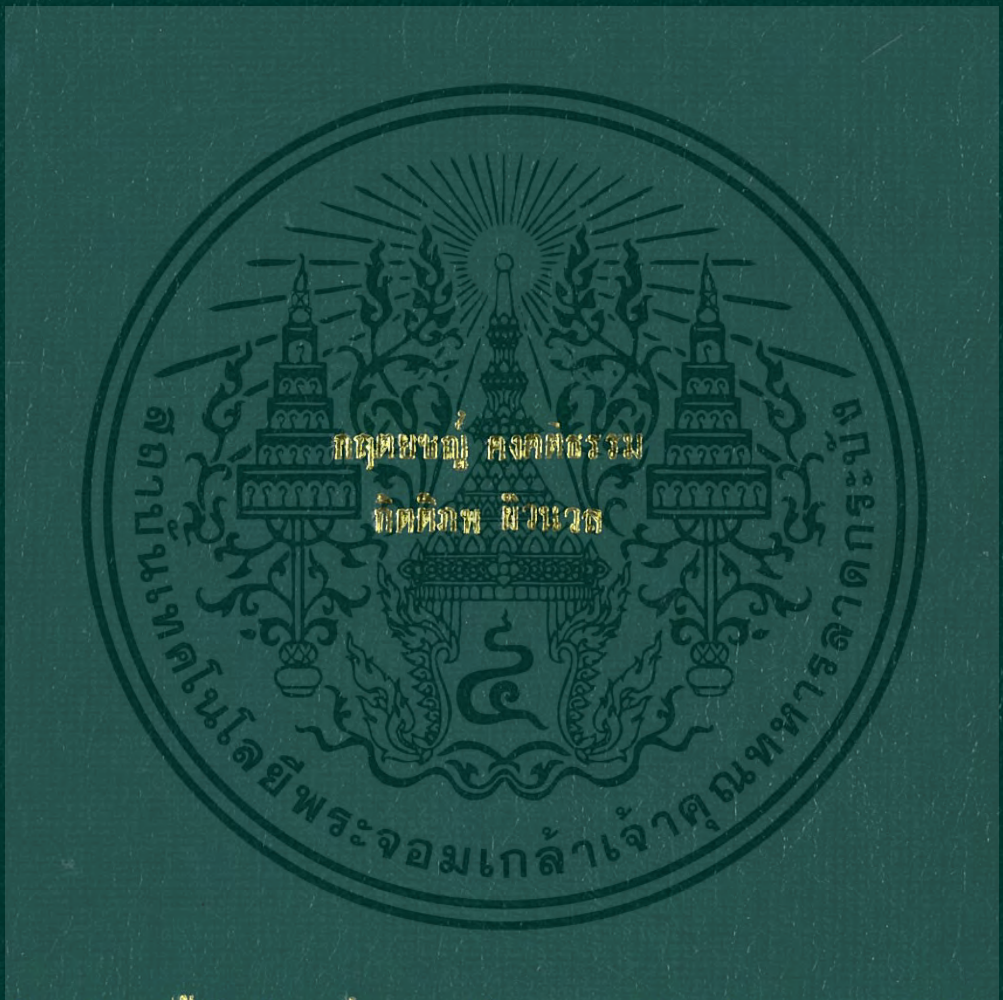
การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำตัวอักษรบนอุปกรณ์พกพา

MODIFIED ADAPTIVE THRESHOLDING FOR PRE-PROCESSING OF TEXT RECOGNITION ON MOBILE DEVICET



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2558

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

MODIFIED ADAPTIVE THRESHOLDING FOR PRE-PROCESSING OF TEXT RECOGNITION ON MOBILE DEVICES

T146185

โดย

กฤตยชญ์ คงคติธรรม

กิตติภพ ผิวนวล

อาจารย์ที่ปรึกษา

ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา

อาจารย์ที่ปรึกษาร่วม

ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาต พสุภา

แขหมู่ 146185
เลขทะเบียน
วัน เดือน ปี 25 เม.ย. 2560

b.12841043
i.

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2558

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

MODIFIED ADAPTIVE THRESHOLDING FOR PRE-PROCESSING OF TEXT RECOGNITION ON MOBILE DEVICES

โดย

กฤตยชญ์ คงคติธรรม

กิตติภพ ผิวนวล

อาจารย์ที่ปรึกษา

ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา

อาจารย์ที่ปรึกษาร่วม

ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาต พสุภา

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2558

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MODIFIED ADAPTIVE THRESHOLDING FOR PRE-PROCESSING OF TEXT RECOGNITION ON MOBLE DEVICES

KITTAYOCH KONGKATITHAM

KITTIPOP PEUWNUAN

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

2/2015

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2016

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองปริญญานิพนธ์ ประจำปีการศึกษา 2558

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

MODIFIED ADAPTIVE THRESHOLDING FOR PRE-PROCESSING OF TEXT RECOGNITION ON MOBLE DEVICES

ผู้จัดทำ

1. นายกฤตยชญ์ คงคติธรรม รหัสนักศึกษา 55070001
2. นายกิตติภพ ผิวนวล รหัสนักศึกษา 55070010

.....อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา)

.....อาจารย์ที่ปรึกษาร่วม

(ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาต พสุภา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อโครงงาน	การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา
นักศึกษา	นายกฤตยชญ์ คงคติธรรม รหัสนักศึกษา 55070001
	นายกิตติภพ ผิวนวล รหัสนักศึกษา 55070010
ปริญญา	วิทยาศาสตรบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2558
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาต พสุภา

บทคัดย่อ

หลังจากที่ประเทศไทยได้ก้าวเข้าสู่ประชาคมเศรษฐกิจอาเซียน การติดต่อสื่อสารกับประเทศเพื่อนบ้านและประเทศคู่ค้าในภูมิภาคใกล้เคียงทำให้เราได้รับข้อมูล ข่าวสารในภาษาต่างประเทศมากขึ้น นอกจากนี้ยังรวมไปถึงสื่อสิ่งพิมพ์ โฆษณา ประกาศ และอื่นๆ ที่ใช้ภาษาต่างประเทศ ซึ่งบางคนต้องการเครื่องมือสำหรับแปลข้อความเหล่านั้นเพื่อการรับรู้ข้อมูลที่ถูกต้อง

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตามการพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และปัญหาจากแสงและเงาซึ่งเป็นอุปสรรคต่อการรู้จำข้อความ โครงงานชุดนี้จึงพัฒนาแอปพลิเคชันโดยศึกษาและพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ แอปพลิเคชันนี้ไม่เพียงแต่ใช้การรู้จำเพื่อนำไปแปลความหมายเท่านั้น แต่ยังออกแบบให้สามารถนำข้อความที่รู้จำได้ไปใช้ต่อในอุปกรณ์พกพาของผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Project	Modified Adaptive Thresholding for Pre-Processing of Text Recognition on Mobile Devices	
Student	Mr. Kittayoch Kongkatitham	Student ID 55070001
	Mr. Kittipop Peuwnuan	Student ID 55070010
Degree	Bachelor of Science	
Program	Information Technology	
Academic Year	2015	
Advisor	Assistant Professor Dr.Kuntpong Woraratpanya	
Co-Advisor	Assistant Professor Dr.Kitsuchart Pasupa	

ABSTRACT

After Thailand has joined in ASEAN Economic Community, communication with trading partners brings information in foreign languages to this country. Furthermore, there are many publications, advertisements, announcements, etc. that people need translation tools to help them better understanding.

There are many translation applications developed for translating foreign languages into Thai. However, there are still many challenges of applications for text translation from natural scene, because of diversity of components in an image diversity of character patterns and especially problems from shadow and lighting conditions. This project is to developing an application by researching and developing how to adjust an image to support text recognition. This application is not only used for recognition for translation, but also designed for using the recognized text on user's mobile devices.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงงานนี้สำเร็จได้ด้วยความกรุณาจาก ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา ซึ่งเป็นอาจารย์ที่ปรึกษา และ ผู้ช่วยศาสตราจารย์ ดร.กิติ์สุชาติ พสุภา ซึ่งเป็นอาจารย์ที่ปรึกษาร่วม ที่ให้ความช่วยเหลือ ให้คำชี้แนะ ตรวจสอบแก้ไขโครงงาน แสดงความเอาใจใส่ตลอดจนให้ความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า

ขอขอบพระคุณคณาจารย์คณะเทคโนโลยีสารสนเทศซึ่งถ่ายทอดความรู้ให้ ทั้งที่ได้ใช้โดยทางตรงและทางอ้อมในการพัฒนาโครงงานนี้

ขอขอบพระคุณบิดามารดาและครอบครัวที่เป็นแรงผลักดันสำคัญโดยการสนับสนุนด้านการศึกษา และให้กำลังใจในการมุ่งมั่นจนประสบความสำเร็จ

ขอบคุณทุกกำลังใจและความช่วยเหลือจากบุคคลรอบข้างทั้งที่ระลึกได้และระลึกไม่ได้

จึงขอแสดงความขอบพระคุณเป็นอย่างยิ่งไว้ ณ โอกาสนี้

กฤตยชญ์ คงคติธรรม

กิตติภพ ผิวนวล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ.....III

สารบัญ..... IV

สารบัญตาราง VII

สารบัญรูป.....VIII

บทที่ 1 บทนำ..... 1

1.1 ความเป็นมาของปัญหา..... 1

1.2 ความมุ่งหมายและวัตถุประสงค์..... 1

1.3 วิธีการดำเนินงาน..... 1

1.4 ขอบเขตของโครงงาน..... 1

1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 2

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง..... 3

2.1 การรู้จำข้อความ..... 3

2.1.1 การประมวลผลเบื้องต้น (Pre-Processing)..... 3

2.1.2 การรู้จำข้อความ (Text recognition)..... 4

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)..... 4

2.2 การแก้ไขปัญหาสำหรับการรู้จำข้อความ..... 5

บทที่ 3 การดำเนินงานวิจัย..... 7

3.1 การศึกษากระบวนการรู้จำข้อความ (Text Recognition)..... 7

3.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา..... 7

3.3 การศึกษาความต้องการต่อการใช้แอปพลิเคชัน..... 7

3.4 การศึกษาการปรับปรุงภาพต้นฉบับ..... 7

3.4.1 ขั้นตอนในการศึกษาการปรับปรุงภาพต้นฉบับมีดังนี้..... 7

3.4.2 นิยามลักษณะปัญหา..... 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

3.4.3 การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะ 8
3.5 ผลการดำเนินการ 10
3.5.1 การศึกษากระบวนการรู้จำข้อความ (Text Recognition) 10
3.5.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา 10
3.5.3 การศึกษาการปรับปรุงภาพต้นฉบับ 11
3.6 การออกแบบระบบใหม่ 12
3.6.1 ยูสเคสโมเดล 12
3.6.2 แผนภาพกิจกรรม (Activity Diagram) 16
3.6.3 แผนภาพคลาส (Class Diagram) 17

บทที่ 4 กระบวนการทำงานของระบบ 18
4.1 แอปพลิเคชัน 18
4.1.1 ส่วนที่ user ใช้งานแอปพลิเคชัน 18
4.1.1.1 หน้าจอการเลือกรูปภาพที่มีหรือเลือกถ่ายรูปภาพใหม่ 18
4.1.1.2 หน้าจอการถ่ายภาพของแอปพลิเคชัน 19
4.1.1.3 หน้าจอการเลือกรูปภาพหรือถ่ายภาพที่ต้องการเสร็จแล้ว 20
4.1.1.4 หน้าจอการเลือกภาษาที่ต้องการรู้จำตัวอักษร 21
4.1.1.5 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อ 22
4.1.1.6 หน้าจอการเลือกรูปภาพและเลือกภาษาที่ต้องการสกัดข้อความ 23
4.1.1.7 หน้าจอเมื่อเลือกรูปภาพและภาษาที่ต้องการจะสกัดข้อความแล้ว 23
4.1.1.8 หน้าจอการแจ้งเตือนการเลือกภาษาและรูปภาพที่ต้องการจะสกัดข้อความ 24
4.1.1.9 หน้าจอการสกัดข้อความออกมาและแปลภาษา 24
4.2 การทดลองและการสรุปผล 24
4.2.1 ตัวอย่างรูปที่ใช้ในการทดลอง 25
4.2.2 สรุปผลการทดลอง 26

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 5 สรุปผลโครงงานและข้อเสนอแนะ 28

5.1 สรุปผลการดำเนินงาน 28

5.2 ปัญหาและอุปสรรคที่พบ 28

5.3 ข้อจำกัดของแอปพลิเคชัน 28

5.4 ข้อเสนอแนะ 29

บรรณานุกรม 30

ภาคผนวก 31

ภาคผนวก ก. ผลการสำรวจผู้ใช้งานแอปพลิเคชันเพื่อการแปลภาษา 32

ภาคผนวก ข. รูปที่ใช้ในการทดลองแอปพลิเคชัน 35

ประวัติผู้เขียน 45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

3.1 รายละเอียดของยูสเคสการสกัดข้อความจากภาพ 13

3.2 รายละเอียดของยูสเคสการแปลข้อความที่พบ 14

3.3 รายละเอียดของยูสเคสการนำข้อความไปใช้บนอุปกรณ์ 15

4.1 การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชัน(โดยเฉลี่ย) 25

4.2 การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชันเฉพาะภาพที่มีปัญหาแสงและเงา (โดยเฉลี่ย) 25

4.3 การจำแนกปัญหาของที่ปรากฏในแต่ละรูปภาพที่เป็นชุดทดสอบ 25

4.4 ตารางแสดงรูปภาพกรณีที่ Recall สูงขึ้น 26

4.5 ตารางแสดงผลลัพธ์กรณีที่ Recall สูงขึ้น 27

ข.1 ตารางแสดงรูปภาพที่ใช้ในการทดลอง 36

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2.1 ขั้นตอนการทำงานของระบบ Optical character recognition 3
2.2 ภาพการพิจารณาพื้นที่ย่อยในรูปภาพ 6
3.1 ตัวอย่างภาพที่มีปัญหาด้านเงา 8
3.2 ตัวอย่างภาพที่มีปัญหาแสงสว่าง 8
3.3 แผนภาพยูสเคสของระบบรู้จำข้อความจากฉากธรรมชาติ 12
3.4 แผนภาพกิจกรรมของระบบรู้จำข้อความจากฉากธรรมชาติ 16
3.5 แผนภาพคลาสของระบบรู้จำข้อความจากฉากธรรมชาติ 17
4.1 หน้าจอการเลือกรูปภาพที่มีหรือเลือกถ่ายรูปภาพใหม่บนsmartphone 18
4.2 หน้าจอการถ่ายภาพของแอปพลิเคชันบนsmartphone 19
4.3 หน้าจอการเลือกรูปภาพหรือถ่ายภาพที่ต้องการเสร็จแล้วบนsmartphone 20
4.4 หน้าจอการเลือกภาษาที่ต้องการรู้จำตัวอักษรบนsmartphone 21
4.5 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อบนsmartphone 22
4.6 หน้าจอการเลือกรูปภาพและเลือกภาษาที่ต้องการสกัดข้อความบนsurface 23
4.7 หน้าจอเมื่อเลือกรูปภาพและภาษาที่ต้องการจะสกัดข้อความแล้วบนsurface 23
4.8 หน้าจอการแจ้งเตือนการเลือกภาษาและรูปภาพที่ต้องการจะสกัดข้อความบนsurface 24
4.9 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อบนsurface 24
ก.1 ผลของการสำรวจผู้ใช้งานแอปพลิเคชันเพื่อการแปลภาษา 33
ก.2 แผนภูมิของผลของการสำรวจผู้ใช้งานแอปพลิเคชันเพื่อการแปลภาษา 34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาของปัญหา

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตามการพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และโดยเฉพาะปัญหาแสงเงาที่กระทบต่อการรู้จำข้อความ

1.2 ความมุ่งหมายและวัตถุประสงค์

1) เพื่อศึกษาและปรับปรุงกระบวนการรู้จำข้อความ (Text Recognition) บนอุปกรณ์พกพา (Mobile Devices) ให้มีประสิทธิภาพดียิ่งขึ้น
2) เพื่อพัฒนาแอปพลิเคชันที่สามารถรู้จำข้อความจากฉากธรรมชาติ เพื่อความสะดวกต่อการนำข้อความที่ได้ไปใช้ต่อบนอุปกรณ์พกพา

1.3 วิธีการดำเนินงาน

1) ศึกษากระบวนการรู้จำข้อความ (Text Recognition) และกระบวนการประมวลผลเบื้องต้น (Pre-Processing)
2) ศึกษาปัญหาที่เกิดในการรู้จำข้อความ และหาวิธีการแก้ไข
3) พัฒนาเป็นแอปพลิเคชันที่สามารถรู้จำข้อความจากฉากธรรมชาติ

1.4 ขอบเขตของโครงงาน

1) ศึกษาและพัฒนากระบวนการรู้จำข้อความ (Text Recognition) บนอุปกรณ์พกพา (Mobile Devices)
2) ศึกษาและสร้างกระบวนการการแก้ไขปัญหาจากการรู้จำข้อความ โดยการปรับปรุงลักษณะของภาพต้นฉบับ
3) พัฒนาบนระบบปฏิบัติการ Windows/Windows Phone

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1) ได้เรียนรู้กระบวนการรู้จำข้อความ (Text Recognition) และแก้ไขปัญหาที่เกิดในการรู้จำข้อความเพื่อให้กระบวนการมีประสิทธิภาพยิ่งขึ้น
2) ได้พัฒนาแอปพลิเคชันเพื่อการรู้จำข้อความ สำหรับนำข้อความที่ได้ไปใช้ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การรู้จำข้อความ

OCR หรือ Optical Character Recognition เป็นกระบวนการแปลงข้อมูลภาพของข้อความทั้งตัวพิมพ์และตัวเขียน (ลายมือ) ไปสู่ข้อความในแบบที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้ ซึ่งก็คือการแปลงข้อมูลในเอกสาร สื่อสิ่งพิมพ์ต่าง ๆ ให้อยู่ในรูปแบบข้อมูลดิจิทัลที่สามารถนำไปแก้ไข ค้นหา จัดหมวดหมู่ จัดเก็บ แสดงผล และประมวลผลได้ เช่น การแปลภาษา การสังเคราะห์เสียงพูดจากข้อความ การจัดทำฐานข้อมูล ฯลฯ

OCR มีความเกี่ยวข้องกับทั้งการศึกษาการรู้จำรูปแบบ (Pattern Recognition) ปัญญาประดิษฐ์ (Artificial Intelligence) และคอมพิวเตอร์วิทัศน์ (Computer Vision)
ขั้นตอนการทำงานของ OCR มี 3 ขั้นตอนหลัก

รูปที่ 2.1 ขั้นตอนการทำงานของระบบ Optical character recognition

2.1.1 การประมวลผลเบื้องต้น (Pre-Processing)

กระบวนการเบื้องต้นก่อนที่จะนำรูปไปประมวลผล คือ การนำรูปภาพมาปรับแต่ง และจัดเตรียมให้เหมาะสมก่อนจะถูกนำไปประมวลผล ซึ่งหากเกิดความผิดพลาดจะมีผลกระทบต่อขั้นตอนการประมวลผลต่อไป การปรับแต่งภาพต้นฉบับมีหลายวิธีการ ดังตัวอย่างต่อไปนี้

2.1.1.1 De-skew

หากข้อความที่อยู่ในรูปมีการเอียง จะทำการปรับองศาเพื่อให้รูปตรง โดยหมุนตามเข็มนาฬิกา หรือ ทวนเข็มนาฬิกา ขึ้นอยู่กับว่าหมุนด้านไหนใช้องศาน้อยที่สุด

2.1.1.2 Despeckle

เป็นการนำเอา positive และ negative spot ออก ทำให้ขอบมีความเรียบเนียนมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.3 Binarization

แปลงภาพจากภาพต้นฉบับให้เป็น ภาพโทนสีเทา (Grayscale Image) และแปลงต่อให้เป็นภาพขาวดำ (Binary Image) เพื่อให้การแยกตัวอักษรจากพื้นหลังกระทำได้ง่ายขึ้น

2.1.2 การรู้จำข้อความ (Text recognition)

การรู้จำข้อความจากฉากธรรมชาติ (Text Recognition in Natural Scenes) เป็นกระบวนการประมวลผลข้อมูลภาพถ่ายที่มีข้อความปรากฏอยู่ เพื่อสกัดข้อความนั้น ๆ ออกมาเป็นข้อความในรูปแบบดิจิทัลที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้

กระบวนการที่สำคัญในการรู้จำข้อความจากฉากธรรมชาติได้แก่ การดำเนินการเบื้องต้น การตรวจจับข้อความ การรู้จำข้อความ และการดำเนินการหลังการประมวลผล
ประเภทของหลักการ OCR พื้นฐานแบ่งออกได้เป็น 2 รูปแบบ ดังต่อไปนี้

2.1.2.1 Matrix matching

เป็นการเปรียบเทียบรูปภาพแบบอักษร โดยที่การเปรียบเทียบจะเทียบโดย พิกเซลต่อพิกเซล หรือเรียกได้ว่า "pattern matching", "pattern recognition" หรือ "image correlation" จะขึ้นอยู่กับอินพุทซึ่งแยกออกมาอย่างถูกต้องจากส่วนอื่นของภาพ และในตัวอักษรเฉพาะที่เก็บไว้ ซึ่งมีฟอนต์ ที่ใกล้เคียงกันและขนาดเดียวกัน เทคนิคนี้ทำงานได้ดีกับตัวพิมพ์ และทำงานได้ไม่ดีเมื่อพบฟอนต์ ใหม่ๆ

2.1.2.2 Feature extraction

จะแยกตัวอักษรเฉพาะ ออกเป็นลักษณะเช่น เส้น เส้นปิด ทิศทางของเส้น และจุดตัดของเส้น ซึ่งเป็นตัวแทนลักษณะของตัวอักษร กระบวนการ Nearest neighbor classifiers เช่น k-nearest neighbors algorithm ถูกใช้ในการเปรียบเทียบคุณลักษณะของรูปภาพกับตัวอักษรเฉพาะ และเลือกรูปแบบที่ใกล้เคียงที่สุด

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)

ความแม่นยำของกระบวนการ OCR สามารถเพิ่มขึ้นได้ถ้าหากผลลัพธ์ถูกควบคุมโดยรายการของศัพท์เฉพาะที่อนุญาตให้ใช้ในเอกสาร(lexicon) ซึ่งเป็นรายการของคำที่สามารถปรากฏในเอกสารได้ ตัวอย่างเช่น คำในภาษาอังกฤษ ทั้งหมด และศัพท์เทคนิคเฉพาะซึ่งเทคนิคนี้อาจเป็นปัญหาได้ถ้าในเอกสารมีข้อความที่ไม่ได้อยู่ใน lexicon นอกจากนี้ยังสามารถใช้วิธีการ Near-neighbor analysis แก้ไขคำผิดได้โดยการเปรียบเทียบกับตัวอักษรที่มักจะปรากฏในข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เดียวกัน และไวยากรณ์ภาษายังช่วยในการระบุคำที่พบว่าควรจะเป็นคำนาม คำกริยา หรืออื่น ๆ เป็นต้น

