บทที่ 2

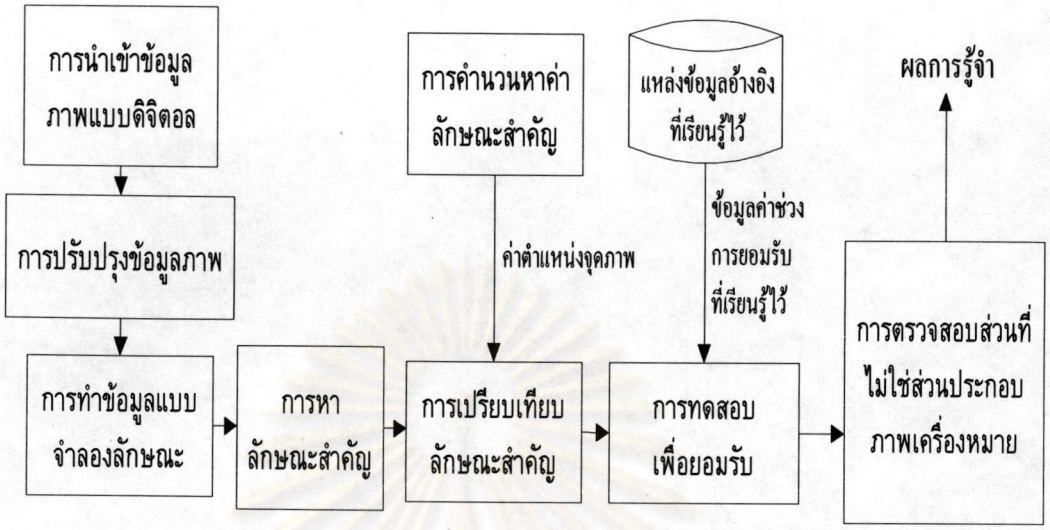
แนวคิดและทฤษฎี

จากที่ได้กล่าวมาแล้วในความเป็นมาของปัญหาการนำวิธีการประมวลผลภาพดิจิตอลมาใช้นั้น จะทำให้มีวิธีการที่แตกต่างไปจากการตรวจคำตอบแบบที่ใช้แสงความถี่ต่ำกว่าสีแดง โดยนำข้อมูลภาพแบบดิจิตอลที่ได้จากเครื่องสแกนเนอร์ ซึ่งจะเป็นข้อมูลค่าความเข้มของสีเทาของแต่ละจุดภาพของเครื่องหมายคำตอบมาวิเคราะห์ในรายละเอียดซึ่งจะมากกว่าการวัดค่าแสงความถี่ต่ำกว่าสีแดงจากภาพเครื่องหมายโดยตรงจากการที่มีข้อมูลค่าความเข้มของสีเทาของแต่ละจุดภาพจำนวนมาก ทำให้สามารถแบ่งแยกการพิจารณาเป็นส่วน ๆ และใช้เทคนิคการรู้จำในการวิเคราะห์ว่ามีภาพการทำเครื่องหมายอยู่ที่ตำแหน่งใดของกระดาษคำตอบ

2.1 การรู้จำภาพเครื่องหมาย

ทฤษฎีเบื้องต้นของการรู้จำ โดยธรรมชาติของมนุษย์เราแล้วเมื่อได้พบเห็นหรือรับฟังสิ่งใดก็ตามจะสามารถเรียนรู้และจดจำสิ่งนั้นไว้ จากนั้นเมื่อได้พบเห็นได้รับฟังสิ่งเดิมที่เคยเรียนรู้ไว้อีกครั้งหนึ่งสามารถบอกได้ว่าคือสิ่งเดียวกันหรือไม่ ซึ่งเรียกว่าการรู้จำ (recognition)[10] จากหลักการที่กล่าวมาได้นำมาใช้กับการรู้จำภาพเครื่องหมายคำตอบโดยการตรวจลักษณะข้อมูลภาพที่ต้องการวิเคราะห์ว่าเป็นภาพที่มีค่าคุณสมบัติเฉพาะใกล้เคียงหรือตรงกับข้อมูลอ้างอิงที่ได้เรียนรู้ไว้หรือไม่ และมีการกำหนดค่าช่วงการยอมรับ (tolerance band for acceptance)[12] เอาไว้เพื่อแยกภาพที่ถูกต้อง การรู้จำภาพเครื่องหมายสามารถแบ่งออกเป็นขั้นตอนใหญ่ได้ 7 ขั้นตอนได้แก่

- การนำเข้าข้อมูลภาพแบบดิจิตอล
- การปรับปรุงข้อมูลภาพ
- การทำข้อมูลแบบจำลองลักษณะ
- การหาคุณสมบัติ
- การเปรียบเทียบคุณสมบัติ
- การทดสอบเพื่อการยอมรับ
- การตรวจสอบส่วนที่ไม่ใช่ส่วนประกอบภาพเครื่องหมาย



รูปที่ 2.1 ผังการทำงานของการรู้จำภาพเครื่องหมาย

2.2 ขั้นตอนการรู้จำภาพเครื่องหมาย

2.2.1 การนำเข้าข้อมูลภาพแบบดิจิตอล

ใช้เครื่องสแกนเนอร์ในการแปลงข้อมูลภาพจากข้อมูลความเข้มแสงแต่ละจุดภาพให้มาอยู่ในรูปแบบของฟังก์ชันความเข้มแสง (image function)[8] ที่จัดแถวลำดับ (array) หรือเมตริกซ์ (matrix) มีระยะห่างเท่า ๆ กันโดยมีแถวแนวนอน (row) และแถวแนวตั้ง (column) เป็นตัวบ่งชี้ตำแหน่งโดยในแต่ละตำแหน่งจะมี image element เรียกชื่อย่อว่า พิกเซล (pixel) ซึ่งค่าของพิกเซลเป็นการแทนค่าระดับความเข้มของสีเทา (gray level) ของจุดที่อยู่ตรงตำแหน่งนั้น ภาพที่ได้นี้เรียกว่า ภาพดิจิตอล (digital image)

