่งแบบอิงลักษณะเด่นต่อไป

b) cvOpticalFlowPyrLK

ใช้สำหรับจับคู่จุดควบคุมที่สอดคล้องกันระหว่างภาพ 2 ภาพด้วยวิธีการไหลเชิงแสงแบบพีรามิดของลูคัส-คานาเด่ (Lucas-Kanade Optical Flow Method) [x] โดยมีความแม่นยำในระดับใต้จุดภาพ (Sub-pixel Accuracy) ตัวแปรเสริมที่สำคัญในฟังก์ชันนี้ได้แก่ ขนาดของหน้าต่างค้นหา ระดับของพีรามิด

c) cvLaplace

ทำหน้าที่คูณประสานกับภาพตั้งต้นด้วยตัวดำเนินการลาปลาซหรือตัวดำเนินการอนุพันธ์เต็มหน่วยขนาด 3×3 (ตัวดำเนินการ C ในแบบจำลองของปัญหา)

d) cvSmooth

ทำหน้าที่คูณประสานกับภาพตั้งต้นด้วยตัวดำเนินการพรมัวหรือเคอร์เนลแบบเกาส์ โดยสามารถเลือกขนาดของเคอร์เนลและค่าเบี่ยงเบนมาตรฐานได้ ฟังก์ชันนี้ทำหน้าที่เป็นตัวดำเนินการ B ในสมการแบบจำลองของปัญหา

e) cvWarpPerspective

ทำหน้าที่แปลงภาพแบบทัศนมิติโดยย้ายพิกัดจุดภาพเดิมไปสู่พิกัดจุดภาพใหม่ $P(x, y) \Rightarrow P(u, v)$ ด้วยเมตริกซ์แปลงภาพขนาด 3×3 ตามสมการ

$$u = \frac{H_{11} * x + H_{12} * y + H_{13}}{H_{31} * x + H_{32} * y + H_{33}}, v = \frac{H_{21} * x + H_{22} * y + H_{23}}{H_{31} * x + H_{32} * y + H_{33}} \quad (5.1)$$

เมื่อ H_{ij} คือสมาชิกในแถวที่ i และสดมภ์ที่ j ตามลำดับ โดยฟังก์ชันนี้จะชดเชยค่าจุดภาพที่ขาดหายไปโดยอัตโนมัติ ภาพที่ได้จากการแปลงด้วยฟังก์ชันนี้จึงมีความต่อเนื่อง และไม่มีตำแหน่งที่ข้อมูลจุดภาพขาดหายไป โดยฟังก์ชันนี้ทำหน้าที่เป็นตัวดำเนินการ W ในสมการแบบจำลองของปัญหา

f) cvGetPerspectiveTransform

เป็นอินเวอร์สฟังก์ชันของฟังก์ชัน cvWarpPerspective โดยทำหน้าที่คำนวณเมตริกซ์แปลงภาพจากคู่ของจุดควบคุมจำนวน 4 คู่ ซึ่งในงานวิจัยนี้ได้นำมาใช้เป็นขั้นตอนหนึ่งของฟังก์ชันซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติที่เขียนขึ้นเอง ซึ่งต้องคำนวณเมตริกซ์แปลงภาพจากคู่ควบคุมที่สอดคล้องตั้งแต่ 4 ถึง 16 คู่เพื่อให้มีความแม่นยำมากที่สุด

g) cvSolve

ทำหน้าที่แก้ระบบสมการ $AX = B$ ด้วยระเบียบวิธีย่อยสลายค่าเอกฐาน (Singular Value Decomposition) โดยสามารถใช้ได้กับระบบสมการขนาดใหญ่ ซึ่งในงานวิจัยนี้ได้นำมาใช้เพื่อคำนวณเมตริกซ์แปลงภาพโดยการแก้สมการ (4.1)

5.2.2 ฟังก์ชันที่เขียนขึ้นมาใหม่

a) RoundedWarp

ทำหน้าที่แปลงภาพแบบทัศนมิติโดยอาศัยจุดควบคุม 4 คู่เช่นเดียวกับฟังก์ชัน `cvWarpPerspective` โดยฟังก์ชันนี้จะทำการปิดเศษทศนิยมของค่าตำแหน่งให้เป็นจำนวนเต็มโดยอัตโนมัติ และไม่มีการประมาณค่าในช่วงสำหรับตำแหน่งจุดภาพที่ไม่มีข้อมูล เพื่อให้การประมวลผลสอดคล้องกับแบบจำลองของปัญหาที่ตั้งไว้ ทำหน้าที่สร้างเมตริกซ์ถ่วงน้ำหนักสำหรับแต่ละจุดภาพในขั้นตอนรวมภาพ

b) UpSampling & DownSampling

ทำหน้าที่เสมือนตัวดำเนินการชักสุมตัวอย่างขาขึ้นและขาลงแบบกระจายจุดตามลำดับดังที่ได้แสดงไว้ในรูปที่ 3-9 เพื่อขยายขนาดภาพ

c) MeanAdd_Mask

ทำหน้าที่รวมภาพ 8 ภาพแบบเฉลี่ยถ่วงน้ำหนักโดยทำงานร่วมกับแผ่นบัง (Mask) แบบถ่วงน้ำหนักซึ่งมีข้อมูลจำนวนภาพความละเอียดต่ำในแต่ละจุดภาพ ทำให้จุดภาพที่ได้ข้อมูลจากภาพความละเอียดต่ำจำนวนมากกว่ามีผลกระทบต่อภาพลัพธ์มากกว่าจุดภาพที่มีข้อมูลจากภาพความละเอียดต่ำจำนวนน้อย

d) ImageRegister

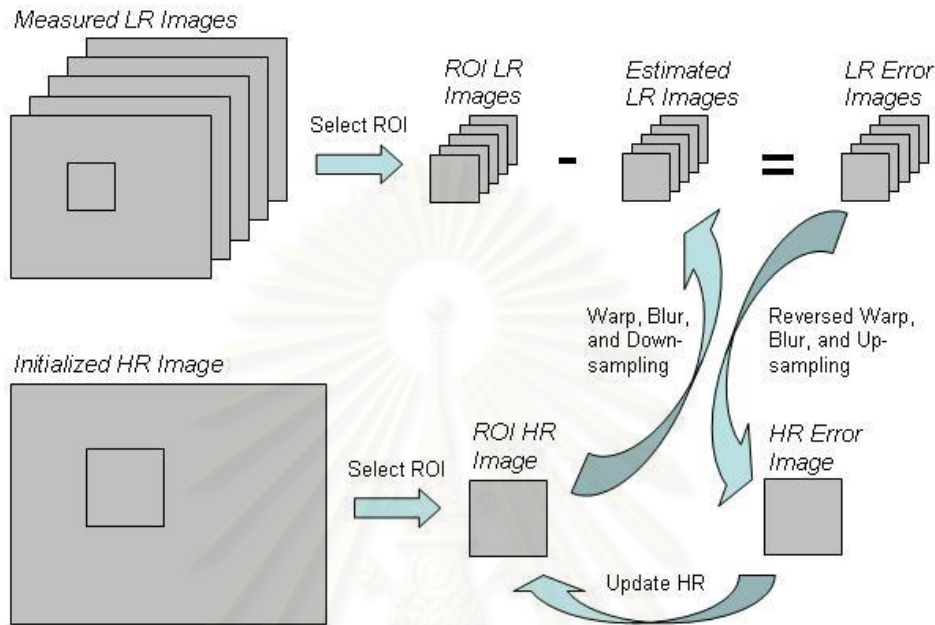
ทำหน้าที่ซ้อนทับภาพโดยกำหนดตำแหน่งระหว่างภาพขาวดำแบบ 8 บิตโดยอัตโนมัติ โดยฟังก์ชันจะทำการค้นหาจุดควบคุมในภาพ จับคู่จุดควบคุมด้วยฟังก์ชันการไหลเชิงแสง และคัดกรองคู่ของจุดควบคุมที่ไม่เข้าพวกทั้งด้วยระเบียบวิธีคัดกรองแรงแชค และคืนค่าเป็นเมตริกซ์แปลงภาพแบบทัศนมิติ 2 เมตริกซ์ได้แก่ เมตริกซ์แปลงภาพไปข้างหน้า (Forward Transformation Matrix) และเมตริกซ์แปลงภาพย้อนกลับ (Inverse Transformation Matrix)

e) PSNR_Calculation

ทำหน้าที่คำนวณค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR : คูณรายละเอียดในบทที่ 6) ระหว่างภาพ 2 ภาพ ในที่นี้ได้แก่ความละเอียดสูงต้นแบบและภาพที่ผ่านการขยายรายละเอียด เพื่อใช้เป็นดัชนีเปรียบเทียบคุณภาพของภาพลัพธ์ที่ได้

5.3 การออกแบบโปรแกรม

ขั้นตอนการทำงานของโปรแกรมขยายรายละเอียดภาพ อธิบายเป็นแผนภาพได้ตามรูปที่ 5-1 ซึ่งเป็นกระบวนการโดยสังเขปต่อการประมวลผล 1 รอบและต่อ 1 ช่องสัญญาณสี

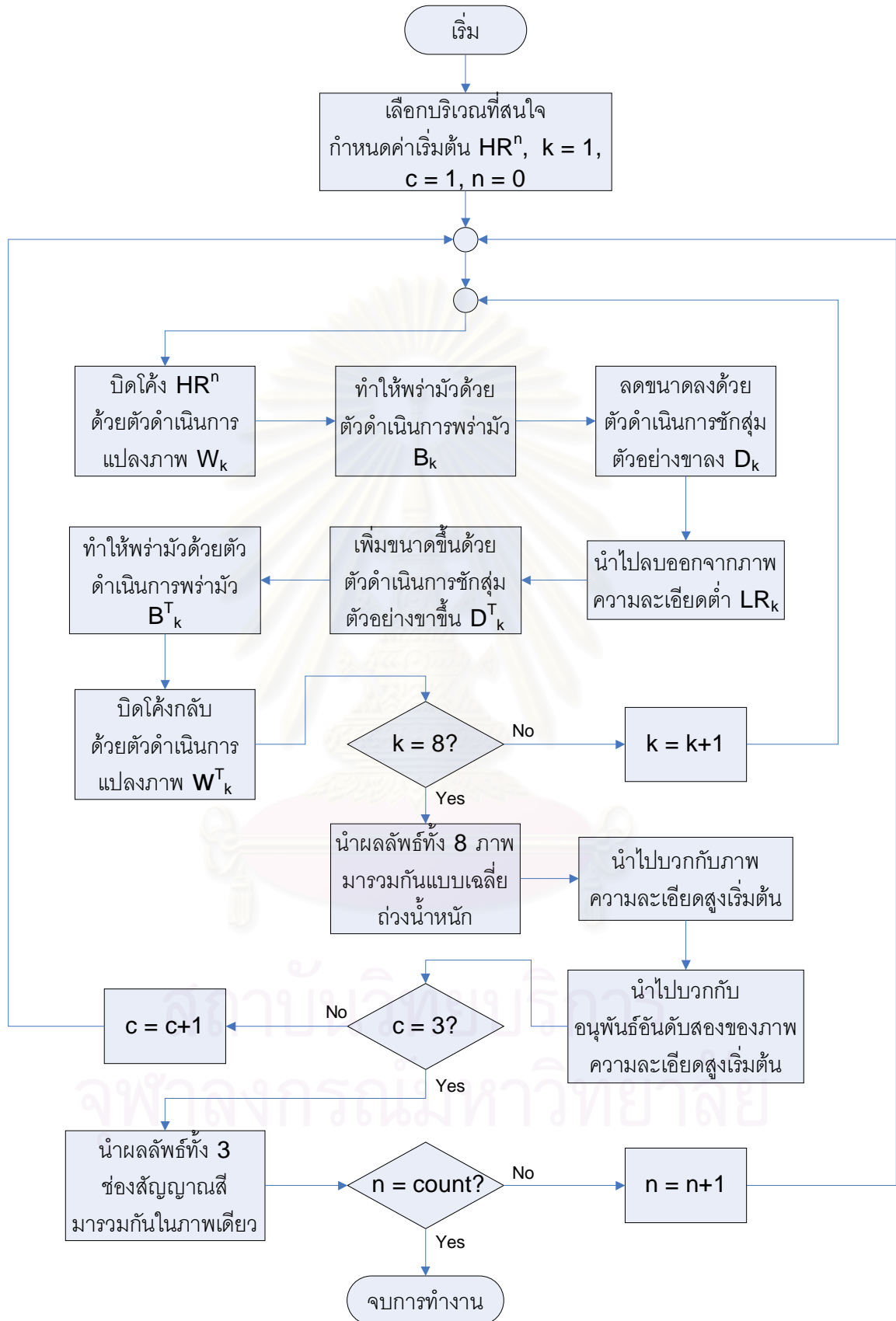


รูปที่ 5-1 แผนภาพการทำงานของโปรแกรมขยายรายละเอียดภาพที่น่าเสนอ

การทำงานโดยละเอียดของโปรแกรมแสดงเป็นผังงานได้ตามรูปที่ 5-2 โดยเริ่มจากผู้ใช้เล่นกลับเพิ่มภาพวิดีโอที่ค้นความละเอียดต่ำ จากนั้นเลือกหยุดเล่น ณ ตำแหน่งเฟรมที่ต้องการและเลือกบริเวณที่สนใจ (Region of Interest: ROI) ซึ่งในที่นี้คือบริเวณที่ต้องการขยายรายละเอียด โปรแกรมจะตัดลอกภาพในบริเวณที่สนใจในเฟรมก่อนหน้าจำนวน 4 เฟรมและเฟรมถัดไปจำนวน 4 เฟรม นั่นคือโปรแกรมจะใช้ภาพความละเอียดต่ำจำนวน 8 ภาพในการสร้างกลับภาพความละเอียดสูง 1 ภาพ

