

การจัดการแบบฟอร์มเพื่อการรู้จำตัวอักษรภาษาไทย

FORMS MANAGEMENT FOR THAI CHARACTER RECOGNITION

สุปัญญา อภิวังศ์โสภณ

SUPANYA APHIWONGSOPHON

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2547

ISBN 974-9700-58-9

การจัดการแบบฟอร์มเพื่อการรู้จำตัวอักษรภาษาไทย

FORMS MANAGEMENT FOR THAI CHARACTER RECOGNITION

สุปัญญา อภิวังศ์โสภณ

SUPANYA APHIWONGSOPHON

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2547

ISBN 974-9700-58-9

FORMS MANAGEMENT FOR THAI CHARACTER RECOGNITION

SUPANYA APHIWONGSOPHON

A THESIS SUBMITTED IN PARTIAL FUFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2004

ISBN 974-9700-58-9

COPYRIGHT 2004

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

หัวข้อวิทยานิพนธ์	การจัดการแบบฟอร์มเพื่อการรู้จำตัวอักษรภาษาไทย
ชื่อนักศึกษา	นางสาว สุปัญญา อภิวังศ์โสภณ
รหัสประจำตัว	42067031
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
พ.ศ.	2547
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ. ดร. วิเชียร เปรมชัยสวัสดิ์
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	รศ. ดร. นุชรี เปรมชัยสวัสดิ์

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้เสนอวิธีการจัดการแบบฟอร์ม เพื่อใช้สำหรับการรู้จำตัวอักษรภาษาไทย โดยสามารถนำรูปภาพแบบฟอร์มที่มีการกรอกข้อมูลแล้วมากำหนดตำแหน่งที่ต้องการรู้จำได้โดยอัตโนมัติ (Automatic Mark) หรือผู้ใช้สามารถกำหนดตำแหน่งที่ต้องการรู้จำได้เอง และผู้ใช้สามารถกำหนดรายละเอียดของฟิลด์ในแบบฟอร์มเปล่า (ชื่อฟิลด์ ประเภทข้อมูล หรือขอบเขตของข้อมูล) สำหรับแต่ละตำแหน่งที่มีการกรอกข้อมูล เพื่อสร้างฐานข้อมูลของแบบฟอร์มเปล่า ตำแหน่งข้อมูลที่ถูกกำหนดโดยอัตโนมัติจะถูกนำไปใช้ในกระบวนการรู้จำตัวอักษร โดยผลที่ได้จากการรู้จำตัวอักษรซึ่งประมวลผลจากโปรแกรมอื่น จะถูกนำมาตรวจสอบความถูกต้องกับประเภทข้อมูลที่กำหนดไว้ก่อนที่จะเก็บข้อมูลลงในฐานข้อมูล

Thesis Title	Forms Management for Thai Character Recognition
Student	Miss Supanya Aphiwongsophon
Student ID.	42067031
Degree	Master of Science
Programme	Information Technology
Year	2004
Thesis Advisor	Assoc.Prof.Dr. Wichian Premchaiswadi
Thesis Co-Advisor	Assoc.Prof.Dr. Nucharee Premchaiswadi

ABSTRACT

This thesis presents Forms Management technique for Thai character recognition. The Recognition areas can be marked automatically by the system or by users. User can define description of fields such as field-name field-data type that to be record in the database. The characters in marked area are passed to the recognition part that process by another program. Then, the result of the recognition part is verified and store in the database for the later used.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สามารถสำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำ และคำปรึกษาจาก รศ. ดร. วิเชียร เปรมชัยสวัสดิ์ ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ และ รศ. ดร. นุชรี เปรมชัยสวัสดิ์ อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ผู้วิจัยรู้สึกซาบซึ้งในความอนุเคราะห์เป็นอย่างมากและขอกราบขอบพระคุณท่านเป็นอย่างสูง

ขอกราบขอบพระคุณ บิดา มารดาของข้าพเจ้าที่คอยให้กำลังใจ ให้ความห่วงใย และให้คำปรึกษาที่ดีตลอดมา และยังสนับสนุนทุนทรัพย์ในการศึกษามาโดยตลอด และขอบคุณน้อง ๆ ทุกคนที่คอยเป็นกำลังใจ และช่วยกันหาแบบฟอร์มต่าง ๆ และกรอกข้อความลงในแบบฟอร์ม

ขอกราบขอบคุณ คุณครู อาจารย์ทุกท่านที่เคยสอนสั่งและให้ความรู้แก่ข้าพเจ้า ตั้งแต่ข้าพเจ้าเริ่มต้นศึกษาเล่าเรียนในสมัยเด็กจนกระทั่งปัจจุบัน

ขอขอบคุณอาจารย์ประจำคณะเทคโนโลยีสารสนเทศทุกท่าน ที่ได้สั่งสอนวิชาพื้นฐานต่าง ๆ ซึ่งเป็นการวางรากฐานจนมาสู่ความสำเร็จของงานวิจัยนี้ รวมทั้งขอขอบคุณคณะเทคโนโลยีสารสนเทศที่ให้การสนับสนุนเครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง รวมทั้งเจ้าหน้าที่ซึ่งคอยอำนวยความสะดวกในด้านต่าง ๆ ในระหว่างการศึกษาเป็นอย่างดี

ขอบคุณเพื่อน ๆ พี่ ๆ ทุกคนที่คอยเป็นกำลังใจเสมอมา

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

น.ศ. สุปัญญา อภิวงศ์โสภณ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญรูป	VII
บทที่ 1 ความเป็นมา	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	2
1.3 สมมติฐานของการศึกษา	2
1.4 ขอบเขตการวิจัย	2
1.5 ขั้นตอนการศึกษา	3
1.6 โครงสร้างวิทยานิพนธ์	3
1.7 ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 ทฤษฎีและแนวคิดที่เกี่ยวข้อง	5
2.1 การกำหนดบรรทัดข้อความในภาพเอกสาร	5
2.2 การลบเส้นบรรทัดในเอกสาร	10
2.2.1 LWTA (Line Width Thresholding Approach)	11
2.2.2 LWTA with heuristic a-priori information	12
2.2.3 LWTA และการพิจารณาผลลัพธ์จากการรู้	13
บทที่ 3 แนวทางที่ใช้ในงานวิจัย	15
3.1 การกำหนดกรอบโดยอัตโนมัติ (Automatic Mark)	16
3.2 การกำหนดประเภทของข้อมูลโดยผู้ใช้	24
3.3 การโหลดกรอบโดยอัตโนมัติ	28
3.4 การตรวจสอบความถูกต้องของข้อความ	31
3.5 การเก็บข้อมูลที่สามารถรู้จำได้ลงในฐานข้อมูลที่ถูกต้อง	31

สารบัญ(ต่อ)

	หน้า
บทที่ 4 ผลการทดลอง	32
4.1 ผลการทดลอง	32
บทที่ 5 สรุปผลการทดลองและข้อเสนอแนะ	54
5.1 สรุปผลการทดลอง	54
5.2 ข้อเสนอแนะและแนวทางพัฒนาต่อในอนาคต	55
เอกสารอ้างอิง	56
ภาคผนวก บทความและผลงานวิจัยที่ได้รับการตีพิมพ์	57
ประวัติผู้เขียน	65

สารบัญตาราง

ตารางที่	หน้า
4.1 เวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม	41

สารบัญรูป

รูปที่	หน้า	
2.1	ทิศทางการพิจารณาจุดสีดำในแบบฟอร์ม	5
2.2	แสดงภาพแบบฟอร์มที่ต้องการกำหนดบรรทัดโดยการหาค่าฮิสโตแกรมตามแนวนอน.....	7
2.3	ภาพแบบฟอร์มที่กำหนดบรรทัดโดยคำนวณจากฮิสโตแกรมและสมการที่ 2.1 และสมการที่ 2.2	8
2.4	ภาพแบบฟอร์มที่มีเส้นที่ใช้ค้นข้อความและกราฟความถี่ของฮิสโตแกรมแต่ละบรรทัด	9
2.5	ประเภทของเส้นที่ซ้อนทับกับตัวอักษรทั้ง 3 ประเภท	10
2.6	เส้นแนวนอนใน Array ของพิกเซลสีดำที่ยาวติดต่อกัน	11
2.7	ผลลัพธ์ของการลบเส้นด้วย LWTA โดยการใช้ค่าคงที่ต่างกัน	12
2.8	การเปรียบเทียบภาพที่เกิดจากการลบเส้นด้วยวิธี Basic LWTA กับวิธี LWTA with heuristic a-priori information.....	13
2.9	ผลลัพธ์ที่เกิดจากการลบเส้นด้วยวิธี LWTA with heuristic a-priori information และความน่าจะเป็นที่สามารถคำนวณได้ในแต่ละภาพ	14
3.1	ขั้นตอนการทำงานในระบบการจัดการแบบฟอร์มสำหรับการรู้จำตัวอักษรภาษาไทย	15
3.2	แบบฟอร์มเปล่าพร้อมทั้งภาพความถี่จุดสีดำในแนวนอน (Horizontal projection)	16
3.3	ผลลัพธ์ของการแยกบรรทัดของตัวอักษรในแบบฟอร์มเปล่า	17
3.4	ส่วนหนึ่งของแบบฟอร์มที่มีเส้นลักษณะที่เป็นจุดประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	18
3.5	ส่วนหนึ่งของแบบฟอร์มที่มีเส้นประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	18
3.6	ส่วนหนึ่งของแบบฟอร์มที่มีเส้นที่ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	18
3.7	ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.4	19
3.8	ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.5	19
3.9	ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.6	19
3.10	ความถี่ของจุดสีดำในแนวตั้งและแนวนอนของเครื่องหมาย \bigcirc , \square และพิกัดจุดต่าง ๆ ที่กำหนดเป็นกรอบข้อมูล	20
3.11	ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายวงกลม \bigcirc ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	21
3.12	ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายสี่เหลี่ยม \square ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	21

สารบัญรูป(ต่อ)

รูปที่	หน้า
3.13 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายวงเล็บ () ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	22
3.14 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมาย [] ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง	22
3.15 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.11	22
3.16 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.12	22
3.17 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.13	22
3.18 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความของรูปที่ 3.14	23
3.19 แบบฟอร์มเปล่าพร้อมด้วยกรอบที่ถูกกำหนดขึ้นมาโดยอัตโนมัติ	24
3.20 การเปลี่ยนชื่อฟิลด์โดยผู้ใช้	25
3.21 การกำหนดประเภทของข้อมูลการกำหนดขอบเขตข้อมูลโดยผู้ใช้	26
3.22 การแก้ไขประเภทข้อมูลเชิงตรรกศาสตร์โดยผู้ใช้	27
3.23 การเปลี่ยนประเภทข้อมูลเป็นตัวเลข โดยผู้ใช้	28
3.24 การกำหนดขอบเขตของข้อมูลอายุที่สนใจโดยผู้ใช้	28
3.25 ส่วนหนึ่งของแบบฟอร์มที่กรอกข้อความและได้โหลดกรอบโดยอัตโนมัติ	29
3.26 แสดงแบบฟอร์มที่กรอกข้อความแล้วพร้อมทั้งการโหลดกรอบโดยอัตโนมัติ	30
4.1 แบบฟอร์มเปล่าที่มีการกำหนดกรอบที่จะกรอกข้อมูลโดยอัตโนมัติ	33
4.2 การกำหนดชื่อฟิลด์ของข้อมูลลงในฐานข้อมูล	33
4.3 การเปลี่ยนแปลงประเภทข้อมูลเป็น Date ก่อนบันทึกลงในฐานข้อมูล	34
4.4 การเปลี่ยนแปลงประเภทข้อมูลในช่องที่สอง ก่อนบันทึกลงในฐานข้อมูล	34
4.5 การเปลี่ยนแปลงประเภทข้อมูลเป็น Month ก่อนบันทึกลงในฐานข้อมูล	35
4.6 การเปลี่ยนแปลงประเภทข้อมูลในช่องที่สาม ก่อนบันทึกลงในฐานข้อมูล	35
4.7 การเปลี่ยนแปลงประเภทข้อมูลเป็น Year ก่อนบันทึกลงในฐานข้อมูล	36
4.8 การเปลี่ยนแปลงประเภทข้อมูลในช่องที่สี่ ก่อนบันทึกลงในฐานข้อมูล	36
4.9 การเปลี่ยนชื่อฟิลด์แต่คงกำหนดประเภทข้อมูลเป็น Text ก่อนบันทึกลงในฐานข้อมูล	37
4.10 การกำหนดประเภทข้อมูลเป็น Boolean (Yes/No) ก่อนบันทึกลงในฐานข้อมูล	37
4.11 แบบฟอร์มเปล่าที่กำหนดกรอบที่จะบันทึกข้อมูลโดยอัตโนมัติเสร็จสิ้นทั้งหมด	38
4.12 ภาพแบบฟอร์มที่มีการกรอกข้อมูลกับแบบฟอร์มเปล่าที่มีความคล้ายคลึงกันมากที่สุด	39

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.13 ภาพแบบฟอร์มที่กรอกข้อมูลเรียบร้อยแล้วโหลดกรอบข้อมูลขึ้นมาโดยอัตโนมัติ	39
4.14 ภาพเต็มทั้งแบบฟอร์มที่กรอกข้อมูลเรียบร้อยแล้วโหลดกรอบข้อมูลขึ้นมาโดยอัตโนมัติ ..	40
4.15 ภาพกรอบข้อความที่ตัดจากบางส่วนของแบบฟอร์มที่กรอกข้อมูลแล้ว	52
4.16 ภาพกรอบข้อความของตัวอักษรที่ตัดเส้นบรรทัดออก	52

บทที่ 1

ความเป็นมา

เอกสารแบบฟอร์มต่าง ๆ ที่ใช้กันในปัจจุบัน มักอยู่บนกระดาษ (Paper Document) ซึ่งมีการบันทึกข้อความ ข้อมูลต่าง ๆ ลงบนกระดาษเหล่านั้น โดยการเขียน ซึ่งหากต้องการนำข้อมูลมาเก็บเป็นรูปแบบอิเล็กทรอนิกส์ (Electronic Document) เพื่อให้สามารถนำข้อมูล (Data) มาประมวลผลเป็นสารสนเทศ (Information) ที่สามารถนำไปใช้ประโยชน์ในงานต่าง ๆ ได้ อาทิเช่น งานวิจัยทางวิทยาศาสตร์และสังคมศาสตร์ งานเอกสารทางธุรกิจ โดยกระบวนการกรอกข้อมูลจากกระดาษลงในเครื่องคอมพิวเตอร์จะต้องอาศัยมนุษย์เป็นผู้กระทำ

ในกระบวนการกรอกข้อมูลจากกระดาษลงในเครื่องคอมพิวเตอร์ อาจเกิดความผิดพลาดในระหว่างการกรอกข้อมูล (Human Typing Error) เช่น การพิมพ์ผิด พิมพ์ขาดท่อน พิมพ์เกิน ซึ่งทำให้ข้อมูลที่ได้อาจไม่ตรงกับข้อเท็จจริง เกิดความผิดพลาดเมื่อมีการนำข้อมูลเหล่านั้นไปใช้งาน

1.1 ความเป็นมาและความสำคัญของปัญหา

เอกสารที่มีการใช้งานมีลักษณะและรูปแบบที่แตกต่างกัน เอกสารที่ใช้ในงานแต่ละประเภทย่อมมีลักษณะเฉพาะตัว ซึ่งการนำข้อมูลที่มีอยู่ในเอกสารมาใช้งานก็มีความแตกต่างกันตามรูปแบบข้อมูล อาทิเช่น ใบสมัครงานประกอบด้วย ชื่อ-นามสกุล ที่อยู่ บ้านเลขที่ ถนน ตำบล อำเภอ จังหวัด วุฒิการศึกษา ตำแหน่งงานที่ต้องการสมัคร จำนวนเงินเดือนที่ต้องการได้รับ ความสามารถพิเศษ หรือใบประกาศผลการศึกษาที่ประกอบด้วย ชื่อวิชาที่มีการเรียนการสอน รหัสนักศึกษาที่ลงทะเบียนเรียนวิชานั้น เกรดของนักศึกษา เป็นต้น ซึ่งหากมีการแยกประเภทข้อมูลเหล่านั้นได้อย่างเหมาะสม จะสามารถนำไปใช้งานได้โดยมีประสิทธิภาพ

งานวิจัยของกฤษฎา วิไลลักษณ์^[2] มีการกล่าวถึงการกำหนดกรอบข้อมูลเพื่อการรู้จำตัวอักษรภาษาไทย แต่เป็นการกำหนดโดยผู้ใช้งาน เพื่อให้สามารถนำข้อมูลที่อยู่ในกรอบนั้นมาผ่านกระบวนการรู้จำตัวอักษร ซึ่งในการกำหนดกรอบข้อมูลแต่ละตำแหน่งนั้นผู้ใช้งานจะต้องเสียเวลา กำหนดกรอบขอบบริเวณข้อมูลที่ต้องการ

วิทยานิพนธ์นี้ต้องการลดเวลาที่ใช้ในการกำหนดขอบเขตข้อมูลที่จะทำการรู้จำ จึงพัฒนาระบบที่สามารถกำหนดกรอบข้อมูลเพื่อการรู้จำตัวอักษรได้โดยอัตโนมัติ ทำให้สามารถทำงานได้สะดวกรวดเร็วยิ่งขึ้น ผู้ใช้งานสามารถเลือกที่จะบันทึกกรอบข้อมูลที่กำหนดขึ้นโดยอัตโนมัติได้ตามความต้องการ และยังสามารถกำหนดชนิดของข้อมูลที่ต้องการจะบันทึกลงในฐานข้อมูล หลังจากผ่านกระบวนการรู้จำเรียบร้อยแล้ว เพื่อให้สามารถแบ่งประเภทของข้อมูลได้อย่างชัดเจน และยังเป็นการลดขอบเขตในการจำแนกตัวอักษรต่าง ๆ หากบริเวณใดที่มีการกำหนดกรอบบันทึก

ข้อมูลที่ผู้ใช้งานไม่ต้องการกำหนด ผู้ใช้งานสามารถยกเลิกการบันทึกกรอบข้อมูลเหล่านั้นได้ หลังจากที่ผ่านมาขั้นตอนการกำหนดกรอบบันทึกข้อมูลเรียบร้อยแล้ว กรอบข้อมูลเหล่านั้นจะถูกส่งไปผ่านกระบวนการรู้จำตัวอักษรต่อไป และเมื่อผ่านกระบวนการรู้จำตัวอักษรจาก โปรแกรมอื่นแล้วจะมีการตรวจสอบข้อมูลเพื่อความถูกต้องมากยิ่งขึ้น และส่งผ่านข้อมูลที่ผ่านมาการตรวจสอบเรียบร้อยแล้วไปยังฐานข้อมูลที่ทำกรกำหนดไว้แล้วในส่วนของการกำหนดกรอบข้อมูลโดยอัตโนมัติ

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

เพื่อลดเวลาที่ใช้ในการกำหนดกรอบข้อมูลที่จะส่งไปรู้จำ
 เพื่อศึกษาลักษณะต่าง ๆ ที่เป็นบริเวณที่มีการกรอกข้อมูลของแบบฟอร์ม
 เพื่อศึกษาลักษณะข้อมูลเฉพาะตัวที่กรอกลงในแบบฟอร์ม
 เพื่อพัฒนาระบบการจัดการแบบฟอร์ม โดยการหากรอบตำแหน่งข้อมูลที่ต้องการบันทึกโดยอัตโนมัติ

1.3 สมมติฐานของการศึกษา

กรอบตำแหน่งข้อมูลที่มีการกำหนดโดยอัตโนมัติมีส่วนช่วยลดเวลาที่ใช้ในการกำหนดขอบเขตของข้อมูล

