

## ข้อสรุปและข้อเสนอแนะ

## ข้อสรุป

จากการศึกษาการรู้จำตัวอักษรพิมพ์ภาษาไทยในงานวิจัยนี้เน้นการแก้ปัญหาความแม่นยำในการรู้จำของตัวอักษรพิมพ์ภาษาไทยให้ได้สูงขึ้นโดยมีการวิเคราะห์ และพิจารณาถึงปัญหาและ อุปสรรคที่ได้เคยเกิดขึ้นกับกระบวนการวิธีการของงานวิจัยที่ได้เคยผ่านมา อันได้แก่ปัญหากับตัวอักษรประเภทที่มีบางส่วนของตัวอักษรเป็นเส้นหยักมากพอสมควร เช่น ตัวอักษร ข, ค, ม, ท เป็นต้น เส้นหยักของตัวอักษรประเภทนี้จะกลายเป็นเส้นตรงเมื่อผ่านกระบวนการทำให้บางแบบดั้งเดิม (Classical Thinning) พบว่า ที่บริเวณหยักของภาพจะเห็นว่าการเว้าหรือนูนของเส้นออกจากเส้นปกติของตัวอักษรเพียง 1 จุดภาพเท่านั้น ทำให้เมื่อผ่านการทำให้บางแล้วปรากฏว่าได้ออกมาเป็นเส้นตรง เช่น ตัวอักษร ข ก็จะถูกรู้จำเป็นตัวอักษร ข

สระอือ ในวิทยานิพนธ์สนธยา เมรินทร์ [4] ที่ผ่านมา สามารถรู้จำได้เพียง 5 ตัวของจำนวนสระอือทั้งหมด 14 ตัว ทั้งนี้เพราะถ้าความสูงของเส้นตรงเส้นด้านซ้ายบนที่มีลักษณะเป็นขีดเส้นตรงของสระอือมีความสูงของเส้นไม่เกิน 4 จุดภาพ หลังการทำให้บางวิธีเดิมเส้นตรงเส้นนี้จะหายไป หรือเหลือเพียงหยักนูนขึ้นด้านบนเพียงเล็กน้อย และเหลือเพียงเส้นตรงที่อยู่ด้านขวามือเพียงเส้นเดียวเท่านั้น ซึ่งเป็นผลให้สระอือถูกรู้จำเป็นสระอี

งานวิจัยนี้จึงเน้นวิเคราะห์ไปในส่วนของขั้นตอนกระบวนการที่มีผลมากในการรู้จำตัวอักษรพิมพ์ภาษาไทยที่ผ่านมาคือส่วน Preprocessing ซึ่งเป็นส่วนที่จะนำข้อมูลภาพมาเตรียมความพร้อมเพื่อแปลงเป็นข้อมูลรายละเอียดมากพอในการรู้จำ ฉะนั้นงานวิจัยนี้ได้มีการนำเสนอเทคนิควิธีการใหม่โดยการพิจารณารูปแบบเงื่อนไขของวิธีการ Safe Point Thinning Algorithm มาประยุกต์รวมเข้ากับวิธีการทำให้บางแบบดั้งเดิม (Classical Thinning) ซึ่งข้อดีของวิธีการนี้คือ จะไม่ตัดรายละเอียดของภาพทิ้งในส่วนด้านบนของแนวนอน ซึ่งเหมาะกับตัวอักษรภาษาไทย เช่น ลักษณะหัวหยักเช่น ข , ลักษณะภาพที่จำนวนจุดภาพที่ปรากฏไม่เกิน 4 จุดภาพ เช่น สระอือ และนอกจากนี้ยังแสดงเส้นบนของ ส และ ศ ให้เกิดความยาวมากขึ้นกว่าชัดเจนกว่าเดิมด้วย

ส่วนการนำเทคนิคพีซีโลจิกมาใช้นั้นเป็นการช่วยเสริมในการเปรียบเทียบจากวิธีซินแทกติกให้สามารถรู้จำตัวอักษรที่ไม่สามารถรู้จำได้ และยังแก้ไขตัวอักษรที่มีการรู้จำผิดที่มีลักษณะของความผิดพลาดที่เกิดขึ้นอย่างเด่นชัด เช่น จำนวนวงกลมแตกต่างกัน ,จำนวนส่วนย่อยแตกต่างกัน เป็นต้น ซึ่งผลการทดสอบ ผลการรู้จำจากตัวอย่างตัวอักษร 2 รูปแบบ ๆ ละ 6 ขนาดตัวอักษร รวมเป็นตัวอักษรทั้งสิ้น 1106 ตัวอักษรพบว่ามีอัตราการรู้จำได้เท่ากับ 99.64% ใช้เวลาในการรู้จำเฉลี่ยต่อตัวอักษร 0.89 วินาทีบนเครื่องไมโครคอมพิวเตอร์ 486Dx2-66 (ความเร็ว 66 เมกกะเฮิร์ตซ์) สำหรับเวลาในการรู้จำเฉลี่ยต่อตัวอักษรของวิทยานิพนธ์สนธยา เมรินทร์ (2537) [4] เท่ากับ 1.09 วินาทีบนเครื่อง IBM-PC 486sx ที่ความเร็ว 33 เมกกะเฮิร์ตซ์ ซึ่งถ้าเทียบเวลากันโดยไม่คำนึงสมรรถนะของเครื่องไมโครคอมพิวเตอร์มาเป็นปัจจัย ในวิทยานิพนธ์ฉบับนี้เวลาในการรู้จำเฉลี่ยต่อตัวอักษรเร็วกว่าวิทยานิพนธ์สนธยา เมรินทร์ (2537) [4] แต่ถ้าคำนึงถึงสมรรถนะของเครื่องไมโครคอมพิวเตอร์มาเป็นปัจจัย ในวิทยานิพนธ์ฉบับนี้เวลาในการรู้จำเฉลี่ยต่อตัวอักษรช้ากว่าวิทยานิพนธ์สนธยา เมรินทร์ (2537) [4] ซึ่งเกิดจากขบวนการต่าง ๆ ที่ปรับปรุงเข้าไปใช้เพื่อให้อัตราการรู้จำสูงขึ้น



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

### ข้อเสนอแนะ

แนวทางในการพัฒนาที่น่าจะทำต่อไปคือ สามารถนำไปใช้ในการรู้จำตัวอักษรพิมพ์ภาษาไทยให้มีรูปแบบตัวอักษรหลากหลายมากขึ้น โดยพัฒนาตัวอักษรภาษาไทยต้นแบบที่มีความเป็นกลางของตัวอักษรภาษาไทยหลายรูปแบบเพื่อใช้ในการรู้จำได้อย่างแท้จริง โดยไม่ได้ยึดติดกับรูปแบบตัวอักษรรูปแบบใดรูปแบบหนึ่งหรือขนาดใดขนาดหนึ่งโดยเฉพาะ เช่นการนำตัวอักษรภาพของภาษาไทยหลายรูปแบบในแต่ละตัวอักษรมาหาค่าเฉลี่ยจุดภาพของตัวอักษรนั้นๆ ก่อนนำมาทำเป็นอักษรต้นแบบ เพื่อให้ได้ความเป็นกลางมากที่สุด และ เวลาการรู้จำควรพัฒนาให้ต่ำลงไป โดยใช้เทคนิคทางโปรแกรมมิ่ง เพื่อสามารถนำไปประยุกต์ใช้งานได้จริง หรือการนำเทคนิคนิวรอลเน็ตเวิร์กเข้ามาช่วยในการเรียนรู้รูปแบบตัวอักษรที่หลากหลายได้ และถ้ามีการพิจารณารูปแบบของคำศัพท์ต่าง ๆ ประกอบก็อาจจะทำให้ความถูกต้องของการรู้จำมีค่าสูงขึ้นมากได้



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย