การแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านทางกล้องเว็บแคม

English to Thai word Translation via Webcam

ห้องสมุดคณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

นายยุทธพงษ์	สนั่นนาม	รหัส	51362138
นายพฤษภรรค	กาญจนโชติ	รหัส	51364873

ห้องสมุดคณะวิศวกรรมศาสตร์
วันที่รับ - 2 ต.ค. 2556
เลขทะเบียน 16280006
เลขเรียกหนังสือ ปร.
มหาวิทยาลัยนเรศวร ย356 ก 2554

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2554

ใบรับรองปริญญานิพนธ์

หัวข้อโครงงาน	การแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านทางกล้องเว็บแคม
ผู้ดำเนินโครงงาน	นายยุทธพงษ์ สนั่นนาม รหัส 51362138
	นายพฤษภรรค กาญจนโชติ รหัส 51364873
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2554

....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้ปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะกรรมการสอบโครงงานวิศวกรรม

.... ประธานกรรมการ
(อาจารย์รัฐภูมิ วรานุสาสน์)

.... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล)

.... กรรมการ
(อาจารย์จิราพร พุกสุข)

.... กรรมการ
(อาจารย์ภาณุพงศ์ สอนคม)

หัวข้อโครงงาน	การแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านทางกล้องเว็บแคม
ผู้ดำเนินโครงงาน	นายยุทธพงษ์ สนั่นนาม รหัส 51362138
	นายพฤษภรรค กาญจนโชติ รหัส 51364873
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2554

บทคัดย่อ

ปัจจุบันภาษาอังกฤษจัดได้ว่าเป็นภาษาสากลที่มักพบบ่อยในชีวิตประจำวัน การเข้าใจในความหมาย หรือความสามารถในการแปลคำศัพท์ภาษาอังกฤษจึงเป็นสิ่งสำคัญ ดังนั้นผู้จัดทำโครงงานจึงได้สร้างทางเลือกในการให้ความช่วยเหลือในการแปลคำศัพท์ภาษาอังกฤษ ผ่านทางคอมพิวเตอร์ ที่มีกล้องเว็บแคม

เมื่อมีการจับภาพผู้ใช้ต้องทำการกำหนดตำแหน่งพื้นที่ของภาพคำศัพท์ที่ต้องการ จากนั้นทำการแยกตัวอักษรภาพ และทำการเปรียบเทียบกับรูปแบบตัวอักษรที่ถูกเก็บไว้ภายในตัวโปรแกรม และจะนำไปเปรียบเทียบกับคำศัพท์ที่ถูกเก็บไว้ในฐานข้อมูลพจนานุกรม ถ้าพบจะแสดงคำศัพท์พร้อมกับความหมายของคำๆ นั้น แต่ถ้าไม่พบก็จะแสดงคำที่มีความใกล้เคียง ด้วยกระบวนการแนะนำคำศัพท์

ผลการทำงานของโปรแกรม พบว่าโปรแกรมสามารถใช้งานได้จริง อย่างไรก็ตามการผลลัพธ์จะขึ้นอยู่กับความสมบูรณ์ของภาพ ความคมชัดในการจับภาพและถ่ายภาพของกล้องเว็บแคม และความถูกต้องตามรูปแบบตัวอักษรที่ถูกเก็บไว้ภายในตัวโปรแกรม

Project Title	English to Thai word Translation via webcam.
Name	Mr. Yuttapong Sanannam ID. 51362138
	Mr. Prutsaprak Knajanachote ID. 51364873
Project Advisor	Assistant Professor Panomkhawn Riyamongkol, Ph.D.
Major	Computer Engineering.
Department	Electrical and Computer Engineering.
Academic year	2011

....

Abstract

Nowadays, English is the universal language that is common in daily life. Understanding of meaning and ability to translate words in English is essential. Therefore, this project has created an alternatives to assist in translating English words through the computer with a webcam.

When an image is captured user must define the area of the visual vocabulary. After that, the visual vocabulary area is separated into character images and then compared with the character templates that is stored within the program. The next step is comparing the output characters with words that are stored in the dictionary database. If it is found, the program will display word and meaning. Otherwise, the program will show close words by word suggestion.

The results of the program show that the program works properly. However, it depends on the integrity of the image, the contrast of the capturing and shooting of the camera, and how close to the template formats that is stored within the program.

กิตติกรรมประกาศ

โครงงานวิศวกรรมฉบับนี้ สำเร็จได้ด้วยดีก็เนื่องจากความอนุเคราะห์ของอาจารย์ที่ปรึกษาโครงงานคือ ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล รวมถึงอาจารย์ทุกท่านที่กรุณาสละเวลาอันมีค่าเพื่อชี้แนะแนวทาง ให้คำปรึกษา และให้ความช่วยเหลือ ในโอกาสนี้คณะผู้จัดทำโครงงานจึงขอขอบพระคุณทุกๆท่าน ที่มีส่วนช่วยให้โครงงานนี้ประสบความสำเร็จได้ด้วยดี สุดท้ายขอขอบพระคุณบิดา และมารดาของคณะผู้จัดทำที่เป็นผู้สนับสนุนในทุกๆด้านโดยเฉพาะด้านการศึกษาตลอดมา คณะผู้จัดทำขอนอบระลึกในพระคุณ และขอกราบขอบพระคุณมา ณ ที่นี้

นายพฤษภรรค กาญจนโชติ
นายยุทธพงษ์ สนั่นนาม

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญรูป	ฉ
สารบัญตาราง	ฌ
บทที่ 1 บทนำ	
1.1 ที่มาและความสำคัญของโครงงาน	1
1.2 วัตถุประสงค์ของโครงงาน	2
1.3 ขอบเขตของโครงงาน	2
1.4 แผนการดำเนินการ	3
1.5 ผลที่คาดว่าจะได้รับ	4
1.6 งบประมาณโครงงาน	4
บทที่ 2 หลักการและทฤษฎี	
2.1 การประมวลผลรูปภาพดิจิตอล (Digital Image Processing)	5
2.2 การแปลงภาพสีเป็นภาพระดับเทา	8
2.3 การปรับขนาดภาพ (Resize image)	8
2.4 การแยกภาพ (Image Segmentation)	9
2.5 การประมวลผลภาพกับรูปร่างและโครงสร้าง (Morphological Image Processing)	14
2.6 การรู้จำตัวอักษรโอซีอาร์ (Optical character recognition)	18
2.7 กระบวนการตรวจสอบตัวสะกด (Spell Check)	25
2.8 ฐานข้อมูลเบื้องต้น (Foundation of Database)	26
2.9 การติดต่อกับตัวกลางในการติดต่อกับฐานข้อมูล (ODBC)	29
บทที่ 3 ขั้นตอนการดำเนินงาน	
3.1 โครงสร้างของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม	30
3.2 ขั้นตอนการนำภาพเข้าสู่โปรแกรม	32

สารบัญ (ต่อ)

หน้า

3.3 ขั้นตอนการประมวลผลขั้นต้น (Preprocessing)....33

3.4 การรู้จำอักขระทางภาพ (Optical Character Recognition (OCR))....36

3.5 การตรวจสอบคำสะกดและค้นหาคำในฐานข้อมูล (Spell Check and Database))....41

3.6 การจัดเก็บข้อมูล....44

บทที่ 4 ผลการทดลอง

4.1 ส่วนของของโปรแกรม....48

4.2 ผลการทดลองส่วนของการเตรียมข้อมูลภาพ....49

4.3 ผลการทดลองส่วนของการหาตำแหน่งของคำที่ต้องการแปล....50

4.4 ผลการทดลองส่วนของการแปลงภาพเป็นตัวอักษร....51

4.5 ผลการทดลองในส่วนของการแนะนำคำใกล้เคียง....57

4.6 ผลการทดลองส่วนของการแปลคำศัพท์....60

บทที่ 5 สรุป

5.1 สรุปผลการทดลอง....62

5.2 ปัญหาและอุปสรรค....63

5.3 แนวทางการพัฒนาในอนาคต....63

เอกสารอ้างอิง....64

ภาคผนวก ก. การติดตั้งโปรแกรมและการใช้โปรแกรม....65

ภาคผนวก ข. ภาพรูปแบบ (Templates) ในโปรแกรม....71

ประวัติผู้เขียนโครงงาน....73

สารบัญรูป

รูปที่ หน้า

2.1 ภาพสี และค่าในแต่ละพิกเซล 5

2.2 คุณสมบัติของภาพระดับเทา 6

2.3 ภาพแบบไบนารี และค่าในแต่ละพิกเซล 7

2.4 เมตริกซ์ของรูปขนาด 256*256 7

2.5 แสดงวิธีการทำไบริเนียอินเทอโพเรชั่น (Bilinear Interpolation) 8

2.6 แสดงค่า k ตั้งแต่ 1 ถึง 255 10

2.7 แสดงค่า i ตั้งแต่ 1 ถึง k 11

2.8 แสดงค่า i ตั้งแต่ k+1 ถึง L 11

2.9 แสดงช่วงความแปรปรวนของกลุ่ม 12

2.10 แสดงค่าความแปรปรวมสูงสุด 13

2.11 แสดงการทำ Vertical Projection Profile เพื่อหาภาพตัวอักษร 14

2.12 (ก) แสดงรูปภาพเริ่มต้น (Original image) 15

2.12 (ข) แสดงรูปภาพย่อย (Structuring Element) 15

2.13 แสดงผลลัพธ์ที่ได้จากการย่อภาพ (Erosion) 15

2.14 (ก) แสดงรูปภาพเริ่มต้น (Original image) 16

2.14 (ข) แสดงรูปภาพย่อย (Structuring Element) 16

2.15 แสดงผลลัพธ์ที่ได้จากการทำขยายภาพ (Dilation) 16

2.16 (ก) แสดงรูปภาพเริ่มต้น (Original image) 17

2.16 (ข) แสดงรูปภาพย่อย (Structuring Element) 17

2.17 แสดงผลลัพธ์ที่ได้จากการทำการย่อภาพ (Erosion)) 17

2.18 แสดงผลลัพธ์ของการทำโอเปอเรชั่นโอเพนนิ่ง (opening) 17

2.19 (ก) ภาพ (template) รูป A จากภาพถ่ายและผ่านการปรับขนาด(resize)ให้เท่ากับภาพ (ข) 20

2.19 (ข) ภาพรูปแบบ (template) ของรูป B ที่ถูกเก็บไว้ภายในโปรแกรม 20

2.20 (ก) ภาพ (template) รูปA จากภาพถ่ายและผ่านการปรับขนาด (resize) ให้เท่าภาพ (ข) 22

2.20 (ข) แสดงภาพรูปแบบ (template) ของรูป B ที่ถูกเก็บไว้ภายในโปรแกรม 22

2.21 (ก) ภาพ (template) รูป A จากภาพถ่ายและผ่านการปรับขนาด (resize) ให้เท่ากับภาพ (ข) 24

2.21 (ข) ภาพรูปแบบ (template) รูป B ที่ถูกเก็บไว้ภายในโปรแกรม 24

สารบัญรูป (ต่อ)

รูปที่ หน้า

3.1 ภาพโดยรวมของขั้นตอนการดำเนินโครงการ (System Overview) 30
3.2 แสดงขั้นตอนการทำงานของโปรแกรม 31
3.3 ภาพย่อย (Sub Image)....................32
3.4 ขั้นตอนการทำงานของการประมวลผลขั้นต้น (Preprocessing)....................33
3.5 แปลงภาพสีแบบ RGB เป็นภาพแบบระดับเทา....................34
3.6 ภาพที่ผ่านกระบวนการเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ....................35
3.7 ภาพที่ผ่านกระบวนการ โอเพนนิ่ง (Opening)....................35
3.8 ขั้นตอนการรู้จำอักขระทางภาพ (OCR)....................36
3.9 แสดงการตัดพื้นที่ ให้เหลือเฉพาะพื้นที่ตัวอักษรที่ต้องการ)....................37
3.10 ขั้นตอนตัดพื้นที่บริเวณรอบคำ (Crop Image) 37
3.11 ภาพแสดงฮิสโทแกรมแนวแกนตั้งของตัวอักษร 38
3.12 แสดงการคำนวณหาค่าสัมประสิทธ์สหสัมพันธ์แบบง่าย....................39
3.13 แสดงการคำนวณหาค่าสัมประสิทธ์สหสัมพันธ์แบบง่ายๆ 40
3.14 ขั้นตอนตรวจสอบคำสะกดและค้นหาข้อมูลในฐานข้อมูล....................41
3.15 ภาพแสดงคำใกล้เคียงในกรณีจับภาพตกตัวอักษร y 42
3.16 ภาพแสดงโครงสร้างในฐานข้อมูล....................43
3.17 แสดงความสัมพันธ์ของ ER Diagram 44
3.18 แสดงการสร้างฐานข้อมูลใหม่....................45
3.19 แสดงการเลือก Design View....................45
3.20 แสดงหน้ากำหนดค่าเอนทิตี้ ของตาราง 46
3.21 แสดงเอนทิตี้ (Entity) ทั้งหมดในตาราง 46
3.22 แสดงการกลับไปที่ Datasheet View 47
3.23 แสดงการป้อนข้อมูลคำศัพท์....................47
4.1 แสดง GUI ของโปรแกรม 48
4.2 (ก) แสดงภาพที่ถ่ายด้วยกล้องเว็บแคมที่ไม่สามารถปรับโฟกัสของภาพได้....................49
4.2 (ข) แสดงภาพที่ถ่ายด้วยกล้องเว็บแคมที่สามารถปรับโฟกัสของภาพได้ 49
4.3 แสดงภาพที่ถูกจับภาพ (capture) แล้ว 50

สารบัญรูป (ต่อ)

รูปที่ หน้า

4.4 (ก) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างไม่เพียงพอ........................50

4.4 (ข) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างมากเกินไป........................50

4.4 (ค) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างพอดี.................................50

4.5 แสดงภาพที่ใช้ทดลองในการแปลงคำทั้งประโยค...54

4.6 แสดงกราฟผลลัพธ์ของความคมชัดของตัวอักษร 56

4.7 แสดงกราฟผลลัพธ์ของความเอียงของตัวอักษร 57

ก-1 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม 65

ก-2 ขั้นตอนการทำการจับภาพ..65

ก-3 ขั้นตอนการคลิก Select Word 66

ก-4 ขั้นตอนเลือกคำที่ต้องการ...66

ก-5 ขั้นตอนการทำให้คุณภาพของภาพดีขึ้น 67

ก-6 ขั้นตอนการคลิก Convert to Text...67

ก-7 ขั้นตอนการคลิก Translate เพื่อหาความหมาย...68

ก-8 แสดงผลลัพธ์ของความหมาย 68

ก-9 แสดงการเตือนเมื่อการแปลงภาพเป็นตัวอักษรเกิดความผิดพลาด 69

ก-10 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม..69

ก-11 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม..70

ก-12 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม..70

ข-1 แสดงตัวอักษรรูปแบบ A ไปจนถึง J 71

ข-2 แสดงตัวอักษรรูปแบบ K ไปจนถึง T 72

ข-3 แสดงตัวอักษรรูปแบบ K ไปจนถึง T 72

สารบัญตาราง

ตารางที่ หน้า

1.1 ตารางขั้นตอน และแผนการดำเนินโครงงาน 3
2.1 แสดงความหมายของค่าสัมประสิทธิ์สหสัมพันธ์....................19
3.1 ตารางแสดงรอบการแยกตัวอักษรโดยวิธีโปรเจคชั่นแนวตั้ง....................38
3.2 ตารางการระบุตัวอักษรจากค่าสูงสุดของค่าสัมประสิทธ์สหสัมพันธ์....................40
4.1 ทดสอบความสามารถในการระบุอักษรภาษาอังกฤษตัวพิมพ์ใหญ่....................51
4.2 ทดสอบความสามารถในการระบุอักษรภาษาอังกฤษตัวพิมพ์เล็ก....................52
4.3 แสดงการเปรียบเทียบการแปลงภาพเป็นตัวอักษรโดยใช้รูปแบบอักษรที่ต่างกัน 53
4.4 การทดลองแปลงภาพเป็นตัวอักษรทั้งประโยค....................55
4.5 การทดลองแปลงภาพเป็นตัวอักษรในมุมเอียงที่ต่างกัน 56
4.6 แสดงการแนะนำคำใกล้เคียงของคำว่า chapter 58
4.7 แสดงการแนะนำคำใกล้เคียงของคำว่า control....................58
4.8 แสดงการแนะนำคำใกล้เคียงของคำว่า detail 58
4.9 แสดงการแนะนำคำใกล้เคียงของคำว่า stability 59
4.10 แสดงการแนะนำคำใกล้เคียงของคำว่า analysis 59
4.11 แสดงการแนะนำคำใกล้เคียงของคำว่า dynamic 59
4.12 แสดงการแนะนำคำใกล้เคียงของคำว่า transient 60
4.15 แสดงความหมายของคำเมื่อกระบวนการแปลงภาพเป็นตัวอักษรถูกต้อง 61

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของโครงงาน

เนื่องจากปัจจุบันภาษาอังกฤษได้รับการยอมรับให้เป็นภาษาสากลช่วยส่งเสริมการเรียนรู้ แบ่งปันประสบการณ์กับคนต่างเชื่อชาติและการเรียนการสอนในสถานศึกษาต่างๆได้มีการนำเอกสารหรือหนังสือเรียนที่เป็นภาษาอังกฤษมาใช้ ซึ่งคำศัพท์เป็นสิ่งสำคัญอย่างยิ่งที่จะช่วยให้สามารถเข้าใจความหมายในบทเรียนที่เรียนได้

การหาความหมายของคำศัพท์นั้นอาจมีได้หลายวิธี ไม่ว่าจะเป็นใช้พจนานุกรมอังกฤษ-ไทยที่เป็นเล่ม อินเตอร์เน็ตเป็นต้น ซึ่งเป็นวิธีที่ง่ายที่สุด เพื่อเป็นอีกหนึ่งทางเลือกสำหรับผู้ใช้ในการค้นหาความหมายของคำศัพท์ โดย อาศัยคอมพิวเตอร์ หรือ แลปท็อป ที่มีกล้องเว็บแคม จึงเป็นที่มาของแนวคิดการแปลคำศัพท์ผ่านกล้องเว็บแคมเพื่อช่วยในการค้นหาความหมายของคำศัพท์ในเอกสารการเรียนการสอนภาษาอังกฤษ

ดังนั้น ผู้จัดทำโครงงานจึงได้ทำการศึกษาหลักการประมวลผลภาพต่างๆเพื่อนำมาประยุกต์ใช้ในการสร้างโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านทางกล้องเว็บแคมเพื่ออำนวยความสะดวกและเป็นอีกทางเลือกให้แก่ผู้ใช้ในการค้นหาความหมายของคำศัพท์ภาษาอังกฤษ

1.2 วัตถุประสงค์ของโครงงาน

1.2.1 เพื่อเป็นทางเลือกให้กับผู้ใช้ในการหาคำศัพท์นอกจากเล่มพจนานุกรมหรืออินเตอร์เน็ต

1.2.2 เพื่อสร้างโปรแกรมแปลภาษาที่ผู้ใช้สามารถใช้ง่ายได้ง่ายและเป็นประโยชน์

1.3 ขอบเขตของโครงงาน

1.3.1 ตัวอักษรที่ใช้จำนวน 52 ตัว ประกอบด้วย

- ตัวอักษรภาษาอังกฤษตัวพิมพ์ใหญ่จำนวน 26 ตัว
- ตัวอักษรภาษาอังกฤษตัวพิมพ์เล็กจำนวน 26 ตัว

1.3.2 รูปแบบอักษรที่ใช้คือ Times New Roman ขนาด 16 หน่วยขึ้นไป

1.3.3 ตัวอักษรมีสีดำ, ไม่เอียง, ไม่มีเส้นใต้ และสัญลักษณ์พิเศษอื่นๆ

1.3.4 ฐานข้อมูลคำศัพท์อังกฤษ-ไทยของ NECTEC (LEXITRON)

1.4 แผนการดำเนินงาน

ตารางที่ 1.1 ตารางขั้นตอนและแผนการดำเนินโครงงาน

กิจกรรม	ปี 2554							ปี 2555	
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.
1.ศึกษาค้นคว้าข้อมูลในการทำโครงงาน	↔								
2.ศึกษาโครงสร้างภาษาแมทแลบ(MATLAB)	↔	↔							
3.ศึกษาข้อมูลของการรู้จำอักขระทางภาพและข้อมูลต่างๆ เกี่ยวกับการประมวลผลภาพ		↔	↔						
4.ศึกษาการเชื่อมโยงโปรแกรมกับฐานข้อมูล			↔	↔					
5.ทำการออกแบบโปรแกรมและกำหนดส่วนประกอบ				↔	↔				
6.เขียนโปรแกรม						↔	↔	↔	
7.ทดสอบการใช้งานและแก้ไขข้อบกพร่อง								↔	
8.สรุปผลการทำโครงงานและจัดทำรายงาน									↔

1.5 ผลที่คาดว่าจะได้รับ

1.5.1 ผู้ใช้งานสามารถใช้โปรแกรมเพื่อหาความหมายของคำศัพท์ได้สะดวก

1.6 งบประมาณของโครงงาน

1.6.1 ค่าอุปกรณ์ในการดำเนินโครงงาน	1,000 บาท
1.6.2 ค่าเอกสารที่ใช้ในการดำเนินโครงงาน	200 บาท
1.6.3 ค่าหมึกพิมพ์	300 บาท
1.6.4 ค่าเข้าเล่มโครงงาน	500 บาท
รวมเป็นเงินทั้งสิ้น	2,000 บาท

<u>หมายเหตุ</u> ขออนุมัติถัวเฉลี่ยทุกรายการ

มหาวิทยาลัยนเรศวร

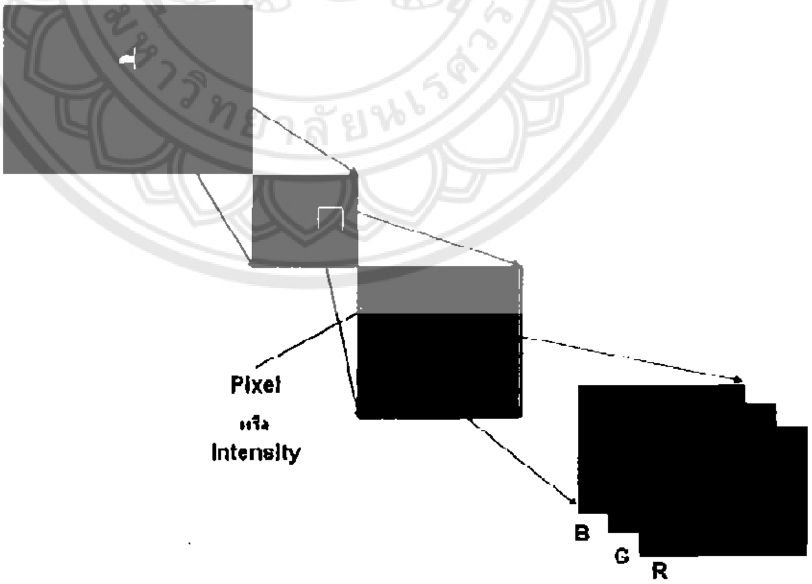
บทที่ 2
หลักการและทฤษฎีที่เกี่ยวข้อง

2.1 การประมวลผลรูปภาพดิจิตอล (Digital Image Processing)[1]

การประมวลผลรูปภาพดิจิตอล คือ การใช้อัลกอริทึมของคอมพิวเตอร์เพื่อใช้ในการประมวลผลรูปภาพ บนรูปภาพดิจิตอลที่เกี่ยวกับการประมวลผลภาพที่เกี่ยวกับการคำนวณ และการจัดการสีในระดับพิกเซล หรือรูปแบบของสี มีการเปลี่ยนแปลงของสีได้ เพื่อวัตถุประสงค์ที่ต้องการ

2.1.1 ภาพสี (Color Image)

ค่าในแต่ละพิกเซลของภาพสี จะประกอบไปด้วยเวกเตอร์ที่แสดงถึงค่าของสีแดง สีเขียว และสีน้ำเงิน อย่างละ 8 บิต (สีแต่ละสีมีความเข้มแสง 0-255 ค่า) ดังนั้น ภาพสี 1 พิกเซล จะประกอบไปด้วยจำนวนบิตทั้งหมด 24 บิต ทำให้ภาพสีมีจำนวนสีที่เป็นไปได้ทั้งหมด 2^{24} สี ดังรูปที่ 2.1