2.2 การแก้ไขปัญหาสำหรับการรู้จำข้อความ

Chucai Yi และYingli Tian [1] กล่าวไว้ว่า สำหรับการสกัดข้อความจากฉากธรรมชาติด้วยอุปกรณ์พกพานั้น การตรวจจับข้อความอย่างอัตโนมัติและมีประสิทธิผล และอัลกอริทึมสำหรับการรู้จำถือเป็นสิ่งจำเป็นอย่างยิ่ง อย่างไรก็ตามการสกัดข้อความจากภาพฉากธรรมชาติมีอุปสรรคท้าทายหลายประการเช่น ความหลายหลายของฉากหลังที่มีสิ่งรบกวนและสิ่งที่ไม่ใช่ข้อความปะปนอยู่ และความหลากกหลายของรูปแบบตัวอักษรเช่น ประเภทตัวอักษร แบบตัวหนังสือ และขนาด ดังนั้นการแก้ไขปัญหาสำหรับการรู้จำข้อความจึงเป็นความพยายามในการทำให้ความหลากหลายของฉากหลังและความหลากหลายของรูปแบบตัวอักษรมีผลต่อกระบวนการต่าง ๆ เช่น การตรวจจับข้อความ และการรู้จำข้อความให้น้อยที่สุด เพื่อให้ผลลัพธ์มีความถูกต้องและแม่นยำมากขึ้น สอดคล้องกับที่ Woraratpanya, K และคณะ [6] กล่าวไว้ว่า อย่างไรก็ตามประสิทธิผลของอัตราการรู้จำขึ้นอยู่กับการระบุตำแหน่งของข้อความ ความบริสุทธิ์ของกระบวนการแยกข้อความออกจากฉากหลังที่มากจะนำไปสู่อัตราความแม่นยำของการรู้จำที่มากขึ้น

ในปี ค.ศ. 1993 Pierre D. Wellner [5] ได้นำเสนอวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding) สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) โดยการพิจารณาภาพโทนสีเทาจากภาพต้นฉบับ จากนั้นเปรียบเทียบข้อมูลค่าความสว่างในแต่ละตำแหน่งของรูปภาพกับค่าเฉลี่ยความสว่างของพื้นที่ย่อยที่ตำแหน่งนั้น ๆ ปรากฏอยู่ เพื่อใช้เป็นเกณฑ์กำหนดสำหรับการสร้างภาพขาวดำ วิธีการนี้สามารถนำมาใช้กับภาพถ่ายของเอกสารที่มีเงาตกกระทบเพื่อแก้ปัญหาข้อความที่ถูกเงาบังและไม่สามารถตรวจจับหรือรู้จำได้ และยังสามารถประมวลผลได้อย่างรวดเร็ว

การพิจารณาพื้นที่ย่อยในวิธีการของ Wellner จะเริ่มจากการพิจารณารูปภาพแต่ละรูปเป็นแถวของพิกเซลที่เรียงตัวเป็นแถวเดียว ความกว้างของพื้นที่ย่อย (s) แต่ละช่วงจะมีขนาดเป็น 1 ใน 8 ของความกว้างภาพต้นฉบับ จากนั้นพิจารณาผลรวมของค่าความสว่างจากภาพโทนสีเทาของภาพต้นฉบับภายในแต่ละพื้นที่ย่อยได้ด้วยสมการต่อไปนี้

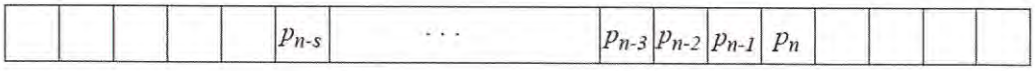
$$f_s(n) = \sum_{i=0}^{s-1} P_{n-1} \qquad (2.1)$$

เมื่อ $f_s(n)$ เป็นผลรวมของค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n

P_n คือค่าความสว่างของพิกเซล ณ ตำแหน่ง n ของภาพโทนสีเทาจากภาพต้นฉบับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

s คือความกว้างของพื้นที่ย่อย



รูปที่ 2.2 ภาพการพิจารณาพื้นที่ย่อยในรูปภาพ

ภาพผลลัพธ์จากการประมวลผลคำนวณได้จากการกำหนดค่าในแต่ละพิกเซลของภาพผลลัพธ์ด้วยเงื่อนไขต่อไปนี้

$$T(n)=\begin{cases} 0, \text{ เมื่อ } P_n < \left(\frac{f_s(n)}{s}\right)\left(\frac{100\text{-}t}{100}\right) \\ 1, \text{ กรณีอื่น ๆ} \end{cases} \quad (2.2)$$

เมื่อ t เป็นเปอร์เซ็นต์ที่ค่าความสว่าง ณ ตำแหน่ง n น้อยกว่าค่าเฉลี่ยของพื้นที่ย่อยที่พิจารณา (Wellner เสนอให้ค่า t เป็น15) โดยที่เลข 0 แทนสีดำ และเลข 1 แทนสีขาว
ซึ่งผลลัพธ์ที่ได้จะเป็นภาพขาวดำ (Binarization Image)

วิธีการของ Wellner เป็นเพียงรูปแบบหนึ่งของวิธีการกำหนดเกณฑ์แบบปรับเหมาะสำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ ยังมีวิธีการที่แตกต่างออกไปอีกมากมาย เช่น ในปี ค.ศ. 2007 Derek Bradley และคณะ [2] ได้นำเสนอการปรับปรุงวิธีการของ Wellner เพื่อเพิ่มประสิทธิภาพในการแก้ปัญหาในบทความ Adaptive Thresholding Using Integral Image โดยปรับเปลี่ยนวิธีการพิจารณาพื้นที่ย่อยในขั้นตอนการเปรียบเทียบค่าความสว่างของแต่ละตำแหน่งในรูปภาพ

นอกเหนือจากการใช้วิธีการใดวิธีการหนึ่งในการประมวลผลแต่ละรูปภาพ แนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการแสดงให้เห็นผลลัพธ์ที่ดีกว่าดังที่ J. Sauvola และคณะ [3] ได้นำเสนอไว้ในบทความในปี ค.ศ. 2000 โดยดำเนินการประมวลผลจำแนกข้อความและพื้นหลังจากภาพถ่ายจากนั้นได้แบ่งวิธีการประมวลผลออกเป็นสองวิธี วิธีการหนึ่งเป็นการแยกองค์ประกอบพื้นหลัง และอีกวิธีการหนึ่งเป็นการแปลงภาพเป็นภาพขาวดำเพื่อการสกัดข้อความออกจากพื้นหลังภายใต้ปัญหาที่ส่งผลต่อการตรวจจับข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การดำเนินงานวิจัย

3.1 การศึกษากระบวนการรู้จำข้อความ (Text Recognition)

สำหรับการศึกษากระบวนการรู้จำข้อความ (Text Recognition) ทำโดยการค้นคว้าจากทฤษฎีทางด้านการประมวลผลรูปภาพ (Image Processing) และการค้นคว้าจากบทความทางวิชาการที่เกี่ยวข้องกับการรู้จำข้อความ, การแก้ปัญหาด้านต่าง ๆ ที่เป็นอุปสรรคต่อการตรวจจับข้อความ และการรู้จำข้อความ รวมถึงบทความที่เกี่ยวข้องกับการรู้จำข้อความบนอุปกรณ์พกพา (Mobile Devices)

3.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

สำหรับการศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา ได้ศึกษาการพัฒนาแอปพลิเคชันบน Windows Platform และได้ทดลองพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย เพื่อทำการศึกษาวิธีการพัฒนาการรู้จำข้อความบนอุปกรณ์พกพาและเรียนรู้ปัญหาหรืออุปสรรคที่เกิดขึ้น

3.3 การศึกษาความต้องการต่อการใช้แอปพลิเคชัน

สำหรับการศึกษาความต้องการต่อการใช้แอปพลิเคชัน เป็นการสำรวจความต้องการของกลุ่มผู้ใช้เป้าหมายเพื่อเป็นตัวกำหนดคุณลักษณะ (Feature) สำหรับแอปพลิเคชันที่พัฒนา

3.4 การศึกษาการปรับปรุงภาพต้นฉบับ

การศึกษาการปรับปรุงภาพต้นฉบับต้องการแก้ปัญหาการตรวจจับข้อความในฉากธรรมชาติดังต่อไปนี้ คือ ปัญหาด้านเงา และ ปัญหาด้านแสงสว่าง

3.4.1 ขั้นตอนในการศึกษาการปรับปรุงภาพต้นฉบับมีดังนี้

1. ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ
2. ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา
3. ปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชัน
4. ทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2 นิยามลักษณะปัญหา

3.4.2.1 ปัญหาด้านเงา

1. ภาพที่มีเงาบดบังบางส่วนของบริเวณที่มีข้อความ
2. ภาพที่มีเงาบดบังบริเวณที่มีข้อความทั้งหมด

3.4.2.2. ปัญหาด้านแสงสว่าง

1 ภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบางส่วนของบริเวณที่มีข้อความ

2 ภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบริเวณที่มีข้อความทั้งหมด

รูปที่ 3.1 ตัวอย่างภาพที่มีปัญหาด้านเงา

รูปที่ 3.2 ตัวอย่างภาพที่มีปัญหาแสงสว่าง

3.4.3 การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะ

จากการศึกษากระบวนการแก้ปัญหาของ Pierre D. Wellner [5] ดังที่ได้อธิบายวิธีการในหัวข้อ 2.2 พบว่าสามารถประมวลผลได้อย่างรวดเร็วเนื่องจากกระบวนการมีความซับซ้อนที่น้อย ทว่ามีข้อจำกัด ข้อบกพร่อง และข้อสังเกตในการประมวลผลดังนี้

1) ให้ผลลัพธ์ที่ดีเฉพาะกับภาพที่มีปัญหาเงาในระดับปานกลางถึงรุนแรง
2) ให้ผลลัพธ์ที่ไม่ดีกับกรณีทั่วไป ได้แก่ ปัญหาแสง ปัญหาอื่น ๆ ที่นอกเหนือจากนิยาม และกรณีภาพที่ไม่มีปัญหา
3) หากใช้ภาพส่วนกลับโทนสีเทา (Complement-grayscale image) ของภาพต้นฉบับในการดำเนินการจะสามารถให้ผลลัพธ์ที่ดีกับภาพที่มีปัญหาแสงในระดับปานกลางถึงรุนแรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากแนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการดังที่ J. Sauvola และคณะ [3] ได้ดำเนินการศึกษาและพัฒนากระบวนการแก้ปัญหาต่อการรู้จำข้อความ ทำให้เชื่อว่าอาจมีวิธีการดัดแปลงให้การแก้ปัญหาของ Wellner สามารถแก้ไขปัญหาแสงได้โดยการประยุกต์ใช้วิธีการร่วม

สำหรับการประยุกต์ใช้วิธีการร่วมเพื่อแก้ปัญหาของรูปภาพนั้น K. Peuwnuan และคณะ [4] ได้นำเสนอวิธีการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะโดยใช้ภาพรวมผล (Integral Image) สำหรับการตรวจจับข้อความ ซึ่งเป็นการดัดแปลงกระบวนการของ Derek Bradley และคณะ [2] โดยมีการแบ่งพื้นที่รูปภาพออกเป็นส่วนต่าง ๆ (Image Segmentation) คือพื้นที่ที่มีค่าความสว่างสูงและพื้นที่ที่มีค่าความสว่างปานกลางถึงต่ำ จากนั้นจึงดำเนินการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในแต่ละพื้นที่ด้วยวิธีการที่แตกต่างกัน

โดยอาศัยแนวคิดข้างต้น ในการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพาจึงได้ดัดแปลงวิธีการของ Wellner ในลักษณะดังต่อไปนี้

3.4.3.1 กระบวนการจำแนกพื้นที่

ก่อนดำเนินการกำหนดเกณฑ์แบบปรับเหมาะจำเป็นต้องมีการจำแนกพื้นที่ออกเป็นพื้นที่ที่มีค่าความสว่างสูงและค่าความสว่างต่ำ เนื่องจากพื้นที่ที่จำแนกได้นั้นจะใช้วิธีการประมวลผลเพื่อการแปลงภาพขาวดำที่แตกต่างกัน โดยเงื่อนไขที่ใช้จำแนกพื้นที่ของรูปภาพแสดงโดยสมการที่ (3.1)

$$P(n) = \begin{cases} \text{พื้นที่ความสว่างสูง, เมื่อ } P_n > M + SD \\ \text{พื้นที่ความสว่างต่ำ, กรณีอื่น ๆ} \end{cases} \quad (3.1)$$

เมื่อ P_n คือค่าความสว่างของพิกเซลที่พิจารณาจากภาพโทนสีเทา

M คือค่าเฉลี่ยความสว่างของทั้งรูปภาพ

SD คือส่วนเบี่ยงเบนมาตรฐานของค่าความสว่างทั้งรูปภาพ

3.4.3.2 การกำหนดเกณฑ์แบบปรับเหมาะ

หลังจากได้แบ่งพื้นที่รูปภาพตามเงื่อนไขในสมการ (3.1) แล้วจึงดำเนินการกำหนดเกณฑ์แบบปรับเหมาะสำหรับแต่ละพื้นที่ซึ่งจะมีความแตกต่างกันดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P'_n \geq \left(\frac{f'_s(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (3.2)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำซึ่งเป็นผลลัพธ์

P'_n คือค่าความสว่างของพิกเซลที่ n ของภาพส่วนกลับโทนสีเทา

$f'_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพส่วนกลับโทนสีเทา

สำหรับสมการ (3.2) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างสูงตามที่จำแนกได้ในหัวข้อ 3.4.3.1 และ

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P_n \geq \left(\frac{f_s(n)}{s}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (3.3)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำซึ่งเป็นผลลัพธ์

P_n คือค่าความสว่างของพิกเซลที่ n ของภาพโทนสีเทา

$f_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพโทนสีเทา

สำหรับสมการ (3.3) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างปานกลางถึงต่ำตามที่จำแนกได้ในหัวข้อ 3.4.3.1

3.5 ผลการดำเนินการ

3.5.1 การศึกษากระบวนการรู้จำข้อความ (Text Recognition)

ผลการศึกษาฯ ทำให้มีความรู้ความเข้าใจในกระบวนการรู้จำข้อความมากยิ่งขึ้น เพียงพอต่อการนำไปใช้เพื่อพัฒนาแอปพลิเคชันต่อไป รวมถึงเข้าใจในปัญหาและอุปสรรคต่อการตรวจจับและการรู้จำข้อความ

3.5.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

ผลการศึกษาและทดลองฯ ทำให้สามารถพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย และได้ออกแบบการทดลองและบันทึกผลการรู้จำข้อความจากภาพตัวอย่างที่มีความหลากหลายขององค์ประกอบในภาพ เพื่อคำนวณความแม่นยำของแอปพลิเคชันก่อนจะพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.3 การศึกษาการปรับปรุงภาพต้นฉบับ

3.5.3.1. ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ

จากการศึกษาพบว่าการปรับปรุงภาพต้นฉบับด้วยวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding) สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) สามารถประมวลผลได้เร็วเหมาะแก่การพัฒนาบนอุปกรณ์พกพาโดยเฉพาะวิธีการของ Wellner ซึ่งเหมาะอย่างยิ่งสำหรับกรณีภาพที่มีปัญหาเงาปานกลางถึงรุนแรง และสามารถดัดแปลงเพื่อให้ใช้แก้ปัญหาแสงได้เช่นกัน

3.5.3.2. ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา

มีการทดลองใช้วิธีการของ Wellner เพื่อดูผลในการแก้ปัญหาเงาที่มีผลต่อการตรวจจับข้อความ พบว่าสามารถแก้ปัญหาเงาระดับปานกลางถึงสูงได้ดี แต่จะส่งผลลบกับการประมวลผลหากนำไปใช้กับภาพที่มีปัญหาแสงหรือปัญหาอื่น ๆ กล่าวคือจะลดความแม่นยำ ความถูกต้อง และความครบถ้วนของผลลัพธ์ลงอย่างมาก แต่สามารถปรับให้ใช้แก้ไขปัญหาแสงได้หากใช้ภาพส่วนกลับของภาพโทนสีเทา (Complement-Grayscale Image) ในการดำเนินการ

3.5.3.3. ปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชัน

มีการปรับใช้วิธีการของ Wellner เพื่อให้สามารถแก้ไขปัญหาของภาพได้ทั้งเงาและแสงโดยการสร้างเงื่อนไขเพื่อพิจารณาพื้นที่ของรูปภาพที่เป็นส่วนความสว่างต่ำถึงปานกลาง และความสว่างสูง และใช้วิธีการดำเนินการในแต่ละพื้นที่ที่แตกต่างกันซึ่งให้ผลลัพธ์ที่ดีขึ้นสำหรับกรณีปัญหาแสงที่เกิดกับรูปภาพ

3.5.3.4. ทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น

ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงาแต่มีปัญหาในกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้ ซึ่งตัวเลขผลการทดลอง การเปรียบเทียบ และการอภิปรายจะได้กล่าวถึงในรายละเอียด ในหัวข้อ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 การออกแบบระบบใหม่

3.6.1 ยูสเคสโมเดล

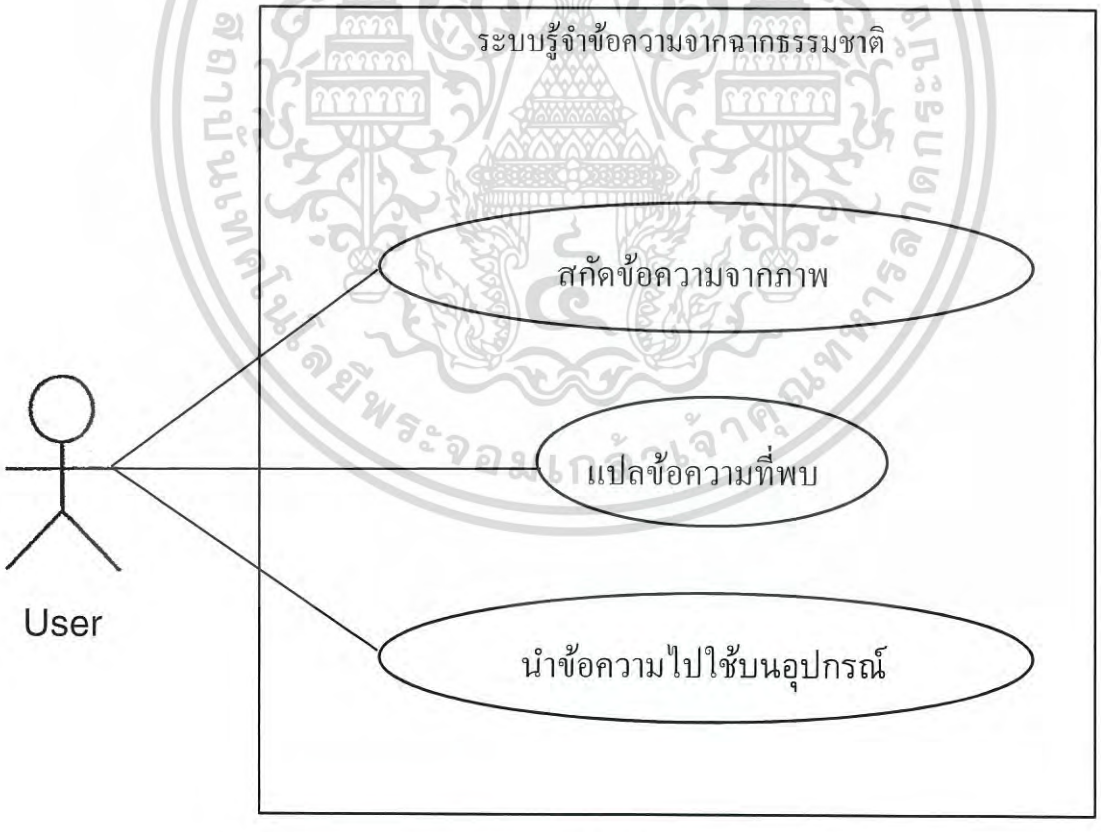
3.6.1.1 ผู้ที่เกี่ยวข้องในระบบ (Actor)

1. User

3.6.1.2 ฟังก์ชันการทำงานหลัก

1. การสกัดข้อความจากธรรมชาติ
2. การแปลข้อความที่พบ
3. นำข้อความไปใช้บนอุปกรณ์พกพา (Mobile Devices)

3.6.1.3 แผนภาพยูสเคส (Use Case Diagram)



รูปที่ 3.3 แผนภาพยูสเคสของระบบรู้จำข้อความจากฉากธรรมชาติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6.1.4 รายละเอียดของยูสเคส (Use Case Description)

ตารางที่ 3.1 รายละเอียดของยูสเคสการสกัดข้อความจากภาพ

Use Case Name: สกัดข้อความจากฉากธรรมชาติ	**ID :** 1	**Importance Level :** high
Primary Actor : user	**Use Case Type :** Detail, Essential	
Stakeholders and Interests : user ต้องการสกัดข้อความจากฉากธรรมชาติ		
Brief Description : ผู้ใช้สกัดข้อความจากภาพโดยนำเข้าอินพุทเป็นภาพถ่าย		
Relationships Association: user Include:- Extend:- Generalization:-		
Pre-Condition user ต้องนำเข้าอินพุทเป็นภาพถ่าย		
Post-Condition -		
Normal Flow of Events 1. user นำเข้ารูปภาพที่ต้องการจะสกัดข้อความจากฉากธรรมชาติ 2. user เลือกภาษาที่ต้องการให้สกัดข้อความจากฉากธรรมชาติ 3. โปรแกรมจะสกัดข้อความที่ user ต้องการ 4. user ได้ข้อความที่สกัดจากภาพ		
Alternate/Exceptional Flows: 2a. user ต้องการเปลี่ยนรูปภาพ 1. user นำเข้ารูปภาพใหม่		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 รายละเอียดของยูสเคสการแปลข้อความที่พบ

Use Case Name: แปลข้อความที่พบ	**ID :** 2	**Importance Level :** high
Primary Actor : user	**Use Case Type :** Detail, Essential	
Stakeholders and Interests : user ต้องการแปลข้อความ		
Brief Description : user ต้องการแปลข้อความที่พบในภาพ		
Relationships Association: user Include:- Extend:- Generalization:-		
Pre-Condition user ต้องมีข้อความที่สกัดได้จากฉากธรรมชาติก่อน		
Post-Condition -		
Normal Flow of Events 1. user กดใช้ฟังก์ชันการแปลภาษา 2. user เลือกภาษาที่ต้องการจะแปล 3. user ได้ข้อความที่แปลออกมา		
Alternate/Exceptional Flows:		

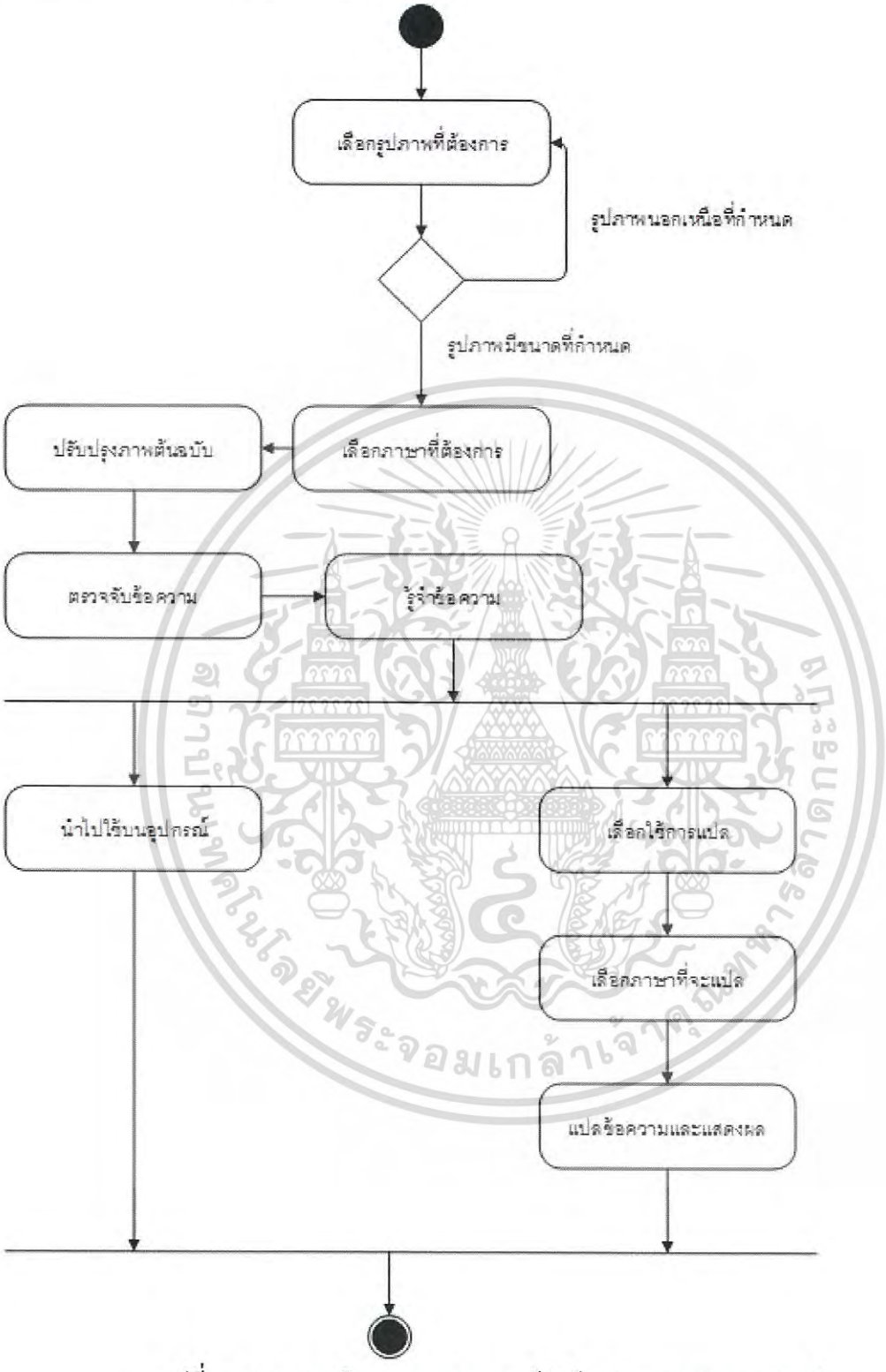
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 รายละเอียดของยูสเคสการนำข้อความไปใช้บนอุปกรณ์

Use Case Name: นำข้อความที่ได้ไปใช้บนอุปกรณ์	**ID :** 3	**Importance Level :** high
Primary Actor : user	**Use Case Type :** Detail, Essential	
Stakeholders and Interests : user ต้องการนำข้อความที่ได้ไปใช้ใน smartphone		
Brief Description : user ต้องการนำข้อความที่ได้ไปใช้ใน smartphone		
Relationships Association: user Include:- Extend:- Generalization:-		
Pre-Condition user ต้องมีข้อความที่สกัดได้จากฉากธรรมชาติก่อน		
Post-Condition -		
Normal Flow of Events 1. user กดใช้ฟังก์ชันการคัดลอกข้อความ 2. user จะได้ข้อความที่จะนำไปใช้ใน smartphone		
Alternate/Exceptional Flows:		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6.2 แผนภาพกิจกรรม (Activity Diagram)

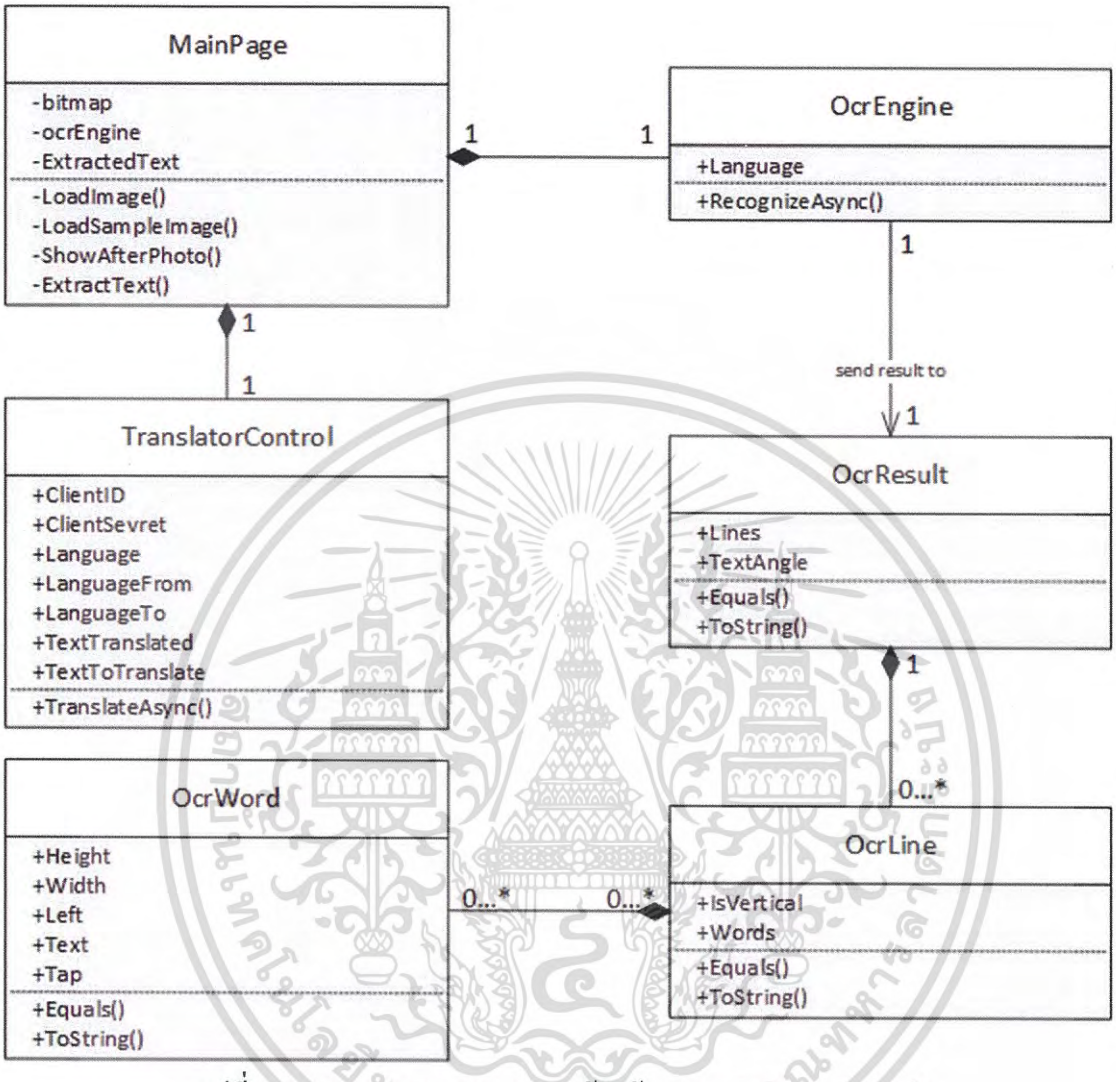


รูปที่ 3.4 แผนภาพกิจกรรมของระบบรู้จำข้อความจากฉากธรรมชาติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

3.6.3 แผนภาพคลาส (Class Diagram)



รูปที่ 3.5 แผนภาพคลาสของระบบรู้จำข้อความจากฉากธรรมชาติ

146185

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

กระบวนการทำงานของระบบ

4.1 แอปพลิเคชัน

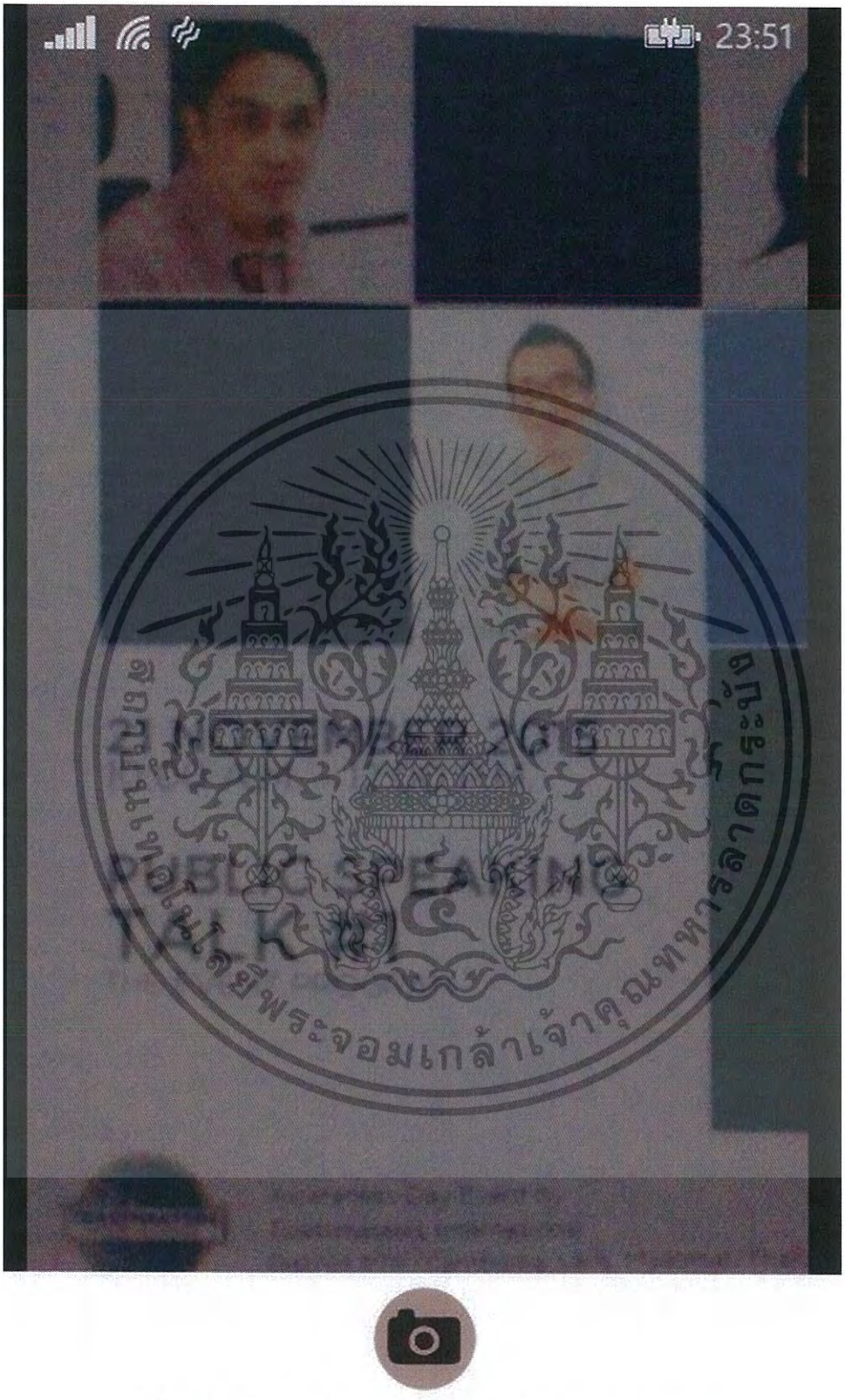
4.1.1 ส่วนที่ user ใช้งานแอปพลิเคชัน

4.1.1.1 หน้าจอการเลือกรูปภาพที่มีหรือเลือกถ่ายรูปภาพใหม่

รูปที่ 4.1 หน้าจอการเลือกรูปภาพที่มีหรือเลือกถ่ายรูปภาพใหม่บนsmartphone

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.2 หน้าจอการถ่ายภาพของแอปพลิเคชัน



รูปที่ 4.2 หน้าจอการถ่ายภาพของแอปพลิเคชันบนsmartphone

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.3 หน้าจอการเลือกรูปภาพหรือถ่ายภาพที่ต้องการเสร็จแล้ว

รูปที่ 4.3 หน้าจอการเลือกรูปภาพหรือถ่ายภาพที่ต้องการเสร็จแล้วบนsmartphone

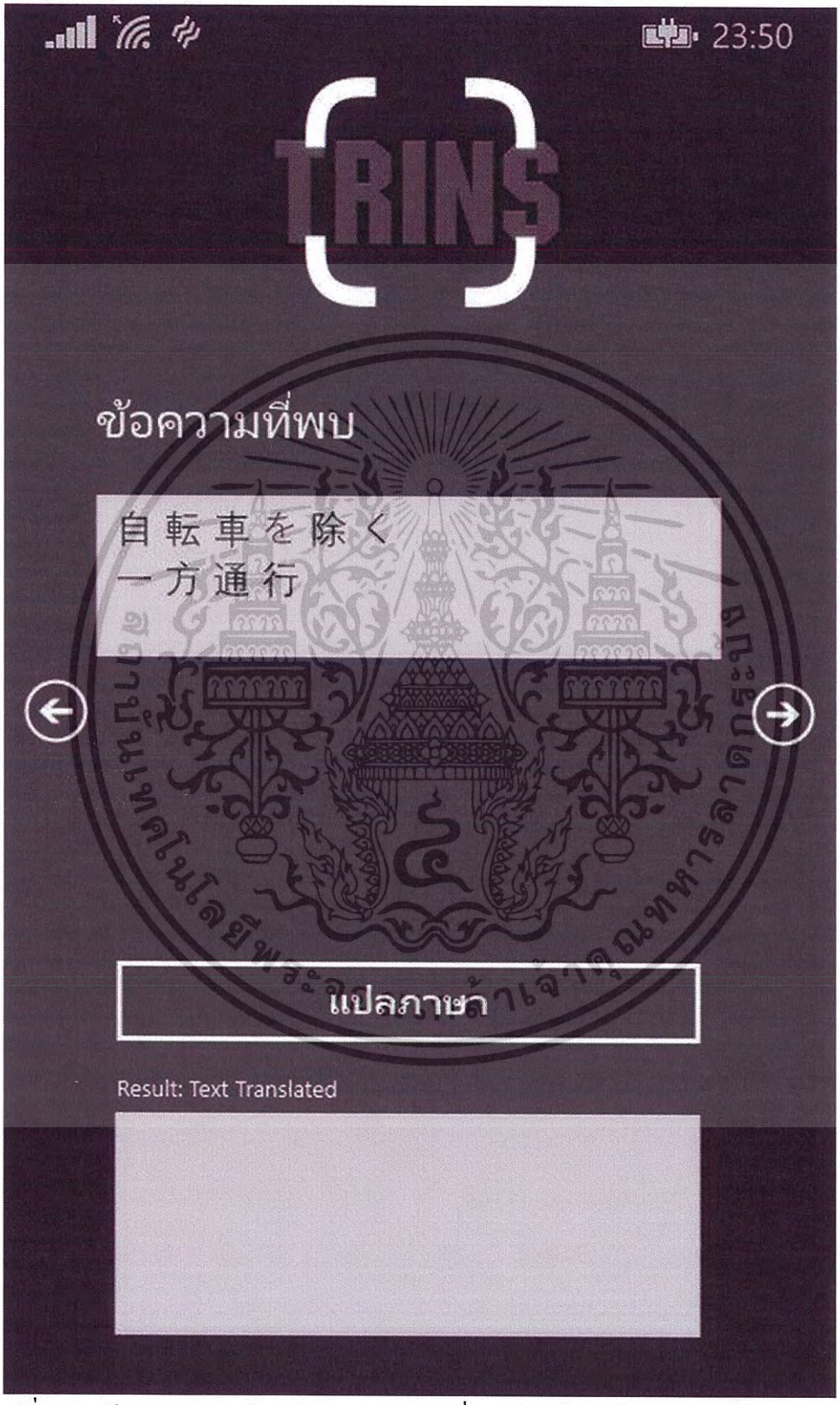
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.4 หน้าจอการเลือกภาษาที่ต้องการรู้จำตัวอักษร

รูปที่ 4.4 หน้าจอการเลือกภาษาที่ต้องการรู้จำตัวอักษรบนsmartphone

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.5 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อ



รูปที่ 4.5 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อบนsmartphone

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.6 หน้าจอการเลือกรูปภาพและเลือกภาษาที่ต้องการสกัดข้อความ

รูปที่ 4.6 หน้าจอการเลือกรูปภาพและเลือกภาษาที่ต้องการสกัดข้อความบนsurface

4.1.1.7 หน้าจอเมื่อเลือกรูปภาพและภาษาที่ต้องการจะสกัดข้อความแล้ว

รูปที่ 4.7 หน้าจอเมื่อเลือกรูปภาพและภาษาที่ต้องการจะสกัดข้อความแล้วบนsurface

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.8 หน้าจอการแจ้งเตือนการเลือกภาษาและรูปภาพที่ต้องการจะสกัดข้อความ

รูปที่ 4.8 หน้าจอการแจ้งเตือนการเลือกภาษาและรูปภาพที่ต้องการจะสกัดข้อความบนsurface

4.1.1.9 หน้าจอการสกัดข้อความออกมาและแปลภาษา

รูปที่ 4.9 หน้าจอการสกัดข้อความออกมาพร้อมที่จะแปลหรือนำไปใช้ต่อบนsurface

4.2 การทดลองและการสรุปผล

การทดลองมีจุดประสงค์เพื่อเปรียบเทียบผลลัพธ์จากการทำงานของแอพพลิเคชั่นก่อนและหลังการเพิ่มเติมการประมวลผลเบื้องต้น โดยชุดทดสอบที่ใช้เป็นภาพฉากธรรมชาติจากการแข่งขันของ ICDAR 2013 จำนวน 229 ภาพ ซึ่งเป็นภาพที่มีปัญหาแสงและ/หรือเงาเพียง 59 ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชัน(โดยเฉลี่ย)

การประมวลผล	ตัวชี้วัดการประเมินผล		
	precision(%)	**recall(%)**	**correctness(%)**
ก่อนเพิ่มการประมวลผลเบื้องต้น	82.66	74.44	82.18
หลังเพิ่มการประมวลผลเบื้องต้น	40.55	30.64	39.99

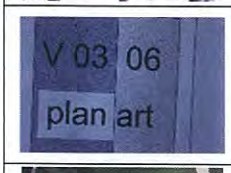
ตารางที่ 4.2 การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชันเฉพาะภาพที่มีปัญหาแสงและเงา (โดยเฉลี่ย)

การประมวลผล	ตัวชี้วัดการประเมินผล		
	precision(%)	**recall(%)**	**correctness(%)**
ก่อนเพิ่มการประมวลผลเบื้องต้น	80.90	74.03	80.63
หลังเพิ่มการประมวลผลเบื้องต้น	39.18	32.69	38.73

4.2.1 ตัวอย่างรูปที่ใช้ในการทดลอง

รูปที่ใช้ในการทดลองจำแนกออกตามนิยามปัญหาได้ ในการทดลองจำแนกออกตามนิยามปัญหาได้ 2 ประเภทคือ แสงและเงา และกรณีอื่น ๆ เพิ่มเข้ามาด้วย

ตารางที่ 4.3 การจำแนกปัญหาของที่ปรากฏในแต่ละรูปภาพที่เป็นชุดทดสอบ

ปัญหาเงา	ปัญหาแสง	กรณีอื่น ๆ
		

4.2.2 สรุปผลการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทดลอง

ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงามีจุดบกพร่องกับกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้

ตารางที่ 4.4 ตารางแสดงรูปภาพกรณีที่ Recall สูงขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 ตารางแสดงผลลัพธ์กรณีที่ Recall สูงขึ้น

รูป		ตัวชี้วัดการประเมินผล		
		precision(%)	**recall(%)**	**correctness(%)**
1	ก่อนเพิ่มการประมวลผลเบื้องต้น	50	13.04	100
	หลังเพิ่มการประมวลผลเบื้องต้น	92	100	100
2	ก่อนเพิ่มการประมวลผลเบื้องต้น	85.71	85.71	100
	หลังเพิ่มการประมวลผลเบื้องต้น	87.5	100	100
3	ก่อนเพิ่มการประมวลผลเบื้องต้น	100	60	66.67
	หลังเพิ่มการประมวลผลเบื้องต้น	71.43	100	100
4	ก่อนเพิ่มการประมวลผลเบื้องต้น	100	52.63	100
	หลังเพิ่มการประมวลผลเบื้องต้น	100	100	100
5	ก่อนเพิ่มการประมวลผลเบื้องต้น	100	65.22	100
	หลังเพิ่มการประมวลผลเบื้องต้น	100	69.57	87.5
6	ก่อนเพิ่มการประมวลผลเบื้องต้น	100	44.90	100
	หลังเพิ่มการประมวลผลเบื้องต้น	96.97	65.31	93.94

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลโครงงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

การศึกษาการระบบการรู้จำข้อความจากฉากธรรมชาติ ได้มีการนำไลบรารี OCR เข้ามาใช้แทนระบบการรู้จำข้อความเพื่อให้การประมวลผลบนอุปกรณ์พกพาประเภทสมาร์ทโฟนทำได้รวดเร็วยิ่งขึ้นและใช้ทรัพยากรที่มีจำกัดของอุปกณ์ได้อย่างมีประสิทธิภาพ

การศึกษาภาพการปรับปรุงต้นฉบับได้ใช้หลักการการประมวลผลรูปภาพ (Image Processing) เพื่อศึกษาทดลองวิธีการปรับปรุงภาพต้นฉบับก่อนเข้าสู่กระบวนการรู้จำข้อความ พบว่าการปรับปรุงภาพต้นฉบับต้องพัฒนาอัลกอริทึมเพิ่มเติมที่สามารถระบุวิธีการปรับปรุงภาพต้นฉบับที่จำเพาะกับภาพนั้นๆ เนื่องจากภาพถ่ายจากฉากธรรมชาติมีความหลากหลายมากมาย ทำให้กระบวนการปรับปรุงภาพต้นฉบับไม่สามารถใช้รูปแบบใดรูปแบบหนึ่งได้กับทุกภาพ

5.2 ปัญหาและอุปสรรคที่พบ

1. การปรับปรุงภาพต้นฉบับเพื่อให้เอื้อต่อการรู้จำนั้นดำเนินการได้ยาก เนื่องจากภาพถ่ายจากฉากธรรมชาติมีความหลากหลายมากมาย ทำให้กระบวนการปรับปรุงภาพต้นฉบับไม่สามารถใช้รูปแบบใดรูปแบบหนึ่งได้กับทุกภาพ
2. การรู้จำข้อความมีข้อจำกัดทางภาษาที่จะทำการรู้จำ
3. วิธีที่ได้จากการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะยังมีข้อบกพร่องเนื่องมาจากข้อบกพร่องเดิมของวิธีการต้นแบบที่ได้ดัดแปลงมา ซึ่งยังไม่สามารถแก้ไขได้

5.3 ข้อจำกัดของแอปพลิเคชัน

1. รองรับภาษาที่จะทำการrecognition 21 ภาษา คือ Czech, Danish, Dutch, English, Finnish, French, German, Hungarian, Italian, Norwegian, Polish, Portuguese, Spanish, Swedish, Chinese Simplified, Greek, Japanese, Russian, Turkish, Chinese Traditional and Korean
2. การแปลภาษาต้องใช้งานอินเตอร์เน็ต
3. ความแม่นยำของการจดจำตัวอักษร ขึ้นอยู่กับคุณภาพของรูป หากรูปเบลอก็ทำให้ความแม่นยำลดลง, อักษรตัวเขียนจะส่งผลเมื่อขนาดฟอนต์น้อยกว่า 15 พิกเซลของภาษาตะวันตก หรือเล็กกว่า 20 พิกเซลของภาษาเอเชียกลาง, ภาพพื้นหลังที่ซับซ้อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นภาพที่ใกล้เคียงกับอักษร หรือทำการแยกระหว่างอักษรกับพื้นหลังได้ยาก,แสง-เงา ,มุมมองที่ทำให้เห็นไม่ชัด อักษรที่อยู่ต่ำกว่าตัวอื่นๆ เหนือตัวอื่นๆ หรือตัวที่ถูกขีดฆ่า เหล่านี้คือปัจจัยที่ทำให้ความแม่นยำในการรู้จำลดลง

5.4 ข้อเสนอแนะ

1. สามารถต่อยอดพัฒนาลงบนอุปกรณ์อื่น ๆ นอกจากอุปกรณ์พกพาได้
2. สามารถต่อยอดพัฒนาให้ทำงานแบบ Real Time ได้
3. สามารถเพิ่มประสิทธิภาพการทำงานของแอปพลิเคชันให้ดีขึ้นได้โดยการปรับปรุงกระบวนการประมวลผลเบื้องต้นเพิ่มเติม
4. กระบวนการที่ได้ศึกษาและพัฒนามุ่งเน้นที่ความรวดเร็วในการประมวลผลซึ่งโดยปกติจะทำให้ประสิทธิผลลดลง การพัฒนาเพิ่มเติมควรเพิ่มประสิทธิผลโดยให้กระทบกับความเร็วในการประมวลผลน้อยที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Chucai Yi and Yingli Tian, "**Scenes Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration,**" IEEE Transactions on Image Processing, Vol. 23, No. 7, July 2014.

[2] Derek Bradley and Gerhard Roth, "**Adaptive Thresholding Using the Integral,**" Journal of Graphics, GPU, and Game Tools Volume 12, Issue 2, 2007.

[3] J. Sauvola and M. PietikaKinen, "**Adaptive Document Image Binarization,**" Machine Vision and Media Processing Group, Infotech Oulu, University of Oulu, P.O. BOX 4500, FIN-90401 Oulu, Finland Received 29 April 1998; accepted 21 January 1999

[4] K. Peuwnuan, K. Woraratpanya, K. Pasupa, "**Modified Adaptive Thresholding Using Integral Image,**" unpublished.

[5] Pierre D. Wellner," **Adaptive Thresholding for the DigitalDesk,**" Technical Report EPC-1993-110.

[6] K. Woraratpanya, K. Pasupa, U. Suttapakti, P. Boonchukusol, T. Titijaroonroj, R. Hokking, et al., "**Text-background Decomposition for Thai Text Localization and Recognition in Natural Scenes,**" Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on 7-8 Oct. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

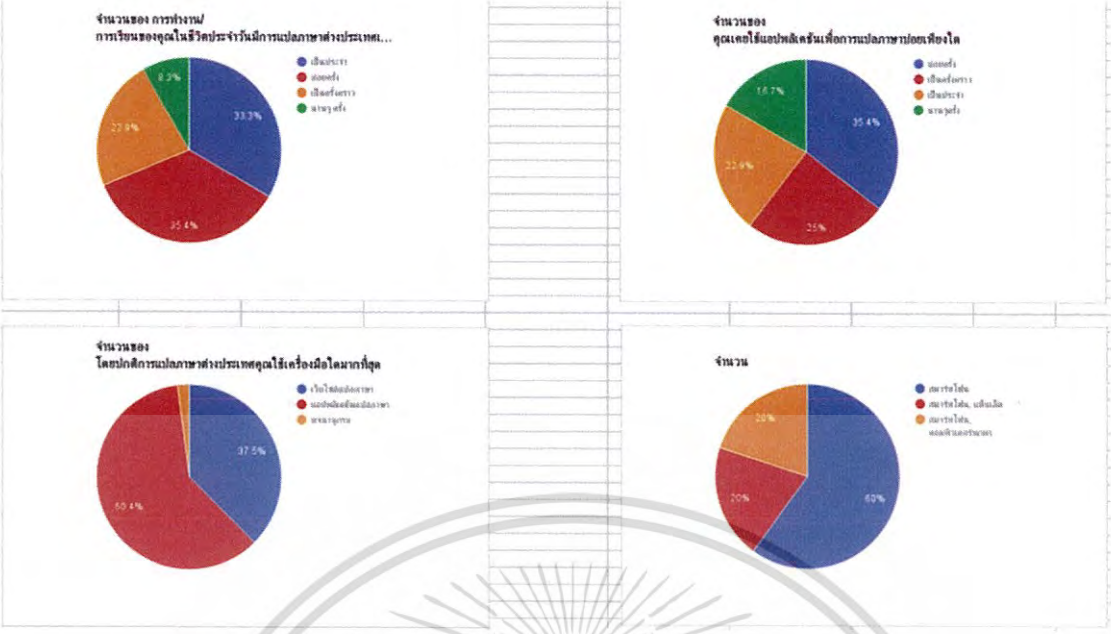
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก
ผลการสำรวจผู้ใช้งานแอปพลิเคชันเพื่อการแปลภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก ผลของการวิเคราะห์ข้อมูลแบบสอบถามออนไลน์

ประทับเวลา	การทำงานการเรียนของคุ	โดยปกติการแปลภาษาด้วย	คุณเคยใช้แอปพลิเคชันเพื่อ	คุณมีความคิดที่จะใช้แอปพ	เหตุผลที่คุณไม่เคยใช้แอป	แอปพลิเคชันเพื่อการแปล	การใช้แอปพลิเคชันเพื่อการแปลภาษาของคุณโดยทั่วไปใช้งานอย่างไร (ตอบได้หลายข้อ)
3/11/2015, 15:15:34	เป็นประจำ	เว็บไซต์แปลภาษา	บ่อยครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:15:44	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์ตั	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:16:05	เป็นครั้งคราว	เว็บไซต์แปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:16:12	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:17:05	เป็นครั้งคราว	เว็บไซต์แปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:17:10	เป็นครั้งคราว	แอปพลิเคชันแปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
3/11/2015, 15:19:05	เป็นครั้งคราว	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:24:44	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 15:30:58	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์ตั	พิมพ์ข้อความและดูคำแปล, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
3/11/2015, 16:02:43	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
3/11/2015, 16:19:06	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ
3/11/2015, 17:02:06	เป็นครั้งคราว	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์ตั	พิมพ์ข้อความและดูคำแปล
3/11/2015, 17:30:43	เป็นครั้งคราว	แอปพลิเคชันแปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ
3/11/2015, 17:36:23	นานๆ ครั้ง	แอปพลิเคชันแปลภาษา	นานๆครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 17:53:09	นานๆ ครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 18:42:40	บ่อยครั้ง	เว็บไซต์แปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:04:45	บ่อยครั้ง	พจนานุกรม	บ่อยครั้ง			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ, ตรวจแกรมมา, เช็คคำสะกดผิด
3/11/2015, 21:06:49	เป็นครั้งคราว	เว็บไซต์แปลภาษา	นานๆครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:06:50	บ่อยครั้ง	เว็บไซต์แปลภาษา	นานๆครั้ง			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:09:20	บ่อยครั้ง	เว็บไซต์แปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:10:29	เป็นครั้งคราว	เว็บไซต์แปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:11:28	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:20:06	เป็นประจำ	เว็บไซต์แปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล
3/11/2015, 21:58:56	บ่อยครั้ง	เว็บไซต์แปลภาษา	นานๆครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล, คัดลอกจากเว็บ วางในช่องเพื่อแปล
3/11/2015, 22:03:59	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:04:53	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:05:19	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:06:14	เป็นประจำ	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			แท็บเล็ต, คอมพิวเตอร์ตาพ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:16:33	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:19:34	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:21:09	เป็นประจำ	แอปพลิเคชันแปลภาษา	นานๆครั้ง			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:25:22	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:26:02	นานๆ ครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, แท็บเล็ต, คอม	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
3/11/2015, 22:28:41	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ
3/11/2015, 22:40:45	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:41:31	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:52:48	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 22:57:42	บ่อยครั้ง	เว็บไซต์แปลภาษา	นานๆครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 23:01:18	เป็นประจำ	เว็บไซต์แปลภาษา	เป็นครั้งคราว			คอมพิวเตอร์พกพา	พิมพ์ข้อความและดูคำแปล
3/11/2015, 23:05:32	นานๆ ครั้ง	เว็บไซต์แปลภาษา	นานๆครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล
3/11/2015, 23:46:05	บ่อยครั้ง	เว็บไซต์แปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล, นำเข้ารูปภาพเพื่อแปลข้อความที่พบ, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
3/11/2015, 23:49:17	บ่อยครั้ง	เว็บไซต์แปลภาษา	นานๆครั้ง			คอมพิวเตอร์ตั้งโต๊ะ	พิมพ์ข้อความและดูคำแปล
4/11/2015, 0:15:56	เป็นประจำ	เว็บไซต์แปลภาษา	เป็นครั้งคราว				
4/11/2015, 0:24:49	เป็นครั้งคราว	เว็บไซต์แปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
4/11/2015, 7:37:58	เป็นประจำ	แอปพลิเคชันแปลภาษา	เป็นประจำ			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล
4/11/2015, 8:18:43	เป็นครั้งคราว	เว็บไซต์แปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน, แท็บเล็ต	พิมพ์ข้อความและดูคำแปล
4/11/2015, 11:04:59	เป็นครั้งคราว	แอปพลิเคชันแปลภาษา	เป็นครั้งคราว			สมาร์ทโฟน	พิมพ์ข้อความและดูคำแปล, ใช้กล้องถ่ายภาพแล้วแปลข้อความแบบ real time
5/11/2015, 15:39:45	บ่อยครั้ง	แอปพลิเคชันแปลภาษา	บ่อยครั้ง			สมาร์ทโฟน, คอมพิวเตอร์พ	พิมพ์ข้อความและดูคำแปล



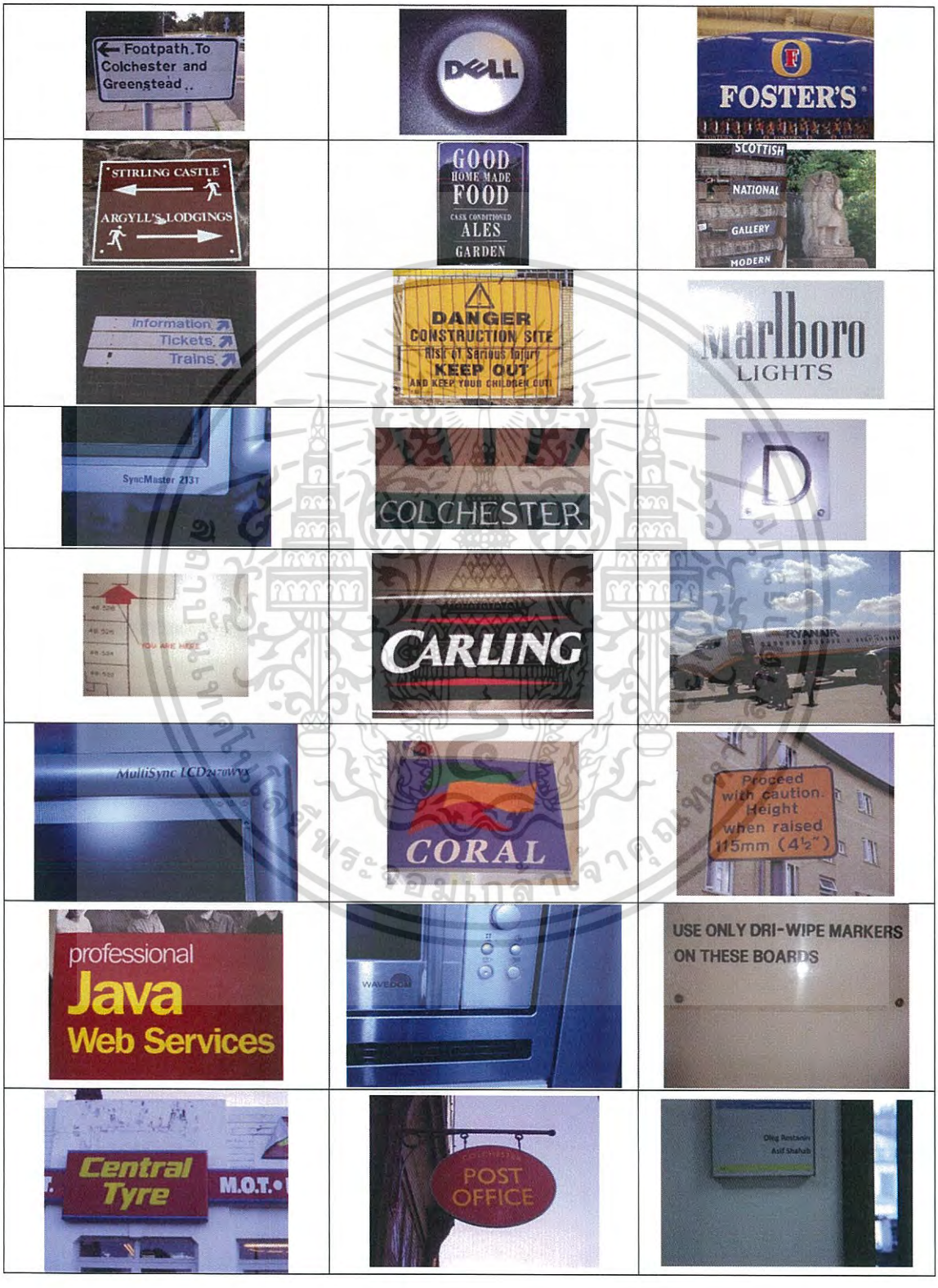
รูปที่ ก.2 แผนภูมิของผลของการสำรวจผู้ใช้งานแอปพลิเคชันเพื่อการแปลภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

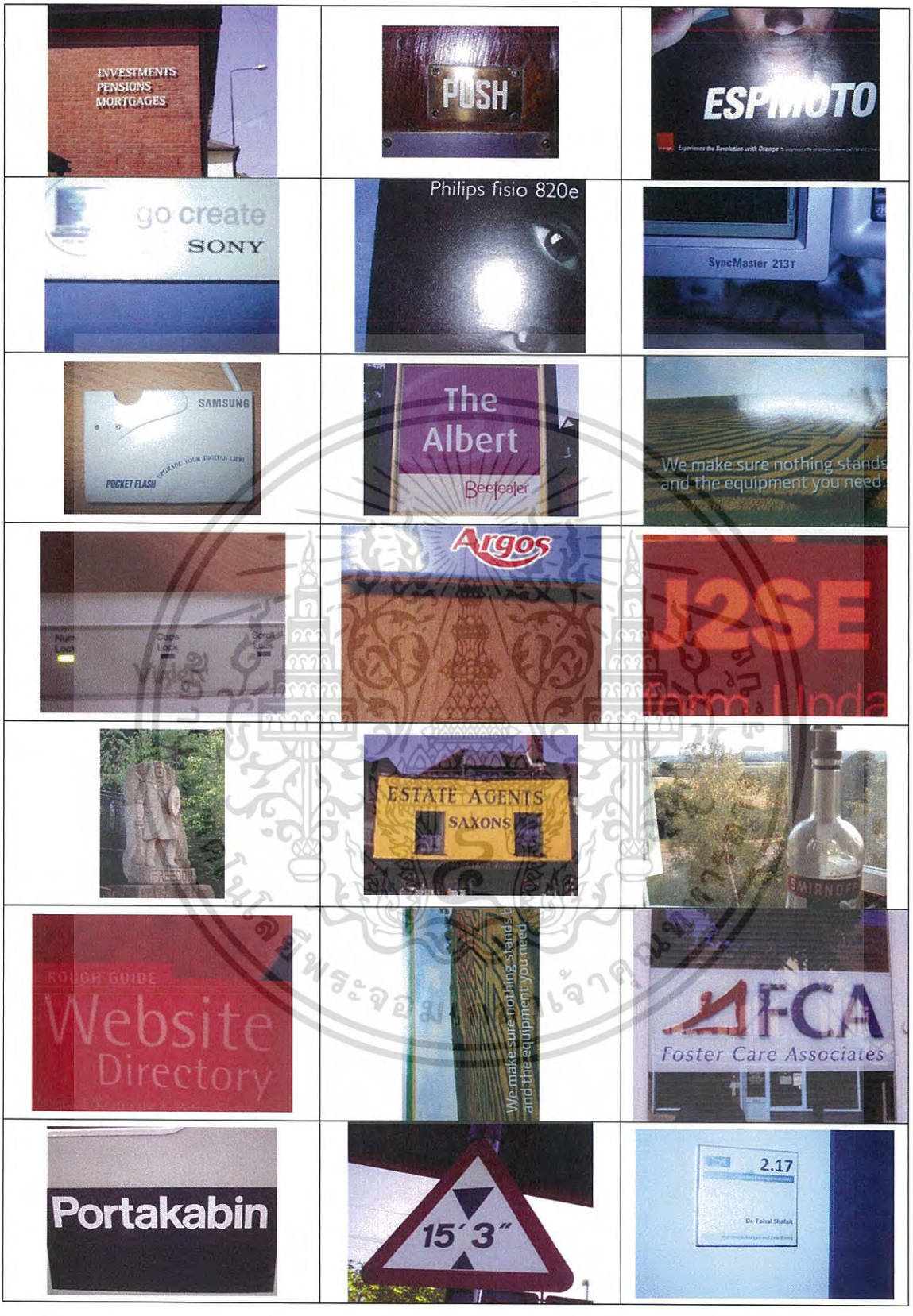
ภาคผนวก ข
รูปที่ใช้ในการทดลองแอปพลิเคชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

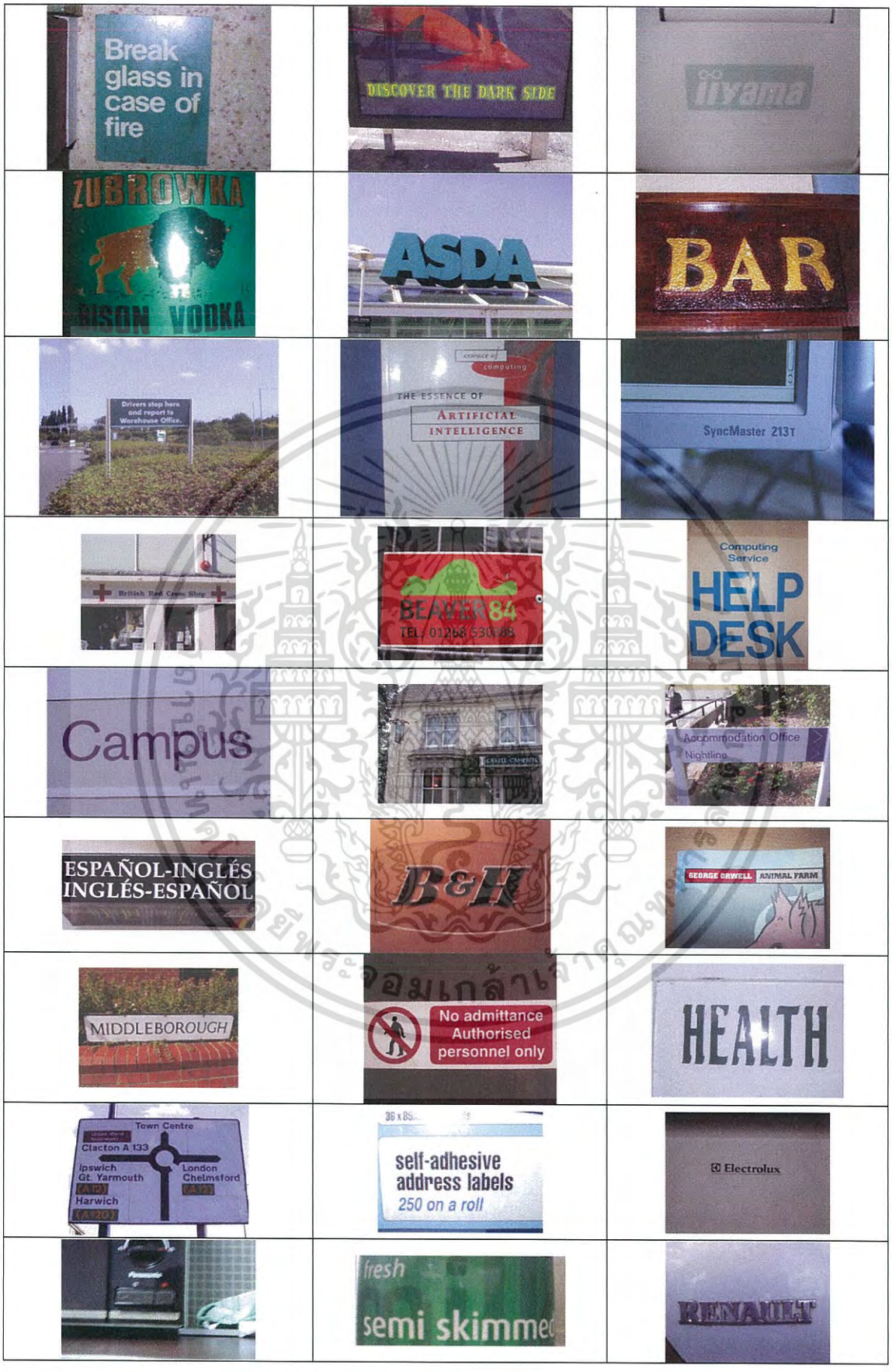
ตารางที่ ข.1 ตารางแสดงรูปภาพที่ใช้ในการทดลอง



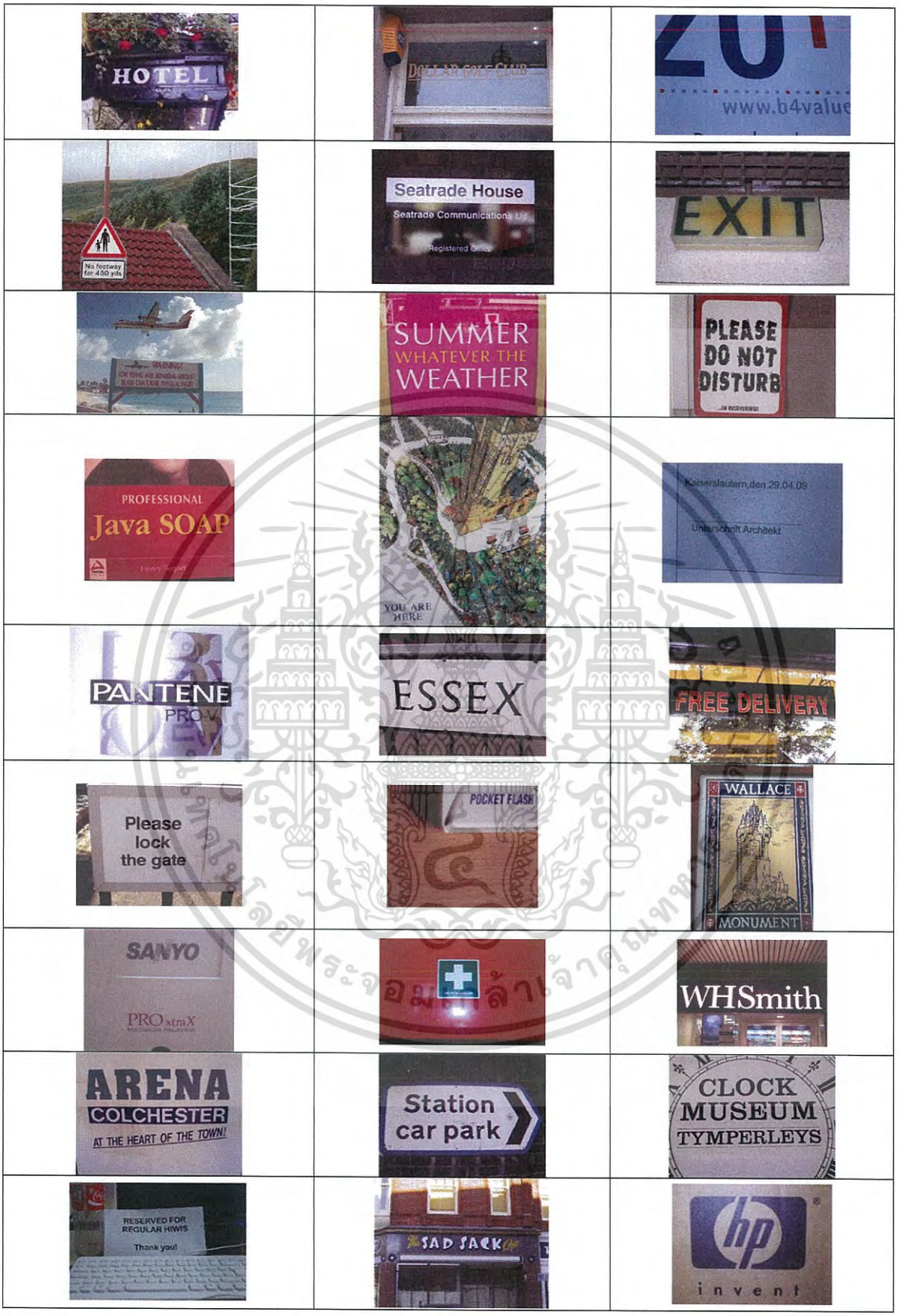
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