2.2.2 การปรับปรุงภาพข้อมูลภาพ

ข้อมูลที่ได้จากขั้นตอนการนำเข้าจะมีค่าระดับความเข้มของสีเทาอยู่หลายระดับทั้งนี้เพื่อช่วยวิเคราะห์ในกรณีที่ภาพเครื่องหมายมีสีจางโดยการพิจารณาทางด้านจำนวนของข้อมูลที่มีค่าระดับความเข้มของสีเทาใกล้เคียงกัน แต่ในการวิเคราะห์เพื่อการรู้จำข้อมูลภาพเครื่องหมายต้องการข้อมูลสองกลุ่มเท่านั้นคือ กลุ่มข้อมูลที่แสดงค่าของภาพพื้น (background image) และกลุ่มข้อมูลที่แสดงค่าของภาพที่สนใจ (object image) ซึ่งต้องมีการเลือกใช้ค่าขีดแบ่ง (threshold)[4] ที่ได้จากการคำนวณเพื่อให้ได้ค่ากึ่งกลางระหว่างกลุ่มข้อมูลส่วนใหญ่สองกลุ่มดังกล่าว มาเป็นตัวแบ่งกลุ่มของข้อมูลภาพ และ

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

หลังจากการแบ่งกลุ่มของข้อมูลภาพแล้ว สามารถช่วยในด้านการกำจัดสัญญาณรบกวนและลดสัญญาณรบกวนที่เกิดขึ้นได้ในกระบวนการอ่านข้อมูลของเครื่องสแกนเนอร์ และในขณะที่ทำการส่งข้อมูลมายังเครื่องไมโครคอมพิวเตอร์ ซึ่งการรบกวนนั้นมีลักษณะเป็นแบบมีค่าความเข้มแสงแตกต่างจากพิกเซลข้างเคียง

2.2.3 การทำข้อมูลแบบจำลองลักษณะ (model representation)[12]

การทำข้อมูลแบบจำลองลักษณะคือ การหาค่าของข้อมูลภาพที่สามารถทดแทนลักษณะโดยรวมของภาพเช่น ค่าตำแหน่งของเส้นรอบรูป ค่าตำแหน่งที่เป็นแกนกลางของภาพ ค่าตำแหน่งจุดปลายของภาพ ค่าตำแหน่งจุดกึ่งกลางของภาพ ค่าตำแหน่งที่เป็นขอบเขตสูงสุดและต่ำสุดของภาพเป็นต้น ซึ่งค่าข้อมูลเหล่านี้นำไปเป็นค่าที่ใช้เพื่อวัดค่าลักษณะสำคัญของภาพเครื่องหมาย และนำผลที่ได้ไปวิเคราะห์เพื่อการตัดสินใจในขั้นตอนถัดไป นอกจากนี้ค่าข้อมูลแบบจำลองลักษณะยังนำมาใช้เพื่อการแสดงผลลัพธ์ของการรู้จำ โดยใช้ค่าตำแหน่งจุดกึ่งกลางของภาพมาเป็นจุดที่บอกว่าเป็นตำแหน่งที่ทำการรู้จำได้

2.2.4 การหาข้อมูลลักษณะสำคัญ (feature detection)[12]

เนื่องจากข้อมูลที่ได้มาจากขั้นตอนการทำแบบจำลองลักษณะยังมีจำนวนมาก ในการวิเคราะห์เพื่อหาข้อมูลลักษณะสำคัญมีความต้องการ การแบ่งแยกเพื่อให้สามารถประมวลผลในแต่ละส่วน และนำค่าที่ได้มาเป็นข้อมูลเพื่อใช้ในการตัดสินใจ ซึ่งสามารถแบ่งการพิจารณาเพื่อหาข้อมูลลักษณะสำคัญนี้ออกได้ 3 กลุ่มดังนี้

1. ลักษณะทั่วไป (global feature)
2. ลักษณะพื้นฐาน (local feature)
3. ลักษณะความสัมพันธ์ (relation feature)

ข้อมูลลักษณะทั่วไปได้แก่ ขนาด ข้อมูลความเข้มของสีเทา และค่าพื้นที่ของภาพเครื่องหมายเป็นต้น ซึ่งข้อมูลส่วนนี้แสดงลักษณะโดยรวมของภาพเครื่องหมายที่ทำการวิเคราะห์ โดยที่ค่าข้อมูลเหล่านี้สามารถคำนวณได้จากข้อมูลในแต่ละพิกเซล เช่น ค่าความเข้มของสีเทา ค่าตำแหน่งของพิกเซลในแนวตั้งและแนวนอน การหาข้อมูลขนาดของภาพเครื่องหมายทำได้โดยใช้ค่าตำแหน่งที่เป็นจุดเริ่ม-ต้นซึ่งจะมีค่าตำแหน่งน้อยที่สุดในกลุ่มข้อมูลของภาพเครื่องหมายนั้น ๆ และค่าตำแหน่งที่จุดสิ้นสุดซึ่งมีค่าตำแหน่งมากที่สุดในกลุ่มข้อมูลของภาพเครื่องหมายนั้น ๆ จากค่าระยะตำแหน่งที่ได้ทำให้สามารถ

