

เทคนิคการตรวจหาเส้นบรรทัดจากลายมือเขียนโดยอัตโนมัติ

กรัณย์ ต้นไม้ทอง และ สุภาวรรณ อันนันหนับ

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Emails: golfkungjaa@gmail.com, supawan@it.kmitl.ac.th

บทคัดย่อ

การปรับปรุงลายมือให้มีความสวยงาม (Handwriting beautification) ประกอบด้วยขั้นตอนสำคัญ 2 ขั้นตอนได้แก่ 1. ขั้นตอนก่อนการประมวลผล (Preprocessing) 2. ขั้นตอนการปรับลายเส้นตัวอักษรให้มีความสวยงาม (Beautification) ขั้นตอนก่อนการประมวลผลเป็นขั้นตอนสำหรับการค้นหาเส้นบรรทัดเพื่อใช้เป็นเส้นอ้างอิงสำหรับปรับแนวของตัวอักษร ก่อนเข้ากระบวนการปรับแต่งให้อ่านง่าย ขั้นตอนการปรับลายเส้นตัวอักษรให้มีความสวยงามเป็นขั้นตอนสำหรับปรับลายเส้นตัวอักษรให้มีความชัดเจน มีความสวยงามและทำให้ตัวอักษรอ่านได้ง่ายยิ่งขึ้น ในบทความฉบับนี้นำเสนอเฉพาะขั้นตอนก่อนการประมวลผลด้วยการค้นหาเส้นบรรทัดจำนวนทั้งสิ้น 3 วิธีได้แก่ 1. การค้นหาเส้นบรรทัดด้วยฮิสโตแกรม 2. การค้นหาเส้นบรรทัดด้วยฮิสโตแกรมแบบประยุกต์ 3. การค้นหาเส้นบรรทัดด้วยโครงข่ายประสาทเทียม

คำสำคัญ – โครงข่ายประสาท; เส้นบรรทัด; ลายมือเขียน; คอนโวลูชัน; ฮิสโตแกรม; ตัวอักษร

Abstract

Handwriting beautification includes with two main processes. First, preprocessing step which is a method for line segmentation. In order to fit characters perfectly within the lines, line segmentation must be accurately. Second, beautification step which is a method to beautify the characters. In this paper we present interesting and widely used preprocessing techniques. Three methods are introduced, Projection Profile method, Generalized Profile method and Neural Network method.

Keyword – neural networks; line segmentation; handwriting; convolution; histogram; character thai

1. บทนำ

วิธีการการปรับปรุงลายมือเขียนให้มีความสวยงาม (Handwriting Beautification) [1] ได้รับความสนใจและวิจัยอย่างกว้างขวางในหลาย ๆ ภาษาซึ่งข้อได้เปรียบของวิธีการปรับปรุงลายมือเขียนให้มีความสวยงามมากยิ่งขึ้นเมื่อเทียบกับวิธีการรู้จำตัวอักษรจากลายมือเขียนพบว่า

อยู่หลายอย่างด้วยกัน เช่น เป็นวิธีที่ไม่ต้องอาศัยข้อมูลตัวอย่างทำให้ไม่สิ้นเปลืองทรัพยากรในการจัดเก็บข้อมูล และเป็นวิธีการที่อนุญาตให้ผู้เขียนสามารถจดบันทึกสิ่งต่าง ๆ ที่ไม่ใช่ตัวอักษรหรือพยัญชนะได้ การปรับปรุงลายมือเขียนให้มีความสวยงามประกอบด้วย 2 กระบวนการย่อยได้แก่ 1. กระบวนการก่อนการประมวลผล (Preprocessing) 2. กระบวนการปรับ