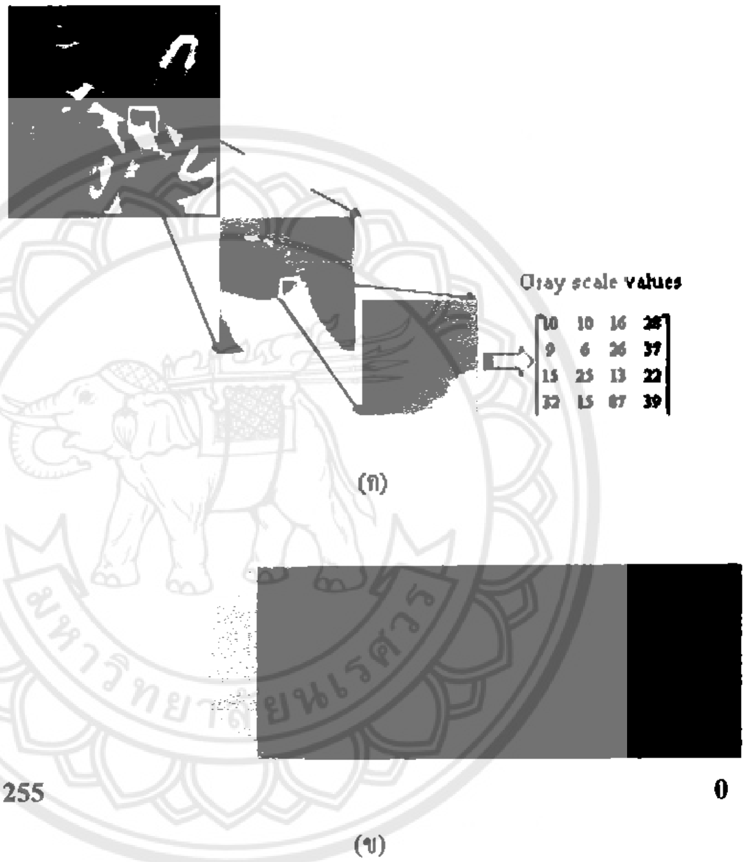


รูปที่ 2.1 ภาพสี และค่าในแต่ละพิกเซล

มหาวิทยาลัยนเรศวร

2.1.2 ภาพระดับเทา (Gray Image)

ค่าในแต่ละพิกเซลของภาพระดับเทา คือค่าความเข้มของแสง ณ แต่ละตำแหน่งของพิกเซล ซึ่งจะอยู่ในรูปของค่าระดับความเข้ม ดังรูปที่ 2.2 (ก) ค่าที่เป็นไปได้ของค่าระดับความเข้มจะขึ้นอยู่กับจำนวนบิตที่ใช้ ตัวอย่างเช่น ภาพระดับเทาจะมีค่าระดับความเข้ม 8 บิต หรือทั้งหมด 256 ระดับ ดังรูปที่ 2.2 (ข)



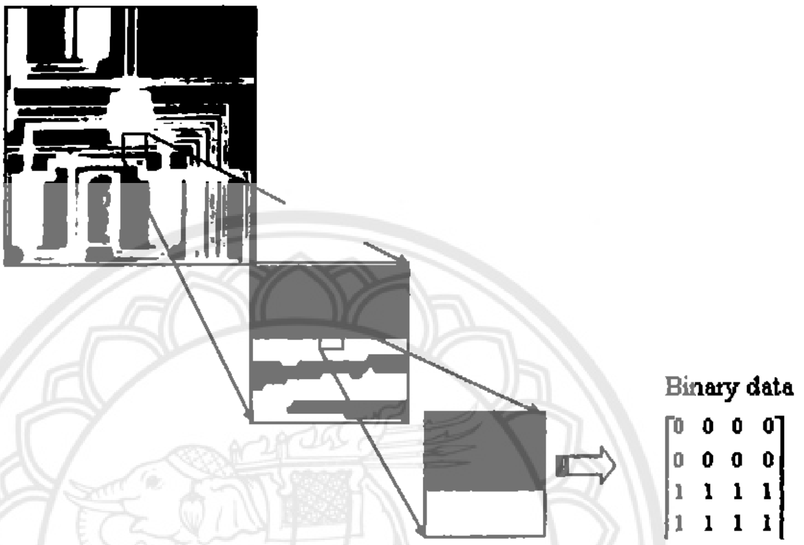
รูปที่ 2.2 คุณสมบัติของภาพระดับเทา

(ก) ภาพระดับเทา และค่าในแต่ละพิกเซล

(ข) ค่าระดับความเข้ม 8 บิต แสดงการไล่ระดับเฉดสีขาวที่สุดไปดำที่สุด (Gray-Scale)

2.1.3 ภาพแบบไบนารี (Binary Image)

ภาพแบบไบนารี คือ ภาพที่มีระดับความเข้มเพียง 1 บิต หรือ 2 ระดับ คือ ในหนึ่งพิกเซลมีค่าเพียงสองค่าคือ 0 และ 1 เท่านั้น โดย พิกเซลที่มีค่าเท่ากับ 1 จะมีสีขาว และ พิกเซลที่มีค่าเท่ากับ 0 จะมีสีดำ ดังรูปที่ 2.3



รูปที่ 2.3 ภาพแบบไบนารี และค่าในแต่ละพิกเซล

2.1.4 ลักษณะของแฟ้มการเก็บภาพให้อยู่ในระบบข้อมูลดิจิตอล

ลักษณะของแฟ้มข้อมูลภาพจะมีการกำหนดตำแหน่งเหมือนกับ**เมตริกซ์โดยที่จำนวนคอลั**มน์ (Column) และแถว (Row) จะแทนจำนวนจุดภาพของภาพ เช่น ภา**พขนาด** 256*256 **จะหมาย**ถึงแฟ้มข้อมูลของภาพดังกล่าวมีขนาด 256 คอลัมน์และ 256 แถว

$$A = \begin{bmatrix} f(1,1) & f(1,2) & f(1,3) & \text{----} & f(1,256) \\ f(2,1) & f(2,2) & f(2,3) & \text{-----} & f(2,256) \\ f(3,1) & f(3,2) & f(3,3) & \text{----} & f(3,256) \\ - & - & - & & - \\ f(256,1) & f(256,2) & f(256,3) & \text{-----} & f(256,256) \end{bmatrix}$$

รูปที่ 2.4 เมตริกซ์ของรูปขนาด 256*256

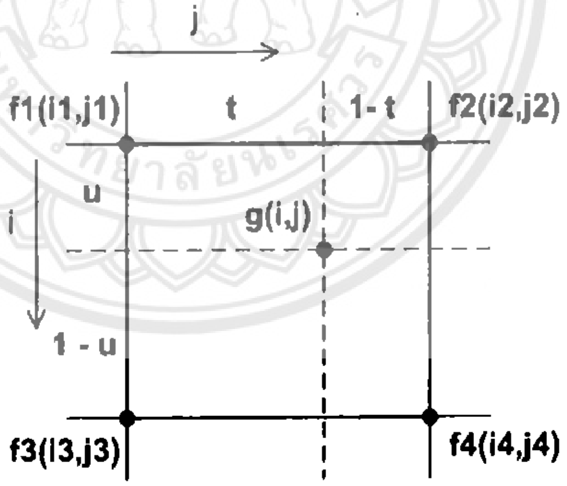
จากรูปที่ 2.4 จะเห็นว่าแฟ้มข้อมูลภาพจะประกอบไปด้วยสมาชิกของเมตริกซ์ทั้งหมด i * j ตัว และที่ตำแหน่งในเมตริกซ์ต่างๆ ก็จะมี f(i,j) ซึ่งเป็นค่าความเข้มแสง (Gray Level) ของภาพที่จุดนั้น

2.2 การแปลงภาพสีเป็นภาพระดับเทา[2]

ภาพที่มนุษย์มองเห็นโดยทั่วไป จัดเป็นภาพสี ซึ่งคุณสมบัติภาพสีนี้จะมี 3 ระนาบ ทำให้การวิเคราะห์ภาพมีความซับซ้อน และยากต่อการพิจารณา การลดทอนความเป็นสีสามารถทำได้โดยการแปลงให้เป็นภาพระดับเทา ขั้นตอนดังกล่าวนั้นมีหลายวิธี เช่น เลือกระนาบใดระนาบหนึ่งจากภาพสี หรือ นำทั้ง 3 ระนาบมาทำการหาค่าเฉลี่ยเพื่อให้เหลือเพียงระนาบเดียว เป็นต้น ซึ่งภาพระดับเทาที่ได้จะซับซ้อนน้อยกว่าภาพสีเมื่อนำมาวิเคราะห์วัตถุต่างๆ ภายในภาพ

2.3 การปรับขนาดภาพ (Resize image)[3]

ในโครงงานนี้ มีการใช้ขั้นตอนการปรับขนาดของภาพ (resize) ด้วยการใช้เทคนิคไบริเนียอินเทอโพเรชั่น (Bilinear Interpolation) ซึ่งเป็นการกำหนดค่าพิกเซลในตำแหน่งใหม่ โดยใช้ค่าพิกเซลหรือค่าความเข้มแสง (intensity) เฉลี่ยของ 4 พิกเซลที่อยู่รอบซึ่งสามารถอธิบายดังรูปที่ 2.5



รูปที่ 2.5 แสดงวิธีการทำไบริเนียอินเทอโพเรชั่น (Bilinear Interpolation)

จากรูปที่ 2.5 จะสามารถสร้างสมการได้ดังสมการที่ 2.1

$$g(i,j) = (1-t)*(1-u)*f1(i1,j1) + t*(1-u)*f2(i2,j2) + (1-t)*u*f3(i3,j3) + t*u*f4(i4,j4) \quad (2.1)$$

2.4 การแยกภาพ (Image Segmentation)[4]

การแยกภาพ คือ กระบวนการแยกองค์ประกอบ หรือแยกวัตถุออกจากภาพ มักจะทำหลังจากปรับปรุงคุณภาพของภาพให้ดีขึ้นแล้ว ซึ่งมีความสำคัญมาก เพราะการวิเคราะห์จะไม่สามารถกระทำกับภาพทั้งภาพเพื่อให้ได้ผลลัพธ์ที่ดีได้ เพราะเนื่องจากจะช้า และจำนวนข้อมูลในการวิเคราะห์มาก แต่จะวิเคราะห์ภาพเป็นส่วนๆ ไป เช่น การหาวัตถุที่ควรจะเป็นตัวอักษร ว่าอยู่ตำแหน่งใดภายในภาพ เป็นต้น ซึ่งการแยกภาพนี้เป็นขั้นตอนสำคัญของการวิเคราะห์ภาพ

2.4.1 ประโยชน์ของการแยกภาพ

2.4.1.1 ลดจำนวนข้อมูลในรูปภาพที่ไม่จำเป็นในการวิเคราะห์ลง เนื่องจากการทำการแยกภาพ เป็นการแยกแยะระหว่างส่วนที่สนใจ เช่น ตัวอักษร กับส่วนที่ไม่ต้องการนั่นคือ ฉากหลัง เมื่อตัดข้อมูลในส่วนที่ไม่ต้องการออกไป จำนวนข้อมูลที่จำเป็นในการวิเคราะห์จะลดลงอย่างมาก และยังทำให้การประมวลผลไวขึ้นอีกด้วย

2.4.1.2 จัดระเบียบข้อมูลในรูปภาพให้เป็นกลุ่มได้ดีขึ้น ข้อมูลภาพที่ผ่านการแบ่งแยกแล้ว จะมีโครงสร้างที่ชัดเจนขึ้นและนำไปใช้งานได้สะดวกขึ้น

2.4.1.3 แสดงข้อมูลในรูปแบบที่เข้าใจง่ายยิ่งขึ้น ซึ่งข้อมูลภาพที่ผ่านการแบ่งแยกแล้วจะทำให้สามารถแยกแยะความแตกต่างและความเหมือนของรูปภาพได้

การทำการแยกภาพ เพื่อแยกข้อมูลภาพในส่วนที่ต้องการออกมาได้ วิธีการที่ใช้สำหรับการแยกภาพ ในโครงงานนี้ คือ การพิจารณาความสว่างของภาพสำหรับภาพระดับเทาด้วยการสร้างภาพแบบไบนารีด้วยการทำเทรชโฮลด์ (Thresholding) โดยอธิบายได้ดังหัวข้อที่ 2.3.2

2.4.2 การสร้างภาพแบบไบนารีด้วยการทำเทรชโฮลด์ (Thresholding)[6]

ในการสร้างภาพแบบไบนารีด้วยวิธีการทำเทรชโฮลด์นั้น มีหลักการพิจารณาคือ จะพิจารณาจุดของภาพว่าจุดใดควรจะเป็นจุดดำหรือจุดขาว โดยเปรียบเทียบระหว่างค่าในพิกเซลเริ่มต้นกับค่าคงที่ค่าหนึ่งหรือค่าเทรชโฮลด์ นั่นเอง ซึ่งวิธีเทรชโฮลด์ นิยมใช้มากในกรณีที่วัตถุ และพื้นหลัง ของภาพมีความแตกต่างกันอย่างชัดเจน

บางครั้งการเลือกค่าเทรชโฮลด์ที่ไม่เหมาะสม คือ เลือกค่าเทรชโฮลด์ที่มากหรือน้อยจนเกินไป จะทำให้ผลลัพธ์ที่ได้ อาจจะขาดความคมชัดและรายละเอียดบางส่วนขาดหายไป กล่าวคือภาพที่ได้อาจจะมืดเกินไป หรือสว่างจนเกินไป หรือบางทีอาจมีสิ่งรบกวนเกิดขึ้น ดังนั้นจะแก้ปัญหาเหล่านี้ด้วยการเลือกค่าเทรชโฮลด์ที่เหมาะสมด้วยวิธีการของออทซึ (Otsu's Thresholding method) ซึ่งเป็นวิธีที่นิยมมากที่สุด อีกทั้งยังนำมาใช้ในแมทแลป (MATLAB) ด้วย โดยทั่วไปแล้ว

ตั้งอยู่บนสมมุติฐานที่ว่าความเข้มแสงของบริเวณที่เป็นวัตถุที่สนใจ และบริเวณที่เป็นฉากหลังมีความแตกต่างกันพอประมาณ ซึ่งค่าเทรชโฮลด์ที่เลือกใช้นั้น จะต้องสามารถแบ่งวัตถุและฉากหลังออกจากกันได้เป็นอย่างดี

2.4.2.1 การหาค่าเทรชโฮลด์ด้วยวิธีการของออทซึ (Otsu's Thresholding method)

วิธีการของออทซึเป็นวิธีการหาค่าความแปรปรวนระหว่างกลุ่ม โดยใช้ k เป็นค่าในการแบ่งกลุ่ม เริ่มจาก k =1 ถึง k=255 ดังรูปที่ โดยมีสูตรการคำนวณดังสมการที่ 2.2

รูปที่ 2.6 แสดงค่า k ตั้งแต่ 1 ถึง 255

$$\sigma_B^2 = \omega_0\omega_1(\mu_0 - \mu_1)^2 \quad (2.2)$$

$$P_i \geq 0, \quad \sum_{i=1}^{L} P_i = 1 \quad (2.3)$$

P_i คือ การ Normalization ค่าให้อยู่ระหว่าง 0 ถึง 1

P_i คือ จำนวนพิกเซล ณ ระดับที่ i

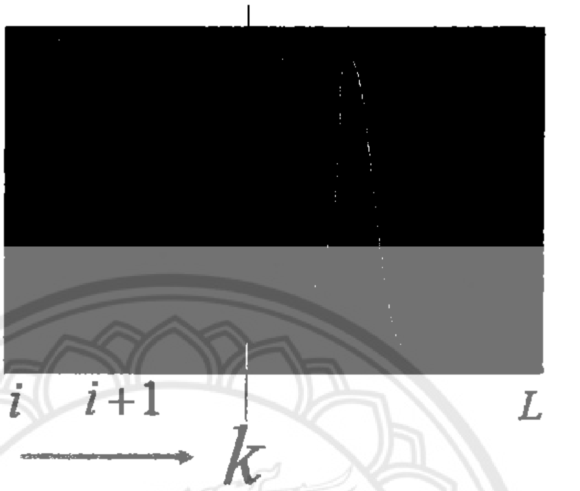
N คือ ผลรวมของจำนวนพิกเซลในแต่ละระดับ $N = n_1 + n_2 + ... + n_L$

L คือ พิกเซลของภาพสีเทา $\{1,2,...,L\}$

$$\omega_0 = \sum_{i=1}^{K} P_i = \omega(k) \quad (2.4)$$

ω_0 คือ ค่าความน่าจะเป็นของกลุ่มที่ 1

$\sum_{i=1}^{K} P_i$ คือ ผลรวมของ P_i ตั้งแต่ i = 1 ถึง k

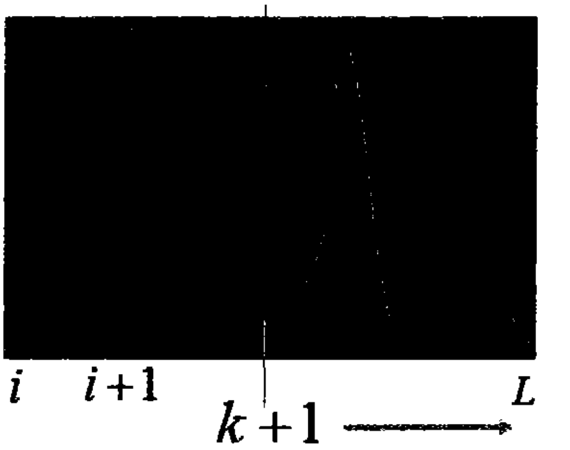


รูปที่ 2.7 แสดงค่า i ตั้งแต่ 1 ถึง k

$$\omega_1 = \sum_{i=k+1}^{L} P_i \qquad (2.5)$$

ω_1 คือ ค่าความน่าจะเป็นกลุ่มที่ 2

$\sum_{i=k+1}^{L} P_i$ คือ ผลรวมของ P_i ตั้งแต่ $i = K + 1$ ถึง L



รูปที่ 2.8 แสดงค่า i ตั้งแต่ k+1 ถึง L

$$\mu_0 = \mu(k) / \omega_0 \tag{2.6}$$

โดย μ_0 คือ ค่าเฉลี่ยของกลุ่มที่ 1

$\mu_k = \sum_{i=1}^{K} ip_i$ คือ ผลรวมของ $i * P_i$ ตั้งแต่ $i = 1$ ถึง k

$$\mu_1 = \sum_{i=K+1}^{L} ip_i / \omega_1 \tag{2.7}$$

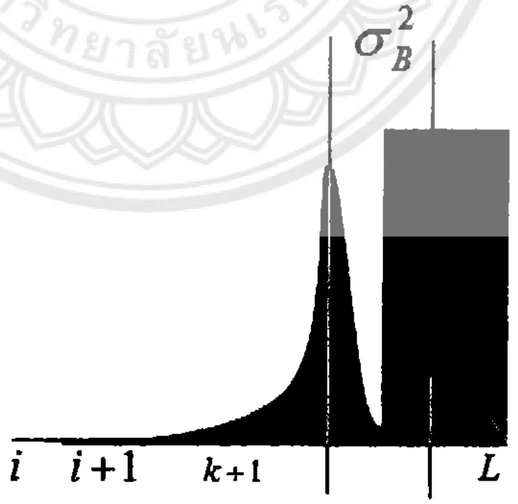
โดย μ_1 คือ ค่าเฉลี่ยของกลุ่มที่ 2

$\sum_{i=K+1}^{L} ip_i / \omega_1$ คือ ผลรวมของ $i * P_i / \omega_1$ ตั้งแต่ $i = k+1$ ถึง L

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_0 - \mu_1)^2 \tag{2.8}$$

σ_B^2 คือ ค่าความแปรปรวนของกลุ่มตั้งแต่รอบที่ 1 – 255

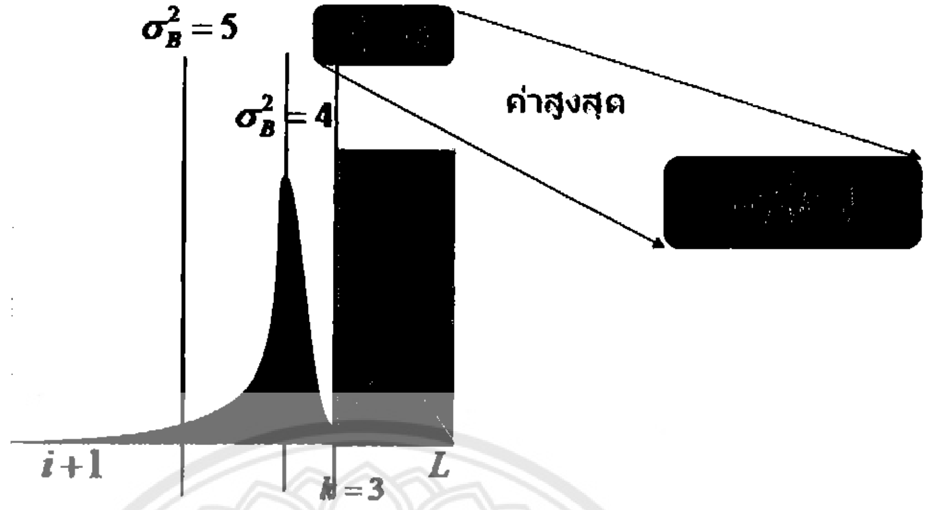
โดย $k = 1 - 255$



รูปที่ 2.9 แสดงช่วงความแปรปรวนของกลุ่ม

$$\sigma_B^2(k^*) = \max_{1 \le k < L} \sigma_B^2(k) \tag{2.9}$$

$\sigma_B^2(k^*)$ คือ ค่าความแปรปรวนสูงสุดของกลุ่ม ตั้งแต่รอบที่ k = 1 - 255



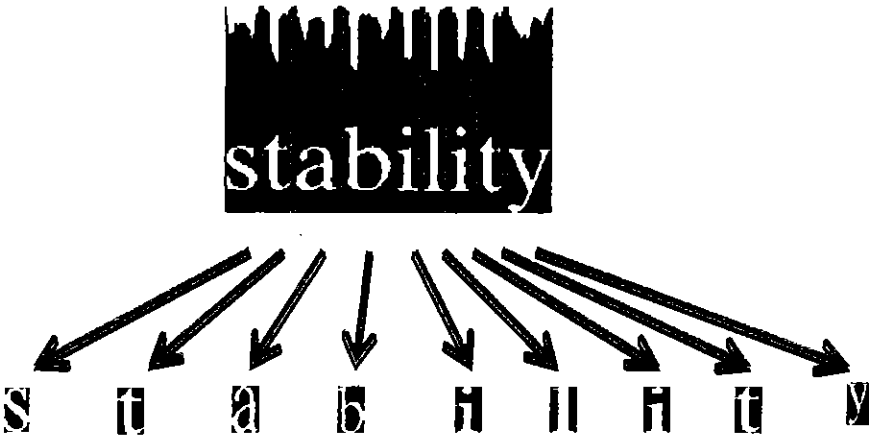
รูปที่ 2.10 แสดงค่าความแปรปรวมสูงสุด

กำหนดให้ค่า Threshold คือ ค่า K ที่ทำให้เกิดค่าความแปรปรวนระหว่างกลุ่มสูงสุด T = K และเปลี่ยนค่าพิกเซล ดังสมการที่ 2.10

$$f_{thr}(i,j) = \begin{cases} 1 & if \quad f(i,j) < T \\ 0 & if \quad f(i,j) \geq T \end{cases} \tag{2.10}$$

2.4.3 การแยกภาพตัวอักษร (Character Segmentation)

เป็นการทำงานเพื่อดึงข้อมูลของตัวอักษรแต่ละตัวออกมาจากข้อความโดยการดูค่าจากฮิสโทแกรม แล้วพิจารณาหาช่องว่างของตัวอักษร ในประโยคใช้การทำโปรเจคชั่นในแนวแกนตั้ง (Vertical Projection Profile) โดยแผนผังฮิสโทแกรมจะแสดงจำนวนของพิกเซลที่มีสถานะเป็น ON ในแต่ละสดมภ์ของข้อความในบรรทัดนั้น ซึ่งค่าแอมปลิจูด จะมีค่าต่ำระหว่างตัวอักษร มีค่าสูงภายในตัวอักษร และมีค่าสูงสุดตามความสูงของตัวอักษรที่สูงสุดในคำนั้น สำหรับค่าแอมปลิจูด ที่ไม่มีสัญญาณรบกวนค่าความสูงที่เป็น 0 จะบอกถึงตำแหน่งของช่องว่างระหว่างตัวอักษร แต่ถ้าภาพที่มีสัญญาณรบกวนอาจทำให้ตัวอักษร 2 ตัวติดกันได้ หรือ ตัวอักษรตัวเดียวขาดจากกันได้ ดังรูปที่ 2.11



รูปที่ 2.11 แสดงการทำโปรเจคชั่นในแนวแกนตั้ง เพื่อหาภาพตัวอักษร

2.5 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ (Morphological Image Processing)[5,7]

การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ คือ การประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ โอเปอเรชั่นพื้นฐานโดยทั่วไปมักจะกระทำกับภาพแบบไบนารี แต่ที่กระทำกับภาพระดับเทานั้นก็มี ส่วนในโครงงานนี้จะกล่าวถึงแค่ส่วนที่กระทำกับภาพแบบไบนารีเท่านั้น มีหลายโอเปอเรชั่น ได้แก่