คำนวณหาค่าขนาดของภาพเครื่องหมายได้ ส่วนในการหาค่าพื้นที่ของภาพเครื่องหมายทำได้โดยใช้การนับจำนวนของพิกเซลที่มีค่าเป็น 1 ซึ่งได้หลังจากการปรับปรุงข้อมูลภาพและเป็นพิกเซลที่อยู่ภายในเส้นรอบรูปของภาพเครื่องหมาย โดยกำหนดให้ 1 พิกเซลมีพื้นที่เท่ากับ 1 หน่วย

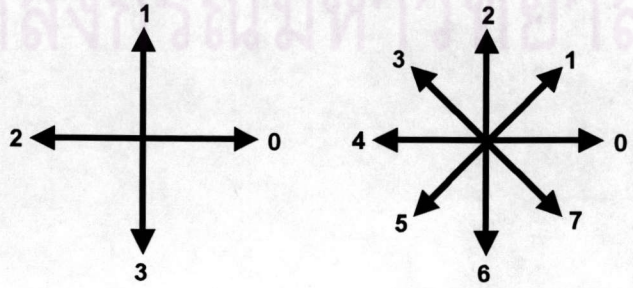
ข้อมูลลักษณะพื้นฐานได้แก่ เส้นรอบวง เส้นรอบรูป เส้นแกนกลาง ทิศทาง ความเอียงเป็นต้น การวิเคราะห์มีทฤษฎีที่เกี่ยวข้องได้แก่ รหัสลูกโซ่ (chain code)[10] และการหาค่า ความเอียง (slope)[12] ซึ่งการหาค่าความเอียงนี้สามารถนำไปคำนวณหา ค่าตำแหน่งของจุดภาพที่เป็น ส่วนทดแทนเส้นตรงใด ๆ (poly line representation)[12] หรือเรียกอีกอย่างหนึ่งว่า ค่าตำแหน่งของจุดภาพที่เป็นส่วนประกอบของเส้นตรง โดยทฤษฎีทั้งสองจะกล่าวรายละเอียดในตอนถัดไป

ข้อมูลลักษณะความสัมพันธ์ได้แก่ อัตราส่วนต่าง ๆ เช่น อัตราส่วนพื้นที่ต่อความยาวเส้นรอบวงของวงกลมเป็นต้น ซึ่งการหาความสัมพันธ์นี้จะคำนวณได้จากสมการ 2.1

$$\frac{4\pi * \text{พื้นที่ของวงกลม}}{(\text{เส้นรอบวง})^2} = 1 \qquad (2.1)$$

จากผลของความสัมพันธ์ดังกล่าว สามารถนำมาประยุกต์ใช้ในการวิเคราะห์ความสมบูรณ์ของรูปวงกลมโดยดูจากค่าความผิดพลาดที่แตกต่างไปจากค่า 1 และพิจารณาแบบค่าสัมบูรณ์ (absolute)

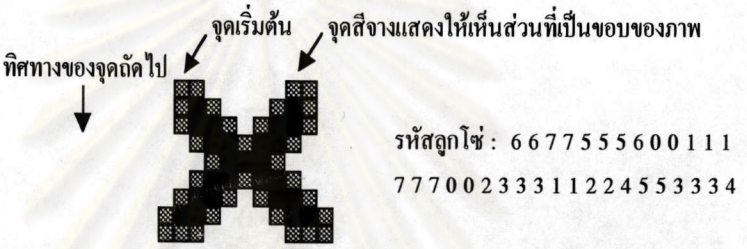
รหัสลูกโซ่ คือ ค่ารหัสที่ใช้แสดงค่าแทนจุดภาพที่มีตำแหน่งเป็นเส้นรอบรูป โดยรหัสลูกโซ่จะทดแทน ทิศทางทีละจุดภาพโดยต่อเนื่องไปจนครบเส้นรอบรูปของภาพเครื่องหมาย รหัสลูกโซ่โดยทั่วไปจะมีอยู่ 2 แบบ คือ แบบ 4 ทิศทางและแบบ 8 ทิศทางดังรูปที่ 2.2



(ก) แบบ 4 ทิศทาง (ข) แบบ 8 ทิศทาง

รูปที่ 2.2 ค่าตัวเลขของรหัสลูกโซ่

รหัสลูกโซ่นี้นิยมใช้ในการประมวลผลภาพแบบดิจิตอล โดยนำมาเป็นข้อมูลทดแทนในส่วนที่เป็นขอบของภาพที่สนใจ ซึ่งมีข้อดีคือง่ายต่อการคำนวณหาค่าทิศทางของเส้นภาพ และการคำนวนหาค่าความยาวเส้นรอบรูป ส่วนข้อเสียคือ ในกรณีที่ภาพมีขนาดใหญ่ทำให้มีค่ารหัสที่ยาวมากซึ่งการใช้วิธีการทดแทนแบบอื่น ๆ เช่น ตำแหน่งจุดเปลี่ยนมุมจะทำให้ได้ข้อมูลที่มีจำนวนน้อยกว่า และในกรณีที่ภาพมีข้อมูลที่เป็นจุดภาพรบกวนมาก ทำให้การเข้ารหัสลูกโซ่ได้ค่าตัวเลขรหัส ของส่วนที่เป็นสิ่งรบกวนปะปนไปด้วย มีผลทำให้การวิเคราะห์ผิดพลาดได้



