

ชื่อเรื่องวิทยานิพนธ์ การรู้จำอักษรไทยโดยใช้วิธีการติดตามลายเส้น

ชื่อผู้เขียน นายกฤษดา ยิ่งขยัน

วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า

คณะกรรมการสอบวิทยานิพนธ์ :

รองศาสตราจารย์ ขจรศักดิ์ คັນทรนิต	ประธานกรรมการ
ผู้ช่วยศาสตราจารย์ ดร. กิติ ลิขิตอนุรักษ์	กรรมการ
นายธนวิชญ์ ชุลิกาวิทย์	กรรมการ

บทคัดย่อ

วิทยานิพนธ์นี้เป็นการพัฒนาโปรแกรมวิเคราะห์ภาพอักษรไทย โดยใช้ทฤษฎีโทโปโลยีคอลเฟียเจอร์เอ็กซ์แทรคชัน (Topological Feature Extraction) ในการค้นหาลักษณะโครงร่างของลายเส้น

การวิเคราะห์เริ่มจากผู้ใช้งานทำการกวาดภาพอักษรตัวพิมพ์ภาษาไทยจากเอกสารลงในหน่วยความจำโดยใช้เครื่องกวาดภาพ (Scanner) กระบวนการรู้จำจะแยกบรรทัดข้อความออกจากภาพบทความ และแยกตัวอักษรออกจากบรรทัด ต่อจากนั้นส่งภาพอักษรที่ถูกตัดไปยังภาควิเคราะห์ ซึ่งภาควิเคราะห์จะทำการหาขอบภาพของตัวอักษร แล้ววิเคราะห์หารหัสเฉพาะของตัวอักษร (A-Code) และส่งรหัส A-Code ไปตรวจสอบเพื่อหารหัสแอสกี (ASCII) ของภาพอักษร แล้วสร้างเท็กซ์ไฟล์สำหรับเอกสารที่วิเคราะห์

โปรแกรมที่พัฒนาขึ้นมานี้ เขียนโดยใช้โปรแกรมเดลไฟ (Delphi) ทำงานบนระบบปฏิบัติการวินโดวส์ 3.11 ที่สามารถแสดงผลภาษาไทยได้

ผลการทดสอบโปรแกรมพบว่า ความถูกต้องของข้อมูลเท็กซ์ไฟล์ เมื่อทดสอบตามรูปแบบของภาพอักษรในรูปแบบต่าง ๆ แล้ว จะมีความถูกต้องประมาณร้อยละ 61 และได้เพิ่มโมดูลเพื่ออำนวยความสะดวกในการคัดป้อนตัวอักษรที่ผิด เพื่อให้สามารถนำไปใช้งานจริงได้

มหาวิทยาลัยเชียงใหม่  
Chiang Mai University

**Thesis Title** Thai Optical Character Recognition (OCR) using Line Segmentation Algorithm

**Author** Mr. Krisda Yingkayun

**M.Eng.** Electrical Engineering

**Examining Committee :**

Associate Prof. Kajomsak Kantapanit	Chairman
Assistant Prof. Dr. Kiti Likit-anurucks	Member
Mr. Dhanavich Chulikavit	Member

**Abstract**

This thesis presents the development of Thai Optical Character Recognition (OCR) using Line Segmentation Algorithm. The analytical process is carried out according to Topological Feature Extraction theory. The analytical process begins with the user scans Thai characters' image into memories by a scanner. The recognition process then extracts the characters' images from this, and proceeds to analytical process to find out the A-code. The A-code is then sent to the checking process to find out the ASCII code. Finally, the ASCII code generated into text file.

The software program is written in Delphi version 1.0 to run on Windows 3.11 (Thai Edition) operating system.

The conclusion shows that the accuracy of the text file data with the change in the style of fonts' images is about 61 %. A module is then added so that the user can keyin the mistaken characters for practical application.

มหาวิทยาลัยเชียงใหม่  
Chiang Mai University