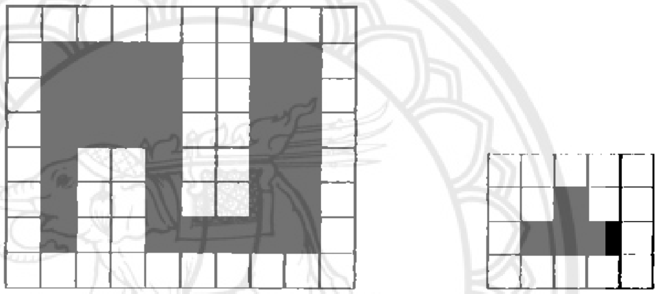
- การขยายภาพ (Dilation) คือ การขยายภาพโดยมีสัดส่วนเท่ากันทั่วทั้งภาพ
- การย่อภาพ (Erosion) คือ การย่อภาพโดยมีสัดส่วนเท่ากันทั่วทั้งภาพ
- การโครซซิ่ง (Closing คือ การขยายภาพ (Dilation) แล้วตามด้วยการทำย่อภาพ (Erosion)
- การโอเพนนิ่ง (Opening) คือ การทำการย่อภาพ (Erosion) แล้วตามด้วยการขยายภาพ (Dilation)
- การทำให้ภาพบาง (Thinning) คือ การลบขอบของวัตถุ

ในโครงงานนี้จะกล่าวถึง เฉพาะกระบวนการย่อภาพ (Erosion) การขยายภาพ (Dilation) และการทำโอเพนนิ่ง(Opening) เท่านั้น

2.5.1 การย่อภาพ (Erosion)

การย่อภาพเป็นลักษณะของการลบข้อมูลภาพบริเวณขอบของภาพ การย่อภาพสามารถทำได้มีลักษณะคล้ายกับการขยายภาพโดยการสร้างรูปภาพย่อย (Structuring Element) ขึ้นแล้วนำรูปภาพย่อย (Structuring Element) ไปสแกนตามข้อมูลภาพ

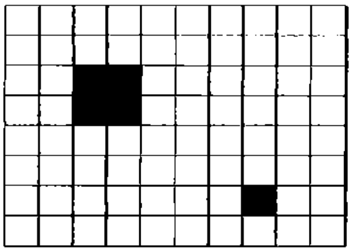
สำหรับทุกตำแหน่งที่เลื่อนรูปภาพย่อย (Structuring Element) ไปบนภาพก็จะมีการเปรียบเทียบกับข้อมูลภาพ ถ้าข้อมูลภาพมีค่าเหมือนกับรูปภาพย่อย (Structuring Element) จะทำการกำหนดค่าข้อมูลภาพในตำแหน่งที่ตรงกับจุดเริ่มต้น (Origin) ของรูปภาพย่อย (Structuring Element) ถูกกำหนดให้มีค่าเท่ากับ 1



(ก) (ข)

รูปที่ 2.12 (ก) แสดงรูปภาพเริ่มต้น (Original image)
(ข) แสดงรูปภาพย่อย (Structuring Element)

ซึ่งผลที่ได้จะมีเพียง 5 ตำแหน่งเท่านั้นที่มีค่าเหมือนกับรูปแบบ (template)

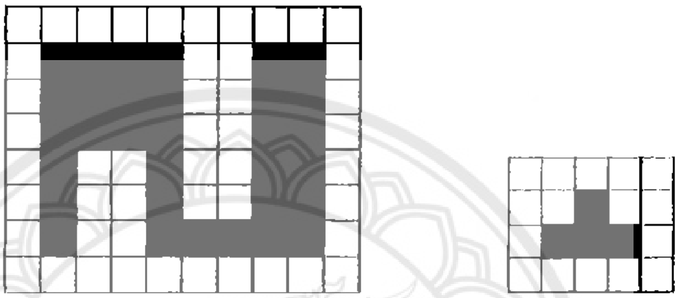


รูปที่ 2.13 แสดงผลลัพธ์ที่ได้จากการย่อภาพ (Erosion)

ผลที่ได้ตามรูปที่ 2.13 ข้อมูลภาพที่ผ่านการทำโอเปอเรชั่นกับรูปภาพย่อย (Structuring Element) แล้วพบว่ารูปภาพย่อย (Structuring Element) จะเป็นตัวกำหนดขนาดของผลลัพธ์ที่ได้

2.5.2 การขยายภาพ (Dilation)

การขยายภาพจะทำได้โดยกำหนดรูปแบบ (template) และนำรูปภาพย่อย (Structuring Element) นี้สแกนไปบนข้อมูลภาพตามลำดับตลอดทั้งภาพซึ่งในขณะที่จุดเริ่มต้น (Origin) ของรูปภาพย่อย (Structuring Element) ตรงกับตำแหน่งข้อมูลภาพที่พิกเซลมีค่าเท่ากับ 1 นั้นก็จะทำการยูเนียนรูปภาพย่อย (Structuring Element) นี้เข้ากับข้อมูลภาพดังตัวอย่าง

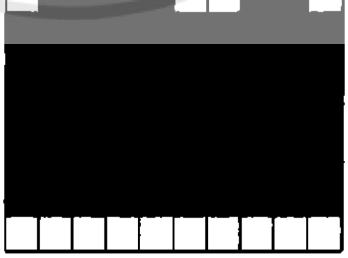


(ก) (ข)

รูปที่ 2.14 (ก) แสดงรูปภาพเริ่มต้น (Original image)

(ข) แสดงรูปภาพย่อย (Structuring Element)

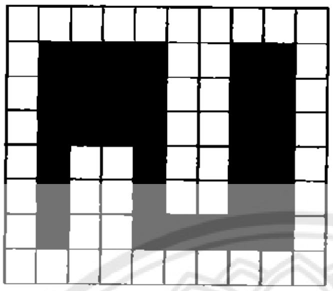
ซึ่งผลลัพธ์ที่ได้จะเป็นไปตามรูปที่ 2.15



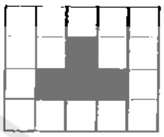
รูปที่ 2.15 แสดงผลลัพธ์ที่ได้จากการทำขยายภาพ (Dilation)

2.5.3 การทำโอเปอเรชั่นโอเพนนิ่ง (Opening)

การทำโอเปอเรชั่นโอเพนนิ่ง คือการนำข้อมูลภาพ ผ่านการทำย่อภาพ (Erosion) แล้วตามด้วยการขยายภาพ (Dilation) โดยใช้ รูปภาพย่อย (Structuring Element) ชุดเดียวกัน



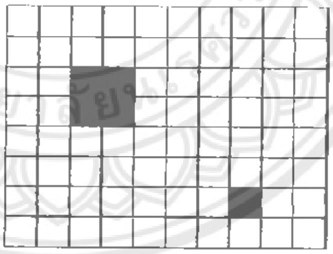
(ก)



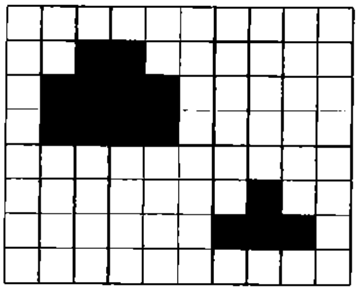
(ข)

รูปที่ 2.16 (ก) แสดงรูปภาพเริ่มต้น (Original image)

(ข) แสดงรูปภาพย่อย (Structuring Element)



รูปที่ 2.17 แสดงผลลัพธ์ที่ได้จากการทำการย่อภาพ (Erosion)



รูปที่ 2.18 แสดงผลลัพธ์ของการทำโอเปอเรชั่นโอเพนนิ่ง (opening)

2.6 การรู้จำตัวอักษรโอซีอาร์ (Optical character recognition)

OCR (Optical character recognition) การรู้จำอักขระทางภาพ (อังกฤษ: optical character recognition) หรือมักเรียกอย่างย่อว่า โอซีอาร์ (อังกฤษ: OCR) คือกระบวนการทางกลไกหรือทางอิเล็กทรอนิกส์เพื่อแปลภาพของข้อความจากการเขียนหรือจากการพิมพ์ ไปเป็นข้อความที่สามารถแก้ไขได้โดยเครื่องคอมพิวเตอร์ การจับภาพอาจทำโดยเครื่องสแกนเนอร์ กล้องดิจิทัล

โอซีอาร์เป็นสาขาวิจัยในการรู้จำแบบ, ปัญญาประดิษฐ์, และคอมพิวเตอร์วิทัศน์ แม้การวิจัยเชิงวิชาการในสาขายังคงดำเนินอยู่ แต่จุดเน้นในสาขาโอซีอาร์ได้เปลี่ยนไปสู่การสร้างระบบที่ใช้ได้จริงจากเทคนิคที่พิสูจน์แล้ว การรู้จำอักขระทางแสง (optical character recognition การใช้เทคนิคทางแสง เช่นกระจกและเลนส์) การรู้จำอักขระทางดิจิทัล (digital character recognition การใช้เทคนิคทางดิจิทัล เช่นสแกนเนอร์และอัลกอริธึมคอมพิวเตอร์) เดิมเคยเป็นสาขาที่แยกจากกัน แต่เนื่องจากเหลือการใช้งานน้อยมากที่ใช้เฉพาะเทคนิคทางแสง คำว่า โอซีอาร์ ในปัจจุบันจึงกินความกว้างถึงการประมวลผลภาพทางดิจิทัลด้วยเช่นกัน

2.6.1 วิธีการเข้าคู่รูปแบบ (Template matching)

วิธีการเข้าคู่รูปแบบเป็นวิธีการแรกๆ ที่มาใช้ในการรู้จำตัวอักษร หลักการโดยทั่วไปคือจะต้องมีรูปแบบ (template) ที่สร้างขึ้นมาสำหรับอ่านตัวอักษร โดยมีการกำหนดตำแหน่งสำคัญที่สามารถใช้แยกแยะความแตกต่างระหว่างตัวอักษรแต่ละตัว เวลาทำงานก็ให้นำรูปภาพที่ต้องการอ่านไปทาบบนแบบเพื่อวัดความคล้ายคลึงกันของภาพกับตัวแบบ จากนั้นก็ระบุว่าเป็นรหัสตัวอักษรอะไร โดยใช้ค่าผ่านระดับหรือวิธีการบางอย่างในการตัดสิน วิธีการนี้จะค่อนข้างอ่อนไหวต่อข้อมูลแทรกซ้อน ขนาด และการเอียงของตัวอักษร จึงจำเป็นต้องมีขั้นตอนการปรับแต่งข้อมูลที่ดี นอกจากนั้นขั้นตอนการเปรียบเทียบก็ไม่ใช่ว่าสามารถเทียบกันแบบจุดต่อจุดได้ เพราะในทางปฏิบัติตัวอักษรที่ส่งเข้าสามารถมีความแปรปรวนได้หลายรูปแบบ ดังนั้นวิธีการเทียบก็ต้องมีประสิทธิภาพเพียงพอที่จะรองรับกับปัญหาดังกล่าวได้

2.6.2 สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)

สำหรับการทำวิธีการเข้าคู่รูปแบบ (Template matching) จะใช้สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ในการเปรียบเทียบค่าระหว่างค่าพิกเซลของรูปแบบ (template) ของภาพที่ทำการเลือกมาเพื่อที่จะทำการแปลงเป็นตัวอักษรกับพิกเซลของรูปแบบ (template) ที่สร้างขึ้นไว้เพื่อใช้ในการเปรียบเทียบ

สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) **คือ ตัววัดความสัมพันธ์ระหว่างตัวแปรในเชิงเส้นตรง ค่าของสัมประสิทธิ์สหสัมพันธ์บอกถึงระดับความสัมพันธ์ว่ามากหรือน้อย โดยจะมีสูตรการหาค่าสัมประสิทธ์สหสัมพันธ์** (Correlation Coefficient) **ระหว่างรูปภาพ 2 รูป ดัง** สมการที่ 2.11

$$r = \frac{\sum_{m}\sum_{n}(A_{mn}-\overline{A})(B_{mn}-\overline{B})}{\sqrt{\left(\sum_{m}\sum_{n}(A_{mn}-\overline{A})^{2}\right)\left(\sum_{m}\sum_{n}(B_{mn}-\overline{B})^{2}\right)}} \qquad (2.11)$$

โดยที่ $\sum_{m}\sum_{n}(A_{mn}-\overline{A})(B_{mn}-\overline{B})$

คือ ผลรวมของผลต่างระหว่างค่า A แต่ละค่า กับ ค่าเฉลี่ย A คูณกับ **ผลรวมของผลต่างระห**ว่างค่า B แต่ละค่า กับ ค่าเฉลี่ย B

$$\left(\sum_{m}\sum_{n}(A_{mn}-\overline{A})^{2}\right)$$

คือ ผลรวมของกำลังสอง ของผลต่างระหว่างค่า A แต่ละค่า กับ ค่าเ**ฉลี่ย A**

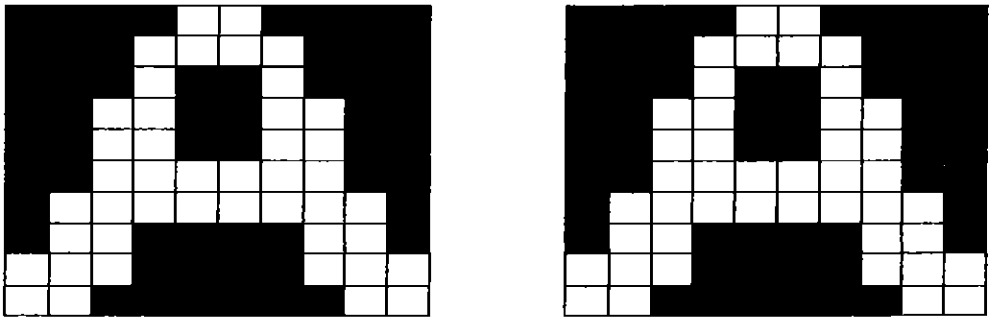
$$\left(\sum_{m}\sum_{n}(B_{mn}-\overline{B})^{2}\right)$$

คือ ผลรวมของกำลังสอง ของผลต่างระหว่างค่า B แต่ละค่า กับ ค่าเ**ฉลี่ย B**

ตารางที่ 2.1 แสดงความหมายของค่าสัมประสิทธิ์สหสัมพันธ์

ค่าสัมประสิทธิ์สหสัมพันธ์	ความหมาย
|r| > 1	คำนวณผิดให้คำนวณใหม่
r เข้าใกล้ -1 หรือ 1	ตัวแปรทั้งสองมีสหสัมพันธ์ชัดเจน
r ยิ่งเข้าใกล้ 0	ตัวแปรทั้งสองมีสหสัมพันธ์ยิ่งน้อยลง
r = 1 หรือ r = -1	ตัวแปรทั้งสองมีสหสัมพันธ์เป็นเส้นตรง 1 เส้น

2.6.2.1 ในกรณีที่ภาพทั้งสองมีความเหมือนกัน



(ก) (ข)

รูปที่ 2.19 (ก) ภาพ (template) รูป A จากภาพถ่ายและผ่านการปรับขนาด (resize) ให้เท่ากับภาพ (ข)
(ข) ภาพรูปแบบ (template) ของรูป B ที่ถูกเก็บไว้ภายในโปรแกรม

การคำนวณหาค่าค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ระหว่างรูปภาพ 2 รูปจะได้ ดังนี้

ก. การคำนวณส่วนด้านบนของสมการ

ค่าเฉลี่ยของรูป A เท่ากับ 0.44 (เพราะว่ามีจุดสีขาวที่มีค่าเป็นหนึ่งทั้งหมด 44 จุด)

ค่าเฉลี่ยของรูป B เท่ากับ 0.44 (เพราะว่ามีจุดสีขาวที่มีค่าเป็นหนึ่งทั้งหมด 44 จุด)

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ 1-0.44 เท่ากับ 0.56

ผลต่างระหว่างจุดดำ (มีค่าเป็นศูนย์) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ -0.44

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป B กับค่าเฉลี่ยของรูป B เท่ากับ 1-0.44 เท่ากับ 0.56

ผลต่างระหว่างจุดดำ (มีค่าเป็นศูนย์) ของรูป B กับค่าเฉลี่ยของรูป B เท่ากับ -0.44

ดังนั้นในการคำนวณจะได้ว่า จุดที่มีค่าสีขาวเหมือนกัน จะได้ค่าเท่ากับ (0.56)(0.56) และ จุดที่มีค่าสีดำเหมือนกันได้ค่าเท่ากับ (-0.44)(-0.44)

ซึ่งในภาพ A และภาพ B ถ้าดูจากรูปด้านบนจะพบว่า มีจุดสีขาวเหมือนกันทั้งหมด 44 จุด จะได้ค่าเท่ากับ (0.56)(0.56)(44) ดังนั้นจะได้ในส่วนนี้เท่ากับ 13.7984 และจุดสีดำเหมือนกัน 56 จุด ค่าที่ได้จะเท่ากับ (-0.44)(-0.44)(56) ดังนั้นจะได้ในส่วนนี้เท่ากับ 10.8416

ผลรวมในส่วนด้านบนของสมการจะเท่ากับ 24.64 ในการคำนวณส่วนด้านล่างของสมการจะนำค่าผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป A กับค่าเฉลี่ยของรูป A มายกกำลังสอง จะได้เท่ากับ 0.3136

ค่าผลต่างระหว่างจุดสีดำ (มีค่าเป็นศูนย์) ของรูป A กับค่าเฉลี่ยของรูป A มายกกำลังสอง จะได้เท่ากับ 0.1936 และหาผลรวมระหว่างจุดของภาพ A ซึ่งจะพบว่ามีจุดที่มีสีขาว 44 จุด จะได้ค่าผลรวมของจุดสีขาวเป็น (0.3136)(44) เท่ากับ 13.7984 และมีจุดสีดำทั้งหมด 56 จุด จะได้ค่าผลรวมของจุดสีดำเป็น (0.1936)(56) เท่ากับ 10.8416 นำผลลัพธ์ทั้งสองมารวมกันจะได้เป็นผลรวมระหว่างจุดของภาพ A ซึ่งมีค่าเท่ากับ 24.64

ข. การคำนวณส่วนด้านล่างของสมการ

ในการคำนวณส่วนด้านล่างของสมการจะนำค่าผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป B กับค่าเฉลี่ยของรูป B มายกกำลังสอง จะได้เท่ากับ 0.3136

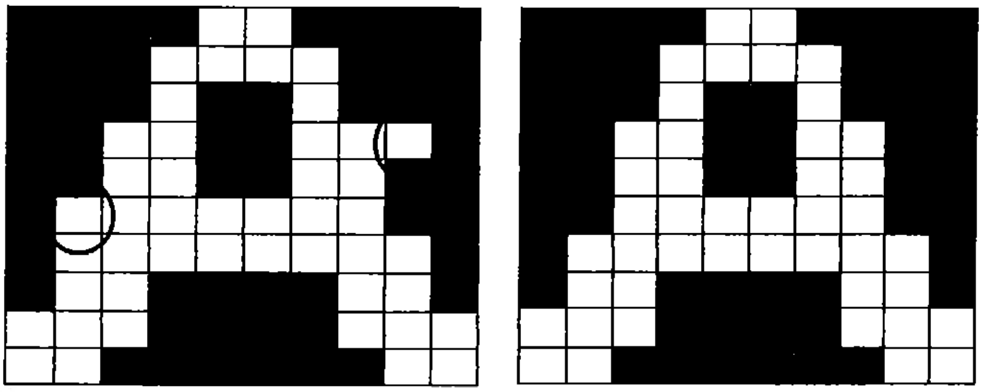
นำค่าผลต่างระหว่างจุดสีดำ (มีค่าเป็นศูนย์) ของรูป B กับค่าเฉลี่ยของรูป B มายกกำลังสอง จะได้เท่ากับ 0.1936

จากนั้นหาผลรวมระหว่างจุดของภาพ B ซึ่งจะพบว่ามีจุดที่มีสีขาว 44 จุด จะได้ค่าผลรวมของจุดสีขาวเป็น (0.3136)(44) เท่ากับ 13.7984 และมีจุดสีดำทั้งหมด 56 จุด จะได้ค่าผลรวมของจุดสีดำเป็น (0.1936)(56) เท่ากับ 10.8416 นำผลลัพธ์ทั้งสองมารวมกันจะได้เป็นผลรวมระหว่างจุดของภาพ B ซึ่งมีค่าเท่ากับ 24.64

หลังจากนั้นนำค่าผลรวมของภาพทั้งสองมาคูณกันจะได้ค่าเท่ากับ 607.1296 ทำการหาค่ารากที่สองของ 607.1296 จะได้เท่ากับ 24.62

จากนั้นนำค่าของด้านบนของสมการมาหารด้วยค่าด้านล่างของสมการจะได้ค่า r เท่ากับ 1

2.6.2.2 ในกรณีที่ภาพทั้งสองมีความแตกต่างกันเล็กน้อย



(ก) (ข)

รูปที่ 2.20 (ก) ภาพ (template) รูปA จากภาพถ่ายและผ่านการปรับขนาด (resize) ให้เท่าภาพ (ข)

(ข) แสดงภาพรูปแบบ (template) ของรูป B ที่ถูกเก็บไว้ภายในโปรแกรม

การคำนวณหาค่าค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ระหว่างรูปภาพ 2 รูปจะได้ดังนี้

ก. การคำนวณส่วนด้านบนของสมการ

ค่าเฉลี่ยของรูป A เท่ากับ 0.46 (เพราะว่ามีจุดสีขาวที่มีค่าเป็นหนึ่งทั้งหมด 46 จุด)

ค่าเฉลี่ยของรูป B เท่ากับ 0.44 (เพราะว่ามีจุดสีขาวที่มีค่าเป็นหนึ่งทั้งหมด 44 จุด)

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ 1-0.46 เท่ากับ 0.54

ผลต่างระหว่างจุดดำ (มีค่าเป็นศูนย์) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ -0.46

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป B กับค่าเฉลี่ยของรูป B เท่ากับ 1-0.44 เท่ากับ 0.56

ดังนั้นในการคำนวณจะได้ว่า จุดที่มีค่าสีขาวเหมือนกันจะได้ค่าเท่ากับ (0.54)(0.56) และ จุดที่มีค่าสีดำเหมือนกันจะได้ค่าเท่ากับ (-0.46)(-0.44)

ซึ่งในภาพ A และภาพ B ถ้าดูจากรูปด้านบนจะพบว่า มีจุดสีขาวเหมือนกันทั้งหมด 44 จุด จะได้ว่า (0.54)(0.56)(44) = 13.3056 และจุดดำเหมือนกัน 54 จุดจะได้ (-0.46)(-0.44)(54) = 10.9296 ดังนั้น 10.9296 + 13.3056 = 24.2352

โดยมีจุดที่มีค่าสีแตกต่างกันสองจุดคือจุด (4,9) และ (6,2) การคำนวณจะเป็นดังนี้ จุด (4,9) ภาพ A เป็นสีขาว ภาพ B เป็นสีดำ จะได้ว่า (0.54)(-0.44) จุด (6,2) จะได้ A เป็นสีขาว B เป็นสีดำ จะได้ (0.54)(-0.44)

ดังนั้นในสมการด้านบนจะได้ 24.2352 + (-0.54)(0.44)(2) = 24.2352 – 0.4752 = 23.76

ข. การคำนวณส่วนด้านล่างของสมการ

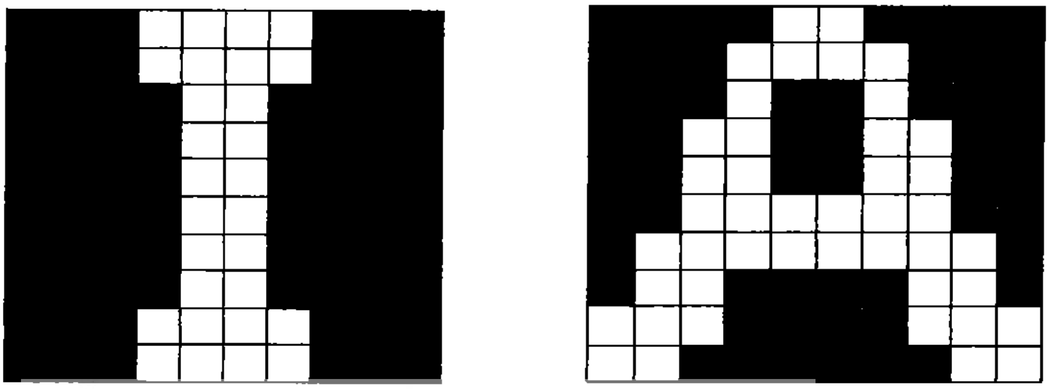
A มีจุดดำ 54 จุด ขาว 46 จุด จะได้ในส่วนของผลรวมกำลังสองของ A เป็น $((-0.46)^2 * (54)) + ((0.54)^2 * (46))$ เท่ากับ 24.84

B มีจุดดำ 56 จุด ขาว 44 จุด จะได้ในส่วนของผลรวมกำลังสองของ B เป็น $((-0.44)^2 * (56)) + ((0.56)^2 * (44))$ เท่ากับ 24.64

นำค่าผลรวมของ A และ B มาคูณกัน (24.84)(24.64) = 612.0576 หาค่ารากที่สอง ได้ 24.7398

จากนั้นนำ 23.76/24.7398 เท่ากับ 0.9604

2.6.2.3 ในกรณีที่ภาพทั้งสองมีความแตกต่างกันมาก



(ก) (ข)

รูปที่ 2.21 (ก) ภาพ (template) รูป A จากภาพถ่ายและผ่านการปรับขนาด (resize) **ให้เท่ากับภาพ (ข)**

(ข) ภาพรูปแบบ (template) รูป B ที่ถูกเก็บไว้ภายในโปรแกรม

การคำนวณหาค่าค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) **ระหว่างรูปภาพ 2 รูปจะได้ดังนี้**

ก. การคำนวณส่วนด้านบนของสมการ

ค่าเฉลี่ยของรูป A เท่ากับ 0.28 (เพราะว่ามีจุดสีขาวที่มีค่าเป็น**หนึ่งทั้งหมด 28 จุด)**

ค่าเฉลี่ยของรูป B เท่ากับ 0.44 (เพราะว่ามีจุดสีขาวที่มีค่าเป็น**หนึ่งทั้งหมด 44 จุด)**

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ 1-0.28 เท่ากับ 0.72

ผลต่างระหว่างจุดดำ (มีค่าเป็นศูนย์) ของรูป A กับค่าเฉลี่ยของรูป A เท่ากับ -0.28

ผลต่างระหว่างจุดสีขาว (มีค่าเป็นหนึ่ง) ของรูป B กับค่าเฉลี่ยของรูป B เท่ากับ 1-0.44 เท่ากับ 0.56

ผลต่างระหว่างจุดดำ (มีค่าเป็นศูนย์) ของรูป B กับค่าเฉลี่ยของรูป B เท่ากับ -0.44

ดังนั้นในการคำนวณจะได้ว่า จุดที่มีค่าสีขาวเหมือนกันได้ค่าเท่ากับ (0.72)(0.56)

และจุดที่มีค่าสีดำเหมือนกัน จะได้ค่าเท่ากับ (-0.28)(-0.44) ซึ่งในภาพ A และภาพ B ถ้าดูจากรูปด้านบนจะพบว่า มีจุดสีขาวเหมือนกันทั้งหมด 10 จุด จะได้ว่า (0.72)(0.56)(11) = 4.4352 และจุดดำเหมือนกัน36 จุดจะได้ (-0.28)(-0.44)(36) = 4.4352 ดังนั้น 4.4352 + 4.4352 = 8.8704

มีจุดที่มีค่าสีแตกต่างกันโดยที่ ภาพ A เป็นสีขาว ภาพ B เป็นสีดำ 18 จุด จะได้ว่า (0.72)(-0.44)(18) = - 5.7024 และในส่วนที่ภาพ A เป็นสีดำ B เป็นสีขาว 36 จุด จะได้ (-0.28) (0.56)(36) = -5.6448

ดังนั้นในสมการด้านบนจะได้ 8.8704 - 5.7024 - 5.6448 = -2.4768

ข. การคำนวณส่วนด้านล่างของสมการ

A มีจุดดำ 72 จุด ขาว 28 จุด จะได้ในส่วนของผลรวมกำลังสองของ A เป็น

$((-0.72)^2 * (28)) + ((0.28)^2 * (72))$ เท่ากับ 20.16

B มีจุดดำ 56 จุด ขาว 44 จุด จะได้ในส่วนของผลรวมกำลังสองของ B เป็น

$((-0.44)^2 * (56)) + ((0.56)^2 * (44))$ เท่ากับ 24.64

นำผลรวมของ A และ B มาคูณกัน (20.16)(24.64) = 612.0576 จากนั้นหารากที่สองจะได้ 22.2877

จากนั้นนำ -2.4768/22.2877 เท่ากับ -0.1111

2.7 กระบวนการตรวจสอบตัวสะกด (Spell Check)

กระบวนการตรวจสอบตัวสะกด หมายถึง โปรแกรมที่เขียนขึ้นเพื่อใช้ช่วยในการตรวจสอบการสะกดของข้อความที่พิมพ์ โดยเทียบกับพจนานุกรม (โดยเฉพาะภาษาอังกฤษ) มักใช้ร่วมกับโปรแกรมประมวลคำ (Word processing) ในปัจจุบันได้มีการพัฒนาไปมากนอกจากจะทักท้วงถ้าพบตัวที่สะกดผิดแล้ว ยังแสดงคำที่สะกดถูกให้ด้วย

2.8 ฐานข้อมูลเบื้องต้น (Foundation of Database)[8]

2.8.1 ความหมายของฐานข้อมูล

ฐานข้อมูล หมายถึง กลุ่มของข้อมูลที่ถูกรวบรวมเก็บไว้ โดยมีความสัมพันธ์ซึ่งกันและกัน โดยไม่ได้บังคับว่าข้อมูลทั้งหมดจะต้องถูกเก็บอยู่ในแฟ้มข้อมูลเดียวกัน หรือแยกเก็บหลายๆ แฟ้มข้อมูล ที่สำคัญคือ ฐานข้อมูลจะต้องมีการเก็บข้อมูลอย่างเป็นระบบ และจะต้องสร้างความสัมพันธ์ระหว่างระเบียน (Record) และสามารถเรียกใช้ความสัมพันธ์นั้นได้ ในฐานข้อมูลนั้นย่อมมีความซ้ำซ้อนของข้อมูล จำเป็นต้องกำจัดความซ้ำซ้อนของข้อมูลออกเสีย และจะต้องมีการกำหนดระบบความปลอดภัยของข้อมูลโดยการมีการจำกัดสิทธิ์ของผู้เข้าใช้ข้อมูล เพราะบางข้อมูลอาจจะมีการใช้ที่จำกัดสิทธ์ และบางข้อมูลสามารถใช้ร่วมกันได้ ระบบฐานข้อมูล หรือที่เรียกว่า DBMS (Data Base Management System) มีหน้าที่ช่วยให้ผู้ใช้เข้าถึงข้อมูลโดยง่าย สะดวก และมีประสิทธิภาพ การเข้าถึงข้อมูลนั้นอาจจะเป็น การสร้าง การแก้ไขฐานข้อมูลเป็นต้น

2.8.2 โครงสร้างข้อมูล

2.8.2.1 บิท (Bit) หมายถึง หน่วยของข้อมูลที่มีขนาดเล็กที่สุด

2.8.2.2 ไบท์ (Byte) หมายถึง หน่วยของข้อมูลที่เกิดจากการนำบิทมารวมกันเป็นตัวอักขระ

2.8.2.3 เขตข้อมูล (Field) หมายถึง หน่วยของข้อมูลที่ประกอบขึ้นจากตัวอักขระตั้งแต่หนึ่งตัวขึ้นไปมารวมกันแล้วได้ความหมายของสิ่งใดสิ่งหนึ่ง เช่น ชื่อ ที่อยู่ เป็นต้น

2.8.2.4 ระเบียน (Record) หมายถึง หน่วยของข้อมูลที่เกิดจากการนำเอาเขตข้อมูลหลายๆ เขตข้อมูลมารวมกัน เพื่อเกิดเป็นข้อมูลเรื่องใดเรื่องหนึ่ง เช่น ข้อมูลของนักศึกษา 1 ระเบียน (1 คน) จะประกอบด้วย

ก. รหัสประจำตัวนักศึกษา 1 เขตข้อมูล

ข. ชื่อนักศึกษา 1 เขตข้อมูล

ค. ที่อยู่ 1 เขตข้อมูล

2.8.2.5 แฟ้มข้อมูล (File) หมายถึง หน่วยของข้อมูลที่เกิดจากการนำเอาข้อมูลหลายๆ ระเบียนที่เป็นเรื่องเดียวกันมารวมกัน เช่น แฟ้มข้อมูลนักศึกษา แฟ้มข้อมูลลูกค้า แฟ้มข้อมูลพนักงาน

2.8.2.6 เอนทิตี้ (Entity) หมายถึง ชื่อของสิ่งใดสิ่งหนึ่ง ได้แก่ คน สิ่งของ การกระทำ สถานที่ ซึ่งต้องการจัดเก็บข้อมูล

2.8.2.7 แอททริบิวต์ (Attribute) หมายถึง รายละเอียดข้อมูลที่แสดงลักษณะและคุณสมบัติของเอนทิตี้หนึ่งๆ เช่น เอนทิตี้นักศึกษา ประกอบด้วย

ก. แอทริบิวต์รหัสนักศึกษา

ข. แอททริบิวต์ชื่อนักศึกษา

ค. แอททริบิวต์ที่อยู่นักศึกษา

2.8.2.8 ความสัมพันธ์ (Relationship) หมายถึง ความสัมพันธ์ระหว่างเอนทิตี้ เช่น ความสัมพันธ์ระหว่างเอนทิตี้นักเรียนและเอนทิตี้โรงเรียน เป็นลักษณะว่า นักเรียนแต่ละคนเรียนอยู่โรงเรียนใด ความสัมพันธ์ระหว่างเอนทิตี้ แบ่งออกเป็น 3 ประเภท ดังนี้

ก. ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (One-to-one Relationship) เป็นการแสดงความสัมพันธ์ของข้อมูลในเอนทิตี้หนึ่งที่มีความสัมพันธ์กับข้อมูลในอีกเอนทิตี้หนึ่ง ในลักษณะหนึ่งต่อหนึ่ง (1:1)

ข. ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (One-to-many Relationship) เป็นการแสดงความสัมพันธ์ของข้อมูลในเอนทิตี้หนึ่ง ที่มีความสัมพันธ์กับข้อมูลหลายๆ ข้อมูลในอีกเอนทิตี้หนึ่ง ในลักษณะ (1:m)

ค. ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (Many-to-many Relationship) เป็นการแสดงความสัมพันธ์ของข้อมูลสองเอนทิตี้ในลักษณะกลุ่มต่อกลุ่ม (m:n)

จากคำศัพท์ต่างๆ ที่เกี่ยวข้องกับระบบฐานข้อมูลที่ได้กล่าวมาแล้วข้างต้น จึงอาจให้ นิยามของฐานข้อมูลในอีกลักษณะได้ว่า "ฐานข้อมูล" อาจหมายถึง โครงสร้างสารสนเทศ ที่ประกอบด้วยหลายๆเอนทิตี้ที่มีความสัมพันธ์กัน

2.8.3 รูปแบบของระบบฐานข้อมูล

รูปแบบของระบบฐานข้อมูล มีอยู่ด้วยกัน 3 ประเภท คือ

2.8.3.1 ฐานข้อมูลเชิงสัมพันธ์ (Relational Database)

เป็นการเก็บข้อมูลในรูปแบบที่เป็นตาราง หรือเรียกว่า รีเลชั่น (Relation) มีลักษณะเป็น 2 มิติ คือเป็นแถว และเป็นคอลัมน์ การเชื่อมโยงข้อมูลระหว่างตาราง จะเชื่อมโยงโดยใช้แอททริบิวต์ (attribute) หรือคอลัมน์ที่เหมือนกันทั้งสองตารางเป็นตัวเชื่อมโยงข้อมูล ฐานข้อมูลเชิงสัมพันธ์นี้จะเป็นรูปแบบของฐานข้อมูลที่นิยมใช้ในปัจจุบัน

2.8.3.2 ฐานข้อมูลแบบเครือข่าย (Network Database)

ฐานข้อมูลแบบเครือข่ายจะเป็นการรวมระเบียนต่าง ๆ และความสัมพันธ์ระหว่างระเบียนแต่จะต่างกับฐานข้อมูลเชิงสัมพันธ์ คือ ในฐานข้อมูลเชิงสัมพันธ์จะแฝงความสัมพันธ์เอาไว้ โดยระเบียนที่มีความสัมพันธ์กันจะต้องมีค่าของข้อมูลในแอททริบิวต์ใดแอททริบิวต์หนึ่งเหมือนกัน แต่ฐานข้อมูลแบบเครือข่าย จะแสดงความสัมพันธ์อย่างชัดเจน ตัวอย่างเช่น

2.8.3.3 ฐานข้อมูลแบบลำดับชั้น (Hierarchical Database)

ฐานข้อมูลแบบลำดับชั้น เป็นโครงสร้างที่จัดเก็บข้อมูลในลักษณะความสัมพันธ์แบบพ่อ-ลูก (Parent-Child Relationship Type: PCR Type) หรือเป็นโครงสร้างรูปแบบต้นไม้ (Tree) ข้อมูลที่จัดเก็บในที่นี้ คือ ระเบียน (Record) ซึ่งประกอบด้วยค่าของเขตข้อมูล (Field) ของเอนทิตี้หนึ่ง ๆ ฐานข้อมูลแบบลำดับชั้นนี้คล้ายคลึงกับฐานข้อมูลแบบเครือข่าย แต่ต่างกันที่ฐานข้อมูลแบบลำดับชั้น มีกฎเพิ่มขึ้นมาอีกหนึ่งประการ คือ ในแต่ละกรอบจะมีลูกศรวิ่งเข้าหาได้ไม่เกิน 1 หัวลูกศร

2.8.4 ความสำคัญของระบบฐานข้อมูล

2.8.4.1 จัดเก็บและบันทึกข้อมูล (Data Storage)

2.8.4.2 ลดความซ้ำซ้อนของข้อมูล (Reduce Data Redundancy)

2.8.4.3 สามารถใช้ข้อมูลร่วมกันได้ (Data Concurrency)

2.8.4.4 ลดความขัดแย้งหรือแตกต่างกันของข้อมูล (Reduce Data Inconsistency)

2.8.4.5 ป้องกันการแก้ไขข้อมูลต่างๆ (Protect Data Editing)

2.8.4.6 ความถูกต้องของข้อมูลมีมากขึ้น (Data Accuracy)

2.8.4.7 สะดวกในการสืบค้นข้อมูล (Data Retrieval or Query

2.8.4.8 ป้องกันการสูญหายของข้อมูล หรือฐานข้อมูลถูกทำลาย (Data Security)

2.8.5 คีย์ในฐานข้อมูลเชิงสัมพันธ์

ในระบบฐานข้อมูลเชิงสัมพันธ์จะต้องกำหนดชนิดของคีย์ เพื่อเป็นแอททริบิวต์พิเศษที่ทำหน้าที่บางอย่างเช่น เป็นตัวแทนของตาราง ซึ่งถูกกำหนดขึ้นมาเพื่อความสะดวกในการอ้างถึงทฤษฎีเกี่ยวกับการทำนอร์มัลไรเซชั่น (Normalization) ซึ่งมีดังนี้

2.8.5.1 คีย์หลัก (Primary Key) จะเป็นฟิลด์ที่ไม่ซ้ำกันเลยในแต่ละเร็คคอร์ดในตารางนั้นและสามารถใช้คีย์นี้เป็นตัวแทนของตารางนั้นได้ทันที

2.8.5.2 คีย์คู่แข่ง (Candidate Key) เป็นฟิลด์หนึ่งหรือหลายฟิลด์ที่พอเอามารวมกันแล้วคุณสมบัติ เป็นคีย์หลัก และไม่ได้ถูกใช้เป็นคีย์หลัก

2.8.5.3 คีย์ส่วนประกอบ (Composite Key) บางตารางหาฟิลด์ที่ไม่ซ้ำกันไม่ได้เลย จึงต้องใช้หลายๆฟิลด์มารวมกันเป็นคีย์หลัก ซึ่งฟิลด์ที่ใช้รวมกันนี้เรียกว่าคีย์หลัก

2.8.5.4 คีย์นอก (Foreign Key) เป็นฟิลด์ใดๆ ในตารางหนึ่ง (ฝั่ง Many) ที่มีความสัมพันธ์กับฟิลด์ที่เป็นคีย์หลัก ในอีกตารางหนึ่ง (ฝั่ง One) โดยที่ทั้งสองตารางมีความสัมพันธ์แบบหนึ่งต่อกลุ่ม

2.8.6 โปรแกรมฐานข้อมูลที่ใช้

โปรแกรมฐานข้อมูล เป็นโปรแกรมหรือซอฟแวร์ที่ช่วยจัดการข้อมูลหรือรายการต่าง ๆ ที่อยู่ในฐานข้อมูล ไม่ว่าจะเป็นการจัดเก็บ การเรียกใช้ การปรับปรุงข้อมูลโปรแกรมฐานข้อมูล จะช่วยให้ผู้ใช้สามารถค้นหาข้อมูลได้อย่างรวดเร็ว ซึ่งโปรแกรมฐานข้อมูลที่นิยมใช้มีอยู่ด้วยกันหลายตัวโดยแต่ละโปรแกรมจะมีความสามารถต่างกัน บางโปรแกรมใช้ง่ายแต่จะจำกัดขอบเขตการใช้งาน บางโปรแกรมใช้งานยากกว่า แต่จะมีความสามารถในการทำงานมากกว่า

โปรแกรม Microsoft Access นับเป็นโปรแกรมที่นิยมใช้กันมากในขณะนี้ เนื่องจาก Microsoft Access เป็นโปรแกรมฐานข้อมูลที่หาใช้ได้ง่ายเพราะอยู่ในMicrosoft Office และในระบบฐานข้อมูลขนาดใหญ่Microsoft Access สามารถสร้างแบบฟอร์มที่ต้องการจะเรียกดูข้อมูลในฐานข้อมูล หลังจากบันทึกข้อมูลในฐานข้อมูลเรียบร้อยแล้ว จะสามารถค้นหาหรือเรียกดูข้อมูลจากเขตข้อมูลใดก็ได้ นอกจากนี้Microsoft Access ยังมีระบบรักษาความปลอดภัยของข้อมูล โดยการกำหนดรหัสผ่านเพื่อป้องกันความปลอดภัยของข้อมูลในระบบได้ด้วย ซึ่งในโครงงานนี้ได้เลือกใช้ Microsoft Access เป็นฐานข้อมูลให้กับโปรแกรม

2.9 การติดต่อกับตัวกลางที่ใช้ในการติดต่อกับฐานข้อมูล [9]

ตัวกลางในการติดต่อกับฐานข้อมูลต่างๆ(Open Database Connectivity :ODBC)เป็น ช่องทางการเชื่อต่อ (Application Programming Interface : API) คือ มาตรฐานในการเข้าถึงข้อมูลที่ถูกกำหนดขึ้น ที่อนุญาตให้เชื่อมต่อฐานข้อมูลอื่นๆ ได้

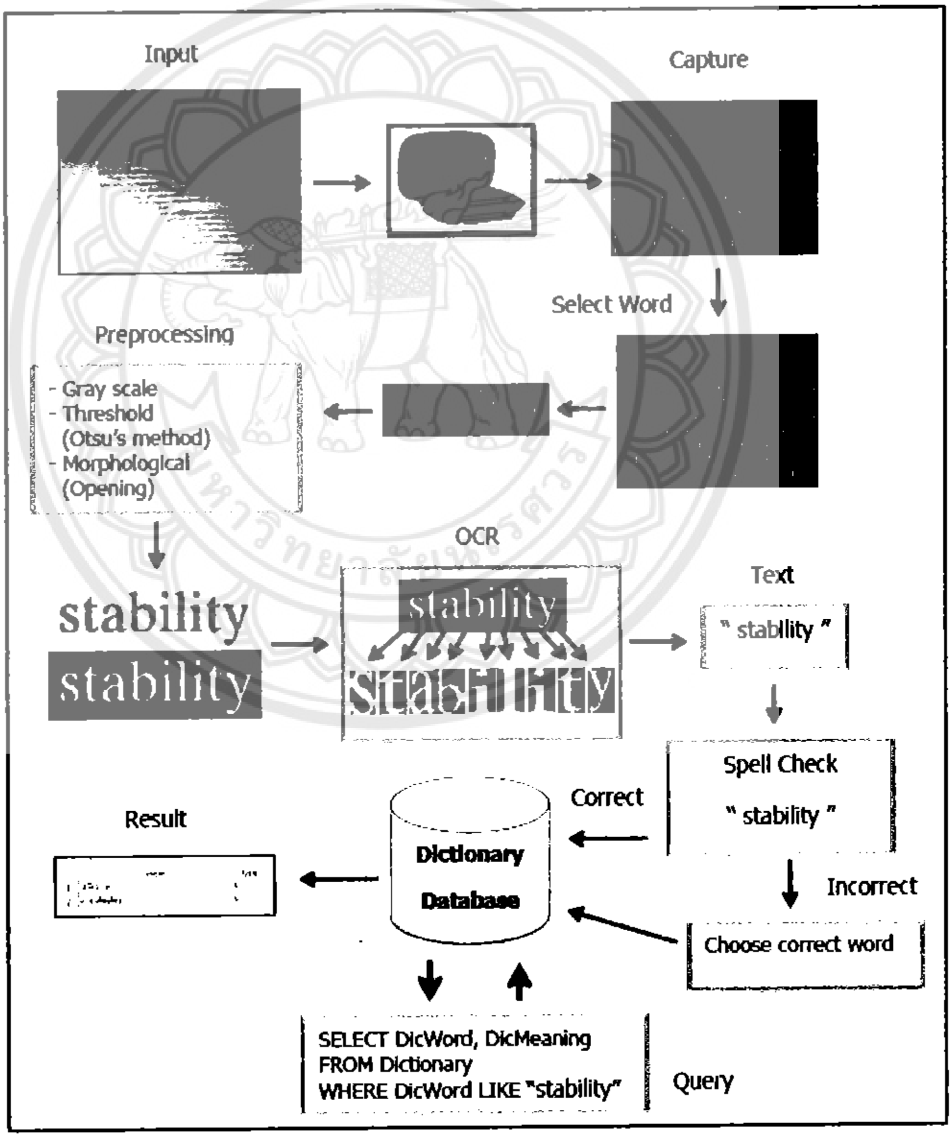
ตัวกลางในการติดต่อกับฐานข้อมูลมีพื้นฐานและการจัดตำแหน่งใกล้เคียงกับ Open Group มาตรฐานของภาษา SQL ซึ่ง ตัวกลางในการติดต่อกับฐานข้อมูลเป็นระบบอินเตอร์เฟส โดยการใช้โปรแกรมนี้จะทำให้สามารถใช้คำสั่ง SQL เข้าถึงมาตรฐานข้อมูลโดยไม่จำเป็นต้องรู้จักคุณสมบัติการอินเตอร์เฟสของฐานข้อมูลตัวกลางในการติดต่อกับฐานข้อมูล จะรับรู้ภาษา SQL และแปลงเป็นคำสั่งของฐานข้อมูลแต่ละระบบ

ตัวกลางในการติดต่อกับฐานข้อมูลจะใช้ในการติดต่อกับฐานข้อมูลในลักษณะระบบไคลเอนท์เซิฟเวอร์(Client/Server)โดยการติดต่อกับฐานข้อมูลใดๆจะไม่ขึ้นอยู่กับชนิดของโปรแกรมระบบฐานข้อมูลหน้าที่หลักๆของ ตัวกลางในการติดต่อกับฐานข้อมูลก็คือ การติดต่อกับฐานข้อมูล ทำการจัดเตรียมคำสั่ง SQL เพื่อทำการส่งให้ระบบจักการฐานข้อมูล *(Database Management System* :DBMS)อีกที แล้วก็จัดการประมวลผลของการกระทำที่ใช้งานกลุ่มคำสั่งปฏิบัติในฐานข้อมูล (Transaction)ทำการรับ-ส่ง ข้อมูลที่ได้จาก ฐานข้อมูล กลับไปยังโปรแกรม พร้อมทั้งแจ้งข้อผิดพลาดถ้ารูปแบบคำสั่งนั้นผิดและความสามารถในลักษณะเช่นนี้ถือว่าเป็นคุณลักษณะเด่นที่สำคัญที่สุดของตัวกลางในการติดต่อกับฐานข้อมูล

