



การปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองเกาส์เซียนหลายระดับ
แบบปรับเหมาะสำหรับการตรวจจับบริเวณที่น่าสนใจของภาพ



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
ปีการศึกษา 2562
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

การปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองเกาส์เซียนหลายระดับ
แบบปรับเหมาะสำหรับการตรวจจับบริเวณที่น่าสนใจของภาพ



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
ปีการศึกษา 2562
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

วิทยานิพนธ์ เรื่อง "การปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองเกาส์เซียน
หลายระดับแบบปรับเหมาะสำหรับการตรวจจับบริเวณที่น่าสนใจของภาพ"

ของ เจนจิรา แจ่มศิริ

ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการสอบวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรัตน์ปัญญา)

..... ประธานที่ปรึกษาวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร.จรัสศรี รุ่งรัตนอุบล)

..... กรรมการผู้ทรงคุณวุฒิภายใน

(รองศาสตราจารย์ ดร.ไกรศักดิ์ เกษร)

..... กรรมการผู้ทรงคุณวุฒิภายใน

(ดร.สุธาสินี จิตต่อนันต์)

อนุมัติ

.....
(ศาสตราจารย์ ดร.ไพศาล มุณีสว่าง)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรอง เกาส์เซียนหลายระดับแบบปรับเหมาะสำหรับการตรวจจับบริเวณที่ น่าสนใจของภาพ
ผู้วิจัย	เจนจิรา แจ่มศิริ
ประธานที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. จรัสศรี รุ่งรัตนอุบล

ประเภทสารนิพนธ์ วิทยานิพนธ์ วท.ม. สาขาวิชาเทคโนโลยีสารสนเทศ, มหาวิทยาลัย
นเรศวร, 2562

คำสำคัญ ขนาดและพื้นที่ของตัวกรองเกาส์เซียน; ค่าเอนโทรปีแบบท้องถิ่น; การ
วิเคราะห์แบบระดับ; การตรวจจับบริเวณที่น่าสนใจ

บทคัดย่อ

การตรวจจับบริเวณที่น่าสนใจเป็นกระบวนการที่สำคัญโดยเฉพาะในการค้นหาบริเวณที่สำคัญบนภาพ ซึ่งเป็นที่นิยมในการประมวลผลภาพ การตรวจจับบริเวณที่น่าสนใจได้รับการปรับปรุงและประยุกต์ใช้งานในหลาย ๆ ด้านของระบบคอมพิวเตอร์วิทัศน์โดยคอมพิวเตอร์พยายามที่จะทำงาน เข้าใจหรือเลียนแบบการมองเห็นของมนุษย์ โดยมีงานวิจัยหลายงานพยายามปรับปรุงวิธีการตรวจจับเพื่อนำเสนอแบบจำลองที่มีประสิทธิภาพในการตรวจหาบริเวณที่มีความสำคัญที่สอดคล้องกับการรับรู้ของมนุษย์ ดังนั้น งานวิจัยนี้จึงมุ่งเน้นไปที่การพัฒนาและปรับปรุงกระบวนการตรวจจับบริเวณที่น่าสนใจของภาพให้ดีขึ้น โดยนำเสนอการปรับปรุงขนาดและพื้นที่ของตัวกรองสัญญาณเกาส์เซียนที่สามารถลดบริเวณที่ไม่สนใจออกจากพื้นที่ของบริเวณที่น่าสนใจได้ ส่งผลให้บริเวณที่น่าสนใจของภาพปรากฏเด่นชัดขึ้น รวมถึงมีการใช้เกณฑ์เอนโทรปีแบบท้องถิ่นเพื่อเลือกภาพบริเวณที่น่าสนใจที่ได้จากปริภูมิสี RGB และ Lab ที่มีเหมาะสมมากขึ้น จากผลการทดลองแสดงให้เห็นว่าประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้มีประสิทธิภาพค่า AUC โดยรวมที่ดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ได้แก่ วิธีการของ HFT, Itti, SAL, SR และ SUN ตามลำดับ และภาพผลลัพธ์ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้มีความใกล้เคียงกับภาพเฉลี่ยที่มนุษย์สร้างขึ้นอีกด้วย

Title	MODIFIED SCALE-SPACE ANALYSIS IN FREQUENCY DOMAIN BASED ON ADAPTIVE MULTISCALE GAUSSIAN FILTER FOR SALIENCY DETECTION
Author	JENJIRA JAEMSIRI
Advisor	Assistant Professor Jaratsri Rungrattanaubol, Ph.D.
Academic Paper	Thesis M.S. in Information Technology, Naresuan University, 2019
Keywords	Scale-and-space Gaussian filter; local entropy; scale- space analysis; saliency detection

ABSTRACT

Saliency detection is a process to find a significant region on the images, which is popular in the image processing area. It has been extended and used in many applications in computer vision systems. Many researches have attempted to propose an effective model to detect the saliency area which is corresponding to human perception. Therefore, this research focuses on improving the saliency detection process by proposing the improved adaptive spectrum scale-space. The main contributions of the proposed method include (i) scale-and-space Gaussian filter and (ii) the new method for saliency map selection based on local entropy. Firstly, the Gaussian filter is used to suppress the non-saliency amplitude spectrum to extract the saliency map. Then, the best saliency map is selected from the results of the RGB and Lab color images by using the local entropy criteria. Then, the experimental results of the AUC overall performance of the proposed method are higher than that of the HFT, Itti, SAL, SR, Itti, and SUN methods, respectively. The saliency maps of the proposed method are closely corresponding to the ground truth labeled by humans.

ประกาศคุณูปการ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงอย่างสมบูรณ์ได้ ด้วยความอนุเคราะห์และความช่วยเหลือเป็นอย่างดียิ่งจากผู้มีพระคุณหลายท่าน ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงในความกรุณาของ ผู้ช่วยศาสตราจารย์ ดร.จรัสศรี รุ่งรัตนอุบล อาจารย์ที่ปรึกษาวิทยานิพนธ์ และกรรมการผู้ทรงคุณวุฒิทุกท่าน ที่ให้ความรู้ พร้อมทั้งให้คำแนะนำต่าง ๆ ตลอดระยะเวลาในการทำวิทยานิพนธ์ ตลอดจนในการตรวจทานแก้ไขวิทยานิพนธ์ฉบับนี้ และกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์อันประกอบไปด้วย รองศาสตราจารย์ ดร.ไกรศักดิ์ เกษร ดร.สุธาสินี จิตอนันต์ และ ผู้ช่วยศาสตราจารย์ ดร. กันต์พงษ์ วรรณปัญญญา กรรมการผู้ทรงคุณวุฒิภายนอก ที่ได้ให้คำแนะนำและให้แนวคิดในวิทยานิพนธ์ฉบับนี้ ทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้อย่างสมบูรณ์

กราบขอบพระคุณ ดร.ธราวิเชษฐ์ ธิติจรูญโรจน์ ที่ได้กรุณาเสียสละเวลาอันมีค่า เป็นผู้เชี่ยวชาญให้คำปรึกษาและแนะนำในการทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงสมบูรณ์

ขอขอบคุณคณะวิทยาศาสตร์และภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ที่ให้ความช่วยเหลือในด้านต่าง ๆ

ขอขอบคุณพี่ ๆ เพื่อน ๆ ในภาควิชาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ทุกคนที่คอยเป็นกำลังใจในการทำงานวิจัยตลอดมา

ท้ายที่สุดแห่งความสำเร็จในการศึกษานี้ขอขอบคุณพระคุณบุคคลในครอบครัวอันเป็นที่รัก ซึ่งเป็นบุคคลที่ทำให้กำลังใจผู้วิจัยมาโดยตลอด รวมไปถึงอาจารย์และผู้มีพระคุณทุกท่านที่ให้การสนับสนุนผู้วิจัยขอกราบขอบพระคุณด้วยความรักและความเคารพเป็นอย่างสูง

เจนจิรา แจ่มศิริ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
ประกาศคุุณูปการ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ญ
สารบัญรูปภาพ.....	ฎ
บทที่ 1.....	1
บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 จุดมุ่งหมายของการศึกษา.....	4
1.3 ขอบเขตของงานวิจัย.....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.5 นิยามศัพท์เฉพาะ.....	5
1.6 แผนการดำเนินงาน.....	5
บทที่ 2.....	6
เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
2.1 แบบจำลองสี (Color Model).....	6
2.1.1 RGB.....	6
2.1.2 Lab.....	7

2.1.3 ภาพระดับสีเทา (Grayscale Image).....	8
2.1.4 การแปลงภาพสีจากแบบจำลองสี RGB ไปเป็นภาพระดับสีเทา (Gray Scale). 9	
2.1.5 การแปลงภาพระดับสีเทา (Gray Level) ไปเป็นภาพไบนารี (Binary Image)... 9	
2.2. การแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform).....	10
2.2.1 การแปลงฟูเรียร์ (Fourier Transform).....	10
2.2.2 การแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform)... 11	
2.3 การตรวจจับบริเวณที่น่าสนใจของภาพ (Saliency Detection).....	12
2.3.1 อ้างอิงตามคุณสมบัติบนโดเมนตำแหน่ง	13
2.3.2 อ้างอิงตามสเปกตรัมบนโดเมนความถี่	13
2.4 การกรองภาพ (Image Filtering).....	14
2.4.1 ตัวกรองความถี่ต่ำ (Low-pass filter)	14
2.4.2 ตัวกรองความถี่สูง (High-pass filter).....	15
2.4.3 ตัวกรองสัญญาณเกาส์เซียน (Gaussian filter).....	15
2.4.4 ตัวดำเนินการแบบคอนโวลูชัน (Convolution Operation Filtering).....	16
2.5 งานวิจัยที่เกี่ยวข้อง	17
บทที่ 3	21
วิธีดำเนินงานวิจัย.....	21
3.1 การเตรียมข้อมูลภาพ	21
3.2 การออกแบบและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1	23
3.3 การออกแบบและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2	27
3.4 การประเมินประสิทธิภาพของการวิจัย.....	36
3.4.1 พื้นที่ใต้เส้นโค้ง Area Under Curve (AUC)	36

3.4.2 กราฟ Receiver Operator Characteristic (ROC)	37
บทที่ 4	39
ผลการวิจัย	39
4.1 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1	39
4.2 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2	40
4.3 ผลการวิเคราะห์ผลการทดลองเชิงเปรียบเทียบโดยจำแนกตามหมวดหมู่ของภาพ 6 ประเภท	43
4.4 ผลการเปรียบเทียบประสิทธิภาพด้านระยะเวลาในการประมวลผลของการตรวจจับ บริเวณที่น่าสนใจของภาพ	46
บทที่ 5	49
บทสรุป	49
5.1 สรุปผลการวิจัย	49
5.2 อภิปรายผล	49
5.3 ข้อเสนอแนะ	50
บรรณานุกรม	51
ภาคผนวก	53
ภาคผนวก ก	54
ภาคผนวก ข	61
ประวัติผู้วิจัย	74

สารบัญตาราง

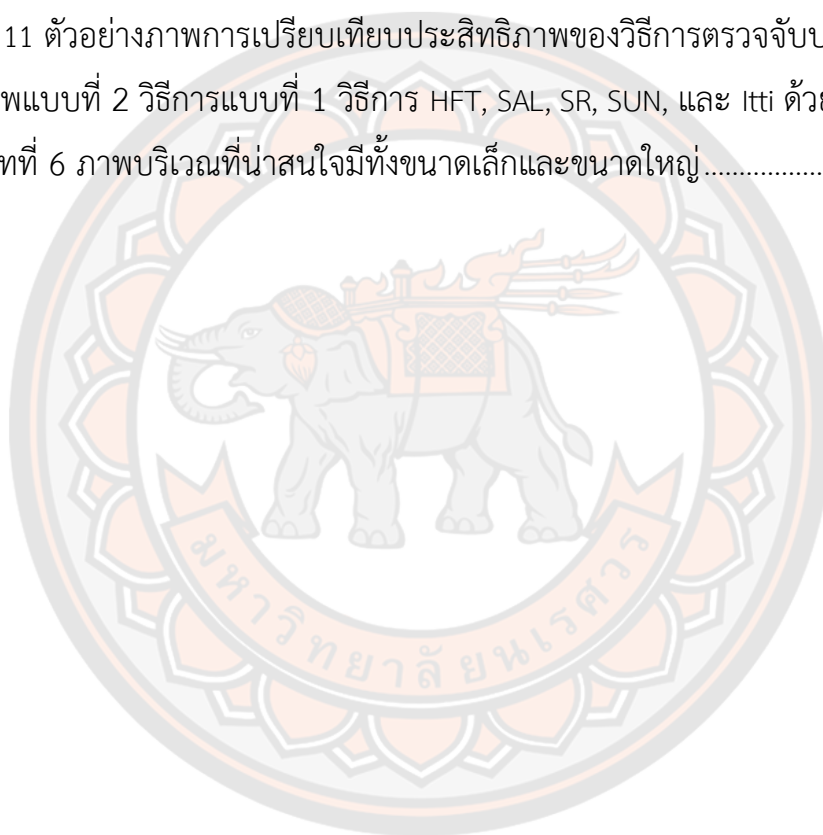
หน้า

ตาราง 1 แผนการดำเนินงาน	5
ตาราง 2 เปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 กับวิธีการ Itti, SUN, SR, SAL, และ HFT ด้วยค่า AUC	39
ตาราง 3 เปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 กับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, Itti, SUN, SR, SAL, และ HFT ด้วย ค่า AUC	41
ตาราง 4 การเปรียบเทียบประสิทธิภาพการเลือกใช้ตัวเลือกต่าง ๆ ของวิธีการตรวจจับ บริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ด้วยค่า AUC	42
ตาราง 5 เปรียบเทียบประสิทธิภาพการวิธีตรวจจับบริเวณที่น่าสนใจของภาพด้วยค่า AUC กับระยะเวลาในการประมวลผลของวิธีการ Itti, SUN, SR, SAL, HFT, การตรวจจับบริเวณที่ น่าสนใจของภาพแบบที่ 1 และการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2	47
ตาราง 6 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน ประเภทที่ 1 ภาพบริเวณที่น่าสนใจมีขนาดใหญ่	55
ตาราง 7 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน ประเภทที่ 2 ภาพบริเวณที่น่าสนใจมีขนาดปานกลาง	56
ตาราง 8 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน ประเภทที่ 3 ภาพบริเวณที่น่าสนใจมีขนาดเล็ก	57

ตาราง 9 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ
ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน
ประเภทที่ 4 ภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน58

ตาราง 10 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ
ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน
ประเภทที่ 5 ภาพพื้นหลังมีความซับซ้อน.....59

ตาราง 11 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจ
ของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ใน
ประเภทที่ 6 ภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่.....60



สารบัญรูปภาพ

	หน้า
ภาพ 1 ตัวอย่างบริเวณที่น่าสนใจของภาพ	1
ภาพ 2 ผลลัพธ์การตรวจจับบริเวณที่น่าสนใจ	2
ภาพ 3 แบบจำลองสี RGB.....	7
ภาพ 4 แบบจำลองสี Lab.....	7
ภาพ 5 ภาพระดับสีเทา.....	8
ภาพ 6 การไล่ค่าสีของภาพระดับสีเทา	9
ภาพ 7 ภาพไบนารี.....	10
ภาพ 8 กราฟของสัญญาณในโดเมนเวลาและโดเมนความถี่.....	11
ภาพ 9 ผลลัพธ์จากการแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์.....	12
ภาพ 10 ตำแหน่งของความถี่	14
ภาพ 11 ผลลัพธ์เมื่อภาพผ่านตัวกรองความถี่ต่ำ	15
ภาพ 12 ผลลัพธ์เมื่อภาพผ่านตัวกรองความถี่สูง	15
ภาพ 13 ตัวกรองสัญญาณเกาส์เซียนแบบ 2 มิติ	16
ภาพ 14 แสดงการทำงานของคอนโวลูชัน	16
ภาพ 15 ตัวอย่างผลลัพธ์การตรวจจับบริเวณที่น่าสนใจโดยวิธีการของ Jain.....	19
ภาพ 16 ผลลัพธ์ของภาพบริเวณที่น่าสนใจของภาพที่แสดงในปริภูมิสี RGB และ Lab.....	20
ภาพ 17 ภาพบริเวณที่น่าสนใจมีขนาดใหญ่.....	21
ภาพ 18 ภาพบริเวณที่น่าสนใจมีขนาดปานกลาง.....	21
ภาพ 19 ภาพบริเวณที่น่าสนใจมีขนาดเล็ก	22

ภาพ 20 ภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน	22
ภาพ 21 ภาพพื้นหลังมีความซับซ้อน.....	22
ภาพ 22 ภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่.....	22
ภาพ 23 ภาพรวมการทำงานของการทำงานของตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1	24
ภาพ 24 กระบวนการทำงานการกรองสเปกตรัมของวิธีการแบบที่ 1	25
ภาพ 25 ภาพรวมการทำงานของตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2	28
ภาพ 26 กระบวนการทำงานการกรองสเปกตรัมของวิธีการแบบที่ 2.....	30
ภาพ 27 ผลลัพธ์ของภาพที่แสดงบริเวณที่น่าสนใจจากปริภูมิสี RGB และ Lab.....	34
ภาพ 28 ผลลัพธ์ของสเปกตรัมที่ได้จากวิธีการ (ก) HFT (ข) วิธีการแบบที่ 1 และ (ค) วิธีการแบบที่ 2	35
ภาพ 29 ตัวอย่างภาพผลลัพธ์และภาพเฉลี่ยที่นำมาใช้ในการวัดประสิทธิภาพด้วยค่า AUC	36
ภาพ 30 ตัวอย่างการเปรียบเทียบระหว่างภาพผลลัพธ์และภาพเฉลี่ย.....	38
ภาพ 31 ตารางแจกแจงความถูกต้องของการจำแนก (Confusion Matrix).....	38
ภาพ 32 ตัวอย่างภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ	44
ภาพ 33 เปรียบเทียบภาพบริเวณที่น่าสนใจที่ได้จากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 แบบที่ 1, HFT, Itti, SAL, SR และ SUN	48

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

บริเวณที่น่าสนใจหรือลักษณะเด่น (Saliency Area) คือ บริเวณที่โดดเด่นจากบริเวณอื่น ๆ ของสิ่งใดสิ่งหนึ่งไม่ว่าจะเป็น วัตถุ บุคคล สิ่งของ ฯลฯ หรือสามารถกล่าวได้ว่า คือ บริเวณที่มีลักษณะแตกต่างจากบริเวณอื่น ๆ อาทิเช่น สี รูปร่าง หรือลวดลาย เป็นต้น โดยทั่วไปมักเกิดขึ้นจากความแตกต่างระหว่างบริเวณที่เป็นลักษณะเด่นกับบริเวณที่อยู่รอบข้าง (Itti, Koch, & Niebur, 1998) มนุษย์เราสามารถมองเห็น และรับภาพผ่านดวงตา จึงสามารถตีความหมายของภาพและรู้จำวัตถุที่เป็นองค์ประกอบภายในภาพได้ทันทีที่มองไปยังรูปภาพ การศึกษาวิจัยการตรวจจับบริเวณที่น่าสนใจของภาพ (Saliency Detection) ซึ่งได้รับความสนใจอย่างมากในระบบภาพ (Vision System) เป็นการทำงานด้านกระบวนการการมองเห็นของมนุษย์ในการรู้จำภาพด้วยกระบวนการคอมพิวเตอร์ ที่เป็นการเลียนแบบการมองเห็นของมนุษย์เพื่อตรวจจับบริเวณที่น่าสนใจ

บริเวณที่น่าสนใจของภาพ คือ บริเวณที่มีลักษณะโดดเด่นมากที่สุดในภาพและเป็นบริเวณที่ดึงดูดสายตาของมนุษย์เมื่อมองเข้าไปในภาพดังแสดงในภาพ 1



ภาพ 1 ตัวอย่างบริเวณที่น่าสนใจของภาพ

ในภาพ 1 บริเวณที่น่าสนใจหรือลักษณะเด่นของภาพในที่นี้ คือ จักรยาน ดอกไม้ และรถยนต์ ตามลำดับจากภาพด้านซ้ายมาขวา

การตรวจจับบริเวณที่น่าสนใจของภาพ คือ การประมวลผลภาพเพื่อค้นหาส่วนต่าง ๆ ที่มีลักษณะเด่นที่สุดในภาพ ซึ่งบริเวณที่น่าสนใจหรือบริเวณที่มีลักษณะเด่นจะเกี่ยวข้องกับกระบวนการมองเห็นของมนุษย์ (Human Vision System) โดยที่มนุษย์มีความสามารถในการกรองบริเวณที่โดดเด่นในภาพได้อย่างรวดเร็วและง่ายดาย กล่าวคือ เมื่อมนุษย์มองไปในรูปภาพ มนุษย์จะสามารถบอกได้ว่าบริเวณไหนหรือพื้นที่ตรงไหนของภาพเป็นจุดเด่นที่น่าสนใจที่สุดได้ทันที การตรวจจับ

บริเวณที่น่าสนใจของภาพสามารถให้ข้อมูลที่เป็นประโยชน์เกี่ยวกับตำแหน่งของวัตถุหรือบริเวณที่เป็นลักษณะเด่นที่สุดในภาพ ดังแสดงในภาพ 2 การตรวจจับบริเวณที่น่าสนใจของภาพได้รับการนำมาประยุกต์ใช้อย่างกว้างขวาง ตัวอย่างการนำการตรวจจับบริเวณที่น่าสนใจของภาพมาประยุกต์ใช้ ดังเช่น การแบ่งส่วนของภาพ (Achanta, Hemami, Estrada, & Susstrunk, 2009) การติดตามวัตถุ (Object tracing) การรู้จำวัตถุ (Object recognition) (Walther, Itti, Riesenhuber, Poggio, & Koch, 2002) และการตรวจจับวัตถุที่เคลื่อนไหว (Rokszin et al., 2010) เป็นต้น



ภาพ 2 ผลลัพธ์การตรวจจับบริเวณที่น่าสนใจ

การออกแบบการตรวจจับบริเวณที่น่าสนใจในภาพได้รับความสำคัญจากการศึกษาทางด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) เป็นเวลาหลายทศวรรษ ทำให้มีงานวิจัยออกมาหลายรูปแบบ โดยการตรวจจับบริเวณที่น่าสนใจของภาพสามารถแยกออกได้เป็นสองกลุ่มหลักที่ใช้เทคนิคที่ต่างกัน (Jaemsiri, Titijaroonroj, & Rungrattanaubol, 2019) ได้แก่ 1) อ้างอิงตามคุณสมบัติบนโดเมนตำแหน่ง (Spatial Domain) 2) อ้างอิงตามสเปกตรัมบนโดเมนความถี่ (Frequency Domain) วิธีการกลุ่มแรกอ้างอิงตามคุณสมบัติบนโดเมนตำแหน่งโดย Itti และคณะ (Itti et al., 1998) ได้นำเสนอวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ โดยใช้หลักการในการมองเห็นของมนุษย์ ซึ่งวิธีการนี้จะพิจารณาความแตกต่างระหว่างพื้นที่ที่ทำการพิจารณาให้เป็นบริเวณที่น่าสนใจกับพื้นที่รอบ ๆ โดยที่ขนาดของภาพแตกต่างกัน และใช้คุณสมบัติของภาพ ได้แก่ สี ทิศทาง และความเข้มของแสง เพื่อทำการสกัดบริเวณที่น่าสนใจของภาพ ในปี 2008 กรอบแนวคิดแบบเบย์ (Zhang, Tong, Marks, Shan, & Cottrell, 2008) สำหรับการตรวจจับบริเวณที่น่าสนใจของได้ถูกเสนอโดย Zhang และคณะ วิธีการนี้สามารถประมาณความน่าจะเป็นของบริเวณที่น่าสนใจของภาพทั้งหมด โดยขึ้นอยู่กับคุณสมบัติของภาพ ได้แก่ ทิศทางและองศา โดยจะต้องนำคุณสมบัติเหล่านี้ให้คอมพิวเตอร์เรียนรู้ก่อนที่จะทำการตรวจจับบริเวณที่น่าสนใจ จะเห็นได้ว่าวิธีการในกลุ่มอ้างอิงตามคุณสมบัติบนโดเมนตำแหน่งต้องการความรู้ก่อนหน้าและชุดภาพฝึกสอนเพื่อให้คอมพิวเตอร์ได้เรียนรู้ก่อนนำมาตรวจจับบริเวณที่น่าสนใจของภาพ หากไม่มีการฝึกสอนอาจทำงานไม่ได้หรือทำให้ผลลัพธ์ที่ได้เกิดความผิดพลาด

วิธีการในกลุ่มที่สองอ้างอิงตามสเปกตรัมบนโดเมนความถี่ วิธีการในกลุ่มนี้ได้ทดลองตรวจจับบริเวณที่น่าสนใจในภาพโดยมีจุดประสงค์เพื่อให้ได้ภาพที่เก็บเฉพาะบริเวณที่น่าสนใจ (Saliency Map) โดยใช้สเปกตรัมบนโดเมนความถี่ เช่น วิธีการสเปกตรัมส่วนที่เหลือ โดย Hou และคณะ (Hou & Zhang, 2007) เสนอวิธีการแปลงข้อมูลจากโดเมนตำแหน่งไปอยู่บนโดเมนความถี่ แล้วนำภาพไปหาค่าเฉลี่ยของสเปกตรัม จากนั้นนำภาพต้นฉบับที่แปลงเป็นสัญญาณความถี่กลับกับภาพที่หาค่าเฉลี่ยของสเปกตรัมจะได้ผลลัพธ์เป็นส่วนที่เหลือและแปลงส่วนที่เหลือกลับมาโดเมนตำแหน่ง ซึ่งผลลัพธ์ที่ได้ คือ ภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ ต่อมาในปี 2012 Hou และคณะ (Hou, Harel, & Koch, 2012) ได้ประยุกต์ใช้วิธีการแปลงแบบโคไซน์ไม่ต่อเนื่อง (Discrete Cosine Transform: DCT) สำหรับการแยกสัญญาณของแต่ละองค์ประกอบซึ่งเทียบเท่ากับองค์ประกอบเฟส (Phase) ในโดเมนความถี่ ซึ่งสามารถประมาณพื้นที่ที่มีลักษณะเด่นในเบื้องหน้า (ฉากหน้าของภาพ) เพื่อพิจารณาให้เป็นบริเวณที่สำคัญสำหรับการแยกบริเวณที่เป็นที่น่าสนใจออกจากพื้นหลังที่มีความซับซ้อนได้ ในปี 2013 Jain และคณะ (Jian li, Levine, An, Xu, & He, 2012) ได้นำเสนอการวิเคราะห์หลายระดับบนโดเมนความถี่โดยใช้การแปลงฟูเรียร์ไฮเปอร์คอมเพล็กซ์ (Hypercomplex Fourier Transform: HFT) วิธีการนี้จะทำการระงับสเปกตรัมที่มีรูปแบบซ้ำซ้อน โดยใช้ตัวกรองสัญญาณเกาส์เซียนหลายระดับร่วมกับเกณฑ์วัดค่าความปะปนกันของข้อมูลหรือเกณฑ์เอนโทรปี แต่วิธีการของ Jain ยังคงใช้ขนาดของตัวกรองสัญญาณเกาส์เซียนขนาดเดียวกันทั้งหมดในการกรองภาพทุกภาพ ต่อมาในปี 2019 Jaemsiri และคณะ (Jaemsiri et al., 2019) ได้นำเสนอการตรวจจับบริเวณที่น่าสนใจของภาพโดยใช้ตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะ (AMG filter) เพื่อทำงานระงับรูปแบบที่ซ้ำซ้อนของสเปกตรัม วิธีการนี้ทำให้ได้ประสิทธิภาพของการตรวจจับมากขึ้นด้วยการปรับพื้นที่ของตัวกรองสัญญาณเกาส์เซียนเป็นแบบอัตโนมัติ แต่ถึงอย่างไรก็ตาม แม้ว่าจะมีการปรับพื้นที่ของตัวกรองสัญญาณเกาส์เซียนแบบอัตโนมัติแล้วแต่ก็ยังคงใช้ขนาดของตัวกรองสัญญาณเกาส์เซียนเพียงขนาดเดียวเท่านั้น

งานวิจัยเกี่ยวกับการตรวจจับบริเวณที่น่าสนใจของภาพโดยส่วนใหญ่นั้นเป็นการใช้ปริภูมิสี RGB ในการพัฒนาวิธีการ แต่ในปี 2015 Xiang และคณะ (Xiang & Zhong, 2015) ได้ทำการใช้ปริภูมิสี RGB ร่วมกับ Lab ในการสร้างวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ เนื่องจาก Xiang และคณะ เชื่อว่าลักษณะที่น่าสนใจหรือบริเวณที่เป็นวัตถุเด่นนั้น สามารถปรากฏขึ้นได้ทั้งในปริภูมิสี RGB และ Lab ซึ่งการปรากฏในแต่ละปริภูมิสีจะไม่เหมือนกันแล้วแต่ลักษณะของบริเวณที่เป็นวัตถุเด่นของรูปแต่ละรูปที่แตกต่างกันออกไป

จากที่กล่าวมา ผู้วิจัยจึงเล็งเห็นแนวทางในการปรับปรุงขั้นตอนวิธีในการตรวจจับบริเวณที่น่าสนใจของภาพให้มีประสิทธิภาพมากยิ่งขึ้น บนแนวคิดของการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองเกาส์เซียนหลายระดับแบบปรับเหมาะ

1.2 จุดมุ่งหมายของการศึกษา

1. เพื่อศึกษาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ
2. เพื่อปรับปรุงและพัฒนากระบวนการตรวจจับบริเวณที่น่าสนใจของภาพให้มีประสิทธิภาพ
3. เพื่อเปรียบเทียบประสิทธิภาพของกระบวนการตรวจจับบริเวณที่น่าสนใจของภาพที่พัฒนาขึ้นกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพกับวิธีอื่น ได้แก่ วิธีการของ Itti, HFT, SR และ SUN

1.3 ขอบเขตของงานวิจัย

1.3.1 ขอบเขตข้อมูล

ผู้วิจัยได้ทำการวิจัย โดยใช้ข้อมูลดังต่อไปนี้

ภาพทั่วไปเป็นภาพนิ่ง จำนวน 235 รูป ขนาด 640 × 480 พิกเซล จากชุดข้อมูล ชื่อ Salmap (J. Li, Levine, An, Xu, & He, 2013) ประกอบด้วย 6 ประเภท ดังนี้

- 1) ภาพบริเวณที่น่าสนใจมีขนาดใหญ่ (Large salient region) จำนวน 50 รูป
- 2) ภาพบริเวณที่น่าสนใจมีขนาดปานกลาง (Intermediate salient regions) จำนวน 80 รูป
- 3) ภาพบริเวณที่น่าสนใจมีขนาดเล็ก (Small salient regions) จำนวน 60 รูป
- 4) ภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน (Repeating distractors) จำนวน 15 รูป
- 5) ภาพพื้นหลังมีความซับซ้อน (Cluttered backgrounds) จำนวน 15 รูป
- 6) ภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่ (Both large and small salient regions) จำนวน 15 รูป

