

บทที่ 3

วิธีดำเนินการวิจัย

ในการดำเนินงานวิจัยประกอบด้วย 6 ส่วน คือ 1. องค์ประกอบด้านซอฟต์แวร์ 2. ขั้นตอนการทำงานของกระบวนการผลขั้นต้นและการรู้จำ 3. ขั้นตอนการประมวลผลขั้นปลายด้วยเทคนิคแอลซีเอส 4. เทคนิคและการประยุกต์ 5. การออกแบบการทดลองโปรแกรม 6. แผนผังการดำเนินงานและขั้นตอนการดำเนินงานวิจัย

องค์ประกอบด้านซอฟต์แวร์

องค์ประกอบด้านซอฟต์แวร์ที่ใช้ในงานวิจัยประกอบด้วย

1. Eclipse Kepler
2. Microsoft Windows 7 (64 bit)

ขั้นตอนการทำงานของกระบวนการผลขั้นต้นและการรู้จำ

ก่อนที่จะนำชุดตัวอักษรที่ได้จากกระบวนการรู้จำมาเปรียบเทียบและวิเคราะห์ความเหมือนรวมถึงการเรียงตัวของสายอักขระ โดยในส่วนการประมวลผลขั้นต้นและการรู้จำไม่ได้อยู่ในขอบเขตของงานวิจัย ซึ่งผู้วิจัยจะพูดถึงขั้นตอนพื้นฐานของการได้มาซึ่งชุดอักขระจากการประมวลผลขั้นต้นและการรู้จำ ประกอบด้วยขั้นตอนดังนี้

1. การประมวลผลขั้นต้น

1.1 การกรองภาพ (Noise Filtering) เป็นการลดสิ่งรบกวนภายในภาพก่อนที่จะนำมาประมวลผล โดยนำภาพของป้ายทะเบียนรถยนต์มาทำการแปลงให้เป็นภาพแบบเกรย์สเกล (Gray Scale Image) และแปลงให้เป็นภาพขาวดำ (Thresholding) เพื่อลดสิ่งรบกวนภายในภาพ



ภาพที่ 3.1 ขั้นตอนการการกรองภาพ (Noise Filtering)

1.2 การตัดแบ่งภาพ (Cropping) และการปรับขนาดภาพ (Resize)

1.2.1 การตัดโครงสร้างที่ไม่เกี่ยวข้องกับป้ายทะเบียนออก เช่น ตัวรถ ถนน เป็นต้น เพื่อให้เหลือเฉพาะภาพป้ายทะเบียนรถยนต์

1.2.2 ปรับขนาดภาพป้ายทะเบียนให้มีอัตราส่วน 34:15 โดยมีอัตราส่วนเท่ากับภาพที่ใช้ในกระบวนการรู้จำ เพื่อความถูกต้องในการวิเคราะห์ชื่อจังหวัดจากกระบวนการรู้จำ

1.2.3 การตัดภาพแบ่งให้เหลือเฉพาะภาพส่วนที่เป็นชื่อจังหวัดเพื่อนำไปวิเคราะห์ตัวอักษรในส่วนการรู้จำ

Thresholding



↓
Crop
(Image License plate)



↓
Resize (34:15)



↓
Crop
(Image Province Name)

กรุงเทพมหานคร

ภาพที่ 3.2 ขั้นตอนการตัดแบ่งภาพ และการปรับขนาดภาพ

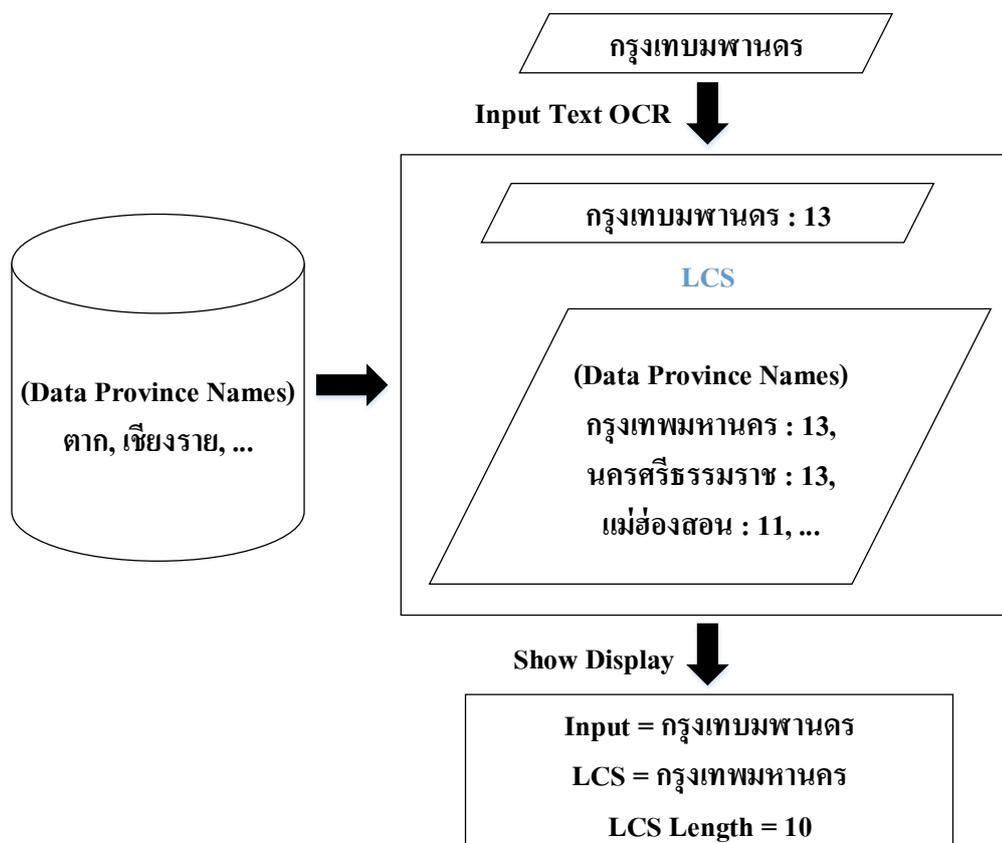
ขั้นตอนการประมวลผลขั้นปลายด้วยเทคนิคแอลซีเอส

การประมวลผลขั้นปลายโดยใช้การโปรแกรมเชิงพลวัตด้วยเทคนิคแอลซีเอส เพื่อใช้ในการเปรียบเทียบและวิเคราะห์ความเหมือนรวมถึงการเรียงตัวของสายอักขระของชื่อจังหวัดบนป้ายทะเบียนรถยนต์ สามารถแบ่งการทำงานออกเป็น 3 ขั้นตอน ดังนี้

1. การกรองชื่อจังหวัดด้วยความยาวของอักขระที่ได้จากกระบวนการรู้จำ เพื่อเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัด เนื่องจากชุดอักขระจากอักขระที่ได้จากกระบวนการรู้จำอาจมีจำนวนอักขระที่อ่านจากป้ายทะเบียนรถยนต์ไม่ครบหรืออ่านได้เกินกว่าจำนวนอักขระในชื่อจังหวัดจริง เมื่อไม่มีการกำหนดเกณฑ์ความยาวของอักขระทำให้เกิดความผิดพลาดของการวิเคราะห์ความเหมือนของชื่อจังหวัด ซึ่งส่งผลต่อประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัด อีกทั้งการกรองชื่อจังหวัดยังเป็นการลดปัญหาความยาวของลำดับย่อยร่วมยาวสุดที่เท่ากัน ซึ่งทำให้เกิดความผิดพลาดของการอ่านชื่อจังหวัดอีกเช่นกัน โดยการกรองชื่อจังหวัดจากจำนวนของอักขระที่ใกล้เคียงกับฐานข้อมูลชื่อจังหวัดที่ถูกต้อง ซึ่งค่าที่ใช้เป็นเกณฑ์บอกความยาวของชื่อจังหวัดที่ใกล้เคียงในงานวิจัยนี้จะใช้ค่า $\pm \varepsilon = \{1, 2\}$ เท่านั้น ตัวอย่างเช่น อักขระที่ได้จากกระบวนการรู้จำ คือ “บระนศรศรียุธา” มีจำนวนอักขระทั้งหมด 15 ตัวอักษร ทำให้ต้องวิเคราะห์ความเหมือนของชื่อจังหวัดที่มีความยาวของอักขระใกล้เคียงเพียงแค่ 4 ชื่อจังหวัดเท่านั้น คือ “กรุงเทพมหานคร”, “นครศรีธรรมราช”, “ประจวบคีรีขันธ์” และ “พระนครศรีอยุธยา” โดยไม่จำเป็นต้องวิเคราะห์ความเหมือนกับชื่อจังหวัดทั้งหมด