รูปที่ 2.3 รหัสลูกโซ่แบบ 8 ทิศทางของภาพเครื่องหมายกากบาท

จากรูปที่ 2.3 แสดงตัวอย่างการเข้ารหัสลูกโซ่แบบ 8 ทิศทางของภาพเครื่องหมายกากบาทโดยจุดตั้งต้นเริ่มจากจุดมุมบนซ้ายของขอบภาพ และใช้ค่ารหัสที่มีตัวเลขมากที่สุดในทิศทางทวนเข็มนาฬิกาซึ่งจะได้รหัส 6 6 7..... 4 ตามลำดับ จากการได้รหัสลูกโซ่สามารถนำมาคำนวณหาค่าระยะเส้นรอบรูปได้ โดยกำหนดให้ค่าทิศทาง 0 2 4 6 มีค่าระยะทางเท่ากับ 1 หน่วยและในทิศทาง 1 3 5 7 มีค่าระยะทางเท่ากับ 1.414 หน่วยหรือเท่ากับ $\sqrt{2}$ ซึ่งจะเป็นไปตามวิธีทางคณิตศาสตร์ในการหาความยาวด้านของรูปสามเหลี่ยมที่มีด้านตรงข้ามมุมและด้านประชิดมุมยาว 1 หน่วย

2.2.5 <u>การเปรียบเทียบลักษณะสำคัญ (feature matching)</u>[12]

การเปรียบเทียบลักษณะสำคัญทำโดยนำค่าที่ได้จากการวิเคราะห์หาข้อมูลลักษณะสำคัญในด้านตำแหน่งจุดภาพ มาเปรียบเทียบกับข้อมูลลักษณะสำคัญที่ได้จากการคำนวนซึ่งค่านี้ใช้วิธีทางคณิตศาสตร์ในการหาค่าตำแหน่งจุดภาพเช่น ตำแหน่งของจุดภาพของวงกลมซึ่งสามารถคำนวณได้ดังสมการที่ 2.2 และสมการที่ 2.3 ส่วนการคำนวนหาค่าตำแหน่งของจุดภาพของเส้นตรง จะกล่าวในรายละเอียดและวิธีการคำนวนในหัวข้อ 2.5 ในการเปรียบเทียบข้อมูลนั้นจะทำการหาค่าผลความแตกต่าง

เพื่อนำใช้ในขั้นตอนการทดสอบเพื่อการยอมรับ ซึ่งค่าผลความแตกต่างก็คือค่าความผิดพลาด (error) และในการเปรียบเทียบลักษณะสำคัญมีทฤษฎีที่เกี่ยวข้องได้แก่

การคำนวณหาตำแหน่งของจุดภาพของวงกลมใช้การแทนค่ามุมของจุดภาพของวงกลมโดยเริ่มตั้งแต่ 0 องศาจนถึง 360 องศาในสมการ 2.2 และ 2.3 ทำให้สามารถหาค่าตำแหน่งในแนวระนาบ xy ที่ประกอบเป็นภาพวงกลมในกรณีที่ค่ารัศมีในแนวแกน x และรัศมีในแนวแกน y เท่ากัน ส่วนในกรณีที่ค่ารัศมีในแนวแกน x และแนวแกน y ไม่เท่ากันจะทำให้ได้ตำแหน่งจุดภาพที่วงรี

$$x = rx * \cos\theta + x \qquad (2.2)$$

$$y = ry * \sin\theta + y \qquad (2.3)$$

โดยที่

x, y = ตำแหน่งของจุดภาพของวงกลมในแนวแกน X และ Y

rx = รัศมีในแนวแกน X

ry = รัศมีในแนวแกน Y

θ = มุมของจุดภาพของวงกลม

x = ตำแหน่งในแนวแกน X ของจุดศูนย์กลางของวงกลม

y = ตำแหน่งในแนวแกน Y ของจุดศูนย์กลางของวงกลม

การหาค่า ระยะห่าง (distance)[12] ของตำแหน่งจุดสองจุดที่ในระนาบ xy ซึ่งใช้เพื่อการหาค่าความแตกต่างของตำแหน่งภาพ โดยสามารถคำนวณได้ตามสมการที่ 2.4

$$D = \sqrt{(x_i - x_n)^2 + (y_i - y_n)^2} \qquad (2.4)$$

โดยที่

D = ระยะห่างของจุดสองจุดที่ในระนาบ xy

x_i = ตำแหน่งในแนวแกน X ของจุดตั้งต้น

x_n = ตำแหน่งในแนวแกน X ของจุดปลาย

y_i = ตำแหน่งในแนวแกน Y ของจุดตั้งต้น

y_n = ตำแหน่งในแนวแกน Y ของจุดปลาย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.2.6 การทดสอบเพื่อยอมรับ (acceptance test)

การทดสอบเพื่อยอมรับคือ การวิเคราะห์ค่าความผิดพลาดที่ได้จากขั้นตอนการเปรียบเทียบลักษณะสำคัญนำมาเปรียบเทียบกับข้อมูลค่าความผิดพลาดที่ได้เรียนรู้และเก็บเป็นข้อมูลอ้างอิงไว้ ว่ามีค่าอยู่ในช่วงที่ยอมรับได้ หรือนอกเหนือจากช่วงที่ยอมรับได้ ในกรณีที่มีค่าความผิดพลาดอยู่ในช่วงที่ยอมรับได้ก็จะสามารถให้ผลการตัดสินใจในขั้นแรกได้ว่า ข้อมูลภาพที่นำมาวิเคราะห์นั้นเป็นข้อมูลของภาพที่มีลักษณะสำคัญที่ตรงกันกับลักษณะสำคัญของข้อมูลภาพที่ได้เรียนรู้ไว้ โดยการหาค่าที่จะนำมากำหนดเป็นค่าช่วงการยอมรับได้เพื่อใช้ในการตัดสินใจ มีขั้นตอนดังต่อไปนี้

