

บทที่ 2

หลักการและทฤษฎีที่เกี่ยวข้อง

การศึกษาอัลกอริทึมเรื่องการแยกตัวอักษรลายมือเขียนที่เขียนสัมผัสกันเป็นการหาจุดสัมผัสระหว่างตัวอักษรเพื่อหาจุดตัดในการแบ่งแยกตัวอักษรออกจากกันเพื่อให้ได้ตัวอักษรที่ต้องการและสมบูรณ์ ผู้วิจัยจึงได้ศึกษาอัลกอริทึมที่เหมาะสมกับการแยกตัวอักษรที่เขียนสัมผัส เช่นการนำโครงข่ายประสาทเทียม (neural network) ทำงานร่วมกับโปรแกรมโครงข่ายประสาทเทียม (neural network) ทำงานร่วมกับโปรแกรมโครงข่ายประสาทเทียม โดยขั้นตอนในการแยกตัวอักษรสัมผัสกันจะใช้วิธีการทำโปรแกรมโครงข่ายประสาทเทียมในแนวนอนก่อนเพื่อที่จะทำการแยกตัวอักษรระหว่างบรรทัด จากนั้นก็จะใช้วิธีการทำโปรแกรมโครงข่ายประสาทเทียมในแนวตั้งซึ่งจะได้ภาพที่เป็นฮิสโตแกรม จากนั้นทำการหาค่าโปรแกรมโครงข่ายประสาทเทียมในแนวตั้งซึ่งจะได้ภาพที่เป็นฮิสโตแกรม จากนั้นทำการหาค่าโปรแกรมโครงข่ายประสาทเทียมในแนวตั้งซึ่งจะได้ภาพที่เป็นฮิสโตแกรม จากนั้นทำการหาค่าโปรแกรมโครงข่ายประสาทเทียมในแนวตั้งซึ่งจะได้ภาพที่เป็นฮิสโตแกรม จากนั้นทำการหาค่าโปรแกรมโครงข่ายประสาทเทียมในแนวตั้งซึ่งจะได้ภาพที่เป็นฮิสโตแกรม

หลังจากนั้นทำให้ตัวอักษรที่ได้เป็นตัวพอมและแปลงเป็นข้อมูลเข้าสู่โครงข่ายประสาทเทียมเพื่อทำการพยากรณ์ว่าเป็นตัวอักษรอะไร โดยวิธีการดังกล่าวข้างต้นจะให้ความถูกต้องในการแยกตัวอักษรถึง 94.94 เปอร์เซ็นต์ งานวิจัยของ Rui และคณะ ได้นำเอาวิธี Drop-fall algorithm มาทำงานร่วมกับ Pattern-Oriented strategy ในการแยกตัวอักษรลายมือเขียนที่สัมผัสกัน ในขั้นแรกจะทำการจัดกลุ่มลักษณะของตัวอักษรสัมผัสกันออกเป็น 2 กลุ่มคือกลุ่มแรกการสัมผัสกันแบบง่าย (Simple touching) คือมีจุดที่ต้องทำการแยกเพียงที่เดียว (single-segment touching) หรือมีจุดที่สัมผัสกันเพียงจุดเดียว (single-point touching) กลุ่มที่สองคือกลุ่มที่มีจุดสัมผัสกันหลายจุด (multiple touching) วิธีการนี้จะเริ่มต้นโดยทำการหาพื้นที่หลังระหว่างตัวอักษรโดยใช้วิธีการ water reservoirs concept จากตัวเลขที่สัมผัสกันจะทำให้มีพื้นที่ในการรองรับน้ำไม่ว่าจะเป็นปลายเปิดด้านบน ด้านล่างหรือด้านข้าง จากนั้นนำเอาลักษณะของรูปแบบการสัมผัสกันมาทำการแยกโดยใช้ Descending-left algorithm, Descending-right algorithm, Ascending-left algorithm และ Ascending-right algorithm มาช่วยในการพิจารณาจุดแยกของตัวเลขซึ่งสามารถให้ความถูกต้องได้ถึง 95.60 เปอร์เซ็นต์ งานวิจัยของ Yang และคณะ ได้ทำการศึกษาและปรับปรุงระบบรู้จำตัวอักษรเขียนภาษาจีนที่มีลักษณะกลุ่มตัวอักษรที่ขาดจากกัน ลักษณะกลุ่มตัวอักษรที่ห่างกันและลักษณะ