บทที่ 3
ขั้นตอนการดำเนินงาน

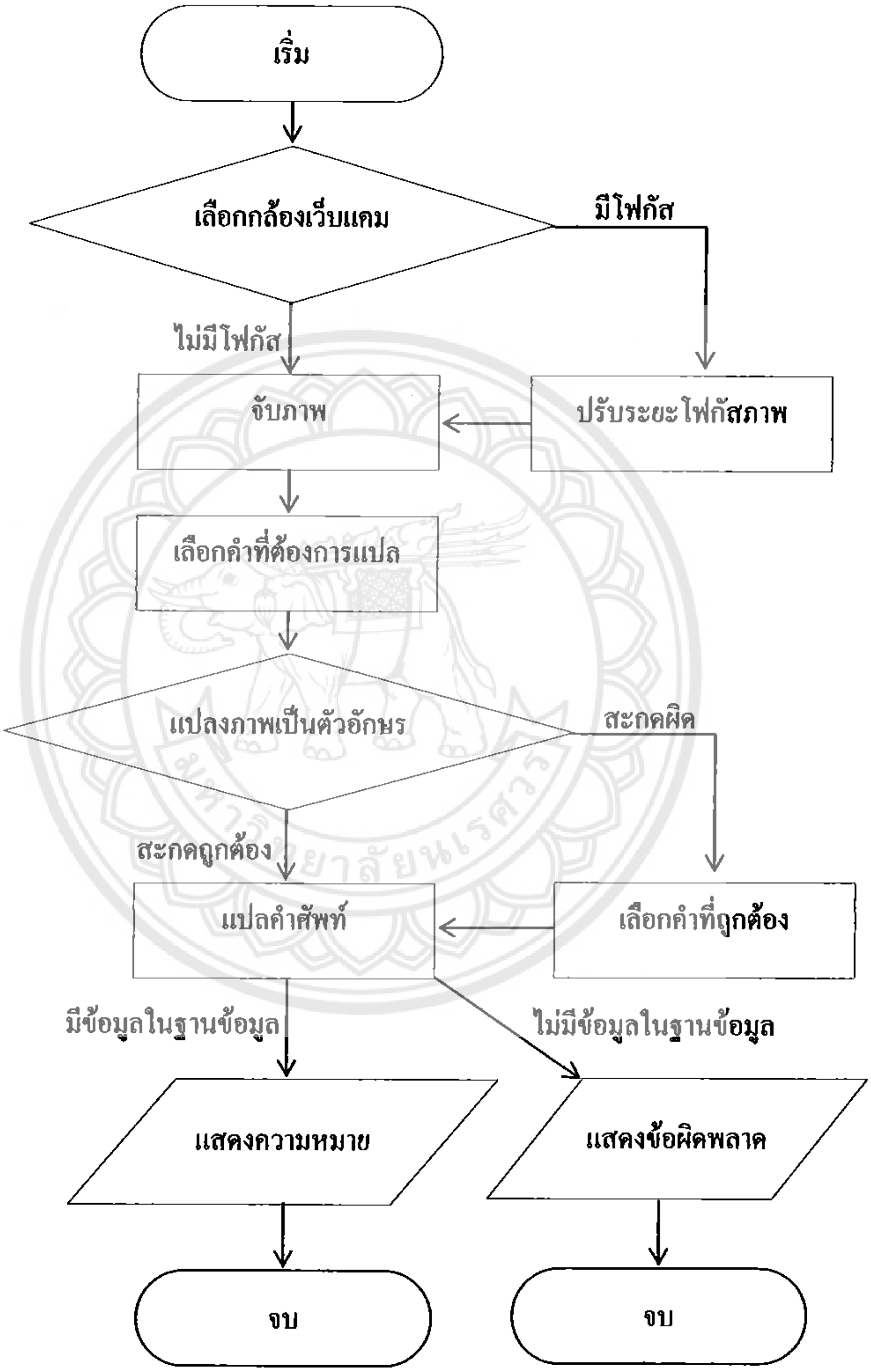
3.1 โครงสร้างของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม

โปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทย ผ่านกล้องเว็บแคมประกอบไปด้วยข้อมูลภาพถ่ายซึ่งตัวอักษรภาษาอังกฤษที่ต้องการแปลและส่วนของโปรแกรมที่ใช้ในการประมวลผลภาพและส่วนสุดท้ายคือส่วนของฐานข้อมูลในการเก็บความหมายของคำศัพท์ภาษาอังกฤษ



รูปที่ 3.1 ภาพโดยรวมของขั้นตอนการดำเนินโครงการ (System Overview)

ในส่วนของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม มีขั้นตอนการทำงานทั้งหมด ดังรูปที่ 3.2



รูปที่ 3.2 แสดงขั้นตอนการทำงานของโปรแกรม

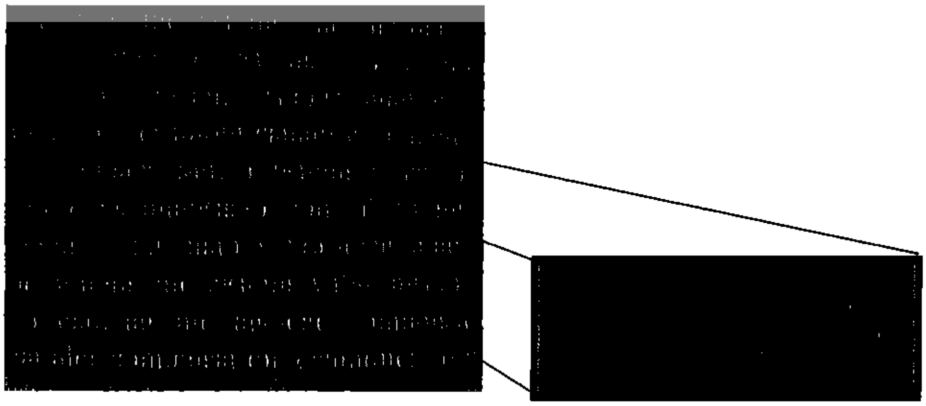
โดยขั้นแรกจะเป็นขั้นตอนของการเตรียมภาพถ่ายโดย โดยภาพถ่ายที่จะเป็นอินพุตให้กับระบบนี้จะจับภาพโดยกล้องเว็บแคม โดยจะเลือกจับภาพเฉพาะคำที่ต้องการเท่านั้น (คำที่อยู่ในกรอบสี่เหลี่ยม)หลังจากนั้นเมื่อได้รูปภาพของคำที่ต้องการเข้ามาในระบบแล้วก็จะทำการปรับภาพให้เป็นภาพระดับเทา เมื่อได้ภาพระดับเทาแล้วก็จะแปลงภาพให้เป็นภาพขาวดำ 2 ระดับ ซึ่งจะเหลือข้อมูลภาพเพียง 2 ระดับเท่านั้นคือ 0 และ 1 หลังจากนั้นก็จะเข้าสู่กระบวนการแยกคำให้เป็นตัวอักษรแล้วจึงนำรูปภาพตัวอักษรแต่ละตัวไปเปรียบเทียบกับภาพแม่แบบ (Template) แต่ละตัว เมื่อเสร็จสิ้นกระบวนการเปรียบเทียบภาพแม่แบบ (Template) ก็จะได้คำที่ต้องการค้นหาสุดท้ายก็จะนำคำที่นี้ไปเปรียบเทียบกับคำในฐานข้อมูล เพื่อหาความหมายของคำศัพท์

3.2 ขั้นตอนการนำภาพเข้าสู่โปรแกรม

3.2.1 ภาพย่อย (Sub Image)

ขั้นตอนแรกจะเป็นการเตรียมข้อมูลภาพเพื่อใช้ในการประมวลผล เนื่องจากโปรแกรมสามารถแปลคำศัพท์ได้ทีละคำ จึงต้องทำการ จับภาพเฉพาะคำที่ต้องการแปลเท่านั้น (Sub Image) เพื่อนำไปผ่านกระบวนการประมวลผลขั้นต้น (Preprocessing)

การหาตำแหน่งของภาพคำศัพท์ที่ต้องการแปลนั้นสามารถทำได้โดยใช้คำสั่ง getrect เพื่อระบุจุดที่ต้องการในรูปด้วยการลากตีกรอบด้วยเมาส์เพื่อครอบคลุมตำแหน่งของคำที่ต้องการจะแปล ด้วยกรอบสี่เหลี่ยมนี้จะทำให้สามารถหาตำแหน่งของภาพที่ต้องการจะนำมาแปลความหมาย ดังรูปที่ 3.3



รูปที่ 3.3 ภาพย่อย (Sub Image)

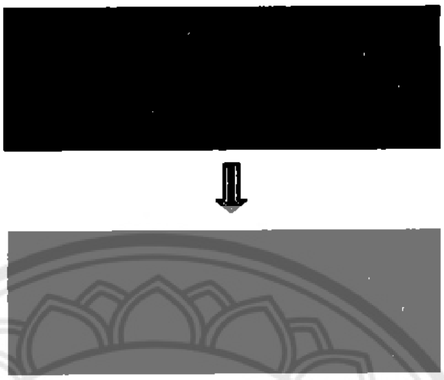
3.3 ขั้นตอนการประมวลผลขั้นต้น (Preprocessing)

ในส่วนของขั้นตอนการประมวลผลขั้นต้นของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม มีขั้นตอนการทำงานทั้งหมด ดังรูปที่ 3.4

รูปที่ 3.4 ขั้นตอนการทำงานของการประมวลผลขั้นต้น (Preprocessing)

3.3.1 ภาพระดับเทา (Gray Scale)

ซึ่งข้อมูลภาพที่ได้จะเป็นภาพระดับสี (RGB) เนื่องจากภาพระดับสีมีความซับซ้อนมาก จึงต้องทำการแปลงภาพให้เป็นภาพระดับเทาเสียก่อน เพื่อลดความซับซ้อนลงดังรูปที่ 3.5



รูปที่ 3.5แปลงภาพสีแบบ RGB เป็นภาพแบบระดับเทา

เมื่อได้ภาพระดับเทาแล้ว จะแปลงภาพให้เป็นภาพขาวดำ 2 ระดับ ด้วยการเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ เนื่องจากภายในรูปภาพที่ทำการแปลงเป็นภาพระดับเทาแล้ว อาจจะมีความเข้มของแสงหลายๆค่า ถ้าบางตัวอักษรมีค่าความเข้มแสงน้อยแต่ไม่เท่ากับ 0 ระบบก็อาจจะมองว่าเป็นพื้นหลังของภาพก็ได้

3.3.2 การเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ (Otsu's Thresholding method)

การทำเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ จะช่วยลดจำนวนข้อมูลที่ไม่จำเป็นในการวิเคราะห์ได้ ซึ่งจะเหลือข้อมูลภาพเพียงสองระดับคือ 0 และ 1 โดย 0 จะเป็นพื้นหลัง ส่วน 1 จะเป็นตัวอักษรในรูปภาพ จึงทำให้สามารถแยกตัวอักษรออกจากรูปภาพได้ ซึ่งจะทำให้ง่ายต่อการนำไปเปรียบเทียบกับรูปแบบ (templates) ที่ได้เตรียมไว้แล้วดังรูปที่ 3.6

รูปที่ 3.6 ภาพที่ผ่านกระบวนการเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ

3.3.3 การกำจัดสัญญาณรบกวนโดยวิธีการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ โอเปเรชั่นโอเพนนิ่ง (Opening)

เนื่องจากในบางครั้งการจับภาพคำอาจจะมี สัญญาณรบกวนเกิดขึ้นได้ ไม่ว่าจะเป็นจากการเลือกคำ ที่มีพื้นที่ตัวอักษรของคำอื่นติดมาด้วยเล็กน้อย หรือ คุณภาพของกล้องเว็บแคมด้อยประสิทธิภาพ ถ้าไม่กำจัดพื้นที่นี้ออกจะส่งผลกระทบต่อการแยกคำเพื่อนำมาประมวลผลได้

โดยจะเริ่มจากการ กลับสี (Invert) ของภาพที่ผ่านกระบวนการเทรชโฮลด์อัตโนมัติตามหลักการของออทซึ แล้วนำไปผ่านกระบวนการโอเพนนิ่ง (Opening) ในที่นี้จะกำหนดไว้ว่า พื้นที่ที่มี ขนาดใหญ่ไม่เกิน 15 หน่วย จะถูกตัดออกไป ในภาพตัวอย่างจะทำการเพิ่มสัญญาณรบกวนให้กับภาพเพื่อใช้ทดสอบ ดังรูปที่ 3.7

รูปที่ 3.7 ภาพที่ผ่านกระบวนการโอเพนนิ่ง (Opening)

3.4 การรู้จำอักขระทางภาพ (Optical Character Recognition (OCR))

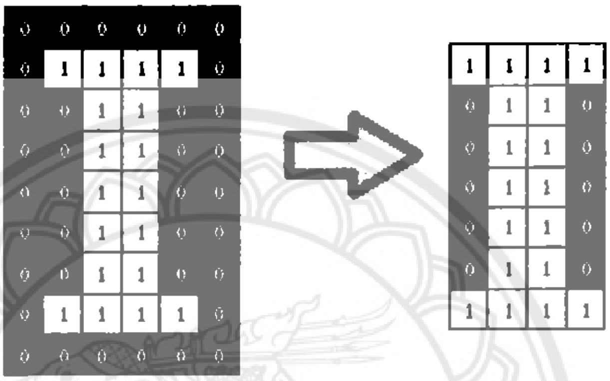
ในส่วนของขั้นตอนการรู้จำอักขระทางภาพของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม มีขั้นตอนการทำงานทั้งหมด ดังรูปที่ 3.8

รูปที่ 3.8 ขั้นตอนการรู้จำอักขระทางภาพ (OCR)

3.4.1 การตัดพื้นที่บริเวณรอบคำ (Crop Image)

เนื่องจากการจับภาพเฉพาะคำที่ต้องการแปลนั้น จะมีพื้นที่ที่เป็นกรอบรอบๆ คำ ซึ่งไม่มีส่วนที่จำเป็นที่ใช้ในการประมวลผลในขั้นตอนการรู้จำอักขระ จึงตัดพื้นที่รอบทิ้งไปก่อนที่จะนำไปประมวลผลในลำดับขั้นตอนต่อไป

โดยพิจารณาตำแหน่งสูงสุดและต่ำสุดของแกนตั้ง และตำแหน่งขวาสุด และซ้ายสุดของแกนนอน ในที่นี้จะพิจารณาพิกเซลสีขาว ซึ่งเป็นพื้นที่ของตัวอักษร ดังรูปที่ 3.9 และ 3.10



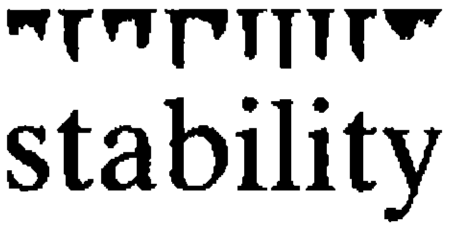
รูปที่ 3.9แสดงการตัดพื้นที่ ให้เหลือเฉพาะพื้นที่ตัวอักษรที่ต้องการ

รูปที่ 3.10 ตัวอย่างการตัดพื้นที่บริเวณรอบคำ (Crop Image)

3.4.2 การตัดตัวอักษรโดยวิธีโปรเจคชั่นในแนวแกนตั้ง (Vertical Projection Profile)

เมื่อทำการตัดพื้นที่ที่ไม่ต้องการเสร็จเรียบร้อยแล้ว ก็จะทำการหาค่ามาเก็บไว้เป็นภาพขาวดำ 2 ระดับ แบบอาร์เรย์หนึ่งมิติ ที่เรียกว่า ฮิสโทแกรม หรือ โปรเจคชั่นค่าของฮิสโทแกรม คือ ผลรวมพิกเซลสีดำ (หรือสีขาวในกรณีกลับค่าสี) เพื่อนำมาแยกตัวอักษรซึ่งพิจารณา ค่าฮิสโทแกรม ที่

มีสีดำ (หรือขาว) ที่เป็นพื้นที่ของตัวอักษรโดยพื้นที่ที่มีค่าเป็นสีตรงข้าม หรือมีค่าฮิสโทแกรมเป็น 0 จะจัดว่าพื้นที่นั้นเป็นช่องว่างระหว่างตัวอักษร ดังรูปที่ 3.11



รูปที่ 3.11 ภาพแสดงฮิสโทแกรมแนวแกนตั้งของตัวอักษร

โดยในโปรแกรมจะตัดตัวอักษรตัวแรกที่ถูกแยกออกมา เพื่อนำไปเปรียบเทียบกับแม่แบบ (Template) สำหรับระบุว่าตัวอักษรในภาพคือตัวอะไรและตัวอักษรที่เหลือ โปรแกรมจะทำการหาตัวอักษรทีละตัวเป็นรอบๆ จนครบทุกตัว ซึ่งมีการทำงานในแต่ละรอบ ดังตารางที่ 3.1

ตารางที่ 3.1 ตารางแสดงรอบการแยกตัวอักษรโดยวิธีโปรเจคชั่นแนวตั้ง

รอบที่	ตัวอักษรแรก	ตัวอักษรที่เหลือ	ปรับขนาด [42x24]
1	s	tability	s
2	t	ability	t
3	a	bility	a
4	b	ility	b
5	i	lity	i
6	l	ity	l
7	i	ty	i
8	t	y	t
9	y	-	y

จากการทำงานในขั้นตอนแยกตัวอักษรโดยวิธีโปรเจคชั่นแนวแกนตั้ง จะสามารถแยกตัวอักษรได้แล้วจะทำการปรับขนาด (Resize) ให้ภาพที่ถูกแยกออกมามีขนาดเท่ากับ ภาพแม่แบบ (Template) ที่ได้ทำขึ้นเพื่อเอาไว้เปรียบเทียบระบุตัวอักษรในขั้นตอนต่อไป

3.4.3 การระบุตัวอักษรโดยวิธีหาค่าสัมประสิทธิ์สหสัมพันธ์ 2 มิติ (2-D Correlation Coefficient)

การหาค่าสัมประสิทธิ์สหสัมพันธ์ คือ การนำภาพที่ผ่านการแยกออกมาเป็นเฉพาะตัวอักษรแต่ละตัวแล้ว มีทำการปรับขนาดให้เท่ากับ ขนาดของภาพแม่แบบ (Template) ซึ่งจะทำให้สามารถนำมาใช้คำนวณได้ตามสมการในทฤษฎี โดยจะเอาภาพที่แยกออกมาเป็นหลักแล้วทำการไล่คำนวณไปทุกตัวอักษรในภาพแม่แบบ (Template) แล้ว เลือกค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสูงสุดคือ ภาพมีความใกล้เคียงภาพแม่แบบ (Template) ตัวนั้นมากที่สุด แล้วมาเทียบดูว่าตัวอักษรนั้นคือตัวอักษรอะไร (ในกรณีที่มีค่าสัมประสิทธิ์สหสัมพันธ์ เท่ากับ 1 คือ ภาพที่แยกมากับภาพรูปแบบ (Template) มีความเหมือนกัน 100%) ซึ่งมีรูปแบบการคำนวณดังรูปที่ 3.12 และ 3.13 และได้ผลการคำนวณดังตารางที่ 3.2

A

0	1	1
0	0	1
0	1	1

Mean (A) = 5/9

B

0	0	1
0	0	1
0	0	1

Mean (B) = 3/9

Correlation Coefficient

$$r = \frac{\sum_N \sum_N (A_{NN} - \bar{A})(B_{NN} - \bar{B})}{\sqrt{((\sum_N \sum_N (A_{NN} - \bar{A}))^2)((\sum_N \sum_N (A_{NN} - \bar{A}))^2)}}$$

$$r = \frac{4\left(-\frac{5}{9}\right)\left(-\frac{3}{9}\right) + 2\left(\frac{4}{9}\right)\left(-\frac{3}{9}\right) + 3\left(\frac{4}{9}\right)\left(\frac{6}{9}\right)}{\sqrt{\left[4\left(-\frac{5}{9}\right)^2 + 5\left(\frac{4}{9}\right)^2\right]\left[6\left(-\frac{3}{9}\right)^2 + 3\left(\frac{6}{9}\right)^2\right]}}$$

$r \approx 0.624$

รูปที่ 3.12 แสดงการคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

A

0	1	1
0	0	1
0	1	1

B

0	1	1
0	0	1
0	1	1

Mean A = 5/9 Mean B = 3/9

Correlation Coefficient

$$r = \frac{\sum_M \sum_N (A_{NN} - \bar{A})(B_{NN} - \bar{B})}{\sqrt{((\sum_M \sum_N (A_{NN} - \bar{A}))^2)((\sum_M \sum_N (A_{NN} - \bar{A}))^2)}}$$

$$r = \frac{4\left(-\frac{5}{9}\right)^2 + 5\left(\frac{4}{9}\right)^2}{\sqrt{\left[4\left(-\frac{5}{9}\right)^2 + 5\left(\frac{4}{9}\right)^2\right]\left[4\left(-\frac{5}{9}\right)^2 + 5\left(\frac{4}{9}\right)^2\right]}}$$

$r = 1$

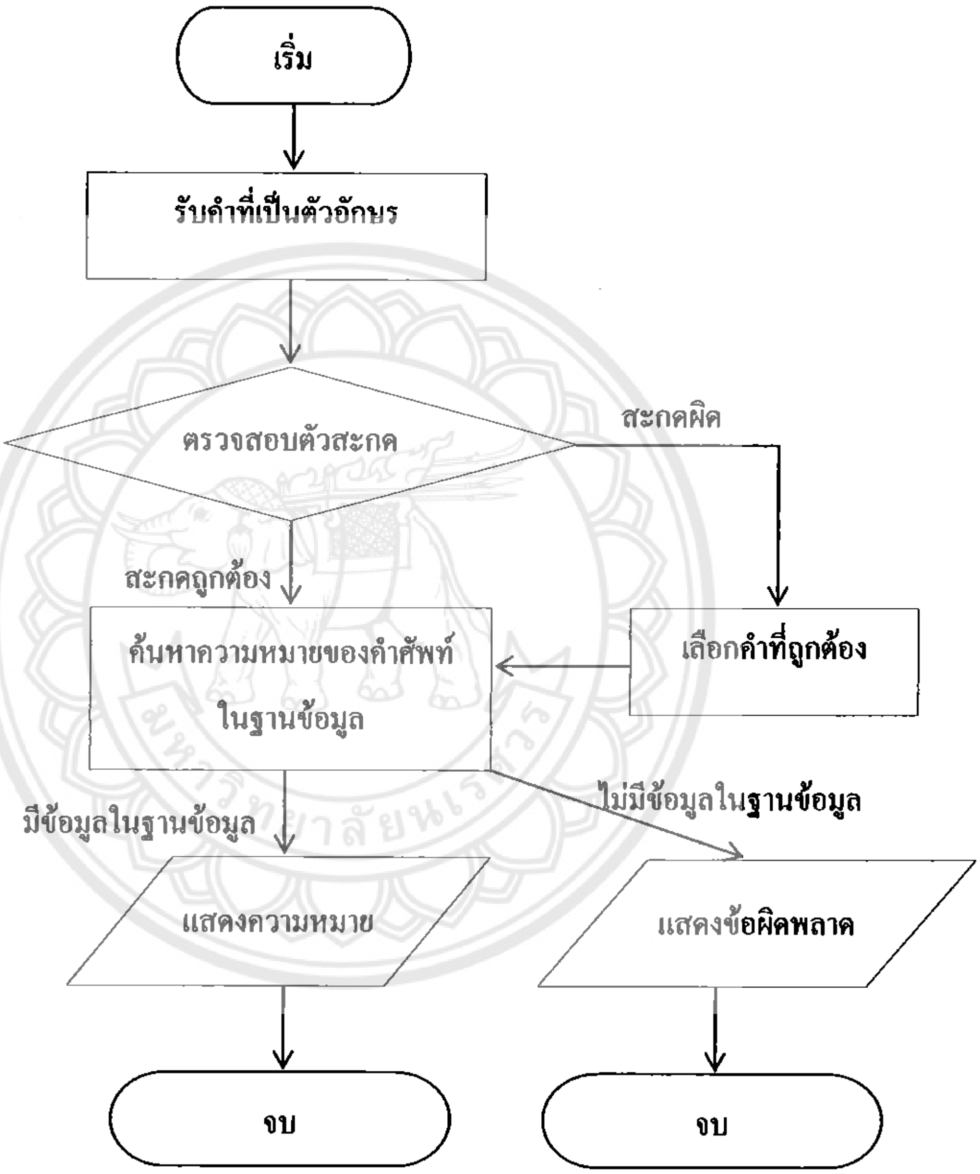
รูปที่ 3.13 แสดงการคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์แบบง่ายๆ

ตารางที่ 3.2 ตารางการระบุตัวอักษรจากค่าสูงสุดของค่าสัมประสิทธิ์สหสัมพันธ์

ตัวอักษรที่	ปรับภาพ [42x24]	แม่แบบ [42x24]	ค่าสัมประสิทธิ์สหสัมพันธ์สูงสุด (-1≤ r≤1)	ตัวอักษร
1			0.832572	s
2			0.739486	t
3			0.681661	a
4			0.699952	b
5			0.742953	i
6			0.878912	l
7			0.869865	i
8			0.710935	t
9			0.600350	y

3.5 การตรวจสอบคำสะกดและค้นหาคำในฐานข้อมูล (Spell Check and Database)

ในส่วนของขั้นตอนการตรวจสอบคำสะกดและค้นหาข้อมูลในฐานข้อมูลของโปรแกรมแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านกล้องเว็บแคม มีขั้นตอนการทำงานทั้งหมด ดังรูปที่ 3.14

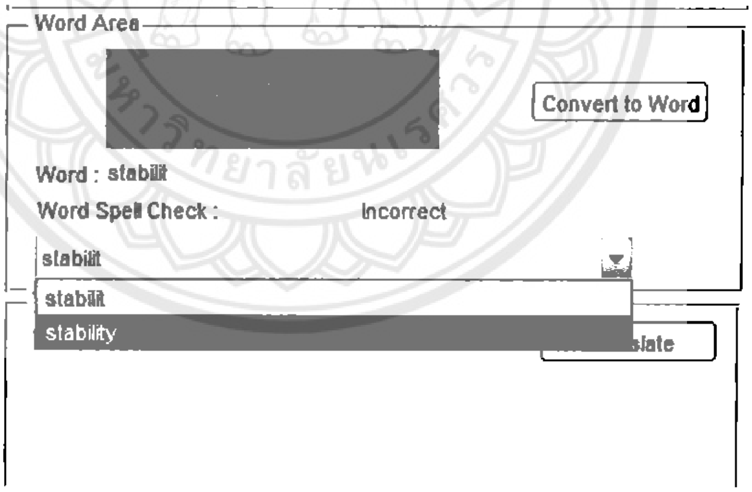


รูปที่ 3.14 ขั้นตอนตรวจสอบคำสะกดและค้นหาข้อมูลในฐานข้อมูล

3.5.1 การตรวจสอบคำสะกด (Spell Check)

