

แม้ว่ามีการศึกษาวิจัยการรู้จำตัวอักษรและตัวเลขภาษาไทยอย่างกว้างขวาง แต่การศึกษาวิจัยด้านการรู้จำลายมือเขียนยังไม่ประสบความสำเร็จมากนัก โดยเฉพาะอย่างยิ่งระบบที่ไม่จำกัดความหลากหลายของลายมือเขียน งานวิจัยนี้เป็นการพัฒนาขั้นตอนวิธีเพื่อการรู้จำลายมือเขียนตัวเลขไทยโดยใช้การจัดกลุ่มแบบฟัซซี่ การรู้จำจะเริ่มจากการนำภาพตัวเลขไทยแบบตัวอักษรโดดซึ่งเป็นภาพขาว-ดำ มาหาขอบของภาพ จากนั้นนำจุดพิกัดบนขอบภาพทั้งหมดที่ได้มาทำการแปลงฟูเรียร์ ผลลัพธ์ที่ได้คือ Fourier Descriptor (FD) ต่อมาในขั้นตอนการเรียนรู้ก็นำเวกเตอร์ที่ได้จากภาพตัวเลขแต่ละตัวมาจัดกลุ่มด้วยวิธีค่าเฉลี่ย-ซีแบบไม่ชัดเจน (Fuzzy C-Means) ผลลัพธ์ที่ได้จะเรียกว่า Training Set ในขั้นตอนการทดสอบจะนำภาพตัวเลขที่ต้องการทดสอบมาหาขอบภาพและคำนวณหาค่า FD ผลลัพธ์ที่ได้จะเรียกว่า Testing Set การทดลองจะพิจารณาจากระยะทางสั้นที่สุดระหว่าง Testing Set กับ Training Set ผลการทดลองพบว่าการจำแนกลายมือเขียนแบบขึ้นกับผู้เขียนจะมีค่าความถูกต้องที่ 93.00% ในขณะที่การจำแนกลายมือเขียนแบบไม่ขึ้นกับผู้เขียนจะมีค่าความถูกต้องที่ 56.85%

Although many research papers on Thai printed character and numeric recognition have been widely conducted, much less paper on Thai handwritten recognition has been studied, especially on non-limited writer-independent systems. In this thesis, an algorithm for Thai handwritten numeric recognition using Fuzzy C-Means was developed. The recognition began by binarizing isolated Thai numeric pictures. Then an image contour was gathered. Next, features, Fourier Descriptors (FD), were extracted by transforming co-ordinate values of each pixel on the contour from the spatial domain to the frequency domain using the Fourier Transform. In learning stage, the FD feature vectors of each group of numeric picture were utilized to calculate mean. These mean values were then used as the initialized values for Fuzzy C-Means clustering. In testing stage, an unknown numeric picture was acquired and binarized. In FD feature vector was then detected. To assign a group to this numeric picture, the shortest distance between the training feature set and the testing feature set was considered. The Results from the experiments shown that the writer-dependent clustering gave 93.00% of accuracy whereas the writer-independent clustering gave only 56.85% of accuracy.