2. การวิเคราะห์ความเหมือนรวมถึงการเรียงตัวของสายอักขระเพื่อหาลำดับย่อยร่วมยาวสุดของชื่อจังหวัดจากกระบวนการรู้จำและชื่อจังหวัดที่ถูกต้องในฐานข้อมูลที่ได้จากกระบวนการกรองชื่อจังหวัดในขั้นตอนแรก

3. แสดงอักขระจากกระบวนการรู้จำ และแสดงผลพัทธ์ชื่อจังหวัดที่มีความเหมือนมากที่สุด รวมถึงแสดงความยาวของลำดับย่อยร่วมยาวสุดของคำตอบ



ภาพที่ 3.5 ขั้นตอนการประมวลผลขั้นปลายด้วยเทคนิคแอลซีเอส

เทคนิคและการประยุกต์

ผู้วิจัยได้นำเทคนิค LCS มาประยุกต์ใช้กับงานวิจัยโดยมีขั้นตอนดังนี้

1. นิยามแบบจำลองปัญหา

นิยามแบบจำลองปัญหา โดยกำหนดตัวแปรใน Algorithm 2 ให้ P_i เป็นชื่อจังหวัดในเซต P ตัวอย่างเช่น $P = \{\text{ตาก, กรุงเทพมหานคร, เชียงใหม่, ...}\}$ โดยที่ l_i เป็นจำนวนอักขระของชื่อจังหวัด P_i และชื่อจังหวัดในเซต P จะถูกเปรียบเทียบกับชื่อจังหวัด p_{ocr} ที่ถูกอ่านมาจากป้ายทะเบียนรถยนต์ โดยมีความยาวเท่ากับ l_{ocr}

2. การประยุกต์เทคนิค LCS

ก่อนการประยุกต์เทคนิค LCS เพื่อหาคำตอบนั้นผู้วิจัยได้เพิ่มขั้นตอนการกรองความยาวชื่อจังหวัดด้วยจำนวนของอักขระที่ใกล้เคียงกับฐานข้อมูลชื่อจังหวัดที่ถูกต้อง ก่อนที่นำมาวิเคราะห์เพื่อหาความเหมือนรวมถึงการเรียงตัวของสายอักขระ ซึ่งค่าที่ใช้เป็นเกณฑ์บอกความยาวของชื่อจังหวัดที่ใกล้เคียงคือ $\pm \varepsilon$ ในงานวิจัยนี้เราใช้ค่า $\pm \varepsilon = \{1, 2\}$ เท่านั้น ในขั้นตอนนี้สามารถเขียนเป็นอัลกอริทึมได้ดัง ภาพที่ 3.6

Algorithm ALCS (The Applied LCS)Input: P_{ocr} Output: $P_{\max(LCS_p)}$

- 1: **For** $P_i \in P$ **do**
- 2: **If** $l_{ocr} > l_i - \varepsilon$ **and** $l_{ocr} < l_i + \varepsilon$ **then**
- 3: $P_{result} = P_{result} \cup p_i$
- 4: $LCS_p = LCS_p \cup lcs(p_i)$
- 5: **Return** $P_{Max(LCS_p)}$

ภาพที่ 3.6 Algorithm ALCS

จากภาพที่ 3.6 Algorithm ALCS (The Applied LCS) โดยกำหนดให้ Input คือ ชื่อจังหวัดที่อ่านมาจากป้ายทะเบียนรถยนต์ด้วยกระบวนการรู้จำ (OCR) และ Output คือ ชื่อจังหวัดที่มีความเหมือนกับชื่อจังหวัดในฐานข้อมูลมากที่สุด เริ่มจากเมื่อ p_i เป็นชื่อจังหวัดในเซต P และพบว่าชื่อจังหวัด p_i ที่มีความยาวในเกณฑ์ที่กำหนดจะถูกนำไปใส่ไว้ในเซต P_{result} แล้วนำไปหาค่าความเหมือนรวมถึงการเรียงตัวของอักขระด้วยฟังก์ชัน $lcs(p_i)$ และเก็บไว้ในเซตของ LCS_p โดยแสดงผลลัพธ์ของชื่อจังหวัดที่มีความเหมือนที่มากที่สุด $Max(LCS_p)$ ซึ่งเป็นคำตอบของชื่อจังหวัดที่อ่านมาจากป้ายทะเบียนรถยนต์ด้วยกระบวนการรู้จำ

การออกแบบการทดลองโปรแกรม

การออกแบบการทดลองโปรแกรมเพื่อหาประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์ด้วยการรู้จำอักขระทางภาพจากเทคนิคแอลซีเอส ในการทดลองจะใช้ชื่อจังหวัดทั้งหมดบนป้ายทะเบียนรถยนต์ตามกฎกระทรวง กำหนดขนาด ลักษณะ และสีของแผ่นป้ายทะเบียนรถยนต์ พ.ศ. 2554 จำนวนทั้งหมด 78 ชื่อจังหวัด โดยการทดลองจะจำลองความผิดพลาดของอักขระที่ได้จากกระบวนการรู้จำ ทำการทดลอง 10 ครั้งต่อ 1 ช่วงความผิดพลาดของอักขระที่กำหนดและแบ่งการทดลองด้วยตำแหน่งความผิดพลาดและเปอร์เซ็นต์ความผิดพลาดของอักขระออกเป็น 4 แบบคือ

1. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนหน้า
2. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนกลาง
3. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนท้าย
4. ตำแหน่งของความผิดพลาดของตัวอักษรในชื่อจังหวัดแบบสุ่ม

นิยามตำแหน่งความผิดพลาดของตัวอักษร

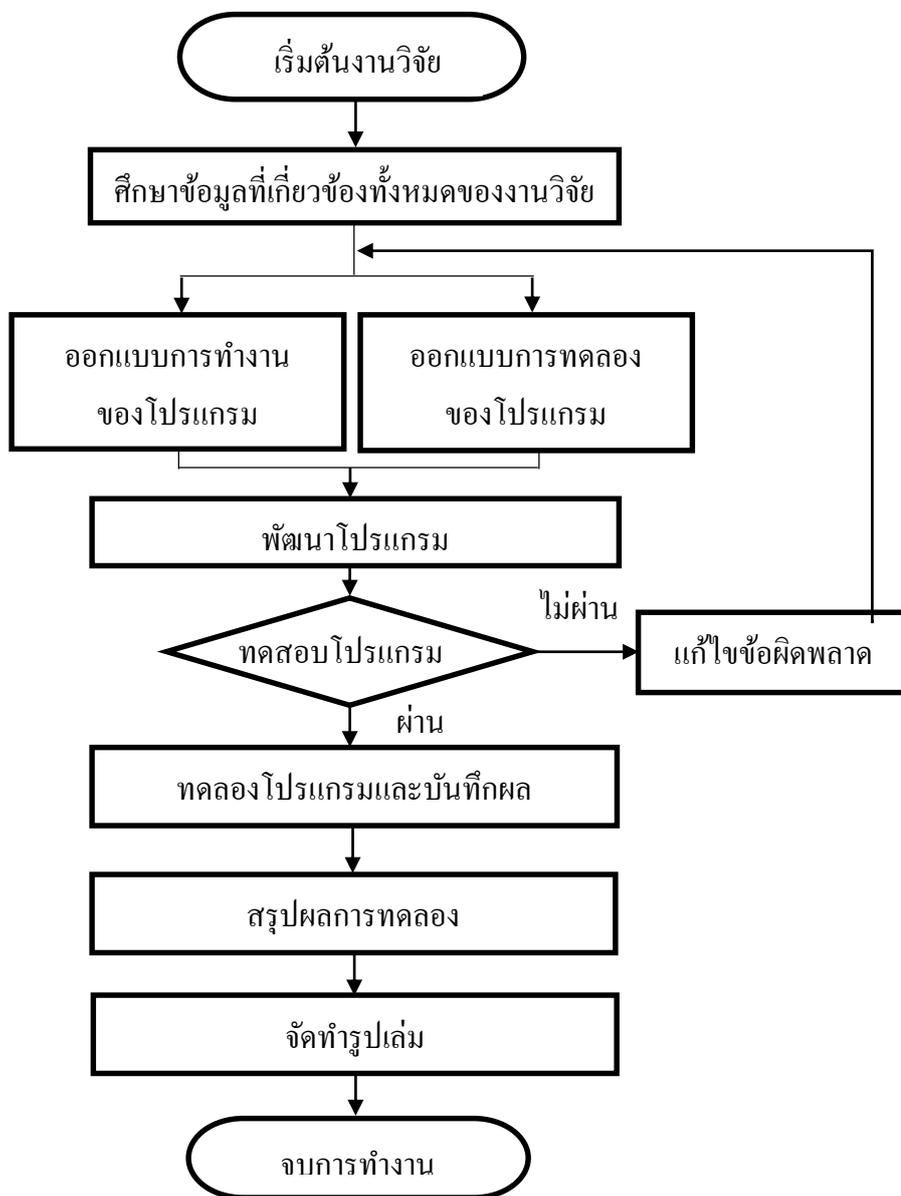
1. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนหน้า คือ ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดที่อยู่ระหว่าง 1-50% ของจำนวนตัวอักษรทั้งหมดในชื่อจังหวัด
2. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนกลาง คือ ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดที่อยู่ระหว่าง 31-80% ของจำนวนตัวอักษรทั้งหมดในชื่อจังหวัด
3. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดส่วนท้าย คือ ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดที่อยู่ระหว่าง 51-100% ของจำนวนตัวอักษรทั้งหมดในชื่อจังหวัด
4. ตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดแบบสุ่ม คือ มีตำแหน่งความผิดพลาดของตัวอักษรในชื่อจังหวัดหลายส่วน แต่มีความผิดพลาดของตัวอักขระรวมแล้วไม่เกิน 50% ของจำนวนตัวอักษรทั้งหมดในชื่อจังหวัด

วิธีการวัดผลการทดลอง

วัดผลการทดสอบด้วยสมการ (1) โดยกำหนดให้ np_{OCR} คือ จำนวนชื่อจังหวัดที่อ่านถูกต้อง, Np_i คือ จำนวนชื่อจังหวัดที่ใช้ในการทดลอง ดังนี้

$$\text{ความถูกต้องของการอ่านชื่อจังหวัด (\%)} = \left(\frac{np_{ocr}}{Np_i} \right) \times 100\% \quad (1)$$

แผนผังการดำเนินงานวิจัย



ภาพที่ 3.7 แผนผังการดำเนินงานวิจัย

ขั้นตอนการดำเนินงานวิจัย

1. ศึกษาปัญหาของงานวิจัยและงานวิจัยที่เกี่ยวข้องกับการเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์
2. ออกแบบการทำงานของโปรแกรมและออกแบบการทดลองของโปรแกรม
3. พัฒนาโปรแกรมเพื่อเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์
4. ทดลองประสิทธิภาพของโปรแกรมและบันทึกผล
5. สรุปผลการทดลอง
6. จัดทำรูปเล่ม

ตารางที่ 3.1 แผนการดำเนินงานวิจัย

รายการ	ระยะเวลาดำเนินการ					
	ส.ค. 2558	ก.ย. 2558	ต.ค. - ธ.ค. 2558	ม.ค. - ก.พ. 2559	มี.ค. 2559	เม.ย. 2559
ศึกษาปัญหาของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์และงานวิจัยที่เกี่ยวข้องกับการเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์						
ออกแบบการทำงานของโปรแกรมและออกแบบการทดลองของโปรแกรม						

ตารางที่ 3.1 (ต่อ) แผนการดำเนินงานวิจัย

รายการ	ระยะเวลาดำเนินการ					
	ส.ค. 2558	ก.ย. 2558	ต.ค. - ธ.ค. 2558	ม.ค. - ก.พ. 2559	มี.ค. 2559	เม.ย. 2559
พัฒนาโปรแกรมเพื่อเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์			↔			
ทดลองประสิทธิภาพของโปรแกรมและบันทึกผลการทดลอง				↔		
สรุปผลการทดลอง					↔	
จัดทำรูปเล่ม						↔

จากตารางที่ 3.1 ในเดือนสิงหาคม 2558 ผู้วิจัยได้ศึกษาปัญหาของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์และงานวิจัยที่เกี่ยวข้องกับการเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์ ในเดือนกันยายน 2558 ผู้วิจัยออกแบบการทำงานของโปรแกรมและออกแบบการทดลองของโปรแกรม ในเดือนตุลาคมถึงเดือนธันวาคม 2558 ผู้วิจัยได้ทำการพัฒนาโปรแกรมเพื่อเพิ่มประสิทธิภาพความถูกต้องของการอ่านชื่อจังหวัดบนป้ายทะเบียนรถยนต์ด้วยการรู้จักอักขระทางภาพจากเทคนิคแอลซีเอส หลังจากนั้นในเดือนมกราคมถึงเดือนกุมภาพันธ์ 2559 ผู้วิจัยทำการทดลองประสิทธิภาพของโปรแกรมและบันทึกผลการทดลอง ในเดือนมีนาคม 2559 ผู้วิจัยทำการสรุปผลการทดลอง และในเดือนเมษายน 2559 ผู้วิจัยจัดทำรูปเล่มวิจัยเพื่อนำเสนอต่อคณะกรรมการวิจัย