ลายเส้นตัวอักษรให้มีความสวยงาม (Beautification) กระบวนการก่อนการประมวลผลเป็นขั้นตอนสำหรับการค้นหาเส้นบรรทัด (Line Segmentation) [2] เพื่อใช้เป็นแนวทางสำหรับการปรับตัวอักษรให้อยู่ในระดับเดียวกัน ก่อนนำไปใช้ประโยชน์ในขั้นตอนถัดไป กระบวนการปรับลายเส้นตัวอักษรให้มีความสวยงามเป็นขั้นตอนสำหรับปรับลายเส้นของตัวอักษรให้มีความชัดเจน มีความเป็นระเบียบ มีความสวยงามและให้อ่านตัวอักษรได้ง่ายยิ่งขึ้น ในบทความฉบับนี้กล่าวถึงขั้นตอนสำหรับการค้นหาเส้นบรรทัดด้วยวิธีการต่าง ๆ 3 วิธีได้แก่

1. การค้นหาเส้นบรรทัดด้วยฮิสโตแกรม [3]
2. การค้นหาเส้นบรรทัดด้วยฮิสโตแกรมแบบประยุกต์ [3]
3. การค้นหาเส้นบรรทัดด้วยโครงข่ายประสาทเทียม [4]

เพื่อให้ง่ายต่อการทำความเข้าใจของผู้อ่าน ผู้เขียนจึงลำดับการเขียนบทความด้วยลำดับดังต่อไปนี้

ส่วนที่ 2 อธิบายถึงลักษณะตัวอักษรภาษาอังกฤษ และการเก็บข้อมูลลายมือเขียนดิจิทัล

ส่วนที่ 3 อธิบายถึงการค้นหาเส้นบรรทัดด้วยวิธีการต่าง ๆ

ส่วนที่ 4 เป็นส่วนสรุปสุดท้ายของบทความฉบับนี้

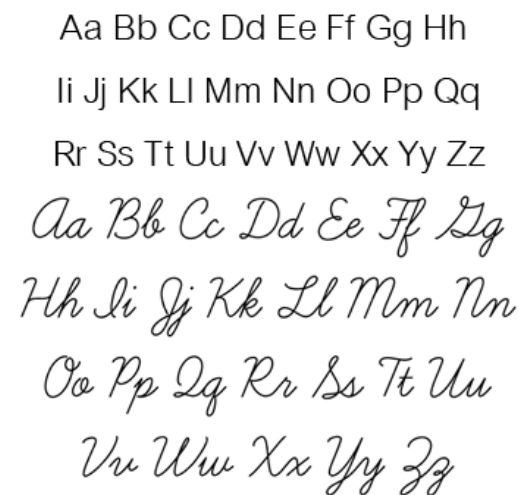
2. ลักษณะตัวอักษรภาษาอังกฤษและการเก็บข้อมูลลายมือเขียนดิจิทัล

2.1. ลักษณะโดยทั่วไปของตัวอักษรภาษาอังกฤษ

ตัวอักษรภาษาอังกฤษแบ่งออกเป็น 4 แบบได้แก่

1. ตัวพิมพ์ใหญ่ (Capital letters)
2. ตัวพิมพ์เล็ก (small letters)
3. ตัวเขียนใหญ่ (cursive uppercase letters)
4. ตัวเขียนเล็ก (cursive lowercase letters)

ดังรูปที่ 1 การฝึกเขียนตัวอักษรภาษาอังกฤษทั้ง 4 แบบในวัยเด็กจะมีเส้นบรรทัด 2 เส้นกำกับเพื่อเป็นแนวทางในการเขียนตัวอักษรให้มีความสวยงามและได้ขนาดที่เท่ากัน