โปรแกรมแปลงรูปภาพเป็นตัวอักษรจะทำงานได้ถูกต้องและมีประสิทธิภาพมากยิ่งขึ้นถ้ามีการตรวจสอบการสะกดคำและแสดงคำที่ถูกต้องให้เมื่อมีการผิดพลาดของการแปลงรูปภาพเป็นตัวอักษรกระบวนการตรวจสอบตัวสะกดนี้จะช่วยเพิ่มเปอร์เซ็นต์ความถูกต้องของโปรแกรมการแปลภาษาไทยเป็นอังกฤษผ่านทางเว็บแคมให้มีมากยิ่งขึ้น เพราะถ้าการแปลงรูปภาพเป็นตัวอักษรเกิดความผิดพลาดขึ้นจะทำให้คำที่จะนำไปค้นหาความหมายในฐานข้อมูลผิดพลาดและไม่พบในฐานข้อมูล จึงต้องมีการเสนอแนะคำที่ใกล้เคียงกับคำที่ได้จากการแปลงที่ผิด โดยมีการพิจารณาโดยใช้หลักที่ว่าตัวอักษรแรกจะต้องถูก และมีการเพิ่มและลดความยาวของคำได้ไม่เกิน 2 ตัวอักษร โดยจะเลือกแสดงคำที่มีความคล้ายคลึงกับคำที่แปลงได้จากรูปภาพมากที่สุด

ในส่วนของการแสดงผลคำใกล้เคียงในโปรแกรมได้ใช้การตรวจสอบคำจาก Microsoft Word ซึ่งเป็นโปรแกรมสำเร็จรูปสำหรับการสร้างเอกสารต่างๆ โดยอาศัยอัลกอริทึมการตรวจสอบคำจาก Microsoft Word โดยตรง ซึ่งโปรแกรมจะแสดงคำใกล้เคียงมาให้เลือก ตัวอย่างเช่น ถ้าเกิดจับภาพตกตัวอักษร y ไป 1 ตัวก็จะมีคำที่ถูกต้องแสดงมาให้เลือก ดังรูปที่3.15



รูปที่ 3.15 ภาพแสดงคำใกล้เคียงในกรณีจับภาพโดยตัดตัวอักษร y

การตรวจสอบคำสะกด จะใช้การเชื่อมต่อโปรแกรม (Application Program Interface: API) ของ Microsoft Word ในการตรวจสอบคำ และแสดงคำแนะนำกลับมาเพื่อทำการเลือกคำที่ถูกต้อง

3.5.2 การค้นหาคำในฐานข้อมูล (Database)

การค้นหาคำศัพท์ในฐานข้อมูลเมื่อได้คำที่แปลงจากรูปภาพแล้ว จะนำคำที่ได้ไปค้นหาในฐานข้อมูล (Database) โดยการใช้คำสั่งในการเรียกดูข้อมูลจากฐานข้อมูลมาแสดงออกทางจอภาพ (Query) ในแบบทั่วไป โดยการเรียกดูข้อมูลจากฐานข้อมูลมาแสดงออกทางจอภาพ (Query) ที่นำมาใช้ค้นหาความหมายและประเภทของคำ คือ

```
SELECT    tentry ,ecat
FROM      entry
WHERE     esearch= 'word'
```

โดยที่ฐานข้อมูลมีโครงสร้างดังตัวอย่างในรูปที่ 3.16

entry

esearch	eentry	tentry	ecat	id	ethai	esyn	eant
dog	dog	สุนัข	N	19775	หมา, สัตว์ตระกูล	canine; hound	
dogfish	dogfish	ปลาฉลามหนู	N	19776			
dogged	dogged	ดื้อรั้น	ADJ	19777	ไม่เชื่อฟัง	stubborn	
doggedly	doggedly	อย่างดื้อรั้น	ADV	19778	อย่างไม่เชื่อฟัง		
doggy	doggy	ที่เหมือนสุนัข	ADJ	19779	หมา	doggie	
doggy	doggy	สุนัข(ภาษาเด็ก)	N	19780	หมา	doggie	
doghole	doghole	ที่อยู่ที่คับแคบและสกปรก	N	19781		doghutch	
doghouse	doghouse	บ้านสุนัข	N	19782	ที่พักของสุนัข		

รูปที่ 3.16 ภาพแสดงโครงสร้างในฐานข้อมูล

โดยที่ entry คือ ชื่อของฐานข้อมูล

esearch คือ คอลัมน์ที่แสดงคำศัพท์ทั้งหมด

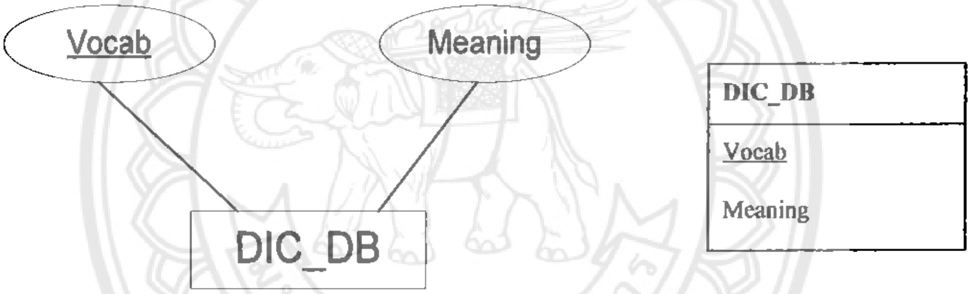
tentry คือ คอลัมน์ที่แสดงความหมายของคำศัพท์

ecat คือ คอลัมน์ที่แสดงประเภทของความหมาย

3.6 การจัดเก็บข้อมูล

3.6.1 การออกแบบฐานข้อมูล (Database Design)

ต้องการสร้างฐานข้อมูลเพื่อเก็บรวบรวมคำศัพท์ทั้งหมดเพื่อนำมาประยุกต์ใช้กับโปรแกรม แปลคำศัพท์ หรือใช้งานอื่นๆ ขั้นตอนการออกแบบ คือ ต้องกำหนดเอนทิตี้ (Entity) หรือถ้าในการเขียนโปรแกรมอาจจะเรียกว่าการประกาศตัวแปร ในการประกาศเอนทิตี้นั้นก็เพื่อเป็นการสร้างตัวแปรมารองรับข้อมูลที่จะนำไปเก็บไว้นั่นเอง ในที่นี้จะกำหนดเอนทิตี้ (Entity) ทั้งหมด 2 ตัวคือ ตัวแปร Vocab ใช้เก็บคำศัพท์ และตัวแปร Meaning เพื่อใช้เก็บความหมาย เมื่อได้เอนทิตี้ (Entity) มาแล้ว สิ่งต่อไปนั้นคือการกำหนดความสำคัญของ เอนทิตี้ ก็คือการกำหนดคีย์หลัก (Primary Key) ซึ่งจะกำหนดให้ ตัวแปร Vocab เป็นคีย์หลัก (Primary Key) สามารถเขียนแสดง ER Diagram ได้ดังรูปที่ 3.17

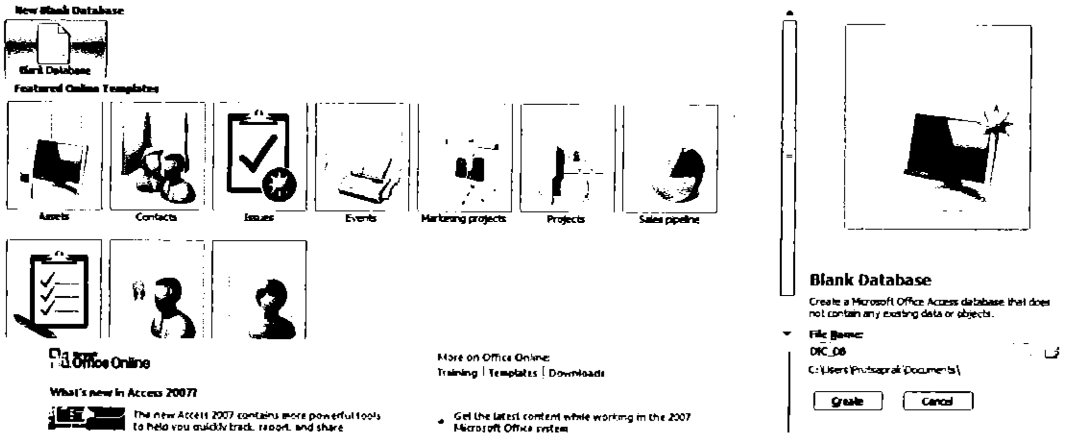


รูปที่ 3.17แสดงความสัมพันธ์ของ ER Diagram

3.6.2 วิธีการสร้างฐานข้อมูล

ฐานข้อมูลถือเป็นแหล่งเก็บรวบรวมคำศัพท์และความหมาย เป็นส่วนที่ Applicationนี้ใช้อ้างอิงถึงเพื่อค้นหา คำที่ต้องการแปล และมีการติดต่อระหว่าง Application กับฐานข้อมูล (database) ในโครงงานนี้ฐานข้อมูลที่ใช้จะเป็นฐานข้อมูลอย่างง่าย ไม่ซับซ้อนมากมาย มีวิธีการสร้างดังนี้

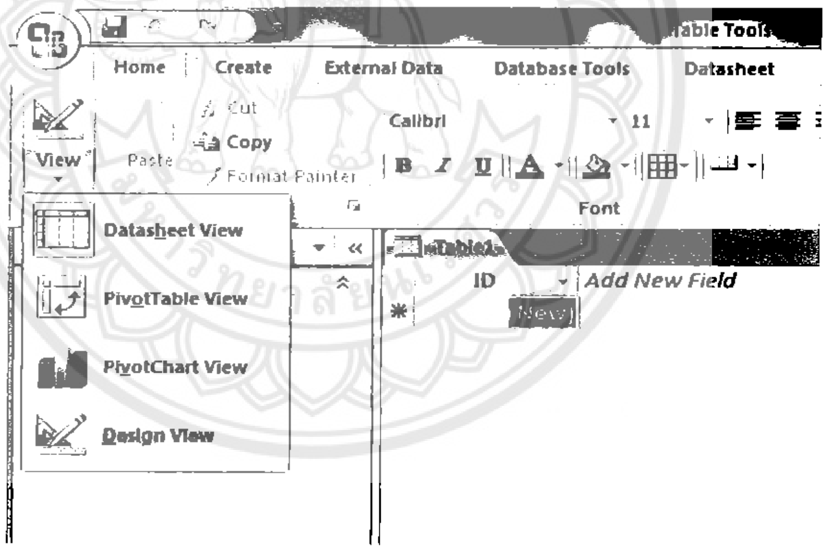
เมื่อเข้าสู่โปรแกรม Microsoft Access 2007 จะพบกับหน้าต่างหน้าแรกขึ้นมา ให้เลือกที่ New Blank Databaseเสร็จแล้วทางด้านขวาจะมีให้ตั้งชื่อฐานข้อมูล (database) ให้ตั้งชื่อเป็น DIC_DB จากนั้นเลือกที่ปุ่มสร้าง (Create) ดังรูปที่ 3.18



รูปที่ 3.18 แสดงการสร้างฐานข้อมูลใหม่

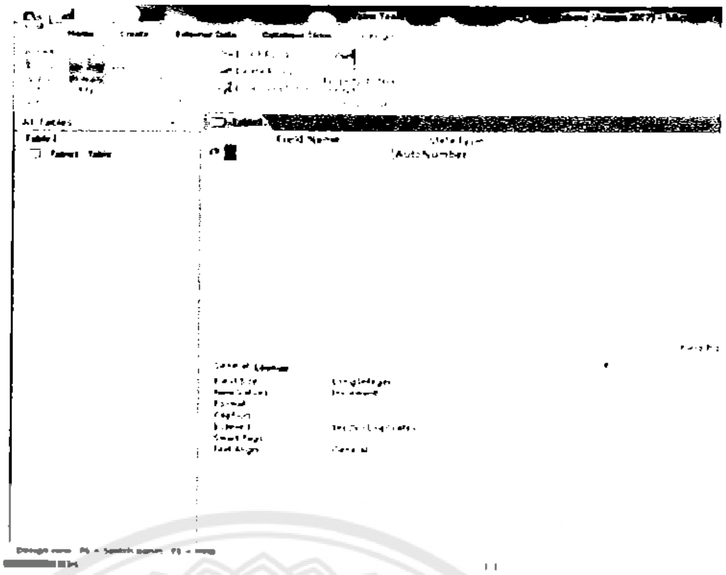
3.6.2.1 ตั้งชื่อให้กับฐานข้อมูล โดยควรจะตั้งชื่อให้สื่อความหมายดังรูปที่ 3.18

3.6.2.2 คลิกที่ปุ่มสร้าง (Create) ก็จะได้ฐานข้อมูลใหม่ตามที่ตั้งชื่อไว้ดังรูปที่ 3.18



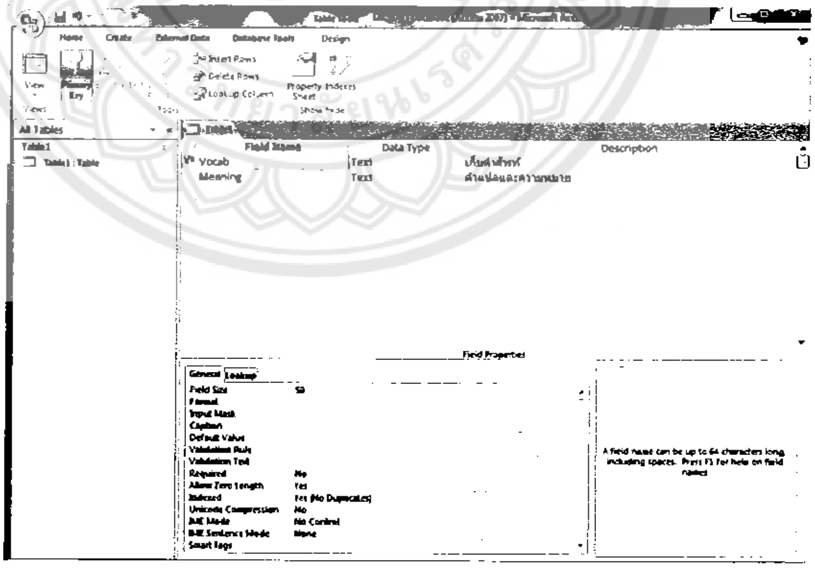
รูปที่ 3.19 แสดงการเลือกมุมมองการออกแบบ (Design View)

3.6.2.3 หลังจากนั้นให้กดที่ View แล้วเลือก Design View ดังรูปที่ 3.19 โดยเมื่อกดไปแล้วจะมีการถามว่าต้องการบันทึกชื่อตารางว่าอะไร ให้ตั้งชื่อตามใจชอบ

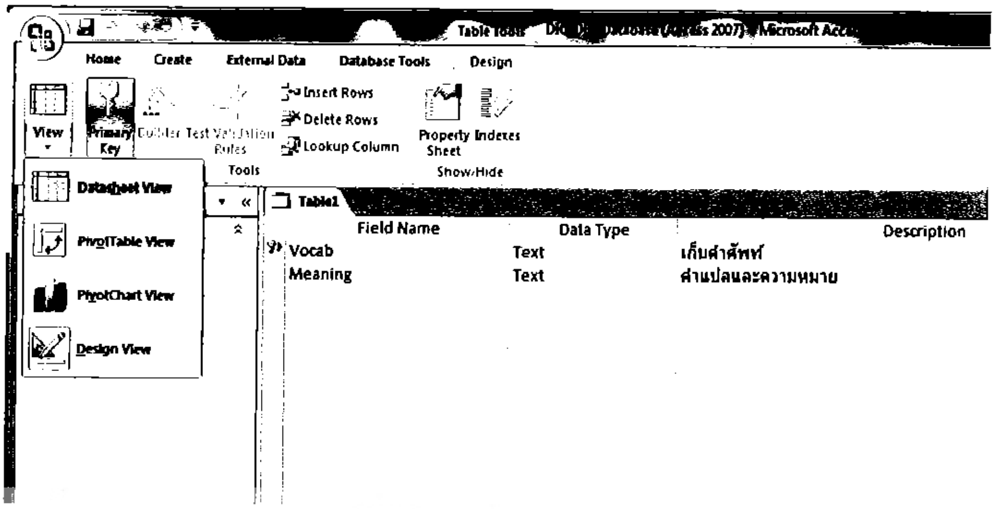


รูปที่ 3.20 แสดงหน้ากำหนดค่าเอนทิตี้ของตาราง

3.6.2.4 หลังจากนั้นก็จะได้หน้าจอ Design View ดังรูปที่ 3.20 **ให้ทำการกำหนดค่าเอนทิตี้** (Entity) และคีย์หลัก (Primary Key) ในที่นี้ จะกำหนดให้ Vocab **เป็นคีย์หลัก** (Primary Key) ดังรูปที่ 3.21

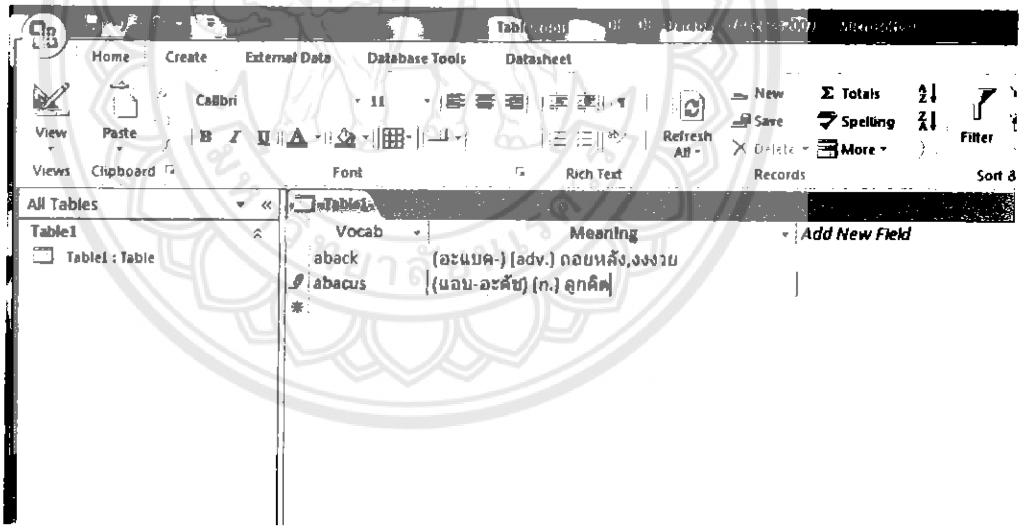


รูปที่ 3.21 แสดงเอนทิตี้ (Entity) ทั้งหมดในตาราง



รูปที่ 3.22 แสดงการกลับไปที่ Datasheet View

3.6.2.5 จากนั้นให้ทำการคลิกที่ View แล้วเลือก Datasheet View ดังรูปที่ 3.22 แล้วเลือกบันทึกข้อมูลในตาราง



รูปที่ 3.23 แสดงการป้อนข้อมูลคำศัพท์

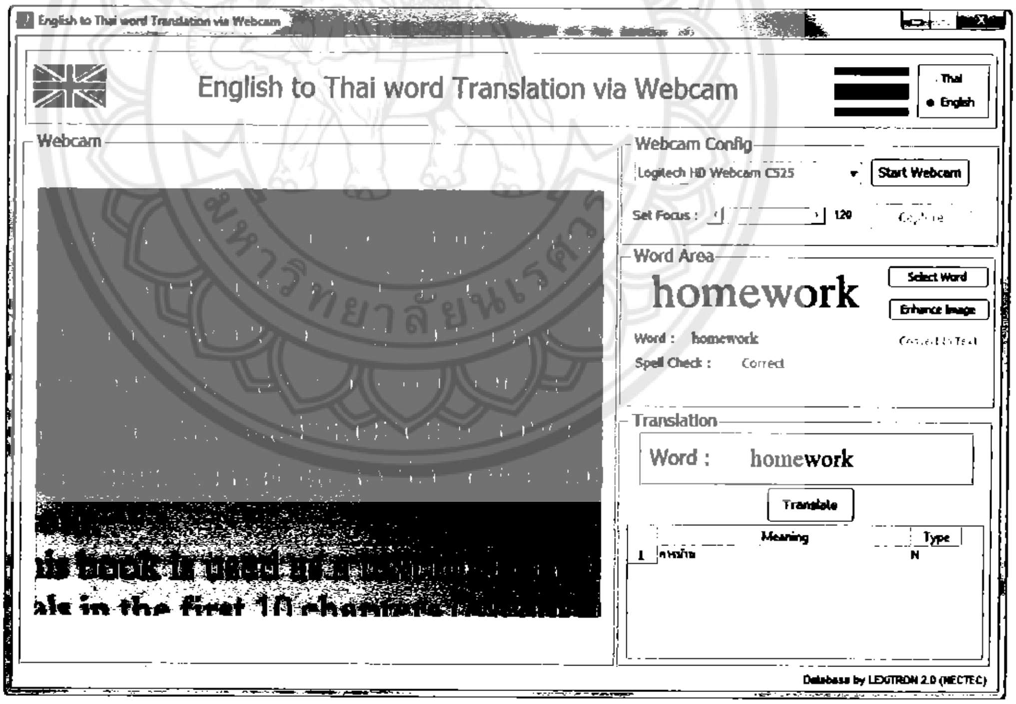
3.6.2.6 หลังจากนั้นให้ทำการป้อนคำศัพท์จนครบ แล้วทำการบันทึกข้อมูลดังรูป รูปที่ 3.23 ก็จะได้ไฟล์ฐานข้อมูลพจนานุกรม

บทที่ 4

ผลการทดลอง

4.1 ส่วนของโปรแกรม

การแปลคำศัพท์ภาษาอังกฤษเป็นภาษาไทยผ่านทางกล้องเว็บแคมโดยรวบรวมกระบวนการต่างๆ เริ่มตั้งแต่ การเตรียมภาพ การหาตำแหน่งของคำที่ต้องการแปล การแปลงภาพเป็นแบบไบนารี่ การแปลงรูปภาพเป็นตัวอักษร การแปลคำศัพท์ และสุดท้ายเป็นการแนะนำคำใกล้เคียงเมื่อมีความผิดพลาดในการแปลงรูปภาพเป็นตัวอักษร ซึ่งโปรแกรมนี้นำส่วนต่างๆมาเชื่อมโยงกัน และจัดหน้าตาของโปรแกรมให้ใช้งานได้ง่าย ดูเป็นระเบียบเรียบร้อย จะทำได้โดยการสร้าง GUI ดังรูปที่ 4.1



รูปที่ 4.1 แสดง GUI ของโปรแกรม

4.2 ผลการทดลองส่วนของการเตรียมข้อมูลภาพ

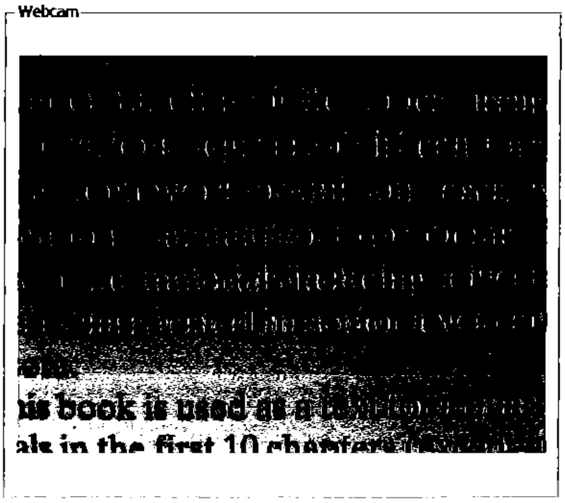
ในส่วนของการเตรียมข้อมูลภาพจะเป็นส่วนที่ทำการจับภาพ (capture) รูปภาพ ผ่านทางกล้องเว็บแคม โดย ในการทดลองขั้นแรกจะเป็นส่วนของการเตรียมภาพ โดยจะมีการจับภาพ (capture) ผ่านทางเว็บแคม เพื่อตรวจสอบดูว่าภาพที่ทำการถ่ายมานั้น ตรงกับขอบเขตภาพที่ต้องการหรือไม่ โดยปัญหาที่มักจะพบเจอคือ กล้องเว็บแคมที่มีขนาดความละเอียดน้อยและไม่สามารถปรับโฟกัสได้ ดังรูปที่ 4.2 (ก) เป็นภาพที่ถูกถ่ายด้วยกล้องเว็บแคมที่มีความละเอียดต่ำและไม่สามารถปรับโฟกัสของภาพได้ จึงทำให้รายละเอียดของภาพที่จำเป็นต่อการแปลงเป็นตัวอักษรขาดหายไปอย่างมาก ดังนั้นจึงต้องมีการนำกล้องเว็บแคมที่มีความละเอียดเพียงพอและสามารถปรับโฟกัสได้ดังรูปที่ 4.2 (ข) ซึ่งจะทำให้ภาพที่ได้มีรายละเอียดที่จำเป็นต่อการแปลงเป็นตัวอักษรครบถ้วน

(ก) (ข)

รูปที่ 4.2 (ก) แสดงภาพที่ถ่ายด้วยกล้องเว็บแคมที่ไม่สามารถปรับโฟกัสของภาพได้

(ข) แสดงภาพที่ถ่ายด้วยกล้องเว็บแคมที่สามารถปรับโฟกัสของภาพได้

การทดลองจะใช้ภาพที่ประกอบไปด้วยตัวอักษรที่มีรูปแบบอักษร Times New Roman ขนาด 16 ขึ้นไป และอยู่ในเงื่อนไขขอบเขตของโครงงาน เมื่อได้ภาพที่ต้องการแล้ว ขั้นต่อไปคือการนำภาพเข้าสู่ระบบของโปรแกรมโดยการจับภาพ (capture) ดังรูปที่ 4.3



รูปที่ 4.3 แสดงภาพที่ถูกจับภาพ (capture)

4.3 ผลการทดลองส่วนของการหาตำแหน่งของคำที่ต้องการแปล

เมื่อทำการจับภาพ (Capture) แล้ว ขั้นตอนต่อไปคือการเลือกเฉพาะคำที่ต้องการภายในภาพ ด้วยการคลิกเลือกตำแหน่งมุมบนและบนล่างของภาพ และนำภาพที่ได้นี้ไปทำการเทรชโฮลด์แบบอัตโนมัติต่อไป ซึ่งจะเห็นว่าภาพหลังจากการทำเทรชโฮลด์แบบอัตโนมัติแล้ว จะแสดงให้เห็นถึงความสว่างของภาพที่พอดีหรือไม่ ดังรูปที่ 4.4

(ก) (ข) (ค)

รูปที่ 4.4 แสดงความสว่างของแสงที่มีผลต่อการทำเทรชโฮลด์ (Thresholding)

(ก) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างไม่เพียงพอ

(ข) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างมากเกินไป

(ค) แสดงการทำเทรชโฮลด์ (Thresholding) ของภาพที่มีแสงสว่างพอดี

4.4 ผลการทดลองส่วนของการแปลงภาพเป็นตัวอักษร

4.4.1 การทดลองแปลงภาพเป็นตัวอักษรทีละตัว

การทดลองนี้จะทำการทดสอบความสามารถในการแปลงภาพเป็นตัวอักษรแบบ Times New Roman เพื่อวัดประสิทธิภาพในการระบุตัวอักษร ซึ่งในการทดลองจะจำกัดสภาพแวดล้อมเดียวกัน ขนาดอักษร 16 หน่วย ทำการทดลองตัวอักษรภาษาอังกฤษพิมพ์ใหญ่และตัวอักษรภาษาพิมพ์เล็ก ตัวอักษรละ 40 ครั้ง ได้ผลการทดลอง ดังตารางที่ 4.1 และ 4.2

ตารางที่ 4.1 ทดสอบความสามารถในการระบุอักษรภาษาอังกฤษตัวพิมพ์ใหญ่

ตัวอักษร	ระบุตัวอักษรได้(ร้อยละ)	ตัวอักษรที่ผิดพลาด(ร้อยละ)
A	100%	-
B	95%	E (5%)
C	100%	-
D	90%	O (10%)
E	90%	B (10%)
F	85%	P (12.5%), E (2.5%)
G	90%	O (10%)
H	95%	R (5%)
I	75%	l (แอลเล็ก) (25%)
J	95%	I (5%)
K	95%	R (5%)
L	100%	-
M	100%	-
N	100%	-
O	100%	-
P	100%	-
Q	100%	-
R	97.5%	B (2.5%)
S	100%	-
T	100%	-
U	100%	-

ตัวอักษร	ระบุตัวอักษรได้(ร้อยละ)	ตัวอักษรที่ผิดพลาด(ร้อยละ)
V	97.5%	U (2.5%)
W	100%	-
X	100%	-
Y	92.5%	V (7.5%)
Z	100%	-

ตารางที่ 4.2 ทดสอบความสามารถในการระบุอักษรภาษาอังกฤษตัวพิมพ์เล็กตัวพิมพ์เล็ก

ตัวอักษร	ระบุตัวอักษรได้(ร้อยละ)	ตัวอักษรที่ผิดพลาด(ร้อยละ)
a	95%	s (5%)
b	100%	-
c	100%	-
d	100%	-
e	85%	c (15%)
f	85%	t (15%)
g	100%	-
h	90%	b (10%)
i	100%	-
j	95%	i (5%)
k	100%	-
l	85%	I (15%)
m	100%	-
n	97.5%	o (2.5%)
o	100%	-
p	100%	-
q	100%	-
r	95%	l (แอลเล็ก) (5%)
s	100%	-
t	92.5%	l (แอลเล็ก) (7.5%)
u	97.5%	o (2.5%)

ตัวอักษร	ระบุตัวอักษรได้(ร้อยละ)	ตัวอักษรที่ผิดพลาด(ร้อยละ)
v	97.5%	u (2.5%)
w	100%	-
x	100%	-
y	95%	v (5%)
z	100%	-

จากผลการทดลองการระบุอักษรตัวพิมพ์ใหญ่จะพบว่าสามารถระบุตัวอักษรคิดเป็นร้อยละ 96.07 และสามารถระบุอักษรตัวพิมพ์เล็ก คิดเป็นร้อยละ 96.53 ซึ่งตัวอักษรที่มีโอกาสผิดพลาดนั้นจะมาความคล้ายคลึงกัน ส่งผลให้การระบุตัวอักษรนั้นผิดพลาดได้ โดยตัวพิมพ์ใหญ่ที่ผิดพลาดคิดเป็นร้อยละ 3.83 และตัวพิมพ์เล็กที่ผิดพลาดคิดเป็นร้อยละ 3.47

4.4.2 การทดลองแปลงภาพเป็นตัวอักษรโดยใช้รูปแบบอักษรที่ต่างกัน

การทดลองนี้จะทดสอบความสามารถในการระบุตัวอักษร ในกรณีที่แบบตัวหนังสือมีความแตกต่างกัน ในการทดลองจะทำการความคุมสภาพแวดล้อมภายนอกเหมือนกัน ขนาดแบบอักษร 16 หน่วย ทำการทดลองจำนวน 10 ครั้งต่อหนึ่งตัวอักษร จะได้ค่าเฉลี่ยและตัวอักษรที่ระบุผิดพลาด ดังตารางที่ 4.3

ตารางที่ 4.3 แสดงการเปรียบเทียบการแปลงภาพเป็นตัวอักษรโดยใช้รูปแบบอักษรที่ต่างกัน

รูปแบบตัวอักษร	ความถูกต้อง(ร้อยละ)	ตัวพิมพ์เล็กที่ผิดพลาด	ตัวพิมพ์ใหญ่ที่ผิดพลาด
Serif Font			
Times New Roman	100%	-	-
Georgia	98%	e	-
Book Antiqua	94.2%	e f	G
Sans-serif Font			
Arial	67.3%	e u n y g	H E Q U I B R F P L D G
Tahoma	61.5%	a e d b r n f m p y g	H E U B F P L D G
Cordia New	51.9%	h e u b r n f m p y d g	H E Q U I B R W F P L D G

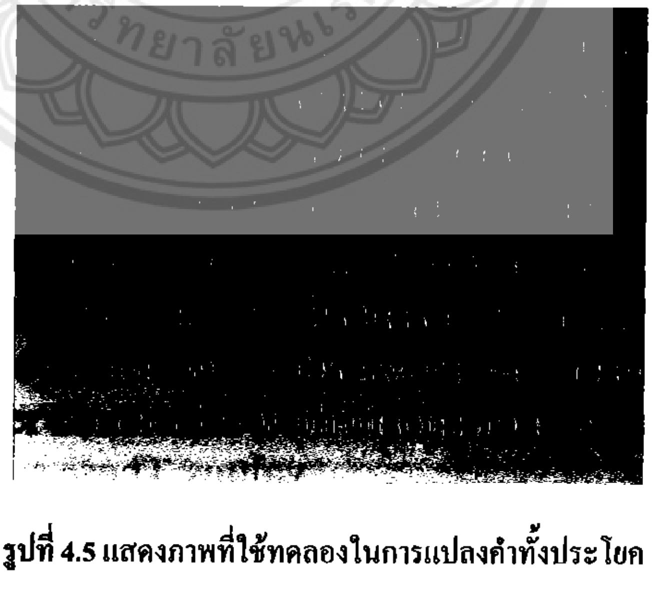
จากการทดลองสามารถสรุปได้ว่าตัวอักษรประเภท serif (แบบมีหัว)จะสามารถระบุตัวอักษรได้มากกว่าทั้งนี้เพราะ อักษรแบบ serif นั้นจะมีลักษณะคล้ายคลึงกับแบบ Template ของแบบอักษร Times New Roman ซึ่งก็เป็นประเภท serif เหมือนกัน แต่ในอีกประเภท คือ Sans-serif (ไม่มีหัว) จะสามารถระบุได้เพียงบ้างตัวเท่านั้น

4.4.3 การทดลองแปลงคำศัพท์ทั้งประโยค

การทดลองนี้จะอาศัย Gaussian blur ซึ่งมีสูตรการคำนวณ ดังนี้

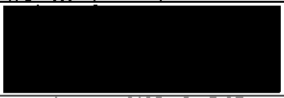
$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{4.1}$$

เพื่อช่วยในการปรับภาพให้มีความเบลอในระดับต่างๆ โดยอาศัยค่า σ (sigma) เป็นตัวกำหนดความเบลอของภาพ ใช้ภาพ ดังรูปที่ 4.5 โดยเลือกคำศัพท์ที่สมบูรณ์ โดย ควบคุมสภาพแวดล้อมเดียวกัน และทำการทดลองกับคำศัพท์ 30 คำ จำนวนการทดลอง 3 ครั้ง จะได้ค่าเฉลี่ยดังตารางที่ 4.4

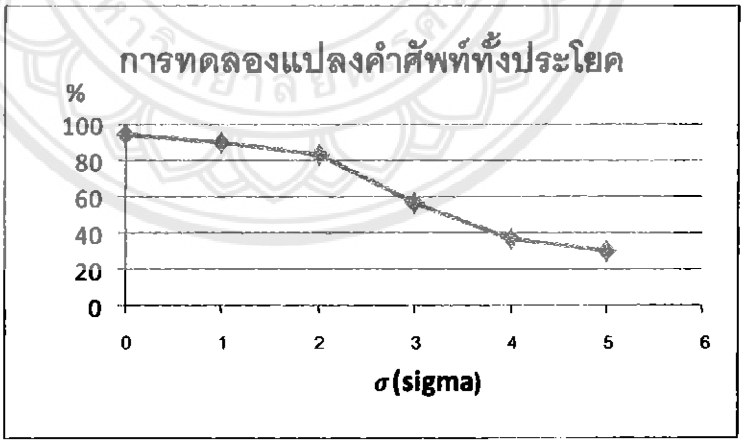


รูปที่ 4.5 แสดงภาพที่ใช้ทดลองในการแปลงคำทั้งประโยค

ตารางที่ 4.4 การทดลองแปลงภาพเป็นตัวอักษรทั้งประโยค

ค่า σ (sigma)	ภาพตัวอย่าง	ระบุคำได้(%)
0		94.4%
1		90%
2		83.3%
3		56.6%
4		36.6%
5		25%

การทดลองนี้ทำเพื่อแสดงให้เห็นถึงความสามารถในการแปลงภาพเป็นตัวอักษร ในสภาพที่ตัวอักษรมีความคมชัดแตกต่างกัน สรุปการทดลองแปลงคำศัพท์ทั้งประโยครูปได้ดังนี้



รูปที่ 4.6 แสดงกราฟผลลัพธ์ของความคมชัดของตัวอักษร

จากผลการทดลองพบว่าเมื่อทำการเพิ่มค่า sigma เพื่อให้ตัวอักษรมีความเบลอ จะทำให้การประมวลผลภาพเพื่อแปลงเป็นตัวอักษรมีความผิดพลาดขึ้นตามระดับความเบลอที่สูงขึ้น เนื่องจากกระบวนการประมวลผลขั้นต้นจะทำการปรับค่า เทรชโฮลด์ ทำให้ตัวอักษรที่มีความเบลอนั้น ติดกัน ส่งผลให้ขั้นตอนการตัดภาพนั้นระบุว่าเป็นตัวอักษรเดียวกัน

4.4.4 การทดลองระบุคำที่มีความเอียง

การทดลองนี้เป็นการปรับมุมเอียงเพื่อทดสอบความสามารถในการระบุตัวอักษรว่าโปรแกรมจะสามารถแยกแยะตัดอักษรได้ในมุมเอียงใด ซึ่งการทดลองจะทำการปรับมุมเอียง โดยใช้ฟังก์ชั่นหมุนภาพในองศาต่างๆ แล้วทำการประมวลผลภาพ เป็นคำศัพท์ 30 คำ จำนวน 5 ครั้งจะได้ค่าเฉลี่ยออกมาดังตารางที่ 4.5

ตารางที่ 4.5 การทดลองแปลงภาพเป็นตัวอักษรในมุมเอียงที่ต่างกัน

มุมเอียง(องศา)	ภาพตัวอย่าง	ผลลัพธ์
-10	[redacted]	0%
-8	[redacted]	6.6%
-6	[redacted]	33.3%
-4	[redacted]	76.6%
-2	[redacted]	100%
0	[redacted]	100%
2	[redacted]	96.6%
4	[redacted]	90%
6	[redacted]	56.6%
8	[redacted]	**10%**
10	[redacted]	**6.6%**

สรุปผลการทดลองแปลงคำศัพท์ที่มีความเอียง ได้ดังรูปที่

รูปที่ 4.7 แสดงกราฟผลลัพธ์ของความเอียงของตัวอักษร

จากการทดลองทำการเอียงภาพคำศัพท์จะพบว่าเมื่อองศาเปลี่ยนแปลงไปมากเท่าไรก็จะส่งผลให้ความสามารถในการแยกแยะตัวอักษรมีความผิดพลาดมากขึ้น เนื่องด้วยคำศัพท์บ้างคำที่มีจำนวนตัวอักษรเยอะและมีตัวอักษรที่มีความสูง เช่น l, k, h f, เป็นต้น จะทำให้มุมเอียงที่สามารถระบุได้น้อยลงไปด้วย มุมเอียงที่มีมุม -4 ถึ 4 องศา ถือว่าเป็นมุมที่การประมวลยังคงมีประสิทธิภาพ

4.5 ผลการทดลองในส่วนของการแนะนำคำใกล้เคียง

การทดลองในส่วนของการแนะนำคำใกล้เคียง เพื่อแสดงให้ทราบว่าเมื่อเกิดข้อผิดพลาดในการจับภาพหรือการประมวลผลภาพ ซึ่งสาเหตุอาจเกิดจากผู้ใช้จับภาพไม่สมบูรณ์ หรือสภาพแวดล้อมไม่เหมาะสม จะทำให้การระบุตัวอักษรผิดพลาดได้ ดังตารางที่ 4.6 ถึง ตารางที่ 4.12

ตารางที่ 4.6 แสดงการแนะนำคำใกล้เคียงของคำว่า chapter

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
chapter	ถูก	-	-	-	-
chapte	ผิด	เจอ	chapter	chaste	chapters
hapter	ผิด	เจอ	halter	hater	chapter
chapt	ผิด	เจอ	chap	chat	chapter

ตารางที่ 4.7 แสดงการแนะนำคำใกล้เคียงของคำว่า control

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
control	ถูก	-	-	-	-
contro	ผิด	เจอ	control	contra	contort
ontrol	ผิด	เจอ	control	onto	-
contr	ผิด	เจอ	contra	control	contras

ตารางที่ 4.8 แสดงการแนะนำคำใกล้เคียงของคำว่า detail

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
detail	ถูก	-	-	-	-
detai	ผิด	เจอ	detail	detain	details
etail	ผิด	เจอ	entail	email	detail
deta	ผิด	เจอ	data	delta	detail

ตารางที่ 4.9 แสดงการแนะนำคำใกล้เคียงของคำว่า stability

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
stability	ถูก	-	-	-	-
stabilit	ผิด	เจอ	stability	-	-
tability	ผิด	เจอ	stability	ability	debility
stabili	ผิด	เจอ	stabile	stabiles	stability

ตารางที่ 4.10 แสดงการแนะนำคำใกล้เคียงของคำว่า analysis

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
analysis	ถูก	-	-	-	-
analysi	ผิด	เจอ	analysis	analyze	analyses
nalysis	ผิด	เจอ	analysis	analyses	-
analys	ผิด	เจอ	analyst	analyze	analysis

ตารางที่ 4.11 แสดงการแนะนำคำใกล้เคียงของคำว่า dynamic

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
dynamic	ถูก	-	-	-	-
dynami	ผิด	เจอ	dynamic	dynamo	dynamics
ynamic	ผิด	-	-	-	-
dynam	ผิด	เจอ	dynamo	dynamic	dynamos

ตารางที่ 4.12 แสดงการแนะนำคำใกล้เคียงของคำว่า transient

คำศัพท์	สะกดคำ	ผลลัพธ์	คำแนะนำที่ 1	คำแนะนำที่ 2	คำแนะนำที่ 3
transient	ถูก	-	-	-	-
transien	ผิด	เจอ	transient	transience	transients
ransient	ผิด	-	-	-	-
transie	ผิด	เจอ	transit	transient	transited

จากตารางที่ 4.6 ถึง 4.12 พบว่า ผลของการทดลองในส่วนของการแนะนำคำใกล้เคียง โดยการทดลองตัดตัวอักษรของคำที่ต้องการแปลไปทีละตัว โดยเลือกตัดในส่วนของตัวหน้าสุดและตัวหลังสุด จะพบว่าการตัดตัวอักษรในส่วนหลังสุด การแนะนำคำใกล้เคียงของโปรแกรมจะสามารถแนะนำคำที่ต้องการที่จะค้นหาได้ทุกคำ แต่ในส่วนของการตัดตัวอักษรในตัวแรกสุดออกไป จะทำให้การแนะนำคำใกล้เคียงของโปรแกรมทำงานได้ไม่ถูกต้องเท่าที่ควร

4.6 ผลการทดลองส่วนของการแปลคำศัพท์

การหาความหมายของคำศัพท์ สามารถทำได้โดยการนำคำศัพท์ที่ได้จากการแปลงรูปภาพให้เป็นตัวอักษรไปเปรียบเทียบกับคำศัพท์ที่มีอยู่ในฐานข้อมูล ซึ่งถ้าคำที่ต้องการค้นหามีอยู่ในฐานข้อมูล ความหมายของคำศัพท์ก็จะแสดงขึ้นมาทันที ดังตารางที่ 4.13

ตารางที่ 4.13 แสดงความหมายของคำเมื่อกระบวนการแปลงภาพเป็นตัวอักษรถูกต้อง

คำศัพท์	ความหมายคำที่ 1	ความหมายคำที่ 2	ความหมายคำที่ 3
chapter	บท (ของหนังสือหรืองานเขียน)	ส่วนที่สำคัญ (ของประวัติศาสตร์หรือชีวิตคน)	สาขา (ขององค์กรสโมสรหรือสมาคม)
control	การจำกัด	ควบคุมดูแล	เครื่องควบคุม
detail	รายละเอียด	อธิบายโดยละเอียด	-
stability	เสถียรภาพ	ความเด็ดเดี่ยว	-
analysis	การวิเคราะห์	จิตวิเคราะห์	ผลการวิเคราะห์
dynamic	เต็มไปด้วยพลังและความคิดสร้างสรรค์	-	-
transient	ชั่วคราว	ชั่วประเดี๋ยว	เกิดผลกระทบนอกจิตใจ

ถ้าคำที่ต้องการค้นหาความหมายไม่มีอยู่ในฐานข้อมูล ซึ่งอาจจะเกิดสาเหตุมาจากกระบวนการแปลงรูปภาพเป็นตัวอักษรผิดพลาด ด้วยเหตุนี้จะทำให้ไม่สามารถทราบความหมายของคำศัพท์ที่ต้องการได้ ปัญหาเหล่านี้สามารถแก้ไขได้โดยการใช้ การแนะนำคำใกล้เคียง (Word suggestion)

บทที่ 5

บทสรุป

5.1 วิเคราะห์และสรุปผลการทดลอง

เนื่องจากโครงงานนี้ ใช้การประมวลผลภาพเป็นหลัก ทำให้การใช้งานโปรแกรมค่อนข้างมีข้อจำกัด และปัจจัยที่ส่งผลกระทบต่อการประมวลผลภาพมีอยู่หลากหลายรูปแบบไม่ว่าจะเป็นสภาพแวดล้อมภายนอก คุณภาพของเอกสารที่ต้องการแปล คุณภาพของกล้องเว็บแคม และการจับภาพ เป็นต้น เมื่อปัจจัยเหล่านี้ส่งผลกระทบจะทำให้การระบุตัวอักษรมีความถูกต้องน้อยลง ผลลัพธ์ที่ได้ก็จะผิดพลาดไปด้วย

ในการทดลองจะทำการตรวจสอบประสิทธิภาพของการประมวลผลภาพและระบุตัวอักษรเป็นหลัก โดยอาศัยการจำลองปัจจัยที่มีผลกระทบเบื้องต้น เพื่อทดสอบว่าผลลัพธ์ของโปรแกรมว่ามีขีดจำกัดเท่าใด ที่จะทำให้การประมวลผลนั้นยังมีความถูกต้องสมบูรณ์อยู่

จากผลการทดลองพบว่า ปัจจัยที่ส่งผลกระทบมากที่สุดในการประมวลผลภาพ คือ ความคมชัดของตัวอักษร ซึ่งสาเหตุส่วนใหญ่ที่ทำให้การประมวลผลภาพมีความผิดพลาด เกิดจากความละเอียดของกล้องเว็บแคมน้อย ไม่สามารถปรับโฟกัสได้ ทำให้ตัวอักษรที่ทำการจับภาพไม่สามารถระบุได้ว่าตัวอักษรนั้นคือตัวใด ซึ่งมาจากการทดลองใช้การเบลอภาพ โดยวิธีของเกาซ์เซี่ยน(Gaussian blur) จำลองความเบลอของตัวอักษร และอีกหนึ่งปัจจัยที่ส่งผลกระทบคือ มุมเอียงต่างๆ ซึ่งสาเหตุอาจเกิดจากขั้นตอนการรับภาพผู้ใช้อาจจับเอกสารไม่ตรง ทำให้เกิดมุมเอียงซึ่งจะส่งผลต่อการระบุตัวอักษร โดยมุมที่โปรแกรมสามารถระบุตัวอักษรได้ถูกต้องสมบูรณ์จะอยู่ในช่วง มุม -4 ถึง 4 องศา ปัจจัยสุดท้ายที่ส่งผลกระทบคือ ตัวอักษรที่มีความคล้ายคลึงกัน สาเหตุเช่นเดียวกับการทดลองข้างต้น โดยโปรแกรมจะสามารถระบุตัวอักษรได้ร้อยละ 96.3 ซึ่งถือว่ามีประสิทธิภาพ ส่วนตัวอักษรที่มักเกิดข้อผิดพลาดคือตัว l (แอลเล็ก) และ I (ไอใหญ่) ซึ่งตัวอักษรมีลักษณะใกล้เคียงกัน และรูปแบบอักษรที่กำหนดคือ Times New Roman เมื่อเปรียบเทียบกับแบบอักษรอื่นที่เป็นประเภท serif (มีหัว) เหมือนกันจะสามารถระบุตัวอักษรได้มากกว่า แบบอักษรประเภท Sans-serif (ไม่มีหัว) ส่วนการแนะนำคำศัพท์นั้น จะพบว่าตัวอักษรแรกมีความสำคัญมากสุดในการแนะนำคำศัพท์ที่ใกล้เคียง และการแสดงความหมายของคำศัพท์เป็นภาษาไทยก็จะขึ้นอยู่กับฐานข้อมูลที่ใช้ว่ามีข้อมูลมากเท่าไร โปรแกรมจึงจะทำงานได้อย่างสมบูรณ์

5.2 ปัญหาและแนวทางแก้ไข

จากการทดลองเพื่อหาปัจจัยที่ส่งผลกระทบต่อการทำงานของโปรแกรม ทำให้ทราบถึงสาเหตุต่างๆ ที่ทำให้เกิดข้อผิดพลาดขึ้น โดยส่งผลให้ไม่สามารถทำการแปลภาษาอังกฤษเป็นภาษาไทยได้อย่างสมบูรณ์ ซึ่งพบปัญหาอุปสรรคในการใช้งานและแนวทางแก้ไขปัญหาดังนี้

ปัญหาและอุปสรรค	แนวทางการแก้ไข
1. สภาพแวดล้อมที่มีแสงสว่างไม่เพียงพอ	1. หาสถานที่ที่มีความสว่างมาเพียงพอ
2.ความละเอียดของกล้องเว็บแคมไม่เพียงพอ	2. เพิ่มความละเอียดของกล้องเว็บแคม
3.ระบุตัวอักษรได้ไม่สมบูรณ์ 100%	3. เพิ่มการประมวลผลขั้นต้นและขั้นตอนการรู้จำให้มีประสิทธิภาพมากยิ่งขึ้น
4. มุมเอียงมากเกินไป	4. เพิ่มการปรับภาพมุมเอียงกลับเป็นปกติ

ดังนั้นในทางปฏิบัติถ้าสามารถขจัดปัญหาและอุปสรรคดังกล่าวได้ก็จะสามารถเพิ่มประสิทธิภาพการทำงานของโปรแกรมได้สมบูรณ์มากขึ้น

5.3 แนวทางการพัฒนาในอนาคต

1. พัฒนาโปรแกรมให้สามารถรองรับภาพที่มีรายละเอียดต่ำและไม่สมบูรณ์ให้ดีขึ้น

2.ปรับปรุงในส่วนของการประมวลผลขั้นต้นและการรู้จำตัวอักษรให้ดีขึ้น

3. ปรับปรุงให้สามารถใช้กับรูปแบบอักษรที่หลากหลายมากขึ้น

4.ปรับปรุงให้สามารถแปลภาษาไทยเป็นภาษาอังกฤษได้

เอกสารอ้างอิง

[1] Gonzalez, R.C. and Richard E. Woods. (2001). Digital Image Processing. New Jersey : Prentice-Hall. Inc.