การกำหนดชนิดของข้อมูลหลังจากที่มีการกำหนดกรอบข้อมูลมีส่วนช่วยให้สามารถจำแนกข้อมูลและสามารถรู้จำได้ดีขึ้น

การตรวจสอบข้อมูลหลังจากที่ผ่านมากระบวนการรู้จำแล้ว สามารถทำให้ผลการรู้จำตัวอักษรมีความถูกต้องมากยิ่งขึ้น

1.4 ขอบเขตการวิจัย

รูปภาพแบบฟอร์มที่นำมาใช้จะต้องผ่านการสแกน (Scan) โดยกำหนดความละเอียดในการสแกนเป็น 200 จุดต่อนิ้ว (Dot per inch: dpi) ในแบบระดับสีเทา (Gray scale) แล้วนำภาพมาผ่านกระบวนการแปลงภาพสีสองระดับ (Duo tone) ได้เป็นภาพสีขาวดำ (Binary Image) จากนั้นจึงเข้าสู่กระบวนการจับคู่แบบฟอร์มที่กรอกข้อมูล (Fill Form) กับแบบฟอร์มเปล่า (Empty Form) ซึ่งจะไม่กล่าวถึงรายละเอียดในที่นี้

เมื่อทำการจับคู่แบบฟอร์มได้แล้วจะทำการเรียกกรอบตำแหน่งข้อมูลที่ต้องการบันทึกโดยอัตโนมัติ (Automatic Mark) และกำหนดฐานข้อมูลที่จะทำการบันทึกข้อมูล

งานวิจัยนี้ทำหน้าที่จัดการแบบฟอร์มเท่านั้น เพื่อส่งบริเวณที่มีการกรอกข้อมูลให้โปรแกรมที่
ทำหน้าที่รู้จำตัวอักษรเขียนภาษาไทย ^[2]

1.5 ขั้นตอนการศึกษา

ผู้วิจัยมีขั้นตอนในการศึกษางานวิจัยดังต่อไปนี้

- ศึกษาค้นคว้างานวิจัยและเอกสารทางวิชาการที่เกี่ยวข้องกับการทำวิทยานิพนธ์ โดยเริ่มจากการศึกษาผลงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับการรู้จำตัวอักษร ในวิชาสัมมนา 1 และสัมมนา 2 ตามลำดับ
- เก็บรวบรวมข้อมูลแบบฟอร์มต่าง ๆ และลายมือเขียนในแบบฟอร์มเหล่านั้นที่เกิดจากหลายบุคคล เพื่อความหลากหลายของรูปแบบลายมือสำหรับการใช้ในการทดลอง
- ศึกษาลักษณะ โครงสร้างของบริเวณที่มีการกรอกข้อมูลในแบบฟอร์มแต่ละประเภท
- กำหนดสมมติฐานที่ใช้ในการระบุตำแหน่งที่มีการกรอกข้อมูลในแบบฟอร์มต่าง ๆ
- ผู้วิจัยเขียน โปรแกรมดำเนินการทดสอบตามสมมติฐานที่กำหนดไว้
- ผู้วิจัยประมวลผลการทดลองและสรุปผลการดำเนินงาน
- เขียนบทความทางวิชาการ และเนื้อหาวิทยานิพนธ์
- สรุปผลการทดลองและข้อเสนอแนะที่จะสามารถพัฒนาต่อได้ในอนาคต
- นำเสนอวิทยานิพนธ์

1.6 โครงสร้างวิทยานิพนธ์

วิทยานิพนธ์นี้ได้แบ่งเนื้อหาออกเป็น 5 ส่วนประกอบด้วย

บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของวิทยานิพนธ์ ความมุ่งหมายและวัตถุประสงค์ของการศึกษา สมมติฐานของการศึกษา ขอบเขตการวิจัย และขั้นตอนการศึกษา

บทที่ 2 อธิบายทฤษฎีและแนวคิดที่เกี่ยวข้องกับวิทยานิพนธ์ที่ผู้วิจัยได้ทำการศึกษาและนำมาประยุกต์ใช้งาน

บทที่ 3 จะกล่าวถึงทฤษฎีและหลักการต่าง ๆ ที่ใช้วิทยานิพนธ์นี้

บทที่ 4 ผลการทดลองจากการนำทฤษฎีที่ได้กล่าวไว้ในบทที่สองและบทที่สาม มาเขียนโปรแกรมเพื่อทดสอบสมมติฐานและการทำงาน

บทที่ 5 สรุปผลการดำเนินงานและข้อเสนอแนะที่จะพัฒนาต่อได้ในอนาคต

1.7 ประโยชน์ที่คาดว่าจะได้รับ

ระบบการจัดการแบบฟอร์ม โดยการหากรอบตำแหน่งข้อมูลที่ต้องการบันทึกโดยอัตโนมัติ เพื่อลดปัญหาความผิดพลาดของข้อมูลจากการพิมพ์โดยอาศัยคน

กรอบตำแหน่งข้อมูลที่มีการกำหนดโดยอัตโนมัติมีส่วนช่วยลดเวลาที่ใช้ในการกำหนดขอบเขตของข้อมูล

การกำหนดชนิดของข้อมูลหลังจากที่มีการกำหนดกรอบข้อมูลมีส่วนช่วยให้สามารถจำแนกข้อมูลและสามารถรู้จำได้ดียิ่งขึ้น

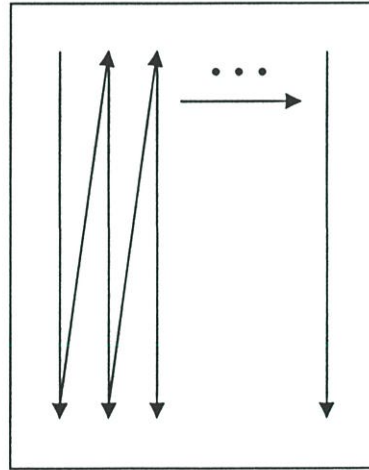
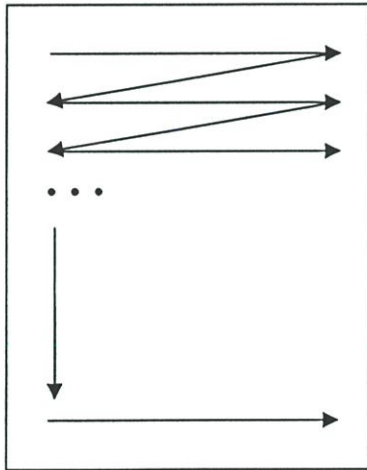
บทที่ 2

ทฤษฎีและแนวคิดที่เกี่ยวข้อง

ในบทนี้อธิบายถึงแนวคิด เทคนิค กระบวนการต่าง ๆ ที่เกี่ยวข้องกับการทำงานวิจัยนี้ และผู้วิจัยได้เลือกกระบวนการที่สำคัญซึ่งใช้ในการทำงานวิจัยมาแสดงอาทิเช่น การกำหนดบรรทัดข้อความในภาพเอกสารที่อาศัยหลักการพิจารณาค่าฮิสโตแกรม และการลบเส้นบรรทัดในภาพเอกสาร โดยวิธีการ Line Width Thresholding Approach [LWTA] ของ Y.Zheng และคณะ [5]

2.1 การกำหนดบรรทัดข้อความในภาพเอกสาร

วิธีการหาส่วนที่เป็นบรรทัดในภาพเอกสารแบบฟอร์ม โดยใช้วิธีการหาค่าฮิสโตแกรมตามแนวนอน (Horizontal Histogram) โดยการนับค่าจุดสีดำในภาพแบบฟอร์มตามแนวนอน ซึ่งพิจารณาจากทางซ้ายไปขวาแล้ววาดเส้นกราฟความถี่ ซึ่งทำตั้งแต่ต้นเอกสารจนหมดเอกสารจากบนลงล่างดังรูปที่ 2.1 (a) จากนั้นก็ทำการกำหนดค่าคงที่ (Threshold) ค่าหนึ่งเพื่อใช้ในการแบ่งบรรทัด



(a) ทิศทางการหาค่าฮิสโตแกรมตามแนวนอน (b) ทิศทางการหาค่าฮิสโตแกรมตามแนวตั้ง

รูปที่ 2.1 ทิศทางการพิจารณาค่าจุดสีดำในแบบฟอร์ม

หลังจากแบ่งบรรทัดได้ สามารถแบ่งหาเฉพาะส่วนที่เป็นเนื้อความในบรรทัดได้โดยการหาค่าฮิสโตแกรมในแนวตั้ง (Vertical Histogram) โดยทำการนับค่าจุดสีดำในแบบฟอร์มตามแนวตั้ง ซึ่งพิจารณาจากบนลงล่างแล้ววาดเส้นกราฟความถี่ โดยทำจากซ้ายไปขวาดังรูปที่ 2.1 (b) จากนั้นให้

กำหนดค่าคงที่ค่าหนึ่งเพื่อใช้แบ่งส่วนที่เป็นเนื้อความออกมาจากพื้นหลัง โดยสามารถคำนวณหาค่าฮิสโตแกรมในแนวนอนจากสมการที่ (2.1) และฮิสโตแกรมในแนวตั้ง ตามสมการที่ (2.2)

$$\text{Horizontal Histogram}(y) = \sum_{x=0}^{x \max} P(x, y) \quad (2.1)$$

$$\text{Vertical Histogram}(x) = \sum_{y=0}^{y \max} P(x, y) \quad (2.2)$$

$$\text{โดยที่ค่า } P(x, y) = \begin{cases} 0; & \text{white pixel} \\ 1; & \text{black pixel} \end{cases}$$

ตัวอย่างการกำหนดบรรทัดในภาพเอกสารแบบฟอร์มที่มีการกรอกข้อความแล้ว โดยการหาค่าฮิสโตแกรมตามแนวนอน (Horizontal Histogram) โดยการนับค่าจุดสีดำในภาพแบบฟอร์มตามแนวนอน และวาดเส้นกราฟความถี่ได้ดังรูปที่ 2.2

ใบสมัครขอรับทุนฝึกอบรมบุคลากรภาคอุตสาหกรรม

ข้าพเจ้า(นาย/นางนางสาว/.....)
 ชื่อ..... สุพินญา..... นามสกุล..... อินทร์ไผ่คง
 ตำแหน่ง..... ผู้ช่วยผู้จัดการ..... ระยะเวลาการปฏิบัติงาน..... 4 ปี
 วุฒิกการศึกษา..... ปริญญาตรี..... สาขา..... อหาม..... (วิทยาการคอมพิวเตอร์)
 สถาบัน..... มหาวิทยาลัยราชภัฏวชิรเวศน์
 บริษัท/หน่วยงาน.....
 ที่อยู่บริษัท/หน่วยงาน..... ๑7/๖ ม. ๗ อ.สุรนาคอ้ง อ.คลองกระเบื้อง อ.เมือง จ.บุรีรัมย์
 ผลิตภัณฑ์..... เหล็กหล่อ
 โทรศัพท์..... 0-๖๖๕1-๖๖40..... โทรสาร..... ๐-๖๖๕1-๖๖40..... e-mail..... supinya.o@icmail.com

มีความประสงค์จะสมัครเข้ารับทุนฯ (ผู้สมัคร 1 ท่านสามารถสมัครได้เพียง 1 หลักสูตร)
 หลักสูตรที่ 1 การทลอบขางและผลิตภัณฑ์
 หลักสูตรที่ 2 ความรู้เบื้องต้นจำหรับงานหล่อโลหะ
 และได้แนบเอกสารประกอบการสมัคร ดังนี้
 หลักฐานทางการศึกษา (Transcript)
 จดหมายรับรองการทำงานจากผู้บังคับบัญชา
 เอกสารแนบโครงการ (อยู่ด้านหลังใบสมัคร)

ทั้งนี้ข้าพเจ้าขอรับรองว่า มีคุณสมบัติตรงตามเกณฑ์การพิจารณาและยินดีปฏิบัติตามเงื่อนไขที่ระบุไว้

ลงชื่อผู้สมัคร..... สุพินญา
 (..... สุพินญา อินทร์ไผ่คง.....)
 วันที่..... /..... /.....

ส่งใบสมัครและเอกสารประกอบการสมัครได้ที่
 งานฝึกอบรม เอ็มเทค
 โทรสารหมายเลข 0 2564 6505

รูปที่ 2.2 แสดงภาพแบบฟอร์มที่ต้องการกำหนดบรรทัดโดยการหาค่าฮิสโตแกรมตามแนวนอน

เมื่อได้ค่าความถี่ของจุดสีดำในแนวนอนจะสามารถนำไปกำหนดค่าคงที่ค่าหนึ่งเพื่อทำการแบ่งบรรทัดได้ โดยอาศัยค่าเฉลี่ยระยะห่างระหว่างบรรทัด และนำค่าต่าง ๆ ที่ได้มาคำนวณตามสมการที่ 2.1 และ สมการที่ 2.2 เพื่อได้ส่วนที่เป็นเนื้อความออกมาจากพื้นหลัง ดังรูปที่ 2.3

ใบสมัครขอรับทุนฝึกอบรบบุคลากรภาคอุตสาหกรรม

ข้าพเจ้า(นาย/คุณ/นางสาว.....)

ชื่อ..... สุวิมลภา..... นามสกุล..... อึ้งวงศ์ไผ่ยศ.....

ตำแหน่ง..... ผู้ช่วยผู้จัดการ..... ระยะเวลาการปฏิบัติงาน..... 4.....ปี

วุฒิการศึกษา..... บัลชีวิทยาศาสร์..... สาขา..... อ.ท.ช. (วิชาทหารอุตสาหกรรม)

สถาบัน..... ม. วิชาญนิตวณิชพระเกษม.....

บริษัท/หน่วยงาน.....

ที่อยู่บริษัท/หน่วยงาน..... ๖7/๖ ม.17 อ.สุวิภาจวงษ์ ต.คลองขจรเหนือองศา อ.เมือง จ.ฉะเชิงเทรา

ผลิตภัณฑ์..... เหล็กถ่วง.....

โทรศัพท์..... 0-๖๖๕1-๗๖40..... โทรสาร..... 0-๖๖๕1-๗๖40..... e-mail..... supanya.a@hotmail.com

มีความประสงค์จะสมัครเข้ารับทุนฯ (ผู้สมัคร 1 ท่านสามารถสมัครได้เพียง 1 หลักสูตร)

หลักสูตรที่ 1 การทดสอบยางและผลิตภัณฑ์

หลักสูตรที่ 2 ความรู้เบื้องต้นสำหรับงานหล่อโลหะ

และได้แนบเอกสารประกอบการสมัคร ดังนี้

หลักฐานทางการศึกษา (Transcript)

จดหมายรับรองการทำงานจากผู้บังคับบัญชา

เอกสารแนบโครงการ (อยู่ด้านหลังใบสมัคร)

ทั้งนี้ข้าพเจ้าขอรับรองว่า มีคุณสมบัติตรงตามเกณฑ์การพิจารณาและยินดีปฏิบัติตามเงื่อนไขที่ระบุไว้

ลงชื่อผู้สมัคร..... สุวิมลภา.....

(..... สุวิมลภา อึ้งวงศ์ไผ่ยศ.....)

วันที่..... /..... /.....

ส่งใบสมัครและเอกสารประกอบการสมัครได้ที่

งานฝึกอบรม เริ่มเทศ

โทรสารหมายเลข 0 2564 6505

รูปที่ 2.3 ภาพแบบฟอร์มที่กำหนดบรรทัดโดยคำนวณจากฮิสโตแกรมและสมการที่ 2.1 และสมการที่ 2.2

ในกรณีที่มีเส้นในแบบฟอร์มเปล่า การพิจารณาเพื่อกำหนดให้เส้นนั้นเป็นส่วนที่ถูกกำหนดเป็นกรอบที่ใช้สำหรับกรอกข้อความในแบบฟอร์มหรือเป็นเส้นที่ใช้คั่นระหว่างข้อความส่วนหนึ่งออกจากข้อความอื่น จะพิจารณาจากระยะห่างเฉลี่ยระหว่างบรรทัดปกติในแบบฟอร์มนั้น ๆ หากเส้นนั้นมีระยะห่างระหว่างบรรทัดที่อยู่ข้างเคียงมากกว่าระยะห่างระหว่างบรรทัดเฉลี่ยในแบบฟอร์ม จะกำหนดให้เส้นนั้นเป็นเส้นที่ใช้คั่นข้อความ มิใช่เส้นบรรทัดที่มีการกรอกข้อมูล

ตัวอย่างภาพแบบฟอร์มที่ประกอบด้วยเส้นที่ใช้ในการคั่นข้อความที่อยู่กลางหน้าของแบบฟอร์มเพื่อแบ่งแยกข้อความส่วนบนกับส่วนล่างออกจากกัน พร้อมทั้งค่าความถี่ของจุดสีดำของข้อความในแต่ละบรรทัดในแบบฟอร์ม แสดงในรูปที่ 2.4

มหาวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
แบบฟอร์มคำร้องขอเปลี่ยนแปลงหัวข้อวิทยานิพนธ์
(กรณีไม่กระทบกระเทือนโครงสร้างวิทยานิพนธ์ที่ได้รับอนุมัติแล้ว)

วันที่ เดือน พ.ศ.