จากนั้นโปรแกรมจะแปลงภาพเฟรมปัจจุบันซึ่งกำหนดเป็นเฟรมอ้างอิง และเฟรมข้างเคียงทั้ง 8 เฟรมให้เป็นภาพขาวดำ และใช้ฟังก์ชันตรวจหาลักษณะเด่นเพื่อค้นหาจุดควบคุมในเฟรมปัจจุบันจำนวน 16 จุด และใช้ฟังก์ชันการไหลเชิงแสงในการตรวจหาจุดควบคุมที่สอดคล้องในเฟรมข้างเคียงทั้ง 8 เฟรม โดยการคัดกรองด้วยระเบียบวิธีแรนแซค ตามที่ได้กล่าวถึงในบทที่ 4 โดยขั้นตอนคำนวณเมตริกซ์แปลงภาพนี้จะทำเพียงครั้งเดียวโดยพิจารณาว่าสัญญาณสีของภาพทั้ง 3 ช่องมีการแปลงภาพแบบเดียวกัน

จากนั้นผู้ใช้จะกำหนดตัวประกอบการขยายที่ต้องการโดยการคลิกที่ภาพความละเอียดต่ำ ซึ่งจะทำให้ภาพดังกล่าวมีขนาดใหญ่ขึ้นเป็น 2 3 และ 4 เท่าตามลำดับ เมื่อผู้ใช้คลิกเลือกค่าตัวประกอบการขยายเสร็จแล้ว โปรแกรมจะเก็บค่าดังกล่าวไว้คำนวณเมตริกซ์สำหรับแปลงภาพที่ถูกขยายต่อไป



รูปที่ 5-2 ผังงานการทำงานของโปรแกรม

เมื่อได้ตัวดำเนินการที่จำเป็นครบถ้วนแล้ว การประมวลผลของโปรแกรมจะเริ่มจากกำหนดช่องสัญญาณสีที่ต้องการประมวลผลก่อน โดยในงานวิจัยนี้จะเริ่มจากช่องสัญญาณสีน้ำเงินของภาพเริ่มต้นและเริ่มประมวลผลสำหรับภาพความละเอียดต่ำเฟรมที่ 1 โดยกำหนดค่าเริ่มต้นของภาพความละเอียดสูงแบบ 32 บิตซึ่งเป็นรูปแบบของภาพที่ใช้ประมวลผลได้ จากนั้นจึงผ่านตัวดำเนินการลดทอนคุณภาพต่างๆ ได้แก่ ตัวดำเนินการบิดโค้งหรือเมตริกซ์แปลงภาพ ตัวดำเนินการพหามัว และตัวดำเนินการซักร่มตัวอย่างขาลง เพื่อให้ได้ภาพความละเอียดต่ำเสมือนจากนั้นจึงนำไปเปรียบเทียบกับภาพความละเอียดต่ำเฟรมที่ 1 ที่มีอยู่เพื่อหาส่วนที่แตกต่าง และนำส่วนต่างดังกล่าวผ่านกระบวนการลดทอนคุณภาพย้อนกลับ จากนั้นดำเนินการซ้ำสำหรับภาพความละเอียดต่ำเฟรมที่ 2 ถึง 8 และนำผลลัพธ์ซึ่งเป็นส่วนต่างที่ได้ทั้งหมดไปบวกเข้ากับภาพความละเอียดสูงเริ่มต้นแบบเฉลี่ยถ่วงน้ำหนัก เป็นอันเสร็จกระบวนการขยายรายละเอียดภาพสำหรับ 1 ช่องสัญญาณสี และสำหรับ 1 รอบการประมวลผล

จากนั้นจึงคัดลอกช่องสัญญาณสีเขียวและสีแดงเพื่อดำเนินการตามขั้นตอนข้างต้น และนำผลลัพธ์ที่ได้ทั้ง 3 ช่องสัญญาณสีมาแปลงกลับเป็นรูปแบบ 8 บิตและนำมารวมกันในภาพเดียว ทำให้ได้ภาพสีที่ผ่านการประมวลผลแล้ว 1 รอบ

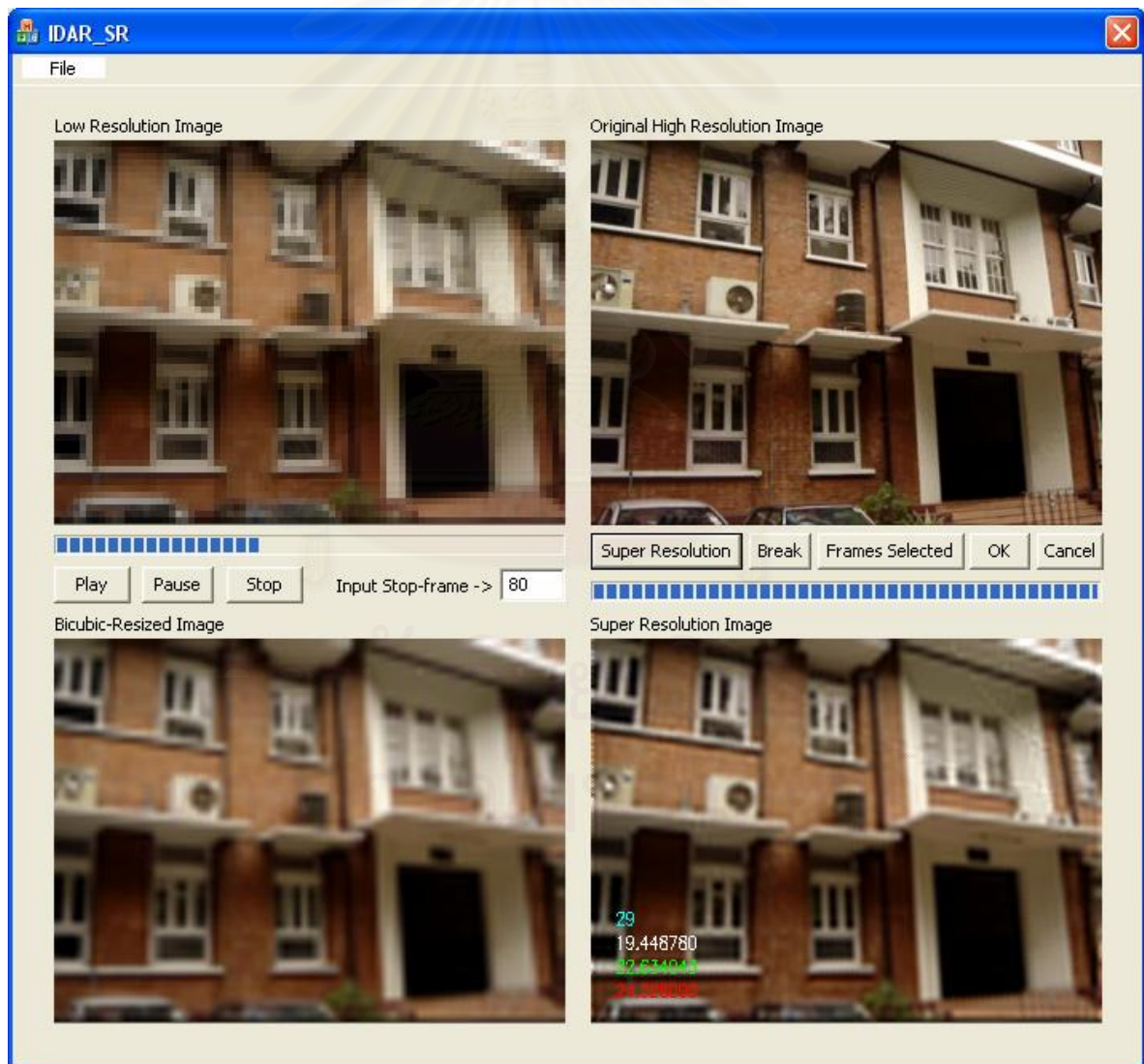
จากนั้นดำเนินการซ้ำขั้นตอนทั้งหมดสำหรับการประมวลผลรอบต่อไปโดยนำภาพสีผลลัพธ์ที่ได้ไปเป็นภาพเริ่มต้นสำหรับรอบถัดไป โดยทั่วไปจำนวนรอบการทำซ้ำที่เหมาะสมอยู่ที่ 50-100 และใช้เวลาในการประมวลผลประมาณ 1 วินาทีต่อรอบ สำหรับการขยายรายละเอียดภาพจากภาพสีขนาด 80x60 จุดภาพไปเป็นภาพสีขนาด 320x240 จุดภาพ (ตัวประกอบขยายเท่ากับ 4)

5.4 โปรแกรมขยายรายละเอียดภาพที่นำเสนอ

โปรแกรมขยายรายละเอียดภาพมีรูปแบบของหน้าต่างแสดงผลดังแสดงในรูปที่ 5-3 ซึ่งโปรแกรมถูกออกแบบมาให้เล่นกลับแฟ้มภาพวิดีโอที่คืนความละเอียดสูงไปพร้อมๆ กับแฟ้มภาพความละเอียดต่ำเพื่อเปรียบเทียบผลลัพธ์ที่ได้กับภาพความละเอียดสูงต้นแบบ โดยผู้ใช้สามารถเลือกตำแหน่งของเฟรมที่ต้องการขยายรายละเอียดได้โดยระบุเป็นตัวเลขลำดับที่ของเฟรม ซึ่งโปรแกรมจะหยุดเล่นกลับเมื่อวิ่งมาถึงตำแหน่งเฟรมที่เลือกไว้โดยอัตโนมัติ นอกจากนี้โปรแกรมจะแสดงภาพความละเอียดสูงที่ได้จากวิธีไบควิกเพื่อเปรียบเทียบกับวิธีที่นำเสนออีกด้วย โดยโปรแกรมจะแสดงค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (ดูรายละเอียดในบทที่ 7) ของแต่ละภาพเทียบกับภาพความละเอียดสูงต้นแบบที่บริเวณมุมขวาล่างของภาพ

เมื่อผู้ใช้เลือกแฟ้มภาพเคลื่อนไหวที่ต้องการแล้วกดปุ่ม Play โปรแกรมจะเล่นกลับภาพให้จนเมื่อถึงตำแหน่งเฟรมที่ระบุไว้ก็จะหยุด ซึ่ง ณ จุดนี้ผู้ใช้สามารถเลือกดูภาพความละเอียดต่ำจำนวน 8 ภาพที่