1.3.2 ขอบเขตระบบ

ตรวจจับบริเวณที่น่าสนใจในรูปภาพสองมิติที่เป็นภาพนิ่ง เพื่อให้ได้ผลลัพธ์เป็นภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้รับความรู้เกี่ยวกับเรื่องการตรวจจับบริเวณที่น่าสนใจของภาพ
2. ได้กระบวนการสำหรับตรวจจับของบริเวณที่น่าสนใจของภาพ
3. ได้ศึกษาความแตกต่างของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่ได้พัฒนากับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ได้แก่ วิธีการของ Itti, HFT, SR และ SUN รวมถึงประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจที่พัฒนามีประสิทธิภาพเมื่อเปรียบเทียบกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพวิธีอื่น โดยใช้วิธีวัดประสิทธิภาพด้วยค่า AUC

[illegible]

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยทฤษฎีพื้นฐานต่าง ๆ ที่นำมาประยุกต์ในการวิจัยประกอบด้วยเนื้อหาดังต่อไปนี้

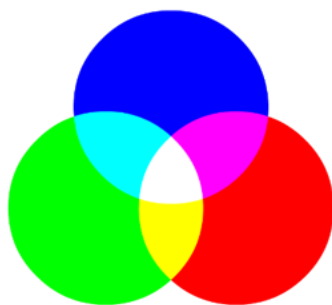
- 2.1 แบบจำลองสี (Color Model)
- 2.2 การแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform)
- 2.3 การตรวจจับบริเวณที่น่าสนใจของภาพ (Saliency Detection)
- 2.4 การกรองภาพ (Image Filtering)
- 2.5 งานวิจัยที่เกี่ยวข้อง

2.1 แบบจำลองสี (Color Model)

แบบจำลองสี คือ รูปแบบการผสมสีด้วยเทคนิคหรือวิธีการต่าง ๆ เพื่อให้ได้มาซึ่งสีที่ใช้ในการแสดงออกทางจอภาพหรือเครื่องพิมพ์ โดยในงานวิจัยนี้ได้ใช้แบบจำลองสีดังต่อไปนี้

2.1.1 RGB

ระบบสี RGB เป็นระบบสีของแสงซึ่งเกิดจากการหักเหของแสงผ่านแท่งแก้วปริซึม จะเกิดแถบสีที่เรียกว่า สเปกตรัม (Spectrum) ซึ่งแยกสีตามที่ยาวตามองเห็นได้ 7 สี คือ แดง แสด เหลือง เขียว น้ำเงิน คราม ม่วง ซึ่งเป็นพลังงานอยู่ในรูปของรังสีที่มีช่วงคลื่นที่สายตาสามารถมองเห็นได้ แสงสีม่วงมีความถี่คลื่นสูงสุด คลื่นแสงที่มีความถี่สูงกว่าแสงสีม่วงเรียกว่า อัลตราไวโอเล็ต (Ultra Violet) และคลื่นแสงสีแดงมีความถี่คลื่นต่ำที่สุด คลื่นแสงที่ต่ำกว่าแสงสีแดงเรียกว่า อินฟราเรด (Infrared) คลื่นแสงที่มีความถี่สูงกว่าสีม่วง และต่ำกว่าสีแดงนั้น สายตาของมนุษย์ไม่สามารถรับได้ และเมื่อศึกษาดูแล้วแสงสีทั้งหมดเกิดจากแสงสี 3 สี คือ สีแดง (Red) สีน้ำเงิน (Blue) และสีเขียว (Green) ทั้ง 3 สีถือเป็นแม่สีของแสง เมื่อนำมาฉายรวมกันจะทำให้เกิดสีใหม่ อีก 3 สี คือ สีแดงมาเจนต้า สีฟ้าไซแอนและสีเหลือง ดังแสดงในภาพ 3 และถ้าฉายแสงสีทั้งหมดรวมกันจะได้แสงสีขาว จากคุณสมบัติของแสงนี้เราสามารถนำมาใช้ประโยชน์ทั่วไป ได้แก่ ในการฉายภาพยนตร์ การบันทึกภาพ วิดีโอ ภาพโทรทัศน์ การสร้างภาพเพื่อการนำเสนอทางจอคอมพิวเตอร์ และการจัดแสงสีในการแสดง เป็นต้น



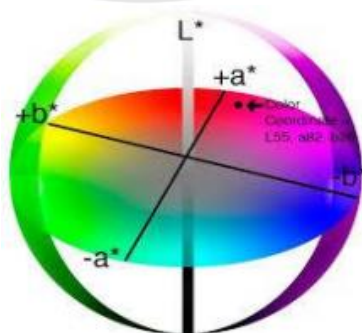
ภาพ 3 แบบจำลองสี RGB

2.1.2 Lab

ระบบสีแบบ Lab คือระบบสีที่ไม่อิงอุปกรณ์ (Device Independent Color) เป็นอีกหนึ่งระบบสีที่นิยมกันมากในการนำมาใช้วัดค่าสีและใช้กันอย่างแพร่หลายในหลาย ๆ งาน โดยหน่วยสีนี้เป็นประเภทที่มีสเกลสม่ำเสมอ (Uniform) และเป็นระบบสีที่สัมพันธ์กับการมองเห็นของมนุษย์ไม่ขึ้นกับอุปกรณ์ใด ๆ กล่าวคือ เมื่อวัดค่าสีได้เท่ากันแล้ว อุปกรณ์ต่าง ๆ จะแสดงสีที่เห็นเหมือนกันในสภาวะแวดล้อมอย่างเดียวกัน เช่น กล้อง สแกนเนอร์ จอภาพ เครื่องพิมพ์ เป็นต้น

สถาบัน Commission Internationale de l'Eclairage (CIE) พัฒนาระบบสี CIE Lab ในปี ค.ศ. 1976 แสดงดังภาพ 4 และได้ทำการใช้ตัวเลขในการแทนค่าสีต่าง ๆ ดังนี้

- L (Lightness) ใช้กำหนดค่าความสว่างและความมืด ถ้า L มีค่าเท่ากับ 0 จะมีค่าเป็นโทนสีดำ และ L มีค่าเท่ากับ 100 จะมีค่าเป็นโทนสีขาว
- a (Red – Green) ใช้กำหนดค่าโทนสีแดงและโทนสีเขียว ถ้า a มีค่ามากกว่า 0 จะมีค่าเป็นโทนสีแดง และ a มีค่าน้อยกว่า 0 จะมีค่าเป็นโทนสีเขียว
- b (Blue – Yellow) ใช้กำหนดค่าโทนสีน้ำเงินและโทนสีเหลือง ถ้า b มีค่ามากกว่า 0 จะมีค่าเป็นโทนสีเหลือง และ b มีค่าน้อยกว่า 0 จะมีค่าเป็นโทนสีน้ำเงิน



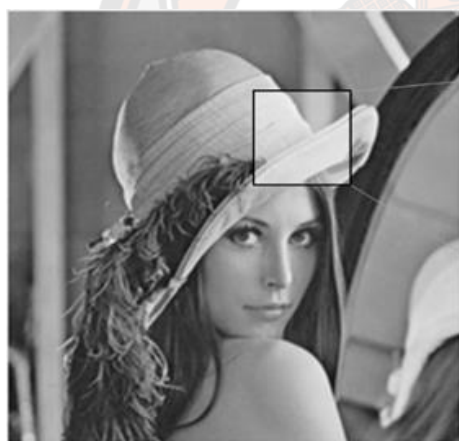
ภาพ 4 แบบจำลองสี Lab

ที่มา: (มหาวิทยาลัยศรีนครินทรวิโรฒ, 2558)

ระบบสีไม่อิงกับอุปกรณ์ (Device Independent Color) คือ ระบบสีที่มีค่าสีไม่เปลี่ยนแปลงตามอุปกรณ์ที่ใช้งาน เช่น ระบบสี CIE Lab ระบบสี CIE Lch เป็นต้น ค่าสีของระบบนี้ประมวลผลจากข้อมูลสเปกตรัมของแสง ดังนั้นจึงนำค่าสีไปใช้เป็นมาตรฐานของระบบการจัดการสี (Color Management) เพื่อใช้สื่อสารค่าสีระหว่างอุปกรณ์ต่าง ๆ สีอิงอุปกรณ์ (Device Dependent Color) และ สีไม่อิงกับอุปกรณ์ (Device Independent Color)

2.1.3 ภาพระดับสีเทา (Grayscale Image)

ภาพระดับสีเทา (ไทยนิมิต, 2010) เป็นภาพที่ความเข้มแสงของพิกเซลมีเจดสีอยู่ในช่วงสีดำ-เทา-ขาว ค่าความเข้มแสงของแต่ละพิกเซลมีค่าอยู่ระหว่าง 0 (สีดำ) ถึง 255 (สีขาว) ในการประมวลผลภาพระดับสีเทา ค่าความเข้มแสงของภาพจะถูกเก็บในเมตริกซ์ 2 มิติ ขนาดเท่ากับ ความละเอียดของภาพ แสดงดังภาพ 5 เป็นตัวอย่างของภาพระดับสีเทาพร้อมทั้งตัวอย่างค่าความเข้มแสงของพิกเซลภายในภาพ



200	198	159	160	140
195	198	182	154	169
180	187	132	145	120
170	195	189	170	80
114	108	54	60	96

ภาพ 5 ภาพระดับสีเทา

ภาพ 5 จะแสดงให้เห็นว่า บริเวณที่มีความเข้มแสงน้อยหรือบริเวณที่มีแสงทึบ อาทิเช่น บริเวณเงาตรงรอยต่อระหว่างหมวกและเส้นผม ค่าความเข้มแสงบริเวณนั้นจะมีค่าน้อย ขณะที่หากค่าความเข้มแสงของบริเวณนั้นมากหรือบริเวณนั้นมีความสว่าง ค่าความเข้มแสงบริเวณนั้นจะมีค่ามาก

2.1.4 การแปลงภาพสีจากแบบจำลองสี RGB ไปเป็นภาพระดับสีเทา (Gray Scale)

ภาพระดับสีเทาเป็นภาพที่แสดงค่าความเข้มของแสงในแต่ละพิกเซลของภาพ ซึ่งภาพระดับสีเทามีค่า 0 - 255 (0 คือสีดำ 255 คือสีขาว) ดังแสดงในภาพ 6 โดยแสดงภาพตัวอย่างในการแปลงภาพสีจากแบบจำลองสี RGB ไปเป็นภาพระดับสีเทา



ภาพ 6 การไล่ค่าสีของภาพระดับสีเทา

ภาพระดับสีเทาจะมีการไล่ระดับความอ่อนแก่ของสี ซึ่งอยู่ระหว่างสีขาวและสีอย่าง ต่อเนื่อง โดยที่ค่าแต่ละพิกเซลของภาพจะหมายถึงความเข้มแสงแต่ละตำแหน่งของพิกเซลที่อยู่ในรูป ระดับสีเทา การเปลี่ยนภาพจากระบบสีอาร์จีบี (RGB) เป็นระดับสีเทา (Gray Scale) โดยอาศัย สมการที่ 1

$$I_g(x, y) = 0.3R(x, y) + 0.59G(x, y) + 0.11B(x, y) \quad (1)$$

เมื่อ $I_g(x, y)$ คือ ค่าระดับสีเทา ณ พิกเซลตำแหน่ง (x, y)
 $R(x, y)$ คือ ค่าสีแดง ณ พิกเซลตำแหน่ง (x, y)
 $G(x, y)$ คือ ค่าสีเขียว ณ พิกเซลตำแหน่ง (x, y)
 $B(x, y)$ คือ ค่าสีน้ำเงิน ณ พิกเซลตำแหน่ง (x, y)

2.1.5 การแปลงภาพระดับสีเทา (Gray Level) ไปเป็นภาพไบนารี (Binary Image)

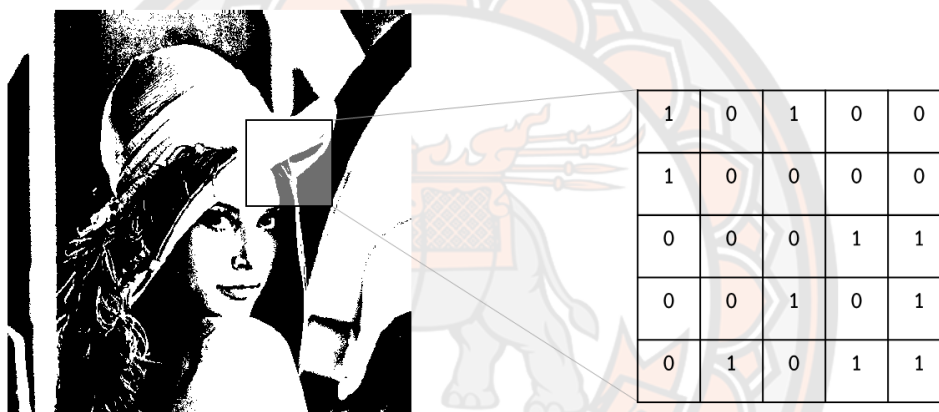
ภาพไบนารี คือภาพที่มีค่าความเข้มแสงเท่ากับ 0 หรือ 1 โดยที่ 0 คือสีดำ และ 1 คือสีขาว โดยกำหนดค่าขีดแบ่ง (Threshold) เพื่อแยกค่าความเข้มแสงของพิกเซลในภาพระดับสีเทา โดยพิกเซลที่มีมากกว่าค่าขีดแบ่งจะถูกปรับเป็น 1 และพิกเซลที่มีน้อยกว่าค่าขีดแบ่งจะถูกปรับเป็น 0 และค่าความเข้มแสงประจำพิกเซลแต่ละพิกเซลของภาพไบนารีมีการจัดเก็บโดยใช้หน่วยความจำ จำนวน 1 บิต โดยการแปลงภาพระดับสีเทาเป็นภาพไบนารีอาศัยสมการที่ 2

$$I_{bw}(x, y) = \begin{cases} 1 & \text{if } I_g(x, y) \geq T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

เมื่อ $I_{bw}(x, y)$ คือ ภาพไบนารี

$I_g(x, y)$ คือ ภาพระดับสีเทา ณ ตำแหน่ง (x, y)

ตัวอย่างภาพที่เหมาะสมในการจัดเก็บภาพแบบไบนารี ได้แก่ ภาพลายนิ้วมือ ภาพถ่ายตัวอักษรทั้งแบบที่พิมพ์ขึ้นมาหรือแบบลายมือเขียน และ ภาพถ่ายแบบเลเอาต์ของสิ่งก่อสร้าง เป็นต้น ดังแสดงในภาพ 7 แสดงตัวอย่างภาพไบนารี ที่จัดเก็บพิกเซลของวัตถุหรือพิกเซลที่เราสนใจด้วย 0 และ 1



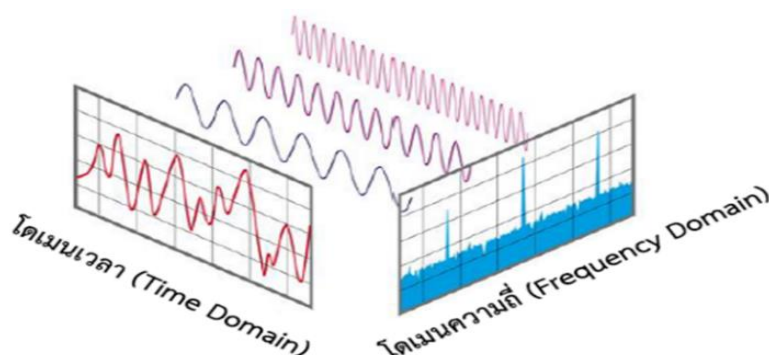
ภาพ 7 ภาพไบนารี

2.2. การแปลงไฮเปอร์คอมเพล็กซ์ฟูรีเยร์ (Hypercomplex Fourier Transform)

2.2.1 การแปลงฟูรีเยร์ (Fourier Transform)

การแปลงฟูรีเยร์ คือ เครื่องมือสำหรับการประมวลผลภาพดิจิทัลที่สำคัญซึ่งถูกใช้ในการแปลงข้อมูลจากโดเมนเวลาหรือโดเมนตำแหน่งมายังโดเมนความถี่ ซึ่งฟูรีเยร์จะทำหน้าที่ในการแปลงข้อมูลโดยกระจายข้อมูลออกเป็นความถี่ที่ต่างกัน และแยกสัญญาณให้เห็นองค์ประกอบของคลื่นความถี่ต่าง ๆ ที่ประกอบกันขึ้นมาเป็นสัญญาณนั้น ๆ ดังแสดงในภาพ 8

การแปลงฟูรีเยร์จะทำการสกัดคุณลักษณะภาพจากปริภูมิสี RGB ให้เป็นภาพระดับสีเทา ก่อนที่จะทำการแปลงภาพไปยังโดเมนความถี่ และฟูรีเยร์จะแสดงในรูปของขนาด (Magnitude) หรือสเปกตรัม (Spectrum) บนโดเมนความถี่ ดังนั้น จึงสามารถแสดงและวิเคราะห์สเปกตรัมได้บนโดเมนความถี่ (ไทยนิมิต, 2010)



ภาพ 8 กราฟของสัญญาณในโดเมนเวลาและโดเมนความถี่

ในภาพ 8 แสดงกราฟสัญญาณในโดเมนเวลาและโดเมนความถี่ โดยกราฟสัญญาณในโดเมนเวลาเมื่อพิจารณาจะเห็นว่ากราฟมีความไม่เป็นระเบียบ ยากต่อการประมวลผล จึงทำให้การวิเคราะห์สัญญาณในโดเมนเวลาทำได้ยาก แต่เมื่อแปลงสัญญาณดังกล่าวให้อยู่ในโดเมนความถี่ด้วยการแปลงฟูเรียร์ จากกราฟจะเห็นได้ว่าสัญญาณดังกล่าวประกอบด้วยคลื่นความถี่หลักสามคลื่น ความถี่ที่มีขนาดแตกต่างกัน ทำให้สามารถวิเคราะห์และประมวลผลสัญญาณนี้ได้ง่ายในโดเมนความถี่

2.2.2 การแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform)

เป็นการแปลงข้อมูลไปยังโดเมนความถี่เช่นเดียวกันกับฟูเรียร์ แต่ฟูเรียร์รับข้อมูลได้เพียงมิติเดียว คือ ภาพระดับสีเทา ไม่สามารถรับข้อมูลที่เป็น 3 มิติ แบบภาพปริภูมิ RGB ได้ ไฮเปอร์คอมเพล็กซ์ฟูเรียร์จะนำคุณลักษณะของภาพ ได้แก่ สี ที่ได้จากการสกัดออกจากรูปภาพมา แปลงไปบนโดเมนความถี่ โดยในงานวิจัยนี้จะทำการสกัดจากปริภูมิสี RGB และ Lab ตามวิธีการ HFT (J. Li et al., 2013) เพื่อแปลงคุณลักษณะของสีไปยังโดเมนความถี่และนำไปทำการวิเคราะห์ในขั้นตอนต่อไป โดยผลลัพธ์ที่ได้จากการแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์แสดงดังภาพ 9 สามารถแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ได้โดยอาศัยสมการที่ 3 – 7

$$F_h = HFT(f_h) \quad (3)$$

เมื่อ	F_h	คือ เมทริกซ์ไฮเปอร์คอมเพล็กซ์บนโดเมนความถี่
	HFT	คือ ไฮเปอร์คอมเพล็กซ์ฟูเรียร์ทรานฟอร์ม
	f_h	คือ ข้อมูลที่อยู่บนโดเมนตำแหน่ง

$$f_h = (0.5 \times I) + (0.25 \times RG) + (0.25 \times BY) \quad (4)$$

เมื่อ	I	คือ ภาพระดับสีเทา
-------	-----	-------------------

RG	คือ ความแตกต่างระหว่างสีแดงและสีเขียว
BY	คือ ความแตกต่างระหว่างสีน้ำเงินและสีเหลือง

$$A(u, v) = ||F_h(u, v)|| \quad (5)$$

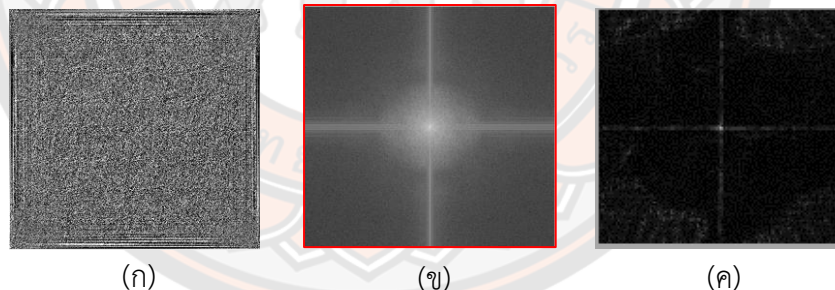
เมื่อ $A(u, v)$	คือ ขนาดของข้อมูลบนโดเมนความถี่
$ \cdot $	คือ ตัวดำเนินการสำหรับการหาขนาดของเมทริกซ์ไฮเปอร์คอมเพล็กซ์

$$P(u, v) = \tan^{-1} \left(\frac{F_{h,v}(u, v)}{F_{h,s}(u, v)} \right) \quad (6)$$

เมื่อ $P(u, v)$	คือ ระยะความถี่ของข้อมูล
$F_{h,v}(u, v)$	คือ ส่วนเวกเตอร์เมทริกซ์ไฮเปอร์คอมเพล็กซ์บนโดเมนความถี่
$F_{h,s}(u, v)$	คือ ส่วนจำนวนปริมาณเมทริกซ์ไฮเปอร์คอมเพล็กซ์บนโดเมนความถี่

$$X(u, v) = \frac{F_{h,v}(u, v)}{||F_{h,v}(u, v)||} \quad (7)$$

เมื่อ $X(u, v)$	คือ ควอเทอร์เนียนเมทริกซ์
-----------------	---------------------------



ภาพ 9 ผลลัพธ์จากการแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์

จากภาพ 9 แสดงถึงผลลัพธ์ของการแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ โดยภาพ 9 (ก) คือ ภาพระยะความถี่ของข้อมูล ภาพ 9 (ข) คือ ภาพขนาดของข้อมูล และภาพ 9 (ค) คือ ควอเทอร์เนียนเมทริกซ์ (Quaternion matrix)

2.3 การตรวจจับบริเวณที่น่าสนใจของภาพ (Saliency Detection)

การตรวจจับบริเวณที่น่าสนใจ เป็นการประมวลผลภาพเพื่อค้นหาส่วนต่าง ๆ ในภาพที่มีลักษณะเด่นที่สุดที่เป็นบริเวณที่ดึงดูดสายตามนุษย์ โดยวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ

สามารถแบ่งออกได้เป็นสองกลุ่ม ได้แก่ อ้างอิงตามคุณสมบัติบนโดเมนตำแหน่งและอ้างอิงตามสเปกตรัมบนโดเมนความถี่

2.3.1 อ้างอิงตามคุณสมบัติบนโดเมนตำแหน่ง

โดเมนตำแหน่งหรือโดเมนเวลาเป็นโดเมนที่เรามองเห็นภาพในลักษณะปกติซึ่งการตรวจจับบริเวณที่น่าสนใจของภาพโดยวิธีการอ้างอิงตามคุณสมบัติบนโดเมนตำแหน่งต้องการความรู้ก่อนหน้า (Prior knowledge) หรือกล่าวได้ว่าต้องการชุดภาพฝึกสอน (Judd, Ehinger, Durand, & Torralba, 2009) และวิธีการอ้างอิงตามคุณสมบัติบนโดเมนตำแหน่งต้องอาศัยการคำนวณโดยใช้คุณสมบัติพื้นฐานของภาพ (Low level feature) อาทิเช่น สี ลายพื้นผิว หรือ มุมในภาพ เพื่อตรวจจับ และตัดแบ่งส่วนที่เป็นวัตถุหรือบริเวณที่น่าสนใจที่สุดในภาพ (Itti et al., 1998) การตรวจจับบริเวณที่น่าสนใจของภาพโดยวิธีการอ้างอิงตามคุณสมบัติบนโดเมนตำแหน่ง หากไม่มีการใช้ความรู้ก่อนหน้าหรือชุดภาพฝึกสอน อาจทำให้ผลลัพธ์ที่ได้หรือภาพบริเวณที่น่าสนใจที่ได้มีความผิดพลาดและไม่สมบูรณ์ได้

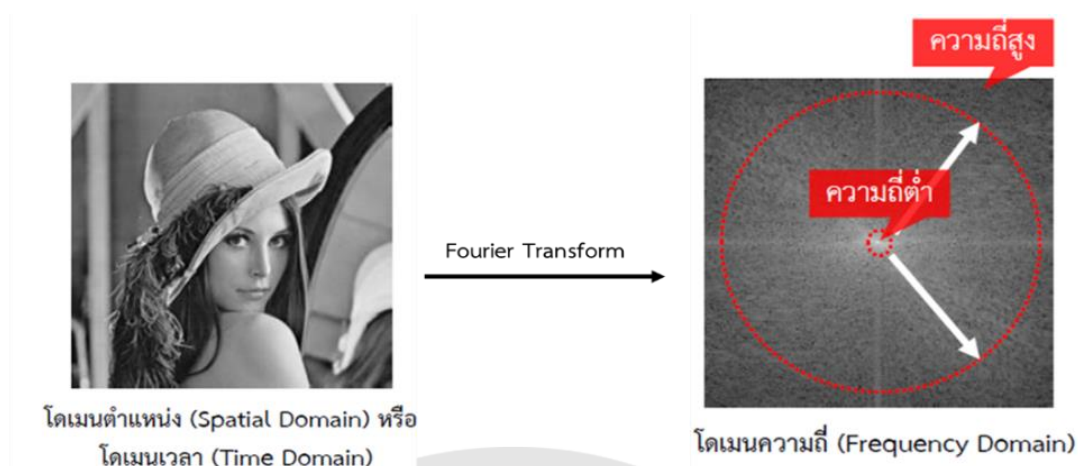
2.3.2 อ้างอิงตามสเปกตรัมบนโดเมนความถี่

โดเมนความถี่จะประกอบด้วยสัญญาณที่เกี่ยวข้องกับความถี่ โดยโดเมนความถี่จะมีรูปแบบการจัดเรียงข้อมูลตามความถี่ และบนโดเมนความถี่จะบ่งบอกถึงข้อมูลบนโดเมนตำแหน่งที่สอดคล้องและเกี่ยวข้องกับความถี่ ณ ตำแหน่งนั้น การแปลงข้อมูลจากโดเมนตำแหน่งมายังโดเมนความถี่จะใช้เครื่องมือหรือตัวแปลงข้อมูลที่เรียกว่า การแปลงฟูเรียร์ ซึ่งการแปลงฟูเรียร์จะทำการกระจายข้อมูลภาพออกเป็นความถี่ที่แตกต่างกันไป โดยโดเมนความถี่สามารถแบ่งเป็นคุณลักษณะของความถี่ได้ดังนี้

2.3.2.1 คลื่นความถี่

ข้อมูลภาพประกอบด้วยความถี่ สามารถออกได้เป็น 2 กลุ่ม คือ ความถี่ต่ำ (Low Frequency) คือ บริเวณที่ข้อมูลไม่มีการเปลี่ยนแปลง ได้แก่ บริเวณที่เป็นเนื้อสีบนโดเมนตำแหน่ง

ความถี่สูง (High Frequency) คือ บริเวณที่ข้อมูลมีการเปลี่ยนแปลงมาก ได้แก่ เส้นขอบ ลวดลายและรูปร่างบนโดเมนตำแหน่ง ดังแสดงในภาพ 10 แสดงถึงตำแหน่งของความถี่ต่ำและความถี่สูงเมื่อภาพถูกแปลงจากโดเมนตำแหน่งมายังโดเมนความถี่ จากภาพสังเกตเห็นได้ว่า ความถี่จะถูกจัดเรียงจากความถี่ต่ำไปยังความถี่สูงหรือกล่าวได้ว่าความถี่จะถูกจัดเรียงจากกึ่งกลางของภาพไปยังขอบภาพ



ภาพ 10 ตำแหน่งของความถี่

2.3.2.2 ความถี่ในการเกิด

ความถี่ในการเกิดน้อย คือ เกิดรูปแบบของข้อมูลนั้นซ้ำเป็นจำนวนน้อย

ความถี่ในการเกิดมาก คือ เกิดรูปแบบของข้อมูลนั้นซ้ำเป็นจำนวนมากหรือกล่าวได้ว่าเกิดรูปแบบซ้ำซ้ำ (Repeated patterns) ของข้อมูลภาพ

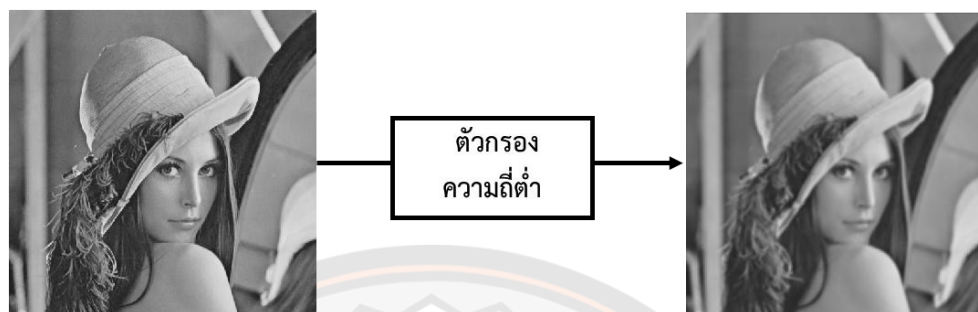
วิธีการในกลุ่มแบบอ้างอิงสเปกตรัมบนโดเมนโดเมนความถี่ไม่อาศัยความรู้ก่อนหน้าหรือชุดภาพฝึกสอนเหมือนวิธีการในกลุ่มอ้างอิงคุณสมบัติบนโดเมนความถี่ โดยวิธีการกลุ่มอ้างอิงคุณสมบัติบนโดเมนความถี่จะอาศัยสเปกตรัมบนโดเมนความถี่เพื่อใช้ในการตรวจจับบริเวณที่น่าสนใจของภาพ (Hou et al., 2007) เพราะรูปภาพที่นำเข้ามาประมวลผลจะแบ่งออกเป็นสองส่วนหลัก ๆ คือ ส่วนบริเวณที่เกิดรูปแบบซ้ำซ้ำหรือส่วนที่เป็นบริเวณที่ไม่น่าสนใจ (Non-saliency) และส่วนที่เป็นบริเวณที่น่าสนใจ บริเวณที่เกิดรูปแบบซ้ำซ้ำในโดเมนตำแหน่งจะมีตำแหน่งอยู่บนโดเมนความถี่ จึงทำให้สเปกตรัมมีความสูง หรืออาจกล่าวได้ว่าทำให้เกิดยอดแหลม (Spike) บนโดเมนความถี่