เรียน คณะบดีบัณฑิตวิทยาลัย

ข้าพเจ้า นาย นาง นางสาว อื่นๆ รับผิดชอบ

ตำแหน่ง สาขาวิชา

มีความประสงค์ขอเปลี่ยนแปลงหัวข้อวิทยานิพนธ์ จาก

เป็น

เนื่องจาก

จึงเรียนมาเพื่อโปรดพิจารณาดำเนินการต่อไปด้วย จักขอบคุณยิ่ง

ขอแสดงความนับถือ

ลงชื่อ

ที่อยู่ติดต่อได้

โทรศัพท์บ้าน โทรศัพท์ที่ทำงาน โฉนดถ้ำ

ผู้รับใช้

ความเห็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ผู้ควบคุมวิทยานิพนธ์ร่วม เดิม

ลงชื่อ

ความเห็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ผู้ควบคุมวิทยานิพนธ์ร่วม ใหม่

ลงชื่อ

ความเห็นของคณะกรรมการควบคุมวิทยานิพนธ์บัณฑิตวิทยาลัย

ลงชื่อ

คณะกรรมการควบคุมวิทยานิพนธ์บัณฑิตวิทยาลัย

รูปที่ 2.4 ภาพแบบฟอร์มที่มีเส้นที่ใช้คั่นข้อความและกราฟความถี่ของฮิสโตแกรมแต่ละบรรทัด

2.2 การลบเส้นบรรทัดในเอกสาร ^[5]

ในแบบฟอร์มปกติ มักมีตัวอักษรที่เหลื่อมล้ำกับเส้นบรรทัด วิธีการ Line Width Thresholding Approach [LWTA] ใช้ค่าคงที่ของความกว้างในเส้นบรรทัดในการลบเส้นออกไป

Line Width Thresholding Approach [LWTA] นำค่าคงที่ (Threshold) ที่มีค่าต่ำ ๆ สำหรับเส้น (Run-lengths) ที่อยู่ภายในตัวอักษร และค่าคงที่ (Threshold) ที่มากขึ้นสำหรับช่องระหว่างตัวอักษร การเหลื่อมล้ำระหว่างตัวอักษรกับเส้นแบ่งเป็นสามลักษณะคือ

1. เส้นสัมผัสกับตัวอักษร (Contact) ง่ายต่อการลบเส้นได้โดยไม่ส่งผลกระทบต่อตัวอักษรที่จะทำการรู้จำ
2. เส้นผ่ากลางตัวอักษร (Intersection) เส้นลักษณะนี้สามารถลบได้ยากมาก แต่ในการวิเคราะห์รูปภาพ ทำให้เราสามารถลบเส้น
3. เส้นทับบนส่วนหนึ่งของตัวอักษร (Superposition) สามารถลบได้ยากที่สุด แต่บางจุดที่อยู่ในบริเวณที่ซ้อนทับกันไม่สามารถที่จะแยกตัวอักษรออกจากขอบเขตเส้นหรือเส้นเชื่อมของตัวอักษรด้วยการวิเคราะห์ภาพโดยตรง ในบางครั้งข้อมูลก่อนหน้า (Prior Information) อาทิเช่น ข้อมูลไวยากรณ์และความหมายจึงต้องนำมาใช้ในกรณีนี้เพื่อทำให้ได้ข้อมูลที่ถูกต้อง



(a) contact

(b) intersection

(c) superposition

รูปที่ 2.5 ประเภทของเส้นที่ซ้อนทับกับตัวอักษรทั้ง 3 ประเภท

หลักการที่ใช้ในการลบเส้นออกจากตัวอักษร

1. เริ่มด้วยการลบเส้นเกือบทั้งหมดออกก่อน
2. ใช้คุณสมบัติของพื้นที่ซ้อนทับกัน เช่น ทิศทางของเส้นในการเขียนหนังสือที่เชื่อมต่ออยู่
3. เติมเต็มส่วนของเส้นในการเขียนหนังสือที่ขาดหายไป

หลักการที่กล่าวมานี้ใช้ได้กับเส้นสัมผัสกับตัวอักษร (Contact) และ เส้นผ่ากลางตัวอักษร (Intersection) แต่ไม่สามารถใช้ได้กับเส้นทับบนส่วนหนึ่งของตัวอักษร (Superposition) เนื่องจากหลังจากที่ทำการลบเส้นกรอบไปแล้วจะทำให้ข้อมูลที่จำเป็นหายไปด้วย โดยที่ไม่สามารถกู้ข้อมูลคืนกลับด้วยการใช้ทิศทางและจุดเชื่อมต่อของส่วนที่เหลือ

2.2.1 LWTA (Line Width Thresholding Approach)

เริ่มจากการสร้างกรอบของเส้นเก็บใน Array ของพิกเซลสีดำที่มีความยาวติดต่อกัน (Run-lengths) โดยกำหนดค่าคงที่ (Threshold) สูง ๆ สำหรับใช้ภายในตัวอักษรที่มีเส้นทับและกำหนดค่าคงที่ต่ำ ๆ สำหรับใช้ระหว่างตัวอักษรที่มีเส้นทับ

วิธีการนี้เป็นการกำหนดค่าคงที่ มีส่วนสำคัญมาก หากกำหนดค่าน้อยเกินไปก็จะทำให้ลบเส้นออกได้ไม่หมด แต่หากกำหนดค่ามากเกินไปก็จะทำให้ลบส่วนของตัวอักษรออกไปด้วย โดยปกติแล้วจะกำหนดค่าคงที่ (Threshold) ให้มากกว่าค่าเฉลี่ยของความกว้างของกรอบเส้น 2-3 พิกเซล แต่หากมีการเอียงหรือมีสัญญาณรบกวนจะเกิดความยาวมาก ๆ เรียกว่าความยาวที่ผิดปกติ (Abnormal Run-lengths)

หลักการของ LWTA ใช้เทคนิคการกรองค่ากลางเพื่อจะทำการแทรกความยาวที่ผิดปกติ โดยกำหนด นิยามของความยาวปกติคือ

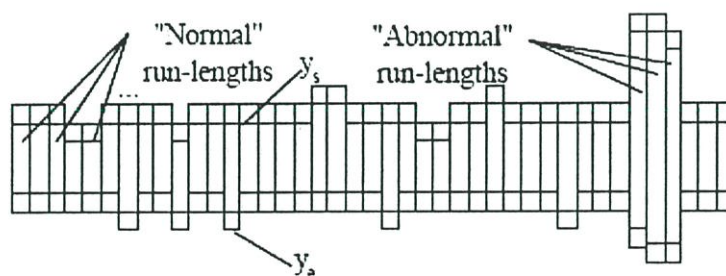
$$R_{norm} = \{R_i \mid y_{e_i} - y_{s_i} + 1 < 2l_{mid}\} \quad (2.3)$$

โดยที่

R_i = Array ของพิกเซลสีดำที่มีความยาวต่อเนื่องกัน

l_{mid} = ค่ากลางของความยาวทั้งหมด

ความกว้างของเส้นกรอบประมาณด้วยค่าเฉลี่ยของความยาวที่ต่อเนื่องแบบปกติทั้งหมด



รูปที่ 2.6 เส้นแนวนอนใน Array ของพิกเซลสีดำที่ยาวติดต่อกัน

เทคนิคที่ใช้ในการลบเส้นด้วยค่าคงที่สองค่า คือใช้ค่าคงที่สูงใช้กับส่วนที่อยู่ภายในตัวอักษรที่มีเส้นทับ และกำหนดค่าคงที่ต่ำ ๆ ใช้ระหว่างตัวอักษรที่มีเส้นทับ ซึ่งแสดงดังรูปที่ 2.7 (a), 2.7 (b), 2.7 (c), 2.7 (d)

¥416600.00

(a) ภาพต้นแบบของตัวอักษรซ้อนที่มีเส้นทับกับตัวอักษร

¥416600.00

(b) ผลลัพธ์การลบเส้นด้วยค่าคงที่ (Threshold) เพียงค่าเดียว

¥416600.00

(c) ผลลัพธ์ของการแบ่งตัวอักษร ซึ่งอยู่ในกรอบสี่เหลี่ยมกรอบตัวอักษร

¥416600.00

(d) ผลลัพธ์ที่เกิดจากการลบเส้นด้วยการใช้ค่าคงที่ (Threshold) 2 ค่า
รูปที่ 2.7 ผลลัพธ์ของการลบเส้นด้วย LWTA โดยใช้ค่าคงที่ต่างกัน

2.2.2 LWTA with heuristic a-priori information

การลบเส้นที่ติดกับตัวเลข โดยอาศัยข้อมูลส่วนก่อนหน้า ในการตัดสินใจที่จะลบ เนื่องจากวิธีการ Basic LWTA ที่กล่าวมาในข้างต้นให้ผลลัพธ์ผิดพลาด ในกรณีที่ลบเส้นส่วนล่างของตัวเลข 0, 2 และ 8 ทำให้ข้อมูลที่เหลืออยู่เกิดความคลาดเคลื่อน

ในกรณีที่เป็นตัวเลข 0 มีส่วนที่ขาดหายไปในส่วนบนและส่วนล่าง การรู้จำจะแสดงผลออกมาเป็นตัวเลข 11 [()]

หากตัวเลข 2 มีการลบส่วนเส้นในแนวนอนส่วนล่างทิ้งไป การรู้จำจะแสดงผลเป็นตัวเลข 7

ส่วนบนของตัวเลข 2 และ 7 มีความแตกต่างกันเพียงบางส่วนคือ ส่วนบนของเลข 2 จะเกิดจากการลากเส้นโค้งครึ่งวงกลมที่ราบรื่นกว่าเลข 7 ซึ่งเลข 7 เกิดจากการลากเส้นแนวนอนและตามด้วยเส้นในแนวตั้ง

หากเลขส่วนทางด้านขวาในจุดสิ้นสุดของเลข 2 ซ้อนทับบนเส้นกรอบและถูกลบไป จะใช้การพิจารณาในส่วนบนว่ามีส่วนของเส้นตรงในแนวนอนและแนวตั้งเพื่อบอกว่าเป็นเลข 7

หลังจากที่พบว่ามีการซ้อนทับของตัวอักษรบนเส้นให้ทำการรวมส่วนที่ขาดติดต่อกันไว้ถ้ามีค่าสั้นกว่าค่าคงที่แล้วให้ทำการลบ แต่ถ้ามีความยาวมากกว่าทั้ง 2 ด้านไม่ลบ ตัวอย่างการลบเส้นแสดงผลลัพธ์ดังรูป 2.8

¥70822.00

(a) ภาพต้นแบบของตัวเลขที่มีเส้นซ้อนทับอยู่

¥70822.00

(b) ผลลัพธ์ของการลบเส้นด้วยวิธีการ basic LWTA

¥70822.00

(c) ผลลัพธ์ของการลบเส้นด้วยวิธีการ LWTA with heuristic a-priori information

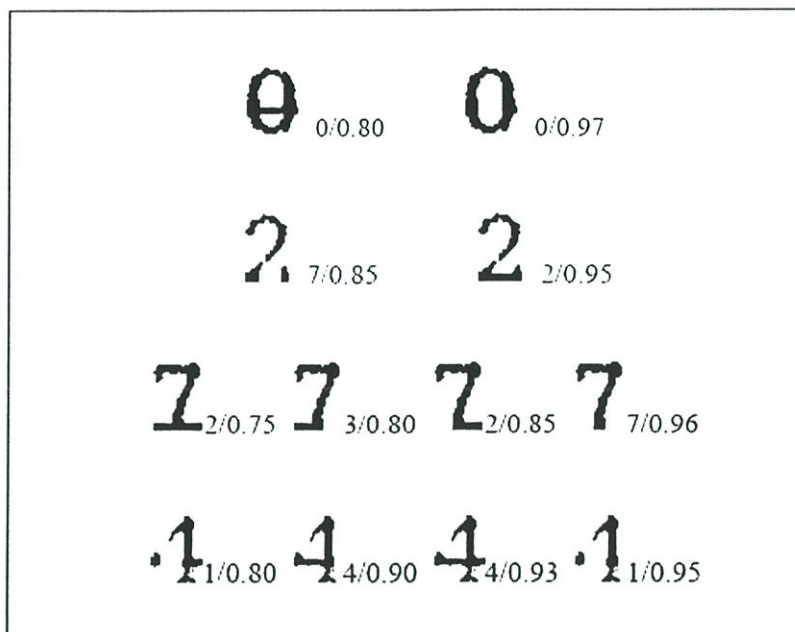
รูปที่ 2.8 การเปรียบเทียบภาพที่เกิดจากการลบเส้นด้วยวิธี Basic LWTA กับวิธี LWTA with heuristic a-priori information

หลักการการทำงานของ LWTA with heuristic a-priori information

1. สร้างฮิสโตแกรมทางแนวตั้ง
 2. หาจุดที่มีความถี่สูงกว่าบริเวณรอบข้างทั้งสองด้านของกรอบเส้น
 3. หากบริเวณที่มีความถี่สูงกว่าบริเวณรอบข้าง อยู่ใกล้กันมากกว่าส่วนอื่น แสดงว่าไม่เกิดการซ้อนทับสามารถลบเส้นกรอบออกได้
 4. แต่หากในบริเวณที่มีความถี่สูงกว่าบริเวณรอบข้างทั้ง 2 ด้าน อยู่ห่างไกลกันจากกรอบเส้น
 - 4.1 ถ้ามีส่วนที่ซ้อนทับกันแบบผ่ากลางตัวอักษร (Intersection) อยู่ ให้ทำ Basic LWTA
 - 4.2 ถ้ามีส่วนที่ทับบนเส้น (Superposition) ใช้การปรับปรุงวิธีการทำงาน
- สำหรับการซ้อนทับแบบเชื่อมต่อกันของเลข 1, 4, 7 และ 9 สามารถลบเส้นได้ แต่ถ้ามีส่วนที่ทับบนเส้น (Superposition) ของเลข 0, 2, 3, 5, 6 และ 8 ห้ามลบเส้นด้านล่าง

2.2.3 LWTA และการพิจารณาผลลัพธ์จากการรู้จำ

ใช้วิธีการลบเส้นแบบ LWTA with heuristic a-priori information และทำการรู้จำตัวอักษรของ Array ของภาพ ในแต่ละขั้นของการลบเส้นแบบต่าง ๆ พร้อมทั้งคิดค่าความน่าจะเป็นของผลลัพธ์ โดยเรียงลำดับจากน้อยไปมาก ให้ค่าที่มีค่าความน่าจะเป็นที่น่าเชื่อถือที่สุดเป็นรูปที่อยู่ทางขวาสุดตามตัวอย่างในรูปที่ 2.9



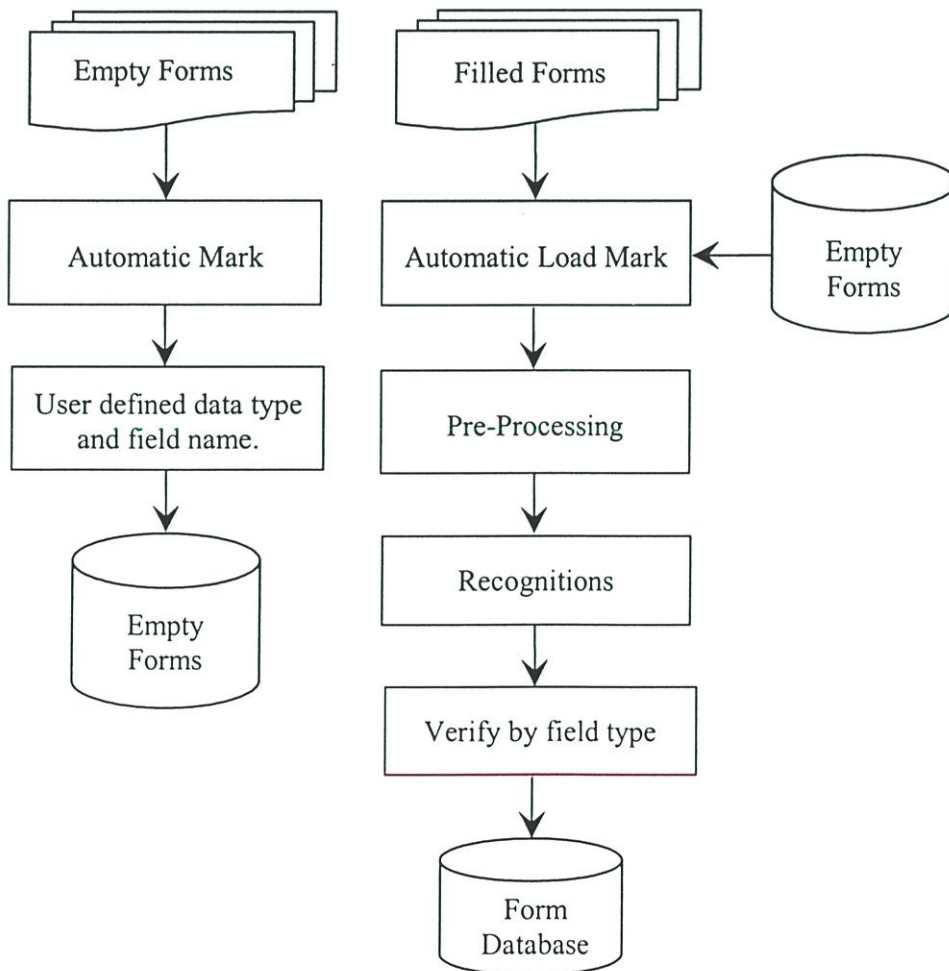
รูปที่ 2.9 ผลลัพธ์ที่เกิดจากการลบเส้นด้วยวิธี LWTA with heuristic a-priori information และความน่าจะเป็นที่สามารถคำนวณได้ในแต่ละภาพ

บทที่ 3

แนวทางที่ใช้ในการวิจัย

การจัดการแบบฟอร์ม เพื่อใช้สำหรับการรู้จำตัวอักษรภาษาไทย ประกอบด้วยสองกระบวนการทำงานหลักคือ ส่วนที่หนึ่ง การจัดการแบบฟอร์มเปล่าโดยการนำแบบฟอร์มต่าง ๆ ที่ยังมีได้มีการกรอกข้อมูลลงไป มากำหนดกรอบตำแหน่งข้อมูลโดยอัตโนมัติ และกำหนดรายละเอียดโดยผู้ใช้งาน ในส่วนที่สอง เป็นการจัดการแบบฟอร์มที่มีการกรอกข้อมูลแล้ว โดยการโหลดข้อมูลที่มีการกำหนดขอบเขตที่ต้องการรู้จำตัวอักษรภาษาไทยจากฐานข้อมูลแบบฟอร์มเปล่าที่เก็บไว้ ผ่านกระบวนการจัดการก่อนเข้าสู่กระบวนการรู้จำ เพื่อให้สามารถนำข้อมูลมาผ่านกระบวนการรู้จำตัวอักษรภาษาไทยซึ่งประมวลผลด้วยโปรแกรมอื่น และสามารถตรวจสอบความถูกต้องก่อนจัดเก็บข้อมูลลงในฐานข้อมูล

สามารถแสดงรายละเอียดของกระบวนการทั้งหมดในระบบดังแผนภาพในรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนการทำงานในระบบการจัดการแบบฟอร์มสำหรับการรู้จำตัวอักษรภาษาไทย

3.1 การกำหนดกรอบโดยอัตโนมัติ (Automatic Mark)

ในส่วนนี้อธิบายถึงขั้นตอนการกำหนดกรอบโดยอัตโนมัติที่สามารถหาตำแหน่งกรอบที่ต้องการบันทึกข้อมูลได้โดยอัตโนมัติ

ขั้นตอนแรก แยกบรรทัดในแบบฟอร์ม โดยการหาความถี่ของจุดในแนวนอน (Horizontal projection) เพื่อค้นหาตำแหน่งที่มีการกรอกข้อมูล

ตัวอย่างภาพแบบฟอร์มเปล่าพร้อมทั้งความถี่ของจุดในแนวนอน ที่ใช้ในการแยกบรรทัดของแบบฟอร์ม แสดงดังรูปที่ 3.2

แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน

วันที่ เดือน พ.ศ.

เรียน

ข้าพเจ้า ชื่อ นามสกุล

อยู่บ้านเลขที่ หมู่ที่ ตำบล/แขวง

อำเภอ/เขต จังหวัด

รหัสไปรษณีย์ โทรศัพท์

มีความประสงค์จะใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก
หน่วยงานของท่าน ดังต่อไปนี้

1.

2.

โดยต้องการ ขอตรวจดู ขอสำเนา

ขอสำเนาที่มีคำรับรองถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามที่กำหนดไว้ (ถ้ามี)

ลงชื่อ ผู้ยื่นคำขอ
(.....)

รูปที่ 3.2 แบบฟอร์มเปล่าพร้อมทั้งภาพความถี่จุดสีดำในแนวนอน (Horizontal projection)

จากการหาภาพความถี่จุดสีดำในแนวนอนของแบบฟอร์มในรูปตัวอย่างที่ 3.2 สามารถนำไปใช้ในการแยกบรรทัดของตัวอักษรที่ปรากฏอยู่ในแบบฟอร์ม ซึ่งได้ผลลัพธ์แสดงดังรูปที่ 3.3

แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน

วันที่ เดือน พ.ศ.

เรียน

ข้าพเจ้า ชื่อ นามสกุล

อยู่บ้านเลขที่ หมู่ที่ ตำบล/แขวง

อำเภอ/เขต จังหวัด

รหัสไปรษณีย์ โทรศัพท์

มีความประสงค์ใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก
หน่วยงานของท่าน ดังต่อไปนี้

1.

2.

โดยต้องการ ขอตรวจสอบ ขอสำเนา

ขอสำเนาที่มีคำรับรองถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามที่กำหนดไว้ (ถ้ามี)

ลงชื่อ ผู้ยื่นคำขอ

.....