กลุ่มตัวอักษรที่สัมพันธ์กัน โดยการนำทฤษฎีฟัซซีมาปรับใช้ โดยกลุ่มของตัวอักษรที่มีลักษณะที่ขาดจากกัน และกลุ่มตัวอักษรที่มีลักษณะตัวอักษรที่ห่างกันให้ทำการรวมตัวอักษรให้มีความถูกต้อง ในส่วนของกลุ่มตัวอักษรที่มีลักษณะที่สัมพันธ์กันให้ทำการแยกตัวอักษรออกจากกัน ขั้นแรกทำการแยกตัวอักษรแบบหยาบโดยใช้วิธีการทางสถิติมาช่วยในการหาขนาดความกว้างและความสูงของตัวอักษร ขั้นตอนต่อมาทำการจัดกลุ่มของตัวอักษรที่ได้ออกเป็นกลุ่มตัวอักษรที่ขาดจากกัน กลุ่มตัวอักษรที่ห่างกันและกลุ่มตัวอักษรที่สัมพันธ์กัน จากนั้นนำค่าเทรสโวลที่มีค่ามากที่สุดของแต่ละกลุ่มนำมาสร้างค่าความเป็นสมาชิกของแต่ละกลุ่ม (membership functions) และในขั้นตอนสุดท้ายเป็นการนำเอาข้อมูลที่ได้ทำการทดสอบโดยใช้ fuzzy threshold method และทำการตรวจสอบความถูกต้องในแก้ปัญหาดังกล่าวโดยผู้ทำงานวิจัย ซึ่งผลที่ได้มีการแก้ปัญหาดังกล่าวดีขึ้น งานวิจัยของ Alamri และคณะ ได้ทำการแยกตัวอักษรเขียนที่เป็นตัวเลขอาระบิกที่สัมพันธ์กันและในส่วนของความรู้จำได้ทำการเลือกใช้ทฤษฎีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) ขั้นตอนแรกเริ่มต้นด้วยการแยกตัวอักษรเขียนที่เป็นตัวเลขอาระบิกสัมพันธ์กันโดยหาขนาดความยาวทั้งหมดของตัวอักษรที่สัมพันธ์กันทั้งสองตัวเลข แทนด้วยตัวแปร W_f จากนั้นใช้วิธีการทางสถิติเข้ามาช่วยโดยทำการสร้างค่าพารามิเตอร์ขึ้นมาสองค่าคือพารามิเตอร์สำหรับความยาวตัวอักษรตัวแรกแทนด้วย (α) และพารามิเตอร์สำหรับความยาวระหว่างตัวอักษรตัวแรกและตัวที่สองแทนด้วย (β) จากนั้นทำการคำนวณความยาวของตัว อักษรตัวแรกจากสมการ $W_f = \alpha \cdot W_1$ และความยาวของตัวอักษรที่สองจากสมการ $W_s = W_1 - \alpha \cdot W_1 - \beta$ จากนั้นนำตัวอักษรที่ได้ไปเข้าสู่กระบวนการดึงคุณสมบัติของตัวอักษรและทำการเรียนรู้จำโดยใช้ทฤษฎีซัพพอร์ตเวกเตอร์แมชชีนทำให้ประสิทธิภาพในการรู้จำมีความถูกต้องถึง 98.48 เปอร์เซ็นต์เป็นต้น

ในงานวิจัยครั้งนี้ได้ศึกษาทฤษฎีพื้นฐานที่เกี่ยวข้องกับงานวิจัยได้แก่การประมวลผลภาพเบื้องต้น (Image Processing) การประมวลผลภาพถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหาเกี่ยวกับการมองเห็น ซึ่งการมองเห็นของมนุษย์เป็นสิ่งสำคัญและมีกลไกที่ซับซ้อนอย่างหนึ่ง การที่จะได้ผลลัพธ์ของการรู้จำให้มีประสิทธิภาพนั้นขั้นตอนนี้มีความสำคัญอย่างมากเพราะคุณภาพของภาพตัวอักษรที่จะนำไปเรียนรู้จำนั้น ถ้าคุณภาพของภาพตัวอักษรดีผลลัพธ์ที่ได้ก็ดี ถ้าคุณภาพของภาพตัวอักษร ไม่ดีผลลัพธ์ที่ได้ก็จะไม่ดีตามไปด้วย

2.1 การแปลงภาพ

การแปลงภาพเป็นขั้นตอนที่ใช้แปลงจากภาพสี ให้เป็นภาพขาวดำ การแปลงภาพสีเป็นภาพสีเทาทำได้โดยการแยกระดับสีแต่ละพิกเซลออกจากกัน ในรูปแบบภาพสีจากนั้นนำค่าจากภาพสีมาเข้าสู่สมการเพื่อคำนวณหาค่าสีเทา และนำค่าที่ได้ไปแทนในจุดพิกเซลเดิม

2.1.1 การแปลงภาพสี เป็นภาพสีเทา

การแปลงภาพสีเป็นภาพสีเทาคือกระบวนการในการแปลงภาพสีแต่ละพิกเซลให้กลายเป็นสีเทาและนำค่าที่ได้ไปแทนที่จุดพิกเซลเดิม โดยคำนวณได้จากสมการที่ 1 และสมการที่ 2 ดังต่อไปนี้

$$G' = 0.3R + 0.59G + 0.11B \quad (1)$$

หรือ

$$G' = \frac{R + G + B}{3} \quad (2)$$

โดยกำหนดให้

- G' คือค่าระดับสีเทา
- R คือค่าระดับสีแดง
- G คือค่าระดับสีเขียว
- B คือค่าระดับสีน้ำเงิน

2.1.2 การแปลงภาพสีเทาเป็นภาพขาวดำ

วิธีที่นิยมใช้แปลงภาพสีเทาเป็นภาพขาวดำนั้นวิธีการที่นิยมได้แก่ Otsu's method โดยอาศัยวิธีการทำเรดโซลด์ ซึ่งเป็นวิธีที่ใช้ในการแปลงภาพสีเทาให้อยู่ในรูปแบบของภาพไบนารีทำให้สามารถ

จำแนกวัตถุและพื้นหลังออกจากกันได้ จากฮิสโตแกรมภาพที่ 2.1 แสดงระดับความเข้มของภาพนั้นจะเห็นว่าส่วนที่เป็นวัตถุจะมีเข้มมากกว่าส่วนที่เป็นพื้นหลังทำให้สามารถแยกกลุ่มของวัตถุและพื้นหลังออกจากกันได้อย่างชัดเจน การเลือกค่าขีดโหดต้องทำการเลือกค่าที่มีระดับความเข้มอยู่ระหว่างกลุ่มที่เป็นวัตถุและกลุ่มที่เป็นพื้นหลัง

จากนั้นทำการเปรียบเทียบค่าของแต่ละจุดภาพหากค่า $f_{old}(i,j)$ น้อยกว่าค่าขีดโหดจุดภาพนั้นจะถูกปรับให้เป็นสีดำหรือส่วนของวัตถุ และหากค่า $f_{old}(i,j)$

$$f_{new}(i,j) = \begin{cases} 1 & \text{if } f_{old}(i,j) < T \\ 0 & \text{if } f_{old}(i,j) \geq T \end{cases} \quad (3)$$

โดยกำหนดให้

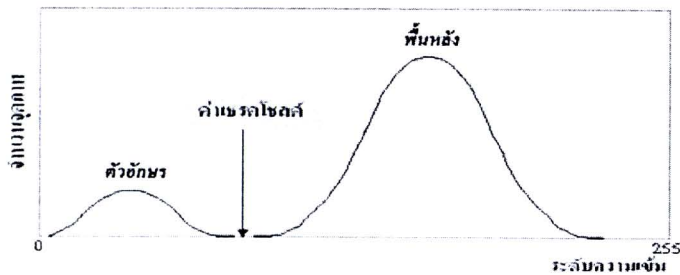
1 คือ สีดำ ซึ่งเป็นส่วนของตัวอักษร

0 คือ สีขาว ซึ่งเป็นส่วนของพื้นหลัง

$f_{new}(i,j)$ คือค่าใหม่ของภาพในลำดับ (i,j)

$f_{old}(i,j)$ คือค่าเก่าของภาพในลำดับ (i,j)

T คือค่าขีดโหดของภาพ



ภาพที่ 2.1 ฮิสโตแกรมแสดงค่าขีดรอยขีด

เมื่อได้ภาพสีขาวดำ จะนำภาพเหล่านั้นมาทำการตัดแยกบรรทัดและตัวอักษร ซึ่งเป็นวิธีการหนึ่ง
ในกระบวนการประมวลผลภาพ

2.2 สมการถดถอย (Linear Regression)

สมการถดถอยนั้นเป็นการดูค่าตัวแปร 2 ตัวแปรที่เป็นคู่ๆ กันว่ามีความสัมพันธ์กันในลักษณะใด
โดยการสร้างแผนภาพการกระจาย (Scatter Diagram) ระหว่างตัวแปร 2 ตัวนั้น ซึ่งภาพที่ได้นี้อาจจะเป็น
เส้นตรง, เส้นโค้ง, ไม่เป็นเส้นตรง ฯลฯ

ในกรณีที่ แผนภาพการกระจายนั้นเป็นเส้นตรง เราสามารถสร้างเป็นความสัมพันธ์ในรูปแบบ
ของสมการที่ 4 ได้ดังนี้

$$Y_i = \alpha + \beta X_i + \varepsilon_i \quad (4)$$

ซึ่งสมการดังกล่าวเรียกว่าสมการถดถอยอย่างง่าย

โดยกำหนดให้

Y_i คือค่าของตัวแปรตาม

α คือค่าของ Y_i เมื่อ X_i เป็นศูนย์ (Y- intercept)

β คือค่าความชัน (Slope) ของสมการ

ϵ_i คือค่าความคลาดเคลื่อนระหว่างค่าจริงกับค่าประมาณของ Y_i

วิธีหาค่า α และค่า β จะประมาณโดยวิธีกำลังสองน้อยที่สุด ซึ่งเป็นวิธีที่ได้รับความนิยมกันมาก โดยมีหลักการคือ จะทำให้ผลรวมกำลังสองของความแตกต่างระหว่าง Y_i และ Y ที่ได้จากการประมาณมีค่าต่ำที่สุดซึ่งสมการประมาณการจะมีรูปแบบดังสมการที่ 5

$$\hat{Y}_i = a + bX_i \quad (5)$$

โดยกำหนดให้

a คือค่า Y-intercept เป็นค่าประมาณของ α

b คือค่า Slope เป็นค่าประมาณของ β

\hat{Y}_i คือค่าประมาณการของ Y_i เมื่อตัวแปรอิสระคือ X_i

ค่า a สามารถหาได้จากสมการที่ 6

$$a = \bar{y} - b\bar{x} \quad \text{หรือ} \quad a = \frac{\sum Y_i}{n} - b \frac{\sum X_i}{n} \quad (6)$$

ค่า b สามารถหาได้จากสมการที่ 7 สมการที่ 8 และสมการที่ 9

$$b = \frac{S_{xy}}{S_{xx}} \quad (7)$$

$$S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n} \quad (8)$$

$$S_{xy} = \sum X_i Y_i - \frac{\sum X_i Y_i}{n} \quad (9)$$

2.3 เวกเตอร์

เวกเตอร์เป็นปริมาณที่กำหนดขึ้นเพื่อใช้อธิบายเกี่ยวกับขนาดและทิศทาง โดยจะต้องบอกทั้งขนาดและทิศทางจึงจะได้ความหมายสมบูรณ์ เช่น ความเร่ง ความเร็ว แรง โมเมนตัม เป็นต้น การหาผลลัพธ์ของปริมาณเวกเตอร์ ต้องอาศัยวิธีการทางเวกเตอร์ โดยต้องหาผลลัพธ์ทั้งขนาดและทิศทาง ขนาดเวกเตอร์แทนด้วยความยาวเส้นตรง ทิศทางเวกเตอร์แทนด้วยหัวลูกศร ใช้สัญลักษณ์ \overline{AB} แทนส่วนของเส้นตรงที่ระบุทิศทางจาก A ไป B อ่านว่า เวกเตอร์เอบี เรียก A ว่าจุดเริ่มต้น (initial point) ของ \overline{AB} เรียก B ว่าจุดสิ้นสุด (terminal point) ของ \overline{AB} หัวลูกศรจะอยู่ที่จุดสิ้นสุด ขนาดของเวกเตอร์ \overline{AB} คือความยาวของส่วนของเส้นตรง AB ใช้สัญลักษณ์ $|\overline{AB}|$ ขนาดนี้จะเป็นปริมาณสเกลาร์ อาจใช้สัญลักษณ์ u แทนเวกเตอร์ได้

การคูณเชิงสเกลาร์

การคูณเชิงสเกลาร์ของเวกเตอร์สองเวกเตอร์ผลลัพธ์ที่ได้เป็นจำนวนจริงไม่ใช่เวกเตอร์ เช่น เวกเตอร์ A เขียนแทนด้วย \vec{A} คูณกับเวกเตอร์ B เขียนแทนด้วย \vec{B} ในปริภูมิสองมิติผลคูณเชิงสเกลาร์ของ \vec{A} และ \vec{B} เขียนแทนด้วย $\vec{A} \cdot \vec{B}$ และคำนวณได้จากสมการที่ 10 สมการที่ 11 สมการที่ 12 และสมการที่ 13

$$\vec{A} \cdot \vec{B} = \|\vec{A}\| \|\vec{B}\| \cos \theta \quad (10)$$

โดยกำหนดให้

$$\vec{A} \cdot \vec{B} = (x_a \times x_b) + (y_a \times y_b) \quad (11)$$

$$\|\vec{A}\| = \sqrt{x_a^2 + y_a^2} \quad (12)$$

$$\|\vec{B}\| = \sqrt{x_b^2 + y_b^2} \quad (13)$$

ดังนั้นการหามุมระหว่างเวกเตอร์สองเวกเตอร์จะได้ดังสมการที่ 14

$$\theta = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| \|\vec{B}\|} \quad (14)$$