2.4 การกรองภาพ (Image Filtering)

การกรองภาพ คือ การนำภาพไปผ่านตัวกรองสัญญาณเพื่อให้ได้ภาพผลลัพธ์ออกมา ภาพผลลัพธ์ที่ได้จะมีคุณสมบัติแตกต่างจากภาพเริ่มต้น วัตถุประสงค์หลักของการกรองข้อมูลภาพคือการเน้น (Enhance) หรือลดทอน (Attenuate) คุณสมบัติบางประการของภาพ เพื่อให้ได้ภาพที่มีคุณสมบัติตามที่ต้องการ โดยตัวอย่างของตัวกรองมีดังต่อไปนี้

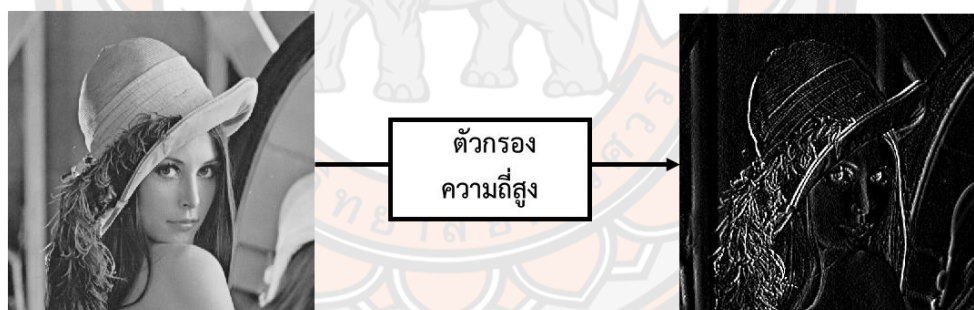
2.4.1 ตัวกรองความถี่ต่ำ (Low-pass filter) เป็นตัวกรองภาพโดยให้บริเวณที่มีความถี่ต่ำสามารถผ่านไปได้นั้น และบริเวณที่มีความถี่สูงจะถูกกรองให้ลดความถี่ลง ใช้สำหรับการลดสัญญาณรบกวนในภาพ ทำให้รายละเอียดเล็ก ๆ น้อย ๆ ภายในภาพหลังจากการกรองถูกกำจัด

ออกไป ขอบของวัตถุภายในภาพจะมีความพร่ามัวไม่คมชัดเท่าเดิม จึงสามารถใช้ตัวกรองชนิดนี้ทำให้ภาพเรียบ (Smoothing filter) หรือภาพเบลอได้ ดังแสดงในภาพ 11



ภาพ 11 ผลลัพธ์เมื่อภาพผ่านตัวกรองความถี่ต่ำ

2.4.2 ตัวกรองความถี่สูง (High-pass filter) เป็นตัวกรองภาพโดยให้บริเวณที่มีความถี่สูงสามารถผ่านไปได้เท่านั้น ซึ่งจะทำให้ภาพเกิดความคมชัดมากขึ้น และสามารถนำไปใช้เน้นส่วนของขอบภาพ หรือหาขอบภาพได้ ดังแสดงในภาพ 12



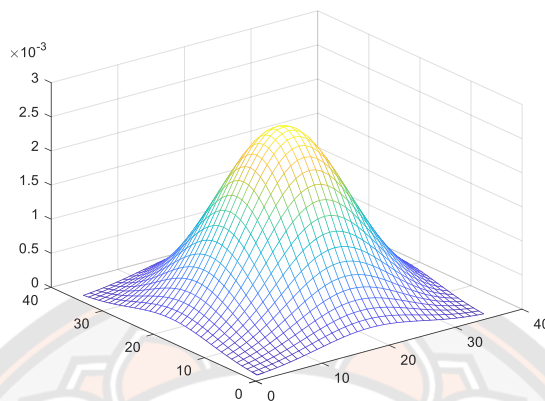
ภาพ 12 ผลลัพธ์เมื่อภาพผ่านตัวกรองความถี่สูง

2.4.3 ตัวกรองสัญญาณเกาส์เซียน (Gaussian filter) เป็นตัวกรองความถี่ต่ำที่มีคุณลักษณะรูปทรงคล้ายระฆังคว่ำ ดังแสดงในภาพ 13 เมื่อภาพผ่านตัวกรองชนิดนี้ภาพที่ได้จะพร่ามัวและมีความคมชัดน้อยลง สามารถใช้สำหรับลดสัญญาณรบกวน (Noise) ภายในภาพได้ โดยอาศัยสมการตัวกรองสัญญาณเกาส์เซียนที่ 8

$$G(u, v, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(u^2+v^2)}{2\sigma^2}} \quad (8)$$

เมื่อ $G(u, v, \sigma)$ คือ ตัวกรองสัญญาณเกาส์เซียน

(u, v) คือ ตำแหน่งที่พิจารณาในการกรองบนภาพ
 σ คือ ส่วนเบี่ยงเบนมาตรฐาน



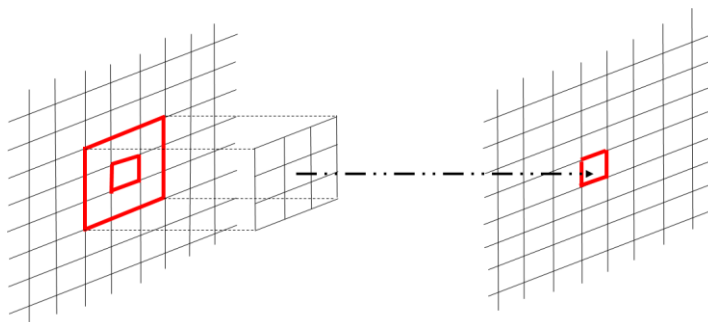
ภาพ 13 ตัวกรองสัญญาณเกาส์เซียนแบบ 2 มิติ

2.4.4 ตัวดำเนินการแบบคอนโวลูชัน (Convolution Operation Filtering)

คือ ตัวดำเนินการเชิงเส้นสำหรับประมวลผลสัญญาณ ซึ่งเป็นกระบวนการกรองภาพ โดยการเฉลี่ยค่าของพิกเซล (Pixel) และกระทำการกรองทุกพิกเซลทั่วทั้งภาพ ดังแสดงในภาพ 14 ซึ่งการคอนโวลูชันอาศัยสมการที่ 9

$$h(x, y) = f(x, y) * g(x, y) \quad (9)$$

เมื่อ $h(x, y)$ คือ ภาพผลลัพธ์จากการคอนโวลูชัน
 $f(x, y)$ คือ ภาพนำเข้า
 $g(x, y)$ คือ เมทริกซ์ขนาด $n \times n$
 $*$ คือ ตัวดำเนินการคอนโวลูชัน



ภาพ 14 แสดงการทำงานของคอนโวลูชัน

ภาพ 14 แสดงการทำงานของคอนโวลูชัน โดยขั้นตอนของการคอนโวลูชัน ประกอบด้วยการเลื่อน บวก และคูณ ในทุก ๆ พิกเซลของภาพ เราสามารถใช้การคอนโวลูชันในการประมวลผลภาพได้ในหลายลักษณะ เช่น กรองสัญญาณภาพ การหาขอบภาพ หรือการหารูปทรงของวัตถุในภาพ เป็นต้น

2.5 งานวิจัยที่เกี่ยวข้อง

Itti และคณะ ได้นำเสนอวิธีตรวจจับบริเวณที่น่าสนใจของภาพตามหลักการอ้างอิงคุณสมบัติบนโดเมนตำแหน่งและใช้หลักการของความสนใจในการมองเห็นและกระบวนการในการมองเห็นของมนุษย์ (Itti et al., 1998) วิธีการนี้จะคำนวณบริเวณที่น่าสนใจของภาพโดยการจำแนกภาพออกเป็นสามส่วนที่ต่างกัน คือ สี ความเข้มของแสง และทิศทาง (Orientation) ซึ่งสีจะประกอบด้วย 2 คู่สี ได้แก่ สีแดงและสีเขียว สีน้ำเงินและสีเหลือง ความเข้มของแสงประกอบด้วยภาพระดับสีเทา และทิศทางประกอบด้วยกำหนดยาทิศทางเชิงพื้นที่ที่ร่วมกับมุม 0, 45, 90 และ 135 องศา และได้ทำการแบ่งสเกลออกเป็น 9 สเกล ด้วยตัวกรองกาบอร์ร่วมกับการสุ่มตัวอย่างหลังจากนั้นทำการสกัดหาคุณลักษณะของภาพด้วยการพิจารณาขอบ ๆ จุดกึ่งกลาง (Center-Surround) ของสเกล การพิจารณาขอบ ๆ จุดกึ่งกลางของสเกลนั้นจะเป็นตัวกำหนดความแตกต่างระหว่างสเกลที่สามารถนำไปประมวลผลต่อในกระบวนการต่อไปและสเกลที่ไม่สามารถนำไปประมวลผลต่อได้ จากการดำเนินการเหล่านี้จะทำให้เกิด 42 คุณลักษณะ ได้แก่ ความเข้มของแสง 6 คุณลักษณะ สี 12 คุณลักษณะ และทิศทาง 24 คุณลักษณะ จากนั้นจะทำการรวมคุณลักษณะทั้งสามคุณลักษณะเข้าด้วยกันและสกัดออกมาซึ่งจะทำให้เกิดภาพที่มีจุดเด่นชัดเจนหรือกล่าวได้ว่าทำให้เกิดภาพที่เป็นบริเวณที่น่าสนใจ

Zhang และคณะ (Zhang et al., 2008) ได้นำเสนอวิธีการ A Bayesian framework for saliency using natural statistics: SUN โดยกำหนดให้การตรวจจับบริเวณที่น่าสนใจของภาพตามหลักการอ้างอิงคุณสมบัติบนโดเมนตำแหน่งเป็นการรวมกันของสามองค์ประกอบ องค์ประกอบแรกเป็นข้อมูลของภาพโดยทำการพิจารณาคุณลักษณะเฉพาะของภาพแค่หนึ่งตำแหน่งในภาพเท่านั้นทำให้คุณสมบัติเฉพาะถือว่าเป็นส่วนประกอบของบริเวณที่น่าสนใจ องค์ประกอบที่สองเป็นชุดข้อมูลของคุณลักษณะเชิงเส้น องค์ประกอบที่สามประกอบด้วยชุดข้อมูลของความน่าจะเป็นที่เกี่ยวข้องกับตำแหน่งของลักษณะเด่นในภาพ หรือกล่าวได้ว่าองค์ประกอบที่สองและสามจะใช้ข้อมูลชุดภาพฝึกสอน และบริเวณที่น่าสนใจของภาพจะถูกคำนวณโดยใช้ความแตกต่างของตัวกรองสัญญาณเกาส์เซียน (Difference of Gaussians: DOG) ร่วมกับทั้งสามองค์ประกอบ

Hou และคณะ (Hou & Zhang, 2007) ได้นำเสนอวิธีการสเปกตรัมส่วนที่เหลือตามหลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ ซึ่ง Hou และคณะได้อธิบายว่ารูปภาพที่นำเข้ามาประมวลผลจะแบ่งออกเป็นสองส่วนหลัก ๆ คือ ส่วนบริเวณที่เกิดรูปแบบซ้ำซ้ำ (Repeated patterns) หรือส่วนที่

เป็นบริเวณที่ไม่น่าสนใจ (Non-saliency) และส่วนที่เป็นบริเวณที่น่าสนใจ โดยบริเวณที่เกิดรูปแบบซ้ำซ้ำในโดเมนตำแหน่งจะมีตำแหน่งอยู่บนโดเมนความถี่ จึงเกิดการสะสมของสเปกตรัมทำให้สเปกตรัมมีความสูงหรือเรียกอีกอย่างว่าทำให้เกิด Spike ในโดเมนความถี่ Hou และคณะได้นำแนวคิดนี้ไปประยุกต์ใช้ร่วมกับการตรวจจับบริเวณที่น่าสนใจของภาพโดยใช้ตัวกรองสัญญาณเกาส์เซียนในการประมาณสเปกตรัมของบริเวณที่เกิดรูปแบบซ้ำซ้ำ ซึ่งวิธีการ คือ แปลงข้อมูลจากโดเมนตำแหน่งไปอยู่บนโดเมนความถี่โดยใช้เครื่องมือแปลงข้อมูลฟูเรียร์ จากนั้นนำภาพไปหาค่าเฉลี่ยของสเปกตรัมด้วยตัวกรองสัญญาณเกาส์เซียน จากนั้นนำภาพต้นฉบับที่แปลงเป็นสัญญาณความถี่กลับกับภาพที่หาค่าเฉลี่ยของสเปกตรัม โดยจะได้ผลลัพธ์เป็นส่วนที่เหลือของสเปกตรัม เมื่อแปลงส่วนที่เหลือของสเปกตรัมกลับมาโดเมนตำแหน่งจะได้ผลลัพธ์เป็นภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ

Jain และคณะ (J. Li et al., 2013) ได้นำเสนอการวิเคราะห์สเปกตรัมหลายระดับบนโดเมนความถี่ตามหลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ โดยใช้การแปลงฟูเรียร์ไฮเปอร์คอมเพล็กซ์ (HFT) ร่วมด้วยกระบวนการสกัดคุณลักษณะของสีออกจากภาพเพื่อให้ได้ผลต่างของสี ได้แก่ R-G B-Y และ I โดย R คือ สีแดง G คือ สีเขียว B คือ สีน้ำเงิน Y คือ สีเหลือง และ I คือ ความเข้มของแสงหรือภาพระดับสีเทา โดยอาศัยสมการที่ 10 - 12

$$I = \frac{r + g + b}{3} \quad (10)$$

เมื่อ	I	คือ ภาพระดับสีเทา
	r	คือ ส่วนสีแดงของแบบจำลองสี RGB
	g	คือ ส่วนสีเขียวของแบบจำลองสี RGB
	b	คือ ส่วนสีน้ำเงินของแบบจำลองสี RGB

$$RG = \left(r - \left(\frac{g + b}{2} \right) \right) - \left(g - \left(\frac{r + b}{2} \right) \right) \quad (11)$$

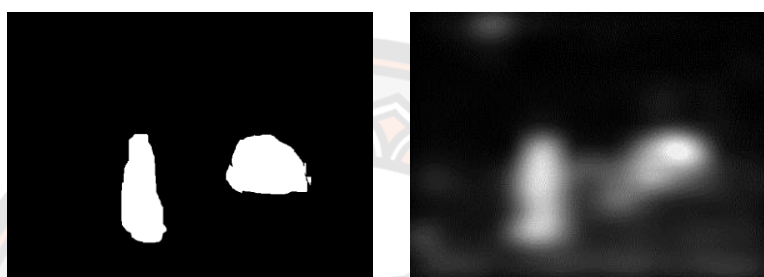
เมื่อ	RG	คือ ความแตกต่างระหว่างสีแดงและสีเขียว
-------	------	---------------------------------------

$$BY = \left(b - \left(\frac{r + g}{2} \right) \right) - \left(\left(\frac{g + r}{2} \right) \right) - \left(\frac{|r - g|}{2} \right) - b \quad (12)$$

เมื่อ	BY	คือ ความแตกต่างระหว่างสีน้ำเงินและสีเหลือง
-------	------	--

และนำคุณลักษณะของสีที่สกัดออกมาไปทำการแปลงข้อมูลจากโดเมนตำแหน่งไปยังโดเมนความถี่โดยใช้การแปลงฟูเรียร์ไฮเปอร์คอมเพล็กซ์และใช้คุณสมบัติของสเปกตรัมที่ได้จากการแปลง

คุณลักษณะของสี โดยทำการระงับหรือกดทับสเปกตรัมที่เป็นรูปแบบที่ซ้ำ ๆ ด้วยตัวกรองสัญญาณเกาส์เซียน 8 ระดับ เมื่อแปลงสเปกตรัมที่ถูกกรอง 8 ระดับ กลับมายังโดเมนตำแหน่งจะใช้เกณฑ์เอนโทรปีหรือเกณฑ์วัดความปะปนกันของข้อมูลภาพที่มีค่าต่ำที่สุดเพื่อสกัดให้ได้ภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ แต่เมื่อนำวิธีการนี้มาใช้งานกับข้อมูลภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกันและภาพที่มีพื้นหลังซับซ้อน พบว่าให้ผลลัพธ์เป็นภาพบริเวณที่น่าสนใจที่มีพื้นหลังหรือบริเวณอื่นปะปนอยู่ด้วยในภาพ แสดงดังภาพ 15



(ก) ภาพเฉลี่ย

(ข) ภาพผลลัพธ์

ภาพ 15 ตัวอย่างผลลัพธ์การตรวจจับบริเวณที่น่าสนใจโดยวิธีการของ Jain

ภาพ 15 แสดงถึงผลลัพธ์ของการตรวจจับบริเวณที่น่าสนใจของภาพโดยวิธีการของ Jain ภาพ (ก) คือ ภาพเฉลี่ยและภาพ (ข) คือ ภาพผลลัพธ์ของวิธีการนี้ เมื่อพิจารณาภาพผลลัพธ์จะพบว่าผลลัพธ์ของวิธีการนี้ยังมีพื้นหลังหรือบริเวณอื่นปะปนอยู่กับบริเวณที่น่าสนใจของภาพ

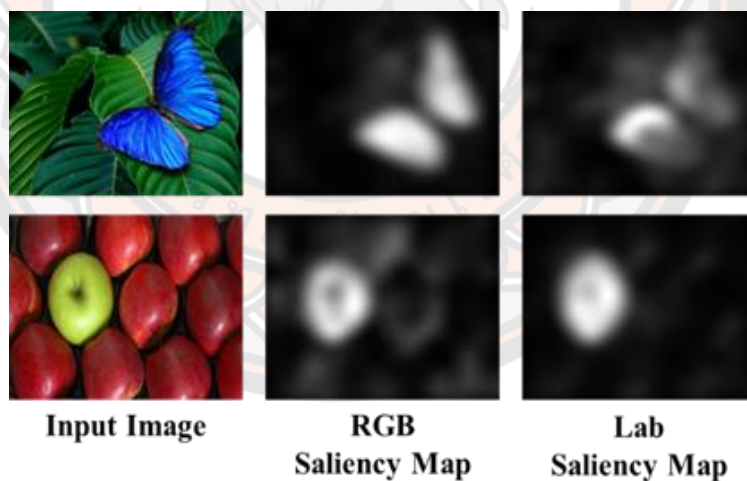
Jaemsiri และคณะ (Jaemsiri et al., 2019) ได้นำเสนอการตรวจจับบริเวณที่น่าสนใจของภาพตามหลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ โดยได้พัฒนาและปรับปรุงมาจากวิธีการ HFT ของ (J. Li et al., 2013) ในส่วนของกระบวนการกรองสเปกตรัม ซึ่งจะทำการกรองสเปกตรัมด้วยตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะ ซึ่งตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะได้มีแนวคิดมาจากตัวกรองจุดกึ่งกลางแบบปรับเหมาะ (Adaptive Median filter) (Chan, Chung-Wa, & Nikolova, 2005) โดยตัวกรองจุดกึ่งกลางแบบปรับเหมาะเป็นตัวกรองทางสถิติที่สามารถลดอิทธิพลของสัญญาณรบกวนในภาพ โดยอาศัยการเลือกค่ากลางของข้อมูลหรือค่าสถิติจากบริเวณนั้น ๆ ซึ่ง Jaemsiri และคณะได้มีแนวคิดที่ว่า สเปกตรัมที่เกิดรูปแบบซ้ำซ้ำในภาพมีความคล้ายคลึงกันกับสัญญาณรบกวน ในวิธีการนี้จึงได้นำตัวกรองจุดกึ่งกลางแบบปรับเหมาะมาประยุกต์ใช้ร่วมกับตัวกรองสัญญาณเกาส์เซียน เพื่อทำการระงับรูปแบบที่เกิดซ้ำซ้ำของสเปกตรัมหรือบริเวณที่เป็นพื้นหลังของรูปภาพ

Xiang และคณะ (Xiang & Zhong, 2015) ได้ทำการตรวจจับบริเวณที่น่าสนใจของภาพตามหลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ โดยการนำปริภูมิสี Lab และ ปริภูมิสี RGB มาใช้ร่วมกับวิธีการ HFT ของ (J. Li et al., 2013) ที่ทำการกรองสเปกตรัมด้วยเกาส์เซียนหลายระดับ โดย Xiang และคณะ เชื่อว่าบริเวณที่เป็นวัตถุเด่นสามารถปรากฏขึ้นในพื้นที่อื่นสีอื่น ๆ ได้ นอกจากปริภูมิสี RGB เช่น ปริภูมิสี HSV และปริภูมิ Lab เป็นต้น

งานวิจัยของ Xiang ได้นำปริภูมิสี Lab มาใช้กับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ เพราะ Xiang ได้ทดลองและค้นพบบริเวณที่มีลักษณะเด่นหรือบริเวณที่น่าสนใจของภาพในปริภูมิสี Lab ดังแสดงในภาพ 16 โดยการแปลงอาศัยสมการที่ 13

$$f(x, y) = w_1 f_1 i + w_2 f_2 j + w_3 f_3 k \quad (13)$$

เมื่อ $f_1 = L, f_2 = a, f_3 = b$ คือ คุณลักษณะของสี
 $w_1, \dots, w_3$ คือ ค่าน้ำหนักของคุณลักษณะของสี โดยมีน้ำหนักเท่ากันทั้งหมด
 i, j, k คือ พิกัดพิกัดในภาพ



ภาพ 16 ผลลัพธ์ของภาพบริเวณที่น่าสนใจของภาพที่แสดงในปริภูมิสี RGB และ Lab

ภาพ 16 แสดงให้เห็นถึงบริเวณที่เป็นลักษณะเด่นสามารถปรากฏได้ชัดเจนในปริภูมิสีที่ต่างกัน ดังภาพจะสังเกตได้ว่า ภาพผีเสื้อปรากฏเด่นชัดในปริภูมิสี RGB และภาพผลแอปเปิ้ลปรากฏเด่นชัดในปริภูมิสี Lab

บทที่ 3

วิธีดำเนินงานวิจัย

ในการดำเนินการทำงานวิจัยการปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองเกาส์เซียนหลายระดับแบบปรับเหมาะสำหรับการตรวจจับบริเวณที่น่าสนใจของภาพ เพื่อทำการพัฒนาวิธีการการตรวจจับบริเวณที่น่าสนใจของภาพ โดยมีรายละเอียดวิธีดำเนินการวิจัย ดังนี้

3.1 การเตรียมข้อมูลภาพ

3.2 การออกแบบและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 โดยผู้วิจัยได้ทำการพัฒนามาจากวิธีการ HFT (J. Li et al., 2013)

3.3 การออกแบบและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

3.4 การประเมินประสิทธิภาพของการตรวจจับบริเวณที่น่าสนใจของภาพ

3.1 การเตรียมข้อมูลภาพ

การพัฒนาระบบการตรวจจับบริเวณที่น่าสนใจของภาพนั้น ในงานวิจัยนี้ได้ใช้ฐานข้อมูล Salmap (Li et al., 2013) เป็นฐานข้อมูลมาตรฐาน โดยฐานข้อมูลมีภาพทั้งหมด 235 ภาพ ตัวอย่างของภาพแสดงดังภาพ 17 – 22 ซึ่งแบ่งเป็นหมวดหมู่ทั้ง 6 ประเภท



ภาพ 17 ภาพบริเวณที่น่าสนใจมีขนาดใหญ่



ภาพ 18 ภาพบริเวณที่น่าสนใจมีขนาดปานกลาง



ภาพ 19 ภาพบริเวณที่น่าสนใจมีขนาดเล็ก



ภาพ 20 ภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน



ภาพ 21 ภาพพื้นหลังมีความซับซ้อน



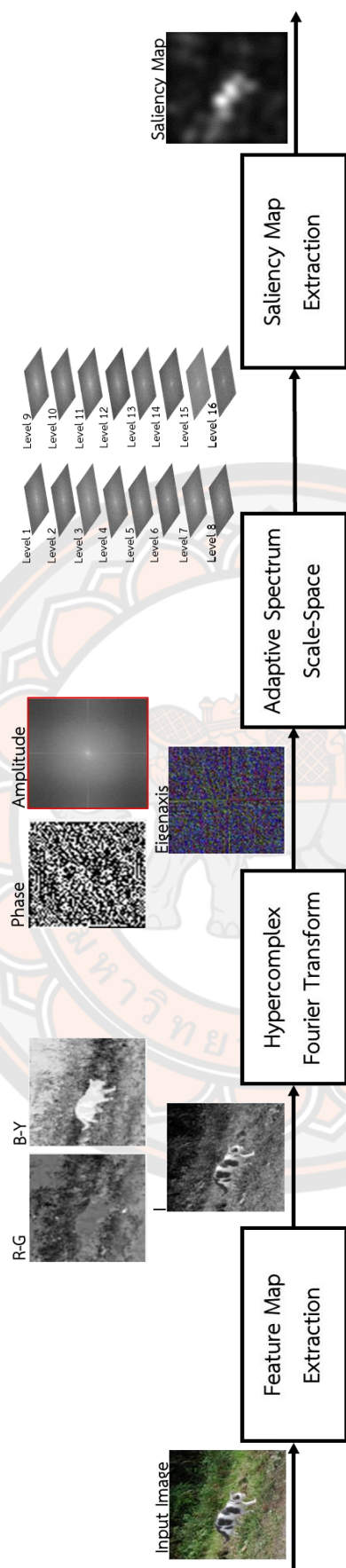
ภาพ 22 ภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่

3.2 การออกแบบและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1

ในการพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ได้ทำการปรับปรุงมาจากวิธีการ HFT (J. Li et al., 2013) ถึงแม้ว่าวิธีการ HFT จะประสบผลสำเร็จในการตรวจจับบริเวณที่น่าสนใจของภาพ แต่อย่างไรก็ตาม ประสิทธิภาพการตรวจจับบริเวณที่น่าสนใจของภาพและการกรองสเปกตรัมก็ขึ้นอยู่กับความเหมาะสมของตัวกรองที่ใช้ไม่สามารถปรับได้โดยอัตโนมัติ ภาพ 23 แสดงกระบวนการในการทำงานของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 หรือวิธีการแบบที่ 1 ที่ได้รับการปรับปรุงมาจากวิธีการ HFT โดยการแทนที่จากตัวกรองสัญญาณเกาส์เซียนเดิมเป็นตัวกรองสัญญาณเกาส์เซียนที่ปรับพื้นที่ตัวกรองแบบอัตโนมัติเพื่อกรองสเปกตรัมที่เหมาะสมสำหรับแต่ละบริเวณในแต่ละภาพบนโดเมนความถี่ วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 สามารถเลือกตัวกรองที่เหมาะสมสำหรับแต่ละพื้นที่ของแต่ละภาพโดยอัตโนมัติตามวิธีการแก้ปัญหาจากการวิเคราะห์ทางสถิติ การตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 มี 4 ขั้นตอน ดังนี้

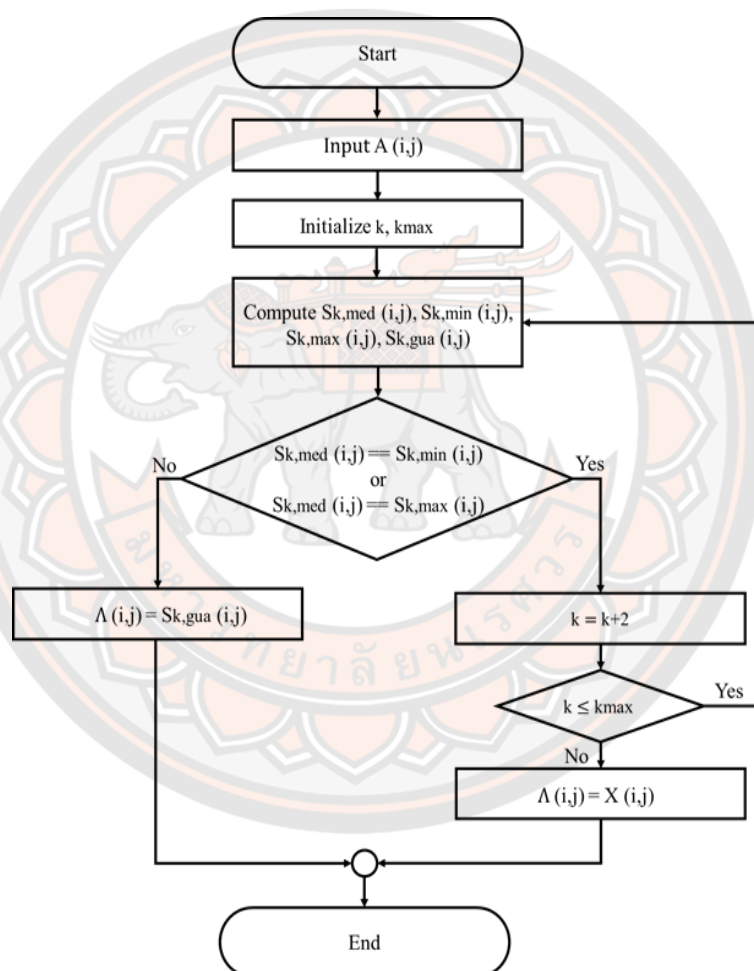
3.2.1 การสกัดคุณลักษณะของภาพ (Feature Map Extraction) เป็นขั้นตอนในการแยกหรือสกัดคุณลักษณะของภาพ โดยการสกัดคุณลักษณะนั้น สามารถทำการสกัดคุณลักษณะของภาพออกมาได้เป็น สี เส้นขอบ ทิศทาง มุม หรือ องศา แต่ในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ที่ได้พัฒนามาจากวิธีการ HFT (J. Li et al., 2013) จะทำการสกัดคุณลักษณะของภาพออกมาเป็นคุณลักษณะของปริภูมิสี RGB โดยจะแบ่งออกเป็น 3 ชนิด ด้วยกัน คือ I, RG, และ BY โดยที่ RG คือ ความแตกต่างระหว่างสีแดงและสีเขียว BY คือ ความแตกต่างระหว่างสีน้ำเงินและสีเหลือง และ I คือ ภาพระดับสีเทาหรือความเข้มของแสง (Intensity) (Itti et al., 1998) โดยการสกัดคุณลักษณะภาพเพื่อให้คุณลักษณะของสีทำได้โดยอาศัยสมการที่ 10 – 12 ดังแสดงในหัวข้อ 2.5 งานวิจัยที่เกี่ยวข้องในบทที่ 2

3.2.2 การแปลงฟูเรียร์ไฮเปอร์คอมเพล็กซ์ (Hypercomplex Fourier Transform) เป็นขั้นตอนในการแปลงข้อมูลภาพจากโดเมนตำแหน่งให้ไปอยู่บนโดเมนความถี่ โดยใช้เครื่องมือในการแปลง คือ ไฮเปอร์คอมเพล็กซ์ อาศัยสมการที่ 3 – 7 ดังแสดงในหัวข้อ 2.2.2 การแปลงไฮเปอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform) ในบทที่ 2