ใช้เป็นสัญญาณเข้าสำหรับประมวลผลได้โดยการคลิกที่ปุ่ม **Frames Selected** ซึ่งโปรแกรมจะแสดงภาพความละเอียดให้ที่ละภาพแบบวนในช่องล่างขวา จากนั้นผู้ใช้จะเลือกตัวประกอบการขยายโดยการคลิกที่ภาพความละเอียดต่ำ โดยขนาดของภาพจะเปลี่ยนไปเป็น 2 เท่า 3 เท่า และ 4 เท่าตามลำดับ เมื่อได้ค่าตัวประกอบการขยายที่ต้องการแล้วจึงกดปุ่ม **Superresolution** โปรแกรมจะทำการประมวลผลโดยการแสดงผลภาพผลลัพธ์ที่ได้จากการคำนวณในแต่ละรอบ โดยผู้ใช้จะสามารถทราบได้ว่าโปรแกรมใกล้จะประมวลผลเสร็จแล้วหรือไม่จากแถบ **Progress Bar** หลังจากการประมวลผลเสร็จสิ้นแล้ว เมื่อผู้ใช้กดปุ่ม **OK** โปรแกรมจะทำการบันทึกภาพทุกภาพที่แสดงในหน้าต่างแสดงผลพร้อมทั้งแฟ้มบันทึกข้อมูลของแฟ้มภาพซึ่งได้แก่ชื่อแฟ้ม ตำแหน่งของเฟรมที่เลือก ค่าตัวแปรเสริมเช่นคอร์เนลความพรวดมัว ตัวคูณลากรางจ์ ระยะเวลา ค่า **PSNR** โดยวิธีข้างเคียงใกล้สุด วิธีไบคิวบิก และวิธีที่นำเสนอของภาพเหล่านี้ในแต่ละรอบการทำซ้ำ รวมถึงเวลาที่ใช้ในการประมวลผลไว้ในไดเรกทอรีที่กำหนดไว้โดยอัตโนมัติ



รูปที่ 5-3 หน้าต่างแสดงผลของโปรแกรมขยายรายละเอียดภาพ

5.5 สรุปท้ายบท

โปรแกรมขยายรายละเอียดภาพในงานวิจัยนี้ถูกออกแบบด้วยโปรแกรม **Visual Studio 2005** ในภาษา **C++** โดยใช้ร่วมกับไลบรารี **OpenCV** ฟังก์ชันที่จำเป็นในไลบรารีถูกเลือกมาใช้ในขั้นตอนการประมวลผลของโปรแกรม แต่ยังไม่สามารถตอบสนองของความต้องการของขั้นตอนวิธีที่เลือกใช้ได้ จึงได้มีการเขียนฟังก์ชันขึ้นเองเพิ่มเติม ขั้นตอนการขยายรายละเอียดภาพเริ่มจากสมมุติภาพความละเอียดสูงเริ่มต้น จากนั้นจึงผ่านตัวดำเนินการตัดทอนคุณภาพแบบต่างๆ กันเพื่อให้ได้ภาพความละเอียดต่ำ จากนั้นจึงนำภาพที่ได้ไปเปรียบเทียบกับภาพความละเอียดต่ำที่มีอยู่ และนำผลต่างผ่านตัวดำเนินการลดทอนคุณภาพย้อนกลับ เพื่อไปปรับปรุงกลับที่ภาพความละเอียดสูงที่สมมุติในตอนแรก โดยสามารถทำซ้ำได้เพื่อให้ภาพมีคุณภาพดียิ่งขึ้น ในโปรแกรมขยายรายละเอียดภาพที่นำเสนอสามารถนำภาพความละเอียดสูงต้นแบบมาใช้เปรียบเทียบกับภาพที่ได้ผ่านขยายรายละเอียดแล้วเพื่อทดสอบคุณภาพของโปรแกรมที่นำเสนอได้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 6

ผลการทดสอบโปรแกรม

ในบทนี้จะเป็นการรายงานผลการทดสอบโปรแกรมขยายรายละเอียดภาพด้วยวิธีที่นำเสนอกับเพิ่มภาพวีดิทัศน์ต่างๆ โดยจะเปรียบเทียบผลลัพธ์ที่ได้กับวิธีการขยายภาพแบบไบคิวบิก ทั้งในส่วนของภาพวัตถุ สถานที่ และบุคคล โดยทดสอบกับเพิ่มภาพสังเคราะห์และเพิ่มภาพทั่วไป

6.1 ดัชนีที่ใช้ในการวัดผลการทดสอบ

เพื่อความสะดวกในการเปรียบเทียบผลของการขยายรายละเอียดภาพด้วยโปรแกรมที่นำเสนอกับการขยายขนาดภาพด้วยวิธีทั่วไป จึงได้กำหนดดัชนีสำหรับการเปรียบเทียบคือ อัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR: Peak Signal-to-Noise Ratio) มีหน่วยเป็นเดซิเบล โดยมีนิยามสำหรับภาพแบบ 8 บิตตามสมการ (6.1)

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right] = 20 \log_{10} \left[\frac{255}{\sqrt{MSE}} \right] \quad (6.1)$$

เมื่อ 255 คือค่าสูงสุดของจุดภาพซึ่งเป็นจำนวนเต็มบวก และ MSE คือค่าผิดพลาดเฉลี่ยกำลังสอง (Mean Square Error) ซึ่งมีนิยามสำหรับภาพสีตามสมการ (6.2)

$$MSE = \frac{\sum_{c=1}^3 \sum_{m=1}^{rM} \sum_{n=1}^{rN} [\hat{x}_c(m,n) - x_c(m,n)]^2}{3r^2MN} \quad (6.2)$$

เมื่อ x คือภาพความละเอียดสูงต้นแบบ และ \hat{x} คือภาพความละเอียดสูงที่ได้จากการขยายรายละเอียดภาพ c คือดัชนีระบุนของสัญญาณสี (แดง เขียว น้ำเงิน) (m,n) คือดัชนีบ่งชี้ตำแหน่งจุดภาพในภาพ r^2MN คือจำนวนจุดภาพของภาพความละเอียดสูงขนาด $rM \times rN$ จุดภาพ ซึ่งถูกขยายขนาดมาจากภาพความละเอียดต่ำขนาด $M \times N$ จุดภาพ และ r คือตัวประกอบการขยาย ค่า PSNR ยิ่งมากหมายถึงภาพผลลัพธ์ที่ผ่านกระบวนการสร้างกลับมีความใกล้เคียงกับภาพความละเอียดสูงต้นแบบมาก กล่าวคือภาพผลลัพธ์ที่ได้จะมีรายละเอียดที่ดีกว่านั่นเอง

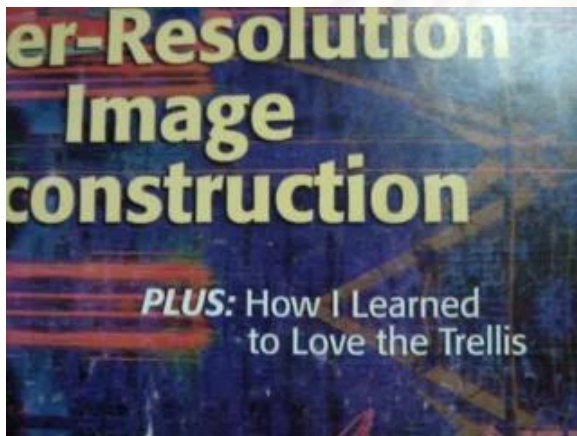
6.2 การทดสอบโปรแกรมกับเพิ่มภาพสังเคราะห์

ในหัวข้อนี้จะเป็นการทดสอบโปรแกรมกับเพิ่มภาพความละเอียดต่ำที่เตรียมจากเพิ่มภาพความละเอียดสูง ทำให้สามารถเปรียบเทียบภาพผลลัพธ์ที่ได้กับภาพความละเอียดสูงต้นแบบได้

6.2.1 กรณีการเคลื่อนที่แบบเลื่อนขนานทั้งหมด

เพิ่มภาพวีดิทัศน์ในการทดลองแรกนี้ได้มาจากการถ่ายปกวารสาร IEEE ฉบับหนึ่ง ในระยะใกล้ ด้วยกล้องดิจิทัลความละเอียด 640 เส้นที่ระยะห่างประมาณ 20 เซนติเมตร โดยให้วัตถุเป้าหมายอยู่นิ่ง แล้วเลื่อนกล้องบันทึกจากขวาไปซ้าย จากนั้นจึงนำไปลดขนาดให้เหลือ 320x240 จุดภาพสำหรับเป็นภาพความละเอียดสูงต้นแบบ และลดขนาดลงเป็น 80x60 จุดภาพเพื่อเป็นภาพความละเอียดต่ำสำหรับการทำการขยายรายละเอียดภาพ โดยเปรียบเทียบภาพผ่านการขยายรายละเอียดด้วยวิธีการที่นำเสนอกับวิธีวิธีไบคิวบิก (Bicubic) ดังแสดงในรูปที่ 6-1 พร้อมค่าอัตราสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) ซึ่งเปรียบเทียบกับภาพความละเอียดสูงต้นแบบ

ตัวแปรเสริมที่ใช้ในการทดลองแรกคือ $a = 0.0015$, $b = 2$ ขนาดของเคอร์เนลความพรั่มเท่ากับ 5×5 โดยที่ $s = 1.5$ และจำนวนรอบการทำซ้ำเท่ากับ 100 รอบโดยใช้ภาพเริ่มต้นเป็นภาพเปล่า (Null Image) จากผลการทดลองจะเห็นว่าภาพผลลัพธ์โดยวิธีการที่นำเสนอมีค่า PSNR สูงกว่าภาพที่ได้จากวิธีไบคิวบิกประมาณ 1.6 เดซิเบล และข้อความสีขาวในภาพชัดเจนขึ้นจนสามารถอ่านได้



(ก) ภาพความละเอียดสูงต้นแบบ

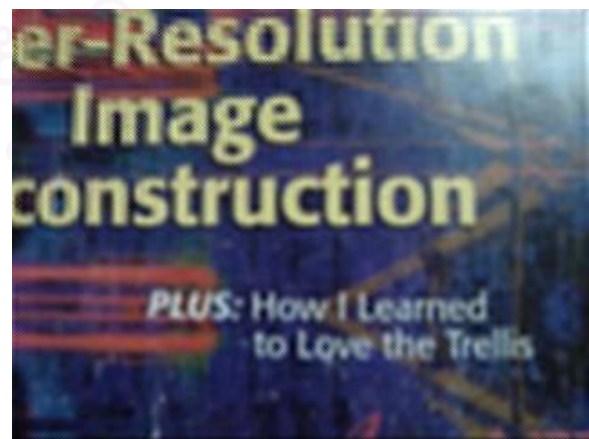


(ข) ภาพความละเอียดต่ำ



(ค) ภาพความละเอียดสูงโดยวิธีไบคิวบิก

PSNR = 22.30 dB



(ง) ภาพความละเอียดสูงโดยวิธีที่นำเสนอ

PSNR = 23.95 dB

รูปที่ 6-1 ผลการทดลองที่ 1

การทดลองที่ 2 เป็นการบันทึกภาพอาคาร 1 ในคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยจัดเป็นภาพที่มีการเคลื่อนที่แบบเลื่อนขนานทั้งหมดเช่นเดียวกับการทดลองที่ 1 แต่เกิดจากการหมุนกวาดกล้องจากขวาไปซ้าย และใช้ภาพความละเอียดต่ำเริ่มต้นและตัวประกอบการขยายเช่นเดียวกับการทดลองที่ 1 โดยมีค่าตัวแปรเสริมดังนี้ $a = 0.003$, $b = 2$ ขนาดของเคอร์เนลความพรมัวเท่ากับ 5×5 โดยที่ $s = 5$ และจำนวนรอบการทำซ้ำเท่ากับ 30 รอบโดยมีผลการทดลองดังแสดงในรูป 6-2 โดยจะสังเกตได้ว่า บริเวณขอบวัตถุ เช่นหน้าต่าง ประตู มีความคมชัดและความเปรียบต่างมากขึ้น



