

การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้กล้องและเซ็นเซอร์ภาพของ
เครื่องบันทึกการตามองคนไข้
PREDICTION OF HUMAN EMOTIONS STIMULATED BY IMAGES
USING IMAGE FEATURES AND EYE TRACKING DEVICE



มีลิขสิทธิ์เป็นของตนเองและสงวนสิทธิ์ในสิ่งที่ปรากฏในเอกสารฉบับนี้โดยไม่มีส่วนเกี่ยวข้องกับ
มหาวิทยาลัยราชภัฏบุรีรัมย์

นางสาววิมลมาศ โนนสีดา ภาควิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยราชภัฏบุรีรัมย์

ภาคเรียนที่ 2 ปีการศึกษา 2567

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้ลักษณะเด่นของภาพและ
เครื่องมือติดตามการมองเห็น

PREDICTION OF HUMAN EMOTIONS STIMULATED BY IMAGES
USING IMAGE FEATURES AND EYE TRACKING DEVICE



T144560

โดย

โชติรส วุฒิเลิศเดชา

CHOTIROS WUTTILERTDESHAR

ภาณวี ฉัตรคำจุนเจริญ

PANAWEE CHATKAMJUNCHAROEN

อาจารย์ที่ปรึกษา

ผู้ช่วยศาสตราจารย์ ดร. กิติ์สุชาติ พลุภา

เลขหมู่.....144560
เลขทะเบียน.....
วัน,เดือน,ปี..25 11๑..2559

600268137
b. 42814490
i.....

600268137

ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในห้องสมุดเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้ลักษณะเด่นของภาพและ
เครื่องมือติดตามการมองเห็น

**PREDICTION OF HUMAN EMOTIONS STIMULATED BY IMAGES
USING IMAGE FEATURES AND EYE TRACKING DEVICE**

โดย



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในเท่านั้น มิใช่เพื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**PREDICTION OF HUMAN EMOTIONS STIMULATED BY IMAGES
USING IMAGE FEATURES AND EYE TRACKING DEVICE**

**CHOTIROS WUTTILERTDESHAR
PANAWEE CHATKAMJUNCHAROEN**



**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2/2014



COPYRIGHT 2015

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ขออนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองปริญญาโท ประจำปีการศึกษา 2557

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้ลักษณะเด่นของภาพ
และเครื่องมือติดตามการมองเห็น

PREDICTION OF HUMAN EMOTIONS STIMULATED BY
IMAGES USING IMAGE FEATURES AND EYE TRACKING
DEVICE

ผู้จัดทำ

1. นางสาวโชติรส วุฒิเลิศเดชา รหัสนักศึกษา 54070019
2. นางสาวภาณวี นัฏรคำจุนเจริญ รหัสนักศึกษา 54070068

..... *Kiradee P.* อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร. กิติ์สุชาติ พสุภา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Project Title	Prediction of Human Emotions Stimulated By Image Using Image Feature And Eye Tracking Device		
Student	Ms. Chotiros Wuttitertdeshar	Student ID	54070019
	Ms. Panawee Chatkamjuncharoen	Student ID	54070068
Degree	Bachelor of Science		
Programme	Information Technology		
Academic Year	2014		
Advisor	Asst. Prof.Dr. Kitsuchart Pasupa		

ABSTRACT

Nowadays, a computer method—emotional semantic image retrieval—enables users to access images that they want in a database. However, different users may experience different emotions depending on where, in the image, they are gazing on. This paper presents an improved prediction method by taking into account the users' eye movement while they are looking at the image. Our experimental results show that using both eye movement tracking and image feature together to determine users' emotion gave more accurate predictions than using image feature alone.

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จไปได้ด้วยดี ด้วยความกรุณาของอาจารย์ที่ปรึกษาโครงการ ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาติ พสุภา และนาย ธราวิชญ์ ธิติจรูญโรจน์ นักศึกษาปริญญาเอกที่ได้สละเวลาให้ความช่วยเหลือ คำแนะนำในการปรับปรุงแก้ไขปัญหาต่าง ๆ ที่เกิดขึ้น ตลอดจนให้ความรู้และประสบการณ์ที่ดีในการทำงาน และขอขอบพระคุณอย่างสูง

ขอกราบขอบพระคุณคณาจารย์คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่านที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า คอยให้คำปรึกษาและชี้แนะแนวทางในการแก้ไขปัญหาดังกล่าว ให้สำเร็จลุล่วงไปได้

ขอขอบคุณพี่ เพื่อน และน้อง คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกคนที่ให้ความช่วยเหลือ คำแนะนำ และกำลังใจเสมอมา

ขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้า ที่ให้กำลังใจ และให้การสนับสนุนในทุกๆเรื่อง ทำให้ข้าพเจ้าสามารถทำโครงการฉบับนี้สำเร็จลุล่วงด้วยดี คุณค่าและประโยชน์อันพึงมาจากโครงการฉบับนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

โชติรส วุฒิเลิศเดชา
ภาณวี วัชรคำจุนเจริญ

สารบัญ

หน้า

บทคัดย่อ	I
ABSTRACT	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญรูป	VIII

บทที่

1. บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	2
1.3 สมมติฐานของการศึกษา	2
1.4 ขอบเขตการศึกษา	2
1.5 ขั้นตอนการปฏิบัติโครงการ	2
1.6 ประโยชน์ที่คาดว่าจะได้รับ	3
2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 ทฤษฎีที่เกี่ยวข้อง	4
2.2 งานวิจัยที่เกี่ยวข้อง	6
2.3 เครื่องมือที่ใช้ในการดำเนินงาน	8
2.4 เทคโนโลยีที่เกี่ยวข้อง	8
3. การดำเนินงานวิจัย	9
3.1 รูปภาพ	9
3.2 การเก็บอารมณ์รูปภาพ	10
3.3 เครื่องมือติดตามการมองเห็น	12
3.4 โปรแกรมเก็บข้อมูล	12
3.5 กระบวนการเตรียมข้อมูล	15
3.6 การสกัดคุณลักษณะเด่นรูปภาพ	17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา **IV** ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

3.7 Leave One-Out CV	20
4. ผลการทดลอง	21
4.1 การดำเนินการทดลอง	21
4.2 ผลการทดลอง	21
4.3 สรุปผลการทดลอง	28
5. สรุปผลและข้อเสนอแนะ	29
5.1 สรุปผลการทดลอง	29
5.2 ข้อดีข้อด้อยของการพัฒนา	29
5.3 ข้อเสนอแนะ	29
บรรณานุกรม	30
ภาคผนวก ก	32
ประวัติผู้เขียน	46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

หน้า

ตารางที่

4.1 ค่าเฉลี่ยการทำนายอารมณ์ทั้ง 20 คนในแต่ละคุณลักษณะ	23
4.2 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์	23
4.3 แสดงค่าเฉลี่ย True Positive ทั้ง 4 อารมณ์ของแต่ละคุณลักษณะ	24
4.4 ค่าเฉลี่ยการทำนายอารมณ์แต่ละผู้ร่วมเก็บข้อมูล	27
ก.1 ผลการเก็บอารมณ์รูปภาพแต่ละผู้ร่วมเก็บข้อมูล	33
ก.2 แสดงข้อมูลดิบทั้ง 20 คน ของตาราง 4.1	34
ก.3 แสดงข้อมูลดิบทั้ง 20 คน ของตาราง 4.3	34
ก.4 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากสี	35
ก.5 p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากสี	35
ก.6 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากรูปร่าง	36
ก.7 p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากรูปร่าง	36
ก.8 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากพื้นผิว	36
ก.9 p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากพื้นผิว	36
ก.10 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากสีและรูปร่าง	37
ก.11 p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากสีและรูปร่าง	37
ก.12 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากสีและพื้นผิว	37
ก.13 p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากสีและพื้นผิว	38
ก.14 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากรูปร่างและพื้นผิว	38
ก.15 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากรูปร่างและพื้นผิว	38
ก.16 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินทั้ง 3 คุณลักษณะ	39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา VI ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

หน้า

ตารางที่

ก.17 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์ทั้ง 3 คุณลักษณะ	39
ก.18 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิตทั้ง 3 คุณลักษณะ	39
ก.19 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์ทั้ง 3 คุณลักษณะ	40



สารบัญรูป

หน้า

รูปที่

3.1 แสดงขั้นตอนการดำเนินงานวิจัย	9
3.2 ตัวอย่างรูปภาพนามธรรม	9
3.3 การเก็บการเคลื่อนไหวของตาโดยเครื่องติดตามการมอง และอารมณ์ของผู้ร่วมการ ทดลองที่มีต่อรูปภาพ	10
3.4 ตัวอย่างรูปภาพ และส่วนที่ถูกลมองของผู้เข้าร่วมที่มีอารมณ์แตกต่างกัน	11
3.5 หน้าจอกรอกข้อมูลผู้ร่วมเก็บข้อมูลอารมณ์รูปภาพ	12
3.6 หน้าจอตรวจสอบตำแหน่งก่อนการปรับเทียบ (Calibration) ของอุปกรณ์การติดตาม การมอง	13
3.7 ทดสอบการปรับเทียบ (Calibration) ของอุปกรณ์การติดตามการมอง	13
3.8 หน้าจอตรวจสอบประสิทธิภาพการมองแต่ละจุดของอุปกรณ์ Eye Tracking	14
3.9 จุดโฟกัสให้ผู้เข้าร่วมเก็บข้อมูลมองก่อนเริ่มดูรูปภาพ	14
3.10 แสดงรูปภาพให้ผู้เข้าร่วมเก็บข้อมูล	15
3.11 (F1) Original Image	15
3.12 (F2) Eye Movements Image	16
3.13 Eye Movements with Gaussian Function Image	17
3.14 ตัวกรองกาบองฟังก์ชัน	18
3.15 การฝึกฝนข้อมูลแบบ Leave-One-Out CV	20
4.1 True Positive อารมณ์มีความสุข	25
4.2 True Positive อารมณ์เศร้า	25
4.3 True Positive อารมณ์โกรธ	25
4.4 True Positive อารมณ์กลัว	26
4.5 ค่าเฉลี่ยของ True Positive ของอารมณ์ในแต่ละคุณลักษณะ	26
ก.1 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากสี	40
ก.2 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากรูปร่าง	41
ก.3 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากพื้นผิว	41
ก.4 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากสีและรูปร่าง	42
ก.5 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากสีและพื้นผิว	42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา VIII ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

ก.6 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากรูปร่างและพื้นผิว .43	
ก.7 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ User Model ในแต่ละอารมณ์ทั้ง 3	
คุณลักษณะ.....	44
ก.8 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ Global Model ในแต่ละอารมณ์ทั้ง 3	
คุณลักษณะ.....	44



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา **IX** ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1. ความเป็นมาและความสำคัญของปัญหา

การค้นคืนภาพ (Image Retrieval) คือ การค้นหารูปภาพที่ผู้ใช้ต้องการจากรูปภาพที่มีในฐานข้อมูล ในอดีตระบบการค้นคืนภาพจะเป็นการค้นคืนโดยใช้คำสำคัญ (Keyword) ที่ถูกแท็ก (Tag) กับรูปภาพแต่ละรูปในฐานข้อมูล ในการแท็กรูปภาพทั้งหมดที่มีอยู่นั้นฐานข้อมูล ต้องใช้แรงงานมาก จึงส่งผลให้มีการพัฒนาระบบเพื่อมีการแท็กรูปภาพอัตโนมัติ [1] หรือใช้การค้นคืนภาพแบบอิงเนื้อหา (Content Based Image Retrieval; CBIR) โดยการเปรียบเทียบภาพคำถาม (Query Image) กับรูปภาพทั้งหมดที่อยู่ในฐานข้อมูล ตัวอย่างของระบบ CBIR คือ PICSOM [2], Google Image Search [3] เป็นต้น จากนั้นในปี 2006 มีการนำเสนอระบบการค้นคืนภาพจากอารมณ์ของรูปภาพ (Emotion Sematic Image Retrieval; ESIR) [4] และเริ่มมีงานวิจัยที่ศึกษาเกี่ยวกับ ESIR มากขึ้น เช่น การศึกษาคุณลักษณะ (Feature) ในการจำแนกอารมณ์รูปภาพโดยใช้ทฤษฎีด้านจิตวิทยาและศิลปะในการหาอารมณ์รูปภาพ [5] การศึกษาอารมณ์ของมนุษย์ที่มีต่อรูปภาพจากคุณลักษณะพื้นฐานของรูปภาพนามธรรม (Abstract Art) [6] และการศึกษาคุณลักษณะของรูปภาพที่มีผลต่อการจำแนกอารมณ์โดยใช้ Multiple Kernel Learning (MKL) โดยพื้นฐานของคุณลักษณะที่ใช้คือ สี, รูปร่างและพื้นผิว [7]

ในระบบ CBIR จะต้องมีการป้อนกลับความเกี่ยวข้อง (Relevance Feedback) เพื่อประเมินตัวอย่างรูปภาพที่ค้นคืนโดยระบบโดยผู้ใช้ นั่นคือการคลิกเมาส์ (Click) ซึ่งการคลิกเมาส์นั้นจัดเป็นการป้อนกลับโดยชัดเจน (Explicit Feedback) ซึ่งเกิดจากความตั้งใจของผู้ใช้ ซึ่งต้องใช้แรงงานในการป้อนกลับ ในปี 2009 ได้มีการเริ่มพัฒนาระบบค้นคืนโดยเริ่มจากการเรียงลำดับของรูปภาพโดยใช้การเคลื่อนไหวของตา [8,9] งานวิจัยนี้ใช้ประโยชน์ของการป้อนกลับโดยปริยาย (Implicit Feedback) ซึ่งเกิดจากความไม่ตั้งใจในขณะที่กำลังเรียงรูปภาพ ในการเพิ่มประสิทธิภาพของระบบ จากนั้นจึงมีการพัฒนาระบบค้นคืนภาพที่ใช้ทั้งการป้อนกลับโดยชัดเจนและการป้อนกลับโดยปริยายที่เรียกว่าระบบ PinView [10]

จากสำนวน “Beauty is in the eye of the beholder” หรือ “ความสวยขึ้นอยู่กับคนมอง” จึงทำให้เกิดคำถามวิจัยที่ว่า รูปภาพหนึ่งรูป และมีคนมองหลายคน อารมณ์ที่ถูกกระตุ้นโดยรูปภาพของแต่ละผู้มองนั้นอาจไม่เหมือนกัน ขึ้นอยู่กับพื้นที่ที่ผู้ใช้กำลังมองรูปภาพนั้นอยู่ งานวิจัยนี้จึงได้นำเสนอการจำแนกอารมณ์ของรูปภาพโดยใช้คุณลักษณะของภาพและการเคลื่อนไหวของตา โดยมีสมมติฐานที่ว่า การใช้คุณลักษณะของภาพร่วมกับข้อมูลการเคลื่อนไหวของตานี้สามารถเพิ่มเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสิทธิภาพในการจำแนกอารมณ์ของผู้ใช้ที่มีต่อรูปภาพจากการใช้คุณลักษณะของภาพอย่างเดียว

1.2. ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

งานวิจัยชิ้นนี้ทำขึ้นเพื่อศึกษาเทคโนโลยี Eye Tracking ซึ่งเป็นเทคโนโลยีใหม่นำมาใช้เพื่อเพิ่มประสิทธิภาพในการทำนายอารมณ์ของรูปภาพให้มีความแม่นยำมากขึ้นและสามารถนำงานวิจัยชิ้นนี้ไปเป็นแนวคิดในการแสดงรูปภาพให้กับผู้ใช้ (Image Retrieval) แบบ Emotional Semantic Image Retrieval ได้

1.3. สมมติฐานของการศึกษา

การใช้ลักษณะเด่นของรูปภาพร่วมกับเครื่องมือติดตามการมองเห็นที่เก็บข้อมูลการป้อนกลับโดยปริยาย (Implicit Feedback) ใช้ในการแบ่งอารมณ์รูปภาพได้ดี สามารถเพิ่มประสิทธิภาพการทำนายอารมณ์ของรูปภาพได้มากขึ้น เนื่องจากพฤติกรรมการมองแต่ละคนต่างกัน ทำให้มีอารมณ์ต่อรูปภาพต่างกัน ดังนั้นหากรู้ลักษณะของรูปภาพที่ทำให้เกิดอารมณ์จะมีการทำนายอารมณ์ได้เพิ่มมากขึ้น

1.4. ขอบเขตการศึกษา

- ใช้เครื่องมือติดตามการมองเห็น และคุณลักษณะภาพในการทำนายอารมณ์ของรูปภาพ
- ศึกษาว่า Eye Tracking สามารถเพิ่มประสิทธิภาพการทำนายอารมณ์ได้แม่นยำหรือไม่
- รูปภาพที่ใช้ในการเก็บข้อมูล คือ รูปภาพนามธรรม
- แบ่งลักษณะอารมณ์เป็น 4 อารมณ์คือ มีความสุข (Happiness), เสียใจ (Sadness), โกรธ (Anger) และกลัว (Fear)
- คุณลักษณะรูปภาพที่ใช้มี สี (color), รูปร่าง (shape) และพื้นผิว (Texture)

1.5. ขั้นตอนการปฏิบัติโครงการ

1. ศึกษาหลักการทำงานและเทคโนโลยี Eye Tracking ที่สังเกตการเคลื่อนไหวของดวงตา
2. พัฒนาโปรแกรมที่ใช้ร่วมกับ Eye Tracking เก็บข้อมูลการเคลื่อนไหวของดวงตาเพื่อนำไปวิเคราะห์ลักษณะเด่นของรูปภาพและอารมณ์
3. วิเคราะห์ข้อมูลหาลักษณะเด่นและอารมณ์ของรูปภาพ
4. สรุปผลการวิจัยและนำเสนอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6. ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นข้อมูลในการวิจัยเกี่ยวกับระบบการค้นคืนรูปภาพ (Image Retrieval) แบบ Emotional Sematic Image Retrieval (ESIR) ได้
2. นำข้อมูลไปพัฒนาระบบแบบ Emotional Sematic Image Retrieval (ESIR) ได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. ทฤษฎีที่เกี่ยวข้อง

2.1.1 Relevance Feedback

เป็นการให้ข้อมูลป้อนกลับสู่ระบบค้นคืน เพื่อประเมินว่าข้อมูลที่ถูกรับคืนจากระบบนั้นมีความเกี่ยวข้องกับสิ่งที่ต้องการหรือไม่ นำไปปรับเปลี่ยนเพื่อเพิ่มประสิทธิภาพของการค้นคืนข้อมูลที่ต้องการในครั้งต่อไป สามารถแบ่งออกเป็น 2 ประเภท คือ [10][11][12]

1. การป้อนกลับโดยชัดเจน เป็นการป้อนกลับของผู้ใช้ระบบที่เกิดขึ้นโดยความตั้งใจ ซึ่งจำเป็นต้องใช้เวลาและแรงงาน เช่น การคลิกเมาส์การเคลื่อนไหวของตาโดยความตั้งใจ เสียง และ ท่าทาง
2. การป้อนกลับโดยปริยาย เป็นการป้อนกลับของผู้ใช้ระบบที่เกิดจากความไม่ตั้งใจ ซึ่งเกิดโดยอัตโนมัติขณะที่มนุษย์กำลังทำงานใด ๆ อยู่ การป้อนกลับชนิดนี้จำเป็นต้องถูกประมวลผลก่อนนำไปใช้ เช่น อัตราการเต้นของหัวใจ ความดันโลหิต อุณหภูมิของร่างกาย และ การเคลื่อนไหวของตาโดยความไม่ตั้งใจ (ขณะที่กำลังอ่านหนังสือ หรือ หาข้อมูลบนเว็บเบราว์เซอร์)

2.1.2 Emotional Foundation

อารมณ์คือความรู้สึกที่ได้รับผลกระทบจากสิ่งเร้ารอบ ๆ ตัว ซึ่งการแบ่งอารมณ์พื้นฐานของมนุษย์สามารถอธิบายได้ 2 แนวทาง คือ

1. Discrete Approach – อารมณ์แต่ละอารมณ์นั้นแตกต่างกันอย่างสิ้นเชิง โดยแต่ละอารมณ์จะมีระดับความเข้มของตัวเอง ทฤษฎีต่าง ๆ ที่อธิบายอารมณ์โดยใช้แนวทางนี้ เช่น [11,12] ในงานวิจัยนี้จะแบ่งอารมณ์โดยใช้แนวทางของ Ekman (1972) ซึ่งแบ่งอารมณ์ของมนุษย์ออกเป็น 6 อารมณ์ นั่นคือ โกรธ (Anger), ขยะแขยง (Disgust), กลัว (Fear), ดีใจ (Happiness), เสียใจ (Sadness) และประหลาดใจ (Surprise) [13] อารมณ์ดังกล่าวได้ถูกจัดกลุ่มใหม่เป็น 4 อารมณ์ คือ ดีใจ, เสียใจ, โกรธและ กลัว โดยแบ่งตามกล้ามเนื้อการแสดงสีหน้าของแต่ละอารมณ์ โดยพบว่ากล้ามเนื้อของการแสดงสีหน้าของมนุษย์จากอารมณ์ประหลาดใจมีลักษณะคล้ายกับกลัว เช่นเดียวกับอารมณ์ขยะแขยงและ โกรธ [14]
2. Dimensional Approach – อารมณ์ต่าง ๆ เกิดขึ้นได้จากการผสมกันระหว่างอารมณ์พื้นฐาน ซึ่งอารมณ์พื้นฐานแต่ละอารมณ์จะมีระดับความเข้มของอารมณ์ด้วย ตัวอย่างเช่น Plutchik (1980) ได้แบ่งอารมณ์พื้นฐานออกเป็น 8 อารมณ์ คือ สุข (Joy), เศร้า (Sadness), โกรธ (Anger), กลัว (Fear), ไว้วางใจ (Trust), รังเกียจ (Disgust), ประหลาดใจ (Surprise) และ คาดหวัง