1. การเก็บตัวอย่างข้อมูลภาพเครื่องหมายเพื่อนำมาวัดค่าลักษณะสำคัญของภาพ
2. การหาค่าสูงสุดและต่ำสุดของค่าลักษณะสำคัญของภาพเพื่อนำมากำหนดเป็นค่าช่วงการยอมรับได้
3. การปรับค่าช่วงการยอมรับได้ ในกรณีของค่าที่ได้มีการแปรผันตรงกับขนาดของภาพให้อยู่ในรูปแบบที่ไม่แปรผัน โดยการใช้เป็นแบบอัตราร้อยละ

การเก็บตัวอย่างข้อมูลภาพเครื่องหมายทำโดยออกแบบฟอร์มเพื่อให้ผู้ที่ทำการให้ตัวอย่างเขียนเครื่องหมายที่สมบูรณ์ด้วยลายมือในแต่ละชนิดของเครื่องหมายด้วยจำนวนที่เหมาะสม โดยดูจากค่าความแปรปรวนและค่าเบี่ยงเบนมาตรฐานของค่าลักษณะสำคัญของภาพในกลุ่มข้อมูลตัวอย่าง ว่าค่าทางสถิติทั้งสองนั้นมีความใกล้เคียงกันในแต่ละกลุ่มข้อมูลตัวอย่างที่ทำการเก็บข้อมูล โดยกลุ่มของข้อมูลใช้การแบ่งแยกหลาย ๆ แบบเช่น การแบ่งเป็นแต่ละคนของผู้ที่ทำการให้ตัวอย่าง การแบ่งตามจำนวนภาพเครื่องหมาย ซึ่งค่าที่ได้ถ้าใกล้เคียงกันจะหมายถึง จำนวนของตัวอย่างนั้นเพียงพอที่จะนำมาใช้เป็นข้อมูลอ้างอิง จากนั้นจะเป็นขั้นตอนการหาค่าเฉลี่ยของค่าลักษณะสำคัญ ซึ่งการหาค่าเฉลี่ยนี้ทำโดยการค้นหาค่าลักษณะสำคัญที่มีค่าสูงที่สุดและต่ำที่สุดจากข้อมูลตัวอย่างทั้งหมด เพื่อนำมากำหนดเป็นค่าช่วงการยอมรับได้ ค่าที่ได้จากขั้นตอนนี้ในบางค่าลักษณะสำคัญจะมีการแปรผันตรงกับขนาดของภาพซึ่งทำให้มีความต้องการในการปรับค่าช่วงการยอมรับได้ให้อยู่ในรูปแบบที่ไม่แปรผันกับขนาดของภาพโดยการใช้ค่าช่วงการยอมรับได้เป็นแบบอัตราร้อยละ

2.2.7 การตรวจสอบส่วนที่นอกเหนือจากส่วนประกอบของภาพเครื่องหมาย

จากขั้นตอนทั้งหมดที่กล่าวมา ทำให้พอสรุปได้ว่า ถ้าข้อมูลภาพที่สามารถผ่านการทดสอบเพื่อการยอมรับได้หมายความว่าข้อมูลภาพเครื่องหมายที่วิเคราะห์นั้นเป็นภาพเดียวกันกับที่ได้เรียนรู้ไว้ แต่การวิเคราะห์ทั้งหมดที่ผ่านมา จะยังไม่สามารถครอบคลุมกรณีที่ภาพเครื่องหมายมีส่วนจุดภาพที่นอกเหนือปะปนอยู่มากเกินไปจนทำให้ภาพเครื่องหมายนั้นไม่สมบูรณ์ ดังตัวอย่างเช่นการวิเคราะห์

ภาพเครื่องหมายกากบาทด้วยขั้นตอนที่กล่าวมา ผลลัพธ์ที่ได้จากการวิเคราะห์ภาพสี่เหลี่ยมทึบในรูปที่ 2.4 (ก) และภาพเครื่องหมายกากบาทในรูปที่ 2.4 (ข) มีค่าที่เท่ากันระหว่างภาพทั้งสอง เหตุที่เป็นเช่นนี้เนื่องจากข้อมูลภาพสี่เหลี่ยมทึบมีจุดภาพที่ครอบคลุมจุดภาพของเครื่องหมายกากบาทโดยครบถ้วน สำหรับปัญหานี้จะใช้วิธีการวิเคราะห์ส่วนที่ไม่ใช่ส่วนประกอบของภาพเครื่องหมายซึ่งทำโดยตรวจดูจุดข้อมูลภาพที่เป็นส่วนที่นอกเหนือจากส่วนประกอบของภาพเครื่องหมายว่ามีเกินกว่าจำนวนที่ได้เรียนรู้ไว้หรือไม่ และถ้าในกรณีผลการวิเคราะห์มีจุดภาพเกิน การตัดสินใจในขั้นตอนสุดท้ายนี้จะถือว่าภาพที่ทำการรู้จำนี้ไม่ใช่ภาพเครื่องหมายที่ถูกต้อง ทั้งนี้การวิเคราะห์ส่วนที่ไม่ใช่ส่วนประกอบของภาพเครื่องหมายจะทำในทุกส่วนที่เป็นช่องว่าง โดยการนับจำนวนจุดภาพที่มีค่าเป็น 1 และเป็นจุดภาพที่ไม่ใช่ส่วนของภาพเครื่องหมายซึ่งได้จากการลบข้อมูลภาพ ในส่วนของการตรวจสอบส่วนที่นอกเหนือจากส่วนประกอบของภาพเครื่องหมายอื่น ๆ เช่น เครื่องหมายวงกลม เครื่องหมายถูก ก็จะทำด้วยหลักการเดียวกัน