(ก) ภาพความละเอียดสูงต้นแบบ



(ข) ภาพความละเอียดต่ำ



(ค) ภาพความละเอียดสูงโดยวิธีไบคิวบิก

$PSNR = 22.60 \text{ dB}$



(ง) ภาพความละเอียดสูงโดยวิธีที่นำเสนอ

$PSNR = 24.40 \text{ dB}$

รูปที่ 6-2 ผลการทดลองที่ 2

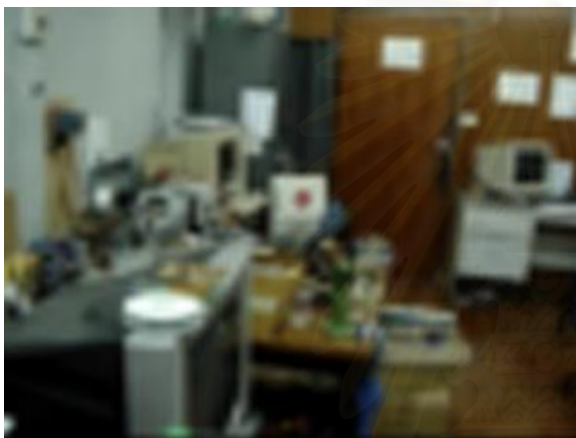
การทดลองที่ 3 เป็นการบันทึกภาพภายในห้องปฏิบัติการวิจัยและประยุกต์วงจรรวมโดยการหมุนกวาดกล้องไปรอบๆ ทำให้ได้ภาพที่มีการเคลื่อนไหวแบบพ้องรูปเชิงระนาบ โดยมีภาพความละเอียดต่ำเริ่มต้นและตัวประกอบการขยายเช่นเดียวกับการทดลองที่ 1 โดยมีค่าตัวแปรเสริมดังนี้ $a = 0.003$, $b = 2$ ขนาดของเคอร์เนลความพรมัวเท่ากับ 5×5 โดยที่ $s = 5$ และจำนวนรอบการทำซ้ำเท่ากับ 50 รอบโดยมีผลการทดลองดังแสดงในรูป 6-3



(ก) ภาพความละเอียดสูงต้นแบบ

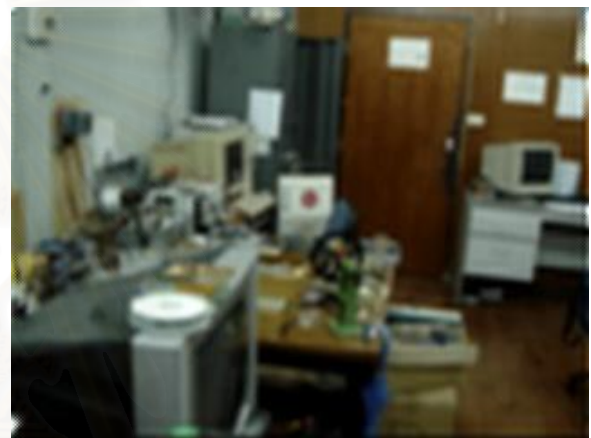


(ข) ภาพความละเอียดต่ำ



(ค) ภาพความละเอียดสูงโดยวิธีไปควิบิค

PSNR = 26.16 dB



(ง) ภาพความละเอียดสูงโดยวิธีที่นำเสนอ

PSNR = 28.19 dB

รูปที่ 6-3 ผลการทดลองที่ 3

จากผลการทดลองจะเห็นว่า สามารถแยกแยะวัตถุต่างๆ ในภาพลัทธิ เช่น มิเตอร์วัดค่าไฟฟ้า สายไฟ กิ่งของไม้คอมแพ็คติสก็์ได้ดีขึ้น และขอบวัตถุ เช่น ขอบป้ายประกาศ โทรทัศน์ ขอบประตู และ กิ่งของกระดาษ มีความคมชัดมากขึ้น

การทดลองสุดท้าย เป็นการทดสอบโปรแกรมที่นำเสนอกับภาพบุคคล ซึ่งเป็นภาพของผู้เขียนที่ ถูกบันทึกในระยะเวลาใกล้ภายในห้องปฏิบัติการวิจัยและประยุกต์วงจรรวมโดยการหมุนกวาดกล้องจากขวา ไปซ้ายอย่างช้าๆ ซึ่งบุคคลและฉากหลังอยู่นิ่ง โดยมีความละเอียดของภาพความละเอียดต่ำเริ่มต้นและ ตัวประกอบการขยายเช่นเดียวกับการทดลองที่ 1 โดยมีค่าตัวแปรเสริมดังนี้ $a = 0.003$, $b = 2$ ขนาด ของเคอร์เนลความพรมัวเท่ากับ 5×5 โดยที่ $s = 5$ และจำนวนรอบการทำซ้ำเท่ากับ 50 รอบโดยมีผล การทดลองดังแสดงในรูป 6-4



(ก) ภาพความละเอียดสูงต้นแบบ



(ข) ภาพความละเอียดต่ำ



(ค) ภาพความละเอียดสูงโดยวิธีไบคิวบิก

PSNR = 27.10 dB



(ง) ภาพความละเอียดสูงโดยวิธีที่นำเสนอ

PSNR = 29.17 dB

รูปที่ 6-4 ผลการทดลองที่ 4

จากการทดลองกับแฟ้มภาพที่มีการเคลื่อนไหวแบบเลื่อนขนานทั้ง 4 การทดลองข้างต้นสามารถสรุปเป็นตารางได้ดังนี้

ตารางที่ 6-1 สรุปผลการทดลองการเพิ่มภาพที่มีการเคลื่อนไหวแบบเลื่อนขนานทั้งหมด

ชุดภาพ	No. of Iteration	Time (sec.)	Bicubic Resize PSNR (dB)	Superresolution PSNR (dB)
IEEE	26	11.8	22.30	23.81
Red Castle	35	16.8	22.60	24.40
Inside IDAR	51	25.0	26.16	28.19
IDAR Man	18	7.9	27.10	29.17

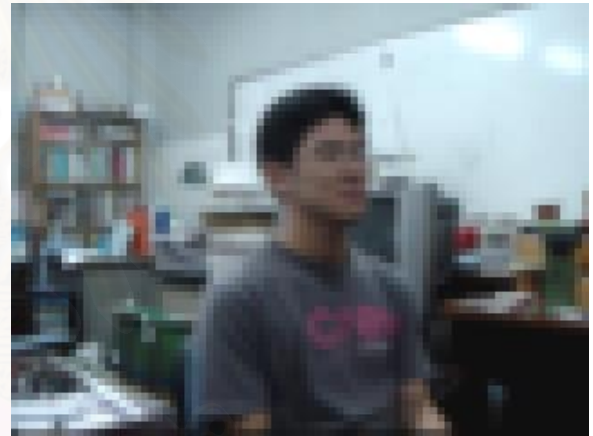
6.2.2 กรณีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่

โดยทั่วไป ภาพเคลื่อนไหวที่ได้จากกล้องวิดีโอที่สนั้นส่วนใหญ่จะมีการเคลื่อนที่ของวัตถุในภาพเป็นแบบเลื่อนขนานเฉพาะที่ เนื่องจากวัตถุที่ถูกบันทึกโดยส่วนใหญ่มักจะเป็นบุคคลหรือยานพาหนะซึ่งมีการเคลื่อนไหวอยู่ตลอดเวลา จึงแทบไม่มีโอกาสเกิดการเคลื่อนไหวแบบเลื่อนขนานทั้งหมดในภาพ การทดสอบกับแฟ้มภาพซึ่งมีการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่นี้ จึงเป็นการทดสอบว่าโปรแกรมที่นำเสนอนี้สามารถนำไปใช้งานจริงในระบบเฝ้าระวังได้ดีเพียงใด

ในการทดลองที่ 5 นี้เป็นการบันทึกภาพผู้เขียนในระยะใกล้ภายในห้องปฏิบัติการวิจัยและประยุกต์วงจรรวมโดยฉากหลังอยู่หนึ่งและตัวบุคคลเคลื่อนไหวแบบเลื่อนขนาน โดยมีค่าตัวแปรเสริมดังนี้ $a = 0.003$, $b = 2$ ขนาดของคอร์เนลความพรั้มเท่ากับ 5×5 โดยที่ $s = 5$ และจำนวนรอบการทำซ้ำเท่ากับ 50 รอบโดยมีผลการทดลองดังแสดงในรูป 6-5



(ก) ภาพความละเอียดสูงต้นแบบ



(ข) ภาพความละเอียดสูงโดยวิธีข้างเคียงใกล้เคียงสุด



(ค) ภาพความละเอียดสูงโดยวิธีไบควิบิค



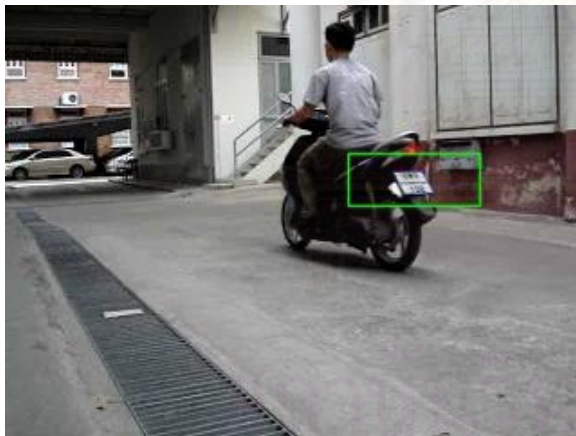
(ง) ภาพความละเอียดสูงโดยวิธีที่นำเสนอ

รูปที่ 6-5 ผลการทดลองที่ 5

ในการทดลองห้วงนี้จะเห็นได้ว่าไม่การแสดงค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด ซึ่งเป็นการวัดผลการทดสอบแบบอัตวิสัย (Subjective Measurement) โดยเป็นการวัดผลโดยความรู้สึกของผู้สังเกตที่มีต่อภาพลัพธ์ที่ได้ ซึ่งต่างจากการทดลองในหัวข้อแรกที่มีค่าดัชนีระบุสำหรับเปรียบเทียบ ซึ่งจัดเป็นการวัดแบบวัตถุวิสัย (Objective Measurement) เนื่องจากในการทดลองกับการเคลื่อนที่แบบเลื่อนขนานเฉพาะที่นี้ ไม่สามารถหาภาพความละเอียดสูงต้นแบบมาเปรียบเทียบได้เนื่องจากภาพลัพธ์ที่ผ่านการขยายรายละเอียดมานั้น ในบริเวณนอกเหนือจากบริเวณที่สนใจในภาพ การซ้อนทับภาพจากภาพความละเอียดต่ำแต่ละภาพลงบนภาพอ้างอิงจะเป็นไปอย่างสุ่ม และไม่ได้รับการป้อนกลับที่ถูกต้อง ทำให้ภาพในบริเวณรอบๆ วัตถุที่สนใจผิดเพี้ยนไป เมื่อนำมาคำนวณค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด ทำให้ได้ค่าต่ำมากเนื่องจากมีค่าความคลาดเคลื่อนในบริเวณรอบๆ วัตถุที่สนใจมีค่าสูง

6.3 การทดสอบโปรแกรมกับแฟ้มภาพทั่วไป

ในหัวข้อนี้จะเป็นการทดสอบโปรแกรมขยายรายละเอียดภาพกับแฟ้มภาพวิดีโอทั้งหมด ซึ่งมีความละเอียดเชิงพื้นที่ในบริเวณที่สนใจค่อนข้างต่ำ และไม่ทราบค่าของภาพความละเอียดสูงสำหรับเปรียบเทียบ โดยในการทดลองที่ 6 เป็นแฟ้มภาพเคลื่อนไหวของรถจักรยานยนต์ซึ่งกำลังเคลื่อนที่ โดยเลือกขยายรายละเอียดบริเวณแผ่นป้ายทะเบียน โดยมีผลการทดลองดังแสดงในรูปที่ 6-6



