หัวข้อวิทยานิพนธ์ การใช้ครอสคอร์รีเลชันนิวรอลเน็ตเวิร์คสำหรับการรู้จำตัวอักษรพิมพ์ภาษาไทยแบบฟอนต์ไม่มีหัว

นักศึกษา นางสาว พฤกษา ดวงผาสุข

รหัสนักศึกษา 44067036

ปริญญา วิทยาศาสตรมหาบัณฑิต

สาขาวิชา เทคโนโลยีสารสนเทศ

พ.ศ. 2549

อาจารย์ผู้ควบคุมวิทยานิพนธ์ รศ. ดร. อาริต ธรรมโน

บทคัดย่อ

งานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรภาษาไทยที่ผ่านมามุ่งเน้นในการเพิ่มความสามารถในระบบการรู้จำตัวอักษรภาษาไทยด้วยการพิจารณาโครงสร้างลักษณะเด่นของตัวอักษรที่มีรูปแบบตัวอักษรมาตรฐานตามคุณลักษณะตัวอักษรภาษาไทยแบบดั้งเดิมที่ประกอบไปด้วย เส้นตรง เส้นโค้ง เส้นซิกแซก และการเริ่มต้นตัวอักษรด้วยลูปวงกลมหรือที่เรียกว่าหัวตัวอักษร ส่วนกลุ่มรูปแบบตัวอักษรที่โครงสร้างของตัวอักษรปราศจากส่วนหัวตัวอักษรเช่น “**ก**” “ภ” “ถ” “**ด**” “**ต**” “**น**” “**บ**” “**ข**” “**ข**” “**ช**” “**ซ**” จากตัวอย่างดังกล่าวตัวอักษรปราศจากส่วนหัวตัวอักษรนั้นจะเห็นว่าตัวอักษรมีรูปแบบคล้ายกันมากขึ้น และปัญหาที่ตามมาคือความสามารถในการรู้จำตัวอักษรภาษาไทยรูปแบบตัวอักษรที่ไม่มีหัวยังไม่ถูกต้องเท่าที่ควร

ดังนั้นวิทยานิพนธ์นี้จึงได้นำเสนอแนวคิดวิธีการใหม่โดยใช้การแบ่งแยกตามลำดับชั้นในโครงข่ายประสาทเทียม Hierarchical Normalization Cross-Correlation ARTMAP สำหรับการรู้จำรูปแบบตัวอักษรพิมพ์ภาษาไทยที่ไม่มีหัว และนำเอาหลักการของครอสคอร์รีเลชันมาใช้เป็นฟังก์ชันสำหรับการวัดค่ามาตรฐานความคล้ายคลึงระหว่าง Input Vector กับ Weight Vector ข้อดีของการใช้ ครอสคอร์รีเลชันคือสามารถวัดค่าความคล้ายคลึงกันของรูปแบบในแต่ละส่วนออกมาได้โดยไม่ขึ้นกับขนาด เพราะเป็นการเปรียบเทียบความเหมือนในแต่ละส่วนของตัวอักษร โดยไม่จำเป็นต้องมีขนาดเริ่มต้นที่เท่ากัน กลุ่มข้อมูลที่ใช้ในการ Train และ Test ประกอบด้วยรูปแบบตัวอักษรพิมพ์ภาษาไทยไม่มีหัว 39 แบบและตัวอักษรบนภาพจากป้ายทะเบียนรถไทย 506 ป้าย ผลการทดลองที่ได้จากการรู้จำด้วยวิธีใหม่ที่นำเสนอโดยเฉลี่ยเท่ากับ 88.49% นอกจากนี้ยังนำกลุ่มข้อมูลไปทดสอบกับโปรแกรม ArnThai, ThaiOCR และ Fuzzy ARTMAP