



หัวข้อวิทยานิพนธ์

การรู้จำตัวอักษรภาษาไทยแบบลายมือเขียนโดยใช้การแปลงเวฟเลตที่สามารถปรับเปลี่ยนรูปร่างได้

หน่วยกิตของวิทยานิพนธ์

12 หน่วย

โดย

นางสาวอรรวรรณ เผ่าพนัส

อาจารย์ที่ปรึกษา

รศ.ดร. โกสินทร์ จ่านงไทย

ระดับการศึกษา

วิศวกรรมศาสตรมหาบัณฑิต

ภาควิชา

วิศวกรรมไฟฟ้า

ปีการศึกษา

2544

บทคัดย่อ

ในลายมือเขียนภาษาไทยนั้นมีรูปแบบการเขียนที่ต่างกันไป ตัวอักษรตัวเดียวกันอาจเขียนต่างกันได้ หรือตัวอักษรที่เป็นคนละตัวกันก็มีลักษณะคล้ายกันจะแตกต่างกันก็เพียงในส่วนเล็กน้อยเท่านั้น ดังนั้นเพื่อที่จะแยกแยะตัวอักษรแต่ละตัวให้ถูกต้องเหมาะสมกับลายมือเขียนของคนเรามากที่สุดวิทยานิพนธ์นี้จึงเสนอการรู้จำตัวอักษรภาษาไทยแบบลายมือเขียน โดยใช้วิธีการแปลงเวฟเลตที่สามารถปรับเปลี่ยนรูปร่างได้ ซึ่งจะใช้เส้นโครงร่างของตัวอักษรภาษาไทยแบบตัวพิมพ์มาทำเป็นตัวอักษรต้นแบบโดยเปลี่ยนการแสดงเส้นโครงร่างของตัวอักษรในรูปแบบพิกเซลให้อยู่ในรูปของแกนเวลา 1 มิติ หลังจากนั้นก็ทำการหาช่วงการปรับเปลี่ยนรูปร่างของตัวอักษรต้นแบบจากกลุ่มของตัวอักษรที่ใช้ในการสอนระบบ ซึ่งใช้เส้นโครงร่างของตัวอักษรภาษาไทยแบบลายมือเขียนหลายๆ แบบที่อยู่ในแกนเวลา 1 มิติแล้วมาทำการแปลงเวฟเลต หลังจากนั้นก็นำค่าสัมประสิทธิ์ของเวฟเลตที่ได้มาคำนวณหาค่าเบี่ยงเบนมาตรฐานที่ทำให้ตัวอักษรต้นแบบสามารถปรับเปลี่ยนรูปร่างไปได้ภายในค่าเบี่ยงเบนมาตรฐานนี้ เพื่อให้พอดีกับตัวอักษรที่ได้รับเข้ามาเพื่อทำการรู้จำและจะทำการวัดค่าความแตกต่างของตัวอักษรต้นแบบกับตัวอักษรที่ได้รับเข้ามานี้ ตัวอักษรต้นแบบตัวใดที่ให้ค่าความแตกต่างน้อยที่สุดก็จะถือว่าตัวอักษรตัวนั้นเป็นแบบเดียวกันกับตัวอักษรที่ได้รับเข้ามา แสดงออกมาเป็นผลของการรู้จำตัวอักษร ในการทดสอบนั้นจะแบ่งเป็น 2 ส่วน ส่วนแรกจะเป็นเปรียบเทียบผลการรู้จำเมื่อใช้ลายมือเขียนแบบบรรจงกับลายมือเขียนแบบหวัด โดยเราได้ใช้ตัวอักษรในภาษาไทยทั้ง 44 ตัว ที่เขียนโดยคน 10 คน (440 ตัว) เป็นกลุ่มตัวอักษรสำหรับสอนระบบ และใช้ลายมือเขียนแบบบรรจงและแบบหวัดอีกอย่างละ 10 คน (880 ตัว) เป็นตัวอักษรที่จะทำการรู้จำ ผลการทดสอบแสดงถึงความถูกต้องเมื่อใช้ลายมือเขียนแบบบรรจงเป็นร้อยละ 98.64 และเมื่อใช้ลายมือเขียนแบบหวัดเป็นร้อยละ 89.1 ในการทดสอบส่วนที่

สองเป็นการเปรียบเทียบผลการรู้จำเมื่อใช้วิธีการปรับรูปร่างจากค่าสัมประสิทธิ์ของการแปลงฟูเรียร์กับวิธีที่เสนอขึ้น ผลปรากฏว่าวิธีที่เสนอขึ้นนี้ให้อัตราการรู้จำถูกต้องกว่าถึงร้อยละ 10

คำสำคัญ (Keywords) : การรู้จำตัวอักษร / การแปลงเวฟเลตที่สามารถปรับเปลี่ยนรูปร่างได้ / ตัวอักษรต้นแบบ / ช่วงการปรับรูปร่าง / ค่าเบี่ยงเบนมาตรฐาน

Abstract

Handwritten Thai characters have complex structure and unconstrained shape depending on writing style. In order to increase recognition rate and use low computation, this thesis proposes handwritten Thai characters recognition using deformable wavelet descriptor. In the recognition, contours of Thai characters are utilized to determine templates in term of temporal domain. Then, range of deformation is obtained by standard deviation of wavelet descriptor coefficients from character training set. To fit the templates with input characters, all templates are deformed to obtain the best fit by determined deformation range. The best fit of each character template is represented by score, and the character template with highest score is the recognition result. The experiment consists of two part. The first is experiment for comparing the recognition rates of the system for constrained and unconstrained handwritten Thai character. We use ten styles of characters (440 categories) as the training set and twenty styles (880 categories) as the input characters. The results show 98.64 percent recognition rate for constrained shapes and 89.1 percent for unconstrained shapes. The second one is the experiment for comparing the recognition rates of the system using deformable Fourier descriptor with that of the proposed method. The results show that the recognition rate of the proposed method is about 10 percent better.

Keywords : Character Recognition / Deformable Wavelet Descriptor / Template / Deformable Range / Standard Deviation