(ก) ภาพปกติและบริเวณที่สนใจ



(ข) บริเวณที่สนใจที่ผ่านการขยายด้วยวิธีทั่วไป



(ค) บริเวณที่สนใจที่ผ่านการขยายรายละเอียดด้วยโปรแกรมที่นำเสนอ

รูปที่ 6-6 ผลการทดลองที่ 6

ในรูปที่ 6-6 (ก) แสดงภาพทั้งเฟรมของแฟ้มภาพวิดีโอที่บันทึกตำแหน่งเฟรมที่ต้องการและบริเวณที่สนใจ รูปที่ 6-6 (ข) บนและล่างแสดงภาพบริเวณที่สนใจความละเอียดสูงที่ได้จากการขยายด้วยวิธีข้างเคียงใกล้สุด (Nearest Neighborhood) และวิธีไบควบิตตามลำดับ และรูปที่ 6-6 (ค) แสดงภาพบริเวณที่สนใจที่ได้จากโปรแกรมขยายรายละเอียดภาพที่นำเสนอ ซึ่งสามารถอ่านค่าตัวเลขบนแผ่นป้ายทะเบียนได้ชัดเจนมากกว่ารูปที่ 6-6 (ข)

การทดลองที่ 7 เป็นการทดสอบกับแฟ้มภาพของรถยนต์ที่กำลังเคลื่อนที่และเลือกขยายรายละเอียดในบริเวณแผ่นป้ายทะเบียนด้านหน้า โดยมีผลลัพธ์ดังแสดงในรูปที่ 6-7 และมีลำดับการนำเสนอภาพ เช่นเดียวกับการทดลองที่ 6



(ก) ภาพปกติและบริเวณที่สนใจ



(ข) บริเวณที่สนใจที่ผ่านการขยายด้วยวิธีทั่วไป



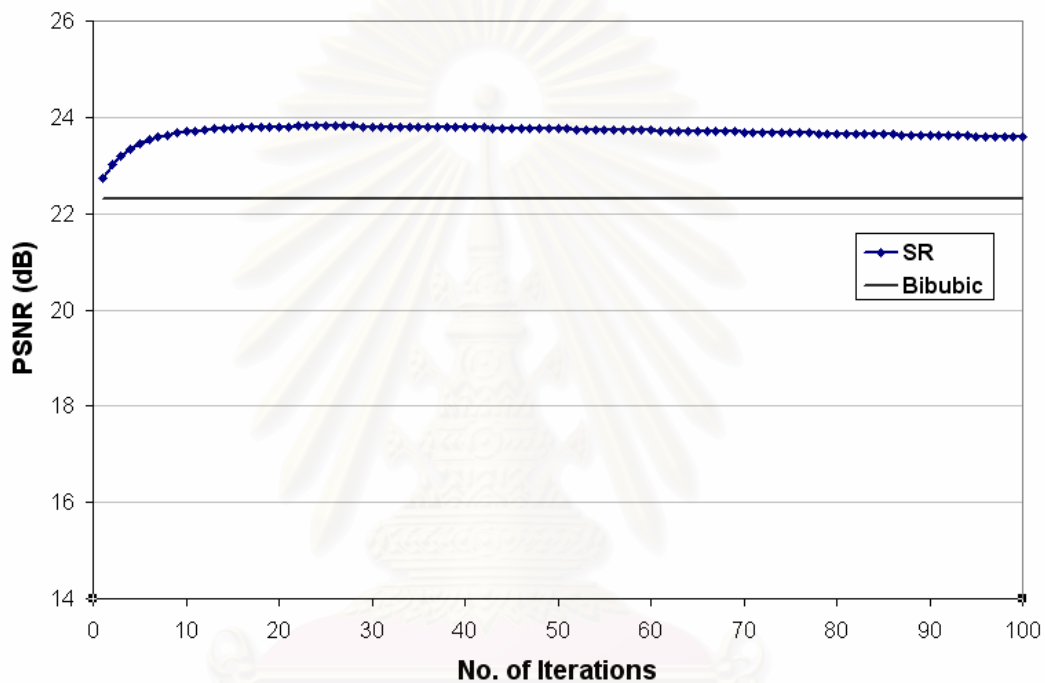
(ค) บริเวณที่สนใจที่ผ่านการขยายรายละเอียดด้วยโปรแกรมที่นำเสนอ

รูปที่ 6-7 ผลการทดลองที่ 7

6.4 เวลาที่ใช้ในการประมวลผล

จากผลการทดลองพบว่า สำหรับการขยายรายละเอียดภาพ จากภาพความละเอียดต่ำขนาด 80x60 จุดภาพ จำนวน 8 ภาพ เป็นภาพความละเอียดสูงขนาด 320x240 จุดภาพ จำนวน 1 ภาพ นั้นใช้เวลาประมวลผลสำหรับขั้นตอนซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติประมาณ 2 วินาที และใช้เวลาสำหรับขั้นตอนการทำซ้ำประมาณ 0.5 วินาทีต่อ 1 รอบการประมวลผล โดยหากเลือกค่าตัวประกอบการขยายที่ต่ำกว่านี้ หรือเลือกครอบตัดบริเวณที่สนใจที่เล็กกว่านี้จะทำให้ใช้เวลาในการประมวลผลน้อยลง

ในรูปที่ 6-8 แสดงถึงตัวอย่างกราฟความสัมพันธ์ระหว่างจำนวนรอบทำซ้ำกับค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุดสำหรับการทดลองที่ 1 ซึ่งคำนวณแบบทำซ้ำทั้งหมด 100 รอบ และใช้เวลาประมวลผลทั้งหมด 45.59 วินาที คิดเป็นค่าเฉลี่ย 0.45 วินาทีต่อ 1 รอบ โดยจากกราฟจะเห็นว่า ค่า PSNR ของภาพผลลัพธ์จะเริ่มมีค่าคงตัวที่รอบทำซ้ำประมาณรอบที่ 15-20 และมีค่าสูงสุดที่รอบทำซ้ำที่ 26 ซึ่งหากกำหนดให้โปรแกรมหยุดประมวลผลเมื่อค่า PSNR มีค่าสูงสุด จะใช้เวลาคำนวณประมาณ 11.8 วินาที ซึ่งการที่ค่า PSNR มีค่าลดลงจากจุดสูงสุดนี้ทางผู้วิจัยจะได้ทำการวิเคราะห์เพื่อปรับปรุงต่อไป



รูปที่ 6-8 ค่า PSNR ที่รอบทำซ้ำต่างๆ ในการทดลองที่ 1

6.5 สรุปท้ายบท

ในบทนี้ได้กล่าวถึงผลการทดสอบโปรแกรมขยายรายละเอียดภาพกับแฟ้มภาพวีดิทัศน์ที่มีการเคลื่อนไหวภายในภาพเป็นแบบเลื่อนขนาน โดยจากผลการทดลองในกรณีเลื่อนขนานทั้งหมดพบว่าภาพที่ผ่านการขยายรายละเอียดด้วยวิธีการที่นำเสนอมีค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด สูงกว่าภาพที่ผ่านการขยายรายละเอียดด้วยวิธีไปคิวบิคโดยเฉลี่ย 1.8 เดซิเบล และจากผลการทดลองในกรณีเลื่อนขนานทั้งหมดและการเลื่อนขนานเฉพาะที่กับแฟ้มภาพวีดิทัศน์สังเคราะห์และแฟ้มภาพวีดิทัศน์ทั่วไป พบว่าภาพผลลัพธ์ที่ได้มีรายละเอียดชัดเจนมากขึ้น และมีความเปรียบเทียบมากขึ้น ซึ่งแสดงให้เห็นว่าโปรแกรมขยายรายละเอียดภาพที่นำเสนอสามารถปรับปรุงคุณภาพของภาพให้ดีขึ้นได้ โดยใช้เวลาในการคำนวณต่อ 1 ภาพสีแบบ 8 บิตที่ขนาด 320x240 จุดภาพ ประมาณ 10-12 วินาทีสำหรับผลลัพธ์ที่ดีที่สุด

บทที่ 7

ข้อสรุปและข้อเสนอแนะ

7.1 ข้อสรุป

วิทยานิพนธ์นี้นำเสนอเทคนิคการขยายรายละเอียดภาพวีดิทัศน์ในบริเวณที่สนใจสำหรับระบบเฝ้าระวังในรูปแบบของโปรแกรมภาษา C++ โดยอาศัยชุดภาพความละเอียดต่ำจากแฟ้มภาพวีดิทัศน์ที่ต่อเนื่องกันจำนวนภาพ และให้ภาพผลลัพธ์เป็นภาพสีหรือขาวดำที่มีความละเอียดเชิงพื้นที่สูงขึ้นโดยมีตัวประกอบขยายในช่วง 2-4 เท่า

การทำงานของขั้นตอนวิธีที่เลือกใช้เริ่มจากการตั้งแบบจำลองปัญหาโดยมองว่าภาพความละเอียดต่ำนั้นเกิดจากภาพความละเอียดสูงที่ผ่านกระบวนการลดทอนคุณภาพ โดยอาศัยแบบจำลองปัญหาย้อนกลับ ร่วมกับการประมาณค่าตัวแปรเสริมและตัวดำเนินการต่างๆ ทำให้สามารถคำนวณหาคำตอบของสมการซึ่งเป็นภาพผลลัพธ์ความละเอียดสูงได้ ในขั้นตอนการคำนวณได้เลือกใช้ระเบียบวิธีเชิงกำหนดแบบคุ่มค่า ซึ่งเป็นการหาคำตอบของสมการ (ภาพความละเอียดสูง) ซึ่งเมื่อผ่านกระบวนการลดทอนคุณภาพแล้ว จะมีความแตกต่างจากภาพความละเอียดต่ำที่มีอยู่เดิมน้อยที่สุด

การขยายรายละเอียดภาพหรืออีกนัยหนึ่งคือปัญหาการสร้างกลับภาพความละเอียดสูงหนึ่งภาพจากภาพความละเอียดต่ำหลายๆ ภาพ โดยทั่วไปมักเป็นปัญหาแบบกำหนดขาดเนื่องจากจำนวนข้อมูลที่ได้จากภาพความละเอียดต่ำเข้ามามีน้อยกว่าจำนวนข้อมูลที่ต้องการในการสร้างกลับภาพความละเอียดสูง ทำให้เข้าข่ายเป็นปัญหาลักษณะเลว การแก้ไขสามารถทำได้โดยการใส่พจน์คุ่มค่าเพื่อปรับลักษณะสมบัติของคำตอบของสมการให้ดีขึ้น โดยพจน์คุ่มค่าที่เลือกใช้ในงานวิจัยนี้คือพจน์คุ่มค่าแบบทริโคโนฟซึ่งเป็นการจำกัดความราบเรียบของภาพผลลัพธ์ ซึ่งสามารถแก้ปัญหาจุดภาพขาดหายได้โดยพจน์ดังกล่าวจะทำหน้าที่ประมาณค่าจุดภาพให้ สิ่งที่ต้องคำนึงในการเลือกใช้พจน์คุ่มค่าคือขนาดของตัวคูณลากรางจ์ซึ่งทำหน้าที่ถ่วงดุลระหว่างพจน์รายละเอียดและพจน์คุ่มค่า

ขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งนับเป็นขั้นตอนที่มีความสำคัญมากในการขยายรายละเอียดภาพ ซึ่งการซ้อนทับภาพไม่ตรงตำแหน่งจะทำให้คุณภาพของภาพผลลัพธ์แยกลงอย่างมากในงานวิจัยนี้ได้เลือกใช้การซ้อนทับภาพโดยกำหนดตำแหน่งแบบอิงลักษณะเด่นร่วมกับการคัดกรองด้วยระเบียบวิธีเรนแซคซึ่งจะทำให้ได้จุดคุ่มค่าที่มีความแม่นยำถูกต้องสำหรับคำนวณเมตริกซ์แปลงภาพแบบทศมิติ โดยแฟ้มภาพวีดิทัศน์ที่สามารถนำมาขยายรายละเอียดได้ต้องมีการเคลื่อนไหวในภาพเป็นแบบเลื่อนขนานทั้งหมดหรือแบบเลื่อนขนานเฉพาะที่ โดยต้องไม่มีปัญหาการบังบังของวัตถุที่สนใจ และวัตถุดังกล่าวมีความคงรูป เพื่อให้การซ้อนทับภาพเป็นไปอย่างถูกต้อง

สำหรับขั้นตอนการคำนวณในงานวิจัยนี้ได้เลือกใช้ระเบียบวิธีทำซ้ำแบบลาดลงขั้นที่สุด ซึ่งเมื่อซ้อนทับภาพโดยกำหนดตำแหน่งอย่างถูกต้องแล้ว จะทำให้ภาพผลลัพธ์ที่คำนวณได้ในแต่ละรอบการทำซ้ำมีคุณภาพดีขึ้นเรื่อยๆ โดยสิ่งสำคัญที่ควรคำนึงในขั้นตอนนี้ได้แก่การเลือกค่าระยะก้าว เนื่องจากหากระยะก้าวมีขนาดใหญ่จะทำให้ลู่อู่เข้าสู่คำตอบได้รวดเร็วแต่อาจเกิดปัญหาเสถียรภาพของคำตอบ

ในงานวิจัยนี้ได้เลือกโปรแกรม Microsoft Visual Studio 2005 สำหรับออกแบบโปรแกรมร่วมกับไลบรารีโอเพนซีวี ซึ่งภายในบรรจุฟังก์ชันที่อำนวยความสะดวกต่อการประมวลผลภาพ โดยบางฟังก์ชันที่ใช้ในการออกแบบโปรแกรมนี้นักเขียนขึ้นมาใหม่เพื่อให้ครอบคลุมความต้องการของขั้นตอนวิธีที่เลือกใช้

ผลการทดสอบโปรแกรมขยายรายละเอียดภาพแสดงให้เห็นว่าภาพที่ผ่านการขยายรายละเอียดด้วยวิธีที่นำเสนอมีคุณภาพดีกว่าภาพที่ผ่านการขยายรายละเอียดด้วยวิธีข้างเคียงใกล้สุดและวิธีไบคิวบิก โดยมีค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุดโดยเฉลี่ย 5.5 เดซิเบลและ 1.8 เดซิเบลตามลำดับ โดยใช้เวลาประมวลผลเฉลี่ย 0.5 วินาทีต่อ 1 รอบการทำซ้ำ และคำตอบจะเริ่มลู่อู่เข้าเมื่อผ่านการประมวลผลไปประมาณ 15-20 รอบ

7.2 ข้อเสนอแนะ

เพื่อให้โปรแกรมขยายรายละเอียดภาพมีประสิทธิภาพมากยิ่งขึ้น และสามารถนำไปใช้งานได้จริง จึงได้เสนอจุดที่ควรปรับปรุงและข้อเสนอแนะสำหรับการพัฒนาโปรแกรมในอนาคต ดังต่อไปนี้

1. การเคลื่อนที่ของวัตถุที่สนใจในภาพจำกัดอยู่เพียงการเคลื่อนที่แบบเลื่อนขนานและเป็นแบบคงรูปเท่านั้น ทำให้ไม่สามารถนำไปใช้กับวัตถุที่มีการเคลื่อนที่แบบเกิดการบิดบัง หรือเป็นวัตถุไม่คงรูป เช่นกรณีที่ต้องการติดตามภาพใบหน้าคนได้ ซึ่งปัญหานี้อาจแก้ไขได้บางส่วนโดยการเพิ่มจำนวนจุดควบคุมในขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่ง ร่วมกับแบบจำลองเรขาคณิตที่ครอบคลุมมากกว่าการแปลงทัศนมิติ เช่น การแปลงแบบไม่เชิงเส้น หรืออาศัยแบบจำลอง 3 มิติของใบหน้าคนโดยเฉพาะ รวมถึงการออกแบบให้โปรแกรมสามารถรู้จำโดยอาศัยฐานข้อมูลเกี่ยวกับใบหน้าบุคคลที่มีอยู่ก่อนได้
2. เวลาที่ใช้ในการประมวลผลโดยเฉลี่ยอยู่ที่ประมาณ 7-10 วินาทีต่อ 1 ภาพความละเอียดสูง ซึ่งสามารถทำให้ลดลงได้โดยการออกแบบขั้นตอนวิธีใหม่ เช่น ซ้อนทับภาพโดยกำหนดตำแหน่งแบบครั้งเดียวเสร็จ (One Step) จากนั้นจึงประมาณค่าในช่วงและกำจัดความพัวมัวด้วยการคำนวณแบบทำซ้ำ ตามที่แนะนำใน [2] ซึ่งทำให้ลดระยะเวลาในการคำนวณลงไปได้มาก

3. ค่าตัวแปรเสริมต่างๆ เช่น ตัวดำเนินการพหุคูณ ตัวคูณลากรางจ์ และระยะก้าวในการทำซ้ำยังเป็นแบบที่ผู้ใช้กำหนดเอง ทำให้ไม่สะดวกในการนำไปใช้งานจริง ซึ่งปัญหานี้อาจแก้ไขได้โดยการใช้การคูณประสานกลับแบบบอด (Blind Deconvolution) สำหรับประมาณความพหุคูณ ใช้การทำให้ถูกแบบไขว้ทั่วไป (General Cross Validation) สำหรับคำนวณค่าตัวคูณลากรางจ์ที่เหมาะสมที่สุด และเลือกใช้ระยะก้าวแบบปรับตัวได้ (Adaptive Step Size) เป็นต้น
4. โปรแกรมที่นำเสนอในงานวิจัยนี้ใช้ได้กับแฟ้มภาพวีดิทัศน์ชนิด AVI เท่านั้น เนื่องจากเป็นแฟ้มภาพเคลื่อนไหวชนิดเดียวที่ไลบรารีโอเพ่นซีวีรองรับ การแก้ไขที่เป็นไปได้คือการศึกษาค้นคว้าในไลบรารีโอเพ่นซีวีแล้วนำมาเขียนเป็นฟังก์ชันใหม่ที่รองรับแฟ้มภาพเคลื่อนไหวชนิดอื่นๆ
5. ออกแบบโปรแกรมให้มีความสามารถตรวจจับวัตถุเคลื่อนไหวในภาพและทำการขยายรายละเอียดได้เอง ทำให้ผู้ใช้ไม่จำเป็นต้องเลือกบริเวณที่สนใจสำหรับขยายรายละเอียดภาพ เนื่องจากในบางครั้งหากผู้ใช้กำหนดกรอบบริเวณที่สนใจได้ไม่เหมาะสม อาจทำให้ขั้นตอนการซ้อนทับภาพผิดพลาดได้ ซึ่งอาจปรับปรุงได้โดยใช้การทำนายทิศทางการเคลื่อนที่ของวัตถุที่สนใจเพื่อกำหนดบริเวณครอบตัดที่เหมาะสมในเฟรมข้างเคียงได้

รายการอ้างอิง

- [1] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-Resolution Image Reconstruction: A Technical Overview. IEEE Signal Processing Magazine. (May 2003): 21-36.
- [2] Sina Farsiu, M. Dirk Robinson, Michael Elad, and Peyman Milanfar. Fast and Robust Multiframe Super Resolution. IEEE Transactions on Image Processing. 13, 10 (October 2004): 1327-1344.
- [3] Andrew J. Patti, M. Ibrahim Sezan, and A. Murat Tekalp. Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time. IEEE Transactions on Image Processing. 6, 8 (August 1997): 1064-1076.
- [4] Deepesh Jain. Superresolution using Papoulis-Gerchberg Algorithm. EE329J Digital Video Processing. Stanford University, Stanford, CA.
- [5] David Capel and Andrew Zisserman. Computer Vision: Applied to Super Resolution. IEEE Signal Processing Magazine. (May 2003): 75-86.
- [6] Timothy F. Gee, Thomas P. Karnowski, and Kenneth W. Tobin. Multiframe Combination and Blur Deconvolution of Video Data. Image and Video Communications and Processing. 3974 (April 2000): 788-795.
- [7] Nhat Nguyen, Peyman, and Gene Golub. A Computationally Efficient Superresolution Image Reconstruction Algorithm. IEEE Transactions on Image Processing. 10, 4 (April 2001): 573-583.
- [8] Barbara Zitova and Jan Flusser. Image Registration Methods: A Survey. Image and Vision Computing. 21 (2003): 977-1000.
- [9] Wikipedia. Well-posed Problem. Numerical Analysis: Partial Differential Equations. (May 2007).
- [10] Wikipedia. YCbCr. Color Space. (September 2007).
- [11] Gady Agam, Introduction to Programming with OpenCV. Department of Computer Science Illinois Institute of Technology. (January 2006).

บรรณานุกรม

- [1] Lei Liang and Yuanchang Xu. Adaptive Landweber Method to Deblur Images. IEEE Signal Processing Letters. 10, 5 (May 2003): 129-132.
- [2] Michael Elad and Yacov Hel-Or. A Fast Super-Resolution Reconstruction Algorithm for Pure Translational Motion and Common Space-Invariant Blur. IEEE Transactions on Image Processing. 10, 8 (August 2001): 1187-1193.
- [3] Michael Elad and Arie Feuer. Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images. IEEE Transaction on Image Processing. 6, 12 (December 1997): 1646-1658.
- [4] Celine Mancas-Thillou and Majid Mirmehdi. An Introduction to Super-Resolution Text. (2006).
- [5] Richard R. Schultz, and Robert L. Stevenson. Extraction of High-Resolution Frames from Video Sequences. IEEE Transactions on Image Processing, 5, 6 (June 1996): 996-1011.
- [6] Jan Biemond, Reginal L. Lagendijk, and Russell M. Mersereau. Iterative Methods for Image Deblurring. Proceedings of the IEEE. 78, 5 (May 1990): 856-883.
- [7] Deepa Kundur and Dimitrios Hatzinakos. Blind Image Deconvolution. IEEE Signal Processing Magazine (May 1996): 43-64.
- [8] Nhat Nguyen. Peyman Milanfar, and Gene Golub. Efficient Generalized Cross-Validation with Applications to Parametric Image Restoration and Resolution Enhancement. IEEE Transaction on Image Processing, 10, 9 (September 2001): 1299-1308.
- [9] Paulo Jorge S. G. Ferreira. Interpolation and the Discrete Papoulis-Gerchberg Algorithm. IEEE Transactions on Signal Processing. 42, 10 (October 1994): 2596-2606.
- [10] Shinji Umeyama. Least-square Estimation of Transformation Parameters Between Two Point Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 13, 4 (April 1991): 376-380.

- [11] Mario A. T. Figueiredo and Robert D. Nowak, An EM Algorithm for Wavelet-Based Image Restoration. IEEE Transactions on Image Processing. 12, 8 (August 2003): 906-916.
- [12] Xin Li, and Michael T. Orchard. New Edge-Directed Interpolation. IEEE Transactions on Image Processing. 10, 10 (October 2001): 1521-1527.
- [13] Subhasis Chaudhuri, and Dipti Rani Taur. High-Resolution Slow-Motion Sequencing: How to generate a slow-motion sequence from a bit stream. IEEE Signal Processing Magazine. (March 2005): 16-24.
- [14] Fred L. Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence. 11, 6 (June 1989): 567-585.
- [15] Weiming Hu, Tieniu Tan, Liang Wang, and Steve Maybank. A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews. 34, 3 (August 2004): 334-352.
- [16] จันทริจรีรา สิ้นทนะโยธิน. การประมวลผลภาพด้วย C++ Builder: ตอนที่9 การซ้อนทับภาพโดยกำหนดตำแหน่ง. สาร NECTEC. (มีนาคม - เมษายน 2547): 49-55.



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก.
บทความที่ได้รับการตีพิมพ์
ในการประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 30



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การขยายรายละเอียดภาพวิดีโอพร้อมการซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ

Video Image Super-resolution with Auto-registration

โกศล ปัญญาโสภณเลิศ และ สุรีย์ พุ่มรินทร์

ห้องปฏิบัติการวิจัยออกแบบและประยุกต์วงจรรวม

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ถ. พญาไท แขวงวังใหม่ เขตปทุมวัน กรุงเทพฯ 10330

โทร. 089 665 3850, E-mail: link82ee@hotmail.com, suree.p@chula.ac.th

บทคัดย่อ

บทความนี้นำเสนอโปรแกรมการขยายรายละเอียดภาพจากเพิ่มภาพวิดีโอซึ่งเขียนด้วยภาษา C++ ร่วมกับไลบรารีโอเพ่นซีวีซึ่งมีจุดเด่นตรงที่สามารถซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติได้โดยใช้ระเบียบวิธีแรนแซคร่วมกับการแปลงภาพแบบทัศนมิติ และใช้ระเบียบวิธีทำซ้ำในปริภูมิเชิงกำหนดแบบควมค่าในปริภูมิ L_2 สำหรับขั้นตอนประมาณค่าในช่วงและขั้นตอนกำจัดความพร่ามัว โดยสามารถเลือกตัวประกอบการขยายได้ในช่วง 2-4 เท่า

คำสำคัญ: ไลบรารีโอเพ่นซีวี, ซ้อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ, แรนแซค, การแปลงทัศนมิติ, ระเบียบวิธีทำซ้ำเชิงกำหนดแบบควมค่า, ปริภูมิ L_2

Abstract

This paper proposes the Video Image Super-resolution Software in C++ language using OpenCV library. Main advantage of the software is that it can process registration step automatically using RANSAC method for perspective image transformation. The regularized iterative deterministic method in L_2 -space is used for interpolation and deblurring steps. The enhancement factor can be adjusted in the range of 2-4

Keywords: OpenCV Library, Auto-registration, RANSAC, Perspective Transformation, Regularized Iterative Deterministic Method, L_2 -Space.

1. คำนำ

การขยายรายละเอียดภาพเป็นการนำข้อมูลจากภาพของวัตถุหรือฉากเหตุการณ์เดียวกันซึ่งบันทึกในขณะเวลาต่างๆ กันหรือจากกล้องหลายตัว มาบรรจุอยู่ในภาพเพียงภาพเดียวซึ่งมีความละเอียดเชิงพื้นที่สูงกว่า ทำให้ได้ภาพที่มีรายละเอียดมากขึ้น เมื่อนำหลักการนี้ไปใช้กับเพิ่มภาพวิดีโอซึ่งเปรียบเสมือนชุดของภาพนิ่งในช่วงเวลาต่อเนื่องกัน จึงสามารถปรับปรุงรายละเอียดที่สำคัญภายในภาพได้ การขยายรายละเอียด

ภาพนี้นับเป็นทางเลือกหนึ่งซึ่งสามารถแก้ไขปัญหาละเอียดเชิงพื้นที่ไม่เพียงพอได้โดยไม่ต้องปรับเปลี่ยนอุปกรณ์ในระบบ

ในงานวิจัยที่ผ่านมาเช่น [1] มักกล่าวถึงการขยายรายละเอียดภาพในกรณีที่การเคลื่อนที่สัมพันธ์ระหว่างเฟรมอ้างอิงกับเฟรมข้างเคียงเป็นที่ทราบค่าแน่นอนอยู่ก่อนแล้ว ซึ่งทำให้ขั้นตอนการซ้อนทับภาพโดยกำหนดตำแหน่งเป็นแบบสมบูรณ์ (Perfect Image Registration) หรือใน [2-3] ซึ่งทำการทดลองกับชุดของภาพนิ่งจำนวนมากที่ถูกบันทึกโดยการเปลี่ยนมุมกล้องไปที่ละน้อยซึ่งจัดเป็นกรณีกำหนดเกิน (Over-determined Case) จึงไม่สามารถนำไปประยุกต์ใช้กับเพิ่มภาพวิดีโอซึ่งมีภาพขาเข้าจำนวนมากจำกัดและจัดเป็นปัญหาที่กำหนดขาด (Under-determined) และมีการเคลื่อนที่สัมพันธ์ที่ไม่ทราบค่าได้ จึงจำเป็นต้องเลือกใช้ระเบียบวิธีและการประมาณค่าตัวแปรเสริมต่างๆ ที่เหมาะสม

โดยการเขียนโปรแกรมภาษา C++ ร่วมกับฟังก์ชันในไลบรารีโอเพ่นซีวี [4] ในการออกแบบโปรแกรมเพื่อให้สามารถซ้อนทับภาพโดยกำหนดตำแหน่งแบบอัตโนมัติร่วมกับระเบียบวิธีคัดกรองแรนแซค (RANSAC) [2] ทำให้สามารถลดความคลาดเคลื่อนที่เกิดจากการซ้อนทับภาพผิดตำแหน่ง และนำไปโปรแกรมไปใช้งานกับเพิ่มภาพวิดีโอได้อย่างน่าเชื่อถือมากยิ่งขึ้น โดยวิธีการที่นำเสนอนี้ครอบคลุมการแปลงรูปแบบพ้องรูปเชิงระนาบ (Planar Homography Transformation) [2] ซึ่งเป็นการแปลงในกรณีที่ไม่มีปัญหาการบังตัวเอง (Self-occlusion) และความไม่คงรูปของวัตถุ (Non-rigidity) เข้ามาเกี่ยวข้อง

2. แบบจำลองของปัญหา

โดยทั่วไปนิยมตั้งแบบจำลองของปัญหาการขยายรายละเอียดภาพด้วยแบบจำลองการบูรณะภาพไปข้างหน้า (Forward Image Restoration Model) [5] ตามสมการ (1)

$$y_k = H_k * x + h_k, \quad k = 1, 2, 3, \dots, N \quad (1)$$

โดยที่ x เป็นภาพความละเอียดสูง คุณประสาน (Convolute) กับเมตริกซ์ลดทอนคุณภาพ H_k h_k เป็นสัญญาณรบกวน และ y_k เป็นภาพความละเอียดต่ำ โดยที่ k คือดัชนีระนาบลำดับของภาพจากทั้งหมด N ภาพ เมตริกซ์ H_k เป็นผลคูณประสานของตัวดำเนินการลดทอนคุณภาพ

ซึ่งได้แก่ การบิดโค้ง (Warping W_k), ความพร่ามัว (Blurring B_k), และการซັกรุ่นขาลง (Down-sampling D_k) ตามสมการ (2)

$$H_k = D_k * B_k * W_k \quad (2)$$

ภาพความละเอียดสูง x ประมาณได้จากค่า \hat{x} ที่ทำให้พจน์ความคลาดเคลื่อน (พจน์แรกทางขวามือ) ของสมการ (3) มีค่าน้อยที่สุด

$$x = \underset{\hat{x}}{\text{ArgMin}} \sum_{k=1}^N \|y_k - H_k \hat{x}\|_2^2 + a \|\hat{x}\|_2^2 \quad (3)$$

โดยพจน์ที่สองทางขวามือของสมการคือพจน์กุ่มค่า (Regularization Term) ซึ่งใส่เพิ่มเติมเพื่อปรับแก้สถานะเลว (Ill-conditioned) ซึ่งมักพบในปัญหาการบูรณะภาพอันเนื่องมาจากตัวดำเนินการพร่ามัว และแก้ปัญหาข้อมูลขาดหายอันเนื่องมาจากจำนวนภาพขาเข้าไม่เพียงพอ a เป็นตัวคูณลากรางจ์ (Lagrange Multiplier) ซึ่งทำหน้าที่ถ่วงดุลระหว่างพจน์ความคลาดเคลื่อนและพจน์กุ่มค่า โดยการคำนวณทั้งหมดกระทำในปริภูมิ l_2

3. การซ่อนทับภาพโดยกำหนดตำแหน่งอัตโนมัติ

ในขั้นตอนนี้ได้เลือกระเบียบวิธีอิงลักษณะเด่น (Feature-based Method) เพื่อคำนวณหาเมตริกซ์แปลงภาพแบบทัศนมิติ (Perspective Transformation Matrix) ควบคู่กับการคัดกรองด้วยระเบียบวิธีแรนแซค (RANSAC) [2] ซึ่งทำให้ได้เมตริกซ์แปลงภาพที่มีความแม่นยำในระดับต่ำกว่าจุดภาพ (Sub-pixel) โดยมีฟังก์ชันแสดงขั้นตอนดังในรูปที่ 1

การคำนวณเริ่มจากค้นหาตำแหน่งของลักษณะเด่นหรือจุดควบคุม (Control Point) ภายในบริเวณที่สนใจในเฟรมอ้างอิง เช่น เส้นขอบหรือมุมของวัตถุ จากนั้นจึงใช้ฟังก์ชันติดตามการไหลเชิงแสง (Optical Flow Tracking) เพื่อคำนวณหาตำแหน่งของจุดควบคุมที่เปลี่ยนไปในเฟรมข้างเคียง

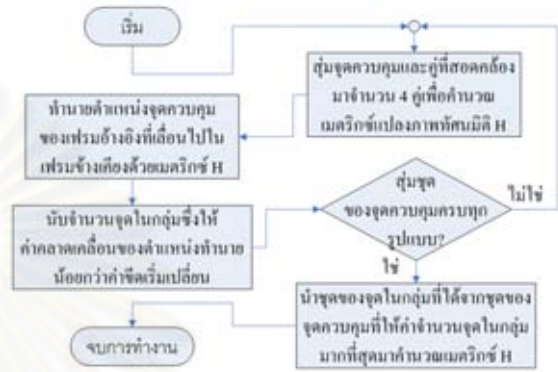
จากนั้นจึงสุ่มเลือกจุดควบคุมในเฟรมอ้างอิงพร้อมกับจุดควบคุมในเฟรมข้างเคียงที่สอดคล้องกันจำนวน 4 คู่มาคำนวณหาเมตริกซ์แปลงภาพแบบทัศนมิติโดยอาศัยความสัมพันธ์การแปลงที่กัก $P(x, y) \Rightarrow P(u, v)$ ตามสมการ (4)

$$P(u, v) = P \left(\frac{h_{11} \cdot x + h_{12} \cdot y + h_{13}}{h_{31} \cdot x + h_{32} \cdot y + 1}, \frac{h_{21} \cdot x + h_{22} \cdot y + h_{23}}{h_{31} \cdot x + h_{32} \cdot y + 1} \right) \quad (4)$$

โดย h_{ij} คือสมาชิกในแถวที่ i และสดมภ์ที่ j ของเมตริกซ์แปลงภาพ H และ $h_{33} = 1$ เมื่อได้เมตริกซ์แปลงภาพแล้วจึงนำมาคำนวณตำแหน่งโดยประมาณของจุดควบคุมในภาพข้างเคียง และบันทึกค่าผิดพลาด (ระยะห่าง) ระหว่างจุดควบคุมจริงกับจุดควบคุมที่คำนวณจากเมตริกซ์แปลงภาพในภาพข้างเคียง โดยหากค่าผิดพลาดมีค่าต่ำกว่าค่าขีดเริ่มเปลี่ยนที่กำหนดไว้จะถือว่าเป็นจุดในกลุ่ม (Inlier) แต่หากค่าผิดพลาด

มีค่าสูงกว่าจะถือว่าเป็นจุดนอกกลุ่ม (Outlier) จากนั้นบันทึกผลรวมจำนวนจุดในกลุ่มที่ได้สำหรับแต่ละชุดของจุดควบคุมที่สุ่มมา

ขั้นต่อไปจึงสุ่มชุดของจุดควบคุมชุดใหม่และทำซ้ำตามขั้นตอนที่เดียวจนครบทุกรูปแบบของชุดจุดควบคุม นำชุดของจุดควบคุมที่มีจำนวนจุดในกลุ่มมากที่สุด และจุดในกลุ่มและคู่ที่สอดคล้องทั้งหมดมาคำนวณหาเมตริกซ์แปลงภาพอีกครั้งด้วยระเบียบวิธีกำลังสองน้อยสุด



รูปที่ 1 ฟังก์ชันของระเบียบวิธีคัดกรองแรนแซค

ด้วยวิธีการในขั้นตอนนี้จะทำให้ได้เมตริกซ์แปลงภาพที่มีความถูกต้องมากยิ่งขึ้น และความผิดพลาดอันเกิดจากการซ่อนทับภาพผิดพลาด (Misalignment) ลดลง จึงสามารถนำไปปรับปรุงและพัฒนาเพื่อนำไปใช้กับเพิ่มภาพวิดีโอที่สนับทั้งแบบเชื่อมต่อตรง (Online) และไม่เชื่อมต่อตรง (Offline) ได้โดยมีความน่าเชื่อถือมากขึ้น

4. การประมาณค่าในช่วงและการกำจัดความพร่ามัว

ในขั้นตอนนี้ได้เลือกใช้ระเบียบวิธีทำซ้ำเชิงกำหนดแบบกุ่มค่า [6] ซึ่งเมื่อนำไปใช้กับสมการ (3) แล้ว จะสามารถหาผลเฉลยในรูปแบบปิดสำหรับการทำซ้ำในแต่ละรอบได้เป็น