(ก.) ภาพสี่เหลี่ยมทึบ (ข.) ภาพเครื่องหมายกากบาทตัวอย่าง

รูปที่ 2.4 แสดงพื้นที่ช่องว่างที่ต้องการตรวจสอบส่วนที่นอกเหนือจากส่วนประกอบภาพ

การลบข้อมูลภาพ (image subtraction)[10] มีวิธีการคือ ใช้ค่าความเข้มของจุดภาพที่ตำแหน่งเดียวกันของภาพที่เป็นต้นแบบลบกับค่าความเข้มของจุดภาพที่เป็นภาพที่ต้องการวิเคราะห์ เพื่อให้เหลือเฉพาะส่วนต่างดังสมการที่ 2.5

$$g(x,y) = i(x,y) - b(x,y) \qquad (2.5)$$

โดยที่

x,y = ตำแหน่งจุดภาพในระนาบ xy

g(x,y) = ข้อมูลภาพที่เป็นผลลัพธ์จากการลบภาพ

i(x,y) = ข้อมูลภาพที่เป็นภาพเครื่องหมาย ที่ทำการวิเคราะห์

b(x,y) = ข้อมูลภาพที่เป็นภาพพื้นของกระดาษคำตอบ

2.3 การแบ่งแยกกลุ่มข้อมูลตามความเข้มแสงของภาพ

การปรับปรุงข้อมูลภาพเพื่อแบ่งกลุ่มข้อมูลตามค่าความเข้มแสงของข้อมูลภาพด้วยวิธีของ Otsu 1979[4] จะใช้การคำนวณหาค่าที่เป็นจุดเชื่อมของข้อมูลที่พิจารณาจาก ฮีสโตแกรมของค่าเข้มของสีเทา (gray level histogram)[4] สองกลุ่มในด้านต่ำสุด เพื่อใช้เป็นจุดแบ่งระหว่างกลุ่มข้อมูลภาพพื้นหลังของกระดาษคำตอบ และกลุ่มข้อมูลภาพเครื่องหมายที่สนใจและต้องการวิเคราะห์ ซึ่งมีขั้นตอนการคำนวณดังต่อไปนี้

ขั้นตอนแรก ทำการหาค่าผลรวมของแต่ละพิกเซลของข้อมูลภาพเครื่องหมายที่มีระดับความเข้มของแสงสีเทาที่เท่ากัน ดังค่าตัวอย่างในตารางที่ 2.1 สดมภ์ที่ 2 จากนั้นทำการรวมค่าผลบวกต่อเนื่องของจุดข้อมูลภาพที่ได้ใน สดมภ์ที่ 3 โดยมีผลลัพธ์เป็นค่าใน สดมภ์ที่ 4 ซึ่งวิธีการคำนวณสามารถทำได้ดังสมการที่ 2.6

$$t(g) = \sum_{i=0}^{g} f(i) \qquad (2.6)$$

โดยที่

t(g) = ผลรวมต่อเนื่องของข้อมูลภาพที่มีค่าความเข้มแสง จากค่า 0 ถึง g

g = ระดับความเข้มของแสงสีเทา

f(i) = จำนวนพิกเซลในแต่ละ ระดับความเข้มของแสงสีเทา i

i = ระดับความเข้มของแสงสีเทาตั้งแต่ 0 ถึงค่า g

ขั้นตอนที่สอง ทำการหาค่าเฉลี่ยของระดับความเข้มแสงสีเทาในแต่ละพิกเซลที่มีค่าระดับความเข้มแสงสีเทาอยู่ในช่วง 0 ถึง g ซึ่งการคำนวณทำได้ดังสมการที่ 2.7

$$m(g) = \frac{\sum_{i=0}^{g} g * f(i)}{t(g)} \qquad (2.7)$$

โดยที่

m(g) = ค่าเฉลี่ยของระดับความเข้มแสงสีเทาในแต่ละพิกเซลที่มีค่าระดับความเข้มแสงสีเทาอยู่ในช่วง 0 ถึง g

t(g) = ผลรวมต่อเนื่องของข้อมูลภาพที่มีค่าความเข้มแสง จากค่า 0 ถึง g

g = ระดับความเข้มของแสงสีเทา

f(i) = จำนวนพิกเซลในแต่ละ ระดับความเข้มของแสงสีเทา i

<u>ขั้นตอนที่สาม</u> คำนวณหาค่าระดับความเข้มแสงที่เป็นจุดแบ่งของข้อมูลภาพเพื่อแยกกลุ่มข้อมูลภาพเครื่องหมาย และกลุ่มข้อมูลที่เป็นภาพพื้นหลัง สามารถทำการคำนวณได้ดังสมการที่ 2.8 โดยที่การคำนวณจะคิดภายในขอบเขตของส่วนภาพเครื่องหมายทั้งหมดจำนวน P พิกเซล ซึ่งหาได้จากค่าผลคูณของความยาว M ของภาพเครื่องหมาย และความกว้าง N ของภาพเครื่องหมาย

$$T = max \left\{ \frac{t(g)}{P - t(g)} * [\, m(g) - m(G-1) \,]^2 \right\} - 1$$

$$T = max \left\{ A * B \right\} - 1 \qquad (2.8)$$

โดย $A = \dfrac{t(g)}{P - t(g)}$

และ $B = [\, m(g) - m(G-1) \,]^2$

โดยที่

T = ระดับความเข้มแสงที่ใช้เป็นค่าขีดแบ่ง

m(G-1) = ค่าสูงสุดของ m(g)