รูปที่ 3.3 ผลลัพธ์ของการแยกบรรทัดของตัวอักษรในแบบฟอร์มเปล่า

ขั้นตอนที่สอง หาความถี่ของจุดในแนวตั้ง (Vertical projection) ของตัวอักษรในแต่ละบรรทัด เพื่อหาบริเวณที่ควรเป็นตำแหน่งสำหรับการกรอกข้อความ และสร้างเป็นกรอบที่ต้องการรู้จำ

การกำหนดตำแหน่งกรอบที่จะบันทึกข้อมูลโดยอัตโนมัติ กรณีที่แบบฟอร์มมีช่องว่างที่จะเติมข้อความที่รองรับข้อความด้วยเส้นบรรทัด จะกำหนดเป็นกรอบของเส้นตรง (Straight line mark) และประเภทที่มีการบันทึกเครื่องหมายลงในช่องว่างที่กำหนดขึ้นในแบบฟอร์ม จะกำหนดเป็นกรอบที่ใช้ในการตอบคำถามที่มีคำตอบสองลักษณะคือใช่/ไม่ (Optional mark)

กรอบของเส้นตรง (Straight line mark) ประกอบด้วยลักษณะดังนี้ จุด เส้นปะ เส้นทึบ สำหรับเติมข้อมูลประเภทตัวอักษรหรือตัวเลข ในกรณีนี้จะพิจารณาจากบริเวณที่มีอัตราส่วนระหว่างความกว้างต่อความสูงของเส้นมีค่าแตกต่างกันมาก (ความกว้างมีค่ามากกว่าความสูงมาก ๆ)

ตัวอย่างส่วนหนึ่งของภาพแบบฟอร์มที่ประกอบด้วยส่วนที่มีลักษณะเป็นจุด เส้นปะ เส้นทึบ พร้อมทั้งแสดงภาพความถี่ของจุดสีดำในแนวตั้งของแต่ละภาพตัวอย่าง แสดงดังรูปที่ 3.4, 3.5 และ 3.6 ตามลำดับ และผลที่ได้จากการกำหนดกรอบโดยอัตโนมัติจากรูปที่ 3.4, 3.5 และ 3.6 แสดงดังรูปที่ 3.7, 3.8 และ 3.9 ตามลำดับ

วันที่ เดือน พ.ศ.

จาก เหตุ พ.ศ.

รูปที่ 3.4 ส่วนหนึ่งของแบบฟอร์มที่มีเส้นลักษณะที่เป็นจุดประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง

วันที่ เดือน พ.ศ.

จาก เหตุ พ.ศ.

รูปที่ 3.5 ส่วนหนึ่งของแบบฟอร์มที่มีเส้นปะประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง

วันที่ เดือน พ.ศ.

จาก เหตุ พ.ศ.

รูปที่ 3.6 ส่วนหนึ่งของแบบฟอร์มที่มีเส้นทึบประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำในแนวตั้ง

วันที่.....	เดือน.....	พ.ศ.....
-------------	------------	----------

รูปที่ 3.7 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.4

วันที่.....	เดือน.....	พ.ศ.....
-------------	------------	----------

รูปที่ 3.8 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.5

วันที่.....	เดือน.....	พ.ศ.....
-------------	------------	----------

รูปที่ 3.9 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.6

กรอบที่ใช้ในการตอบคำถามที่มีคำตอบสองลักษณะคือใช่/ไม่ (Optional mark) ประกอบด้วย \bigcirc , \square , $()$ และ $[\]$ สำหรับเติมข้อมูลประเภทเครื่องหมายถูก/ผิด หรือไม่เติมเครื่องหมายใด ๆ

ลักษณะที่ใช้ในการพิจารณากรอบที่ใช้ในการตอบคำถามที่มีคำตอบสองลักษณะคือใช่/ไม่

1. ปรากฏสัญลักษณ์ \bigcirc , \square , $()$ หรือ $[\]$ ที่บริเวณต้นบรรทัด
2. มีพื้นที่ว่างระยะหนึ่งก่อนที่จะมีตัวอักษรเป็นคำอธิบายตามมา

คุณลักษณะของสัญลักษณ์ \bigcirc , \square ให้พิจารณาจากจุดสูงสุดของความถี่ของจุดสีดำในแนวตั้ง ซึ่งมีจุดสูงสุดของความถี่ของจุดสีดำ 2 จุดสูงเท่า ๆ กันบริเวณปลายทั้งสองข้าง

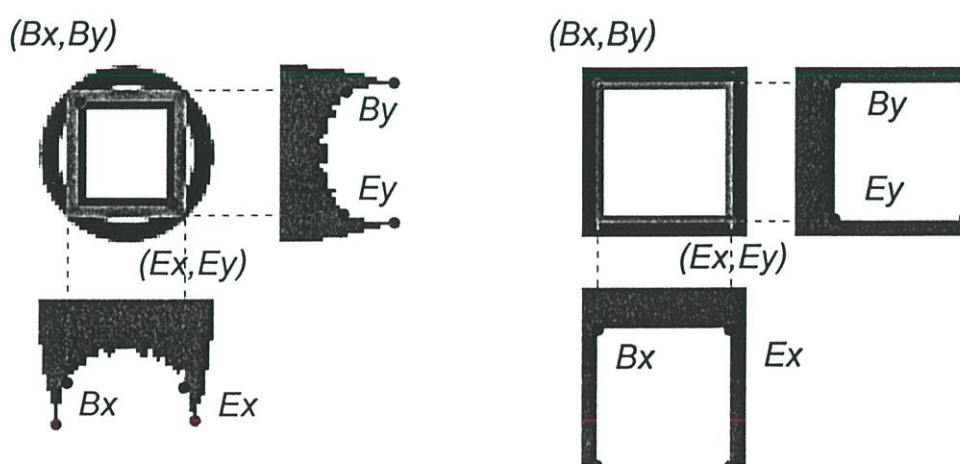
ความแตกต่างระหว่างกรอบสี่เหลี่ยมและกรอบวงกลมคือวงกลมจะมีลักษณะของความถี่ของจุดสีดำเป็นเส้นโค้ง แต่กรอบสี่เหลี่ยมจะมีลักษณะของความถี่ของจุดสีดำเป็นจุดสูงสุดที่ปลายทั้งสองด้านและทำมุมฉากกับความถี่ของจุดสีดำของเส้นที่อยู่ช่วงกลาง

เมื่อพบว่าบริเวณที่พิจารณามีสัญลักษณ์ \bigcirc , \square ปรากฏอยู่ จะพิจารณาบริเวณดังกล่าวตามขั้นตอนดังต่อไปนี้

1. หาค่าความถี่ของจุดสีดำในแนวนอน และหาจุดสูงสุดของค่าความถี่ (Peak) ที่พบบริเวณปลายทั้งสองด้าน
2. กำหนดจุดที่มีค่าความถี่สูงสุดจุดแรกที่พบเป็น B_y และกำหนดจุดที่มีค่าความถี่สูงสุดจุดสุดท้ายเป็น E_y

3. พิจารณาจุดที่ต่ำกว่าจุดสูงสุดที่พบในข้อ 2 ประมาณ 30% กำหนดให้จุดนั้นเป็นจุด B_y ใหม่ และพิจารณาจุดสุดท้ายที่มีค่าความถี่ต่ำกว่าค่าจุดความถี่สูงสุดสุดท้ายในข้อ 2 ประมาณ 30% กำหนดให้เป็นจุด E_y ใหม่
4. หาค่าความถี่ของจุดสีดำในแนวตั้ง และหาจุดสูงสุดของค่าความถี่ (Peak) ที่พบบริเวณปลายทั้งสองด้าน
5. กำหนดจุดที่มีค่าความถี่สูงสุดจุดแรกที่พบเป็น B_x และกำหนดจุดที่มีค่าความถี่สูงสุดจุดสุดท้ายเป็น E_x
6. พิจารณาจุดที่ต่ำกว่าจุดสูงสุดที่พบในข้อ 5 ประมาณ 30% กำหนดให้จุดนั้นเป็นจุด B_x ใหม่ และพิจารณาจุดสุดท้ายที่มีค่าความถี่ต่ำกว่าค่าจุดความถี่สูงสุดสุดท้ายในข้อ 5 ประมาณ 30% กำหนดให้เป็นจุด E_x ใหม่
7. เมื่อได้ค่าพิกัด (B_x, B_y) และ (E_x, E_y) กำหนดกรอบสี่เหลี่ยมตามพิกัดที่ได้
8. ตรวจสอบเส้นกรอบสี่เหลี่ยมที่กำหนดขึ้นจะต้องไม่พบจุดสีดำบนเส้นเหล่านั้น
 - 8.1. หากยังพบจุดสีดำบนเส้นกรอบ ให้ลดขนาดกรอบลงหนึ่ง Pixel
 - 8.2. ตรวจสอบจุดสีดำบนเส้นกรอบที่กำหนดซ้ำอีกครั้ง และทำซ้ำไปเรื่อย ๆ จนกระทั่งไม่พบจุดสีดำบนเส้นกรอบที่กำหนดขึ้นใหม่
9. บันทึกค่าพิกัดของกรอบที่กำหนดลงในฐานข้อมูลเพื่อใช้ในการกำหนดกรอบโดยอัตโนมัติในแบบฟอร์มที่มีการกรอกข้อมูลแล้ว

จากวิธีการที่กล่าวมาข้างต้น สามารถแสดงภาพตัวอย่างการกำหนดกรอบในสัญลักษณ์ \circ , \square ดังรูปที่ 3.10



รูปที่ 3.10 ความถี่ของจุดสีดำในแนวตั้งและแนวนอนของเครื่องหมาย \circ , \square และพิกัดจุดต่าง ๆ ที่กำหนดเป็นกรอบข้อมูล

การพิจารณาความถี่ของจุดสีด้าในแนวตั้งของเครื่องหมายขลิขิตหรือวงเล็บ () และเครื่องหมาย [] จะแสดงให้เห็นว่

1. มีช่วงแคบ ๆ ที่มีความถี่ของจุดสีด้าที่สูงมาก แล้วตามด้วยช่องว่างของตัวอักษรที่ซึ่งจะต้องมีค่าอยู่ในระหว่างค่าคงที่ค่าหนึ่ง ก่อนที่จะพบช่วงความถี่ของจุดสีด้าที่มีสูงมากที่จุดสีด้าสิ้นสุด
2. จะต้องพบช่องว่างซึ่งมีค่าไม่เกินค่าคงที่อีกค่าหนึ่ง ก่อนที่จะพบความถี่ของตัวอักษรที่เป็นคำอธิบายต่อไป

หากไม่เป็นไปตามเงื่อนไขที่กล่าวมาจะถือว่าบริเวณนั้นไม่ใช่บริเวณที่จะต้องกำหนดกรอบ (mark) เพื่อจะบันทึกข้อมูล

การกำหนดประเภทข้อมูลเบื้องต้น (Default data type) ของกรอบข้อความที่ปรากฏสัญลักษณ์ที่กล่าวมาข้างต้นนี้จะกำหนดประเภทข้อมูลเป็น Boolean

ตัวอย่างส่วนหนึ่งของภาพแบบฟอร์มเปล่าที่ประกอบด้วยเครื่องหมาย ○, □, () และ [] แสดงดังรูปที่ 3.11, 3.12, 3.13 และ 3.14 ตามลำดับ และผลที่ได้จากการกำหนดกรอบโดยอัตโนมัติในรูป 3.11, 3.12, 3.13 และ 3.14 แสดงดังรูปที่ 3.15, 3.16, 3.17 และ 3.18 ตามลำดับ

15-25 ปี

☐ 15-25 ปี

รูปที่ 3.11 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายวงกลม ○ ประกอบอยู่พร้อมด้วยความถี่ของจุดสีด้าในแนวตั้ง

15-25 ปี

☐ 15-25 ปี

รูปที่ 3.12 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายสี่เหลี่ยม □ ประกอบอยู่พร้อมด้วยความถี่ของจุดสีด้าในแนวตั้ง

15-25 ปี

๒ ๖ ๒๓-๓๓ ๓๓

รูปที่ 3.13 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมายวงเล็บ () ประกอบอยู่พร้อมด้วยความถี่ของจุด
สีดำในแนวตั้ง

15-25 ปี

๒ ๖ ๒๓-๓๓ ๓๓

รูปที่ 3.14 ส่วนหนึ่งของแบบฟอร์มที่มีเครื่องหมาย [] ประกอบอยู่พร้อมด้วยความถี่ของจุดสีดำใน
แนวตั้ง

15-25 ปี

รูปที่ 3.15 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.11

15-25 ปี

รูปที่ 3.16 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.12

15-25 ปี

รูปที่ 3.17 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.13



รูปที่ 3.18 ตำแหน่งและพื้นที่ของกรอบสำหรับเติมข้อความจากรูปที่ 3.14

ในกรอบที่ใช้ในการตอบคำถามที่มีคำตอบสองลักษณะคือใช่/ไม่ (Optional mark) การเติมข้อมูลประเภทนี้ จะใช้เครื่องหมาย ถูกหรือกากบาท (\checkmark, \times) หรือการไม่เติมเครื่องหมายแทนความหมายว่าไม่มีหรือไม่ได้เติมข้อมูล หรืออาจกำหนดให้เครื่องหมายถูกมีความหมายถึงการเลือกข้อมูลส่วนนั้น แต่หากมีการใช้เครื่องหมายกากบาทจะมีความหมายว่าไม่ต้องการเลือกบันทึกข้อมูลส่วนนั้น โดยลักษณะที่กำหนดในแต่ละแบบฟอร์มนั้นจะถูกกำหนดในขั้นตอนการกำหนดประเภทข้อมูลของฟิลด์กรอบข้อมูลในแบบฟอร์มที่มีการกรอกข้อมูลแล้ว ซึ่งจะได้อ้างถึงรายละเอียดในหัวข้อถัดไป

กรอบที่ปรากฏในแบบฟอร์มที่กรอกข้อความแล้วและแบบฟอร์มเปล่าจะทำการพิจารณาจากตำแหน่งความสัมพันธ์ของจุดอ้างอิง (Relative reference position) ค่าตำแหน่งความสัมพันธ์ของจุดอ้างอิงจะใช้ความสัมพันธ์ของระยะห่าง ที่คำนวณด้วยสมการของ Euclidean distance แสดงในสมการที่ (3.1) ซึ่งจะทำการคำนวณระยะห่างระหว่างตำแหน่งทางเรขาคณิตของจุดเริ่มต้นของบรรทัดในแบบฟอร์มและบรรทัดที่สนใจ

$$d(b_i, b_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3.1)$$

โดยที่ $d(b_i, b_j)$ คือระยะห่างระหว่างบรรทัดเริ่มต้น b_i ซึ่งมีพิกัดอยู่ที่จุด (x_i, y_i) และบรรทัดที่กำลังพิจารณา b_j ซึ่งมีพิกัดอยู่ที่จุด (x_j, y_j)

ตัวอย่างผลลัพธ์ของกระบวนการกำหนดกรอบบริเวณสำหรับกรอกข้อความโดยอัตโนมัติในแบบฟอร์มเปล่า แสดงตัวอย่างตำแหน่งและพื้นที่ของกรอบบริเวณที่ต้องกรอกข้อความ ดังรูปที่ 3.19

แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน

วันที่ เดือน พ.ศ.

เรื่อง

ข้าพเจ้า นามสกุล

อยู่บ้านเลขที่ หมู่ที่ ตำบล/แขวง

อำเภอ จังหวัด

รหัสไปรษณีย์ โทรศัพท์

มีความประสงค์ใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อกขอข้อมูลข่าวสารจาก
หน่วยงานของท่าน ดังต่อไปนี้

โดยต้องการ ขอตรวจสอบ ขอค้นหา

ขอค้นหาที่มีสำเนาถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามที่กำหนดไว้ (ถ้ามี)

ลงชื่อ ผู้ยื่นคำขอ

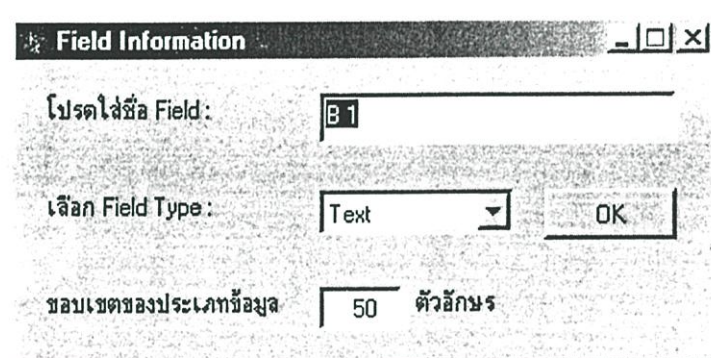
รูปที่ 3.19 แบบฟอร์มเปล่าพร้อมด้วยกรอบที่ถูกกำหนดขึ้นมาโดยอัตโนมัติ

3.2 การกำหนดประเภทของข้อมูลโดยผู้ใช้

เมื่อระบบทราบตำแหน่งและพื้นที่ของกรอบที่จะกรอกข้อมูลพร้อมด้วยการกำหนดชื่อฟิลด์ประเภทของข้อมูล ขอบเขตของข้อมูล (มีหรือไม่มีก็ได้) และชื่อฐานข้อมูลที่ถูกกำหนดโดยผู้ใช้งาน

กรอบของเส้นตรง ระบบจะทำการกำหนดประเภทของฟิลด์ให้เป็น Text ให้โดยอัตโนมัติ แต่ผู้ใช้อาจยังสามารถเปลี่ยนประเภทของฟิลด์เป็น Date, Month, Year, Age, Number และ Province ตามที่ต้องการได้

ตัวอย่างการกำหนดชื่อฟิลด์ และการกำหนดประเภทของข้อมูลโดยผู้ใช้ หลังจากทีระบบกำหนดข้อมูลให้แล้ว แต่ผู้ที่มีความประสงค์ที่จะต้องการแก้ไขเปลี่ยนแปลง ในรูปที่ 3.20 เป็นการแสดงการเปลี่ยนชื่อฟิลด์โดยผู้ใช้สามารถพิมพ์ชื่อฟิลด์ใหม่ที่ต้องการลงในช่องแรกที่ปรากฏแถบสี



รูปที่ 3.20 การเปลี่ยนชื่อฟิลด์โดยผู้ใช้

นอกจากการกำหนดชื่อฟิลด์โดยผู้ใช้ ระบบยังยินยอมให้ผู้ใช้สามารถกำหนดประเภทของข้อมูลได้ตามที่ต้องการสำหรับแต่ละฟิลด์ เนื่องจากข้อมูลในแต่ละฟิลด์มีความแตกต่างกัน ในบางฟิลด์อาจเป็นข้อความธรรมดา หรืออาจเป็นข้อมูลประเภทตัวเลข และระบบยังมีประเภทข้อมูลพิเศษที่เป็นข้อมูลเฉพาะมาให้เลือกด้วย อาทิเช่น ข้อมูลที่เป็นวัน, เดือน, ปี, อายุ, จังหวัด, รหัสไปรษณีย์ ฯลฯ ซึ่งประเภทข้อมูลเหล่านี้จะเป็นประโยชน์ในการตรวจสอบความถูกต้องหลังจากที่มีการรู้จำตัวอักษร เพื่อเป็นการตรวจสอบข้อมูลให้อยู่ในช่วงตามค่าที่กำหนด

ตัวอย่างภาพที่แสดงในรูปที่ 3.21 (a), 3.21 (b) และ 3.21 (c) เป็นการกำหนดประเภทของข้อมูลอื่นที่มีใช้ประเภทที่ระบบได้กำหนดไว้ให้ตั้งแต่แรก โดยผู้ใช้เป็นผู้เลือกเองจากที่มีอยู่ในรายการประเภทข้อมูล

Field Information [] [] [X]

โปรดใส่ชื่อ Field :

เลือก Field Type : [v]

(a) การกำหนดประเภทของข้อมูลเป็น Date โดยผู้ใช้

Field Information [] [] [X]

โปรดใส่ชื่อ Field :

เลือก Field Type : [v]

ขอบเขตของประเภทข้อมูล

เดือน

มกราคม - ธันวาคม

ม.ค. - ธ.ค.