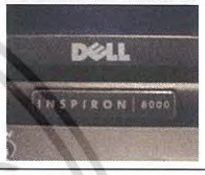


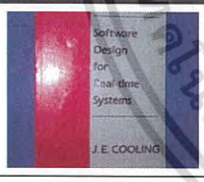
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



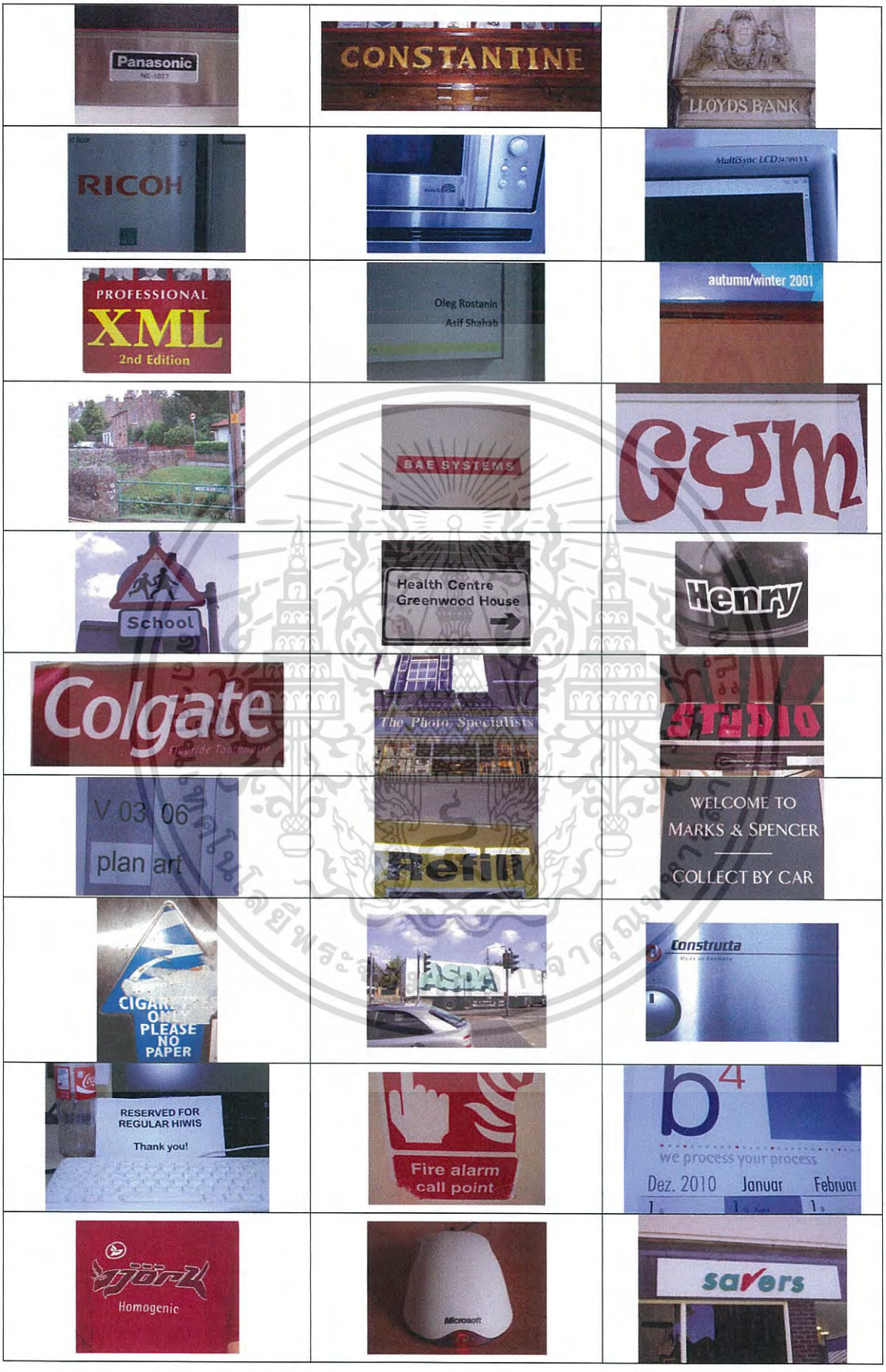
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

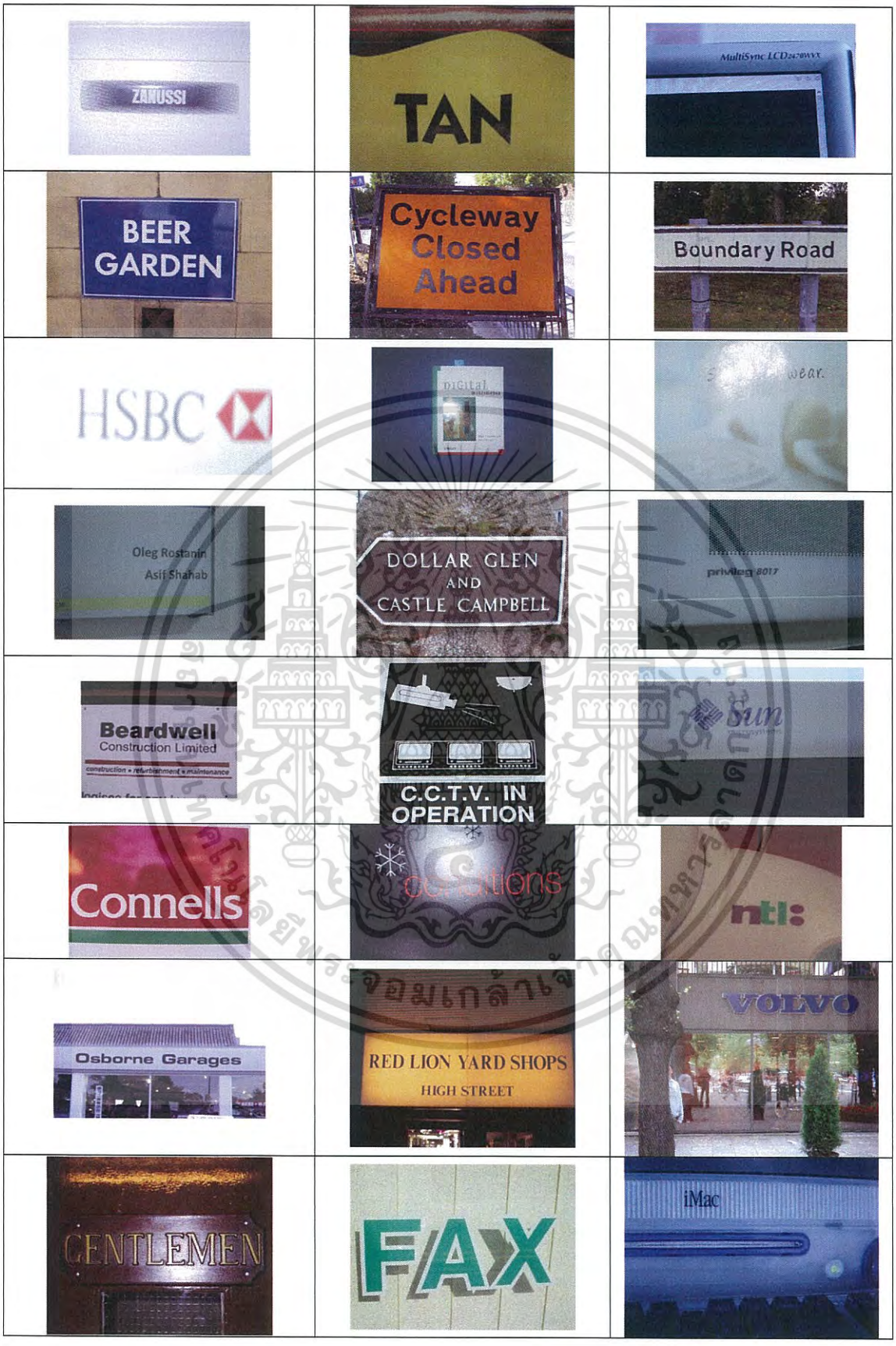




เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นายกฤตยชญ์ คงคติธรรม
วัน เดือน ปี เกิด	6 มิถุนายน 2537
ที่อยู่	22/129 หมู่6 หมู่บ้าน มัณฑนา ซอย ทวีวัฒนา59 ถนน ทวีวัฒนา ตำบล ศาลายา อำเภอ พุทธมณฑล จังหวัด นครปฐม 73170
ประวัติการศึกษา	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ชื่อ	นายกิตติภพ ผิวนวล
วัน เดือน ปี เกิด	12 ตุลาคม 2536
ที่อยู่	259/246 ซอยปรีดีพนมยงค์ 15 ถนนสุขุมวิท 71 แขวงพระโขนงเหนือ เขตวัฒนา กรุงเทพฯ 10110
ประวัติการศึกษา	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

กฤตยชญ์ คงคติธรรม[1] กิตติภพ ผิวนวล[2] กิตติภพ ผิวนวล[2]

ผศ.ดร. กิติ์สุชาต พสุภา[3] และ ผศ.ดร. กันต์พงษ์ วรรัตน์ปัญญา[4]

[1]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: k.kittayoch@gmail.com[1], plewpleng@hotmail.com[2],
kitsuchart@it.kmitl.ac.th[3], kuntpong@it.kmitl.ac.th[4]

บทคัดย่อ

หลังจากที่ประเทศไทยได้ก้าวเข้าสู่ประชาคมเศรษฐกิจอาเซียน การติดต่อสื่อสารกับประเทศเพื่อนบ้านและประเทศคู่ค้าในภูมิภาคใกล้เคียงทำให้เราได้รับข้อมูล ข่าวสารในภาษาต่างประเทศมากขึ้น นอกจากนี้ยังรวมไปถึงสื่อสิ่งพิมพ์ โฆษณา ประกาศ และอื่นๆ ที่ใช้ภาษาต่างประเทศ ซึ่งบางคนต้องการเครื่องมือสำหรับแปลข้อความเหล่านั้นเพื่อการรับรู้ข้อมูลที่ถูกต้อง

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตาม การพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และปัญหาจากแสงและเงาซึ่งเป็นอุปสรรคต่อการรู้จำข้อความโครงงานชุดนี้จึงพัฒนาแอปพลิเคชันโดยศึกษาและพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ แอปพลิเคชันนี้ไม่เพียงแต่ใช้การรู้จำเพื่อนำไปแปลความหมายเท่านั้น แต่ยังออกแบบให้สามารถนำข้อความที่รู้จำได้ไปใช้ต่อในอุปกรณ์พกพาของผู้ใช้

คำสำคัญ – การประมวลผลเบื้องต้น; การตรวจจับข้อความ; การรู้จำข้อความ; การแปลภาษา

1. บทนำ

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตามการพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และโดยเฉพาะปัญหาแสงเงาที่กระทบต่อการรู้จำข้อความ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. การรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OCR หรือ Optical Character Recognition เป็นกระบวนการแปลงข้อมูลภาพของข้อความทั้งตัวพิมพ์และตัวเขียน (ลายมือ) ไปสู่ข้อความในแบบที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้ ซึ่งก็คือการแปลงข้อมูลในเอกสาร สื่อสิ่งพิมพ์ต่าง ๆ ให้อยู่ในรูปแบบข้อมูลดิจิทัลที่สามารถนำไปแก้ไข ค้นหา จัดหมวดหมู่ จัดเก็บ แสดงผล และประมวลผลได้ เช่น การแปลภาษา การสังเคราะห์เสียงพูดจากข้อความ การจัดทำฐานข้อมูล ฯลฯ ทั้งนี้ OCR มีความเกี่ยวข้องกับทั้งการศึกษาการรู้จำรูปแบบ (Pattern Recognition) ปัญญาประดิษฐ์ (Artificial Intelligence) และคอมพิวเตอร์วิทัศน์ (Computer Vision)

ขั้นตอนการทำงานของ OCR มี 3 ขั้นตอนหลัก

รูปที่ 1. ขั้นตอนการทำงานของระบบ OCR

2.1.1 การประมวลผลเบื้องต้น (Pre-Processing)
กระบวนการเบื้องต้นก่อนที่จะนำรูปภาพไปประมวลผล คือ การนำรูปภาพมาปรับแต่งและจัดเตรียมให้เหมาะสมก่อนจะถูกนำไปประมวลผล ซึ่งหากเกิดความผิดพลาดจะมีผลกระทบต่อขั้นตอนการประมวลผลต่อไป การปรับแต่งภาพต้นฉบับมีหลายวิธีการ เช่น การปรับองศาของรูปภาพ การเบลอหรือทำให้ภาพคมชัดขึ้น การแปลงภาพเป็นภาพขาวดำ การลดทอนแสงเงาในภาพ ฯลฯ

2.1.2 การรู้จำข้อความ (Text recognition)
การรู้จำข้อความจากฉากธรรมชาติ (Text Recognition in Natural Scenes) เป็นกระบวนการประมวลผลข้อมูลภาพถ่ายที่มีข้อความปรากฏอยู่ เพื่อสกัดข้อความนั้น ๆ ออกมาเป็นข้อความในรูปแบบดิจิทัลที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้กระบวนการที่สำคัญในการรู้จำข้อความจากฉากธรรมชาติได้แก่ การดำเนินการเบื้องต้น การตรวจจับข้อความ การรู้จำข้อความ และการดำเนินการหลังการประมวลผล

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)
ความแม่นยำของกระบวนการ OCR สามารถเพิ่มขึ้นได้ถ้าหาก ผลลัพธ์ถูกควบคุมโดย รายการของศัพท์เฉพาะที่อนุญาตให้ใช้ในเอกสาร (lexicon) ซึ่งเป็นรายการของคำที่สามารถปรากฏในเอกสารได้ ตัวอย่างเช่น คำในภาษาอังกฤษ ทั้งหมด และศัพท์เทคนิคเฉพาะซึ่งเทคนิคนี้อาจเป็นปัญหาได้ถ้าในเอกสารมีข้อความที่ไม่ได้อยู่ใน lexicon นอกจากนี้ยังสามารถใช้วิธีการ Near-neighbor analysis แก้ไขคำผิดได้โดยการเปรียบเทียบกับตัวอักษรที่มักจะปรากฏในข้อความเดียวกัน และไวยากรณ์ภาษายังช่วยในการระบุคำที่พบว่าควรจะเป็นคำนาม คำกริยา หรืออื่น ๆ เป็นต้น

2.2 การแก้ไขปัญหาสำหรับการรู้จำข้อความ
Chucai Yi และYingli Tian [1] กล่าวไว้ว่า สำหรับการสกัดข้อความจากฉากธรรมชาติด้วยอุปกรณ์พกพานั้น การตรวจจับข้อความอย่างอัตโนมัติและมีประสิทธิผล และอัลกอริทึมสำหรับการรู้จำถือเป็นสิ่งจำเป็นอย่างยิ่ง อย่างไรก็ตามการสกัดข้อความจากภาพฉากธรรมชาติมีอุปสรรคท้าทายหลายประการเช่น ความหลายหลายของฉากหลังที่มีสิ่งรบกวนและสิ่งที่ไม่ใช่ข้อความปะปนอยู่ และความหลากหลายของรูปแบบตัวอักษรเช่น ประเภทตัวอักษร แบบตัวหนังสือ และขนาด ดังนั้นการแก้ไขปัญหาสำหรับการรู้จำข้อความจึงเป็นความพยายามในการทำให้ความหลากหลายของฉากหลังและความหลากหลายของรูปแบบตัวอักษรมีผลต่อกระบวนการต่าง ๆ เช่น การตรวจจับข้อความ และการรู้จำข้อความให้น้อยที่สุด เพื่อให้ผลลัพธ์มีความถูกต้องและแม่นยำมากขึ้น สอดคล้องกับที่ Woraratpanya, K และคณะ [6] กล่าวไว้ว่า อย่างไรก็ตามประสิทธิผลของอัตราการรู้จำขึ้นอยู่กับการระบุตำแหน่งของข้อความ ความบริสุทธิ์ของกระบวนการแยกข้อความออกจากฉากหลังที่มากจะนำไปสู่อัตราความแม่นยำของการรู้จำที่มากขึ้น
ในปี ค.ศ. 1993 Pierre D. Wellner [5] ได้นำเสนอวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) โดยการพิจารณาภาพโทนสีเทาจากภาพต้นฉบับ จากนั้นเปรียบเทียบข้อมูลค่าความสว่างในแต่ละตำแหน่งของรูปภาพกับค่าเฉลี่ยความสว่างของพื้นที่ย่อยที่ตำแหน่งนั้น ๆ ปรากฏอยู่ เพื่อใช้เป็นเกณฑ์กำหนดสำหรับการสร้างภาพขาวดำ วิธีการนี้สามารถนำมาใช้กับภาพถ่ายของเอกสารที่มีเงาตกกระทบเพื่อแก้ปัญหาข้อความที่ถูกเงาบังและไม่สามารถตรวจจับหรือรู้จำได้ และยังสามารถประมวลผลได้อย่างรวดเร็ว การพิจารณาพื้นที่ย่อยในวิธีการของ Wellner จะเริ่มจากการพิจารณารูปภาพแต่ละรูปเป็นแถวของพิกเซลที่เรียงตัวเป็นแถวเดียว ความกว้างของพื้นที่ย่อย (s) แต่ละช่วงจะมีขนาดเป็น 1 ใน 8 ของความกว้างภาพต้นฉบับ จากนั้นพิจารณาผลรวมของค่าความสว่างจากภาพโทนสีเทาของภาพต้นฉบับภายในแต่ละพื้นที่ย่อยได้ด้วยสมการต่อไปนี้

$$f_s(n) = \sum_{i=0}^{s-1} P_{n-1} \quad (1)$$

เมื่อ $f_s(n)$ เป็นผลรวมของค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n P_n คือค่าความสว่างของพิกเซล ณ ตำแหน่ง n ของภาพโทนสี เทาจากภาพต้นฉบับ และ s คือความกว้างของพื้นที่ย่อย

รูปที่ 2. ภาพการพิจารณาพื้นที่ย่อยในรูปภาพ

ภาพผลลัพธ์จากการประมวลผลคำนวณได้จากการกำหนดค่าในแต่ละพิกเซลของภาพผลลัพธ์ด้วยเงื่อนไขต่อไปนี้

$$T(n) = \begin{cases} 0, \text{ เมื่อ } P_n < \left(\frac{f_s(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ 1, \text{ กรณีอื่น ๆ} \end{cases} \quad (2)$$

เมื่อ t เป็นเปอร์เซ็นต์ที่ค่าความสว่าง ณ ตำแหน่ง n น้อยกว่าค่าเฉลี่ยของพื้นที่ย่อยที่พิจารณา (Wellner เสนอให้ค่า t เป็น15) โดยที่เลข 0 แทนสีดำ และเลข 1 แทนสีขาว ซึ่งผลลัพธ์ที่ได้จะเป็นภาพขาวดำ (Binarization Image)

วิธีการของ Wellner เป็นเพียงรูปแบบหนึ่งของวิธีการกำหนดเกณฑ์แบบปรับเหมาะสำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ ยังมีวิธีการที่แตกต่างออกไปอีกมากมาย เช่น ในปี ค.ศ. 2007 Derek Bradley และคณะ [2] ได้นำเสนอการปรับปรุงวิธีการของ Wellner เพื่อเพิ่มประสิทธิภาพในการแก้ปัญหาในบทความ Adaptive Thresholding Using Integral Image โดยปรับเปลี่ยนวิธีการพิจารณาพื้นที่ย่อยในขั้นตอนการเปรียบเทียบค่าความสว่างของแต่ละตำแหน่งในรูปภาพ

นอกเหนือจากการใช้วิธีการใดวิธีการหนึ่งในการประมวลผลแต่ละรูปภาพ แนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการแสดงให้เห็นผลลัพธ์ที่ดีกว่าดังที่ J. Sauvola และคณะ [3] ได้นำเสนอไว้ในบทความในปี ค.ศ. 2000 โดยดำเนินการประมวลผลจำแนกข้อความและพื้นหลังจากภาพถ่ายจากนั้นได้แบ่งวิธีการประมวลผลออกเป็นสองวิธี วิธีการหนึ่งเป็นการแยกองค์ประกอบพื้นหลัง และอีกวิธีการหนึ่งเป็นการแปลงภาพเป็นภาพขาวดำเพื่อการสกัดข้อความออกจากพื้นหลังภายใต้ปัญหาที่ส่งผลต่อการตรวจจับข้อความ

3. การดำเนินงานวิจัย

3.1. การศึกษากระบวนการรู้จำข้อความ

สำหรับการศึกษากระบวนการรู้จำข้อความ (Text Recognition) ทำโดยการค้นคว้าจากทฤษฎีทางด้านการประมวลผลรูปภาพ (Image Processing) และการค้นคว้าจากบทความทางวิชาการที่เกี่ยวข้องกับการรู้จำข้อความ, การแก้ปัญหาด้านต่าง ๆ ที่เป็นอุปสรรคต่อการตรวจจับข้อความ และการรู้จำข้อความ รวมถึงบทความที่เกี่ยวข้องกับการรู้จำข้อความบนอุปกรณ์พกพา (Mobile Devices)

3.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์เคลื่อนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา ได้ศึกษาการพัฒนาแอปพลิเคชันบน **Windows Platform** และได้ทดลองพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย เพื่อทำการศึกษาวิธีการพัฒนาการรู้จำข้อความบนอุปกรณ์พกพาและเรียนรู้ปัญหาหรืออุปสรรคที่เกิดขึ้น

3.3 การศึกษาความต้องการต่อการใช้แอปพลิเคชัน

สำหรับการศึกษาความต้องการต่อการใช้แอปพลิเคชันเป็นการสำรวจความต้องการของกลุ่มผู้ใช้เป้าหมายเพื่อเป็นตัวกำหนดคุณลักษณะ (**Feature**) สำหรับแอปพลิเคชันที่พัฒนา

3.4 การศึกษาการปรับปรุงภาพต้นฉบับ

การศึกษาการปรับปรุงภาพต้นฉบับต้องการแก้ปัญหาการตรวจจับข้อความในฉากธรรมชาติดังต่อไปนี้ คือ ปัญหาด้านเงา และ ปัญหาด้านแสงสว่าง

3.4.1 ขั้นตอนในการศึกษาการปรับปรุงภาพต้นฉบับ

ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา การปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชัน และการทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น

3.4.2 นิยามลักษณะปัญหา

3.4.2.1 ปัญหาด้านเงา

ภาพที่มีเงาบดบังบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีเงาบดบังบริเวณที่มีข้อความทั้งหมด

3.4.2.2. ปัญหาด้านแสงสว่าง

ภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบริเวณที่มีข้อความทั้งหมด

รูปที่ 3. ตัวอย่างภาพที่มีปัญหาด้านเงา

รูปที่ 4. ตัวอย่างภาพที่มีปัญหาแสงสว่าง

3.4.3 การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะ

จากการศึกษากระบวนการแก้ปัญหาของ Pierre D. Wellner [5] ดังที่ได้อธิบายวิธีการในหัวข้อ 2.2 พบว่าสามารถประมวลผลได้อย่างรวดเร็วเนื่องจากกระบวนการมีความซับซ้อนที่น้อย ทว่ามีข้อจำกัด ข้อบกพร่อง และข้อสังเกตในการประมวลผล เช่น ให้ผลลัพธ์ที่ดีเฉพาะกับภาพที่มีปัญหาเงาในระดับปานกลางถึงรุนแรง ให้ผลลัพธ์ที่ไม่ดีกับกรณีทั่วไป ได้แก่ ปัญหาแสง ปัญหาอื่น ๆ ที่นอกเหนือจากนิยาม และกรณีภาพที่ไม่มีปัญหา นอกจากนี้ยังพบว่าหากใช้ภาพส่วนกลับโทนสีเทา (Complement-grayscale image) ของภาพต้นฉบับในการดำเนินการจะสามารถให้ผลลัพธ์ที่ดีกับภาพที่มีปัญหาแสงในระดับปานกลางถึงรุนแรง

จากแนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการดังที่ J. Sauvola และคณะ [3] ได้ดำเนินการศึกษาและพัฒนากระบวนการแก้ปัญหาต่อการรู้จำข้อความทำให้เชื่อว่าอาจมีวิธีการดัดแปลงให้การแก้ปัญหาของ Wellner สามารถแก้ไขปัญหาแสงได้โดยการประยุกต์ใช้วิธีการร่วม สำหรับการประยุกต์ใช้วิธีการร่วมเพื่อแก้ปัญหาของรูปภาพนั้น K. Peuwnuan และคณะ [4] ได้นำเสนอวิธีการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะโดยใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพรวมผล (Integral Image) สำหรับการตรวจจับข้อความ ซึ่งเป็นการดัดแปลงกระบวนการของ Derek Bradley และคณะ [2] โดยมีการแบ่งพื้นที่รูปภาพออกเป็นส่วนต่าง ๆ (Image Segmentation) คือพื้นที่ที่มีค่าความสว่างสูง และพื้นที่ที่มีค่าความสว่างปานกลางถึงต่ำ จากนั้นจึงดำเนินการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในแต่ละพื้นที่ด้วยวิธีการที่แตกต่างกัน โดยอาศัยแนวคิดข้างต้น ในการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพาจึงได้ดัดแปลงวิธีการของ Wellner โดย ก่อนดำเนินการกำหนดเกณฑ์แบบปรับเหมาะจำเป็นต้องมีการจำแนกพื้นที่ออกเป็นพื้นที่ที่มีค่าความสว่างสูงและค่าความสว่างต่ำ เนื่องจากพื้นที่ที่จำแนกได้นั้นจะใช้วิธีการประมวลผลเพื่อการแปลงภาพขาวดำที่แตกต่างกันโดยเงื่อนไขที่ใช้จำแนกพื้นที่ของรูปภาพแสดงโดยสมการที่ (3)

$$P(n) = \begin{cases} \text{พื้นที่ความสว่างสูง, เมื่อ } P_n > M + SD \\ \text{พื้นที่ความสว่างต่ำ, กรณีอื่น ๆ} \end{cases} \quad (3)$$

เมื่อ P_n คือค่าความสว่างของพิกเซลที่พิจารณาจากภาพโทนสีเทา M คือค่าเฉลี่ยความสว่างของทั้งรูปภาพ *และ* SD คือส่วนเบี่ยงเบนมาตรฐานของค่าความสว่างทั้งรูปภาพ

หลังจากแบ่งพื้นที่รูปภาพตามเงื่อนไขในสมการ (3) แล้วจึงดำเนินการกำหนดเกณฑ์แบบปรับเหมาะสำหรับแต่ละพื้นที่ซึ่งจะมีความแตกต่างกันดังนี้

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P'_n \geq \left(\frac{f'_S(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (4)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P'_n คือค่าความสว่างของพิกเซลที่ n ของภาพส่วนกลับโทนสีเทา และ $f'_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพส่วนกลับโทนสีเทา สำหรับสมการ (4) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างสูงตามที่จำแนกได้และ

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P_n \geq \left(\frac{f_S(n)}{s}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (5)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P_n คือค่าความสว่างของพิกเซลที่ n ของภาพโทนสีเทา และ $f_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพโทนสีเทา สำหรับสมการ (5) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างปานกลางถึงต่ำตามที่จำแนกได้

3.5 ผลการดำเนินงาน

3.5.1 การศึกษากระบวนการรู้จำข้อความ

ผลการศึกษาฯ ทำให้มีความรู้ความเข้าใจในกระบวนการรู้จำข้อความมากยิ่งขึ้น เพียงพอต่อการนำไปใช้เพื่อพัฒนาแอปพลิเคชันต่อไป รวมถึงเข้าใจในปัญหาและอุปสรรคต่อการตรวจจับและการรู้จำข้อความ

3.5.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

ผลการศึกษาและทดลองฯ ทำให้สามารถพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย และได้ออกแบบการทดลองและบันทึกผลการรู้จำข้อความจากภาพตัวอย่างที่มีความหลากหลายขององค์ประกอบในภาพเพื่อคำนวณความแม่นยำของแอปพลิเคชันก่อนจะพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ

3.5.3 การศึกษาการปรับปรุงภาพต้นฉบับ

3.5.3.1. ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ

จากการศึกษาพบว่าการปรับปรุงภาพต้นฉบับด้วยวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding) สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) สามารถประมวลผลได้เร็วเหมาะแก่การพัฒนาบนอุปกรณ์พกพาโดยเฉพาะวิธีการของ Wellner ซึ่งเหมาะอย่างยิ่งสำหรับกรณีภาพที่มีปัญหาเงาปานกลางถึงรุนแรง และสามารถดัดแปลงเพื่อให้ใช้แก้ปัญหาแสงได้เช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.3.2. ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา มีการทดลองใช้วิธีการของ Wellner เพื่อดูผลในการแก้ปัญหาเงาที่มีผลต่อการตรวจจับข้อความ พบว่าสามารถแก้ปัญหาเงาระดับปานกลางถึงสูงได้ดี แต่จะส่งผลลบกับการประมวลผลหากนำไปใช้กับภาพที่มีปัญหาแสงหรือปัญหาอื่น ๆ กล่าวคือจะลดความแม่นยำ ความถูกต้อง และความครบถ้วนของผลลัพธ์ลงอย่างมาก แต่สามารถปรับให้ใช้แก้ไขปัญหาแสงได้หากใช้ภาพส่วนกลับของภาพโทนสีเทา (Complement-Grayscale Image) ในการดำเนินการ

3.5.3.3. ปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชันอยู่ มีการปรับใช้วิธีการของ Wellner เพื่อให้สามารถแก้ไขปัญหาของภาพได้ทั้งเงาและแสงโดยการสร้างเงื่อนไขเพื่อพิจารณาพื้นที่ของรูปภาพที่เป็นส่วนความสว่างต่ำถึงปานกลาง และความสว่างสูง และใช้วิธีการดำเนินการในแต่ละพื้นที่ที่แตกต่างกันซึ่งให้ผลลัพธ์ที่ดีขึ้นสำหรับกรณีปัญหาแสงที่เกิดกับรูปภาพ

3.5.3.4. ทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงาแต่มีปัญหาในกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้ซึ่งตัวเลขผลการทดลอง การเปรียบเทียบและการอภิปรายจะได้กล่าวถึงในรายละเอียด ในหัวข้อ 4.2

บทที่ 4 กระบวนการทำงานของระบบ

4.1 แอปพลิเคชัน

การพัฒนาแอปพลิเคชันได้พัฒนาบนอุปกรณ์สองประเภทคือบนสมาร์ทโฟน และบนแทบเล็ต สำหรับตัวอย่างหน้าจอการทำงานภายในแอปพลิเคชันมีดังต่อไปนี้

รูปที่ 5. หน้าจอแอพลิเคชันบนสมาร์ทโฟน

รูปที่ 6. หน้าจอแอพลิเคชันบนแทบเล็ต

4.2 การทดลองและการสรุปผล

การทดลองมีจุดประสงค์เพื่อเปรียบเทียบผลลัพธ์จากการทำงานของแอพพลิเคชั่นก่อนและหลังการเพิ่มเติมการประมวลผลเบื้องต้น โดยชุดทดสอบที่ใช้เป็นภาพฉากธรรมชาติจากการแข่งขันของ ICDAR 2013 จำนวน 229 ภาพ ซึ่งเป็นภาพที่มีปัญหาแสงและ/หรือเงาเพียง 59 ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 ตัวอย่างรูปที่ใช้ในการทดลอง

รูปที่ใช้ในการทดลองจำแนกออกตามนิยามปัญหาได้ ในการทดลองจำแนกออกตามนิยามปัญหาได้ 2 ประเภทคือ แสงและเงา และกรณีอื่น ๆ เพิ่มเข้ามาด้วย

ตารางที่ 1. การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชัน(โดยเฉลี่ย)

Shadow	Lighting	Others

ตารางที่ 2. การจำแนกปัญหาของที่ปรากฏในแต่ละรูปภาพที่เป็นชุดทดสอบ

การประมวลผล	ตัวชี้วัดการประเมินผล		
	precision	recall	correctness
ก่อน	82.66%	74.44%	82.18%
หลัง	40.55%	30.64%	39.99%

ตารางที่ 3. ตารางแสดงรูปภาพกรณีที่ Recall สูงขึ้น

1.	3.

2.	4.

4.2.2 สรุปผลการทดลอง

ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงามีจุดบกพร่องกับกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้

ตารางที่ 4. ตารางแสดงผลลัพธ์กรณีที่ Recall สูงขึ้น

รูป		ตัวชี้วัดการประเมินผล		
		Precision (%)	Recall (%)	Correctness (%)
1	ก่อน	50	13.04	100
	หลัง	92	100	100
2	ก่อน	100	60	66.67
	หลัง	71.43	100	100
3	ก่อน	100	52.63	100
	หลัง	100	100	100
4	ก่อน	100	44.90	100
	หลัง	96.97	65.31	93.94

บทที่ 5 สรุปผลโครงงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

การศึกษาการระบบการรู้จำข้อความจากฉากธรรมชาติ ได้มีการนำไลบรารี OCR เข้ามาใช้แทนระบบการรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้การประมวลผลบนอุปกรณ์พกพาประเภทสมาร์ทโฟนทำได้รวดเร็วยิ่งขึ้นและใช้ทรัพยากรที่มีจำกัดของอุปกณ์ได้อย่างมีประสิทธิภาพ

การศึกษาภาพการปรับปรุงต้นฉบับได้ใช้หลักการการประมวลผลรูปภาพ (Image Processing) เพื่อศึกษาทดลองวิธีการปรับปรุงภาพต้นฉบับก่อนเข้าสู่กระบวนการรู้จำข้อความ พบว่าการปรับปรุงภาพต้นฉบับต้องพัฒนาอัลกอริทึมเพิ่มเติมที่สามารถระบุวิธีการปรับปรุงภาพต้นฉบับที่จำเพาะกับภาพนั้นๆ เนื่องจากภาพถ่ายจากฉากธรรมชาติมีความหลากหลายมากมาย ทำให้กระบวนการปรับปรุงภาพต้นฉบับไม่สามารถใช้รูปแบบใดรูปแบบหนึ่งได้กับทุกภาพ

ข้อเสนอแนะ

1. สามารถต่อยอดพัฒนาลงบนอุปกรณ์อื่น ๆ นอกจากอุปกรณ์พกพาได้
2. สามารถต่อยอดพัฒนาให้ทำงานแบบ Real Time ได้
3. สามารถเพิ่มประสิทธิภาพการทำงานของแอปพลิเคชันให้ดีขึ้นได้โดยการปรับปรุงกระบวนการประมวลผลเบื้องต้นเพิ่มเติม

กระบวนการที่ได้ศึกษาและพัฒนามุ่งเน้นที่ความรวดเร็วในการประมวลผลซึ่งโดยปกติจะทำให้ประสิทธิผลลดลง การพัฒนาเพิ่มเติมควรเพิ่มประสิทธิผลโดยให้กระทบกับความเร็วในการประมวลผลน้อยที่สุด

เอกสารอ้างอิง

[1] Chucai Yi and Yingli Tian, "**Scenes Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration**", IEEE Transactions on Image Processing, Vol. 23, No. 7, July 2014.

[2] Derek Bradley and Gerhard Roth, "**Adaptive Thresholding Using the Integral**", Journal of Graphics, GPU, and Game Tools Volume 12, Issue 2, 2007.

[3] J. Sauvola and M. PietikaKinen, "**Adaptive Document Image Binarization**", Machine Vision and Media Processing Group, Infotech Oulu, University of Oulu, P.O. BOX 4500, FIN-90401 Oulu, Finland Received 29 April 1998; accepted 21 January 1999

[4] K. Peuwnuan, K. Woraratpanya, K. Pasupa, "**Modified Adaptive Thresholding Using Integral Image**", (unpublished).

[5] Pierre D. Wellner," **Adaptive Thresholding for the DigitalDesk**", Technical Report EPC-1993-110.

[6] Woraratpanya, K.Fac. of Inf. Technol., King Mongkut's Inst. of Technol. Ladkrabang, Bangkok, Thailand Pasupa, K. ; Suttapakti, U. ; Boonchukusol, P. ; Titijaroonroj, T. ; Hokking, R. ; Kuroki, Y. ; Kato, Y. , "**Text-background Decomposition for Thai Text Localization and Recognition in Natural Scenes**", Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on 7-8 Oct. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

กฤตยชญ์ คงคติธรรม[1] กิตติภพ ผิวนวล[2] กิตติภพ ผิวนวล[2]

ผศ.ดร. กิติ์สุชาต พสุภา[3] และ ผศ.ดร. กันต์พงษ์ วรรัตน์ปัญญา[4]

[1]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: k.kittayoch@gmail.com[1], plewplerng@hotmail.com[2],
kitsuchart@it.kmitl.ac.th[3], kuntpong@it.kmitl.ac.th[4]

บทคัดย่อ

หลังจากที่ประเทศไทยได้ก้าวเข้าสู่ประชาคมเศรษฐกิจอาเซียน การติดต่อสื่อสารกับประเทศเพื่อนบ้านและประเทศคู่ค้าในภูมิภาคใกล้เคียงทำให้เราได้รับข้อมูล ข่าวสารในภาษาต่างประเทศมากขึ้น นอกจากนี้ยังรวมไปถึงสื่อสิ่งพิมพ์ โฆษณา ประกาศ และอื่นๆ ที่ใช้ภาษาต่างประเทศ ซึ่งบางคนต้องการเครื่องมือสำหรับแปลข้อความเหล่านั้นเพื่อการรับรู้ข้อมูลที่ถูกต้อง

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตาม การพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และปัญหาจากแสงและเงาซึ่งเป็นอุปสรรคต่อการรู้จำข้อความโครงงานชุดนี้จึงพัฒนาแอปพลิเคชันโดยศึกษาและพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ แอปพลิเคชันนี้ไม่เพียงแต่ใช้การรู้จำเพื่อนำไปแปลความหมายเท่านั้น แต่ยังออกแบบให้สามารถนำข้อความที่รู้จำได้ไปใช้ต่อในอุปกรณ์พกพาของผู้ใช้

คำสำคัญ – การประมวลผลเบื้องต้น; การตรวจจับข้อความ; การรู้จำข้อความ; การแปลภาษา

1. บทนำ

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตามการพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และโดยเฉพาะปัญหาแสงเงาที่กระทบต่อการรู้จำข้อความ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. การรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OCR หรือ Optical Character Recognition เป็นกระบวนการแปลงข้อมูลภาพของข้อความทั้งตัวพิมพ์และตัวเขียน (ลายมือ) ไปสู่ข้อความในแบบที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้ ซึ่งก็คือการแปลงข้อมูลในเอกสาร สื่อสิ่งพิมพ์ต่าง ๆ ให้อยู่ในรูปแบบข้อมูลดิจิทัลที่สามารถนำไปแก้ไข ค้นหา จัดหมวดหมู่ จัดเก็บ แสดงผล และประมวลผลได้ เช่น การแปลภาษา การสังเคราะห์เสียงพูดจากข้อความ การจัดทำฐานข้อมูล ฯลฯ ทั้งนี้ OCR มีความเกี่ยวข้องกับทั้งการศึกษาการรู้จำรูปแบบ (Pattern Recognition) ปัญญาประดิษฐ์ (Artificial Intelligence) และคอมพิวเตอร์วิทัศน์ (Computer Vision)

ขั้นตอนการทำงานของ OCR มี 3 ขั้นตอนหลัก

รูปที่ 1. ขั้นตอนการทำงานของระบบ OCR

2.1.1 การประมวลผลเบื้องต้น (Pre-Processing)
กระบวนการเบื้องต้นก่อนที่จะนำรูปภาพไปประมวลผล คือ การนำรูปภาพมาปรับแต่งและจัดเตรียมให้เหมาะสมก่อนจะถูกนำไปประมวลผล ซึ่งหากเกิดความผิดพลาดจะมีผลกระทบต่อขั้นตอนการประมวลผลต่อไป การปรับแต่งภาพต้นฉบับมีหลายวิธีการ เช่น การปรับองศาของรูปภาพ การเบลอหรือทำให้ภาพคมชัดขึ้น การแปลงภาพเป็นภาพขาวดำ การลดทอนแสงเงาในภาพ ฯลฯ

2.1.2 การรู้จำข้อความ (Text recognition)
การรู้จำข้อความจากฉากธรรมชาติ (Text Recognition in Natural Scenes) เป็นกระบวนการประมวลผลข้อมูลภาพถ่ายที่มีข้อความปรากฏอยู่ เพื่อสกัดข้อความนั้น ๆ ออกมาเป็นข้อความในรูปแบบดิจิทัลที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้กระบวนการที่สำคัญในการรู้จำข้อความจากฉากธรรมชาติได้แก่ การดำเนินการเบื้องต้น การตรวจจับข้อความ การรู้จำข้อความ และการดำเนินการหลังการประมวลผล

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)
ความแม่นยำของกระบวนการ OCR สามารถเพิ่มขึ้นได้ถ้าหาก ผลลัพธ์ถูกควบคุมโดย รายการของศัพท์เฉพาะที่อนุญาตให้ใช้ในเอกสาร (lexicon) ซึ่งเป็นรายการของคำที่สามารถปรากฏในเอกสารได้ ตัวอย่างเช่น คำในภาษาอังกฤษ ทั้งหมด และศัพท์เทคนิคเฉพาะซึ่งเทคนิคนี้อาจเป็นปัญหาได้ถ้าในเอกสารมีข้อความที่ไม่ได้อยู่ใน lexicon นอกจากนี้ยังสามารถใช้วิธีการ Near-neighbor analysis แก้ไขคำผิดได้โดยการเปรียบเทียบกับตัวอักษรที่มักจะปรากฏในข้อความเดียวกัน และไวยากรณ์ภาษายังช่วยในการระบุคำที่พบว่าควรจะเป็นคำนาม คำกริยา หรืออื่น ๆ เป็นต้น

2.2 การแก้ไขปัญหาสำหรับการรู้จำข้อความ
Chucai Yi และYingli Tian [1] กล่าวไว้ว่า สำหรับการสกัดข้อความจากฉากธรรมชาติด้วยอุปกรณ์พกพานั้น การตรวจจับข้อความอย่างอัตโนมัติและมีประสิทธิผล และอัลกอริทึมสำหรับการรู้จำถือเป็นสิ่งจำเป็นอย่างยิ่ง อย่างไรก็ตามการสกัดข้อความจากภาพฉากธรรมชาติมีอุปสรรคท้าทายหลายประการเช่น ความหลายหลายของฉากหลังที่มีสิ่งรบกวนและสิ่งที่ไม่ใช่ข้อความปะปนอยู่ และความหลากหลายของรูปแบบตัวอักษรเช่น ประเภทตัวอักษร แบบตัวหนังสือ และขนาด ดังนั้นการแก้ไขปัญหาสำหรับการรู้จำข้อความจึงเป็นความพยายามในการทำให้ความหลากหลายของฉากหลังและความหลากหลายของรูปแบบตัวอักษรมีผลต่อกระบวนการต่าง ๆ เช่น การตรวจจับข้อความ และการรู้จำข้อความให้น้อยที่สุด เพื่อให้ผลลัพธ์มีความถูกต้องและแม่นยำมากขึ้น สอดคล้องกับที่ Woraratpanya, K และคณะ [6] กล่าวไว้ว่า อย่างไรก็ตามประสิทธิผลของอัตราการรู้จำขึ้นอยู่กับการระบุตำแหน่งของข้อความ ความบริสุทธิ์ของกระบวนการแยกข้อความออกจากฉากหลังที่มากจะนำไปสู่อัตราความแม่นยำของการรู้จำที่มากขึ้น

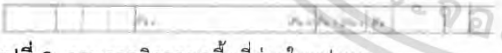
ในปี ค.ศ. 1993 Pierre D. Wellner [5] ได้นำเสนอวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) โดยการพิจารณาภาพโทนสีเทาจากภาพต้นฉบับ จากนั้นเปรียบเทียบข้อมูลค่าความสว่างในแต่ละตำแหน่งของรูปภาพกับค่าเฉลี่ยความสว่างของพื้นที่ย่อยที่ตำแหน่งนั้น ๆ ปรากฏอยู่ เพื่อใช้เป็นเกณฑ์กำหนดสำหรับการสร้างภาพขาวดำ วิธีการนี้สามารถนำมาใช้กับภาพถ่ายของเอกสารที่มีเงาตกกระทบเพื่อแก้ปัญหาข้อความที่ถูกเงาบังและไม่สามารถตรวจจับหรือรู้จำได้ และยังสามารถประมวลผลได้อย่างรวดเร็ว การพิจารณาพื้นที่ย่อยในวิธีการของ Wellner จะเริ่มจากการพิจารณารูปภาพแต่ละรูปเป็นแถวของพิกเซลที่เรียงตัวเป็นแถวเดียว ความกว้างของพื้นที่ย่อย (s) แต่ละช่วงจะมีขนาดเป็น 1 ใน 8 ของความกว้างภาพต้นฉบับ จากนั้นพิจารณาผลรวมของค่าความสว่างจากภาพโทนสีเทาของภาพต้นฉบับภายในแต่ละพื้นที่ย่อยได้ด้วยสมการต่อไปนี้

$$f_s(n) = \sum_{i=0}^{s-1} P_{n-1} \quad (1)$$

เมื่อ $f_s(n)$ เป็นผลรวมของค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n P_n คือค่าความสว่างของพิกเซล ณ ตำแหน่ง n ของภาพโทนสี เทาจากภาพต้นฉบับ และ s คือความกว้างของพื้นที่ย่อย