(Anticipation) [15] ซึ่งอารมณ์พื้นฐานเหล่านี้สามารถผสมเป็นอารมณ์ใหม่ได้ เช่น สุข-ไว้ว่างใจ จะได้ อารมณ์รัก (Love)

แต่จากงานวิจัยของ Rachael E. Jack,¹ Oliver G.B. Garrod,¹ and Philippe G. Schyns [23] ที่ศึกษาเกี่ยวกับลักษณะหน้าแต่ละอารมณ์พบว่า การแสดงออกความรู้สึกทางอารมณ์ผ่านสีหน้าของคนแต่ละทวีปต่างกัน และลักษณะหน้าของอารมณ์ระหว่างกลัวและเซอร์ไพรส์ กับอารมณ์โกรธและขยะแขยง มีความคล้ายคลึงกันมาก งานวิจัยชิ้นนี้จึงสรุปว่าอารมณ์สามารถแบ่งออกเป็น 4 อารมณ์ คือ ความสุข, เศร้า, กลัวหรือเซอร์ไพรส์, โกรธหรือขยะแขยง

2.1.3 Gaussian function

ใช้คำนวณความน่าจะเป็นการกระจายของตำแหน่งในรูปภาพ เพื่อสร้างชุดข้อมูล โดยใช้ เกล็ดเขียน 3 ขนาด คือ 100x100, 125x125 และ 150x150 นำไปคำนวณกับตำแหน่งของรูปภาพที่ผู้ร่วมเก็บข้อมูลมอง ใช้สร้างชุดข้อมูล Eye movements Image และ Eye movements with Gaussian Function Image จากสมการที่ (2.1)

$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \quad (2.1)$$

โดยที่ σ คือ ส่วนเบี่ยงเบนมาตรฐาน; x คือ ตำแหน่งของรูปภาพในแกนอน; $g(x)$ คือ ค่าตำแหน่งของรูปภาพที่ได้จากการคำนวณ

2.1.4 รูปภาพนามธรรม

รูปภาพที่แสดงถึงอารมณ์หรือความรู้สึกของผู้วาด Wassily Kandinsky ผู้คิดศิลปะนามธรรม กล่าวว่า สิ่งที่สำคัญสำหรับศิลปะนามธรรม คือ สีและรูปทรงที่ใช้ในการถ่ายทอดโดยคำนึงถึงความรู้สึกภายนอกและภายใน กล่าวคือ การถ่ายทอดรูปทรงต่าง ๆ ให้กลมกลืนด้วยสี ลักษณะผิว หรือ ความเด่นชัดของรูปภาพ เป็นต้น เพื่อให้เกิดอารมณ์และความรู้สึก [16] ลักษณะของรูปนามธรรมจะไม่สื่อความหมายโดยตรงเป็นรูป ผู้ดูต้องใช้ความรู้สึกและจินตนาการ อารมณ์ที่เกิดจากการกระตุ้นจากรูปนามธรรมนั้นเกิดได้จากการตีความหมายของแต่ละผู้มอง ซึ่งขึ้นอยู่กับพื้นที่ของรูปภาพที่ผู้มองต้องการจะมองด้วย ดังนั้นรูปภาพรูปเดียวกัน ผู้มองคนละคน อาจมีอารมณ์ที่ต่างกัน ดังนั้นอุปกรณ์การติดตามการมองจึงถูกนำมาใช้เพื่อเก็บข้อมูลพฤติกรรมการมองของผู้ร่วมเก็บข้อมูล และนำมาวิเคราะห์

2.1.5 ซัพพอร์ตเวกเตอร์แมตชีน

เป็นอัลกอริทึมการเรียนรู้ของเครื่องชนิดหนึ่งที่ได้รับการนิยามอย่างแพร่หลายในการจำแนกประเภทข้อมูลและถูกนำมาใช้ในการศึกษาอารมณ์ของรูปภาพ [5, 6, 7, 17] SVM จะแบ่งแยกข้อมูลโดยใช้ระนาบหรือไฮเปอร์เพลน (Hyperplane) แบ่งข้อมูลแบบเชิงเส้น

กำหนดให้ x_i คือเวกเตอร์ตัวอย่างที่ i ที่คู่กับค่าเป้าหมาย y_i ซึ่ง $y_i \in \{-1, 1\}$, w คือเวกเตอร์ค่าน้ำหนัก, b คือค่าความโน้มเอียง (Bias), ξ_i คือค่าความผิดพลาดของการจำแนกประเภท, และ C คือตัวแปรเรกิวเลอเรชัน โดยระนาบนั้นจะถูกสร้างขึ้นโดยการจะพยายามทำให้ระยะห่างระหว่างสองกลุ่มนั้นมากที่สุด และ ความผิดพลาดในการแบ่งกลุ่มน้อยที่สุดดังสมการที่ (2.2)

$$\arg \min_{w, b, \xi} \frac{1}{2} w^2 + C \sum_i \xi_i \quad (2.2)$$

ซึ่งอยู่ภายใต้ $y_i(w \cdot x_i + b) \geq 1 - \xi_i$ และ $\xi_i \geq 0$ โดยที่มีฟังก์ชันในการตัดสินใจคือ

$$f(x) = \text{sign}(w^T x + b) \quad (2.3)$$

ในโลกแห่งความเป็นจริงข้อมูลไม่ได้เป็นเชิงเส้นเสมอไป จึงมีการใช้เทคนิคที่เรียกว่า “เคอร์เนล” (Kernel) โดยใช้ฟังก์ชันเพื่อแปลงจากปริภูมิคุณลักษณะ (Feature Space) ปัจจุบันไปยังปริภูมิคุณลักษณะใหม่ที่สามารถแบ่งข้อมูลด้วยระนาบได้ เช่น Polynomial, Radial Basis Function เป็นต้น

2.2. งานวิจัยที่เกี่ยวข้อง

ในปัจจุบันได้มีงานวิจัยที่เกี่ยวกับระบบการค้นคืนรูป (Image Retrieval) มากมาย เช่น การค้นคืนรูปภาพ Content- Based Image Retrieval (CBIR) โดยใช้ลักษณะรูปภาพ (Image Feature) ในการค้นหาและแสดงรูปภาพ เพื่อดึงรูปภาพให้ตรงกับความต้องการผู้ใช่มากที่สุด นอกจากนี้ยังมีแนวคิดระบบการค้นคืนรูปภาพแบบ Emotional Sematic Image Retrieval (ESIR) เป็นงานวิจัยที่หาอารมณ์รูปภาพจากลักษณะของรูปและยังมีการวิจัยเกี่ยวกับ Eye Tracking ที่นำมาใช้ในงานวิจัย

ในปี 2009 งานวิจัย [2] ศึกษาการพัฒนาการค้นหารูปภาพนำโดยใช้คุณลักษณะรูปภาพและ Eye movements ของผู้ร่วมเก็บข้อมูล จากอุปกรณ์ Eye Tracker ที่สามารถเก็บข้อมูลการป้อนกลับ

โดยปริยาย (Implicit Feedback) ซึ่งเป็นข้อมูลที่ได้จากพฤติกรรมของผู้ใช้ เมื่อดูรูปภาพแต่ละรูปนำมาวิเคราะห์ร่วมกันเพื่อช่วยในการจัดลำดับการแสดงผลรูปภาพให้มีประสิทธิภาพมากขึ้น

ในปี 2010 งานวิจัย [1] ได้ศึกษาเกี่ยวกับ การป้อนกลับโดยปริยาย (Implicit Feedback) ที่ได้จากการเคลื่อนไหวของดวงตาในขณะระหว่างที่ผู้ใช้งานกำลังค้นหารูปภาพในระบบการค้นหาแบบ Content-Based Image Retrieval นำข้อมูลดังกล่าวมาวิเคราะห์ความสัมพันธ์ระหว่างคุณลักษณะรูปภาพและ Eye movements ที่ผู้ชมมองภาพที่สนใจ เพื่อพัฒนาค้นคืนภาพที่ผู้ใช้งานต้องการให้มีประสิทธิภาพมากขึ้น

ในปี 2010 งานวิจัยของ [5] ศึกษาการพัฒนาวิธีการสกัดคุณลักษณะรูปภาพเพื่อหาอารมณ์รูปภาพและจัดกลุ่มอารมณ์รูปภาพ โดยใช้คุณลักษณะพื้นฐานของรูปภาพ เช่น สี (Color), พื้นผิว (Texture), ตำแหน่งองค์ประกอบ (Composition), เนื้อหา (Content) สามารถนำไปใช้ในการค้นคืนรูปภาพ การทดลองมีรูปภาพทั้งหมด 3 แบบ คือ รูปภาพ, รูปถ่ายนามธรรม และ International Affective Picture System (IAPS) และแบ่งอารมณ์เป็น 8 อารมณ์ คือ Amusement, Anger, Awe, Contentment, Disgust, Excitement, Fear และ Sad จากผลการทดลองพบว่าวิธีการสกัดคุณลักษณะรูปภาพมีประสิทธิภาพต่างกันขึ้นอยู่กับประเภทอารมณ์และชุดข้อมูล

ในปี 2011 งานวิจัย [4] ศึกษาอารมณ์ของคนที่มีต่อรูปภาพ Abstract Art จากลักษณะของรูปภาพแบบพื้นฐาน โดยในการสกัดคุณลักษณะรูปภาพ ใช้ Compound Transform ที่มีการทดลองว่ามีประสิทธิภาพในการหาคุณลักษณะมากกว่า และใช้ Fisher score สำหรับการจัดลำดับคุณลักษณะที่มีผลต่ออารมณ์แต่ละอารมณ์ จากผลการทดลองพบว่าลักษณะรูปภาพพื้นฐาน สามารถอธิบายอารมณ์ของคนได้และใช้ในการจัดหมวดหมู่อารมณ์รูปภาพได้ดี

ในปี 2013 งานวิจัย [3] ศึกษาการจำแนกอารมณ์ของรูปถ่ายนามธรรมที่สามารถนำไปใช้ในการค้นคืนรูป (Image Retrieve) โดยใช้ Multiple Kernel Learning (MKL) ทำนายอารมณ์ของรูปภาพที่มีประสิทธิภาพมากกว่าและยังมีการเลือกคุณลักษณะที่สำคัญให้อัตโนมัติ [8] ลักษณะของงานวิจัยมีการเก็บอารมณ์แต่ละรูปภาพจากคนทั่วไปผ่านเว็บไซต์ และสกัดคุณลักษณะโดยใช้ Picosom System เป็นการหาคุณลักษณะพื้นฐานรูปภาพจะมีการหาคุณลักษณะสี (color), รูปร่าง (shape), พื้นผิว (texture) การหาคุณลักษณะรูปภาพมีทั้งแบบ global image และ local image ที่แบ่งรูปภาพออกเป็น 5 โซน จากนั้นหาคุณลักษณะรูปหาในแต่ละโซน ซึ่งผลการทดลองเมื่อเทียบประสิทธิภาพการทำนายพบว่า MKL มีผลการทำนายดีกว่ามากกว่าแบบ single kernel learning นอกจากนี้พบว่าคุณลักษณะของสีและรูปร่างเป็นข้อมูลที่มีคุณสมบัติต่อการรับรู้อารมณ์มาก

งานวิจัยชิ้นนี้เราจะเสนอการหาอารมณ์รูปจากลักษณะรูปภาพโดยใช้รูปถ่ายนามธรรมหาคุณลักษณะพื้นฐานของรูป เช่น สี (Color), รูปร่าง (Shape) และ พื้นผิว (Texture) ซึ่งเป็นวิธีเดียวกับระบบค้นคืนภาพแบบ CBIR แต่งานวิจัยชิ้นนี้มีอุปกรณ์ Eye Tracking ที่จับดวงตาและจัดเก็บข้อมูล

การมองตำแหน่งของรูปภาพนำไปทำข้อมูลที่มีการมองซึ่งเป็นข้อมูลการป้อนกลับโดยปริยายการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Implicit Feedback) ที่ได้จากผู้ร่วมเก็บข้อมูล มาวิเคราะห์ข้อมูล ได้ถูกต้องแม่นยำและมีประสิทธิภาพ สามารถนำไปใช้ในการแสดงข้อมูลรูปภาพที่มีลักษณะอารมณ์ความคล้ายคลึงกันในระบบการค้นคืนรูปภาพ (Image Retrieval) แบบ ESIR และแบ่งอารมณ์เป็น 4 อารมณ์ คือ ความสุข (Happiness), เสียใจ (Sadness), โกรธ (Anger) และกลัว (Fear)

2.3. เครื่องมือในการดำเนินงาน

Software

1. Visual Studio 2013 โปรแกรมใช้พัฒนาโปรแกรมเก็บข้อมูลจากผู้ร่วมเก็บข้อมูล
2. Matlab R2014a โปรแกรมใช้วิเคราะห์ข้อมูลลักษณะรูปภาพ

Hardware

1. เครื่องมือติดตามการมองเห็น อุปกรณ์ติดตามและเก็บข้อมูลดวงตาของผู้ร่วมเก็บข้อมูล

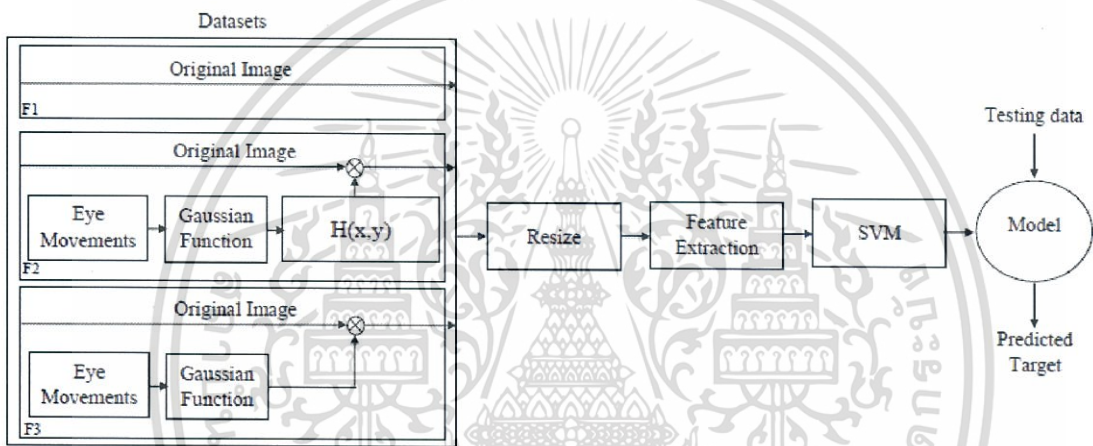
2.4. เทคโนโลยีที่เกี่ยวข้อง

1. C# ภาษาที่ใช้ในการพัฒนาโปรแกรมเก็บข้อมูลจาก User

บทที่ 3

การดำเนินงานวิจัย

ขั้นตอนงานวิจัยเพื่อเทียบประสิทธิภาพการทำนายอารมณ์คนที่ต่อรูปภาพโดยใช้อุปกรณ์ Eye Tracking เริ่มจากการหารูปภาพเพื่อนำมาใช้ในงานวิจัยและเก็บข้อมูลอารมณ์รูปภาพจากผู้ร่วมเก็บข้อมูล จากนั้นนำเตรียมชุดข้อมูลรูปภาพเพื่อสกัดคุณลักษณะเด่นของรูป นำข้อมูลภาพไปฝึกฝนและแบ่งประเภทข้อมูลอารมณ์รูปภาพ จากนั้นทดสอบ โมเดลอารมณ์รูปภาพ เพื่อเทียบประสิทธิภาพการทำนายอารมณ์แต่ละชุดข้อมูล



รูปที่ 3.1 แสดงขั้นตอนการดำเนินงานวิจัย

3.1. รูปภาพ



รูปที่ 3.2 ตัวอย่างรูปภาพนามธรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับศึกษาเชิงวิชาการเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิจัยชิ้นนี้ใช้รูปนามธรรมจาก [4] ซึ่งประกอบไปด้วยรูป 228 รูป แบ่งออกเป็น 8 อารมณ์ คือความสนุก (Amusement), ตื่นเต้น (Excitement), พอใจ (Contentment), เสียใจ (Sad), โกรธ (Anger), น่าเกรงขาม (Awe), กลัว (Fear), และขยะแขยง (Disgust) เราจึงจัดกลุ่มอารมณ์รูปภาพที่มีอารมณ์ใกล้เคียงกันให้เหลือเพียง 4 อารมณ์ เพื่อลดความซับซ้อนให้กับผู้ร่วมเก็บข้อมูลในการตัดสินใจ โดยอ้างอิงจาก Parrot (2001) ซึ่งแบ่งอารมณ์อย่างมีโครงสร้างเป็น 3 ลำดับ คือ Primary Emotions, Secondary Emotions และ Tertiary Emotions [19] โดยรวมรูปภาพที่มีอารมณ์คล้ายคลึงกัน ที่อยู่ในชั้น Primary Emotions เดียวกัน ให้เป็นอารมณ์พื้นฐาน 4 อารมณ์ เช่นเดียวกับ [14] คือ (1) ดีใจ – {สนุก, ตื่นเต้น, พอใจ}, (2) เสียใจ (3) โกรธ (4) กลัว – {กลัว, ขยะแขยง} เนื่องจากรูปที่มีอารมณ์น่าเกรงขามไม่ได้ถูกกล่าวไว้ใน [19] เราจึงไม่ได้สนใจรูปในกลุ่มนี้

จากนั้นจึงเลือกรูปภาพจำนวน 100 รูป โดยเลือกรูปที่มีอารมณ์ทั้ง 4 อารมณ์ อารมณ์ละ 25 รูป วิธีการเลือกรูปภาพนั้น เลือกจากคะแนนสูงสุดของแต่ละอารมณ์จากการโหวตของผู้ใช้งานงานวิจัยชิ้นนั้น

3.2. การเก็บอารมณ์รูปภาพ



รูปที่ 3.3 การเก็บการเคลื่อนไหวของตาโดยเครื่องติดตามการมอง และอารมณ์ของผู้ร่วมการทดลองที่มีต่อรูปภาพ

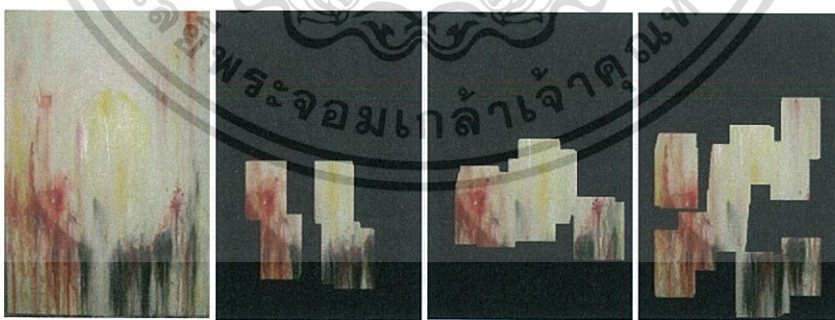
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการทดลอง ได้มีการเก็บข้อมูลอารมณ์ของมนุษย์ที่มีต่อรูปภาพ และการเคลื่อนไหวของตา เพื่อใช้ในการวิเคราะห์ จากนักศึกษาระดับปริญญาตรี ซึ่งมีอายุ 18-22 ปี จำนวนทั้งหมด 20 คน แบ่งเป็น ชาย 10 คน หญิง 10 คน คนละ 100 รูป ผู้ร่วมการทดลองมีสายตาที่ปกติ กล่าวคือไม่ได้ใส่แว่นตา

การเก็บข้อมูลจะมีโปรแกรมที่พัฒนาขึ้นเพื่อแสดงรูปภาพจากฐานข้อมูล โดยทำงานร่วมกับอุปกรณ์ติดตามการมองเห็น โดยจะให้ผู้ร่วมเก็บข้อมูลนั่งประจำเครื่องคอมพิวเตอร์ขนาดพกพา 13 นิ้ว โดยการแสดงผลมีความละเอียดเท่ากับ 1366 x 768 พิกเซล และผู้ควบคุมการเก็บข้อมูลจะอธิบายขั้นตอนและวิธีการเก็บข้อมูลให้กับผู้ร่วม โดยผู้เข้าร่วมเก็บข้อมูลจะบอกอารมณ์ที่เกิดขึ้นจากการมองแต่ละรูปภาพที่แสดงบนหน้าจอ และผู้ควบคุมการทดลองจะบันทึกข้อมูลที่ได้โดยใช้คีย์บอร์ด ดังรูปที่ 3.3

ก่อนเริ่มการเก็บข้อมูล มีการแจ้งให้ผู้ร่วมการทดลองพยายามให้ศีรษะอยู่นิ่ง ๆ และจะต้องผ่านการปรับเทียบ (Calibration) ของอุปกรณ์การติดตามการมอง โดยใช้ 9 จุด (กลางหน้าจอ 1 จุด และ 8 จุดที่ขอบหรือมุมหน้าจอ) ดังรูปที่ 2(ก) และก่อนที่จะให้ผู้ร่วมการทดลองดูรูปภาพใหม่ จะมีการแสดงรูปที่มีจุดอ้างอิง (Reference) ตรงกลางหน้าจอเป็นเวลา 3 วินาทีทุกครั้ง ดังรูปที่ 2(ข) เพื่อให้การเคลื่อนไหวของตาเริ่มต้นที่จุด ๆ เดียวกันทุกครั้ง

จากการเก็บข้อมูลนั้นพบว่ารูปเดียวกัน ดังรูปที่ 3.4(ก) คนมองต่างกัน อารมณ์ที่ได้ต่างกัน เช่นกัน จากการมองรูปภาพของ ผู้เข้าร่วม-1 ให้ผลอารมณ์รูปภาพ คือ กลัว ดังรูปที่ 3.4(ข) ผู้เข้าร่วม-2 ให้ผลอารมณ์รูปภาพ คือ โกรธ ดังรูปที่ 4(ค) และ ผู้เข้าร่วม-3 ให้ผลอารมณ์รูปภาพ คือ เสียใจ ดังรูปที่ 3.4(ง)



(ก) รูปต้นฉบับ (ข) กลัว (ค) โกรธ (ง) เสียใจ

รูปที่ 3.4 ตัวอย่างรูปภาพ และส่วนที่ถูกมองของผู้เข้าร่วมที่มีอารมณ์แตกต่างกัน

3.3. เครื่องมือติดตามการมองเห็น

เทคโนโลยี Eye Tracking รองรับการใช้งานบนคอมพิวเตอร์ระบบปฏิบัติการ Windows (7 และ 8), OS X และสมาร์ทโฟนหรือแท็บเล็ต ที่ใช้งานบนระบบปฏิบัติการ Android จับการเคลื่อนไหวของดวงตาด้วย USB 3.0 สามารถใช้ในการควบคุมการทำงานได้เหมือนเมาส์ (Mouse) มีอัตราการเก็บข้อมูลเท่ากับ 30 Hz และมีความแม่นยำที่ 0.5° - 1° ใช้ JSON Formatted ในการแลกเปลี่ยนข้อมูล เราสามารถจะขอข้อมูลจากอุปกรณ์ Eye Tracker โดยส่งความต้องการในการดึงข้อมูล

ปัจจุบันมีการนำ Eye Tracker มาทดลองใช้เล่นเกม Fruit Ninja ในแท็บเล็ต ซึ่งจะประเมินและประมวลผลการมองของ User ในขณะที่กำลังมองหน้าจอ เพื่อส่งการทำงานต่อไป นอกจากนี้ อุปกรณ์ Eye Tracking ยังสามารถบันทึกตำแหน่งการมองหน้าจอของผู้ใช้ได้ จึงมีการนำอุปกรณ์ดังกล่าวไปใช้ในการทำวิจัย เช่น ใช้ประโยชน์ในการทำโฆษณาบนอินเทอร์เน็ตหรือการจัดเนื้อหาบนเว็บไซต์ แต่ในงานวิจัยนี้เรานำอุปกรณ์ Eye Tracking เข้ามาช่วยในการจัดเก็บพฤติกรรมการมองรูปภาพของผู้ใช้ และนำข้อมูลดังกล่าวไปแสดงเป็นข้อมูลในรูปแบบแผนที่ความร้อนเพื่อเอาข้อมูลไปวิเคราะห์ช่วยเพิ่มประสิทธิภาพในการทำงานอารมณ์รูปภาพแม่นยำมากยิ่งขึ้น

3.4. โปรแกรมเก็บข้อมูล

พัฒนาโปรแกรมเก็บข้อมูลอารมณ์ให้ทำงานร่วมกับอุปกรณ์ Eye Tracker สำหรับจัดเก็บข้อมูลตำแหน่งของรูปภาพที่ผู้ร่วมเก็บข้อมูลมองและนำข้อมูลดังกล่าวให้เป็นชุดข้อมูล เพื่อนำข้อมูลภาพเฉพาะส่วนไปสกัดคุณลักษณะของรูปภาพ ส่วนภาษาที่ใช้ในการพัฒนาโปรแกรมดังกล่าวคือ ภาษา C#

1. กรอกรายละเอียดของผู้เข้าร่วมเก็บข้อมูล

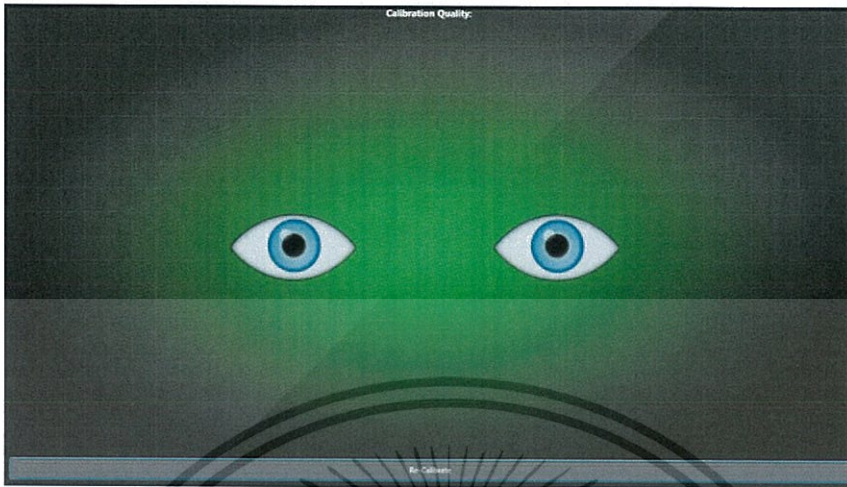
The image shows a screenshot of a software application window titled "UI Project". The window contains a registration form with the following elements:

- A text input field labeled "Your name :".
- A "Gender :" section with two radio button options: "Male" and "Female".
- An "Age :" section with a dropdown arrow.
- A "Select Type :" section with two radio button options: "Abstract" and "Photograph".
- An "OK" button at the bottom center.

รูปที่ 3.5 หน้าจอกรอกข้อมูลผู้ร่วมเก็บข้อมูลอารมณ์รูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การเขียนในเชิงการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การตรวจสอบประสิทธิภาพการทำงานของEye Tracking กับผู้ใช้



รูปที่ 3.6 หน้าจอตรวจสอบตำแหน่งก่อนการปรับเทียบ (Calibration) ของอุปกรณ์การติดตามการมอง

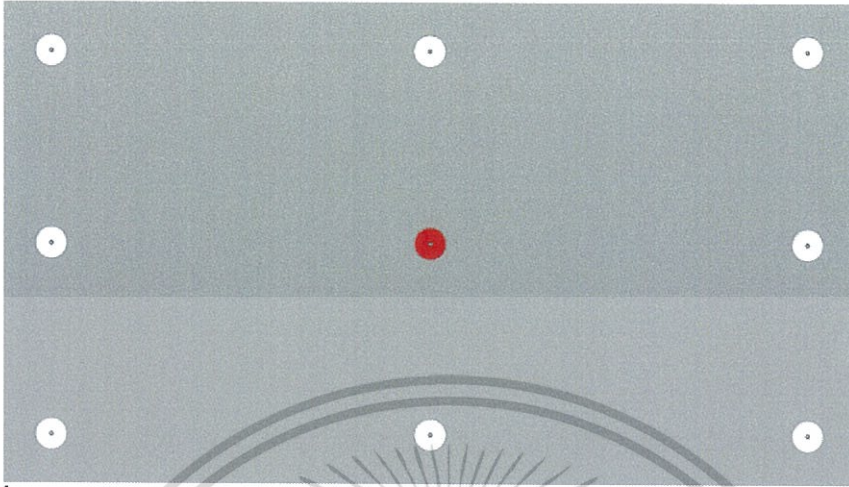
3. ทดสอบการทำงานของอุปกรณ์ Eye Tracking โดยให้ผู้เข้าร่วมเก็บข้อมูลมองตามจุดสีขาว



รูปที่ 3.7 ทดสอบการปรับเทียบ (Calibration) ของอุปกรณ์การติดตามการมอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ผู้ร่วมเก็บข้อมูลมองจุดตามตำแหน่งต่างๆ เพื่อดูว่าเซนเซอร์สามารถจับดวงตาได้หรือไม่



รูปที่ 3.8 หน้าจอตรวจสอบประสิทธิภาพการมองแต่ละจุดของอุปกรณ์ Eye Tracking

5. เริ่มการเก็บข้อมูลโดยให้ผู้ร่วมเก็บข้อมูลมองจุดโฟกัสที่จุดเดียวกัน เพื่อให้การเก็บข้อมูลจาก Eye Tracker มีประสิทธิภาพ



รูปที่ 3.9 จุดโฟกัสให้ผู้เข้าร่วมเก็บข้อมูลมองก่อนเริ่มดูรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. หน้าจอแสดงรูปภาพของโปรแกรมเก็บข้อมูล



รูปที่ 3.10 แสดงรูปภาพให้ผู้เข้าร่วมเก็บข้อมูล

3.5. กระบวนการเตรียมข้อมูล (Preprocessing)

ชุดข้อมูลรูปภาพในงานนี้มีทั้งหมด 3 แบบ คือ Original Image ส่วน Eye Movements Image และ Eye Movements with Gaussian function Image ขึ้นอยู่กับการมองแต่ละผู้ร่วมเก็บข้อมูล

1. Original Image

เป็นรูปภาพนามธรรมที่นำมาจากงานวิจัย [5] ดังรูปที่ 3.11



รูปที่ 3.11 (F1) Original Image

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Eye Movements Image

ภาพในส่วนนี้จะขึ้นอยู่กับตำแหน่งการมองของแต่ละผู้ร่วมเก็บข้อมูล โดยใช้ฟังก์ชันความน่าจะเป็นการกระจายของตัวแปรสุ่มเกาส์เซียนที่มีขนาดต่าง ๆ คือ 100×100 , 125×125 และ 150×150 กำหนดให้ $I(x, y)$ คือ รูปภาพ โดยที่มี (x, y) เป็นพิกัดของรูปภาพ และ $E(x, y)$ คือ ข้อมูลการมองของรูปภาพ จากนั้นจึงใช้ฟังก์ชันเกาส์เซียนกับข้อมูลการมองของรูปภาพจาก

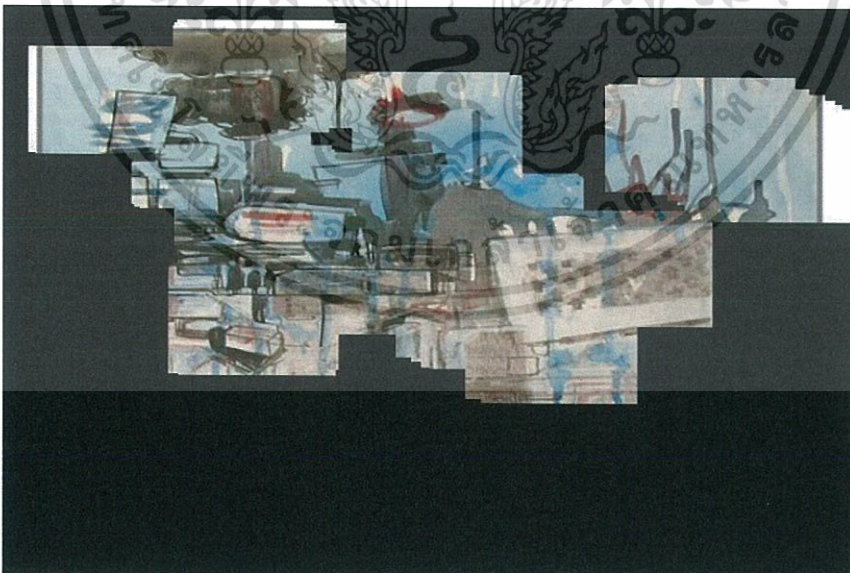
$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \quad (3.1)$$

โดยที่ σ คือ ส่วนเบี่ยงเบนมาตรฐานของการกระจายเกาส์เซียน จะได้ข้อมูลการมองของรูปภาพที่มีการกระจายตัวแบบเกาส์เซียน $G(x, y)$ จากนั้นจึงนำไปแทนในฟังก์ชัน (5) เพื่อสร้างตัวกรองสำหรับตำแหน่งรูปภาพที่ถูกผู้ใช้งาน

$$H(x, y) = \begin{cases} 1 & \text{if } G(x, y) > 0 \\ 0 & \text{if } G(x, y) = 0 \end{cases} \quad (3.2)$$

ดังนั้นรูปที่ผ่านตัวกรองคือ

$$I_{F2}(x, y) = H(x, y) \times I(x, y) \quad (3.3)$$



รูปที่ 3.12 (F2) Eye movements Image

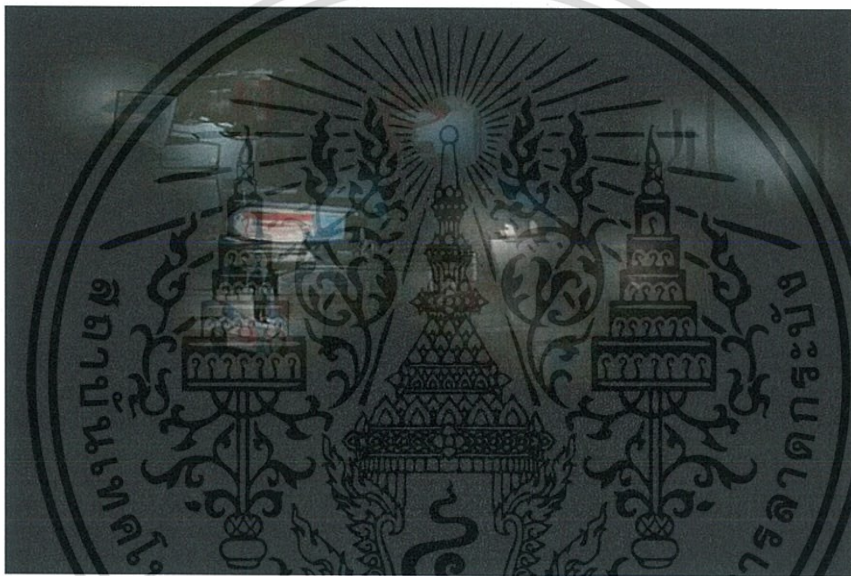
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. Eye Movements with Gaussian function Image

ในส่วนนี้จะมีวิธีการสร้างเช่นเดียวกับรูปภาพในหัวข้อ 2. โดยใช้ $G(x, y)$ เป็นตัวกรองในการสร้างรูปใหม่

$$I_{F3}(x, y) = G(x, y) \times I(x, y) \tag{3.4}$$

$G(x, y)$ นั้นจะมีค่าตั้งแต่ 0 ถึง 1 หากค่าของตัวกรองตำแหน่งนั้นเข้าใกล้ศูนย์ ตำแหน่งภาพนั้นจะมีมืด แต่ถ้าค่าของตัวกรองตำแหน่งนั้นเข้าใกล้ 1 ตำแหน่งรูปภาพนั้นจะสว่าง



รูปที่ 3.13 (F3) Eye Movements with Gaussian Function Image

3.6. การสกัดคุณลักษณะเด่นรูปภาพ

การทดลองนี้เราใช้ฮิสโตแกรมในการอธิบายลักษณะรูปภาพ โดยมีลักษณะพื้นฐานของรูป 3 ประเภท คือ สี, รูปร่าง และ พื้นผิว ส่วนข้อมูลรูปภาพที่ใช้ในการสกัดคุณลักษณะเด่นรูปภาพมี 3 แบบ คือ Original Image (F1) ส่วน Eye Movements Image (F2) และ Eye Movements with Gaussian Function Image (F3) การสกัดคุณลักษณะของรูปภาพขึ้นอยู่กับมุมมองของแต่ละผู้ร่วมเก็บข้อมูล

1. Color Histogram ใช้ฮิสโตแกรมสีแบบรวมองค์ประกอบของทั้งสามสีเป็นฮิสโตแกรมเดียว ในการแจกแจงความถี่ของสี RGB space ที่มีอยู่ในรูปภาพ โดยทดลองใช้ขนาดของช่วงต่างกัน นั้นคือ {8, 16, 32, 64}-bin ในการทดลอง โดยที่ขนาดของคุณลักษณะที่ได้จะเท่ากับ

$$n_{\text{feature}} = n_{\text{bin}}^3 \tag{3.5}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Histogram of Sobel แสดงจำนวนของขนาดและปริมาณเส้นขอบที่มีอยู่ในรูปภาพ

Sobel Edge Detection เป็นวิธีการหาเส้นขอบของรูปภาพเป็นการหาคุนต์ลักษณะของเส้นขอบเพื่อให้ได้รูปร่าง (Shape) ในรูปภาพ โดยนำรูปภาพมา Convolution กับมาร์ค (Mask) ขนาด 3x3 พิกเซล 2 แบบ เพื่อหาเส้นขอบในแนวตั้งและแนวนอนของรูปภาพ จากนั้นจะหาขนาดเส้นขอบวัตถุในรูปภาพจาก จากสูตรสมการ

$$M = \sqrt{\Delta A^2 + \Delta B^2} \quad (3.6)$$

โดยที่ A คือ ขนาดของเส้นขอบแนวนอน

B คือ ขนาดของเส้นขอบแนวตั้ง

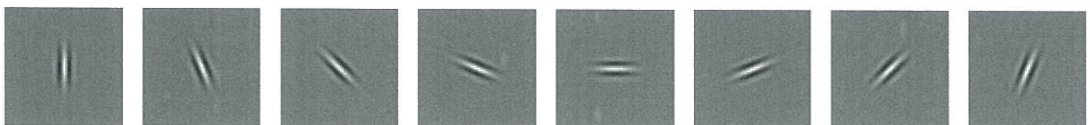
และการหาทิศทางของเส้นขอบรูปภาพสามารถหาจากสมการที่

$$\theta = \tan^{-1} \left[\frac{\Delta B}{\Delta A} \right] \quad (3.7)$$

เมื่อได้ค่าขนาดเส้นขอบ (Magnitude) และ ทิศทาง (gradient) จากนั้นจะรวมค่า Magnitude ที่มีทิศทางตามบินต่างๆ เพื่อนำไปสร้างฮิสโตแกรม โดยที่ขนาดของคุนต์ลักษณะที่ได้จะเท่ากับ

$$n_{\text{feature}} = n_{\text{bin}} \quad (3.8)$$

3. Histogram of Gabor แสดงจำนวนลักษณะทิศทางที่มีอยู่ในรูปภาพตามตัวกรองแต่ละทิศทาง ตัวกรองกาบองฟังก์ชัน เราใช้ตัวกรองความถี่สูงในแต่ละทิศทางเพื่อหาเส้นขอบในรูปภาพ ที่แสดงถึงพื้นผิวของรูปภาพ ขนาด 1 ทิศทาง 8 ดังรูปที่ 3.14



รูปที่ 3.14 ตัวกรองกาบองฟังก์ชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยหาคุณลักษณะดังกล่าวจากสมการ

$$G(x, y) = \frac{f^2}{\pi\gamma\eta} \exp(j2\pi fx' + \phi) \quad (3.9)$$

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta$$

โดยที่ f คือ องค์ประกอบความถี่ของคลื่นลูก \sin ; θ คือ ทิศทางของตัวกรองคาบของ ฟังก์ชัน; ϕ คือ ค่าความคลาดเคลื่อน; σ คือ ส่วนเบี่ยงเบนมาตรฐาน; γ คือ อัตราการ เบี่ยงเบนองศา

และหาขนาด (Magnitude) ของพื้นผิวดังกล่าวจากสมการ

$$M = \log(1 + \sqrt{im^2 + re^2}) \quad (3.10)$$

โดย $\log(x)$ เพื่อลดขนาดคุณลักษณะของภาพของฟังก์ชันให้มีความใกล้เคียงกัน; im คือ ค่าสัมประสิทธิ์จำนวนจินตภาพ; re คือ ค่าสัมประสิทธิ์จำนวนจริง

เมื่อได้ค่าขนาด (Magnitude) ของแต่ละทิศทางตามตัวกรองคาบของทั้ง 8 ทิศ จากนั้นจะรวม ค่าMagnitude ที่มีทิศทางตามบินซึ่งแต่ละบินก็คือทิศทางตัวกรองทิศต่างๆ เพื่อนำไปสร้างฮิสโตแกรม โดยที่ขนาดของคุณลักษณะที่ได้จะเท่ากับสมการเดียวกับ 3.8

เนื่องจากสกัดคุณลักษณะรูปภาพมีค่าต่างกันจึงมีทำการ นอมอลไรซ์เซชันคุณลักษณะ รูปภาพตามสูตรที่ 3.11 เพื่อให้ค่าของคุณลักษณะทั้งหมดมีค่าใกล้เคียงกัน ให้มีค่าเฉลี่ยเท่ากับศูนย์ และค่าเบี่ยงเบนมาตรฐานเท่ากับหนึ่ง

$$z = \frac{(x - \mu)}{\sigma} \quad (3.11)$$

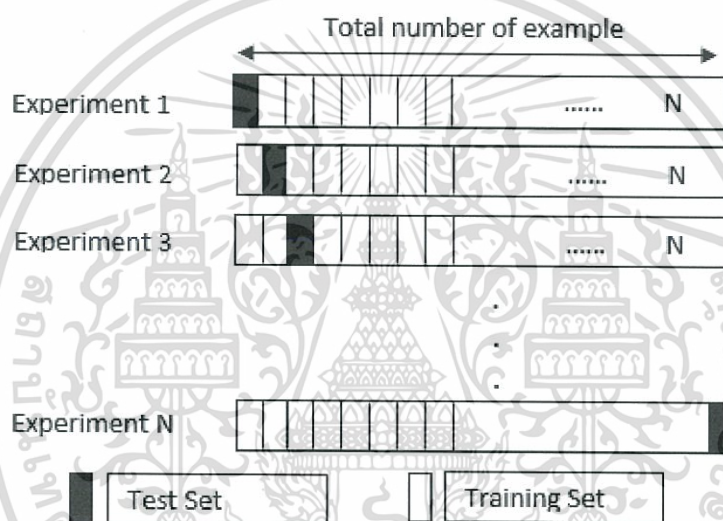
โดยที่ x คือ ค่าของข้อมูลทั้งหมด; μ คือ ค่าเฉลี่ยของข้อมูล σ ; คือ ส่วน เบี่ยงเบนมาตรฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.7. Leave-one-out cross validation

เทคนิคในการฝึกฝนข้อมูลมีหลายวิธี เช่น วิธีการเลือกสุ่มแบบ K-Fold Cross Validation หรือวิธีแยกทีละหนึ่ง (Leave-one-out cross validation; LOOCV) ในงานวิจัยนี้เราเลือกใช้แบบวิธีแยกทีละหนึ่ง (LOOCV)

