

การประกอบคืนชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวม
ความยาวแกนหลักวงรีและฮัฟทรานส์ฟอร์ม

TORN DOCUMENT RECONSTRUCTION USING HISTOGRAM OF
ACCUMULATED MAJOR AXIS LENGTH OF FITTED ELLIPSE (HOARE) AND
HOUGH TRANSFORM

ภัทรภร วัฒนาชีพ

BHATTARABHORN WATTANACHEEP

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าคุณทหารลาดกระบัง

พ.ศ. 2558

KMITL-2015-EN-M-070-195

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การประกอบคืนชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวม
ความยาวแกนหลักวงรีและฮัฟทรานสฟอร์ม

TORN DOCUMENT RECONSTRUCTION USING HISTOGRAM OF
ACCUMULATED MAJOR AXIS LENGTH OF FITTED ELLIPSE (HOARE) AND
HOUGH TRANSFORM



T141257

ภัทรภร วัฒนาชีพ

BHATTARABHORN WATTANACHEEP

เลขหมู่.....
เลขทะเบียน..... 141257
วัน.เดือน.ปี..... ๒๐ สิงหาคม ๒๕๕๙

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.2558

KMITL-2015-EN-M-070-195

TORN DOCUMENT RECONSTRUCTION USING HISTOGRAM OF
ACCUMULATED MAJOR AXIS LENGTH OF FITTED ELLIPSE (HOARE) AND
HOUGH TRANSFORM

BHATTARABHORN WATTANACHEEP

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ENGINEERING IN COMPUTER ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2015
KMITL-2015-EN-M-070-195

COPYRIGHT 2015

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การประกอบคืนชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวม
ความยาวแกนหลักวงรีและฮัฟทรานส์ฟอร์ม

Thesis Title Torn Document Reconstruction using Histogram of Accumulated Major Axis
Length of Fitted Ellipse (HoARE) and Hough Transform

นักศึกษา นางสาวภัทรภร วัฒนาชีพ

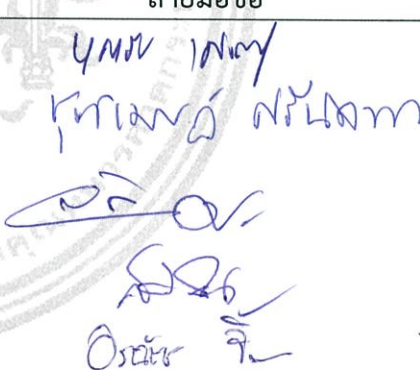
รหัสประจำตัว 56601126

ปริญญา วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา วิศวกรรมคอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์ รศ.ดร.อรฉัตร จิตต์โสภักดิ์

หมายเลขวิทยานิพนธ์ KMITL-2015-EN-M-070-195

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.ดร.บุญธีร์	เครือข่าย	
ผศ.ดร.ชุตินิเมษฐ์	ศรีนิลทา	
ผศ.ดร.จักรี	ศรีนนท์ฉัตร	
ผศ.ดร.สมศักดิ์	วัลย์รัชต์	
รศ.ดร.อรฉัตร	จิตต์โสภักดิ์	

วัน / เดือน / ปี ที่สอบ วันจันทร์ที่ 23 พฤศจิกายน พ.ศ. 2558 เวลา 10.30-12.00 น.
สถานที่สอบ ณ อาคาร A ชั้น 5 ห้องประชุม 1

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิศวกรรมศาสตร์ รับรองแล้ว


(รองศาสตราจารย์ ดร. คมสัน มาลีสี)

คณบดี คณะวิศวกรรมศาสตร์
วันที่ 23 พฤศจิกายน พ.ศ. 2558

หัวข้อวิทยานิพนธ์	การประกอบชิ้นชิ้นส่วนเอกสารโดยใช้คุณลักษณะของ
	ฮิสโทแกรมผลรวมความยาวแกนหลักวงรีและฮัฟทรานสฟอร์ม
นักศึกษา	นางสาวภัทรกร วัฒนาชีพ
รหัสนักศึกษา	56601126
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
พ.ศ.	2558
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.ดร.อรฉัตร จิตต์โสภักตร์

บทคัดย่อ

การประกอบชิ้นส่วนเอกสารที่ถูกฉีกขาดอัตโนมัติ สามารถนำมาประยุกต์ใช้ในงานที่แตกต่างกัน เช่น นำไปใช้ประกอบในงานของนักวิทยาศาสตร์ นักโบราณคดีและนักกฎหมาย โดยเอกสารอาจถูกทำให้ขาดเนื่องจากการเก็บรักษาไม่ดีหรือเอกสารถูกฉีก เพื่อให้ข้อมูลในเอกสารไม่สามารถอ่านได้ หากมีชิ้นส่วนที่ขาดออกเป็นจำนวนมากนักต่อจิ๊กซอว์จะไม่สามารถประกอบชิ้นส่วนได้ เพราะต้องมีค่าใช้จ่ายสูงและเวลาที่ต้องใช้มากจนเกินไป การประกอบเอกสารที่ถูกฉีกขาดสามารถแยกเป็นเทคนิคที่มีพื้นฐานโดยการวิเคราะห์รูปร่าง (ไม่มีเนื้อหาภายในชิ้นส่วน) และเทคนิคที่วิเคราะห์เนื้อหาภายใน (ที่เป็นภาพวาดหรือภาพประกอบ) ของเอกสารที่แยกออกเป็นชิ้นๆ หากใช้เฉพาะเทคนิคการวิเคราะห์รูปร่างชิ้นส่วน จะทำให้ชิ้นส่วนไม่สามารถตั้งตรงหรืออยู่ในระนาบที่ต้องการได้ ซึ่งอาจก่อให้เกิดอัตราความผิดพลาดที่เพิ่มมากขึ้นได้ ดังนั้นวิธีการผสมผสานเทคนิคการวิเคราะห์รูปร่างและเทคนิคการวิเคราะห์เนื้อหาเอกสารเป็นสิ่งจำเป็นในการประกอบกลับคืนชิ้นส่วนที่ถูกฉีกขาดจำนวนมากเช่นกัน การประกอบเอกสารที่ถูกฉีกขาดอัตโนมัติโดยทั่วไปนั้น การจัดระนาบชิ้นส่วนให้ตั้งตรงเป็นส่วนแรกที่ต้องทำก่อนการประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาด หากการจัดเรียงชิ้นส่วนสามารถทำได้อย่างถูกต้อง จะมีส่วนช่วยให้การประกอบชิ้นส่วนกระดาษมีความถูกต้องมากยิ่งขึ้น ดังนั้นงานวิจัยนี้จึงนำเสนอการประกอบชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวมความยาวแกนหลักวงรีและฮัฟทรานสฟอร์ม โดยใช้ทิศทางการเรียงตัวของตัวอักษรในการจัดระนาบชิ้นส่วนให้ตั้งตรง โดยเทคนิคที่นำเสนอสามารถลดอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วน (ROE) 5.07 เปอร์เซ็นต์ และเพิ่มประสิทธิภาพของทิศทางการเรียงตัวหลักแบบค่าความถูกต้อง (Precision) 24.11 เปอร์เซ็นต์ นอกจากนี้ยังนำเสนอการใช้ฮัฟทรานสฟอร์ม (Hough Transform) หาแนวหลักของเส้นขอบ และนำพารามิเตอร์ของเส้นตรงฮัฟมาใช้เป็นคุณลักษณะรูปร่าง เพื่อเป็นข้อมูลที่สามารถนำมาใช้ในการประกอบชิ้นส่วนที่ถูกฉีกขาด ซึ่งจากผลการทดลองแสดงให้เห็นถึงการปรับปรุงประสิทธิภาพอย่างมีนัยสำคัญของอัลกอริธึมที่นำเสนอ

Thesis Title	Torn Document Reconstruction Using Histogram of Accumulated Major Axis Length of Fitted Ellipse (HoARE) and Hough Transform
Student	Miss Bhattarabhorn Wattanacheep
Student ID	56601126
Degree	Master of Engineering
Program	Computer Engineering
Year	2015
Thesis Advisor	Assoc. Prof. Dr. Orachat Chitsobhuk

ABSTRACT

The automated torn document reconstruction can be applied in a different ways such as in the work of scientists, archaeologists or lawyers. The documents may be torn due to poor preservation or intended destruction. If the number of the snippets is large, it would be difficult for a human to reconstruct the puzzles since it requires excessive cost and time consuming. The automated torn document alignment is based on technical analysis of shape (not for painting or illustration) and its inner content (for a painting or illustration) of each snippets. Using only shape analysis, parts cannot be placed in the correct position. This may result in an increment of the error rate. Thus, a combination of shape and content analysis techniques is needed for effective reconstruction of a large number of snippets. Snippet alignment is a process to arrange pieces of the torn document in the same direction according to the direction of the alphabet line. It is the prerequisite to assure for the effective reconstruction. The higher the performance of the snippet alignment, the greater the opportunity for successful reconstruction. Therefore, this paper presents torn document reconstruction using Histogram of Accumulated Major Axis Length of fitted Ellipse (HoARE) and Hough Transform is proposed. The proposed technique analyzes the contents inside the snippet such as the direction of the character alignment based on the histogram of the accumulated major axis length of the fitted ellipses. The direction result is then used to revert the snippet to its original position. The proposed technique can achieve approximately 5.07 percent greater in relative orientation error and 24.11 percent increment in reverting precision. Moreover, Hough transform based local descriptor is extracted as shape feature. These parameters are helpful for accurate reconstruction. With the proposed snippet alignment and Hough transform based reconstruction, this can demonstrate the significant performance improvement of the proposed algorithm.

กิตติกรรมประกาศ

คุณงามความดีและประโยชน์ที่ได้จากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแต่บิดามารดาและครูอาจารย์ที่เคารพรักทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้และประสบการณ์ที่ดีอันเป็นประโยชน์แก่ข้าพเจ้า

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงตามวัตถุประสงค์ได้ด้วยความกรุณาของรองศาสตราจารย์ ดร.อรฉัตร จิตต์โสภาคย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งให้คำปรึกษา ข้อชี้แนะและแนวทางในการแก้ปัญหาต่างๆ ทั้งด้านวิชาการ ด้านการดำเนินชีวิต ตลอดจนให้ความรู้ และความช่วยเหลือในหลายสิ่งหลายอย่างจนกระทั่งลุล่วงไปได้ด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงมา ณ. ที่นี้

ข้าพเจ้าขอขอบพระคุณครูอาจารย์ที่เคารพทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้ อีกทั้งถ่ายทอดประสบการณ์อันเป็นประโยชน์ต่างๆ ทั้งโดยตรง และโดยอ้อม ตลอดจนนักวิจัยทุกท่านที่เอื้อเฟื่องานวิจัย จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอกราบขอบพระคุณคณะกรรมการการสอบที่ให้ความกรุณาในการชี้แนะและแก้ไขข้อบกพร่องต่างๆ ของงานวิจัย

ขอบคุณ พี่ๆ เพื่อน และน้องภาควิชาวิศวกรรมคอมพิวเตอร์ทุกคน ที่คอยสอบถามด้วยความห่วงใย ให้ความช่วยเหลือ คำแนะนำ กำลังใจที่ดีตลอดมาและขอขอบคุณพิเศษสำหรับ “พี่จุ่ม” คุณรัตติกร สมบัติแก้ว “พี่กุล” คุณกุลวลัญช์ วรณสิน “พินพ” คุณนพพล น้อยแก้ว “พี่เบน” คุณจตุรนต์ เงินพลับพลา และ “พี่แจ้ว” คุณวงศ์สวรรค์ ศรีมนตรีสง่า รวมถึงผู้มีพระคุณทุกท่านที่มีได้เอ่ยนามไว้ ณ. ที่นี้

ผู้วิจัยจึงขอกราบขอบพระคุณเป็นอย่างสูง มา ณ โอกาสนี้

ภัทรภร วัฒนาชีพ

สารบัญ

หน้า

บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VIII
สารบัญรูป	XI
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	2
1.3 สมมติฐานของการศึกษา	2
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย	3
1.5 ขอบเขตงานวิจัย	3
1.6 ขั้นตอนการศึกษา	4
1.7 เครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัย	4
1.8 โครงสร้างของวิทยานิพนธ์	5
บทที่ 2 งานวิจัยที่เกี่ยวข้อง	6
2.1 การทำชิ้นส่วนให้ตั้งตรง	7
2.2 การประกอบชิ้นส่วน	8
2.2.1 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ของ H. FREEMAN และ L. GARDER [13]	8
2.2.2 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ของ Feng-Hui Yao และ Gui-Feng Shao [18]	11
2.2.3 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์และการประกอบชิ้นส่วนที่ถูกฉีดยึดขาด ด้วยวิธีของ Daniel J. Ho และ Peter J. Olver [19]	12

สารบัญ (ต่อ)

หน้า

2.2.4 การประกอบชิ้นส่วนกระดาดที่ถูกฉีกขาดด้วยวิธีของ Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade [20]	14
บทที่ 3 ความรู้พื้นฐานในเรื่องที่เกี่ยวข้องกับงานวิจัย	16
3.1 การแปลงภาพสีเป็นขาวดำ (Image Binarization).....	16
3.2 การคำนวณขนาดและทิศทางของเวกเตอร์การเปลี่ยนแปลง (Magnitude and Orientation of Gradient Vectors).....	18
3.3 การหาขอบภาพโดยเทคนิคแคนนี่ [24].....	21
3.3.1 การปรับภาพให้เรียบด้วยตัวกรองเกาส์เซียน.....	22
3.3.2 การคำนวณค่าขนาดและทิศทางของค่าการเปลี่ยนแปลงความเข้มแสง	22
3.3.3 การกำจัดค่าขนาดการเปลี่ยนแปลงความเข้มแสงที่ไม่ใช่ค่าสูงสุดเฉพาะที่ (Nonmaxima Suppression).....	22
3.3.4 การทำฮิสเทียรีซิส (Hysteresis) เพื่อคัดกรองให้เหลือแต่พิกเซลเส้นขอบที่แท้จริง	23
3.4 การประมวลผลรูปร่างและโครงร่างของวัตถุในภาพ (Morphological Image Processing) [25].....	23
3.4.1 เทคนิคฮิตและมิส (Hit and Miss).....	23
3.4.2 การขยายองค์ประกอบ (Dilation).....	24
3.4.3 การกร่อนองค์ประกอบ (Erosion)	25
3.4.4 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบเปิด (Opening)	26
3.4.5 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบปิด (Closing)	27
3.4.6 การเปลี่ยนองค์ประกอบให้บางลง (Thinning).....	28
3.5 การทำ Labeling and Measuring Objects	29
3.6 การคำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (Histogram of accumulated Major Axis Length of ellipses; HoARE)	31
3.7 การแปลงเรขาคณิตแบบเชิงเส้น (Geometrical Linear Transformation) [28].....	32
3.7.1 การหมุนภาพ (Rotation)	32
3.7.2 การเคลื่อนย้ายภาพ (Translation).....	33
3.8 ฮัฟฟรานสฟอร์ม [30]	34

สารบัญ (ต่อ)

	หน้า
3.9 การวิเคราะห์มุม.....	35
บทที่ 4 งานวิจัยที่นำเสนอ	37
4.1 ภาพรวมของงานวิจัยที่นำเสนอ.....	37
4.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร [29].....	38
4.2.1 การวิเคราะห์การหมุน.....	39
4.2.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร.....	40
4.3 การหมุนกลับคืนชิ้นส่วน	47
4.4 การวิเคราะห์แนวหลักของเส้นขอบชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม.....	49
4.5 การประกอบชิ้นส่วน	52
บทที่ 5 ผลการทดลองและการวิเคราะห์	58
5.1 รายละเอียดฐานข้อมูลภาพ	59
5.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร.....	60
5.3 การวิเคราะห์ประสิทธิภาพการหมุนกลับคืนชิ้นส่วน	64
5.4 การวิเคราะห์ผลลัพธ์จากการหาทิศทางหลักของขอบชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม (Hough Transform).....	68
5.5 การวิเคราะห์พารามิเตอร์ที่ใช้ในการประกอบชิ้นส่วน	72
5.6 การวิเคราะห์ประสิทธิภาพการประกอบคืนชิ้นส่วนที่ถูกฉีกขาด	74
บทที่ 6 สรุปผลและแนวทางในการพัฒนา.....	76
6.1 การวิเคราะห์และสรุปผลการดำเนินงานวิจัย	76
6.1.1 การวิเคราะห์และสรุปผลการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรง.....	76
6.1.2 การวิเคราะห์และสรุปผลการประกอบชิ้นส่วน.....	76
6.2 ข้อจำกัดและขอบเขตของงานวิจัย.....	77
6.3 แนวทางในการพัฒนา	77
เอกสารอ้างอิง.....	78

สารบัญ (ต่อ)

	หน้า
ภาคผนวก ก ผลงานวิจัยในระหว่างการศึกษาที่ได้รับการตีพิมพ์เผยแพร่.....	81
ภาคผนวก ข ข้อมูลมมมมมมจากมมมต่างๆของแต่ละชั้นส่วนที่ได้จาก HoARE.....	88
ภาคผนวก ค ข้อมูลค่า ROE และ Precision แต่ละชั้นส่วนจากมมมมมมต่างๆ.....	93
ภาคผนวก ง ข้อมูลการประกอบของแต่ละมมมในแต่ละเงื่อนไข.....	96
ประวัติผู้เขียน	121

สารบัญตาราง

ตารางที่	หน้า
4.1	แสดงข้อมูลสำหรับการประกอบชิ้นส่วน52
5.1	จำนวนวงรีที่ล้อมรอบรูปภาพของชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน).....62
5.2	จำนวนวงรีที่ล้อมรอบเส้นคั่นของชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ62
5.3	จำนวนวงรีที่วงรูปภาพหรือเส้นตาราง.....62
5.4	จำนวนวงรีที่ล้อมรอบเส้นตารางของชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง63
5.5	ชิ้นส่วนจากงานวิจัย [2].....64
5.6	ผลจากการวัดความถูกต้องด้วยวิธีการวัดอัตราความผิดพลาดของทิศการเรียงตัวของชิ้นส่วน (ROE) และวิธีการวัดประสิทธิภาพทิศการเรียงตัวหลัก (Precision)67
5.7	เงื่อนไขในการพิจารณาการเข้าคู่กันของชิ้นส่วน72
5.8	จำนวนมุมที่จับคู่ถูก-คู่ผิดของชุดชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน).....73
5.9	แสดงการจับคู่มุมของชุดชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)74
5.10	การวัดประสิทธิภาพการประกอบชิ้นส่วนที่ถูกฝึกขาของแต่ละชุดการทดสอบชิ้นส่วน.....75

สารบัญรูป

รูปที่	หน้า
2.1	กระดาษที่ถูกฉีกด้วยมือ 7
2.2	แผนผังเชนเอนโค้ดดิ้ง (Chain encoding scheme)..... 9
2.3	ขั้นตอนการประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ด้วยวิธีของ H. FREEMAN และ L. GARDER 10
2.4	มุมเชื่อมที่เชื่อมด้วยชิ้นส่วนสามชิ้นและสี่ชิ้น 11
2.5	ขั้นตอนการประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ด้วยวิธีของ Feng-Hui Yao และ Gui-Feng Shao 12
2.6	ภาพชิ้นส่วน18 แบบ..... 12
2.7	การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์และชิ้นส่วนที่ถูกฉีกขาดด้วยวิธีของ Daniel J. Ho และ Peter J. Olver..... 13
2.8	ภาพแสดงผลรวมของผลต่างของการจับคู่ขอบ..... 13
2.9	การประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาดด้วยวิธีของ Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade..... 15
3.1	ระบบสี RGB..... 16
3.2	ระบบสีเกรสเกล..... 17
3.3	ภาพในระบบสีเกรสเกล 17
3.4	ภาพที่แบ่งเรซโซลด์..... 17
3.5	กราฟเกรสเกล 17
3.6	ขั้นตอนการหาลักษณะเด่นของภาพด้วย SIFT 18
3.7	การหาค่า DoG ของแผ่นภาพในแต่ละช่วงข้อมูล 19
3.8	การพิจารณาจุดของภาพเพื่อเลือกเป็นจุดสนใจ 20
3.9	ลักษณะเด่นของภาพที่ได้จากวิธี SIFT..... 21
3.10	ตัวอย่างการขยายองค์ประกอบ 25
3.11	ตัวอย่างการกร่อนองค์ประกอบ..... 26
3.12	ตัวอย่างการกร่อนองค์ประกอบที่ได้จากเทมเพลตใหม่ 26
3.13	ผลลัพธ์จากการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบเปิด 27
3.14	ผลลัพธ์จากการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบปิด 28
3.15	รูปแบบเทมเพลตที่นำมาพิจารณา..... 28

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.16 เหมเพลดตามตัวอย่าง (ก) $p_3 = 0$ (ข) $p_5 = 0$ (ค) $p_1 = 0$ และ $p_7 = 0$	29
3.17 เหมเพลดตามตัวอย่าง (ก) $p_1 = 0$ (ข) $p_7 = 0$ (ค) $p_4 = 0$ และ $p_5 = 0$	29
3.18 8-connectivity.....	30
3.19 ต้นไม้ตัดสินใจที่ใช้ตรวจสอบหา 8-connectivity.....	30
3.20 แกนหลักและแกนรองของวงรี.....	31
3.21 การหมุน (x, y) รอบจุดกำเนิดไปที่จุด (u, v) ด้วยมุม θ	32
3.22 ตัวอย่างการหมุนภาพ.....	33
3.23 ตัวอย่างการย้ายภาพ $t_x = 20, t_y = 20$	33
3.24 ค่า ρ และ θ ของเส้นตรงในพิกัดเชิงขั้ว	34
3.25 การทำฮัฟทรานสฟอร์ม.....	35
3.26 ลักษณะมุมโค้งเข้าและโค้งออกของจิกซอร์.....	36
4.1 การจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดให้อยู่ในระนาบเดียวกัน.....	38
4.2 กราฟฮิสโทแกรมของแนวทิศการวางตัวด้วย Spline Interpolation.....	40
4.3 ขั้นตอนการหาขอบของชิ้นส่วนเพื่อให้เหลือเฉพาะเนื้อหาภายในชิ้นส่วน	41
4.4 การกำจัดขอบของชิ้นส่วน	41
4.5 วิธีการคัดเลือกให้พื้นที่ที่สนใจมีแต่ตัวอักษร.....	42
4.6 วงรีล้อมรอบเฉพาะตัวอักษร	43
4.7 การเปรียบเทียบทิศทางการเรียงตัวหลักของชิ้นส่วนที่ทำมุม -45 องศา	45
4.8 การเปรียบเทียบผลลัพธ์ของการหาทิศทางของตัวหนังสือของชิ้นส่วนเอียง -45 องศา	47
4.9 จุดที่ตำแหน่ง x, y ถูกหมุนไปยังจุด x', y' โดยถูกหมุนด้วยมุม θ รอบจุด $(0,0)$	48
4.10 ภาพแสดงการตีกรอบล้อมรอบขอบของชิ้นส่วนด้วยวิธี Hough Transform.....	50
4.11 (ก) ภาพมุมของชิ้นส่วนที่ถูกระบุโดยช่องสี่เหลี่ยม (ข) เส้นขอบของชิ้นส่วนจากค่าสูงสุดของ ฮัฟสเปซในแต่ละช่วงโรลและเซต้า	51
4.12 ภาพแสดงรูปร่างของชิ้นส่วน.....	51
4.13 (ก) รูปร่างชิ้นส่วนของการเก็บข้อมูลแบบพิกเซลของขอบ (ข) รูปร่างชิ้นส่วนโดยใช้คุณลักษณะ ฮัฟทรานสฟอร์ม	53
4.14 การเข้าคู่กันของชิ้นส่วน	54
4.15 การพิจารณามุมที่เข้าคู่กันของชิ้นส่วน.....	55

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.15 การพิจารณามุมที่เข้าคู่กันของชั้นส่วน.....	55
4.16 กลุ่มชั้นส่วนที่ผ่านการพิจารณาในเงื่อนไขหนึ่ง.....	56
4.17 กลุ่มชั้นส่วนที่ผ่านการพิจารณาในเงื่อนไขสอง.....	56
4.18 กลุ่มชั้นส่วนที่ผ่านการพิจารณาในเงื่อนไขสาม.....	57
5.1 ชั้นส่วนตัวอย่างในแต่ละชุดการทดสอบ.....	59
5.2 ชั้นส่วนตัวอย่างในแต่ละชุดการทดสอบ.....	60
5.3 การตีกรอบตัวอักษรบนชั้นส่วนที่มีรูปภาพและเป็นตาราง.....	61
5.4 การเปรียบเทียบทิศทางการเรียงตัวหลักของชั้นส่วนเอียงที่ 45 องศา.....	65
5.5 การหาขอบของชั้นส่วนโดยการใช้ค่า $\Delta\rho$ ต่างๆ.....	69
5.6 การหาขอบของชั้นส่วนโดยการปรับค่า $\Delta\theta$	69
5.7 การหาขอบของชั้นส่วนโดยการปรับค่าเรซโซลต์.....	70
5.8 ภาพการคัดเลือกเส้นฮัฟ.....	71
5.9 ภาพการสร้างเส้นเชื่อมให้เป็นรูปปิด.....	71

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการประกอบต้นฉบับหรือหนังสือที่ได้รับการถูกทำลายให้กลับคืนสภาพใกล้เคียงกับต้นฉบับมากที่สุดนั้นเป็นสิ่งที่จำเป็นในการคงคุณค่าทางประวัติศาสตร์และวัฒนธรรมของเอกสารเหล่านั้นไว้ ไม่ว่าการทำลายจะเกิดขึ้นโดยอุบัติเหตุหรือเกิดจากการล่มสลายของประวัติศาสตร์หรือสงคราม แต่เนื่องจากชิ้นส่วนของเอกสารที่ถูกทำลายมีเป็นจำนวนมากจึงทำให้การประกอบชิ้นส่วนด้วยคนต้องใช้เวลาและเสียค่าใช้จ่ายจำนวนมาก ทำให้การประกอบเอกสารที่ถูกฉีกขาดอัตโนมัติเป็นอีกทางเลือกที่สำคัญในการลดข้อจำกัดนี้

การประกอบเอกสารที่ถูกฉีกขาดอัตโนมัติโดยทั่วไปนั้นมีความซับซ้อน กล่าวคือ การประกอบชิ้นส่วนฉีกขาดยังมีข้อจำกัดในเรื่องรูปแบบของชิ้นส่วนจึงทำให้ผลลัพธ์ที่ได้จากการประกอบชิ้นส่วนมีความผิดพลาดมากซึ่งยังคงเป็นปัญหาของงานวิจัยทางด้านนี้ [1] หากมีขั้นตอนการจัดเตรียมรูปแบบของชิ้นส่วนก่อนการประกอบจะช่วยลดความผิดพลาดของการประกอบชิ้นส่วนผิดได้ โดยขั้นตอนการจัดเตรียมชิ้นส่วนเริ่มจากการจัดระนาบชิ้นส่วนให้ตั้งตรงซึ่งเป็นส่วนแรกที่มีความสำคัญที่ควรทำการประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาด หากการจัดเรียงชิ้นส่วนสามารถทำได้ถูกต้องจะต้องมีส่วนช่วยให้การประกอบชิ้นส่วนกระดาษมีความถูกต้องมากยิ่งขึ้น ในการจัดระนาบชิ้นส่วนให้ตั้งตรงหากพิจารณาเฉพาะขอบของชิ้นส่วนจะไม่สามารถทำให้ชิ้นส่วนตั้งตรงได้เพราะขอบของชิ้นส่วนไม่สามารถบอกทิศทางของชิ้นส่วน นอกจากนี้ในกรณีที่ชิ้นส่วนมีรูปร่างลักษณะเหมือนกันการพิจารณาเฉพาะขอบจะทำให้ตำแหน่งชิ้นส่วนสลับตำแหน่งกันได้ จึงส่งผลให้ไม่สามารถประกอบชิ้นส่วนได้ถูกต้อง การจัดระนาบชิ้นส่วนให้ตั้งตรงจึงจำเป็นต้องพิจารณาเนื้อหาภายในชิ้นส่วนร่วมด้วย การพิจารณาโดยใช้เฉดสี เช่น สีกระดาษ สีพื้นหลัง หรือสีตัวอักษรไม่สามารถทำให้กระดาษตั้งตรงได้เนื่องจากสีเป็นข้อมูลที่ไม่อาจกำหนดทิศทางได้ โดยส่วนใหญ่การพิจารณาโดยใช้เฉดสีจะถูกนำมาใช้เป็นเงื่อนไขหนึ่งสำหรับพิจารณาการจับคู่ชิ้นส่วน (ชิ้นส่วนที่มีสีเดียวกันเป็นชิ้นส่วนที่อยู่ติดกันแต่ชิ้นส่วนที่มีสีต่างกัมนั้นอาจไม่ใช่ชิ้นส่วนที่อยู่ใกล้กันหรือคู่กัน) ดังนั้นการทำให้ชิ้นส่วนตั้งตรงนั้นจึงมีความจำเป็นต้องพิจารณาทิศทางเรียงตัวของตัวหนังสือภายในชิ้นส่วนว่าส่วนใหญ่แล้วตัวหนังสืออยู่ในทิศใด ถ้าทิศของตัวหนังสือในชิ้นส่วนนั้นตั้งตรงหมายความว่าชิ้นส่วนนั้นอยู่ในระนาบที่ตั้งตรงด้วยเช่นกัน

ดังนั้นเพื่อลดข้อจำกัดที่ได้กล่าวมาข้างต้น ในงานวิจัยนี้จึงนำเสนอเทคนิคการจัดเรียงระนาบของชิ้นส่วนโดยพิจารณารูปร่างของขอบและเนื้อหาภายในชิ้นส่วน เริ่มจากการหาทิศทางเรียงตัวหลักของชิ้นส่วนเอกสารที่ถูกฉีกขาดจากแนวของตัวหนังสือภายในชิ้นส่วนเพื่อหมุนกลับให้ตัวหนังสือตั้งตรง โดยวิเคราะห์ค่าทางสถิติด้วยค่าของฮิสโทแกรมทิศทางของผลรวมความยาวแกนหลักวงรีที่อยู่ในทิศทางเดียวกัน หลังจากขั้นตอนการหมุนกลับจะทำให้ชิ้นส่วนที่ถูกฉีกขาดเหล่านี้อยู่ในระนาบ

เดียวกัน (ตั้งตรง) จากนั้นจึงใช้วิธีฮัฟทรานสฟอร์ม (Hough Transform) ซึ่งมีบทบาทสำคัญในการเรียนรู้รูปร่างของวัตถุ (Shape) โดยวิธีการตรวจหาเส้นตรงที่มีภายในภาพซึ่งเป็นแนวหลักของเส้นขอบของชิ้นส่วนเพื่อนำข้อมูลที่ได้มาประยุกต์ใช้ในการประกอบชิ้นส่วนที่ถูกฉีกขาดให้กลับคืนใกล้เคียงกับต้นฉบับมากที่สุด

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

- 1.2.1 ศึกษาเทคนิคและการจัดเรียงและประกอบภาพชิ้นส่วน
- 1.2.2 ปรับปรุงวิธีการโดยอาศัยสมมติฐาน ความเข้าใจ และทฤษฎีอื่นประกอบการทดลอง
- 1.2.3 ทดลองการจัดเรียงและประกอบภาพชิ้นส่วนที่ถูกฉีกขาดจริง เพื่อทดสอบความเป็นไปได้และวัดประสิทธิภาพในการนำไปประยุกต์ใช้งานจริง
- 1.2.4 เพื่อเป็นแหล่งข้อมูลอ้างอิงสำหรับผู้สนใจงานวิจัยด้านนี้ต่อไป

1.3 สมมติฐานของการศึกษา

ในขั้นตอนการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงจำเป็นต้องพิจารณาทิศการเรียงตัวของตัวอักษรภายในชิ้นส่วนเนื่องจากการทำชิ้นส่วนให้อยู่ในระนาบตั้งตรงด้วยวิธีการหาฮิสโทแกรมทิศทาง การเปลี่ยนแปลงค่าระดับความเข้มแสงนั้นเป็นการหาทิศทางของชิ้นส่วนที่มีข้อจำกัดคือจุดสูงสุดของกราฟจะไม่ใช่ทิศทางของชิ้นส่วนที่แท้จริงเมื่อขอบของชิ้นส่วนถูกฉีกขาดเป็นเส้นตรง ซึ่งข้อจำกัดนี้สามารถแก้ไขโดยใช้วงรีที่ล้อมรอบเนื้อหาภายในชิ้นส่วน เนื้อหาภายในชิ้นส่วนนั้นสามารถประกอบด้วยตัวอักษร รูปภาพ และเส้นตาราง จึงจำเป็นต้องนำเฉพาะวงรีที่ล้อมรอบตัวอักษรมาพิจารณาหาทิศทางการเรียงตัวของตัวหนังสือส่วนใหญ่ในชิ้นส่วน เพื่อให้ทิศทางการเอียงของชิ้นส่วนเป็นทิศทางที่แท้จริง ในการนำวงรีที่ล้อมรอบตัวหนังสือมาพิจารณาหาทิศทางของชิ้นส่วนนั้นสามารถคำนวณวิธีทางสถิติได้หลายวิธี เช่น การนับจำนวนวงรีล้อมรอบคำที่มีทิศทางเดียวกัน แต่วิธีนี้ยังมีข้อจำกัดคือ ไม่สามารถหาทิศทางของชิ้นส่วนได้หากตัวหนังสือภายในชิ้นส่วนมีน้อย (2-3 ตัวอักษร) ซึ่งการหาผลรวมของรัศมีของวงรีที่มีทิศทางเดียวกันนั้นจะสามารถแก้ปัญหาข้อจำกัดนั้นได้

ดังนั้นในงานวิจัยนี้จึงนำเสนอวิธีการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงที่มีประสิทธิภาพ มีความถูกต้องและรวดเร็ว รวมทั้งนำเสนอคุณลักษณะหรือตัวแทนรูปร่างของชิ้นส่วนจากฮัฟทรานสฟอร์มซึ่งสามารถนำไปประยุกต์ใช้ประกอบชิ้นส่วนจริงได้ โดยใช้เทคนิคการหาผลรวมของรัศมีของวงรีที่มีทิศทางเดียวกัน ซึ่งทำให้ผลลัพธ์ที่ได้มีความถูกต้องมากยิ่งขึ้นและสามารถใช้กับชิ้นส่วนที่มีลักษณะต่างๆได้ และนำวิธีฮัฟทรานสฟอร์มมาใช้เป็นตัวแทนข้อมูลขอบของชิ้นส่วน ได้แก่ ค่าระยะที่วัดจากจุดกำเนิดไปตั้งฉากกับเส้นตรง (Rho ; ρ) ค่าของมุมระหว่างเวกเตอร์ ρ กับแกน x (Theta ; θ) และค่าเรโซลต์ (Threshold) ซึ่งพารามิเตอร์เหล่านี้จะใช้เป็นตัวแทนสำหรับการหาขอบของชิ้นส่วนในรูปแบบคู่อันดับของเส้นตรง (ρ, θ) จึงทำให้สามารถนำข้อมูลเส้นขอบมาคำนวณหามุมของชิ้นส่วน

เพื่อให้ได้ขนาดของมุม ความยาวเส้นประกอบมุมและองศาของเส้นประกอบมุมซึ่งสามารถนำไปประยุกต์ใช้ในการประกอบชิ้นส่วนที่ถูกฉีกขาดได้อย่างมีประสิทธิภาพต่อไป

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย

วิธีการจัดเรียงชิ้นส่วนให้อยู่ในระนาบเดียวกันหรือตั้งตรงในงานวิจัยนี้ อาศัยทฤษฎีและหลักการดังต่อไปนี้

- 1.4.1 การแปลงภาพสีเป็นขาวดำ (Image Binarization)
- 1.4.2 การคำนวณขนาดและทิศทางของเวกเตอร์การเปลี่ยนแปลง (magnitude and orientation of gradient vectors) ถูกนำมาประยุกต์ใช้กับการหาทิศทางทั้งหมดของชิ้นส่วน
- 1.4.3 การประมวลผลรูปร่างและโครงสร้างของภาพ (Morphological image processing) ถูกนำมาใช้เปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ
- 1.4.4 การทำลาเบลลิ่ง (Labeling and Measuring Objects) ถูกนำมาประยุกต์ใช้กับการตีกรอบวงรีล้อมรอบเฉพาะตัวอักษรภายในชิ้นส่วน
- 1.4.5 การคำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (Histogram of accumulated Major Axis Length of ellipses; HoARE) ถูกนำมาประยุกต์ใช้กับการหาผลรวมรัศมีของวงรีที่ล้อมรอบตัวอักษรเพื่อหามุมที่ชิ้นส่วนเอียงให้หมุนกลับมาตั้งตรง
- 1.4.6 การแปลงเรขาคณิตแบบเชิงเส้น (Geometrical Linear Transformation) ถูกนำมาใช้กับการหมุนและการเคลื่อนย้ายภาพ
- 1.4.7 ฮัฟทรานสฟอร์ม (Hough transform) ถูกนำมาประยุกต์ใช้กับการหาขอบและข้อมูลที้นำมาใช้ในการประกอบชิ้นส่วน
- 1.4.8 เทคนิคการประกอบชิ้นส่วนแบบฮิวริสติกเสิร์จ (Heuristic search)

1.5 ขอบเขตงานวิจัย

วิทยานิพนธ์นี้ได้ทำการศึกษา และพัฒนาวิธีการจัดเรียงชิ้นส่วนกระดาษที่ถูกฉีกขาดให้อยู่ในระนาบตั้งตรง ซึ่งมีขอบเขตการวิจัยดังนี้

- 1.5.1 ลักษณะการขาดของกระดาษอาจเกิดเป็นชั้นที่แตกต่างกันจะทำให้เกิดการหลวมกัน บางส่วนของขอบด้านนอกของชิ้นส่วน การพิจารณาขอบของชิ้นส่วนจึงพิจารณาขอบด้านในของชิ้นส่วน
- 1.5.2 ชุดข้อมูลที่ใช้ในวิทยานิพนธ์นี้เป็นชุดข้อมูลชิ้นส่วนที่ถูกฉีกขาดด้วยมือที่ถูกจัดทำขึ้นเพื่อทดสอบในงานวิจัยนี้ โดยเป็นภาพสีที่มีเนื้อหาภายในชิ้นส่วนประกอบด้วยตัวอักษร รูปภาพ และตาราง ที่นำมาจากเอกสารต้นฉบับ 4 ชุด และชุดข้อมูลชิ้นส่วน

มาจากงานวิจัยที่มีผู้ทำอยู่เดิมเพื่อทำการทดสอบและวัดประสิทธิภาพเชิงเปรียบเทียบกับงานวิจัยอ้างอิง

- 1.5.3 การทดลองเพื่อเปรียบเทียบประสิทธิภาพของการหมุนกลับของชิ้นส่วนโดยวัดความถูกต้องจากชิ้นส่วนที่มีมุมเอียง -45, 45 และ 90 องศา กับงานวิจัยที่มีผู้ทดลองไว้ในลักษณะเดียวกัน [2]
- 1.5.4 การระบุรูปร่างขอบของชิ้นส่วนโดยวิธีฮัฟทรานสฟอร์ม เป็นการตรวจหาเส้นตรงที่มีในภาพจึงทำให้สามารถหาแนวหลักของเส้นขอบของชิ้นส่วน ซึ่งรูปร่างที่ได้จะมีรูปร่างคล้ายกับชิ้นส่วนจริงเพื่อนำข้อมูลที่ได้มาใช้ในการเชื่อมต่อชิ้นส่วน
- 1.5.5 ข้อมูลของชิ้นส่วนสำหรับการประกอบคือ ขนาดมุม ความยาวของด้านประกอบมุม และองศาของเส้นประกอบมุม

1.6 ขั้นตอนการศึกษา

- 1.6.1 กำหนดวัตถุประสงค์และขอบเขตของงานวิจัย ว่าต้องการศึกษางานวิจัยในหัวข้อนี้อย่างไร
- 1.6.2 ศึกษาวิธีการทำการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงที่เคยมีผู้ทำไว้ก่อนหน้านี้ เปรียบเทียบข้อดีและข้อด้อยของแต่ละวิธีการ เพื่อนำมาปรับปรุงและประยุกต์ใช้ให้เกิดประโยชน์กับงานวิจัยนี้
- 1.6.3 ตั้งสมมติฐานตามทฤษฎีหรือวิธีการที่ได้ศึกษาและกำหนดแนวทางในการวิจัย
- 1.6.4 เตรียมฐานข้อมูลงานวิจัยเพื่อนำมาใช้ทดลองในงานวิจัย โดยฉีกชิ้นส่วนเอกสารด้วยมือ และเก็บเป็นข้อมูลภาพด้วยเครื่องสแกน
- 1.6.5 นำฐานข้อมูลชิ้นส่วนที่เตรียมไว้มาทำการทดลอง โดยใช้โปรแกรมที่ได้พัฒนาขึ้น ทำการเก็บข้อมูลและผลลัพธ์ที่ได้จากการทดลอง เพื่อวิเคราะห์และวัดประสิทธิภาพ เพื่อนำผลการปรับปรุงเป็นแนวทางการพัฒนางานวิจัยต่อไป
- 1.6.6 ทำการสรุปผลการทดลองโดยเปรียบเทียบผลการทดลองในด้านความถูกต้องและเวลาที่ใช้

1.7 เครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัย

เครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย มีดังนี้

- 1.7.1 เครื่องสแกน Brother DCP-J315W
- 1.7.2 เครื่องคอมพิวเตอร์ส่วนบุคคลใช้หน่วยความจำ Intel CORE i5 จำนวน 1 เครื่อง
- 1.7.3 ระบบปฏิบัติการ Windows 8.1
- 1.7.4 โปรแกรม Adobe Photoshop CS6
- 1.7.5 โปรแกรม Matlab2013
- 1.7.6 ตัวอย่างกระดาษที่ถูกฉีกขาด

1.8 โครงสร้างของวิทยานิพนธ์

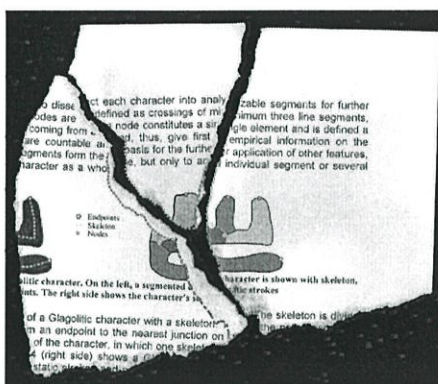
- วิทยานิพนธ์ฉบับนี้แบ่งเนื้อหาออกเป็น 6 บท แต่ละบทประกอบด้วยเนื้อหาดังนี้
- บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา ความมุ่งหมายและวัตถุประสงค์ของการศึกษา สมมติฐานของการศึกษา ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย ขอบเขตงานวิจัย ขั้นตอนการศึกษา และเครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัย
- บทที่ 2 กล่าวถึงงานวิจัยที่เกี่ยวข้องกับการจัดเรียงชิ้นส่วนให้ตั้งตรงและการประกอบ รวมทั้งข้อดีและข้อจำกัดของแต่ละวิธี
- บทที่ 3 กล่าวถึงความรู้พื้นฐานในเรื่องที่เกี่ยวข้องกับงานวิจัยได้แก่ การคำนวณขนาดและทิศทางของเวกเตอร์การเปลี่ยนแปลง (magnitude and orientation of gradient vectors), คำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (Histogram of accumulated Major Axis Length of ellipses ; HoARE) และฮัฟทรานสฟอร์ม (Hough transform)
- บทที่ 4 กล่าวถึงภาพรวมและขั้นตอนการดำเนินการในงานวิจัยนี้ โดยมีขั้นตอนการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงของงานวิจัยนี้ เริ่มจากการนำขอบของชิ้นส่วนออกเพื่อหาทิศการเรียงตัวของตัวหนังสือภายในชิ้นส่วน และทำการหมุนกลับชิ้นส่วนเพื่อให้อยู่ในระนาบตั้งตรง ขั้นตอนการหาตัวแทนรูปร่างของชิ้นส่วนโดยใช้ฮัฟทรานสฟอร์ม (Hough transform) และขั้นตอนการนำค่าตัวแทนฮัฟมาประยุกต์ในการประกอบชิ้นส่วน
- บทที่ 5 ผลการทดลองและการวิเคราะห์
- บทที่ 6 สรุปผลการทดลอง และแนวทางในการพัฒนา

บทที่ 2

งานวิจัยที่เกี่ยวข้อง

การประกอบข้อมูลที่ถูกฉีกขาดกลับคืนอาศัยหลักการพื้นฐานที่มีความสัมพันธ์คล้ายคลึงกับเทคนิคที่ใช้ในเกมส์ปริศนาโบราณที่เป็นการประกอบภาพปริศนาสองมิติ [3] และเกมส์ปริศนาจิ๊กซอว์ [4,5] อย่างไรก็ตามความแตกต่างสำคัญของเอกสารที่ถูกฉีกขาดกับเกมส์ปริศนาจิ๊กซอว์คือ รูปร่างของชิ้นส่วนที่ถูกนำมาประกอบ โดยรูปร่างของชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์จะมีลักษณะเหมือนหรือคล้ายกันกับชิ้นส่วนอื่นภายในชุด เนื่องจากแต่ละชิ้นส่วนเกิดจากการกำหนดรูปแบบโดยผู้คิดเกมส์ปริศนาไว้ก่อนแล้ว แต่ชิ้นส่วนที่ถูกฉีกขาดนั้นรูปร่างของชิ้นส่วนแต่ละชิ้นมีไม่แน่นอนและมีความแตกต่างกันเป็นส่วนใหญ่ เนื่องจากรูปร่างของชิ้นส่วนที่ถูกฉีกไม่ได้ถูกกำหนดรูปแบบของการฉีกไว้ อีกประเด็นหนึ่งคือ เนื้อหาภายในชิ้นส่วน โดยเนื้อหาภายในชิ้นส่วนของเกมส์ปริศนาจิ๊กซอว์มีลักษณะเป็นภาพ แต่เนื้อหาภายในชิ้นส่วนของเอกสารที่ถูกฉีกขาดประกอบด้วยตัวหนังสือเป็นส่วนใหญ่และอาจมีภาพหรือเส้นตารางประกอบด้วย โดยทั่วไปปริศนาการประกอบชิ้นส่วนสามารถแบ่งออกเป็นสองมิติและสามมิติ (เช่น การประกอบเครื่องบินดินเผาที่แตกให้คืนสภาพ [6], [7]) แม้ว่าการประกอบคืนเอกสารที่ถูกฉีกขาดจะใช้ลักษณะการประมวลผลเช่นเดียวกับปริศนาสองมิติ แต่อาจเกิดความซับซ้อนในการประกอบเนื่องจากลักษณะการขาดของกระดาษอาจเกิดเป็นชั้นที่แตกต่างกันซึ่งจะทำให้เกิดการเหลื่อมกันบางส่วนของชิ้นส่วน (ดูรูปที่ 2.1) [2] ขอบด้านนอกและขอบด้านในของรอยต่อสามชิ้นส่วนมีการทำเครื่องหมายสีที่ต่างกัน (หรือจุด/เส้นประ) เมื่อนำขอบของชิ้นส่วนมาประกอบกันขอบด้านนอกของชิ้นส่วนไม่สามารถประกอบหรือซ้อนทับกับชิ้นส่วนอื่นได้สนิท หากพิจารณาขอบด้านในของชิ้นส่วนจะพอดีกับขอบของชิ้นส่วนที่เกี่ยวข้อง ดังนั้นการพิจารณาขอบของชิ้นส่วนจึงควรพิจารณาขอบด้านในของชิ้นส่วนควบคู่ไปด้วย

ปริศนาสองมิติสามารถแบ่งออกเป็นวิธีการประกอบกลับภาพและการประกอบกลับที่ไม่เป็นภาพ [8] ในปัญหาการประกอบกลับข้อมูลที่ไม่เป็นภาพจะใช้เพียงรูปร่างของชิ้นส่วนเท่านั้นที่นำมาใช้พิจารณาในการจัดเรียง เนื่องจากชิ้นส่วนที่ไม่เป็นภาพไม่มีเนื้อหาหรือรูปภาพภายในชิ้นส่วนให้พิจารณา จึงไม่จำเป็นต้องนำรายละเอียดภายในชิ้นส่วนมาร่วมวิเคราะห์สำหรับการประกอบ เมื่อเทียบกับปัญหาการประกอบกลับข้อมูลที่เป็นภาพ รูปร่างและข้อมูลที่ถูกพิมพ์ลงบนชิ้นส่วน (เช่น ภาพพิมพ์หรือข้อความ) จะถูกนำมาพิจารณา หากขอบของชิ้นส่วนประกอบกันได้สนิทแต่เนื้อหาภายในชิ้นส่วนเช่น เฉดสีหรือข้อความภายในชิ้นส่วนบนคู่ชิ้นส่วนนั้น เมื่อพิจารณาแล้วไม่สอดคล้องกันหรือไม่ใช่ชิ้นส่วนที่เป็นคู่ประกอบกันที่แท้จริงจะไม่ถือว่าเป็นคำตอบที่ถูกต้อง ในทางกลับกันเมื่อชิ้นส่วนต่อกันสนิทและเนื้อหาภายในชิ้นส่วนนั้นสอดคล้องกันแสดงว่าชิ้นส่วนคู่กันเป็นคู่ชิ้นส่วนกันจริง ดังนั้นเนื้อหาภายในชิ้นส่วนจึงเป็นสิ่งสำคัญที่ควรนำมาร่วมวิเคราะห์เพื่อให้ชิ้นส่วนประกอบคำตอบที่ถูกต้อง



รูปที่ 2.1 กระดาษที่ถูกฉีกด้วยมือ

การพิจารณาเฉพาะรูปร่างหรือขอบของชิ้นส่วนปริศนาจะทำให้ผลลัพธ์การประกอบมีความผิดพลาด เนื่องจากหากเป็นชิ้นส่วนที่มีรูปร่างเหมือนกันและขนาดเท่ากัน จะทำให้ไม่สามารถแยกได้ว่าชิ้นส่วนนั้นควรอยู่ตำแหน่งใดหรือมาจากเอกสารต้นฉบับไหน จึงจำเป็นต้องใช้วิธีการประกอบกลับที่เป็นภาพ เพื่อแก้ปัญหาปริศนาทุกชิ้นส่วนที่รูปร่างเหมือนกัน (เช่นสีเหลี่ยมเช่น [9], ปริศนาเอกสารที่ถูกหั่น [10]) ดังนั้นการพิจารณาขอบและเนื้อหาภายในชิ้นส่วนจะช่วยแก้ปัญหาการประกอบเอกสารที่มีลักษณะเป็นแถบและมีขนาดเท่ากันให้ได้ผลลัพธ์ถูกต้องมากขึ้น

การใช้เทคนิคการวิเคราะห์การจัดเรียงระนาบของชิ้นส่วนเอกสารที่ถูกฉีกขาดให้ตั้งตรง เป็นงานส่วนหนึ่งที่สำคัญของการประกอบเอกสารที่ถูกฉีกขาดอัตโนมัติ ซึ่งสามารถนำมาประยุกต์ใช้ในงานที่แตกต่างกัน ในทางวิทยาศาสตร์ เช่น นิติวิทยา [11], [12] ในทางโบราณคดี เช่น นำไปใช้ประกอบในงานของนักนิรุกติศาสตร์ นักโบราณคดีและผู้เชี่ยวชาญด้านนิติเวชศาสตร์ นำไปใช้ในการประกอบต้นฉบับลายมือโบราณที่ได้รับการแยกส่วนเนื่องจากการเก็บรักษาที่ไม่ดี การประกอบต้นฉบับหรือหนังสือที่ได้รับการถูกทำลายโดยอุบัติเหตุให้กลับคืนเนื่องจากการล่มสลายของประวัติศาสตร์ทำให้ชิ้นเก็บวัตถุและเอกสารโบราณถูกทำลายรวมกันเป็นจำนวนมาก (เช่น [1]) และการประกอบกลับคืนแฟ้มข้อมูลที่ถูกฉีกด้วยมือในประเทศเยอรมันเพื่อสืบสวนประวัติศาสตร์เกี่ยวกับด้านหลักฐานการสอบสวน การฉีกเอกสารเพื่อการสืบสวนปราบปรามอาชญากรรมทางธุรกิจ [12], [13] เป็นต้น ซึ่งต้องการการแก้ไขอัตโนมัติเพื่อรักษาคุณค่าทางประวัติศาสตร์และวัฒนธรรม

2.1 การทำชิ้นส่วนให้ตั้งตรง

นอกจากการพิจารณารูปร่างภายนอกและเนื้อหาภายในชิ้นส่วนเป็นสิ่งจำเป็นในการจัดเรียงชิ้นส่วนให้มีประสิทธิภาพแล้ว การจัดเรียงชิ้นส่วนให้ตั้งตรงก่อนที่จะนำไปประกอบสามารถช่วยให้การประกอบกลับมีความรวดเร็วและถูกต้องมากขึ้น ดังนั้นการประกอบเอกสารที่ถูกฉีกขาดอัตโนมัติโดยทั่วไปนั้น การจัดระนาบชิ้นส่วนให้ตั้งตรงเป็นส่วนแรกๆ ที่ควรทำก่อนการประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาด หากการจัดเรียงชิ้นส่วนสามารถทำได้อย่างถูกต้อง จะมีส่วนช่วยให้การประกอบชิ้นส่วนกระดาษมีความถูกต้องมากยิ่งขึ้น ในการจัดระนาบชิ้นส่วนให้ตั้งตรงหากพิจารณาเฉพาะขอบของชิ้นส่วน เมื่อชิ้นส่วนมีรูปร่างลักษณะเหมือนกัน [14] จะไม่สามารถทราบได้ว่าชิ้นส่วนนั้นตั้งตรง

หรือไม่ จึงเป็นการยากที่จะประกอบชิ้นส่วนได้ถูกต้อง การทำชิ้นส่วนให้อยู่ในระนาบที่ตั้งตรงจึงจำเป็นต้องพิจารณาเนื้อหาภายในชิ้นส่วนด้วย การพิจารณาเนื้อหาภายในชิ้นส่วนโดยใช้เฉดสี เช่น สีกระดาษ สีพื้นหลังหรือสีตัวอักษร เป็นการยากใช้เป็นตัวช่วยทำให้กระดาษตั้งตรงได้ เนื่องจากเฉดสีสามารถจำแนกเฉพาะว่าชิ้นส่วนนี้น่าจะนำมาจัดกลุ่มกับชิ้นส่วนใด ชิ้นส่วนที่มีสีเหมือนกันน่าจะอยู่ในกลุ่มที่มาจากต้นฉบับเดียวกัน แต่เฉดสีไม่มีรูปแบบกำหนดหรือไม่สามารถบอกทิศทางที่สามารถระบุได้ว่าชิ้นส่วนตั้งตรงหรือไม่ การทำให้ชิ้นส่วนตั้งตรงนั้นจำเป็นต้องพิจารณาทิศการเรียงตัวของตัวหนังสือภายในชิ้นส่วนว่าส่วนใหญ่แล้วตัวหนังสืออยู่ในทิศใด ถ้าทิศของตัวหนังสือในชิ้นส่วนนั้นตั้งตรง มีแนวโน้มที่ชิ้นส่วนนั้นอยู่ในระนาบที่ตั้งตรงด้วยเช่นกัน

การทำให้ชิ้นส่วนอยู่ในระนาบที่ตั้งตรงนั้น หากใช้วิธีการประมาณค่าทิศทางการจัดเรียงตัวเบื้องต้นจากค่าฮิสโทแกรมของขนาดการเปลี่ยนแปลงของค่าเกรเดียนท์ (Magnitude Gradient) และทิศทางการเปลี่ยนแปลงของค่าเกรเดียนท์ (Direction of Gradient) [15] ทิศการเรียงตัวเบื้องต้นที่ได้อาจยังไม่ใช่ทิศการเรียงตัวหลักที่แท้จริง เนื่องจากการหาทิศทางการจัดเรียงตัวเบื้องต้นนี้อาจเกิดข้อผิดพลาดในกรณีที่ขอบของชิ้นส่วนถูกฉีกเป็นแนวเส้นตรงยาว จะทำให้ทิศการเรียงตัวหลักของชิ้นส่วนที่ได้เป็นทิศการเรียงตัวของขอบนั้นไม่ใช่ทิศของตัวหนังสือที่เราต้องการ และยังมีงานวิจัยที่นำแนวของตัวหนังสือภายในชิ้นส่วนมาวิเคราะห์เพื่อหาทิศการเรียงตัวหลักของตัวหนังสือโดยคำนวณค่าฮิสโทแกรมทิศทางของวงรี (Histogram of Number of fitted Ellipses; HoNE) [2] ซึ่งวิธีนี้สามารถหาแนวทิศการเรียงตัวของตัวหนังสือได้จริง แต่จะไม่สามารถใช้กับชิ้นส่วนที่มีเนื้อหาภายในน้อย เช่น สองถึงสามคำ และไม่สามารถหาทิศการเรียงตัวของชิ้นส่วนที่แท้จริงได้หากภายในชิ้นส่วนมีข้อมูลเป็นภาพมากกว่าตัวหนังสือหรือมีเส้นตาราง ด้วยข้อจำกัดดังกล่าวงานวิจัยฉบับนี้จึงได้นำเสนอวิธีการจัดเรียงให้ชิ้นส่วนอยู่ในระนาบที่ตั้งตรงโดยเริ่มจากการคำนวณฮิสโทแกรมผลรวมของความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (Histogram of Accumulated Major Axis Length of fitted Ellipse; HoARE) ซึ่งสามารถแก้ปัญหาข้อมูลภายในชิ้นส่วนที่มีปริมาณน้อย มีรูปภาพมากกว่าตัวหนังสือและเนื้อหาภายในชิ้นส่วนมีเส้นตารางที่กล่าวมาข้างต้นได้ หลังจากขั้นตอนนี้ชิ้นส่วนที่ถูกฉีกขาดเหล่านี้จะอยู่ในระนาบเดียวกัน(ตั้งตรง) จากนั้นใช้วิธีฮัฟทรานสฟอร์ม (Hough Transform) ซึ่งมีบทบาทสำคัญในการรู้รูปร่างของชิ้นส่วน (Shape) โดยวิธีการตรวจหาเส้นตรงที่มีในภาพ จึงทำให้สามารถหาแนวหลักของเส้นขอบของชิ้นส่วนเพื่อนำข้อมูลที่ได้มาใช้ในการเชื่อมชิ้นส่วนที่ถูกฉีกขาด

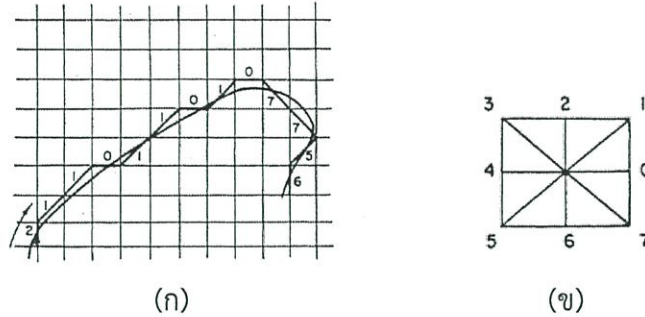
2.2 การประกอบชิ้นส่วน

ในหัวข้อนี้จะกล่าวถึงภาพรวมของการประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์และการประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาดในงานวิจัย [13,18,19,20] รวมทั้งข้อจำกัดและลักษณะของชิ้นส่วน

2.2.1 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ของ H. FREEMAN และ L. GARDER [13]

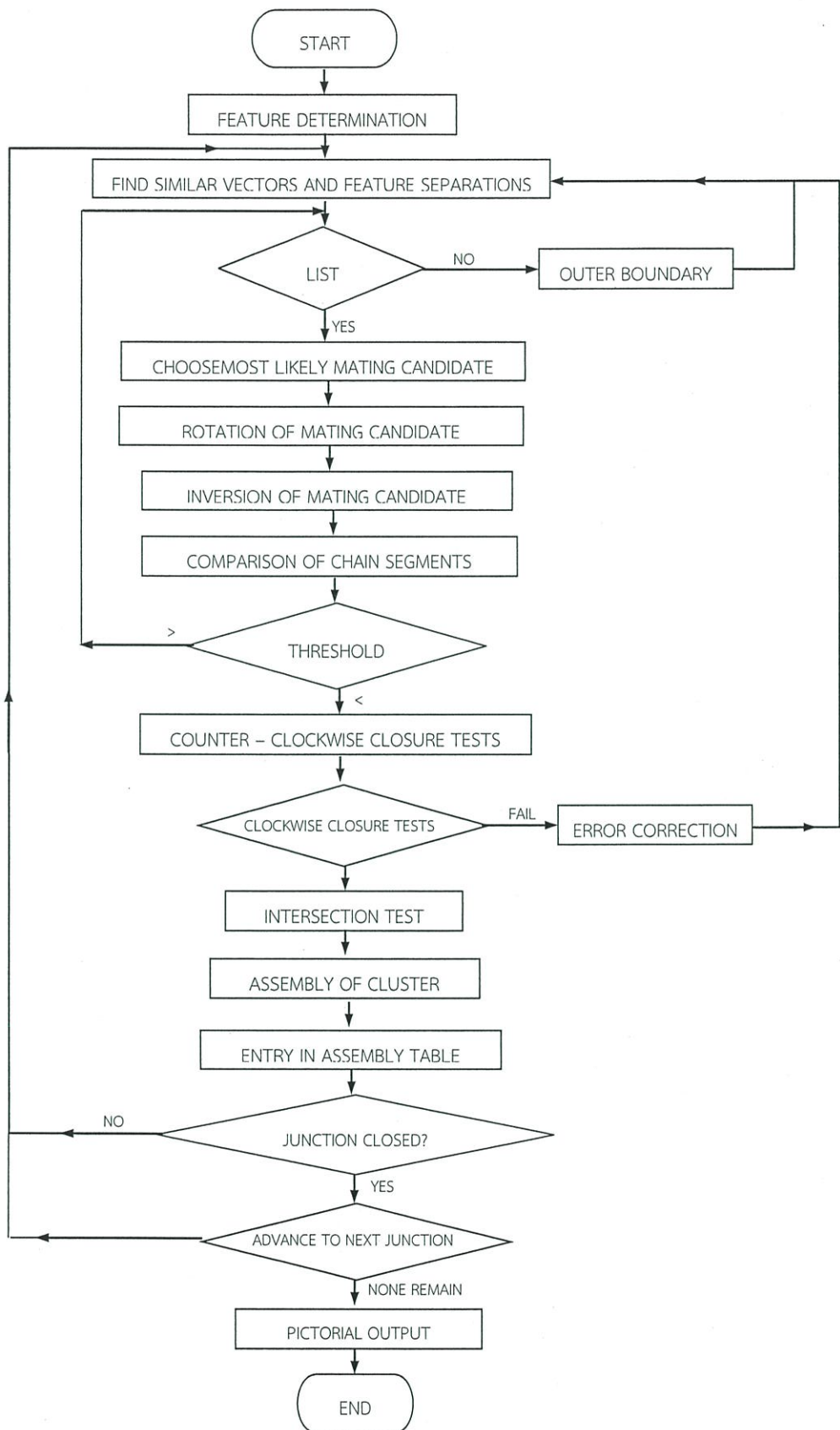
การเชื่อมชิ้นส่วนนี้เป็นการเชื่อมชิ้นส่วนแบบไม่นำรายละเอียดภายในชิ้นส่วนมาร่วมวิเคราะห์ จะพิจารณาเฉพาะรูปร่างของชิ้นส่วน โดยทุกชิ้นส่วนจะระบุการหาขอบด้วยเทคนิคเซนเอนโคคด์

(Chain-encoding technique) ด้านที่เป็นส่วนโค้งของขอบของชิ้นส่วนเมื่อผ่านวิธีของเซนแล้วจะถูกแทนด้วยเส้นตรง (ดูรูปที่ 2.2(ก)) ซึ่งเส้นตรงแต่ละส่วนสามารถแสดงได้โดยลำดับเลขแปดตัว (ดูรูปที่ 2.2(ข)) ในทิศตามเข็มนาฬิกา ดังนั้นข้อมูลขอบของรูปที่ 2.2(ก) สามารถระบุทิศของขอบด้วยเซนโค้ดเป็น 2110110107756 (ตามทิศของลูกศร) ข้อมูลขอบที่ได้จากวิธีเซนเอนโค้ดดังจะถูกเก็บเพื่อทำการเปรียบเทียบหาชิ้นส่วนที่เหมาะสมต่อไป

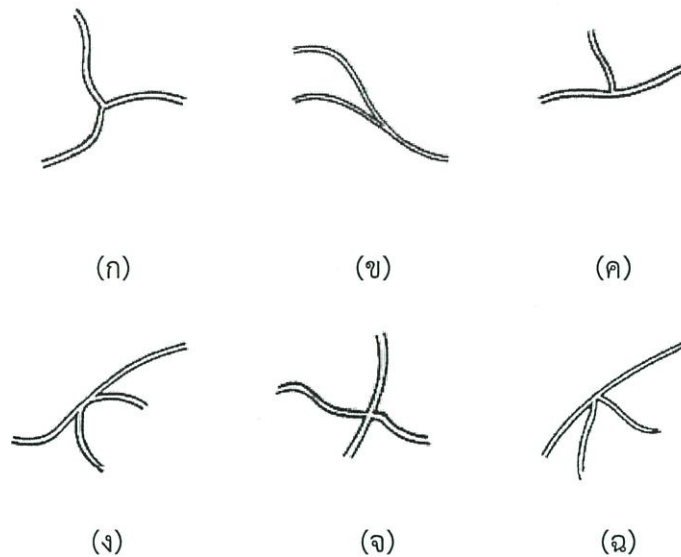


รูปที่ 2.2 แผนผังเซนเอนโค้ดตั้ง (Chain encoding scheme)

ในงานวิจัยของ H. FREEMAN และ L. GARDER ได้นำเสนอขั้นตอนการประกอบชิ้นส่วนจิ๊กซอว์แสดงดังรูปที่ 2.3 โดยเริ่มแรกจะพิจารณาแต่ละเซนเลท (chainlet) ซึ่งประกอบด้วยความยาวเซน (chain length) ความยาวของดิสแทนเวกเตอร์ (distance vector) จากจุดเริ่มต้นถึงจุดปลายมุมที่ทำกับแกน x ของเวกเตอร์ พื้นที่ระหว่างเซนเลทและดิสแทนเวกเตอร์ และระยะทางที่ยาวสุดสั้นสุดจากดิสแทนเวกเตอร์ถึงเซนเลท จากนั้นจะใช้คุณลักษณะที่กล่าวมาข้างต้นมาทำการจัดกลุ่ม แล้วทำการจับคู่ชิ้นส่วนที่มีความเป็นไปได้ที่จะเข้าคู่กัน โดยพิจารณาจากความยาวแต่ละด้านของขอบชิ้นส่วน มุมของแต่ละชิ้นส่วนที่นำมาประกบกันสนิทหรือไม่ แต่งานวิจัยของ H. FREEMAN และ L. GARDER ยังมีข้อจำกัดคือ ไม่สามารถประกอบชิ้นส่วนที่มีเนื้อหภายในได้ถูกต้องทุกชิ้นเนื่องจากเนื้อหภายในชิ้นส่วนไม่ได้ถูกนำมาวิเคราะห์ร่วมกับการประกอบ และในมุมที่เป็นจุดเชื่อมของชิ้นส่วนที่นำมาต่อกันต้องมีจำนวนไม่เกินสามชิ้น (รูปที่ 2.4(ก), 2.4(ข), 2.4(ค)) หากที่จุดเชื่อมนั้นประกอบด้วยชิ้นส่วนมากกว่าสามชิ้นแล้วจะไม่สามารถประกอบได้ (รูปที่ 2.4(ง), 2.4(จ), 2.4(ฉ))



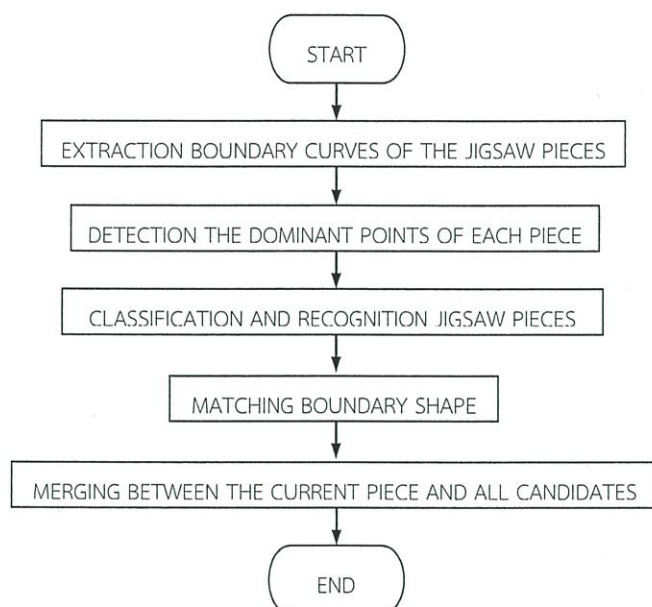
รูปที่ 2.3 ขั้นตอนการประกอบชิ้นส่วนเกมส์ปริศนาจิกซอร์ด้วยวิธีของ H. FREEMAN และ L. GARDER



รูปที่ 2.4 มุมเชื่อมที่เชื่อมด้วยชิ้นส่วนสามชิ้นและสี่ชิ้น

2.2.2 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์ของ Feng-Hui Yao และ Gui-Feng Shao [18]

การประกอบชิ้นส่วนเกมส์ปริศนาโดยใช้เรขาคณิตในการคำนวณรูปร่างส่วนโค้งออกและโค้งเข้าของชิ้นส่วนด้วยหลักการคณิตศาสตร์และวิเคราะห์เนื้อหา (เฉดสี) ภายในชิ้นส่วนเพื่อให้สามารถประกอบจิ๊กซอว์ได้ โดยขั้นตอนของกระบวนการประกอบแสดงดังรูปที่ 2.5 รูปแบบชิ้นส่วนที่นำมาใช้ทดสอบมีทั้งหมด 18 แบบ (รูปที่ 2.6) โดยมีชิ้นส่วน 12 แบบที่เป็นชิ้นส่วนที่มีจริงบนโลก เริ่มแรกชิ้นส่วนจะถูกแยกโดยสีและส่วนโค้งของชิ้นส่วน จากนั้นมุมของชิ้นส่วนสามารถหาได้จากมุมภายในของแต่ละมุมจะมีขนาดใกล้เคียงกัน ระยะทางจากจุดศูนย์กลางถึงมุมเป็นระยะทางที่เท่ากัน ส่วนโค้งสามารถหาได้จากการนำมุมแต่ละมุมมาหาจุดกึ่งกลาง ด้านที่มีส่วนที่โค้งเข้าจะมีระยะทางระหว่างส่วนโค้งเข้ากับจุดศูนย์กลางสั้นที่สุด ในทางกลับกันด้านที่มีส่วนที่โค้งออกระยะทางจากจุดศูนย์กลางชิ้นส่วนถึงส่วนโค้งออกจะเป็นระยะทางที่ยาวที่สุด เมื่อได้มุมและส่วนโค้งเข้าโค้งออกของแต่ละชิ้นส่วนแล้ว ชิ้นส่วนจะถูกจัดกลุ่ม 18 กลุ่มตามรูปแบบชิ้นส่วน 18 แบบ และถูกคำนวณความกว้างและความสูงของส่วนโค้งเข้าโค้งออกเพื่อหาส่วนโค้งที่ต่อพอดีกัน จากนั้นทำการวิเคราะห์เนื้อหา (เฉดสี) บริเวณรอบส่วนโค้งเข้าและโค้งออกของชิ้นส่วนที่ต่อกันสนิท หากช่วงสีที่อยู่บริเวณส่วนโค้งเข้าและโค้งออกนั้นอยู่ในช่วงเรขาคณิตเดียวกันหมายความว่า การเชื่อมต่อกันของชิ้นส่วนคู่นั้นถูกต้องสุดท้ายชิ้นส่วนจะถูกเชื่อมกัน



รูปที่ 2.5 ขั้นตอนการประกอบชิ้นส่วนเกมสปริศนาจิ๊กซอว์ด้วยวิธีของ Feng-Hui Yao และ Gui-Feng Shao



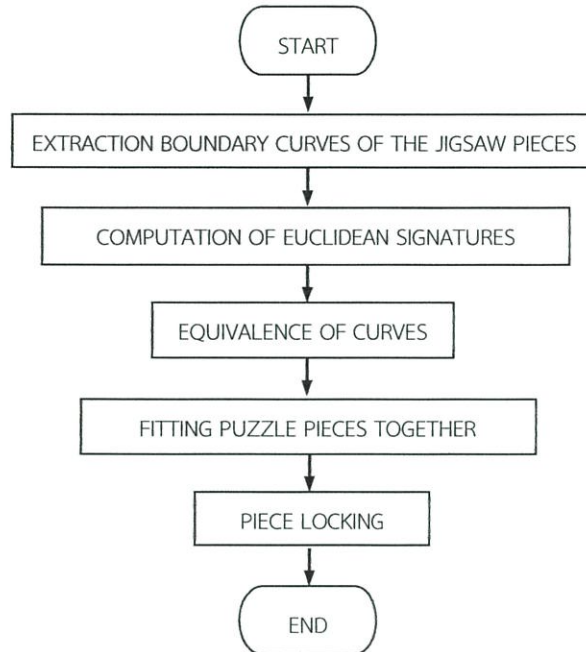
รูปที่ 2.6 ภาพชิ้นส่วน 18 แบบ

ผลลัพธ์ของการประกอบชิ้นส่วนของ Feng-Hui Yao และ Gui-Feng Shao มีความถูกต้องสูง เหมาะสำหรับการประกอบชิ้นส่วนตามรูปแบบที่งานวิจัยกำหนด (18 แบบ) ซึ่งไม่สามารถใช้กับการประกอบชิ้นส่วนที่มีรูปแบบที่นอกเหนือจากงานวิจัยหรือชิ้นส่วนที่มีรูปแบบไม่ตายตัวได้ และหากเฉดสีภายในชิ้นส่วนมีความใกล้เคียงกันมากหลายชิ้นส่วนจะส่งผลให้ชิ้นส่วนมีการจับคู่ผิด คำตอบที่ได้จะไม่ถูกต้อง

2.2.3 การประกอบชิ้นส่วนเกมสปริศนาจิ๊กซอว์และการประกอบชิ้นส่วนที่ถูกฉีดยาด้วยวิธีของ Daniel J. Ho และ Peter J. Olver [19]

ขั้นตอนการประกอบชิ้นส่วนปริศนาจิ๊กซอว์และชิ้นส่วนที่ถูกฉีดยาด้วยวิธี Daniel J. Ho และ Peter J. Olver แสดงดังรูปที่ 2.7 เป็นการประกอบชิ้นส่วนซึ่งสามารถประกอบรูปแบบชิ้นส่วนที่ไม่เป็นรูปแบบตายตัวได้โดยไม่มีข้อจำกัดของลักษณะชิ้นส่วน เนื่องจากวิธีการประกอบของงานวิจัยนี้เกิดจากการนำขอบชิ้นส่วนมาระบุเป็นพิกเซลรูปร่างทำให้ได้รูปร่างใกล้เคียงกับชิ้นส่วนมากที่สุด จากนั้นจะทำการสุ่มชิ้นส่วนชิ้นแรกมาเพื่อจะนำพิกเซลขอบของชิ้นส่วนนั้นมาคำนวณความเข้ากันกับชิ้นส่วนอื่นๆ เพื่อจับคู่ความเข้ากันได้ระหว่างขอบของชิ้นส่วนสองชิ้น โดยการวางขอบทั้งสองด้านของแต่ละชิ้นคู่ของขอบต่อกัน และนับพิกเซลในพื้นที่ที่มีความแตกต่างกัน (ดูรูปที่ 2.8) พื้นที่สีเขียวแสดง

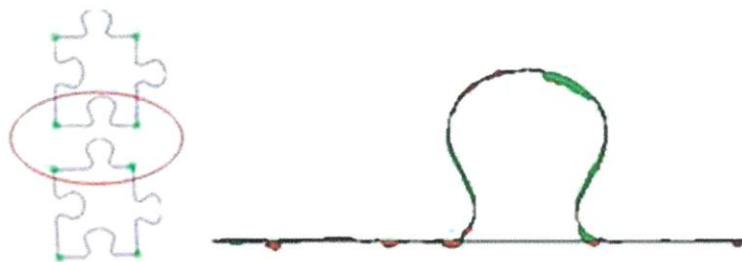
ถึงพื้นที่ในขอบชั้น 1 ที่ไม่อยู่ในขอบชั้น 2 ในขณะที่พื้นที่สีแดงแสดงถึงพื้นที่ในขอบชั้น 2 ที่ไม่อยู่ในขอบชั้น 1 จากนั้นผลรวมของทั้งสองพื้นที่ (เขียวและแดง) จะถูกเก็บไว้เพื่อนำมาเปรียบเทียบกับการหาผลรวมของขอบชั้น 1 กับชั้นอื่นๆ หากผลรวมคู่ขอบของชั้น 1 กับชั้นใดๆต่ำ แสดงว่าคู่ขอบนั้นคือคู่ขอบที่แท้จริง สุดท้ายชิ้นส่วนจะถูกทำให้พอดีและถูกยึดติดกัน



รูปที่ 2.7 การประกอบชิ้นส่วนเกมส์ปริศนาจิ๊กซอว์และชิ้นส่วนที่ถูกฝึกขาดด้วยวิธีของ Daniel J. Ho และ Peter J. Olver

Green area = 218

Red area = 270



รูปที่ 2.8 ภาพแสดงผลรวมของผลต่างของการจับคู่ขอบ

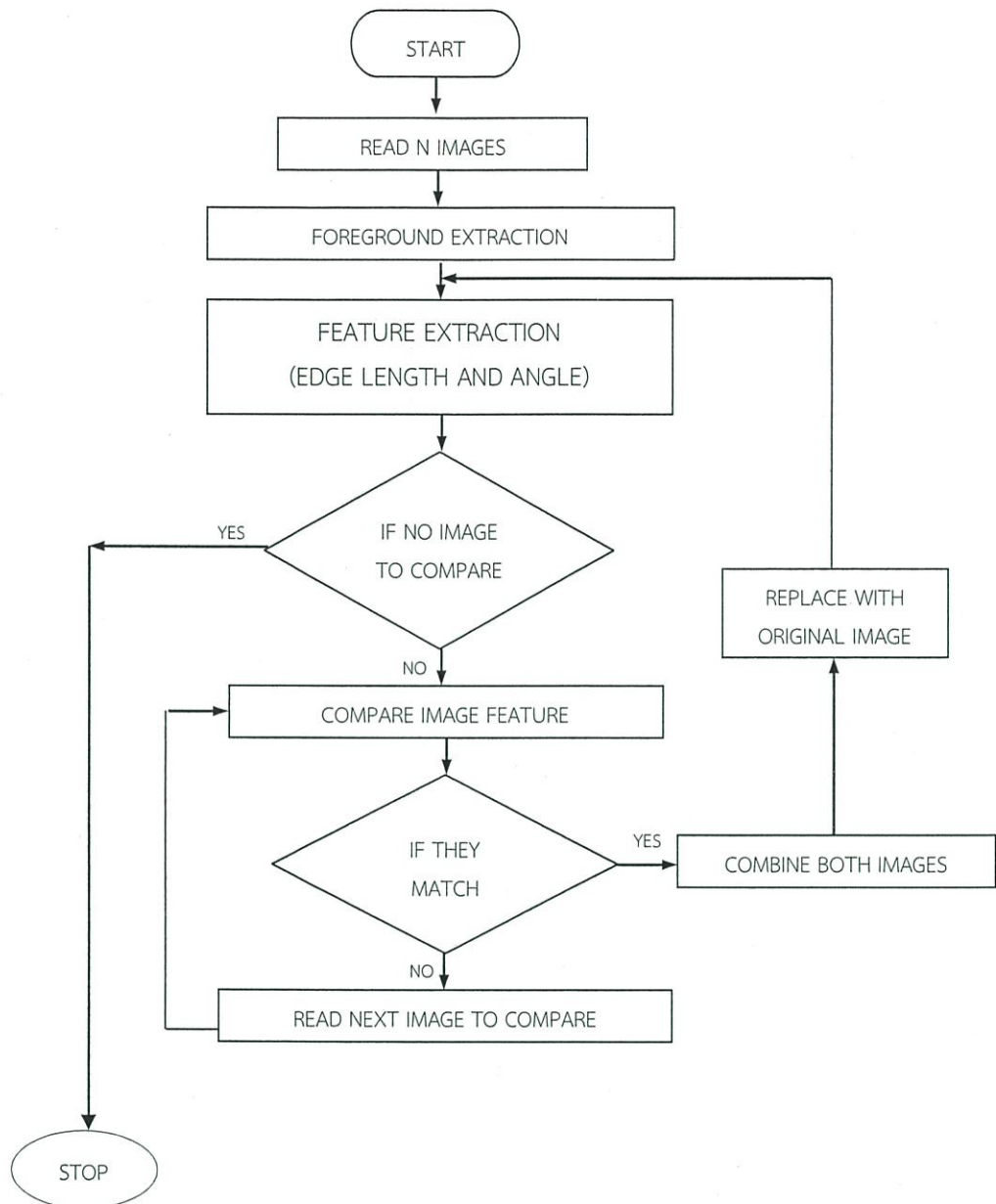
การประกอบชิ้นส่วนด้วยวิธีของ Daniel J. Ho และ Peter J. Olver สามารถแก้ปัญหาชิ้นส่วนที่มีรูปแบบไม่ปกติหรือมีหลากหลายรูปแบบได้ แต่ชิ้นส่วนที่นำมาทดลองนั้นเป็นชิ้นส่วนที่สร้างขึ้นจากโปรแกรมคอมพิวเตอร์ ซึ่งจะทำให้วิธีนี้ใช้ได้กับเฉพาะชิ้นส่วนที่คู่ชิ้นส่วนมีการซ้อนทับกัน

สนิทจริง ข้อมูลภาพที่รับเข้าไปต้องเป็นพิกเซลขอบที่ปิด คือจุดต้นและจุดสุดท้ายของพิกเซลขอบต้องเป็นจุดเดียวกันหรือเป็นจุดที่อยู่ถัดจากกัน นอกจากนี้ระยะเวลาที่ใช้ในการประกอบค่อนข้างนาน

2.2.4 การประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาดด้วยวิธีของ Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade [20]

ขั้นตอนการประกอบชิ้นส่วนปริศนาจิ๊กซอว์และชิ้นส่วนที่ถูกฉีกขาดด้วยวิธี Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade แสดงดังรูปที่ 2.9 งานวิจัยนี้ทำการประกอบชิ้นส่วนเอกสารฉีกขาด ซึ่งวิธีการประกอบเกิดจากการพิจารณาเฉพาะขอบของชิ้นส่วน โดยเริ่มแรกชิ้นส่วนจะถูกนำเนื้อหาภายในชิ้นส่วนออก จากนั้นขอบจะถูกวิเคราะห์หามุมและความยาวด้านประกอบแต่ละมุมของชิ้นส่วน หากชิ้นส่วนไม่มีชิ้นส่วนที่จะนำมาเปรียบเทียบ กระบวนการประกอบจะสิ้นสุด แต่ถ้าชิ้นส่วนมีชิ้นส่วนสำหรับเปรียบเทียบชิ้นส่วนคู่นั้นจะถูกวิเคราะห์ว่าเข้าคู่กันหรือไม่โดยการพิจารณามุมและความยาวด้านประกอบมุมของชิ้นส่วน หากชิ้นส่วนที่นำมาเปรียบเทียบนั้นเข้าคู่กัน ชิ้นส่วนคู่นั้นจะถูกประกอบกลายเป็นชิ้นส่วนหนึ่งชิ้น แล้วถูกนำไปหามุมและความยาวด้านประกอบแต่ละมุมของชิ้นส่วนใหม่ แต่หากคู่ชิ้นส่วนนั้นไม่เข้าคู่กัน ชิ้นส่วนชิ้นอื่นจะถูกนำมาพิจารณาการเข้าคู่กัน