[2] Baxes, Gregory A. (1994). Digital Image Processing. Canada : John Wiley & Sons. Inc.

[3] "Digital Image Fundamental." [online]. Available : http://www.ecpe.nu.ac.th/panomkhawn/imagepro/pdf/ch02-part2.pdf. 2554.

[4] "Image Segmentation." [online]. Available : Http://fivedots.coe.psu.ac.th/~montri/Teaching/image/segment.DOC. 2554.

[5] "Morphological Image Processing." [online]. Available : Http://www.ecpe.nu.ac.th/panomkhawn/imagepro/pdf/ch09.pdf. 2554.

[6] "Thresholding." [online]. Available : Http://www.wbi.msu.ac.th/file/648/doc_45.ppt. 2554.

[7] "Morphological Image Processing." [online]. Available : Http://fivedots.coe.psu.ac.th/~montri/Teaching/image/morph.DOC. 2554.

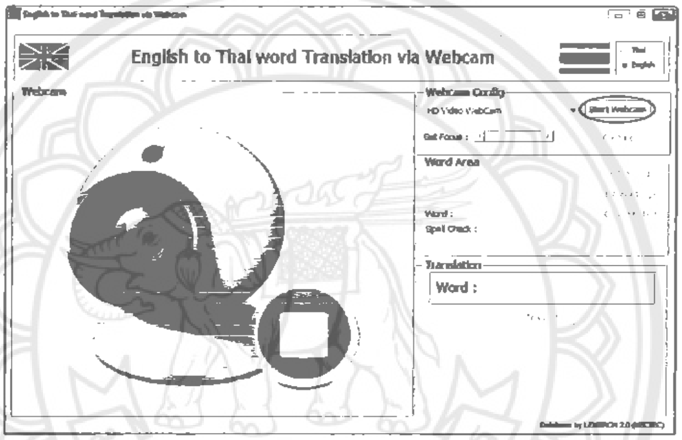
[8] "ระบบฐานข้อมูล." [online]. Available : Http://www.chandra.ac.th/office/ict/document/it/it04/page01.html. 2554.

[9] "ODBC." [online]. Available : Http://www.phpstreet.com/ tutorials/php/php_db_odbc.php. 2554.

ภาคผนวก ก.
การใช้งานโปรแกรม

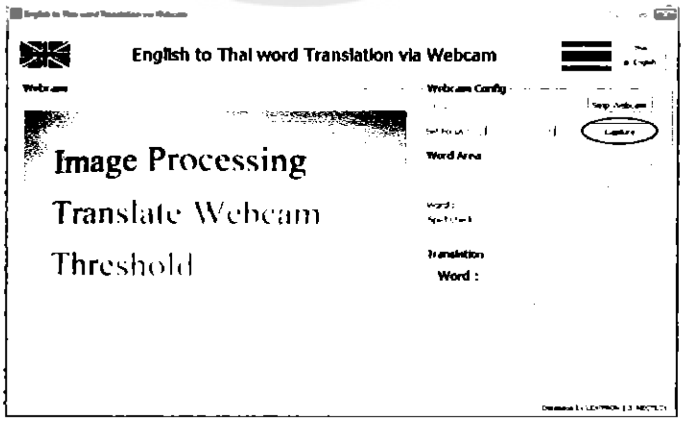
ขั้นตอนการใช้งานของโปรแกรม

1. คลิกที่ Start Webcam เพื่อทำการเริ่มเปิดกล้อง



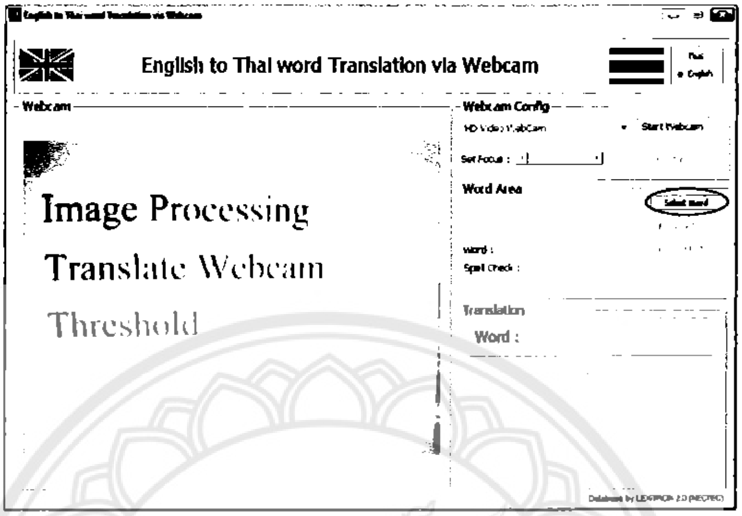
รูปที่ ก-1 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม

2. คลิกที่ Capture เพื่อทำการจับภาพ



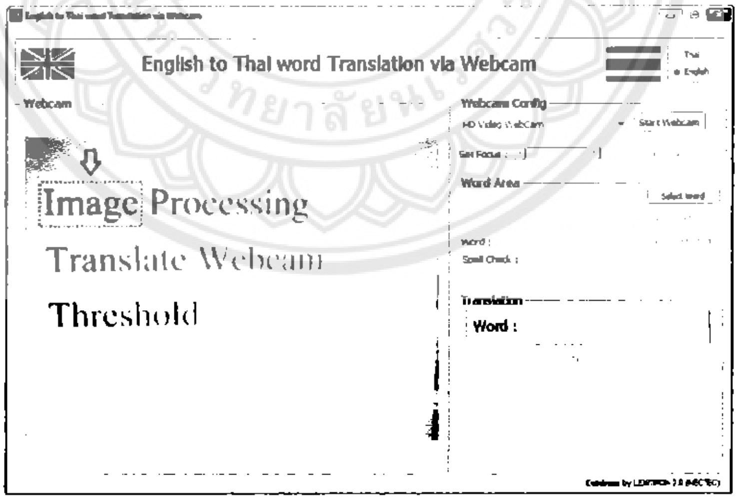
รูปที่ ก-2 ขั้นตอนการทำการจับภาพ

3. คลิกที่ Select Word เพื่อเลือกตัดเฉพาะคำที่ต้องการ



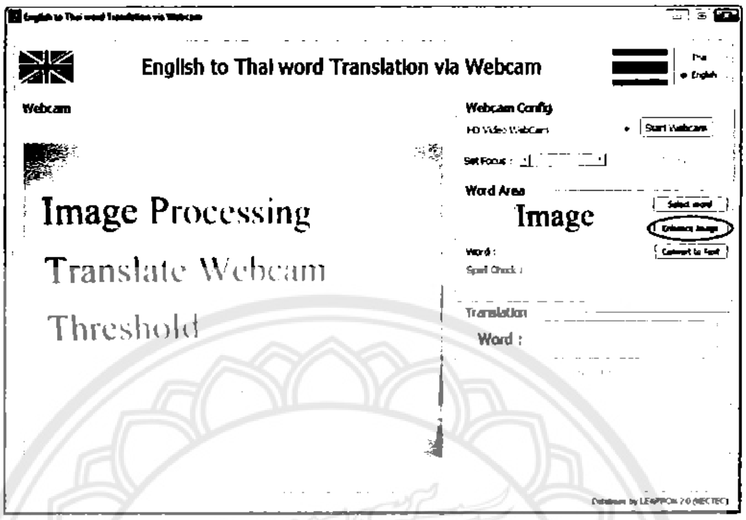
รูปที่ ก-3 ขั้นตอนการคลิก Select Word

4. ทำการลากตัดเฉพาะคำที่ต้องการ



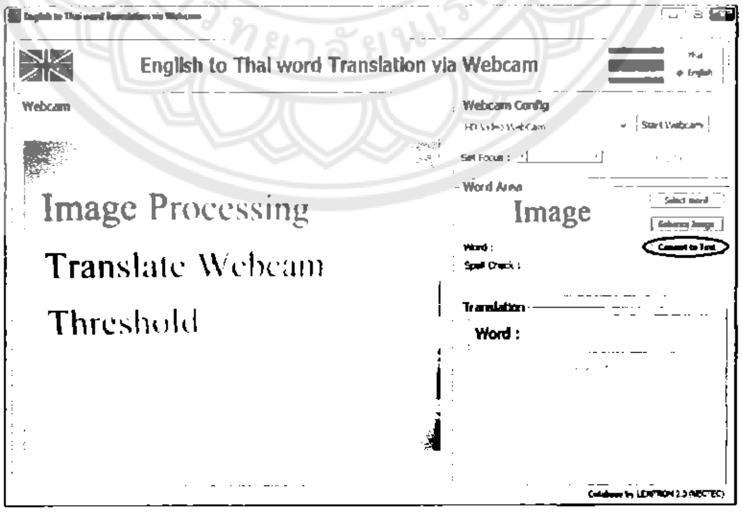
รูปที่ ก-4 ขั้นตอนเลือกคำที่ต้องการ

5. คลิก Enhance Image เพื่อทำให้คุณภาพของภาพดีขึ้น



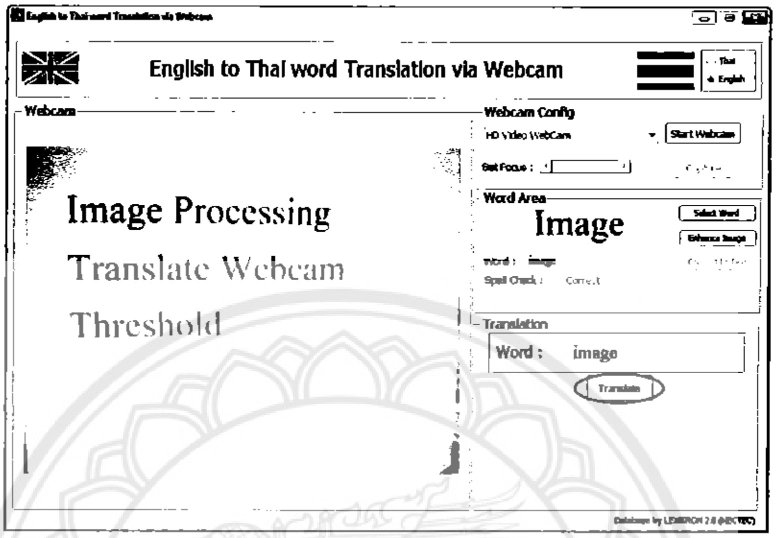
รูปที่ ก-5 ขั้นตอนการทำให้คุณภาพของภาพดีขึ้น

6. คลิกที่ Convert to Text เพื่อทำการแปลงภาพเป็นตัวอักษร



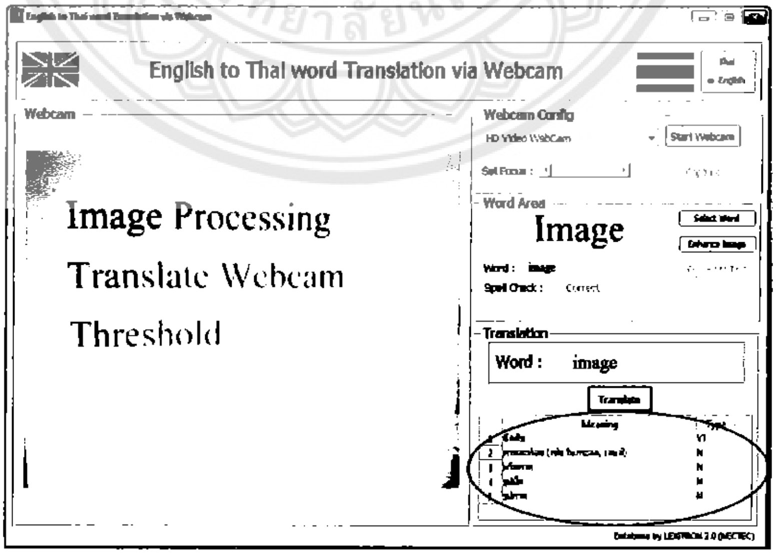
รูปที่ ก-6 ขั้นตอนการคลิก Convert to Text

7. เมื่อได้คำที่ผ่านการแปลงที่ถูกต้องแล้ว ให้ทำการคลิกที่ Translate เพื่อหาความหมาย



รูปที่ ก-7 ขั้นตอนการคลิก Translate เพื่อหาความหมาย

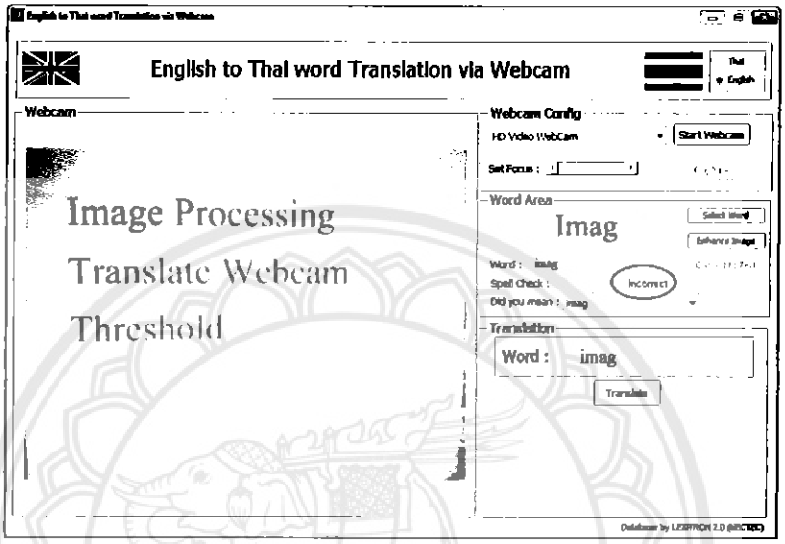
8. แสดงผลลัพธ์ของการแสดงความหมาย



รูปที่ ก-8 แสดงผลลัพธ์ของความหมาย

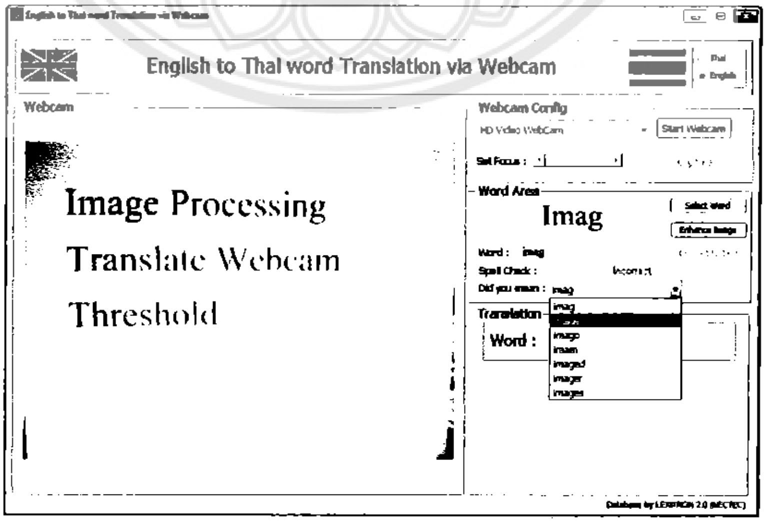
ขั้นตอนการใช้โปรแกรมในกรณีที่การแปลงภาพเป็นตัวอักษรผิดพลาด

1. ถ้าแถบ Spell Check เป็น Incorrect แสดงว่าการแปลงภาพเป็นตัวอักษรผิดพลาด



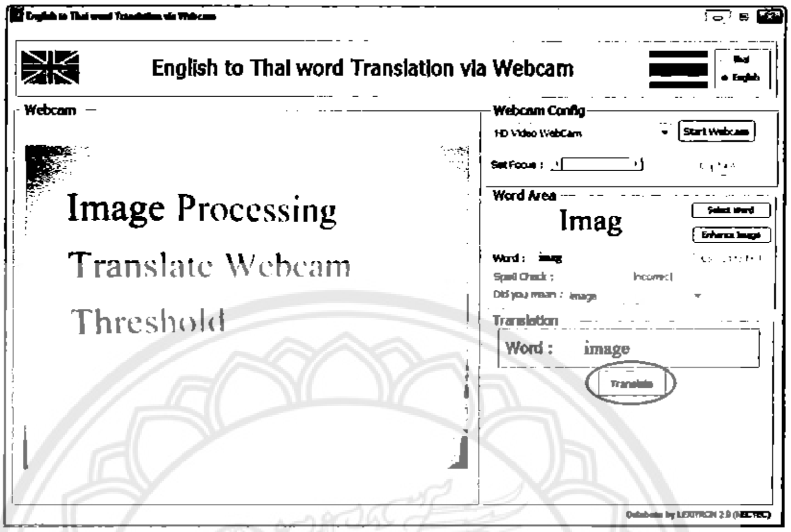
รูปที่ ก-9 แสดงการเตือนเมื่อการแปลงภาพเป็นตัวอักษรเกิดความผิดพลาด

2. ทำการคลิกเลือกพ็อพอัพเมนู (pop-up menu) เพื่อทำการเลือกคำที่ต้องการจะค้นหา



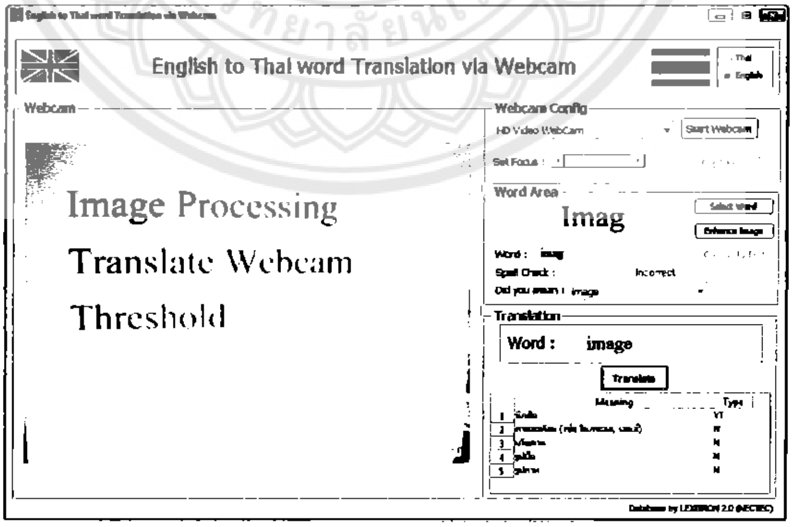
รูปที่ ก-10 ขั้นตอนการเริ่มการทำงานของกล้องเว็บแคม

3. คลิก Translate เพื่อทำการแปลความหมายของคำศัพท์



รูปที่ ก-11 ขั้นตอนการคลิกปุ่ม Translate เพื่อแปลความหมาย

4. ผลลัพธ์การแสดงความหมายของคำศัพท์



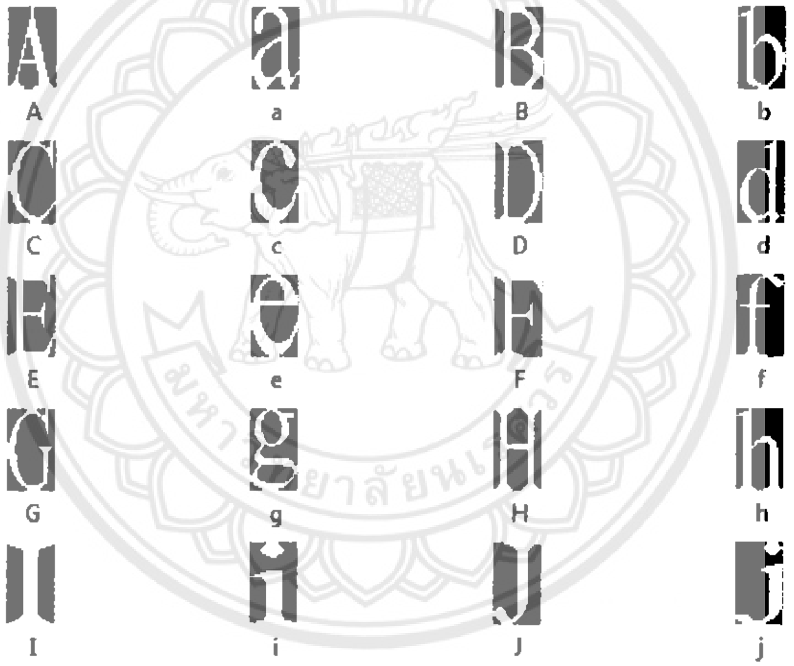
รูปที่ ก-12 แสดงความหมายของคำศัพท์

ภาคผนวก ข.

ภาพรูปแบบ (Templates) ในโปรแกรม

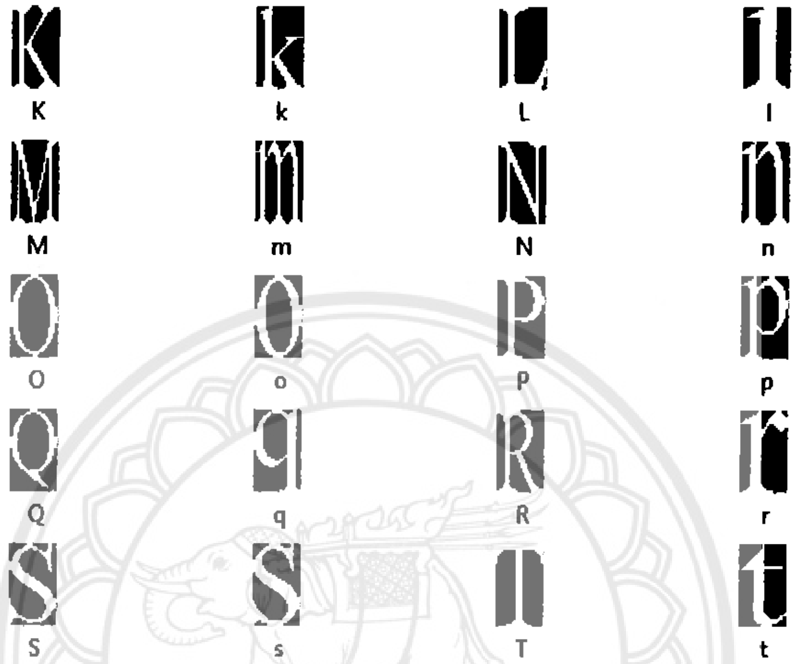
การแสดงภาพรูปแบบที่เก็บไว้ภายในโปรแกรม

1. ภาพรูปแบบทั้งตัวอักษรใหญ่และเล็ก ตั้งแต่ตัว A ไปจนถึง J



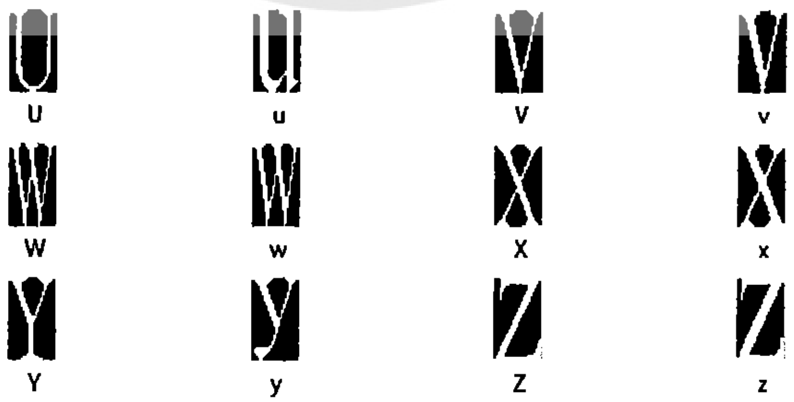
รูปที่ ข-1 แสดงตัวอักษรรูปแบบ A ไปจนถึง J

2. ภาพรูปแบบทั้งตัวอักษรใหญ่และเล็ก ตั้งแต่ตัว K ไปจนถึง T



รูปที่ ข-2 แสดงตัวอักษรรูปแบบ K ไปจนถึง T

3. ภาพรูปแบบทั้งตัวอักษรใหญ่และเล็ก ตั้งแต่ตัว U ไปจนถึง Z



รูปที่ ข-3 แสดงตัวอักษรรูปแบบ K ไปจนถึง T