โดยทั่วไปการฝึกฝนข้อมูลจะแบ่งชุดข้อมูลออกเป็น 2 แบบ คือ Training set ใช้ในการฝึกข้อมูลเพื่อสร้างโมเดลจำลองและ Test set ใช้ในการทดสอบการทำนายผลของโมเดล วิธีการแบบ LOOCV จะแบ่งข้อมูล Test set ออกมาทีละหนึ่ง เช่น มีข้อมูลทั้งหมด N รูป แบ่งเป็น Training data เท่ากับ $N-1$ และ Test set เท่ากับ 1 จะเทรนข้อมูลทั้งหมด N ครั้ง ดังรูปที่ 3.13



รูปที่ 3.15 การฝึกฝนข้อมูลแบบ Leave-One-Out CV

บทที่ 4

ผลการทดลอง

4.1. การดำเนินการทดลอง

ในงานวิจัยนี้มีการทดลองทั้งหมด 2 การทดลอง เพื่อเปรียบเทียบคุณลักษณะชุดข้อมูลทั้ง 3 ประเภท การทดลองแรกเป็นโมเดลการทำนายอารมณ์ของผู้ร่วมเก็บข้อมูลแต่ละคน จำนวน 20 คน โดยมีการทดลองเปลี่ยนค่าจำนวนช่วงของการแจกแจงความถี่แต่ละคุณลักษณะเป็น 8, 16, 32 และ 64 บิน และขนาดของเกาส์เซียนคือ 100x100, 125x125 และ 150x150 จากนั้นข้อมูลได้ถูกเรียนรู้และทดสอบโดยใช้ SVM พร้อมกับ Linear Kernel โดยมีพารามิเตอร์ที่ต้องปรับค่าเพียงตัวเดียวนั้นคือ C ในช่วงตั้งแต่ 10^{-6} – 10^4 เพื่อให้ได้ผลการทำนายที่ดีที่สุด มีการใช้เทคนิค LOOCV ในการแบ่งข้อมูลก่อนการทดลอง ส่วนการทดลองที่สองเป็น Global Model New User โดยมีเงื่อนไขการทดลองเช่นเดียวกับการทดลองแรก ต่างกันที่ใช้เทคนิค Leave-one user-out Cross-validation

สำหรับ Global Model New User ใช้คุณลักษณะที่มีค่าเฉลี่ยความถูกต้องมากที่สุด นั่นคือการรวมคุณลักษณะทั้ง 3 ลักษณะ ได้แก่ สี, รูปร่าง และพื้นผิว โดยการนำเวกเตอร์คุณลักษณะดังกล่าวมาต่อกัน

4.2. ผลการทดลอง

การทดลองที่ 1 โมเดลการทำนายอารมณ์ของผู้ร่วมเก็บข้อมูลแต่ละคน (User Model)

ประสิทธิภาพของการจำแนกอารมณ์ผู้รายงาน โดยใช้ค่าร้อยละความถูกต้องเฉลี่ยรวมของผู้ใช้ทั้ง 20 คน โดยเลือกจากพารามิเตอร์ C, ขนาดเกาส์เซียน และจำนวนบิน พิจารณาจากร้อยละความถูกต้องเฉลี่ยจากคุณลักษณะพื้นฐานของรูปภาพ ซึ่งตารางที่ 4.1 แสดงค่าเฉลี่ยรวมการทำนายแต่ละคุณลักษณะในแต่ละชุดข้อมูล นอกจากนี้ยังมีการทดสอบนัยสำคัญของผลการทดลองโดยใช้ Paired *t*-test สามารถดูค่าเฉลี่ยการทำนายแต่ละบินและผลการทดสอบสมมติฐานการวิจัยดังกล่าวเพิ่มเติมได้ในภาคผนวก ก.4-ก.9

เนื่องจากอารมณ์ทั้ง 4 อารมณ์ของผู้ร่วมการทดลองนั้นไม่เท่ากัน คุณลักษณะชุดข้อมูลทั้ง 3 ประเภท จึงถูกเปรียบเทียบกับ Baseline (BL) เป็นกรณีที่ใช้ SVM ทำการทำนายได้เป็นเพียงอารมณ์เดียว (Single label)

จากการทดลองการทำนายอารมณ์ของคุณลักษณะสีพบว่า F1, F2 และ F3 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.05$) ยกเว้น F1 ในกรณี 32-bin และ F2, F3 มีผลการทำนายดีกว่า F1 ในทุกกรณีเช่นเดียวกัน ยิ่งกว่านั้นยังมีกรณี 16-bin ของ F3 กับ {32, 64}-bin ของ F2, F3 ที่มีนัยสำคัญทางเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถิติที่ระดับ $p < 0.05$ แสดงว่าการทำนายอารมณ์ในบิตดังกล่าวจากคุณลักษณะสีที่มีการเคลื่อนไหวของดวงตามีผลดีกว่าการทำนายแบบรูปภาพต้นฉบับเพียงอย่างเดียวจริง

ส่วนคุณลักษณะรูปร่างพบว่า F1, F2 และ F3 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.05$) ยกเว้น F1 ดังตารางที่ ก.7 และ F2, F3 มีผลการทำนายดีกว่า F1 ในทุกกรณีเช่นเดียวกัน อย่างไรก็ตามมีเพียงกรณี {16, 32}-bin ของ F2 และ 64-bin ของ F2, F3 ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$ แสดงว่าการทำนายอารมณ์ที่มีการเคลื่อนไหวของดวงตามีประสิทธิภาพการทำนายดีกว่า

สุดท้ายคุณลักษณะพื้นผิวพบว่า F1, F2 และ F3 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.05$) ยกเว้น 8-bin ของ F1, F2 และ F3 ดังตารางที่ ก.9 และมี {32, 64}-bin ของ F2 ที่มีผลความถูกต้องเฉลี่ยสูงสุด นอกจากนี้ยังมีกรณี 64-bin ของ F2 ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$ แสดงว่าการทำนายอารมณ์จากคุณลักษณะพื้นผิวที่มีการเคลื่อนไหวของดวงตาสามารถเพิ่มประสิทธิภาพการทำนายได้จริง

จากค่าเฉลี่ยความถูกต้องจากการทำนายทั้ง 3 คุณลักษณะและจากการทดสอบสมมติฐานการวิจัยโดยใช้ Paired t -test ข้างต้นจะเห็นได้ว่าการทำนายในแต่ละคุณลักษณะ F1, F2, และ F3 นั้นให้ผลการทดลองที่ดีกว่า BL ($p < 0.05$) และการใช้ประโยชน์ของการเคลื่อนไหวของตา (F2, F3) นั้นสามารถเพิ่มประสิทธิภาพจาก F1 ได้ ($p < 0.05$) อีกด้วย

ทั้งนี้เพื่อเพิ่มความถูกต้องของการทำนายอารมณ์จากคุณลักษณะรูปร่างให้มากขึ้น เราจะทดลองทำนายอารมณ์โดยการรวมคุณลักษณะของสีร่วมกับรูปร่าง, สีร่วมกับพื้นผิว, รูปร่างร่วมกับพื้นผิว และสีร่วมกับรูปร่างและพื้นผิว โดยนำเวกเตอร์คุณลักษณะแต่ละตัวมาต่อกัน สามารถดูค่าเฉลี่ยการทำนายแต่ละบิตและผลการทดสอบสมมติฐานการวิจัยโดยใช้ Paired t -test เพิ่มเติมได้ในภาคผนวก ก.10-ก.17

จากการรวมคุณลักษณะสีกับรูปร่างพบว่า F1, F2 และ F3 ให้ประสิทธิภาพดีกว่า BL ในทุกกรณี อีกทั้งยังมี {32, 64}-bin ของ F2, F3 ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$

คุณลักษณะสีกับพื้นผิวพบว่า F1, F2 และ F3 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.05$) ดังตารางที่ ก.13 ยิ่งไปกว่านั้น F2, F3 มีผลการทำนายดีกว่า F1 ในทุกกรณีเช่นเดียวกัน อย่างไรก็ตามมีกรณี 64-bin ของ F2 ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$

คุณลักษณะรูปร่างกับพื้นผิวพบว่า F1, F2 และ F3 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.01$) ยกเว้น {16, 32, 64}-bin ของ F1 ดังตารางที่ ก.15 และ F2, F3 มีผลการทำนายดีกว่า F1 ในทุกกรณีเช่นเดียวกัน นอกจากนี้ยังมีกรณี {16, 64}-bin ของ F2 ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$

สุดท้ายผลการทำนายจากการรวมคุณลักษณะทั้ง 3 พบว่า F1, F2 และ F3 ให้ประสิทธิภาพดีกว่า BL ในทุกกรณี ($p < 0.01$) ดังตารางที่ ก.17 และ F2, F3 มีผลการทำนายดีกว่า F1 ในทุกกรณี ยกเว้น 8-bin ของ F2, F3 อย่างไรก็ตามมีกรณี 64-bin ของ F2, F3 เท่านั้น ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$

ดังนั้นเห็นได้ว่าการทำนายในการรวมคุณลักษณะพบว่า สามารถเพิ่มประสิทธิภาพการทำนายได้ดีกว่า F1, F2 และ F3 นั้นให้ผลการทดลองที่ดีกว่า BL เช่นเดิม ($p < 0.01$) และการใช้ประโยชน์ของการเคลื่อนไหวของตา (F2, F3) นั้นสามารถเพิ่มประสิทธิภาพจาก F1 ได้ ($p < 0.05$) แสดงว่าการทำนายอารมณ์ข้างต้นจากคุณลักษณะต่างๆ ที่มีการเคลื่อนไหวของดวงตามีผลดีกว่าการทำนายแบบรูปภาพต้นฉบับเพียงอย่างเดียวจริง

ตารางที่ 4.1 ค่าเฉลี่ยการทำนายอารมณ์ทั้ง 20 คนในแต่ละคุณลักษณะ

Feature	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
Color		43.35	45.85	45.70
Shape		41.95	45.80	44.80
Texture		43.15	42.20	42.15
Color + Shape	36.35	46.60	47.20	46.95
Color + Texture		45.65	47.60	45.65
Shape + Texture		41.40	44.60	44.45
Color + Shape + Texture		47.10	48.05	47.20
Average	36.35	44.17	45.90	45.26

ตารางที่ 4.2 ค่า p -value จากการทำ Pair t -test ของ ประสิทธิภาพในการทำนายอารมณ์

Average	p -value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
Total	0.001	0.001	0.001	0.032	0.124	0.054

จากตารางที่ 4.1 ค่าเฉลี่ยความถูกต้องการทำนาย จากคุณลักษณะพื้นฐาน ชุดข้อมูล F2 และ F3 มีประสิทธิภาพดีกว่า F1 ทุกกรณี ยกเว้น Histogram of Gabor ที่ F1 มีค่าเฉลี่ยผลทำนายดีกว่า นอกจากนี้การรวมคุณลักษณะยังสามารถเพิ่มประสิทธิภาพการทำนายได้มากขึ้นอีกด้วย

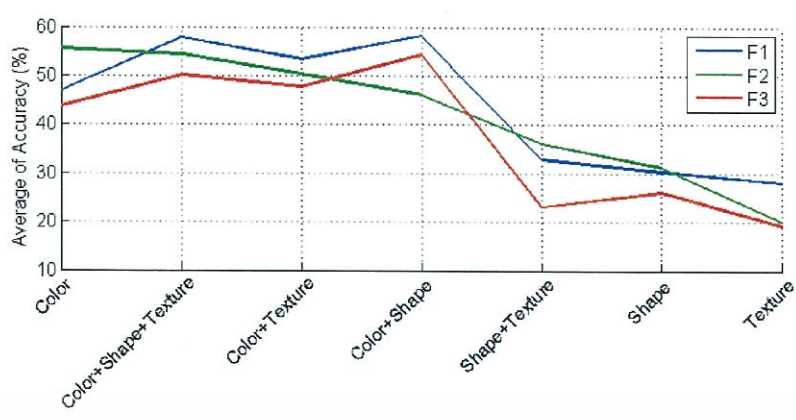
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ประโยชน์เฉพาะเท่านั้น ไม่อนุญาตให้เผยแพร่ข้อมูลด้านการศึกษา
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองจะเห็นได้ว่าการรวมคุณลักษณะเพื่อการรวมคุณลักษณะเพื่อทำนายอารมณ์ โดยรวมทั้ง 3 คุณลักษณะ มีผลค่าเฉลี่ยการทำนายอารมณ์ที่ดีที่สุดเท่ากับร้อยละ 48.05 รองลงมาเป็นการทำนายโดยรวม 2 คุณลักษณะ ซึ่งผลที่ดีที่สุดคือ การรวมลักษณะสีกับพื้นผิวและสีกับรูปร่างเท่ากับ ร้อยละ 47.60 และ 47.20 ตามลำดับ แต่ในการรวมคุณลักษณะรูปร่าง และพื้นผิวให้ผลการทำนายอารมณ์ได้น้อยกว่าการทำนายโดยใช้คุณลักษณะสีเพียงอย่างเดียว ดังนั้นจากตารางดังกล่าวจะสังเกตได้ว่าการรวมคุณลักษณะที่มีสีร่วมอยู่ด้วยนั้นจะมีประสิทธิภาพสูงกว่าการไม่มีสีอยู่ ทำให้สันนิษฐานได้ว่าคุณลักษณะสีมีผลต่อการทำนายอารมณ์รูปภาพได้ดีที่สุด นอกจากนี้การทำนายทั้ง 3 ชุดข้อมูลมีประสิทธิภาพดีกว่า BL ทั้งหมด และจากตารางที่ 2. แสดงค่า p -value พบว่า ชุดข้อมูล F1, F2 และ F3 ให้ผลดีกว่า BL ในทุกกรณี ยิ่งไปกว่านั้น F2, F3 ยังมีประสิทธิภาพดีกว่า F1 อย่างไรก็ตามมีเพียง F2 ที่มีนัยสำคัญทางสถิติที่ $p = 0.032$

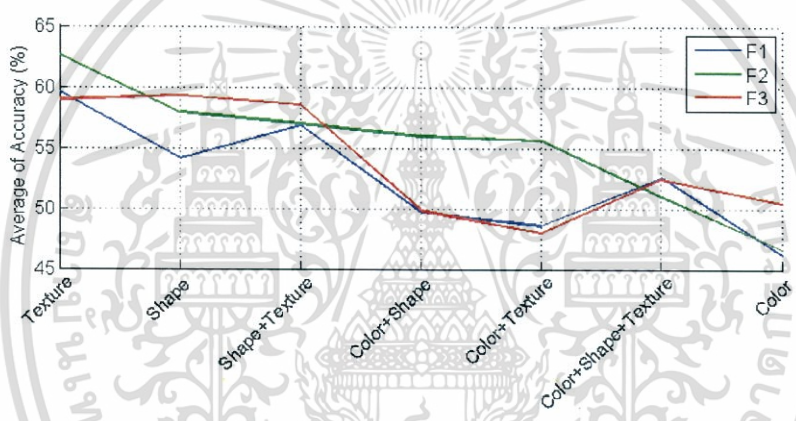
ตารางที่ 4.3 แสดงค่าเฉลี่ย True Positive ทั้ง 4 อารมณ์ของแต่ละคุณลักษณะ

Feature	ความถูกต้องเฉลี่ย (%)		
	F1	F2	F3
Color	41.30	44.30	44.00
Shape	39.90	44.00	43.30
Texture	40.60	39.10	40.00
Color + Shape	43.80	45.70	46.00
Color + Texture	43.20	45.30	44.00
Shape + Texture	38.10	42.30	42.50
Color + Shape + Texture	44.30	46.50	45.40
Total Average	41.60	43.90	43.60

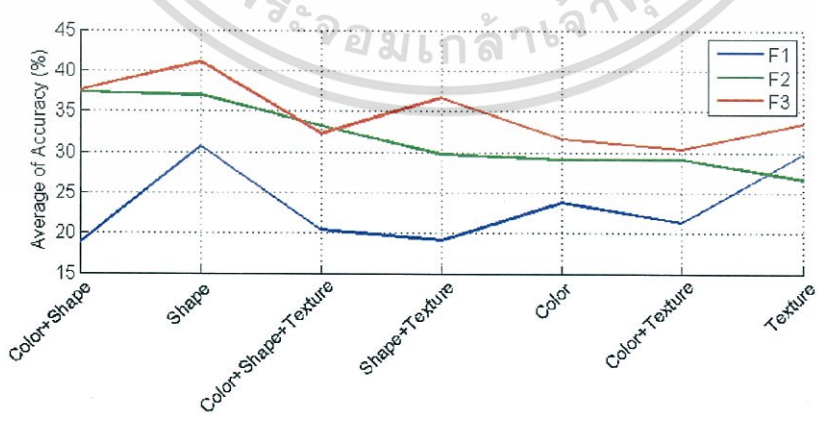
จากตารางที่ 4.3 แสดงค่าเฉลี่ยการทำนายอารมณ์ถูกต้องทั้ง 4 อารมณ์ พบว่า ค่าเฉลี่ยรวมของ F2 มีค่าสูงสุดเท่ากับร้อยละ 43.90 รองลงมาคือ F3 และ F1 ซึ่งมีค่าเท่ากับร้อยละ 43.60 และ 41.60 ตามลำดับ ซึ่ง F2 และ F3 ประสิทธิภาพดีกว่า F1 ในทุกคุณลักษณะ ยกเว้น Histogram of Gabor



รูปที่ 4.1 True Positive อารมณ์มีความสุข

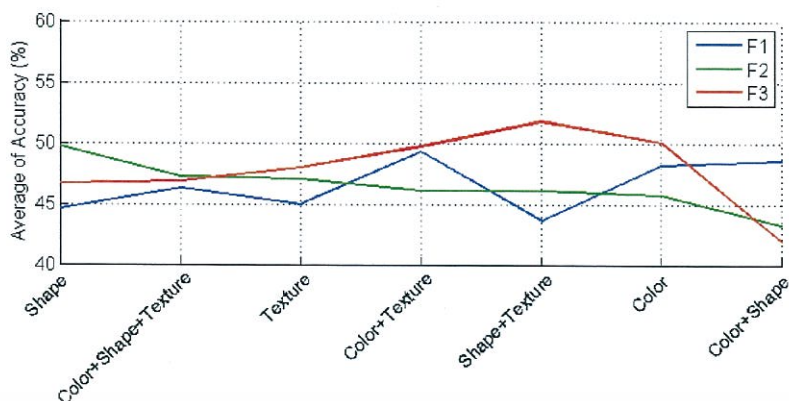


รูปที่ 4.2 True Positive อารมณ์เศร้า



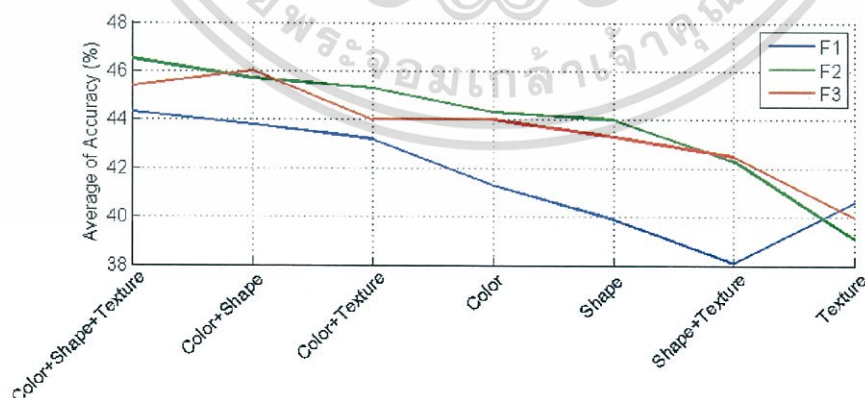
รูปที่ 4.3 True Positive อารมณ์โกรธ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 True Positive อารมณ์กลัว

จะเห็นได้ว่า F1 ให้ผลการทำนายดีที่สุดในการมโนอารมณ์ความสุขและกลัว ดังรูปที่ 4.1 และ 4.4 แต่อย่างไรก็ตามในการมโนเสียใจและโกรธ ส่วนข้อมูลที่มีการเคลื่อนไหวของตา (F2, F3) สามารถทำนายอารมณ์ได้มีประสิทธิภาพที่สุด ดังรูปที่ 4.2 และ 4.3 นอกจากนี้จากค่าเฉลี่ยทั้ง 4 อารมณ์จากทุกคุณลักษณะ พบว่า ชุดข้อมูลที่มีการเคลื่อนไหวดวงตา (F2, F3) มีประสิทธิภาพมากที่สุด ดังรูปที่ 4.5 ทำให้สันนิษฐานได้ว่า การทำนายอารมณ์รูปโดยมีการเคลื่อนไหวของดวงตาของผู้ร่วมเก็บข้อมูลมีประสิทธิภาพการทำนายแม่นยำกว่าการใช้ข้อมูลรูปภาพเพียงอย่างเดียวทำให้สันนิษฐานได้ว่า การทำนายอารมณ์รูปโดยมีการเคลื่อนไหวของดวงตาของผู้ร่วมเก็บข้อมูลในการทำนายอารมณ์ มีประสิทธิภาพการทำนายแม่นยำกว่าการใช้ข้อมูลรูปภาพเพียงอย่างเดียว และในแต่ละชุดข้อมูลจะมีผลการทำนายดีที่สุดในแต่ละอารมณ์ต่างกัน ข้อมูลแสดงผลการทำนายอารมณ์ที่ถูกตั้งอยู่ในแต่ละอารมณ์ สามารถดูรูปภาพเพิ่มเติมได้ในภาคผนวก ก.1- ก.7



รูปที่ 4.5 ค่าเฉลี่ยของ True Positive ของอารมณ์ในแต่ละคุณลักษณะ

จากรูปที่ 4.5 แสดงค่าเฉลี่ย True Positive ของอารมณ์ในแต่ละคุณลักษณะโดยเรียงจากคุณลักษณะที่มีผลการทำนายดีที่สุดไปยังน้อยที่สุด ซึ่งค่า True Positive คือ กรณีที่ผลการทำนายเอกลีขานเป็นเอกลักษณ์ที่ส่งมอบให้หรือการเขียนเพื่อที่ทุกคนได้เห็น เมื่อผู้ดูแลระบบใช้ระบบเอกลีขานการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ออกมาตรงกับอารมณ์เป้าหมายที่เก็บจากผู้เข้าร่วมเก็บข้อมูล จะเห็นได้ว่าในทุกคุณลักษณะ F2 มีผลการทำนาย ดีกว่า F1 ในทุกกรณี ยกเว้นคุณลักษณะพื้นผิว นอกจากนี้การรวมคุณลักษณะที่มีประสิทธิภาพเพิ่มประสิทธิภาพการทำนายได้ดี และการรวมคุณลักษณะทั้ง 3 มีประสิทธิภาพการทำนายดีที่สุดเท่ากับร้อยละ 46.50

การทดลองที่ 2 Global Model New User

จากการทดลองแรกพบว่า F2 มีผลการทำนายดีกว่า F3 อาจเป็นผลมาจาก F3 มีการสูญเสียของข้อมูลเนื่องจากการเบลอ อีกทั้งการทำนายโดยใช้ทั้ง 3 คุณลักษณะ มีประสิทธิภาพมากที่สุด ดังนั้นการทดลองนี้จึงเลือกใช้ F2 และการรวมคุณลักษณะทั้ง 3 ซึ่งในการรวมคุณลักษณะจะนำเวกเตอร์คุณลักษณะแต่ละตัวมาต่อกัน โดยใช้ โดยใช้ข้อมูลของผู้เก็บข้อมูลทั้งหมด 10 คน โดยเลือกจากการสุ่มและใช้เทคนิคการแบ่งข้อมูลแบบ Leave-one-user-out Cross-validation ในการทดสอบ

ตารางที่ 4.4 ค่าเฉลี่ยการทำนายอารมณ์แต่ละผู้ร่วมเก็บข้อมูล

User	ความถูกต้องเฉลี่ย (%)	
	F1	F2
1	48	50
2	51	55
3	59	65
4	47	57
5	49	48
6	62	56
7	55	55
8	52	53
9	55	61
10	56	57
Total Average	53.40	55.70

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.4 พบว่า F2 มีค่าเฉลี่ยความถูกต้องดีกว่า F1 ยกเว้น ผู้ร่วมเก็บข้อมูลที่ 5-7 อาจเป็นผลมาจากพฤติกรรมกรรมการมองที่แตกต่างกันและมีอารมณ์รูปต่างกัน ทำให้ข้อมูลไม่สอดคล้องกับผู้ร่วมเก็บข้อมูลคนอื่น ๆ เท่ากับ 55.70 ซึ่งมีค่าเฉลี่ยการทำนายมากกว่า User Model อาจเป็นเพราะในโมเดลนี้มีตัวอย่างข้อมูลในการฝึกฝนมากกว่า ส่วนค่าเฉลี่ยการทำนายแต่ละบินและผลการทดสอบสมมติฐานการวิจัยโดยใช้ Paired *t*-test สามารถดูตารางได้ในภาคผนวก ก.18-ก.19 ดังนั้นจากผลการทดลองนี้ สันนิษฐานได้ว่าการใช้ข้อมูลการเคลื่อนไหวของดวงตามีผลต่อการทำนายอารมณ์สามารถเพิ่มประสิทธิภาพการทำนายผลได้

4.3. สรุปผลการทดลอง

จากการทดลองทำนายอารมณ์รูปภาพโดยใช้ลักษณะเด่นของภาพร่วมกับข้อมูลการเคลื่อนไหวดวงตาเพื่อเปรียบเทียบกับการใช้ลักษณะเด่นของภาพเพียงอย่างเดียว

ในการทดลอง User Model ในแต่ละคุณลักษณะจาก 3 คุณลักษณะพบว่า สามีผลการทำนายที่ดีที่สุด ยิ่งกว่านั้นในการรวมคุณลักษณะพบว่า หากมีการรวมกันกับคุณลักษณะสีจะส่งผลให้มีผลการทำนายอารมณ์ดีขึ้น จึงสรุปได้ว่าคุณลักษณะสีมีความสำคัญในการทำนายอารมณ์รูปภาพมากที่สุด ยิ่งกว่านั้นการรู้พฤติกรรมกรรมการมองของแต่ละคน จะสามารถทำนายอารมณ์ได้ดี

จากการทดลองทั้ง 2 การทดลองพบว่า Global Model มีผลการทำนายอารมณ์ดีกว่า User Model เนื่องจากมีตัวอย่างข้อมูลในการฝึกฝนมากกว่า ซึ่งแสดงว่าการที่มีข้อมูลมากจะสามารถเพิ่มประสิทธิภาพการทำนายอารมณ์ได้ อย่างไรก็ตาม Global Model ในผู้ร่วมเก็บข้อมูลที่ 5-7 ข้อมูลการเคลื่อนไหวของดวงตาไม่สามารถเพิ่มประสิทธิภาพการทำนายได้ อาจเป็นเพราะข้อมูลการมองของคนดังกล่าวแตกต่างกันและอาจมีอารมณ์ของรูปภาพที่ไม่สอดคล้องกันกับผู้ร่วมเก็บข้อมูลคนอื่น

ดังนั้นถ้าหากเรามีตัวอย่างข้อมูลของ User Model มากกว่านี้ ก็จะสามารถเพิ่มประสิทธิภาพการทำนายอารมณ์ของ User Model ได้

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1. สรุปการทดลอง

งานวิจัยนี้นำเสนอการจำแนกประเภทของรูปภาพเชิงอารมณ์ คือ ดีใจ, เสียใจ, โกรธ และ กลัว โดยใช้ข้อมูล (ก) การเคลื่อนไหวของตาที่มีต่อรูปภาพ ร่วมกับ (ข) คุณลักษณะพื้นฐานจากรูปภาพ คือ สี, รูปร่าง และ พื้นผิว ผลทดลองสรุปได้ว่า การใช้ข้อมูลการเคลื่อนไหวของตาที่มีต่อรูปภาพนั้น สามารถเพิ่มประสิทธิภาพของการจำแนกประเภทของภาพเชิงอารมณ์สำหรับโมเดลของแต่ละผู้ใช้ได้อย่างมีนัยสำคัญทางสถิติ

จากผลการทดลองเปรียบเทียบค่าความแม่นยำดังกล่าว พบว่าการใช้ประโยชน์จากการเคลื่อนไหวของดวงตา สามารถเพิ่มประสิทธิภาพการทำนายอารมณ์ได้ดีกว่าการทำนายแบบใช้คุณลักษณะเพียงอย่างเดียว โดยมีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$ ดังนั้นสมมติฐานของงานวิจัยนี้ และการรวมคุณลักษณะรูปภาพเพื่อการทำนาย พบว่าการรวมคุณลักษณะที่มีสีด้วย สามารถเพิ่มผลการทำนายให้ดียิ่งขึ้น จึงสรุปได้ว่า สีมีผลต่อการทำนายมากที่สุด

5.2. ข้อดีข้อด้อยของการพัฒนา

การทดลองนี้สามารถนำไปพัฒนาการค้นคืนภาพแบบ Emotional Semantic Image Retrieval ได้ และทำให้รู้คุณลักษณะพื้นฐานของรูปที่มีผลต่ออารมณ์ของมนุษย์ แต่มีข้อด้อยที่การเก็บข้อมูลของผู้ร่วมอาจคลาดเคลื่อนเนื่องจากอารมณ์ของมนุษย์มีความซับซ้อน

5.3. ข้อเสนอแนะ

งานวิจัยนี้มีผลการทำนายเกี่ยวกับอารมณ์ของมนุษย์ที่มีต่อรูปภาพซึ่งเป็นเรื่องปัจเจกบุคคลดังนั้นคนแต่ละคนอาจมีอารมณ์ต่อรูปภาพนั้นๆต่างกันอาจขึ้นอยู่กับอายุ, เพศ, การรับรู้, พื้นหลัง, การศึกษาหรือ วัฒนธรรม ของบุคคลแต่ในการทดลองนี้เป็นเพียงกลุ่มคนหนึ่งที่มีความคล้ายคลึงกัน ดังนั้นควรศึกษาการทดลองที่มีหลายกลุ่มบุคคลด้วย

บรรณานุกรม

- [1] L. Ye, P. Ogunbona and J. Wang “**Image Content Annotation Based on Visual Features**”, In: Proceeding of International Symposium on Multimedia (ISM’2006), 11-13 Dec 2006, San Diego, USA, pp. 62–69, 2006.
- [2] J. Laaksonen, M. Koskela, and E. Oja “**PicSOM Self-organizing Image Retrieval with MPEG-7 Content Descriptors**”, IEEE Transactions on Neural Networks, 13(4), pp. 841–853, 2002.
- [3] R. Datta, J. Li, and J. Z. Wang, “**Content-based Image Retrieval – Approaches and Trends of the New Age**”, In: Proceedings of ACM International Workshop on Multimedia Information Retrieval (MIR’2005), 10-11 Nov 2005, Singapore, pp. 253–262, 2005.
- [4] W. Weining, Y. Yinlin, and J. Shengming, “**Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction**”, In: Proceeding of IEEE International Conference on Systems, Man and Cybernetics (SMC’2006), 8-11 Oct 2006, Taipei, Taiwan, pp. 3534–3539, 2006.
- [5] J. Machajdik, and A. Hanbury, “**Affective Image Classification Using Features Inspired by Psychology and Art Theory**,” In: Proceeding of ACM International Conference on Multimedia (MM’2010), 25-29 Oct 2010, Firenze, Italy, pp. 83–92, 2010.
- [6] H. Zhang, E. Augilius, T. Honkela, J. Laaksonen, H. Gamper, and H. Alene, “**Analyzing Emotional Semantics of Abstract Art Using Low-level Image Features**”, In: Proceeding of International Symposium on Intelligent Data Analysis (IDA’2011), 29-31 Oct 2011, Porto, Portugal, pp. 413–423, 2011.
- [7] H. Zhang, M. Gönen, Z. Yang and E. Oja, “**Predicting Emotional States of Images Using Bayesian Multiple Kernel Learning**”, In: Proceeding of International Conference on Neural Information Processing (ICONIP’2013), 3-7 Nov 2013, Daegu, Korea, pp. 274–282, 2013.
- [8] K. Pasupa, C. J. Saunders, S. Szedmak, A. Klami, S. Kaski, and S. R. Gunn, “**Learning to Rank Images from Eye movements**”, In: Proceeding of 2009 IEEE 12th International Conference on Computer Vision (ICCV’2009) Workshops on Human-Computer Interaction (HCI’2009), 27 Sep-4 Oct 2009, Kyoto, Japan, pp. 2009–2016, 2009.

บรรณานุกรม (ต่อ)

- [9] D. R. Hardoon and K. Pasupa “**Image Ranking with Implicit Feedback from Eye Movements**” In: Proceedings of the 6th Biennial Symposium on Eye Tracking Research & Applications (ETRA'2010), 22-24 Mar 2010, Austin, USA, pp. 291–298, 2010.
- [10] P. Auer, Z. Hussain, S. Kaski, A. Klami, J. Kujala, J. Laaksonen, A. P. Leung, K. Pasupa, and J. Shawe-Taylor “**Pinview: Implicit Feedback in Content-Based Image Retrieval**” In: Proceeding of Workshop on Applications of Pattern Analysis (WAPA'2010), 1-2 Sep 2010, Cumberland Lodge, UK, pp 51–57, 2010.
- [11] C. E. Izard “**Basic Emotions, Relations Among Emotions, and Emotion-Cognition Relations**”, Psychological Review, 99(3), pp. 561–565, 1992.
- [12] K. Vytal and S. Hamann “**Neuroimaging Support for Discrete Neural Correlates of Basic Emotions: A Voxel-based Meta-analysis**”, Cognitive Neuroscience, 22(12), pp. 2864–2885, 2010.
- [13] P. Ekman “**Universals and Cultural Differences in Facial Expressions of Emotion**” Nebraska Symposium on Motivation, 19, pp. 207–282, 1972.
- [14] R. E. Jack, O. G.B. Garrod, and P. G. Schyns “**Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals Over Time**”, Current Biology, 24(2), pp. 187–192, 2014.
- [15] R. Plutchik and H. Kellerman, “**Emotion: Theory, Research and Experience**” Psychological Medicine, 11(1), pp. 207, 1980.
- [16] D. W. Galenson “**Two Paths to Abstract Art Kandinsky and Malevich**”, Technical Report, National Bureau of Economic Research, No. 12403, 2006.
- [17] Y. Wu, C. Bauckhage and C. Thurau “**The Good, the Bad, and the Ugly: Predicting Aesthetic Image Labels**”, In: Proceedings of 20th International Conference on Pattern Recognition (ICPR'2010). 23-26 Aug 2010, Istanbul, Turkey. pp. 1586–1589, 2010.
- [18] The Eye Tribe Aps, “**The Eye Tribe**”, Available at <https://theeyetribe.com>.
- [19] Changing Minds, “**Basic Emotion**”, Available at

ภาคผนวก ก.

การแสดงผลการทำนายอารมณ์จากคุณลักษณะในแต่ละบิน และผลการทดสอบสมมติฐานการวิจัย

โดยใช้ Paired *t*-test

ตารางที่ ก.1 ผลการเก็บอารมณ์รูปภาพแต่ละผู้ร่วมเก็บข้อมูล

User	ดีใจ	เสียใจ	โกรธ	กลัว
1	9	38	7	46
2	22	36	14	28
3	21	25	23	31
4	15	32	19	34
5	28	24	21	27
6	35	30	19	16
7	31	20	23	26
8	20	49	21	10
9	33	31	10	26
10	22	26	27	25
11	26	46	8	20
12	30	32	2	36
13	27	33	17	23
14	24	26	12	38
15	16	34	19	31
16	29	34	21	16
17	19	31	19	31
18	18	52	10	20
19	30	26	18	26
20	43	31	9	17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.2 แสดงข้อมูลคิบบททั้ง 20 คน ของตาราง 4.1

U S E R	BL	RGB - F1	RGB - F2	RGB-F3	Sobel-F1	Sobel-F2	Sobel- F3	Gabor - F1	Gabor - F2	Gabor - F3	RGB + Sobel F1	RGB + Sobel F2	RGB + Sobel F3	RGB + Gabor F1	RGB + Gabor F2	RGB + Gabor F3	Sobel + Gabor F1	Sobel + Gabor F2	Sobel + Gabor F3	RGB + Sobel + Gabor F1	RGB + Sobel + Gabor F2	RGB + Sobel + Gabor F3
1	46	54	58	55	52	58	51	53	58	56	55	53	53	53	57	54	57	54	56	57	54	52
2	36	39	38	40	39	43	37	45	37	34	48	45	41	44	42	43	54	48	39	48	45	40
3	31	35	38	36	40	40	39	38	34	33	39	36	34	41	39	37	44	42	41	40	41	35
4	34	38	41	41	48	46	51	43	40	42	44	41	38	38	45	41	35	44	50	44	41	38
5	28	45	46	46	34	46	49	36	37	34	46	45	44	50	45	45	37	38	38	48	48	45
6	35	43	48	45	39	43	49	43	35	40	50	48	48	47	48	47	36	42	44	50	48	48
7	31	48	46	42	37	42	41	35	34	37	55	52	44	49	48	43	36	40	39	55	50	46
8	49	50	53	53	51	50	49	54	51	52	52	50	52	49	56	54	53	52	49	53	53	51
9	33	37	45	45	34	42	38	36	41	39	37	46	49	40	54	44	26	42	41	37	48	49
10	27	40	41	37	44	41	39	37	38	43	35	45	36	44	41	38	35	37	38	35	41	39
11	46	50	51	51	46	50	52	45	46	42	54	50	52	51	49	49	45	51	50	54	51	52
12	36	52	55	48	47	48	44	48	47	45	56	56	56	57	53	51	44	46	51	56	57	54
13	33	36	41	43	37	39	47	45	39	42	38	44	44	35	43	43	31	38	39	38	49	48
14	38	46	49	43	40	54	49	39	48	45	50	55	50	51	54	42	45	54	47	53	55	48
15	34	43	40	46	36	52	40	34	43	44	44	43	47	42	39	43	33	44	43	47	45	47
16	34	40	48	50	45	44	43	38	39	43	50	50	51	46	54	42	46	41	42	50	53	52
17	31	41	43	50	36	46	39	43	36	40	45	48	49	46	46	48	38	42	43	44	46	48
18	52	52	54	58	52	54	54	52	52	52	52	53	59	52	54	58	52	53	53	52	53	59
19	30	37	39	46	43	35	42	51	43	35	40	41	51	37	43	44	40	40	44	41	39	52
20	43	41	43	39	39	43	43	48	46	45	42	43	41	41	42	39	41	44	42	40	44	41

ตารางที่ ก.3 แสดงข้อมูลคิบบททั้ง 20 คน ของตาราง 4.3

U S E R	BL	RGB - F1	RGB - F2	RGB-F3	Sobel-F1	Sobel-F2	Sobel- F3	Gabor - F1	Gabor - F2	Gabor - F3	RGB + Sobel F1	RGB + Sobel F2	RGB + Sobel F3	RGB + Gabor F1	RGB + Gabor F2	RGB + Gabor F3	Sobel + Gabor F1	Sobel + Gabor F2	Sobel + Gabor F3	RGB + Sobel + Gabor F1	RGB + Sobel + Gabor F2	RGB + Sobel + Gabor F3
1	31	37	42	29	34	28	31	34	33	32	44	44	30	36	39	34	31	32	33	47	45	52
2	34	35	33	37	36	35	37	30	29	42	43	37	37	38	39	47	40	38	46	40	38	40
3	34	36	35	37	38	38	35	32	31	38	35	36	39	36	36	42	41	40	39	41	34	35
4	34	38	39	45	42	49	38	31	34	41	41	39	34	41	39	31	38	41	41	37	33	38
5	46	44	44	34	47	48	35	37	36	45	44	43	49	45	43	37	38	40	47	48	44	45
6	33	40	37	35	35	50	41	31	36	42	39	38	37	40	38	36	36	46	42	39	38	48
7	47	44	41	35	41	41	34	33	39	53	49	42	46	45	41	37	40	39	53	47	44	46
8	35	32	33	30	28	26	35	27	29	29	34	40	33	38	43	31	32	26	30	38	29	51
9	32	36	42	32	39	33	33	41	38	31	47	42	34	49	41	23	36	34	30	47	43	49

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการเข้าถึงเพื่อการศึกษาเท่านั้น เมื่อผู้จัดทำนำไปเผยแพร่โดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.3 แสดงข้อมูลดิบทั้ง 20 คน ของตาราง 4.3 (ต่อ)

U S E R	BL	RGB - F1	RGB - F2	RGB-F3	Sobel-F1	Sobel-F2	Sobel- F3	Gabor - F1	Gabor - F2	Gabor - F3	RGB + Sobel F1	RGB + Sobel F2	RGB + Sobel F3	RGB + Gabor	RGB + Gabor	RGB + Gabor	Sobel + Gabor	Sobel + Gabor	Sobel + Gabor	RGB + Sobel +	RGB + Sobel +	RGB + Sobel +
10	40	39	36	43	39	37	35	37	42	35	43	35	43	40	37	35	36	37	35	40	38	39
11	36	37	32	27	34	34	30	33	28	35	39	38	38	35	33	27	41	33	35	38	38	52
12	39	42	36	35	36	34	36	35	45	43	43	43	44	41	38	34	34	38	43	44	40	54
13	37	41	44	35	39	46	47	39	46	37	44	44	34	39	43	29	33	37	38	49	47	48
14	37	45	35	32	52	45	34	41	36	42	46	42	42	46	33	37	48	37	44	46	40	48
15	39	34	42	33	48	33	28	37	40	41	41	44	39	35	42	35	40	38	46	42	41	47
16	38	44	47	38	45	42	33	36	38	43	46	48	41	50	44	41	38	45	43	51	49	52
17	40	42	48	36	47	39	42	33	36	46	49	49	47	46	45	34	41	42	45	47	45	48
18	25	37	39	25	31	28	25	25	25	25	26	45	25	37	39	25	30	27	25	26	45	59
19	33	36	45	44	34	43	50	41	36	36	42	49	33	39	41	40	37	44	37	36	51	52
20	33	31	24	30	37	40	34	41	32	34	34	25	32	28	31	27	28	34	32	38	26	41

ตารางที่ ก.4 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากที

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8		42.50	43.95	43.00
16	36.35	40.45	42.35	43.35
32		39.0	43.00	42.65
64		38.95	44.30	43.85

ตารางที่ ก.5 ค่า *p*-value จากการทำ Pair *t*-test ของประสิทธิภาพในการทำนายอารมณ์

<i>nbin</i>	<i>P</i> -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.064	0.661	0.319
16	0.010	0.001	0.001	0.060	0.003	0.365
32	0.066	0.001	0.001	0.001	0.013	0.738
64	0.036	0.001	0.001	0.001	0.001	0.694

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.6 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากรูปร่าง

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	36.35	38.90	40.05	40.90
16		38.00	42.85	40.65
32		37.95	41.40	41.20
64		35.15	42.50	41.05

ตารางที่ ก.7 ค่า *p*-value จากการทำ Pair *t*-test ของประสิทธิภาพในการทำนายอารมณ์จากรูปร่าง

<i>nbin</i>	<i>P</i> -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.087	0.001	0.001	0.462	0.203	0.541
16	0.153	0.001	0.001	0.003	0.086	0.177
32	0.207	0.001	0.001	0.040	0.060	0.877
64	0.366	0.001	0.001	0.001	0.001	0.208

ตารางที่ ก.8 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบินจากพื้นผิว

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	36.35	37.85	38.0	38.15
16		40.45	39.90	40.10
32		39.95	40.95	39.65
64		39.75	40.55	40.20

ตารางที่ ก.9 ค่า *p*-value จากการทำ Pair *t*-test ของประสิทธิภาพในการทำนายอารมณ์จากพื้นผิว

<i>nbin</i>	<i>P</i> -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.204	0.070	0.128	0.898	0.814	0.855
16	0.001	0.003	0.001	0.643	0.806	0.838
32	0.014	0.001	0.005	0.547	0.866	0.213
64	0.022	0.001	0.001	0.001	0.771	0.665

ตารางที่ ก.10 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิตจากสี่และรูปร่าง

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	36.35	45.40	45.15	44.20
16		43.85	43.75	43.85
32		41.85	44.50	44.60
64		41.30	44.20	45.20

ตารางที่ ก.11 ค่า *p*-value จากการทำ Pair *t*-test ของประสิทธิภาพในการทำนายอารมณ์จากสี่และรูปร่าง

<i>nbin</i>	<i>P</i> -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.857	0.415	0.475
16	0.001	0.001	0.001	0.924	1	0.918
32	0.001	0.001	0.001	0.026	0.040	0.927
64	0.001	0.001	0.001	0.001	0.001	0.381

ตารางที่ ก.12 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิตจากสี่และพื้นผิว

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	36.35	42.95	44.60	43.70
16		41.85	43.20	43.00
32		42.55	44.95	43.00
64		42.75	45.20	43.85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.13 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากสีและพื้นผิว

$nbin$	P -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.131	0.516	0.420
16	0.003	0.001	0.001	0.239	0.314	0.841
32	0.001	0.001	0.001	0.077	0.743	0.095
64	0.001	0.001	0.001	0.001	0.293	0.103

ตารางที่ ก.14 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิตจากรูปร่างและพื้นผิว

$nbin$	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8		39.30	41.15	40.60
16		38.25	40.85	41.10
32	36.35	38.55	41.25	40.25
64		38.20	42.50	40.30

ตารางที่ ก.15 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์จากรูปร่างและพื้นผิว

$nbin$	P -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.209	0.354	0.534
16	0.165	0.001	0.001	0.046	0.123	0.833
32	0.105	0.001	0.001	0.058	0.244	0.374
64	0.168	0.001	0.001	0.001	0.236	0.139

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.16 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิตทั้ง 3 คุณลักษณะ

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	36.35	45.35	44.85	44.25
16		43.85	44.40	43.90
32		43.30	44.40	44.60
64		41.75	45.45	45.60

ตารางที่ ก.17 ค่า *p*-value จากการทำ Pair *t*-test ของประสิทธิภาพในการทำนายอารมณ์ ทั้ง 3 คุณลักษณะ

<i>nbin</i>	<i>P-Value</i>					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.706	0.494	0.609
16	0.001	0.001	0.001	0.653	0.972	0.720
32	0.001	0.001	0.001	0.282	0.366	0.862
64	0.002	0.001	0.001	0.001	0.010	0.888

การทดลองที่ 2

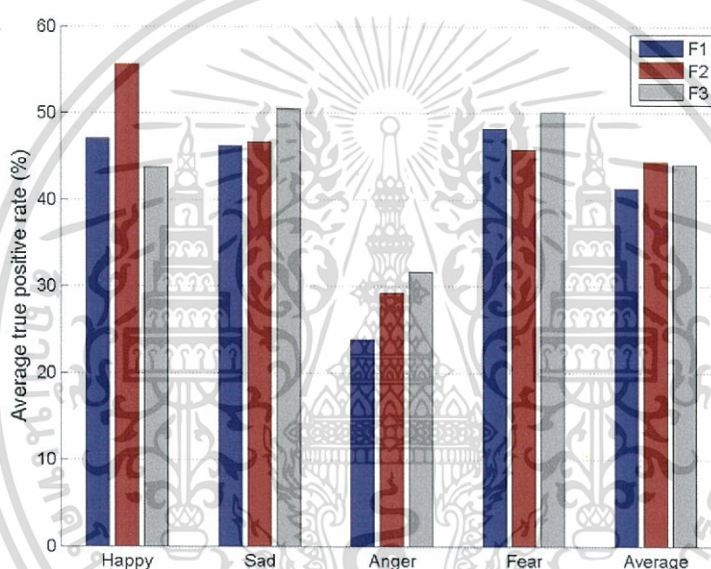
ตารางที่ ก.18 ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์แต่ละบิต ทั้ง 3 คุณลักษณะ

<i>nbin</i>	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
8	39.70	51.90	52.00	47.60
16		52.50	52.60	48.70
32		51.30	55.00	49.80
64		50.70	53.80	49.90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

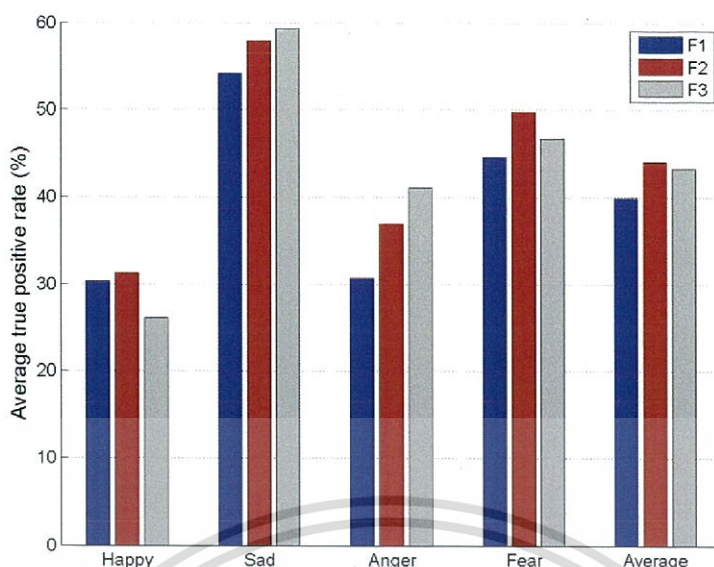
ตารางที่ ก.19 ค่า p -value จากการทำ Pair t -test ของประสิทธิภาพในการทำนายอารมณ์ ทั้ง 3
คุณลักษณะ

n_{bin}	P -Value					
	F1-vs-BL	F2-vs-BL	F3-vs-BL	F1-vs-F2	F1-vs-F3	F2-vs-F3
8	0.001	0.001	0.001	0.955	0.044	0.009
16	0.001	0.001	0.003	0.949	0.010	0.026
32	0.004	0.001	0.001	0.034	0.264	0.003
64	0.006	0.001	0.001	0.001	0.678	0.015



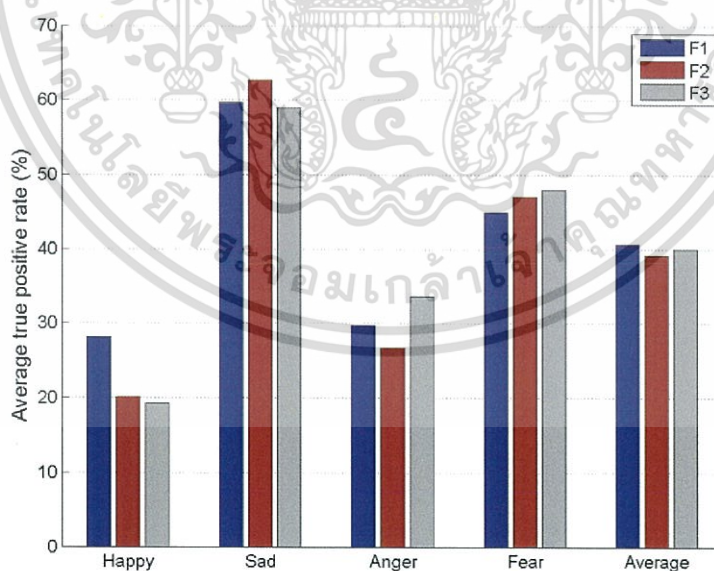
รูปที่ ก.1 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จาก 3

รูปที่ ก.1 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F2 จะมีผลดีกว่า F1 ยกเว้นอารมณ์กลัว ส่วน F3 ดีกว่า F1 เพียงอารมณ์เสียใจ และ โกรธ อย่างไรก็ตามในภาพรวมนั้น F2 นั้นให้ผลดีที่สุด ตามมาด้วย F3 และ F1 ตามลำดับ



รูปที่ ก.2 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากรูปร่าง

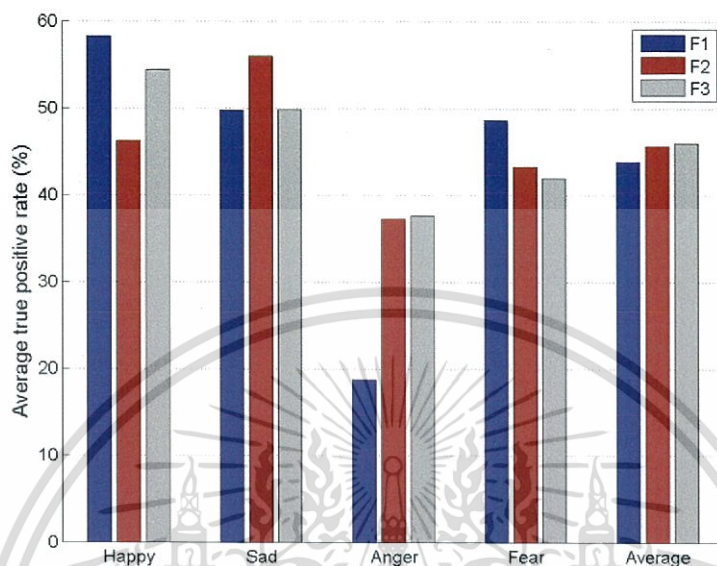
รูปที่ ก.2 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F3 จะมีผลดีกว่า F1 ยกเว้นอารมณ์ความสุข ส่วน F2 ดีกว่า F1 เพียงอารมณ์เสียใจ และ กลัว อย่างไรก็ตามในภาพรวมนั้น F3 ให้ผลดีที่สุด ตามมาด้วย F2 และ F1 ตามลำดับ



รูปที่ ก.3 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากพื้นผิว

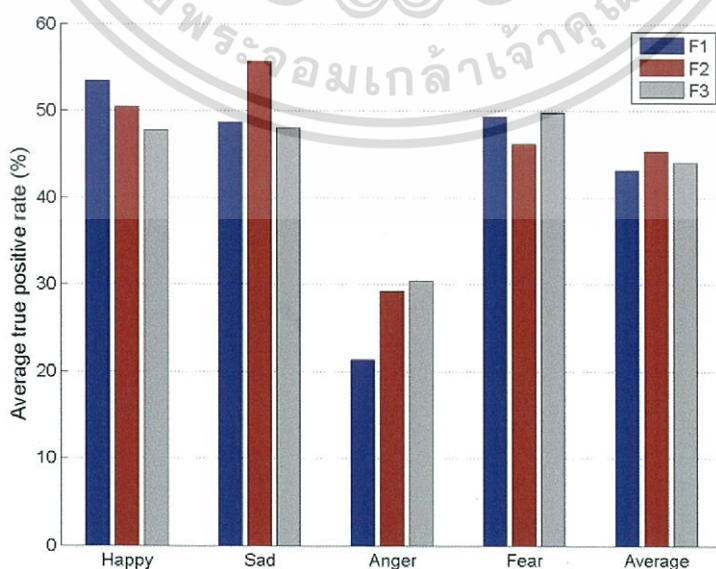
รูปที่ ก.3 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F3 จะมีผลดีกว่า F1 เพียงอารมณ์เสียใจ และ โกรธ ส่วน F2 ดีกว่า F1 เพียงเอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์เพื่อการเรียนเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ขึ้นต้นการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อารมณ์เสียใจ และ กลัว อย่างไรก็ตามในภาพรวมนั้น F1 ให้ผลดีที่สุด ตามมาด้วย F3 และ F2 ตามลำดับ



รูปที่ ก.4 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากสี และรูปร่าง

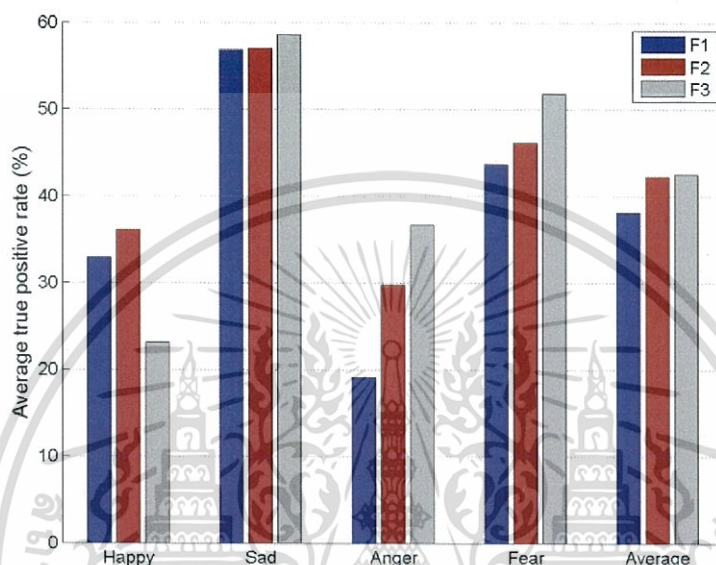
รูปที่ ก.4 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F2 จะมีผลดีกว่า F1 เพียงอารมณ์เสียใจ และ โกรธ ส่วน F3 ดีกว่า F1 เพียงอารมณ์โกรธ อย่างไรก็ตามในภาพรวมนั้น F2 ให้ผลดีที่สุด ตามมาด้วย F1 และ F3 ตามลำดับ



รูปที่ ก.5 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากสี และพื้นผิว

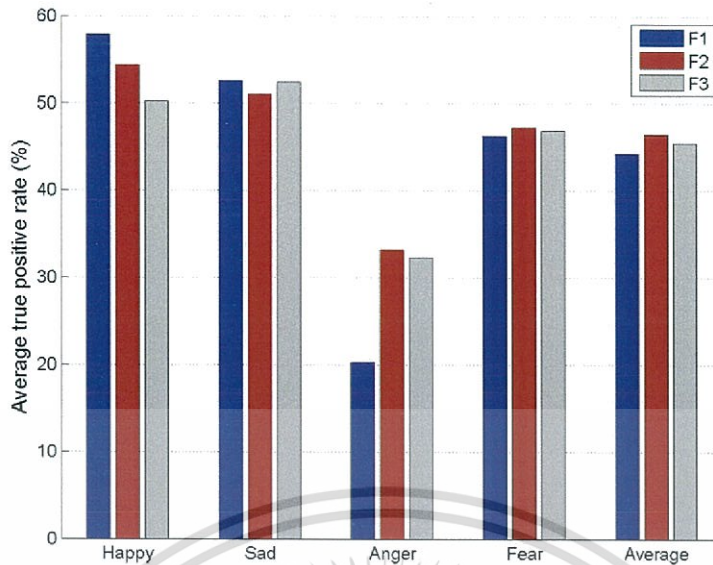
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่หรือนำไปใช้ในการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ก.5 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F2 จะมีผลดีกว่า F1 เพียงอารมณ์เสียใจ และ โกรธ ส่วน F3 ดีกว่า F1 เพียงอารมณ์เสียใจ และ โกรธ อย่างไรก็ตามในภาพรวมนั้น F2 ให้ผลดีที่สุด ตามมาด้วย F1 และ F3 ตามลำดับ



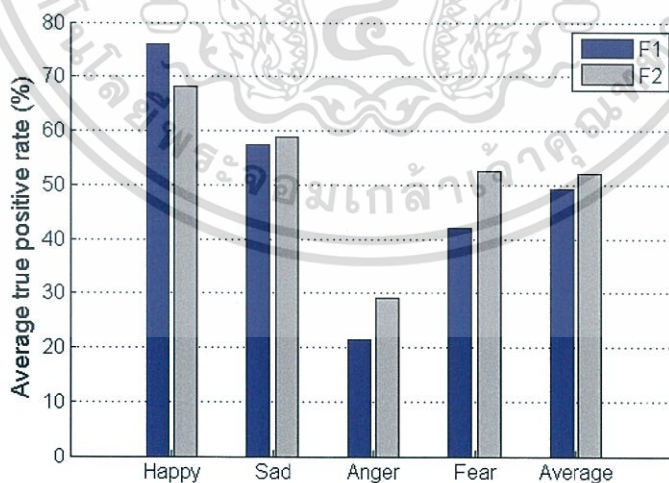
รูปที่ ก.6 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากรูปร่าง และพื้นผิว

รูปที่ ก.6 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F2 จะมีผลดีกว่า F1 ในทุกกรณี ยกเว้นเพียงอารมณ์เสียใจ ส่วน F3 ดีกว่า F1 เพียงอารมณ์โกรธ และกลัว อย่างไรก็ตามในภาพรวมนั้น F2 ให้ผลดีที่สุด ตามมาด้วย F3 และ F1 ตามลำดับ



รูปที่ ก.7 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ User Model ในแต่ละอารมณ์ทั้ง 3 คุณลักษณะ

รูปที่ ก.7 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1, F2 และ F3 จะเห็นได้ว่า F2 จะมีผลดีกว่า F1 เพียงอารมณ์โกรธ และกลัว ส่วน F3 ดีกว่า F1 เพียงอารมณ์โกรธ อย่างไรก็ตามในภาพรวมนั้น F2 ให้ผลดีที่สุด ตามมาด้วย F1 และ F3 ตามลำดับ



รูปที่ ก.8 ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ Global Model ในแต่ละอารมณ์ทั้ง 3 คุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ก.8 แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และทุกอารมณ์ของทั้ง F1 และ F2 จะเห็นได้ว่า F2 มีประสิทธิภาพดีกว่า F1 ในทุกอารมณ์ ยกเว้นอารมณ์ความสุข นอกจากนี้ในภาพรวมนั้น F2 มีประสิทธิภาพดีที่สุดอีกด้วย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ผู้เขียน	นางสาว โชติรส วุฒิเลิศเดชา
วันเดือนปีเกิด	23 ตุลาคม 2535
สถานที่เกิด	จังหวัด กรุงเทพมหานคร
ปริญญา	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ประสบการณ์ทำงาน	ไม่มี
งานวิจัยที่ตีพิมพ์	ไม่มี

ผู้เขียน	นางสาว ภาณวี ฉัตรคำจุนเจริญ
วันเดือนปีเกิด	28 มกราคม 2536
สถานที่เกิด	จังหวัด กรุงเทพมหานคร
ปริญญา	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ประสบการณ์ทำงาน	ไม่มี
งานวิจัยที่ตีพิมพ์	ไม่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้คุณลักษณะของภาพและ เครื่องมือติดตามการมองเห็น

ภาณวี ฉัตรคำจุนเจริญ¹, โชติรส วุฒิเลิศเดชา², ผศ.ดร. กิตส์สุชาติ พสุภา³

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: {¹panawee.c@gmail.com, ²chotiroswuttillertdeshar@gmail.com, ³kitsuchart@it.kmit.ac.th}

บทคัดย่อ

ปัจจุบันการสืบค้นเชิงอารมณ์เป็นส่วนที่จะสนับสนุนให้ผู้ใช้ได้เข้าถึงข้อมูลในคลังข้อมูลโดยครอบคลุมความต้องการของผู้ใช้ได้มากขึ้น ถึงแม้ว่ารูปภาพรูปเดียวกันก็ตาม ผู้ใช้ต่างกัน อารมณ์ของผู้ใช้ที่ถูกกระตุ้นโดยรูปภาพนั้นอาจจะแตกต่างกัน ขึ้นอยู่กับส่วนของภาพที่ผู้ชมมองและสนใจ บทความนี้นำเสนอการเพิ่มประสิทธิภาพของการจำแนกประเภทรูปภาพเชิงอารมณ์ของผู้ใช้ โดยการใช้ประโยชน์จากข้อมูลการเคลื่อนไหวของตาที่เก็บจากผู้ใช้ขณะมองรูปภาพ กับคุณลักษณะพื้นฐานของรูปภาพจากการทดลองพบว่าการเคลื่อนไหวของตาที่มีต่อรูปภาพนั้นสามารถเพิ่มประสิทธิภาพในการจำแนกประเภทอารมณ์ได้ดีกว่าการใช้เพียงคุณลักษณะพื้นฐาน และการรวมคุณลักษณะสี รูปร่าง และพื้นผิวเข้าด้วยกันสามารถเพิ่มประสิทธิภาพในการทำนายได้ดีกว่าใช้เพียงคุณลักษณะเดียว ยิ่งไปกว่านั้นยังพบว่าคุณลักษณะสีมีผลต่อการทำนายมากที่สุด

คำสำคัญ—Eye Movements; Implicit Feedback; Emotion; Image Retrieval

1. บทนำ

การค้นคืนภาพ (Image Retrieval) คือ การค้นหารูปภาพที่ผู้ใช้ต้องการจากรูปภาพที่มีในฐานข้อมูล ในอดีตระบบการค้นคืนภาพจะเป็นการค้นหาโดยใช้คำสำคัญ (Keyword) ที่ถูกแท็ก (Tag) กับรูปภาพแต่ละรูปในฐานข้อมูล ในการแท็กรูปภาพทั้งหมดที่มีอยู่ในฐานข้อมูล ต้องใช้แรงงานมาก จึงส่งผลให้มีการพัฒนาระบบเพื่อมีการแท็กรูปภาพอัตโนมัติ [1] หรือใช้การค้นหาแบบอิงเนื้อหา (Content Based Image Retrieval; CBIR) โดยการเปรียบเทียบภาพคำถาม (Query Image) กับรูปภาพทั้งหมดที่อยู่ในฐานข้อมูล ตัวอย่างของระบบ CBIR คือ PICSOM [2], Google Image Search [3] เป็นต้น จากนั้นในปี 2006 มีการนำเสนอระบบการค้นคืนภาพจากอารมณ์ของรูปภาพ (Emotion Sematic Image Retrieval; ESIR) [4] และเริ่มมีงานวิจัยที่ศึกษาเกี่ยวกับ ESIR มากขึ้น เช่น การศึกษาคุณลักษณะ (Feature) ในการจำแนกอารมณ์รูปภาพโดยใช้ทฤษฎีด้านจิตวิทยาและศิลปะในการหาอารมณ์รูปภาพ [5] การศึกษา

ของรูปภาพนามธรรม (Abstract Art) [6] และการศึกษาคุณลักษณะของรูปภาพที่มีผลต่อการจำแนกอารมณ์โดยใช้ Multiple Kernel Learning (MKL) โดยพื้นฐานของคุณลักษณะที่ใช้คือ สี, รูปร่างและพื้นผิว [7]

ในระบบ CBIR จะต้องมีการป้อนกลับความเกี่ยวข้อง (Relevance Feedback) เพื่อประเมินตัวอย่างรูปภาพที่ค้นคืนโดยระบบโดยผู้ใช้ นั่นคือการคลิกเมาส์ (Click) ซึ่งการคลิกเมาส์นั้นจัดเป็นการป้อนกลับโดยชัดเจน (Explicit Feedback) ซึ่งเกิดจากความตั้งใจของผู้ใช้ ซึ่งต้องใช้แรงงานในการป้อนกลับ ในปี 2009 ได้มีการเริ่มพัฒนาระบบค้นคืนโดยเริ่มจากการเรียงลำดับของรูปภาพโดยใช้การเคลื่อนไหวของตา [8,9] งานวิจัยนี้ใช้ประโยชน์ของการป้อนกลับโดยปริยาย (Implicit Feedback) ซึ่งเกิดจากความไม่ตั้งใจในขณะที่กำลังเรียงรูปภาพ ในการเพิ่มประสิทธิภาพของระบบ จากนั้นจึงมีการพัฒนาระบบค้นคืนภาพที่ใช้ทั้งการป้อนกลับโดยชัดเจนและการป้อนกลับโดยปริยายที่เรียกว่าระบบ PinView [10]

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ การใช้งานโดยไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสำนวน “Beauty is in the eye of the beholder” หรือ “ความสวยขึ้นอยู่กับคนมอง” จึงทำให้เกิดคำถามวิจัยที่ว่า รูปภาพหนึ่งรูป และมีคนมองหลายคน อารมณ์ที่ถูกกระตุ้นโดยรูปภาพของแต่ละผู้มองนั้นอาจไม่เหมือนกัน ขึ้นอยู่กับพื้นที่ที่ผู้ใช้กำลังมองรูปภาพนั้นอยู่ งานวิจัยนี้จึงได้นำเสนอการจำแนกอารมณ์ของรูปภาพโดยใช้คุณลักษณะของภาพและการเคลื่อนไหวของตา โดยมีสมมติฐานที่ว่า การใช้คุณลักษณะของภาพร่วมกับข้อมูลการเคลื่อนไหวของตานั้นสามารถเพิ่มประสิทธิภาพในการจำแนกอารมณ์ของผู้ใช้ที่มีต่อรูปภาพจากการใช้คุณลักษณะของภาพอย่างเดียวได้

2. ทฤษฎีที่เกี่ยวข้อง

2.1. อารมณ์ (Emotions)

อารมณ์คือความรู้สึกที่ได้รับการกระทบจากสิ่งเร้า ซึ่งการแบ่งอารมณ์พื้นฐานของมนุษย์สามารถอธิบายได้ 2 แนวทาง คือ

1) Discrete Approach – อารมณ์แต่ละอารมณ์นั้นแตกต่างกันอย่างสิ้นเชิง โดยแต่ละอารมณ์จะมีระดับความเข้มของตัวเอง ทฤษฎีต่าง ๆ ที่อธิบายอารมณ์โดยใช้แนวทางนี้ เช่น [11,12] ในงานวิจัยนี้จะแบ่งอารมณ์โดยใช้แนวทางของ Ekman (1972) ซึ่งแบ่งอารมณ์ของมนุษย์ออกเป็น 6 อารมณ์ นั่นคือ โกรธ (Anger), ขยะแขยง (Disgust), กลัว (Fear), ดีใจ (Happiness), เสียใจ (Sadness) และประหลาดใจ (Surprise) [13] อารมณ์ดังกล่าวได้ถูกจัดกลุ่มใหม่เป็น 4 อารมณ์ คือ ดีใจ, เสียใจ, โกรธและ กลัว โดยแบ่งตามกลไกเนื้อการแสดงสีหน้าของแต่ละอารมณ์ โดยพบว่ากลไกเนื้อของการแสดงสีหน้าของมนุษย์จากอารมณ์ประหลาดใจมีลักษณะคล้ายกับกลัว เช่นเดียวกับอารมณ์ขยะแขยงและโกรธ [14]

2) Dimensional Approach –ซึ่งอารมณ์ต่าง ๆ เกิดขึ้นได้จากการผสมกันระหว่างอารมณ์พื้นฐาน อารมณ์พื้นฐานแต่ละอารมณ์จะมีระดับความเข้มของอารมณ์ด้วย ตัวอย่างเช่น Plutchik (1980) ได้แบ่งอารมณ์พื้นฐานออกเป็น 8 อารมณ์ คือ สุข (Joy), เสียใจ (Sadness), โกรธ (Anger), กลัว (Fear), ไว้วางใจ (Trust), รังเกียจ (Disgust), ประหลาดใจ (Surprise), และ คาดหวัง (Anticipation)

[15] ซึ่งอารมณ์พื้นฐานเหล่านี้สามารถผสมเป็นอารมณ์ใหม่ได้ เช่น สุข-ไว้วางใจ จะได้ อารมณ์รัก (Love)

2.2. รูปภาพนามธรรม (Abstract Art)

รูปภาพนามธรรม คือ รูปภาพที่แสดงถึงอารมณ์หรือความรู้สึกของผู้วาด Wassily Kandinsky ผู้คิดศิลปะนามธรรมกล่าวว่า สิ่งที่สำคัญสำหรับศิลปะนามธรรม คือ สีและรูปทรงที่ใช้ในการถ่ายทอดโดยคำนึงถึงความรู้สึกภายนอกและภายใน กล่าวคือ การถ่ายทอดรูปทรงต่าง ๆ ให้กลมกลืนด้วยสี ลักษณะผิว หรือ ความเด่นชัดของรูปภาพ เป็นต้น เพื่อให้เกิดอารมณ์และความรู้สึก [16] ลักษณะของรูปนามธรรมจะไม่สื่อความหมายโดยตรงเป็นรูป ผู้ดูต้องใช้ความรู้สึกและจินตนาการ

อารมณ์ที่เกิดจากการกระตุ้นจากรูปนามธรรมนั้นเกิดได้จากการตีความหมายของแต่ละผู้มอง ซึ่งขึ้นอยู่กับพื้นที่ของรูปภาพที่ผู้มองต้องการจะมองด้วย ดังนั้นรูปภาพรูปเดียวกัน ผู้มองคนละคน อาจมีอารมณ์ที่ต่างกัน ดังนั้นอุปกรณ์การติดตามการมองจึงถูกนำมาใช้เพื่อเก็บข้อมูลพฤติกรรมการมองของผู้ร่วมเก็บข้อมูล และนำมาวิเคราะห์

2.3. การป้อนกลับความเกี่ยวข้อง (Relevance Feedback)

เป็นการให้ข้อมูลป้อนกลับสู่ระบบค้นคืน เพื่อประเมินว่าข้อมูลที่ถูกค้นคืนจากระบบนั้นมีความเกี่ยวข้องกับคำสอบถาม (Query) หรือไม่ เพื่อนำไปปรับเปลี่ยนคำสอบถามเพื่อเพิ่มประสิทธิภาพของการค้นคืนข้อมูลที่ต้องการในครั้งต่อไป สามารถแบ่งออกเป็น 2 ประเภท คือ

1) การป้อนกลับโดยชัดเจน เป็นการป้อนกลับของผู้ใช้ระบบที่เกิดขึ้นโดยความตั้งใจ ซึ่งจำเป็นต้องใช้เวลาและแรงงาน เช่น การคลิกเมาส์ การเคลื่อนไหวของตาโดยความตั้งใจ เสียง และ ท่าทาง

2) การป้อนกลับโดยปริยาย เป็นการป้อนกลับของผู้ใช้ระบบที่เกิดจากความไม่ตั้งใจ ซึ่งเกิดโดยอัตโนมัติขณะที่มนุษย์กำลังทำงานใด ๆ อยู่ การป้อนกลับชนิดนี้จำเป็นต้องถูกประมวลผลก่อนนำไปใช้ เช่น อัตราการเต้นของหัวใจ ความดันโลหิต อุณหภูมิของร่างกาย และ การเคลื่อนไหว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของตาโดยความไม่ตั้งใจ (ขณะที่กำลังอ่านหนังสือ หรือ หาข้อมูลบนเว็บเบราว์เซอร์)

2.4. ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine; SVM)

เป็นอัลกอริทึมการเรียนรู้ของเครื่องชนิดหนึ่งที่ได้รับค่านิยมอย่างแพร่หลายในการจำแนกประเภทข้อมูลและถูกนำมาใช้ในการศึกษาอารมณ์ของรูปภาพ [5, 6, 7, 17] SVM จะแบ่งแยกข้อมูลโดยใช้ระนาบหรือไฮเปอร์เพลน (Hyperplane) แบ่งข้อมูลแบบเชิงเส้น

กำหนดให้ x_i คือเวกเตอร์ตัวอย่างที่ i ที่คู่กับค่าเป้าหมาย y_i ซึ่ง $y_i \in \{-1, 1\}$, w คือเวกเตอร์ค่าน้ำหนัก, b คือค่าความโน้มเอียง (Bias), ξ_i คือค่าความผิดพลาดของการจำแนกประเภท, และ C คือตัวแปรเรกิวลารีเซชัน โดยระนาบนั้นจะถูกสร้างขึ้นโดยการจะพยายามทำให้ระยะห่างระหว่างสองกลุ่มนั้นมากที่สุด และ ความผิดพลาดในการแบ่งกลุ่มน้อยที่สุดดังสมการที่ (1)

$$\arg \min_{w, b, \xi} \frac{1}{2} w^2 + C \sum_i \xi_i \quad (1)$$

ซึ่งอยู่ภายใต้ $y_i (w \cdot x_i + b) \geq 1 - \xi_i$ และ $\xi_i \geq 0$ โดยที่มีฟังก์ชันในการตัดสินใจ คือ

$$f(x) = \text{sign}(w^T x + b) \quad (2)$$

ในโลกแห่งความเป็นจริงข้อมูลไม่ได้เป็นเชิงเส้นเสมอไป จึงมีการใช้เทคนิคที่เรียกว่า “เคอร์เนล” (Kernel) โดยใช้ฟังก์ชันเพื่อแปลงจากปริภูมิคุณลักษณะ (Feature Space) ปัจจุบันไปยังปริภูมิคุณลักษณะใหม่ที่สามารถแบ่งข้อมูลด้วยระนาบได้ เช่น Polynomial, Radial Basis Function เป็นต้น

3. การเก็บข้อมูล

ในการทดลอง ได้มีการเก็บข้อมูลอารมณ์ของมนุษย์ที่มีต่อรูปภาพ และการเคลื่อนไหวของตา เพื่อใช้ในการวิเคราะห์จากนักศึกษาระดับปริญญาตรี ซึ่งมีอายุ 18-22 ปี จำนวนทั้งหมด 20 คน แบ่งเป็น ชาย 10 คน หญิง 10 คน คนละ 100 รูป ผู้ร่วมการทดลองมีสายตาที่ปกติ กล่าวคือไม่ได้ใส่แว่นตา

ในการเก็บข้อมูล จะมีโปรแกรมที่พัฒนาขึ้นเพื่อแสดง

รูปภาพพื้นฐานข้อมูล โดยทำงานร่วมกับอุปกรณ์ติดตามการศึกษานาน ไมออนุญาตเห็นไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การมองเห็น โดยจะให้ผู้ร่วมเก็บข้อมูลนั่งประจำเครื่องคอมพิวเตอร์ขนาดพกพา 13 นิ้ว โดยการแสดงผลมีความละเอียดเท่ากับ 1366 x 768 พิกเซล และผู้ควบคุมการเก็บข้อมูลจะอธิบายขั้นตอนและวิธีการเก็บข้อมูลให้กับผู้เข้าร่วม โดยผู้เข้าร่วมเก็บข้อมูลจะบอกอารมณ์ที่เกิดขึ้นจากการมองแต่ละรูปภาพที่แสดงบนหน้าจอ และผู้ควบคุมการทดลองจะบันทึกข้อมูลที่ได้อาศัยคีย์บอร์ด ดังรูปที่ 1.



รูปที่ 1. การเก็บการเคลื่อนไหวของตาโดยเครื่องติดตามการมอง และ อารมณ์ของผู้ร่วมการทดลองที่มีต่อรูปภาพ

ก่อนเริ่มการเก็บข้อมูล มีการแจ้งให้ผู้ร่วมการทดลองพยายามให้ศีรษะอยู่นิ่ง ๆ และจะต้องผ่านการปรับเทียบ (Calibration) ของอุปกรณ์การติดตามการมอง โดยใช้ 9 จุด (กลางหน้าจอ 1 จุด และ 8 จุดที่ขอบหรือมุมหน้าจอ) ดังรูปที่ 2(ก) และก่อนที่จะให้ผู้ร่วมการทดลองดูรูปภาพใหม่ จะมีการแสดงรูปที่มีจุดอ้างอิง (Reference) ตรงกลางหน้าจอเป็นเวลา 3 วินาทีทุกครั้ง ดังรูปที่ 2(ข) เพื่อให้การเคลื่อนไหวเคลื่อนไหวของตาเริ่มต้นที่จุด ๆ เดียวกันทุกครั้ง



รูปที่ 2. ตัวอย่างหน้าจอ (ก) 9-point Calibration (ขวา) จุดอ้างอิงเพื่อให้ผู้ทดสอบเริ่มมองที่จุดเดียวกัน

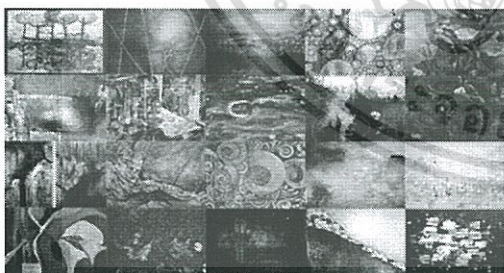
อุปกรณ์การติดตามการมองที่ใช้คือ Eye Tribe Tracker สามารถเชื่อมต่อกับคอมพิวเตอร์หรือแท็บเล็ตด้วย USB3.0 โดยมีอัตราของการเก็บข้อมูลเท่ากับ 30 Hz และมีความแม่นยำที่ 0.5°-1° [18]

งานวิจัยชิ้นนี้ใช้รูปนามธรรมจาก [4] ซึ่งประกอบไปด้วยรูป 228 รูป แบ่งออกเป็น 8 อารมณ์ คือความสนุก (Amusement), ตื่นเต้น (Excitement), พอใจ (Contentment), เสียใจ (Sad), โกรธ (Anger), น่าเกรงขาม (Awe), กลัว (Fear), และขยะแขยง (Disgust) เราจึง

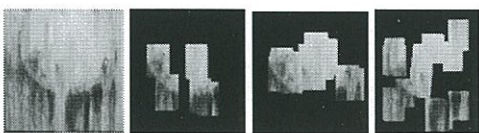
จัดกลุ่มอารมณ์รูปภาพที่มีอารมณ์ใกล้เคียงกันให้เหลือเพียง 4 อารมณ์ เพื่อลดความซับซ้อนให้กับผู้ร่วมเก็บข้อมูลในการตัดสินใจ โดยอ้างอิงจาก Parrot (2001) ซึ่งแบ่งอารมณ์อย่างมีโครงสร้างเป็น 3 ลำดับ คือ Primary Emotions, Secondary Emotions และ Tertiary Emotions [19] โดยรวมรูปภาพที่มีอารมณ์คล้ายคลึงกัน ที่อยู่ในชั้น Primary Emotions เดียวกัน ให้เป็นอารมณ์พื้นฐาน 4 อารมณ์ เช่นเดียวกับ [14] คือ (1) ดีใจ – {สนุก, ตื่นเต้น, พอใจ}, (2) เสียใจ (3) โกรธ (4) กลัว – {กลัว, ขยะแขยง} เนื่องจากรูปที่มีอารมณ์น่าเกรงขามไม่ได้ถูกกล่าวไว้ใน [19] เราจึงไม่ได้สนใจรูปในกลุ่มนี้

จากนั้นจึงเลือกรูปภาพจำนวน 100 รูป โดยเลือกรูปที่มีอารมณ์ทั้ง 4 อารมณ์ อารมณ์ละ 25 รูป วิธีการเลือกรูปภาพนั้น เลือกจากคะแนนสูงสุดของแต่ละอารมณ์จากการโหวตของผู้ใช้ในงานวิจัยชั้นนั้น ตัวอย่างของรูปภาพที่ถูกเลือกมาบางส่วน แสดงดังรูปที่ 3.

จากการเก็บข้อมูลนั้นพบว่ารูปเดียวกัน ดังรูปที่ 4(ก) คนมองต่างกัน อารมณ์ที่ได้ต่างกันเช่นกัน จากการมองรูปภาพของ ผู้เข้าร่วม-1 ให้ผลอารมณ์รูปภาพ คือ กลัว ดังรูปที่ 4(ข) ผู้เข้าร่วม-2 ให้ผลอารมณ์รูปภาพ คือ โกรธ ดังรูปที่ 4(ค) และ ผู้เข้าร่วม-3 ให้ผลอารมณ์รูปภาพ คือ เสียใจ ดังรูปที่ 4(ง)



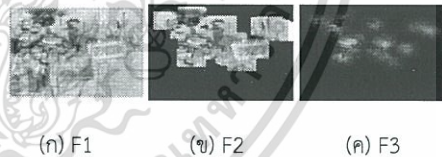
รูปที่ 3. ตัวอย่างรูปนามธรรมที่ถูกเลือกจากงานวิจัย



(ก) รูปต้นฉบับ (ข) กลัว (ค) โกรธ (ง) เสียใจ
รูปที่ 4. ตัวอย่างรูปภาพ และส่วนที่ถูกมองของผู้เข้าร่วมที่มีอารมณ์แตกต่างกัน

4. การสกัดคุณลักษณะ

ในงานวิจัยนี้เราสกัดคุณลักษณะของรูปภาพโดยใช้การแจกแจงความถี่ของคุณลักษณะเด่นสี (Color) รูปร่าง (Shape) และ พื้นผิว (Texture) ของรูปภาพ โดยชนิดของคุณลักษณะเด่นที่นิยมใช้ของสีคือ RGB ซึ่งจะแสดงปริมาณความเข้มของสีแดง เขียว และน้ำเงินในแต่ละพิกเซล โดยมีค่าตั้งแต่ 0-255 ส่วนรูปร่างถูกอธิบายโดยใช้โซเบล (Sobel) ที่ใช้แสดงขนาดทิศทางของเส้นขอบของรูปภาพ ขณะที่ขนาดและทิศทางของพื้นผิวได้ถูกอธิบายโดยกาบอร์ (Gabor) ซึ่งทั่วไปมีทั้งหมด 5 ขนาด 8 ทิศทาง แต่ในงานวิจัยครั้งนี้เราใช้เพียง 1 ขนาด 8 ทิศทาง ที่เป็น high-pass filter นอกจากนี้เพื่อเพิ่มประสิทธิภาพในการทำนายอารมณ์ จึงมีการรวมคุณลักษณะเหล่านี้เข้าด้วยกัน โดยการนำเวกเตอร์คุณลักษณะดังกล่าวมาต่อกัน คุณลักษณะเหล่านี้ได้ถูกนำมาใช้อธิบายรูปภาพทั้ง 3 ชนิด ได้แก่ (ก) รูปภาพต้นฉบับ (Original Image) (ข) รูปภาพที่มีข้อมูลการมอง (Original Image + Eye Movements) และ (ค) รูปภาพที่มีการมองโดยการเบลอด้วยฟังก์ชันเกาส์เซียน (Original Image + Eye Movements + Gaussian Blur) ดังรูปที่ 5.



รูปที่ 5. ตัวอย่างรูปภาพ (ก) Original Image (F1) (ข) Original Image + Eye Movements (F2) และ (ค) Original Image + Eye Movements + Gaussian Blur (F3)

4.1 รูปต้นฉบับ

คุณลักษณะของแต่ละรูปถูกสกัดโดยใช้การแจกแจงความถี่ของสี RGB, โซเบล และ กาบอร์ ของ โดยทดลองใช้ขนาดของช่วงต่างกัน นั่นคือ {8, 16, 32, 64}-bin ในการทดลอง โดยที่ขนาดของคุณลักษณะของสีที่ได้จะเท่ากับ

$$n_{\text{feature}} = n_{\text{bin}}^3 \quad (3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขณะที่ขนาดของคุณลักษณะของการแจกแจงความถี่ของรูปร่างและพื้นผิวของรูปจะเท่ากับ

$$n_{\text{feature}} = n_{\text{bin}} \quad (4)$$

4.2 รูปภาพที่มีข้อมูลการมอง

ภาพในส่วนนี้จะขึ้นอยู่กับตำแหน่งการมองของแต่ละผู้ร่วมเก็บข้อมูล โดยใช้ฟังก์ชันความน่าจะเป็นการกระจายของตัวแปรสุ่มเกาส์เซียนที่มีขนาดต่าง ๆ คือ 100×100 , 125×125 และ 150×150 กำหนดให้ $I(x, y)$ คือ รูปภาพ โดยที่ (x, y) เป็นพิกัดของรูปภาพ และ $E(x, y)$ คือ ข้อมูลการมองของรูปภาพ จากนั้นจึงใช้ฟังก์ชันเกาส์เซียนกับข้อมูลการมองของรูปภาพจาก

$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \quad (5)$$

โดยที่ σ คือ ส่วนเบี่ยงเบนมาตรฐานของการกระจายเกาส์เซียน จะได้ข้อมูลการมองของรูปภาพที่มีการกระจายตัวแบบเกาส์เซียน $G(x, y)$ จากนั้นจึงนำไปแทนในฟังก์ชัน (5) เพื่อสร้างตัวกรองสำหรับตำแหน่งรูปภาพที่ถูกผู้ใช้มอง

$$H(x, y) = \begin{cases} 1 & \text{if } G(x, y) > 0 \\ 0 & \text{if } G(x, y) = 0 \end{cases} \quad (6)$$

ดังนั้นรูปที่ผ่านตัวกรองคือ

$$I_{F2}(x, y) = H(x, y) \times I(x, y) \quad (7)$$

4.3 รูปภาพที่มีการมองโดยการเบลอด้วยฟังก์ชันเกาส์เซียน

ในส่วนนี้จะมีการสร้างเช่นเดียวกับรูปภาพในหัวข้อ 4.2 โดยใช้ $G(x, y)$ เป็นตัวกรองในการสร้างรูปใหม่

$$I_{F3}(x, y) = G(x, y) \times I(x, y) \quad (8)$$

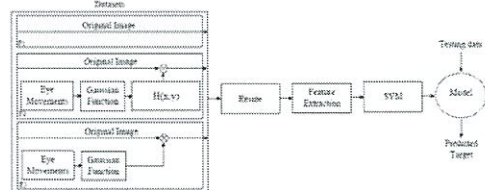
$G(x, y)$ นั้นจะมีค่าตั้งแต่ 0 ถึง 1 หากค่าของตัวกรองตำแหน่งนั้นเข้าใกล้ศูนย์ ตำแหน่งภาพนั้นจะมีมืด แต่ถ้าค่าของตัวกรองตำแหน่งนั้นเข้าใกล้ 1 ตำแหน่งรูปภาพนั้นจะสว่าง

5. การทดลอง

ในงานนี้เรามีการทดลองทั้งหมด 2 การทดลอง เพื่อเปรียบเทียบคุณลักษณะทั้ง 3 ชนิด คือการแจกแจงความถี่ของสี RGB โซเบล และกาบอร์ จากชุดรูปภาพทั้ง 3 ชนิด คือ รูปภาพต้นฉบับ รูปภาพที่มีข้อมูลการมอง และรูปภาพที่มีการมองโดยการเบลอด้วยฟังก์ชันเกาส์เซียน

การทดลองแรกเป็นโมเดลของผู้ใช้แต่ละคน จำนวน 20 คน กล่าวคือ เป็นการทำนายอารมณ์จากข้อมูลของแต่ละผู้ร่วมเก็บข้อมูลจำนวน 20 โมเดล โดยทดลองเปลี่ยนค่าจำนวนช่วงของการแจกแจงความถี่แต่ละคุณลักษณะ เป็น 8, 16, 32 และ 64 บิน และ ขนาดของเกาส์เซียน คือ 100×100 , 125×125 และ 150×150 จากนั้นข้อมูลได้ถูกเรียนรู้และทดสอบโดยใช้ SVM พร้อมกับ Linear Kernel โดยมีพารามิเตอร์ที่ต้องการปรับค่าเพียงตัวเดียวสำหรับ SVM นั่นคือ C ซึ่งมีค่าตั้งแต่ 10^{-6} – 10^4 เพื่อให้ได้ผลการทำนายที่ดีที่สุด ในการทดลองนี้เราได้ใช้เทคนิค Leave-one-out Cross-validation (LOO-CV) ในการทดสอบ ส่วนการทดลองที่สอง เป็นการทดลองโมเดลรวมทุกผู้ใช้ ซึ่งเป็นการทำนายอารมณ์จากข้อมูลผู้ร่วมเก็บข้อมูลทั้งหมด โดยใช้เงื่อนไขการทดลองเช่นเดียวกับการทดลองแรก ต่างกันที่ใช้เทคนิค Leave-one-user-out Cross-validation กล่าวคือ การทดสอบโมเดลกับผู้ใช้ใหม่

เนื่องจากข้อมูลประกอบไปด้วย 4 ประเภท คือ ดีใจ, เสียใจ, โกรธ และกลัว เราจึงเลือกการแบ่งกลุ่มข้อมูลโดยใช้ Multiclass Classification ชนิด 1-vs-all ข้อมูลที่ได้ถูกนำมา نرمัลไลเซชัน (Normalization) ให้มีค่าเฉลี่ยเท่ากับ ศูนย์และค่าเบี่ยงเบนมาตรฐานเท่ากับหนึ่ง กระบวนการดำเนินงานสามารถดูได้ดังรูปที่ 6



รูปที่ 6. ขั้นตอนการดำเนินงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. ผลการทดลอง

6.1 โมเดลการทำนายอารมณ์ของแต่ละผู้ร่วมเก็บข้อมูล (User Model)

ประสิทธิภาพของการจำแนกอารมณ์ได้ถูกรายงานโดยใช้ค่าร้อยละความถูกต้องเฉลี่ยของโมเดลผู้ใช้ทั้ง 20 คน

เนื่องจากอารมณ์ทั้ง 4 อารมณ์ของผู้ร่วมการทดลองนั้นไม่สมดุลกัน (Unbalanced Data) คุณลักษณะทั้ง 3 ประเภท และคุณลักษณะที่เกิดจากการรวมอีก 4 ประเภท บนชุดข้อมูลทั้ง 3 ชุด ที่นำเสนอในหัวข้อที่ 4 ได้ถูกเปรียบเทียบกับกรณี SVM ทำการทำนายได้เพียงอารมณ์เดียว ในที่นี้เรากำหนดให้เป็น Baseline (BL)

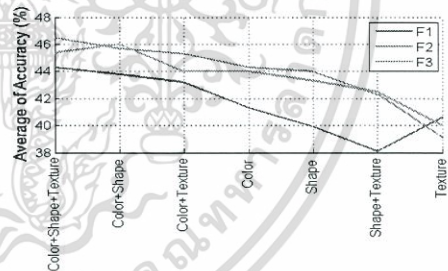
ตารางที่ 1. ค่าเฉลี่ยการทำนายอารมณ์ทั้ง 20 คนในแต่ละคุณลักษณะ

Feature	ความถูกต้องเฉลี่ย (%)			
	BL	F1	F2	F3
Color		43.35	45.85	45.70
Shape		41.95	45.80	44.80
Texture		43.15	42.20	42.15
Color+Shape	36.35	46.60	47.20	46.95
Color+Texture		45.65	47.60	45.65
Shape+Texture		41.40	44.60	44.45
Color+Shape+Texture		47.10	48.05	47.20
Average	36.35	44.17	45.90	45.26

จากตารางที่ 1 จะเห็นได้ว่าค่าเฉลี่ยความถูกต้องจากการทำนายโดยใช้คุณลักษณะพื้นฐาน F1 F2 และ F3 นั้น ต่ำกว่า BL และ การใช้ F2 และ F3 นั้นมีประสิทธิภาพดีกว่า F1 โดยเฉลี่ย ยกเว้นการใช้ลักษณะเด่นพื้นผิวที่ F1 ที่มีค่าเฉลี่ยผลทำนายดีกว่า F2 และ F3 นอกจากนี้การรวมคุณลักษณะยังสามารถเพิ่มประสิทธิภาพการทำนายได้มากขึ้นอีกด้วย ในการทดลองนี้เราได้รวมคุณลักษณะโดยกำหนดให้จำนวนbinของแต่ละคุณลักษณะเท่ากัน และทดลองเปลี่ยนค่าตั้งแต่ {8,16,32,64}-bin จากตารางที่ 1 จะเห็นได้ว่าการรวมคุณลักษณะเพื่อทำนายอารมณ์โดยรวมทั้ง 3 คุณลักษณะมีผลค่าเฉลี่ยการทำนายอารมณ์ดีที่สุดเท่ากับร้อยละ 48.05% รองลงมาเป็นการทำนายโดยรวม 2 คุณลักษณะ ซึ่งผลที่ดีที่สุดคือ การรวมลักษณะสีกับพื้นผิว

และสีกับรูปร่างเท่ากับร้อยละ 47.60% และ 47.20% ตามลำดับ แต่ในการรวมคุณลักษณะรูปร่างและพื้นผิวให้ผลการทำนายอารมณ์ได้น้อยกว่าการทำนายโดยใช้คุณลักษณะสีเพียงอย่างเดียว ดังนั้นจากผลการทดลองจะสังเกตได้ว่าการรวมคุณลักษณะที่มีส่วนอยู่ด้วยนั้นจะมีประสิทธิภาพสูงกว่าการไม่มีสีอยู่ ทำให้สันนิษฐานได้ว่าคุณลักษณะสีมีผลต่อการทำนายอารมณ์รูปภาพได้ดีที่สุด และการทำนายอารมณ์รูปโดยมีการเคลื่อนไหวของดวงตาของผู้ร่วมเก็บข้อมูล มีประสิทธิภาพการทำนายแม่นยำกว่าการใช้ข้อมูลรูปภาพเพียงอย่างเดียว

จากนั้นเราได้แสดงค่าเฉลี่ยของ True Positive (TP) ของอารมณ์ทั้งหมดในแต่ละคุณลักษณะโดยเรียงจากคุณลักษณะที่มีค่าเฉลี่ยของ TP ดีที่สุดไปยังน้อยที่สุด ซึ่งค่า TP คือ ค่าที่แสดงถึงกรณีที่ทำนายออกมาตรงกับอารมณ์เป้าหมายที่เก็บจากผู้เข้าร่วมเก็บข้อมูล จะเห็นได้ว่าผลการทดลองนั้นมีลักษณะเหมือนกับผลข้างต้น คือ ในทุกคุณลักษณะ F2 มีผลการทำนาย ดีกว่า F1 ในทุกกรณี ยกเว้นคุณลักษณะพื้นผิว



รูปที่ 7. ค่าเฉลี่ย True Positive ของอารมณ์ในแต่ละคุณลักษณะ

6.2 โมเดลสำหรับผู้ใช้ทุกคน (Global Model - New User)

จากการทดลองแรกพบว่า F2 มีผลการทำนายดีกว่า F3 อาจเป็นผลมาจาก F3 มีการสูญเสียของข้อมูลเนื่องจากการเบลอ อีกทั้งการทำนายโดยใช้ทั้ง 3 คุณลักษณะ มีประสิทธิภาพมากที่สุด ดังนั้นการทดลองนี้จึงเลือกใช้ F2 และการรวมคุณลักษณะทั้ง 3 ชนิด โดยใช้ข้อมูลของผู้เก็บข้อมูลทั้งหมด 10 คน จากการสุ่มและใช้เทคนิคการแบ่งข้อมูลแบบ Leave-one-user-out Cross-validation ในการทดสอบ

ตารางที่ 2. ค่าเฉลี่ยรวมแต่ละชุดข้อมูล

User	ความถูกต้องเฉลี่ย (%)		
	BL	F1	F2
1	46.00	48.00	50.00
2	28.00	51.00	55.00
3	35.00	59.00	65.00
4	49.00	47.00	57.00
5	33.00	49.00	48.00
6	46.00	62.00	56.00
7	36.00	55.00	55.00
8	38.00	52.00	53.00
9	34.00	55.00	61.00
10	52.00	56.00	57.00
Average	39.70	53.40	55.70

เมื่อดูค่าเฉลี่ยของการทำนายของ Global Model จากตารางที่ 2. และของ User Model จากตารางที่ 1. จะพบว่า ค่าเฉลี่ยของ Global Model นั้นสูงกว่า ทั้งนี้เนื่องจากจำนวนข้อมูลที่ใช้ในการสอนของ Global Model นั้นเท่ากับ 900 รูปภาพ ซึ่งมากกว่า User Model ที่ใช้เพียง 99 รูปภาพในการสอนเท่านั้น ตารางที่ 3. ได้แสดงว่าประสิทธิภาพในการทำนายโดยใช้ F1 และ F2 นั้นดีกว่า BL อีกทั้งค่าเฉลี่ยรวมของ F2 ยังสูงกว่า F1 อีกด้วย นอกจากนี้พบว่า ส่วนใหญ่ F2 มีผลความถูกต้องเฉลี่ยดีกว่า F1 เท่ากับร้อยละ 55.70 % ยกเว้น ผู้ร่วมเก็บข้อมูลที่ 5 และ 6 อาจเป็นเพราะว่าพฤติกรรมอารมณ์ และการเลือกอารมณ์ของกลุ่มนี้แตกต่างจากผู้ร่วมเก็บข้อมูลคนอื่น ๆ ทำให้ Global Model นั้นไม่มีประสิทธิภาพสำหรับกลุ่มนี้

ตารางที่ 3. แสดงค่าเฉลี่ยการทำนายในแต่ละbin และใช้ Pair t-test ในการทดสอบค่านัยสำคัญซึ่งแสดงอยู่ในตารางที่ 4. เพื่อเปรียบเทียบประสิทธิภาพของชุดข้อมูลทั้ง 3 ชุด

ตารางที่ 3. ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของผลการทำนายอารมณ์

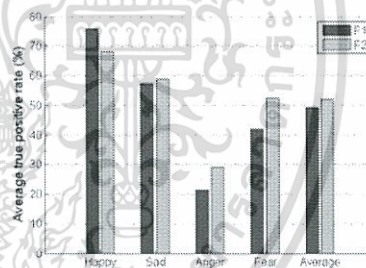
nbin	ความถูกต้องเฉลี่ย (%)		
	BL	F1	F2
8	39.70±7.96	51.90±4.23	52.00±6.13
16		52.50±5.46	52.60±5.97
32		51.30±5.14	55.00±5.35
64		50.70±5.14	53.80±4.76

ตารางที่ 4. ค่า p-value จากการทำ Pair t-test ของประสิทธิภาพในการทำนายอารมณ์ ทั้ง 3คุณลักษณะ

nbin	p-Value		
	F1-vs-BL	F2-vs-BL	F1-vs-F2
8	0.001	0.001	0.955
16	0.002	0.001	0.949
32	0.005	0.001	0.034
64	0.006	0.001	0.001

จากตารางที่ 3. พบว่า F1, F2 ให้ประสิทธิภาพที่ดีกว่า BL ในทุกกรณี ($p < 0.01$ ดังตารางที่ 4) และ F2 มีผลการทำนายดีกว่า F1 ในทุกกรณีเช่นเดียวกัน มีเพียงกรณี {32, 64}-bin ที่มีนัยสำคัญทางสถิติที่ระดับ $p < 0.05$

จากรูปที่ 8. แสดงค่าเฉลี่ยของ True Positive ของแต่ละอารมณ์และเฉลี่ยทุกอารมณ์ของทั้ง F1 และ F2 จะเห็นได้ว่า F2 มีประสิทธิภาพดีกว่า F1 ในอารมณ์เศร้า, โกรธ, และกลัว นอกจากนี้ในภาพรวมนั้น F2 ให้ประสิทธิภาพที่ดีที่สุด



รูปที่ 8. ค่าเฉลี่ย True Positive ของการทำนายอารมณ์ในแต่ละอารมณ์จากการรวมคุณลักษณะทั้ง 3

7. สรุปผลการทดลอง

บทความนี้นำเสนอการจำแนกประเภทของรูปภาพเชิงอารมณ์ คือ ดีใจ เสียใจ โกรธ และกลัว โดยใช้ข้อมูล (ก) การเคลื่อนไหวของตาที่มีต่อรูปภาพ ร่วมกับ (ข) คุณลักษณะพื้นฐาน คือ สี, รูปร่าง และพื้นผิว จากการทดลองทำนายอารมณ์รูปภาพโดยใช้ลักษณะเด่นของภาพร่วมกับข้อมูลการเคลื่อนไหวดวงตาเพื่อเปรียบเทียบกับการใช้ลักษณะเด่นของภาพเพียงอย่างเดียว อีกทั้งยังพบว่าการรวมคุณลักษณะทั้ง 3 เข้าด้วยกันแบบพื้นฐาน ยังเพิ่มประสิทธิภาพการทำนายได้ โดยคุณลักษณะสีมีผลในการช่วยให้การทำนายนั้นมีประสิทธิภาพเพิ่มขึ้น นอกจากนี้การ

เพิ่มจำนวนข้อมูลในการทดสอบ (Global Model) สามารถเพิ่มประสิทธิภาพในการทำนาย แต่พฤติกรรมของผู้ใช้นั้น ๆ ควรจะต้องคล้ายกัน มิฉะนั้นประสิทธิภาพจะลดลง นั่นคือ การโมเดลของผู้ใช้ (User-modelling) และการปรับสภาพตามผู้ใช้ (User-adaptation) เป็นสิ่งที่ควรจะต้องคำนึงต่อไปในอนาคต

เอกสารอ้างอิง

- [1] L. Ye, P. Ogunbona and J. Wang "Image Content Annotation Based on Visual Features", In: Proceeding of International Symposium on Multimedia (ISM'2006), 11-13 Dec 2006, San Diego, USA, pp. 62-69, 2006.
- [2] J. Laaksonen, M. Koskela, and E. Oja "PicSOM Self-organizing Image Retrieval with MPEG-7 Content Descriptors", IEEE Transactions on Neural Networks, 13(4), pp. 841-853, 2002.
- [3] R. Datta, J. Li, and J. Z. Wang, "Content-based Image Retrieval - Approaches and Trends of the New Age", In: Proceedings of ACM International Workshop on Multimedia Information Retrieval (MIR'2005), 10-11 Nov 2005, Singapore, pp. 253-262, 2005.
- [4] W. Weining, Y. Yinlin, and J. Shengming, "Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction", In: Proceeding of IEEE International Conference on Systems, Man and Cybernetics (SMC'2006), 8-11 Oct 2006, Taipei, Taiwan, pp. 3534-3539, 2006.
- [5] J. Machajdik, and A. Hanbury, "Affective Image Classification Using Features Inspired by Psychology and Art Theory," In: Proceeding of ACM International Conference on Multimedia (MM'2010), 25-29 Oct 2010, Firenze, Italy, pp. 83-92, 2010.
- [6] H. Zhang, E. Augilius, T. Honkela, J. Laaksonen, H. Gamper, and H. Alene, "Analyzing Emotional Semantics of Abstract Art Using Low-level Image Features", In: Proceeding of International Symposium on Intelligent Data Analysis (IDA'2011), 29-31 Oct 2011, Porto, Portugal, pp. 413-423, 2011.
- [7] H. Zhang, M. Gonen, Z. Yang and E. Oja, "Predicting Emotional States of Images Using Bayesian Multiple Kernel Learning", In: Proceeding of International Conference on Neural Information Processing (ICONIP'2013), 3-7 Nov 2013, Daegu, Korea, pp. 274-282, 2013.
- [8] K. Pasupa, C. J. Saunders, S. Szedmak, A. Klami, S. Kaski, and S. R. Gunn, "Learning to Rank Images from Eye movements", In: Proceeding of 2009 IEEE 12th International Conference on Computer Vision (ICCV'2009) Workshops on Human-Computer Interaction (HCI'2009), 27 Sep-4 Oct 2009, Kyoto, Japan, pp. 2009-2016, 2009.
- [9] D. R. Hardoon and K. Pasupa "Image Ranking with Implicit Feedback from Eye Movements," In: Proceedings of the 6th Biennial Symposium on Eye Tracking Research & Applications (ETRA'2010), 22-24 Mar 2010, Austin, USA, pp. 291-298, 2010.
- [10] P. Auer, Z. Hussain, S. Kaski, A. Klami, J. Kujala, J. Laaksonen, A. P. Leung, K. Pasupa, and J. Shawe-Taylor "Pinview: Implicit Feedback in Content-Based Image Retrieval, In: Proceeding of Workshop on Applications of Pattern Analysis (WAPA'2010), 1-2 Sep 2010, Cumberland Lodge, UK, pp 51-57, 2010.
- [11] C. E. Izard "Basic Emotions, Relations Among Emotions, and Emotion-Cognition Relations", Psychological Review, 99(3), pp. 561-565, 1992.
- [12] K. Vytal and S. Hamann "Neuroimaging Support for Discrete Neural Correlates of Basic Emotions: A Voxel-based Meta-analysis," Cognitive Neuroscience, 22(12), pp. 2864-2885, 2010.
- [13] P. Ekman "Universals and Cultural Differences in Facial Expressions of Emotion," Nebraska Symposium on Motivation, 19, pp. 207-282, 1972.
- [14] R. E. Jack, O. G.B. Garrod, and P. G. Schyns "Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals Over Time", Current Biology, 24(2), pp. 187-192, 2014.
- [15] R. Plutchik and H. Kellerman, "Emotion: Theory, Research and Experience," Psychological Medicine, 11(1), pp. 207, 1980.
- [16] D. W. Galenson "Two Paths to Abstract Art Kandinsky and Malevich", Technical Report, National Bureau of Economic Research, No. 12403, 2006.
- [17] Y. Wu, C. Bauckhage and C. Thurau "The Good, the Bad, and the Ugly: Predicting Aesthetic Image Labels", In: Proceedings of 20th International Conference on Pattern Recognition (ICPR'2010). 23-26 Aug 2010, Istanbul, Turkey. pp. 1586-1589, 2010.
- [18] The Eye Tribe Aps, The Eye Tribe, Available at <https://theeyetribe.com>.
- [19] Changing Minds, Basic Emotion, Available at <http://changingminds.org/explanations/emotions/basic%20emotions.htm>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้