ดังรูปที่ 1 ตัวอักษรภาษาอังกฤษแบบต่าง ๆ

แต่เมื่อตัวอักษรถูกสร้างด้วยระบบดิจิทัลแทนการเขียนด้วยมือมนุษย์จำนวนเส้นบรรทัดที่ช่วยในการกำหนดความสวยงามจะเพิ่มขึ้นเป็น 4 ระดับได้แก่ 1.เส้นบรรทัดฐาน (Baseline) 2. เส้นบรรทัดกลาง (Meanline) 3. เส้นบรรทัดบน (Ascender) 4. เส้นบรรทัดล่าง (Descender) (ดังรูปที่ 2) ตัวพิมพ์ใหญ่และตัวเขียนใหญ่จะถูกเขียนอยู่ระหว่างเส้นบรรทัดฐานและเส้นบรรทัดบน ในขณะที่ตัวพิมพ์เล็กและตัวเขียนเล็กจะมีขนาดความสูงเท่ากับ x-height คืออยู่ระหว่างเส้นบรรทัดฐานและเส้นบรรทัดกลาง แต่ตัวอักษรเล็กบางตัวจะมีเส้นบางส่วนของตัวอักษรเกินความสูง x-height เรียกว่า Overshoot เส้นเหล่านี้จะไปสิ้นสุดที่เส้นบรรทัดบนหรือเส้นบรรทัดล่างขึ้นอยู่กับตัวอักษร ดังรูปที่ 2 [5]



รูปที่ 2 ตัวอักษรภาษาอังกฤษกับบรรทัดของเส้นบรรทัด

2.2. ข้อมูลลายมือเขียนดิจิทัลและการเก็บข้อมูลลายมือเขียนดิจิทัล

เมื่อเขียนตัวอักษรบนอุปกรณ์เก็บข้อมูลอย่างเช่นแท็บเล็ตทำให้เกิดลายเส้นจำนวนหลากหลาย ซึ่งลายเส้นต่าง ๆ เหล่านี้เกิดจากการเรียงตัวกันของพิกเซล โดยที่พิกเซลแต่ละพิกเซลมีข้อมูลประจำตำแหน่งคือค่าพิกัดบนแกน x และแกน y (รูปที่ 3 แสดงให้เห็นถึงการเรียงตัวของพิกเซล และค่าพิกัดบนระนาบแกน x และแกน y ของแต่ละพิกเซล) การเก็บข้อมูลลายมือเขียนดิจิทัลอาศัยการเก็บรวบรวมค่าพิกัดบนระนาบแกน x และแกน y ของแต่ละพิกเซลและนำข้อมูลค่าพิกัดเหล่านี้มาใช้ประมวลผลหาเส้นบรรทัด



รูปที่ 3 ลายเส้นเกิดจากการเรียงตัวกันของพิกเซล

3. เทคนิคการค้นหาเส้นบรรทัด

3.1 การค้นหาเส้นบรรทัดด้วยฮิสโตแกรม

3.1.1. ฮิสโตแกรม

ฮิสโตแกรม คือ รูปแบบของการนำเสนอข้อมูลแบบหนึ่ง ซึ่งแกนแนวดิ่งแสดงค่าของข้อมูลและแกนแนวนอนแสดงตำแหน่งค่า y นั้นๆ ของข้อมูล รูปที่ 5 นำเสนอฮิสโตแกรมของลายมือเขียนในรูปที่ 4 ด้วยสมการที่ 1 [2]

