

**ชื่อเรื่อง การจำแนกชนิดภาพเอกสารหนังสือราชการ ไทยด้วยเทคนิคการประมวลผลภาพ
ร่วมกับโครงข่ายภาษาเทียน**

ผู้วิจัย ภานุวัฒน์ เกตุวงศ์	ปริญญา วท.ม. (เทคโนโลยีคอมพิวเตอร์เพื่อการศึกษา)
อาจารย์ที่ปรึกษา	ผศ. สิทธิชัย บุญหมื่น
	ผศ. กาญจนा คำสมบัติ
	ผศ. พัชรพงษ์ พันธุ์มณี

มหาวิทยาลัยราชภัฏมหาสารคาม 2554

บทคัดย่อ

ถึงแม้ว่าการติดต่อสื่อสารแบบอิเล็กทรอนิกส์จะเพิ่มมากขึ้นในปัจจุบัน สาระสำคัญของเอกสารจำนวนมากที่ยังคงถูกบันทึกลงบนกระดาษ ดังนี้ในความพยายามที่จะลดปริมาณการใช้กระดาษในสำนักงานจึงมีการจัดเก็บเอกสารจากที่เป็นกระดาษในรูปแบบของข้อมูลภาพ

หนังสือราชการเป็นเอกสารสำคัญที่ใช้ในการติดต่อสื่อสารระหว่างหน่วยงานซึ่งสามารถแบ่งออกเป็นหลายชนิดตามวัตถุประสงค์ของเอกสารและองค์ประกอบอื่น ๆ เช่น ผู้แต่ง ดังนั้นงานวิจัยนี้จึงนำเสนอการพัฒนาเครื่องมือการจำแนกชนิดภาพเอกสารหนังสือราชการ ไทยด้วยหลักการทางการประมวลผลภาพร่วมกับโครงข่ายภาษาเทียน โดยมีลักษณะกระบวนการดำเนินการ 3 กระบวนการ คือ 1) การประมวลผลเบื้องต้น เป็นการกระบวนการเพื่อเตรียมข้อมูลให้เหมาะสมกับการทดลอง มีขั้นตอน 2 ขั้นตอน คือ การสร้างภาพใบเริร์จจากหลักการพิจารณาจากอิสโตร์นิค และการปรับความอิ่งของภาพด้วยหลักการฟาร์มนัสฟอร์ม 2) การสกัดลักษณะสำคัญ เป็นกระบวนการลักษณะสำคัญที่ทำการสกัดลักษณะสำคัญที่เป็นเอกลักษณ์ของภาพแต่ละชนิดเพื่อนำลักษณะสำคัญนั้นมาเป็นข้อมูลนำเข้าสำหรับโครงข่ายภาษาเทียน ซึ่งใช้หลักการของวินโดว์ไฟเจอร์ 3) การสร้างตัวจำแนกด้วยโครงข่ายภาษาเทียน

การทดสอบการจำแนกได้ใช้ภาพจำนวน 125 ภาพ จากหนังสือราชการ 5 ชนิด มาทดสอบในรูปแบบของวินโดว์และจำนวนโหนดในชั้นช่องของโครงข่ายภาษาเทียนที่แตกต่างกัน ผลการทดสอบ พบว่า การจำแนกชนิดภาพเอกสารที่ให้ผลจำแนกถูกต้องสูงสุดคือร้อยละ 82.40 จากการใช้รูปแบบวินโดว์ 4x5 และใช้จำนวนโหนดในชั้นช่องจำนวน 3 โหนด

Title : Type Classification of Thai government official document images by using

Image Processing technique and Artificial Neural Network

Author : Panuwat ketwong Degree : M.Sc. (Computing Technology for Education)

Advisors : Asst. Prof. Sittichai Bussaman **Chairman**

Asst. Prof. Kanjana Khamsombat Committee

Asst. Prof. Natthapong Phanmanee Committee

Rajabhat Maha Sarakham University, 2011

ABSTRACT

In spite of the explosive increase of electronic communication in recent years, a significant amount of documents is still printed on paper. In an attempt to move toward a paper-less office, large quantities of printed documents are often scanned and archived as images.

Thai government official document is the document used in communication among government organizations. It can be divided into many types, depending on their objectives and some factors such as a class of the composer who generates the document.

In this research proposed methods for categorizing Thai government official document images into their classes by using Image Processing technique and Artificial Neural Network. The research was operated in 3 procedures : 1) Data preparation, by using the process of histogram-based binarization to convert multi-level images to bi-level images and Hough Transform for skew correction; 2) Feature extraction, , by using the window feature to indentify characteristics of images; 3) Implement classifier, implementing classification scheme by using artificial neural network.

125 document images of Thai government official document images from 5 types of documents were tested by using difference window formats and different nodes of hidden layer. The classification yielded the highest rate of accuracy: 82.40 per cent from window 4x5 with 3 nodes of hidden layer.