ภาพ 23 ภาพรวมการทำงานของขั้นตอนการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1

3.2.3 การพัฒนาสเปกตรัมหลายระดับแบบปรับเหมาะ (Adaptive Spectrum Scale-Space) เป็นขั้นตอนในการกรองสเปกตรัมของบริเวณที่เกิดการทำซ้ำของข้อมูล โดยวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ได้แนวคิดตัวกรองสัญญาณเกาส์เซียนหลายระดับมาจากตัวกรองสัญญาณเกาส์เซียนมัลติสเกลแบบปรับเหมาะ (Chan et al., 2005) ซึ่งเป็นตัวกรองที่มีประสิทธิภาพในการกำจัดสัญญาณรบกวนออกจากภาพ ผู้วิจัยได้มีแนวคิดที่ว่า สเปกตรัมของบริเวณที่เกิดการทำซ้ำมีความคล้ายคลึงกับสัญญาณรบกวนในภาพ โดยการพัฒนาสเปกตรัมหลายระดับแบบปรับเหมาะซึ่งกระบวนการทำงาน ดังแสดงในภาพ 24



ภาพ 24 กระบวนการทำงานการกรองสเปกตรัมของวิธีการแบบที่ 1

ภาพ 24 แสดงกระบวนการทำงานของการกรองสเปกตรัม โดยกำหนดให้ A คือ สเปกตรัม G_k คือ ตัวกรองสัญญาณเกาส์เซียน และ Λ คือ สเปกตรัมที่ถูกกรองโดยตัวกรองสัญญาณเกาส์เซียน ซึ่งสามารถอธิบายรายละเอียดของกระบวนการทำงานได้ดังนี้

1) **ขั้นตอนเริ่มต้นการทำงาน (Initialize step)** กำหนดค่าเริ่มต้นให้พารามิเตอร์ k ของตัวกรองสัญญาณเกาส์เซียน เท่ากับ 3 จากนั้นสร้างตัวกรองสัญญาณเกาส์เซียน โดยอาศัยสมการที่ 8 ดังแสดงในหัวข้อ 2.4.3 ตัวกรองสัญญาณเกาส์เซียน (Gaussian filter) ในบทที่ 2

2) **ขั้นตอนการค้นหาค่าสูงสุด ต่ำสุด และค่ามัธยฐานของสเปกตรัม (Find maximum, minimum, and median step)** ทำการคำนวณค่าสูงสุด ต่ำสุด และค่ามัธยฐานของสเปกตรัม ดังสมการที่ 14 – 16

$$S_{k,max}(i,j) = \max \left\{ \begin{array}{l} A(a,b) | -\frac{(i-k)}{2} \leq a \leq +\frac{(i+k)}{2} \\ \text{and } \frac{(j-k)}{2} \leq b \leq \frac{(j+k)}{2} \end{array} \right\} \quad (14)$$

$$S_{k,min}(i,j) = \min \left\{ \begin{array}{l} A(a,b) | -\frac{(i-k)}{2} \leq a \leq +\frac{(i+k)}{2} \\ \text{and } \frac{(j-k)}{2} \leq b \leq \frac{(j+k)}{2} \end{array} \right\} \quad (15)$$

$$S_{k,med}(i,j) = \text{median} \left\{ \begin{array}{l} A(a,b) | -\frac{(i-k)}{2} \leq a \leq +\frac{(i+k)}{2} \\ \text{and } \frac{(j-k)}{2} \leq b \leq \frac{(j+k)}{2} \end{array} \right\} \quad (16)$$

โดยที่ $S_{k,max}(i,j)$ คือ ค่าสูงสุดของสเปกตรัม ณ ตำแหน่ง (i,j)
 $S_{k,min}(i,j)$ คือ ค่าต่ำสุดของสเปกตรัม ณ ตำแหน่ง (i,j)
 $S_{k,med}(i,j)$ คือ ค่ามัธยฐานของสเปกตรัม ณ ตำแหน่ง (i,j)

3) **ขั้นตอนการปรับส่วนเบี่ยงเบนมาตรฐาน (Adjusting sigma step)** กำหนดให้พารามิเตอร์ k เพิ่มค่าครั้งละ 2 เมื่อค่ามัธยฐานของสเปกตรัมมีค่าเท่ากับค่าสูงสุดและค่าต่ำสุดของสเปกตรัม และพารามิเตอร์ k น้อยกว่าค่า k_{max} ให้กลับไปทำยังขั้นตอน 2) ขั้นตอนการค้นหาค่าสูงสุด ต่ำสุด และค่ามัธยฐานของสเปกตรัม

4) **ขั้นตอนในการกรองสเปกตรัม (Suppressing amplitude spectrum step)** กำหนดให้กรองสเปกตรัมเมื่อค่ามัธยฐานของสเปกตรัมไม่เท่ากับค่าสูงสุดและค่าต่ำสุดของสเปกตรัม โดยอาศัยสมการที่ 17

$$S_{k,gua} = G_k * A \quad (17)$$

เมื่อ $S_{k,qua}$ คือ สเปกตรัมที่ผ่านการกรองโดยตัวกรองสัญญาณเกาส์เซียน หรือสามารถกล่าวได้ว่าเป็นตัวเลือกของภาพบริเวณที่น่าสนใจ

3.2.4 การสกัดหาบริเวณที่น่าสนใจของภาพ (Saliency Map Extraction) เป็นกระบวนการแยกหรือสกัดเอาข้อมูลที่สำคัญของภาพออกมาโดยใช้เกณฑ์วัดความปะปนกันของข้อมูลภาพหรือเกณฑ์เอนโทรปีที่มีค่าน้อยที่สุด เพื่อให้ได้ภาพผลลัพธ์ที่แสดงเฉพาะบริเวณที่น่าสนใจในการสกัดหาบริเวณที่น่าสนใจของภาพใช้เกณฑ์เอนโทรปีที่มีค่าน้อยที่สุดเป็นตัวตัดสินใจเพราะว่าเกณฑ์เอนโทรปีจะเป็นเครื่องมือวัดความปะปนกันของข้อมูล หากเกณฑ์เอนโทรปีมีค่าที่สูง ภาพผลลัพธ์ที่แสดงเฉพาะบริเวณที่น่าสนใจจะมีพื้นหลังมาปะปนมาก ดังนั้นจึงใช้เกณฑ์เอนโทรปีที่มีค่าต่ำในการเลือกภาพผลลัพธ์ที่แสดงเฉพาะบริเวณที่น่าสนใจ โดยอาศัยสมการที่ 18 -19

$$S_k = g_\sigma * ||IHFT(\Lambda e^{XP})||^2 \quad (18)$$

เมื่อ g_σ คือ ตัวกรองสัญญาณเกาส์เซียนที่มีส่วนเบี่ยงเบนมาตรฐานมีค่าเป็น 0.05 (J. Li et al., 2013)

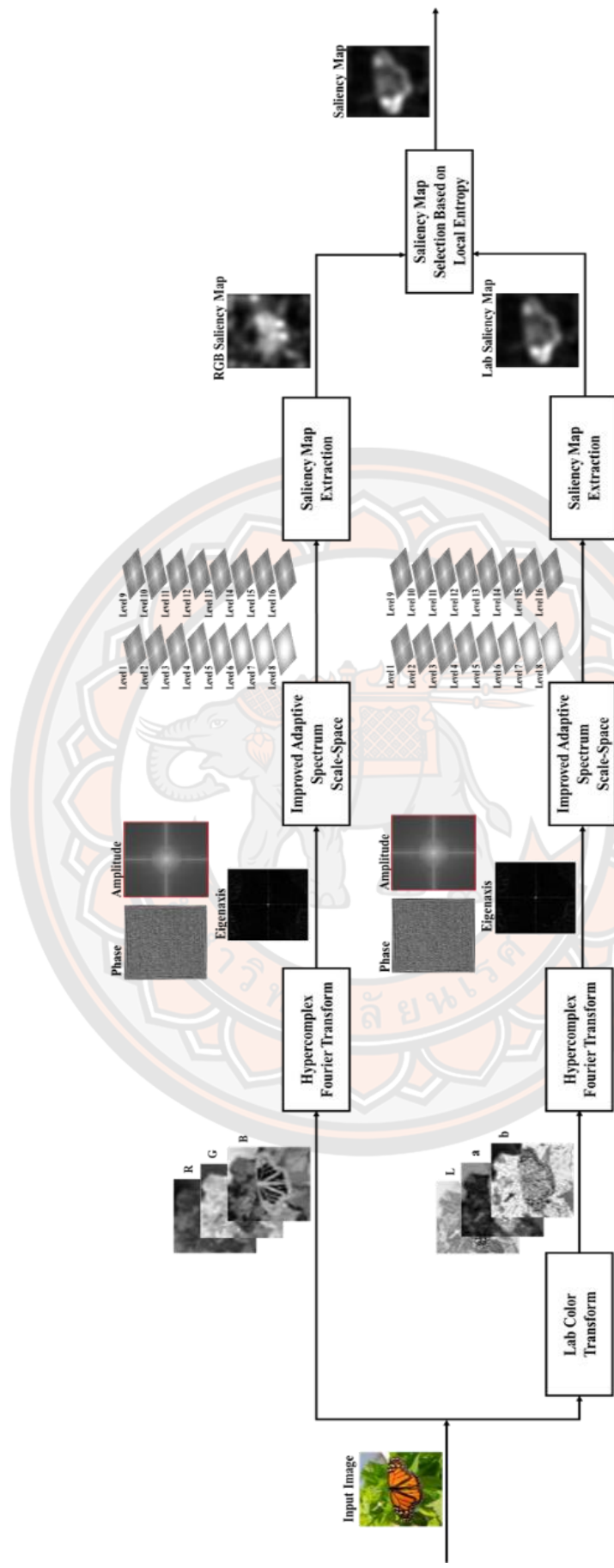
$IHFT(\cdot)$ คือ การแปลงฟูเรียร์แบบย้อนกลับ

$$S = \arg \min_k (S_k) \quad (19)$$

เมื่อ S คือ ภาพบริเวณที่น่าสนใจที่เหมาะสมที่สุด

3.3 การออกแบบและพัฒนาระบบการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

การออกแบบและพัฒนาระบบการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 โดยทำการปรับปรุงมาจากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 โดยผู้วิจัยได้เล็งเห็นว่า ถึงแม้วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 จะมีประสิทธิภาพในการตรวจจับบริเวณที่น่าสนใจของภาพ แต่การที่มีพารามิเตอร์ในตัวกรองสัญญาณเกาส์เซียนคงที่ เช่น ขนาดของตัวกรองสัญญาณเกาส์เซียนคงที่ ซึ่งไม่สามารถปรับขนาดเพิ่มลดอัตโนมัติในแต่ละบริเวณของภาพ อาจทำให้ไม่สามารถกรองบริเวณที่เกิดข้อมูลซ้ำซ้ำได้อย่างมีประสิทธิภาพ และการใช้เกณฑ์เอนโทรปีแบบเดิมที่อาจจะนำมาซึ่งปัญหาในการเลือกภาพผลลัพธ์ที่แสดงเฉพาะบริเวณที่น่าสนใจที่ไม่ถูกต้อง จากปัญหาเหล่านี้ผู้วิจัยจึงได้ปรับปรุงและพัฒนาระบบแบบที่ 2 ขึ้นมา ซึ่งประกอบด้วย 5 ขั้นตอน แสดงดังภาพ 25 โดยแสดงถึงภาพรวมการทำงานทั้งหมดของระบบ



ภาพ 25 ภาพรวมการทำงานของโครงข่ายประสาทเทียมที่นำสนใจของภาพแบบที่ 2

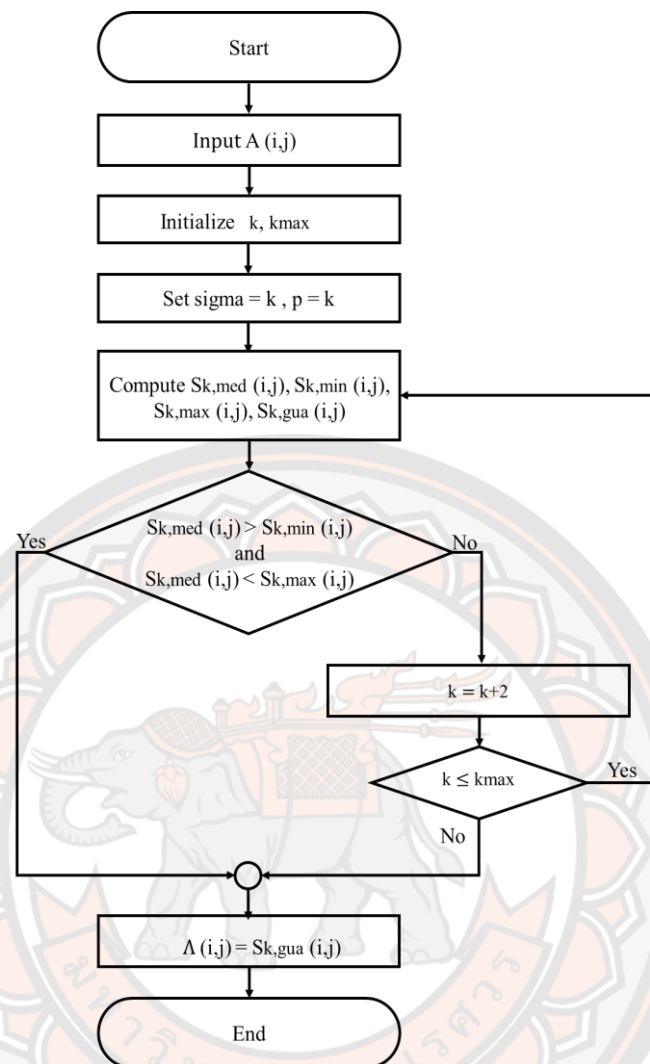
3.3.1 การแปลงภาพสี Lab (Lab color Transform) เป็นขั้นตอนการแปลงภาพสี Lab ซึ่งบริเวณที่เป็นลักษณะเด่นของภาพนั้นสามารถปรากฏได้ทั้งในปริภูมิสี RGB และ Lab ขึ้นอยู่กับลักษณะของภาพแต่ละภาพ (Xiang & Zhong, 2015) โดยอาศัยสมการที่ 13 ดังแสดงในหัวข้อ 2.5 งานวิจัยที่เกี่ยวข้องของในบทที่ 2

3.3.2 การแปลงฟูเรียร์ไฮเพอร์คอมเพล็กซ์ (Hypercomplex Fourier Transform) เป็นกระบวนการที่ใช้สำหรับการแปลงข้อมูลจากโดเมนเวลาไปอยู่บนโดเมนความถี่ ดังแสดงในสมการที่ 3 – 7 ดังแสดงในหัวข้อ 2.2.2 การแปลงไฮเพอร์คอมเพล็กซ์ฟูเรียร์ (Hypercomplex Fourier Transform) ในบทที่ 2

3.3.3 การปรับปรุงสเปกตรัมหลายระดับแบบปรับเหมาะ (Improved Adaptive Spectrum Scale-Space) เป็นกระบวนการในการระงับ (Suppress) สเปกตรัมที่มีรูปแบบซ้ำซ้ำของรูปภาพออก โดยได้ทำการปรับปรุงมาจากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบ 1 ซึ่งตัวกรองในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 มีประสิทธิภาพในการกรองภาพ แต่อย่างไรก็ตาม ตัวกรองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ยังคงต้องกำหนดขนาดของตัวกรองให้คงที่ซึ่งตัวกรองมีขนาดเท่ากับภาพ อีกทั้งตัวกรองไม่สามารถขยายขนาดไปในแต่ละบริเวณของภาพได้อัตโนมัติ ผู้วิจัยจึงมีแนวคิดว่าจะทำให้ไม่สามารถกรองข้อมูลบริเวณที่เกิดรูปแบบซ้ำซ้ำในแต่ละบริเวณของภาพได้อย่างมีประสิทธิภาพ ดังนั้น ผู้วิจัยจึงมีแนวคิดว่าจะหาตัวกรองสามารถขยายขนาดและทำการกรองในแต่ละบริเวณของภาพได้อย่างอัตโนมัติ ประสิทธิภาพของกระบวนการตรวจจับบริเวณที่น่าสนใจของภาพน่าจะเพิ่มมากขึ้นกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 เพราะตัวกรองจะสามารถกรองสเปกตรัมที่มีรูปแบบซ้ำซ้ำในแต่ละบริเวณของภาพได้ด้วยตัวกรองหลายขนาด ซึ่งไม่เหมือนกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ที่ตัวกรองภาพจะกรองทั้งภาพได้เพียงขนาดเดียว ผู้วิจัยจึงเพิ่มในส่วนของการขยายขนาดของตัวกรองแบบอัตโนมัติ ซึ่งจากเดิมในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบตรวจจับบริเวณที่น่าสนใจของภาพที่ 1 มีเพียงการปรับพื้นที่แบบอัตโนมัติเพียงอย่างเดียว โดยการปรับปรุงสเปกตรัมหลายระดับแบบปรับเหมาะ แสดงกระบวนการดังภาพ 26 ซึ่งมีขั้นตอนดังต่อไปนี้

1) ขั้นตอนเริ่มต้นการทำงาน (Initialize step) เป็นการกำหนดค่าเริ่มต้นให้พารามิเตอร์ k เท่ากับ 3 (Jaemsiri et al., 2019)

2) ขั้นตอนการอัปเดตพารามิเตอร์ (Update Gaussian parameter step) เป็นการกำหนดพารามิเตอร์ของตัวกรองสัญญาณเกาส์เซียน σ และ ρ ให้เท่ากับ k โดยที่กำหนดให้ k น้อยกว่าหรือเท่ากับ k_{max} และกำหนดให้ σ คือ ส่วนเบี่ยงเบนมาตรฐาน และ ρ คือ ขนาดของตัวกรองสัญญาณเกาส์เซียน



ภาพ 26 กระบวนการทำงานการกรองสเปกตรัมของวิธีการแบบที่ 2

3) ขั้นตอนการสร้างตัวกรองสัญญาณเกาส์เซียน (Generate Gaussian filter step) เป็นการสร้างตัวกรองสัญญาณเกาส์เซียนโดยใช้สมการที่ 8 ดังแสดงในหัวข้อ 2.4.3 ตัวกรองสัญญาณเกาส์เซียน (Gaussian filter) ในบทที่ 2

4) ขั้นตอนการกรองสเปกตรัม (Suppress amplitude spectrum step) เป็นการทำการคำนวณหาสมาชิกของสเปกตรัม $\tilde{A}(u, v)$ โดยอาศัยสมการที่ 20

$$\tilde{A}(u, v) = A(u, v) * G(u, v, \sigma) \quad (20)$$

โดยที่ $A(u, v)$ คือ ขนาดของข้อมูลบนโดเมนความถี่หรือแอมปริจูดสเปกตรัม
 $G(u, v, \sigma)$ คือ ตัวกรองสัญญาณเกาส์เซียน

5) ขั้นตอนการค้นหาค่าสูงสุด ต่ำสุด และค่ามัธยฐานของสเปกตรัม (Find maximum, minimum, and median step) ทำการคำนวณค่าสูงสุด ต่ำสุด และค่ามัธยฐานของสเปกตรัม ดังสมการที่ 21 - 23

$$s_{k,max}(u, v) = \max \left\{ \begin{array}{l} A(a, b) | u - \frac{(\rho - 1)}{2} \leq a \leq u + \frac{(\rho - 1)}{2} \\ \text{and } v - \frac{(\rho - 1)}{2} \leq b \leq v + \frac{(\rho - 1)}{2} \end{array} \right\} \quad (21)$$

$$s_{k,min}(u, v) = \min \left\{ \begin{array}{l} A(a, b) | u - \frac{(\rho - 1)}{2} \leq a \leq u + \frac{(\rho - 1)}{2} \\ \text{and } v - \frac{(\rho - 1)}{2} \leq b \leq v + \frac{(\rho - 1)}{2} \end{array} \right\} \quad (22)$$

$$s_{k,med}(u, v) = \text{median} \left\{ \begin{array}{l} A(a, b) | u - \frac{(\rho - 1)}{2} \leq a \leq u + \frac{(\rho - 1)}{2} \\ \text{and } v - \frac{(\rho - 1)}{2} \leq b \leq v + \frac{(\rho - 1)}{2} \end{array} \right\} \quad (23)$$

โดยที่ $s_{k,max}(u, v)$ คือ ค่าสูงสุดของสเปกตรัม ณ ตำแหน่ง (u, v)
 $s_{k,min}(u, v)$ คือ ค่าต่ำสุดของสเปกตรัม ณ ตำแหน่ง (u, v)
 $s_{k,med}(u, v)$ คือ ค่ามัธยฐานของสเปกตรัม ณ ตำแหน่ง (u, v)

6) ขั้นตอนการกำหนดสเปกตรัมที่ถูกกรองแล้ว (Define the smooth amplitude spectrum step) ถ้า $s_{k,med}$ มีค่ามากกว่า $s_{k,min}$ และ $s_{k,max}$ และมีค่าน้อยกว่า k_{max} จะทำการกำหนดให้ตัวกรองสัญญาณเกาส์เซียนมีขนาดเท่าเดิม ในขณะที่หากพารามิเตอร์ k เพิ่มขึ้นครึ่งละ 2 จะทำการกลับไปยังขั้นตอน 3) เพื่อสร้างตัวกรองสัญญาณเกาส์เซียนจนกระทั่งพารามิเตอร์ k เท่ากับ k_{max} อย่างไรก็ตาม หากพารามิเตอร์ k มีค่ามากกว่า k_{max} สเปกตรัมจะถูกกรองด้วยตัวกรองสัญญาณเกาส์เซียนที่มีค่าพารามิเตอร์เป็น k_{max} ลำดับสุดท้าย จากภาพ 28 จะสามารถเห็นได้อย่างชัดเจนว่าสเปกตรัมที่ได้จากวิธีการที่นำเสนอมีความเรียบมากกว่าวิธีการ HFT และ วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ซึ่งแสดงให้เห็นว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 สามารถกรองสเปกตรัมที่เป็นพื้นหลังบริเวณที่ไม่สำคัญออกได้อย่างมีประสิทธิภาพ

3.2.4 การสกัดหาบริเวณที่น่าสนใจของภาพ (Saliency Map Extraction หรือ SME)

เป็นกระบวนการแยกหรือสกัดเอาข้อมูลที่สำคัญของภาพออกมาโดยใช้เกณฑ์วัดความปะปนกันของข้อมูลภาพหรือเกณฑ์เอนโทรปีแบบท้องถิ่นที่มีค่าน้อยที่สุด (Minimum local entropy) เพื่อให้

ภาพได้ผลลัพธ์ที่แสดงเฉพาะส่วนของสิ่งเด่น โดยตัวเลือกภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ S ได้รับจากการแปลงสเปกตรัมที่ถูกกรองโดยตัวกรองสัญญาณเกาส์เซียน Λ ซึ่งแปลงกลับจากโดเมนความถี่มายังโดเมนตำแหน่งโดยใช้ไฮเปอร์คอมเพล็กซ์ฟูเรียร์ผกผัน

$$S = \{S_k^G, S_{kmax}^A\} \quad (24)$$

เมื่อ S_k^G และ S_{kmax}^A คือ ตัวเลือกของภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ ซึ่งกำหนดไว้ดังสมการที่ 25 - 27

$$S_k^G = SME(\Lambda_k^G) \quad (25)$$

$$S_{kmax}^A = SME(\Lambda_{kmax}^A) \quad (26)$$

$$SME(\Lambda) = g_\sigma * ||IHFT(\Lambda e^{XP})||^2 \quad (27)$$

เมื่อ g_σ คือ ตัวกรองสัญญาณเกาส์เซียนที่มีส่วนเบี่ยงเบนมาตรฐานมีค่าเป็น 0.05
 Λ_k^G คือ สเปกตรัมที่ถูกกรองด้วยตัวกรองสัญญาณเกาส์เซียนที่มีขนาด k เมื่อ k มีค่า เป็น 1,2,3,...,8 (J. Li et al., 2013)
 Λ_{kmax}^A คือ สเปกตรัมที่ถูกกรองด้วยตัวกรองสัญญาณเกาส์เซียน AS^2G ที่มีขนาด k เมื่อ k มีค่าเป็น 3,5,7,...,17

เกณฑ์เอนโทรปีที่มีค่าต่ำสุดใช้งานในการเลือกภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ ในแต่ละปริภูมิสี เพื่อหาภาพที่แสดงเฉพาะบริเวณที่น่าสนใจที่เหมาะสม S' โดยอาศัยสมการที่ 28 และ 29

$$S' = \arg \min \{H_{Entropy}(S)\} \quad (28)$$

เมื่อ $H_{Entropy}(\cdot)$ คือ ฟังก์ชันเอนโทรปีของภาพบริเวณที่เป็นลักษณะเด่น
 S' คือ ภาพบริเวณที่น่าสนใจที่ถูกเลือก

$$H_{Entropy}(\cdot) = - \sum_{i=1}^n h_i \log_2 h_i \quad (29)$$

เมื่อ h_i คือ ฮิสโตแกรมในรูปแบบมาตรฐาน โดย i เป็นความถี่สะสมของโอกาสในการเกิดข้อมูล

3.2.5 การเลือกบริเวณที่น่าสนใจของภาพโดยใช้เกณฑ์เอนโทรปีแบบท้องถิ่น (Saliency Map Selection Based on Local Entropy) ในการเลือกภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ นั้นหากเป็นวิธีแบบเดิมที่ใช้เกณฑ์เอนโทรปีจะทำให้เกิดการเลือกภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ ที่ไม่ถูกต้อง แสดงดังภาพ 27 จะเห็นได้ว่าภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ โดยใช้เกณฑ์เอนโทรปี แบบท้องถิ่นใกล้เคียงกับภาพเฉลี่ย แต่ภาพที่แสดงเฉพาะบริเวณที่น่าสนใจที่ได้จากเกณฑ์เอนโทรปี แบบท้องถิ่นไม่ถูกเลือกเมื่อใช้วิธีการแบบเดิม เพื่อปรับปรุงการเลือกภาพที่แสดงเฉพาะบริเวณที่ น่าสนใจที่มีความเหมาะสมที่สุด S_{final} ในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 จึงได้ เลือกใช้เกณฑ์เอนโทรปีแบบท้องถิ่นมาช่วย โดยอาศัยสมการที่ 30 - 32

$$S_{final} = \arg \min \{H_{LE}(S'_{RGB}), H_{LE}(S'_{Lab})\} \quad (30)$$

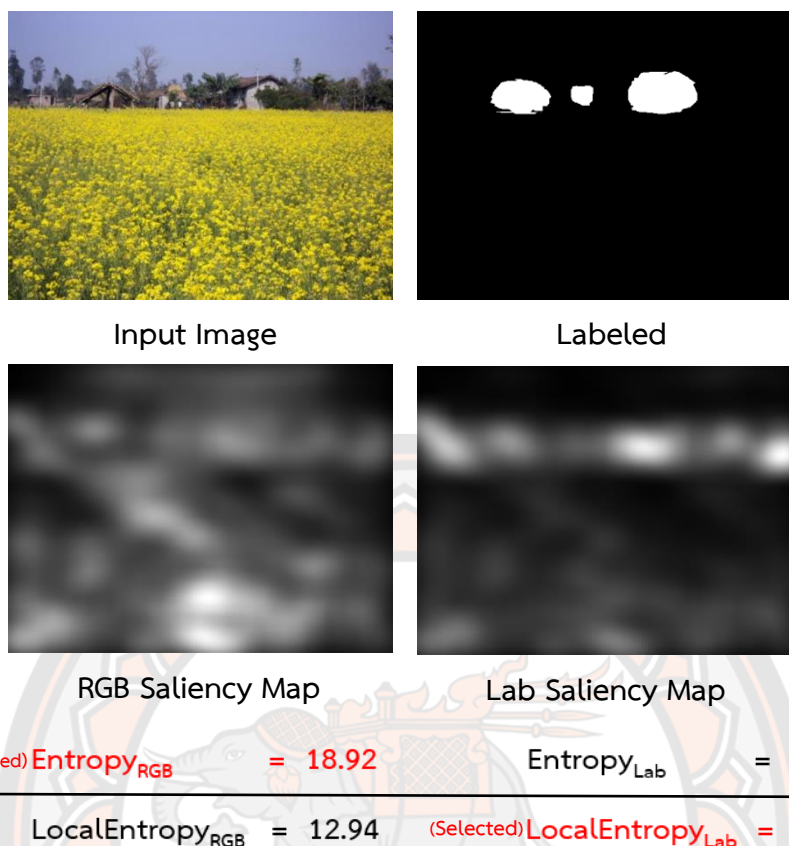
เมื่อ $H_{LE}(\cdot)$ คือ ฟังก์ชันเอนโทรปีแบบท้องถิ่น
 S'_{RGB} คือ ภาพบริเวณที่เป็นลักษณะเด่นในปริภูมิสี RGB
 S'_{Lab} คือ ภาพบริเวณที่เป็นลักษณะเด่นในปริภูมิสี Lab

$$H_{LE} = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y H_{Entropy}(x, y) \quad (31)$$

เมื่อ X คือ ความกว้างของภาพนำเข้า
 Y คือ ความสูงของภาพนำเข้า

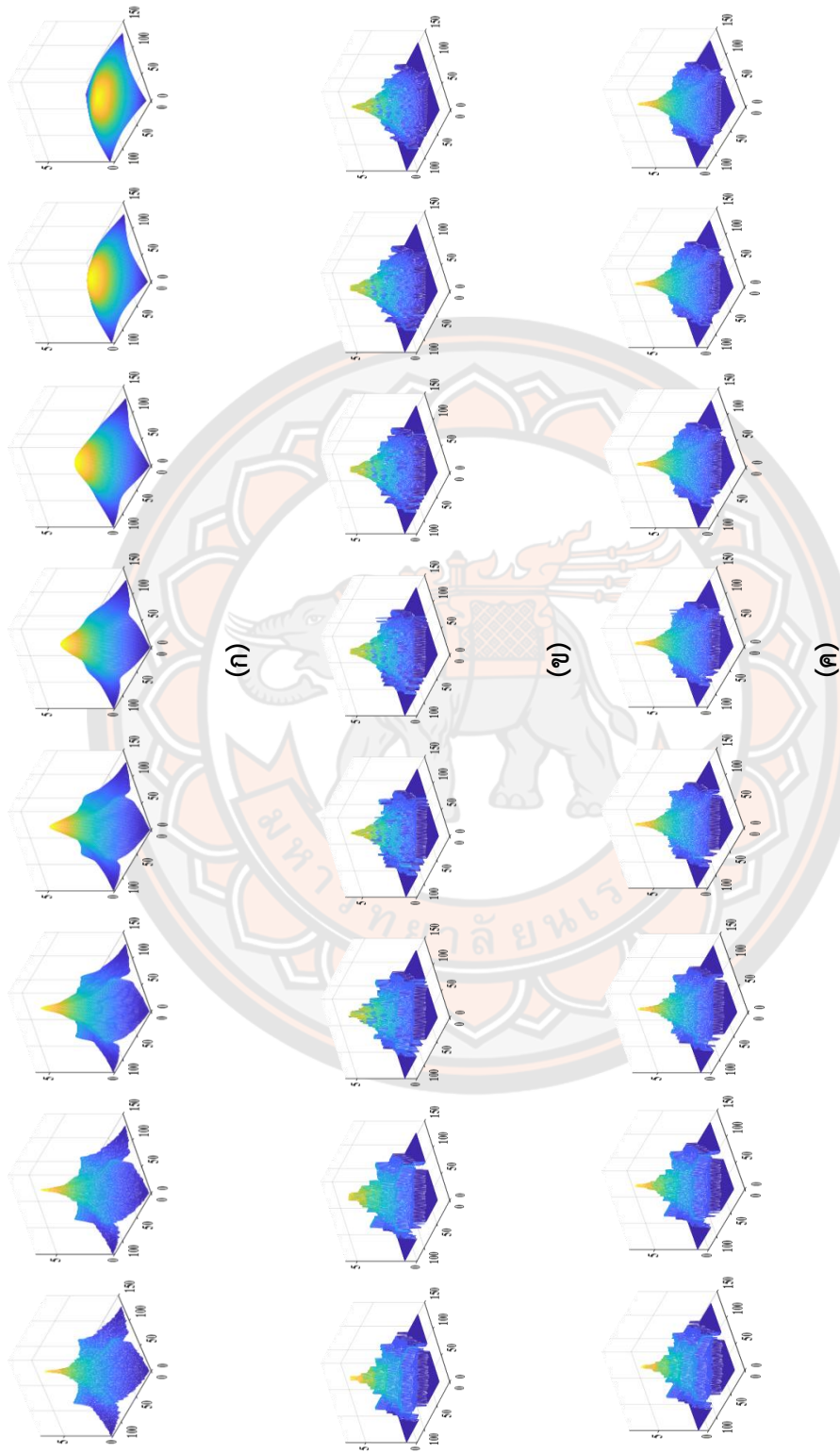
$$H_{Entropy}(x, y) = - \sum_{i=1}^{256} h_i(x, y) \times \log_2 h_i(x, y) \quad (32)$$

เมื่อ $h_i(x, y)$ คือ ฮิสโตแกรมในรูปแบบมาตรฐานโดย i เป็นความถี่สะสมของ โอกาสในการเกิดข้อมูล ณ ตำแหน่ง (x, y)



ภาพ 27 ผลลัพธ์ของภาพที่แสดงบริเวณที่น่าสนใจจากปริภูมิสี RGB และ Lab

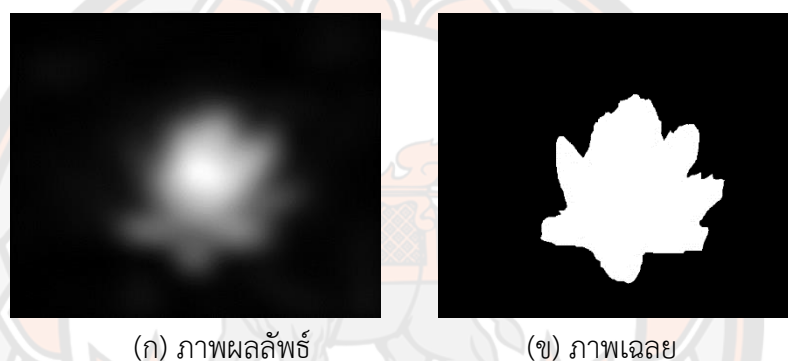
ภาพที่ 27 แสดงถึงผลลัพธ์ของภาพที่แสดงเฉพาะบริเวณที่น่าสนใจ โดยแสดงให้เห็นว่าหากใช้เกณฑ์เอนโทรปีเลือกจะได้ภาพที่แสดงบริเวณที่น่าสนใจจากปริภูมิสี RGB ซึ่งเมื่อเปรียบเทียบกับภาพเฉลี่ยแล้วพบว่าภาพที่แสดงบริเวณที่น่าสนใจจากปริภูมิสี Lab ที่ใช้เกณฑ์เอนโทรปีแบบท้องถิ่นในการคัดเลือกมีความใกล้เคียงกับภาพเฉลี่ยมากกว่า



ภาพ 28 ผลลัพธ์ของสเปกตรัมที่ได้จากวิธีการ (ก) HFT (ข) วิธีการแบบที่ 1 และ (ค) วิธีการแบบที่ 2

3.4 การประเมินประสิทธิภาพของการวิจัย

การทดสอบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพนี้ ผู้วิจัยได้ใช้วิธีวัดประสิทธิภาพของการตรวจจับบริเวณที่น่าสนใจของภาพที่นิยมในปัจจุบัน คือ กราฟความสัมพันธ์ระหว่างอัตราค่าความถูกต้องเชิงบวกกับอัตราความผิดพลาดเชิงบวก หรือเรียกอีกอย่างว่า กราฟ Receiver Operator Characteristic (ROC) และได้ใช้พื้นที่ใต้เส้นโค้ง Area Under Curve (AUC) ช่วยในการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ ซึ่งเป็นเกณฑ์ที่นิยมใช้ในการวัดประสิทธิภาพงานในด้านการตรวจจับบริเวณที่น่าสนใจของภาพ (J. Li et al., 2013), (Achanta, Hemami, Estrada, & Susstrunk, 2009) แสดงดังภาพ 29



ภาพ 29 ตัวอย่างภาพผลลัพธ์และภาพเฉลยที่นำมาใช้ในการวัดประสิทธิภาพด้วยค่า AUC

ภาพ 29 แสดงถึงภาพตัวอย่างระหว่างภาพ 29 (ก) ภาพผลลัพธ์ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพและภาพ 29 (ข) ภาพเฉลยที่มนุษย์สร้างขึ้น โดยการนำภาพผลลัพธ์หรือภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพมาทำการเปรียบเทียบกับภาพเฉลยที่มนุษย์สร้างขึ้น (Ground Truth)

3.4.1 พื้นที่ใต้เส้นโค้ง Area Under Curve (AUC)

พื้นที่ใต้เส้นโค้งจากกราฟความสัมพันธ์ระหว่างอัตราค่าความถูกต้องเชิงบวกกับอัตราความผิดพลาดเชิงบวก (ROC) ใช้สำหรับการวัดและเปรียบเทียบประสิทธิภาพในแต่ละวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ ซึ่งสามารถใช้โปรแกรมทางสถิติคำนวณได้ ค่า AUC จะเป็นตัวบ่งชี้ถึงความแม่นยำ โดยค่า AUC จะอยู่ในช่วง 0 – 1 หากค่า AUC เข้าใกล้ 1 ยิ่งแสดงถึงประสิทธิภาพของวิธีการที่สูง เพราะแสดงให้เห็นถึงความน่าจะเป็นที่ชุดข้อมูลที่ทดสอบนั้นจะให้ผลลัพธ์ใกล้เคียงและสอดคล้องกับภาพเฉลย โดยการวัดประสิทธิภาพ AUC จะอาศัยสมการที่ 33

$$AUC = \int TPR d(FPR) \quad (33)$$

โดยที่ TPR คือ อัตราค่าความถูกต้องเชิงบวก ดังสมการที่ 34
 FPR คือ อัตราค่าความผิดพลาดเชิงบวก ดังสมการที่ 35

3.4.2 กราฟ Receiver Operator Characteristic (ROC)

กราฟ ROC คือ การสร้างความสัมพันธ์ระหว่างอัตราค่าความถูกต้องเชิงบวกกับอัตราความผิดพลาดเชิงบวก แสดงดังภาพ 30 - 31 สามารถคำนวณโดยอาศัยสมการที่ 34 - 35

$$TPR = \frac{TP}{TP + FN} \quad (34)$$

$$FPR = \frac{FP}{FP + TN} \quad (35)$$

โดยที่ TP คือ ค่าความถูกต้องเชิงบวก (True Positive) ผลลัพธ์ที่ได้จากวิธีการทำนายว่าเป็นพื้นที่ของบริเวณที่น่าสนใจ แล้วเป็นพื้นที่ของบริเวณที่น่าสนใจจริง

TN คือ ค่าความถูกต้องเชิงลบ (True Negative) ผลลัพธ์ที่ได้จากวิธีการทำนายว่าเป็นพื้นที่หลังบริเวณที่น่าสนใจ แล้วเป็นพื้นที่ของพื้นที่หลังจริง

FP คือ ค่าความผิดพลาดเชิงบวก (False Positive) ผลลัพธ์ที่ได้จากวิธีการทำนายว่าเป็นพื้นที่ของบริเวณที่น่าสนใจ ความจริงแล้วเป็นพื้นที่ของพื้นที่หลัง

FN คือ ค่าความผิดพลาดเชิงลบ (False Negative) ผลลัพธ์ที่ได้จากวิธีการทำนายว่าเป็นพื้นที่ของพื้นที่หลัง แต่ความจริงเป็นพื้นที่ของบริเวณที่น่าสนใจ (Le Meur & Baccino, 2012)

TN	TN	TP	FN
TN	TP	TP	TN
FP	TP	TP	TN
TP	TP	TP	TN

(ก) ภาพผลลัพธ์

(ข) ภาพเฉลย

ภาพ 30 ตัวอย่างการเปรียบเทียบระหว่างภาพผลลัพธ์และภาพเฉลย

ภาพ 30 แสดงตัวอย่างของการเปรียบเทียบเพื่อหาความสัมพันธ์ระหว่างอัตราค่าความถูกต้องเชิงกับอัตราความผิดพลาดเชิงบวกแบบพิกเซลต่อพิกเซลระหว่างภาพ 30 (ก) คือ ภาพผลลัพธ์และภาพ 30 (ข) คือ ภาพเฉลย จากภาพแสดงให้เห็นว่าเมื่อทำการเปรียบเทียบมี 2 ตำแหน่งที่ไม่ตรงกับภาพเฉลย ได้แก่ ตำแหน่งที่เป็น FN (ความผิดพลาดเชิงลบ) ผลลัพธ์ที่ได้จากวิธีการทำนายว่าเป็นพื้นที่ของพื้นหลัง แต่ความจริงเป็นพื้นที่ของบริเวณที่น่าสนใจ และ ตำแหน่งที่เป็น FP (ความผิดพลาดเชิงบวก) ทำนายว่าเป็นพื้นที่ของบริเวณที่น่าสนใจ ความจริงแล้วเป็นพื้นที่ของพื้นหลัง

ค่าจริง (Actual value)

ค่าจากการทำนาย (Predict outcome)		Salient	Background
	Salient	8 TP	1 FP
	Background	1 FN	6 PN

ภาพ 31 ตารางแจกแจงความถูกต้องของการจำแนก (Confusion Matrix)

ภาพ 31 แสดงถึงตารางแจกแจงความถูกต้องของการจำแนก (Confusion Matrix) โดยค่าในตารางแจกแจงความถูกต้องของการจำแนกได้นำมาจากภาพ 30 ตัวอย่างของการเปรียบเทียบแบบพิกเซลต่อพิกเซลระหว่างภาพผลลัพธ์และภาพเฉลย

บทที่ 4

ผลการวิจัย

งานวิจัยเรื่องการปรับปรุงการวิเคราะห์แบบระดับบนโดเมนความถี่ด้วยตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะสำหรับการตรวจจับบริเวณที่น่าสนใจของภาพ ผู้วิจัยได้ทำการเปรียบเทียบผลการทดลอง โดยมีรายละเอียดผลการทดลองของวิจัยดังนี้

4.1 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1

4.2 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

4.3 ผลการวิเคราะห์ผลการทดลองเชิงเปรียบเทียบโดยจำแนกตามหมวดหมู่ของภาพทั้งหมด 6 ประเภท

4.4 ผลการเปรียบเทียบประสิทธิภาพด้านระยะเวลาในการประมวลผลของการตรวจจับบริเวณที่น่าสนใจของภาพ

4.1 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1

จากการทดสอบและวัดประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 โดยใช้วิธีการวัดประสิทธิภาพด้วยค่า AUC เปรียบเทียบกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพจากงานวิจัยทั้ง 5 แบบ ได้แก่ วิธีการ Itti (Itti et al., 1998), SUN (Zhang, Tong, Marks, Shan, & Cottrell, 2008), SR (Hou & Zhang, 2007), SAL (Hou et al., 2012) และ HFT (J. Li et al., 2013) แสดงผลการทดลองดังตาราง 2

ตาราง 2 เปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 กับวิธีการ Itti, SUN, SR, SAL, และ HFT ด้วยค่า AUC

วิธีการ	ประเภท 1	ประเภท 2	ประเภท 3	ประเภท 4	ประเภท 5	ประเภท 6	ภาพรวม
Itti	0.9111	0.9155	0.9349	0.8435	0.9165	0.9419	0.9147
SUN	0.7423	0.7645	0.8365	0.6812	0.7081	0.7938	0.7719
SR	0.8131	0.8512	0.9285	0.7558	0.7911	0.8996	0.8574
SAL	0.8693	0.9028	0.9320	0.8604	0.9050	0.9472	0.9037
HFT	0.9432	0.9135	0.9221	0.9440	0.9200	0.9530	0.9244
แบบที่ 1	0.9376	0.9231	0.9230	0.9564	0.9307	0.9566	0.9283

จากตาราง 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ แสดงให้เห็นว่าประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ทั้งหมดด้วยค่า AUC ในภาพรวม โดยเหตุผลหลักที่ทำให้วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 มีประสิทธิภาพดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบอื่น ๆ ได้แก่ การใช้หลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ที่สามารถนำสเปกตรัมมาใช้งานได้ทันทีโดยไม่ต้องใช้ความรู้ก่อนหน้าและชุดภาพฝึกสอน อีกทั้งมีการนำสเปกตรัมมาใช้งานร่วมกับตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะหรือตัวกรองสัญญาณเกาส์เซียน AMG ที่มีการปรับพื้นที่แบบอัตโนมัติ โดยสามารถเพิ่มขยายพื้นที่ของตัวกรองได้อย่างอัตโนมัติในแต่ละบริเวณของภาพ ซึ่งตัวกรองสัญญาณเกาส์เซียน AMG เป็นเครื่องมือในการกรองสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ จึงสามารถลบบริเวณที่ไม่เกี่ยวข้องหรือพื้นหลังของภาพออกทำให้บริเวณที่น่าสนใจของภาพปรากฏเด่นชัดขึ้น ในขณะที่วิธีการ HFT, SAL และ SR ที่ใช้หลักการอ้างอิงสเปกตรัมบนโดเมนความถี่และมีหลักการทำงานบนโดเมนความถี่เช่นเดียวกันมีประสิทธิภาพด้อยกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 1 เนื่องจากการกรองสเปกตรัมในแต่ละวิธีการได้ใช้ตัวกรองที่มีขนาดเดียวซึ่งไม่เหมาะสมกับแต่ละบริเวณของภาพ ทำให้ลบบริเวณที่ไม่เกี่ยวข้องหรือพื้นหลังของภาพได้อย่างไม่มีประสิทธิภาพและมีบริเวณอื่นปะปนอยู่ในภาพผลลัพธ์บริเวณที่น่าสนใจของภาพ ส่งผลให้ผลลัพธ์ภาพบริเวณที่น่าสนใจอาจไม่สมบูรณ์และปรากฏไม่เด่นชัด

ในขณะที่วิธีการของ Itti และ SUN เป็นวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่ได้ใช้หลักการอ้างอิงคุณสมบัติบนโดเมนตำแหน่งซึ่งมีประสิทธิภาพด้อยกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่ใช้หลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ เนื่องจากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่ใช้หลักการคุณสมบัติบนโดเมนตำแหน่งนั้น จะต้องใช้ความรู้ก่อนหน้าและชุดภาพฝึกสอนในกระบวนการตรวจจับบริเวณที่น่าสนใจของภาพโดยให้คอมพิวเตอร์ได้เรียนรู้ หากภาพทดลองภาพใด ๆ มีความแตกต่างไปจากความรู้ก่อนหน้าและชุดภาพฝึกสอน เช่น ภาพฝึกสอนเป็นภาพแมวที่มีอิริยาบถกำลังยืน แต่ภาพทดลองนั้นเป็นภาพแมวที่มีอิริยาบถกำลังนอน เป็นต้น ผลลัพธ์ภาพบริเวณที่น่าสนใจของภาพนั้น ๆ อาจมีความผิดเพี้ยนไม่สมบูรณ์และไม่ชัดเจน เพราะคอมพิวเตอร์ไม่รู้จำเนื่องจากไม่มีในความรู้ก่อนหน้าและชุดภาพฝึกสอน ซึ่งอาจส่งผลให้เกิดข้อผิดพลาดในตรวจจับบริเวณที่น่าสนใจของภาพ

4.2 ผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

จากการทดสอบและวัดประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 โดยใช้วิธีการวัดประสิทธิภาพด้วยค่า AUC เปรียบเทียบกับวิธีการตรวจจับบริเวณที่น่าสนใจจากงานวิจัยทั้ง 6 แบบ วิธีการ Itti (Itti et al., 1998), SUN (Zhang et al., 2008), SR (Hou &

Zhang, 2007), SAL (Hou et al., 2012), HFT (J. Li et al., 2013) และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 แสดงผลการทดลองดังตาราง 3

ตาราง 3 เปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 กับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, Itti, SUN, SR, SAL, และ HFT ด้วยค่า AUC

วิธีการ	ประเภท 1	ประเภท 2	ประเภท 3	ประเภท 4	ประเภท 5	ประเภท 6	ภาพรวม
Itti	0.9111	0.9155	0.9349	0.8435	0.9165	0.9419	0.9147
SUN	0.7423	0.7645	0.8365	0.6812	0.7081	0.7938	0.7719
SR	0.8131	0.8512	0.9285	0.7558	0.7911	0.8996	0.8574
SAL	0.8693	0.9028	0.9320	0.8604	0.9050	0.9472	0.9037
HFT	0.9432	0.9135	0.9221	0.9440	0.9200	0.9530	0.9244
แบบที่ 1	0.9376	0.9231	0.9230	0.9564	0.9307	0.9566	0.9283
แบบที่ 2	0.9399	0.9262	0.9308	0.9633	0.9287	0.9559	0.9311

จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแสดงให้เห็นว่าประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ทั้งหมดด้วยค่า AUC ในภาพรวม โดยเหตุผลหลักที่ทำให้ประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ได้แก่ ตัวกรองเกาส์เซียนหลายระดับแบบปรับเหมาะหรือตัวกรองสัญญาณเกาส์เซียน AS^2G ที่ปรับขนาดและพื้นที่ของตัวกรองได้อย่างอัตโนมัติต่างกับตัวกรองสัญญาณเกาส์เซียน AMG ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ที่ขยายพื้นที่อัตโนมัติได้เพียงอย่างเดียว จึงทำให้สามารถเพิ่มขยายขนาดและพื้นที่ของตัวกรองไปในแต่ละบริเวณของภาพได้อย่างอัตโนมัติส่งผลให้กรองสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ ได้อย่างมีประสิทธิภาพ อีกทั้งยังมีการใช้เกณฑ์เอนโทรปีแบบท้องถิ่นในการเลือกภาพบริเวณที่น่าสนใจที่เหมาะสมจากปริภูมิสี RGB และปริภูมิสี Lab รวมไปถึงการนำแนวคิดบริเวณที่น่าสนใจสามารถปรากฏได้ในปริภูมิสีอื่น ๆ ดังที่ผู้วิจัยได้อธิบายไว้ในบทที่ 2 หัวข้องานวิจัยที่เกี่ยวข้อง โดยได้นำปริภูมิสี Lab มาใช้ร่วมกับปริภูมิสี RGB ในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

ในการปรับปรุงและพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 จากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 นั้น ผู้วิจัยได้ทำการทดลองด้วยการเลือกใช้ตัวเลือกต่าง ๆ ในการปรับปรุงและพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ โดยมีตัวเลือกในการปรับปรุงและพัฒนาวิธีการ 4 ตัวเลือก ได้แก่ 1) ปริภูมิสี RGB และปริภูมิสี Lab 2) ตัวกรองสัญญาณเกาส์

เขียน AS^2G 3) เกณฑ์เอนโทรปี และ 4) เกณฑ์เอนโทรปีแบบท้องถิ่น ซึ่งการเลือกใช้ตัวเลือกต่าง ๆ ในการปรับปรุงและพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพผู้วิจัยได้ทำการแบ่งการเลือกใช้ตัวเลือกต่าง ๆ ออกเป็น 6 รูปแบบด้วยกัน โดยผลประสิทธิภาพการทดลองใช้ตัวเลือกของรูปแบบต่าง ๆ ในการพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพด้วยค่า AUC แสดงดังตาราง 4

ตาราง 4 การเปรียบเทียบประสิทธิภาพการเลือกใช้ตัวเลือกต่าง ๆ ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ด้วยค่า AUC

รูปแบบ	ปริภูมิสี RGB และ Lab	ตัวกรอง AS^2G	เอนโทรปี	เอนโทรปี แบบท้องถิ่น	AUC
1 (วิธีการที่ 1)	×	×	×	×	0.9283
2	✓	×	✓	×	0.9291
3	✓	×	×	✓	0.9301
4	×	✓	×	×	0.9269
5	✓	✓	✓	×	0.9290
6 (วิธีการที่ 2)	✓	✓	×	✓	0.9311

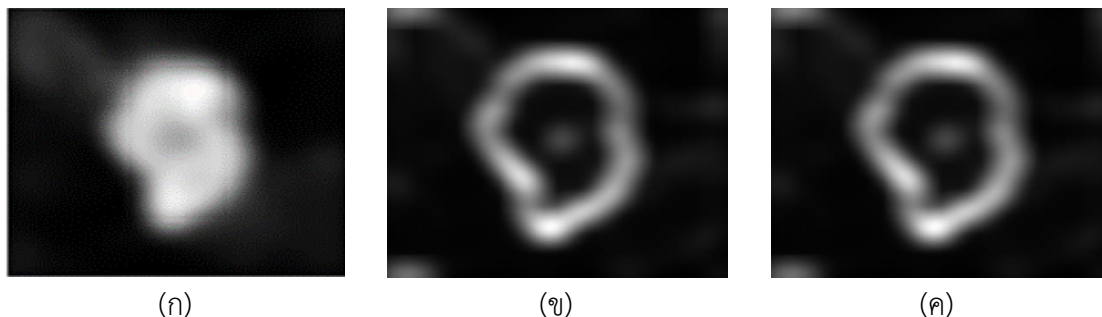
จากตารางที่ 4 ผลการเปรียบเทียบประสิทธิภาพการเลือกใช้ตัวเลือกต่าง ๆ ในการพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพด้วยค่า AUC ในการปรับปรุงและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 จากการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 แสดงให้เห็นว่าเมื่อเลือกใช้ปริภูมิสี RGB และ Lab ร่วมกับเกณฑ์เอนโทรปีแสดงดังรูปแบบ 2 พบว่าประสิทธิภาพสูงกว่าการใช้เพียงปริภูมิสี RGB เท่านั้น แสดงดังรูปแบบ 1 ซึ่งกล่าวได้ว่าการใช้ปริภูมิสีทั้งสองสีสามารถเพิ่มประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ ยิ่งไปกว่านั้นประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพจะดียิ่งขึ้นเมื่อใช้ปริภูมิสี RGB และ Lab ร่วมกับเกณฑ์เอนโทรปีแบบท้องถิ่นแสดงดังรูปแบบ 3 ถึงแม้ว่าประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพจะลดลงเมื่อใช้เพียงตัวกรอง AS^2G แสดงดังรูปแบบ 4 แต่จะได้รับการประสิทธิภาพที่สูงขึ้นเมื่อใช้ตัวกรอง AS^2G ร่วมกับปริภูมิสี RGB และ Lab และเกณฑ์เอนโทรปีแบบท้องถิ่นแสดงดังรูปแบบ 6 จึงสามารถสรุปได้ว่า การปรับปรุงและพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจของภาพโดยใช้ตัวเลือกในรูปแบบต่าง ๆ นั้น ประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพดีขึ้นด้วยสามประเด็นหลักได้แก่ ประเด็นแรกการใช้ปริภูมิสี RGB และ Lab ประเด็นที่สองการใช้ตัวกรอง AS^2G เพื่อกรองสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ และประเด็นสุดท้ายการเลือกภาพบริเวณที่น่าสนใจโดยใช้เกณฑ์เอนโทร

ปีแบบท้องถิ่น ดังนั้น ผู้วิจัยจึงเลือกใช้รูปแบบที่ 6 ในการปรับปรุงและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

4.3 ผลการวิเคราะห์ผลการทดลองเชิงเปรียบเทียบโดยจำแนกตามหมวดหมู่ของภาพ 6 ประเภท

การวิเคราะห์ผลการทดลองเชิงเปรียบเทียบของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ สามารถแจกแจงผลการทดลองออกเป็นภาพทั้งหมด 6 ประเภท ดังนี้

4.3.1 ผลการทดลองประเภทภาพบริเวณที่น่าสนใจมีขนาดใหญ่ เป็นภาพบริเวณที่น่าสนใจมีขนาดใหญ่ที่มีความเด่นชัดที่สุดในภาพและเป็นองค์ประกอบหลักของภาพ อาจมีพื้นหลังหรือวัตถุอื่น ๆ เป็นส่วนประกอบในภาพ จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่า ภาพประเภท 1 หรือภาพบริเวณที่น่าสนใจมีขนาดใหญ่มีประสิทธิภาพดีที่สุดในวิธีการ HFT, วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2, วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, Itti, SAL, SR และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการ HFT ได้ใช้ตัวกรองสัญญาณเกาส์เซียนเพียงขนาดเดียวและพื้นที่เดียวในการกรองต่อหนึ่งภาพซึ่งภาพบริเวณที่น่าสนใจมีขนาดใหญ่อาจมีความเหมาะสมกับตัวกรองสัญญาณเกาส์เซียนของวิธีการ HFT มากกว่าตัวกรองสัญญาณเกาส์เซียนของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 เพราะตัวกรองสัญญาณเกาส์เซียนของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 นั้น สามารถกรองภาพได้ละเอียดมากกว่า โดยจะกรองไปในแต่ละบริเวณของภาพรวมทั้งปรับขนาดและพื้นที่ของตัวกรองอัตโนมัติทำให้ลดบริเวณที่ไม่เกี่ยวข้องหรือพื้นหลังออกไปได้มากกว่าตัวกรองสัญญาณเกาส์เซียนของวิธีการ HFT ซึ่งจากการกรองที่มีความละเอียดมากกว่าของตัวกรองสัญญาณเกาส์เซียนในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 อาจทำให้ผลลัพธ์ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 มีพื้นที่ของบริเวณที่น่าสนใจลดลงหรือบริเวณที่สำคัญที่สุดของภาพบริเวณที่น่าสนใจลดลง แสดงดังภาพ 32 จากเหตุผลข้างต้นจึงทำให้ประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ดีกว่าวิธีการ HFT โดยมีตัวอย่างผลการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพบริเวณที่น่าสนใจมีขนาดใหญ่ แสดงดังภาคผนวก ก ตาราง 6



ภาพ 32 ตัวอย่างภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ

ภาพ 32 แสดงตัวอย่างภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพ ภาพ 32 (ก) วิธีการ HFT ภาพ 32 (ข) วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 ภาพ 32 (ค) วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 จากภาพแสดงให้เห็นว่าผลลัพธ์ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 มีบริเวณที่สำคัญหรือพื้นที่ของบริเวณที่น่าสนใจหายไป

4.3.2 ผลการทดลองประเภทภาพบริเวณที่น่าสนใจมีขนาดปานกลาง เป็นภาพที่มีบริเวณที่น่าสนใจมีพื้นที่ขนาดเล็กกลางมาจากภาพประเภท 1 ซึ่งมีขนาดพื้นที่ไม่เล็กมากและไม่ใหญ่มากมีบริเวณที่น่าสนใจเป็นองค์ประกอบที่เด่นชัดที่สุดในภาพและมีพื้นหลังหรือบริเวณอื่น ๆ เป็นองค์ประกอบในภาพ จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่าภาพประเภท 2 หรือภาพบริเวณที่น่าสนใจมีขนาดปานกลางมีประสิทธิภาพดีที่สุดในวิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2, วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, Itti, HFT, SAL, SR และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 สามารถกรองภาพบริเวณที่น่าสนใจมีขนาดปานกลางได้อย่างมีประสิทธิภาพ ลดบริเวณที่ไม่สำคัญหรือพื้นหลังออกได้โดยไม่ส่งผลกระทบต่อพื้นที่ของบริเวณที่น่าสนใจ ยิ่งไปกว่านั้นได้มีการใช้ปริภูมิสี RGB ร่วมกับ Lab ทำให้ผลลัพธ์ภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 มีความสมบูรณ์และบริเวณที่น่าสนใจปรากฏชัด ส่งผลให้วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 มีประสิทธิภาพดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบอื่น ๆ ตัวอย่างผลภาพการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพแสดงดังภาคผนวก ก ตาราง 7

4.3.3 ผลการทดลองประเภทภาพบริเวณที่น่าสนใจมีขนาดเล็ก เป็นภาพบริเวณที่น่าสนใจมีพื้นที่ขนาดเล็ก ซึ่งองค์ประกอบของภาพโดยส่วนใหญ่เป็นพื้นหลังหรือบริเวณอื่น ๆ และมีบริเวณที่

สนใจเป็นองค์ประกอบเด่นชัดที่สุดในภาพ จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่าภาพประเภท 3 หรือภาพบริเวณที่น่าสนใจมีขนาดเล็กมีประสิทธิภาพดีที่สุดในวิธีการ Itti, SAL, วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2, SR, วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, HFT และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการ Itti ใช้หลักการอ้างอิงคุณสมบัติบนโดเมนตำแหน่งที่มีการใช้คุณสมบัติพื้นฐานของภาพ ได้แก่ ทิศทาง องศา และใช้คุณลักษณะของภาพที่สกัดมาจากตัวกรองกาบอร์ซึ่งเป็นตัวกรองที่มีคุณสมบัติเด่นในเรื่องของทิศทาง และเส้น สามารถตรวจจับบริเวณที่น่าสนใจที่มีขนาดเล็กได้อย่างมีประสิทธิภาพ ซึ่งวิธีการ Itti ได้นำชุดภาพฝึกสอนเป็นคุณลักษณะของภาพจากตัวกรองกาบอร์ให้คอมพิวเตอร์เรียนรู้ จึงทำให้วิธีการตรวจจับ Itti มีความแม่นยำมากกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบอื่น ๆ ตัวอย่างผลภาพการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพแสดงดังภาคผนวก ก ตาราง 8

4.3.4 ผลการทดลองประเภทภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน เป็นภาพที่มีลักษณะของพื้นหลังและมีลักษณะของรูปแบบที่ซ้ำกัน โดยองค์ประกอบในภาพส่วนใหญ่เป็นรูปแบบที่ซ้ำกันและมีบริเวณที่น่าสนใจปรากฏเด่นชัดที่สุด จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่าภาพประเภทที่ 4 หรือภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกันมีประสิทธิภาพดีที่สุดในวิธีการตรวจจับบริเวณที่น่าสนใจของวิธีการแบบที่ 2, วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 1, HFT, Itti, SAL, SR, และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 สามารถกรองภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกันได้อย่างมีประสิทธิภาพ ลดบริเวณที่เกิดรูปแบบซ้ำ ๆ ได้โดยไม่ส่งผลกระทบต่อถึงพื้นที่ของบริเวณที่น่าสนใจ อีกทั้งยังมีการใช้เกณฑ์เอนโทรปีแบบท้องถิ่นเลือกภาพบริเวณที่น่าสนใจจากปริภูมิสี RGB และปริภูมิ Lab ทำให้ผลลัพธ์ภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 มีความสมบูรณ์และบริเวณที่น่าสนใจปรากฏชัดมากขึ้น ส่งผลให้วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 มีประสิทธิภาพดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบอื่น ๆ ตัวอย่างผลภาพการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพแสดงดังภาคผนวก ก ตาราง 9

4.3.5 ผลการทดลองประเภทภาพพื้นหลังมีความซับซ้อน เป็นภาพที่มีลักษณะของพื้นหลังมีความวุ่นวายและซับซ้อนมีบริเวณพื้นหลังบางส่วนปะปนกับบริเวณที่น่าสนใจของภาพ องค์ประกอบในภาพมีความสับสนแต่ยังสามารถสังเกตบริเวณที่น่าสนใจได้ จากตาราง 3 ผลการเปรียบเทียบ

ประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่า ภาพประเภท 5 หรือภาพพื้นหลังมีความซับซ้อนมีประสิทธิภาพดีที่สุดในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, วิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2, HFT, Itti, SAL, SR, และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 สามารถกรองภาพพื้นหลังที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ สามารถลดบริเวณพื้นหลังที่ซับซ้อนและปะปนกับบริเวณที่น่าสนใจออกได้เพราะตัวกรองของวิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 1 ปรับปรุงมาจากตัวกรองจุดกึ่งกลางแบบปรับเหมาะที่มีประสิทธิภาพในการกำจัดสิ่งรบกวนในภาพซึ่งภาพพื้นหลังที่ซับซ้อนคล้ายกับสิ่งรบกวนภายในภาพ จึงทำให้ผลลัพธ์ภาพบริเวณที่น่าสนใจของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 มีประสิทธิภาพดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจแบบอื่น ๆ ตัวอย่างผลภาพการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพแสดงดังภาคผนวก ก ตาราง 10

4.3.6 ผลการทดลองประเภทภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่ เป็นภาพที่มีลักษณะพื้นที่ของบริเวณที่น่าสนใจขนาดเล็กและขนาดใหญ่ปะปนอยู่ด้วยกัน องค์ประกอบของภาพโดยส่วนใหญ่เป็นบริเวณที่น่าสนใจทั้งขนาดเล็กและขนาดใหญ่ อาจมีบริเวณอื่นหรือพื้นหลังปะปนอยู่ จากตาราง 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC แสดงให้เห็นว่าภาพประเภท 6 หรือภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่มีประสิทธิภาพดีที่สุดในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1, การตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2, HFT, SAL, Itti, SR และ SUN ตามลำดับ จากการวิเคราะห์ผลการทดลองพบว่า วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 สามารถกรองภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่ได้อย่างมีประสิทธิภาพ โดยกรองบริเวณที่เป็นพื้นหลังออกได้โดยไม่มีผลกระทบกับบริเวณที่น่าสนใจ ตัวอย่างผลภาพการทดลองและเปรียบเทียบประสิทธิภาพด้วยค่า AUC ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ในประเภทภาพแสดงดังภาคผนวก ก ตาราง 11

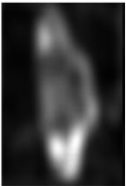
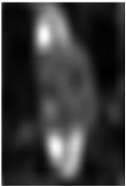
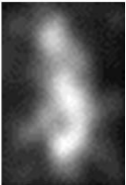
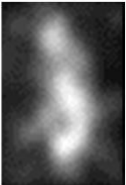
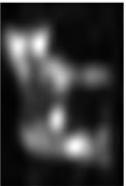
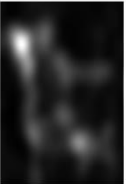
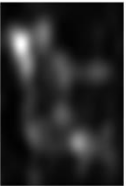
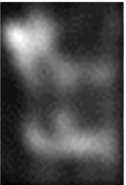
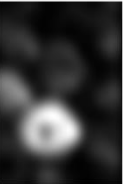
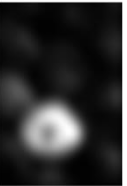
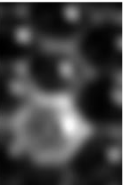
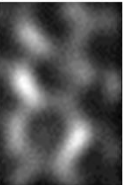
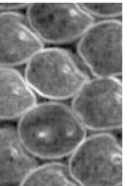
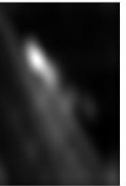
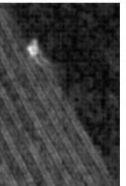
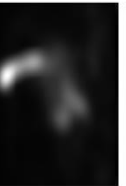
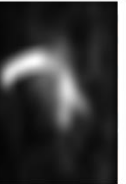
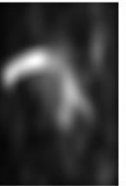
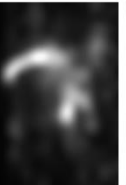
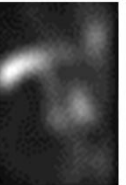
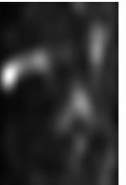
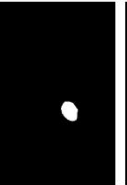
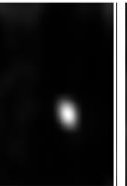
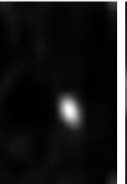
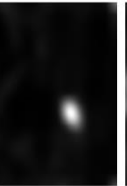
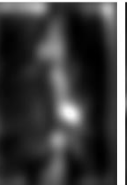
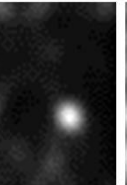
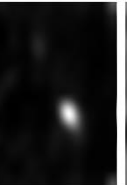
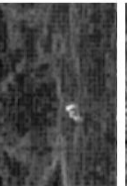
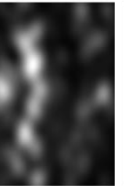
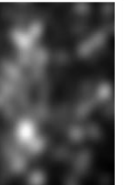
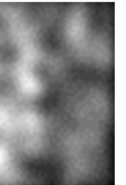
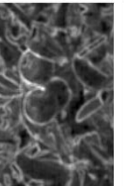
4.4 ผลการเปรียบเทียบประสิทธิภาพด้านระยะเวลาในการประมวลผลของการตรวจจับบริเวณที่น่าสนใจของภาพ

ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพการตรวจจับบริเวณที่น่าสนใจของภาพด้วยค่า AUC กับเวลาในการประมวลผลของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ แสดงดังตาราง 5

ตาราง 5 เปรียบเทียบประสิทธิภาพการวิธีตรวจจับบริเวณที่น่าสนใจของภาพด้วยค่า AUC กับระยะเวลาในการประมวลของวิธีการ Itti, SUN, SR, SAL, HFT, การตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2

วิธีการ	Itti	SUN	SR	SAL	HFT	แบบที่ 1	แบบที่ 2
เวลาในการประมวล (วินาที)	0.1324	0.6498	0.7256	0.0635	0.7357	1.9286	2.7278
AUC	0.9147	0.7719	0.8547	0.9237	0.9244	0.9283	0.9311

จากตาราง 5 ผลการเปรียบเทียบประสิทธิภาพการทำงานของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ ด้วยค่า AUC กับระยะเวลาในการประมวลผลการทำงานของ การตรวจจับบริเวณที่น่าสนใจของภาพ แสดงให้เห็นว่าวิธีการ SAL ใช้เวลาในการประมวลผลน้อยที่สุด และวิธีการ Itti, SUN, SR, HFT, วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ใช้เวลาในการประมวลผลตามลำดับ จะเห็นได้ว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ใช้เวลาในการประมวลผลมากที่สุดเนื่องจากต้องใช้เวลาค่อนข้างมากสำหรับการสกัดคุณลักษณะของภาพจากปริภูมิสี RGB และ Lab อีกทั้งยังต้องคำนวณภาพบริเวณที่น่าสนใจจากปริภูมิสี RGB และ Lab ก่อนที่จะนำมาเปรียบเทียบเพื่อเลือกภาพบริเวณที่น่าสนใจที่เหมาะสมมากที่สุดโดยใช้เกณฑ์เอนโทรปีแบบท้องถิ่นในการเลือก รวมไปถึงการกรองสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ ในแต่ละบริเวณของภาพ แต่ถึงอย่างไรก็ตามความแม่นยำของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 มีประสิทธิภาพสูงกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ดังนั้น ผู้วิจัยจึงเลือกใช่วิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 ในการพัฒนาและปรับปรุงการตรวจจับบริเวณที่น่าสนใจของภาพ และมีตัวอย่างของผลการทดลองเป็นภาพเชิงประจักษ์ของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 เปรียบเทียบกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 Itti, SUN, SR, และ HFT แสดงดังภาพ 33 ซึ่งจะสามารถสังเกตเห็นได้อย่างชัดเจนว่า ภาพบริเวณที่น่าสนใจที่ได้จากวิธีการตรวจจับบริเวณที่น่าสนใจแบบที่ 2 นั้น มีบริเวณที่น่าสนใจปรากฏขึ้นอย่างเด่นชัด มีบริเวณอื่น ๆ ปะปนกับบริเวณที่น่าสนใจปริมาณที่น้อยกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ เช่น ภาพผีเสื้อ ภาพผลแอปเปิ้ล และภาพแมว เป็นต้น

ภาพต้นฉบับ	ภาพเฉลี่ย	แบบที่ 2	แบบที่ 1	HFT	Itti	SAL	SR	SUN
								
								
								
								
								
								
								

ภาพ 33 เปรียบเทียบภาพบริเวณที่น่าสนใจที่ได้จากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 แบบที่ 1, HFT, Itti, SAL, SR และ SUN

บทที่ 5

บทสรุป

งานวิจัยฉบับนี้ได้นำเสนอวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่ปรับปรุงและพัฒนากระบวนการจากวิธีการตรวจจับบริเวณที่น่าสนใจของภาพที่มีอยู่ในปัจจุบัน โดยความแตกต่างในวิธีการตรวจจับบริเวณที่น่าสนใจของภาพของงานวิจัยฉบับนี้กับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ คือ งานวิจัยนี้สามารถนำเสนอบริเวณที่น่าสนใจได้อย่างชัดเจน โดยลบบริเวณที่เป็นพื้นหลังหรือบริเวณที่ไม่น่าสนใจ

5.1 สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์ในการศึกษาการตรวจจับบริเวณที่น่าสนใจของภาพ เพื่อปรับปรุงและพัฒนากระบวนการตรวจจับบริเวณที่น่าสนใจของภาพและเปรียบเทียบผลการทดลองของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในวิธีการแบบต่าง ๆ จากผลการทดลองพบว่า ผลการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบต่าง ๆ แสดงให้เห็นถึงประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพของงานวิจัยนี้ดีกว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบอื่น ๆ ทั้งหมดด้วยค่า AUC ในภาพรวม ซึ่งงานวิจัยนี้ใช้หลักการอ้างอิงสเปกตรัมบนโดเมนความถี่ในการปรับปรุงและพัฒนาวิธีการตรวจจับบริเวณที่น่าสนใจ ซึ่งวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้สามารถสกัดและค้นหาบริเวณที่น่าสนใจหรือบริเวณที่เป็นลักษณะเด่นของภาพและสามารถแยกแยะบริเวณที่ไม่ใช่ลักษณะเด่นหรือพื้นหลังออกจากบริเวณที่น่าสนใจได้อย่างมีประสิทธิภาพ

5.2 อภิปรายผล

ประสิทธิภาพของการตรวจจับบริเวณที่น่าสนใจของภาพที่ได้ปรับปรุงและพัฒนาพบว่าประสิทธิภาพของงานวิจัยวิธีการตรวจจับบริเวณที่น่าสนใจของภาพโดยใช้ฐานข้อมูล Salmap (J. Li et al., 2013) มีประสิทธิภาพดีที่สุดเมื่อเปรียบเทียบกับวิธีการตรวจจับบริเวณที่น่าสนใจของภาพวิธีการอื่น ๆ ด้วยค่า AUC ในภาพรวม ได้แก่ วิธีการ HFT (J. Li et al., 2013), SAL (Hou et al., 2012), SR (Hou & Zhang, 2007), Itti (Itti et al., 1998) และ SUN (Zhang et al., 2008), ซึ่งวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้ได้ปรับปรุงตัวกรองสัญญาณเกาส์เซียนหลายระดับแบบปรับเหมาะมาจากตัวกรองสัญญาณเกาส์เซียนมัลติสเกลแบบปรับเหมาะ (Chan et al., 2005) โดยได้แนวคิดที่ว่าตัวกรองสัญญาณเกาส์เซียนมัลติสเกลแบบปรับเหมาะมีประสิทธิภาพในการกำจัดสัญญาณรบกวนใน ซึ่งสัญญาณรบกวนในภาพมีความคล้ายคลึงกับสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ

หรือบริเวณที่ไม่น่าสนใจในภาพ จากแนวคิดนี้ผู้วิจัยจึงนำมานำมาปรับปรุงและพัฒนาและพบว่าวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้สามารถการกรองสเปกตรัมที่เกิดรูปแบบซ้ำ ๆ หรือพื้นที่ที่เป็นบริเวณที่เป็นพื้นหลังออกจากพื้นที่ของบริเวณที่น่าสนใจได้อย่างมีประสิทธิภาพ อีกทั้งตัวกรองสัญญาณเกาส์เซียนยังมีการปรับขนาดและพื้นที่แบบอัตโนมัติโดยจะปรับให้เหมาะสมกับพื้นที่แต่ละบริเวณในภาพ รวมการใช้เกณฑ์เอนโทรปีแบบท้องถิ่นในการช่วยเลือกภาพที่เป็นบริเวณที่น่าสนใจจากภาพที่เป็นบริเวณที่น่าสนใจจากปริภูมิสี RGB และ Lab ซึ่งสอดคล้องกับงานวิจัยของ (Xiang & Zhong, 2015) ที่ว่าบริเวณที่น่าสนใจสามารถปรากฏขึ้นได้ในปริภูมิสีอื่น ๆ นอกจากปริภูมิสี RGB ซึ่งในงานวิจัยนี้ได้ใช้ปริภูมิสี Lab ในการปรับปรุงและพัฒนาการตรวจจับบริเวณที่น่าสนใจของภาพ ผลลัพธ์ของงานวิจัยนี้แสดงให้เห็นว่าประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพมีประสิทธิภาพ จึงสรุปได้ว่า วิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้สามารถสกัดบริเวณที่น่าสนใจของภาพได้ใกล้เคียงกับภาพเฉลี่ยที่มนุษย์สร้างขึ้น และสามารถนำวิธีการตรวจจับบริเวณที่น่าสนใจของภาพไปประยุกต์ใช้กับงานอื่นให้เกิดประโยชน์ได้

5.3 ข้อเสนอแนะ

วิธีการในงานวิจัยนี้มีกระบวนการในการตรวจจับบริเวณที่น่าสนใจของภาพโดยเริ่มต้นจากกระบวนการการสกัดคุณลักษณะของภาพ ซึ่งวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้ได้ทำการสกัดคุณลักษณะของสีทั้งสองปริภูมิสี คือ ปริภูมิสี RGB และปริภูมิสี Lab ก่อนที่จะเข้าสู่กระบวนการกรองสเปกตรัมที่มีการกรองในแต่ละบริเวณของภาพ และกระบวนการเปรียบเทียบและเลือกภาพที่เป็นบริเวณที่น่าสนใจที่เหมาะสมที่สุดโดยใช้เกณฑ์เอนโทรปีแบบท้องถิ่น ซึ่งกระบวนการสกัดคุณลักษณะของสีทั้งสองปริภูมิสีใช้เวลาเยอะเป็นสองเท่าในกระบวนการสกัด หากมีวิธีที่สามารถนำปริภูมิสีทั้งสองสีมาใช้ได้โดยไม่ต้องผ่านการสกัดคุณลักษณะของภาพ อาจช่วยให้วิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้ลดเวลาในการคำนวณลงได้ และวิธีการตรวจจับบริเวณที่น่าสนใจของภาพในงานวิจัยนี้มีการใช้สเปกตรัมหลายระดับ อาจจะดีกว่าหากใช้สเปกตรัมเพียงระดับเดียวในกระบวนการ

บรรณานุกรม

- Achanta, R., Hemami, S., Estrada, F., & Susstrunk, S. (2009). **Frequency-tuned salient region detection**. Paper presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition.
- Chan, R. H., Chung-Wa, H., & Nikolova, M. (2005). **Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization**. IEEE Transactions on Image Processing, 14(10), 1479-1485.
- Hou, X., Harel, J., & Koch, C. (2012). **Image Signature: Highlighting Sparse Salient Regions**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(1), 194-201.
- Hou, X., & Zhang, L. (2007). **Saliency Detection: A Spectral Residual Approach**. Paper presented at the 2007 IEEE Conference on Computer Vision and Pattern Recognition.
- Itti, L., Koch, C., & Niebur, E. (1998). **A model of saliency-based visual attention for rapid scene analysis**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11), 1254-1259.
- Jaemsiri, J., Titijaroonroj, T., & Rungrattanaubol, J. (2019). **Modified Scale-Space Analysis in Frequency Domain Based on Adaptive Multiscale Gaussian Filter for Saliency Detection**. Paper presented at the 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE).
- Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). **Learning to predict where humans look**.
- Le Meur, O., & Baccino, T. (2012). **Methods for comparing scanpaths and saliency maps: Strengths and weaknesses**. Behavior research methods.
- Li, J., Levine, M. D., An, X., Xu, X., & He, H. (2013). **Visual Saliency Based on Scale-Space Analysis in the Frequency Domain**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(4), 996-1010.
- Rokszin, A., Márkus, Z., Braunitzer, G., Berényi, A., Benedek, G., & Nagy, A. (2010). **Visual pathways serving motion detection in the mammalian brain**. Sensors (Basel,

Switzerland), 10(4), 3218-3242.

Walther, D., Itti, L., Riesenhuber, M., Poggio, T., & Koch, C. (2002). **Attentional Selection for Object Recognition - a Gentle Way.**

Xiang, D., & Zhong, B. (2015). **Scale-space saliency detection in combined color space.** Paper presented at the 2015 Chinese Automation Congress (CAC).

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). **SUN: A Bayesian framework for saliency using natural statistics.** Journal of Vision, 8(7), 32-32.

ไทย์นิต, ส. (2010). **การประมวลผลภาพดิจิทัลเบื้องต้นด้วย MATLAB.** กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์ คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมไฟฟ้า.

มหาวิทยาลัยศรีนครินทรวิโรฒ, ภ. ค. (2558). **คณะวิศวกรรมศาสตร์ มศว.** ปีที่13 ฉบับที่ 1 ตุลาคม 2557-มีนาคม 2558





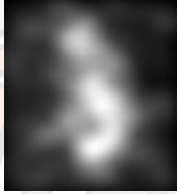
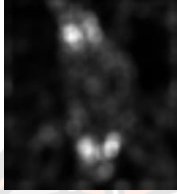
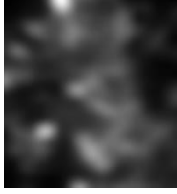
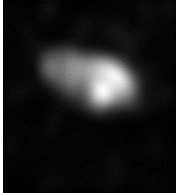
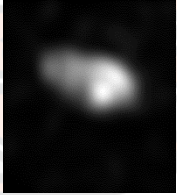
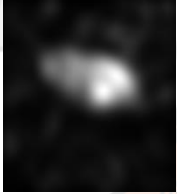
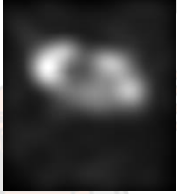
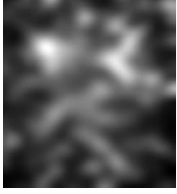
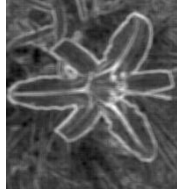
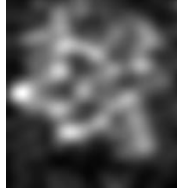
ภาคผนวก



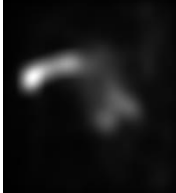
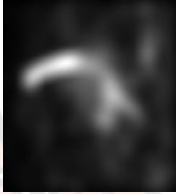
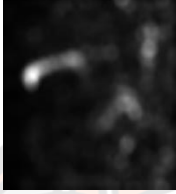
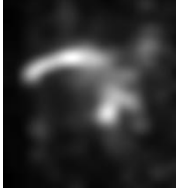
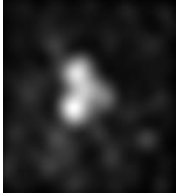
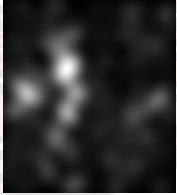
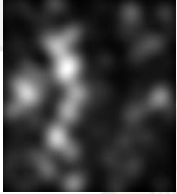
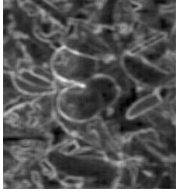
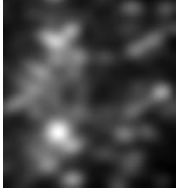
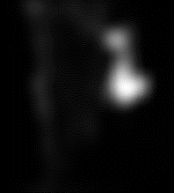
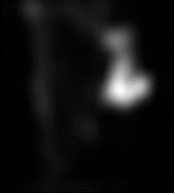
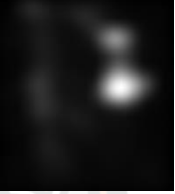
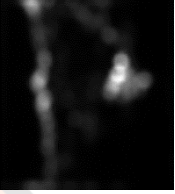
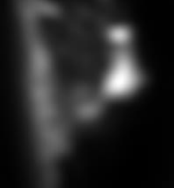
ภาคผนวก ก

ตัวอย่างภาพการเปรียบเทียบผลการทดลองการตรวจจับบริเวณที่น่าสนใจของภาพ

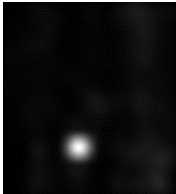
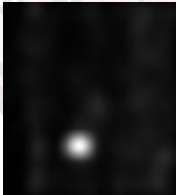
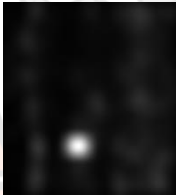
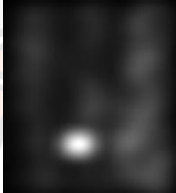
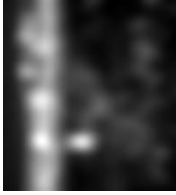
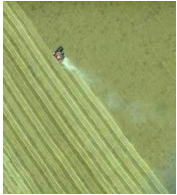
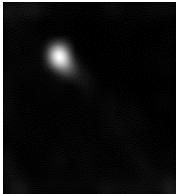
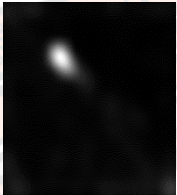
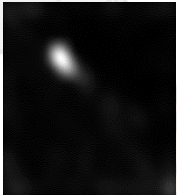
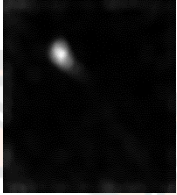
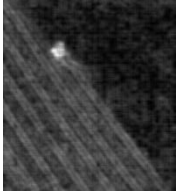
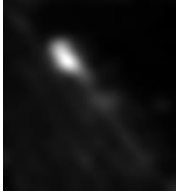
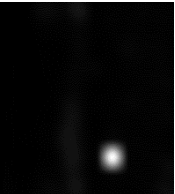
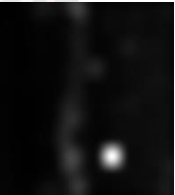
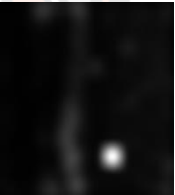
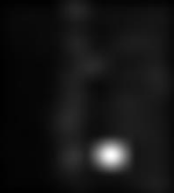
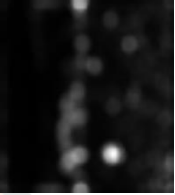
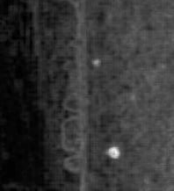
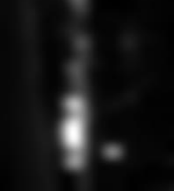
ตาราง 6 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 1 ภาพบริเวณที่น่าสนใจมีขนาดใหญ่

ภาพต้นฉบับ	วิธีการแบบที่ 2	วิธีการแบบที่ 1	HFT	SAL	SR	SUN	Itti
							
AUC	0.9906	0.9890	0.9892	0.9863	0.9798	0.9747	0.9788
							
AUC	0.9954	0.9953	0.9951	0.9938	0.8654	0.9735	0.9799
							
AUC	0.9886	0.9884	0.9882	0.9705	0.9374	0.9796	0.9783

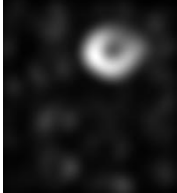
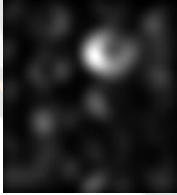
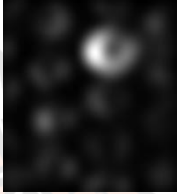
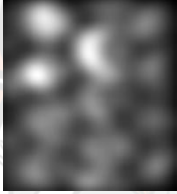
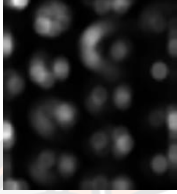
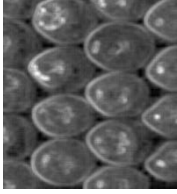
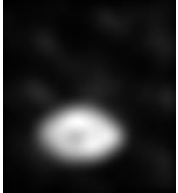
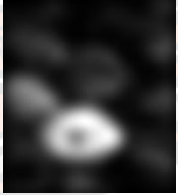
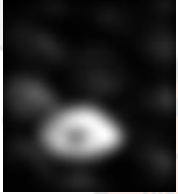
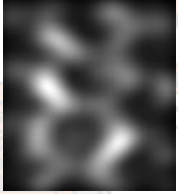
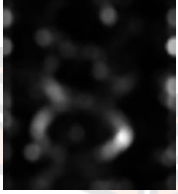
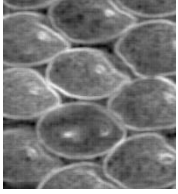
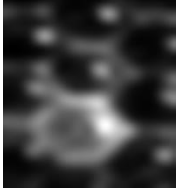
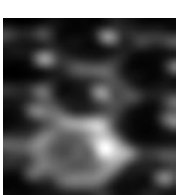
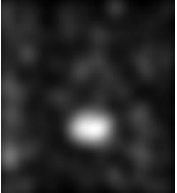
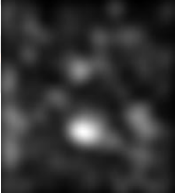
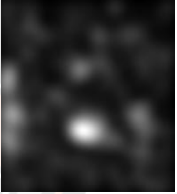
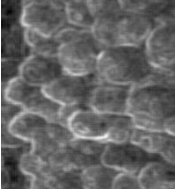
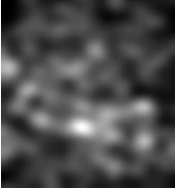
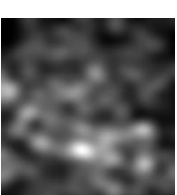
ตาราง 7 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 2 ภาพบริเวณที่น่าสนใจมีขนาดปานกลาง

ภาพต้นฉบับ	วิธีการแบบที่ 2	วิธีการแบบที่ 1	HFT	SAL	SR	SUN	Itti
							
AUC	0.9921	0.9891	0.9869	0.981	0.9666	0.8958	0.9840
							
AUC	0.9969	0.9704	0.9651	0.9328	0.8663	0.9256	0.9467
							
AUC	0.9977	0.9974	0.9960	0.9965	0.9862	0.9874	0.9942

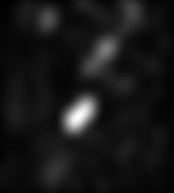
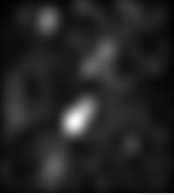
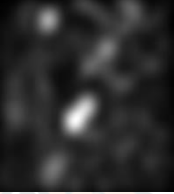
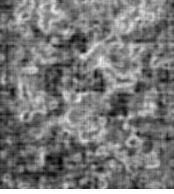
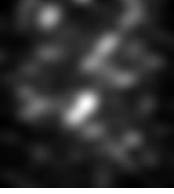
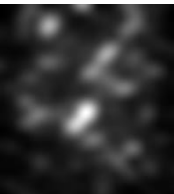
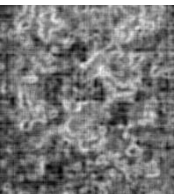
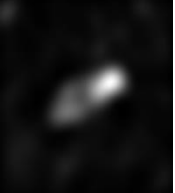
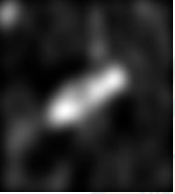
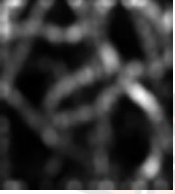
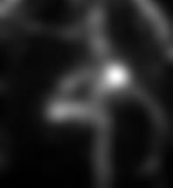
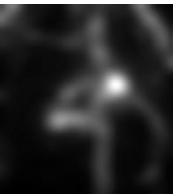
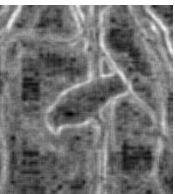
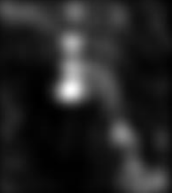
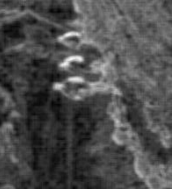
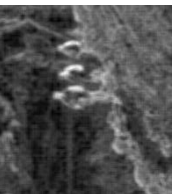
ตาราง 8 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 3 ภาพบริเวณที่น่าสนใจมีขนาดเล็ก