$$P(i) = \sum_j I(i, j) \quad (1)$$



รูปที่ 4 ตัวอย่างลายมือเขียน



รูปที่ 5 ฮิสโตแกรมนำเสนอข้อมูลจำนวนพิกเซลในแต่ละแถวของรูปที่ 4

3.1.2. การค้นหาเส้นบรรทัดด้วยฮิสโตแกรม

การหาเส้นบรรทัดสามารถหาได้จากฮิสโตแกรม (Projection Profile) [2] เมื่อทำการโปรเจกชันพบว่า ตำแหน่งเริ่มต้นของข้อมูลบนฮิสโตแกรมมักเป็นส่วนของเส้นบรรทัดบนหรือเส้นบรรทัดกลาง และตำแหน่งสิ้นสุดของข้อมูลบนฮิสโตแกรมมักเป็นส่วนของเส้นบรรทัดฐานหรือเส้นบรรทัดล่าง (ดังรูปที่ 5) การค้นหาเส้นบรรทัดจากฮิสโตแกรมสามารถทำได้โดยกำหนดค่าเทรชโวล์ (Threshold) ที่เหมาะสม (ค่าเทรชโวล์ที่เหมาะสมนั้นหาได้จากการทดลอง) เมื่อทำการโปรเจกชันในระนาบแนวนอนให้นำค่าเทรชโวล์ดังกล่าวตรวจสอบกับปริมาณพิกเซล ณ ตำแหน่งปัจจุบันและตำแหน่งถัดไปหากปริมาณพิกเซลของทั้งสองตำแหน่งมีค่ามากกว่าเทรชโวล์ที่กำหนดและตำแหน่งก่อนหน้ามีปริมาณพิกเซลที่ต่ำกว่าค่าเทรชโวล์ที่กำหนดสามารถประมาณการได้ว่าตำแหน่งดังกล่าวเป็นตำแหน่งของเส้นบรรทัด อีกกรณีหนึ่งหากตำแหน่งปัจจุบันและตำแหน่งก่อนหน้ามีปริมาณพิกเซลมากกว่าค่าเทรชโวล์ที่กำหนดรวมถึงตำแหน่งถัดไปมีปริมาณพิกเซลที่ต่ำกว่าค่าเทรชโวล์ที่กำหนดสามารถประมาณการได้ว่าตำแหน่งดังกล่าวเป็นตำแหน่งของเส้นบรรทัดได้เช่นกัน ซึ่งการค้นหาเส้นบรรทัดทั้งสองกรณีเกิดจากสมมติฐานที่ว่าตำแหน่งของเส้นบรรทัดบนฮิสโตแกรมมักเป็นตำแหน่งที่มีการเปลี่ยนแปลง ค่าความถี่ของปริมาณพิกเซลอย่างกะทันหัน

จากการวิเคราะห์สิ่งที่ได้จากการโปรเจกชันด้วยค่าเทรชโวล์ทำให้ได้เส้นบรรทัดจำนวนน้อยกว่าหรือเท่ากับ 4 เส้นด้วยกันได้แก่ 1. เส้นบรรทัดฐาน 2. เส้นบรรทัดกลาง 3. เส้นบรรทัดบน 4. เส้นบรรทัดล่าง

การค้นหาเส้นบรรทัดด้วยฮิสโตแกรมเป็นวิธีที่เรียบง่ายแต่มีข้อจำกัดอยู่ที่การกำหนดค่าเทรชโวล์ที่เหมาะสม ซึ่งค่าเทรชโวล์ที่เหมาะสมไม่มีกฎกำหนดที่ตายตัวแต่สามารถค้นหาได้จากการทดลอง จากสาเหตุดังกล่าวส่งผลให้ผลการค้นหาเส้นบรรทัดด้วยฮิสโตแกรมเกิด

ข้อผิดพลาดได้ง่าย อีกทั้งหากข้อมูลลายมือเขียนที่นำมาโปรเจกชันมีปริมาณของพิกเซลอันไม่ถึงประสงค์ (Noise) เป็นจำนวนมากส่งผลให้การค้นหาเส้นบรรทัดด้วยวิธีนี้เกิดข้อผิดพลาดได้เช่นกัน

3.2. การค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรม

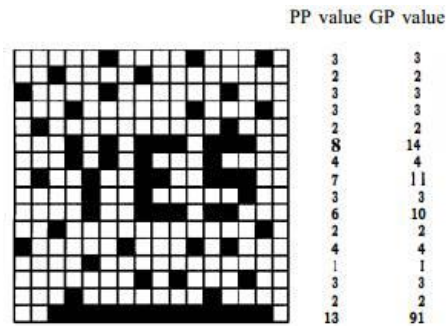
การค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรม (Generalised Projections) [2] เป็นวิธีการที่ประยุกต์มาจากวิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรม ทั้งนี้เพื่อแก้ปัญหาข้อผิดพลาดที่เกิดจากปริมาณของพิกเซลอันไม่ถึงประสงค์ที่เกิดขึ้นกับวิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรม

การค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรมใช้หลักการปรับปริมาณของจำนวนพิกเซลใหม่ของพิกเซลที่มีพิกเซลด้านข้างติดกันให้มีปริมาณที่มากขึ้น ทั้งนี้เพื่อแยกพิกเซลที่ต้องการกับพิกเซลอันไม่ถึงประสงค์ (Noise) ออกจากกันโดยชัดเจน การปรับปริมาณของจำนวนพิกเซลใหม่เป็นไปดังสมการที่ 2

$$GP(j) = \sum \mathfrak{N}(I(i,j)) \quad (2)$$

ภายหลังการปรับปริมาณของจำนวนพิกเซลใหม่ทำให้ปริมาณของพิกเซลที่มีพิกเซลด้านข้างติดกันมีปริมาณที่สูงขึ้นเมื่อเทียบกับวิธีการโปรเจกชันแบบดั้งเดิม (ดังหัวข้อที่ 3.1.1.) ซึ่งรูปที่ 6 แสดงการเปรียบเทียบระหว่างปริมาณพิกเซลของคำว่า YES ด้วยวิธี Projection Profile (คำนวณปริมาณพิกเซลด้วยสมการที่ 1) กับ Generalised Projections (คำนวณปริมาณพิกเซลด้วยสมการที่ 2)

การค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรมใช้หลักการการเดียวกันกับการค้นหาเส้นบรรทัดด้วยฮิสโตแกรม กล่าวคือภายหลังการสร้างฮิสโตแกรมด้วยวิธี Generalised Projections ให้กำหนดค่าเทรชโวล์ที่เหมาะสมและนำค่าเทรชโวล์ดังกล่าวไปหาเส้นบรรทัดตามหลักการค้นหาเส้นบรรทัดในหัวข้อที่ 3.1.2.



รูปที่ 6 เปรียบเทียบระหว่างปริมาณพิกเซลของคำว่า YES ด้วยวิธี Projection Profile กับ Generalised Projections

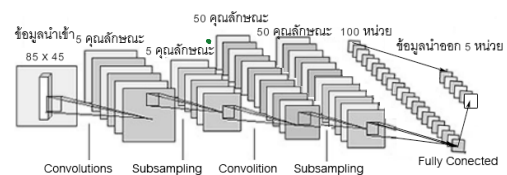
ถึงแม้วิธีการค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรม จะสามารถแก้ปัญหาข้อผิดพลาดที่เกิดจากปริมาณของพิกเซลอื่นไม่ถึงประสงค์ได้แต่วิธีการดังกล่าวยังมีข้อจำกัดในการกำหนดค่าเทรชโวล์ที่เหมาะสมเช่นเดียวกับการค้นหาเส้นบรรทัดด้วยฮิสโตแกรมดังหัวข้อที่ 3.1.

3.3. การ ค้นหาเส้นบรรทัด ด้วย โครงข่ายประสาทเทียม

3.3.1. โครงข่ายประสาทเทียม

โครงข่ายประสาทเป็นเทคนิคที่นิยมนำไปใช้สำหรับการรู้จำวัตถุต่าง ๆ ซึ่งโครงข่ายประสาทเป็นวิธีที่จำลองการทำงานมาจากสมองของมนุษย์ออกมาในลักษณะของการคำนวณทางคณิตศาสตร์ มีส่วนประกอบสำคัญ 2 ส่วน ได้แก่ 1. ส่วนนำเข้า (Input) 2. ส่วนนำออก (Output) หลักการทำงานของวิธีโครงข่ายประสาทให้นำส่วนนำเข้าคูณกับค่าน้ำหนัก (Weight) ต่าง ๆ ที่ถูกกำหนดขึ้น หากผลลัพธ์หรือค่านำออกที่ได้ออกมามีค่ามากกว่าค่าขีดแบ่ง (Threshold) ให้นำค่านำออกดังกล่าวส่งไปประมวลผลยังบัพ (Node) ถัดไปและทำเช่นนี้ไปเรื่อย ๆ จนครบทุกบัพ จะได้ผลลัพธ์ตามที่ต้องการ (ดังรูปที่ 7) ซึ่งค่าขีดแบ่ง

และค่าน้ำหนักที่ใช้ในการกำหนดขอบเขตของการทำงานของโครงข่ายประสาทเกิดจากการนำภาพผลลัพธ์ที่ทราบอยู่แล้วกับภาพที่ต้องการหาผลลัพธ์เข้าสู่กระบวนการโครงข่ายประสาทพร้อมทั้งบันทึกค่าความคลาดเคลื่อนที่เกิดจากความแตกต่างของผลลัพธ์ระหว่างภาพทั้งสองในแต่ละรอบการประมวลผล ทำเช่นนี้ไปเรื่อย ๆ จนกว่าความคลาดเคลื่อนจะอยู่ในจุดสมดุลหรือเริ่มเข้าสู่ค่าคงที่ค่าหนึ่ง กระบวนการหาค่าขีดแบ่งและค่าน้ำหนักดังกล่าวเราเรียกว่า “กระบวนการเรียนรู้ (Train)”

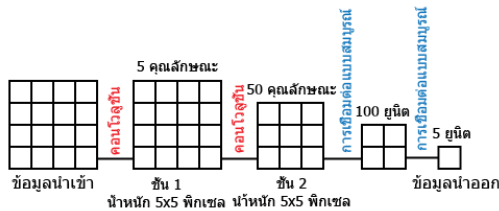


รูปที่ 7 ภาพจำลองการทำงานของโครงข่ายประสาท

ลักษณะของโครงข่ายประสาทแบบคอนโวลูชัน[4] สำหรับการรู้จำระดับของเส้นบรรทัด (Convolution Neural Network) เริ่มด้วยภาพที่มีขนาด 85 x 45 พิกเซล มาเชื่อมต่อกันเป็น 3 ชั้น (Layer) ชั้นแรกภาพประกอบด้วย 5 คุณลักษณะ (Feature) ชั้นที่สองภาพประกอบด้วย 50 คุณลักษณะ ชั้นที่สาม 100 หน่วย และสุดท้ายได้ข้อมูลนำออก 5 หน่วย จากรูปที่ 7[6] การเชื่อมต่อระหว่างชั้นด้วย convolution นั้นเปรียบเหมือนการคูณด้วยค่าถ่วงน้ำหนักขนาด 5x5 พิกเซลหากผลลัพธ์ใดมีค่าเกินค่าขีดแบ่ง จะเข้าสู่การเชื่อมต่อถัดไปซึ่งยังอยู่ในชั้นเดิมและเป็นการเชื่อมต่อแบบ subsampling ซึ่งเป็นการนำส่วนที่ผลลัพธ์เกินขีดแบ่งมาทำการสุ่มตัวอย่างออกมาให้ได้จำนวนคุณลักษณะเท่ากันสำหรับชั้นเดียวกัน ต่อจากนั้นเป็นการเชื่อมต่อสองชั้นสุดท้าย คือ การเชื่อมต่อแบบ fully connected และจะให้ผลลัพธ์สุดท้ายเป็นข้อมูล 5 ค่าที่จะนำไปใช้งานและอธิบายในส่วนต่อไป

3.3.2. ค้นหาจุดต่ำสุดและจุดสูงสุดของตัวอักษร ด้วยโครงข่ายประสาทแบบคอนโวลูชัน

การใช้โครงข่ายประสาทจำเป็นต้องกำหนดขนาดของข้อมูลนำเข้าและขนาดของลักษณะที่อยู่ในเลเยอร์ต่าง ๆ โดยปกติแล้วมักกำหนดให้ข้อมูลนำเข้ามีขนาดเท่ากับ 85×45 พิกเซล เชื่อมต่อกับชั้นที่ 1 และ 2 ด้วยคอนโวลูชัน ในแต่ละชั้นมีขนาดเท่ากับ 5 คุณลักษณะ และ 50 คุณลักษณะ ตามลำดับสองชั้นสุดท้าย เป็นการเชื่อมต่ออย่างสมบูรณ์ที่โดยเลเยอร์ที่ 3 มีขนาดเท่ากับ 100 หน่วย และขนาดของข้อมูลนำออกจะมีขนาดเท่ากับ 5 หน่วยซึ่งแต่ละหน่วยเป็นค่าเส้นล่าง, เส้นฐาน, เส้นอื่น ๆ, เส้นกลาง และเส้นบน ตามลำดับ [5] (ดังรูปที่ 8) ข้อมูลที่ได้จากโครงข่ายประสาทแบบคอนโวลูชันนี้เป็นค่าจุดต่ำสุดและจุดสูงสุดที่เชื่อถือได้ เราจะนำไปเข้าสมการเพื่อคำนวณเส้นบรรทัดออกมา ซึ่งจะอธิบายในหัวข้อถัดไป



รูปที่ 8 แสดงการทำงานของโครงข่ายประสาทแบบคอนโวลูชัน

3.3.3. ตรวจสอบเส้นบรรทัดของลายมือเขียนด้วยจุดต่ำสุดและจุดสูงสุดของตัวอักษร

จากขั้นตอนที่ผ่านมาทำให้ทราบถึงตำแหน่งที่ชัดเจนของจุดต่าง ๆ ซึ่งจุดสูงสุดหมายถึงจุดที่อยู่ในตำแหน่งเส้นกลางและเส้นบน จุดต่ำสุดหมายถึงจุดที่อยู่ในตำแหน่งเส้นฐานและเส้นล่างโดยที่จุดต่าง ๆ เหล่านี้มีค่าประจำจุดคือค่าพิกัดบนระนาบแกน x และแกน y ให้นำค่าพิกัดบนระนาบแกน x และแกน y ของทุก ๆ จุดเข้าสมการที่ (3) จะได้เส้นบรรทัดตามที่ต้องการ

$$y = ax + b \quad (3)$$

โดย

$$a = \frac{T_{xy} - \frac{1}{N} T_x T_y}{T_{xx} - \frac{1}{N} T_x T_x} \quad (4)$$

และ

$$b = \frac{1}{N} (T_y - T_x a) \quad (5)$$

ซึ่ง a คือ ค่าความชัน, b คือ ค่าคงตัว

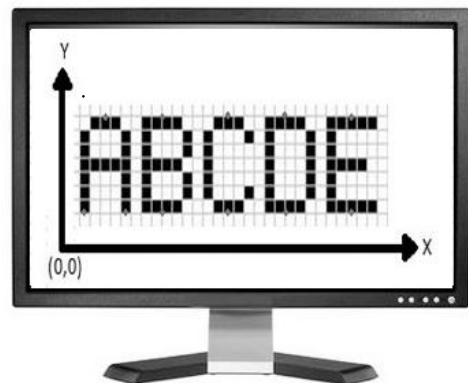
T_x คือ ผลรวมของค่าพิกัดแนวแกน x ทั้งหมด

T_y คือ ผลรวมของค่าพิกัดแนวแกน y ทั้งหมด

T_{xy} คือ ผลรวมของค่าพิกัดในแนวแกน x ซึ่งคูณกับค่าพิกัดในแนวแกน y ของจุดหนึ่ง ๆ ทั้งหมด

T_{xx} คือ ผลรวมของค่าพิกัดในแนวแกน x ซึ่งคูณกับค่าพิกัดในแนวแกน x ของจุดหนึ่ง ๆ ทั้งหมด

N คือ จำนวนของจุดทั้งหมด



รูปที่ 9 ค่าพิกัดของจุดในแนวแกน x และ y ที่นำมาคำนวณ



รูปที่ 10 ตัวอย่างเส้นบรรทัดที่ได้จากลายมือเขียน

4. บทสรุป

วิธีการค้นหาเส้นบรรทัดมีอยู่หลากหลายวิธี วิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรมเป็นหนึ่งในวิธีการค้นหาเส้นบรรทัดที่อาศัยการวิเคราะห์ข้อมูลจากฮิสโตแกรมโดยการกำหนดเทรตโซว์ที่เหมาะสมและนำค่าเทรตโซว์ดังกล่าวไปเปรียบเทียบกับจำนวนพิกเซลที่ถูกฉายบนฮิสโตแกรม วิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรมเป็นวิธีที่เรียบง่ายแต่มีข้อจำกัดในการกำหนดเทรตโซว์ที่ไม่มีกฎตายตัวซึ่งขึ้นอยู่กับผลการทดลองและหากข้อมูลที่นำมาโปรเจกชันมีปริมาณข้อมูลของพิกเซลอันไม่พึงประสงค์จำนวนมากจะส่งผลให้เกิดความผิดพลาดกับการค้นหาเส้นบรรทัดได้ง่าย

เพื่อลดข้อจำกัดเรื่องปริมาณข้อมูลของพิกเซลอันไม่พึงประสงค์จำนวนมากที่อาจก่อให้เกิดข้อผิดพลาดขึ้นกับวิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรม จึงเกิดวิธีใหม่ที่เรียกว่าการค้นหาเส้นบรรทัดด้วยการประยุกต์ใช้ฮิสโตแกรม ซึ่งเป็นวิธีที่นำหลักการเพิ่มความถี่ให้แก่พิกเซลที่มีพิกเซลด้านข้างติดกันเพื่อคัดแยกพิกเซลที่ต้องการออกจากพิกเซลอันไม่พึงประสงค์ให้ชัดเจนยิ่งขึ้น จากการเพิ่มความถี่ของพิกเซลที่มีพิกเซลด้านข้างติดกันช่วยลดปัญหาปริมาณข้อมูลของพิกเซลอันไม่พึงประสงค์จำนวนมาก แต่วิธีการดังกล่าวยังคงมีข้อจำกัดในการกำหนดเทรตโซว์ที่ไม่มีกฎตายตัวเช่นเดียวกับวิธีการค้นหาเส้นบรรทัดด้วยฮิสโตแกรม

การค้นหาเส้นบรรทัดด้วยโครงข่ายประสาทเทียมเป็นอีกหนึ่งวิธีสำหรับการค้นหาเส้นบรรทัดโดยไม่จำเป็นต้องกำหนดค่าเทรตโซว์และโปรเจกชัน การค้นหาเส้นบรรทัดด้วยโครงข่ายประสาทเทียมอาศัยการค้นหาจุดต่ำสุดของตัวอักษรและจุดสูงสุดของตัวอักษรและนำค่าคุณลักษณะต่าง ๆ ของจุดต่ำสุดและจุดสูงสุดของตัวอักษรไปเข้าสมการเพื่อหาเส้นบรรทัดออกมา วิธีการค้นหาเส้นบรรทัดด้วยโครงข่ายประสาทเทียมเป็นวิธีที่แม่นยำและมีประสิทธิภาพแต่ยังคงมีข้อจำกัดในเรื่องของการค้นหาจุดต่ำสุดและจุดสูงสุดของตัวอักษรที่ต้องอาศัยจำนวนข้อมูล

ตัวอย่างเป็นจำนวนมากจึงจะทำให้ผลลัพธ์ที่ได้ออกมามีความถูกต้องและแม่นยำ

เอกสารอ้างอิง

- [1] P. Y. Simard, D. Steinkraus, and M. Agrawala, "Ink normalization and beautification," , *ICDAR*, pp. 1182–1187, 2005.
- [2] R. Chamchong and C. Fung, "Segmentation of Thai handwritten text for automatic document retrieval," *PEECS2008*, 2008
- [3] G. Nicchiotti and C. Scagliola, "Generalised projections: a tool for cursive handwriting normalisation," *Proc. Fifth Int. Conf. Doc. Anal. Recognition. ICDAR '99 (Cat. No.PR00318)*, 1999.
- [4] C. M. Bishop, *Neural Networks for Pattern Recognition*, vol. 92. CLARENDONPRESS OXFORD, p. 482, 1995.
- [5] N. Chapman and J. Chapman, "Text and Typography," *Digital multimedia*, 3rd edition, 2009.
- [6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.