



การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์

โดย

นางสาวอรรวรรณ เลียบศิริ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
สาขาวิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
ปีการศึกษา 2559
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์

โดย

นางสาวอรรณณ เลียบศิริ



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

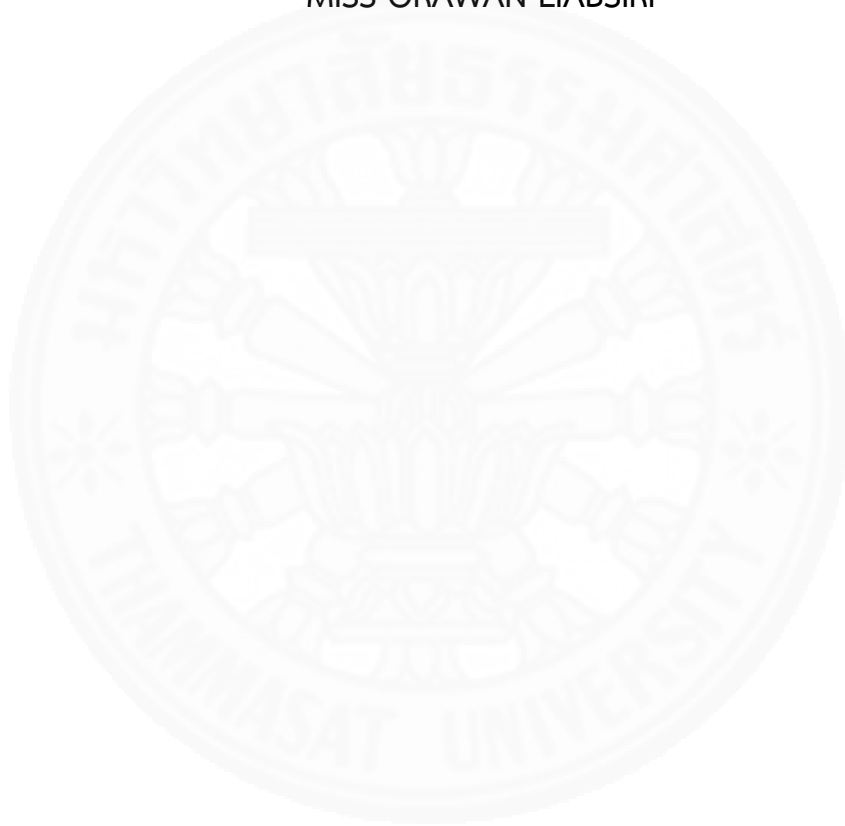
ปีการศึกษา 2559

ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

CLASSIFICATION OF FACIAL EXPRESSIONS OF CARTOONS USING
HISTOGRAM OF ORIENTED GRADIENT

BY

MISS ORAWAN LIABSIRI



A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE (COMPUTER SCIENCE)

DEPARTMENT OF COMPUTER SCIENCE
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY

ACADEMIC YEAR 2016

COPYRIGHT OF THAMMASAT UNIVERSITY

มหาวิทยาลัยธรรมศาสตร์
คณะวิทยาศาสตร์และเทคโนโลยี

วิทยานิพนธ์

ของ

นางสาวอรรณณ เลียบศิริ

เรื่อง

การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

เมื่อ วันที่ 31 กรกฎาคม พ.ศ. 2560

ประธานกรรมการสอบวิทยานิพนธ์



(ผู้ช่วยศาสตราจารย์ ดร. ณัฐนันท์ หงส์วิทธิธร)

กรรมการและอาจารย์ที่ปรึกษาวิทยานิพนธ์



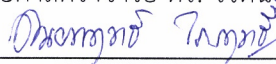
(ผู้ช่วยศาสตราจารย์ ดร. เสาวลักษณ์ วรรณภา)

กรรมการสอบวิทยานิพนธ์



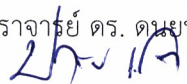
(ผู้ช่วยศาสตราจารย์ ดร. วิรัตน์ จาริวงศ์ไพบูลย์)

กรรมการสอบวิทยานิพนธ์



(รองศาสตราจารย์ ดร. ดนุพฤกษ์ ไกรฤทธิ)

คณบดี



(รองศาสตราจารย์ ปกรณ์ เสริมสุข)

หัวข้อวิทยานิพนธ์	การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์
ชื่อผู้เขียน	นางสาวอรรณณ เลียบศิริ
ชื่อปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
สาขาวิชา/คณะ/มหาวิทยาลัย	สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร. เสาวลักษณ์ วรรณภา
ปีการศึกษา	2559

บทคัดย่อ

งานวิจัยนี้นำเสนอวิธีการจำแนกอารมณ์จากการแสดงออกทางใบหน้าของตัวการ์ตูนแอนิเมชันเป็นสามกลุ่มอารมณ์ ได้แก่ โกรธ มีความสุข และ เศร้า โดยเริ่มจากการตรวจจับใบหน้าการ์ตูนในภาพด้วยเทคนิคฮาร์ไลค์ฟีเจอร์ (Haar-like Feature) แล้วใช้เทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient: HOG) ในการสกัดคุณลักษณะเฉพาะของภาพใบหน้า จากนั้นนำไปจำแนกอารมณ์ด้วยตัวจำแนก โดยในงานวิจัยนี้ทำการเปรียบเทียบประสิทธิภาพของตัวจำแนกสามแบบ ได้แก่ (1) ต้นไม้ตัดสินใจ (Decision - Tree: DT) (2) ตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors: k -NN) และ (3) ตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

ชุดข้อมูลที่ใช้ในการทดลองประกอบด้วยภาพการ์ตูนที่เห็นส่วนประกอบบริเวณใบหน้าตัวการ์ตูนชัดเจนและแสดงสามอารมณ์หลัก รวม 300 ภาพ จากแอนิเมชัน 20 เรื่อง การทดลองจะแบ่งออกเป็นสี่การทดลองตามตำแหน่งขององค์ประกอบใบหน้าที่จะนำมาใช้สกัดคุณลักษณะ HOG ได้แก่ (1) ทั้งใบหน้า (2) คิ้ว ตา ปาก (3) คิ้ว ปาก และ (4) ตา ปาก ซึ่งผลการทดลองแสดงให้เห็นว่าตัวจำแนกแบบ DT ให้ค่าความถูกต้องร้อยละ 59.0, 68.3, 68.7 และ 63.7 ตามลำดับ ตัวจำแนกแบบ k -NN โดยกำหนดให้ $k=3$ ให้ค่าความถูกต้องร้อยละ 72.3, 83.3, 85.3 และ 72.0 ตามลำดับ และตัวจำแนกแบบ SVM ให้ค่าความถูกต้องสูงสุดที่ร้อยละ 88.7, 93.0, 90.3 และ 80.0 ตามลำดับ

คำสำคัญ: การ์ตูนแอนิเมชัน, ฮาร์ไลค์ฟีเจอร์, การสกัดคุณลักษณะเฉพาะ, ค่าความถี่ของทิศทางตามค่า เกรเดียนท์.

Thesis Title	CLASSIFICATION OF FACIAL EXPRESSIONS OF CARTOONS USING HISTOGRAM OF ORIENTED GRADIENT
Author	Miss Orawan Liabsiri
Degree	Master of Science (Computer Science) Computer Science
Major Field/Faculty/University	Faculty of Science and Technology Thammasat University
Thesis Advisor	Asst. Prof. Dr. Saowaluk Watanapa
Academic Years	2016

ABSTRACT

This research presents techniques to classify facial expressions of animation cartoons. Into 3 main classes, which are anger, happiness and sadness. Firstly, we detect face of cartoons by Haar-like Feature technique. After that, the histogram of oriented gradient (HOG) technique is used to extract features. Then, the decision tree (DT) classifier, the k- nearest neighbors (k-NN) classifier and the support vector machine (SVM) were applied to classify the facial expressions of cartoons.

The data set used in the experiments consists of 300 images from 20 animated movies. Each image presents a face with clear expression. Four experiments are conducted based on how to extract HOG Feature from (1) the whole face (2) eyebrow-eye-mouth (3) eyebrow-mouth and (4) eye-mouth. From the experimental results, DT yields 59.0%, 68.3%, 70.3% and 63.7%. k-NN yields 72.3%, 83.3%, 86.0% and 72.0%. And SVM yields 88.7%, 93.0%, 90.0% and 80.0% accuracy.

Keywords: Cartoon Animation, Haar-like Feature, Feature Extraction, Histogram of oriented gradient.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความอนุเคราะห์ของบุคคลหลายท่าน อาจจะนำมากล่าวถึงได้ไม่ทั้งหมด บุคคลแรกที่ขอขอบพระคุณ คือ ผศ.ดร. เสาวลักษณ์ วรรณภา อาจารย์ที่ปรึกษา ที่ให้คำแนะนำตรวจทานและแก้ไขข้อบกพร่องของวิทยานิพนธ์ด้วยความดูแลและเอาใจใส่ในทุกขั้นตอน เพื่อให้วิทยานิพนธ์ฉบับนี้ มีความสมบูรณ์แบบมากที่สุด

ขอขอบคุณ ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยธรรมศาสตร์ที่อำนวยความสะดวกและช่วยเหลือในเรื่องต่าง ๆ ของการจัดทำวิทยานิพนธ์ฉบับนี้

ขอขอบคุณ เพื่อนรหัส 55 ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยธรรมศาสตร์ ที่คอยให้กำลังใจ ให้คำปรึกษาและคอยช่วยเหลือในเรื่องต่าง ๆ ตลอดมา

ขอขอบคุณ พี่อ่า หัวหน้า และพี่โรจน์ที่คอยเข้าใจและอำนวยความสะดวกในทุกเรื่องเกี่ยวกับการทำวิทยานิพนธ์ฉบับนี้มาตลอด

ขอขอบคุณน้องนุก น้องสาวที่แสนดีและน่ารัก ที่คอยช่วยเหลือทุก ๆ ด้าน ทั้งการจัดเตรียมข้อมูลและการให้กำลังใจ เพื่อให้การจัดทำวิทยานิพนธ์ฉบับนี้ ผ่านลุล่วงไปได้

ขอขอบคุณน้องบอม น้องสาวที่คอยสนับสนุน ในการจัดหาข้อมูล อีกทั้งยังคอยให้คำปรึกษาและคอยกำลังใจอยู่เสมอ

ขอขอบคุณ อีฟ แพร่า และรุ่ง ที่คอยให้กำลังใจตลอดระยะเวลาที่ได้ทำการจัดทำวิทยานิพนธ์ฉบับนี้

สุดท้าย ขอขอบพระคุณคุณพ่อ คุณแม่ พี่สาวและน้องสาว สมาชิกครอบครัวเลียบศิริ ที่เป็นเบื้องหลังความสำเร็จในครั้งนี้ ที่คอยให้กำลังใจและสนับสนุนในทุก ๆ ด้าน เพื่อให้การจัดทำวิทยานิพนธ์ประสบความสำเร็จ

นางสาวอรรวรรณ เลียบศิริ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	(1)
บทคัดย่อภาษาอังกฤษ	(2)
กิตติกรรมประกาศ	(3)
สารบัญตาราง	(8)
สารบัญภาพ	(11)
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของงานวิจัย	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 คำจำกัดความของคำศัพท์เฉพาะ	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	5
2.1 ความรู้พื้นฐานเกี่ยวกับอารมณ์ (Emotion Basic Knowledge)	5
2.1.1 การจำแนกอารมณ์จากการแสดงอารมณ์บนใบหน้า (Classification of emotion from facial expressions)	5
2.1.2.1 อารมณ์สับสน (confusion)	6
2.1.2.2 อารมณ์อับอาย (Shame)	6

2.1.2.3	อารมณ์แปลกใจ (Surprise)	7
2.1.2.4	อารมณ์มุ่งมั่น (Focus)	8
2.1.2.5	อารมณ์อ่อนเพลีย (Exhaustion)	8
2.1.2.6	อารมณ์ยั่วชวน (Seduction)	9
2.1.2.7	อารมณ์โกรธ (Anger)	9
2.1.2.8	อารมณ์กลัว (Fear)	10
2.1.2.9	อารมณ์เศร้า (Sadness)	11
2.1.2.10	อารมณ์มีความสุข (Happiness)	11
2.2	ความรู้พื้นฐานเกี่ยวกับการ์ตูน (Cartoon Basic Knowledge)	12
2.2.1	ประเภทของการ์ตูน	12
2.2.1.1	ประเภทของการ์ตูนที่แบ่งตามลักษณะในการเคลื่อนไหว	12
2.2.1.2	ประเภทของการ์ตูนที่แบ่งตามอรรถบทยที่พบเห็นในประเทศไทย	12
2.2.2	การวาดภาพใบหน้าตัวการ์ตูนขั้นพื้นฐาน	13
2.2.3	เทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูน	15
2.3	ความรู้พื้นฐานเกี่ยวกับการประมวลผลและวิเคราะห์ภาพดิจิทัล (Digital Image Processing & Analysis Fundamental)	23
2.3.1	รูปภาพดิจิทัล (Digital Images)	23
2.3.2	ประเภทของภาพ (Image Types)	25
2.3.3	แบบจำลองสี (Color Model)	28
2.3.4	การประมวลผลและวิเคราะห์ภาพ (Image Processing and Analysis)	33
2.4	การวิเคราะห์ภาพด้วยคุณสมบัติการแยกลักษณะเด่นของวัตถุ (Feature Extraction for Image Analysis)	34
2.4.1	คุณลักษณะเกี่ยวกับสี (Color - Based Features)	35
2.4.2	คุณลักษณะเกี่ยวกับพื้นผิว (Texture - Based Features)	35
2.4.3	คุณลักษณะเกี่ยวกับรูปร่าง (Shape - Based Features)	35
2.4.4	ค่าความถี่ของทิศทางตามค่าเกรเดียนต์ (Histograms of Oriented Gradients: HOG)	36
2.4.5	ลักษณะเฉพาะที่ดี (Good Features)	39

2.5 การจำแนกประเภทของข้อมูล (Classification)	39
2.5.1 ตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด K ตัว (K – Nearest Neighbors)	40
2.5.2 ตัวจำแนกแบบต้นไม้การตัดสินใจ (Decision Tree)	41
2.5.3 ตัวจำแนกแบบประเภทข้อมูลแบบเครือข่ายใยประสาท (Neural Network)	41
2.5.4 ตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)	42
2.5.5 ตัวจำแนกแบบเบย์ส (Bayes Classifier)	43
2.6 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง (Theory and Related Research)	46
2.6.1 การตรวจจับใบหน้า (Face Detection)	46
2.6.2 การจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้า (Facial Emotion Classification)	55
บทที่ 3 วิธีการดำเนินงานวิจัย	66
3.1 ระเบียบวิธีในการดำเนินงานวิจัย	66
3.2 ขั้นตอนการดำเนินงาน	67
3.2.1 การตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน	67
3.2.1.1 สมมติฐานของงานวิจัยในการตรวจจับบริเวณใบหน้าของตัวการ์ตูน	68
3.2.1.2 การเตรียมข้อมูลการทดลอง	68
3.2.1.3 วิธีการทดลอง	69
3.2.2 การจำแนกอารมณ์จากใบหน้าของตัวการ์ตูน	70
3.2.2.1 สมมติฐานของงานวิจัยในการจำแนกอารมณ์จากใบหน้าของตัวการ์ตูน	72
3.2.2.2 การเตรียมข้อมูลการทดลอง	72
3.2.2.3 วิธีการทดลอง	75
บทที่ 4 ผลการวิจัยและอภิปรายผล	76
4.1 ผลของการทดลอง	76

4.1.1 การทดลองการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน	76
4.1.2 การทดลองการจำแนกอารมณ์จากใบหน้าของตัวการ์ตูน	83
4.1.2.1 ขั้นตอนการสกัดคุณลักษณะทั้งหน้า	84
4.1.2.2 ขั้นตอนการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	88
4.1.2.3 ขั้นตอนการสกัดคุณลักษณะแยกส่วนคิ้วและปาก	91
4.1.2.4 ขั้นตอนการสกัดคุณลักษณะแยกส่วนตาและปาก	95
บทที่ 5 สรุปและข้อเสนอแนะ	101
5.1 สรุปงานวิจัย	101
5.1.1 สรุปงานวิจัยด้านการตรวจจับใบหน้าการ์ตูน	101
5.1.2 สรุปงานวิจัยด้านการจำแนกอารมณ์บนใบหน้าของการ์ตูน	102
5.1.3 สรุปงานวิจัยด้านการประมวลผลภาพ	104
5.1.4 สรุปแบบสอบถามชุด Ground Truth	104
5.2 ปัญหาที่พบ	108
5.3 ข้อเสนอแนะ	109
รายการอ้างอิง	111
ภาคผนวก	
ภาคผนวก ก การจำแนกอารมณ์จากความรู้สึภายใน	115
ประวัติผู้เขียน	121

สารบัญตาราง

ตารางที่	หน้า
2.1 ตารางแสดงการเปรียบเทียบอาร์มการแสดงผลบนใบหน้าของมนุษย์และการ์ตูน	22
2.2 ค่าระดับสีของแบบจำลอง HSV	31
2.3 ข้อมูลสำหรับการสร้างตัวแบบจำแนกประเภทข้อมูลแบบเบย์ส	44
2.4 แสดงการ Combination ของทั้งหกอาร์ม	61
2.5 สรุปเทคนิคในการตรวจจับบริเวณใบหน้าและ เทคนิคในการจำแนกอาร์มพื้นฐานของงานวิจัยที่เกี่ยวข้อง	64
3.1 ใบหน้าและส่วนสำคัญของใบหน้าที่น่าจะมีส่วนช่วยในการจำแนกอาร์ม บนใบหน้าของตัวการ์ตูน	71
3.2 ตัวอย่างโดยสรุปของการจำแนกอาร์มจากใบหน้าที่การ์ตูนด้วยมนุษย์จำนวน 10 คน	73
4.1 ชุดฝึกสอนและทดสอบระบบของการ์ตูนในแต่ละระดับ	76
4.2 ผลการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนด้วยฮาร์โลคฟีเจอร์แบบมาตรฐาน	77
4.3 ผลการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนด้วย ฮาร์โลคฟีเจอร์แบบปรับปรุงชุดฝึกสอน	78
4.4 ตารางแสดงตัวอย่างภาพของการแสดงออกบนใบหน้าของการ์ตูนในแต่ละอาร์ม	83
4.5 เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะทั้งหน้า	85
4.6 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัด คุณลักษณะทั้งหน้า	85
4.7 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัด คุณลักษณะทั้งหน้า	86
4.8 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนก เพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน(SVM) แบบสกัดคุณลักษณะทั้งหน้า	86
4.9 ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าความถี่ของทิศทาง ตามค่าเกรเดียนท์ แยกส่วน คิ้ว ตา ปาก	88
4.10 เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะ แยกส่วน คิ้ว ตา ปาก	89
4.11 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$	

แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	89
4.12 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	89
4.13 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนก เพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	90
4.14 ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าความถี่ของทิศทาง ตามค่าเกรเดียนท์ แยกส่วนคิ้วและปาก	92
4.15 เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะ แยกส่วนคิ้วและปาก	92
4.16 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก	93
4.17 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก	93
4.18 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนก เพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก	93
4.19 ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าความถี่ของทิศทาง ตามค่าเกรเดียนท์ แยกส่วนตาและปาก	96
4.20 เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะ แยกส่วนตาและปาก	96
4.21 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วนตาและปาก	97
4.22 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วนตาและปาก	97
4.23 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนก เพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วนตาและปาก	97
4.24 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT) และตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k=3$	

แบบสกัดคุณลักษณะแยกส่วนคิ้ว ตา ปาก และ คิ้วและปาก	100
5.1 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะทั้งสี่ขั้นตอน	103
5.2 เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มี ค่าความเห็นที่ตรงกัน 100%	105
5.3 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มี ค่าความเห็นที่ตรงกัน 100%	106
5.4 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มี ค่าความเห็นที่ตรงกัน 100%	106
5.5 เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นที่ตรงกัน 100%	106

สารบัญภาพ

ภาพที่	หน้า
2.1 การแสดงออกทางใบหน้าของอารมณ์สับสน	6
2.2 การแสดงออกทางใบหน้าของอารมณ์อับอาย	7
2.3 การแสดงออกทางใบหน้าของอารมณ์แปลกใจ	7
2.4 การแสดงออกทางใบหน้าของอารมณ์มึนงง	8
2.5 การแสดงออกทางใบหน้าของอารมณ์อ่อนเพลีย	9
2.6 การแสดงออกทางใบหน้าของอารมณ์ขี้หวุ่น	9
2.7 การแสดงออกทางใบหน้าของอารมณ์โกรธ	10
2.8 การแสดงออกทางใบหน้าของอารมณ์กลัว	10
2.9 การแสดงออกทางใบหน้าของอารมณ์เศร้า	11
2.10 การแสดงออกทางใบหน้าของอารมณ์มีความสุข	11
2.11 เส้นร่างวงกลมและกากบาท	13
2.12 หน้าการ์ตูนแบบดัดแปลงอย่างมาก	14
2.13 หน้าการ์ตูนแบบแปลงและประดิษฐ์	14
2.14 หน้าการ์ตูนแบบดัดแปลงเล็กน้อย	15
2.15 หน้าการ์ตูนแบบเหมือนจริง	15
2.16 กล้ามเนื้อบนใบหน้า	16
2.17 กล้ามเนื้อใบหน้าเมื่อตกใจ	16
2.18 การวาดภาพการแสดงออกทางสีหน้าของตัวการ์ตูน 50 อารมณ์	17
2.19 การวาดภาพสีหน้าอารมณ์มีความสุข	19
2.20 ระดับการยิ้มของอารมณ์มีความสุข	19
2.21 ตัวอย่างสีหน้ามีความสุขของตัวการ์ตูนในแต่ละระดับการยิ้ม	20
2.22 การวาดภาพสีหน้าอารมณ์เศร้า	20
2.23 ระดับความเศร้าของอารมณ์เศร้า	20
2.24 ตัวอย่างสีหน้าเศร้าของตัวการ์ตูนในแต่ละระดับความเศร้า	21
2.25 การวาดภาพสีหน้าอารมณ์โกรธ	21
2.26 ระดับความไม่พอใจของอารมณ์โกรธ	22
2.27 ตัวอย่างสีหน้าโกรธของตัวการ์ตูนในแต่ละระดับความไม่พอใจ	22

2.28	ภาพแบบบิตแมปและภาพแบบเวกเตอร์	23
2.29	ภาพขยายแบบบิตแมปเมื่อเทียบกับภาพต้นฉบับ	24
2.30	ภาพขยายแบบเวกเตอร์เมื่อเทียบกับภาพต้นฉบับ	25
2.31	ภาพระดับความเข้มเทา	26
2.32	ภาพสี	26
2.33	ภาพไบนารีหรือภาพขาวดำ	27
2.34	ภาพดัชนี	28
2.35	แบบจำลองสี RGB	29
2.36	สมการในการแปลงค่าสี RGB ให้เป็น CMY	30
2.37	สมการของแบบจำลอง CMYK	30
2.38	ภาพโมเดลสี CMYK	30
2.39	รูปแบบของ Hue, Saturation และ Value	32
2.40	แบบจำลองสี HSV	33
2.41	ขั้นตอนในการประมวลผลภาพ	33
2.42	รูปแสดงการประมวลผลภาพเฉพาะส่วนที่สนใจเพื่อนำข้อมูลไปพัฒนาต่อ	34
2.43	การเข้ารหัสของวัตถุในภาพ	36
2.44	ภาพรวมของการสกัดหาคุณลักษณะเฉพาะในการตรวจจับวัตถุ	36
2.45	ตัวอย่างการแบ่งภาพเป็นเซลล์และบล็อก	37
2.46	เซลล์และบล็อกบริเวณที่ทำการตรวจจับ HOG	38
2.47	ลักษณะของพีเจอร์ที่ดี (Good Feature)	39
2.48	การจำแนกข้อมูล X โดยใช้วิธี K - Nearest Neighbors	40
2.49	ตัวอย่างต้นไม้อัตโนมัติสำหรับการจำแนกประเภทของข้อมูล	41
2.50	โครงสร้างเครือข่ายแบบคาดการณ์สถานการณ์ล่วงหน้า (Feed forward Network)	42
2.51	แสดงโมเดล SVM ที่แบ่งข้อมูลออกเป็น 2	43
2.52	ขั้นตอนในการตรวจจับส่วนที่เป็นใบหน้าของตัวการ์ตูน	46
2.53	ตัวอย่างการหาบริเวณที่เป็นใบหน้าของการ์ตูนด้วยเทคนิค Jaw Contour	47
2.54	(ก) ภาพการ์ตูนที่ดัดแปลงอย่างมาก	
	(ข) ภาพการ์ตูนที่มีความคล้ายหรือเหมือนกันกับมนุษย์	49
2.55	แสดงผลลัพธ์ที่ผิดพลาด จากการใช้ชุดการ Training ระบบที่มีขนาดเล็ก	49
2.56	ผลลัพธ์จากการปรับปรุงเทคนิคในการใช้ชุดฝึกอบรมที่เพิ่มมากขึ้น	50

2.57	รูปแบบของพีเจอร์ 14 ชนิดที่ใช้สำหรับการตรวจจับลักษณะแบบต่าง ๆ บนใบหน้า	50
2.58	ตัวอย่างการใช้ Haar Like Feature ตรวจจับลักษณะต่าง ๆ เพื่อหาส่วนที่เป็นใบหน้าของมนุษย์	51
2.59	รูปตัวอย่างในการคำนวณหา Integral Image	52
2.60	แสดงขั้นตอนการ รวมตัวจำแนกกลุ่มแบบเรียงต่อกัน	53
2.61	ภาพที่เก็บในฐานข้อมูลของอารมณ์โกรธที่แสดงออกทางสีหน้าของมนุษย์	54
2.62	ภาพที่เก็บในฐานข้อมูลของอารมณ์รังเกียจที่แสดงออกทางสีหน้าของมนุษย์	54
2.63	ภาพที่เก็บในฐานข้อมูลของอารมณ์มีความสุขที่แสดงออกทางสีหน้าของมนุษย์	55
2.64	แสดงให้เห็นถึงสถาปัตยกรรมพื้นฐานสำหรับเครือข่ายแบบ RBF	56
2.65	แสดงผลลัพธ์โดยประมาณของทั้ง 3 อารมณ์	57
2.66	แสดงผลการทำนายอารมณ์ของภาพที่รับเข้ามา	58
2.67	เทคนิคสำหรับการรู้จำในการแสดงออกทางสีหน้าของมนุษย์ ตามแบบการรู้จำขั้นพื้นฐาน	58
2.68	แสดงวิธีการหา Feature หลักของการแยกแยะอารมณ์ บนใบหน้ามนุษย์ด้วยเทคนิค AU	60
2.69	แสดงพีเจอร์หลักที่สนใจส่วนบนและล่างของใบหน้า	60
2.70	ภาพแสดงการตรวจจับมนุษย์ด้วย HOG	61
2.71	ภาพท่ามือที่ใช้ในการทดลอง	62
2.72	ภาพตัวอย่างการโอเวอร์แลปบล็อก และการนอร์มอลไลซ์ค่าความถี่ของทิศทางตามค่าเกรเดียนท์	63
3.1	ผังแสดงขั้นตอนการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน	68
3.2	ตัวอย่างการ์ตูนทั้งสี่ระดับ	69
3.3	ภาพที่ไม่ใช่ใบหน้า	69
3.4	ผังแสดงขั้นตอนการจำแนกอารมณ์บนใบหน้าของตัวการ์ตูน	70
3.5	แสดงภาพการตรวจจับบริเวณที่เป็น คิ้ว ตา ปาก บนใบหน้าการ์ตูน	71
3.6	ตัดแยกเฉพาะส่วนที่เป็น (ก) คิ้ว (ข) ตา (ค) ปาก	71
3.7	ตัวอย่างแบบสอบถามที่ให้ผู้ชมทดสอบ	73
3.8	ตัวอย่างชุดข้อมูลฝึกสอนระบบ ที่ตัดเฉพาะส่วนที่เป็นบริเวณใบหน้าของการ์ตูน	75
4.1	ภาพ (ก) ก่อนปรับชุดฝึกสอนตรวจจับไม่ได้ (ข) หลังปรับชุดฝึกสอนตรวจจับได้	79
4.2	ภาพ (ก) ก่อนปรับชุดฝึกสอน (ข) หลังปรับชุดฝึกสอน ไม่สามารถตรวจจับได้	80

4.3 ภาพ (ก) ก่อนปรับชุดฝึกสอนตรวจจับได้ (ข) หลังปรับชุดฝึกสอนตรวจจับไม่ได้	81
4.4 ตัวอย่างภาพการ์ตูนแอนิเมชันทั้งหมด 20 เรื่อง	82
4.5 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า ของการสกัดคุณลักษณะทั้งหน้า	87
4.6 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะทั้งหน้า	87
4.7 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะทั้งหน้า	87
4.8 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ จากโกรธเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	90
4.9 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	91
4.10 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก	91
4.11 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก	94
4.12 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก	94
4.13 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก	95
4.14 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า	

ของการสกัดคุณลักษณะแยกส่วนตาและปาก	98
4.15 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข	
(ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า	
ของการสกัดคุณลักษณะแยกส่วนตาและปาก	98
4.16 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า	
(ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข	
ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก	99
5.1 ผลการจำแนกอารมณ์ที่ผิดพลาดของระบบ	104
5.2 ผลการจำแนกอารมณ์ที่ไม่ตรงกันของมนุษย์	105
5.3 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ	
จากโกรธเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก	
จากภาพที่มีค่าความเห็นตรงกัน 100%	107
5.4 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข	
จากมีความสุขเป็นโกรธ ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก	
จากภาพที่มีค่าความเห็นตรงกัน 100%	107
5.5 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า	
(ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข	
ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก	
จากภาพที่มีค่าความเห็นตรงกัน 100%	108
5.6 ตัวอย่างภาพการ์ตูนที่มีเส้นผมปกคลุมบริเวณใบหน้า	108
5.7 ชุดข้อมูลภาพที่ต่อเนื่องกันของการ์ตูน	110
ก.1 ตัวอย่างของการเกิดอารมณ์จากภายใน	
หลังจากมีการเปลี่ยนแปลงทางสรีระทางร่างกาย	117
ก.2 เกิดอารมณ์จากภายในหลังจากมีการเปลี่ยนแปลง	
ทางสรีระทางร่างกายตามทฤษฎีของแคนนอน - บาร์ด	118
ก.3 ทฤษฎีทางอารมณ์ของ แซตเตอร์ - ซิงเกอร์	119
ก.4 แสดงการเกิดของอารมณ์ตามทฤษฎีทั้ง 3 ทฤษฎี	119

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของงานวิจัย

การแยกแยะอารมณ์ของตัวการ์ตูนเป็นงานจำแนกประเภทวัตถุ (Classification) ในรูปแบบนามธรรมของการค้นคืนภาพ นอกจากนี้ยังสามารถนำไปใช้ทำในส่วนของการดัชนีภาพหรือดัชนีวิดีโอได้ด้วย การทำดัชนีของภาพนั้น มีอยู่ด้วยกัน 3 ระดับ ระดับที่หนึ่ง เป็นระดับพื้นฐาน สามารถพบเห็นได้ทั่วไป คือ ดัชนีเกี่ยวกับ สี พื้นผิว และรูปร่าง ระดับที่สอง คือ ระดับการแยกแยะความหมายเบื้องต้นของภาพ เช่น การทำดัชนีเกี่ยวกับ ภาพคน ภาพสัตว์ และภาพสิ่งของ เป็นต้น ระดับสุดท้าย ระดับที่สาม เป็นระดับที่จะทำกับงานวิจัยนี้ คือ การทำดัชนีของภาพที่เป็นนามธรรม เช่น การทำดัชนีเกี่ยวกับภาพสื่ออารมณ์ ภาพสื่อความหมายเกี่ยวกับศาสนาหรือวัฒนธรรม เป็นต้น การพัฒนาระบบต้นแบบนี้ขึ้นมาเพื่อให้ระบบสามารถนำไปประยุกต์ใช้ในการทำดัชนีอารมณ์ของตัวการ์ตูนในภาพยนตร์แทนมนุษย์ได้

การผลิตภาพยนตร์ในปัจจุบันมีความหลากหลาย ไม่ใช่แค่ภาพยนตร์ที่มนุษย์แสดงเท่านั้นแต่มีภาพยนตร์ที่สร้างขึ้นมาจากการ์ตูนแอนิเมชันรวมเข้ามาด้วย ผู้วิจัยจึงมีแนวคิดที่จะนำภาพการ์ตูนแอนิเมชันมาใช้ในการวิเคราะห์และประมวลผลภาพเพื่อให้เกิดประโยชน์กับวงการวิจัยทางด้านคอมพิวเตอร์ และมีแนวคิดว่าจะทำอย่างไรที่ทำให้คอมพิวเตอร์สามารถแยกแยะอารมณ์จากการแสดงออกทางใบหน้าของตัวการ์ตูนในภาพยนตร์แอนิเมชันได้เหมือนกับการแยกแยะอารมณ์จากการแสดงออกทางใบหน้าของมนุษย์ จากการศึกษาของงานวิจัยที่เกี่ยวข้องกับการแยกแยะอารมณ์บนใบหน้าที่เห็นกันโดยทั่วไปนั้น งานวิจัยที่พบจะมุ่งเน้นไปที่การแยกแยะอารมณ์บนใบหน้าของมนุษย์เป็นหลัก ผู้วิจัยจึงเล็งเห็นถึงความน่าสนใจและประโยชน์ของงานวิจัยนี้

งานวิจัยนี้จะสามารถช่วยแก้ปัญหาในการรู้จำอารมณ์บนใบหน้าการ์ตูนได้ ด้วยการนำภาพการ์ตูนมาวิเคราะห์โดยใช้ความรู้ด้านการประมวลผลภาพ (Image Processing) และทัศนศาสตร์ทางคอมพิวเตอร์ (Computer Vision) ซึ่งระบบจะนำภาพใบหน้าการ์ตูนไปทำการวิเคราะห์หาคุณสมบัติ (Feature) ของอารมณ์ (Emotion) ว่าภาพการ์ตูนที่นำไปประมวลผลนั้น เป็นอารมณ์ประเภทใด โดยในงานวิจัยนี้จะนำอารมณ์พื้นฐานของมนุษย์สามประเภท คือ อารมณ์โกรธ (Anger) อารมณ์เศร้า (Sadness) และอารมณ์มีความสุข (Happiness) มาใช้ในการวิเคราะห์อารมณ์บนใบหน้าตัวการ์ตูน

วิทยานิพนธ์ฉบับนี้เสนอแนวทางในการแยกแยะอารมณ์ของภาพการ์ตูน โดยเริ่มจากรับภาพการ์ตูนเข้ามาในระบบและทำการประมวลผลภาพ (Image Processing) จากนั้นทำการตรวจจับบริเวณใบหน้า (Face Detection) ของการ์ตูน เมื่อได้ส่วนที่เป็นใบหน้าของการ์ตูนแล้ว ทำการวิเคราะห์ภาพด้วยการใช้คุณสมบัติของการแยกลักษณะเด่นของวัตถุ (Feature Extraction) เพื่อหาบริเวณที่เป็น คิ้ว ตา ปาก และจมูก ของการ์ตูน จากนั้นนำคุณลักษณะในแต่ละส่วนที่ได้ไปทำการจำแนกประเภทของอารมณ์ (Emotion Classification) โดยใช้ตัวแบบการจำแนกประเภท (Training Data) ที่หลากหลายมาทำการเปรียบเทียบกันว่า ตัวจำแนกประเภทใดสามารถวิเคราะห์หาอารมณ์ที่แสดงบนใบหน้าของตัวการ์ตูนได้ดีที่สุด

1.2 วัตถุประสงค์ของงานวิจัย

งานวิจัยนี้มีเป้าหมายเพื่อศึกษาวิธีการที่สามารถนำไปใช้ในการแยกแยะอารมณ์ของตัวการ์ตูนจากภาพ จึงได้กำหนดวัตถุประสงค์ของงานวิจัยดังนี้

1. เพื่อศึกษาและหาอัลกอริทึมในการตรวจจับที่เป็นบริเวณใบหน้าของการ์ตูนในภาพ
2. เพื่อศึกษาคุณลักษณะที่เกี่ยวข้องกับการแสดงอารมณ์พื้นฐานที่เหมาะสมและนำมาสร้างแบบจำลองให้สามารถจำแนกอารมณ์บนใบหน้าการ์ตูนได้
3. เพื่อเปรียบเทียบประสิทธิภาพการจำแนกอารมณ์บนใบหน้าการ์ตูนจากเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ได้แก่ ต้นไม้การตัดสินใจ (Decision Tree) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k-Nearest Neighbor Algorithm)

1.3 ขอบเขตของงานวิจัย

ประเภทของภาพการ์ตูนที่ใช้ เป็นการ์ตูนแอนิเมชัน (Animation Cartoon) ที่มีใบหน้าคล้ายหรือเหมือนกับใบหน้าของมนุษย์

1.4 คำจำกัดความของคำศัพท์เฉพาะ

เพื่อให้เข้าใจความความหมายเบื้องต้นของคำศัพท์เฉพาะทางด้านเทคนิค ผู้วิจัยจึงได้จัดทำส่วนของคำจำกัดความของคำศัพท์เฉพาะขึ้นมา เพื่ออธิบายดังนี้

1. การ์ตูนภาพเคลื่อนไหว หรือ แอนิเมชัน หรือ การ์ตูนภาพยนตร์ (Animated Cartoons) หมายถึง การ์ตูนที่มีการลำดับในการเคลื่อนไหวภาพและเรื่องราวที่ต่อเนื่องกันคล้ายภาพยนตร์ มีความเปลี่ยนแปลงลักษณะท่าทางและอิริยาบถ โดยที่การวาดภาพการเคลื่อนไหวของการ์ตูนนั้น เป็นการวาดจากภาพหนึ่งไปยังอีกภาพหนึ่งเพื่อให้เห็นถึงความแตกต่างและเกิดเป็นภาพเคลื่อนไหวเกิดขึ้น (สุตามาศ ผลดี, 2553)

2. การวาดภาพใบหน้าตัวการ์ตูนขั้นพื้นฐาน (Basic Cartoon Face) หมายถึง การวาดภาพโดยใช้รูปทรงเรขาคณิตมาเป็นส่วนของการร่างภาพ ซึ่งในส่วนของใบหน้าก็จะมีขั้นตอนในการวาดที่ง่าย โดยรูปทรงเรขาคณิตที่เป็นจุดกำเนิดของการวาดใบหน้า ประกอบไปด้วยวงกลมและกากบาท การวาดภาพใบหน้าตัวการ์ตูนที่เห็นในปัจจุบันมีอยู่ด้วยกันสี่ระดับ ได้แก่ การวาดใบหน้าการ์ตูนแบบตัดแปดอย่างมา การวาดใบหน้าการ์ตูนแบบตัดแปดและประดิษฐ์ การวาดใบหน้าการ์ตูนแบบตัดแปดเล็กน้อย และการวาดใบหน้าการ์ตูนแบบเหมือนจริง (ศิริพร ชัยมานะ, 2555)

3. อารมณ์โกรธ (Anger) เป็นอารมณ์เดิมหรืออารมณ์หลัก (Primary Emotion) เป็นอารมณ์ที่แสดงออกได้อย่างชัดเจน สามารถรับรู้ได้แน่นอน คือ คิ้วจะขมวด เปลือกตาจะตึงและขวาง ความโกรธเกิดจากสถานการณ์ที่ไม่พึงประสงค์ ซึ่งเกิดจากความโกรธเคืองหรือเกิดจากความผิดหวัง (Dobrin, 2013)

4. อารมณ์มีความสุข (Happiness) เป็นอารมณ์ที่เป็นสากลที่สุดซึ่งสามารถมองเห็นได้กับทุกวัฒนธรรม พฤติกรรมการแสดงออกทางอารมณ์นี้ติดตัวมาตั้งแต่กำเนิด อารมณ์ความสุขนี้จะแสดงออกได้ชัดเจนจากรอยยิ้มซึ่งเป็นรอยยิ้มที่แตกต่างออกมาจากอารมณ์อื่น (Dobrin, 2013)

5. อารมณ์เศร้า (Sadness) ความโศกเศร้ามักจะเกิดขึ้นร่วมกับความรู้สึกสูญเสียและทำอะไรไม่ถูก คุณลักษณะของความโศกเศร้า จะทำให้ใบหน้าของมนุษย์หุบลงมีความเหี่ยวเฉา การแสดงออกทางสีหน้าไม่สดใส (Dobrin, 2013)

1.5 ประโยชน์ที่คาดว่าจะได้รับ

หลังจากดำเนินการวิจัยตามวัตถุประสงค์ ภายใต้ขอบเขตของการวิจัยที่ได้กำหนดไว้แล้ว ผู้วิจัยคาดว่าจะงานวิจัยจะให้ประโยชน์ดังนี้

1. ได้อัลกอริทึมที่เหมาะสมในการตรวจจับใบหน้าของตัวการ์ตูนในภาพ
2. ทำให้ทราบถึงเทคนิคการเรียนรู้ของเครื่องที่เหมาะสมกับการจำแนกอารมณ์พื้นฐานบนใบหน้าของตัวการ์ตูน และสามารถนำข้อมูลจากการศึกษางานวิจัยพื้นฐาน เพื่อเป็นแนวทางในการประยุกต์ใช้กับงานวิจัยอื่นที่เกี่ยวข้องกับงานด้านการจำแนกอารมณ์ของภาพได้

3. ทำให้ทราบถึงคุณลักษณะ (Feature) ทางใบหน้าการันตุนที่นำไปใช้ในการจำแนก
อารมณ์จากใบหน้าของตัวการ์ตูนได้อย่างมีประสิทธิภาพ

4. นำไปประยุกต์ใช้ในการทำดัชนีเกี่ยวกับอารมณ์ ซึ่งจัดเป็นดัชนีในระดับนามธรรม
เพื่อใช้ในการสืบค้นข้อมูลจากฐานข้อมูลภาพ และฐานข้อมูลวิดีโอได้



บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง ประกอบด้วยความรู้พื้นฐานเกี่ยวกับอารมณ์ (Emotion Basic Knowledge) ความรู้พื้นฐานเกี่ยวกับการ์ตูน (Cartoon Basic Knowledge) ความรู้พื้นฐานเกี่ยวกับการประมวลผลและวิเคราะห์ภาพดิจิทัล (Digital Image Processing & Analysis) ซึ่งรวมไปถึง เทคนิคการปรับปรุงภาพและการกำจัดสัญญาณรบกวนภายในภาพ (Image Enhancement) การตรวจจับใบหน้า (Face Detection) และการวิเคราะห์ภาพด้วยการใช้คุณสมบัติของการแยกลักษณะเด่นของวัตถุ (Feature Extraction & Analysis) ในส่วนงานวิจัยที่เกี่ยวข้อง จะเป็นการ ทบทวนวรรณกรรมที่เกี่ยวกับการตรวจจับใบหน้าของมนุษย์จริง ใบหน้าของการ์ตูน และการแยกแยะ อารมณ์บนใบหน้ามนุษย์

2.1 ความรู้พื้นฐานเกี่ยวกับอารมณ์ (Emotion Basic Knowledge)

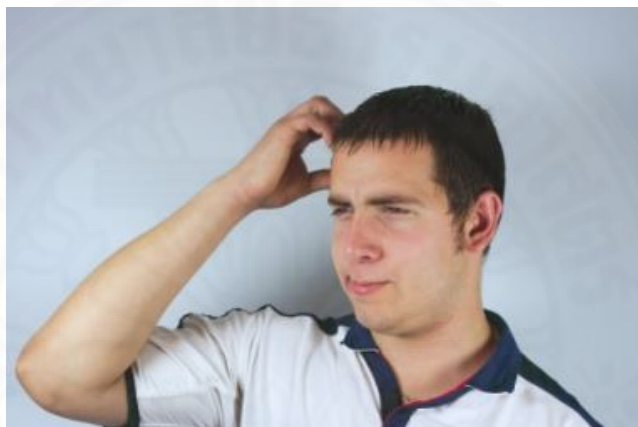
อารมณ์ เกิดจากความรู้สึกที่เกิดขึ้นจากการที่สภาวะร่างกายถูกสิ่งยั่วยุ ในแต่ละ สถานการณ์ จนเกิดเป็นการเปลี่ยนแปลงทางการแสดงออกทางสรีรวิทยาของมนุษย์ เช่น ใจสั่น หน้า แดง ซิพจรเต้นเร็วขึ้น เป็นต้น บางครั้งอารมณ์ก็โยงไปถึงความรู้สึกภายใน ที่เกิดขึ้นหลังจากสภาวะ ทางร่างกายถูกยั่วยุจากสถานการณ์ เช่น เกิดความพอใจหรือไม่พอใจ มนุษย์แสดงอาการของอารมณ์ เหล่านั้นออกมาด้วย สีหน้า ท่าทาง คำพูด หรือ น้ำเสียง เป็นต้น ซึ่งสามารถสรุปได้ว่าอารมณ์เป็นได้ ทั้งแรงจูงใจและเป้าประสงค์ในชีวิต ถ้าเป็นอารมณ์ที่น่าพึงพอใจก็เป็นเป้าประสงค์เชิงบวก ถ้าไม่น่าพึง พพอใจก็เป็นเป้าประสงค์เชิงลบ (พริ้มเพรา ดิษยวณิช, 2548)

2.1.1 การจำแนกอารมณ์จากการแสดงอารมณ์บนใบหน้า (Classification of emotion from facial expressions)

การจำแนกอารมณ์จากการแสดงอารมณ์บนใบหน้า (Classification of emotion from facial expressions) การแสดงอารมณ์จะมองเห็นได้บ่อยจากการแสดงออกทางสี หน้า ท่าทาง และน้ำเสียง งานวิจัยของนักจิตวิทยาที่ผ่านมาจะศึกษาเกี่ยวกับการแสดงออกทางสีหน้า ของมนุษย์เป็นส่วนใหญ่ Mehrabian ระบุว่า การสื่อสารออกทางตัวอักษรมีผลต่อการจำแนกอารมณ์ เพียง 7% เท่านั้น และการสื่อสารด้วยน้ำเสียงมีผล 38% เมื่อเทียบกับการแสดงออกทางสีหน้าของผู้ พูด ซึ่งมีผลมากที่สุดถึง 55% จึงสามารถสรุปได้ว่าการแสดงออกทางสีหน้าของมนุษย์คือสิ่งที่สำคัญที่สุด

ในการจำแนกอารมณ์ (Mehrabian, 1968) ซึ่ง Dobrin นักจิตวิทยาชาวอเมริกันกล่าวถึงการสื่อสารทางร่างกายที่แสดงออกทางสีหน้าโดยแบ่งเป็นอารมณ์ทางสีหน้า 10 ชนิด ดังนี้ (Dobrin, 2013)

2.1.1.1. อารมณ์สับสน (confusion) การแสดงออกของความสับสนจะถูกกำหนดโดยจมูกและหน้าผากที่ขมวด บางครั้งอาจจะมีคิ้วยกสูงขึ้นและอาจจะมีการแหม้มริมฝีปากร่วมด้วย แต่ถึงอย่างไรก็ตามความสับสนสามารถเห็นได้ชัดมากที่สุดจากการแสดงออกทางจมูกและดวงตานอกจากนี้ยังอาจจะแสดงท่าทางประกอบไปด้วย เช่น ยกมือเกาศีรษะหรือเอียงคอ เป็นต้น ดังภาพที่ 2.1 ซึ่งมนุษย์สามารถเข้าใจความหมายของอารมณ์นี้ได้



ภาพที่ 2.1 การแสดงออกทางใบหน้าของอารมณ์สับสน จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.2. อารมณ์อับอาย (Shame) เป็นการแสดงออกที่มองเห็นได้ง่ายและเป็นที่ยอมรับสากลสำหรับการรู้จำ ซึ่งประกอบไปด้วยการหลบสาย ตาที่มองลงด้านล่างและมีลักษณะเศร้าหรือกังวลร่วมด้วย ศีรษะก็มักจะคว่ำลงปากอาจจะมู่ยหรือปกติก็ได้ ความอับอายมักมีความสัมพันธ์ร่วมกันกับความละอายใจในสังคม ที่มีแต่ความซบซ้อน ความพ่ายแพ้จากการแข่งขัน อาจจะทำให้บุคคลคนเกิดความอับอายที่มาจากความพ่ายแพ้ในการแข่งขันโดยที่ไม่ได้คาดคิดมาก่อน หรือ เกิดจากความรู้สึกที่ตัวเองเอาไปเปรียบเทียบกับคนอื่น โดยเกิดมาจากการตั้งเงื่อนไขให้ตนเอง แต่ความหมายของความอับอายทั้งหมดเกิดมาจากรู้สึกผิดหวังหรือสูญเสีย การแสดงออกทางอารมณ์อับอายนี้ อาจจะมีท่าทางประกอบไปด้วย เช่น ก้มหน้า ยกมือปิดหน้าปิดตา เป็นต้น ดังภาพที่ 2.2



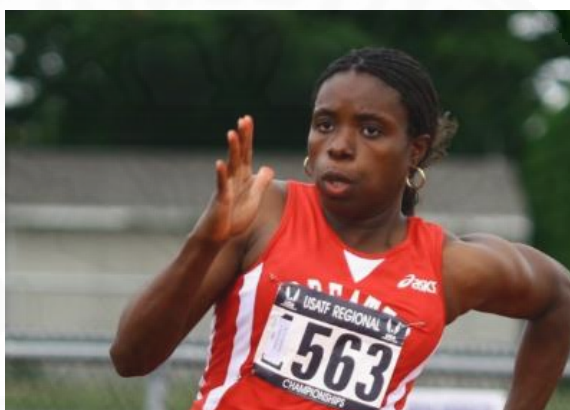
ภาพที่ 2.2 การแสดงออกทางใบหน้าของอารมณ์อับอาย จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.3. อารมณ์แปลกใจ (Surprise) สามารถมองเห็นภาพได้โดยง่าย ด้วยดวงตาที่โตกว่าปกติและปากที่เบิกกว้าง อารมณ์เกิดจากรู้สึกประหลาดใจหรือตกใจ ซึ่งอารมณ์นี้มีการแสดงลักษณะทางใบหน้าที่มีความใกล้เคียงหรือคล้ายกับอารมณ์กลัว ใบหน้าประหลาดใจเป็นใบหน้าหนึ่งที่ถูกแสดงออกมากที่สุดของมนุษย์ การแสดงสีหน้าส่วนใหญ่เกิดขึ้นเองโดยอัตโนมัติ โดยที่มนุษย์ไม่รู้ตัวมาก่อนเป็นปฏิกิริยาที่เกิดขึ้นทันทีกับสถานการณ์ในตอนนั้น ปฏิกิริยาสามารถเกิดขึ้นได้กับทั้งคนและสัตว์ เมื่อพบเจอสิ่งที่ไม่คาดคิด ดวงตามักจะขยาย ม่านตาจะเบิกกว้างและร่างกายจะตอบสนองต่อสถานการณ์นั้นได้อย่างรวดเร็ว นอกจากนี้ยังสามารถมีท่าทางร่วมด้วย เช่น เอามือปิดปาก ปิดแก้ม เป็นต้น ดังภาพที่ 2.3



ภาพที่ 2.3 การแสดงออกทางใบหน้าของอารมณ์แปลกใจ จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.4. อารมณ์มุ่งมั่น (Focus) การแสดงออกทางสีหน้ามุ่งมั่นขึ้นอยู่กับสถานการณ์ที่แตกต่าง บุคคลที่มีความมุ่งมั่นจะแสดงออกทางสายตาที่แน่วแน่และมั่นคง อาจจะมีมุ่งมั่นไปที่เป้าหมายที่อยู่ทางด้านหน้า หากมีการกระพริบตาที่อาจเกิดขึ้นเพียงเล็กน้อยเท่านั้น สิ่งที่น่าสนใจสำหรับอารมณ์นี้ คือ เมื่อบุคคลหนึ่งคนมุ่งมั่นที่จะบรรลุเป้าหมาย จะถูกสั่งการด้วยสมอง อาจแสดงออกทางร่างกาย เช่น ลิ้นจะบิตตัวและย้ายไปทางด้านข้าง การแสดงออกทางอารมณ์ลักษณะนี้จะเกิดโดยไม่รู้ตัว การแสดงอารมณ์มุ่งมั่นนี้ ไม่ได้มีการแสดงออกเพียงแค่การบิตตัวของลิ้นไปทางด้านข้างเท่านั้น บางคนยังแสดงออกจากการกระพริบตานั้นน้อยลงด้วย เป็นต้น ลักษณะท่าทางของการแสดงออกทางสีหน้าอารมณ์มุ่งมั่น ดังภาพที่ 2.4



ภาพที่ 2.4 การแสดงออกทางใบหน้าของอารมณ์มุ่งมั่น จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.5. อารมณ์อ่อนเพลีย (Exhaustion) เป็นอารมณ์ที่เห็นได้ชัดจากการที่มนุษย์พยายามแสดงลักษณะท่าทางในการยกคิ้วให้สูงขึ้น เพื่อพยายามเปิดตาให้กว้างขึ้น เสมือนปลุกร่างกายให้ตื่นจากความเมื่อยล้า ซึ่งความเหนื่อยล้าที่เกิดขึ้นมาจากการที่ใช้พลังงานทางร่างกายอย่างหักโหม การแสดงออกทางสีหน้าที่เห็นได้ชัดของมนุษย์สามารถบ่งบอกถึงพลังงานภายในร่างกายที่ลดน้อยลงของเขาได้ดีอีกด้วย ดังภาพที่ 2.5 จะเป็นการแสดงออกทางใบหน้าของอารมณ์อ่อนเพลีย อาจจะมีการแสดงออกทางท่าทางโดยอาจจะหายใจแผ่วลง กิริยาท่าทางในการเคลื่อนไหวทางสรีระ ก็อาจจะทำได้ช้าลงไปด้วย



ภาพที่ 2.5 การแสดงออกทางใบหน้าของอารมณ์อ่อนเพลีย จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.6. อารมณ์ยั่ววน (Seduction) เป็นการแสดงทางอารมณ์ที่มีศิลปะเพื่อหลอกล่อให้ชวนหลงใหล อารมณ์นี้จะถูกแสดงออกเพื่อกระตุ้นอารมณ์ทางเพศเป็นส่วนใหญ่ การแสดงออกทางสีหน้าที่มีเสน่ห์ชวนหลงใหลเป็นสิ่งที่สามารถมองเห็นได้ชัดจากดวงตาที่จับจ้องไปยังบุคคลตรงหน้า ริมฝีปากที่เฝิ้มหรือแยกห่างออกจากกันเล็กน้อย ศีรษะที่เอียงไปข้างใดข้างหนึ่งเพียงเล็กน้อย ทำให้มองเห็นลำคอมากขึ้นเพื่อเป็นการกระตุ้นให้เกิดความยั่ววน แววตาจะเป็นประกายเมื่อมองไปยังบุคคลที่อยู่ตรงหน้า สายตาที่มองจะมุ่งเน้นไปที่ริมฝีปากและลำคอ บวกกับการกระพริบตาที่ถี่ขึ้นและรอยยิ้มอ่อน ๆ ที่ส่งผลให้การแสดงออกทางอารมณ์ยั่ววนประสบความสำเร็จ ดังภาพที่ 2.6



ภาพที่ 2.6 การแสดงออกทางใบหน้าของอารมณ์ยั่ววน จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.7. อารมณ์โกรธ (Anger) เป็นอารมณ์ที่แสดงออกได้อย่างชัดเจน สามารถรับรู้ได้แน่นอน คือ คิ้วจะขมวด เปลือกตาจะตึงและขวาง หัวคิ้วมักจะลดลงเล็กน้อยและสายตาในการมองจะหริ่ง ด้วยการมองลอดผ่านหัวคิ้ว ความโกรธเกิดจากสถานการณ์ที่ไม่พึงประสงค์

ซึ่งเกิดจากความโกรธเคืองหรือเกิดจากความผิดหวัง การแสดงออกทางสีหน้าของความโกรธทำให้มองเห็นถึงความเครียดบนใบหน้าที่ตั้งและการแสดงออกทางสีร้ระร่วมด้วย เช่น เมื่อมีคนหนึ่งในห้องกำลังโกรธ คนที่อยู่ในห้องก็สามารถรับรู้ได้อย่างชัดเจน เป็นต้น การแสดงออกทางสีหน้าของอารมณ์โกรธ ดังภาพที่ 2.7



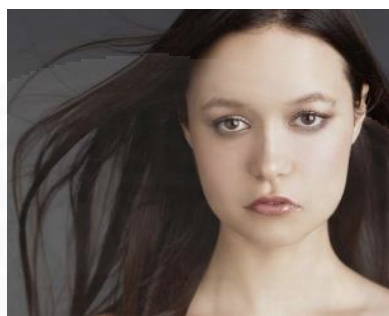
ภาพที่ 2.7 การแสดงออกทางใบหน้าของอารมณ์โกรธ จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.8. อารมณ์กลัว (Fear) การแสดงออกทางสีหน้าของอารมณ์นี้จะเห็นได้อย่างชัดเจน คือ ตาที่เบิกกว้างและคิ้วที่เปื้อน ปากมักจะถูกเปิดออกในระดับหนึ่ง อารมณ์กลัวการแสดงออกทางสีหน้ามีความใกล้เคียงและมีความคล้ายกันกับอารมณ์ประหลาดใจ มีรากฐานทางอารมณ์ที่ใกล้เคียงกันมากและบ่งบอกถึงความปรารถนาที่จะหลบเลี่ยงและหนีบางสิ่งบางอย่าง นักวิจัยทางจิตวิทยาค้นพบว่าผู้ที่มีความกลัวจะหายใจแรงมากขึ้นกว่าปกติและยังมีการแสดงของตาที่เบิกกว้างมากขึ้นร่วมด้วย สามารถสรุปได้โดยย่อว่าการแสดงออกทางสีหน้ามาจากการรับรู้ทางประสาทสัมผัส ดังภาพที่ 2.8



ภาพที่ 2.8 การแสดงออกทางใบหน้าของอารมณ์กลัว จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.9. อารมณ์เศร้า (Sadness) การแสดงออกทางสีหน้า คือ คิ้วเฉียงและขมวด ความโศกเศร้ามักจะเกิดขึ้นร่วมกับความรู้สึกสูญเสียและทำอะไรไม่ถูก คุณลักษณะของความโศกเศร้า จะทำให้ใบหน้าของมนุษย์หุบลงมีความเหี่ยวเฉา การแสดงออกทางสีหน้าไม่สดใส โดยที่ไม่ได้เน้นบริเวณใดบริเวณหนึ่ง ดังภาพที่ 2.9



ภาพที่ 2.9 การแสดงออกทางใบหน้าของอารมณ์เศร้า จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.1.1.10. อารมณ์มีความสุข (Happiness) เป็นอารมณ์ที่เป็นสากลที่สุดซึ่งสามารถมองเห็นได้กับทุก ๆ วัฒนธรรม การแสดงอารมณ์นี้สามารถระบุได้จากรอยยิ้มและดวงตาที่เป็นรูปพระจันทร์ครึ่งเสี้ยว ซึ่งพฤติกรรมการแสดงออกทางอารมณ์นี้ติดตัวมาตั้งแต่กำเนิด อารมณ์ความสุขนี้จะแสดงออกได้ชัดเจนจากรอยยิ้มซึ่งเป็นรอยยิ้มที่แตกต่างออกมาจากอารมณ์อื่น รอยยิ้มที่เบิกกว้างจนเห็นฟันยังสามารถไปถึงสู่สภาพที่ดีของมนุษย์อีกด้วย ดังภาพที่ 2.10



ภาพที่ 2.10 การแสดงออกทางใบหน้าของอารมณ์มีความสุข จาก *Facial Expressions: Universal vs. Cultural*, โดย Dobrin, A., 2013.

2.2 ความรู้พื้นฐานเกี่ยวกับการ์ตูน (Cartoon Basic Knowledge)

คำว่า "การ์ตูน" ในภาษาไทย เขียนมาจากคำทับศัพท์ในภาษาอังกฤษว่า "Cartoon" และคำว่าการ์ตูนในภาษาอังกฤษสันนิษฐานว่ามาจากคำว่า คาโทเน (Cartone) ในภาษาอิตาลี ซึ่งมีความหมายว่า แผ่นกระดาษที่มีภาพวาดอยู่ด้านใน (สุพัฒน์, 2556) แต่บทความบางฉบับกล่าวว่า การ์ตูน เป็นวรรณกรรมประเภทหนึ่งที่ถ่ายทอดความรู้สึกและความเข้าใจด้วยการวาดภาพ ซึ่งภาพที่เขียนขึ้นนั้นมาจากการวาดภาพจำลองสถานการณ์ ทำให้มนุษย์สามารถเข้าใจเรื่องราวและเหตุการณ์ได้ ในส่วนของการวาดภาพการ์ตูนนั้น ไม่จำเป็นต้องมีความประณีตมาก เพียงแต่วาดให้สามารถมองเห็นได้ว่าสิ่งที่วาดนั้นคืออะไร (นิรุทธิ์ อุทาพิทย์, 2553)

2.2.1 ประเภทของการ์ตูน

การแบ่งประเภทของการ์ตูนสามารถแบ่งได้สองลักษณะ คือ แบ่งตามลักษณะในการเคลื่อนไหว และ แบ่งตามอรรถบท (Theme) ที่พบเห็นในประเทศไทย ซึ่งจะอธิบายตามลำดับดังต่อไปนี้ (สุดามาศ ผลดี, 2553)

2.2.1.1 ประเภทของการ์ตูนที่แบ่งตามลักษณะในการเคลื่อนไหว

สามารถแบ่งเป็นสองลักษณะ ดังนี้

1. การ์ตูนภาพนิ่ง (Static Cartoons) หมายถึง การ์ตูนที่ไม่มีการเปลี่ยนแปลงอิริยาบถ ซึ่งการ์ตูนในลักษณะนี้จะเป็นการบอกเล่าหรือเล่าเรื่องราว ซึ่งจะไม่มีการดำเนินเรื่องของการ์ตูน

2. การ์ตูนภาพเคลื่อนไหว (Dynamic Cartoons) หมายถึง การ์ตูนที่มีความเปลี่ยนแปลงลักษณะท่าทางและอิริยาบถ โดยที่การวาดภาพการเคลื่อนไหวของการ์ตูนนั้นเป็นการวาดจากภาพหนึ่งไปยังอีกภาพหนึ่งเพื่อให้เห็นถึงความแตกต่างและเกิดเป็นภาพเคลื่อนไหวเกิดขึ้น

2.2.1.2 ประเภทของการ์ตูนที่แบ่งตามอรรถบทที่พบเห็นในประเทศไทย

ลักษณะของการวาดภาพการ์ตูนในประเทศไทยแบ่งตามอรรถบทที่พบเห็นได้บ่อย จากหลากหลายสื่อในการนำเสนอ เช่น สื่อสิ่งพิมพ์ หรือ สื่อทางโทรทัศน์ เป็นต้น เพื่อให้เข้าใจเหตุผลของการวาดการ์ตูนในประเทศไทยเพิ่มมากขึ้น สามารถแบ่งประเภทของการ์ตูนตามอรรถบทที่พบเห็นในประเทศไทยได้เป็นห้าลักษณะ ดังนี้

1. การ์ตูนการเมือง (Political Cartoons) เป็นการ์ตูนที่วาดขึ้นมาเพื่อมุ่งเน้นการล้อเลียน เสียดสี ประชดประชันบุคคล หรือ เหตุการณ์ที่สำคัญภายในประเทศ นิยมล้อ

เหตุการณ์ทางการเมือง เพื่อกระตุ้นให้ผู้อ่านเกิดความสนใจ ลักษณะการ์ตูนชนิดนี้อาจมีคำบรรยายหรือไม่มีก็ได้

2. การ์ตูนขำขัน (Gag Cartoons) เป็นการ์ตูนภาพเคลื่อนไหว ที่มุ่งเน้นความตลกขบขันเป็นหลัก การ์ตูนลักษณะนี้จะนิยมนำเหตุการณ์ใกล้ตัวมาเขียนเพื่อให้ผู้อ่านได้ อรรถรสในการอ่านมากขึ้น ซึ่งเป็นการ์ตูนที่ได้รับความนิยมมากในสังคมไทย

3. การ์ตูนเรื่องยาว (Comicorserial Cartoons) คือ การ์ตูนที่เขียนขึ้นมาเป็นเรื่องราวต่อเนื่องกันจนจบ การ์ตูนชนิดนี้ปรากฏอยู่ในนิตยสาร และหนังสือพิมพ์ เรียกว่า Comics Strips แต่ถ้านำมาพิมพ์รวมเล่ม เรียกว่า Comics Books เช่น การ์ตูนเล่มของญี่ปุ่นและฝรั่ง ส่วนของการ์ตูนไทยนั้นนิยมนำเรื่องจากวรรณคดี นิทานพื้นบ้าน จักร ๆ วงศ์ ๆ เป็นต้น

4. การ์ตูนประกอบเรื่อง (Illustrated Cartoons) เป็นการ์ตูนที่วาดขึ้นมาเพื่อใช้ประกอบกับข้อเขียน หรือ ประกอบโฆษณาเพื่อขยายความ หรือ การ์ตูนประกอบการศึกษา เป็นต้น

5. การ์ตูนภาพเคลื่อนไหว หรือ แอนิเมชัน หรือ การ์ตูนภาพยนตร์ (Animated Cartoons) เป็นการ์ตูนที่มีการลำดับในการเคลื่อนไหวภาพและเรื่องราวที่ต่อเนื่องกัน คล้ายภาพยนตร์

2.2.2 การวาดภาพใบหน้าการ์ตูนขั้นพื้นฐาน

การวาดใบหน้าขั้นพื้นฐานนั้นสามารถทำได้ไม่ยาก เพราะ การวาดภาพโดยปกติ จะใช้รูปทรงเรขาคณิตมาเป็นส่วนของการร่างภาพ ซึ่งในส่วนของใบหน้าก็จะมีขั้นตอนในการวาดที่ง่าย โดยรูปทรงเรขาคณิตที่เป็นจุดกำเนิดของการวาดใบหน้านั้น ประกอบไปด้วยวงกลมและกากบาท ซึ่งจะเรียกสิ่งเหล่านี้ว่า “เส้นนำทาง” จุดกำเนิดของใบหน้าตัวการ์ตูน ถือกำเนิดมาจากการที่นักเขียนลงมือขีดเขียนเส้นซ้อน กันลงไป ซึ่งรวมถึงเส้นร่างที่เป็นวงกลมและกากบาทด้วย เส้นเหล่านี้ทำหน้าที่เป็นเครื่องนำทางในการวาดเส้นโครงรอบนอกและองค์ประกอบโดยรวมบนใบหน้า ดังภาพที่ 2.11 (ศิริพร ชัยมานะ, 2555)



ภาพที่ 2.11 เส้นร่างวงกลมและกากบาท จาก การวาดใบหน้าขั้นพื้นฐาน, โดย ศิริพร ชัยมานะ, 2555

การวาดภาพใบหน้าตัวการ์ตูนที่เห็นในปัจจุบันมีอยู่ด้วยกันสี่ระดับ ได้แก่ การวาดใบหน้าการ์ตูนแบบดัดแปลงอย่างมาก การวาดใบหน้าการ์ตูนแบบดัดแปลงและประติษฐ์ การวาดใบหน้าการ์ตูนแบบดัดแปลงเล็กน้อย และการวาดใบหน้าการ์ตูนแบบเหมือนจริง ศิริพร อธิบายลักษณะของแต่ละระดับตามลำดับดังต่อไปนี้ (ศิริพร ชัยมานะ, 2555)

1. แบบดัดแปลงอย่างมาก ถือว่าเป็นการวาดการ์ตูนแบบง่าย โดยที่มีการวาดภาพดัดแปลงหน้าของตัวการ์ตูนเพื่อให้สามารถวาดได้ง่ายขึ้น การวาดการ์ตูนรูปแบบนี้ จะพบได้ในการ์ตูนเรื่องยาวที่มีคำบรรยายได้ภาพแบบสี่กรอบหรือภาพวาดเดี่ยวที่มีคำบรรยายประกอบ โดยที่ไม่เน้นรูปร่างหน้าตาตามความจริงทางกายวิภาค การ์ตูนลักษณะนี้จะมีดวงตาโตกว่าตาของมนุษย์ปกติ ไม่มีจมูก ดังภาพที่ 2.12



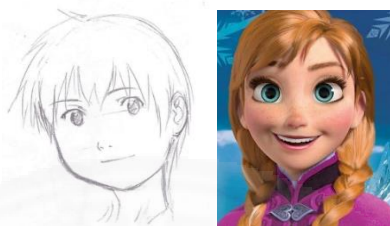
ภาพที่ 2.12 หน้าการ์ตูนแบบดัดแปลงอย่างมาก จาก การวาดใบหน้าขั้นพื้นฐาน, โดย ศิริพร ชัยมานะ, 2555

2. แบบดัดแปลงและประติษฐ์ รูปแบบนี้จะพบได้ในการ์ตูนสำหรับเด็ก การดัดแปลงตัวละครหรือทำให้ตัวละครวาดง่ายขึ้นทำให้สามารถสร้างความหลากหลายได้ไม่สิ้นสุดและเป็นเทคนิคที่เป็นพื้นฐานและใช้ได้ผลดีที่สุดในการแสดงออกถึงอารมณ์ให้มองเห็นได้ง่าย หน้าการ์ตูนลักษณะจะมีดวงตาที่กลมโตและมีประกายภายในตา ปรากฏเส้นโค้งไว้ในส่วนของการแสดงจมูก ดังภาพที่ 2.13



ภาพที่ 2.13 หน้าการ์ตูนแบบแปลงและประติษฐ์ จาก การวาดใบหน้าขั้นพื้นฐาน, โดย ศิริพร ชัยมานะ, 2555

3. แบบดัดแปลงเล็กน้อย คือ รูปแบบนี้พบได้ในการ์ตูนสำหรับผู้ใหญ่จำพวก การ์ตูนญี่ปุ่น การ์ตูนแอนิเมชัน เป็นต้น ซึ่งลักษณะในการวาดภาพการ์ตูนนั้นมีการดัดแปลงใบหน้าของตัวการ์ตูนเพียงเล็กน้อยเท่านั้น รายละเอียดของลายจะเส้นน้อยกว่าแบบเหมือนจริงเพียงเล็กน้อย ดังภาพที่ 2.14



ภาพที่ 2.14 หน้าการ์ตูนแบบดัดแปลงเล็กน้อย จาก *การวาดใบหน้าขั้นพื้นฐาน*, โดย ศิริพร ชัยมานะ, 2555

4. แบบเหมือนจริง คือ รูปแบบนี้จะใช้กับการวาดภาพเหมือนจริงในการ์ตูนแนวสัจนิยมที่มีเนื้อเรื่องแนวชีวิต มักใช้กับการ์ตูนแนวภาพยนตร์เล่าเรื่องราว การ์ตูนในลักษณะนี้แบบใบหน้าที่จะมีความคล้ายหรือเหมือนกับใบหน้าของมนุษย์เป็นอย่างมาก ดังรูปที่ 2.15

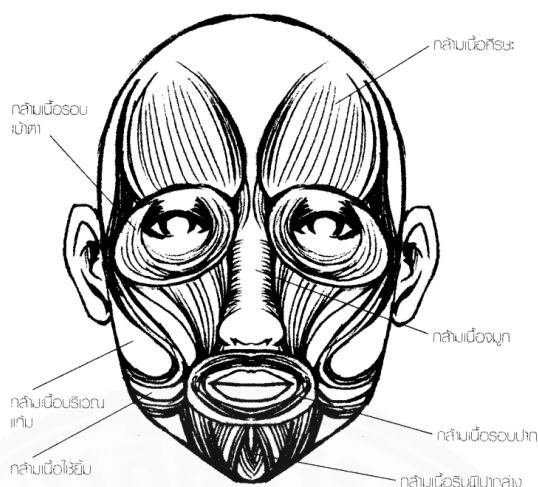


ภาพที่ 2.15 หน้าการ์ตูนแบบเหมือนจริง จาก *การวาดใบหน้าขั้นพื้นฐาน*, โดย ศิริพร ชัยมานะ, 2555

จากการศึกษาข้อมูลเบื้องต้นพบว่า ตัวการ์ตูนจากการ์ตูนภาพเคลื่อนไหว (Animated Cartoons) นั้นภาพใบหน้าของตัวละครในการ์ตูนมักมีส่วนประกอบและโครงสร้างบนใบหน้าที่มีความคล้ายหรือเหมือนมนุษย์ดังภาพที่ 2.13, 2.14 และ 2.15

2.2.3 เทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูน

การวาดสีหน้า มีความสุข เศร้า โกรธ จำเป็นต้องเข้าใจลักษณะการเคลื่อนไหวของอวัยวะบนใบหน้าให้เข้าใจ การถ่ายทอดอารมณ์จะแสดงผ่านการขยับกล้ามเนื้อใบหน้าอย่างไร ก่อนจะศึกษาวิธีการวาด ในอันดับแรกผู้วาดต้องทำความรู้จักเกี่ยวกับกล้ามเนื้อและมัดบนใบหน้าเสียก่อน (ปิยะบุตร สุทธิดาราร, 2558) ดังภาพที่ 2.16



ภาพที่ 2.16 กล้ามเนื้อบนใบหน้า จาก Manga Bible 1: ครบทุกพื้นฐาน การหัตถศาสตร์ตุน (พิมพ์ครั้งที่ 1), โดย ปิยะบุตร สุทธิตารา, 2558

ภาพกล้ามเนื้อและมัดบนใบหน้า ประกอบไปด้วย กล้ามเนื้อรอบเบ้าตา กล้ามเนื้อคิริษะ กล้ามเนื้อบริเวณแก้ม กล้ามเนื้อจมูก กล้ามเนื้อไช้ยืม กล้ามเนื้อรอบปาก และ กล้ามเนื้อริมฝีปากล่าง เมื่อทำความรู้จักกับกล้ามเนื้อทั้งหมดบนใบหน้าแล้ว ขั้นตอนต่อไปจะอธิบาย ในส่วนของกล้ามเนื้อส่งผลต่อการแสดงอารมณ์



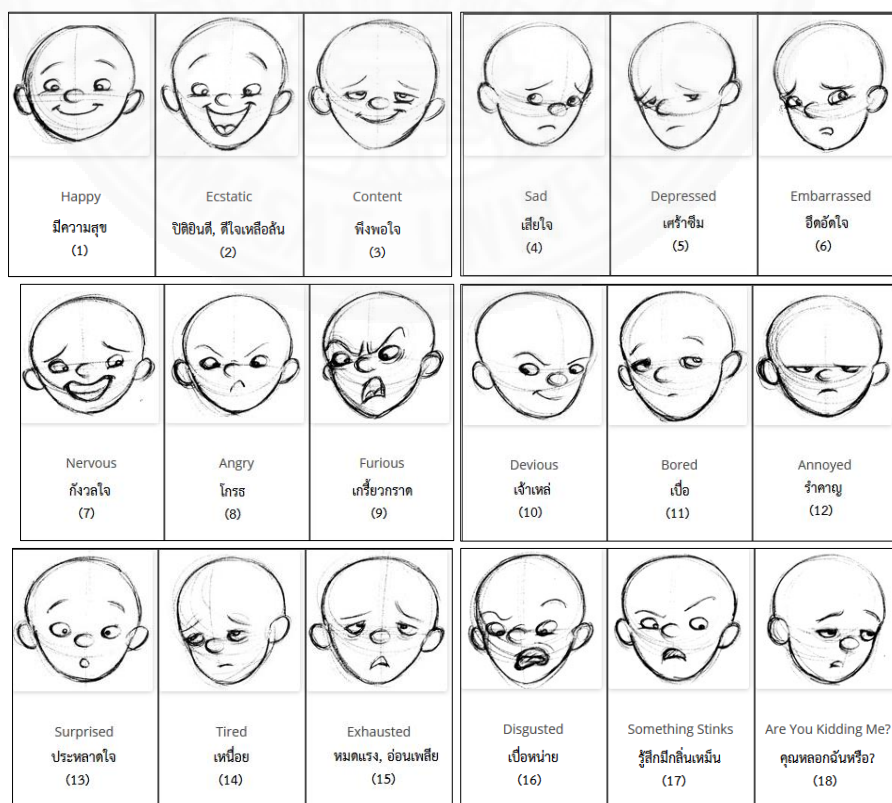
ภาพที่ 2.17 กล้ามเนื้อใบหน้าเมื่อตกใจ จาก Manga Bible 1: ครบทุกพื้นฐาน การหัตถศาสตร์ตุน (พิมพ์ครั้งที่ 1), โดย ปิยะบุตร สุทธิตารา, 2558

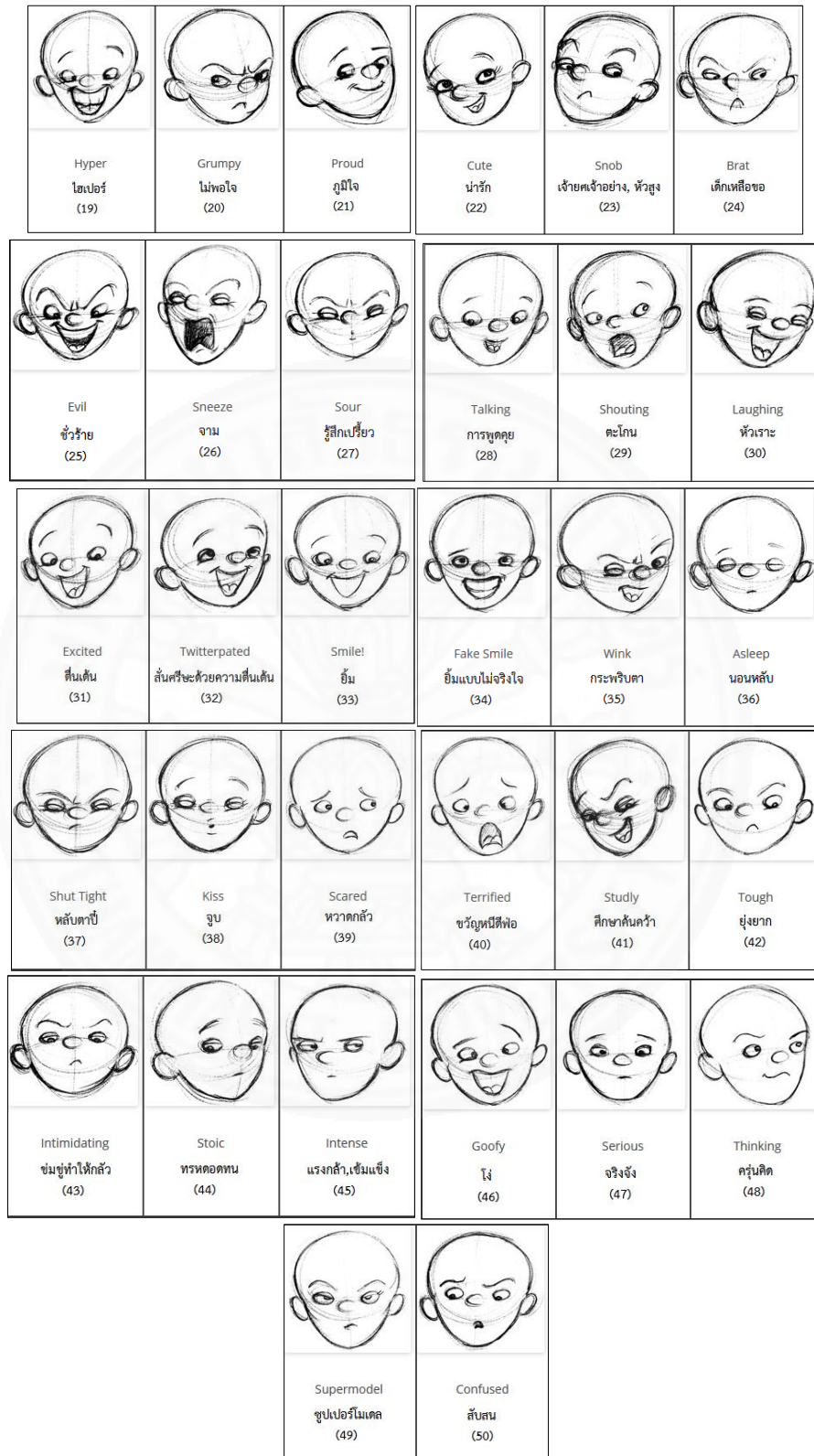
จากภาพที่ 2.17 เป็นภาพแสดงการขยับของกล้ามเนื้อและมัดบนใบหน้าเมื่อตกใจ กล้ามเนื้อคิริษะทั้งสองข้างจะยกสูงขึ้น กล้ามเนื้อรอบเบ้าตาจะเปิดกว้างและขยายมากกว่าปกติ กล้ามเนื้อบริเวณแก้ม กล้ามเนื้อจมูก กล้ามเนื้อไช้ยืม กล้ามเนื้อรอบปาก และกล้ามเนื้อริมฝีปากล่าง จะหดลง ซึ่งจากภาพทำให้เห็นถึงการเคลื่อนไหวของกล้ามเนื้อและมัดบนใบหน้า จึงทำให้สามารถที่จะรับรู้และเข้าใจอารมณ์ที่แสดงออกบนใบหน้าได้ (ปิยะบุตร สุทธิตารา, 2558)

Jones ได้อธิบายว่าการวาดภาพการแสดงออกทางสีหน้าของตัวการ์ตูน คุณสมบัติที่สำคัญของการวาดตัวการ์ตูนที่มีใบหน้าคล้ายมนุษย์จะต้องประกอบไปด้วย องค์ประกอบหลักสามประการดังนี้ (Jones, 2007)

1. ตาและคิ้ว เป็นคุณลักษณะที่สำคัญที่สุดของการวาดภาพสื่ออารมณ์ของตัวการ์ตูน เพราะ ผู้วาดส่วนมากจะใช้แววตาและคิ้วในการแสดงผลลัพธ์ทางอารมณ์เป็นหลัก
2. แก้ม สามารถวาดให้ขยายหรือหดขนาดได้ ซึ่งจะส่งผลถึงลักษณะตำแหน่งของดวงตาและคิ้วได้
3. ปาก รูปทรงของปากก็มีความสำคัญเป็นอย่างมาก ปากส่งผลถึงการขยายของแก้มและรูปร่างทั้งหมดของใบหน้าด้วย

ในการวาดภาพการ์ตูนให้คำนึงถึงว่าเมื่อมีการเคลื่อนย้ายรูปร่างหรือตำแหน่งของแต่ละพีเจอร์ (Feature) จะส่งผลต่อทุกอย่างบนภาพซึ่งไม่มีอะไรตายตัวแน่ชัด การวาดภาพตัวการ์ตูนที่ดีจะต้องมีการแสดงออกทางสีหน้าที่ชัดเจน เช่น วาดภาพตัวละครที่มีความสุขด้วยความเปี่ยมล้นหรือวาดภาพตัวละครที่กำลังโกรธอย่างเกรี้ยวกราด เป็นต้น ตัวอย่างดังภาพที่ 2.18 (2) และ ภาพที่ 2.18 (9) ตามลำดับ การวาดภาพต้องวาดให้เสมือนมีกระจกเงาสท้อนอยู่ เช่น เมื่อเราแสดงออกทางสีหน้า มักจะพบว่าใบหน้าของตัวเองเราจะมักจะเคลื่อนไหวในลักษณะที่แปลกโดยที่เราไม่รู้ตัว





ภาพที่ 2.18 การวาดภาพการแสดงออกทางสีหน้าของตัวการ์ตูน 50 อารมณ์ จาก *Tops for Drawing Facial Expressions*, โดย Jones, D., 2007.

งานวิจัยนี้จะมุ่งเน้นที่การวิเคราะห์อารมณ์สามอารมณ์ ได้แก่ อารมณ์มีความสุข (Happiness) อารมณ์เศร้า (Sadness) และอารมณ์โกรธ (Anger) เพราะอารมณ์ทั้งสามเป็นอารมณ์พื้นฐานของมนุษย์ (วิลโลว์รอน ศรีสงคราม, สุขัญญา รัตนัญญา, โรจน์รวี พจน์พัฒนาพล, พีรพล เทพประสิทธิ์, 2549), (ทรงพล ภูมิพัฒน์, 2538) และเป็นการแสดงออกที่พบบ่อยในภาพยนตร์หรือแอนิเมชัน

1. การวาดภาพสีหน้าอารมณ์มีความสุข (Happiness)

การวาดภาพสีหน้าของอารมณ์มีความสุขนั้นจะประกอบไปด้วยสามส่วน คือ คิ้ว ตา ปาก ในส่วนของการวาดคิ้วสำหรับอารมณ์นี้ คิ้วโก่ง กล้ามเนื้อศีรษะและกล้ามเนื้อรอบตาจะดันคิ้วให้สูง หาดตาจะตกลงเล็กน้อย ตาหรือครึ่งหนึ่ง กล้ามเนื้อที่ใช้ยิ้มจะขยายออก ทำให้ริมฝีปากโค้งขึ้น ยาวขึ้น ดังภาพที่ 2.19 ซึ่งจะแสดงสีหน้าของอารมณ์มีความสุขทั้งสามมุมมอง ประกอบไปด้วย หน้าตรง (ก) หน้าเอียงข้างและ (ค) หน้าหันข้าง (ปิยะบุตร สุทธิดาราร, 2558)



(ก) หน้าตรง (ข) หน้าเอียงข้าง (ค) หน้าหันข้าง

ภาพที่ 2.19 การวาดภาพสีหน้าอารมณ์มีความสุข จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัตวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดาราร, 2558

อารมณ์มีความสุขนั้นส่วนใหญ่จะแสดงออกด้วยการยิ้ม สำหรับการยิ้มนั้นมีอยู่ด้วยกันทั้งหมดสามระดับ ได้แก่ (ก) มีความสุขเล็กน้อย (ข) มีความสุขมาก (ค) มีความสุขที่สุด ดังภาพที่ 2.20 และตัวอย่างดังภาพที่ 2.21 (ปิยะบุตร สุทธิดาราร, 2558)



(ก) มีความสุขเล็กน้อย (ข) มีความสุขมาก (ค) มีความสุขที่สุด

ภาพที่ 2.20 ระดับการยิ้มของอารมณ์มีความสุข จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัตวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดาราร, 2558



(ก) มีความสุขเล็กน้อย (ข) มีความสุขมาก (ค) มีความสุขที่สุด

ภาพที่ 2.21 ตัวอย่างสีหน้ามีความสุขของตัวการ์ตูนในแต่ละระดับการยิ้ม จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัตวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดาราร, 2558

2. การวาดภาพสีหน้าอารมณ์เศร้า (Sadness)

การวาดภาพสีหน้าของอารมณ์เศร้าประกอบไปด้วยสามส่วน คือ คิ้ว ตา ปาก ในส่วนของการวาดคิ้วสำหรับอารมณ์นี้ หัวคิ้วจะยกขึ้น กล้ามเนื้อหน้าผากดึงเข้าหาระหว่างคิ้ว หัวคิ้วขมวด หาดตาจะตกขอบตามีน้ำตาคลอ กล้ามเนื้อที่ใช้ยิ้มจะดึงตัวลง มุมปากจึงเหยียดออกและตก ดังภาพที่ 2.22 ซึ่งจะแสดงสีหน้าของอารมณ์เศร้าทั้งสามมุมมอง ประกอบไปด้วย หน้าตรง (ข) หน้าเอียงข้างซ้ายและ (ค) หน้าหันข้าง (ปิยะบุตร สุทธิดาราร, 2558)



(ก) หน้าตรง (ข) หน้าเอียงข้าง (ค) หน้าหันข้าง

ภาพที่ 2.22 การวาดภาพสีหน้าอารมณ์เศร้า จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัตวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดาราร, 2558

อารมณ์เศร้านั้นมีทั้งแบบเก็บกดและระเบิดอารมณ์ออกมา หากมีการระเบิดอารมณ์ออกมาจะมีน้ำตาด้วย อารมณ์เศร้ามียู่ด้วยกันทั้งหมดสามระดับ ได้แก่ (ก) เสียใจ (ข) ร้องไห้ (ค) ปลั่งโฮ โฮ ดังภาพที่ 2.23 และตัวอย่างดังภาพที่ 2.24 (ปิยะบุตร สุทธิดาราร, 2558)



(ก) เสียใจ (ข) ร้องไห้ (ค) ปลั่งโฮโฮ

ภาพที่ 2.23 ระดับความเศร้าของอารมณ์เศร้า จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัตวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดาราร, 2558



(ก) เสียใจ



(ข) ร้องไห้



(ค) ปล่อยโฮ

ภาพที่ 2.24 ตัวอย่างสีหน้าเศร้าของตัวการ์ตูนในแต่ละระดับความเศร้า จาก *Manga Bible 1: ครบทุกพื้นฐาน การวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดารา, 2558

3. การวาดภาพสีหน้าอารมณ์โกรธ (Anger)

การวาดภาพสีหน้าของอารมณ์โกรธประกอบไปด้วยสามส่วน คือ คิ้ว แก้ม ปาก ในส่วนของการวาดคิ้วสำหรับอารมณ์นี้ หัวคิ้วกดลงหางคิ้วชี้ขึ้น กล้ามเนื้อหน้าผากย่นเข้าหาหว่างคิ้ว หัวคิ้วขมวดเข้าหากัน กล้ามเนื้อแก้มเคลื่อนเข้าหาระหว่างคิ้ว ตาจ้องเขม็ง กล้ามเนื้อที่ใช้ยึดดึงตัวขึ้น ทำให้มุมปากลดต่ำ ดึงริมฝีปากให้ยาวขึ้น ดังภาพที่ 2.25 ซึ่งจะแสดงสีหน้าของอารมณ์เศร้าทั้งสามมุมมอง ประกอบไปด้วย หน้าตรง (ข) หน้าเอียงข้างและ (ค) หน้าหันข้าง (ปิยะบุตร สุทธิดารา, 2558)



(ก) หน้าตรง



(ข) หน้าเอียงข้าง



(ค) หน้าหันข้าง

ภาพที่ 2.25 การวาดภาพสีหน้าอารมณ์โกรธ จาก *Manga Bible 1: ครบทุกพื้นฐาน การวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิดารา, 2558

อารมณ์โกรธนั้นเน้นที่คิ้วและตา กล้ามเนื้อจะขมวดระหว่างตาและคิ้ว ระดับความโกรธมีอยู่ด้วยกันทั้งหมดสามระดับ ได้แก่ (ก) ไม่พอใจ (ข) ประท้วง (ค) โกรธมาก ดังภาพที่ 2.26 และตัวอย่างสีหน้าโกรธของตัวการ์ตูนในแต่ละระดับความไม่พอใจดังภาพที่ 2.27 (ปิยะบุตร สุทธิดารา, 2558)



(ก) ไม่พอใจ



(ข) ประท้วง



(ค) โกรธมาก

ภาพที่ 2.26 ระดับความไม่พอใจของอารมณ์โกรธ จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัดวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิตารา, 2558



(ก) ไม่พอใจ



(ข) ประท้วง









(ค) โกรธมาก

ภาพที่ 2.27 ตัวอย่างสีหน้าโกรธของตัวการ์ตูนในแต่ละระดับความไม่พอใจ จาก *Manga Bible 1: ครบทุกพื้นฐาน การหัดวาดการ์ตูน (พิมพ์ครั้งที่ 1)*, โดย ปิยะบุตร สุทธิตารา, 2558

ตารางที่ 2.1 แสดงตัวอย่างภาพการแสดงอารมณ์บนใบหน้า สามอารมณ์หลัก ได้แก่ อารมณ์โกรธ อารมณ์เศร้า อารมณ์มีความสุข โดยเทียบระหว่างภาพใบหน้ามนุษย์และภาพการ์ตูนที่มีการตัดแปลงเล็กน้อย

ตารางที่ 2.1

ตารางแสดงการเปรียบเทียบอารมณ์การแสดงอารมณ์บนใบหน้าของมนุษย์และการ์ตูน

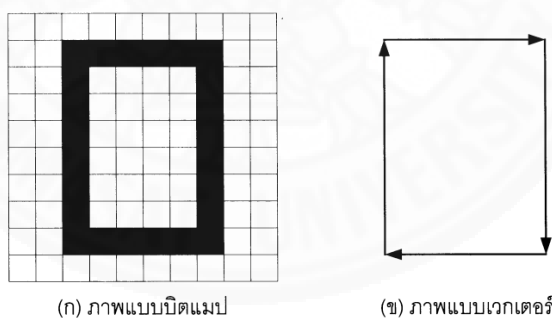
อารมณ์	มนุษย์	การ์ตูน
อารมณ์โกรธ		
อารมณ์เศร้า		
อารมณ์มีความสุข		

2.3 ความรู้พื้นฐานเกี่ยวกับการประมวลผลและวิเคราะห์ภาพดิจิทัล (Digital Image Processing & Analysis Fundamental)

งานวิจัยฉบับนี้ จำเป็นต้องศึกษาเกี่ยวกับทฤษฎีที่เกี่ยวข้องกับความรู้พื้นฐานในการประมวลผลภาพดิจิทัล ซึ่งประกอบไปด้วย รูปภาพดิจิทัล (Digital Image) ภาพระดับสีเทา (Grayscale Image) ภาพสี (Color Image) ซึ่งรายละเอียดในแต่ละทฤษฎีจะอธิบายแยกตามหัวข้อ ดังนี้

2.3.1 รูปภาพดิจิทัล (Digital Images)

ภาพที่มนุษย์มองเห็นได้ทั่วไปนั้นเกิดจากการที่แสงตกกระทบกับวัตถุแล้วเกิดการสะท้อนผ่านเลนส์บนดวงตาของมนุษย์เข้าสู่การแปลงและประมวลผลภาพในกลไกทางสมองของมนุษย์ ภาพดิจิทัลถูกพัฒนาขึ้นมาเพื่อช่วยมนุษย์ในการบันทึกจดจำเรื่องราว จากสิ่งแวดล้อมให้อยู่ในรูปแบบของสื่อทางอิเล็กทรอนิกส์ เช่น รูปถ่าย สิ่งพิมพ์ และพิมพ์เขียว เป็นต้น ภาพที่ใช้งานบนเครื่องคอมพิวเตอร์โดยทั่วไปสามารถแบ่งรูปแบบตามการใช้งานได้สองประเภท คือ รูปภาพแบบบิตแมป (Bitmap Image) และรูปภาพแบบเวกเตอร์ (Vector Image) ดังภาพที่ 2.28 ตามลำดับ (สมเกียรติ อุดมธรรษากุล, 2554)



ภาพที่ 2.28 ภาพแบบบิตแมปและภาพแบบเวกเตอร์ จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

1. ภาพบิตแมป (Bitmap Image)

ภาพบิตแมปเป็นภาพที่เกิดขึ้นมาจากการประกอบกันของจุดภาพขนาดเล็ก หรือจะเรียกว่าภาพที่เกิดจากพิกเซล (Pixel) จำนวนมากที่เรียงต่อกันจนเกิดเป็นภาพภาพหนึ่ง เพื่อให้เห็นภาพชัดเจน ให้นึกถึงการสร้างภาพบนตารางสี่เหลี่ยมขนาดเล็ก โดยทำการแต้มสีลงไปในช่องสี่เหลี่ยมแต่ละช่องเพื่อให้เกิดภาพที่สมบูรณ์ที่มีขนาดใหญ่ ภาพบิตแมปนี้จะมีจำนวนจุดขนาดเล็กบน

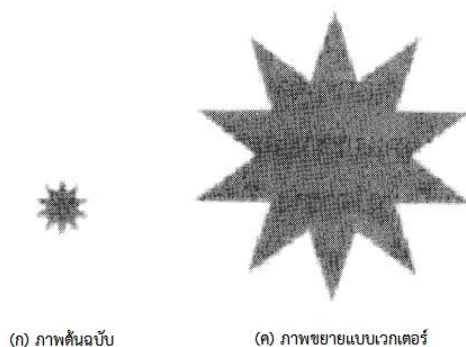
ภาพเป็นจำนวนมาก ดวงตาของมนุษย์ไม่สามารถมองเห็นและแยกแยะรายละเอียดย่อยของจุดภาพเหล่านี้ได้ หากลองขยายภาพก็จะทำให้มนุษย์สามารถมองเห็นรายละเอียดของจุดภาพนั้นได้ แต่มนุษย์ก็ไม่สามารถตอบได้ว่าภาพนั้นเป็นภาพของอะไร ดังภาพที่ 2.29 (ข) ภาพแบบบิตแมปนิยมใช้กันมากในภาพถ่ายหรือภาพวาด เนื่องจากภาพในลักษณะนี้สามารถไลโทนสีของแสงและเงาได้เหมือนจริงที่สุด (สมเกียรติ อุตมหาราชกุล, 2554)



ภาพที่ 2.29 ภาพขยายแบบบิตแมปเมื่อเทียบกับภาพต้นฉบับ จาก *การประมวลผลภาพดิจิทัล: Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุตมหาราชกุล, 2554

2. ภาพแบบเวกเตอร์ (Vector Image)

ภาพแบบเวกเตอร์จะมีคุณสมบัติที่แตกต่างกับภาพแบบบิตแมป คือ ภาพลักษณะนี้ไม่ว่าจะทำการขยายภาพให้ใหญ่ขนาดเท่าใดก็ตาม ภาพจะยังคงรายละเอียดและความคมชัดไว้เหมือนอย่างเดิมไม่มีการเปลี่ยนแปลงหรือผิดเพี้ยนไปจากภาพต้นฉบับ เนื่องจากภาพแบบเวกเตอร์นั้นจะประกอบไปด้วย เส้นโค้ง เส้นตรง และรูปทรงที่แตกต่าง ภาพที่ได้จะเกิดขึ้นมาจากการใช้คำสั่งที่บอกถึงลักษณะของภาพในรูปทรงเลขาคณิตด้วยสมการทางคณิตศาสตร์ จุดเด่นของแบบภาพเวกเตอร์ คือ ไม่ว่าจะทำการขยายภาพให้มีขนาดใหญ่เท่าไรก็ตาม คอมพิวเตอร์สามารถคำนวณค่าให้ได้ใหม่ทุกครั้ง ดังภาพที่ 2.30 (ค) ภาพลักษณะนี้จึงเหมาะกับการออกแบบโลโก้ การสร้างภาพสามมิติ เป็นต้น (สมเกียรติ อุตมหาราชกุล, 2554)



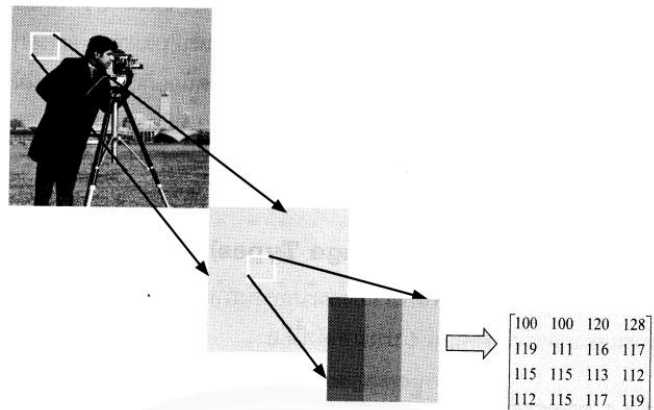
ภาพที่ 2.30 ภาพขยายแบบเวกเตอร์เมื่อเทียบกับภาพต้นฉบับ จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

2.3.2 ประเภทของภาพ (Image Types)

สามารถแบ่งประเภทของตามคุณสมบัติการแสดงผลของสีภาพ ออกเป็นสี่ประเภท ดังนี้

1. ภาพระดับความเข้มแสงเทาหรือภาพระดับเทา (Intensity Image or Gray Scale Image)

ภาพลักษณะนี้ ในแต่ละพิกเซลจะมีค่าความเข้มของแสงในระดับที่แตกต่างกันออกไป โดยจะแบ่งระดับความเข้มของแสงตั้งแต่ระดับสีเทาไปยังระดับสีขาว ภาพระดับสีเทาจะมีค่าความเข้มเท่ากับ 8 บิต (Bit) ค่าความเข้มของแสงจึงถูกแบ่งออกเป็น 256 ระดับ เมื่อค่าความเข้มเทามีค่าเป็น 0 ให้ความหมายว่าจุดภาพนั้นมีความเข้มของแสงต่ำ จุดภาพจะเป็นสีดำ ในทางกลับกัน หากค่าระดับความเข้มเทามีค่าเป็น 1 หมายความว่าจุดภาพนั้นมีความเข้มของแสงสูง จุดภาพที่ได้จะเป็นสีขาว ซึ่งแสงสีขาวจะถูกแทนด้วยค่าความเข้มเท่ากับ 255 (11111111) และสีดำจะถูกแทนค่าความเข้มเท่ากับ 0 (00000000) ส่วนค่าความเข้มแสงระหว่างจุดภาพสีดำกับจุดภาพสีขาวจะมีเฉดสีไล่จาก 0 – 255 ตัวอย่างภาพระดับความเข้มเทาดังภาพที่ 2.31 (สมเกียรติ อุดมธรรษากุล, 2554)



ภาพที่ 2.31 ภาพระดับความเข้มเทา จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมทรรษากุล, 2554

2. ภาพสี (Color Image)

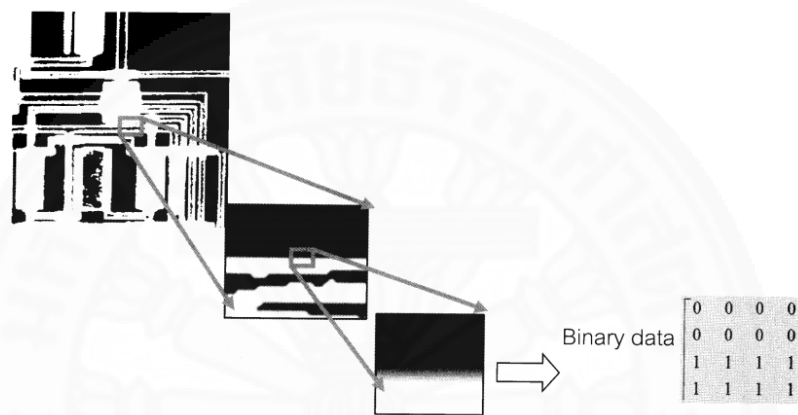
ภาพสี เป็นภาพที่ใช้ระดับความเข้มเทาของแต่ละแถบแสงของแม่สีหลักทั้ง 3 สีที่ซ้อนกันอยู่ประกอบไปด้วย แสงสีแดง (Red) แสงสีเขียว (Green) และแสงสีน้ำเงิน (Blue) ซึ่งในแต่ละพิกเซลนั้นจะแสดงผลของค่าสีแต่ละพิกเซลตามระดับความเข้มในแต่ละแถบแสงสีนั้น ดังภาพที่ 2.32 พิจารณาจุดภาพบนมุมซ้ายสุดของภาพที่ขยาย จะพบว่าระดับความเข้มของแถบ แสงสีแดง สีน้ำเงิน และสีเขียว จะมีค่าเท่ากับ 99, 65 และ 10 ตามลำดับ (สมเกียรติ อุดมทรรษากุล, 2554)



ภาพที่ 2.32 ภาพสี จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมทรรษากุล, 2554

3. ภาพไบนารี (Binary Image)

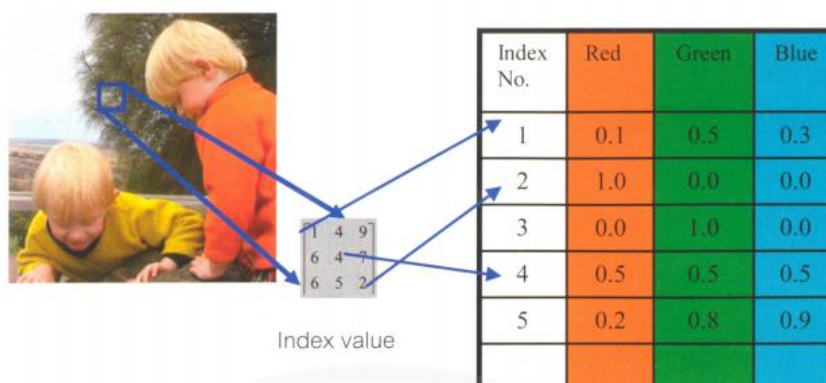
ภาพไบนารีจะแสดงข้อมูลของภาพในรูปแบบขาวดำ คือ ในแต่ละพิกเซลของภาพจะถูกแสดงด้วยค่าแบบไบนารี (Binary) ที่มี 1 บิต ซึ่งประกอบไปด้วยค่า 1 และ ค่า 0 โดยค่าที่เป็น 1 จะมีจุดภาพเป็นสีขาว และค่าที่เป็น 0 จะมีจุดภาพเป็นสีดำ ดังภาพที่ 2.33 ภาพไบนารีถูกนำไปใช้กับงานที่เกี่ยวกับตัวอักษร (Text) ภาพเกี่ยวกับลายนิ้วมือ (Finger Print) เป็นต้น (สมเกียรติ อุดมหาระชากุล, 2554)



ภาพที่ 2.33 ภาพไบนารีหรือภาพขาวดำ จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมหาระชากุล, 2554

4. ภาพแบบดัชนี (Index Image)

ภาพประเภทนี้ ในแต่ละพิกเซลของภาพจะเก็บค่าดัชนี (Index Number) ตัวเลขที่จัดเก็บจะเป็นตัวเลขจำนวนเต็ม ซึ่งค่าดัชนีดังกล่าวจะถูกนำไปเปรียบเทียบกับตารางสี (Color Table) ในตารางสีจะแสดงค่าของ แสงสีแดง แสงสีเขียว และแสงสีน้ำเงิน ค่าดัชนีจะเป็นตัวบ่งชี้ให้เห็นว่าภาพในแต่ละตำแหน่งพิกเซลใด ๆ มีค่าอัตราส่วนของแม่แสง 3 สี ในอัตราส่วนละเท่าใด ดังภาพที่ 2.34 จุดภาพที่เราพิจารณามีค่าดัชนีเท่ากับ 2 จากตารางสี จึงทำให้ทราบว่าจุดที่กำลังพิจารณาอยู่นั้นเป็นสีแดงสด เนื่องจากจุดภาพไม่มีปริมาณของแสงสีเขียวและแสงสีน้ำเงินอยู่เลย (สมเกียรติ อุดมหาระชากุล, 2554)



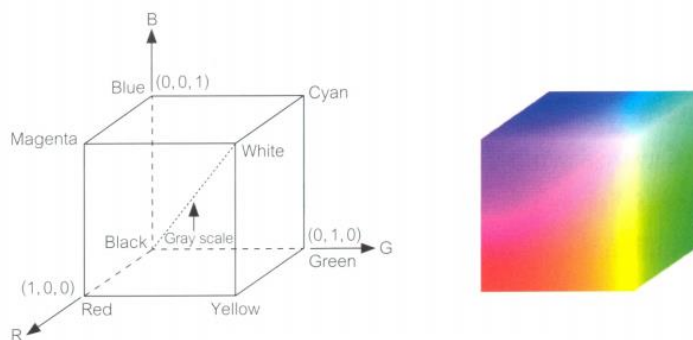
ภาพที่ 2.34 ภาพดัชนี จาก การประมวลผลภาพดิจิทัล: Fundamentals of Digital Image Processing, โดย สมเกียรติ อุดมทรรษากุล, 2554

2.3.3 แบบจำลองสี (Color Model)

แบบจำลองสีหรือเรียกอีกอย่างหนึ่งว่าโมเดลของสี โมเดลของสีนั้นถูกนำมาใช้กำหนดสีต่าง ๆ ให้เป็นไปตามแบบมาตรฐาน ซึ่งแบบจำลองสีในแต่ละแบบจะมีคุณสมบัติที่แตกต่างกันออกไป จึงทำให้เหมาะกับการใช้งานที่แตกต่างกันออกไป ในแบบจำลองสีนั้นจะใช้แม่สีหลัก ๆ มาผสมผสานกันเพื่อให้เกิดการแสดงค่าสีอื่น ๆ โดยแม่สีหลักนั้นจะแตกต่างกันไปตามแบบจำลองแต่ละชนิด (สมเกียรติ อุดมทรรษากุล, 2554)

1. แบบจำลองสี RGB

แบบจำลองสีนี้เป็นแบบจำลองของแม่สีหลัก (Primary Color) 3 สี เป็นสีที่เกิดจากการรวมกันของแสง (Additive Color) ซึ่งประกอบไปด้วยแม่สีที่สำคัญทั้ง 3 สี ดังนี้ คือ สีแดง (Red) สีเขียว (Green) และน้ำเงิน (Blue) แบบจำลองสีในลักษณะนี้ถูกนำมาใช้ในงานด้านการแสดงผลข้อมูลบนจอภาพ รวมไปถึงการเก็บข้อมูลภาพในระบบงานทางด้านคอมพิวเตอร์ แบบจำลองสีลักษณะนี้ เป็นแบบจำลองสีที่ใกล้เคียงกับการมองเห็นของมนุษย์มากที่สุด แต่ก็ยังมีข้อจำกัดในเรื่องการพิมพ์ภาพออกมาทางเครื่องพิมพ์ เพราะสีบางชนิดที่มนุษย์สามารถมองเห็นด้วยจอภาพนั้น สีบางชนิดไม่สามารถสร้างขึ้นมาจริงๆ ด้วยการใช้หมึกพิมพ์ได้ เราสามารถแสดงคุณสมบัติของการผสมสีในแบบจำลองนี้ได้ดังภาพที่ 2.35



ภาพที่ 2.35 แบบจำลองสี RGB จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

จากภาพแบบจำลองสี RGB สามารถอธิบายได้ดังนี้ จุดกำเนิด $(R, G, B) = (0, 0, 0)$ จะมีการแสดงค่าของสีเป็นสีดำ ส่วนตำแหน่งที่ $(1, 1, 1)$ จะมีการแสดงค่าของสีเป็นสีขาว ส่วนค่าในแต่ละแกนนั้นจะแสดงปริมาณความเข้มของแต่ละสี ในการแสดงสีที่เรียกว่าระดับสีเทาจะแสดงในเส้นทแยงมุมจากจุดกำเนิด $(0, 0, 0)$ ที่มีสีดำ มาที่ตำแหน่ง $(1, 1, 1)$ ที่เป็นสีขาว (สมเกียรติ อุดมธรรษากุล, 2554)

2. แบบจำลองสี CMY

แบบจำลองสี CMY จะมีแม่สีหลักอยู่ 3 สี ซึ่งประกอบไปด้วย สีเขียวน้ำเงิน (Cyan) สีม่วงแดง (Magenta) และสีเหลือง (Yellow) แบบจำลองสีนี้ นิยมใช้กับงานพิมพ์เป็นหลัก เพราะคุณสมบัติในการแสดงสีจะใช้ลักษณะการลบออก คือ ถ้าแสงสีขาวตกกระทบหมึกที่เป็นสีเหลือง (สีเหลืองเกิดจากการผสมระหว่างสีแดงกับสีเขียวในแบบจำลอง RGB) แสงที่สะท้อนออกมาจากสีเหลืองจะไม่มีสีน้ำเงิน เนื่องจากสีน้ำเงินจะถูกดูดซับเอาไว้โดยหมึกสีเหลือง (สีน้ำเงินถูกลบออกจากแสงสีขาว ทำให้มนุษย์มองเห็นแสงที่สะท้อนออกมาเป็นสีเหลือง) แต่เนื่องจากเมื่อสามสีหลัก ๆ ในแบบจำลองนี้ ไม่สามารถผสมออกมาแล้วเกิดสีดำได้ ดังนั้นแบบจำลองชนิดนี้จึงเพิ่มสีหลักเข้ามาอีกหนึ่งสี นั่นก็คือ สีดำ ซึ่งนิยมเรียกแบบจำลองนี้ว่า CMYK (K ย่อมาจาก Key Color) การแปลงแบบจำลอง RGB ให้อยู่ในรูปของแบบจำลอง CMY สามารถทำได้โดยการใช้สมการดังภาพที่ 2.36 โดยค่า RGB จะถูกกระทำด้วยการนอร์มอลไลซ์ (Normalization) ให้มีค่าอยู่ในช่วง $[0,1]$ ส่วนแบบจำลอง CMYK สามารถเขียนแทนได้ด้วยสมการดังภาพที่ 2.37

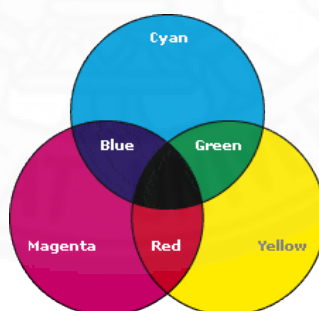
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

ภาพที่ 2.36 สมการในการแปลงค่าสี RGB ให้เป็น CMY จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

$$\begin{aligned} K &= \min(C, M, Y) \\ C &= C - K \\ M &= M - K \\ Y &= Y - K \end{aligned}$$

ภาพที่ 2.37 สมการของแบบจำลอง CMYK จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

จากภาพที่ 2.37 อธิบายได้ดังนี้ ค่า $\min(C, M, Y)$ คือค่าที่น้อยที่สุดเมื่อเทียบค่ากันระหว่าง ค่าสีเขียวน้ำเงิน (C) สีม่วงแดง (M) และ สีเหลือง (Y) ซึ่งสามารถแสดงภาพแบบจำลองสี CMY ได้ดังภาพที่ 2.38 (สมเกียรติ อุดมธรรษากุล, 2554)



ภาพที่ 2.38 ภาพโมเดลสี CMYK จาก ระบบสี (Color Model), โดย พิชญานานฎ ธีรรัช, 2553

3. แบบจำลอง HSV

เนื่องจากแบบจำลองสีที่กล่าวมาในข้างต้นสามารถแสดงถึงความสว่างในระดับหนึ่ง ๆ ได้ยากมาก เพื่อให้สามารถลดหรือเพิ่มค่าความสว่างของสีใดสีหนึ่งได้ จึงได้มีการออกแบบ แบบจำลอง HSV ขึ้นมา โดยมีองค์ประกอบหลักสามส่วนดังนี้ (สมเกียรติ อุดมธรรษากุล, 2554)

ก. H (Hue) คือ ค่าของแม่สีหลัก (สีแดง สีเขียว และสีน้ำเงิน) ในทางปฏิบัติ ค่าของสีจะอยู่ระหว่าง 0 – 1 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนที่ด้วยสีแดง และเมื่อ Hue มีค่าเพิ่มขึ้นเรื่อย ๆ สีก็จะเปลี่ยนแปลงไปตามเสปคตรัมของสีจนถึงค่า 1 จึงจะกลับมาเป็นสีแดงอีกครั้ง โดยแสดงตามตารางที่ 2.2 จากตารางจะเป็นค่าของระดับสีต่าง ๆ ตามค่าของ Hue ตั้งแต่ 0 – 1 ซึ่งรูปแบบสีต่าง ๆ ของ Hue จะถูกแสดงไว้ดังภาพที่ 2.39 (ก)

ตารางที่ 2.2

ค่าระดับสีของแบบจำลอง HSV

สี	ค่าระดับสี (Hue)
แดง	0
เหลือง	0.1667
เขียว	0.333
น้ำเงินเขียว	0.5
น้ำเงิน	0.6667
ม่วงแดง	0.8333

หมายเหตุ. จาก การประมวลผลภาพดิจิทัล, โดย สมเกียรติ อุดมทรรษากุล, 2554.

ข. S (Saturation) คือค่าความบริสุทธิ์ของสี ซึ่งค่า Saturation มีค่าเท่ากับ 0 แล้วสีที่ได้จะไม่มีความอิ่มตัว ซึ่งเป็นสีขาวล้วน แต่ถ้าค่า Saturation มีค่าเท่ากับ 1 แสดงว่าไม่มีแสงสีขาวผสมอยู่เลย หรือสามารถอธิบายได้ว่าค่าความอิ่มตัว (Saturation) คือปริมาณของสีที่ถูกทำให้เจือจางด้วยแสงสีขาว หากมีปริมาณสีขาวมากขึ้นเท่าไร ค่าความอิ่มตัวจะยิ่งมีน้อย เช่น สีแดงมีค่าความอิ่มตัวสูงกว่าสีชมพู ซึ่งรูปแบบค่าความอิ่มตัวที่มีค่าต่างระดับกันจะถูกแสดงในภาพที่ 2.39 (ข)

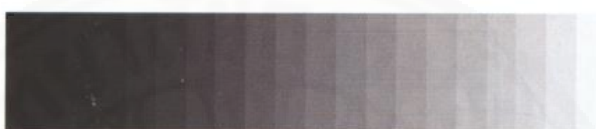
ค. V (Value) คือค่าความสว่างของสี ซึ่งสามารถวัดได้จากค่าความเข้มของความสว่างของแต่ละสีที่ประกอบกัน รูปแบบของค่าความสว่างที่มีค่าต่างระดับกันจะถูกแสดงไว้ดังภาพที่ 2.39 (ค)



(ก) รูปแบบสีต่างๆ ของ Hue



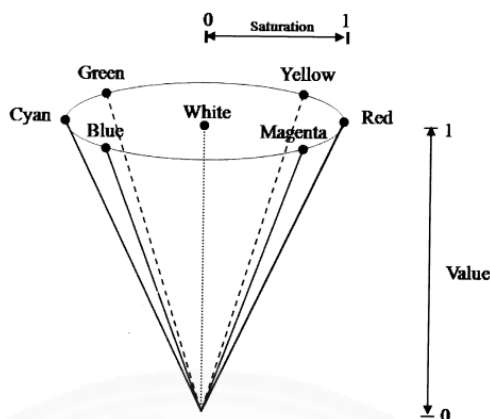
(ข) รูปแบบค่าความอิ่มตัวที่มีค่าต่างระดับกันของสีน้ำเงิน



(ค) รูปแบบของค่าความสว่างที่มีค่าต่างระดับกัน

ภาพที่ 2.39 รูปแบบของ Hue, Saturation และ Value จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมธรรษากุล, 2554

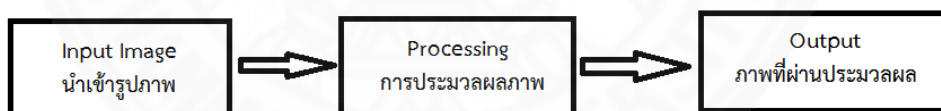
แบบจำลองนี้ความเป็นไปได้มีมากกว่าแบบจำลองอื่น ๆ เพราะความเข้มข้นของแต่ละสีนั้นไม่ได้ขึ้นอยู่กับตัวเนื้อสีเพียงอย่างเดียวแต่ยังเกี่ยวกับความเข้มของแสงสว่างด้วย ดังนั้นแบบจำลองนี้สามารถเพิ่มหรือลดความสว่างของสีใดสีหนึ่งได้และแบบจำลองนี้ค่อนข้างมีประโยชน์สำหรับการประมวลผลภาพเป็นอย่างมาก หลักการของแบบจำลองนี้ให้มองในมุมมองเสมือนมองลูกบาศก์ในแบบจำลอง RGB ในทิศทางขนานกับแกนระดับสีเทา จากตำแหน่งสีขาวเข้าไป เราสามารถมองภาพของแบบจำลองให้เป็นรูปทรงกรวย ดังภาพที่ 2.40 ซึ่งทุกจุดบนผิวหน้าของรูปได้แสดงถึงค่าความอิ่มตัว ส่วนในแกนรัศมีบ่งบอกถึงค่าความอิ่มตัวของสี ส่วนแกนความสูงจะเป็นแกนบอกถึงค่าความสว่าง สุดท้ายมุมมองของแกนรัศมีจะบอกถึงค่าของสีใด ๆ (สมเกียรติ อุดมธรรษากุล, 2554)



ภาพที่ 2.40 แบบจำลองสี HSV จาก การประมวลผลภาพดิจิทัล: *Fundamentals of Digital Image Processing*, โดย สมเกียรติ อุดมทรรษากุล, 2554

2.3.4 การประมวลผลและวิเคราะห์ภาพ (Image Processing and Analysis)

การประมวลผลภาพ (Image Processing) หมายถึง การกระทำอย่างใดอย่างหนึ่งกับภาพต้นฉบับ (Input Image) ให้ได้ภาพผลลัพธ์ที่ต้องการ (Output Image) เพื่อนำไปทำการวิเคราะห์และประมวลผลหาสิ่งที่ต้องการในลำดับถัดไป (สมเกียรติ อุดมทรรษากุล, 2554) ดังภาพที่ 2.41



ภาพที่ 2.41 ขั้นตอนในการประมวลผลภาพ

โดยมีขั้นตอนที่สำคัญ คือ การทำให้ภาพมีความคมชัดมากขึ้น (Contrast Stretching) การกำจัดสัญญาณรบกวนออกจากภาพ (Noise Filtering) การแบ่งส่วนของวัตถุที่เราสนใจออกมาจากภาพ (Segmentation) เพื่อนำภาพวัตถุที่ได้ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ จากนั้นเราสามารถนำข้อมูลเชิงปริมาณเหล่านี้ไปวิเคราะห์ และสร้างเป็นระบบ เพื่อใช้ประโยชน์ในงานด้านต่าง ๆ เช่น ระบบรู้จำลายนิ้วมือเพื่อตรวจสอบว่าภาพลายนิ้วมือที่มีอยู่นั้นเป็นของผู้ใด ระบบตรวจสอบคุณภาพของผลิตภัณฑ์ในกระบวนการผลิตของโรงงานอุตสาหกรรม ระบบคัดแยกเกรดหรือคุณภาพของพืชผลทางการเกษตร ระบบอ่านรหัสไปรษณีย์อัตโนมัติ เพื่อคัดแยกปลายทางของจดหมายที่มีจำนวนมากในแต่ละวันโดยใช้

ภาพถ่ายของรหัสไปรษณีย์ที่อยู่บนซอง ระบบเก็บข้อมูลรถที่เข้าและออกอาคารโดยใช้ภาพถ่ายของป้ายทะเบียนรถเพื่อประโยชน์ในด้านความปลอดภัย ระบบดูแลและตรวจสอบสภาพการจราจรบนท้องถนนโดยการนับจำนวนรถบนท้องถนนในภาพถ่ายด้วยกล้องวงจรปิดในแต่ละช่วงเวลา ระบบรู้จำใบหน้าเพื่อเฝ้าระวังผู้ก่อการร้ายในอาคารสถานที่สำคัญหรือในเขตคนเข้าเมือง ดังภาพที่ 2.42 (เศรษฐพงศ์ มะลิสุวรรณ, 2552)



ภาพที่ 2.42 รูปแสดงการประมวลผลภาพเฉพาะส่วนที่สนใจเพื่อนำข้อมูลไปพัฒนาต่อ จาก *Image processing เทคโนโลยีการประมวลผลภาพ*, โดย เศรษฐพงศ์ มะลิสุวรรณ, 2552

จะเห็นได้ว่าระบบเหล่านี้จำเป็นต้องมีการประมวลผลภาพจำนวนมาก และเป็นกระบวนการที่ต้องทำซ้ำกันในรูปแบบเดิมเป็นส่วนใหญ่ ซึ่งงานในลักษณะเหล่านี้ หากให้มนุษย์วิเคราะห์เอง มักต้องใช้เวลามากและใช้แรงงานสูง อีกทั้งหากจำเป็นต้องวิเคราะห์ภาพเป็นจำนวนมาก ผู้วิเคราะห์ภาพเองอาจเกิดอาการล้า ส่งผลให้เกิดความผิดพลาดขึ้นได้ ดังนั้นคอมพิวเตอร์จึงมีบทบาทสำคัญในการทำหน้าที่เหล่านี้แทนมนุษย์ อีกทั้ง เป็นที่ทราบโดยทั่วกันว่า คอมพิวเตอร์มีความสามารถในการคำนวณและประมวลผลข้อมูลจำนวนมากได้ในเวลาอันสั้น จึงมีประโยชน์อย่างมากในการเพิ่มประสิทธิภาพการประมวลผลภาพและวิเคราะห์ข้อมูลที่ได้จากภาพในระบบดังกล่าวข้างต้น

2.4 การวิเคราะห์ภาพด้วยคุณสมบัติการแยกลักษณะเด่นของวัตถุ (Feature Extraction for Image Analysis)

ภาพจะถูกวิเคราะห์หาลักษณะเฉพาะที่สำคัญ (Feature) เพื่อหาตัวแทนของข้อมูลสำหรับใช้แสดงถึงภาพนั้น ๆ ซึ่งการหาลักษณะเฉพาะของภาพ ในงานวิจัยนี้จำเป็นต้องศึกษาทฤษฎีเกี่ยวกับการวิเคราะห์รูปภาพโดยใช้คุณสมบัติการแยกลักษณะเฉพาะของวัตถุ (Feature Extraction)

เข้ามาใช้หาคุณลักษณะเฉพาะหลัก เพื่อนำมาใช้ในการวิเคราะห์ลักษณะเฉพาะของรูปภาพการ์ตูน และนำมาใช้ในการวิเคราะห์อารมณ์ของการ์ตูนได้อย่างอัตโนมัติ ซึ่งวิธีการแทนรูปภาพด้วย ลักษณะเฉพาะของภาพโดยทั่วไปที่นิยมใช้คือ คุณสมบัติของสี (Color) ลักษณะพื้นผิว (Texture) รูปทรง (Shape) และลักษณะเฉพาะที่ได้จากการกระจายตัวของเส้นขอบ (Histograms of Oriented Gradients: HOG) เป็นต้น ซึ่งจะอธิบายแยกตามหัวข้อดังต่อไปนี้ (จิรา แก้วสุวรรณ, 2549)

2.4.1 คุณลักษณะเกี่ยวกับสี (Color - Based Features)

ถือเป็นคุณลักษณะที่มนุษย์สามารถมองเห็นได้เด่นชัดที่สุด (Arnheim, 1974) แบบจำลองสี คือ โครงสร้างที่กำหนดขึ้นเพื่อจัดระบบสีให้เป็นระเบียบ โดยมีมาตราส่วน (Scale) บอกระยะ เพื่อช่วยให้จำแนกและควบคุมการเปลี่ยนแปลงสีให้ง่ายขึ้น อีกทั้งยังถูกสร้างขึ้นมาเพื่อช่วยในการกำหนดสีให้เป็นไปตามมาตรฐานเดียวกัน แบบจำลองสีที่ใช้อยู่ในปัจจุบันได้สร้างขึ้นมาเพื่อรองรับอุปกรณ์หรือโปรแกรมประยุกต์ โดยแบบจำลองสีที่นิยมใช้ ได้แก่ แบบจำลองสี RGB ซึ่งเป็นรูปแบบสำหรับใช้ในงานแสดงผลทางหน้าจอแสดงผล (Monitor) ของกล้องถ่ายภาพ เป็นต้น แบบจำลองสี HSV เป็นแบบจำลองสีที่สามารถแยกองค์ประกอบสีออกจากค่าความเข้มแสงได้ แบบจำลองสีแบบ HSV จึงใกล้เคียงกับการรับรู้และแปลความหมายสีของมนุษย์มากที่สุด (ประภาพร กุลลิ้มรัตน์ชัย, 2556)

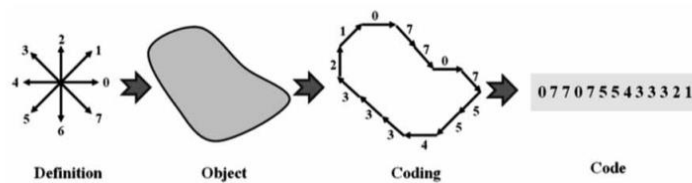
2.4.2 คุณลักษณะเกี่ยวกับพื้นผิว (Texture - Based Features)

มีผู้ให้นิยามคุณลักษณะพื้นผิวที่ต่างกันออกไปโดย Parker (Parker, 1997) กล่าวว่าลักษณะเด่นของการเป็นพื้นผิว คือ การมีรูปแบบที่ซ้ำกันในบริเวณหนึ่ง ๆ อย่างไรก็ตามแม้ว่าพื้นผิวนั้นจะเป็นพื้นผิวสี แต่ในการวิเคราะห์พื้นผิวจะเน้นการทำงานกับพื้นผิวที่เป็นระดับสีเทาเท่านั้น Tamura และคณะ (Tamura, Mori & Yamawaki, 1987) ได้ค้นพบความสัมพันธ์ระหว่างการรับรู้ทางการมองเห็นของมนุษย์ และคำนวณหาความแตกต่างระหว่าง พื้นผิว 6 คุณลักษณะ ได้แก่ ความหยาบ (Coarseness) ความตรงกันข้าม (Contrast) ทิศทาง (Directionality) ความเหมือนกันของเส้น (Line-Likeness) ความสม่ำเสมอ (Regularity) และความขรุขระ (Roughness) (ประภาพร กุลลิ้มรัตน์ชัย, 2556)

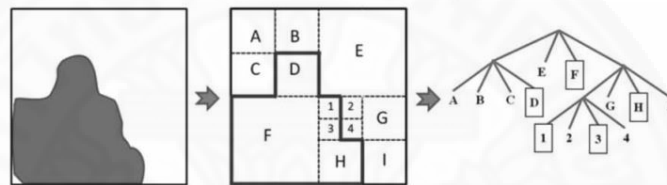
2.4.3 คุณลักษณะเกี่ยวกับรูปร่าง (Shape - Based Features)

วิธีที่ใช้ในการวิเคราะห์รูปร่างของวัตถุที่ปรากฏในภาพนั้น มีการใช้อยู่สองวิธี (Watt & Policarpo, 1998) วิธีแรกคือแบบขอบเขต (Boundary - Based) และวิธีที่สอง คือแบบบริเวณ (Region - Based) ซึ่งทั้งสองวิธีนี้จะให้หลักการเข้ารหัสภาพ (image coding scheme) เพื่อ

แทนเนื้อหาของภาพ หลักการวิเคราะห์รูปร่างของทั้งสองวิธีแสดงดังภาพที่ 2.43 (ประภาพร กุลลิมรัตน์ชัย, 2556)



(ก) แบบขอบเขต (Boundary-Based)

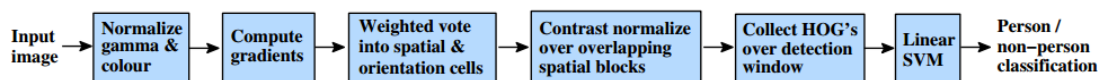


(ข) แบบบริเวณ (Region-Based)

ภาพที่ 2.43 การเข้ารหัสของวัตถุในภาพ จาก *ความรู้พื้นฐานเกี่ยวกับการค้นคืนภาพ Image Retrieval Fundamentals*, โดย ประภาพร กุลลิมรัตน์ชัย, 2556, วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเชีย ปีที่ 7 ฉบับที่ 2 ประจำเดือน กรกฎาคม - ธันวาคม

2.4.4 ค่าความถี่ของทิศทางตามค่าเกรเดียนต์ (Histograms of Oriented Gradients: HOG)

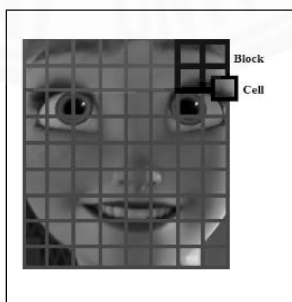
ภาพรวมของการหาค่าความถี่ของทิศทางตามค่าเกรเดียนต์ หรือเรียกว่า HOG มีด้วยกันทั้งหมดหกขั้นตอน ดังภาพที่ 2.44



ภาพที่ 2.44 ภาพรวมของการสกัดหาคุณลักษณะเฉพาะในการตรวจจับวัตถุ จาก *Histograms of Oriented Gradients for Human Detection*, โดย Dalal, N., & Triggs, B. 2005, NRIA Rhone-Alps, 655 avenue de l'Europe, Montbonnot 38334, France

จากภาพที่ 2.44 Dalal และ Triggs อธิบายว่าหน้าต่างในการตรวจจับจะเรียงต่อกันเป็นแผ่นที่ทับซ้อนกัน ซึ่งใช้ค่าความถี่ตามทิศทางเกรเดียนต์มาเป็นคุณสมบัติในการจำแนกคุณลักษณะเฉพาะ การรวมตัวกันของเวกเตอร์จะใช้ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) จำแนกประเภทที่เป็นวัตถุและไม่ใช่วัตถุ หน้าต่างในการตรวจจับจะถูกแสกนทุกตำแหน่งของและขนาดของภาพ Dalal และ Triggs มุ่งเน้นไปที่ขั้นตอนในการสกัดคุณลักษณะเฉพาะของภาพ (Dalal, N., & Triggs, B., 2005) โดยอธิบายว่าเมื่อรับภาพเข้ามาในระบบ ขั้นตอนแรก ทำการนอร์มอลไลซ์จากค่าความสว่าง (Gamma) และสี (Colour) ขั้นตอนที่สอง ทำการคำนวณค่าเกรเดียนต์ ขั้นตอนที่สาม ทำการโหวตค่าถ่วงน้ำหนักของระยะห่างที่สอดคล้องกันของเซลล์ (Cells) ขั้นตอนที่สี่ ทำการนอร์มอลไลซ์พื้นที่ที่ทับซ้อนกันของบล็อก (Block) ขั้นตอนที่ห้า รวบรวมหน้าต่างของ HOG สำหรับการตรวจจับทั้งหมด ขั้นตอนที่หก ทำการจำแนกคุณลักษณะเฉพาะด้วยซัพพอร์ตเวกเตอร์แมชชีนแบบเชิงเส้น (Linear)

ค่าความถี่ของทิศทางตามค่าเกรเดียนต์ เป็นวิธีการดึงคุณลักษณะเฉพาะของวัตถุ ซึ่งสามารถดึงรูปร่างภายในภาพโดยใช้การกระจายตัวของความเข้มเกรเดียนต์หรือทิศทางของเส้นขอบ การดึงคุณลักษณะของ HOG จะทำได้โดยการแบ่งภาพออกเป็นส่วนย่อยขนาดเล็ก หรือ เรียกว่า เซลล์ (Cells) โดยแต่ละเซลล์จะรวบรวมฮิสโตแกรมของทิศทางเกรเดียนต์หรือทิศทางของขอบภายในเซลล์ที่มีขนาดหนึ่งมิติ (1-D) โดยที่จะมีการรวมฮิสโตแกรมนั้นเข้าด้วยกัน เพื่อแสดงถึงคุณลักษณะเฉพาะของวัตถุที่สนใจ เพื่อให้มีประสิทธิภาพของความถูกต้องเพิ่มมากขึ้น สามารถนำฮิสโตแกรมมาทำนอร์มอลไลซ์ด้วยการคำนวณตัวชี้วัดของค่าความเข้มทั่วทั้งพื้นที่ขนาดใหญ่ของภาพ หรือ เรียกอีกอย่างว่าบล็อก (Block) (ชุมพล เสนาพันธ์ และ คำรณ สุนด์, 2556) ดังภาพที่ 2.45



ภาพที่ 2.45 ตัวอย่างการแบ่งภาพเป็นเซลล์และบล็อก

การทำนอร์มอลไลซ์ ต้องทำทุกเซลล์ภายในบล็อก ผลลัพธ์จากการทำนอร์มอลไลซ์ จะทำให้ผลกระทบจากการเปลี่ยนแปลงของแสงสว่างและเงาเล็กน้อย และสามารถหาคุณลักษณะเฉพาะของวัตถุได้ดีมากขึ้น โดยขั้นตอนของ HOG ดังนี้

1. คำนวณค่าเกรเดียนต์แนวนอนจากสมการที่ (1)

$$D_H = [-1 \ 0 \ 1] \quad (1)$$

2. ทำการคำนวณค่าเกรเดียนต์แนวตั้งจากสมการที่ (2)

$$D_V = [-1 \ 0 \ 1]^T \quad (2)$$

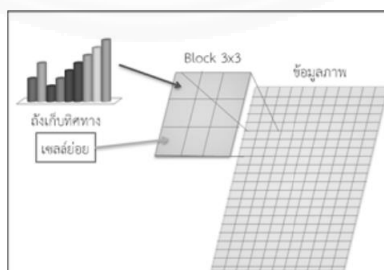
3. ทำการคำนวณหาขนาดของเกรเดียนต์จากสมการที่ (3)

$$M_G(X, Y) = \sqrt{D_H(X, Y)^2 + D_V(X, Y)^2} \quad (3)$$

4. ทำการคำนวณมุมทิศทางเกรเดียนต์จากสมการที่ (4)

$$O_G(X, Y) = \tan^{-1} \left(\frac{D_V(X, Y)}{D_H(X, Y)} \right) \quad (4)$$

5. เก็บค่าความถี่ของทิศทางเกรเดียนต์ โดยการเก็บทิศทางนั้นจะใช้บล็อกขนาด 3×3 เซลล์ โดยบล็อก จะประกอบไปด้วยเซลล์ย่อยทั้งหมด 9 เซลล์ ดังภาพที่ 2.46 แต่ละเซลล์จะมีถึงเก็บทิศทาง 8 bins โดยแต่ละ bins จะเก็บค่าทิศทางของมุม



ภาพที่ 2.46 เซลล์และบล็อกบริเวณที่ทำการตรวจจับหา HOG จาก การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนต์ร่วมกับการวิเคราะห์องค์ประกอบหลักและซัพพอร์ตเวกเตอร์แมชชีน, โดย อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร, 2557, การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 10

2.4.5 ลักษณะเฉพาะที่ดี (Good Features)

ลักษณะเฉพาะที่ดีต้องมีความสามารถในการจำแนก (High Discrimination Power) ช่วยในการจำแนกวัตถุแต่ละชนิดออกจากกันได้ ต้องไม่ควรขึ้นกับตำแหน่งในภาพ ขนาด การหมุนของภาพและการสะท้อนกลับของภาพ (Invariant to translation, scaling, rotation and reflection) มีความคงทน, คงเส้นคงวา (Stability) มีค่าที่แน่นอน และควรจัดการได้ง่าย (Ease of implement) ไม่มีความยุ่งยากซับซ้อน ดังภาพที่ 2.47 (Gonzalez, R. C. & Woods, E. R., 2009)

ลำดับ	Object#	#Feature1	#Feature2	#Feature3
1	ก	5	40	10
2	ฮ	100	41	5
3	ก	10	39	70
4	ฮ	98	40	68

ภาพที่ 2.47 ลักษณะของฟีเจอร์ที่ดี (Good Feature)

จากรูป มีวัตถุ (Object) อยู่สี่ตัว ได้แก่ ก และ ฮ อย่างละสองตัว และมีฟีเจอร์ทั้งหมดสามชนิด #Feature1 เป็นฟีเจอร์ที่ดี มีค่าที่คงที่โดยสังเกตได้จาก Stability ที่มีค่าไม่แตกต่างกันมากของ เมื่อมีการเปรียบเทียบกับตัวอักษร ก ใน Object 1 กับ Object 3 และ ตัวอักษร ฮ ใน Object 2 กับ Object 4 ในส่วนของ #Feature2 ค่า Discrimination Power ไม่ดี สังเกตได้จาก Object 1 - 4 ไม่สามารถจะจำแนกแยกแยะออกจากกันได้เพราะมีค่าของตัวเลขที่ใกล้เคียงกันมาก และ #Feature3 เป็นฟีเจอร์ที่ไม่ค่อยดีเพราะมีค่า Stability ของ Object 1 กับ Object 2 และ Object 3 กับ Object 4 มีค่าไม่มีแตกต่างกันจึงทำให้ไม่สามารถแยกแยะตัวอักษร ก ออกจาก ฮ ได้ เป็นต้น

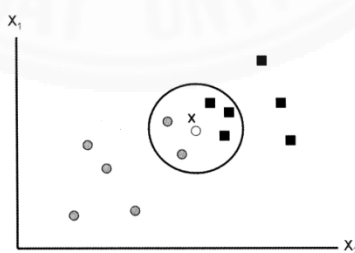
2.5 การจำแนกประเภทของข้อมูล (Classification)

การจำแนกประเภทของข้อมูล เป็นงานที่สำคัญงานหนึ่งในวิเคราะห์อัตโนมัติโดยหลังจากทำการสกัดคุณลักษณะหรือฟีเจอร์ของภาพออกมาได้แล้ว จะนำฟีเจอร์เหล่านั้นมาเป็นข้อมูล

ข่าวของตัวจำแนก (Classifier) เพื่อให้ระบบทำนายว่าภาพนั้นควรจัดอยู่ในประเภทใด ตัวจำแนกประเภทข้อมูลมีหลายรูปแบบ เช่น ต้นไม้ตัดสินใจ (Decision Tree) เครือข่ายประสาท (Neural Network) ตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors) ตัวแบบจำแนกประเภทแบบเบย์ส (Bayes Classifier) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เป็นต้น การสร้างตัวแบบจำแนกประเภทข้อมูลจะพยายามให้เกิดความผิดพลาดในการจำแนกให้น้อยที่สุดบนข้อมูลเรียนรู้ (Training Data) อย่างไรก็ตาม เมื่อนำตัวแบบจากการเรียนรู้ไปใช้ในการจำแนกประเภทข้อมูลที่ไม่เคยเห็นมาก่อน ตัวแบบอาจจะจำแนกข้อมูลไม่ดีเท่ากับการใช้ข้อมูลเรียนรู้ จึงต้องมีการใช้ข้อมูลอีกชุดหนึ่งเพื่อมาใช้ในการตรวจสอบ เรียกข้อมูลชุดนี้ว่า ข้อมูลทดสอบ (Testing Data) ซึ่งจะอธิบายประเภทของการจำแนกประเภทข้อมูลดังต่อไปนี้ (สุรพงศ์ เอื้อวัฒนามงคล, 2559)

2.5.1 ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors)

วิธีการนี้ เป็นวิธีการจำแนกประเภทของข้อมูลโดยการเปรียบเทียบข้อมูลกับข้อมูลตัวอย่างที่เก็บไว้ในฐานข้อมูล คือ ข้อมูลที่ถูกเก็บไว้เป็นตัวอย่างในฐานข้อมูล เมื่อต้องจำแนกประเภทข้อมูลใหม่ จะค้นหาตัวอย่างของข้อมูลในฐานข้อมูลที่มีลักษณะใกล้เคียงข้อมูลใหม่มากที่สุด โดยจำนวน k ตัว มาใช้ในการตัดสินใจว่าข้อมูลใหม่นี้ควรมีประเภทใดบ้าง การจำแนกประเภทลักษณะนี้จะใช้หลักการที่ว่า ข้อมูลใหม่ที่ต้องการจะทราบประเภทมีลักษณะใกล้เคียงกับข้อมูลใดในฐานข้อมูล ข้อมูลใหม่นั้นก็ควรมีประเภทเช่นเดียวกับประเภทของข้อมูลที่มีลักษณะใกล้เคียงกับข้อมูลนั้น การตัดสินใจว่าข้อมูลใหม่จะมีประเภทใด อาจใช้วิธีการของ Majority Vote คือ ข้อมูลใหม่จะมีประเภทเดียวกับข้อมูลที่มีจำนวนมากที่สุดในโดยอ้างอิงจากข้อมูลเพื่อนบ้าน K ตัวนั้น ดังภาพที่ 2.48 (สุรพงศ์ เอื้อวัฒนามงคล, 2559)

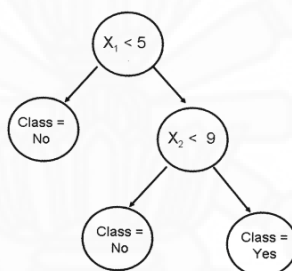


ภาพที่ 2.48 การจำแนกข้อมูล X โดยใช้วิธี K - Nearest Neighbors จาก การทำเหมืองข้อมูล, โดยสุรพงศ์ เอื้อวัฒนามงคล, 2559, สถาบันบัณฑิตพัฒนบริหารศาสตร์, สำนักพิมพ์สถาบันบัณฑิตพัฒนบริหารศาสตร์

จากภาพที่ 2.48 ข้อมูลที่ต้องการจำแนกประเภทคือ X โดยการใช้วิธี k - Nearest Neighbors มาช่วยในการจำแนก โดยกำหนดให้ $k = 5$ จะเห็นได้ว่า ข้อมูลเพื่อนบ้านที่ใกล้ที่สุด 5 ตัว มีประเภทที่แสดงด้วยจุดสีเหลี่ยมเป็นส่วนใหญ่ (3 ใน 5) ดังนั้นข้อมูล X ควรจะมีประเภทเดียวกับเพื่อนบ้านส่วนใหญ่โดยจากภาพ คือ แบบจุดสีเหลี่ยม

2.5.2 ต้นไม้การตัดสินใจ (Decision Tree)

วิธีการนี้จะทำการจำแนกข้อมูลเป็นรูปแบบต้นไม้การตัดสินใจ แต่ละโหนดของต้นไม้จะประกอบไปด้วยเงื่อนไขที่ใช้ตัวแปรใดตัวแปรหนึ่งของข้อมูลในการตัดสินใจเลือกโหนดลูกใดเพื่อใช้เป็นโหนดต่อไปในการตัดสินใจ การตัดสินใจจะเริ่มจากโหนดราก (Root) ของต้นไม้และไล่ไปยังโหนดลูกจนถึงโหนดใบ ซึ่งเงื่อนไขของแต่ละโหนดจะสามารถบอกประเภทของข้อมูลนั้นได้ ตัวอย่างการสร้างข้อมูลต้นไม้การตัดสินใจ ดังภาพที่ 2.49 โดยมีตัวแปรคือ X_1 และ X_2 และข้อมูลอาจจะมีประเภทเป็น Yes หรือ No ก็ได้ (สุรพงษ์ เอื้อวัฒนามงคล, 2559)

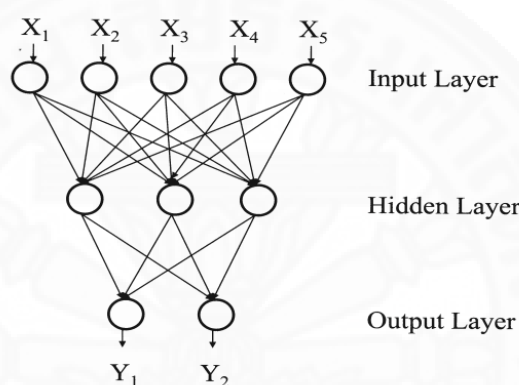


ภาพที่ 2.49 ตัวอย่างต้นไม้การตัดสินใจสำหรับการจำแนกประเภทของข้อมูล จาก การทำเหมืองข้อมูล, โดย สุรพงษ์ เอื้อวัฒนามงคล, 2559, สถาบันบัณฑิตพัฒนบริหารศาสตร์, สำนักพิมพ์สถาบันบัณฑิตพัฒนบริหารศาสตร์

2.5.3 ตัวแบบจำแนกประเภทข้อมูลแบบเครือข่ายประสาท (Neural Network)

เครือข่ายประสาทเป็นแบบจำลองที่จำลองการทำงานของเครือข่ายประสาทของมนุษย์ เพื่อให้สามารถตัดสินใจปัญหาบางอย่างได้เช่นเดียวกับมนุษย์ ถึงแม้แบบจำลองแบบนี้จะยังไม่สามารถทำงานได้อย่างสมบูรณ์ เนื่องจากความสลับซับซ้อนของเครือข่ายประสาทของมนุษย์ ถึงกระนั้นวิธีการนี้ก็สามารถตัดสินใจปัญหาได้อย่างดี โดยเฉพาะเรื่องการจำแนกประเภทของข้อมูลที่ให้ความถูกต้องแม่นยำในการจำแนกประเภทสูง นอกจากนี้ตัวจำแนกประเภทนี้ยังสามารถทำงานได้ดีกับข้อมูลที่มีสัญญาณรบกวนหรือนอยซ์ (Noise) ปะปนอยู่ เครือข่ายประสาทที่ใช้จำแนกข้อมูล จะเป็นเครือข่ายแบบ เครือข่ายแยกประเภทข้อมูลแบบคาดการณ์สถานการณ์ล่วงหน้า (Feed forward Network) ประกอบด้วยจำนวนชั้นของเซลล์ประสาทหลายชั้น เซลล์ประสาทชั้นที่หนึ่งจะทำการรับ

ค่าผลลัพธ์จากเซลล์ประสาทที่อยู่ในชั้นก่อนหน้า ทำการคำนวณด้วยฟังก์ชันแบบไม่เชิงเส้นและให้ผลลัพธ์ที่ส่งออกไปยังเซลล์ประสาททุกเซลล์ในชั้นถัดไปผ่านเส้นเชื่อมโยง ดังภาพที่ 2.50 เครือข่ายประสาทจะต้องมีชั้นอย่างน้อย 2 ชั้น ชั้นแรก คือ ชั้นรับข้อมูล (Input Layer) ซึ่งในชั้นนี้จะประกอบไปด้วย เซลล์ประสาท ซึ่งทำหน้าที่รับข้อมูลและส่งต่อไปยังเซลล์ประสาทในชั้นถัดไป ชั้นถัดไป คือ ชั้นผลลัพธ์ (Output Layer) ซึ่งประกอบไปด้วยเซลล์ประสาทที่รับผลลัพธ์จากเซลล์ประสาทชั้นก่อนหน้า จากนั้นจึงทำการคำนวณเมื่อได้ผลลัพธ์แล้วจึงส่งผลลัพธ์ออกไปเพื่อเป็นผลลัพธ์สุดท้ายของการคำนวณของเครือข่าย (สุรพงศ์ เอื้อวัฒนามงคล, 2559)



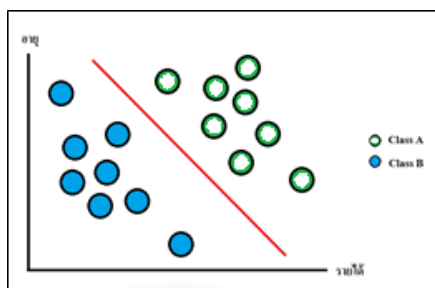
ภาพที่ 2.50 โครงสร้างเครือข่ายแบบการคำนวณสถานะการณ่วงหน้า (Feed forward Network) จาก การทำเหมืองข้อมูล, โดย สุรพงศ์ เอื้อวัฒนามงคล, 2559, สถาบันบัณฑิตพัฒนบริหารศาสตร์, สำนักพิมพ์สถาบันบัณฑิตพัฒน บริหารศาสตร์

จากภาพที่ 2.49 สำหรับชั้นของเครือข่ายที่อยู่ตรงกลางระหว่างชั้นรับข้อมูลและชั้นส่งผลลัพธ์ เรียกว่า ชั้นปกปิดหรือชั้นซ่อน (Hidden Layer) อาจจะมีหรือไม่มีก็ได้และอีกกรณีหนึ่งคืออาจมีกี่ชั้นก็ได้ จำนวนชั้นปกปิดของเครือข่ายประสาทและจำนวนเซลล์ที่อยู่ในชั้นปกปิดถือเป็นตัวแปรที่จะต้องปรับค่า เพื่อให้ได้ผลลัพธ์ในการจำแนกประเภทของข้อมูลให้ได้ถูกต้องมากที่สุด สำหรับจำนวนเซลล์ในชั้นรับข้อมูลและส่งผลลัพธ์จะขึ้นกับจำนวนข้อมูลนำเข้า จำนวนค่าของผลลัพธ์และวิธีการแทนค่าข้อมูลนำเข้าและค่าผลลัพธ์

2.5.4 ตัวจำแนกข้อมูลแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

ตัวจำแนกข้อมูลแบบซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้ในการจำแนกค่าคุณลักษณะของชุดข้อมูลสองกลุ่มที่ออกจากกันโดยการสร้างเส้นแบ่งที่เป็นเส้นตรงเพื่อให้ทราบว่า

เส้นตรงที่แบ่งข้อมูลออกจากกันเส้นตรงใดเป็นเส้นตรงที่ดีที่สุด ดังภาพที่ 2.51 (อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร, 2557), (Manning, Raghavan and Schütze, 2008)



ภาพที่ 2.51 แสดงโมเดล SVM ที่แบ่งข้อมูลออกเป็น 2 กลุ่ม

โดยเลือกแบบจำลองเส้นตรง (Linear Model) ที่มีระยะห่างระหว่าง 2 กลุ่มที่ห่างกันมากที่สุด ซึ่งสามารถนำมาเขียนเป็นสมการในการแก้ปัญหา โดยนำข้อมูลมาวางไว้ในพื้นที่คุณลักษณะกลุ่มข้อมูลในรูปแบบเวกเตอร์จะได้ ดังสมการที่ (5)

$$X = ((x_1 - y_1), \dots, (x_i - y_i)) \quad (5)$$

กำหนดสมการตัวแบ่งข้อมูลเส้นแบ่ง (Hyperplane) ด้วยการใช้เส้นขอบ (Margin) ของเส้นแบ่งเป็นเส้นที่สัมผัสกับค่าข้อมูลในพื้นที่คุณลักษณะที่ใกล้สุดเส้นขอบของเส้นทั้งสอง ซึ่งถูกแทนด้วย $w^T x + b \geq y \geq 1$ ถ้า $y = 1$ และ $w^T x + b \leq y \leq -1$ ถ้า $y = -1$ หากเส้นขอบใดมีค่ากว้างมากที่สุดจะแสดงให้เห็นข้อมูลที่แยกกันได้ดีชัดเจน ดังนั้นเส้นขอบที่มีความกว้างมากที่สุดจะเป็นเส้นแบ่งข้อมูลที่ดีที่สุด โดยเรียกเส้นแบ่งข้อมูลทั้งสองว่าเส้นขอบ (อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร, 2557)

2.5.5 ตัวแปรการจำแนกข้อมูลแบบเบย์ส (Bayes Classifier)

วิธีการนี้ เป็นตัวแบบที่คำนวณค่าความน่าจะเป็นที่ข้อมูลจะมีประเภทใดประเภทหนึ่ง เมื่อกำหนดค่าตัวแปรต่าง ๆ ของข้อมูล หรือ (สุรพงศ์ เอื้อวัฒนามงคล, 2559)

$$P(C_j | A_1, A_2, \dots, A_m)$$

โดยที่ C_j เป็นประเภทของข้อมูลที่ j

A_1, A_2, \dots, A_m เป็นค่าตัวแปรต่าง ๆ ของข้อมูลที่กำหนดให้ตามลำดับ

m เป็นจำนวนของตัวแปร

ถ้าสามารถคำนวณค่าความน่าจะเป็นข้างต้นสำหรับทุกประเภทของข้อมูล อาจอนุมานว่าข้อมูลที่กำหนดให้มีประเภทข้อมูลที่ให้ค่าความน่าจะเป็นดังกล่าวสูงสุด หรือ $\text{Max}P_j(C_j|A_1, A_2, \dots, A_m)$ การสร้างตัวแบบเบย์สจะคำนวณหาค่าความน่าจะเป็นดังกล่าว โดยตรงจากข้อมูลดังนี้

จากทฤษฎีของเบย์ส จะได้ว่า

$$P(C_j|A_1, A_2, \dots, A_m) = (P(A_1, A_2, \dots, A_m|C_j) * P(C_j)) / P(A_1, A_2, \dots, A_m)$$

เนื่องจากค่าความน่าจะเป็นของทุกประเภทของข้อมูลตามสมการข้างต้น มีตัวหารที่เป็นค่าตัวเดียวกัน การหา C_j ที่ให้ค่าความน่าจะเป็นตามสมการข้างต้น สามารถทำได้โดยการคำนวณเฉพาะค่าผลคูณที่เป็นตัวตั้งของสมการข้างต้น แล้วนำผลคูณของแต่ละประเภทมาเปรียบเทียบ เพื่อหาค่าสูงสุดสำหรับค่า $P(C_j)$ เป็น Prior Probability ซึ่งสามารถคำนวณได้โดยตรงจากข้อมูลเรียนรู้

การหาค่า $P(A_1, A_2, \dots, A_m|C_j)$ จากข้อมูลเรียนรู้อาจทำได้ยาก ถ้าข้อมูลเรียนรู้ไม่มีข้อมูลที่มีค่าตัวแปรต่าง ๆ ตรงตามที่กำหนด อย่างไรก็ตาม ถ้ากำหนดสมมติฐานที่ตัวแปรต่าง ๆ ของข้อมูลเป็นอิสระต่อกัน ค่าความน่าจะเป็นนี้สามารถคำนวณได้ด้วยสมการต่อไปนี้ ซึ่งความน่าจะเป็นแต่ละตัวทางขวาของสมการสามารถหาได้โดยตรงจากข้อมูลเรียนรู้

$$P(A_1, A_2, \dots, A_m|C_j) = P(A_1|C_j) * P(A_2|C_j) * \dots * P(A_m|C_j)$$

โดยที่

$$P(A_i|C_j) = P(A_i \cap C_j) / P(C_j)$$

ตารางที่ 2.3

ข้อมูลสำหรับการสร้างตัวแบบจำแนกประเภทข้อมูลแบบเบย์ส

A ₁	A ₂	A ₃	A ₄	Class
No	No	No	No	Positive
Yes	No	No	No	Positive
No	Yes	No	No	Positive
Yes	Yes	No	No	Positive
No	No	Yes	No	Positive

A_1	A_2	A_3	A_4	Class
Yes	No	Yes	No	Negative
No	Yes	Yes	No	Negative
Yes	Yes	Yes	No	Negative
No	No	No	Yes	Negative
Yes	No	No	Yes	Positive
No	Yes	No	Yes	Negative

หมายเหตุ. จาก การทำเหมืองข้อมูล, โดย สุรพงศ์ เอื้อวัฒนามงคล, 2559, สถาบันบัณฑิตพัฒนบริหารศาสตร์, สำนักพิมพ์สถาบันบัณฑิตพัฒนบริหารศาสตร์

กำหนดให้ข้อมูลเรียนรู้ 11 ตัวอย่างดังตารางที่ 2.3 ถ้าต้องการทราบประเภทของข้อมูล X ต่อไปนี้ $X = (A_1 = \text{Yes}, A_2 = \text{Yes}, A_3 = \text{Yes}, A_4 = \text{Yes})$ จากข้อมูลที่กำหนดให้สามารถคำนวณหาค่าความน่าจะเป็นของแต่ละประเภท ได้ดังต่อไปนี้

$$\begin{aligned}
 P(X|\text{Positive}) &= P(A_1 = \text{Yes}|\text{Positive}) * P(A_2 = \text{Yes}|\text{Positive}) * \\
 &\quad P(A_3 = \text{Yes}|\text{Positive}) * P(A_4 = \text{Yes}|\text{Positive}) \\
 &= 3/6 * 2/6 * 1/6 * 1/6 \\
 &= 0.005
 \end{aligned}$$

$$\begin{aligned}
 P(X|\text{Negative}) &= P(A_1 = \text{Yes}|\text{Negative}) * P(A_2 = \text{Yes}|\text{Negative}) * \\
 &\quad P(A_3 = \text{Yes}|\text{Negative}) * P(A_4 = \text{Yes}|\text{Negative}) \\
 &= 2/5 * 3/5 * 3/5 * 2/5 \\
 &= 0.058
 \end{aligned}$$

$$P(\text{Positive}) = 6/11 = 0.545$$

$$P(\text{Negative}) = 5/11 = 0.455$$

$$P(X|\text{Positive}) * P(\text{Positive}) = 0.005 * 0.545 = 0.003$$

$$P(X|\text{Negative}) * P(\text{Negative}) = 0.058 * 0.455 = 0.026$$

$$P(X|\text{Positive}) * P(\text{Positive}) < P(X|\text{Negative}) * P(\text{Negative})$$

ดังนั้น X น่าจะมีประเภทเป็น Negative

2.6 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง (Theory and Related Research)

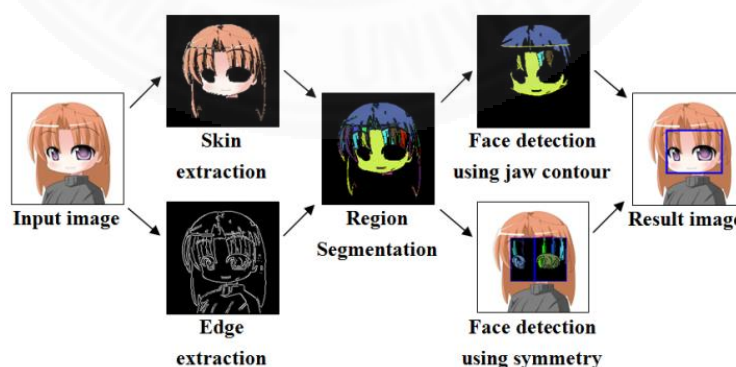
ในส่วนองงานวิจัยเรื่อง การจำแนกอารมณ์จากใบหน้าที่การตูนด้วยระบบอัตโนมัติ จำเป็นต้องศึกษางานวิจัยที่เกี่ยวข้องเพื่อนำมาเป็นแนวทางในการพัฒนาระบบ ในหัวข้อนี้จะประกอบได้ด้วยทฤษฎีและงานวิจัยเกี่ยวกับ การตรวจจับใบหน้าที่ (Face Detection) และการจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้า (Facial Emotion Classification) ซึ่งจะอธิบายตามหัวข้อดังต่อไปนี้

2.6.1 การตรวจจับใบหน้าที่ (Face Detection)

การตรวจจับใบหน้าที่เป็นเทคโนโลยีที่ถูกนำไปใช้งานอย่างแพร่หลายที่พัฒนาขึ้นมาเพื่อเอาไว้ใช้ระบุหน้าของมนุษย์จากภาพถ่ายดิจิทัล (Frischholz, 2016) การตรวจจับใบหน้าที่หมายรวมไปถึงกระบวนการทางจิตวิทยา โดยที่มนุษย์สามารถที่จะเข้าใจความหมายของแต่ละอารมณ์ที่แสดงออกมาได้ (Lewis and Ellis, 2003) การตรวจจับใบหน้าที่มีหลายเทคนิค ได้แก่ การตรวจจับใบหน้าที่ด้วยคุณลักษณะสีและขอบของวัตถุ การตรวจจับใบหน้าที่ด้วยฮาร์ไลค์ฟีเจอร์ (Haar-like Feature) และการตรวจจับใบหน้าที่ด้วยแบบจำลองอวัยวะหลักบนหน้า

การตรวจจับใบหน้าที่ด้วยคุณลักษณะสีและขอบของวัตถุ

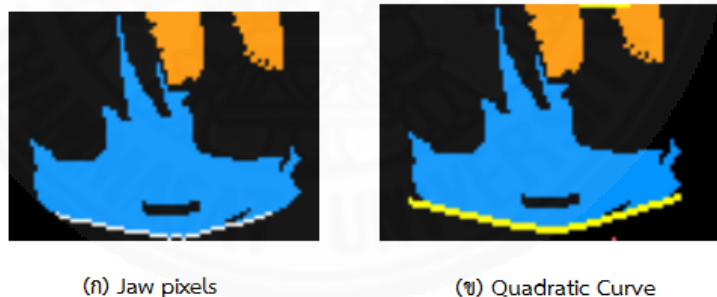
Takayama, Johan และ Nishita (2012) กล่าวว่า การตรวจจับใบหน้าที่ของตัวการ์ตูนนั้น สามารถหาส่วนที่เป็นบริเวณใบหน้าที่ได้ด้วยการหาบริเวณที่เป็นสีผิวและสกัดหาขอบ (Skin Color and Edge Extraction) ของภาพการ์ตูน ดังภาพที่ 2.52



ภาพที่ 2.52 ขั้นตอนในการตรวจจับส่วนที่เป็นใบหน้าที่ของตัวการ์ตูน จาก *FACE DETECTION AND FACE RECOGNITION OF CARTOON CHARACTERS USING FEATURE EXTRACTION*, โดย Takayama, K., Johan, H., & Nishita, T., 2012, Proceedings of the IIEEJ Image Electronics and Visual Computing Workshop Kuching, Malaysia, November 21-24

จากภาพที่ 2.52 ทำการรับภาพเข้ามาในระบบ จากนั้นหาส่วนที่เป็นสีผิวของการ์ตูน พร้อมทั้งทำการหาขอบ (Edge Extraction) จากนั้นทำการเซ็กเมนต์ (Segmentation) เพื่อนำมาเลือกเฉพาะส่วนของภาพที่มีข้อมูลที่สนใจ จากนั้นนำข้อมูลที่ได้มาทำการหาบริเวณที่เป็นใบหน้าด้วยการใช้เทคนิค Jaw Contour และ Symmetry ซึ่งจะอธิบายตามลำดับการทำงานในแต่ละเทคนิคตามลำดับ

Jaw Contour เป็นเทคนิคที่ช่วยหาส่วนที่เป็นบริเวณของกรามและคางบนใบหน้า ซึ่งมนุษย์สามารถที่จะแยกแยะได้ว่าส่วนใดคือบริเวณใบหน้าของตัวการ์ตูนในภาพ วิธีทางคอมพิวเตอร์จึงได้นำเทคนิคของ Jaw Contour และ Symmetry มาทำการหาเงื่อนไข เพื่อให้คอมพิวเตอร์สามารถที่จะแยกแยะส่วนที่เป็นบริเวณของใบหน้าของการ์ตูนได้ อันดับแรก คือ คำนิยามหาบริเวณที่เป็นใบหน้าของการ์ตูน โดยการกำหนดค่าพิกเซลเริ่มต้นเป็น $(X_i, Y_i) = (i = 1, \dots, n)$ โดยที่ X คือ ขอบเขตของข้อมูลที่ต้องการหาและ n คือ จำนวนพิกเซลทั้งหมด ดังภาพที่ 2.53 (ก) กรณีที่พบเห็นโดยส่วนใหญ่ เส้นโค้งของบริเวณกรามและคาง (Jaw Contour) ของการ์ตูนจะถูกเส้นโค้งควอดราติก (Quadratic Curve) ปิดทับอยู่ดังภาพที่ 2.53 (ข) ซึ่งเซลล์ของ Quadratic Curve คือ การยกกำลังสองของเส้นโค้ง (R^2)



ภาพที่ 2.53 ตัวอย่างการหาบริเวณที่เป็นใบหน้าของการ์ตูนด้วยเทคนิค Jaw Contour จาก *FACE DETECTION AND FACE RECOGNITION OF CARTOON CHARACTERS USING FEATURE EXTRACTION*, โดย Takayama, K., Johan, H., & Nishita, T., 2012, Proceedings of the IIEEJ Image Electronics and Visual Computing Workshop Kuching, Malaysia, November 21-24

จากภาพที่ 2.53 (ก) จะสังเกตเห็นได้ว่า ส่วนที่เป็นบริเวณใบหน้าจะแทนด้วยสีน้ำเงิน หลังจากนั้นจึงทำการแบ่งส่วนของขอบพิกเซลด้วยสีขาว ภาพที่ 2.53 (ข) คือ นำค่าตัวเลขจากเส้น

ขอปสี่ขวามาทำการยกกำลังสองและสร้างเส้นเหลืองขึ้นมา กำหนดให้ R^2 คือ สัมประสิทธิ์ของการตัดลีนใจ Y_a คือค่าเฉลี่ยของ Y_i และผลรวมของความคลาดเคลื่อนยกกำลังสอง คือ S จากนั้นจะทำการประมาณ เพื่อหาค่าที่ดีที่สุดโดยพิจารณาที่ค่า R^2 เข้าใกล้ 1 เมื่อ R^2 มีค่ามากกว่า 0.85 แล้ว จะพิจารณาได้ว่ามีความเป็นไปได้ที่จะเป็น Jaw Contour ในการหา Jaw Contour สามารถคำนวณหา ค่าได้จากสมการ ดังสมการที่ (6)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_a)^2}{S}. \quad (6)$$

หลังจากที่ได้ค่า Jaw Contour เรียบร้อยแล้ว จากนั้นก็นำค่าที่ได้มาหาบริเวณที่คาดว่าจะเป็นใบหน้าต่อด้วยเทคนิค Face Detection Using Symmetry ในส่วนของ Symmetry จะเป็นการพิจารณาหาส่วนที่เป็นสมมาตรของใบหน้า โดยการพิจารณาหาส่วนที่เป็นสมมาตรของใบหน้านั้น จะพิจารณาจากการหาจุดสมมาตร โดยการวัดสมมาตรจากบริเวณที่เป็นดวงตาทั้งสองข้าง อาจจะพิจารณาในส่วนของโรมที่บริเวณหน้าผากร่วมเข้ามาด้วย อันดับแรก คือ ทำการสร้างกรอบ Bounding Box ขึ้นมาเพื่อแสดงให้เห็นทราบถึงส่วนที่เป็นบริเวณใบหน้าของการ์ตูน โดยใช้ทฤษฎี k - Nearest Neighbors ในการหาบริเวณที่เป็นใบหน้าของการ์ตูนโดยพิจารณาจากสีผิวเป็นหลัก ในการทดลองของ Takayama, Johan และ Nishita (2012) นี้ กำหนดค่า $k = 3$ เมื่อหาบริเวณที่เป็นดวงตาและเส้นผมที่สมมาตรแล้ว จะทำการกำหนดสีหลักของ Regions ตรวจสอบค่า Regions จากทั้งสองกลุ่มว่ามีค่าของตัวจำแนกประเภทมีความคล้ายกับเพื่อนบ้านกลุ่มใด ถ้าอยู่ในกลุ่มใดมากที่สุด ระบบก็จะสามารถทำนายได้ว่าส่วนนี้อาจจะเป็นบริเวณของใบหน้าการ์ตูน

Chau (2015) กล่าวว่า ในการหาส่วนที่เป็นบริเวณใบหน้าของการ์ตูนภาพวาดนั้น งานวิจัยบางส่วนก็ได้ใช้เทคนิคของ Haar like Feature มาทำการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูน การใช้เทคนิค Haar like Feature นี้มีประสิทธิภาพในระดับหนึ่ง สามารถตรวจจับบริเวณที่เป็นใบหน้าการ์ตูนได้ บางครั้งอาจจะมีค่าแปรปรวนของการตรวจจับบ้าง เพราะขึ้นอยู่กับรูปร่างรูปทรงของการ์ตูนในภาพที่ทำให้ระบบอาจจะตรวจจับบริเวณที่เป็นส่วนของใบหน้าของการ์ตูนผิดพลาดบ้างในบางกรณี เพราะโปรแกรมนี้ ใช้ในการตรวจจับใบหน้าของมนุษย์เป็นหลัก ซึ่งไม่ได้นิยมนำมาใช้ในการวิเคราะห์หาบริเวณที่เป็นใบหน้าของการ์ตูน แต่ในทางตรงกันข้าม ผู้วิจัยสามารถนำเทคนิคนี้มาใช้ในการวิเคราะห์ตัวการ์ตูนที่มีใบหน้าคล้ายคลึงกับมนุษย์ได้ ดังภาพที่ 2.54



(ก) ภาพการ์ตูนที่ดัดแปลงอย่างมาก

(ข) ภาพการ์ตูนที่มีความคล้ายหรือเหมือนกันกับมนุษย์

ภาพที่ 2.54 (ก) ภาพการ์ตูนที่ดัดแปลงอย่างมาก (ข) ภาพการ์ตูนที่มีความคล้ายหรือเหมือนกันกับมนุษย์ จาก *FACE DETECTION AND RECOGNITION OF DRAWN CHARACTERS*, โดย Chau, H., 2015.

Chau (2015) ได้ปรับปรุงเทคนิคในการตรวจจับหาบริเวณที่เป็นใบหน้าของการ์ตูน ด้วยวิธีการนำภาพเดียวกันมาการสอนระบบ (Train) ด้วยเทคนิคของ Haar like feature โดยเลือกภาพที่เป็น Positive Training มาประมาณ 100 ภาพ เพื่อมาใช้ในการสอนระบบ ซึ่งในภาพต้องเห็นบริเวณที่เป็นใบหน้าชัดเจน และจากนั้นเลือกภาพที่เป็น Negative Training มาประมาณ 100 ภาพ เพื่อนำมาสอนระบบด้วยเช่นกัน ซึ่งภาพทั้งสองชุดนี้ต้องมีฉากประกอบรวมอยู่ด้วย เมื่อโปรแกรมทำการตรวจจับภาพใบหน้าของตัวการ์ตูนในครั้งแรกเสร็จ จะพบว่าเกิดข้อผิดพลาดในการตรวจจับบริเวณที่เป็นใบหน้าการ์ตูน แสดงให้เห็นดังภาพที่ 2.55



ภาพที่ 2.55 แสดงผลลัพธ์ที่ผิดพลาด จากการใช้ชุดการ Training ระบบที่มีขนาดเล็ก จาก *FACE DETECTION AND RECOGNITION OF DRAWN CHARACTERS*, โดย Chau, H., 2015.

หลังจากนั้นจึงเพิ่มชุดฝึกอบรมเพิ่มจากเดิม 100 ภาพ เป็น 650 ภาพ โดยภาพต้องเห็นบริเวณที่เป็นใบหน้าชัดเจน และชุด Negative Training เพิ่มปริมาณเป็นประมาณ 200 ภาพ ระบบสามารถตรวจจับบริเวณที่เป็นส่วนหน้าฉากได้มากที่สุดถึงแม้บางภาพจะเห็นใบหน้าไม่ชัดเจนก็ตาม ภาพผลลัพธ์ที่ได้จากการ Training ระบบด้วยชุดฝึกอบรมมากขึ้น ความถูกต้องในการตรวจจับบริเวณที่เป็นส่วนใบหน้าการ์ตูนของภาพ แสดงดังภาพที่ 2.56

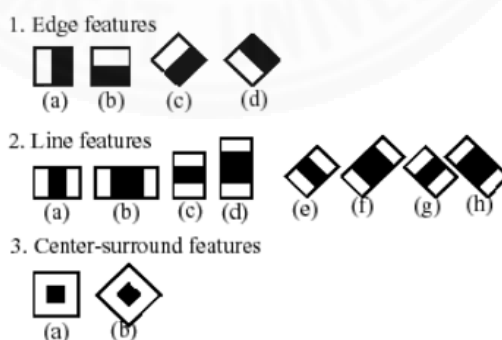


ภาพที่ 2.56 ผลลัพธ์จากการปรับปรุงเทคนิคในการใช้ชุดฝึกอบรมที่เพิ่มมากขึ้น จาก FACE DETECTION AND RECOGNITION OF DRAWN CHARACTERS, โดย Chau, H., 2015.

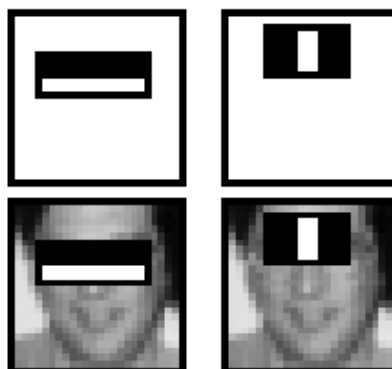
จากวิธีการของ Chau (2015) จะเห็นได้ว่า เพื่อให้ได้ผลลัพธ์ที่ดีในการตรวจจับหาใน ส่วนที่เป็นบริเวณใบหน้าของการ์ตูน ผู้ทำการวิจัยสามารถเลือกรูปภาพที่เหมาะสม มาใช้ในการฝึกสอนระบบและตรวจจับบริเวณที่เป็นส่วนของใบหน้าได้

ตรวจจับใบหน้าด้วยวิธีของ Haar Like Features

หลักในการทำ Haar like-Features ตามวิธีของ Viola และ Jones เป็นวิธีการ ตรวจจับและตีความจากวัตถุภายในภาพด้วยหลักการของ Haar Wavelet สำหรับการสร้างรูปที่ เหลี่ยม เพื่อนำมาใช้เป็นตัวกำหนดคุณลักษณะพิเศษของแต่ละฟีเจอร์ (Feature) รูปที่ 2.57 นี้จะ แสดงถึงผลต่างระหว่างพื้นที่ส่วนที่เป็นสีขาวและพื้นที่ส่วนที่เป็นสีดำ ซึ่งแต่ละฟีเจอร์สามารถ เปลี่ยนแปลงขนาดและตำแหน่งได้ ฟีเจอร์ที่ถูกกำหนดมานั้น ถูกนำไปใช้สำหรับการตรวจจับหา ลักษณะบนภาพแบบต่าง ๆ เช่น เพื่อใช้ในการค้นหาบริเวณที่เป็นส่วนของใบหน้ามนุษย์ (Viola and Jones, 2001) ดังรูป 2.58



ภาพที่ 2.57 รูปแบบของฟีเจอร์ 14 ชนิดที่ใช้สำหรับการตรวจจับลักษณะแบบต่าง ๆ บนใบหน้า จาก Rapid Object Detection using a Boosted Cascade of Simple Features, โดย Viola, P., & Jones, M. J., 2001, ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001



ภาพที่ 2.58 ตัวอย่างการใช้ Haar Like Feature ตรวจสอบลักษณะต่าง ๆ เพื่อหาส่วนที่เป็นใบหน้าของมนุษย์ จาก *Rapid Object Detection using a Boosted Cascade of Simple Features*, โดย Viola, P., & Jones, M. J., 2001, ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001

หลักการพื้นฐานของการตรวจจับใบหน้าของ Viola และ Jones คือ การนำภาพที่ต้องการตรวจหา มาแบ่งเป็นภาพย่อยๆ จากภาพย่อยดังกล่าวจะถูกพิจารณาเป็นภาพรับเข้า (Input Image) ของกระบวนการตรวจหาใบหน้า ซึ่งเทคนิคโดยทั่ว ๆ ไป ที่พบเห็นในการตรวจหาบริเวณที่เป็นใบหน้า ทำได้โดยปรับขนาดของภาพที่รับเข้าจาก ภาพที่มีความแตกต่างกันหลาย ๆ แบบและจากนั้นจึงใช้ตัวตรวจจับ (Detector) ที่มีขนาดคงที่มาใช้ในการค้นหาวัตถุ ซึ่งวิธีการตรวจจับแบบนี้มักมีข้อเสีย คือ ระยะเวลาในการคำนวณนั้นไม่คงที่ Viola และ Jones จึงเสนอเทคนิคในการตรวจจับภาพแบบใหม่ โดยใช้แบบจำลอง Haar like Feature มาเป็นตัวจับ โดยการปรับขนาดของตัวตรวจจับใบหน้าหลาย ๆ รอบ โดยแต่ละรอบใช้ขนาดของตัวจับที่แตกต่างกันออกไป ซึ่งเมื่อเทียบกับวิธีการเดิมแล้ว แสดงให้เห็นว่าเวลาที่ใช้ในการคำนวณนั้นไม่ได้แตกต่างกันมาก แต่เวลาในการตรวจจับบริเวณที่เป็นใบหน้าในแต่ละครั้งมีความคงที่ เทคนิคที่ Viola และ Jones นำเสนอ คือ การคำนวณการจำลองรูปแบบ Haar like Feature ด้วย Integral Image (Viola & Jones, 2001)

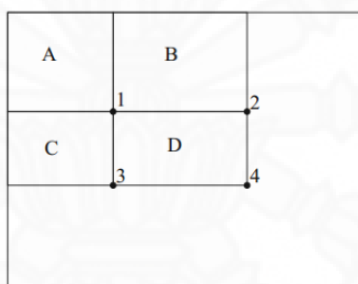
ขั้นตอนในการคำนวณหา Integral Image พิจารณาจากแบบจำลองของ Haar like Feature ที่ใช้ในการตรวจจับบริเวณที่เป็นใบหน้าของ Viola และ Jones (2001) ดังภาพที่ 2.47 โดยในแต่ละฟีเจอร์จะประกอบไปด้วยสี่เหลี่ยม 2 ประเภท คือ ส่วนที่แรเงาและส่วนที่ไม่ได้แรเงา ในการพิจารณาค่าการจำลองตามรูปแบบของ Haar like Feature คือ การหาผลต่างระหว่างความเข้มในส่วนที่แรเงากับส่วนที่ไม่ได้แรเงา จากนั้นนำผลลัพธ์ที่ได้ไปเปรียบเทียบกับค่า Thresholds กับ Polarity ที่ใช้ในการตัดสินใจว่าภาพที่รับเข้ามานั้น ควรจัดให้อยู่ในประเภทบวก (ภาพใบหน้า) หรือ

ลบ (ภาพที่ไม่ใช่ใบหน้า) เมื่อได้ผลลัพธ์เรียบร้อยแล้ว ผลลัพธ์นั้นจะถูกนำไปพิจารณาในลำดับถัดไป (รูสตี สุทธีวีร์กุล และ วิไลพร แซ่ลี, 2554)

การคำนวณแบบ Integral Image นั้นจะใช้สมการที่ใช้ในการหาดังสมการที่ (7) ซึ่ง Integral image คือ การหาผลรวมความเข้มของค่าในทุกพิกเซลของพื้นที่สี่เหลี่ยมใด ๆ ที่ตำแหน่ง $i(x, y)$ ใด ๆ ซึ่งมีเวลาการทำงานเป็น $O(1)$ ทำให้การคำนวณแต่ละพิกเซลนั้นทำได้เร็ว (Chen, Georganas & Petriu, 2007)

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (7)$$

Integral Image จะเป็นพีเจอร์ที่มีลักษณะเป็นสี่เหลี่ยมผืนผ้า ดังภาพที่ 2.59 สามารถใช้เวลาในการคำนวณหาค่าได้อย่างรวดเร็ว โดยการใช้ตัวแทนของพีเจอร์ นั้น ๆ ในการหาส่วนที่สนใจ เพื่อนำไปใช้ในการหาส่วนที่เป็นบริเวณของใบหน้าที่ต้องการ (Viola and Jones, 2001)



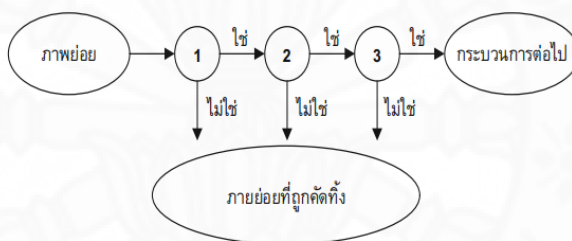
ภาพที่ 2.59 รูปตัวอย่างในการคำนวณหา Integral Image จาก *Rapid Object Detection using a Boosted Cascade of Simple Features*, โดย Viola, P., & Jones, M. J., 2001, ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001

ผลรวมของพิกเซลภายในกรอบสี่เหลี่ยม ดังภาพที่ 2.59 สามารถคำนวณหาด้วยอาเรย์ (Array) ทั้งสี่ตำแหน่งที่ใช้ในการอ้างอิงค่าใดค่าหนึ่งของ Integral Image ที่ตำแหน่งที่ 1 คือ ผลรวมของรูปสี่เหลี่ยมผืนผ้าของพิกเซล A ตำแหน่งที่ 2 คือ ค่า A+B ตำแหน่งที่ 3 คือ A+C และตำแหน่งที่ 4 คือ A+B+C+D หากต้องการคำนวณหาผลรวมของสี่เหลี่ยม D สามารถคำนวณได้จากนำตำแหน่งที่ 4 + 1 - (2 + 3) ตำแหน่ง $ii(x,y)$ เป็น Integral Image และ $i(x,y)$ เป็นตำแหน่งของภาพต้นฉบับ ในการคำนวณหา Integral Image จะใช้สมการคู่นี้ ดังสมการที่ (8) และ (9) (Viola and Jones, 2001)

$$s(x,y) = s(x,y-1) + i(x,y) \quad (8)$$

$$ii(x,y) = ii(x-1,y) + s(x,y) \quad (9)$$

จากสมการค่า $s(x,y)$ คือ ผลรวมเฉลี่ยสะสมของตำแหน่งที่ $s(x, -1) = 0$ และ ตำแหน่งที่ $ii(-1, y) = 0$ ภาพต้นฉบับที่ใช้ในการคำนวณหา Integral Image ใช้ภาพเพียงหนึ่งภาพเท่านั้น หลักการหาผลรวมจากรูปสี่เหลี่ยมผืนผ้าใด ๆ สามารถคำนวณได้จากตำแหน่งทั้งสี่ ที่กำหนดไว้ (Viola and Jones, 2001) ดังภาพที่ 2.58 Viola และ Jones ได้นำเสนอเทคนิคที่ใช้ในการเพิ่มประสิทธิภาพในการตรวจจับให้มีความถูกต้อง และใช้เวลาในการคำนวณที่ลดลง โดยแนวความคิดของขั้นตอนนี้คือ การสร้างการรวมตัวจำแนกกลุ่มแบบต่อเรียง (Cascaded Classifier) โดยเทคนิคนี้จะนำตัวจำแนก (Classifier) หลาย ๆ ตัวมาเรียงต่อกันตามลำดับ ดังภาพที่ 2.60 (รูสตี สุทธีวีร์กุล และ วิไลพร แซ่ลี, 2554)



ภาพที่ 2.60 แสดงขั้นตอนการ รวมตัวจำแนกกลุ่มแบบเรียงต่อกัน จาก การตรวจจับใบหน้าด้วยวิธีการพื้นฐานของการจำลองรูปแบบ Haar-like, โดย รูสตี สุทธีวีร์กุล และ วิไลพร แซ่ลี, 2554, วารสารวิศวกรรมศาสตรมหาวิทาลัยศรีนครินทรวิโรฒปีที่ 6 ฉบับที่ 2 เดือนกรกฎาคม - ธันวาคม

โดยที่ตัวจำแนกในลำดับแรกมักมีความซับซ้อนที่น้อยกว่า อาจสร้างมาจากตัวจำแนกแบบอ่อนแอกว่าไม่ก็ตัวเพื่อลดปริมาณการคำนวณและลดระดับของค่า Threshold ลงเพื่อลดความผิดพลาดในการตรวจจับใบหน้าผิดตำแหน่ง ซึ่งภาพถ่ายที่ถูกจัดประเภทว่าไม่ใช่ใบหน้า (Non - Face) จะถูกส่งต่อไปยังตัวจำแนกตัวถัดไป ซึ่งมีความซับซ้อนสูงขึ้นตามลำดับ และกล่าวได้ว่า ยังมีจำนวนชั้นของตัวจำแนกมากเท่าใด โอกาสที่ภาพถ่ายจะเป็นใบหน้าจะยิ่งมีมากขึ้น (รูสตี สุทธีวีร์กุล และ วิไลพร แซ่ลี, 2554)

การตรวจจับใบหน้าด้วยแบบจำลองอวัยวะหลักบนหน้า

Arumugam (2012) อธิบายว่าการที่จะสอนให้ระบบรู้จำอารมณ์บนใบหน้าของมนุษย์ได้นั้น ต้องตัดปัญหาหลัก สามประการ (ได้แก่ เพศ อายุ และ เชื้อชาติ) ออกไปจากการนำข้อมูลภาพ

มาทำการวิเคราะห์ระบบ กล่าวคือ การตรวจจับบริเวณที่เป็นใบหน้าที่ดี จะต้องสามารถวิเคราะห์เพื่อหาคุณลักษณะเฉพาะบนใบหน้าและนำคุณลักษณะเด่นต่าง ๆ ในสภาพแวดล้อมของบุคคลที่มีความแตกต่างกันออกไปทางสรีระ ระบบควรจะมีความยืดหยุ่นกับสภาพแวดล้อมที่แตกต่างกันออกไป เช่น การเปลี่ยนแปลงของทรงผม ขนใบหน้า หนวด เครา ฯลฯ และระบบก็ต้องสามารถที่จะเติมชิ้นส่วนที่ขาดหายของใบหน้าและสร้างส่วนที่หายไปขึ้นมาให้ครบทั้งใบหน้าได้ นอกจากนี้ยังควรดำเนินการวิเคราะห์การแสดงออกของใบหน้าที่สามารถที่จะมีการเคลื่อนไหวไปในทิศทางใด ระบบก็สามารถจะอ้างอิงและประมวลผลภาพได้เหมือนกับการมองเห็นของมนุษย์ ขั้นตอนแรก คือ การเตรียมข้อมูลภาพ (Image Detection) ซึ่งภาพที่จะใช้งานหลัก เป็นภาพมาจากการใช้ Real Time Database ภาพที่เก็บเป็นภาพ .jpg ภาพโดยส่วนใหญ่ต้องปราศจาก แวนตา หนวดเครา และการเคลื่อนไหวใด ๆ ศีรษะต้องตั้งตรง จากนั้นนำภาพที่ได้ไปเก็บในฐานข้อมูล แยกเก็บภาพตามประเภทของอารมณ์ดังภาพที่ 2.61, 2.62 และ 2.63 ตามลำดับ



ภาพที่ 2.61 ภาพที่เก็บในฐานข้อมูลของอารมณ์โกรธที่แสดงออกทางสีหน้าของมนุษย์ จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7



ภาพที่ 2.62 ภาพที่เก็บในฐานข้อมูลของอารมณ์รังเกียจที่แสดงออกทางสีหน้าของมนุษย์ จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7



ภาพที่ 2.63 ภาพที่เก็บในฐานะข้อมูลของอารมณ์มีความสุขที่แสดงออกทางสีหน้าของมนุษย์ จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7

สร้างแบบจำลองในการคำนวณหาส่วนที่เป็นบริเวณใบหน้า ซึ่งประสิทธิภาพในการตรวจหาบริเวณที่เป็นใบหน้าส่วนใหญ่ขึ้นอยู่กับข้อมูลความถูกต้องภายในบริเวณหน้า เช่น คิ้ว ตา ปาก เป็นต้น หลังจากตรวจพบบริเวณที่เป็นใบหน้าแล้ว จะมีวิธีการแยกแยะเพื่อหาคุณสมบัติหลัก ออกเป็นสองวิธี คือ วิธีการหาแบบองค์รวม และการเชิงวิเคราะห์ วิธีการหาแบบองค์รวม จะใช้ภาพใบหน้าดิบเข้ามาเพื่อทำการวิเคราะห์หาบริเวณที่เป็นใบหน้า และจะทำการสร้าง Block ที่มีขนาด 50 เพื่อเก็บข้อมูลภาพที่รับเข้ามาในระบบ หลังจากที่ได้รับภาพเข้ามาในระบบแล้วข้อมูลภาพเหล่านั้นจะถูกป้อนไปเก็บไว้ในฐานข้อมูล

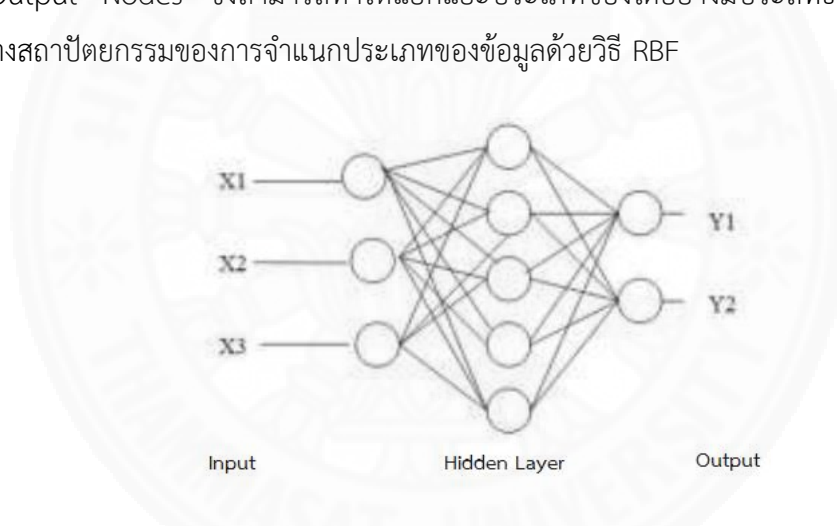
2.6.2 การจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้า (Facial Emotion Classification)

การแสดงอารมณ์ออกทางสีหน้าของมนุษย์นั้นมีความสำคัญมากในความสัมพันธ์ระหว่างบุคคล เพราะวิธีการแสดงออกแบบนี้แสดงให้เห็นถึงการถ่ายทอดของข้อมูลจำนวนมาก จากการมองเห็นได้ด้วยสายตา มากกว่าคำพูด ถึงแม้ว่ามนุษย์สามารถที่จะรู้จำการแสดงออกทางสีหน้าได้อย่างรวดเร็ว ผู้วิจัยจึงมีความเชื่อว่า เครื่องคอมพิวเตอร์ ก็มีความสามารถในการรู้จำได้เหมือนกับมนุษย์ ซึ่งแนวความคิดนี้เป็นสิ่งที่มีความท้าทายและน่าสนใจมาก (Arumugam, 2011)

การจำแนกอารมณ์บนใบหน้าด้วยคุณลักษณะทางสถิติ

Arumugam (2011) อธิบายว่าวิธีการคำนวณหาคุณลักษณะเฉพาะโดยใช้วิธีการคณิตศาสตร์ (Statistical Method) หลังจากอ่านภาพเข้ามาในระบบ จำเป็นต้องทำสำเนาภาพเพื่อนำมาใช้ในการวิเคราะห์ โดยการหาความสัมพันธ์ของเมตริกซ์ เทียบกับข้อมูลภาพที่จัดเก็บไว้ในตาราง

เมตริกซ์ โดยใช้วิธีการในการคำนวณหาทั้งหมดสามวิธี ดังนี้ วิธีที่หนึ่ง คือ การคำนวณหาค่า Eigen vector และ Eigen Value วิธีที่สอง คือ การคำนวณหาแบบเชิงเส้นด้วยเทคนิคของ Fisher's Linear Discriminant วิธีที่สามจำแนกหาพีเจอร์ด้วยลักษณะเฉพาะแบบ Singular Value Decomposition (SVD) จากนั้นนำคุณลักษณะที่ได้มาทำการจำแนกประเภทข้อมูลด้วยเครือข่ายประสาท (Neural networks) ซึ่งการใช้โครงข่ายประสาทเป็นประโยชน์อย่างยิ่งกับการวิเคราะห์ข้อมูลที่มีขนาดใหญ่ โดยเฉพาะข้อมูลที่เกี่ยวข้องกับการรู้จำใบหน้า ซึ่งในการจำแนกประเภทข้อมูลด้วยเครือข่ายประสาทนั้น Arumugam เลือกรูปแบบการจำแนกแบบ RBF (Radial Basis Function Network) เทคนิคของ RBF คือ ในแต่ละเซลล์จะมีรัศมีที่สมมาตรของฟังก์ชันพื้นฐานและฟังก์ชันที่มีการเปลี่ยนแปลงวัตถุประสงค์ของ เซลล์ประสาทแบบซ่อน (Hidden Neurons) คือ การจัดกลุ่มข้อมูลที่รับเข้าและทำการลดขนาดของข้อมูลลง จึงทำให้พบว่าสมมูลของค่าที่ดีที่สุดนั้นจะอยู่ระหว่าง Hidden neurons และ Output Nodes ซึ่งสามารถทำให้แยกแยะประเภทของได้อย่างมีประสิทธิภาพ ภาพที่ 2.64 ตัวอย่างสถาปัตยกรรมของการจำแนกประเภทของข้อมูลด้วยวิธี RBF



ภาพที่ 2.64 แสดงให้เห็นถึงสถาปัตยกรรมพื้นฐานสำหรับเครือข่ายแบบ RBF จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, *International Journal of Advanced Computer Science and Applications*, Vol. 2, No. 7

ขั้นตอนถัดไป คือ ขั้นตอนในการฝึกสอนระบบ (Training Phase) ขั้นตอนนี้จะทำการเปลี่ยนค่านำหนักของข้อมูลในแต่ละการเชื่อมต่อ เพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพ โดยป้อนข้อมูลภาพไว้สำหรับใช้สอนระบบลงไปเก็บไว้ในฐานข้อมูล ซึ่งในแต่ละภาพนั้นต้องเป็นบริเวณที่เป็นใบหน้าของมนุษย์ จะเรียกข้อมูลภาพชุดนี้ว่า ภาพใบหน้า จากนั้นจะทำการหาสกัดหาคุณลักษณะเด่นของภาพที่สำคัญและจัดเก็บค่านั้นไว้ในฐานข้อมูล ค่าเหล่านี้จะเรียกว่า พีเจอร์ โดย Arumugam

กำหนดให้พีเจอรที่ใช้ในระบบสองพีเจอร คือ ϕ_1 และ ϕ_2 ตามลำดับ ซึ่ง เวกเตอร์ทั้งสองอันนี้จะเป็นตัวแทนของคุณลักษณะที่สำคัญบนใบหน้าของมนุษย์ ซึ่งค่าที่สูงที่สุดของ Eigen values จะถูกเลือกนำมาใช้ สองข้อได้เปรียบสำหรับวิธีการนี้ คือ อันดับแรก ข้อมูลภาพทั้งหมดจะถูกลดขนาดและเลือกเอาเฉพาะพีเจอรที่สำคัญมาใช้งาน และอันดับที่สอง คือ การเลือกภาพที่เห็นโครงสร้างของข้อมูลในพิกเซลมากที่สุดมาเป็นพีเจอร การฝึกอบรมทำให้สามารถระบบสามารถวิเคราะห์และแยกแยะอารมณ์ความรู้สึกบนใบหน้าของมนุษย์ได้ จากการฝึกอบรมระบบ ทำให้ได้ข้อสรุปว่า ค่าสำหรับอารมณ์โกรธมีค่าเกือบ 0.1 ค่าสำหรับอารมณ์เบื่อหน่ายมีค่าเกือบ 0.2 และค่าสำหรับอารมณ์มีความสุขมีค่าเกือบ 0.3 รูปภาพที่ 2.65 จะแสดงให้เห็นผลลัพธ์โดยประมาณสำหรับสามอารมณ์พื้นฐาน

Inode for Angry Emotion		Angry	
-0.0012	0.0043	0.1000	
-0.0045	-0.0058	0.1000	=0.1
0.0026	-0.0024	0.1000	
Inode for Disgust Emotion		Disgust	
-0.0012	0.0043	0.1000	
-0.0045	-0.0058	0.1000	
0.0026	-0.0024	0.1000	= 0.2
-0.0070	0.0008	0.2000	
-0.0071	0.0008	0.2000	
Inode for Happy Emotion		Happy	
-0.012	0.0043	0.1000	
-0.0045	-0.0058	0.1000	=0.3
0.0026	-0.0024	0.2000	
-0.0070	-0.0008	0.3000	

ภาพที่ 2.65 แสดงผลลัพธ์โดยประมาณของทั้ง 3 อารมณ์ จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7

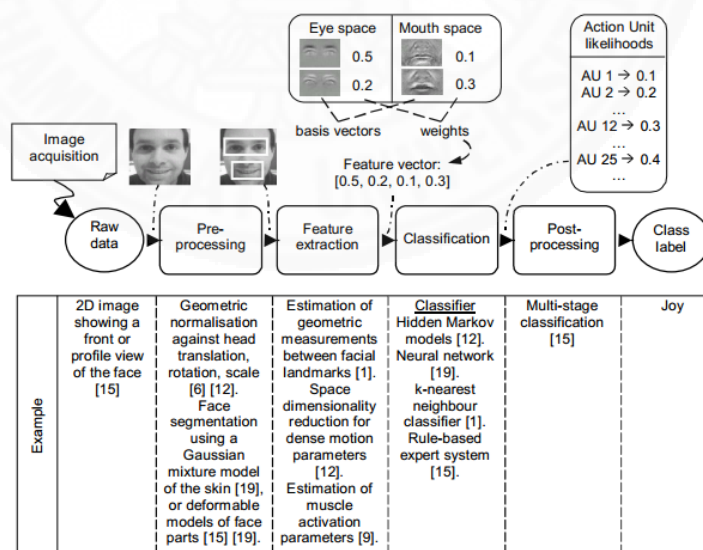
ขั้นตอนสุดท้าย คือ ขั้นตอนการทดสอบระบบ (Testing phase) หลังจากทำการสอนระบบไปแล้ว ในลำดับต่อไปจะทำการทดสอบความถูกต้องในการการแยกแยะอารมณ์บนใบหน้าของมนุษย์ ในขั้นตอนนี้ จะนำภาพดิบเข้ามาในระบบ โดยภาพที่นำเข้ามา นั้น ไม่ได้บ่งบอกถึงอารมณ์ความรู้สึก และภาพที่นำมาใช้ มาจากการนำเข้าภาพถ่ายแบบ Real Time จากนั้นนำภาพที่ได้มาทำการหาบริเวณที่เป็นใบหน้า และสกัดหาพีเจอรหลัก ๆ เพื่อนำไปใช้งาน เมื่อได้ผลลัพธ์ที่ต้องการแล้วก็นำไปตรวจสอบกับค่าที่จัดเก็บไว้ในฐานข้อมูล แล้วก็แสดงผลออกมาว่าภาพใบหน้ามนุษย์ที่รับเข้ามาในระบบนั้นมีอารมณ์ความรู้สึกใด ดังภาพที่ 2.66



ภาพที่ 2.66 แสดงผลการทำนายอารมณ์ของภาพที่รับเข้ามา จาก *Emotion Classification Using Facial Expression*, โดย Arumugam, D., & Purushothaman, B. E., 2011, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7

การตรวจจับอารมณ์บนใบหน้าด้วยวิธีของ Action Units

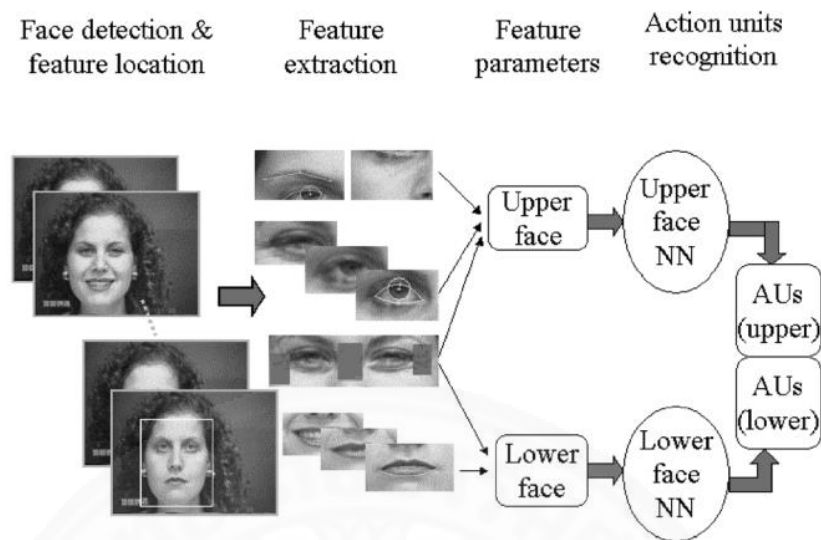
ในงานวิจัยของ Chibelushi และ Bourel (2002) จะทำการวิเคราะห์ในขั้นสูง ซึ่งเกี่ยวกับการทำดัชนีของภาพที่เป็นแบบนามธรรม มาใช้สำหรับการแยกแยะทางอารมณ์ของการแสดงออกทางสีหน้าของมนุษย์ด้วยระบบงานอัตโนมัติ องค์ประกอบหลักของการออกแบบระบบประกอบไปด้วย สถาปัตยกรรมทั่วไป (Generic Architecture) จะใช้ลำดับขั้นตอนในการประมวลผลภาพ ดังภาพที่ 2.67



ภาพที่ 2.67 เทคนิคสำหรับการรู้จำในการแสดงออกทางสีหน้าของมนุษย์ตามแบบการรู้จำขั้นพื้นฐาน จาก *Facial Expression Recognition: A Brief Tutorial Overview*, โดย Chibelushi, C., & Bourel, F., 2002.

จากภาพที่ 2.67 สามารถอธิบายการทำงานหลักของระบบได้ดังนี้ รับภาพเข้ามา (Image Acquisition) ในระบบ โดยภาพที่รับเข้ามานั้นจะเป็นภาพขนาด 2D และเป็นภาพขาวดำ และเห็นรายละเอียดบนใบหน้าครบถ้วน แต่อย่างไรก็ตาม ภาพสีก็นิยมใช้กันแพร่หลายเพราะ ภาพสีสามารถนำมาเป็นตัวจับอารมณ์ได้จากการแสดงออกทางสีหน้า เช่น อารมณ์โกรธ ใบหน้าจะแดง ลำดับถัดไป คือ การเตรียมข้อมูลก่อนการประมวลผล (Pre-processing) โดยทำการแบ่งตัดมาเฉพาะในส่วนที่สนใจ ที่จะนำมาใช้ในระบบ ตัวอย่างดัง ภาพที่ 2.67 ได้ทำการแยกส่วนที่สนใจคือ ตา และ ปากมาใช้ในลำดับถัดไป จากนั้นก็ทำการหาคุณลักษณะเฉพาะของภาพ (Feature Extraction) แสดงออกมา เป็นค่าของพีเจอร์นั้น ๆ และจัดเก็บข้อมูลไว้ในระบบ โดยทำการจัดเก็บในหมวดหมู่ของแต่ละพีเจอร์ เมื่อได้ข้อมูลดังกล่าวแล้วก็มาทำการวิเคราะห์และแยกแยะ (Classification) อารมณ์ด้วยเทคนิค Action Units (AU) ในลำดับถัดไป การจำแนกอารมณ์พื้นฐานอัตโนมัติ นั้น Chibelushi และ Bourel (2002) ได้กำหนดให้มีคลาส (Class) ในการจำแนกอารมณ์ออกเป็นหกแบบ ได้แก่ อารมณ์มีความสุข อารมณ์เศร้า อารมณ์ประหลาดใจ อารมณ์โกรธ อารมณ์กลัวและอารมณ์รังเกียจ

การรู้จำอารมณ์บนใบหน้าของมนุษย์นั้นถือว่าเป็นผลงานที่น่าสนใจเป็นอย่างมาก เทคนิคการรู้จำบนใบหน้าของ Action Units จึงเป็นอีกวิธีหนึ่งที่น่าสนใจ ซึ่งหลักในการทำงานของ Action Units นั้นส่วนใหญ่จะดูที่ความถี่ของชั้นข้อมูล ถ้ามีความถี่ของชั้นข้อมูลมาก ค่าความถูกต้องในการทำนายผลลัพธ์ก็จะมีประสิทธิภาพมากขึ้น ในการทำงานของ Action Units จะใช้เทคนิคของ Bayesian Network ในการหาผลลัพธ์ สามารถอธิบายหลักในการทำงานของ Action Units ได้ดังภาพที่ 2.68 (Tian, Kanade and Cohn, 2001) เริ่มแรกรับภาพเข้ามาในระบบและทำการตรวจจับหาบริเวณที่เป็นส่วนของใบหน้า จากนั้นหา Feature Extraction โดยทำการแยกพีเจอร์ (Feature) หลักออกสองส่วน คือ 1) ส่วนบนของใบหน้า ดังภาพที่ 2.69 (ก) ซึ่งประกอบไปด้วยดวงตาและคิ้ว 2) ส่วนล่างของใบหน้า ดังภาพที่ 2.69 (ข) ซึ่งประกอบไปด้วยจมูกและปาก (KalaiSelvi, Kavitha and Shunmuganathan, 2014)



ภาพที่ 2.68 แสดงวิธีการหา Feature หลักของการแยกแยะอารมณ์บนใบหน้ามนุษย์ด้วยเทคนิค AU จาก *Recognizing Action Units for Facial Expression Analysis*, โดย Tian, Y. L., Kanade, T., & Cohn, F. J., 2001, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL.23, NO.2, FEBRUARY.



(ก) ส่วนบนของใบหน้า

AU2



(ข) ส่วนล่างของใบหน้า

AU17

ภาพที่ 2.69 แสดงฟีเจอร์หลักที่สนใจส่วนบนและล่างของใบหน้า จาก *Capturing Facial Actions in Video to Revive Expressions of Humans*, โดย KalaiSelvi, R., Kavitha, P., & Shunmuganathan, K. L., 2014, International Conference on Innovations in Engineering and Technology.

เมื่อได้สองส่วนที่ต้องการจึงนำสองส่วนทั้งบนและล่างของใบหน้ามาทำการควมรวมกัน และบันทึกข้อมูลลงตาราง ดังตารางที่ 2.4 เพื่อสอนให้ระบบสามารถรู้จำอารมณ์พื้นฐานด้วยเทคนิคของเบย์ส (Bayes Classifier) ต่อไป

ตารางที่ 2.4

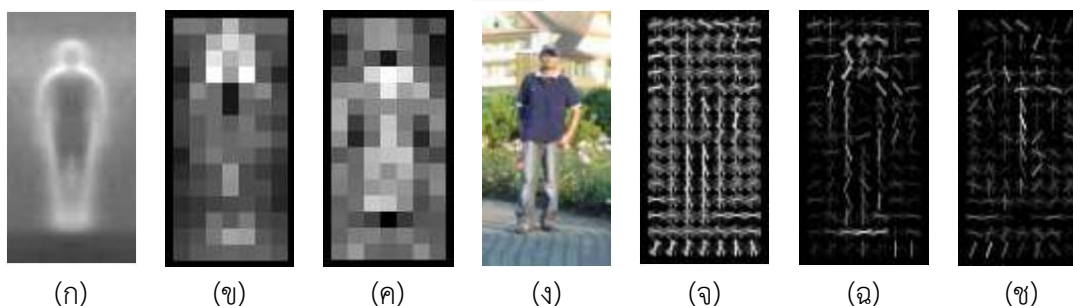
แสดงการ Combination ของทั้งหกอารมณ์

อารมณ์ (Emotion)	ส่วนที่สอดคล้องกันของ AU (Corresponding Action Units)
เศร้า (Sad)	AU2,AU17
มีความสุข (Happy)	AU6,AU12
ประหลาดใจ (Surprise)	AU1,AU5,AU27,AU25
โกรธ (Anger)	AU9,AU4,AU23
กลัว (Fear)	AU1,AU7,AU4
ปกติ (Normal)	AU25,AU7

หมายเหตุ. จาก *Capturing Facial Actions in Video to Revive Expressions of Humans*, โดย KalaiSelvi, R., Kavitha, P., & Shunmuganathan, K. L., 2014, *International Conference on Innovations in Engineering and Technology*

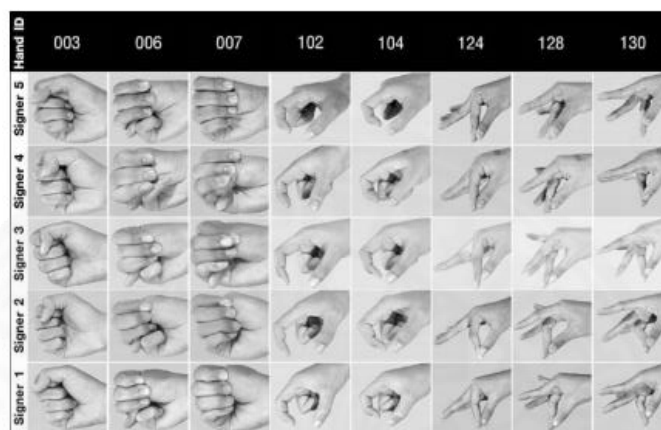
การหาคุณลักษณะเฉพาะที่ได้จากการกระจายตัวของเส้นขอบ (HOG)

Dalal และ Triggs (2005) กล่าวว่า การตรวจจับด้วย HOG จะเน้นไปที่เงาของภาพเป็นส่วนใหญ่ (โดยเฉพาะบริเวณหัว หัวไหล่ และเท้า) บล็อก (Blocks) ที่ใช้งานส่วนใหญ่จะอยู่ตรงกลางของพื้นหลังของภาพ ซึ่งสามารถอธิบายขั้นตอนในการหาคุณลักษณะเฉพาะของภาพได้ดังภาพที่ 2.70 ตามลำดับดังนี้ (ก) ภาพการไล่ระดับสีเฉื่อยของชุดฝึกสอน (ข) แต่ละพิกเซลแสดงค่าน้ำหนักที่เป็นบวกมากที่สุดของ SVM ที่จุดศูนย์กลางของพิกเซลในแต่บล็อก (ค) แสดงภาพที่เป็นค่าน้ำหนักเชิงลบของ SVM (ง) รูปภาพทดสอบ (จ) อธิบายถึงการคำนวณ R-HOG (ฉ,ช) โดยอธิบายตามค่าน้ำหนักที่เป็นบวกและลบของ SVM ตามลำดับ



ภาพที่ 2.70 ภาพแสดงการตรวจจับมนุษย์ด้วย HOG จาก *Histograms of Oriented Gradients for Human Detection*, โดย Dalal, N., & Triggs, B., 2005, NRIA Rhone-Alps, 655 avenue de l'Europe, Montbonnot 38334, France

อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร (2557) ใช้ HOG ร่วมกับ PCA (Principal Component Analysis) ในการรู้จำท่ามือภาษาไทย ผลที่ใช้ HOG ในการสกัดคุณลักษณะเฉพาะของภาพท่ามือ โดยใช้ข้อมูลในการเก็บค่าทิศทางเกรเดียนท์ $0^\circ - 180^\circ$ ค่าทิศทางจำนวน 6 bins และกำหนดค่ากริด (Grid) เป็น 8×8 ดังภาพที่ 2.71

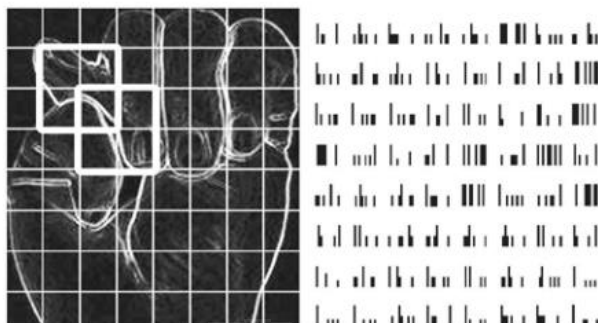


ดังภาพที่ 2.71 ภาพท่ามือที่ใช้ในการทดลอง จาก การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับการวิเคราะห์องค์ประกอบหลักและซัพพอร์ตเวกเตอร์แมชชีน โดย อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร, 2557, การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 10

เพื่อให้ได้ความแปรปรวนที่เหมาะสมผู้ทดลองได้ กำหนดให้ค่า v และ v' เป็นค่าปกติและไม่ปกติตามลำดับ โดยค่าสัมบูรณ์ v เป็นค่าแสดง L2 - norm ของเวกเตอร์ และ e เป็นค่าคงที่ป้องกันข้อผิดพลาดที่เกิดจากการหารด้วยศูนย์ดังสมการที่ (10)

$$v' = v / (||v|| + \epsilon) \quad (10)$$

และนอร์มอลไลซ์ด้วยการคำนวณตัวชี้วัดค่าความเข้มจากโอเวอร์แลปของเซลล์ภายในบล็อก ทำการยุบเนียนเซลล์ขนาด 2×2 เพื่อใช้ในการลดผลกระทบจากการเปลี่ยนแปลงของแสงและเงาให้น้อยลง ค่าฮิสโตแกรมที่ถูกเก็บทิศทางค่าเกรเดียนท์ไว้จะใช้เป็นตัวแทนแสดงคุณลักษณะในแต่ละเซลล์ และเมื่อเสร็จสิ้นกระบวนการ การเปลี่ยนรูป (Transformation) จะได้มิติของข้อมูล (Dimension) ขนาด 384 มิติซึ่งได้มาจากจำนวนเซลล์ย่อยคูณกับค่าทิศทาง ดังภาพที่ 2.72



ภาพที่ 2.72 ภาพตัวอย่างการโอเวอร์แลปบล็อกและการนอร์มอลไลซ์ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ จาก การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับการวิเคราะห์องค์ประกอบหลักและซัพพอร์ตเวกเตอร์แมชชีน โดย อนุสรณ์ อุ้นท้าว และ สมปอง เวฬุวนาธร, 2557, การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 10

จากการทดลองปรากฏว่าขั้นตอนวิธีการ HOG - PCA - SVM มีประสิทธิภาพการรู้จำสูงสุด ให้ความถูกต้องมากที่สุดถึง 96.00% และมีผลรวมระยะเวลาที่ใช้้น้อยสุดคือ 59.84 วินาที (อนุสรณ์ อุ้นท้าว และ สมปอง เวฬุวนาธร, 2557)

หลังจากที่ได้ทำการศึกษาทฤษฎีที่เกี่ยวข้อง พร้อมทั้งหลักในการทำงานของการตรวจจับบริเวณใบหน้าและการจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้าแล้ว ผู้วิจัยได้แสดงตารางสรุปเทคนิคและคุณลักษณะที่ใช้ในงานวิจัยที่เกี่ยวข้องดังตารางที่ 2.5 ทำให้ทราบขั้นตอนการรู้จำอารมณ์บนใบหน้าของมนุษย์นั้น มีหลักการการทำงานที่คล้ายกัน คือ รับภาพเข้ามาประมวลผล ตรวจจับบริเวณใบหน้า จากนั้นทำการหาคุณลักษณะเฉพาะ เมื่อได้คุณลักษณะเฉพาะในรูปแบบตัวเลข และจัดเก็บลงในฐานข้อมูลของระบบแล้ว จากนั้นก็ทำการฝึกสอนตัวจำแนกเกี่ยวกับการรู้จำอารมณ์ เมื่อสอนระบบเรียบร้อยแล้วจึงทำการทดสอบระบบและวัดผล หลักการและขั้นตอนเหล่านี้ น่าจะนำมาใช้งานกับการจำแนกอารมณ์พื้นฐานบนใบหน้าของตัวการ์ตูนระดับตัดแปดได้ด้วย ซึ่งผู้วิจัยคิดว่าการจำแนกอารมณ์จากใบหน้าที่การ์ตูนด้วยระบบอัตโนมัติ นั้น เป็นประโยชน์ให้กับบุคคลที่สนใจเกี่ยวกับงานทางด้านการประมวลผลภาพเชิงดัชนีระดับสูง ที่เกี่ยวกับการแยกแยะอารมณ์พื้นฐานของตัวการ์ตูน

ตารางที่ 2.5

สรุปเทคนิคในการตรวจจับบริเวณใบหน้าและเทคนิคในการจำแนกอารมณ์พื้นฐานของงานวิจัยที่เกี่ยวข้อง

ผู้วิจัย	ลักษณะข้อมูลเข้า		ลักษณะบริเวณใบหน้า		เทคนิคในการตรวจจับใบหน้า	จำแนกอารมณ์		คุณลักษณะที่ใช้ในการจำแนก	ตัวจำแนก (Classifiers)
	ภาพคนจริง	ภาพการ์ตูน	ไม่มีสิ่งบดบัง	มีสิ่งบดบังใบหน้า เช่น หน้าม้า ไรผม ฯลฯ		ไม่	ใช่		
Takayama, Johan และ Nishita		/		/	คุณลักษณะเกี่ยวกับสี และขอบภาพ	/		-	-
Chau		/		/	ฮาร์โลคฟ์พีเจอร์	/		-	-
Arumugam	/		/		แบบจำลองอวัยวะหลักบนใบหน้า เช่น ตา จมูก ปาก ฯลฯ		/ (โกรธ, รังเกียจ, มีความสุข)	Eigen vector, Eigen Value, Fisher's Linear Discriminant, Singular Value Decomposition (SVD)	Neural Network
Chibelushi และ Bourel	/		/		คุณลักษณะเกี่ยวกับ รูปทรง การเคลื่อนไหว สี ลักษณะพื้นผิว และ ตำแหน่ง		/ (มีความสุข, เศร้า, ประหลาดใจ, โกรธ, กลัว, รังเกียจ)	Action units Recognition	Bayes Classifier
งานวิจัยนี้		/		/	ฮาร์โลคฟ์พีเจอร์		/ (โกรธ, มีความสุข, เศร้า)	HOG	Decision tree k-NN SVM

จากการศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้องกับการตรวจจับใบหน้า (Face Detection) แล้วผู้วิจัย จึงมองเห็นถึงเทคนิคที่จะนำมาใช้กับงานในการจำแนกอารมณ์จากใบหน้าที่การรู้ตัวด้วยระบบอัตโนมัติ ในส่วนของการตรวจจับบริเวณที่เป็นใบหน้า ผู้วิจัยมีความสนใจในเทคนิคของ Haar like Feature เป็นอย่างมาก เพราะจากการที่ได้ศึกษางานวิจัยที่เกี่ยวข้องกับเทคนิคของ Haar-like feature ให้ผลในการตรวจจับบริเวณใบหน้าได้ดี ซึ่งเทคนิคของ Haar like Feature โดยส่วนใหญ่ ถูกนำมาใช้ตรวจจับบริเวณที่เป็นส่วนของใบหน้าของมนุษย์เท่านั้น ผู้วิจัยจึงคิดว่าเป็นแนวทางที่น่าสนใจที่จะนำมาประยุกต์ใช้ในการตรวจจับบริเวณที่เป็นใบหน้าของการรู้ตัว และผู้วิจัยมีความสนใจที่จะนำเทคนิคของ HOG มาใช้ในการสกัดหาคุณลักษณะเฉพาะของใบหน้าที่การรู้ตัว เนื่องจากจากการที่ได้ศึกษางานวิจัยที่เกี่ยวข้อง ทำให้ทราบถึงประสิทธิภาพในการหาคุณลักษณะของ HOG นั้นมีความละเอียดผู้วิจัยจึงมีแนวคิดว่าจะนำ HOG มาช่วยในการสกัดคุณลักษณะบนใบหน้าที่การรู้ตัวจะสามารถทำได้ดี อีกทั้งยังสามารถนำคุณลักษณะเฉพาะที่ได้ไปทำการจำแนกข้อมูลและอาจจะให้ค่าความถูกต้องในการจำแนกได้ดีเหมือนกับการจำแนกอารมณ์บนใบหน้าของมนุษย์ อีกทั้งยังมีประโยชน์ต่อผู้ที่สนใจจะนำไปพัฒนาต่ออีกด้วย ซึ่งในบทความต่อไป ผู้วิจัยจะกล่าวถึง วิธีในการดำเนินงานวิจัย

บทที่ 3

วิธีการดำเนินงานวิจัย

ในงานวิจัยนี้จะทำการศึกษาปัจจัยที่มีผลต่อประสิทธิภาพของเครื่องมือในการตรวจจับและจำแนกอารมณ์บนใบหน้าของตัวการ์ตูนที่มีลักษณะใบหน้าคล้ายหรือเหมือนกันกับมนุษย์ทางด้านข้อมูลชุดฝึกสอนระบบ ด้านคุณลักษณะเด่นบนใบหน้า และด้านการปรับปรุงอัลกอริธึมในการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน เนื้อหาในบทนี้จะแสดงให้เห็นถึงแนวทางในการพัฒนาระบบ และกล่าวสรุปเกี่ยวกับระเบียบวิธีที่ใช้ในการทำวิจัยและการออกแบบระบบให้สามารถทำงานได้ตามวัตถุประสงค์

3.1 ระเบียบวิธีในการดำเนินงานวิจัย

การวิจัยเพื่อพัฒนา ระบบการจำแนกอารมณ์จากใบหน้าที่การ์ตูนด้วยระบบอัตโนมัติ มีขั้นตอนการดำเนินการแบ่งเป็นสี่ขั้นตอน สรุปได้ดังนี้

ขั้นตอนที่ 1 การศึกษาความรู้พื้นฐานงานวิจัยและทฤษฎีที่เกี่ยวข้องในหัวข้อต่อไปนี้

- 1.1 ความรู้พื้นฐานเกี่ยวกับการ์ตูน (Cartoon Basic Knowledge)
- 1.2 ความรู้พื้นฐานเกี่ยวกับอารมณ์ (Emotion Basic Knowledge)
- 1.3 ความรู้พื้นฐานเกี่ยวกับการประมวลผลและวิเคราะห์ภาพดิจิทัล (Digital Image Processing & Analysis Fundamental)
- 1.4 ความรู้พื้นฐานเกี่ยวกับการวิเคราะห์ลักษณะเฉพาะของภาพด้วยคุณสมบัติการแยกลักษณะเด่นของวัตถุ (Feature Extraction for Image Analysis)
- 1.5 การจำแนกประเภทของข้อมูล (Classification)
- 1.6 ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการตรวจจับใบหน้า (Face Detection)
- 1.7 ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้า (Emotion Classification)
- 1.8 ซอฟต์แวร์ที่จำเป็นที่ต้องนำมาใช้ในการพัฒนาและวิเคราะห์ระบบการจำแนกอารมณ์จากใบหน้าที่การ์ตูนด้วยระบบอัตโนมัติ

ขั้นตอนที่ 2 วิเคราะห์และออกแบบโครงสร้างของระบบ

2.1 เตรียมข้อมูลภาพการ์ตูนแอนิเมชัน ได้แก่ ภาพการ์ตูนที่มีการแสดงออกทางสีหน้าด้วย อารมณ์โกรธ อารมณ์มีความสุข และอารมณ์เศร้า

2.2 ออกแบบระบบการตรวจจับบริเวณใบหน้าของตัวการ์ตูน

2.3 ออกแบบระบบการจำแนกอารมณ์จากใบหน้าการ์ตูนด้วยระบบอัตโนมัติ

2.4 จัดเตรียมชุดข้อมูลภาพเพื่อนำมาสอนระบบ และสอนระบบให้รู้จำการจำแนกอารมณ์พื้นฐานสามอารมณ์ ได้แก่ อารมณ์โกรธ อารมณ์มีความสุข และอารมณ์เศร้า

2.5 จัดเตรียมชุดข้อมูลภาพเพื่อนำมาทดสอบระบบ และทดสอบระบบในการจำแนกอารมณ์บนใบหน้าของตัวการ์ตูน

ขั้นตอนที่ 3 ทำการทดลองเพื่อวัดความถูกต้องของระบบ

3.1 วัดความถูกต้องของระบบ เพื่อวิเคราะห์ถึงข้อจำกัดของการทดลองและความสามารถในการประมวลผลของระบบที่ได้

3.2 ทำการทดลองจากชุดข้อมูลที่จัดเตรียมไว้สำหรับการทดสอบระบบ เพื่อแสดงผลลัพธ์ที่ได้ จากการฝึกสอนให้ระบบรู้จำและจำแนกอารมณ์บนใบหน้าของตัวการ์ตูน

3.3 วิเคราะห์และสรุปผลการทดลอง

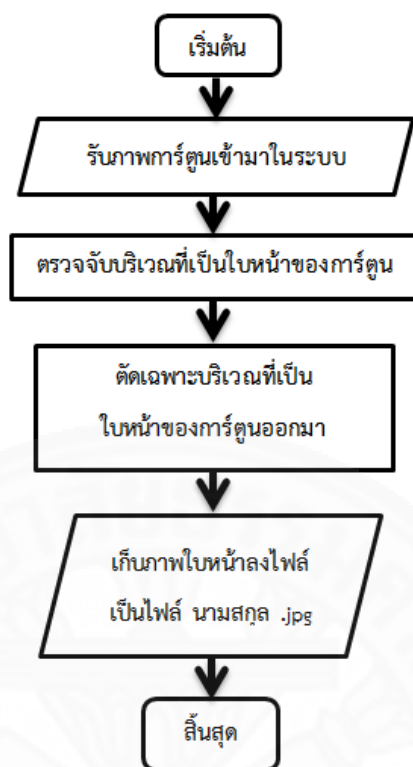
ขั้นตอนที่ 4 จัดทำเอกสารวิทยานิพนธ์ ตีพิมพ์ผลงานและเผยแพร่หลักการในการทำงานวิจัย

3.2 ขั้นตอนการดำเนินงาน

ขั้นตอนในการทำงานของระบบ สามารถอธิบายโดยแบ่งออกเป็นสองขั้นตอน ได้แก่ (1) ขั้นตอนในการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูนจากภาพในวิดีโอและภาพสลิปคั่นจากอินเทอร์เน็ต และ (2) ขั้นตอนในการจำแนกอารมณ์จากใบหน้าการ์ตูน จะอธิบายแยกตามหัวข้อดังนี้

3.2.1 การตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน (Face Detection Of Cartoons)

ในการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูนนั้น มีขั้นตอนการดำเนินงานโดยสรุป ซึ่งสามารถอธิบายได้ดังภาพที่ 3.1 เริ่มจากทำการรับภาพการ์ตูนเข้ามาในระบบ หลังจากนั้น ทำการตรวจจับหาบริเวณที่เป็นส่วนของใบหน้าการ์ตูน เมื่อได้ส่วนที่เป็นบริเวณของใบหน้าตัวการ์ตูนแล้วจึงทำการตัดเอาเฉพาะบริเวณที่เป็นใบหน้าออกมา เพื่อนำไปใช้ในการสอนจำแนกอารมณ์ (หัวข้อ 3.2.2) ต่อไป



ภาพที่ 3.1 แสดงขั้นตอนการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน

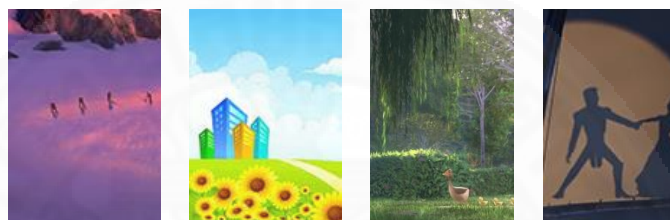
3.2.1.1 สมมติฐานของงานวิจัยในการตรวจจับบริเวณใบหน้าของตัวการ์ตูน

จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องในบทที่สอง ผู้วิจัยได้ตั้งสมมติฐานที่คาดว่าน่าจะส่งผลต่อปัจจัยในด้านประสิทธิภาพของเครื่องมือในการตรวจจับบริเวณใบหน้าของตัวการ์ตูนได้ดังนี้

1. เครื่องมือที่ใช้ในการตรวจจับบริเวณที่เป็นใบหน้าที่ใช้ Haar like Feature มาตรฐานซึ่งใช้ในการตรวจจับใบหน้านามมนุษย์จริงได้ผลดีกว่าเมื่อนำมาตรวจจับบริเวณใบหน้าของการ์ตูนทั้งสี่ระดับ
2. การเปลี่ยนชุดข้อมูลการฝึกสอน จากใบหน้านามมนุษย์จริงเป็นภาพใบหน้าการ์ตูนทั้งสี่ระดับ จะช่วยให้เครื่องมือสามารถตรวจจับบริเวณใบหน้าของการ์ตูนได้ดีขึ้น

3.2.1.2 การเตรียมข้อมูลการทดลอง

งานวิจัยนี้ทำการแบ่งข้อมูลในการทดลองกับการ์ตูนทั้งสี่ระดับ ได้แก่ (1) การ์ตูนแบบตัดแปดอย่างมาก (2) การ์ตูนแบบตัดแปดและประติษฐ์ (3) การ์ตูนแบบตัดแปดเล็กน้อย และ (4) การ์ตูนแบบเหมือนจริง โดยแต่ละระดับจะแบ่งออกเป็นสองชุด ชุดแรกเป็นภาพการ์ตูนแต่ละระดับ อย่างละ 99 ภาพ ดังภาพที่ 3.2 รวมเป็น 396 ภาพ และชุดข้อมูลที่ไม่ใช่ภาพใบหน้า (Negative Face) จำนวน 200 ภาพ ดังภาพที่ 3.3



3.2.1.3 วิธีการทดลอง

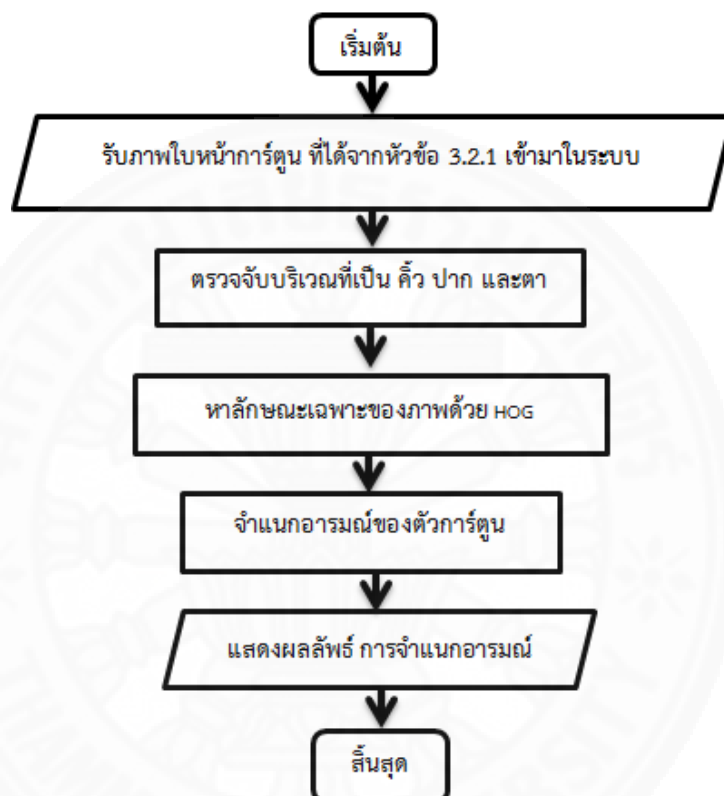
การทดลองด้านการตรวจจับบริเวณใบหน้าของตัวการ์ตูน โดยใช้ Haar Like Feature คือ คำนวณหาค่าตำแหน่งที่เป็นบริเวณใบหน้าจากชุดข้อมูลฝึกสอน จากนั้นทำการคำนวณหาค่าตำแหน่งของภาพที่เป็นบริเวณใบหน้าจากชุดข้อมูลทดสอบระบบ ระบบจะแสดงกรอบสี่เหลี่ยมสีแดง ณ ตำแหน่งที่คำนวณว่าเป็นบริเวณใบหน้า โดยทำการทดลองออกเป็น 2 ขั้นตอน

1. วิธีการทดลอง ใช้ชุดข้อมูลฝึกสอนระบบ Haar like Feature มาตรฐานที่ไม่ได้มีการเปลี่ยนแปลงชุดข้อมูลในการฝึกสอน จากนั้นใช้ข้อมูลทดสอบระบบด้วยภาพใบหน้าของการ์ตูนที่ทั้งสี่ระดับ บันทึกผลการทดลองในรูปของตารางคอนฟิวส์ชันเมทริก (Confusion Matrix) และสรุปผลเป็นค่าความถูกต้อง

2. วิธีการทดลอง ใช้ชุดข้อมูลฝึกสอนระบบ Haar like Feature มาตรฐานที่ได้มีการเปลี่ยนแปลงชุดข้อมูลในการฝึกสอน จากนั้นใช้ข้อมูลทดสอบระบบด้วยภาพใบหน้าของการ์ตูนที่ทั้งสี่ระดับ บันทึกผลการทดลองในรูปของตารางคอนฟิวส์ชันเมทริก (Confusion Matrix) และสรุปผลเป็นค่าความถูกต้อง

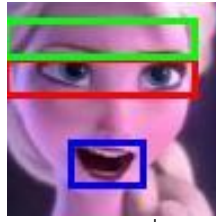
3.2.2 การจำแนกอารมณ์จากใบหน้าของตัวการ์ตูน (Classification Emotion Of Cartoons)

ในการจำแนกอารมณ์จากใบหน้าของตัวการ์ตูนนั้น มีขั้นตอนการดำเนินงานโดยสรุป ซึ่งสามารถอธิบายได้ดังภาพที่ 3.4

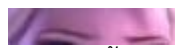


ภาพที่ 3.4 ผังแสดงขั้นตอนการจำแนกอารมณ์บนใบหน้าของตัวการ์ตูน

เริ่มจากการรับภาพการ์ตูนที่ได้จากการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูนในหัวข้อที่ 3.2.1 เข้ามาในระบบ ทำการตรวจจับหาบริเวณที่เป็น ปาก ตา บนใบหน้าของการ์ตูน จากนั้นทำการตีกรอบบริเวณที่เป็นปากและตา เมื่อได้บริเวณที่เป็นตาของตัวการ์ตูนแล้ว ผู้วิจัยได้ทำการเขียนอัลกอริธึมเพิ่มเติมขึ้นมาเพื่อหาบริเวณที่เป็นคิ้วของตัวการ์ตูน ซึ่งอ้างอิงจากตำแหน่งของดวงตาทั้งสองข้างของตัวการ์ตูนเป็นหลัก และทำการตีกรอบบริเวณที่เป็นคิ้ว เมื่อได้บริเวณ คิ้ว ตา ปาก ดังภาพที่ 3.5 แล้วจึงทำการตัดทั้งสามส่วนแยกออกมา ดังภาพที่ 3.6



ภาพที่ 3.5 แสดงภาพการตรวจจับบริเวณที่เป็น คิ้ว ตา ปาก บนใบหน้าการ์ตูน



(ก) คิ้ว



(ข) ตา



(ค) ปาก

ภาพที่ 3.6 ตัดแยกเฉพาะส่วนที่เป็น (ก) คิ้ว (ข) ตา (ค) ปาก

จากนั้นนำสามส่วนที่ตัดออกมาได้ มาทำการประมวลผลภาพ (Image Processing) เมื่อประมวลผลภาพเรียบร้อยแล้วจึงนำมาหาลักษณะเฉพาะของภาพ (Feature Extraction) เพื่อนำค่าตัวเลขของชุดข้อมูล ไปทำการฝึกสอนและทดสอบระบบ เพื่อให้ระบบสามารถจำแนกอารมณ์บนใบหน้าของการ์ตูนด้วยเทคนิคในการจำแนกข้อมูล (Classification) ในลำดับถัดไป อารมณ์พื้นฐานในการจำแนก มีสามคลาส (Classes) โดยแสดงตัวอย่างใบหน้าและส่วนสำคัญของใบหน้าที่น่าจะมีส่วนช่วยในการจำแนกอารมณ์บนใบหน้าของตัวการ์ตูนแสดง ดังตารางที่ 3.1

ตารางที่ 3.1

ใบหน้าและส่วนสำคัญของใบหน้าที่น่าจะมีส่วนช่วยในการจำแนกอารมณ์บนใบหน้าของตัวการ์ตูน

ภาพใบหน้า	คิ้ว	ตา	ปาก	Class
				Sadness
				Happiness
				Anger

3.2.2.1 สมมติฐานของงานวิจัยในการจำแนกอารมณ์จากใบหน้าของตัว

การ์ตูน

หลังจากที่ได้ทำการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องเกี่ยวกับการจำแนกอารมณ์บนใบหน้าของการ์ตูน พบว่าหากทำการจำแนกลักษณะเฉพาะของภาพใบหน้าการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์นั้นคาดว่าจะสามารถนำมาใช้ในการจำแนกลักษณะเฉพาะได้ดี เนื่องจากการแสดงอารมณ์บนใบหน้าที่มีการเคลื่อนไหวของมดกล่อมเนื้อบริเวณใบหน้า โดย HOG เก็บทิศทางค่าความสว่างของภาพอย่างละเอียด ผู้วิจัยจึงได้ตั้งสมมติฐานดังนี้

1. ลักษณะเฉพาะ (Feature Extraction) ของ HOG ช่วยทำให้สามารถจำแนกอารมณ์จากใบหน้าของตัวการ์ตูนแบบเหมือนหรือคล้ายมนุษย์ได้ผลดี

จากเทคนิคการวาดภาพอารมณ์บนใบหน้าของตัวการ์ตูนนั้นจะประกอบไปด้วยสามองค์ประกอบหลักที่สำคัญ อารมณ์โกรธจะประกอบไปด้วย คิ้ว แก้ม ปาก ส่วนอารมณ์มีความสุขและเศร้าจะประกอบไปด้วย คิ้ว ตา ปาก จึงเป็นที่มาของสมมติฐานข้อที่สอง

2. การสกัดคุณลักษณะของ HOG จากภาพทั้งใบหน้าที่แยกส่วนเฉพาะอวัยวะบนใบหน้าที่สำคัญ น่าจะให้ผลการจำแนกที่แตกต่างกัน

และจากงานวิจัยที่เกี่ยวข้องเกี่ยวกับ HOG พบว่าเมื่อนำ HOG ไปใช้งานร่วมกับตัวจำแนกแบบ SVM สามารถให้ผลในการจำแนกได้ดี จึงเป็นที่มาของสมมติฐานข้อที่สาม

3. ตัวจำแนกแบบ SVM สามารถจำแนกอารมณ์บนใบหน้าของการ์ตูนได้ดีที่สุดเมื่อเปรียบเทียบกับตัวจำแนกแบบ DT และ k-NN

3.2.2.2 การเตรียมข้อมูลการทดลอง

งานวิจัยนี้ ทำการทดลองกับภาพใบหน้าตัวการ์ตูนเป็นหลัก ดังนั้นเพื่อให้ตอบสนองสมมติฐานที่ได้ตั้งไว้ ข้อมูลที่ผู้วิจัยจะนำมาใช้จำเป็นต้องมีชุดข้อมูลสำหรับการทดลอง ซึ่งประกอบด้วย ชุดข้อมูล Ground Truth Data และ ชุดข้อมูลที่ใช้ฝึกสอนและทดสอบระบบ ซึ่งจะอธิบายแยกตามหัวข้อดังนี้

1. ชุดข้อมูล Ground Truth Data

งานวิจัยนี้ ชุดข้อมูล Ground Truth Data คือ ชุดข้อมูลทดสอบผลของค่าความความเห็นที่ตรงกันในการจำแนกอารมณ์จากใบหน้าการ์ตูน ซึ่งข้อมูลชุดนี้จะใช้มนุษย์เป็นผู้ตอบแบบสอบถาม เพื่อนำภาพที่ได้ไปทำการทดลองในขั้นตอนต่อไป ด้านการออกแบบแบบสอบถามสำหรับชุดข้อมูลนี้ จะทำการออกแบบโดยการใช้ Google Sheet เพื่อสร้างแบบสอบถามและจะแบ่งข้อมูลออกเป็นสามชุดจำแนกตามอารมณ์ ชุดละ 100 ภาพ ซึ่งใช้ผู้ทดสอบประมาณ 10 คนมาช่วยในงานวัดผล เพื่อให้สอดคล้องกับสมมติฐานที่ได้ตั้งไว้ ตัวอย่างแบบสอบถามดังภาพที่ 3.7



- อารมณ์โกรธ
 อารมณ์เศร้า
 อารมณ์มีความสุข
 ไม่สามารถจำแนกได้

ภาพที่ 3.7 ตัวอย่างแบบสอบถามที่ให้นัก্ষัตริย์ทดสอบ

จากแบบสอบถามผู้วิจัยจะวัดความคิดเห็นที่ตรงกันที่ได้จากจำนวนผู้ทำแบบสอบถาม งานวิจัยนี้ได้นำภาพการ์ตูนแอนิเมชันมาใช้ในการทดลอง ซึ่งในการทดลองผู้วิจัยได้จัดทำแบบสอบถามเพื่อให้นัก্ষัตริย์ทดสอบและหลังจากนั้นจึงเลือกภาพที่มีความเห็นที่ตรงกันไม่ต่ำกว่า 7 จาก 10 คน มาใช้ในการทดลอง โดยแยกตามประเภทของอารมณ์เป็นอารมณ์ละ 100 ภาพ รวมทั้งสิ้น 300 ภาพ โดยตารางสรุปผลจากแบบสอบถามที่ให้นัก্ষัตริย์เข้ามาทำการทดสอบผลแสดงได้ดังตารางที่ 3.2

ตารางที่ 3.2

ตัวอย่างโดยสรุปของการจำแนกอารมณ์จากใบหน้าการ์ตูนด้วยมนุษย์จำนวน 10 คน

คน รูป	1	2	3	4	5	6	7	8	9	10	Anger (A)	Happy (H)	Sad (S)	None (N)	%
1	N	A	A	A	A	A	A	A	A	A	9	-	-	1	90%
2	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
3	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
4	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
5	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
6	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
7	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
8	A	A	A	A	A	A	A	A	A	A	10	-	-	-	100%
.
.
.
100	S	A	A	A	N	A	A	A	S	A	7	2	-	1	70%
101	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
102	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
103	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%

คน รูป	1	2	3	4	5	6	7	8	9	10	Anger (A)	Happy (H)	Sad (S)	None (N)	%
	104	H	H	H	H	H	H	H	H	H	H	-	10	-	-
105	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
106	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
107	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
108	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
.
.
.
200	H	H	H	H	H	H	H	H	H	H	-	10	-	-	100%
201	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
202	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
203	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
204	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
205	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
206	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
207	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
208	S	S	S	S	S	S	S	S	S	S	-	-	10	-	100%
.
.
.
300	S	S	S	S	H	N	S	S	S	S	-	1	8	1	80%

จากตารางที่ 3.2 แสดงให้เห็นถึงการจำแนกอารมณ์จากใบหน้า การ์ตูนด้วยมนุษย์จำนวน 10 คน ซึ่งผลในการทดสอบมนุษย์สามารถจำแนกอารมณ์บนใบหน้าการ์ตูนจากภาพการ์ตูนได้มากที่สุดถึง 100% หรือ 10 คนมีความเห็นที่ตรงกัน และจำแนกอารมณ์บนใบหน้าการ์ตูนจากภาพการ์ตูนได้น้อยที่สุดคือ 70% หรือ 7 จาก 10 คนมีความเห็นที่ตรงกัน ซึ่งจากทดสอบแบบสอบถามของมนุษย์ สรุปได้ว่าภาพการ์ตูนที่นำมาใช้ในการทดลองสามารถนำมาใช้ในการฝึกสอนและทดสอบระบบได้

2. ชุดข้อมูลฝึกสอนและทดสอบระบบ

งานวิจัยนี้ ชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบระบบ คือ ชุดข้อมูลภาพที่ได้มาจากระบบการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูนในหัวข้อที่ 3.2.1 ซึ่งลักษณะของภาพจะเป็นการตัดภาพออกมาเฉพาะบริเวณที่เป็นใบหน้าของการ์ตูน เห็นลักษณะเด่นเฉพาะของภาพ (Feature) ชัดเจน ซึ่งประกอบด้วยสามส่วนหลัก คือ คิ้ว ตา และ ปาก แสดงดังภาพที่ 3.8



ภาพที่ 3.8 ตัวอย่างชุดข้อมูลฝึกสอนระบบ ที่ตัดเฉพาะส่วนที่เป็นบริเวณใบหน้าของการ์ตูน

ผู้วิจัยได้ทำการแบ่งหมวดหมู่ของสามองค์ประกอบ ออกเป็นทั้งหมด สามอารมณ์ ได้แก่ อารมณ์เศร้า (Sadness) อารมณ์โกรธ (Anger) อารมณ์สนุกสนาน (Joy)

3.2.2.3 วิธีการทดลอง

ในการทดลองจะใช้ ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histograms of Oriented Gradients: HOG) ในการหาค่าคุณลักษณะเฉพาะของภาพใบหน้าการ์ตูน ในงานวิจัยนี้จะแบ่งการทดลองออกเป็นสองส่วน ได้แก่ การหาค่าคุณลักษณะเฉพาะของภาพใบหน้าการ์ตูนที่ได้จากหัวข้อ 3.2.1 และการหาค่าคุณลักษณะเฉพาะของภาพใบหน้าการ์ตูนที่ได้ทำการตัดส่วนที่เป็นคิ้ว ตา ปาก จากรูปที่ได้ในหัวข้อ 3.2.1 ด้วย

เมื่อได้คุณลักษณะเฉพาะของชุดข้อมูลที่ใช้ในการฝึกสอนระบบแล้ว จากนั้นทำการสอนระบบจำแนกตามแต่ละอารมณ์ เพื่อให้ระบบเข้าใจขั้นตอนในการจำแนกอารมณ์ ผู้วิจัยจึงได้กำหนดชุดของข้อมูลที่ใช้ในการสอนระบบแยกเป็นอารมณ์สำหรับภาพใบหน้าการ์ตูนใช้อารมณ์ละ 100 ภาพ และภาพจากการคุณลักษณะเฉพาะของภาพใบหน้าการ์ตูนที่ได้ทำการตัดส่วนที่เป็น คิ้ว ตา ปาก อารมณ์ละ 100 ภาพ รวมเป็น 300 ภาพ จากนั้นใช้ตัวจำแนก (Classifiers) แบบต่าง ๆ มาทำการสอนระบบให้รู้จำ ซึ่งในงานวิจัยนี้ จะใช้ต้นไม้การตัดสินใจ (Decision Tree) ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k – Nearest Neighbors) และซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine) มาทำการจำแนกประเภทของอารมณ์ เมื่อได้ผลลัพธ์ในการจำแนกอารมณ์แล้ว จะนำผลที่ได้มาบันทึกลงตารางคอนฟิวชันเมทริกซ์ (Confusion Matrix) เพื่อทำการเปรียบเทียบประสิทธิภาพในการจำแนกกว่า ตัวจำแนกรูปแบบใดที่ให้ผลร้อยละของค่าความถูกต้องมากที่สุดในการจำแนกอารมณ์ ซึ่งในบทถัดไปผู้วิจัยจะอธิบายถึง ผลการวิจัยและอภิปรายผล

บทที่ 4

ผลการวิจัยและอภิปรายผล

4.1 ผลของการทดลอง

ขั้นตอนในการทำการทดลอง สามารถอธิบายโดยแบ่งออกเป็นสองขั้นตอน ได้แก่ ขั้นตอนในการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน และ ขั้นตอนในการจำแนกอารมณ์จากใบหน้าที่การ์ตูน ซึ่งจะอธิบายแยกตามหัวข้อดังนี้

4.1.1 การทดลองการตรวจจับบริเวณที่เป็นใบหน้าของตัวการ์ตูน

ตารางที่ 4.1

ชุดฝึกสอนและทดสอบระบบของการ์ตูนในแต่ละระดับ

จำนวนรอบ ชุดที่	รอบที่ 1	รอบที่ 2	รอบที่ 3
ชุดที่ 1	ชุดฝึกสอน	ชุดฝึกสอน	ชุดทดสอบ
ชุดที่ 2	ชุดฝึกสอน	ชุดทดสอบ	ชุดฝึกสอน
ชุดที่ 3	ชุดทดสอบ	ชุดฝึกสอน	ชุดฝึกสอน

จากสมมุติฐานที่ได้ตั้งไว้ในหัวข้อที่ 3.2.1.1 ได้แก่ (1) เครื่องมือที่ใช้ในการตรวจจับบริเวณที่เป็นใบหน้าที่ใช้ Haar like Feature มาตรฐานสามารถใช้ในการตรวจจับใบหน้ามนุษย์จริงได้ผลดีกว่าเมื่อนำมาตรวจจับบริเวณใบหน้าของการ์ตูนทั้งสี่ระดับ และ (2) หลังจากทำการเปลี่ยนชุดข้อมูลการฝึกสอน จากใบหน้ามนุษย์จริงเป็นภาพใบหน้าที่การ์ตูนทั้งสี่ระดับ จะช่วยให้เครื่องมือสามารถตรวจจับบริเวณใบหน้าของการ์ตูนได้ดีขึ้น การทดลองแบ่งการทำงานออกเป็นสองส่วน ส่วนที่หนึ่ง คือ การตรวจจับบริเวณที่เป็นใบหน้าที่การ์ตูนด้วย Haar like Feature แบบมาตรฐาน และส่วนที่สอง คือ การตรวจจับบริเวณที่เป็นใบหน้าที่การ์ตูนด้วย Haar like Feature แบบปรับปรุงชุดข้อมูลฝึกสอน โดยทั้งสองส่วนจะใช้วิธี 3 – fold validation คือ แบ่งชุดการทดลองออกเป็น 3 ชุด กล่าวคือ นำภาพการ์ตูนที่จัดเตรียมไว้ในแต่ละระดับมาทำการแบ่งส่วนออกเป็นสามส่วน โดยใช้สอง

ส่วนเป็นชุดฝึกและส่วนที่เหลือเป็นชุดทดสอบ ดังตัวอย่างในตารางที่ 4.1 จากนั้นใช้ข้อมูลทดสอบ มาทดสอบระบบบันทึกผลการทดลองลงตารางและหาค่าเฉลี่ยของการทำงานทั้งสามรอบมาสรุปผลเป็นค่าความถูกต้อง

ผลการทดลอง

ในการทดลองจะแบ่งออกเป็นสองขั้นตอน ได้แก่ ขั้นตอนทดลองก่อนการเปลี่ยนแปลงชุดฝึกสอนระบบ (Before Train) และขั้นตอนทดลองหลังการเปลี่ยนแปลงชุดฝึกสอนระบบ (After Train)

1. นำชุดข้อมูลทดสอบ มาทดสอบระบบในการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนด้วยวิธีของ Haar like feature แบบมาตรฐาน โดยที่ไม่มีมีการเปลี่ยนแปลงชุดฝึกสอนระบบ หรือเรียกว่าชุดทดสอบระบบก่อนการเปลี่ยนแปลงชุดฝึกสอน ทำการทดลองแบบ 3 – Fold Validation และบันทึกผลการทดลองลงตารางพร้อมสรุปผลเปอร์เซ็นต์ความถูกต้อง ดังตารางที่ 4.2 ของการ์ตูนแต่ละระดับ

ตารางที่ 4.2

ผลการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนด้วยฮาร์ไลค์พีเจอร์แบบมาตรฐาน

ประเภทการ์ตูน	การตรวจจับ	ตรวจจับได้	เปอร์เซ็นต์ความถูกต้อง
ระดับที่ 1	ตัดแปลงอย่างมาก	8 / 99	8.1%
ระดับที่ 2	ตัดแปลงและประดิษฐ์	33 / 99	33.3%
ระดับที่ 3	การ์ตูนแบบตัดแปลงเล็กน้อย	77 / 99	77.8%
ระดับที่ 4	การ์ตูนแบบเหมือนจริง	83 / 99	83.8%
	ใบหน้ามนุษย์จริง	99 / 99	100.0%

2. นำชุดข้อมูลฝึกสอนเข้าไปทำการสอนระบบและทดสอบระบบ ทำการสลับชุดในการสอนและสทดสอบระบบทั้งหมด 3 ครั้ง และบันทึกผลการทดลองลงตาราง พร้อมสรุปผลเปอร์เซ็นต์ความถูกต้อง ดังตารางที่ 4.3

ตารางที่ 4.3

ผลการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนด้วยฮาร์ไลค์พีเจอร์แบบปรับปรุงชุดฝึกสอน

ประเภทการ์ตูน	การตรวจจับ	ตรวจจับได้	เปอร์เซ็นต์ ความถูกต้อง
ระดับที่ 1 ดัดแปลงอย่างมาก		59 / 99	59.6%
ระดับที่ 2 ดัดแปลงและประดิษฐ์		56 / 99	56.6%
ระดับที่ 3 การ์ตูนแบบดัดแปลงเล็กน้อย		68 / 99	68.7%
ระดับที่ 4 การ์ตูนแบบเหมือนจริง		62 / 99	62.6%
ใบหน้ามนุษย์จริง		47 / 99	47.5%

จากผลการทดลองหลังจากที่ได้ปรับชุดในการฝึกสอนการ์ตูนในระดับที่ 1 (การ์ตูนแบบดัดแปลงอย่างมาก) และระดับที่ 2 (การ์ตูนแบบดัดแปลงและประดิษฐ์) ช่วยให้ผลดีขึ้นจากร้อยละ 8.1 เป็นร้อยละ 59.6 และจากร้อยละ 33.3 เป็นร้อยละ 56.6 ตามลำดับ ซึ่งเป็นไปตามสมมุติฐานที่ 2 ในหัวข้อ 3.2.1.1 แต่สำหรับการ์ตูนระดับที่ 3 (การ์ตูนแบบดัดแปลงเล็กน้อย) และระดับที่ 4 (การ์ตูนแบบเหมือนจริง) ซึ่งมีความคล้ายมนุษย์ค่อนข้างมาก การปรับชุดฝึกสอนกลับทำให้ผลในการตรวจจับแย่ง ไม่เป็นไปตามสมมุติฐานที่ 2 ในหัวข้อที่ 3.2.1.1 จากผลการทดลอง ผู้วิจัยจึงแนะนำว่า หากต้องการตรวจจับใบหน้าการ์ตูนในระดับ 3 และ 4 ควรใช้ Haar like Feature แบบมาตรฐาน หลังจากที่ได้ทำการทดลอง สามารถแบ่งการตรวจจับใบหน้าของการ์ตูนแบ่งออกเป็นทั้งหมดสามกรณี ดังนี้

1. ก่อนปรับชุดในการฝึกสอนระบบ ระบบไม่สามารถตรวจจับใบหน้าการ์ตูนได้ แต่หลังจากที่ทำการปรับชุดฝึกสอนระบบ ระบบสามารถตรวจจับใบหน้าการ์ตูนได้ โดยจะแสดงภาพการ์ตูนในแต่ละระดับของการตรวจจับบริเวณที่เป็นใบหน้าการ์ตูน ดังภาพที่ 4.1

(ก) ก่อนปรับชุดฝึกสอน

(ข) หลังปรับชุดฝึกสอน

ระดับ 1



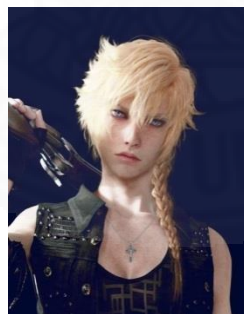
ระดับ 2



ระดับ 3



ระดับ 4



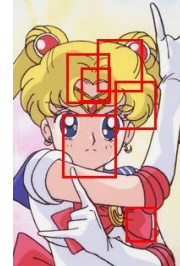
ภาพที่ 4.1 ภาพ (ก) ก่อนปรับชุดฝึกสอนตรวจจับไม่ได้ (ข) หลังปรับชุดฝึกสอนตรวจจับได้

2. ก่อนปรับและหลังปรับชุดฝึกสอนในระบบ ระบบไม่สามารถตรวจจับบริเวณที่เป็นใบหน้าได้ ดังภาพที่ 4.2

(ก) ก่อนปรับชุดฝึกสอน

(ข) หลังปรับชุดฝึกสอน

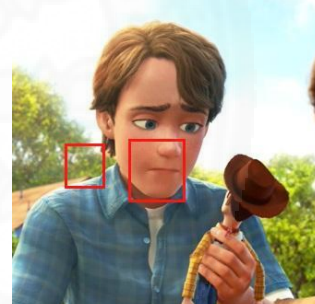
ระดับ 1



ระดับ 2



ระดับ 3



ระดับ 4



ภาพที่ 4.2 ภาพ (ก) ก่อนปรับชุดฝึกสอน (ข) หลังปรับชุดฝึกสอน ไม่สามารถตรวจจับได้

3. ก่อนปรับชุดฝึกสอนในระบบสามารถตรวจจับใบหน้าได้ แต่หลังปรับชุดฝึกสอนในระบบไม่สามารถตรวจจับใบหน้าได้ ดังภาพที่ 4.7



ภาพที่ 4.3 ภาพ (ก) ก่อนปรับชุดฝึกสอนตรวจจับได้ (ข) หลังปรับชุดฝึกสอนตรวจจับไม่ได้

สรุปและอภิปราย

จากการทดลอง การปรับปรุงชุดฝึกสอนระบบทำให้ประสิทธิภาพในการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนในระดับที่ 3 และ 4 (การ์ตูนแบบดัดแปลงเล็กน้อย และ การ์ตูนแบบเหมือนจริง ตามลำดับ) มีการตรวจจับได้น้อยลง ผลที่ได้จากการทดลองตรวจจับใบหน้าการ์ตูนทำให้เข้าใจได้ว่า ขั้นตอนในการตรวจจับบริเวณที่เป็นใบหน้าด้วยเทคนิคของ Haar like Feature แบบ

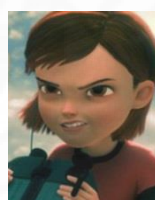
ปรับปรุงชุดข้อมูลฝึกสอนสามารถตรวจจับใบหน้าการ์ตูนในระดับที่ 1 และ 2 (การ์ตูนแบบตัดแปง
 อย่างมาก และ การ์ตูนแบบตัดแปงและประดิษฐ์ ตามลำดับ) สามารถตรวจจับบริเวณใบหน้าได้ดีขึ้น
 แต่ก็ยังอยู่ในระดับปานกลางและยังไม่สามารถนำมาทำการทดลองในการจำแนกอารมณ์บนใบหน้า
 การ์ตูนได้ เพื่อให้งานวิจัยสามารถดำเนินการต่อได้ ผู้วิจัยจึงได้เลือกภาพใบหน้าการ์ตูนระดับที่ 3
 (การ์ตูนแบบตัดแปงเล็กน้อย) ซึ่งพบมากในภาพยนตร์แอนิเมชันทั่วไป มาใช้ในการจำแนกอารมณ์
 โดยรวบรวมภาพการ์ตูนแอนิเมชันทั้งหมด 20 เรื่อง มาใช้ในการทดลอง ดังนี้ ได้แก่ Astro boy, Big
 hero, Bolt, Brave, frozen, How to Train Your Dragon, Inside out, Moana, Mr.Peabody
 and Sherman, Rise of guardians, Tangled, The boss baby, The good dinosaur, The
 princess twins of legendale, The Son of Bigfoot, Toy story, Vir the robot boy, Wreck-it
 ralph, คุณทองแดง และ นางในบาดาล ดังภาพที่ 4.4



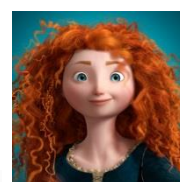
Astro boy



Big hero



Bolt



Brave

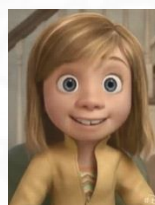


frozen



How to Train Your

Dragon



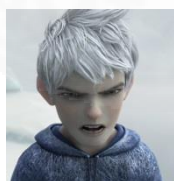
Inside out



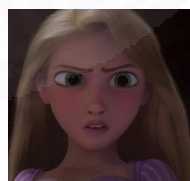
Moana



Mr.Peabody and Sherman



Rise of guardians



Tangled



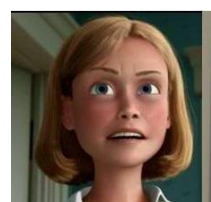
The boss baby



The good dinosaur

The princess twins of
legendale

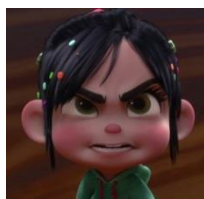
The Son of Bigfoot



Toy story



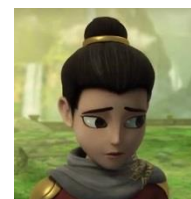
Vir the robot boy



Wreck-it ralph



คุณทองแดง



นางในบาดาล

ภาพที่ 4.4 ตัวอย่างภาพการ์ตูนแอนิเมชันทั้งหมด 20 เรื่อง

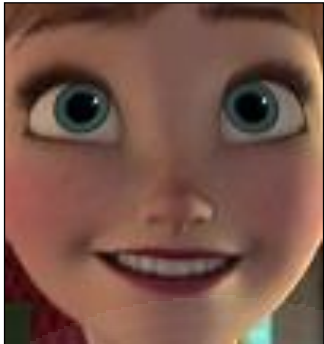
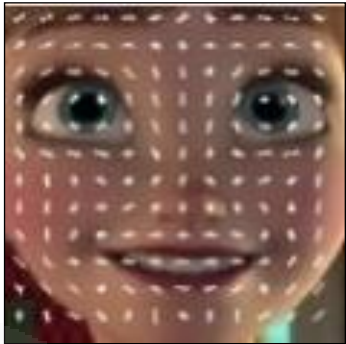
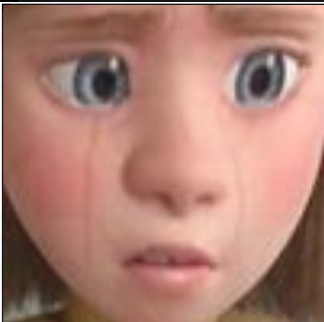

4.1.2 การทดลองการจำแนกอารมณ์จากใบหน้าของตัวการ์ตูน

จากสมมุติฐานในหัวข้อที่ 3.2.2.1 ซึ่งผู้วิจัยได้ทำการตั้งสมมุติฐานในการจำแนกอารมณ์ออกเป็นสามข้อ ได้แก่ (1) ลักษณะเฉพาะ (Feature Extraction) ของ HOG ส่งผลต่อตัวจำแนก ทำให้สามารถจำแนกอารมณ์จากใบหน้าของตัวการ์ตูนแบบเหมือนหรือคล้ายมนุษย์ได้ (2) การสกัดคุณลักษณะของ HOG จากภาพทั้งใบหน้าที่แยกส่วนเฉพาะอวัยวะที่สำคัญ น่าจะให้ผลการจำแนกที่แตกต่างกัน และ (3) ตัวจำแนกแบบ SVM สามารถจำแนกอารมณ์บนใบหน้าของการ์ตูนได้ดีที่สุดเมื่อเปรียบเทียบกับตัวจำแนกแบบ DT และ k-NN งานวิจัยนี้รวบรวมภาพการ์ตูนจากภาพยนตร์แอนิเมชัน 20 เรื่องมาใช้ในการทดลอง โดยเลือกเฉพาะภาพที่เป็นบริเวณใบหน้าของการ์ตูน โดยตัดภาพใบหน้าจากพื้นหลังด้วยเทคนิคฮาร์ไลค์ฟีเจอร์ (Haar-like Feature) และจำแนกประเภทของอารมณ์ออกเป็น 3 ประเภท คือ โกรธ มีความสุข และเศร้า จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลอง มาสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์ (Histogram of Oriented Gradient) ดังแสดงในตารางที่ 4.4 โดยจะได้ค่าคุณลักษณะออกมา 4,356 คุณลักษณะต่อหนึ่งภาพ

ตารางที่ 4.4

ตารางแสดงตัวอย่างภาพของการแสดงออกบนใบหน้าของการ์ตูนในแต่ละอารมณ์

ประเภทอารมณ์	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัดคุณลักษณะด้วย HOG
โกรธ (Anger)		

ประเภท อารมณ์	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัดคุณลักษณะ ด้วย HOG
มีความสุข (Happiness)		
เศร้า (Sadness)		

การสกัดคุณลักษณะเฉพาะของใบหน้าการ์ตูน จะแบ่งการทดลองออกเป็นสี่วิธี (1) สกัดคุณลักษณะทั้งหน้า (2) สกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก (3) สกัดคุณลักษณะแยกส่วน คิ้ว ปาก และ (4) สกัดคุณลักษณะแยกส่วนตา ปาก ซึ่งจะทำการแบ่งการสกัดคุณลักษณะโดยแต่ละอารมณ์จะมี 100 ภาพ รวมทั้งสิ้น 300 ภาพ และปรับขนาดของภาพให้เป็น 100×100 พิกเซล เพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง ขั้นตอนการจำแนกอารมณ์บนใบหน้าการ์ตูนนั้น ผู้วิจัยได้ทำการทดลองโดยใช้ตัวจำแนกสามประเภท คือ (1) ต้นไม้การตัดสินใจ (Decision Tree: DT) (2) ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors: KNN) โดยจากการทดลองเบื้องต้นพบว่า กำหนดค่า $k = 3$ จะให้ผลการทดลองได้ดีกว่าจำนวนอื่น และ (3) ซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine: SVM) จากนั้นทำการเปรียบเทียบประสิทธิภาพการจำแนกประเภทของทั้งสี่วิธี

4.1.2.1 ขั้นตอนการสกัดคุณลักษณะทั้งหน้า

จากที่ได้ทำการศึกษาเทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูน ในหัวข้อที่ 2.2.3 พบว่าการแสดงออกทางอารมณ์บนใบหน้าของตัวการ์ตูนนั้นจะประกอบไปด้วย สามองค์ประกอบหลักๆ ได้แก่ คิ้ว ตา ปาก และรูปภาพการ์ตูนในขั้นตอนนี้ เห็นรายละเอียดบนใบหน้าครบทั้งสามส่วน ผู้วิจัยจึงได้ทำการทดลองนี้ ในขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการ

ทดลองมาสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient) โดยใช้ทั้งหน้าที่ทำการตรวจจับมาได้ด้วยวิธีของ Haar-like Feature

ผลการจำแนกประเภทการสกัดคุณลักษณะทั้งหน้า

ผลการจำแนกจากตัวจำแนกทั้งสามประเภท ของการสกัดคุณลักษณะทั้งหน้า แสดงได้ดังตารางที่ 4.5, 4.6, 4.7 และ 4.8 ตามลำดับ

ตารางที่ 4.5

เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะทั้งหน้า

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	57	15	28	57.0%
Happiness	11	55	34	55.0%
Sadness	13	22	65	65.0%
ค่าเฉลี่ยความถูกต้อง (%)				59.0%

ตารางที่ 4.6

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะทั้งหน้า

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	67	10	23	67.0%
Happiness	10	74	16	74.0%
Sadness	17	7	76	76.0%
ค่าเฉลี่ยความถูกต้อง (%)				72.3%

ตารางที่ 4.7

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะทั้งหมด
ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	90	3	7	90.0%
Happiness	5	86	9	86.0%
Sadness	6	4	90	90.0%
ค่าเฉลี่ยความถูกต้อง (%)				88.7%

ตารางที่ 4.8

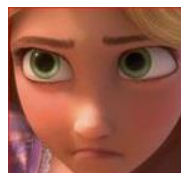
เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะทั้งหมด

ตัวจำแนก	ผลการทดลอง
DT	59.0%
k-NN (k=3)	72.3%
SVM	88.7%

สรุปผลการทดลองสกัดคุณลักษณะทั้งหมด

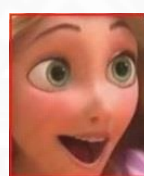
การทดลองพบว่าผลการจำแนกประเภทด้วยวิธีแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 59.0 โดยแสดงผลการทดลองในรูปแบบคอนฟิวชันเมทริกซ์ (Confusion Matrix) ตารางที่ 4.5 พบว่า อารมณ์มีความสุขผิดพลาดมากที่สุดโดย การจำแนกประเภทด้วยวิธีแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 72.3 ดังตารางที่ 4.6 ซึ่งจะพบว่าระบบทำนายอารมณ์โกรธมีความผิดพลาดมากที่สุด และการจำแนกประเภทด้วยวิธีแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) สามารถจำแนกอารมณ์แต่ละประเภทได้ดีที่สุด ให้ผลลัพธ์ในการจำแนกสูงถึงร้อยละ 88.7 ดังตารางที่ 4.7 ซึ่งระบบจำแนกอารมณ์โกรธและเศร้าได้เกือบจะถูกต้องทั้งหมด จากการทดลองทำให้ทราบว่าตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์

แมชชีนมีประสิทธิภาพในการจำแนกที่ดีที่สุด ผู้วิจัยจึงขอแสดงตัวอย่างผลการจำแนกที่ผิดพลาดของ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) ดังภาพที่ 4.5, 4.6 และ 4.7 ตามลำดับ



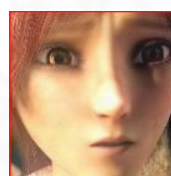
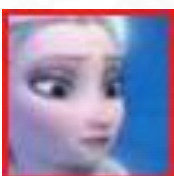
(ก) อารมณ์โกรธเป็นมีความสุข (ข) อารมณ์โกรธเป็นเศร้า

ภาพที่ 4.5 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า ของการสกัดคุณลักษณะทั้งหน้า



(ก) อารมณ์มีความสุขเป็นโกรธ (ข) อารมณ์มีความสุขเป็นเศร้า

ภาพที่ 4.6 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะทั้งหน้า



(ก) อารมณ์เศร้าเป็นโกรธ (ข) อารมณ์เศร้าเป็นมีความสุข

ภาพที่ 4.7 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะทั้งหน้า

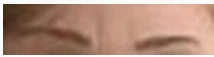

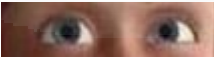
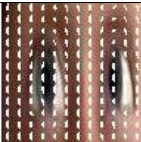

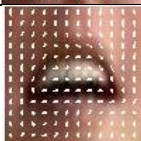
จากภาพที่ 4.5 (ก) จากแบบสอบถามที่ให้มนุษย์จำแนก จากมนุษย์ 10 มีเพียงหนึ่งคน จำแนกผิดพลาดจากอารมณ์โกรธเป็นอารมณ์มีความสุข ภาพที่ 4.5 (ข) มีเพียงหนึ่งคน จำแนกผิดพลาดจากอารมณ์โกรธเป็นอารมณ์เศร้า ซึ่งตรงกับผลการจำแนกของระบบ ภาพที่ 4.6 (ก) และภาพที่ 4.6 (ข) จากแบบสอบถามที่ให้มนุษย์จำแนก มนุษย์ทุกคนตอบตรงกันว่าเป็นอารมณ์มีความสุข ภาพที่ 4.7 (ก) จากแบบสอบถามที่ให้มนุษย์จำแนก มนุษย์ทุกคนตอบตรงกันว่าเป็นอารมณ์เศร้า และภาพที่ 4.7 (ข) มีหนึ่งคนจำแนกว่าไม่มีคำตอบ

4.1.2.2 ขั้นตอนการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

จากการศึกษาเทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูนในหัวข้อที่ 2.2.3 พบว่าการวาดภาพสีหน้าอารมณ์มีความสุภาพและเศร้า จะประกอบไปด้วยองค์ประกอบหลักบนใบหน้า คือ คิ้ว ตา ปาก ซึ่งในการทดลองนี้ผู้วิจัยจะทำการทดลองที่แตกต่างจากหัวข้อที่ 4.1.2.1 คือ จะนำมาทำการตัดแยกเฉพาะส่วนที่สำคัญ ทั้งสามออกมา ซึ่งในขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลอง มาสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient) ดังแสดงในตารางที่ 4.9 โดยจะได้ค่าคุณลักษณะออกมา 13,067 คุณลักษณะต่อหนึ่งภาพ จะทำการแบ่งการสกัดคุณลักษณะโดยแต่ละอารมณ์จะมี 100 ภาพ รวมทั้งสิ้น 300 ภาพ และปรับขนาดของภาพให้เป็น 100×100 พิกเซลเพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง

ตารางที่ 4.9

ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แยกส่วน คิ้ว ตา ปาก

คุณลักษณะเฉพาะแยกส่วน	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัดคุณลักษณะด้วย HOG
คิ้ว		
ตา		
ปาก		

ผลการจำแนกประเภทการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

ผลการจำแนกจากตัวจำแนกทั้งสามประเภทของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก แสดงดังตารางที่ 4.10, 4.11, 4.12 และ 4.13 ตามลำดับ

ตารางที่ 4.10

เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	69	12	19	69.0%
Happiness	13	71	16	71.0%
Sadness	27	8	65	65.0%
ค่าเฉลี่ยความถูกต้อง (%)				68.3%

ตารางที่ 4.11

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	73	9	18	73.0%
Happiness	2	90	8	90.0%
Sadness	9	4	87	87.0%
ค่าเฉลี่ยความถูกต้อง (%)				83.3%

ตารางที่ 4.12

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	94	0	6	94.0%
Happiness	4	93	3	93.0%
Sadness	6	2	92	92.0%
ค่าเฉลี่ยความถูกต้อง (%)				93.0%

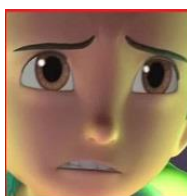
ตารางที่ 4.13

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

ตัวจำแนก	ผลการทดลอง
DT	68.3%
k-NN (k=3)	83.3%
SVM	93.0%

สรุปผลการทดลองการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

การทดลองพบว่าการจำแนกประเภทด้วยวิธีแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 68.3 โดยแสดงผลการทดลองในรูปแบบคอนฟิวชันเมทริกซ์ (Confusion Matrix) ตารางที่ 4.10 พบว่า อารมณ์เศร้าผิดพลาดมากที่สุดโดย การจำแนกประเภทด้วยวิธีแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 83.3 ดังตารางที่ 4.11 ซึ่งจะพบว่าระบบทำนายอารมณ์โกรธมีความผิดพลาดมากที่สุด และการจำแนกประเภทด้วยวิธีแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) สามารถจำแนกอารมณ์แต่ละประเภทได้ดีที่สุด ให้ผลลัพธ์ในการจำแนกสูงถึงร้อยละ 93.0 ดังตารางที่ 4.12 ซึ่งระบบจำแนกอารมณ์โกรธได้ถูกต้องมากที่สุดจากอารมณ์อื่น จากการทดลองทำให้ทราบว่าตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพในการจำแนกดีที่สุดในผู้วิจัยจึงขอแสดงตัวอย่างผลการจำแนกที่ผิดพลาดของซัพพอร์ตเวกเตอร์แมชชีน (SVM) ดังภาพที่ 4.8, 4.9 และ 4.10 ตามลำดับ



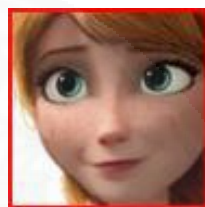
อารมณ์โกรธเป็นเศร้า

ภาพที่ 4.8 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ จากโกรธเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก



(ก) อารมณ์มีความสุขเป็นโกรธ (ข) อารมณ์มีความสุขเป็นเศร้า

ภาพที่ 4.9 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก



(ก) อารมณ์เศร้าเป็นโกรธ (ข) อารมณ์เศร้าเป็นมีความสุข

ภาพที่ 4.10 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก

จากภาพที่ 4.8 จากแบบสอบถามที่ให้มนุษย์มีเพียงหนึ่งคนจำแนกผิดพลาดจากอารมณ์โกรธเป็นเศร้า ซึ่งตรงกับผลการจำแนกของระบบ ภาพที่ 4.9 (ก) และภาพที่ 4.9 (ข) มีหนึ่งคนจำแนกว่าไม่มีคำตอบ ภาพที่ 4.10 (ก) จากแบบสอบถามที่ให้มนุษย์จำแนก มนุษย์ทุกคนตอบตรงกันว่าเป็นอารมณ์เศร้า และภาพที่ 4.10 (ข) มีสองคนจำแนกว่าไม่มีคำตอบ

4.1.2.3 ขั้นตอนการสกัดคุณลักษณะแยกส่วนคิ้วและปาก

จากการศึกษาเทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูนในหัวข้อที่ 2.2.3 พบว่าการวาดภาพสีหน้าอารมณ์โกรธ ประกอบไปด้วยสามองค์ประกอบหลัก ได้แก่ ตา แก้ม ปาก ในส่วนของแก้มนั้นไม่ได้นำมาใช้ในการสกัดคุณลักษณะเนื่องจากไม่ใช่คุณลักษณะเด่นสำหรับการจำแนกอารมณ์ ผู้วิจัยจึงได้ทำการทดลองนี้ โดยใช้สกัดคุณลักษณะเฉพาะแยกส่วน คิ้ว และ ปาก ขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลอง มาสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient) ดังแสดงในตารางที่ 4.14 โดยจะได้ค่าคุณลักษณะออกมาประมาณ 8,712 คุณลักษณะต่อหนึ่งภาพ จะทำการแบ่งการสกัดคุณลักษณะโดยแต่ละอารมณ์จะมี 100 ภาพ รวมทั้งสิ้น 300 ภาพ และปรับขนาดของภาพให้เป็น 100×100 พิกเซล เพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง

ตารางที่ 4.14

ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แยกส่วนคิ้วและปาก

คุณลักษณะเฉพาะแยกส่วน	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัดคุณลักษณะด้วย HOG
คิ้ว		
ปาก		

ผลการจำแนกประเภทการสกัดคุณลักษณะแยกส่วนคิ้วและปาก

ผลการจำแนกจากตัวจำแนกทั้งสามประเภทของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก แสดงดังตารางที่ 4.15, 4.16, 4.17 และ 4.18 ตามลำดับ

ตารางที่ 4.15

เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	65	11	24	65.0%
Happiness	18	71	11	71.0%
Sadness	21	9	70	70.0%
ค่าเฉลี่ยความถูกต้อง (%)				68.7%

ประเภทที่ถูกจัด

ตารางที่ 4.16

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก

		ประเภทที่ระบบทำนาย			
		Anger	Happiness	Sadness	%
ประเภทที่ถูกต้อง	Anger	79	4	17	79.0%
	Happiness	4	90	6	90.0%
	Sadness	5	8	87	87.0%
	ค่าเฉลี่ยความถูกต้อง (%)				85.3%

ตารางที่ 4.17

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก

		ประเภทที่ระบบทำนาย			
		Anger	Happiness	Sadness	%
ประเภทที่ถูกต้อง	Anger	90	1	9	90.0%
	Happiness	3	91	6	91.0%
	Sadness	7	3	90	90.0%
	ค่าเฉลี่ยความถูกต้อง (%)				90.3%

ตารางที่ 4.18

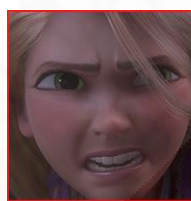
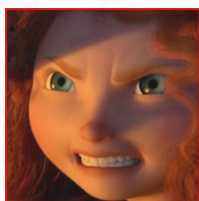
เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วนคิ้วและปาก

ตัวจำแนก	ผลการทดลอง
DT	68.7%

ตัวจำแนก	ผลการทดลอง
k-NN (k=3)	85.3%
SVM	90.3%

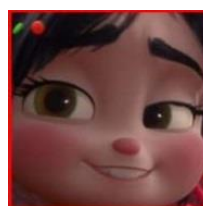
สรุปผลการทดลองการสกัดคุณลักษณะแยกส่วนคิ้วและปาก

การทดลองพบว่าการจำแนกประเภทด้วยวิธีแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 68.7 โดยแสดงผลการทดลองในรูปแบบคอนฟิวชันเมทริกซ์ (Confusion Matrix) ตารางที่ 4.15 พบว่า อารมณ์มีความสุขผิดพลาดมากที่สุดโดย การจำแนกประเภทด้วยวิธีแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 85.3 ดังตาราง ที่ 4.16 ซึ่งจะพบว่าระบบทำนายอารมณ์โกรธมีความผิดพลาดมากที่สุด และการจำแนกประเภทด้วยวิธีแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) สามารถจำแนกอารมณ์แต่ละประเภทได้ดีที่สุด ให้ผลลัพธ์ในการจำแนกสูงถึงร้อยละ 90.3 ดังตารางที่ 4.17 ซึ่งระบบจำแนกอารมณ์โกรธถูกต้องมากที่สุด จากการทดลองทำให้ทราบว่าตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพในการจำแนกดีที่สุดในผู้วิจัยจึงขอแสดงตัวอย่างผลการจำแนกที่ผิดพลาดของซัพพอร์ตเวกเตอร์แมชชีน (SVM) ดังภาพที่ 4.11, 4.12 และ 4.13 ตามลำดับ



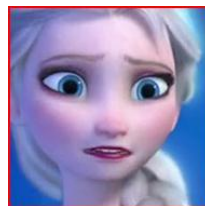
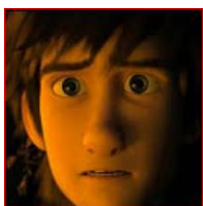
(ก) อารมณ์โกรธเป็นมีความสุข (ข) อารมณ์โกรธเป็นเศร้า

ภาพที่ 4.11 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก



(ก) อารมณ์มีความสุขเป็นโกรธ (ข) อารมณ์มีความสุขเป็นเศร้า

ภาพที่ 4.12 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก



(ก) อารมณ์เศร้าเป็นโกรธ (ข) อารมณ์เศร้าเป็นมีความสุข

ภาพที่ 4.13 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วนคิ้วและปาก

จากภาพที่ 4.11 (ก) และ 4.11 (ข) จากแบบสอบถามมนุษย์จำแนกว่า เป็นอารมณ์โกรธทั้งหมด ภาพที่ 4.12 (ก) มนุษย์จำแนกว่าเป็นอารมณ์มีความสุขทั้งหมดและภาพที่ 4.12 (ข) มีหนึ่งคนจำแนกว่าไม่มีคำตอบ ภาพที่ 4.13 (ก) จากแบบสอบถามที่ให้มนุษย์จำแนกมีหนึ่งคนจำแนกว่าไม่มีคำตอบและภาพที่ 4.13 (ข) มนุษย์ทุกคนตอบตรงกันว่าเป็นอารมณ์เศร้า

4.1.2.4 ขั้นตอนการสกัดคุณลักษณะแยกส่วนตาและปาก

ภาพใบหน้าของตัวการ์ตูนแอนิเมชันที่สืบค้นมาจากอินเทอร์เน็ตนั้น มีข้อจำกัดของชุดข้อมูล เนื่องจากการ์ตูนแอนิเมชันส่วนใหญ่ มักมีผมปกคลุมบริเวณใบหน้า ทำให้เห็นรายละเอียดบริเวณหน้าผากที่มีคิ้วเป็นองค์ประกอบอยู่ได้ไม่ครบถ้วน ซึ่งจากการศึกษาเทคนิคการวาดภาพอารมณ์บนใบหน้าตัวการ์ตูน ในหัวข้อที่ 2.2.3 การวาดภาพสีหน้าอารมณ์มีโกรธ มีความสุข เศร้า นั้นจะมีคิ้วเป็นองค์หลักในการแสดงอารมณ์บนใบหน้า จึงได้ทำการทดลองนี้ขึ้นมาเพื่อตรวจสอบว่า คิ้ว นั้นมีความสำคัญต่อการจำแนกอารมณ์บนใบหน้าของการ์ตูนหรือไม่อย่างไร ในขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลอง มาทำการสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์ (Histogram of Oriented Gradient) โดยใช้การสกัดคุณลักษณะเฉพาะแยกส่วน ตาและปาก ดังแสดงในตารางที่ 4.19 โดยจะได้ค่าคุณลักษณะออกมาประมาณ 8,712 คุณลักษณะต่อหนึ่งภาพ จะทำการแบ่งการสกัดคุณลักษณะโดยแต่ละอารมณ์จะมี 100 ภาพ รวมทั้งสิ้น 300 ภาพ และปรับขนาดของภาพให้เป็น 100×100 พิกเซล เพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง

ตารางที่ 4.19

ตารางแสดงตัวอย่างภาพการสกัดคุณลักษณะเฉพาะตามค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ แยกส่วนตาและปาก

คุณลักษณะเฉพาะ แยกส่วน	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัด คุณลักษณะด้วย HOG
ตา		
ปาก		

ผลการจำแนกประเภทการสกัดคุณลักษณะแยกส่วนตาและปาก

ผลการจำแนกจากตัวจำแนกทั้งสามประเภทของการสกัดคุณลักษณะแยกส่วนตาและปาก แสดงดังตารางที่ 4.20, 4.21, 4.22 และ 4.23 ตามลำดับ

ตารางที่ 4.20

เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะแยกส่วนตาและปาก

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	61	9	30	61.0%
Happiness	18	66	16	66.0%
Sadness	29	7	64	64.0%
ค่าเฉลี่ยความถูกต้อง (%)				63.7%

ประเภทที่ถูกต้อง

ตารางที่ 4.21

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วนตาและปาก

		ประเภทที่ระบบทำนาย			
		Anger	Happiness	Sadness	%
ประเภทที่ถูกต้อง	Anger	41	10	49	41.0%
	Happiness	3	93	4	93.0%
	Sadness	9	9	82	82.0%
	ค่าเฉลี่ยความถูกต้อง (%)				72.0%

ตารางที่ 4.22

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วนตาและปาก

		ประเภทที่ระบบทำนาย			
		Anger	Happiness	Sadness	%
ประเภทที่ถูกต้อง	Anger	73	4	23	73.0%
	Happiness	4	87	9	87.0%
	Sadness	11	9	80	80.0%
	ค่าเฉลี่ยความถูกต้อง (%)				80.0%

ตารางที่ 4.23

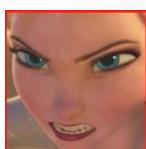
เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วนตาและปาก

ตัวจำแนก	ผลการทดลอง
DT	63.7%

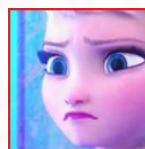
ตัวจำแนก	ผลการทดลอง
k-NN (k=3)	72.0%
SVM	80.0%

สรุปผลการทดลองการสกัดคุณลักษณะแยกส่วนตาและปาก

การทดลองพบว่าการจำแนกประเภทด้วยวิธีแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 63.7 โดยแสดงผลการทดลองในรูปแบบคอนฟิวชันเมทริกซ์ (Confusion Matrix) ตารางที่ 4.20 พบว่า อารมณีสัญญาพิลาตมากที่สุดโดย การจำแนกประเภทด้วยวิธีแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 72.0 ดังตาราง ที่ 4.21 ซึ่งจะพบว่าระบบทำนายอารมณ์โกรธมีความผิดพลาดมากที่สุด และการจำแนกประเภทด้วยวิธีแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) สามารถจำแนกอารมณ์แต่ละประเภทได้ดีที่สุด ให้ผลลัพธ์ในการจำแนกสูงถึงร้อยละ 80.0 ดังตารางที่ 4.22 ซึ่งระบบจำแนกอารมณ์โกรธถูกต้องมากที่สุด จากการทดลองทำให้ทราบว่าตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพในการจำแนกดีที่สุด ผู้วิจัยจึงขอแสดงตัวอย่างผลการจำแนกที่ผิดพลาดของซัพพอร์ตเวกเตอร์แมชชีน (SVM) ดังภาพที่ 4.14, 4.15 และ 4.16 ตามลำดับ



(ก) อารมณ์โกรธเป็นมีความสุข

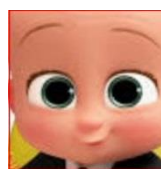


(ข) อารมณ์โกรธเป็นเศร้า

ภาพที่ 4.14 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ (ก) อารมณ์โกรธเป็นมีความสุข และ (ข) อารมณ์โกรธเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนตาและปาก

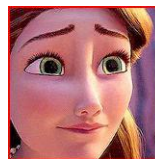
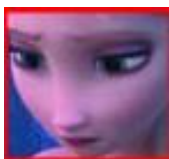


(ก) อารมณ์มีความสุขเป็นโกรธ



(ข) อารมณ์มีความสุขเป็นเศร้า

ภาพที่ 4.15 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข (ก) อารมณ์มีความสุขเป็นโกรธ และ (ข) อารมณ์มีความสุขเป็นเศร้า ของการสกัดคุณลักษณะแยกส่วนตาและปาก



(ก) อารมณ์เศร้าเป็นโกรธ (ข) อารมณ์เศร้าเป็นมีความสุข

ภาพที่ 4.16 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วนตาและปาก

จากภาพที่ 4.14 (ก) จากแบบสอบถามมนุษย์จำแนกว่าเป็นอารมณ์โกรธทั้งหมด และ 4.14 (ข) จากแบบสอบถามมนุษย์จำแนกว่าเป็นอารมณ์เศร้าสองคน ภาพที่ 4.15 (ก) และภาพที่ 4.15 (ข) มีหนึ่งคนจำแนกว่าไม่มีคำตอบ ภาพที่ 4.16 (ก) และภาพที่ 4.16 (ข) มนุษย์ทุกคนตอบตรงกันว่าเป็นอารมณ์เศร้า จากการทดลองนี้ทำให้ทราบว่าหากขาด คิ้ว ซึ่งเป็นองค์ประกอบหลักของการแสดงอารมณ์บนใบหน้า จะทำให้ผลในการจำแนกอารมณ์ทำได้น้อยลงของทุกตัวจำแนกที่ใช้ในการทดลอง

สรุปผลการทดลองจากทั้งสี่วิธี

สรุปผลการทดลองจากรูปภาพตัวอย่างของทั้งสี่วิธี ผลการทดลองความผิดพลาดในการจำแนกประเภทของตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน SVM ในแต่ละอารมณ์มีผลมาจากการแสดงออกทางสีหน้าของการ์ตูนที่มีลักษณะใกล้เคียงกัน จึงส่งผลให้ค่าการกระจายตัวความเข้มของเกรเดียนที่มีค่าใกล้เคียงกันทำให้การจำแนกประเภทอารมณ์ของระบบเกิดความผิดพลาดได้ เนื่องจากภาพบางภาพที่นำมาใช้ในการทดลอง มนุษย์ยังไม่สามารถที่จะจำแนกได้เช่นเดียวกับกับระบบ

หลังจากที่ได้ดำเนินการทดลองในการตรวจจับบริเวณที่เป็นใบหน้าของการ์ตูนแล้วพบว่าของ Haar like Feature แบบมาตรฐานสามารถตรวจจับการ์ตูนในระดับที่ 3 และ 4 (การ์ตูนแบบดัดแปลงเล็กน้อย และ การ์ตูนแบบเหมือนจริง ตามลำดับ) ได้ดี แต่หลังจากที่ปรับปรุงชุดข้อมูลฝึกสอนระบบสามารถตรวจจับใบหน้าการ์ตูนในระดับที่ 1 และ 2 (การ์ตูนแบบดัดแปลงอย่างมาก และ การ์ตูนแบบดัดแปลงและประติษฐ์ ตามลำดับ) ได้ดีขึ้นแต่ก็ยังอยู่ในระดับปานกลาง ให้ค่าความถูกต้องอยู่ที่ประมาณร้อยละ 50 – 60 เท่านั้น จึงยังไม่สามารถนำมาทำการทดลองในการจำแนกอารมณ์บนใบหน้าการ์ตูนได้ เพื่อให้งานวิจัยสามารถดำเนินการต่อไป

ผู้วิจัยจึงได้เลือกภาพใบหน้าการ์ตูนระดับที่ 3 (การ์ตูนแบบดัดแปลงเล็กน้อย) มาใช้ในการจำแนกอารมณ์จากการแสดงออกทางสีหน้าของการ์ตูนทั้งสี่วิธี ได้แก่ (1) สกัดคุณลักษณะทั้งหน้า (2) สกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก (3) สกัดคุณลักษณะแยกส่วนคิ้วและปาก และ (4) สกัดคุณลักษณะแยกส่วนตาและปาก ทำให้สามารถเข้าใจได้ว่า ตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์

แมชชีน (SVM) สามารถจำแนกประเภทของอารมณ์ได้ดีที่สุดซึ่งให้ผลสูงที่สุดกับทุกกรณีในการทดลอง การสกัดคุณลักษณะทั้งหน้า สามารถจำแนกอารมณ์ได้ร้อยละ 88.7 ส่วนการสกัดคุณลักษณะแยก ส่วน คิ้ว ตา ปาก สามารถจำแนกอารมณ์ได้ถึงร้อยละ 93.0 การสกัดคุณลักษณะแยกส่วนคิ้วและปาก สามารถจำแนกอารมณ์ได้ร้อยละ 90.3 และ การสกัดคุณลักษณะแยกส่วนตาและปาก สามารถ จำแนกอารมณ์ได้ร้อยละ 80.0 จากผลการจำแนกทำให้สามารถเข้าใจได้ว่าการจำแนกอารมณ์ที่ดีที่สุด บนใบหน้าการจำแนกจำเป็นต้องสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก มาช่วยในการจำแนกอารมณ์ เพื่อให้ได้ผลของค่าความถูกต้องที่ดีที่สุด แต่ในการทำงานวิจัยอาจจะมีค่าที่ใกล้เคียงกันของการจำแนก อารมณ์บนใบหน้าการจำแนกจากการสกัดคุณลักษณะเฉพาะแยกส่วนคิ้ว ตา ปาก และ คิ้ว ปาก ของตัว จำแนกแบบต้นไม้ตัดสินใจ (DT) และตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) ดัง ตารางที่ 4.24

ตารางที่ 4.24

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT) และตัวจำแนกเพื่อนบ้านที่ ใกล้เคียงที่สุด (k-NN) $k=3$ แบบสกัดคุณลักษณะแยกส่วนคิ้ว ตา ปาก และ คิ้วและปาก

ตัวจำแนก	ผลการทดลอง	
	การจำแนกคุณลักษณะแยกส่วน คิ้ว ตา ปาก	การจำแนกคุณลักษณะแยกส่วน คิ้วและปาก
DT	68.3%	68.7%
k-NN (k=3)	83.3%	85.3%

จากตารางที่ 4.24 ตัวจำแนกแบบต้นไม้ตัดสินใจ (DT) และตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกที่ใกล้เคียงกันมาก การสกัดคุณลักษณะแบบแยกส่วนคิ้วและปาก ค่าความถูกต้องในการจำแนกอารมณ์มีค่ามากกว่าการสกัดคุณลักษณะแบบแยกส่วนคิ้ว ตา ปาก อยู่เล็กน้อย ซึ่งปัญหาดังกล่าวอาจจะเกิดมาจากการ Overfitting ของระบบ ในบทถัดไปผู้วิจัยจะอธิบายถึง สรุปและข้อเสนอแนะ

บทที่ 5

สรุปและข้อเสนอแนะ

5.1 สรุปงานวิจัย

5.1.1 สรุปงานวิจัยด้านการตรวจจับใบหน้าการ์ตูน

การจำแนกอารมณ์จากใบหน้าการ์ตูน จากทฤษฎีและงานวิจัยที่เกี่ยวข้องสามารถอธิบายได้ว่าใบหน้าการ์ตูนชั้นพื้นฐานมีอยู่ด้วยกันทั้งหมด 4 ระดับ ได้แก่ (1) การ์ตูนแบบดัดแปลงอย่างมาก (2) การ์ตูนแบบดัดแปลงและประดิษฐ์ (3) การ์ตูนแบบดัดแปลงเล็กน้อย และ (4) การ์ตูนแบบเหมือนจริง ซึ่งเทคนิคที่นำมาใช้ในการตรวจจับบริเวณที่เป็นใบหน้าการ์ตูน ผู้วิจัยได้เลือกใช้ฮาร์ไลค์ฟีเจอร์ (Haar like Feature) ซึ่งจากผลการวิจัยพบว่าการ์ตูนในระดับที่ 1 และ 2 ก่อนการปรับปรุงชุดฝึกสอนระบบ ให้ผลการตรวจจับใบหน้าได้น้อยเพียงร้อยละ 8.1 และ 33.3 ตามลำดับ แต่หลังจากทำการปรับปรุงชุดฝึกสอนระบบแล้วผลการตรวจจับใบหน้าการ์ตูนสามารถทำได้ดีขึ้น ให้ผลการตรวจจับถึงร้อยละ 59.6 และ 56.6 แต่ก็ยังอยู่ในระดับปานกลางและยังไม่สามารถนำมาทำการทดลองในการจำแนกอารมณ์บนใบหน้าการ์ตูนได้ ส่วนการ์ตูนในระดับที่ 3 และ 4 ก่อนการปรับปรุงชุดฝึกสอนระบบ ให้ผลการตรวจจับสูงสุดถึงร้อยละ 77.8 และ 83.8 ตามลำดับ แต่หลังจากทำการปรับปรุงชุดฝึกสอนระบบแล้วผลการตรวจจับใบหน้าการ์ตูนสามารถทำได้น้อยลง ให้ผลการตรวจจับร้อยละ 68.7 และ 62.6 ตามลำดับ

เนื่องจากชุดข้อมูลที่ใช้ในการฝึกสอนระบบ ส่งผลต่อการตรวจจับใบหน้าของตัวการ์ตูน ซึ่งชุดในการฝึกสอนระบบประกอบไปด้วย บริเวณที่เป็นใบหน้าการ์ตูน (Positive Objects) และไม่ใช่ใบหน้าการ์ตูน (Negative Objects) โดยทำการฝึกสอนและทดสอบระบบแบบการแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของโมเดล (Fold Cross Validation) โดยการทดลองในการตรวจจับบริเวณที่เป็นใบหน้าเลือกใช้ 3 Fold Cross Validation หลังจากที่ได้ทำการทดลองทำให้ทราบว่าชุดในการฝึกสอนระบบในงานวิจัยนี้อาจจะยังไม่เพียงพอต่อการทดลอง จึงทำให้ผลการทดลองออกมาอยู่ในระดับที่ยังไม่สามารถนำไปทำงานวิจัยต่อได้

เพื่อให้งานวิจัยสามารถดำเนินการต่อได้ ผู้วิจัยจึงได้เลือกภาพใบหน้าการ์ตูนแบบดัดแปลงเล็กน้อย (ระดับ 3) ซึ่งพบมากในภาพยนตร์แอนิเมชันทั่วไปมาใช้ในการจำแนกอารมณ์บนใบหน้าการ์ตูน ส่วนการ์ตูนแบบเหมือนจริง (ระดับ 4) ในงานวิจัยนี้จะไม่นำมาใช้ เนื่องจากภาพใบหน้าการ์ตูนแบบเหมือนจริงมีความเหมือนกันกับมนุษย์ หากจะทำการจำแนกอารมณ์น่าจะไม่มี ความแตกต่างกันกับการจำแนกอารมณ์จากภาพใบหน้าของมนุษย์

5.1.2 สรุปงานวิจัยด้านการจำแนกอารมณ์บนใบหน้าของการ์ตูน

ในส่วนของการจำแนกอารมณ์จากใบหน้าที่การ์ตูน สำหรับงานวิจัยนี้ ผู้วิจัยเลือกอารมณ์บนใบหน้าที่การ์ตูนสามอารมณ์หลัก ได้แก่ อารมณ์โกรธ (Anger) อารมณ์มีความสุข (Happiness) และอารมณ์เศร้า (Sadness) เนื่องจากการแสดงออกทางสีหน้าของทั้งสามอารมณ์พื้นฐานนี้ เป็นอารมณ์สากลสามารถพบเห็นได้ทั่วไป ในการจำแนกอารมณ์บนใบหน้าที่การ์ตูน ผู้วิจัยเลือกเทคนิคการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient: HOG) มาใช้ในการสกัดหาคุณลักษณะเฉพาะของภาพ และเลือกใช้ตัวเปรียบเทียบประสิทธิภาพของการจำแนกสามประเภท ได้แก่ ต้นไม้การตัดสินใจ (Decision Tree: DT) ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors: KNN) โดยกำหนดค่า $k = 3$ เนื่องจากให้ผลในการจำแนกที่ดีที่สุด และซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine: SVM) โดยที่แบ่งข้อมูลในการจำแนกออกเป็นสี่วิธี คือ 1) การจำแนกคุณลักษณะทั้งหน้า 2) การจำแนกคุณลักษณะแยกส่วนคิ้ว ตา ปาก 3) การจำแนกคุณลักษณะแยกส่วนคิ้วและปาก และ 4) การจำแนกคุณลักษณะแยกส่วนตาและปาก

จากผลการทดลองสรุปได้ว่า การจำแนกคุณลักษณะทั้งหน้า ด้วยตัวจำแนกประเภทแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 59.0 ตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 72.3 และตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) ให้ผลการจำแนกคุณลักษณะทั้งหน้าได้ดีที่สุดถึงร้อยละ 88.7

การสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก ด้วยตัวจำแนกประเภทแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 68.3 ตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 83.3 และตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) ให้ผลการจำแนกคุณลักษณะแยกส่วน คิ้ว ตา ปาก ได้ดีที่สุดถึงร้อยละสามารถจำแนกอารมณ์ได้ถึงร้อยละ 93.0

การสกัดคุณลักษณะแยกส่วนคิ้วและปาก ด้วยตัวจำแนกประเภทแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 68.7 ตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 85.3 และตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) ให้ผลการจำแนกคุณลักษณะแยกส่วนคิ้วและปากได้ดีที่สุดถึงร้อยละ 90.3

การสกัดคุณลักษณะแยกส่วนตาและปาก ด้วยตัวจำแนกประเภทแบบต้นไม้ตัดสินใจ (DT) ให้ผลลัพธ์ในการจำแนกร้อยละ 63.7 ตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) $k = 3$ ให้ผลลัพธ์ในการจำแนกร้อยละ 72.0 และตัวจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) ให้ผลการจำแนกคุณลักษณะแยกส่วนตาและปากได้ดีถึงร้อยละ 80.0 ดังตารางที่ 5.1

ตารางที่ 5.1

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k-NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะทั้งสี่ขั้นตอน

ตัวจำแนก	ผลการทดลอง การจำแนก คุณลักษณะ ทั้งหน้า	ผลการทดลอง การจำแนก คุณลักษณะแยกส่วน คิ้ว ตา ปาก	ผลการทดลอง การจำแนก คุณลักษณะแยกส่วน คิ้วและปาก	ผลการทดลอง การจำแนก คุณลักษณะแยก ส่วน ตาและปาก
DT	59.0%	68.3%	68.7%	63.7%
k-NN (k=3)	72.3%	83.3%	85.3%	72.0%
SVM	88.7%	93.0%	90.3%	80.0%

จากตารางที่ 5.1 ทำให้ทราบว่าผลการจำแนกของทั้งสี่วิธี ตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) สามารถจำแนกอารมณ์บนใบหน้าการ์ตูนได้ดีที่สุด และหลังจากที่ได้ทำการศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้องกับการจำแนกอารมณ์บนใบหน้าพบว่า การจำแนกอารมณ์นั้นจำเป็นต้องสกัดคุณลักษณะแยกส่วน คิ้ว ตา ปาก มาช่วยในการจำแนกอารมณ์เพื่อให้ได้ผลของความถูกต้องดีที่สุดในที่นี้ คิ้ว ตา และ ปาก เป็นองค์ประกอบหลักที่สำคัญสำหรับการจำแนกอารมณ์จริง และหลังจากที่ได้ทำการทดลองการจำแนกคุณลักษณะตาและปาก ทำให้ทราบว่า คิ้ว เป็นองค์ประกอบที่สำคัญสำหรับการแสดงอารมณ์บนใบหน้าสำหรับการจำแนกอารมณ์บนใบหน้าของการ์ตูน เพราะการทดลองการจำแนกคุณลักษณะคิ้วและปากยังคงให้ผลในการจำแนกอารมณ์ด้วยซัพพอร์ตเวกเตอร์แมชชีนได้ดี แสดงว่าองค์ประกอบหลักบนใบหน้านั้นหากทำการตัดส่วนของ คิ้ว ออกจากการฝึกสอนระบบ ทำให้มีผลต่อการจำแนกอารมณ์ และเมื่อทำการทดลองครบทุกวิธีทำให้ทราบว่าคุณลักษณะเฉพาะที่ได้จากเทคนิคการหาค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) น่าจะใช้ในการจำแนกอารมณ์บนใบหน้าการ์ตูนได้ดี

ผู้วิจัยเลือกใช้ค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) มาใช้ในการสกัดหาคุณลักษณะเฉพาะ เนื่องจากการแสดงออกทางอารมณ์บนใบหน้ามีความซับซ้อน ค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) น่าจะสามารถนำมาใช้กับงานทางด้านความรู้จำอารมณ์ได้ดี ค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) เก็บทิศทางตามค่าเกรเดียนท์อย่างละเอียดและใช้เวลาใน

การคำนวณที่น้อย อีกทั้งเมื่อนำค่าคุณลักษณะเฉพาะไปทำการจำแนกอารมณ์ยังให้ค่าความถูกต้องในการจำแนกได้ดีเมื่อใช้ร่วมกับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM)

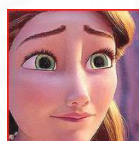
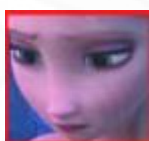
5.1.3 สรุปงานวิจัยด้านการประมวลผลภาพ

การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์สามารถนำไปทำดัชนีภาพในการจำแนกอารมณ์ของตัวการ์ตูนในภาพยนตร์แทนมนุษย์ได้ เนื่องจากมีความถูกต้องในการจำแนกถึง 93% เมื่อใช้ตัวจำแนกแบบ SVM

5.1.4 สรุปแบบสอบถามชุด Ground Truth

จากแบบสอบถามที่ให้มนุษย์เข้ามาตอบโดยที่วัดจากค่าความเห็นที่ตรงกันของมนุษย์ จาก 7 ใน 10 คน โดยผู้วิจัยจะสรุปออกได้เป็น 2 กรณี ดังนี้

1. มนุษย์ 10 คนเห็นตรงกันว่าเป็นอารมณ์นั้น ๆ แต่ระบบไม่สามารถจำแนกได้ ตรงกับการจำแนกของมนุษย์ ยกตัวอย่างเช่น ภาพที่ 5.1 (ก) และ (ข) โดยผลการจำแนกมนุษย์ทั้ง 10 คนจำแนกว่าเป็นอารมณ์เศร้าแต่ระบบทำนายผิด ภาพที่ 5.1 (ก) จากเศร้าเป็นโกรธ และ ภาพที่ 5.1 (ข) จากอารมณ์เศร้าเป็นมีความสุข

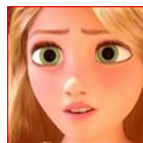
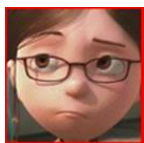


(ก) อารมณ์เศร้าเป็นโกรธ (ข) อารมณ์เศร้าเป็นมีความสุข

ภาพที่ 5.1 ผลการจำแนกอารมณ์ที่ผิดพลาดของระบบ

จากภาพที่ 5.1 สามารถสรุปได้ว่าเนื่องจากภาพใบหน้าการ์ตูนที่นำมาใช้ในการฝึกสอนและทดสอบระบบ หลังจากที่ได้ทำการคำนวณค่าคุณลักษณะเฉพาะด้วยค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) ค่าคุณลักษณะเฉพาะที่ได้ อาจจะมีคุณลักษณะที่ใกล้เคียงใกล้เคียงสำหรับบางภาพจึงทำให้ผลในการจำแนกอาจมีความผิดพลาดไปบ้าง เพราะค่าความเข้มและสว่างของภาพการ์ตูนอาจจะมีโทนสีที่ใกล้เคียงกัน

2. มีเพียงมนุษย์มนุษย์ 7 หรือ 8 จาก 10 คน เห็นตรงกันว่าเป็นอารมณ์ชนิดใด ซึ่งจากผลการฝึกสอนและทดสอบ ระบบสามารถจำแนกได้อารมณ์บนใบหน้าการ์ตูนได้จากกรณีดังกล่าวได้ดี ได้สูงสุดถึง 100% ยกตัวอย่างภาพที่มนุษย์จำแนกไม่ตรงกันแต่ระบบสามารถจำแนกได้ถูกต้องตามชุดฝึกสอนระบบ ดังภาพที่ 5.2 (ก) มนุษย์มองว่าเป็นอารมณ์เศร้าจำนวน 2 ใน 10 คน และภาพที่ 5.2 (ข) มนุษย์ไม่สามารถจำแนกอารมณ์ได้จำนวน 1 ใน 10 คน



(ก) อารมณ์โกรธเป็นเศร้า (ก) อารมณ์เศร้าเป็นไม่สามารถจำแนกได้

ภาพที่ 5.2 ผลการจำแนกอารมณ์ที่ไม่ตรงกันของมนุษย์

จากสองกรณีที่กำลังมาในข้างต้นทำให้ผู้วิจัยเกิดแนวความคิดที่ว่า หากนำภาพที่มนุษย์เห็นตรงกัน 10 คนมาใช้ในการทดลอง จะให้ผลในการจำแนกอารมณ์ของระบบเป็นอย่างไร ซึ่งอ้างอิงจากข้อมูลแบบสอบถามที่มนุษย์ตอบตรงกัน 10 คนแยกตามอารมณ์ดังนี้ โกรธ (Anger) ทั้งหมด 47 ภาพ มีความสุข (Happiness) ทั้งหมด 58 ภาพ เศร้า (Sadness) ทั้งหมด 38 ภาพ จากนั้นข้อมูลมาทำฝึกสอนและทดสอบระบบด้วย 5 - fold cross validation โดยใช้ตัวจำแนกสามชนิด ได้แก่ ต้นไม้การตัดสินใจ (Decision Tree: DT) ตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k - Nearest Neighbors: KNN) กำหนดค่า $k = 3$ และซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine: SVM) ซึ่งได้ผลในการจำแนกอารมณ์แยกตามตัวจำแนกดังนี้ ดังแสดงในตารางที่ 5.2, 5.3, 5.4 และ 5.5 ตามลำดับ

ตารางที่ 5.2

เปรียบเทียบผลการจำแนกด้วยต้นไม้ตัดสินใจ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นที่ตรงกัน 100%

ประเภทที่ระบบทำนาย

ประเภท	Anger	Happiness	Sadness	%
Anger	30	11	6	63.8%
Happiness	9	39	10	67.2%
Sadness	5	8	25	65.8%
ค่าเฉลี่ยความถูกต้อง (%)				65.6%

ประเภทที่ถูกต้อง

ตารางที่ 5.3

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด $k = 3$ แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นที่ตรงกัน 100%

ประเภทที่ระบบทำนาย

ประเภทที่ถูกต้อง	ประเภท	Anger	Happiness	Sadness	%
	Anger	38	6	3	80.9%
	Happiness	3	55	0	94.8%
	Sadness	4	4	30	78.9%
ค่าเฉลี่ยความถูกต้อง (%)					84.9%

ตารางที่ 5.4

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นที่ตรงกัน 100%

ประเภทที่ระบบทำนาย

ประเภทที่ถูกต้อง	ประเภท	Anger	Happiness	Sadness	%
	Anger	46	1	0	97.9%
	Happiness	3	55	0	94.8%
	Sadness	2	2	34	89.5%
ค่าเฉลี่ยความถูกต้อง (%)					94.1%

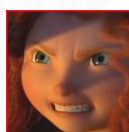
ตารางที่ 5.5

เปรียบเทียบผลการจำแนกด้วยตัวจำแนกแบบต้นไม้ตัดสินใจ (DT), ตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด (k -NN) และตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) แบบสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นที่ตรงกัน 100%

ตัวจำแนก	ผลการทดลอง
DT	65.6%

ตัวจำแนก	ผลการทดลอง
k-NN (k=3)	84.9%
SVM	94.1%

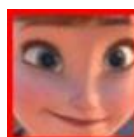
ซึ่งจากผลการทดลองพบว่า ภาพที่มนุษย์เห็นตรงกันทั้งหมดให้ค่าความถูกต้องเพิ่มมากกว่าเดิมเพียงเล็กน้อยสำหรับทุกตัวจำแนก ซึ่งจากขั้นตอนในการทดลองนี้ทำให้ทราบว่าภาพที่นำมาใช้มีผลต่อการจำแนกอารมณ์บนใบหน้าการ์ตูนซึ่งอาจจะเป็นผลมาจากค่าความสว่างภายในภาพมีค่าที่ใกล้เคียงกัน จึงทำให้ผลการจำแนกอาจเกิดความผิดพลาดในการจำแนกอารมณ์ได้ เนื่องจากในการทดลองนี้ซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องในการจำแนกมากที่สุด ผู้วิจัยจึงขอแสดงผลการจำแนกที่ผิดพลาดของซัพพอร์ตเวกเตอร์แมชชีน (SVM) แยกตามอารมณ์ ดังภาพที่ 5.3, 5.4 และ 5.5 ตามลำดับ



โกรธเป็นมีความสุข

ภาพที่ 5.3 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์โกรธ จากโกรธเป็นมีความสุขของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นตรงกัน 100%

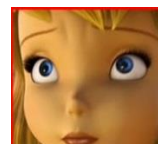
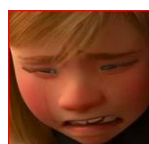
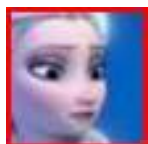
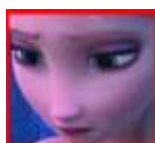
จากภาพ 5.3 อาจจะทำให้เกิดความสับสนบริเวณที่เป็นปากของตัวการ์ตูนเนื่องจากปากขยายออกและโค้งขึ้นเป็นรูปพระจันทร์ครึ่งเสี้ยว จึงทำให้ได้ค่าคุณลักษณะเฉพาะตามทิศทางเกรเดียนท์ที่ใกล้เคียงกับอารมณ์มีความสุข จึงทำให้ส่งผลการจำแนกอารมณ์บนใบหน้าการ์ตูน



มีความสุขเป็นโกรธ

ภาพที่ 5.4 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์มีความสุข จากมีความสุขเป็นโกรธของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นตรงกัน 100%

จากภาพ 5.4 อาจเกิดความสับสนบริเวณที่เป็นคิ้วของตัวการ์ตูน เนื่องจากหัวคิ้วตกลงทางคิ้วยกสูงขึ้น จึงทำให้ได้ค่าคุณลักษณะเฉพาะตามทิศทางเกรเดียนท์ที่ใกล้เคียงกับอารมณ์โกรธ จึงส่งผลต่อการจำแนกอารมณ์บนใบหน้าการ์ตูน



(ก) เศร้าเป็นโกรธ

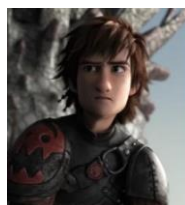
(ข) เศร้าเป็นมีความสุข

ภาพที่ 5.5 แสดงการทำนายอารมณ์ที่ผิดพลาดของ SVM ในอารมณ์เศร้า (ก) อารมณ์เศร้าเป็นโกรธ และ (ข) อารมณ์เศร้าเป็นมีความสุข ของการสกัดคุณลักษณะแยกส่วน คิ้ว ตา และ ปาก จากภาพที่มีค่าความเห็นตรงกัน 100%

จากภาพที่ 5.5 (ก) อาจเกิดความสับสนบริเวณคิ้วของตัวการ์ตูน เนื่องจากหัวคิ้วตกลงทางคิ้วยกสูงขึ้น จึงทำให้ได้ค่าคุณลักษณะเฉพาะตามทิศทางเกรเดียนท์ที่ใกล้เคียงกับอารมณ์โกรธ และ ภาพที่ 5.5 (ข) บริเวณคิ้วของตัวการ์ตูนมีความก้ำกึ่งกันระหว่างการยกหัวคิ้วให้สูงขึ้นเพียงเล็กน้อยของอารมณ์มีความสุข ซึ่งอาจตรงกับค่าคุณลักษณะเฉพาะตามทิศทางเกรเดียนท์ จึงทำให้ผลการจำแนกอารมณ์บนใบหน้าการ์ตูนเกิดความผิดพลาดได้

5.2 ปัญหาที่พบ

ปัญหาที่พบจากการทดลอง คือ ภาพการ์ตูนมีข้อจำกัดมากสำหรับการตรวจจับบริเวณใบหน้าโดยระบบอัตโนมัติ เนื่องจากการ์ตูนมักมีเส้นผมหรือสิ่งปกคลุมบนใบหน้า ดังภาพที่ 5.6



ภาพที่ 5.6 ตัวอย่างภาพการ์ตูนที่มีเส้นผมปกคลุมบริเวณใบหน้า

จึงส่งผลต่อการนำภาพมาใช้ และการจัดเตรียมชุดข้อมูลใบหน้าการ์ตูนแยกตามอารมณ์นั้นสามารถทำได้ยาก เนื่องจากการ์ตูนส่วนใหญ่ใบหน้าที่ใช้ในการแสดงอารมณ์ โกรธและอารมณ์เศร้าจะมีน้อย เมื่อเปรียบเทียบกับอารมณ์มีความสุข

5.3 ข้อเสนอแนะ

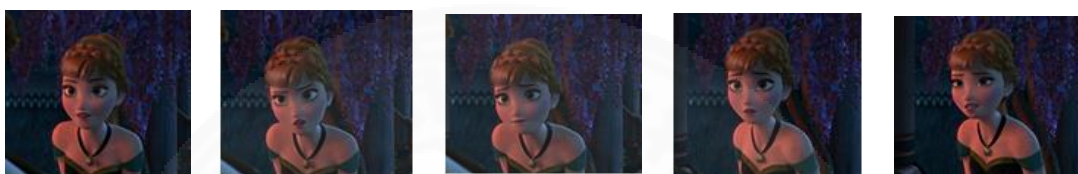
การจำแนกอารมณ์พื้นฐานสำหรับงานวิจัยนี้เป็นเพียงการทดลองเบื้องต้นสำหรับการจำแนกอารมณ์บนใบหน้าการ์ตูนยังมีข้อจำกัดในการตรวจจับใบหน้าของเทคนิคฮาร์โลคฟีเจอร์แบบมาตรฐาน ซึ่งสามารถตรวจจับบริเวณใบหน้าของการ์ตูนได้ดีเฉพาะการ์ตูนแบบดัดแปลงเล็กน้อย (ระดับ 3) และการ์ตูนแบบเหมือนจริง (ระดับ 4) โดยให้ค่าความถูกต้องถึงร้อยละ 77.8 และ 83.8 ตามลำดับ แต่สำหรับการ์ตูนแบบดัดแปลงอย่างมาก (ระดับ 1) และการ์ตูนแบบแปลงและประดิษฐ์ (ระดับ 2) ผลการตรวจจับใบหน้าการ์ตูนสามารถทำได้น้อย การทดลองก่อนการปรับปรุงชุดฝึกสอนระบบ (ใช้เทคนิคของฮาร์โลคฟีเจอร์แบบมาตรฐาน) สามารถตรวจจับได้ร้อยละ 8.1 และ 33.3 ตามลำดับ และหลังการปรับปรุงชุดในการฝึกสอนระบบสามารถตรวจจับได้ดีขึ้นเพียง 59.6 และ 56.6 ตามลำดับ

ก่อนที่จะทำการนำข้อมูลภาพไปสกัดหาคุณลักษณะเฉพาะบนใบหน้าของการ์ตูนด้วยค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) เพื่อให้เกิดประโยชน์สูงสุดในการจำแนกอารมณ์นั้น ต้องขึ้นอยู่กับว่าการแสดงออกทางสีหน้าของการ์ตูนในแต่ละอารมณ์ใช้อวัยวะใดเป็นหลัก งานวิจัยนี้เลือกใช้อารมณ์พื้นฐานสามอารมณ์ ได้แก่ อารมณ์โกรธ (Anger) อารมณ์มีความสุข (Happiness) และอารมณ์เศร้า (Sadness) โดยองค์ประกอบของอวัยวะหลักบนใบหน้าที่แสดงให้เห็นว่าการ์ตูนที่มองเห็นนั้นแสดงอารมณ์อย่างไร อารมณ์มีความสุขและอารมณ์เศร้า ประกอบด้วยอวัยวะหลักในการแสดงอารมณ์ได้แก่ คิ้ว ตา และปาก ส่วนอารมณ์โกรธ อวัยวะหลักที่ใช้ในการแสดงอารมณ์ได้แก่ คิ้ว แก้ม ปาก เป็นต้น

นอกจากนี้ยังอาจปรับปรุงอัลกอริธึมสำหรับการสกัดหาคุณลักษณะเฉพาะ โดยอาจจะใช้เทคนิคในการหาคุณลักษณะเฉพาะที่แตกต่างไปจากค่าฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) เช่น Eigen vector, Eigen Value, Fisher's Linear Discriminant, Singular Value, Decomposition (SVD) และ Action units เป็นต้น ซึ่งวิธีการเหล่านี้อาจจะช่วยเพิ่มประสิทธิภาพการจำแนกอารมณ์บนใบหน้าการ์ตูนได้ดีขึ้น ทั้งนี้เพื่อให้เกิดประโยชน์สูงสุดเกี่ยวกับงานทางด้านทัศนศาสตร์คอมพิวเตอร์ ยังสามารถนำไปประยุกต์ใช้ทำดัชนีเพื่อค้นคืนภาพหรือวิดีโอ สำหรับการจำแนกอารมณ์จากการแสดงออกบนใบหน้าของการ์ตูนแทนมนุษย์ได้ โดยอาจจะเพิ่มอารมณ์บนใบหน้าตัวการ์ตูนให้มากขึ้น เช่น อารมณ์ตื่นเต้น อารมณ์ตกใจ อารมณ์มึนงง เป็นต้น ซึ่งอาจจะใช้

องค์ประกอบหลัก อื่น ๆ บนที่นอกเหนือไปจาก คิ้ว ตา และปาก อาจจะมีส่วนของท่าทางจาก มือรวมเข้ามาด้วย

เพิ่มการจำแนกอารมณ์จากชุดข้อมูลภาพที่ต่อเนื่องกัน หรือ อาจจะมาจกเฟรมที่ใกล้เคียงกัน ดังภาพที่ 5.7 เพื่อเพิ่มประสิทธิภาพในการจำแนกอารมณ์บนใบหน้าของการ์ตูน เช่น อารมณ์โกรธ และ อารมณ์เศร้า การแสดงสีหน้าของทั้งสองอารมณ์มีความใกล้เคียงกัน ทำให้ไม่ว่าจะเป็นมนุษย์หรือระบบก็สามารถสับสนและจำแนกอารมณ์ผิดพลาดได้



เฟรมที่ (1)

เฟรมที่ (2)

เฟรมที่ (3)

เฟรมที่ (4)

เฟรมที่ (5)

ภาพที่ 5.7 ชุดข้อมูลภาพที่ต่อเนื่องกันของการ์ตูน

จากภาพที่ 5.7 ถ้าหากมองครบทุกเฟรมของภาพจะพบว่าการ์ตูนตัวนี้กำลังแสดงอารมณ์โกรธอยู่ แต่เนื่องจากบางเฟรมอาจจะมีอาการกำกวมของอารมณ์ เช่น เฟรมที่ (3) และ (4) หากดูเพียงหนึ่งภาพอาจจะเข้าใจได้ว่าการ์ตูนตัวนั้นกำลังแสดงอารมณ์เศร้า เป็นต้น

รายการอ้างอิง

หนังสือ

- Christopher D. M., Prabhakar R., & Schütze, H. (2008) *Introduction to Information Retrieval*. Cambridge University Press.
- Gonzalez, R. C. & Woods, E. R., (2009). *Digital Image Processing*. Prentice Hall Upper Saddle River : New Jersey 07458.
- Mehrabian A. (1968). *Communication without words*. Psychology Today 2. p.53–56.
- ทรงพล ภูมิพัฒน์. (2538). *จิตวิทยาทั่วไป*. น.120 – น.121.
- ปิยะบุตร สุทธิธิดารา. (2558). *Manga Bible 1 ครบทุกพื้นฐาน การหัตถ์วาดการ์ตูน*. (พิมพ์ครั้งที่ 1). วิไลวรรณ ศรีสงคราม, สุชัญญา รัตนสัญญา, โรจน์รวี พจน์พัฒนพล และ พีรพล เทพประสิทธิ์. (2549). *จิตวิทยาทั่วไป: General Psychology*. (พิมพ์ครั้งที่ 1). น.189 – น.191.
- สมเกียรติ อุดมธรรษากุล. (2554). *การประมวลผลภาพดิจิทัล: Fundamentals of Digital Image Processing*.
- สุรพงศ์ เอื้อวัฒนามงคล. (2559). *การทำเหมืองข้อมูล*. สถาบันบัณฑิตพัฒนบริหารศาสตร์ : สำนักพิมพ์สถาบันบัณฑิตพัฒนบริหารศาสตร์.

บทความวารสาร

- Arumugam, D., & Purushothaman, B. E. (2011). Emotion Classification Using Facial Expression. *International Journal of Advanced Computer Science and Applications*. Vol. 2. No. 7.
- Chau, H. (2015). FACE DETECTION AND RECOGNITION OF DRAWN CHARACTERS.
- Chen, Q., Georganas, N., & Petriu, E. (2007). Real-time vision-based hand gesture recognition using haar-like feature. *IEEE International Conference on Instrumentation and Measurement Technology*.
- Chibelushi, C., & Bourel, F. (2002). Facial Expression Recognition, *A Brief Tutorial Overview*.

- Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. *NRIA Rhone-Alps*. 655 avenue de l'Europe, Montbonnot 38334. France.
- KalaiSelvi, R., Kavitha, P., & Shunmuganathan, K. L. (2014). Capturing Facial Actions in Video to Revive Expressions of Humans. *International Conference on Innovations in Engineering and Technology*.
- Said, Y., Atri, M., & Tourki, R. (2011). Human Detection Based on Integral Histograms of Oriented Gradients and SVM. *International Conference Communications. Computing and Control Applications (CCCA)*.
- Takayama, K., Johan, H., & Nishita, T. (2012). FACE DETECTION AND FACE RECOGNITION OF CARTOON CHARACTERS USING FEATURE EXTRACTION. *Proceedings of the IEEE Image Electronics and Visual Computing Workshop Kuching*. Malaysia. November 21-24.
- Tian, Y. L., Kanade, T., & Cohn, F. J. (2001). Recognizing Action Units for Facial Expression Analysis. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*. VOL.23. NO.2. FEBRUARY.
- Viola, P., & Jones, M. J. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. *ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*.
- ชุมพล เสนาพันธ์ และ คำรณ สุนันต์. (2556). การตรวจจับมนุษย์โดยใช้คุณลักษณะตามการแปลงค่า การเก็บรวบรวมข้อมูลร่วมกับค่าความถี่ของค่าทิศทางตามค่าเกรเดียนท์และตัวแบบการเรียนรู้เครื่องจักรเรียนรู้เอ็กซ์ทรีมสำหรับเรเดียลเบสิสทั่วไปรุ่นปรับปรุง. *การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 9*.
- ประภาพร กุลลิมรัตน์ชัย. (2556). ความรู้พื้นฐานเกี่ยวกับการค้นคืนภาพ Image Retrieval Fundamentals. *วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเชีย*. ปีที่ 7. ฉบับที่ 2. ประจำเดือน กรกฎาคม - ธันวาคม.
- รุสลี่ สุทธิวีร์กุล และ วิไลพร แซ่ลี. (2554). การตรวจจับใบหน้าด้วยวิธีการพื้นฐานของการจำลองรูปแบบ Haar-like. *วารสารวิศวกรรมศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒปีที่ 6*. ฉบับที่ 2. เดือนกรกฎาคม - ธันวาคม.

อนุสรณ์ อุ๋นท้าว และ สมปอง เวฬุวนาธร. (2557). การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับการวิเคราะห์องค์ประกอบหลักและซัพพอร์ตเวกเตอร์แมชชีน. *การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 10.*

วิทยานิพนธ์

จิรา แก้วสุวรรณ. (2549). *การตรวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน.* สาขาวิชาเทคโนโลยีคอมพิวเตอร์. สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

สื่ออิเล็กทรอนิกส์

Dobrin, A. (2013). Facial Expressions: Universal vs. Cultural. *Psychology Today*. Retrieved Oct 10, 2015, from <https://www.psychologytoday.com/blog/am-i-right/201306/facial-expressions-universal-vs-cultural>

Frischholz, R. (2016). The Face Detection, Retrieved Oct 10, 2015, <https://facedetection.com>

Jones, D. (2007). Tops for Drawing Facial Expressions. Retrieved Oct 10, 2015, from http://danidraws.com/media/danidraws_facialexpressions.pdf

Lewis, M. B., & Ellis, H. D. (2003). How we detect a face. *A survey of psychological evidence*. DOI: 10.1002/ima.10040

นิรุทธิ์ อุทาทิพย์. (2553). ความหมายคำว่าการ์ตูน และแอนิเมชัน. สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <https://namofree.wordpress.com/about/>

พริ้มเพรา ดิษยวณิช. (2548). อารมณ์ (Emotion). สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <http://www.chamlongclinic-psych.com/document/emotion/>

พิชชยานาฏ ริรักษ์. (2553). ระบบสี (Color Model). สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <https://sites.google.com/site/wbicomputergraphics/rabb-si-color-model>

ศิริพร ชัยมานะ. (2555). การวาดใบหน้าขั้นพื้นฐาน. สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก http://siripornchaimana.blogspot.com/p/blog-page_7639.html

เศรษฐพงศ์ มะลิสวรรณ. (2552). Image processing เทคโนโลยีการประมวลผลภาพ. สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <http://jaratcyberu.blogspot.com/2009/10/image-processing.html>

สุตามาศ พลดี. (2553). ประเภทการ์ตูน. สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <http://pirun.ku.ac.th/~b5310300635/tpye.html>

สุพัฒน์. (2556). ความหมายของการ์ตูน. สืบค้นเมื่อวันที่ 10 ตุลาคม 2559, จาก <http://poosampoo002.blogspot.com/2013/02/blog-post.html>



ภาคผนวก



ภาคผนวก ก
การจำแนกอารมณ์จากความรู้สึกภายใน
(Classification of emotion from feeling inside)

การศึกษาเกี่ยวกับการจำแนกประเภทของอารมณ์มีความหลายหลาย ซึ่งนักจิตวิทยาแต่ละบุคคลมีการจำแนกอารมณ์พื้นฐานของมนุษย์ที่แตกต่างกันออกไป นักจิตวิทยาแต่ละท่าน พยายามที่จะตั้งข้อสันนิษฐานต่าง ๆ ขึ้นมาเพื่อ อธิบายเรื่องของอารมณ์ว่าเกิดขึ้นได้อย่างไร เกี่ยวข้องกับสิ่งใดบ้าง ซึ่งผู้วิจัยจะอธิบายรายละเอียดในแต่ละทฤษฎีทางอารมณ์ดังนี้ (วิไลวรรณ ศรีสงคราม, สุขัญญา รัตันสัญญา, โรจน์รวี พจน์พัฒนาผล และ พีรพล เทพประสิทธิ์, 2549)

เจมส์-แลงก์ (James – Lange Theory) กล่าวว่า อารมณ์ของมนุษย์เกิดขึ้นมาจากสิ่งเร้ารอบๆ ภายของมนุษย์ เมื่อได้รับสิ่งเร้าแล้ว ร่างกายจะเกิดสภาวะการเปลี่ยนแปลง ทำให้เกิดพลังงานในระบบเส้นประสาททางสมองของมนุษย์และส่งผลทำให้เกิดสภาวะทางอารมณ์ขึ้นมา วิลเลียม เจมส์ (William James) ได้เสนอแนวความคิดว่าปัจจัยสำคัญที่ทำให้บุคคลมีความรู้สึกว่ามีอารมณ์เกิดขึ้นก็คือการรู้สึกในการเปลี่ยนแปลงต่าง ๆ ทางสรีระที่เกิดขึ้นในขณะที่มีการตอบสนองต่อสถานการณ์น่ากลัวหรือตกใจ ซึ่งวิลเลียมได้อธิบายต่อไปอีกว่า ให้เปรียบกับการนำรถเทียมม้ามาเตรียมไว้ก่อนม้า กล่าวคือ มนุษย์เราจะเกิดความกลัวก็เพราะการวิ่ง มนุษย์เราจะเกิดความโกรธเพราะเกิดความปั่นป่วนภายในสรีระ นอกจากนี้ เจมส์ยังมีความเห็นที่สอดคล้องกันกับ วิลเลียม จึงได้จัดทำทฤษฎีร่วมกันขึ้น

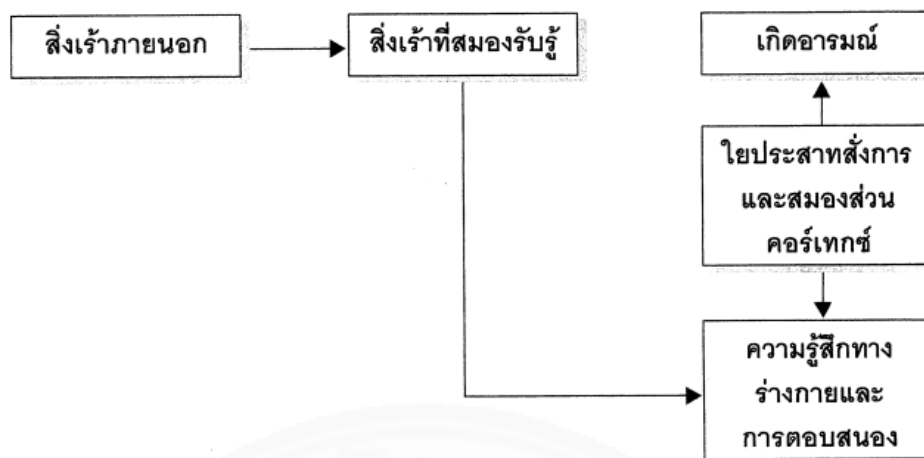
ทฤษฎีนี้จะอธิบายให้เห็นว่า มนุษย์จะรับรู้สิ่งเร้าเป็นอันดับที่หนึ่ง หลังจากนั้นจะเกิดการเปลี่ยนแปลงทางร่างกายและความรู้สึก และในทฤษฎีนี้ก็เสนอแนวความคิดอีกว่าอารมณ์จะเกิดขึ้นภายหลังจากมีปฏิกิริยาหรือการเปลี่ยนแปลงของร่างกาย ฉะนั้นกระบวนการเกิดจะแสดงลำดับขั้นตอนดังภาพที่ 1 ตามลำดับดังนี้



ภาพที่ ก.1 ตัวอย่างของการเกิดอารมณ์จากภายในหลังจากมีการเปลี่ยนแปลงทางสรีระทางร่างกาย จาก จิตวิทยาทั่วไป: *General Psychology* (พิมพ์ครั้งที่ 1), โดย วิไลวรรณ ศรีสงคราม, สุขัญญา รัตนสัญญา, โรจน์รวี พจน์พัฒนาพล, พีรพล เทพประสิทธิ์, 2549, น.189 - น.191.

ยกตัวอย่างการเกิดอารมณ์จากภาพที่ ก.1 เช่น เมื่อมนุษย์เดินสะดุดขึ้นบันได มนุษย์ มักจะใช้มือจับราวบันไดแบบอัตโนมัติเพื่อช่วยในการทรงตัว เหตุการณ์ที่เกิดขึ้นดังกล่าว ทำให้อารมณ์ความรู้สึกภายในร่างกายของมนุษย์เกิดขึ้นมาด้วยตามลำดับดังนี้ เกิดความรู้สึกตกใจ หัวใจเต้นแรง หายใจหอบ มีความรู้สึกสั่นหรือแขนขาอ่อนแรง เป็นต้น

แคนนอน-บาร์ด (Cannon-Bard Theory) มีความเชื่อที่ตรงกันข้ามกับเจมส์-แลงก์ โดยแคนนอนเชื่อว่า สมองส่วนไฮโปทาลามัส (Hypothalamus) จะควบคุมอารมณ์ของสิ่งมีชีวิต โดยเมื่อมีสิ่งเร้ามากระตุ้น จะมีการส่งผลลัพธ์นี้ผ่านกระแสประสาทไปยังสมองส่วนนี้ หลังจากนั้นกระแสประสาทจะถูกแยกออกเป็นสองทาง ทางแรก คือ ส่งไปยังอวัยวะตอบโต้เกิดเป็นพฤติกรรมขึ้น อีกทางหนึ่งจะส่งไปยังสมองส่วนคอร์เทกซ์ (Cortex) ส่งผลให้เกิดอารมณ์ขึ้นในเวลาทีใกล้เคียงกัน กระบวนการเกิดจะแสดงลำดับขั้นตอนดังภาพที่ ก.2 ตามลำดับดังนี้



ภาพที่ ก.2 เกิดอารมณ์จากภายในหลังจากมีการเปลี่ยนแปลงทางสรีระทางร่างกายตามทฤษฎีของ แคนนอน - บาร์ต จาก จิตวิทยาทั่วไป: *General Psychology* (พิมพ์ครั้งที่ 1), โดย วิลโลรณ ศรีสงคราม, สุขัญญา รัตนสัญญา, โรจน์รวี พจน์พัฒนาพล, พีรพล เทพประสิทธิ์, 2549, น.189 - น.191.

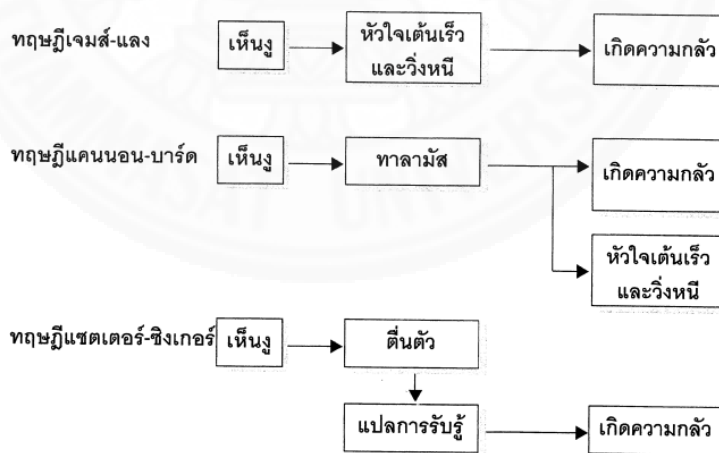
จากภาพที่ 2 จะอธิบายถึงสาระสำคัญของทั้งสองทฤษฎีนี้ยังเป็นที่ถกเถียงกันอยู่ในบางสถานการณ์ ยกตัวอย่างเช่น ในกรณีที่มนุษย์ประสบอันตรายอย่างกะทันหัน การเกิดอารมณ์นั้นจะเกิดขึ้นภายหลังจากการแสดงออกทางพฤติกรรมนี้อย่างอัตโนมัติ และในบางสถานการณ์ อารมณ์ของมนุษย์มักจะเกิดพร้อมกันกับการเปลี่ยนแปลงทางสรีระและเกิดก่อนพฤติกรรมที่เกิดอัตโนมัติ เช่น อารมณ์กลัวเกิดจนลูกแล้วจึงวิ่งหนี เป็นต้น

ส่วนแซตเตอร์ - ซิงเกอร์ (Schachter - Singer Theory) เน้นความสำคัญของการตีความหมายของสิ่งเร้าหรือสถานการณ์ที่พบเห็นว่ามีสาเหตุมาจากสิ่งใด ซึ่งทฤษฎีนี้จัดอยู่ในกลุ่มของพุทธินิยม (Cognitive Theory) ทฤษฎีเสนอว่าสมองไม่มีความจำเป็นต้องมีการตอบสนองทางอารมณ์ด้วยพฤติกรรมที่เฉพาะตายตัว แต่การรับรู้แล้วเข้าใจต่างหากที่ก่อให้เกิดการสร้างแบบแผนปฏิกิริยาทางอารมณ์ให้แตกต่างกันออกไป โดยขึ้นอยู่กับตัวกระตุ้นประสบการณ์ของแต่ละบุคคล ดังนั้นทฤษฎีนี้จึงเน้นว่า อารมณ์ คือ ความรู้สึกที่เกิดจากการประเมินของสิ่งเร้าที่สมองรับรู้ซึ่งเป็นการทำหน้าที่ควบคุมสมองขั้นต้นเพื่อป้องกันรักษาชีวิต เช่น บทเพลงบางเพลงบางคนฟังแล้วร้องไห้ อาจเป็นเพราะเนื้อเพลงมีความคล้ายคลึงกับชีวิตของตนเอง หรือ ขณะที่กำลังเชียร์กีฬาอยู่นั้น เมื่อความสนใจของมนุษย์พุ่งเป้าไปที่เกมส์นั้น ๆ จะพบเห็นว่ามีมนุษย์หายใจถี่เร็ว หัวใจเต้นแรง มีเหงื่อออก สิ่งเหล่านี้เกิดมาจากอิทธิพลของการเข้าใจและรับรู้คุณค่าของการแข่งขันนั้น จึงเรียกทฤษฎีนี้ได้อีกอย่างว่า “ทฤษฎีอารมณ์ตู้เพลง” (Jukebox Theory) เพราะปุ่มเพลงที่กด เปรียบได้กับจังหวะหรือท่วงทำนองหรือแนวทางที่แต่ละบุคคลตีความหมายผสมผสานกับประสบการณ์ในอดีตของแต่ละบุคคลนั่นเอง กระบวนการเกิดจะแสดงลำดับขั้นตอนดังภาพที่ ก.3 ตามลำดับดังนี้



ภาพที่ ก.3 ทฤษฎีทางอารมณ์ของ แซตเตอร์ – ซิงเกอร์ จาก จิตวิทยาทั่วไป: *General Psychology* (พิมพ์ครั้งที่ 1), โดย วิลาวรรณ ศรีสงคราม, สุชัยญา รัตนสัญญา, โรจน์วี พจน์พัฒน์พล, พีรพล เทพประสิทธิ์, 2549, น.189 - น.191.

จากภาพที่ ก.3 แสดงให้เห็นว่า อารมณ์เป็นกระบวนการทางสมอง ซึ่งสมองเป็นศูนย์กลางอวัยวะแหล่งสำคัญที่รับผิดชอบของการเกิดอารมณ์ เป็นฐานให้มีปฏิกิริยาทางร่างกายเพื่อตอบสนองและเพื่อปรับสมดุลในร่างกาย ทั้งนี้ทฤษฎีทางอารมณ์ทั้งสามนี้กล่าวได้ว่าไม่มีทฤษฎีใดที่สามารถให้ความกระจ่างและอธิบายได้แน่ชัดว่าอารมณ์เกิดขึ้นมาได้อย่างไร ถึงกระนั้น ทั้งสามทฤษฎีข้างต้นก็พอที่จะอธิบายข้อสงสัยในเรื่องการเกิดอารมณ์ได้ในระดับหนึ่ง ซึ่งจะเปรียบเทียบการเกิดอารมณ์ของทั้ง 3 ทฤษฎีดังภาพที่ ก.4



ภาพที่ ก.4 แสดงการเกิดของอารมณ์ตามทฤษฎีทั้ง 3 ทฤษฎี จาก จิตวิทยาทั่วไป: *General Psychology* (พิมพ์ครั้งที่ 1), โดย วิลาวรรณ ศรีสงคราม, สุชัยญา รัตนสัญญา, โรจน์วี พจน์พัฒน์พล, พีรพล เทพประสิทธิ์, 2549, น.189 - น.191.

ประเภทของอารมณ์

อารมณ์โดยหลักของมนุษย์ มีอยู่ด้วยกันทั้งหมดสามอย่าง คือ อารมณ์รัก อารมณ์โกรธ และอารมณ์กลัว แต่จริงแล้วมนุษย์มีอารมณ์อยู่ด้วยกันหลายอย่าง โสเครตีส (Socrates) นักปรัชญาชาวกรีก กล่าวว่า อารมณ์ของมนุษย์นั้นมีอยู่ด้วยกัน 6 อย่าง คือ ความปรารถนาอยากได้ ความเกลียดชัง ความแปลกใจ ความร่าเริงยินดี ความโศกเศร้าและความรัก ส่วนเครชและครัชฟิลด์ (Krech and Crutchfield) แบ่งอารมณ์ออกเป็นหกชนิด ตามลักษณะ ดังนี้ (ทรงพล ภูมิพัฒน์, 2538)

1. อารมณ์เดิมหรืออารมณ์หลัก (Primary Emotion) เช่น อารมณ์สนุก (Joy) อารมณ์ร่าเริงมีความสุข (Happiness) อารมณ์ยินดี (Pleasure) อารมณ์กลัว (Fear) อารมณ์โกรธ (Anger) อารมณ์เศร้าหมอง (Sadness)
2. อารมณ์ที่เกี่ยวข้องกับการกระตุ้นประสาทสัมผัส (Emotions Pertaining to sensory stimulation) เช่น อารมณ์เจ็บปวด (Pain) อารมณ์ขยะแขยง (Disgust) อารมณ์หวาดกลัว (Panic)
3. อารมณ์ที่เกี่ยวข้องกับความรู้สึกภายในของตนเอง - การวัดผลตนเอง (Emotions Pertaining to self - appraisal) เช่น อารมณ์ภูมิใจ (Proud) อารมณ์อาย (Shy) อารมณ์สำนึกผิด (Repent)
4. อารมณ์ที่เกี่ยวข้องกับผู้อื่น (Emotions Pertaining to other people) เช่น อารมณ์เกลียด (Hate) อารมณ์สงสาร (Compassion)
5. อารมณ์ที่เกิดจากความรู้สึกจิตใจได้สำนึก (Appreciative Emotions) เช่น ประหลาดใจ (Surprise) ตื่นเต้น (Excited) พิศวง (Amazing) งงงวย (Puzzled)
6. อารมณ์ความรู้สึก (Moods) หรือเรียกว่า อารมณ์พื้นหรืออารมณ์ค้างอยู่เดิม เช่น อารมณ์ซึม (Depressed) อารมณ์เศร้า (Sadness) อารมณ์ปิติใจ (Joy)

ประวัติผู้เขียน

ชื่อ	นางสาวอรวรรณ เลียบศิริ
วันเดือนปีเกิด	3 พฤศจิกายน 2529
วุฒิการศึกษา	ปีการศึกษา 2552: วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ
ตำแหน่ง	เสมียน กวก.พบ. กรมแพทย์ทหารบก
ผลงานวิชาการ	อรวรรณ เลียบศิริ และ เสาวลักษณ์ วรรณภา. (กรกฎาคม 2560). การจำแนกอารมณ์ของตัวการ์ตูนโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนต์: <i>Cartoonized Facial Expression Classification Using Histogram of Oriented Gradient (HOG)</i> . The 13th National Conference on Computing and Information Technology (NCCIT), กรุงเทพฯ.
ประสบการณ์ทำงาน	2554 – ปัจจุบัน: เสมียน กวก.พบ. กรมแพทย์ทหารบก 2552 – 2554: พนักงานเขียนโปรแกรมชั้น 4 โรงพยาบาลพระมงกุฎเกล้า