รูปที่ 2. ภาพการพิจารณาพื้นที่ย่อยในรูปภาพ

ภาพผลลัพธ์จากการประมวลผลคำนวณได้จากการกำหนดค่าในแต่ละพิกเซลของภาพผลลัพธ์ด้วยเงื่อนไขต่อไปนี้

$$T(n) = \begin{cases} 0, \text{ เมื่อ } P_n < \left(\frac{f_s(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ 1, \text{ กรณีอื่น ๆ} \end{cases} \quad (2)$$

เมื่อ t เป็นเปอร์เซ็นต์ที่ค่าความสว่าง ณ ตำแหน่ง n น้อยกว่าค่าเฉลี่ยของพื้นที่ย่อยที่พิจารณา (Wellner เสนอให้ค่า t เป็น15) โดยที่เลข 0 แทนสีดำ และเลข 1 แทนสีขาว ซึ่งผลลัพธ์ที่ได้จะเป็นภาพขาวดำ (Binarization Image)

วิธีการของ Wellner เป็นเพียงรูปแบบหนึ่งของวิธีการกำหนดเกณฑ์แบบปรับเหมาะสำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ ยังมีวิธีการที่แตกต่างออกไปอีกมากมาย เช่น ในปี ค.ศ. 2007 Derek Bradley และคณะ [2] ได้นำเสนอการปรับปรุงวิธีการของ Wellner เพื่อเพิ่มประสิทธิภาพในการแก้ปัญหาในบทความ Adaptive Thresholding Using Integral Image โดยปรับเปลี่ยนวิธีการพิจารณาพื้นที่ย่อยในขั้นตอนการเปรียบเทียบค่าความสว่างของแต่ละตำแหน่งในรูปภาพ

นอกเหนือจากการใช้วิธีการใดวิธีการหนึ่งในการประมวลผลแต่ละรูปภาพ แนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการแสดงให้เห็นผลลัพธ์ที่ดีกว่าดังที่ J. Sauvola และคณะ [3] ได้นำเสนอไว้ในบทความในปี ค.ศ. 2000 โดยดำเนินการประมวลผลจำแนกข้อความและพื้นหลังจากภาพถ่ายจากนั้นได้แบ่งวิธีการประมวลผลออกเป็นสองวิธี วิธีการหนึ่งเป็นการแยกองค์ประกอบพื้นหลัง และอีกวิธีการหนึ่งเป็นการแปลงภาพเป็นภาพขาวดำเพื่อการสกัดข้อความออกจากพื้นหลังภายใต้ปัญหาที่ส่งผลต่อการตรวจจับข้อความ

3. การดำเนินงานวิจัย

3.1. การศึกษากระบวนการรู้จำข้อความ

สำหรับการศึกษากระบวนการรู้จำข้อความ (Text Recognition) ทำโดยการค้นคว้าจากทฤษฎีทางด้านการประมวลผลรูปภาพ (Image Processing) และการค้นคว้าจากบทความทางวิชาการที่เกี่ยวข้องกับการรู้จำข้อความ, การแก้ปัญหาด้านต่าง ๆ ที่เป็นอุปสรรคต่อการตรวจจับข้อความ และการรู้จำข้อความ รวมถึงบทความที่เกี่ยวข้องกับการรู้จำข้อความบนอุปกรณ์พกพา (Mobile Devices)

3.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์เคลื่อนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา ได้ศึกษาการพัฒนาแอปพลิเคชันบน **Windows Platform** และได้ทดลองพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย เพื่อทำการศึกษาวิธีการพัฒนาการรู้จำข้อความบนอุปกรณ์พกพาและเรียนรู้ปัญหาหรืออุปสรรคที่เกิดขึ้น

3.3 การศึกษาความต้องการต่อการใช้แอปพลิเคชัน

สำหรับการศึกษาความต้องการต่อการใช้แอปพลิเคชันเป็นการสำรวจความต้องการของกลุ่มผู้ใช้เป้าหมายเพื่อเป็นตัวกำหนดคุณลักษณะ (**Feature**) สำหรับแอปพลิเคชันที่พัฒนา

3.4 การศึกษาการปรับปรุงภาพต้นฉบับ

การศึกษาการปรับปรุงภาพต้นฉบับต้องการแก้ปัญหาการตรวจจับข้อความในฉากธรรมชาติดังต่อไปนี้ คือ ปัญหาด้านเงา และ ปัญหาด้านแสงสว่าง

3.4.1 ขั้นตอนในการศึกษาการปรับปรุงภาพต้นฉบับ

ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา การปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชัน และการทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น

3.4.2 นิยามลักษณะปัญหา

3.4.2.1 ปัญหาด้านเงา

ภาพที่มีเงาบดบังบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีเงาบดบังบริเวณที่มีข้อความทั้งหมด

3.4.2.2. ปัญหาด้านแสงสว่าง

ภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบริเวณที่มีข้อความทั้งหมด

รูปที่ 3. ตัวอย่างภาพที่มีปัญหาด้านเงา

รูปที่ 4. ตัวอย่างภาพที่มีปัญหาแสงสว่าง

3.4.3 การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะ

จากการศึกษากระบวนการแก้ปัญหาของ Pierre D. Wellner [5] ดังที่ได้อธิบายวิธีการในหัวข้อ 2.2 พบว่าสามารถประมวลผลได้อย่างรวดเร็วเนื่องจากกระบวนการมีความซับซ้อนที่น้อย ทว่ามีข้อจำกัด ข้อบกพร่อง และข้อสังเกตในการประมวลผล เช่น ให้ผลลัพธ์ที่ดีเฉพาะกับภาพที่มีปัญหาเงาในระดับปานกลางถึงรุนแรง ให้ผลลัพธ์ที่ไม่ดีกับกรณีทั่วไป ได้แก่ ปัญหาแสง ปัญหาอื่น ๆ ที่นอกเหนือจากนิยาม และกรณีภาพที่ไม่มีปัญหา นอกจากนี้ยังพบว่าหากใช้ภาพส่วนกลับโทนสีเทา (Complement-grayscale image) ของภาพต้นฉบับในการดำเนินการจะสามารถให้ผลลัพธ์ที่ดีกับภาพที่มีปัญหาแสงในระดับปานกลางถึงรุนแรง

จากแนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการดังที่ J. Sauvola และคณะ [3] ได้ดำเนินการศึกษาและพัฒนากระบวนการแก้ปัญหาต่อการรู้จำข้อความทำให้เชื่อว่าอาจมีวิธีการดัดแปลงให้การแก้ปัญหาของ Wellner สามารถแก้ไขปัญหาแสงได้โดยการประยุกต์ใช้วิธีการร่วม สำหรับการประยุกต์ใช้วิธีการร่วมเพื่อแก้ปัญหาของรูปภาพนั้น K. Peuwnuan และคณะ [4] ได้นำเสนอวิธีการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะโดยใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพรวมผล (Integral Image) สำหรับการตรวจจับข้อความ ซึ่งเป็นการดัดแปลงกระบวนการของ Derek Bradley และคณะ [2] โดยมีการแบ่งพื้นที่รูปภาพออกเป็นส่วนต่าง ๆ (Image Segmentation) คือพื้นที่ที่มีค่าความสว่างสูง และพื้นที่ที่มีค่าความสว่างปานกลางถึงต่ำ จากนั้นจึงดำเนินการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในแต่ละพื้นที่ด้วยวิธีการที่แตกต่างกัน โดยอาศัยแนวคิดข้างต้น ในการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพาจึงได้ดัดแปลงวิธีการของ Wellner โดย ก่อนดำเนินการกำหนดเกณฑ์แบบปรับเหมาะจำเป็นต้องมีการจำแนกพื้นที่ออกเป็นพื้นที่ที่มีค่าความสว่างสูงและค่าความสว่างต่ำ เนื่องจากพื้นที่ที่จำแนกได้นั้นจะใช้วิธีการประมวลผลเพื่อการแปลงภาพขาวดำที่แตกต่างกันโดยเงื่อนไขที่ใช้จำแนกพื้นที่ของรูปภาพแสดงโดยสมการที่ (3)

$$P(n) = \begin{cases} \text{พื้นที่ความสว่างสูง, เมื่อ } P_n > M + SD \\ \text{พื้นที่ความสว่างต่ำ, กรณีอื่น ๆ} \end{cases} \quad (3)$$

เมื่อ P_n คือค่าความสว่างของพิกเซลที่พิจารณาจากภาพโทนสีเทา M คือค่าเฉลี่ยความสว่างของทั้งรูปภาพ *และ SD* คือส่วนเบี่ยงเบนมาตรฐานของค่าความสว่างทั้งรูปภาพ

หลังจากแบ่งพื้นที่รูปภาพตามเงื่อนไขในสมการ (3) แล้วจึงดำเนินการกำหนดเกณฑ์แบบปรับเหมาะสำหรับแต่ละพื้นที่ซึ่งจะมีความแตกต่างกันดังนี้

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P'_n \geq \left(\frac{f'_S(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (4)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P'_n คือค่าความสว่างของพิกเซลที่ n ของภาพส่วนกลับโทนสีเทา และ $f'_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพส่วนกลับโทนสีเทา สำหรับสมการ (4) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างสูงตามที่จำแนกได้และ

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P_n \geq \left(\frac{f_S(n)}{s}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (5)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P_n คือค่าความสว่างของพิกเซลที่ n ของภาพโทนสีเทา และ $f_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพโทนสีเทา สำหรับสมการ (5) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างปานกลางถึงต่ำตามที่จำแนกได้

3.5 ผลการดำเนินงาน

3.5.1 การศึกษากระบวนการรู้จำข้อความ

ผลการศึกษาฯ ทำให้มีความรู้ความเข้าใจในกระบวนการรู้จำข้อความมากยิ่งขึ้น เพียงพอต่อการนำไปใช้เพื่อพัฒนาแอปพลิเคชันต่อไป รวมถึงเข้าใจในปัญหาและอุปสรรคต่อการตรวจจับและการรู้จำข้อความ

3.5.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

ผลการศึกษาและทดลองฯ ทำให้สามารถพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย และได้ออกแบบการทดลองและบันทึกผลการรู้จำข้อความจากภาพตัวอย่างที่มีความหลากหลายขององค์ประกอบในภาพเพื่อคำนวณความแม่นยำของแอปพลิเคชันก่อนจะพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ

3.5.3 การศึกษาการปรับปรุงภาพต้นฉบับ

3.5.3.1. ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ

จากการศึกษาพบว่าการปรับปรุงภาพต้นฉบับด้วยวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding) สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) สามารถประมวลผลได้เร็วเหมาะแก่การพัฒนาบนอุปกรณ์พกพาโดยเฉพาะวิธีการของ Wellner ซึ่งเหมาะอย่างยิ่งสำหรับกรณีภาพที่มีปัญหาเงาปานกลางถึงรุนแรง และสามารถดัดแปลงเพื่อให้ใช้แก้ปัญหาแสงได้เช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.3.2. ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา มีการทดลองใช้วิธีการของ Wellner เพื่อดูผลในการแก้ปัญหาเงาที่มีผลต่อการตรวจจับข้อความ พบว่าสามารถแก้ปัญหาเงาระดับปานกลางถึงสูงได้ดี แต่จะส่งผลลบกับการประมวลผลหากนำไปใช้กับภาพที่มีปัญหาแสงหรือปัญหาอื่น ๆ กล่าวคือจะลดความแม่นยำ ความถูกต้อง และความครบถ้วนของผลลัพธ์ลงอย่างมาก แต่สามารถปรับให้ใช้แก้ไขปัญหาแสงได้หากใช้ภาพส่วนกลับของภาพโทนสีเทา (Complement-Grayscale Image) ในการดำเนินการ

3.5.3.3. ปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชันอยู่ มีการปรับใช้วิธีการของ Wellner เพื่อให้สามารถแก้ไขปัญหาของภาพได้ทั้งเงาและแสงโดยการสร้างเงื่อนไขเพื่อพิจารณาพื้นที่ของรูปภาพที่เป็นส่วนความสว่างต่ำถึงปานกลาง และความสว่างสูง และใช้วิธีการดำเนินการในแต่ละพื้นที่ที่แตกต่างกันซึ่งให้ผลลัพธ์ที่ดีขึ้นสำหรับกรณีปัญหาแสงที่เกิดกับรูปภาพ

3.5.3.4. ทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงาแต่มีปัญหาในกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้ซึ่งตัวเลขผลการทดลอง การเปรียบเทียบและการอภิปรายจะได้กล่าวถึงในรายละเอียด ในหัวข้อ 4.2

บทที่ 4 กระบวนการทำงานของระบบ

4.1 แอปพลิเคชัน

การพัฒนาแอปพลิเคชันได้พัฒนาบนอุปกรณ์สองประเภทคือบนสมาร์ทโฟน และบนแทบเล็ต สำหรับตัวอย่างหน้าจอการทำงานภายในแอปพลิเคชันมีดังต่อไปนี้

รูปที่ 5. หน้าจอแอพลิเคชันบนสมาร์ทโฟน

รูปที่ 6. หน้าจอแอพลิเคชันบนแทบเล็ต

4.2 การทดลองและการสรุปผล

การทดลองมีจุดประสงค์เพื่อเปรียบเทียบผลลัพธ์จากการทำงานของแอพพลิเคชั่นก่อนและหลังการเพิ่มเติมการประมวลผลเบื้องต้น โดยชุดทดสอบที่ใช้เป็นภาพฉากธรรมชาติจากการแข่งขันของ ICDAR 2013 จำนวน 229 ภาพ ซึ่งเป็นภาพที่มีปัญหาแสงและ/หรือเงาเพียง 59 ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 ตัวอย่างรูปที่ใช้ในการทดลอง

รูปที่ใช้ในการทดลองจำแนกออกตามนิยามปัญหาได้ ในการทดลองจำแนกออกตามนิยามปัญหาได้ 2 ประเภทคือ แสงและเงา และกรณีอื่น ๆ เพิ่มเข้ามาด้วย

ตารางที่ 1. การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชัน(โดยเฉลี่ย)

Shadow	Lighting	Others

ตารางที่ 2. การจำแนกปัญหาของที่ปรากฏในแต่ละรูปภาพที่เป็นชุดทดสอบ

การประมวลผล	ตัวชี้วัดการประเมินผล		
	precision	recall	correctness
ก่อน	82.66%	74.44%	82.18%
หลัง	40.55%	30.64%	39.99%

ตารางที่ 3. ตารางแสดงรูปภาพกรณีที่ Recall สูงขึ้น

1.	3.

2.	4.

4.2.2 สรุปผลการทดลอง

ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงามีจุดบกพร่องกับกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้

ตารางที่ 4. ตารางแสดงผลลัพธ์กรณีที่ Recall สูงขึ้น

รูป		ตัวชี้วัดการประเมินผล		
		Precision (%)	Recall (%)	Correctness (%)
1	ก่อน	50	13.04	100
	หลัง	92	100	100
2	ก่อน	100	60	66.67
	หลัง	71.43	100	100
3	ก่อน	100	52.63	100
	หลัง	100	100	100
4	ก่อน	100	44.90	100
	หลัง	96.97	65.31	93.94

บทที่ 5 สรุปผลโครงงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

การศึกษาการระบบการรู้จำข้อความจากฉากธรรมชาติ ได้มีการนำไลบรารี OCR เข้ามาใช้แทนระบบการรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้การประมวลผลบนอุปกรณ์พกพาประเภทสมาร์ทโฟนทำได้รวดเร็วยิ่งขึ้นและใช้ทรัพยากรที่มีจำกัดของอุปกณ์ได้อย่างมีประสิทธิภาพ

การศึกษาภาพการปรับปรุงต้นฉบับได้ใช้หลักการการประมวลผลรูปภาพ (Image Processing) เพื่อศึกษาทดลองวิธีการปรับปรุงภาพต้นฉบับก่อนเข้าสู่กระบวนการรู้จำข้อความ พบว่าการปรับปรุงภาพต้นฉบับต้องพัฒนาอัลกอริทึมเพิ่มเติมที่สามารถระบุวิธีการปรับปรุงภาพต้นฉบับที่จำเพาะกับภาพนั้นๆ เนื่องจากภาพถ่ายจากฉากธรรมชาติมีความหลากหลายมากมาย ทำให้กระบวนการปรับปรุงภาพต้นฉบับไม่สามารถใช้รูปแบบใดรูปแบบหนึ่งได้กับทุกภาพ

ข้อเสนอแนะ

1. สามารถต่อยอดพัฒนาลงบนอุปกรณ์อื่น ๆ นอกจากอุปกรณ์พกพาได้
2. สามารถต่อยอดพัฒนาให้ทำงานแบบ Real Time ได้
3. สามารถเพิ่มประสิทธิภาพการทำงานของแอปพลิเคชันให้ดีขึ้นได้โดยการปรับปรุงกระบวนการประมวลผลเบื้องต้นเพิ่มเติม

กระบวนการที่ได้ศึกษาและพัฒนามุ่งเน้นที่ความรวดเร็วในการประมวลผลซึ่งโดยปกติจะทำให้ประสิทธิผลลดลง การพัฒนาเพิ่มเติมควรเพิ่มประสิทธิผลโดยให้กระทบกับความเร็วในการประมวลผลน้อยที่สุด

เอกสารอ้างอิง

[1] Chucai Yi and Yingli Tian, "**Scenes Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration**", IEEE Transactions on Image Processing, Vol. 23, No. 7, July 2014.

[2] Derek Bradley and Gerhard Roth, "**Adaptive Thresholding Using the Integral**", Journal of Graphics, GPU, and Game Tools Volume 12, Issue 2, 2007.

[3] J. Sauvola and M. PietikaKinen, "**Adaptive Document Image Binarization**", Machine Vision and Media Processing Group, Infotech Oulu, University of Oulu, P.O. BOX 4500, FIN-90401 Oulu, Finland Received 29 April 1998; accepted 21 January 1999

[4] K. Peuwnuan, K. Woraratpanya, K. Pasupa, "**Modified Adaptive Thresholding Using Integral Image**", (unpublished).

[5] Pierre D. Wellner," **Adaptive Thresholding for the DigitalDesk**", Technical Report EPC-1993-110.

[6] Woraratpanya, K.Fac. of Inf. Technol., King Mongkut's Inst. of Technol. Ladkrabang, Bangkok, Thailand Pasupa, K. ; Suttapakti, U. ; Boonchukusol, P. ; Titijaroonroj, T. ; Hokking, R. ; Kuroki, Y. ; Kato, Y. , "**Text-background Decomposition for Thai Text Localization and Recognition in Natural Scenes**", Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on 7-8 Oct. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพา

กฤตยชญ์ คงคติธรรม[1] กิตติภพ ผิวนวล[2] กิตติภพ ผิวนวล[2]

ผศ.ดร. กิติ์สุชาต พสุภา[3] และ ผศ.ดร. กันต์พงษ์ วรรัตน์ปัญญา[4]

[1]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: k.kittayoch@gmail.com[1], plewplerng@hotmail.com[2],
kitsuchart@it.kmitl.ac.th[3], kuntpong@it.kmitl.ac.th[4]

บทคัดย่อ

หลังจากที่ประเทศไทยได้ก้าวเข้าสู่ประชาคมเศรษฐกิจอาเซียน การติดต่อสื่อสารกับประเทศเพื่อนบ้านและประเทศคู่ค้าในภูมิภาคใกล้เคียงทำให้เราได้รับข้อมูล ข่าวสารในภาษาต่างประเทศมากขึ้น นอกจากนี้ยังรวมไปถึงสื่อสิ่งพิมพ์ โฆษณา ประกาศ และอื่นๆ ที่ใช้ภาษาต่างประเทศ ซึ่งบางคนต้องการเครื่องมือสำหรับแปลข้อความเหล่านั้นเพื่อการรับรู้ข้อมูลที่ถูกต้อง

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตาม การพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และปัญหาจากแสงและเงาซึ่งเป็นอุปสรรคต่อการรู้จำข้อความโครงงานชุดนี้จึงพัฒนาแอปพลิเคชันโดยศึกษาและพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ แอปพลิเคชันนี้ไม่เพียงแต่ใช้การรู้จำเพื่อนำไปแปลความหมายเท่านั้น แต่ยังออกแบบให้สามารถนำข้อความที่รู้จำได้ไปใช้ต่อในอุปกรณ์พกพาของผู้ใช้

คำสำคัญ – การประมวลผลเบื้องต้น; การตรวจจับข้อความ; การรู้จำข้อความ; การแปลภาษา

1. บทนำ

การพัฒนาแอปพลิเคชันสำหรับการแปลข้อความจากภาษาต่างประเทศเป็นภาษาไทยเป็นสิ่งที่มีอยู่แพร่หลาย อย่างไรก็ตามการพัฒนาแอปพลิเคชันเพื่อการแปลข้อความจากฉากธรรมชาติยังคงเป็นสิ่งที่มีความท้าทายอยู่หลายประการ เช่น ความหลากหลายขององค์ประกอบในภาพ ความหลากหลายของรูปแบบตัวอักษร และโดยเฉพาะปัญหาแสงเงาที่กระทบต่อการรู้จำข้อความ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. การรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OCR หรือ Optical Character Recognition เป็นกระบวนการแปลงข้อมูลภาพของข้อความทั้งตัวพิมพ์และตัวเขียน (ลายมือ) ไปสู่ข้อความในแบบที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้ ซึ่งก็คือการแปลงข้อมูลในเอกสาร สื่อสิ่งพิมพ์ต่าง ๆ ให้อยู่ในรูปแบบข้อมูลดิจิทัลที่สามารถนำไปแก้ไข ค้นหา จัดหมวดหมู่ จัดเก็บ แสดงผล และประมวลผลได้ เช่น การแปลภาษา การสังเคราะห์เสียงพูดจากข้อความ การจัดทำฐานข้อมูล ฯลฯ ทั้งนี้ OCR มีความเกี่ยวข้องกับทั้งการศึกษาการรู้จำรูปแบบ (Pattern Recognition) ปัญญาประดิษฐ์ (Artificial Intelligence) และคอมพิวเตอร์วิทัศน์ (Computer Vision)

ขั้นตอนการทำงานของ OCR มี 3 ขั้นตอนหลัก

รูปที่ 1. ขั้นตอนการทำงานของระบบ OCR

2.1.1 การประมวลผลเบื้องต้น (Pre-Processing)
กระบวนการเบื้องต้นก่อนที่จะนำรูปภาพไปประมวลผล คือ การนำรูปภาพมาปรับแต่งและจัดเตรียมให้เหมาะสมก่อนจะถูกนำไปประมวลผล ซึ่งหากเกิดความผิดพลาดจะมีผลกระทบต่อขั้นตอนการประมวลผลต่อไป การปรับแต่งภาพต้นฉบับมีหลายวิธีการ เช่น การปรับองศาของรูปภาพ การเบลอหรือทำให้ภาพคมชัดขึ้น การแปลงภาพเป็นภาพขาวดำ การลดทอนแสงเงาในภาพ ฯลฯ

2.1.2 การรู้จำข้อความ (Text recognition)
การรู้จำข้อความจากฉากธรรมชาติ (Text Recognition in Natural Scenes) เป็นกระบวนการประมวลผลข้อมูลภาพถ่ายที่มีข้อความปรากฏอยู่ เพื่อสกัดข้อความนั้น ๆ ออกมาเป็นข้อความในรูปแบบดิจิทัลที่คอมพิวเตอร์สามารถเข้าใจ จดจำ และประมวลผลได้กระบวนการที่สำคัญในการรู้จำข้อความจากฉากธรรมชาติได้แก่ การดำเนินการเบื้องต้น การตรวจจับข้อความ การรู้จำข้อความ และการดำเนินการหลังการประมวลผล

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)
ความแม่นยำของกระบวนการ OCR สามารถเพิ่มขึ้นได้ถ้าหาก ผลลัพธ์ถูกควบคุมโดย รายการของศัพท์เฉพาะที่อนุญาตให้ใช้ในเอกสาร (lexicon) ซึ่งเป็นรายการของคำที่สามารถปรากฏในเอกสารได้ ตัวอย่างเช่น คำในภาษาอังกฤษ ทั้งหมด และศัพท์เทคนิคเฉพาะซึ่งเทคนิคนี้อาจเป็นปัญหาได้ถ้าในเอกสารมีข้อความที่ไม่ได้อยู่ใน lexicon นอกจากนี้ยังสามารถใช้วิธีการ Near-neighbor analysis แก้ไขคำผิดได้โดยการเปรียบเทียบกับตัวอักษรที่มักจะปรากฏในข้อความเดียวกัน และไวยากรณ์ภาษายังช่วยในการระบุคำที่พบว่าควรจะเป็นคำนาม คำกริยา หรืออื่น ๆ เป็นต้น

2.2 การแก้ไขปัญหาสำหรับการรู้จำข้อความ
Chucai Yi และYingli Tian [1] กล่าวไว้ว่า สำหรับการสกัดข้อความจากฉากธรรมชาติด้วยอุปกรณ์พกพานั้น การตรวจจับข้อความอย่างอัตโนมัติและมีประสิทธิผล และอัลกอริทึมสำหรับการรู้จำถือเป็นสิ่งจำเป็นอย่างยิ่ง อย่างไรก็ตามการสกัดข้อความจากภาพฉากธรรมชาติมีอุปสรรคท้าทายหลายประการเช่น ความหลายหลายของฉากหลังที่มีสิ่งรบกวนและสิ่งที่ไม่ใช่ข้อความปะปนอยู่ และความหลากหลายของรูปแบบตัวอักษรเช่น ประเภทตัวอักษรแบบตัวหนังสือ และขนาด ดังนั้นการแก้ไขปัญหาสำหรับการรู้จำข้อความจึงเป็นความพยายามในการทำให้ความหลากหลายของฉากหลังและความหลากหลายของรูปแบบตัวอักษรมีผลต่อกระบวนการต่าง ๆ เช่น การตรวจจับข้อความ และการรู้จำข้อความให้น้อยที่สุด เพื่อให้ผลลัพธ์มีความถูกต้องและแม่นยำมากขึ้น สอดคล้องกับที่ Woraratpanya, K และคณะ [6] กล่าวไว้ว่า อย่างไรก็ตามประสิทธิผลของอัตราการรู้จำขึ้นอยู่กับการระบุตำแหน่งของข้อความ ความบริสุทธิ์ของกระบวนการแยกข้อความออกจากฉากหลังที่มากจะนำไปสู่อัตราความแม่นยำของการรู้จำที่มากขึ้น

ในปี ค.ศ. 1993 Pierre D. Wellner [5] ได้นำเสนอวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) โดยการพิจารณาภาพโทนสีเทาจากภาพต้นฉบับ จากนั้นเปรียบเทียบข้อมูลค่าความสว่างในแต่ละตำแหน่งของรูปภาพกับค่าเฉลี่ยความสว่างของพื้นที่ย่อยที่ตำแหน่งนั้น ๆ ปรากฏอยู่ เพื่อใช้เป็นเกณฑ์กำหนดสำหรับการสร้างภาพขาวดำ วิธีการนี้สามารถนำมาใช้กับภาพถ่ายของเอกสารที่มีเงาตกกระทบเพื่อแก้ปัญหาข้อความที่ถูกเงาบังและไม่สามารถตรวจจับหรือรู้จำได้ และยังสามารถประมวลผลได้อย่างรวดเร็ว การพิจารณาพื้นที่ย่อยในวิธีการของ Wellner จะเริ่มจากการพิจารณารูปภาพแต่ละรูปเป็นแถวของพิกเซลที่เรียงตัวเป็นแถวเดียว ความกว้างของพื้นที่ย่อย (s) แต่ละช่วงจะมีขนาดเป็น 1 ใน 8 ของความกว้างภาพต้นฉบับ จากนั้นพิจารณาผลรวมของค่าความสว่างจากภาพโทนสีเทาของภาพต้นฉบับภายในแต่ละพื้นที่ย่อยได้ด้วยสมการต่อไปนี้

$$f_s(n) = \sum_{i=0}^{s-1} P_{n-1} \quad (1)$$

เมื่อ $f_s(n)$ เป็นผลรวมของค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n P_n คือค่าความสว่างของพิกเซล ณ ตำแหน่ง n ของภาพโทนสี เทาจากภาพต้นฉบับ และ s คือความกว้างของพื้นที่ย่อย

รูปที่ 2. ภาพการพิจารณาพื้นที่ย่อยในรูปภาพ

ภาพผลลัพธ์จากการประมวลผลคำนวณได้จากการกำหนดค่าในแต่ละพิกเซลของภาพผลลัพธ์ด้วยเงื่อนไขต่อไปนี้

$$T(n) = \begin{cases} 0, \text{ เมื่อ } P_n < \left(\frac{f_s(n)}{s}\right)\left(\frac{100\text{-}t}{100}\right) \\ 1, \text{ กรณีอื่น ๆ} \end{cases} \quad (2)$$

เมื่อ t เป็นเปอร์เซ็นต์ที่ค่าความสว่าง ณ ตำแหน่ง n น้อยกว่าค่าเฉลี่ยของพื้นที่ย่อยที่พิจารณา (Wellner เสนอให้ค่า t เป็น15) โดยที่เลข 0 แทนสีดำ และเลข 1 แทนสีขาว ซึ่งผลลัพธ์ที่ได้จะเป็นภาพขาวดำ (Binarization Image)

วิธีการของ Wellner เป็นเพียงรูปแบบหนึ่งของวิธีการกำหนดเกณฑ์แบบปรับเหมาะสำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ ยังมีวิธีการที่แตกต่างออกไปอีกมากมาย เช่น ในปี ค.ศ. 2007 Derek Bradley และคณะ [2] ได้นำเสนอการปรับปรุงวิธีการของ Wellner เพื่อเพิ่มประสิทธิภาพในการแก้ปัญหาในบทความ Adaptive Thresholding Using Integral Image โดยปรับเปลี่ยนวิธีการพิจารณาพื้นที่ย่อยในขั้นตอนการเปรียบเทียบค่าความสว่างของแต่ละตำแหน่งในรูปภาพ

นอกเหนือจากการใช้วิธีการใดวิธีการหนึ่งในการประมวลผลแต่ละรูปภาพ แนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการแสดงให้เห็นผลลัพธ์ที่ดีกว่าดังที่ J. Sauvola และคณะ [3] ได้นำเสนอไว้ในบทความในปี ค.ศ. 2000 โดยดำเนินการประมวลผลจำแนกข้อความและพื้นหลังจากภาพถ่ายจากนั้นได้แบ่งวิธีการประมวลผลออกเป็นสองวิธี วิธีการหนึ่งเป็นการแยกองค์ประกอบพื้นหลัง และอีกวิธีการหนึ่งเป็นการแปลงภาพเป็นภาพขาวดำเพื่อการสกัดข้อความออกจากพื้นหลังภายใต้ปัญหาที่ส่งผลต่อการตรวจจับข้อความ

3. การดำเนินงานวิจัย

3.1. การศึกษากระบวนการรู้จำข้อความ

สำหรับการศึกษากระบวนการรู้จำข้อความ (Text Recognition) ทำโดยการค้นคว้าจากทฤษฎีทางด้านการประมวลผลรูปภาพ (Image Processing) และการค้นคว้าจากบทความทางวิชาการที่เกี่ยวข้องกับการรู้จำข้อความ, การแก้ปัญหาด้านต่าง ๆ ที่เป็นอุปสรรคต่อการตรวจจับข้อความ และการรู้จำข้อความ รวมถึงบทความที่เกี่ยวข้องกับการรู้จำข้อความบนอุปกรณ์พกพา (Mobile Devices)

3.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์เคลื่อนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา ได้ศึกษาการพัฒนาแอปพลิเคชันบน **Windows Platform** และได้ทดลองพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย เพื่อทำการศึกษาวิธีการพัฒนาการรู้จำข้อความบนอุปกรณ์พกพาและเรียนรู้ปัญหาหรืออุปสรรคที่เกิดขึ้น

3.3 การศึกษาความต้องการต่อการใช้แอปพลิเคชัน

สำหรับการศึกษาความต้องการต่อการใช้แอปพลิเคชันเป็นการสำรวจความต้องการของกลุ่มผู้ใช้เป้าหมายเพื่อเป็นตัวกำหนดคุณลักษณะ (**Feature**) สำหรับแอปพลิเคชันที่พัฒนา

3.4 การศึกษาการปรับปรุงภาพต้นฉบับ

การศึกษาการปรับปรุงภาพต้นฉบับต้องการแก้ปัญหาการตรวจจับข้อความในฉากธรรมชาติดังต่อไปนี้ คือ ปัญหาด้านเงา และ ปัญหาด้านแสงสว่าง

3.4.1 ขั้นตอนในการศึกษาการปรับปรุงภาพต้นฉบับ

ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ ทดลองดำเนินการตามวิธีการแก้ปัญหาที่ได้ศึกษา การปรับใช้วิธีการที่ได้ศึกษาสำหรับส่วนการประมวลผลเบื้องต้นของแอปพลิเคชัน และการทดลองและเปรียบเทียบผลลัพธ์การทำงานของแอปพลิเคชันก่อนและหลังการเพิ่มส่วนการประมวลผลเบื้องต้น

3.4.2 นิยามลักษณะปัญหา

3.4.2.1 ปัญหาด้านเงา

ภาพที่มีเงาบดบังบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีเงาบดบังบริเวณที่มีข้อความทั้งหมด

3.4.2.2. ปัญหาด้านแสงสว่าง

ภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบางส่วนของบริเวณที่มีข้อความ หรือภาพที่มีแสงสว่าง/แสงสะท้อนตกกระทบบริเวณที่มีข้อความทั้งหมด

รูปที่ 3. ตัวอย่างภาพที่มีปัญหาด้านเงา

รูปที่ 4. ตัวอย่างภาพที่มีปัญหาแสงสว่าง

3.4.3 การปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะ

จากการศึกษากระบวนการแก้ปัญหาของ Pierre D. Wellner [5] ดังที่ได้อธิบายวิธีการในหัวข้อ 2.2 พบว่าสามารถประมวลผลได้อย่างรวดเร็วเนื่องจากกระบวนการมีความซับซ้อนที่น้อย ทว่ามีข้อจำกัด ข้อบกพร่อง และข้อสังเกตในการประมวลผล เช่น ให้ผลลัพธ์ที่ดีเฉพาะกับภาพที่มีปัญหาเงาในระดับปานกลางถึงรุนแรง ให้ผลลัพธ์ที่ไม่ดีกับกรณีทั่วไป ได้แก่ ปัญหาแสง ปัญหาอื่น ๆ ที่นอกเหนือจากนิยาม และกรณีภาพที่ไม่มีปัญหา นอกจากนี้ยังพบว่าหากใช้ภาพส่วนกลับโทนสีเทา (Complement-grayscale image) ของภาพต้นฉบับในการดำเนินการจะสามารถให้ผลลัพธ์ที่ดีกับภาพที่มีปัญหาแสงในระดับปานกลางถึงรุนแรง

จากแนวความคิดในการใช้วิธีการประมวลผลร่วมระหว่างหลายวิธีการดังที่ J. Sauvola และคณะ [3] ได้ดำเนินการศึกษาและพัฒนากระบวนการแก้ปัญหาต่อการรู้จำข้อความทำให้เชื่อว่าอาจมีวิธีการดัดแปลงให้การแก้ปัญหาของ Wellner สามารถแก้ไขปัญหาแสงได้โดยการประยุกต์ใช้วิธีการร่วม สำหรับการประยุกต์ใช้วิธีการร่วมเพื่อแก้ปัญหาของรูปภาพนั้น K. Peuwnuan และคณะ [4] ได้นำเสนอวิธีการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะโดยใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพรวมผล (Integral Image) สำหรับการตรวจจับข้อความ ซึ่งเป็นการดัดแปลงกระบวนการของ Derek Bradley และคณะ [2] โดยมีการแบ่งพื้นที่รูปภาพออกเป็นส่วนต่าง ๆ (Image Segmentation) คือพื้นที่ที่มีค่าความสว่างสูง และพื้นที่ที่มีค่าความสว่างปานกลางถึงต่ำ จากนั้นจึงดำเนินการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในแต่ละพื้นที่ด้วยวิธีการที่แตกต่างกัน โดยอาศัยแนวคิดข้างต้น ในการปรับปรุงการกำหนดเกณฑ์แบบปรับเหมาะสำหรับการประมวลผลเบื้องต้นของการรู้จำข้อความบนอุปกรณ์พกพาจึงได้ดัดแปลงวิธีการของ Wellner โดย ก่อนดำเนินการกำหนดเกณฑ์แบบปรับเหมาะจำเป็นต้องมีการจำแนกพื้นที่ออกเป็นพื้นที่ที่มีค่าความสว่างสูงและค่าความสว่างต่ำ เนื่องจากพื้นที่ที่จำแนกได้นั้นจะใช้วิธีการประมวลผลเพื่อการแปลงภาพขาวดำที่แตกต่างกันโดยเงื่อนไขที่ใช้จำแนกพื้นที่ของรูปภาพแสดงโดยสมการที่ (3)

$$P(n) = \begin{cases} \text{พื้นที่ความสว่างสูง, เมื่อ } P_n > M + SD \\ \text{พื้นที่ความสว่างต่ำ, กรณีอื่น ๆ} \end{cases} \quad (3)$$

เมื่อ P_n คือค่าความสว่างของพิกเซลที่พิจารณาจากภาพโทนสีเทา M คือค่าเฉลี่ยความสว่างของทั้งรูปภาพ *และ* SD คือส่วนเบี่ยงเบนมาตรฐานของค่าความสว่างทั้งรูปภาพ

หลังจากแบ่งพื้นที่รูปภาพตามเงื่อนไขในสมการ (3) แล้วจึงดำเนินการกำหนดเกณฑ์แบบปรับเหมาะสำหรับแต่ละพื้นที่ซึ่งจะมีความแตกต่างกันดังนี้

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P'_n \geq \left(\frac{f'_s(n)}{s}\right)\left(\frac{100-t}{100}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (4)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P'_n คือค่าความสว่างของพิกเซลที่ n ของภาพส่วนกลับโทนสีเทา และ $f'_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพส่วนกลับโทนสีเทา สำหรับสมการ (4) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างสูงตามที่จำแนกได้และ

$$T(n) = \begin{cases} \text{ดำ, ถ้า } P_n \geq \left(\frac{f_s(n)}{s}\right) \\ \text{ขาว, กรณีอื่น ๆ} \end{cases} \quad (5)$$

เมื่อ T_n คือฟังก์ชันแสดงค่าของพิกเซลที่ n ในภาพขาวดำ ซึ่งเป็นผลลัพธ์ P_n คือค่าความสว่างของพิกเซลที่ n ของภาพโทนสีเทา และ $f_s(n)$ คือผลรวมค่าความสว่างภายใต้พื้นที่ย่อย ณ จุด n ของภาพโทนสีเทา สำหรับสมการ (5) จะใช้ในการกำหนดเกณฑ์สำหรับการแปลงภาพขาวดำในบริเวณของภาพที่มีความสว่างปานกลางถึงต่ำตามที่จำแนกได้

3.5 ผลการดำเนินงาน

3.5.1 การศึกษากระบวนการรู้จำข้อความ

ผลการศึกษาฯ ทำให้มีความรู้ความเข้าใจในกระบวนการรู้จำข้อความมากยิ่งขึ้น เพียงพอต่อการนำไปใช้เพื่อพัฒนาแอปพลิเคชันต่อไป รวมถึงเข้าใจในปัญหาและอุปสรรคต่อการตรวจจับและการรู้จำข้อความ

3.5.2 การศึกษาและทดลองกระบวนการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

ผลการศึกษาและทดลองฯ ทำให้สามารถพัฒนาแอปพลิเคชันเพื่อการสกัดข้อความจากรูปภาพอย่างง่าย และได้ออกแบบการทดลองและบันทึกผลการรู้จำข้อความจากภาพตัวอย่างที่มีความหลากหลายขององค์ประกอบในภาพเพื่อคำนวณความแม่นยำของแอปพลิเคชันก่อนจะพัฒนาวิธีการปรับรูปภาพให้เอื้อต่อการรู้จำข้อความ

3.5.3 การศึกษาการปรับปรุงภาพต้นฉบับ

3.5.3.1. ศึกษาวิธีการแก้ปัญหาสำหรับการตรวจจับข้อความจากฉากธรรมชาติ

จากการศึกษาพบว่าการปรับปรุงภาพต้นฉบับด้วยวิธีการกำหนดเกณฑ์แบบปรับเหมาะ (Adaptive Thresholding) สำหรับกระบวนการแปลงภาพเป็นภาพขาวดำ (Image Binariztion) สามารถประมวลผลได้เร็วเหมาะแก่การพัฒนาบนอุปกรณ์พกพาโดยเฉพาะวิธีการของ Wellner ซึ่งเหมาะอย่างยิ่งสำหรับกรณีภาพที่มีปัญหาเงาปานกลางถึงรุนแรง และสามารถดัดแปลงเพื่อให้ใช้แก้ปัญหาแสงได้เช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 ตัวอย่างรูปที่ใช้ในการทดลอง

รูปที่ใช้ในการทดลองจำแนกออกตามนิยามปัญหาได้ ในการทดลองจำแนกออกตามนิยามปัญหาได้ 2 ประเภทคือ แสงและเงา และกรณีอื่น ๆ เพิ่มเข้ามาด้วย

ตารางที่ 1. การประเมินผลลัพธ์ของการประมวลผลในแอปพลิเคชัน(โดยเฉลี่ย)

Shadow	Lighting	Others

ตารางที่ 2. การจำแนกปัญหาของที่ปรากฏในแต่ละรูปภาพที่เป็นชุดทดสอบ

การประมวลผล	ตัวชี้วัดการประเมินผล		
	precision	recall	correctness
ก่อน	82.66%	74.44%	82.18%
หลัง	40.55%	30.64%	39.99%

ตารางที่ 3. ตารางแสดงรูปภาพกรณีที่ Recall สูงขึ้น

1.	3.

2.	4.

4.2.2 สรุปผลการทดลอง

ผลการทดลองและการเปรียบเทียบผลลัพธ์แสดงให้เห็นประสิทธิภาพโดยเฉลี่ยที่ลดลงจากการเพิ่มส่วนประมวลผลเบื้องต้น เนื่องมาจากการปรับใช้วิธีการของ Wellner ที่โดดเด่นในการแก้ปัญหาเงามีจุดบกพร่องกับกรณีอื่น ๆ แม้ปรับวิธีการก็ยังไม่สามารถกำจัดจุดบกพร่องที่ได้มาจากวิธีการของ Wellner แต่ในบางรูปภาพสามารถให้ผลลัพธ์ที่ดีขึ้นได้

ตารางที่ 4. ตารางแสดงผลลัพธ์กรณีที่ Recall สูงขึ้น

รูป		ตัวชี้วัดการประเมินผล		
		Precision (%)	Recall (%)	Correctness (%)
1	ก่อน	50	13.04	100
	หลัง	92	100	100
2	ก่อน	100	60	66.67
	หลัง	71.43	100	100
3	ก่อน	100	52.63	100
	หลัง	100	100	100
4	ก่อน	100	44.90	100
	หลัง	96.97	65.31	93.94

บทที่ 5 สรุปผลโครงงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

การศึกษาการระบบการรู้จำข้อความจากฉากธรรมชาติ ได้มีการนำไลบรารี OCR เข้ามาใช้แทนระบบการรู้จำข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้การประมวลผลบนอุปกรณ์พกพาประเภทสมาร์ทโฟนทำได้รวดเร็วยิ่งขึ้นและใช้ทรัพยากรที่มีจำกัดของอุปกณ์ได้อย่างมีประสิทธิภาพ

การศึกษาภาพการปรับปรุงต้นฉบับได้ใช้หลักการการประมวลผลรูปภาพ (Image Processing) เพื่อศึกษาทดลองวิธีการปรับปรุงภาพต้นฉบับก่อนเข้าสู่กระบวนการรู้จำข้อความ พบว่าการปรับปรุงภาพต้นฉบับต้องพัฒนาอัลกอริทึมเพิ่มเติมที่สามารถระบุวิธีการปรับปรุงภาพต้นฉบับที่จำเพาะกับภาพนั้นๆ เนื่องจากภาพถ่ายจากฉากธรรมชาติมีความหลากหลายมากมาย ทำให้กระบวนการปรับปรุงภาพต้นฉบับไม่สามารถใช้รูปแบบใดรูปแบบหนึ่งได้กับทุกภาพ

ข้อเสนอแนะ

1. สามารถต่อยอดพัฒนาลงบนอุปกรณ์อื่น ๆ นอกจากอุปกรณ์พกพาได้
2. สามารถต่อยอดพัฒนาให้ทำงานแบบ Real Time ได้
3. สามารถเพิ่มประสิทธิภาพการทำงานของแอปพลิเคชันให้ดีขึ้นได้โดยการปรับปรุงกระบวนการประมวลผลเบื้องต้นเพิ่มเติม

กระบวนการที่ได้ศึกษาและพัฒนามุ่งเน้นที่ความรวดเร็วในการประมวลผลซึ่งโดยปกติจะทำให้ประสิทธิผลลดลง การพัฒนาเพิ่มเติมควรเพิ่มประสิทธิผลโดยให้กระทบกับความเร็วในการประมวลผลน้อยที่สุด

เอกสารอ้างอิง

[1] Chucai Yi and Yingli Tian, "**Scenes Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration**", IEEE Transactions on Image Processing, Vol. 23, No. 7, July 2014.

[2] Derek Bradley and Gerhard Roth, "**Adaptive Thresholding Using the Integral**", Journal of Graphics, GPU, and Game Tools Volume 12, Issue 2, 2007.

[3] J. Sauvola and M. PietikaKinen, "**Adaptive Document Image Binarization**", Machine Vision and Media Processing Group, Infotech Oulu, University of Oulu, P.O. BOX 4500, FIN-90401 Oulu, Finland Received 29 April 1998; accepted 21 January 1999

[4] K. Peuwnuan, K. Woraratpanya, K. Pasupa, "**Modified Adaptive Thresholding Using Integral Image**", (unpublished).

[5] Pierre D. Wellner," **Adaptive Thresholding for the DigitalDesk**", Technical Report EPC-1993-110.

[6] Woraratpanya, K.Fac. of Inf. Technol., King Mongkut's Inst. of Technol. Ladkrabang, Bangkok, Thailand Pasupa, K. ; Suttapakti, U. ; Boonchukusol, P. ; Titijaroonroj, T. ; Hokking, R. ; Kuroki, Y. ; Kato, Y. , "**Text-background Decomposition for Thai Text Localization and Recognition in Natural Scenes**", Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on 7-8 Oct. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้