P = จำนวนพิกเซลทั้งหมดของบริเวณภาพ ที่ทำการวิเคราะห์

m(g) = ค่าเฉลี่ยของระดับความเข้มแสงสีเทาในแต่ละพิกเซลที่มีค่าระดับความเข้มแสงสีเทาอยู่ในช่วง 0 ถึง g

t(g) = ผลรวมต่อเนื่องของข้อมูลภาพที่มีค่าความเข้มแสง จากค่า 0 ถึง g

g = ระดับความเข้มของแสงสีเทา

f(i) = จำนวนพิกเซลในแต่ละ ระดับความเข้มของแสงสีเทา i

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากสมการ 2.7 ค่าของระดับความเข้มแสงที่ใช้เป็นขีดแบ่ง T จะสามารถหาได้จากค่าที่สูงสุดของผลคูณของค่ากลุ่มฟังก์ชันในส่วนแรกและกลุ่มฟังก์ชันในส่วนที่สองในแต่ละระดับของความเข้มแสงลบด้วย 1 ดังรายละเอียดในตารางที่ 1 สดมภ์ที่ 9 (A*B) สมการที่ 2.8 เป็นการแสดงฟังก์ชันของสมการแบบรวมกลุ่มเพื่อให้ง่ายต่อการเข้าใจ วิธีการที่กล่าวมาสามารถแสดงตัวอย่างค่าการคำนวณตัวอย่างในตารางที่ 2.1 ซึ่งเป็นข้อมูลภาพแบบที่มีความเข้มขนาด 9 ระดับและมีจำนวนจุดภาพตามข้อมูลตารางที่ 1 ในสดมภ์ที่ 1 (Histogram) ผลลัพธ์ที่ได้จากข้อมูลตัวอย่างนี้ค่าระดับความเข้มแสงที่เป็นขีดแบ่งคือ 4 โดยได้จากค่าระดับความเข้มแสงที่ 5 มีค่า A* B สูงที่สุดซึ่งเท่ากับ 5.44 จากนั้นนำค่าระดับความเข้มแสงที่ได้ลบด้วย 1 ตามวิธีการในสมการ 2.8 ทำให้ได้ค่าเท่ากับ 4

ตารางที่ 2.1 ค่าตัวอย่างการคำนวณจุดแบ่งของข้อมูลภาพ

Histogram	f(g)	t(g)	g*f(g)	$\sum$g*f(g)	m(g)	A	B	A * B
0 ****	4	4	0	0	0	0.08	23.04	0.18
1 **	2	6	2	2	0.3	0.13	20.25	2.83
2 *****	5	11	10	12	1.1	0.27	13.69	3.70
3 *********	9	20	27	39	2	0.63	7.84	4.94
4 *****	5	25	20	59	2.4	0.93	5.76	5.36
5 *****	5	30	25	84	2.8	1.36	4.00	5.44(MAX)
6 *****	5	35	30	114	3.3	2.06	2.25	4.64
7 *****	5	40	35	149	3.7	3.33	1.21	4.03
8 *********	9	49	72	221	4.5	16.33	0.09	1.47
9 ***	3	52	27	248	4.8	∞		

2.4 การหมุนภาพดิจิตอล

การหมุนภาพดิจิตอล (image rotation)[10] ซึ่งใช้แก้ไขและปรับปรุงในกรณีที่ข้อมูลภาพเอียงอันเนื่องมาจากการนำภาพเข้าสู่เครื่องสแกนเนอร์ไม่ขนานกับแนวของ เส้นสแกน (scan line) เพื่อให้การ วิเคราะห์ทำได้ถูกต้องมากยิ่งขึ้น จึงต้องหมุนภาพเพื่อให้ลักษณะข้อมูลภาพกลับมาอยู่ในแนวตรงขนานกับเส้นสแกนซึ่งทำได้โดยสมการที่ 2.9 และสมการที่ 2.10

$$x_i = x*\cos\theta - y*\sin\theta \qquad (2.9)$$

$$y_i = x*\sin\theta + y*\cos\theta \qquad (2.10)$$

โดยที่

θ = ค่าองศา ที่ต้องการหมุน

x, y = ค่าตำแหน่งของสมาชิกข้อมูลภาพ

x_i, y_i = ค่าตำแหน่งใหม่ของข้อมูลภาพ ที่หมุนไป θ องศา

2.5 การวิเคราะห์เพื่อหาจุดภาพที่เป็นส่วนประกอบของภาพเส้นตรงใด ๆ

การแทนค่าตำแหน่งจุดภาพที่เป็นส่วนประกอบของเส้นตรงใด ๆ (polyline representation)[12] ในข้อมูลภาพที่อยู่ในระนาบ xy ซึ่งภาพเส้นตรงจะเกิดขึ้นจากจุดภาพหลาย ๆ จุดภาพอยู่ในตำแหน่งที่ติดต่อกัน โดยเริ่มจากจุดเริ่มต้นไปยังจุดปลายของเส้นตรงใด ๆ ซึ่งสามารถแสดงค่าเป็นฟังก์ชันได้ดังนี้

$$\text{ตำแหน่งพิกเซลของเส้นตรง} = \{(x_1, y_1,), (x_2, y_2), \ldots\ldots (x_n, y_n)\} \qquad (2.11)$$