01 - 12

(b) การกำหนดประเภทของข้อมูลเป็น Month และการกำหนดขอบเขตข้อมูลโดยผู้ใช้

Field Information [] [] [X]

โปรดใส่ชื่อ Field :

เลือก Field Type : [v]

ขอบเขตของประเภทข้อมูล

ปี

พ.ศ. 2300 - 2599

ค.ศ. 1900 - 2100

(c) การกำหนดประเภทของข้อมูลเป็น Year และการกำหนดขอบเขตข้อมูลโดยผู้ใช้

รูปที่ 3.21 การกำหนดประเภทของข้อมูลและการกำหนดขอบเขตข้อมูลโดยผู้ใช้

กรอบที่ใช้ในการตอบคำถามที่มีคำตอบสองลักษณะคือใช่/ไม่ ระบบจะทำการกำหนดประเภทของฟิลด์ให้เป็นข้อมูลเชิงตรรกศาสตร์ (Option) ให้โดยอัตโนมัติ

ตัวอย่างการกำหนดประเภทข้อมูลเชิงตรรกศาสตร์ โดยระบบซึ่งผู้ใช้สามารถเปลี่ยนแปลงแก้ไขประเภทข้อมูลลักษณะนี้ได้โดยเลือกรูปแบบที่ระบบได้กำหนดไว้ให้ แสดงในรูปที่ 3.22

Field Information [Close] [OK]

โปรดใส่ชื่อ Field :

เลือก Field Type :

ขอบเขตของประเภทข้อมูล

Boolean = Yes/No = True/False

มีเครื่องหมาย = จริง ไม่มีเครื่องหมาย = เท็จ

เครื่องหมายถูก = จริง เครื่องหมายผิด = เท็จ

เครื่องหมายถูก = จริง ไม่มีเครื่องหมาย = เท็จ

เครื่องหมายผิด = จริง ไม่มีเครื่องหมาย = เท็จ

(a) การแก้ไขชื่อฟิลด์ของข้อมูลเชิงตรรกศาสตร์โดยผู้ใช้

Field Information [Close] [OK]

โปรดใส่ชื่อ Field :

เลือก Field Type :

ขอบเขตของประเภทข้อมูล

Boolean = Yes/No = True/False

มีเครื่องหมาย = จริง ไม่มีเครื่องหมาย = เท็จ

เครื่องหมายถูก = จริง เครื่องหมายผิด = เท็จ

เครื่องหมายถูก = จริง ไม่มีเครื่องหมาย = เท็จ

เครื่องหมายผิด = จริง ไม่มีเครื่องหมาย = เท็จ

(b) การแก้ไขประเภทข้อมูลเป็นข้อมูลเชิงตรรกศาสตร์โดยผู้ใช้

รูปที่ 3.22 การแก้ไขประเภทข้อมูลเชิงตรรกศาสตร์โดยผู้ใช้

ในกรณีที่ข้อมูลส่วนนั้นเป็นตัวเลขที่เกิดจากการเรียงลำดับ สามารถเปลี่ยนแปลงให้เป็นประเภทข้อมูลตัวเลข ดังแสดงในรูปที่ 3.23

รูปที่ 3.23 การเปลี่ยนประเภทข้อมูลเป็นตัวเลขโดยผู้ใช้

นอกจากการกำหนดประเภทข้อมูลแล้ว ระบบยังยินยอมให้สามารถกำหนดขอบเขตของข้อมูลได้ตามที่ต้องการ ตัวอย่างเช่นการกำหนดขอบเขตอายุที่สนใจให้อยู่ในช่วง 15-20 ปี ดังแสดงในรูปที่ 3.24

รูปที่ 3.24 การกำหนดขอบเขตของข้อมูลอายุที่สนใจโดยผู้ใช้

3.3 การโหลดกรอบโดยอัตโนมัติ (Automatic load mark)

เมื่อระบบรับภาพแบบฟอร์มที่กรอกข้อความแล้วเข้ามา ก็จะทำการโหลดกรอบที่มีการกรอกข้อมูลขึ้นมาให้โดยอัตโนมัติ โดยใช้ข้อมูลจากแบบฟอร์มเปล่าที่มีความเหมือนกับแบบฟอร์มที่มีการกรอกข้อมูลนั้น ๆ มากที่สุด

กระบวนการโหลดกรอบจะใช้ตำแหน่งความสัมพันธ์กับแบบอ้างอิงของแบบฟอร์มเปล่าเพื่อที่จะหาตำแหน่งที่จะกำหนดกรอบในแบบฟอร์มที่กรอกข้อความแล้ว

การกำหนดกรอบโดยอัตโนมัติในแบบฟอร์มที่มีการกรอกข้อมูลแล้ว ระบบจะโหลดกรอบข้อมูลเหล่านั้น โดยอาศัยค่าพิกัดต่าง ๆ ที่บันทึกไว้ในขณะที่มีการกำหนดกรอบข้อมูลในแบบฟอร์ม

เปล่าและมีการตรวจสอบกรอบในบริเวณที่กำหนดกรอบอีกครั้ง เพื่อให้แน่ใจว่ากรอบข้อมูลที่กำหนดครอบคลุมตัวอักษรทั้งหมด โดยใช้หลักการหาขอบภาพตัวอักษรเปรียบเทียบกับพิกัดกับเส้นกรอบของข้อความที่กำหนดขึ้นมาโดยอัตโนมัติ

ตัวอย่างจากบางส่วนของแบบฟอร์มที่นำมากรอกข้อความแล้วให้ระบบทำการโหลดกรอบโดยอัตโนมัติ เพื่อทำการกำหนดกรอบของบริเวณที่กรอกข้อความแสดงผลดังรูปที่ 3.25

วันที่18.....	เดือน	...กุมภาพันธ์.....	พ.ศ.	...2546.....
วันที่	11	เดือน	เมษายน	พ.ศ.	2546
วันที่	23	เดือน	มิถุนายน	พ.ศ.	2546
<input checked="" type="checkbox"/>	15-25 ปี	<input checked="" type="checkbox"/>	15-25 ปี		
<input checked="" type="checkbox"/>	15-25 ปี	<input checked="" type="checkbox"/>	15-25 ปี		
<input checked="" type="checkbox"/>	15-25 ปี	<input checked="" type="checkbox"/>	15-25 ปี		
<input checked="" type="checkbox"/>	15-25 ปี	<input checked="" type="checkbox"/>	15-25 ปี		

รูปที่ 3.25 ส่วนหนึ่งของแบบฟอร์มที่กรอกข้อความและได้โหลดกรอบโดยอัตโนมัติ

จากรูปที่ 3.25 พบว่ากรอบของข้อความ “กุมภาพันธ์”, “เมษายน” และ “มิถุนายน” มีการขยายความสูงของกรอบตามค่าความถี่ของจุดสีดำที่ปรากฏอยู่บนจุดสูงสุดตามแนวแกน y ที่พบในแต่ละบรรทัด

เมื่อแบบฟอร์มที่มีการกรอกข้อมูลสามารถจับคู่กับแบบฟอร์มเปล่าที่มีอยู่ในฐานข้อมูลได้อย่างถูกต้อง จากนั้นระบบทำการโหลดกรอบโดยอัตโนมัติ เพื่อทำการกำหนดกรอบของบริเวณที่กรอกข้อความแสดงผลดังรูปที่ 3.26

แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน

วันที่ เดือน พ.ศ.

เรียน

ข้าพเจ้า ชื่อ นามสกุล

อยู่บ้านเลขที่ หมู่ที่ ตำบล/แขวง

อำเภอ/เขต จังหวัด

รหัสไปรษณีย์ โทรศัพท์

มีความประสงค์ใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. ๒๕๔๐ เพื่อขอข้อมูลข่าวสารจาก
หน่วยงานของท่าน ดังต่อไปนี้

1.

2.

โดยต้องการ ขอตรวจดู ขอสำเนา

ขอสำเนาที่มีคำรับรองถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามที่กำหนดไว้ (ถ้ามี)

ลงชื่อ ผู้ยื่นคำขอ

รูปที่ 3.26 แสดงแบบฟอร์มที่กรอกข้อความแล้วพร้อมทั้งการไหลครบโดยอัตโนมัติ

3.4 การตรวจสอบความถูกต้องของข้อความ (Data verification)

กระบวนการตรวจสอบความถูกต้องของข้อความ เป็นกระบวนการสุดท้ายหลังจากการรู้จำตัวอักษร ประเภทของฟิลด์สามารถช่วยลดขอบเขตความน่าจะเป็นของตัวอักษรและใช้เพื่อในการแก้ไขข้อความให้ถูกต้อง

ขอบเขตของข้อมูลต่าง ๆ ที่ระบบได้กำหนดนี้ผู้ใช้สามารถแก้ไขหรือเปลี่ยนแปลงได้ตามที่ต้องการ อาทิตัวอย่างเช่น

ประเภทข้อมูลเป็นตัวเลข ได้แก่

อายุ มีค่าอยู่ระหว่าง 1-150

วันที่ มีค่าอยู่ระหว่าง 1-31

ปี มีค่าอยู่ระหว่าง 1900-2200 ในปีคริสต์ศักราช และ 2400-2600 ในปีพุทธศักราช

ประเภทข้อมูลเป็นข้อความ ได้แก่

จังหวัด (ชื่อทั้ง 76 จังหวัด)

เดือน (มกราคม-ธันวาคม ทั้ง 12 เดือน)

เดือนที่ระบุเป็นตัวอักษรย่อ (ม.ค.-ธ.ค.)

3.5 การเก็บข้อมูลที่สามารถรู้จำได้ลงในฐานข้อมูลที่ถูกต้อง

หลังจากที่มีการตรวจสอบความถูกต้องและความเป็นไปได้ของข้อมูลเรียบร้อยแล้ว ระบบจะทำการบันทึกข้อมูลต่าง ๆ ที่สามารถรู้จำได้ลงในฐานข้อมูลที่มีการกำหนดไว้ก่อนที่จะมีการรู้จำข้อมูล ซึ่งโดยมากจะกำหนดชื่อฐานข้อมูลที่จัดเก็บเป็นชื่อเดียวกันกับชื่อไฟล์ของแบบฟอร์มเปล่า เพื่ออำนวยความสะดวกในการใช้งาน

บทที่ 4

ผลการทดลอง

ในงานวิจัยนี้ผู้วิจัยทำการวิจัยโดยการเขียน โปรแกรมตามหลักการต่าง ๆ ที่ได้กล่าวมาแล้วในบทก่อนหน้านี้ เพื่อทำการทดลองด้วยแบบฟอร์มต่าง ๆ ที่มีความหลากหลาย โดยใช้แบบฟอร์มเปล่าในการทดลองประมาณ 120 แบบฟอร์ม และแบบฟอร์มที่กรอกข้อความแล้วอีกประมาณ 300 แบบฟอร์ม ซึ่งแบบฟอร์มต่าง ๆ ได้ใช้ลายมือของผู้ที่อยู่ในช่วงอายุ 10-60 ปีจำนวน 4 คนเป็นผู้กรอกข้อมูลทั้งหมดเพื่อทดสอบการทำงานของระบบ และได้นำผลการทดลองในส่วนของเวลาที่ได้มาแสดงผลตามหัวข้อย่อต่าง ๆ ดังต่อไปนี้

4.1 ผลการทดลอง

ในงานวิจัยนี้ทำการทดลองบนเครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผลเป็น Intel Pentium IV 2.4 GHz หน่วยความจำขนาด 512 Mbytes และใช้ Microsoft Visual Basic 6.0 ในการพัฒนาโปรแกรม

การทำงานขั้นตอนแรกคือ เมื่อรับภาพแบบฟอร์มเปล่าเข้ามา ระบบจะแบ่งบรรทัดด้วยค่าความถี่ของจุดสีดำที่พบในแบบฟอร์มดังที่ได้กล่าวมาแล้วในบทก่อนหน้านี้ จากนั้นระบบจะกำหนดกรอบบริเวณที่จะบันทึกข้อมูลโดยอัตโนมัติ

จากนั้นผู้ใช้งานจะต้องเป็นผู้กำหนดชื่อฟิลด์หรือกำหนดเปลี่ยนแปลงประเภทของข้อมูลตามที่ต้องการ และผู้ใช้งานยังสามารถเลือกที่จะกำหนดการบันทึกกรอบข้อมูลในแต่ละกรอบ หรือจะไม่นับที่ก็ได้

ตัวอย่างกรอบข้อมูลที่จะไม่นับที่กลงในฐานะข้อมูลสำหรับการรู้จำตัวอักษรคือ กรอบของลายเซ็นชื่อ เนื่องจากการบันทึกลายเซ็นไม่สามารถนำไปรู้จำลายมือตัวอักษรออกมาเป็นตัวอักษรแต่ละตัวได้ แต่ในบางกรณีผู้ใช้งานอาจกำหนดให้เก็บกรอบข้อมูลของลายเซ็นด้วยก็ได้สำหรับประโยชน์อื่นตามลักษณะการใช้งาน และสุดท้ายจะเก็บรายละเอียดเหล่านั้นที่ผู้ใช้งานได้กำหนดไว้ลงในฐานข้อมูลแบบฟอร์มเปล่า

จากการทดลองเมื่อระบบรับภาพแบบฟอร์มเปล่าเข้ามาผ่านกระบวนการต่าง ๆ แล้วกำหนดกรอบที่จะกรอกข้อมูลโดยอัตโนมัติได้แสดงตัวอย่างดังรูปที่ 4.1 จากนั้นทำการกำหนดชื่อฟิลด์โดยผู้ใช้งานได้แสดงไว้ในรูปที่ 4.2 และทำการเปลี่ยนแปลงประเภทข้อมูลโดยผู้ใช้งานได้แสดงไว้ในรูปที่ 4.3 – รูปที่ 4.10 และทำการบันทึกที่รายละเอียดต่าง ๆ ที่กำหนดขึ้นลงในฐานข้อมูลโดยใช้ชื่อฐานข้อมูลเป็นชื่อเดียวกันกับชื่อแบบฟอร์มแบบฟอร์มเปล่า นั้น ๆ โดยอัตโนมัติ

Form Management - [แบบฟอร์มข้อมูลข่าวสารของกรมพัฒนาเกษตร] _ | ๒ | X

File Empty Form Fill Form Zoom About

แบบฟอร์มขอข้อมูลข่าวสารของกรมพัฒนาเกษตร

วันที่ เดือน พ.ศ.

.....

หมู่บ้าน ตำบล อำเภอ จังหวัด

โทรศัพท์

พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก

.....

.....

รูปที่ 4.1 แบบฟอร์มเปล่าที่มีการกำหนดกรอบที่จะกรอกข้อมูลโดยอัตโนมัติ

Form Management - [แบบฟอร์มข้อมูลข่าวสารของกรมพัฒนาเกษตร] _ | ๒ | X

File Empty Form Fill Form Zoom About

แบบฟอร์มขอข้อมูลข่าวสารของกรมพัฒนาเกษตร

วันที่ เดือน พ.ศ.

.....

หมู่บ้าน ตำบล อำเภอ จังหวัด

โทรศัพท์

พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก

.....

.....

รูปที่ 4.2 การกำหนดชื่อฟิลด์ของข้อมูลลงในฐานข้อมูล

Form Management - [แบบกำหนดข้อมูลประชากรกรมพัฒนาฯ (ipg)]

File Empty Form Fill Form Zoom About

แบบกำหนดข้อมูลประชากรกรมพัฒนาฯ

วันที่ เดือน พ.ศ.

หมู่บ้าน

จังหวัด

โทรศัพท์

พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก

Field Information

ไปดใช้ Field: Date

เลือก Field Type: Text

ประเภทของประเภทข้อมูล: Text, Yes/No, Numeric, Age, Year, ZipCode, Date, Month

OK

รูปที่ 4.3 การเปลี่ยนแปลงประเภทข้อมูลเป็น Date ก่อนบันทึกลงในฐานข้อมูล

Form Management - [แบบกำหนดข้อมูลประชากรกรมพัฒนาฯ (ipg)]

File Empty Form Fill Form Zoom About

แบบกำหนดข้อมูลประชากรกรมพัฒนาฯ

วันที่ เดือน พ.ศ.

นาม

ตำบล/แขวง

จังหวัด

โทรศัพท์

ลข่าวสารของราชการ พ.ศ. 2540 เพื่อขอข้อมูลข่าวสารจาก

Field Information

ไปดใช้ Field: Month

เลือก Field Type: Text

ประเภทของประเภทข้อมูล: Text, Yes/No, Numeric, Age, Year, ZipCode, Date, Month

OK

รูปที่ 4.4 การเปลี่ยนแปลงประเภทข้อมูลในช่องที่สอง ก่อนบันทึกลงในฐานข้อมูล

Form Management - [แบบกำหนดข้อมูลสำหรับกรมสรรพากร] - [X] X

File Empty Form Fill Form Zoom About

วันที่ เดีย

เรียน _____

Field Information - [X] X

โปรดใส่ชื่อ Field:

เลือก Field Type:

ขนาดของประเภทข้อมูล:

อยู่บ้านเลขที่ น
 อำเภอ/เขต ตำบล/แขวง
 รหัสไปรษณีย์ จังหวัด
 มีความประสงค์ใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ
 หน่วยงานของท่าน ดังต่อไปนี้

1.
2.

รูปที่ 4.9 การเปลี่ยนชื่อฟิลด์แต่คงกำหนดประเภทข้อมูลเป็น Text ก่อนบันทึกลงในฐานข้อมูล

Form Management - [แบบกำหนดข้อมูลสำหรับกรมสรรพากร] - [X] X

File Empty Form Fill Form Zoom About

2.

โดยต้องการ ขอตรวจดู () ขอสำเนา
 () ขอสำเนาที่มีคำรับรองถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามอัตราที่กำหนดไว้ (ถ้ามี)

Field Information - [X] X

โปรดใส่ชื่อ Field:

เลือก Field Type:

ขนาดของประเภทข้อมูล: Boolean = Yes/No = True/False

<input checked="" type="radio"/> มีเครื่องหมาย = จริง	<input type="radio"/> ไม่มีเครื่องหมาย = เท็จ
<input type="radio"/> เครื่องหมายถูก = จริง	<input type="radio"/> เครื่องหมายผิด = เท็จ
<input type="radio"/> เครื่องหมายถูก = จริง	<input type="radio"/> ไม่มีเครื่องหมาย = เท็จ
<input type="radio"/> เครื่องหมายผิด = จริง	<input type="radio"/> ไม่มีเครื่องหมาย = เท็จ

รูปที่ 4.10 การกำหนดประเภทข้อมูลเป็น Boolean(Yes/No) ก่อนบันทึกลงในฐานข้อมูล

แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน

วันที่ ๑ เดือน สิงหาคม พ.ศ. ๒๕๕๖

เรียน อธิบดีกรมพัฒนาที่ดิน

ข้าพเจ้า ชื่อ นาง พจนน นามสกุล แสงสว่างฉายา
 อยู่บ้านเลขที่ ๒๕ หมู่ที่ ๖ ตำบล/แขวง ศรีโพธิ์
 อำเภอ/เขต เมือง จังหวัด เชียงใหม่
 รหัสไปรษณีย์ ๕๕๐๐๐ โทรศัพท์ ๐-๖๖๕๑-๑๕๖๖

มีความประสงค์ใช้สิทธิตาม พ.ร.บ. ข้อมูลข่าวสารของราชการ พ.ศ. ๒๕๔๐ เพื่อขอข้อมูลข่าวสารจาก
 หน่วยงานของท่าน ดังต่อไปนี้

1. เอกสารแนบคำสั่งกรมพัฒนาที่ดิน ปี ๒๕๕๖
2. ที่อยู่สำนักงานเขตติดต่อของสำนักงานพัฒนาที่ดินเขตที่ ๑๕๖๖

โดยต้องการ ขอตรวจดู ขอสำเนา
 ขอสำเนาที่มีคำรับรองถูกต้อง

โดยยินยอมเสียค่าธรรมเนียมตามอัตราที่กำหนดไว้ (ถ้ามี)

ลงชื่อ นาง พจนน ผู้ยื่นคำขอ
นาง พจนน

รูปที่ 4.14 ภาพเต็มทั้งแบบฟอร์มที่กรอกข้อมูลเรียบร้อยแล้ว โหลดกรอบข้อมูลขึ้นมาโดยอัตโนมัติ

เมื่อแบบฟอร์มที่มีการกรอกข้อมูลแล้ว สามารถจับคู่กับแบบฟอร์มเปล่าที่มีอยู่ ต่อมาระบบจะทำการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติ ซึ่งข้อมูลเวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์มที่มีการกรอกข้อมูลแล้ว ในการทดลองนี้พบว่าเวลาที่ใช้โหลดมีค่าไม่เกิน 1 วินาที ดังแสดงรายละเอียดในตารางที่ 4.1 ต่อไปนี้

ตารางที่ 4.1 เวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อความ(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.141107	fill สัมมนา SMEs
0.187113	ใบสมัครสมาชิกสถาบัน EDI วนิษา
0.187988	ใบสมัครสมาชิกสถาบัน EDI กุลธร
0.188102	ใบสมัครสมาชิกสถาบัน EDI รุ่งนภา
0.217875	ใบสมัคร EDI อติเรก
0.234000	Request for unofficial transcript Silviara ABAC
0.234863	Request for unofficial transcript Weena ABAC
0.235227	Request for unofficial transcript Watcharapol ABAC
0.250227	Request for unofficial transcript Panomwan ABAC
0.311875	NEC หญิง29 2
0.311875	แบบเสนอโครงการวิจัย2 องอาจ kmitl
0.312000	คำร้องทั่วไป รโหฐาน มน18
0.312000	บันทึกข้อความ สุรกี บว 25.2 kku
0.312113	ใบประเมินผลร่วมโครงการ ABAC แอโรบิก
0.312238	NEC หญิง32 3
0.312238	คำร้องขอพักการเรียน นัฐพงศ์ Ben
0.312863	NEC ชาย35 2
0.312863	แบบคำร้องขอรับบริการข้อมูลข่าวสาร จรูญ
0.312863	แบบเสนอโครงการวิจัย1 พิมพ์รณ kmitl
0.313102	ใบแจ้งผลการตรวจสอบหนี้สินของนศ ชลดา มข
0.313102	ใบมอบฉันทะ วิชญ์ สปส
0.313227	แบบแจ้งผลการสอบประมวลความรู้ เกษตร 1-44 บว8
0.327863	คำร้องขอลาออก วิกา Ben

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.327863	คำร้องขอหนังสือรับรองผลการศึกษา สุทธิพงษ์ kmitl
0.327863	คำร้องทั่วไป ส่งเข้า วิภา kmitl
0.327863	แบบขออนุมัติเปลี่ยนแปลง-โครงการวิทยานิพนธ์ เควิทย์ มอ
0.327863	แบบขออนุมัติเปลี่ยนแปลง-โครงการวิทยานิพนธ์ วรรณวิไล มอ
0.327863	แบบขออนุมัติเปลี่ยนแปลง-โครงการวิทยานิพนธ์ เอนก มอ
0.327863	แบบขออนุมัติเปลี่ยนระดับการศึกษา อลงกรณ์ บว13 มอ
0.327863	แบบคำร้องขอรับบริการข้อมูลข่าวสาร หนึ่งฤดี
0.327863	ใบแจ้งผลการตรวจสอบหนังสือของนศ โชติรส มข
0.327863	ใบสมัครวิศวะ มข นิติวินน์
0.327863	ใบสมัครอบรมสัมมนา วลัยพร EQStar
0.327875	แต่งตั้งอาจารย์บัณฑิต 3-3 เจริญศักดิ์ บววป15 kmitl
0.327875	แบบการตรวจวิทยานิพนธ์ กัลยารัตน์ บว6-1 มอ
0.327875	ใบประเมินผลร่วมโครงการ ABAC Dummy
0.327875	ศิษย์เก่าดีเด่น2 รุ่งนภา มหิดล
0.327875	หนังสือสละสิทธิการเช่าที่ราชพัสดุ อรอุมา
0.327988	การส่งต่อผู้ป่วยวัลโรค เสรี
0.327988	คำร้องทั่วไป ส่งเข้า สถิตาพร kmitl
0.327988	แบบการตรวจวิทยานิพนธ์ ทองเจริญ บว6-1 มอ
0.327988	แบบแจ้งผลการสอบวิทยานิพนธ์ สิทธิโชค บว7 มอ
0.327988	ใบสมัครวิศวะ มข วรรณณา
0.327988	หนังสือรับรองการเป็นอ.ที่ปรึกษา อัครพล บววป1 kmitl
0.327988	หนังสือสละสิทธิการถอนเงินประกันการเช่าที่ราชพัสดุ พงศกร
0.328000	คำร้องขอมีสถิตีสอบปลายภาค คูสิต Ben
0.328000	ใบแจ้งผลการตรวจสอบหนังสือของนศ พัชรพันธ์ มข
0.328000	ใบมอบฉันทะ มนตรี สุปส
0.328000	ใบสมัครขอรับทุนราชกรีกิจาสมอสร์ วีระเดช ku
0.328000	ใบสมัครอบรมสัมมนา จักรพรรดิ EQStar

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.328000	หนังสือสละสิทธิการเช่าที่ราชพัสดุ นันทนีย์
0.328113	NEC ซาช35 3
0.328113	NEC3
0.328113	แบบฟอร์มการส่งต่อผู้ป่วยวัณโรค ทองดี
0.328113	ใบมอบคันทะ สุเชค สปส
0.328113	ใบสมัครอบรมสัมมนา ถักรสุตา EQStar
0.328113	ศิษย์เก่าดีเด่น1 วิทยา มหิดล
0.328227	คำร้องขอพักการเรียน ชัยสิทธิ์ บถ
0.328227	คำร้องลาออก นพพร
0.328227	แบบขออนุมัติย้ายสาขาวิชา อธิป บว12 มอ
0.328227	แบบคำร้องขอหนังสือส่งตัวกลับคืนสังกัด ยุพวรรณ คำร้อง4 มอ
0.328227	แบบแจ้งผลการสอบประมวลความรู้ เกษคร 1-45 บว8
0.328227	แบบแจ้งผลการสอบวิทยานิพนธ์ วนิดา บว7 มอ
0.328227	แบบตรวจฟอร์มโครงการวิทยานิพนธ์ คุณัญญิณีพนธ์ ซาลิสสา ku
0.328227	ใบแจ้งผลการตรวจสอบหนังสือของนศ รักเกียรติ มข
0.328227	ใบประเมินผลร่วมโครงการ ABAC Freshy Night
0.328238	NEC หลิง29 3
0.328238	NEC2
0.328238	คำร้องขอพักการเรียน ประพัฒน์ Ben
0.328238	คำร้องขอมิติสิทธิสอบปลายภาค วุฒิชัย Ben
0.328238	คำร้องขอหนังสือรับรองผลการศึกษา เจษฎา สจล
0.328238	แบบการตรวจวิทยานิพนธ์ กัทรพล บว6-1 มอ
0.328238	แบบขออนุมัติย้ายสาขาวิชา รัตติกาล
0.328238	แบบคำร้องขอหนังสือส่งตัวกลับคืนสังกัด รุจิโรจน์ คำร้อง4 มอ
0.329102	คำร้องขอลาออก วรรณภา Ben
0.329102	คำร้องทั่วไป ชีรชัย มน18
0.329102	บันทึกข้อความ อังตรา บว 25.2 kku

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.329102	แบบคำร้องขอรับบริการข้อมูลข่าวสาร นิพนธ์
0.329102	แบบคำร้องขอรับบริการข้อมูลข่าวสาร เบลูจพล
0.329102	แบบประเมินผลการพิจารณา EECON24 ทินกรณ
0.329102	ใบสมัครขอรับทุนสนับสนุนการทำวิทยานิพนธ์ วัฒนา kmitl
0.329102	ใบสมัครวิสวะ มช ทรงภพ
0.329102	ใบสมัครอบรมสัมมนา eq พูนทรัพย์
0.342875	แบบการตรวจวิทยานิพนธ์ ภัทรพล บว6-2 มอ
0.342875	ใบมอบฉันทะ เสนาะ สปส
0.342875	ใบสมัครขอรับทุนราชกรีฑาสโมสร ภัทรพล
0.342875	ใบสมัครอบรมสัมมนา ชนินตรา มธ ส่วนภูมิภาค
0.342875	ประวัติเพื่อแต่งตั้งบุคคลเป็นอ.บัณฑิตวิทยาลัย ภาคภูมิ บว11-2 มอ
0.342875	รับรองผลงานวิชาการ บว ว7 บงกช kmitl
0.342875	หนังสือรับรองการเป็นอ.ที่ปรึกษา นันทนา บววป1 kmitl
0.342875	หนังสือสละสิทธิการเข้าที่ราชพัสดุ แม่ฮ่องสอน
0.343000	คำร้องขอลาออก ภรมาศ บวทบ6 kmitl
0.343000	คำร้องทั่วไป มน18 ปิยะนุช
0.343000	แต่งตั้งอาจารย์บัณฑิต 2-3 เจริญศักดิ์ บววป15 kmitl
0.343000	แบบขออนุมัติเทียบโอนรายวิชา พนม บว10 มอ
0.343000	แบบประวัติอ.ประจำหรืออ.พิเศษ สาขา อุบลวรรณ Ku
0.343000	แบบประวัติอาจารย์ประจำหรืออาจารย์พิเศษ
0.343000	ประวัติเพื่อแต่งตั้งบุคคลเป็นอ.บัณฑิตวิทยาลัย พิชญ่า บว11-2 มอ
0.343000	รับรองผลงานวิชาการ บว ว7 วิไลวรรณ kmitl
0.343000	ศิษย์เก่าดีเด่น3 รุ่งนภา มหิดล
0.343000	หนังสือสละสิทธิการขอถอนเงินประกันการเข้าที่ราชพัสดุ อวรรณ
0.343863	แบบขออนุมัติเทียบโอนรายวิชา ชลดา บว10 มอ
0.343863	แบบขออนุมัติเปลี่ยนระดับการศึกษา ชลดา บว13 มอ
0.343863	แบบประเมินผลการพิจารณา EECON24 Net

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการไหลครบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการไหลครบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.343863	ใบสมัครขอรับทุนราชกรีฑาสโมสร ลักษณะ ku
0.343863	ใบสมัครอบรมสัมมนา ชัยยศ มธ1
0.343863	ศิษย์เก่าดีเด่น1 รัตนาวลี มหิดล
0.343863	ศิษย์เก่าดีเด่น3 รัตนาวลี มหิดล
0.343863	หนังสือขอความร่วมมือ ลัดดาวรรณ บว16 มน
0.343988	คำร้องขอหนังสือรับรองผลการศึกษา จันทรา kmil
0.343988	คำร้องทั่วไป ส่งเข้า อรุณ kmil
0.343988	แบบแจ้งผลการสอบวิทยานิพนธ์ กนกวรรณ บว7 มอ
0.343988	แบบนำเสนอบทความ EECON-24 วิชุดา
0.343988	ใบสมัครขอรับทุนราชกรีฑาสโมสร ภัทรกร ku
0.344102	NEC หญิง32 2
0.344102	ขอรับทุนอุดหนุนและส่งเสริมวิทยานิพนธ์ กมลชนก Ku
0.344102	แบบเสนอโครงการวิจัย2 พิมพ์ธรรม kmil
0.344102	หนังสือขอความร่วมมือ อนุวัฒน์ บว16 มน
0.344102	หนังสือถึงผู้เชี่ยวชาญ สุพจน์ บว12 มน
0.344113	คำร้องขอใบคะแนนและหนังสือรับรอง อรอนงค์ บว8 Ku
0.344113	คำร้องขอพักการเรียน ชัยวัฒน์ Ben
0.344113	แบบการตรวจวิทยานิพนธ์ ทองเจริญ บว6-2 มอ
0.344113	แบบขออนุมัติเทียบโอนรายวิชา หาญชัย บว10 มอ
0.344113	แบบขออนุมัติเปลี่ยนระดับการศึกษา อนงนาถ บว13 มอ
0.344113	แบบตรวจฟอร์มโครงการวิทยานิพนธ์ คุณกัญญา ชัยสุนทร ku
0.344113	ใบสมัครอบรมสัมมนา มธ รักเกียรติ
0.344113	ศิษย์เก่าดีเด่น1 รุ่งนภา มหิดล
0.344113	หนังสือขอความร่วมมือ ศิริวิทย์ บว13 มน
0.344227	การส่งต่อผู้ป่วยวิน โรค แก้วขวัญ
0.344227	คำร้องขอมิติทริสสอบปลายภาค ขนิษฐา Ben
0.344227	คำร้องขอสอบกรณีพิเศษ คำรงค์ Ben

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการโหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.344227	คำร้องทั่วไป ดิเรก บวทบ1 kmitl
0.344227	คำร้องส่งรายงานฉบับสมบูรณ์ล่าช้า พิเศษ kmitl
0.344227	แบบขออนุมัติย้ายสาขาวิชา อรรถพ บว12 มอ
0.344227	ใบสมัครอบรมสัมมนา จักรพรรดิ มธ1
0.344227	ใบสมัครอบรมสัมมนา ส่วนภูมิภาค ชูติมา มธ
0.344238	แบบขอสอบวิทยานิพนธ์ ภัทรพล บว5 มอ
0.344238	แบบขออนุมัติเทียบโอนรายวิชา เมธาวิ บว10 มอ
0.344238	แบบนำส่งบทความ EECON-24 วัฒนชัย
0.344238	หนังสือถึงผู้เชี่ยวชาญ อนุวัฒน์ บว15 มน
0.358875	NEC หญิง32 3t
0.358875	แบบประวัติอ.ประจำหรืออ.พิเศษ สาขา จตุรงค์ Ku
0.358875	ใบสมัครอบรมสัมมนา นิตยา มธ ส่วนภูมิภาค
0.358988	คำร้องขอมีสถิตีสอบปลายภาค ไพโรจน์ Ben
0.358988	คำร้องขอสอบกรณีพิเศษ มาโนชญ์ Ben
0.358988	แบบขอสอบวิทยานิพนธ์ ปณิศา บว5 มอ
0.358988	แบบนำส่งบทความ EECON-24 ทินกรณั
0.358988	หนังสือถึงผู้เชี่ยวชาญ เบญจกุล บว12 มน
0.359000	แบบคำร้องขอหนังสือส่งตัวกลับคืนสังกัด ปราณี คำร้อง4 มอ
0.359000	แบบตรวจฟอร์มโครงการวิทยานิพนธ์ ดุษฎีนิพนธ์ กาญจนา ku
0.359000	ใบสมัครขอรับทุนสนับสนุนการทำวิทยานิพนธ์ พิมพ์พรณ kmitl
0.359000	ใบสมัครขอรับทุนอุดหนุนและส่งเสริมวิทยานิพนธ์ป. โท-เอก ยอดขวัญ Ku
0.359000	ใบสมัครอบรมสัมมนา สิริินภา มธ1
0.359000	หนังสือขอความร่วมมือ เบญจกุล บว13 มน
0.359000	หนังสือสละสิทธิการขอถอนเงินประกันการเข้าที่ราชพัสดุ วัลลภ
0.359113	ใบสมัครอบรม มธ.1 สาโรจน์
0.359113	ใบสมัครอบรมสัมมนา รุ่งนภา มธ
0.359238	คำร้องขอใบคะแนนและหนังสือรับรอง รัตนชัย บว8 Ku

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.359863	คำร้องขอสอบกรณีพิเศษ แสง โสม Ben
0.359863	คำร้องทั่วไป ธิดา มน18
0.359863	บันทึกข้อความ ภัทรพล บว 25.2 kku
0.359863	แบบแจ้งผลการสอบประมวลผลความรู้ 2-44 บว8 มอ
0.359863	แบบประเมินผลการพิจารณา EECON24 Doc
0.359863	หนังสือขอความร่วมมือ มัคคิกา บว16 มน
0.360102	คำร้องขอสอบกรณีพิเศษ เสวารส
0.360102	ใบสมัครอบรมสัมมนา วณิชชา มธ
0.360102	ใบสมัครอบรมสัมมนา วรรณทอง มธ ส่วนภูมิภาค
0.360102	หนังสือถึงผู้เชี่ยวชาญ สุรสิทธิ์ บว15 มน
0.360227	ศิษย์เก่าดีเด่น2 วิทยา มหิดล
0.374863	ศิษย์เก่าดีเด่น2 รัตนาวลี มหิดล
0.375000	คำร้องขอใบคะแนนและหนังสือรับรอง เทียมจันทร์ บว8 Ku
0.375000	แบบประวัติ.ประจำหรือ.พิเศษ สายข จิตรাত্র Ku
0.375000	หนังสือรับรองการเป็นอ.ที่ปรึกษา เจษฎา บววป1 kmitl
0.375102	คำร้องทั่วไป ส่งเข้า วิทยารายณ์ kmitl
0.375113	คำร้องขอหนังสือรับรองผลการศึกษา kmitl
0.375113	ใบสมัครอบรม มธ. อารีชา
0.375227	หนังสือขอความร่วมมือ ปริญา บว16 มน
0.375238	คำร้องขอลาออก พัชราภรณ์ Ben
0.375238	รับรองผลงานวิชาการ บว วป7 ธีรรัตน์ kmitl
0.390238	คำขอตรวจค้นคำขอรับสิทธิบัตร ไพรวลัย
0.390988	คำร้องขอเปลี่ยนชื่อ-ชื่อสกุล วัน เดือน ปีเกิด ชศ ศรีศักดิ์
0.391102	แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน นงนุช
0.391113	Plan-it Watcharapol
0.391227	ประวัติเพื่อแต่งตั้งบุคคลเป็นอ.บัณฑิตวิทยาลัย ธานี บว11-2 มอ
0.405863	ขอทุนการศึกษา 4-7 บว วป9 kmitl

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.405863	คำร้องขอ รบ. ม.ต้น พงศธร Ben
0.405863	คำร้องขอ รบ. ม.ปลาย กัจจกร Ben
0.405863	คำร้องขอเปลี่ยนชื่อ พจนเป็นวรรณศาสตร์ ben
0.405988	ข้อมูลการเรียนและความก้าวหน้าในการทำวิทยานิพนธ์ อุบลวรรณ คำร้อง4-1 มอ
0.405988	คำร้องขอปรับสภาพเป็นนักศึกษาสามัญ ประเทือง บวทป7 kmitl
0.406000	ความเห็นอาจารย์ที่ปรึกษา วาสิณี มอ
0.406000	คำรับรองจากอาจารย์ที่ปรึกษา พรพิมล บววป2 kmitl
0.406000	หนังสือมอบอำนาจ วรรณภา
0.406000	หนังสือรับรองของผู้ปกครอง 7-7 รศศริน บววป9 kmitl
0.406238	ขอดอนเงินประกันของเสียหยา3 ประพันธ์ kmitl
0.406238	หนังสือให้ความยินยอมของกลุ่มสมรส สุรพล
0.407102	คำร้องการไฟฟ้า สรอรอด
0.407102	แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน วชิรา
0.407227	NEC ชาย35 I
0.407227	คำร้องขอ รบ. ม.ปลาย สุจิรา Ben
0.407227	หนังสือมอบอำนาจ ธวัช
0.421000	NEC หญิง32 I
0.421000	คำร้องขอเปลี่ยนแปลงชื่อผู้ควบคุมวิทยานิพนธ์ พรพิมล บววป3 kmitl
0.421000	แบบฟอร์มคำร้องขอเปลี่ยนแปลงอาจารย์ผู้ควบคุมวิทยานิพนธ์ บัณฑิต
0.421863	interview result kitiyapa ABAC
0.421863	ขอทุนการศึกษา 2-7 บว วป9 kmitl
0.421863	หนังสือให้ความยินยอมของกลุ่มสมรส สัญชัย
0.421863	หนังสือให้ความยินยอมของกลุ่มสมรส อภิสิทธิ์
0.421875	ขอสอบประมวลความรู้ สมศักดิ์ บววป11 kmitl
0.421875	คำร้องขอ รบ. ม.ต้น สิริวัฒนา Ben
0.421875	คำร้องขอส่งวิทยานิพนธ์ล่าช้า สิริภรณ์ บวทป14 kmitl
0.421875	คำรับรองจากอาจารย์ที่ปรึกษา เคนชัย บววป2 kmitl

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.421875	แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน อมตะ
0.421988	ขอลดเงินประกันของเสียหาย2 สุทธิพงษ์ kmitl
0.421988	ข้อมูลการเรียนและความก้าวหน้าในการทำวิทยานิพนธ์ ปราณีย์ คำร้อง4-1 มอ
0.421988	คำร้องขอให้ตรวจสอบและบำรุงรักษาหม้อแปลงเฉพาะราย สมพร
0.422102	ขอลดเงินประกันของเสียหาย1 เกษราภรณ์ kmitl
0.422102	คำขอดูข้อมูลข่าวสารกระทรวงคมนาคม บัณฑิต
0.422102	คำร้องขอเปลี่ยนหัวข้อวิทยานิพนธ์ ไพลิน บววป5 kmitl
0.422113	ขอแต่งตั้ง.บัณฑิต บววป15 ปสุด kmitl
0.422113	คำร้องขอขยายระยะเวลาการศึกษา คำร้อง2-2 วนิดา มอ
0.422113	หนังสือรับรองของผู้ปกครอง 7-7 อิศรา บววป9 kmitl
0.422227	mtec sul
0.422227	ขอลดเงินประกันของเสียหาย1 สุทธิพงษ์ kmitl
0.422227	คำร้องการไฟฟ้า นิกรณ
0.422227	คำรับรองจากอาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม วัชรพล
0.422227	ใบคำร้องของมคึกงาน วัชรพล ABAC
0.422227	หนังสือให้ความยินยอมของผู้สมรส วริศรา
0.422238	ความเห็นอาจารย์ที่ปรึกษา สิริเพ็ญ มอ
0.422238	คำร้องขอเปลี่ยนชื่อ พิมลรัตน์ ben
0.422238	แบบแจ้งรายชื่อผู้เข้าอบรม ใช้เครน ชูพงษ์
0.436875	ขอทุนการศึกษา 5-7 บว วป9 kmitl
0.436875	ใบคำร้องของมคึกงาน ชิตินันท์ ABAC
0.436875	หนังสือมอบอำนาจ เจียมจรัส
0.436875	หนังสือรับรองของผู้ปกครอง 7-7 บว วป9 kmitl
0.436994	interview result karaked ABAC
0.437000	ขอทุนการศึกษา 3-7 บว วป9 วรรณดี kmitl
0.437000	คำร้องขอ รบ. ม.ต้น สุรสิต Ben
0.437000	หนังสือรับรองของผู้ปกครอง 7-7 สมชาย

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการไหลครบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

เวลาที่ใช้ในการไหลครบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.437113	การตีพิมพ์ผลงานวิทยานิพนธ์ อรรถวุฒิ วพ4 ku
0.437113	ขอลอนเงินประกันของเสียหาย2 เกษราภรณ์ kmitl
0.437113	ขอเปลี่ยน.ที่ปรึกษา ประพิศ บววป2 kmitl
0.437113	คำร้องขอปรับสภาพเป็นนักศึกษาสามัญ กิตติวิทย์ บวทบ7 kmitl
0.437238	การขอสอบวัดคุณสมบัติ ราชฤดี บววป12 kmitl
0.437238	ขอสอบประมวลความรู้ อัครพล บววป11 kmitl
0.437238	คำรับรองจาก.ผู้ควบคุมวิทยานิพนธ์ร่วม ภัสรี บววป4-2 kmitl
0.437863	การตีพิมพ์ผลงานวิทยานิพนธ์ ชลภิรา วพ4 ku
0.437863	ขอทุนการศึกษา 3-7 บว วป9 kmitl
0.437863	คำร้องขอ รบ. ม.ปลาย ศิรินาด
0.437863	คำร้องขอปรับสภาพเป็นนักศึกษาสามัญ कमเดช บวทบ7 kmitl
0.437863	คำรับรองจากอาจารย์ที่ปรึกษา อารีรัตน์ บววป2 kmitl
0.437988	NECI
0.437988	ขอลอนเงินประกันของเสียหาย1 ประพันธ์ kmitl
0.437988	คำร้องการไฟฟ้า วิษณุ
0.437988	คำร้องขอ รบ. ม.ต้น ธนาธร Ben
0.437988	คำร้องขอเปลี่ยนหัวข้อวิทยานิพนธ์ วโรดม บววป5 kmitl
0.437988	คำรับรองจาก.ผู้ควบคุมวิทยานิพนธ์ร่วม กิตติ บววป4-2 kmitl
0.437988	คำรับรองจากอาจารย์ที่ปรึกษา นันทนา บววป2 kmitl
0.437988	ศรีสมัช บววป16 kmitl
0.438102	interview result anusorn ABAC
0.438102	คำร้องขอ รบ. ม.ปลาย กิตติภพ Ben
0.438227	กรมทรัพย์สินทางปัญญา โอภาส
0.438227	คำร้องขอ รบ. ม.ต้น วลัย Ben
0.438227	คำร้องขอส่งวิทยานิพนธ์ล่าช้า อารยา บววป14 kmitl
0.438227	คำร้องขอหนังสือรับรอง kmitl
0.438227	คำร้องขอหนังสือรับรอง ชาดิษา สงล

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

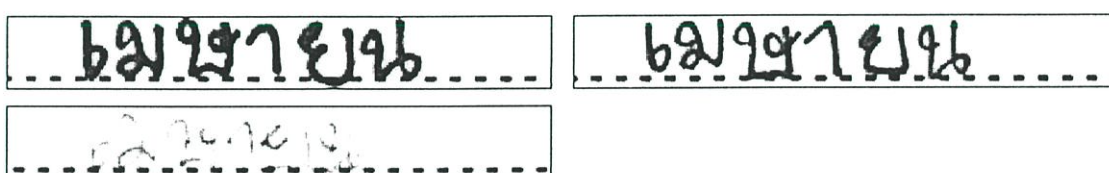
เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.438227	คำร้องขอให้ตรวจสอบและบำรุงรักษาหม้อแปลงเฉพาะราช กรรณิกา
0.438227	แบบขอสอบประมวลความรู้ เนตรนภา บววป11 kmitl
0.438227	แบบคำขอข้อมูลข่าวสารของกรมพัฒนาที่ดิน ธนพล
0.452863	คำร้องขอเปลี่ยนแปลง.ผู้ควบคุมวิทยานิพนธ์ นันทนา บววป3 kmitl
0.452863	คำร้องขอหนังสือรับรอง กุลธิดา kmitl
0.452875	ความก้าวหน้าในการทำวิทยานิพนธ์ 4-1 มอ
0.453000	interview result pompem ABAC
0.453000	ขอแต่งตั้ง.บัณฑิต บววป15 ธนาธิป kmitl
0.453000	ขอทุนการศึกษา 5-7 บว วป9 วรรณดี kmitl
0.453113	ขอทุนการศึกษา 2-7 บว วป9 วรรณดี kmitl
0.453113	ขอทุนการศึกษา 4-7 บว วป9 วรรณดี kmitl
0.453113	คำขอข้อมูลข่าวสารกระทรวงคมนาคม สุวภัทร
0.453113	หนังสือมอบอำนาจ รัตน์
0.453227	Plan-it Suchada
0.453227	ขอถอนเงินประกันของเสียหยา2 ประพันธ์ kmitl
0.453227	คำขอข้อมูลข่าวสารกระทรวงคมนาคม นิก
0.453227	หนังสือรับรองของผู้ปกครอง 7-7 ฉนากร บววป9 kmitl
0.453238	คำร้องขอส่งวิทยานิพนธ์ล่าช้า ณัฐวดี บวทป14 kmitl
0.453238	ใบคำร้องของมฝึกงาน วิภา ABAC
0.454102	ขอถอนเงินประกันของเสียหยา รัตน์ kmitl
0.454102	ข้อมูลการเรียนและความก้าวหน้าในการทำวิทยานิพนธ์ ทิพวรรณ
0.454102	คำร้องขอให้ตรวจสอบและบำรุงรักษาหม้อแปลงเฉพาะราช วิชญ
0.454102	ใบคำร้องของมฝึกงาน อลิสา
0.467875	คำร้องการไฟฟ้า กมล
0.468000	คำร้องขอหนังสือรับรอง สุทธิพงษ์ kmitl
0.468988	กรีฑาพล บววป16 Kmitl
0.469102	ประวัติ.บัณฑิตวิทยาลัย ชวลิต

ตารางที่ 4.1 (ต่อ) เวลาที่ใช้ในการ โหลดกรอบข้อความขึ้นมาโดยอัตโนมัติของแต่ละแบบฟอร์ม

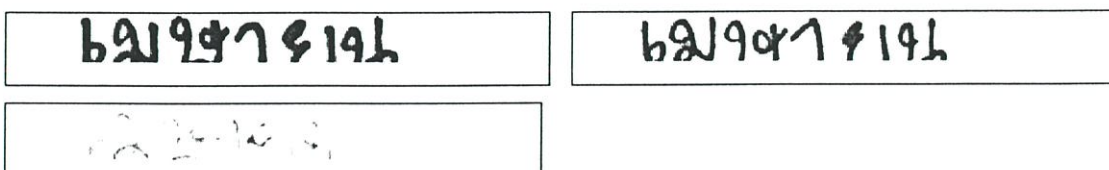
เวลาที่ใช้ในการโหลดกรอบข้อมูล(วินาที)	ชื่อไฟล์ที่ใช้ในการทดลอง
0.469238	ขอแต่งตั้ง.บัณฑิต อมรศักดิ์ บววป15 kmitl
0.483875	ขอถอนเงินประกันของเสี่ยหาข3 สุทธิพงษ์ kmitl
0.484113	การตีพิมพ์ผลงานวิทยานิพนธ์ นันทพร วพ4 ku
0.484113	ขอถอนเงินประกันของเสี่ยหาข3 เกษราภรณ์ kmitl
0.485227	คำร้องขอเปลี่ยนหัวข้อวิทยานิพนธ์ เอมอร บววป5 kmitl
0.499875	คำร้องขอเปลี่ยนชื่อ บิซอง ben
0.500113	คำร้องขอ Transcript สุทธิพงษ์ kmitl

หลังจากการโหลดกรอบบริเวณที่มีการกรอกข้อมูลโดยอัตโนมัติแล้วระบบจะนำกรอบที่ได้ในแต่ละกรอบมาบันทึกเป็นไฟล์รูปภาพชนิด *.bmp และกำหนดชื่อของภาพนั้นตามชื่อฟิลด์ที่กำหนดไว้ในฐานข้อมูล

จากนั้นระบบจะตัดเส้นบรรทัดที่พบในแต่ละกรอบข้อความด้วยหลักการทำงานที่ได้อธิบายไว้ในหัวข้อที่ 2.2 แต่ละภาพกรอบของข้อความประกอบด้วยข้อความที่มีลักษณะเส้นหนาหรือเส้นบางตามลักษณะของปากกาที่ใช้ในการกรอกข้อมูล ซึ่งค่าคงที่ที่ใช้ในการตัดเส้นบรรทัดจะมีความแตกต่างตามขนาดของเส้นที่พบในแต่ละแบบฟอร์ม ตัวอย่างภาพกรอบข้อความที่ตัดจากบางส่วน of แบบฟอร์มที่กรอกข้อมูลแล้วระบบได้บันทึกเป็นไฟล์ *.bmp แสดงในรูปที่ 4.15 และตัวอย่างภาพกรอบข้อความของตัวอักษรที่ตัดเส้นบรรทัดออกเรียบร้อยแล้วได้แสดงในรูปที่ 4.16



รูปที่ 4.15 ภาพกรอบข้อความที่ตัดจากบางส่วน of แบบฟอร์มที่กรอกข้อมูลแล้ว



รูปที่ 4.16 ภาพกรอบข้อความของตัวอักษรที่ตัดเส้นบรรทัดออก

จากรูปที่ 4.16 พบว่าเส้นบรรทัดที่ตัดออกของตัวอักษรที่เขียนด้วยเส้นขนาดบาง ๆ ตัวอักษรจะถูกตัดออกไปน้อยมาก แต่ตัวอักษรที่เขียนด้วยเส้นที่มีความหนามาก ตัวอักษรจะถูกตัดออกไปบางส่วน ซึ่งเป็นปัญหาต่อการรู้จำตัวอักษร แต่หากค่าคงที่ที่ใช้ในการตัดเส้นบรรทัดเปลี่ยนไปจะทำให้ประสิทธิภาพในการตัดเส้นบรรทัดเปลี่ยนแปลงตามไปด้วย ปัญหาเส้นของตัวอักษรขาดไปบางส่วนจะมีผลทำให้การรู้จำตัวอักษรไม่สามารถทำได้

ต่อจากนั้นระบบจะส่งภาพกรอบที่กรอกข้อมูลของแบบฟอร์มที่ผ่านกระบวนการลบเส้นบรรทัดแล้ว ไปสู่กระบวนการจัดการก่อนเข้าสู่กระบวนการรู้จำตัวอักษรซึ่งเป็นกระบวนการกำจัดสัญญาณรบกวน (Noise Removal) และปรับเอียงให้อยู่ในแนวตรง (Skew Correction) ก่อนที่จะนำภาพในกรอบข้อความที่กำหนดไว้ไปสู่กระบวนการรู้จำซึ่งใช้เทคนิคการรู้จำโดยการแบ่งครึ่งตัวอักษรและลักษณะเด่นของตัวอักษร

สุดท้ายระบบจะนำข้อมูลที่สามารรู้จำได้มาตรวจสอบความถูกต้องกับประเภทและขอบเขตข้อมูลที่ผู้ใช้ได้กำหนดไว้ล่วงหน้า ก่อนที่จะบันทึกลงในฐานข้อมูล

บทที่ 5

สรุปผลการทดลองและข้อเสนอแนะ

จากผลการทดลองในงานวิจัยนี้สามารถจัดการแบบฟอร์มที่มีการกรอกข้อมูลให้สามารถนำไปใช้ในกระบวนการรู้จำตัวอักษรภาษาไทยได้ และยังสามารถตรวจสอบความถูกต้องของข้อมูลก่อนที่จะมีการเก็บข้อมูลลงในฐานข้อมูลตามที่กำหนด

5.1 สรุปผลการทดลอง

ข้อมูลแบบฟอร์มเปล่าที่นำมาใช้ในงานวิจัยนี้เป็นแบบฟอร์มที่ได้มาจากหน่วยงานเอกชน อาทิ เช่น บริษัท, มหาวิทยาลัยเอกชน ฯลฯ และหน่วยงานของรัฐ อาทิเช่น มหาวิทยาลัย, โรงเรียน, กระทรวง, การไฟฟ้าส่วนภูมิภาค, โรงพยาบาล ฯลฯ

แบบฟอร์มเปล่าที่นำมาใช้ในงานวิจัยนี้มีประมาณ 120 แบบฟอร์ม โดยแบบฟอร์มที่ใช้มักเป็นแบบฟอร์มการสมัครเข้าร่วมการสัมมนา ใบคำร้องต่าง ๆ ของแต่ละหน่วยงาน หรือใบมอบอำนาจ งานวิจัยนี้ใช้คนที่กรอกแบบฟอร์มมีอายุอยู่ในช่วง 10-60 ปี จำนวน 4 คน กรอกแบบฟอร์มประมาณ 300 แบบฟอร์ม

ระบบการจัดการแบบฟอร์มเพื่อการจำแนกตัวอักษรภาษาไทยที่กล่าวมาแล้ว ทำหน้าที่จัดการและเตรียมการแบบฟอร์มที่มีการกรอกข้อมูลแล้ว ให้สามารถผ่านเข้าสู่กระบวนการจำแนกตัวอักษรได้อย่างมีประสิทธิภาพ โดยสามารถกำหนดกรอบบริเวณที่มีการกรอกข้อมูลได้โดยอัตโนมัติ และยังให้ผู้ใช้สามารถเปลี่ยนแปลงชนิดของข้อมูลในแต่ละฟิลด์ หรือกำหนดขอบเขตข้อมูลได้ตามความเหมาะสมของงานแต่ละประเภท

เวลาที่ใช้ในการกำหนดกรอบข้อมูลในแต่ละแบบฟอร์ม จะใช้เวลาไม่เท่ากัน ระยะเวลาที่ใช้จะมีความยาวนานเพียงใดขึ้นอยู่กับความซับซ้อนของแบบฟอร์ม แต่เวลาที่ใช้ในการโหลดกรอบข้อมูลโดยอัตโนมัติหลังจากที่ทราบว่าเป็นแบบฟอร์มที่มีการกรอกข้อมูลแล้วตรงกับแบบฟอร์มเปล่าใดมากที่สุด จะใช้ระยะเวลาไม่เกิน 1 วินาที ดังที่ได้แสดงไว้ในตารางที่ 4.1 ในส่วนของบทที่ 4 ที่ได้กล่าวมาแล้ว

ในแบบฟอร์มที่กรอกข้อมูลแล้วมีบางแบบฟอร์มที่ผู้กรอกข้อมูลเขียนด้วยลายมือที่มีขนาดใหญ่กว่าขนาดตัวอักษรปกติของแบบฟอร์มมาก ๆ หรือเขียนข้อความล้นไปในส่วนระหว่างบรรทัด จะมีผลทำให้ข้อความในบริเวณที่เป็นกรอบข้อมูลไม่สามารถรู้จำตัวอักษรเหล่านั้น ได้ทั้งหมด

ลายมือที่กรอกข้อความเพื่อใช้ในงานวิจัยนี้จะต้องเขียนด้วยลายมือบรรจงและเขียนด้วยตัวตรงไม่เอียงไปมา เพื่อความสะดวกในการนำข้อมูลไปใช้ในส่วนของการรู้จำตัวอักษรต่อไป

งานวิจัยนี้สามารถใช้จัดการแบบฟอร์มภาษาไทย เพื่อให้สามารถนำส่วนที่เป็นกรอบบริเวณที่มีการกรอกข้อมูล ส่งต่อไปยังส่วนของการรู้จำตัวอักษรซึ่งเป็นโปรแกรมที่อยู่นอกเหนือจากขอบเขตงานวิจัย และระบบในงานวิจัยนี้สามารถพัฒนาให้มีความสามารถจัดการแบบฟอร์มที่เป็นภาษาอังกฤษได้ แต่ระบบในงานวิจัยนี้ยังไม่มีความสามารถในการรู้จำตัวอักษรทั้งภาษาไทยและภาษาอังกฤษ เนื่องจากอยู่นอกเหนือขอบเขตของงานวิจัย

5.2 ข้อเสนอแนะและแนวทางพัฒนาต่อในอนาคต

จากผลงานวิจัยนี้ เมื่อสามารถกำหนดกรอบบริเวณที่ต้องการกรอกข้อมูลได้โดยอัตโนมัติ ปัญหาที่พบในงานวิจัยนี้คือ

- ภาพที่ส่งต่อไปยัง โปรแกรมรู้จำตัวอักษร ยังมีปัญหาเรื่องตัวอักษรขาดหายไปบางส่วนที่เกิดจากการตัดเส้นบรรทัดที่ซ้อนทับกับข้อความที่เขียนลงไปแบบฟอร์ม

งานวิจัยที่สามารถพัฒนาต่อไปในอนาคตคือ

- พัฒนาให้สามารถใช้ในแบบฟอร์มที่มีหลายภาษาในแบบฟอร์มเดียวกันได้
- อาจมีการนำงานวิจัยไปพัฒนาต่อในแบบฟอร์มของภาษาต่างประเทศอื่น ๆ
- สามารถนำไปพัฒนาส่วนของประเภทข้อมูลหรือขอบเขตของข้อมูลให้มีความหลากหลายยิ่งขึ้น

เอกสารอ้างอิง

- [1] สมหมาย บัวแย้มแสง, “การรู้จำตัวพิมพ์อักษรไทยโดยการพิจารณาโครงสร้างของขอบตัวอักษร” *การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 22*, กรุงเทพฯ, 2542, หน้า 143-146.
- [2] กฤษฎา วิไลลักษณ์, “การรู้จำลายมือเขียนภาษาไทยในแบบฟอร์ม” *วิทยานิพนธ์, มหาวิทยาลัยขอนแก่น*, 2541.
- [3] กฤษฎา วิไลลักษณ์, วิเชียร เปรมชัยสวัสดิ์ และ นุชรี เปรมชัยสวัสดิ์ “การรู้จำอักษรไทยโดยใช้ลักษณะเด่นของตัวอักษร” *การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 21*, กรุงเทพฯ, 2541, หน้า 90-93.
- [4] กฤษฎา ลิ้มปานานท์ และ ดร. โกสินทร์ จ่านงไทย “วิธีการรู้จำลายมือเขียนภาษาไทยโดยใช้กรอบมาตรฐาน” *การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 16*, กรุงเทพฯ, 2536, หน้า 630-635.
- [5] Y. Zheng, C. Liu, and X. Ding, “Form Frame Line Removal with Line Width Threshold Approach,” *Pattern Recognition and Artificial Intelligence (Chinese)*, Vol. 14, No. 2, June 2001, pp. 210-214.
- [6] Surapan AIRPHAIBOON and Shozo KONDO, “Off-line Handwritten Thai Characters Recognition from Word Script,” *Proceedings of the RESTECS'96*, Bangkok, 1996, pp. C55-C59.
- [7] Wicha Panich, Somchai Jitapunlul and Prasert Choruengwiwat, “Segmentation Of Connected Characters Using Distinctive Features Of Thai Characters In Thai Character Recognition System,” *20th Electrical Engineering Conference*, Bangkok, 1997, pp. 338-342.
- [8] Suzanne Liebowitz Taylor, Richard Fritson, and Jon A. Pastor, “Extraction of Data from Preprinted Forms,” *Machine Vision and Applications*, Vol. 5, 1992, pp. 211-222.
- [9] Supanya Aphiwongsophon, Wichian Premchaiswadi and Nucharee Premchaiswadi, “Forms Management for Thai Character Recognition”, *In Proceedings of the 2003 International Conference on Information and Communication Technologies (ICT 2003)*, Bangkok, 2003, pp.366-372.

ภาคผนวก

บทความและผลงานวิจัยที่ได้รับการตีพิมพ์

บทความที่ตีพิมพ์คือ

Supanya Aphiwongsophon, Wichian Premchaiswadi and Nucharee Premchaiswadi, "Forms Management for Thai Character Recognition," *In Proceedings of the 2003 International Conference on Information and Communication Technologies (ICT 2003)*, Bangkok, 2003, pp.366-372.

Forms Management for Thai Character Recognition

Supanya Aphiwongsophon and Wichian Premchaiswadi
*Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang
Chalongkrung Road, Ladkrabang Bangkok 10520, Thailand*

Nucharee Premchaiswadi
*Faculty of Information Technology, Dhurakijpundit University
110/1-4 Prachacheun Road, Laksi, Bangkok 10210, Thailand*

Abstract

This research presents forms management technique for Thai character recognition. The recognition areas can be marked automatically by the system or by users. The user can define description of fields such as field name and its data type to be recorded in the database. Characters in the marked area are passed to the recognition process. The result of the recognition process is verified and stored in the database automatically for later use.

Keywords: Automatic mark, Automatic load mark, Relative reference position, Verify data.

1. Introduction

Usually form identification is used to identify the used form automatically. However, the boundary for recognition is defined by users. Therefore, if the process can be processed automatically, it would be benefit for the practical used. This research presents the automatic marking process to define the boundary of data to be recognized and verify the recognition data before update to database. The objective is reduce the time to define data boundary to be recognized. The marking area is not only useful for recognition process, but also for verify data to be input into database as well.

User or system can select a marking area and assign field name together with its data type to be recorded in a database. The data type is used to reduce the scope of the recognition process and verify the recognition result before recorded into database.

Automatic mark can define the boundary of data to be recognition and the data type such as text, number, and optional answer (yes/no).

The overall process of the proposed system is shown in Fig. 1.

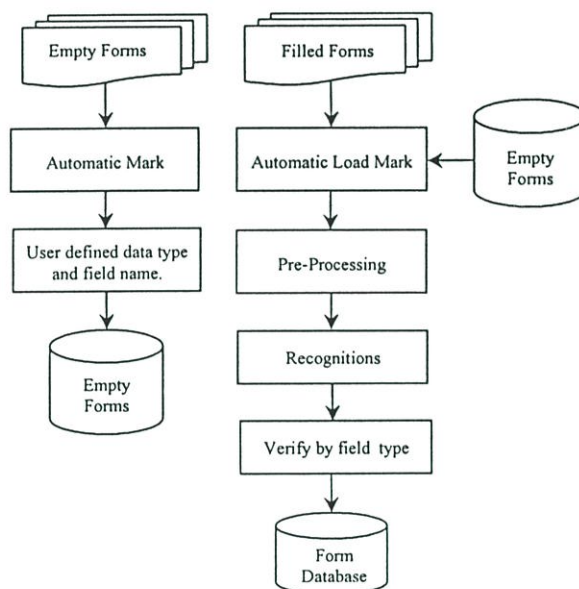


Figure 1. The steps in forms management for Thai character recognition.

2. Automatic mark and user define type

In this article we describe the processing steps of automatic mark that find the mark position which desire to save automatically. The first step is to find horizontal projection of form image for separate character line. Next the system uses vertical projection in character line for search filled data area to mark.

The picture illustrates the empty form and the horizontal projection that use to separate character lines shown in Figure 2.



Figure 2. An example of empty forms with horizontal projection.

For example, the horizontal projection in Figure 2 is separate character line and the result is shown in Figure 3.

In empty form, automatic mark defines the mark area that composes straight line mark or optional mark for fill data.

The straight line mark is dot, dash, thick line of fill data for character or number. These cases consider the black pixel area that ratio between width per height is very different (width is greater than height). An example of dot, dash and thick line are shown in Figures 4-1, 5-1, and 6-1. The result of automatic mark from Figure 4-1, 5-1 and 6-1 are shown in Figures 4-2, 5-2, and 6-2.

The optional mark is ○, □, (), [] for sign in. Normally the optional mark usually locate at begin of line and have more space before text describe. Optional mark's characteristic is peak of vertical projection. The square and circle mark have a pair of peak projection on the both side. Distinguish between square mark and circle mark is circle mark's projection has curve but square mark's projection has a right angle.

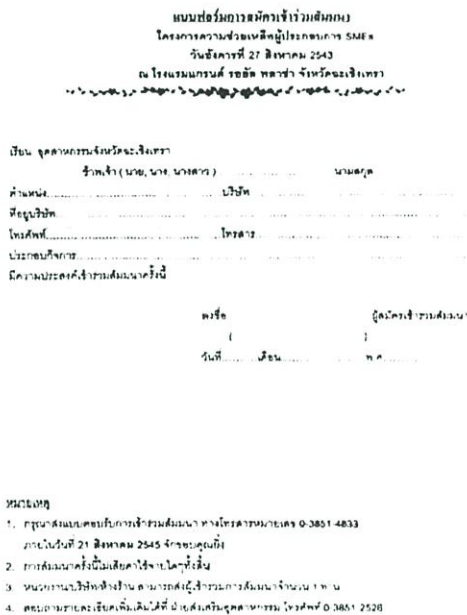


Figure 3. The result of separates character line in empty form.



Figure 5-1. Dash line with vertical projection.

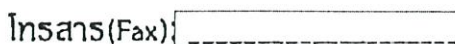


Figure 5-2. The filled data position and area of Figure 5-1.



Figure 6-1. Thick line with vertical projection.

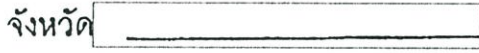


Figure 6-2. The filled data position and area of Figure 6-1.

In vertical projection of parenthesis or blanket show narrow period of height frequency then follow character space with empty frequency and narrow period of height frequency at the end.

An example of circle, square, parenthesis and blanket mark are shown in Figures 7-1, 8-1, 9-1 and 10-1. The result of automatic mark from Figures 7-1, 8-1, 9-1 and 10-1 are shown in Figures 7-2, 8-2, 9-2 and 10-2.

- 0-5,000 Baht
- 5,001-10,000 Baht
-
-

Figures 7-1. Circle mark with vertical projection.

- 0-5,000 Baht
- 5,001-10,000 Baht

Figure 7-2. The filled data position and area of Figure 7-1.

- Newspaper Internet
-

Figures 8-1. Square mark vertical projection.

- Newspaper Internet

Figure 8-2. The filled data position and area of Figure 8-1.

- 20 - 50 ล้านบาท
-

Figure 9-1. Parentheses mark with vertical projection.

- 20 - 50 ล้านบาท

Figure 9-2. The filled data position and area of Figure 9-1.

- คณะเทคโนโลยีสารสนเทศ
-

Figure 10-1. Blanket mark with vertical projection.



Figure 10-2. The filled data position and area of Figure 10-1.

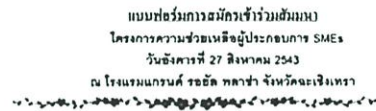
In optional mark has two values are space and true or false mark (✓,×). The space in blank that mean nothing or it doesn't have value. The true or false marks that mean have value.

The mark in fill form and empty form will match with relative reference position. The relative position use relative distance that defines by the Euclidean distance formula between the geometric position of the beginning line in form and the interest line

$$d(b_i, b_j) = (x_i - x_j)^2 + (y_i - y_j)^2 \quad (1)$$

where $d(b_i, b_j)$ is distance between the beginning line $b_i: (x_i, y_i)$ and the interest line $b_j: (x_j, y_j)$.

Automatic mark process of empty form that shown the frame of filled data position and area. From Figure 3 uses automatic mark to define the frame of filled data position and area as shown in Figure 11.



เป็นบุคลากรประจำหน่วยงาน
 ชื่อ (นาม, นามสกุล) : นามสกุล
 ตำแหน่ง : บริษัท :
 ที่อยู่บริษัท :
 โทรศัพท์ : โทรสาร :
 ครอบคลุมกิจการ :
 มีความประสงค์เข้าร่วมแข่งขันหรือไม่

ลงชื่อ : ผู้สมัครเข้าร่วมแข่งขัน
 ()
 วันที่ เดือน พ.ศ.

- หมายเหตุ
- กรุณาส่งมอบใบการเข้าร่วมแข่งขันมา ทางโทรสารหมายเลข 0-3851-4833 ภายในวันที่ 21 สิงหาคม 2545 ถึงรอบสมัคร
 - การสมัครครั้งนี้ไม่เสียค่าใช้จ่ายใดๆทั้งสิ้น
 - หน่วยงานบริษัทข้างต้น สามารถส่งผู้เข้าร่วมการแข่งขันจำนวน 1 ท่าน
 - สอบถามรายละเอียดเพิ่มเติมได้ที่ ฝ่ายส่งเสริมอุตสาหกรรม โทรศัพท์ 0-3851-2526

Figure 11. An empty form that automatic marks to find filled data position and area.

When the system gets the position and area of mark, the mark area define the data field's name, data type, range of data (optional) and database's name that sets by user. In the optional mark, the system defines field type as option. The straight line

Table 3. The character's set of vowel consonant and number with segment at half vertical in upper and middle zone.

# up	# low	Character	Vowel	Consonant	Num
1	1		๑๑๑		
1	2	๓			
2	1	๗๗๗๗๗๗			
2	2	๓			
3	2	๗๗			

Table 4. The character's set of vowel consonant and number with segment at half vertical in lower zone.

# up	# low	Character	Vowel	Consonant	Num
1	1		,		
2	1		,		

Table 5. The character's set of vowel consonant and number with segment at half vertical in middle and lower zone.

# u	# low	Character	Vowel	Consonant	Num
1	1		๑		
1	2	๓๓๓๓๓			
2	1	๓			
2	2	๗			

4.2 Feature Extraction

The set of character is decided by the character's head (clockwise head [q, b], counterclockwise head [p, d] or no head), character's curl, structure of write (up, down, left, right) and the ratio of width and height that is illustrated in Tables 6 to 10.

Table 6. Feature of one up piece one low piece.

Level	Character	Feature
Middle	๓	up piece has q head
	๑	no head and up piece is curve
	1	no head and width/height < 0.4
	7	no head and width/height > 0.4
	2	up piece is curve and low piece write to left-right
	4	up piece write to up-down
	5	up piece write to left-right and low piece is curve
	3	up piece is curve, left side is higher than right side, low piece is curve and left side is shorter than right

		side
	0	up piece is curve, left side is equal right side and low piece is curve, left side is equal right side
	8	up and lower piece is close circle
	๗	up piece write to left-right and low piece has d head
	๑	up piece write to left and low piece has d head
	๑	low piece has b head and width/height < 0.4
	6	low piece has b head and width/height > 0.4
	9	up piece has q head and low piece is curve
Up Middle	๑	up piece write to left-right and low piece has b head
	๑	up piece is curve and low piece has b head
	๑	up piece write to up-down and low piece has b head
Up	๓	up and low piece is curve
	๓	up curve and low curve
	๓	vertical straight line
	๓	change direction in vertical and horizontal
	๓	p head
	๓	q head
	๓	b head
Low	๓	q head

Table 7. Feature of character level that is one up piece and two low pieces.

Level	Character	Feature
Middle	๓	no head
	๓	low piece has b head
	๓	low piece has d head
Up Middle	๓	low piece has b head
Middle Low	๓	low piece has d head and a curl
	๓	low piece has d head and 2 curl
	๑	low piece has b head
	๑	low piece has d head

Table 8. Feature of character level that is two up pieces and one low piece.

Level	Character	Feature
Middle Middle	ก	up piece has q head and width/height < 0.4
	ข	up piece has q head and width/height > 0.4
	ค	up piece has p head
	ด	up piece has q head and right is greater than u
	จ	up piece has q head and low piece has b head
	ฉ	up piece has q and curl head low piece has d head
	ช	up piece has q head and curve
	ฉ	up piece has q head and low piece has b head
	ค	up piece has p head
	ด	no head
	น	up piece has q head and low piece has b head
	ม	up piece has q head and low piece has d head
	ผ	up piece has p head and low piece is curl
	ข	p head and two piece
Up Middle	ก	up piece has p head
	ข	up piece has q and curl head
	ป	up piece has q head and width/height (left) > 0.4
	ผ	up piece has q head and low piece is a curl
Up	ค	up and low piece have curve and a vertical line
	ด	up and low piece have curve
	จ	up and low piece have curve and two vertical line
	ข	q head in left hand
Low	ก	low piece has q head

Table 9. Feature of character level that is two up pieces and two low pieces.

Level	Character	Feature
Middle	ค	up piece has p head
	ด	up piece has p head and curl
	ค	up piece has q head
	ด	up piece has q head and curl
	ฉ	low piece has b head and middle is d head
	ญ	low piece has b head

	ณ	low piece has b head and right is b head
	ล	low piece has b head and right is straight line
	ท	up piece left q head
Up Middle	ห	up piece left q head and right is p head
	ศ	up piece has p head
Up	ฃ	low piece has b head and curl

Table 10. Feature of character level that is three up pieces and two low pieces.

Level	Character	Feature
Middle	ฅ	up piece has q head and curl
	พ	up piece has q head and low piece has a curl
Up Middle	ฟ	up piece has q head
	ฟ	up piece has q head right side has greater than ฟ

5. Data verification

Verify recognition is post process after character recognition. Field type can be help to reduce scope of character probability and use to correct character recognition.

For example: The number type of database are age (range in 1-150), date (range in 1-31), year (range in 1900-2200 and 2400-2600). The range can change by user define. In table 11 illustrate the character level and the number of character in each Thai's month.

Table 11. Verify Thai months with level and number of character.

Character level	Thai month	Number of characters
Center	มกราคม	6
	เมษายน	6
Upper + Center	มีนาคม	6
	สิงหาคม	7
	กันยายน	7
	ธันวาคม	7
Center + Lower	ตุลาคม	6
	พฤษภาคม	7
	กรกฎาคม	7
Upper + Center + Lower	มิถุนายน	8
	พฤศจิกายน	9
	กุมภาพันธ์	10

The text types are province (Thailand's provinces are 76 names in database), month (The year has 12 months namely January-December or month abbreviation form or the number 1-12).

After the character verification is complete, the process finishes and saves the information into desired databases.

6. Experimental results

The system is implemented by using Microsoft Visual Basic 6.0. The forms are scanned into the system with a resolution of 200 dpi and saved in binary format. The empty forms that are used in this research are about 80 forms and the filled forms are written by five different persons.

The system can define automatic mark in each empty form within 1 second. User defines data type for database record and range of each data.

When the system receives the filled form, the system uses automatic load mark in filled form. Then, the mark area is used for recognition by the method described previously. The last process verifies the correctness of the data before to update the database.

7. Conclusion

Form management for Thai character recognition can automatically define the boundary of filled data

area. The user sets data type for verification and reduces the scope of data recognition.

The experimental results show that the system can mark data boundary correctly within a reasonable time.

References

- [1] Krisada Wililak, "Thai character handwritten recognition in form", Thesis, Khon Kaen University, 1998.
- [2] Krisada Wililak, Wichian Premchaiswadi, and Nucharee Premchaiswadi, "Thai character handwritten recognition by character feature", 21st Electrical Engineering Conference, Bangkok, 1998, pp. 90-93.
- [3] Krisada Limpanon and Kosin Chamnongthai, "Thai character in handwritten method by use of standard frame", 16th Electrical Engineering Conference, Bangkok, 1993, pp. 630-635.
- [4] Surapan Airphaiboon and Shozo Kondo, "Off-line Handwritten Thai Characters Recognition from Word Script" Proceedings of the RESTECS'96, Bangkok, 1996, pp. C55-C59.
- [5] Wicha Panich, Somchai Jitapunlul, and Prasert Choruengwiwat "Segmentation of Connected Characters using Distinctive Features of Thai Characters in Thai Character Recognition System" 20th Electrical Engineering Conference, Bangkok, 1997, pp. 338-342.

ประวัติผู้เขียน

ชื่อผู้เขียน	นางสาวสุปัญญา อภิวงศ์โสภณ
วันเดือนปีเกิด	วันที่ 4 เดือนพฤษภาคม พ.ศ. 2522
สถานที่เกิด	จังหวัดกรุงเทพมหานคร
การศึกษาระดับปริญญาตรี	วิทยาศาสตร์บัณฑิต (วิทยาการคอมพิวเตอร์) ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ ปีการศึกษา 2541
ผลงานด้านวิชาการ	Supanya Aphiwongsophon, Wichian Premchaiswadi and Nucharee Premchaiswadi, "Forms Management for Thai Character Recognition", <i>In Proceedings of the 2003 International Conference on Information and Communication Technologies (ICT 2003)</i> , Bangkok, 2003, pp.366-372.