การประกอบชิ้นส่วนเอกสารกระดาษที่ฉีกขาดด้วยวิธีของ Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade สามารถแก้ปัญหาชิ้นส่วนที่มีรูปแบบไม่ปกติหรือมีหลากหลายรูปแบบได้ แต่ชิ้นส่วนที่นำมาทดลองนั้นเป็นชิ้นส่วนที่ขอบของรอยฉีกมีลักษณะเป็นเส้นตรง



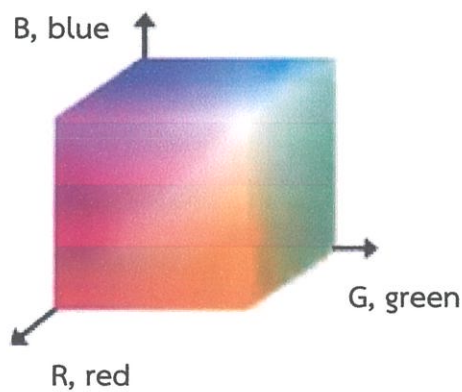
รูปที่ 2.9 การประกอบชิ้นส่วนกระดาษที่ถูกฉีกขาดด้วยวิธีของ Ms.Poonam D.Kesarkar, Ms.Mridula RC. Prasad และ Prof.S.L.Tade

บทที่ 3

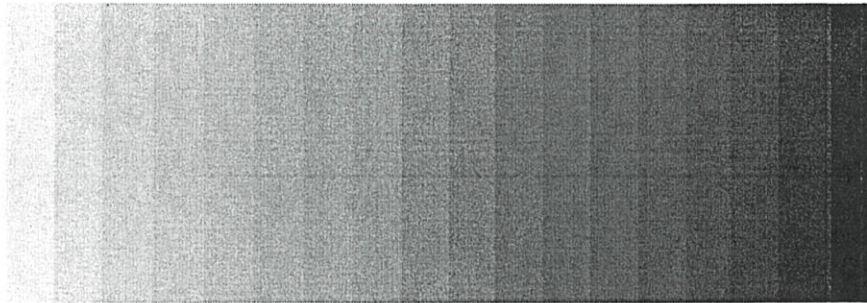
ความรู้พื้นฐานในเรื่องที่เกี่ยวข้องกับงานวิจัย

3.1 การแปลงภาพสีเป็นขาวดำ (Image Binarization)

การแปลงภาพสีให้เป็นภาพขาวดำนั้นต้องแปลงภาพสีเป็นเกรสเกล (Gray scale) ก่อนแล้วจึงแปลงภาพเกรสเกลเป็นภาพขาวดำอีกที การแปลงภาพเกรสเกลเป็นภาพขาวดำจำเป็นต้องกำหนดค่าธรชโฮลด์ (Threshold) ซึ่งมีไว้สำหรับกำหนดค่าความสว่างของแต่ละพิกเซล (Pixel) หรือจุดภาพ ในภาพเมื่อแปลงเป็นภาพขาวดำแล้วพิกเซลดังกล่าวควรจะเป็นสีดำหรือสีขาว เพื่อให้ได้ภาพขาวดำที่เหมาะสม เช่น ค่าธรชโฮลด์ถูกกำหนดเป็น 128 แล้วอ่านค่าความสว่างในพิกเซลแรกของรูปได้ 150 ดังนั้นเมื่อทำการแปลงเป็นภาพขาวดำแล้วพิกเซลแรกนี้จะต้องเป็นสีขาว เป็นต้น จากตัวอย่างข้างต้นจะได้ว่าระบบสี RGB คือระบบที่มีค่าของสีแดง (R) สีเขียว (G) และสีน้ำเงิน (B) ค่าใดค่าหนึ่งหรือหลายๆ ค่ารวมกัน โดยแต่ละสีจะมีค่าตั้งแต่ 0 ถึง 255 ดังรูปที่ 3.1 ภาพเกรสเกลไม่ใช่ภาพขาวดำเนื่องจากยังมีระดับของความสว่างที่แตกต่างกันอยู่หลายระดับ ซึ่งจำนวนระดับจะขึ้นกับจำนวนบิตต่อพิกเซล โดยทั่วไประบบรับภาพถ่ายจะกำหนดให้ข้อมูลพิกเซลถูกแทนด้วย 8 บิตต่อพิกเซล ซึ่งสัมพันธ์กับค่าจำนวนระดับในช่วง 0 ถึง 255 กล่าวคือภาพเกรสเกลเป็นภาพที่มีค่าสี R G B เท่ากันหมดทำให้ภาพออกมาในโทนสีขาว-เทา-ดำ ดังรูปที่ 3.2 แต่ภาพขาวดำ(Binary Image) เป็นภาพที่มีค่าเพียงสองระดับคือ 0 หรือ 1 เท่านั้น



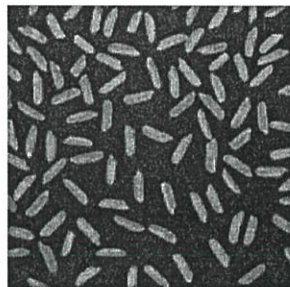
รูปที่ 3.1 ระบบสี RGB



รูปที่ 3.2 ระบบสีเกรสเกล

เธอร์ชโพลด์เป็นการแปลงภาพเมล็ดข้าวในระบบเกรสเกลดังรูปที่ 3.3 ซึ่งมีค่าอยู่ระหว่าง 0-255 ให้เป็นภาพขาวดำดังรูปที่ 3.4 โดยมีเงื่อนไขว่า ถ้าความเข้มแสงของจุดภาพใดมีค่าต่ำกว่าหรือเท่ากับค่าเธอร์ชโพลด์ที่กำหนดไว้จะให้จุดภาพนั้นมีค่าเป็น 0 หรือเป็นสีดำ และจุดภาพใดที่มีค่าสูงกว่าค่าเธอร์ชโพลด์ให้จุดภาพนั้นมีค่าเป็น 1 หรือสีขาว เธอร์ชโพลด์เป็นการทำค่าความเข้มแสงของพิกเซลให้มีค่าเพียง 2 สถานะคือขาวและดำ ดังตัวอย่างในสมการ 3.1 โดยสามารถแสดงความสัมพันธ์แบบกราฟได้ดังรูปที่ 3.5

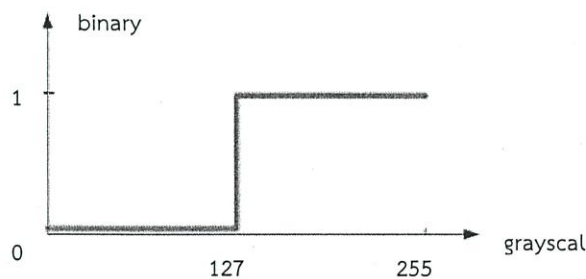
$$\text{Grayscale} = \begin{cases} 255 & ; \text{if grayscale} > 127 \\ 0 & ; \text{else} \end{cases} \quad (3.1)$$



รูปที่ 3.3 ภาพในระบบสีเกรสเกล



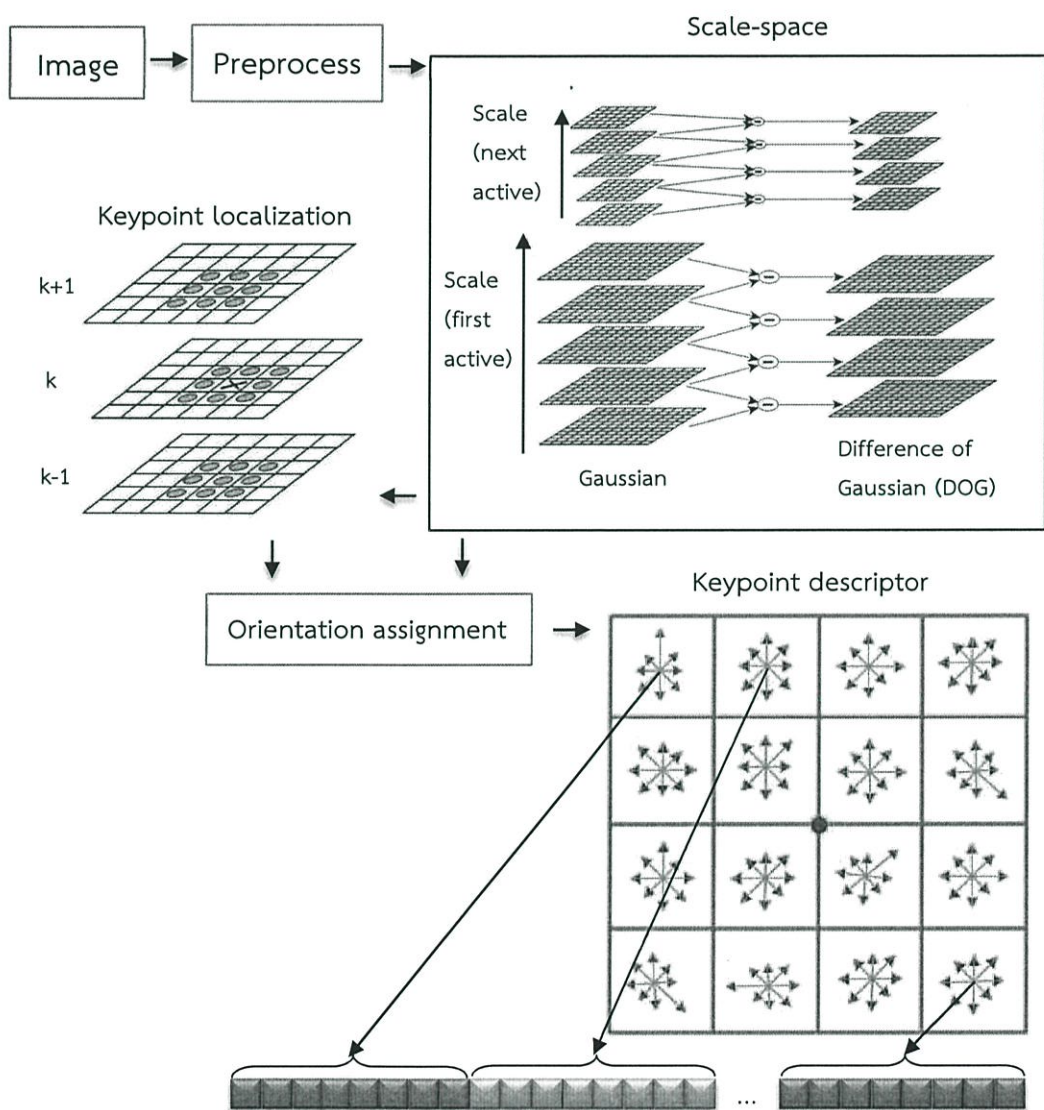
รูปที่ 3.4 ภาพที่แบ่งเธอร์ชโพลด์



รูปที่ 3.5 กราฟเกรสเกล

3.2 การคำนวณขนาดและทิศทางของเวกเตอร์การเปลี่ยนแปลง (Magnitude and Orientation of Gradient Vectors)

การคำนวณขนาดและทิศทางของเวกเตอร์เป็นการใช้เทคนิคที่ได้ตัวอย่างมาจากอัลกอริทึม SIFT ของ Lowe โดยวิธีการ SIFT ย่อมาจาก Scale Invariant Feature Transform เป็นวิธีการหา ลักษณะเด่นแบบท้องถิ่นที่ไม่ขึ้นกับการเปลี่ยนขนาดและมุมของภาพ จากงานวิจัย [21,22] วิธีการ SIFT ประกอบด้วย 4 ขั้นตอนสำคัญ คือ 1) การสร้างปริภูมิค่าในมิติขนาดระยะทาง 2) การกำหนด ตำแหน่งจุดสนใจ 3) การกำหนดทิศทางของจุดสนใจ 4) การสร้างคำอธิบายลักษณะเด่นของภาพ ดัง รูปที่ 3.6



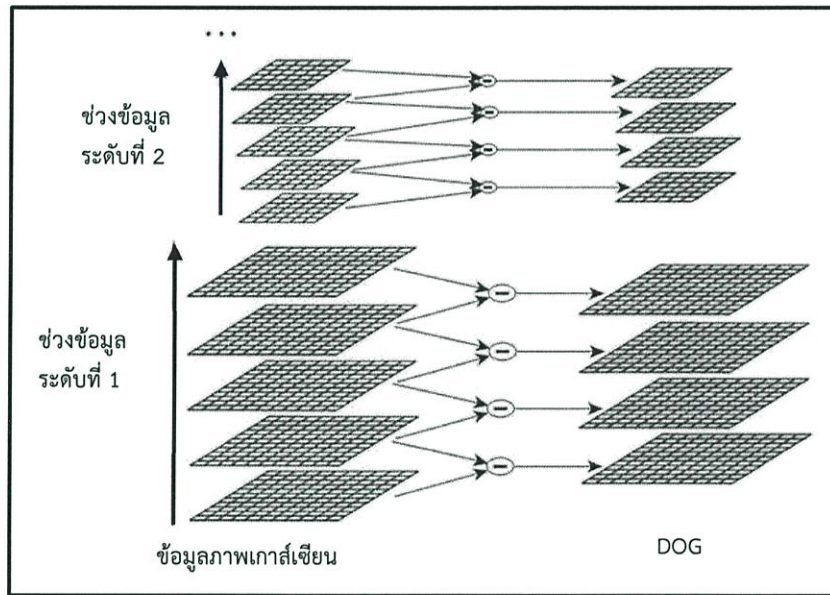
รูปที่ 3.6 ขั้นตอนการหาลักษณะเด่นของภาพด้วย SIFT

- 1) ขั้นตอนของการค้นหาจุดสนใจของภาพ สามารถนิยามได้ด้วยค่า $L(x, y, \sigma)$ ซึ่งได้จากการคอนโวลูชันภาพ $I(x, y)$ ด้วยตัวกรองแบบเกาส์เซียนหลายระดับ โดยสามารถหาได้จากสมการที่ 3.2 และสมการที่ 3.3

$$L(x, y, \sigma) = (G(x, y, \sigma) * I(x, y)) \quad (3.2)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-((x^2 + y^2) / 2\sigma^2)} \quad (3.3)$$

เมื่อ $G(x, y)$ เป็นตัวกรองแบบเกาส์เซียน รายละเอียดการกำหนดค่า σ อธิบายใน [23] ข้อมูลภาพที่ผ่านการกรองแล้วจะต้องหาค่าความแตกต่างของแผ่นข้อมูลภาพระหว่างชั้นของข้อมูลที่อยู่ติดกัน ด้วยวิธีการ Difference-of-Gaussian (DoG) ดังรูปที่ 3.7 โดยค่า DoG สามารถหาได้จากแผ่นข้อมูลภาพ $L(x, y, \sigma_1)$ ที่กรองด้วยค่าตัวกรองเกาส์เซียน σ_1 ในสมการ 3.4 และแผ่นข้อมูลภาพ $L(x, y, \sigma_2)$ ที่กรองด้วยค่าตัวกรองเกาส์เซียน σ_2 ในสมการ 3.5 และค่า DOG ที่ได้แสดงดังสมการที่ 3.6



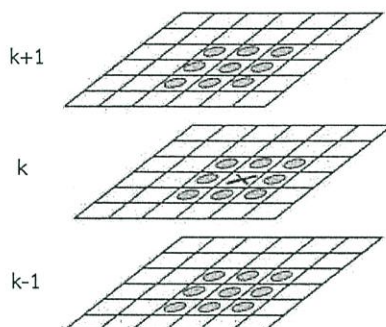
รูปที่ 3.7 การหาค่า DoG ของแผ่นภาพในแต่ละช่วงข้อมูล

$$L(x, y, \sigma_1) = (G(x, y, \sigma_1) * I(x, y)) \quad (3.4)$$

$$L(x, y, \sigma_2) = (G(x, y, \sigma_2) * I(x, y)) \quad (3.5)$$

$$DOG = L(x, y, \sigma_2) - L(x, y, \sigma_1) \quad (3.6)$$

- 2) ขั้นตอนต่อไปคือการหาค่าสูงสุดและต่ำสุดเพื่อเลือกจุดสนใจ จุดข้อมูลภาพจะถูกเทียบค่ากับจุดรอบข้าง 8 จุดและ 9 จุดในชั้นที่อยู่ติดกัน แสดงดังรูปที่ 3.8 จุดที่พิจารณาจะเป็นจุดในชั้นที่ k จุดสนใจที่เลือกจะเป็นค่าสูงสุดหรือค่าต่ำสุดในบริเวณที่พิจารณาเท่านั้น



รูปที่ 3.8 การพิจารณาจุดของภาพเพื่อเลือกเป็นจุดสนใจ

เมื่อได้จุดสนใจแล้ว จะทำการหาขนาด $m(x, y)$ และทิศทาง $\theta(x, y)$ ของจุดภาพที่ผ่านการกรองทั้งหมดเพื่อนำไปสร้างลักษณะเด่นของภาพดังสมการ 3.7 และ 3.8

$$m(x, y) = \sqrt{(Lx)^2 + (Ly)^2} \quad (3.7)$$

$$\theta(x, y) = \arctan((Ly)/(Lx)) \quad (3.8)$$

โดย Lx และ Ly คือค่าเกรเดียนต์ในแนวตั้งและแนวนอนตามลำดับ

หลังจากนั้นลักษณะเด่นของภาพจะถูกสร้างโดยให้พื้นที่ที่สนใจมีจุดสนใจเป็นจุดศูนย์กลางแล้วแบ่งพื้นที่สนใจเป็นพื้นที่ย่อยขนาด 4×4 ดังรูปที่ 3.9 การคำนวณแ่งความถี่หรือฮิสโทแกรมของทิศทางเกรเดียนต์ตามขนาดของเกรเดียนต์ในแต่ละพื้นที่ย่อยนั้น ทิศทางของจุดจะถูกแบ่งออกเป็น 8 กลุ่มทิศทาง หรือฮิสโทแกรมบinnedดังสมการที่ 3.9

$$h_{r(l,m)}(k) = \sum_{x,y \in r(l,m)} m(x, y)(1 - |\theta_k(x, y) - c_k|/\Delta_k) \quad (3.9)$$

เมื่อ $\theta_k(x, y)$ เป็นสมาชิกของกลุ่ม k

c_k คือ ทิศทางของศูนย์กลางของกลุ่ม k

Δ_k คือ ช่วงทิศทางของกลุ่ม k

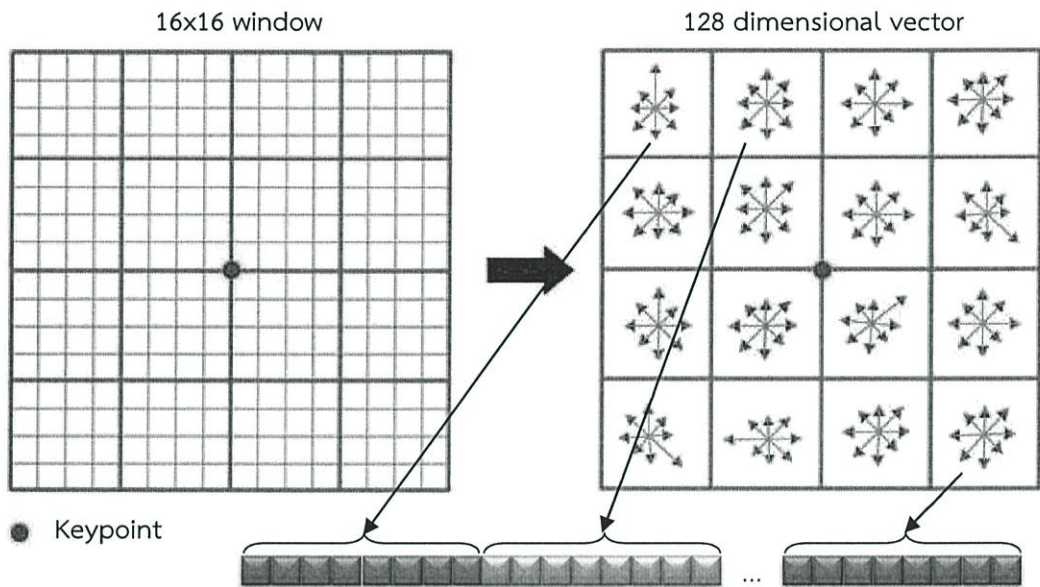
(x, y) คือ จุดของภาพในพื้นที่ย่อย

เมื่อคำนวณครบทุกพื้นที่ย่อยแล้วนำค่าแ่งความถี่ของทิศทางเกรเดียนต์มาเรียงต่อกันจะได้เวกเตอร์ลักษณะเด่นของภาพที่หาจากวิธีการ SIFT (SIFT vector) ขนาด 128 ข้อมูล (128 คือจำนวนฮิสโทแกรมของพื้นที่ 4×4 พื้นที่ละ 8 ทิศ = $4 \times 4 \times 8$) ดังสมการที่ 3.10

$$u = [h_{r(1,1)}, \dots, h_{r(l,m)}, \dots, h_{r(4,4)}] \quad (3.10)$$

หลังจากนั้นจะทำการแปลงค่าลักษณะเด่นที่ได้ให้อยู่ในรูปของค่ามาตรฐานหนึ่งหน่วยดังรูปที่

3.9



รูปที่ 3.9 ลักษณะเด่นของภาพที่ได้จากวิธี SIFT

3.3 การหาขอบภาพโดยเทคนิคแคนนี่ [24]

Canny Edge Detection คือ อัลกอริธึมการหาเส้นขอบภาพแบบแคนนี่ถูกนำเสนอในบทความเรื่อง “A Computational Approach to Edge Detection” โดย John Canny ในปี 1986 มีเงื่อนไข 3 ประการ คือ 1) การตรวจจับเส้นขอบภาพโดยตรวจจับเส้นขอบให้ผิดพลาดน้อยที่สุด กล่าวคือ เส้นขอบที่ได้ในภาพผลลัพธ์ ขอบของวัตถุที่สำคัญต้องไม่ขาดหายและไม่นำเสนอส่วนที่ไม่ใช่เส้นขอบที่แท้จริง 2) การระบุตำแหน่ง ระยะห่างระหว่างพิกเซลเส้นขอบที่ตรวจจับได้กับพิกเซลเส้นขอบที่แท้จริงต้องหาน้อยที่สุด 3) กรองการตอบสนองแบบหลายตำแหน่ง (Multiple Responses) คือการกรองจากหลายตำแหน่งของขอบของวัตถุให้เหลืออย่างน้อยที่สุด (One Response)

วิธีการตรวจจับเส้นขอบภาพแบบแคนนี่สามารถแบ่งได้เป็น 4 ขั้นตอนหลัก คือ การปรับภาพให้เรียบด้วยตัวกรองเกาส์เซียน (Gaussian Filter) เพื่อกำจัดจุดรบกวนเบื้องต้น จากนั้นคำนวณค่าขนาด (Magnitude) และทิศทาง (Direction) ของค่าการเปลี่ยนแปลงความเข้มแสง (Gradient) จะถูกนำมาคำนวณ ขั้นตอนต่อมาเพื่อให้ได้เส้นขอบภาพที่คมชัดมากขึ้นจึงทำการกำจัดค่าขนาดการเปลี่ยนแปลงความเข้มแสงที่ไม่ใช่ค่าสูงสุดเฉพาะที่ออก (Nonmaxima Suppression) และขั้นตอนสุดท้ายคือการทำฮิสเทอรีซิส (Hysteresis) คือการคัดกรองให้เหลือแต่พิกเซลเส้นขอบที่แท้จริง โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

3.3.1 การปรับภาพให้เรียบด้วยตัวกรองเกาส์เซียน

ขั้นตอนแรกของการตรวจจับเส้นขอบภาพแบบแคนนี่ คือ การนำตัวกรองเกาส์เซียนมากรอง จุตรบกวน (Noise) ออกจากภาพดั้งเดิม (Original Image) หน้ากากตัวกรองเกาส์เซียน ($G_\sigma(x, y)$) จะทำการคอนโวลูชันฟังก์ชันภาพ ($f(x, y)$) ด้วยดังสมการ 3.11

$$g(x, y) = (x, y) * f(x, y) \quad (3.11)$$

โดย

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2+y^2}{2\sigma^2}\right] \quad (3.12)$$

เมื่อ $g(x, y)$ คือ ภาพผลลัพธ์การกรองด้วยตัวกรองเกาส์เซียน

G_σ คือ หน้ากากตัวกรองเกาส์เซียน

σ คือ ตัวควบคุมระดับการกรอง (Spread of the Gaussian)

จะเห็นได้ว่าการคอนโวลูชันระหว่างภาพกับหน้ากากตัวกรองเกาส์เซียน สามารถสร้างตัวดำเนินการได้ง่ายและมีทิศทางที่ถูกต้อง สำหรับหน้ากากตัวกรองที่นิยมใช้มักจะเป็นวงกลมที่มีขนาดเล็ก และเล็กกว่าภาพดั้งเดิมเนื่องจากความทนทานต่อการตรวจจับจุตรบกวนลดน้อยลงหากเลือกหน้ากากตัวกรองที่มีขนาดใหญ่

3.3.2 การคำนวณค่าขนาดและทิศทางของค่าการเปลี่ยนแปลงความเข้มแสง

การคำนวณค่าขนาดและทิศทางของค่าการเปลี่ยนแปลงความเข้มแสงสามารถทำได้โดยนำ ภาพผลลัพธ์ที่ผ่านการกรองแบบเกาส์เซียนจาก 3.3.1 มาทำคอนโวลูชันกับเทมเพลตตัวกรอง หรือนำไปหาอนุพันธ์อันดับหนึ่ง

3.3.3 การกำจัดค่าขนาดการเปลี่ยนแปลงความเข้มแสงที่ไม่ใช่ค่าสูงสุดเฉพาะที่ (Nonmaxima Suppression)

การกำจัดค่าขนาดการเปลี่ยนแปลงความเข้มแสงที่ไม่ใช่ค่าสูงสุดเฉพาะที่สามารถทำได้โดยการกำจัด (Suppression) ค่าอื่นๆที่ไม่ใช่ค่าสูงสุดเฉพาะที่ทิ้งไปและคงแต่ค่าการเปลี่ยนแปลงความเข้มแสงของพิกเซลที่เป็นค่าสูงสุดเฉพาะที่ (Local Maxima) ไว้ ซึ่งจุดประสงค์หลักของขั้นตอนนี้ เพื่อปรับเส้นขอบภาพที่ไม่คมชัด (Blurred Edge) ให้เป็นเส้นขอบภาพที่มีความคมชัด (Sharp Edge) โดยขั้นแรกจะพิจารณาทิศทางการเปลี่ยนแปลงความเข้มแสงของพิกเซลที่สนใจว่าสัมพันธ์กับทิศการเปลี่ยนแปลงความเข้มแสงค่าสูงสุดเฉพาะที่ จากนั้นเปรียบเทียบค่าขนาดการเปลี่ยนแปลงความเข้มแสงของพิกเซลที่สนใจกับพิกเซลเพื่อนบ้าน ถ้าค่าขนาดการเปลี่ยนแปลงความเข้มแสงของพิกเซลเพื่อนบ้านมีค่ามากกว่าให้คงค่าเดิมไว้ เพราะถือว่าค่าดังกล่าวเป็นค่าสูงสุดเฉพาะที่แล้ว แต่ถ้ามีค่าน้อยกว่าให้แทนค่าขนาดของพิกเซลที่สนใจด้วย 0

3.3.4 การทำฮิสทีเรซิส (Hysteresis) เพื่อคัดกรองให้เหลือแต่พิกเซลเส้นขอบที่แท้จริง

เนื่องจากภาพผลลัพธ์ที่ได้หลังการกำจัดค่าขนาดการเปลี่ยนแปลงความเข้มแสงที่ไม่ใช่ค่าสูงสุดเฉพาะที่ ยังคงประกอบด้วยพิกเซลของสัญญาณรบกวน (Noise) หรือบริเวณพื้นผิวที่ขรุขระ (Rough Surface) ซึ่งไม่ใช่เส้นขอบที่แท้จริงปรากฏอยู่ ทำให้เกิดปัญหาเส้นลาย (Streaking) เนื่องจากภาพถูกผ่านการทำค่าเรชโฮลด์ (Thresholding) เพื่อเลือกขอบวัตถุ โดยปัญหา Streaking สามารถแก้ได้ด้วยการทำฮิสทีเรซิสซึ่งมีขั้นตอนคือ กำหนดค่าขีดแบ่งขึ้นมา 2 ค่า คือ ค่าเรชโฮลด์แบ่งต่ำ (Low Threshold) และเรชโฮลด์แบ่งสูง (High Threshold) จากนั้นเปรียบเทียบขนาดการเปลี่ยนแปลงความเข้มแสงของแต่ละพิกเซลกับค่าขีดแบ่งทั้งสอง ถ้าพิกเซลใดมีค่าขนาดการเปลี่ยนแปลงความเข้มแสงมากกว่าเรชโฮลด์แบ่งสูงจะถูกกำหนดให้เป็นพิกเซลเส้นขอบที่แท้จริง เนื่องจากจากจุดภาพที่มีค่าเกรเดียนสูงซึ่งจะเป็นบริเวณที่เป็นเส้นขอบที่แท้จริง (ถูกปรับค่าเป็น 1) แต่ถ้าน้อยกว่าเรชโฮลด์แบ่งต่ำจะถูกกำหนดให้เป็นพิกเซลสัญญาณรบกวน (ถูกปรับค่าเป็น 0) เรียกอัลกอริธึมการกำหนดค่าขีดแบ่งแบบคู่ (Double Thresholding) เพื่อวิเคราะห์และจำแนกพิกเซลเส้นขอบแท้จริงออกจากพิกเซลที่เป็นสัญญาณรบกวน

3.4 การประมวลผลรูปร่างและโครงสร้างของวัตถุในภาพ (Morphological Image Processing) [25]

การประมวลผลรูปร่างและโครงสร้างของภาพเป็นการประมวลผลโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของวัตถุในภาพ โอเปอเรชันพื้นฐานโดยทั่วไป ได้แก่ การขยายองค์ประกอบ (Dilation) คือการขยายวัตถุในภาพโดยมีสัดส่วนเท่ากันทั่วทั้งวัตถุ (Uniform) การกร่อนองค์ประกอบ (Erosion) คือการย่อวัตถุในภาพ และการเปลี่ยนองค์ประกอบให้บางลง (Thinning) ส่วนการเปลี่ยนองค์ประกอบเป็นโครงกระดูก (Skeleton) เป็นการหาโครงสร้างหลักของวัตถุ นอกจากนี้โอเปอเรชันพื้นฐานที่ได้กล่าวในข้างต้นแล้ว ยังมีโอเปอเรชันอื่นๆอีก ได้แก่ การประมวลผลภาพกับรูปร่างและโครงสร้างของวัตถุในภาพแบบเปิด (Opening) และการประมวลผลภาพกับรูปร่างและโครงสร้างของวัตถุในภาพแบบปิด (Closing) เป็นต้น ซึ่งจะกล่าวต่อจากนี้

3.4.1 เทคนิคฮิตและมิส (Hit and Miss)

โอเปอเรชันพื้นฐานสำหรับการกระทำกับรูปร่างหรือโครงสร้างของภาพไม่ว่าจะเป็น การย่อหรือการขยายภาพ จำเป็นที่จะต้องมีการนำเอาเทคนิคการทำฮิตและมิสมาใช้ โดยจะมีเทมเพลตโครงรูปร่าง (Structuring element) ซึ่งทำหน้าที่เป็นเมตริก (Template) ที่เป็นไบนารีอิมเมจ (0 กับ 1) ซึ่งมีหลายลักษณะ เช่น Disk, Circle, Diamond, Square, Line, Pair เป็นต้น โดยเมื่อนำ Structuring element ไปวางยังตำแหน่งพิกเซลที่เราสนใจ และจะต้องมีการอ้างอิงกับพิกเซลรอบข้างที่เกี่ยวข้องตามลักษณะของเมตริก โดยการเปรียบเทียบจะทำกับทุกวัตถุที่สนใจในภาพตั้งแต่ต้นภาพจนถึงท้ายภาพ ถ้าข้อมูลตำแหน่งพิกเซลที่สนใจของวัตถุในภาพมีลักษณะเหมือนกับเมตริกดังกล่าว เอาพิกเซลที่ได้จะขึ้นอยู่กับพิกเซลที่เป็นศูนย์กลางของเมตริกซึ่งจะถูกกำหนดให้เป็นค่าตาม

ต้องการ (1 หรือ 0) แต่ถ้าข้อมูลในเมตริกไม่เหมือนกับข้อมูลภาพข้อมูลเอาต์พุตที่ได้จะมีค่าตรงกันข้าม

3.4.2 การขยายองค์ประกอบ (Dilation)

การขยายองค์ประกอบในที่นี้จะพิจารณาสำหรับข้อมูลภาพที่เป็นแบบไบนารีโดยการใช้เทคนิคการทำอิทและมัส ตามที่ได้กล่าวไว้ในหัวข้อ 3.4.1 การขยายองค์ประกอบสามารถทำได้โดยกำหนดเทมเพลต (ซึ่งสามารถสร้างได้จาก * และ 1 โดยมีจุดเริ่มต้นที่กำหนดโดยวงกลม) และนำเทมเพลตนี้สแกนไปบนข้อมูลภาพตามลำดับตลอดทั้งภาพ หากจุดในเทมเพลตตรงกับตำแหน่งข้อมูลภาพที่มีพิกเซลมีค่าเท่ากับ 1 นั้นจะทำการยูเนียนเทมเพลตนี้เข้ากับข้อมูลภาพดังกล่าวอย่างรูปที่ 3.10

ตัวอย่าง

ข้อมูลภาพ										ลักษณะเทมเพลต		
*	*	*	*	*	*	1	*	*	1	*		
*	*	*	*	*	*	1	*	*	*	1		
*	*	*	*	*	1	1	*	1	1	*		
*	*	*	*	1	1	1	1	1	1	1	1	*
*	*	*	*	1	1	1	1	1	*	1	1	*
*	*	*	*	1	1	1	1	1	1	1		
*	*	*	*	1	1	1	1	1	1	1		

ข้อมูลแถวแรกของภาพเป็นดังนี้

* * * * * 1 * * 1 *

เมื่อทำการยูเนียนกับเทมเพลต ณ ตำแหน่งข้อมูลภาพที่พิกเซลเท่ากับ 1 ในแถวแรก

* * * * * 1 * * * *
* * * * * 1 1 * 1 *

และเมื่อยูเนียนกับเทมเพลตเข้ากับพิกเซลที่มีค่าเท่ากับ 1 ณ ตำแหน่งพิกเซลที่สองในแถวแรก

* * * * * 1 * * * *
* * * * * 1 1 * 1 1

และเมื่อทำการยูเนียนทั้งภาพจะได้ภาพสุดท้ายดังนี้

```

* * * * * 1 * * 1 * *
* * * * * 1 1 * 1 1 *
* * * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1

```

รูปที่ 3.10 ตัวอย่างการขยายองค์ประกอบ

3.4.3 การกร่อนองค์ประกอบ (Erosion)

การกร่อนองค์ประกอบเป็นลักษณะของการลบข้อมูลภาพบริเวณขอบของภาพ การกร่อนองค์ประกอบภาพมีลักษณะคล้ายกับการขยายองค์ประกอบโดยการสร้างเทมเพลตขึ้นมาแล้วนำเทมเพลตไปสแกนตามข้อมูลภาพ เทมเพลตจะเลื่อนไปแต่ละตำแหน่งบนภาพ จากนั้นจะมีการเปรียบเทียบกับข้อมูลภาพ ถ้าข้อมูลภาพมีค่าเหมือนกับเทมเพลตจะทำการกำหนดค่าข้อมูลภาพในตำแหน่งที่ตรงกับจุดเริ่มต้น (Origin) ของเทมเพลตให้มีค่าเท่ากับ 1 ดังแสดงในรูปที่ 3.11

ตัวอย่าง

ข้อมูลภาพ	ลักษณะเทมเพลต
* * * * * 1 * * 1 *	
* * * * * 1 * * * 1	
* * * * * 1 1 * 1 1 *	
* * * * 1 1 1 1 1 1 1	1 *
* * * * 1 1 1 1 1 * 1	1 1
* * * * 1 1 1 1 1 1 1	
* * * * 1 1 1 1 1 1 1	

ผลที่ได้จะมีเพียง 3 ตำแหน่งเท่านั้นที่มีค่าเหมือนกับเทมเพลต

```

* * * * * * * * * *
* * * * * * * * * *
* * * * * * * * * *
* * * * * * 1 * * 1 *
* * * * * * * * * *
* * * * * * * * 1 * *
* * * * * * * * * *
    
```

รูปที่ 3.11 ตัวอย่างการกร่อนองค์ประกอบ

ผลที่ได้ตามรูปที่ 3.11 ข้อมูลภาพที่ผ่านการทำโอเพอเรชันกับเทมเพลตแล้วพบว่ามีข้อมูลของภาพเพียง 3 ตำแหน่งเท่านั้นที่เหมือนกับเทมเพลต จึงมีตำแหน่งที่ถูกแทนด้วย 1 เพียงสามตำแหน่ง

1 1

ถ้ามีการเปลี่ยนเทมเพลตเป็น 1 1 ผลที่ได้มีลักษณะดังรูปที่ 3.12 คือ

```

* * * * * * * * * *
* * * * * * * * * *
* * * * * * * * * *
* * * * * 1 * * 1 * *
* * * * 1 1 1 1 1 * *
* * * * 1 1 1 1 * * *
* * * * 1 1 1 1 1 1 *
    
```

รูปที่ 3.12 ตัวอย่างการกร่อนองค์ประกอบที่ได้จากเทมเพลตใหม่

ผลที่ได้ตามรูปที่ 3.12 จะเห็นว่าจะเป็นการย่อขนาดของภาพแต่สามารถย่อขนาดได้น้อยกว่า

1 *

การใช้เทมเพลต 1 1 ซึ่งได้ผลเป็นที่น่ายอมรับมากกว่า ดังนั้นในการเลือกเทมเพลตเป็นสิ่งสำคัญอย่างหนึ่งในการย่อและขยายภาพ

3.4.4 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบเปิด (Opening)

กำหนดให้ $OPEN(I, T)$ เป็นการทำ Opening ของภาพ I โดยใช้เทมเพลต T ซึ่งมีลักษณะดังสมการ 3.13 ต่อไปนี้

$$OPEN(I, T) = D(E(I)) \tag{3.13}$$

จากสมการจะเห็นว่า การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบเปิดคือการนำข้อมูลภาพ I ผ่านการทำขยายภาพ (Erosion) แล้วตามด้วยการย่อภาพ (Dilation) โดยใช้เทมเพลตชุดเดียวกันคือ T

ตัวอย่าง

ภาพเริ่มต้น	ลักษณะเทมเพลต	ภาพผลลัพธ์
0 0 0 0 0 0 0 0		0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 4		0 0 0 0 0 0 1 4
0 0 0 0 1 2 4 5		0 0 0 0 1 2 4 4
0 0 1 1 2 5 5 4	1	0 0 0 0 1 2 4 4
0 0 0 0 0 1 2 3	1	0 0 0 0 0 1 2 3
0 0 0 0 0 0 1 3		0 0 0 0 0 0 1 3
0 0 0 0 0 1 2 3		0 0 0 0 0 1 2 3
0 0 0 0 1 2 2 3		0 0 0 0 0 1 2 3

รูปที่ 3.13 ผลลัพธ์จากการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบเปิด

3.4.5 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบปิด (Closing)

กำหนดให้ $CLOSE(I, T)$ เป็นการกระทำการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบปิดของภาพ I โดยใช้ Template T ซึ่งมีลักษณะดังสมการที่ 3.14

$$CLOSE(I, T) = E(D(I)) \quad (3.14)$$

จากสมการจะเห็นว่า การทำโอเพอเรชั่น CLOSE คือการนำข้อมูลภาพ I ผ่านการทำการย่อภาพ (Dilation) แล้วตามด้วยการขยายภาพ (Erosion) โดยใช้เทมเพลตชุดเดียวกันคือ T

ตัวอย่าง

ภาพต้นฉบับ	ลักษณะเทมเพลต	ภาพผลลัพธ์
	1	
	1	

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	4	0	0	0	0	0	0
0	0	0	0	0	1	2	4	5	0	0	0	0	1	2
0	0	1	1	2	5	5	4	0	0	1	1	2	5	5
0	0	0	0	0	0	1	2	3	0	0	0	0	0	1
0	0	0	0	0	0	0	1	3	0	0	0	0	0	1
0	0	0	0	0	0	1	2	3	0	0	0	0	0	1
0	0	0	0	1	2	2	3	0	0	0	0	1	2	2

รูปที่ 3.14 ผลลัพธ์จากการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพแบบปิด

3.4.6 การเปลี่ยนองค์ประกอบให้บางลง (Thinning)

วิธีการทำองค์ประกอบให้บางลงสามารถทำได้โดยอาศัยกฎ 2 ข้อคือ P1 และ P2 ซึ่งมี 2 ขั้นตอน คือขั้นแรกจะใช้ กฎ P1 โดยการนำเทมเพลตขนาด 3×3 สแกนไปตามข้อมูลภาพและทำการพิจารณาพิกเซลบริเวณขอบภาพว่าสามารถลบได้หรือไม่ ถ้าลบได้ให้หมายเหตุไว้แต่ยังไม่ต้องลบ หลังจากทีสแกนทั่วทั้งภาพแล้วจึงทำการลบข้อมูลภาพดังที่ได้หมายเหตุไว้ ขั้นที่สองใช้กฎ P2 และดำเนินการเหมือนการใช้กฎข้อที่ P1 เมื่อทำการลบข้อมูลภาพที่มีไว้ในหมายเหตุแล้ว จากนั้นทำซ้ำต่อไปเรื่อยๆจนกระทั่งไม่สามารถลบข้อมูลภาพออกได้อีก

ลักษณะของข้อมูลที่นำมาพิจารณาจะมีขนาดเท่ากับ 3×3 ดังรูปที่ 3.15 โดยที่พิกเซลปัจจุบันคือพิกเซลตรงกลางหรือ p_0

p_8	p_1	p_2
p_7	p_0	p_3
p_6	p_5	p_4

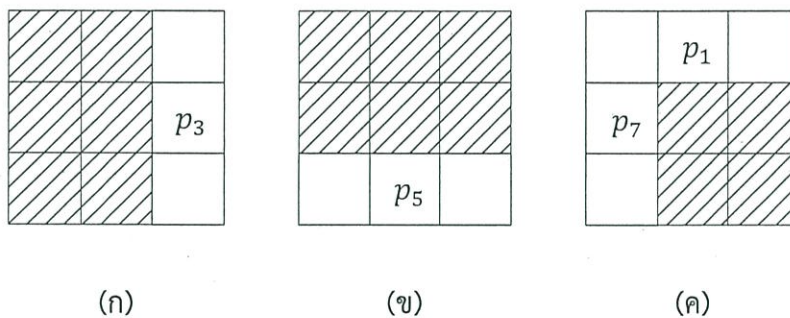
รูปที่ 3.15 รูปแบบเทมเพลตที่นำมาพิจารณา

กำหนดให้ $N(p_0) = \sum_{i=1}^8 p_i$ เป็นจำนวนของพิกเซลรอบ p_0 เมื่อ $p_0 = 0, 1$ และ $i = 1, 2, \dots, 8$ $T(p_0)$ แสดงถึงการเปลี่ยนแปลงของข้อมูลจาก 0 ไปเป็น 1 เมื่อพิจารณาข้อมูลใน $p_1, p_2, p_3, \dots, p_8, p_1$ ตามลำดับ สำหรับกฎ P1 และ P2 จะมีตรรกะดังนี้คือ

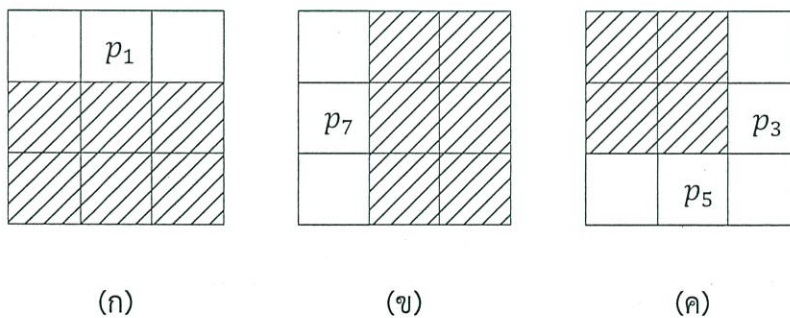
$$P1 : (2 \leq N(p_0) \leq 6) \&\& (T(p_0) = 1) \&\& (p_1, p_3, p_5 = 0) \&\& (p_3, p_5, p_7 = 0)$$

$$P2 : (2 \leq N(p_0) \leq 6) \&\& (T(p_0) = 1) \&\& (p_1, p_3, p_7 = 0) \&\& (p_1, p_5, p_7 = 0)$$

สำหรับเงื่อนไขแรกของ P1 และ P2 คือพิกเซลตรงกลาง (p_0) สามารถลบได้ก็ต่อเมื่อค่าของ p_1 ถึง p_8 มีค่าเท่ากับ 1 (เป็นสีดำ) ซึ่งมีมากกว่า 1 ตำแหน่งแต่ต้องไม่มากกว่า 6 ตำแหน่ง สำหรับเงื่อนไขที่ 2 กำหนดให้ $T(p_0) = 1$ กล่าวคือต้องมีการเปลี่ยนแปลงของข้อมูลจาก 0 ไปเป็น 1 ของข้อมูล $p_1, p_2, p_3, \dots, p_8, p_1$ เพียงครั้งเดียวเท่านั้น และในเงื่อนไขที่ 3 และ 4 คือ ($p_1 \cdot p_2 \cdot p_3 = 0$) และ ($p_3 \cdot p_5 \cdot p_7 = 0$) สามารถใช้ได้สำหรับเงื่อนไขที่ $p_3 = 0$ หรือ $p_5 = 0$ หรือ ถ้า $p_1 = 0$ และ $p_7 = 0$ ตามตัวอย่างดังกล่าวนี้แสดงไว้ดังรูปที่ 3.16 ซึ่งจะเห็นว่าเมื่อ $p_3 = 0$ จะเป็นลักษณะของขอบด้านขวา เมื่อ $p_5 = 0$ จะเป็นขอบด้านล่าง และเมื่อ $p_1 = 0$ และ $p_7 = 0$ จะเป็นมุมบนซ้าย สำหรับใน P2 นั้นเราสามารถลบพิกเซล p_0 ได้ก็ต่อเมื่อ $p_1 = 0$ หรือ $p_7 = 0$ หรือ (p_3 และ $p_5 = 0$) ดังแสดงในรูปที่ 3.17 ซึ่งจะเห็นว่าเมื่อ $p_1 = 0$ จะเป็นขอบบน $p_7 = 0$ จะเป็นขอบด้านซ้าย และเมื่อ ($p_3 = 0$ และ $p_5 = 0$) จะเป็นมุมล่างขวา



รูปที่ 3.16 เทมเพลตตามตัวอย่าง (ก) $p_3 = 0$ (ข) $p_5 = 0$ (ค) $p_1 = 0$ และ $p_7 = 0$



รูปที่ 3.17 เทมเพลตตามตัวอย่าง (ก) $p_1 = 0$ (ข) $p_7 = 0$ (ค) $p_3 = 0$ และ $p_5 = 0$

3.5 การทำ Labeling and Measuring Objects

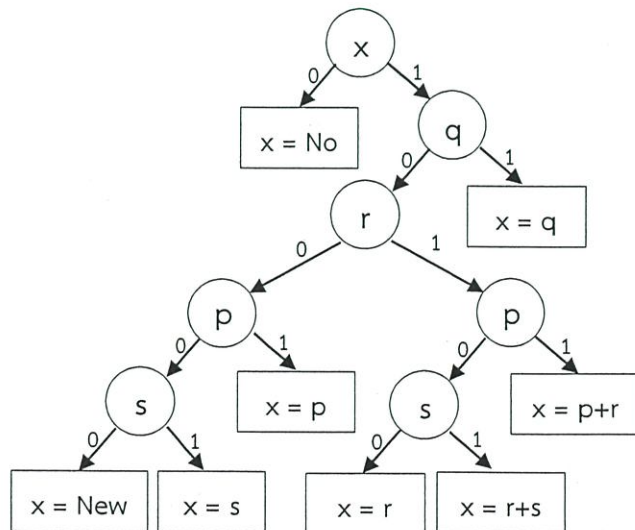
การทำลาเบลลิ่ง (Labeling) คือวิธีการค้นหาพิกเซลในวัตถุที่ติดกันหรือเป็นกลุ่มเดียวกัน โดยวิธีการค้นหาพิกเซลของวัตถุที่ติดกันในภาพ 2 มิติ นิยมใช้วิธี 8-connectivity ดังแสดงในรูปที่ 3.18 ซึ่งมีการนำเสนอครั้งแรกโดย Rosenfeld และ Pfaltz. ในปี 1966 [26] ที่เรียกว่า Two-scans labeling มีการทำงานที่สามารถแบ่งออกเป็น 3 ขั้นตอนดังนี้

1. First image scan: เป็นการใช้น้ำกาก (mask) มาสแกนแล้วทำการค้นหาพิกเซลภาพที่เป็นวัตถุ (Foreground) จากนั้นจะกำหนดค่า Label แบบชั่วคราว (Provisional label) ให้กับพิกเซลภาพที่เป็นวัตถุ และทำการบันทึกค่าของ Provisional Label ลงใน Equivalences Table
2. Equivalences Resolution: เป็นการนำ Provisional Label ใน Equivalences Table มาทำการจัดกลุ่ม และเลือกค่าน้อยที่สุดของแต่ละกลุ่มเพื่อนำมาใช้เป็นตัวแทนของกลุ่ม
3. Second image scan: ทำการสแกนในรอบที่สอง เพื่อนำตัวแทนของแต่ละกลุ่มในตาราง Equivalences Table มาทำการ Label ลงที่ภาพคำตอบ

$p_{(x-1,y-1)}$	$p_{(x,y-1)}$	$p_{(x+1,y-1)}$
$p_{(x-1,y)}$	$p_{(x,y)}$	$p_{(x+1,y)}$
$p_{(x-1,y+1)}$	$p_{(x,y+1)}$	$p_{(x+1,y+1)}$

รูปที่ 3.18 8-connectivity

ในปี 2005 Wu และคณะ [27] ได้เสนอวิธีการต้นไม้การตัดสินใจ (Decision Tree) ในการค้นหาพิกเซลที่ติดกันในขั้นตอน First image scan ซึ่งทำให้ลดจำนวนครั้งของการตรวจสอบพิกเซลภาพในรอบแรกของ Two-scans labeling ดังแสดงในรูปที่ 3.19



รูปที่ 3.19 ต้นไม้ตัดสินใจที่ใช้ตรวจสอบหา 8-connectivity

จากรูปที่ 3.19 มีการทำงานอยู่ 4 ประเภทดังนี้

1. No action: พิกเซล “x” เป็น Background ไม่มีการกำหนดค่าให้กับพิกเซล “x”
2. New label: พิกเซล “x” เป็น Foreground และพิกเซลข้างเคียงเป็น Background ให้กำหนดค่า Provisional Label ค่าใหม่ให้กับพิกเซล “x”
3. Assign action: พิกเซล “x” เป็น Foreground และพิกเซลข้างเคียงทุกพิกเซลเป็น Foreground ที่มี Provisional Label ค่าเดียวกัน
4. Merge action: พิกเซล “x” เป็น Foreground และพิกเซลข้างเคียงมากกว่า 1 พิกเซลเป็น Foreground และมี Provisional Label ที่ไม่ซ้ำกัน

3.6 การคำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (Histogram of accumulated Major Axis Length of ellipses; HoARE)

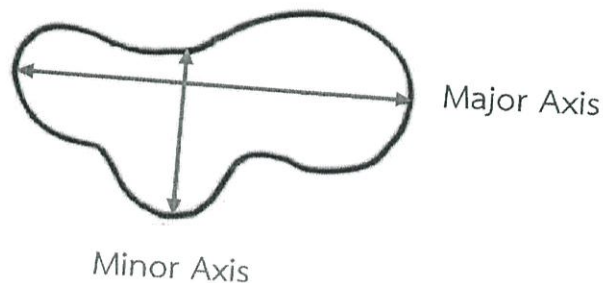
การคำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน คือการคำนวณทิศทางของวงรีที่ล้อมรอบกลุ่มตัวอักษรภายในชิ้นส่วนที่พิจารณา โดยวงรีสามารถหาได้จากการหาจุดศูนย์กลาง Centroid คือ จุดศูนย์กลางของมวลของ Region หรือกลุ่มตัวอักษรใดๆ ซึ่งวงรีประกอบด้วย แกนหลักของวงรี (Major Axis) คือ จุดปลายสองจุด (Ending Point) ที่ยาวที่สุดใน Region และแกนรองของวงรี (Minor Axis) คือ จุดปลายสองจุดที่ยาวที่สุดและตั้งฉากกับแกนหลักของวงรีดังรูปที่ 3.20 โดยระยะและมุมแกนหลักของวงรีสามารถคำนวณได้จากสมการที่ 3.15 และ 3.16 ตามลำดับ

$$\text{Major Axis Length} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.15)$$

$$\text{Major Axis Angel} = \tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \quad (3.16)$$

และระยะแกนรองของวงรีสามารถหาได้จากสมการที่ 3.17

$$\text{Minor Axis Length} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.17)$$



รูปที่ 3.20 แกนหลักและแกนรองของวงรี

3.7 การแปลงเรขาคณิตแบบเชิงเส้น (Geometrical Linear Transformation) [28]

การแปลงทางเรขาคณิตแบบเชิงเส้นเป็นขบวนการพื้นฐานในการประมวลผลภาพแบบดิจิทัลที่เกี่ยวข้องกับการสเกล (Scaling) การเคลื่อนย้ายภาพ (Translation) การหมุนภาพ (Rotation) และการเฉือนภาพ (Shear) การแปลงเรขาคณิตจะทำในระบบพิกัดแบบ Cartesian นั้นคือจุดกำเนิด (0,0) อยู่ที่มุมล่างซ้ายของภาพ แต่สำหรับภาพดิจิทัลจุดกำเนิด (1,1) อยู่ที่มุมขวาบนของภาพ

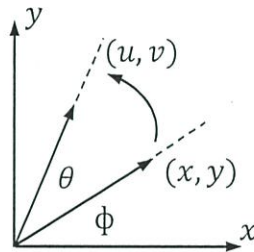
3.7.1 การหมุนภาพ (Rotation)

การหมุนภาพรอบจุดกำเนิดสามารถทำได้โดยการคำนวณดังนี้

พิจารณการหมุนของจุด (x, y) รอบจุดกำเนิดไปที่ (u, v) ด้วยมุม θ รูปที่ 3.21 ซึ่งสามารถแทนจุด (x, y) และจุด (u, v) ในพิกัดโพลาร์ได้ดังสมการที่ 3.18 และ 3.19

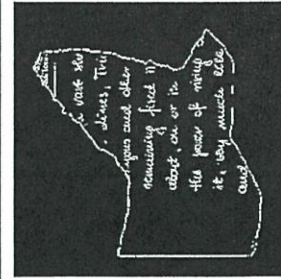
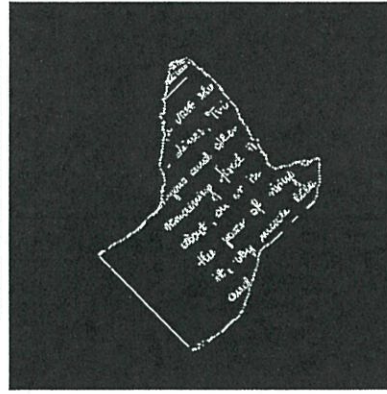
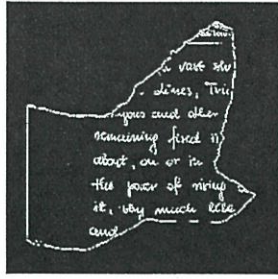
$$x = r \cos \phi \qquad y = r \sin \phi \qquad (3.18)$$

$$u = r \cos(\theta + \phi) \qquad v = r \sin(\theta + \phi) \qquad (3.19)$$



รูปที่ 3.21 การหมุน (x, y) รอบจุดกำเนิดไปที่จุด (u, v) ด้วยมุม θ

โดยที่แทนมุมของการหมุนในทิศทางตามเข็มนาฬิกาเมื่อเทียบกับแนวนอน (Horizontal Axis) สมการการหมุนภาพ (3.20) แทนการหมุนภาพรอบจุดกำเนิดในระบบพิกัด Cartesian การหมุนภาพรอบจุดใดๆ สามารถทำได้โดยการเลื่อนจุดหมุนไปที่จุดกำเนิดแล้วจึงทำการหมุนภาพสุดท้ายเลื่อนจุดหมุนกลับไปดั้งเดิมตัวอย่างการหมุนภาพในรูปที่ 3.22



(ก) ภาพต้นฉบับ

(ข) การหมุนภาพ 45 องศา

(ค) การหมุนภาพ 90 องศา

รูปที่ 3.22 ตัวอย่างการหมุนภาพ

$$u = x \cos \theta - y \sin \theta \quad (3.20)$$

$$v = x \sin \theta + y \cos \theta \quad (3.21)$$

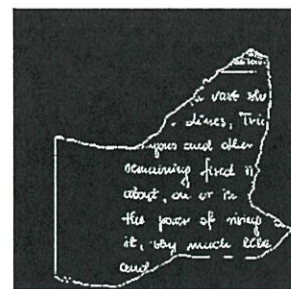
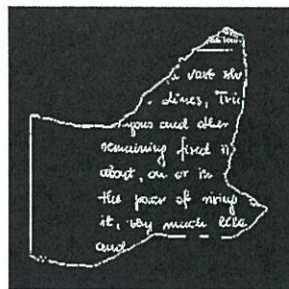
3.7.2 การเคลื่อนย้ายภาพ (Translation)

การเคลื่อนย้ายภาพ $f(x, y)$ เทียบกับจุดกำเนิด Cartesian ไปเป็นภาพ $g(u, v)$ เกี่ยวข้องกับการคำนวณตำแหน่งออฟเซ็ทสัมพัทธ์ (Relative Offset Address) ของภาพทั้งสอง ซึ่งสามารถแสดงความสัมพันธ์ระหว่างตำแหน่งที่เคลื่อนย้ายได้ดังสมการที่ 3.22 และ 3.23 และแสดงตัวอย่างการย้ายภาพในรูปที่ 3.23

$$u = x + t_x \quad (3.22)$$

$$v = y + t_y \quad (3.23)$$

โดยที่ t_x และ t_y แทนระยะของการเคลื่อนย้ายตำแหน่ง



(ก) ก่อนเคลื่อนย้ายภาพ

(ข) หลังเคลื่อนย้ายภาพ

รูปที่ 3.23 ตัวอย่างการย้ายภาพ $t_x = 20, t_y = 20$

3.8 ฮัฟทรานสฟอร์ม [29]

ฮัฟทรานสฟอร์มเป็นวิธีการในการตรวจหาเส้นตรงแต่ละเส้นที่แฝงอยู่ในภาพ โดยภาพตั้งต้นที่จะนำมาตรวจสอบเพื่อหาเส้นตรงด้วยวิธีฮัฟทรานสฟอร์มจำเป็นต้องเป็นภาพที่ผ่านเทคนิคการหาขอบมาก่อน สำหรับแนวคิดเบื้องต้นของฮัฟทรานสฟอร์มคือ การเชื่อมจุด (x, y) ที่อยู่ในภาพ โดยการตรวจสอบว่าจุดดังกล่าวอยู่บนเส้นตรงใดๆ ซึ่งสามารถมีเส้นตรงผ่านได้จำนวนไม่จำกัดและเส้นตรงแต่ละเส้นจะมีสมการเส้นตรงประจำตัวดังสมการที่ 3.24

$$y = mx + b \quad (3.24)$$

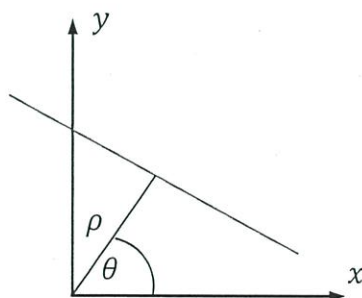
เมื่อ m คือ ความชันของเส้นตรง และ b คือจุดตัดที่เส้นตรงทำกับแกน y

จุดใดๆที่อยู่บนเส้นตรงเดียวกันจะมีค่า m และ b เท่ากัน โดยสมการที่ 3.25 สามารถแปลงสมการใหม่ได้เป็น

$$b = y - mx \quad (3.25)$$

หากจุดใดๆ มีค่า m และ b เท่ากัน แสดงว่าจุดใดๆนั้นอยู่บนเส้นตรงเดียวกัน พิกัด (m, b) จะถูกเก็บความถี่ของจุดในช่องอาร์เรย์เดียวกัน ดังนั้นในแต่ละช่องของ (m, b) จะเก็บจำนวนจุดที่อยู่บนเส้นตรงเดียวกันคือ จุดที่มี ค่า m และ b เท่ากัน ซึ่งอาร์เรย์ (m, b) นี้เรียกว่า ฮัฟสเปซ (Hough Space) ดังนั้นฮัฟสเปซ คือตารางแสดงอาร์เรย์ที่ใช้เก็บจำนวนจุดที่อยู่บนแต่ละเส้นตรงในภาพ

จากสมการเส้นตรงจะสังเกตได้ว่ายังคงมีข้อจำกัดในกรณีที่เส้นตรงทำมุม 90 องศา ซึ่งจะทำให้มีค่าความชันเป็นอนันต์จึงยังไม่สามารถนำมาใช้งานได้จริง ดังนั้นจึงต้องมีวิธีการเก็บค่าใหม่ที่สามารถแก้ปัญหากรณีนี้ คือการคำนวณพิกัดเชิงขั้วดังแสดงในรูปที่ 3.24



รูปที่ 3.24 ค่า ρ และ θ ของเส้นตรงในพิกัดเชิงขั้ว

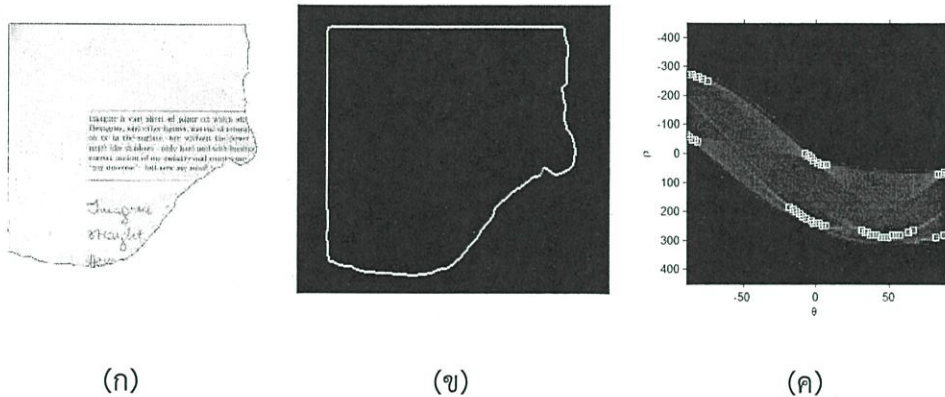
สมการที่ 3.26 แสดงสมการพิกัดเชิงขั้วดังนี้

$$\rho = x \cos \theta + y \sin \theta \quad (3.26)$$

โดย ρ คือ ความยาวของเส้นที่ลากจากจุด (x, y) ไปตั้งฉากกับเส้นตรง

θ คือ มุมระหว่างเส้นที่ลากจากจุด (x, y) ไปตั้งฉากกับเส้นตรงกับแกน x

ฮัฟเฟิลสำหรับพิกัดเชิงขั้วนี้คือ (ρ, θ) ซึ่งทำให้สามารถเก็บข้อมูลได้ 360 องศา โดยจะพิจารณาแต่ละพิกเซลของภาพขาวดำ (Binary Image) ที่ผ่านการหาขอบมาแล้ว ตำแหน่งพิกเซลที่มีค่าเป็น 1 คือพิกเซลของขอบซึ่งจะเป็นองค์ประกอบของเส้นใดเส้นหนึ่งในภาพ โดย (x, y) คือจุดแต่ละพิกเซลในภาพที่พิจารณา ในงานวิจัยนี้ใช้ θ เป็นมุมระหว่าง -90 ถึง 90 องศา แบ่งเป็นช่วงๆ ละ 1 องศา และกำหนดค่า ρ เป็น 8 เมื่อแทนค่าในสมการที่ 3.25 แล้ว ในแต่ละอาเรย์ของฮัฟเฟิลประกอบไปด้วยจำนวนพิกเซลในแต่ละ (ρ, θ) ซึ่งพิกเซลที่อยู่บนเส้นตรงเดียวกันจะอยู่ในตำแหน่ง (ρ, θ) เดียวกัน เมื่อนำฮัฟเฟิลมาแสดงเป็นภาพจะเห็นว่าตำแหน่งใดที่มีความถี่สูงตำแหน่งนั้นจะสว่างมาก ดังรูปที่ 3.25 (ค)



รูปที่ 3.25 การทำฮัฟเฟิลทรานสฟอร์ม

(ก) ภาพตั้งต้น

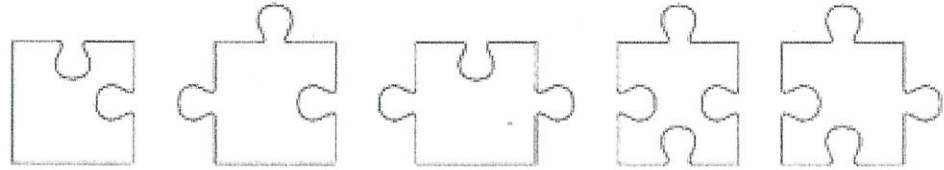
(ข) ภาพขอบของภาพ (ก)

(ค) ภาพฮัฟเฟิลที่ได้จากการทำฮัฟเฟิลทรานสฟอร์มของภาพ (ข)

3.9 การวิเคราะห์มุม

มุม (angle) เกิดจากเส้น 2 เส้นลากเชื่อมกันที่จุดๆ หนึ่ง และหน่วยในการวัดมุมอาจมีหน่วยเป็นองศาซึ่งเขียนในสัญลักษณ์ " $^{\circ}$ " หรือในหน่วยเรเดียน มุมมีหลายขนาดและมีชื่อเรียกแตกต่างกันไป ได้แก่ มุมขนาด 90° (หรือ $\pi/2$ เรเดียน หรือเรียกว่า มุมฉาก (right angle) มุมที่เล็กกว่ามุมฉาก (น้อยกว่า 90°) เรียกว่า มุมแหลม (acute angle) มุมที่ใหญ่กว่ามุมฉากแต่เล็กกว่าสองมุมฉาก (ระหว่าง 90° กับ 180°) เรียกว่า มุมป้าน (obtuse angle) มุมที่มีขนาดเท่ากับสองมุมฉาก (180°) เรียกว่า มุมตรง (straight angle) และมุมที่ใหญ่กว่ามุมตรงแต่น้อยกว่ารูปร่างกลมเต็มวง (ระหว่าง 180° กับ 360°) เรียกว่า มุมกลับ (reflex angle)

สำหรับการวิเคราะห์มุมเพื่อประกอบปริศนาชิ้นส่วนจิ๊กซอว์ ส่วนใหญ่ชิ้นส่วนที่อยู่ในชุดเดียวกันจะมีมุมโค้งเข้าไปภายในชิ้นส่วนและมุมโค้งออกมาภายนอกชิ้นส่วนที่มีขนาดเท่ากันหมด จึงทำให้ชิ้นส่วนมีมุมที่เท่ากันหลายชิ้น มุมโค้งเข้าจะพอดีกับมุมโค้งออกดังรูปที่ 3.26 แต่สำหรับชิ้นส่วนกระดาดชิ้นนั้นมุมจะมีขนาดที่แตกต่างกัน โดยแต่ละมุมของชิ้นส่วนสามารถหาได้จากสมการที่ 3.27



รูปที่ 3.26 ลักษณะมุมโค้งเข้าและโค้งออกของจิ๊กซอว์

$$\cos \theta = \frac{\overline{A_1} \cdot \overline{A_2}}{\|A_1\| \|A_2\|} \quad (3.27)$$

เมื่อเส้นประกอบมุมมีเวกเตอร์แสดงทิศทางเป็น $\overline{A_1}$ และ $\overline{A_2}$ แต่สมการนี้สามารถหาได้เฉพาะมุมตั้งแต่ 0 องศาถึง 180 องศาเท่านั้นจึงไม่สามารถใช้สมการนี้ในการหามุมของชิ้นส่วนได้ เนื่องจากมุมของชิ้นส่วนมีขนาดระหว่าง 0 องศาถึง 360 องศา ดังนั้นจึงต้องใช้สมการที่ 3.28

$$angle = atan2(det, dot) \quad (3.28)$$

เมื่อ det คือ ดีเทอร์มิแนนต์ (determinant) ของเส้นตรงทั้งสองเส้น

dot คือ ผลคูณจุด (dot product)

โดยค่า det และ dot สามารถหาได้จากสมการ 3.29 และ 3.30 ตามลำดับ

$$det = veca(x) * vecb(x) + veca(y) * vecb(y) \quad (3.29)$$

$$dot = veca(x) * vecb(y) - veca(y) * vecb(x) \quad (3.30)$$

เมื่อ $veca$ คือ ค่าเวกเตอร์เส้นตรงเส้นที่หนึ่ง

$vecb$ คือ ค่าเวกเตอร์เส้นตรงเส้นที่สอง

งานวิจัยที่นำเสนอ

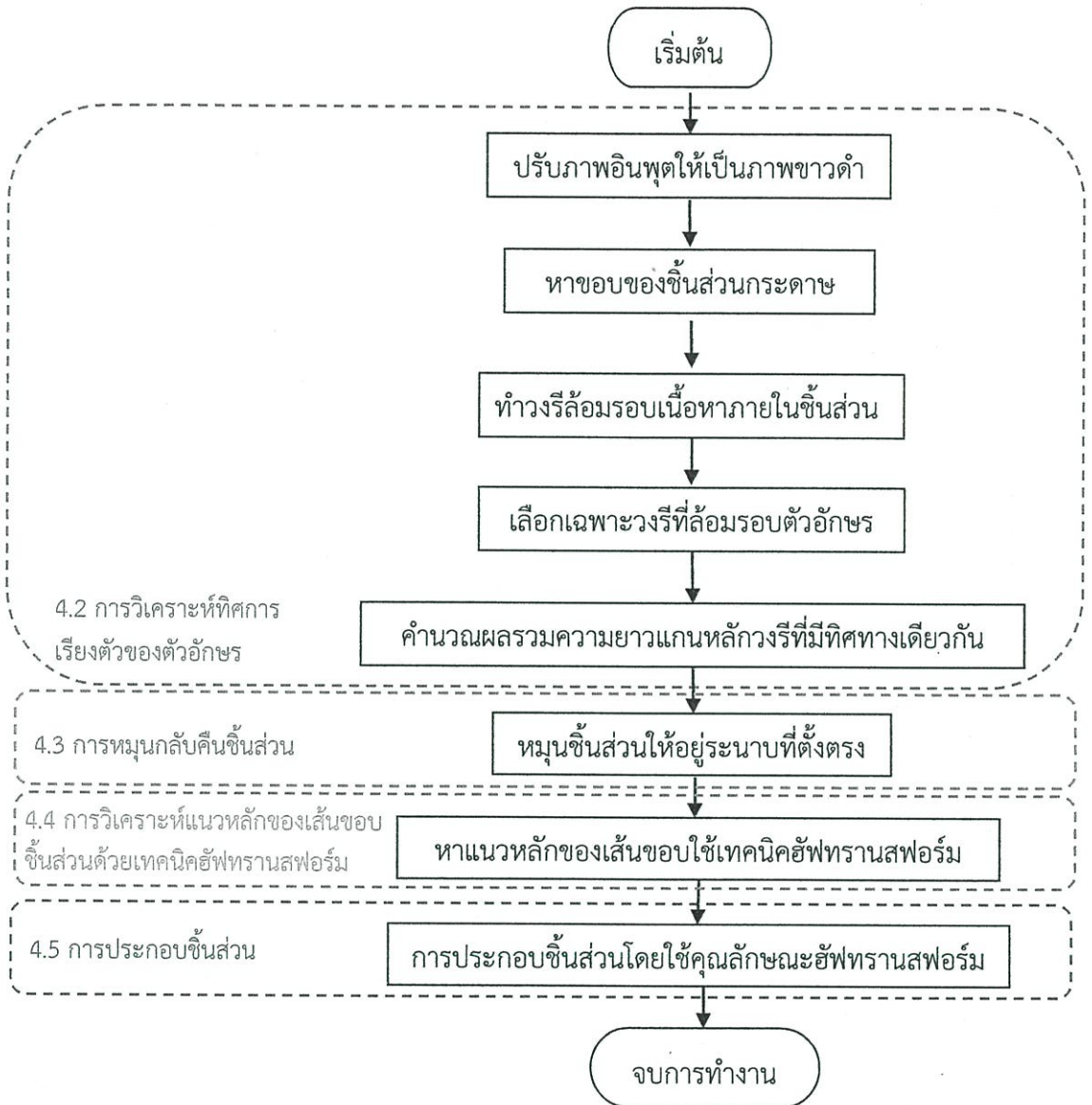
บทนี้จะกล่าวถึงเทคนิคการจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดให้อยู่ในระนาบตั้งตรง และการประยุกต์ใช้คุณลักษณะฮัฟทรานสฟอร์มเพื่อประกอบชิ้นส่วนเอกสาร ขั้นตอนการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตรงนี้เป็นการทำงานที่คาดการณ์ล่วงหน้าของชิ้นส่วน โดยงานวิจัยส่วนใหญ่จะประมาณค่าจากจุดสูงสุดของฮิสโทแกรมของทิศทางการเปลี่ยนแปลงค่าระดับความเข้มแสง อย่างไรก็ตามจุดสูงสุดของฮิสโทแกรมที่ได้นั้นอาจยังไม่ใช่ทิศทางการวางตัวหลักที่แท้จริงของชิ้นส่วน จึงมีงานวิจัยที่พัฒนาต่อโดยการประมาณค่าแนวทิศทางการวางตัวหลักของเนื้อหาภายในชิ้นส่วนจากฮิสโทแกรมผลรวมของจำนวนวงรีที่มีทิศทางเดียวกัน ซึ่งวิธีนี้ไม่สามารถทำให้ชิ้นส่วนตั้งตรงได้หากเนื้อหาภายในชิ้นส่วนมีปริมาณน้อยหรือเนื้อหาภายในชิ้นส่วนมีรูปภาพหรือเส้นตรงเป็นส่วนประกอบอยู่ด้วย ดังนั้นเพื่อให้การประมาณค่าทิศทางการวางตัวหลักของชิ้นส่วนที่แท้จริงมีค่าถูกต้องมากยิ่งขึ้น ในงานวิจัยนี้จะนำเสนอการคำนวณฮิสโทแกรมผลรวมของความยาวแกนหลักวงรีที่ล้อมรอบแต่ละกลุ่มของตัวอักษรที่มีทิศทางเดียวกันซึ่งจะพิจารณาเฉพาะวงรีที่ล้อมรอบตัวหนังสือชิ้นส่วนนั้น ส่วนวงรีที่ล้อมรอบรูปภาพหรือเส้นตรงจะไม่ถูกนำมาคิดคำนวณร่วมด้วย เมื่อได้ค่าประมาณของแนวทิศทางการวางตัวหลักของตัวหนังสือในชิ้นส่วนแล้วจะทำการหมุนกลับให้ชิ้นส่วนอยู่ในระนาบตั้งตรง จากนั้นจะทำการหาแนวหลักของเส้นขอบของชิ้นส่วนที่ตั้งตรงโดยใช้เทคนิคฮัฟทรานสฟอร์มในการรู้รูปร่างของชิ้นส่วนโดยการตรวจหาเส้นตรงที่มีในภาพ จึงทำให้สามารถหาแนวหลักของเส้นขอบของชิ้นส่วนเพื่อนำข้อมูลเส้นขอบที่ได้มาทำให้อยู่ในรูปแบบคู่อันดับของเส้นตรงคือจุดเริ่มต้นและจุดปลายสำหรับแต่ละด้านของขอบ จากนั้นหาจุดตัดหรือจุดปลายของเส้นที่สามารถลากเชื่อมต่อกันของแต่ละเส้นขอบ ตำแหน่งที่จุดตัดหรือจุดปลายของเส้นทั้งสองเชื่อมกันนั้นคือ มุมของชิ้นส่วน แต่ละมุมของชิ้นส่วนจะถูกเก็บขนาด ขอบแต่ละด้านของชิ้นส่วนจะถูกเก็บความยาวและมุมเขี้ยวของเส้นประกอบมุมซึ่งสามารถนำไปประยุกต์ใช้เป็นข้อมูลในขั้นตอนการประกอบชิ้นส่วนต่อไป

4.1 ภาพรวมของงานวิจัยที่นำเสนอ

การจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดให้อยู่ในระนาบเดียวกันนั้นเป็นขั้นตอนหนึ่งที่ทำให้การต่อเชื่อมชิ้นส่วนย่อยเป็นชิ้นส่วนขนาดใหญ่ได้อย่างมีประสิทธิภาพยิ่งขึ้น โดยขั้นตอนการจัดเรียงชิ้นส่วนเอกสารให้อยู่ในระนาบเดียวกันในงานวิจัยนี้แสดงดังรูปที่ 4.1

เริ่มจากนำภาพชิ้นส่วนเอกสารที่ถูกฉีกขาดมาแปลงจากภาพสีเป็นภาพขาวดำ จากนั้นนำขอบของชิ้นส่วนออกให้เหลือเฉพาะเนื้อหาหรือข้อมูลภายในชิ้นส่วนซึ่งประกอบด้วย ตัวอักษร เส้นตารางและรูปภาพ แล้วทำการหาทิศทางการเรียงตัวของตัวอักษรด้วยการตีกรอบวงรีล้อมรอบเนื้อหาภายในชิ้นส่วน ซึ่งจะต้องมีการกรองให้เหลือแต่วงรีที่ล้อมรอบกลุ่มตัวอักษรเท่านั้น หลังจากนั้นจึงคำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) เพื่อหาทิศทางการเอียงของชิ้นส่วนสำหรับนำมาใช้ในการจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดให้อยู่ในระนาบตั้งตรง เมื่อได้ชิ้นส่วน

ที่ตั้งตรงแล้วขอบของชิ้นส่วนที่ตั้งตรงจะถูกนำมาหาแนวหลักของเส้นขอบเพื่อให้ได้รูปร่างของชิ้นส่วน ซึ่งค่าของมุม ความยาวและเซตค่าขอบแต่ละด้านของชิ้นส่วนที่ได้จะถูกเก็บไว้เป็นข้อมูลสำหรับนำไปใช้ในการประกอบของชิ้นส่วน



รูปที่ 4.1 การจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดให้อยู่ในระนาบเดียวกัน

4.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร [29]

ในหัวข้อนี้เป็นการหาทิศทางการเอียงของชิ้นส่วนแล้วทำการจัดเรียงชิ้นส่วนเอกสารที่ถูกฉีกขาดเพื่อให้อยู่ในระนาบที่ตั้งตรง โดยมีหัวข้อ 4.2.1 การวิเคราะห์การหมุน ซึ่งกล่าวถึงวิธีการวิเคราะห์การหมุนชิ้นส่วนด้วยเทคนิค Spline Interpolation ซึ่งยังมีข้อจำกัดคือทิศทางเอียงของชิ้นส่วนที่ได้มีความผิดพลาดเมื่อชิ้นส่วนมีขอบที่ถูกฉีกในลักษณะของเส้นตรง งานวิจัยนี้จึงนำเสนอวิธีที่สามารถแก้ไขข้อจำกัดนี้ได้ในหัวข้อ 4.2.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษรด้วยวิธีการ

คำนวณผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ซึ่งสามารถหาทิศทางการเอียงของชิ้นส่วนได้อย่างมีประสิทธิภาพโดยไม่จำกัดรูปร่างและชนิดของเนื้อหาภายในชิ้นส่วน (ประกอบด้วยตัวอักษร เส้นตารางและรูปภาพ)

4.2.1 การวิเคราะห์การหมุน

เทคนิคการประมาณค่าการหมุนของชิ้นส่วนเอกสารที่ถูกฉีกขาด ได้ตัวอย่างมาจากอัลกอริทึม SIFT ของ Lowe [15] โดยนำมาประยุกต์ใช้ในการคำนวณขนาดการเปลี่ยนแปลง $m(x, y)$ และทิศทางการเปลี่ยนแปลง $\theta(x, y)$ ของชิ้นส่วนย่อยดังสมการต่อไปนี้

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (4.1)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (4.2)$$

เมื่อ $L(x, y)$ เป็นภาพอินพุตที่จะนำมาใช้ในการคำนวณค่าฮิสโทแกรมของแนวทิศการวางตัวของชิ้นส่วน โดยแต่ละบิตจะสัมพันธ์กับทิศทางของการเปลี่ยนแปลง $\theta(x, y)$ ดังนั้นค่าทิศทางของการเปลี่ยนแปลงของแต่ละพิกเซลจะถูกเก็บสะสมลงในบิตตรงกับทิศทางที่สัมพันธ์กัน และถ่วงน้ำหนักด้วยขนาดการเปลี่ยนแปลง $m(x, y)$ ของพิกเซลนั้นๆ จุดสูงสุดในกราฟฮิสโทแกรมแสดงให้เห็นทิศการเรียงตัวของชิ้นส่วนเอกสารที่ถูกฉีกขาดที่กำลังพิจารณา

เนื่องจากเวกเตอร์ของขนาดการเปลี่ยนแปลง ($m(x, y)$) ที่ถูกสะท้อนที่ขอบจะทำให้สามารถระบุได้ว่าขอบเป็นสีดำหรือสีขาวดำ อย่างไรก็ตามค่านี้ไม่สามารถนำมาใช้ในการบ่งบอกถึงทิศการเรียงตัวในพื้นที่สนใจได้ถูกต้องเสมอไป ดังนั้นค่าทิศทางของการเปลี่ยนแปลง ($\theta(x, y)$) จะถูกคำนวณบนช่วง $[-\pi/2 \pi/2]$ เวกเตอร์การเปลี่ยนแปลงที่ถูกสะท้อนจะเก็บสะสมไว้ที่บิตเดียวกันในฮิสโทแกรมของแนวทิศการวางตัวเพื่อนำไปสู่การปรับปรุงผลลัพธ์ของจุดสูงสุดของฮิสโทแกรมที่ดีขึ้น

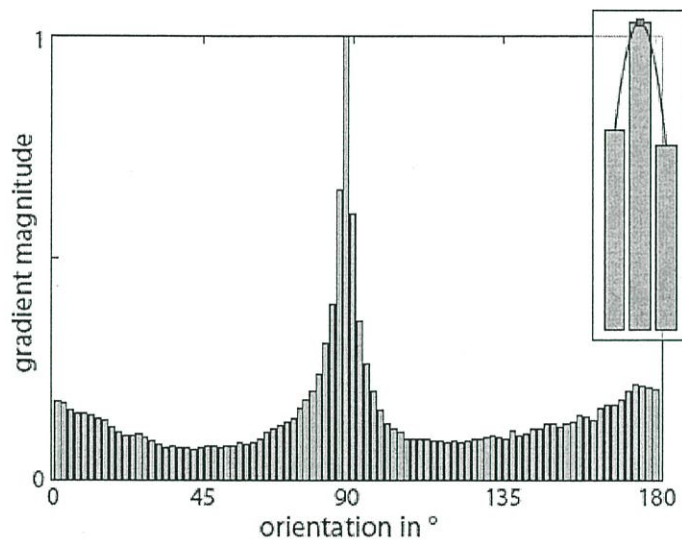
การประมาณค่าทิศการเรียงตัวของชิ้นส่วนไม่สามารถนำเฉพาะค่าบิตของฮิสโทแกรมสูงสุดมาพิจารณาเท่านั้น เนื่องจากอาจเกิดเหตุการณ์ที่เส้นขอบมีลักษณะฉีกขาดเป็นเส้นตรง จะทำให้จุดสูงสุดของฮิสโทแกรมกลายเป็นทิศของบริเวณเส้นขอบซึ่งไม่ใช่ทิศทางการเรียงตัวของชิ้นส่วนที่แท้จริงของข้อความในชิ้นส่วนนั้น ดังนั้นเพื่อแก้ปัญหาที่บิตที่มีค่าสูงสุดเมื่อเทียบกับเพื่อนบ้านหรือบิตข้างเคียงของมันจะถูกนำมาพิจารณาแทนที่การพิจารณาเฉพาะจุดสูงสุดหลักของฮิสโทแกรมของแนวทิศการวางตัว เพื่อให้เป็นไปตามวัตถุประสงค์นี้ ฮิสโทแกรมจะถูกทำให้เรียบด้วยเกาส์เซียน (Gaussian; $\sigma = 3$) จึงทำให้สามารถกำจัดตำแหน่งสูงสุดในพื้นที่ย่อยขนาดเล็กซึ่งไม่ใช่ตำแหน่งที่ให้ค่าสูงสุดหลักจริง จากนั้นทำการตรวจหาตำแหน่งสูงสุดในพื้นที่ย่อย (the local maxima: $l(x_i)$) จากฮิสโทแกรมที่ถูกทำให้เรียบแล้ว ตำแหน่งที่ตรวจพบจะถูกระบุลงในฮิสโทแกรมต้นฉบับก่อนถูกทำให้เรียบ เพื่อนำไปพิจารณาหาจุดสูงสุดหลักที่เป็นคำตอบตามสมการที่ 4.3

$$p = \max(l(x_i) - \text{median}[l(x_{i-j})l(x_{i+j})]) \quad (4.3)$$

เมื่อ $l(x_i)$ เป็นตำแหน่ง i^{th} ของฮิสโทแกรมที่พิจารณา

p คือ ตำแหน่งจุดสูงสุดหลัก และ j คือ ช่วงของเพื่อนบ้านที่นำมา พิจารณาในกรณีนี้คือ 11

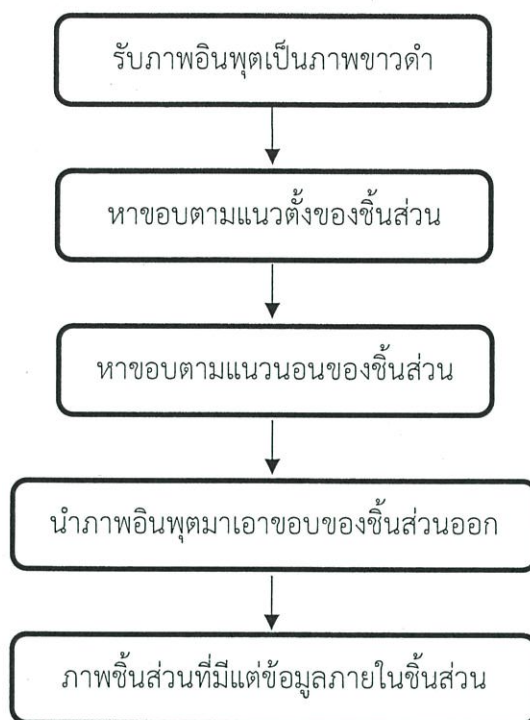
ความถูกต้องของทิศการเรียงตัวหลักที่พิจารณาตามสมการที่ 4.3 นั้น จะขึ้นอยู่กับจำนวนของบินในฮิสโทแกรมของแนวทิศการวางตัว (ตัวอย่างเช่น ช่วงของบิน 1 องศา จะมีจำนวนฮิสโทแกรมทั้งหมด 180 บิน ตั้งแต่มุมในช่วง $[-\pi/2 \pi/2]$) เพื่อเพิ่มความถูกต้องโดยไม่ขึ้นกับจำนวนบินของฮิสโทแกรมทิศการเรียงตัว ดังนั้นจึงนำเทคนิค Spline Interpolation (พหุนามลำดับที่ 2) มาใช้เพื่อประมาณค่าในช่วงของตำแหน่งค่าสูงสุดกับเพื่อนบ้านทั้งสองข้างของมันดังรูปที่ 4.2 โดยทิศทางการเรียงตัวหลักของชิ้นส่วนย่อยจะเป็นตำแหน่งที่ให้ค่าพหุนามสูงสุด



รูปที่ 4.2 กราฟฮิสโทแกรมของแนวทิศการวางตัวด้วย Spline Interpolation

4.2.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร

ดังที่ได้กล่าวในส่วนก่อนหน้านี้ว่าฮิสโทแกรมของแนวทิศการวางตัวจะเป็นตัวกำหนดทิศทางการเรียงตัวหลักภายในช่วง $[-\pi/2 \pi/2]$ ดังนั้นเพื่อทำการประมาณค่าควอเทรนต์ที่ถูกต้องจึงจำเป็นต้องพิจารณาเนื้อหาของชิ้นส่วนเอกสารที่ถูกฉีกขาด โดยจะทำการคำนวณรูปร่างที่ใช้เป็นกรอบของเนื้อหาที่พอดีที่สุด อย่างไรก็ตามการกำหนดเส้นขอบของชิ้นส่วนที่จะส่งผลให้เกิดความผิดพลาดในการวิเคราะห์กรอบวงรีของเนื้อหาเอกสาร ในงานวิจัยนี้จะนำชิ้นส่วนเอกสารที่ถูกฉีกขาดมาทำการปรับเป็นภาพขาวดำสองระดับ (Binarization) ด้วยอัลกอริทึมของอัทสุ (Otsu) [16] จากนั้นทำการหาขอบของชิ้นส่วนตามขั้นตอนในรูปที่ 4.3 และทำการกำจัดขอบของชิ้นส่วนออกไปเพื่อให้เหลือแต่เนื้อหาภายในชิ้นส่วน ดังตัวอย่างในรูปที่ 4.4



รูปที่ 4.3 ขั้นตอนการหาขอบของชิ้นส่วนเพื่อให้เหลือเฉพาะเนื้อหาภายในชิ้นส่วน



(ก) ภาพขาวดำสองระดับ

(ข) ภาพขอบของชิ้นส่วน

(ค) ภาพข้อมูลภายในชิ้นส่วน

รูปที่ 4.4 การกำจัดขอบของชิ้นส่วน

จากรูปที่ 4.4 (ข) แสดงขอบของชิ้นส่วนที่ได้จากการทำงานตามขั้นตอนในรูปที่ 4.3 โดยหาได้จากการหาขอบตามแนวตั้งในทุกๆ คอลัมน์ของชิ้นส่วนแล้วจึงหาขอบตามแนวนอนในทุกๆ แถวของชิ้นส่วน ทำให้ขอบของภาพชิ้นส่วนที่ได้แสดงมุมทึบทุกมุมและด้านทุกด้านของชิ้นส่วน จากนั้นนำขอบออกให้เหลือแต่ข้อมูลภายในของชิ้นส่วน เพื่อทำการหาทิศทางการเรียงตัวหลักของตัวอักษรในขั้นตอนต่อไป

จากการตัดขอบของชิ้นส่วนออกก่อนหน้านี้นี้นำผลลัพธ์ที่ได้มาระบุทิศทางการเรียงตัวหลักของชิ้นส่วนที่ถูกต้อง โดยใช้วิธีการหาทิศทางของวงรีซึ่งจะถูกเก็บสะสมลงในฮิสโทแกรมในแต่ละทิศทางของการหมุน ดังนั้นค่าที่ได้จากการคำนวณทิศทางในแต่ละบินของฮิสโทแกรมทิศทางของวงรี

ที่มีทิศทางการหมุนเดียวกันจะมีความสัมพันธ์กับจำนวนของวงรี ซึ่งผลลัพธ์ที่ได้ไม่ใช่ทิศการเรียงตัวของชิ้นส่วนที่แท้จริง กล่าวคือมีชิ้นส่วนที่ไม่สามารถตั้งตรงได้จากวิธีการหาทิศทางของวงรีโดยพิจารณาผลรวมของจำนวนวงรีที่มีทิศทางเดียวกัน เนื่องจากพื้นที่ที่สนใจส่วนใหญ่จะมีหลายองค์ประกอบทั้งตัวอักษร รูปภาพและเส้นตาราง ดังนั้นจึงต้องมีการคัดเลือกพื้นที่ที่สนใจเพื่อให้ได้เฉพาะวงรีที่ล้อมรอบตัวอักษรเท่านั้น เนื่องจากวงรีที่ล้อมรอบรูปภาพหรือเส้นตารางเมื่อนำมาพิจารณาร่วมกับวงรีที่ล้อมรอบตัวอักษรแล้วจะทำให้แนวทิศการเรียงตัวของชิ้นส่วนมีความผิดพลาดเกิดขึ้นได้ ในการคัดเลือกพื้นที่ที่สนใจจะพิจารณาจากขนาดและอัตราส่วนระหว่างแนวตั้งแนวนอนที่ถูกกำหนดให้แต่ละรูปวงรี โดยการคัดเลือกเฉพาะพื้นที่ที่สนใจที่นับว่าเป็นตัวอักษรเนื้อความนั้นจะคัดเลือกวงรีที่มีความยาวแกนรองระหว่าง 3 ถึง 15 เท่านั้น สำหรับบางพื้นที่ที่มีจุดครบวงซึ่งมีขนาดแกนรองต่ำกว่า 3 จะไม่นำมาพิจารณา และพื้นที่ซึ่งอาจเป็นรูปภาพหรือเส้นตารางจะไม่นำมาพิจารณาเช่นกัน โดยวงรีที่ถูกพิจารณาว่าเป็นรูปภาพจะมีขนาดแกนรองของวงรีมากกว่า 15 และวงรีที่ถูกพิจารณาว่าเป็นเส้นตารางจะมีขนาดของความยาวแกนหลักต้องมากกว่า 1.5 เท่าของความยาวแกนรอง โดยเงื่อนไขการพิจารณาวงรีที่อาจเป็นพื้นที่ที่มีจุดครบวง รูปภาพ และเส้นตารางแสดงดังรูปที่ 4.5 ผลลัพธ์ที่ได้หลังจากผ่านการกรองวงรีของข้อความที่ปราศจากจุดครบวง รูปภาพ และเส้นตารางแสดงในรูปที่ 4.6

INPUT:

ellipse property = {MajorAxis, MinorAxis, Orientation}

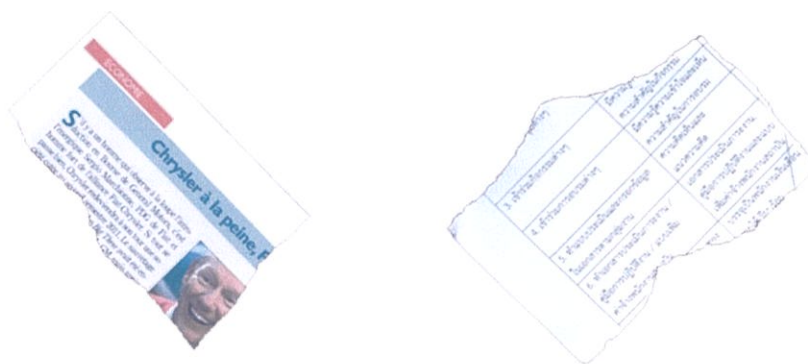
MajorAxis is length of ellipse's Major axis.

MinorAxis is length of ellipse's Minor axis.

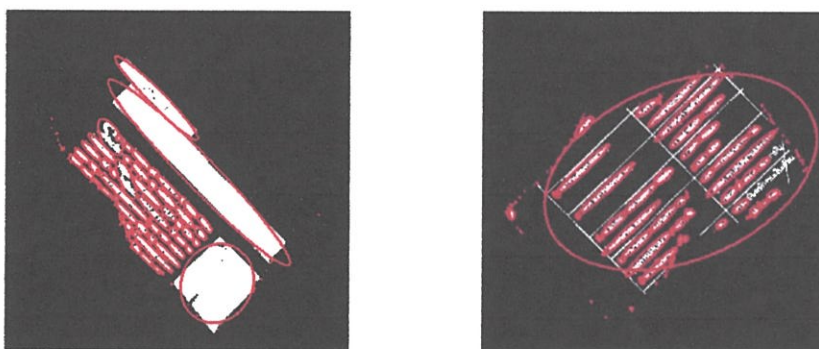
Orientation is orientation of ellipse (-90 to 90 degree).

1. FOR (Orientation of all ellipse)
2. IF (MajorAxis > 1.5*MinorAxis)
3. IF (MinorAxis > 3) && (MinorAxis < 15)
4. KEEP ellipse property ;
5. ENDIF
6. ENDIF
7. ENDFOR

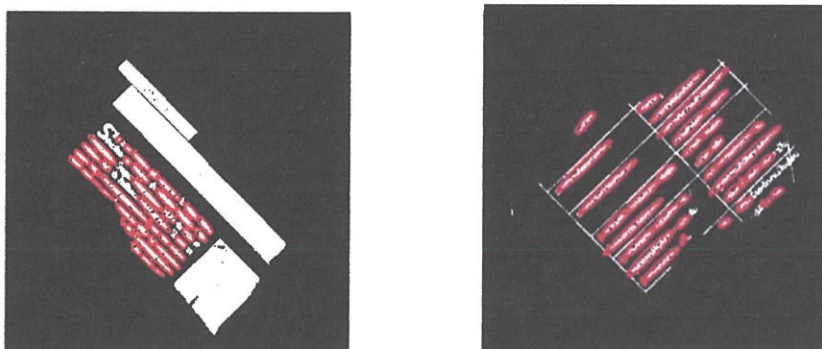
รูปที่ 4.5 วิธีการคัดเลือกให้พื้นที่ที่สนใจมีแต่ตัวอักษร



(ก) ภาพต้นฉบับ



(ข) วงรีล้อมรอบเนื้อหาภายในชิ้นส่วน



(ค) วงรีล้อมรอบเฉพาะตัวอักษร

รูปที่ 4.6 วงรีล้อมรอบเฉพาะตัวอักษร

รูปที่ 4.6 (ข) เป็นรูปภาพชิ้นส่วนข่าวคำที่มีรูปภาพและตัวอักษร (รูปที่ 4.6(ก)(ซ้าย)) กับ เส้นตารางและตัวอักษร (รูปที่ 4.6(ก)(ขวา)) เป็นเนื้อหาภายในชิ้นส่วนและทิศทางการเอียงของชิ้นส่วนคือ -45 องศาซึ่งถูกล้อมรอบโดยวงรีที่ไม่ได้คัดเลือกเฉพาะวงรีที่ล้อมรอบตัวอักษรจึงส่งผลให้ทิศทางหลักชิ้นส่วนที่ได้จากกราฟฮิสโทแกรมผลรวมของจำนวนวงรีที่มีทิศทางเดียวกันนั้นผิดพลาด เนื่องจากเส้นตารางเป็นเส้นที่มีความยาวมากกว่าผลรวมความยาวตัวอักษรทั้งหมด จึงทำให้วงรีที่ล้อมรอบเส้น

ตารางนั้นมีความยาวแกนหลักมากที่สุด ผลลัพธ์ที่ได้จึงเป็นมุมการเอียงของเส้นตารางซึ่งไม่ใช่ทิศทางการเอียงของชิ้นส่วนเมื่อหมุนกลับแล้วชิ้นส่วนจึงยังไม่ตั้งตรง

ดังนั้นในงานวิจัยนี้จึงนำเสนอการทำชิ้นส่วนให้อยู่ในระนาบตั้งตรงตามทิศการเรียงตัวหลักของตัวอักษรด้วยวิธีการคำนวณฮิสโทแกรมผลรวมขนาดความยาวแกนหลักวงรีของแนวทิศการวางตัวของตัวอักษรตามสมการที่ 4.4 ดังนี้

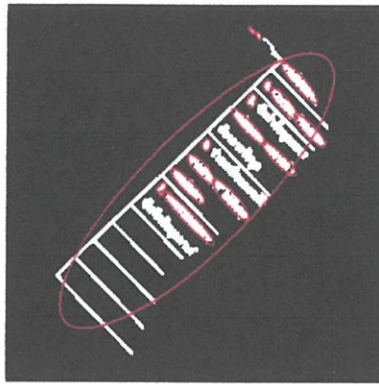
$$s = \max_{-90 \leq x_i \leq 90} (\sum_{j \in N} r_j(x_i)) \quad (4.4)$$

เมื่อ r คือ รัศมีของวงรี

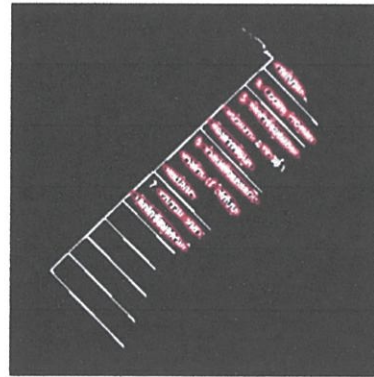
x_i เป็นตำแหน่ง i^{th} ของทิศทางการเปลี่ยนแปลงที่พิจารณา

s คือ ทิศที่มีผลรวมขนาดรัศมีของวงรีมากที่สุดซึ่งเป็นทิศหลักของการเรียงตัวของตัวอักษร

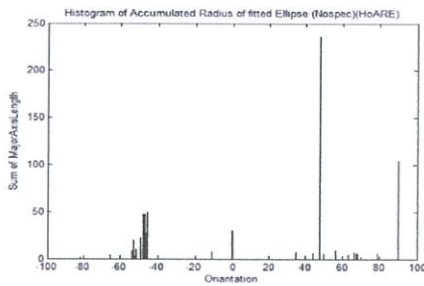
ผลรวมของขนาดทิศทางการเปลี่ยนแปลงในแต่ละทิศการเปลี่ยนแปลงคือ ความยาวแกนหลักของแต่ละวงรีที่ล้อมรอบพื้นที่ที่สนใจเฉพาะตัวอักษรจะถูกเก็บลงในบินตรงกับทิศทางที่สัมพันธ์กัน ทิศที่มีผลรวมขนาดความยาวแกนหลักของวงรีมากที่สุดจะถูกพิจารณาให้เป็นทิศหลักของการเรียงตัวของตัวอักษรซึ่งหมายถึงทิศทางการเรียงตัวหลักของชิ้นส่วนนั่นเอง ขั้นตอนสุดท้ายคือ การหมุนชิ้นส่วนให้อยู่ในระนาบที่ตั้งตรงตามทิศการเรียงตัวหลักของตัวอักษร โดยนำผลรวมของความยาวแกนหลักของวงรีที่มีทิศทางการเปลี่ยนแปลงหลักที่ได้จากสมการที่ 4.4 มาใช้เป็นทิศทางในการหมุนชิ้นส่วนกลับให้อยู่ในระนาบที่ตั้งตรง ซึ่งสามารถแสดงผลลัพธ์ทิศทางการเรียงตัวหลักของตัวอักษรของรูปที่ 4.7 (ข) ที่ได้จากสมการ 4.4 ดังรูปที่ 4.7 (ง) โดยผลลัพธ์ของทิศทางการเรียงตัวหลักของตัวอักษรที่ได้นั้นถูกต้องคือ -45 องศา ซึ่งช่วยแก้ปัญหาการวิเคราะห์ทิศการเรียงตัวหลักที่ผิดของรูปที่ 4.7 (ก) ที่ได้ผลลัพธ์ทิศการเรียงตัวหลักเป็น 47 องศา แสดงดังรูปที่ 4.7 (ค) คือกราฟฮิสโทแกรมผลรวมขนาดความยาวแกนหลักวงรีของเนื้อหาภายในชิ้นส่วน



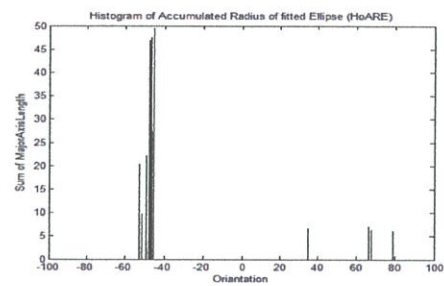
(ก)



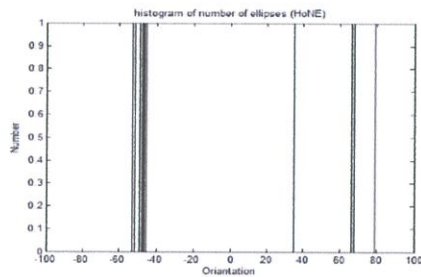
(ข)



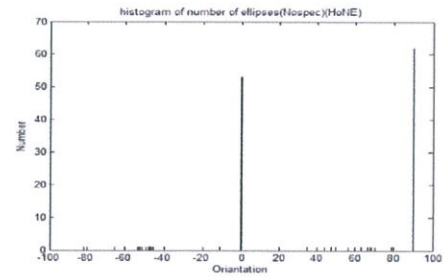
(ค)



(ง)



(จ)



(ฉ)

รูปที่ 4.7 การเปรียบเทียบทิศทางการเรียงตัวหลักของชิ้นส่วนที่ทำมุม -45 องศา

(ก) การตีกรอบวงรีล้อมรอบเนื้อหาภายในชิ้นส่วนที่มีตัวอักษรและเส้นตาราง

(ข) การตีกรอบวงรีล้อมรอบเฉพาะตัวอักษร

(ค) กราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ของรูป (ก)

(ง) กราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ของรูป (ข)

(จ) กราฟฮิสโทแกรมจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE) ของรูป (ก)

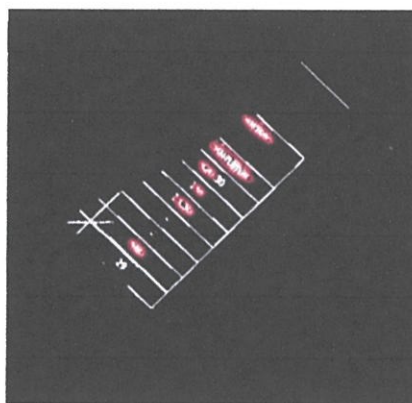
(ฉ) กราฟฮิสโทแกรมจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE) ของรูป (ข)

จากรูปที่ 4.7 เป็นการเปรียบเทียบการหาทิศทางการเรียงตัวของชิ้นส่วนเอียงที่ -45 องศาด้วยการนำวงรีที่ล้อมรอบเนื้อหาทั้งหมดภายในชิ้นส่วนและวงรีที่ล้อมรอบเฉพาะตัวอักษรมาพิจารณาโดยการคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) และการคำนวณกราฟฮิสโทแกรมผลรวมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) ซึ่งทิศการเรียงตัวของชิ้นส่วนที่ได้จากการนำวงรีที่ล้อมรอบเนื้อหาทั้งหมดภายในชิ้นส่วนที่ได้จากการคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ในรูปที่ 4.7 (ค) และวิธีคำนวณกราฟฮิสโทแกรมผลรวมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) ในรูปที่ 4.7 (จ) คือมุม 47 องศาและ 90 องศาตามลำดับ ส่วนการหาทิศการเรียงตัวของชิ้นส่วนที่ได้จากการนำวงรีที่ล้อมรอบเฉพาะตัวอักษรด้วยวิธีคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ในรูปที่ 4.7 (ง) และวิธีคำนวณกราฟฮิสโทแกรมผลรวมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) ในรูปที่ 4.7 (ฉ) คือมุม -45 องศา ซึ่งจะเห็นได้ว่าการหาทิศทางการเรียงตัวของชิ้นส่วนด้วยการนำวงรีที่ล้อมรอบเฉพาะตัวอักษรมาพิจารณานั้นให้ผลลัพธ์ที่มีความผิดพลาดน้อยกว่าการนำวงรีที่ล้อมรอบเนื้อหาทั้งหมดภายในชิ้นส่วน

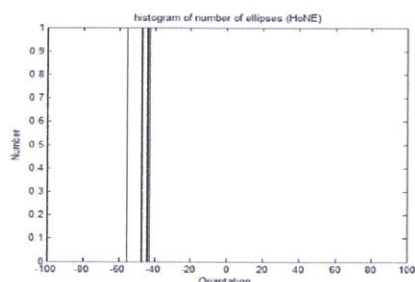
การหาทิศทางที่แท้จริงของตัวอักษรเพื่อให้ชิ้นส่วนตั้งตรงนั้น หากภายในชิ้นส่วนมีตัวอักษรน้อยจะส่งผลให้แนวหลักที่ได้มีมากกว่า 1 แนว ดังรูปที่ 4.8 (ค) เนื่องจากวิธีฮิสโทแกรมผลรวมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) เป็นการนับจำนวนวงรีที่ล้อมรอบแต่ละกลุ่มคำเมื่อตัวหนังสือในชิ้นส่วนมีปริมาณน้อยจึงทำให้ทิศที่มีจำนวนวงรีมากที่สุดเท่ากันหลายทิศคือ -43, -44, -45, -47, -48, -49 และ -55 องศา จำนวนเท่ากันหมด ด้วยข้อจำกัดนี้สามารถแก้ไขได้โดยวิธีการคำนวณฮิสโทแกรมผลรวมของขนาดความยาวแกนหลักวงรีของแนวทิศการวางตัวของตัวหนังสือ (HoARE) ในรูปที่ 4.8 (ง) โดยหาทิศทางหลักของตัวหนังสือจากผลรวมของความยาวแกนหลักวงรีที่มากที่สุด ทำให้แม้มีจำนวนวงรีน้อยแต่ทิศทางที่ได้ก็เป็นทิศทางของกลุ่มที่มีผลรวมรัศมียาวที่สุดคือ -45 องศาซึ่งให้ทิศทางที่ถูกต้อง



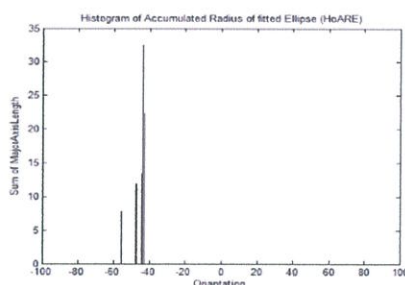
(ก)



(ข)



(ค)



(ง)

รูปที่ 4.8 การเปรียบเทียบผลลัพธ์ของการหาทิศทางของตัวหนังสือของชิ้นส่วนเอียง -45 องศา

(ก) ภาพต้นฉบับที่เป็นตารางและตัวหนังสือ (พิมพ์และเขียน)

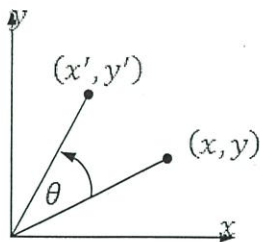
(ข) วงรีล้อมรอบกลุ่มคำ

(ค) ฮิสโทแกรมของจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE)

(ง) ฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่ล้อมรอบกลุ่มคำ (HoARE)

4.3 การหมุนกลับคืนชิ้นส่วน

การหมุนภาพจะต้องมีจุดศูนย์กลางของการหมุน เนื่องจากการหมุนด้วยจุดศูนย์กลางที่ต่างกันผลลัพธ์ที่ออกมาจะแตกต่างกัน โดยปกติจุดศูนย์กลางของการหมุนจะอยู่ที่ตำแหน่ง $(0,0)$ แต่สามารถใช้เทคนิคการเลื่อนแกนมาช่วยได้หากต้องการหมุนรอบจุดใดๆ ดังรูปที่ 4.9 ถ้าจุดในภาพอยู่ที่ตำแหน่ง (x, y) ถูกหมุนไปยังจุด (x', y') โดยถูกหมุนด้วยมุม θ รอบจุด $(0,0)$



รูปที่ 4.9 จุดที่ตำแหน่ง (x, y) ถูกหมุนไปยังจุด (x', y') โดยถูกหมุนด้วยมุม θ รอบจุด $(0, 0)$

สำหรับการหมุนมีสมการดังนี้

$$x'_n = x_n \cos \theta - y_n \sin \theta \quad (4.4)$$

$$y'_n = x_n \sin \theta + y_n \cos \theta \quad (4.5)$$

โดย (x', y') คือจุดคู่อันดับของมุม (x, y) ถูกหมุนด้วยมุม θ ในทิศทวนเข็มนาฬิกาการรอบจุด $(0, 0)$ สมการที่ 4.4 และ 4.5 สามารถนำมาเขียนร่วมกันให้อยู่ในรูปเมตริกซ์ ดังแสดงในสมการที่ 4.6

$$\begin{bmatrix} x'_n \\ y'_n \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (4.6)$$

อีกรูปแบบหนึ่งคือการหมุนย้อนกลับในทิศทาง $-\theta$ สมการและเมตริกซ์ที่ได้ เป็นดังนี้

$$x'_n = x_n \cos(-\theta) - y_n \sin(-\theta) \quad (4.7)$$

$$y'_n = x_n \sin(-\theta) + y_n \cos(-\theta) \quad (4.8)$$

$$\begin{bmatrix} x'_n \\ y'_n \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (4.9)$$

การหมุนหากมุมเป็นบวกภาพจะหมุนไปในทิศทวนเข็มนาฬิกา ในทางกลับกันหากมุมการหมุนเป็นลบภาพที่ได้จะถูกหมุนไปในทิศตามเข็มนาฬิกา การหมุนภาพของงานวิจัยนี้จะนำภาพต้นฉบับที่เป็นภาพขาวดำสองระดับที่ได้จากในขั้นตอนแรกมาทำการหมุนด้วยค่ามุมที่ได้จากการวิเคราะห์ทิศการหมุนหลักเพื่อให้ชิ้นส่วนตั้งตรง ซึ่งเป็นผลลัพธ์จากการวิเคราะห์ฮิสโทแกรมผลรวมความยาวแกนของวงรีที่ล้อมรอบกลุ่มคำดังที่กล่าวข้างต้น

4.4 การวิเคราะห์แนวหลักของเส้นขอบชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม

ฮัฟทรานสฟอร์มมีบทบาทสำคัญในการรู้รูปร่างของวัตถุ (Shape) โดยการตรวจหาเส้นตรง ส่วนโค้ง และวงกลมที่มีอยู่ในภาพ โดยก่อนที่จะนำภาพเข้ากระบวนการฮัฟทรานสฟอร์มได้นั้น ภาพอินพุตต้องผ่านการหาขอบของภาพมาก่อน

แนวคิดเบื้องต้นของฮัฟทรานสฟอร์มคือ ที่จุดใดๆของภาพสามารถมีเส้นตรงผ่านได้เป็นจำนวนไม่จำกัด ซึ่งเส้นตรงแต่ละเส้นจะใช้พิกัดเชิงขั้ว (polar equation) ดังนี้

$$\rho = x \sin \theta + y \cos \theta \quad (4.10)$$

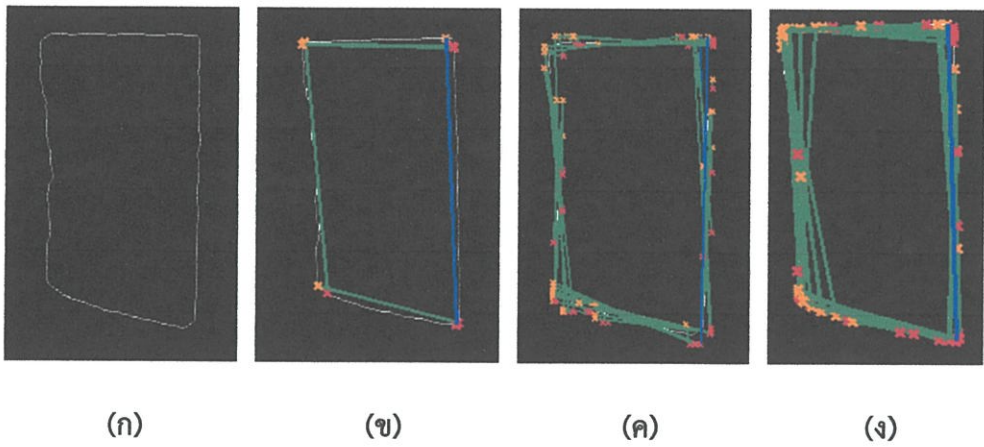
เมื่อ x, y คือ พิกัดของจุดภาพบนระนาบภาพ x, y

ρ คือ ระยะทางที่ตั้งฉากจากเส้นตรงถึงจุดกำเนิดในระนาบภาพ x, y

θ คือ มุมจากแกน x ถึงเส้นตั้งฉาก

จากพิกัดเชิงขั้วนี้ คือ (ρ, θ) จะทำให้สามารถเก็บข้อมูลได้ทั้ง 360 องศา (x, y) ของจุดภาพใดๆ ที่อยู่บนเส้นตรงเส้นเดียวกันจะสะสมอยู่ในค่าความถี่ของจุดช่องเดียวกัน โดยมีค่า ρ และ θ เท่ากัน ดังนั้นถ้าจำนวนที่สะสมในค่าความถี่ของจุดใดที่มีค่ามากกว่าค่าเรซโฮลด์ (threshold) ที่กำหนดไว้ จะถูกแปลงกลับไปเป็นพิกัด (x, y) ซึ่งจะทำให้เราทราบว่าเส้นตรงดังกล่าวมีพิกัดที่จุดใดบ้างบนระนาบ (x, y)

ข้อมูลที่ได้จากฮัฟทรานสฟอร์มสามารถนำไปประยุกต์ใช้ในการรู้รูปร่างของชิ้นส่วนได้ โดยมีพารามิเตอร์สำคัญที่จำเป็นต้องพิจารณาได้แก่ ค่าโรล ρ ค่าเซต้า θ และค่าเรซโฮลด์ ซึ่งมีผลต่อจำนวนเส้นและทิศทางของเส้นที่มาตีเพื่อเป็นตัวแทนของขอบของชิ้นส่วน หากเลือกค่าของพารามิเตอร์แต่ละตัวไม่เหมาะสม จะทำให้ได้เส้นที่ตีรอบขอบของชิ้นส่วนมีมากเกินไปหรือไม่ครบทุกด้านของชิ้นส่วนทำให้รูปร่างชิ้นส่วนไม่ถูกต้องดังรูปที่ 4.10 แสดงการตีกรอบเพื่อประมาณรูปร่างของชิ้นส่วนด้วยการปรับพารามิเตอร์ทั้งสาม ถ้ามีการปรับพารามิเตอร์ให้มีค่าน้อยจนเกินไปจะได้ผลลัพธ์การตีกรอบเพื่อกำหนดรูปร่างของชิ้นส่วนเป็นดังรูปที่ 4.10 (ข) จะเห็นว่าจำนวนเส้นฮัฟทรานสฟอร์มที่ล้อมรอบขอบของชิ้นส่วนมีน้อยจึงทำให้รูปร่างของชิ้นส่วนผิดรูปไปจากข้อมูลจริง (ground-truth) จากชิ้นส่วนมีสองด้านกลายเป็นชิ้นส่วนมีเพียงด้านเดียว (เส้นสีเขียวด้านซ้าย) จึงทำให้มุมของชิ้นส่วนหายไป และหากมีการปรับพารามิเตอร์ให้มีค่ามากเกินไปจะได้ผลลัพธ์การตีกรอบเพื่อกำหนดรูปร่างของชิ้นส่วนแสดงดังรูปที่ 4.10 (ง) ซึ่งจำนวนเส้นฮัฟที่ล้อมรอบขอบของชิ้นส่วนมีมากเกินไปทำให้รูปร่างของชิ้นส่วนไม่ชัดเจนและยากต่อการทำให้ได้เส้นขอบที่เหมาะสมในขั้นตอนต่อไป ดังนั้นการปรับพารามิเตอร์ทั้งสามที่เหมาะสมนั้น ควรปรับให้ผลลัพธ์การตีเส้นขอบเพื่อกำหนดรูปร่างของชิ้นส่วนนั้นมีมุมและด้านของชิ้นส่วนครบด้วยจำนวนเส้นที่ไม่น้อยเกินไปหรือมากเกินไปซึ่งจะทำให้รูปร่างของชิ้นส่วนเปลี่ยนไปดังรูปที่ 4.10 (ค) เป็นการตีเส้นขอบที่เหมาะสม คือมีมุมและด้านของชิ้นส่วนครบ เมื่อกรองเอาเส้นขอบตัวแทนของแต่ละด้านออกมาแล้วทำให้รูปร่างของชิ้นส่วนถูกต้องมากที่สุด



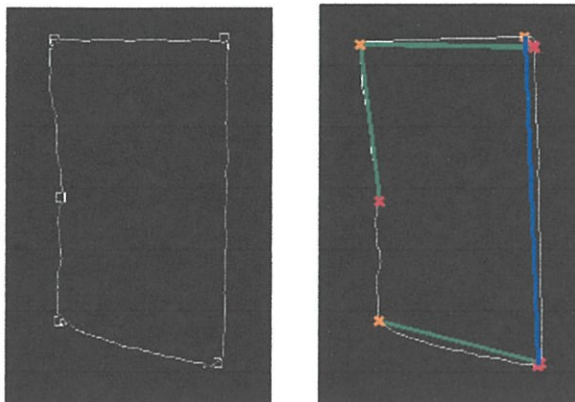
รูปที่ 4.10 ภาพแสดงการตีกรอบล้อมรอบขอบของชิ้นส่วนด้วยวิธี Hough Transform

- (ก) ภาพขาวดำของขอบชิ้นส่วน
- (ข) การตีขอบของชิ้นส่วนด้วยจำนวนเส้นที่น้อยเกินไป
- (ค) การตีขอบของชิ้นส่วนด้วยจำนวนเส้นที่เหมาะสม
- (ง) การตีขอบของชิ้นส่วนด้วยจำนวนเส้นที่มากเกินไป

ในการประกอบชิ้นส่วนจำเป็นต้องใช้ข้อมูลสำคัญคือ มุมและด้านของชิ้นส่วน หากมีการปรับพารามิเตอร์ทั้งสามไม่เหมาะสมจะส่งผลให้การประมวลรูปร่างของชิ้นส่วนผิดพลาดได้ กล่าวคือมุมและด้านที่ได้จากการปรับพารามิเตอร์ให้มีค่าน้อยจะทำให้ได้มุมและด้านของชิ้นส่วนไม่ครบ ส่วนการปรับค่าพารามิเตอร์มากเกินไปจะทำให้ชิ้นส่วนที่ถูกตีด้วยเส้นฮัฟนั้นมีค่าเซต้าของเส้นขอบหลักที่ทำให้รูปร่างของชิ้นส่วนผิดไปจากเดิมและหากมีเส้นมากเกินไปจะทำให้ขั้นตอนการเลือกเส้นตัวแทนใช้เวลานาน ดังนั้นจึงต้องมีการพิจารณาปรับพารามิเตอร์เพื่อให้ได้จำนวนเส้นเหมาะสมซึ่งจะทำให้ได้รูปร่างชิ้นส่วนถูกต้อง

เมื่อมีการกำหนดรูปร่างของชิ้นส่วนด้วยวิธีฮัฟทรานสฟอร์มโดยการเลือกค่าพารามิเตอร์ที่เหมาะสมแล้ว ยังพบว่าจำนวนเส้นที่ได้ในแต่ละด้านของชิ้นส่วนนั้นมีมากกว่าหนึ่งเส้นจะส่งผลให้เส้นขอบด้านนั้นมีมากกว่าหนึ่งเส้น เส้นเหล่านั้นจึงยังไม่ใช่อุปรูปที่ถูกต้อง ขั้นตอนต่อไปคือ วิธีการเลือกเส้นขอบโดยการกรองเส้นฮัฟในแต่ละด้านให้เหลือเส้นหลักด้านละหนึ่งเส้น จึงต้องมีการคำนวณหาจำนวนมุมของชิ้นส่วนในการรู้รูปร่างของชิ้นส่วน เนื่องจากความรู้พื้นฐานของรูปปิดนั้นมุมและด้านจะมีจำนวนเท่ากัน การที่จะรู้รูปร่างของชิ้นส่วนได้ถูกต้องนั้นสามารถทำได้โดยการแบ่งช่วงของค่าโรลและค่าเซต้าของฮัฟสเปซด้วยเงื่อนไขจากการคำนวณหาจำนวนมุมของชิ้นส่วน กล่าวคือถ้าจำนวนมุมของชิ้นส่วนมีจำนวนมากแล้วจำนวนช่วงที่ถูกแบ่งในฮัฟสเปซของค่าเซต้าและค่าโรลจะถูกแบ่งเป็นกลุ่มมากเช่นกัน ในทางกลับกันหากจำนวนมุมของชิ้นส่วนน้อยจำนวนช่วงของค่าโรลและเซต้าจะถูกแบ่งจำนวนช่วงน้อย สุดท้ายค่าของฮัฟสเปซที่มากที่สุดของแต่ละช่วงที่ถูกแบ่งนั้นจะถูกเลือกนำมาเป็นเส้นขอบหลักในแต่ละด้านชิ้นส่วน ดังนั้นเส้นหลักที่ถูกเลือกเป็นตัวแทนในแต่ละด้านของชิ้นส่วนจะมีจำนวนเส้นน้อยกว่าหรือเท่ากับจำนวนมุมของชิ้นส่วน ดังแสดงในรูปที่ 4.11 (ก) คือภาพที่ได้จากการคำนวณหามุมของชิ้นส่วน ซึ่งแต่ละมุมของชิ้นส่วนจะถูกตีกรอบด้วยสี่เหลี่ยมสีขาว และรูปที่ 4.11

(ข) คือภาพการตีขอบของชิ้นส่วนหลังจากวิธีการเลือกค่าฮัฟสเปซที่สูงที่สุดในแต่ละช่วงของค่าโรลและค่าเซต้า



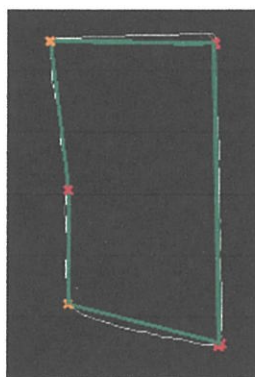
(ก)

(ข)

รูปที่ 4.11 (ก) ภาพมุมของชิ้นส่วนที่ถูกระบุโดยช่องสี่เหลี่ยม

(ข) เส้นขอบของชิ้นส่วนจากค่าสูงสุดของฮัฟสเปซในแต่ละช่วงโรลและเซต้า

จำนวนเส้นที่ได้จากขั้นตอนการเลือกเส้นขอบด้วยค่าสูงสุดของฮัฟสเปซในแต่ละช่วงค่าโรลและค่าเซตานั้นยังคงทำให้ด้านของขอบชิ้นส่วนไม่ครบ เนื่องจากในการแบ่งช่วงค่าโรลและค่าเซต้าของฮัฟสเปซนั้นมีช่วงหนึ่งที่ถูกแบ่งแล้วค่าฮัฟสเปซมีค่าเป็นศูนย์ทั้งหมดซึ่งทำให้ช่วงนั้นไม่มีค่าที่สูงที่สุดที่เป็นเส้นขอบออกมา จึงต้องมีการสร้างเส้นตรงสำหรับเชื่อมต่อจุดปลายของเส้นตรงที่ยังไม่มีจุดตัดกันกับเส้นตรงอีกเส้น และมีการเชื่อมเส้นตรงในตำแหน่งที่เส้นทั้งสองตัดกันด้วยพิกัดของจุดตัดนั้นเพื่อให้ได้รูปร่างของชิ้นส่วนครบทุกมุมและทุกด้านดังแสดงในรูปที่ 4.12 ซึ่งเป็นภาพที่เกิดจากการเชื่อมจุดปลายของรูปที่ 4.11



รูปที่ 4.12 ภาพแสดงรูปร่างของชิ้นส่วน

เมื่อได้จุดต้นจุดปลายของเส้นตรงในแต่ละด้านของขอบชิ้นส่วนแล้ว จึงนำมาคำนวณหาขนาดของแต่ละมุม ความยาวของแต่ละด้านประกอบมุม และมุมเซต้าของด้านประกอบมุม เพื่อนำมาเป็นข้อมูลในการประกอบชิ้นส่วน ซึ่งรูปแบบของข้อมูลที่เตรียมสำหรับการประกอบแสดงดังตารางที่ 4.1

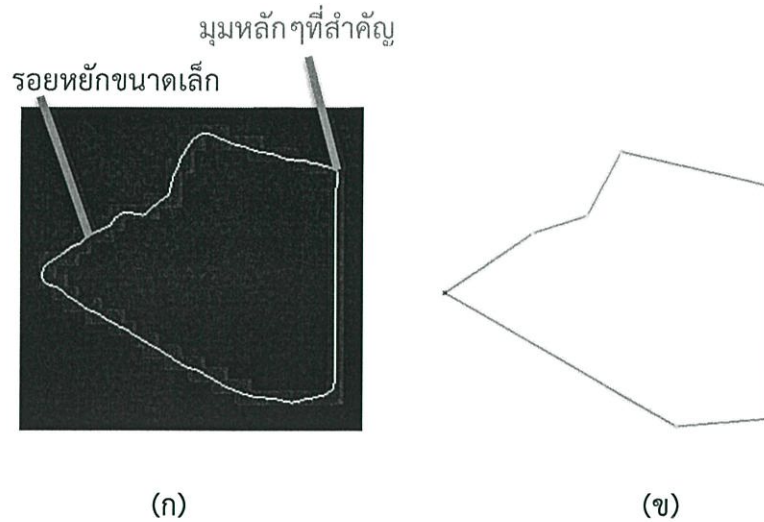
ตารางที่ 4.1 แสดงข้อมูลสำหรับการประกอบชิ้นส่วน

ลำดับ มุม	ขนาดมุม	ความยาวเส้น ก่อนมุม	ความยาวเส้นหลัง มุม	เซต้าเส้นก่อน มุม	เซต้าเส้นหลัง มุม
1	90	118	219	0	90
2	74.78742	219	118.1397	90	-164.787
3	102.4535	118.1397	83.09633	-164.787	-87.2409
4	187.1169	83.09633	105.3043	-87.2409	-94.357
5	85.64302	105.3043	118	-94.357	0

4.5 การประกอบชิ้นส่วน

ในขั้นตอนการประกอบชิ้นส่วนจะไม่มีกรรมชิ้นส่วนเพื่อพิจารณาการเข้าคู่กันของชิ้นส่วน เนื่องจากชิ้นส่วนนั้นถูกทำให้ตั้งตรงเรียบร้อยแล้ว จากขั้นตอนการจัดเรียงชิ้นส่วนให้ตั้งตรงนั้นทำให้ผลของการเข้าคู่ของชิ้นส่วนนั้นมีประสิทธิภาพมากขึ้นและระยะเวลาที่ใช้ในการประกอบชิ้นส่วนลดลงเพราะในการประกอบชิ้นส่วน ขั้นตอนที่ใช้เวลาในการคำนวณมากที่สุดคือขั้นตอนการเทียบว่าชิ้นส่วนเข้าคู่กันหรือไม่ ส่วนขั้นตอนการหมุนชิ้นส่วนนั้นใช้ระยะเวลานั้นซึ่งถ้าไม่จัดเรียงชิ้นส่วนให้ตั้งตรงก่อน จะต้องหมุนชิ้นส่วนที่ละหนึ่งองศาแล้วเทียบการเข้าคู่กันจึงทำให้ต้องใช้เวลาสำหรับพิจารณาการเข้าคู่ของชิ้นส่วนหนึ่งคู่คือ จำนวนครั้งที่หมุน*(ระยะเวลาการหมุน+ระยะเวลาพิจารณาการเข้าคู่) แต่ถ้ามีการทำให้ชิ้นส่วนตั้งตรงก่อนจะทำให้ระยะเวลาการเข้าคู่ของชิ้นส่วนหนึ่งคู่เป็น ระยะเวลาการหมุนชิ้นส่วนให้ตั้งตรง+ระยะเวลาพิจารณาการเข้าคู่

การประกอบชิ้นส่วนนั้นมีการเก็บข้อมูลของชิ้นส่วนเพื่อนำมาพิจารณารูปร่างของชิ้นส่วนหลากหลายวิธี เช่น การเก็บข้อมูลในรูปของพิกเซลของขอบชิ้นส่วน ซึ่งวิธีนี้จะเป็นการเก็บรูปร่างของชิ้นส่วนที่มีความละเอียดสูงเนื่องจากมุมทุกมุม (รอยหยักขนาดเล็กหรือมุมขนาดใหญ่) และด้านทุกด้าน (ส่วนเว้า ส่วนโค้งหรือเส้นตรง) ของชิ้นส่วนจะถูกเก็บรายละเอียดครบหมด แต่วิธีการเก็บข้อมูลแบบนี้ทำให้ระยะเวลาที่ใช้ในการประมวลผลสำหรับการประกอบชิ้นส่วนนานมาก และอาจทำให้ผลลัพธ์ของการเข้าคู่ของชิ้นส่วนมีความผิดพลาดมากด้วยเช่นกัน เนื่องจากรายละเอียดของมุมที่เป็นรอยหยักขนาดเล็กซึ่งมีในแต่ละด้านของชิ้นส่วนจะทำให้ผลลัพธ์ของการเข้าคู่เกิดการเข้าคู่ผิดได้มากกว่าการพิจารณาเฉพาะมุมหลักๆที่สำคัญของชิ้นส่วนดังรูปที่ 4.13 (ก)



รูปที่ 4.13 (ก) รูปร่างชิ้นส่วนของการเก็บข้อมูลแบบพิกเซลของขอบ
(ข) รูปร่างชิ้นส่วนโดยใช้คุณลักษณะฮัฟทรานสฟอร์ม

งานวิจัยนี้จึงเลือกวิธีการประกอบชิ้นส่วนโดยใช้คุณลักษณะฮัฟทรานสฟอร์ม ซึ่งนำโครงรูปร่างชิ้นส่วนที่เป็นด้านและมุมหลักมาใช้ในการพิจารณา ทำให้ลดระยะเวลาในการประกอบและลดโอกาสเกิดความผิดพลาดของการเข้าคู่ดังรูปที่ 4.13 (ข) ข้อมูลที่ใช้ในการประกอบชิ้นส่วนโดยใช้คุณลักษณะฮัฟทรานสฟอร์มประกอบด้วย ขนาดมุม ความยาวของด้านประกอบมุม และมุมเซต้าของด้านประกอบมุม โดยหลักการพิจารณาการเข้าคู่ของชิ้นส่วนนั้นจะเลือกพิจารณาคู่มุมของชิ้นส่วนที่ผ่านเงื่อนไขทั้งสามกรณีดังนี้

เงื่อนไขที่หนึ่งคือ ผลรวมของขนาดคู่มุมที่พิจารณารวมกันต้องมีค่าต่างจาก 180 หรือ 360 ไม่เกินค่าที่กำหนด (แทนด้วยตัวแปร ANG) เนื่องจากการเข้าคู่กันของชิ้นส่วนที่เป็นคู่กันนั้นมุมที่ประกอบกันจะมีผลรวมมุมเป็น 180 และ 360 องศา ดังแสดงในสมการที่ 4.11 และ 4.12

$$|180 - (a_a + a_b)| \leq ANG \quad (4.11)$$

$$|360 - (a_a + a_b)| \leq ANG \quad (4.12)$$

เมื่อ a_a คือ มุมชิ้นส่วนตั้งต้น

a_b คือ มุมของชิ้นส่วนอีกชิ้นที่นำมาพิจารณาร่วม

ANG คือ ค่าความต่างผลรวมขนาดคู่มุมที่พิจารณาจาก 180 หรือ 360

เงื่อนไขที่สองคือ ความยาวด้านประกอบมุมซึ่งด้านประกอบมุมนั้นมีสองด้านคือเส้นก่อนมุม (ด้านที่อยู่ก่อนมุม เช่นด้าน g_1 หรือ x_1 ในรูปที่ 4.14) และเส้นหลังมุม (ด้านที่อยู่หลังมุม เช่นด้าน g_2 หรือ x_2 ในรูปที่ 4.14) โดยความยาวด้านประกอบมุมของมุมคู่ชิ้นส่วนที่พิจารณาจะต้องมีความยาวต่างกันไม่เกินค่าที่กำหนด (แทนด้วยตัวแปร DIS) ดังแสดงในสมการที่ 4.13

$$|(d_{p_a} - d_{n_b})| + |(d_{n_a} - d_{p_b})| \leq DIS \quad (4.13)$$

เมื่อ d_{p_a} คือ ความยาวของด้านที่อยู่ก่อนมุมของชิ้นส่วนตั้งต้น

d_{n_a} คือ ความยาวของด้านที่อยู่หลังมุมของชิ้นส่วนตั้งต้น

d_{p_b} คือ ความยาวของด้านที่อยู่ก่อนมุมของชิ้นส่วนที่นำมาพิจารณาร่วม

d_{n_b} คือ ความยาวของด้านที่อยู่หลังมุมของชิ้นส่วนที่นำมาพิจารณาร่วม

DIS คือ ค่าความต่างของความยาวด้านประกอบมุมของคู่ชิ้นส่วน

และเงื่อนไขที่สามคือ เซตของด้านประกอบมุมนั้นต้องต่างกันไม่เกินค่าที่กำหนด (แทนด้วยตัวแปร $THETA$) โดยเซตของด้านประกอบมุมหนึ่งมุมนั้นมีสองค่าคือเซตด้านก่อนมุม (เซตของด้านที่อยู่ก่อนมุม) และเซตด้านหลังมุม (เซตของด้านที่อยู่หลังมุม) ดังแสดงในสมการที่ 4.14

$$|180 - |\theta_{p_a}| + |\theta_{n_b}|| + |180 - |\theta_{p_b}| + |\theta_{n_a}|| \leq THETA \quad (4.14)$$

เมื่อ θ_{p_a} คือ เซตของด้านที่อยู่ก่อนมุมของชิ้นส่วนตั้งต้น

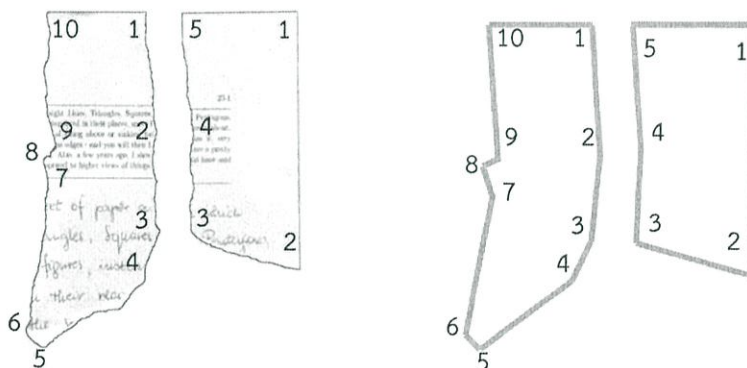
θ_{n_a} คือ เซตของด้านที่อยู่หลังมุมของชิ้นส่วนตั้งต้น

θ_{p_b} คือ เซตของด้านที่อยู่ก่อนมุมของชิ้นส่วนที่นำมาพิจารณาร่วม

θ_{n_b} คือ เซตของด้านที่อยู่หลังมุมของชิ้นส่วนที่นำมาพิจารณาร่วม

$THETA$ คือ ค่าความต่างระหว่างเซตของด้านประกอบมุม

ซึ่งการพิจารณาการเข้าคู่ทั้งสามกรณีนี้แสดงดังรูปที่ 4.14



(ก) ภาพต้นฉบับ

(ข) รูปร่างชิ้นส่วนจากเทคนิคฟทรานสฟอร์ม

รูปที่ 4.14 การเข้าคู่กันของชิ้นส่วน

จากรูปที่ 4.14 ในการพิจารณาการเข้าคู่กันของกลุ่มชิ้นส่วน กลุ่มที่สามารถประกอบกันได้ นั้นจะถูกพิจารณาจากสามเงื่อนไข โดยเงื่อนไขที่หนึ่ง คือการพิจารณาคู่มุมที่มีผลรวมของขนาดมุมของชิ้นส่วนซึ่งมีผลรวมคู่มุมต่างจาก 180 หรือ 360 ไม่เกินค่า ANG ที่กำหนดไว้ จากนั้นคู่มุมที่ผ่านเงื่อนไขที่หนึ่งจะถูกนำมาพิจารณาในกรณีที่สอง คือการพิจารณาผลรวมของผลต่างความยาวด้าน

ทั้งสองข้างต้นจะถูกนำมาพิจารณาในเงื่อนไขที่สาม คือการพิจารณาผลรวมของผลต่างเซตด้านประกอบมุมคู่ที่พิจารณา ซึ่งผลรวมที่ได้มีค่าไม่เกินค่า $THETA$ ที่กำหนดไว้ จากทั้งสามกรณีจะได้ว่ากลุ่มพิจารณาที่ตรงตามเงื่อนไขทั้งสามกรณีนั้นจะเป็นมุมที่สามารถเข้าคู่กัน ดังรูปที่ 4.15

สำหรับการเข้าคู่กันของกลุ่มชิ้นส่วนจากรูปที่ 4.14 (ข) เมื่อนำมาพิจารณาด้วยเงื่อนไขในรูปที่ 4.15 ซึ่งกลุ่มตัวอย่างคือ มุม 2 (รูปซ้ายของ 4.14 (ข)) และมุม 4 (รูปขวาของ 4.14 (ข)) สามารถแสดงได้ดังรูปที่ 4.16, 4.17 และ 4.18

INPUT:

Geometric shape property = {angle, d_p , d_n , θ_p , θ_n }

angle is the angle of corner

d_p is length of boundary (Previous)

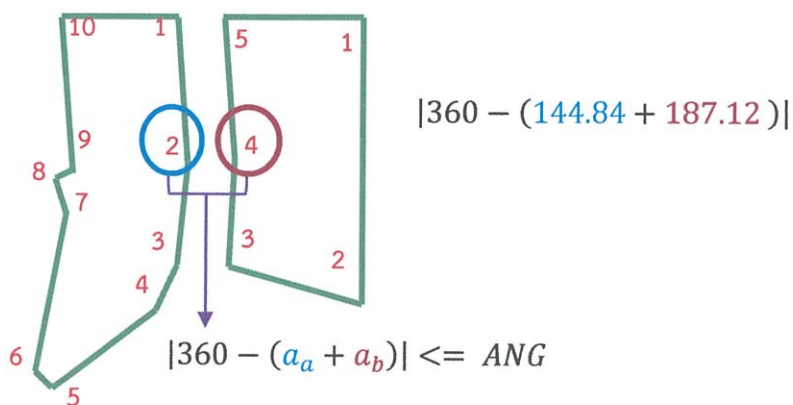
d_n is length of boundary (Next)

θ_p is theta of boundary (Previous)

θ_n is theta of boundary (Next)

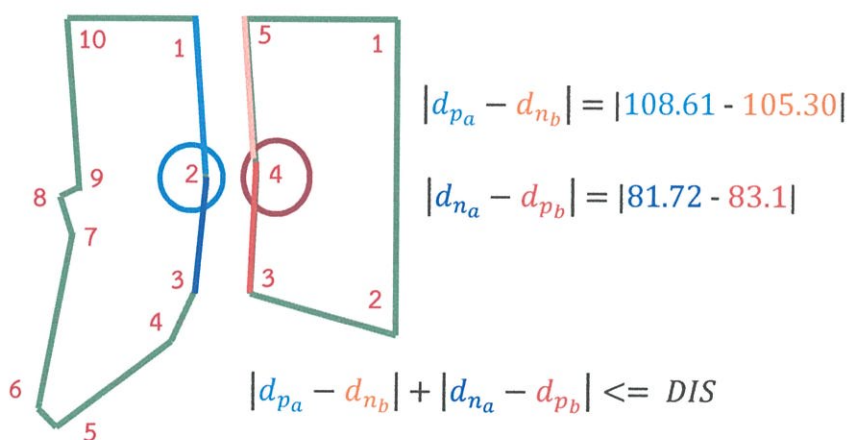
1. FOR (angle of corner)
2. IF $|180 - (a_a + a_b)| \leq ANG \parallel |360 - (a_a + a_b)| \leq ANG$
3. IF $|d_{p_a} - d_{n_b}| + |d_{n_a} - d_{p_b}| \leq DIS$
4. IF $|180 - |\theta_{p_a}| + |\theta_{n_b}|| + |180 - |\theta_{p_b}| + |\theta_{n_a}|| \leq THETA$
5. KEEP Geometric shape property;
6. END
7. END
8. END
9. END

รูปที่ 4.15 การพิจารณามุมที่เข้าคู่กันของชิ้นส่วน



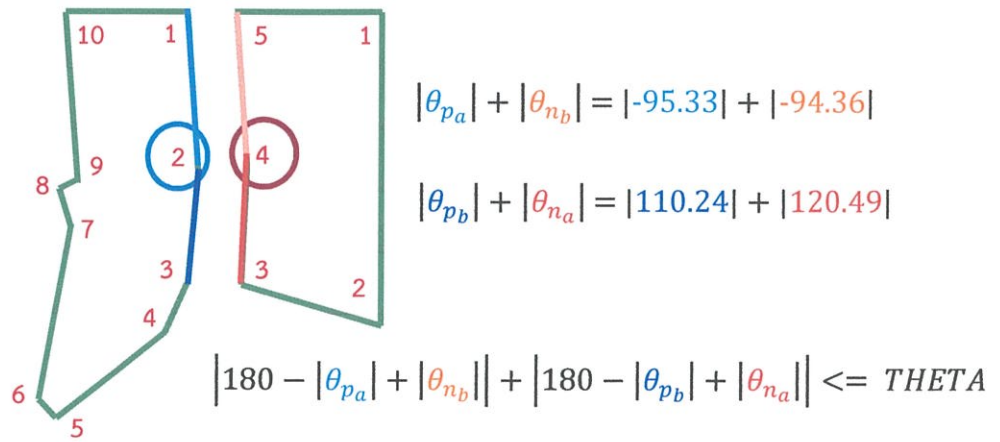
รูปที่ 4.16 คุ่มุมชิ้นส่วนที่ผ่านการพิจารณาในเงื่อนไขหนึ่ง

จากรูปที่ 4.16 มุม 2 (รูปซ้าย) และมุม 4 (รูปขวา) มีผลต่างของผลรวมขนาดมุมกับ 360 ไม่เกินค่า ANG ที่กำหนด ซึ่งถือว่าคุ่มุมนี้อยู่ในเงื่อนไขที่หนึ่ง จึงสามารถนำไปพิจารณาต่อในเงื่อนไขที่สอง ดังรูปที่ 4.17



รูปที่ 4.17 คุ่มุมชิ้นส่วนที่ผ่านการพิจารณาในเงื่อนไขสอง

จากรูปที่ 4.17 มุม 2 (รูปซ้าย) และมุม 4 (รูปขวา) มีผลรวมของผลต่างความยาวด้านประกอบมุมไม่เกินค่า DIS ที่กำหนด ซึ่งถือว่าคุ่มุมนี้อยู่ในเงื่อนไขที่สอง จึงสามารถนำไปพิจารณาต่อในเงื่อนไขที่สาม ดังรูปที่ 4.18



รูปที่ 4.18 กลุ่มชิ้นส่วนที่ผ่านการพิจารณาในเงื่อนไขสาม

จากรูปที่ 4.18 มุม 2 (รูปซ้าย) และมุม 4 (รูปขวา) มีผลรวมของผลต่างเซตด้านประกอบมุมไม่เกินค่า *THETA* ที่กำหนด ซึ่งถือว่าคู่มุมนี้อยู่ในเงื่อนไขที่สาม ดังนั้นคู่มุมนี้อยู่ในเงื่อนไขที่กำหนด จึงได้ว่ามุม 2 (รูปซ้าย) และมุม 4 (รูปขวา) สามารถประกอบกันได้

ผลการทดลองและการวิเคราะห์

ในบทนี้จะกล่าวถึงการทดลอง ผลการทดลอง และการวิเคราะห์ผลที่ได้จากการทดสอบการทำงานของการทำงานที่ขึ้นส่วนให้อยู่ในระนาบตั้งตรงจากกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) โดยจะทำการเปรียบเทียบกับงานวิจัย [2] ที่มีการศึกษาการเรียงตัวของขึ้นส่วนจากกราฟฮิสโทแกรมของจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE) เนื่องจากเทคนิคดังกล่าวให้ทิศทางการเอียงของขึ้นส่วนที่มีความถูกต้องสูง งานวิจัยนี้ได้นำเสนอการศึกษาการเรียงตัวของตัวอักษรจากกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) โดยการหาขอบของขึ้นส่วนเพื่อนำขอบของขึ้นส่วนออกให้เหลือแต่เนื้อหาภายในขึ้นส่วนแล้วจึงหาทิศทางการเรียงตัวหลักของตัวอักษร แต่เนื่องจากเนื้อหาภายในขึ้นส่วนนั้นประกอบด้วยตัวอักษร รูปภาพ หรือตาราง จึงจำเป็นต้องมีการทดลองสำหรับการนำวงรีที่ไม่ได้ล้อมรอบตัวอักษรออกเสียก่อนด้วยการกำหนดความยาวของแกนรองและอัตราส่วนระหว่างแกนหลักกับแกนรองของวงรีเพื่อให้ได้ทิศทางการเรียงตัวหลักของขึ้นส่วนที่ถูกต้อง ซึ่งการทดสอบประสิทธิภาพการหาทิศทางการเรียงตัวของขึ้นส่วนที่ได้นำเสนอในงานวิจัยนี้จะถูกประเมินออกมาอยู่ในรูปความถูกต้องแม่นยำของมุมของทิศทางการเรียงตัวของขึ้นส่วนที่จะทำให้ขึ้นส่วนหมุนกลับแล้วอยู่ในระนาบที่ตั้งตรง โดยมีการประเมินประสิทธิภาพด้วยการวัดอัตราของความผิดพลาด (Rate Of Error ; ROE) และการวัดค่าความถูกต้อง (Precision) เพื่อให้ได้ขึ้นส่วนที่ตั้งตรงซึ่งเป็นการเตรียมขึ้นส่วนสำหรับการประกอบ ในงานวิจัยนี้ได้นำผลลัพธ์จากการหาทิศทางหลักของขึ้นส่วนมาหาขอบขึ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม กล่าวคือการนำขอบขึ้นส่วนที่ตั้งตรงแล้วมาหารูปร่างของขึ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์มของเส้นตรง จากการทดลองการหาค่าพารามิเตอร์ที่ปรับเปลี่ยนได้ตามจำนวนมุมของขึ้นส่วนเพื่อให้ได้จำนวนมุมและจำนวนด้านของขึ้นส่วนจากเทคนิคฮัฟทรานสฟอร์มเหมือนกับรูปร่างของขึ้นส่วนจริงมากที่สุด จากการรู้รูปร่างของขึ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์มจึงทำให้ได้พารามิเตอร์ที่ใช้ในการประกอบขึ้นส่วน และกระบวนการสุดท้าย คือ การประกอบขึ้นส่วน ในงานวิจัยนี้จึงนำเสนอการจับคู่มุมของขึ้นส่วนที่เป็นคู่กันโดยประสิทธิภาพการประกอบคืนขึ้นส่วนที่ถูกฉีกขาดจะถูกวัดด้วยการวัดค่าความถูกต้อง (Precision)

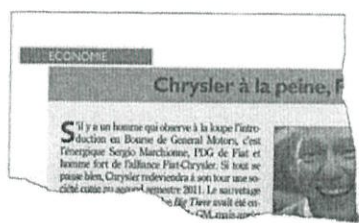
สำหรับส่วนของการทดลองและผลการทดลองแบ่งออกเป็น 6 หัวข้อหลักๆ ได้แก่ รายละเอียดฐานข้อมูลภาพขึ้นส่วน การวิเคราะห์ทิศทางการเรียงตัวของตัวอักษร การวิเคราะห์ประสิทธิภาพการหมุนกลับคืนขึ้นส่วน การวิเคราะห์ผลลัพธ์จากการหาทิศทางหลักของขอบขึ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม การวิเคราะห์พารามิเตอร์ที่ใช้ในการประกอบขึ้นส่วน และส่วนสุดท้ายคือการวิเคราะห์ประสิทธิภาพการประกอบคืนขึ้นส่วนที่ถูกฉีกขาด ดังรายละเอียดในแต่ละหัวข้อได้ดังนี้

5.1 รายละเอียดฐานข้อมูลภาพ

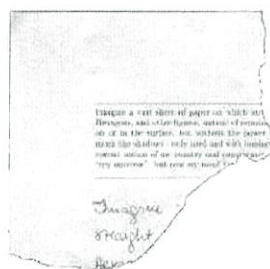
ในงานวิจัยนี้ได้สร้างข้อมูลจำลองความจริง (ground-truth) สำหรับใช้เป็นข้อมูลภาพพื้นฐานสำหรับเปรียบเทียบประสิทธิภาพของผลลัพธ์ที่ได้จากการจัดเรียงชิ้นส่วนของเอกสารที่ถูกฉีกขาดที่นำเสนอ โดยชิ้นส่วนที่นำมาใช้ในการทดลองมีรูปร่างและเนื้อหาภายในที่หลากหลายประกอบด้วยเอกสารฉีกขาดทั้งหมด 45 ชิ้น จากชุดการทดสอบ 7 ชุด โดยทดสอบให้ชิ้นส่วนแต่ละชิ้นอยู่ในทิศเอียง $[-90^{\circ} 90^{\circ}]$ มีขนาดตั้งแต่ 3 ตารางเซนติเมตรถึง 378 ตารางเซนติเมตร ใช้ความละเอียดในการสแกน 600 จุดภาพต่อความยาว 1 นิ้ว ในแต่ละชุดการทดสอบมีรูปแบบของชิ้นส่วนดังนี้

1. ชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ(แมกกาซีน) 8 ชิ้น
2. ชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ 7 ชิ้น
3. ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทย และตาราง 14 ชิ้น
4. ชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง 13 ชิ้น
5. ชิ้นส่วนที่มีตัวพิมพ์และลายมือมีทิศการเรียงตัวที่ต่างกัน
6. ชิ้นส่วนที่มีเฉพาะตัวพิมพ์
7. ชิ้นส่วนที่มีข้อมูลข้างในน้อยมาก (6 ตัวอักษร)

จากตัวอย่างของชุดข้อมูลภาพในแต่ละชุดการทดสอบแสดงตัวอย่างในรูปแบบที่ 5.1 ประกอบด้วยชิ้นส่วน 5 ชิ้นของแต่ละชุดการทดสอบ โดยชิ้นส่วน (ก) คือชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน) (ข) คือชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ (ค) คือชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง ในรูปที่ 5.2 (ก) คือชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง และรูปที่ 5.2 (ข) คือชิ้นส่วนที่มีข้อมูลข้างในน้อยมาก



(ก)



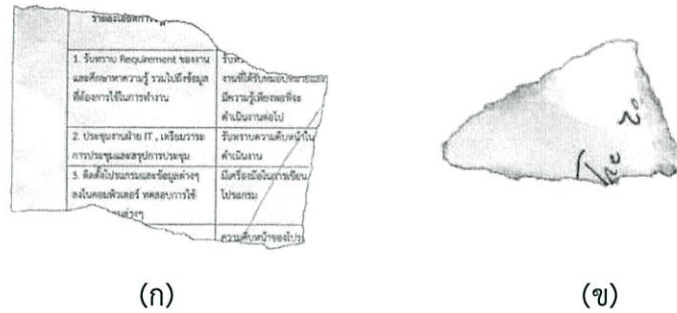
(ข)



(ค)

รูปที่ 5.1 ชิ้นส่วนตัวอย่างในแต่ละชุดการทดสอบ

- (ก) ชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ(แมกกาซีน)
 (ข) ชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ
 (ค) ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง



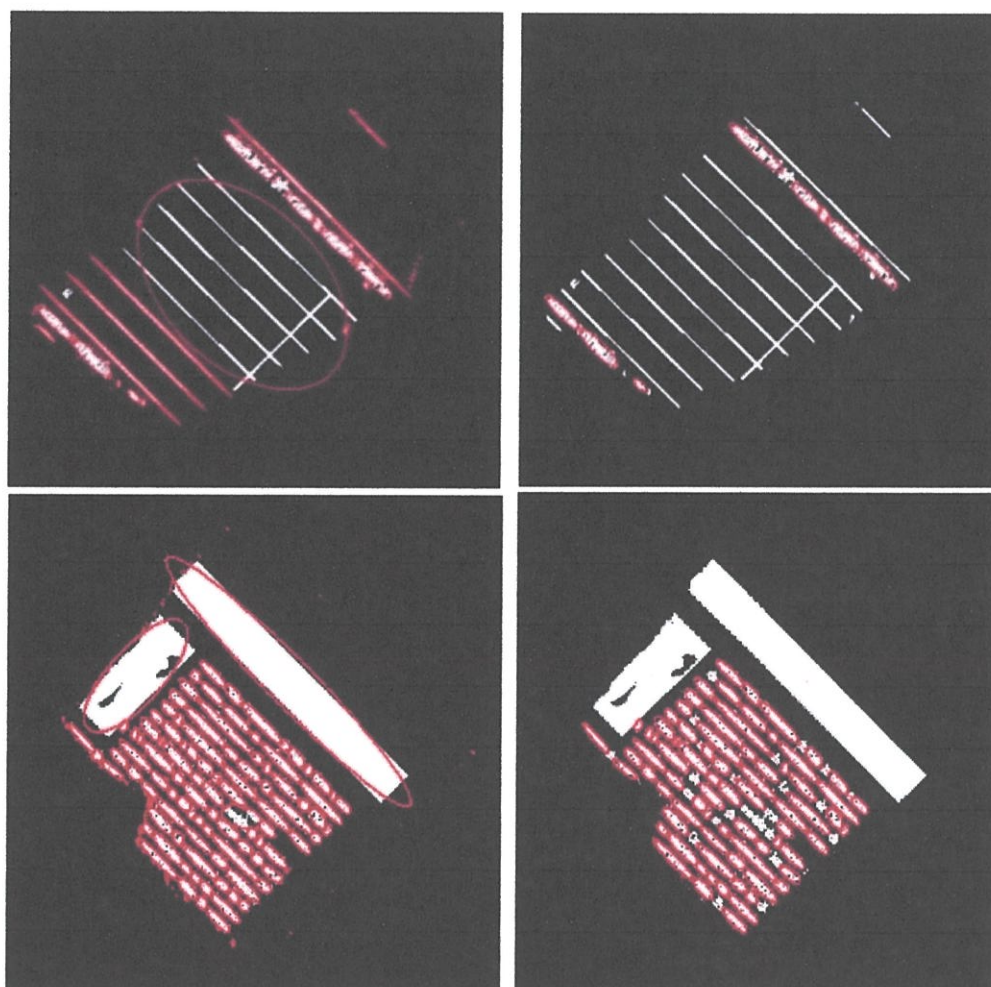
รูปที่ 5.2 ชิ้นส่วนตัวอย่างในแต่ละชุดการทดสอบ

- (ก) ชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง
(ข) ชิ้นส่วนที่มีข้อมูลข้างในน้อยมาก

5.2 การวิเคราะห์ทิศการเรียงตัวของตัวอักษร

การทดลองเพื่อวิเคราะห์ทิศการเรียงตัวของตัวอักษรนั้น คือการทดลองให้วงรีที่ใช้ในการหาทิศทางการเรียงตัวของชิ้นส่วนนั้นตีกรอบล้อมรอบเฉพาะตัวอักษรภายในชิ้นส่วน โดยวงรีที่ไม่ได้ล้อมรอบตัวอักษรซึ่งได้แก่ วงรีที่ตีกรอบล้อมรอบรูปภาพหรือตารางนั้นจะไม่ถูกนำมาคิดคำนวณหาทิศการเรียงตัวของชิ้นส่วน ขั้นตอนการวิเคราะห์จะเริ่มจากการนำภาพชิ้นส่วนที่ได้กำจัดขอบของชิ้นส่วนออกแล้วด้วยวิธีการขั้นแรกทำการประมวลผลภาพกับรูปร่างและโครงสร้างของภาพในรูปแบบการขยายองค์ประกอบ (Dilation) และทำลาเบลลิ่ง ต่อมาตีกรอบวงรีล้อมรอบกลุ่มเนื้อหาภายในชิ้นส่วน จากนั้นจึงทำการกรองเพื่อกำจัดการตีกรอบวงรีที่ไม่ได้ล้อมรอบตัวอักษรออกและนำผลที่ได้ไปเปรียบเทียบกับทิศทางการเรียงตัวของเนื้อหาภายในชิ้นส่วนทั้งหมด (มีทั้งวงรีล้อมรอบตัวอักษร รูปภาพและตาราง)

ในการวัดประสิทธิภาพการตีกรอบตัวอักษรด้วยวงรีของงานวิจัยนี้ใช้การวัดประสิทธิภาพจากจำนวนวงรีที่ตีล้อมรอบรูปภาพหรือเส้นตาราง โดยการทดลองการกรองวงรีที่ไม่ใช่วงรีที่ล้อมรอบตัวอักษรมีสองขั้นตอนคือ การกำหนดความยาวแกนรองของวงรีที่เป็นวงรีล้อมรอบตัวอักษร และการปรับอัตราส่วนระหว่างแกนหลักและแกนรองของวงรี เพื่อให้วงรีที่ได้หลังจากการกรองแล้วนั้นเป็นวงรีที่ตีล้อมรอบเฉพาะตัวอักษรภายในชิ้นส่วนเท่านั้น ถ้ายังเหลือวงรีที่ตีล้อมรอบรูปภาพหรือเส้นตารางอยู่แสดงว่าการกรองด้วยความยาวแกนรองของวงรีและการปรับอัตราส่วนระหว่างแกนหลักกับแกนรองของวงรียังไม่เหมาะสมเพียงพอ แสดงตัวอย่างผลการเปรียบเทียบการตีกรอบวงรีโดยไม่มีกรองวงรีและมีการกรองวงรีดังรูปที่ 5.3



(ก)

(ข)

รูปที่ 5.3 การตีกรอบตัวอักษรบนชิ้นส่วนที่มีรูปภาพและเป็นตาราง

(ก) การตีวงรีโดยไม่มีการกรองวงรี

(ข) การตีกรอบวงรีที่มีการกรองด้วยการกำหนดความยาวแกนรองของวงรี และการกรองด้วย อัตราส่วนระหว่างแกนหลักและแกนรองของวงรี

จากการทดลองในชุดการทดสอบทั้ง 5 ชุด แต่ละชุด ชิ้นส่วนสามารถกรองวงรีที่ไม่ได้ล้อมรอบตัวอักษรออกได้ทุกชิ้น ซึ่งแสดงผลที่ได้จากการทดลองในตารางด้านล่างของแต่ละชุดการทดสอบดังตารางที่ 5.1 5.2 5.3 5.4 และ 5.5 โดยในคอลัมน์ก่อนการกรองวงรีคือจำนวนวงรีที่ล้อมรอบจุดรบกวน รูปภาพหรือเส้นตารางก่อนมีการกรองวงรีให้เหลือแต่วงรีที่ล้อมรอบตัวอักษร และคอลัมน์หลังการกรองวงรี คือจำนวนวงรีที่ล้อมรอบรูปภาพหรือเส้นตารางหลังจากมีการกรองวงรีให้เหลือแต่วงรีที่ล้อมรอบตัวอักษร

ตารางที่ 5.1 จำนวนวงรีที่ล้อมรอบรูปภาพของชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)

ชิ้นส่วนที่	ก่อนการกรอวงรี	หลังการกรอวงรี
1	2	0
2	2	0
3	0	0
4	1	0
5	3	0
6	2	0
7	1	0
8	1	0
รวม	12	0

ตารางที่ 5.2 จำนวนวงรีที่ล้อมรอบเส้นค้นของชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ

ชิ้นส่วนที่	ก่อนการกรอวงรี	หลังการกรอวงรี
1	0	0
2	3	0
3	1	0
4	0	0
5	0	0
6	0	0
7	0	0
รวม	4	0

ตารางที่ 5.3 จำนวนวงรีที่วางรูปภาพหรือเส้นตาราง

ชิ้นส่วนที่	ก่อนการกรอวงรี	หลังการกรอวงรี
1	1	0
2	2	0
3	2	0
4	1	0
5	1	0
6	1	0
7	1	0
8	2	0

ตารางที่ 5.3 (ต่อ)

9	1	0
10	4	0
11	2	0
12	1	0
13	1	0
14	1	0
รวม	21	0

ตารางที่ 5.4 จำนวนวงรีที่ล้อมรอบเส้นตารางของชั้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง

ชั้นส่วนที่	ก่อนการกรอวงรี	หลังการกรอวงรี
1	4	0
2	5	0
3	3	0
4	4	0
5	2	0
6	2	0
7	4	0
8	1	0
9	1	0
10	1	0
11	2	0
12	2	0
13	1	0
รวม	32	0

ตารางที่ 5.5 ชิ้นส่วนจากงานวิจัย [2]

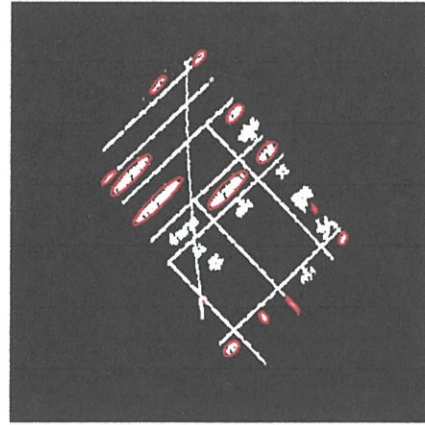
ชิ้นส่วนที่	ก่อนการกรองวงรี	หลังการกรองวงรี
1	0	0
2	2	0
3	2	0
รวม	2	0

5.3 การวิเคราะห์ประสิทธิภาพการหมุนกลับคืนชิ้นส่วน

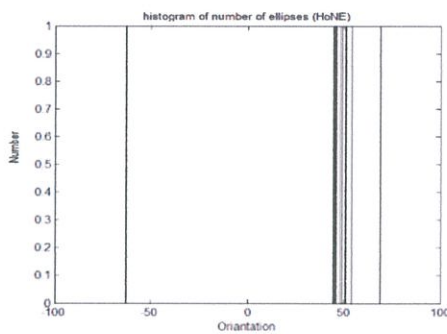
หลังจากขั้นตอนการตีกรอบวงรีล้อมรอบเฉพาะตัวอักษรแล้ว วงรีที่ได้จะถูกนำมาคำนวณหาทิศการเรียงตัวของตัวอักษร ในการทดลองได้ทำการเปรียบเทียบความถูกต้องของทิศทางการเรียงตัวหลักของตัวอักษรที่ทำให้การหมุนกลับคืนของชิ้นส่วนที่ได้ด้วยการหมุนกลับมามีค่าตั้งตรงจากงานวิจัย [2] กับอัลกอริทึมที่นำเสนอ โดยวิธีการของงานวิจัย [2] ใช้หลักการหาทิศทางการเรียงตัวหลักของตัวอักษรจากกราฟฮิสโทแกรมของจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE) โดยเป็นการหาทิศทางการเรียงตัวหลักของตัวอักษรจากทิศทางของวงรีที่มีจำนวนวงรีที่มากที่สุด ซึ่งวิธีการนี้มีประสิทธิภาพการหาทิศการเรียงตัวหลักของชิ้นส่วนได้ดี แต่ยังคงมีข้อจำกัดในกรณีที่เนื้อหภายในชิ้นส่วนมีจำนวนน้อย เช่น สองถึงสามคำ จะส่งผลให้จำนวนของวงรีที่มากที่สุดมีมากกว่าหนึ่งทิศทำให้ไม่สามารถหาทิศทางการเรียงตัวของชิ้นส่วนได้ ดังนั้นจากวิธีการที่นำเสนอในงานวิจัยนี้ด้วยการหาทิศทางการเรียงตัวหลักของตัวอักษรจากการคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ซึ่งการหาทิศทางการเรียงตัวหลักของตัวอักษรจากทิศทางของวงรีที่มีผลรวมความยาวแกนหลักวงรีที่มากที่สุด ซึ่งวิธีการนี้มีประสิทธิภาพการหาทิศการเรียงตัวหลักของชิ้นส่วนได้ดีและสามารถหาทิศทางหลักของตัวอักษรได้แม้เนื้อหภายในชิ้นส่วนมีจำนวนน้อย โดยสามารถแก้ไขข้อจำกัดของงานวิจัย [2] ได้ ดังแสดงตัวอย่างเปรียบเทียบประสิทธิภาพการหมุนกลับคืนชิ้นส่วนในรูปที่ 5.4



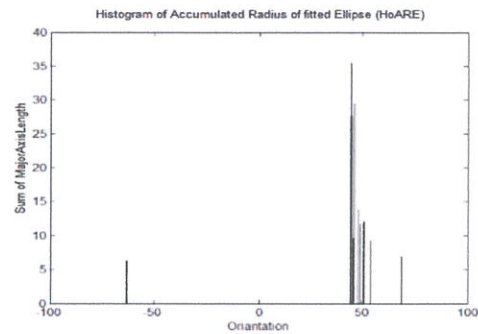
(ก)



(ข)



(ค)



(ง)

รูปที่ 5.4 การเปรียบเทียบทิศทางการเรียงตัวหลักของชิ้นส่วนเอียงที่ 45 องศา

(ก) ภาพต้นฉบับที่มีทิศการเรียงตัวของชิ้นส่วน 45 องศา

(ข) แสดงการตีกรอบวงรีล้อมรอบตัวอักษร

(ค) กราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ของรูปที่ 5.4 (ข)

(ง) กราฟฮิสโทแกรมของจำนวนวงรีที่มีทิศทางเดียวกัน (HoNE) ของรูป (ข)

จากรูปที่ 5.4 เป็นการเปรียบเทียบการหาทิศทางการเรียงตัวหลักของชิ้นส่วนเอียงที่ 45 องศาด้วยการนำวงรีที่ล้อมรอบเฉพาะตัวอักษรมาพิจารณาโดยการคำนวณกราฟฮิสโทแกรมผลรวมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) และการคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ซึ่งทิศการเรียงตัวหลักของชิ้นส่วนที่ได้จากการนำวงรีที่ล้อมรอบตัวอักษรภายในชิ้นส่วนที่ได้จากวิธีคำนวณกราฟฮิสโทแกรมจำนวนของวงรีที่มีทิศทางเดียวกัน (HoNE) ในรูปที่ 5.4 (ค) และการคำนวณกราฟฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่มีทิศทางเดียวกัน (HoARE) ในรูปที่ 5.4 (ง) คือมุม 53 องศาและ 45 องศาตามลำดับ

การวัดประสิทธิภาพการหมุนคืนกลับของชิ้นส่วนการหาทิศการเรียงตัวของตัวหนังสือในงานวิจัยนี้จะเป็นการประมวลผลการวิเคราะห์การหมุนกลับคืนชิ้นส่วนที่อธิบายไว้ในหัวข้อ 4.2 โดย

นำทิศทางการเรียงตัวหลักของชิ้นส่วนที่ได้จากอัลกอริทึมที่นำเสนอเทียบกับทิศการเอียงของชิ้นส่วนที่ถูกต้องด้วยผู้ทดลอง (Ground truth) แล้วทำการวัดค่าความถูกต้องในการทำให้ชิ้นส่วนตั้งตรงด้วยวิธีการวัดประสิทธิภาพทิศการเรียงตัวหลักแบบค่าความถูกต้อง (Precision) และวิธีการวัดอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วน (ROE) [2] โดยวิธีการวัดประสิทธิภาพของทิศการเรียงตัวหลักแบบค่าความถูกต้องสามารถหาได้จากสมการที่ 5.1 และวิธีการวัดอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วนสามารถหาได้จากสมการที่ 5.2 และ 5.3

$$P = \sum_{i=1}^N \left(\frac{D_i}{N} \right) \quad (5.1)$$

เมื่อ P คือ เปอร์เซ็นต์ความถูกต้องของจำนวนชิ้นส่วนในการหมุนกลับของแต่ละมุมที่ทดสอบ

D_i คือ เปอร์เซ็นต์ของจำนวนชิ้นส่วนที่มีทิศทางการเรียงตัวของชิ้นส่วนที่ถูกต้องในแต่ละชุดทดสอบ i

N คือ จำนวนชุดของชิ้นส่วนทั้งหมด

วิธีการวัดอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วนสามารถหาได้จาก

$$e_1 = \|\text{mod}(\alpha_c, 90) - \text{mod}(\alpha_m, 90)\| \quad (5.2)$$

$$e = \min(e_1, 90 - e_1) \quad (5.3)$$

เมื่อ α_m เป็นทิศทางมุมตามค่าจำลองความจริงที่กำหนด α_c เป็นมุมที่ประมาณค่าได้จากการทดลอง และ e เป็นค่าความผิดพลาดสุดท้าย การคำนวณความผิดพลาดนี้พิจารณาเพียงความแตกต่างของมุมภายในหนึ่งควอดแรนต์ $[0 \pi/2]$ ดังนั้นวิธีการที่อธิบายไว้ไม่ได้พิจารณาว่าชิ้นส่วนกลับหัวหรือไม่

ตารางที่ 5.6 ผลจากการวัดความถูกต้องด้วยวิธีการวัดอัตราความผิดพลาดของทิศการเรียงตัวของ
 ชั้นส่วน (ROE) และวิธีการวัดประสิทธิภาพทิศการเรียงตัวหลัก (Precision)

ชุดข้อมูล	จำนวน (ชั้น)	RoE [2]	RoE propose	Precision [2]	Precision propose
1. ชั้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)	8	12.39	2.06	29.17	91.67
2. ชั้นส่วนที่มีตัวพิมพ์และลายมือ ภาษาอังกฤษ	7	10.54	0.53	42.86	80.95
3. ชั้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทย และเป็นตาราง	14	2.93	0.09	80.95	95.24
4. ชั้นส่วนที่มีตัวพิมพ์ภาษาไทยและ เป็นตาราง	13	5.39	0.05	79.49	100
5. ชั้นส่วนที่มีตัวพิมพ์และลายมือมีทิศ การเรียงตัวที่ต่างกัน [2]	1	0.04	0	100	100
6. ชั้นส่วนที่มีเฉพาะตัวพิมพ์ [2]	1	0.49	0	100	100
7. ชั้นส่วนที่มีข้อมูลข้างในน้อยมาก [2]	1	6.43	0	66.67	100
ค่าเฉลี่ย	45	5.46	0.39	71.30	95.41

จากตารางที่ 5.6 แสดงผลจากการวัดความถูกต้องด้วยวิธีการวัดอัตราความผิดพลาดของทิศการเรียงตัวของชั้นส่วน (ROE) และวิธีการวัดประสิทธิภาพทิศการเรียงตัวหลักของแต่ละชุดชั้นส่วน (Precision) ที่ได้จากการทดลองของวิธีการที่นำเสนอในงานวิจัยฉบับนี้เทียบกับงานวิจัย [2] ในชุดทดสอบที่มีความผิดพลาดของทิศการเรียงตัวของชั้นส่วน (ROE) มากที่สุดคือ ชั้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน) เนื่องจากในชั้นส่วนดังกล่าวมีรูปภาพที่มีตัวอักษรอยู่ภายใน ซึ่งตัวอักษรที่อยู่ภายในรูปภาพนั้นเป็นกลุ่มตัวอักษรของภาพกราฟิกที่อยู่ในแนวเอียงจากแนวตัวอักษรปกติ จึงให้ค่าความผิดพลาดของทิศทางการหมุนกลับของชั้นส่วนให้อยู่ในระนาบตั้งตรงสูงกว่าชุดการทดสอบอื่นๆ แต่จะเห็นได้ว่าอัตราความผิดพลาดของทิศการเรียงตัวของชั้นส่วน (ROE) ของงานวิจัยที่นำเสนอ นั้นมีค่าความผิดพลาดน้อยกว่าวิธีการของงานวิจัย [2] อยู่ 5.07 ในขณะที่ประสิทธิภาพทิศการเรียงตัวหลักของแต่ละชุดชั้นส่วน (Precision) ของงานวิจัยฉบับนี้มากกว่างานวิจัย [2] อยู่ 24.11 เปอร์เซ็นต์

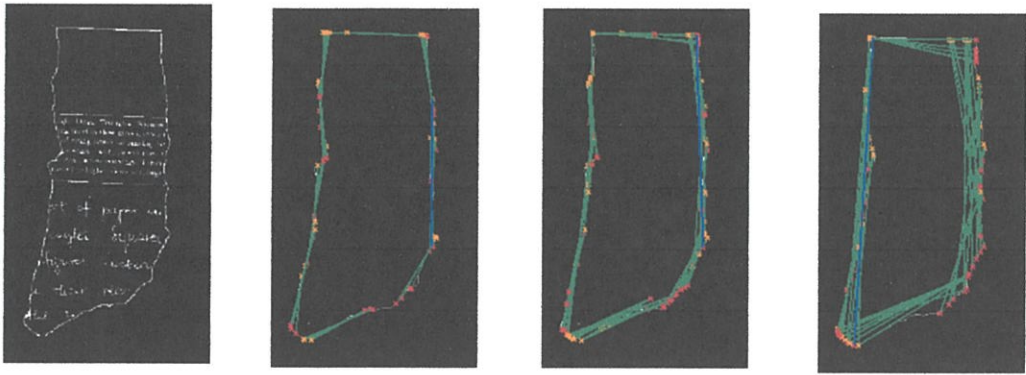
จากผลการทดลองพบว่าการจัดเรียงชั้นส่วนให้อยู่ในระนาบตั้งตรงด้วยวิธีการในงานวิจัย [2] นั้น ถึงแม้ว่าจะมีความถูกต้องของจำนวนชั้นส่วนที่มีผลลัพธ์ทิศทางการเรียงตัวของชั้นส่วนสูง แต่เมื่อเปรียบเทียบกับค่าความผิดพลาดของชั้นส่วนที่ได้จากผลต่างของทิศทางที่แท้จริงนั้นมีความผิดพลาดค่อนข้างสูง เมื่อนำมาเปรียบเทียบกับวิธีการที่งานวิจัยนี้นำเสนอ พบว่าอัลกอริธึมที่นำเสนอสามารถ

เพิ่มประสิทธิภาพของจำนวนชิ้นส่วนที่มีทิศทางหมุนกลับคืนที่ถูกต้องได้สูงกว่าและค่าความผิดพลาดของผลต่างทิศทางที่คำนวณได้กับทิศทางที่แท้จริงนั้นมีความผิดพลาดน้อยกว่าอีกด้วย

5.4 การวิเคราะห์ผลลัพธ์จากการหาทิศทางหลักของขอบชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม (Hough Transform)

หลังจากที่ชิ้นส่วนถูกหมุนกลับคืนให้อยู่ในระนาบตั้งตรงแล้ว ขอบของชิ้นส่วนที่ตั้งตรงจะถูกนำไปหารูปร่างของชิ้นส่วนด้วยการทำเทคนิคฮัฟทรานสฟอร์มของขอบของชิ้นส่วน ในขั้นตอนนี้เพื่อให้ได้ข้อมูลในฮัฟสเปซ (Hough space) ที่ทำให้ได้ขอบของชิ้นส่วนครบทุกด้าน ซึ่งจะทำได้มุมของชิ้นส่วนครบทุกมุมนั้นสามารถทำได้โดยการปรับพารามิเตอร์ทั้งสามของเทคนิคฮัฟทรานสฟอร์มเพื่อให้ได้เส้นฮัฟทรานสฟอร์มที่ตีรอบขอบชิ้นส่วนให้จำนวนเส้นขอบและมุมของชิ้นส่วนมีครบทุกด้านทุกมุม ในขั้นตอนนี้เส้นที่ดีในแต่ละด้านของขอบชิ้นส่วนด้วยฮัฟทรานสฟอร์มจะยังคงถูกตีด้วยด้านละหลายเส้น เนื่องจากการกำหนดการรูปร่างชิ้นส่วนโดยให้จำนวนเส้นฮัฟที่ตีบนชิ้นส่วนนั้นมีเพียงด้านละเส้นจะส่งผลให้มุมบางมุมและด้านบางด้านที่สำคัญของชิ้นส่วนหายไป และหากให้จำนวนเส้นฮัฟที่ตีบนชิ้นส่วนนั้นมีมากเกินไป เส้นฮัฟจะส่งผลให้รูปร่างของชิ้นส่วนเปลี่ยนไปไม่เหมือนรูปร่างที่แท้จริง ดังนั้นจึงจำเป็นต้องกำหนดพารามิเตอร์ของฮัฟทรานสฟอร์มที่เหมาะสมที่จะทำให้ชิ้นส่วนถูกตีด้วยจำนวนเส้นฮัฟที่เหมาะสมที่ทำให้ได้มุมและด้านครบทุกมุมทุกด้าน อีกทั้งต้องทำให้รูปร่างชิ้นส่วนเหมือนจริงมากที่สุด เมื่อได้จำนวนเส้นที่ตีล้อมรอบขอบชิ้นส่วนแล้วจะทำการคัดเลือกเส้นในแต่ละด้านของขอบชิ้นส่วนให้เหลือมากที่สุดเพียงด้านละเส้นด้วยการแบ่งช่วงของฮัฟสเปซ เนื่องจากฮัฟสเปซของเส้นฮัฟที่ปรากฏบนขอบชิ้นส่วน เส้นฮัฟที่อยู่บนขอบด้านเดียวกันจะมีเซตค่าของเส้นฮัฟที่ใกล้เคียงกันจึงทำให้อยู่ในตำแหน่งของฮัฟสเปซที่ใกล้กัน ในทางกลับกันเส้นฮัฟที่อยู่บนขอบคนละด้านจะมีเซตค่าของเส้นฮัฟที่ต่างกันจึงทำให้ตำแหน่งในฮัฟสเปซอยู่ไกลกันออกไป ซึ่งการแบ่งช่วงของฮัฟสเปซนั้นสามารถแบ่งได้จากการหาจำนวนมุมของชิ้นส่วนให้เป็นตัวกำหนดสำหรับการแบ่งช่วงฮัฟสเปซเพื่อให้ด้านที่ถูกตีด้วยเส้นฮัฟเหลือด้านละมากที่สุดหนึ่งเส้น จากนั้นจึงลากเชื่อมจุดปลายของเส้นที่ติดกันหรืออยู่ใกล้กันเพื่อให้ได้เป็นรูปปิด สุดท้ายรูปร่างของชิ้นส่วนจะถูกพิจารณาเก็บข้อมูลขนาดของมุม ความยาวแต่ละด้าน และเซตค่าของแต่ละด้านของชิ้นส่วนเพื่อนำไปใช้เป็นข้อมูลในการประกอบชิ้นส่วน

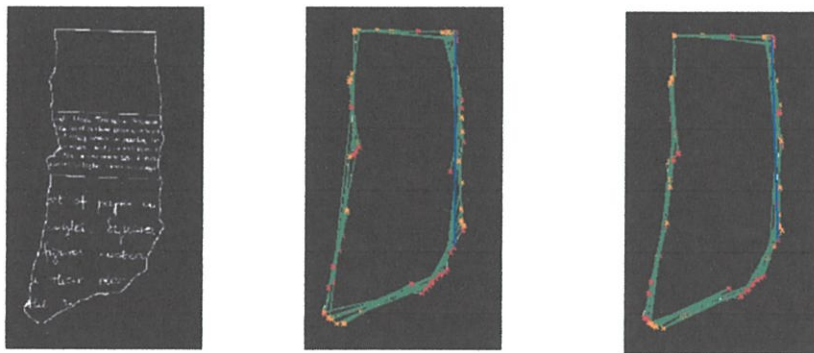
งานวิจัยฉบับนี้นำเสนอพารามิเตอร์ที่เหมาะสมสำหรับชิ้นส่วน 45 ชิ้นที่ทำให้เส้นฮัฟสามารถตีบนทุกด้านทุกมุมบนชิ้นส่วน ซึ่งจะทำให้รูปร่างที่ได้นั้นเหมือนชิ้นส่วนจริงมากที่สุดด้วยพารามิเตอร์ทั้งสามตัว พารามิเตอร์ตัวแรก คือ $\Delta\rho$ โดยจะทดลองค่าพารามิเตอร์นี้ในช่วง $[2, 20]$ ถ้าค่า $\Delta\rho$ น้อยจำนวนเส้นฮัฟที่ตีบนขอบของชิ้นส่วนจะน้อย ในทางกลับกันหากค่า $\Delta\rho$ มีค่ามากจำนวนเส้นฮัฟที่ตีบนขอบของชิ้นส่วนจะมาก ดังรูปที่ 5.5 แสดงตัวอย่างค่าพารามิเตอร์ $\Delta\rho$ ต่างๆ ในการหาขอบของชิ้นส่วน



(ก) ภาพต้นฉบับ

(ข) ค่า $\Delta\rho = 2$ (ค) ค่า $\Delta\rho = 8$ (ง) ค่า $\Delta\rho = 20$ รูปที่ 5.5 การหาขอบของชิ้นส่วนโดยการใช้ค่า $\Delta\rho$ ต่างๆ

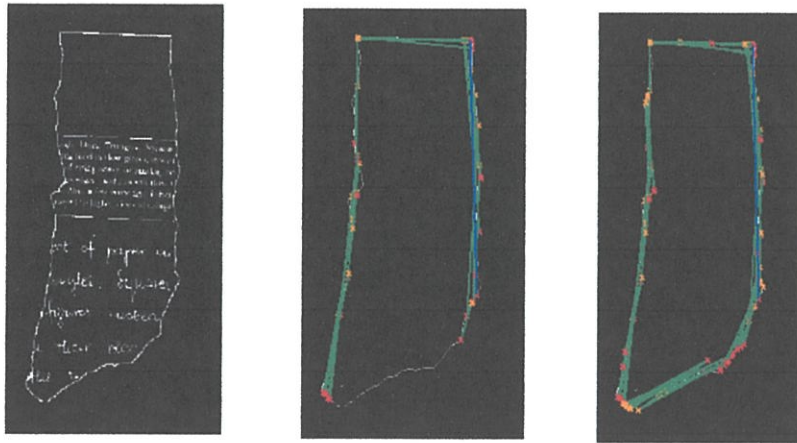
พารามิเตอร์ตัวที่สอง คือ $\Delta\theta$ ซึ่งจะทดลองค่าพารามิเตอร์นี้ในช่วง $[0.1, 0.7]$ ถ้าค่า $\Delta\theta$ น้อยจำนวนเส้นฮัพที่ติดขอบของชิ้นส่วนจะน้อย ในทางกลับกันหากค่า $\Delta\theta$ มีค่ามากจำนวนเส้นฮัพที่ติดขอบของชิ้นส่วนจะมาก ดังรูปที่ 5.6 แสดงตัวอย่างค่าพารามิเตอร์ $\Delta\theta$ ต่างๆ ในการหาขอบของชิ้นส่วน



(ก) ภาพต้นฉบับ

(ข) ใช้ค่า $\Delta\theta = 0.7$ (ค) ใช้ค่า $\Delta\theta = 0.1$ รูปที่ 5.6 การหาขอบของชิ้นส่วนโดยการปรับค่า $\Delta\theta$

พารามิเตอร์ตัวสุดท้ายคือค่าเรชโฮลด์ซึ่งในการพิจารณาเลือกค่านั้นจะมีผลต่อคุณภาพของผลการทดลอง หากเลือกค่าเรชโฮลด์ (Threshold) ต่ำ จะทำให้ได้เส้นที่ติดขอบของชิ้นส่วนไม่ครบทุกด้านของชิ้นส่วน ค่าเรชโฮลด์ (Threshold) ที่ทดลองจะอยู่ในช่วง $[10, 70]$ ซึ่งค่าเรชโฮลด์ที่เหมาะสมกับชุดทดสอบของงานวิจัยนี้อยู่ในช่วง $[40, 55]$ ตัวอย่างผลการทดลองเลือกตัดด้วยค่าเรชโฮลด์ (Threshold) ต่างๆ ดังรูปที่ 5.7



(ก) ภาพต้นฉบับ

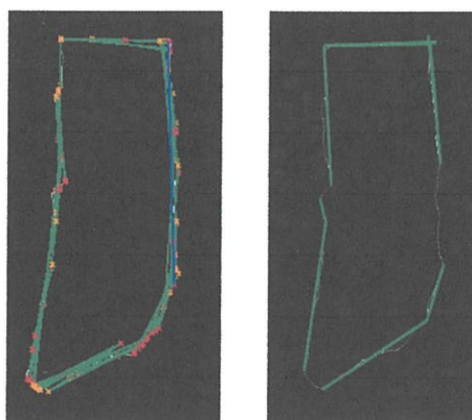
(ข) ค่าเรซโซลต์ = 20

(ค) ค่าเรซโซลต์ = 45

รูปที่ 5.7 การหาขอบของชิ้นส่วนโดยการปรับค่าเรซโซลต์

จากรูปที่ 5.5 จะเห็นว่าเมื่อค่า $\Delta\rho$ น้อยกว่าหรือมากกว่า 8 จะทำให้ได้ขอบของข้อมูลไม่ครบ หรือมีจำนวนเส้นที่ตีขอบของชิ้นส่วนซ้อนกันมากเกินไป ในรูปที่ 5.6 เมื่อกำหนดค่า $\Delta\theta$ มากกว่า 0.1 จะทำให้เส้นที่มีมุมเอียงน้อยๆ ยากที่จะถูกตีกรอบโดยใช้เพียงเส้นตรงไม่กี่เส้น และในรูปที่ 5.7 เมื่อกำหนดค่าเรซโซลต์ (Threshold) ด้วยค่าที่ต่ำเกินไปจะทำให้ไม่สามารถตีกรอบชิ้นส่วนได้ทุกด้าน ดังนั้นค่าที่เหมาะสมที่สุดสำหรับชิ้นส่วนทั้ง 45 ชิ้นของค่า $\Delta\rho$ คือ 8 , $\Delta\theta$ คือ 0.1 และค่าเรซโซลต์ (Threshold) คือ [40, 55]

เมื่อได้จำนวนเส้นฮัฟที่ล้อมรอบขอบของชิ้นส่วนด้วยจำนวนที่เหมาะสมดังรูปที่ 5.8 (ก) แล้วจึงทำการหาจำนวนมุมของชิ้นส่วนเพื่อใช้เป็นตัวแบ่งช่วงให้กับฮัฟสเปซ โดยจากคุณสมบัติของจุดที่อยู่บนเส้นตรงเดียวกันจะมีค่า ρ และ θ เท่ากัน ซึ่งในที่นี้ใช้ช่วงการเปลี่ยน (Step Size) ค่า $\Delta\rho = 8$ และ $\Delta\theta = 0.1$ ตามลำดับ เมื่อจุดเหล่านั้นบวกความถี่เพิ่มขึ้นในฮัฟสเปซช่องเดียวกันจะทำให้ฮัฟสเปซในช่องนั้นมีค่าสูง เมื่อได้จำนวนมุมแล้วจะแบ่งจำนวนช่วงของฮัฟสเปซเท่ากับจำนวนมุมของแต่ละชิ้นส่วน จากนั้นจะเลือกค่าฮัฟสเปซที่สูงที่สุดของแต่ละช่วงนั้นออกมาเพื่อให้ได้เส้นหลักของขอบแต่ละด้านดังรูปที่ 5.8 (ข)



(ก)

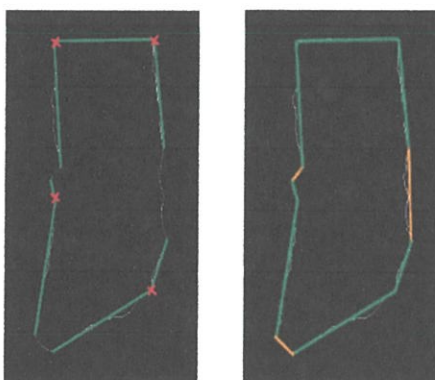
(ข)

รูปที่ 5.8 ภาพการคัดเลือกเส้นฮัพ

(ก) ภาพชิ้นส่วนที่ถูกตีขอบด้วยจำนวนเส้นฮัพที่เหมาะสม

(ข) ภาพการคัดเลือกเส้นหลักของรูป (ก)

จากรูปที่ 5.8 (ข) จะเห็นว่าขอบของชิ้นส่วนที่ถูกคัดเลือกให้เหลือเส้นฮัพอย่างมากด้านละเส้น แต่มีบางด้านที่ยังไม่ถูกตีด้วยเส้นฮัพทรานสฟอร์ม เพื่อให้ได้รูปปิดจึงต้องมีการเชื่อมเส้นโดยวิธีการตรวจสอบว่าบนเส้นตรงนั้นมีจุดตัดหรือเส้นตรงเส้นอื่นลากผ่านหรือไม่ หากไม่มีแล้วจำเป็นต้องสร้างเส้นเชื่อมโดยการนำจุดปลายของเส้นตรงนั้นเป็นจุดเริ่มต้นของเส้นเชื่อมและทำการหาจุดเริ่มต้นหรือจุดปลายของเส้นใกล้เคียงเพื่อนำมาเป็นจุดปลายของเส้นเชื่อมของเส้นที่สร้างขึ้นมาใหม่นั้นเอง ซึ่งจะแทนเส้นเชื่อมที่สร้างจากขั้นตอนนี้ด้วยสีเส้นเหลืองและหากตรวจสอบพบว่าบนเส้นตรงนั้นมีจุดตัดหรือเส้นตรงเส้นอื่นลากผ่านแล้ว จะให้ตำแหน่งจุดคู่อันดับที่ติดกันเป็นจุดเชื่อมตรงมุมของเส้นทั้งสองนั้นดังรูปที่ 5.9



(ก)

(ข)

รูปที่ 5.9 ภาพการสร้างเส้นเชื่อมให้เป็นรูปปิด

(ก) ภาพที่ได้จากการคัดเลือกเส้นหลัก

(ข) ภาพหลังจากการสร้างเส้นเชื่อมเส้นหลักของรูป (ก)

5.5 การวิเคราะห์พารามิเตอร์ที่ใช้ในการประกอบชิ้นส่วน

ในการประกอบชิ้นส่วนข้อมูลของชิ้นส่วนที่จำเป็นต้องนำมาพิจารณาการเข้าคู่ของชิ้นส่วนมีสามอย่างคือ ขนาดมุม ความยาวของด้านประกอบมุม และเซตของด้านประกอบมุม จึงต้องมีการกำหนดพารามิเตอร์ทั้งสามตัวในการคัดเลือกกลุ่มของชิ้นส่วนที่นำมาพิจารณาร่วม ได้แก่ พารามิเตอร์แรกคือ *ANG* เป็นพารามิเตอร์สำหรับกำหนดค่าความต่างระหว่างผลรวมของขนาดมุมที่พิจารณาร่วมกันต้องมีค่าต่างจาก 180 หรือ 360 ไม่เกินค่าที่กำหนด พารามิเตอร์ที่สองคือ *DIS* ใช้สำหรับการกำหนดค่าความต่างของความยาวด้านประกอบมุมของคู่ชิ้นส่วน ซึ่งด้านประกอบมุมนั้นมีสองด้านคือ เส้นก่อนมุม (ด้านที่อยู่ก่อนมุม) และเส้นหลังมุม (ด้านที่อยู่หลังมุม) โดยความยาวด้านประกอบมุมของมุมคู่ชิ้นส่วนที่พิจารณาจะต้องมีความยาวต่างกันไม่เกินค่าที่กำหนด และพารามิเตอร์สุดท้าย คือ *THETA* ซึ่งเป็นการกำหนดค่าความต่างระหว่างเซตของด้านประกอบมุมของคู่ชิ้นส่วนที่นำมาพิจารณาโดยจะต้องต่างกันไม่เกินค่าที่กำหนด ค่าเซตของด้านประกอบมุมหนึ่งมุมนั้นมีสองค่าคือ เซตเส้นก่อนมุม (เซตของด้านที่อยู่ก่อนมุม) และเซตเส้นหลังมุม (เซตของด้านที่อยู่หลังมุม) ซึ่งในการวิเคราะห์พารามิเตอร์ที่ใช้ในการประกอบชิ้นส่วนนั้นจะพิจารณาเฉพาะ 4 ชุดการทดสอบ ได้แก่

1. ชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)
2. ชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ
3. ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง
4. ชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง

เนื่องจากชุดที่ 5, 6 และ 7 เป็นชิ้นส่วนจากงานวิจัย [2] จำนวน 3 ชิ้นนั้นไม่ใช่ชิ้นส่วนที่เป็นคู่ชิ้นส่วนกัน

งานวิจัยนี้ได้ทำการทดสอบการเข้าคู่กันของชิ้นส่วนด้วยการกำหนดพารามิเตอร์ทั้งสามทั้งหมด 9 กรณีดังตารางที่ 5.7

ตารางที่ 5.7 เงื่อนไขในการพิจารณาการเข้าคู่กันของชิ้นส่วน

กรณีที่	<i>ANG</i>	<i>DIS</i>	<i>THETA</i>
1	25	150	50
2	20	150	50
3	15	150	50
4	25	100	50
5	20	100	50
6	15	100	50
7	70	200	30
8	70	150	30
9	60	150	30

ตารางที่ 5.8 จำนวนมุมที่จับคู่ถูก-คู่ผิดของชุดชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)

กรณีที่	จำนวนมุมที่จับคู่ถูก	จำนวนมุมที่จับคู่ผิด	คู่ถูก-คู่ผิด
1	33	85	-52
2	34	78	-44
3	35	64	-29
4	33	53	-20
5	33	48	-15
6	34	41	-7
7	40	87	-47
8	40	72	-32
9	40	72	-38

จากตารางที่ 5.7 เมื่อนำค่าพารามิเตอร์ทั้งสามมาคำนวณการเข้าสู่ของชิ้นส่วนในแต่ละกรณีของชุดชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน) จะได้ผลลัพธ์การจับคู่ถูกและการจับคู่ผิดดังตารางที่ 5.8

จากผลการทดลองจะเห็นว่ากรณีที่มีการจับคู่ของมุมของชิ้นส่วนน้อยที่สุดคือการใช้พารามิเตอร์กรณีที่ 6 คือ การกำหนดค่าผลรวมของขนาดคู่มุมที่พิจารณารวมกันต้องมีค่าต่างจาก 180 หรือ 360 ไม่เกิน (*ANG*) 15 ความยาวด้านประกอบมุมทั้งสองไม่เกิน (*DIS*) 100 และผลต่างเซต้าของด้านประกอบมุมไม่เกิน (*THETA*) 50 จะทำให้ได้การจับคู่มุมถูก 34 คู่ และมีการจับคู่มุมผิด 41 คู่ ดังนั้นมีการจับคู่มุมผิดมากกว่าการจับคู่มุมถูก 7 คู่ ทำให้ค่าคู่ถูก-คู่ผิดที่ได้มีค่าเป็น -7 ถึงแม้การจับคู่มุมผิดของชิ้นส่วนจะมากกว่าการจับคู่มุมถูกของชิ้นส่วนแต่ไม่ได้หมายความว่า การจับคู่ชิ้นส่วนนั้นจะมีความผิดพลาดสูง เนื่องจากผลการจับคู่มุมนั้นมุมหนึ่งมุมมีมุมได้ไม่จำกัด หากคู่มุมที่พิจารณาแล้วผ่านเงื่อนไขของกรณีจะทำให้การจับคู่ของมุมนั้นออกมาเป็นผลลัพธ์ได้นั่นเอง ดังนั้นจากขั้นตอนนี้จึงจำเป็นต้องมีการพิจารณาคู่มุมที่ผ่านเงื่อนไขแล้วมีเฉพาะการจับคู่มุมถูก กล่าวคือไม่มีการจับคู่ผิดกับมุมอื่นจึงจะถือว่ามุมนั้นเป็นมุมคู่ชิ้นส่วนที่แท้จริงซึ่งจำนวนชิ้นส่วนที่เข้าคู่มุมถูกนั้นแสดงดังตารางที่ 5.9

ตารางที่ 5.9 แสดงการจับคู่ของชุดชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)

ชิ้นส่วนที่	มุมที่	ผลลัพธ์คู่ชิ้นส่วนของงานวิจัยที่ นำเสนอ	มุมผลลัพธ์คู่ชิ้นส่วนของงานวิจัยที่ นำเสนอ
1	2	2	6
	5	3	1
	6	3	6
2	3	3	4
	4	3	3
3	3	4	11
	4	5	1
	5	5	5
4	11	3	3
	6	6	4
	7	6	3
	8	6	2
5	5	3	5
	3	7	1
6	2	4	8
	3	4	7
	4	4	6
	6	8	2
	7	8	1
7	1	5	3
	6	5	4
8	1	6	7
	2	6	6
	3	6	5

5.6 การวิเคราะห์ประสิทธิภาพการประกอบชิ้นส่วนที่ถูกฉีกขาด

การวิเคราะห์ประสิทธิภาพการประกอบชิ้นส่วนที่ถูกฉีกขาดสำหรับงานวิจัยนี้ได้นำเสนอการจับคู่ชิ้นส่วนแบบวิธีฮิวลิสติกเสิร์จ คือมีการวิเคราะห์การจับคู่ของมุมกับมุมทุกมุม แล้วจึงเลือกมุมที่มีการเข้าคู่ที่ถูกต้อง การวัดประสิทธิภาพของผลการทดลองประกอบชิ้นส่วนที่ถูกฉีกขาดด้วยสมการที่ 5.4 โดยใช้การวัดประสิทธิภาพจากจำนวนชิ้นส่วนที่เป็นคู่ชิ้นส่วนซึ่งได้จากอัลกอริทึมเทียบ

กับจำนวนของคู่ชิ้นส่วนที่ถูกต้องด้วยผู้ทดลอง (ground-truth) แล้วทำการวัดประสิทธิภาพการประกอบชิ้นชิ้นส่วนที่ถูกฉีกขาด

$$P = \left(\frac{N_A}{N_G} \right) \times 100 \quad (5.4)$$

เมื่อ P คือ ความถูกต้องของการประกอบชิ้นส่วน

N_A คือ จำนวนคู่ชิ้นส่วนที่อัลกอริทึมหาได้

N_G คือ จำนวนของคู่ชิ้นส่วนที่ถูกต้องด้วยผู้ทดลอง

จากสมการสามารถคำนวณการวัดประสิทธิภาพการประกอบชิ้นชิ้นส่วนที่ถูกฉีกขาดของแต่ละชุดการทดสอบชิ้นส่วนได้ดังตารางที่ 5.10

ตารางที่ 5.10 การวัดประสิทธิภาพการประกอบชิ้นชิ้นส่วนที่ถูกฉีกขาดของแต่ละชุดการทดสอบชิ้นส่วน

ชุดการทดสอบ	จำนวนคู่ชิ้นส่วนที่ถูก (ชิ้น)	จำนวนคู่ชิ้นส่วนที่ไม่ถูก (ชิ้น)	ประสิทธิภาพการประกอบชิ้นชิ้นส่วนที่ถูกฉีกขาด (%)
ชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ(แมกกาซีน)	13	11	54
ชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ	13	6	68.4
ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง	27	14	65.85
ชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง	23	12	65.7

จากตารางที่ 5.10 ประสิทธิภาพในการประกอบชิ้นชิ้นส่วนที่ถูกฉีกขาดของชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ(แมกกาซีน) คือ 54 เปอร์เซ็นต์ ชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ 68.4 เปอร์เซ็นต์ ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง 65.85 เปอร์เซ็นต์ และชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง 65.7 เปอร์เซ็นต์

สรุปผลและแนวทางในการพัฒนา

วิทยานิพนธ์เล่มนี้ได้นำเสนออัลกอริธึมการประกอบชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวมความยาวแกนหลักวงรีและฮัฟทรานสฟอร์ม โดยใช้ทิศทางการเรียงตัวของตัวอักษรเป็นทิศทางการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรง เพื่อเพิ่มประสิทธิภาพและความเร็วของขั้นตอนการประกอบชิ้นส่วนเอกสารที่ถูกฉีกขาดจากการออกแบบอัลกอริธึมการจัดเรียงชิ้นส่วนด้วยการคำนวณฮิสโทแกรมผลรวมความยาวแกนหลักวงรีของตัวอักษรที่มีทิศทางเดียวกันและได้ทำการทดลองด้วยชุดข้อมูลจำนวน 45 ชิ้น ซึ่งมีการตรวจสอบประสิทธิภาพโดยใช้การวัดค่าอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วนและมีการวัดประสิทธิภาพของการหาทิศทางการเรียงตัวหลักของชิ้นส่วนของอัลกอริธึมการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงด้วยการคำนวณฮิสโทแกรมผลรวมความยาวแกนหลักวงรีที่นำเสนอ และมีการทดลองปรับพารามิเตอร์ในการรูปร่างของชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์ม เพื่อนำข้อมูลรูปร่างของชิ้นส่วนไว้สำหรับขั้นตอนการประกอบ

6.1 การวิเคราะห์และสรุปผลการดำเนินงานวิจัย

6.1.1 การวิเคราะห์และสรุปผลการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรง

สำหรับงานวิจัยนี้ได้นำเสนอการจัดเรียงชิ้นส่วนให้ตั้งตรงเพื่อนำข้อมูลคุณลักษณะชิ้นส่วนใช้ในการประกอบชิ้นส่วนที่ถูกฉีกขาด โดยใช้ข้อมูลร่วระหว่างรูปร่างของขอบและเนื้อความภาพในชิ้นส่วน เพื่อลดระยะเวลาและความผิดพลาดในการประกอบชิ้นส่วนให้มากที่สุด ซึ่งเทคนิคที่ได้แนะนำเสนอคือเทคนิคการประมาณค่าระนาบการวางตัวของชิ้นส่วนโดยวิเคราะห์ค่าการหมุนหลักของชิ้นส่วนจากแนวทิศการวางตัวของตัวอักษรด้วยเทคนิคการคำนวณฮิสโทแกรมผลรวมความยาวแกนหลักวงรีของตัวอักษรที่มีทิศทางเดียวกัน จากผลการทดลองแสดงให้เห็นว่าการหาทิศทางการเรียงตัวของชิ้นส่วนเป็นเทคนิคที่เหมาะสมที่สุดสำหรับการจัดเรียงชิ้นส่วนให้อยู่ในระนาบตั้งตรงที่ได้แนะนำเสนอ เนื่องจากให้ค่าอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วนน้อยที่สุดและให้ผลประสิทธิภาพของการหาทิศทางการเรียงตัวหลักของชิ้นส่วนสูงที่สุด ซึ่งเมื่อเทียบกับการหาทิศทางการเรียงตัวของชิ้นส่วนด้วยวิธีการคำนวณฮิสโทแกรมของจำนวนวงรีที่มีทิศทางเดียวกัน จากผลการทดลองแสดงให้เห็นว่าเทคนิคที่นำเสนอสามารถแก้ไขข้อจำกัดของงานวิจัย [2] และให้ค่าอัตราความผิดพลาดของทิศทางการเรียงตัวหลักของชิ้นส่วนน้อยกว่า 5.07 และให้ประสิทธิภาพของการหาทิศทางการเรียงตัวหลักของชิ้นส่วนสูงกว่า 24.11 เปอร์เซ็นต์

6.1.2 การวิเคราะห์และสรุปผลการประกอบชิ้นส่วน

การวัดประสิทธิภาพการประกอบชิ้นส่วน ในงานวิจัยที่นำเสนอนี้มีการหารูปร่างของชิ้นส่วนด้วยเทคนิคฮัฟทรานสฟอร์มเพื่อนำข้อมูลรูปร่างของชิ้นส่วนมาใช้ในการประกอบ เนื่องจากข้อมูลสำคัญที่ใช้ในการประกอบ คือขนาดของมุมและด้านประกอบมุมของชิ้นส่วน ดังนั้นการรูปร่างด้วย

เทคนิคฮัฟทรานสฟอร์มจึงเป็นเทคนิคที่ทำให้ได้รูปร่างของชิ้นส่วนคล้ายกับรูปร่างของชิ้นส่วนจริงมากที่สุด สำหรับข้อมูลที่ใช้ในการพิจารณาการเข้าสู่ของชิ้นส่วนคือขนาดมุม ความยาวด้านประกอบมุม และเซตด้านของด้านประกอบมุม

การวัดประสิทธิภาพของผลการทดลองการเข้าสู่ของการประกอบชิ้นส่วน จากการทดลองพบว่าความถูกต้องของประสิทธิภาพการประกอบชิ้นส่วนที่ถูกฉีกขาดของชิ้นส่วนให้ผลถูกต้องเฉลี่ยเป็น 63.49 เปอร์เซ็นต์ โดยมีรูปแบบของชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ ให้ประสิทธิภาพการประกอบสูงที่สุด ค่าประสิทธิภาพการประกอบยังไม่สูงนักเนื่องจากการการต่อกันของชิ้นส่วนสามารถประกอบกันได้มากกว่าสองชิ้น กล่าวคือ การพิจารณาการเข้าสู่ของชิ้นส่วนไม่จำเป็นต้องจำกัดการเข้าสู่ของชิ้นส่วนเพียงสองชิ้น เพราะการเข้าสู่ของชิ้นส่วนสามารถเป็นการประกอบของชิ้นส่วนจำนวนกี่ชิ้นก็ได้

6.2 ข้อจำกัดและขอบเขตของงานวิจัย

ในงานวิจัยฉบับนี้เป็นการประกอบชิ้นส่วนเอกสารโดยใช้คุณลักษณะของฮิสโทแกรมผลรวมความยาวแกนหลักวงรีและฮัฟทรานสฟอร์ม โดยชิ้นส่วนต้องมีเนื้อหาภายในชิ้นส่วนเป็นตัวอักษรจึงจะสามารถหาทิศทางการเรียงตัวของชิ้นส่วนได้อย่างมีประสิทธิภาพ และเนื่องจากทิศทางการเอียงของชิ้นส่วนที่พิจารณาคือช่วง $[-\frac{\pi}{2} \frac{\pi}{2}]$ จึงทำให้ขั้นตอนการจัดเรียงชิ้นส่วนนั้นยังไม่ครอบคลุมถึงกรณีที่ชิ้นส่วนกลับหัว

6.3 แนวทางในการพัฒนา

เนื่องจากงานวิจัยนี้เป็นการประกอบเอกสารที่ถูกฉีกขาดที่มีการวิเคราะห์การเข้าสู่ของชิ้นส่วนเพียงขั้นตอนเดียว ซึ่งอาจจะทำให้ประสิทธิภาพของการประกอบไม่ถูกต้อง 100 เปอร์เซ็นต์ ดังนั้นการพัฒนาอัลกอริธึมโดยการเพิ่มวิธีการการคัดเลือกที่เหมาะสมต่อการประกอบชิ้นส่วนกระดาษฉีกขาดหรือเพิ่มอัลกอริธึมในการรู้จำเข้ามาอาจจะทำให้ได้อัลกอริธึมที่เหมาะสมต่อการใช้งานมากขึ้น ซึ่งจะสามารถเพิ่มความถูกต้องของการประกอบชิ้นส่วนให้สูงขึ้น

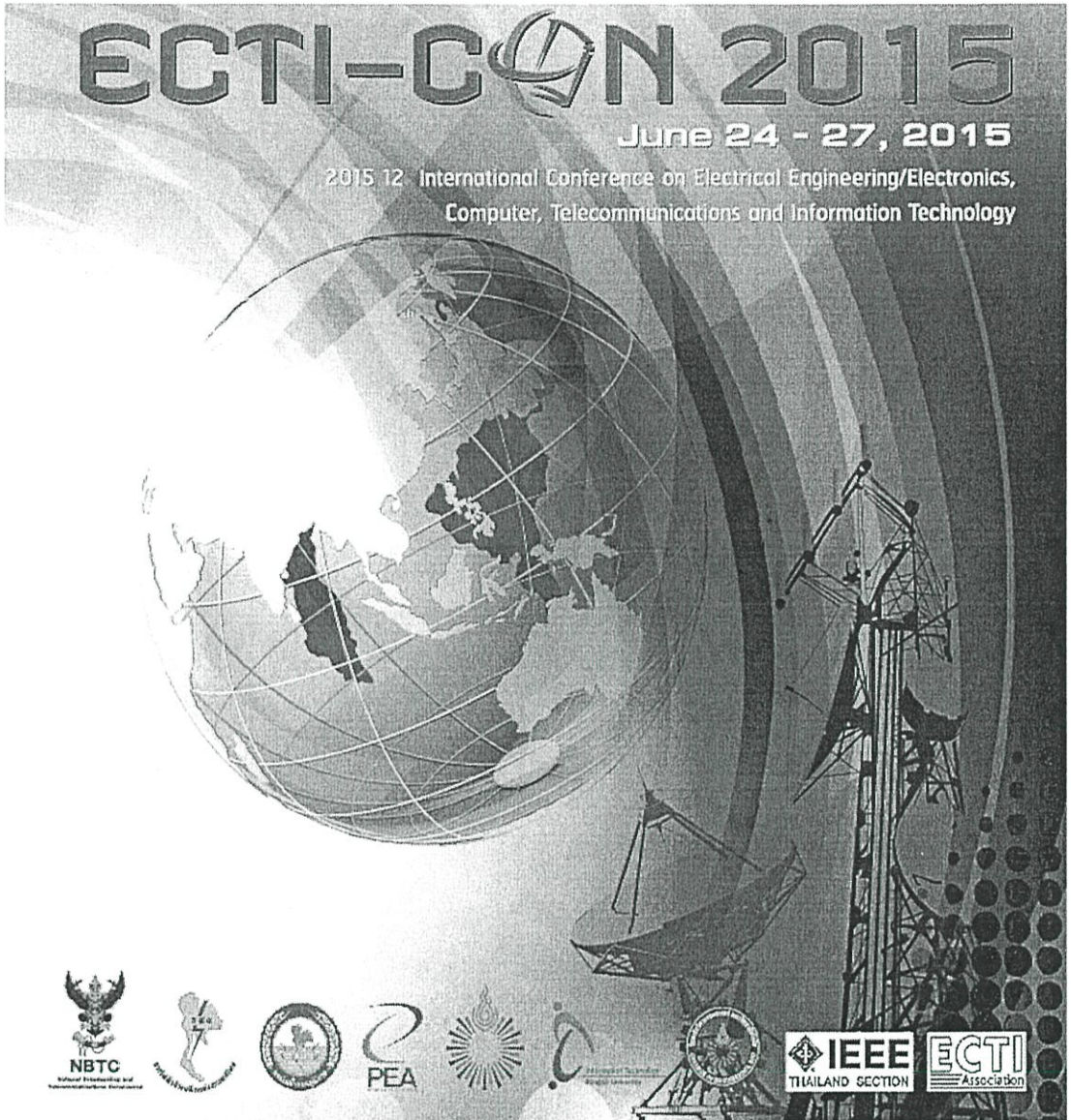
เอกสารอ้างอิง

- [1] A. Curry, "Archive collapse disaster for historians," Spiegel online international, 04th march 2009, <http://www.spiegel.de/international/germany/0,1518,611311,00.html>.
- [2] K. Florian, M. Diem and R. Sablatnig, "Torn Document Analysis as a Prerequisite for," in IEEE international conference on Virtual Systems and Multimedia, Austria, 2009, pp. 143-148.
- [3] E. D. Demaine and M. L. Demaine, "Jigsaw puzzles, edge matching, and polyomino packing: Connections and complexity," Graphs and Combinatorics, vol. 23, no. 1, 2007, pp. 195-208.
- [4] M. G. Chung, M. Fleck, and D. Forsyth, "Jigsaw puzzle solver using shape and color," 4th International Conference on Signal Processing Proceedings (ICSP '98), vol. 2, 1998, pp. 877-880.
- [5] R. Tybon, "Generating solutions to the jigsaw puzzle problem," Ph.D. dissertation, Griffith University, Australia, 2004.
- [6] M. Kampel and R. Sablatnig, "On 3d mosaicing of rotationally symmetric ceramic fragments," in Proc. of 17th International Conference on Pattern Recognition, Cambridge, UK, vol. 2, IEEE Computer Society, 2004, pp. 265-268.
- [7] R. Kasturi, D. Laurendeau, and C. Suen, Eds., "Bayesian pot-assembly from fragments as problems in perceptual-grouping and geometric-learning," in Proc. of 16th International Conference on Pattern Recognition, Quebec City, vol. 1, 2002, pp. 297-302.
- [8] C. Papaodysseus, T. Panagopoulos, M. Exarhos, C. Triantafillou, D. Fragoulis, and C. Doumas, "Contour-shape based reconstruction of fragmented, 1600 bc wall paintings," IEEE Transactions on Signal Processing, vol. 50, no. 6, pp. 1277-1288, Jun 2002.
- [9] M. Debajyoti, Y. Wang and S. Durocher, "Robust Solvers for Square Jigsaw Puzzles," in IEEE international Conference, Canada, 2013.
- [10] A. Ukovich and G. Ramponi, "Features for the reconstruction of shredded notebook paper," IEEE International Conference on Image Processing (ICIP 2005), vol. 3, Sept. 2005, pp. III-93-6.

- [11] BStU Berlin. Rekonstruktion von Unterlagen (german). accessed 11th december 2009.http://www.bstu.bund.de/cln_012/nn_714874/DE/Archiv/Rekonstruktion/rekonstruktion__node.html__nnn=true.
- [12] B. Nickolay and J. Schneider, Virtuelle Rekonstruktion "vorvernichteter" Stasi-Unterlagen. Technologische Machbarkeit und Finanzierbarkeit - Folgerungen für Wissenschaft, Kriminaltechnik und Publizistik, 2nd ed. Schriftenreihe des Berliner Landesbeauftragten für die Unterlagen des Staatssicherheitsdienstes der ehemaligen DDR (German), Berlin, 2007, vol. 21, ch. Automatische virtuelle Rekonstruktion "vorvernichtender" Stasi-Unterlagen - Machbarkeit, Systemlösung Potenziale, pp. 11–28.
- [13] H. Freeman and L. Garder, "Apictorial jigsaw puzzles: The computer solution of a problem in pattern recognition," IEEE Transactions on Computers, vol. EC-13, no. 2, pp. 118–127, April 1964.
- [14] Y. L. Huei, C. Wen and C. Fan, "Reconstruction of shredded document based on image feature matching," An International Journal in Expert Systems with Applications 39, USA, pp. 3324-3332, Sep. 2011.
- [15] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, Jan. 2004.
- [16] N. Otsu, "A threshold selection method from grey level histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, pp. 62–66, Jan. 1979.
- [17] A. Ukovich and G. Ramponi. "Features for the reconstruction of shredded notebook paper," in IEEE International Conference on Image Processing (ICIP 2005), Sep. 2005, pp. III-93-6.
- [18] F.H. Yao and S. F. Gui, "A shape and image merging technique to solve jigsaw puzzles," in Elsevier Science B.V., Japan, 2002.
- [19] E. S. A. K. Y. L. H. Wolfson, "Solving jigsaw puzzles by computer," ACM, vol. 12, no. 1-4, Feb. 1988, pp. 51-64, 1988.
- [20] S.L.Tade. Ms.Poonam D.Kesarkar, "RECONSTRUCTION OF TORN PAGE USING CORNER AND SEGMENT MATCHING," in International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, India, 2013.
- [21] Rafael C. Gonzalez, Richard E. Woods, and Steven L.Eddins, "Digital Image Processing using Matlab" Pearson Prentice Hall: Upper Saddle River, New Jersey, 2004.
- [22] ว. ดอนพรพันธ์, ส. เขียวชาญวัฒนา and ค. สุนัต, "Keypoints detection with Dop-PPRBF for SIFT feature descriptor on Low Complexity Processor," Thailand.

- [23] S. Vivier, G. Bossco and C. Nguyen, "Multiresolution approach for image processing," Campus scientifique. Loria.
- [24] John Canny. "A Computational Approach to Edge Detection". IEEE Transactions on Pattern Annalysis and Machine Intelligence, pp. 679-698, 1986.
- [25] Rafael Gonzalez and Richard Woods. 2010, *Digital Image Processing*, Third Edition. : Person Education
- [26] ไพศาล สุธีบรรเจิด, และ วิเชียร เปรมชัยสวัสดิ์, "Connected Components Labeling in 3-Dimension," in EECON, Thailand, 2554.
- [27] K. Wu, E. Otoo and A. Shoshani, "Optimizing Connected Component Labeling Algorithms," Proceedings of the SPIE, Volume 5747, pp. 1965-1976, 2005.
- [28] รศ.ดร. อรฉัตร จิตต์โสภักตร์, *ทฤษฎีการประมวลผลภาพดิจิทัล*, พิมพ์ครั้งที่ 1, กรุงเทพมหานคร: บริษัทสงวนกิจ พรินท์ แอนด์ মিเดีย, 2552.
- [29] ศิวพร ชุณหวิทย์ธีระ, "การจัดเรียงภาพย่อยเพื่อประกอบเป็นภาพใหญ่โดยใช้เทคนิคการแปลงข้อมูลแบบฮัฟ เฟสคอลลีเรชัน และวิธีการค้นหาแบบเลือกค่าที่ดีที่สุดก่อน," วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2550.

ภาคผนวก ก.
ผลงานที่ได้รับการตีพิมพ์



Plane alignment algorithm for torn document reconstruction

Bhattacharhorn Wattanacheep
 Department of Computer Engineering,
 Faculty of Engineering, King Mongkut's Institute of
 Technology Ladkrabang,
 Ladkrabang Bangkok 10520 Thailand.
 nroskool2@gmail.com

Assoc. Prof. Dr. Orachat Chitsobhuk
 Department of Computer Engineering,
 Faculty of Engineering, King Mongkut's Institute of
 Technology Ladkrabang,
 Ladkrabang Bangkok 10520 Thailand.
 orachate@yahoo.com

Abstract—Snippet alignment is a process to arrange pieces of the torn document to the original positions according to the direction of the alphabet line. It is a prerequisite to assure the effective reconstruction. The higher the performance of the snippet alignment, the greater the opportunity for successful reconstruction. Therefore, this paper presents a plane alignment algorithm for torn document reconstruction. The proposed technique analyzes the contents inside the snippet such as the direction of the character alignment based on the histogram of the accumulated radius of the fitted ellipses. The direction result is then used to revert the snippet to its original position. Hough transform based local descriptor is extracted as shape feature. These parameters are helpful for accurate reconstruction. The proposed technique can achieve approximately 5.07 decrease in relative orientation error thus increase 24.11 percent in reverting precision. This can demonstrate the significant performance improvement of the proposed algorithm.

Keywords—torn document

I. INTRODUCTION

Reconstruction of torn document is related to a technique for solving the 2D pictorial card board puzzles in the antique puzzle game [1]. However, the snippet shape of the torn document is essentially different from that of the puzzle since its shape is irregular and the content may contain text lines with different directions. Even though torn document can be processed similar to 2D puzzles, it is possible that the paper is torn in different layers causing overlapping parts of pieces [2]. With different torn layers, the boundary of the snippet may contain the outer edge called the border edge and the content edge called the internal edge as shown in Fig. 1. The border edges and the internal edges of the three connecting fragments are marked with different types of lines (dotted/dash lines) and colors.

Only the internal edges are needed to be considered since it can be used to properly align the snippet with its neighbors. In contrast, considering the border edges may result in additional errors in assembling process. Techniques for reconstructing the snippets can be categorized into two groups as pictorial or apictorial methods [3]. Mostly apictorial reconstruction depends on only a single feature such as the shape of fragments

while both shape and printed information are taken into account for the pictorial reconstruction.

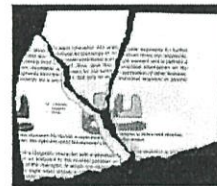


Fig. 1. Snippet from torn document with the border edge and the internal edge.

For apictorial reconstruction, an automatic reassembling of the large jigsaw puzzles was proposed by Wulfson et al. [4], which considered the curves of the jigsaw edges and analyzed them using geometry relation. The curvature of the boundary as symbolic representation of 2D shape is the derivative of the tangent angle to the curve, which is parametrized as a function of the arc length. However, the algorithm can be applied only on limited shapes of fragment formats such as triangle, rectangle, etc. To alleviate this problem, Goldberg et al. [5] proposed puzzle reconstruction using the geometric information such as convex and concave curve. Even though this technique can support reconstruction of various shapes, it still cannot distinguish the pieces with the same shapes (such as squares [6] or shredded [7]) and sizes by considering only the shape information.

Plain alignment is an alternative process to revert the snippet back to its original angle before reconstruction for faster and more accurate. It helps to distinguish snippets with similar shapes but different orientation. To align the snippet, there is a need to determine not only its shape but also its content. Analyzing only continuous shade of background color may be adequate for jigsaw arrangement. However, for torn document reconstruction, the direction of characters is important. In [2], the main orientation of the snippet is estimated based on the orientation histogram of gradient magnitude. The primary direction can be incorrectly detected

in the case of long straight boundary which dominates the direction of the characters.

Therefore, the plane alignment for torn document reconstruction based on Histogram of Accumulated Radius of fitted Ellipse (HoARE) is proposed. The received information such as snippet boundary, pictures, and table lines is eliminated prior to estimating the preliminary orientation of the ellipses fitted on text line segments. Once the main orientation parameter of each snippet has been achieved, it is reverted to its original position. This process ensures that all the snippets are ready to be assembled. It can greatly accelerate the reconstruction thus deliver higher accuracy. Then, Hough transform based local descriptor is extracted as shape feature and used later in reconstruction process. This paper is organized as follows: The proposed algorithm is detailed in section II while section III presents the experimental results. Finally, summarize and conclusion is presented in Section IV.

II. PROPOSED ALGORITHM

This paper proposes a plane alignment technique for torn document reconstruction, which analyzes and assembles the snippets based on their inner edges and content. Both features help to distinguish the snippets with similar shape thus positioning in different directions. The proposed plane alignment process consists of snippet reverting process and local descriptor extraction based on Hough transform as illustrated in Fig. 2.

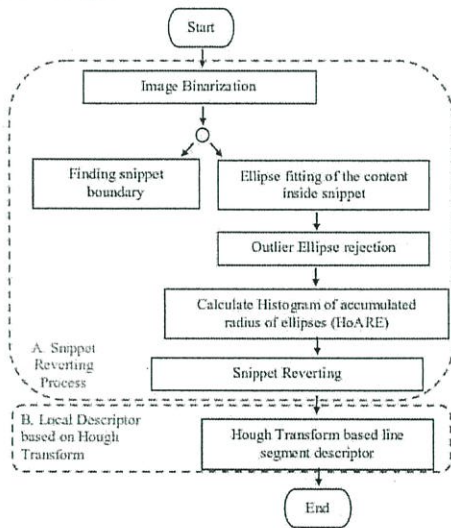


Fig. 2. The overview of the proposed plane alignment for torn document reconstruction.

In reverting process, the snippets are transformed back to their original position. The transform parameter in term of orientation is generally estimated from the maximum strength

of gradient [2] of an orientation histogram with spline interpolation. The estimation is performed both on boundary and the characters. If the snippet contains boundary with a long straight line, it can dominate the main orientation of the characters.

The fitted ellipse provides good geometric primitives of an object and can be used to represent text line segment. Evaluating dominant orientation can be done using the Histogram of Number of oriented Ellipses (HoNE) fitted on text line segments. Nevertheless, the limitation is in the case of the snippet containing a mixture of typed and handwriting document or a few text information, which results in more than one main orientation as illustrated in Fig. 3.

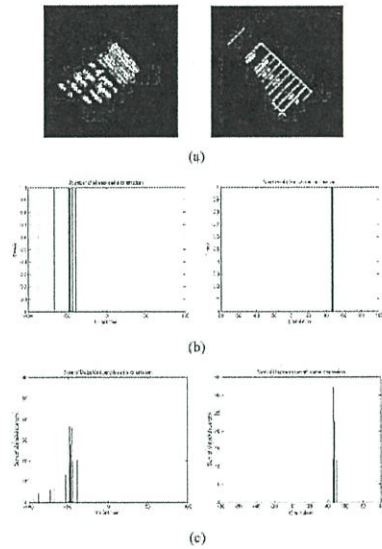


Fig. 3. A Comparison of Ellipse Histogram: (a) Input from mixture of Thai typed and handwritten letters and a few text content with table, (b) Histogram of Number of fitted Ellipses (HoNE) and (c) Histogram of Accumulated Radius of fitted Ellipse (HoARE)

With noisy information, the dominant orientation can better be represented by analyzing the histogram of accumulated radius of the fitted ellipses as HoARE. With a few fitted ellipse, the dominant direction can still be identified according the direction of the group with the longest total radius. The snippet is then reverted to the original position according to the estimated dominant orientation. In second process, the local shape descriptor of each snippet is determined using Hough transform, which will be used later in the reconstruction process. Hough transform is performed to extract primary line segments of the snippet boundary. Even though Hough transform is robust to noise, it requires high computation complexity. Suitable Hough parameters are crucial and should be evaluated in order to achieve the efficient representation.

A. Snippet reverting process

First, binary image conversion is performed on each snippet using Otsu thresholding technique [8]. In order to obtain the content inside the snippet, the boundary edges are eliminated using vertical and horizontal boundary edge scanning as shown in Fig. 4. Removing the boundary edges help to accurately approximate the quadrant of snippet orientation. The remaining content edges are then fitted by the ellipses. The length of ellipse radius with the same direction is then summarized and kept in the HoARE as presented in (1).

$$s = \max_{-90 \leq x_i \leq 90} \left(\sum_{j \in N} r_j(x_i) \right) \quad (1)$$

where $r_j(x_i)$ is the radius of the j^{th} fitted ellipses having direction within x_i , x_i is the i^{th} orientation bin of the HoARE and s is the peak of the HoARE.

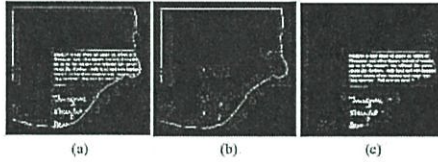


Fig. 4. Snippet boundary elimination: (a) Binary image, (b) boundary edges and (c) content edges.

The peak of the HoARE is considered as the primary orientation of the snippet, which is used to revert the snippet back to the original position. However, the primary orientation may be deviated due to the mixture content of text line segments, pictures, and tables, which sometimes align in different directions as seen in Fig. 5. In the proposed technique, the direction of text line segment is the main concentration. The fitted ellipse with the axial ratio of major and minor axis greater than 1.5 and the minor axis length between four and ten are considered as text line segment. Otherwise, it is categorized as pictures or tables and then discarded from the consideration. Eliminating the ellipses of pictures and tables can help to substantially alleviate the problem.

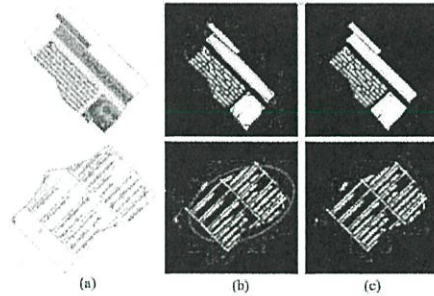


Fig. 5. Ellipse fitting on detected objects: (a) original images, (b) The fitted ellipses of detected objects, (c) Ellipse fitted on only text line segments.

B. Local descriptor extraction based on Hough transform [9]

Hough Transform (HT) is widely used to represent shape of a detected object in terms of lines, curves, or circles. It collects edge pixels corresponding to the similar parameters of the predetermined shape and summarizes them into the two dimensional histogram. Summation or integration operation in HT has been proven to be robust to noise thus provides reliable results under noisy conditions. Due to its high computational complexity, typically only the linear curve has been chosen instead of high order curve.

A linear curve can be represented in the form of polar coordinates as follows (2):

$$\rho = x \sin \theta + y \cos \theta \quad (2)$$

where x, y is the coordinate of the image on x, y plane, ρ is the distance perpendicular from the origin of the x, y image's plane to the straight line and θ is the angle from the x axis to the perpendicular line

Using the polar coordinate allows HT to collect lines of all directions. Edge pixels are quantized to the accumulated HT array according to their straight line parameters (ρ and θ). In addition to ρ and θ , another important parameter is a threshold. The appropriate threshold should capture the main straight line representing the perimeter of the snippet.

III. EXPERIMENTAL RESULTS

This paper presents a plane alignment and Hough transform based feature extraction for tom document reconstruction. In order to illustrate the performance of the proposed algorithm, a set of 45 pieces of the snippets have been prepared with various parameters such as the size ranging from 3 cm^2 to 378 cm^2 at a resolution of 600 dpi, the type of message (Print/ Handwriting), the format of the paper (blank, lined, receipt) and the angle of rotation. The snippet categories consist of letters and pictures from magazine (8 pieces), English typed and handwritten letters (7 pieces), Thai handwritten letters and table (14 pieces), Thai typed and table (13 pieces), a dataset from [2] including English typed and handwritten letters (1 pieces), English typed letters (1 piece), and very low content (1 piece).

A. Performance analysis of the snippet reverting process.

This section focuses on evaluation of the reverting results using the proposed technique compared to that of [2]. For performance comparison, precision and relative orientation error (RoE) are chosen to measure the effectiveness of the reverting technique. The relative orientation error adopted in [2] is as followed:

$$e_1 = \|\text{mod}(\alpha_c, 90) - \text{mod}(\alpha_m, 90)\| \quad (3)$$

$$e = \min(e_1, 90 - e_1) \quad (4)$$

where α_m is the groundtruth angle, α_c is the resulted angle and e is the final relative orientation error. However, considering only the angle within a single quadrant $[0 \pi/2]$ still cannot determine whether a piece is upside down or not.

TABLE I. MEASUREMENT RESULTS OF PRECISION AND ROE FOR EACH SNIPPET CATEGORY

Dataset	# (pcs)	RoE [2]	RoE propose	Precision [2]	Precision propose
letters and pictures from magazine	8	12.39	2.06	29.17	91.67
English typed and handwritten letters	7	10.54	0.53	42.86	80.95
Thai handwritten letters and table	14	2.93	0.09	80.95	95.24
Thai typed and table	13	5.39	0.05	79.49	100
English typed and handwritten letters [2]	1	0.04	0	100	100
English typed letters [2]	1	0.49	0	100	100
Very low content [2]	1	6.43	0	66.67	100
AVERAGE	45	5.46	0.39	71.30	95.41

Table 1 shows the measurement results of RoE and reverting precision for each snippet category. From the experimental results, large error occurs in the category of letters and pictures from magazine since there is an irregular piece containing text line with various directions inside the picture. Text in picture can unintentionally dominant the direction of the actual document text. It can be seen that the performance measurement of the proposed technique in RoE is decreased approximately 5.07 while the reverting precision is increased 24.11 percent compared to that of [2].

B. Analysis of Hough transform based feature extraction.

Once the snippets are reverted, their features are then extracted and used later for reconstruction. In this paper, the linear hough transform of the snippet boundary is extracted. Due to its high computational complexity, Hough parameters ($\Delta\rho$, $\Delta\theta$, th) should be investigated in order to optimize between the quality of the descriptor and the processing time. The greater quantized precision of the parameter, the better detail of the line boundary segment. The excessive precision may result in the over-detected line segments due to noisy edges thus causing inefficient execution time. In the experiments, the quantized step sizes of the parameters are chosen as $\Delta\rho = 8$ and $\Delta\theta = 0.1$. The examples of the detected line segments with the variation of $\Delta\rho$ and $\Delta\theta$ are presented in Fig. 6 and Fig. 7, respectively.

Choosing the value of the threshold is crucial since it affects the performance of the geometric representation of the snippets. If the selected threshold is too low, the detected line segments will not cover all edges of the piece. Inversely, the high threshold can lead to over-detected line segments and time consuming. In this paper, the suitable threshold is approximately between 40 and 55. Fig. 8 shows the line segments resulted from sample thresholds.

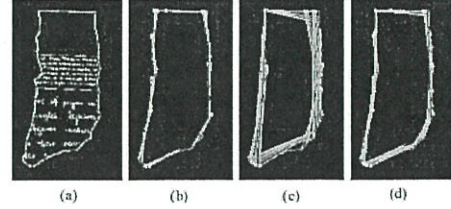


Fig. 6. The detected line segments with the variation of $\Delta\rho$ (a) Binary image, (b) $\Delta\rho = 2$, (c) $\Delta\rho = 20$ and (d) $\Delta\rho = 8$.

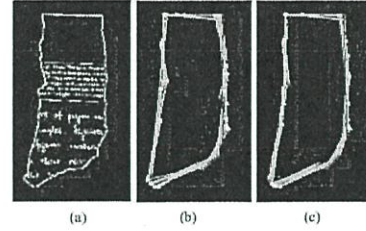


Fig. 7. The detected line segments with the variation of $\Delta\theta$: (a) Binary image, (b) $\Delta\theta = 0.7$ and (c) $\Delta\theta = 0.1$.

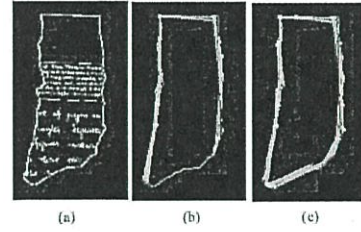


Fig. 8. The line segments resulted from sample thresholds: (a) Binary image, (b) $th = 20$ and (c) $th = 45$.

IV. CONCLUSION

This article presents a plane alignment algorithm for torn document reconstruction with the analysis of the orientation of text line segments inside each snippet. The accumulated radius of the fitted ellipses of text line segments in term of HoARE is proposed to distinguish the snippets with ambiguous properties such as similar shape but different orientation especially in the case of a mixture of typed and handwriting document, pictures, tables, and a few text information. The experimental results illustrate that, the average errors in term of RoE is decreased by 5.07 while the reverting precision is increased approximately 24.11 percent compared to the reference. This can demonstrate the significant performance improvement of the proposed algorithm. Estimating geometric shape representation in term of Hough transform descriptor is analyzed and used later in the

reconstruction process. Performing plane alignment prior to reconstruction can help to accelerate the reconstruction with higher success.

REFERENCES

- [1] E. D. Demaine and M. L. Demaine, "Jigsaw puzzles, edge matching, and polyomino packing: Connections and complexity," *Graphs and Combinatorics*, vol. 23, no. 1, pp. 195–208, 2007. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] K. Florian, M. Diem and R. Sablatnig, "Tom Document Analysis as a Prerequisite for," in *IEEE international conference*, Austria, 2009.
- [3] H. Freeman and L. Garder, "Apictorial jigsaw puzzles: The computer solution of a problem in pattern recognition," *Computers, IEEE Transactions on*, vol. EC-13, no. 2, pp. 118–127, April 1964.
- [4] WOLFSON, H., SCHONBERG, E., KALVIN, A., & LAMDAN, Y. (1988). *Solving jigsaw puzzles by computer.*, (pp. 51-64). New York.
- [5] Goldberg, D., Malon, C., & Bern, M. (n.d.). *A Global Approach to Automatic Solution of Jigsaw Puzzles*. USA.
- [6] M. Debajyoti, Y. Wang and S. Durocher, "Robust Solvers for Square Jigsaw Puzzles," in *IEEE international Conference*, Canada, 2013.
- [7] A. Ukovich and G. Ramponi, "Features for the reconstruction of shredded notebook paper," *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 3, pp. III–93–6, Sept. 2005.
- [8] N. Otsu, "A threshold selection method from grey level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, 1979.
- [9] Gonzalez, R. C., Woods, R. E., and Eddins, S. L. [2009]. *Digital Image Processing Using MATLAB*, 2nd ed., Gatesmark Publishing, Knoxville, TN.

ภาคผนวก ข.

ข้อมูลมุ่มหมุ่นจากมุ่มต่างๆของแต่ละชั้นส่วนที่ได้จาก HoARE

ตาราง ข.1 ข้อมูลมุมมองหมุนจากมุมต่างๆของแต่ละชั้นส่วนที่ได้จาก HoARE

ชุดข้อมูล	#(ชั้น)	ground-truth (องศา)																		
		-45	1	3	5	10	20	25	45	90										
1. ชั้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)	8																			
	1	-45.46	1.03	-3	5.32	10.03	-29.85	25.39	44.54	89.57										
	2	-45.22	1.64	3.97	5.48	10.64	20.83	25.2	44.8	89.73										
	3	-45.31	3.3	3.01	5.35	10.04	29.97	24.97	45.48	89.79										
	4	-44.66	1.02	3.73	5	10.02	19	25	45.35	89.55										
	5	-45.83	1	2	5.02	10	20	25	44.52	89.8										
	6	-45	-1.27	4	5.11	10	20	25.09	45	90										
	7	-45.28	18	8.58	4.54	30.28	42.58	51	71	72.58										
	8	-45	1.19	2.83	5.39	10.01	20	25.2	45	90										
2. ชั้นส่วนที่มีตัวพิมพ์และลายมือ ภาษาอังกฤษ	7																			
	1	-46.03	0.48	3.19	5	10.05	19.42	24.03	45.39	-89.85										
	2	-45.28	0.99	3	4.97	9.99	20	25	44.88	90										
	3	-45.64	0.56	3	5.2	10	20.04	25	45.2	89.83										
	4	-45	1.04	4	4.91	10	20	25	45	90										
	5	-45	0.23	3	5	10	20.12	25	45	90										

ตาราง ข.1 (ต่อ) ข้อมูลเมทริกซ์ความแตกต่างของแต่ละชิ้นส่วนที่ได้จาก HoARE

ชุดข้อมูล	#(ชิ้น)	ground-truth (องศา)									
		-45	1	3	5	10	20	25	45	90	
	6	-47.05	1.38	-3	5	10.38	20	25	44.03	-89.42	
	7	47.03	10	3.8	5	10	20	25	42.97	-89.51	
3. ชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทย และเป็นตาราง	14										
	1	-45.02	-1.23	-3.42	4.93	11.2	19.99	24.91	45.01	90	
	2	-45.04	-0.66	3	-5	10.12	20	25	45.02	-89.97	
	3	-46.2	4.16	3.04	5	10	19.99	25	44.91	89.99	
	4	-45.91	-0.36	3	5	10.27	19.99	25	45	89.98	
	5	-45.04	-1.76	3.57	5	10	20	25.32	44.97	89.97	
	6	-45.02	-1.64	3	4.88	10.31	20.79	25.48	45	89	
	7	-45	-12.85	3.38	5	9.66	20.55	25.35	45	90	
	8	-45.08	-1.31	2	5.01	10	20	25	45.02	-90	
	9	-45.12	6.77	2.63	5.09	10.04	29.97	25.02	45	-90	
	10	-45	1.98	2	7	10	20.01	25.11	44.98	90	
	11	-45.08	-1.43	3	5.2	10.22	20.73	24.8	44.96	90	
	12	-44.99	-0.23	3	5	10	20	25	45	90	
	13	-45	1.64	4	5	10.08	20	25.02	45	90.04	

ตาราง ข.1 (ต่อ) ข้อมูลมุมมองหมุนจากมุมต่างๆของแต่ละชิ้นส่วนที่ได้จาก HoARE

ชุดข้อมูล	#(ชิ้น)	ground-truth (องศา)									
		-45	1	3	5	10	20	25	45	90	
	14	-45	1.32	3.42	4.98	10.83	20.8	24.52	45	90	
4. ชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง	13										
	1	-45	-0.79	-2.85	5	10.46	20.57	24.54	45	90	
	2	-45	0.78	3	5.2	10	20	25	45.11	90.01	
	3	-45	0.67	2.97	4.96	10.28	20	24.88	45	90	
	4	-45.27	3.69	2.9	5	-10.02	20	25.01	45	89.99	
	5	-45	4.3	2.98	5	10.08	20	24.96	45.32	90	
	6	-45	2.03	2	5	10.08	20.03	25.2	45	90.12	
	7	-45.38	-2.33	2.98	5	10	20	25	45	90	
	8	-45	2.92	3	4.03	10	20	25	45.09	90	
	9	-45	-3.4	2.89	4.8	10.91	19.98	25	45	89.99	
	10	-45.08	-2.02	3	5	10	20	25	45.2	90.03	
	11	-45.01	-1.98	3.42	5.02	10	20	24.98	45.2	90	
	12	-45	10	4	4.52	10	20	25	45	90	
	13	-45	1.31	3	5	12.05	19.42	24.03	45	90	

ตาราง ข.1 (ต่อ) ข้อมูลมุมมองจากมุมต่างๆของแต่ละชั้นส่วนที่ได้จาก HOARE

ชุดข้อมูล	#(ชั้น)	ground-truth (องศา)								
		-45	1	3	5	10	20	25	45	90
5. ชั้นส่วนที่มีตัวพิมพ์และลายมือมีทิศ การเรียงตัวที่ต่างกัน (2)	1	-45	1.65	3.99	5	10	20	25	45	90
6. ชั้นส่วนที่มีเฉพาะตัวพิมพ์ (2)	1	-45	-1.7	3	5	10	20	25	45	90
7. ชั้นส่วนที่มีข้อมูลข้างในน้อยมาก (2)	1	-45	0.36	3.79	4.98	10	20	25	45	90

ภาคผนวก ค.

ข้อมูลค่า ROE และ Precision แต่ละชั้นส่วนจากมุมมองต่างๆ

ตาราง ค.1 ค่า ROE และ Precision แต่ละชั้นส่วนจากมูมหมุนต่างๆ

ชุดข้อมูล	#(ชั้น)	ROE	Precision
1. ชั้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน)	8		
	1	0.45	1
	2	0.23	1
	3	0.33	1
	4	0.38	1
	5	0.50	1
	6	0	1
	7	14.57	0.33
	8	0	1
ค่าเฉลี่ย (กลุ่ม 1)		2.06	0.9167
2. ชั้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ	7		
	1	0.52	0.67
	2	0.13	1
	3	0.34	1
	4	0	1
	5	0	1
	6	1.2	0.67
	7	1.52	0.33
ค่าเฉลี่ย (กลุ่ม 2)		0.53	0.8095
3. ชั้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง	14		
	1	0.01	1
	2	0.03	1
	3	0.43	0.67
	4	0.31	1
	5	0.03	1
	6	0.34	0.67
	7	0	1
	8	0.03	1
	9	0.04	1

ตาราง ค.1 (ต่อ) ค่า ROE และ Precision แต่ละชั้นส่วนจากมูมหมุ่นต่างๆ

ชุดข้อมูล	#(ชั้น)	ROE	Precision
	10	0.01	1
	11	0.04	1
	12	0.003	1
	13	0.01	1
	14	0	1
ค่าเฉลี่ย (กลุ่ม 3)		0.09	0.9524
4. ชั้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง	13		
	1	0	1
	2	0.04	1
	3	0	1
	4	0.09	1
	5	0.11	1
	6	0.04	1
	7	0.13	1
	8	0.03	1
	9	0.003	1
	10	0.10	1
	11	0.07	1
	12	0	1
	13	0	1
ค่าเฉลี่ย (กลุ่ม 4)		0.05	1
5. ชั้นส่วนที่มีเฉพาะตัวพิมพ์ (2)	1	0	1
6. ชั้นส่วนที่มีเฉพาะตัวพิมพ์ (2)	1	0	1
7. ชั้นส่วนที่มีข้อมูลข้างในน้อยมาก (2)	1	0	1
ค่าเฉลี่ย (7 กลุ่ม)		0.39	0.9541

ภาคผนวก ง.

ข้อมูลการประกอบของแต่ละมุมในแต่ละเงื่อนไข

ในส่วนของภาคผนวก ง. นี้ ประกอบด้วยตาราง 4 ตาราง ได้แก่

ตาราง ง.1 ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์และรูปภาพ (แมกกาซีน) ซึ่งเป็นตารางแสดง การจับคู่มุมของแต่ละมุมของชุดชิ้นส่วนแมกกาซีนในแต่ละกรณี

ตาราง ง.2 ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษ ซึ่งเป็นตารางแสดง การจับคู่มุมของแต่ละมุมของชุดชิ้นส่วนที่มีตัวพิมพ์และลายมือภาษาอังกฤษในแต่ละกรณี

ตาราง ง.3 ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง ซึ่งเป็นตารางแสดง การจับคู่มุมของแต่ละมุมของชุดชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตารางในแต่ละกรณี

ตาราง ง.4 ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตาราง ซึ่งเป็นตารางแสดง การจับคู่มุมของแต่ละมุมของชุดชิ้นส่วนที่มีตัวพิมพ์ภาษาไทยและเป็นตารางในแต่ละกรณี

โดยการเปรียบเทียบความถูกต้องจาก ground-truth โดยจำนวนมุมที่จับคู่ถูกและการจับคู่ผิดของมุมที่พิจารณาอาจมีได้มากกว่าหนึ่งมุม

กรณี 1 คือ กรณีที่ให้ค่า $ANG = 25, DIS = 150. THETA = 50$

กรณี 2 คือ กรณีที่ให้ค่า $ANG = 20, DIS = 150. THETA = 50$

กรณี 3 คือ กรณีที่ให้ค่า $ANG = 15, DIS = 150. THETA = 50$

กรณี 4 คือ กรณีที่ให้ค่า $ANG = 25, DIS = 100. THETA = 50$

กรณี 5 คือ กรณีที่ให้ค่า $ANG = 20, DIS = 100. THETA = 50$

กรณี 6 คือ กรณีที่ให้ค่า $ANG = 15, DIS = 100. THETA = 50$

กรณี 7 คือ กรณีที่ให้ค่า $ANG = 70, DIS = 150. THETA = 30$

กรณี 8 คือ กรณีที่ให้ค่า $ANG = 70, DIS = 150. THETA = 30$

กรณี 9 คือ กรณีที่ให้ค่า $ANG = 60, DIS = 150. THETA = 30$

ตาราง ๖.3 (ต่อ) ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง

ชั้นที่	จำนวนมุม	มุมที่	คู่ถูก									คู่ผิด								
			กรณี 1	กรณี 2	กรณี 3	กรณี 4	กรณี 5	กรณี 6	กรณี 7	กรณี 8	กรณี 9	กรณี 1	กรณี 2	กรณี 3	กรณี 4	กรณี 5	กรณี 6	กรณี 7	กรณี 8	กรณี 9
		3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		4	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0
		5	0	0	0	0	0	1	1	1	1	1	1	1	3	2	2	4	4	4
		6	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2	2
		7	1	1	1	1	1	1	1	1	1	1	1	1	3	2	1	0	3	2
		8	1	1	1	1	1	1	1	1	1	1	1	1	4	3	2	1	2	2
11	9	1	1	1	1	1	1	1	1	1	1	1	1	3	3	2	2	2	2	3
		2	0	0	0	0	0	1	1	1	1	1	1	2	2	1	1	1	1	1
		3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1
		4	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0
		5	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0	0	0
		6	0	0	0	0	0	0	0	0	0	0	0	3	2	2	2	5	2	2
		7	1	1	1	1	1	1	1	1	1	1	1	4	4	3	2	3	3	3
		8	1	1	1	0	0	1	1	1	1	1	1	3	3	2	2	4	3	3
		9	1	1	1	1	1	1	1	1	1	1	1	2	2	2	0	0	2	2
12	10	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

ตาราง ๓.๓ (ต่อ) ตารางจำนวนคู่ถูก-คู่ผิดแต่ละมุมของชิ้นส่วนที่มีตัวพิมพ์ ลายมือภาษาไทยและเป็นตาราง

ชิ้นที่	จำนวนมุม	มุมที่	คู่ถูก									คู่ผิด													
			กรณี 1	กรณี 2	กรณี 3	กรณี 4	กรณี 5	กรณี 6	กรณี 7	กรณี 8	กรณี 9	กรณี 1	กรณี 2	กรณี 3	กรณี 4	กรณี 5	กรณี 6	กรณี 7	กรณี 8	กรณี 9					
		8	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0	0
14	9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	2	2	3	3	3
		2	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	2	2	0	0	0	0	2
		3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	0	0	0	0	3
		4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1
		5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1
		6	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	1	1
		7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
รวม	122		74	75	78	74	74	74	74	76	76	92	91	91	91	172	161	139	105	98	89	200	145	145	

ประวัติผู้เขียน

ชื่อ-นามสกุล

นางสาวภัทรภร วัฒนาชีพ

ที่อยู่

71 หมู่ 8 ต.อิสาน อ.เมือง จ.บุรีรัมย์ 31000

ประวัติการศึกษา

พ.ศ. 2555

จบการศึกษาระดับปริญญาตรี คณะวิศวกรรมศาสตร์ สาขาคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง