

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

#### สรุปผลการวิจัย

โปรแกรมโอซีอาร์ภาษาไทยที่ได้จากวิทยานิพนธ์ ประกอบด้วยขั้นตอนหลักคือ การประมวลผลภาพ การตัดแยกบรรทัด การตัดแยกตัวอักษร การรู้จำตัวอักษร การแก้ไขผลลัพธ์ที่ได้จากขั้นตอนการรู้จำ การสร้างบรรทัด และการแก้ไขคำผิด ผลการรู้จำเอกสาร 48 หน้า ซึ่งมีตัวอักษรแบบ AngsanaUPC, BrowalliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC และ FreesiaUPC แต่ละแบบของตัวอักษรจะประกอบด้วยตัวอักษรขนาด 14, 16, 18, 20, 22, 24, 28 และ 36 จุด ทำการสแกนภาพที่ความละเอียด 300 จุดต่อนิ้ว รวมเป็นจำนวน 71832 ตัวอักษร พบว่าเปอร์เซ็นต์ความผิดพลาดจากผลลัพธ์ของโปรแกรมโอซีอาร์ เท่ากับ 1.85% และ เปอร์เซ็นต์ความผิดพลาดหลังจากทำการแก้ไขคำผิดที่ไม่เป็นคำ เท่ากับ 1.47% และ เปอร์เซ็นต์ความผิดพลาดหลังจากทำการแก้ไขคำผิดที่ไม่เป็นคำและคำผิดที่เป็นคำ เท่ากับ 1.50%

#### ข้อเสนอแนะ

ในการวิจัยนี้มีข้อเสนอแนะบางประการที่น่าจะนำมาทำการปรับปรุงประสิทธิภาพของโปรแกรมโอซีอาร์ภาษาไทยให้ดีขึ้นดังนี้

1. ปรับปรุงอัลกอริทึมในขั้นตอนต่างๆ เพื่อให้โปรแกรมโอซีอาร์สามารถทำงานได้เร็วขึ้นกว่าเดิม
2. ปรับปรุงวิธีการกำจัดจุดภาพรบกวนขนาดใหญ่ และวิธีการเชื่อมตัวอักษรเมื่อเส้นของตัวอักษรบางจนขาดออกจากกัน
3. ปรับปรุงวิธีการตัดแยกตัวอักษรที่สามารถตัดแยกตัวอักษรที่ติดกันให้แยกออกจากกันได้อย่างถูกต้อง
4. ปรับปรุงวิธีการตัดแยกบรรทัดเพื่อให้สามารถใช้งานกับรูปเอกสารที่มีความละเอียดสูงหรือต่ำกว่านี้ หรือเอกสารที่มีช่องว่างระหว่างบรรทัดที่แคบมากๆ เช่น หนังสือพิมพ์ เป็นต้น
5. เพิ่มเติมวิธีการกำหนดขอบเขตของเอกสารในการรู้จำ เพื่อให้สามารถรู้จำข้อความบางส่วนบนเอกสารได้ เช่น กำหนดขอบเขตล้อมรอบเฉพาะข้อความโดยไม่รวมรูปที่ปรากฏบนเอกสาร
6. ปรับปรุงการหมุนภาพของเอกสาร โดยให้เครื่องสามารถคำนวณหาองศาความเอียงและทำการหมุนภาพให้กับผู้ใช้
7. ปรับปรุงการแก้ไขคำผิดให้สามารถแก้ไขคำผิดที่เป็นภาษาอังกฤษได้
8. เพิ่มเติมคำศัพท์ในพจนานุกรมโดยเฉพาะคำทับศัพท์, ชื่อเฉพาะ, คำในภาษาพูด และคำศัพท์ภาษาอังกฤษ