ค่าในแต่ละจุดของเส้นตรงนั้นมีค่าความเอียงที่เท่ากันและตลอดแนว จากหลักการที่กล่าวมาจะสามารถนำมาคำนวณหาตำแหน่งของพิกเซลที่เป็นส่วนประกอบของเส้นตรงได้ โดยใช้ค่าตำแหน่งแนวตั้ง y และค่าตำแหน่งแนวนอน x ของจุดเริ่มต้นและจุดปลายของเส้นตรง จากนั้นนำค่าที่ได้มาคำนวณหาค่าตำแหน่งได้จากสมการที่ 2.12

จากการคำนวณเพื่อเปรียบเทียบค่าความเอียง (slope) ในระหว่างตำแหน่งจุดสามจุดซึ่งถ้าจุดทั้งสามอยู่ในแนวเส้นตรงเดียวกันจะได้ค่าเท่ากันตามสมการที่ 2.12

$$\frac{x - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \qquad (2.12)$$

โดยที่

x_1 = ตำแหน่งทางแนวนอนของจุดเริ่มต้น

x_2 = ตำแหน่งทางแนวนอนของจุดปลาย

y_1 = ตำแหน่งทางแนวตั้งของจุดเริ่มต้น

y_2 = ตำแหน่งทางแนวตั้งของจุดปลาย

x = ตำแหน่งทางแนวนอนของจุดที่ต้องการวิเคราะห์ว่าเป็นส่วนของเส้นตรง

y = ตำแหน่งทางแนวตั้งของจุดที่ต้องการวิเคราะห์ว่าเป็นส่วนของเส้นตรง

จากสมการที่ 2.12 สามารถจัดให้อยู่ในรูปของผลคูณจะได้ดังสมการที่ 2.13

$$x*(y_1 - y_2) + y*(x_2 - x_1) + y_2*x_1 - y_1*x_2 = 0 \qquad (2.13)$$

ในสมการที่ 2.13 เป็นสมการแบบ 2 ตัวแปรคือ x และ y การแก้สมการเพื่อหาค่าจากตัวแปรทำโดยกำหนดค่าที่ทราบให้กับตัวแปรตัวใดตัวหนึ่งจะสามารถหาค่าของตัวแปรอีกตัวหนึ่งได้ จากหลักการที่กล่าวมาสามารถหาค่าตำแหน่งของพิกเซลที่เป็นส่วนประกอบของเส้นตรงได้ โดยการกำหนดค่าตำแหน่งในแนวนอน x เปลี่ยนค่าไปทีละช่วง ตั้งแต่จุดเริ่มต้นจนถึงจุดปลาย จากนั้นทำการคำนวณหาค่าตำแหน่งในแนวนอน y ดังตัวอย่างในสมการที่ 2.14

จุดเริ่มต้นที่ $x_1 = 1$ $y_1 = 1$

จุดปลายที่ $x_2 = 6$ $y_2 = 6$

กำหนดค่า $x = 3$

$$Y = \frac{(-1 *(x *(y_1 - y_2)) - y_2*x_1 + y_1*x_2)}{(x_2 - x_1)} \qquad (2.14)$$

$$3 = \frac{(-1*(3*(1-6)) - 6*1 + 1*6)}{(6-1)}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากการแทนค่าสมการที่ 2.14 จะได้ค่า y เท่ากับ 3 และเมื่อนำค่าที่ได้ไปแทนค่าในสมการที่ 2.13 ทำให้สมการนั้นเป็นจริง ซึ่งแสดงว่าจุดนั้นเป็นส่วนประกอบของเส้นตรง

2.6 การวิจัยที่เกี่ยวข้อง

Seong Whan Lee[6] แห่งมหาวิทยาลัย Chungbuk National University ประเทศเกาหลีได้เสนอผลงานวิจัยเรื่องการรู้จำภาพเครื่องหมายสัญลักษณ์ทางวงจรไฟฟ้าที่เขียนด้วยลายมือ (Recognizing Hand Drawn Electrical Circuit Symbols) ที่ประกอบไปด้วยสัญลักษณ์แอนเกท (AND gate) สัญลักษณ์ออเกท (OR gate) สัญลักษณ์น็อทเกท (NOT gate) โดยการพิจารณาค่าตัวเลขที่ได้จากการวัดส่วนสำคัญพื้นฐาน (numerical measurement of the local structure) ของข้อมูลภาพเครื่องหมาย ซึ่งประกอบด้วย ความยาว ความโค้ง จุดตัด จุดต่อ และมุม แล้วนำมาเปรียบเทียบกับสิ่งที่เรียนรู้ไว้ โดยใช้วิธีการทางสถิติในการแบ่งแยกความถูกต้อง ผลการวิจัยในครั้งนั้นมีความถูกต้อง 80.7% จากตัวอย่างในการทดสอบ 300 ตัวอย่างโดยใช้ตัวอย่างจากคน 7 คนเพื่อให้ตัวอย่างคนละ 10 ตัวอย่างเพื่อการเรียนรู้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย