

บทคัดย่อ

งานวิจัยนี้เป็นการเสนอแนวทางการจัดเตรียมตัวอักษรจากแฟ้มข้อมูลที่เป็นรูปภาพที่มีรูปแบบการเก็บแบบพีซีเอกซ์ โดยได้มาจากการเครื่องกวาดภาพเพื่อใช้เป็นข้อมูลให้กับการประมวลผล การเรียนรู้และจดจำอักขระ ระบบที่พัฒนาขึ้นมานี้เริ่มต้นจากการเปลี่ยนการจัดเก็บเอกสารแบบรูปภาพให้มาอยู่ในรูปของเท็กซ์ก่อน จากนั้นจะมีการตรวจสอบความเอียงของภาพที่นำมาวิเคราะห์ ถ้าภาพที่นำมาวิเคราะห์มีการเอียงก็จะทำการปรับภาพให้ตรงโดยอัตโนมัติ แยกแต่ละบรรทัดออกจากกัน แล้วทำการแยกอักขระในแต่ละบรรทัดออกเป็นแต่ละอักขระ ทำอักขระให้มีลักษณะเป็นโครงร่าง โดยมีการแทนจุดสีดำและจุดสีขาวของภาพด้วย “1” และ “0” ตามลำดับ แล้วทำการปรับขนาดของภาพอักขระที่ได้ให้มีรูปแบบของตัวอักษรในลักษณะของแมทริกซ์ ขนาด 8×20 จากการทดลองพบว่าวิธีการที่นำเสนอนี้สามารถใช้แยกตัวอักษรได้เป็นอย่างดี และยังสามารถปรับขนาดของผลลัพธ์เพื่อให้เหมาะสมกับการใช้งานได้ด้วย

วิทยานิพนธ์การงาน

อาจารย์ รัชต์ ผู้อำนวยการงานจัดการจัด

ABSTRACT

This paper present the technique for input processing of Thai printed charecter recognition. The document is input into computer using a scanner in PCX picture format. Then the file is converted to a text file. The inclination of the scanned document is determine and adjust to the correct direction automatically. Then each line of document is separated and each character in each line is also separated from each other. The out-line of each character is constructed by represent black and white pixel by "1" and "0" respectively. The last step is to scale each character to fit into a 8 X 20 block size. The experimental results show that the proposed technique can be applied to Thai character efficiently and a size of output also be adjusted to a suitable size.