ภาพต้นฉบับ	วิธีการแบบที่ 2	วิธีการแบบที่ 1	HFT	SAL	SR	SUN	Itti
							
AUC	0.9992	0.9991	0.9991	0.9989	0.9967	0.9790	0.9630
							
AUC	0.9998	0.9994	0.9993	0.9987	0.9996	0.9794	0.9982
							
AUC	0.9950	0.9874	0.9928	0.9946	0.9915	0.9509	0.9772

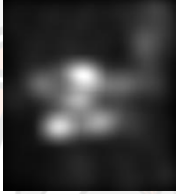
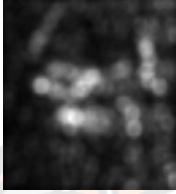
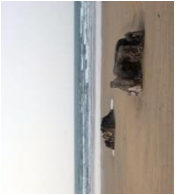
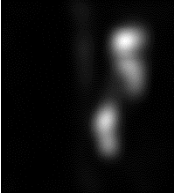
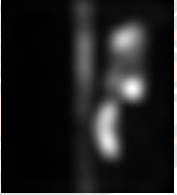
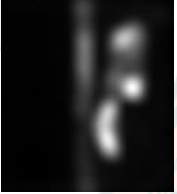
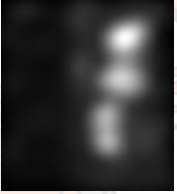
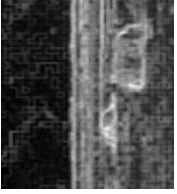
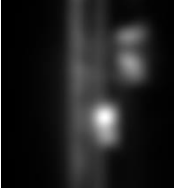
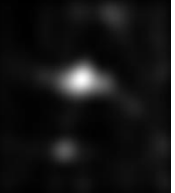
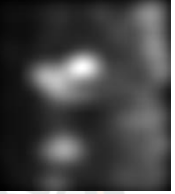
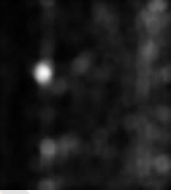
ตาราง 9 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 4 ภาพที่มีสิ่งรบกวนในรูปแบบที่ซ้ำกัน

ภาพต้นฉบับ	วิธีการแบบที่ 2				วิธีการแบบที่ 1				HFT	SAL	SR	SUN	Itti
													
AUC	0.9998	0.9992	0.9995	0.9566	0.9522	0.9037	0.9936						
													
AUC	0.9998	0.9992	0.9995	0.9566	0.9522	0.9037	0.9936						
													
AUC	0.9998	0.9975	0.9967	0.9915	0.9116	0.9278	0.9901						

ตาราง 10 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 2 วิธีการแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 5 ภาพพื้นหลังมีความซับซ้อน

ภาพต้นฉบับ	วิธีการแบบที่ 2					วิธีการแบบที่ 1					HFT	SAL	SR	SUN	Itti
															
AUC	0.9965	0.9965	0.9965	0.9965	0.9965	0.9965	0.9955	0.9836	0.9466	0.9261	0.9959				
															
AUC	0.9989	0.9951	0.9963	0.9726	0.9558	0.9192	0.9751								
															
AUC	0.9925	0.9778	0.9767	0.9898	0.9677	0.9497	0.9836								

ตาราง 11 ตัวอย่างภาพการเปรียบเทียบประสิทธิภาพของวิธีการตรวจจับบริเวณที่น่าสนใจของภาพแบบที่ 1 วิธีการ HFT, SAL, SR, SUN, และ Itti ด้วยค่า AUC ในประเภทที่ 6 ภาพบริเวณที่น่าสนใจมีทั้งขนาดเล็กและขนาดใหญ่

ภาพต้นฉบับ	วิธีการแบบที่ 2	วิธีการแบบที่ 1	HFT	SAL	SR	SUN	Itti
							
AUC	0.9981	0.9906	0.9933	0.9906	0.9892	0.9607	0.9938
							
AUC	0.9969	0.9939	0.9933	0.9958	0.9934	0.9609	0.9893
							
AUC	0.9959	0.9946	0.9915	0.9908	0.9614	0.8944	0.9901

ภาคผนวก ข

บทความที่นำเสนอผลงานในงานประชุมวิชาการ

1. บทความเรื่อง “Modified Scale-Space Analysis in Frequency Domain Based on Adaptive Multiscale Gaussian Filter for Saliency Detection” นำเสนอในงานประชุมวิชาการ The 16th International Joint Conference on Computer Science and Software Engineering (JCSSE2019) จัดที่โรงแรม Amari Pattaya จังหวัดชลบุรี ประเทศไทย ระหว่างวันที่ 10-12 กรกฎาคม พ.ศ.2562 บทความนี้ตีพิมพ์ไว้ใน Proceedings of 2019 - 16th International Joint Conference on Computer Science and Software Engineering หน้า 212-217

2. บทความเรื่อง “Improved Adaptive Spectrum Scale-Space in Frequency Domain for Saliency Detection” นำเสนอในงานประชุมวิชาการ The 12th International Conference on Knowledge and Smart Technology (KST-2020) จัดที่โรงแรม Amari Pattaya จังหวัดชลบุรี ประเทศไทย ระหว่างวันที่ 29 มกราคม – 1 กุมภาพันธ์ พ.ศ.2563 บทความนี้ตีพิมพ์ไว้ใน Proceedings of 2020 - 12th International Conference on Knowledge and Smart Technology หน้า 97-102

Modified Scale-Space Analysis in Frequency Domain Based on Adaptive Multiscale Gaussian Filter for Saliency Detection

Jenjira Jaemsiri *, Taravichet Titijaroonroj[†] and Jaratsri Rungrattanaubol[‡]

^{*}[†]Faculty of Science, Department of Computer Science and Information Technology,
Naresuan University, Phitsanulok, 65000, THAILAND

Email: jenjiraja61@email.nu.ac.th* and jaratsrir@nu.ac.th[‡]

[†]Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang,
Bangkok, 10520, THAILAND. Email: taravichet@it.kmitl.ac.th[†]

Abstract—The salient region is an area in the image that is paid attention from a human. It is a distinctive feature from neighbors such as color, shape or pattern. Saliency detection is a model that imitates the human visual system to perceive the scene. It has been widely used in many vision systems. Many papers use the smoothing or suppressing technique to extract a desirable output as so-called saliency map. Even though most of these researches achieve saliency detection based on the filter, the size of the filter is fixed. This leads to a filter ineffective when applying to the whole area of each image. In order to solve this issue, an adaptive multiscale Gaussian filter (MSS) for scale-space analysis in the frequency domain is proposed. The proposed filter is extended from an adaptive median filter which is a powerful method to remove the noise from the input image. This paper proposes the method that offers the appropriate filter to suppress the repeated pattern spectrum of each region in each image before extracting the saliency map. The experimental result shows that the proposed method outperforms the baseline methods, which includes HFT, Itti, SAL, SR and SUN.

Keywords—Adaptive multiscale Gaussian filter; scale-space analysis; saliency detection; saliency map

I. INTRODUCTION

Saliency detection is a process for detecting the candidate area paid attention by the human on an image. It is widely used in many applications such as image segmentation, detection and recognition to reduce the computational cost [1].

Based on the best of my knowledge, the researches of the saliency detection can be separated into two groups including (I) spatial feature-based and (II) frequency spectrum-based approaches. In the first group, Itti et al. presented a model of saliency-based visual attention as explained in [2]. This method uses three main features—color, intensity and orientation which are extracted by using the four broadly-tuned color channels, the color double-opponent and the Gabor pyramid filter. In 2007, Bayesian framework [3] for saliency was presented by Zhang et al. This method can estimate the probability of a whole given image based on the visual feature and the presence of a target, represented in a form of self-information and log likelihood, respectively.

In the second group, it aims to detect the saliency map in the given image by using the spectrum on the frequency domain. A spectral residual approach [4], for example, was

presented in 2007 by Hou et al., which extracts the saliency map by using spectral residual, the input image is transformed to the frequency domain for analyzing the difference between log spectrum and average log spectrum to construct a saliency map. After that, an image signature introduced in 2012 by Hou et al. as described in [5] uses the discrete cosine transform (DCT) for extracting the sign of each component which is equivalent to the phase component in the frequency domain. It can approximate foreground which can be considered as a salient region for figure-ground separation. In 2013, Jain and his colleagues [6] proposed the scale-space analysis in the frequency domain by using hypercomplex Fourier transform (HFT). This method suppresses the repeated pattern spectrum by using the multiscale Gaussian filter in conjunction with a local entropy. The advantage of frequency spectrum-based is that it doesn't require the prior knowledge and the training image whereas they are needed in the spatial feature-based. Moreover it is an adaptive approach to the characteristic of the input image. Therefore, this paper focuses on the frequency spectrum-based approach. Although Jain's approach is successful for the object region prediction which humans pay attention, its performance depends on the suitable of filter used.

In order to discover the appropriate filter for each region in each image, Modified Scale-Space analysis (MSS) in the frequency domain based on adaptive multiscale Gaussian filter is proposed here as depicted in Fig.1. This proposed method extends from Jain's method by replacing the conventional Gaussian filter to the adaptive multiscale Gaussian filter. It can select the appropriate filter for each area of each image automatically based on a solution from statistical analysis.

The remainder of this paper is organized as follows. In section II, related works are briefly described. The detail of the proposed method is given in section III and experiments are discussed in section IV. Finally, section V presents the conclusion.

II. RELATED WORKS

This section briefly explains the overview of the existing approaches for the saliency detection in the frequency domain as follows.

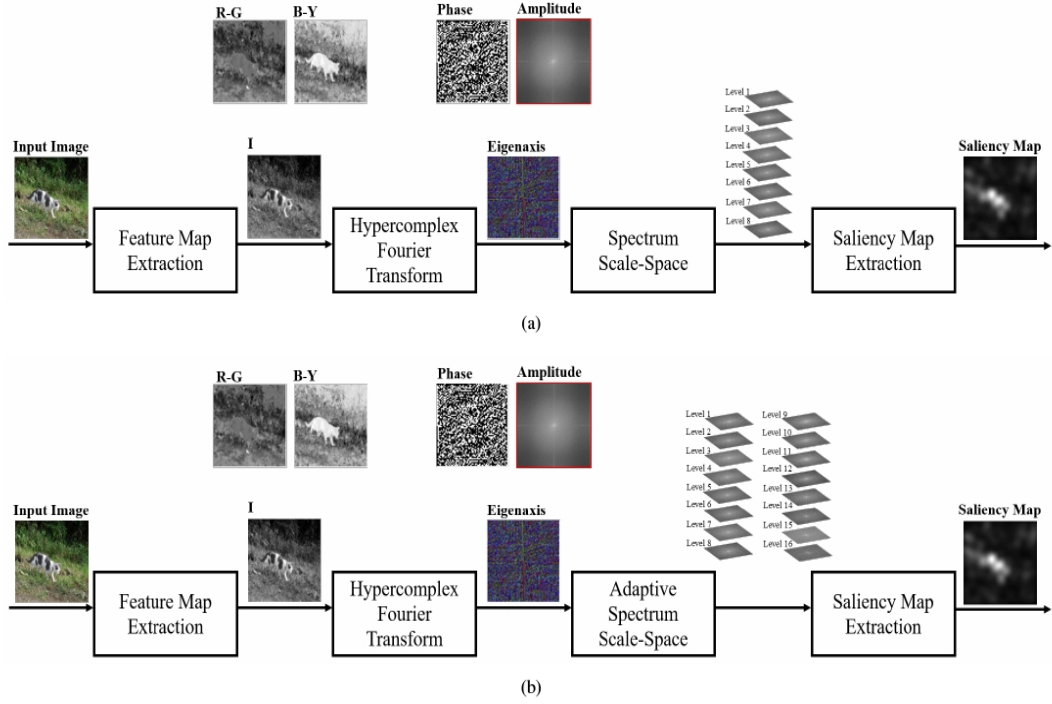


Fig. 1: A framework of visual saliency scale-space analysis in the frequency domain, (a) Jain's approach (b) proposed method (MSS).

A. Spectral Residual Approach

In 2007, the spectral residual approach [4] presented by Hou et al. described that the input image has two main areas—repeated pattern area and saliency area, which relate to the spectrum in the frequency domain. The repeated pattern area in the spatial domain is located in the same position in the frequency domain. This makes its spectrum so high as a so-called spike in the frequency domain. Meanwhile, the saliency area is extracted by eliminating the repeated pattern spectrum from the original spectrum named spectral residue. The Hou's approach uses this concept to detect the saliency area by using the Gaussian filter to approximate the spectrum of the repeated area. Firstly, the input image is transformed into the frequency domain from the spatial domain as defined in Eq. (1).

$$F(u, v) = DFT(f(x, y)) \quad (1)$$

where F and f are the input image in the frequency and spatial domain, respectively. $DFT(\cdot)$ is a discrete Fourier function. And then, the log spectrum L is computed by using F as defined in Eq. (2).

$$L(u, v) = \log(\text{real}(F(u, v))) \quad (2)$$

where $\log(\cdot)$ and $\text{real}(\cdot)$ denote the logarithm function and the real part of F , respectively. The repeated pattern spectrum A_r is approximated from the log spectrum by using the Gaussian filter, G , with size 3×3 and k is set to "4" as shown in (3) and (4).

$$A_r = G_k * L \quad (3)$$

$$G_k(u, v) = \frac{1}{0.5 \times \sqrt{2\pi} 2^{k-1}} e^{-(u^2+v^2)/(0.5 \times 2^{2k-1})} \quad (4)$$

where $*$ is a convolution operation. The spectral residual in the frequency domain, R , is computed by using Eq. (5).

$$R(u, v) = L(u, v) - A_r(u, v) \quad (5)$$

The spectral residual in the spatial domain, defined as S , denotes the saliency map when an inverse discrete Fourier transform applied to R .

B. Scale-Space Analysis Approach

In 2013, the scale-space analysis approach in the frequency domain [6] was presented by Jian et al. This paper extracts the saliency map by suppressing the repeated pattern spectrum by means of smoothing the amplitude spectrum. Scale-space analysis approach consists of four main steps as follows.

1) Feature Map Extraction: This process consists of three components including I , RG and BY are extracted by using Eqs. (6) to (8), respectively.

$$I = \frac{r + g + b}{3} \quad (6)$$

$$RG = (r - (\frac{g+b}{2})) - (g - (\frac{r+b}{2})) \quad (7)$$

$$BY = (b - (\frac{r+g}{2})) - ((\frac{g+r}{2}) - (\frac{|r-g|}{2}) - b) \quad (8)$$

where r , g and b are the color channels of the input image including red, green and blue, respectively, while I is a gray scale image. RG and BY are the color double-opponent in red-green and blue-yellow, respectively, as explained in [2].

2) Hypercomplex Fourier Transform: It is a transformation method for converting the input from the spatial domain to the frequency domain.

$$F_h = HFT(f_h) \quad (9)$$

where $HFT(\cdot)$ is the hypercomplex Fourier transform. F_h and f_h are a hypercomplex matrix in the frequency and spatial domain, respectively, which f_h is formed by Eq. (10).

$$f_h = (0.5 \times I) + (0.25 \times RG) + (0.25 \times BY) \quad (10)$$

And then, the amplitude, the phase and the eigenaxis spectra are calculated by using Eqs. (11) to (13), respectively.

$$A(u, v) = \|F_h(u, v)\| \quad (11)$$

$$P(u, v) = \tan^{-1}\left(\frac{F_{h,v}(u, v)}{F_{h,s}(u, v)}\right) \quad (12)$$

$$X(u, v) = \frac{F_{h,v}(u, v)}{\|F_{h,v}(u, v)\|} \quad (13)$$

where A , P and X are the amplitude, phase and eigenaxis spectra, respectively. $F_{h,s}$ and $F_{h,v}$ are the scale and vector parts of F_h as explained in [6] while $\|\cdot\|$ is a hypercomplex matrix modulus.

3) Spectrum Scale-Space: the repeated pattern spectrum is suppressed by applying Gaussian filter G_k at scale k to the amplitude spectrum as shown in Eq. (14).

$$\Lambda_k = G_k * A \quad (14)$$

where Λ_k is a smoothed amplitude spectrum. The Gaussian filter at scale $k \in \{1, 2, 3, \dots, 8\}$ is defined in Eq. (4).

4) Saliency Map Extraction: the candidate saliency map at scale k is computed by Eq. (15).

$$S_k = g_\sigma * \|IHFT(\Lambda_k e^{XP})\|^2 \quad (15)$$

where $IHFT(\cdot)$ is the inverse hypercomplex Fourier transform and g_σ is the Gaussian filter with $\sigma = 0.05$ and S_k is the saliency map at scale k in the spatial domain. The final saliency map is defined by selecting the minimum local entropy which is computed from S_k .

III. PROPOSED METHOD

The method proposed is divided into 2 main stages (A and B) and named as MSS (Modified Scale-Space Analysis).

A. Adaptive Multiscale Gaussian Filter for Suppressing Amplitude Spectrum

In order to suppress the spectrum of the repeating area, adaptive multiscale Gaussian filter is proposed. It is extended from an adaptive median filter, which is a powerful method to remove the noise from the input image since the amplitude spectrum of the repeating area is similar to the noise image. There are many spikes on the image. The flowchart of the adaptive multiscale Gaussian filter is graphically depicted in

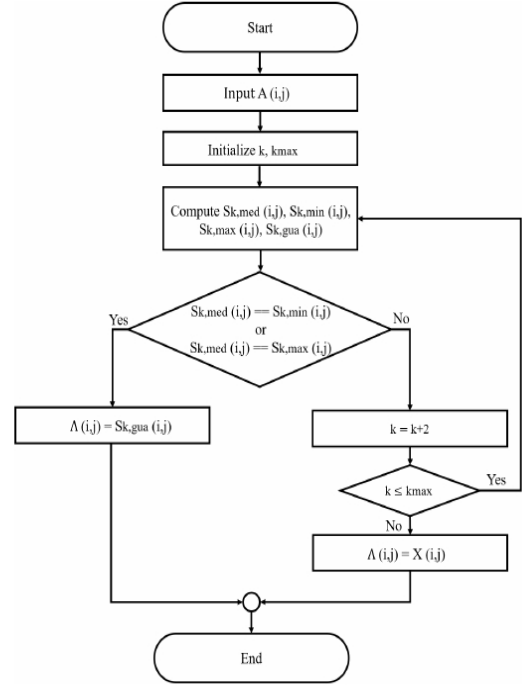


Fig. 2: The flowchart of adaptive multiscale Gaussian filter.

Fig. 2. Let A be the original amplitude spectrum and G_k be the Gaussian filter, Λ is the smoothed amplitude spectrum. The flowchart can be described step by step as follows.

1) Initializing step: the parameter k of the Gaussian filter in Eq. (4) is set to "3".

2) Computing median, max and min step: the local-min, local-max and local-median are defined in Eqs. (16) to (18).

$$S_{k,min}(i, j) = \min\{A(a, b) \mid \frac{i-k}{2} \leq a \leq \frac{i+k}{2} \text{ and } \frac{j-k}{2} \leq b \leq \frac{j+k}{2}\} \quad (16)$$

$$S_{k,max}(i, j) = \max\{A(a, b) \mid \frac{i-k}{2} \leq a \leq \frac{i+k}{2} \text{ and } \frac{j-k}{2} \leq b \leq \frac{j+k}{2}\} \quad (17)$$

$$S_{k,med}(i, j) = \text{median}\{A(a, b) \mid \frac{i-k}{2} \leq a \leq \frac{i+k}{2} \text{ and } \frac{j-k}{2} \leq b \leq \frac{j+k}{2}\} \quad (18)$$

where $S_{k,med}(i, j)$, $S_{k,max}(i, j)$ and $S_{k,min}(i, j)$ are the local-median, local-max and local-min, respectively, in a sigma k .

3) Adjusting sigma step: the parameter k is added by two when $S_{k,med}(i, j)$ is equal to $S_{k,max}(i, j)$ or $S_{k,min}(i, j)$ and if k is less than k_{max} where k_{max} is a maximum scale then, go to the second step.

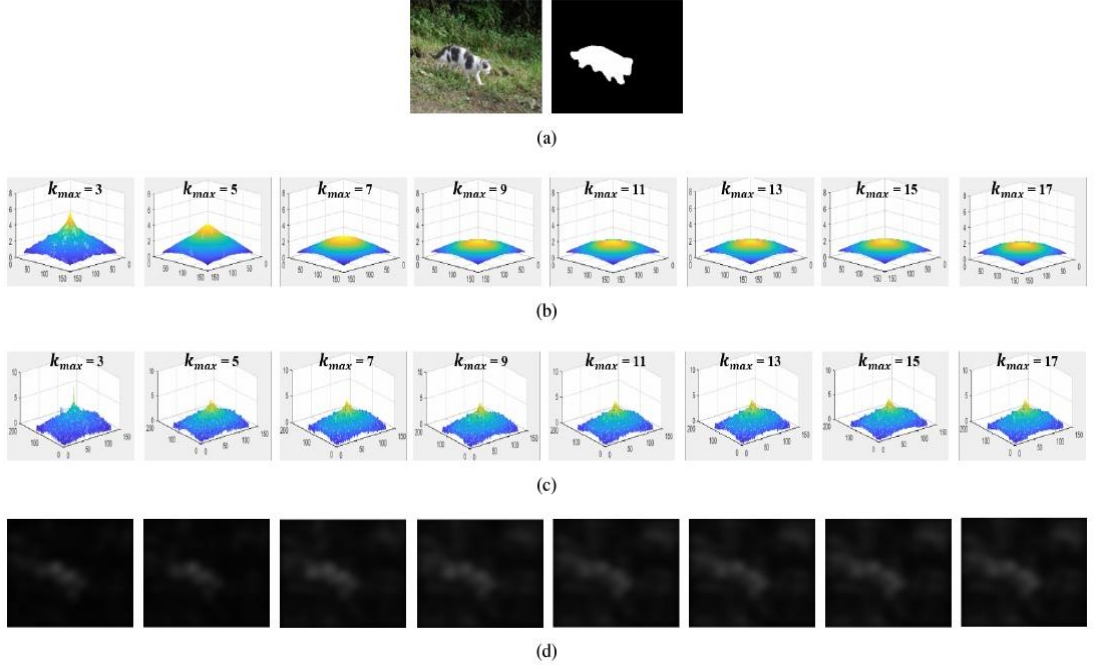


Fig. 3: An illustration of varying the parameter k of the adaptive multiscale Gaussian filter in different k_{max} : (a) given image and ground truth, (b) adaptive multiscale Gaussian Filter, (c) smoothed amplitude spectrum and (d) saliency map.

TABLE I: A comparison of performance of the proposed method and baseline methods based on AUC values.

Model	Category I	Category II	Category III	Category IV	Category V	Category VI	Overall
Itti	0.9111	0.9155	0.9349	0.8435	0.9165	0.9419	0.9147
SUN	0.7423	0.7645	0.8365	0.6812	0.7081	0.7938	0.7719
SR	0.8131	0.8512	0.9285	0.7558	0.7911	0.8996	0.8574
SAL	0.8693	0.9028	0.9320	0.8604	0.9050	0.9472	0.9037
HFT	0.9432	0.9135	0.9221	0.9440	0.9200	0.9530	0.9244
MSS	0.9376	0.9231	0.9230	0.9564	0.9307	0.9566	0.9283

4) Suppressing amplitude spectrum step: $\Lambda(i, j)$ is set to $S_{k, gau}(i, j)$ when $S_{k, med}(i, j)$ is not equal to $S_{k, max}(i, j)$ or $S_{k, min}(i, j)$. $S_{k, gau}(i, j)$ is computed by using Eq. (19).

$$S_{k, gau} = G_k * A \quad (19)$$

However, $\Lambda(i, j)$ is set by $A(i, j)$ when k is higher than k_{max} . Fig. 3 illustrates the smoothed amplitude spectrum and the saliency map in different k_{max} .

B. Modified Scale-Space Analysis Method

The proposed method MSS extended from Jian's method [6] consists of three main steps. The method uses the adaptive multiscale Gaussian filter in conjunction with the conventional Gaussian filter which can be explained as follows.

1) Feature Map Extraction and Hypercomplex Fourier Transform: the amplitude, phase and eigenaxis spectra are given by using Eqs. (11) to (13), respectively, as explained in section II(B).

2) Adaptive Spectrum Scale-Space: the adaptive multiscale Gaussian filter as explained in section III(A) is used

to suppress the amplitude spectrum in conjunction with the conventional Gaussian filter in [6]. The result is a set of sixteen smoothed amplitude spectra which the first eight spectra are extracted from the conventional Gaussian filter, while the residual spectra are obtained from the proposed filter. The spectra forward to saliency map extraction. Furthermore, the scale k_{max} of the adaptive multiscale Gaussian filter is set by $2k + 1$ of the conventional Gaussian filter. For example, in the first round, the k_{max} of the proposed method is set to "3" while the scale k of conventional Gaussian filter is set to "1".

3) Saliency Map Extraction: all candidate saliency maps are computed by using Eq. (15), and then they are selected to be the saliency maps based on the minimum local entropy.

IV. EXPERIMENT AND DISCUSSION

In order to evaluate the performance of the proposed method for the object region prediction which humans pay attention by comparing the performance with the baseline methods, these methods are Itti [2], SUN [3], SR [4], SAL [5] and HFT [6]. The experiment is set up by using the Salmap

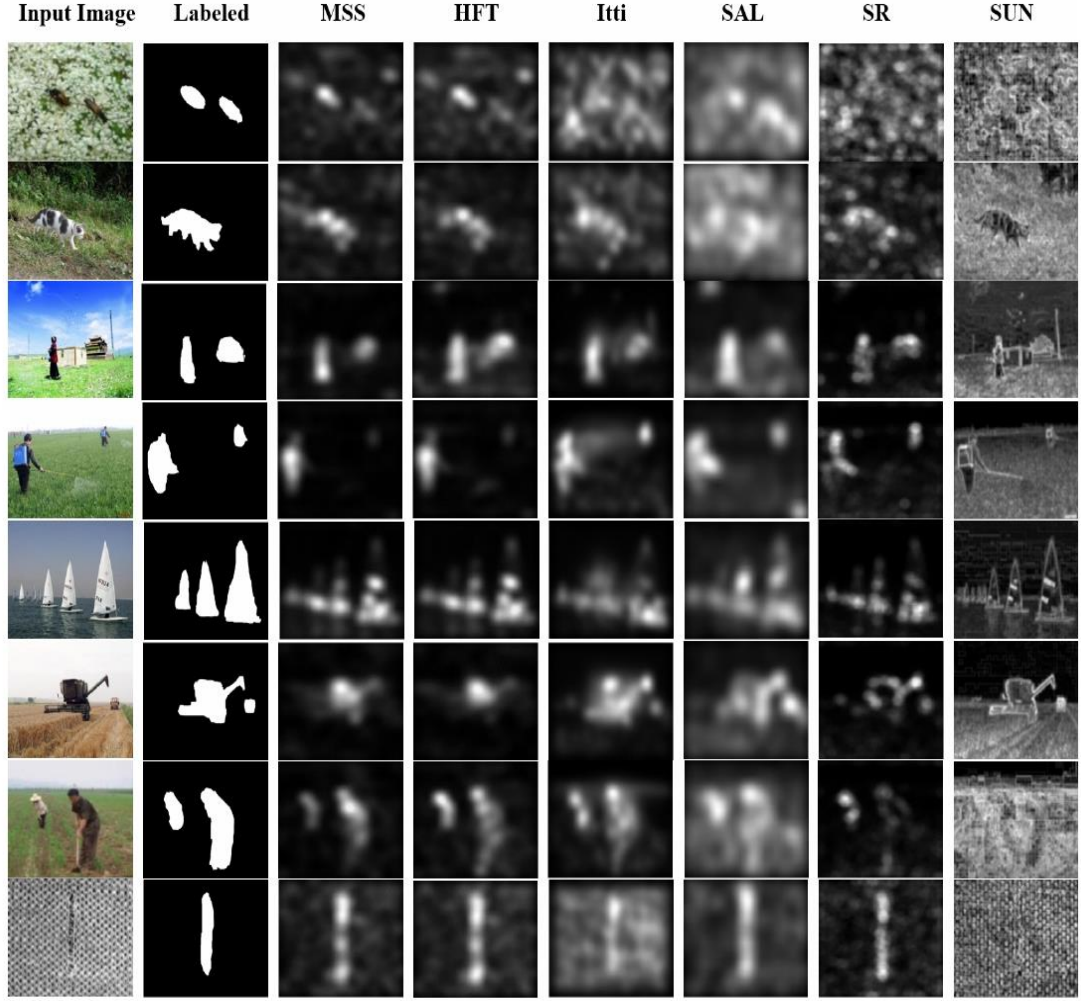


Fig. 4: A comparison of the saliency map obtained from the proposed method (MSS) and baseline methods including HFT [6], Itti [2], SAL [5], SR [4] and SUN [3] methods, respectively.

data set containing 235 images with 640×480 pixels. The data set is divided into six categories—(I) large salient region, (II) intermediate salient regions, (III) small salient regions, (IV) cluttered backgrounds, (V) repeating distractors and (VI) both large and small salient regions. The performance of the proposed method and the baseline methods is evaluated by comparing between the saliency map extracted from each method and salient region labeled by human observers as defined in [6]. Moreover, the performance of the proposed and baseline methods is evaluated by using an area under the curve (AUC) (or know as receiver operating characteristic curve (ROC)). The range of AUC is between 0 and 1 where "0" means that the saliency map is not corresponding to salient region labeled by human observers while "1" means that it is corresponding to human observers.

In table I, it shows that the AUC of the proposed method is higher than all of the baseline methods, in four categories—

(II) intermediate salient regions, (IV) cluttered backgrounds, (V) repeating distractors and (VI) both large and small salient regions with the difference of 0.0095, 0.0124, 0.0107 and 0.0036, respectively. While the proposed method performs a little lower than HFT in (I) large salient regions, and is the third place in (III) small salient regions, which HFT and Itti performs best and second place in this categories. In (I) large salient regions, HFT is more suitable than our proposed adaptive Gaussian filter method. It is probably because the large area is quite straightforward to be detect, then HFT with a fixed-parameter Gaussian filter is effective here. Similarly to in (III) small salient regions, where the single small area is easy to detect with the specific selected filter in Itti. Even though the proposed method does not perform best in all categories, the overall performance of the proposed method increases by 0.0136, 0.1564, 0.0709, 0.0246 and 0.0039 when compared to the Itti, SUN, SR, SAL and HFT methods, respectively.

Moreover, Fig. 4 shows that the saliency map obtained from the proposed method provides better performance than the baseline methods because the amplitude spectrum of the repeated pattern is suppressed by the proposed filter more than that from the HFT's filter. The appropriate scale of the proposed filter is selected based on the characteristic of the spectrum in each area, thus making it adaptive local processing.

V. CONCLUSION

In this paper, the modified scale-space analysis (MSS) in the frequency domain based on adaptive multiscale Gaussian filter is proposed. The performance of the proposed method is evaluated by using the Salmat dataset in terms of the receiver operating characteristic curve. The experimental result shows that the overall of the proposed method outperforms the baseline methods—Itti, SUN, SR, SAL and HFT, especially, in case of the intermediate salient regions, cluttered backgrounds, repeating distractors and both large and small salient regions which are high complexity. The contribution in this paper is the use of the adaptive multiscale Gaussian filter for suppressing the amplitude spectrum to obtain the saliency map. The proposed filter can select the scale automatically to proper each region in amplitude spectrum based on the statistical solution.

REFERENCES

- [1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 1597–1604.
- [2] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
- [3] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *Journal of Vision*, vol. 8, no. 7, pp. 32.1–20, Dec. 2008.
- [4] X. Hou and L. Zhang, "Saliency Detection: A Spectral Residual Approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp. 1–8.
- [5] X. Hou, J. Harel, and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, Jan. 2012.
- [6] J. Li, M. D. Levine, X. An, X. Xu, and H. He, "Visual Saliency Based on Scale-Space Analysis in the Frequency Domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, pp. 996–1010, Apr. 2013.

Improved Adaptive Spectrum Scale-Space in Frequency Domain for Saliency Detection

Taravichet Titijaroonroj[†], Jenjira Jaemsiri* and Jaratsri Rungrattanaubol[‡]

*[†]Faculty of Science, Department of Computer Science and Information Technology,
Naresuan University, Phitsanulok, 65000, THAILAND

Email: jenjiraja61@email.nu.ac.th* and jaratsrir@nu.ac.th[‡]

[†]Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang,
Bangkok, 10520, THAILAND. Email: taravichet@it.kmitl.ac.th[†]

Abstract—Saliency detection is a process to find the significant region on the images, which is popular in the image processing area. It has been extended and used in many applications in computer vision systems. Many researches have attempted to propose effective model to detect the saliency area which is corresponding to human perception. Therefore, this research focuses on improving the saliency detection process by proposing the improved adaptive spectrum scale-space (IASSS). The main contributions of the proposed method include (i) scale-and-space Gaussian filter (AS²G filter) and (ii) the new method for saliency map selection based on local entropy. Firstly, the AS²G filter is used to suppress the non-saliency amplitude spectrum to extract the saliency map. Then, the best saliency map is selected from the results of the RGB and Lab color images by using the local entropy criteria. Then, the experimental results based on 235 images show that the overall performance of the proposed IASSS method outperforms the baseline methods including Itti, SR, SUN, SAL, HFT, and MSS methods.

Keywords—Scale-and-space Gaussian filter; local entropy; scale-space analysis; saliency detection

I. INTRODUCTION

Saliency detection on the natural scenery images is to find the most outstanding characteristics which is normally related to human perception. People can immediately point out the objects or regions which are distinctive and different from the surrounding area. The result of the saliency detection is the saliency area which contains important information such as shape, size, and position. The saliency detection is useful in many applications such as image segmentation [1], object tracking [2], and object recognition [3].

Many researches [4]–[10] have attempted to simulate the model which imitates a visual system of human perceptual abilities. For example, in 1998, Itti et al. [4] proposed a model of saliency-based visual attention for rapid scene analysis by using three main feature maps (e.g. color, intensity, and orientations). Each feature map is computed based on a linear operator as so-called "center-surround", it is employed to extract the saliency map. In 2007, Hou et al. [6] presented the spectral residual (SR) approach based on the log-spectrum of an input image on the frequency domain, which found that the difference between the original log-spectrum and its smoothed log-spectrum by the Gaussian filter as so-called "spectrum residual" is equivalent to the saliency map in the spatial domain. Hence, the Gaussian filter plays an important role to obtain the smoothed log-spectrum. In other words,

the good saliency map is received if the Gaussian filter can suppress the non-saliency area or the repeated patterns. After that, in 2012, Hou et al. [7] introduced the image signature based on discrete cosine transform (DCT), which approximated the location and area of a foreground hidden in a complex background based on a sign function of the given image. In 2012, Li et al. [8] proposed the solution to extract the saliency map from the RGB color image based on spectrum scale space, which consists of four main steps—feature map extraction, hypercomplex Fourier transform, spectrum scale-space, and saliency map extraction, as shown in Fig. 1(a). The saliency map is approximated from suppressing the amplitude spectrum by using the Gaussian filter. Although, the performance of Li's method is better than the selected baseline method on the Salmat standard dataset, the conventional Gaussian filter used is not adaptive, and all parameters of filters are fixed. Later on, Jaemsiri et al. [10] proposed the modified spectrum scale space based on adaptive multiscale Gaussian filter (MSS) as depicted in Fig. 1(b), which aimed to provide adaptive multiscale Gaussian (AMG) filter in order to suppress the repeat pattern on the amplitude spectrum of each region. The performance of MSS outperformed some selected baseline methods. However, the size of Gaussian filter is still fixed.

In order to improve the performance, adaptive scale-and-space Gaussian filter, which is extended from the AMG filter is proposed here. It provides an automatically adaptive scale and space filter instead of being fixed size. From our review, we found that many researches have tried to detect the saliency map from a single color space [4]–[8] but, recently Xiang et al. [9] proposed an approach to construct the saliency map from two color spaces, which are the RGB and Lab color images. The result illustrated that the saliency maps extracted from the RGB and Lab color image are different depending on the characteristic of each image, even though the two images look identical. Xiang's method is quite effective, however, it lacks the effective criteria in selecting the accurate or final saliency map between the maps constructed from RGB and Lab color. To improve the performance under this issue, this paper proposed a local entropy as a criteria for selecting the map in our saliency detection. The contribution in this paper consists of two main points—(i) the adaptive scale-and-space Gaussian filter and (ii) the saliency map selecting based on the local entropy as explained in section II(B) and II(C), respectively.

The rest of the paper is organized as follows. Section II

gives details of our proposed method, followed by the experiment and discussion in section III. Finally, the conclusion is discussed in section IV.

II. PROPOSED METHOD

In this section, we proposed our saliency detection model, which includes three main stages—(A) proposed saliency detection, (B) improved adaptive spectrum scale-space, and (C) saliency map selection based on local entropy. Fig. 1 displays the workflow of our proposed method called IASSS in Fig. 1(c), with a comparison with the other two methods, HFT and MSS, depicted as Fig. 1(a) and 1(b), respectively.

A. Proposed Saliency Detection

Our proposed method (IASSS) extended from the MSS method [10] aims to improve the performance of the saliency detection. Its process is divided into five steps as follows.

1) Lab Color Transform: The input image in RGB color space is converted into Lab color formats, then both RGB and Lab color images are used in saliency detection since the characteristics of saliency detection from RGB and Lab color images are different in color space [9] as depicted in Fig. 2.

2) Hypercomplex Fourier Transform: The RGB and Lab color images are transformed to the frequency domain by using hypercomplex Fourier transform as defined in Eq. (1).

$$F_h = HFT(f_h) \quad (1)$$

where $HFT(\cdot)$ is the hypercomplex Fourier transform, F_h and f_h are a hypercomplex matrix in the frequency and the spatial domains, respectively, as explained in [8]. The amplitude spectrum (A), phase (P) and eigenaxis (X) spectrum are computed based on F_h as explained in [8], [10]. In this paper, the amplitude spectrum is essential information that can be used to locate the repeated pattern and suppress it to detect the saliency map.

3) Improved Adaptive Spectrum Scale-Space: The amplitude spectrum is suppressed by using adaptive scale-and-space Gaussian filter to smooth amplitude spectrum. It is used to obtain the candidates of the saliency map. The detail of this step is given in section II(B).

4) Saliency Map Extraction: The smooth amplitude spectrum Λ is converted by using the inverse hypercomplex Fourier transform $IHFT(\cdot)$ to obtain the saliency map candidates. A set of the saliency maps S is defined in Eq. (2).

$$S = \{S_k^G, S_{kmax}^A\} \quad (2)$$

where S_k^G and S_{kmax}^A are the saliency map candidates as defined in Eqs. (3) - (5).

$$S_k^G = SME(\Lambda_k^G) \quad (3)$$

$$S_{kmax}^A = SME(\Lambda_{kmax}^A) \quad (4)$$

$$SME(\Lambda) = g_\sigma * ||IHFT(\Lambda e^{XP})||^2 \quad (5)$$

where g_σ is a fixed scale Gaussian filter ($\sigma = 0.05$). Λ_k^G and Λ_{kmax}^A are a smooth amplitude spectrum suppressed by using the conventional Gaussian filter as explained in [8] at scale $k \in \{1, 2, 3, \dots, 8\}$ and the proposed AS²G filter as described in section II(B) at scale $kmax \in \{3, 5, 7, \dots, 17\}$, respectively.

Algorithm 1: Adaptive Scale-and-Space Gaussian Filter

Input: The amplitude spectrum A and $kmax$;

Output: The smoothed amplitude Λ .

```

1 for  $u = 1$  to  $U$  do
2   for  $v = 1$  to  $V$  do
3     Set  $k = 3$ ;
4     while  $k \leq k_{max}$  do
5       Set  $\sigma = k$  and  $\rho = k$ ;
6       Generate Gaussian filter  $G(\sigma, \rho)$ ;
7       Suppressing amplitude  $\tilde{A}(u, v)$ ;
8       Find maximum spectrum  $S_{k,max}(u, v)$ ;
9       Find minimum spectrum  $S_{k,min}(u, v)$ ;
10      Find median spectrum  $S_{k,med}(u, v)$ ;
11      if  $S_{k,med}(u, v) \geq S_{k,min}(u, v)$  &
          $S_{k,med}(u, v) \leq S_{k,max}(u, v)$  then
12        Set  $\tilde{A}(u, v)$  to  $\Lambda(u, v)$ ;
13        Break;
14      else
15         $k = k + 2$ ;
16        if  $k \leq k_{max}$  then
17          Continue;
18        else
19          Set  $\tilde{A}(u, v)$  to  $\Lambda(u, v)$ ;
20        end
21      end
22    end
23  end
24 end
```

The minimal entropy is used to determine the optimal saliency map from the set of the saliency map candidates in each color space. The minimum of saliency map is defined in Eq. (6).

$$S' = \arg \min \{H_{Entropy}(S)\} \quad (6)$$

where $H_{Entropy}(\cdot)$ is an entropy function of the saliency map and S' is a selected saliency map. S'_{RGB} and S'_{Lab} are the saliency maps, which are resulted from this step, when RGB and Lab color images are used, respectively.

5) Saliency Map Selection Based on Local Entropy: The final saliency map is selected from the RGB and Lab saliency maps by minimal local entropy as computed by Eq. (7).

$$S_{final} = \arg \min \{H_{LE}(S'_{RGB}), H_{LE}(S'_{Lab})\} \quad (7)$$

where $H_{LE}(\cdot)$ is a local entropy function as explained in section II(C). S'_{RGB} and S'_{Lab} are the RGB and Lab color saliency map, respectively.

B. Improved Adaptive Spectrum Scale-Space

This section aims to improve the performance of the adaptive spectrum scale-space for saliency detection presented by Jaemsiri et al. as explained in [10]. The achievement of Jaemsiri's method is suppressing the spectrum by using adaptive multiscale Gaussian filter. Although its performance is higher than that of the baseline methods, its process is still lack of parameter-free setting such as filter size. Hence, we proposed the adaptive spectrum scale-space by modifying the

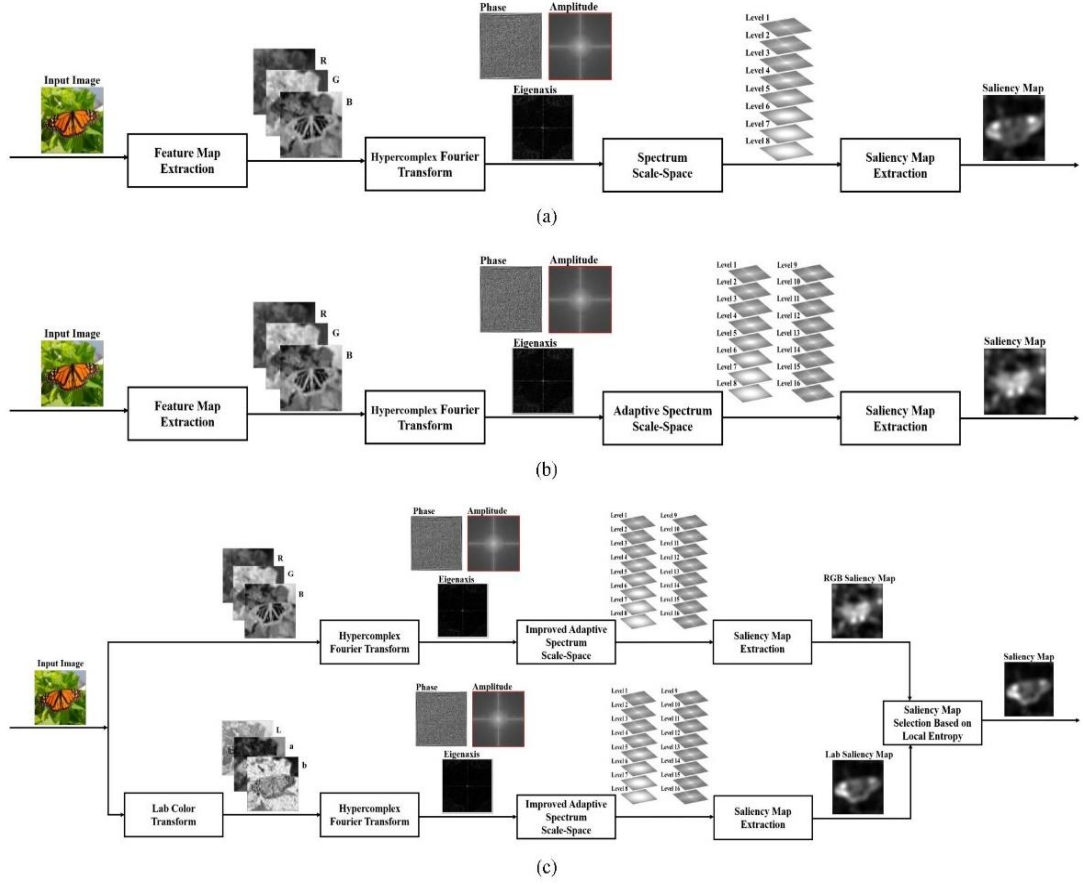


Fig. 1: The workflow of (a) HFT [8], (b) MSS [10] and (c) proposed IASSS methods.

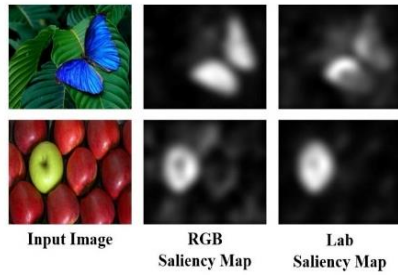


Fig. 2: A comparison of the saliency map extracted from RGB and Lab color spaces.

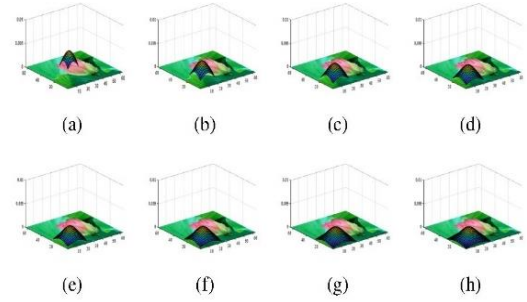


Fig. 3: An illustration of the adaptive scale-and-space Gaussian filter.

adaptive multiscale Gaussian filter based on the adaptive scale and space as so-called adaptive scale-and-space Gaussian filter (AS²G filter). The maximum, minimum, and median of the amplitude spectrum are used for automatically adjusting the parameters. As illustrated in Fig. 3, Fig. 3(a) is an initial AS²G filter. Then, its parameters including the sigma and size are

increased until they are equal or higher than the maximum boundary k_{max} as shown in Algorithm 1 at line 11. The result of parameter adjustment is demonstrated on Figs. 3(b) - 3(h), respectively.

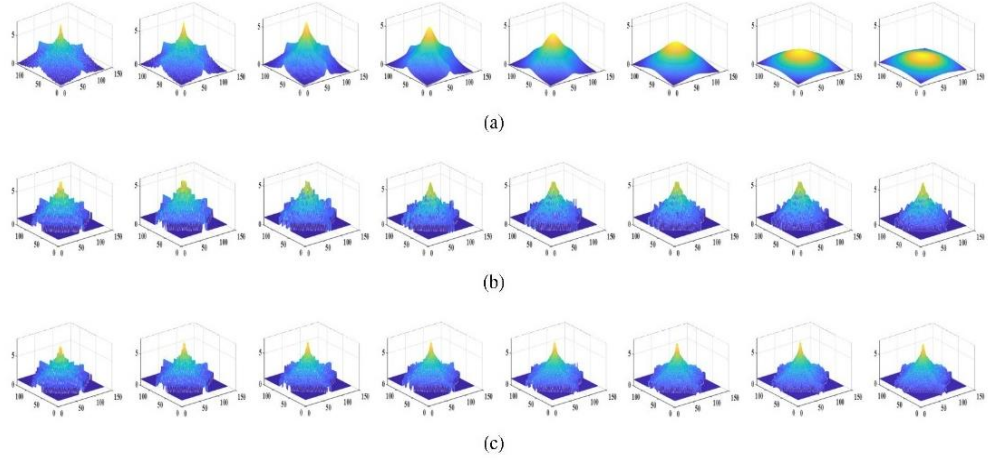


Fig. 4: The smooth amplitude spectrum obtained from (a) HFT, (b) MSS, and (c) proposed IASSS methods.

The original amplitude spectrum A with $U \times V$ resolution is suppressed by using the AS^2G filter to obtain the smooth amplitude spectrum $\hat{A}(u, v)$. The pseudo code of the proposed AS^2G filter is displayed in Algorithm 1, which can be explained step by step as follows.

1) Initialize step: Set k to "3" as recommended in [10].

2) Update Gaussian parameter step: Set the parameters σ and ρ to k if k is lower than or equal to k_{max} where σ is a standard deviation and ρ is window size of the Gaussian filter.

3) Generate Gaussian filter step: Generate filter by using Eq. (8).

$$G(u, v, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/(2\sigma^2)} \quad (8)$$

where (u, v) is a position located at the filter's sizes $\rho \times \rho$.

4) Suppress amplitude spectrum step: Compute the candidate of the smooth amplitude spectrum ($\tilde{A}$) by using Eq. (9).

$$\tilde{A}(u, v) = A(u, v) * G(u, v, \sigma) \quad (9)$$

where A is the amplitude spectrum and $*$ is a convolution operation.

5) Find maximum, minimum, and median step: The maximum, minimum, and median of the amplitude spectrum are defined in Eqs. (10) to (12).

$$S_{k,max}(u, v) = \max\{A(a, b) \mid \epsilon \leq a \leq \tilde{\epsilon} \text{ and } \omega \leq b \leq \tilde{\omega}\} \quad (10)$$

$$S_{k,min}(u, v) = \min\{A(a, b) \mid \epsilon \leq a \leq \tilde{\epsilon} \text{ and } \omega \leq b \leq \tilde{\omega}\} \quad (11)$$

$$S_{k,med}(u, v) = \text{median}\{A(a, b) \mid \epsilon \leq a \leq \tilde{\epsilon} \text{ and } \omega \leq b \leq \tilde{\omega}\} \quad (12)$$

where ϵ is $u - \frac{(\rho-1)}{2}$, $\tilde{\epsilon}$ is $u + \frac{(\rho-1)}{2}$, ω is $v - \frac{(\rho-1)}{2}$ and $\tilde{\omega}$ is $v + \frac{(\rho-1)}{2}$.

6) Define the smooth amplitude spectrum step: As shown in Algorithm 1 at line 11, if $S_{k,med}(u, v)$ is higher than $S_{k,min}(u, v)$ and $S_{k,med}(u, v)$ is lower than $S_{k,max}(u, v)$ set $\hat{A}(u, v)$ to $\tilde{A}(u, v)$. Whereas the parameter k is added by two and then go back to the update Gaussian parameter step if k is lower than or equal to k_{max} . Otherwise, $\hat{A}(u, v)$ is set to $\tilde{A}(u, v)$ when k is higher than k_{max} . The smooth amplitude spectrum $\hat{A}$ obtained from the HFT, MSS, and IASSS methods is depicted in Fig. 4. It can be seen that the artifact on the amplitude spectrum obtained from the AS^2G filter is lower than the MSS method.

C. Saliency Map Selection Based on Local Entropy

This section aims to point out the classical saliency map selection problem which leads to an incorrect selection as depicted in Fig. 5. It can be seen that the best saliency map can be obtained but it is not selected. This problem comes from the global entropy cannot capture the best saliency map. In order to improve the saliency map selection, the local entropy H_{LE} is required due to the fact that it can present the complexity contained in a given neighborhood of the given image as defined in Eq. (13).

$$H_{LE} = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y H_{Entropy}(x, y) \quad (13)$$

where X and Y are the width and height of the input image. $H_{Entropy}(\cdot)$ is the entropy of the saliency map as defined in Eq. (14).

$$H_{Entropy}(x, y) = - \sum_{i=1}^{256} h_i(x, y) \times \log_2 h_i(x, y) \quad (14)$$

where $h_i(x, y)$ is a normalized histogram counts of the sub-image which its center is located at the position (x, y) and

TABLE I: The performance based on AUC value of the proposed IASSS and the baseline methods including Itti, SUN, SR, SAL, HFT, and MSS.

Model	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6	Overall
Itti	0.9111	0.9155	0.9349	0.8435	0.9165	0.9419	0.9147
SUN	0.7423	0.7645	0.8365	0.6812	0.7081	0.7938	0.7719
SR	0.8131	0.8512	0.9285	0.7558	0.7911	0.8996	0.8574
SAL	0.8693	0.9028	0.9320	0.8604	0.9050	0.9472	0.9037
HFT	0.9432	0.9135	0.9221	0.9440	0.9200	0.9530	0.9244
MSS	0.9376	0.9231	0.9230	0.9564	0.9307	0.9566	0.9283
IASSS	0.9399	0.9262	0.9308	0.9633	0.9287	0.9559	0.9311

TABLE II: The comparison of using various options from MSS to IASSS method

Pattern	RGB and Lab color space	AS ² G filter	Global Entropy	Local Entropy	AUC
1	×	×	×	×	0.9283
2	✓	×	✓	×	0.9291
3	✓	×	×	✓	0.9301
4	×	✓	×	×	0.9269
5	✓	✓	✓	×	0.9290
6	✓	✓	×	✓	0.9311

Note: ✓ using the options, which are RGB and Lab color space, AS²G filter, Global Entropy and Local Entropy. × means not using the options.

its nearest neighbour 9×9 pixels on bin No. i . Based on our empirical experiment, the appropriate size of the neighborhood is set to 9 when comparing to the size of neighborhood 3, 5 and 7. For saliency map detection, the high entropy value means that there are many noises mixed in the saliency map whereas the low entropy value means that there are a few noises mixed in the saliency map.

III. EXPERIMENT AND DISCUSSION

This section aims to evaluate the performance of the proposed method extended from the MSS method [10]. It consists of two main parts—(i) experimental setup and (ii) result and discussion as follows.

A. Experimental Setup

The performance of the proposed method is evaluated by using 235 images from the Salmap dataset which is a standard and well-known dataset. This dataset can be divided into six categories—(1) large salient regions with 50 images, (2) intermediate salient regions with 80 images, (3) small salient regions with 60 images, (4) repeating distractors with 15 images, (5) cluttered backgrounds with 15 images and (6) both large and small salient regions with 15 images. All saliency maps extracted from the proposed IASSS, MSS [10], HFT [8], Itti [4], SAL [7], SR [6], and SUN [5] methods are measured for their performance by using an area under the curve (AUC) with the ground truth labeled by humans.

B. Result and Discussion

In the first part, we will explain the improvement of the proposed method, which is extended from the MSS method. Based on TABLE I, it shows that the performances of the proposed method comparing to MSS in the categories (1) to (4) are better than that of the MSS method, with the percentage difference of 0.24%, 0.33%, 0.84%, and 0.72%, respectively. Even though the performance of IASSS in the categories (5)

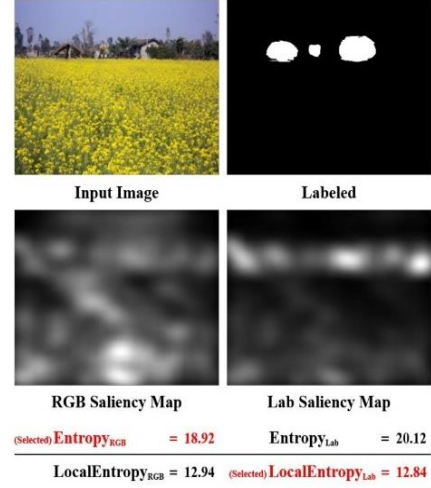


Fig. 5: An example of the incorrect saliency map selection.

and (6) are 0.22% and 0.8%, respectively, lower than that of the MSS method, the overall value indicates that the performance of the IASSS is 0.30% higher than the MSS method.

In the second part, we aim to evaluate the performance of the proposed method with the baseline methods. The experimental results in TABLE I show that the performance of the proposed method is the best in the category 2 and 4 and performs the second best in categories 1, 5 and 6, but is the third in category 3. However, the overall of the performance of the proposed method is 0.0164, 0.1592, 0.0737, 0.0274, 0.0067, and 0.0028 higher than that of the Itti, SUN, SR, SAL, HFT, and MSS methods, respectively. Moreover, Fig. 6 shows that the saliency map extracted from the proposed method outperforms the baseline methods. It has minimal noise and is closely corresponding to the ground truth labeled by humans.

TABLE II shows that the performance of using the RGB-and-LAB color space with the global entropy criteria (Pattern 2) is higher than that of using only RGB color space (Pattern 1, i.e. MSS). It means that using both color spaces can increase the performance of the proposed method. Moreover, its performance can be improved when the RGB-and-LAB color space is used with local entropy criteria as shown in (Pattern 3) from TABLE II. This is a prove-by-experiment that the local entropy criteria can improve the efficiency of the proposed method. Although our performance is reduced when

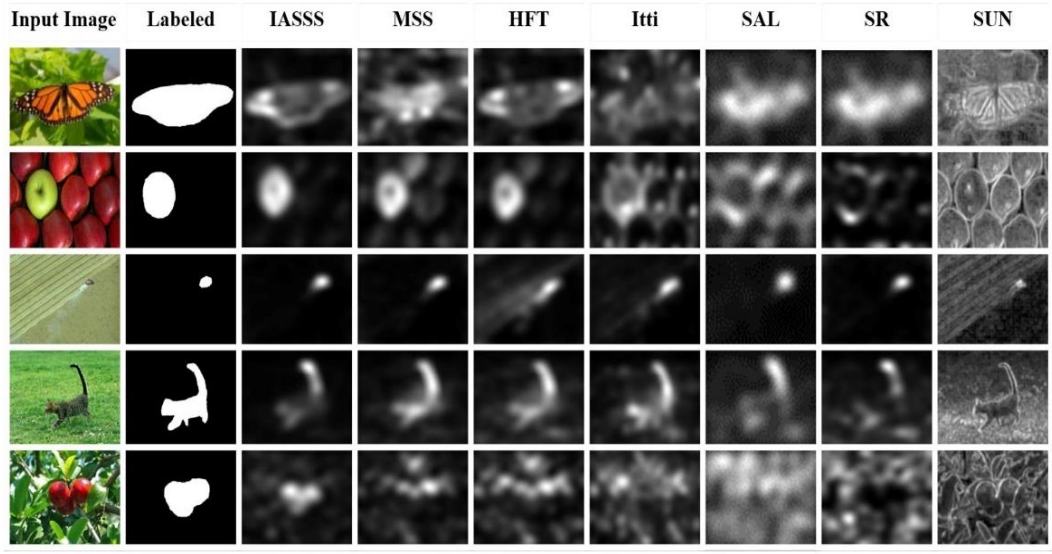


Fig. 6: A comparison of the saliency map extracted from IASSS, MSS, HFT, Itti, SAL, SR, and SUN.

the AS^2G filter is employed in (Pattern 4), it can increase when it is used with the RGB-and-LAB color space and the local entropy criteria in (Pattern 6, i.e. IASSS). We can conclude that the improvement of the proposed method derives from three main points including (i) incorporating useful salient information in the RGB and Lab color spaces, (ii) using the AS^2G filter to suppress the amplitude spectrum of the repeated pattern. It can automatically adjust the scale and space filter to be suitable for in each image (i.e. parameter-free), and (iii) avoiding the incorrect saliency map selection to obtain the map which has a few noise mixed. As mentioned before, this leads to obtaining the best saliency map from the proposed method which is closely similar to the ground truth labeled by a human.

IV. CONCLUSION

This paper proposes the improved spectrum scale space (IASSS) for saliency detection. The achievement of this method comes from (i) the adaptive scale-and-space Gaussian (AS^2G) filter for suppressing the amplitude spectrum of the repeated pattern and (ii) the local entropy criteria for saliency map selection from the RGB and Lab color image. The AS^2G filter is an adaptive filter in scale (σ) and space (ρ) for each image automatically. Moreover, to avoid the incorrect selection problem, the local entropy is used. They are a key factor to improve the performance of the saliency detection. The experimental results show that the overall performance based on AUC of the proposed ISSA method is higher than that of the baseline methods including Itti, SAL, SUN, SR, HFT, and MSS methods in the Salmap dataset.

REFERENCES

- [1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 1597–1604.
- [2] G. Zhang, Z. Yuan, N. Zheng, X. Sheng, and T. Liu, "Visual Saliency Based Object Tracking," in *Computer Vision ACCV 2009*, ser. Lecture Notes in Computer Science, H. Zha, R.-i. Taniguchi, and S. Maybank, Eds. Springer Berlin Heidelberg, 2010, pp. 193–203.
- [3] Z. Ren, S. Gao, L. Chia, and I. W. Tsang, "Region-Based Saliency Detection and Its Application in Object Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 5, pp. 769–779, May 2014.
- [4] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
- [5] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *Journal of Vision*, vol. 8, no. 7, pp. 32–32, May 2008.
- [6] X. Hou and L. Zhang, "Saliency Detection: A Spectral Residual Approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp. 1–8.
- [7] X. Hou, J. Harel, and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, Jan. 2012.
- [8] J. li, M. Levine, X. An, X. Xu, and H. He, "Visual Saliency Based on Scale-Space Analysis in the Frequency Domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 996–1010, Nov. 2012.
- [9] D. Xiang and B. Zhong, "Scale-space saliency detection in combined color space," in *2015 Chinese Automation Congress (CAC)*, Nov. 2015, pp. 726–731.
- [10] J. Jaemsiri, T. Titjaroonroj, and J. Rungrattanaubol, "Modified Scale-Space Analysis in Frequency Domain Based on Adaptive Multiscale Gaussian Filter for Saliency Detection," in *2019 International Joint Conference on Computer Science and Software Engineering IJCSSE*, Jul. 2019, pp. 212–217.