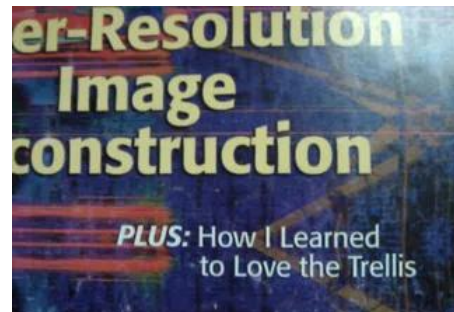
$$x_{n+1} = x_n + b \left[\sum_{k=1}^N H_k^T (y_k - H_k x_n) - a C^T C x_n \right] \quad (5)$$

โดยภาพเริ่มต้น x_0 อาจเป็นภาพเปล่า (Null Image) หรือเป็นภาพความละเอียดต่ำที่ถูกขยายด้วยการประมาณค่าในช่วงวิธีใดวิธีหนึ่งก็ได้ H_k มีนิยามตามสมการ (2) โดยสมมุติให้ใช้ตัวดำเนินการซັกรุ่นขาลงและตัวดำเนินการพร่ามัวร่วมกันสำหรับทุกค่า k กล่าวคือ $\forall k, D_k = D$ และ $B_k = B$ [6] โดยที่ b คือระยะก้าว (Step Size) ในแต่ละรอบการทำซ้ำ สำหรับ C ในพจน์กุ่มค่าได้เลือกใช้ตัวดำเนินการลาปลาซขนาด 3×3 ซึ่งเป็นตัวดำเนินการตรวจจับขอบ (Edge Detector) เพื่อคุมให้คำตอบมีความราบเรียบโดยมีนิยามตามสมการ (6)

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (6)$$



รูปที่ 2-ก ภาพความละเอียดสูงต้นแบบ



รูปที่ 3-ก ภาพความละเอียดสูงต้นแบบ



รูปที่ 2-ข ภาพความละเอียดต่ำที่ได้จากการลดขนาดภาพในรูป 2-ก



รูปที่ 3-ข ภาพความละเอียดต่ำที่ได้จากการลดขนาดภาพในรูป 3-ก



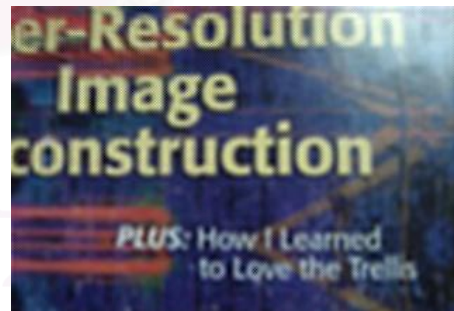
รูปที่ 2-ค ภาพความละเอียดสูงโดยวิธีไบคิวบิก: PSNR = 22.60 dB



รูปที่ 3-ค ภาพความละเอียดสูงโดยวิธีไบคิวบิก: PSNR = 22.30 dB



รูปที่ 2-ง ภาพความละเอียดสูงโดยวิธีการที่นำเสนอ: PSNR = 24.25 dB



รูปที่ 3-ง ภาพความละเอียดสูงโดยวิธีการที่นำเสนอ: PSNR = 23.81 dB

5. ผลการทดสอบ

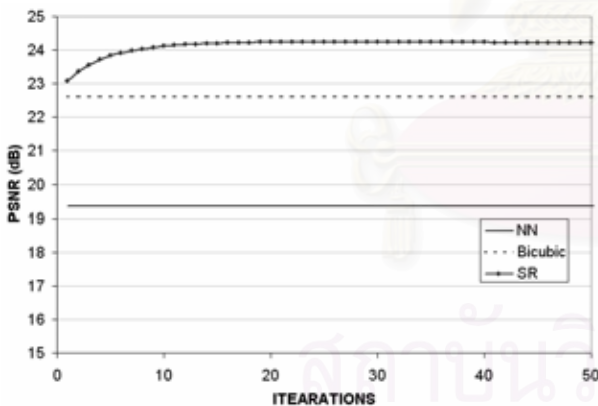
หัวข้อนี้เป็นการแสดงผลการทดสอบ โปรแกรมกับเพิ่มภาพวิดีโอต้นขนาด 320×240 จุดภาพซึ่งถูกลดขนาดลงเป็น 80×60 จุดภาพ โดยในการทดลองแรกเป็นภาพเคลื่อนไหวซึ่งบันทึกด้วยการหมุนกล้องอยู่กับที่โดยกวาดจากขวาไปซ้าย โดยรูปที่ 2-ก เป็นภาพความละเอียดสูงต้นแบบภาพหนึ่งจากชุดภาพเคลื่อนไหวทั้งหมด รูปที่ 2-ข แสดงภาพความละเอียดต่ำภาพหนึ่งจากทั้งหมด 8 ภาพ ซึ่งเมื่อผ่านการเพิ่มขนาดด้วยวิธี

ไบคิวบิก (Bicubic Resize) จะได้ผลลัพธ์ดังรูปที่ 2-ค โดยมีตัวประกอบการขยายเท่ากับ 4 รูปที่ 2-ง แสดงภาพความละเอียดสูงที่ถูกสร้างกลับโดยวิธีการที่นำเสนอ พร้อมแสดงค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) เทียบกับภาพในรูปที่ 2-ก ตัวแปรเสริมต่างๆ ที่ใช้ในการทดลองมีค่าดังนี้: $N = 8$, $a = 0.003$, $b = 2$ และเลือกใช้ B เป็นตัวดำเนินการพรวามัวแบบเกาส์ขนาด 5×5 ซึ่งมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 2

ในการทดลองที่ 2 เป็นกรณีที่ภาพเคลื่อนไหวถูกบันทึกโดยการเคลื่อนกล้องแบบขนานจากขวาไปซ้าย ทำให้การเคลื่อนไหวในภาพเป็นแบบเลื่อนขนานทั้งหมด (Global Translational Motion) โดยรูปที่ 3-ก ถึง 3-ง มีลำดับการแสดงผลภาพเช่นเดียวกับการทดลองแรก ตัวแปรเสริมที่ใช้ในการทดลองมีค่าดังนี้: $a = 0.003$, $b = 1$ โดยที่ B เป็นตัวดำเนินการพรา้มัวแบบเกาส์ขนาด 5×5 ซึ่งมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 3 ทั้งสองการทดลองใช้ภาพที่ขยายด้วยวิธีไบคิวบิกเป็นภาพเริ่มต้น x_0

จากผลการทดลองทั้งสอง จะเห็นได้ว่าภาพผลลัพธ์ที่ได้จากวิธีการที่นำเสนอมีรายละเอียดชัดเจนขึ้นและความพรา้มัวลดลงเมื่อเทียบกับภาพความละเอียดต่ำและมีอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด สูงกว่าภาพความละเอียดสูงจากวิธีไบคิวบิกประมาณ 1.5 ถึง 1.6 dB

โปรแกรมภาษา C++ ที่นำเสนอใช้เวลาในการประมวลผลประมาณ 0.45 วินาทีต่อ 1 รอบทำซ้ำสำหรับการสร้างภาพสี่เหลี่ยมความละเอียดสูงขนาด 320×240 จุดภาพ โดยคำตอบจะดูและเริ่มเสถียรที่รอบทำซ้ำที่ 15 ถึง 20 คิดเป็นเวลาประมาณ 7-10 วินาที ดังแสดงในรูปที่ 4 ซึ่งเป็นกราฟแสดงค่า PSNR ที่แต่ละรอบทำซ้ำของการทดลองที่ 1 เปรียบเทียบกับวิธีไบคิวบิกและวิธีข้างเคียงใกล้สุด (Nearest Neighborhood : NN) จากการทดลองปรับเปลี่ยนค่าระยะก้าวในช่วง 1.0-4.0 พบว่าให้ภาพผลลัพธ์ที่แทบจะไม่แตกต่างกันโดยค่าระยะก้าวที่มากกว่าจะทำให้คำตอบดูเข้าได้รวดเร็วกว่า อย่างไรก็ตามหากเลือกค่าระยะก้าวมากกว่า 4.0 ขึ้นไปจะทำให้จุดภาพเริ่มอึมครึมเมื่อทำซ้ำไปได้ระยะหนึ่ง



รูปที่ 4 ค่า PSNR ในแต่ละรอบทำซ้ำของการทดลองที่ 1

6 สรุป

จากผลการทดลองจะเห็นได้ว่าวิธีการที่นำเสนอสามารถปรับปรุงรายละเอียดภายในแฟ้มภาพเคลื่อนไหวให้ดีขึ้นได้ อย่างไรก็ตามวิธีการดังกล่าวยังคงจำกัดอยู่ที่กรณีของภาพเคลื่อนไหวที่ไม่มีปัญหาการบดบังตัวเองและปัญหาความไม่คงรูปของวัตถุ นอกจากนี้ยังต้องอาศัยการปรับค่าตัวแปรเสริมด้วยมือสำหรับการขจัดความพรา้มัวเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด ซึ่งข้อจำกัดเหล่านี้สามารถปรับปรุงได้โดยการศึกษาแบบจำลองการแปลงภาพที่เหมาะสมกับวัตถุในภาพ เช่น ใบหน้าคน

เพื่อให้สามารถประยุกต์ใช้ในระบบเฟื่อระวังเพื่อตรวจจับบุคคลได้ และใช้วิธีทูลประสานกลับแบบบอด (Blind Deconvolution) เพื่อลดความซับซ้อนในการปรับค่าตัวแปรเสริม

เอกสารอ้างอิง

- [1] Michael Elad and Yacov Hel-Or, "A Fast Super-Resolution Reconstruction Algorithm for Pure Translational Motion and Common Space-Invariant Blur," IEEE Transactions on Image Processing vol. 10, no. 8, August 2001.
- [2] David Capel and Andrew Zisserman, "Computer Vision: Applied to Super Resolution," IEEE Signal Processing Magazine, pp. 75-86, May 2003.
- [3] Andrew J. Patti, M Ibrahim Sezan, and A. Murat Tekalp, "Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time," IEEE Transactions on Image Processing vol. 6, no. 8, August 1997.
- [4] Intel Corporation, "Open Source Computer Vision Library," www.intel.com/technology/computing/opencv/.
- [5] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-Resolution Image Reconstruction: A Technical Overview," IEEE Signal Processing Magazine, pp. 21-36, May 2003.
- [6] Sina Farsiu, M Dirk Robinson, Michael Elad, and Peyman Milanfar, "Fast and Robust Multiframe Super Resolution," IEEE Transactions on Image Processing vol. 13, no. 10, October 2004.

ประวัติผู้เขียน



โกศล ปัญญาโสภณเลิศ จบการศึกษาระดับปริญญาตรี สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย ปัจจุบันกำลังศึกษาต่อระดับปริญญาโท สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย มีความสนใจทางด้าน การประมวลผลภาพและคอมพิวเตอร์วิทัศน์



สุรีย์ พุ่มรินทร์ จบการศึกษาระดับปริญญาตรี สาขาวิศวกรรมอิเล็กทรอนิกส์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง ระดับปริญญาโท สาขาวิศวกรรมไฟฟ้า มหาวิทยาลัยอริโซนาสเตท และระดับปริญญาเอก สาขาวิศวกรรมไฟฟ้า มหาวิทยาลัยยวชิงตัน ปัจจุบันดำรงตำแหน่งอาจารย์ประจำภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย มีความสนใจทางด้าน การประมวลผลภาพ คอมพิวเตอร์วิทัศน์ และระบบฝังตัว

ประวัติผู้เขียนวิทยานิพนธ์

นายโกศล ปัญญาโสภณเลิศ เกิดเมื่อวันที่ 20 เมษายน พ.ศ.2524 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า จากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2544 จากนั้นจึงเข้าทำงานในตำแหน่งวิศวกรแผนกรับประกันคุณภาพ บริษัทโซนี่ ดีไวซ์ เทคโนโลยี (ประเทศไทย) จำกัด จนถึงปี 2548 ได้ลาออกมาศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2548



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย